

# NFA: Introduction and Application of Data for Fire and Emergency Services

## Lesson 3: Big, Small, and Raw Data

### Learning Objectives

Upon completion of this lesson, you will be able to:

- Discriminate among big data, small data, and raw data.
- Identify validity and reliability of data sources.
- Identify standard data sources relevant to fire and emergency services.
- Import raw data into Excel.

This lesson should take approximately 45 minutes to complete.

### Resources

- National Fire Academy courses website
- U.S. Fire Administration website
- Center for Public Safety Excellence – 21<sup>st</sup> Century Fire and Emergency Services

### Preparing for the Lesson

Before we learn about various data, please read pages 13-15 of the 21st Century Fire and Emergency Services white paper – Section titled “Critical Issue C: Robust Use of Data.”

Reference: Center for Public Safety Excellence & International City/County Management Association (2020). *21st Century Fire and Emergency Services* [White Paper].

### Introduction to Big, Small, and Raw Data (1 of 2)

Think about all the data that is collected by your agency.

- Are the numbers and types increasing or decreasing over the years?

The numbers and types are likely increasing each year.

Think back when you started your career.

- What were the types and the amount of data being collected by your agency?
- What was done with them?
- Have they changed over the span of your career?

The types and the amount of data have likely exploded in most cases.

Before learning to analyze and manage data, you need to first understand why the data is being collected. Without the complete and full understanding of the reason, you may inaccurately examine the data or examine the wrong data, ultimately leading to inaccurate or irrelevant results.

## Introduction to Big, Small, and Raw Data (2 of 2)

Think about these questions:

- How does your agency measure the changes in raw numbers?
- How does your agency show the effectiveness of the accomplishments as an emergency organization?
- Why is this important?

More and more, the policy makers and our stakeholders want to know not only what we are doing as an agency, but what our plan is to measure and show improvement in efficiency metrics.

It is important to understand the difference between outputs and outcomes.

## Outputs vs. Outcomes (1 of 3)

### Outputs

Think about an agency's typical end-of-the-year report on activities. Generally, it will include numbers and types of responses, number of training hours by topic, and perhaps public education contacts. There are many additional data that may be reported, but let's take the three examples mentioned and examine what they really tell us:

- Responded to 2,622 requests for medical assistance.
- Trained 3,562 staff hours on emergency medical procedures.
- Conducted 342 presentations to the public.

It certainly looks like this agency was busy, but what do these numbers really tell us? They show the outputs of various efforts – what the agency performed. However, what these numbers cannot tell is us is how well they performed in these areas.

Response times are an output. They show how fast or slow the department arrived, but what was accomplished as a result is not measured. Fast response time does not necessarily lead to positive impact.

Put yourself on the receiving end – would you care more about how fast the response was or the quality of the response?

## Outputs vs. Outcomes (2 of 3)

### Outcomes

If we really want to see an agency's effectiveness, or the difference it made in the community, then we need to develop outcome-based metrics or data. These outcome-based data must be collected, analyzed, and shown to indicate the effectiveness rather than just the efficiency of the efforts.

For example, in a request for medical assistance – what was the outcome? In other words, what difference did the agency make in that response?

Cardiac survival rate is a classic metric that all agencies should be recording. The collected data allows you to track changes to the metrics, showing the effectiveness of your program. If the data shows a direction to a negative outcome, then you need to analyze the data to determine the cause of

ineffectiveness. Conversely, an increased effectiveness also warrants a similar analysis, so that the efforts that caused positive outcome can be continued.

### Outputs vs. Outcomes (3 of 3)

#### **Outcomes**

By using outcomes as a reporting mechanism, we can allow our policy makers and stakeholders to be informed of and determine how well an agency is performing. It can also serve as a baseline to present the impact of changes in service delivery, whether they are positive or negative.

In any emergency services organizations, you should be tracking the performance of your team using some of these measures.

### Definitions – Big Data, Small Data, and Raw Data

Let's examine the differences among big, small, and raw data.

#### **Big Data**

Big datasets are large data that may have several thousand entries and cover long periods of time. Both the National Fire Incident Reporting System (NFIRS) and the National Emergency Medical Services Information System (NEMSIS) are examples of large datasets.

Lists of data can be exported, as either a text (.txt) or comma-separated values (.csv) file and then converted into a spreadsheet format (e.g., Microsoft Excel) or other database formats. Once it is in the converted format, you can control how that data will be analyzed and reported, if you know the capabilities of the program you are using.

Microsoft Excel, as of 2020, has 16,384 columns and 1,048,576 rows available on each spreadsheet. That means it has over 17 billion cells available on each worksheet. While Excel is usually a small flat file, Microsoft Access is a relational data base that combines data held in different tables, and allows for deeper and more efficient analysis of larger files.

#### **Small Data**

Small dataset can have less than 500 entries in data fields. However, there is no set number that says any set of data is necessarily big or small. If a dataset fits on the screen or a couple of tabs, it could be considered a small dataset. The actual data manipulation using the various tools found in Microsoft Excel or other spreadsheet database programs is the same whether the dataset is large or small. The advantage of a small dataset is usually found in its ease of analysis. Extraneous data from a database can easily be removed to only allow for those data elements that need to be analyzed to be viewed.

#### **Raw Data**

Raw data is not formatted or organized prior to performing any type of data analysis.

This format is used often, because it's easy to export raw data by creating relatively smaller files and allows for an easy and efficient transfer of data between different programs.

### Knowledge Review

Which of the following best describes raw data?

- Data that is not formatted or organized prior to performing any type of data analysis
- Large data that may have several thousand entries and cover long periods of time
- Dataset with less than 500 entries in data fields
- Data found in the NFIRS and the NEMSIS

**Correct Answer:** Data that is not formatted or organized prior to performing any type of data analysis

**Additional Information:** Raw data is not formatted or organized prior to performing any type of data analysis. Raw data are easy to export and transfer between different programs.

Revisit “Definitions – Big Data, Small Data, and Raw Data” for more information.

### Importing and Exporting Data

In this course, we’ll practice using a standard set of data and provide you with an opportunity to manipulate it from a CSV file to an Excel file. You will then perform several options to both analyze and present the data.

Before starting this activity, download and open the file 1 - Fire in the US 1993 – 2017 Ready for Excel File Conversion.csv using notepad. If not set as the default application for CSV files, you may have to right-click on the file name, select “Open with,” then select “Notepad”.

This dataset is much simpler than the larger one reviewed earlier, and the content of the data can be identified easily. This is a list of fire deaths and fire injuries taken from various versions of USFA’s Fire in the United States reports. Let’s try converting this file to something that we can work with.

#### Activity 3.1: Converting Raw File into a Spreadsheet Format (1 of 7)

1. Open the Excel program.
2. Select cell A1. From the Data ribbon, select the pull-down arrow on the first icon on the left, “Get Data.” Select “From File”, then select “From Text/CSV.”

#### Activity 3.1: Converting Raw File into a Spreadsheet Format (2 of 7)

3. From the Import Data dialog box, select and open the 1 - Fire in the US 1993 – 2017 Ready for Excel File Conversion.csv file, then select the “Import” button.

#### Activity 3.1: Converting Raw File into a Spreadsheet Format (3 of 7)

4. A preview of the selected csv file will appear in a dialog box. At the bottom right, select the “Load” button.

#### Activity 3.1: Converting Raw File into a Spreadsheet Format (4 of 7)

5. The data from the CSV raw file should now be populated in the spreadsheet. On the right side, Queries & Connections information may display automatically. If so, you can select the small “x” button on the right hide this information for now.

#### Activity 3.1: Converting Raw File into a Spreadsheet Format (5 of 7)

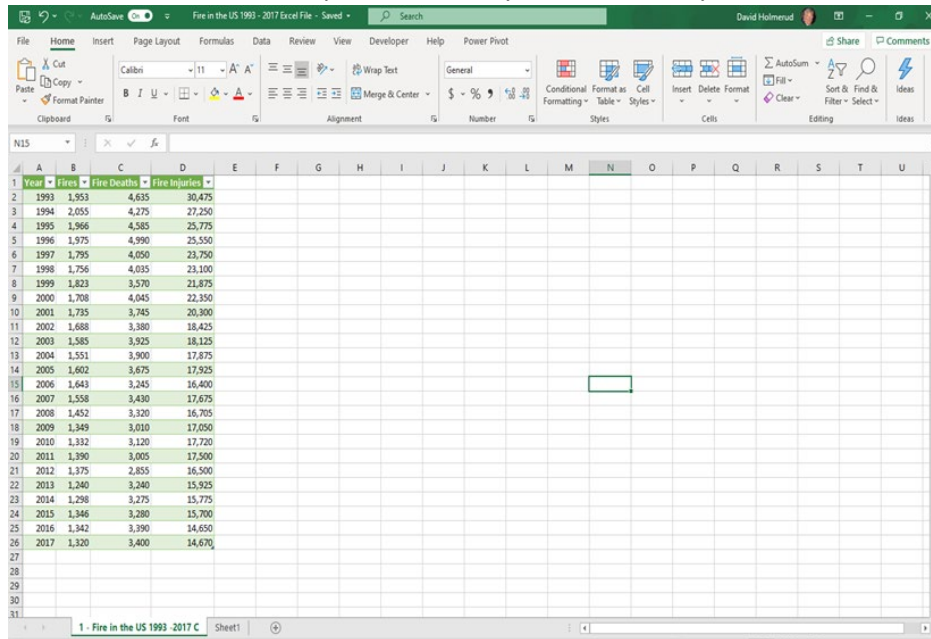
6. From the File ribbon, select “Save As,” then select the “Browse” option. In the “Save As” pop-up window that appears, change the file name to “2 - Fire in the US 1993 – 2017 Excel File” and select the “Save” button.

### Activity 3.1: Converting Raw File into a Spreadsheet Format (6 of 7)

7. We now have a table of the Number of Fires (in thousands), Fire Deaths, and Fire Injuries for a 25-year period from 1993 to 2017. One final formatting needs to be completed – Select and highlight all the values, except the “Years” from cell B2 through D26. Once highlighted, right-click (or Shift+F10 for Windows or Fn + up arrow + F10 for Mac) anywhere in the highlighted area. Select “Format Cells.” The Format Cells dialog box will appear. Under “Category”, select the “Number” option. Change the “Decimal places:” to 0 and place a checkmark in the box next to “Use 1000 Separator (,).”

### Activity 3.1: Converting Raw File into a Spreadsheet Format (7 of 7)

8. Here is what the finished spreadsheet may look like. Save your file.



Year	Fires	Fire Deaths	Fire Injuries
1993	1,953	4,635	30,475
1994	2,055	4,275	27,250
1995	1,966	4,585	25,775
1996	1,975	4,990	25,550
1997	1,795	4,050	23,750
1998	1,756	4,035	23,100
1999	1,823	3,570	21,875
2000	1,708	4,045	22,350
2001	1,775	3,745	20,300
2002	1,688	3,380	18,425
2003	1,585	3,925	18,125
2004	1,551	3,900	17,875
2005	1,602	3,675	17,925
2006	1,643	3,245	16,400
2007	1,558	3,430	17,675
2008	1,452	3,320	16,705
2009	1,349	3,010	17,050
2010	1,332	3,120	17,720
2011	1,390	3,005	17,500
2012	1,375	2,855	16,500
2013	1,240	3,240	15,925
2014	1,298	3,275	15,775
2015	1,346	3,280	15,700
2016	1,342	3,390	14,650
2017	1,320	3,400	14,670

**Alt Text:** Screenshot of the final converted spreadsheet which includes a column for the years ranging from 1993 to 2017; number of fires, fire deaths, and fire injuries in each of those years.

### Knowledge Review

According to the converted data, how many U.S. fire deaths were there in year 2000?

- 1708
- 3925
- 4045
- 15775

**Correct Answer:** 4045

**Additional Information:** Raw data is not formatted or organized prior to performing any type of data analysis. Raw data are easy to export and transfer between different programs.

Revisit “Definitions – Big Data, Small Data, and Raw Data” for more information.

### Activity 3.2: Converting a Spreadsheet File into a CSV Format (1 of 4)

Now that we have completed converting a raw CSV file into a Spreadsheet (Excel) format, let's try converting it back to a CSV file.

1. From the File ribbon, select "Export," then "Change File Type." Select CSV (Comma delimited), and then "Save As."

### Activity 3.2: Converting a Spreadsheet File into a CSV Format (2 of 4)

2. The "Save As" dialog box should appear. At the end of the file name that is automatically filled out in the File name field, add the word "Conversion." The file name should now read, "3 - Fire in the United States 1993 – 2017 Excel File Conversion Back to CSV" and the 'Save as type' should be defaulted to CSV (Comma delimited). If a warning box shows up select, "OK".
3. Open the Notepad application (commonly found under Windows Accessories).

### Activity 3.2: Converting a Spreadsheet File into a CSV Format (3 of 4)

4. Within the Notepad application, select "File," then "Open." Select the 3 - Fire in the United States 1993 – 2017 Excel File Conversion Back to CSV Excel file. You may need to change the Text Documents (.txt) option in the lower right to "All Files" in order to view the file to open.

### Activity 3.2: Converting a Spreadsheet File into a CSV Format (4 of 4)

5. Your conversion back to a .csv file is complete.

## Lesson 3 Summary

This lesson introduced the types of data including big, small, and raw data.

Here are the key points discussed in this lesson:

- Developing outcome-based data is important for presenting the effectiveness of an organization or a program.
- Big datasets are large data that may have several thousand entries and cover long periods of time. While there is no set number to define big or small data, a small dataset may fit on the screen or a couple of tabs.
- Raw data is not formatted or organized prior to performing any type of data analysis.
- It's important to ensure that you're collecting and producing accurate and meaningful data. When using any data source and using the collected data for reporting, it is important to properly cite the source and identify any data errors including data suppression.
- There are major sources of reliable national data such as the NFIRS and the NFPA Survey of Fire Departments for U.S. Fire Experience. However, both sources do have rooms for improvements.
- While raw data format such as CSV is difficult to interpret and use for analysis, it can be converted to a manageable spreadsheet format using tools like Microsoft Excel.

Congratulations! You have completed the lesson.