Emergency Resource Deployment Planning (SOC)

ERDP (SOC)-Student Manual

1st Edition, 6th Printing-October 2022

FEMA

FEMA/USFA/NFA
ERDP (SOC)-SM
October 2022
1st Edition, 6th Printing

Emergency Resource Deployment Planning (SOC)

Emergency Resource Deployment Planning (SOC)

ERDP (SOC)-Student Manual

1st Edition, 6th Printing-October 2022

FEMA

U.S. DEPARTMENT OF HOMELAND SECURITY

UNITED STATES FIRE ADMINISTRATION

NATIONAL FIRE ACADEMY

FOREWORD

The U.S. Fire Administration (USFA), an important component of the Department of Homeland Security (DHS), serves the leadership of this nation as the DHS's fire protection and emergency response expert. The USFA is located at the National Emergency Training Center (NETC) in Emmitsburg, Md., and includes the National Fire Academy (NFA), National Fire Data Center (NFDC), and the National Fire Programs (NFP). The USFA also provides oversight and management of the Noble Training Center in Anniston, Ala. The mission of the USFA is to save lives and reduce economic losses due to fire and related emergencies through training, research, data collection and analysis, public education, and coordination with other federal agencies and fire protection and emergency service personnel.

The USFA's NFA offers a diverse course delivery system, combining resident courses, off-campus deliveries in cooperation with state training organizations, weekend instruction, and online courses. The USFA maintains a blended learning approach to its course selections and course development. Resident courses are delivered at both the Emmitsburg campus and the Noble facility. Off-campus courses are delivered in cooperation with state and local fire training organizations to ensure this nation's firefighters are prepared for the hazards they face.

The USFA's organizational values include: integrity, communication, honesty, accountability, respect and trust. Integrity: We adhere to our code of ethics and controls which govern conduct and performance. Communication: We consistently share and provide access to information throughout the USFA to enhance collaboration and to eliminate ambiguity, frustration and uncertainty. Honesty: We embrace fairness and equity as paramount to all human capital and business affairs. Accountability: We are obligated and willing to accept responsibility and to answer for the results of our performance and conduct. Respect: We consider all USFA members worthy of high regard and have a sincere desire to see others succeed. Trust: We optimistically rely on the character, ability and strength of each member to contribute wholeheartedly to the success of the USFA.

The "Emergency Resource Deployment Planning (SOC)" training uses these values as a baseline. The goal of the course is to give students the ability to develop a comprehensive Standards of Cover (SOC) for their fire organization. The purpose of this course is to provide the skills and knowledge needed to develop an SOC for your fire organization.

This page intentionally left blank.

TABLE OF CONTENTS

This page intentionally left blank.

COURSE PURPOSE

The purpose of this course is to provide the knowledge, skills and tools needed to develop an SOC for your fire organization.

COURSE GOAL

The goal of this course is for students to develop a comprehensive SOC for their fire organization.

COURSE OBJECTIVES

The students will be able to:

1. Develop a process for creating a community SOC.
2. Describe the process for evaluating community expectations, risk and performance measures.
3. Select, use and apply data analysis tools and techniques appropriate to the creation of risk assessment, deployment analysis, and performance measurement.
4. Use National Fire Incident Reporting System (NFIRS) and common data analysis tools to evaluate an organization's abilities to mitigate and respond to risk (Plan, Mitigate, Respond and Recover).
5. Conduct and apply the results of a risk assessment, deployment analysis, and performance measurement to identify responses to community risks.
6. Develop a template for SOC which addresses community needs.

TARGET AUDIENCE

The target audience for this course consists of fire and emergency service personnel in positions of authority who have an opportunity to exercise leadership. The students must, minimally, be assigned to a supervisory level position, e.g., company officer.

PREREQUISITES

1. NFIRS Self Study Q0494.
2. Geographic Information System (GIS) Tutorial — plus Captivate from the USFA website.
3. Two Quick Source Guides — Microsoft Excel, Microsoft Excel Advanced — plus UMD online tutorial.

This page intentionally left blank.

SCHEDULE

TIME	DAY 1	DAY 2
	Unit 0: Welcome	Unit 2: Data Analysis and Techniques (cont'd)
	Break	*Break*
	Unit 0: Welcome (cont'd)	Unit 2: Data Analysis and Techniques (cont'd)
	Break	*Break*
	Unit 1: Expectations, Standards of Cover, Performance Measures and Risks	Unit 2: Data Analysis and Techniques (cont'd)
	Lunch Break	*Lunch Break*
	Unit 1: Expectations, Standards of Cover, Performance Measures and Risks (cont'd)	Unit 2: Data Analysis and Techniques (cont'd)
	Break	*Break*
	Unit 2: Data Analysis and Techniques	Unit 2: Data Analysis and Techniques (cont'd) Unit 3: Working with Statistics

Evening Icebreaker Activity

TIME	DAY 3	DAY 4
	Unit 3: Working With Statistics (cont'd)	Unit 3: Working With Statistics (cont'd)
	Break	*Break*
	Unit 3: Working With Statistics (cont'd)	Unit 4: Geographic Information Technologies
	Break	*Break*
	Unit 3: Working With Statistics (cont'd)	Unit 4: Geographic Information Technologies (cont'd)
	Lunch Break	*Lunch Break*
	Unit 3: Working With Statistics (cont'd)	Unit 4: Geographic Information Technologies (cont'd)
	Break	*Break*
	Unit 3: Working With Statistics (cont'd)	Unit 4: Geographic Information Technologies (cont'd)

TIME	DAY 5	DAY 6
	Unit 4: Geographic Information Technologies (cont’d)	Unit 5: Bringing It Full Circle
	Break	*Break*
	Unit 4: Geographic Information Technologies (cont’d)	Unit 5: Bringing It Full Circle (cont’d)
	Break	*Break*
	Unit 4: Geographic Information Technologies (cont’d)	Unit 5: Bringing It Full Circle (cont’d)
	Lunch Break	*Lunch Break*
	Unit 4: Geographic Information Technologies (cont’d)	Unit 5: Bringing It Full Circle (cont’d)
	Break	*Break*
	Unit 4: Geographic Information Technologies (cont’d)	Unit 5: Bringing It Full Circle (cont’d)

This page intentionally left blank.

FIREFIGHTER CODE OF ETHICS

Background

The Fire Service is a noble calling, one which is founded on mutual respect and trust between firefighters and the citizens they serve. To ensure the continuing integrity of the Fire Service, the highest standards of ethical conduct must be maintained at all times.

Developed in response to the publication of the Fire Service Reputation Management White Paper, the purpose of this National Firefighter Code of Ethics is to establish criteria that encourages fire service personnel to promote a culture of ethical integrity and high standards of professionalism in our field. The broad scope of this recommended Code of Ethics is intended to mitigate and negate situations that may result in embarrassment and waning of public support for what has historically been a highly respected profession.

Ethics comes from the Greek word ethos, meaning character. Character is not necessarily defined by how a person behaves when conditions are optimal and life is good. It is easy to take the high road when the path is paved and obstacles are few or non-existent. Character is also defined by decisions made under pressure, when no one is looking, when the road contains land mines, and the way is obscured. As members of the Fire Service, we share a responsibility to project an ethical character of professionalism, integrity, compassion, loyalty and honesty in all that we do, all of the time.

We need to accept this ethics challenge and be truly willing to maintain a culture that is consistent with the expectations outlined in this document. By doing so, we can create a legacy that validates and sustains the distinguished Fire Service institution, and at the same time ensure that we leave the Fire Service in better condition than when we arrived.

FIREFIGHTER CODE OF ETHICS

I understand that I have the responsibility to conduct myself in a manner that reflects proper ethical behavior and integrity. In so doing, I will help foster a continuing positive public perception of the fire service. Therefore, I pledge the following…

- Always conduct myself, on and off duty, in a manner that reflects positively on myself, my department and the fire service in general.
- Accept responsibility for my actions and for the consequences of my actions.
- Support the concept of fairness and the value of diverse thoughts and opinions.
- Avoid situations that would adversely affect the credibility or public perception of the fire service profession.
- Be truthful and honest at all times and report instances of cheating or other dishonest acts that compromise the integrity of the fire service.
- Conduct my personal affairs in a manner that does not improperly influence the performance of my duties, or bring discredit to my organization.
- Be respectful and conscious of each member's safety and welfare.
- Recognize that I serve in a position of public trust that requires stewardship in the honest and efficient use of publicly owned resources, including uniforms, facilities, vehicles and equipment and that these are protected from misuse and theft.
- Exercise professionalism, competence, respect and loyalty in the performance of my duties and use information, confidential or otherwise, gained by virtue of my position, only to benefit those I am entrusted to serve.
- Avoid financial investments, outside employment, outside business interests or activities that conflict with or are enhanced by my official position or have the potential to create the perception of impropriety.
- Never propose or accept personal rewards, special privileges, benefits, advancement, honors or gifts that may create a conflict of interest, or the appearance thereof.
- Never engage in activities involving alcohol or other substance use or abuse that can impair my mental state or the performance of my duties and compromise safety.
- Never discriminate on the basis of race, religion, color, creed, age, marital status, national origin, ancestry, gender, sexual preference, medical condition or handicap.
- Never harass, intimidate or threaten fellow members of the service or the public and stop or report the actions of other firefighters who engage in such behaviors.
- Responsibly use social networking, electronic communications, or other media technology opportunities in a manner that does not discredit, dishonor or embarrass my organization, the fire service and the public. I also understand that failure to resolve or report inappropriate use of this media equates to condoning this behavior.

Developed by the National Society of Executive Fire Officers

A Student Guide to End-of-course Evaluations

Say What You Mean ...

Ten Things You Can Do to Improve the National Fire Academy

The National Fire Academy takes its course evaluations very seriously. Your comments and suggestions enable us to improve your learning experience.

Unfortunately, we often get end-of-course comments like these that are vague and, therefore, not actionable. We know you are trying to keep your answers short, but the more specific you can be, the better we can respond.

Actual quotes from student evaluations:	Examples of specific, actionable comments that would help us improve the course:
1 "Update the materials."	• The (ABC) fire video is out-of-date because of the dangerous tactics it demonstrates. The available (XYZ) video shows current practices. • The student manual references building codes that are 12 years old.
2 "We want an advanced class in (fill in the blank)."	• We would like a class that enables us to calculate energy transfer rates resulting from exposure fires. • We would like a class that provides one-on-one workplace harassment counseling practice exercises.
3 "More activities."	• An activity where students can physically measure the area of sprinkler coverage would improve understanding of the concept. • Not all students were able to fill all ICS positions in the exercises. Add more exercises so all students can participate.
4 "A longer course."	• The class should be increased by one hour per day to enable all students to participate in exercises. • The class should be increased by two days so that all group presentations can be peer evaluated and have written abstracts.
5 "Readable plans."	• The plans should be enlarged to 11 by 17 and provided with an accurate scale. • My plan set was blurry, which caused the dotted lines to be interpreted as solid lines.
6 "Better student guide organization," "manual did not coincide with slides."	• The slide sequence in Unit 4 did not align with the content in the student manual from slides 4-16 through 4-21. • The instructor added slides in Unit 4 that were not in my student manual.
7 "Dry in spots."	• The instructor/activity should have used student group activities rather than lecture to explain Maslow's Hierarchy. • Create a pre-course reading on symbiotic personal relationships rather than trying to lecture on them in class.
8 "More visual aids."	• The text description of V-patterns did not provide three-dimensional views. More photographs or drawings would help me imagine the pattern. • There was a video clip on NBC News (date) that summarized the topic very well.
9 "Re-evaluate pre-course assignments."	• The pre-course assignments were not discussed or referenced in class. Either connect them to the course content or delete them. • The pre-course assignments on ICS could be reduced to a one-page job aid rather than a 25-page reading.
10 "A better understanding of NIMS."	• The instructor did not explain the connection between NIMS and ICS. • The student manual needs an illustrated guide to NIMS.

UNIT 0:
WELCOME

TERMINAL OBJECTIVE

The students will be able to:

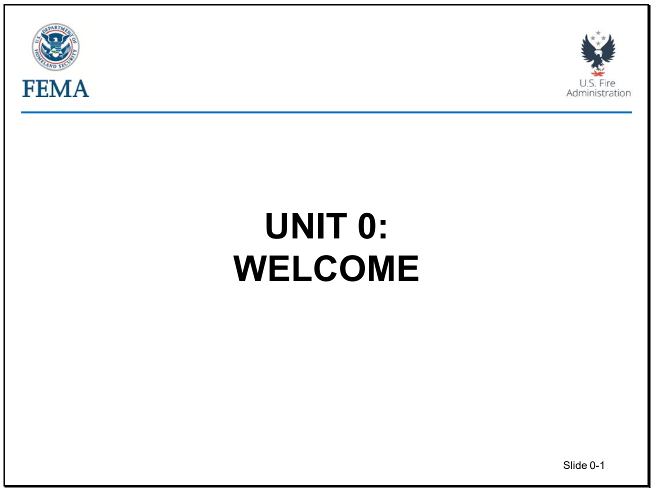
0.1 Prepare themselves for training after reviewing administrative matters and making introductions.

ENABLING OBJECTIVE

The students will be able to

0.1 Introduce themselves and list personal expectations for the course.

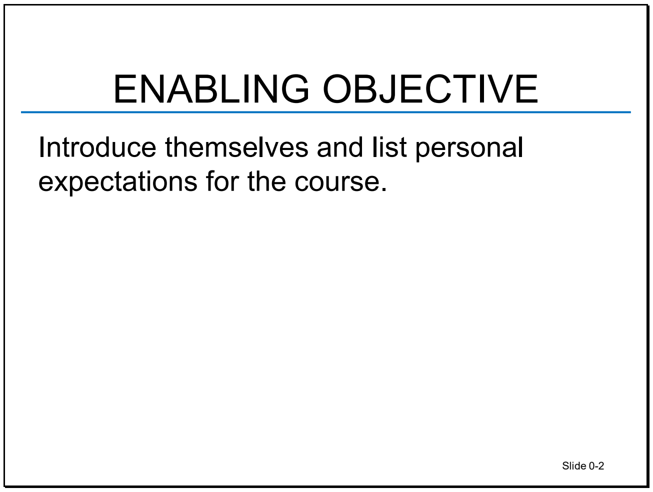
This page intentionally left blank



FEMA

U.S. Fire Administration

UNIT 0:
WELCOME



ENABLING OBJECTIVE

Introduce themselves and list personal expectations for the course.

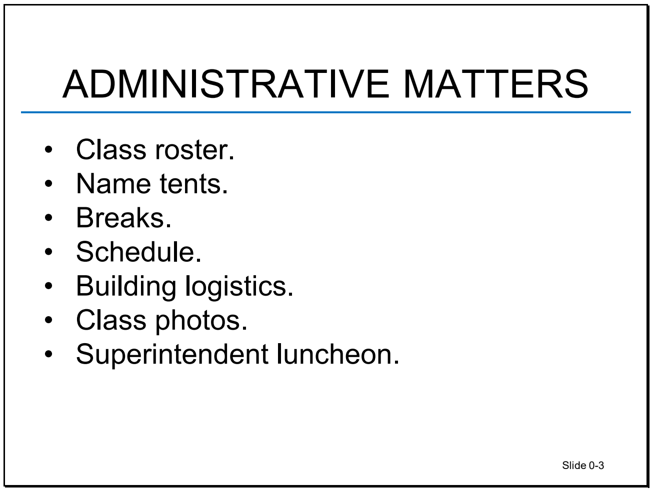
I. WELCOME AND INTRODUCTIONS

A. Instructor introductions.

B. Overview.

  1. Welcome and introductions.
  2. Administrative matters.
  3. Student introductions and expectations.
  4. Course purpose, goal and objectives.
  5. Overview of Standards of Cover (SOC).

II. ADMINISTRATIVE MATTERS



ADMINISTRATIVE MATTERS

- Class roster.
- Name tents.
- Breaks.
- Schedule.
- Building logistics.
- Class photos.
- Superintendent luncheon.

A. Class roster — circulate the roster through the class to correct spelling of names, addresses, etc., and give it to the training specialist.

B. Name tents.

C. Breaks.

D. Schedule.

E. Building logistics:

   1. Restrooms.

   2. Fire exits — note fire exits and relevant procedures.

F. Class photos.

G. Superintendent luncheon.

III. CD ON DESK

CD ON DESK CONTENTS

- Firefighter Life Safety Initiatives.
- Credit for National Fire Academy (NFA) Courses.
- Emergency Management and Response-Information Sharing and Analysis Center (EMR-ISAC) Fact Sheet.
- Evaluation of NFA Courses.
- Fire Corps Brochure and Letter.
- First Responders — Pandemic.
- NFA Course Catalog and Schedule.
- Transcript Request.

A. Contents.

1. Firefighter Life Safety Initiatives.

2. Credit for National Fire Academy (NFA) Courses.

3. Emergency Management and Response-Information Sharing and Analysis Center (EMR-ISAC) Fact Sheet.

4. Evaluation of NFA Courses.

5. Fire Corps Brochure and Letter.

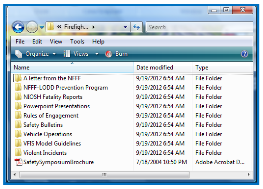
6. First Responders — Pandemic.

7. NFA Course Catalog and Schedule.

8. Transcript Request.

FIREFIGHTER LIFE SAFETY INITIATIVES

- A Letter from the National Fallen Firefighters Foundation (NFFF).
- NFFF-Line-of-Duty Death (LODD) Prevention Program.
- National Institute of Occupational Safety and Health (NIOSH) Fatality Reports.
- Rules of Engagement.
- Safety Bulletins.
- Vehicle Operations.
- Volunteer Firemen's Insurance Services (VFIS) Model Guidelines.
- Violent Incidents.



Name	Date modified	Type
A letter from the NFFF	9/19/2012 6:54 AM	File Folder
NFFF-LODD Prevention Program	9/19/2012 6:54 AM	File Folder
NIOSH Fatality Reports	9/19/2012 6:54 AM	File Folder
Powerpoint Presentations	9/19/2012 6:54 AM	File Folder
Rules of Engagement	9/19/2012 6:54 AM	File Folder
Safety Bulletins	9/19/2012 6:54 AM	File Folder
Vehicle Operations	9/19/2012 6:54 AM	File Folder
VFIS Model Guidelines	9/19/2012 6:54 AM	File Folder
Violent Incidents	9/19/2012 6:54 AM	File Folder
SafetySymposiumBrochure	7/18/2004 10:50 PM	Adobe Acrobat D...

B. Firefighter Life Safety Initiatives

1. A Letter from the National Fallen Firefighters Foundation (NFFF).
2. NFFF-Line-of-Duty Death (LODD) Prevention Program.
3. National Institute of Occupational Safety and Health (NIOSH) Fatality Reports.
4. Rules of Engagement.
5. Safety Bulletins.
6. Vehicle Operations.
7. Volunteer Firemen's Insurance Services (VFIS) Model Guidelines.
8. Violent Incidents.

ACTIVITY 0.1

Student Introductions and Expectations

Purpose

To give you an opportunity to introduce yourself and provide your expectations for the course.

Directions

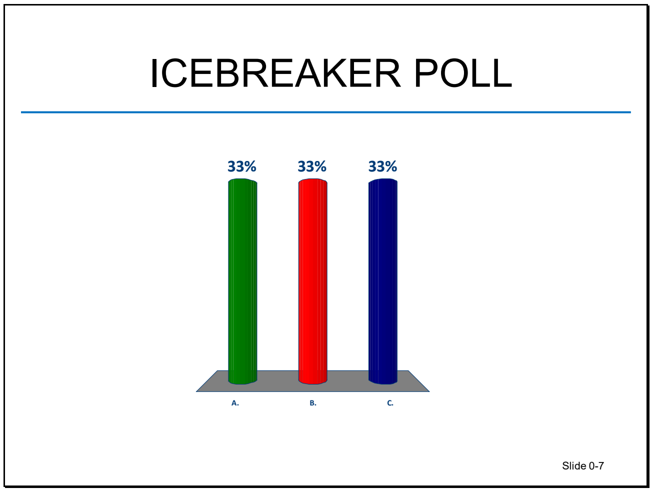
Part 1

1. Briefly introduce yourselves to one another at your table.
2. Produce a list of at least five reasons that collectively indicate your expectations for the course. (Scribe as you go on the easel pad, you have 10 minutes.)

Part 2

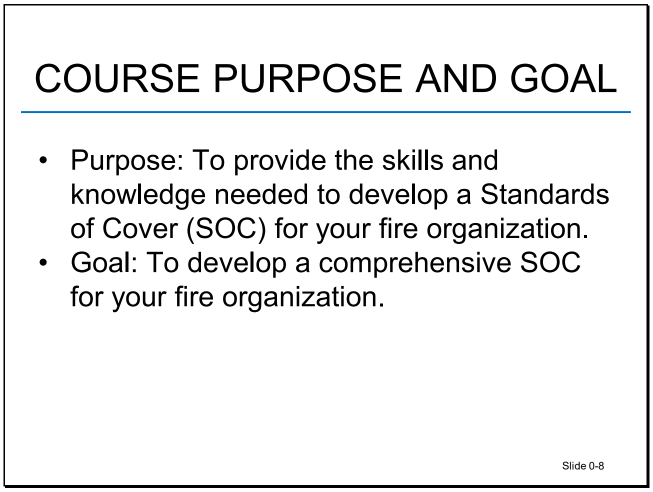
1. We will then go around the room and allow you to briefly introduce yourself to the other members of the class and share your own expectations for the course.
2. What is your agency's current status with regard to fire service or Emergency Medical Services (EMS) accreditation, and/or developing an SOC?

This page intentionally left blank



ICEBREAKER POLL

IV. COURSE OVERVIEW



COURSE PURPOSE AND GOAL

- Purpose: To provide the skills and knowledge needed to develop a Standards of Cover (SOC) for your fire organization.
- Goal: To develop a comprehensive SOC for your fire organization.

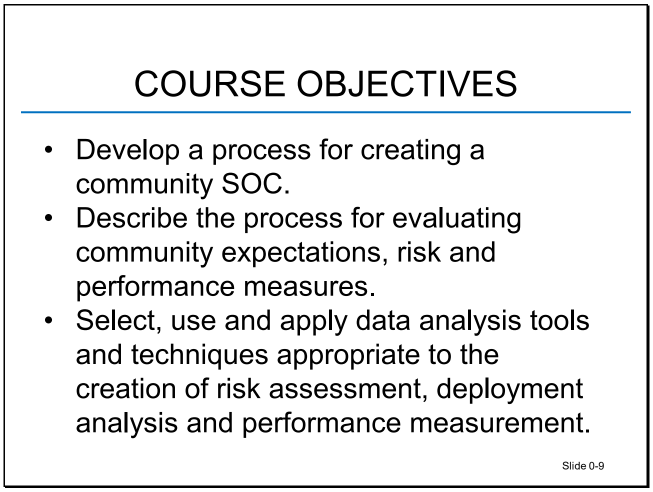
A. Course purpose, goal and objectives.

1. Purpose.

   The course purpose is to provide the skills and knowledge needed to develop an SOC for your fire organization.

2. Goal.

   The course goal is to develop a comprehensive SOC for your fire organization.



COURSE OBJECTIVES

- Develop a process for creating a community SOC.
- Describe the process for evaluating community expectations, risk and performance measures.
- Select, use and apply data analysis tools and techniques appropriate to the creation of risk assessment, deployment analysis and performance measurement.

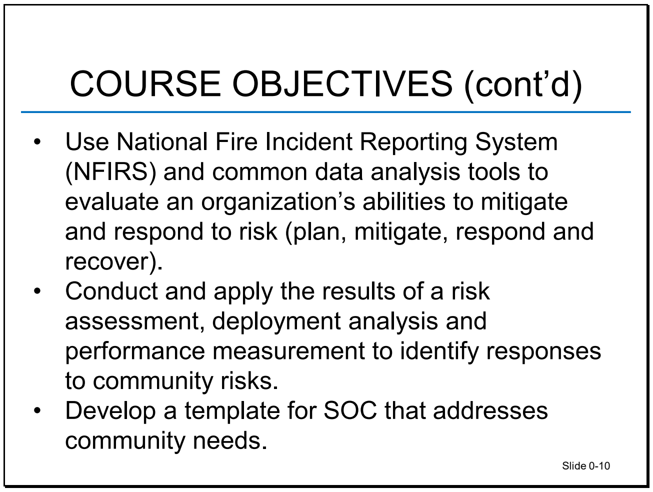
3. Objectives.

The students will be able to:

a. Develop a process for creating a community SOC.

b. Describe the process for evaluating community expectations, risk and performance measures.

c. Select, use and apply data analysis tools and techniques appropriate to the creation of risk assessment, deployment analysis and performance measurement.



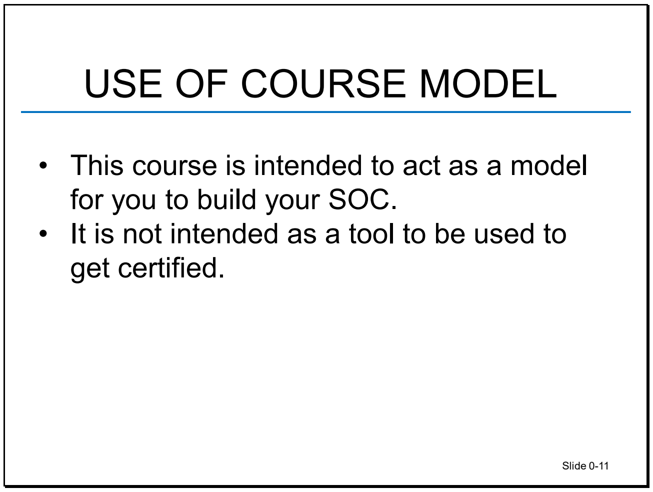
COURSE OBJECTIVES (cont'd)

- Use National Fire Incident Reporting System (NFIRS) and common data analysis tools to evaluate an organization's abilities to mitigate and respond to risk (plan, mitigate, respond and recover).
- Conduct and apply the results of a risk assessment, deployment analysis and performance measurement to identify responses to community risks.
- Develop a template for SOC that addresses community needs.

d. Use National Fire Incident Reporting System (NFIRS) and common data analysis tools to evaluate an organization's abilities to mitigate and respond to risk (plan, mitigate, respond and recover).

e. Conduct and apply the results of a risk assessment, deployment analysis and performance measurement to identify responses to community risks.

f. Develop a template for SOC that addresses community needs.



USE OF COURSE MODEL

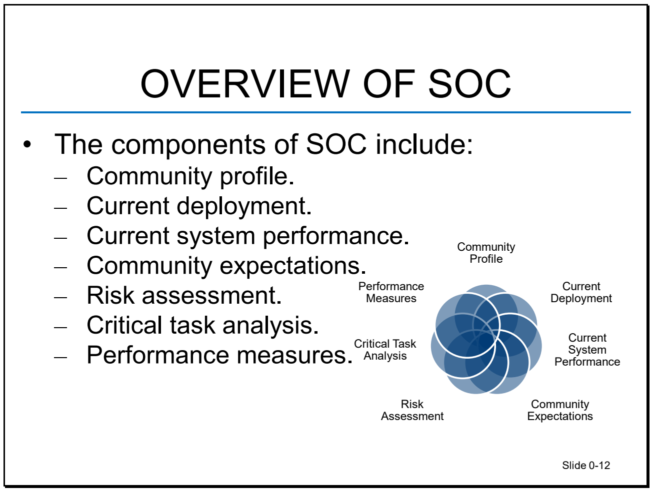
- This course is intended to act as a model for you to build your SOC.
- It is not intended as a tool to be used to get certified.

Slide 0-11

B. Use of course model.

This course is intended to act as a model for you to build your SOC. It is not intended as a tool to be used to get certified.

V. OVERVIEW OF SOC



OVERVIEW OF SOC

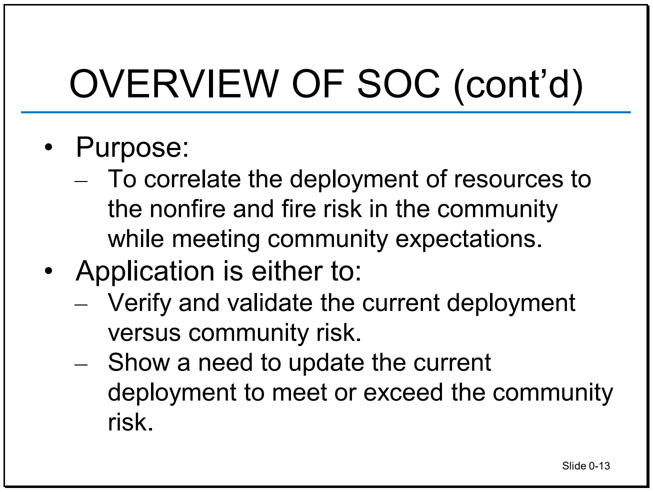
- The components of SOC include:
  - Community profile.
  - Current deployment.
  - Current system performance.
  - Community expectations.
  - Risk assessment.
  - Critical task analysis.
  - Performance measures.

Slide 0-12

Components, purpose and application of SOC

A. The components of SOC include:

1. Community profile.
2. Current deployment.
3. Current system performance.
4. Community expectations.
5. Risk assessment.
6. Critical task analysis.
7. Performance measures.



OVERVIEW OF SOC (cont'd)

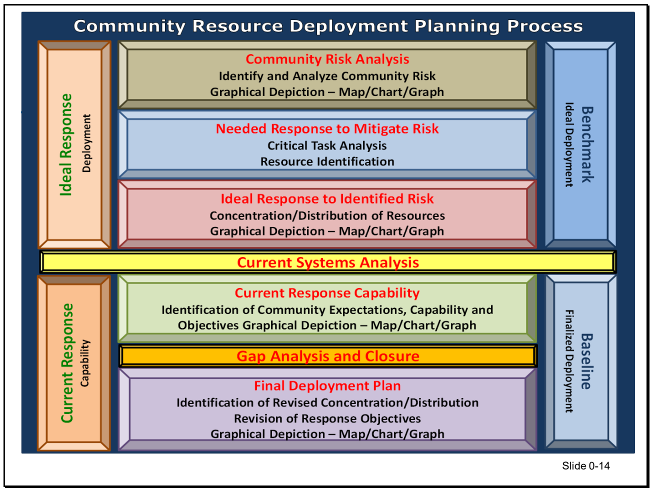
- Purpose:
  - To correlate the deployment of resources to the nonfire and fire risk in the community while meeting community expectations.
- Application is either to:
  - Verify and validate the current deployment versus community risk.
  - Show a need to update the current deployment to meet or exceed the community risk.

B. Purpose.

The purpose of creating an SOC is to correlate the deployment of resources to the nonfire and fire risk in the community with meeting community expectations.

C. Application of SOC is either to:

1. Verify and validate the current deployment versus community risk.
2. Show a need to update the current deployment to meet or exceed the community risk.



Community Resource Deployment Planning Process

Ideal Response
Deployment

Community Risk Analysis
Identify and Analyze Community Risk
Graphical Depiction – Map/Chart/Graph

Needed Response to Mitigate Risk
Critical Task Analysis
Resource Identification

Ideal Response to Identified Risk
Concentration/Distribution of Resources
Graphical Depiction – Map/Chart/Graph

Benchmark
Ideal Deployment

Current Systems Analysis

Current Response
Capability

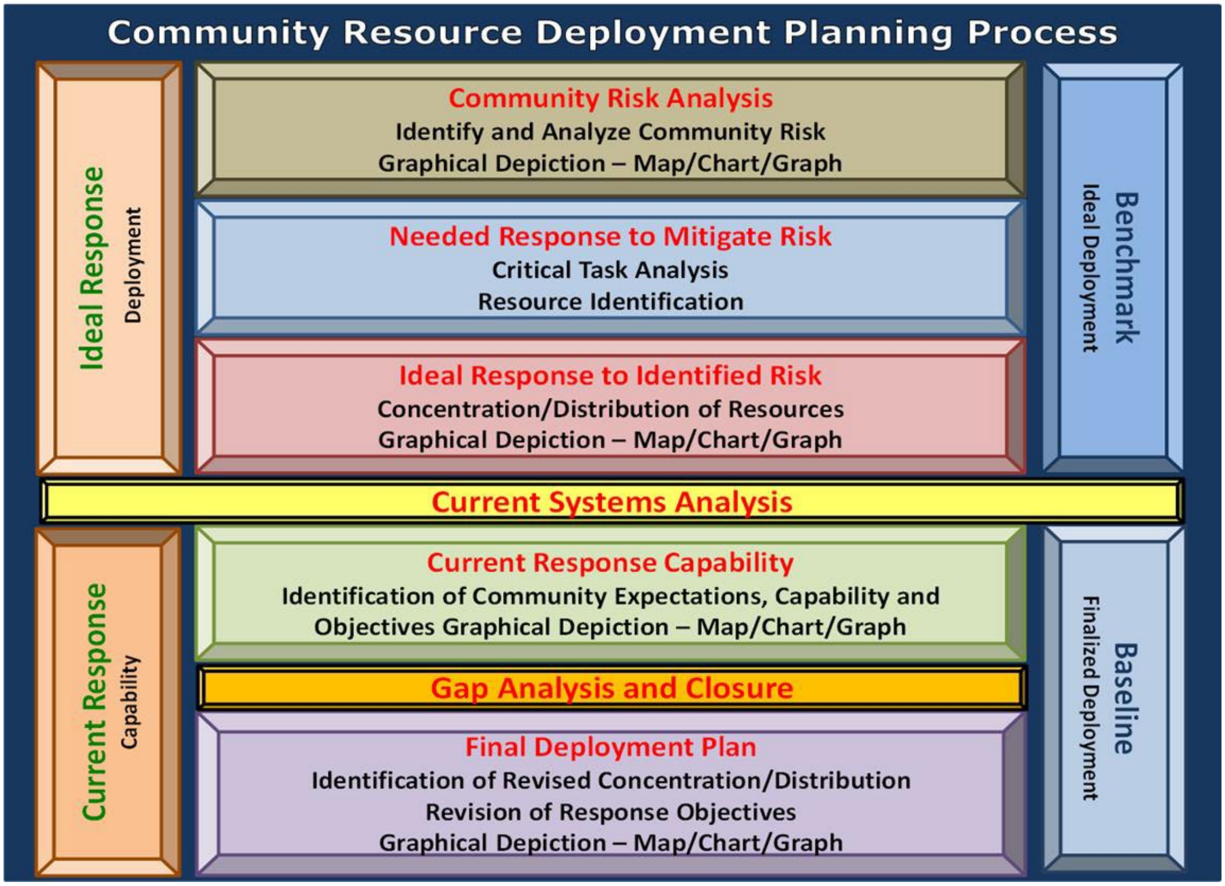
Current Response Capability
Identification of Community Expectations, Capability and Objectives Graphical Depiction – Map/Chart/Graph

Gap Analysis and Closure

Final Deployment Plan
Identification of Revised Concentration/Distribution
Revision of Response Objectives
Graphical Depiction – Map/Chart/Graph

Baseline
Finalized Deployment

Slide 0-14



Community Resource Deployment Planning Process

Ideal Response
Deployment

Community Risk Analysis
Identify and Analyze Community Risk
Graphical Depiction – Map/Chart/Graph

Needed Response to Mitigate Risk
Critical Task Analysis
Resource Identification

Ideal Response to Identified Risk
Concentration/Distribution of Resources
Graphical Depiction – Map/Chart/Graph

Benchmark
Ideal Deployment

Current Systems Analysis

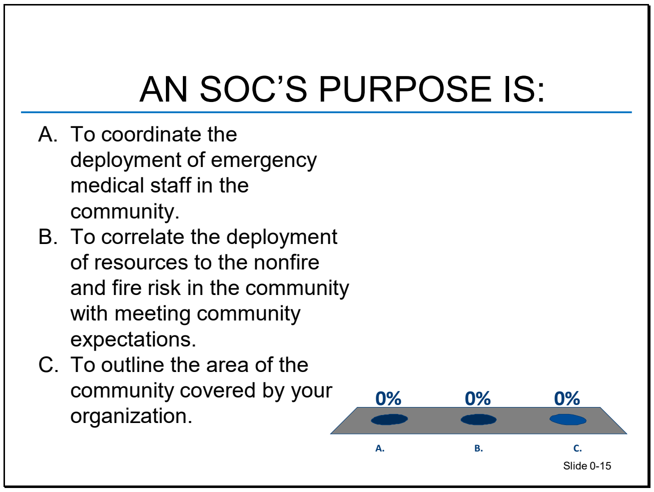
Current Response
Capability

Current Response Capability
Identification of Community Expectations, Capability and Objectives Graphical Depiction – Map/Chart/Graph

Gap Analysis and Closure

Final Deployment Plan
Identification of Revised Concentration/Distribution
Revision of Response Objectives
Graphical Depiction – Map/Chart/Graph

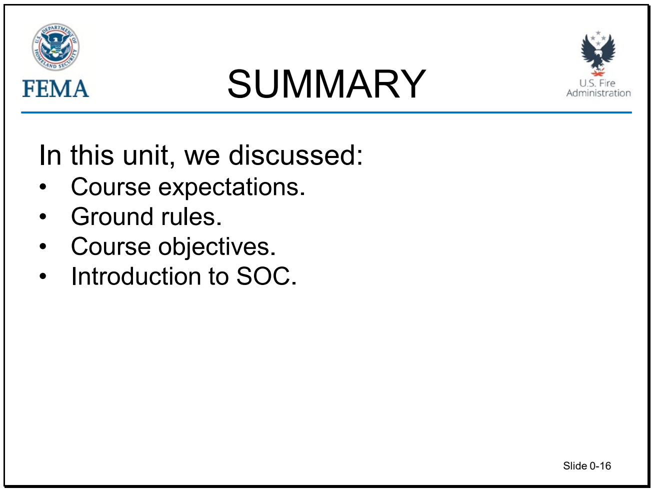
Baseline
Finalized Deployment



AN SOC'S PURPOSE IS:

A. To coordinate the deployment of emergency medical staff in the community.
B. To correlate the deployment of resources to the nonfire and fire risk in the community with meeting community expectations.
C. To outline the area of the community covered by your organization.

VI. SUMMARY



SUMMARY

In this unit, we discussed:

- Course expectations.
- Ground rules.
- Course objectives.
- Introduction to SOC.

UNIT 1:

EXPECTATIONS, STANDARDS OF COVER, PERFORMANCE MEASURES AND RISK

TERMINAL OBJECTIVES

The students will be able to:

1.1 At the conclusion of the unit of instruction, analyze stakeholder expectations, Standards of Cover (SOC) components and risk assessment factors required for system performance improvement.

ENABLING OBJECTIVES

The students will be able to:

1.1 Define SOC and its components.

1.2 Critique the component advantages of an SOC.

1.3 Describe core performance measures used in an SOC.

1.4 Apply study data/results for use in an SOC.

1.5 Discuss the elements of a community profile, including key stakeholders, demographics, and community specific elements, which may impact SOC.

1.6 Determine your community's expected level of service.

1.7 Appraise deployment of resources and compare emergency versus nonemergency deployments.

1.8 Interpret, document and incorporate community expectations in terms of performance and outcome measures for inclusion in an SOC.

1.9 *Compare performance measures and how they are used.*

1.10 *Explain the elements that must be included in developing a service-level objective.*

1.11 *Discuss risk and its components.*

1.12 *Explain risk assessment methodologies.*

1.13 *Employ data to prepare for use in a risk assessment for an SOC.*

1.14 *Discuss community risks and hazards as they relate to emergency services.*

1.15 *Explain the concept of strategic planning and how it relates to risk analysis.*

FEMA

U.S. Fire Administration

UNIT 1: EXPECTATIONS, STANDARDS OF COVER, PERFORMANCE MEASURES AND RISK

ENABLING OBJECTIVES

- Define Standards of Cover (SOC) and its components.
- Critique the component advantages of a SOC.
- Describe core performance measures used in an SOC.
- Apply study data/results for use in an SOC.

ENABLING OBJECTIVES (cont'd)

- Discuss the elements of a community profile, including key stakeholders, demographics, and community specific elements, which may impact SOC.
- Determine your community's expected level of service.
- Appraise deployment of resources and compare emergency versus nonemergency deployments.

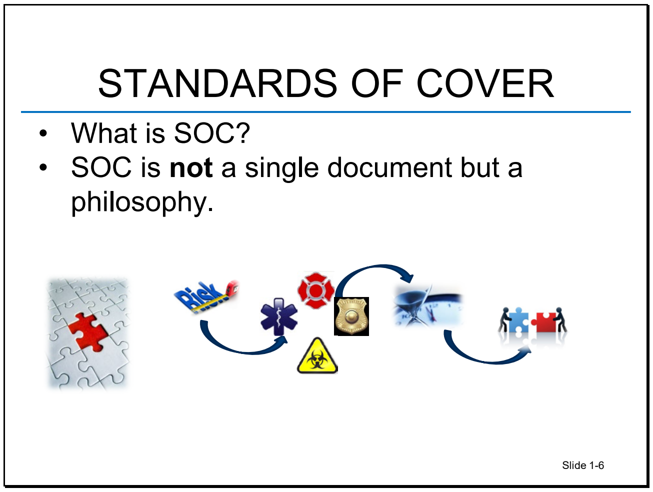
ENABLING OBJECTIVES (cont'd)

- Interpret, document and incorporate community expectations in terms of performance and outcome measures for inclusion in an SOC.
- Compare performance measures and how they are used.
- Explain the elements that must be included in developing a service-level objective.
- Discuss risk and its components.

ENABLING OBJECTIVES (cont'd)

- Explain risk assessment methodologies.
- Employ data to prepare for use in a risk assessment for an SOC.
- Discuss community risks and hazards as they relate to emergency services.
- Explain the concept of strategic planning and how it relates to risk analysis.

I. UNIT INTRODUCTION



STANDARDS OF COVER

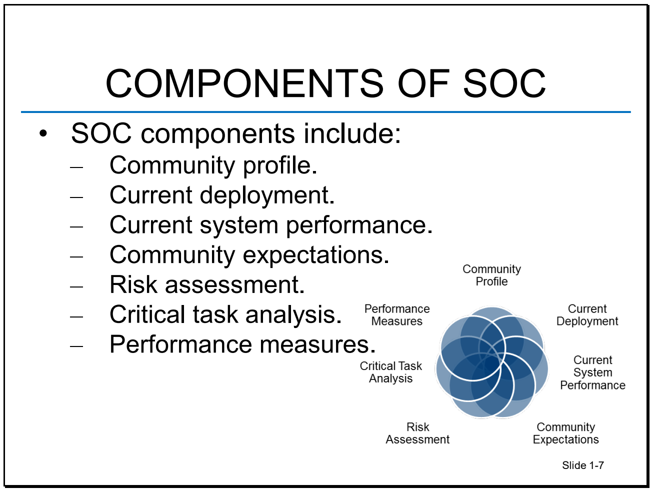
- What is SOC?
- SOC is **not** a single document but a philosophy.

II. STANDARDS OF COVER

A. What is an SOC?

1. **SOC** is a concept that allows for analysis of current deployment of resources compared to the risk assessment in the community.

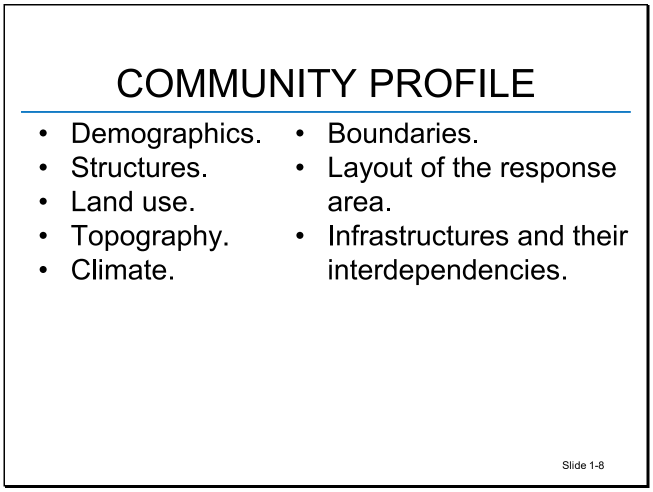
2. SOC is not a single document but a philosophy of meeting the risk with the right resources in a certain amount of time to mitigate the incident that has occurred.



COMPONENTS OF SOC

- SOC components include:
  - Community profile.
  - Current deployment.
  - Current system performance.
  - Community expectations.
  - Risk assessment.
  - Critical task analysis.
  - Performance measures.

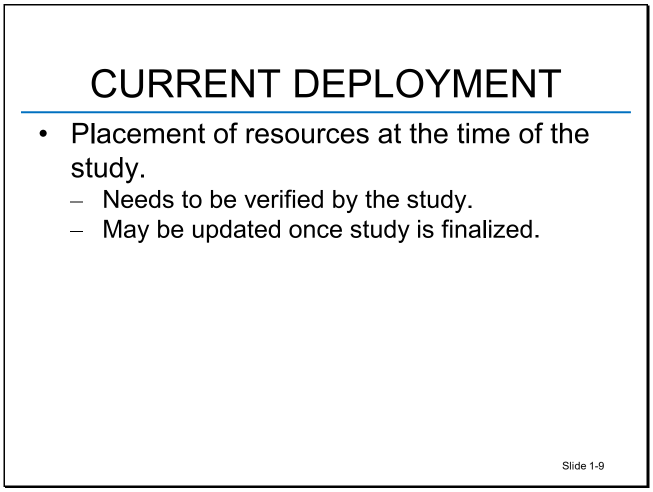
B. The components of SOC include: community profile, current deployment, current system performance, community expectations, risk assessment, critical task analysis, and performance measures.



COMMUNITY PROFILE

- Demographics.
- Structures.
- Land use.
- Topography.
- Climate.
- Boundaries.
- Layout of the response area.
- Infrastructures and their interdependencies.

1. **Community profile** is the demographics, structures, land use, topography, climate, boundaries, and layout of the response area, infrastructure, and its interdependencies.



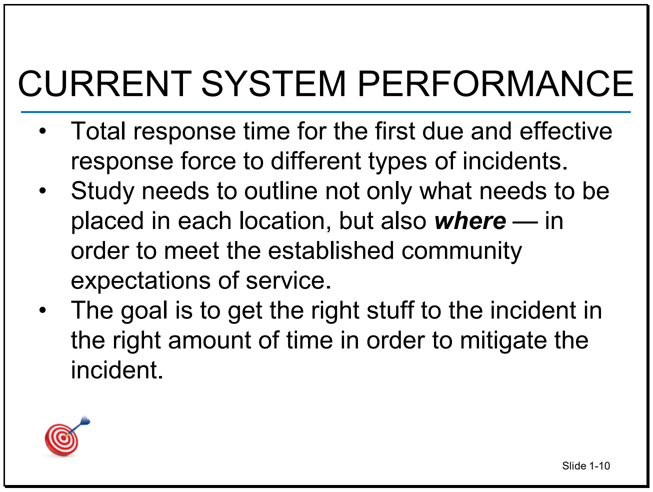
CURRENT DEPLOYMENT

- Placement of resources at the time of the study.
  - Needs to be verified by the study.
  - May be updated once study is finalized.

2. **Current deployment** is the placement of resources at the time of the study.

   a. This is the placement of the stations, personnel and equipment at the beginning of the study that will set the SOC.

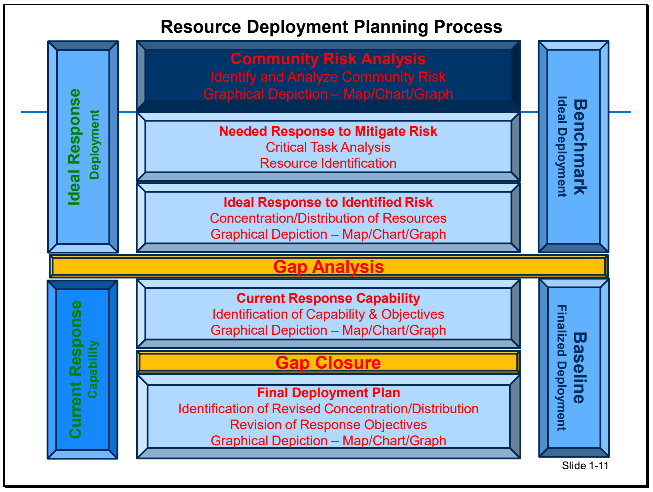
   b. This is what needs to be verified by the study; if resources need to be moved, that will occur after the conclusions have been reached and the study finalized.



CURRENT SYSTEM PERFORMANCE

- Total response time for the first due and effective response force to different types of incidents.
- Study needs to outline not only what needs to be placed in each location, but also ***where*** — in order to meet the established community expectations of service.
- The goal is to get the right stuff to the incident in the right amount of time in order to mitigate the incident.

3. **Current system performance** is the total response time for the first due and effective response force to different types of incidents.



Resource Deployment Planning Process

Ideal Response Deployment

Community Risk Analysis
Identify and Analyze Community Risk
Graphical Depiction – Map/Chart/Graph

Needed Response to Mitigate Risk
Critical Task Analysis
Resource Identification

Ideal Response to Identified Risk
Concentration/Distribution of Resources
Graphical Depiction – Map/Chart/Graph

Benchmark Ideal Deployment

Gap Analysis

Current Response Capability

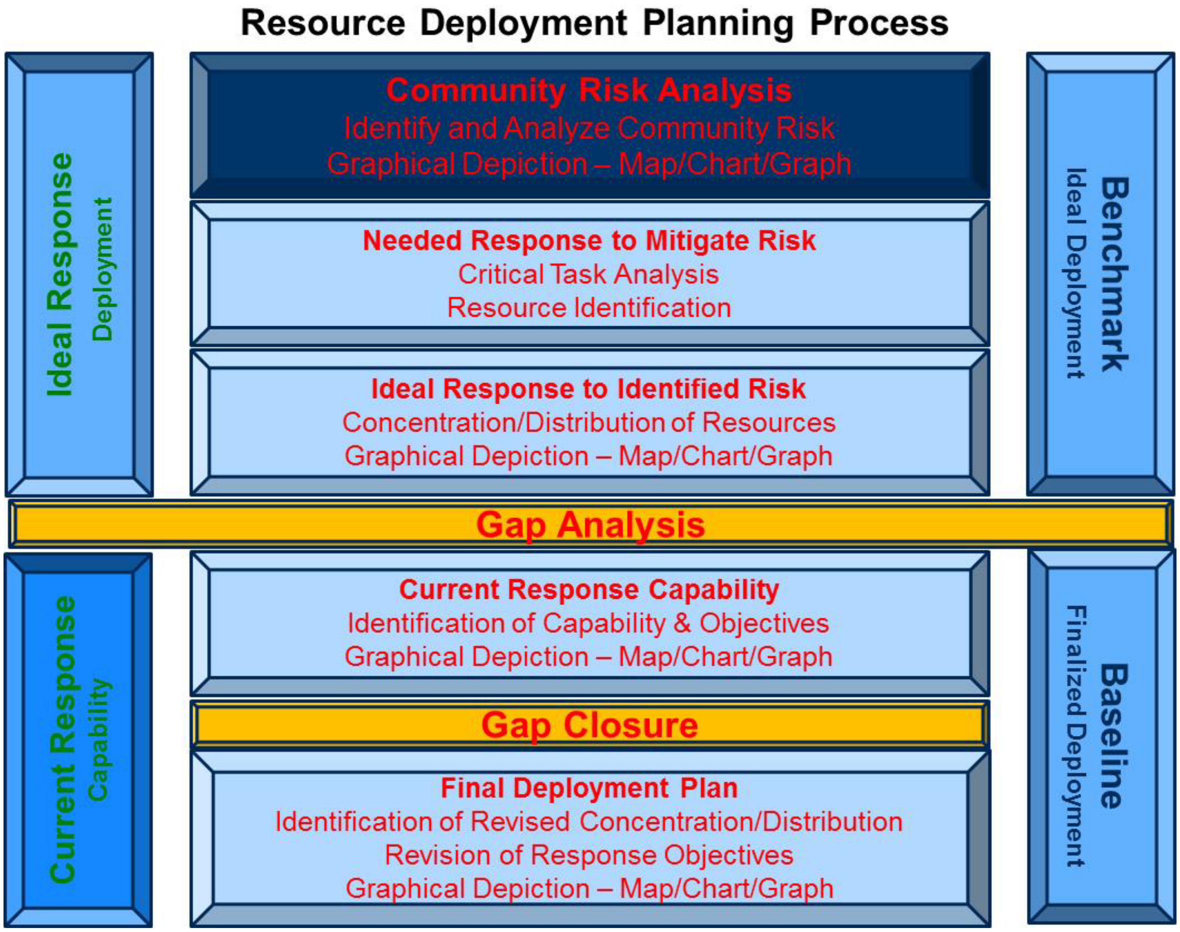
Current Response Capability
Identification of Capability & Objectives
Graphical Depiction – Map/Chart/Graph

Gap Closure

Final Deployment Plan
Identification of Revised Concentration/Distribution
Revision of Response Objectives
Graphical Depiction – Map/Chart/Graph

Baseline Finalized Deployment

Slide 1-11



Resource Deployment Planning Process

Ideal Response Deployment

Community Risk Analysis
Identify and Analyze Community Risk
Graphical Depiction – Map/Chart/Graph

Needed Response to Mitigate Risk
Critical Task Analysis
Resource Identification

Ideal Response to Identified Risk
Concentration/Distribution of Resources
Graphical Depiction – Map/Chart/Graph

Benchmark Ideal Deployment

Gap Analysis

Current Response Capability

Current Response Capability
Identification of Capability & Objectives
Graphical Depiction – Map/Chart/Graph

Gap Closure

Final Deployment Plan
Identification of Revised Concentration/Distribution
Revision of Response Objectives
Graphical Depiction – Map/Chart/Graph

Baseline Finalized Deployment

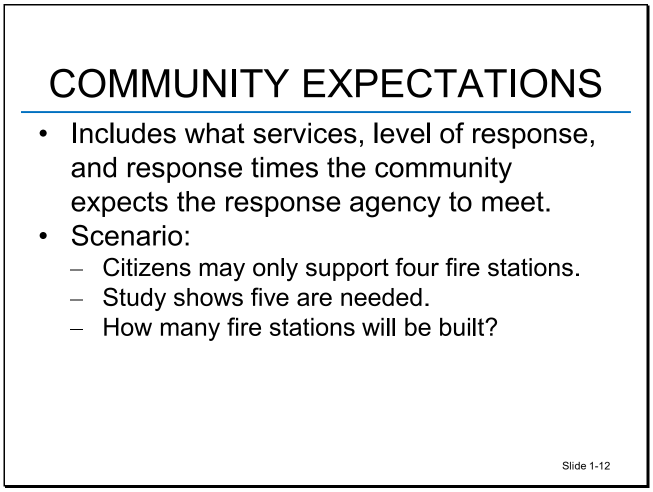
a. The performance of the resources is just as important as the placement of the resources.

b. The study needs to outline not only what needs to be placed in each location, but also where those resources need to be placed in order to meet the established community expectations for service.

c. The goal is to get the right resources to the incident in the right amount of time in order to mitigate the incident.

d. The risk assessment will identify where all of the various risks are located; once the risk is known, the resources can be placed in the area to respond more effectively to the identified risk.

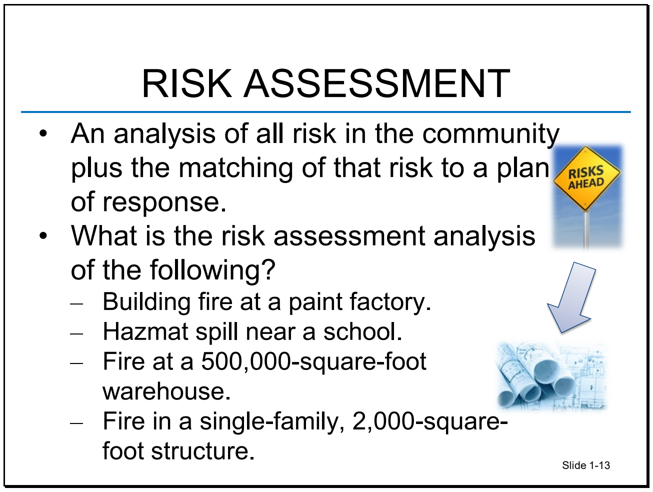
e. Once the right resources are in the right place to respond to the identified risk in a timely manner, you have created a proper deployment of resources.



COMMUNITY EXPECTATIONS

- Includes what services, level of response, and response times the community expects the response agency to meet.
- Scenario:
  - Citizens may only support four fire stations.
  - Study shows five are needed.
  - How many fire stations will be built?

4. **Community expectations** incorporate what services, level of response, and response times the community expects the response agency to meet.



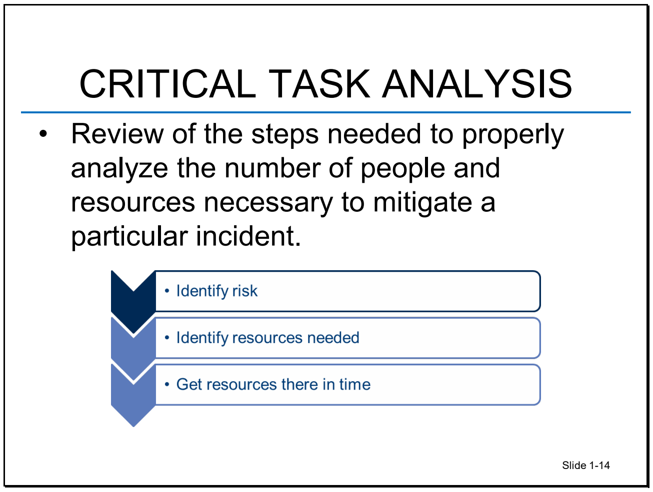
RISK ASSESSMENT

- An analysis of all risk in the community plus the matching of that risk to a plan of response.
- What is the risk assessment analysis of the following?
  - Building fire at a paint factory.
  - Hazmat spill near a school.
  - Fire at a 500,000-square-foot warehouse.
  - Fire in a single-family, 2,000-square-foot structure.

5. **Risk assessment** is an analysis of all the risk in the community plus the matching of that risk to a plan of response.

    a. This is a major component of the SOC.

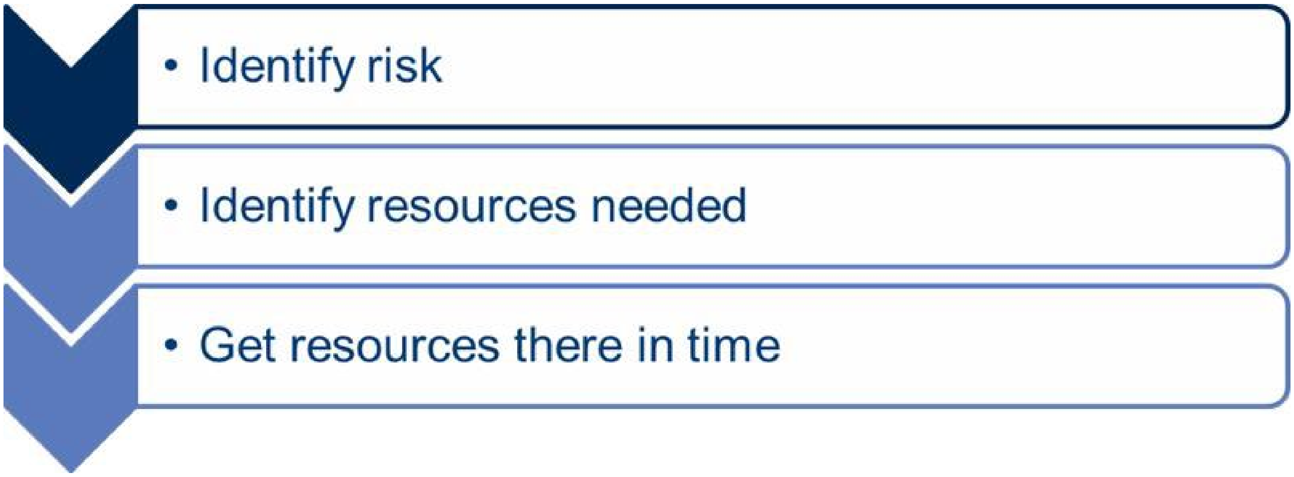
    b. There needs to be consideration of the risk associated with a car accident on a side street, on a major thoroughfare, or on the interstate.



CRITICAL TASK ANALYSIS

- Review of the steps needed to properly analyze the number of people and resources necessary to mitigate a particular incident.
    - Identify risk
    - Identify resources needed
    - Get resources there in time

Slide 1-14



6. **Critical task analysis** is a review of the steps needed to properly analyze the number of people and resources necessary to mitigate a particular incident.

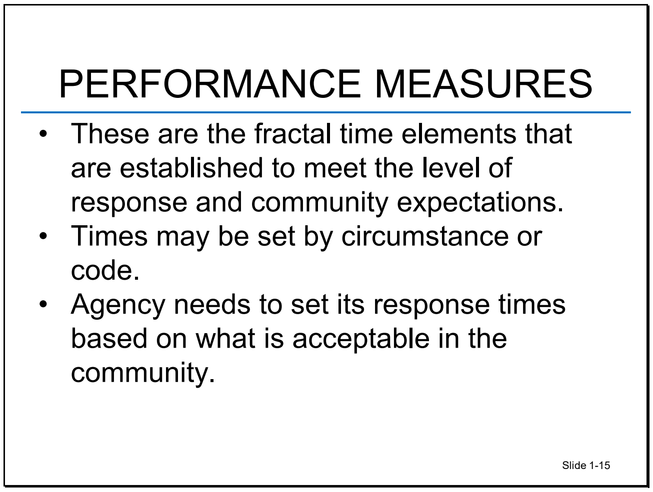
    a. The critical task analysis is a major piece of the study.

    b. The goal is to identify the risk, figure out the needed resources for that risk, and get the identified resources there in a timely manner.

c. The number of resources needed is determined by outlining the steps necessary to mitigate the situation.

d. Once the steps of mitigation are known, the number of personnel needed to carry out each task can be determined.

e. Once you know the number of personnel, you know how many apparatus are needed to respond to that incident.



PERFORMANCE MEASURES

- These are the fractal time elements that are established to meet the level of response and community expectations.
- Times may be set by circumstance or code.
- Agency needs to set its response times based on what is acceptable in the community.

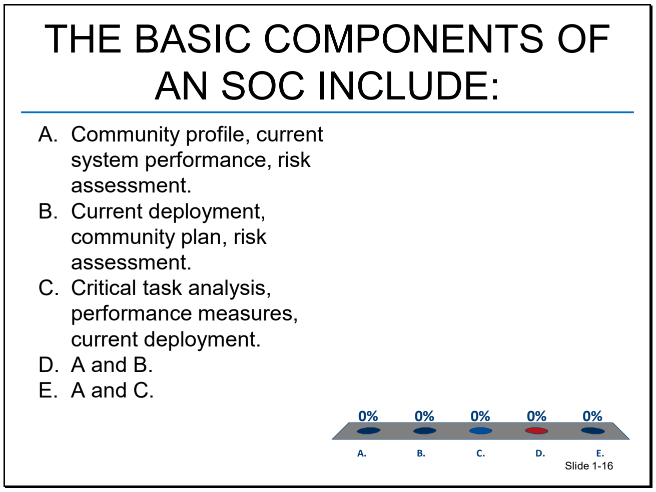
7. **Performance measures** are the fractal time elements that are established to meet the level of response and community expectations.

a. These are the times that are acceptable for a response to a certain type of incident — National Fire Protection Association (NFPA) 1710, *Standard for the Organization and Deployment of Fire Suppression Operations, Emergency Medical Operations, and Special Operations to the Public by Career Fire Departments*, sets the response times at:

- First due = four minutes.

- Effective response force = eight minutes.

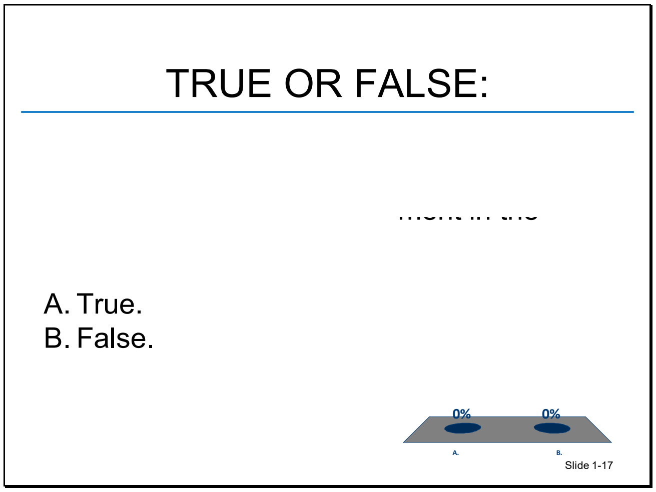
b. Accreditation sets the times based on type of service area — urban, suburban or rural. Each level has a different response time.



THE BASIC COMPONENTS OF AN SOC INCLUDE:

A. Community profile, current system performance, risk assessment.
B. Current deployment, community plan, risk assessment.
C. Critical task analysis, performance measures, current deployment.
D. A and B.
E. A and C.

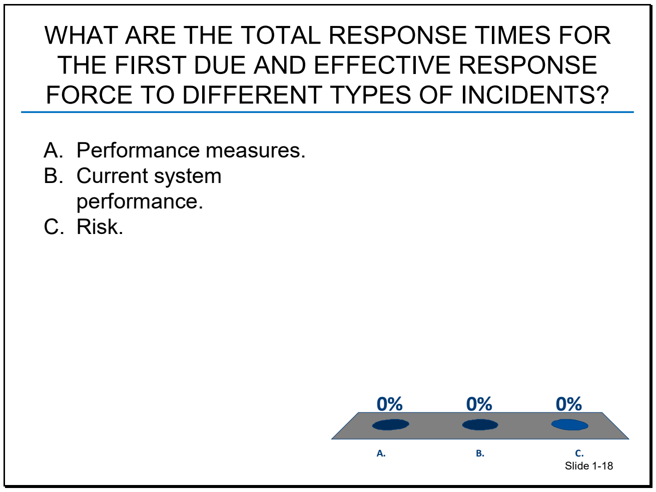
Slide 1-16



TRUE OR FALSE:

A. True.
B. False.

Slide 1-17



WHAT ARE THE TOTAL RESPONSE TIMES FOR THE FIRST DUE AND EFFECTIVE RESPONSE FORCE TO DIFFERENT TYPES OF INCIDENTS?

A. Performance measures.
B. Current system performance.
C. Risk.

Slide 1-18

This page intentionally left blank.

ACTIVITY 1.1

Understanding a Standards of Cover

Purpose

To identify the information required to begin development of an SOC.

Directions

1. You will work in groups; elect a group representative for debriefing.
2. Spend 10 to 15 minutes developing a list of needed information in order to begin the development of an SOC.
3. You will use the worksheet to record items.
4. Your group's representative will share your group's list with the class.
5. Keep this worksheet to use in subsequent units.

This page intentionally left blank.

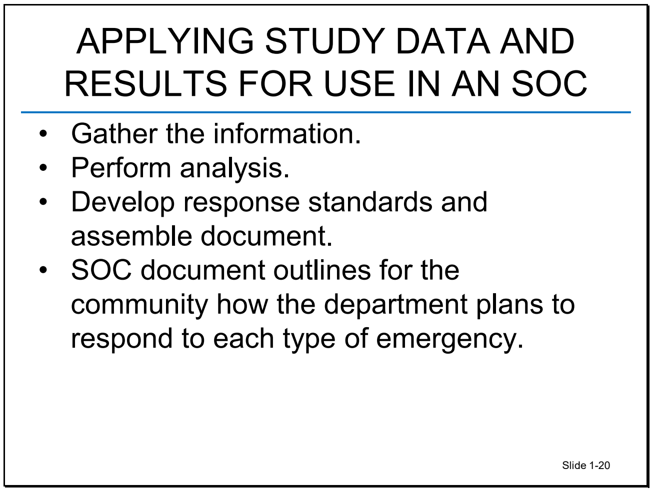
ACTIVITY 1.1 (cont'd)

Understanding a Standards of Cover

Information	Purpose for Using an SOC

This page intentionally left blank.

II. STANDARDS OF COVER (cont'd)



APPLYING STUDY DATA AND RESULTS FOR USE IN AN SOC

- Gather the information.
- Perform analysis.
- Develop response standards and assemble document.
- SOC document outlines for the community how the department plans to respond to each type of emergency.

C. Applying study data and results for use in an SOC.

1. Steps to incorporate data and results into your SOC include the following:

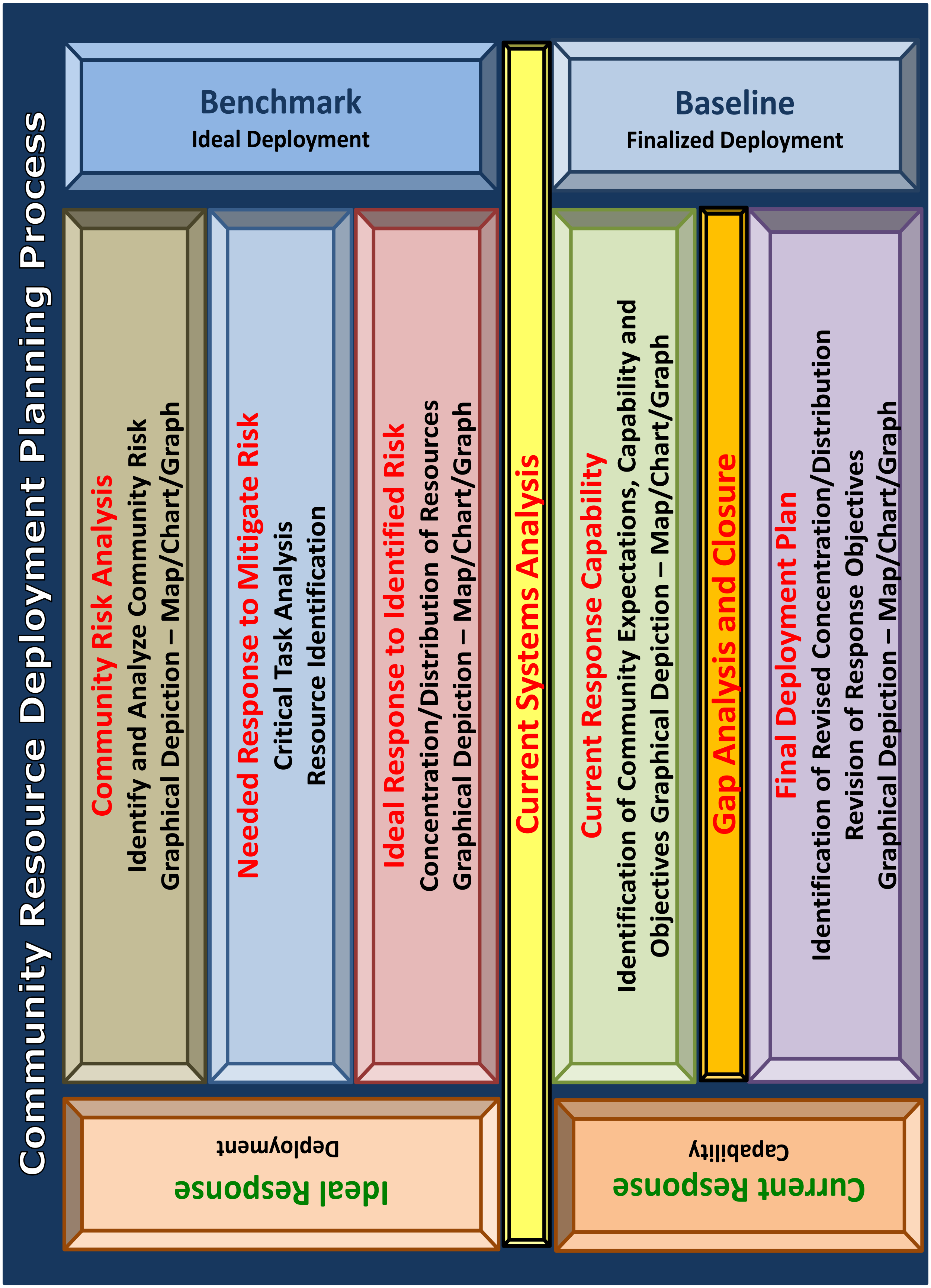
   a. Gathering the information together will help you start to develop a plan of how you are going to meet the SOC in your jurisdiction.

   b. Through analysis of the information, the agency should be able to develop a strong plan on how to respond to various emergencies.

   c. Once information is put together and response standards developed, the document will need to be assembled.

2. The SOC document will outline for the community how the department plans to respond to each type of emergency.

3. The final step will be to have the authority with jurisdiction approve the plan and commit to supporting that level of response.



Community Resource Deployment Planning Process

Benchmark
Ideal Deployment

Community Risk Analysis
Identify and Analyze Community Risk
Graphical Depiction – Map/Chart/Graph

Needed Response to Mitigate Risk
Critical Task Analysis
Resource Identification

Ideal Response to Identified Risk
Concentration/Distribution of Resources
Graphical Depiction – Map/Chart/Graph

Ideal Response
Deployment

Current Systems Analysis

Baseline
Finalized Deployment

Current Response Capability
Identification of Community Expectations, Capability and Objectives Graphical Depiction – Map/Chart/Graph

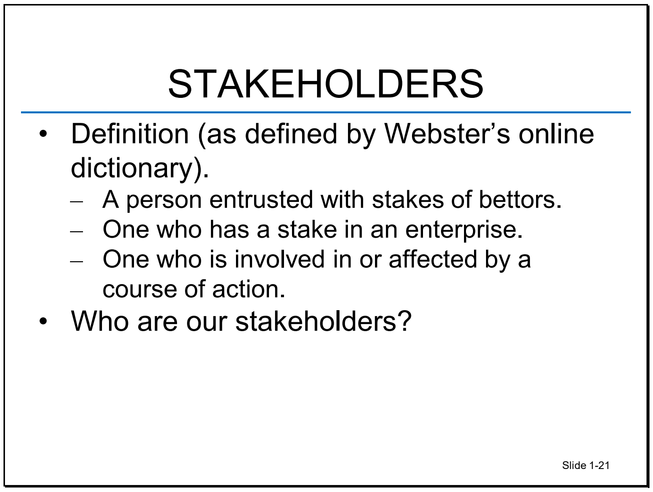
Gap Analysis and Closure

Final Deployment Plan
Identification of Revised Concentration/Distribution
Revision of Response Objectives
Graphical Depiction – Map/Chart/Graph

Current Response
Capability

III. COMMUNITY EXPECTATIONS

A. Introduction.



STAKEHOLDERS

- Definition (as defined by Webster's online dictionary).
  - A person entrusted with stakes of bettors.
  - One who has a stake in an enterprise.
  - One who is involved in or affected by a course of action.
- Who are our stakeholders?

1. According to Webster's online dictionary, a stakeholder can be any of the following:

   a. A person entrusted with the stakes of bettors.

   b. One who has a stake in an enterprise.

   c. One who is involved in or affected by a course of action.

2. Stakeholders can be any of the following:

   a. Citizens.

   b. Businesses.

   c. Employees.

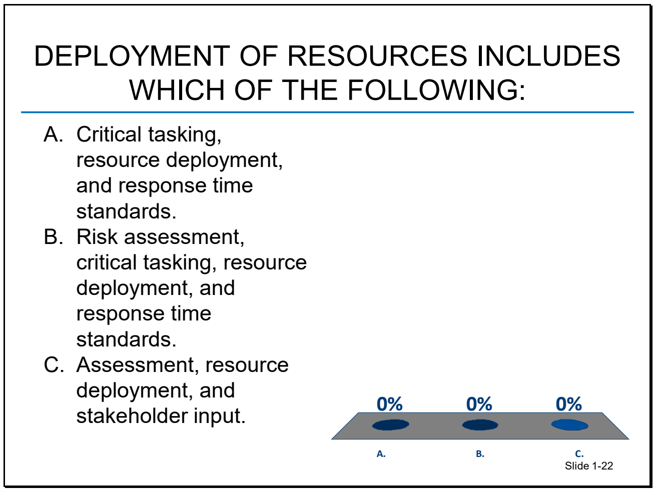
   d. Visitors.

   e. Media.

   f. Other governments.

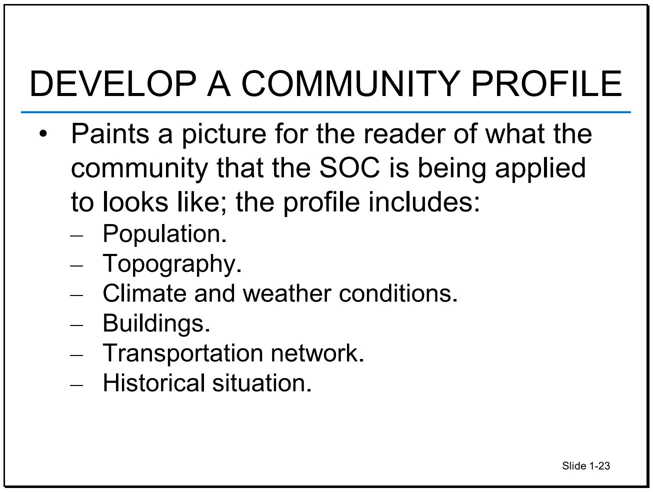
   g. Governing body.

3. Stakeholders can be users, participants, or individuals, or organizations affected by the work that is done by the fire department.



DEPLOYMENT OF RESOURCES INCLUDES WHICH OF THE FOLLOWING:

A. Critical tasking, resource deployment, and response time standards.
B. Risk assessment, critical tasking, resource deployment, and response time standards.
C. Assessment, resource deployment, and stakeholder input.



DEVELOP A COMMUNITY PROFILE

- Paints a picture for the reader of what the community that the SOC is being applied to looks like; the profile includes:
  - Population.
  - Topography.
  - Climate and weather conditions.
  - Buildings.
  - Transportation network.
  - Historical situation.

B. Develop a community profile: a community profile paints a picture for the reader of what the community that the SOC is being applied to looks like. The profile must include information on:

1. Population:

   a. Census population.

   b. Workday versus bedtime.

   c. Heavy visitor load.

   d. Hotel population.

2. Topography:

   a. Square miles.

    b. Natural obstacles.

3. Climate and weather condition issues:

    a. Snow.

    b. Rain.

    c. Ice.

    d. Tornadoes.

    e. Hurricanes.

    f. High wind.

    g. Flooding.

    h. Droughts.

4. Buildings of the area:

    a. How much residential?

    b. How much business?

    c. How much commercial?

    d. How much industrial?

    e. How much undeveloped?

5. Transportation network:

    a. Airport.

    b. Train tracks.

    c. Highways.

    d. Roadways.

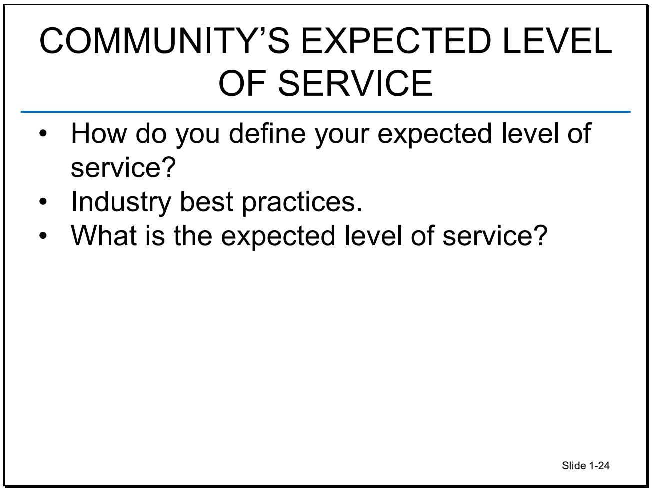
    e. Waterways.

6. Historical situation of the department:

   a. Legal basis for the department.

   b. How the department began.

   c. Major milestones in development.



COMMUNITY’S EXPECTED LEVEL OF SERVICE

- How do you define your expected level of service?
- Industry best practices.
- What is the expected level of service?

Slide 1-24

C. Community’s expected level of service.

Potential ways to define expected level of service include using:

1. Focus groups:

   a. Most accurate way for defining expected level of service.

   b. Allows for interaction with the citizens.

   c. Must provide them with accepted standards for industry.

2. Written survey:

   a. Easier, but less accurate.

   b. No interaction with people, will see only what they write.

   c. Still needs to provide accepted standards.

   d. Still needs to be random selection.

3. Industry best practices include:

   a. NFPA 1710, NFPA 1720, *Standard for the Organization and Deployment of Fire Suppression Operations, Emergency Medical Operations, and Special Operations to the Public by Volunteer Fire Departments*, and NFPA 1221, *Standard for the Installation, Maintenance, and Use of Emergency Services Communications Systems*.

   b. Commission of Fire Accreditation International (CFAI) Response Standards.

   c. National Institute of Standards and Technology (NIST) Fireground Field Study.

   d. Survey similarly situated fire departments.

      - Same population, number of stations, or response area.

   e. Look at 10 departments and come up with service-level objectives.

      - The easiest method: look back at the last several years of response data and determine what your response has been in the past without citizens complaining. This could be your acceptable level of response.

   f. Submit the results to your focus group for citizen input.

4. Level of service covers many items, including:

   a. Number of stations.

   b. Number of engines.

   c. Number of aerial devices.

   d. Number of rescues/ambulances.

   e. Number of specialty vehicles.

   f. Number of personnel on each of these apparatus.

   g. Number of apparatus for each type of incident.

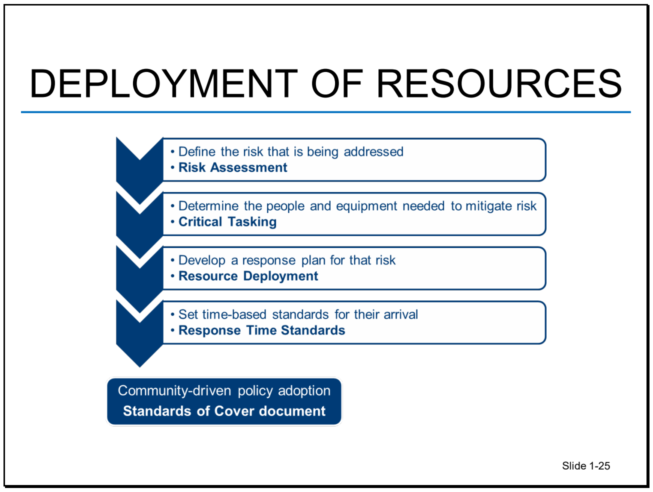
   h. Dispatch handling time.

   i. Turn-out time.

j. Travel time for first due unit.

k. Travel time for effective response force.

5. The idea of level of service is to specify how many personnel and how many pieces of equipment are needed to mitigate the situation that is being addressed.



DEPLOYMENT OF RESOURCES

- Define the risk that is being addressed
- **Risk Assessment**

- Determine the people and equipment needed to mitigate risk
- **Critical Tasking**

- Develop a response plan for that risk
- **Resource Deployment**

- Set time-based standards for their arrival
- **Response Time Standards**

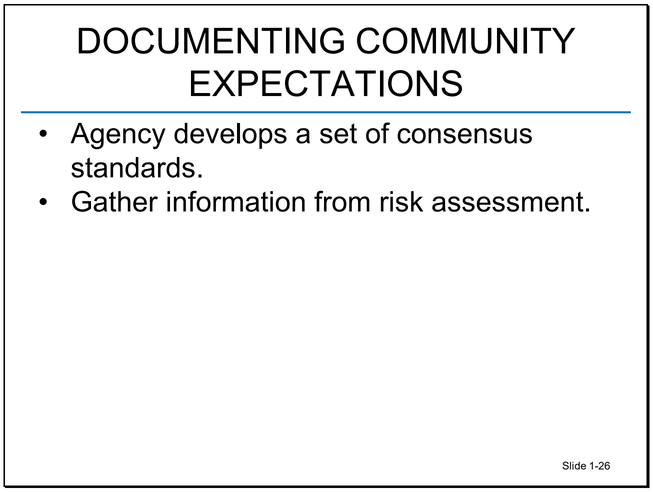
Community-driven policy adoption
Standards of Cover document

D. Deployment of resources.

1. Deployment of resources involves understanding how and why you place your stations, apparatus and personnel where you do when you do.

2. This is not a tactics course, but the aerial device needs to be placed as close as possible to the area where the need for the aerial has been identified, based on the risk assessment that the agency has conducted.

3. The resources need to be placed where they can respond best to meet the first due and effective response force needs identified through the risk assessment and critical tasking. Put your equipment where you need it when you need it.



DOCUMENTING COMMUNITY EXPECTATIONS

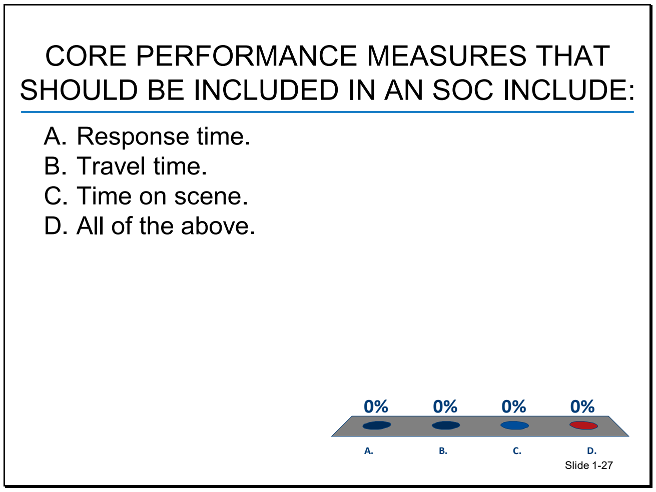
- Agency develops a set of consensus standards.
- Gather information from risk assessment.

Slide 1-26

E. Documenting community expectations.

1. Your agency develops a set of consensus standards that you have determined to be acceptable.

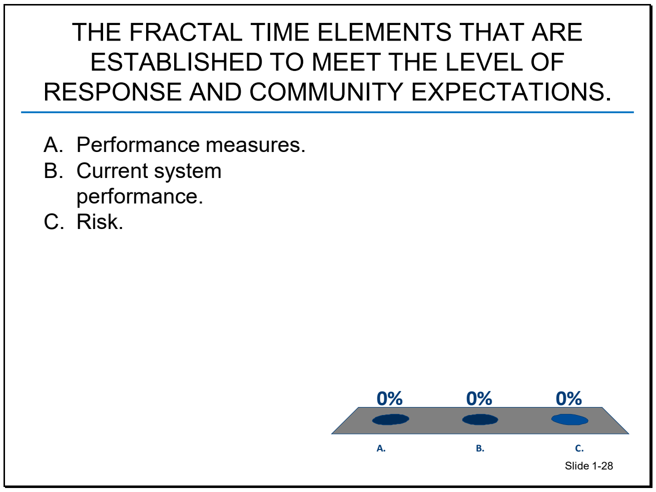
2. Gather the information from the risk assessment, critical tasking, and community expectation, and develop the standards that your agency will use as your SOC.



CORE PERFORMANCE MEASURES THAT SHOULD BE INCLUDED IN AN SOC INCLUDE:

A. Response time.
B. Travel time.
C. Time on scene.
D. All of the above.

Slide 1-27



THE FRACTAL TIME ELEMENTS THAT ARE ESTABLISHED TO MEET THE LEVEL OF RESPONSE AND COMMUNITY EXPECTATIONS.

A. Performance measures.
B. Current system performance.
C. Risk.

ACTIVITY 1.2

Community Expected Level of Service

Purpose

To begin to develop the community expectations outline for an SOC.

Directions

1. You will work in groups; elect a group representative for debriefing.
2. Spend 15 to 20 minutes developing an outline for an SOC.
3. You will use the worksheet to record items.
4. Your group's representative will share your group's outline with the class.
5. Keep this worksheet to use in subsequent units.

This page intentionally left blank.

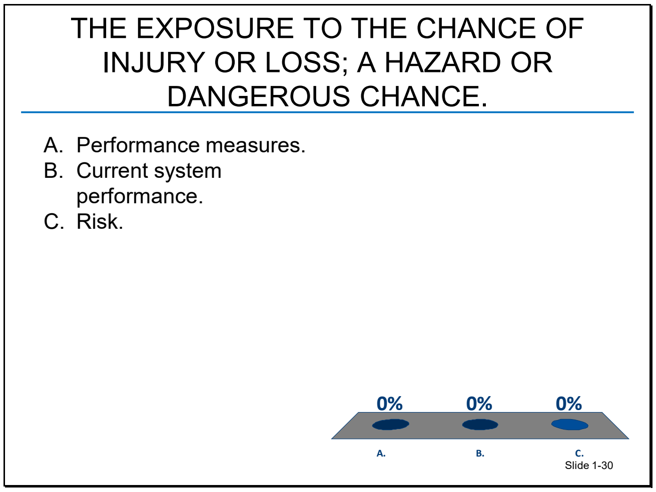
ACTIVITY 1.2 (cont'd)

Community Expected Level of Service

Heading	Subheadings

This page intentionally left blank.

IV. PERFORMANCE MEASURES



THE EXPOSURE TO THE CHANCE OF INJURY OR LOSS; A HAZARD OR DANGEROUS CHANCE.

A. Performance measures.
B. Current system performance.
C. Risk.

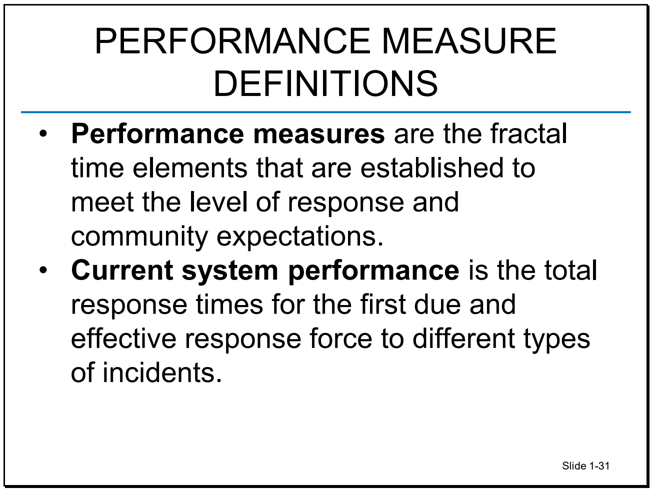
Slide 1-30

A. Introduction.

National/International expectations:

1. These are the only organizations that contain expectations on department response and staffing — NFPA, CFAI and Incident Safety Officer (ISO).

2. Community expectations should be the priority.



PERFORMANCE MEASURE DEFINITIONS

- **Performance measures** are the fractal time elements that are established to meet the level of response and community expectations.
- **Current system performance** is the total response times for the first due and effective response force to different types of incidents.

Slide 1-31

B. What are performance measures?

1. **Performance measures** are the fractal time elements that are established to meet the level of response and community expectations.

   a. These are the times that are acceptable for a response to a certain type of incident.

        b. NFPA 1710 sets the response at four minutes for first due and eight minutes for effective response force.

        c. Accreditation sets the times based on type of service area — urban, suburban and rural; each level has a different response time.

        d. The important thing here is that the agency needs to set its response times based on what is acceptable in its community.

    2. **Current system performance** is the total response times for the first due and effective response force to different types of incidents.

        a. The performance of the resources is just as important as the placement of the resources.

        b. The study needs to outline not only what needs to be placed in each location, but also where those resources are placed, in order to meet the community expectations for services that have been established.

        c. The goal is to get the right stuff to the incident in the right amount of time in order to mitigate the incident.

        d. The risk assessment will identify where all of the various risks are located; once the risk is known, the resources can be placed in the area to respond better to the identified risk.

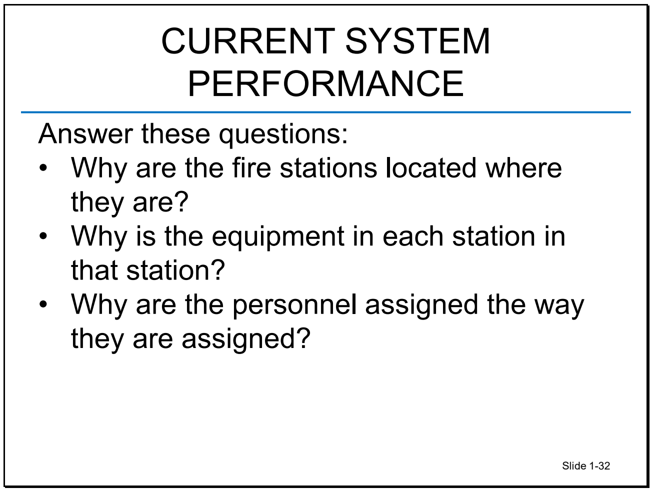
        e. Once the right resources are in the right place to respond to the identified risk in a timely manner, you have created a proper deployment of resources.

C. How are performance measures used?

    1. SOC is not about a quick response to a call or about where to put fire stations and trucks.

    2. Failure to perform a proper risk assessment is the reason that fire departments have difficulty explaining why their resources are so important.

    3. Elected officials question why a fire department needs more people and equipment than it has because the fire service has failed to tie the request to an analysis that shows a need.



CURRENT SYSTEM PERFORMANCE

Answer these questions:

- Why are the fire stations located where they are?
- Why is the equipment in each station in that station?
- Why are the personnel assigned the way they are assigned?

4. **With current system performance** you are **trying to figure out why you are doing what you are doing**.

5. Answer these questions:

   a. Why are the fire stations located where they are?

   b. Why is the equipment in each station in that station?

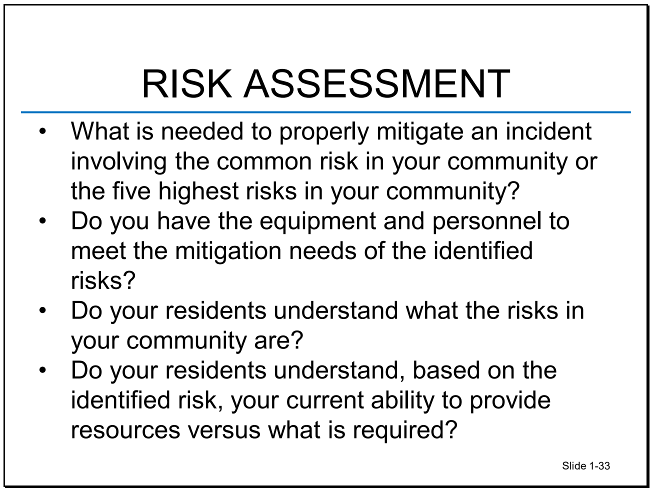
   c. Why are the personnel assigned the way they are assigned?

6. When you study how you got to be where you are today, you will learn of some pretty smart reasons that this station was placed here or there. You will also learn of historical community expectations.

7. What services do you provide to your community today?

8. A full explanation of the services you provide to your community and what they mean is an important step in this process.

9. It is the department's responsibility to determine what the community expectations have been in the past and whether or not they need to be re-evaluated.



RISK ASSESSMENT

- What is needed to properly mitigate an incident involving the common risk in your community or the five highest risks in your community?
- Do you have the equipment and personnel to meet the mitigation needs of the identified risks?
- Do your residents understand what the risks in your community are?
- Do your residents understand, based on the identified risk, your current ability to provide resources versus what is required?

10. **Risk assessment** is an analysis of your response area in order to determine what types of risk you may respond to on any given day. This is usually done using real-world factors, specifically those that the agency and the community have agreed upon as representing the community's risk level. The following questions must be asked to gain proper insight:

    a. What will you need (equipment and personnel) to properly mitigate an incident involving the common risk in your community?

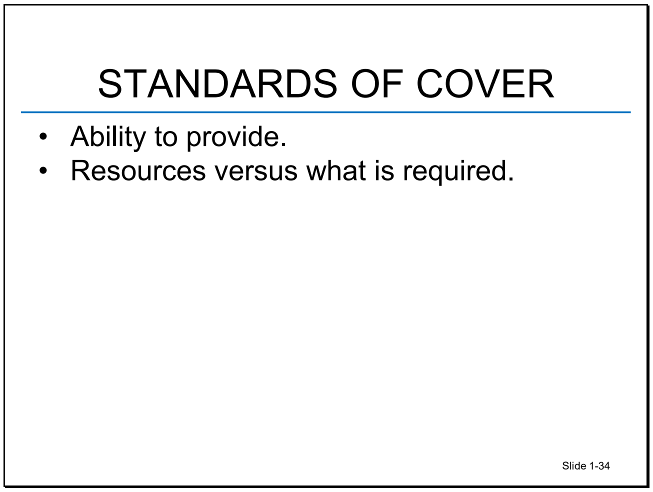
    b. What will you need (equipment and personnel) to properly mitigate an incident involving the five highest risks in your community?

    c. Do you have the equipment and personnel to meet the mitigation needs of the identified risks?

    d. Do your residents understand what the risks in your community are, how you describe the different levels of risk, and how much equipment and personnel you will need to mitigate the various levels of risk?

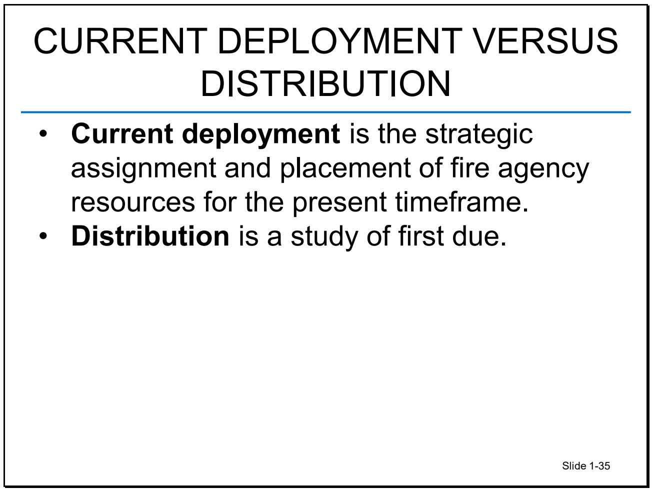
    e. Do your residents understand, based on the identified risk, your current ability to provide resources versus what is required?

11. These questions are answered through the concept of SOC.



STANDARDS OF COVER

- Ability to provide.
- Resources versus what is required.



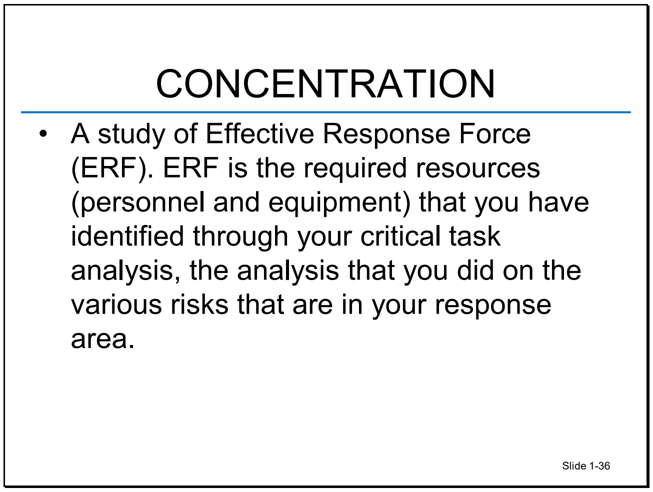
CURRENT DEPLOYMENT VERSUS DISTRIBUTION

- **Current deployment** is the strategic assignment and placement of fire agency resources for the present timeframe.
- **Distribution** is a study of first due.

12. **Current deployment** is the strategic assignment and placement of fire agency resources (such as fire companies, fire stations, and specific staffing levels for those companies) for the present timeframe (i.e., within the last calendar or budget year).

13. **Distribution** is a study of first due. What is the timeframe that your first due has been arriving on calls over the past three to five years?

14. Distribution is an important data set to look at because this is a study of the initial interventions, initial size-up, first shot of mitigation performance, etc.



CONCENTRATION

- A study of Effective Response Force (ERF). ERF is the required resources (personnel and equipment) that you have identified through your critical task analysis, the analysis that you did on the various risks that are in your response area.

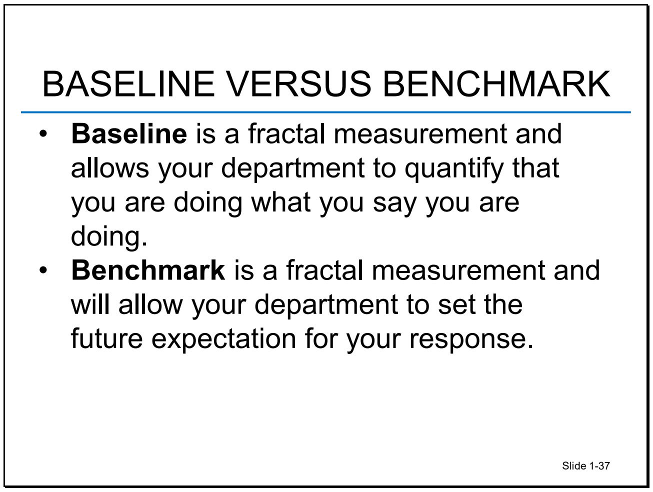
Slide 1-36

15. **Concentration** is a study of Effective Response Force (ERF). ERF is the required resources (personnel and equipment) that you have identified through your critical task analysis, the analysis that you did on the various risks that are in your response area.

16. Concentration is an important area of study; you are fighting a losing battle if you are not matching appropriate resources to the resource needs that you have identified through your critical task analysis for the identified risk.

17. If your highest identified risk critical tasking requires 26 personnel in the first 10 minutes of the incident and you are only matching that with 14 personnel in the first 10 minutes, you are not stopping the forward progress of the incident.

18. Once all of the above has been considered, it is time to quantify (measure) what you are doing and where your department is headed. It is now time to develop organizational expectations. Such expectations are addressed by risk and incident, type and within two timeframes: baseline (current) and benchmark (future).



BASELINE VERSUS BENCHMARK

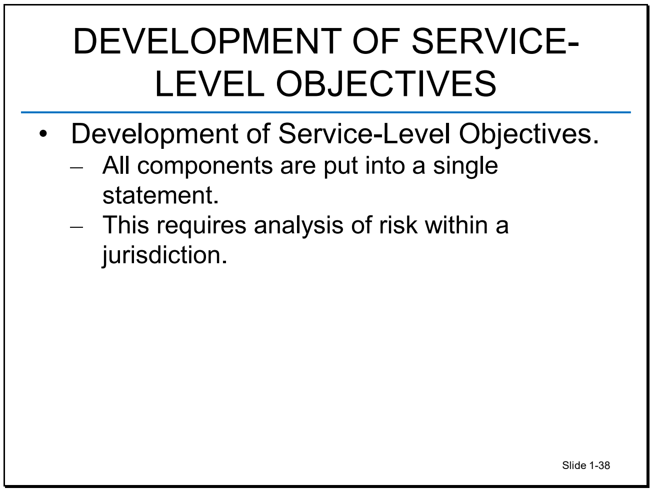
- **Baseline** is a fractal measurement and allows your department to quantify that you are doing what you say you are doing.
- **Benchmark** is a fractal measurement and will allow your department to set the future expectation for your response.

19. **Baseline** (today) performance expectation for first due and effective response force for suppression, Emergency Medical Services (EMS), Hazmat, and Tech rescue. This is a fractal measurement and allows your department to quantify that you are doing what you say you are doing. The numbers are in minutes and seconds and are based on all of the information that you have gathered.

20. **Benchmark** (future) performance expectation for first due and effective response force for suppression, EMS, Hazmat, and Tech rescue. This is also a fractal measurement and will allow your department to set the future expectation for your response.

21. Once all of this information is put together, you can determine whether you are sending the right equipment to the identified type of risk within the identified timeframe and how often you are meeting those requirements.

D. Core performance measures in an SOC.



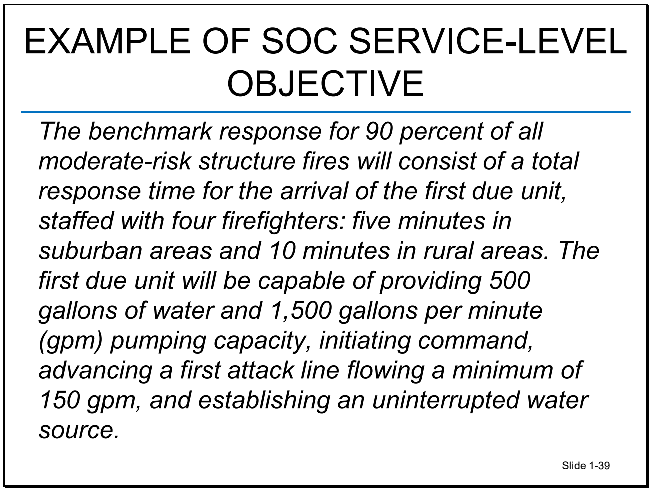
DEVELOPMENT OF SERVICE-LEVEL OBJECTIVES

- Development of Service-Level Objectives.
  - All components are put into a single statement.
  - This requires analysis of risk within a jurisdiction.

The SOC must contain a service-level objective.

1. It is in the SOC that all of the components are put into a single statement and the level of service that will be provided in a certain circumstance is stated.

2. When all of the pieces are put together, the objective should look like this example:



EXAMPLE OF SOC SERVICE-LEVEL OBJECTIVE

The benchmark response for 90 percent of all moderate-risk structure fires will consist of a total response time for the arrival of the first due unit, staffed with four firefighters: five minutes in suburban areas and 10 minutes in rural areas. The first due unit will be capable of providing 500 gallons of water and 1,500 gallons per minute (gpm) pumping capacity, initiating command, advancing a first attack line flowing a minimum of 150 gpm, and establishing an uninterrupted water source.

Slide 1-39

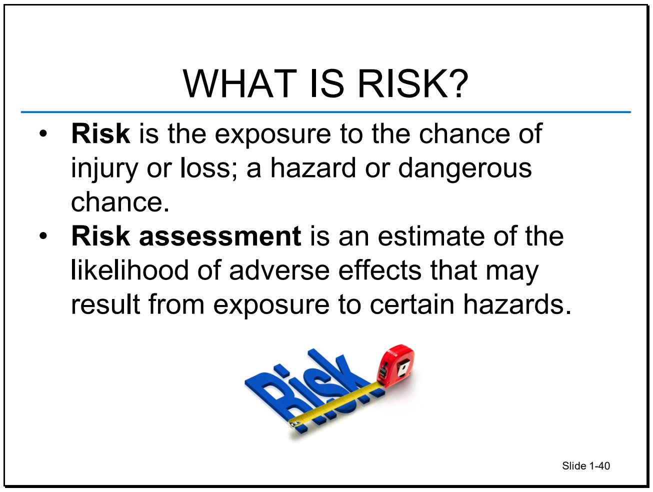
3. *The benchmark response for 90 percent of all moderate-risk structure fires will consist of a total response time for the arrival of the first due unit, staffed with four firefighters; it will be five minutes in suburban areas and 10 minutes in rural areas. The first due unit will be capable of providing 500 gallons of water and 1,500 gallons per minute (gpm) pumping capacity; initiating command; advancing a first attack line flowing a minimum of 150 gpm; and establishing an uninterrupted water source.*

Core Performance Measures
Service-Level Objective Checklist

To help identify core performance measures used in the SOC Service-Level Objective, address the following questions:

1. **What is the type of risk?** Is this a high-, moderate-, or low-risk building, or is this a special-risk building such as a dynamite factory?
2. **What is the type of community?** Is it an urban area, a suburban area, a rural area, or wildland interface? What is the population density?
3. **Is the performance measure for a first due statement (distribution of resources)?** Distribution is the placement of resources throughout the community to get the first unit there in a timely manner. The concentration is the measure of how many resources need to be put how close together to respond to mitigate the incident, and of how to get the effective response force there in a timely manner.
4. **Is the performance measure for an ERF statement (concentration of resources)?**
5. **Is it a benchmark statement (goal)?** A benchmark statement is a goal that an agency is trying to reach — where the agency wants to be in the future.
6. **Is it a baseline statement (minimal acceptable performance level of the agency)?** This is the response that has to be met for the agency to be successful at what it is doing.
7. **What are the critical tasks that have to be completed by personnel identified in the objective?** How many firefighters or responders are there going to be on the apparatus that is responding to mitigate the situation? Is the first due staffed with four firefighters or three? Is the effective response force 15 or 21? The objective needs to have this information in it. Is the first due unit going to establish command, lay a hoseline, or establish a water supply? What tasks are required? What tasks are going to be accomplished by the effective response force?
8. **How much time will be required to complete this response?** The time allowed will be set by the agency, as will the measurement of the time, from what point to what point. The fractal level will be what percentage of the time this performance measure will be expected. It should never be 100 percent, because there are factors beyond the agency's control. The usual fractal level is 90 percent.

V. RISK ASSESSMENT



WHAT IS RISK?

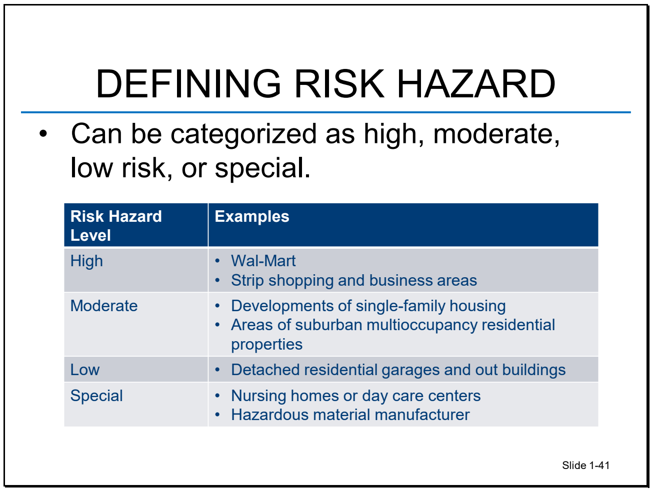
- **Risk** is the exposure to the chance of injury or loss; a hazard or dangerous chance.
- **Risk assessment** is an estimate of the likelihood of adverse effects that may result from exposure to certain hazards.

A. Introduction.

1. **Risk** is defined as the exposure to the chance of injury or loss; a hazard or dangerous chance.

2. **Risk assessment** is an estimate of the likelihood of adverse effects that may result from exposure to certain hazards.

   (ISO 31000 — Risk management: http://www.iso.org/iso/home/standard/iso31000.htm)



DEFINING RISK HAZARD

- Can be categorized as high, moderate, low risk, or special.

Risk Hazard Level	Examples
High	• Wal-Mart • Strip shopping and business areas
Moderate	• Developments of single-family housing • Areas of suburban multioccupancy residential properties
Low	• Detached residential garages and out buildings
Special	• Nursing homes or day care centers • Hazardous material manufacturer

Risk Hazard Level	Examples
High	• Wal-Mart • Strip shopping and business areas
Moderate	• Developments of single-family housing • Areas of suburban multioccupancy residential properties
Low	• Detached residential garages and out buildings
Special	• Nursing homes or day care centers • Hazardous material manufacturer

B. Defining risk hazard: Risk hazards can be categorized into four areas: high, moderate, low risk, or special risk.

1. **High risk** is defined as high concentration of property presenting a substantial risk of life loss, a severe financial impact on the community, or unusual potential damage to property. Examples include:

   a. Wal-Mart — one of the highest retail sales tax providers to a municipality. You lose this building and you lose approximately $1.2 million a year to the community.

   b. Strip shopping and business areas, consisting of either single or multistory properties, offering some degree of a major fire problem.

   c. Concentration of hospital and medical facilities.

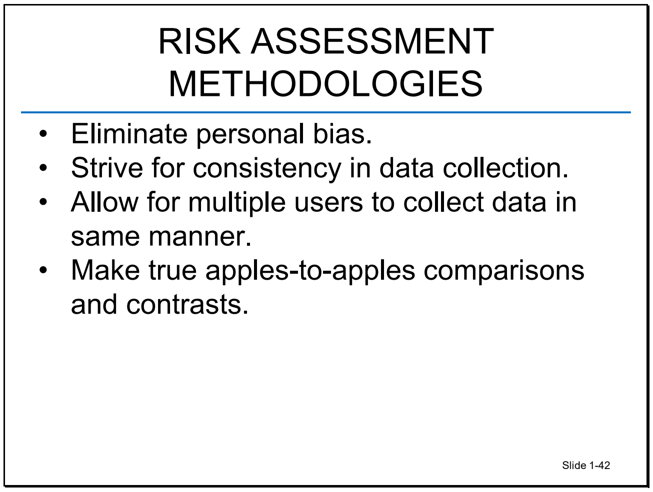
   d. Concentrations of older multistory property offering substantial amounts of exposure to life loss potential.

   e. Apartment buildings more than two stories in height, buildings with low occupant load but which store high fire load materials or high hazard materials.

   f. Infrastructure facilities, such as schools, city, state or federal facilities, and fire stations.

   g. Industrial buildings containing some high-risk occupancies.

2. **Moderate risk** is assigned when the situation contains built-up areas, where the risk of life loss or damage to property in the event of a fire in a single structure is usually limited to the occupants. Examples include:
    a. Developments of single-family housing, and small multistory dwellings.
    b. Areas of older, multifamily, two-story dwellings.
    c. Areas of suburban multioccupancy residential properties.
    d. Mixed low-risk industrial and residential areas.
    e. Industrial or commercial areas under 10,000 square feet without high-hazard or high fire load contents.
3. **Low risk** is assigned for small commercial structures that are remote from other buildings, such as detached residential garages and out buildings.
4. **Special risk** situations are those identified by the agency as being outside the above definitions. The idea is you send extra people and equipment to these incidents and consider them of concern to your agency. Examples include:
    a. Nursing homes or day care centers.
    b. Government or infrastructure risks.
    c. Hospitals.
    d. Hazardous materials manufacturers and all buildings where available water supply is less than projected fire flow.



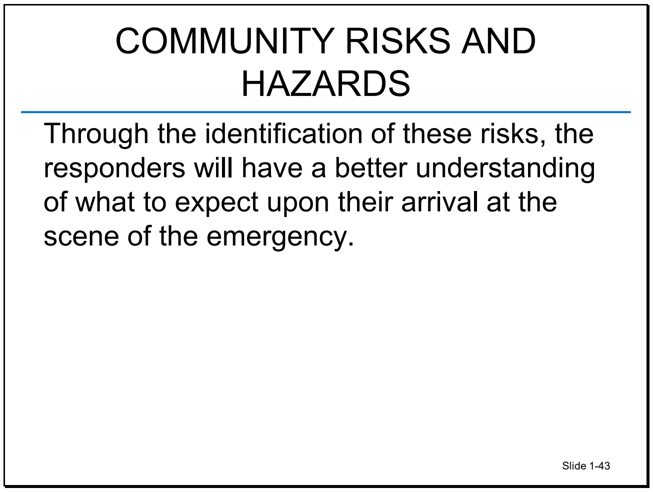
RISK ASSESSMENT METHODOLOGIES

- Eliminate personal bias.
- Strive for consistency in data collection.
- Allow for multiple users to collect data in same manner.
- Make true apples-to-apples comparisons and contrasts.

Slide 1-42

C. Risk assessment methodologies serve to:

1. Eliminate personal bias.
2. Strive for consistency in data collection.
3. Allow for multiple users to collect data in same manner.
4. Make true apples-to-apples comparisons and contrasts.



COMMUNITY RISKS AND HAZARDS

Through the identification of these risks, the responders will have a better understanding of what to expect upon their arrival at the scene of the emergency.

D. Community risks and hazards.

1. Through the identification of these risks the responders will have a better understanding of what to expect upon their arrival at the scene of the emergency.

2. Through a proper risk analysis the agency will better understand what types and amounts of equipment need to be sent to each type of emergency that is seen.

   If an agency can get the proper equipment (Hazmat Unit to Hazmat call) and the correct numbers (15 firefighters to a single family residence fire) then the agency will be closer to mitigating that incident.

3. The final focus needs to be that if the leaders of the agency, whether formal or informal leaders:

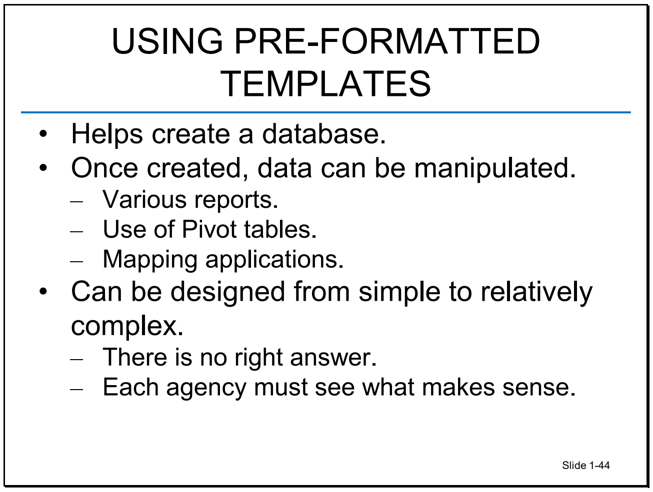
   a. Plan for the four main types of emergencies, (fire, EMS, hazmat, and rescue).

   b. Plan at each of the different types of risk (special, high, moderate, low).

   c. Set the response with what the critical tasking requires for that risk (number of personnel and number of each equipment type).

d. Prepare to mitigate any type of emergency that arises at these locations.

4. If an agency utilizes these guidelines and responds to the incident with what they have determined they will need at that location ahead of time, they have put their responders in a better, safer position to properly mitigate the situation.



USING PRE-FORMATTED TEMPLATES

- Helps create a database.
- Once created, data can be manipulated.
  - Various reports.
  - Use of Pivot tables.
  - Mapping applications.
- Can be designed from simple to relatively complex.
  - There is no right answer.
  - Each agency must see what makes sense.

Slide 1-44

E. Using pre-formatted templates.

1. Using pre-formatted templates reduces the chance of data being subjective instead of objective.

2. Objective data are completely unbiased, while subjective data cannot be verified using concrete facts and figures.

3. Pre-formatted templates help create a database; once the database is created, data can be manipulated for:

a. Various reports.

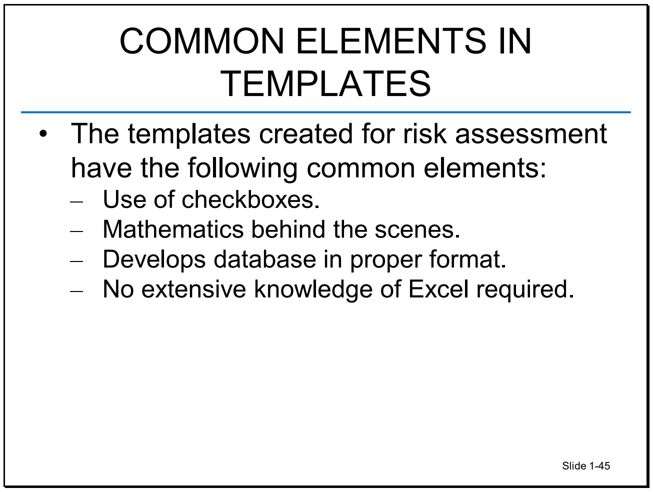
b. Use of Pivot tables.

c. Mapping applications.

4. These templates can also be designed to be anywhere from simple to relatively complex, keeping in mind that:

a. There is no right answer.

b. Each agency must see what makes sense for it.



COMMON ELEMENTS IN TEMPLATES

- The templates created for risk assessment have the following common elements:
  - Use of checkboxes.
  - Mathematics behind the scenes.
  - Develops database in proper format.
  - No extensive knowledge of Excel required.

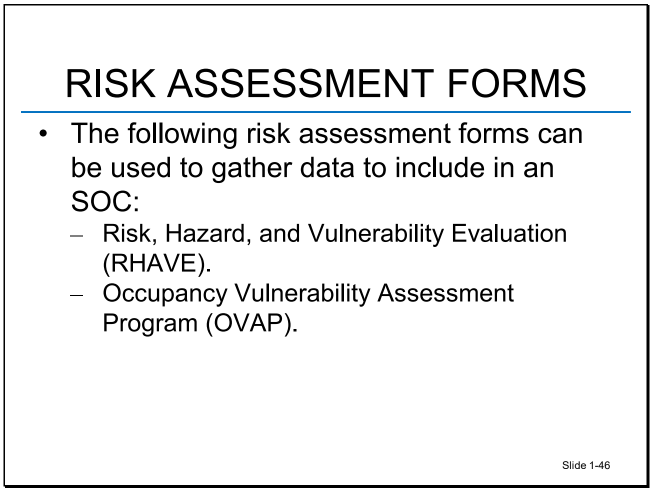
5. The templates created for risk assessment have the following common elements:

   a. Use of checkboxes.

   b. Mathematics behind the scenes.

   c. Develops database in proper format.

   d. No extensive knowledge of Excel required.



RISK ASSESSMENT FORMS

- The following risk assessment forms can be used to gather data to include in an SOC:
  - Risk, Hazard, and Vulnerability Evaluation (RHAVE).
  - Occupancy Vulnerability Assessment Program (OVAP).

F. Risk assessment forms.

1. Introduction.

   The following risk assessment forms can be used to gather data to include in an SOC:

a. Risk, Hazard, and Vulnerability Evaluation (RHAVE).

b. Occupancy Vulnerability Assessment Program (OVAP).

RISK ASSESSMENT FORMS (cont'd)

- Emergency Response Risk Assessment Form — Size, Height, Use, and Probability (SHUP).
- SPEED Vulnerability and Risk Rating Form (SPEED).
- Risk Assessment Form — Emergency Response (RAFER).

c. Emergency Response Risk Assessment Form — Size, Height, Use, and Probability (SHUP).

d. SPEED Vulnerability and Risk Rating Form (SPEED).

e. Risk Assessment Form — Emergency Response (RAFER).

RHAVE

2. **RHAVE** is a Home Security Assessment Profile. It evaluates risk, hazard, and vulnerability for:

a. Building.

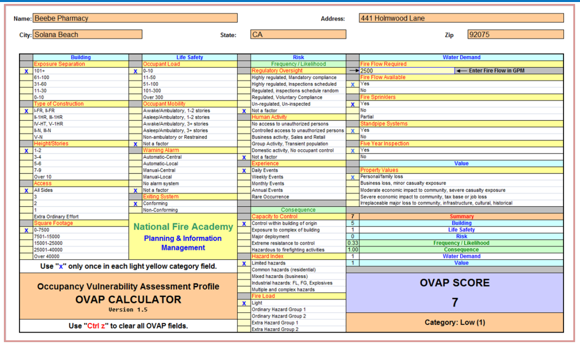
b. Life safety.

c. Risk.

d. Water demand.

e. Value.

OVAP



Name: Beebe Pharmacy
Address: 441 Holmwood Lane
City: Solana Beach
State: CA
Zip 92075

Building

Exposure Separation
☒ 101+
☐ 61-100
☐ 31-60
☐ 11-30
☐ 0-10

Type of Construction
☒ I-FR, II-FR
☐ II-1HR, III-1HR
☐ IV-HT, V-1HR
☐ II-N, III-N
☐ V-N

Height/Stories
☒ 1-2
☐ 3-4
☐ 5-6
☐ 7-9
☐ Over 10

Access
☒ All Sides
☐ 3
☐ 2
☐ 1
☐ Extra Ordinary Effort

Square Footage
☒ 0-7500
☐ 7501-15000
☐ 15001-25000
☐ 25001-40000
☐ Over 40000

Life Safety

Occupant Load
☒ 0-10
☐ 11-50
☐ 51-100
☐ 101-300
☐ Over 300

Occupant Mobility
☐ Awake/Ambulatory, 1-2 stories
☐ Asleep/Ambulatory, 1-2 stories
☐ Awake/Ambulatory, 3+ stories
☐ Asleep/Ambulatory, 3+ stories
☐ Non-ambulatory or Restrained
☒ Not a factor

Warning Alarm
☐ Automatic-Central
☐ Automatic-Local
☐ Manual-Central
☐ Manual-Local
☐ No alarm system
☒ Not a factor

Exiting System
☒ Conforming
☐ Non-Conforming

National Fire Academy
Planning & Information Management

Use "x" only once in each light yellow category field.

Occupancy Vulnerability Assessment Profile
OVAP CALCULATOR
Version 1.5

Use "Ctrl z" to clear all OVAP fields.

Risk

Frequency / Likelihood

Regulatory Oversight
☐ Highly regulated, Mandatory compliance
☐ Highly regulated, Inspections scheduled
☐ Regulated, inspections schedule random
☐ Regulated, Voluntary Compliance
☐ Un-regulated, Un-inspected
☒ Not a factor

Human Activity
☐ No access to unauthorized persons
☐ Controlled access to unauthorized persons
☐ Business activity, Sales and Retail
☐ Group Activity, Transient population
☐ Domestic activity, No occupant control
☒ Not a factor

Experience
☒ Daily Events
☐ Weekly Events
☐ Monthly Events
☐ Annual Events
☐ Rare Occurrence

Consequence

Capacity to Control
☒ Control within building of origin
☐ Exposure to complex of building
☐ Major deployment
☐ Extreme resistance to control
☐ Hazardous to firefighting activities

Hazard Index
☒ Limited hazards
☐ Common hazards (residential)
☐ Mixed hazards (business)
☐ Industrial hazards: FL, FG, Explosives
☐ Multiple and complex hazards

Fire Load
☒ Light
☐ Ordinary Hazard Group 1
☐ Ordinary Hazard Group 2
☐ Extra Hazard Group 1
☐ Extra Hazard Group 2

Water Demand

Fire Flow Required
→ 2500 ← Enter Fire Flow in GPM

Fire Flow Available
☒ Yes
☐ No

Fire Sprinklers
☒ Yes
☐ No
☐ Partial

Standpipe Systems
☒ Yes
☐ No

Five Year Inspection
☒ Yes
☐ No

Value

Property Values
☒ Personal/family loss
☐ Business loss, minor casualty exposure
☐ Moderate economic impact to community, severe casualty exposure
☐ Severe economic impact to community, tax base or job loss
☐ Irreplaceable major loss to community, infrastructure, cultural, historical

7	Summary
5	Building
1	Life Safety
0	Risk
0.33	Frequency / Likelihood
1.00	Consequence
1	Water Demand
1	Value

OVAP SCORE
7

Category: Low (1)

3. **OVAP** is a Homeland Security component. It calculates risk for:

a. Building.

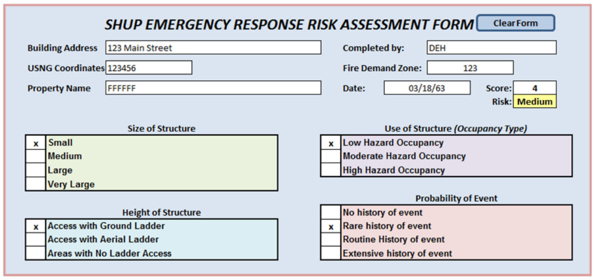
b. Life safety.

c. Risk.

d. Water demand.

e. Value.

SHUP



SHUP EMERGENCY RESPONSE RISK ASSESSMENT FORM

Building Address: 123 Main Street
Completed by: DEH
USNG Coordinates: 123456
Fire Demand Zone: 123
Property Name: FFFFFF
Date: 03/18/63
Score: 4
Risk: Medium

Size of Structure
☒ Small
☐ Medium
☐ Large
☐ Very Large

Use of Structure *(Occupancy Type)*
☒ Low Hazard Occupancy
☐ Moderate Hazard Occupancy
☐ High Hazard Occupancy

Height of Structure
☒ Access with Ground Ladder
☐ Access with Aerial Ladder
☐ Areas with No Ladder Access

Probability of Event
☐ No history of event
☒ Rare history of event
☐ Routine History of event
☐ Extensive history of event

4. **SHUP** is derived from the SOC, fifth edition document. It allows you to identify structure data including:

   a. Address, including map coordinates.

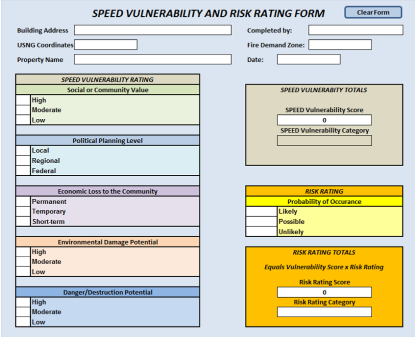
   b. Demand zone.

   c. Size and height.

   d. Occupancy type.

   e. Probability of event.

SPEED



SPEED VULNERABILITY AND RISK RATING FORM

Clear Form

Building Address ☐
USNG Coordinates ☐
Property Name ☐
Completed by: ☐
Fire Demand Zone: ☐
Date: ☐

SPEED VULNERABILITY RATING

Social or Community Value
☐ High
☐ Moderate
☐ Low

Political Planning Level
☐ Local
☐ Regional
☐ Federal

Economic Loss to the Community
☐ Permanent
☐ Temporary
☐ Short-term

Environmental Damage Potential
☐ High
☐ Moderate
☐ Low

Danger/Destruction Potential
☐ High
☐ Moderate
☐ Low

SPEED VULNERABITY TOTALS

SPEED Vulnerability Score
0
SPEED Vulnerability Category

RISK RATING

Probability of Occurance
☐ Likely
☐ Possible
☐ Unlikely

RISK RATING TOTALS

Equals Vulnerability Score x Risk Rating

Risk Rating Score
0
Risk Rating Category

5. **SPEED** takes coordinates from the United States National Grid (USNG) and allows you to determine the risk rating. It collects data on the following and gives total and risk rating as well as probability:

   a. Social or community value.

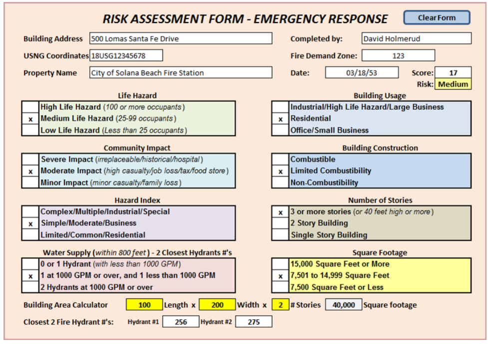
   b. Political planning level.

   c. Economic loss to the community.

   d. Environmental damage.

   e. Potential danger/damage potential.

RAFER



RISK ASSESSMENT FORM - EMERGENCY RESPONSE Clear Form

Building Address: 500 Lomas Santa Fe Drive
USNG Coordinates: 18USG12345678
Property Name: City of Solana Beach Fire Station
Completed by: David Holmerud
Fire Demand Zone: 123
Date: 03/18/53
Score: 17
Risk: Medium

Life Hazard

- ☐ High Life Hazard (*100 or more occupants*)
- ☒ Medium Life Hazard (*25-99 occupants*)
- ☐ Low Life Hazard (*Less than 25 occupants*)

Community Impact

- ☐ Severe Impact (*irreplaceable/historical/hospital*)
- ☒ Moderate Impact (*high casualty/job loss/tax/food store*)
- ☐ Minor Impact (*minor casualty/family loss*)

Hazard Index

- ☐ Complex/Multiple/Industrial/Special
- ☒ Simple/Moderate/Business
- ☐ Limited/Common/Residential

Water Supply (*within 800 feet*) - 2 Closest Hydrants #'s

- ☐ 0 or 1 Hydrant (*with less than 1000 GPM*)
- ☒ 1 at 1000 GPM or over, and 1 less than 1000 GPM
- ☐ 2 Hydrants at 1000 GPM or over

Building Usage

- ☐ Industrial/High Life Hazard/Large Business
- ☒ Residential
- ☐ Office/Small Business

Building Construction

- ☐ Combustible
- ☒ Limited Combustibility
- ☐ Non-Combustibility

Number of Stories

- ☒ 3 or more stories (*or 40 feet high or more*)
- ☐ 2 Story Building
- ☐ Single Story Building

Square Footage

- ☐ 15,000 Square Feet or More
- ☒ 7,501 to 14,999 Square Feet
- ☐ 7,500 Square Feet or Less

Building Area Calculator: 100 Length x 200 Width x 2 # Stories 40,000 Square footage
Closest 2 Fire Hydrant #'s: Hydrant #1 256 Hydrant #2 275

6. **RAFER** collects data and then calculates the risk level. It collects data on the following:

   a. Life hazard.

   b. Community impact.

   c. Hazard index.

   d. Water supply.

   e. Building usage.

   f. Building construction.

   g. Number of stories.

   h. Square footage.

RISK LEVELS

- Probability and Consequence (P&C).
- Response requirements.
- Calculators.
  - P&C.
  - Low risk.
  - Medium risk.
  - High risk.

7. **Risk levels** worksheets help determine the Probability and Consequence (P&C) of situations based on data entered and help calculate response requirements. Risk calculators include:

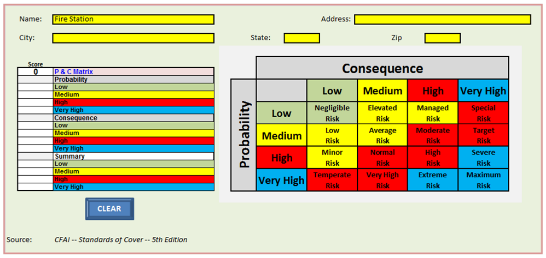
   a. P&C.

   b. Low risk.

   c. Medium risk.

   d. High risk.

PROBABILITY AND CONSEQUENCE



Name: Fire Station

Address:

City:

State:

Zip

Score	
0	P & C Matrix
	Probability
	Low
	Medium
	High
	Very High
	Consequence
	Low
	Medium
	High
	Very High
	Summary
	Low
	Medium
	High
	Very High

		Consequence			
		Low	Medium	High	Very High
Probability	Low	Negligible Risk	Elevated Risk	Managed Risk	Special Risk
	Medium	Low Risk	Average Risk	Moderate Risk	Target Risk
	High	Minor Risk	Normal Risk	High Risk	Severe Risk
	Very High	Temperate Risk	Very High Risk	Extreme Risk	Maximum Risk

Source: *CFAI -- Standards of Cover -- 5th Edition*

NOAA

- National Oceanic and Atmospheric Administration assessment matrices include:
  - Vulnerability Assessment Matrix.
  - Opportunity Assessment Matrix.

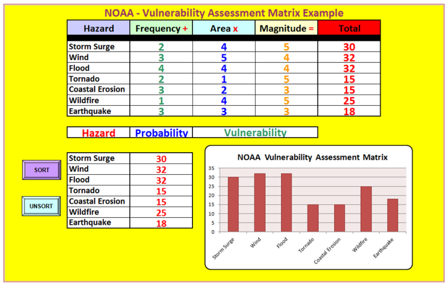
8. **The National Oceanic and Atmospheric Administration** (NOAA) matrices help determine risk levels based on natural disasters. Note that these matrices have little to do with development of an SOC. The assessment matrices include:

   a. Vulnerability Assessment Matrix — Natural disasters.

b. Opportunity Assessment Matrix — Risk Program.

9. As you define purchases and programs in a department, these matrices may be able to help you get the most for your money (for planning for natural disasters).

NOAA — VULNERABILITY ASSESSMENT MATRIX



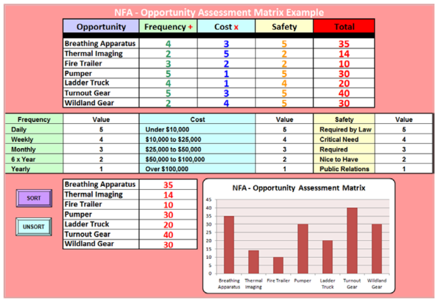
NOAA - Vulnerability Assessment Matrix Example

Hazard	Frequency +	Area x	Magnitude =	Total
Storm Surge	2	4	5	30
Wind	3	5	4	32
Flood	4	4	4	32
Tornado	2	1	5	15
Coastal Erosion	3	2	3	15
Wildfire	1	4	5	25
Earthquake	3	3	3	18

Hazard	Probability	Vulnerability

Storm Surge	30
Wind	32
Flood	32
Tornado	15
Coastal Erosion	15
Wildfire	25
Earthquake	18

NOAA — OPPORTUNITY ASSESSMENT MATRIX

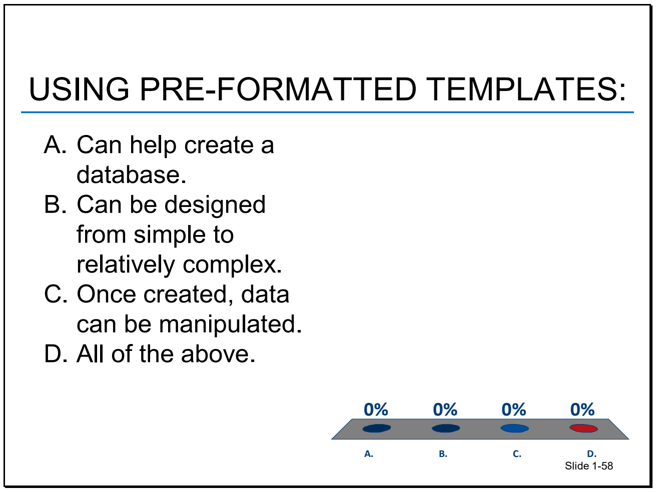


NFA - Opportunity Assessment Matrix Example

Opportunity	Frequency +	Cost x	Safety	Total
Breathing Apparatus	4	3	5	35
Thermal Imaging	2	5	2	14
Fire Trailer	3	2	2	10
Pumper	5	1	5	30
Ladder Truck	4	1	4	20
Turnout Gear	5	3	5	40
Wildland Gear	2	4	5	30

Frequency	Value	Cost	Value	Safety	Value
Daily	5	Under $10,000	5	Required by Law	5
Weekly	4	$10,000 to $25,000	4	Critical Need	4
Monthly	3	$25,000 to $50,000	3	Required	3
6 x Year	2	$50,000 to $100,000	2	Nice to Have	2
Yearly	1	Over $100,000	1	Public Relations	1

Breathing Apparatus	35
Thermal Imaging	14
Fire Trailer	10
Pumper	30
Ladder Truck	20
Turnout Gear	40
Wildland Gear	30



USING PRE-FORMATTED TEMPLATES:

A. Can help create a database.
B. Can be designed from simple to relatively complex.
C. Once created, data can be manipulated.
D. All of the above.

Slide 1-58

ACTIVITY 1.3

Risk Assessment

Purpose

To identify how risk assessment tools can aid in planning.

Directions

1. You will work in groups; elect a group representative for debriefing.
2. Spend 10 to 15 minutes discussing the following amongst the group:
   a. Identifying tools used in planning.
   b. Identifying sources of information (examples, inspection reports, fire data, or site inspections).
   c. Explaining how data can impact your planning.
3. You will use the worksheet to record items.
4. Your group's representative will share your group's findings with the class.

This page intentionally left blank.

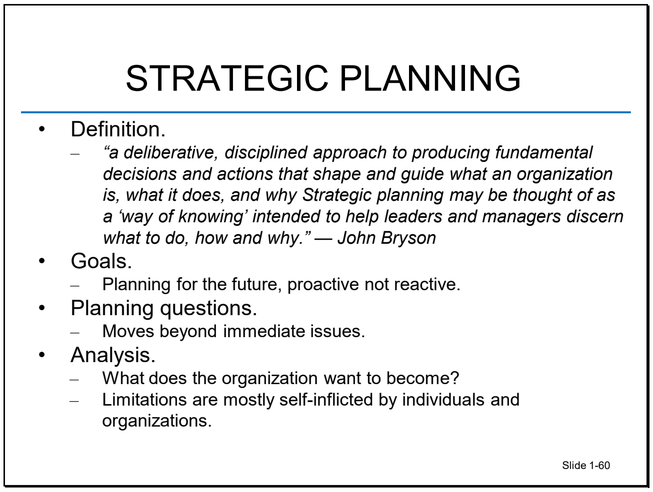
ACTIVITY 1.3 (cont'd)

Risk Assessment

Identify and list which tools you will use (from the list of tools just discussed or existing tools you have in-house) to perform a risk assessment.
Identify and list the sources of information you will use in conjunction with these tools.
Armed with this information, does this change your planning? Explain how.

This page intentionally left blank.

VI. STRATEGIC PLANNING



STRATEGIC PLANNING

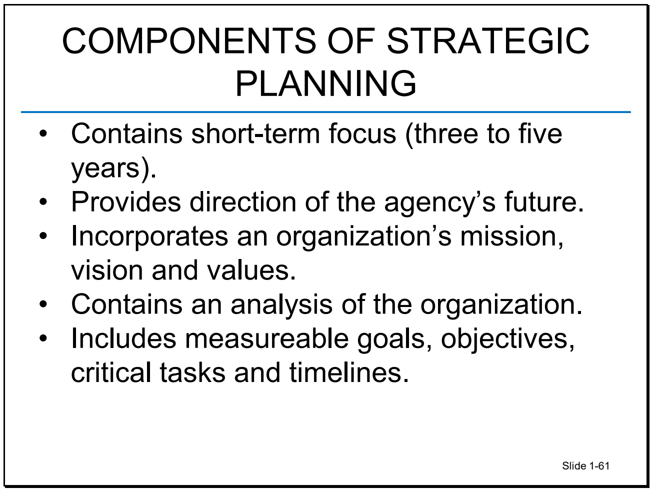
- Definition.
  - *"a deliberative, disciplined approach to producing fundamental decisions and actions that shape and guide what an organization is, what it does, and why Strategic planning may be thought of as a 'way of knowing' intended to help leaders and managers discern what to do, how and why." — John Bryson*
- Goals.
  - Planning for the future, proactive not reactive.
- Planning questions.
  - Moves beyond immediate issues.
- Analysis.
  - What does the organization want to become?
  - Limitations are mostly self-inflicted by individuals and organizations.

A. Introduction.

1. Strategic planning is "a deliberative, disciplined approach to producing fundamental decisions and actions that shape and guide what an organization (or other entity) is, what it does, and why Strategic planning may be thought of as a 'way of knowing' intended to help leaders and mangers discern what to do, how, and why." (Source: M. Bryson, *Strategic Planning for Public and Non-Profit Organizations*, 2011.)

2. Strategic planning allows organizations to be proactive rather than reactive. If an organization is always working to handle immediate issues, it may not be prepared for larger issues coming its way.

3. By thinking long term (or strategically), organizations can plan for the future. This is exactly what preparing an SOC works toward.



COMPONENTS OF STRATEGIC PLANNING

- Contains short-term focus (three to five years).
- Provides direction of the agency's future.
- Incorporates an organization's mission, vision and values.
- Contains an analysis of the organization.
- Includes measureable goals, objectives, critical tasks and timelines.

B. Components of strategic planning.

1. A strategic plan does the following:

a. Contains a short-term focus of three to five years.

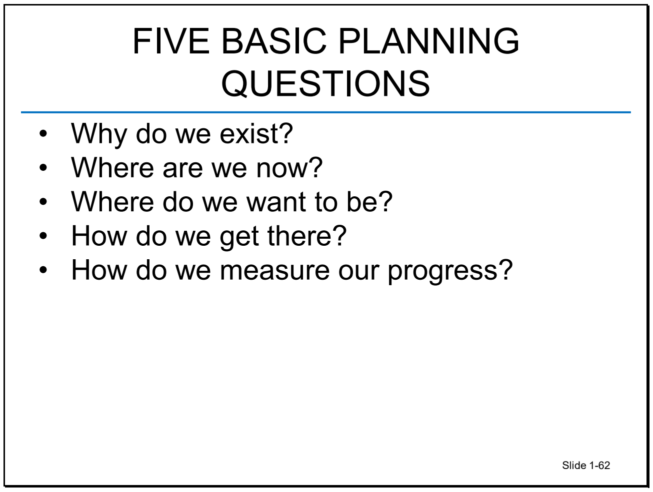
b. Provides direction of the agency's future to all stakeholders.

c. Incorporates an organization's mission, vision and values.

d. Contains an analysis of the organization's strengths, weaknesses, threats and opportunities.

e. Involves input from external and internal stakeholders.

f. Includes measurable goals, objectives, critical tasks and timelines.



FIVE BASIC PLANNING QUESTIONS

- Why do we exist?
- Where are we now?
- Where do we want to be?
- How do we get there?
- How do we measure our progress?

Slide 1-62

2. There are five basic planning questions that should be addressed when developing a strategic plan. They are the following:

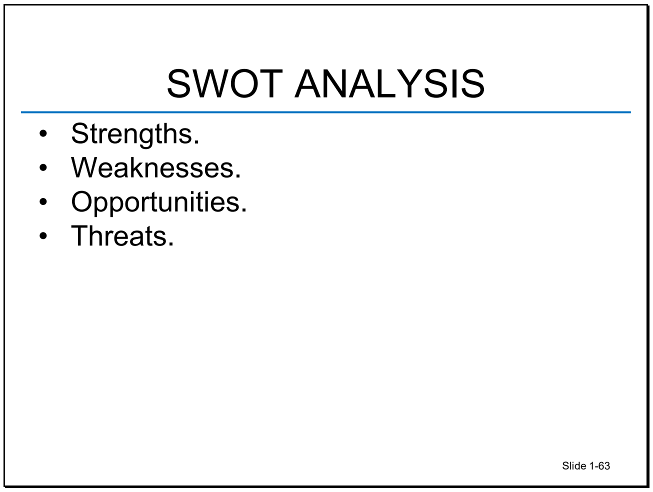
a. Why do we exist? Outlined by our mission and values.

b. Where are we now? Outlined by services and programs.

c. Where do we want to be? Outlined by vision.

d. How do we get there? Outlined by goals, objectives, tasks and timelines.

e. How do we measure our progress? Outlined by performance measures.



SWOT ANALYSIS

- Strengths.
- Weaknesses.
- Opportunities.
- Threats.

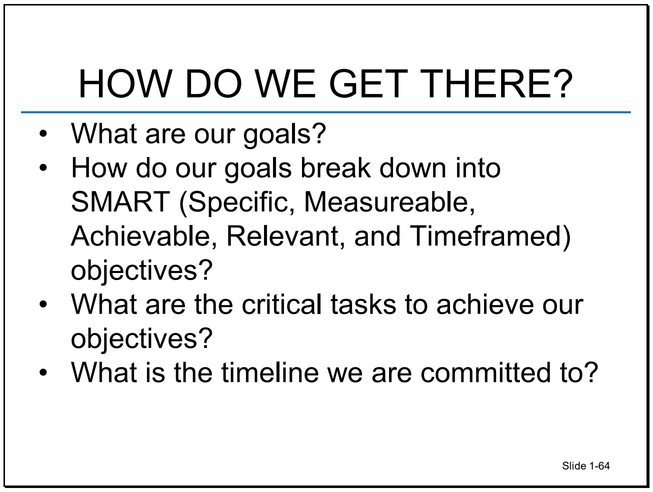
3. It is recommended that your organization perform a Strengths, Weaknesses, Opportunities, and Threats (SWOT) analysis when developing a strategic plan. The SWOT analysis can help identify strengths as well as areas of opportunity.

   a. **Strengths and weaknesses** — looking internal to your organization and includes:

      - Policies.
      - Facilities.
      - Training.
      - Culture.
      - Organizational structure.

   b. **Opportunities and threats** — looking external to your organization and includes:

      - Politics.
      - Legal issues.
      - Demographic changes.
      - Technological developments.
      - Economic variables.



HOW DO WE GET THERE?

- What are our goals?
- How do our goals break down into SMART (Specific, Measureable, Achievable, Relevant, and Timeframed) objectives?
- What are the critical tasks to achieve our objectives?
- What is the timeline we are committed to?

4. How do we get there? The following questions should be addressed in order to develop a strong strategic plan for your organization:

   a. What are our goals?

   b. How do our goals break down into SMART (Specific, Measurable, Achievable, Relevant, and Timeframed) objectives?

   c. What are the critical tasks to achieve our objectives?

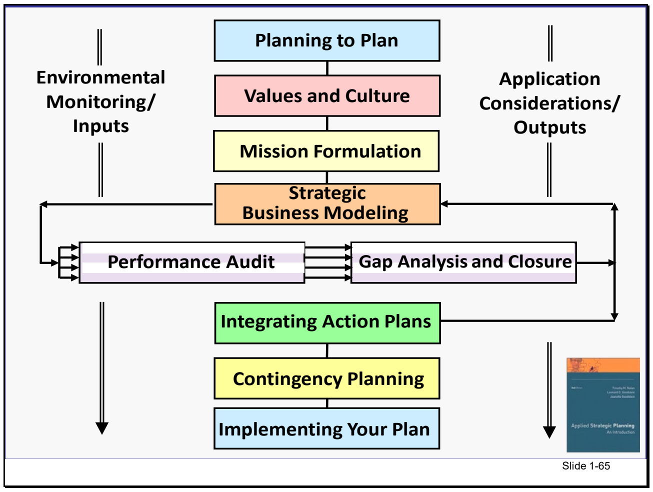
   d. What is the timeline we are committed to?

5. To apply this risk analysis and emergency response, first you must have an understanding of the community in order to properly prepare the risk analysis.

   a. The plan should cover the goal of analyzing the community.

   b. The needs that are identified in the risk analysis will need to be addressed in the plan.

   c. The plan is your “roadmap” of where you are headed as an agency.



Environmental Monitoring/ Inputs

Planning to Plan

Values and Culture

Mission Formulation

Strategic Business Modeling

Performance Audit

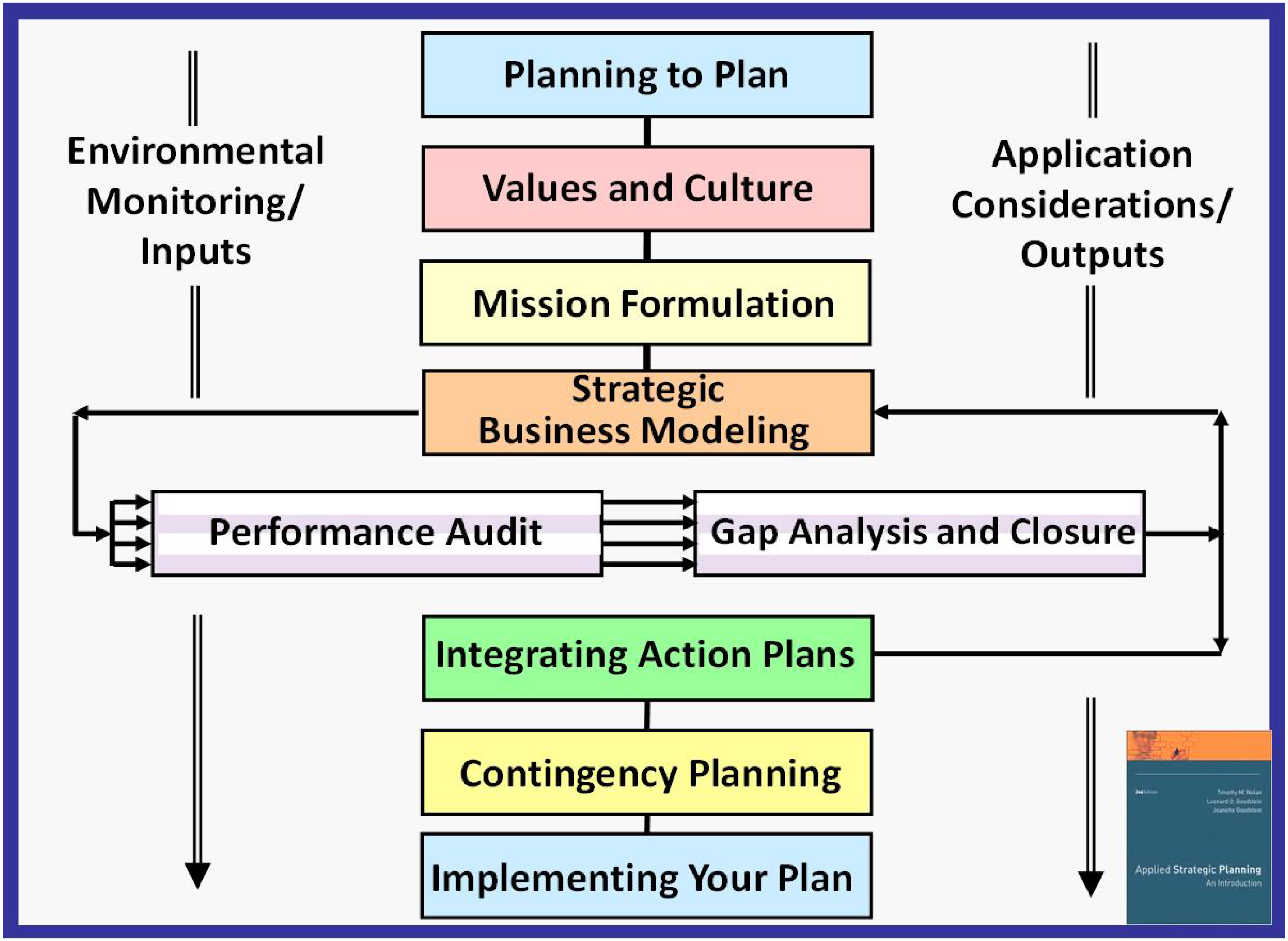
Gap Analysis and Closure

Integrating Action Plans

Contingency Planning

Implementing Your Plan

Application Considerations/ Outputs



Environmental Monitoring/ Inputs

Planning to Plan

Values and Culture

Mission Formulation

Strategic Business Modeling

Performance Audit

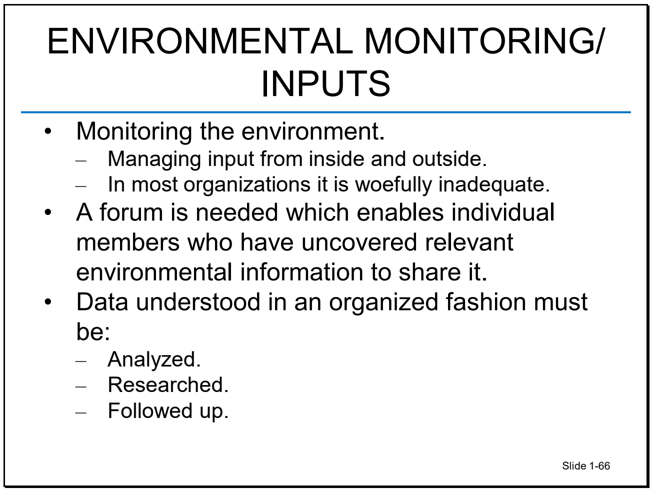
Gap Analysis and Closure

Integrating Action Plans

Contingency Planning

Implementing Your Plan

Application Considerations/ Outputs



ENVIRONMENTAL MONITORING/ INPUTS

- Monitoring the environment.
  - Managing input from inside and outside.
  - In most organizations it is woefully inadequate.
- A forum is needed which enables individual members who have uncovered relevant environmental information to share it.
- Data understood in an organized fashion must be:
  - Analyzed.
  - Researched.
  - Followed up.

C. Environmental monitoring/inputs.

1. Monitoring the environment is managing input from inside and outside the organization. When conducted by most organizations, it is woefully inadequate.

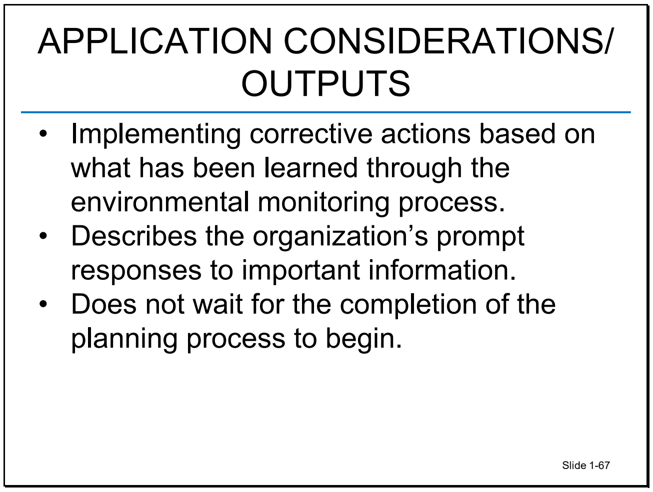
2. A forum, which enables individual members who have uncovered relevant environmental information to share, needs to be established.

3. Data understood in an organized fashion must be:

   a. Analyzed.

   b. Researched.

   c. Followed up.



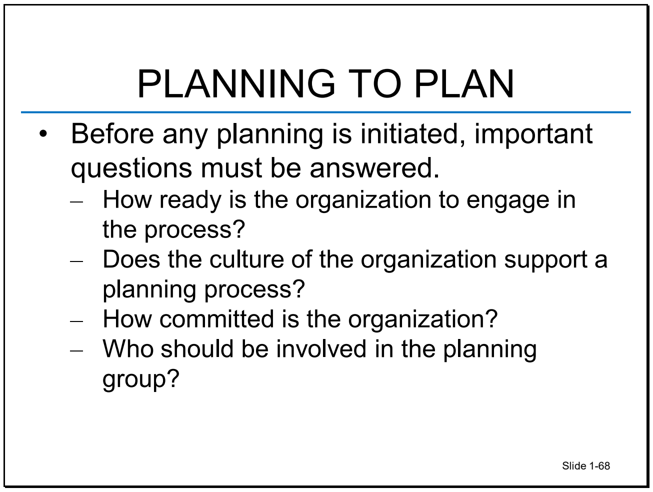
APPLICATION CONSIDERATIONS/ OUTPUTS

- Implementing corrective actions based on what has been learned through the environmental monitoring process.
- Describes the organization’s prompt responses to important information.
- Does not wait for the completion of the planning process to begin.

D. Application considerations/outputs.

1. Implementing corrective actions based on what you learn through the environmental monitoring process is an important piece of the Applied Strategic Planning model. Instead of waiting to complete the planning process, organizations should respond promptly to important information.

2. This is called application considerations and is the second process in the model.



PLANNING TO PLAN

- Before any planning is initiated, important questions must be answered.
  - How ready is the organization to engage in the process?
  - Does the culture of the organization support a planning process?
  - How committed is the organization?
  - Who should be involved in the planning group?

E. Planning to plan.

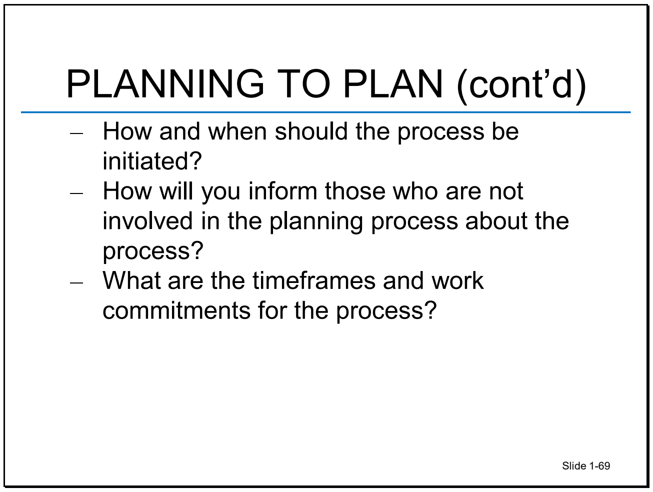
1. You must consider several important questions before beginning any planning.

   a. How ready is the organization to engage in the process?

   b. Does the culture of the organization support a planning process?

   c. How much time and energy does the organization have to commit to the process?

   d. Who should be involved in the planning group?



PLANNING TO PLAN (cont'd)

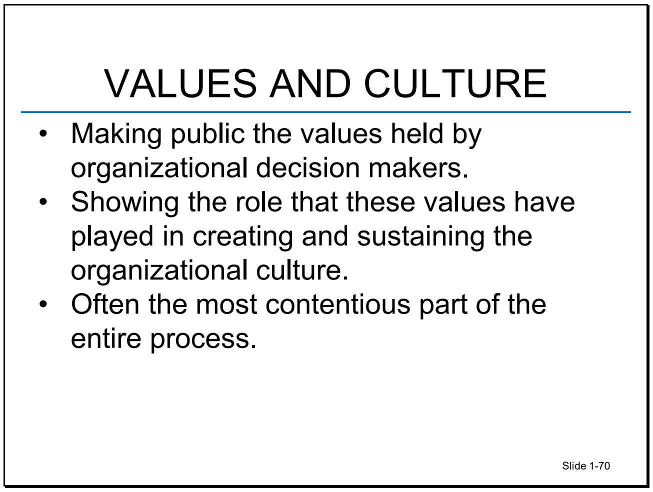
- How and when should the process be initiated?
- How will you inform those who are not involved in the planning process about the process?
- What are the timeframes and work commitments for the process?

e. How and when should the process be initiated?

f. How will you inform those who are not involved in the planning process about the process?

g. What are the timeframes for the process?

2. The planning group is one of the most important considerations. This group should consist of seven to 11 people who broadly represent the management level of the organization. This group must be prepared to commit time and energy to the process while still accomplishing their daily tasks.

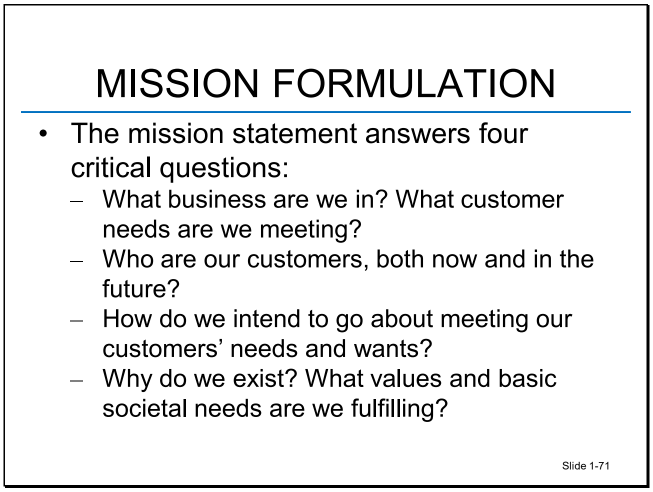


VALUES AND CULTURE

- Making public the values held by organizational decision makers.
- Showing the role that these values have played in creating and sustaining the organizational culture.
- Often the most contentious part of the entire process.

F. Values and culture.

Establishing organizational values to guide individual and organizational decision-making is an integral part of the Applied Strategic Planning process. Because establishing and clarifying these values can be very contentious, a highly skilled facilitator is often required to mediate the process.



MISSION FORMULATION

- The mission statement answers four critical questions:
  - What business are we in? What customer needs are we meeting?
  - Who are our customers, both now and in the future?
  - How do we intend to go about meeting our customers' needs and wants?
  - Why do we exist? What values and basic societal needs are we fulfilling?

G. Mission formulation/clarification.

1. Upon establishing organizational goals, it is time to formulate a future vision for the organization. This vision guides the drafting of a mission statement.

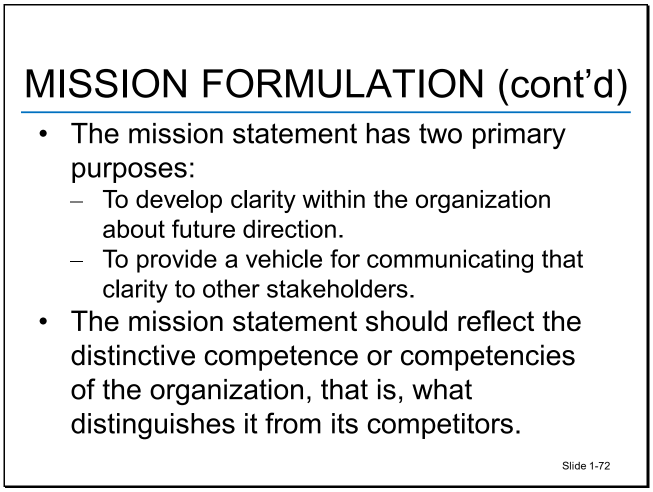
2. The mission statement answers four questions:

   a. What business are we in? What customer needs are we meeting?

   b. Who are our customers, both now and in the future?

   c. How do we intend to go about meeting our customers' needs and wants?

   d. Why do we exist? What values and basic societal needs are we fulfilling?

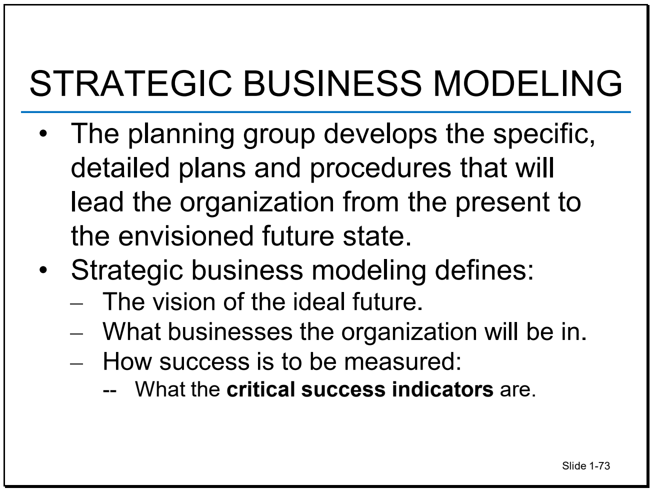


MISSION FORMULATION (cont'd)

- The mission statement has two primary purposes:
  - To develop clarity within the organization about future direction.
  - To provide a vehicle for communicating that clarity to other stakeholders.
- The mission statement should reflect the distinctive competence or competencies of the organization, that is, what distinguishes it from its competitors.

3. The mission statement provides clarity within the organization about future direction and provides a method of communicating that direction to other stakeholders. The mission statement should distinguish the organization from its competitors.

4. Often, organizations already have mission statements. Since that is the case, the task may shift to clarifying the existing mission statement.



STRATEGIC BUSINESS MODELING

- The planning group develops the specific, detailed plans and procedures that will lead the organization from the present to the envisioned future state.
- Strategic business modeling defines:
  - The vision of the ideal future.
  - What businesses the organization will be in.
  - How success is to be measured:
    - -- What the **critical success indicators** are.

H. Strategic business modeling.

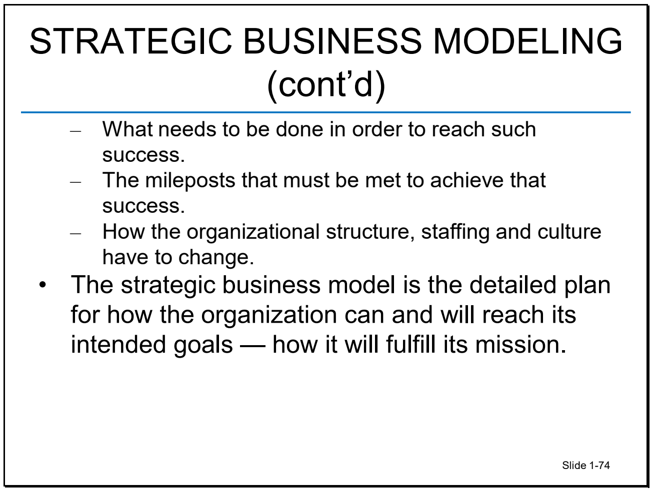
1. Strategic business modeling is the next step in the planning process. In this step, the planning group develops plans and procedures that will guide the organization to accomplish the mission statement. Creating or clarifying the mission statement is the first step in this process.

2. Strategic business modeling defines:

   a. The vision of the ideal future.

b. What businesses the organization will be in.

c. Critical success indicators.



STRATEGIC BUSINESS MODELING (cont'd)

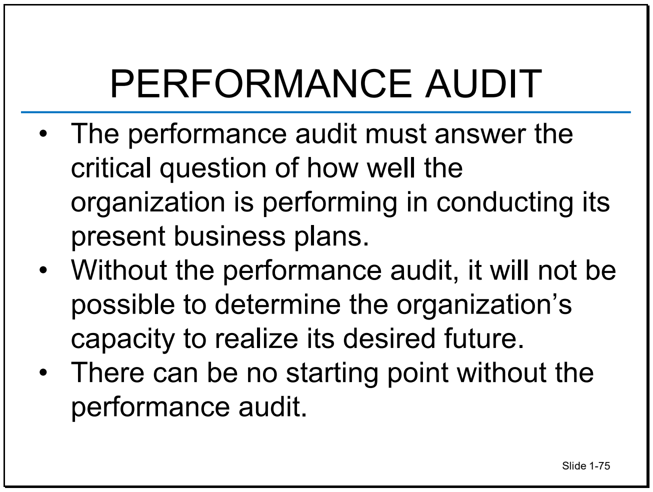
  - What needs to be done in order to reach such success.
  - The mileposts that must be met to achieve that success.
  - How the organizational structure, staffing and culture have to change.
- The strategic business model is the detailed plan for how the organization can and will reach its intended goals — how it will fulfill its mission.

d. Steps to reach success.

e. Mileposts to achieve success.

f. Organizational structure, staffing and culture changes that need to happen.

3. The strategic business model is a plan for how to reach its intended goals or mission. This process requires a creative approach to establishing goals and reaching them.

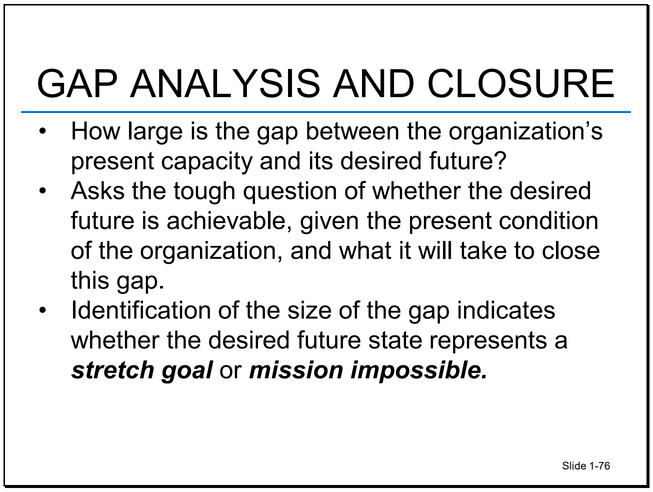


PERFORMANCE AUDIT

- The performance audit must answer the critical question of how well the organization is performing in conducting its present business plans.
- Without the performance audit, it will not be possible to determine the organization's capacity to realize its desired future.
- There can be no starting point without the performance audit.

I. Performance audit.

After articulating a mission statement and plan, the planning group must conduct an objective, unbiased performance audit. The performance audit will explore how well the organization is performing in conducting its present business plans. This is necessary to provide a baseline to begin achieving the mission established earlier.



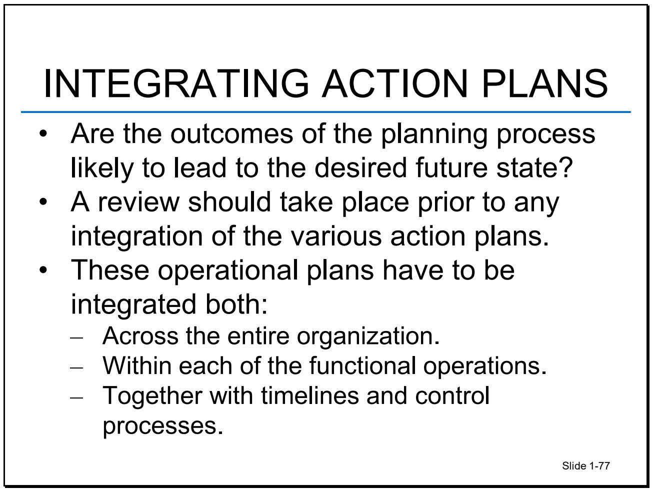
GAP ANALYSIS AND CLOSURE

- How large is the gap between the organization's present capacity and its desired future?
- Asks the tough question of whether the desired future is achievable, given the present condition of the organization, and what it will take to close this gap.
- Identification of the size of the gap indicates whether the desired future state represents a ***stretch goal*** or ***mission impossible.***

Slide 1-76

J. Gap analysis and closure.

After establishing a desired future state and a clear picture of the current state, the planning group must define the gap between the organization's present capacity and its desired future. Gap analysis and closure asks whether the desired future is achievable given the present condition of the organization, and what it will take to close this gap. The size of the gap indicates whether the future state represents a **stretch goal** or **mission impossible.**



INTEGRATING ACTION PLANS

- Are the outcomes of the planning process likely to lead to the desired future state?
- A review should take place prior to any integration of the various action plans.
- These operational plans have to be integrated both:
  - Across the entire organization.
  - Within each of the functional operations.
  - Together with timelines and control processes.

Slide 1-77

K. Finalizing strategic direction/Integrating action plans.

1. Prior to integrating any action plans, review your desired outcomes and consider whether the action plans will lead to the desired future state. Then, finalize the strategic direction or return to some earlier steps in the planning process if you decide the action plans will not yield the desired results.

2. In order to implement a strategic plan, you must create and integrate several operational plans. You must include:

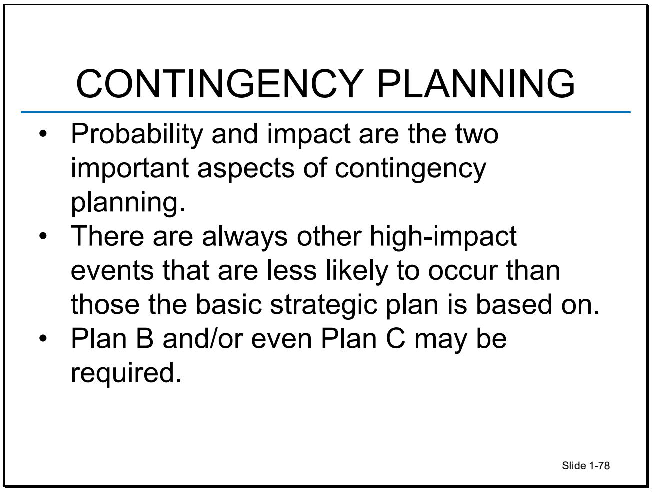
   a. One plan for each line of business.

   b. Plans for:

      - Product or services development.
      - Human resources.
      - Marketing.
      - Capital acquisition and budgeting.
      - Acquisition of technical resources.

   c. A plan for creating an adequate communications network.

3. These operational plans must be integrated across the organization and within each of the functional operations, and must include timelines and control processes.



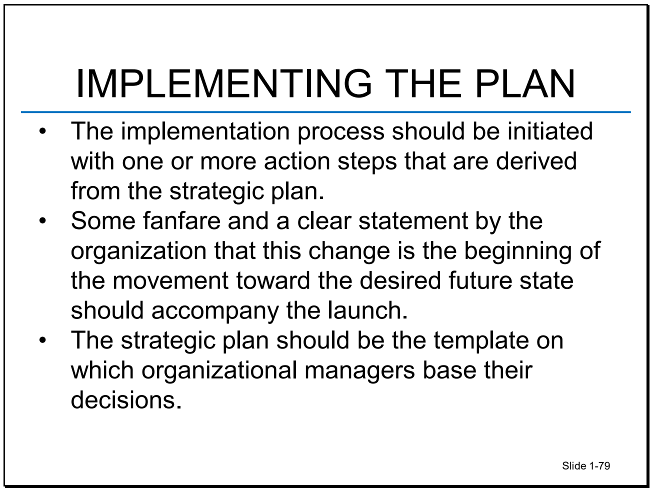
CONTINGENCY PLANNING

- Probability and impact are the two important aspects of contingency planning.
- There are always other high-impact events that are less likely to occur than those the basic strategic plan is based on.
- Plan B and/or even Plan C may be required.

Slide 1-78

L. Contingency planning.

Probability and impact are the two important aspects of contingency planning. The basic strategic plan depends upon the scenario with the highest probability of successful implementation. However, there are other high-impact events that are less likely to occur but must be considered. Develop contingency plans for these events. If this cannot be accomplished, the strategic plan should include methods for tracking these alternative events so that the organization's strategic plan can be re-examined and changed when necessary.



IMPLEMENTING THE PLAN

- The implementation process should be initiated with one or more action steps that are derived from the strategic plan.
- Some fanfare and a clear statement by the organization that this change is the beginning of the movement toward the desired future state should accompany the launch.
- The strategic plan should be the template on which organizational managers base their decisions.

Slide 1-79

M. Implementing the plan.

1. Strategic planning must be followed by implementation. The strategic plan should guide organizational managers in their decision-making processes. The implementation process should be initiated with one or more action steps derived from the strategic plan. This could mean:

    a. Restructuring the organization.

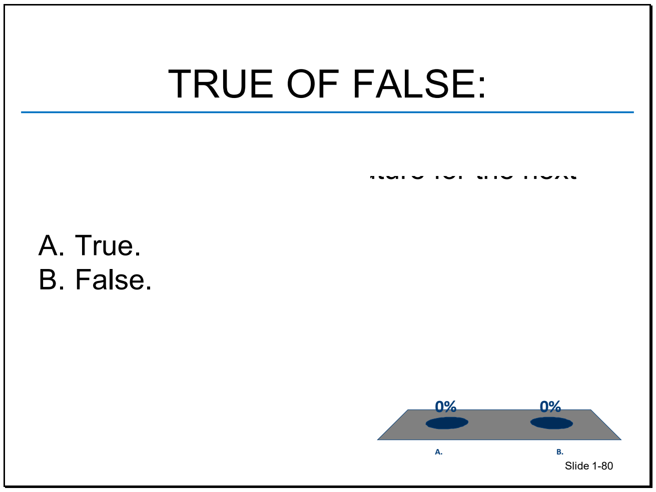
    b. Launching a new product or service.

    c. Changing some senior management positions.

    d. Developing a new training program.

2. A good communications or marketing campaign when the plan is launched will contribute to the success of the plan.

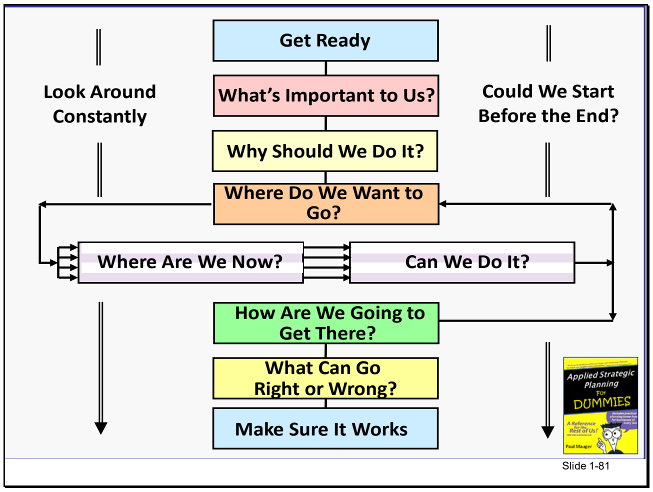
3. With this brief overview of the model in mind, we turn our attention to the role of the consultant in the planning process.



TRUE OF FALSE:

ture for the next

A. True.
B. False.



Look Around Constantly

Get Ready

What's Important to Us?

Why Should We Do It?

Where Do We Want to Go?

Where Are We Now?

Can We Do It?

How Are We Going to Get There?

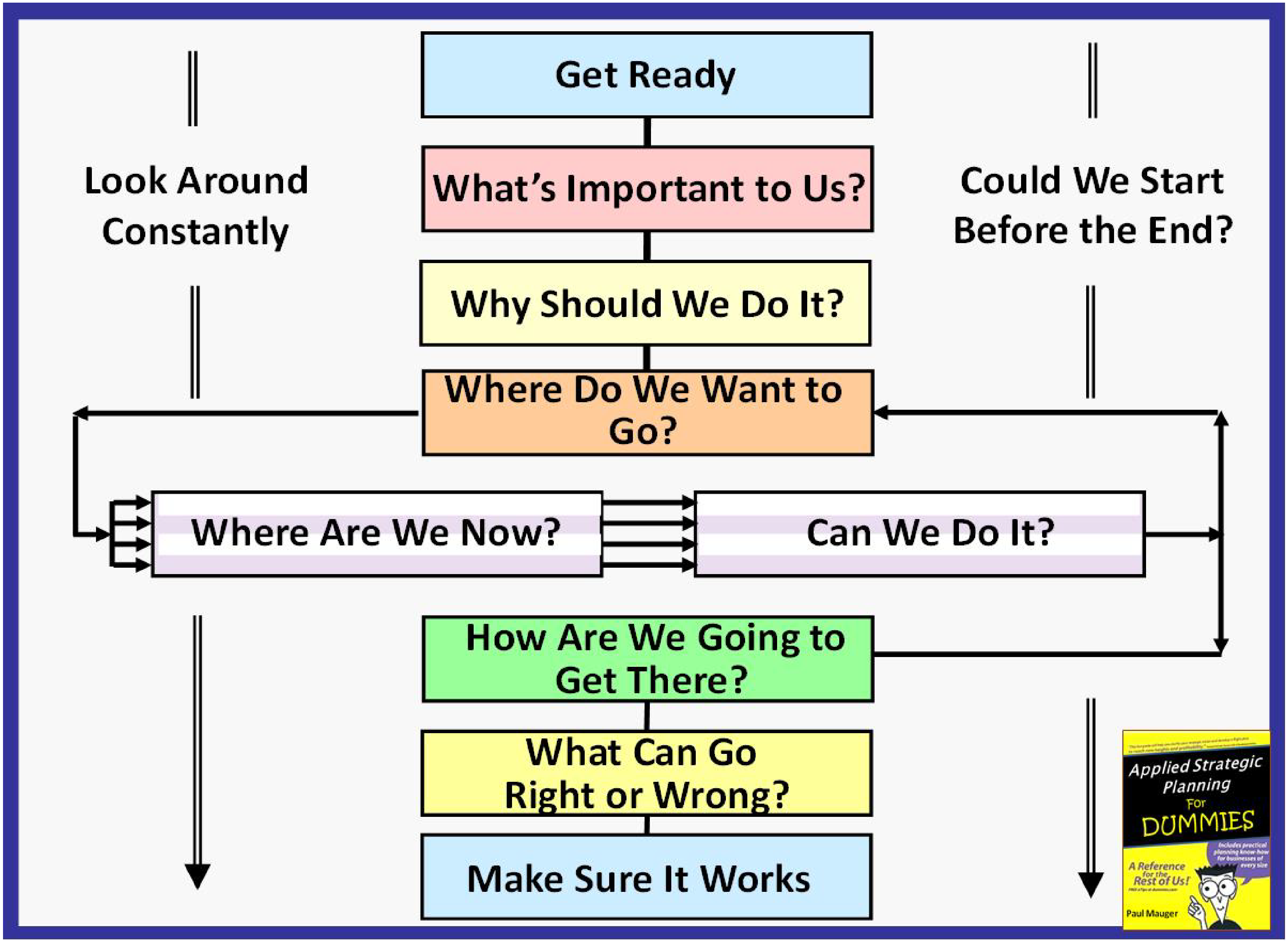
What Can Go Right or Wrong?

Make Sure It Works

Could We Start Before the End?

Transformational Model

Transformational Model



Get Ready

What's Important to Us?

Why Should We Do It?

Where Do We Want to Go?

Where Are We Now?

Can We Do It?

How Are We Going to Get There?

What Can Go Right or Wrong?

Make Sure It Works

Look Around Constantly

Could We Start Before the End?

Transformational Model

Resources

Fire

EMS

Services

Outcomes

Feedback

Transformational Model

Resources

Services

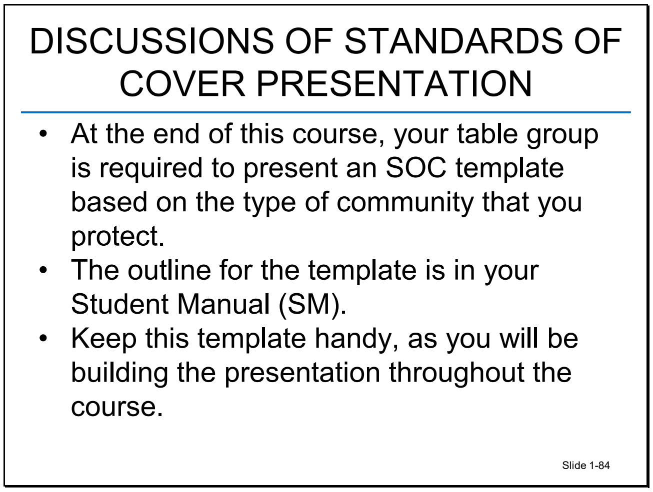
Outcomes

Feedback

Fire

EMS

VII. DISCUSSIONS OF STANDARDS OF COVER PRESENTATION



DISCUSSIONS OF STANDARDS OF COVER PRESENTATION

- At the end of this course, your table group is required to present an SOC template based on the type of community that you protect.
- The outline for the template is in your Student Manual (SM).
- Keep this template handy, as you will be building the presentation throughout the course.

Slide 1-84

A. In your table group, you will develop an SOC template that is based on the type of community that you protect.

B. As we cover the various pieces of the document, you will want to create your SOC and begin to work on how you will put the information into the document to show where you are at currently.

C. You have the freedom, as a group, to lay the document out as you see fit. The final document should be a template that only needs the correct information to be inserted into the correct place. Each member of your group should be able to take your template, go home, and plug the information in where needed to develop a document that demonstrates how your agency provides the proper coverage for the identified risk.

For example:

1. Under the heading of “Area Description,” there is a subheading of “Population.”

2. You would describe how you are going to list your population (daytime, nighttime, seasonal) and where you would get this information (census data, GIS data, county assessor’s office, etc.).

3. For the various headings, you will, as a group, need to determine what information you want to include in this section and where best to obtain the information.

4. Please use the final project outline as a guide to lead you through the development of this final project document.

STANDARDS OF COVER PRESENTATION OUTLINE

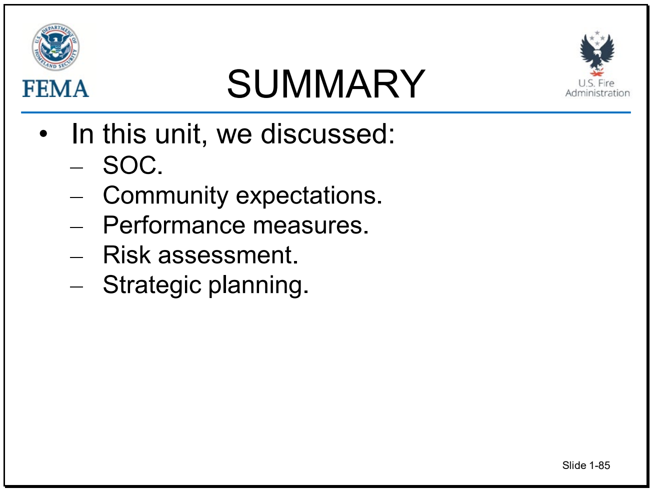
This page provides you with an outline for your SOC presentation that is to be presented at the end of this course.

I. Introduction
II. Table of Contents
III. Executive Summary
IV. Description of Community Served
    1. Legal Basis
    2. History of the Agency
    3. Financial
    4. Area Description
        A. Topography
        B. Climate
        C. Population
        D. Disaster Potentials
        E. Area Development
        F. Demographic Features
V. Services Provided
    1. Service Delivery Programs
        A. Fire Suppression
        B. Rescue
        C. Medical
        D. Hazardous Materials
    2. Current Deployment
        A. Resources
        B. Response Areas
    3. Community Response History
        A. Overall Response Area
        B. Station Response Areas
        C. Other Response Areas
VI. Community Expectations and Performance Goals
    1. Community Expectations
    2. Mission Statement
    3. Performance Goals
    4. Community Service Expectations
    5. Community Service Priorities
VII. Community Risk Assessment and Risk Levels
    1. Risk Assessment Methodology
        A. Methodology
        B. Planning Areas/Zones
    2. Risk Assessment
        A. Fire Suppression Services
        B. Emergency Medical Services
        C. Hazardous Materials Services
        D. Rescue Services

VIII. Historical Perspective and Summary of System Performance

1. Distribution Factors
2. Concentration Factors
3. Reliability Factors
4. Comparability Factors

VIII. SUMMARY



SUMMARY

- In this unit, we discussed:
  - SOC.
  - Community expectations.
  - Performance measures.
  - Risk assessment.
  - Strategic planning.

UNIT 2:
DATA ANALYSIS AND TECHNIQUES

TERMINAL OBJECTIVES

The students will be able to:

2.1 At the conclusion of the unit of instruction, create a task analysis, incorporating elements of risk assessment, including National Fire Incident Reporting System (NFIRS) data, and geographic information system (GIS) software programs.

ENABLING OBJECTIVES

The students will be able to:

2.1 Discuss deployment analysis methodology.

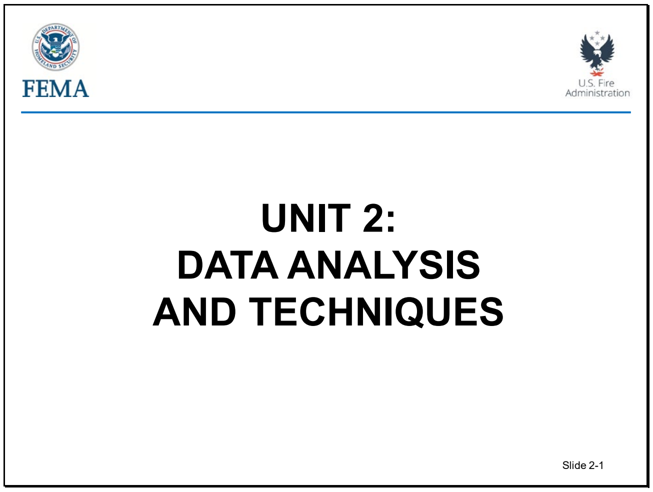
2.2 Create a deployment analysis.

2.3 Examine a critical task analysis methodology.

2.4 Create a critical task analysis.

2.5 Support NFIRS data and how they are used in risk assessment.

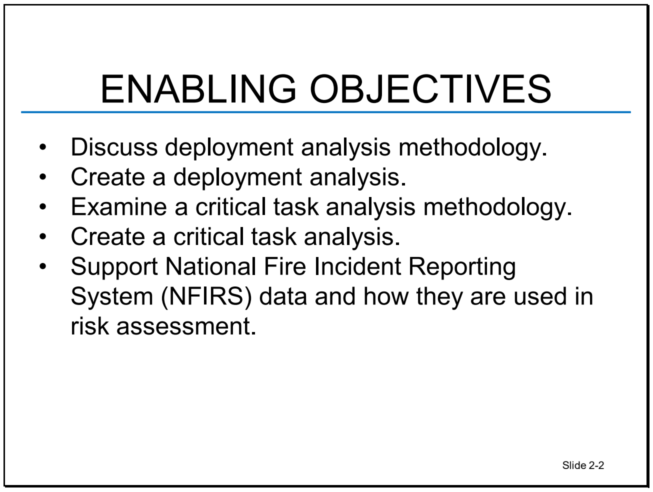
This page intentionally left blank.



FEMA

U.S. Fire Administration

UNIT 2:
DATA ANALYSIS
AND TECHNIQUES

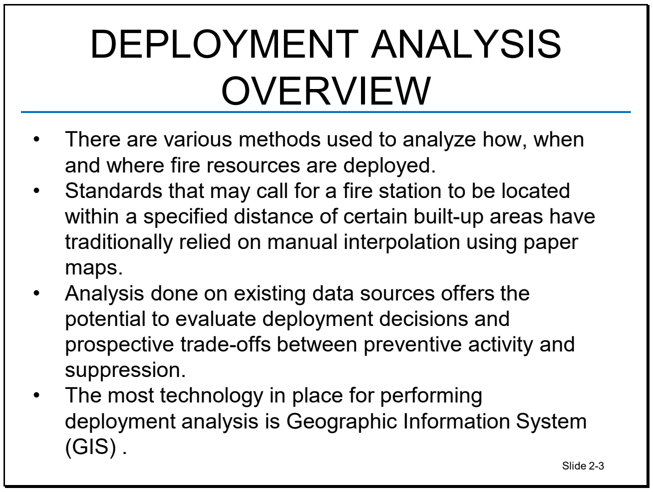


ENABLING OBJECTIVES

- Discuss deployment analysis methodology.
- Create a deployment analysis.
- Examine a critical task analysis methodology.
- Create a critical task analysis.
- Support National Fire Incident Reporting System (NFIRS) data and how they are used in risk assessment.

I. UNIT INTRODUCTION

II. DEPLOYMENT ANALYSIS OVERVIEW



DEPLOYMENT ANALYSIS OVERVIEW

- There are various methods used to analyze how, when and where fire resources are deployed.
- Standards that may call for a fire station to be located within a specified distance of certain built-up areas have traditionally relied on manual interpolation using paper maps.
- Analysis done on existing data sources offers the potential to evaluate deployment decisions and prospective trade-offs between preventive activity and suppression.
- The most technology in place for performing deployment analysis is Geographic Information System (GIS) .

Definition of deployment analysis.

A. Various methods are used to analyze how, when and where fire resources are deployed.

B. Standards that may call for a fire station to be located within a specified distance of certain built-up areas traditionally relied on manual interpolation using paper maps. Simple techniques such as drawing radii representing distances from fire stations sufficed as a means of assessing the adequacy of service delivery or Standards of Cover (SOC).

C. Analysis done on existing data sources offers the potential to evaluate deployment decisions and prospective trade-offs between preventive activity and suppression.

D. The most current and common technology in place for performing deployment analysis is Geographic Information System (GIS).

ACTIVITY 2.1

Creating a Critical Task Analysis — Defining Tasks

Purpose

To develop a list of critical tasks for four types of incidents.

Directions

1. This is a group activity.

2 Spend five to seven minutes reviewing the types of incidents listed on the worksheet.

3. List the tasks that are required to mitigate the incident in the right column labeled Tasks in the worksheet.

4. Keep in mind the following points:

   a. What tasks are critical to mitigating the incident?

   b. What tasks may not be necessary?

5. Be prepared to debrief as a class.

This page intentionally left blank.

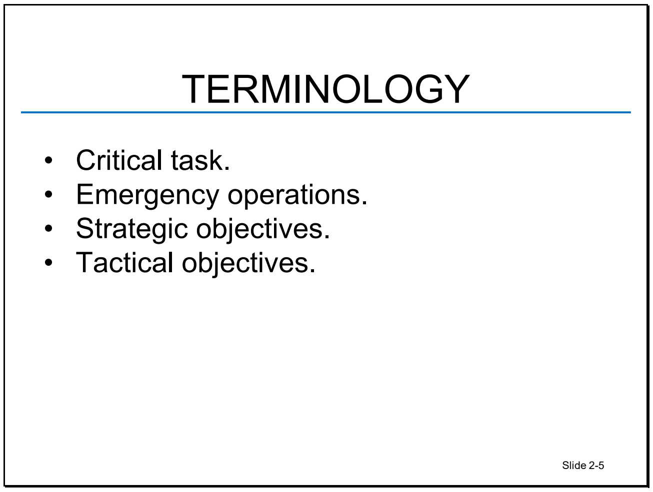
ACTIVITY 2.1 (cont'd)

Creating a Critical Task Analysis — Defining Tasks

Type of Incident	Tasks
Technical Rescue (Group Define)	
Cardiac Arrest	
Motor Vehicle Crash	
Hazmat Incident	

This page intentionally left blank.

III. TERMINOLOGY



TERMINOLOGY

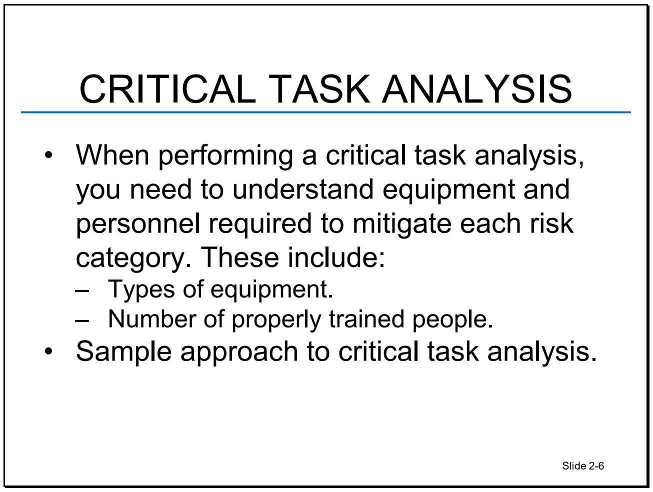
- Critical task.
- Emergency operations.
- Strategic objectives.
- Tactical objectives.

Slide 2-5

A. Key terms.

The following terminology should be defined before we proceed:

Term	Definition
Critical task	Specific tasks required, within appropriate time, to mitigate the emergency.
Emergency operations	Activities of the fire department relating to: • Rescue • Fire suppression • Emergency medical care • Special operations
Strategic objectives	These never change; there are three strategic objectives: • Life safety • Incident stabilization • Property conservation
Tactical objectives	Immediate short-term desired result of a given activity: • Rescue • Exposures • Confinement • Extinguishment • Overhaul • Ventilation • Salvage



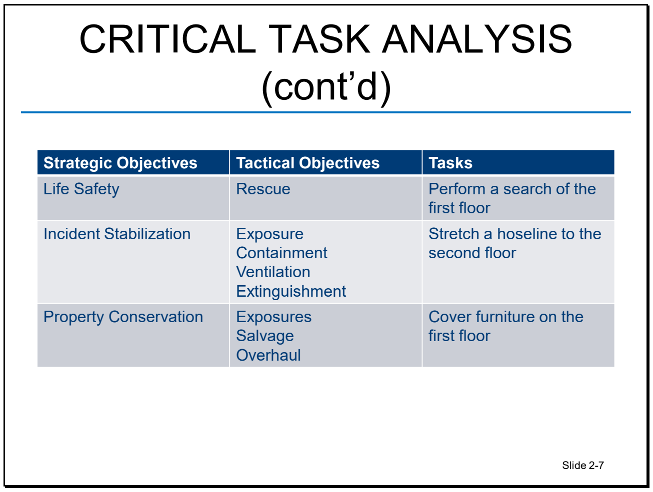
CRITICAL TASK ANALYSIS

- When performing a critical task analysis, you need to understand equipment and personnel required to mitigate each risk category. These include:
  - Types of equipment.
  - Number of properly trained people.
- Sample approach to critical task analysis.

B. Critical task analysis.

1. To respond effectively to an identified risk or level of risk, it is necessary to have an understanding of what types of equipment and numbers of properly trained people are needed to mitigate each risk category (Center for Public Safety Excellence, 2008).

2. **Identifying critical tasks** — There are some critical tasks that must be conducted by firefighters at structure fires, vehicle extrications and emergency medical incidents. To create a level of response, an assessment must be made locally to determine the capabilities of the arriving companies and individual responders. When identifying critical tasks, responder safety is the top priority.



CRITICAL TASK ANALYSIS (cont'd)

Strategic Objectives	Tactical Objectives	Tasks
Life Safety	Rescue	Perform a search of the first floor
Incident Stabilization	Exposure Containment Ventilation Extinguishment	Stretch a hoseline to the second floor
Property Conservation	Exposures Salvage Overhaul	Cover furniture on the first floor

3. The process begins with understanding the strategic and tactical objectives developed in the fire service. Tasks are individual units of work designed to meet the objectives.

4. To perform a critical task analysis for a fire incident:

    a. Begin with strategic objectives, since these never change.

    b. Identify the tactical objective that applies in this incident.

    c. Identify and list the individual tasks that are required to meet the strategic and tactical objective.

    d. Repeat steps 1-3 for each strategic objective.

This page intentionally left blank.

ACTIVITY 2.2

Creating a Critical Task Analysis — Mapping Tasks to Objectives

Purpose

To develop a critical task analysis for a fire incident.

Directions

1. Work as a group.
2. Task analysis is being done on a fire incident of a 1.5-story, wood frame, 2,000-square-foot dwelling.
3. Spend 10-12 minutes reviewing the types of incidents listed on the worksheet.
4. List the applicable tactical tasks that are required to mitigate the incident and record your answers in the Tasks column (use the worksheet below to capture items).
5. Be prepared to debrief as a class.

This page intentionally left blank.

ACTIVITY 2.2 (cont'd)

Creating a Critical Task Analysis — Mapping Tasks to Objectives

Strategic Objective	Tasks
Life Safety	
Incident Stabilization	
Property Conservation	

Use back of worksheet if more space is needed.

This page intentionally left blank.

IV. WHERE DO WE START?

WHERE DO WE START?

- Methodology.
  - Strategic objectives.
  - Tactical objectives.
  - Tasks.
- Best practices.
  - Target hazard.
  - Fire flows.
  - Pre-plans.
  - Experience.
  - National Fire Protection Association (NFPA) standards.

A. Introduction.

1. Review the methodology behind the development of a critical task analysis, specifically:

   a. Strategic objectives.

   b. Tactical objectives.

   c. Tasks.

2. Discuss best practices related to:

   a. Target hazard.

   b. Fire flows.

   c. Pre-plans.

   d. Experience.

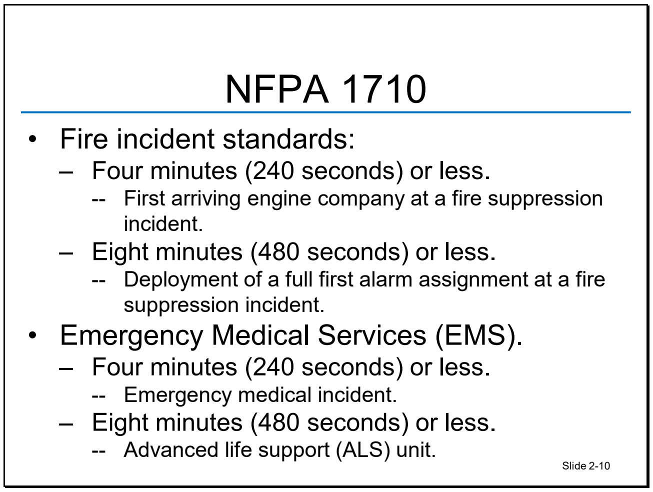
   e. National Fire Protection Association (NFPA) standards.

B. NFPA standards.

NFPA has created two standards that we follow:

1. NFPA 1710.

2. NFPA 1720.



NFPA 1710

- Fire incident standards:
  - Four minutes (240 seconds) or less.
    - First arriving engine company at a fire suppression incident.
  - Eight minutes (480 seconds) or less.
    - Deployment of a full first alarm assignment at a fire suppression incident.
- Emergency Medical Services (EMS).
  - Four minutes (240 seconds) or less.
    - Emergency medical incident.
  - Eight minutes (480 seconds) or less.
    - Advanced life support (ALS) unit.

C. NFPA 1710.

1. Explain that the standard should be treated as a benchmark or best business practice. Although generally many organizations do not meet this standard, organizations should strive to meet it through realistic planning.

2. According to NFPA 1710, the following standards apply:

   a. Fire incident standards:

      - Four minutes (240 seconds) or less.

        -- First arriving engine company at a fire suppression incident.

      - Eight minutes (480 seconds) or less.

        -- Deployment of a full first alarm assignment at a fire suppression incident.

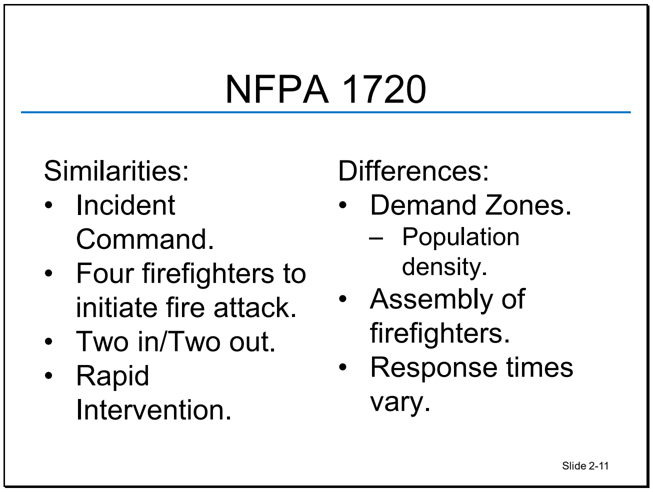
   b. Emergency Medical Services (EMS).

      - Four minutes (240 seconds) or less.

        -- Emergency medical incident.

      - Eight minutes (480 seconds) or less.

        -- Advanced life support (ALS) unit.



NFPA 1720

Similarities:
- Incident Command.
- Four firefighters to initiate fire attack.
- Two in/Two out.
- Rapid Intervention.

Differences:
- Demand Zones.
  - Population density.
- Assembly of firefighters.
- Response times vary.

D. Comparison of NFPA 1710 and NFPA 1720.

1. The following standards are the same for both NFPA 1710 and NFPA 1720:

   a. Incident Command.

   b. Four firefighters to initiate fire attack.

   c. Two in/Two out.

   d. Rapid Intervention.

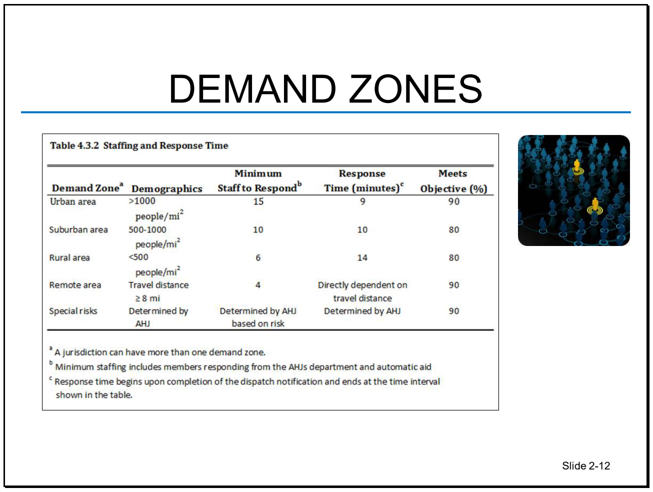
2. The following standards differ in NFPA 1710 and NFPA 1720:

   a. Demand Zone — Population density.

   b. Assembly of firefighters.

   c. Response times vary.

3. Requirement **4.2 Community Risk Management** applies to this section.



DEMAND ZONES

Table 4.3.2 Staffing and Response Time

Demand Zone[a]	Demographics	Minimum Staff to Respond[b]	Response Time (minutes)[c]	Meets Objective (%)
Urban area	>1000 people/mi^2	15	9	90
Suburban area	500-1000 people/mi^2	10	10	80
Rural area	<500 people/mi^2	6	14	80
Remote area	Travel distance ≥ 8 mi	4	Directly dependent on travel distance	90
Special risks	Determined by AHJ	Determined by AHJ based on risk	Determined by AHJ	90

[a] A jurisdiction can have more than one demand zone.

[b] Minimum staffing includes members responding from the AHJs department and automatic aid

[c] Response time begins upon completion of the dispatch notification and ends at the time interval shown in the table.

E. Demand Zones (from NFPA 1720).

1. Demand Zones are zones based on the population per square mile. The following are the types of Demand Zones and the requirements to determine the correct zone.

Table 4.3.2 Staffing and Response Time

Demand Zone[a]	Demographics	Minimum Staff to Respond[b]	Response Time (minutes)[c]	Meets Objective (%)
Urban area	>1000 people/mi^2	15	9	90
Suburban area	500-1000 people/mi^2	10	10	80
Rural area	<500 people/mi^2	6	14	80
Remote area	Travel distance ≥ 8 mi	4	Directly dependent on travel distance	90
Special risks	Determined by AHJ	Determined by AHJ based on risk	Determined by AHJ	90

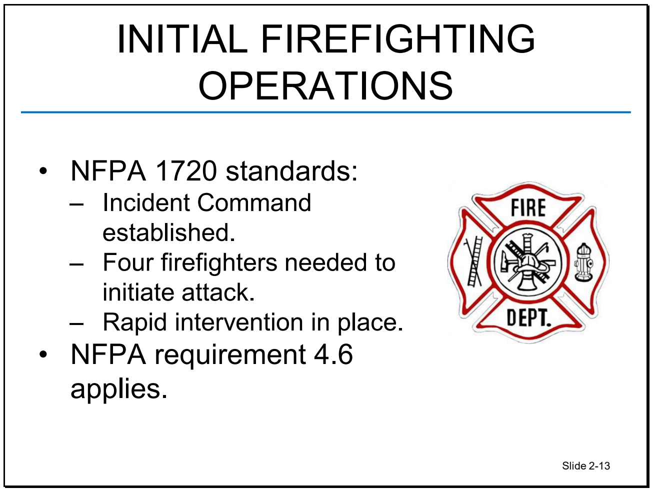
[a] A jurisdiction can have more than one demand zone.

[b] Minimum staffing includes members responding from the AHJs department and automatic aid

[c] Response time begins upon completion of the dispatch notification and ends at the time interval shown in the table.

2. For instance, the population of Urban Demand Zones is greater than 1,000 people per square mile, while that of Rural Demand Zones is less than 500 people per square mile.

3. Staffing and deployment are based on Demand Zones. See the chart below for more detail.



INITIAL FIREFIGHTING OPERATIONS

- NFPA 1720 standards:
  - Incident Command established.
  - Four firefighters needed to initiate attack.
  - Rapid intervention in place.
- NFPA requirement 4.6 applies.

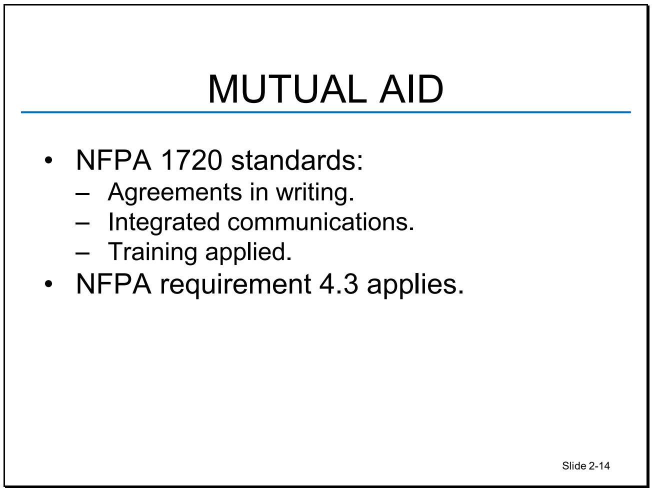
F. Initial firefighting operations.

1. The standard for initial firefighting operations as based on NFPA 1720 is:

   a. Four firefighters to initiate interior attack.

   b. Two in/Two out.

2. Requirement **4.6 Initial Fire-Fighting Operations** applies to this section.



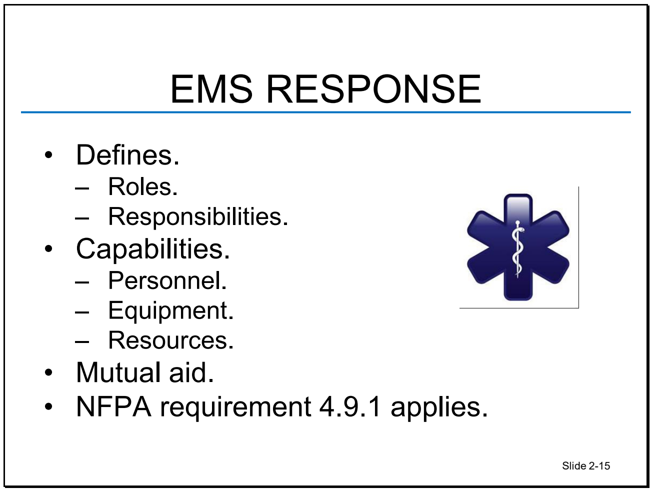
MUTUAL AID

- NFPA 1720 standards:
  - Agreements in writing.
  - Integrated communications.
  - Training applied.
- NFPA requirement 4.3 applies.

G. Mutual aid.

1. Mutual aid procedures should be in place and in writing. They should also identify common communication requirements, as well as training.

2. Requirement **4.3 Mutual Aid** applies to this section.



EMS RESPONSE

- Defines.
  - Roles.
  - Responsibilities.
- Capabilities.
  - Personnel.
  - Equipment.
  - Resources.
- Mutual aid.
- NFPA requirement 4.9.1 applies.

H. EMS response.

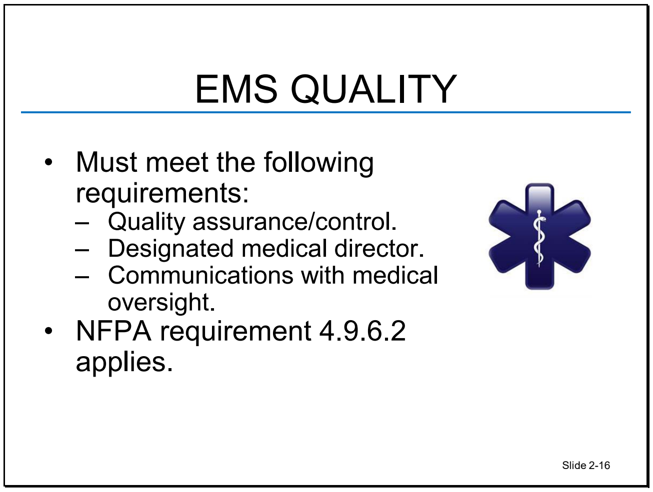
1. EMS response:

   a. Defines roles and responsibilities.

   b. Capabilities including personnel, equipment and resources.

   c. Mutual aid.

2. Requirement **4.9.1 EMS Response** applies to this section.



EMS QUALITY

- Must meet the following requirements:
  - Quality assurance/control.
  - Designated medical director.
  - Communications with medical oversight.
- NFPA requirement 4.9.6.2 applies.

I. EMS quality.

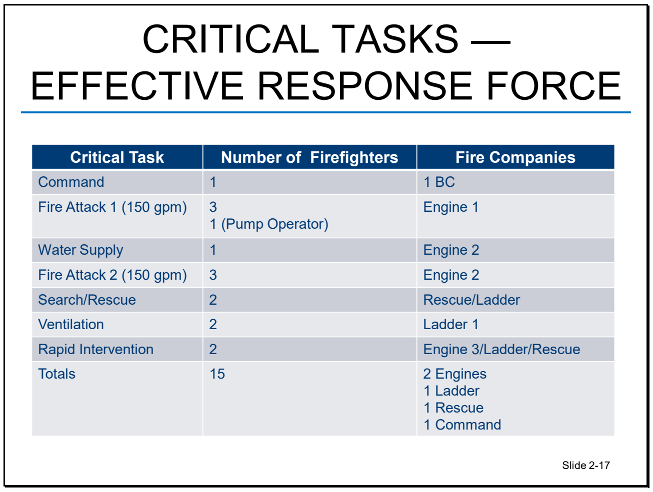
1. Organizations must meet the following requirements:

   a. Review actions of first responders and basic life support (BLS).

   b. Identify a medical director for fire departments with ALS.

c. Provide a mechanism for immediate communication with EMS.

2. Requirement **4.9.6.2 EMS Quality** applies to this section.



CRITICAL TASKS —
EFFECTIVE RESPONSE FORCE

Critical Task	Number of Firefighters	Fire Companies
Command	1	1 BC
Fire Attack 1 (150 gpm)	3 1 (Pump Operator)	Engine 1
Water Supply	1	Engine 2
Fire Attack 2 (150 gpm)	3	Engine 2
Search/Rescue	2	Rescue/Ladder
Ventilation	2	Ladder 1
Rapid Intervention	2	Engine 3/Ladder/Rescue
Totals	15	2 Engines 1 Ladder 1 Rescue 1 Command

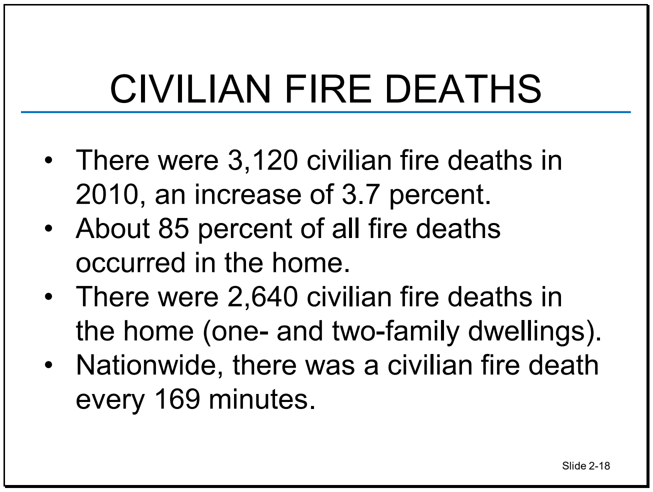
Slide 2-17

J. Effective response force.

1. Review with the class the critical task list in a minimum response to a 2,000-square-foot residential dwelling. Keep in mind that fire flows are minimal and do not reflect the increased occupancy load now being experienced in modern residential structures.

Critical Task	Number of Firefighters	Fire Companies
Command	1	1 BC
Fire Attack 1 (150 gpm)	3 1 (Pump Op)	Engine 1
Water Supply	1	Engine 2
Fire Attack 2 (150 gpm)	3	Engine 2
Search/Rescue	2	Rescue/Ladder
Ventilation	2	Ladder 1
Rapid Intervention	2	Engine 3/Ladder/ Rescue
Totals	15	2 Engines 1 Ladder 1 Rescue 1 Command

2. Requirement **5.2.4.2 2 Initial Full Alarm Assignment Capability** applies to this section.



CIVILIAN FIRE DEATHS

- There were 3,120 civilian fire deaths in 2010, an increase of 3.7 percent.
- About 85 percent of all fire deaths occurred in the home.
- There were 2,640 civilian fire deaths in the home (one- and two-family dwellings).
- Nationwide, there was a civilian fire death every 169 minutes.

Slide 2-18

K. Civilian fire deaths.

1. According to NFPA, the civilian fire death statistics included the following:

    a. There were 3,120 civilian fire deaths in 2010, an increase of 3.7 percent.

    b. About 85 percent of all fire deaths occurred in the home.

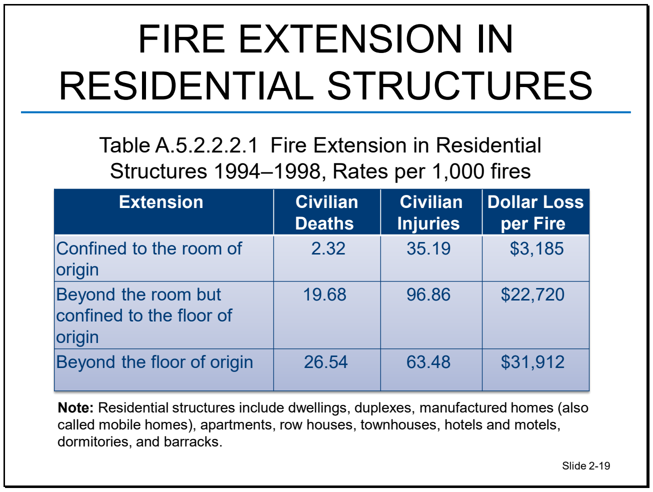
    c. There were 2,640 civilian fire deaths in the home (one- and two-family dwellings).

    d. Nationwide, there was a civilian fire death every 169 minutes.

2. Fire extension in residential fires can help explain the higher cost in terms of human life and money.

3. Refer to the chart below for more details.

    (Source: Karter, Michael J. “Fire Loss in the United States During 2010.” NFPA, September 2011.)



FIRE EXTENSION IN RESIDENTIAL STRUCTURES

Table A.5.2.2.2.1 Fire Extension in Residential Structures 1994–1998, Rates per 1,000 fires

Extension	Civilian Deaths	Civilian Injuries	Dollar Loss per Fire
Confined to the room of origin	2.32	35.19	$3,185
Beyond the room but confined to the floor of origin	19.68	96.86	$22,720
Beyond the floor of origin	26.54	63.48	$31,912

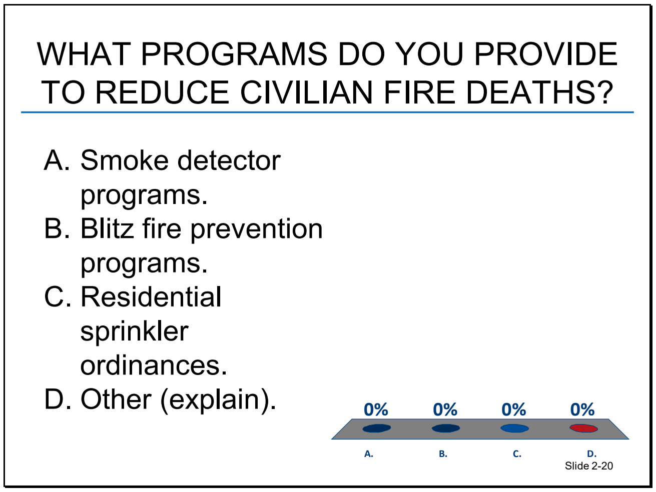
Note: Residential structures include dwellings, duplexes, manufactured homes (also called mobile homes), apartments, row houses, townhouses, hotels and motels, dormitories, and barracks.

Fire Extension in Residential Structures
1994-1998, Rates per 1,000 fires

Extension	Civilian Deaths	Civilian Injuries	Dollar Loss per Fire
Confined to the room of origin	2.32	35.19	$3,185
Beyond the room but confined to the floor of origin	19.68	96.86	$22,720
Beyond the floor of origin	26.54	63.48	$31,912

Note: Residential structures include dwellings, duplexes, manufactured homes (also called mobile homes), apartments, row houses, townhouses, hotels and motels, dormitories, and barracks.

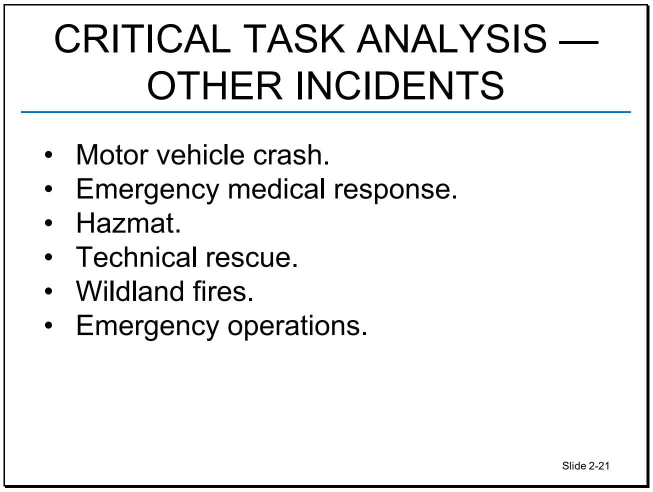
(Sources: “NFPA Annual Fire Experience Survey and NFIRS.”)



WHAT PROGRAMS DO YOU PROVIDE TO REDUCE CIVILIAN FIRE DEATHS?

A. Smoke detector programs.
B. Blitz fire prevention programs.
C. Residential sprinkler ordinances.
D. Other (explain).

V. CRITICAL TASK ANALYSIS — OTHER INCIDENTS



CRITICAL TASK ANALYSIS — OTHER INCIDENTS

- Motor vehicle crash.
- Emergency medical response.
- Hazmat.
- Technical rescue.
- Wildland fires.
- Emergency operations.

A. There are other types of incidents that require critical task analysis:

1. Motor vehicle crash.
2. Emergency medical response.
3. Hazmat.
4. Technical rescue.
5. Wildland fires.
6. Emergency operations.

B. Other methods of developing critical task analysis will be discussed below.

C. Motor vehicle crash — objectives and tasks may include:

  1. Command.

  2. EMS.

  3. Extrication.

  4. Fire prevention.

D. How does the following affect strategic planning:

  1. Change in community demographics?

  2. Upgrades to transportation?

  3. Rapid development?

E. Cardiac arrest — objectives and tasks may include:

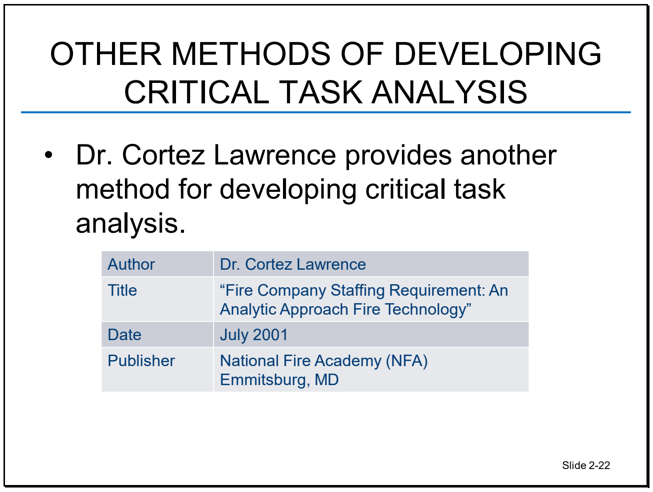
  1. How to operate at a cardiac arrest.

  2. Benchmarks available include:

    a. Response.

    b. Treatment.

    c. Transport.



OTHER METHODS OF DEVELOPING CRITICAL TASK ANALYSIS

- Dr. Cortez Lawrence provides another method for developing critical task analysis.

Author	Dr. Cortez Lawrence
Title	“Fire Company Staffing Requirement: An Analytic Approach Fire Technology”
Date	July 2001
Publisher	National Fire Academy (NFA) Emmitsburg, MD

F. Other methods of developing critical task analysis.

1. Dr. Cortez Lawrence provides another method for developing critical task analysis in his article titled "Fire Company Staffing Requirement: An Analytic Approach Fire Technology" published July 2001, Volume 37, Issue 1, p. 199-218, by the National Fire Academy (NFA), Emmitsburg, Md.

   a. There has been considerable discussion and concern about fire company staffing managers. Most of the studies undertaken have lacked objectivity in that the researcher had a vested interest in justifying higher numbers of fire personnel. Exacerbating the issue has been the development and implementation of NFPA Standard 1500, which includes recommendations for fire attack staffing.

   b. The City of Auburn Fire Division [Auburn Alabama] has the mission of offensively controlling pre-flashover fires involving a room and contents. Eight years ago, the city purchased replacement pumpers on custom chassis. These were 2,000-gpm units with two-man cabs. To provide firefighters at the scene required personnel to ride the rear step or "tail board." Court cases and common sense proscribed rear-step riding.

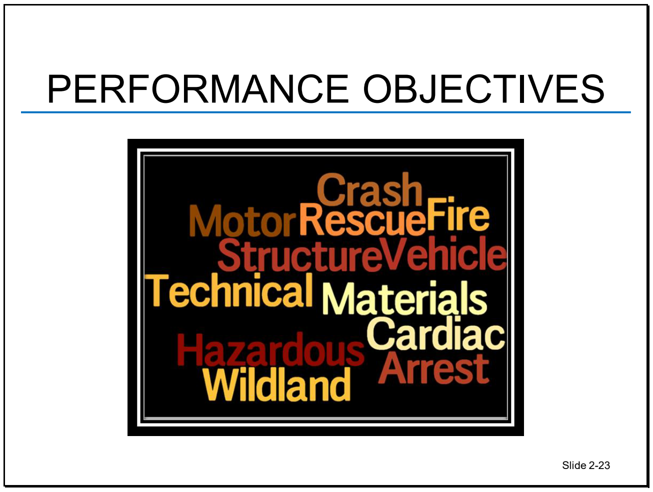
   c. The city was faced with the question, "How many firefighters do we need at most incidents and how do we get them there?" The last part of the question was arbitrarily answered by the purchase of a carry-all vehicle for a citywide crew response vehicle. The first parts of the question were the issues addressed in Lawrence's article.

   d. To determine optimal staffing, tests were conducted which would simulate operations most common to the division's mission. Sixty-eight weeks of calls were analyzed to determine the types of action most often required at emergency scenes. The independent variable in this quasi-experimental design was crew size. Goal achievement times were the dependent variable. Constants included tasks, equipment, and training levels. Crews were assembled from 25 fire recruits who were physically fit, trained in each task, and provided with professional, experienced leadership. No tasks were beyond the firefighters' skill levels.

e. Based upon the results of the testing, the most efficient staffing level for specific situations became apparent. The test results showed an inverse relationship between crew sizes and task performance time. For most tasks, however, the differences between three- and four-man crews were not as significant as between two- and three-man crews. While time is not the only selection criterion on staffing, the test resulted in the recommendation of three-firefighter crew staffing for the citywide response vehicle.

2. Another method of developing critical task analysis comes from “Fireground Field Experiments” published by the National Institute of Standards and Technology (NIST).

VI. PERFORMANCE OBJECTIVES



PERFORMANCE OBJECTIVES

Slide 2-23

A. Objectives are developed for the SOC document:

1. Strategic objectives.

2. Tactical objectives.

3. Tasks.

B. Each organization should develop clearly stated objectives, which should be:

1. Specific.

2. Measureable.

3. Achievable.

4. Relevant.

5. Timeframed.

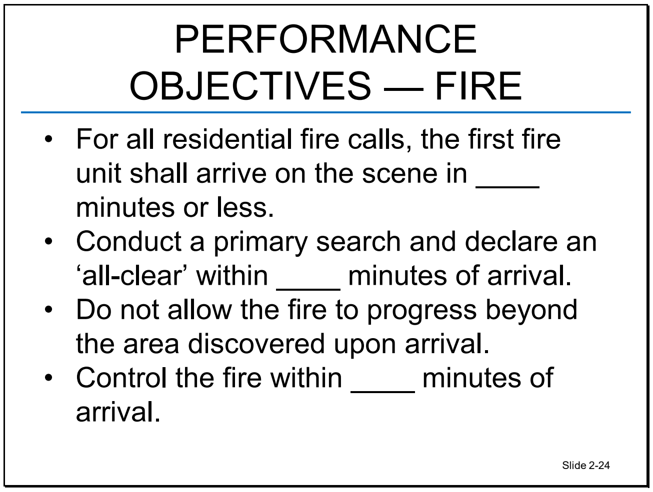
C. Objectives may be based on several factors, such as:

1. Best practices.

2. Standards.

3. Community want or need.

4. Financial considerations.



PERFORMANCE OBJECTIVES — FIRE

- For all residential fire calls, the first fire unit shall arrive on the scene in ____ minutes or less.
- Conduct a primary search and declare an ‘all-clear’ within ____ minutes of arrival.
- Do not allow the fire to progress beyond the area discovered upon arrival.
- Control the fire within ____ minutes of arrival.

D. Performance objectives — fire.

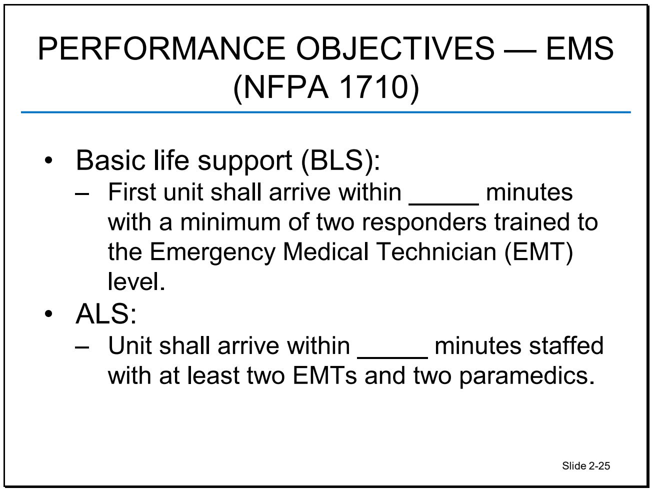
Sample performance objectives for fire include:

1. For all residential fire calls, the first fire unit shall arrive on the scene in _____ minutes or less.

2. Conduct a primary search and declare an ‘all-clear’ within _____ minutes of arrival.

3. Do not allow the fire to progress beyond the area discovered upon arrival.

4. Control the fire within _____ minutes of arrival.



PERFORMANCE OBJECTIVES — EMS (NFPA 1710)

- Basic life support (BLS):
  - First unit shall arrive within _____ minutes with a minimum of two responders trained to the Emergency Medical Technician (EMT) level.
- ALS:
  - Unit shall arrive within _____ minutes staffed with at least two EMTs and two paramedics.

Slide 2-25

E. Performance objectives — EMS

Concentration of performance objectives for EMS includes:

1. Low.
    a. First responder.
    b. CPR training.
2. Medium.

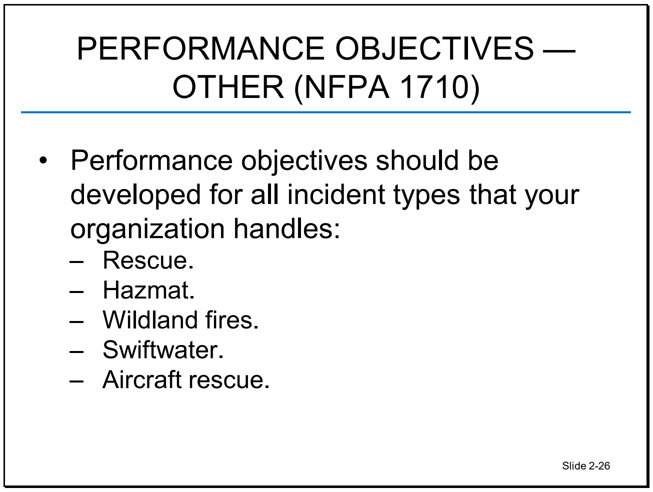
    BLS.
3. High.

    ALS.

F. Requirement **4.9 EMS** applies to this section.

G. Sample performance objectives for EMS include:

1. BLS: “First unit shall arrive within _____ minutes with a minimum of two responders trained to the Emergency Medical Technician (EMT) level.”
2. ALS: “Unit shall arrive within _____ minutes staffed with at least two EMTs and two paramedics.”
3. Refer to the following NFPA 1710 standards — **5.3 EMS** — for the requirements.



PERFORMANCE OBJECTIVES — OTHER (NFPA 1710)

- Performance objectives should be developed for all incident types that your organization handles:
  - Rescue.
  - Hazmat.
  - Wildland fires.
  - Swiftwater.
  - Aircraft rescue.

Slide 2-26

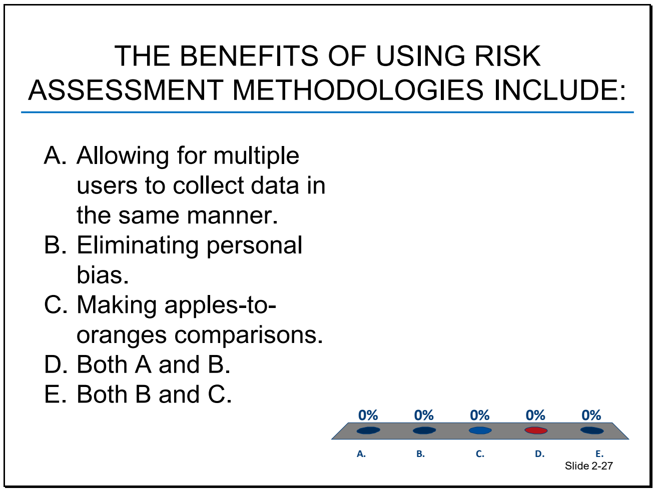
H. Performance objectives — other.

Performance objectives should be developed for all incident types that your organization handles. For example:

1. Rescue.
   a. Motor vehicle crashes.
   b. Confined spaces.
   c. High- or low-angle rescue.
   d. Building collapses.
2. Hazmat.
   a. Industry.
   b. Transportation.
   c. Terror.
   d. Vulnerable infrastructure.
3. Wildland fires.
   a. Forests.
   b. Climate.
   c. Interface.

4. Swiftwater.
    a. Rivers.
    b. Flooding.
5. Aircraft rescue.

Airports.



THE BENEFITS OF USING RISK ASSESSMENT METHODOLOGIES INCLUDE:

A. Allowing for multiple users to collect data in the same manner.
B. Eliminating personal bias.
C. Making apples-to-oranges comparisons.
D. Both A and B.
E. Both B and C.

This page intentionally left blank.

ACTIVITY 2.3

Creating a Critical Task Analysis — Use in Planning

Purpose

To walk through critical task analysis.

Directions

1. This is a group activity.
2. The instructor will provide you with the parameters of this task analysis:
   a. Dwelling fire.
   b. 1.5 stories.
   c. Wood frame.
   d. 2,000 square feet.
3. Spend 15-20 minutes identifying how your organization will respond to this incident and listing the critical tasks.
4. Use the worksheet provided in your SMs.
5. Fill out an individual worksheet for your situation.
6. Keep your worksheets to use in subsequent units.

This page intentionally left blank.

ACTIVITY 2.3 (cont'd)

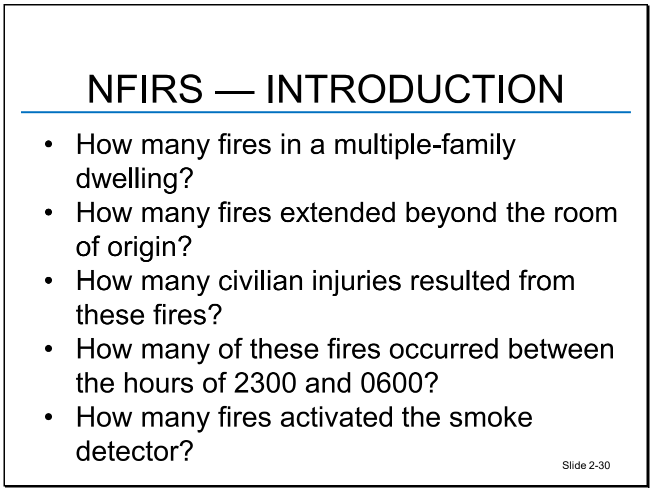
Creating a Critical Task Analysis — Use in Planning

Strategic Objectives	Tactical Objectives	Apparatus	Personnel
Life Safety	Rescue		
Incident Stabilization	Exposure Containment Ventilation Extinguishment		
Property Conservation	Exposures Salvage Overhaul		

Use back of worksheet if more space is needed.

This page intentionally left blank.

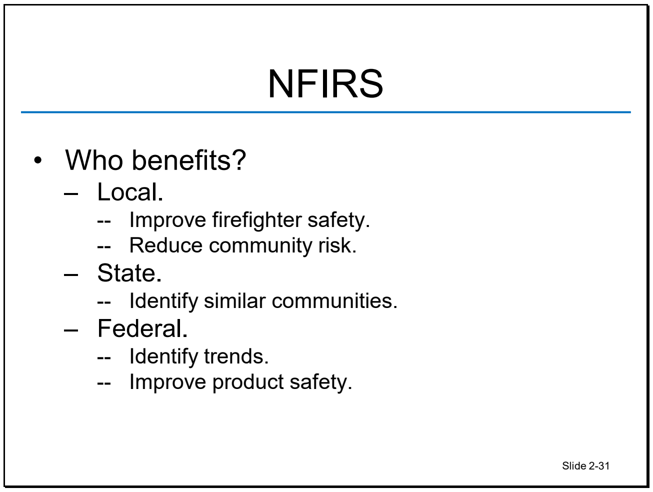
VII. NATIONAL FIRE INCIDENT REPORTING SYSTEM — INTRODUCTION



NFIRS — INTRODUCTION

- How many fires in a multiple-family dwelling?
- How many fires extended beyond the room of origin?
- How many civilian injuries resulted from these fires?
- How many of these fires occurred between the hours of 2300 and 0600?
- How many fires activated the smoke detector?

VIII. NATIONAL FIRE INCIDENT REPORTING SYSTEM — SYSTEM OVERVIEW



NFIRS

- Who benefits?
  - Local.
    - Improve firefighter safety.
    - Reduce community risk.
  - State.
    - Identify similar communities.
  - Federal.
    - Identify trends.
    - Improve product safety.

A. NFIRS provides valuable information on all types of incidents. The data can be used in multiple ways, from simply totaling data from the system to performing an in-depth analysis on the data to look for trends.

B. The benefits increase when a database that combines the fire experience of many local fire departments is used.

C. Who benefits from incident reporting and the sharing of information?

1. Local:

   a. Improve firefighter safety.

   b. Reduce community risk.

2. State:

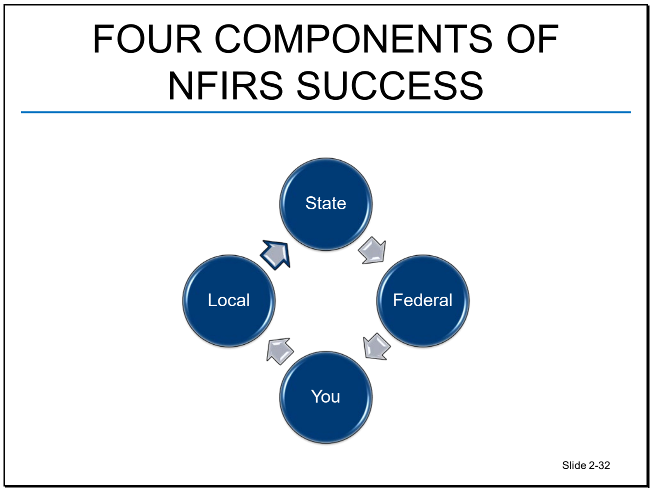
   Identify similar communities.

3. Federal:

   a. Identify trends.

   b. Improve product safety.

D. NFIRS components.



FOUR COMPONENTS OF NFIRS SUCCESS

Slide 2-32

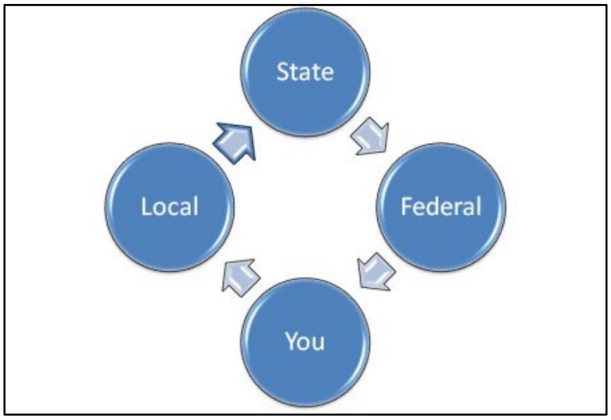
The success of the system depends on commitment at four levels:

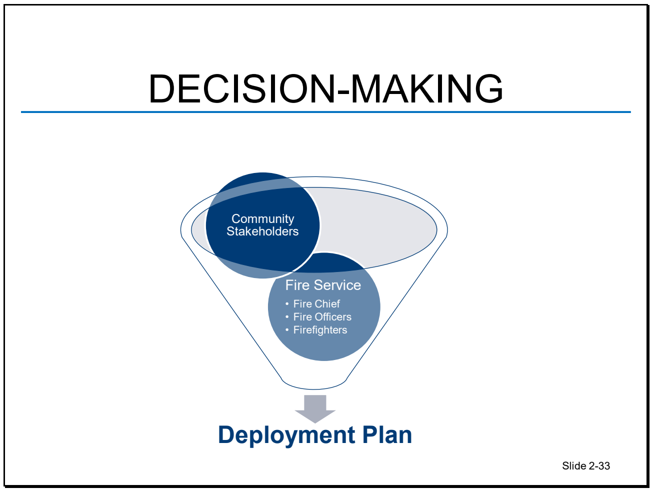
1. **Local** — NFIRS success begins with the commitment of the local fire department to collect accurate data in a timely manner.

2. **State** — The state fire marshal or another agency that manages NFIRS must support the efforts of fire departments and encourage the participation of non-reporting departments.

3. **Federal** — The U.S. Fire Administration (USFA) and the NFA provide NFIRS training opportunities.

4. **Program Manager** — Working at the local fire agency level, the NFIRS Program Manager is responsible for managing the system, ensuring quality reports, providing feedback to their department and increasing participation in NFIRS.





DECISION-MAKING

Community Stakeholders

Fire Service
- Fire Chief
- Fire Officers
- Firefighters

Deployment Plan

E. Use of NFIRS data.

1. The main use of NFIRS data is to aid in decision-making and the development of the deployment plan, specifically in the attempt to:

   a. Identify incident patterns.

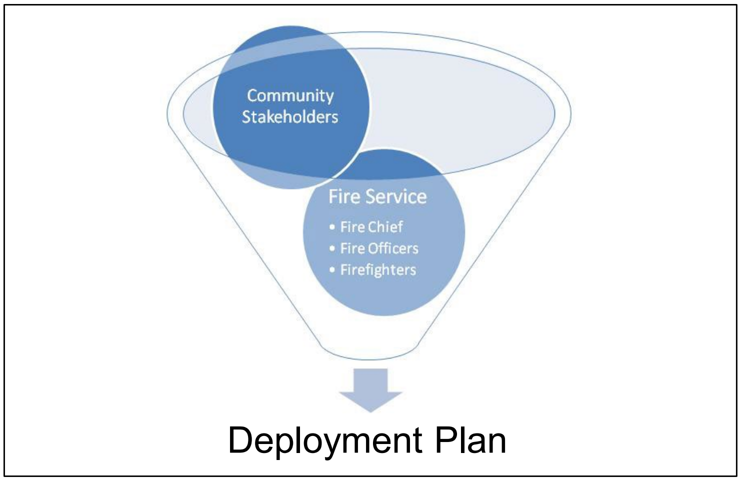
   b. Identify common areas for prevention.

   c. Identify high-risk targets.

   d. Identify geographic areas.

      - Response times.

      - Deployment strategy.



2. Decision-making can be influenced by community stakeholders as well as the fire service.

3. Community stakeholders can be any of the following:

   a. Individuals within the community.

   b. Local politicians.

   c. People who make budget decisions for fire departments.

4. Decisions can be made about the staffing, deployment and station placement with regard to:

   a. Fires.

   b. EMS.

   c. Hazmat.

   d. Rescue.

COMMUNITY BENEFITS

- Describe a community's fire problem.
- Strategic planning.
  - Standards of Cover (SOC).
- Plan future fire protection needs.
  - Support budget request.
    - Staffing.
    - Apparatus.
    - Stations.

F. Community benefits of using data.

1. Data are used to describe the community’s fire problem:

   a. What types of fires do we have?

   b. Where do we have them?

   c. What resources do we use?

2. Data describe a community’s fire problem and drive a needs assessment.

3. Data assist in strategic planning in terms of SOC through the development of objectives and serve as a guide for the operation of the organization.

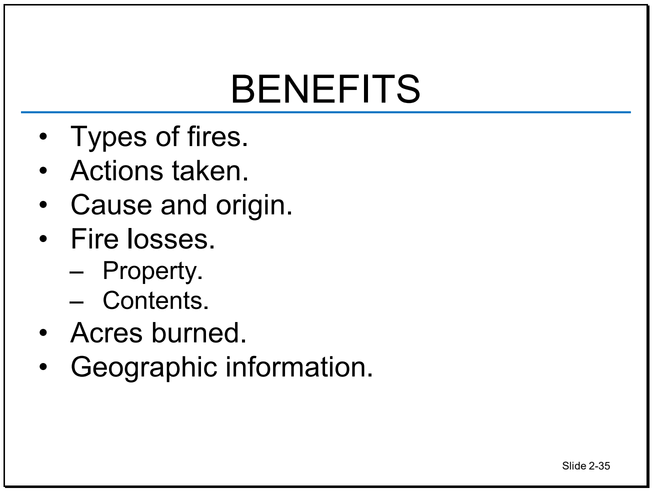
4. Data aid in planning for future fire protection needs, such as supporting budget requests regarding:

   a. Staffing.

   b. Apparatus.

   c. Stations.

5. Data are developed in association with the community, stakeholders and fire service.



BENEFITS

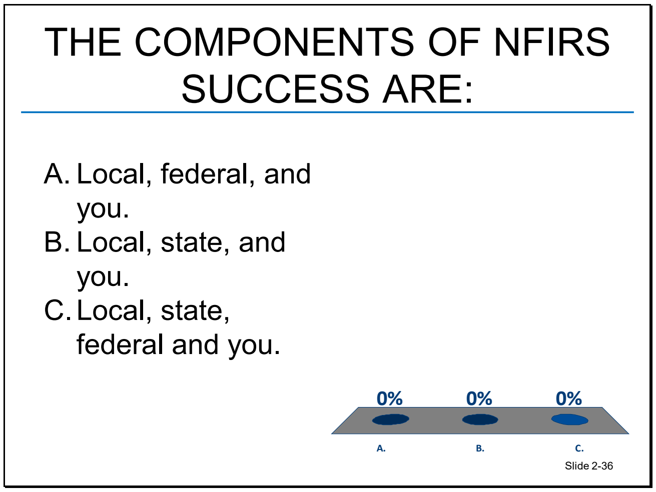
- Types of fires.
- Actions taken.
- Cause and origin.
- Fire losses.
  - Property.
  - Contents.
- Acres burned.
- Geographic information.

G. Why report incidents using NFIRS?

1. NFIRS 5.0 allows officers to report multiple actions taken at scenes. It is not unusual to perform extinguishment, overhaul, rescue and EMS at one scene. Get credit for all tasks performed.

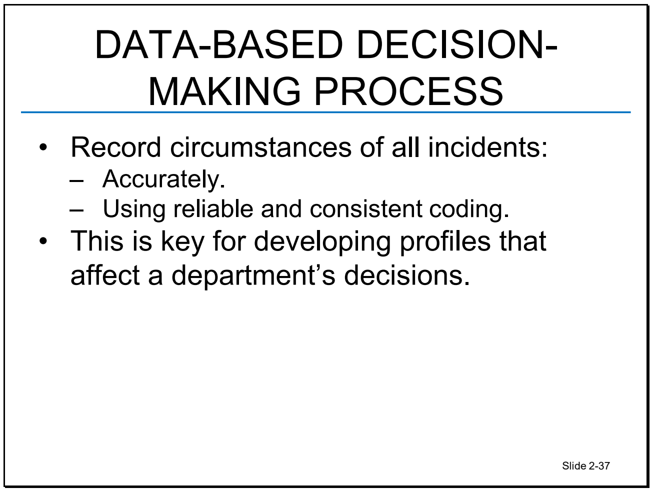
2. With utilization of easy-to-understand terms, a fire officer can provide an explanation for unusual circumstances.

3. NFIRS 5.0 provides the means to document factors that affected response and fire suppression or drove fire management strategies, such as burglar bars, structural collapse, light-weight trusses, balloon construction, access and traffic delays, natural conditions, and much more that will help paint the picture of exactly what happened.



THE COMPONENTS OF NFIRS SUCCESS ARE:

A. Local, federal, and you.
B. Local, state, and you.
C. Local, state, federal and you.



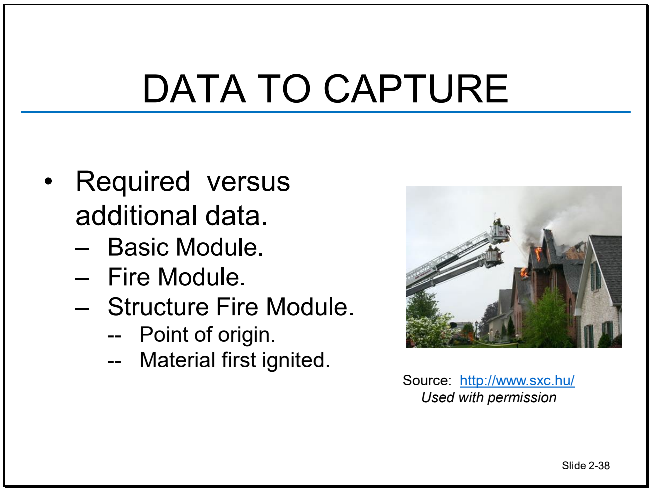
DATA-BASED DECISION-MAKING PROCESS

- Record circumstances of all incidents:
  - Accurately.
  - Using reliable and consistent coding.
- This is key for developing profiles that affect a department's decisions.

H. Data-based decision-making process.

1. The first step in the data reporting process is for fire personnel to record the circumstances of all incidents **accurately**, using a **reliable and consistent** coding methodology. This is **key** for developing profiles that affect a department's decisions.

2. Local agencies can send incident data to the state, where the information is combined with data from other fire departments into a statewide database.



DATA TO CAPTURE

- Required versus additional data.
  - Basic Module.
  - Fire Module.
  - Structure Fire Module.
    - -- Point of origin.
    - -- Material first ignited.

Source: http://www.sxc.hu/
Used with permission

I. Which incident data to capture in NFIRS?

1. The required data include:

   a. Basic Module.

   b. Fire Module.

c. Structure Fire Module, including point of origin and material first ignited.

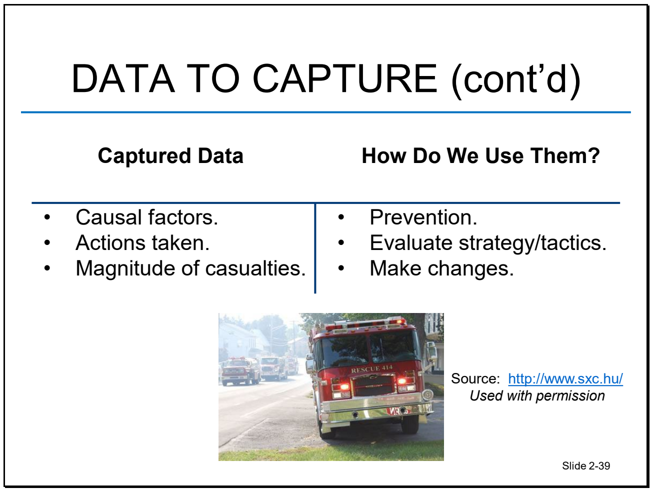
2. The report should include a description of the circumstances related to the incident, including:

   a. Causal factors contributing to the magnitude of the incident.

   b. Actions taken by the fire department to mitigate the incident.

   c. Description of the casualties or the damage resulting from the incident.

3. The chart below explains how each piece of data gathered is used:



DATA TO CAPTURE (cont'd)

Captured Data	How Do We Use Them?
Causal factors.	Prevention.
Actions taken.	Evaluate strategy/tactics.
Magnitude of casualties.	Make changes.

Source: http://www.sxc.hu/
Used with permission

Captured Data	How We Use Them?
Causal factors	Prevention
Actions taken	Evaluate strategy and/or tactics
Magnitude of casualties	Make changes

ACTIVITY 2.4

Uncaptured Data in Your Organization

Purpose

To evaluate the understanding of incident data that may not be captured by an organization, but that would be important for the organization to have.

Directions

1. Work individually.
2. Spend five to seven minutes completing the worksheet. Do the following:
    a. What uncaptured data elements do you think would be important to your organization?
    b. What would you use the data for?
3. Be prepared to debrief as a class.

This page intentionally left blank.

ACTIVITY 2.4 (cont'd)

Uncaptured Data in Your Organization

What uncaptured data elements do you think would be important to your organization?	
1	
2	
3	
4	
5	

What would you use the data for?	
1	
2	
3	
4	
5	

Use back of worksheet if more space is needed.

This page intentionally left blank.

ACTIVITY 2.4 (cont'd)

Uncaptured Data in Your Organization

What uncaptured data elements do you think would be important to your organization?	
1	
2	
3	
4	
5	

What would you use the data for?	
1	
2	
3	
4	
5	

Use back of worksheet if more space is needed.

This page intentionally left blank.

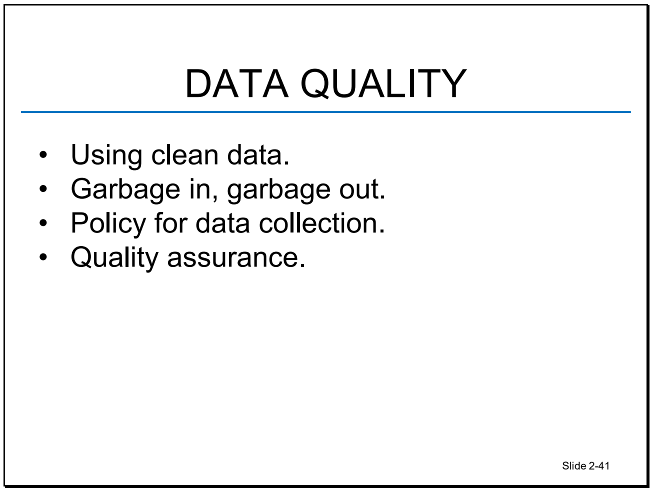
IX. DATA COLLECTION AND REQUIREMENTS

A. Data quality is critical to ensuring that the findings from a data analysis are accurate.

B. The NFPA 1710 and 1720 requirements set forth standards for:

1. Pre-dispatch data.

2. Call times.



DATA QUALITY

- Using clean data.
- Garbage in, garbage out.
- Policy for data collection.
- Quality assurance.

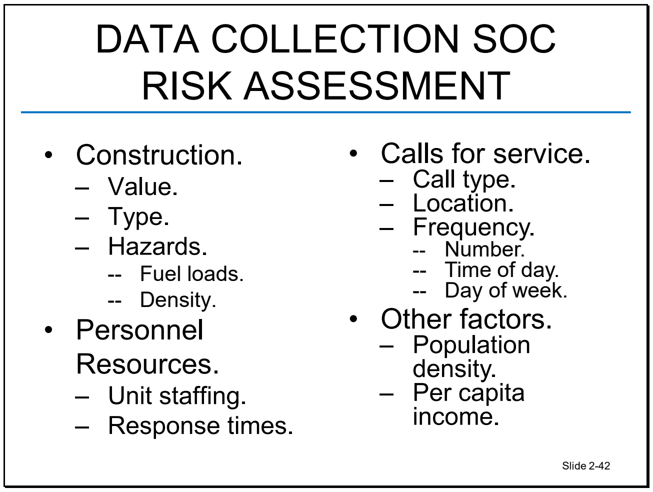
C. Data quality.

1. Using clean data is important. Remember the adage, “garbage in, garbage out.”

2. Each organization should have policies in place for:

    a. Data collection best practices.

    b. Quality assurance practices.



DATA COLLECTION SOC RISK ASSESSMENT

- Construction.
  - Value.
  - Type.
  - Hazards.
    - Fuel loads.
    - Density.
- Personnel Resources.
  - Unit staffing.
  - Response times.
- Calls for service.
  - Call type.
  - Location.
  - Frequency.
    - Number.
    - Time of day.
    - Day of week.
- Other factors.
  - Population density.
  - Per capita income.

D. Data collection for SOC and risk assessment.

1. Data from incident reporting are used to determine a community's historical risk based on its calls for service loads.

2. Data are used to analyze the following issues:

a. Construction.

- Value.

- Type.

- Hazards — fuel loads, density.

b. Calls for service.

- Call type.

- Location.

- Frequency — number, time of day, day of week.

c. Personnel resources.

- Unit staffing.

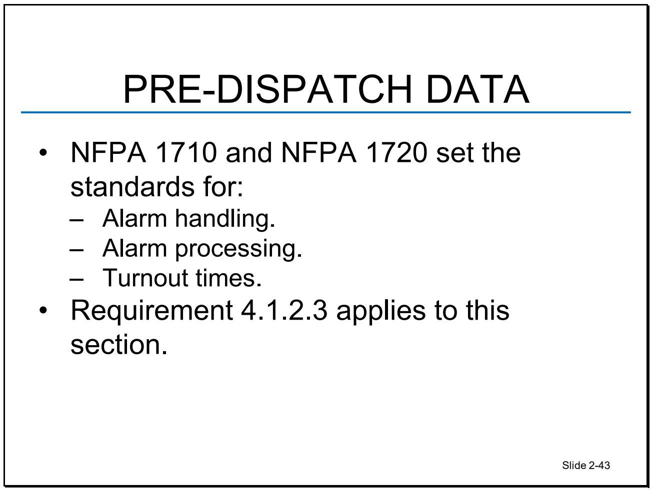
- Response times.

d. Other factors.

- Population density.

- Per capita income.

(Source: "Center of Public Safety Excellence, 2008.")



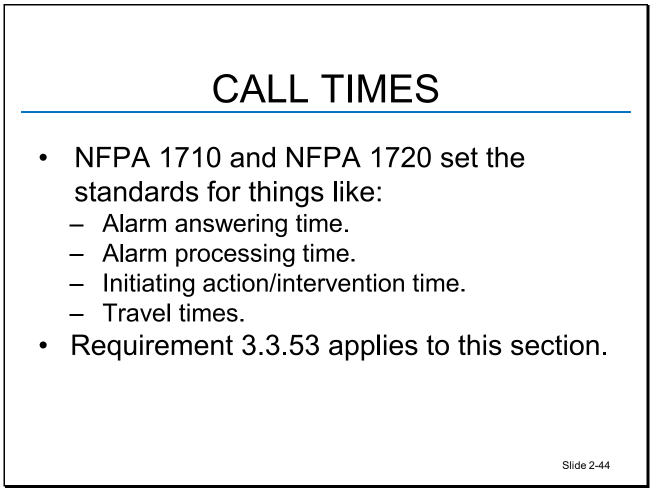
PRE-DISPATCH DATA

- NFPA 1710 and NFPA 1720 set the standards for:
  - Alarm handling.
  - Alarm processing.
  - Turnout times.
- Requirement 4.1.2.3 applies to this section.

Slide 2-43

E. Pre-dispatch data.

1. NFPA 1710 and NFPA 1720 set the standards for:
   a. Alarm handling.
   b. Alarm processing.
   c. Turnout times — including station construction.
2. Requirement **4.1.2.3 Alarm Handling** applies to this section.



CALL TIMES

- NFPA 1710 and NFPA 1720 set the standards for things like:
  - Alarm answering time.
  - Alarm processing time.
  - Initiating action/intervention time.
  - Travel times.
- Requirement 3.3.53 applies to this section.

Slide 2-44

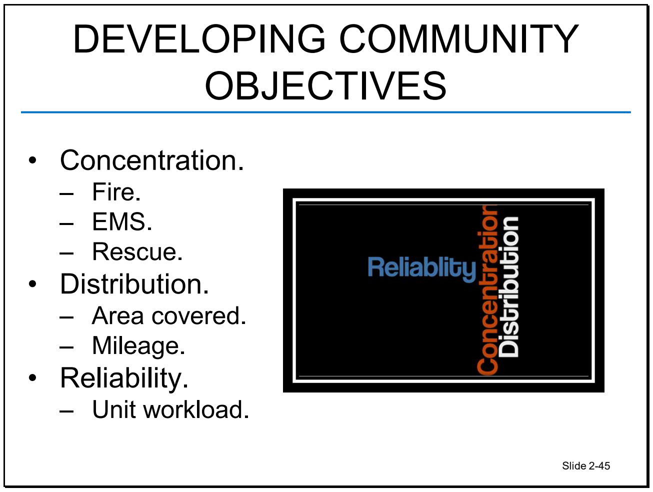
F. Call times.

1. NFPA 1710 and NFPA 1720 set the standards for call times, including:
    a. Alarm answering time.
    b. Alarm handling time.
    c. Alarm processing time.
    d. Alarm transfer time.
    e. Initiating action/intervention time.
    f. Turnout times.
    g. Travel times.
    h. Total response times.
2. Requirement **3.3.53 Time** applies to this section.

X. USING NATIONAL FIRE INCIDENT REPORTING SYSTEM FOR DATA COLLECTION

A. Topics discussed in this section include:

1. Terminology for developing community objectives.
2. NFIRS modules:
    a. Basic.
    b. Fire.
    c. Structure Fire.



DEVELOPING COMMUNITY OBJECTIVES

- Concentration.
  - Fire.
  - EMS.
  - Rescue.
- Distribution.
  - Area covered.
  - Mileage.
- Reliability.
  - Unit workload.

B. Developing community objectives terminology.

1. The terms and concepts discussed here relate to the NFIRS Basic Module and include:

   a. Concentration.

   b. Distribution.

   c. Reliability.

2. **Concentration:**

   a. The arrangement of resources spaced close enough together to create an effective response force that can be assembled at the scene within the adopted-community timeframe (Center of Public Safety Excellence, 2008).

   b. Increased risk requires increased concentration.

3. **Distribution:**

   a. Locating geographically distributed first due resources for all-risk initial intervention. Station locations need to assure rapid deployment of resources to minimize and terminate routine emergencies (Center of Public Safety Excellence, 2008).

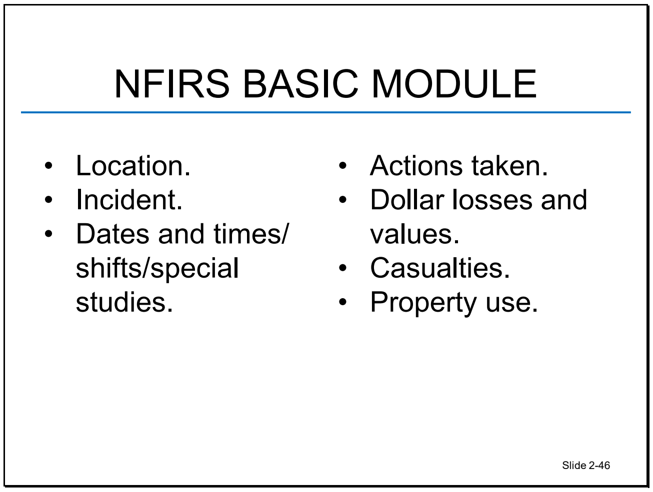
   b. Distribution is evaluated by the percentage of the population covered in the unit's first due area.

4. **Reliability:**

   a. The measure of historical performance in accordance with adopted performance measures.

   b. Revolves around time and distance from incident.

   c. Relies upon the availability and capabilities of the resource.



NFIRS BASIC MODULE

- Location.
- Incident.
- Dates and times/shifts/special studies.
- Actions taken.
- Dollar losses and values.
- Casualties.
- Property use.

Slide 2-46

C. NFIRS Basic Module.

1. The NFIRS Basic Module includes the following data elements:

   a. Location.

   b. Incident.

   c. Dates and times/shifts/special studies.

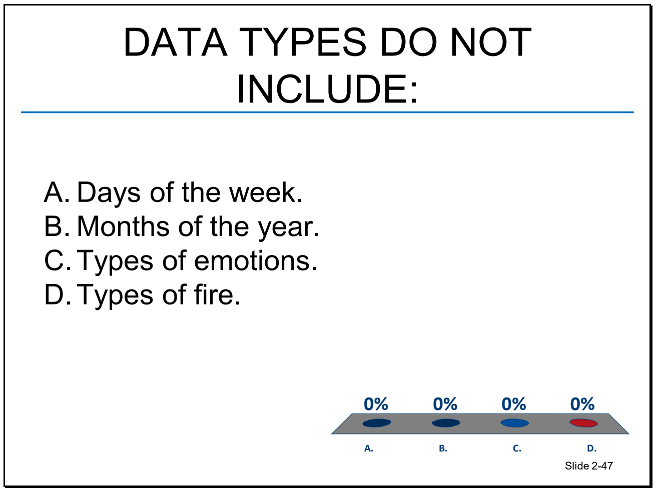
   d. Actions taken.

   e. Dollar losses and values.

   f. Casualties.

   g. Property use.

2. Each of these data elements can be used to help develop community service objectives.



DATA TYPES DO NOT INCLUDE:

A. Days of the week.
B. Months of the year.
C. Types of emotions.
D. Types of fire.

This page intentionally left blank.

ACTIVITY 2.5

Scenarios and NFIRS Modules

Purpose

To complete necessary NFIRS modules based on scenarios provided.

Directions

1. Work in a group.
2. Read the first scenario and complete the necessary NFIRS modules for that scenario.
3. Use the worksheet below to record your work.
4. Read the second scenario and complete the necessary NFIRS modules for that scenario.
5. Use the worksheet below to record your work.
6. Be prepared to debrief as a class.

Scenario 1 — Incident # 9900332 (Toaster)

The Jonesville, Wis., Fire Department, Station 1, (FDID TR300) is called at 0156 hours on July 4, 2002 to respond to a fire in a single-family dwelling. The first unit, Engine 3, arrives at 0202 hours and discovers heavy smoke and fire coming from the house. A family of four occupied the house: a father, a mother, and two children, ages 3 and 7.

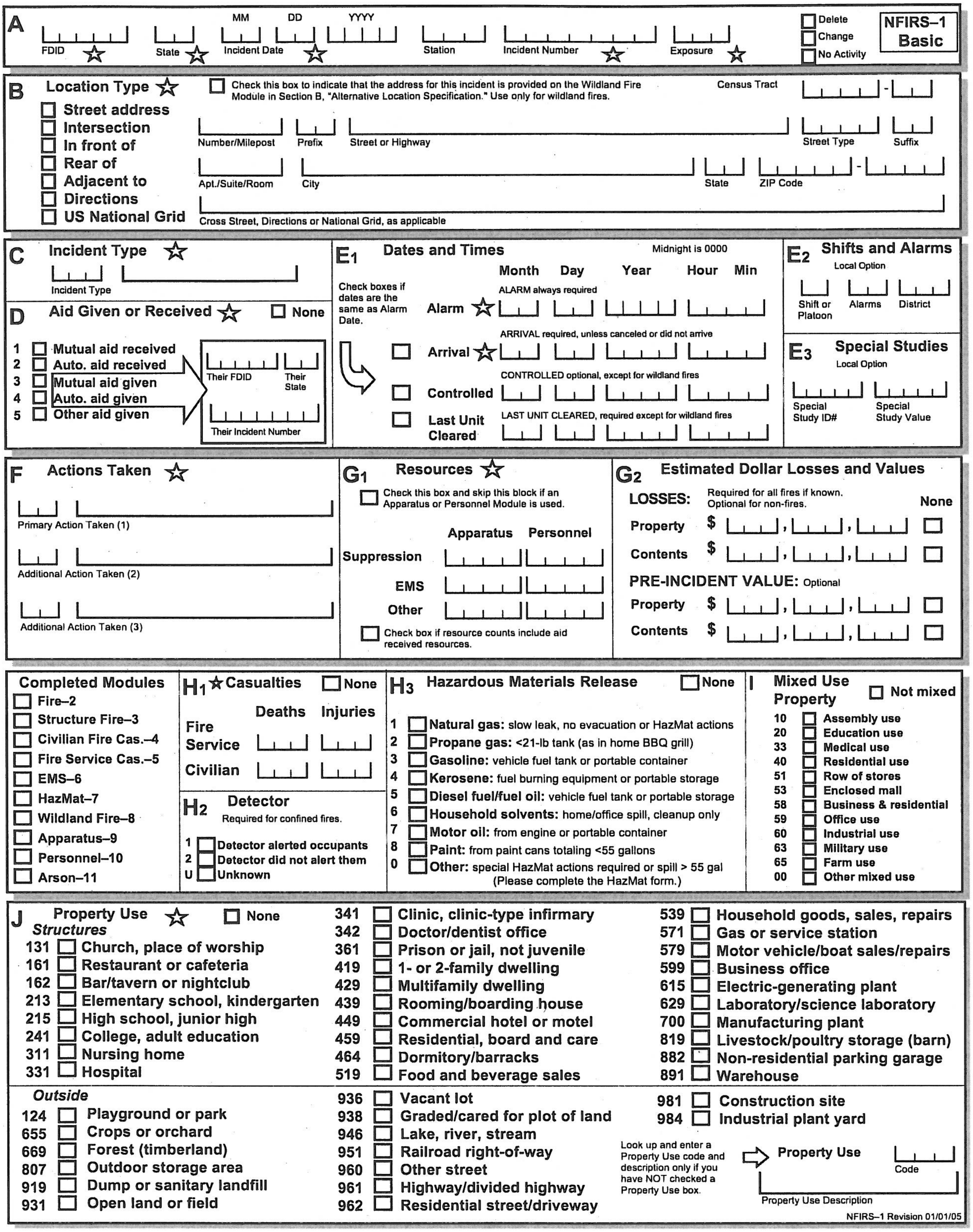
Two crew members from Engine 3 conducted a primary search for victims, located the family in bedrooms on the second floor, and rescued all the family members from the structure.

The rest of the crew brought a hoseline into the house. The fire was confined to the first floor, brought under control, and extinguished at 0215 hours. There was significant fire damage to two rooms: (1) the kitchen, where the fire originated from a defective toaster, which ignited a fire that went up a wall, and (2) the dining room. The brand of the toaster was a Toastwell, Model #ZX2, Serial #567X. The toaster was manufactured in 1985. The incident number was #9900332.

This page intentionally left blank.

ACTIVITY 2.5 (cont'd)

Scenarios and NFIRS Modules



A FDID ☆ | State ☆ | Incident Date (MM DD YYYY) ☆ | Station | Incident Number ☆ | Exposure ☆

☐ Delete
☐ Change
☐ No Activity

NFIRS–1 Basic

B Location Type ☆

☐ Check this box to indicate that the address for this incident is provided on the Wildland Fire Module in Section B, "Alternative Location Specification." Use only for wildland fires.

Census Tract

- ☐ Street address
- ☐ Intersection
- ☐ In front of
- ☐ Rear of
- ☐ Adjacent to
- ☐ Directions
- ☐ US National Grid

Number/Milepost | Prefix | Street or Highway | Street Type | Suffix

Apt./Suite/Room | City | State | ZIP Code

Cross Street, Directions or National Grid, as applicable

C Incident Type ☆

Incident Type

D Aid Given or Received ☆ ☐ None

1 ☐ Mutual aid received
2 ☐ Auto. aid received
3 ☐ Mutual aid given
4 ☐ Auto. aid given
5 ☐ Other aid given

Their FDID | Their State

Their Incident Number

E1 Dates and Times Midnight is 0000

Check boxes if dates are the same as Alarm Date.

	Month	Day	Year	Hour	Min
Alarm ☆ (ALARM always required)					
☐ Arrival ☆ (ARRIVAL required, unless canceled or did not arrive)					
☐ Controlled (CONTROLLED optional, except for wildland fires)					
☐ Last Unit Cleared (LAST UNIT CLEARED, required except for wildland fires)					

E2 Shifts and Alarms Local Option

Shift or Platoon | Alarms | District

E3 Special Studies Local Option

Special Study ID# | Special Study Value

F Actions Taken ☆

Primary Action Taken (1)

Additional Action Taken (2)

Additional Action Taken (3)

G1 Resources ☆

☐ Check this box and skip this block if an Apparatus or Personnel Module is used.

	Apparatus	Personnel
Suppression		
EMS		
Other		

☐ Check box if resource counts include aid received resources.

G2 Estimated Dollar Losses and Values

LOSSES: Required for all fires if known. Optional for non-fires.

		None
Property	$ ___,___,___	☐
Contents	$ ___,___,___	☐
PRE-INCIDENT VALUE: Optional		
Property	$ ___,___,___	☐
Contents	$ ___,___,___	☐

Completed Modules

- ☐ Fire–2
- ☐ Structure Fire–3
- ☐ Civilian Fire Cas.–4
- ☐ Fire Service Cas.–5
- ☐ EMS–6
- ☐ HazMat–7
- ☐ Wildland Fire–8
- ☐ Apparatus–9
- ☐ Personnel–10
- ☐ Arson–11

H1 ☆ Casualties ☐ None

	Deaths	Injuries
Fire Service		
Civilian		

H2 Detector Required for confined fires.

1 ☐ Detector alerted occupants
2 ☐ Detector did not alert them
U ☐ Unknown

H3 Hazardous Materials Release ☐ None

1 ☐ **Natural gas:** slow leak, no evacuation or HazMat actions
2 ☐ **Propane gas:** <21-lb tank (as in home BBQ grill)
3 ☐ **Gasoline:** vehicle fuel tank or portable container
4 ☐ **Kerosene:** fuel burning equipment or portable storage
5 ☐ **Diesel fuel/fuel oil:** vehicle fuel tank or portable storage
6 ☐ **Household solvents:** home/office spill, cleanup only
7 ☐ **Motor oil:** from engine or portable container
8 ☐ **Paint:** from paint cans totaling <55 gallons
0 ☐ **Other:** special HazMat actions required or spill > 55 gal (Please complete the HazMat form.)

I Mixed Use Property ☐ Not mixed

10 ☐ Assembly use
20 ☐ Education use
33 ☐ Medical use
40 ☐ Residential use
51 ☐ Row of stores
53 ☐ Enclosed mall
58 ☐ Business & residential
59 ☐ Office use
60 ☐ Industrial use
63 ☐ Military use
65 ☐ Farm use
00 ☐ Other mixed use

J Property Use ☆ ☐ None

Structures

131 ☐ Church, place of worship
161 ☐ Restaurant or cafeteria
162 ☐ Bar/tavern or nightclub
213 ☐ Elementary school, kindergarten
215 ☐ High school, junior high
241 ☐ College, adult education
311 ☐ Nursing home
331 ☐ Hospital
341 ☐ Clinic, clinic-type infirmary
342 ☐ Doctor/dentist office
361 ☐ Prison or jail, not juvenile
419 ☐ 1- or 2-family dwelling
429 ☐ Multifamily dwelling
439 ☐ Rooming/boarding house
449 ☐ Commercial hotel or motel
459 ☐ Residential, board and care
464 ☐ Dormitory/barracks
519 ☐ Food and beverage sales
539 ☐ Household goods, sales, repairs
571 ☐ Gas or service station
579 ☐ Motor vehicle/boat sales/repairs
599 ☐ Business office
615 ☐ Electric-generating plant
629 ☐ Laboratory/science laboratory
700 ☐ Manufacturing plant
819 ☐ Livestock/poultry storage (barn)
882 ☐ Non-residential parking garage
891 ☐ Warehouse

Outside

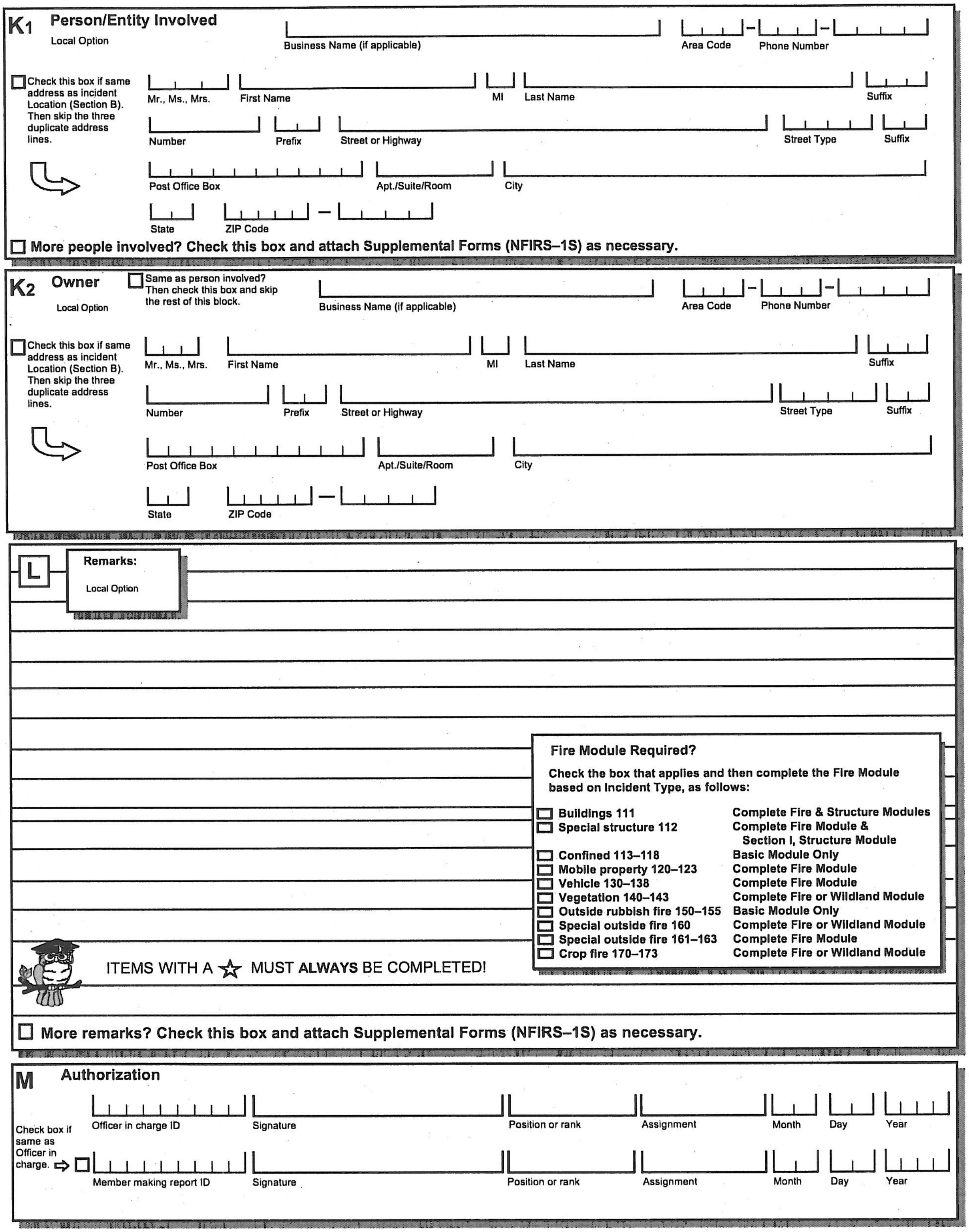
124 ☐ Playground or park
655 ☐ Crops or orchard
669 ☐ Forest (timberland)
807 ☐ Outdoor storage area
919 ☐ Dump or sanitary landfill
931 ☐ Open land or field
936 ☐ Vacant lot
938 ☐ Graded/cared for plot of land
946 ☐ Lake, river, stream
951 ☐ Railroad right-of-way
960 ☐ Other street
961 ☐ Highway/divided highway
962 ☐ Residential street/driveway
981 ☐ Construction site
984 ☐ Industrial plant yard

Look up and enter a Property Use code and description only if you have NOT checked a Property Use box. ⇨ Property Use

Code

Property Use Description

NFIRS–1 Revision 01/01/05



K1 Person/Entity Involved

Local Option

Business Name (if applicable)

Area Code – Phone Number

☐ Check this box if same address as incident Location (Section B). Then skip the three duplicate address lines.

Mr., Ms., Mrs. | First Name | MI | Last Name | Suffix

Number | Prefix | Street or Highway | Street Type | Suffix

Post Office Box | Apt./Suite/Room | City

State | ZIP Code –

☐ **More people involved? Check this box and attach Supplemental Forms (NFIRS–1S) as necessary.**

K2 Owner

Local Option

☐ Same as person involved? Then check this box and skip the rest of this block.

Business Name (if applicable)

Area Code – Phone Number

☐ Check this box if same address as incident Location (Section B). Then skip the three duplicate address lines.

Mr., Ms., Mrs. | First Name | MI | Last Name | Suffix

Number | Prefix | Street or Highway | Street Type | Suffix

Post Office Box | Apt./Suite/Room | City

State | ZIP Code –

L Remarks:

Local Option

Fire Module Required?

Check the box that applies and then complete the Fire Module based on Incident Type, as follows:

☐ Buildings 111	Complete Fire & Structure Modules
☐ Special structure 112	Complete Fire Module & Section I, Structure Module
☐ Confined 113–118	Basic Module Only
☐ Mobile property 120–123	Complete Fire Module
☐ Vehicle 130–138	Complete Fire Module
☐ Vegetation 140–143	Complete Fire or Wildland Module
☐ Outside rubbish fire 150–155	Basic Module Only
☐ Special outside fire 160	Complete Fire or Wildland Module
☐ Special outside fire 161–163	Complete Fire Module
☐ Crop fire 170–173	Complete Fire or Wildland Module

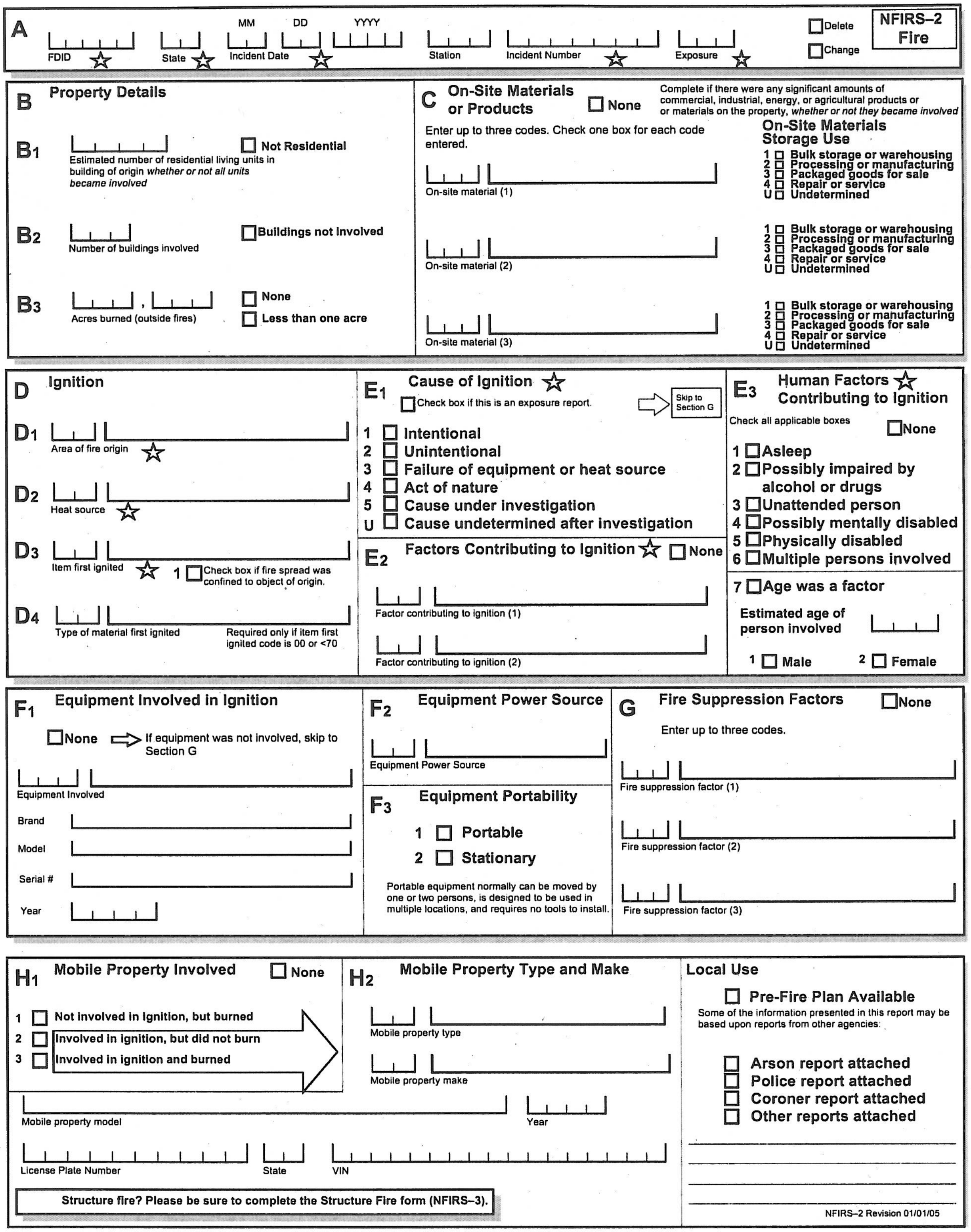
ITEMS WITH A ☆ MUST **ALWAYS** BE COMPLETED!

☐ **More remarks? Check this box and attach Supplemental Forms (NFIRS–1S) as necessary.**

M Authorization

Officer in charge ID | Signature | Position or rank | Assignment | Month | Day | Year

Check box if same as Officer in charge. ⇨ ☐ Member making report ID | Signature | Position or rank | Assignment | Month | Day | Year



A

MM DD YYYY

FDID ☆ State ☆ Incident Date ☆ Station Incident Number ☆ Exposure ☆

☐ Delete
☐ Change

NFIRS–2
Fire

B Property Details

B1 Estimated number of residential living units in building of origin *whether or not all units became involved* ☐ **Not Residential**

B2 Number of buildings involved ☐ **Buildings not involved**

B3 Acres burned (outside fires) ☐ **None** ☐ **Less than one acre**

C On-Site Materials or Products ☐ **None**

Complete if there were any significant amounts of commercial, industrial, energy, or agricultural products or materials on the property, *whether or not they became involved*

Enter up to three codes. Check one box for each code entered.

On-Site Materials Storage Use

On-site material (1)
1 ☐ **Bulk storage or warehousing**
2 ☐ **Processing or manufacturing**
3 ☐ **Packaged goods for sale**
4 ☐ **Repair or service**
U ☐ **Undetermined**

On-site material (2)
1 ☐ **Bulk storage or warehousing**
2 ☐ **Processing or manufacturing**
3 ☐ **Packaged goods for sale**
4 ☐ **Repair or service**
U ☐ **Undetermined**

On-site material (3)
1 ☐ **Bulk storage or warehousing**
2 ☐ **Processing or manufacturing**
3 ☐ **Packaged goods for sale**
4 ☐ **Repair or service**
U ☐ **Undetermined**

D Ignition

D1 Area of fire origin ☆

D2 Heat source ☆

D3 Item first ignited ☆ 1 ☐ Check box if fire spread was confined to object of origin.

D4 Type of material first ignited Required only if item first ignited code is 00 or <70

E1 Cause of Ignition ☆

☐ Check box if this is an exposure report. ⇨ Skip to Section G

1 ☐ **Intentional**
2 ☐ **Unintentional**
3 ☐ **Failure of equipment or heat source**
4 ☐ **Act of nature**
5 ☐ **Cause under investigation**
U ☐ **Cause undetermined after investigation**

E2 Factors Contributing to Ignition ☆ ☐ **None**

Factor contributing to ignition (1)

Factor contributing to ignition (2)

E3 Human Factors Contributing to Ignition ☆

Check all applicable boxes ☐ **None**

1 ☐ **Asleep**
2 ☐ **Possibly impaired by alcohol or drugs**
3 ☐ **Unattended person**
4 ☐ **Possibly mentally disabled**
5 ☐ **Physically disabled**
6 ☐ **Multiple persons involved**

7 ☐ **Age was a factor**

Estimated age of person involved

1 ☐ **Male** 2 ☐ **Female**

F1 Equipment Involved in Ignition

☐ **None** ⇨ If equipment was not involved, skip to Section G

Equipment Involved

Brand

Model

Serial #

Year

F2 Equipment Power Source

Equipment Power Source

F3 Equipment Portability

1 ☐ **Portable**
2 ☐ **Stationary**

Portable equipment normally can be moved by one or two persons, is designed to be used in multiple locations, and requires no tools to install.

G Fire Suppression Factors ☐ **None**

Enter up to three codes.

Fire suppression factor (1)

Fire suppression factor (2)

Fire suppression factor (3)

H1 Mobile Property Involved ☐ **None**

1 ☐ **Not involved in ignition, but burned**
2 ☐ **Involved in ignition, but did not burn**
3 ☐ **Involved in ignition and burned**

H2 Mobile Property Type and Make

Mobile property type

Mobile property make

Mobile property model

Year

License Plate Number

State

VIN

Structure fire? Please be sure to complete the Structure Fire form (NFIRS–3).

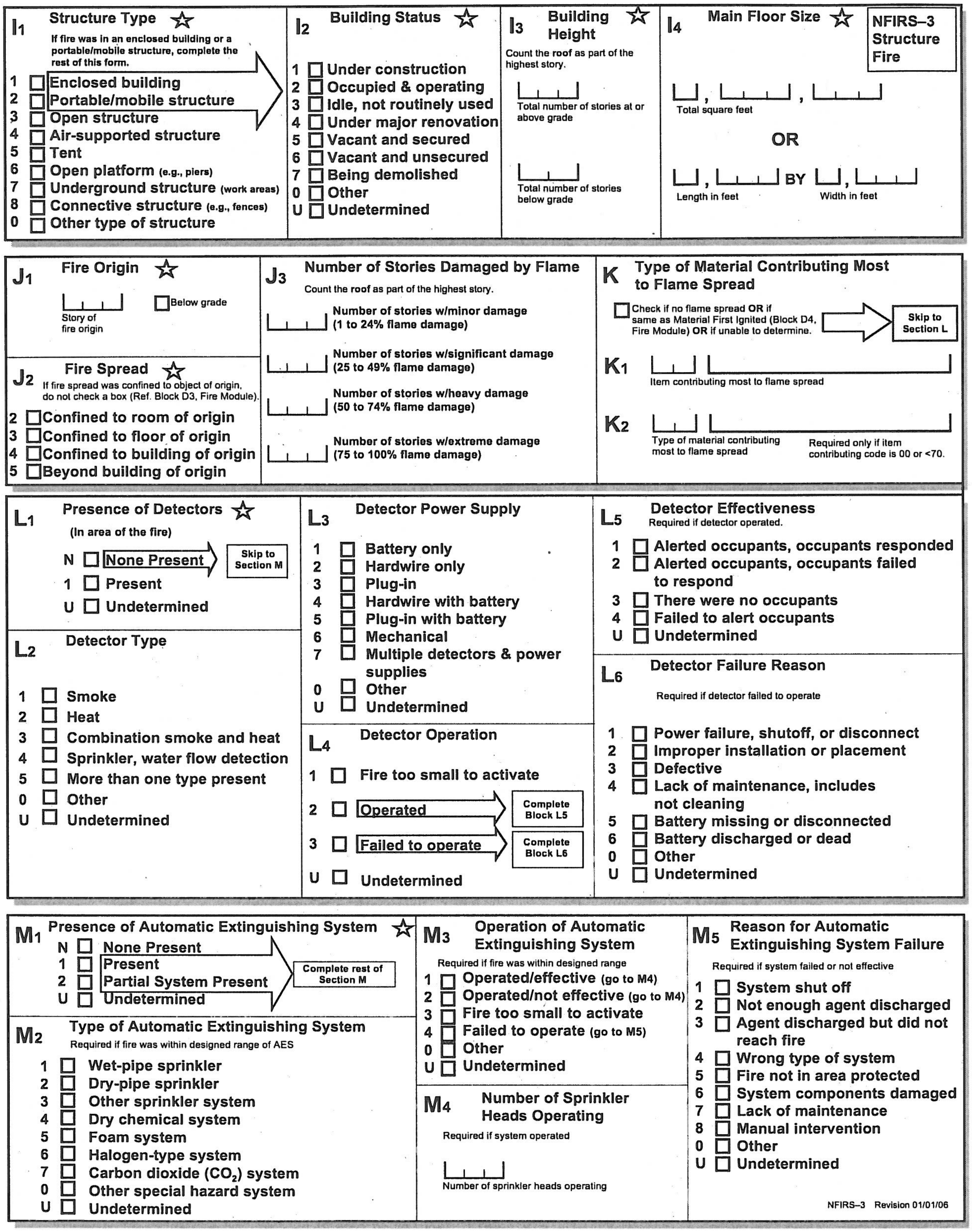
Local Use

☐ **Pre-Fire Plan Available**

Some of the information presented in this report may be based upon reports from other agencies:

☐ **Arson report attached**
☐ **Police report attached**
☐ **Coroner report attached**
☐ **Other reports attached**

NFIRS–2 Revision 01/01/05



NFIRS–3 Structure Fire

I1 Structure Type ☆

If fire was in an enclosed building or a portable/mobile structure, complete the rest of this form.

1 ☐ Enclosed building
2 ☐ Portable/mobile structure
3 ☐ Open structure
4 ☐ Air-supported structure
5 ☐ Tent
6 ☐ Open platform (e.g., piers)
7 ☐ Underground structure (work areas)
8 ☐ Connective structure (e.g., fences)
0 ☐ Other type of structure

I2 Building Status ☆

1 ☐ Under construction
2 ☐ Occupied & operating
3 ☐ Idle, not routinely used
4 ☐ Under major renovation
5 ☐ Vacant and secured
6 ☐ Vacant and unsecured
7 ☐ Being demolished
0 ☐ Other
U ☐ Undetermined

I3 Building Height ☆

Count the roof as part of the highest story.

Total number of stories at or above grade

Total number of stories below grade

I4 Main Floor Size ☆

Total square feet

OR

Length in feet BY Width in feet

J1 Fire Origin ☆

Story of fire origin ☐ Below grade

J2 Fire Spread ☆

If fire spread was confined to object of origin, do not check a box (Ref. Block D3, Fire Module).

2 ☐ Confined to room of origin
3 ☐ Confined to floor of origin
4 ☐ Confined to building of origin
5 ☐ Beyond building of origin

J3 Number of Stories Damaged by Flame

Count the roof as part of the highest story.

Number of stories w/minor damage (1 to 24% flame damage)
Number of stories w/significant damage (25 to 49% flame damage)
Number of stories w/heavy damage (50 to 74% flame damage)
Number of stories w/extreme damage (75 to 100% flame damage)

K Type of Material Contributing Most to Flame Spread

☐ Check if no flame spread OR if same as Material First Ignited (Block D4, Fire Module) OR if unable to determine. → Skip to Section L

K1 Item contributing most to flame spread

K2 Type of material contributing most to flame spread — Required only if item contributing code is 00 or <70.

L1 Presence of Detectors ☆

(In area of the fire)

N ☐ None Present → Skip to Section M
1 ☐ Present
U ☐ Undetermined

L2 Detector Type

1 ☐ Smoke
2 ☐ Heat
3 ☐ Combination smoke and heat
4 ☐ Sprinkler, water flow detection
5 ☐ More than one type present
0 ☐ Other
U ☐ Undetermined

L3 Detector Power Supply

1 ☐ Battery only
2 ☐ Hardwire only
3 ☐ Plug-in
4 ☐ Hardwire with battery
5 ☐ Plug-in with battery
6 ☐ Mechanical
7 ☐ Multiple detectors & power supplies
0 ☐ Other
U ☐ Undetermined

L4 Detector Operation

1 ☐ Fire too small to activate
2 ☐ Operated → Complete Block L5
3 ☐ Failed to operate → Complete Block L6
U ☐ Undetermined

L5 Detector Effectiveness

Required if detector operated.

1 ☐ Alerted occupants, occupants responded
2 ☐ Alerted occupants, occupants failed to respond
3 ☐ There were no occupants
4 ☐ Failed to alert occupants
U ☐ Undetermined

L6 Detector Failure Reason

Required if detector failed to operate

1 ☐ Power failure, shutoff, or disconnect
2 ☐ Improper installation or placement
3 ☐ Defective
4 ☐ Lack of maintenance, includes not cleaning
5 ☐ Battery missing or disconnected
6 ☐ Battery discharged or dead
0 ☐ Other
U ☐ Undetermined

M1 Presence of Automatic Extinguishing System ☆

N ☐ None Present
1 ☐ Present
2 ☐ Partial System Present
U ☐ Undetermined
→ Complete rest of Section M

M2 Type of Automatic Extinguishing System

Required if fire was within designed range of AES

1 ☐ Wet-pipe sprinkler
2 ☐ Dry-pipe sprinkler
3 ☐ Other sprinkler system
4 ☐ Dry chemical system
5 ☐ Foam system
6 ☐ Halogen-type system
7 ☐ Carbon dioxide (CO_2) system
0 ☐ Other special hazard system
U ☐ Undetermined

M3 Operation of Automatic Extinguishing System

Required if fire was within designed range

1 ☐ Operated/effective (go to M4)
2 ☐ Operated/not effective (go to M4)
3 ☐ Fire too small to activate
4 ☐ Failed to operate (go to M5)
0 ☐ Other
U ☐ Undetermined

M4 Number of Sprinkler Heads Operating

Required if system operated

Number of sprinkler heads operating

M5 Reason for Automatic Extinguishing System Failure

Required if system failed or not effective

1 ☐ System shut off
2 ☐ Not enough agent discharged
3 ☐ Agent discharged but did not reach fire
4 ☐ Wrong type of system
5 ☐ Fire not in area protected
6 ☐ System components damaged
7 ☐ Lack of maintenance
8 ☐ Manual intervention
0 ☐ Other
U ☐ Undetermined

NFIRS–3 Revision 01/01/06

ACTIVITY 2.5 (cont'd)

Scenarios and NFIRS Modules

Scenario 2 — Incident #9900211 (Dryer)

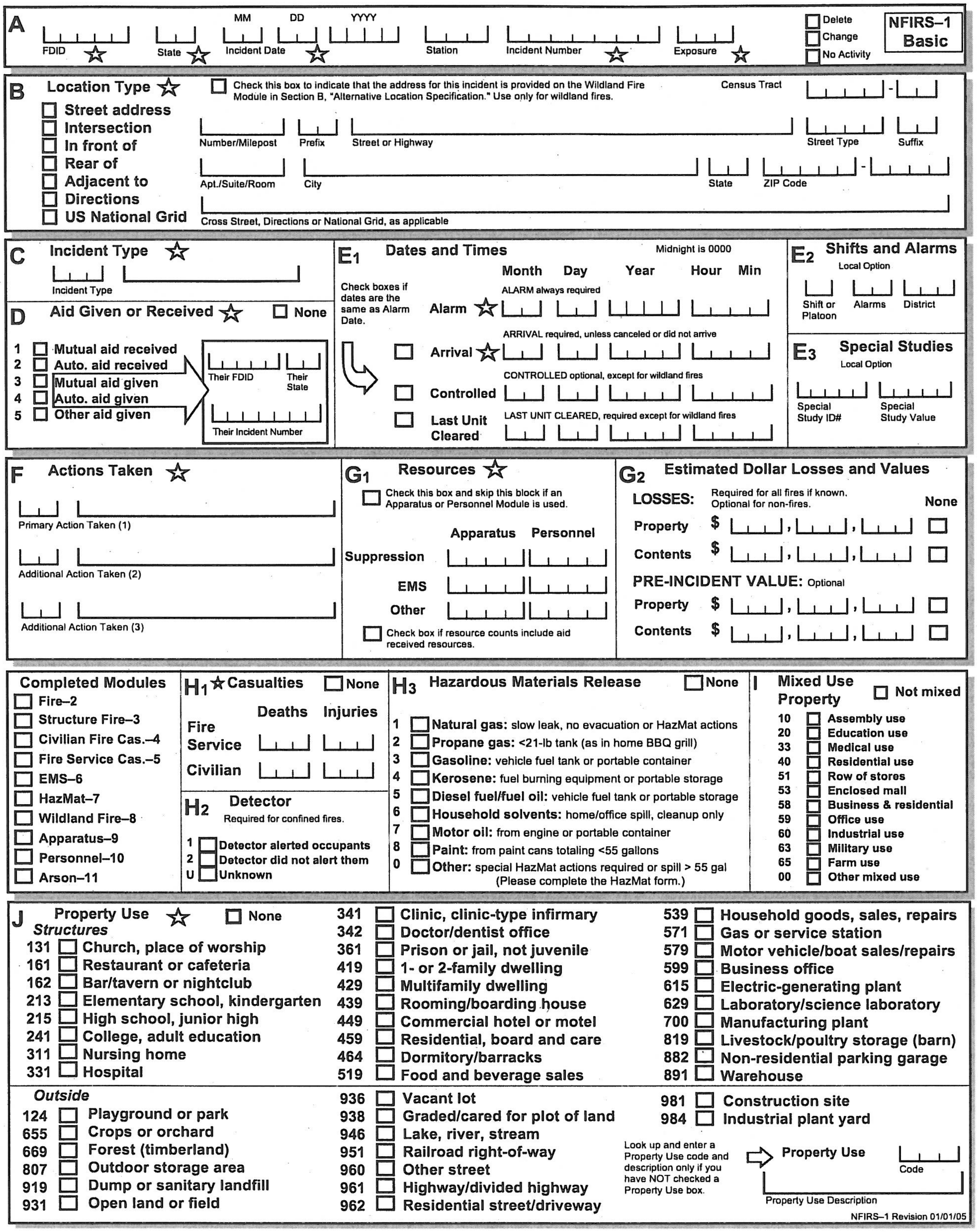
On July 1, 2002, at 1338 hours, Station 1 of the Jonesville, Wis., Fire Department (FDID TR100) responded to a call from a police officer who observed heavy smoke in a laundromat in a multiple-use commercial structure. Engine 2 arrived at 1400 hours and immediately called for a second alarm. The other businesses in the structure were occupied, and the laundromat was open for business but no one was present. A lumber supply company was located next to the structure. The laundromat was 50 percent involved with smoke and heavy visible fire.

The fire service personnel on the scene conducted a primary search for victims, and found none. They evacuated all employees in the other offices, performed horizontal ventilation at the site of the fire, and protected the exposure in the adjacent office and at the lumberyard. The fire was extinguished at 1430 hours. The fire was caused by flammable fibers caught in an improperly ventilated natural gas clothes dryer, a GCM Model 1992G, Serial Number 688599332C. The dryer was manufactured in 1992. There was extensive damage to the Laundromat. Most of the equipment and all of the furniture was destroyed. The incident was reported as #9900211.

This page intentionally left blank.

ACTIVITY 2.5 (cont'd)

Scenarios and NFIRS Modules



A

FDID ☆ | State ☆ | MM DD YYYY Incident Date ☆ | Station | Incident Number ☆ | Exposure ☆

☐ Delete
☐ Change
☐ No Activity

NFIRS–1 Basic

B **Location Type** ☆

☐ Check this box to indicate that the address for this incident is provided on the Wildland Fire Module in Section B, "Alternative Location Specification." Use only for wildland fires.

Census Tract ____ - ____

☐ Street address
☐ Intersection
☐ In front of
☐ Rear of
☐ Adjacent to
☐ Directions
☐ US National Grid

Number/Milepost | Prefix | Street or Highway | Street Type | Suffix

Apt./Suite/Room | City | State | ZIP Code

Cross Street, Directions or National Grid, as applicable

C **Incident Type** ☆

Incident Type

D **Aid Given or Received** ☆ ☐ None

1 ☐ Mutual aid received
2 ☐ Auto. aid received
3 ☐ Mutual aid given
4 ☐ Auto. aid given
5 ☐ Other aid given

Their FDID | Their State

Their Incident Number

E1 **Dates and Times**

Midnight is 0000

Month Day Year Hour Min

Check boxes if dates are the same as Alarm Date.

Alarm ☆ — ALARM always required

☐ Arrival ☆ — ARRIVAL required, unless canceled or did not arrive

☐ Controlled — CONTROLLED optional, except for wildland fires

☐ Last Unit Cleared — LAST UNIT CLEARED, required except for wildland fires

E2 **Shifts and Alarms**

Local Option

Shift or Platoon | Alarms | District

E3 **Special Studies**

Local Option

Special Study ID# | Special Study Value

F **Actions Taken** ☆

Primary Action Taken (1)

Additional Action Taken (2)

Additional Action Taken (3)

G1 **Resources** ☆

☐ Check this box and skip this block if an Apparatus or Personnel Module is used.

	Apparatus	Personnel
Suppression		
EMS		
Other		

☐ Check box if resource counts include aid received resources.

G2 **Estimated Dollar Losses and Values**

LOSSES: Required for all fires if known. Optional for non-fires.

		None
Property	$ ___,___,___	☐
Contents	$ ___,___,___	☐

PRE-INCIDENT VALUE: Optional

Property	$ ___,___,___	☐
Contents	$ ___,___,___	☐

Completed Modules

☐ Fire–2
☐ Structure Fire–3
☐ Civilian Fire Cas.–4
☐ Fire Service Cas.–5
☐ EMS–6
☐ HazMat–7
☐ Wildland Fire–8
☐ Apparatus–9
☐ Personnel–10
☐ Arson–11

H1 ☆ **Casualties** ☐ None

	Deaths	Injuries
Fire Service		
Civilian		

H2 **Detector**

Required for confined fires.

1 ☐ Detector alerted occupants
2 ☐ Detector did not alert them
U ☐ Unknown

H3 **Hazardous Materials Release** ☐ None

1 ☐ **Natural gas:** slow leak, no evacuation or HazMat actions
2 ☐ **Propane gas:** <21-lb tank (as in home BBQ grill)
3 ☐ **Gasoline:** vehicle fuel tank or portable container
4 ☐ **Kerosene:** fuel burning equipment or portable storage
5 ☐ **Diesel fuel/fuel oil:** vehicle fuel tank or portable storage
6 ☐ **Household solvents:** home/office spill, cleanup only
7 ☐ **Motor oil:** from engine or portable container
8 ☐ **Paint:** from paint cans totaling <55 gallons
0 ☐ **Other:** special HazMat actions required or spill > 55 gal (Please complete the HazMat form.)

I **Mixed Use Property** ☐ Not mixed

10 ☐ Assembly use
20 ☐ Education use
33 ☐ Medical use
40 ☐ Residential use
51 ☐ Row of stores
53 ☐ Enclosed mall
58 ☐ Business & residential
59 ☐ Office use
60 ☐ Industrial use
63 ☐ Military use
65 ☐ Farm use
00 ☐ Other mixed use

J **Property Use** ☆ ☐ None

Structures

131 ☐ Church, place of worship
161 ☐ Restaurant or cafeteria
162 ☐ Bar/tavern or nightclub
213 ☐ Elementary school, kindergarten
215 ☐ High school, junior high
241 ☐ College, adult education
311 ☐ Nursing home
331 ☐ Hospital
341 ☐ Clinic, clinic-type infirmary
342 ☐ Doctor/dentist office
361 ☐ Prison or jail, not juvenile
419 ☐ 1- or 2-family dwelling
429 ☐ Multifamily dwelling
439 ☐ Rooming/boarding house
449 ☐ Commercial hotel or motel
459 ☐ Residential, board and care
464 ☐ Dormitory/barracks
519 ☐ Food and beverage sales
539 ☐ Household goods, sales, repairs
571 ☐ Gas or service station
579 ☐ Motor vehicle/boat sales/repairs
599 ☐ Business office
615 ☐ Electric-generating plant
629 ☐ Laboratory/science laboratory
700 ☐ Manufacturing plant
819 ☐ Livestock/poultry storage (barn)
882 ☐ Non-residential parking garage
891 ☐ Warehouse

Outside

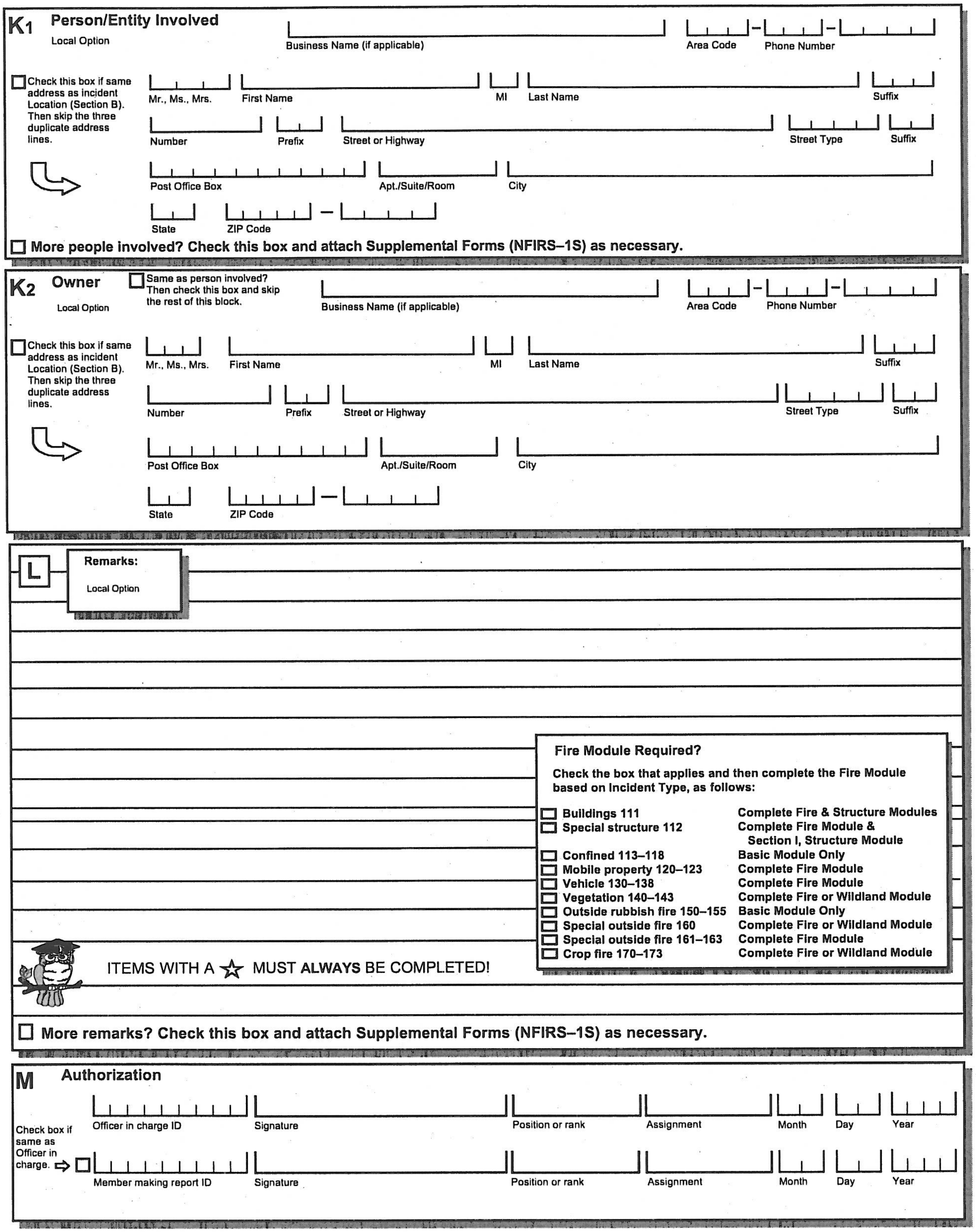
124 ☐ Playground or park
655 ☐ Crops or orchard
669 ☐ Forest (timberland)
807 ☐ Outdoor storage area
919 ☐ Dump or sanitary landfill
931 ☐ Open land or field
936 ☐ Vacant lot
938 ☐ Graded/cared for plot of land
946 ☐ Lake, river, stream
951 ☐ Railroad right-of-way
960 ☐ Other street
961 ☐ Highway/divided highway
962 ☐ Residential street/driveway
981 ☐ Construction site
984 ☐ Industrial plant yard

Look up and enter a Property Use code and description only if you have NOT checked a Property Use box. ⇨ Property Use

Code

Property Use Description

NFIRS–1 Revision 01/01/05



K1 **Person/Entity Involved**
Local Option

Business Name (if applicable)
Area Code – Phone Number

☐ Check this box if same address as incident Location (Section B). Then skip the three duplicate address lines.

Mr., Ms., Mrs. | First Name | MI | Last Name | Suffix
Number | Prefix | Street or Highway | Street Type | Suffix
Post Office Box | Apt./Suite/Room | City
State | ZIP Code –

☐ **More people involved? Check this box and attach Supplemental Forms (NFIRS–1S) as necessary.**

K2 **Owner**
Local Option

☐ Same as person involved? Then check this box and skip the rest of this block.

Business Name (if applicable)
Area Code – Phone Number

☐ Check this box if same address as incident Location (Section B). Then skip the three duplicate address lines.

Mr., Ms., Mrs. | First Name | MI | Last Name | Suffix
Number | Prefix | Street or Highway | Street Type | Suffix
Post Office Box | Apt./Suite/Room | City
State | ZIP Code –

L **Remarks:**
Local Option

Fire Module Required?

Check the box that applies and then complete the Fire Module based on Incident Type, as follows:

☐ Buildings 111	Complete Fire & Structure Modules
☐ Special structure 112	Complete Fire Module & Section I, Structure Module
☐ Confined 113–118	Basic Module Only
☐ Mobile property 120–123	Complete Fire Module
☐ Vehicle 130–138	Complete Fire Module
☐ Vegetation 140–143	Complete Fire or Wildland Module
☐ Outside rubbish fire 150–155	Basic Module Only
☐ Special outside fire 160	Complete Fire or Wildland Module
☐ Special outside fire 161–163	Complete Fire Module
☐ Crop fire 170–173	Complete Fire or Wildland Module

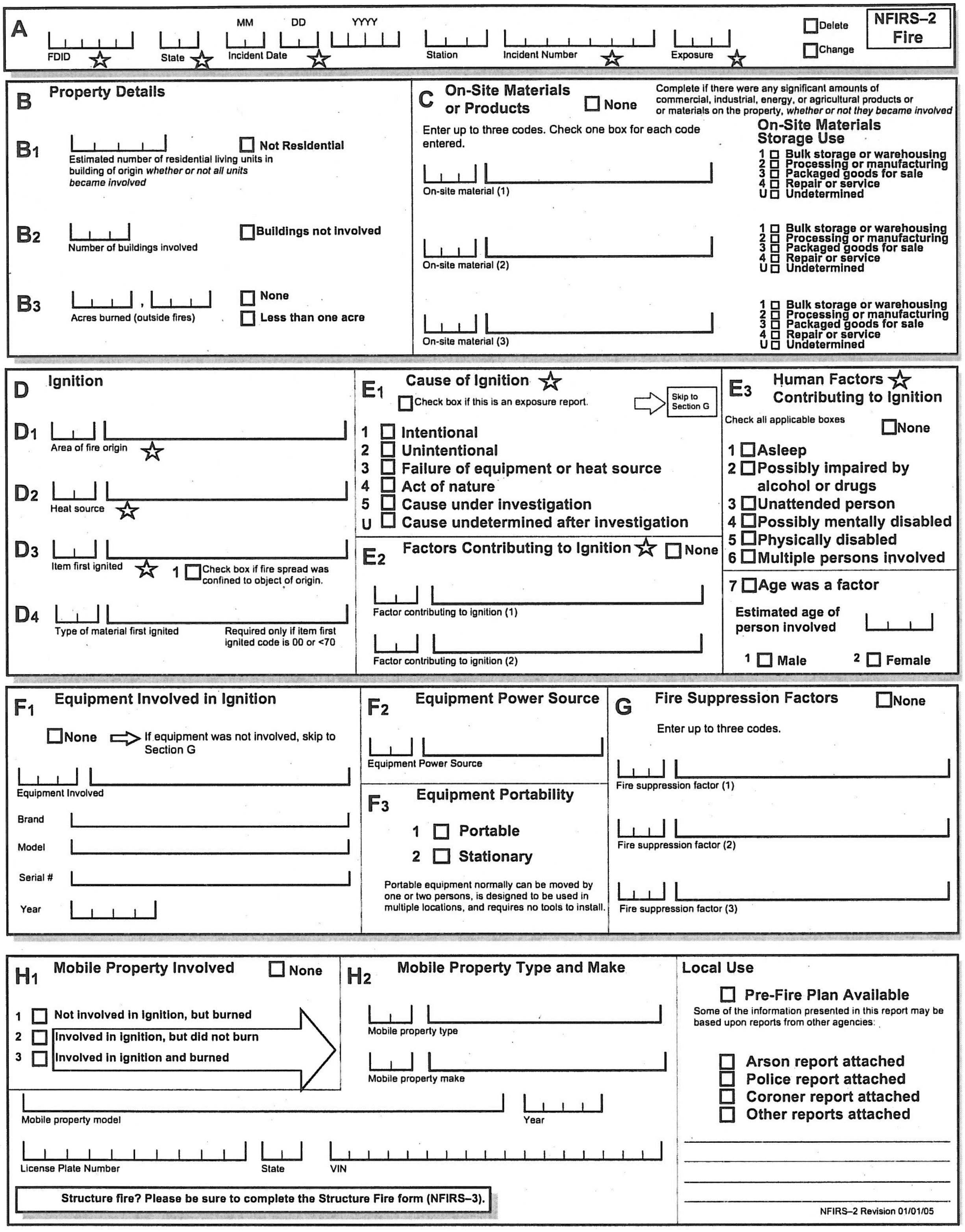
ITEMS WITH A ☆ MUST **ALWAYS** BE COMPLETED!

☐ **More remarks? Check this box and attach Supplemental Forms (NFIRS–1S) as necessary.**

M **Authorization**

Officer in charge ID | Signature | Position or rank | Assignment | Month | Day | Year

Check box if same as Officer in charge. ⇨ ☐ Member making report ID | Signature | Position or rank | Assignment | Month | Day | Year



NFIRS–2 Fire

☐ Delete
☐ Change

A
FDID ☆ | State ☆ | Incident Date (MM DD YYYY) ☆ | Station | Incident Number ☆ | Exposure ☆

B Property Details

B1 Estimated number of residential living units in building of origin *whether or not all units became involved* ☐ Not Residential

B2 Number of buildings involved ☐ Buildings not involved

B3 Acres burned (outside fires) ☐ None ☐ Less than one acre

C On-Site Materials or Products ☐ None

Complete if there were any significant amounts of commercial, industrial, energy, or agricultural products or materials on the property, *whether or not they became involved*

Enter up to three codes. Check one box for each code entered.

On-site material (1)

On-Site Materials Storage Use
1 ☐ Bulk storage or warehousing
2 ☐ Processing or manufacturing
3 ☐ Packaged goods for sale
4 ☐ Repair or service
U ☐ Undetermined

On-site material (2)

1 ☐ Bulk storage or warehousing
2 ☐ Processing or manufacturing
3 ☐ Packaged goods for sale
4 ☐ Repair or service
U ☐ Undetermined

On-site material (3)

1 ☐ Bulk storage or warehousing
2 ☐ Processing or manufacturing
3 ☐ Packaged goods for sale
4 ☐ Repair or service
U ☐ Undetermined

D Ignition

D1 Area of fire origin ☆

D2 Heat source ☆

D3 Item first ignited ☆ 1 ☐ Check box if fire spread was confined to object of origin.

D4 Type of material first ignited — Required only if item first ignited code is 00 or <70

E1 Cause of Ignition ☆

☐ Check box if this is an exposure report. ⇨ Skip to Section G

1 ☐ Intentional
2 ☐ Unintentional
3 ☐ Failure of equipment or heat source
4 ☐ Act of nature
5 ☐ Cause under investigation
U ☐ Cause undetermined after investigation

E2 Factors Contributing to Ignition ☆ ☐ None

Factor contributing to ignition (1)

Factor contributing to ignition (2)

E3 Human Factors Contributing to Ignition ☆

Check all applicable boxes ☐ None

1 ☐ Asleep
2 ☐ Possibly impaired by alcohol or drugs
3 ☐ Unattended person
4 ☐ Possibly mentally disabled
5 ☐ Physically disabled
6 ☐ Multiple persons involved

7 ☐ Age was a factor

Estimated age of person involved

1 ☐ Male 2 ☐ Female

F1 Equipment Involved in Ignition

☐ None ⇨ If equipment was not involved, skip to Section G

Equipment Involved

Brand

Model

Serial #

Year

F2 Equipment Power Source

Equipment Power Source

F3 Equipment Portability

1 ☐ Portable
2 ☐ Stationary

Portable equipment normally can be moved by one or two persons, is designed to be used in multiple locations, and requires no tools to install.

G Fire Suppression Factors ☐ None

Enter up to three codes.

Fire suppression factor (1)

Fire suppression factor (2)

Fire suppression factor (3)

H1 Mobile Property Involved ☐ None

1 ☐ Not involved in ignition, but burned
2 ☐ Involved in ignition, but did not burn
3 ☐ Involved in ignition and burned

H2 Mobile Property Type and Make

Mobile property type

Mobile property make

Mobile property model

Year

License Plate Number

State

VIN

Structure fire? Please be sure to complete the Structure Fire form (NFIRS–3).

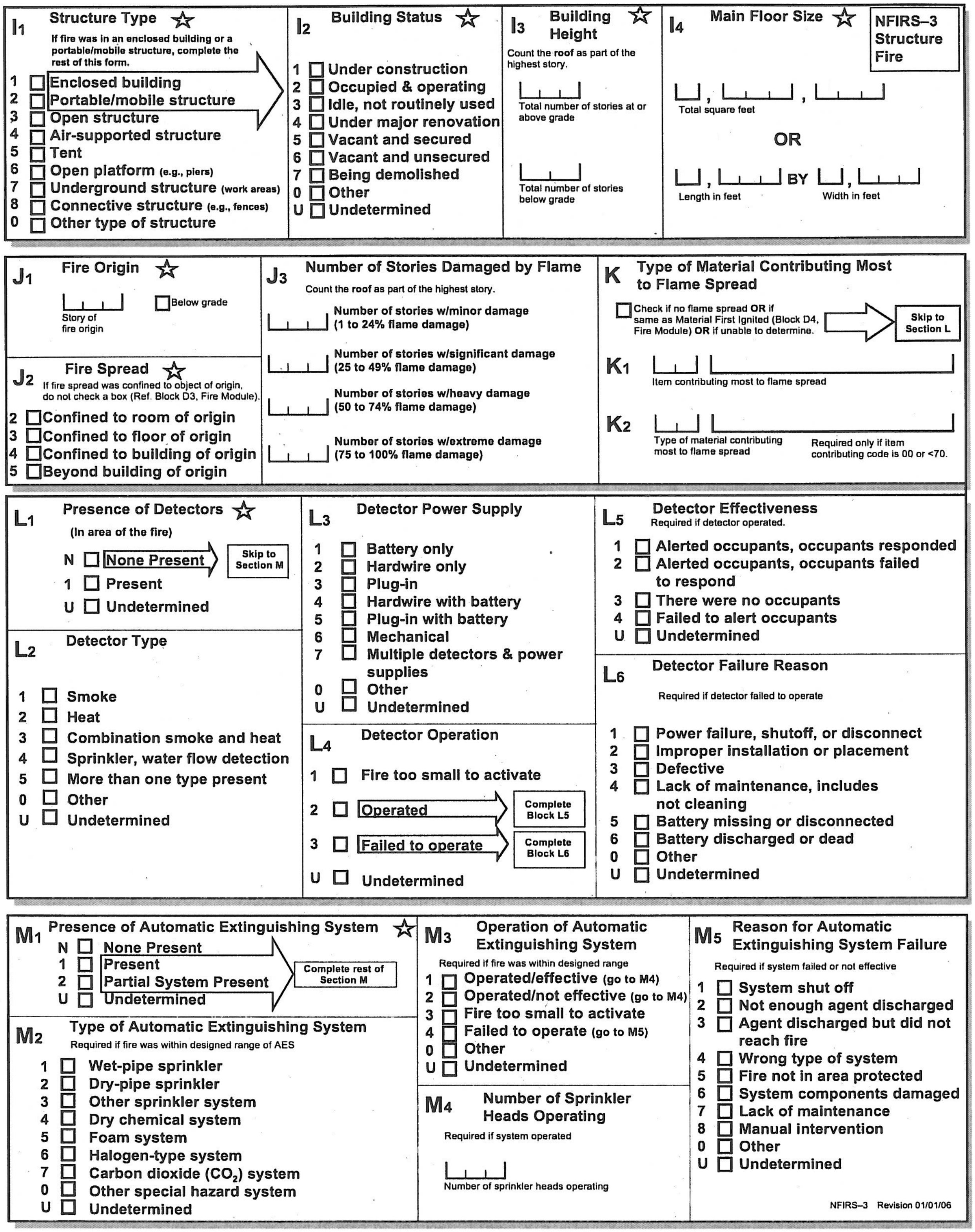
Local Use

☐ Pre-Fire Plan Available

Some of the information presented in this report may be based upon reports from other agencies:

☐ Arson report attached
☐ Police report attached
☐ Coroner report attached
☐ Other reports attached

NFIRS–2 Revision 01/01/05



NFIRS–3 Structure Fire

I1 Structure Type ☆

If fire was in an enclosed building or a portable/mobile structure, complete the rest of this form.

1 ☐ Enclosed building
2 ☐ Portable/mobile structure
3 ☐ Open structure
4 ☐ Air-supported structure
5 ☐ Tent
6 ☐ Open platform (e.g., piers)
7 ☐ Underground structure (work areas)
8 ☐ Connective structure (e.g., fences)
0 ☐ Other type of structure

I2 Building Status ☆

1 ☐ Under construction
2 ☐ Occupied & operating
3 ☐ Idle, not routinely used
4 ☐ Under major renovation
5 ☐ Vacant and secured
6 ☐ Vacant and unsecured
7 ☐ Being demolished
0 ☐ Other
U ☐ Undetermined

I3 Building Height ☆

Count the roof as part of the highest story.

Total number of stories at or above grade

Total number of stories below grade

I4 Main Floor Size ☆

Total square feet

OR

Length in feet BY Width in feet

J1 Fire Origin ☆

Story of fire origin ☐ Below grade

J2 Fire Spread ☆

If fire spread was confined to object of origin, do not check a box (Ref. Block D3, Fire Module).

2 ☐ Confined to room of origin
3 ☐ Confined to floor of origin
4 ☐ Confined to building of origin
5 ☐ Beyond building of origin

J3 Number of Stories Damaged by Flame

Count the roof as part of the highest story.

Number of stories w/minor damage (1 to 24% flame damage)

Number of stories w/significant damage (25 to 49% flame damage)

Number of stories w/heavy damage (50 to 74% flame damage)

Number of stories w/extreme damage (75 to 100% flame damage)

K Type of Material Contributing Most to Flame Spread

☐ Check if no flame spread OR if same as Material First Ignited (Block D4, Fire Module) OR if unable to determine. → Skip to Section L

K1 Item contributing most to flame spread

K2 Type of material contributing most to flame spread — Required only if item contributing code is 00 or <70.

L1 Presence of Detectors ☆

(In area of the fire)

N ☐ None Present → Skip to Section M
1 ☐ Present
U ☐ Undetermined

L2 Detector Type

1 ☐ Smoke
2 ☐ Heat
3 ☐ Combination smoke and heat
4 ☐ Sprinkler, water flow detection
5 ☐ More than one type present
0 ☐ Other
U ☐ Undetermined

L3 Detector Power Supply

1 ☐ Battery only
2 ☐ Hardwire only
3 ☐ Plug-in
4 ☐ Hardwire with battery
5 ☐ Plug-in with battery
6 ☐ Mechanical
7 ☐ Multiple detectors & power supplies
0 ☐ Other
U ☐ Undetermined

L4 Detector Operation

1 ☐ Fire too small to activate
2 ☐ Operated → Complete Block L5
3 ☐ Failed to operate → Complete Block L6
U ☐ Undetermined

L5 Detector Effectiveness

Required if detector operated.

1 ☐ Alerted occupants, occupants responded
2 ☐ Alerted occupants, occupants failed to respond
3 ☐ There were no occupants
4 ☐ Failed to alert occupants
U ☐ Undetermined

L6 Detector Failure Reason

Required if detector failed to operate

1 ☐ Power failure, shutoff, or disconnect
2 ☐ Improper installation or placement
3 ☐ Defective
4 ☐ Lack of maintenance, includes not cleaning
5 ☐ Battery missing or disconnected
6 ☐ Battery discharged or dead
0 ☐ Other
U ☐ Undetermined

M1 Presence of Automatic Extinguishing System ☆

N ☐ None Present
1 ☐ Present
2 ☐ Partial System Present
U ☐ Undetermined
→ Complete rest of Section M

M2 Type of Automatic Extinguishing System

Required if fire was within designed range of AES

1 ☐ Wet-pipe sprinkler
2 ☐ Dry-pipe sprinkler
3 ☐ Other sprinkler system
4 ☐ Dry chemical system
5 ☐ Foam system
6 ☐ Halogen-type system
7 ☐ Carbon dioxide (CO_2) system
0 ☐ Other special hazard system
U ☐ Undetermined

M3 Operation of Automatic Extinguishing System

Required if fire was within designed range

1 ☐ Operated/effective (go to M4)
2 ☐ Operated/not effective (go to M4)
3 ☐ Fire too small to activate
4 ☐ Failed to operate (go to M5)
0 ☐ Other
U ☐ Undetermined

M4 Number of Sprinkler Heads Operating

Required if system operated

Number of sprinkler heads operating

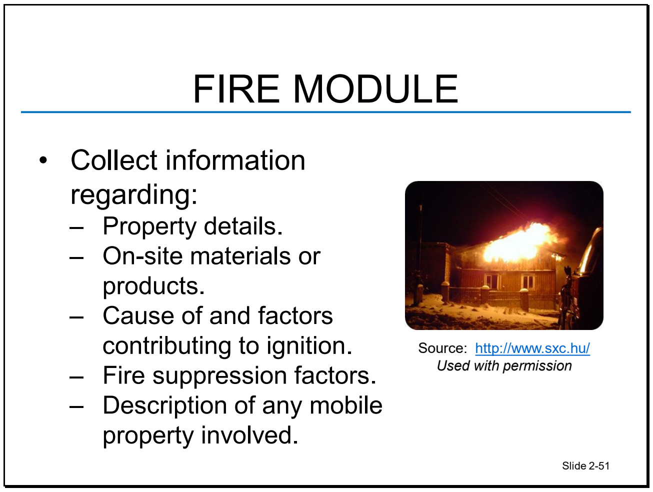
M5 Reason for Automatic Extinguishing System Failure

Required if system failed or not effective

1 ☐ System shut off
2 ☐ Not enough agent discharged
3 ☐ Agent discharged but did not reach fire
4 ☐ Wrong type of system
5 ☐ Fire not in area protected
6 ☐ System components damaged
7 ☐ Lack of maintenance
8 ☐ Manual intervention
0 ☐ Other
U ☐ Undetermined

NFIRS–3 Revision 01/01/06

X. USING NATIONAL FIRE INCIDENT REPORTING SYSTEM FOR DATA COLLECTION (cont'd)



FIRE MODULE

- Collect information regarding:
  - Property details.
  - On-site materials or products.
  - Cause of and factors contributing to ignition.
  - Fire suppression factors.
  - Description of any mobile property involved.

Source: http://www.sxc.hu/
Used with permission

D. Fire Module.

1. The Fire Module allows you to collect information regarding:

   a. Property details.

   b. On-site materials or products.

   c. Cause of and factors contributing to ignition.

   d. Fire suppression factors.

   e. Description of any mobile property involved.

2. Each piece of this information can be helpful when developing the appropriate elements in the SOC plan, including:

   a. Property details.

   b. On-site materials.

   c. Ignition, including:

      - Area.

      - Source of ignition.

      - Material ignited.

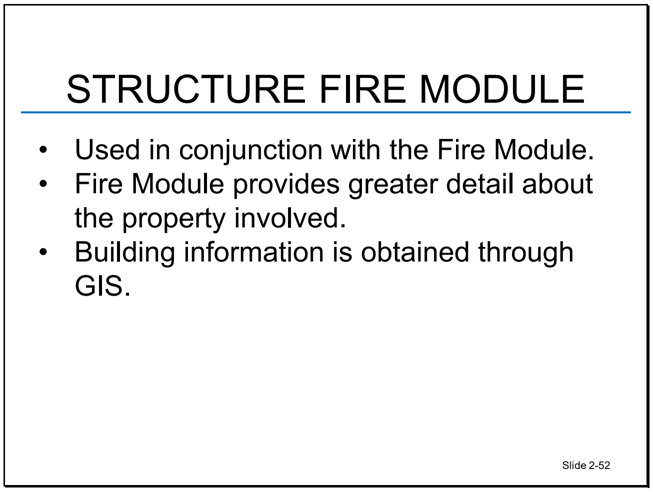
- Contributing factors.
- Human issues.
- Equipment involved.

d. Human factors involved.

e. Mobile property description.

f. Fire origin and spread description.

g. Fire suppression factors.



STRUCTURE FIRE MODULE

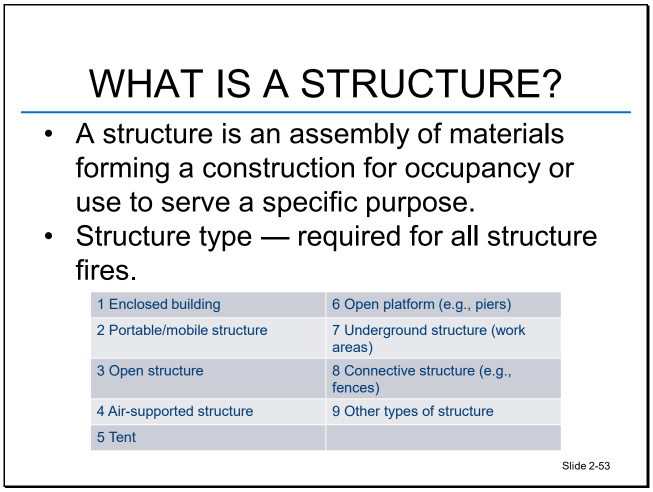
- Used in conjunction with the Fire Module.
- Fire Module provides greater detail about the property involved.
- Building information is obtained through GIS.

E. Structure Fire Module.

1. The Structure Fire Module is used in conjunction with the Fire Module for building fires which extend beyond a noncombustible container (Incident Types 111 and 120s).

2. The Fire Module provides greater detail about the property involved, whereas the Structure Fire Module furnishes information regarding the buildings involved in the fire, how the fire started, and detection and suppression equipment present.

3. Building information is obtained through GIS.



WHAT IS A STRUCTURE?

- A structure is an assembly of materials forming a construction for occupancy or use to serve a specific purpose.
- Structure type — required for all structure fires.

1 Enclosed building	6 Open platform (e.g., piers)
2 Portable/mobile structure	7 Underground structure (work areas)
3 Open structure	8 Connective structure (e.g., fences)
4 Air-supported structure	9 Other types of structure
5 Tent	

F. Structure fires.

1. Topics discussed in this section include:

    a. Definition of a structure.

    b. Structure type.

    c. Capturing square footage.

    d. Area of fire origin.

    e. Fire spread.

    f. Floors involved.

    g. Detection involved.

    h. Automatic extinguishing system (AES) involved.

2. A structure is an assembly of materials forming a construction for occupancy or used to serve a specific purpose.

    Structure types for fires can include:

    a. Enclosed building.

    b. Portable or mobile structure.

    c. Open structure.

    d. Air-supported structure.

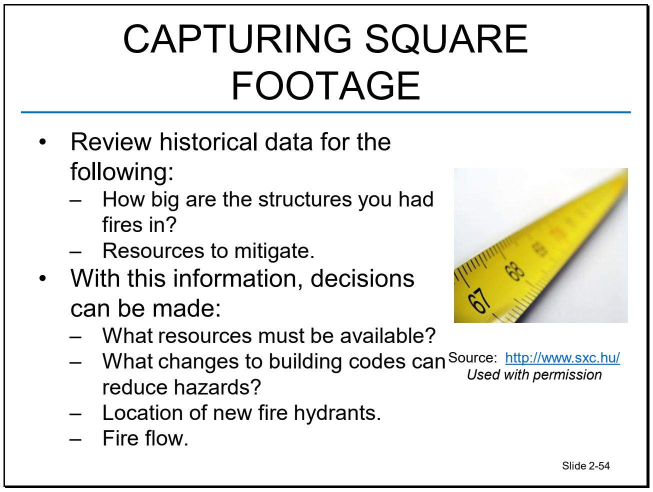
e. Tent.

f. Open platform (e.g., piers).

g. Underground structure (work areas).

h. Connective structure (e.g., fences).

i. Other types of structure.



CAPTURING SQUARE FOOTAGE

- Review historical data for the following:
  - How big are the structures you had fires in?
  - Resources to mitigate.
- With this information, decisions can be made:
  - What resources must be available?
  - What changes to building codes can reduce hazards?
  - Location of new fire hydrants.
  - Fire flow.

Source: http://www.sxc.hu/
Used with permission

G. Capturing square footage.

1. By reviewing historical data, organizations can plan for the future.

   Pertinent historical data that can help them plan for the future include:

   a. How big are the structures you had fires in?

   b. Resources to mitigate.

      - Fire flow?

      - Fuel load?

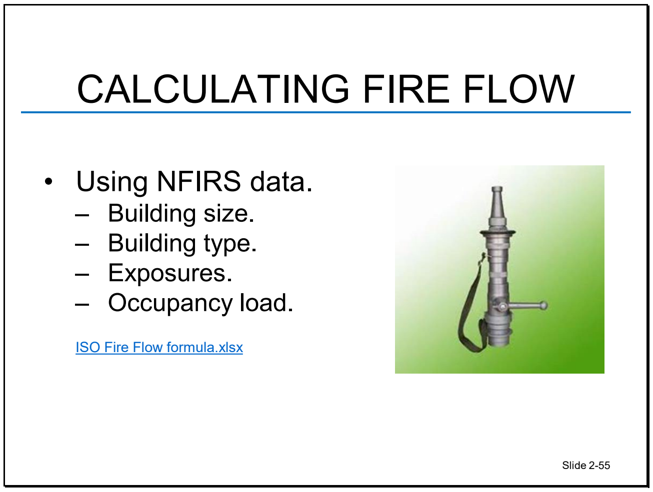
2. With this information, decisions can be made, such as:

   a. What resources must be available?

   b. What changes to building codes can reduce hazards (including location and number of sprinklers)?

   c. Location of new fire hydrants.

d. Fire flow.



CALCULATING FIRE FLOW

- Using NFIRS data.
  - Building size.
  - Building type.
  - Exposures.
  - Occupancy load.

ISO Fire Flow formula.xlsx

H. Calculating fire flow.

1. Data from NFIRS reports allow the planner to evaluate fire flows of fire streams. NFIRS data includes:

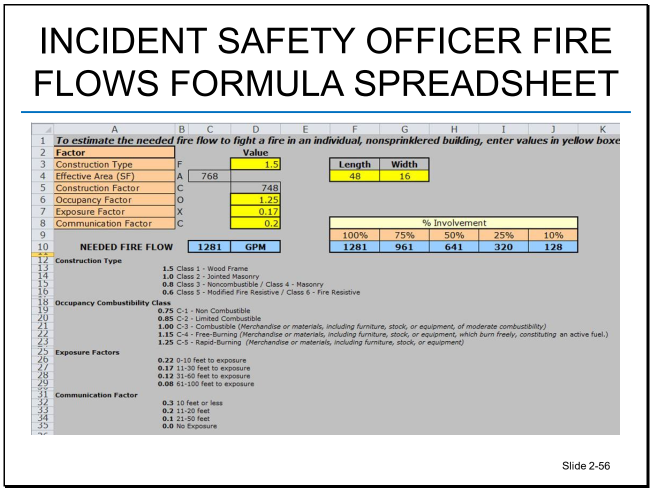
   a. Building size.

   b. Building type.

   c. Exposures.

   d. Occupancy load.

2. Based on data, it can be determined whether enough hose lines are being used.



INCIDENT SAFETY OFFICER FIRE FLOWS FORMULA SPREADSHEET

To estimate the needed fire flow to fight a fire in an individual, nonsprinklered building, enter values in yellow boxe

Factor			Value
Construction Type	F		1.5
Effective Area (SF)	A	768	
Construction Factor	C		748
Occupancy Factor	O		1.25
Exposure Factor	X		0.17
Communication Factor	C		0.2
NEEDED FIRE FLOW		**1281**	**GPM**

Length	Width
48	16

% Involvement				
100%	75%	50%	25%	10%
1281	**961**	**641**	**320**	**128**

Construction Type

1.5 Class 1 - Wood Frame
1.0 Class 2 - Jointed Masonry
0.8 Class 3 - Noncombustible / Class 4 - Masonry
0.6 Class 5 - Modified Fire Resistive / Class 6 - Fire Resistive

Occupancy Combustibility Class

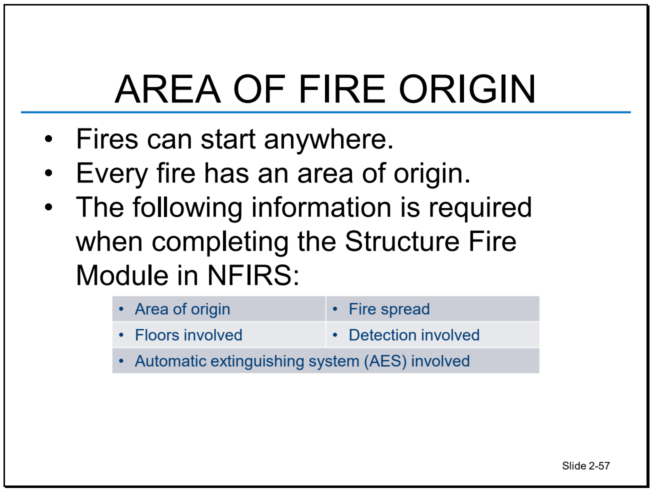
0.75 C-1 - Non Combustible
0.85 C-2 - Limited Combustible
1.00 C-3 - Combustible (*Merchandise or materials, including furniture, stock, or equipment, of moderate combustibility)*
1.15 C-4 - Free-Burning *(Merchandise or materials, including furniture, stock, or equipment, which burn freely, constituting* an active fuel.)
1.25 C-5 - Rapid-Burning *(Merchandise or materials, including furniture, stock, or equipment)*

Exposure Factors

0.22 0-10 feet to exposure
0.17 11-30 feet to exposure
0.12 31-60 feet to exposure
0.08 61-100 feet to exposure

Communication Factor

0.3 10 feet or less
0.2 11-20 feet
0.1 21-50 feet
0.0 No Exposure



AREA OF FIRE ORIGIN

- Fires can start anywhere.
- Every fire has an area of origin.
- The following information is required when completing the Structure Fire Module in NFIRS:

• Area of origin	• Fire spread
• Floors involved	• Detection involved
• Automatic extinguishing system (AES) involved	

Slide 2-57

I. Area of fire origin.

1. Fires can start anywhere; every fire has an area of origin. This information is required to be entered into NFIRS.

2. According to the USFA, most fires are started:

    a. In a cooking area.

    b. Between 1500 and 1800 hours.

    c. By children under 14 years of age.

3. How do you plan to support your data?

    a. Planning.

    b. Mitigation.

    c. Response.

    d. Recovery.

4. The following information is required when completing the Structure Fire Module in NFIRS:

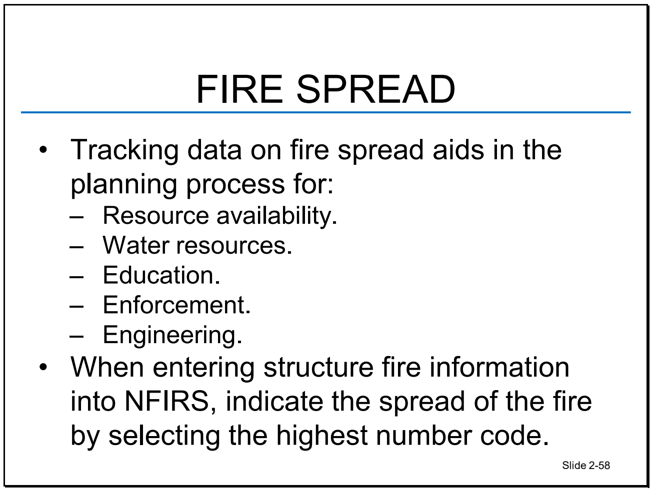
    a. Area of fire origin.

    b. Fire spread.

    c. Floors involved.

d. Detection involved.

e. AES involved.



FIRE SPREAD

- Tracking data on fire spread aids in the planning process for:
  - Resource availability.
  - Water resources.
  - Education.
  - Enforcement.
  - Engineering.
- When entering structure fire information into NFIRS, indicate the spread of the fire by selecting the highest number code.

J. Fire spread.

1. Tracking data on fire spread aids in the planning process. Data can be tracked on fire spread, floors involved, detection involved, and AES or sprinklers.

2. When entering structure fire information into NFIRS, indicate the spread of the fire by selecting the highest number code that applies:

   a. Confined to object of origin.

   b. Confined to room of origin.

   c. Confined to floor of origin.

   d. Confined to building of origin.

   e. Beyond building of origin.

3. Data gathered on the floors involved in a fire, such as:

Data	Plan
Type of building construction • Single-family • Two-story • Balloon construction • Lightweight	• Resource availability • Water resources • Fire flows
Fire spread	

4. Data gathered on smoke detectors can influence planning:

Data	Plan
• Presence of detectors	• Education
• Detector type	• Enforcement
• Detector operation	• Engineering
• Detector performance	

5. AES or sprinkler system-involved data are also useful in planning:

Data	Plan
• Type of AES	• Education
• Operation	• Enforcement
• Performance	• Engineering

ACTIVITY 2.6

NFIRS — Structure Fire Module

Purpose

To allow you to complete NFIRS — Structure Fire Module form based on call information provided.

Directions

1. Work in a group.
2. Read the call information and complete the NFIRS — Structure Fire Module.
3. Identify what is lacking in the NFRIS — Structure Fire Module form.
4. Be prepared to debrief as a class.

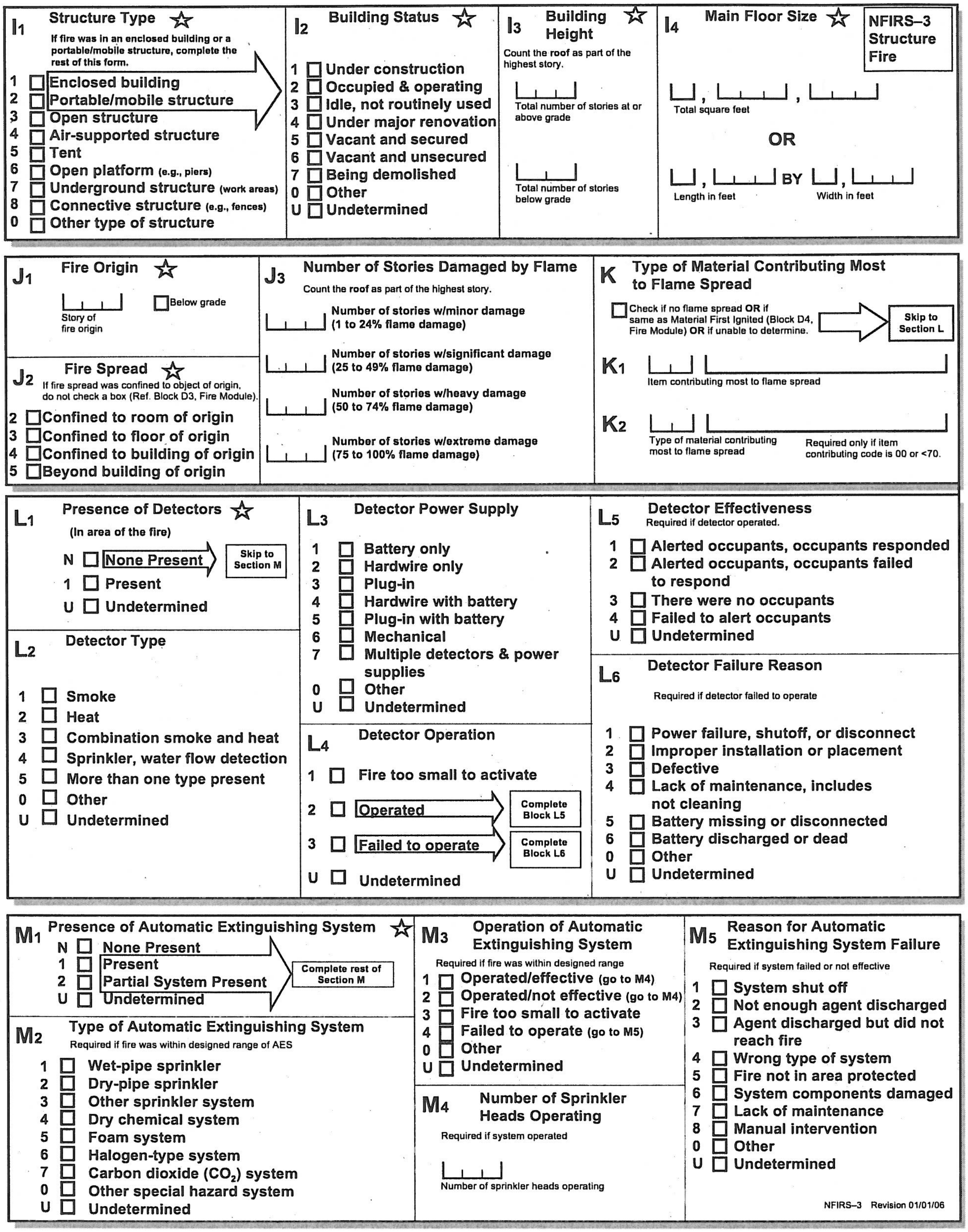
Call Information

A smoke detector in the first-floor hallway alerted the residents of a single-family dwelling of a possible problem. They quickly exited out the front door and reported seeing smoke coming from the basement. Children playing with matches had started a fire in a small stack of newspapers that were in the 30 by 50 foot basement of a ranch-style home. Luckily they were uninjured. There was fire damage in the basement and smoke damage on the first floor. The detector was hardwired with a battery backup. There was a residential wet-pipe sprinkler system installed throughout the home. One sprinkler head activated and extinguished the fire.

This page intentionally left blank.

ACTIVITY 2.6 (cont'd)

NFIRS — Structure Fire Module



I1 Structure Type ☆
If fire was in an enclosed building or a portable/mobile structure, complete the rest of this form.
1 ☐ Enclosed building
2 ☐ Portable/mobile structure
3 ☐ Open structure
4 ☐ Air-supported structure
5 ☐ Tent
6 ☐ Open platform (e.g., piers)
7 ☐ Underground structure (work areas)
8 ☐ Connective structure (e.g., fences)
0 ☐ Other type of structure

I2 Building Status ☆
1 ☐ Under construction
2 ☐ Occupied & operating
3 ☐ Idle, not routinely used
4 ☐ Under major renovation
5 ☐ Vacant and secured
6 ☐ Vacant and unsecured
7 ☐ Being demolished
0 ☐ Other
U ☐ Undetermined

I3 Building Height ☆
Count the roof as part of the highest story.
____ Total number of stories at or above grade
____ Total number of stories below grade

I4 Main Floor Size ☆
NFIRS–3 Structure Fire
___, ____, ____ Total square feet
OR
___, ____ Length in feet BY ___, ____ Width in feet

J1 Fire Origin ☆
____ Story of fire origin ☐ Below grade

J2 Fire Spread ☆
If fire spread was confined to object of origin, do not check a box (Ref. Block D3, Fire Module).
2 ☐ Confined to room of origin
3 ☐ Confined to floor of origin
4 ☐ Confined to building of origin
5 ☐ Beyond building of origin

J3 Number of Stories Damaged by Flame
Count the roof as part of the highest story.
____ Number of stories w/minor damage (1 to 24% flame damage)
____ Number of stories w/significant damage (25 to 49% flame damage)
____ Number of stories w/heavy damage (50 to 74% flame damage)
____ Number of stories w/extreme damage (75 to 100% flame damage)

K Type of Material Contributing Most to Flame Spread
☐ Check if no flame spread OR if same as Material First Ignited (Block D4, Fire Module) OR if unable to determine. → Skip to Section L
K1 ____ ________ Item contributing most to flame spread
K2 ____ ________ Type of material contributing most to flame spread
Required only if item contributing code is 00 or <70.

L1 Presence of Detectors ☆
(In area of the fire)
N ☐ None Present → Skip to Section M
1 ☐ Present
U ☐ Undetermined

L2 Detector Type
1 ☐ Smoke
2 ☐ Heat
3 ☐ Combination smoke and heat
4 ☐ Sprinkler, water flow detection
5 ☐ More than one type present
0 ☐ Other
U ☐ Undetermined

L3 Detector Power Supply
1 ☐ Battery only
2 ☐ Hardwire only
3 ☐ Plug-in
4 ☐ Hardwire with battery
5 ☐ Plug-in with battery
6 ☐ Mechanical
7 ☐ Multiple detectors & power supplies
0 ☐ Other
U ☐ Undetermined

L4 Detector Operation
1 ☐ Fire too small to activate
2 ☐ Operated → Complete Block L5
3 ☐ Failed to operate → Complete Block L6
U ☐ Undetermined

L5 Detector Effectiveness
Required if detector operated.
1 ☐ Alerted occupants, occupants responded
2 ☐ Alerted occupants, occupants failed to respond
3 ☐ There were no occupants
4 ☐ Failed to alert occupants
U ☐ Undetermined

L6 Detector Failure Reason
Required if detector failed to operate
1 ☐ Power failure, shutoff, or disconnect
2 ☐ Improper installation or placement
3 ☐ Defective
4 ☐ Lack of maintenance, includes not cleaning
5 ☐ Battery missing or disconnected
6 ☐ Battery discharged or dead
0 ☐ Other
U ☐ Undetermined

M1 Presence of Automatic Extinguishing System ☆
N ☐ None Present
1 ☐ Present
2 ☐ Partial System Present
U ☐ Undetermined
→ Complete rest of Section M

M2 Type of Automatic Extinguishing System
Required if fire was within designed range of AES
1 ☐ Wet-pipe sprinkler
2 ☐ Dry-pipe sprinkler
3 ☐ Other sprinkler system
4 ☐ Dry chemical system
5 ☐ Foam system
6 ☐ Halogen-type system
7 ☐ Carbon dioxide (CO_2) system
0 ☐ Other special hazard system
U ☐ Undetermined

M3 Operation of Automatic Extinguishing System
Required if fire was within designed range
1 ☐ Operated/effective (go to M4)
2 ☐ Operated/not effective (go to M4)
3 ☐ Fire too small to activate
4 ☐ Failed to operate (go to M5)
0 ☐ Other
U ☐ Undetermined

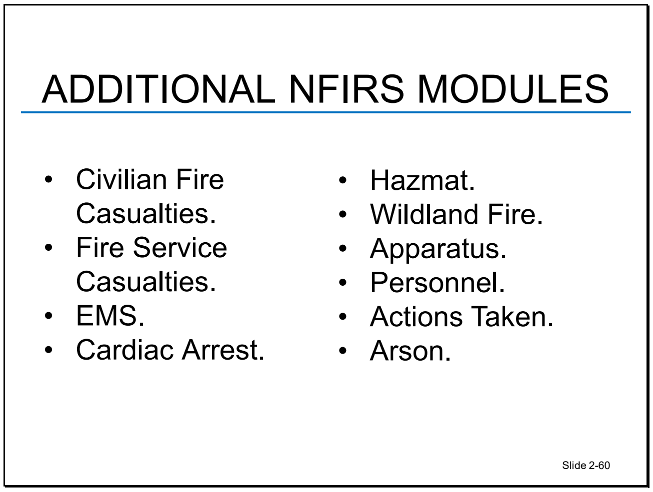
M4 Number of Sprinkler Heads Operating
Required if system operated
____ Number of sprinkler heads operating

M5 Reason for Automatic Extinguishing System Failure
Required if system failed or not effective
1 ☐ System shut off
2 ☐ Not enough agent discharged
3 ☐ Agent discharged but did not reach fire
4 ☐ Wrong type of system
5 ☐ Fire not in area protected
6 ☐ System components damaged
7 ☐ Lack of maintenance
8 ☐ Manual intervention
0 ☐ Other
U ☐ Undetermined

NFIRS–3 Revision 01/01/06

This page intentionally left blank.

XI. ADDITIONAL NATIONAL FIRE INCIDENT REPORTING SYSTEM MODULES



ADDITIONAL NFIRS MODULES

- Civilian Fire Casualties.
- Fire Service Casualties.
- EMS.
- Cardiac Arrest.
- Hazmat.
- Wildland Fire.
- Apparatus.
- Personnel.
- Actions Taken.
- Arson.

A. Introduction.

1. Other NFIRS modules that may require completion include:

    a. Civilian Fire Casualties.

    b. Fire Service Casualties.

    c. EMS.

    d. Cardiac Arrest.

    e. Hazardous Materials.

    f. Wildland Fire.

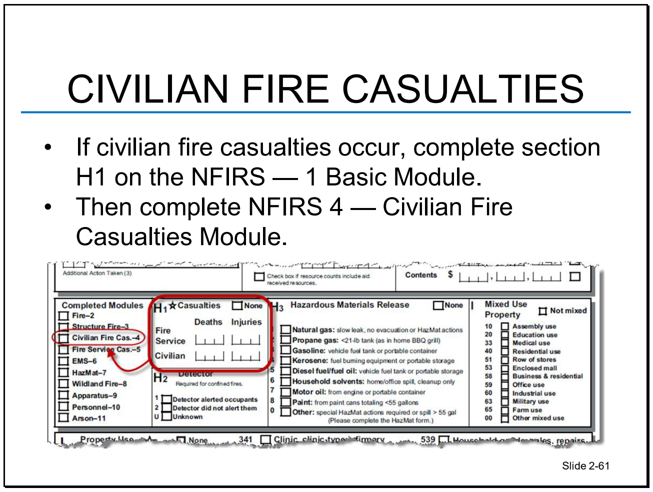
    g. Apparatus.

    h. Personnel.

    i. Actions Taken.

    j. Arson.

2. All data gathered in the incident reporting system can be used for planning purposes.



CIVILIAN FIRE CASUALTIES

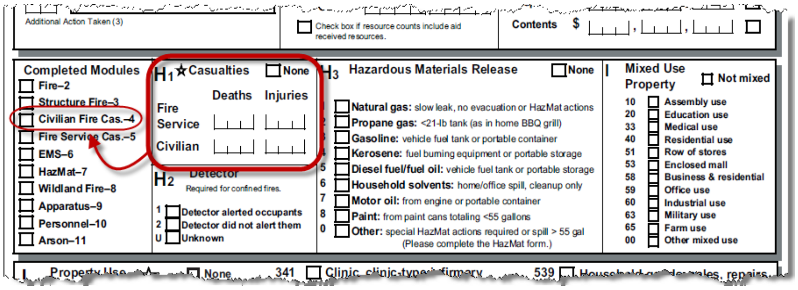
- If civilian fire casualties occur, complete section H1 on the NFIRS — 1 Basic Module.
- Then complete NFIRS 4 — Civilian Fire Casualties Module.

Slide 2-61

B. Civilian fire casualties.

1. The Civilian Fire Casualty Module captures data regarding any civilian (nonfire service) casualty associated with fire-related incidents. If civilian injuries or deaths are recorded in H1 of the Basic Module, then you are required to complete the Civilian Fire Casualty Module.

2. A civilian fire casualty can be a private citizen, emergency medical responder (nonfire department), or police officer who dies or is physically injured as the result of a fire-related incident. This description is not meant to exclude other people who fall into this category. For instance, any number of public and emergency services personnel may be on the scene of an emergency, such as public works personnel, state highway personnel, and other federal, state, or local employees/officials.



Additional Action Taken (3)

☐ Check box if resource counts include aid received resources.

Contents $ ☐

Completed Modules
☐ **Fire–2**
☐ **Structure Fire–3**
☐ **Civilian Fire Cas.–4**
☐ **Fire Service Cas.–5**
☐ **EMS–6**
☐ **HazMat–7**
☐ **Wildland Fire–8**
☐ **Apparatus–9**
☐ **Personnel–10**
☐ **Arson–11**

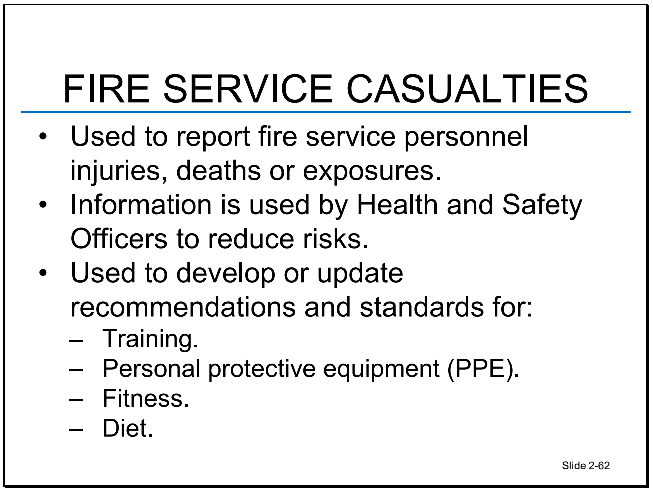
H1 ★Casualties ☐ **None**

	Deaths	Injuries
Fire Service		
Civilian		

H2 Detector
Required for confined fires.
1 ☐ **Detector alerted occupants**
2 ☐ **Detector did not alert them**
U ☐ **Unknown**

H3 Hazardous Materials Release ☐ **None**
☐ **Natural gas:** slow leak, no evacuation or HazMat actions
☐ **Propane gas:** <21-lb tank (as in home BBQ grill)
☐ **Gasoline:** vehicle fuel tank or portable container
☐ **Kerosene:** fuel burning equipment or portable storage
5 ☐ **Diesel fuel/fuel oil:** vehicle fuel tank or portable storage
6 ☐ **Household solvents:** home/office spill, cleanup only
7 ☐ **Motor oil:** from engine or portable container
8 ☐ **Paint:** from paint cans totaling <55 gallons
0 ☐ **Other:** special HazMat actions required or spill > 55 gal
(Please complete the HazMat form.)

I Mixed Use Property ☐ **Not mixed**
10 ☐ **Assembly use**
20 ☐ **Education use**
33 ☐ **Medical use**
40 ☐ **Residential use**
51 ☐ **Row of stores**
53 ☐ **Enclosed mall**
58 ☐ **Business & residential**
59 ☐ **Office use**
60 ☐ **Industrial use**
63 ☐ **Military use**
65 ☐ **Farm use**
00 ☐ **Other mixed use**



FIRE SERVICE CASUALTIES

- Used to report fire service personnel injuries, deaths or exposures.
- Information is used by Health and Safety Officers to reduce risks.
- Used to develop or update recommendations and standards for:
  - Training.
  - Personal protective equipment (PPE).
  - Fitness.
  - Diet.

C. Fire service casualties.

1. The Fire Service Casualty Module is used to report fire service personnel injuries, deaths or exposures while on duty. This casualty information is used by Health and Safety Officers to reduce the risks associated with all types of work-related casualties. The Fire Service Casualty Module is also used to collect information about protective equipment that failed and contributed to the injury.

2. Researchers, educators, equipment makers, design engineers and governmental regulatory agencies may use the specific information provided to make various determinations, such as which specific pieces of equipment are involved in casualties. Complete information must be collected for each individual casualty in order to provide the data needed to make determinations related to improving job safety.

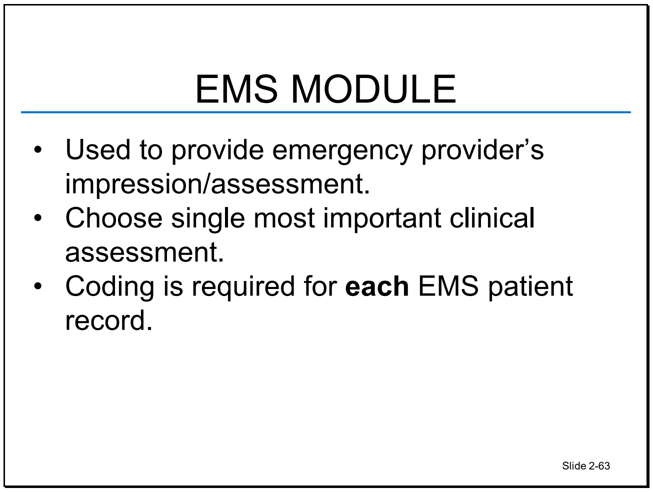
3. The historical data on fire service casualties can be used to develop or update a strategic plan and can include recommendations or standards for:

   a. Training.

   b. Personal protective equipment (PPE).

   c. Fitness.

   d. Diet.



EMS MODULE

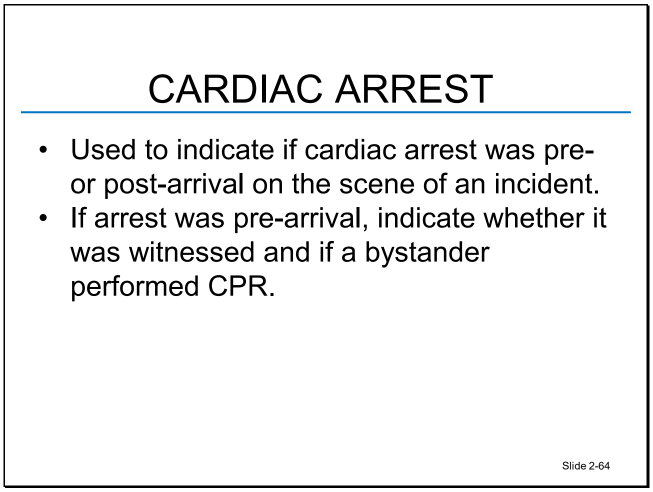
- Used to provide emergency provider's impression/assessment.
- Choose single most important clinical assessment.
- Coding is required for **each** EMS patient record.

D. EMS Module.

1. When completing the EMS Module in NFIRS, check one box that best describes the emergency provider's impression/assessment. When more than one choice is applicable to the patient, choose the single most important clinical assessment that drove the choice of treatment.

2. The available codes for EMS Module are:

Code	Description	Code	Description
10	Abdominal pain	26	Hypovolemia
11	Airway obstruction	27	Inhalation injury, toxic gases
12	Allergic reaction, excludes stings and venomous bite	28	Obvious death
13	Altered level of consciousness	29	Overdose/Poisoning
14	Behavioral — mental status, psychiatric disorder	30	Pregnancy/OB
15	Burns	31	Respiratory arrest
16	Cardiac arrest	32	Respiratory distress
17	Cardiac dysrhythmia	33	Seizure
18	Chest pain	34	Sexual assault
19	Diabetic symptom	35	Sting/Bite
20	Do not resuscitate	36	Stroke/CVA
21	Electrocution	37	Syncope, fainting
22	General illness	38	Trauma
23	Hemorrhaging/Bleeding	00	Other impression/assessment
24	Hyperthermia	NN	None/No patient or refused treatment
25	Hypothermia		



CARDIAC ARREST

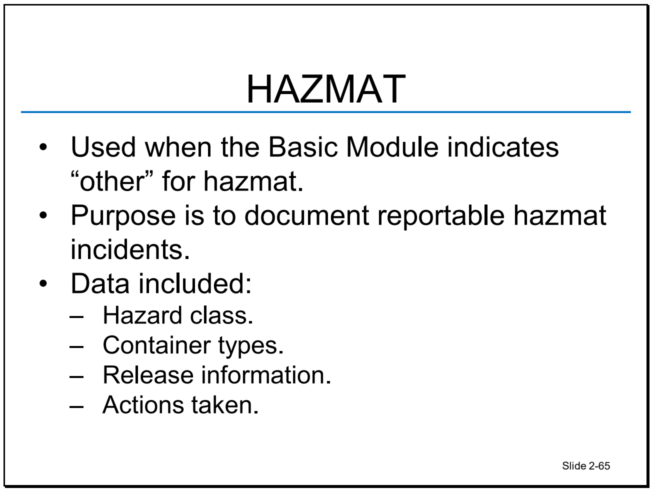
- Used to indicate if cardiac arrest was pre- or post-arrival on the scene of an incident.
- If arrest was pre-arrival, indicate whether it was witnessed and if a bystander performed CPR.

E. Cardiac arrest.

1. This section is used to indicate whether patient cardiac arrest was pre- or post-arrival on the scene of an incident. If it occurred pre-arrival, you should indicate whether or not it was witnessed and/or if bystanders performed CPR.

2. You also should record the initial arrest rhythm by checking the box next to either V-Fib/V-Tach, Other, or Undetermined.

3. Data from this section are used to evaluate pre-hospital CPR and the effect of cardiac care on reducing morbidity.



HAZMAT

- Used when the Basic Module indicates “other” for hazmat.
- Purpose is to document reportable hazmat incidents.
- Data included:
  - Hazard class.
  - Container types.
  - Release information.
  - Actions taken.

F. Hazmat.

1. Use the optional Hazardous Materials Module when the Basic Module (Block H3 — Hazardous Materials Release) indicates “other” for hazmat. Its purpose is to document reportable hazmat incidents.

2. Data collected include:

   a. Hazard class.

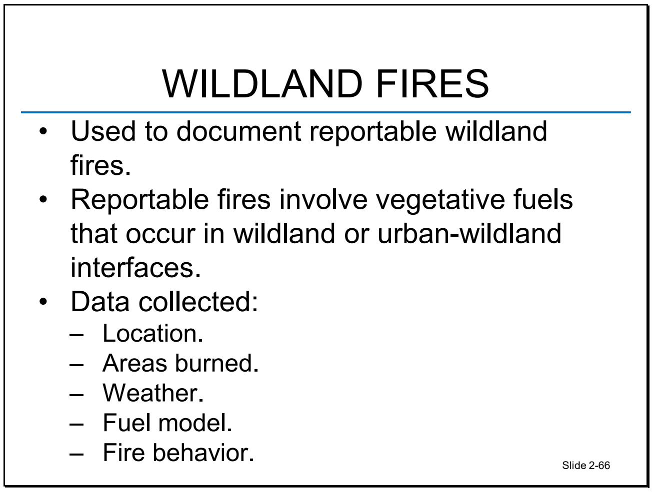
   b. Container types — fixed sites or transportation.

   c. Release information.

   d. Actions taken.

3. A reportable hazmat incident is one in which specialized hazmat resources were dispatched or used, or should have been dispatched or used, for assessing, mitigating, or managing the situation.

4. The Hazardous Materials Module is also used when an incident involves a release or spill of hazmat that exceeds 55 gallons.



WILDLAND FIRES

- Used to document reportable wildland fires.
- Reportable fires involve vegetative fuels that occur in wildland or urban-wildland interfaces.
- Data collected:
  - Location.
  - Areas burned.
  - Weather.
  - Fuel model.
  - Fire behavior.

Slide 2-66

G. Wildland fires.

1. The purpose of the Wildland Fire Module is to document reportable wildland fires. A reportable wildland fire is any fire involving vegetative fuels that occurs in wildland or urban-wildland interface areas, including those fires that threaten or consume structures. To understand better the role of fire on the wildland ecosystem, planned fires also are included in this definition of reportable fires.

2. Data collected include:

   a. Location.

   b. Areas burned.

   c. Weather.

   d. Fuel model.

   e. Fire behavior.

3. The Wildland Fire Module permits wildland fires to be profiled in-depth for resource allocation, incident management and fire impact analysis. In addition, aggregated data on wildland fires will provide invaluable information that can be used by policymakers developing codes and standards, zoning ordinances, and forest management plans.

4. **Using GIS for Wildland Planning:**

   a. This section documents the geographical location of the wildland fire — use it in place of Section B of the Basic Module when traditional addressing methods or the U.S. National Grid standard are not used (www.fgdc.gov/usng).

   b. Enter both the latitude and longitude of the fire location or the Township, Range, Section, Subsection, and Meridian. This information may be of value to local authorities for contacting the owner in connection with the fire and in making a long-term analysis of wildland fires in similar areas or on property under the same ownership.

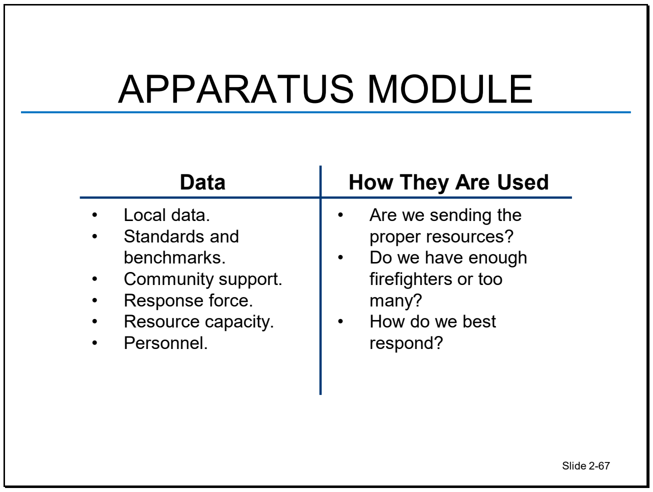
5. The GIS data on wildland fires can be used in planning for:

   a. Resource planning.

   b. Mitigation practices.

   c. Incident management.

   d. Recovery.



APPARATUS MODULE

Data	How They Are Used
Local data.	Are we sending the proper resources?
Standards and benchmarks.	Do we have enough firefighters or too many?
Community support.	How do we best respond?
Response force.	
Resource capacity.	
Personnel.	

H. Apparatus Module.

1. The Apparatus or Resources Module (NFIRS-9) is an optional module that is used to help manage and track apparatus and resources used on incidents. The Personnel Module (NFIRS-10) should be used when details about apparatus and personnel are needed. Resource counts are not needed on the Basic Module (G1) if either the Apparatus or Resources Module, or the Personnel Module is used.

2. The Apparatus or Resources Module is used as a local option to identify the apparatus and personnel sent to an incident. If this module is used, it is not necessary to use the Personnel Module.

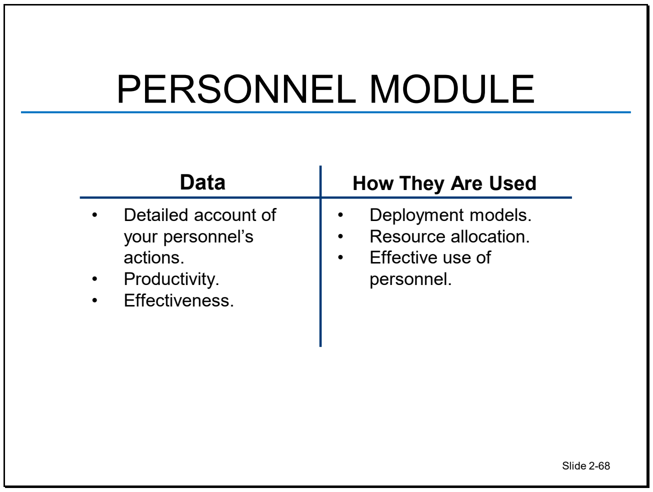
3. On the paper form, lines are available to document nine pieces of apparatus, and additional sheets can be used. This will document all apparatus that were used to control the incident, alarms and district resources.

4. The information is used to answer the following questions (for future incidents):

   a. Are we sending the proper resources?

   b. Do we have enough firefighters or too many?

   c. How do we best respond?



PERSONNEL MODULE

Data	How They Are Used
• Detailed account of your personnel's actions. • Productivity. • Effectiveness.	• Deployment models. • Resource allocation. • Effective use of personnel.

I. Personnel Module.

1. The Personnel Module (NFIRS-10) is an optional module that is used to help manage and track personnel and resources used on incidents. This module can be used in place of the Apparatus or Resources Module (NFIRS-9) if more detail on personnel is needed.

   Note: Either the Personnel Module, or Apparatus or Resources Module may be used — not both.

2. Data collected in this module include:

   a. Detailed account of your personnel's actions.

   b. Productivity.

   c. Effectiveness.

3. The data are used for:

   a. Deployment models.

   b. Resource allocation.

   c. Effective use of personnel.

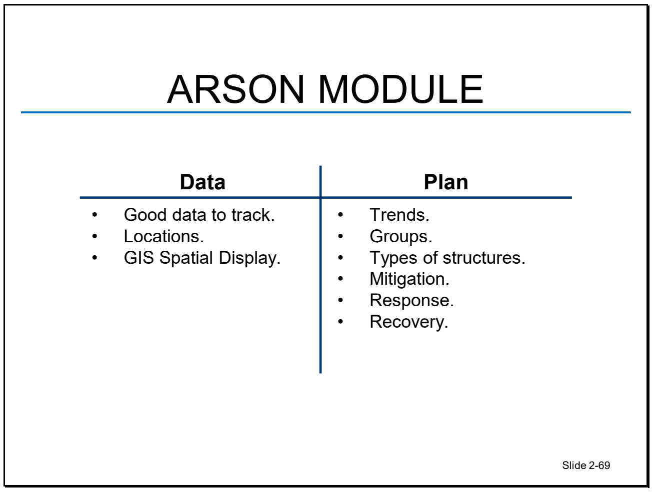
4. Using standards and benchmarks, the planner can develop a response force strategy that builds off the critical task analysis.

5. The planning uses the three strategic objectives:

   a. Life safety.

b. Incident stabilization.

c. Property conservation.



ARSON MODULE

Data

- Good data to track.
- Locations.
- GIS Spatial Display.

Plan

- Trends.
- Groups.
- Types of structures.
- Mitigation.
- Response.
- Recovery.

J. Arson Module.

1. An indispensable tool in the war against arson is the ability to identify when and where the crime takes place, what form it takes, and the characteristics of its targets and perpetrators. Armed with such information, fire service and law enforcement agencies can develop and implement arson prevention initiatives, allowing them to use their resources in the most efficient and effective manner. The NFIRS 5.0 Arson Module (NFIRS-11) was developed with this goal in mind.

2. The data are used to track:

   a. Trends.

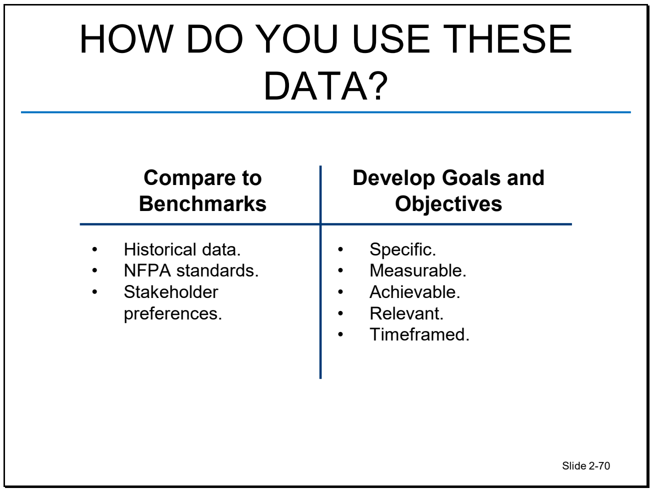
   b. Groups.

   c. Types of structures.

   d. Mitigation.

   e. Response.

   f. Recovery.



HOW DO YOU USE THESE DATA?

Compare to Benchmarks	**Develop Goals and Objectives**
• Historical data. • NFPA standards. • Stakeholder preferences.	• Specific. • Measurable. • Achievable. • Relevant. • Timeframed.

K. How do you use these data?

1. Using the data will help develop goals and objectives to meet the needs of the community.

2. The process should involve many stakeholders in the community. Support of the objectives is needed.

3. The data can be compared to benchmarks such as:

    a. Historical data.

    b. NFPA standards.

    c. Stakeholder preferences.

4. The data can be used to develop goals and objectives that are:

    a. Specific.

    b. Measurable.

    c. Achievable.

    d. Relevant.

    e. Timeframed.

This page intentionally left blank.

ACTIVITY 2.7

NFIRS 5.0 Knowledge Check

Purpose

To allow you to gauge your knowledge of the material studied in class so far.

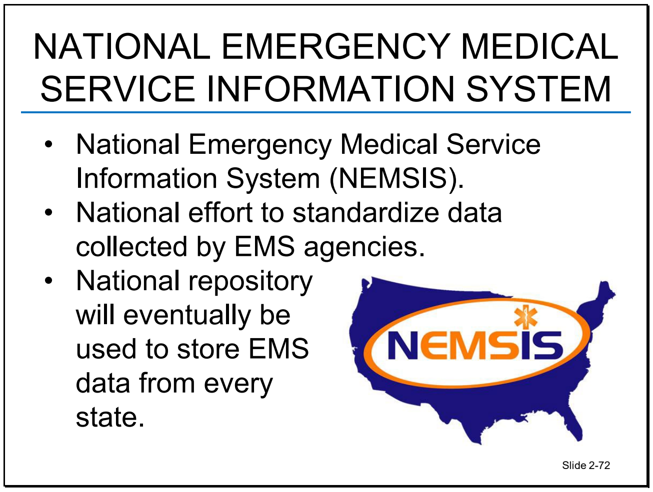
Directions

1. Work as a group.
2. Recall six ways data collected through the NFIRS 5.0 system are being used.
3. Use the table below to record your answers.
4. Be prepared to debrief as a class.

Number	Ways Data Collected Through NFIRS 5.0 are Being Used
1.	
2.	
3.	
4.	
5.	
6.	

This page intentionally left blank.

XII. NATIONAL EMERGENCY MEDICAL SERVICE INFORMATION SYSTEM

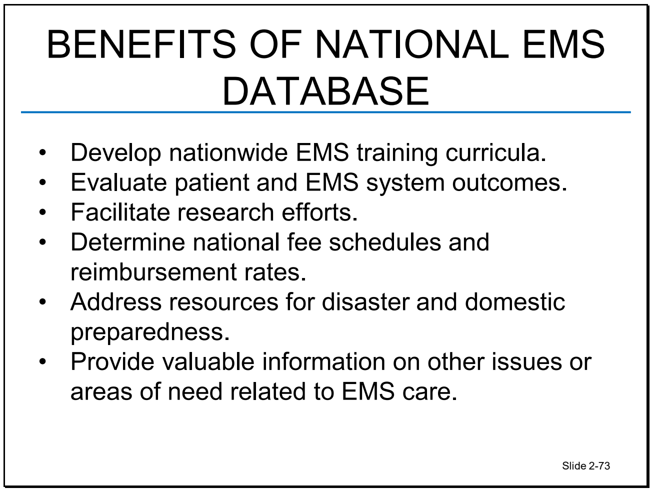


NATIONAL EMERGENCY MEDICAL SERVICE INFORMATION SYSTEM

- National Emergency Medical Service Information System (NEMSIS).
- National effort to standardize data collected by EMS agencies.
- National repository will eventually be used to store EMS data from every state.

A. The National Emergency Medical Service Information System (NEMSIS) is a national repository used to store EMS data from every state in the nation.

B. Since the 1970s, the need for EMS information systems and databases has been well established, and many statewide data systems have been created. However, these EMS systems vary in their ability to collect patient and systems data and allow analysis at a local, state, and national level.



BENEFITS OF NATIONAL EMS DATABASE

- Develop nationwide EMS training curricula.
- Evaluate patient and EMS system outcomes.
- Facilitate research efforts.
- Determine national fee schedules and reimbursement rates.
- Address resources for disaster and domestic preparedness.
- Provide valuable information on other issues or areas of need related to EMS care.

C. Benefits of national EMS database.

The data collected in this central national EMS database will be useful when:

1. Developing nationwide EMS training curricula.

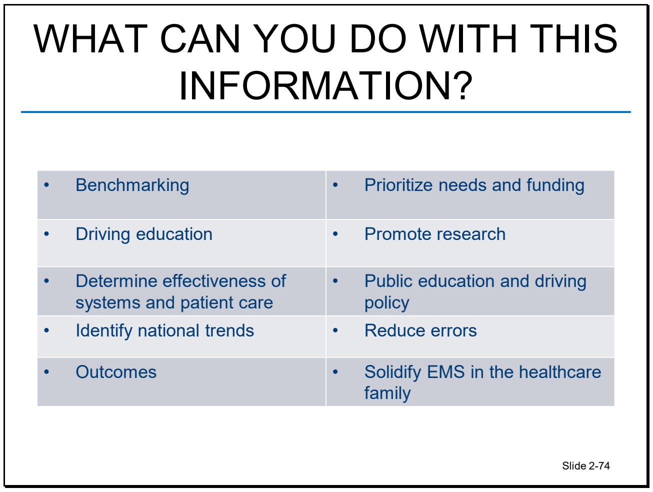
2. Evaluating patient and EMS system outcomes.

3. Facilitating research efforts.

4. Determining national fee schedules and reimbursement rates.

5. Addressing resources for disaster and domestic preparedness.

6. Providing valuable information on other issues or areas of need related to EMS care.



WHAT CAN YOU DO WITH THIS INFORMATION?

• Benchmarking	• Prioritize needs and funding
• Driving education	• Promote research
• Determine effectiveness of systems and patient care	• Public education and driving policy
• Identify national trends	• Reduce errors
• Outcomes	• Solidify EMS in the healthcare family

D. What can be done with NEMSIS information?

Here are just a few of the possibilities that the information collected and stored in NEMSIS can be used for:

1. Benchmarking.

2. Business structure and management.

3. Determining effectiveness of systems and patient care.

4. Driving education.

5. EMS research hypothesis.

6. Identifying national trends.

7. Outcomes.

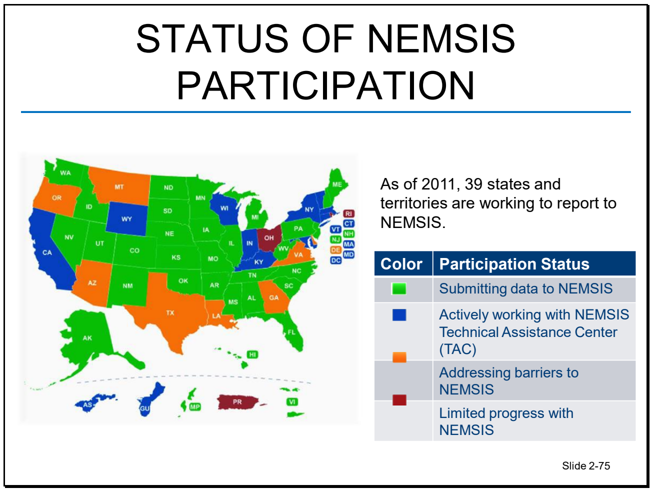
8. Prioritizing needs and funding.

9. Promoting research.

10. Public education and driving policy.

11. Reducing errors.

12. Solidifying EMS in the healthcare family.



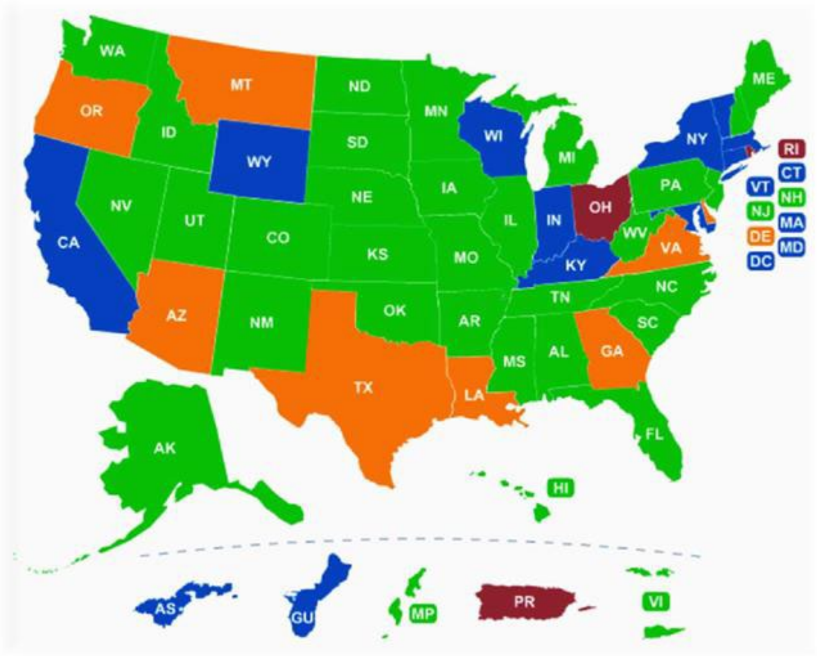
STATUS OF NEMSIS PARTICIPATION

As of 2011, 39 states and territories are working to report to NEMSIS.

Color	Participation Status
🟩	Submitting data to NEMSIS
🟦	Actively working with NEMSIS Technical Assistance Center (TAC)
🟧	Addressing barriers to NEMSIS
🟥	Limited progress with NEMSIS

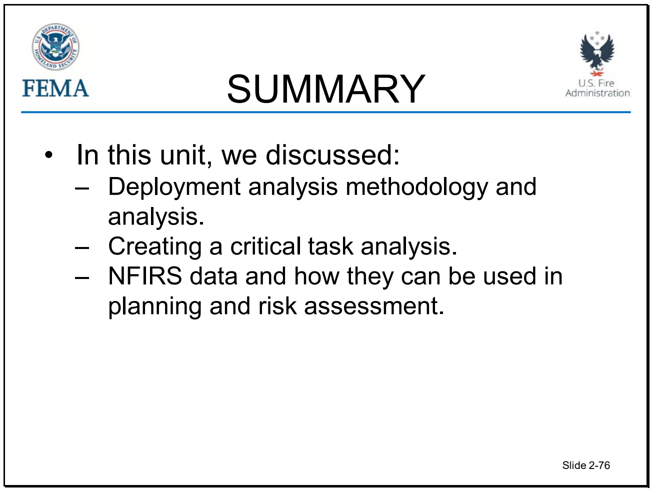
E. Status of NEMSIS Participation.

As of 2011, 39 states and territories, or approximately 70 percent of the nation, are using NEMSIS to report EMS data.



Color	Participation Status
🟩	Submitting data to NEMSIS
🟦	Actively working with NEMSIS Technical Assistance Center (TAC)
🟧	Addressing barriers to NEMSIS
🟥	Limited progress with NEMSIS

XIII. SUMMARY



SUMMARY

- In this unit, we discussed:
  - Deployment analysis methodology and analysis.
  - Creating a critical task analysis.
  - NFIRS data and how they can be used in planning and risk assessment.

Slide 2-76

UNIT 3:
WORKING WITH STATISTICS

TERMINAL OBJECTIVES

The students will be able to:

3.1 Compare the properties of data types.

3.2 Assemble the quantitative components needed for Standards of Cover (SOC).

3.3 Assemble the predictive components needed for an SOC.

ENABLING OBJECTIVES

The students will be able to:

3.1 Explain the properties of data types.

3.2 Differentiate between data types.

3.3 Determine the purpose of descriptive statistics in SOC analysis.

3.4 Calculate descriptive statistics.

3.5 Apply manipulation and summarization techniques to calculate deviations of data.

3.6 Prepare extracted data.

3.7 Modify data parameters.

3.8 Explain the ethical use of data.

3.9 Explain the purpose of predictive statistics and demonstrate their use.

3.10 Calculate coefficients of correlation.

3.11 *Explain use of correlation tools.*

3.12 *Demonstrate the application of discrete application models.*

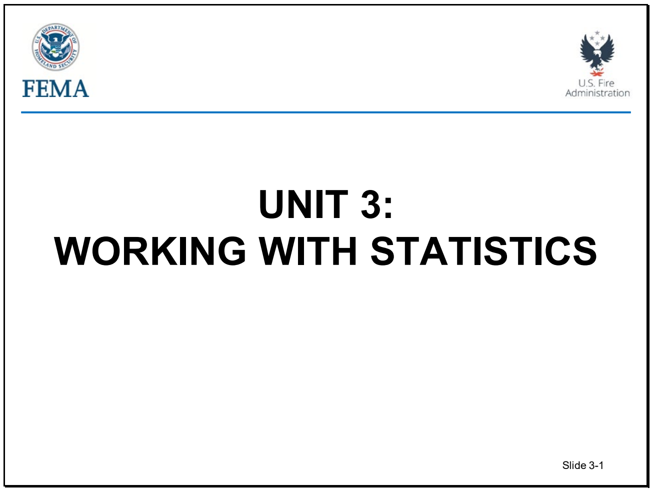
3.13 *Apply interpolative method.*

3.14 *Demonstrate appropriate survey instrument construction skills.*

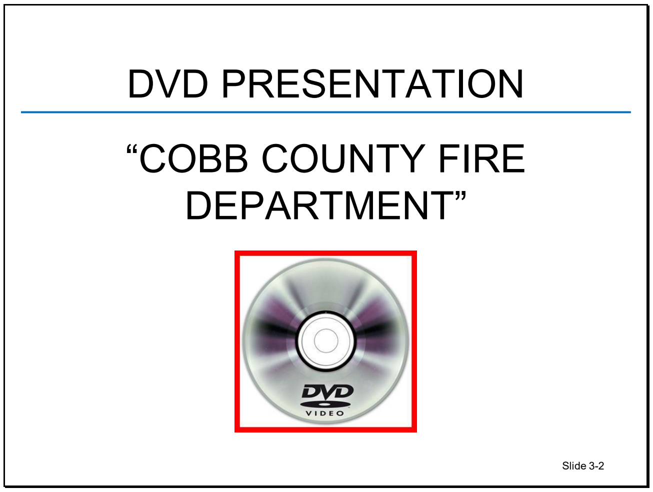
3.15 *Demonstrate assigning relative values — Size, Height, Use, and Probability (SHUP) and Risk, Hazard and Value Evaluation (RHAVE).*

3.16 *Given a scenario, accurately estimate the probability of occurrence within acceptable standard of deviation.*

3.17 *Test the use of risk analysis programs for application within their organization.*

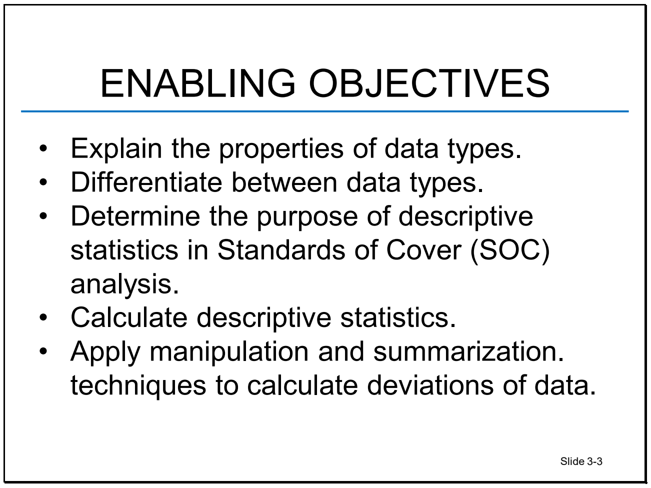


UNIT 3:
WORKING WITH STATISTICS



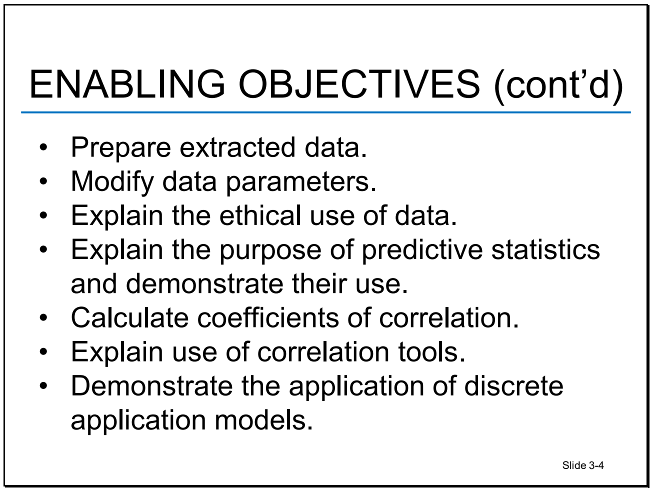
DVD PRESENTATION

“COBB COUNTY FIRE DEPARTMENT”



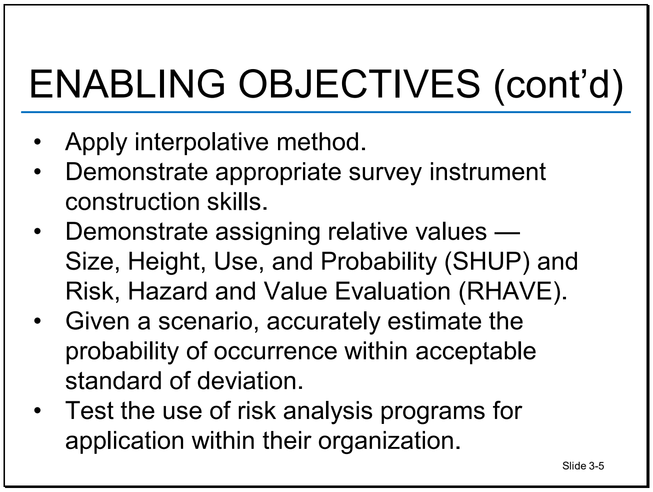
ENABLING OBJECTIVES

- Explain the properties of data types.
- Differentiate between data types.
- Determine the purpose of descriptive statistics in Standards of Cover (SOC) analysis.
- Calculate descriptive statistics.
- Apply manipulation and summarization. techniques to calculate deviations of data.



ENABLING OBJECTIVES (cont'd)

- Prepare extracted data.
- Modify data parameters.
- Explain the ethical use of data.
- Explain the purpose of predictive statistics and demonstrate their use.
- Calculate coefficients of correlation.
- Explain use of correlation tools.
- Demonstrate the application of discrete application models.

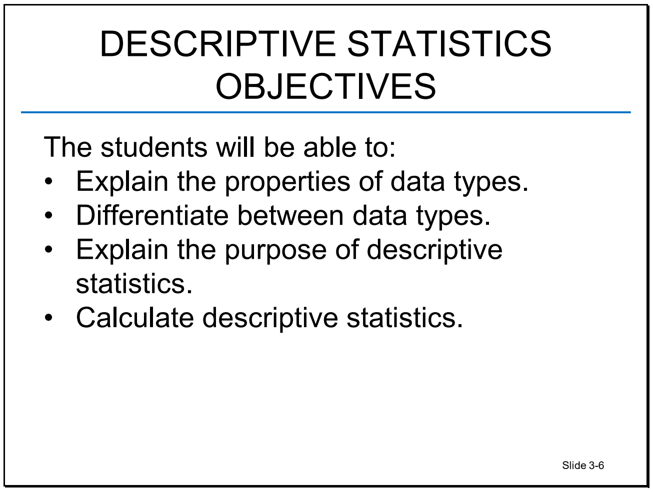


ENABLING OBJECTIVES (cont'd)

- Apply interpolative method.
- Demonstrate appropriate survey instrument construction skills.
- Demonstrate assigning relative values — Size, Height, Use, and Probability (SHUP) and Risk, Hazard and Value Evaluation (RHAVE).
- Given a scenario, accurately estimate the probability of occurrence within acceptable standard of deviation.
- Test the use of risk analysis programs for application within their organization.

I. INTRODUCTION

II. DESCRIPTIVE STATISTICS INTRODUCTION



DESCRIPTIVE STATISTICS OBJECTIVES

The students will be able to:

- Explain the properties of data types.
- Differentiate between data types.
- Explain the purpose of descriptive statistics.
- Calculate descriptive statistics.

A. Descriptive statistics objectives.

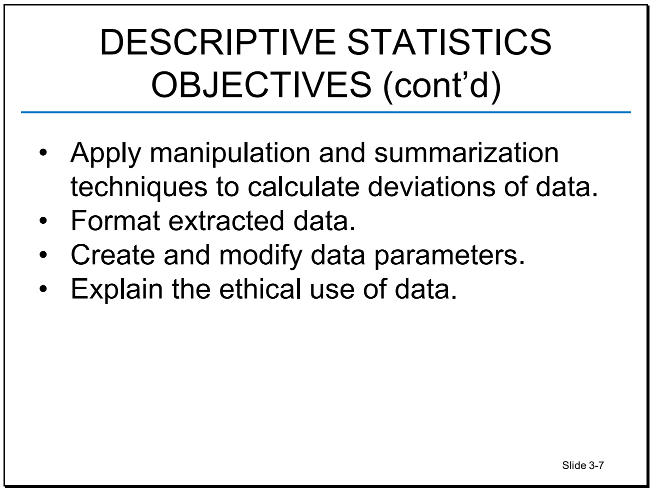
The students will be able to:

1. Explain the properties of data types.

2. Differentiate between data types.

3. Explain the purpose of descriptive statistics.

4. Calculate descriptive statistics.



DESCRIPTIVE STATISTICS OBJECTIVES (cont'd)

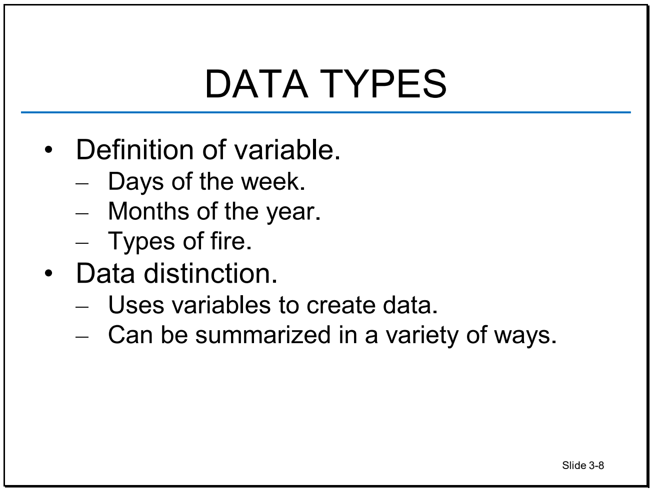
- Apply manipulation and summarization techniques to calculate deviations of data.
- Format extracted data.
- Create and modify data parameters.
- Explain the ethical use of data.

5. Apply manipulation and summarization techniques to calculate deviations of data.

6. Format extracted data.

7. Create and modify data parameters.

8. Explain the ethical use of data.



DATA TYPES

- Definition of variable.
  - Days of the week.
  - Months of the year.
  - Types of fire.
- Data distinction.
  - Uses variables to create data.
  - Can be summarized in a variety of ways.

B. Data types.

1. Definition of a variable: a characteristic that varies or changes.

2. Types of variables.

   a. Days of the week vary from Sunday through Saturday.

   b. Months vary from January through December.

   c. Types of fire vary — structure, vehicle, residential and so on.

3. **Data distinction**.

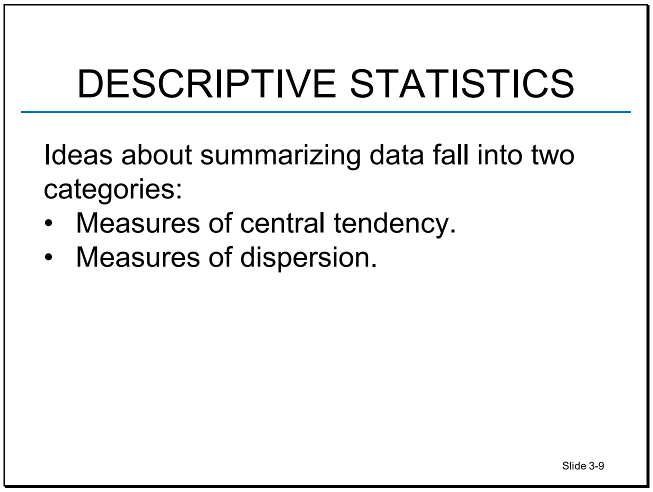
   a. Whenever observations are made on a variable, data are created to be analyzed. Each time a National Fire Incident Reporting System (NFIRS) report is completed, data for the variables listed are created. By listing the day of the week, hour of the day, month, and type of situation found, and values for all other applicable variables in the NFIRS Basic Incident module, data are created.

   b. Data can be summarized in a variety of ways, such as tables, graphs and charts.

   c. There are two types of variables that we focus on: qualitative and quantitative.

4. **Qualitative** variables are classified into groups or categories, also known as categorical variables (data), since they are not measured in quantity but segregated into groups. For example, fires can be broken down into structure, vehicle, explosions, etc. Categorical data in fire service can include property use, cause of ignition, extent of flame damage, etc.

5. On the other hand, **quantitative** variables are variables that reflect some type of measurement.



DESCRIPTIVE STATISTICS

Ideas about summarizing data fall into two categories:

- Measures of central tendency.
- Measures of dispersion.

C. Descriptive statistics.

1. Measures of central tendency.

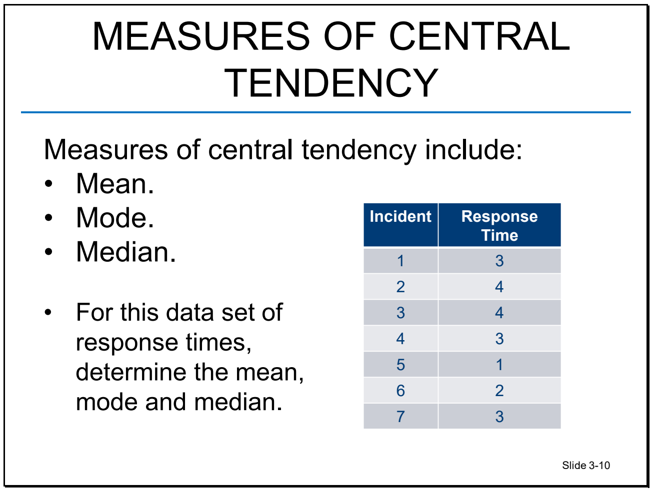
a. Ideas of summarizing data include six basic descriptive measures:

- **Measures of central tendency** provide a summary of the values in a distribution. **Mode** is the only measure of central tendency that can be used for qualitative data. **Mean** is the best choice for a measure of central tendency. It is a response to the exact position of each value in a distribution. **Median** is a better choice than mode for a measure of central tendency but cannot be used for qualitative data.

- Percentiles are generally set at four levels, 1st, 2nd, 3rd and 4th. The 2nd percentile is actually the mean.

- **Fractiles** are percentiles defined outside the usual limits of percentiles. For example, you can have a **fractile** of 90 percent, which means that, of the values in the data set being analyzed, 90 percent of the values will be in the subset created.

- **Measures of dispersion** provide a summary of the variability or spread in a distribution. Measures of dispersion express quantitatively the extent to which the values in a distribution scatter about or cluster together. Measures of dispersion include range, variance and standard deviation.



MEASURES OF CENTRAL TENDENCY

Measures of central tendency include:

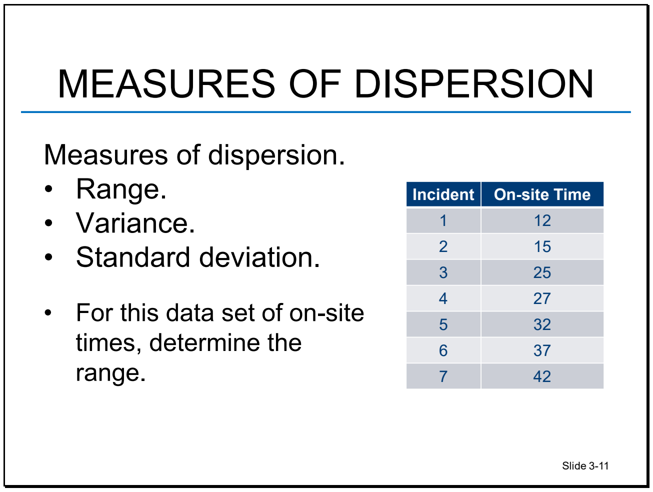
- Mean.
- Mode.
- Median.

- For this data set of response times, determine the mean, mode and median.

Incident	Response Time
1	3
2	4
3	4
4	3
5	1
6	2
7	3

b. Measures of central tendency include:

- **Mean** — the average of a group of numbers (in arithmetic terms).
- **Mode** — the value that occurs most frequently in a distribution.
- **Median** — the middle value (50th percentile) of a distribution.



MEASURES OF DISPERSION

Measures of dispersion.

- Range.
- Variance.
- Standard deviation.

- For this data set of on-site times, determine the range.

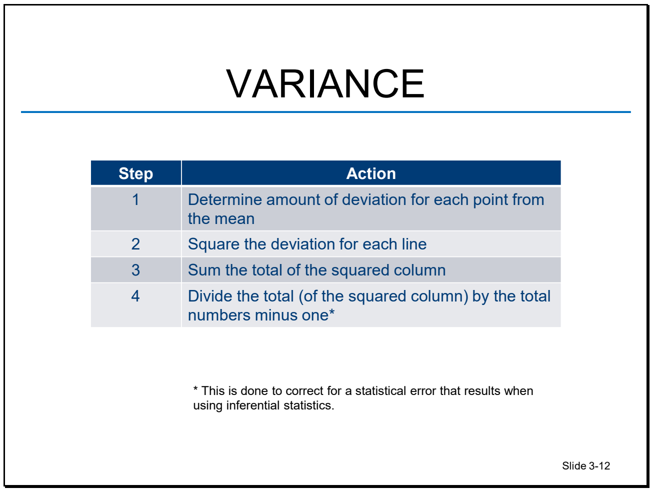
Incident	On-site Time
1	12
2	15
3	25
4	27
5	32
6	37
7	42

2. Measures of dispersion.

a. Measures of dispersion provide a summary of the variability or spread in a distribution. Measures of dispersion express quantitatively the extent to which the values in a distribution scatter about or cluster together.

b. Measures of dispersion include:

- **Range** — most basic measure of dispersion. It is the difference between the lowest and highest value in a distribution.
- **Variance** — the average of the squared differences from the mean.



VARIANCE

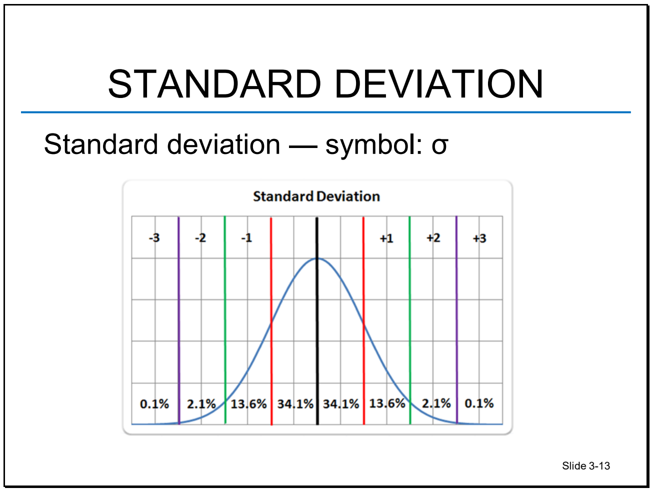
Step	Action
1	Determine amount of deviation for each point from the mean
2	Square the deviation for each line
3	Sum the total of the squared column
4	Divide the total (of the squared column) by the total numbers minus one*

* This is done to correct for a statistical error that results when using inferential statistics.

Slide 3-12

c. The steps to calculate the variance are:

- Determine the amount of deviation for each point from the mean.
- After figuring out the deviation for each line, square the deviation.
- Sum the total of the squared column.
- Divide the total (of the squared column) by the total numbers minus one. This is done to correct for a statistical error that results when using inferential statistics.

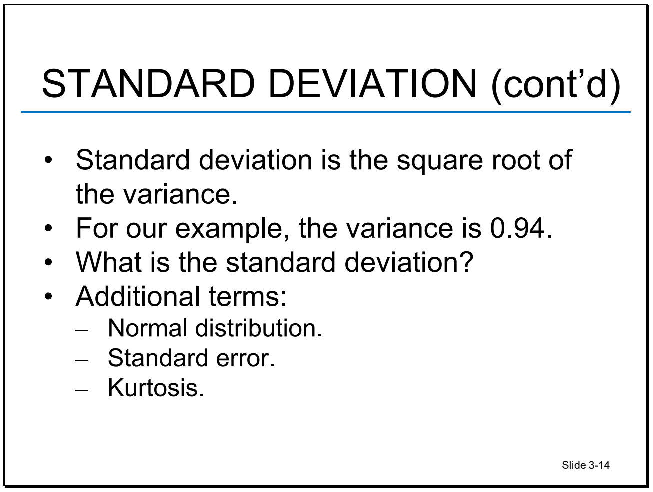


STANDARD DEVIATION

Standard deviation — symbol: σ

3. **Standard deviation**.

a. A measure of how spread out numbers are (in a distribution), as in different kinds of bell curve — steeper or flatter. The bell curve is the shape created when a normal distribution of data is plotted.



STANDARD DEVIATION (cont'd)

- Standard deviation is the square root of the variance.
- For our example, the variance is 0.94.
- What is the standard deviation?
- Additional terms:
  - Normal distribution.
  - Standard error.
  - Kurtosis.

b. Additional terms you will need to be aware of include:

- In a **normal distribution,** the mean, median and mode are all the same. Most statisticians assume a normal distribution, but not all are normal or symmetrical.

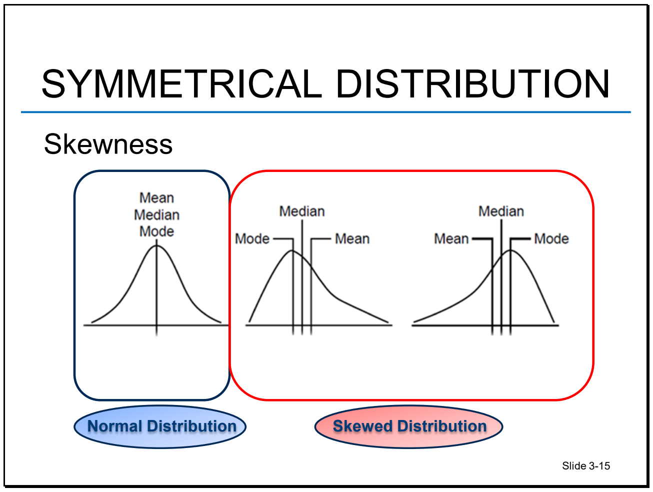
- **Standard error** is the estimated standard deviation of a statistic.

- **Kurtosis** is a measure of whether the data are peaked or flat relative to a normal distribution.

- **High kurtosis** tends to have a distinct peak near the mean, to decline rather rapidly, and to have a heavy tail.

- **Low kurtosis** tends to have a flat top near the mean rather than a sharp peak.

- A **uniform distribution** would be the extreme case.



SYMMETRICAL DISTRIBUTION

Skewness

Normal Distribution

Skewed Distribution

4. Skewness is when the distribution of data is not normal. It can be skewed positively or negatively.

   a. In a **symmetrical distribution** (also called a normal distribution) the mean, median and mode are all the same. Most statisticians assume a normal distribution, but not all are normal or symmetrical.

   b. In a **positive skew**, the extreme scores are at the positive end of the distribution. This is displayed when the "tail" is on the right side and pulls the mean to the right. Since the median and the mode are less responsive to the extreme scores, they remain to the left of the mean. What this means is that the mean has the highest value, the median is in the middle, and the mode has the lowest value.

   c. In a **negative skew**, the extreme scores are at the negative end of the distribution. This is displayed when the "tail" is on the left side and pulls the mean to the left. What this means is that the mean has the lowest value, the median is in the middle, and the mode has the largest value.

This page intentionally left blank.

ACTIVITY 3.1

Basic Statistics

Purpose

To calculate basic statistics.

Directions

1. Work individually or in pairs.
2. Perform the two activities:
   a. Measures of Dispersion.
   b. Standard Deviation.
3. Use the following Excel worksheets:
   a. FDA_Handbook_Ch4_Basic_Stats_partA.xlsx.
   b. FDA_Handbook_Ch4_Basic_Stats_partB.xlsx.
4. Your instructor will provide examples and explain how to use the worksheets and then give you the opportunity to practice.

This page intentionally left blank.

II. DESCRIPTIVE STATISTICS INTRODUCTION (cont'd)

D. Sorting, manipulating, and summarizing large data sets.

This page intentionally left blank.

ACTIVITY 3.2

Data Analysis

Purpose

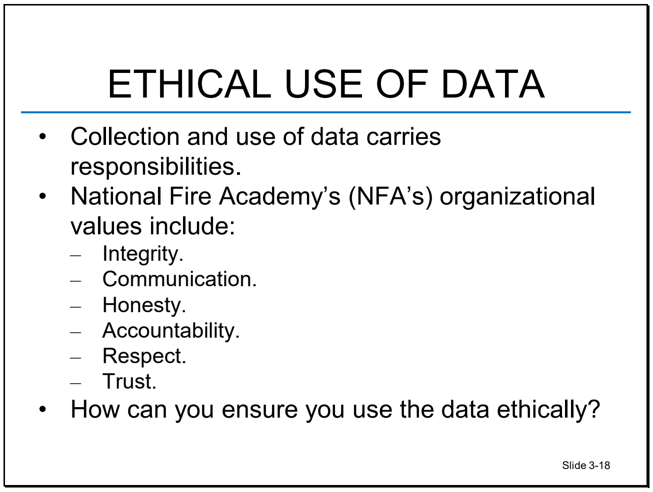
To perform data analysis on a given set of data.

Directions

1. Work individually or in pairs.
2. Perform data analysis using Excel worksheets.
3. Use the following Excel worksheets:
    a. EDA_Ch1_Excel_Lists.xlsx.
    b. EDA_Ch2_Cleaning_Data.xlsx.
    c. EDA_Ch3_Database_Functions.xlsx.
    d. EDA_Ch4_Statistical_Functions.xlsx.
    e. EDA_Ch5_Data_Analysis_ToolPak.xlsx.
4. Your instructor will provide examples and explain how to use the worksheets and then give you the opportunity to practice.

This page intentionally left blank.

II. DESCRIPTIVE STATISTICS INTRODUCTION (cont'd)



ETHICAL USE OF DATA

- Collection and use of data carries responsibilities.
- National Fire Academy's (NFA's) organizational values include:
  - Integrity.
  - Communication.
  - Honesty.
  - Accountability.
  - Respect.
  - Trust.
- How can you ensure you use the data ethically?

E. Ethical use of data.

1. The collection and use of data comes with responsibilities, the main responsibility being ensuring that the data is used ethically.

2. As a reminder, the NFA's organizational values are:

   a. **Integrity:** We adhere to our code of ethics and controls which govern conduct and performance.

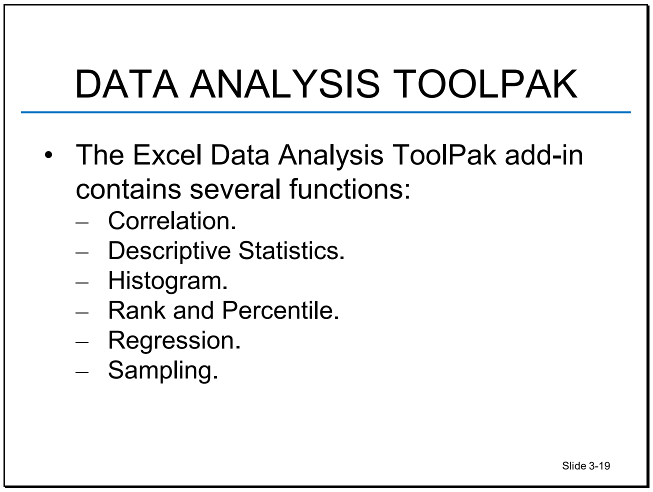
   b. **Communication:** We consistently share and provide access to information throughout the U.S. Fire Administration (USFA) to enhance collaboration and to eliminate ambiguity, frustration and uncertainty.

   c. **Honesty:** We embrace fairness and equity as paramount to all human capital and business affairs.

   d. **Accountability:** We are obligated and willing to accept responsibility and to answer for the results of our performance and conduct.

   e. **Respect:** We consider all USFA members worthy of high regard and have a sincere desire to see others succeed.

   f. **Trust:** We optimistically rely on the character, ability, and strength of each member to contribute wholeheartedly to the success of the USFA.



DATA ANALYSIS TOOLPAK

- The Excel Data Analysis ToolPak add-in contains several functions:
  - Correlation.
  - Descriptive Statistics.
  - Histogram.
  - Rank and Percentile.
  - Regression.
  - Sampling.

F. Data Analysis ToolPak.

1. The Data Analysis ToolPak is a group of Excel spreadsheets that allows you to perform descriptive statistical calculations easily using pre-formatted worksheet.

2. The table below describes the tools and their features in the ToolPak:

Tool	Description
Correlation	Quantifies the relationship between two sets of data
Descriptive Statistics	Summarizes large data sets and calculates key values
Histogram	Creates a frequency distribution and histogram chart
Rank and Percentile	Calculates rank and percentile
Regression	Returns regression statistics and formula values
Sampling	Randomly selects items from data sets

3. The ToolPak also includes instructions for installation and a blank worksheet for your use.

   a. **Descriptive Statistics tool** — contains two sections:

      - The first section is for checking to see if the Data Analysis ToolPak has been installed.

- In the second section, how to use the Data Analysis ToolPak is shown. There is a graphic that shows the command as it should appear in Excel. This list shows approximately one-half of the tools available. Scrolling down will reveal the balance of the content of this tool.

b. **Correlation** — This quantifies the relationship between two sets of data. There is a graphic that shows the command as it should appear in Excel. This will return a value which is the Pearson Correlation or R. A short data set is provided for your practice with the answers shown below. Step-by-step instructions are given.

c. **Descriptive Statistics** — This tool summarizes data sets and calculates key values. There is a graphic that shows the command as it should appear in Excel.

d. **Histogram** — This tool shows how values in the data set are distributed across various categories. There is a graphic that shows the command as it should appear in Excel.

e. **Rank and Percentile** — This tool calculates rank and percentile information for the values in a given data set. There is a graphic that shows the command as it should appear in Excel.

f. **Regression** — This tool returns regression statistics and will generate formula values which will allow you to create what-ifs for specific values. There is a graphic that shows the command as it should appear in Excel.

g. **Sampling** — This tool will allow you to randomly select items from the data set. It could be used on a very large data set to develop a manageable data set to either use or present. There is a graphic that shows the command as it should appear in Excel.

This page intentionally left blank.

ACTIVITY 3.3

Manipulating Data

Purpose

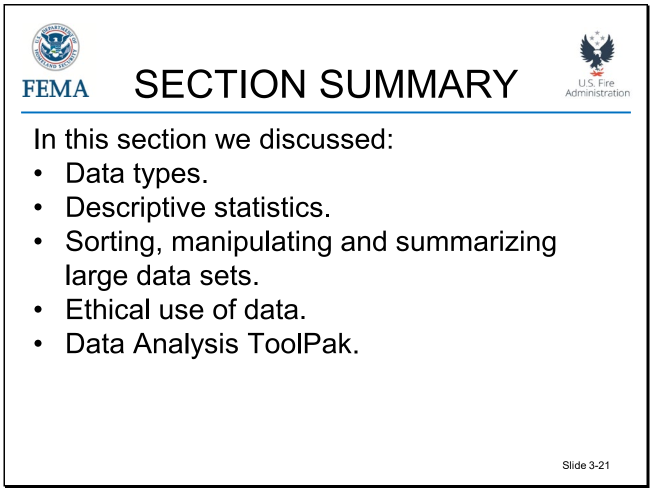
To practice data manipulation using pre-formatted Excel worksheets.

Directions

1. Work individually or in pairs.
2. Display data using Excel worksheets.
3. Use the following Excel worksheets:
   a. FDA_Handbook_Charts.xlsx.
   b. FDA_Handbook_Ch2_Histograms.xlsx.
   c. FDA_Workbook_Ch2_Histograms.xlsx.
   d. FDA_Handbook_Ch3_Charts.xlsx.
   e. FDA_Workbook_Ch3_Charts.xlsx.
   f. FDA_Handbook_Ch5_Analysis_of_Tables.xlsx.
   g. FDA_Workbook_Ch5_Analysis_of_Tables.xlsx.
   h. FDA_Handbook_Ch9_Queueing_Analysis.xlsx.
   i. FDA_Workbook_Ch9_Queueing_Analysis.xlsx.
4. Your instructor will provide examples and explain how to use the worksheets and then give you the opportunity to practice.

This page intentionally left blank.

II. DESCRIPTIVE STATISTICS INTRODUCTION (cont'd)



FEMA

U.S. Fire Administration

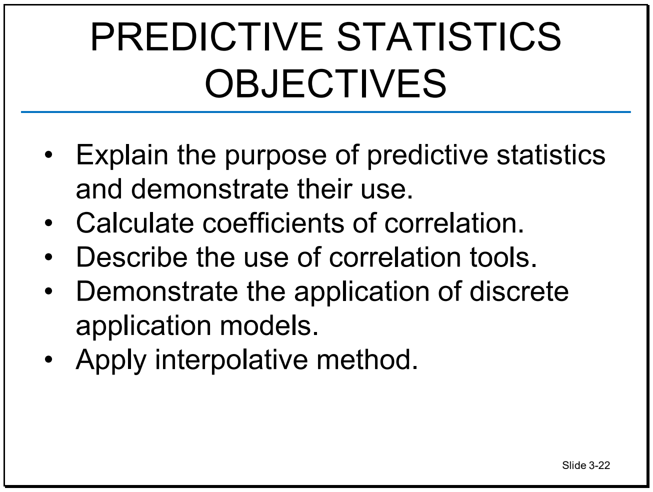
SECTION SUMMARY

In this section we discussed:
- Data types.
- Descriptive statistics.
- Sorting, manipulating and summarizing large data sets.
- Ethical use of data.
- Data Analysis ToolPak.

G. Section summary.

In this section, we discussed data types, working with descriptive statistics, ethical use of data, and the Data Analysis ToolPak. Time was also provided for hands-on activities to practice with data sets and pre-formatted Excel worksheets.

III. PREDICTIVE STATISTICS INTRODUCTION



PREDICTIVE STATISTICS OBJECTIVES

- Explain the purpose of predictive statistics and demonstrate their use.
- Calculate coefficients of correlation.
- Describe the use of correlation tools.
- Demonstrate the application of discrete application models.
- Apply interpolative method.

A. Predictive statistics objectives.

The students will be able to:

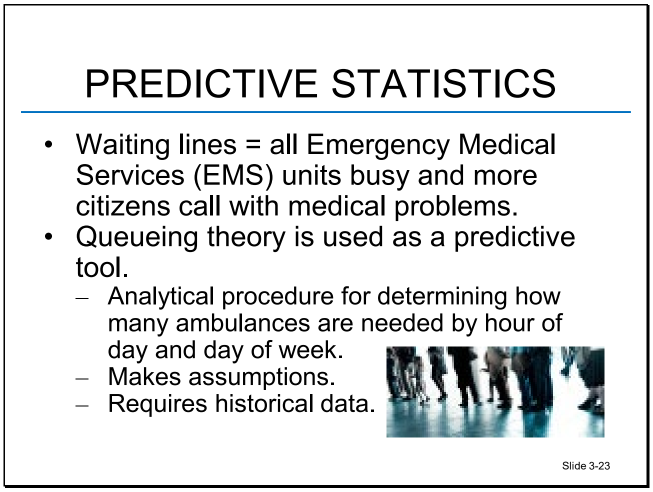
1. Explain the purpose of predictive statistics and demonstrate their use.

2. Calculate coefficients of correlation.

3. Describe the use of correlation tools.

4. Demonstrate the application of discrete application models.

5. Apply interpolative method.



PREDICTIVE STATISTICS

- Waiting lines = all Emergency Medical Services (EMS) units busy and more citizens call with medical problems.
- Queueing theory is used as a predictive tool.
  - Analytical procedure for determining how many ambulances are needed by hour of day and day of week.
  - Makes assumptions.
  - Requires historical data.

Slide 3-23

B. Predictive statistics.

1. Waiting lines — No one likes waiting in lines, whether it is at the post office, grocery store, or for emergency services.

   a. Within fire departments we face the challenge of determining how many Emergency Medical Services (EMS) units we need to adequately handle requests from citizens for EMS.

   b. A “waiting line” develops when all EMS units are busy and more citizens call with medical problems.

   c. Queueing theory provides methods to analyze whether waiting lines will occur and what the consequences of waiting lines will be.

2. Queueing theory:

   a. Is used as a predictive tool.

   b. Is an analytical procedure for determining how many ambulances are needed by hour of day and day of week.

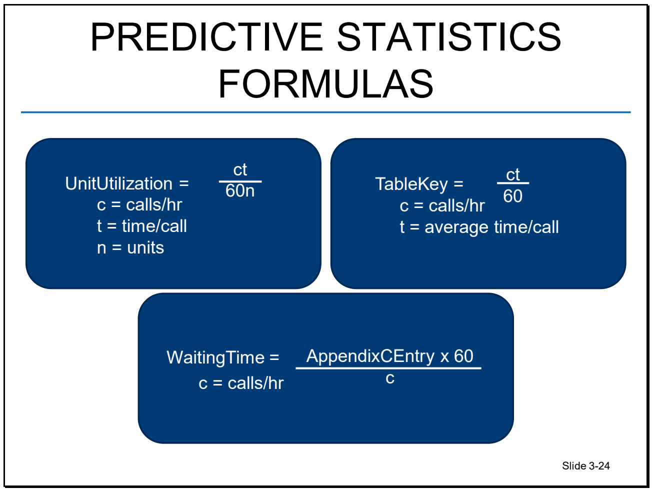
   c. Assumes a fixed number of EMS units in the field and estimates their performance measures.

d. Can provide assistance in managing and operating the emergency medical services for a fire department. Potential applications include:

- Estimation of how busy EMS units will be based on an expected workload of EMS calls.
- Estimation of the probability that a citizen's call for medical services will have to wait because all EMS units are busy.
- Estimation of the average number of citizens waiting for medical service.
- Estimation of the average waiting time for these citizens.
- Estimation of the number of EMS units needed to satisfy objectives established by a fire department.

e. Has three formulas that can be used to help determine utilization and make decisions:

- Unit utilization.
- Table key.
- Waiting time.



PREDICTIVE STATISTICS FORMULAS

UnitUtilization = $\frac{ct}{60n}$
c = calls/hr
t = time/call
n = units

TableKey = $\frac{ct}{60}$
c = calls/hr
t = average time/call

WaitingTime = $\frac{\text{AppendixCEntry x 60}}{\text{c}}$
c = calls/hr

3. Unit utilization.

Begin with historical data and make assumptions based on those data. In our example, we will make the following assumptions:

a. Time period: Saturday evenings from 8 p.m. to midnight.

b. Average call rate = 2 per hour.

c. Mean time for a call = 40 minutes.

d. Two EMS units are available in the field to respond to citizens' calls.

  - Next calculate how busy the two EMS units will be for the given time period using the following calculation:

$$\mathbf{UnitUtilization} = \frac{\mathbf{ct}}{\mathbf{60n}}$$

c = calls/hour
t = time/call
n = units

  - For our example, use the following information:

    -- Two units provide 120 minutes of unit time each hour (2 units x 60 minutes).

    -- Calls require 80 minutes each hour (2 calls x 40 minutes of unit time).

e. What is the unit utilization for three EMS units?

$$\mathbf{UnitUtilization} = \frac{\mathbf{2 \times 40}}{\mathbf{3 \times 60}} = \mathbf{44.4}\%$$

4. Table key.

a. By applying queueing theory, we can estimate several other system performance measures. Three of the most common are:

  - The probability that a person calling for service will have to wait.

  - The average number of citizens waiting for service.

  - The average waiting time for these citizens.

b. These performance measures take into account the inevitable variation in workload.

c. Formulas for the probability of delay and the length of a waiting line are complicated. Appendices C and D are provided to ease the calculation burden.

d. An example of another formula that can be used to measure system performance waiting lines is the table key:

$$\mathbf{TableKey} = \frac{\mathbf{ct}}{\mathbf{60}}$$

c = calls/hour
t = average time/call

e. The table key is actually the amount of work we expect each hour.

f. In our previous example, we use the following parameters:

- c = 2 calls per hour.
- t = 40 minutes.

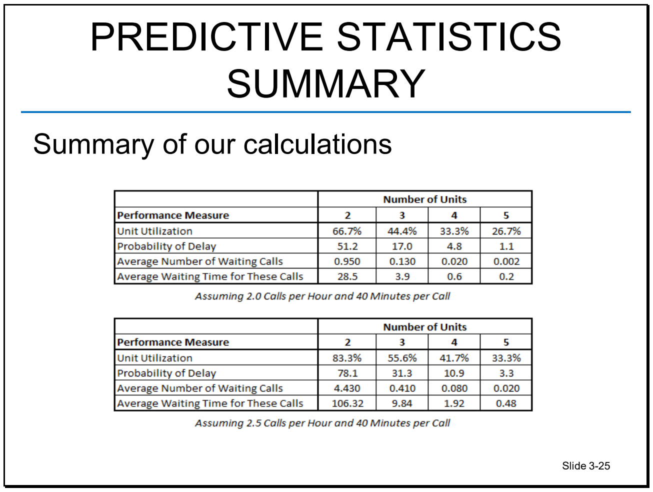
5. Waiting time.

a. This formula allows us to determine the amount of time that waiting citizens will have to wait before a dispatch occurs.

b. Use the formula below to calculate the waiting time:

$$\mathbf{WaitingTime} = \frac{\mathbf{AppendixCEntry} \times \mathbf{60}}{\mathbf{c}}$$

c = calls/hour



PREDICTIVE STATISTICS SUMMARY

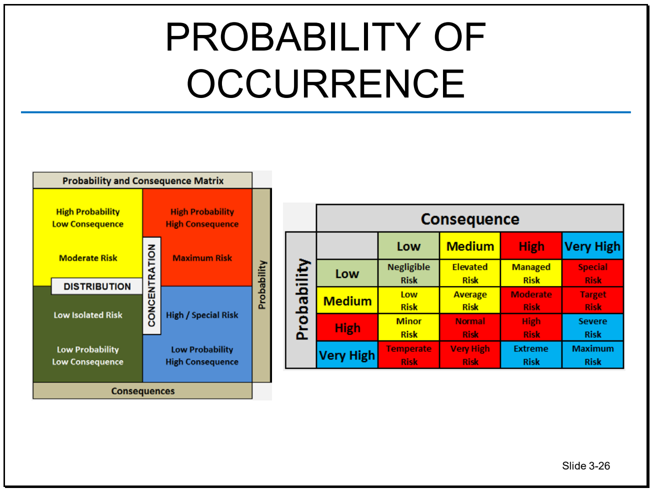
Summary of our calculations

	Number of Units			
Performance Measure	**2**	**3**	**4**	**5**
Unit Utilization	66.7%	44.4%	33.3%	26.7%
Probability of Delay	51.2	17.0	4.8	1.1
Average Number of Waiting Calls	0.950	0.130	0.020	0.002
Average Waiting Time for These Calls	28.5	3.9	0.6	0.2

Assuming 2.0 Calls per Hour and 40 Minutes per Call

	Number of Units			
Performance Measure	**2**	**3**	**4**	**5**
Unit Utilization	83.3%	55.6%	41.7%	33.3%
Probability of Delay	78.1	31.3	10.9	3.3
Average Number of Waiting Calls	4.430	0.410	0.080	0.020
Average Waiting Time for These Calls	106.32	9.84	1.92	0.48

Assuming 2.5 Calls per Hour and 40 Minutes per Call



PROBABILITY OF OCCURRENCE

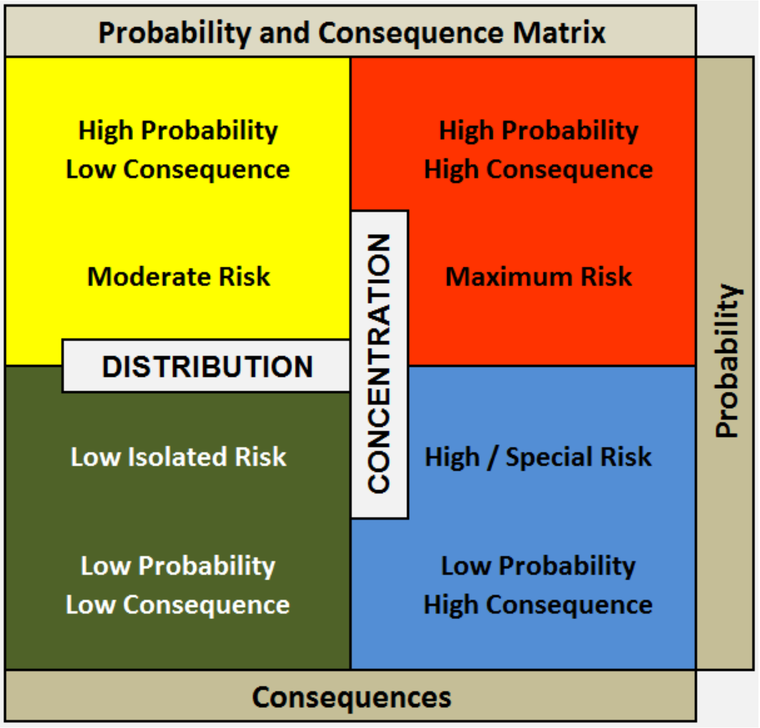
		Consequence			
		Low	Medium	High	Very High
Probability	Low	Negligible Risk	Elevated Risk	Managed Risk	Special Risk
	Medium	Low Risk	Average Risk	Moderate Risk	Target Risk
	High	Minor Risk	Normal Risk	High Risk	Severe Risk
	Very High	Temperate Risk	Very High Risk	Extreme Risk	Maximum Risk

C. Probably of occurrence (probability and consequence).

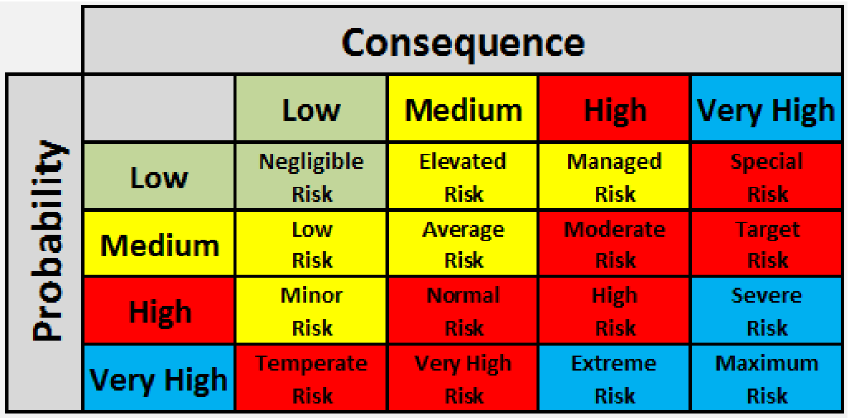
1. A probability provides a quantitative description of the likely occurrence of a particular event.

2. Risk plays a part in probability of occurrence and looks at the degree of harm that might occur.

3. Probability of occurrence assesses the risk and harm; below is a diagram that shows the Probability and Consequence Matrix.



4. Below is the rating matrix of probability and consequence.



		Consequence			
		Low	Medium	High	Very High
Probability	Low	Negligible Risk	Elevated Risk	Managed Risk	Special Risk
	Medium	Low Risk	Average Risk	Moderate Risk	Target Risk
	High	Minor Risk	Normal Risk	High Risk	Severe Risk
	Very High	Temperate Risk	Very High Risk	Extreme Risk	Maximum Risk

This page intentionally left blank.

ACTIVITY 3.4

Probability and Consequence Matrix Calculator

Purpose

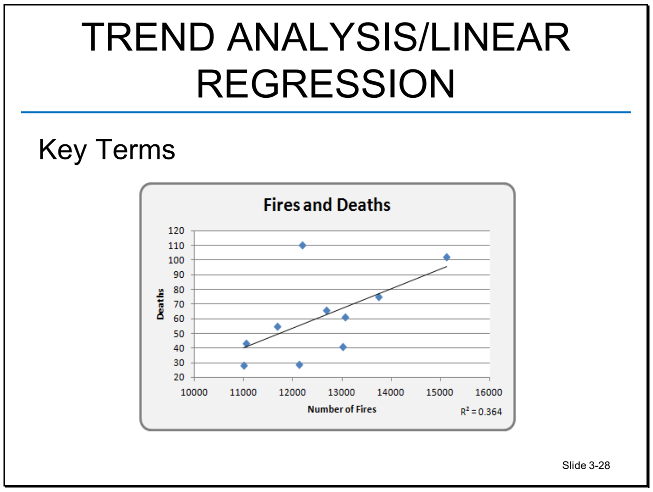
To calculate the probability of occurrence.

Directions

1. Work in pairs for this activity.
2. Your instructor will guide you through the use of each Risk Levels Excel workbook and provide data for use with each scenario.
3. You will have 40 to 45 minutes per topic to practice using the workbook.
4. You will debrief between each topic.
5. Use an easel pad to list issues with workbooks, if needed.

This page intentionally left blank.

III. PREDICTIVE STATISTICS INTRODUCTION (cont'd)



TREND ANALYSIS/LINEAR REGRESSION

Key Terms

Slide 3-28

D. Trend analysis/linear regression.

1. Key terms.

a. **Correlation** is a statistical measure which indicates the degree to which one variable changes with respect to another variable's change. For example, calls for EMS generally increase with population growth.

b. **Positive correlation** is indicated when a population increases; more medical service calls would be expected.

- Generally, the greater the association (correlation) between two variables, the more accurately a prediction can be made of the standing in one variable from the standing in the other.

c. **Scatter Diagram**.

- The scatter diagram displayed on the slide shows the number of fire deaths and the number of other residential fires for the 10-year period 1989 to 1998.

- As seen in the diagram, as the number of fires increases so does the number of deaths; although a correlation is shown, it is **not perfect.**

d. **Not perfect** refers to the fact that the points do not fall on a straight line.

e. **Variables** are labeled to help depict and explain the relationships. In this diagram the following variables are noted as:

- Independent variable = fires.
- Dependent variable = deaths.
- The number of fires influences the number of fire-related deaths.

2. Correlation coefficient.

a. Correlation coefficient measures the strength of association between two variables.

b. The correlation is always between -1 and +1.

c. A correlation of exactly -1 or +1 is called a **perfect correlation** and means that all of the points fall on a straight line.

d. A correlation of zero indicates no relationship between the variables and would be represented on a scatter diagram as random points with no discernible direction.

e. As a correlation coefficient moves from zero in either direction, the strength of the association between the two variables increases.

f. Also note that correlations are not arithmetically related to each other. For example, a correlation of 0.6 is not twice as strong as a correlation of 0.3. Although it is obvious that a correlation of 0.6 reflects a stronger association than a correlation of 0.3, there is no exact specification of the difference.

g. In order to make direct comparisons between correlations, the correlation coefficient must be converted to a **coefficient of determination.**

3. Calculating the correlation.

a. Calculating the correlation is much easier today than it was previously. Many pocket calculators and statistical software (including Excel) contain the formula.

b. However, if you need to calculate by hand, the deviation-score method is used. The most commonly used deviation-score method is the **Pearson product-moment correlation coefficient**.

c. The steps to use the Pearson product-moment correlation coefficient method are as follows:

- List the pairs of scores in two columns. The order in which the pairs are listed makes no difference in the value of the correlation. However, if one raw score is shifted, the one it is paired with must be shifted as well.

- Find the mean for the raw scores of each variable.

- Convert each score in both variables to a deviation score by subtracting the respective mean from each.

- Calculate the standard deviation for both variables. Since the deviation scores are already done, they only need to be squared and summed. Divide each of these totals by the number of pairs (in this case 10) and take the square root of each.

- Multiply each pair of deviation scores, known as the cross-product, and total the results.

- Next, multiply the two standard deviations by each other and multiply that result by the number of pairs (10).

- Divide the results of Step 5 by the results of Step 6. The result is the Pearson product-moment correlation coefficient.

- Square this for the coefficient of determination.

This page intentionally left blank.

ACTIVITY 3.5

Correlation Coefficient

Purpose

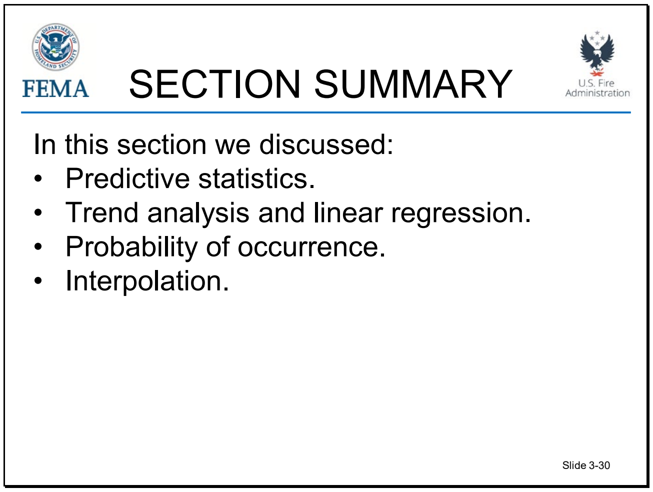
To calculate correlation coefficients using Excel worksheets.

Directions

1. Work individually or in pairs.
2. Display data using Excel worksheets.
3. Use the following Excel worksheets:
    a. FDA_Handbook_Ch6_Correlation.xlsx.
    b. FDA_Workbook_Ch6_Correlation.xlsx.
4. Your instructor will provide examples and explain how to use the worksheets and then give you the opportunity to practice.

This page intentionally left blank.

III. PREDICTIVE STATISTICS INTRODUCTION (cont'd)



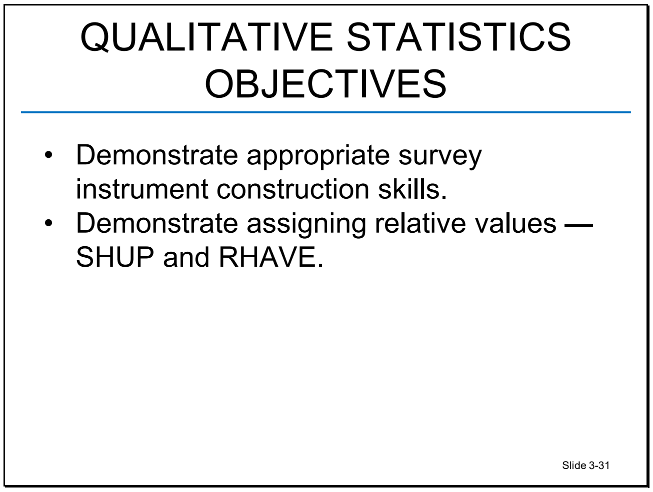
SECTION SUMMARY

In this section we discussed:
- Predictive statistics.
- Trend analysis and linear regression.
- Probability of occurrence.
- Interpolation.

E. Section summary.

In this section, we covered working with predictive statistics. The topics discussed included "waiting lines," queueing theory, trend analysis, and correlation coefficient. The students were given the opportunity to practice working with predictive statistics manually.

IV. WORKING WITH QUALITATIVE STATISTICS INTRODUCTION



QUALITATIVE STATISTICS OBJECTIVES

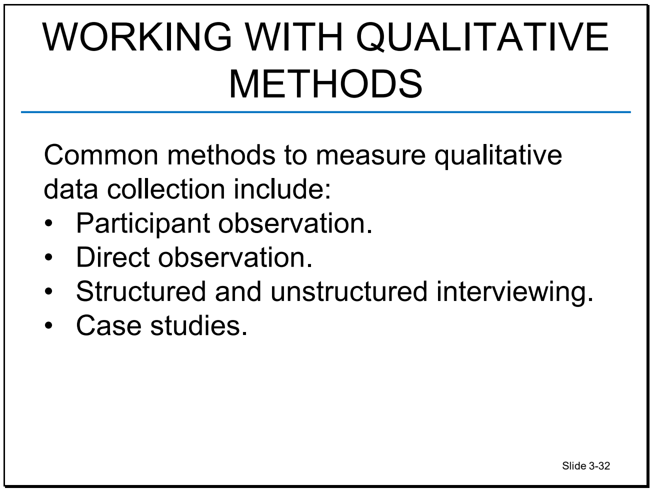
- Demonstrate appropriate survey instrument construction skills.
- Demonstrate assigning relative values — SHUP and RHAVE.

A. Objectives.

The students will be able to:

1. Demonstrate appropriate survey instrument construction skills.

2. Demonstrate assigning relative values — Size, Height, Use, and Probability (SHUP) and Risk, Hazard and Value Evaluation (RHAVE).



WORKING WITH QUALITATIVE METHODS

Common methods to measure qualitative data collection include:

- Participant observation.
- Direct observation.
- Structured and unstructured interviewing.
- Case studies.

B. Working with qualitative methods.

1. Information gathered using this method is gathered and then analyzed in an interpretive manner, subjective or diagnostic manner.

2. There are a number of methods that are common in qualitative measurement; commonly used methods discussed here include:

    a. Participant observation.

        - This is one of the most common methods for qualitative data collection, but it is also one of the most demanding.

        - This method requires the researcher to become a participant in the culture or context being observed.

        - There is extensive literature on:

            -- Entering the context.

            -- The role of the researcher as a participant.

            -- The collection and storage of field notes.

            -- The analysis of field data.

- This method may require months or years of intensive work because the researcher needs to become accepted as a natural part of the culture in order to assure that the observations are representative of the natural situation.

b. Direct observation differs from participant observation in the following ways:

- A direct observer does not typically try to become a participant in the context. That said, a direct observer does strive to be as unobtrusive as possible so as not to bias the observations.

- Direct observation is viewed as a more detached perspective because the researcher is watching instead of participating. Therefore, technology can be a useful part of direct observation. For example, someone can videotape an observation or watch the interaction from behind one-way mirrors.

- Direct observation tends to be more focused than participant observation. This method allows the researcher to focus on certain situations or people rather than trying to become immersed in the entire context. For example, the researcher might observe child-mother interactions under specific circumstances in a laboratory setting from behind a one-way mirror, looking especially for the nonverbal cues being used.

c. Structured interviewing.

This method involves direct interaction between the researcher and a respondent or group. This method is categorized by:

- Using the same protocol or list of questions for every interview conducted.

- Asking the questions in the same order every time.

- Using the answer scale to ensure an apples-to-apples comparison.

d. Unstructured interviewing.

- This method involves direct interaction between the researcher and a respondent or group. This method differs from the more traditional structured interview method as follows:
    - Although the researcher may have some initial guiding questions or core concepts to ask about, there is no formal structured instrument or protocol. This provides more flexibility.
    - This method allows the interviewer to move the conversation freely in any direction of interest that may come up. This allows a topic to be explored more broadly if deemed necessary by the researcher.
- The lack of structure can provide challenges. Because each interview tends to be unique with no predetermined set of questions asked of all respondents, it is usually more difficult to analyze unstructured interview data, especially when synthesizing across respondents.

e. Case studies.

- A case study is an intensive study of a specific individual or specific context. For example, Freud developed case studies of several individuals as the basis for the theory of psychoanalysis, and Piaget did case studies of children to study developmental phases.
- There is no single way to conduct a case study, and a combination of methods can be used (such as direct observation and unstructured interviewing).

ACTIVITY 3.6

Assigning Relative Values

Purpose

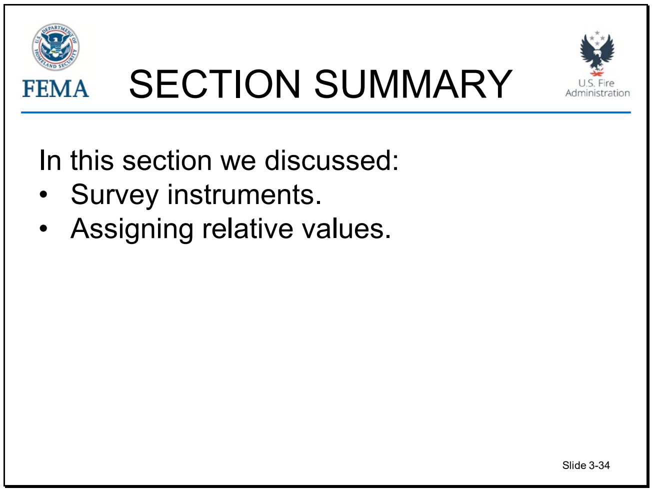
To assign relative values to qualitative data using the given tools.

Directions

You will work exclusively in the SHUP and RHAVE worksheets with the data provided.

This page intentionally left blank.

IV. WORKING WITH QUALITATIVE STATISTICS INTRODUCTION (cont'd)



SECTION SUMMARY

In this section we discussed:
- Survey instruments.
- Assigning relative values.

C. Section summary.

In this section, we reviewed the current tools available for assigning relative values to qualitative data and provided an opportunity for the students to experiment using the tools.

This page intentionally left blank.

ACTIVITY 3.7

Analysis of Risk-calculator Tools

Purpose

This activity is a review of the various calculators used in this unit. You are encouraged to once again review the operation of each of the calculators and to consider potential use by your own agency.

Directions

The calculators in this unit consist of six different instruments. A review of each follows:

1. Vulnerability Assessment Matrix (NOAA – VAM & OAM.xlsx)

   This calculator looks at potential man-made and natural disasters that may affect a community or region. It looks at three different elements which generate an overall score. The goal of using this tool is to rank those disasters using a systematic approach. A final portion of this tool sorts the results from high to low and produces a graph with the sorted information.

2. Opportunity Assessment Matrix (NOAA – VAM & OAM.xlsx)

   This calculator follows the same format as a Vulnerability Assessment Matrix, only this time using it to rank purchasing or acquisition requests. It also uses three elements to generate an overall score. The goal of using this tool is to rank those purchasing decisions using a systematic approach. As with the Vulnerability Assessment Matrix, the final portion of this tool sorts the results from high to low and produces a graph with the sorted information.

3. Occupancy Vulnerability Assessment Profile (OVAP.xlsx)

   This tool is based on the R.H.A.V.E. Program produced in the late 1990s by the U.S. Fire Administration. The elements used by the tool look at the building, life safety element, risk component, including frequency and likelihood of the water demand, and finally the property value to the community. A score is generated in a building classified into one of four risk categories. A database and mapping element are included.

4. SHUP (SHUP Form.xlsx)

   This is the simplest of any of the calculators. It simply looks at the size of the structure, the height of the structure, the use of the structure, and the probability of an event occurring. A score is generated to allow comparison between other buildings classified using this tool. A database is created using this tool.

5. SPEED (SPEED Form.xlsx)

   This calculator is based on the STAPLE/E process used in community predisaster planning. It looks at the vulnerability rating of the building from five different factors. These factors include social or community values, political planning level, economic loss to the community, and environmental damage potential in danger or damage potential. Then the tool looks at a risk rating using the probability of an occurrence. Those two elements will generate a score which again can be used as a comparison tool between other buildings that have been classified. A database is created using this tool.

6. RAFER (RAFER Form.xlsx and RAFER Form Expanded.xlsx)

   This calculator uses eight elements which will generate a score and classify occupancy. The elements include: life hazard, community impact, hazard index, water supply, building usage, building construction, number of stories, and the square footage of the building. In addition to the RAFER Form, there is an expanded version that allows entry of up to 25 occupancies as well as a database development function for import or export to mapping applications.

UNIT 4: PART A
GEOGRAPHIC INFORMATION TECHNOLOGIES

TERMINAL OBJECTIVES

The students will be able to:

4A.1 Assemble a risk assessment, deployment analysis and performance measurement to support a Standard of Cover (SOC) assessment.

4A.2 Given National Fire Incident Reporting System (NFIRS) and data analysis software, evaluate an organization's abilities to mitigate and respond to risk (Plan, Mitigate, Respond and Recover).

4A.3 Integrate the United States National Grid (USNG) System with a given maps coordinate and landmark system.

ENABLING OBJECTIVES

The students will be able to:

4A.1 Articulate the characteristics of a geographic information system (GIS), Remote Sensing and GPS methods for application in conducting an SOC for a given jurisdiction.

4A.2 Explain basic analysis techniques.

4A.3 Convert GPS points for use in GIS.

4A.4 Translate NFIRS data for use in GIS.

4A.5 Given a feature on a scaled map product, determine the USNG coordinate.

4A.6 Given a USNG coordinate, distinguish the location of features on a scaled map product.

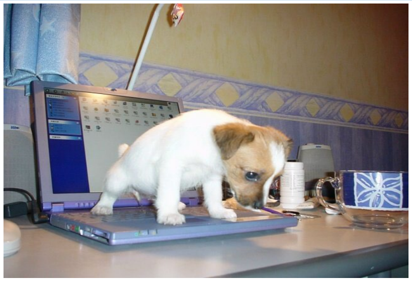
This page intentionally left blank.

FEMA

U.S. Fire Administration

UNIT 4: PART A
GEOGRAPHIC INFORMATION TECHNOLOGIES

IF AT FIRST YOU DON'T SUCCEED…

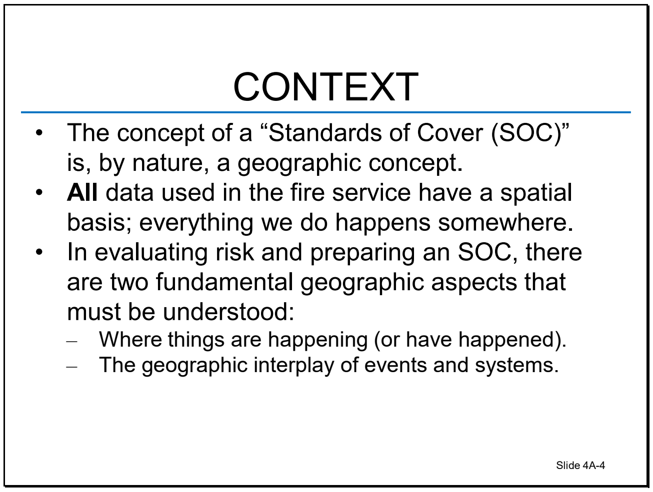


Computer hardware and software packages break, malfunction and are otherwise cantankerous. Gremlins invade. Be patient. Ask for help before hurling keyboards.

ENABLING OBJECTIVES

- Articulate the characteristics of a geographic information system (GIS), Remote Sensing and GPS methods for application in conducting a Standard of Cover (SOC) for a given jurisdiction.
- Explain basic analysis techniques.
- Convert GPS points for use in GIS.
- Translate National Fire Incident Reporting System (NFIRS) data for use in GIS.
- Given a feature on a scaled map product, determine the United States National Grid (USNG) coordinate.
- Given a USNG coordinate, distinguish the location of features on a scaled map product.

I. UNIT INTRODUCTION



CONTEXT

- The concept of a “Standards of Cover (SOC)” is, by nature, a geographic concept.
- **All** data used in the fire service have a spatial basis; everything we do happens somewhere.
- In evaluating risk and preparing an SOC, there are two fundamental geographic aspects that must be understood:
  - Where things are happening (or have happened).
  - The geographic interplay of events and systems.

A. Context.

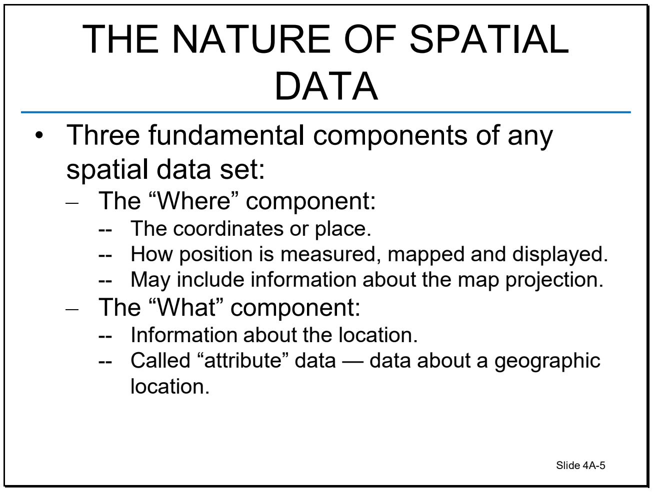
1. The concept of a “Standards of Cover (SOC)” is, by nature, a geographic concept.

2. **All** data used in the fire service have a spatial basis; everything we do happens somewhere.

3. In evaluating risk and preparing an SOC, there are two fundamental geographic aspects that must be understood:

   a. Where things are happening (or have happened).

   b. The geographic interplay of events and systems.

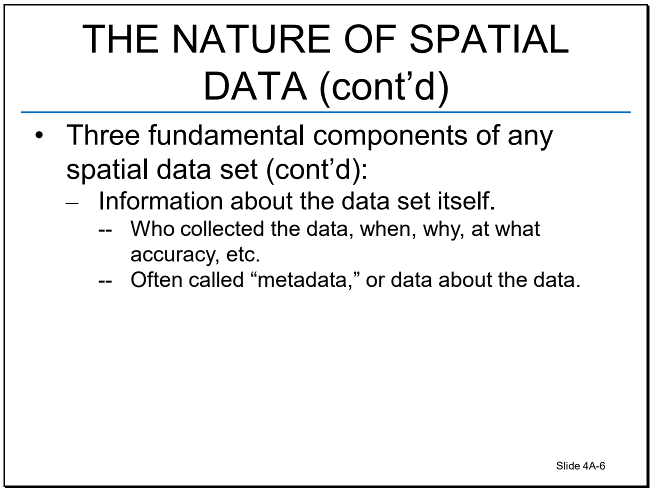


THE NATURE OF SPATIAL DATA

- Three fundamental components of any spatial data set:
  - The “Where” component:
    - The coordinates or place.
    - How position is measured, mapped and displayed.
    - May include information about the map projection.
  - The “What” component:
    - Information about the location.
    - Called “attribute” data — data about a geographic location.

B. The nature of spatial data.

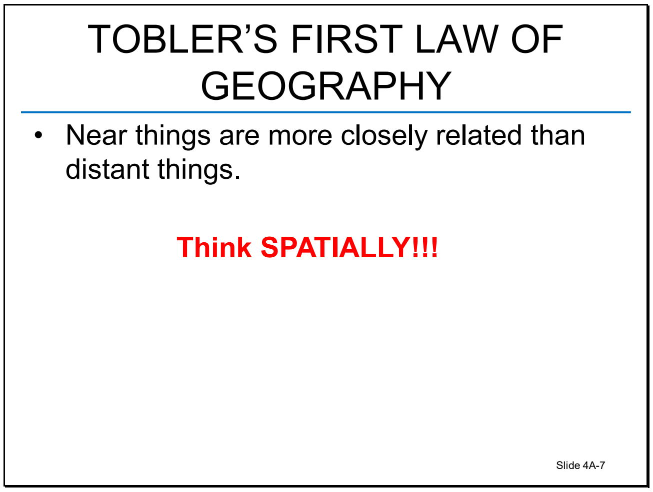
The nature of spatial data consists of three fundamental components (see table below):



THE NATURE OF SPATIAL DATA (cont'd)

- Three fundamental components of any spatial data set (cont'd):
  - Information about the data set itself.
    - Who collected the data, when, why, at what accuracy, etc.
    - Often called "metadata," or data about the data.

Component	What Does It Contain?
Where	• The coordinates or place. • How position is measured, mapped and displayed. • May include information about the map projection.
What	• Information about the location. • Called "attribute" data (about a geographic location).
Information about data set itself (metadata)	• Who collected the data, when, why, at what level of accuracy, etc.



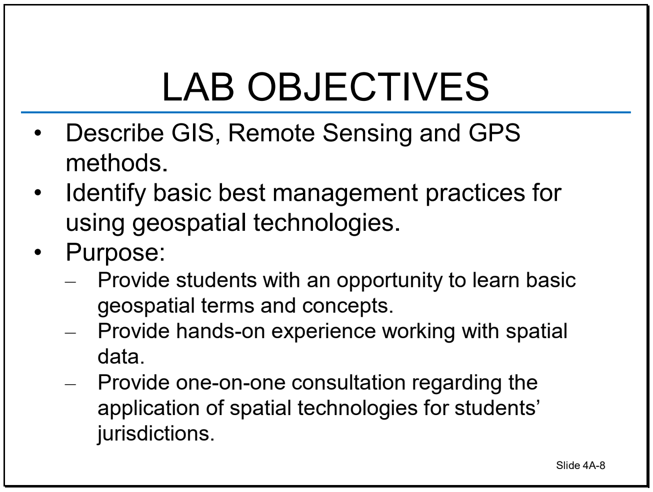
TOBLER'S FIRST LAW OF GEOGRAPHY

- Near things are more closely related than distant things.

Think SPATIALLY!!!

C. Tobler's First Law of Geography.

Near things are more closely related than distant things.



LAB OBJECTIVES

- Describe GIS, Remote Sensing and GPS methods.
- Identify basic best management practices for using geospatial technologies.
- Purpose:
  - Provide students with an opportunity to learn basic geospatial terms and concepts.
  - Provide hands-on experience working with spatial data.
  - Provide one-on-one consultation regarding the application of spatial technologies for students' jurisdictions.

Slide 4A-8

D. Lab objectives.

1. The students will be able to:

   a. Describe Geographic Information System (GIS), Remote Sensing and GPS methods.

   b. Identify basic best management practices for using geospatial technologies.

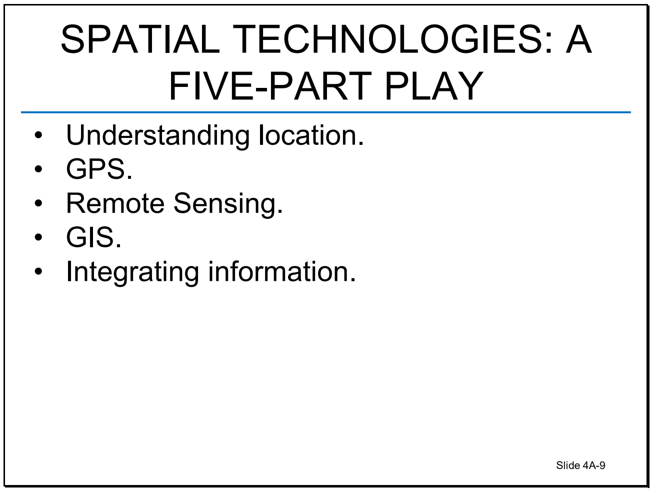
2. The purpose of this lab is to:

   a. Provide students with an opportunity to learn basic geospatial terms and concepts.

   b. Provide hands-on experience working with spatial data.

   c. Provide one-on-one consultation regarding the application of spatial technologies for students' jurisdictions.

II. SPATIAL TECHNOLOGIES INTRODUCTION



SPATIAL TECHNOLOGIES: A FIVE-PART PLAY

- Understanding location.
- GPS.
- Remote Sensing.
- GIS.
- Integrating information.

Slide 4A-9

A. Spatial technologies.

1. **Spatial technologies** is the term used to describe loosely the collective use of five allied technical approaches to understanding the relationship between spaces and places:

   a. **Understanding location.**

   b. **GPS** is used for coordinate-based data capture.

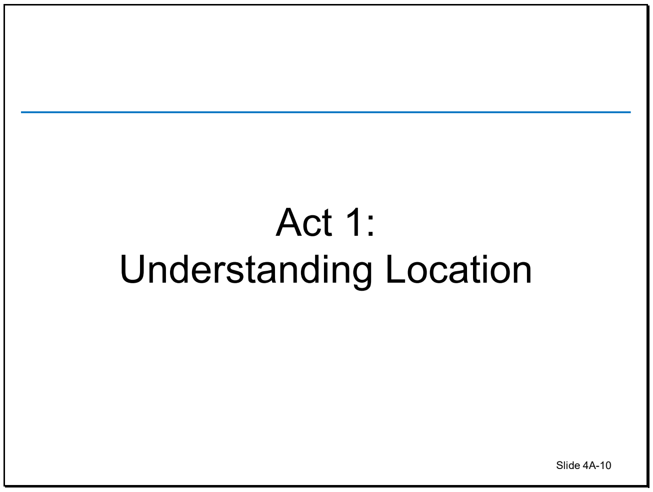
   c. **Remote Sensing** captures images using photographic equipment or sensors that measure different kinds of light.

   d. **GIS** is used to integrate GPS, Remote Sensing, and other data types for viewing, analysis and other purposes.

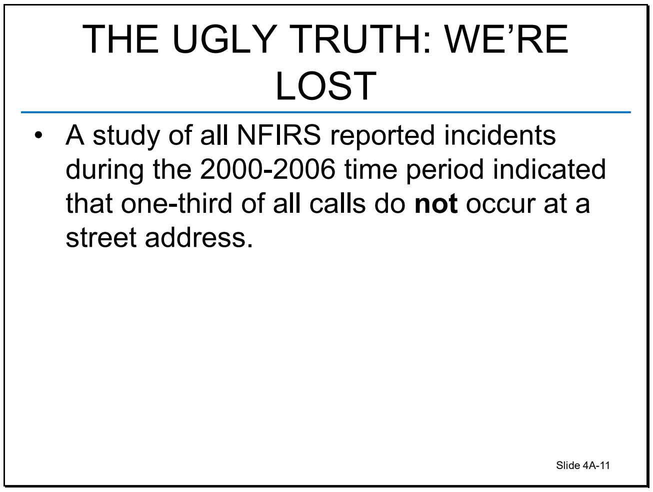
   e. **Integrating information.**

2. Each of these approaches fills a different need and produces different types of information.

III. ACT 1: UNDERSTANDING LOCATION



Act 1:
Understanding Location



THE UGLY TRUTH: WE'RE LOST

- A study of all NFIRS reported incidents during the 2000-2006 time period indicated that one-third of all calls do **not** occur at a street address.

A. The Ugly Truth: We're Lost.

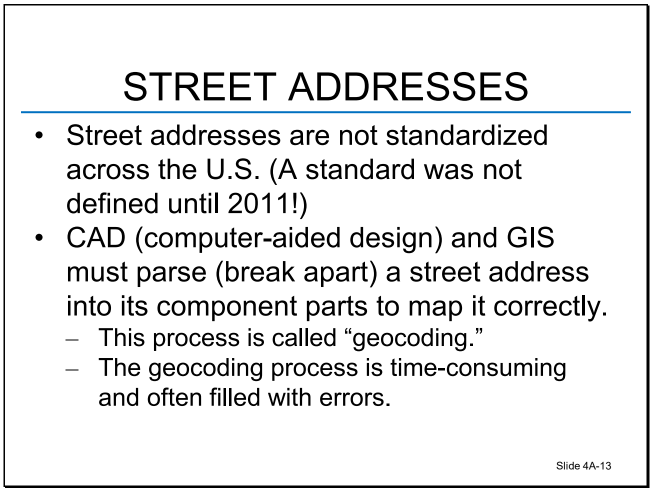
A study of all National Fire Incident Reporting System (NFIRS) reported incidents during the 2000-2006 time period indicated that one-third of all calls do **not** occur at a street address.

“THE LOST SEQUEL”

Lost fire truck

B. “The Lost Sequel.”

This video by One True Media shows an engine company that is obviously having a hard time finding a location.



STREET ADDRESSES

- Street addresses are not standardized across the U.S. (A standard was not defined until 2011!)
- CAD (computer-aided design) and GIS must parse (break apart) a street address into its component parts to map it correctly.
  - This process is called “geocoding.”
  - The geocoding process is time-consuming and often filled with errors.

C. Street Addresses.

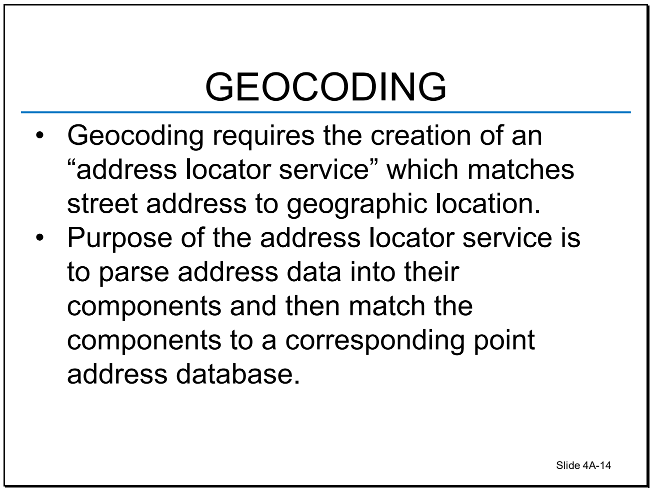
1. Street addresses are not standardized across the U.S.

2. CAD (computer-aided design) and GIS must parse (break apart) a street address into its component parts to map it correctly:

   a. This process is called “geocoding.”

   b. The geocoding process is time-consuming and often filled with errors.

3. The Federal Emergency Management Agency (FEMA) supported the adoption of the United States National Grid (USNG) as a standard of horizontal reference mapping in 2001.



GEOCODING

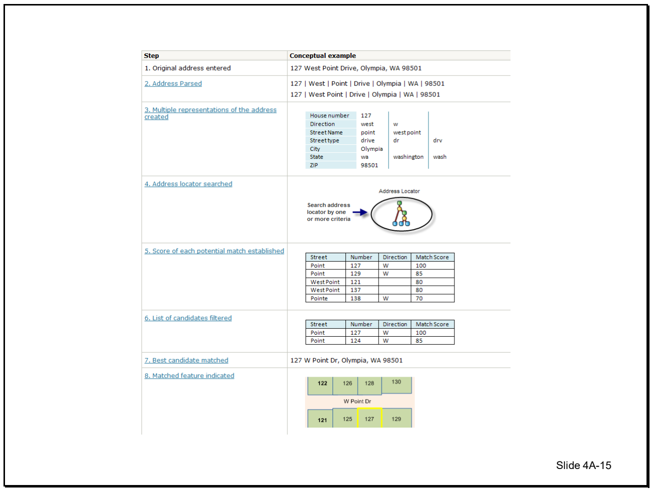
- Geocoding requires the creation of an “address locator service” which matches street address to geographic location.
- Purpose of the address locator service is to parse address data into their components and then match the components to a corresponding point address database.

D. Geocoding.

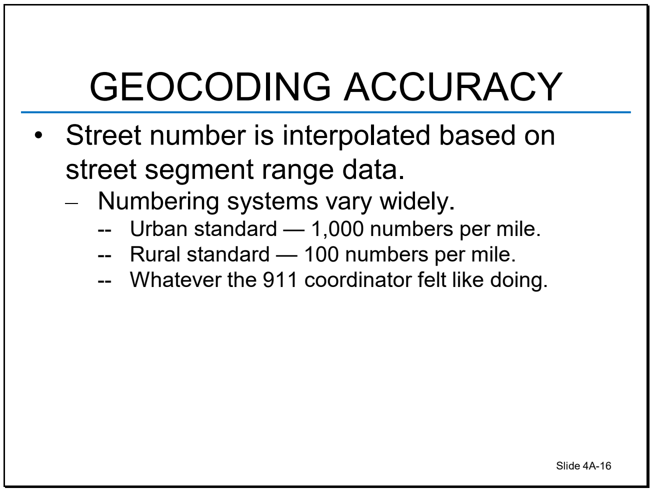
1. Geocoding, based on street address, requires the creation of an “address locator service.”

2. The purpose of the address locator service is to parse address data into their components and then match the components to a corresponding point address, street or parcel data.

3. Since NFIRS accepts such a wide variety of location description types, it may be necessary to subset call data into address description types for mapping purposes.



<table>
<tr><th>Step</th><th>Conceptual example</th></tr>
<tr><td>1. Original address entered</td><td>127 West Point Drive, Olympia, WA 98501</td></tr>
<tr><td>2. Address Parsed</td><td>127 | West | Point | Drive | Olympia | WA | 98501
127 | West Point | Drive | Olympia | WA | 98501</td></tr>
<tr><td>3. Multiple representations of the address created</td><td>
<table>
<tr><td>House number</td><td>127</td><td></td><td></td></tr>
<tr><td>Direction</td><td>west</td><td>w</td><td></td></tr>
<tr><td>Street Name</td><td>point</td><td>west point</td><td></td></tr>
<tr><td>Street type</td><td>drive</td><td>dr</td><td>drv</td></tr>
<tr><td>City</td><td>Olympia</td><td></td><td></td></tr>
<tr><td>State</td><td>wa</td><td>washington</td><td>wash</td></tr>
<tr><td>ZIP</td><td>98501</td><td></td><td></td></tr>
</table>
</td></tr>
<tr><td>4. Address locator searched</td><td>Address Locator
Search address locator by one or more criteria</td></tr>
<tr><td>5. Score of each potential match established</td><td>
<table>
<tr><th>Street</th><th>Number</th><th>Direction</th><th>Match Score</th></tr>
<tr><td>Point</td><td>127</td><td>W</td><td>100</td></tr>
<tr><td>Point</td><td>129</td><td>W</td><td>85</td></tr>
<tr><td>West Point</td><td>121</td><td></td><td>80</td></tr>
<tr><td>West Point</td><td>137</td><td></td><td>80</td></tr>
<tr><td>Pointe</td><td>138</td><td>W</td><td>70</td></tr>
</table>
</td></tr>
<tr><td>6. List of candidates filtered</td><td>
<table>
<tr><th>Street</th><th>Number</th><th>Direction</th><th>Match Score</th></tr>
<tr><td>Point</td><td>127</td><td>W</td><td>100</td></tr>
<tr><td>Point</td><td>124</td><td>W</td><td>85</td></tr>
</table>
</td></tr>
<tr><td>7. Best candidate matched</td><td>127 W Point Dr, Olympia, WA 98501</td></tr>
<tr><td>8. Matched feature indicated</td><td>122 126 128 130
W Point Dr
121 125 127 129</td></tr>
</table>



GEOCODING ACCURACY

- Street number is interpolated based on street segment range data.
  - Numbering systems vary widely.
    - Urban standard — 1,000 numbers per mile.
    - Rural standard — 100 numbers per mile.
    - Whatever the 911 coordinator felt like doing.

Slide 4A-16

E. Geocoding accuracy.

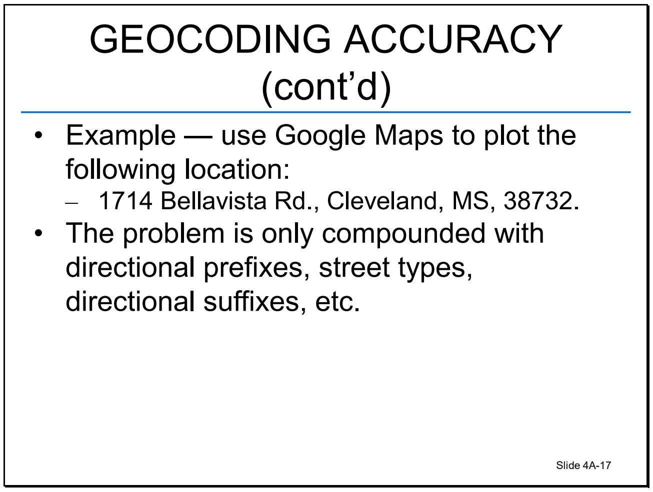
1. A street number is interpolated based on street segment range data.

2. The street numbering systems vary widely:

   a. Urban standard — 1,000 numbers per mile.

   b. Rural standard — 100 numbers per mile.

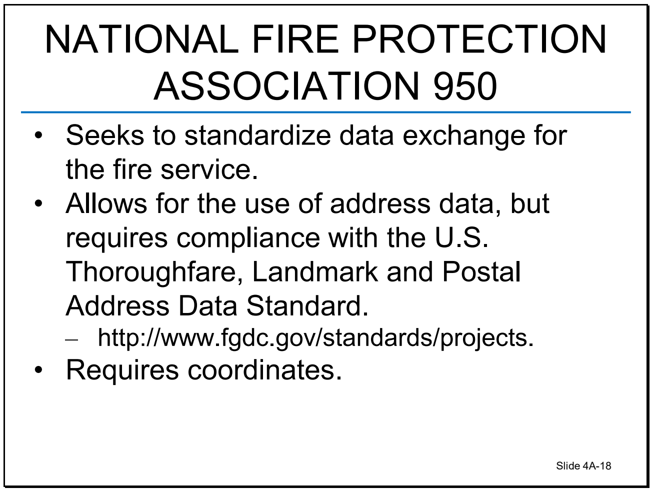
   c. Whatever the 911 coordinator felt like doing.



GEOCODING ACCURACY (cont'd)

- Example — use Google Maps to plot the following location:
  - 1714 Bellavista Rd., Cleveland, MS, 38732.
- The problem is only compounded with directional prefixes, street types, directional suffixes, etc.

Slide 4A-17



NATIONAL FIRE PROTECTION ASSOCIATION 950

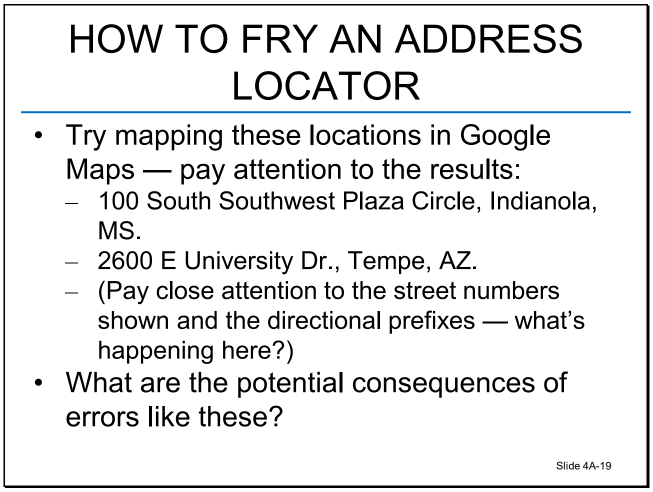
- Seeks to standardize data exchange for the fire service.
- Allows for the use of address data, but requires compliance with the U.S. Thoroughfare, Landmark and Postal Address Data Standard.
  - http://www.fgdc.gov/standards/projects.
- Requires coordinates.

F. National Fire Protection Association (NFPA) 950, *Standard for Data Development and Exchange for the Fire Service*.

1. Seeks to standardize data exchange for the fire service.

2. Allows for the use of address data, but requires compliance with the U.S. Thoroughfare, Landmark and Postal Address Data Standard.

3. Requires coordinates.



HOW TO FRY AN ADDRESS LOCATOR

- Try mapping these locations in Google Maps — pay attention to the results:
  - 100 South Southwest Plaza Circle, Indianola, MS.
  - 2600 E University Dr., Tempe, AZ.
  - (Pay close attention to the street numbers shown and the directional prefixes — what’s happening here?)
- What are the potential consequences of errors like these?

BILLY GOLDFEDDER: "CAD SAYS"

- Senior Corporal Victor Lozada escorts the Clinton motorcade to the airport and wrecks his motorcycle.
- It takes 11 minutes to get an ambulance on scene (CAD short-circuited because of limited access and the inability to describe a street address).

WHAT'S THE ADDRESS HERE?

THE BASIS OF LOCATION: COORDINATE SYSTEMS

- So many coordinate systems, so many different ways to identify locations.
- A coordinate system may be comprised of five or more components and is typically named based on its datum (base reference point).
- Types of coordinate systems:
  - Spherical (often referred to as "Geographic").
  - Grid (sometimes called "Planar").

G. Coordinate systems.

1. Coordinates are typically used to describe the location of an object on the Earth's surface.

2. A coordinate system may be comprised of up to five components. Of importance for this course are:

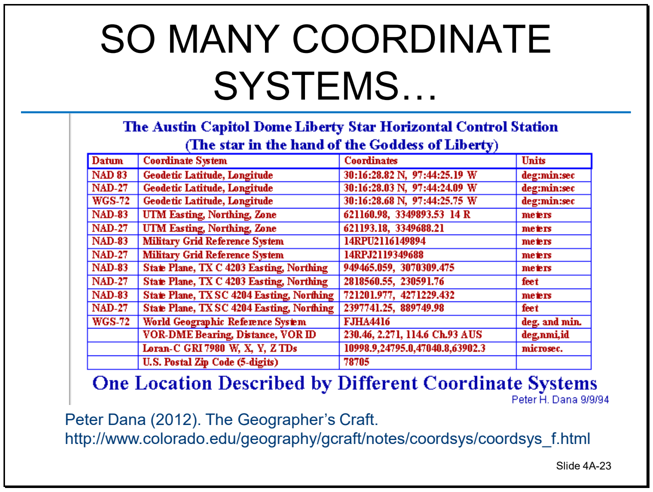
   a. Horizontal datum, which is the lateral origin to which horizontal measurements are referenced.

   b. Vertical datum, which is the base elevation to which vertical measurements are referenced.

   c. Latitudinal displacement, which is the up-down measure.

   d. Longitudinal displacement, which is the left-right measure.

   e. Elevation, which is the measure of height or altitude.



SO MANY COORDINATE SYSTEMS…

The Austin Capitol Dome Liberty Star Horizontal Control Station
(The star in the hand of the Goddess of Liberty)

Datum	Coordinate System	Coordinates	Units
NAD 83	Geodetic Latitude, Longitude	30:16:28.82 N, 97:44:25.19 W	deg:min:sec
NAD-27	Geodetic Latitude, Longitude	30:16:28.03 N, 97:44:24.09 W	deg:min:sec
WGS-72	Geodetic Latitude, Longitude	30:16:28.68 N, 97:44:25.75 W	deg:min:sec
NAD-83	UTM Easting, Northing, Zone	621160.98, 3349893.53 14 R	meters
NAD-27	UTM Easting, Northing, Zone	621193.18, 3349688.21	meters
NAD-83	Military Grid Reference System	14RPU2116149894	meters
NAD-27	Military Grid Reference System	14RPJ2119349688	meters
NAD-83	State Plane, TX C 4203 Easting, Northing	949465.059, 3070309.475	meters
NAD-27	State Plane, TX C 4203 Easting, Northing	2818560.55, 230591.76	feet
NAD-83	State Plane, TX SC 4204 Easting, Northing	721201.977, 4271229.432	meters
NAD-27	State Plane, TX SC 4204 Easting, Northing	2397741.25, 889749.98	feet
WGS-72	World Geographic Reference System	FJHA4416	deg. and min.
	VOR-DME Bearing, Distance, VOR ID	230.46, 2.271, 114.6 Ch.93 AUS	deg,nmi,id
	Loran-C GRI 7980 W, X, Y, Z TDs	10998.9,24795.0,47040.8,63902.3	microsec.
	U.S. Postal Zip Code (5-digits)	78705	

One Location Described by Different Coordinate Systems

Peter H. Dana 9/9/94

Peter Dana (2012). The Geographer's Craft. http://www.colorado.edu/geography/gcraft/notes/coordsys/coordsys_f.html

Slide 4A-23

The Austin Capitol Dome Liberty Star Horizontal Control Station
(The star in the hand of the Goddess of Liberty)

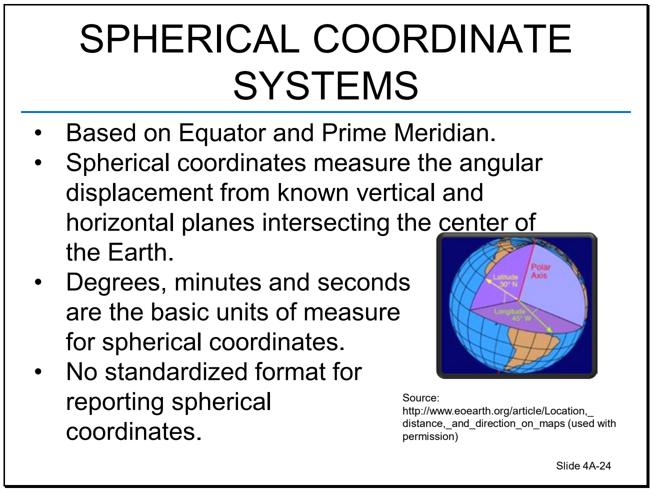
Datum	Coordinate System	Coordinates	Units
NAD 83	Geodetic Latitude, Longitude	30:16:28.82 N, 97:44:25.19 W	deg:min:sec
NAD-27	Geodetic Latitude, Longitude	30:16:28.03 N, 97:44:24.09 W	deg:min:sec
WGS-72	Geodetic Latitude, Longitude	30:16:28.68 N, 97:44:25.75 W	deg:min:sec
NAD-83	UTM Easting, Northing, Zone	621160.98, 3349893.53 14 R	meters
NAD-27	UTM Easting, Northing, Zone	621193.18, 3349688.21	meters
NAD-83	Military Grid Reference System	14RPU2116149894	meters
NAD-27	Military Grid Reference System	14RPJ2119349688	meters
NAD-83	State Plane, TX C 4203 Easting, Northing	949465.059, 3070309.475	meters
NAD-27	State Plane, TX C 4203 Easting, Northing	2818560.55, 230591.76	feet
NAD-83	State Plane, TX SC 4204 Easting, Northing	721201.977, 4271229.432	meters
NAD-27	State Plane, TX SC 4204 Easting, Northing	2397741.25, 889749.98	feet
WGS-72	World Geographic Reference System	FJHA4416	deg. and min.
	VOR-DME Bearing, Distance, VOR ID	230.46, 2.271, 114.6 Ch.93 AUS	deg,nmi,id
	Loran-C GRI 7980 W, X, Y, Z TDs	10998.9,24795.0,47040.8,63902.3	microsec.
	U.S. Postal Zip Code (5-digits)	78705	

One Location Described by Different Coordinate Systems

Peter H. Dana 9/9/94

3. There are so many coordinate systems, and each collects and displays data differently. For example, the Liberty Star on top of the Capitol Dome in Austin, Texas is described 15 different ways by eight coordinate systems.

4. There are two basic types of coordinate systems:



SPHERICAL COORDINATE SYSTEMS

- Based on Equator and Prime Meridian.
- Spherical coordinates measure the angular displacement from known vertical and horizontal planes intersecting the center of the Earth.
- Degrees, minutes and seconds are the basic units of measure for spherical coordinates.
- No standardized format for reporting spherical coordinates.

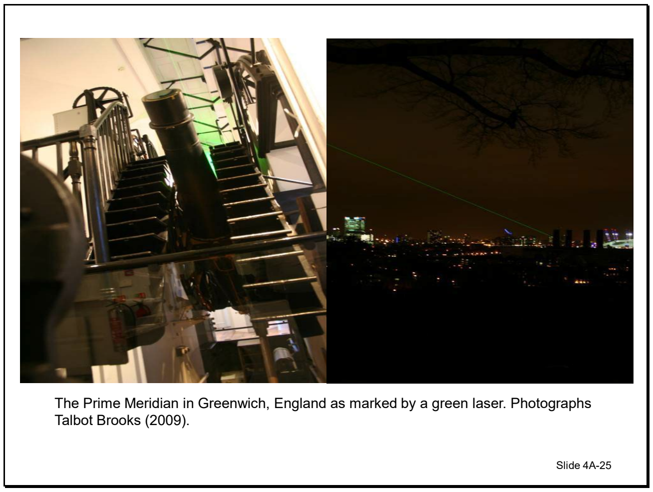
Source: http://www.eoearth.org/article/Location,_distance,_and_direction_on_maps (used with permission)

a. **Spherical** coordinates measure the angular displacement from known vertical and horizontal planes intersecting the center of the Earth. The Trooper 2 accident highlights the problems with this system (see Appendix D).

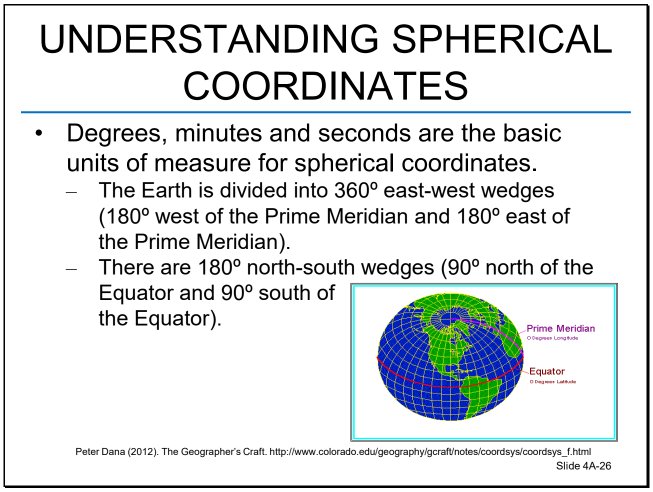
b. **Grid** coordinates use a measured distance from a known set of reference lines for measuring location.

5. What impact does a coordinate system have on fire services?

An analysis of NFIRS for calls that occurred between 2000 and 2006 found that nearly one-third of all fire calls do **not** occur at a street address.



The Prime Meridian in Greenwich, England as marked by a green laser. Photographs Talbot Brooks (2009).



UNDERSTANDING SPHERICAL COORDINATES

- Degrees, minutes and seconds are the basic units of measure for spherical coordinates.
  - The Earth is divided into 360º east-west wedges (180º west of the Prime Meridian and 180º east of the Prime Meridian).
  - There are 180º north-south wedges (90º north of the Equator and 90º south of the Equator).

Peter Dana (2012). The Geographer's Craft. http://www.colorado.edu/geography/gcraft/notes/coordsys/coordsys_f.html

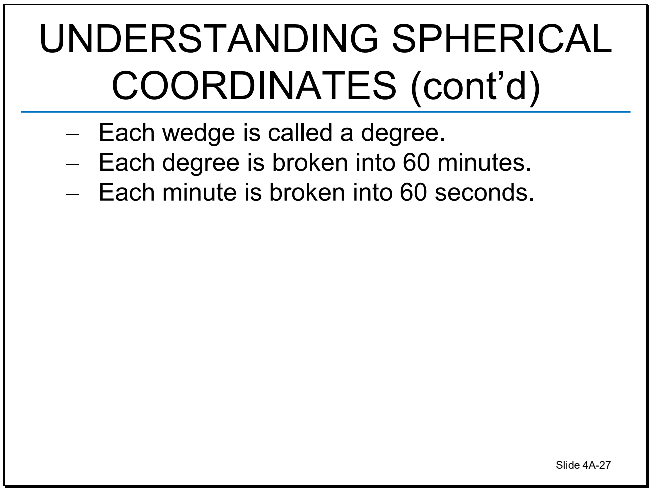
Slide 4A-26

H. Understanding spherical coordinates.

Degrees, minutes and seconds are the basic units of measure for spherical coordinates.

1. The Earth is divided into 360° east-west wedges (180° west of the Prime Meridian and 180° east of the Prime Meridian).

2. There are 180° north-south wedges (90° north of the Equator and 90° south of the Equator).



UNDERSTANDING SPHERICAL COORDINATES (cont'd)

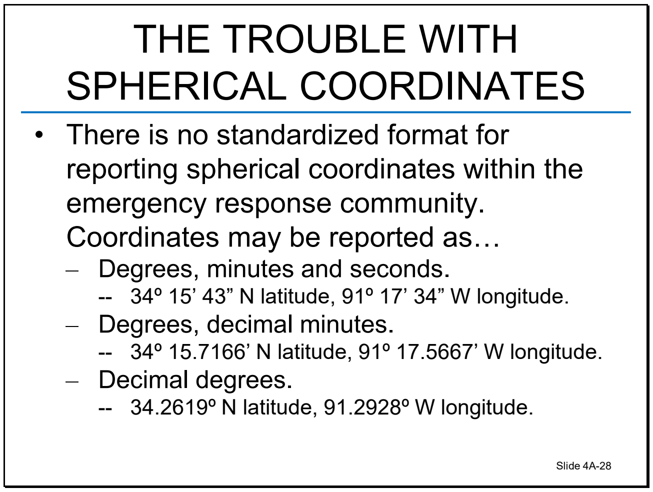
- Each wedge is called a degree.
- Each degree is broken into 60 minutes.
- Each minute is broken into 60 seconds.

Slide 4A-27

3. Each wedge is called a degree.

4. Each degree is broken into 60 minutes.

5. Each minute is broken into 60 seconds.



THE TROUBLE WITH SPHERICAL COORDINATES

- There is no standardized format for reporting spherical coordinates within the emergency response community. Coordinates may be reported as…
  - Degrees, minutes and seconds.
    - 34º 15' 43" N latitude, 91º 17' 34" W longitude.
  - Degrees, decimal minutes.
    - 34º 15.7166' N latitude, 91º 17.5667' W longitude.
  - Decimal degrees.
    - 34.2619º N latitude, 91.2928º W longitude.

I. The trouble with spherical coordinates.

1. There is no standardized format for reporting spherical coordinates within the emergency response community. Coordinates may be reported as:

   a. Degrees, minutes and seconds.

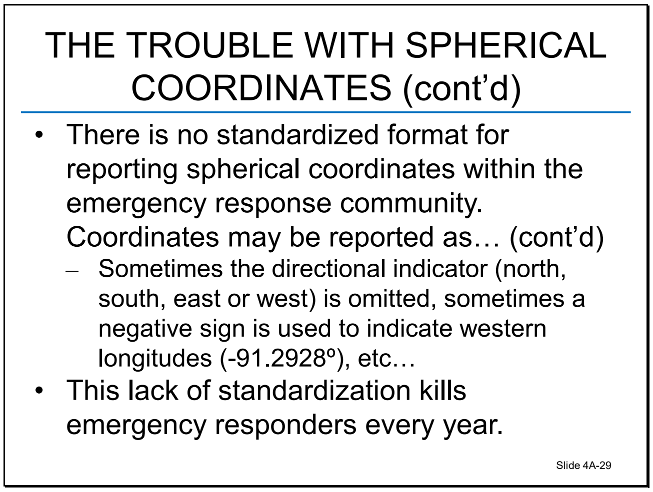
   - 34° 15' 43" N latitude, 91° 17' 34" W longitude.

   b. Degrees, decimal minutes.

   - 34° 15.7166' N latitude, 91° 17.5667' W longitude.

   c. Decimal degrees.

   - 34.2619° N latitude, 91.2928° W longitude.



THE TROUBLE WITH SPHERICAL COORDINATES (cont'd)

- There is no standardized format for reporting spherical coordinates within the emergency response community. Coordinates may be reported as… (cont'd)
  - Sometimes the directional indicator (north, south, east or west) is omitted, sometimes a negative sign is used to indicate western longitudes (-91.2928º), etc…
- This lack of standardization kills emergency responders every year.

d. Sometimes the directional indicator (north, south, east or west) is omitted, sometimes a negative sign is used to indicate western longitudes (-91.2928°), etc.

2. This lack of standardization kills emergency responders every year.

J. Example of problems with spherical coordinates.

MARYLAND TROOPER 2

Slide 4A-30

Slide 4A-31

1. On Sept. 28, 2008, a Maryland State Police rescue helicopter, known as Trooper 2, crashed. This crash highlights the problems with locating the crash site based on location details. This section provides details from two sources: The Washington Post and the National Transportation Safety Board (NTSB), both of which have facts about the case.

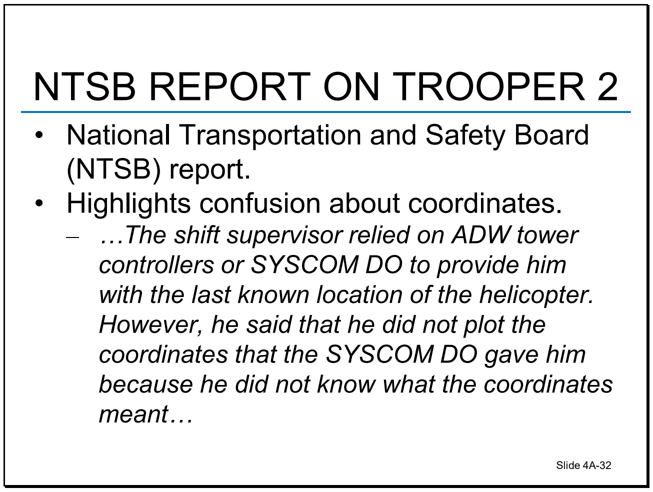
2. Radar contact was lost with Trooper 2 at 2359. Authorities located the crash site at 0158 (almost two hours later).

3. It took until 0158 to find the crash site because of a misunderstanding about latitude and longitude.

4. **The Washington Post Story.**

   a. The Washington Post prepared an outstanding video about the crash of Trooper 2. It is available at http://www.washingtonpost.com/wp-dyn/content/video/2009/08/21/VI2009082102795.html and presents the facts associated with the crash. As with most such fatal accidents, a cascading series of errors led to a catastrophic end and the death of three emergency responders and one patient. This video presents the facts about the mission and crash in a fair and balanced way, but fails to explain why it took emergency responders two hours to find and arrive at the crash scene. Imagine for a moment being trapped in the wreckage of a helicopter crash, soaked in aviation fuel, in a light drizzle during a dark night when the temperature is just a few degrees above freezing.

   b. Keep in mind that this crash occurred in a major metropolitan area and that this medical helicopter system is responsible for some very important people and places, including Camp David.



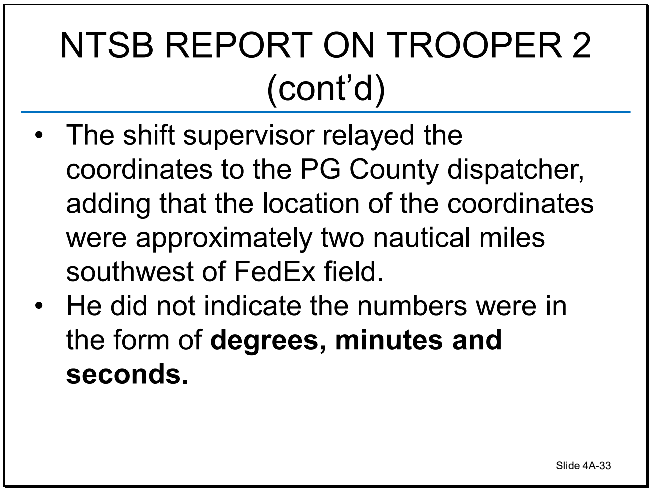
NTSB REPORT ON TROOPER 2

- National Transportation and Safety Board (NTSB) report.
- Highlights confusion about coordinates.
  - *…The shift supervisor relied on ADW tower controllers or SYSCOM DO to provide him with the last known location of the helicopter. However, he said that he did not plot the coordinates that the SYSCOM DO gave him because he did not know what the coordinates meant…*

5. **NTSB Report.**

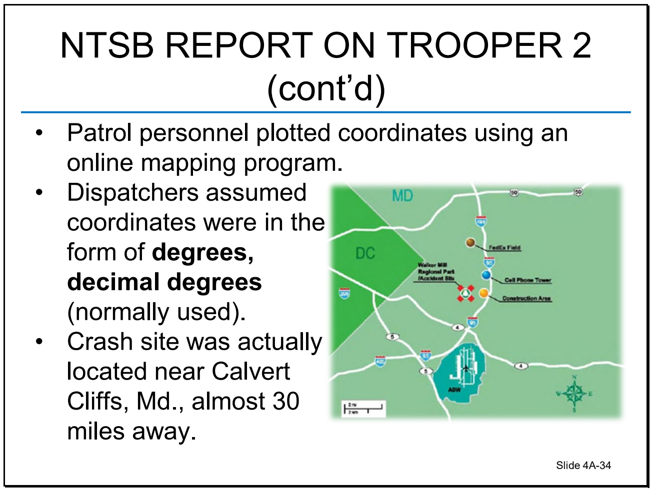
   a. The reason for the length of time required to find the scene becomes apparent in the report published by the NTSB on the incident. The slide above and the next one provide an excerpt from the NTSB report. The following passages are excerpts from the NTSB report:

b. …The shift supervisor relied on ADW tower controllers or SYSCOM to provide him with the last known location of the helicopter. However, he said that he did not plot the coordinates that the SYSCOM DO gave him because he did not know what the coordinates meant. About 00:21, the DO provided PG County dispatchers with Trooper 2's last ADS-B coordinates by reading a string of numbers, three eight, five two one seven, north was seven six five two six. The DO did not indicate that the numbers were in the form of degrees, minutes, seconds. The DO also added that the location of the coordinates was approximately two nautical miles southwest of FedEx Field.



NTSB REPORT ON TROOPER 2 (cont'd)

- The shift supervisor relayed the coordinates to the PG County dispatcher, adding that the location of the coordinates were approximately two nautical miles southwest of FedEx field.
- He did not indicate the numbers were in the form of **degrees, minutes and seconds.**



NTSB REPORT ON TROOPER 2 (cont'd)

- Patrol personnel plotted coordinates using an online mapping program.
- Dispatchers assumed coordinates were in the form of **degrees, decimal degrees** (normally used).
- Crash site was actually located near Calvert Cliffs, Md., almost 30 miles away.

c. **Note:** The report does not speculate as to whether or not additional lives may have been saved had rescue workers arrived on scene much faster. This incident may seem unique, but in fact it happens every day. As emergency responders, we are loathe to admit when we are lost and to face some of the realities associated with delayed response times.

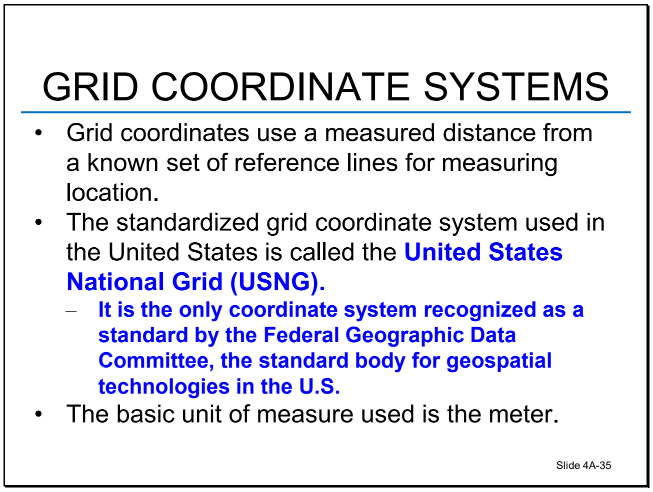
d. PG County dispatchers responded by sending patrol vehicles to the area southwest of FedEx Field. They also plotted the coordinates using an online mapping program, but the dispatchers assumed the coordinates were in the form of degrees, decimal degrees, which they were accustomed to using, so they entered the coordinates in that format. The location returned by the software program was near Calvert Cliffs, Maryland, located about 30 miles southeast of the accident site. This location raised confusion among PG County personnel and, about 00:32, a county dispatcher called SYSCOM to verify the location. An operator at SYSCOM responded, okay I don't know where the duty officer got those [coordinates]…

e. The operator did not communicate with the DO to verify the coordinates given to PG County dispatchers. **The misunderstanding about the format of the coordinates was not discovered**, and confusion about the helicopter being near Calvert Cliffs persisted as the search continued.

f. Sources:

- The Washington Post video Trooper 2's Last Flight. http://www.washingtonpost.com/wp-dyn/content/video/2009/08/21/VI2009082102795.html.

- NTSB. 2009. *Crash During Approach to Landing of Maryland State Police Aerospatiale SA365N1, N92MD, District Heights, Md., Sept. 27, 2008. Aircraft Accident Report NTSB/AAR-09/07. Washington, DC. Excerpt from pages 20-21.*



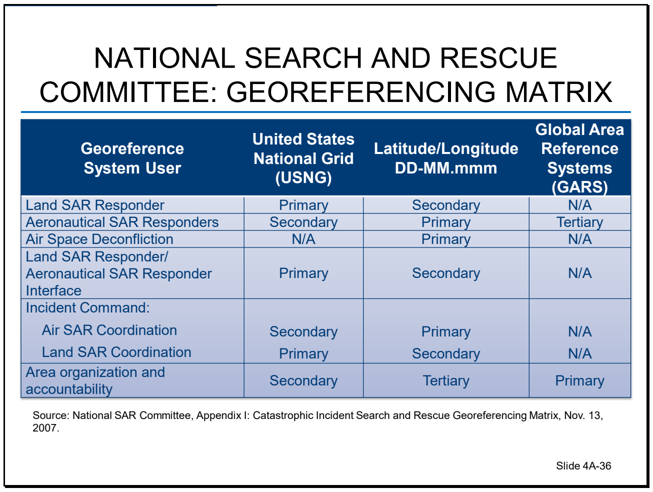
GRID COORDINATE SYSTEMS

- Grid coordinates use a measured distance from a known set of reference lines for measuring location.
- The standardized grid coordinate system used in the United States is called the **United States National Grid (USNG).**
  - **It is the only coordinate system recognized as a standard by the Federal Geographic Data Committee, the standard body for geospatial technologies in the U.S.**
- The basic unit of measure used is the meter.

K. Grid coordinate systems.

1. Grid coordinates use a measured distance from a known set of reference lines for measuring location.

2. Standardized grid coordinate system used in the United States is called the **USNG.** The basic unit of measure used is the meter.



NATIONAL SEARCH AND RESCUE COMMITTEE: GEOREFERENCING MATRIX

Georeference System User	United States National Grid (USNG)	Latitude/Longitude DD-MM.mmm	Global Area Reference Systems (GARS)
Land SAR Responder	Primary	Secondary	N/A
Aeronautical SAR Responders	Secondary	Primary	Tertiary
Air Space Deconfliction	N/A	Primary	N/A
Land SAR Responder/ Aeronautical SAR Responder Interface	Primary	Secondary	N/A
Incident Command:			
Air SAR Coordination	Secondary	Primary	N/A
Land SAR Coordination	Primary	Secondary	N/A
Area organization and accountability	Secondary	Tertiary	Primary

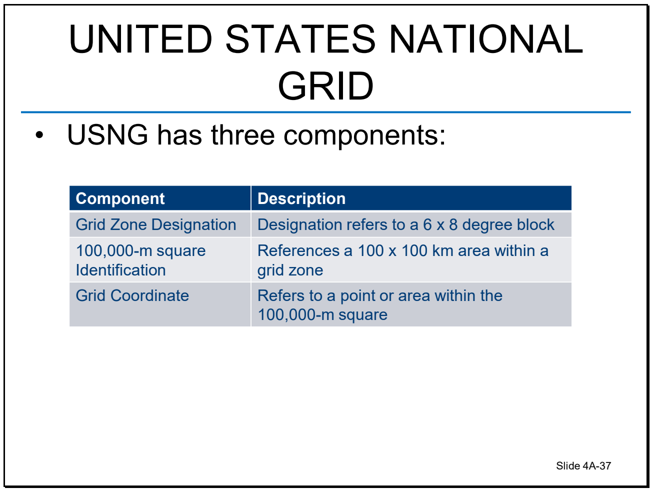
Source: National SAR Committee, Appendix I: Catastrophic Incident Search and Rescue Georeferencing Matrix, Nov. 13, 2007.

3. The National Search and Rescue Committee developed a matrix that shows how each georeference system user is handled by the:

   a. USNG.

   b. Latitude/Longitude.

   c. Global Area Reference System (GARS).



UNITED STATES NATIONAL GRID

- USNG has three components:

Component	Description
Grid Zone Designation	Designation refers to a 6 x 8 degree block
100,000-m square Identification	References a 100 x 100 km area within a grid zone
Grid Coordinate	Refers to a point or area within the 100,000-m square

L. USNG.

Objectives.

1. Terminal Objective.

   Learn how to use the USNG to read and plot coordinates on a map.

2. Enabling Objectives.

   a. Given a feature on a scaled map product, provide the USNG coordinate.

   b. Given a USNG coordinate, locate the feature on a scaled map product.

M. Demo — Map scale and printing.

Reading the USNG.

1. Please visit:

   http://www.fgdc.gov/usng/educational-resources/index_html.

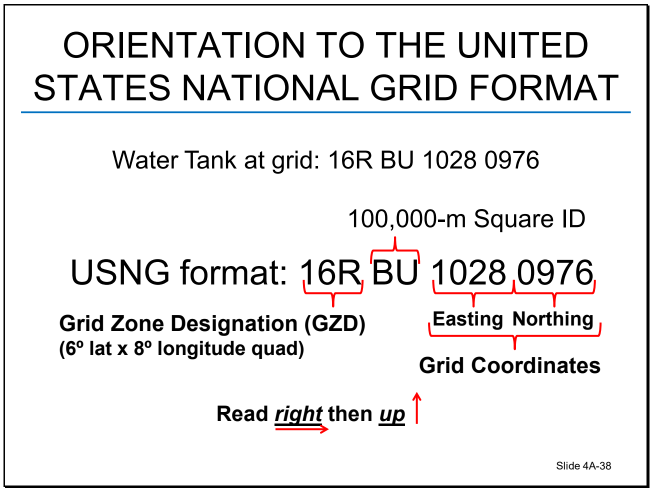
2. Download the USNG 1:24,000 training map and the USNG Grid Reader documents — you will need Adobe's Free PDF reader to view and print them.

3. Print the training map on regular paper, taking care to ensure that "Print to fit" and "Autorotate and center" options are turned **off**.

4. Print the USNG Grid Reader document on an overhead transparency and cut one of the grid readers out.

5. Digits in black along the side and top of the map are called the "principle digits." Gridlines drawn on the map show how principle digits match up top to bottom, left to right.

6. Scale — note the "grid zone designation box" tells both 100,000 meter square ID and the grid zone designation. These are important to note on any map from the National Grid because they are required components for reading a coordinate.



ORIENTATION TO THE UNITED STATES NATIONAL GRID FORMAT

Water Tank at grid: 16R BU 1028 0976

100,000-m Square ID

USNG format: 16R BU 1028 0976

Grid Zone Designation (GZD)
(6° lat x 8° longitude quad)

Easting Northing

Grid Coordinates

Read *right* then *up*

N. Reading the USNG.

1. GPS Accuracy.

a. If you purchase a GPS unit, it will only be accurate to approximately 10 meters, on a good day.

b. This chart depicts the 100 sample locations over a period of 100 minutes from a consumer GPS receiver, listed by miss distance.

- Y values represent the error (miss distance) between what the GPS receiver displayed and the true position of the receiver at horizontal control station GPS112 on the George Mason University (GMU) campus.

- Figure 38. Location samples in temporal sequence.

c. This chart depicts a temporal sequence of the 100 position samples taken at one minute intervals from a consumer GPS receiver.

- Y values represent the error (miss distance) between what the GPS receiver displayed and the true position of the receiver at horizontal control station GPS112 on the GMU campus.

- The average error was only 3.5-m, and 95 percent were within 8.2-m, an amazing capability given the cost and reliability of these consumer devices.

- The outlier excursion was out to 15-m beginning at approximately 67 minutes.

d. In other words, a 1:24,000 scale topographic map sheet matches GPS in accuracy.

e. The lessons:

  - Do not point at a position with your finger, as your finger represents a significant portion of the map.

  - Be aware of false accuracy in GPS.

  - Pay attention to detail when working.

  - A paper map on a scale of 1:24,000 matches the accuracy of a GPS device.

f. Your GPS might give you seven decimal places for latitude and longitude but these numbers are not accurate. If you stand with your GPS in one place for long enough, you will see these numbers start to drift.

2. Orientation to the USNG format.

a. Refer students to the training map.

b. USNG format: three components to a coordinate.

c. Breakdown of components of the coordinates:

  - One — 16R — Grid Zone Designation — a 6-by-8 degree box based in latitude and longitude.

  - Two — BU — within the 6-by-8 degree box are 100,000 meter square IDs.

  - Grid coordinates:

    -- Three — 1028 — “Easting” — always an even number of digits. Distance east.

    -- Four — 0976 — “Northing” — distance north.

d. Universal Transverse Mercator (UTM) format:

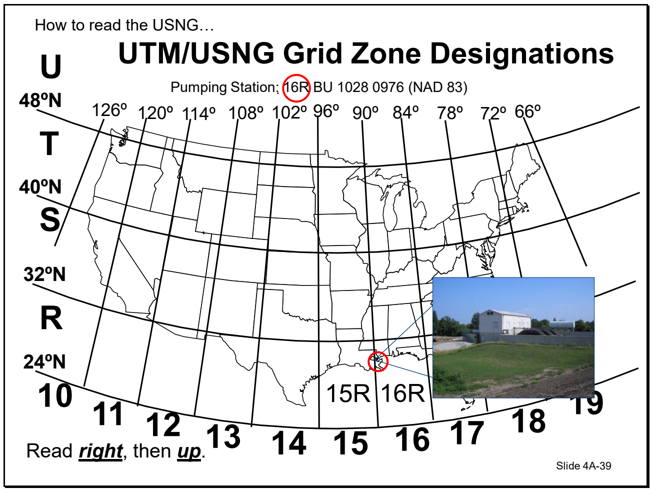
- Land search-and-rescue crews and outdoor enthusiasts are likely familiar with UTM coordinates. The equivalent UTM coordinate for the USNG coordinate shown is provided at the bottom of this slide.

3. Advantage of National Grid for emergency response, search-and-rescue, etc., is the ability to reduce total number of characters recorded or repeated.

   a. A position within 10 meter accuracy with 13 characters.

   b. Whereas UTM requires 16 characters, plus units.

   c. Latitude and longitude also requires use of many characters.



How to read the USNG…

UTM/USNG Grid Zone Designations

Pumping Station; 16R BU 1028 0976 (NAD 83)

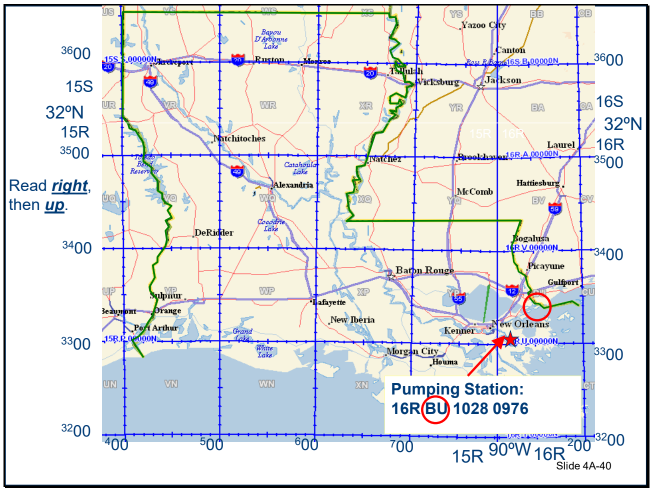
Read ***right***, then ***up***.

Slide 4A-39

O. UTM/USNG Grid Zone Designations.

Grid Zone Designation.

1. The world is further broken down into 8° bands of latitude. Combined with UTM Grid Zones, these 8° latitude by 6° longitude blocks are known as Grid Zone Designations. (**Note:** the U.S. Geological Survey (USGS) does not portray the latitude letter designation.)



Read **right**, then **up**.

Pumping Station: 16R BU 1028 0976

Slide 4A-40

2. 100,000 meter squares.

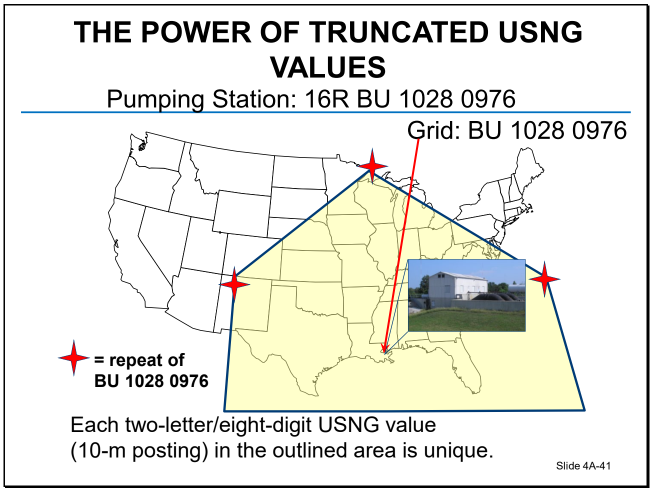
   a. Each 6-by-8 degree block/Grid Zone is further divided into 100,000 meter square blocks (100 kilometer blocks). These are the second component of the USNG coordinate.

   b. Each 100,000 meter square is given a unique two-letter combination. Since the Grid Zones are numbered and the 100,000 meter squares are lettered, now you can see why the letters “O” and “I” have been omitted from the naming scheme — “O” can be confused with zero and “I” can be confused with the number 1.

3. In the Northern Hemisphere, these start with the letter “N” at the equator and proceed consecutively northward in 8° increments (with the omission of the letters “O,” since it can be confused with zero and “I,” which can be confused with the number 1).

4. Military Grid Reference System (MGRS) Grid Zone Designation 18S ranges from 32° N to 40° N and 72° W to 78° W.

5. Indicate how to find the Pumping Station at 16R BU on this map by reading right, then up.



THE POWER OF TRUNCATED USNG VALUES

Pumping Station: 16R BU 1028 0976

Grid: BU 1028 0976

= repeat of BU 1028 0976

Each two-letter/eight-digit USNG value (10-m posting) in the outlined area is unique.

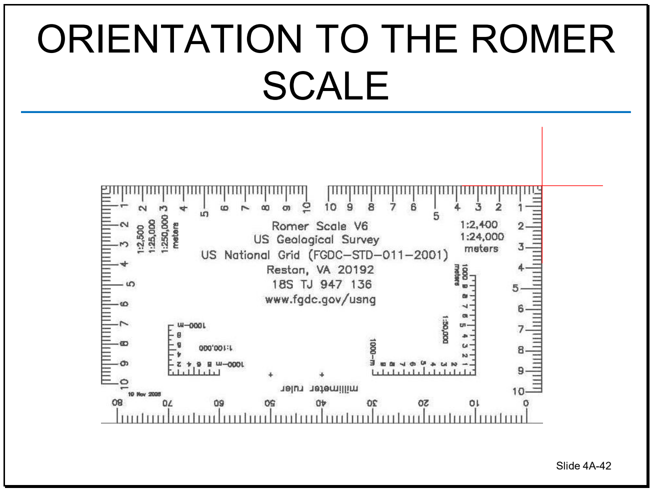
P. The power of truncated USNG values.

1. Benefits of using National Grid with 100,000 meter squares is that we can truncate coordinates. Each two-letter/eight-digit USNG value within a Grid Zone is unique.

2. Each Grid Zone encompasses a very large area. (As indicated on the map, the outlined area contains no repeating 100,000-meter squares.)

3. Washington, DC is located in UTM/ MGRS Grid Zone Designation 18S.

4. I can omit reporting the Grid Zone Designation “16R.”



ORIENTATION TO THE ROMER SCALE

Q. Reading and using the Romer scale.

1. Orientation to the Romer scale.

a. Also called a “Grid Reader.”

b. Work at several different map scales.

- Scales at upper left:
    - 1:2,500 meters.
    - 1:25,000 meters.
    - 1:250,000 meters.
- Scales at upper right:
    - 1:2,400 meters.
    - 1:24,000 meters.
- Scales at lower right:
    - 1:50,000 meters.
- Scales at lower left:
    - 1:100,000 meters.

c. The Romer scale is a good way to check if a map is printed to scale. Look at the Training Map and determine the scale at which it is drawn. Match that scale to the appropriate one on the Romer scale. Place the Romer on the map — the distance between reference lines on the map should equal the distance between zero and 10 on the Romer.

d. Place the corner of the Romer scale on the point that represents the feature for which we are determining the USNG grid coordinates. In this example, we will read the coordinates for the Water Tank for the Pumping Station.

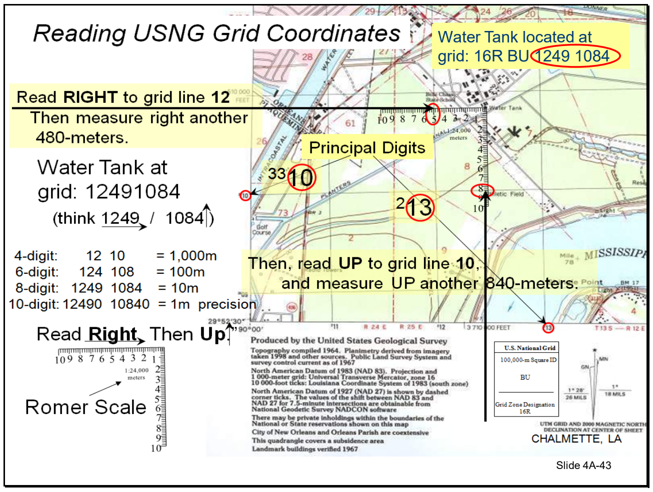
2. Use the 100,000 meter square grid designations in conjunction with the reading on the Romer scale to obtain USNG grid coordinates for a given point on the map.

3. The 100,000 meter square grid designation form the first two digits of each coordinate, locating the point somewhere within a particular 100,000 meter square.

4. For greater precision, use the Romer scale to obtain coordinates within the 100,000 meter square.

a. When the Romer scale is aligned so that the corner is aligned over our point (the Water Tank), look at where the reference line on the map intersects the Grid Reader.

b. The value at this intersection can then be added to the end of the Easting coordinate value.



Reading USNG Grid Coordinates

Water Tank located at grid: 16R BU 1249 1084

Read **RIGHT** to grid line **12**
Then measure right another 480-meters.

Principal Digits

Water Tank at grid: 12491084
(think 1249 / 1084)

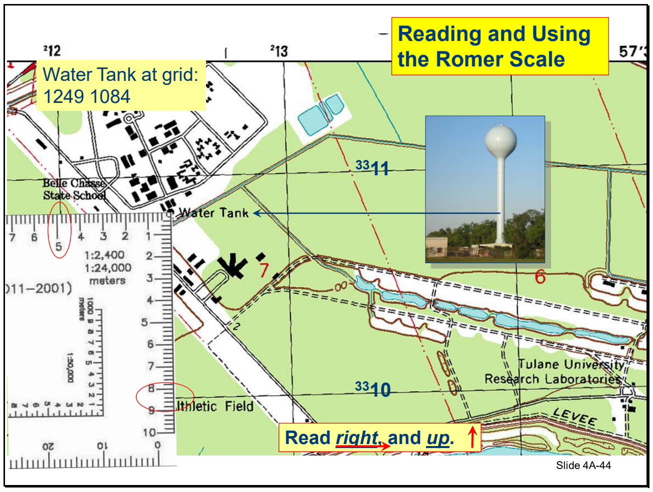
4-digit: 12 10 = 1,000m
6-digit: 124 108 = 100m
8-digit: 1249 1084 = 10m
10-digit: 12490 10840 = 1m precision

Then, read **UP** to grid line **10**, and measure UP another 840-meters.

Read **Right** Then **Up**

Romer Scale

Slide 4A-43



Reading and Using the Romer Scale

Water Tank at grid: 1249 1084

Read *right*, and *up*.

Slide 4A-44

5. Repeat this for the Northing coordinate by reading **up,** starting at the value aligned with the reference line of our 100 meter square, to the corner of the Romer scale. Using the Romer scale and the USNG Training Map, read the USNG grid coordinates of the Water Tank, located at grid: 16R BU 1249 1084.

6. Example:

   a. Read **right** to grid line 12, then measure right another 480 meters.

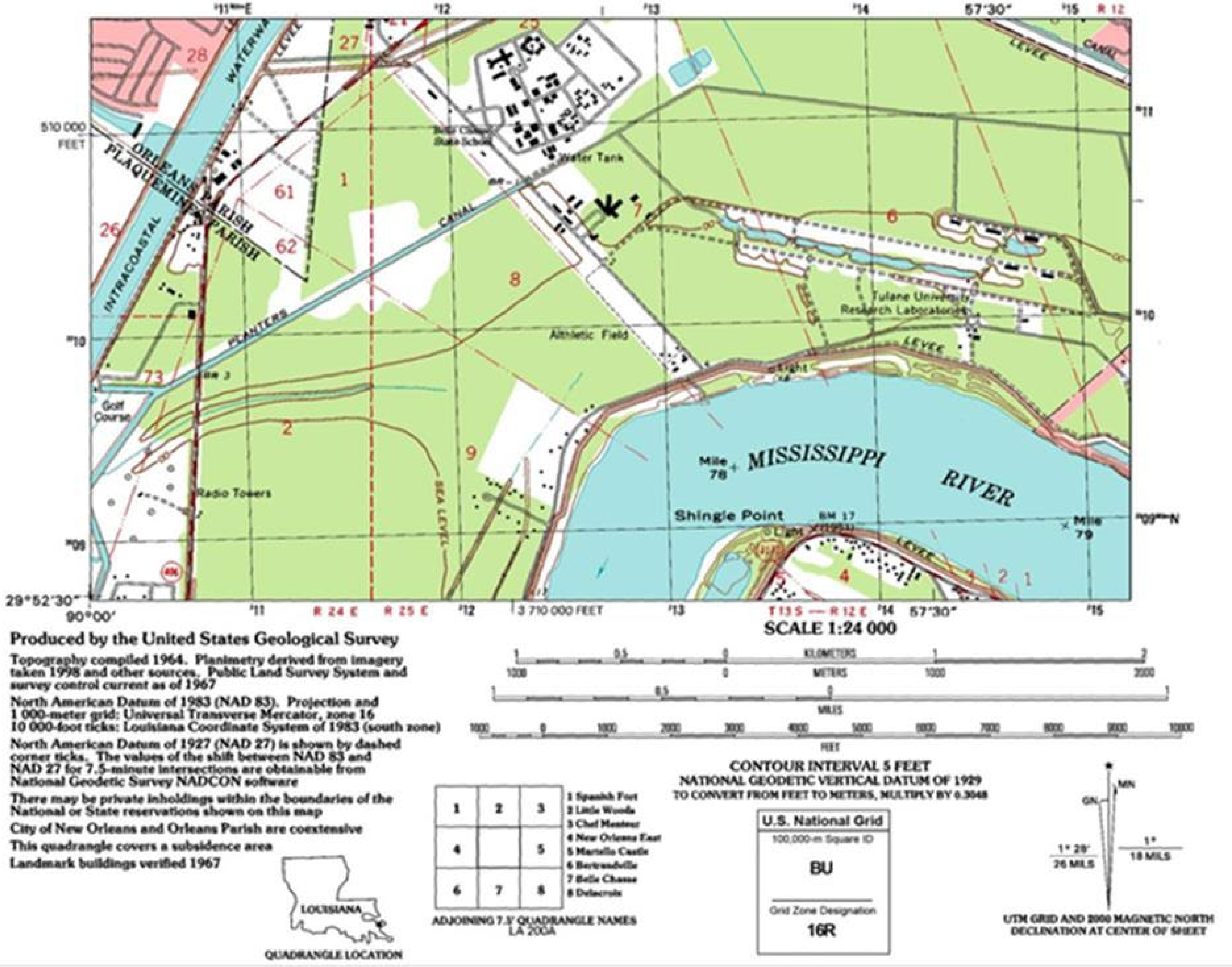
   b. Water Tank at grid: 1249 1084 (think 1249 to the right/1084 up).

c. **Note:** Precision (on the Romer scale):

- Four-digit: 12 10 = 1,000 meters precision.
- Six-digit: 124 108 = 100 meters precision.
- Eight-digit: 1249 1084 = 10 meters precision.
- Ten-digit: 12490 10840 = 1 meter precision.

d. Then read **up** to grid line 10 and measure **up** another 840-meters.

US National Grid Training Map



ACTIVITY 4A.1

Reading the U.S. National Grid

Purpose

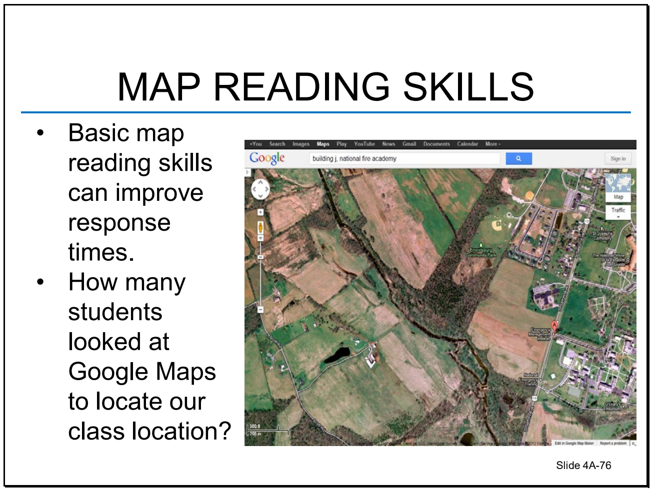
Given a scaled map, a Romer scale and a USNG coordinate, locate the feature at the given coordinates.

Directions

1. Refer to your US National Grid Training Map. Locate the feature at each of the following USNG grid coordinates.
    a. Water tank at 1249 1084.
    b. Building at 145 100.
    c. Building at 1210 1109.
    d. Building at 1215 0924.
    e. Intersection at 1425 0875.
    f. Building at 107 102.
    g. Building at 1028 0976.
2. You will be given a minute to determine the point at each of the given coordinates. Then, after each task, your instructor will explain the answer.

This page intentionally left blank.

III. ACT 1: UNDERSTANDING LOCATION (cont'd)



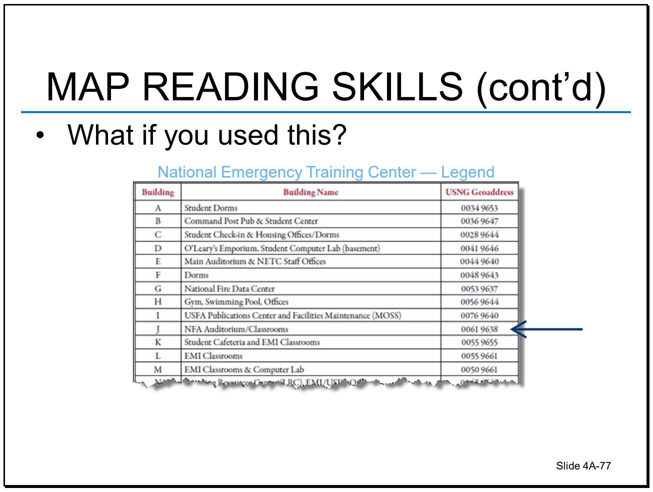
MAP READING SKILLS

- Basic map reading skills can improve response times.
- How many students looked at Google Maps to locate our class location?

Slide 4A-76

R. Map reading skills:

1. Basic map reading skills can improve response times. How many students looked at Google Maps to locate our class location?



MAP READING SKILLS (cont'd)

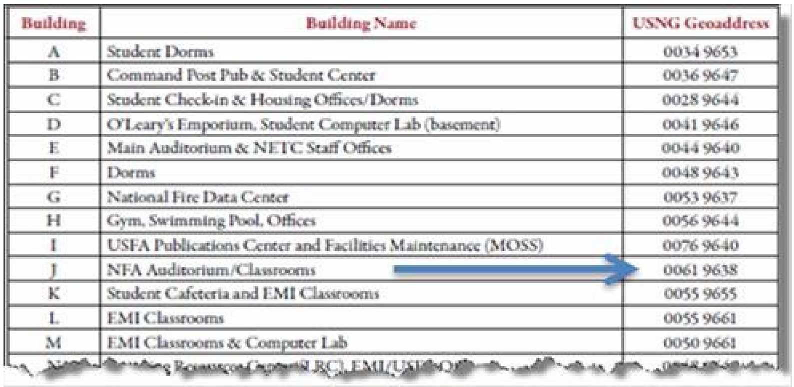
- What if you used this?

National Emergency Training Center — Legend

Building	Building Name	USNG Geoaddress
A	Student Dorms	0034 9653
B	Command Post Pub & Student Center	0036 9647
C	Student Check-in & Housing Offices/Dorms	0028 9644
D	O'Leary's Emporium, Student Computer Lab (basement)	0041 9646
E	Main Auditorium & NETC Staff Offices	0044 9640
F	Dorms	0048 9643
G	National Fire Data Center	0053 9637
H	Gym, Swimming Pool, Offices	0056 9644
I	USFA Publications Center and Facilities Maintenance (MOSS)	0076 9640
J	NFA Auditorium/Classrooms	0061 9638
K	Student Cafeteria and EMI Classrooms	0055 9655
L	EMI Classrooms	0055 9661
M	EMI Classrooms & Computer Lab	0050 9661

Slide 4A-77

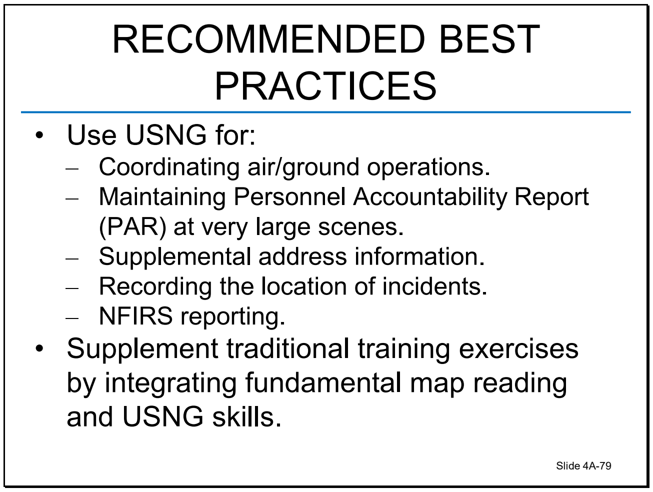
2. What if you used this?



Building	Building Name	USNG Geoaddress
A	Student Dorms	0034 9653
B	Command Post Pub & Student Center	0036 9647
C	Student Check-in & Housing Offices/Dorms	0028 9644
D	O'Leary's Emporium, Student Computer Lab (basement)	0041 9646
E	Main Auditorium & NETC Staff Offices	0044 9640
F	Dorms	0048 9643
G	National Fire Data Center	0053 9637
H	Gym, Swimming Pool, Offices	0056 9644
I	USFA Publications Center and Facilities Maintenance (MOSS)	0076 9640
J	NFA Auditorium/Classrooms	0061 9638
K	Student Cafeteria and EMI Classrooms	0055 9655
L	EMI Classrooms	0055 9661
M	EMI Classrooms & Computer Lab	0050 9661

NATIONAL EMERGENCY TRAINING CENTER

Slide 4A-78



RECOMMENDED BEST PRACTICES

- Use USNG for:
  - Coordinating air/ground operations.
  - Maintaining Personnel Accountability Report (PAR) at very large scenes.
  - Supplemental address information.
  - Recording the location of incidents.
  - NFIRS reporting.
- Supplement traditional training exercises by integrating fundamental map reading and USNG skills.

Slide 4A-79

S. Recommended best practices.

1. Recommended best practices for capturing data:

a. Develop the ability to supplement traditional address and dispatch location information with USNG coordinates.

b. Continue doing business as usual, but have the coordinate available for use in communicating location, just in case better location information is needed.

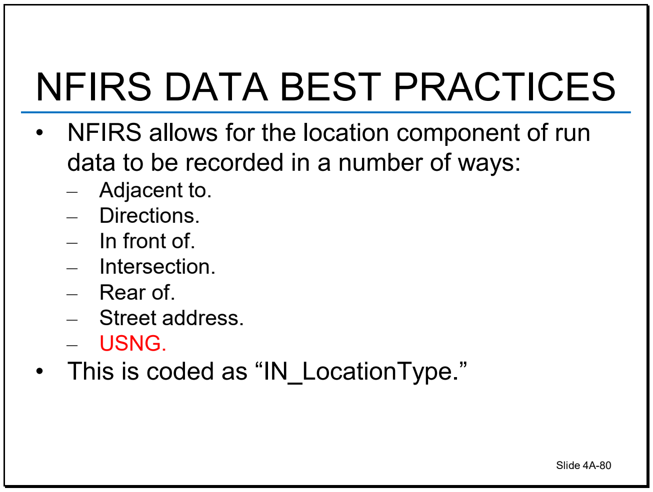
c. Record the final location of all calls for service using USNG, as this will greatly facilitate further SOC analysis practices.

d. Like hose threads, the use of coordinates in the fire service should be standardized to prevent tragedies like the Trooper 2 accident.

e. Use the USNG coordinate system:

- For use in coordinating air/ground operations.
- For maintaining Personnel Accountability Report (PAR) at very large scenes.
- As supplemental address information.
- For recording the location of incidents.
- For NFIRS reporting.

2. Assistance with implementing USNG in your community may be found through the "Contact Us" link at the bottom of http://www.fgdc.gov/ usng.



NFIRS DATA BEST PRACTICES

- NFIRS allows for the location component of run data to be recorded in a number of ways:
  - Adjacent to.
  - Directions.
  - In front of.
  - Intersection.
  - Rear of.
  - Street address.
  - USNG.
- This is coded as "IN_LocationType."

T. NFIRS data.

1. NFIRS allows for the location component of run data to be recorded in a number of ways:

a. Adjacent to.

b. Directions.

c. In front of.

d. Intersection.

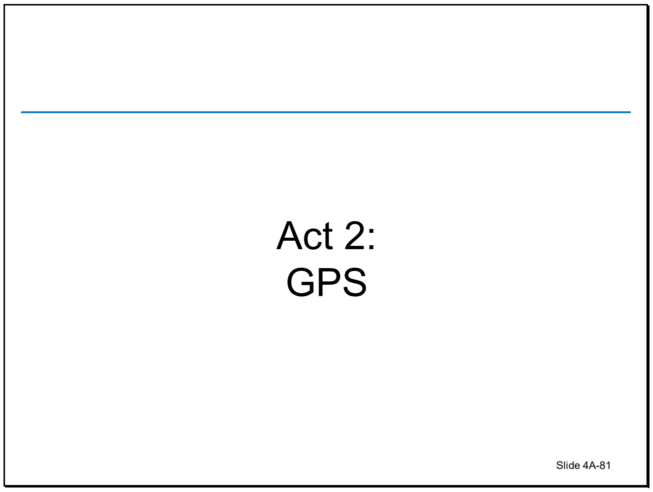
e. Rear of.

f. Street address.

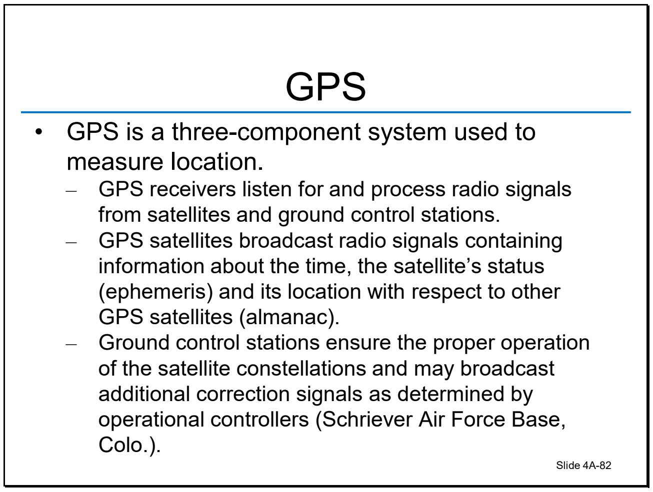
g. USNG.

2. This is coded as “IN_LocationType.”

IV. ACT 2: GPS



Act 2:
GPS



GPS

- GPS is a three-component system used to measure location.
  - GPS receivers listen for and process radio signals from satellites and ground control stations.
  - GPS satellites broadcast radio signals containing information about the time, the satellite’s status (ephemeris) and its location with respect to other GPS satellites (almanac).
  - Ground control stations ensure the proper operation of the satellite constellations and may broadcast additional correction signals as determined by operational controllers (Schriever Air Force Base, Colo.).

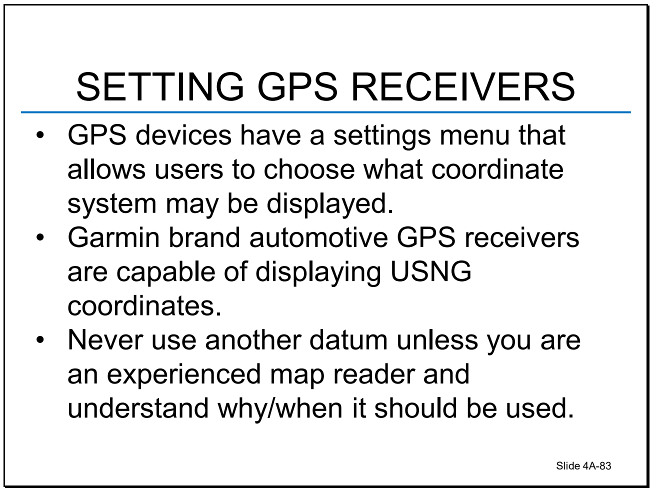
A. Global positioning skills.

1. GPS is a three-component system used to measure location:

a. GPS receivers listen for and process radio signals from satellites and ground control stations.

b. GPS satellites broadcast radio signals containing information about the time, the satellite’s status (ephemeris) and its location with respect to other GPS satellites (almanac).

c. Ground control stations ensure the proper operation of the satellite constellations and may broadcast additional correction signals as determined by operational controllers (Schriever Air Force Base, Colo.).



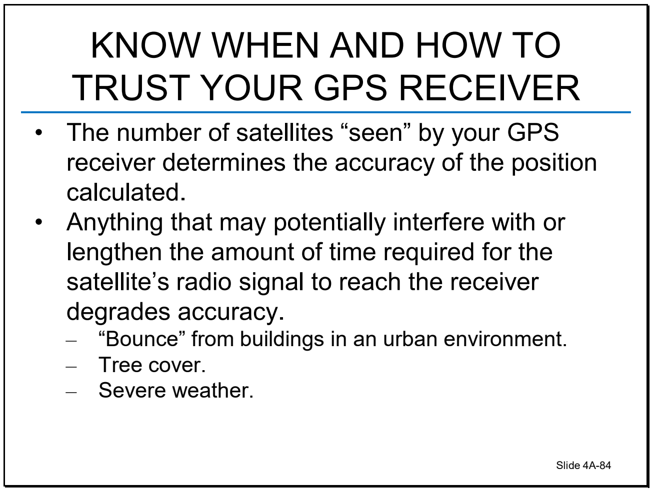
SETTING GPS RECEIVERS

- GPS devices have a settings menu that allows users to choose what coordinate system may be displayed.
- Garmin brand automotive GPS receivers are capable of displaying USNG coordinates.
- Never use another datum unless you are an experienced map reader and understand why/when it should be used.

Slide 4A-83

2. GPS devices have a settings menu that allows users to choose what coordinate system may be displayed. Garmin brand automotive GPS receivers are capable of displaying USNG coordinates.

3. Never use another datum unless you are an experienced map reader and understand why/when it should be used.

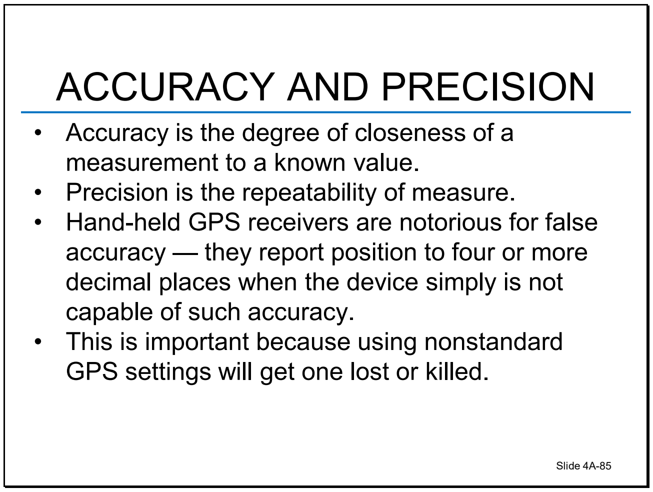


KNOW WHEN AND HOW TO TRUST YOUR GPS RECEIVER

- The number of satellites “seen” by your GPS receiver determines the accuracy of the position calculated.
- Anything that may potentially interfere with or lengthen the amount of time required for the satellite’s radio signal to reach the receiver degrades accuracy.
  - “Bounce” from buildings in an urban environment.
  - Tree cover.
  - Severe weather.

Slide 4A-84

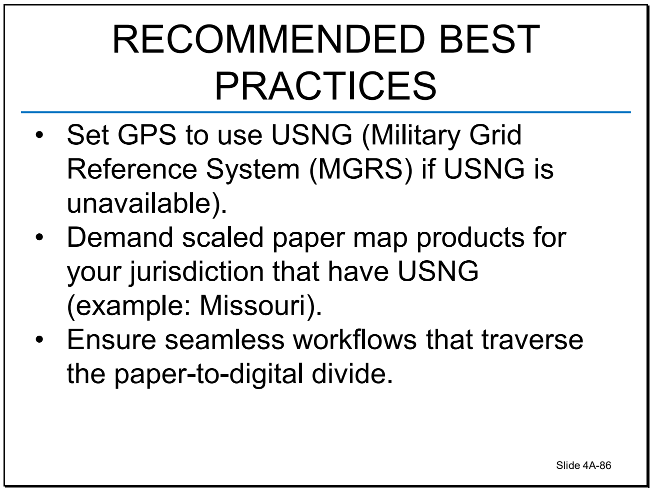
4. Knowing when and how to trust your GPS is important. The number of satellites “seen” by your GPS receiver determines the accuracy of the position calculated. Anything that may potentially interfere with or lengthen the amount of time required for the satellite’s radio signal to reach the receiver degrades accuracy.



ACCURACY AND PRECISION

- Accuracy is the degree of closeness of a measurement to a known value.
- Precision is the repeatability of measure.
- Hand-held GPS receivers are notorious for false accuracy — they report position to four or more decimal places when the device simply is not capable of such accuracy.
- This is important because using nonstandard GPS settings will get one lost or killed.

5. Accuracy is the degree of closeness of a measurement to a known value. Precision is the repeatability of measure. Hand-held GPS receivers are notorious for false accuracy — they report position to four or more decimal places when the device simply is not capable of such accuracy. This is important because using nonstandard GPS settings will get one lost or killed.



RECOMMENDED BEST PRACTICES

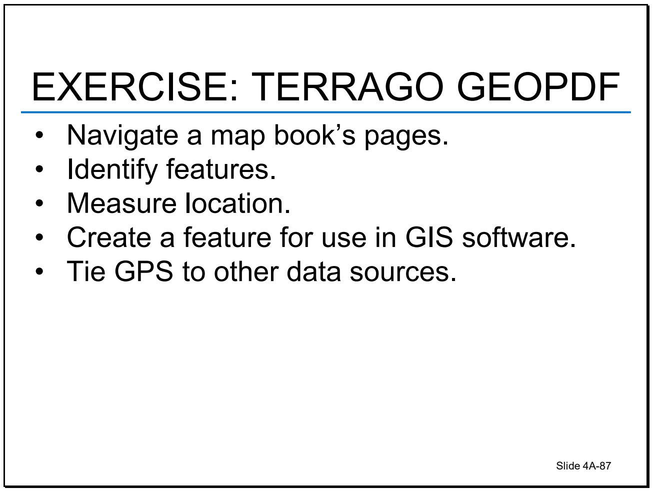
- Set GPS to use USNG (Military Grid Reference System (MGRS) if USNG is unavailable).
- Demand scaled paper map products for your jurisdiction that have USNG (example: Missouri).
- Ensure seamless workflows that traverse the paper-to-digital divide.

B. Recommended best practices.

1. Set GPS to use USNG (MGRS if USNG is unavailable).

2. Demand scaled paper map products for your jurisdiction that have USNG (example: Missouri).

3. Ensure seamless workflows that traverse the paper-to-digital divide.



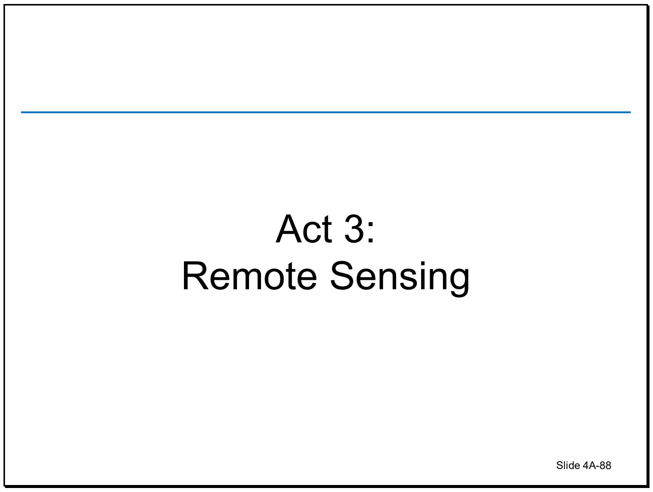
EXERCISE: TERRAGO GEOPDF

- Navigate a map book's pages.
- Identify features.
- Measure location.
- Create a feature for use in GIS software.
- Tie GPS to other data sources.

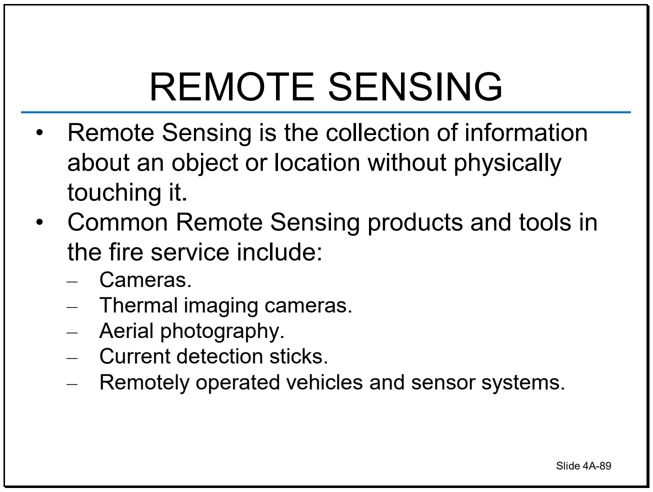
C. TerraGo GeoPDF.

The purpose of this exercise is to provide the students with practice navigating a map, identifying features, measuring locations, creating a feature for use in GIS software, and tying GPS to other data sources.

V. ACT 3: REMOTE SENSING



Act 3:
Remote Sensing



REMOTE SENSING

- Remote Sensing is the collection of information about an object or location without physically touching it.
- Common Remote Sensing products and tools in the fire service include:
  - Cameras.
  - Thermal imaging cameras.
  - Aerial photography.
  - Current detection sticks.
  - Remotely operated vehicles and sensor systems.

A. Introduction.

1. Remote Sensing is the collection of information about an object or location without physically touching it. Common Remote Sensing products and tools in the fire service include:

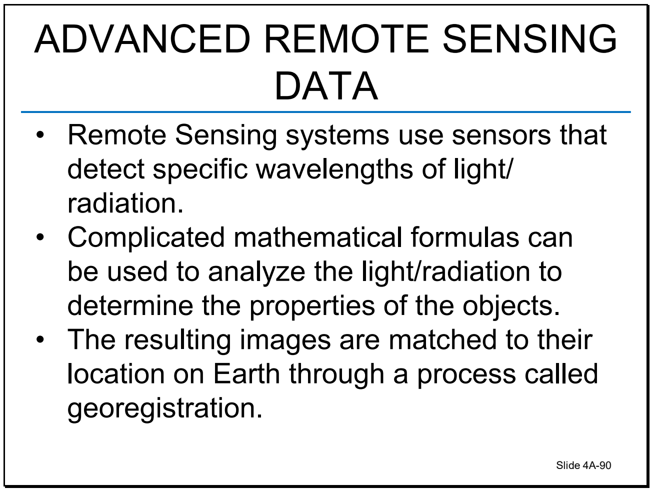
   a. Cameras.

   b. Thermal imaging cameras.

   c. Aerial photography.

   d. Current detection sticks.

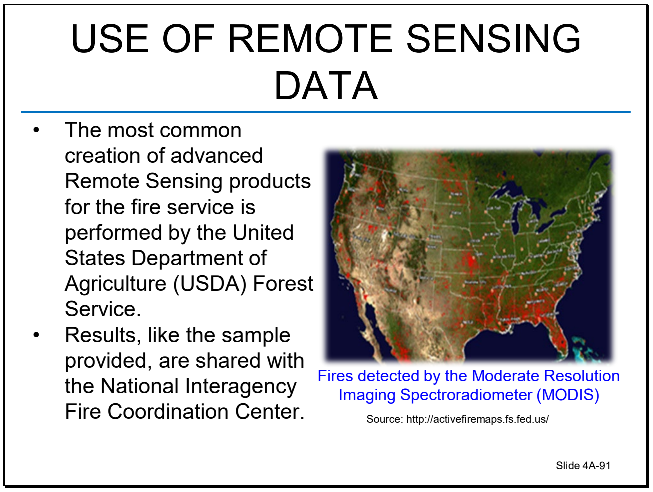
   e. Remotely operated vehicles and sensor systems.



ADVANCED REMOTE SENSING DATA

- Remote Sensing systems use sensors that detect specific wavelengths of light/ radiation.
- Complicated mathematical formulas can be used to analyze the light/radiation to determine the properties of the objects.
- The resulting images are matched to their location on Earth through a process called georegistration.

2. Remote Sensing systems use sensors that detect specific wavelengths of light/radiation. Complicated mathematical formulas can be used to analyze the light/radiation to determine the properties of the objects. The resulting images are matched to their location on Earth through a process called georegistration.



USE OF REMOTE SENSING DATA

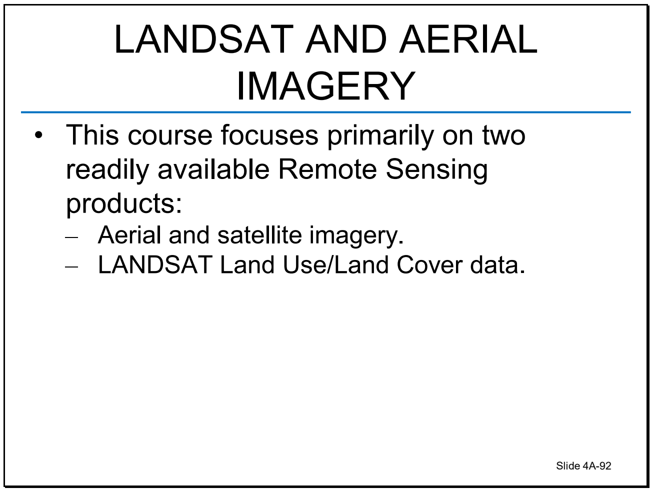
- The most common creation of advanced Remote Sensing products for the fire service is performed by the United States Department of Agriculture (USDA) Forest Service.
- Results, like the sample provided, are shared with the National Interagency Fire Coordination Center.

Fires detected by the Moderate Resolution Imaging Spectroradiometer (MODIS)

Source: http://activefiremaps.fs.fed.us/

B. Use of Remote Sensing data.

The most common creation of advanced Remote Sensing products for the fire service is performed by the U.S. Department of Agriculture (USDA) Forest Service. Results, like the sample provided, are shared with the National Interagency Fire Coordination Center.



LANDSAT AND AERIAL IMAGERY

- This course focuses primarily on two readily available Remote Sensing products:
  - Aerial and satellite imagery.
  - LANDSAT Land Use/Land Cover data.

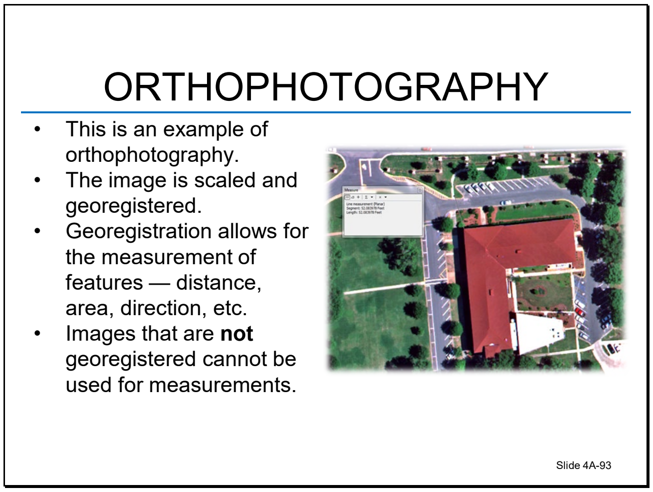
Slide 4A-92

C. LANDSAT and aerial imagery.

1. This course focuses primarily on two readily available Remote Sensing products:

   a. Aerial and satellite imagery.

   b. LANDSAT Land Use/Land Cover data.



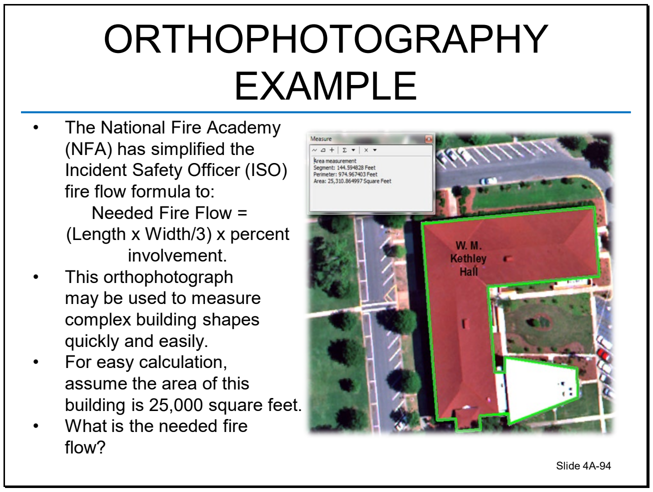
ORTHOPHOTOGRAPHY

- This is an example of orthophotography.
- The image is scaled and georegistered.
- Georegistration allows for the measurement of features — distance, area, direction, etc.
- Images that are **not** georegistered cannot be used for measurements.

Slide 4A-93

2. The image below is an example of orthophotography.

Line measurement (Planar)
Segment: 52.083978 Feet
Length: 52.083978 Feet



ORTHOPHOTOGRAPHY EXAMPLE

- The National Fire Academy (NFA) has simplified the Incident Safety Officer (ISO) fire flow formula to:
  Needed Fire Flow = (Length x Width/3) x percent involvement.
- This orthophotograph may be used to measure complex building shapes quickly and easily.
- For easy calculation, assume the area of this building is 25,000 square feet.
- What is the needed fire flow?

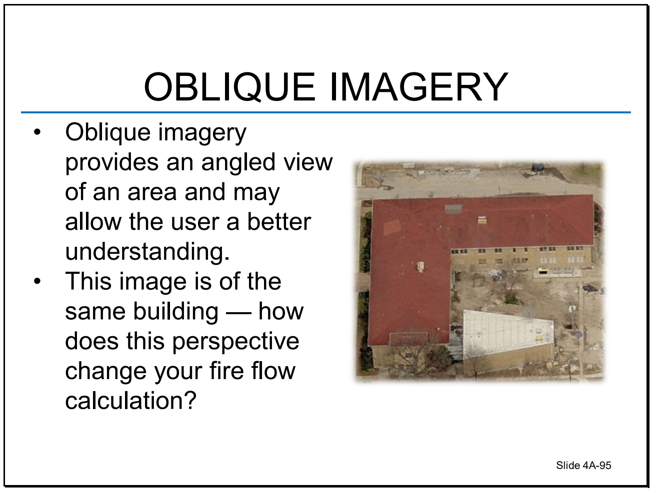
Area measurement
Segment: 144.594828 Feet
Perimeter: 974.967403 Feet
Area: 25,310.864997 Square Feet

Area measurement
Segment: 144.594828 Feet
Perimeter: 974.967403 Feet
Area: 25,310.864997 Square Feet

3. The National Fire Academy (NFA) has simplified the Incident Safety Officer (ISO) fire flow formula to:

   Needed Fire Flow = (Length x Width/3) x percent involvement.

4. This orthophotograph may be used to measure complex building shapes quickly and easily. For easy calculation, assume the area of this building is 25,000 square feet. What is the needed fire flow?

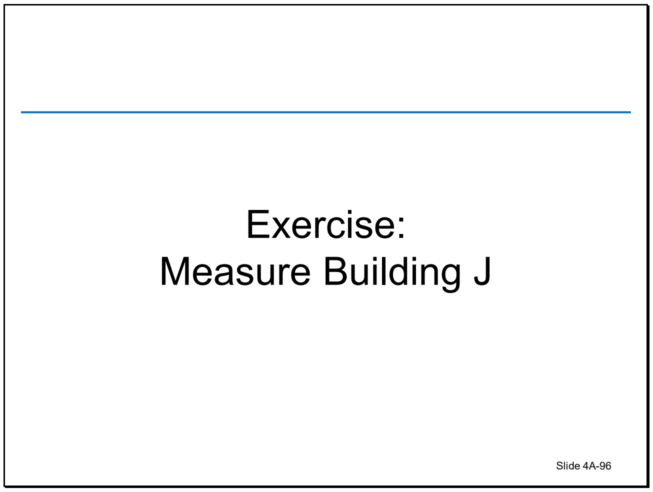


OBLIQUE IMAGERY

- Oblique imagery provides an angled view of an area and may allow the user a better understanding.
- This image is of the same building — how does this perspective change your fire flow calculation?

Slide 4A-95

5. Oblique imagery provides an angled view of an area and may allow the user a better understanding. This image is of the same building — how does this perspective change your fire flow calculation?

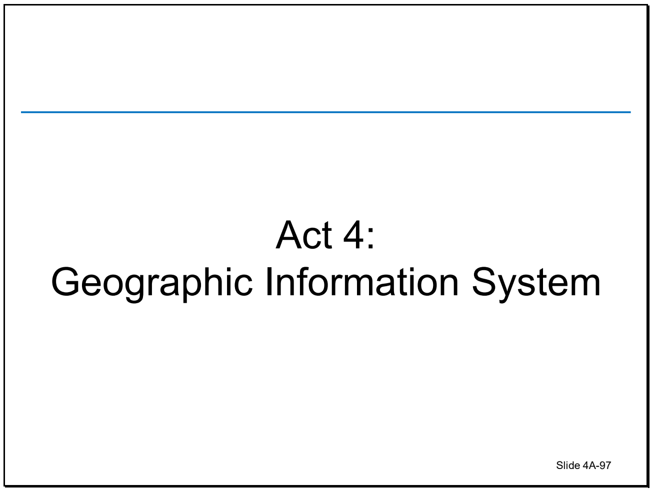


Exercise:
Measure Building J

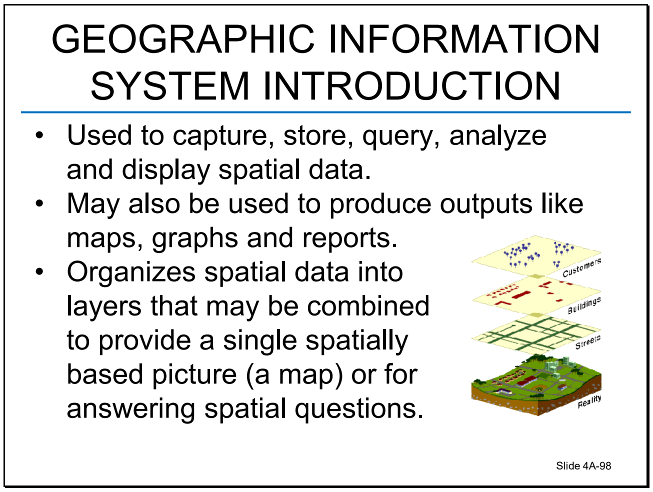
D. **Exercise: Measure Building J.**

This exercise provides you with an opportunity to use ArcMap software.

VI. ACT 4: GEOGRAPHIC INFORMATION SYSTEM



Act 4:
Geographic Information System



GEOGRAPHIC INFORMATION SYSTEM INTRODUCTION

- Used to capture, store, query, analyze and display spatial data.
- May also be used to produce outputs like maps, graphs and reports.
- Organizes spatial data into layers that may be combined to provide a single spatially based picture (a map) or for answering spatial questions.

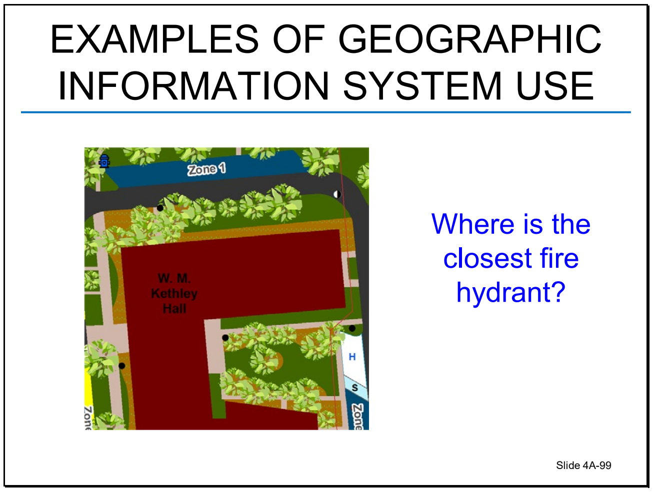
A. GIS introduction.

1. GIS is an outstanding tool for integrating geospatial information for viewing data and modeling information in a spatial context to improve planning and decision-making.

2. At their core, both the creation of an SOC and the accreditation process are plans for self-improvement.

3. By integrating spatial information within GIS, fire service organizations may establish a starting basis and develop consistent means and methods for measuring progress against their plan.

4. While the technical approach may involve an initial skills/knowledge hurdle, a larger return on investment in terms of evaluating and documenting progress will be realized.



EXAMPLES OF GEOGRAPHIC INFORMATION SYSTEM USE

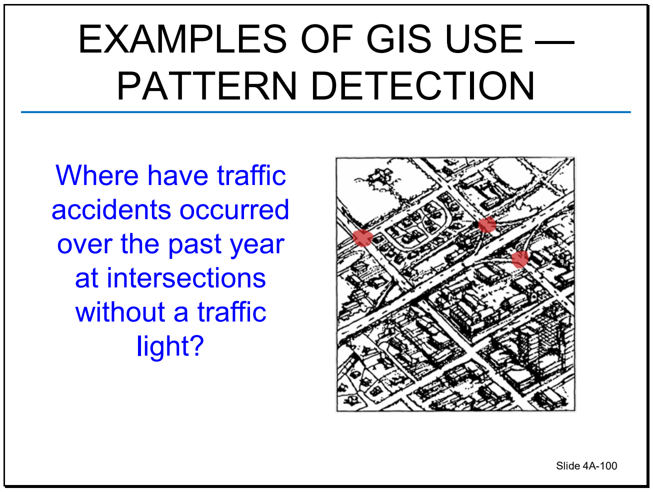
Where is the closest fire hydrant?

B. GIS as a planning tool.

1. GIS may be used as a planning tool.

2. Imagery; infrastructure data, such as roadways and their respective weight capacities; bridge clearances; available parcels; soil types; prevailing winds; population centers; and other information may be integrated within a GIS to identify optimal conditions for land use and development purposes.

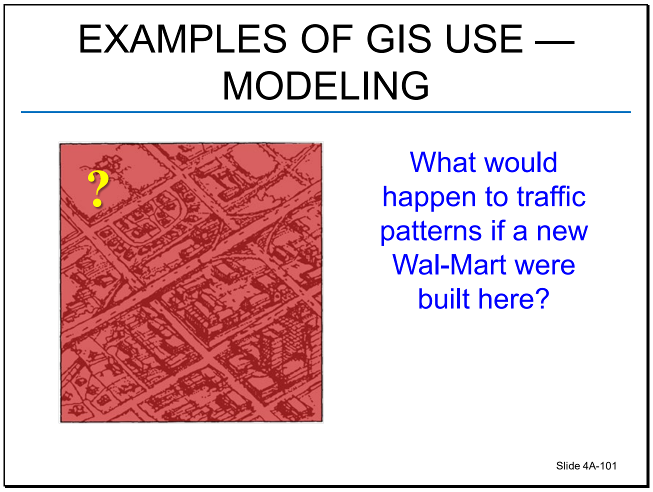
3. Interpretation of aerial and satellite imagery allows for the identification of features and their proximity to other features or hazards.



EXAMPLES OF GIS USE — PATTERN DETECTION

Where have traffic accidents occurred over the past year at intersections without a traffic light?

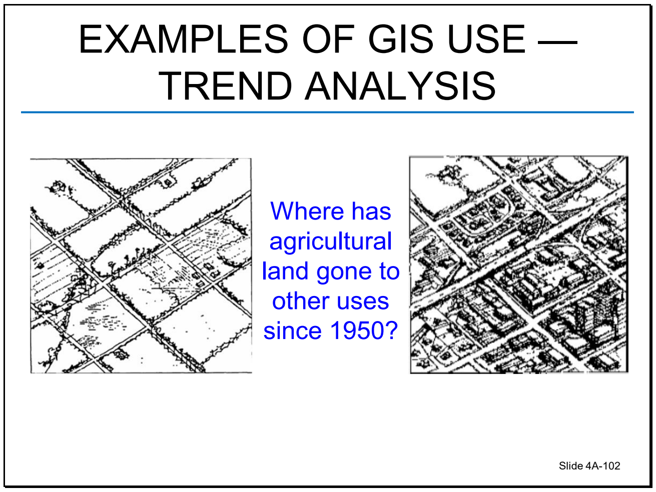
Slide 4A-100



EXAMPLES OF GIS USE — MODELING

What would happen to traffic patterns if a new Wal-Mart were built here?

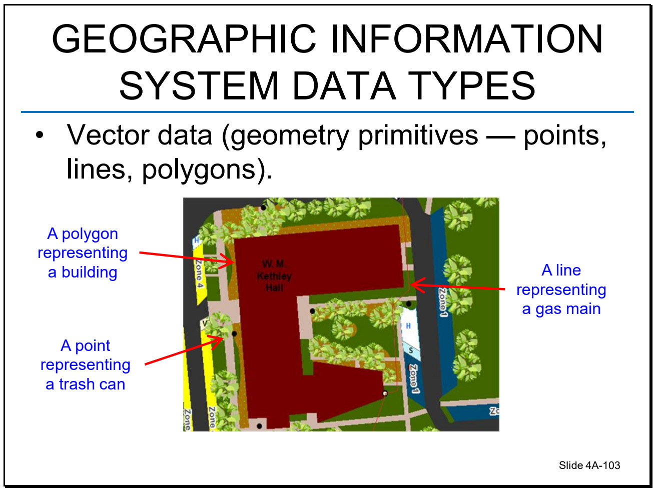
Slide 4A-101



EXAMPLES OF GIS USE — TREND ANALYSIS

Where has agricultural land gone to other uses since 1950?

Slide 4A-102



GEOGRAPHIC INFORMATION SYSTEM DATA TYPES

- Vector data (geometry primitives — points, lines, polygons).

A polygon representing a building

A point representing a trash can

A line representing a gas main

Slide 4A-103

4. GIS has two fundamental data types:

   a. Vector data, using points and lines.

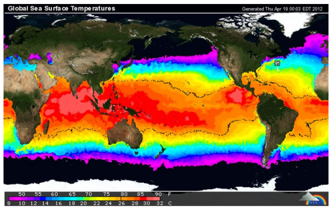
   b. Polygons, representing shapes and spatial relationships among real-world features.

C. Key terms related to GIS.

The following terms are related to GIS:

GEOGRAPHIC INFORMATION SYSTEM DATA TYPES (cont'd)

- Raster data — imagery and derived surfaces like a temperature map or a modeling product.



Source: Global Sea Surface Temperatures from the Weather Underground. http://www.wunderground.com/tropical/

1. **Raster.**

    a. Raster data are imagery and derived surfaces like a temperature or modeling map.

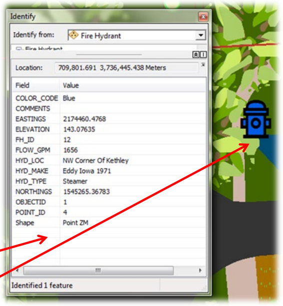
    b. Most Remote Sensing products are called Raster data sets.

    c. Raster data sets are composed of cells arranged in columns and rows. Zoom in tightly in a digital photograph and it begins to "pixelate" — the rows and columns of data become visible.

    d. Raster data sets may also be created through analysis processes, such as interpolation — the estimating of values between known, measured locations.

GEOMETRY VERSUS ATTRIBUTES

- Geometry is the spatially referenced depiction of a feature using a point, line or polygon.
- Attributes are pieces of information attached to geometric objects and are often stored in a table as part of a GIS layer.



Location: 709,801.691 3,736,445.438 Meters

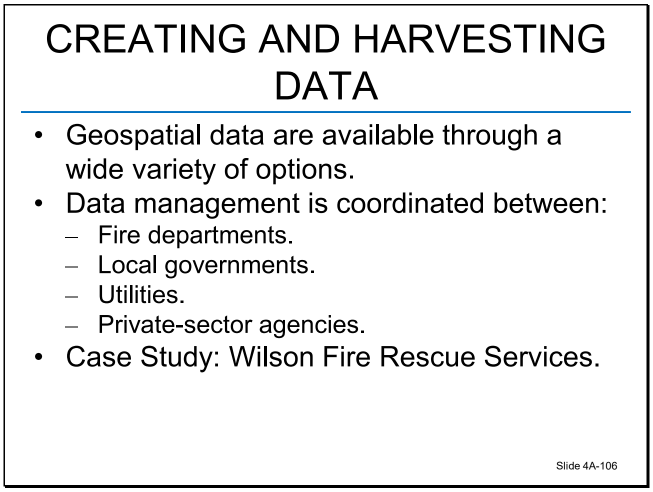
Field	Value
COLOR_CODE	Blue
COMMENTS	
EASTINGS	2174460.4768
ELEVATION	143.07635
FH_ID	12
FLOW_GPM	1656
HYD_LOC	NW Corner Of Kethley
HYD_MAKE	Eddy Iowa 1971
HYD_TYPE	Steamer
NORTHINGS	1545265.36783
OBJECTID	1
POINT_ID	4
Shape	Point ZM

Identified 1 feature

2. **Attributes.**

   a. They are pieces of information, often stored in a table as part of a GIS layer, about a given geographic feature.

   b. They are associated with each feature described in a vector data set. For example, a roadway may be represented by a line. That line may have attributes for the road's name, pavement type, capacity, etc. Attribute information is stored in a table that is related to each geographic feature.



CREATING AND HARVESTING DATA

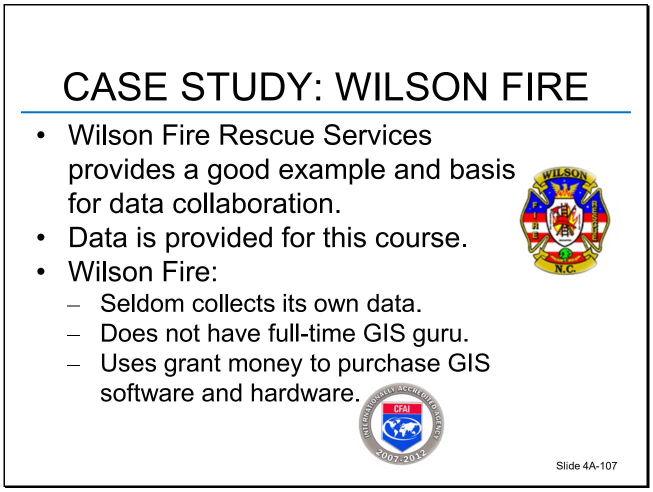
- Geospatial data are available through a wide variety of options.
- Data management is coordinated between:
  - Fire departments.
  - Local governments.
  - Utilities.
  - Private-sector agencies.
- Case Study: Wilson Fire Rescue Services.

D. Creating and harvesting data.

1. Geospatial data are available through a wide variety of options.

2. Where possible, fire departments should work with other governmental, utility and private-sector agencies to obtain data.

3. This unit will provide guidance for preparing typical internal fire service data sets that may be loaded readily into a GIS.



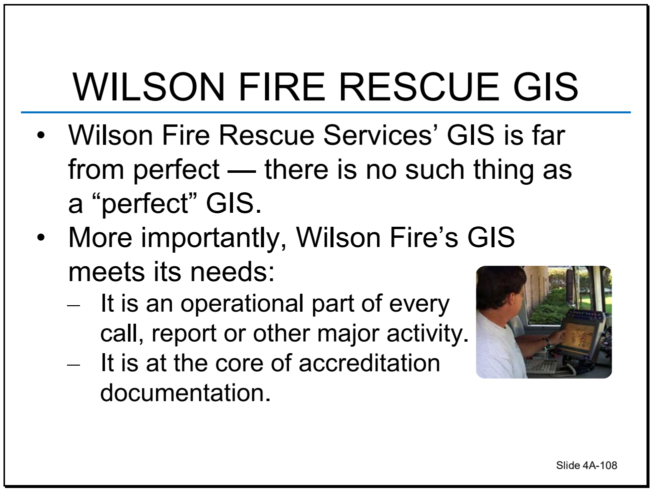
CASE STUDY: WILSON FIRE

- Wilson Fire Rescue Services provides a good example and basis for data collaboration.
- Data is provided for this course.
- Wilson Fire:
  - Seldom collects its own data.
  - Does not have full-time GIS guru.
  - Uses grant money to purchase GIS software and hardware.

E. Wilson Fire Rescue Services.

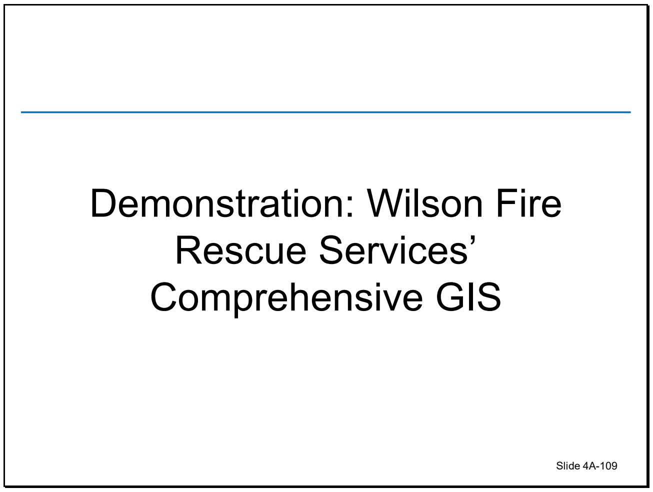
1. Wilson Fire Rescue Services' GIS provides an outstanding example and basis for data collaboration.

2. These data are provided as part of this course to use in exercises and to provide students with example data layers that show what is possible with GIS.



WILSON FIRE RESCUE GIS

- Wilson Fire Rescue Services' GIS is far from perfect — there is no such thing as a "perfect" GIS.
- More importantly, Wilson Fire's GIS meets its needs:
  - It is an operational part of every call, report or other major activity.
  - It is at the core of accreditation documentation.



Demonstration: Wilson Fire Rescue Services' Comprehensive GIS

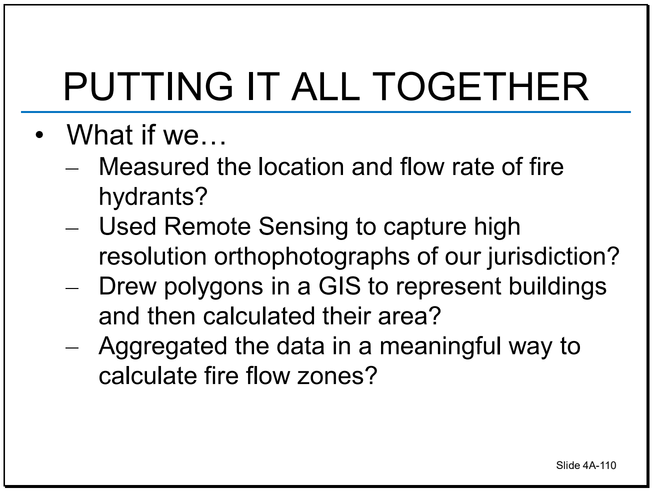
3. Wilson Fire Rescue Services:

   a. Seldom collects its own data — it partners or shares the cost burden with other entities.

   b. Does not have its own full-time GIS guru; rather, it has encouraged the development of GIS skills within its ranks and borrowed GIS help when needed.

   c. Has not paid out huge sums of money for GIS software and hardware. Wilson staffers are aggressive grant writers who work collaboratively to build their GIS.

VII. PUTTING IT ALL TOGETHER



PUTTING IT ALL TOGETHER

- What if we…
  - Measured the location and flow rate of fire hydrants?
  - Used Remote Sensing to capture high resolution orthophotographs of our jurisdiction?
  - Drew polygons in a GIS to represent buildings and then calculated their area?
  - Aggregated the data in a meaningful way to calculate fire flow zones?

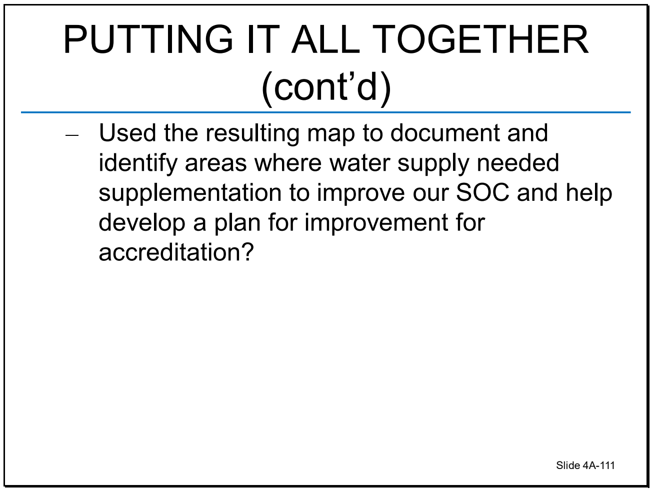
What if we…

A. Measured the location and flow rate of fire hydrants?

B. Used Remote Sensing to capture high-resolution orthophotographs of our jurisdiction?

C. Drew polygons in a GIS to represent buildings and then calculated their area?

D. Aggregated the data in a meaningful way to calculate fire flow zones?

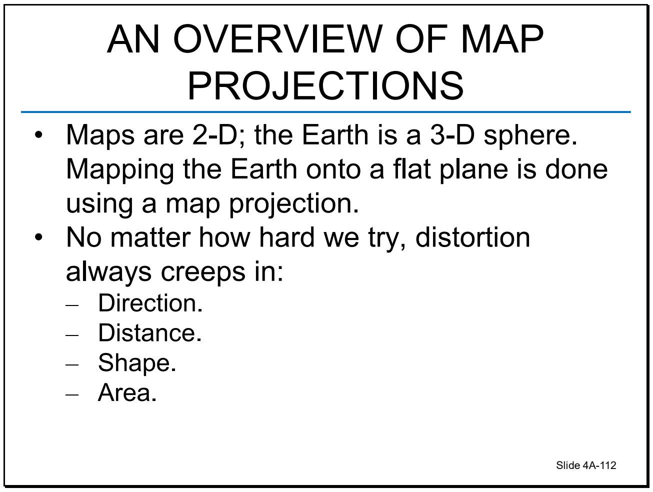


PUTTING IT ALL TOGETHER (cont'd)

- Used the resulting map to document and identify areas where water supply needed supplementation to improve our SOC and help develop a plan for improvement for accreditation?

E. Used the resulting map to document and identify areas where water supply needed supplementation to improve our SOC and help develop a plan for improvement for accreditation?

VIII. MAP PROJECTIONS — OVERVIEW

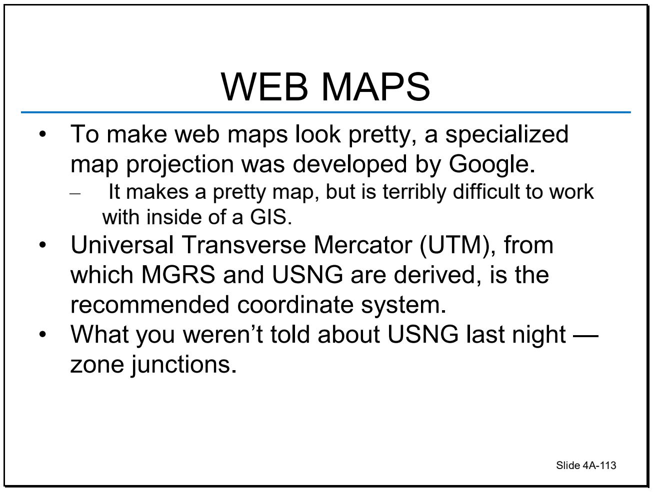


AN OVERVIEW OF MAP PROJECTIONS

- Maps are 2-D; the Earth is a 3-D sphere. Mapping the Earth onto a flat plane is done using a map projection.
- No matter how hard we try, distortion always creeps in:
  - Direction.
  - Distance.
  - Shape.
  - Area.

A. Maps are 2-D; the Earth is a 3-D sphere. Mapping the Earth onto a flat plane is done using a map projection.

B. When using a map projection, direction, distance, shape and area can be distorted.



WEB MAPS

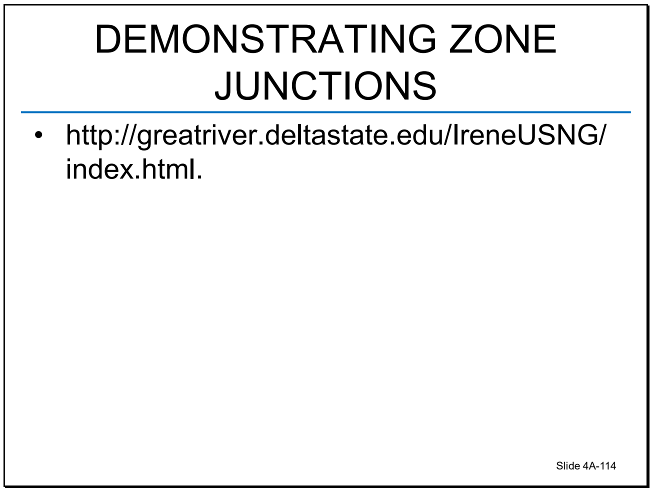
- To make web maps look pretty, a specialized map projection was developed by Google.
  - It makes a pretty map, but is terribly difficult to work with inside of a GIS.
- Universal Transverse Mercator (UTM), from which MGRS and USNG are derived, is the recommended coordinate system.
- What you weren't told about USNG last night — zone junctions.

Slide 4A-113

C. Web maps.

1. To make web maps look pretty, a specialized map projection was developed by Google.

2. UTM, from which MGRS and USNG are derived, is the recommended coordinate system.



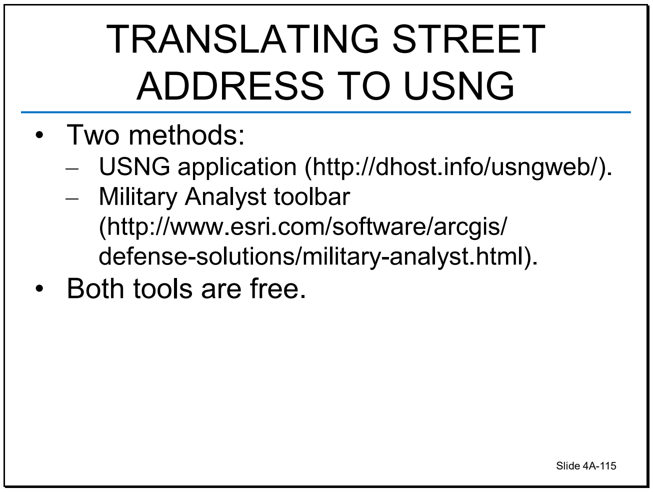
DEMONSTRATING ZONE JUNCTIONS

- http://greatriver.deltastate.edu/IreneUSNG/index.html.

Slide 4A-114

D. Demonstrating zone junctions.

http://greatriver.deltastate.edu/IreneUSNG/index.html.



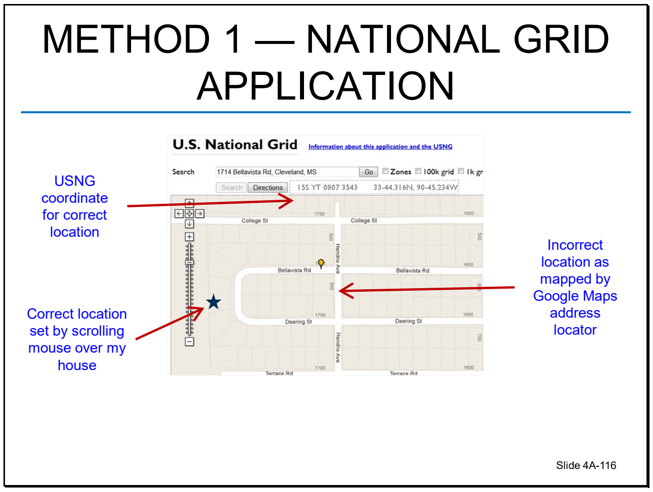
TRANSLATING STREET ADDRESS TO USNG

- Two methods:
  - USNG application (http://dhost.info/usngweb/).
  - Military Analyst toolbar (http://www.esri.com/software/arcgis/defense-solutions/military-analyst.html).
- Both tools are free.

E. Translating street address to USNG.

1. There are two methods for translating street addresses to USNG (the recommended location identifier). They are:

   a. USNG application (http://dhost.info/usngweb/).



METHOD 1 — NATIONAL GRID APPLICATION

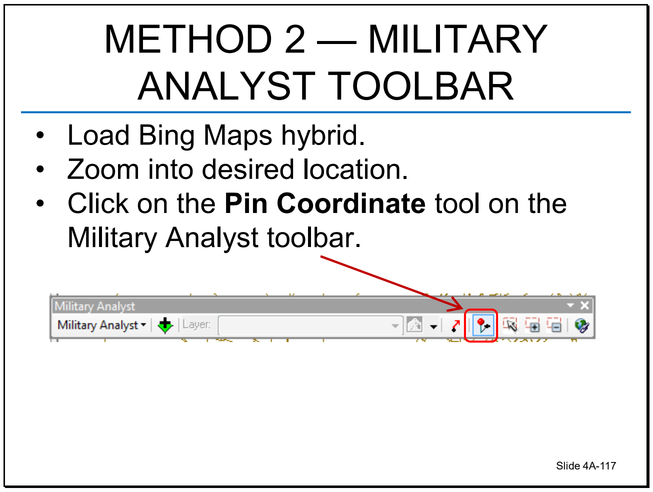
Search: 1714 Bellavista Rd, Cleveland, MS

15S YT 0807 3543 33-44.316N, 90-45.234W

USNG coordinate for correct location

Correct location set by scrolling mouse over my house

Incorrect location as mapped by Google Maps address locator



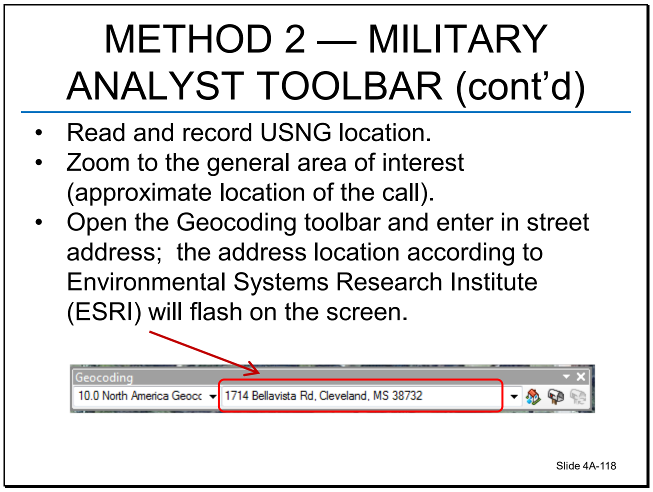
METHOD 2 — MILITARY ANALYST TOOLBAR

- Load Bing Maps hybrid.
- Zoom into desired location.
- Click on the **Pin Coordinate** tool on the Military Analyst toolbar.

b. Military Analyst toolbar (http://www.esri.com/software/arcgis/defense-solutions/military-analyst.html).

c. Both are free.

2. For the **USNG** application, after locating the address, record the USNG coordinate in NFIRS — without spaces.



METHOD 2 — MILITARY ANALYST TOOLBAR (cont'd)

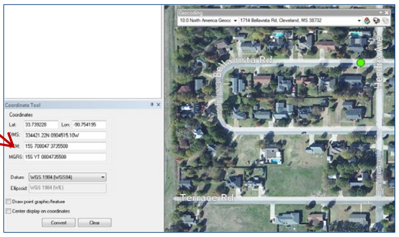
- Read and record USNG location.
- Zoom to the general area of interest (approximate location of the call).
- Open the Geocoding toolbar and enter in street address; the address location according to Environmental Systems Research Institute (ESRI) will flash on the screen.

1714 Bellavista Rd, Cleveland, MS 38732

3. For the **Military Analyst** tool, locate the street address, click the **Pin Coordinate** tool on the toolbar, and then record the USNG location.

METHOD 2 — MILITARY ANALYST TOOLBAR (cont'd)

- Move the mouse over the correct location and copy the MGRS location from the pin coordinate tool readout.



Position the mouse over the correct location and read the coordinate from MGRS line

Geocoding

10.0 North America Geocc

1714 Bellavista Rd, Cleveland, MS 38732

Coordinate Tool

Coordinates

Lat: 33.739228 Lon: -90.754195

DMS: 334421.22N 0904515.10W

UTM: 15S 708047 3735508

MGRS: 15S YT 0804735508

Datum: WGS 1984 (WGS84)

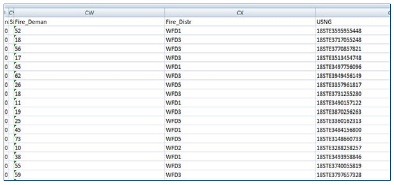
Ellipsoid: WGS 1984 (WE)

☐ Draw point graphic/feature

☐ Center display on coordinates

BULK MAPPING CALL LOCATION

- If all call location data are entered into NFIRS using USNG, they become standardized.
- Standardized data are easy to geocode.
- For our purposes, NFIRS data have been downloaded and converted into Excel spreadsheets.



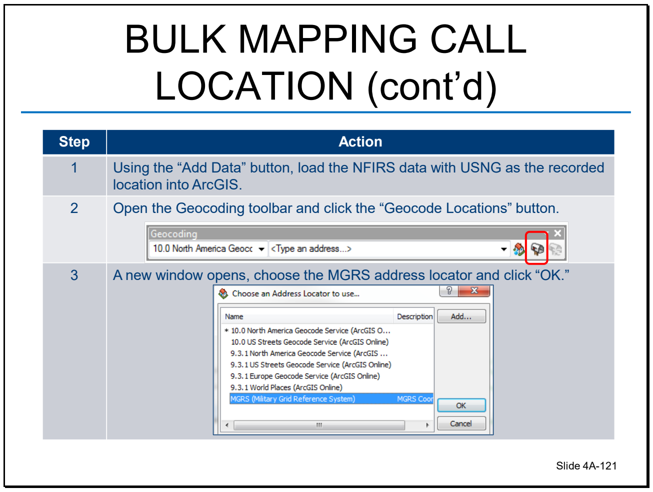
CW	CX	
Fire_Deman	Fire_Distr	USNG
52	WFD1	18STE3595955448
18	WFD3	18STE3717055248
56	WFD3	18STE3770857821
17	WFD3	18STE3513454748
45	WFD1	18STE3497756096
62	WFD3	18STE3949456149
26	WFD5	18STE3357961817
18	WFD3	18STE3731255280
11	WFD1	18STE3490157122
19	WFD3	18STE3870256263
25	WFD5	18STE3360162313
45	WFD1	18STE3484156800
73	WFD5	18STE3148660733
10	WFD2	18STE3288258257
38	WFD1	18STE3493958846
55	WFD3	18STE3740055819
59	WFD3	18STE3797657328

F. Bulk mapping call location.

1. If all call location data are entered into NFIRS using USNG, they become standardized.

2. Standardized data are easy to geocode.

3. To geocode in bulk, perform the following steps:



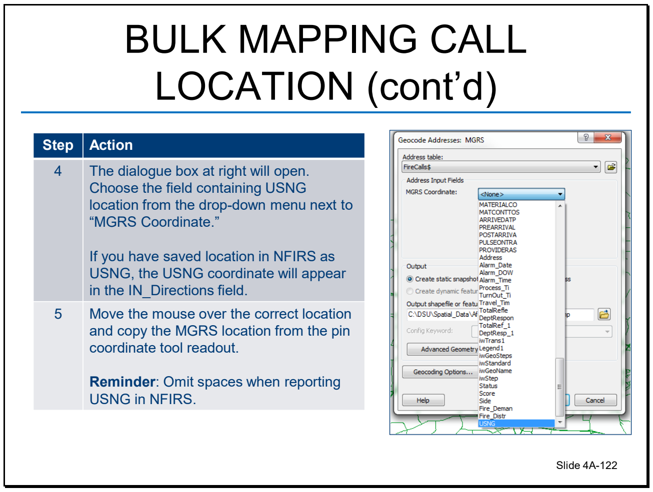
BULK MAPPING CALL LOCATION (cont'd)

Step	Action
1	Using the "Add Data" button, load the NFIRS data with USNG as the recorded location into ArcGIS.
2	Open the Geocoding toolbar and click the "Geocode Locations" button.
3	A new window opens, choose the MGRS address locator and click "OK."

a. Using the "Add Data" button, load the NFIRS data with USNG as the recorded location into ArcGIS.

b. Open the Geocoding toolbar and click the "Geocode Locations" button — the one that looks like a mailbox.

c. A new window will open — choose the MGRS address locator and click "OK."



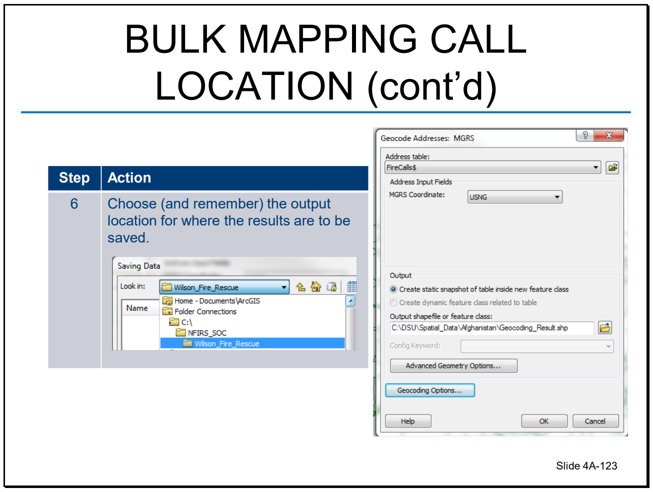
BULK MAPPING CALL LOCATION (cont'd)

Step	Action
4	The dialogue box at right will open. Choose the field containing USNG location from the drop-down menu next to "MGRS Coordinate." If you have saved location in NFIRS as USNG, the USNG coordinate will appear in the IN_Directions field.
5	Move the mouse over the correct location and copy the MGRS location from the pin coordinate tool readout. **Reminder**: Omit spaces when reporting USNG in NFIRS.

d. The "Geocode Addresses: MGRS" dialogue box opens. Choose the field containing USNG location from the drop-down menu next to "MGRS Coordinate."

e. If you have saved the location in NFIRS as USNG, the USNG coordinate will appear in the IN_Directions field.

f. **Reminder:** Omit spaces when reporting USNG in NFIRS.



BULK MAPPING CALL LOCATION (cont'd)

Step	Action
6	Choose (and remember) the output location for where the results are to be saved.

Saving Data

Look in: Wilson_Fire_Rescue

Home - Documents\ArcGIS
Folder Connections
C:\
NFIRS_SOC
Wilson_Fire_Rescue

Geocode Addresses: MGRS

Address table: FireCalls$

Address Input Fields

MGRS Coordinate: USNG

Output

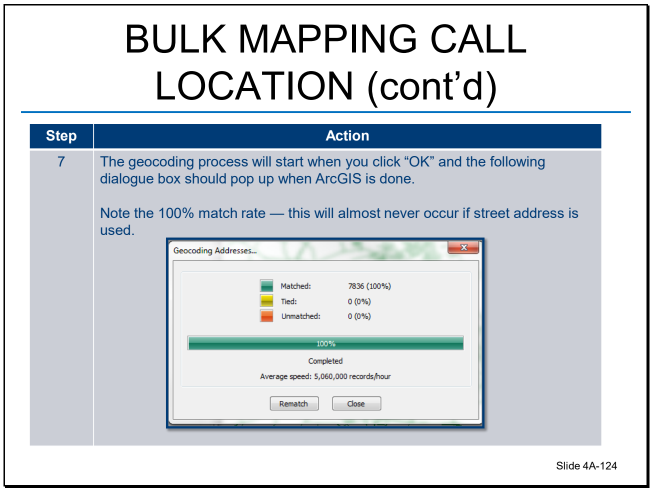
● Create static snapshot of table inside new feature class

○ Create dynamic feature class related to table

Output shapefile or feature class: C:\DSU\Spatial_Data\Afghanistan\Geocoding_Result.shp

Config Keyword:

g. Choose (and remember) the output location for where the results are to be saved.



BULK MAPPING CALL LOCATION (cont'd)

Step	Action
7	The geocoding process will start when you click "OK" and the following dialogue box should pop up when ArcGIS is done. Note the 100% match rate — this will almost never occur if street address is used.

Geocoding Addresses...

🟩	Matched:	7836 (100%)
🟨	Tied:	0 (0%)
🟥	Unmatched:	0 (0%)

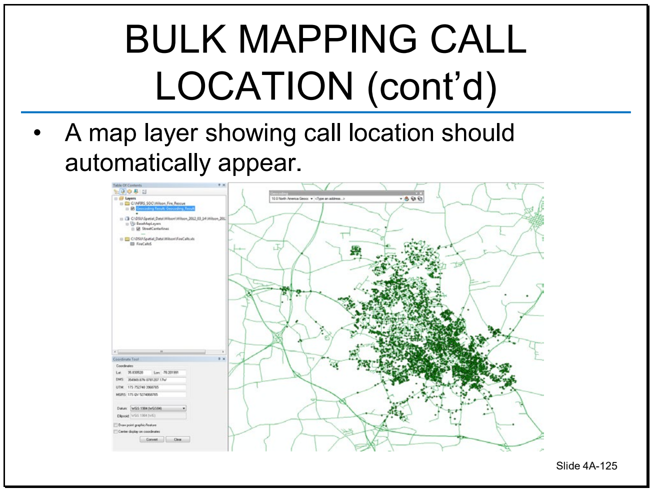
100%

Completed

Average speed: 5,060,000 records/hour

h. The geocoding process will start when you click "OK" and a dialogue should pop up when ArcGIS is done.

i. Note the 100 percent match rate — this will almost never occur if a street address is used.

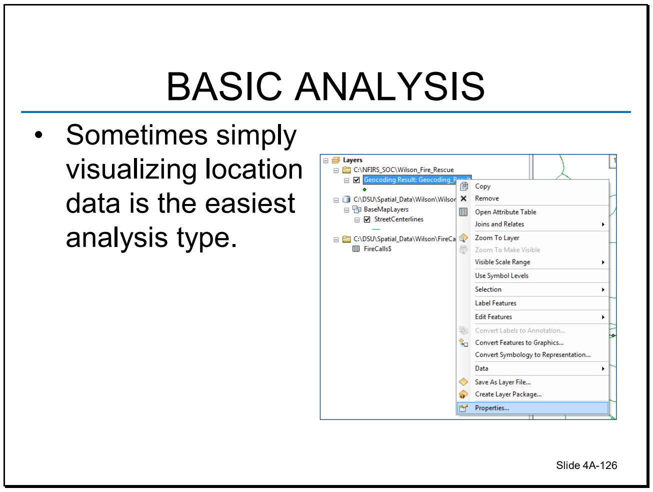


BULK MAPPING CALL LOCATION (cont'd)

- A map layer showing call location should automatically appear.

j. A map layer showing call location should automatically appear (if not, check your saved results).

IX. BASIC ANALYSIS — SYMBOLIZING DATA



BASIC ANALYSIS

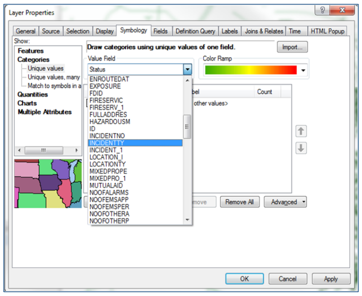
- Sometimes simply visualizing location data is the easiest analysis type.

A. Sometimes simply visualizing location data is the easiest analysis type.

B. Right click on the geocoding result layer and choose "Properties."

SYMBOLIZING DATA

Step	Action
1	Select the “Symbology” tab.
2	Click “Categories” — when the dialogue expands, choose “Unique values.”
3	Choose to symbolize the incident type by NFIRS code by selecting it from the “Value Field” drop-down menu.



C. Symbolizing data.

1. To show data symbolized, perform the following steps:

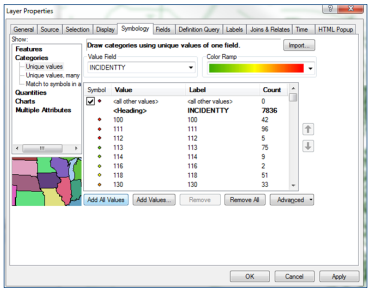
   a. Select the “Symbology” tab.

   b. Click “Categories” — when the dialogue expands, choose “Unique Values.”

   c. Choose to symbolize the incident type by NFIRS code by selecting it from the “Value Field” drop-down menu.

SYMBOLIZING DATA (cont’d)

Step	Action
4	Click on the “Add All Values” to load all call types.
5	Choose a color ramp
6	Select “OK.”



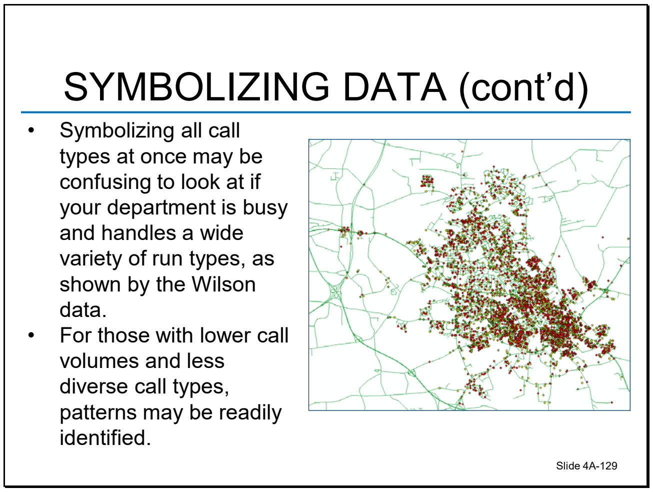
Value Field: INCIDENTTY

Symbol	Value	Label	Count
☑ ◆	<all other values>	<all other values>	0
	<Heading>	**INCIDENTTY**	**7836**
◆	100	100	42
◆	111	111	96
◆	112	112	5
◆	113	113	75
◆	114	114	9
◆	116	116	2
◆	118	118	51
◆	130	130	33

   d. Click on the “Add All Values” to load all call types.

   e. Choose a color ramp.

   f. Select “OK.”

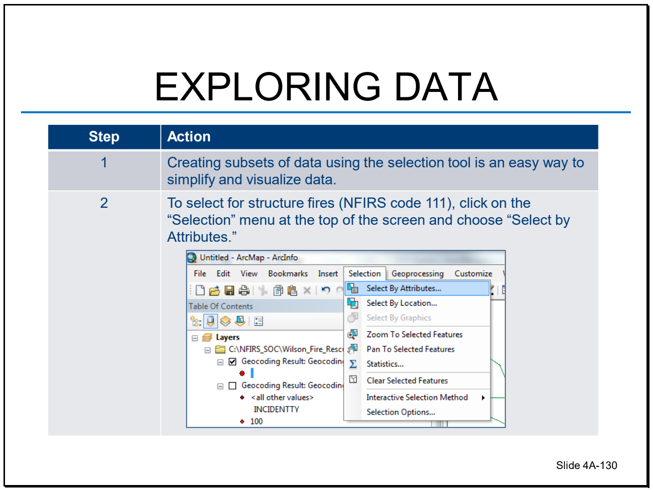


SYMBOLIZING DATA (cont’d)

- Symbolizing all call types at once may be confusing to look at if your department is busy and handles a wide variety of run types, as shown by the Wilson data.
- For those with lower call volumes and less diverse call types, patterns may be readily identified.

2. Symbolizing all call types at once may be confusing to look at if your department is busy and handles a wide variety of run types, as shown by the Wilson data below.

3. For those with lower call volumes and less diverse call types, patterns may be readily identified.



EXPLORING DATA

Step	Action
1	Creating subsets of data using the selection tool is an easy way to simplify and visualize data.
2	To select for structure fires (NFIRS code 111), click on the “Selection” menu at the top of the screen and choose “Select by Attributes.”

D. Exploring data.

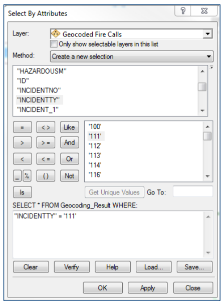
1. Creating subsets of data using the selection tool is an easy way to simplify and visualize data.

2. To select for structure fires (NFIRS code 111), click on the “Selection” menu at the top of the screen and choose “Select by Attributes.”

3. To show and explore data, perform the following steps:

EXPLORING DATA (cont'd)

Step	Action
4	Complete the dialogue box from top to bottom. Start by choosing the layer name containing your geocoded calls.
5	Scroll down in the top window and double click on the attribute that is to be explored – in this case it is IncidentTTY.
6	Click the "Equals" sign button.
7	Click the "Get Unique Values" button.
8	Scroll down in the right center window and double click on '111.' **Note**: ArcGIS is writing an equation, called a query, as you select attributes and values.
9	Click "OK" to run query.



Layer: Geocoded Fire Calls

Method: Create a new selection

SELECT * FROM Geocoding_Result WHERE:

"INCIDENTTY" = '111'

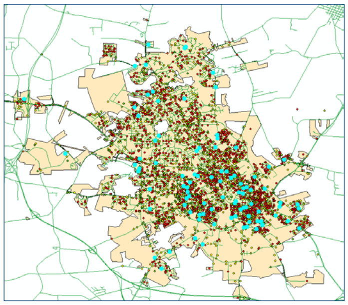
a. Fill out this dialogue box from top to bottom. Start by choosing the layer name containing your geocoded calls.

b. Scroll down in the top window and double click on the attribute that is to be explored — in this case it is IncidentTTY (Incident type).

c. Click the "Equals" sign button.

EXPLORING DATA (cont'd)

The results of the query will be returned as cyan (bright blue) dots.

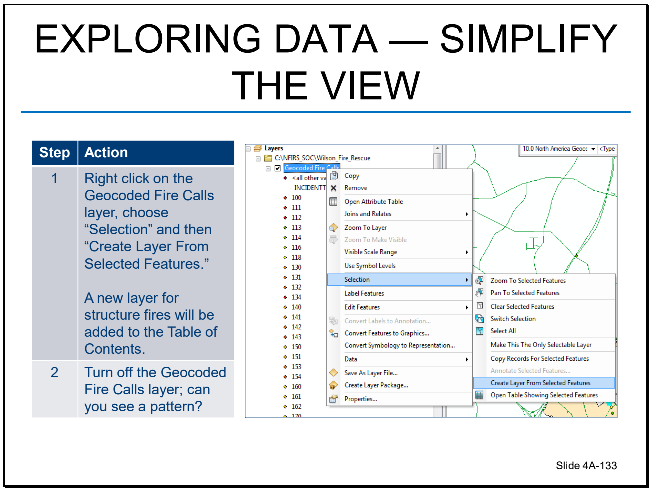


d. Click the "Get Unique Values" button.

e. Scroll down in the right center window and double click on 111.

- **Note:** ArcGIS is writing an equation, called a query, as you select attributes and values.

f. Click "OK" to run the query.



EXPLORING DATA — SIMPLIFY THE VIEW

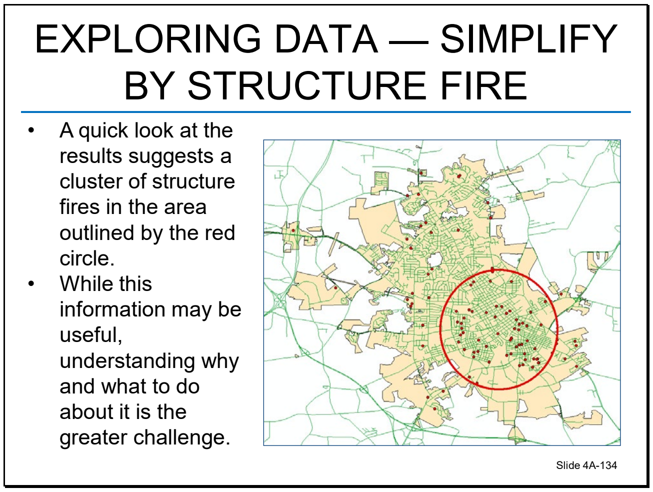
Step	Action
1	Right click on the Geocoded Fire Calls layer, choose “Selection” and then “Create Layer From Selected Features.” A new layer for structure fires will be added to the Table of Contents.
2	Turn off the Geocoded Fire Calls layer; can you see a pattern?

4. To simplify the view by creating a separate layer for structure fires, perform the following steps:

   a. Right click on the Geocoded Fire Calls layer; choose “Selection” and then “Create Layer From Selected Features.”

   b. A new layer for structure fires will be added to the Table of Contents.

   c. Turn off the Geocoded Fire Calls layer; can you see a pattern?

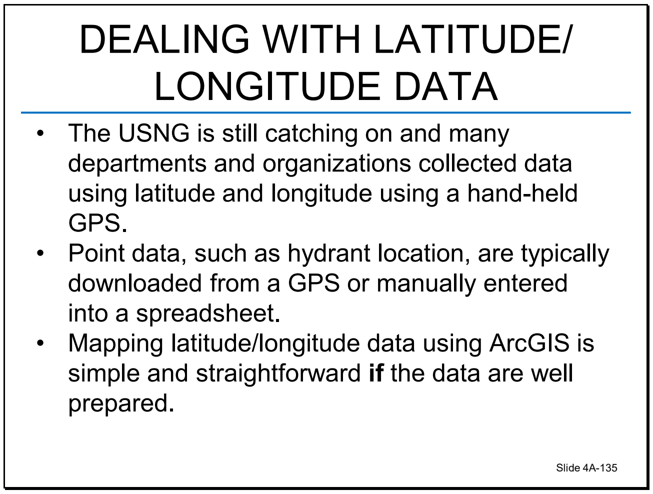


EXPLORING DATA — SIMPLIFY BY STRUCTURE FIRE

- A quick look at the results suggests a cluster of structure fires in the area outlined by the red circle.
- While this information may be useful, understanding why and what to do about it is the greater challenge.

5. A quick look at the results suggests a cluster of structure fires in the area outlined by the red circle.

6. While this information may be useful, understanding why and what to do about it is the challenge.



DEALING WITH LATITUDE/ LONGITUDE DATA

- The USNG is still catching on and many departments and organizations collected data using latitude and longitude using a hand-held GPS.
- Point data, such as hydrant location, are typically downloaded from a GPS or manually entered into a spreadsheet.
- Mapping latitude/longitude data using ArcGIS is simple and straightforward **if** the data are well prepared.

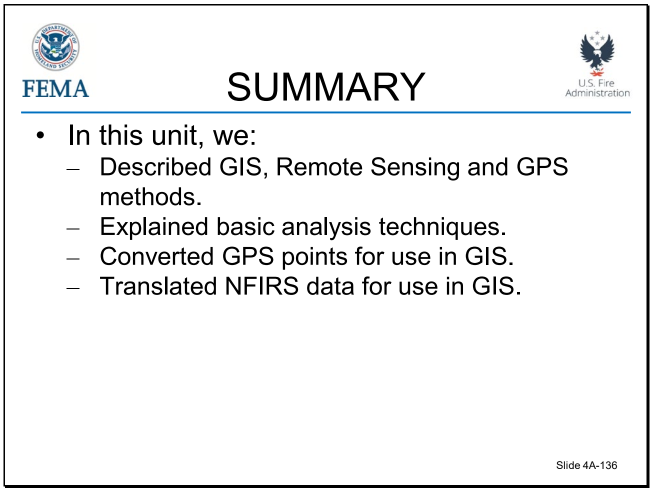
E. Dealing with latitude/longitude data.

1. The USNG is still catching on and many departments and organizations collected data using latitude and longitude from a hand-held GPS.

2. Point data, such as hydrant location, are typically downloaded from a GPS or manually entered into a spreadsheet.

3. Mapping latitude/longitude data using ArcGIS is simple and straightforward **if** the data are well prepared.

X. SUMMARY



FEMA

U.S. Fire Administration

SUMMARY

- In this unit, we:
  - Described GIS, Remote Sensing and GPS methods.
  - Explained basic analysis techniques.
  - Converted GPS points for use in GIS.
  - Translated NFIRS data for use in GIS.

UNIT 4: PART B GEOGRAPHIC INFORMATION TECHNOLOGIES

TERMINAL OBJECTIVES

The students will be able to:

4B.1 Assemble a risk assessment, deployment analysis and performance measurement to support a Standard of Cover (SOC) assessment.

4B.2 Given National Fire Incident Reporting System (NFIRS) and data analysis software, evaluate an organization's abilities to mitigate and respond to risk (Plan, Mitigate, Respond and Recover).

ENABLING OBJECTIVES

The students will be able to:

4B.1 Calculate attribute summary and geographic statistics.

4B.2 Create charts describing attribute data.

4B.3 Compare appropriate interpolation method.

4B.4 Apply the Poisson distribution to describe fire activity.

4B.5 Score risk levels based on map algebra.

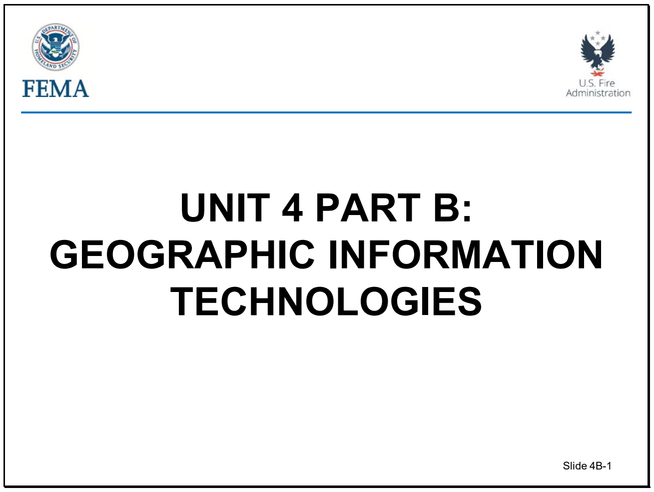
4B.6 Express response times.

4B.7 Express service areas.

4B.8 Create integrative map products which support SOC.

4B.9 Demonstrate the integration of GIS layer information for a given jurisdiction and apply response SOC to evaluate the impact of information on response capabilities.

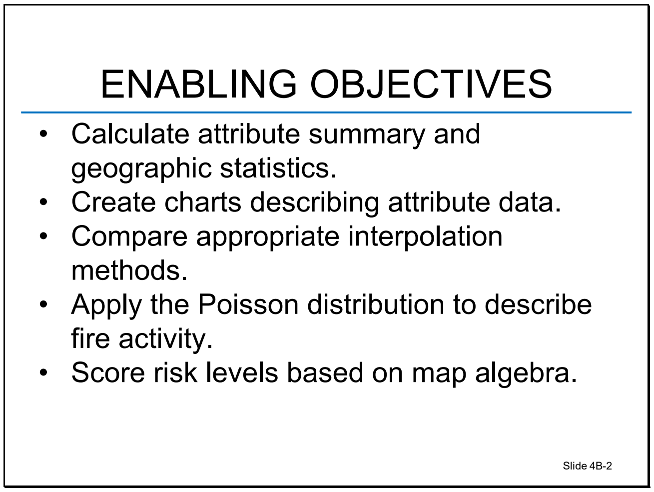
This page intentionally left blank.



FEMA

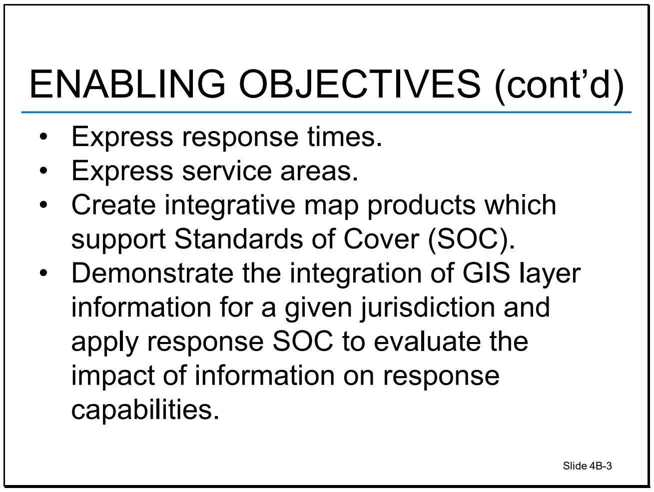
U.S. Fire Administration

UNIT 4 PART B: GEOGRAPHIC INFORMATION TECHNOLOGIES



ENABLING OBJECTIVES

- Calculate attribute summary and geographic statistics.
- Create charts describing attribute data.
- Compare appropriate interpolation methods.
- Apply the Poisson distribution to describe fire activity.
- Score risk levels based on map algebra.

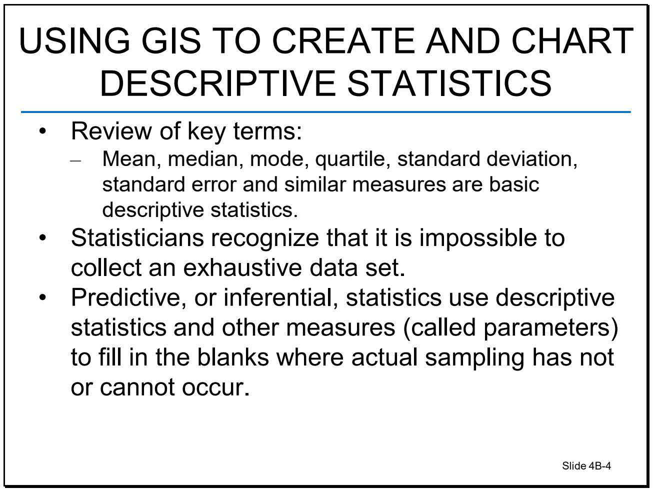


ENABLING OBJECTIVES (cont'd)

- Express response times.
- Express service areas.
- Create integrative map products which support Standards of Cover (SOC).
- Demonstrate the integration of GIS layer information for a given jurisdiction and apply response SOC to evaluate the impact of information on response capabilities.

I. UNIT INTRODUCTION

II. USING GEOGRAPHIC INFORMATION SYSTEM TO CREATE AND CHART DESCRIPTIVE STATISTICS



USING GIS TO CREATE AND CHART DESCRIPTIVE STATISTICS

- Review of key terms:
  - Mean, median, mode, quartile, standard deviation, standard error and similar measures are basic descriptive statistics.
- Statisticians recognize that it is impossible to collect an exhaustive data set.
- Predictive, or inferential, statistics use descriptive statistics and other measures (called parameters) to fill in the blanks where actual sampling has not or cannot occur.

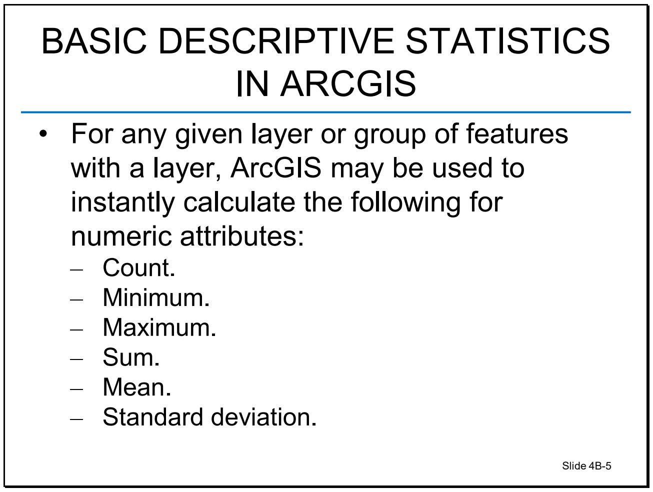
A. Review of basic statistics terms.

1. Descriptive statistics describe the nature of a data set.

2. Mean, median, mode, quartile, standard deviation, standard error and similar measures are basic descriptive statistics.

3. Statisticians recognize that it is impossible to collect an exhaustive data set. For example, finding the true average height of males in America would require measuring the height of **every** male in the country. Instead of undertaking this mammoth task, a statistician would sample a smaller group assumed to be representative of all men. Analysis results about this sample would be used to draw conclusions about the entire male population.

4. Predictive, or inferential, statistics use descriptive statistics and other measures (called parameters) to fill in the blanks where actual sampling has not or cannot occur.



BASIC DESCRIPTIVE STATISTICS IN ARCGIS

- For any given layer or group of features with a layer, ArcGIS may be used to instantly calculate the following for numeric attributes:
  - Count.
  - Minimum.
  - Maximum.
  - Sum.
  - Mean.
  - Standard deviation.

B. Basic descriptive statistics in ArcGIS.

1. For any given layer or group of features with a layer, ArcGIS may be used to instantly calculate the following for numeric attributes:

   a. Count — the number of features used in calculating the descriptive statistic.

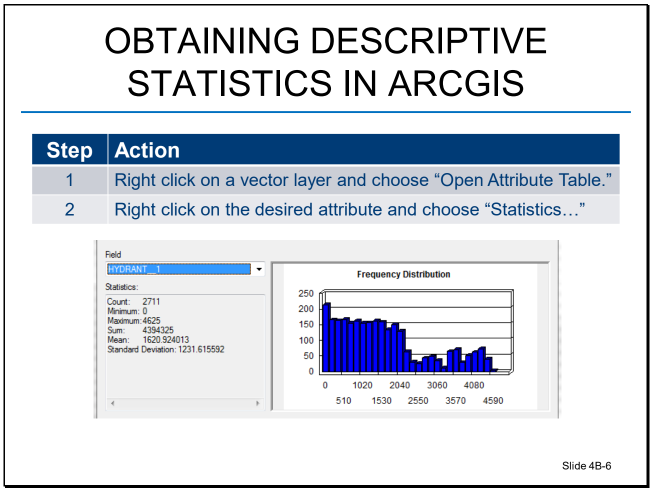
   b. Minimum — the minimum (lowest) value in the data set.

   c. Maximum — the maximum (highest) value in the data set.

   d. Sum — the sum of all values in the data set.

   e. Mean — the mean (average) value of the data set.

   f. Standard deviation.



OBTAINING DESCRIPTIVE STATISTICS IN ARCGIS

Step	Action
1	Right click on a vector layer and choose “Open Attribute Table.”
2	Right click on the desired attribute and choose “Statistics…”

Field: HYDRANT__1

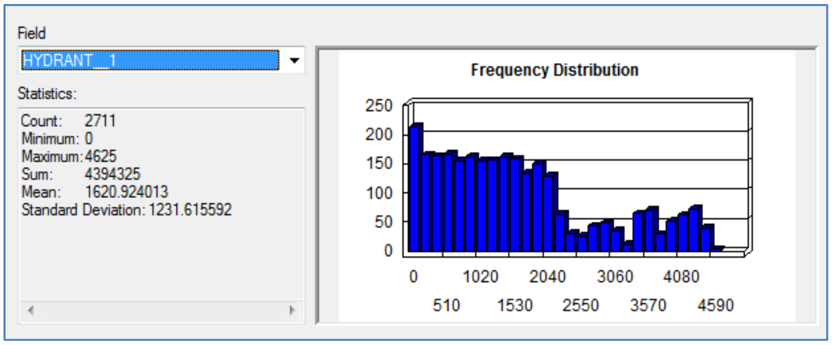
Statistics:

Count: 2711
Minimum: 0
Maximum: 4625
Sum: 4394325
Mean: 1620.924013
Standard Deviation: 1231.615592

2. To obtain basic descriptive statistics:

Step	Action
1	Right click on a vector layer and select "Open Attribute Table."
2	Right click on the desired attribute and choose "Statistics…"

3. A new window appears with descriptive statistical information about the chosen attribute.


Field
HYDRANT__1

Statistics:

Count: 2711
Minimum: 0
Maximum: 4625
Sum: 4394325
Mean: 1620.924013
Standard Deviation: 1231.615592

SYMBOLIZING BY ATTRIBUTES

- By choosing to symbolize our hydrant data based on an attribute value, we can make it easier to understand.
- One example color codes hydrants based upon flow rate.
- The color scheme chosen matches Incident Safety Officer (ISO) hydrant-marking guidance:
  - Class C — less than 500 gallons per minute (gpm) — red.
  - Class B — 500-999 gpm — orange.
  - Class A — 1,000-1,499 gpm — green.
  - Class AA — 1,500 gpm and above — light blue.

C. Symbolizing by attributes.

1. By choosing to symbolize our hydrant data based on an attribute value, we can make it easier to understand.

2. One example color codes hydrants based upon flow rate. The color scheme chosen matches Incident Safety Officer (ISO) hydrant-marking guidance:

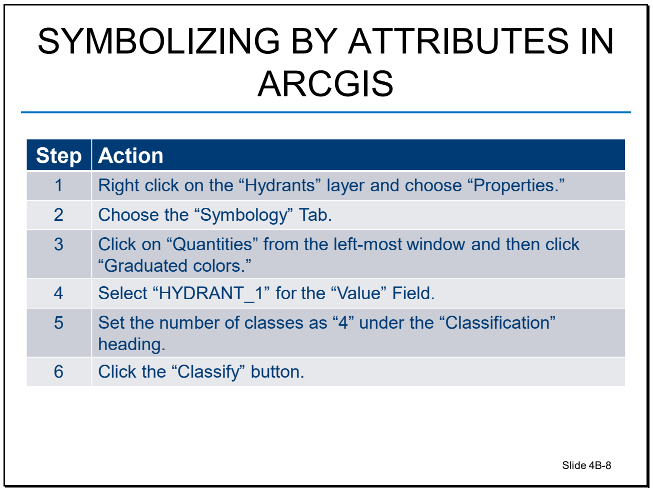
   a. Class C — less than 500 gallons per minute (gpm) — red.

   b. Class B — 500-999 gpm — orange.

   c. Class A — 1,000-1,499 gpm — green.

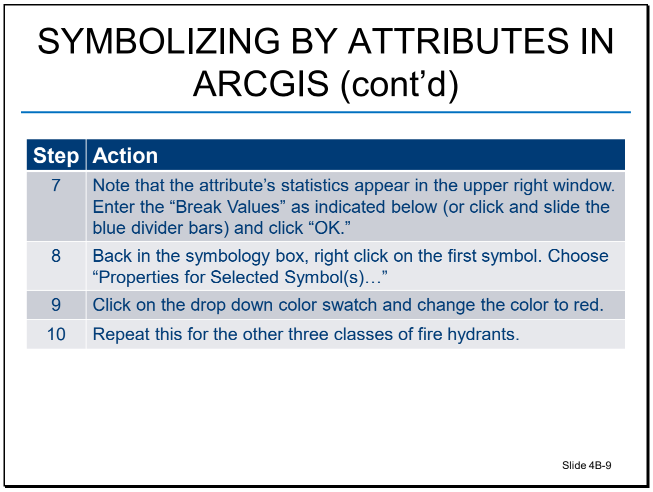
   d. Class AA — 1,500 gpm and above — light blue.

3. To symbolize by attributes:



SYMBOLIZING BY ATTRIBUTES IN ARCGIS

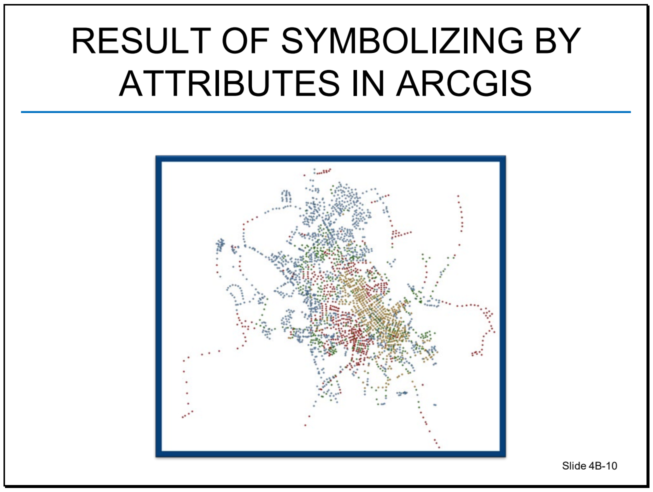
Step	Action
1	Right click on the “Hydrants” layer and choose “Properties.”
2	Choose the “Symbology” Tab.
3	Click on “Quantities” from the left-most window and then click “Graduated colors.”
4	Select “HYDRANT_1” for the “Value” Field.
5	Set the number of classes as “4” under the “Classification” heading.
6	Click the “Classify” button.



SYMBOLIZING BY ATTRIBUTES IN ARCGIS (cont’d)

Step	Action
7	Note that the attribute’s statistics appear in the upper right window. Enter the “Break Values” as indicated below (or click and slide the blue divider bars) and click “OK.”
8	Back in the symbology box, right click on the first symbol. Choose “Properties for Selected Symbol(s)…”
9	Click on the drop down color swatch and change the color to red.
10	Repeat this for the other three classes of fire hydrants.

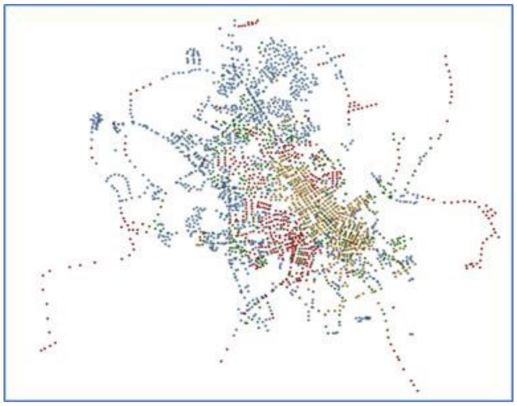
Step	Action
1	Right click on the “Hydrants” layer and choose “Properties.”
2	Choose the “Symbology” Tab.
3	Click on “Quantities” from the left-most window and then click “Graduated colors.”
4	Select “HYDRANT_1” for the “Value” Field.
5	Set the number of classes as “4” under the “Classification” heading.
6	Click the “Classify” button.
7	Note that the attribute’s statistics appear in the upper right window. Enter the “Break Values” as indicated below (or click and slide the blue divider bars) and click “OK.”
8	Back in the symbology box; right click on the first symbol. Choose “Properties for Selected Symbol(s)…”
9	Click on the drop down color swatch and change the color to red.
10	Repeat this for the other three classes of fire hydrants.



RESULT OF SYMBOLIZING BY ATTRIBUTES IN ARCGIS

Slide 4B-10

4. A new window appears with descriptive statistical information about the chosen attribute.



ACTIVITY 4B.1

Water Flow Capacity and First Due Assignments

Purpose

To successfully chart descriptive statistics on water flow and first due assignments using ArcGIS.

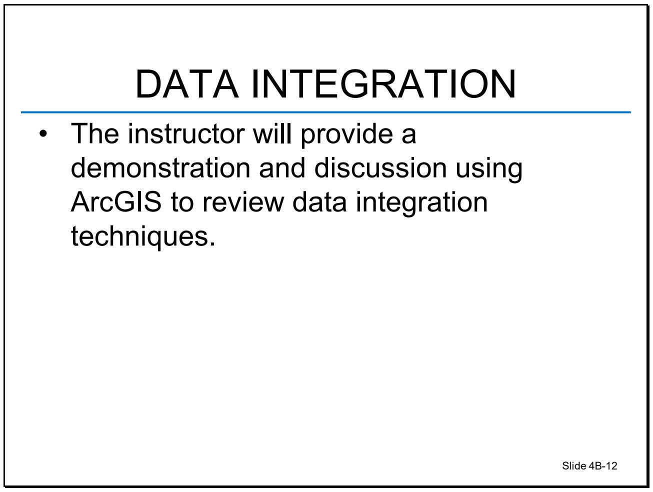
Directions

1. You will work individually.
2. You will use the worksheet provided in your Student Manuals (SMs). You should:
    a. Load structure fires, hydrants and first due areas.
    b. Use symbology to describe potential hydrant water flow.
    c. Use statistics to summarize the number of working fires per first due area.
    d. Use the map to analyze where greater water flow is required.
3. The instructor will be available to answer any questions.

This page intentionally left blank.

III. AFTERNOON ACTIVITY

IV. DATA INTEGRATION



DATA INTEGRATION

- The instructor will provide a demonstration and discussion using ArcGIS to review data integration techniques.

Slide 4B-12

This unit focuses on data integration techniques and the use of ArcGIS to perform analysis.

This page intentionally left blank.

ACTIVITY 4B.2

Data Integration

Purpose

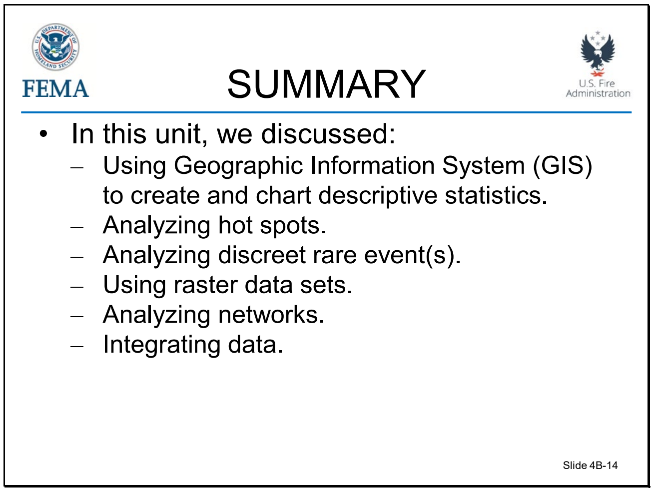
To successfully perform analysis using data integration techniques in ArcGIS for Wilson, N.C. case study.

Directions

1. You will work individually.
2. You will load the following data layers into ArcGIS:
    a. Hydrants.
    b. Building footprints.
    c. First due areas.
    d. Aerial photography.
    e. Street centerlines.
    f. Hazardous materials.
    g. Electrical utility infrastructure.
3. You will create 100-by-100 meter United States National Grid (USNG) polygons for Wilson, N.C.
4. Data layers in step 1 will be normalized to the USNG polygons created in step 2 (attributes representative of the average value for each layer added to a USNG polygon).
5. Map algebra will be applied to create a model for evaluating Standards of Cover (SOC) by numerically combining each layer, with an appropriate weighting factor, to evaluate additive factors.
6. You will discuss your results.

This page intentionally left blank.

V. SUMMARY



FEMA

U.S. Fire Administration

SUMMARY

- In this unit, we discussed:
  - Using Geographic Information System (GIS) to create and chart descriptive statistics.
  - Analyzing hot spots.
  - Analyzing discreet rare event(s).
  - Using raster data sets.
  - Analyzing networks.
  - Integrating data.

This page intentionally left blank.

UNIT 5:
BRINGING IT FULL CIRCLE

TERMINAL OBJECTIVE

The students will be able to:

5.1 Given Geographic Information Systems (GIS) and Standards of Coverage (SOC) data, you shall propose SOC criteria for application in your jurisdiction.

ENABLING OBJECTIVES

The students will be able to:

5.1 Propose standards of coverage IAW accepted standard.

5.2 Determine the information that will need to be developed and applied as data into data to determine the appropriate standards of coverage.

5.3 Determine the NFA core values to be applied to a standards of coverage assessment.

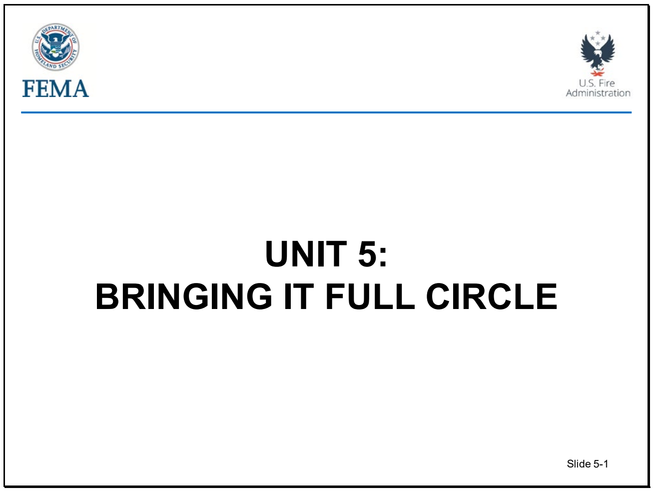
5.4 Articulate one element of applicable knowledge to be applied to the standards of coverage.

5.5 Develop the required elements for a standards of coverage for first alarm assignment.

5.6 Predict the difference between jurisdictional standards of coverage and industry standards of coverage.

5.7 Develop methods for meeting the difference between jurisdictional standards of coverage and industry standards of coverage.

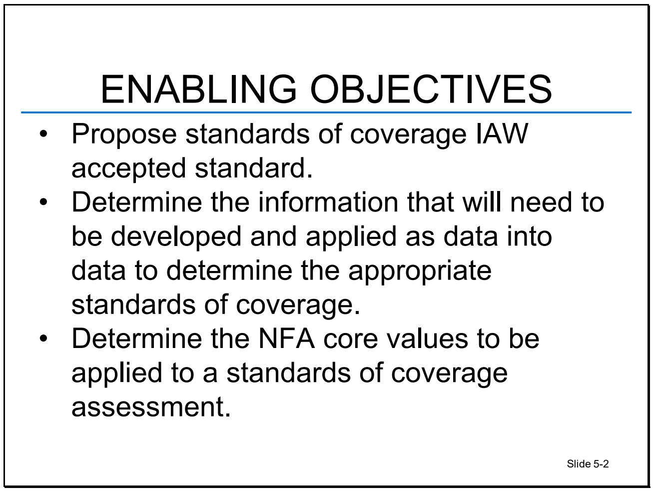
This page intentionally left blank.



FEMA

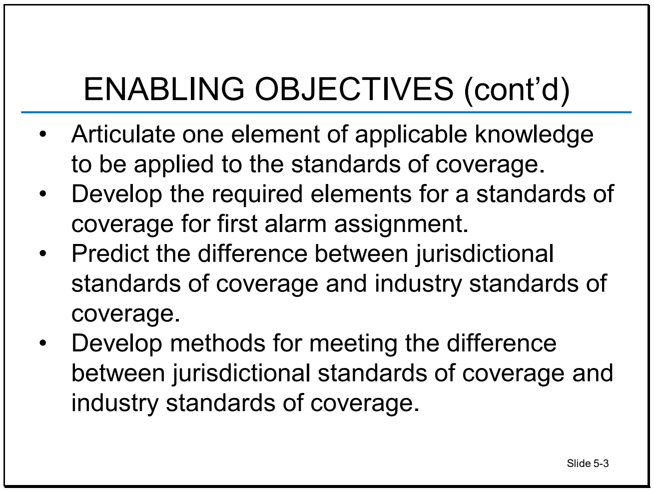
U.S. Fire Administration

UNIT 5:
BRINGING IT FULL CIRCLE



ENABLING OBJECTIVES

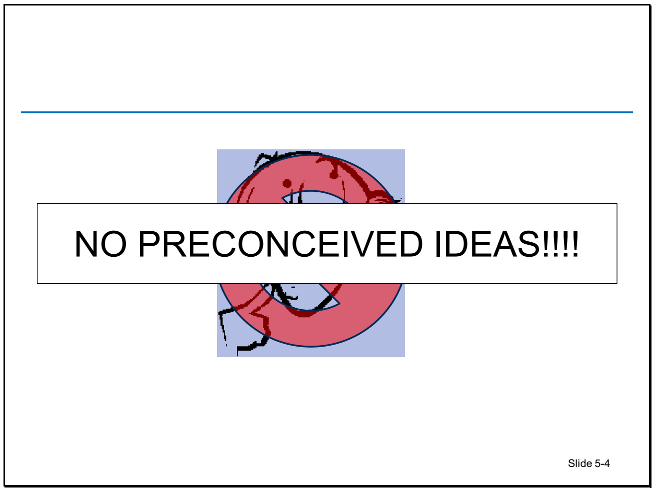
- Propose standards of coverage IAW accepted standard.
- Determine the information that will need to be developed and applied as data into data to determine the appropriate standards of coverage.
- Determine the NFA core values to be applied to a standards of coverage assessment.



ENABLING OBJECTIVES (cont'd)

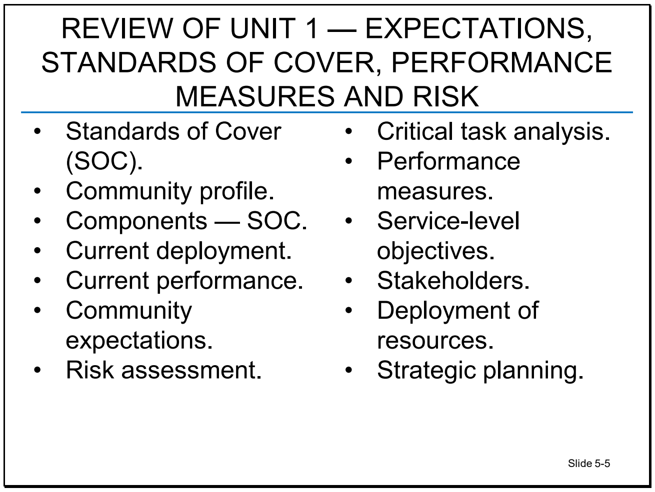
- Articulate one element of applicable knowledge to be applied to the standards of coverage.
- Develop the required elements for a standards of coverage for first alarm assignment.
- Predict the difference between jurisdictional standards of coverage and industry standards of coverage.
- Develop methods for meeting the difference between jurisdictional standards of coverage and industry standards of coverage.

I. UNIT INTRODUCTION



NO PRECONCEIVED IDEAS!!!!

A. No preconceived ideas.



REVIEW OF UNIT 1 — EXPECTATIONS, STANDARDS OF COVER, PERFORMANCE MEASURES AND RISK

- Standards of Cover (SOC).
- Community profile.
- Components — SOC.
- Current deployment.
- Current performance.
- Community expectations.
- Risk assessment.
- Critical task analysis.
- Performance measures.
- Service-level objectives.
- Stakeholders.
- Deployment of resources.
- Strategic planning.

B. Review of Unit 1 — Expectations, Standards of Cover, Performance Measure and Risk.

In Unit 1, we discussed:

1. Standards of Cover (SOC).

2. Community profile.

3. Components — SOC.

4. Current deployment.

5. Current performance.

6. Community expectations.

7. Risk assessment.

8. Critical task analysis.

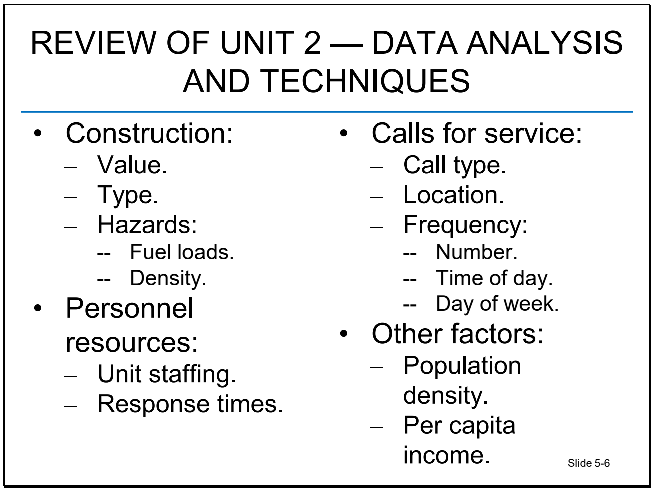
9. Performance measures.

10. Service-level objectives.

11. Stakeholders.

12. Deployment of resources.

13. Strategic planning.



REVIEW OF UNIT 2 — DATA ANALYSIS AND TECHNIQUES

- Construction:
  - Value.
  - Type.
  - Hazards:
    - Fuel loads.
    - Density.
- Personnel resources:
  - Unit staffing.
  - Response times.
- Calls for service:
  - Call type.
  - Location.
  - Frequency:
    - Number.
    - Time of day.
    - Day of week.
- Other factors:
  - Population density.
  - Per capita income.

Slide 5-6

C. Review of Unit 2 — Data Analysis and Techniques.

In Unit 2, we discussed:

1. Construction:

   a. Value.

   b. Type.

   c. Hazards:

      - Fuel loads.

      - Density.

2. Personnel resources:

   a. Unit staffing.

   b. Response times.

3. Calls for service:

   a. Call type.

   b. Location.

   c. Frequency:

      - Number.

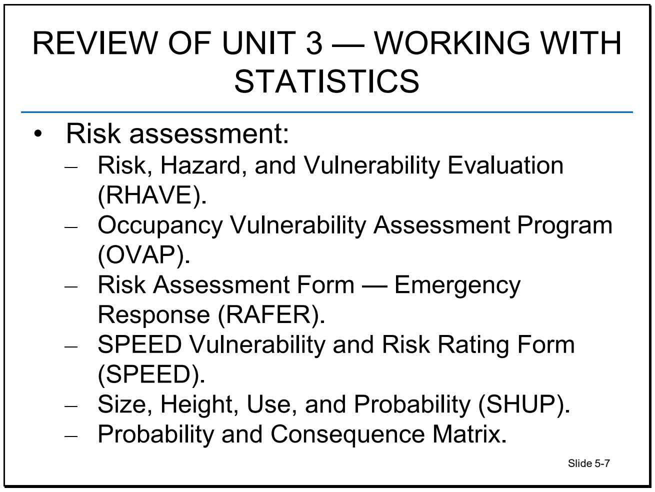
      - Time of day.

      - Day of week.

4. Other factors:

   a. Population density.

   b. Per capita income.



REVIEW OF UNIT 3 — WORKING WITH STATISTICS

- Risk assessment:
  - Risk, Hazard, and Vulnerability Evaluation (RHAVE).
  - Occupancy Vulnerability Assessment Program (OVAP).
  - Risk Assessment Form — Emergency Response (RAFER).
  - SPEED Vulnerability and Risk Rating Form (SPEED).
  - Size, Height, Use, and Probability (SHUP).
  - Probability and Consequence Matrix.

Slide 5-7

D. Review of Unit 3 — Working With Statistics.

In Unit 3, we discussed:

1. Risk assessment.

a. Risk, Hazard, and Vulnerability Evaluation (RHAVE).

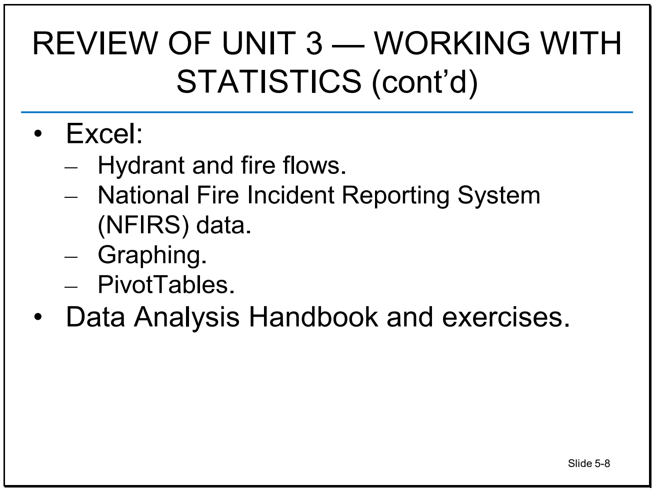
b. Occupancy Vulnerability Assessment Program (OVAP).

c. Risk Assessment Form — Emergency Response (RAFER).

d. SPEED Vulnerability and Risk Rating Form (SPEED).

e. Size, Height, Use, and Probability (SHUP).

f. Probability and Consequence Matrix.



REVIEW OF UNIT 3 — WORKING WITH STATISTICS (cont'd)

- Excel:
  - Hydrant and fire flows.
  - National Fire Incident Reporting System (NFIRS) data.
  - Graphing.
  - PivotTables.
- Data Analysis Handbook and exercises.

Slide 5-8

2. Excel.

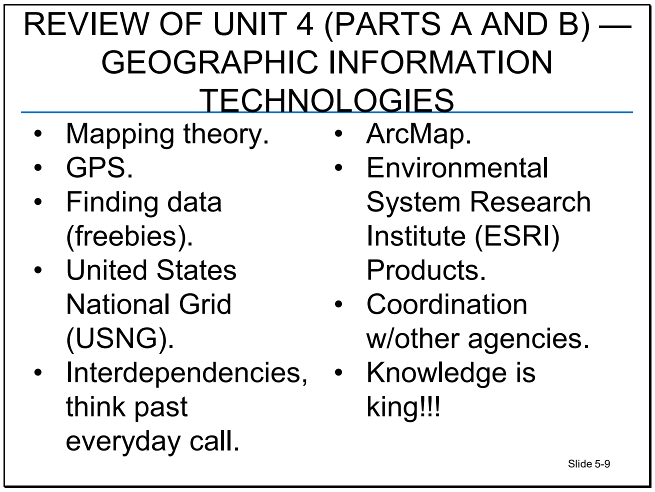
   a. Hydrant and fire flows.

   b. National Fire Incident Reporting System (NFIRS) data.

   c. Graphing.

   d. PivotTables.

3. Data Analysis Handbook and exercises.



REVIEW OF UNIT 4 (PARTS A AND B) — GEOGRAPHIC INFORMATION TECHNOLOGIES

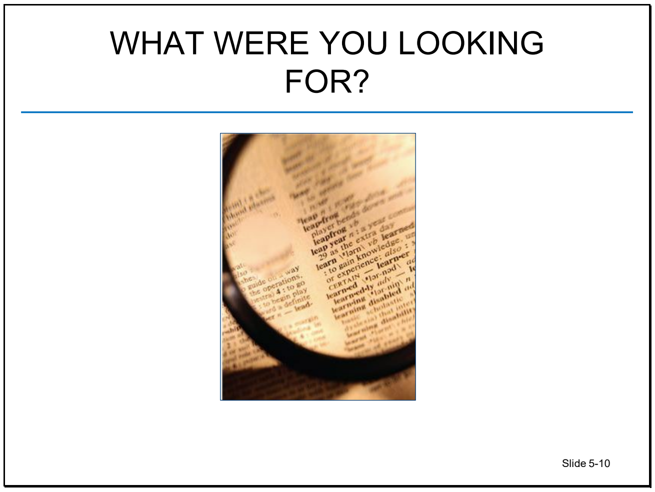
- Mapping theory.
- GPS.
- Finding data (freebies).
- United States National Grid (USNG).
- Interdependencies, think past everyday call.
- ArcMap.
- Environmental System Research Institute (ESRI) Products.
- Coordination w/other agencies.
- Knowledge is king!!!

E. Review of Unit 4 (Parts A and B) — Geographic Information Technologies.

In Unit 4, we discussed:

1. Mapping theory.
2. GPS.
3. Finding data (freebies).
4. United States National Grid (USNG).
5. Interdependencies, think past everyday call.
6. ArcMap.
7. Environmental Systems Research Institute (ESRI) products.
8. Coordination w/other agencies.
9. Knowledge is king!!!

II. WHAT WERE YOU LOOKING FOR?

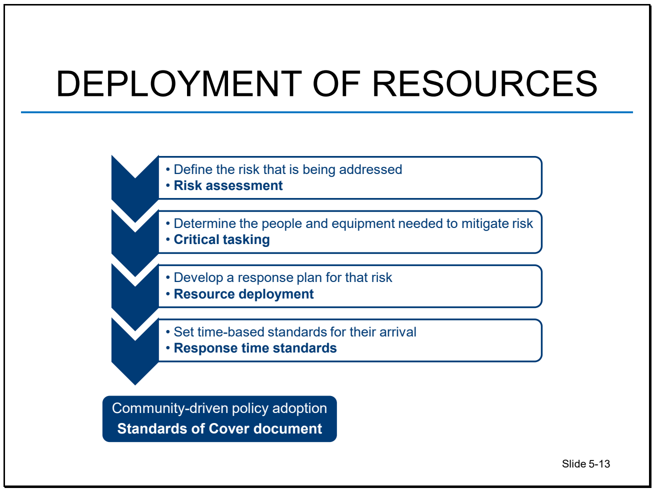


WHAT WERE YOU LOOKING FOR?

TABLE EXPECTATIONS PLACEHOLDER

HOW DID WE DO?

- Table 1
- Table 2
- Table 3
- Table 4



DEPLOYMENT OF RESOURCES

- Define the risk that is being addressed
- **Risk assessment**

- Determine the people and equipment needed to mitigate risk
- **Critical tasking**

- Develop a response plan for that risk
- **Resource deployment**

- Set time-based standards for their arrival
- **Response time standards**

Community-driven policy adoption
Standards of Cover document

Review the deployment of resources:

A. Define the risk that is being addressed.

B. **Risk assessment.**

C. Determine the people and equipment needed to mitigate risk.

D. **Critical tasking.**

E. Develop a response plan for that risk.

F. **Resource deployment.**

G. Set time-based standards for their arrival.

H. **Response time standards.**

ACTIVITY 5.1

Standards of Cover Presentations

Purpose

To provide overview of your SOC developed during this course.

Directions

1. You will be grouped into urban, suburban and rural regions for the presentations.
2. Each group will present the SOC outline you developed during this course, including the following information:
    a. Data developed germane to the group's jurisdiction.
    b. Application of the class to different levels in their department(s).
    c. Association of NFA core values to SOC.
    d. Identification of one "takeaway" (also known as an "aha moment") from working on SOC.
3. You will record key "takeaways" from each presenter on an easel pad.

This page intentionally left blank.

ACTIVITY 5.2

Deployment Analysis

Purpose

To allow you to determine the deployment analysis of your respective agency to a 2,000 square foot residential structure utilizing the effective response force tool.

Directions

1. Review the appropriate sections of the Fire Suppression Rating Schedule, Section 560 and the National Fire Protection Association (NFPA) Standard 1710, relating to "Deployment Analysis."

2. Using the example in the NFPA standard: The initial full alarm assignment to a structure fire in a typical 2,000 square foot, two-story single-family dwelling without basement and with any exposures.

   a. Using the template for effective response force, show your current first alarm response to this identified risk.

   b. Using the same template, what would you like to see as the ideal response to the identified risk?

   c. What would you need to close the gap identified by comparing a. and b.?

   d. Is it possible?

This page intentionally left blank.

ACTIVITY 5.2 (cont'd)

Deployment Analysis

Insurance Services Office — 2012 Fire Suppression Rating Schedule

Section 560 Deployment Analysis

The built-upon area of the fire protection area should have a first-due engine company within 1½ road miles and a ladder-service company within 2½ road miles.

As an alternative to determining the number of needed engine and ladder/service companies through road mile analysis, a fire protection area may use the results of a systematic performance evaluation. This type of evaluation analyzes computer-aided dispatch (CAD) history to demonstrate that, with its current deployment of companies, the fire department meets the time constraints for initial arriving engine and initial full-alarm assignment in accordance with the general criteria of NFPA 1710.

NFPA 1710 — 2010 Edition

ORGANIZATION AND DEPLOYMENT OF FIRE SUPPRESSION OPERATIONS BY CAREER FIRE DEPARTMENTS

Chapter 4 Organization

4.1 Fire Department Organizational Statement.

4.1.1* The authority having jurisdiction (AHJ) shall maintain a written statement or policy that establishes the following:

(1) Existence of the fire department

(2) Services that the fire department is required to provide

(3) Basic organizational structure

(4) Expected number of fire department members

(5) Functions that fire department members are expected to perform

4.1.2* The fire department organizational statement shall provide service delivery objectives, including specific time objectives for each major service component [i.e., fire suppression, emergency medical services (EMS), special operations, aircraft rescue and fire fighting, marine rescue and fire fighting, and/or wildland fire fighting] and objectives for the percentage of responses that meet the time objectives.

4.1.2.1 The fire department shall establish the following objectives:

(1) Alarm handling time to be completed in accordance with 4.1.2.3.

(2) 80 seconds for turnout time for fire and special operations response and 60 seconds turnout time for EMS response

(3)*240 seconds or less travel time for the arrival of the first arriving engine company at a fire suppression incident and 480 seconds or less travel time for the deployment of an initial full alarm assignment at a fire suppression incident

(4) 240 seconds or less travel time for the arrival of a unit with first responder with automatic external defibrillator (AED) or higher level capability at an emergency medical incident

(5) 480 seconds or less travel time for the arrival of an advanced life support (ALS) unit at an emergency medical incident, where this service is provided by the fire department provided a first responder with AED or basic life support (BLS) unit arrived in 240 seconds or less travel time

4.1.2.2 The fire department shall document the initiating action/intervention time.

4.1.2.3 Alarm Handling.

4.1.2.3.1 The fire department shall establish a performance objective of having an alarm answering time of not more than 15 seconds for at least 95 percent of the alarms received and not more than 40 seconds for at least 99 percent of the alarms received, as specified by NFPA 1221.

4.1.2.3.2 When the alarm is received at a public safety answering point (PSAP) and transferred to a secondary answering point or communication center, the agency responsible for the PSAP shall establish a performance objective of having an alarm transfer time of not more than 30 seconds for at least 95 percent of all alarms processed, as specified by NFPA 1221.

4.1.2.3.3 The fire department shall establish a performance objective of having an alarm processing time of not more than 60 seconds for at least 90 percent of the alarms and not more than 90 seconds for at least 99 percent of the alarms, as specified by NFPA 1221.

4.1.2.4 The fire department shall establish a performance objective of not less than 90 percent for the achievement of each turnout time and travel time objective specified in 4.1.2.1.

4.1.2.5 Evaluations.

4.1.2.5.1* The fire department shall evaluate its level of service and deployment delivery and alarm handling time, turnout time, and travel time objectives on an annual basis.

4.1.2.5.2* The evaluations shall be based on emergency incident data relating to level of service, deployment, and the achievement of each time objective in each geographic area within the jurisdiction of the fire department.

4.1.2.6 The fire department shall provide the AHJ with a written report annually.

4.1.2.6.1 The annual report shall define the geographic areas and/or circumstances in which the requirements of this standard are not being met.

4.1.2.6.2 The annual report shall explain the predictable consequences of these deficiencies and address the steps that are necessary to achieve compliance.

4.2 Fire Suppression Services. The fire department organizational statement shall set forth the criteria for the various types of fire suppression incidents to which the fire department is required to respond.

4.3 Emergency Medical Services.

4.3.1 The fire department organizational statement shall set forth the criteria for the various types of emergency medical incidents to which the fire department is required and/or expected to respond.

4.3.2 The fire department organizational statement shall ensure that the fire department's emergency medical response capability includes personnel, equipment, and resources to deploy at the first responder level with AED or higher treatment level.

4.3.3 Where emergency medical services beyond the first responder with AED level are provided by another agency or private organization, the AHJ, based on recommendations from the fire department, shall include the minimum staffing, deployment, and response criteria as required in Section 5.3 in the following:

(1) The fire department organizational statement

(2) Any contract, service agreement, governmental agreement, or memorandum of understanding between the AHJ and the other agency or private organization

4.4 Special Operations.

4.4.1 The fire department organizational statement shall set forth the criteria for the various types of special operations response and mitigation activities to which the fire department is required and/or expected to respond.

4.4.2* The fire department organizational statement shall ensure that the fire department's hazardous materials response capability includes personnel, equipment, and resources to deploy at the first responder operational level as required by 29 CFR 1910.120.

4.4.3 The fire department organizational statement shall ensure that the fire department's confined space response capability includes personnel, equipment, and resources to deploy at the confined space operational level as required by 29 CFR 1910.146.

4.4.4 The fire department organizational statement shall set forth the criteria for the various types of fire department response during natural disasters or terrorism incidents, weapons of mass destruction incidents, or large-scale or mass casualty events.

4.5 Airport Rescue and Fire-Fighting Services. The fire department organizational statement shall set forth the criteria for the various types of airport rescue and fire-fighting incidents to which the fire department is required and/or expected to respond.

4.6 Marine Rescue and Fire-Fighting Services. The fire department organizational statement shall set forth the criteria for the various types of marine rescue and fire-fighting incidents to which the fire department is required and/or expected to respond.

4.7 Wildland Fire Suppression Services. The fire department organizational statement shall set forth the criteria for the various types of wildland fire suppression incidents to which the fire department is required and/or expected to respond.

4.8 Intercommunity Organization.

4.8.1* Mutual aid, automatic aid, and fire protection agreements shall be through a written intergovernmental agreement and shall address issues such as liability for injuries and deaths, disability retirements, cost of service, authorization to respond, staffing, and equipment, including the resources to be made available, availability of interoperable communications, and the designation of the incident commander.

4.8.2 Procedures and training of personnel for all fire departments in mutual aid, automatic aid, and fire protection agreement plans shall be comprehensive to produce an effective fire force and to ensure uniform operations.

Chapter 5 Fire Department Services

5.1 Purpose.

5.1.1 The services provided by the fire department shall include those activities identified by the organizational statement developed as required by Chapter 4.

5.1.2 The procedures involved in providing these services, including operations and deployment, shall be established through written administrative regulations, standard operating procedures (SOPs), and departmental orders.

5.2* Fire Suppression Services.

5.2.1 Fire Suppression Capability.

5.2.1.1 Fire suppression operations shall be organized to ensure that the fire department's fire suppression capability encompasses deployment of personnel, equipment, and resources for an initial arriving company, the initial full alarm assignment, and additional alarm assignments.

5.2.1.2 The fire department shall be permitted to use established automatic aid and mutual aid agreements to comply with the requirements of Section 5.2.

5.2.2* Staffing. The number of on-duty fire suppression personnel shall be sufficient to perform the necessary fire-fighting operations given the expected fire-fighting conditions.

5.2.2.1 These numbers shall be determined through task analyses that take the following factors into consideration:

(1) Life hazard to the populace protected

(2) Provisions of safe and effective fire-fighting performance conditions for the fire fighters

(3) Potential property loss

(4) Nature, configuration, hazards, and internal protection of the properties involved

(5) Types of fireground tactics and evolutions employed as standard procedure, type of apparatus used, and results expected to be obtained at the fire scene

5.2.2.2* On-duty personnel assigned to fire suppression shall be organized into company units and shall have appropriate apparatus and equipment assigned to such companies.

5.2.2.2.1* The fire department shall identify minimum company staffing levels as necessary to meet the deployment criteria required in 5.2.4 to ensure that a sufficient number of members are assigned, on duty, and available to safely and effectively respond with each company.

5.2.2.2.2 Each company shall be led by an officer who shall be considered a part of the company.

5.2.2.2.3* Supervisory chief officers shall be dispatched or notified to respond to all full alarm assignments.

5.2.2.2.4 The supervisory chief officer shall ensure that the incident management system is established as required in Section 6.2.

5.2.2.2.5* Supervisory chief officers shall have staff aides deployed to them for purposes of incident management and accountability at emergency incidents.

5.2.3 Operating Units. Fire company staffing requirements shall be based on minimum levels necessary for safe, effective, and efficient emergency operations.

5.2.3.1 Fire companies whose primary functions are to pump and deliver water and perform basic fire fighting at fires, including search and rescue, shall be known as engine companies.

5.2.3.1.1 These companies shall be staffed with a minimum of four on-duty personnel.

5.2.3.1.2 In jurisdictions with tactical hazards, high-hazard occupancies, high incident frequencies, geographical restrictions, or other pertinent factors as identified by the AHJ, these companies shall be staffed with a minimum of five or six on-duty members.

5.2.3.2 Fire companies whose primary functions are to perform the variety of services associated with truck work, such as forcible entry, ventilation, search and rescue, aerial operations for water delivery and rescue, utility control, illumination, overhaul, and salvage work, shall be known as ladder or truck companies.

5.2.3.2.1 These companies shall be staffed with a minimum of four on-duty personnel.

5.2.3.2.2 In jurisdictions with tactical hazards, high-hazard occupancies, high incident frequencies, geographical restrictions, or other pertinent factors as identified by the AHJ, these companies shall be staffed with a minimum of five or six on-duty personnel.

5.2.3.3 Other Types of Companies.

5.2.3.3.1 Other types of companies equipped with specialized apparatus and equipment shall be provided to assist engine and ladder companies where necessary to support the fire departments' SOPs.

5.2.3.3.2 These companies shall be staffed with the minimum number of on-duty personnel required to deal with the tactical hazards, high-hazard occupancies, high incident frequencies, geographical restrictions, or other pertinent factors as identified by the AHJ.

5.2.3.4 Fire Companies with Quint Apparatus.

5.2.3.4.1 A fire company that deploys with quint apparatus, designed to operate as either an engine company or a ladder company, shall be staffed as specified in 5.2.3.

5.2.3.4.2 If the company is expected to perform multiple roles simultaneously, additional staffing, above the levels specified in 5.2.3, shall be provided to ensure that those operations can be performed as required.

5.2.4 Deployment.

5.2.4.1 Initial Arriving Company.

5.2.4.1.1 The fire department's fire suppression resources shall be deployed to provide for the arrival of an engine company within a 240-second travel time to 90 percent of the incidents as established in Chapter 4.

5.2.4.1.2* Personnel assigned to the initial arriving company shall have the capability to implement an initial rapid intervention crew (IRIC).

5.2.4.2 Initial Full Alarm Assignment Capability.

5.2.4.2.1 The fire department shall have the capability to deploy an initial full alarm assignment within a 480-second travel time to 90 percent of the incidents as established in Chapter 4.

5.2.4.2.2* The initial full alarm assignment to a structure fire in a typical 2000 ft2 (186 m2), two-story single-family dwelling without basement and with no exposures shall provide for the following:

(1) Establishment of incident command outside of the hazard area for the overall coordination and direction of the initial full alarm assignment with a minimum of one individual dedicated to this task

(2) Establishment of an uninterrupted water supply of a minimum of 400 gpm (1520 L/min) for 30 minutes with supply line(s) maintained by an operator

(3) Establishment of an effective water flow application rate of 300 gpm (1140 L/min) from two handlines, each of which has a minimum flow rate of 100 gpm (380 L/min) with each handline operated by a minimum of two individuals to effectively and safely maintain the line

(4) Provision of one support person for each attack and backup line deployed to provide hydrant hookup and to assist in laying of hose lines, utility control, and forcible entry (5) Provision of at least one victim search and rescue team with each such team consisting of a minimum of two individuals

(6) Provision of at least one team, consisting of a minimum of two individuals, to raise ground ladders and perform ventilation

(7) If an aerial device is used in operations, one person to function as an aerial operator and maintain primary control of the aerial device at all times

(8) Establishment of an IRIC consisting of a minimum of two properly equipped and trained individuals

5.2.4.2.3* Fire departments that respond to fires in high-, medium-, or low-hazard occupancies that present hazards greater than those found in the low-hazard occupancy described in 5.2.4.2.2 shall deploy additional resources on the initial alarm.

5.2.4.3 Additional Alarm Assignments.

5.2.4.3.1* The fire department shall have the capability to deploy additional alarm assignments that can provide for additional command staff, personnel, and additional services, including the application of water to the fire; engagement in search and rescue, forcible entry, ventilation, and preservation of property; safety and accountability for personnel; and provision of support activities for those situations that are beyond the capability of the initial full alarm assignment.

5.2.4.3.2 When an incident escalates beyond an initial full alarm assignment or when significant risk is present to the fire fighters due to the magnitude of the incident, the incident commander shall upgrade the IRIC to a full rapid intervention crew(s) (RIC) that consists of an officer and at least three firefighters who are fully equipped and trained in RIC operations.

5.2.4.3.3 An incident safety officer shall be deployed to all incidents that escalate beyond an initial full alarm assignment or when significant risk is present to fire fighters.

5.2.4.3.4 The incident safety officer shall ensure that the safety and health system is established as required in Section 6.1.

5.3* Emergency Medical Services (EMS). The purpose of this section shall be to provide standards for the delivery of EMS by fire departments.

5.3.1 The fire department shall clearly document its role, responsibilities, functions, and objectives for the delivery of EMS.

5.3.1.1 EMS operations shall be organized to ensure that the fire department's emergency medical capability includes personnel, equipment, and resources to deploy the initial arriving company and additional alarm assignments.

5.3.1.2 The fire department shall be permitted to use established automatic aid or mutual aid agreements to comply with the requirements of Section 5.3.

5.3.2* System Components.

5.3.2.1 Treatment Levels.

5.3.2.1.1 The basic treatment levels within an EMS system, for the purposes of this standard, shall be categorized as first responder, basic life support (BLS), and advanced life support (ALS).

5.3.2.1.2 The specific patient treatment capabilities associated with each level shall be determined by the AHJ based on the requirements for approval and licensing of EMS providers within each state or province.

5.3.2.2 Training Levels.

5.3.2.2.1 The minimal level of training for all fire fighters that respond to emergency incidents shall be to the first responder/AED level.

5.3.2.2.2 The AHJ shall determine if further training is required.

5.3.3 EMS System Functions.

5.3.3.1 The AHJ shall determine which of the following components of an EMS system the fire department shall be responsible for providing:

(1) Initial response to provide medical treatment at the location of the emergency (first responder with AED capability or higher)

(2) BLS response

(3) ALS response

(4) Patient transport in an ambulance or alternative vehicle designed to provide for uninterrupted patient care at the ALS or BLS level while en route to a medical facility

(5) Assurance of response and medical care through a quality management program

5.3.3.2 Staffing.

5.3.3.2.1 On-duty EMS units shall be staffed with the minimum personnel necessary for emergency medical care relative to the level of EMS provided by the fire department.

5.3.3.2.2 EMS staffing requirements shall be based on the minimum levels needed to provide patient care and member safety.

5.3.3.2.2.1 Units that provide emergency medical care shall be staffed at a minimum with personnel trained to the first responder/AED level.

5.3.3.2.2.2 Units that provide BLS transport shall be staffed and trained at the level prescribed by the state or provincial agency responsible for providing EMS licensing.

5.3.3.2.2.3 Units that provide ALS transport shall be staffed and trained at the level prescribed by the state or provincial agency responsible for providing EMS licensing.

5.3.3.3 Service Delivery Deployment.

5.3.3.3.1 The fire department shall adopt service delivery objectives based on time standards for the deployment of each service component for which it is responsible.

5.3.3.3.2 The fire department's EMS for providing a first responder with AED shall be deployed to provide for the arrival of a first responder with AED company within a 240-second travel time to 90 percent of the incidents as established in Chapter 4.

5.3.3.3.3* When provided, the fire department's EMS for providing ALS shall be deployed to provide for the arrival of an ALS company within a 480-second travel time to 90 percent of the incidents provided a first responder with AED or BLS unit arrived in 240 seconds or less travel time as established in Chapter 4.

5.3.3.3.4 Personnel deployed to ALS emergency responses shall include a minimum of two members trained at the emergency medical technician–paramedic level and two members trained at the emergency medical technician–basic level arriving on scene within the established travel time.

5.3.4 Quality Management.

5.3.4.1 The fire department shall institute a quality management program to ensure that the service has met time objectives as required in 4.1.2 for all medical responses.

5.3.4.2 Fire Department Medical Personnel Review.

5.3.4.2.1 All first responder and BLS medical care provided by the fire department shall be reviewed by the fire department medical personnel.

5.3.4.2.2 This review process shall be documented.

5.3.4.3 Medical Director Review.

5.3.4.3.1 All fire departments with ALS services shall have a named medical director with the responsibility to oversee and ensure quality medical care in accordance with state or provincial laws or regulations.

5.3.4.3.2 This review process shall be documented.

5.3.4.4 Fire departments providing ALS services shall provide a mechanism for immediate communications with EMS supervision and medical oversight.

5.4 **Special Operations Response**. Special operations shall be organized to ensure that the fire department’s special operations capability includes personnel, equipment, and resources to deploy the initial arriving company and additional alarm assignments providing such services.

5.4.1 The fire department shall be permitted to use established automatic aid or mutual aid agreements to comply with the requirements of Section 5.4.

5.4.2 The fire department shall adopt a special operations response plan and SOPs that specify the roles and responsibilities of the fire department and the authorized functions of members responding to hazardous materials emergency incidents.

5.4.3 All fire department members expected to respond to emergency incidents beyond the first responder operations level for hazardous materials response shall be trained to the applicable requirements of NFPA 472.

5.4.4 All fire department members expected to respond to emergency incidents beyond the confined space operations level for confined space operations shall be trained to the applicable requirements of NFPA 1670.

5.4.5 The fire department shall have the capacity to implement an RIC during all special operations incidents that would subject fire fighters to immediate danger or injury in the event of equipment failure or other sudden events, as required by NFPA 1500.

5.4.6 If a higher level of emergency response is needed beyond the capability of the fire department for special operations, the fire department shall determine the availability of outside resources that deploy these capabilities and the procedures for initiating their response.

5.4.7 The fire department shall limit its activities to only those specific special operations functions for which its personnel have been trained and are correctly equipped.

This page intentionally left blank.

ACTIVITY 5.2 (cont'd)

Deployment Analysis

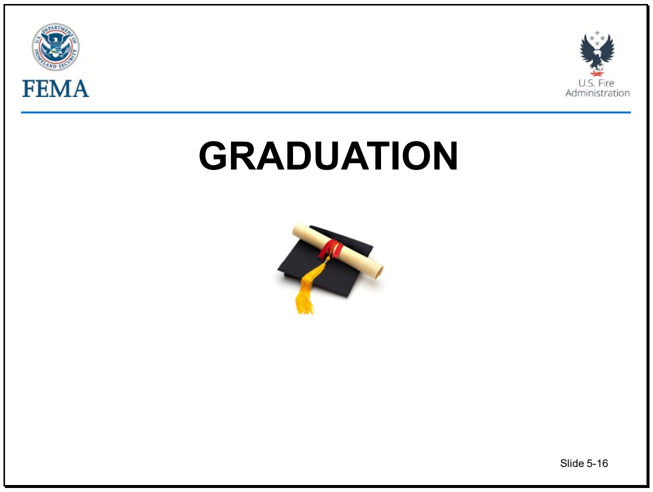
Name: ________________________________ Department: ______________________________

Task A: Your Current Deployment		
Minimum Task Necessary at a 2,000 Square-Foot Residential Structure Fire		
Task Identification	**Firefighters**	**Company Assignment**
Total Personnel		

Task B: Your Ideal Deployment		
Minimum Task Necessary at a 2,000 Square-Foot Residential Structure Fire		
Task Identification	**Firefighters**	**Company Assignment**
Total Personnel		

This page intentionally left blank.

III. SUMMARY



FEMA

U.S. Fire Administration

GRADUATION

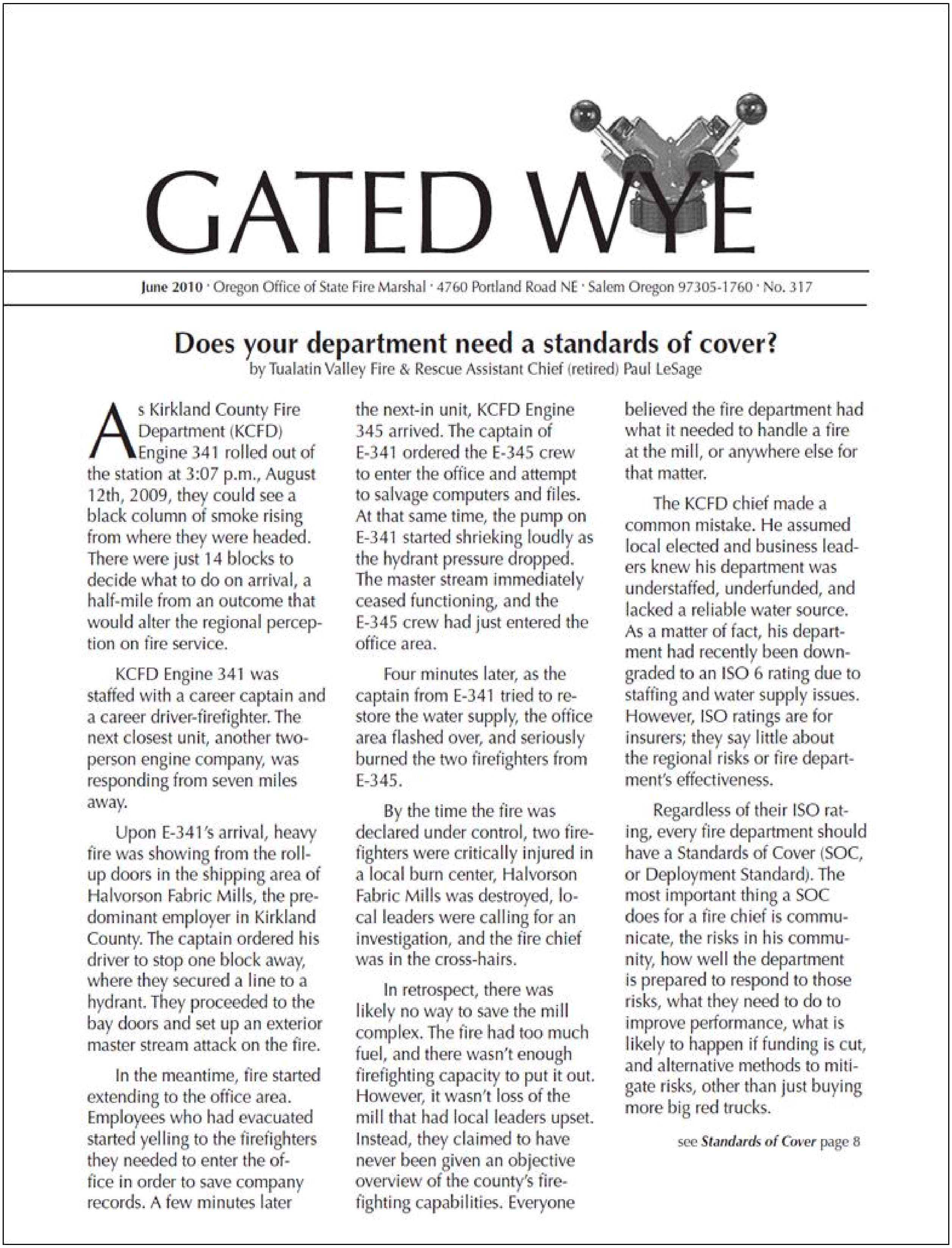
Slide 5-16

This page intentionally left blank.

APPENDIX A

GATED WYE ARTICLE

This page intentionally left blank.



GATED WYE

June 2010 · Oregon Office of State Fire Marshal · 4760 Portland Road NE · Salem Oregon 97305-1760 · No. 317

Does your department need a standards of cover?

by Tualatin Valley Fire & Rescue Assistant Chief (retired) Paul LeSage

As Kirkland County Fire Department (KCFD) Engine 341 rolled out of the station at 3:07 p.m., August 12th, 2009, they could see a black column of smoke rising from where they were headed. There were just 14 blocks to decide what to do on arrival, a half-mile from an outcome that would alter the regional perception on fire service.

KCFD Engine 341 was staffed with a career captain and a career driver-firefighter. The next closest unit, another two-person engine company, was responding from seven miles away.

Upon E-341's arrival, heavy fire was showing from the roll-up doors in the shipping area of Halvorson Fabric Mills, the predominant employer in Kirkland County. The captain ordered his driver to stop one block away, where they secured a line to a hydrant. They proceeded to the bay doors and set up an exterior master stream attack on the fire.

In the meantime, fire started extending to the office area. Employees who had evacuated started yelling to the firefighters they needed to enter the office in order to save company records. A few minutes later the next-in unit, KCFD Engine 345 arrived. The captain of E-341 ordered the E-345 crew to enter the office and attempt to salvage computers and files. At that same time, the pump on E-341 started shrieking loudly as the hydrant pressure dropped. The master stream immediately ceased functioning, and the E-345 crew had just entered the office area.

Four minutes later, as the captain from E-341 tried to restore the water supply, the office area flashed over, and seriously burned the two firefighters from E-345.

By the time the fire was declared under control, two firefighters were critically injured in a local burn center, Halvorson Fabric Mills was destroyed, local leaders were calling for an investigation, and the fire chief was in the cross-hairs.

In retrospect, there was likely no way to save the mill complex. The fire had too much fuel, and there wasn't enough firefighting capacity to put it out. However, it wasn't loss of the mill that had local leaders upset. Instead, they claimed to have never been given an objective overview of the county's firefighting capabilities. Everyone believed the fire department had what it needed to handle a fire at the mill, or anywhere else for that matter.

The KCFD chief made a common mistake. He assumed local elected and business leaders knew his department was understaffed, underfunded, and lacked a reliable water source. As a matter of fact, his department had recently been downgraded to an ISO 6 rating due to staffing and water supply issues. However, ISO ratings are for insurers; they say little about the regional risks or fire department's effectiveness.

Regardless of their ISO rating, every fire department should have a Standards of Cover (SOC, or Deployment Standard). The most important thing a SOC does for a fire chief is communicate, the risks in his community, how well the department is prepared to respond to those risks, what they need to do to improve performance, what is likely to happen if funding is cut, and alternative methods to mitigate risks, other than just buying more big red trucks.

see *Standards of Cover* page 8

Fire & life safety recognition

Photo by Gert Zoutendijk

Lake Oswego Fire Department Lieutenant Steven DeHart receives his Fire and Life Safety Specialist II recognition certificate from Supervising Deputy State Fire Marshal Dave Jones in May.

Photo by Gert Zoutendijk

Albany Fire Department Lieutenant Donnie Schlies receives his Oregon Fire Marshal with Fire Plans Examiner recognition certificate from Supervising Deputy State Fire Marshal Dave Jones at the Oregon Fire Marshals Association annual business meeting in May.

Standards of Cover

continued from front page

In the case of KCFD, an SOC would have pointed out:

- The largest economic risk in the county had no fire suppression system.
- The water system could not provide more than 1,000 GPM for firefighting.
- The fire units were understaffed during the daytime due to a lack of volunteers.
- The department was essentially an 'exterior attack only' department until enough manpower arrived, which could take up to 20 minutes.
- The department's staffing regularly dropped to zero when ambulance calls had to be run (the fire department also ran the local ambulance).
- The department had no prevention or inspection services due to their budget.
- There were no business continuity plans to help larger businesses understand their own responsibilities in the event of a fire (like saving records first).
- There were no analyses of staffing or tasks that demonstrated what a minimum effective firefighting force was for Kirkland County.
- There was no plan of engagement for firefighters that considered NFPA, OSHA, IAFC, IAFF, and other relevant standards.

Do any of these sound familiar? Could a similar story be written about your community?

A standards of cover need not be a book. Many smaller departments have successfully addressed their community risk, ability to respond, rules of engagement, and limits of service in 20 pages or less. When it's complete, policy-makers, the community, firefighters, and the fire chief have an objective document clearly stating what can be expected from the local fire department.

To see some examples, visit the websites for the Oregon Office of State Fire Marshal or the Oregon Fire Chiefs Association. If you have any questions related to developing standards of cover that aren't answered by those documents, please feel free to e-mail me at Paul@cdm-hro.com.

See the DATA Connection column on page six for more information on key elements of standards of cover.

June 2010, GATED WYE, page 8

APPENDIX B

NFPA REQUIREMENTS DEFINITIONS

This page intentionally left blank.

3.3.53 Time.

3.3.53.1 Alarm Answering Time. The time interval that begins when the alarm is received at the communication center and ends when the alarm is acknowledged at the communication center.

3.3.53.2 Alarm Handling Time. The time interval from the receipt of the alarm at the primary PSAP until the beginning of the transmittal of the response information via voice or electronic means to emergency response facilities (ERFs) or the emergency response units (ERUs) in the field.

3.3.53.3 Alarm Processing Time. The time interval from when the alarm is acknowledged at the communication center until response information begins to be transmitted via voice or electronic means to emergency response facilities (ERFs) and emergency response units (ERUs).

3.3.53.4 Alarm Transfer Time. The time interval from the receipt of the emergency alarm at the PSAP until the alarm is first received at the communication center.

3.3.53.5* Initiating Action/Intervention Time. The time interval from when a unit arrives on the scene to the initiation of emergency mitigation.

3.3.53.6* Total Response Time. The time interval from the receipt of the alarm at the primary PSAP to when the first emergency response unit is initiating action or intervening to control the incident.

3.3.53.7 Travel Time. The time interval that begins when a unit is en-route to the emergency incident and ends when that unit arrives at the scene.

3.3.53.8 Turnout Time. The time interval that begins when the emergency response facilities (ERFs) and emergency response units (ERUs) notification process begins by either an audible alarm or visual annunciation or both and ends at the beginning point of travel time.

4.1.2.3 Alarm Handling.

4.1.2.3.1 The fire department shall establish a performance objective of having an alarm answering time of not more than 15 seconds for at least 95 percent of the alarms received and not more than 40 seconds for at least 99 percent of the alarms received, as specified by NFPA 1221.

4.1.2.3.2 When the **alarm** is received at a public safety answering point (PSAP) and transferred to a secondary answering point or communication center, the agency responsible for the PSAP shall establish a performance objective of having an alarm transfer time of not more than 30 seconds for at least 95 percent of all alarms processed, as specified by NFPA 1221.

4.1.2.3.3 The fire **department** shall establish a performance objective of having an alarm processing time of not more than 60 seconds for at least 90 percent of the alarms and not more than 90 seconds for at least 99 percent of the alarms, as specified by NFPA 1221.

4.1.2.4 The fire department shall establish a performance objective of not less than 90 percent for the achievement of each turnout time and travel time objective specified in 4.1.2.1.

4.2 Community Risk Management.

The fire department shall participate in a process that develops a community fire and Emergency Medical Services (EMS) risk management plan.

4.2.1 The specific role of the fire department and other responding agencies shall be defined by the community risk management plan.

4.2.2* The number and type of units assigned to respond to a reported incident shall be determined by risk analysis and/or pre-fire planning.

4.6 Initial Fire-Fighting Operations.

4.6.1 Initial fire-fighting operations shall be organized to ensure that at least four members are assembled before interior fire suppression operations are initiated in a hazardous area.

4.6.2 In the hazardous area, a minimum of two members shall work as a team.

4.6.3* Outside the hazardous area, a minimum of two members shall be present for assistance or rescue of the team operating in the hazardous area.

4.6.3.1 One of the two members assigned outside the hazardous area shall be permitted to be engaged in other activities.

4.6.3.2 The assignment of a member shall not be permitted if abandoning that member’s critical task(s) to perform rescue would jeopardize the safety and health of any fire fighter operating at the incident.

4.8.1 Mutual Aid.

4.8.1* Mutual aid, automatic aid, and fire protection agreements among the affected AHJs shall be in writing and shall address issues such as liabilities for injuries, disabilities, and deaths; cost of service; authorization to respond; staffing; and equipment, including the resources to be made available and the designation of the incident commander.

4.8.2 Procedures and training of personnel for all fire departments in mutual aid, automatic aid, and fire protection agreement plans shall be comprehensive enough to produce an effective force to deal with the emergencies they respond to and to ensure uniform operations at those emergencies.

4.8.3 Companies responding to automatic or mutual aid incidents shall be equipped with communications equipment that allow personnel to communicate with the incident commander, division or group supervisors, or branch directors.

4.9.1 EMS Response.

4.9.1* The provisions of this section shall apply only to those fire departments that are involved in EMS delivery.

4.9.2* The fire department shall clearly document its role, responsibilities, functions, and objectives for the delivery of EMS.

4.9.3 EMS operations shall be organized to ensure the fire department’s emergency medical capability includes personnel, equipment, and resources to deploy the initial arriving company and additional alarm assignments.

4.9.4 The fire department shall be permitted to use established automatic aid or mutual aid agreements to comply with the requirements of Section 4.9.

4.9.6 EMS Quality.

4.9.6.2 All first responder and basic life support (BLS) emergency medical service provided by the fire department shall be reviewed by the fire department medical personnel and that review process shall be documented.

4.9.6.3 All fire departments with ALS services shall have a named medical director with the responsibility to oversee and ensure quality medical care in accordance with state or provincial laws or regulations.

4.9.6.4 Fire departments providing ALS services shall provide a mechanism for immediate communications with EMS supervision and medical oversight.

5.2.4.2 Initial Full Alarm Assignment Capability.

5.2.4.2.1* The fire department shall have the capability to deploy an initial full alarm assignment within an 8-minute response time to 90 percent of the incidents as established in Chapter 4.

5.2.4.2.2 The initial full alarm assignment shall provide for the following:

(1) Establishment of incident command outside of the hazard area for the overall coordination and direction of the initial full alarm assignment. A minimum of one individual shall be dedicated to this task.

(2) Establishment of an uninterrupted water supply of a minimum 1520 L/min (400 gpm) for 30 minutes. Supply line(s) shall be maintained by an operator who shall ensure uninterrupted water flow application.

(3) Establishment of an effective water flow application rate of 1140 L/min (300 gpm) from two handlines, each of which shall have a minimum of 380 L/min (100 gpm). Each attack and backup line shall be operated by a minimum of two individuals to effectively and safely maintain the line.

(4) Provision of one support person for each attack and backup line deployed to provide hydrant hookup and to assist in line lays, utility control, and forcible entry.

(5) A minimum of one victim search and rescue team shall be part of the initial full alarm assignment. Each search and rescue team shall consist of a minimum of two individuals.

(6) A minimum of one ventilation team shall be part of the initial full alarm assignment. Each ventilation team shall consist of a minimum of two individuals.

(7) If an aerial device is used in operations, one person shall function as an aerial operator who shall maintain primary control of the aerial device at all times.

(8) Establishment of an IRIC that shall consist of a minimum of two properly equipped and trained individuals.

5.2.4.3 Additional Alarm Assignments.

5.2.4.3.1 The fire department shall have the capability for additional alarm assignments that can provide for additional personnel and additional services, including the application of water to the fire; engagement in search and rescue, forcible entry, ventilation, and preservation of property; accountability for personnel; and provision of support activities for those situations that are beyond the capability of the initial full alarm assignment.

5.2.4.3.2 When an incident escalates beyond an initial full alarm assignment or when significant risk is present to fire fighters due to the magnitude of the incident, the incident commander shall upgrade the IRIC to a full rapid intervention crew(s) (RICs) that consist(s) of four fully equipped and trained fire fighters.

5.2.4.3.3 An incident safety officer shall be deployed to all incidents that escalate beyond an initial full alarm assignment or when significant risk is present to fire fighters.

5.2.4.3.4 The incident safety officer shall ensure that the safety and health system is established as required in Section 6.1.

5.3* Emergency Medical Services.

5.3.1 Purpose. EMS operations shall be organized to ensure that the fire department's emergency medical capability includes personnel, equipment, and resources to deploy the initial arriving company and additional alarm assignments.

5.3.1.1 The fire department shall be permitted to use established automatic mutual aid or mutual aid agreements to comply with the requirements of Section 5.3.

5.3.1.2 The purpose of this section shall be to provide standards for the delivery of EMS by fire departments.

5.3.1.3 The fire department shall clearly document its role, responsibilities, functions, and objectives for the delivery of EMS.

5.3.2.1 Treatment Levels.

5.3.2.1.1 The basic treatment levels within an EMS system, for the purposes of this standard, shall be categorized as first responder, BLS, and advanced life support (ALS).

5.3.2.1.2 The specific patient treatment capabilities associated with each level shall be determined by the AHJ for the approval and licensing of EMS providers within each state or province.

5.3.2.2 Training Levels.

5.3.2.2.1 The minimal level of training for all fire fighters that respond to emergency incidents shall be to the first responder/AED level.

5.3.2.2.2 The AHJ shall determine if further training is required.

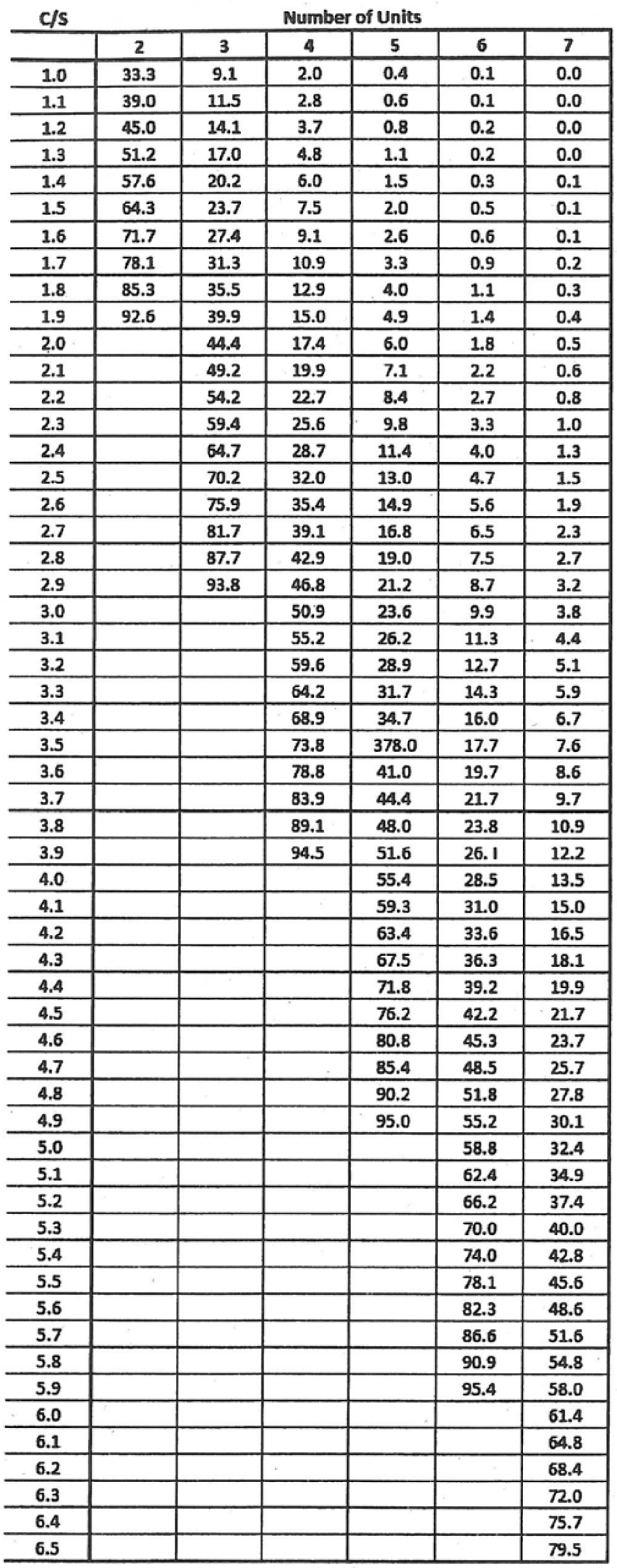
This page intentionally left blank.

APPENDIX C

PROBABILITY OF DELAY TABLE

This page intentionally left blank.

Probability of Delay



C/S	Number of Units					
	2	3	4	5	6	7
1.0	33.3	9.1	2.0	0.4	0.1	0.0
1.1	39.0	11.5	2.8	0.6	0.1	0.0
1.2	45.0	14.1	3.7	0.8	0.2	0.0
1.3	51.2	17.0	4.8	1.1	0.2	0.0
1.4	57.6	20.2	6.0	1.5	0.3	0.1
1.5	64.3	23.7	7.5	2.0	0.5	0.1
1.6	71.7	27.4	9.1	2.6	0.6	0.1
1.7	78.1	31.3	10.9	3.3	0.9	0.2
1.8	85.3	35.5	12.9	4.0	1.1	0.3
1.9	92.6	39.9	15.0	4.9	1.4	0.4
2.0		44.4	17.4	6.0	1.8	0.5
2.1		49.2	19.9	7.1	2.2	0.6
2.2		54.2	22.7	8.4	2.7	0.8
2.3		59.4	25.6	9.8	3.3	1.0
2.4		64.7	28.7	11.4	4.0	1.3
2.5		70.2	32.0	13.0	4.7	1.5
2.6		75.9	35.4	14.9	5.6	1.9
2.7		81.7	39.1	16.8	6.5	2.3
2.8		87.7	42.9	19.0	7.5	2.7
2.9		93.8	46.8	21.2	8.7	3.2
3.0			50.9	23.6	9.9	3.8
3.1			55.2	26.2	11.3	4.4
3.2			59.6	28.9	12.7	5.1
3.3			64.2	31.7	14.3	5.9
3.4			68.9	34.7	16.0	6.7
3.5			73.8	378.0	17.7	7.6
3.6			78.8	41.0	19.7	8.6
3.7			83.9	44.4	21.7	9.7
3.8			89.1	48.0	23.8	10.9
3.9			94.5	51.6	26.1	12.2
4.0				55.4	28.5	13.5
4.1				59.3	31.0	15.0
4.2				63.4	33.6	16.5
4.3				67.5	36.3	18.1
4.4				71.8	39.2	19.9
4.5				76.2	42.2	21.7
4.6				80.8	45.3	23.7
4.7				85.4	48.5	25.7
4.8				90.2	51.8	27.8
4.9				95.0	55.2	30.1
5.0					58.8	32.4
5.1					62.4	34.9
5.2					66.2	37.4
5.3					70.0	40.0
5.4					74.0	42.8
5.5					78.1	45.6
5.6					82.3	48.6
5.7					86.6	51.6
5.8					90.9	54.8
5.9					95.4	58.0
6.0						61.4
6.1						64.8
6.2						68.4
6.3						72.0
6.4						75.7
6.5						79.5

APPENDIX D

QUEUE LENGTH TABLE

This page intentionally left blank.

Queue Length

C/S	Number of Units					
	2	3	4	5	6	7
1	0.33	0.05	0.01	0.00	0.00	0.00
1.1	0.48	0.07	0.01	0.00	0.00	0.00
1.2	0.68	0.09	0.02	0.00	0.00	0.00
1.3	0.95	0.13	0.02	0.00	0.00	0.00
1.4	1.35	0.18	0.03	0.01	0.00	0.00
1.5	1.00	0.03	0.24	0.04	0.00	0.00
1.6	2.84	0.31	0.06	0.01	0.00	0.00
1.7	4.43	0.41	0.08	0.02	0.00	0.00
1.8	7.67	0.53	0.11	0.02	0.00	0.00
1.9	17.59	0.69	0.14	0.03	0.01	0.00
2		0.89	0.17	0.04	0.01	0.00
2.1		1.15	0.22	0.05	0.01	0.00
2.2		1.49	0.28	0.07	0.02	0.00
2.3		1.95	0.35	0.08	0.02	0.00
2.4		2.59	0.43	0.10	0.03	0.01
2.5		3.51	0.53	0.13	0.03	0.01
2.6		4.93	0.66	0.16	0.04	0.01
2.7		7.35	0.81	0.20	0.05	0.01
2.8		12.27	1.00	0.24	0.07	0.02
2.9		27.19	1.23	0.29	0.08	0.02
3			1.53	0.35	0.10	0.03
3.1			1.90	0.43	0.12	0.03
3.2			2.39	0.51	0.15	0.04
3.3			3.03	0.62	0.17	0.05
3.4			3.91	0.74	0.21	0.06
3.5			5.17	0.88	0.25	0.08
3.6			7.09	1.06	0.29	0.09
3.7			10.35	1.26	0.35	0.11
3.8			16.94	1.52	0.41	0.13
3.9			36.86	1.83	0.48	0.15
4				2.22	0.57	0.18
4.1				2.70	0.67	0.20
4.2				3.33	0.78	0.25
4.3				4.15	0.92	0.29
4.4				5.27	1.08	0.34
4.5				6.86	1.26	0.39
4.6				9.29	1.49	0.45
4.7				13.38	1.75	0.53
4.8				21.64	2.07	0.61
4.9				46.57	2.46	0.70
5					2.94	0.81
5.1					3.54	0.94
5.2					4.30	1.08
5.3					5.30	1.25
5.4					6.66	1.44
5.5					8.59	1.67
5.6					11.52	1.94
5.7					16.45	2.26
5.8					26.37	2.65
5.9					56.30	3.11
6						3.68
6.1						4.39
6.2						5.30
6.3						6.48
6.4						8.08
6.5						10.34

This page intentionally left blank.

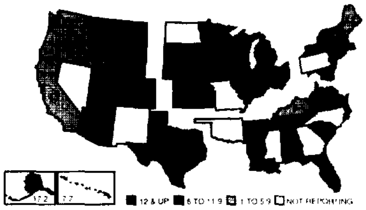
APPENDIX E

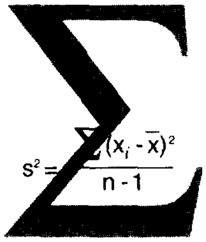
FIRE DATA ANALYSIS HANDBOOK

This page intentionally left blank.

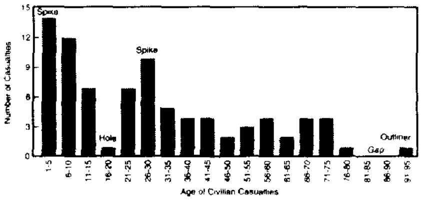
Fire Data Analysis Handbook

By
Tom McEwen
with
Catherine A. Miller





$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$



Federal Emergency Management Agency
United States Fire Administration

Foreword

The fire service exists today in an environment constantly inundated with data, but data are often of little use in the everyday, real world in which first responders live and work. This is no accident. By itself, data are of little use to anyone. Information, on the other hand, is very useful indeed. What's the difference? At the recent Olympic games in Korea, a stadium full of people held up individual, multi-colored squares of cardboard to form a giant image or text which could only be recognized from a distance. This is a good analogy for data and information. The individual squares of cardboard are like data They are very numerous and they all look pretty much alike taken by themselves. The big image formed from the organization of thousands of those cards is like information. It is what can be built from many pieces of data. Information then is an organization of data that makes a point about something.

The fire service of today is changing. More and more, it is not fighting tires as much as it is doing EMS, hazmat, inspections, investigations, prevention, and other non-traditional but important tasks which are vital to the community. Balancing limited resources and justifying daily operations and finances in the face of tough economic times is a scenario that every department can relate to. How well a department can do this depends mainly on how well it uses information.

Turning data into information is neither simple nor easy. It requires some knowledge of the tools and techniques used for this purpose. Historically, the fire service has had few of these tools at its disposal and none of them has been designed with the fire service in mind. This book changes that. It was designed solely for the use of the fire service. The examples and problems were developed from fire data collected from departments all over the nation. This book was also designed to be modular in form. Many departments' information needs can be met by studying only the first few chapters. Others with a more statistical bent may want to dig deeper. The point is, it's up to the reader to decide. This book is just another tool, like a pumper or a ladder, to help do the job.

The United States Fire Administration

Table of Contents

Chapter 1

INTRODUCTION

This handbook has a primary objective to describe statistical techniques for analyzing data typically collected in fire departments. Motivation for the handbook comes from the belief that fire departments collect an immense amount of data, but do very little with the data. Think for a minute about the reports you complete on incidents. You probably document the type of situation found, action taken, time of alarm, time of arrival, time completed, number of engines responding, number of personnel responding, and many other items. For fires, the list grows even longer to include area of fire origin, form of heat of ignition, type of material involved, and other related facts. If civilian or fire fighter injuries occur, you complete other reports.

A compelling reason for these reports is a legal requirement for documenting incidents. Victims, insurance companies, lawyers, and many others want copies of reports. Indeed, fire departments maintain files for retrieval of individual reports.

The reports can, however, provide a more beneficial service to fire departments by providing insight into the nature of fires and injuries in your jurisdiction. Basic information is probably already available. Someone usually tracks the number of fires handled last year, the number of fire-related injuries, and the number of fire deaths. It is another story, however, if you ask more probing questions:

- How many fires took place on Sundays, Mondays, etc.?
- How many fires took place each hour of the day or month of the year?
- What was the average response time to fires? How much did response times vary by fire station areas?
- What was the average time spent at the fire scene and how much did the average vary by type of fire?

In a nutshell, this handbook describes statistical techniques to turn data into information for answering these questions and many others. The techniques range from simple to complex. For example, the next two chapters describe how to develop charts to provide more effective presentations about fire problems. These charts may be beneficial to city or county officials on the activities and needs of your fire department. Chapter 4 tells how to compute simple statistics, such as means, medians, and modes. In Chapter 5, we discuss tables and how to calculate different percentages from tables. Other chapters present more sophisticated techniques, such as correlation, regression, loglinear analysis, and queueing theory. These are all techniques which can tell you more about the nature of fires and injuries.

One way to become more comfortable with analysis is to work with real data. For this handbook, data were obtained from fire departments in sevcral metropolitan areas, including Seattle, Washington; Chicago, Illinois; Detroit, Michigan; Jacksonville, Florida; Los Angeles County, California; Monroe County, New York; Boston, Massachusetts; and Dallas, ‘Texas. Data on medical emergencies were obtained from the fire department in Prince William County, Virginia, which has completed detailed reports on its responses since 1989. By working with real data, it should be easier for you to understand different techniques.

Why Data Analysis?

There still may be a question in your mind as to why we should go to all this trouble to analyze data. Many decisions do not require analysis, such as decisions on personnel, grievance proceedings, promotions, and even decisions on how to handle a fire. It is certainly true that fire departments can continue to operate in the same way they always have without doing a lot of analysis.

On the other hand, we can give three good reasons for looking more closely at your data: (I) to gain insights into fire problems, (2) to improve resource allocation for combatting fires, and (3) to identify training needs. Probably the most compelling is that analysis gives insight into your fire problems which, in turn, can impact operations in your department. You may find, for example, that the average time to fires in an area is 6 minutes, compared to less than 2 minutes overall. This result may assist you in requests for more equipment, more personnel, or justifing another fire station.

As an example of improved resource allocation, statistical analysis of emergency medical calls can determine the impact of providing another paramedic unit in the field. Increasing the number of EMS units from 4 units to 5 units may, for example, *decrease* average response times from 5 minutes to 3 minutes-a change that may save lives. Chapter 9 describes a queueing model for conducting this type of analysis.

Another reason for analysis is to identify training needs. Most training on fire fighting is based on a curriculum which has been in place for many years. It makes sense to see how training matches characteristics of fires in your own jurisdiction. This is not to sap that other training is unimportant, because an exception can always occur. However, knowing more about your fires can improve your training.

In summary, this handbook will help you deal with the volume of data collected on fire incidents. By studying the techniques presented in this handbook, you should be able to improve your skills in collecting data, analyzing data, and presenting results.

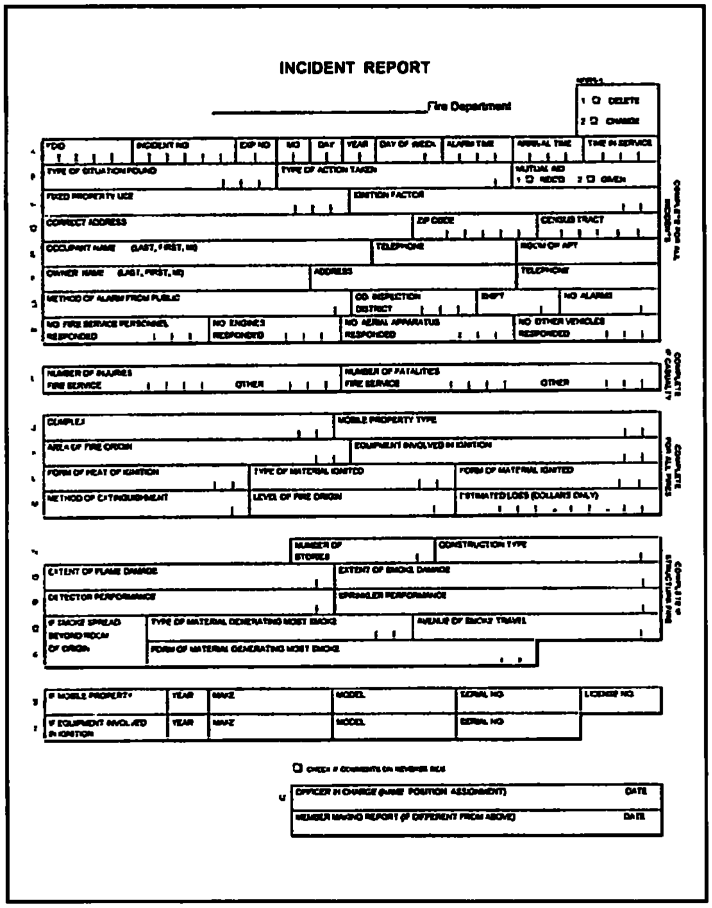
The nest section of this chapter describes the reporting system that serves as the basis for data collection on fires and casualties. ‘The importance of quality control is also discussed.

National Fire Incident Reporting System

The National Fire Incident Reporting System (NFIRS) began over 15 years ago with the aim of collecting and analyzing data on fires from departments across the country. More than 13,000 fire departments now report their fires and injuries to NFIRS.

Exhibit 1-1 shows the basic fire incident report from NFIRS. Your department may use this incident report, or you may have a modified version of it. In either case, the data collected is the same and covers all the elements of fire incidents. Lines A through I are completed on all incidents to which a fire department responds. These lines include incident number, date, day of week, alarm time, arrival time, time in service, type of situation found, and type of action taken.

Exhibit 1-1 Basic Fire Incident Report-NFIRS



INCIDENT REPORT

NFIRS-1

1 ☐ DELETE
2 ☐ CHANGE

_______________ Fire Department

FDID | INCIDENT NO | EXP NO | MO | DAY | YEAR | DAY OF WEEK | ALARM TIME | ARRIVAL TIME | TIME IN SERVICE
TYPE OF SITUATION FOUND | TYPE OF ACTION TAKEN | MUTUAL AID 1 ☐ RECD 2 ☐ GIVEN
FIXED PROPERTY USE | IGNITION FACTOR
CORRECT ADDRESS | ZIP CODE | CENSUS TRACT
OCCUPANT NAME (LAST, FIRST, MI) | TELEPHONE | ROOM OR APT
OWNER NAME (LAST, FIRST, MI) | ADDRESS | TELEPHONE
METHOD OF ALARM FROM PUBLIC | CO. INSPECTION DISTRICT | SHIFT | NO ALARMS
NO FIRE SERVICE PERSONNEL RESPONDED | NO ENGINES RESPONDED | NO AERIAL APPARATUS RESPONDED | NO OTHER VEHICLES RESPONDED

COMPLETE FOR ALL INCIDENTS

NUMBER OF INJURIES FIRE SERVICE OTHER | NUMBER OF FATALITIES FIRE SERVICE OTHER

COMPLETE IF CASUALTY

COMPLEX | MOBILE PROPERTY TYPE
AREA OF FIRE ORIGIN | EQUIPMENT INVOLVED IN IGNITION
FORM OF HEAT OF IGNITION | TYPE OF MATERIAL IGNITED | FORM OF MATERIAL IGNITED
METHOD OF EXTINGUISHMENT | LEVEL OF FIRE ORIGIN | ESTIMATED LOSS (DOLLARS ONLY)

COMPLETE FOR ALL FIRES

NUMBER OF STORIES | CONSTRUCTION TYPE
EXTENT OF FLAME DAMAGE | EXTENT OF SMOKE DAMAGE
DETECTOR PERFORMANCE | SPRINKLER PERFORMANCE
IF SMOKE SPREAD BEYOND ROOM OF ORIGIN | TYPE OF MATERIAL GENERATING MOST SMOKE | AVENUE OF SMOKE TRAVEL
FORM OF MATERIAL GENERATING MOST SMOKE

COMPLETE IF STRUCTURE FIRE

IF MOBILE PROPERTY | YEAR | MAKE | MODEL | SERIAL NO | LICENSE NO
IF EQUIPMENT INVOLVED IN IGNITION | YEAR | MAKE | MODEL | SERIAL NO

☐ CHECK IF COMMENTS ON REVERSE SIDE

OFFICER IN CHARGE (NAME POSITION ASSIGNMENT) | DATE
MEMBER MAKING REPORT (IF DIFFERENT FROM ABOVE) | DATE

Lines J through M are completed on all incidents which arc, in fact, fires. They include type of complex, mobile property type, area of fire origin, equipment involved in ignition, from of heat of ignition, dollar loss, and others. Finally, Lines N through R are completed on all structure fires. They include number of stories, construction type, extent of flame damage, detector performance, sprinkler performance, type of material generating the most smoke, avenue of smoke travel, and form of material generating the most smoke. The last two lines (Lines S and T) are for mobile property and equipment involved.

Most items are recorded as numeric codes with the NFPA 901 Codes[1] serving as the coding source. For example, Exhibit 1-2 shows the codes for extent of flame damage (Line 0) of the incident form.

Exhibit 1-2 Extent of Flame Damage

Categories for Extent of Flame Damage	Code
Confined to Object of Origin	1
Confined to Part of Room or Area of Origin	2
Confined to Room of Origin	3
Confined to Fire-rated Compartment of Origin	4
Confined to Floor of Origin	5
Confined to Structure of Origin	6
Extended Beyond Structure of Origin	7
Undetermined/Not Reported	0

The code numbers have no meaning by themselves, hut instead serve as a way of getting data into a computer in a compact and logical form. In Chaptcr 5, we will discuss these codes further and present how to develop and interpret tables based on them.

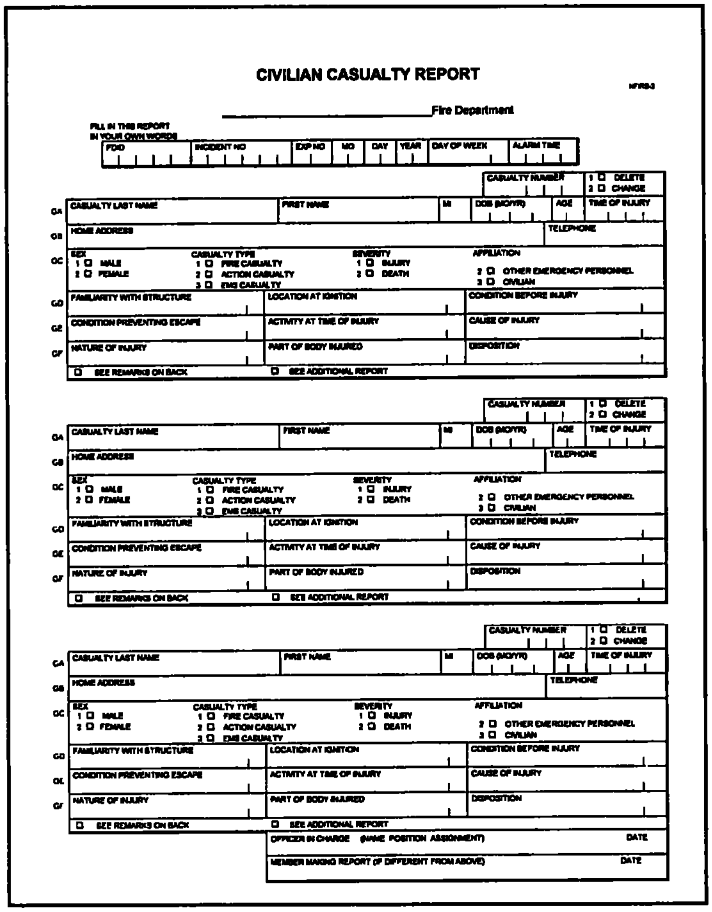
Exhibit 1-3 is the civilian casualty report which is part of NFIRS. Each form allows for recording three casualties. Variables collected for a casualty include incident number, date, day of week, alarm time, casualty name, age, time of injury, sex, severity, familiarity with structure, location at ignition, and condition before injury.

The last form in NFIRS is for fire tighter casualties (Exhibit 1-4). Fire departments complete this form whenever an injury to a fire fighter occurs. This form includes the fire fighter's age, injury severity, part of body injured, activity prior to injury, cause of injury, and medical care provided. In addition, it contains a section on protective equipment for the fire fighter (coat, trousers, boots/shoes, helmet, face protection, breathing apparatus, and gloves).

---

1. For more informatlon on these codes. see *NFPA 901, Uniform Coding for Fire Protection 1976* (National Fire Protection Association. Batterymarch Park. Quincy. Massachusetts. 02269).

Exhibit 1-3 Civilian Casualty Report



CIVILIAN CASUALTY REPORT

NFIRS-3

________________ Fire Department

FILL IN THIS REPORT
IN YOUR OWN WORDS

FDID	INCIDENT NO	EXP NO	MO	DAY	YEAR	DAY OF WEEK	ALARM TIME

CASUALTY NUMBER — 1 ☐ DELETE 2 ☐ CHANGE

GA CASUALTY LAST NAME | FIRST NAME | MI | DOB (MO/YR) | AGE | TIME OF INJURY
GB HOME ADDRESS | TELEPHONE
GC SEX 1 ☐ MALE 2 ☐ FEMALE | CASUALTY TYPE 1 ☐ FIRE CASUALTY 2 ☐ ACTION CASUALTY 3 ☐ EMS CASUALTY | SEVERITY 1 ☐ INJURY 2 ☐ DEATH | AFFILIATION 2 ☐ OTHER EMERGENCY PERSONNEL 3 ☐ CIVILIAN
GD FAMILIARITY WITH STRUCTURE | LOCATION AT IGNITION | CONDITION BEFORE INJURY
GE CONDITION PREVENTING ESCAPE | ACTIVITY AT TIME OF INJURY | CAUSE OF INJURY
GF NATURE OF INJURY | PART OF BODY INJURED | DISPOSITION
☐ SEE REMARKS ON BACK ☐ SEE ADDITIONAL REPORT

CASUALTY NUMBER — 1 ☐ DELETE 2 ☐ CHANGE

GA CASUALTY LAST NAME | FIRST NAME | MI | DOB (MO/YR) | AGE | TIME OF INJURY
GB HOME ADDRESS | TELEPHONE
GC SEX 1 ☐ MALE 2 ☐ FEMALE | CASUALTY TYPE 1 ☐ FIRE CASUALTY 2 ☐ ACTION CASUALTY 3 ☐ EMS CASUALTY | SEVERITY 1 ☐ INJURY 2 ☐ DEATH | AFFILIATION 2 ☐ OTHER EMERGENCY PERSONNEL 3 ☐ CIVILIAN
GD FAMILIARITY WITH STRUCTURE | LOCATION AT IGNITION | CONDITION BEFORE INJURY
GE CONDITION PREVENTING ESCAPE | ACTIVITY AT TIME OF INJURY | CAUSE OF INJURY
GF NATURE OF INJURY | PART OF BODY INJURED | DISPOSITION
☐ SEE REMARKS ON BACK ☐ SEE ADDITIONAL REPORT

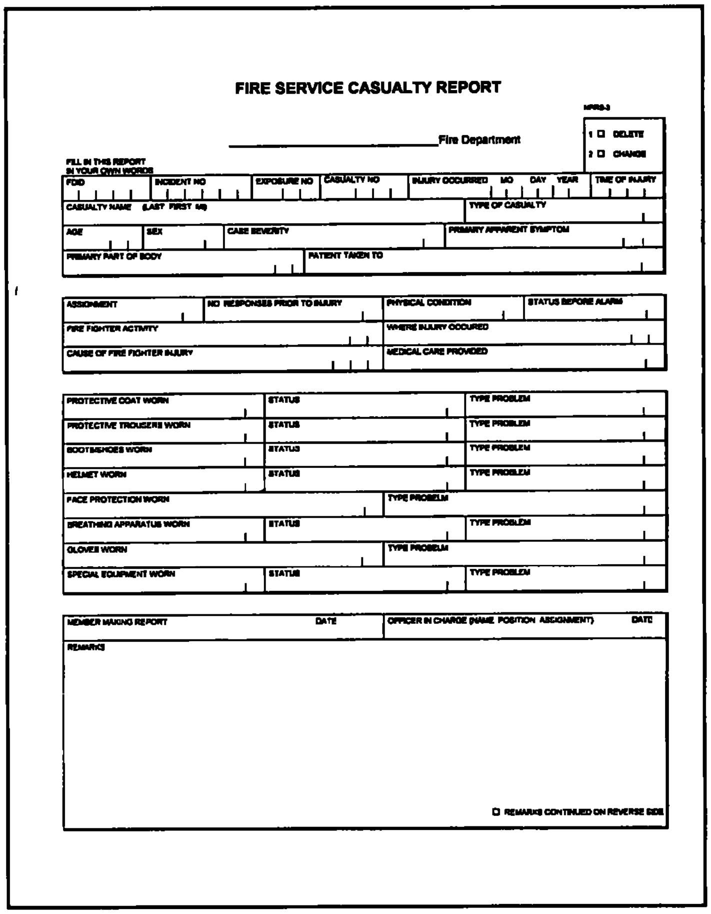
CASUALTY NUMBER — 1 ☐ DELETE 2 ☐ CHANGE

GA CASUALTY LAST NAME | FIRST NAME | MI | DOB (MO/YR) | AGE | TIME OF INJURY
GB HOME ADDRESS | TELEPHONE
GC SEX 1 ☐ MALE 2 ☐ FEMALE | CASUALTY TYPE 1 ☐ FIRE CASUALTY 2 ☐ ACTION CASUALTY 3 ☐ EMS CASUALTY | SEVERITY 1 ☐ INJURY 2 ☐ DEATH | AFFILIATION 2 ☐ OTHER EMERGENCY PERSONNEL 3 ☐ CIVILIAN
GD FAMILIARITY WITH STRUCTURE | LOCATION AT IGNITION | CONDITION BEFORE INJURY
GE CONDITION PREVENTING ESCAPE | ACTIVITY AT TIME OF INJURY | CAUSE OF INJURY
GF NATURE OF INJURY | PART OF BODY INJURED | DISPOSITION
☐ SEE REMARKS ON BACK ☐ SEE ADDITIONAL REPORT
OFFICER IN CHARGE (NAME POSITION ASSIGNMENT) DATE
MEMBER MAKING REPORT (IF DIFFERENT FROM ABOVE) DATE

The state fire marshall's office in each NFIRS state has responsibility for collecting data from its tire departments. They usually collect data in two ways. One way is that fire departments without any data processing capability send their written reports to the fire marshall's office. The office then takes responsibility for keying reports into a computer system. Local departments with data processing capabilities may send their data on microcomputer diskettes or magnetic tapes. In either case, the state fire marshall's office merges all reports into a database.

The state offices have another important responsibility. They create tapes of all fire records (incidents and casualties) and send the tapes on a quarterly basis to the Federal Emergency Management Agency in Washington, D.C. From these tapes, a national database on fires is created each year. The national database for 1990, for example, contains over 941,000 fire incident records and over 18,000 civilian casualty records.

Exhibit 1-4 Fire Service Casualty Report



FIRE SERVICE CASUALTY REPORT

NFIRS-3

__________________________Fire Department

1 ☐ DELETE
2 ☐ CHANGE

FILL IN THIS REPORT IN YOUR OWN WORDS

FDID | INCIDENT NO | EXPOSURE NO | CASUALTY NO | INJURY OCCURRED MO DAY YEAR | TIME OF INJURY
CASUALTY NAME (LAST FIRST MI) | TYPE OF CASUALTY
AGE | SEX | CASE SEVERITY | PRIMARY APPARENT SYMPTOM
PRIMARY PART OF BODY | PATIENT TAKEN TO

ASSIGNMENT | NO RESPONSES PRIOR TO INJURY | PHYSICAL CONDITION | STATUS BEFORE ALARM
FIRE FIGHTER ACTIVITY | WHERE INJURY OCCURED
CAUSE OF FIRE FIGHTER INJURY | MEDICAL CARE PROVIDED

PROTECTIVE COAT WORN | STATUS | TYPE PROBLEM
PROTECTIVE TROUSERS WORN | STATUS | TYPE PROBLEM
BOOTS/SHOES WORN | STATUS | TYPE PROBLEM
HELMET WORN | STATUS | TYPE PROBLEM
FACE PROTECTION WORN | TYPE PROBLEM
BREATHING APPARATUS WORN | STATUS | TYPE PROBLEM
GLOVES WORN | TYPE PROBLEM
SPECIAL EQUIPMENT WORN | STATUS | TYPE PROBLEM

MEMBER MAKING REPORT | DATE | OFFICER IN CHARGE (NAME POSITION ASSIGNMENT) | DATE
REMARKS

☐ REMARKS CONTINUED ON REVERSE SIDE

From a national perspective, this database is vitally important. It allows us to create a picture of different types of fires across the country. The national database has “strength in numbers.” Your department, for example, may have only a few fires, if any, where curtains were ignited (Code 36 for form of material ignited). Nationwide, however, there were 2,395 fires of this type during 1990. ‘These fires can be analyzed to draw conclusions about their causes, extent of damages, dollar Iosses, and other factors.

Data Entry and Data Quality

An assumption throughout the handbook is t-hat data on tire incidents and casualties have been entered into a computer and are available for analysis. While manual analysis is certainly possible, it is usually avoided because the tedious calculations quickly overwhelm our ability to perform analysis in any meaningful manner. The advantage of a computer is that it processes data quickly and accurately.

An immediate problem is how to get data into a computer in the first place. If you are in a large department, this may not be a problem because you probably have a data processing section to enter data. The section may be within the fire department or somewhere else in your government structure. In either case, this section enters NFIRS reports into a computer. Smaller departments usually depend on microcomputers for data collection and analysis. In fact, there are several microcomputer programs specifically for entering incident and casualty data. These programs are not expensive, ranging from $200 to $600.

One word of caution, however, is that any program you purchase should contain a good error checking routine. Data quality is always a problem, and the old adage "Garbage In, Garbage Out" certainly applies to fire department reports. The entry program should, for example, check each item to make sure a valid code has been entered. Whenever the program encounters an error, you should be given an opportunity to correct the error before the data become part of a database. For example, alarm times obviously cannot have hours greater than 23 and minutes greater than 59. An entry program should check hours and minutes for valid numbers, and allow you to make corrections immediately. Similarly, extent of flame damage cannot be coded with an 8 or 9 because these numbers are undefined for this variable (see Exhibit 1-2). Of course, alphabetic characters are also invalid codes.

There is a difference, however, between an "invalid" code and a "wrong" code. By an invalid code, we mean that the code is not on the list of possible codes. A different situation occurs when you enter a "2" instead of a "3" for Extent of Flame Damage. Then you have entered the wrong code.

Wrong codes also occur when blanks or zeroes appear. NFIRS allows blanks because the data are not immediately available and will be determined later. Items particularly susceptible to blanks are the following:

- Mobile property type
- Number of alarms
- Number of fire service personnel
- Number of engines
- Number of aerials
- Complex
- Level of origin
- Number of stories
- Detector performance
- Sprinkler performance

A similar situation exists with coding zeroes. A zero usually indicates that something either could not be determined or was not reported. Items in which zeroes frequently occur are ignition factor, form of heat of ignition, type of material ignited, and form of material ignited.

The problem with entering blanks and zeroes initially is that a fire department may never have an opportunity to conduct a follow-up for the correct codes. The blanks and zeroes become a permanent part of the computer record.

While we are on the subject of quality, it is worth exploring the consequences of errors. Obviously, invalid and wrong codes can result in wrong conclusions. If only a few errors appear in the data, the impact on conclusions may not be substantial. On the other hand, a review of national data shows the number of fire personnel (Line 11) is blank in over 25 percent of the fire records, and this amount of missing data has a substantial impact on analysis of responses to fires.

As another illustration, consider what happens when errors appear in alarm time and arrival time. One problem, which surfaced in virtually all fire department data we analyzed, is an occasional reversal of two times. Suppose the alarm time is 1023 and the arrival time is 1027, but the times are reversed on the incident report so that the computer record shows 1027 as the alarm time and 1023 as the arrival time. If we reviewed this report carefully, we would undoubtedly catch and correct this problem before entry into the computer. However, most entry programs will accept these two times without realizing the error. The adverse consequences can be seen when the computer calculates response time, which is defined as the elapsed time between alarm time and arrival time. The response time should be 4 minutes (from 1023 to 1027), hut the reversal creates a situation where the computer calculates an elapsed time of 1356 minutes!! It acts as if the alarm time is 1027 of one day and the arrival time is 1023 of the next day.

It takes only a few errors of this type to cause the calculation of average response time to be completely erroneous. If one-half of one percent of your records have reversed times, the overall average response time may be increased from a correct value of 3 minutes to an erroneous average of 10 minutes.

The point is that fire departments need to establish *data quality procedures* if they intend to take full advantage of their data. Data quality procedures mean that blanks and zeroes should be checked to see if better entries can be made. It also means that response times more than 15 minutes, for example. should be checked. On-scene times (between arrival time and time in service) should be checked as well.

Quality control also means that reports arc checked for logical inconsistencies. A simple example is that Lines J through M should always contain data if type of situation found is between 10 and 19 , indicating a fire occurred. Similarly, Lines N through R should always contain data if type of situation found is an 11, indicating a structure fire.

In summary, data entry programs should include code checking routines to identify errors in individual items in the report and errors reflected through inconsistencies between items. Because entry programs cannot be expected to find all errors, fire departments also need data quality procedures to ensure that correct data are entered into their systems.

Statistical Packages for Computers

In this handbook, we present many different types of analysis. Chapter 3, for example, discusses several types of charts, including bar charts, column charts, histograms, line charts, and dot charts. Other chapters show how to calculate statistics, such as means and variances, and how to do more advanced calculations such as chi-square tests, correlations, and regression coefficients.

In the future, you will want to depend on computers with analysis programs to perform these calculations instead of doing them manually. For a good understanding of analysis, you need to know what is involved, but you should not continue in a manual mode. There are several good statistical packages available for both microcomputers and mainframe computers:

Exhibit 1-5

BMDP Statistical Software, Inc.
1440 Sepulveda Boulevard
Los Angeles, California 90025
213-479-7799

NCSS
329 North 1000 East
Kaysville, Utah 84037
801-546-0445

SAS Institute, Inc.
Software Sales Department
SAS Campus Drive
Gary, North Carolina 27513
919-677-8200

SPSS, Inc.
444 N. Michigan Avenue
Chicago, Illinois 60611
312-329-3500

Statistical Sciences, Inc.
1700 Westlake Avenue, N.
Suite 500
Seattle, Washington 98109
206-283-8802

STSC, Inc.
2115 East Jefferson St.
Rockville, Maryland 20852
301-984-5123

SYSTAT, Inc.
1800 Sherman Avenue
Evanston, Illinois 60201-3793
708-864-5670

If you intend to apply the techniques in this handbook, you should acquire and learn how to use one of these packages.

Books on Data Analysis

You can also expand your knowledge of data analysis with several good textbooks. The following are basic books intended for general audiences:

Misused Statistics: Straight Talk for Twisted Numbers by A. J. Jaffe and Herbert F. Spirer (Marcel Dekker, Inc., New York. N.Y., 1987).

Say It With Figures by Hans Zeisel (Harper & Row Publishers, Inc., New York, N.Y., 1985)

Say It With Charts by Gene Zelazny (Dow Jones-Irwin, Inc., Homewood, Illinois, 1985)

There are numerous statistics books which provide more details on the subjects of this handbook. These books assume more background in algebra and statistics than we have assumed. A sampling follows:

Statistics: The Exploration and Analysis of Data by Jay Devore and Roxy Peck (West Publishing Company, New York. N.Y., 1986)

Statistics by David Freedman, Robert Pisani, and Roger Purves (W.W. Norton & Company, New York, N.Y., 1978)

Beginning Statistics with Data Analysis by Frederick Mosteller, Stephen E. Fienberg, and Robert E.K. Rourke (Addison-Wesley Publishing Company, Reading, Massachusetts, 1983)

Statistics: Concepts and Applications by William C. Schefler (The Benjamin/Cummings Publishing Company, Inc., Menlo Park, California, 1988)

Statistics and Data Analysis: An Introduction by Andrew F. Siegel (John Wiley & Sons, Inc., New York, N.Y., 1988)

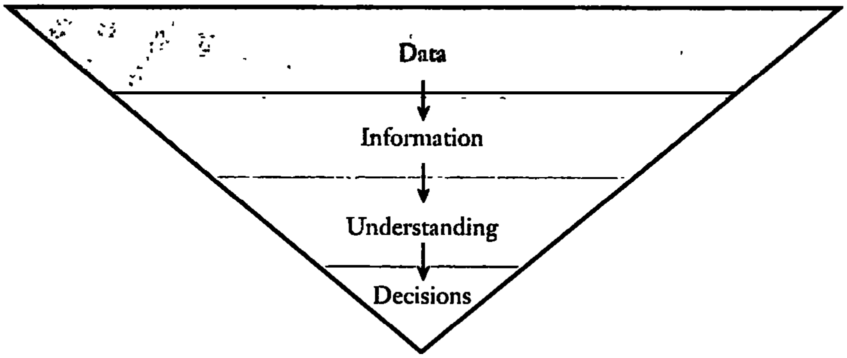
One note of caution. All these books arc general rather than specific to fire departments. However, concepts are clearly explained and can be applied to analysis of fire data. They also give more details about analysis techniques presented in this handbook.

How to Use This Handbook

Data analysis is not an easy process. It requires careful data collection, attention to details, access to statistical programs, and skills on understsnding results. These are not impossible tasks, but requirc time and patience on your part for success. Equally important, you need experience. In the long run, you can only develop capabilities in analysis by applying techniques from this handbook on actual data sets.

As a final note, one way of thinking about analysis is to consider a four-stage process as illustrated in Exhibit 1-6.

Exhibit 1-6 Analysis Steps



Our ultimate objective is to make better and more informed decisions in fire departments. Data have no utility in a vacuum, and fire reports stay as data if we do nothing. *Analysis turns duta into information. We move,* for example, from knowing individual alarm and arrival times to knowing average travel times. Our review of travel times increases our *knowledge* about what is going on with fire incidents which results, in turn, in more informed *decisions* within fire departments.

The remainder of this handbook is organized as follows. We devote Chapters 2 and 3 to descriptions of different types of charts and graphs. Chapter 2 describes histograms, which are probably the easiest and simplest charts to understand. Chapter 3 expands to other types of charts, including bar charts, column charts, pie charts, and dot charts. In Chapter 4, we introduce several basic statistics, including means, medians, modes, and variances. Chapters 5 and 6 discuss analysis of tables, which is particularly important since fire data often comes to us as summaries in the form of tables.

Correlation and regression are the subjects of Chapters 7 and 8. In both chapters, our aim is to present how to perform the calculations associated with these subjects and how to interpret results. Finally, Chapter 9 discusses a modeling technique called queuing theory, which is beneficial for determining the number of emergency medical service units. The numher of units depend on the anticipated workload and on predetermined objectives set by a fire department.

In developing these chapters, we recognized that readers will have varying backgrounds and capabilities. The subject material becomes more difficult as you progress through the handbook. ‘The first five chapters are easy enough to be read by anyone. More technical subjects, such as regression, are more difficult and require knowledge of basic algebra to understand completely. Even in these chapters, however, we have emphasized understanding results rather than concentrating on theory.

More difficult sections in the handbook are identified by asterisks (*) beside the section names. You can skip these sections and still obtain a good understanding of the subject. They contain algebraic equations for calculations associated with a topic. For persons with mathematical backgrounds, these sections will enhance their understanding of the topic.

Finally, problems appear at the end of each chapter. These problems have two purposes. One is to allow you an opportunity to see if you really understand material from the chapter, and the second purpose is to extend your knowledge beyond the basic content of the chapter. As with examples in the chapter, problems include actual data from fire departments.

Chapter 2
HISTOGRAMS

Data as a Descriptive Tool

“A picture is worth a thousand words” is an old saying which applies to numbers as well as words. The task of reaching conclusions from numbers is formidable, particularly when we are looking for trends and patterns in the data. It is for this reason that we turn our attention to histograms and other charts in this chapter and Chapter 3. These tools will assist us in understanding fire data since the human mind appears to comprehend pictures quicker than words and numbers.

The techniques found in these two chapters include:

Histograms
Column charts
Pie charts
Pictograms
Bar charts
Line charts
Dot charts

This chapter describes histograms while Chapter 3 is devoted to the other techniques. With these graphical aids, we can answer several basic questions. When are fires most likely to occur? What are the primary causes of residential fires? vehicle fires? How many civilian injuries occurred last year by month? What are the ages of civilian casualties? What percent of fire incidents have travel times less than 4 minutes? How many structure fires resulted in dollar losses greater than $50,000 last year?

A histogram is a column graph where the height of the columns indicate the relative numbers or frequencies or values of a variable. The following examples show how to organize and display fire data into histograms.

Example 1. One of the most fundamental ways to describe the fire problem is to show how fires are distributed by month, day of week, and hour of day. For example, Exhibit 2-1 shows a **frequency list** of fires by hour of day for Boston, Massachusetts for 1988. A list or array of numbers such as this exhibit is almost always the starting point for a descriptive analysis, but the numbers by themselves are not very useful. It is difficult to get a “feel” for what is happening by scanning a list of numbers.

To grasp what the numbers say in Exhibit 2-1, we can develop a frequency histogram, as shown in Exhibit 2-2. Similarly, Exhibits 2-3 and 2-4 show histograms by day of week and month of the year. Study these exhibits for a few minutes and draw your own conclusions about what they say. Don’t dwell on individual numbers, but instead look for patterns. Ask yourself three questions:

- Where are the low points and high points in the histogram?

- What groups of times (hours, days, or months) have similar frequencies?
- Is there anything in the histogram that runs counter to your experience?

Answers to these questions provide the first insights into your fire data and into conclusions from the data.

Exhibit 2-1 Fires by Hour of Day-Boston-1988

Time Period	Number	Time Period	Number
Midnight - 1 a.m.	478	Noon - 1p.m.	307
1 a.m. - 2 a.m.	420	1 p.m. - 2 p.m.	316
2 a.m. - 3 a.m.	360	2 p.m. - 3 p.m.	363
3 a.m. - 4 a.m.	273	3 p.m. - 4 p.m.	381
4 a.m. - 5 a.m.	192	4 p.m. - 5 p.m.	417
5 a.m. - 6 a.m.	127	5 p.m. - 6 p.m.	433
6 a.m. - 7 a.m.	122	6 p.m. - 7 p.m.	492
7 a.m. - 8 a.m.	139	7 p.m. - 8 p.m.	514
8 a.m. - 9 a.m.	156	8 p.m. - 9p.m.	540
9 a.m. - 10 a.m.	168	9 p.m. - 10 p.m.	622
10 a.m. - 11 a.m.	206	10 p.m. - 11 p.m.	510
11 a.m. - Noon	242	11 p.m. - Midnight	547

It is difficult to get a “feel” for what is happening by scanning a list of numbers.

While these histograms suggest several conclusions, the key ones are:

1. The peak time period for fires in Boston is from 8 pm. to midnight with the hour from 9 p.m. to 10 p.m. having more fires than any other hour.
2. The lowest time period for fires is from 5a.m. to 9 a.m.
3. Weekends are the busiest times for fires while Thursdays are the least busiest days.
4. June and July have more fires than any of the other months while Januay and February have the fewest.

A **histogram** is a column graph where the height of the bars indicate the relative numbers or frequencies for values of a variable. The values may be numeric, such as travel times, or non-numeric, such as days of the week.

With these histograms we begin to see a picture of the fire problem in Boston. Histograms allow for an easy descriptive and analytical procedure without having to think too much about the numbers themselves. Graphical displays should always strive to convey an immediate message describing a particular aspect of the data.

Exhibit 2-2 Fires by Hour of Day–Boston–1988

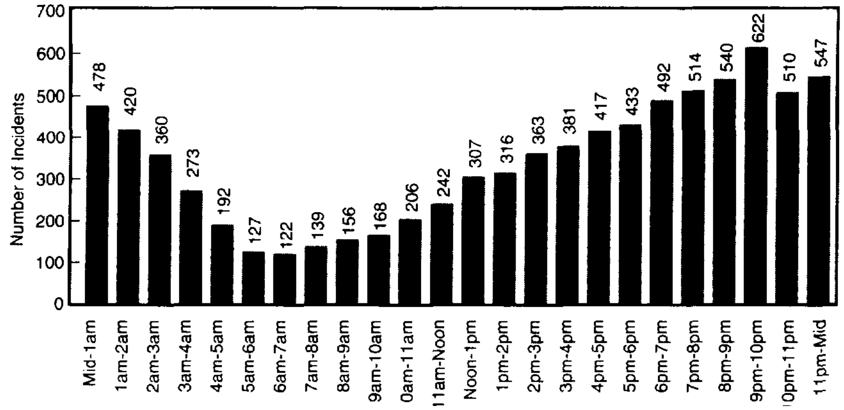


Exhibit 2-3 Fires by Day of Week–Boston–1988

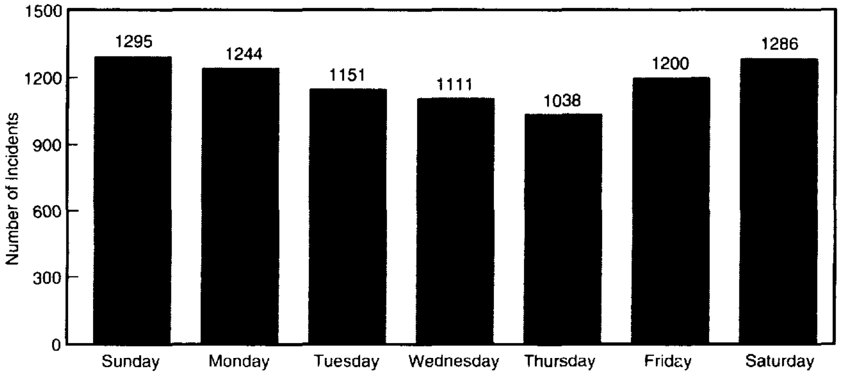
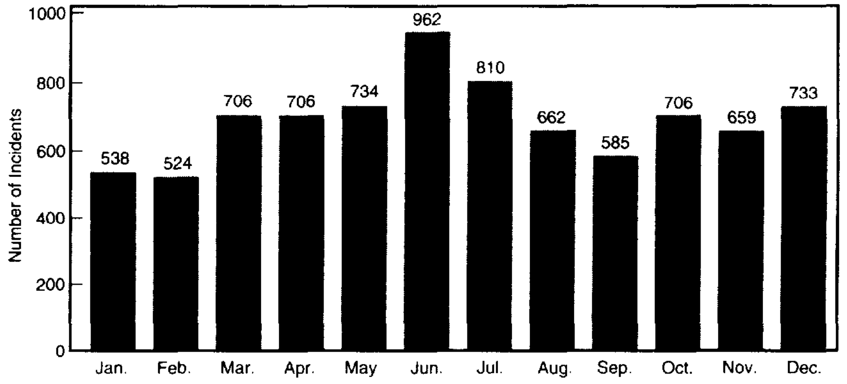


Exhibit 2-4 Fires by Month–Boston–1988

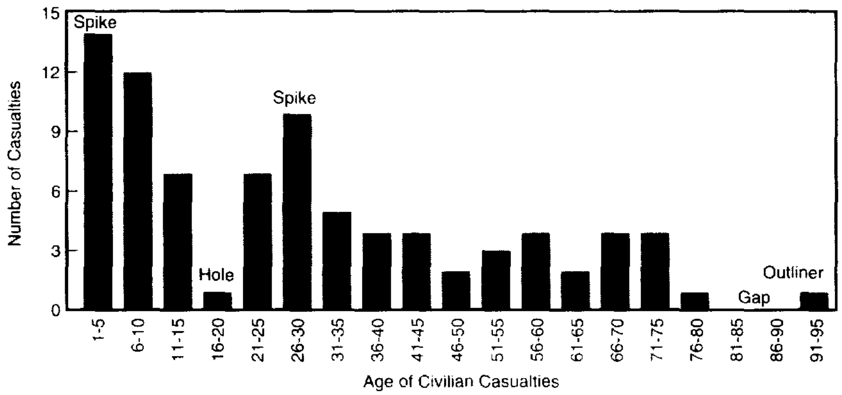


Example 2. Ages of Civilian Casulties. Suppose a fire chief is interested in developing a fire prevention program aimed at reducing civilian injuries and deaths. Descriptive data on civilian casualties is available from the fire reports and there are a number of different descriptions that could be developed from the data. One of the most basic is descriptive data on the ages of civilian casualties.

Exhibit 2-5 shows the ages of civilians injured or killed in fires in Jersey City, New Jersey for 1988. Note that this distribution is considerably different from the previous histograms primarily because it does not have the same “smoothness.” However, the five-year age groups show some interesting patterns. For example, the age group under five years of age accounts for the most civilian casualties, followed in frequency by the 6-to-10 year age group. Also of interest is how the frequency takes a rather sudden drop for the 16-to-20 year age group. A spike in the data occurs with the 26-to-30 year age group. The exhibit also reveals a gap in the data for ages 81-to-90 and an outlier in the last group from 91-to-95 years of age.

> **Spikes** are high or low points that stand out in a histogram. Outliers are extreme values isolated from the body of the data. **Gaps** are spaces in a histogram reflecting low frequencies of data.

Exhibit 2-5 Ages of Civilian Casualties-Jersey City-1988



Note: Age was not provided for 14 casualties. 40% of the civilian casualties were younger than 16 years old. 20% were more than 50 years old.

In histograms and other charts, it is sometimes useful to include comments and conclusions with the chart. In Exhibit 2-5, we provided a note that 14 casualty records did not include age information and were therefore not included in the histogram. Other notes provide summary information on the data such as the percent of casualties under 16 years of age and the percent more than 50 years old. Anyone studying the histogram could reach the same conclusion, but the summary saves time and effort.

Exhibit 2-6 Travel Times–Seattle–1988

Travel Time	Frequency
Less than 1 minute	106
1 to 2 minutes	85
2 to 3 minutes	481
3 to 4 minutes	1,019
4 to 5 minutes	814
5 to 6 minutes	415
6 to 7 minutes	185
7 to 8 minutes	93
8 to 9 minutes	31
9 to 10 minutes	24
10 to 11 minutes	9
11 to 12 minutes	6
12 to 13 minutes	1
13 to 14 minutes	3
14 to 15 minutes	3
15 to 16 minutes	0
16 to 17 minutes	0
17 to 18 minutes	1
18 to 19 minutes	0
19 to 20 minutes	3
Total Fire Calls	3,279

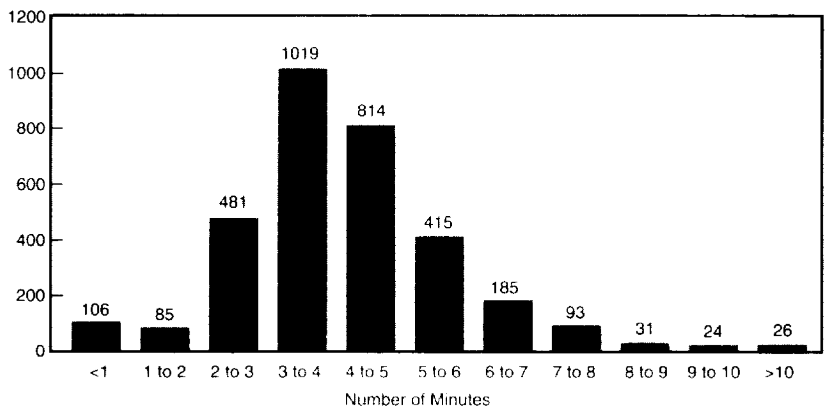
The data for this example included five other calls with travel times of 26 minutes, 47 minutes, 64 minutes, 683 minutes, and 794 minutes respectively. For purposes of presentation, we have assumed that these records are in error. Either the alarm time or the arrival time has probably been incorrectly recorded. This is obviously the case for the last two average times and the first three times are suspect because they greatly exceed the other times in the distribution. In practice, these times should be reviewed and corrected if necessary.

Example 3. Travel Times to Fires. Travel times to tires are one of the most important data sets to study in fire departments. Many fire departments have objectives for average travel times to fires and try to allocate personnel to achieve these travel times. Exhibit 2-6 shows a frequency distribution for travel times to fires in Seattle, Washington in 1988.

Notice in this example that the times are clustered at the low end of the distribution as we would expect since travel times to fires are generally low for most fire departments.

Exhibit 2-7 provides a frequency histogram for this distribution. In this exhibit, we have combined the last few points into a category of 10 minutes or more. A histogram with the shape in this exhibit is sometimes called **skewed to the right** or **skewed toward high values.** What we mean by these terms is that the distribution is not symmetric but instead has a single peak on the left side of the distribution with a long tail toward the right. In fire departments, data on on-scene time (from time of arrival to time back in service) and data on dollar losses at fires also reflect skewness to the right.

Exhibit 2-7 Travel Times–Seattle–1988



Developing a Histogram

Making a histogram is relatively straightforward:

1. Choose the number of groups for classifying the data. You should usually, have between 5 and 10 groups, hut there are exceptions such as histograms by hour of day. Sometimes the groups are natural, as in our exhibits by day of week and month. With other data, you will have to develop appropriate intervals for the data, as we did with the ages of civilian casualties in Exhibit 2-5.
2. Determine the number of events (fires, casualties, etc.) for each of your groups.
3. For data such as ages and travel times, you usually need to define intervals. For these intervals, you should choose convenient whole numbers. That is, avoid fractions in the groups and always make the intervals the same width. In Exhibit 2-5 we used intervals of five years for grouping the data. Data such as day of week do not require this step since their intervals are naturally defined.
4. Determine the number of observations in each group. Statistical packages are particularly, useful for this step since they always include routines for tabulating data.
5. Choose appropriate scales for each axsis to accommodate the data.
6. Display the frequencies with vertical bars.

Do not expect to get a histogram, or any other type of chart, exactly right on the first try. You may have to try several times before you are satistied with the look of the histogram.

The hisgrams presented in the previous section offer good example of different characteristics for describing the data. Mosteller, et. al. (1983) offer the following definitions of features you should try to find in histograms:

1. **Peaks and valleys.** The peaks and valleys in a histogram indicate the values that appear most frequently (peaks) or least frequently (valleys). Exhibit 2-2 shows clear peaks and valleys for incidents by hour of day.
2. **Spikes and holes.** These are high and low points that stand out in the histogram. In Exhibit 2-5, for example, there is a spike for the 26-to-30 year age group, and a hole for the 16-to-20 year age group.
3. **Outliers.** Extreme values arc sometimes called outliers and are points that are isolated from the body of the data. In Exhibit 2-5, there is one outlier in the 91-to-95 year age group.
4. **Gaps.** Spaces may reflect important aspects of a histogram. In Exhibit 2-5, there is a gap in the 81-to-90 year age group.
5. **Symmetry.** Sometimes a histogram will he balanced along a central value. When this happens, the histogram is easier to interpret. The central value is both the average for the distribution and the median (half the data points will be below this value and half will he above).

Cumulative Frequencies

Two other types of distributions which will be important in later chapters are the **cumulative frequency** and the **cumulative percentage frequency.** A cumulative frequency is the number of data points that are less than or equal to a given value. A cumulative percentage frequency converts the cumulative frequencies into percentages.

Example 4. With the data in Exhibit 2-6, we can calculate the cumulative frequency and cumulative percentages for the travel time data from Seattle, Washington.

Exhibit 2-8 Cumulative Travel Times–Seattle–1988

Travel Time	Frequency	Cumulative Frequency	Cumulative Percent
Less than 1 minute	106	106	3.2
1 to 2 minutes	85	191	5.8
2 to 3 minutes	481	672	20.5
3 to 4 minutes	1,019	1,691	51.6
4 to 5 minutes	814	2,505	76.4
5 to 6 minutes	415	2,920	89.1
6 to 7 minutes	185	3,105	94.7
7 to 8 minutes	93	9,198	97.5
8 to 9 minutes	31	3,229	98.5
9 to 10 minutes	24	3,253	99.2
10 or more minutes	26	3,279	100.0
Total		3,279	100.0

A cumulative frequency is the number of data points that are less than or equal to a given value. A cumulative percentage frequency converts the frequencies into percentages.

The first entry under the “Cumulative Frequency” column is 106, which is the same as in the “Frequency” column. The second entry shows 191, which is 106 + 85, the sum of the first two entries in the “Frequency” column. By adding these two number, we can say that I91 incidents have travel times less than 2 minutes. The next entry is 672 (106+85+481) and means that 672 incidents have travel times less than 3 minutes. The cumulative frequencies continue in this manner with the last entry in the column always equal to the total number of incidents in our analysis.

The last column, labeled “Cumulative Percent” merely converts the cumulative frequencies into percentages. This step is accomplished by dividing each cumulative frequency), by 3,279, which is the total number of incidents. The column shows that 3.2 percent of the incidents have travel times less than 1 minute, 5.8 percent less than 2 minutes, 20.5 percent less than 3 minutes, etc.

A **cumulative frequency** is the frequency of data less than or equal to a group. A **cumulative percent** is the cumulative frequency divided by the total number of events.

In general, cumulative percentages describe data in “more than” and “less than” terms. We can conclude, for example, that about half the calls have travel times of less than 4 minutes and about 95 percent have travel times less than 7 minutes. Travel times exceed 9 minutes in only about one percent of the calls.

Summary

A list of numbers is frequently the starting point for analysis. If the question of interest is for specific information, then the list of numbers serves the purpose. For example, Exhibit 2-1 is useful if we are asked about exactly how many fires occurred between 2 a.rn and 3 a.m., or if we want to know the exact difference between the busiest hour and the least busiest hour. On the other hand, Exhibit 2-1 is not very useful for determining, for example, the six busiest hours of the day.

Histograms provide a much better method for getting the feel of a list of numbers and answering several questions about relationships. The patterns in a histogram are especially important. For example, high frequencies and low frequencies are usually important to note. Trends indicated by) spikes, outliers, and paps in a histogram are also important.

Chapter 2
PROBLEMS

1. With the data in Exhibit 2-1, determine the number of fires by four-hour periods (Midnight to 4 a.m., 4 a.m. to 8 a.m., etc.). Develop a histogram for these four hour periods. What are the advantages and disadvantages of this histogram compared to Exhibit 2-2?
2. What do Exhibits 2-2, 2-3, and 2-4 tell us about when we should schedule firefighters if we want to match fire workload with personnel?
3. The following figures and percentages are for the Boston fires occurring in 1990. Compare these distributions to Exhibits 2-2, 2-3, and 2-4. Note first that there is a substantial reduction in the number of fires from 8,325 fires in 1988 to 6,479 fires in 1990. The comparisons you make should determine whether the distribution of fires has changed from 1988 to 1990. That is, are the busy hours during 1990 the same as the busy hours for 1988?

Boston, 1990: Hour of Day

Time Period	Number	Time Period	Number
Midnight - 1 a.m.	386	Noon - 1 p.m.	213
1 a.m. - 2 a.m.	287	1 p.m. - 2 p.m.	265
2 a.m. - 3 a.m.	210	2 p.m. - 3 p.m.	293
3 a.m. - 4 a.m.	194	3 p.m. - 4 p.m.	295
4 a.m. - 5 a.m.	146	4 p.m. - 5 p.m.	354
5 a.m. - 6 a.m.	95	5 p.m. - 6 p.m.	380
6 a.m. - 7 a.m.	78	6 p.m. - 7 p.m.	384
7 a.m. - 8 a.m.	126	7 p.m. - 8 p.m.	432
8 a.m. - 9 a.m.	141	8 p.m. - 9 p.m.	498
9 a.m. -10 a.m.	138	9 p.m. - 10 p.m.	492
10 a.m. -11 a.m.	156	10 p.m. - 11 p.m.	394
11 a.m. - Noon	183	11 p.m. - Midnight	339

Boston, 1990: Day of Week

Day	Number	Percent
Sunday	965	14.9
Monday	885	13.7
Tuesday	960	14.8
Wednesday	912	14.1
Thursday	906	14.0
Friday	944	14.6
Saturday	907	14.0
Total	6,479	100.0

Boston, 1990: Month

Month	Number	Month	Number
January	508	July	798
February	342	August	493
March	529	September	509
April	548	October	436
May	529	November	580
June	702	December	505

4. Answer the following questions based on the data on civilian casualties from Jersey City for 1988:

Jersey City, New Jersey, 1988: Ages of Civilian Casualties

Age Group	Number	Age Group	Number
l-5	14	51-55	3
6-10	12	56-60	4
11-15	7	61-65	2
16-20		66-70	3
21-25	7	71-75	3
26-30	10	76-80	1
31-35	5	81-85	0
36-40	4	86-90	0
41-45	4	91-95	1
46-50	2		

a. Develop a cumulative frequency distribution and cumulative percentage distribution of the data:

b. What percentage of civilians less than 16 years of age were injured or killed in fires?

c. What percentage of civilians were more than 50 years old?

5. The following data are from the national NFIRS system which was discussed in Chapter 1. The data show the distribution of ages, in five year increments, for 2,280 civilian casualties in 1988. Civilian casualties include persons other than firefighters injured or killed during fires. Develop a histogram from these figures and compare the results to the data from Jersey City, New Jersey in Exhibit 2-5.

Age	Number	Percent	Age	Number	Percent
< 5 years	256	11.2	51 - 55 years	69	3.0
6 - 10 years	100	4.4	56 - 60 years	94	4.1
11 - 15 years	90	3.9	61 - 65 years	105	4.6
16 - 20 years	139	6.1	66 - 70 years	76	3.3
21 - 25 years	231	10.1	71 - 75 years	53	2.3
26 - 30 years	275	12.1	76 - 80 years	51	2.2
31 - 35 years	230	10.1	81 - 85 years	47	2.1
36 - 40 years	179	7.9	86 - 90 years	23	1.0
41 - 45 years	139	6.1	91 - 95 years	7	.3
46 - 50 years	114	5.0	96 - 100 years	2	.l

6. For the civilian casualties in the previous problem, the following data show the number of casualties by hour of day. Develop a plot of these casualties and give reasons why this distribution differs from the distribution of fires by hour of day.

Civilian Casualties, 1988: Hour of Day

Time Period	Number	Time Period	Number
Midnight - 1 a.m.	93	Noon - 1 p.m.	66
1 a.m. - 2 a.m.	125	1 p.m. - 2 p.m.	110
2 a.m. - 3 a.m.	118	2 p.m. - 3 p.m.	94
3 a.m. - 4 a.m.	159	3 pm. - -I p.m.	105
4 a.m. - 5 a.m.	101	4 p.m. - 5 p.m.	92
5 a.m. - 6 a.m.	85	5 p.m. - 6 p.m.	112
6 a.m. - 7 a.m.	62	6 p.m. - 7 p.m.	103
7 a.m. - 8 a.m.	69	7 p.m. - 8 p.m.	62
8 a.m. - 9 a.m.	71	8 p.m. - 9 p.m.	88
9 a.m. - 10 a.m.	107	9 p.m. - 10 p.m.	85
10 a.m. - 11 a.m.	108	10 p.m. -11 p.m.	81
11 a.m. - Noon	86	11 p.m. -Midnight	84

NOTE: The time of the fire was not known in 14 cases so that the total for the above figures is 2,266 civilians.

7. How do peaks, holes, spikes and gaps affect cumulative distributions?

This page intentionally left blank.

Chapter 3
CHARTS

Introduction

In this chapter we will extend beyond histograms to other types of charts. The point is that histograms are only one of many different ways of portraying data. As an analyst, you must decide which type of chart best reflects the results you want to present. A histogram may serve as the best vehicle, but other types of charts should be considered such as bar charts, line charts, pie charts, dot charts and pictograms. Each of these will be explained in this chapter.

Two questions to bear in mind throughout this chapter are the following:

- What are the main conclusions from your analysis?
- What is the best way to display the conclusions?

As with the previous chapter, several sets of data will be presented in this chapter. You should study each example carefully and draw your own conclusions about the results. You may, in fact, disagree with what we emphasize or you may identify an aspect of the data we overlooked. In either case, think about how you would present your viewpoints in a graphical format to a given audience. The audience may be an internal group of managers, an outside association or group of citizens, or even your city or county council. Even the audience influences the type of chart selected.

The first step is therefore to determine the key results you see from the data. Once you have reached conclusions, you want to select the best type of chart to convey those conclusions. Often you will want to try different charts to determine the best presentation for your audience and your data.

Each of the following sections describes a different type of chart. At the end of the chapter, we present guidelines on selecting a type of chart suitable for different conclusions.

Bar Charts

A **bar chart** is one of the simplest and most effective ways to display data.

In a bar chart, we simply draw a bar for each category of data allowing for a visual comparison of the results. For example, the figures in Exhibit 3-1 on the following page give the ignition factors (from NFPA 901 codes) for the 7,509 structure fires in Chicago, Illinois for 1990.

Our interest in a list of this type usually centers on how the items compare to each other. What is the leading ignition factor in structure fires? How does misuse compare to mechanical deficiency problems? How big a problem are suspicious fires?

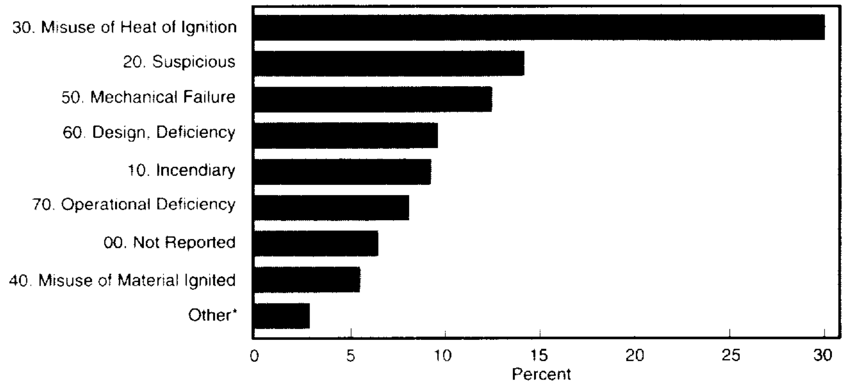
We can determine some results relatively easy from the list of numbers. For example, misuse of heat of ignition is clearly the leading ignition factor followed by suspicious fires and mechanical failures. Operational deficiencies account for less than 10 percent of the total. Design, construction, and installation deficiencies also account for less than 10 percent of the total. However, these results require us to make comparisons mentally with numbers or percentages.

A bar chart overcomes these problems by presenting the data in frequenecy order, as displayed in Exhibit 3-2. The horizontal dimension gives

Exhibit 3-1 Ignition Factors for Structure Fires–Chicago, Illinois–1990

Ignition Factor	Number	Percent
00. Not Reported	496	6.6
10. Incendiary	708	9.4
20. Suspicious	1,077	14.3
30. Misuse of Heat of Ignition	2,273	30.3
40. Misuse of Material Ignited	426	5.7
50. Mechanical Failure	950	12.7
60. Design, Construction, Installation Deficiency	734	9.8
70. Operational Deficiency	620	8.3
80. Natural Conditions	16	.2
90. Other Ignition Factors	209	2.8
Total	7,509	100.0

Exhibit 3-2 Ignition Factors for Structure Fires–Chicago, Illinois–1990



* Other category includes Natural Conditions (80) and Other Ignition Factors (90)

the percent, while the vertical dimension shows the category labels. The bars are presented in numerical order starting with operational deficiencies as the most frequent. Each bar also contains the number of fires for the ignition factor as additional information to a reader.

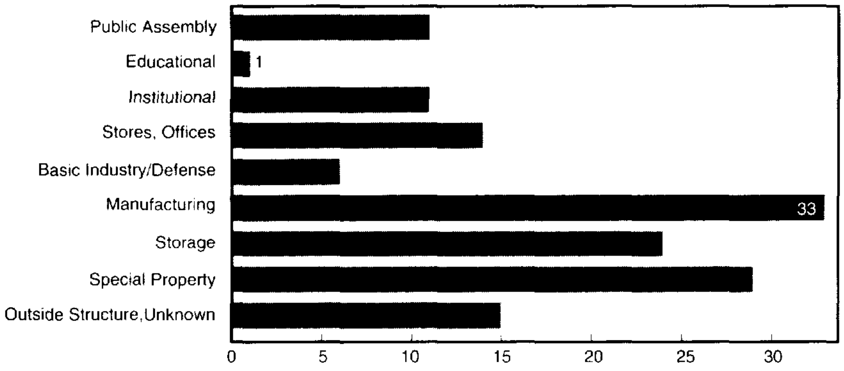
You should also note that the bottom bar is labeled “Other” with 225 structure fires in the bar. This number is actually the combination of the natural conditions (16 fires) and other ignition factors (209 fires). It is not unusual to combine low frequency categories into an “other” category. However, you should accompany the chart with a table, as we have done here, or add a footnote to the chart indicating the combinations.

As a general rule, the horizontal dimension in a bar chart is numeric, such as percentages or other numbers, while the vertical dimension shows the labels for the items in a category. It is not always necessary to include numbers in each bar, but they are sometimes useful to readers unfamiliar with the data. If you omit numbers in the charts, you should provide the total number of incidents either in the title or as a footnote.

We could have arranged the labels in code number order so that they matched the list in Exhibit 3-1. The emphasis in such a chart would be on the individual categories rather than on their ranking. However, in this case, you might want to highlight the bar of the highest ranking item with stripes or a different color than the rest of the graph. Exhibit 3-3 is an example of this type of chart. The exhibit has horizontal bars indicating the number of non-residential deaths by property type during 1990. As emphasized by the solid bar, more deaths occurred in manufacturing facilities than any other type of property.

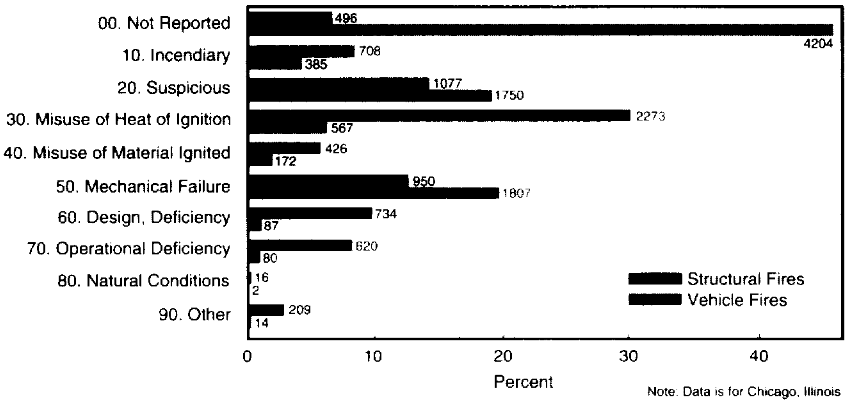
A **grouped bar chart** shows two categories in the same chart. In Exhibit 3-4, for example, we display the ignition factors for structure versus vehicle fires. Since there are two categories, we list the items in code number order.

Exhibit 3-3 Non-Residential Fire Deaths by Property Type-1990



Source: National Fire Information Reporting System, 1990 National Statistics

Exhibit 3-4 Ignition Factors for Structure vs. Vehicle Fires–1990



Note: Data is for Chicago, Illinois

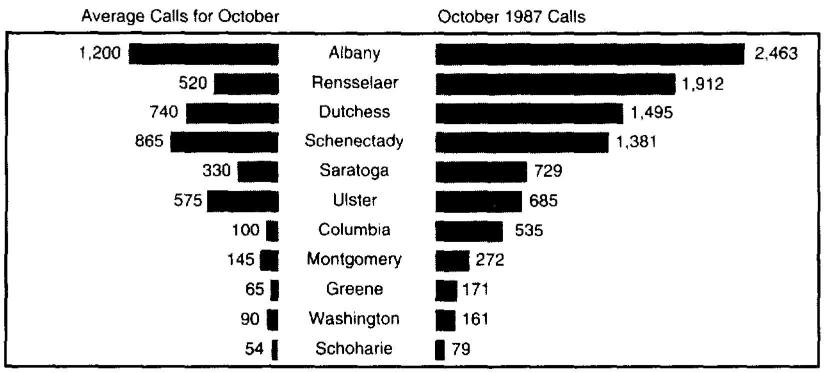
This exhibit shows, for example, that suspicious fires (generally arson fires) are a greater problem with vehicle fires than structure fires. The chart also shows that ignition factors are not reported in more than 4,000 vehicle fires compared to only 496 structure fires. The paired bar chart clearly shows the differences in ignition factors for these two types of fires.

A **paired bar chart** makes item-by-item comparisons. By way of background to an interesting paired bar chart, we provide the following summary of the impact from a snowstorm in New York state in 1987:

> "On Sunday morning, October 4, 1987, the Hudson Valley region of New York State was hit by an unusual, early snowstorm which brought up to two feet of heavy, wet snow. The snow fell in a band from Washington and Saratoga Counties in the north, through the Berkshires and Catskills to the northern part of Westchester County in the south. The sudden impact on fire and other emergency services throughout the region will long be remembered. At least seven deaths were directly attributed to the storm. More than 270 fire departments mobilized over 11,000 firefighters and several hundred pumps, generators, saws and related equipment in what would be, for many, a round-the-clock, week-long outpouring of service to their communities."
>
> (New York State Annual Report, 1987).

Exhibit 3-5 displays the number of calls for the 11 most affected counties. The left side of the exhibit shows the average number of calls for past Octobers while the right side shows the October 1987 calls. With this arrangement, we can immediately see the impact of the snowstorm for each county.

Exhibit 3-5 Calls for October, 1987 vs. Average Calls for Past Octobers



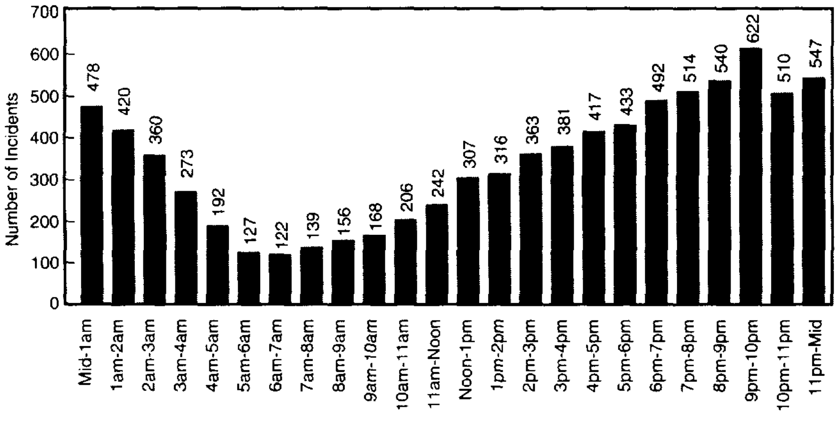
Bar charts emphasize the rankings of items within a group. Item labels provide descriptions for each item in the group. Grouped bar charts show two or more categories in the same chart. Paired bar charts allow for item-by-item comparisons.

Column Charts

We displayed several column charts in Chapter 2. For example, Exhibits 2-2, 2-3, and 2-4 showed Boston fires during 1988 by hour of day, day of week, and month. These are all examples of **time series** presented as **column charts.**

Column charts of this type are particularly useful in demonstrating

Exhibit 3-6 fires by Hour of Day–Boston–1988



change over time. Where is the series increasing, decreasing, or staying about the same? If our analysis shows changes over time, then column charts are particularly beneficial in presenting the changes.

As an example, we repeat the exhibit from Chapter 2 on fires by hour of day in Exhibit 3-6. By moving our eye from left to right we visualize the change in our mind. The horizontal scale scales the hours, but we do not really need that reference to get a feeling for the changes. Calls are low in early morning hours, then increase in the afternoon and evening hours.

> Column charts show frequency distributions that allow for easy identification of trends and other characteristics, particularly with time series data. The horizontal scale defines the natural groupings for the chart and the columns gives the frequencies.

Another good application of column charts is to show comparisons across several sets of data. Exhibit 3-7 lists fire department activities for four sites divided into fires, rescue calls, and other calls. Comparisons across the sites are not easy because the totals differ so much. Site A has 17,576 calls while the other sites have less than 3,000 calls. A simple way to overcome this problem is to develop percentages.

Exhibit 3-7 Comparison of Fire Department Activities–1988

	Site A	Site B	Site C	Site D
Fires	1,390	170	346	368
Rescue Calls	10,242	636	576	1,668
Other	5,944	576	694	879
Total	17,576	1,382	1,616	2,915

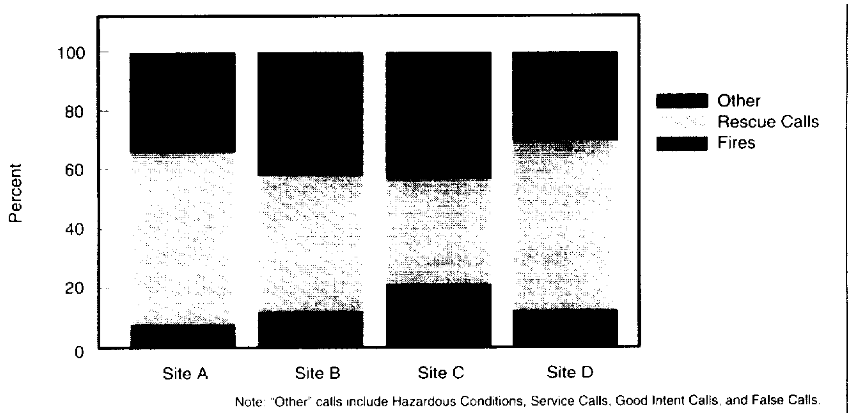
	Site A	Site B	Site C	Site D
Fires	7.9%	12.3%	21.4%	12.6%
Rescue Calls	58.3%	46.0%	35.6%	57.2%
Other	33.8%	41.7%	43.0%	30.2%
Total	100.0%	100.0%	100.0%	100.0%

Note: "Other" calls include Hazardous Conditions, Service calls, Good Intent calls, and False calls

By converting the site figures to percentages, as shown at the bottom of the exhibit, we have a better basis for comparisons. The percentages for each site always add to 100 percent. While there many many conclusions that could he drawn from these percentages, the key conclusions are:

- Fire calls are always the smallest percentage of activity in each site.
- Rescue calls are the predominant type of call in three sites.

Exhibit 3-8 Comparison of Fire Department Activities by Percent–1988



Note: "Other" calls include Hazardous Conditions, Service Calls, Good Intent Calls, and False Calls.

To display this result, we develop **stacked column charts** as shown in Exhibit 3-8 using the percentages for each activity. The columns all have the same height since they add to 100 percent. Different shadings highlight the amount of activity. The results just discussed should he clear from the exhibit.

Line Charts

Effective presentation of time series data may also be developed from **line charts.** Exhibit 3-9 shows a line chart of fires for Detroit, Michigan, for 1988, 1989, and 1990 by month. The line chart immediately highlights problems during the summer months of 1988 when a substantial number of fires occurred. For 1989 and 1990, these summer months are still among the highest for the years, hut do not begin to approach the problems in 1988. Many statisticians believe that a line

Exhibit 3-9 Incidents by Month-Detroit, Michigan–1988–1990

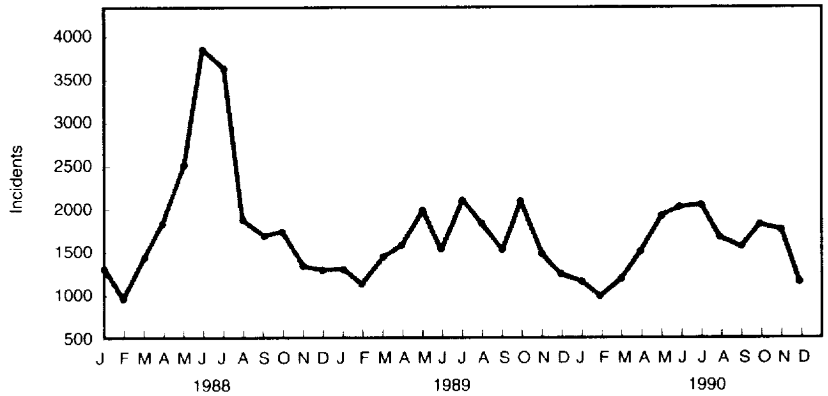


chart is the clearest nay for showing increases, decreases, and fluctuations in a time series.

Line charts give effective presentations of time series data, such as the number of incidents per month for several years. Fluctuations in the data are easily identified by line charts.

Pie Charts

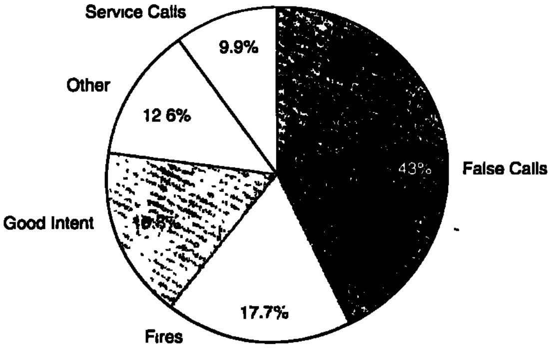
A **pie chart** is an effective way of showing how each component contributes to the whole. In a pie chart, each wedge represents the amount for a given category. The entire pie chart accounts for all the categories.

For example, Exhibit 3-10 shows the activities of the Seattle, Washington Fire Department for 1990 divided into fire calls, false calls, service calls, good intent calls, and other calls. The percentages are inserted in each wedge. Although the percentage numbers are not necessary, they aid in the comparisons of the wedges. The pie chart emphasizes the fact that false calls account for a high percent of incidents in the city. Fire calls and good intent calls account for about the same percent of total incidents.

In developing pie charts, you should follow the following rules:

- Convert your data to percentages.
- Keep the number of wedges to six or less. If you have more than six, try keeping the most important five and group the rest into a sixth category.
- Position the most important wedge starting at the 12 o'clock position.
- Highlight the most important wedge by coloring it the most intense shading.

Exhibit 3-10 Incidents–Seattle, Washington–1990



While pie charts are popular, they are probably the least effective way of displaying your results. For example, it may be hard to compare wedges within a pie to determine their ranks. Similarly, it takes time and effort to compare several pie charts because they are separate figures.

Develop a **pie chart** when the objective is to show how each item contributes to the whole. Pie charts are not effective for comparing several groups of figures.

Dot Charts

Dot charts or **scatter diagrams** emphasize the relationship between two variables. For example, Prince William County, Virginia experienced increases in population and Emergency Medical Services (EMS) calls over the last few years. In 1981, the department responded to 9,538 EMS calls as compared to 12,744 EMS calls in 1991. During these years, the population increased steadily from 152,300 to 223,900. We would expect EMS calls to increase with population, and it is this relationship that we want to depict in a chart.

Exhibit 3-11 is a dot chart for EMS calls versus population for Prince William County for the eleven years from 1981 to 1991. Population is along the horizontal axis while EMS calls are along the vertical axis. The pattern is the important aspect of a dot chart, rather than the individual dots. The horizontal scale of a dot chart should reflect the causation variable while the vertical scale reflects the resulting variable (that is, population causes or creates EMS calls).

A **dot chart** reflects the pattern of one variable with another. The pattern is more important than the individual dots.

Exhibit 3-11 EMS calls & Population-Prince William County–1981–1991

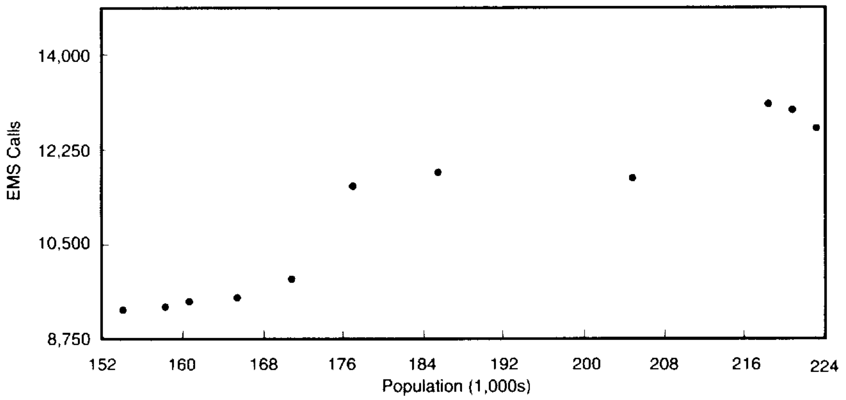
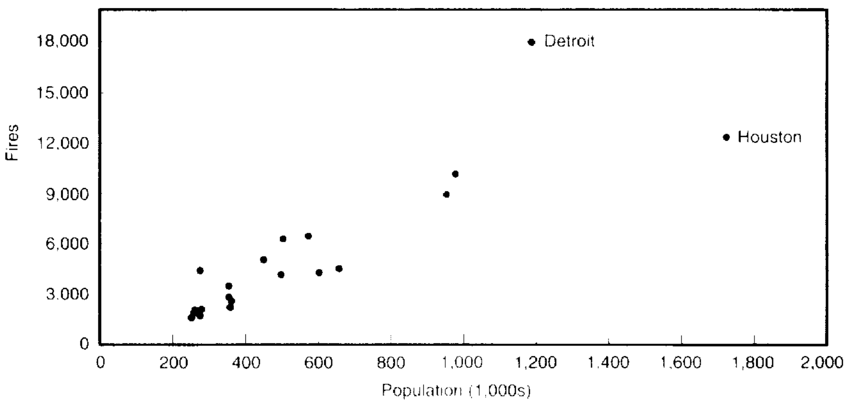


Exhibit 3-12 Total Fires & Population–1988



Another good application of scatter diagrams is to identify outliers in our data. In Chapter 2, we defined outliers as points that are isolated from the body of the data. Exhibit 3-12 shows diagram of population protected and total fires for 20 selected jurisdictions across the country. There is a general pattern showing the obvious fact that fires increase as population increases. However, there are two cities that do not follow this general pattern: Houston, Texas and Detroit. Michigan.

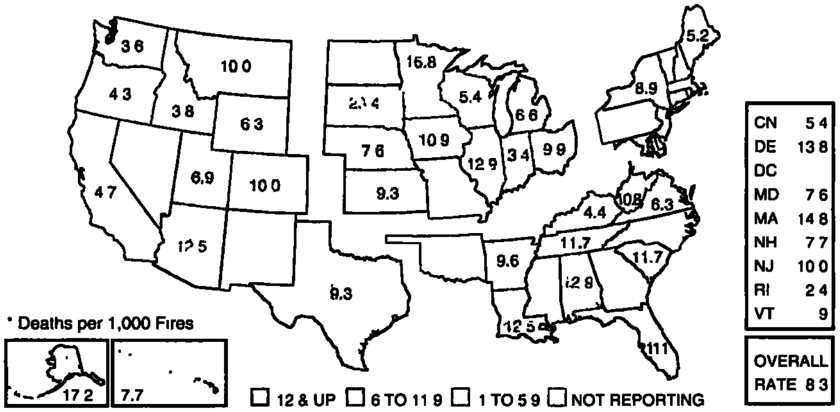
Detroit has more fires than expected based on its population, while Houston has fewer tires than expected based on its population. We make these conclusion because the dots for these two cities differ in location on the exhibit from the general pattern. The dot for Detroit is higher than we would expect from its polulation while Houston is lower than expected based on the general pattern.

In Chapter 5, we consider these two examples in more detail by calculating the correlation and regression line for each example. The correlation and regression line provide greater insight into the strength of association between population. EMS calls, and fires. Because a strong correlation exists, we can apply the regression line for future needs in fire departments.

Pictograms

Our final type of chart takes advantages of pictures to display data. Exhibit 3-13, for example, shows deaths per 1,000 tires for residential structure fires in 1988. Each contributing state has one of three distinquishing patterns reflecting low, average, and high rates. The overall rate of 8.3 injuries per 1,000 residential fires appears at the bottom of the chart. Low rates are in the .1 to 5.9 range (depicted) by light grey), avearge states are in the 6 to 11.9 range (medium grey). and high states have more than 12.0 deaths (dark grey).

Exhibit 3-13 Residential Structure Fire Deaths*–NFIRS Data–1988



The key is that presentation in this manner is more effective than any listing of the death rates. We can easily draw conclusions:

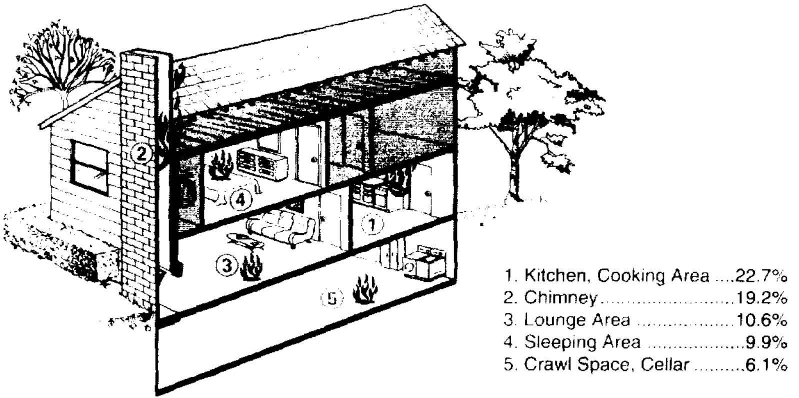
- States with high rates include Arizona, South Dakota, Minnesota, Illinois, Louisiana, Alabama, Delaware, and Massachusetts.
- Low rates predominate in the west (California, Oregon, Washington, and Idaho).
- A group of average states are in the midwest (Montana, Wyoming, Utah, Colorado, Texas, Nebraska, and Kansas).

Other charts for state and local data are easily imagined. At the state level, you may be collecting data from individual counties. A pictogram is 3 good way of depicting the county data by taking a state map showing county boundaries and developing an exhibit similar to Exhibit 3-13. Similarly, if you work for a local jurisdiction, such as a city or county, you may have data for individual fire districts. A jurisdiction map) with fire district boundaries may be an effective way of presenting the data.

As another example, Exhibit 3-14 on the following page shows areas of fire origins for residential structure fires in 1987 for the state of New York (excluding New York City). The percents appear in the tower right corner of the exhibit. The picture gives an effective way of highlighting where fires occurred within the residence.

A **pictogram** takes advantage of the background for the data you want to present. Data by geographical areas, such as counties, census tracts, or fire districts, can be presented on maps showing the boundaries of the areas. Similarly, data on structures, such as residences or manufacturing plants, can be presented on a schematic of the structure.

Exhibit 3-14 Leading Areas of Fire Origin in Residential Property–1987



Note: Excluding New York City Source: New York State Fire Reporting System

Summary

In this chapter we presented six types of charts: bar charts, column charts, line charts, pie charts, dot charts, and pictograms. The primary purpose of using any chart is to indicate your conclusions more quickly and more clearly than is possible with tables and numbers. You may try several types of charts before you hit on the most appropriate. Be sure not to make your final chart too complicated. The message is what is important, so the chart form should not interfere.

As a quick reference guide on chart selection, we suggest the following:

- Use a **bar chart** when you have categorical data and your objective is to show how the items in a category rank. Most fire data is in categories, such as ignition factor, complex, type of ignition, form of ignition. These are reflected by the NFPA 901 codes.
- Use a **column** or **line chart** when you have data with a natural order, such as hours, months, or age groups. The chart will reflect the general pattern and indicate points of special interest, such as spikes, holes, gaps, and outtiers.
- A **pie chart** is beneficial when the objective is to show how the components relate to the whole. However, we suggest caution with pie charts. Keep the number of components to six or less and avoid forming several pie charts for comparison purposes.
- **A dot chart** depicts the relationship between two variables. Generally these variables are continuous rather than categorical. Population, travel times, and ages are examples of continuous variables. The pattern between the two variables is the important aspect for a dot chart.
- **A pictogram** is a pictorial representation of the data. Breakdowns by geographic areas are effectively shown by a pictogram.

Chapter 3
PROBLEMS

1. The following are monthly data for 1990 from Los Angeles County, California showing the number of structure fires and vehicle fires for 1990. Develop charts for these figures and draw conclusions about the differences in the two distributions.

Month	Structure Fires	Vehicle Fires
January	115	159
February	122	151
March	124	173
April	83	136
May	111	142
June	132	183
July	130	176
August	103	171
September	85	187
October	108	163
November	138	128
December	154	145
Total	1,405	1,914

2. Through the NFIRS system, data are also collected nationwide on civilian injuries and deaths. The following data show age groups for male and female civilians injured in fires during 1989. Develop a chart comparing these distributions and draw conclusions about the differences.

Age	Males	Females
Less than 5 years	138	118
6 to 10 years old	64	35
11 to 15 years old	55	35
16 to 20 years old	83	55
21 to 25 years old	142	87
26 to 30 years old	152	123
31 to 35 years old	147	82
36 to 40 years old	125	54
41 to 45 years old	98	41
46 to 50 years old	71	43
51 to 55 years old	53	16
56 to 60 years old	55	39
61 to 65 years old	56	49
Over 65 years old	121	138
Total	1,360	915

(continued from question 2)

a. Would a pie chart be appropriate for comparing the age groups?
b. would you want to separate the data by sex?

3. The following data give the ages of firefighters injured or killed during 1989

Age of Firefighters	Number Injured
Under 20 years old	32
21 to 2 5 years old	172
26 to 30 years old	366
31 to 35 years old	426
36 to 40 years old	402
41 to 45 years old	253
46 to 50 years old	130
51 to 55 years old	60
Over 55 years old	35
Total	1,876

a. Develop a chart for these data and make conclusions from the chart.

b. Now combine the figures into four age groups: up to 30 years of age, 31 to 40 years old, 41 to 50 years old, and than than 50 years old. Graph these groups with a pie chart.

4. Make charts and comparisons from the following data from four cities on types of fires found during 1990.

Type of Fire	Chicago	Dallas	Detroit	Phoenix
Structure	7,509	2,623	8,144	1,681
Vehicle	9,107	2,663	3,788	1,858
Trees, Brush, and Grass	1,154	2,046	435	2,155
Refuse	14,973	2,611	5,277	3,133
Total	32,743	9,943	17,644	8,827

5. The following figures and percentages are for fires in Los Angeles County for 1990 and Seattle, Washington 1990. An advantage of percentages is that they allow for quicker comparisons between distributions. Compare the percentages between Seattle and Los Angeles by day of week and provide some reasons for the similarities and differences.

Los Angeles County, 1990: Day of Week

Day	Number	Percent
Sunday	855	14.7
Monday	929	16.0
Tuesday	817	14.1
Wednesday	816	14.1
Thursday	778	13.4
Friday	767	13.2
Saturday	838	14.5
Total	5,800	100.0

Seattle, Washington, 1990: Day of Week

Day	Number	Percent
Sunday	358	14.6
Monday	338	13.7
Tuesday	359	14.6
Wednesday	368	15.0
Thursday	324	13.2
Friday	350	14.2
Saturday	362	14.7
Total	2,459	100.0

6. In 1990, the Chicago Fire Department responded to 9,107 vehicle fires. The following shows the amount of incident time in 5-minute increments for the fires. Develop a histogram for these incident times.

Minutes	# of Incidents	Percent
5-10	234	2.5
10-15	1,238	13.6
15-20	2,589	28.4
20-25	2,736	30.0
25-30	1,173	12.9
30-35	654	7.2
35-40	176	1.9
40-45	105	1.2
> 45	202	2.2

7. During 1991, the Memphis Fire Department responded to 500 structure fires in which cooking was the cause of the fire. ‘The following is a randomly selected sample of 30 fires from these 500 fires. The data show the incident times (from time of dispatch to time in service) and the dollar losses for these sampled fires.

Incident Time	Dollar Loss	Incident Time	Dollar Loss
6	50	23	500
9	700	25	1300
11	450	26	1500
11	500	27	135
12	1500	28	1800
13	250	29	700
14	500	32	700
15	600	32	1400
16	250	36	4500
17	100	40	4000
17	400	43	1800
18	300	44	15000
19	1000	45	300
19	1500	60	3000
21	300	78	1800

a. Develop a scatter plot for the 30 incidents.
b. Identify one obvious outlier in the data.
c. Redo the scatter plot without the outlier.
d in general, how arc incident times related to dollar losses?

Chapter 4
BASIC STATISTICS

Types of Variables

For purposes of analysis, fire department variables can be divided into two types: categorical variables and continuous variables. Categorical variables are defined as variables that are classified into groups or categories. For example, fires can be classified into structure fires, vehicle fires, refuse tires, explosions, etc. Categorical variables are sometimes called qualitative variables since we are not measuring quantities, but instead are classifying data into groups. Other examples of categorical variables are day of week, hour of day, month, zip code, fixed property use, ignition factor, method of alarm, area of fire origin, equipment involved in ignition, type of material ignited, construction type, extent of flame damage, and method of extinguishment. Most categorical variables are defined by the NFPA 901 codes, which form the foundation for reporting fire data into the NFIRS system, as described in Chapter 1.

Continuous variables always take on numerical values that reflect some type of measurement. Travel time to fires is an example of a continuous variable. We measure the travel time from the time the first unit is dispatched to a fire to the time the first unit arrives at the scene. It should be noted that the first unit to arrive may not be the same as the first unit dispatched. Other examples of continuous variables are on-scene time at fires, incident time at fires (travel time plus on-scene time), and dollar losses of fires. On-scene time is usually measured as the time from arrival of the first unit to the fire until the time back in service of the last unit. The on-scene time can range from practically no time, if the fire has extinguished itself, to several hours for a major fire.

We should note that if we say the on-scene time is 125 minutes, for example, it is highly unlikely that we are precisely correct. The time could be 124.7 minutes, 125.2 minutes or some other value close to 125 minutes. Most communication centers record individual times in hours and minutes, and we are certainly willing to accept on-scene time to the nearest minute for our analysis.

Categorical variables are variables that are divided into group or categories. Days of the week, months and types of fires are examples of categorical variables.
Continuous variables take an numerical values. Travel times, on-scene times, and dollar losses of fires are examples of continuous variables.

In this chapter we will explore will how to summarize categorical and continuous variables. We will discuss what we mean by an average for a set of data and what we mean by variation of our data about an average.

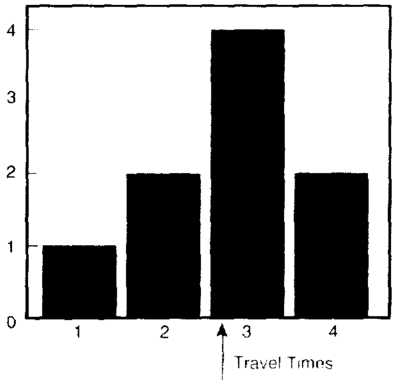
It should be noted that we have been careful with the words **variable** and **data.** A variable is a characteristic that varies or changes. Days of the week vary from Sunday through Saturday; months vary from January through December; and types of fires vary according to NFPA 901 codes such as structure tires, vehicle fires, etc. Whenever we make observations on a variable, we develop data to be analyzed. Each time we complete a fire report, we create data by listing the day of week, hour of day, month, type of situation found, and values for all the other variables in the fire report. The data can then be summarized in a variety of ways, such as histograms and charts. In this chapter, we extend our ideas about summarizing data by introducing averages, standard deviations, box plots, and other techniques. In addition, Chapters 5 and 6 are devoted to analysis of categorical variables through two-way and three-way tables.

Averages: Mean, Median, and Mode

We saw in Chapter 2 and 3 that graphs provide good pictures of an entire set of observations, but we need more descriptive statistics for analysis purposes. In this section, we will develop averages for categorical and continuous variables. An **average** is a single number summarizing a set of data. It represents in a very general manner a “typical” data point. As seen in this section, there is more than one way to calculate an average. In fact, we discuss three different average: mean, median, and mode. Each has advantages and disadvantages, as explained as we go along. Always keep in mind that the aim is to develop a single number, called an average, that best describes the data.

Probably the most commonly known average is the **mean,** or **simple**

Exhibit 4-1 **Hypothetical Travel Times**



average, which is calculated by summing all the data values and dividing by the number of observations. For example, suppose that the travel times to nine incidents are 3 minutes, 2 minutes, 4 minutes, 1 minute, 2 minutes, 3 minutes, 3 minutes, 4 minutes, and 3 minutes. Adding these travel times gives 25 minutes in total and dividing by 9 gives a mean travel time of 2.78 minutes. A histogram of these 9 travel times would look like Exhibit 4-1. Notice that the mean balances the histogram in a seesaw manner.

It should also be noted that a mean can only be calculated with continuous variables, not with categorical variables. We can therefore calculate, for example, mean travel time, mean on-scene time, and mean dollar loss for fires.

> A histogram of continuous data balances at the mean a **mean** can be calculated for continuous variables, but not for categorical variables.

Another type of average is the **median,** which is defined as the middle value (or the 50 percent point) in a group of numbers. For example, we have nine data values for our travel times. If we arrange these in order, they would look as follows: 1, 2, 2, 3, 3, 3, 3, 4, 4. The median is the fifth or middle value, which is 3 minutes in this example. There are four data values to the left of this number, and four values to the right. Another way of saying this is that 50 percent of the data is to the left of the median and 50 percent is to the right. For this reason, the median is also called the 50th percentile.

If we have an even number of data values, then there are two middle values and the median is the mean of the two values. For example, suppose that the on-site times for 10 fire incidents are 12, 15, 17, 25, 27, 29, 32, 35, 37, and 42 minutes. The two middle values are 27 and 29, so the median is 28 minutes (27 + 29 divided by 2). Note that the median again splits the values with five of the data values less than the median and five greater than the median.

As with the mean, the median can only be calculated for continuous variables. WE can determine the median on-scene time at fires and the median dollar loss for fires. However, the “median type of fire” or the “median ignition factor” has no meaning since these are categorical variables.

Two other percentiles frequently calculated from an ordered list of numbers are the 25th and 75th percentiles. Twenty-five percent of the data points are below the 25th percentile and 75 percent are below the 75th percentile. These are called the lower and upper quartiles. The interquartile difference is the difference between these two quartiles (that is, the upper quartile minus the lower quartile). A small interquartile range indicates that the data arc clustered around the median, while a large range reflects a wider spread of the data. In the list of 10 fire incidents above, the 25th percentile is 16 and the 75th percentile is 33.5. The interquartile range is 17.5 which means that half of the data is within a 17.5 point range.

The **median** is the middle value of a group of numbers. The median is also called the 50th percentile since 50% of the data points are lower than this number and 50 percent are higher. The 25th percentile is called the **lower quartile** and the 75th percentile is called the **upper quartile.** the **interquartile difference** is the numerical difference between the upper and lower quartile.

A final type of average is called the **mode,** which is the most frequent data value in your frequency list. It is easily recognized as the peak in a histogram. In Exhibit 4-1, 3 minutes is the mode.

The **mode** is the peak of a histogram.

The mode is the only one of our three averages that can be applied to both categorical and continuous data. With categorical variables, the mode is the group with the highest number of values.

Effects of Extreme Values

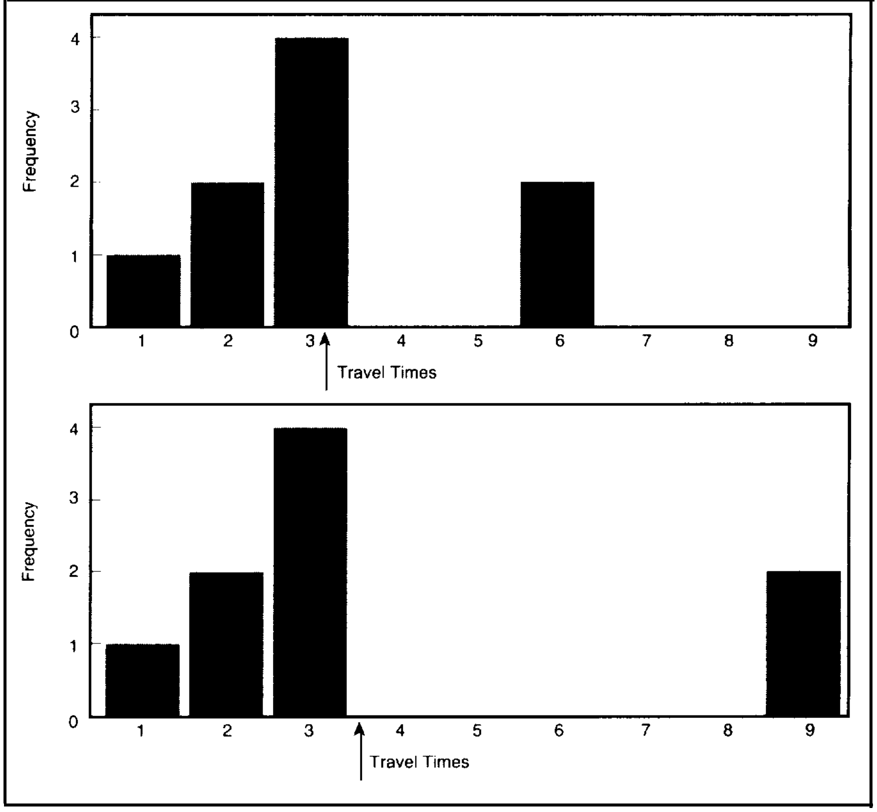
We now want to see how the mean changes with changes in the data. The mean for our nine travel times was 2.78 minutes. In Exhibit 4-2, we have pushed the two travel times of 4 minutes each to 6 minutes. The mean also moves to the right to maintain the balance. Similarly, the bottom part of Exhibit 4-2 shows a mean of almost 4.0 minutes after moving the two travel times to 9 minutes.

The median and mode do not change in Exhibit 4-2. Half the values are still below 3 minutes and half are above 3 minutes, and the peak remains at 3 minutes.

This discussion illustrates that the mean is sensitive to extreme values while the median and mode are not. The sensitivity is especially important with analysis of fire department data. For example, most fires take less than two hours to complete, but a few may take several hours. These extreme fires inflate the mean on-scene time considerably, but have little effect on the median and mode.

If you have this situation, you may want to analyze the fires in two separate groups of on-scene times. In statistical terminology, we use the term **bimodal distribution** to describe data that are really a combination of two types of variables with very different means. Splitting data into two groups is advisable in bimodal situations so that each group can

Exhibit 4-2 Effects of Extreme Data on Averages



be analyzed separately in a more meaningful manner.

Having introduced three different types of averages, the question usually arises as to which average is the best to use. Unfortunately, there is no single answer to this question. A good approach for selection of mean, median, or mode is to look at the distribution of the data, as indicated, for example, by a histogram. If the histogram shows that the data are clustered together with few extreme values, then either the mean or median is a good selection as an average value. On the other hand, we have just seen that extreme values inflate the mean. If extreme values are present in the histogram, then the median or mode may be the best average to represent the data. The mode is particularly useful if a large percentage of the data takes on the value of the mode.

Which is the vest average to **use–mean, median,** or **mode?** The answer is "It depends on the distribution of your data." The average to use is the number that is most representative of the data.

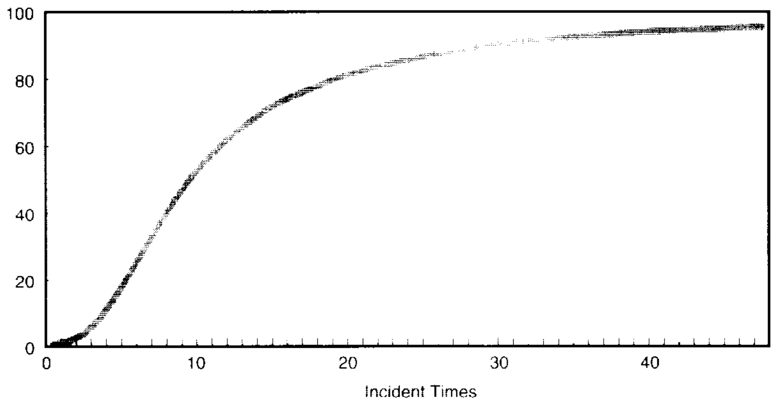
Measuring the Spread of the Data

The purpose of developing an average is to reduce the data to a single representative number. While an average is informative, it is not very satisfactory by itself. Bearing this in mind, statisticians have developed several other measures and graphical techniques to supplement the average. In this section we will explore some of these other measures.

Cumulative Frequencies

Exhibit 4-3 shows **cumulative frequencies** for the incident times from Seattle, Washington for 1990. Incidents include fires, good intent calls, service calls. hazardous conditions calls, and all other types of calls which required a response by the Seattle Fire Department. The department responded to over 13,000 incidents during the year. The incident time is from time of dispatch to time hack in service. It include travel time and on-scene time. For example, suppose the alarm time is 2015, the arrival time is 2019, and the time in service is 2035. The travel time in this example is 4 minutes, the time at the scene is 16 minutes. and the incident time is 20 minutes.

Exhibit 4-3 Cumulative Frequencies for Incident Times-Seattle-1990



You may remember that we introduced cumulative frequencies in Chapter 2. Cumulative frequencies tell you what percent of your data arc less than or equal to a given value. They are also useful for estimating the median of a distribution. In Exhibit 4-3, we can determine the median in the following manner. The left side of the exhibit shows percentages, and the median is the 50 percent mark. Starting at 50 percent we move across until we come to the curve and then move down to the incident time axis. This incident time is the median, which is about 10.2 minutes.

In a similar manner, the 25th percentile is about 6.5 minutes and the 75th percentile is about 17 minutes. This means that 25 percent of the incident times are 6.5 minutes or less and 75 percent are 17 minutes or less.

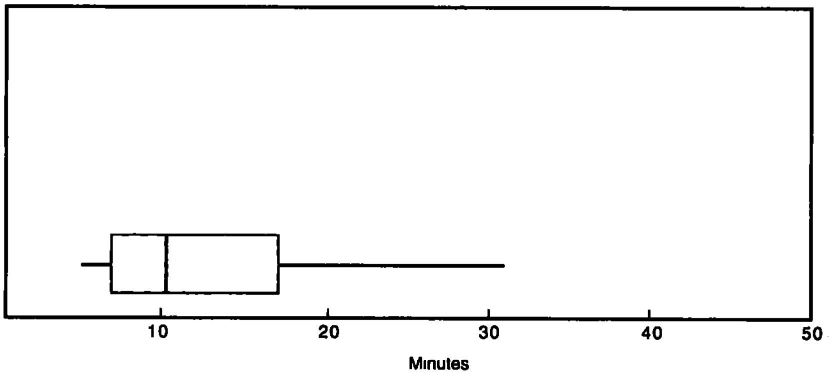
The interquartile range is 10.5 minutes since the upper quartile is 17 minutes and the lower quartile is 6.5 minutes.

Box Plot

A **box plot** is a graphical way of summarizing a continuous variable which takes advantage of percentiles, as shown in Exhibit 4-4.

In this diagram, the left side of the box is the 25th percentile and the right side is the 75th percentile. The vertical line inside the box is the median. Lines extended from each side of the box indicate the 10th and 90th percentiles (about 4.5 minutes and 30 minutes respectively).

Exhibit 4-4 Box Plot of Incident Times

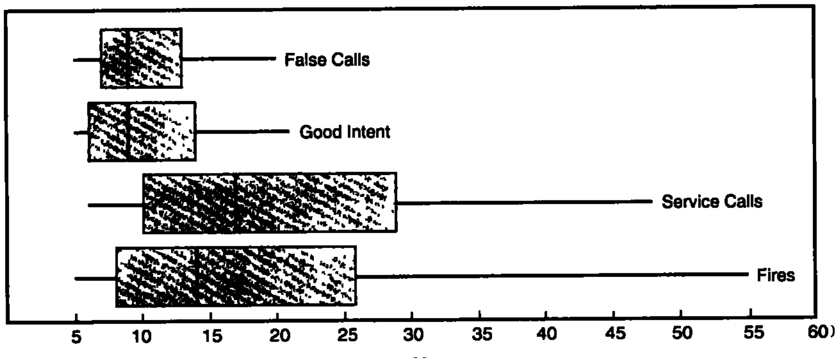


A **box plot** is a graphical summary of a continuous variable. It displays key percentiles to indicate the median and overall spread of the data.

The long "whisker" extending from the right of the box says that 15 percent (from the 75th percentile to the 90th percentile) of the incident times are between 17 minutes and 30 minutes. Another 10 percent of the incident times are more than 30 minutes. Thus, the box plot indicates considerable variability in incident times.

One reason for the variability is the mix of incidents. During 1990, the Seattle, Washington Fire Department responded to 2,459 fires, 5,972 False Calls, 1,371 Service Calls, and 2,340 Good Intent Calls. These incidents have different average times because their on-scene activities differ. Exhibit 4-5 shows box plots for these four types of incidents. Good intent calls and false calls tend to have shorter incident times as indicated by lower medians and less spread in the data. Fire calls and service calls have larger medians and greater variability in their incident times.

Exhibit 4-5 Box Plots of Seattle, Washington Incidents-1990



‘The interquartile range is greater for fires and service calls than the other types of calls.

Variance and Standard Deviation

Another measure of the amount of spread in data is called the sample variance. To illustrate the calculation for sample variance, we go back to the nine hypothetical travel times discussed earlier which had a mean travel time of 2.78 minutes. In the following list, we have subtracted the mean from each individual travel time and then squared the difference.[2]

Exhibit 4-6 Calculation of Variance

Travel Time	Travel Time Mean (2.78)	Squared
1	-1.78	3.17
2	-.78	.61
2	-38	.61
3	.22	.05
3	22	.05
3	.22	.05
3	.22	.05
4	1.22	1.49
4	1.22	1.49
Total	0.00	7.57
Variance		.95

---

[2] The square is the number multlplied by itself. For example. the square of 3 is 9 and the square of .6 is .36.

The middle column displays the amount of deviation from the mean for each point. The first deviation is -1.78 (1 minute minus 2.78 minutes), indicating that this travel time is 1.78 units from the mean and is to the left of the mean (since the sign is negative). Note that the sum of the middle column is zero; that is, the sum of the deviations from a mean adds to zero. In fact, an alternative definition for a mean is that it is the only number with this property; that is, it is the only number where the sum of the deviations is equal to zero.

In the right column, we square each deviation (that is, we multiply each deviation by itself). The sum of the squared deviations is 7.57 and the sample variance is obtained by dividing this sum by 8, which is one less than the total number of points. The variance is therefore 0.95. Since the variance is small compared to the mean, it indicates that the data points are close to the mean.

Finally, a statistic related to the variance is the sample standard deviation, which is the square root of the variance.’ In our example, the standard deviation is .973, since this is the square root of .95. This means that the spread around the average is not very large (in this case less than 1.0 compared to a mean travel time of 2.8 minutes). The mean is therefore a good descriptor of the data in this example.

$$\textbf{Variance} = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

$$\textbf{Standard Deviation} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

For many continuous variables, the sample standard deviation has an interesting property - about 6.5 percent of the data values will be within one standard deviation of the mean. If the mean of a group of numbers is 50 and the standard deviation is 8, then about 65 percent of the data values will be between 42 and 58. Similarly, about 95 percent of the data values will be within two standard deviations.’

The **sample variance** and **sample standard deviation** are measures of the spread of data. A small sample variance indicates that the data points are close to the mean. The sample standard deviation is the square root of the variance.

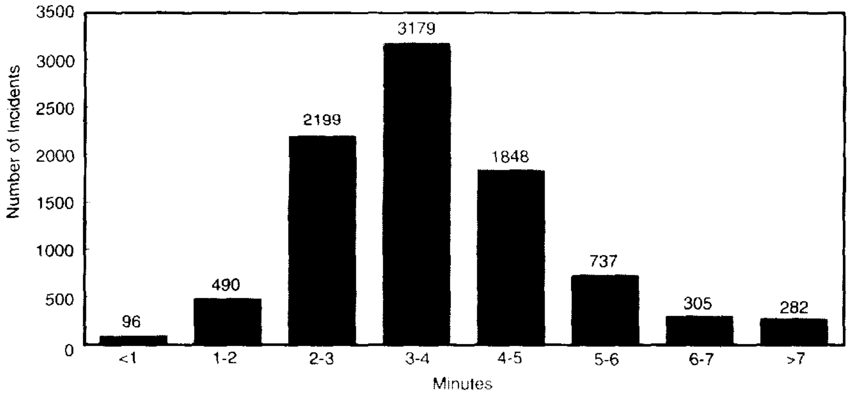
As an example, Exhibit 4-7 shows a histogram of travel times to incidents in Jersey City, New Jersey. The mean is 3.25 minutes and the standard deviation calculates to 1.5 minutes. The data in Exhibit 4-7 generally

---

3. The square root of a number means the number which multiplied by itself gives the original number. The square root of 4 IS 2; the square root of 25 IS .5, etc.

4. Most precisely, in statistical terms, the data values need to follow a normal distribution for this property to be true.

Exhibit 4-7 Travel Times to Incidents-Jersey City, New Jersey



follows the pattern found with normal distributions. One standard deviation about the mean is from 1.75 minutes (3.25-1.5) to 47.5 minutes (3.25+1.5) About 68 percent of the travel times are within this one standard deviation range; about 95 percent of the travel times are within 2 standard deviations of the average.

Sample and Population Variances*

In the previous section, we used the terms sample variance and sample standard deviation. The application of the word sample is important here from a statistical viewpoint. We are assuming that the data we have are a sample of a larger set of data, such as the population of all incidents. When we say we have nine travel times, for example, we mean that we have a sample of nine travel times which we have selected to study. The travel times should be randomly selected so that they are representative of all travel times.

In algebraic terms, the sample variance is expressed as:

$$\textbf{Sample Variance} = s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

The sample standard deviations(s) is the square root of the sample variance

If we are not working with a sample of data, but instead have the entire population, then the variance is calculated in a slightly different manner.

The **population variance,** usually designated as ó^2 is expressed on the following page.

$$\textbf{Population Variance} = \sigma^2 = \frac{\sum (x_i - u)^2}{N}$$

where N is the number of points in the population and a is the mean of the population. The **population standard deviation** (σ) is the square root of the population variance.

Note that the main difference between these two equations is that the sample variance is obtained by dividing by n-1 (the number of points in the sample), while the population variation is obtained by dividing by N (the total number of points).

With the calculation of a sample variance, we are really developing an estimate of the population variance. The reason for dividing by n-1 to obtain the sample variance is that statistical theory tells us that this gives a better estimate of the population variance than dividing by n. This difference may appear to he slight and insignificant, but it is important from a theoretical viewpoint.

If you are using a statistical package on a computer, it is important for you to know whether the output is the sample variance or the population variance. Unfortunately, many statistical packages will merely say “Variance” in the output without indicating whether it is the sample or population variance. One package may give the sample variance while the other may give the population variance. “The documentation to the package usually states which variance is calculated.

Indexes for Time Series*

Development of an index is another approach for understanding a frequency list of numbers. Indexes arc usually calculated for time series data, such as the distribution of fires by hour of day or by day of week. To develop an index, you divide each number in a series by the overall mean for the series. For example, Exhibit 4-8 shows the number of fires in 1990 in Chicago, Illinois. The table shows a total of 33,130 fires for the year which gives a mean of 2,760.8 fires per month. The column on the right is the index for each month. To develop the index for January, for example, we divide the number of fires in January by the mean:

$$\textbf{January Index} = \frac{2,406}{2,760.8} = .87$$

The indexes for the remaining months arc calculated in the same manner.

There are two interesting features of indexes. First, an index greater than I .0 means that the month is above average while an index below 1.0

means the month is below average. This property is easy to see since we divided by the mean to get the index In Exhibit 4-8, we easily see that there are six months above average and six months below average. The second interesting feature of an index is that the index minus one gives the percentage that the month is above or below the mean. For example, July has an index of 1.25 and if we subtract 1.0 from this index, we can say that July is 25 percent above average. Similarly, January is I3 percent below average.

July	1.25 - 1 .00 = .25	25% above the mean
January	.87-1.00=-.13	13% below the mean

In summary, indexes are just another way of getting a feel for data. They are easy to calculate and easy to understand.

Exhibit 4-8 Index for Fires in Chicago–1990

Month	Total Fires	Index
January	2,406	.87
February	1,884	.68
March	2,522	.91
April	2,905	1.05
May	2,796	1.01
June	3,193	1.16
July	3,448	1.25
August	2,744	.99
September	3,118	1.13
October	2,689	.97
November	2,809	1.02
December	2,616	.95
Total	33,130	
Average	2,760.83	

Summary

In this chapter we have introduced several techniques for analyzing data. You should, at this point, understand how to construct a histogram and cumulative frequencies. You should also be able to estimate percentiles from cumulative frequencies and develop box plots.

The most useful statistics from a group of numbers are the mean, median, and mode. Be sure you understand how to calculate and interpret these numbers before you go further in this handbook. You should also understand how to calculate variance and standard deviations. These are particularly important in Chapter 7 on correlation and regression.

Chapter 4

PROBLEMS

1. The following data show travel time, on-scene time, and dollar loss for 27 hotel and motel fires (Fixed Property Use between 440 and 449) in Phoenix, Arizona during 1990.

Travel Time	On-Scene Time	Dollar Loss	Travel Time	On-Scene Time	Dollar Loss
4	26	$1,500	4	10	1,000
3	27	100	5	11	300
6	327	1,000	4	74	20,000
3	94	300	3	35	2,000
6	98	150	5	88	30,000
4	74	15,000	4	112	3,000
4	23	189	4	113	500
1	85	5,000	4	10	2,000
4	22	100	4	69	250
4	83	5,000	5	33	7,000
5	9	50	5	7	1,000
3	33	1,000	1	20	3,000
4	92	500	3	62	10,000
3	48	23,000			

a. Calculate the mean, median, and interquartile range for each variable.

b. Calculate the variance for travel time and dollar loss.

c. How many dollar loss observations are within one standard deviation from the mean? How many within two standard deviations?

d. Why is the variance for travel time so much less than for dollar loss?

e. Comment on the effects of large dollar losses on the mean and median.

2. Construct a histogram with the dollar loss data from the previous exhibit. From the histogram, construct a cumulative frequency distribution. Estimate the l0th, 25th, median, 75th, and 90th percentile from the cumulative frequency distribution.

3. The following data show selected percentile for on-site time and dollar loss for fires in 519 one- and two-family dwellings (Fixed Property Use between 410 and 419) and 154 apartments, tenements, and flats (Fixed Property Use between 420 and 429) in Phoenix, Arizona during 1990.

On-Site Time

	Percentiles				
Type	10th	25th	50th	75th	90th
Dwelling	240	45.0	85.0	123.0	163.0
Apartments	20.5	32.0	71.5	107.0	143.0

Dollar Loss

	Percentiles				
Type	10th	25th	50th	75th	90th
Dwelling	1,500	3,000	10,000	25,000	48,000
Apartments	1,500	2,500	5,000	12,400	40,000

a. Develop box plots for both variable on-site times, and dollar losses.

b. Compare the interquartile ranges of dwellings versus apartments for the two variables.

c. What are your conclusions from the box plots?

d. Based on the box plot, would you expect a greater variance for dollar loss with dwellings or with apartments?

4. In Chapter 2, we discussed skewness and showed that travel times were skewed to the right, i.e. , skewed toward high values. The Pearson coefficient of skewness has been proposed by some authors as a measure of the skewness of a distribution, Denoted by CS, the coefficient of skewness is defined as

$$\text{CS} = \frac{3(\text{Mean - Median})}{\text{Standard Deviation}}$$

a. Calculate the CS for the on-scene times and dollar losses in problem 1 above.

b. When will the CS be zero?

5. A trimmed mean is a compromise average between the mean and median which is useful when outliers are present. A trimmed mean is computed by first ordering the data values from smallest to largest, deleting a selected number of values from each end of the ordered list, and finally averaging the values not deleted.

   a. Using the dollar loss data from problem 1, calculate the trimmed mean by deleting the two largest and two smallest values.

6. The following travel time data arc from Monroe Country, New York.

Cumulative Travel Times: Monroe County-1990

Cumulative Travel Time	Cumulative Frequency	Frequency	Percent
Less than 1 minute	157	157	7.5
1 to 2 minutes	287	444	21.2
2 to 3 minutes	324	768	36.7
3 to 4 minutes	322	1,090	52.2
4 to 5 minutes	271	1,361	65.1
5 to 6 minutes	227	1,588	76.0
6 to 7 minutes	176	1,764	84.4
7 to 8 minutes	120	1,884	90.1
8 to 9 minutes	99	1,983	94.9
9 to 10 minutes	45	2,028	97.0
10 to 11 minutes	25	2,053	98.2
11 to 12 minutes	16	2,069	99.0
12 to 13 minutes	11	2,080	99.5
13 to 16 minutes	10	2,090	100.0
Total	2,090		

   a. Plot the cumulative distribution percentages. Then develop a box plot with estimates from the cumulative distribution.

   b. By looking at your chart, what percent of travel times are under 8 minutes?

   c. Estimate the median using your chart.

Chapter 5
ANALYSIS OF TABLES

Introduction

As we discussed in Chapter 1, the NFPA 901 Codes form the basis for coding fire and injury reports. They cover all the main variables associated with fires, including type of fire, fixed property use, ignition factor, type of complex, area of fire origin, type of material ignited, extent of flame damage, and many others.

For each variable, the 901 Codes provide specific categories, as exemplified by the categories for ignition factors shown in Exhibit 5-1:

Exhibit 5-1 Ignition Factor

Code	Category
1	Incendiary
2	Suspicious
3	Misuse of Heat of Ignition
4	Misuse of Material Ignited
5	Mechanical Failure, Malfunction
6	Design, Construction, Installation Deficiency
7	Operational Deficiency
8	Natural Condition
9	Other Ignition Factors

Note that there is a numeric code for each category. However, from an analysis viewpoint, the code numbers have no statistical meaning. It is not reasonable to say that "Misuse of Heat of Ignition" is one more than "Suspicious" or to observe that "Design, Construction, Installation Deficiency" is twice "Misuse of Heat of Ignition." It is wrong in just the same way to perform statistical analysis with code numbers by calculating means and standard deviations. We cannot, for example, calculate the "Average Ignition Factor. " Thus, the code numbers tell us nothing statistically, but instead serve as a convenience when entering data into a computer. That is, it is much easier to enter a single number instead of entering the category name. There is also a considerable savings in the amount of computer storage required for the data when we can use a single number rather than storing a long name.

We make use of these codes by generating tables showing the number of observations (e.g., fires) for each category. We might find, for example, that the incendiary category accounts for 17 percent of the ignition, factors of fires, suspicious accounts for 23 percent, and so on.

In this chapter we will provide basic techniques for analyzing tables developed from categorical data. The first part of the chapter describes the development and interpretation of percentagcs for categorical data. We then develop a statistical test, called the chi-squared test, for determining whether the percentage distribution from a table differs significantly from a distribution of hypothetical or population percentages.

To summarize the terminology in this chapter, **variable** refers to a characteristic of fires and fire injuries. Each variable has several **categories** with a numeric code for each category. ‘Thc codes have no statistical meaning, but assist in getting the data into a computer for analysis. Tables derived from the categories serve as the basis for calculating percentages and performing chi-squared tests of significance.

A **categorical variable** divides a variable into a group of categories. Each category has a code assigned to it for convenient entry and storage in a computer. **Tables** can be developed showing the number of observations for each category. From the tables, we can calculate percentages and perform chi-squared tests of significance.

Describing Categorical Data

To summarize a categorical variable, we usually report the number of observations in each catepory- and its percentage of the total. For example, consider Exhibit 5-2 for types of situations found in the fires of’ Seattle, Washington during 1989. These percentages arc simple to calculate and easy to understand: 36.3 percent of the fires are structure fires, 25.9 percent are vehicle fires, and so on. As described in Chapter 2, the mode is the category with the largest number of data values. In this example, the mode is structure fires, totaling 1,l95 fires.

Exhibit 5-2 Type of Situations Found-Seattle Fires-1989

Code	Type of Fire	Number	Percent
11	Structure Fires	1,195	38.3
12	Outside of Structure Fires	164	5.0
13	Vehicle Fires	850	25.9
14	Trees, Brush, Grass Fires	525	16.0
15	Refuse Fires	512	15.6
XX	Other Fires	42	1.3
	Total	3,288	100.0

Note that the percentages will not always add up to exactly 100 percent. This is because of rounding the individual calculations. The rounding may cause the total to be slightly more or slightly less than 100.0 percent. In this exhibit, the total is actually 100.1 percent, but we show 100.0 for convenience and consistency.

By way of comparison, Exhibit 5-3 shows the nationwide picture of types of situations found for fires. From a national perspective, structure fires accounted for 32.8 percent of the total, followed by vehicle fires at 23.9 percent, and trees, brush, and grass fires at 2 1.1 percent.

Exhibit 5-3 Type of Situations Found-Nationwide Fires-1989

Code	Type of Fire	Number	Percent
11	Structure Fires	302,708	32.8
12	Outside of Structure Fires	26,436	2.9
13	Vehicle Fires	220,861	23.9
14	Trees, Brush, Grass Fires	195,121	21.1
15	Refuse Fires	162,835	17.6
x x	Other Fires	15,922	1.7
	Total	923,883	100.0

A reasonable question to ask is whether the distribution of fires in Seattle differs from the national picture. We notice some differences by comparing percentages. For example, 36.3 percent of the Seattle fires are structure tires, compared to 32.8 percent nationwide. Similarly, 25.9 percent of the Seattle fires are vehicle fires, compared to 23.9 percent nationwide. We therefore suspect that the distribution of fires in Seattle deviates from the national picture, but a statistical test can be made to test this difference more precisely. In the next section we will provide such a test.

The Chi-squared Test

The chi-squared test is a statistical test to determine whether a sample set of observations has a different distribution from a hypothetical or population set of observations. The test is usually stated in more precise statistical language by defining an hypothesis to be tested. For our purposes, the **null hypothesis, H_0,** is that the percentage distribution from the sample distribution does not differ significantly from the national percentages. The **alternative hypothesis, H_1,** is that there is a significant difference between the two distributions.

To illustrate these ideas, we will deviate from our usual practice of showing examples with fire data. Instead, we will consider a simple experiment where we throw a die over and over again. The resulting data values are the number of dots showing after each throw. The number of dots varies between 1 and 6; that is, we have six possible outcomes. If u-e throw a "fair" die a large number of times, we wou1d expect that about one-sixth of them would result in I dot showing, one-sixth in 2 dots showing, etc. The chi-squared test allows us to determine with some assurance whether we have a fair die.

> The **chi-squared test** determines whether a given distribution differs significantly from an hypothesized or population distribution. The null hypothisis is that no difference exists. while the alternative hypothesis is that there is a difference.

Suppose that we throw the die 90 times and obtain the results in Exhibit 5-4.

Exhibit 5-4 Results of Die Throws

Dots Visible	Number	Percent
One	16	17.8
Two	17	18.9
Three	12	13.3
Four	14	15.6
Five	17	18.9
Six	14	15.6
Total	90	100.0

If the die is a fair die, we should expect to have one dot visible exactly 15 times (one-sixth of the total), two dots visible exactly 15 times, and so on. Our actual results differ from these expected results as shown in Exhibit 5-5

Exhibit 5-5 Actual and Expected Results

Dots Visible	Actual Number	Expected Number
One	16	15
Two	17	15
Three	12	15
Four	14	15
Five	17	15
Six	14	15
Total	90	90

To summarize, we have tossed a die 90 times and obtained the results shown in Exhibit 5-4. Out null hypothesis is that the die is fair, which means that we expect the outcomes to be equally likely at one-sixth of the throws resulting in each possible outcome. The actual results are not the same as the expected either because of variations inherent in throwing a die only 90 times or because the die is not a fair die. The chi-squared test will determine whether the actual results differ significantly from the expected results.

To perform the chi-squared test, we do the following steps:

1. Calculate the **expected number** for each category by multiplying the expected or population percentages by the total sample size. This calculation has already been performed as shown in Exhibit S-S with the "Expected Number" column.
2. For each category, subtract the expected number from the actual number, and then square the result.
3. Divide the results from Step 2 by the expected number.
4. Sum the results from Step 3. This is the **calculated chi-squared statistic.** The larger this number, the more likely there is a significant difference between the observed and expected values. However, the chi-squared statistic also depends on the number of categories, which must be taken into account in the following steps.
5. Find the **degrees of freedom,** which is defined as the number of categories minus 1. In our die example, there are 5 degrees of freedom.
6. Obtain the **critical chi-squared value** from Appendix A by selecting the entry associated with the degrees of freedom. Now compare the computed chi-squared statistic from Step 4 to the critical value.

   If the computed chi-squared statistic is **greater** than the value in the table, then we **reject** the null hypothesis. Otherwise, we accept the null hypothesis. To reject the null hypothesis means that the two distributions differ significantly. To accept the null hypothesis is to say that the two distributions are essentially the same with differences due to sampling or random variations.

The **expected value** is the number we would expect if the null hypothesis were true. The chi-squared value is calculated by first subtracting the expected number from the observed number, squaring the result, and then dividing by the expected number. The results from these calculations are then summed to obtain the **calculated chi-squared value.**

Exhibit 5-6 on the following page summarizes these steps for the die example. The "Diff." column shows the difference between the expected and actual numbers. The "Squared Diff." is the square of the difference obtained by multiplying the number by itself. The right-most column is the

Exhibit 5-6 Actual and Expected Results Die Tossing Experiment

Dots Visible	Actual Number	Expected Number	Diff.	Squared Diff.	Divided by Exp.
One	16	15	1	1	.067
Two	17	15	2	4	.267
Three	12	15	-3	9	.600
Four	14	15	-1	1	.067
Five	17	15	2	4	.267
Six	14	15	-1	1	.067
Total	90	90			1.34
Chi-squared Value			1.34		
Degrees of Freedom = 6 - 1			5		
Critical Chi-squared Value			11.07		

squared difference divided by the expected number; for example, the first figure is .067 obtained from 1 divided by 15.

The chi-squared value is 1.335, which is the sum of the values in the last column. In summary, the chi-squared value is given by:

$$\boxed{\text{Chi-squared Value} = \sum \frac{(\text{observed - expected})^2}{\text{expected}}}$$

From Appendix A, the critical chi-squared value for 5 degrees of freedom is 11.07. Since the calculated chi-squared value of 1.335 is less than this value, we accept the null hypothesis. That is, the results from the ninety throws do not provide evidence that the die is unfair.

Degrees of freedom have been defined as the number of categories minus one. The reason for this definition is as follows. Each category may be considered as contributing one piece of information or one degree of freedom to the chi-squared statistic. The exception is the last category which is not considered to he free because the total sample size is a fixed number. Consequently, the last category can he determined from the total sample size and the numbers in the other categories. Thus, the values in all categories except one can take on any values. We can see this same situation by looking at the calculated percentage. We know that the percentages must total to 100 percent. If the percentages are known except for one category, we can immediately calculate the percentage for the one remaining category?. If we have four categories, the first three percentages are free to vary, but the percentage for the last category is automatically determined. If the first three percentages are 25 percent, 30 percent and 35 percent, then the last category must be 10 percent so that the total sums to 100 percent.

> **Degrees of freedom** for a single list of items are equal to the number of items minus one. The term derives from the fact that all the items can vary except for one item (since the total is fixed).

We can now return to our question on whether the distribution of fires in Seattle differs from the nationwide distribution of fires. We noted differences in some categories; for example, Exhibit 5-2 shows that structure fires account for 36.3 percent of the fires in Seattle compared to 32.8 percent nationwide. Similarly, vehicle fires account for 25.9 percent of the fires in Seattle compared to 23.9 percent nationwide.

However, these are individual comparisons. The chi-squared test allows us to test all categories simultaneously. Our null hypothesis is that “The percentage distribution of fires in Seattle does not differ significanty from the nationwide picture” against the alternative hypothesis that “The percentage distribution of fires in Seattle differs from the nationwide picture.” If the calculated chi-squared value is larger than the appropriate value in Appendix A, we will reject the null hypothesis: otherwise, we cannot reject the hypothesis.

Exhibit 5-7 shows the calculations using the information in Exhibits 5-2 and 5-3.

Exhibit 5-7 Actual and Expected Results-Seattle Fires-1989

Type of Fire	Actual Number	Expected Number	Diff.	Squared Diff.	Divided by Exp.
Structure	1,195	1,078.4	116.6	13,595.6	12.6
Outside	164	95.4	68.6	4,706.O	49.3
Vehicle	850	785.8	64.2	4,121.6	5.2
Grass	525	693.8	-168.8	28,493.4	41.1
Refuse	512	578.7	-66.7	4,448.9	7.7
Other	42	55.9	-13.9	193.2	3.5
Total	3,288	3,288.0			119.4

Chi-squared Value	119.4
Degrees of Freedom=6 - 1	5
Critical Chi-squared Value	11.07

The "Actual Number" column comes directly from Exhibit 5-2. To obtain the expected number, we apply the percentages from Exhibit 5-3 to the 3,288 Seattle fires. For example, 32.8 percent of the nationwide fires were structure fires, which means wc expect 32.8 of the 3,288 fires in Seattle to be structure fires. This calculation gives 1 ,078.4 fires (32.8 percent times 3,288 fires).

The "Diff." column gives the difference between the actual and expected numbers and the next column is the squared difference (the difference multiplied by itself). The last column is the squared difference divided by the expected value. The calculated chi-squared value is the sum of the column, which is I19.4.

In this example, we have six categories of fires, which means we have five degrees of freedom. From Appendix A, the critical ch-squared value is 11.07. Since our calculated chi-squared value of 119.4 is greater than the critical value, we reject the null hypothesis. Our conclusion is that the distribution of fires in Seattle differs from the nationwide picture.

Other information is available from Exhibit 5-7 with regard to these differences. For example, the difference between the actual and expected number of structure fires is 116.6. Squaring this difference and dividing by the expected number gives 12.1, as shown in the last column. Even though there is a fairly large difference in this category, its contribution to the chi-squared value is not large. We expect large categories to have greater numerical variation than small categories, and for this reason, the calculation of the chi-squared value uses counts rather than percentages. On the other hand, the main contributors to the large chi-squared value are outside fires (49.3) and grass fires (41. 1). With Grass Fires, the difference between the actual and expected numbers is large (ignoring the minus sign) relative to the base values and the relatively low volume of fires creates the large contribution to the chi-squared value.

As another example of the ch-squared test, Exhibit 5-8 shows the number and percent of calls by month handled by a group of departments in Florida. This exhibit shows a flat distribution with roughly the same number of incidents each month. An exactly even distribution of calls would assign 8.33 percent of the calls to each month (100.0 percent divided by 12).

We can use the chi-squared procedure to test whether the observed distribution deviates significantly from an expected distribution having the same percent of incidents each month. Exhibit 5-9 shows the calculations for this test. The null hypothesis is that the actual distribution of incidents per month is the same as an expected distribution reflecting the same percent of incidents each month.

---

5 These numbers include all types of calls for which the fire department responded fire calls. rescue calls. hazardous conditions service calls. good Intent calls. false calls. etc

Exhibit 5-8 Fire calls for Selected Cities in Florida-1989

Month	Number	Percent
January	1,193	8.9
February	1,082	8.1
March	1,197	9.0
April	1,109	8.3
May	1,096	8.2
June	1,139	8.5
July	1,060	7.9
August	1,065	8.0
September	1,070	8.0
October	1,123	8.4
November	1,105	8.3
December	1,126	8.4
Total	13,365	100.0

Exhibit 5-9 Actual and Expected Results-Fire Call in Florida-1989

Month	Actual Number	Expected Number	Diff.	Squared Diff.	Divided by Exp.
January	1,193	1,113.7	79.3	6,288.5	5.65
February	1,082	1,113.7	-31.7	1,004.9	0.90
March	1,197	1,113.7	83.3	6,938.9	6.23
April	1,109	1,113.7	-4.7	22.1	0.02
May	1,096	1,113.7	-17.7	313.3	0.28
June	1,139	1,113.7	25.3	640.1	0.57
July	1,060	1,113.7	-53.7	2,883.7	2.59
August	1,065	1.113.7	-48.7	2,371.7	2.13
September	1,070	1,113.7	-43.7	1,909.7	1.71
October	1,123	1,113.7	9.3	86.5	0.08
November	1,105	1,113.7	-8.7	75.7	0.07
December	1,126	1,113.7	12.3	151.3	0.14
Total	13,365	13,365.O			20.37

Chi-squared Value	20.4
Degrees of Freedom = 12 - 1	11
Critical Chi-squared Value	24.7

In this case, the calculated chi-squared value (20.4) is less than the critical chi-squared value (24.7). We therefore accept the null hypothesis.

That is, we conclude that the distribution of incidents by month does not differ significantly from an equal distribution.

This conclusion makes sense because of the weather conditions in Florida. The temperahire does not vary as greatly between winter and summer as in other states. For this reason, there may not he as much seasonal variation in fire department activity.

Technical Notes About the Chi-squared Test*

The chi-squared statistic is based on the **chi-squared distribution,** which is a well-known distribution to theoretical statisticians. If the null hypothesis, Ho, is true and if the sample size is sufficiently large, then the sampling distribution of the calculated chi-squared value is approximately the chi-squared distribution. There is a different chi-squared distribution for each degree of freedom. The primary attribute of the chi-squared statistic is that it tends to be large when the observed percentages are different from the hypothesized values. We reject the null hypothesis, which states they are all equal, when the calculated ch-squared value is “larger than reasonable.”

The table in Appendix A is based on a 5 percent level of significance. To understand what we mean by this level of significance, we need to consider the die throwing experiment again. If we were to throw the die another 90 times, our results would probably be different. In fact, in some instances, the results would lead to a large chi-squared value, leading to a rejection of the null hypothesis of a fair die **even though the die was, in fact, fair.** This is called a Type 1 error. At first glance, this would appear to be a curious result, but is an accepted fact of life in statistical situations where random fluctuations play an important role.

When we state that a test is conducted at a 5 percent level of significance, we are saying that we expect to make an incorrect decision 5 percent of the time. That is, random fluctuations in the sampling procedures will result in an incorrect decision 5 percent of the time.

The choice of a 5 percent level of significance is common among statisticians in social science testing. It is not, however, always necessary) to select the 5 percent level. We can use a Type 1 error of 1 percent level, 10 percent level, or even 20 percent. The selection depends on the amount of risk we are willing to take in drawing an incorrect conclusion. At the I percent level, the critical ch-squared valuer, will be larger since we are saying that we will risk a wrong decision only 1 percent of the time. The critical values for the I percent level are presented in Exhibit 5-10.

Exhibit 5-10 Critical Chi-squared Values-1 Percent Significance Level

Degrees of Freedom	Critical Chi-squared Values
1	6.6
2	9.2
3	11.3
4	13.3
5	15.1
6	16.8
7	18.5
8	20.1
9	21.7
10	23.2
11	24.7
12	26.2

A review of these figures shows that a larger calculated chi-squared value is needed in order to reject the null hypothesis. A 20 percent level provides smaller critical chi-squared values, but selecting such a high level is usually foolish since it means we risk a wrong decision 20 percent of the time.

As a final note, the chi-squared test should not be used with small samples. It is an excellent test when the sample size is large, but is not valid when the sample size is small. As the sample size increases, statistical theory says that the computed chi-squared statistic follows more and more closely to a chi-squared distribution when the null hypothesis is true. The best rule of thumb is to avoid using the chi-squared test when fewer than 5 cases are expected in any category. In some situations, it may be advisable to lump small categories together into an “Other” category prior to the chi-squared test.

Two-way Tables

In this section, we will extend our ideas to two categorical variables. By studying the two variables together rather than separately, we can measure the **statistical association** between the two variables. By association, we mean that knowing the value of one variable gives us information about the other variable. In some instances, there may he no association whatsoever, and we will be able to recognize this situation. Another result may be that the association exists, but is weak. Finally, the association between two variables may be quite strong, and we will measure the strength of this association.

Exhibit 5-11 will serve as the starting point for to to introduce the concepts in this chapter. NFIRS data for 1989 on civilian injuries (not including deaths) was the source for developing the figures in this table. Each record in the database contains information on the location of the person at the time of the injury and the nature of the injury.

There are four categories for location. The first indicates that the person was intimately involved with the fire ignition. A common example is a person who is) burned as a result of accidentially spilling grease on a kitchen stove. This category also includes injuries from ignition of clothing on a person and from ignition of bedding or furniture on which a person is sitting or lying. The next category covers situations whcre the injured person is in the same room or space of the fire. but was not directly involved in the ignition. An example would be smoke inhalation by someone in the kitchen, but not directly at the stove when the grease fire occurred. The last two categories cover when the injured person is either on the same floor as the fire origin or on a different floor of the building.

Exhibit 5-11 Civilian Injuries-1989-Location and Nature of Injury

	Nature of Injury				
Location	Burns Only	Smoke Only	Burns & Smoke	Other	Row Totals
Fire Casualty intimately Involved with Ignition	238	34	116	6	394
Fire Casualty in the Room or Space of Fire Origin	130	100	100	10	340
Fire Casualty on Same Floor as Fire Origin	36	205	110	39	390
Fire Casualty in Same Building as Fire Origin	33	190	79	42	344
Column Totals	437	529	405	97	1,468

The nature of the injury is also divided into four categories (1) burns: only, (2) smoke/asphyxia only, (3) burns and smoke/asphyxia, and (4) other. The first three categories are self-explainatory. “Other” category includes injuries from shock, cuts. dislocations, fraction, and complaints of pain.

Exhibit 5-11 shows that a total of 1,468 injured persons. The top left number means there were 238 persons who were intimately involved in the fire and suffered injuries of burns only. This number is, in fact, the mode of the two-way classification (although the mode does not always have to appear as the first number in the table).

With the exception of identifying the mode, the numbers in the table do not relay much information. In the next section, we will calculate various percentages from this table, which will provide more insight. ‘Then we will calculate a chi-squared value to measure the strength of the relationship between the two variables.

Percentages for Two-way Tables

There arc three different ways to calculate percentages for a two-way table of counts. Each way highlights a different feature of the table. More importantly, each provides a different interpretation of the data and leads to different conclusions about the relationship between the two variables. The three ways of calculating percentages are:

- Joint percentages
- Row percentages
- Column percentages

You select the type of percentage you want depending on what you are trying to show from the data. Joint percentages allow you to compare the entries in the table directly with each other. Ron percentages concentrate on each row of the table with percentages along that row summing to 100.0. In a similar manner, column percentages fix on each column of the table with percentages down the column summing to 100.0.

Three types of percentages are possible with two-way tables: joint percentages, row percentages, and column percentages. **Joint percentages** are formed by dividing each number in the table by the grand total. **Row percentages** are obtained by dividing each number in a row by the row total, and **column percentages** are obtained by dividing each number in a column by the column total.

Joint Percentages

To develop joint percentages, we divide each entry in the table of counts by the overall total. Exhibit 5-12 shows the calculation for the counts from Exhibit .5- 11. The top left entry is simply:

$$\frac{238}{1468} = 16.2 \text{ percent}$$

This tells us that 16.2 percent of the total persons injured were intimately involved in the fire’s ignition and suffer-cd burn injuries. The sum of all the entries in the table is 100.0 percent.

With the calculation of joint percentages, we have increased our knowledge because we can now make more logical comparisons. The table tells us, for example, that 14.0 percent of the persons were on the same floor as the fire origin and had smoke injuries. In a similar manner, only 2.2 percent of the persons were in the same building as the fire origin and had burn injuries.

Exhibit 5-12 Civilian Injuries-1989-Joint Percentages

	Nature of Injury				
Location	Burns Only	Smoke Only	Burns & Smoke	Other	Row Totals
Fire Casualty Intimately Involved with Ignition	16.2	2.3	7.9	.4	26.8
Fire Casualty in the Room or Space of Fire Origin	8.9	6.8	6.8	.7	23.2
Fire Casualty on Same Floor as Fire Origin	2.5	14.0	7.5	2.7	26.6
Fire Casualty in Same Building as Fire Origin	2.2	12.9	5.4	2.9	23.4
Column Totals	29.8	36.0	27.6	6.6	100.0

Exhibit 5-I2 also provides important information from the row and column totals. For example, from the first row we find that 26.8 percent of the persons injured were intimately involved in the ignition. ‘This percent can be derived in two ways. One way is to add the four percentages across the row (7.9 + 16.2 + 2.3 + 0.4 = 26.8). The other way is to divide the row total of 394 (from Exhibit 5-1I) by 1,468 to give 26.8 percent. Exhibit 5-I2 shows a rather equal distribution across the four Iocation catcgorics:

- 26.8 percent intimately involved in ignition
- 23.2 percent in the room or spacc of fire origin
- 26.6 percent on the same floor as the tit-c origin
- 23.4 percent in the same building as fire origin

In a similar manner, we have column percentages which provide information. For example, 29.8 percent of the persons injured suffered from burns only, 36.0 percent from smoke only, 27.6 percent with burns and smoke, and 6.6 percent with other injuries.

While Exhibit 5-12 provides mot-c insight into these two variables, it does not directly address other questions. For example, we cannot immediately compare burn injuries with smoke injuries for persons in the same room or space of fire origin. Similarly, we cannot compare persons located on the same floor with persons located in the same building for other injuries. These comparisons require calculation of row and column percents, as described in the following sections.

Row Percentages

To convert a table of counts into row percentages, we divide each entry in the table of counts by its row total. The top left entry is calculated by:

$$\frac{238}{394} = 60.4 \text{ percent}$$

This tells us that 60.4 percent of the total persons who were intimately involved in the fire ignition suffered burn injuries.

A table of row percentages allows for comparisons among the categories represented by the rows. The total for each row is 100.0 percent, and this figure appears on the right of the table as a reminder that we have developed row percentages.

As indicated, 60.4 percent suffered burn injuries when they were intimately involved with the fire's ignition. A total of 29.4 percent had burn and smoke injuries, 8.6 percent had smoke injuries and only I.5 percent had other injuries. These percentages account for all the injuries of persons intimately involved in the fire's ignition.

Exhibit 5-13 Civilian Injuries-1989-Row Percentages

	Nature of Injury				
Location	Burns Only	Smoke Only	Burns & Smoke	Other	Total
Fire Casualty Intimately Involved with Ignition	60.4	8.6	29.4	1.5	100.0
Fire Casualty in the Room or Space of Fire Origin	38.2	29.4	29.4	2.9	100.0
Fire Casualty on Same Floor as Fire Origin	9.2	52.6	28.2	10.0	100.0
Fire Casualty in Same Building as Fire Origin	9.6	55.2	23.0	12.2	100.0
Overall	29.8	36.0	27.6	6.6	100.0

Selecting the third row, which is for persons injured on the same floor as the fire's origin, a different picture emerges. Burns and smoke injuries account for 52.6 percent of the total, followed by 28.2 percent for burn injuries, and about 10.0 percent for the other two injury categories. Once again, these percentage total to 100.0 percent to account for all persons injured while on the same floor as the tire's origin.

It should he noted that we have repeated the overall percentages along the last row from the table of joint percentages. We can then make comparisons of the categories against the overall figures. For burn injuries, the overall percentage was 29.8 percent, and Exhibit 5- I3 shows that the first two location categories are above this figure while the last two location categories are below it.

Column Percentages

To convert a table of counts into column percentages, we divide each entry) by the total for its column. The top left entry would he calculated as:

$$\frac{238}{437} = 54.5 \text{ percent}$$

This tells us that 54.5 percent of the persons **who received burns** were intimately involved in the fire's ignition.

Exhibit 5-14 Civilian Injuries-1989-Column Percentages

	Nature of Injury				
Location	Burns Only	Smoke Only	Burns & Smoke	Other	Total
Fire Casualty Intimately Involved with Ignition	54.5	6.4	28.6	6.2	26.8
Fire Casualty in the Room or Space of Fire Origin	29.7	18.9	24.7	10.3	23.2
Fire Casualty on Same Floor as Fire Origin	8.2	38.8	27.2	40.2	26.6
Fire Casualty in Same Building as Fire Origin	7.6	35.9	19.5	43.3	23.4
Overall	100.0	100.0	100.0	100.0	100.0

With a table of column percentages, we analyze a particular type of injury across the four locations. With burn injuries, we see that 54.5 percent occurred when the person was intimately involved with the fire's ignition. A total of 29.7 percent occurred when the person was in the room or space of the fire' origin, and less than 10 percent occurred in the other two location categories.

With the "Other" injury category, the picture changes. A total of 43.3 percent occurred when the person was in the same building as the fire's origin, followed closely by 40.2 percent for location on the same floor as the fire's origin. The first two location categories account for 10.3 percent and 6.2 percent, respectively.

Selecting a Percentage Table

The choice of a percentage table depends on what you are trying to conclude from the data. Joint probability tables are beneficial when you want to emphasize the interrelationship between the two variables in the table. Exhibit 5-12 shows that the combination of burns and intimate involvement in the fire's ignition account for 16.2 percent of the total. We can compare this figure to other combinations in the table.

The row percentage table provides a way of emphasizing the type of injury for each location. When the person was in the Same room or space of the fire's origin, Exhibit 5-13 shows 38.2 percent of the injuries were burns, 29.4 percent were smoke, 29.4 percent were burns and smoke, and 2.9 percent were other injuries. These are useful results by themselves, and can be compared to distributions in other rows.

The column percentage table emphasizes the location for each type of injury. For burns only, Exhibit 5-14 shows that 54.5 percent were intimately involved in the fire's ignition, 29.7 percent were in the same room or space of the fire's origin, 8.2 percent on the same floor, and 7.6 percent in the same building.

Testing for Independence in Two-way Tables

In this section we will develop a &i-squared test for testing whether the two variables in a two-way table are independent of each other. As with our prior discussion, we will provide a step-by-step procedure for calculating a chi-squared value in a two-way table. We would like to note at this point, however, that virtually all statstical packages automatically calculate the chi-squared value for you. As you may have concluded with the examples from Seattle and Florida, manual calculation is arduous and time consuming. In practice, it is not advisable to figure out chi-squared values with pencil and paper. However, we go through an example in detail in this section so that you understand what a statistical package is doing when it calculates a chi-squared value. You should appreciate the time saved and the inherent accuracy of these packages in your applications.

Before getting to chi-squared calculations, however, we need to know what we mean by independence. We say that two variables are **independent** if knowledge about one variable does not help us in predicting the outcome of the other variable. In the table of location versus type of injury, we should certainly suspect that the two variables are not independent. If we know, for example, that the person was intimately involved in the tire's ignition, then we can predict that the person probably had burn injuries. As we shall see later, the chi-squared test will confirm the dependence between our two variables.

Independence of two variables means that knowledge of one variable does not help in predicting the other variable. An equivalent definition is that two variables are independent if either their row or column percentages are equal.

As an example of two independent variables, consider the following table showing the relationship between sex and severity of injury in fires. The table shows a total of 1,841 persons who were either injured or killed as a result of a fire. Of this total, there were 1.092 males and 749 females. Of the 1,841 persons, there were 1,561 persons injured and 280 persons killed. This table includes the row percentages, which have rounded for purposes of illustration.

Exhibit 5-15 Sex and Severity of Injuries with Row Percentages-1989

Sex	Injury Severity Injured	 Killed	Total
Male	927	165%	I,092
Row%	85%	15%	100%
Female	634	115	749
Row%	85%	15%	100%
Total	1,561	280	1,841
Row%	85%	15%	100%

This table shows identical row percentages for males and females. That is, 85 percent of the males and 85 percent of the females received injuries. Consequently, knowledge of the sex of a person does not improve our ability to predict injury severity. With either sex, the row percentages have the Same distribution of 85 percent for injured and I5 percent for killed.

This table illustrates an equivalent definition for independence: two variables are independent if either their row percentages or column percentages are the same. When the percentages agree, we have no predictive power.

We cannot always expect that the row or column percentages in a table will be so close that independence is as obvious as Exhibit 5-15. The chi-squared test with two-way tables provides a means to test whether two variables are independent. With a chi-squared test, we can determine in a statistical manner whether the variables are independent.

In order to perform the chi-squared test, we first need to develop **expected values** for the table. The expected values are the counts that would occur if the two variables were independent. After forming a table of expected values, we will be in a position to do the chi-squared test.

Table of Expected Values

We form an expected value table in the following manner. For each number in the original table of counts, we identify the associated row total and column total. The entry in the expected value table is formed by multiplying the row sum by the column sum and dividing by the grand total. We can state this calculation as follows:

$$\text{Expected Value} = \frac{\text{Row Sum x Column Sum}}{\text{Grand Total}}$$

As an example, look back at Exhibit 5-11, which is the table of counts for the location and nature of injury. From our prior analysis with row and column percentages, we strongly suspect an association between location and nature of injury. The chi-squared test allows us to confirm our suspicion statistically.

The top left entry shows 238 persons who were intimately involved in the fire's ignition and had burn injuries. The row sum associated with this value is 394, and the column sum is 437. The expected value is therefore calculated as:

$$\text{Expected Value} = \frac{394 \text{ x } 437}{1468} = 117.29$$

We continue to perform this type of calculate for every entry in the original table of counts. Exhibit 5-16 shows the resulting expected value table.

Exhibit 5-16 Civilian Injuries-1989-Table of expected Values

	Nature of Injury				
Location	Burns Only	Smoke Only	Burns & Smoke	Other	Total
Fire Casualty Intimately Involved with Ignition	117.30	142.00	108.70	26.00	394.0
Fire Casualty in the Room or Space of Fire Origin	101.21	122.52	93.80	22.47	340.0
Fire Casualty on Same Floor as Fire Origin	116.10	140.54	107.60	25.77	390.0
Fire Casualty in Same Building as Fire Origin	102.40	123.96	94.90	22.73	344.0
Overall	437.00	529.00	405.00	97.00	1,468.0

The table of expected values is what we would expect if location and nature of injury were completely independent. It should also be noted that the row totals and column totals arc exactly, the same as the original table of counts. That is, development of the expected value table preserves these totals.

With the table of expected values in place, we can proceed with describing the calculations for the chi-squared test.

Chi-squared Test for Two-way Tables

The chi-squared value is calculated along the same lines as we did for testing just a single categorical variable.

1. Develop the table of expected values, as shown in Exhibit 5-16. Each entry in the table is obtained by multiplying the row total times the column total and then dividing by the grand total.
2. For each table entry, subtract the cxpected value from the corresponding entry in the original table of counts, and then square the result. This difference measures the discrepancy between the actual counts and what we would expect under independence.
3. Divide the results from Step 2 by the expected value. This is an adjustment that allows for the fact that larger expected numbers are usually associated with larger deviations.
4. Sum the results from Step 3. This is the chi-sqnarcd statistic. The larger the chi-squared statistic, the more likely there is a independence between the two variables. However, the &-squared statistic also depends on the number of categories, which must he taken into account in the following steps.
5. Find the **degrees of freedom,** which is calculated for two-way tables by multiplying (number of rows minus one) times (number of columns minus one). For our example, we have four rows and four columns. The number of degrees of freedom is therefore (‘+-I) x (4-1) = 9.
6. Compare the computed A-squared statistic from Step 4 to the value in the chi-squared table in Appendix A using the appropriate degrees of freedom. This table value is called the **critical chi-squared value.**

   If the computed chi-squared statistic is **greater** than the value in the table, then we **reject** the null hypothesis. Otherwise, we cannot reject the null hypothesis.

It is important to keep in mind that the null hypothesis with a two-way table is that the two variables are independent. If we accept the null hypothesis, we are saying that knowing the value of one of the variables does not help in predicting the value of the other variable. In our example, the null hypothesis is that location is independent of the nature of the injury.

Exhibit 5-17 shows the chi-squared entries for our two-way table. These entries are the results after Step 3 above. The top left entry was calculated as follows: Exhibit 5-11 gave an actual count of 238 for this entry and Exhibit 5-16 gave an expected value of 117.30. Subtracting the expected value from the actual count gives 120.7 (238 minus 117.3) and squaring gives 14,568.49. Dividing this number by the expected value, 117.3 provides the entry for the chi-squared table of 124.20.

Exhibit 5-17 Civilian Injuries-1989-Table of Chi-squared Entries

	Nature of Injury			
Location	Burns Only	Smoke Only	Burns & Smoke	Other
Fire Casualty intimately Involved with Ignition	124.20	82.14	0.49	15.39
Fire Casualty in the Room or Space of Fire Origin	8.20	4.13	0.41	6.94
Fire Casualty on Same Floor as Fire Origin	55.26	29.61	0.05	6.75
Fire Casualty in Same Building as Fire Origin	47.04	35.13	2.66	16.41

Total Chi-squared Value = 434.81

To determine the total chi-squared value, we add the numbers from Exhibit 5-17, which gives 434.81.

We are now ready to test our hypothesis about independence of these two variables - location and nature of injury. From Appendix A, we see that the critical chi-squared value for 9 degrees of freedom is 16.92. Our calculated chi-squared value of 434.81 greatly exceeds the critical value.

We therefore reject the null hypothesis and conclude that there is a statistical association between location and nature of injury.

Chi-squared Calculation for 2 x 2 Tables*

One exception to the above steps occurs with 2 x 2 tables. For 2 x 2 tables, we must subtract 0.5 from positive differences between observed and expected counts and add 0.5 from negative differences. This exception is necessary in order for 2 x 2 tables to match the chi-squared distribution under the null hypothesis. In most tables, this adjustment makes little difference in the calculated chi-squared value. However, in some tables it is important and can change the outcome of the test. For this reason, the adjustment should always be made in 2 x 2 tables.

As an illustration of this exception, consider the following table showing sex by part of body injured for the 1989 data.

Exhibit 5-18 Sex and Part of Body Injured with Row Percentaged-1989

	Part of Body Injured		
Sex	Internal	External	Total
Male	366	726	1,092
Row %	33.5	66.5	100.0
Female	338	411	749
Row %	45.1	54.9	100.0
Total	1,561	280	1,841
Row %	38.2	61.8	100.0

In this table, we have collapsed the injuries into internal (including respiratory and heart) and external (head, body, arm, leg, etc.) injuries. The row percentages indicate differences by sex. For males, 33.5 percent had internal injuries and 66.5 percent had external injuries while for females, the percentages are 45. 1 percent and 54.9 percent, respectively.

The chi-squared value from this table is 24.87. which has been adjusted according to the above rule on adding and subtracting 0.5. With a 2 x 2 table, there is always only one degree of freedom. From Appendix A, the critical value with 1 degree of freedom is 3.84. Since our calculated chi-squared value of 24.87 is greater than this critical value, we reject the null hypothesis.

We conclude that there is an association between sex and part of body injured.

Chapter 5
PROBLEMS

1. Suppose that we roll a die 60 times with the following results:

Dots Visible	Number	Percent
One	13	21.7
Two	I2	20.0
Three	7	11.7
Four	12	20.0
Five	9	15.0
Six	7	11.7
Total	60	100.0

Use a chi-squared test to determine whether this die is a “fair” die at the 5 percent level.

2. In New Orleans, Louisiana, the number of fires per month for 1990 was as follows:

January	467
February	291
March	392
April	322
May	319
June	349
July	384
August	374
September	359
October	368
November	298
December	345
Total	4,268

a. With a chi-squared test at the 5 percent level, determine whether there is an even distribution of fires in New Orleans.

b. Which months contribute significantly to deviations from an even distribution?

3. The data on the following pages are for 1990 fires for NFIRS Metropolitan areas and for the remainder of the United States. The NFIRS Metropolitan areas are a group of 25 of the largest cities contributing to NFIRS. The question is whether the distribution of types of fires in the metropolitan areas differs from the rest of the country.

Code	Type of Fire	Metro Areas	Rest of United States
11	Structure Fires	54,189	241,481
12	Outside of Structure Fires	4,378	23,986
13	Vehicle Fires	51,774	174,398
14	Trees, Brush and Grass Fires	28,440	172,217
15	Refuse Fires	59,767	111,291
	Other Fires	2,833	15,353
	Total	201,381	738,726

Source: NFIRS Tally Report 22, 1990.

a. Calculate percentages for each group and develop a tentative answer based only on the percentages.

b Use the chi-squared test to determine whether the two distributions are significantly different at the 5 percent level.

c. Which types of tires account for the main differences between the two groups?

4. The following data for days of the week are for 1990 fires for Denver, Colorado and for the entire United States. The percentages are for the United States. Using a chi-square test, determine whether the distribution by day of week for Denver differs from the entire country.

Day	Denver Fires	Fires in United States	Percent
Sunday	674	134,077	14.37
Monday	688	136,984	14.68
Tuesday	565	130,283	13.96
Wednesday	550	134,886	14.45
Thursday	576	128,729	13.79
Friday	588	129,522	13.88
Saturday	603	138,854	14.88
Total	4,244	933,335	100.0

5. The following data for days of the week are for 1990 fires for Denver, Colorado and for the other metropolitan cities. Using a chi-square test, determine whether the distribution by day of week for Denver differs from the other metropolitan cities.

Day	Denver Fires	Fires in Metro Cities	Percent
Sunday	674	29,382	15.18
Monday	688	28,961	14.96
Tuesday	565	27,200	14.05
Wednesday	550	27,305	14.10
Thursday	576	26,025	13.44
Friday	588	26,226	13.55
Saturday	603	28,505	14.72
Total	4,244	193,604	100.0

Chapter 6
ADVANCED TABLE ANALYSIS

Introduction

In Chapter 5, we focused on statistical tools for analyzing the relationship between two variables. We now want to extend our ideas to situations where we need to understand the relationships among several categorical variables. For example, consider the problem of trying to determine factors that affect the spread of a fire beyond the initial room of origin. We might ask several key questions: If a building has fire detectors, is a fire less likely to spread to other rooms? To what extent does the type of material ignited affect the spread of a fire? If equipment is involved, is there greater likelihood or less likelihood of the fire spreading to other parts of the building?

Each of these variables-fire spread, presence of a fire detector, type of material, involvement of equipment-is a categorical variable. That is, each variable is divided into several categories as defined by the NFPA 901 codes. The box on fire incident reports labeled “Extent of Flame Damage” indicates fire spread according to seven categories defined by the NFPA 901 codes.” Similarly, fire detector performance, type of material, and equipment involvement are also recorded on the fire incident report with NFPA 901 codes.

Loglinear analysis is a statistical approach for analyzing the relationships among several categorical variables. The complexities addressed by loglinear analysis are easily illustrated. We could, for example, develop several two-way tables from the four variables just described:

- Fire spread by fire detector use
- Fire spread by type of material ignited
- Fire spread by involvement of equipment
- Fire detector use by type of material ignited
- Fire detector use by involvement of equipment
- Type of material ignited by involvement of equipment

Each two-way table defines a relationship that may be important in understanding fire spread. We could proceed even further by defining three-way tables, such as fire spread by fire detector use by type of material ignited. This three-way relationship may also be important in explaining the spread of a fire. The point is that we have many potential relationships with these four variables. Loglinear analysis assists in identifying the important relationships.

---

6. The categories are (1) fire confined to object of origin (2) fire confined to part of room, (3) fire confined to room, (4) fire confined to fire-rated compartment of origin, (5) fire confined to floor of origin. (6) fire confined to structure of origln, and (7) fire extended beyond structure of origin.

In this chapter we will present an overview of loglinear analysis. We start by applying loglinear analysis to two-dimensional tables. This discussion serves as an introduction to the ideas behind this approach. Our primary example in this chapter is a table with four variable: extent of flame damage, fire detector performance, type of material ignited, and whether equipment was involved in the fire.

It should he mentioned that this chapter differs from previous chapters because it assumes some knowledge of statistics. You should be able to understand this chapter if you have had an introductory) course in statistics. In particular, an understanding of the standardized normal distribution is required. The calculations for loglinear analysis are also more difficult than the relatively simple calculations in previous chapters. However, all the statistical packages mentioned in Chapter 1 include loglinear analysis procedures which will automatically perform the calculations. Your job will be to interpret the results and select the most appropriate model.

Loglinear analysis is a statistical technique for determining relationships among variables in multi-dimensional tables. From assumptions about interactions among table variables, loglinear analysis develops a model of the table and then tests expected results from the model against the actual table. You can develop several models with different assumptions and then select the model that most appropriately describes the relationships in the table.

Loglinear Analysis of 2 x 2 Tables*

Model of Independence

Exhibit 6-1 is a hypothetical table for two variables that arc mutually, independent. You can verify that the row percentages arc the same for both categories of variable A, which proves its independence with variable B. We have introduced notation in Exhibit 6-1 for identifying the individual cells and the sums along each row and column. For example, x_{11} denotes the first category of variable A and the first category of variable B. Similarly), $^{x}12$ indicates the first category of variable A and the second category of variable B.

The expected value for a cell in the table can be expressed in the following manner:

$$m_{ij} = \frac{(x_{i+})\,(x_{+j})}{x_{++}} \qquad i = 1,2 \quad j = 1,2 \qquad (1)$$

where m_{ij} is the expected value for the cell at row i and column j. Equation (1) says that the expected value, m_{ij} is found by multiplying the total for row

$i(x_{i+})$ by the total for column *$j(x_{+j})$* and then dividing by the grand total *(x_{++})*. This calculation is exactly the same as presented in Chapter 5 for determining expected values.

Exhibit 6-1 Hypothetical Tables for Two Independent Variables

	Variable B		
Variable A	Category B1	Category B2	Total
Category A_1	36 x_{11}	24 x_{12}	60 x_{1+}
Category A2	48 x_{21}	32 x_{22}	80 x_{2+}
Total	84 x_{+1}	56 x_{+2}	140 x_{++}

If we take the **natural** logarithm of both sides of (l), we obtain the following:

$$\log m_{ij} = \log x_{i+} + \log x_{+j} - \log x_{++} \tag{2}$$

To generalize our discussion, we will think in terms of a two-dimensional table with I rows and J columns. Equations (1) and (2) still hold as expressions of expected values and their logarithms.

The formal model of independence is expressed by rewriting (2) in the following manner.

$$\log m_{ij} = u + u_{A(i)} + u_{B(j)} \tag{3}$$

$$u = \frac{1}{\mathrm{IJ}} \sum_i \sum_j \log m_{ij} \tag{4}$$

$$u_{A(i)} + u = \frac{1}{\mathrm{J}} \sum_j \log m_{ij} \qquad i = 1,2,\ldots\mathrm{I} \tag{5}$$

$$u_{B(j)} + u = \frac{1}{\mathrm{I}} \sum_i \log m_{ij} \qquad j = 1,2,\ldots\mathrm{J} \tag{6}$$

The term u is the overall mean of the logarithms of the numbers in the table. Similarly, $u_{A(i)}$ + u gives the mean of the logarithms of the expected counts for the J cells at level i of variable A and uB(j) + u is the mean of the logarithms of the expected counts for the I cells at level j of variable B.

Equations (3) through (6) represent the **loglinear model** of independence for two-dimensional tables. The name loglinear model derives from the fact that we have taken the logarithms of the expected values to form a linear combination. The model is valid for 2 x 2 tables as well as larger tables with I rows and J columns.

We can apply these equations to Exhibit 6-I to get a sense of how the loglinear model of independence operates. The results are shown in Exhibit 6-2. To obtain the value for U, apply Equation 4 to the cell entries:

$$u = \tfrac{1}{4} (\log(36) + \log(24) + \log(48) + \log(32)) = 3{,}525 \qquad (7)$$

Similarly, we obtain $u_{A(1)}$ and $u_{B(1)}$ by applying Equations (5) and (6):

$$u_{A(1)} = \frac{\log(36) + \log(24)}{2} \; 3.525 = -.144 \qquad (8)$$

$$u_{B(1)} = \frac{\log(36) + \log(48)}{2} - 3.525 = .203$$

Exhibit 6-2 Values of Terms for Model of Independence

Term	Term Value
u	3.525
$u_{A(1)}$	-.144
$u_{A(2)}$	.I44
$u_{B(1)}$	.203
$u_{B(2)}$	-.203

An important feature of the model is that the sum of uA(1), and uA(2) is equal to zero. Similarly, the sum of uB(1) and uB(2) is also equal to zero. The reason they sum to zero is that they are deviations from an overall average. We have had other discussions in this handbook where the sum of deviations from a mean is equal to zero. In general, we can state that:

$$\sum_i u_{A(i)} = \sum_j u_{B(j)} = 0 \qquad (9)$$

The values in Exhibit 6-2 can be employed with Equation (3) to obtain Exhibit 6-1. The logarithm of m21, the number of injured females, is shown on the following page.

$$\log m_{21} = u + u_{A(2)} + u_{B(1)} \tag{10}$$

$$\log m_{21} = 3.525 + .144 + .203 = 3.872$$

The antilogarithm gives m_{21} = 48, which agrees with Exhibit 6-1.

Model of Dependence

As we saw in Chapter 5, we frequently encounter tables in which the two variables are not independent. As an example, we analyzed location and nature of injury in Chapter 5 quite extensively. Exhibit 6-3 reproduces the table so that wc can illustrate the extension of our model to two dependent variables.

Exhibit 6-3 Civilian Injuries-1989

	Nature of Injury				
Location	Burns Only	Smoke Only	Burns & Smoke	Other	Row Totals
Fire Casualty Intimately Involved with ignition	238	34	116	6	394
Fire Casualty in the Room or Space of Fire Origin	130	100	100	10	340
Fire Casualty on Same Floor as Fire Origin	36	205	110	39	390
Fire Casualty in Same Building as Fire Origin	33	190	79	42	344
Column Totals	437	529	405	97	1,468

The extension is accomplished by introducing **interaction terms. We** think of location and nature of injury as "interacting" and develop an extended model to take the interactions into account:

$$\log m_{ij} = u + u_{A(i)} + u_{B(j)} + u_{AB(ij)} \tag{11}$$

where $u_{AB(ij)}$ represents the interactions between variables A and B. To calculate $u_{AB(ij)}$, we simply use equation (11) and solve for uAB(ij). Equations (4) thru (6) and Equation (9) still apply, and in addition, we: have:

$$\sum_i u_{AB(ij)} = \sum_j u_{AB(ij)} = 0 \tag{12}$$

Exhibit 6-4 shows the results for location and nature of injury. We have an overall mean, u, of 4.128 for the model. Since location has four categories, we have four $u_{A(i)}$ terms. Similarly, we have four $u_{B(j)}$ terms for the four categories of injury. There are 16 terms for $u_{AB(ij)}$ to cover the 16 combinations of location and nature of injury. The terms $uA(i)$ and $u_{B(j)}$ are called the **main effects** of the model and the terms $u_{AB(ij)}$ are called the **two-way interaction effects.**

These terms operate in exactly the same way as with the independent model. That is, we can obtain an entry in Exhibit 6-3 by applying Equation 11. Exhibit 6-3 shows 205 persons with injuries of smoke only and location on the same floor as the fire's origin (x_{32}). To obtain this value from our model, we calculate as follows:

$$\begin{aligned} \log m_{32} &= u + u_{A(3)} + u_{B(2)} + u_{AB(32)} \\ &= 4.128 + -190 + .548 + .458 \qquad (13) \\ &= 5.324 \end{aligned}$$

The antilogarithm gives the value of 205 persons.

Exhibit 6-4 Values of Terms for Dependent Model

Location and Nature of Injury			
Term	Term Value	Term	Term Value
u	4.128		
$u_{A(1)}$	-.242	$u_{B(1)}$	.227
$u_{A(2)}$	-.033	$u_{B(2)}$	.548
$u_{A(3)}$	.190	$u_{B(3)}$	.479
$U_{A(4)}$	.085	$u_{B(4)}$	-1.254
$u_{AB(11)}$	1.359	$u_{AB(31)}$	-.961
$u_{AB(12)}$	- .907	$u_{AB(32)}$	.458
$u_{AB(13)}$	.388	$u_{AB(33)}$	-.096
$u_{AB(14)}$	- .840	$u_{AB(34)}$	.600
$u_{AB(21)}$	545	$u_{AB(41)}$	-.943
$U_{AB(22)}$	- .038	$u_{AB(42)}$	.487
$u_{AB(23)}$	.031	$u_{AB(43)}$	-.323
$U_{AB(24)}$	- .539	$u_{AB(44)}$	.779

With the model given by Equation (11) for a two-dimensional table, we will always obtain the exact value in the original table. For this reason, the loglinear model with all interaction terms is called a **saturated** model. Later in this chapter, we will develop **unsaturated models** by eliminating terms from the saturated model. The aim with unsaturated models is to obtain good estimates for a table with fewer terms than a saturated model.

While unsaturated models rarely produce the exact values in a table, they sometimes provide estimates close to the table values. The unsaturated model then has a key advantage of having identified the important interactions among the variables.

> For a two-dimensional table, the saturated loglinear model is expressed as log $m_{ij}=u + u_{A(i)}+u_{B(j)}+u_{AB(ij)}$, where u is the overall mean of the logarithms of table entries, $u_{A(i)}$ are the terms for the main effects of variable A, $u_{B(j)}$ are the terms for the main effects of variable B, and $u_{AB(ij)}$ are the interactions terms. If the two variables are independent, then the terms $u_{AB(ij)}$ equal zero and do not appear in the model.

As a final example in this section, Exhibit 6-5 shows a 2 x 2 table on civilian injuries relating sex of the injured person to external and internal injuries from fires. External injuries are usually burns to a part of the body while internal injuries are usually from smoke inhalation. Because the row percentages differ considerably, the two variables are clearly not independent. A saturated model is therefore appropriate for this table.

Exhibit 6-5 Civilian Injuries-Sex by Type of Injury-1989

	External	Internal	Total
Male	362	331	693
Row Percent	52.2	57.8	100.0
Female	211	336	547
Row Percent	38.6	61.4	100.0
Total	573	667	1,240
	46.2	53.8	100.0

u =5.716

$u_{A(1)}$= .131
$u_{A(2)}$ = -.131
$u_{B(1)}$ = - .094
$u_{B(2)}$= .094

$u_{AB(11)}$ = .138
$u_{AB(12)}$ = - .138
$u_{AB(21)}$ = - .138
$u_{AB(22)}$= .138

The right portion of the table shows the terms for the saturated model. Each entry in the table can be derived using Equation (11) with these terms.

One other feature of 2 x 2 tables is important for the remainder of this chapter. Referring back to Exhibit 6-1, suppose that we randomly select an individual from category B_1. The quantity x_{11}/x_{21} represents the **odds** of an individual appearing in category A_1 rather than A_2. Similarly, x_{12}/x_{22} gives the odds of an individual appeari in A_1 rather than A_2 **given** that the individual was from category B_2. In Exhibit 6-5, we have 573 persons with external injuries and the odds are 1.72 of a male rather than a female (362 divided by 211) having external injuries. We usually express the odds as 1.72: 1 of males to females. If internal injuries are involved, the odds are .98:1 (331 divided by 336) of males to females. Odds greater than 1.0 indicate that the first category is larger than the second category, while odds less than 1 .0 indicate the opposite.

The **odds ratio** is the ratio of these two odds. and can be written as:

$$\text{Odds Ratio} = \frac{x_{11}x_{22}}{x_{12}x_{21}} \qquad (14)$$

If the two variables are independent, then the odds ratio will always be equal to 1.0. Conversely, an odds ratio of exactly 1.0 indicates complete independence of two variables. On the other hand, the odds ratio from Exhibit 6-5 is 1.74, which indicates that sex and type of injury are not independent.

Another way to derive the odds ratio is from probabilities. For example, Exhibit 6-6 shows probabilities calculated by dividing each entry in Exhibit 6-5 by the total of 1,240 persons. The odds ratio is obtained by $p_{11}p_{22}/p_{21}p_{12}$. For Exhibit 6-6, the calculation gives 1.74 (.292 x .271 / 170 x .267), which agrees with our previous calculation. The reason for showing this approach is that we may be given the probabilities, rather than the actual data, as our starting table. The calculations show that the results are the same regardless of the starting point.

Exhibit 6-6 Civilian Injuries-Sex by Type of Injury-Probabilities-1989

	External	Internal	Total
Male	.292 p_{11}	.267 p_{I2}	.559
Female	.170 p_{21}	.271 p_{22}	.441
Total	.462	.538	1.000

We will return to the idea of odds ratio later in this chapter when we present results from an example with four variables.

Three-way Tables and Standardized Values*

Suppose we have three variables labeled A, B, and C with I, J, and K categories, respectively. A three-way table with these categories would have / x J x K cells. Each ceil would indicate the number of items with attributes Ai, Bj, and Ck. We consider the items to be a random sample taken from a population from which the talk is derived.

We let wijk = loge xijk, where Xijk is the number of observations in ceil (i,j,k). The saturated model for a three-way table is given by:

$$w_{ijk} = u + u_{A(i)} + u_{B(j)} + u_{C(k)} + u_{AB(ij)} + u_{AC(ik)} + u_{BC(jk)} + u_{ABC(ijk)} \qquad (15)$$

As with the previous discussion, u is the overall mean, and uA(i), uB(i), and uc(k) are the main effects. Similarly, uAB(ij), uAC(ik), and UBC(jk) represent the two-way interactions, and uABC(ijk) are the three-way interactions. For a 2 x 2 x 2 table, we would have an overall mean, 4 terms for each of the two-way interactions, and 8 terms for the three-way interactions.

The Sums of various terms in the model are zero:

$$\sum_i u_{A(i)} = \sum_j u_{B(j)} = \sum_k u_{C(k)} = \sum_i u_{AB(ij)} = \sum_j u_{AB(ij)} = \sum_i u_{AC(ik)} = \sum_k u_{AC(ik)} \qquad (16)$$

$$= \sum_j u_{BC(jk)} = \sum_k u_{BC(jk)} = \sum_i u_{ABC(ijk)} = \sum_j u_{ABC(ijk)} = \sum_k u_{ABC(ijk)}$$

$$= 0$$

While Equations (I5) and (16) appear complex, the calculations are relatively straightforward. For example, u is the average of the natural logarithms of the table entries:

$$u = \frac{1}{IJK} \sum_i \sum_j \sum_k w_{ijk} \qquad (17)$$

To obtain the other values in the saturated model, we need the following definitions:

$$w_{i++} = \frac{1}{JK} \sum_j \sum_k w_{ijk} \qquad (18)$$

$$w_{+j+} = \frac{1}{IK} \sum_i \sum_k w_{ijk} \qquad (19)$$

$$w_{++k} = \frac{1}{IJ} \sum_{i}\sum_{j} w_{ijk} \qquad (20)$$

Then we obtain:

$$u_{A(i)} = w_{i++} - u \qquad i = 1,2 \ldots, I \qquad (21)$$

$$u_{B(j)} = w_{+j+} - u \qquad j = 1,2 \ldots, J \qquad (22)$$

$$u_{C(k)} = w_{++k} - u \qquad k = 1,2 \ldots, K \qquad (23)$$

By extending the model, we develop terms for $u_{AB(ij)}$ and $u_{ABC(ijk)}$ as follows:

$$u_{AB(ij)} = w_{ij+} - w_{i++} - w_{+j+} + u \qquad (24)$$

$$u_{ABC(ijk)} = w_{ijk} - w_{ij+} - w_{i+k} - w_{+jk} + w_{i++} + w_{+j+} + w_{++k} - u \qquad (25)$$

As an example of this saturated model, the value for $u_{AB(11)}$ for a 2 x 2 x 2 table is given by:

$$\begin{aligned} u_{AB(11)} &= w_{11+} - w_{1++} - w_{+1+} + w_{+++} \qquad (26) \\ &= {}^1/_2 (w_{111} + w_{112}) - {}^1/_4 (w_{111} + w_{112} + w_{121} + w_{122}) \\ &\quad - {}^1/_4 (w_{111} + w_{112} + w_{211} + w_{212}) \\ &\quad - {}^1/_8 (w_{111} + w_{112} + w_{121} + w_{122} + w_{211} + w_{212} + w_{221} + w_{222}) \\ &= {}^1/_8 (w_{111}) + {}^1/_8 (w_{112}) - {}^1/_8 (w_{121}) - {}^1/_8 (w_{122}) - {}^1/_8 (w_{211}) \\ &\quad - {}^1/_8 (w_{212}) + {}^1/_8 (w_{221}) + {}^1/_8 (w_{222}) \end{aligned}$$

It should be noted that each number in the last equation is multiplied by either by or ${}^1/_8$ or ${}^1/_8$.

Upton (1978) shows that any parameter in a saturated model can be expressed in the following form:

$$\sum_{i}\sum_{j}\sum_{k} a_{ijk} w_{ijk} \qquad (27)$$

where the aijk are suitably developed constants. In the example just given, the a_{ijk} are either ${}^1/_8$ or $-{}^1/_8$ depending on the particular cell.

Upton also discusses the fact that the variance of any is related to the original frequencies in the table, and can be approximated by:

$$\mathrm{Var}(w_{ijk}) = \frac{1}{x_{ijk}} \qquad (28)$$

The key point is that we can approximate the variance of a parameter with the following relationship:

$$\text{Variance}=\sum_i\sum_j\sum_k \frac{(a_{ijk})^2}{x_{ijk}} \qquad (29)$$

In many applications, our aim will be to identify the most important terms in a model. The estimated variances will not always be the same since they depend on the number of categories. To compare the terms, we therefore need to standardize, which can be done as follows:

$$\text{Standardized Term}=\frac{u^*}{\sqrt{\mathrm{V}(u^*)}} \qquad \mathbf{(30)}$$

where u^* represents any of the terms in the model and V(u*) is the variance of the term according to Equation (27). Goodman (1971) shows that these standardized values follow an approximately normal distribution with an expected mean of 0 and a variance of 1.

The value of the above result is that it allows us to determine the importance of terms in a model. In the next section, we will show the standardized terms for a saturated model in a four-dimensional table and use the standardized terms to identify key interactions among the four variables.

Loglinear Analysis Applied to a Four-Dimensional Table: Fire Data in Chicago, Illinois

Saturated Model

With the background from the previous sections, we are in a position to analyze higher-dimensional tables. Our aim is to determine how several categorical variables relate to each other. That is, we want to examine the interactions among variables as a way of explaining the entire table.

As an example for this section, we use data from 3,548 residential fires that occurred in Chicago, Illinois in 1990.[7] We have selected four variables to study with the following definitions:

Variable	Description
A	Fire was confined to the room of origin, or was not
B	Detector performed, or did not (or was not present)
C	Fire started from fabric material, or did not
D	Equipment was involved in the fire, or was not

---

[7] Residential fires classified as arsons are not included in this analysis.

The first variable (A) indicates whether the fire was confined to the room of origin or whether it extended to another room, another floor in the structure, or perhaps even another structure. Variable A is typically catted the extent of flame damage. The second variable (B) indicates whether a fire detector went off during the fire. The third variable (c) indicates whether the form of material ignited was a fabric (such as cotton, wool, fur, etc.) or another type of material (such as flammable liquids, plastics, wood, or paper). The last variable (D) indicates whether equipment was involved in the fire.8

All four variables are dichotomous variables; that is, they are divided into two categories as either present or not present. We intentional]!. developed dichotomous variables to simptify. the presentation of the loglinear analysis. Extent of flame damage actually has seven categories indicating whether the fire was confined to (1) object of origin. (1) parts of room or area of origin, (3) room of origin, (4) tire-rated compartment of origin, (5) floor or origin, (6) structure of origin, or (7) beyond structure of origin. For our analysis, the first four categories defined A1 (room of origin) and the last three catcgories define A2 (beyond room of origin). It should be mentioned that you can apply loglinear analysis to a table with all seven categories; however, each cell in the table should contain at least five fires in order to assure valid results from the analysis.

Exhibit 6-7 on the following page gives the breakdown of these 3,548 residential fires according to our four-way classification. The first category of each variable is coded aS a one and the second category aS a two. With detector performance, for example, a code of one indicates that a fire detector went off, while a code of two indicates a fire detector did not operate. The exhibit defines a 2 x 2 x 2 x 2 table with a total of 16 cells. Cell (l,l,l,l) indicates 37 fires where the fire remained in the room of origin, the detector went off, fabric was the type of material ignited, and equipment was also involved in the fire. Cell (2,2,2,1) contain 125 fires where the fire extended beyond the room of origin, a detector was not present, fabric material was not involved in the tire, and equipment was involved in the fire.

exhibit 6-8 shows the term values and standardized terms for the saturated model. For each u term, a subscript appears in the “Term” column indicating the variables. The term value and standardized value are shown in the nest two columns. They always correspond to the first category of a variable. The first item in the table is for the first category of variable A confined to room of origin). It shows a term value of uA = 555 and a standardized value of 16.0. As shown in the previous section, term values sum to zero, and similarly, standardized values also sum to zero. We therefore know that the second category of variable A (fire extended beyond the room of origin) has a term value of -.555 and a standardized value of -16.0. The other figures in the exhibit operate in the same manner since all our variables are dichotomous.

---

8 A code of 98 in the Equipment Involved in Ignition indicates that no equipment was involved !n the fire. All four variables have been rededined from the NFPA 901 codes to form dichotomous variables

Exhibit 6-7 Residential Fires in Chicago, Illinois – 1990

						Definition		
Cell				Frequency	Extent of Flame Damage	Fire Detector	Material	Equipment
1	1	1	1	37	Room of Origin	Worked	Fabric	Equipment Involved
1	1	1	2	165	Room of Origin	Worked	Fabric	No Equipment Involved
1	1	2	1	218	Room of Origin	Worked	Other Material	Equipment Involved
1	1	2	2	181	Room of Origin	Worked	Other Material	No Equipment Involved
1	2	1	1	95	Room of Origin	Did Not Work	Fabric	Equipment Involved
1	2	1	2	504	Room of Origin	Did Not Work	Fabric	No Equipment Involved
1	2	2	1	487	Room of Origin	Did Not Work	Other Material	Equipment Involved
1	2	2	2	798	Room of Origin	Did Not Work	Other Material	No Equipment Involved
2	1	1	1	6	Beyond Room of Origin	Worked	Fabric	Equipment Involved
2	1	1	2	55	Beyond Room of Origin	Worked	Fabric	No Equipment Involved
2	1	2	1	35	Beyond Room of Origin	Worked	Other Material	Equipment involved
2	1	2	2	70	Beyond Room of Origin	Worked	Other Material	No Equipment Involved
2	2	1	1	46	Beyond Room of Origin	Did Not Work	Fabric	Equipment involved
2	2	1	2	262	Beyond Room of Origin	Did Not Work	Fabric	No Equipment Involved
2	2	2	1	125	Beyond Room of Origin	Did Not Work	Other Material	Equipment Involved
2	2	2	2	464	Beyond Room of Origin	Did Not Work	Other Material	No Equipment Involved

Exhibit 6-8 Standarized Values for Terms in Standard Model

Chicago Residential Fires-1990

Term	Term Value	Standardized Value	Term	Term Value	Standardized Value
A	555	16.0	BD	.067	1.9
B	-.688	-19.9	CD	-.295	-8.5
C	-.465	-13.4			
D	-.583	-16.8	ABC	.038	1.1
			ABD	.041	1.2
AB	.146	4.2	ACD	-.060	-1.7
AC	-.026	-0.6	BCD	-096	-2.6
AD	.151	4.4			
BC	-.009	-0.3	ABCD	.033	1.0

If we study the standardized values, we can identify the most important variables according to the saturated model. Since the standardized terms follow a normal distribution with mean zero and variance one, **a good rule of thumb is to look at standardized values with magnitudes greater than 2.0 (ignoring the sign).** Any standardized value with magnitude greater than 2.0 indicates a term particularly important to the model. The table shows that all four main effects (A, B, C, and D) are important. In addition, the following interaction effects, in order of absolute magnitude, arc important:

CD	Type of material and equipment involved
AD	Extent of flame damage and equipment involved
AB	Extent of flame damage and detector performance
BCD	Detector performance, type of material, and equipment involved

These results will be useful in the next section where we develop a log-linear model that contains only the interactions of importance.

Continuation of Chicago Example: Hierarchical Models*

By definition, saturated models include every possible interaction in a table. We can derive beneficial conclusions from a saturated model because it is an exact representation of the table. The question we explore in this section is whether we can eliminate terms from the saturated model and still obtain good estimates of the table values. With a reduced model, we will identify more clearly the important interactions in the table. Conversely, terms not included in a reduced model do not contribute to our understanding of the table.

Hierarchical models provide a structured approach to define models with fewer terms than saturated models. The rule with an hierarchical model is that any term included in the model automatically means that higher-level terms are also included. For example, if u_{AB} is included, then u_A and and u_B are automatically included. Of course, the overall mean, u, appears in all models. If u_{ABC} is in the model, then we include all two-way interactions and main effects: u_{AB}, u_{AC}, u_{BC}, u_A, u_B, and u_C. Hierarchical models are unsaturated models because they have fewer terms than saturated models.

With this approach, we are establishing a hierarchy of terms. We are saying that if an interaction term, such as u_{AB} is important, then the main effects, u_A and u_B, are also important. Exhibit 6-9 shows the 9 possible hierarchical models for a table with three variables. The first model is the saturated model which we have already presented. The model A/BC consists of the main effects for all three variables and the two-way interaction term for variables B and C. The model AB/AC consists of the main effects for all three variables and the two interaction effects AB and AC, but excludes BC and ABC.

Exhibit 6-9 Possible Hierarchical Models for Tables with Three Variables

Model	Model Terms	Degrees of Freedom for 2 x 2 x 2 Model
ABC	$u, u_A, u_B, u_C, u_{AB}, u_{AC}, u_{BC}, u_{ABC}$ (Saturated Model)	0
AB/AC/BC	$u, u_A, u_B, u_C, u_{AB}, u_{AC}, u_{BC}$	1
AB/AC	$u, u_A, u_B, u_C, u_{AB}, u_{AC}$	2
AB/BC	$u, u_A, u_B, u_C, u_{AB}, u_{BC}$	2
AC/BC	$u, u_A, u_B, u_C, u_{AC}, u_{BC}$	2
A/BC	u, u_A, u_B, u_C, u_{BC}	3
B/AC	u, u_A, u_B, u_C, u_{AC}	3
C/AB	u, u_A, u_B, u_C, u_{AB}	3
A/B/C	u, u_A, u_B, u_C	4

Note that this table does not include models for which one of the three variables is completely omitted. For example, The model A/B is also considered an hierarchical model with terms u, u_A, and u_B, but not terms involving variable C. Models which do not include one of the variables essentially mean that the omitted variable does not contribute to our understanding of the table, and we can collapse the table over this variable to produce, for example, a two-way table involving only variables A and B.

Also note that the degrees of freedom are the number of terms from the saturated model not included in an hierarchical model. For example, model A/BC omits u_{ac}, u_{bc}, and u_{abc} from the saturated rnodel and therefore has three degrees of freedom.

The computations for the term values for hierarchical models arc usually not straightforward. Instead, the term values are calculated in an iterative manner. For this reason, a statistical package is a requirement for most hierarchical models.

> Saturated models contain every possible interaction in a table. Hierarchical Models contain fewer terms but are developed in a structured manner. The rule with hierarchical models is that any term in the model automatically means that its higher-order terms are also included. If ABD is in a hierarchical model, then A, B, D, AB, AD, and BD must be in the model. The aim of hierarchical models is to develop a loglinear model that provides good estimates for the table with fewer terms than a saturated model.

Any of these models will provide estimates for the cells in our table. To determine how well a model fits the table, we develop a test statistic as follows:

$$Y^2 = 2 \sum O \log \frac{O}{E} \tag{31}$$

where O is the observed or actual count in a given table cell and E is the expected count under a particular hierarchical model. The Y' statistic approximately follows a chi-squared distribution.

For tables with three variable (I x J x K), the Y statistic can be written as:

$$Y^2 = 2 \sum_{i} \sum_{j} \sum_{k} x_{ijk} \frac{\log x_{ijk}}{e_{ijk}} \tag{32}$$

where x_{ijk} is the observed count in cell (i,j,k) and e_{ijk} is the expected count according to the model.

Because the Y^2 statistic follows a chi-quared distribution, we can use Appendix A to test whether a model is appropriate at the 5 percent level. If the calculated Y^2 statistic is less than the entry in Appendix A, then we accept the model as a good fit to the data; conversely, if it is greater, we conclude that our model does not provide a good fit.

We will apple hierarchical models to our example with four dichotomous variables. We will consider the extent of flame damage (fire confined to the room/not confined to the room) as a **response** variable and the other three variables **as explanatory variables.** That is, we want to model how detector performance (Variable B), type of material (Variable C), and involvement of equipment (Variable D) relate to the extent of flame damage (Variable A). Our aim is to select a model which fits the table in a reasonable manner and has fewer terms than the saturated model.

Since variables B, C, and D are explanatory variables, any hierarchical model must include the term BCD, which means any model will also include its higher-order terms (DC, BD, CD, B, C, and D). With these terms as common to all models, we can then concentrate on interactions with our response variable.

Exhibit 6- 10 shows the results from several models. Each contains the BCD term along with selected terms involving the response variable. The degrees of freedom for each model is found by counting the number of terms omitted from the saturated model. For example, the first model, AB/ACD/BCD contains all the terms from the saturated model except ABD, ABC, and ABCD. The model therefore has three degrees of freedom.

Only Models 1 and 7 provide a good fit to the data according to our test. Model 1 is defined as AB/ACD/BCD and Model 7 as AB/AD/BCD. We need to determine which of these two models we want to select as the final model. As a general rule, when choices are available, the model to select is the one with the fewest number of terms. Both models contain u_{BCD} and its higher-order effects. In addition, Model I includes $u_{AC,}$ U_{AC}, $U_{AD,}$ and and u_{AB} while Model 7 includes only u_{AB} and u_{AD}. Model 7 is the better choice since it fits the table reasonably well and has two fewer terms.

Exhibit 6-10 Hierarchical Models for Extent of Flame Damage

Chicago Residential Fires-1990

Model	Terms	Y^2	Degrees of Freedom
1	AB/ACD/BCD	2.09*	3
2	AD/ABC/BCD	11.39	3
3	AC/ABD/BCD	10.96	3
4	ACD/BCD	25.97	4
5	AB/AC/AD/BCD	11.56	4
6	AB/AC/BCD	71.62	5
7	AB/AD/BCD	11.57 *	5
8	AC/AD/BCD	36.91	5

Note: An asterisk indicates that the model is significant at the 5 percent level That is. the model produces good estimates of the actual table values.

We could have anticipated that Model 7 would be a good model because of the results in the previous section with the saturated model. In the saturated model, the primary interaction effects were AB, AD, CD, and BCD. With the hierarchical model, BCD is automatically included along with CD as a higher-order effect. The two remaining effects, AR and AD. complete Model 7.

Exhibit 6-11 shows the term values and standardized values for Model 7. All except one of the terms have standardized values with magnitudes greater than 2.0. In particular, all the terms involving variable A in the model are important.

Exhibit 6-11 Extent of Flame Damage Chicago Residential Fires, 1990

Model AB/AD/BCD

Term	Term Value	Standardized Value
Mean	4.807	N/A
A	.577	21.7
B	-.678	-22.4
C	-.484	-17.9
D	-.591	-19.8
AB	.120	4.9
AD	.169	7.7
BC	.008	.3
BD	.084	3.1
CD	-.326	-12.1
BCD	-.084	-3.1

From these term values, we can obtain expected values as shown in Exhibit 6-12 along with the raw data. Most of the expected values are close to the actual data in the table. Ten expected values are within five percent of their actual values and 14 are within ten percent of their actual values. The model therefore provides reasonably accurate estimates for the cells in the table.

In summary, the hierarchical model (AB/AD/BCD) described in Exhibit 6-11 provides an excellent model for the residential fires in Chicago. The model indicates the following key results:

- Containment of a fire to the room of origin is more likely if fire detectors are present and operated.
- Containment of a fire to the room of origin is more likely if equipment is involved in the fire.
- Containment is not related to whether fabric is the type of material ignited.

Exhibit 6-12 Comparison of Expected Values and Actual Table Values

Chicago Residential Fires-1990

A	B	C	D	Actual Table Value	Model Value
1	1	1	1	37	36.5
1	1	1	2	165	163.3
1	1	2	1	218	215.0
1	1	2	2	181	186.3
1	2	1	1	95	109.6
1	2	1	2	504	490.5
1	2	2	1	487	475.9
1	2	2	2	798	808.0
2	1	1	1	6	6.5
2	1	1	2	55	56.8
2	1	2	1	35	38.0
2	1	2	2	70	64.7
2	2	1	1	46	31.4
2	2	1	2	262	275.5
2	2	2	1	125	136.1
2	2	2	2	464	454.0

Key pointers in using loglinear analysis are as follows:
(1) The most important variables for a table are usually variables in a saturated model with standardized values greater than 2.0.
(2) The Y^2 statistic follows the chi-squared distribution. It can therefore be tested to determine whether a model provides a good fit to the table.
(3) The best model for a table is usually the model that provides a good fit according to the Y^2 statistic. If more than one model provides a good fit, you should usually select the model with the fewest number of terms.

We can verify these results by developing two-way tables as shown in Exhibit 6-13. These tables were derived from the figures in Exhibit 6-7. The first table shows that when a detector operated, the odds of fire confinement to the room of origin were 3.62:1 (601 divided by 166). On the other hand, when a fire detector did not operate, the odds drop to 2.10:1 (1,884 divided by 897). Similarly, if equipment is involved in the fire, the odds are 3.95:1 (837 divided by 212) that the fire will be confined to the room of origin, as compared to 1.94: 1 (1,648 divided by 851) if equipment is not involved. Finally, the third table shows similar column percentages so that we can conclude independence of extent of flame damage and whether fabric is the type of material ignited.

Exhibit 6-13 Two-Way Interaction Tables

Chicago Residential Fires-1990

	Detector Operated	Detector Did Not Operate	Total
Confined to Room	601	1,884	2,485
Column Percent	78.4	67.7	
Extended Beyond Room	166	897	1,063
Column Percent	21.6	32.3	
Total	767	2,781	3,548
	100.0	100.0	

	Equipment Involved	No Equipment Involved	Total
Confined to Room	837	1,648	2,485
Column Percent	79.8	65.9	
Extended Beyond Room	212	851	1,063
Column Percent	20.2	34.1	
Total	1,049	2,499	3,548
	100.0	100.0	

	Fabric Material	Other Material	Total
Confined to Room	801	1684	2,485
Column Percent	68.5	70.8	
Extended Beyond Room	369	694	1,063
Column Percent	31.5	29.2	
Total	1,170	2,378	3,548
	100.0	100.0	

The key point is that loglinear analysis identified the important and unimportant interactions. We did not have to attempt the identifications by analyzing several two-dimensional and three-dimensional tables. The systematic approach provided by loglinear analysis resulted in an excellent model of the table with important insight into the conditions under which fires are contained to the room of origin.

Summary

Loglinear analysis is an systematic approach for analyzing multidimensional tables. The aim of loglinear analysis is to identify the important interactions among the variables in the table. Practical application of loglinear analysis means that the user identifies several different combinations of interactions and then develops a loglinear model for each combination. Each model produces estimates for the cells in the table table these estimates can he compared against the actual table values. The most appropriate model can then he selected as the final representation of the table.

Chapter 6

PROBLEMS

1. Calculate the odds ratio for the table in Exhibit 6-1 to verify independence of the two variables.

2. Calculate the odds ratios for the three tables in Exhibit 6-13.

3. Use the term values in Exhibit 6-11 to verify the model values in Exhibit 6-12 for cells (1,1,1,1), (1,1,2,2), and (2,1,2,1).

4. In the main example for this chapter, it was shown that fabric as the material ignited was not considered an important variable from the loglinear analysis. As an alternative, we substitute the area of origin to give the following four variables

Variable	Description
A	Fire was confined to the room of origin, or was not
B	Detector performed, or did not (or was not present)
C	Fire started in functional area (assembly area, sleeping room, kitchen, dining area, etc.), or did not
D	Equipment was involved in the fire, or was not

The data for Chicago residential fires for 1990 with these definitions is:

A	B	C	D	Frequency	A	B	C	D	Frequency
1	1	1	1	220	2	1	1	1	25
1	1	1	2	26-1	2	1	1	2	78
1	1	2	1	34	2	1	2	1	16
1	1	2	2	68	2	1	2	2	47
1	2	1	2	433	2	2	1	1	119
1	2	1	2	764	2	2	1	2	363
1	2	2	1	147	2	2	2	1	56
1	2	2	2	445	2	2	2	2	329

The saturated model gives the following results:

Term	Term Value	Standardized Value	Term	Term Value	Standardized Value
A	.486	16.2	BD	.094	3.1
B	-.707	-23.6	CD	.101	3.4
C	.413	13.8			
D	-.475	-15.8	ABC	.094	3.1
			ABD	.007	.2
AB	.074	2.5	ACD	.029	1.0
AC	.190	6.3	BCD	-.047	-1.6
AD	.157	5.2			
BC	.105	3.5	ABCD	.042	1.4

a. Analyze the standardized values to determine the most important interactions among the variables.

b. Treat variable A as the response variable and develop several hierarchical models of interest.

5. For the data from Problem 4, the following are model results for selected hierarchical models.

Model	Terms	Y^2	Freedom
1	ABC/AD/BC/BD/CD	3.5	4
2	AB/AC/AD/BC/BD/CD	10.9	5
3	ABC/BC/BD/CD	54.8	5
4	AB/AC/BC/BD/CD	62.4	6
5	AB/AD/BC/BD/CD	52.2	6
6	AC/AD/BC/BD/CD	28.5	6

a. Select the two best models from these alternatives.

b. Which of the two models is the better model?

c. From the data in Problem 4, develop two-way tables for variable A against the other three variables.

d. Calculate the odds and odds ratios for each of the three tables.

Chapter 7

CORRELATION AND REGRESSION

Introduction

In this chapter we present correlation and regression analysis for continuous data. Correlation is a statistical measure which indicates the degree to which one variable changes with another variable. We know, for example, that calls from citizens for Emergency Medical Services (EMS) increase with population growth. That is, as population increases, we normally expect more medical services calls from citizens. Statistically, we say there is a positive correlation between population and EMS calls. The correlation measures the strength of association between the two variables.

With regression we go a step further and create a straight line, with an associated regression equation, through a group of points. The result is as an analytical description of the relationship between two variables. The regression equation defines the straight line algebraically. As we will see in this chapter, population predicts fire department workload fairly accurately with a regression equation. In general, we can say that a regression line summarizes a group of points in a manner similar to the way an average summarizes a group of numbers.

This chapter starts with the scatter diagrams illustrated in Chapter 3 and proceeds with the calculation of correlation. Next, we give an example of a regression equation and discuss several applications. We then discuss how to calculate a regression line. The chapter concludes with a second example on the relationship between increases in population and increases in EMS activity for a fire department.

Scatter Diagram

Exhibit 7-1 shows the data for a scatter diagram presented in Chapter 3 on population protected and number of fires during 1989 for 18 selected jurisdictions.” Exhibit 7-2 is a scatter diagram of the data. The horizontal axis gives population (in thousands) and the vertical axis gives the number of fires. We can see from the exhibit that fire levels are higher with greater population. The general trend is clear although the pattern is not perfect. We use the term “not perfect” to mean that the points do not fall on a straight line.

With relationships depicted in this manner, the usual terminology is to label one variable as the **independent** variable, and the other as the

---

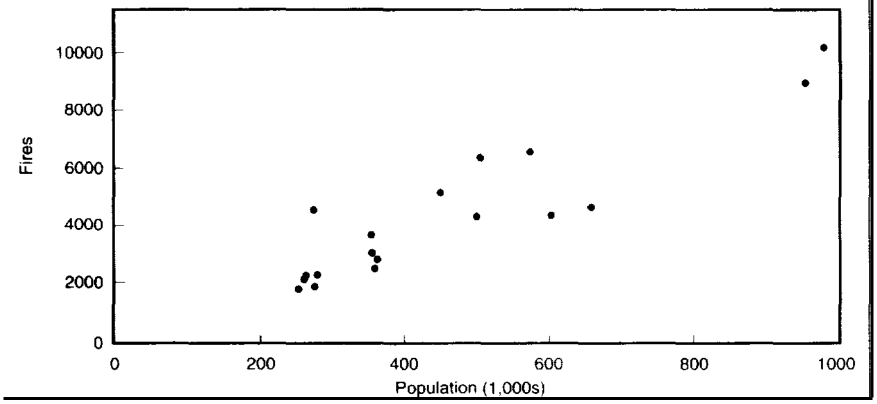
9. In Chapter 3, we identified Houston, Texas and Detroit, Michigan as outliers in the data That is, they have a different relationshIp than the other cities between population and EMS calls For the purposes of this chapter, we have therefore dropped them from the analysis.

Exhibit 7-1 Population and Fires for 1989–Selected Cities

City	Population Protected	Fires
Arlington, Texas	254,500	1,644
Wichita, Kansas	261,000	1,978
St. Paul, Minnesota	264,800	2,041
Corpus Christi, Texas	274,500	1,769
Newark, New Jersey	275,200	4,442
Norfolk, Virginia	280,000	2,140
Toledo, Ohio	354,600	3,597
Minneapolis, Minnesota	356,700	2,897
Omaha, Nebraska	360,000	2,336
Cincinnati, Ohio	364,000	2,645
Fort Worth. Texas	450,100	5,075
Denver, Colorado	500,000	4,244
Cleveland, Ohio	505,600	6,324
Boston, Massachusetts	574,300	6,479
El Paso, Texas	603,900	4,333
Columbus, Ohio	660,000	4,561
Dallas, Texas	982,800	10,210
San Antonio, Texas	956,200	8,957

dependent variable. In Exhibit 7-2, population serves as the independent variable and fires as the dependent variable. The independent variable is viewed as influencing the dependent variable. Obviously, population influences the number of fires; greater population means more structures and more vehicles, which, in turn, may lead to more fires. The question is how strong is the relationship between population and fires.

Exhibit 7-2 Scatter Diagram of Population & Fires-Selected Cities-1989



Correlation measures the strength of association between two variables. One variable is usually called the **independent variable** and the other is called the dependent variable. A *strong association* between the two variables means that knowing the value of the independen Variable helps to predict the value of the dependent variable. conversely, a weak association means that the independent variable does not help much in determining the values of the dependent variable.

Correlation Coefficient

The **correlation coefficient,** more commonly just called **correlation,** measures the strength of association between two variables. In the next section, we will provide the calculation for the correlation coefficient, but here we want to understand it and discuss its key properties.

The first property to know is that a correlation is always between -1 and +1. A correlation of exactly -1 or +l is called a perfect correlation, and means that all the points fall on a straight line. If a correlation is zero, then there is no association between the two variables. That is, one variable does not assist in knowing or predicting the value of the second variable. As a correlation moves from zero towards either +1 or -1, the strength of association between the two variables also increases.

Exhibit 7-3 shows several scatter diagrams with correlations ranging from 0 to +l. Note that when the correlation is zero, the dots are randomly scattered with no pattern. As the correlations increase, the association becomes stronger–the dots begin to cluster together and we can start to visualize a straight line through the points.

The correlations in Exhibit 7-3 are positive because the patterns of the dots always move from the lower left to the upper right. Just the opposite is true with negative correlations, as shown in Exhibit 7-4, where the patterns move from the upper left to the lower right. The ***negative*** relationship means that as the independent variable **increases,** the dependent variable **decreases.** The direction of the dots is the distinguishing feature between negative and positive correlations. A correlation of -.9 is just a strong as a correlation of +.9, but the pattern of the dots is in the opposite direction.

A correlation has no units of measurement associated with it. That is, a correlation is not expressed in terms of the independent or dependent variable. It is dimensionless. The interpretation of a correlation coefficient generally depends on its closeness to -1, 0, or +1. Correlations close to 0 mean there is no association between the two variables while correlations close to -1 or +l indicate strong associations.

Exhibit 7-3 Positive Correlations

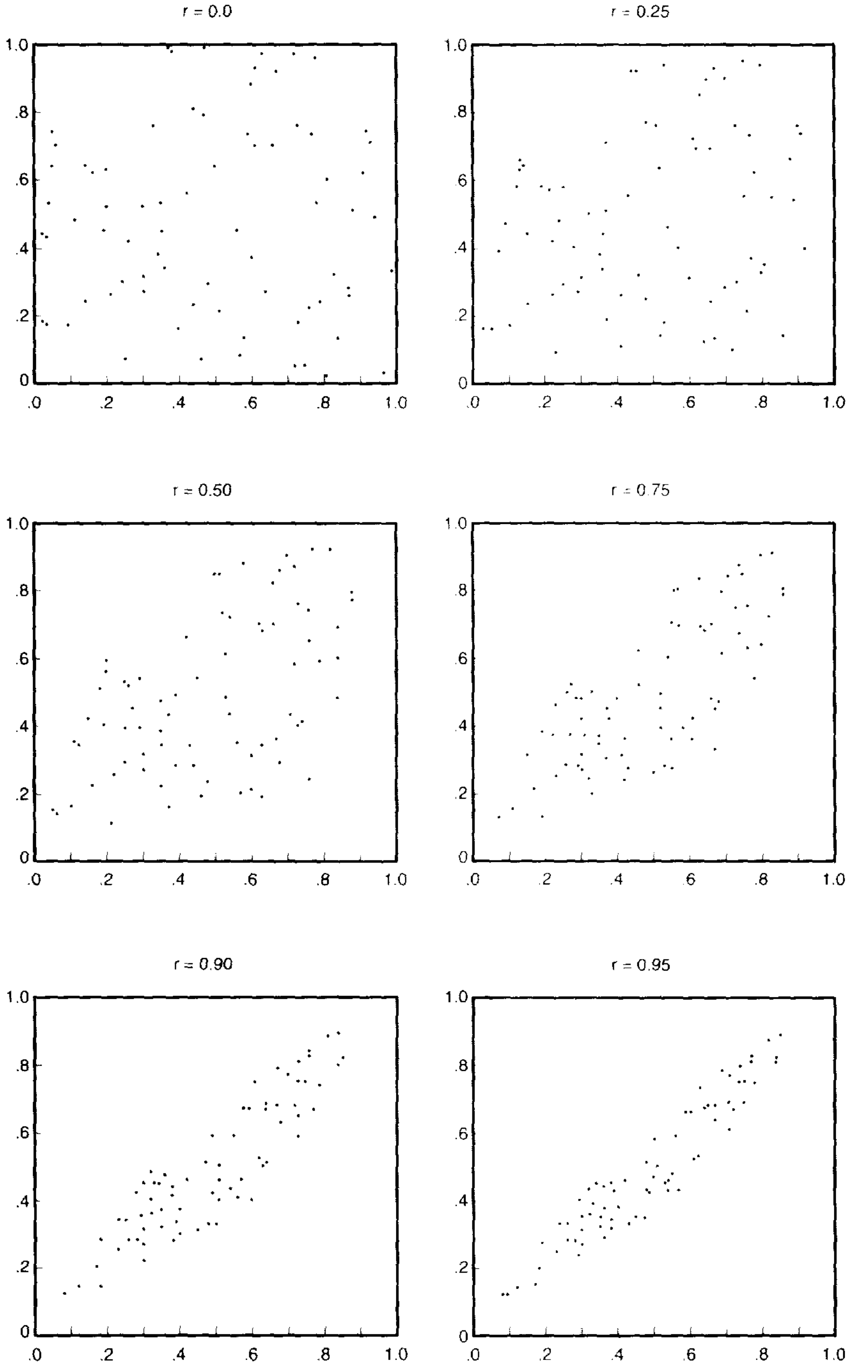
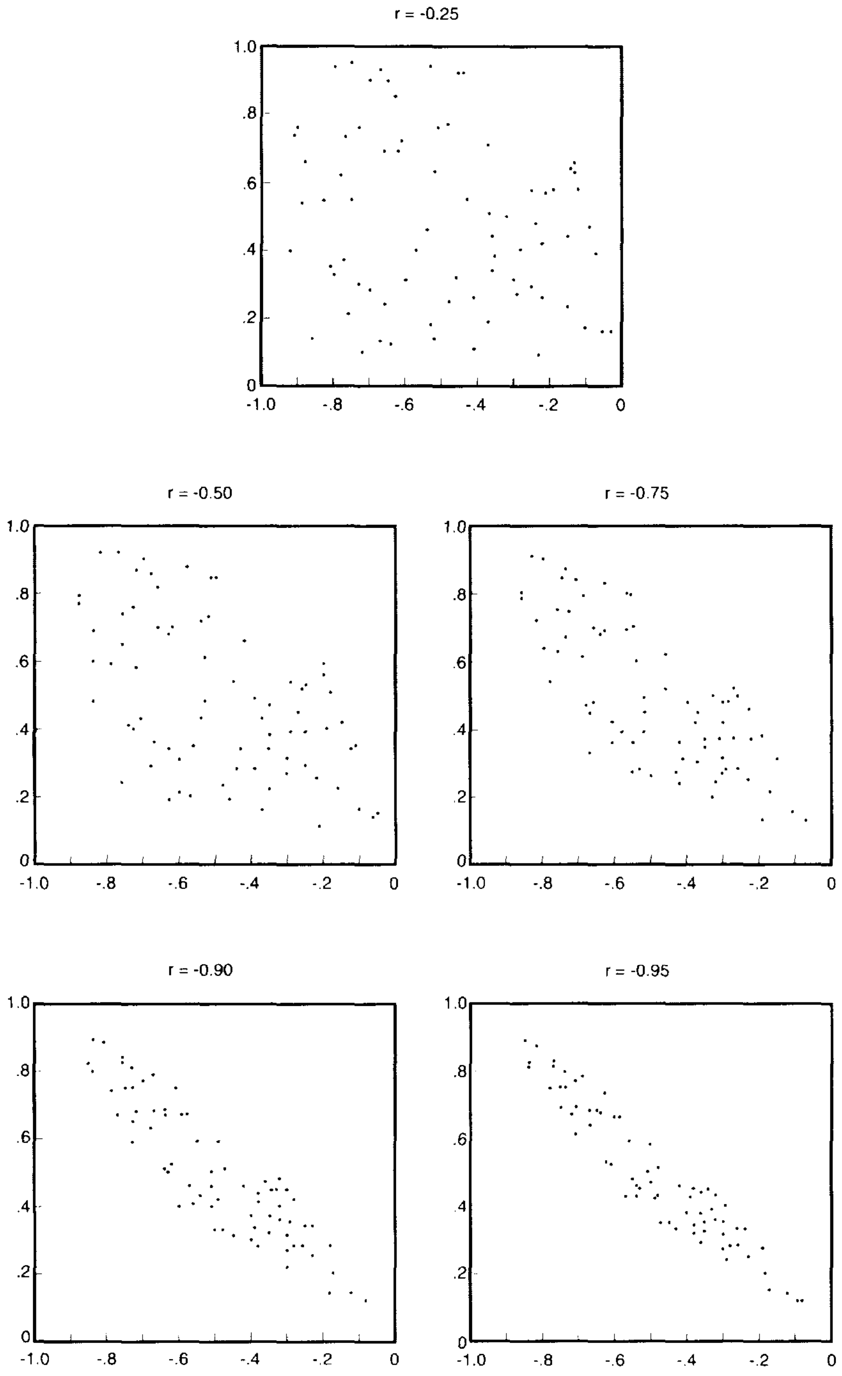


Exhibit 7-4 Negative Correlations



Another important point to know is that correlations are **not** arithmetically related to each other. For example, a correlation of .6 is not twice as strong as a correlation of .3. We can obviously say that a correlation of .6 reflects a stronger association than a correlation of .3, but we must stop short of exact specification of the difference.

Finally, there is no relationship between correlations and percentages. Correlations range between - 1 and +l, but have nothing to do with percentages. Again, the correlations are not arithmetically related to each other.

Correlations are always between -1 and +1. Stronger associations are indicated as the correlations approach-1 and +1 while correlations close to 0 indicate no association. With **positive correlations,** the dots move from the lower left to the upper right while with **negative correlations,** the opposite is true.

Calculating the Correlation

This section will show you how to calculate a correlation. We will use the data from the selected cities, as shown in Exhibit 7-1. As it turns out, there is a correlation of .92 between population and fires. This is a high correlation indicating a strong association between the two variables.

To calculate the correlation, we perform the following steps:

1. Convert the values of both variables to **standard units,** as defined below.
2. Take the product of the standard units for each pair.
3. Sum the resulting values and divide by the number of points minus 1.

To convert a value into standard units, subtract the average"' and divide by the standard deviation. Returning to the numbers in Exhibit 7-1, we can calculate the following information:

Exhibit 7-5 Averages and Standard Deviations for Selected Cities

	Average	Standard Deviation
Population (in thousand)	459.90	224.78
Fires	4,204.00	2,469.43

---

10. In this chapter, the average is the arithmetic mean (rather than the median or mode) and the standard deviation is the sample standard deviation (rather than the population standard deviation).

$$\text{Standard units} = \frac{x_j - \bar{x}}{s}$$

Where $\bar{x}$ = arithmetic mean
s = sample standard deviation

Exhibit 7-6 gives the steps for the correlation calculation. The column labeled "Population Standard Units" lists the standard units for the population of each city. For the first population of 254.5, we obtain the standard unit by subtracting the average of 459.9 and dividing by the standard deviation of 218.45. The calculation proceeds as follows:

$$\underset{\text{(for 254.5)}}{\text{Standard Unit}} = \frac{254.4 - 459.9}{224.78} = \frac{-205.4}{224.78} = -0.91$$

For the standard unit for the first figure for fires, the calculation proceeds in the same manner using the average and standard deviation for fires.

$$\underset{\text{(for 1,644)}}{\text{Standard Unit}} = \frac{1{,}644 - 4{,}204}{2{,}469.43} = \frac{-2{,}560}{2{,}469.43} = -1.04$$

Exhibit 7-6 Correlation Calculation

City	Population (thousands)	Fires	Population Standard Units	Fires Standard Units	Product
Arlington	254.5	1,644	-0.91	-1.04	0.95
Wichita	261.0	1,978	-0.88	-0.90	0.80
St. Paul	264.8	2,041	-0.87	-0.88	0.76
Corpus Christi	274.5	1,769	-0.82	-0.99	0.81
Newark	275.2	4,442	-0.62	0.10	-0.08
Norfolk	280.0	2,140	-0.80	-0.84	0.67
Toledo	354.6	3,597	-0.47	-0.25	0.12
Minneapolis	356.7	2,897	-0.46	-0.53	0.24
Omaha	360.0	2,336	-0.44	-0.76	0.34
Cincinnati	364.0	2,645	-0.43	-0.63	0.27
Ft. Worth	450.1	5,075	-0.04	0.35	-0.02
Denver	500.0	4,244	0.18	0.02	0.00
Cleveland	505.6	6,324	0.20	0.86	0.17
Boston	574.3	6,479	0.51	0.92	0.47
El Paso	603.9	4,333	0.64	0.05	0.03
Columbus	660.0	4,561	0.89	0.14	0.13
Dallas	982.8	10,210	2.33	2.43	5.66
San Antonio	956.2	8,957	2.21	1.92	4.25
Total					15.57

Correlation (Total divided by 17) = .916

The last column gives the product of the two calculations, which is 0.95 (-0.91 times -1.04). These calculations are performed for each of the 18 data points. The sum of this column is 15.57, and the correlation is defined as this total divided by 17, which results in the correlation of .916.

Other Ways to Calculate Correlation*

Several equivalent formulas exist for calculating correlations. We selected the above procedure to illustrate the connections between correlations, averages, and standard deviations. The correlation is an average of the products of standard units, except that we divided our sum by the number of points minus one.

In algebraic terms, the correlation can be seen on the following page.

$$\boxed{\text{Correlation} = \frac{\sum (x_i - \bar{x})(y_i - y)}{(n-1)\, s_x s_y}}$$

where s_x is the sample standard deviation of x s_y is the sample standard deviation of y. This is the equation that was used to Calculate the correlation in Exhibit 7-6.

An equivalent formulation for the correlation is given by the following:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

In this equation, we have taken the prior equation and replaced the standard deviations with their actual formulas.

Another way of expressing the correlation is by the following:

$$r = \frac{\text{Cov}(x,y)}{\sqrt{\text{Var}(x)\,\text{Var}(y)}} \quad \text{where Cov}(x,y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

Var(x) is the sample variance of x (the square of the sample standard deviation for x), and Var(y) is the sample varance of y.

The quantity Cov(x,y), which is called the covariance of x and y, measures the extent to which the two variables rise and fall together. By itself, the covariance is hard to interpret because it is not expressed in units of either x or y and it does not have an upper or lower bound like a correlation. However, dividing the convariance by the two variances standardizes the final result to a correlation, which is always between -1 and + l.

Exhibit 7-7 Correlation Matrix

Variable	Population	Structure Fires	Vehicle Fires	Other Fires	Civilian Injuries	Fire Civilian Fatalities	Fire Service Injuries	Service Fatalities	Dollar Loss
Population	1.00	.82	.73	.94	.25	.42	.53	.2 3	.37
Structure Fires	.82	1.00	.90	.79	.32	.70	.68	-.32	.53
Vehicle Fires	.73	.90	1.00	.79	.18	.86	.77	-.24	.30
Other Fires	.94	.79	.80	1.00	.26	.53	.53	-.22	.29
Civilian Injuries	.25	.32	.18	.26	1.00	.14	.31	-.26	.76
Civilian Fatalities	.42	.70	.86	.53	.14	1.00	.66	-.24	.17
Fire Service Injuries	.53	.68	.77	.53	.31	.66	1.00	-.27	.33
Fire Service Fatalities	-.23	-.32	-.24	-.22	-.26	-.24	-.27	1.00	-.18
Dollar Loss	.37	.53	.30	.29	.76	.17	.33	-.18	1.00

Correlation Matrix*

Exhibit 7-7 on the previous page shows **a correlation matrix** for several variables for the 18 cities discussed in the previous section. Each entry in the matrix is a correlation. For example, the first line shows a correlation of .82 between population and structure fires.

Note that the diagonal of the correlation matrix is always 1.00, since this is the correlation of a variable with itself. Also, the correlation is symmetric about the diagonal. The correlation of .73 (between population and vehicle fires) from the first line also appears in the first column. The lower half of the matrix could, in fact, be omitted without losing any information about the correlations. However, the complete table is usually displayed co you can find specific correlations easier.

There are several high correlations in the matrix. For example, the correlation between population and structure fires is .82, and between population and other fires is .94. On the other hand, there are several low correlations, indicating relatively little relationship between the variables. The correlation between population and civilian injuries is only .25. This means that population does not necessarily provide a good indicator of the number of civilian injuries from fires.

Regression Line

Returning to Exhibit 7-2, we now want to know more about a straight line that best fits the dots. As we will show in the next section, the **regression line** for relating population to fires for these cities is given by:

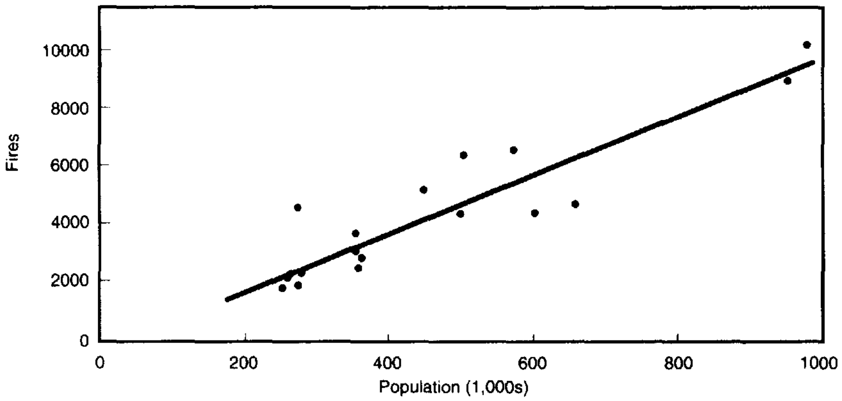
$$\text{Fires} = 10.06 \times \text{Population} -424.03$$

The value 10.06 is called the **regression coefficient,** or **slope,** of the regression line, and the value -424.03 is the **constant** or **intercept.** The next section explains how to calculate these values.

Exhibit 7-8 shows the regression line within the scatter diagram for population and fires. Note that the dots cluster nicely around the regression line. The regression line is a representation of these dots, just as the average is a representation of a single list of numbers.

> The regression line estimates the average value for the dependent variable for a given value of the independent variable. the regression line is the numerical representation of a scatter diagram.

Exhibit 7-8 Scatter Diagram of Population & Fires with Regression Line



To see how to use this regression, consider the population of 500,000 for Denver, Colorado. The number of fires estimated by the regression line is then:[11]

$$\text{Denver Fires} = (10.06 \times 500.0) - 424.03$$
$$= 5{,}030 - 424.03$$
$$= 4{,}607.5$$

Denver actually experienced 4,244 calls. The regression estimate was off by 363.5 fires, or about 8.5 percent.

From the equation, you can see that the number of fires changes by 10.06 every time the population variable changes by 1 (that is, the population changes by 1,000). In general, the regression coefficient reflects the change in the dependent variable with a one unit change in the independent variable.

Another feature of a regression line is that the line always goes through the point created by the averages for the two variables. In our example, the average population (in thousands) is 459.90 (see Exhibit 7-5). For this population average of 459.90, we have the following from the regression line:

$$\text{Denver Fires} = (10.06 \times \text{Population}) - 424.03$$
$$= (10.06 \times 459.90) - 424.03$$
$$= 4{,}204.0$$

The result is 4,204.0 fires, which is the average number of fires for the 18 cities.

---

[11] If you duplicate this calculation, you will not get an answer of exactly 4,607.5 fires because of rounding errors. More precisely, the slope is 10.0613 and the intercept is -424.0344.

The regression line can also be used to estimate the number of fires for cities with other populations. Suppose, for example, that your jurisdiction has population protected of 700,000 persons. How many fires can be expected with this population? The answer is easily, calculated:

$$\begin{aligned} \text{Fires} &= (10.06 \text{ x } 700) -424.03 \\ &= 6{,}618 \end{aligned}$$

Of course, this calculation assumes that your jurisdiction fits the general pattern of these cities and is not an outlier. It is also unlikely, that there will be **exactly** 6,618 fires for a population of 700,000 persons. This is strictly **a point estimate** obtained by applying the regression line. In the next chapter, **we** will describe how to calculate an **interval** estimate around this point estimate.

Calculating the Regression Line

The regression line has the general form:

$$y = mx + b$$

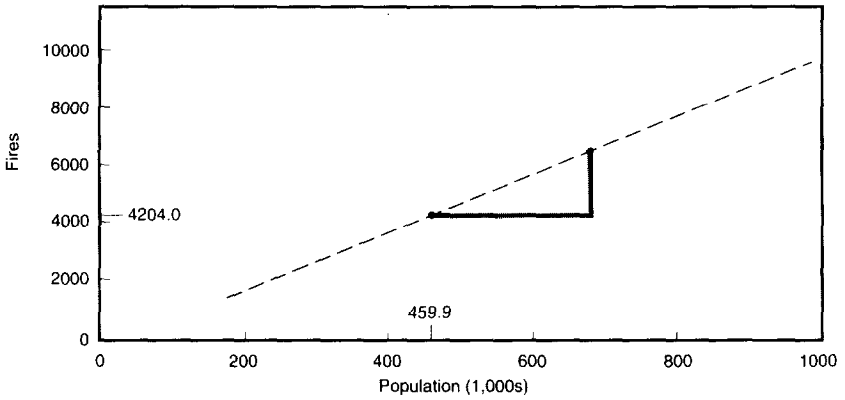
Where y is the dependent variable, x is the independent variable, m is the slope (or regression coefficient), and b is the intercept (or constant). In our example, the dependent variable is the number of fires and the independent variable is population. We will now calculate the slope and intercept for our example. Just as there are several equivalent equations for correlation, there are also several ways to calculate the slope and intercept, all of which result in the same answers. Our approach takes advantage of the information we have about the averages, standard deviations, and correlation for our two variables.

In Exhibit 7-9, we have placed a dot at the pair formed by the two averages (459.9 population and 4,204 fires). As stated in the previous section, we want the regression line to pass through this point. We get a second dot by moving one standard deviation for the population to the right and upward by one standard deviation for fires **times** the correlation. The regression line is then formed graphically by drawing a straight line through these two points.

Algebraically, the slope of the straight line is:

$$\text{Slope} = \frac{\text{r x S.D. of Fires}}{\text{S.D. of Population}} = \frac{(.916 \text{ x } 2{,}469.43)}{224.78} = 10.06$$

Exhibit 7-9 Graphical Formulation of Regression Line



Calculating the intercept takes advantage of the fact that the regression line goes through the point created by the two averages. That is, we determine the intercept, b, by:

Fires Average = m x Population Average + b

4,204.O = 10.06 x 459.90 + b

-424.03 = b

The final result is the regression line:”

Fires = 10.06 x Population -424.03

In a **regression line,** when the independent variable Changes by one standard deviation , the dependent variable changes by **r** standard deviations.

Standard Error of the Regression*

The **standard error** of the regression is an estimate of the accuracy of the regression line you have developed. In this section we will present two equivalent ways to calculate the standard error. One way revolves around the calculation of the **residuals** of the regression line, while the other way is a quicker algebraic formula.

The residuals for a regression are the differences between the data points and the estimates from the regression equation. As we showed earlier, the population protected for Denver was 500,000 so that the estimated number of fires is 4,607.5 (10.06 x 500 - 424.03). Since Denver actually had

---

12 The more precise numbers were used in the calculations to get these results

4,244 fires, the difference is -363.5 (4,244 - 4,606.5). Note that the difference is negative since the estimated number is greater than the actual number. Positive residuals come from points located above the regression line, and negative residuals are from points below the regression line.

Exhibit 7-10 summarizes the residual data for all the points in the regression. The first three columns are from Exhibit 7-1 she-wing population and fires. The fourth column is the estimated number of fires from the regression equation and the next column shows the residuals. It should be noted that the sum of the residuals is zero, a characteristic of residual calculations for regressions. The last column in the exhibit is the square of each residual (the residual multiplied by itself). The total for this last column is 16,681,736.4, which is sometimes referred to as the **sum of squared errors,** or **SSE.**

Exhibit 7-10 Residuals for Regression Line

City	Population (thousands)	Fires	Estimated Fires	Residual	Residual Squared
Arlington	254.5	1,644	2137.0	-493.0	243,081.3
Wichita	261.0	1,978	2,202.4	-224.4	50,374.7
St. Paul	264.8	2,041	2,240.7	-199.7	39,873.3
Corpus Christi	274.5	1,769	2,338.3	-569.3	324,097.2
Newark	275.2	4,442	2,345.3	2,096.7	4,395,985.1
Norfolk	280.0	2,140	2,393.6	-253.6	64,334.6
Toledo	354.6	3,597	3,144.4	452.6	204,890.0
Minneapolis	356.7	2,897	3,165.5	-268.5	72,084.1
Omaha	360.0	2,336	3,198.7	-862.7	744,239.4
Cincinnati	364.0	2,645	3,238.9	-593.9	352,771.4
Ft. Worth	450.1	5,075	4,105.4	969.6	940,160.4
Denver	500.0	4,244	4,607.5	-363.5	132,155.2
Cleveland	505.6	6,324	4,663.9	1660.1	2,755,981.4
Boston	574.3	6,479	5,355.2	1,123.8	1,262,876.3
El Paso	603.9	4,333	5,653.1	-1,320.1	1,742,640.2
Columbus	660.0	4,561	6,217.6	-1656.6	2,744,431.9
Dallas	982.8	10,210	9,466.0	744.0	553,518.6
San Antonio	956.2	8,957	9,198.3	-241.3	58,241.3
Total				0.0	16,681,736.4

We calculate the standard error by the following equation:

$$\text{Standard Error} = \sqrt{\frac{SSE}{n - 2}}$$

where SSE is the Sum of Squared Errors

In this equation, n is the number of points. In our example, we have 18 points, so that n-2 = 16. The mean square error is therefore calculated as:

$$\text{Standard Error} = \sqrt{\frac{16{,}681{,}736.4}{16}} = \sqrt{1{,}042{,}608.5} = 1{,}021.08$$

The standard error tells us how far, on average, the estimates from the regression line deviate from the actual numbers. A small standard error reflects a good fit of the regression line to the data while a large standard error means the regression line is not very representative of the data points. A useful rule of thumb is to see if the standard error is small relative to the dependent variable. In our example, the number of fires ranges between 1,644 and 9,466 with an average of 4,204.O calls. Our standard error of 1,021.08 is relatively small compared to these data values.

The standard error has another interesting feature. From statistical theory, approximately 68 percent of the actual values should be within one standard error and 95 percent within two standard errors. In our example, 13 of the 18 cities have fire figures within one standard error and 17 are within two standard errors. Thus, these results are in line with statistical theory.

The calculation of the standard error in the above manner obviously is a time consuming job when you have a large number of points. A more direct way takes advantage of knowing the correlation coefficient and standard deviation of the dependent variable, as reflected in the following equation:

$$\text{Standard Error} = \sqrt{\frac{n-1}{n-2}} \times \text{S.D. of Fires} \times \sqrt{1 - r^2}$$

Since the standard deviation for the fires is 2,469.43 and the correlation is .916, the standard error can therefore be calculated as:

$$\text{Standard Error} = \sqrt{\frac{n-1}{n-2}} \times \text{S.D. of Fires} \times \sqrt{1 - r^2}$$

$$= \sqrt{\frac{17}{16}} \times 2{,}469.43 \times \sqrt{.161}$$

$$= 1.031 \times 2{,}469.43 \times .401$$

$$= 1{,}021.08$$

The advantage of this equation is that we avoid the arduous residual calculations by knowing the correlation and standard deviations.

Coefficient of Determination: Explained Variation*

The **coefficient of determination** is another way to determine how well a regression line fits the data points. The coefficient of determination is motivated by the following argument. One crude and simple approach for predicting a variable is merely to use the average as the prediction. With this approach, the same prediction (namely, the average) is always made regardless of the value of independent variable. We can then form squared residuals from the average just as we did in the previous section. The result is called the **total sum of squares,** or **SST:**

$$SST = \sum (y - \bar{y})^2$$

where y is the actual number of fires for a city and $\bar{y}$ is the average number of fires.

In our example, the SST is 103,667,214 . We should note that the sample variance is the SST divided by 17. Thus, we can view SST as a measure of variability in the fire figures.

The coefficient of determination is detined as:

$$R^2 = \frac{SST - SSE}{SST}$$

where SST is the total sum of squares and SSE is the sum of squares for errors.

You may recall that SSE is derived from the residuals obtained as the difference between the actual values and the predicted values. In our example, we calculated the SSE (see Exhibit 7-10) as 2,311,594.1. The coefficient of determination for our example is therefore:

$$R^2 = \frac{103{,}667{,}214 - 16{,}681.736}{103{,}667{,}214} = .839$$

We can interpret R^2 in the following way. With the regression equation, the amount of error in the predictions (as measured by the sum of squared errors) is 83.9 percent smaller than when the average is used as the predictor. The value R^2 therefore indicates how much better the linear regression equation is over simply using the average.

Another way to view R^2 is to say that population explains 83.9 percent of the variability in the fire figures. Or conversely, 83.9 percent of the variability in the fire figures is explained by population.

An important feature of the coefficient of determination is that it is equal to the square of the correlation coefficient. In our example, the correlation is .916 and the square of this number is .839. The designation of the coefficient of determination by R^2 is intentional to indicate its relationship to the correlation coefficient. r.

In summary, if we know the correlation coefficient, then a quick calculation shows how much variation will be explained by a regression line. In Exhibit 7-7, we showed correlations for several other variables. The exhibit indicated a correlation of .25 between population and civilian injuries. If we performed a regression of population against civilian injuries, we would explain only 6.25 percent of the variation (since the square of the correlation is .0625 or 6.25 percent). The point is that we could perform such a regression, but the results would not be very satisfying.

An Example with Population and EMS Calls

In this section we present another regression example showing the relationship between population and EMS calls in Prince William County, Virginia. The Prince William County Fire Department has 15 fire stations, of which 4 stations include paramedics for handling EMS calls. Exhibit 7-11 shows the population growth in the county from 1981 to 1991 along with the number of EMS calls handled by paramedics. (You may recall that we showed a scatter diagram of these data in Chapter 3.) One immediate conclusion is that the county is growing as reflected by the population increase of 71,600 over the eleven-year period. The number of EMS calls also increased substantially over the same time period from 9,538 calls in 1981 to 12,744 calls in 199l.

Exhibit 7-11 Population & EMS Calls Prince William County, Virginia

Year	Population	EMS Calls
1981	152,300	9,538
1982	156,700	9,578
1983	159,200	9,657
1984	164,100	9,744
1985	169,700	10,072
1986	176,000	11,703
1987	184,700	11,982
1988	205,000	11,843
1989	221,300	13,074
1990	219,000	13,175
1991	223,900	12,744

Exhibit 7-12 gives the basic statistics and regression line for population and EMS calls in the county.

Exhibit 7-12 Basic Statistics and Regression Results

Variable	Average	Standard Deviation
Population (thousands)	184.72	27.70
EMS Calls	11,191.82	1,491.31

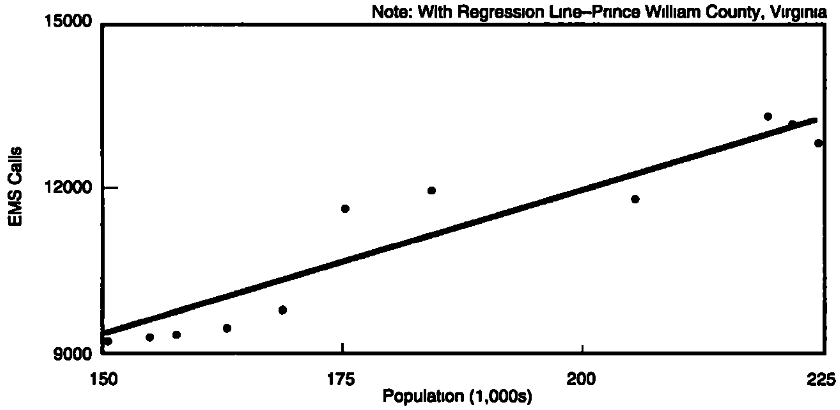
Correlation (r): .946
Coefficient of Determination (R^2): .895
Regression line: EMS Calls = 50.96 x Population + 1,778.6
Standard Error: 506.8

Exhibit 7-13 shows a scatter diagram of our data along with the regression line. In this example, the fit of the regression line is excellent. An application of this regression line is to anticipate the number of EMS calls for future years. Suppose the county expects the population to reach 300,000 residents in the next few years. How many EMS call can be expected with this population? A point estimate can be made with the regression line:

EMS Calls = (50.96 x 300) + 1,778.6
= 17,067

The implication of this increase is that the fire department probably needs to start considering expansion of its EMS program to handle the increased workload.

Exhibit 7-13 Scatter Diagram of Population & EMS Calls



Summary

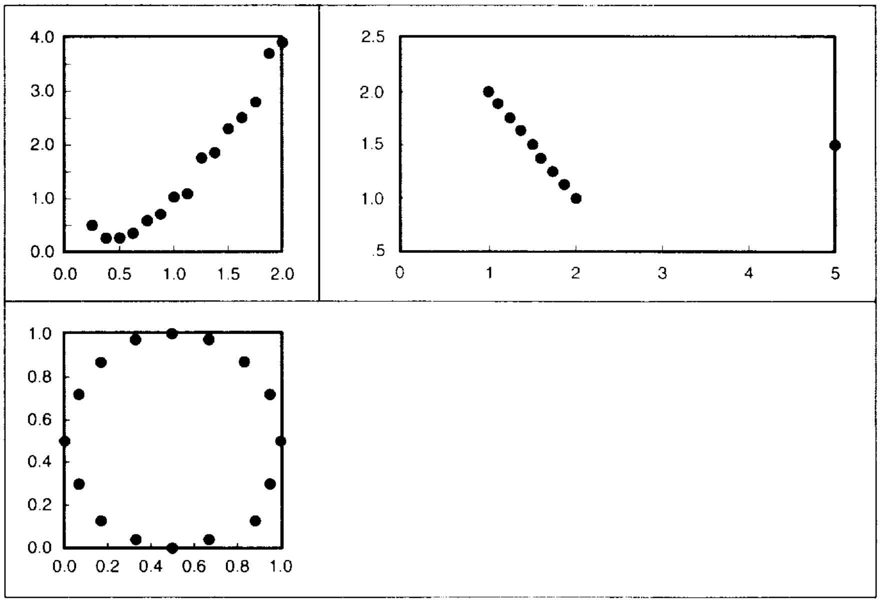
Correlation and regression analysis are two powerful tools for analyzing the relationship between two continuous variables. A correlation close to zero tells you there is no relationship, which means that knowing the value of one of the variables does not tell you much about the value of the other variable. Correlations close to -1 or +1 indicate a strong relationship between two variables. We found, for example, a high correlation between population of jurisdictions and the number of fires. Regression analysis provides a way to quantify the relationship of two variables. The regression line is a representation of the scatter diagram of the two variables. We can apply the regression line to estimate the value of one variable given the other variable. With the regression of population on fires, for example, we can make estimates on the number of fires that a jurisdiction can expect given its particular population.

PROBLEMS

1. Do the following for each of the figures below

a. Estimate the correlation for each figure.

b. The second figure has a clear outlier. What is the correlation without this outlier.



2. The following data, which are from Memphis, Tennessee for 1991, are for structure fires in which the cause of the fire was children playing. All the fires in this list had incident times less than 30 minutes (incident time is from time of dispatch to the scene until time of completion). Develop a scatter diagram and estimate the correlation between incident time and dollar loss for these fires.

Incident Time	Dollar Loss	Incident Time	Dollar Loss
15	700	12	400
18	3,000	16	800
22	250	26	1,000
26	35	22	1,500
14	100	17	5,000
12	500	18	550
28	150	27	250
16	300	11	50

3. The department responded to an additional 73 structure fires caused by children playing in which the incident times were greater than 30 minutes. For these fires the correlation between incident times and dollar losses was .84, and the regression equation relating losses to incident time was as follows:

Fire Loss = 206.2 x Incident Time - 6,266.4

a. Calculate the estimated fire loss for an incident time of 45 minutes and an incident time of 3 hours (180 minutes).

b According to this regression, how much will fire loss increase when incident time increases 10 minutes?

c. In general, fire losses and incident times appear to have a fairly high correlation for incidents requiring more than 30 minutes, but not for incidents less than 30 minutes. Give reasons as to why this result might be true.

4. The number of fires in most jurisdictions has decreased over the last 20 years. However, the following data shows that the total dollar loss due to fires has steadily increased in the United States.

Year	Total Fire Losses (in millions)	Year	Total Fire Losses (in millions)
1967	$1,707	1979	$4,851
1968	1,830	1980	5,579
1969	1,952	1981	5,625
1970	2,328	1982	5,894
1971	2,316	1983	6,320
1972	2,304	1984	7,602
1973	2,639	1985	7,753
1974	3,190	1986	8,488
1975	3,190	1987	8,634
1976	3,558	1988	9,626
1977	3,764	1989	10,210
1978	4,008		

Source: Insurance Information Institute, New York, New York. Insurance *Facts,* annual publications.

a. Develop a scatter diagram showing the year and total dollar loss.

(continued from question 4)

b. The last two digits of the year (67, 68, 69, etc.) can be treated as a variable and we can then develop a regression equation between year and total dollar loss. The correlation between year and dollar loss is very high at .9753, and the means and standard deviations of these variables are as follows:

Variable	Mean	Standard Deviation
Year	78	6.782
Total Fire Losses	4,929.043	2,696.394

Determine the regression line of total fire losses as the dependent variable and year as the dependent variable.

c. With your regression equation, estimate the total fire losses for 1988, 1989, and 1990.

d. Draw the regression line on your scatter diagram. The standard error for this regression equation is 609.06. Draw a line above and below the regression line representing the standard error.

Chapter 8

MULTIPLE REGRESSION

Introduction

Chapter 7 discussed regression with only one independent variable (population). We were able to develop a reliable regression line because a high correlation existed between population and fires in cities. In other situations, however, we may know that several factors are related to a dependent variable. Our approach with these situations is to perform a **multiple regression,** which means that several independent variables are included in the regression.

Fortunately, many concepts from Chapter 7 carry over to multiple regression with little or no modification. One difference, however, is that calculations for multiple regression are more tedious than for single variable regression. A computer is therefore an indispensable tool for multiple regression. The computer's ability to perform many calculations quickly is a primary reason that multiple regression is a predominant analytical tool today. In this chapter, we will not emphasize the calculations, but will instead focus on understanding multiple regression equations. We can only scratch the surface on this subject. Interested readers should refer to the statistics books cited in Chapter 1 for additional information.

> **Multiple regression** is regression with several independent variables rather than a single independent variable. Multiple regression is beneficial when the dependent variable is related to several factors. The resulting multiple regression provides a means of estimating the dependent variable based on values of the several independent variables. Many of the concepts from Chapter 7, such as coefficient of determination and standard error, can be applied to multiple regression.

Boston Fires

The example for this section is based on 1990 census tract information for Boston, Massachusetts combined with data on residential fires for 1989 and 1990 by census tract. A total of 147 Boston census tracts had at least one fire during these two years. A preliminary analysis with correlations showed that residential fires in the Boston census tracts were related to four key independent variables:

- POPULATION: Population in a census tract.
- BOARDED: Number of boarded-up housing units in a census tract.

- FAMTYPE: Number of single-parent households in a census tract with one or more persons under 18 years of age.
- DENSITY: Number of households in a housing unit with one or more persons per room.

As we shall see later, census tracts with large values for these independent variables tended to have more fires, and vice versa. On the basis of these tendencies, the four variables are good candidates for a multiple regression equation with fires as the dependent variable. The resulting multiple regression equation provides insight into reasons why fires are high in one area and low in another area.

Data for the four independent variables were collected as part of the 1990 census conducted by the Bureau of Census. Population is, of course, the number of persons residing in a census tract. The other three variables can be considered as measurements of the socio-economic conditions of census tracts. Census tracts with poor socio-economic conditions usually have more boarded-up buildings, more households headed by single parents, and more persons per room. The parallel between socio-economic conditions and fires is probably no surprise to fire fighters and fire department administrators. The advantage of multiple regression is that we can quantify the relationships between fires and socio-economic conditions rather than depending on personal experiences.

For this analysis the census tract data were combined with residential fire data fire each census tract. Merging the data was possible because the Boston Fire Department records the census tract on each fire incident report. Table 8-1 shows the averages and standard deviations for the variables in this analysis.

Exhibit 8-1 Averages and Standard Deviations for Regression Analysis

Variable	Average	Standard Deviation
Residential Fires	13.1	9.0
POPULATION	1382.2	790.4
BOARDED	10.2	21.6
FAMTYPE	177.6	148.4
DENSITY	95.4	79.2

Note Boston. Mass. Data

The average number of fires for the 147 census tracts was 13.1 with a standard deviation of 9.0. These statistics were derived directly from the Boston Fire Department data. The other averages and standard deviations were calculated from 1990 census data for the 147 census tracts.

Exhibit 8-2 shows the correlation matrix for these five variables. The

DENSITY variable has the highest correlation with residential fires (.74), followed by FAMTYPE (.61), BOARDED (.37), and POPULATION (-35). The correlations among the independent variables are also important to review. As we will discuss in the next section, we do not want a multiple regression with independent variables that are highly correlated with each other. Exhibit 8-2 shows that the largest correlation for the independent variables is .58 between DENSITY and FAMTYPE. Most of the other correlations are low. In fact, the correlation between POPULATION and BOARDED is -.06, which means that these two variables have virtually a random relationship.

Exhibit 8-2 Correlation Matrix-Boston Census Tracts-1990

Variable	FIRES	POPULATION	BOARDED	FAMTYPE	DENSITY
FIRES	1.00	.35	.37		
POPULATION	.35		-.06		
BOARDED	.37	-.06	1.00		
FAMTYPE	.61	.12	.43		
DENSITY	.74	.35	.23		

As previously indicated, computers are a necessity for multiple regression because of the complicated nature of the calculations. We will not attempt to show calculations as we did in Chapter 7, because our objective is on interpreting and applying a multiple regression equation. For the residential fires in Boston, the resulting regression equation was as follows:

$$\begin{aligned}\text{Fires} = 2.0 + .0017 \times \text{POPULATION} + .068 \times \text{BOARDED} \\ + .013 \times \text{FAMTYPE} + .060 \times \text{DENSITY}\end{aligned} \quad (1)$$

We can estimate the number of fires in a census tract by knowing the values of the independent variables. For example, census tract 510 has 1,607 residents, 2 boarded-up units, 154 single-parent households, and 33 households with one or more persons per room. The estimated number of fires for this census tract is as follows:

$$\begin{aligned}\text{Fires} &= 2.0 + .0017 \times 1607 + .068 \times 2 + .013 \times 154 + .060 \times 33 \\ &= 8.92\end{aligned} \quad (2)$$

Census tract 510 experienced 10 fires so that the estimated number of fires is very close to actual experience.

Another interesting feature of the regression is that the coefficients for the independent variables in (1) are always positive. This means that the number of fires increases as these variables increase. According to the multiple regression results, increases in population, boarded-up units, household density, and single-parent households will result in increases in fires.

As with regressions in Chapter 7, we should he interested in how well the regression equation fits the actual data. One measure of fit is the **coefficient of determination,** which has exactly the same definition as in Chapter 7:

$$R^2 = \frac{SST - SSE}{SST} \tag{3}$$

where SST is the total sum of squares and SSE is the **sum of squares for errors.** *Formally,* SST and SSE are defined as follows:

$$SST = \sum (y_i - \bar{y})^2 \tag{4}$$

$$SSE = \sum (y_i - \hat{y}_i)^2 \tag{5}$$

where y_1 are the actual values of the dependent variable, and $\bar{y}$ is the average of the dependent variable and $\hat{y}_i$ are the estimated values from (1). The coefficient of determination, R^2, is always between 0 and 1. A R^2 value close to 1 indicates a good fit while a value close to 0 indicates a poor fit.

R^2 increases whenever we introduce a new independent variable into the regression. However, we need to avoid the temptation to add independent variables just to increase this value. In practice, you will find that a few independent variables will increase R^2 considerably with additional variables adding very little to the R^2 Most computer programs have a procedure for selecting the most important independent variables and omitting variables that have minimal contribution to the regression equation. This procedure is called **stepwise regression** because it introduces independent variables one by one according to their importance until the inclusion of more variables does not significantly improve the equation. Most computer programs for multiple regression include a procedure for stepwise regression.

For the regression with Boston residential fires, SST is 11,868.7 and SSE is 4,333.8 so that R^2 is as follows:

$$\begin{aligned} R^2 &= \frac{11{,}868.7 - 4{,}333.8}{11{,}868.7} \\ &= .63 \end{aligned} \tag{6}$$

This value indicates a fairly good fit, although it is not as large as we would like for this type of analysis.

Another measure of interest is the **standard error,** which we defined in Chapter 7 as indicating how far, on average, the estimates from the regression deviate from the actual numbers. For multiple regression, the equation for the standard error is as follows:

$$\text{Standard Error} = \sqrt{\frac{SSE}{n - k - 1}} \tag{7}$$

where n is the number of points and k is the number of independent variables. For our regression, we calculate the standard error as follows:

$$\text{Standard Error} \equiv \sqrt{\frac{4333.8}{147 - 4 - 1}} \qquad (8)$$
$$= 5.52$$

This standard error means that the estimates of fires will deviate, on average, by 5.52 from the actual number of fires for the census tracts.

In summary, the interesting feature of this regression is that four variables have been identified which can estimate the number of residential fires in a fairly accurate manner. The resulting regression equation could be employed by the Boston Fire Department for planning purposes. For example, other information in the city may be available on expected changes in the four independent variables. The department could therefore estimate its workload for these census tracts in future years.

> The **coefficient of determination,** R^2, determines how well a regression line fits the data points. R^2 is always between 0 and 1. Low values indicate a poor fit to the data while values close to 1 indicate a good fit to the data. The **standard error** indicates how far, on average, the estimates from the regression deviate from the actual values.

Collinearity Between Variables

A problem with multiple regression occurs when two independent variables are highly correlated. When this situation occurs, we need to select **one** of the variables to include in the regression. As an example, we present the data in Exhibit 8-3 from Prince William County, Virginia. The exhibit shows the total number of fires (residential and non-residential) in the county for 1981 through 1989. Two independent variables related to fires are shown in the last two columns. The column labeled "Residences" gives total number of residences (in thousands) and the last column labeled "Non-Residential Space" gives total square feet (in hundred thousands) of non-residential (retail, office, and industrial) space in the county. The data on residences and non-residential square footage are collected on an annual basis by the county.

Exhibit 8-4 shows the correlation matrix for these three variables. All correlations are high. The correlation between fires and residences is .90 and the correlation between fires and non-residential square footage is .91. The exhibit also shows a very high correlation of .98 between residences and non-residential square footage, which means that the pattern of annual increases is virtually the same for these two variables.

Exhibit 8-3 Fires in Prince William County, Virginia-1981-1989

Year	Total Fires	Residences	Non-residential Space
1981	3,313	47.91	106.98
1982	3,003	49.64	107.69
1983	2,938	50.93	108.51
1984	3,157	53.05	112.91
1985	3,631	54.81	125.10
1986	3,877	57.34	144.13
1987	3,761	60.57	162.57
1988	4,256	65.96	182.77
1989	4,156	68.65	206.03

Exhibit 8-4 Correlation Matrix, Prince William County, Virginia

Variable	Fires	Residences	Non-residential Space
Fires	1.00	.90	.91
Residences		1.00	.98
Non-residential Space	.91	.98	1.00

The temptation is to include both variables in a regression equation with fires as the dependent variable. However, we will now show that a regression with only one of the variables gives virtually the same fit to the data as a regression with both variables. Exhibit 8-5 shows the regression equations and coefficients of determination (R^2) for a regression with residences (TOTRES) as the only independent variable, non-residential space (TOTFT) as the only variable, and both variables in a regression.

Exhibit 8-5 Regression Equations for Fires in Prince William County

Variables in Regression	Regression Equation	R^2
TOTRES	Fires = 60.8 x TOTRES + 129.5	.816
TOTFT	Fires = 12.1 x TOTFT + lJ76.9	.829
Both	Fires = 13.9 x TOTRES + 9.4 x TOTFT + 1,469.9	.830

The three regression equations differ considerably because they include different independent variables. However, the coefficients of determination allow us to make comparisons. We have a R^2 value of .816 for the regression with only TOTRES and .829 for the regression with only TOTFT. With both variables in the regression, the coefficient of determination increases very minutely to .830. Consequently, the regression with only TOTFT gives almost exactly the same fit as including both variables. The inclusion of TOTRES with TOTFT contributes virtually nothing to the fit of the regression equation to the data.

The high correlation between TOTRES and TOTFT is the reason that we have such a small increase in the R^2 value when both variables are included. In summary, the best approach in this example is to select the regression equation with TOTFT to estimate fires rather than using the regression equation with both variables.

The term **multicollinearity** is used in most textbooks on statistics to refer to the situation in which high correlations exist between independent variables. In addition to problems with selection of variables, multicollinearity can result in instability of the estimated coefficients. The instability means that estimated coefficients may vary considerably from one sample to the next.

> Multicollinearity is the term used to indicate that two or more independent variables are highly correlated with each other. When multicollinearity occurs, we usually want to select one of the correlated variables to include in the multiple regression. The other correlated variables will not improve the multiple regression and may, in fact, result in unstable coefficients in the equation

Regression with Dummy Variables

The examples in this chapter have been based on continuous variables. The regression for residential fires in Boston included population, boarded-up housing units, single-parent families, and high-density housing units. In previous chapters, however, we have noted that categorical variables are important to fire departments because the 901 codes serve as the basis for completing reports on fires. In this section, we will present an approach that includes categorical variables in a regression.

The example we present is based on data collected during 1990 on Emergency Medical Services (EMS) calls in Prince William County, Virginia. We have selected travel time to EMS calls as our dependent variable. Our aim is to determine the effect of three categorical variables on travel time:

- Area of origin of EMS calls
- Whether delays occurred enroute to calls
- Type of call (Advanced Life support (ALS) or Basic Life Support (BLS) call).

For this analysis, three areas of the county were selected, which we will designate as Area A, Area B, and Area C: Each EMS report indicates the area of the call. In addition, the responding paramedic indicates whether delays were encountered while enroute to the scene. Delays may occur for several reasons, including traffic, weather, and incomplete address information.

Exhibit 8-6 shows average travel times for each variable. The overall average travel time for these 1,620 EMS calls was 6.71 minutes. As seen in the exhibit, the average travel times in the areas differ considerably with a low average time of 5.52 minutes in Area A followed by 7.99 minutes in Area C and 9.33 minutes in Area B. Area B and Area C had longer average travel times primarily because these areas arc larger than Area A. The exhibit also shows only a small difference between travel times to ALS and BLS calls. Travel times to BLS calls averaged 6.52 minutes compared to 7.11 minutes for ALS calls. Finally, the average trawl time to delayed calls was about one minute longer than calls not encountering delays (6.62 minutes compared to 7.69 minutes).

Exhibit 8-6 Average Travel Times-Prince William County-1990

Variable	Average Travel Time (Minutes	Number
Area A	5.52	955
Area B	9.33	218
Area C	7.99	447
ALS Calls	7.11	525
BLS Calls	6.52	1,095
No Delays	6.62	1,473
Delays	7.69	147
Overall	6.71	1,620

While Exhibit 8-6 provides useful information about travel times, it does not give any information on combinations of the variables. We do not know, for example, what travel time to expect in Area A for delayed ALS calls. We can employ regression analysis to provide estimates of travel times for the various combinations.

Since regression analysis requires numerical values, we need to assign numbers to these categorical variables. A convenient approach is to define new variables for each EMS record with appropriate numerical values depending on the type of call, area, and whether delays were encountered. For example, we define a variable called CALLTYPE to indicate the type of call. If an EMS record is for a BLS call, we assign a value of 0 to CALLTYPE. If it is an ALS call, we assign a value of 1 to CALLTYPE. In a similar manner, we define a variable called DEIAYS for each EMS incident. If delays were not encountered, we assign a value of 0 to DELAYS, and if delays were encountered, we assign a value of 1 to DELAYS.

Because we have three areas, we need a slightly different coding approach for them. We define two new variables called AREA1 and AREA2. If an EMS record is for a call in Area 1, we assign a value of 1 to AREA1 and a value of 0 to AREA2. Similarly, if the EMS call is from Area 2 we assign a value of 0 to AREA1 and a value of 1 to AREA2. Finally, if a call is from Area 3, we assign a value of 0 to both AREA1 and AREA2. Note that we would never assign a value of 1 to both AREA1 and AREA2.

In summary, we are defining four new variables in the following manner:

Variable	Definition
CALLTYPE	0 for ALS calls 1 for BLS calls
DELAYS	0 for calls not delayed 1 for delayed calls
AREA1	0 if call is not from Area 1 1 if call is from Area 1
AREA2	0 if call is not from Area 2 1 if call is from Area 2

The terms **dummy variable or indicator variable** are sometimes used to designate variables defined in this manner as either 0 or 1. A dummy variable indicates the presence or absence of a characteristic. The variable DELAYS indicates that a call was either delayed or not delayed; similarly, the variable CALLTYPE indicates that a call is either a BLS call or not a BLS call (that is, it is an ALS call). By assigning the values of 0 or 1, dummy variables transform categorical variables into meaningful numerical variables amenable to statistical analysis. In particular, we can perform a regression analysis with dummy variables, and the interpretation of the regression will be related to the definitions of these dummy variables.

For the regression, the dependent variable is travel time, and the independent variables are the four dummy variables. ‘The resulting regression equation for the data from Prince William County is as follows:

$$\begin{aligned}\text{Travel Time} = 5.27 + 1.38 \times \text{DELAYS} + .36 \times \text{CALLTYPE}\\ + 3.83 \times \text{AREA1} + 2.48 \times \text{AREA2}\end{aligned} \qquad (9)$$

The regression equation indicates that we start with a base travel time of 5.27 minutes and then increase the travel time based on delays, type of call, and area of origin. For example, suppose that we want to estimate the travel time for a delayed ALS call from Area 2. The values of the dummy variables for this example would be as follows:

DELAYS	1
CALLTYPE	0
AREA1	0
AREA2	1

We assign a value of 1 to DELAYS because we are assuming a delayed call. A value of 0 is assigned to CALLTYPE because we are assuming an ALS call. The variable AREA2 is assigned a 1 because we are assuming a call from AREA 2, which means that a value of 0 is assigned to the variable AREA 1

Inserting these values into the regression equation gives the following result:

$$\text{Travel Time} = 5.27 + 1.38 \times 1 + .36 \times 0 + 3.83 \times 0 + 2.48 \times 1$$
$$= 9.13 \text{ minutes}$$

Exhibit 8-7 shows a systematic way of estimating travel times which takes advantage of the fact that we are working with dummy variables. With the exhibit, we start with a travel time of 5.27 minutes and then make additions depending on the values assigned to the dummy variables.

Exhibit 8-7 Travel Times Estimates From Regression

1. Start with a travel time of 5.27 minutes.

2 If the call is delayed, add 1.38 minutes.
If the call is not delayed, add nothing.

3. If the call is a BLS call, add .36 minutes.
If the call is an ALS call, add nothing.

4. If the call is from Area 1, add 3.83 minutes.
If the call is from Area 2, add 2.48 minutes.
If the call is from Area 3, add nothing.

An interesting result can be derived from determining the range of travel times. The smallest travel time will occur for non-delayed BLS calls from Area 3. Calls with these characteristics are estimated to have an average travel time of 5.27 (since we are adding nothing to the base average). On the other hand, delayed BLS calls from Area 1 will have the highest average travel time, which is estimated to he 10.83 minutes.

This example shows that dummy variables enable us to include categorical variables in a regression. We always have one fewer dummy variable than the number of levels in a categorical variable. We needed only one dummy variable to indicate delayed/non-delayed calls since there are only two categories. We needed two dummy variables for areas since we had three areas. You should he aware that a regression with all dummy variables is equivalent to another area of statistical analysis called **analysis of variance.** Details on analysis of variance are beyond the scope of this handbook. The key point is that analysis of variance and dummy variable regression are equivalent statistical procedures.

Finally, it should be mentioned that we can combine continuous and dummy variables in a regression analysis. That is, it is not necessary to have either ail continuous variables or all dummy variables in a multiple regression. The combination of variables in sometimes called **analysis of covariance.**

> Dummy variables or indicator variables are variables that have values of either 0 or 1. A dummy variable indicates the presence or absence of a characteristic. Dummy variables provide a means of converting categorical variables to numerical variables amenable to statistical analysis. The number of dummy variables for a categorical variable is always one less than the number of levels of the category variable.

Summary

Multiple regression means that several independent variables are included in the regression analysis. The advantage of multiple regression is that it allows us to determine the impact of these independent variables on the dependent variable. The independent variables may he continuous or categorical. For categorical variables, we must develop dummy variables, which indicate the presence or absence of a characteristic, for the regression. As with single variable regression in Chapter 7, we can calculate a coefficient of determination and a standard error to determine how well our regression fits the actual data. The equations for these are very similar to their counterparts for single variable regression.

PROBLEMS

1. The regression equation for the Boston census tracts was follows:

$$\text{Fires} = 2.0 + .0017 \times \text{POPUI.ATION} + .068 \times \text{BOARDED} + .013 \times \text{FAMTYPE} + .060 \times \text{DENSITY} \tag{1}$$

Estimate the number of fires for each of the following census tracts and compare your result to the actual number of fires shown in the last column.

Census Tract	Population	Boarded Units	Single Parent families	Housing Density	Actual Fires
709	1386	19	184	100	15
812	980	148	396	180	30
902	664	59	261	104	30
907	1311	3	139	66	10

2. The multiple regression from Prince William County), for travel times to EMS calls is as follows:

$$\text{Travel Time} = 5.27 + 1.38 \times \text{DELAYS} + .36 \times \text{CALLTYPE} + 3.83 \times \text{AREA1} + 2.48 \times \text{AREA2} \tag{9}$$

Estimate the travel times for the following types of calls:

a. A non-delayed BLS call from Area 1

b. A delayed ALS call from Area 2.

c. A non-delayed ALS call from Area 3.

d. A delayed BLS call from Area 3.

This page intentionally left blank.

Chapter 9

QUEUEING ANALYSIS

Applications of Queueing Theory

Waiting lines have become an everyday part of our lives. Think for a moment about your activities during the past few days and how often you have waited in line for some type of service-at a post office, grocery store, theater, hank, airport, or cafeteria. The common feature of these diverse situations is that people arrive for a service that is unavailable because the providers are busy. Waiting lines present a dilemma for managers responsible for the delivery of service. For example, if a bank manager does not have enough tellers, the waiting line of customers may become quite long and they may eventually switch to another bank. On the other hand, if a manager has many tellers on duty, tellers will frequently be idle and bank costs will escalate.

Within many fire departments, we find a similar dilemma in determining how many EMS units we need to adequately handle requests from citizens for emergency medical services. A “waiting line” develops when all EMS units are busy and more citizens call with medical problems. The waiting line is actually in the communications center where the calls must be held until EMS units become available. If there are not enough units in the field, citizens will occasionally have to wait because everyone is busy. On the other hand, if the department fields a large number of EMS units, then costs will increase and units may not be very busy.

Queueing theory provides methods to analyze whether waiting lines will occur and what the consequences of waiting lines will be. It can be applied to diverse situations including the determination of the number of tellers for a bank, the number of ticket agents at an airport, and the number of ambulances for fire departments. In this chapter, we will apply a queueing model to determine the number of ambulances needed.

The following quote from an issue of the **jems** magazine reiterates the points we have been making and serves as an introduction to the topics covered in this chapter.”

> No one likes to wait in a line, and none of the systems with which we work would permit a lengthy wait to occur for a patient while the dispatcher searches for an ambulance to send on the critical call. On the other hand, no system can place an ambulance on every corner and call in a backup crew when each call is dispatched. Some where between having too many or too few ambu-

---

13. Barton. George K. “The Wait for an Ambulance.” jems. December, 1986

> lances a viable solution: queueing theory. This simple mathematical model or set of formulas will enable you to determine the number of ambulances needed, by hour of day and day of week, to meet calls for service in an efficient manner, and will provide objective information to modify previously committed resources.

As indicated in this quotation, queueing theory provides an analytical procedure for determining how many ambulances are needed by hour of day and day of week. In order to make these calculations, however, you must know how many EMS calls you expect and the average amount of time ambulance crews will take on calls. We can make these determinations using the techniques previous described in Chapter 4 of this handbook.

Queueing theory is the study of waiting lines and the consequences of these lines. It has been applied in a variety of situations where waiting lines occur. It is an excellent approach for fire departments in determining how many EMS units they need to field to handle calls from citizens for emergency medical services.

In summary, queueing theory can provide assistance in managing and operating the emergency medical services for a fire department. Potential applications include:

- Estimation of how busy EMS units will be based on an expected workload of EMS calls.
- Estimation of the probability that a citizen's call for medical services will have to wait because all EMS units are busy.
- Estimation of the average number of citizens waiting for medical service.
- Estimation of the average waiting time for these citizens.
- Estimation of the number of EMS units needed to satisfy objectives established by a fire department.

Examples of Results from Queueing Theory

Unit Utilization

As a starting point, suppose you have determined that during a particular time period (e.g., on Saturday evenings from 8 p.m. to midnight), citizen calls for EMS service arrive into the communications center at an average rate of two calls per hour. Assume that 40 minutes is the mean time for a call. The time required for a call starts with the dispatch of an EMS unit to the scene and ends when the unit is available for another call.

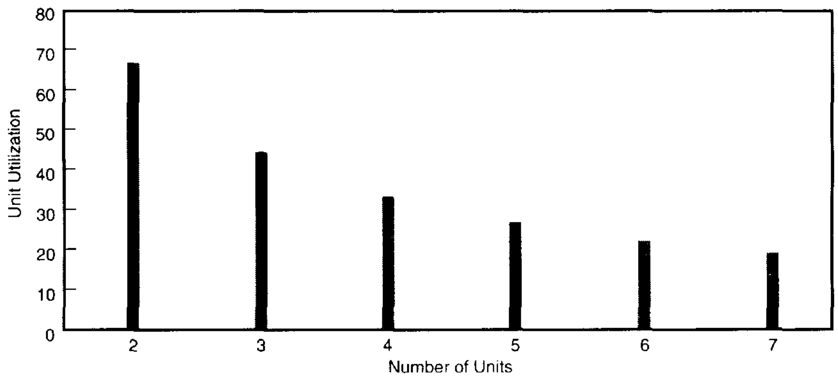
Suppose further that two EMS units are available in the field to respond to citizen calls. As a starting point for our analysis, we can calculate how busy these two units will be. They provide 120 minutes of unit time each hour (2 units times 60 minutes). The calls require 80 minutes each hour (2 calls times 40 minutes each). Thus, each unit will be busy about 66.7 percent of the time (80 minutes of work divided by 120 minutes of unit time). This percentage is called "unit utilization" since it measures the percent of time each unit will be busy during a given time period.

If three units are fielded, unit utilization drops to 44.4 percent, as seen by the following calculation:

$$\text{Unit Utilization} = \frac{2 \times 40}{3 \times 60} = 44.4\%$$

In this equation, the numerator gives the amount of work we expect and the denominator gives the amount of available unit time. You can use this equation to verify that with four units, the unit utilization drops to 33.3 percent, and with five units, it is 26.7 percent. Exhibit 9-1 is a graph of unit utilization for this example. Of course, unit utilization can never equal zero, but it continually decreases as we add more units in the field.

Exhibit 9-1 Unit Utilization



Note: This exhibit assumes an average of 2 calls per hour with each call requiring 40 minutes.

We can state a general formula for unit utilization as follows. Assume that we expect citizen calls for EMS service each hour and that calls average t minutes from time of dispatch to time of completion. If we field n units, the unit utilization is expressed by:

$$\boxed{\textbf{Unit Utilization} = \frac{ct}{60n}}$$

In practice, we can expect considerable variation from what has just been described. Some evenings will be busier than others and some calls will require less time than average, while other calls will take considerably longer. The point is that we will not have exactly the same unit utilization every day; it will fluctuate depending on the actual workload experienced. When we calculate unit utilization, we are looking at what happens “on average.” While we expect fluctuations, we are primarily concerned with overall average performance of our system.

Other Queueing Calculations

Unit utilization is an example **of a system performance measure.** It measures how the EMS system will perform under expected workload conditions. We are not evaluating individual units and we arc not trying to determine the unit utilization of individual units. Instead, we are considering the system as a whole and making estimates on what we expect to occur.

By applying queueing theory, we can estimate several other system performance measures. Three of the most common are:

- The probability that a person calling for service will have to wait.
- The average number of citizens waiting for service.
- The average waiting time for these citizens.

The first measure estimates the probability that all EMS units will be busy and another citizen will call for EMS service. When this situation occurs, the communications center must hold the call until an EMS unit becomes available. The second performance measure estimates how many citizens will be waiting, on average, for the dispatch of an ambulance. Finally, the third performance measure gives the average time these waiting citizens will have to wait until a unit is dispatched.

These performance measures behave in the same manner as unit utilization. For example, the probability of a delay decreases as the number of EMS units increases. Similarly, the number of citizens waiting for service decreases and the amount of waiting time decreases as the number of units increases.

The reason we have these performance measures is because of the inevitable variation in workload. If two units were fielded, no waiting lines would occur if we were assured that we would always have exactly two calls per hour requiring 40 minutes each. In reality, however, workload varies, and waiting lines develop because of these variations. Since queueing models assume considerable variability, they provide excellent estimates on the length of waiting lines and other performance measures.

The formulas for the probability. of delay and the length of a waiting line arc complicated. Interested readers arc invited to study the next section where we provide explanations for both calculations. Appendices B and C have been included in this handbook to ease the calculation burden. They provide a quick way to obtain the probability of delay and length of a waiting line without having to perform many calculations. The tradeoff is that the results may not be quite as precise as the actual formulas because of the rounding of numbers we have to do prior to access into the tables. For most applications, the errors will be small. However, we strongly suggest that you try to use the actual formulas presented later in this chapter if you want to be precise in your analysis.

The key to the tables in Appendices B and C is the following calculation with the average number of calls and average time per call.

$$\textbf{Table Key} = \frac{ct}{60}$$

where c is the number of calls per hour and *t* is the average time per call. Note that division by 60 converts the average time to hours for convenience with the tables. The Table Key is actually the amount of work we expect each hour.

In our example, *c* is 2 calls per hour and *t* is 40 minutes. The Table Key is then 1.33 (2 x 40 divided by 60). We round to 1.3 since the keys in the appendices are to one decimal place. Now turn to Appendix B to find the probability that a call will be delayed. The number of units is shown across the top of the table and the Table Key is shown down the left column. Move down the left column to the line where 1.3 appears as the key. You can now read across this row to obtain the probability of a delay. With two units, the probability is 51.2 percent; with three units, the probability is 17.0 percent, etc.

To determine the average number of citizens waiting for an EMS unit, we use the table in Appendix C in the same manner. The number of units is displayed across the top and the Table Key is down the left column. We go to the line where 1.3 appears as the key and obtain the average number of waiting calls by moving across the row. With two units, we estimate .95 citizens waiting; with three units, only .13 citizens waiting; and with four units, only .02 citizens waiting.

Finally, we want to determine the amount of time that waiting citizens will have to wait before a dispatch occurs. Queueing theory informs us that to determine the average waiting time, we simply divide the entry from Appendix C by the average number of calls per hour:

$$\textbf{Waiting Time} = \frac{\text{Appendix C Entry x 60}}{c}$$

Note that we have multiplied by 60 so that our final answer will be in minutes rather than hours. For example, with two units, we have determined that .95 citizens will be waiting. Since we have 2 calls per hour, the average waiting time will be 28.5 minutes (.95 times 60 divided by 2).

The top of Exhibit 9-2 summarizes the results of our calculations. The averages of 2 calls per hour and 40 minutes per call are based on data from Prince William County, Virginia, for l990. By way of comparison, the bottom portion of the exhibit shows how performance measures change when the average number of calls increases to 2.5 calls per hour. They show dramatic increases. With three units, for example, unit utilization changes from 44.4 percent to 55.6 percent, and the probability of a delay goes from 17.0 to 30.0. On the basis of this exhibit, the county might decide to increase the number of units because of the changes in performance measures caused by the increased workload.

Exhibit 9-2 Summary of Performance Measures

Assuming 2.0 Calls Per Hour and 40 Minutes Per Call

	Number of Units			
Performance Measure	2	3	4	5
Unit Utilization	66.7 %	44.4 %	33.3 %	26.7 %
Probability of a Delay	51.2	17.0	4.8	1.1
Average Number of Waiting Calls	0.95	0.13	0.02	0.005
Average Waiting Time for These Calls (Minutes)	28.5	3.9	0.6	0.2

Assuming 2.5 Calls Per Hour and 40 Minutes Per Call

	Number of Units			
Performance Measure	2	3	4	5
Unit Utilization	83.3 %	55.6 %	41.7 %	33.3 %
Probability of a Delay	75.8	30.0	10.3	3.0
Average Number of Waiting Calls	3.79	0.37	0.07	0.015
Average Waiting Time for These Calls (Minutes)	90.9	9.0	1.8	0.4

Note. For the bottom portion of this exhibit, exact calculations with the queueing theory formulas were used

Determining the Number of EMS Units

The prior analysis presented queueing theory as a descriptive tool. That is, we assumed a fixed number of EMS unit in the field and estimated the performance measures. We can also employ queueing theory in a **prescriptive** manner to determine how many units will be required to achieve predetermined performance measures. That is, we establish objectives for performance and determine how many units will be needed to achieve the objectives. This approach is similar to a management-by-objectives approach in which objectives determine performance standards.

Suppose, for example, that your objectives are to field enough EMS units to insure that:

- Units are busy on EMS calls no more than 40 percent of their shift.
- The probability of a delay is 5 percent or less.

What we want to determine is the number of units needed to obtain these performance standards.

For purposes of illustration, assume that we expect 3.2 calls per hour averaging 37 minutes each. The first objective is on unit utilization. We can use our equation for unit utilization and express it in terms of units needed:

$$\textbf{Units Needed} = \frac{ct}{60 \times \text{Unit Utilization}}$$

Since our objective is 40 percent for unit utilization, we obtain the number of units needed as follows:

$$\text{Unit Needed} = \frac{3.2 \times 37}{60 \times .40} = 4.93$$

We round this result to 5 units since we cannot have a fractional unit. This calculation means that five units will achieve our objective of no more than 40 percent of their time devoted to EMS calls. The result assumes, of course, that the jurisdiction will continue to average 3.2 EMS calls per hour and 37 minutes per call.

To estimate the number of units needed for the second objective, we reference the table in Appendix A. The objective is that the probability of a delay will be 5 percent or less. To access the table, we need the Table Key, which calculates to 1.97 (3.2 times 37 divided by 60). We round this result to 2.0 in order to access the table. We now look at the row for 2.0 in Appendix A, which is shown as Exhibit 9-3 on the following page.

This table indicates that six units will be needed to achieve our objective that the probability of delay will not exceed 5 percent. With three units or four units, the probability of delay is higher than we are willing to tolerate at 44.4 percent and 17.4 percent, respectively. With five units, the probability of delay will be only 6.0 percent, which is close to our objective but does not achieve it. We therefore select six units to he assured of satisfying our objective.

Exhibit 9-3 Probability of Delay

Table Key = 2.0 (3.2 calls averaging 37 minutes each)

Number of Units	Probability of Delay
3	44.4
4	17.4
5	6.0
6	1.8
7	.01

We have now determined how many units are needed for each objective, but we still have one remaining step. We need five units to achieve the first objective, and we need six units to achieve the second objective. To achieve **both** objectives, we select the **maximum** of these two numbers. In other words, we select six units as the final answer. The reason is that six units will satisfy both objectives. It satisfies the second objective because of our manner of selection from the table, and it satisfies the first objective because adding more units decreases unit utilization.

As a final note, it should be mentioned that there are many other queueing models from which to choose. We have selected this particular model because it is one of the most frequently applied and has proven beneficial in many queueing problems similar to what fire departments encounter.

Queueing models can he expanded to cover other situations. For example, we can include call priorities to take into account that some EMS calls are more serious than others. If citizen calls are waiting in the communications center, the more serious calls, such as heart attacks, have higher priority than minor calls, such as bruises. Queueing models exist that include priority systems. They can be applied in the same manner as our basic model. In addition, other queueing models exist that take geography into consideration. These models are more complicated because they require the user to describe the geography of your jurisdiction. However, they may be beneficial to you because they may reflect more accurately the problems of locating EMS units.

Queueing Calculations*

In the literature on queueing theory, the system just described is called a multiple-server queueing model in which call arrivals follow a Poisson distribution and service times follow an exponential distribution. In this section, we will describe the Poisson and exponential distributions and we will give formulas for the queueing calculations needed to derive Appendices A and B.

The Poisson distribution is a probability distribution which has been applied in many diverse situations. To see its application to EMS calls, we will first develop a frequency distribution on the number of calls arriving into a communications center. The frequency distribution will then be approximated by the Poisson distribution.

During some hours, no calls will come into the communications center. In other hours, we will have only one EMS call, and other hours will have more than one EMS call. Over a large number of hours, we can build a frequency table showing the number of hours with zero calls, the number with only one call, the number with two calls, etc.

As an example, Exhibit 9-4 gives information on the number of calls per hour in Prince William County, Virginia, for the hours from 3 p.m. to 7 p.m., January l-May 31, 1990 (a total of 152 days). Since we are looking at four hours during each of these days, we have 608 data points (4 times 152). The exhibit shows that there were 54 hours during which no citizens called for EMS service. For 105 hours, exactly one EMS call came into the communications center; for 123 hours, exactly two EMS calls came into the communications center, and so on. The third column of the exhibit converts the frequencies into percentages. This column shows, for example, that during 14.0 percent of the hours, 4 EMS calls arrived into the communications center.

EMS Calls	Number of Hours	Percent	Poisson
0	54	8.9	5.2
1	105	17.3	15.3
2	123	20.2	22.7
3	109	17.9	22.4
4	65	14.0	16.6
5	51	8.4	9.8
6	37	6.1	4.9
>6	44	7.2	3.2
Total	608	100.0	100.0

Note: Calls are from Januay - May, 1990-3 p.m. to 7 p.m.

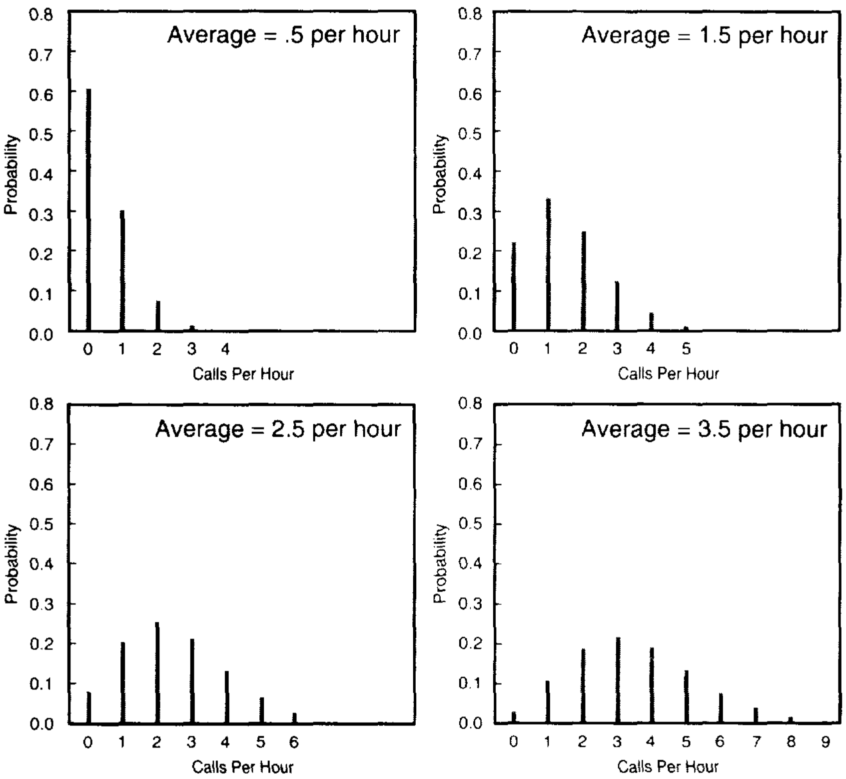
From Exhibit 9-4, we can calculate that the average number of EMS calls per hour is 2.97. We can then develop a Poisson distribution of the expected percentage of calls per hour. The right column of Exhibit 9-4 shows expected frequencies according to a Poisson distribution (we will give the calculations for this column in the following paragraphs). This column can be compared to the actual percentage from our sample. While the expected percentages are not exactly the same, the differences are not large. For example, the Poisson distribution shows 15.3 percent of the hours with exactly one EMS call per hour, as compared to actual experience of 17.3 percent.

The general equation for a Poisson distribution is as follows:

$$P(x=k)=\frac{e^{-c}c^{k}}{k!}$$

where c is the average number of calls per hour, and k is an integer value starting with zero.

Exhibit 9-5 Poisson Distribution



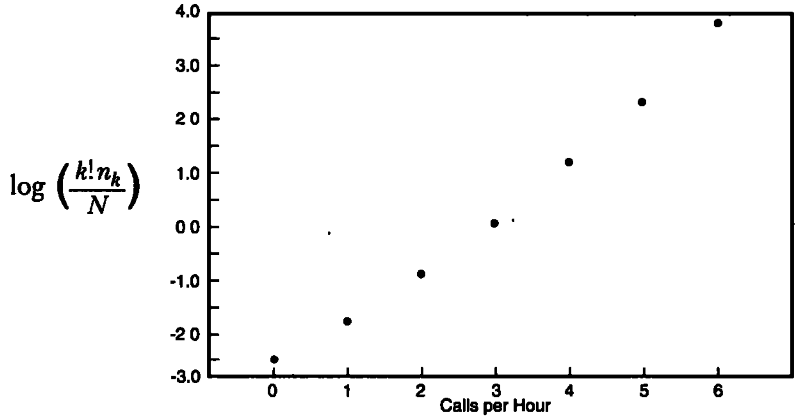
In our example, we have an average of 2.97 calls per hour. To determine the probability of four calls per hour, we compute as seen on the following page.

$$P\ (x=4) = \frac{e^{-2.97} \text{ x } 2.97^4}{4!} = \frac{77.81 \text{ x } .0513}{24} = .166$$

Exhibit 9-5 shows several theoretical Poisson distributions. The shape of the distribution depends on the average. When the average is less than 1.0, the Poisson distribution continually decreases; if the average is greater than 1.0 the distribution increases to the integer portion of the average and then decreases.

The results on queueing theory presented in this chapter are correct provided the arrival of EMS calls into the Communications Center approximately follow a Poisson distribution. Hoaglin, Mosteller, and Tukey (1985) provide a relatively simple graphical technique for determining whether the Poisson distribution is a good selection. In their approach, k represents the number of possible EMS calls in an hour (0, 1, 2, etc.), and n_k represents the observed number of hours with *k* calls. If calls follow a Poisson distribution, you should obtain a straight line by plotting k against the quantity log *(k! nkl* N), where N is the total number of hours (608 in our example). Exhibit 9-6 shows the plot obtained through this procedure. You will note that the dots approximate a straight line, which indicates a relatively good fit between our actual data and what would be expected with a Poisson distribution.

Exhibit 9-6 Plot for Poisson Fit



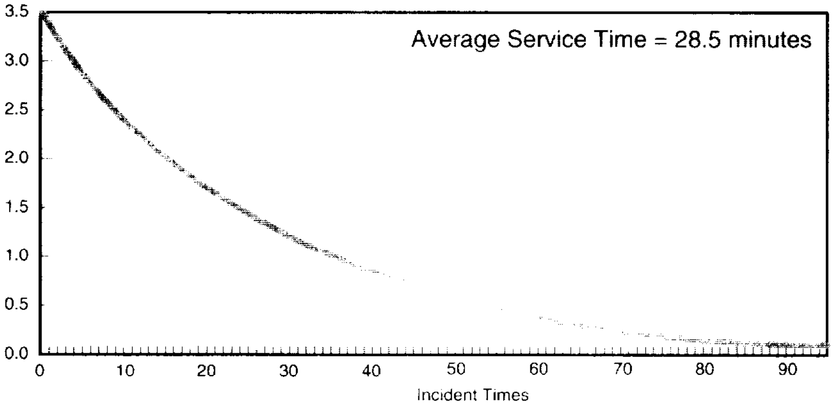
As previously indicated, the queueing model also assumes that the service times follow an exponential distribution. The exponential distribution is given by:

$$\boxed{f(x) = ue^{-ux}}$$

where x represents a particular service time and u is the **service rate.** Suppose, for example, that the average service time for EMS calls is 28.5 minutes, or .475 hours. The service rate is the inverse of the service time; that is, it is 1 divided by the service time. The service rate is therefore 2.11

Exhibit 9-7 shows the exponential distribution assuming a service rate of 2.11. The key feature of the exhibit is that the curve continually decreases. The decrease means that there is a low probability of long service times.

Exhibit 9-7 Exponential Distribution



With this background, we can now provide the equations for the figures appearing in Appendices A and B. We will not derive these equations, since the derivation assumes a considerable amount of background knowledge about queueing theory. However, the equations are of value if you want to perform more exact calculations than provided in the Appendices. While the equations appear complicated, they are fairly easy to develop on most microcomputer spreadsheet program.

In the following equations, we define $r = c/nu$. We assume that $r < 1$, so that the arrival rate does not exceed the maximum service rate. We first calculate P_o, which is the probability that all units are busy:

$$\boxed{P_0 = 1 \Big/ \sum \frac{\left(\frac{c}{u}\right)^k}{k!} + \frac{\left(\frac{c}{u}\right)^n}{n!} \frac{1}{(1-r)}}$$

Then the probability of a delay is expressed as:

$$\boxed{\text{Probability of Delay} = \frac{P_0 \left(\frac{c}{u}\right)^n}{n!\,(1-r)}}$$

Finally, the equations for the queue length, L_q, and for waiting time in the queue, Wq, are given by:

$$\boxed{L_q = \frac{P_0 \left(\frac{c}{u}\right)^n r}{n!\,(1-r)^2}} \qquad \boxed{W_q = \frac{L_q}{c}}$$

Summary

Queueing theory is the analysis of waiting lines and the consequences of these waiting lines. It can be a powerful management tool to address issues of unit utilization, waiting time, call priorities, and other important considerations in your EMS delivery system. In this chapter, we have applied queueing theory to the problem of determining the number of EMS units needed in a fire department. For this determination, we must first know how many EMS calls we expect and what the average time per EMS call will be. We obtain these averages from prior experience using the techniques described in Chapter 4 of this handbook. It is also necessary to establish objectives for the EMS delivery system. We can say, for example, that we want to have enough EMS units so that they devote no more than 50 percent of their time to EMS calls. Other objectives could be established on the probability of a call delay, the average number of waiting calls, and the average that calls will have to wait. We can then apply a queueing model to determine how many units will be needed to achieve these objectives.

This page intentionally left blank.

PROBLEMS

1. Suppose that your department averages 3.5 EMS calls per hour with an average of 45 minutes for each call.

a. Calculate unit utilization assuming you have 3, 4, 5 and 6 units fielded.

b. How much does unit utilization change from 4 units to 5 units? from 5 units to 6 units?

2. For the average of 3.5 calls per hour and 45 minutes per call, use Appendices A and B to answer the following questions assuming 3, 4, 5, and 6 units.

a. What is the probability, of delay?

b. What will be the average number of citizens waiting for service?

c. What is the average waiting time of citizens whose call has been delayed?

3. Suppose that EMS calls have been increasing 6 percent per year and that next year you also expect that calls will average 50 minutes per call, rather than 45 minutes per call.

a. How many calls per hour will you have based on a 6 percent increase assuming a current average of 3.5 calls per hour.

b. Determine unit utilization, probability of delay, and average number of people waiting assuming 4 units in the field.

c. Compare these results to your answers in question 2 for 4 units.

4. How would you determine the actual value of your performance measures in your department today

5. What process would you use to set objectives in your jurisdiction on performance measures?

6. Suppose you expect 4 EMS calls per hour next month with an average of 38 minutes per call. Suppose further that your jurisdiction has set the following objectives:

- Unit utilization not more than 60 percent
- Probability of a delay not more than 3 percent

- Average number of citizens waiting not to exceed .25

a. Determine how many units will he needed to satisfy each of these objectives.

b. How many units will be needed to satisfy all three objectives?

7. Suppose the jurisdiction decides to add an objective that the waiting time should not exceed I5 minutes.

a. How many units will he needed to satisfy this objective?

h. Does this change the total number of units needed to meet all four objectives?

8. In Prince William County, Virginia, the distribution of EMS calls coming into the communications center between 7 a.m.-11 a.m., January-May 1990, was as follows:

EMS Calls	Number of Hours	Percent
0	158	26.0
1	192	3 1.6
2	144	23.7
3	68	11.2
4	34	5.6
5	10	1.6
6	2	0.3
Total	608	100.0

From this distribution, we can calculate an average of 1.46 calls per hour.

a. with the average of 1.46 calls, determine the expected percent of EMS calls under a Poisson distribution.

b. Develop a plot for Poisson tit similar to Exhibit 9-6 to determine whether the Poisson distribution gives a good approximation to the experienced distribution.

c. What is your conclusion?

9. For an exponential distribution, the cumulative distribution is given by the following relationship:

$$F(x) = 1 - e^{-xu}$$

where u is the service rate in our application.

(continued from question 9)

Suppose that the average service time is 28.5 minutes, which means that the service rate is 2.11.

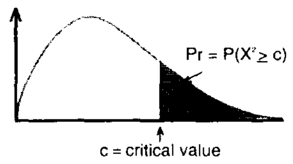
a. Develop a graph of the cumulative distribution assuming this service rate of 2.11.

b. From the graph, estimate the 25th percentile, median, and 75th percentile.

Appendix A

CRITICAL VALUES FOR THE CHI-SQUARED DISTRIBUTION

Degrees of Freedom	Critical Value	Degrees of Freedom	Critical Value
v	.050	v	.050
1	3.841	79	100.7
2	5.991	80	101.9
3	7.815	81	103.0
4	9.488	82	104.1
5	11.07	83	105.3
6	12.59	84	106.4
7	14.07	85	107.5
8	15.51	86	108.6
9	16.92	87	109.8
10	18.31	88	110.9
11	19.68	89	112.0
12	21.03	90	113.1
13	22.36	91	114.3
14	23.68	92	115.4
15	25.00	93	116.5
16	26.30	94	117.6
17	27.59	95	118.8
18	28.87	96	119.9
19	30.14	97	121.0
20	31.41	98	122.1
21	33.67	99	123.2
22	33.92	100	124.3
23	35.17		
24	36.42		
25	37.65		
26	38.88		
27	40.11		
28	41.34		
29	42.56		
30	43.77		
31	44.98		
32	46.19		
33	47.40		
34	48.60		
35	49.80		
36	51.00		
37	52.19		
38	53.38		
39	54.57		
40	55.76		
41	56.94		
42	58.12		
43	59.30		
44	60.48		
45	61.66		
46	62.83		
47	64.00		
48	65.17		
49	66.34		
50	67.50		
75	96.22		
76	97.35		
77	98.48		
78	99.62		



This page intentionally left blank.

Appendix B

PROBABILITY OF DELAY

C/S	Number of Units					
		3	4	5	6	7
1	33.3	9.1	2.0	0.4	0.1	0.0
1.1	19.0	11.5	2.8	0.6	0.1	0.0
1.2	45.0	14.1	3.7	0.8	0.2	0.0
1.3	i1.?	17.0	4.8	1.1	0.2	0.0
1.4	57.6	20.2	6.0	1.5	0.3	0.1
1.5	64.3	23.7	7.5	2.0	0.5	0.1
1.6	71.7	27.4	9.1	2.6	0.6	0.1
1.7	78.1	31.3	10.9	3.3	0.9	0.2
1.8	85.3	35.5	12.9	4.0	1.1	0.3
1.9	92.6	39.9	15.0	4.9	1.4	0.4
2		44.4	17.4	6.0	1.8	0.5
2.1		49.2	19.9	7.1	2.2	0.6
2.2		54.2	22.7	8.4	2.7	0.8
2.3		59.4	25.6	9.8	3.3	1.0
2.4		64.7	28.7	11.4	4.0	1.3
2.5		70.2	32.0	13.0	4.7	1.5
2.6		75.9	35.4	14.9	5.6	1.9
2.7		81.7	39.1	16.8	6.5	2.3
2.8		87.7	42.9	19.0	7.5	2.7
2.9		93.8	46.8	21.2	8.7	3.2
3			50.9	23.6	9.9	3.8
1.1			55.2	26.2	11.3	4.4
3.2			59.6	28.9	12.7	5.1
3.3			64.2	31.7	14.3	5.9
3.4			68.9	34.7	16.0	6.7
3.5			73.8	378	17.7	7.6
3.6			78.8	41.0	19.7	8.6
1.7			83.9	44.4	21.7	9.7
3.8			89.1	48.0	23.8	10.9
3.9			94.5	51.6	26. I	12.2
4				55.4	28.5	13.5
4.1				59.3	31.0	15.0
1.2				63.4	33.6	16.5
1.3				67.5	36.3	18.1
4.4				71.8	39.2	19.9
4.5				76.2	42.2	21.7
4.6				80.8	45.3	23.7
4.7				85.4	48.5	25.7
4.8				90.2	51.8	27.8
4.9				95.0	55.2	30.1
5					58.8	32.4
5.1					62.4	34.9
5.2					66.2	37.4
5.3					70.0	40.0
5.4					74.0	42.8
5.5					78.1	45.6
5.6					82.3	48.6
5.7					86.6	51.6
5.8					90.9	54.8
5.9					95.4	58.0
6						61.4
5.1						64.8
5.2						68.4
6.3						72.0
6.4						75.7
6.5						79.5

Appendix C

QUEUE LENGTH (Lq)

C/S	Number of Units					
	2	3	4	5	6	7
1	0.33	0.05	0.01	0.00	0.00	0.00
1.1	0.48	0.07	0.01	0.00	0.00	0.00
1.2	0.68	0.09	0.02	0.00	0.00	0.00
1.3	0.95	0.13	0.02	0.00	0.00	0.00
1.4	1.35	0.18	0.03	0.01	0.00	0.00
1.5	1.03	0.24	0.04	0.01	0.00	0.00
1.6	2.84	0.31	0.06	0.01	0.00	0.00
1.7	4.43	0.41	0.08	0.02	0.00	0.00
1.8	7.67	0.53	0.11	0.02	0.00	0.00
1.9	17.59	0.69	0.14	0.03	0.01	0.00
2		0.89	0.17	0.04	0.01	0.00
2.1		1.15	0.22	0.05	0.01	0.00
2.2		1.49	0.28	0.07	0.02	0.00
2.3		1.95	0.35	0.08	0.02	0.00
2.4		2.59	0.43	0.10	0.03	0.01
2.5		3.51	0.53	0.13	0.03	0.01
2.6		4.93	0.66	0.16	0.04	0.01
2.7		7.35	0.81	0.20	0.05	0.01
2.8		12.27	1.00	0.24	0.07	0.02
2.9		27.19	1.23	0.29	0.08	0.02
3			1.53	0.35	0.10	0.03
3.1			1.90	0.43	0.12	0.03
3.2			2.39	0.51	0.15	0.04
3.3			3.03	0.62	0.17	0.05
3.4			3.91	0.74	0.21	0.06
3.5			5.17	0.88	0.25	0.08
3.6			7.09	1.06	0.29	0.09
3.7			10.35	1.26	0.35	0.11
3.8			16.94	1.52	0.41	0.13
3.9			36.86	1.83	0.48	0.15
4				2.22	0.57	0.18
4.1				2.70	0.67	0.21
4.2				3.33	0.78	0.25
4.3				4.15	0.92	0.29
4.4				5.27	1.08	0.34
4.5				6.86	1.26	0.39
4.6				9.29	1.49	0.45
4.7				13.38	1.75	0.53
4.8				21.64	2.07	0.61
4.9				46.57	2.46	0.70
5					2.94	0.81
5.1					3.54	0.94
5.2					4.30	1.08
5.3					5.30	1.25
5.4					6.66	1.44
5.5					8.59	1.67
5.6					11.52	1.94
5.7					16.45	2.26
5.8					26.37	2.65
5.9					56.30	3.11
6						3.68
6.1						4.39
6.2						5.30
6.3						6.48
6.4						8.08
6.5						10.34

This page intentionally left blank.

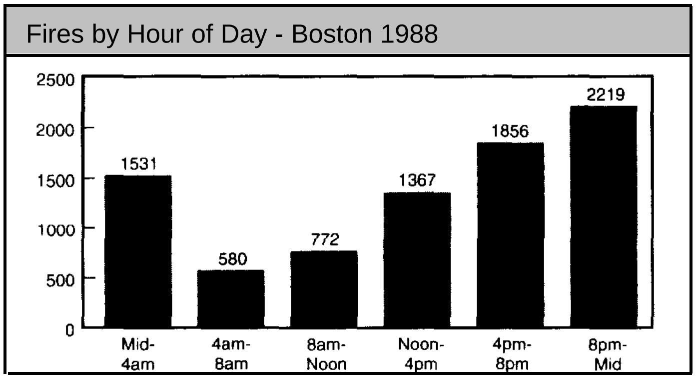
Appendix D

ANSWERS TO PROBLEMS[14]

Chapter 2

1. Fires by Hour of Day - Boston 1988

Midnight - 4 a.m.	1531
4 a.m. - 8 a.m.	580
8 a.m. - Noon	772
Noon - 4 p.m.	1367
4 p.m. - 8 p.m.	1856
8 p.m. - Midnight	2219



This histogram can be used to easily distinguish the morning hours from the afternoon and evening hours. The histogram in Exhibit 2-2 would take a little more concentration. This histogram would be useful when deciding shift assignments. The histogram in Exhibit 2-2 quickly shows which hour of day is the busiest and which hour is the least busiest. Remember, you always lose information when you aggregate data.

2. Exhibit 2-2 indicates that we should schedule more firefighters in the afternoon and evening hours and fewer firefighters in the early morning hours.

Exhibit 2-3 indicates that we need more firefighters on the weekend and not as many during the middle of the week.

---

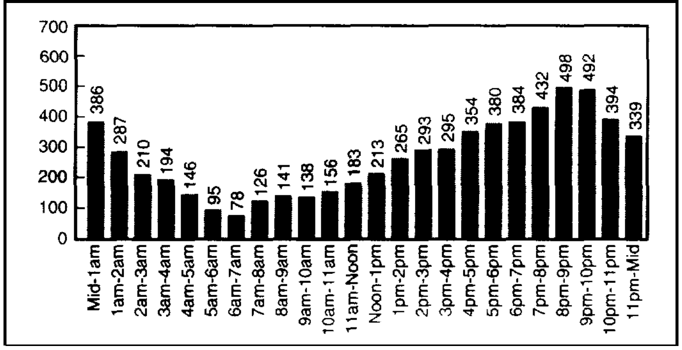
[14] Please note, the answers to the statistical problems were calculated by a scientific calculator. Your answers might not exactly match the answers that we have provided.

Exhibit 2-4 indicates that we should schedule more firefighters in the summer months (June and July). We do not need as many in January and February.

3. The distribution of fires by hour of day is almost identical to the 1988 data. There are more fires in the afternoon and evening hours, and very few fires in the early morning hours.

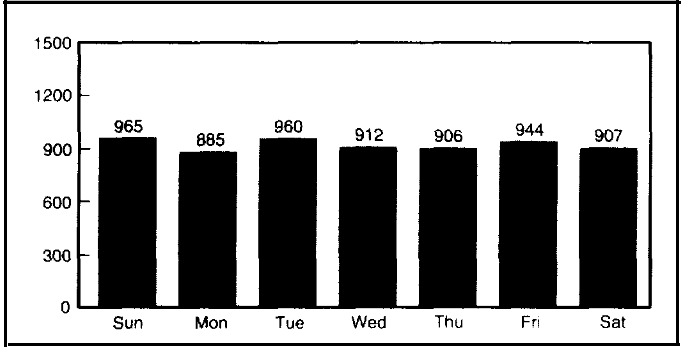
8-9pm was the busiest hour in 1990 and 9-10pm was the busiest hour in 1988.

Fires by Hour of Day – Boston 1990



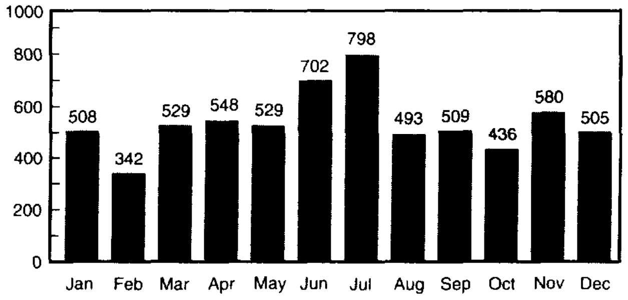
Like Exhibit 2-3, more firefighters would be needed on the weekend. Tuesdays have been a little busier in 1990 than in 1988. Sunday was the busiest in 1990, but Saturday was the busiest in 1988.

Fires by Day of Week-Boston 1990



The 1990 data also indicates that there are more fires in the summer months (June and July). July was the busiest in I990 and June was the busiest in 1988.

Fires by Month-Boston 1990



4. a. Jersey City, New Jersey, 1988: Ages of Civilian Casualties

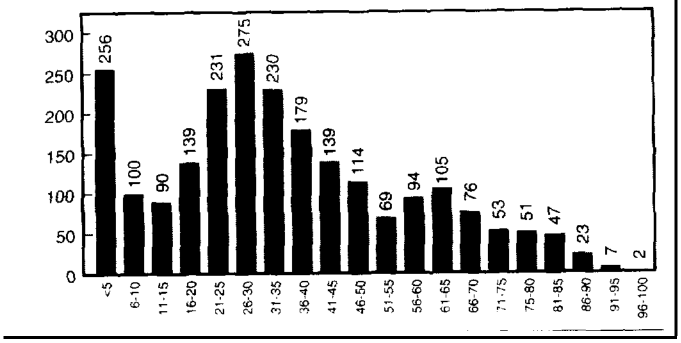
Age Group	Frequency	Cumulative Frequency	Cumulative Percent
1-5	14	14	16.9
6-10	12	26	31.3
11-15	7	33	39.7
16-20	1	34	41.0
21-25	7	41	49.4
26-30	10	51	61.4
31-35	5	56	67.5
36-40	4	60	72.3
41-45	4	64	77.1
46-50	2	66	79.5
51-55	3	69	83.1
56-60	4	73	87.9
61-65	2	75	90.4
66-70	3	78	94.0
71-75	3	81	97.6
76-80	1	82	98.8
81-85	0	82	98.8
86-90	0	82	98.8
91-95	1	83	100.0
Total	83		

b. 39.7% of the civilian casualties were under 16 years of age.

c. 20.5% of the civilian casualties were over 50 years old.
(100.0 - 79.5=20.5)

5. The spikes in this histogram occur at <5 and 26-30. Exhibit 2-5 also has spikes at these age groups.

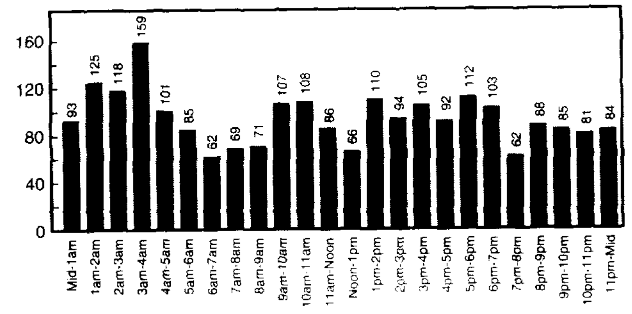
Civilian Casualities-1998



There is no hole or outlier in this group like there was in the Jersey City data.

6. There is no real pattern in this chart. The only thing you can tell is that there are more people killed or injured in the early morning hours.

Casualties by Hour of Day – 1988



7. Peaks cause a cumulative distribution to increase.

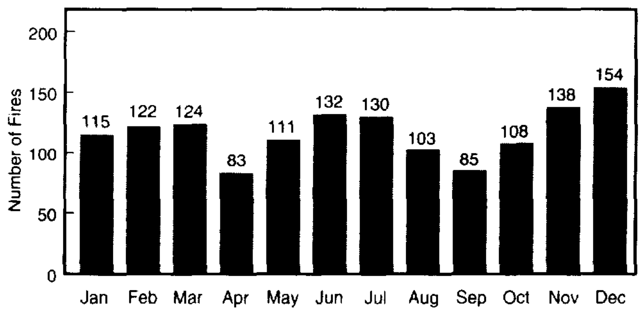
Holes cause a cumulative distribution to increase slightly.

Spikes cause a cumulative distribution to increase sharply.

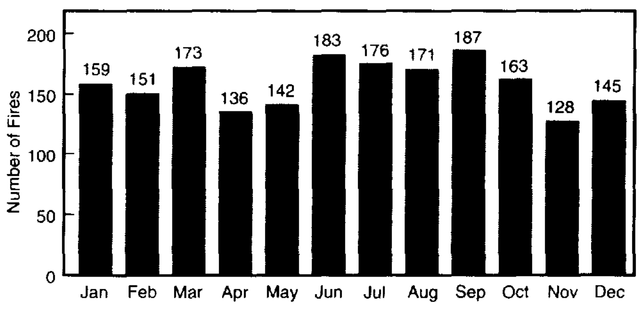
Gaps do not make the cumulative distributions increase. A cumulative distribution is always flat when a gap occurs.

1.

Structure Fires by Months – Los Angels 1990



Vehicle Fires by Months – Los Angels 1990



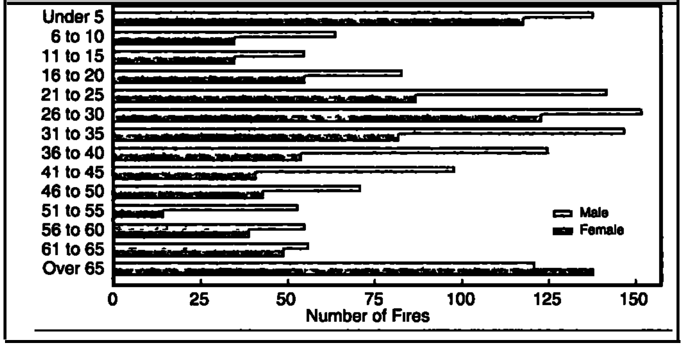
Overall there are more vehicle fires than structure fires. June and September were the months with the highest number of vehicle fires. December had the most structure fires. April and September had the lowest number of structure fires.

2. There are more male casualties than female casualties in all age groups except Over 65.

a. No, because there are too many age groups. You should only use a pie chart to show how components relate to the whole and there should be no more than six components.

b. By separating the data by sex, you can compare female and male casualties for each age group.

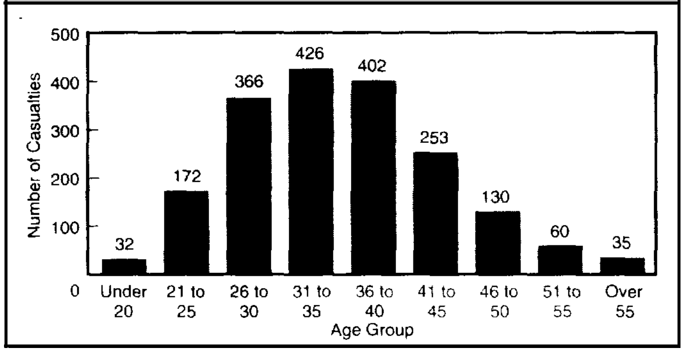
Male and Female Civilian Casualities-1989



3. a. There are more firefighter casualties in the age groups of 26-40. This is probably because the majority of the personnel is in these age groups.

There are very few casualties under the age of 20 and over the age of 55.

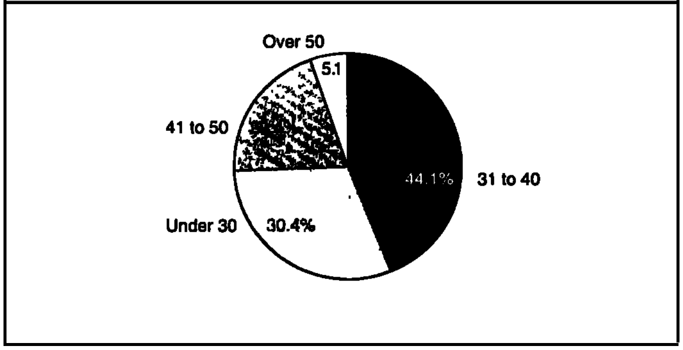
Firefighter Casualties – 1989



b.

Age Group	Number	Percent
Under 30	570	30.4%
31 to 40	828	44.1
41 to 50	383	20.4
Over 51	95	5.1

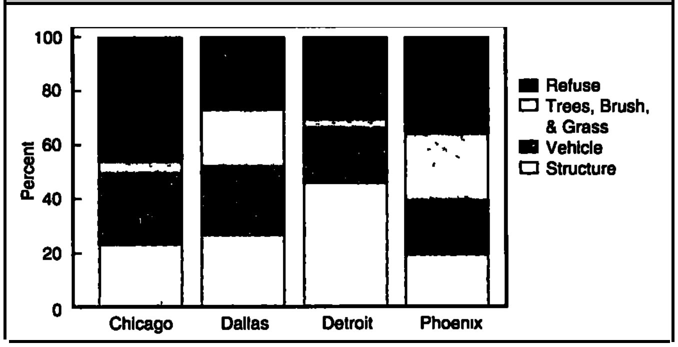
Firefighter Casulalties-1989



4. Trees, brush, and grass fires make up a small percentage of fires in Chicago and Detroit.

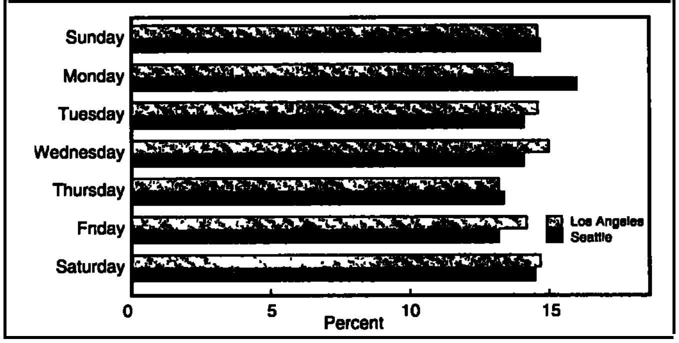
Structure fires are predominant in Detroit where refuse fires are predominant in Chicago.

Comparison of Types of Fires – 1990

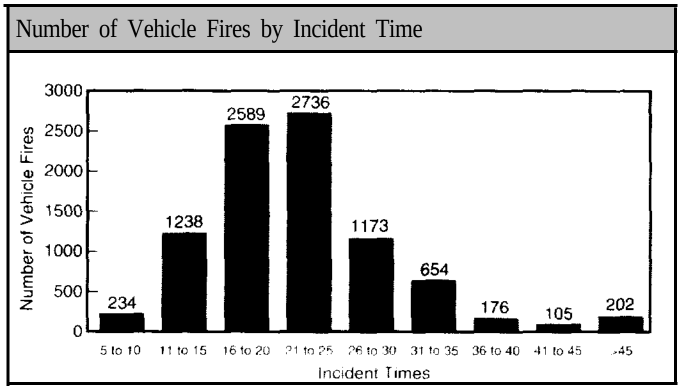


5. The two cities have similar distributions by Day of Week.

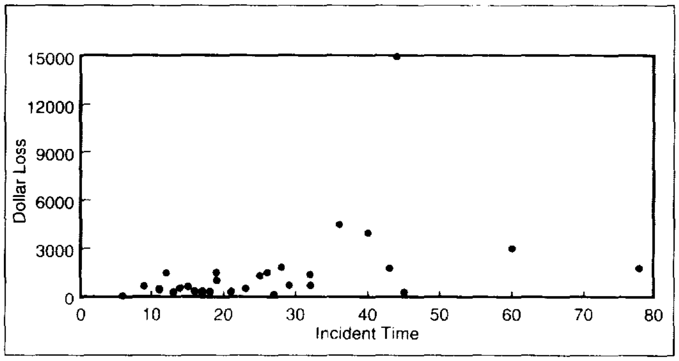
Comparison of Number of Fires by Day of Week – 1990



6.

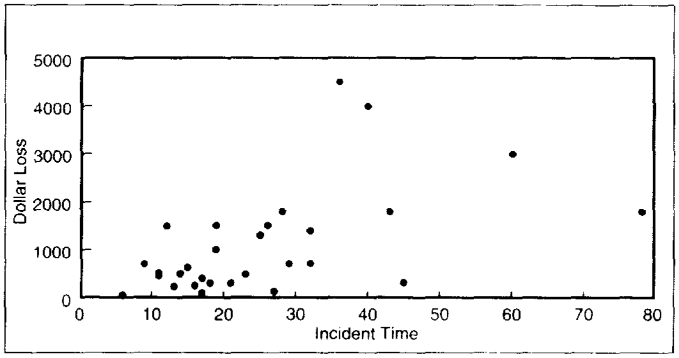

Number of Vehicle Fires by Incident Time

7. a .



h. The outlier is 4-I minutes with $15000 for dollar loss.

c.



d. You can not relate incident time with dollar loss. You would expect the dollar loss to increase with incident time (especially when the incident time is over 30 minutes). As you can see in the scatter plot this is not always true. The incident time of 78 minutes had a lower dollar loss than 36, 40, and 60 minutes.

1. a .

	Travel Time	On-Scene Time	Dollar Loss
Mean	3.89	62.41	4923.67
Median	4	48	1000
Interquartile Range	2	66	4700

- To calculate the mean, sum all the data values and divide by the total number of data values.
- To determine the median, reorder each group of data and take the middle value (also known as the 50th percentile.

	Travel Time New Order	On-Scene Time New Order	Dollar Loss New Order	Cumulative Percent
	1	7	50	3.7
	1	9	100	7.4
	3	10	100	11.1
	3	10	150	14.8
	3	11	189	18.5
	3	20	250	22.2
25%	3	22	300	25.9
	3	23	300	29.6
	4	26	500	33.3
	4	27	500	37.0
	4	33	1000	40.7
	4	33	1000	44.4
	4	35	1000	48.1
50%	4	48	1000	51.9
	4	62	1500	55.6
	4	69	2000	59.2
	4	74	2000	63.0
	4	74	3000	66.7
	4	83	3000	70.4
	4	85	5000	74.1
75%	5	88	5000	77.8
	5	92	7000	81.5
	5	94	10000	85.2
	5	98	15000	88.9
	5	112	20000	92.6
	6	113	23000	96.3
	6	327	30000	100.0

- To determine the interquartile range, find the 25th and 75th percentiles and take the difference. To find the 25th and 75th percentiles, calculate the cumulative percent and pick the value whose cumulative percent is closest to (but not below) the percentile.

  Note: Since we reordered the data and listed all 27 values, the cumulative percent is the same for each of the variables.

b.

	Travel Time	Dollar Loss
Variance	1.41	62750996.1

c. Standard Deviation is the square root of the variance.

	Dollar Loss
Mean	4923.67
Standard Deviation	7021.553

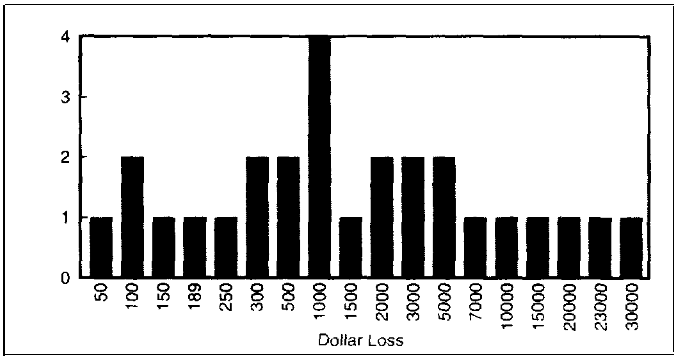
One standard deviation about the mean is from -2997.883 to 12845.243 (mean - one standard deviation) to (mean + one standard deviation). There are 23 (85%) within one standard deviation.

Two standard deviations about the mean is from -10919.43 to 20766.776. There are 25 (93%) within two standard deviations.

d) For Dollar Loss, the large variance means that the spread of data around the mean is very large. For Travel Time, the small variance means that the spread around the mean is small.

e) Large dollar losses increase the mean, but usually have little effect on the median.

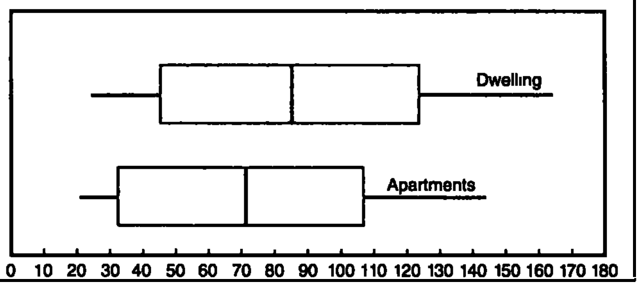
2.



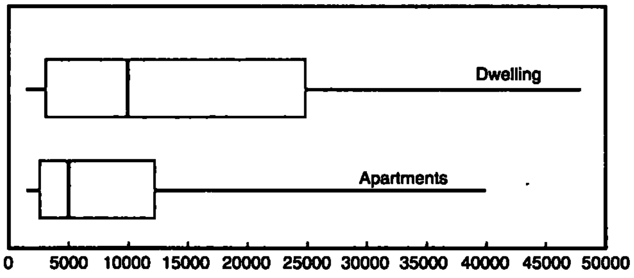
	Dollar Loss	Frequency	Cumulative Distribution	Cumulative Percent
	50	1	1	3.7
10%	**100**	**2**	**3**	**11.1**
	150	1	4	14.8
	189	1	5	18.5
	250	1	6	22.2
25%	**300**	**2**	**8**	**29.6**
	500	2	10	37.0
50%	**1000**	**4**	**14**	**51.9**
	1500	1	15	55.6
	2000	2	17	63.0
	3000	2	19	70.3
75%	**5000**	**2**	**21**	**77.8**
	7000	1	22	81.5
	10000	1	23	85.2
	15000	1	24	88.9
90%	**20000**	**1**	**25**	**92.6**
	23000	1	26	96.3
	30000	1	27	100.0

3. a.

On-Scene Time



Dollar Loss



b. Interquartile Ranges

	On-Scene Time	Dollar Loss
Dwelling	78	22000
Apartments	75	9900

c. Fires in one- and two- family dwellings tend to have longer on-site times and larger dollar losses. For on-site times, the spread of the data is about the same for dwellings and apartments. The interquartile range is 78 for dwellings (123-45) and 75 for apartments (107-32). The dollar losses for dwellings have a greater spread than for apartments. The interquartile range is $22,000 for dwellings compared to $9,900 for apartments.

d. We would expect a greater variance with dwellings because of the larger spread of data.

4. a. On-Scene Time — Dollar Loss

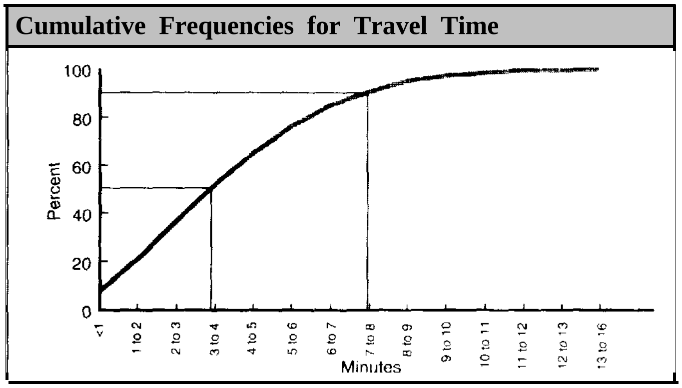
$$CS = \frac{3(62.41 - 48)}{63.36} = .682$$

$$CS = \frac{3(4923.67 - 1000)}{7921.553} = 1.486$$

b. CS will be zero when the mean and median are the same.

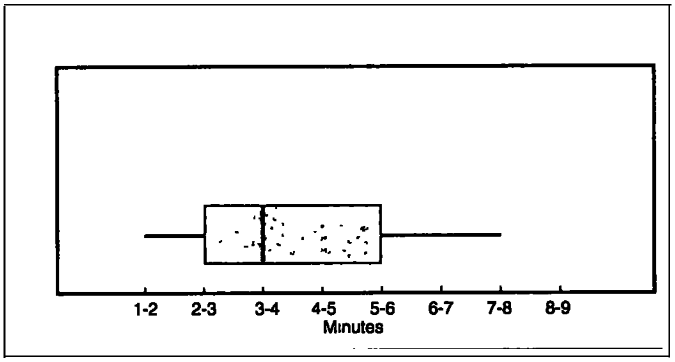
5. a. Drop 50, the 1st 100, 23000, and 30000.
Trimmed Mean = 3469.087

6. a.



Cumulative Frequencies for Travel Time

10% = 1 to 2 minutes
25% = 2 to 3 minutes
50% = 3 to 4 minutes
75% = 5 to 6 minutes
90% = 7 to 8 minutes



b. 90%

c. 3.75 minutes

1. If the die is a fair die, we would expect to have each side visible exactly 10 times (60 ÷ 6).

Step 1. The expected number for each side of the die is 10.

Step 2 and 3.

Dots Visible	Actual Number	Expected Number	Squared Diff.	Divided by Expected
One	13	10	9	.9
Two	11	10	4	.4
Three	7	10	9	.9
Four	12	10	4	.4
Five	9	10	1	.1
Six	7	10	9	.9

Step 4. Chi-squared statistic = sum of the results of the last column = 3.6

Step 5. Degrees of Freedom = *n*- 1 = 5

Step 6. Critical Chi-squared value = 11.07 (look in Appendix A and select the entry associated with 5 degrees of freedom)

2. The chi-squared statistic is less than the critical chi-squared value (3.6<11.07) so we determine that the die is a "fair" die at the 5 percent level.)

	Actual Number	Expected Number	Squared Diff.	Divided by Expected
Jan	467	355.67	12394.37	34.85
Feb	291	355.67	4182.21	11.76
Mar	392	355.67	1319.87	3.71
Apr	322	355.67	1133.67	3.19
May	319	355.67	1344.69	3.78
Jun	349	355.67	44.89	.13
Jul	384	355.67	802.59	2.26
Aug	374	355.67	335.99	.94
Sep	359	355.67	11.09	.03
Oct	368	355.67	152.03	.43
Nov	298	355.67	3325.83	9.35
Dec	348	355.67	58.83	.16

Chi-squared Statistic = 70.59
Degrees of Freedom = 11
Critical Chi-squared Value = 19.68

a. Chi-squared statistic is greater than the critical chi-squared value. This means that the number of fires by month differs significantly from an equal distribution.

b. January, February, and November

3. a .

Type of Fire	Rest of U.S.	Percent	Metros	Percent
Structure	241,481	32.7	54,189	26.9
Outside of Structure	23,986	3.2	4,378	2.2
Vehicle	174,398	23.6	51,774	25.7
Trees, brush, grass	172,217	23.3	28,440	14.1
Refuse	111,291	15.1	59,767	29.7
Other	15,353	2.1	2,833	1.4
Total	738,726	100.00	201,381	100.00

b. Our null hypothesis is that we suspect that the distribution of fires in the Metro Areas does not differ significantly from the rest of the U.S.

Type of Fire	Actual Metros	Expected	Squared Diff.	Divided by Expected
Structure	54,189	65,851.59	136,016,005.51	2,065.49
Outside of Structure	4,378	6,444.19	4,269,141.12	662.48
Vehicle	51,744	47,525.92	17,792,198.89	374.37
Trees, brush, grass	28,440	46,921.77	341,575,822.33	7,279.69
Refuse	59,767	30,408.53	861,919,760.74	28,344.67
Other	2,833	4,229.00	1,948,816.00	460.82
		Calculated chi-squared statistic		39,187.52

Note: To calculate the expected, apply the percentages from the rest of the U.S. For example, 32.7% of the fires for the rest of the U.S. are structure fires. This means we expect that 32.7% of the Metro fires to be structure fires (65,851.59).

Calculated chi-squared statistic = 39,187.52
Degrees of Freedom = 5
Critical chi-squared value = 11.07
Our calculated chi-squared statistic is greater than the critical value.

Our conclusion is that the distribution of fires in the Metro Areas differs significantly from the rest of the U.S.

c. Refuse fires amount for most of the differences in the two distributions. Structure fires and trees, brush, and grass fires also contribute significantly to the chi-squared Statistic.

4.

	Actual	Expected	Squared Diff.	Divided by Expected
Sunday	674	609.86	4,113.94	6.75
Monday	688	623.02	4,222.40	6.78
Tuesday	565	592.46	754.05	1.27
Wednesday	550	613.26	4,001.83	6.53
Thursday	576	585.25	85.56	.15
Friday	588	589.07	1.14	.002
Saturday	603	631.51	812.82	1.29
Total	4,244			22.77

Note: To calculate the expected. apply the percentages from the fires in the U.S. For example, 14.37% of the fires in the U.S. occurred on Sunday. This means we expect 14.37% of the fires to occur on Sunday in Denver. (14.37% times 4244 = 609.86).

Chi-squared statistic = 22.77
Degrees of Freedom = 6
Critical chi-squared value = 12.59

Since our calculated chi-squared statistic is greater than the critical value, our conclusion is that the distribution of fires by day of week for Denver differs from the rest of the country.

5.

	Actual	Expected	Squared Diff.	Divided by Expected
Sunday	674	644.24	885.66	1.37
Monday	688	634.90	2,819.61	4.44
Tuesday	565	596.28	978.44	1.64
Wednesday	550	598.40	2,342.-56	3.91
Thursday	576	570.39	31.47	.055
Friday	588	575.06	167.44	.29
Saturday	603	624.72	471.76	.755
Total	4.2441			12.46

Note: To calculate the expected, apply the percentages from the fires in the other metro cities. For example, 15.18% of the fires in the other metro cities occurred on Sunday-. This means we expect 15.18% of the fires to occur on Sunday,. in Denver. (15.18% times 4244 = 644.24).

Chi-squared statistic = 12.46
Degrees of Freedom = 6
Critical chi-squared value = 12.59

Since our calculated chi-squared statistic is less than the critical value, we can say that the distribution of fires by day of week in Denver does not differ significantly from the other Metro cities.

1.

$$\frac{x_{11}x_{22}}{x_{12}x_{21}} = \frac{36 \times 32}{24 \times 48} = \frac{1552}{1552} = 1$$

Since the Odds Ratio is equal to 1, the two variables have complete independence.

2. 1st Table

$$\frac{601 \times 897 \quad 539{,}097}{1{,}884 \times 166 = 312{,}744} = 1.72$$

2nd Table

$$\frac{837 \times 851 = 712{,}287}{1{,}648 \times 212 \quad 349{,}376} = 2.04$$

3rd Table

$$\frac{801 \times 694 = 555{,}894}{1{,}684 \times 369 \quad 621{,}369} = .89$$

3. In order to verify the model values! you need to sum the mean and term values and take the natural antilogarithm. The term values listed in Exhibit 6-11 are when all 4 variables are present (i, j, k, and 1 are all 1). If a variable is not present (indicated by 2) then the term value is the negative value of the one given in Exhibit 6-11. The easiest way to see this is to develop tables for each of the terms.

Term A	i=1	i=2
	s77	-.577

Term C	k=1	k=2
	-.484	.484

Term B	J=1	j=2
	-.678	.678

Term D	l=1	l=2
	-.591	.591

Term AB	i=1	i=2
j=l	.120	-.120
j=2	-.120	.120

Term BD	j=1	j=2
l=1	.084	-.084
l=2	-.084	.084

Term AD	i=1	i=2
I=1	.169	-.169
I=2	-.169	.169

Term CD	k=1	k=2
l=1	-.326	.326
l=2	.326	-.326

Term BC	j=1	j=2
k=l	.008	-.008
k=2	-.008	.008

Term BCD	j=1			j=2	
	k=1	k=2		k=l	k=2
l=1	-.084	.084	l=1	.084	-.084
l=2	.084	-.084	l=2	-.084	.084

Model (l,l,l,l) = 4.807 + .577 + (-.678) + (-.484) + (-591)
+. 120 + .169 + .008 + .084 + (-.326) + (-.084)
= 3.602
anti log (3.062) = 36.67 Actual = 37

Model (1,1,2,2) = 4.807 + .577 + (-.678) + .484 + 591
+ .120 + (-.169) + (-.008) + (-.084) + (-.326) + (-.084)
= 5.23
anti log (5.23) = 186.79 Actual = 181

Model (2,1,2,1) = 4.807 + (-577) + (-.678) + .484 + (-S91)
+ (-.120) + (-.169) + (-.008) + .084 + .326 + ,084
= 3.642
anti log (3.642) = 38.17 Actual = 3.5

4. a. Look at all standardized values with magnitudes greater than 2.0 (ignoring the sign). All four main effects (A,B,C, and D) are important. Also the following interaction effects are important (in order of magnitude).

AC	(6.3)
AD	(5.2)
BC	(3.5)
CD	(3.4)
BD	(3.1)
ABC	(3.1)
AB	(2.5)

b. Since variable A is the response variable, we want to model how the other three variables relate to A. Since variables B, C, and D are the explanatory variables, any hierarchical models must include the term BCD and any of its higher-order terms (B, C, D, BC, BD, and CD)

Possible hierarchical models would be:

Model	Degrees of Freedom
AB/ACD/BCD	3
AD/ABC/BCD	3
AC/ABD/BCD	3
ACD/BCD	4
AB/AC/AD/BCD	4
AB/AC/BCD	5
AB/AD/BCD	5
AC/AD/BCD	5

5. a. Models 1 and 2. Remember the Y^2 statistic follows a chi-squared distribution. Use Appendix A to test whether a model is less than the entry in Appendix A at the 5 percent level. If the Y^2 statistic is less than the entry in Appendix A, then we accept the model as a good fit to the data.

b. Model 2. When there are two or more models that provide a good fit to the data. select the one with the fewest number of terms.

c.

TABLE 1	Detector Performed	Detector Did Not Perform	Total
Confined to Room	586	1789	2375
Extended Beyond Room	166	867	1033
Total	752	2656	

TABLE 2	Fire Started in Functional Area	Fire Started in Area Non-Functional	Total
Confined to Room	1681	694	2375
Extended Beyond Room	585	448	1033
Total	2266	1142	

TABLE 3	Equipment Involved	No Equipment Involved	Total
Confined to Room	834	1541	2375
Extended Beyond Room	216	817	1033
Total	1050	2358	

d. Table 1

The odds that a fire was confined to the room when the detector performed is 3.53:l. ($x_{11}/x_{21} = 586/166$).

The odds that a fire extended beyond the room when the detector did not perform is 2.06:1.

$$\frac{x_{11}x_{22}}{x_{12}x_{21}} = \frac{586 \times 867}{166 \times 1789} = 1.71$$

Table 2

The odds that a fire was confined to the room when the fire was started in a functional area is 2.87:1.

The odds that a fire extended beyond the room when the fire was started in a non-functional area is 1.55:1.

$$\frac{x_{11}x_{22}}{x_{12}x_{21}} = \frac{1681 \text{ x } 448}{585 \text{ x } 694} = 1.85$$

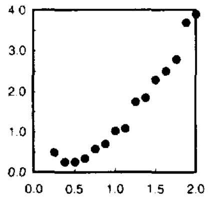
Table 3

'The odds that a fire was confined to the room when equipment was involved is 3.86:1.

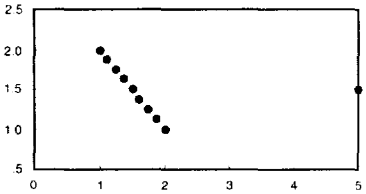
The odds that a fire extended beyond the room when no equipment was involved is 1.89:1.

$$\frac{x_{11}x_{22}}{x_{12}x_{21}} = \frac{834 \text{ x } 817}{216 \text{ x } 1541} = 2.05$$

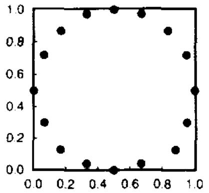
1. a.



r = .957



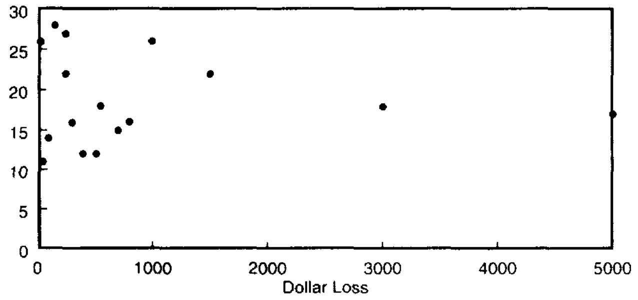
r = .28
As you move the outlier further away from the rest of the points, the correlation gets closer to zero.



r = 0.0
When points are in a circle the correlation is zero.

b. r = 1

2.



Correlation = -.065

Notes:
1) Calculate the average for each variable (sum of the values divided by the total number of values).
2) Calculate the sample standard deviation.

$$\sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

3) Calculate the standard units for each value.

$$\frac{\text{Value - Average}}{\text{Standard Deviation}}$$

4) Calculate the product of the standard units for each pair.
5) Sum the resulting values and divide by the number of points minus 1.

$$\frac{-.98}{16 - 1} = -.065$$

Incident time	Dollar Loss	Incident Time Standard Units	Dollar Loss Standard Units	Product
15	700	-.658	-.221	.145
18	3000	-.132	1.524	-.201
22	250	.570	-.563	-.321
26	35	1.272	-.726	-.923
14	100	-.834	-.677	.565
12	500	.1.185	-.373	.442
28	150	1.623	-.639	-1.037
16	300	-.483	-.525	.254
12	400	.1.185	-.449	.532
16	800	-483	-.145	.070
26	1000	1.272	.006	.008
22	1500	.570	.386	.220
17	5000	-.307	3.042	-.934
18	550	-.132	-.335	.044
27	250	1.448	-.563	-.815
11	50	-1.360	-.714	.971

Average Incident Time = 18.75
Standard Deviation for Incident Time = 5.698
Average Dollar Loss = 991.563
Standard Deviation for Dollar Loss = 1317.791

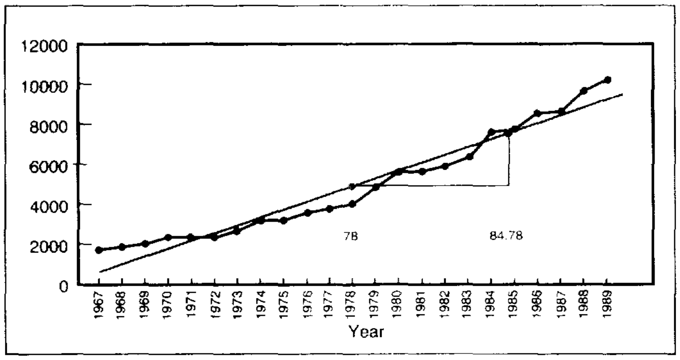
3. a. Estimated Fire Loss = (206.2 x 45) -6266.4
(45 minutes) = 3012.60

Estimated Fire Loss = (206.2 x 180) -6266.4
(180 minutes) =30849.60

b. The slope = 206.2 means that the fire loss increases by 206.2 every time the incident time increases by 1 minute. When the incident time increases by 10 minutes, the fire loss will increase by 2062.00 (206.2 x 10).

c. Incident time is from time of dispatch to time back in service. Sometimes a fire has been going on for a while before the fire department is called, thus resulting in high fire loss even though the incident time is under 30 minutes.

4. a.



$$m = \frac{\text{rxS.D. of Fire Loss}}{\text{S.D. of Years}} \quad \frac{.9753 \times 2696.394}{6.782} = 387.76$$

Determine the intercept:

Fire Loss Average = m x Year Average + b
4929.043 = (387.76 x 78) + b
-25316.237 = b

Regression Line is: Fire Loss = 387.76 x Year -25316.237

Note: If you have Fire Loss along the x-axis and year along the y-axis, your regression line will be: Year = .(0025 x Fire Loss + 65.68

c.

Year	Estimated Fire Loss
1988	8806.64
1989	9194.40
1990	9582.16

Fire Loss = 387.76 x 88 -25316.237

d. We want the regression line to pass through one point that represents the two averages. Put another point by moving one standard deviation for the year to the right and upward by one standard deviation for Fire Loss times the correlation.

This page intentionally left blank.

Chapter 8

1. Fires = 2.0 + (.0017 x POP) + (.068 x BOARDED) +
(.013 x FAMTYPE) + (.060 x DENSITY)

<u>Example</u>
Census Tract 709

Fires = 2.0 + (.0017 x 1386) + (.068 x 19) + (.013 x 184) + (.060 x 100)
= 2.0 + 2.3562 + 1.292 + 2.392 + 6
= 14.0402

Census Tract	Actual Fires	Estimated Fires
709	15	14.04
812	30	29.68
902	30	16.77
907	10	10.20

2. Travel Time = 5.27 + (1.38 x DELAYS) + (.36 x CALLTYPE) +
(3.83 x AREAl) + (2.48 x AREA2)

Remember:	CALLTYPE	0 for ALS Calls 1 for BLS Calls
	DELAYS	0 not delayed 1 delayed
	AREA1	0 not Area 1 1 Area 1
	AREA2	0 not Area 2 1Area2

a. Travel Time = 5.27 + (1.38 x 0) + (.36 x 1) + (3.83 x 1) + (2.48 x 0)
= 9.46

b. Travel Time = 5.27+ (1.38 x 1) + (.36 x 0) + (3.83 x 0) + (2.48 x 1)
= 9.13

c. Travel Time = 5.27+ (1.38 x 0) + (.36 x 0) + (3.83 x 0) + (2.48 x 0)
= 5.27

d. Travel Time = 5.27 + (1.38 x 1) + (.36 x 1) + (3.83 x 0) + (2.48 x 0)
= 7.01

1. a.

Unit Utilization (3) $= \frac{3.5 \times 45}{3 \times 60} = \frac{157.5}{180} = .875$ 87.5%

Unit Utilization (4) = 65.6%
Unit Utilization (5) = 52.5%
Unit Utilization (6) = 43.7%

b. A change from 4 units to 5 units will result in a difference of 13.1%
A change from 5 units to 6 units will result in a difference of 8.8%

2. a. To obtain the Probability of Delay you have to:

1) Calculate the Table Key $= \frac{ct}{60}$, where c is the number of calls per hour and t is the average time per call.
2) Use Appendix B to find the probability that a call will be delayed. The number of units are across the top and the Table Key is shown down the left column.

Table Key $= \frac{ct}{60} = \frac{3.5 \times 45}{60} = 2.625$ 2.6 rounded

Probability of Delay for:

3 Units = 75.9%
4 Units = 35.4%
5 Units = 14.9%
6 Units = 5.6%

b. To determine the average number of citizens waiting, use Appendix C. The number of units is displayed across the top and the Table Key is shown down the left column.

Average Number of Citizens Waiting:

3 Units = 4.93
4 Units = .66
5 Units = .16
6 Units = .04

c. Waiting Time $= \frac{\text{Appendix C Entry} \times 60}{c}$ where c is the number of calls per hour.

3 Units 84.51 minutes (4.93 times 60 divided by 3.5)
4 Units 11.31 minutes
5 Units 2.74 minutes
6 Units .69 minutes

3. a. 6% of 3.5 calls per hour = .21
3.5 + .21 = 3.71
3.71 calls per hour

b.

Unit Utilization $= \frac{ct}{60n} = \frac{3.71 \times 50}{60 \times 4} = =.773$ 77.3%

Table Key$= \frac{ct}{60} = \frac{3.71 \times 50}{60} = 3.09 = 3.1$ rounded

Probability of Delay for 4 units is 55.2%
Average Number People Waiting is 1.9

C. In question #2, the probability of delay for 4 units was 35.4% and the average number of people waiting was .66. The 6% increase in calls per hour and the 5 minute increase on each call greatly increased the probability of delay and tripled the number of people who would be waiting.

4. There has to be an analysis of records available to your department. This could be a computer aided dispatch system (CAD) which would have good time information to determine how busy units are. You should also perform an analysis of schedules in order to determine how many units are fielded.

If you are going to calculate unit utilization, you need to find out how many units are fielded and not scheduled. You might want to look at daily rosters or time cards to determine how many units were actually fielded.

Also read over the section in chapter one about data quality.

5. There is no one way of setting objectives. There needs to be a decision made by the command staff of the fire department. It is also a good idea to involve someone from the city or county when setting the objectives. One popular approach is to perform an analysis of current operations to determine current performance levels. Then determine if the command staff is satisfied with current performance or wants to make improvements. Queuing analysis is a good first step in determining staffing needs for improvements.

6. a.

Unit utilization not more than 60 percent	5 units
Probability of delay not more than 3 percent	7 units
Average number of citizens waiting not to exceed 25	5 units

Unit Utilization $= \frac{ct}{60n}$ $\quad .60 = \frac{4 \times 38}{60n}$ $\quad n = 4.2$

If you had 4 units, the unit utilization would be 63.3%
If you had 5 units, the unit utilization would be 50.67%

Table Key $= \frac{ct}{60} = \frac{4 \times 38}{60} = 2.53$

Probability of Delay for:	Average Number of Citizens Waiting:
4 units = 32%	4 units = .53
5 units = 13%	5 units = .13
6 units = 4.7%	0 units = .03
7 units = 1.5%	7 units = .01

b. 7 units

Unit Utilization (7 units) $= \frac{4x\ 38}{m} =$ .362 36.2%

7. a. 4 units

Waiting Time $= \frac{\text{Appendix C Entry} \times 60}{c}$

Waiting Time (4 units) $= \frac{.53 \times 60}{4} =$ 7.05 minutes

Waiting Time (5 units $= \frac{.13 \times 60}{4}$ 1 .05 minutes

8. a.

EMS Calls	Number of Hours	Percent	Expected Percentage of EMS Calls
0	158	26.0	23.2
1	192	31.6	33.9
2	144	23.6	24.8
3	68	11.2	12.0
4	34	5.6	4.4
5	10	1.6	1.3
6	2	.3	.3
Total	608		

Note: Poisson Distribution: $P(x=k) = \frac{e^{-c}c^k}{k!}$

Where c is the average number of calls per hour and k is an integer value starting with 0.

$$\frac{e^{-1.46}1.46^0}{0!} = \frac{.2322}{0!} = .232 \quad 23.2\%$$

b. If the calls follow a Poisson distribution, we should be able to obtain a straight line by plotting k against the quantity:

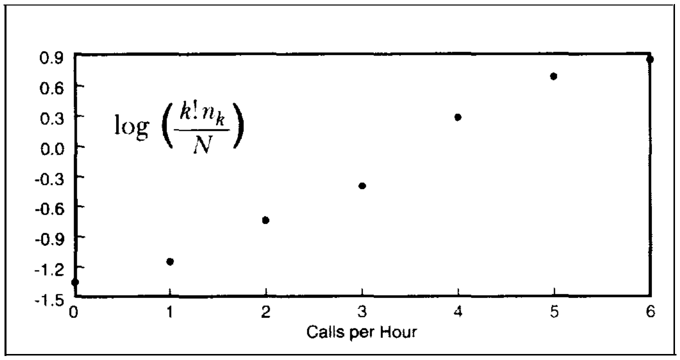
$$\log\left(\frac{k!n_k}{N}\right)$$

where N is the total number of hours.

Note: by definition 0! = 1.

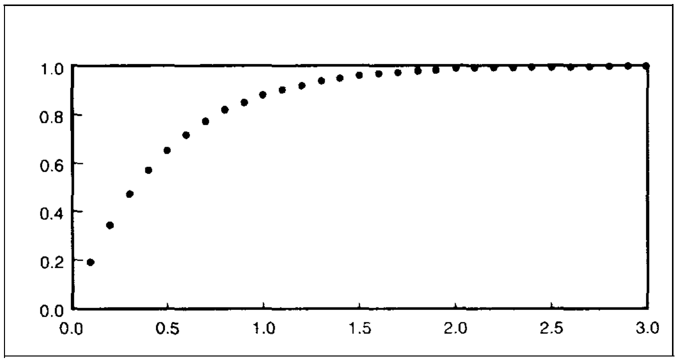
0 calls $\log\left(\frac{0! \times 158}{608}\right) = -1.35$

0 calls -1.35
1 call -1.15
2 calls -.75
3 calls -.40
4 calls .29
5 calls .68
6 calls .86



c. The conclusion is that the Poisson distribution gives a good approximation to the experienced distribution because it produces a relatively straight line.

9. a.



b. 25th percentile = .804
median = .962
75th percentile = .993

This page intentionally left blank.

References

Agresti, Alan, and Barbara Finlay. *Statistical Methods for the Social Sciences.* San Francisco, California: Dellen Publishing Company, 1986.

Cleveland, William S. *The Elements of Graphing Data.* Monterey, California: Wadsworth Advanced Books and Software, 1985.

Devore, Jay and Roxy Peck. *Statistics: The Exploration and Analysis of Data.* New York, New York: West Publishing Company, 1986.

Freedman, David, Robert Pisani, and Roger Purves. *Statistics.* New York, New York: W.W. Norton & Company, 1978.

Goodman, L.A. “The Analysis of Multidimensional Contingency Tables, Stepwise Procedures and Direct Estimation Methods for Building Models for Multiple Classifications.” *Technometrics* (1971): 13, 33-61.

Hoaglin, David D., Federick Mosteller, and John W. Tukey. *Exploring Data Tables, Trends, and Shapes.* New York, New York: John Wiley and Sons, 1985.

Jaffe, A. J. and Herbert F. Spirer. *Statistics: Straight Talk for Twisted Numbers.* New York, New York: Marcel Dekker, Inc., 1987.

Mosteller, Frederick, Stephen E. Fienberg, and Robert E.K. Rourke. *Beginning Statistics with Data Analysis.* Reading, Massachusetts: Addison-Wesley Publishing Company, 1983.

Phillips, John J., Jr. *How to Think About Statistics.* New York, New York: W.H. Freeman and Company, 1988.

Schefler, William C. *Statistics: Concepts and Applications.* Menlo Park, California: The Benjamin/Cummings Publishing Company, Inc., 1988.

Siegel, Andrew F. *Statistics and Data Analysis: An Introduction.* New York, New York: John Wiley & Sons, Inc., 1988.

Tanur, Judith M., Frederick Mosteller, William H. Kruskal, Erich L. Lehmann, Richard F. Link, Richard S. Pieters, and Gerald R. Rising. *Statistics: A Guide to the Unknown.* Pacific Grove, California: Wadsworth & Brooks/Cole Advanced Books & Software (Third Edition), 1988.

Upton, Graham J.G. *The Analysis of Cross-tabulated Data.* New York, New York: John Wiley & Sons, 1978.

Zeisel, Hans. *Say It With Figures.* New York, New York: Harper & Row Publishers, Inc., 1985.

Zelazny, Gene. *Say It With Charts.* Homewood, Illinois: Dow Jones-Irwin, Inc., 1985.

This page intentionally left blank.

Key Equations

Sample Mean $= \bar{x} = \frac{\sum x_i}{n}$

Sample Variance = $s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$

Sample Standard Deviation $= s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$

Chi-squared Value $= \sum \frac{(\text{observed - expected})^2}{\text{expected}}$

Expected Value = $\frac{\text{Row Sum x Column Sum}}{\text{Grand Total}}$

Odds Ratio $= \frac{x_{11}\, x_{22}}{x_{12}\, x_{21}}$

Standard Units $= \frac{x_i - \bar{x}}{s}$

Correlation $= r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{(n-1)\, s_x s_y}$

Slope $= m = \frac{r \text{ x } s_x}{s_y}$

Intercept = $b = \bar{y} - m\bar{x}$

Regression Line $= y = mx + b$

Standard Error = $\sqrt{\frac{SSE}{n-k-1}}$

where k = number of independent variables

Sum of Squared Errors $= SSE = \sum (y_i - \hat{y}_i)^2$

Total Sum of Squares $= SST = \sum (y_i - \bar{y})^2$

Coefficient of Determination = $R^2 = \frac{SST - SSE}{SST}$

Unit Utilization $= \frac{ct}{60n}$

Queuing Table Key = $\frac{ct}{60}$

Waiting Time = $\frac{\text{Queue Length x 60}}{c}$

Poisson Distribution $= P(x = k) = \frac{e^{-c}c^k}{k!}$

Exponential Distribution $= f(x) = ue^{-ux}$

APPENDIX F

FIRE DATA ANALYSIS HANDBOOK — SECOND EDITION

This page intentionally left blank.

Fire Data Analysis Handbook

Second Edition

FA-266/January 2004

U.S. DEPARTMENT OF HOMELAND SECURITY

FEMA

U.S. Fire Administration

Mission Statement

As an entity of the Federal Emergency Management Agency, the mission of the United States Fire Administration is to reduce life and economic losses due to fire and related emergencies, through leadership, advocacy, coordination, and support. We serve the Nation independently, in coordination with other Federal agencies, and in partnership with fire protection and emergency service communities. With a commitment to excellence, we provide public education, training, technology, and data initiatives.

On March 1, 2003, FEMA became part of the U.S. Department of Homeland Security. FEMA's continuing mission within the new department is to lead the effort to prepare the nation for all hazards and effectively manage federal response and recovery efforts following any national incident. FEMA also initiates proactive mitigation activities, trains first responders, and manages Citizen Corps, the National Flood Insurance Program and the U.S. Fire Administration.

FEMA

Foreword

The fire service exists today in an environment constantly inundated with data, but data are seen of little use in the everyday, real world in which first responders live and work. This is no accident. By themselves, pieces of data are of little use to anyone. Information, on the other hand, is very useful indeed. What's the difference? At sporting events, people in stadiums hold up individual, multi-colored squares of cardboard to form a giant image or text, which could be recognized only from a distance. This is a good analogy for data and information. The individual squares of cardboard are like data. They are very numerous and they all look similar taken by themselves. The big image formed from the organization of thousands of those cards is like information. It is what can be built from many pieces of data. Information then is an organization of data that makes a point about something.

The fire service of today is changing. More and more, it is not fighting fires as much as it is doing EMS, HAZMAT, inspections, investigations, prevention, and other nontraditional but important tasks which are vital to the community. Balancing limited resources and justifying daily operations and finances in the face of tough economic times is a scenario that is familiar to every department.

Turning data into information is neither simple nor easy. It requires some knowledge of the tools and techniques used for this purpose. Historically, the fire service has had few of these tools at its disposal and none of them has been designed with the fire service in mind. This book changes that. It was designed solely for the use of the fire service. The examples were developed from the most recent fire data collected from departments all over the Nation. This book also was designed to be modular in form. Many departments' information needs can be met by using only the first few chapters. Others with a more statistical leaning may want to go further. The point is, it's up to the reader to decide. This handbook is just another tool, like a pumper or a ladder, to help do the job.

In this revised edition, the use of statistical symbols and formulas has been eliminated for ease of use and understanding. The problems at the end of each chapter also have been left out. The philosophy behind this is not to discourage anyone seeking immediate results, and to encourage those with a desire for more indepth knowledge of statistical analysis tools.

The United States Fire Administration

This page intentionally left blank.

Table of Contents

CHAPTER 1: INTRODUCTION

The primary objective of this handbook is to describe statistical techniques for analyzing data typically collected in fire departments. Motivation for the handbook stems from the belief that fire departments collect an immense amount of data, but do very little with it. Think for a minute about the reports you complete on incidents. You probably document the type of situation found, action taken, time of alarm, time of arrival, time completed, number of engines responding, number of personnel responding, and many other items. For fires, the list grows even longer to include area of fire origin, form of heat of ignition, type of material involved, and other related facts. Additionally, if civilian or firefighter injuries occur, other reports need to be completed.

A compelling reason for these reports is a legal requirement for documenting incidents. Victims, insurance companies, lawyers, and many others want copies of reports. Indeed, fire departments maintain files for retrieval of individual reports.

The reports can, however, provide a more beneficial service to fire departments by yielding insight into the nature of fires and injuries in their jurisdiction. Basic information probably is available already. Typically, the number of fires handled last year, the number of fire-related injuries, and the number of fire deaths are tracked. It is another story, however, if more probing questions are asked:

- How many fires took place on Sundays, Mondays, etc.?
- How many fires took place each hour of the day or month of the year?
- What was the average response time to fires?
- How much did response times vary by fire station areas?
- What was the average time spent at the fire scene?
- How much did the average time vary by type of fire?

This handbook describes statistical techniques to turn data into information for answering these types of questions and many others. The techniques range from simple to complex. For example, the next two chapters describe how to develop charts to provide more effective presentations about fire problems. These charts may be beneficial to city or county officials on the activities and needs of your fire department. Chapter 4 discusses measures of central tendency (means, medians, modes) and measures of dispersion (range, variance, standard deviation). In Chapter 5, the chi-square statistic and its use in analyzing table data is presented. In Chapter 6 the Pearson correlation coefficient and some additional correlations are discussed. These are all techniques that can tell you more about the nature of fires and injuries.

One way to become more comfortable with analysis is to work with real data. For this handbook, data were obtained from fire departments in several large metropolitan areas. By working with real data, it should be easier to understand the different techniques.

Why Data Analysis?

There still may be a question in your mind as to why we should go to all this trouble to analyze data. Many decisions do not require analysis, such as decisions on personnel, grievance proceedings, promotions, and even decisions on how to handle a fire. It is certainly true that fire departments can continue to operate in the same way they always have without doing a lot of analysis.

On the other hand, there are three good reasons for looking closely at the data: (1) to gain insights into fire problems, (2) to improve resource allocation for combating fires, and (3) to identify training needs. Probably the most compelling is that analysis gives insight into fire problems, which in turn can affect operations in the department. One may find, for example, that the average time to fires in an area is 6 minutes, compared to less than 2 minutes overall. This result may be helpful in requests for more equipment, more personnel, or justifying another fire station.

As an example of improved resource allocation, statistical analysis of emergency medical calls can determine the impact of providing another paramedic unit in the field. Increasing the number of EMS units from four to five may, for example, decrease average response times from 5 minutes to 3 minutes – a change that may save lives.

Another reason for analysis is to identify training needs. Most training on fire-fighting is based on a curriculum that has been in place for many years. It makes sense to see how training matches characteristics of fires in a particular jurisdiction. This is not to say that other training is not important, since an exception can always occur. However, knowing more about the fires in an area can improve the training. Additionally, an analysis of firefighter injuries may indicate a need for certain types of training.

In summary, this handbook will help you deal with the volume of data collected on fire incidents. By using the techniques presented in this handbook, you should be able to improve your skills in collecting data, analyzing data, and presenting the results.

National Fire Incident Reporting System

The National Fire Incident Reporting System (NFIRS) began over 25 years ago with the aim of collecting and analyzing data on fires from departments across the country. More than 14,000 fire departments in 42 States now report their fires and injuries to NFIRS. This makes NFIRS the largest collector of fire-related incident data in the world. NFIRS contributes over 900,000 fire incidents each year to the National Fire Database.

Incident data collection is not new. In 1963 the National Fire Protection Association (NFPA) developed a dictionary of fire terminology and associated numerical codes to encourage fire departments to use a common set of definitions. This dictionary is known as the NFPA 901, *Standard Classifications for Incident Reporting and Fire Protection Data*. The current set of codes used in NFIRS version 5.0 represents the merging of the ideas from NFPA 901 and the many suggested improvements from the users of the NFIRS 4.1 coding system.

Version 5.0 of NFIRS consists of 11 separate modules in which fire departments can report any type of incident that they respond to. The basic module (Module 1), which is required, includes incident number and type, date, day of week, alarm time, arrival time, time in service, and type of action taken. Modules 2 through 5 are required if applicable. If the incident is a fire, the fire module (Module 2) is completed. This includes property details, cause of ignition, human factors, equipment involved, and other information. If it is a structure fire, Module 3 (structure fire) is completed. This would include such things as structure type, main floor size, fire origin, presence of detectors and automatic extinguishment equipment, and other data. If there were civilian casualties or fire service casualties, Modules 4 or 5, respectively, would be filled out. The remaining modules are optional at the local level. They include EMS (Module 6), Hazardous Material (Module 7), Wildland Fire (Module 8), Apparatus or Resources (Module 9), Personnel (Module 10), and Arson (Module 11).

Usually, the State Fire Marshal's office in each NFIRS State has the responsibility for collecting data from its fire departments. They normally collect data in two ways. One way is that fire departments without any data processing capabilities send their paper reports to the fire marshal's office (or cognizant office). The office then enters the reports into a computer system. Local departments with data processing capabilities send their data electronically or on diskettes or tapes. In either case, the State Fire Marshal's office merges all reports onto a database.

This statewide database then is forwarded electronically to the National Fire Data Center (NFDC) at the U.S. Fire Administration (USFA). The NFDC then can compare and contrast statistics from States and large metropolitan departments to develop national public education campaigns, make recommendations for national codes and standards, guide allocations of Federal funds, ascertain consumer product failures, identify the focus for research efforts, and support Federal legislation.

Every fire department is responsible for managing its operations in such a way that firefighters can do the most effective job of fire control and fire prevention in the safest way possible. Effective performance requires careful planning, which can take place only if accurate information about fires and other incidents is available. Patterns that emerge from the analysis of incident data can help departments focus on current problems, predict future problems in their communities, and measure their programs' successes.

The same principle is also applicable at the State and national levels. NFIRS provides a mechanism for analyzing incident data at each level to help meet fire protection management and planning needs. In addition, NFIRS information is used by labor organizations on a variety of matters, such as workloads and firefighter injuries.

Data Entry and Data Quality

An assumption throughout the handbook is that data on fire incidents and casualties have been entered into a computer and are available for analysis. While manual analysis certainly is possible, it usually is avoided because the tedious calculations quickly overwhelm our ability to perform analysis in any meaningful manner. The advantage of a computer is that it processes data quickly and accurately.

Most fire departments have a computer system of some sort ranging from personal computers (PC) for small departments to Local Area Networks (LAN) or Wide Area Networks (WAN) for large metropolitan departments or in regional settings where multiple departments agree to share a system. Whatever the case, the data are entered into either a custom software program that is purchased by the State or local fire department, or the free client tool that is supplied by the USFA to the States. If a custom vendor software is used it must be compatible with the NFIRS program. A list of registered vendors is available from the USFA, but it is the responsibility of the individual States to assure that a vendor's software meets the qualifications. If the USFA client tool is used, it must be supported by the State.

One word of caution, however, is that any program you purchase should contain a good error checking routine. Data quality is always a problem, and the old adage "Garbage In, Garbage Out" certainly applies to fire department reports. The entry program should, for example, check each item to make sure a valid code has been entered. Whenever the program encounters an error, it should give an opportunity to correct the error before it becomes part of the database. For example, alarm times obviously cannot have hours greater than 23 and minutes greater than 59. An entry program should check hours and minutes for valid numbers, and allow corrections to be made immediately.

The data collected to describe an incident are the foundation of the system. Therefore, editing and correcting errors is a system-wide activity, involving local, State, and Federal organizations. All errors resulting from the edit/update process need to be reported to fire departments and the submission of corrections from fire departments is essential. This is especially important for fatal errors, which prevent the data from being included in the NFIRS database.

Fire departments need to establish *data quality procedures* if they intend to take full advantage of their data. There should be a system in place to double check the collection and data entry work. Field edits and relational edits can be built into the system that will reveal unacceptable and unreasonable data. Data management personnel can use these techniques to improve and validate the data.

In summary, data entry programs should include code checking routines to identify errors in individual items in the report and errors reflected through inconsistencies between items. Because entry programs cannot be expected to find all errors, fire departments also need data quality procedures to ensure that correct data are entered into their systems.

Statistical Packages for Computers

In this handbook we present many different types of analysis. Chapter 3, for example, discusses several types of charts, including bar charts, column charts, histograms, line charts, and dot charts. Other chapters show how to calculate statistics, such as means and variances, and how to do more advanced calculations such as chi-square tests, and correlation coefficients.

In the future, you will want to depend on computers with analysis programs to perform these calculations instead of doing them manually. Many of them are time consuming and cumbersome, and the more advanced ones are all but impossible to do manually. For a good understanding of the analysis, you need to know what is involved, but you should not continue in a manual mode. There are several good statistical packages available for both personal and mainframe computers. If you intend to apply the techniques in this handbook, you should acquire and learn how to use one of these packages. They are as follows:

SPSS, Inc.
233 South Wacker Drive, 11th Floor
Chicago, IL 60606-6412
312-651-3000
Web site: www.spss.com

SAS Institute, Inc.
100 SAS Campus Drive
Cary, NC 27513-2414
919-677-8000
Web site: www.sas.com

SYSTAT Software Inc.
501 Suite F
Point Richmond Tech Center
Canal Boulevard
Richmond, CA 94804-2028 U.S.A.
866-797-8288
Web site: www.systat.com

NCSS
329 North 1000 East
Kaysville, UT 84037
800-898-6109
Web site: www.ncss.com

How to Use This Handbook

Data analysis is not an easy process. It requires careful data collection, attention to detail, access to statistical programs, and skills in result interpretation. These are not impossible tasks, but require time and patience on your part for success. Equally important, you need experience. In the long run, you can only develop capabilities in analysis by applying techniques from this handbook on actual data sets.

As a final note, one way of thinking about analysis is to consider it a four stage process. Stage one is to collect the *data*, which is what the NFIRS program does. In and of themselves, the data are meaningless. They must be organized and summarized into *information* that can be analyzed (stage two). In the third stage, the data are analyzed according to whatever problem or issue is being considered. This yields a better *understanding* of the information, which allows appropriate *decisions* to be made (stage four).

Our ultimate objective is to make better and more informed decisions in fire departments. Data have no utility in a vacuum, and fire reports stay as data if we do nothing. *Analysis turns data into information*. We move, for example, from knowing individual alarm and arrival times to knowing average travel times. Our review of travel times increases our *knowledge* about what is going on with fire incidents, which results, in turn, in more informed *decisions* within fire departments.

The remainder of this handbook is organized as follows. Chapters 2 and 3 are devoted to descriptions of different types of charts and graphs. Chapter 2 describes histograms, which are probably the easiest charts to understand. Chapter 3 expands to other types of charts, column charts, pie charts, and dot charts. In Chapter 4 several basic statistics are introduced, including means, medians, modes, variances, and standard deviations. Chapter 5 discusses analysis of tables, which is particularly important since fire data often come to us as summaries in the form of tables. In Chapter 6 correlations and variable relationships are discussed. In both chapters, the goal is to present how to perform the calculations associated with these subjects as well as how to interpret the results.

In developing these chapters, it was recognized that readers will have varying backgrounds and capabilities. Therefore, while a certain understanding of the principles behind the various techniques is presented, in most cases a practical application approach is used. The subject material becomes more difficult as the handbook progresses. The first few chapters are easy enough to understand by anyone. More technical subjects, such as chi-square analysis and correlation, are more difficult and may require knowledge of basic algebra to understand completely. Even in these chapters, however, emphasis has been placed on understanding results rather than concentrating on theory.

It should be noted that every effort has been made to simplify what can be a very complicated topic. While there are many mathematical and statistical symbols normally involved with the formulas and calculations used in this handbook, none are used here. This was done in order to lessen the confusion. This is meant to be a handbook, not a statistical textbook. It is written so that anyone can pick it up and be able to do basic statistical analysis of data. For those who wish for more indepth discussion of the subject matter, a list of recent texts is included.

Books on Statistics and Data Analysis

The following is a sampling of books on data analysis techniques as well as some specific statistical topics handled or referred to in this book. Most are basic or intermediate in scope, but all have more detail than can be presented in this book.

Analyzing Tabular Data: Loglinear and Logistic Models for Social Researchers by Nigel G. Gilbert (UCL Press, London, 1993).

Data Analysis: An Introduction by Michael S. Lewis-Beck (Sage Publications, Thousand Oaks, CA, 1995).

From Numbers to Words: Reporting Statistical Results for the Social Sciences by Tyler R. Harrison, Susan E. Morgan, and Tom Reichert (Allyn and Bacon, Boston, 2002).

Misused Statistics by A. J. Jaffe, Herbert F. Spirer, and Louise Spirer (M. Dekker, New York, 2nd ed., rev. and expanded, 1998).

Say It With Charts by Gene Zelazny (McGraw-Hill, New York, 4th ed., 2001).

Schaum's Outline Theory and Problems of Beginning Statistics by Larry J. Stephens (McGraw-Hill, New York, 1998).

Sorting Data: Collection and Analysis by Anthony P. M. Coxon (Sage Publications, Thousand Oaks, CA, 1999).

Statistics by David Freedman, Robert Pisani, and Roger Perves (W. W. Norton, New York, 3rd ed., 1998).

Statistics: Concepts and Applications by Amir D. Aczel (Irwin, Chicago, 1995).

Statistics and Data Analysis: An Introduction by Charles J. Morgan and Andrew F. Siegel (J. Wiley, New York, 2nd ed., 1996).

Statistics: The Exploration and Analysis of Data by Jay Devore and Roxy Peck, (Brooks/Cole, Pacific Grove, CA, 4th ed., 2001).

Your Statistical Consultant: Answers to Your Data Analysis Questions by Rae R. Newton and Kjell Erik Rudestam (Sage Publications, Thousand Oaks, CA, 1999).

CHAPTER 2: HISTOGRAMS

Data as a Descriptive Tool

“A picture is worth a thousand words” is an old saying which applies to numbers as well as words. The task of reaching conclusions from numbers is a formidable one, particularly when we are looking for trends and patterns in the data. It is for this reason that we turn our attention to histograms and other charts in this chapter and Chapter 3. These tools will assist you in understanding fire data, since the human mind seems to comprehend pictures quicker than words and numbers.

The techniques found in these two chapters include:

- Histograms
- Column charts
- Pie charts
- Pictograms
- Bar charts
- Line charts
- Dot charts

This chapter describes histograms, while Chapter 3 is devoted to the other techniques. With these graphic aids, we can answer several basic questions. When are fires most likely to occur? What are the primary causes of residential fires? Vehicle fires? How many civilian injuries occurred last year by month? What are the ages of civilian casualties? What percent of the fire incidents have travel times less than 4 minutes? How many structure fires resulted in dollar losses greater than $50,000?

A **histogram** is a column graph where the height of the columns indicates the relative numbers or frequencies or values of a variable. The values may be numeric, such as travel times, or non-numeric, such as days of the week. The following examples show how to organize and display fire data into histograms.

Example 1. One of the most fundamental ways to describe the fire problem is to show how fires are distributed by month, day of week, and hour of day. For example, Exhibit 2-1 shows a frequency list of fires by hour of day for Canton, Ohio, for 1999. A list or array of numbers such as this is almost always the starting point for a descriptive analysis, but the numbers by themselves are not very useful. It is difficult to get a “feel” for what is happening by scanning a list of numbers.

To grasp what the numbers say in Exhibit 2-1, we can develop a frequency histogram, as shown in Exhibit 2-2. Similarly, Exhibits 2-3 and 2-4 show histograms by day of week and month of year. Study these exhibits for a few minutes and draw your own conclusions about what they represent. Don't dwell on individual numbers, but instead look for patterns. Ask yourself three questions:

- Where are the low points and high points in the histogram?
- What groups of times (hours, days, or months) have similar frequencies?
- Is there anything in the histogram that runs counter to your experience?

Exhibit 2-1

Fires by Hour of Day - Canton - 1999

Time Period	Number	Time Period	Number
Midnight - 1 a.m.	15	Noon - 1 p.m.	31
1 a.m. - 2 a.m.	15	1 p.m. - 2 p.m.	33
2 a.m. - 3 a.m.	13	2 p.m. - 3 p.m.	39
3 a.m. - 4 a.m.	13	3 p.m. - 4 p.m.	35
4 a.m. - 5 a.m.	13	4 p.m. - 5 p.m.	46
5 a.m. - 6 a.m.	11	5 p.m. - 6 p.m.	39
6 a.m. - 7 a.m.	16	6 p.m. - 7 p.m.	30
7 a.m. - 8 a.m.	11	7 p.m. - 8 p.m.	50
8 a.m. - 9 a.m.	17	8 p.m. - 9 p.m.	32
9 a.m. - 10 a.m.	17	9 p.m. - 10 p.m.	29
10 a.m. - 11 a.m.	19	10 p.m. - 11 p.m.	28
11 a.m. - Noon	19	11 p.m. - Midnight	24

Answers to these questions provide the first insights into your fire data and any conclusions you make from it.

While these histograms suggest several conclusions, the key ones are:

1. Canton has two distinct hourly patterns. The hours from noon to midnight overall have almost twice the fires than the hours from midnight to noon. The hours of 7 p.m. to 8 p.m. and 4 p.m. to 5 p.m. have more fires than any other hours in the day.

2. The lowest time period for fires is from 2 a.m. to 8 a.m.

3. Sunday is the busiest day by far for fires with a continuing tapering off until Friday, the least busy of days.

4. May has the most fires with June, July, and November tied for second. The fewest are in February.

With these histograms we begin to see a picture of the fire problem in Canton. Histograms allow for an easy descriptive and analytical procedure without having to think too much about the numbers themselves. Graphical displays should always strive to convey an immediate message describing a particular aspect of the data.

Exhibit 2-2

Fires by Hour of Day - Canton - 1999

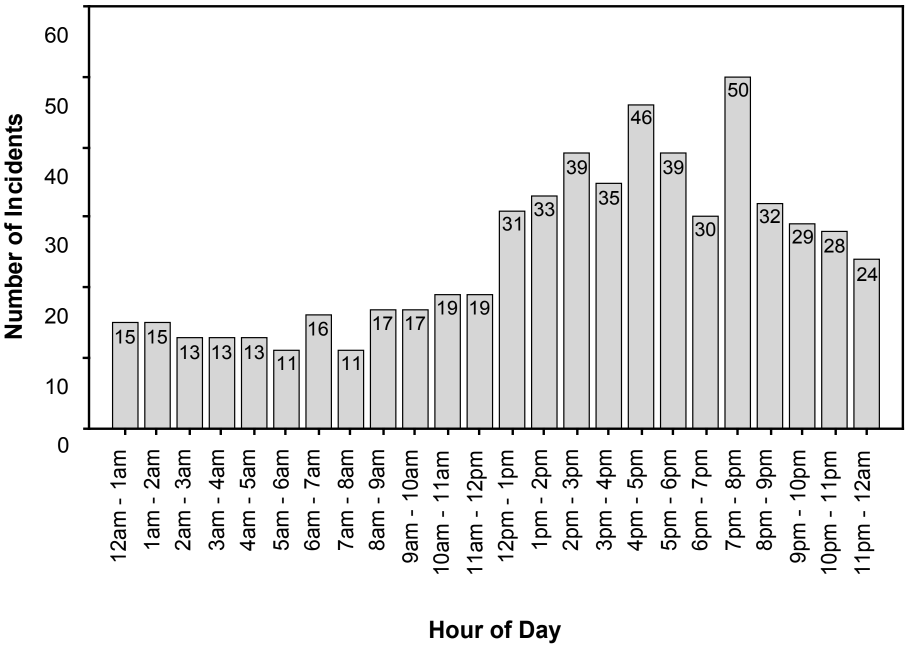


Exhibit 2-3
Fires by Day of Week - Canton - 1999

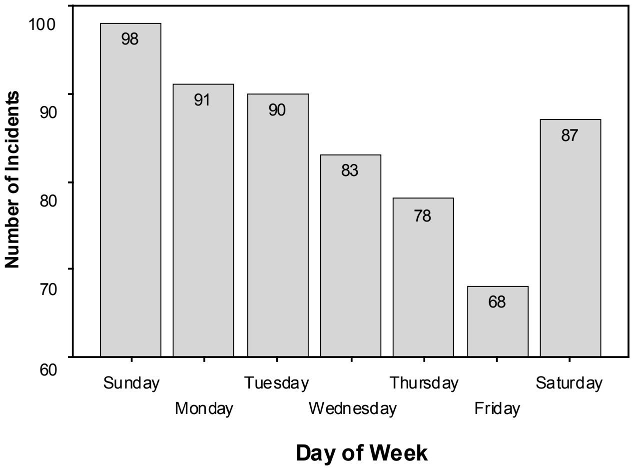
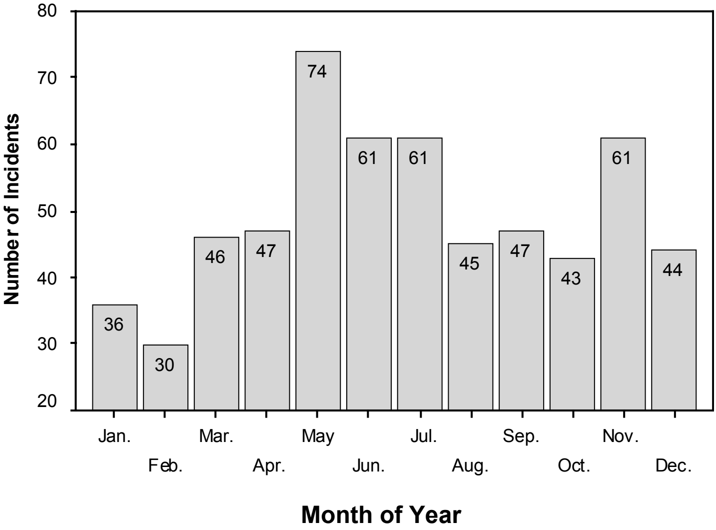


Exhibit 2-4
Fires by Month - Canton - 1999

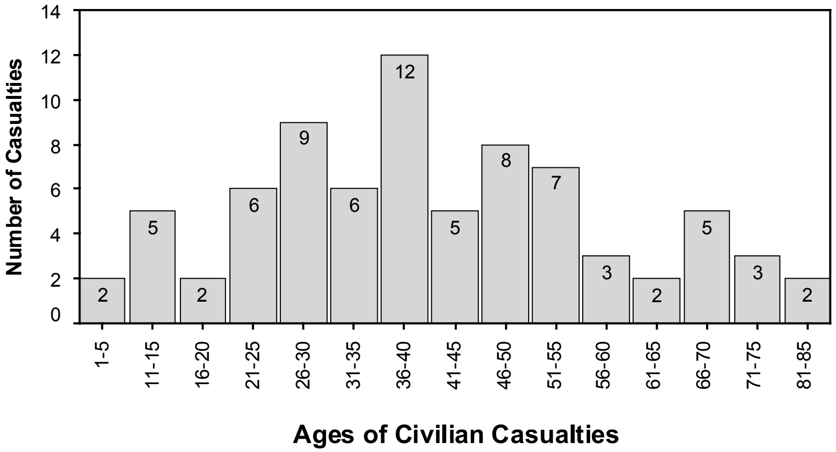


Example 2. Ages of Civilian Casualties. Suppose a fire chief is interested in developing a fire prevention program aimed at reducing civilian injuries and deaths. Descriptive data on civilian casualties is available from the NFIRS reports and there are a number of different descriptions that could be developed from the data. One of the most basic is descriptive data on the ages of civilian casualties.

Exhibit 2-5 shows the ages of civilians injured or killed in fires in Denver, Colorado, for 1999. Note that this distribution is considerably different from the previous histograms primarily because it does not have the same "smoothness." However, the five-year age groups show some interesting patterns. For example, the age group 36 to 40 accounts for the most civilian casualties, followed in frequency by 26 to 30 and 46 to 50, respectively. Also of interest is how the frequency takes a rather sudden drop for both the 16 to 20 and 56 to 60 age groups. Spikes in the data occur at the 26 to 30 and 36 to 40 year age groups. The exhibit also reveals several gaps in the data for ages 6 to 10 and 76 to 80. Due to these gaps at either end of the distribution, two outliers are created in the under five and 81 to 85 age groups.

Exhibit 2-5

Ages of Civilian Casualties - Denver - 1999



Note: Age was not provided for 7 casualties.
52% of the casualties were between 26 and 50 years old.

Spikes are high or low points that stand out in a histogram. **Gaps** are spaces in a histogram reflecting low frequency of data. **Outliers** are extreme values isolated from the body of data.

In histograms and other charts, it is sometimes useful to include comments and conclusions with the chart. In Exhibit 2-5, a note was provided that seven casualty records did not include age information and were therefore not included in the histogram. Other notes provide summary information on the data such as the percent of casualties between the ages 26 and 50 years old. Anyone studying the histogram could reach the same conclusion, but the summary saves time and effort.

Example 3. Response Times to Fires. Response times to fires are one of the most important data sets to study in fire departments. Many fire departments have objectives for average response times to fires and try to allocate personnel to achieve these response times. Exhibit 2-6 shows a frequency distribution for response times to fires in Boston, Massachusetts, in 1999.

Exhibit 2-6

Response Times - Boston - 1999

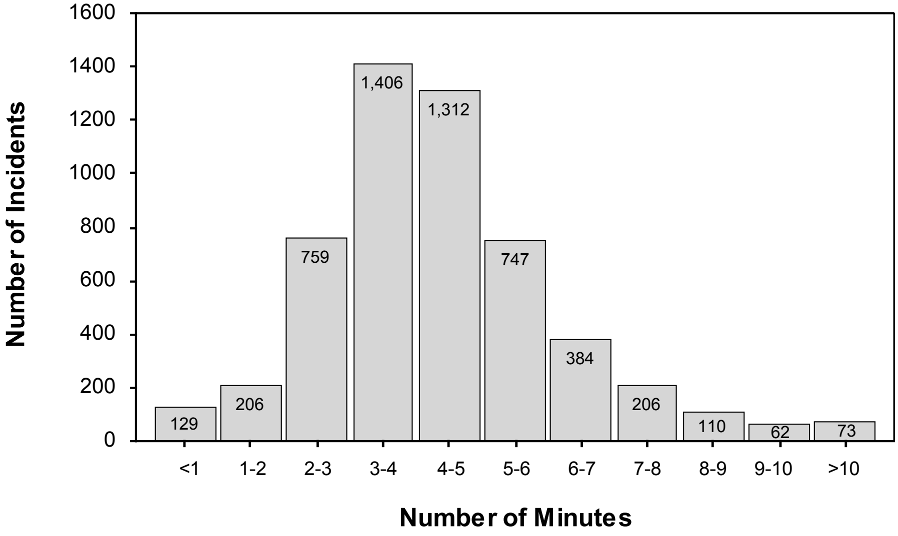
Response Time	Frequency
Less than 1 minute	129
1 to 2 minutes	206
2 to 3 minutes	759
3 to 4 minutes	1,406
4 to 5 minutes	1,312
5 to 6 minutes	747
6 to 7 minutes	384
7 to 8 minutes	206
8 to 9 minutes	110
9 to 10 minutes	62
10 to 11 minutes	18
11 to 12 minutes	15
12 to 13 minutes	15
13 to 14 minutes	10
14 to 15 minutes	5
15 to 16 minutes	6
16 to 17 minutes	2
17 to 18 minutes	0
18 to 19 minutes	1
19 to 20 minutes	1
Total Fire Calls	**5,394**

Notice in this example that the times are clustered at the low end of the distribution as would be expected since response times to fires are generally low for most fire departments.

Exhibit 2-7 provides a frequency histogram for this distribution. In this exhibit, we have combined the last few points into a category of 10 minutes or more. A histogram with the same shape as in this exhibit is said to be **skewed to the right** or **skewed toward high values**. What is meant by these terms is that the distribution is not symmetrical, but instead has a single peak on the left side of the distribution with a long tail toward the right. In fire departments, on scene time data (from time of arrival to time back in service) and fire dollar loss data also reflect values skewed to the right.

Exhibit 2-7

Response Times - Boston - 1999



Developing a Histogram

Making a histogram is relatively straightforward:

1. Choose the number of groups for classifying the data. In most cases, 5 to 10 groups are sufficient, but there are exceptions, such as histograms by hour of day. Sometimes the groups are natural, as in our exhibits by day of week and month. With other data, developing appropriate intervals will be necessary as was done in Exhibit 2-5 with the ages of civilian casualties.

2. Determine the number of events (fires, casualties, etc.) for each of the groups.

3. For data such as ages and response times, intervals usually need to be defined. For these intervals, convenient whole numbers should be chosen. That is, try to avoid the use of fractions in the groups and always make the intervals the same width. In Exhibit 2-5 intervals of 5 years were used for grouping the data. Data such as day of week do not require this step since their intervals are naturally defined.

4. Determine the number of observations in each group. Statistical packages are particularly useful in this step since they usually include routines for tabulating data.

5. Choose appropriate scales for each axis to accommodate the data. Again most statistical packages will do this with a default setting.

6. Display the frequencies with vertical bars.

Do not expect to get a histogram, or any other type of chart, exactly right on the first try. Several tries may be necessary before the look of the histogram is satisfactory.

The histograms presented in the previous section offer good examples of different characteristics for describing the data. In *Beginning Statistics with Data Analysis*, a text by Mosteller, et al., (1983), the following definitions of histogram characteristics are presented:

1. **Peaks and valleys.** The peaks and valleys in a histogram indicate the values that appear most frequently (peaks) or least frequently (valleys). Exhibit 2-2 shows clear peaks and valleys for incidents by hour of day.

2. **Spikes and holes.** These are high and low points that stand out in the histogram. In Exhibit 2-5, for example, there is a spike for the 36 to 40 age group, and a hole for the 16 to 20 age group.

3. **Outliers.** Extreme values are sometimes called outliers and are points that are isolated from the body of the data. In Exhibit 2-5, there are two outliers, in the under 5 and the 81 to 85 age groups.

4. **Gaps.** Spaces may reflect important aspects of a histogram. In Exhibit 2-5, there are gaps in the 6 to 10 and the 76 to 80 age groups.

5. **Symmetry.** Sometimes a histogram will be balanced along a central value. When this happens, the histogram is easier to interpret. The central value is both the mean (average) for the distribution and the median (half the data points will be below this value and half above).

Cumulative Frequencies

Two other types of distributions which will be important in later chapters are the **cumulative frequency** and the **cumulative percentage frequency**. A cumulative frequency is the number of data points that are less than or equal to a given value. A cumulative percentage frequency converts the cumulative frequency into percentages.

Example 4. With the data in Exhibit 2-6, we can calculate the cumulative frequency and cumulative percentages for the response time data from Boston, Massachusetts, found in Exhibit 2-8.

Exhibit 2-8

Cumulative Response Times - Boston - 1999

Response	Frequency	Cumulative Frequency	Cumulative Percent
Less than 1 minute	129	129	2.4
1 to 2 minutes	206	335	6.2
2 to 3 minutes	759	1,094	20.3
3 to 4 minutes	1,406	2,500	46.3
4 to 5 minutes	1,312	3,812	70.7
5 to 6 minutes	747	4,559	84.5
6 to 7 minutes	384	4,943	91.6
7 to 8 minutes	206	5,149	95.5
8 to 9 minutes	110	5,259	97.5
9 to 10 minutes	62	5,321	98.6
10 or more minutes	73	5,394	100.0
Total		**5,394**	**100.0**

The first entry under the “Cumulative Frequency” column is 129, which is the same as in the “Frequency” column. The second entry shows 335, which is 129 + 206, the sum of the first two entries in the “Frequency” column. By adding these two numbers, we can say that 335 incidents have response times less than 2 minutes. The next entry is 1,094 (129 + 206 + 759) and means that 1,094 incidents have response times less than 3 minutes. The cumulative frequencies continue in this manner with the last entry in the column always equal to the total number of incidents in the analysis.

The last column, labeled “Cumulative Percent” merely converts the cumulative frequencies into percentages. This step is accomplished by dividing each cumulative frequency by 5,394, which is the total number of incidents. The column shows that 2.4 percent of the incidents have response times less than 1 minute, 6.2 percent less than 2 minutes, 20.3 percent less than 3 minutes, etc.

In general, cumulative percentages describe data in “more than” and “less than” terms. We can conclude, for example, that about half the calls have response times of less than 4 minutes and about 95 percent have response times less than 8 minutes. Response times exceed 10 minutes in only about 1 percent of the calls.

Summary

A list of numbers is frequently the starting point for analysis. If the question of interest is for specific information, then the list of numbers serves the purpose. For example, Exhibit 2-1 is useful if we are asked about exactly how many fires occurred between 2 a.m. and 3 a.m., or if we want to know the exact difference between the busiest and the least busiest hour. On the other hand, Exhibit 2-1 is not very useful for determining the six busiest hours of the day.

Histograms provide a much better method for getting the feel of a list of numbers and answering several questions about relationships. The patterns in a histogram are especially important, such as high and low frequencies, and trends indicated by spikes, outliers, and gaps. Histograms give quick graphic representations of the data that otherwise would be hidden and hard to dig out of a table of numbers.

CHAPTER 3: CHARTS

Introduction

In this chapter we will extend beyond histograms to other types of charts. Histograms are only one of many different ways of presenting data. As an analyst, you must decide which type of chart best portrays the results you want to represent. A histogram may serve as the best vehicle in some cases, but other types of charts should be considered such as bar charts, line charts, pie charts, dot charts, and pictograms. Each of these will be discussed in this chapter.

Two questions to bear in mind throughout this process:

- What are the main conclusions from your analysis?
- What is the best way to display the conclusions?

As with the previous chapter, several sets of real fire data will be presented. You should study each example carefully and draw your own conclusions about the results. You may, in fact, disagree with what the book emphasizes or you may identify an aspect of the data that was overlooked. In either case, the point is to think about how you would present your viewpoints in a graphical format to a given audience. The audience may be an internal group of managers, an outside association or group of citizens, or even your own city or county council. The audience itself influences the type of chart that is selected.

Therefore, the first step is to determine the key results from the data. Once they have been identified, a selection of the best type of chart to convey them must be made. Often it is helpful to try different charts to determine the best presentation for a particular audience and data set.

Each of the following sections describes a different type of chart. At the end of the chapter, guidelines on selecting a type of chart suitable for different conclusions are presented.

Bar Charts

A **bar chart** is one of the simplest and most effective ways to display data.

In a bar chart, a bar is drawn for each category of data allowing for a visual comparison of the results. For example, the figures in Exhibit 3-1 give the causes of ignition (from NFIRS 5.0 codes) for the 12,600 structure fires in Chicago, Illinois, for 1999.

Interest in a list of this type usually centers on how the items compare to each other. What is the leading cause of ignition in structure fires? How do unintentional causes compare to intentional ones? How many causes are never determined?

Some results can be determined relatively easy from the list of numbers. For example, fires undetermined after investigation are clearly the leading cause of ignition followed by intentional, equipment failure, and unintentional, all close in number. The remaining three not reported, other, and act of nature account for less than one percent combined. While these comparisons can be made from the list, they require mental manipulations and are not easily made or retained in full.

Exhibit 3-1

Cause of Ignition for Structure Fires - Chicago - 1999

Cause of Ignition	Number	Percent
Intentional	2,771	22.0
Unintentional	2,583	20.5
Failure of Equipment or Heat Source	2,654	21.1
Act of Nature	29	.2
Cause, Other	40	.3
Not Reported	55	.4
Cause Undetermined After Investigation	4,468	35.5
Total	**12,600**	**100.0**

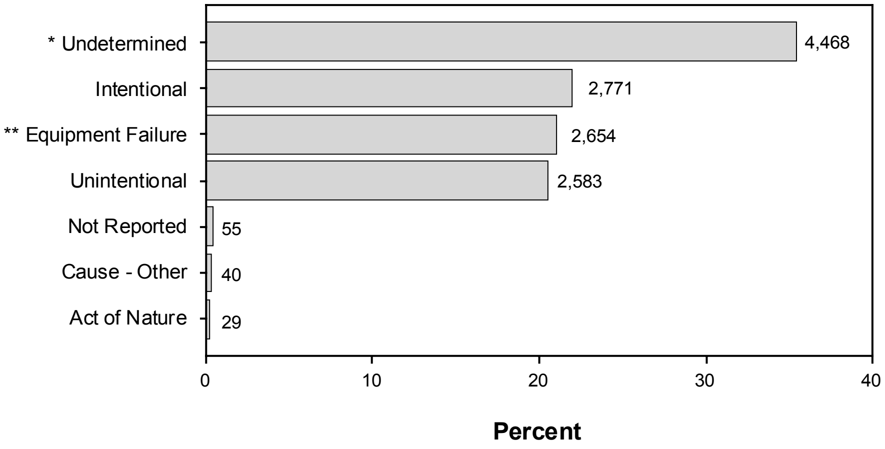
A bar chart overcomes these problems by presenting the data in frequency order as displayed in Exhibit 3-2. The horizontal dimension gives the percent, while the vertical dimension shows the category labels. The bars are presented in numerical order starting with undetermined after investigation as the most frequent. Each bar also contains the number of fires for that cause of ignition as additional information to the reader.

It should also be noted that the category “Cause Under Investigation” had no cases reported, but this fact is mentioned in a footnote since it is a listed option in the NFIRS module. Also in a footnote are the complete titles of two of the categories that were abbreviated in the table listing.

As a general rule, the horizontal dimension in a bar chart is numeric, such as percentages or other numbers, while the vertical dimension shows the labels for the items in a category. It is not always necessary to include numbers in each bar, especially if there is an accompanying table or list, but they can be useful to readers unfamiliar with the data. If the numbers are omitted from the chart, a total number should be provided either in the title or a footnote.

Exhibit 3-2

Cause of Ignition for Structure Fires - Chicago - 1999

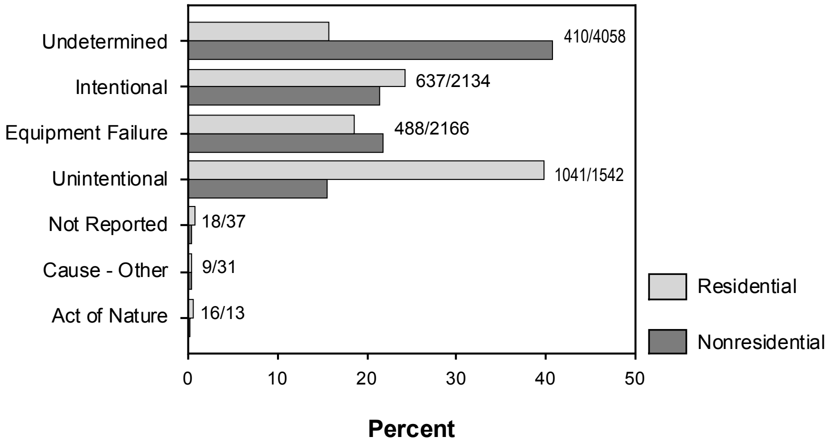


* Undetermined After Investigation/ ** or Heat Source Failure
No cases reported under "Cause Under Investigation" category

A **clustered bar chart** shows two categories in the same chart. In Exhibit 3-3, for example, the causes of ignition for structure fires in Chicago in 1999 are shown in a residential versus nonresidential format. One of the things the exhibit shows is that fires that are undetermined after investigation comprise over 40 percent of the nonresidential fires and only 16 percent of the residential ones. Interestingly, the chart also shows an almost exact ratio of 40 percent and 15 percent for unintentional causes of residential and nonresidential fires respectively. In addition, while the percents are close for residential and nonresidential fires under the unintentional and equipment failure categories, the numbers differ by 3 to 4 times due to the large difference in total fires between residential and nonresidential. The clustered or paired bar chart clearly shows the differences in ignition causes for these two types of structure fires.

Exhibit 3-3

Cause of Ignition Residential Versus
Nonresidential Fires - Chicago - 1999



Column Charts

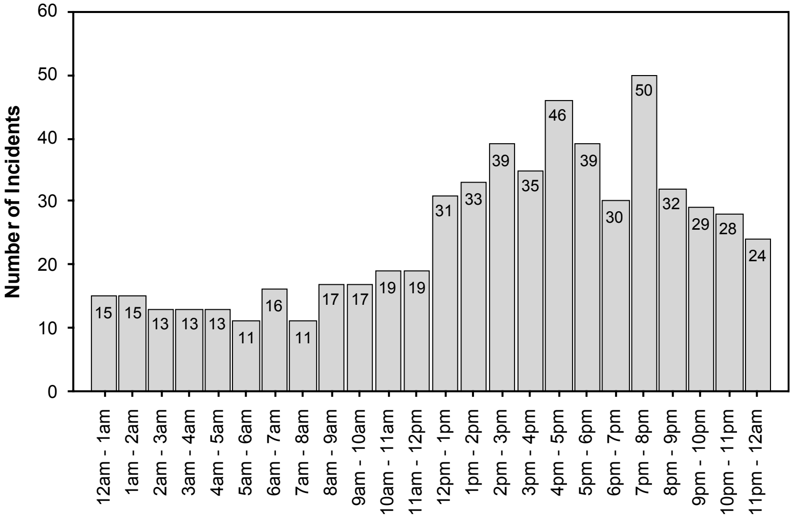
In Chapter 2 several column charts were displayed. For example, Exhibits 2-2, 2-3, and 2-4 showed Canton fires during 1999 by hour of day, day of week, and month respectively. These are all examples of **time series** presented as **column charts**.

Column charts of this type are particularly useful in demonstrating change over time. Where is the series increasing, decreasing, or staying about the same? If the analysis shows change over time, then column charts are particularly beneficial in presenting the changes.

As an example, the exhibit from Chapter 2 on fires by hour of day is repeated in Exhibit 3-4. By looking from left to right one can visualize the change in his or her mind. The horizontal scale shows the hours, but is not really needed to get a “feeling” for the changes. Calls are low in the early morning hours, then increase in the afternoon and evening hours.

Exhibit 3-4

Fires by Hour of Day - Canton - 1999



Hour of Day

Column charts show frequency distributions that allow for easy identification of trends and other characteristics, particularly with time series data. The horizontal scale defines the natural groupings for the chart and the columns give the frequencies.

Another good application of column charts is to show comparisons across sets of data. Exhibit 3-5 lists the causes of ignition from Exhibit 3-3 Residential versus Nonresidential Fires. Due to their small numbers for illustrative purposes, the Not Reported, Causes - Other, and Acts of Nature categories have been combined into “Other.” Comparisons between the venues are not easy because the totals differ so much. Nonresidential fires total just under 10,000 while residential have 2,619. A simple way to overcome this problem is to develop percentages.

By converting the residential and nonresidential figures to percentages, as shown at the bottom of the exhibit, a better comparison can be made. The percentages for both add up to 100 percent. While there are many conclusions that could be drawn from these percentages, the key ones are:

- Intentional, Equipment Failure, and Other account for about the same percentages in both residential and nonresidential fires.
- Unintentional fires account for 40 percent of the residential fires, while 40 percent of the nonresidential fires fall into the Undetermined category.

Exhibit 3-5

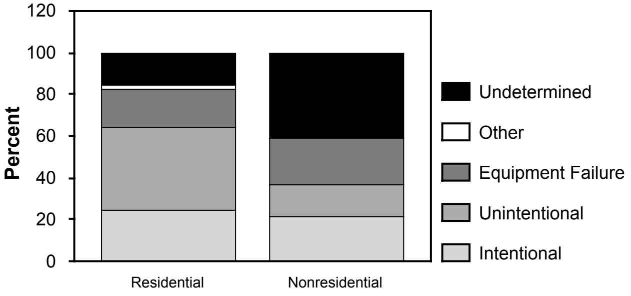
Comparison of Causes of Ignition in Residential Versus. Nonresidential Fires

Cause of Ignition	**Residential**	**Nonresidential**
Intentional	637	2,134
Unintentional	1,041	1,542
Failure of Equipment or Heat Source	488	2,166
Cause Undetermined After Investigation	410	4,058
Other	43	81
Total	**2,619**	**9,981**
Cause of Ignition	**Residential**	**Nonresidential**
Intentional	24.3%	21.4%
Unintentional	39.7%	15.4%
Failure of Equipment or Heat Source	18.6%	21.7%
Cause Undetermined After Investigation	15.7%	40.7%
Other	1.6%	.8%
Total	**100.0**%	**100.0%**

To display this result, **stacked column charts** were developed as shown in Exhibit 3-6 using the percentages for each cause of ignition. The columns have the same height since they both total 100 percent. The colors highlight the differences among the causes of ignition. The results just discussed should be clear from the exhibit.

Exhibit 3-6

Comparison of Causes of Ignition by Percent - Chicago - 1999



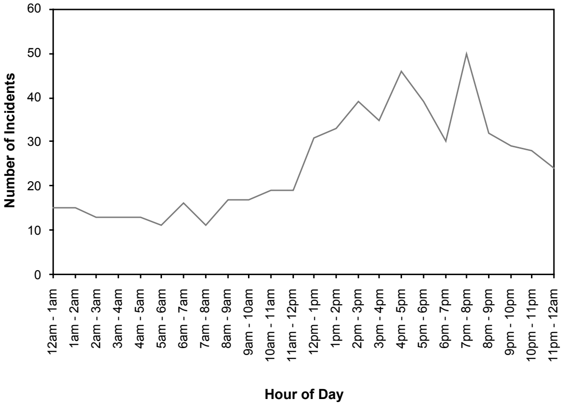
Note: "Other" causes include Not Reported and Act of Nature.

Line Charts

Effective presentation of time series data also may be developed from line charts. Exhibit 3-7 shows a line chart of fires by hour of day for Canton, Ohio, previously displayed as a histogram in Exhibit 2-2. The line chart immediately highlights the jump in fires from a sharp rise in the early afternoon until a peak at around 8:00 p.m. Many statisticians believe that a line chart is the clearest way for showing increases, decreases, and fluctuations in a time series.

Exhibit 3-7

Fires by Hour of Day - Canton - 1999



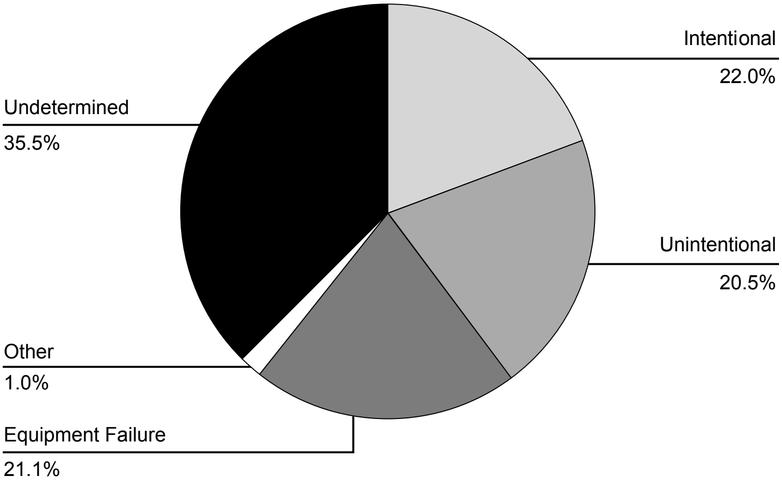
Pie Charts

A **pie chart** is an effective way of showing how each component contributes to the whole. In a pie chart, each wedge represents the amount for a given category. The entire pie chart accounts for all of the categories.

For example, Exhibit 3-8 shows the causes of ignition for structure fires in the Chicago Fire Department for 1999 divided into undetermined, equipment failure, intentional, unintentional, and other. The percentages are included with each wedge label. Although the percentage numbers are not necessary, they aid in comparisons of the wedges. The pie chart emphasizes the fact that the largest percentage of fire causes is undetermined. In addition, intentional, unintentional, and equipment failure all account for about the same percent of the causes.

Exhibit 3-8

Cause of Ignition for Fires - Chicago - 1999



Note: "Other" causes include Not Reported and Act of Nature

In developing pie charts, one should follow these rules:

- Convert data to percentages.
- Keep the number of wedges to six or less. If there are more than six, keep the most important five and group the rest into an “Other” category.
- Position the most important wedge starting at the 12 o’clock position.
- Maintain distinct color differences among the wedges.

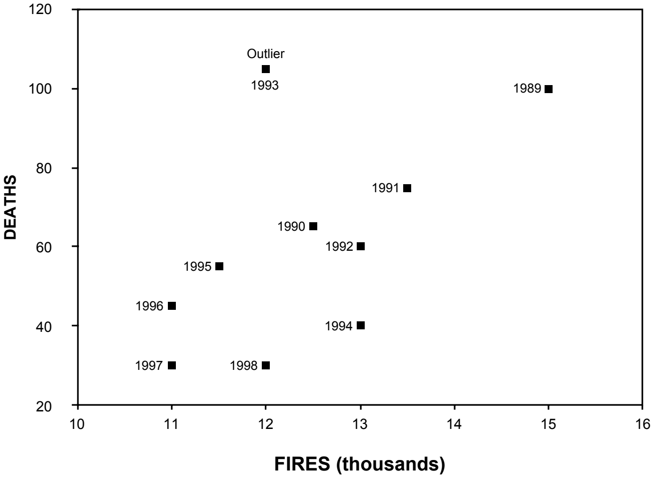
While pie charts are popular, they are probably the least effective way of displaying results. For example, it may be hard to compare wedges within a pie to determine their rank. Similarly, it takes time and effort to compare several pie charts because they are separate figures.

Dot Charts

Dot charts or **scatter diagrams** emphasize the relationship between two variables. For example, the 10-year trend in other residential fires from 1989 to 1998 was generally a decrease from a high in 1989 of 15,000 to a low of 11,000 by 1996. During these years a decrease in fire deaths also occurred. One would expect deaths to decrease with a decrease in fires, and it is this relationship that is depicted in Exhibit 3-9.

Exhibit 3-9

Fires and Deaths - 1989-1998



The exhibit is a dot chart for fires versus deaths for the 10 years from 1989 to 1998. Fires are along the horizontal or x axis, while deaths are along the vertical or y axis. The pattern is the important aspect of a dot chart, rather than the individual dots. The horizontal scale (x axis) should reflect the causation variable (independent variable) while the vertical scale reflects the resulting variable (dependent variable). That is to say that a decrease in fires (the independent variable) results in a decrease in fire-related deaths (the dependent variable).

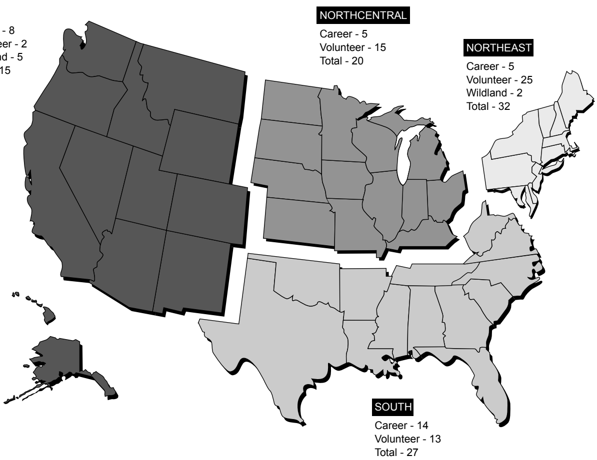
Another useful application of scatter diagrams is to identify outliers in the data. In Chapter 2, outliers were defined as points that are isolated from the body of the data. In Exhibit 3-9 there is a general pattern showing a decrease in deaths over time as fires decrease. However, while the decrease in fires pattern is maintained for 1993, the deaths shoot up to the highest for the period. Therefore, 1993 has many more fire-related deaths than expected, based on its amount of fires. This outlier from the general pattern can be useful in revealing an area of further data analysis that would account for this discrepancy from the rest of the data.

Pictograms

The final type of chart takes advantage of pictures to display data. Data by geographical areas, such as counties, census tracts, or fire districts can be presented on maps showing the boundaries of the areas. Exhibit 3-10, for example, shows firefighter deaths by region for 1997. Each region is broken down by career, volunteer, and, if applicable, wildland department.

Exhibit 3-10

Firefighter Deaths by Region - 1997



The key is that presentation in this manner is more effective than any listing of the death rates. It can be easily seen that:

- Career deaths in the south are two to three times more than in other regions.
- Volunteer deaths in the western region are a fraction of those in the rest of the country.

- The northeast has the most total deaths largely due to a volunteer death number that is almost double the next largest region.

Other pictograms for State and local data are easily imagined. At the State level, data from individual counties may be collected. A pictogram provides a good way of depicting the county data by taking a State map showing county boundaries and developing an exhibit similar to Exhibit 3-10. Similarly, for a local jurisdiction, such as a city or a county, there may be data for individual fire districts. A jurisdiction map with fire district boundaries may be an effective way of presenting the data.

Summary

In this chapter, six types of charts were presented: bar charts, column charts, line charts, pie charts, dot charts, and pictograms. The primary purpose of using any chart is to indicate conclusions more quickly and clearly than is possible with tables or numbers. It may be necessary to try several types of charts before the most appropriate one is found, but in a chart simplicity is the key. The message is what is important, so the chart form should not interfere with it.

As a quick reference guide on chart selection, the following is suggested:

- Use a **bar chart** with categorical data when the objective is to show how the items in a category rank. Most fire data are in categories, such as cause of ignition, property use, area of origin, type of injury, etc. These are reflected in the NFIRS modules.
- Use a **column** or **line chart** for data with a natural order, such as hours, months, or age groups. The chart will reflect the general pattern and indicate points of special interest, such as spikes, holes, gaps, and outliers.
- A **pie chart** is beneficial when the objective is to show how the components relate to the whole. It is recommended that the number of components be kept to six or less and that the forming of several pie charts for comparison purposes be avoided.
- A **dot chart** depicts the relationship between two variables. Generally, these variables are continuous rather than categorical. The pattern between the two variables is the important aspect for a dot chart.
- A **pictogram** is a pictorial representation of the data. Breakdowns by geographic areas, for example, are effectively shown by a pictogram.

This page intentionally left blank.

CHAPTER 4: BASIC STATISTICS

Types of Variables

For purposes of analysis, fire department variables can be divided into two types: qualitative variables and quantitative variables. Qualitative variables are defined as variables that are classified into groups or categories. For example, fires can be broken down into structure fires, vehicle fires, refuse fires, explosions, etc. Qualitative variables are also known as categorical variables (data) since they are not measured in quantity, but segregated into groups. Examples of categorical data in the fire service would include property use, cause of ignition, extent of flame damage, etc. Most categorical variables that will be used in fire data analysis will be found in the modules of the NFIRS system.

Quantitative variables always take on numerical values that reflect some type of measurement. Quantitative variables can be discrete (exact) or analog (continuous). An example of a discrete variable would be the number of days in the month or year (1 through 30 or 1 through 365), but no fractions of days. Whereas, time in hours, minutes, seconds, and infinite fractions of seconds would be analog or continuous. Other examples of quantitative variables would be the number of fires in a district over a period of time (discrete), the response time from alarm to arrival on the scene (analog), and the dollar losses of fires (discrete).

The distinction between a **variable** and **data** should be noted. A variable is a characteristic that varies or changes. Days of the week vary from Sunday through Saturday; months vary from January through December; and types of fires vary, such as structure fires, vehicle fires, residential fires, etc. Whenever observations are made on a variable, data are created to be analyzed. Each time an NFIRS report is completed, data for the variables listed are created. For example, by listing the day of week, hour of day, month, type of situation found, and values for all other applicable variables in the NFIRS Basic Incident module, data are created. The data then can be summarized in a variety of ways, such as tables, graphs, and charts. In this chapter, ideas about summarizing data will be extended by introducing six basic descriptive measures: the mean, mode, and median as well as the range, variance, and standard deviation.

Measures of Central Tendency

Measures of central tendency provide a single summary figure that best describes the central location of an entire distribution. The three most common measures of central tendency are the mode, the median, and the mean. Each will be defined and then the individual properties and uses for each will be discussed.

The **mode** is the value that occurs most frequently in a distribution. It is, therefore, easily recognized since no calculations are necessary.

The **mean** is also known as the arithmetic mean or average. However, since the term *average* is sometimes used indiscriminately for any measure of central tendency, it should be avoided. It is defined as the sum of all values in a distribution divided by the total number of values. For example, suppose that travel times to nine incidents are 3 minutes, 2 minutes, 4 minutes, 1 minute, 2 minutes, 3 minutes, 3 minutes, 4 minutes, and 3 minutes. Adding these travel times gives 25 minutes in total and dividing by 9 yields a mean travel time of 2.78 minutes.

The third measure of central tendency is the **median**, which is defined as the middle value (50th percentile) of a distribution. To determine the median, the data must be ordered. Using the nine travel times from the above example, they would look as follows if arranged in order: 1, 2, 2, 3, 3, 3, 4, 4. The median is the fifth or middle value, which is 3 minutes. There are four data values below and four data values above. In other words, 50 percent of the values lie on either side of the median, placing it at the 50th percentile.

If there had been an even number of data values, then the median would have been the mean of the two middle values. For example, if the onsite times for 10 fire incidents were 12, 15, 17, 25, 27, 29, 32, 35, 37, and 42 minutes, then the two middle values would be 27 and 29. Totaling them and dividing by 2 (calculating the mean value) results in a median value of 28. Again the median splits the values with five below and five above.

Properties and Uses for Measures of Central Tendency

The mode is the only measure of central tendency that can be used for qualitative data. This is really its only redeeming quality other than to serve as an additional qualifier for a distribution. The mode by itself is an unstable measure of central tendency. Equal size samples taken from a distribution are likely to have different modes. Further, on many occasions distributions have more than one mode (bimodal) which adds to the confusion.

The median is a better choice than the mode for a measure of central tendency. Unlike the mode it cannot be used with qualitative data, but with quantitative variables. The median on scene time for fires or the median dollar loss for fires can be determined. However, the “median type of fire” or the “median cause of ignition” has no meaning since these are qualitative variables. Responding to how many values lie above and below, but not to how far away, the median is less sensitive than the mean to the presence of a few extreme values.

Generally, the mean is the best choice for a measure of central tendency. Unlike the mode and the median, the mean is responsive to the exact position of each value in a distribution. It serves as a fulcrum point, balancing all of the values in a distribution. Consequently, the mean is very sensitive to extreme values (outliers) in a distribution. When a measure of central tendency needs to reflect the total of the values, the mean is the best choice since it is the only measure based on this quantity. Another of the more important characteristics of the mean is its stability over samples drawn from a distribution. This becomes especially important when further statistical computation is done.

Measures of Dispersion

While measures of central tendency provide a summary of the values in a distribution, measures of dispersion provide a summary of the variability or spread of the values in a distribution. Measures of dispersion express quantitatively the extent to which the values in a distribution scatter about or cluster together. The three main measures of dispersion are the range, variance, and standard deviation. As with the measures of central tendency, they will first be defined and then their properties and uses will be discussed.

The **range** is the most basic measure of dispersion. Its definition is simply the difference between the lowest and highest value in a distribution. For example, with the 10 onsite times used in the median discussion, the lowest value is 12 minutes and the highest is 42 minutes. Therefore, the range is 30 minutes.

Another measure of the variability of a distribution is the **variance**. In order to calculate the variance it is necessary to first obtain what is known as the deviation values of a distribution. The **deviation values** are the difference between the values in a distribution and its mean. Since the mean is the balance point of the values in a distribution, the total of the deviation values would be zero. Therefore, in order to calculate the variance, it is necessary to square the deviation values to eliminate the negative values.

To illustrate the calculation of a variance, the nine travel times used in the example for the mean will be used. In Exhibit 4-1 the mean of 2.78 has been subtracted from each individual travel time and the result squared.

Exhibit 4-1

Calculation of Variation

Travel Time	Travel Time - Mean (2.78)	Squared
1	-1.78	3.17
2	-.78	.61
2	-.78	.61
3	.22	.05
3	.22	.05
3	.22	.05
3	.22	.05
4	1.22	1.49
4	1.22	1.49
Total	**0.00**	**7.57**
Variance		**.95**

The middle column displays the amount of deviation from the mean for each point. The first deviation is -1.78 (1 minute minus 2.78 minutes), indicating that this travel time is 1.78 units from the mean and is to the left of the mean (since the sign is negative). Note that the sum of the middle column is zero; that is, the sum of the deviations from the mean is zero. In fact, an alternative definition for the mean is that it is the only number with this property.

In the right column is the square of each deviation. The sum of the squared deviations is 7.57 and the variance is obtained by dividing this sum by 8, which is one less than the total number of values. The reason for subtracting one from the total number of values will be discussed shortly. The variance from this calculation is then .95. Since the variance is small compared to the mean, it indicates that the values are close to the mean.

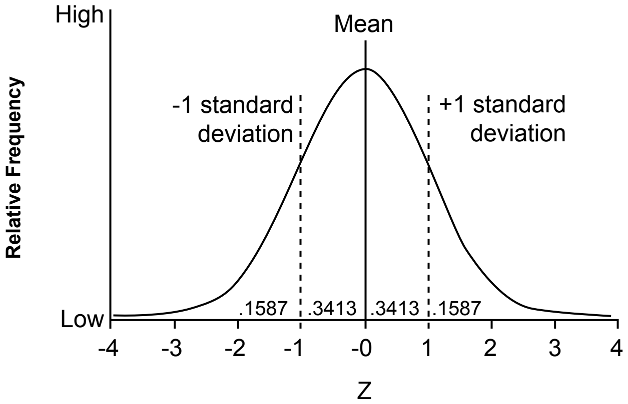
The final measure of dispersion is the **standard deviation**. It is obtained by taking the square root of the variance. In the current example, the standard deviation is .97, since this is the square root of .95. This means that the spread (variability) around the mean is not very large (in this case less than 1.0 compared to a mean travel time of 2.78 minutes). Therefore, the mean is a good descriptor of the data in this example.

Earlier in discussing the calculation for the variance, the sum of the squared deviations was divided by the number of values minus one. This was done to correct for a statistical error that results when using inferential statistics. If the distribution is the entire amount of values being considered, then dividing by that number is perfectly legitimate. However, if the distribution is merely a sample of a larger distribution, which it usually is, then a better representation of the entire population of values can be obtained by subtracting one from the sample distribution.

Normal Distribution and Standard Score

Unless there is a compelling reason otherwise, statisticians usually assume a **normal distribution** for any given set of values. As shown in Exhibit 4-2, a normal distribution is equally spread out in the general shape of a bell. In fact, it is known as the **bell curve**. In a normal distribution the mean, the median, and the mode are the same. Half the values are above the mean and half below. Most of the values, 68 percent, fall within one standard deviation on either side of the mean, within two standard deviations 95 percent, and within three standard deviations fully 99.7 percent of the distribution is represented.

Exhibit 4-2
Normal Distribution



By using a **standard score**, it is possible to compare values from different distributions on an equal basis. A standard score is a derived score that describes how far a given value in a distribution is from some reference point, typically the mean, in terms of standard deviation units. One of the most commonly used standard scores is the z score. Transforming the values of a distribution to z scores changes the mean to zero and the standard deviation to one, but does not change the shape of the distribution. For example, in the travel times used in Exhibit 4-1, a z score of one would be equivalent to a score of 3.75 minutes. That would be calculated by adding the mean of 2.78 to the standard deviation of .973. In another distribution of travel times with a different mean and standard deviation, a z score of one would be totally different. However, using the z scores they could be compared equally without distorting the original distributions.

Properties and Uses for Measures of Dispersion

The range is ideal for preliminary work or in circumstances where precision is not an important requirement. However, it is not sensitive to the total condition of the distribution since only the two outermost values determine its calculation. Therefore, the range is of little use beyond the descriptive level.

Since the variance is the mean of the squares of the deviation values of a distribution, it is responsive to the exact position of each value in a distribution. It can, therefore, be very important in inferential statistics because of its resistance to sampling variation. However, it is of little use in descriptive statistics because it is expressed in squared units.

The standard deviation, like the mean and the variance (from which it is derived), is responsive to the exact position of every value in a distribution. Because it is calculated by using deviations from the mean, the standard deviation increases or decreases as the individual values shift away from or toward the mean. Like the mean, it is influenced by extreme scores, especially with distributions that have a small amount of values. As the number of values increase, each individual value has less ability to shift the mean and the standard deviation. If the mean and the standard deviation of a distribution are known, a fairly accurate picture of the distribution can be obtained.

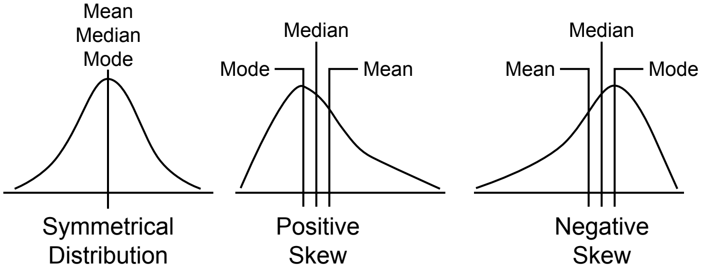
Once again, using the travel time example from Exhibit 4-1, the mean is 2.78 and the standard deviation .973. Assuming a normal distribution, one standard deviation from the mean in both directions should cover 68 percent of the values. In this case, values between 1.807 and 3.753 include 2, 2, 3, 3, 3, and 3. Since there are nine values in the distribution, the six values that fall within one standard deviation from the mean account for 67 percent. Considering its small size, that is an extremely accurate picture of the distribution. It is also a good example of how powerful the combination of the mean and the standard deviation can be. Each are the best measures of their type (central tendency and dispersion) and both are used extensively in more sophisticated statistical calculations.

Skewed Distributions

Even though statisticians assume a normal distribution *prima facie*, not all distributions are normal or symmetrical. As stated before, in a normal distribution the mean, median, and mode are all the same. However, this is not the case with skewed distributions. As shown in Exhibit 4-3, distributions can be skewed positively or negatively. In a **positive skew**, the extreme scores are at the positive end of the distribution. This exhibits the "tail" on the right side and pulls the mean to the right. Since the median and the mode are less responsive to extreme scores, they remain to the left of the mean. So in a positively skewed distribution, the mean has the highest value with the median in the middle and the mode the smallest value. Conversely, in a distribution with a *negative skew* the extreme values and the "tail" are at the negative end, the mean is the smallest value with the median in the middle and the mode the largest value.

Exhibit 4-3

Skewed Distributions



Central Limit Theorem

In order to perform statistical tests and analysis, statisticians rely on their assumption of a normal distribution. However, as we have seen, this is not always the case. Fortunately, there is a rule which allows them to make this assumption even when the distribution is not normal. The **central limit theorem** states that the sampling distribution of means increasingly approximates a normal distribution as the sample size increases. That is a distribution whose individual values are the means of samples drawn from the main distribution (population). The central limit theorem allows inferential statistics to be applied to skewed and otherwise normal distributions.

The central limit theorem is very powerful, and in most situations it works reasonably well with a sample size greater than 10 or 20. Thus, it is possible to closely approximate what the distribution of sample means looks like, even with relatively small sample sizes. The importance of the central limit theorem to statistical thinking cannot be overstated. Most hypothesis testing and sampling theory is based on this theorem.

While there is a mathematical proof for the central limit theorem, it goes beyond the scope of this text to present it. It is discussed here to show that there is a solid statistical base for assuming a normal distribution for the statistical tests used in inferential analysis of fire data. With the proper sample size, the results will be valid even if the population distribution is not normal.

CHAPTER 5: ANALYSES OF TABLES

Introduction

As was discussed previously, most fire data are qualitative (categorical) in nature. Examples of categorical data in the fire service would include property use, cause of ignition, extent of flame damage, etc. Since this type of data cannot be expressed in terms of the mean, median, and standard deviation, the number of each category can be used and listed in table form. It might be found, for example, that arson fires account for 56 percent of all structure fires, equipment failure 23 percent, and so on.

This chapter will provide techniques for analyzing tables developed from categorical data. This will include the development and interpretation of percentages for categorical data and the use of a nonparametric statistical test known as the chi-square. The chi-square is used to determine whether the percentage distribution from a table differs significantly from a distribution of hypothetical or expected percentages.

A **nonparametric** statistical test is one that makes little or no assumptions about the distribution. As stated previously, statisticians assume a normal distribution in their calculations. However, categorical data by nature are not described in this manner, i.e., mean, standard deviation, etc. Therefore, statistical tests that have certain parameters to their use would not be appropriate for this type of data. The chi-square was designed to be used without these parameters and as such is ideal for categorical data.

Describing Categorical Data

Summarizing a categorical variable is usually done by reporting the number of observations in each category and its percentage of the total. For example, consider Exhibit 5-1 for types of situations found in the fires of Lincoln, Nebraska, during 1999. These percentages are simple to calculate and easy to understand: 24.9 percent of the fires are structure fires, 26.9 percent are vehicle fires, and so on. As described in Chapter 2, the mode is the category with the largest number of data values. In this example, the mode is vehicle fires, totaling 175 fires.

Exhibit 5-1

Type of Situations Found - Lincoln Fires - 1999

Type of Fire	Number	Percent
Structure Fires	162	24.9
Outside of Structure Fires	44	6.8
Vehicle Fires	175	26.9
Trees, Brush, Grass Fires	166	25.6
Refuse Fires	88	13.5
Other Fires	15	2.3
Total	**650**	**100.0**

By way of comparison, Exhibit 5-2 shows the nationwide picture of types of situations found for fires. From a national perspective, structure fires accounted for 28.7 percent of the total, closely followed by trees, brush, and grass fires at 27.3 percent, and vehicle fires at 20.2 percent.

Exhibit 5-2

Types of Situations Found - Nationwide Fires - 1999

Type of Fire	Number	Percent
Structure Fires	523,000	28.7
Outside of Structure Fires	64,000	3.5
Vehicle Fires	368,500	20.2
Trees, Brush, Grass Fires	498,000	27.3
Refuse Fires	226,500	12.5
Other Fires	143,000	7.8
Total	**1,823,000**	**100.00**

Looking at these exhibits would prompt the question of whether the distribution of fires in Lincoln differs from the national picture. Some differences can be noticed by comparing percentages. For example, 26.9 percent of the Lincoln fires were vehicle fires, compared to 20.2 percent nationwide. Similarly, 2.3 percent of the Lincoln fires were other fires compared to 7.8 percent nationwide. It would, therefore, seem that the distribution of fires in Lincoln deviates from the national picture. However, a statistical test can be made to test this difference more precisely. The next section provides such a test.

The Chi-Square Test

The chi-square test (pronounced kī) is a statistical test designed to be used with categorical data. Like most statistical tests, it is stated in precise statistical language by defining a hypothesis to be tested. The use of the term **null hypothesis** is commonly seen. The null hypothesis is merely that there is no difference between the two distributions being compared. In this case, the null hypothesis would be that there is no statistical difference between Lincoln and the national percentages in the categories of fires in Exhibits 5-1 and 5-2. This is usually the way it is stated, that there is no difference. If a difference is found, the null hypothesis is rejected. It is sort of like innocent until proven guilty!

Although the chi-square test is conducted in terms of frequencies, it is best viewed conceptually as a test about proportions. To illustrate these ideas, it will be easier at this point to use a format that does not include fire data. Instead, consider a simple experiment where a die is thrown over and over again. The resulting data values are the number of dots showing after each throw. The number of dots varies between 1 and 6; that is, there are six possible outcomes. If a "fair" die is thrown a large number of times, one would expect each number of dots to show up one-sixth of the time. The chi-square test can be used with a certain degree of assurance to determine if, in fact, the die is "fair."

Suppose, for example, that a die is tossed 90 times and the results are as shown in Exhibit 5-3 below.

Exhibit 5-3
Results of Die Throws

Dots Visible	Number	Percent
One	16	17.8
Two	17	18.9
Three	12	13.3
Four	14	15.6
Five	17	18.9
Six	14	15.6
Total	**90**	**100.00**

If the die is a "fair" die, one would expect to have one dot turn up exactly 15 times (one-sixth of the total), two dots visible exactly 15 times, and so on. The actual results differ from these expected results as shown in Exhibit 5-4 below.

Exhibit 5-4
Actual and Expected Results

Dots Visible	Actual Number	Expected Number
One	16	15
Two	17	15
Three	12	15
Four	14	15
Five	17	15
Six	14	15
Total	**90**	**90**

To summarize, a die has been tossed 90 times and obtained the results shown in Exhibit 5-3. The null hypothesis is that the die is "fair," which means that there is no difference between the actual and the expected number of times each number of visible dots appear. The actual results are not the same as the expected either because the die is not "fair" or because of variations inherent in throwing a die only 90 times. The chi-square test will determine whether the actual results differ *significantly* from the expected results.

The following are the steps in performing the chi-square test:

1. Calculate the expected number for each category by multiplying the expected or population percentages by the total sample size. This calculation has already been performed as shown in Exhibit 5-4 with the "Expected Number" column.

2. For each category, subtract the expected number from the actual number, and then square the result.

3. Divide the results from step 2 by the expected number.

4. Sum the results from step 3. This is the calculated chi-square statistic. The larger this number, the more likely there is a significant difference between the actual and expected values.

5. Find the **degrees of freedom**, which is defined as the number of categories minus one. In the die example there are five degrees of freedom.

6. Obtain the **critical chi-square value** from the table in the Appendix by selecting the entry associated with the appropriate degree of freedom. Note that the table includes levels of significance from .05 to .001. Commonly the .05 level is used for most determinations. This indicates that results exceeding the critical value will be statistically significant 95 percent of the time. The other levels are used depending on how critical the results may be. For example, the more stringent .001 level is used in drug testing where lives may depend on the results.

7. If the computed chi-square statistic is greater than the critical value obtained from the table, the null hypothesis is rejected. Otherwise, the null hypothesis is accepted. Rejecting the null hypothesis means there is a significant difference between the two distributions. Conversely, accepting it means that the two distributions are essentially the same with differences due to sampling or random variations.

Exhibit 5-5 summarizes these steps for the die example. The “Difference” Column shows the difference between the expected and actual numbers. The “Squared Difference” is the square of the difference obtained by multiplying the number by itself. The right-most column is the squared difference divided by the expected number; for example, the first figure is .067 obtained from 1 divided by 15. The chi-square value is 1.34, which is the sum of the values in the last column.

Exhibit 5-5

Actual and Expected Results - Die Tossing Experiment

Dots Visible	Actual Number	Expected Number	Difference	Squared Difference	Divided By Exp.
One	16	15	1	1	.067
Two	17	15	2	4	.267
Three	12	15	-3	9	.600
Four	14	15	-1	1	.067
Five	17	15	2	4	.267
Six	14	15	-1	1	.067
Total	**90**	**90**			**1.340**
Chi-Square Value	1.34				
Degrees of Freedom	5.00				
Critical Value	11.07				

From the Appendix, the critical chi-square value for 5 degrees of freedom is 11.07. Since the calculated chi-square value of 1.34 is less, the null hypothesis is accepted. Therefore, the results from the 90 throws do not provide evidence that the die is unfair.

Degrees of freedom have been defined as the number of categories minus one. The rationale for determining degrees of freedom is that each category may be considered as contributing one piece of data to the chi-square statistic. These data are free to vary except for the last category, since it is determined already by what is left. It is, therefore, not free to vary. Thus, the values in all categories except one are free to vary. An illustration of this may be more helpful than an explanation. Suppose you were asked to name any five numbers. In response, you chose 25, 44, 62, 82, and 2. In this case there were no restrictions on the choices. There were five choices and five degrees of freedom. Now suppose you were asked to name any five numbers again. This time you chose 1, 2, 3, and 4, but were stopped at that point and told that the mean of the five numbers must be equal to 4. Now you have no choice for the last number, because it must be 10 (1+2+3+4+10=20 and 20 divided by 5 = 4). The restriction caused you to lose one degree of freedom in your choice. Instead of having 5 degrees of freedom as in the first example, you now have 5 minus 1 or 4 degrees of freedom. Each statistical test of significance has its own built-in degrees of freedom based on the number and type of restrictions it makes. The chi-square has one.

At this time, the question on whether the distribution of fires in Lincoln differs from the nationwide distribution of fires can be dealt with. It was noted that there were differences in some categories; for example, Exhibit 5-1 shows that vehicle fires account for 26.9 percent of the fires in Lincoln compared to 20.2 percent nationwide. Similarly, other fires account for 2.3 percent of the fires in Lincoln compared to 7.8 percent nationwide.

However, these are individual comparisons. The chi-square test allows all categories to be tested simultaneously. The null hypothesis is that “The percentage distribution of fires in Lincoln does not differ significantly from the nationwide picture.” If the calculated chi-square value is larger than the appropriate critical value in the Appendix, then the null hypothesis will be rejected, which would indicate that there was a significant difference. Otherwise, the null hypothesis would be accepted, indicating no significant difference in the two distributions.

Exhibit 5-6 shows the calculations using the information in Exhibits 5-1 and 5-2. The “Actual Number” column comes directly from Exhibit 5-1. To obtain the expected number, the percentages from Exhibit 5-2 are applied to the 650 Lincoln fires. For example, 28.7 percent of the nationwide fires were structure fires, which means we expect 28.7 percent of the 650 fires in Lincoln to be structure fires. This calculation yields 186.6 fires (28.7 percent times 650 fires).

The "Difference" column gives the difference between the actual and expected numbers and the next column is the squared difference (the difference multiplied by itself). The last column is the squared difference divided by the expected value. The calculated chi-square value is the sum of the column, which is 63.8.

In this example, there are six categories of fires, which means there are five degrees of freedom. From the Appendix, the critical chi-square value is 11.07. Since the calculated chi-square value of 63.8 is greater than the critical value, the null hypothesis is rejected. The conclusion is that the distribution of fires in Lincoln differs significantly from those nationwide. As stated before, the table in the Appendix lists the critical values for chi-square at various levels. For the purposes of this type, the .05 level is sufficient, which means that the difference will be significant 95 percent of the time or at the 95 percent confidence level. In this particular example, the obtained chi-square value far exceeds the critical value for even the .001 level, which is 20.52. This means that it is significant 99.9 percent of the time with a chance of error of only one tenth of a percent! In most comparisons, this level of confidence is rarely obtained.

Exhibit 5-6

Actual and Expected Results - Lincoln Fires - 1999

Type of Fire	**Actual Number**	**Expected Number**	**Difference**	**Squared Difference**	**Divided by Expected**
Structure	162	186.6	-24.6	605.16	3.2
Outside	44	22.8	21.2	449.44	19.7
Vehicle	175	131.3	43.7	1,909.69	14.5
Grass	166	177.4	-11.4	129.96	0.7
Refuse	88	81.3	6.7	44.89	0.6
Other	15	50.7	-35.7	1,274.49	25.1
Total	**650**	**650.0**			**63.8**
Chi-Square Value	63.8				
Degrees of Freedom	5.0				
Critical Value	11.07				

Some of the rationale behind the chi-square statistic may be helpful in understanding what it is actually reporting. The dynamics of what contributes to the chi-square value are evident in Exhibit 5-6. For example, the largest difference (regardless of sign) between the actual and expected numbers is 43.7 for vehicle fires. Squaring the difference and dividing by the expected number gives 14.5, as shown in the last column. As can be seen, vehicle fires is only the third largest contributor to the chi-square value even though it has the largest difference between the actual and expected number of fires. The reason for this is that larger categories have greater leeway for numerical variations, since it requires more to account for the same amount of actual change than smaller categories with fewer numbers to begin with. This can readily be seen by looking at the top two categories in contribution weight to the chi-square value. Outside fires with a difference of 21.1 and other fires with a difference of –35.7 contribute 19.7 and 25.1 respectively for a total of 44.8 towards the 63.8 chi-square value. That is over 70 percent of the chi-square value made up of the two smallest categories! While the numerical difference is less than that of vehicle fires, the actual amount of change in those categories is greater, because the numerical difference is greater **proportionally** to the number of fires in those categories. This is why it was stated earlier that “although the chi-square test is conducted in terms of frequencies, it is best viewed conceptually as a test about proportions.”

Two-Way Contingency Tables

Up to this point, chi-square has been applied in cases with only one variable. It also has important application to the analysis of **bivariate** frequency distributions. By studying bivariate distributions with two categorical variables, the **statistical association** between the two variables can be measured. Association allows the gaining of information about one variable by knowing the value of the other. The strength of the association may run from none whatsoever to weak to quite strong. The chi-square measures its existence and strength.

Exhibit 5-7 will be used as the starting point to introduce contingency tables, statistical variable association, and the chi-square statistic’s role in measuring it. The NFPA’s Survey of Fire Departments for U.S. Fire Experience for 2001 was used to develop the exhibit. In order to facilitate the example, 5 of the 10 categories under “Nature of Injury” were eliminated. The “Type of Duty” category is as it appears in the original table.

There are five categories for location or “Type of Duty.” The first is responding to or returning from an incident. The next category, fireground, covers injuries while on site at a fire. Similarly, the third category, nonfire emergency, covers injuries while on site at all nonfire incidents. The training category would be used for any injuries sustained while the firefighter was training for his/her position. The last category covers all injuries not under the other categories, but while still on duty.

The nature of the injury also is divided into five categories. As mentioned above, there were originally 10 categories of injuries, but for simplicity's sake only the top 5 were used. They are: (1) burns, (2) smoke inhalation, (3) wounds/cuts, (4) strains/sprains, and (5) other on duty.

Exhibit 5-7

Firefighter Injuries - 2001 - Type of Duty and Nature of Injury

Type of Duty	**Nature of Injuries**					
	Burns	**Smoke Inhalation**	**Wounds/ Cuts**	**Strains/ Sprains**	**Other**	**Row Totals**
Responding to or from Fire	65	115	960	2,250	710	**4,100**
Fireground	3,255	2,580	9,210	16,410	3,635	**35,090**
Nonfire Emergency	185	185	2,440	8,025	2,725	**13,560**
Training	345	40	1,380	3,860	625	**6,250**
Other On Duty	245	105	2,780	8,185	2,495	**13,810**
Column Totals	**4,095**	**3,025**	**16,770**	**38,730**	**10,190**	**72,810**

Exhibit 5-7 shows that there were a total of 72,810 injured firefighters. The top left number means there were 65 firefighters who were burned either responding to or returning from a fire. Similarly, the number in the second row and fourth column indicates that there were 16,410 firefighters who suffered strains or sprains while on a fire incident. Further, this number is the mode of the contingency table.

Outside of identifying the mode and showing the relative position of each category within its variable, the numbers in the table do not relay much information. Next various percentages will be calculated from the table to provide more insight. Finally, a chi-square value will be calculated to measure the strength of the relationship between the two variables.

Percentages for Two-Way Contingency Tables

There are three ways to calculate percentages for two-way contingency tables of frequencies. Each way highlights a different feature of the table. More importantly, each provides a different interpretation of the data and leads to different conclusions about the relationship between the two variables. The three ways of calculating percentages are:

- Joint percentages
- Row percentages
- Column percentages

The type of percentage used depends upon where the emphasis needs to be placed. Joint percentages allow the direct comparison of table entries with each other. Row percentages concentrate on the individual rows of the table with percentages along the row totaling one hundred. Similarly, column percentages deal with the individual columns of the table with column totals equaling one hundred percent.

Joint Percentages

To calculate joint percentages, each entry in the table is divided by the overall total. Exhibit 5-8 shows the calculation for the counts from Exhibit 5-7. The lower left entry is simply 4,095 divided by 72,810, which equals 5.6 percent. This means that 5.6 percent of the total persons injured suffered burns. The sum of all the entries in the table is 100.0 percent.

More logical comparisons can be made with joint percentages than with just the raw counts. For example, the table shows that 22.5 percent of all injuries were sprains or strains that occurred while the firefighter was on the fireground. In a similar manner, only 1.3 percent of all injuries were wounds or cuts suffered by firefighters responding to or returning from a fire.

Exhibit 5-8 also provides important information from the row and column totals. For example, from the second row it is apparent that nearly half (48.2 percent) of all the injuries were sustained at a fireground. There are two ways to derive this percent. One is to add the five percentages across the row (4.5 + 3.5 + 12.65 + 22.5 + 5.0 = 48.2). The other is to divide the row total of 35,090 (from Exhibit 5-7) by 72,810 to yield the 48.2 percent. (Note: due to rounding, the numbers do not always add up exactly the same both ways.)

Similarly, column percentages provide information about the nature of the injuries involved. For example, only 5.6 percent of persons injured suffered from burns, 4.2 percent from smoke inhalation, 23 percent from wounds or cuts, 14 percent from other injuries, and over half (53.2 percent) from strains or sprains.

Exhibit 5-8

Firefighter Injuries - 2001 - Joint Percentages

Type of Duty	**Nature of Injuries**					
	Burns	**Smoke Inhalation**	**Wounds/ Cuts**	**Strains/ Sprains**	**Other**	**Row Totals**
Responding to or Returning from Fire	.09	.16	1.30	3.1	.98	**5.6**
Fireground	4.50	3.50	12.70	22.5	5.00	**48.2**
Nonfire Emergency	.25	.25	3.35	11.0	3.70	**18.6**
Training	.47	.06	1.90	5.3	.86	**8.6**
Other On Duty	.34	.14	3.80	11.2	3.40	**19.0**
Column Totals	**5.60**	**4.20**	**23.00**	**53.2**	**14.00**	**100.0**

While Exhibit 5-8 provides more insight into these two variables, it does not directly address other questions. For example, direct comparisons between burns and smoke inhalation injuries for any particular type of duty cannot be made. Similarly, comparisons between types of duty for any particular injuries cannot be made. In order to make these types of comparisons, row and column percentage calculations must be made.

Row Percentages

To convert table counts into row percentages, each entry in the table must be divided by its row total. Therefore, the top right entry is calculated by dividing 710 by 4,100. This indicates that 17.3 percent of the total firefighters responding to or returning from an incident sustained other types of injuries.

Exhibit 5-9

Firefighter Injuries - 2001 - Row Percentages

Type of Duty	Nature of Injuries					
	Burns	**Smoke Inhalation**	**Wounds/ Cuts**	**Strains/ Sprains**	**Other**	**Row Totals**
Responding to or Returning from Fire	1.6	2.8	23.4	54.9	17.3	**100.0**
Fireground	9.3	7.3	26.2	46.8	10.4	**100.0**
Nonfire Emergency	1.4	1.4	18.0	59.2	20.0	**100.0**
Training	5.5	.6	22.1	61.8	10.0	**100.0**
Other On Duty	1.8	.8	20.1	59.3	18.0	**100.0**

A table of row percentages allows for comparisons among the categories represented by the rows. The total for each row is 100 percent, and these figures appear on the right of the table as a reminder that row percentages are represented.

As indicated, 17.3 percent suffered other types of injuries when they were responding to or returning from an incident. A total of 1.6 percent had burn injuries, 2.8 percent had smoke injuries, 23.4 percent sustained wounds or cuts, and the vast majority, 54.9 percent, had sprains or strains.

Looking at the second row, which is for firefighters injured at fireground, a somewhat different picture emerges. Burns and smoke inhalations injuries account for 9.3 and 7.3 percent respectively. These are followed by wounds and cuts at 26.2 percent, sprains and strains at 46.8 percent, and 10.4 percent for the other category. Once again, these percentages total 100, accounting for all firefighters injured while at fireground.

Column Percentages

To convert table counts into column percentages each entry in the table must be divided by the total for its column. The top left entry would be calculated by dividing 65 by 4,095 yielding 1.6 percent. This indicates that only 1.6 percent of the firefighters who received burns were responding to or returning from a fire.

Exhibit 5-10

Firefighter Injuries - 2001 - Column Percentages

Type of Duty	Nature of Injuries				
	Burns	**Smoke Inhalation**	**Wounds/ Cuts**	**Strains/ Sprains**	**Other**
Responding to or Returning from Fire	1.6	3.8	5.7	5.8	7.0
Fireground	79.5	85.3	54.9	42.4	35.7
Nonfire Emergency	4.5	6.1	14.6	20.7	26.7
Training	8.4	1.3	8.2	10.0	6.1
Other On Duty	6.0	3.5	16.6	21.1	24.5
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**

The table of column percentages looks at a particular type of injury across the five types of duty. With burn injuries, it can be seen that most, 79.5 percent, occurred at fireground, 8.4 percent during training, 4.5 percent at nonfire emergencies, 6 percent on other types of duty, and only 1.6 percent while responding to or returning from fires. The “Other” injury category shows a very different breakdown. A total of 7 percent occurred while responding to or returning from fires, while 35.7 percent were sustained at the fireground. Nonfire emergencies accounted for 26.7 percent, followed by 24.5 percent for other on duty sites, and lastly 6.1 percent during training.

Selecting a Percentage Table

The choice of a percentage table depends on the uses of the data. Joint percentage tables are beneficial when the emphasis is on the interrelationship between the two variables in the table. For example, Exhibit 5-8 reveals that the combination of burns and fireground account for 4.5 percent of the total. This figure can be compared to other combinations in the table.

The row percentage table provides a way of emphasizing the type of injury for each type of duty. When a firefighter was responding to or returning from a fire, Exhibit 5-9 shows 54.9 percent of the injuries were from strains or sprains, 23.4 percent from wounds or cuts, 17.3 percent from other types of injuries, 2.8 percent from smoke inhalation, and 1.6 percent from burns. These are useful results by themselves, and can be compared to distributions in other rows.

The column percentage table emphasizes the type of duty for each type of injury. For burns only, Exhibit 5-10 shows that 79.5 percent were sustained at fireground, 8.4 percent were during training, 6 percent were on other duty, 4.5 percent were on a nonfire emergency, and 1.6 percent were responding to or returning from a fire. Interestingly, the percent of those burned while responding to or returning from a fire is the same for both the row and the column percentages.

Testing for Independence in a Two-Way Contingency Table

This section will use the chi-square test to determine whether the two variables in a two-way contingency table are independent of each other. As before, a step-by-step procedure for calculating the chi-square value will be provided. It should be noted that, with the chi-square calculations, as with the other calculations that have been performed, virtually all statistical packages automatically calculate the values. As can be seen, manual calculation is arduous and time consuming. Additionally, manual calculations are more subject to error. Therefore, a statistical package should be used whenever possible. However, the details of the computations are shown here in order to enhance the understanding of the underlying principles that are involved.

Before calculating the chi-square, however, a discussion of what is meant by independence is needed. Two variables are said to be **independent** if knowledge about one variable cannot be used in predicting the outcome of the other variable. In general, the **null hypothesis of independence** for a two-way contingency table is equivalent to hypothesizing that in the population the relative frequencies for any row (across the categories of the column variable) are the same for all rows, or that in the population of the relative frequencies for any column (across the categories of the row variable) are the same for all columns. So once again, the hypothesis to be tested by chi-square can be seen as one concerning proportions. For example, there are almost nine times as many injuries sustained on the fireground as there are responding to or returning from a fire, but if the type of duty is unrelated to the number of injuries sustained, then on a **proportional basis** the number of injuries should be the same for each type of duty.

Constructing a Table of Expected Values

In order to calculate the chi-square value, the expected values for each cell must be determined. The **expected values** are the counts that would occur if the two variables were independent. The first step in developing a table of expected values is to calculate the proportion of cases in each cell. This can be done by column or row. Using the column, divide each column total by the grand total. The proportion for the first column, burns, would be calculated as follows: 4,095 divided by 72,810 equals .056. Subsequent column proportions would be smoke inhalation .042, wounds and cuts .23, strains and sprains .532, and other .14. Note that the proportions are the same as the column total percentages calculated for the joint percentages in Exhibit 5-8.

It is a simple matter to calculate the expected cell frequencies from the expected cell proportions. For each cell, multiply the expected column proportion for that cell by the row total for that cell. For example, the cell representing firefighters responding to or returning from a fire who sustained burns would be: .056 (column proportion) times 4,100 (row total) equals 230.6. Exhibit 5-11 shows the results of the remaining expected values.

Exhibit 5-11

Firefighter Injuries - 2001 - Table of Expected Values

Type of Duty	**Nature of Injuries**					
	Burns	**Smoke Inhalation**	**Wounds/ Cuts**	**Strains/ Sprains**	**Other**	**Row Totals**
Responding to or Returning from Fire	230.6	170.3	944.4	2,180.9	573.8	**4,100**
Fireground	1,973.6	1,457.8	8,082.1	18,665.5	4,910.9	**35,090**
Nonfire Emergency	762.6	563.4	3,123.2	7,213.0	1,897.8	**13,560**
Training	351.5	259.7	1,439.5	3,324.6	874.7	**6,250**
Other On Duty	776.7	573.8	3,180.8	7,345.0	1,932.8	**13,810**
Column Totals	**4,095**	**3,025**	**16,770**	**38,730**	**10,190**	**72,810**

The table of expected values is the distribution of proportions in each row (or column) that would be expected in the absence of a dependent relationship between the two variables. In this case, it would mean that the expected values are those that reflect no relationship between the nature of injuries sustained and the type of duty performed. As stated before, the same results could have been obtained by calculating the row proportions and multiplying them by the column totals. It should also be noted that the row and column totals are exactly the same as the original table of counts. That is, the development of the expected value table preserves these totals. However, slight discrepancies may exist due to rounding of decimals.

Calculation of Chi-Square for a Two-Way Contingency Table

The chi-square value for a two-way contingency table is calculated similarly to the one done for a single categorical variable.

1. Develop the table of expected values, as shown in Exhibit 5-11 using the method discussed in the previous section.

2. For each table entry, subtract the expected value from the corresponding entry in the original table of counts, and then square the result. This difference measures the discrepancy between the actual counts and what would be expected if the variables were independent.

3. Divide the results from step 2 by the expected value. This adjustment allows for the larger expected numbers which are usually associated with larger deviations.

4. Sum the results from step 3. This is the chi-square statistic. The larger the chi-square statistic, the more likely that there is a significant statistical association between the two variables. However, the chi-square statistic also depends on the number of categories, which must be taken into account in the following steps.

5. Find the degrees of freedom, which is calculated for a two-way contingency table by multiplying the number of rows minus one times the number of columns minus one. In the current example, there are five rows and five columns. Therefore, the number of degrees of freedom is (5-1) x (5-1) = 16.

6. Compare the computed chi-square statistic from step 4 to the value in the chi-square table in the Appendix using the appropriate degrees of freedom. The table value is called the **critical chi-square value**.

7. If the computed chi-square statistic is greater than the critical value in the table, then the **null hypothesis of independence** is rejected and the variables are related. If the computed chi-square statistic is less than the critical value, the null hypothesis of independence is accepted and the variables are not related.

It is important to keep in mind that in a two-way contingency table the two variables are independent. If the null hypothesis is accepted, it means that knowing the value of one of the variables does not help in predicting the value of the other variable. In the current example, the null hypothesis is that the type of duty engaged in is independent of the nature of the injuries sustained.

Exhibit 5-12

Firefighter Injuries - 2001 - Table of Chi-Square Entries

Type of Duty	Nature of Injuries				
	Burns	Smoke Inhalation	Wounds/ Cuts	Strains/ Sprains	Other
Responding to or Returning from Fire	118.92	17.96	.26	2.19	32.33
Fireground	831.98	863.86	157.40	272.55	331.49
Nonfire Emergency	437.48	254.15	149.45	91.41	360.55
Training	.12	185.86	2.46	86.22	71.28
Other On Duty	363.98	383.01	50.50	96.07	163.53
Total Chi-Square Value = 5325.01	Critical Value = 26.3				

Exhibit 5-12 shows the chi-square entries for the two-way contingency table. These entries are the results after Step 3 above. The top left entry was calculated as follows: Exhibit 5-7 gave an actual count of 65 for this entry and Exhibit 5-11 gave an expected value of 230.6. Subtracting the expected value from the actual count yields a negative 165.6 (65 minus 230.6) and squaring that figure results in 27,423.36. Dividing this number by the expected value, 230.6, provides the chi-square value of 118.92. This value is then entered in Exhibit 5-12 and the procedure is repeated for each of the other entries. When all of the entries are calculated, they are all totaled. This total is the total chi-square value. In Exhibit 5-12, this total is 5,325.01. It is entered at the bottom of the table.

All that remains to test the hypothesis about the independence of the two variables, type of duty, and nature of injury is to compare the total chi-square value to the critical chi-square value from the Appendix. The critical chi-square value for 16 degrees of freedom is 26.3. Since the total chi-square value greatly exceeds this value, the null hypothesis is rejected. Therefore, there is a statistical association between type of duty and nature of injury.

As a cautionary note, remember that a significant outcome of the chi-square test is directly applicable *only to the data taken as a whole*. The chi-square obtained is inseparably a function of the (in this case) twenty-five contributions composing it, one from each cell. Therefore, it cannot be said whether one group is responsible for the finding of significance or whether all are involved.

CHAPTER 6: CORRELATION

Introduction

This chapter deals with the concept of correlation for continuous (quantitative) data. Correlation is a statistical measure which indicates the degree to which one variable changes with another variable. For example, calls for Emergency Medical Services (EMS) generally increase with population growth. That is, as population increases, more medical service calls would be expected. This would indicate a positive correlation between population and EMS calls. The correlation measures the strength of the association between the two variables.

If there is a correlation between two variables, then predictions better than chance can be made from an individual score (or whatever is being measured) on one variable to its predicted score on the correlate variable. Any problem in correlation requires two pairs of corresponding scores, one for each variable. Generally, the greater the association (correlation) between two variables, the more accurately a prediction can be made on the standing in one variable from the standing in the other.

The chapter starts with the scatter diagram illustrated in Chapter 3 and proceeds with a discussion of the correlation coefficient. Next a typical calculation of a correlation is presented for demonstration purposes. The chapter concludes with a discussion of the applicability and uses of a correlation and mention of other types of correlation.

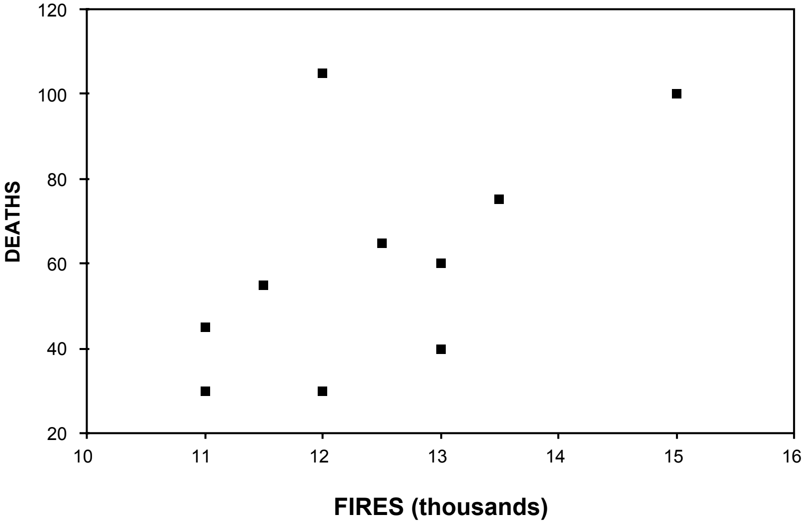
Scatter Diagram

Exhibit 6-1 shows a scatter diagram presented in Chapter 3 on the number of fire deaths and the number of other residential fires for the 10-year period 1989 to 1998. The horizontal axis gives the number of fires (in thousands) and the vertical axis gives the number of deaths. It can be seen from the exhibit that deaths are higher with greater numbers of fires. The general trend is clear even though the pattern is not perfect. The term “not perfect” refers to the fact that the points do not fall on a straight line.

Exhibit 6-1

Fires and Deaths - 1989-1998

Other Residential



With relationships depicted in this manner, the usual terminology is to label one variable as the **independent** variable, and the other as the **dependent** variable. In the case of Exhibit 6-1, “Fires” serves as the independent variable and “Deaths” as the dependent variable. Obviously, the number of fires influences the number of fire-related deaths; the more fires there are the greater number of fire deaths. This represents a positive correlation, since the increase in the independent variable is accompanied by an increase in the dependent variable.

It is important to emphasize two points about correlations. First, correlations assume an underlying linear relationship, that is, a relationship that can be best represented by a straight line. It should be noted, however, that not all relationships are linear. There are, for example, curvilinear relationships where the points on a scatter diagram cluster about a curved line. Secondly, while correlation can be used for prediction, it does not imply causation. The fact that two variables vary together is a necessary, but not a sufficient, condition to conclude that there is a cause-and-effect connection between them. A strong correlation between variables is often the starting point for further research.

Correlation Coefficient

The correlation coefficient measures the strength of association between two variables. The term “correlation coefficient” is used by most statisticians, but is the same as the more commonly used correlation. The correlation is always between -1 and +1. A correlation of exactly -1 or +1 is called a perfect correlation, and means that all the points fall on a straight line. A correlation of zero indicates no relationship between the variables, and would be represented on a scatter diagram as random points with no discernable direction. As a correlation coefficient moves from zero in either direction, the strength of the association between the two variables increases.

As stated before, a positive correlation means that, as the independent variable increases, so does the dependent variable. In a negative correlation, as the independent variable increases, the dependent variable decreases. The sign of the correlation indicates direction, not magnitude. Magnitude is indicated by the size of the number regardless of the sign. Therefore, a correlation of -.82 is greater than a correlation of +.63.

To summarize the relationship between a scatter diagram and the correlation coefficient, the correlation coefficient is a number that indicates how well the data points in a scatter diagram “hug” the straight line of best fit. With perfect correlations, all the data points fall exactly on a straight line that summarizes the relationship, and the value of the coefficient is +1 or -1. When the association between the two variables is less than perfect, the data points show some scatter about the straight line that summarizes the relationship as in Exhibit 7-1, and the absolute value (regardless of sign) of the correlation coefficient is less than 1. The weaker the relationship, the more scatter and the lower the absolute value of the correlation coefficient.

Another important point to know is that correlations are not arithmetically related to each other. For example, a correlation of .6 is not twice as strong as a correlation of .3. Although it is obvious that a correlation of .6 reflects a stronger association than a correlation of .3, there is no exact specification of the difference. Subsequently, there is no relationship between correlations and percentages.

In order to make direct comparisons between correlations, the correlation coefficient must be converted to a **coefficient of determination**. The coefficient of determination is the square of the correlation multiplied by 100. This yields the percentage of association between the two variables. For example, a correlation of .50 would indicate a 25 percent (.50 times .50 equals .25 times 100 equals 25) association between variables. A perfect correlation of 1.00 would be equal to a 100 percent coefficient of determination. So a correlation of 1.00 is four times as strong as a correlation of .50, not twice as strong, as might appear from comparing the correlations directly.

Additionally, the differences between successive correlation coefficient values do not represent equal differences in degree of relationship. For example, the difference between a correlation of .40 and .50 does not represent the same difference as that between correlations of .90 and 1.00. This can be seen more clearly by examining the coefficients of determination and their corresponding correlations in Exhibit 6-2. There is more than double the difference between correlations of .90 and 1.00 than between .40 and .50 when the corresponding coefficients of determination are compared.

Exhibit 6-2

Relationship Between Correlations and
Coefficients of Determination

Correlation Coefficient	Coefficient of Determination
1.00	100%
.90	81%
.80	64%
.70	49%
.60	36%
.50	25%
.40	16%
.30	9%
.20	4%
.10	1%
.00	0%

Calculating the Correlation

Today many pocket calculators include a program to calculate the correlation coefficient. Additionally, virtually all statistical software packages calculate the various types of correlations. However, for those who must calculate a correlation by hand, and in order to show what factors make the coefficient positive or negative and what factors result in a high or low value, the deviation-score method will be used.

Exhibit 6-3 shows the number of fires and civilian fire deaths for the 10-year period from 1991 to 2000. The correlation between these two variables will be computed using the deviation-score method. The most widely used correlation formula is the Pearson. Its full name is the **Pearson product-**

moment correlation coefficient. There are other types of correlations suited for special situations, but the Pearson is by far the most common. In fact, when researchers speak of a correlation coefficient without being specific about which one they mean, it may safely be assumed they are referring to the Pearson product-moment correlation coefficient. The term **moment** is borrowed from physics, and refers to a function of the distance of an object from the center of gravity. With a frequency distribution, the mean is the center of gravity and, thus, deviation scores are the moments. As it will be shown, the Pearson correlation is calculated *by taking the products of the paired moments*.

Exhibit 6-3

Total United States Fires and Civilian Fire Deaths 1991 - 2000

Year	Fires (thousands)	Deaths
1991	2,041.5	4,465
1992	1,964.5	4,730
1993	1,952.5	4,635
1994	2,054.5	4,275
1995	1,965.5	4,585
1996	1,975.0	4,990
1997	1,795.0	4,050
1998	1,755.0	4,035
1999	1,823.0	3,570
2000	1,708.0	4,045
Sum	**19,034.5**	**43,380**

As can be seen in Exhibit 6-3, fires tended to decrease over the 10-year period, while civilian fire deaths seem to have no obvious pattern overall (though the last 4 years have an apparent decrease). From this, it would seem that there is little association between the variables that should result in a low correlation.

The computation of the Pearson correlation using the deviation-score method is illustrated in Exhibit 6-4 and summarized in the following steps:

1. List the pairs of scores in two columns. The order in which the pairs are listed makes no difference in the value of the correlation. However, if one raw score is shifted, the one it is paired with must be shifted as well.

2. Find the mean for the raw scores of each variable.

3. Convert each score in both variables to a deviation score by subtracting the respective mean from each.

4. Calculate the standard deviation for both variables. Since the deviation scores are already done, they need only to be squared and summed. Divide each of these totals by the number of pairs (in this case 10) and take the square root of each.

5. Multiply each pair of deviation scores, known as the cross-product, and total the results.

6. Next multiply the two standard deviations by each other and multiply that result by the number of pairs (10).

7. Divide the results of Step 5 by the results of Step 6. The result is the Pearson product-moment correlation coefficient.

8. Square this for the coefficient of determination.

Exhibit 6-4

Deviation Score Calculation for Pearson Correlation Coefficient

Year	Fires - Mean	Deaths - Mean	Fires - Mean Squared	Deaths - Mean Squared	Cross Product
1991	+138.05	+127	19,057.8	16,129	+17,532.35
1992	+61.05	+392	3,727.1	153,664	+23,931.60
1993	+49.05	+297	2,405.9	88,209	+14,567.85
1994	+151.05	-63	22,816.1	3,969	-9,516.15
1995	+62.05	+247	3,850.2	61,009	+15,326.35
1996	+71.55	+652	5,119.4	425,104	+46,650.60
1997	-108.45	-288	11,761.4	82,944	+31,233.60
1998	-148.45	-303	22,037.4	91,809	+44,980.35
1999	-80.45	-768	6,472.2	589,824	+61,785.60
2000	-195.45	-293	38,200.7	85,849	+57,266.85
Sum	0	0	135,448.2	1,598,510	+303,759.00
Mean	Fires 1,903.45	Deaths 4,338		Correlation Coefficient	Coefficient of Determination
S.D.	116.382	399.814		+.653	42.6%

The correlation obtained in Exhibit 6-4 is relatively high as demonstrated by the coefficient of determination of 42.6 percent. This indicates a measure of relationship between the variables. It does not mean that the relationship is necessarily causal. For example, a high positive correlation probably exists between the amount of beer consumed and the amount of automobile accidents over each year from 1900 to the present. Rather than believe that beer consumption and the number of auto accidents are causally related, however, it is more reasonable to suggest that some condition such as an increase in population accounts for the increase in both beer consumption and automobile accidents.

Since the correlation is positive, it means that as the amount of fires increase/ decrease the number of deaths increases/decreases as well. While on the surface this would seem intuitive, as with the beer/accident example there can be other conditions that would account for the common variance. For example, an increase in fires would be expected as the population and buildings and residences increased. On the other hand, as knowledge and use of fire safety programs and procedures increased over time, the number of fire deaths would be expected to go down. The point is that there are usually many alternate and rational explanations for changes other than a causal one between two simultaneously changing variables.

The next step after obtaining a correlation that shows there is a relationship, is to use it as a predictor. This is done by defining the straight line that the data points cluster around, known as the **regression line**. The regression line is defined algebraically and the formula is used to make the predictions. The predictions become more reliable as the correlation increases. A discussion of the regression method is beyond the scope of this handbook, but is mentioned here to give a fuller meaning to the correlation coefficient.

Other Types of Correlations

While the Pearson correlation is by far the most commonly used, there are other types of correlations derived directly or indirectly from the Pearson. These correlations are used with data that are not continuous and quantitative as with the Pearson. Several of them are presented here with a brief description of their use. Details of their computation and use can be found in some of the texts cited earlier.

1. **Rank-order correlation.** Sometimes it is useful to categorize data by ranking. The largest gets a rank of 1, the second largest a rank of 2, and so on. When both variables consist of ranks, a rank-order correlation coefficient is calculated. It is sometimes called the Spearman rank-order correlation. It is found merely by applying Pearson's procedure to the ranks.

2. **Biserial correlation.** The biserial correlation is suited to cases in which one variable is continuous and quantitative and the other *would* be, except that it has been reduced to just two categories. For example, if the correlation between the number of fires and whether or not the number of civilian fire deaths was above or below the median. This would require the use of the biserial technique, since one variable is continuous and the other is expressed dichotomously.

3. **Point biserial correlation.** This would be used as in the biserial, except that the second variable is qualitative and dichotomous and could not be expressed as continuous and quantitative. For example, a correlation between the number of fires and the number of male and female civilian deaths.

4. **Phi coefficient.** This is the Pearson correlation coefficient for two variables that are both qualitative and dichotomous.

5. **Partial correlation.** The partial correlation shows what the Pearson correlation coefficient between two variables would be in the absence of one or more other variables. For example, with the correlation of fires to deaths the relationship each has to the passage of time could account for the change in each rather than a relationship to each other. By doing a partial correlation between fires and deaths for each month within a given year, time would be held constant. The resulting correlations would reflect a truer picture of the relationship between fires and deaths.

There are other variations of correlations used for determining variable relationships with different circumstances, but these cover most of what is likely to be needed. As stated before, all of these tools along with the ones discussed in the previous chapters are readily available in various statistical packages. Most of them walk the user through the process with clear understandable directions. The purpose of manually calculating these statistics was to give a fuller understanding of what was being done. This should make it easier to interpret the results from using a statistical package. It will also serve as a good foundation for any further study with statistical texts and course work.

APPENDIX: CRITICAL VALUES OF CHI-SQUARE

Level of Significance

df	.05	.025	.01	.005	.001
1	3.84	5.02	6.63	7.88	10.83
2	5.99	7.38	9.21	10.60	13.82
3	7.81	9.35	11.34	12.84	16.27
4	9.49	11.14	13.28	14.86	18.47
5	11.07	12.83	15.09	16.75	20.51
6	12.59	14.45	16.81	18.55	22.46
7	14.07	16.01	18.48	20.28	24.32
8	15.51	17.53	20.09	21.95	26.12
9	16.92	19.02	21.67	23.59	27.88
10	18.31	20.48	23.21	25.19	29.59
11	19.68	21.92	24.73	26.76	31.26
12	21.03	23.34	26.22	28.30	32.91
13	22.36	24.74	27.69	29.82	34.53
14	23.68	26.12	29.14	31.32	36.12
15	25.00	27.49	30.58	32.80	37.70
16	26.30	28.85	32.00	34.27	39.25
17	27.59	30.19	33.41	35.72	40.79
18	28.87	31.53	34.81	37.16	42.31
19	30.14	32.85	36.19	38.58	43.82
20	31.41	34.17	37.57	40.00	45.31
21	32.67	35.48	38.93	41.40	46.80
22	33.92	36.78	40.29	42.80	48.27
23	35.17	38.08	41.64	44.18	49.73
24	36.42	39.36	42.98	45.56	51.18
25	37.65	40.65	44.31	46.93	52.62
26	38.89	41.92	45.64	48.29	54.05
27	40.11	43.19	46.96	49.65	55.48
28	41.34	44.46	48.28	50.99	56.89
29	42.56	45.72	49.59	52.34	58.30
30	43.77	46.98	50.89	53.67	59.70

This page intentionally left blank.

REFERENCES

Aczel, Amir D. *Statistics: Concepts and Applications*. Chicago: Irwin, 1995.

Coxon, Anthony P.M. *Sorting Data: Collection and Analysis*. Thousand Oaks, CA: Sage Publications, 1999.

Devore, Jay, and Roxy Peck. *Statistics*: *The Exploration and Analysis of Data*. 4th ed. Pacific Grove, CA: Brooks/Cole, 2001.

Freedman, David, Robert Pisani, and Roger Perves. *Statistics*. 3rd ed. New York: W.W. Norton, 1998.

Gilbert, Nigel G. *Analyzing Tabular Data*: *Loglinear and Logistic Models for Social Researchers*. London: UCL Press, 1993.

Harrison, Tyler R., Susan E. Morgan, and Tom Reichert. *From Numbers to Words: Reporting Statistical Results for the Social Sciences*. Boston: Allyn and Bacon, 2002.

Jaffe, A.J., Herbert F. Spirer, and Louise Spirer. *Misused Statistics*. 2nd ed. New York: M. Dekker, 1998.

Lewis-Beck, Michael S. *Data Analysis: An Introduction*. Thousand Oaks, CA: Sage Publications, 1995.

Morgan, Charles J., and Andrew F. Siegel. *Statistics and Data Analysis: An Introduction*. 2nd ed. New York: J. Wiley, 1996.

Mosteller, Frederick, Stephen E. Fienberg, and Robert E.K. Rourke. *Beginning Statistics with Data Analysis*. Reading, MA: Addison-Wesley Publishing Company, 1983.

Newton, Rae R., and Kjell Erik Rudestam. *Your Statistical Consultant: Answers to Your Data Analysis Questions*. Thousand Oaks, CA: Sage Publications, 1999.

Stephens, Larry J. *Schaum's Outline Theory and Problems of Beginning Statistics*. New York: McGraw-Hill, 1998.

Zelazny, Gene. *Say It With Charts*. 4th ed. New York: McGraw-Hill, 2001.

REFERENCES

This page intentionally left blank.

REFERENCES

Baton, George. “The Wait for an Ambulance.” *JEMS*, December 1986.

Bryson, M. Strategic Planning for Public and Non-Profit Organizations. (2011).

Center of Public Safety Excellence. (2008).

Chaiken, W., and E. Ignall. Fire Department Deployment Analysis: A Public Policy Analysis CaseStudy, The Rand Fire Project. (1979).

Comparing Quantitative and Qualitative Research. Experiment Resources. (2009).

Fires detected by the Moderate Resolution Imaging Spectroradiometer (MODIS). http://activefiremaps.fs.fed.us.

http://www.eoearth.org/article/Location,_distance,_and_direction_on_maps (used with permission).

Karter, Michael. “Fire Loss in the United States During 2010.” (September 2011).

Lawrence, Cortez. “Optimal Staffing for Fire Attack.” (September 1991).

The Lost Sequel, One True Media.

National Fire Incident Reporting System.

National SAR Committee, Catastrophic Incident Search and Rescue (CIS). (November 2007).

NFPA Annual Fire Experience Survey.

Pidwirny, Michael. The Encyclopedia of Earth. Location, Distance, and Direction on Maps. (October 12, 2006).

USDA Forest Service. Active Fire Mapping Program. Remote Sensing Applications Center.

Webster’s online dictionary.

This page intentionally left blank.

GLOSSARY

This page intentionally left blank.

GLOSSARY

Attributes	Pieces of information, often stored in a table as part of a GIS layer, about a given geographic feature.
CAD	Computer-aided design.
Coordinates	Used to describe the location of an object on the Earth's surface.
FEMA	Federal Emergency Management Agency.
GARS	Global Area Reference System.
GIS	Geographic Information System.
GPS	Global Positioning System.
MODIS	Moderate Resolution Imaging Spectroradiometer.
NEMSIS	National Emergency Medical Service Information System; EMS counterpart to National Fire Incident Reporting System (NFIRS).
NTSB	National Transportation Safety Board.
Remote Sensing	The collection of information about an object or location without physically touching it.
SAR	Search and Rescue.
Standards of Cover	A concept that allows for analysis of current deployment of resources as compared to the risk assessment in the community.
Strategic Planning	A deliberative, disciplined approach to producing fundamental decisions and actions that shape and guide what an organization (or other entity) is, what it does, and why strategic planning may be thought of as a "way of knowing" intended to help leaders and managers discern what to do, how and why.
USNG	United States National Grid.