# *Graphing*

Are the policeconf1 scores normally distributed?

We now know the mean value of our variable, and we can look at the distribution of values around the mean. We can check on the distribution of our variable by using SPSS to create a histogram. Histograms are a way of displaying the distribution of a continuous variable, such as the one we have.

To create a histogram, go to **Graphs**, **Legacy Dialogs**, and then **Histogram**.

Move **policeconf1** from the variable list on the left to the **Variable** box at the top of the dialogue box.

The SPSS Output window should open. Your histogram will look like this:
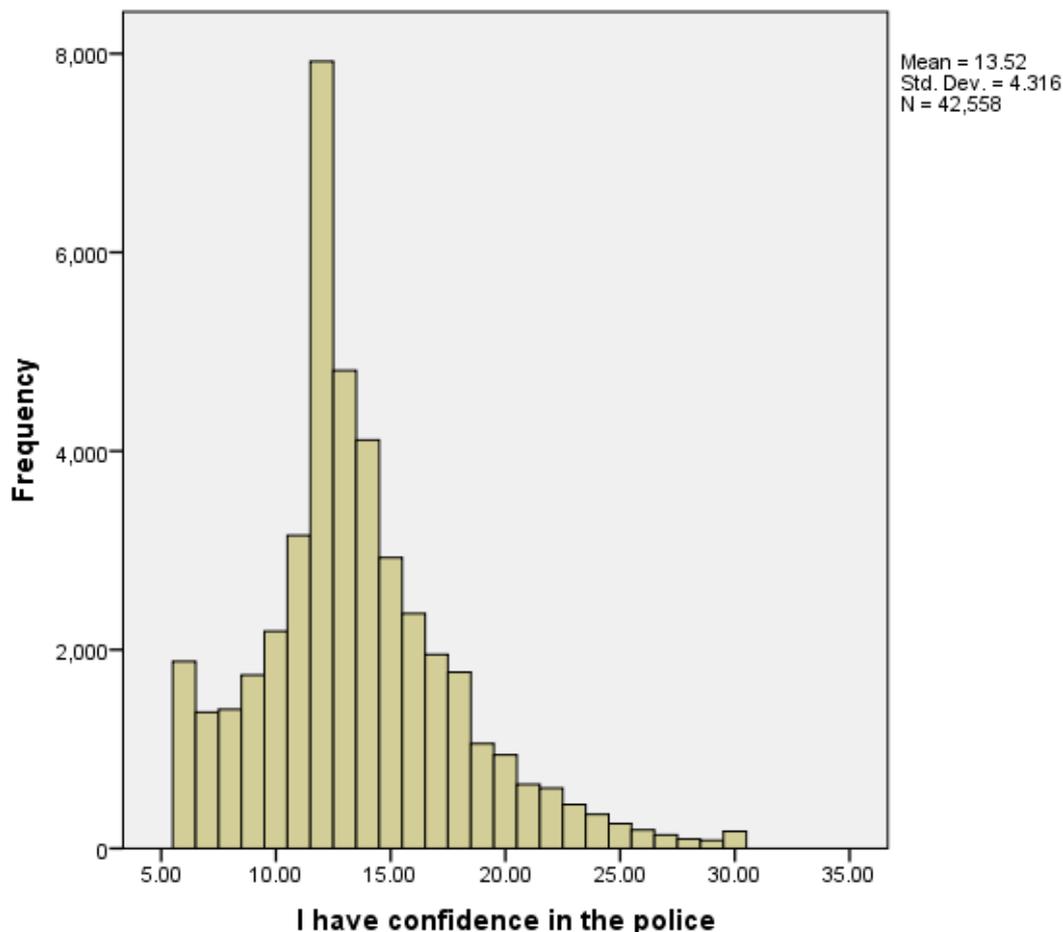
Notice that our histogram is not symmetrical, with the highest numbers of respondents near the left hand side. There is a long 'tail' of values on the right hand side. A distribution of this sort is called a positively skewed distribution. However most of the value points are near the mean of 13.5248, with the number of data points decreasing as they move farther from the mean in either direction. This shows us that the values in our variable **policeconf1** are more likely to be towards the bottom of the range of potential values than towards the top. This is good to know because it is important that you understand your dependent variable before you analyse it.

Looking back at our histogram, why do you think there is a spike in the number of data points at both 5.00 and 30.00, the extremes of the histogram?

This could be due to the fact that our data in **policeconf1** comes from survey respondents who may have answered all six **polatt** questions with only the extreme answers, all "1 = strongly agree" or "5 = strongly disagree."

You can create various types of graphs using the **Graph → Legacy Dialogue** function. Simply select your graph of choice (bar, line, etc.) and then move the variable you wish to illustrate to the **Category Axis** text box in the centre of the graph dialogue box. Other graphs that may be helpful to exploring a continuous variable like **policeconf1** are stem and leaf plots and box plots, which both help to illustrate the data variability for continuous variables.

Summary

You have studied the distribution of policeconf1 in order to better understand this dependent variable before using it in statistical analyses. You've seen that the data in policeconf1 is positively skewed, with most of the data trending toward the low end of the range of values. Considering how we created policeconf1, what might this imply about reported levels of police confidence?