EvalFest How-Tos for Graphical Exploratory Data Analysis in SPSS

Graphical Exploratory Data Analysis in SPSS

Hi, this is Jane. In this video, we will talk about exploring the data graphically using SPSS version 25. This video builds on others in our Data Management series. In the Database Organization and Password Protection and Data Cleaning and Recoding videos, we discuss how to prepare your data for analysis.

Now you might be thinking, why bother? Looking at the data in various graphical forms is another way of ensuring that the data were collected and coded as expected. If you produce a graph of the data that is very unexpected, you will want to go back to the raw data to try to determine what is wrong.

This step is really about looking at the data in a graphical format to ensure that nothing is amiss. If you want more information about formatting graphs or creating other graphs, we suggest Stephanie Evergreen's blog, 'How to Build Data Visualizations in Excel' (http://stephanieevergreen.com/how-to/) and Ann K Emery's online tutorials (http://annkemery.com/tag/tutorials/) as good resources.

Exploring the data graphically

For now, we're still getting to know our data, and so let's get started.

1. Single Bar graph
    a. Let's look at a single bar graph for the variable Gender.
        i. Bar graphs are typically used to display categorical data.
    b. Select on the Graphs menu at the top.
    c. Choose Chart Builder.
    d. You will see in the bottom left of the Chart Builder screen that you can create a variety of different graphs and charts.
    e. For a single bar graph, click on Bar.
    f. We want a simple, single variable bar graph, so choose the first graph at the top left.
    g. Drag that sample graph to the Chart Preview box. Keep in mind that SPSS is not displaying your data in the graph – it is simply using example data.
    h. Next, click and drag the variable for which you want to produce a bar graph, in this case, Gender, to the x-axis.
    i. Click OK.
    j. Just as with numerical summaries, graphical displays allow you to see the

data to ensure that nothing seems off. For example, if we knew that this event was attended by almost all women, this graph wouldn't make sense. If your graph appears very unlike you would expect, you would need to take a closer look at the data before you run any statistical tests.

 k. Because we are simply exploring and getting to know our data, we don't worry about formatting our display at this stage. We recommend reformatting a display such as this before including it in a presentation or report.

2. Side-by-side bar graph
   a. Let's look at OverallCode, but this time, by Gender. For this, we will produce a side-by-side bar graph.
   b. We will go back to Graphs and Chart Builder.
   c. If you have just created a graph or chart, you will want to Reset the Graph Builder window before beginning a new graph.
   d. Click and drag the send graph, the side-by-side bar graph, to the Chart Preview box.
   e. Click and drag OverallCode to the x-axis.
   f. Click and drag the variable on which you want to cluster Overall Code, or Gender, to the Cluster on X box.
   g. Click OK.
   h. Again, look for patterns in the graph that you may not have expected. Does everything look okay? Should you proceed? These are questions you want to ask at this step.
   i. As you can see, Very Good (4) and Excellent (5) are the most frequent ratings, regardless of Gender. If we think back to the mean, median, and standard deviation we computed for OverallCode in the Numeric Exploratory Data Analysis video, these values align with what we see in this graph. The mean of 4.39 is less than the median of 5, the standard deviation is relatively small (.759), and the data are left skewed, showing most of our data at the upper end of the rating scale.