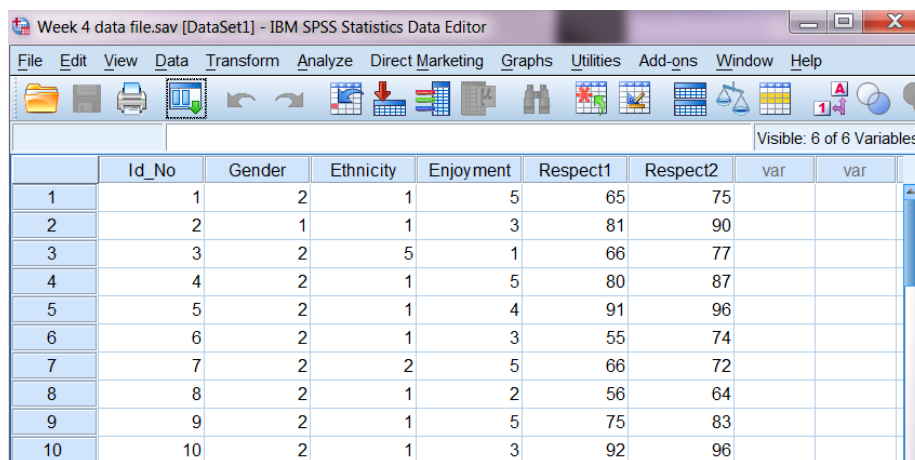# Creating Graphs in SPSS

This tutorial will show you how to explore your data, by producing graphs in SPSS.

Like the 'Descriptive Statistics' tutorial, the data used for this example is loosely based on the evaluation of the Schools Linking Network. As the name suggests, this project links schools in different communities to put 'the contact hypothesis' into practice. The contact hypothesis suggests that by increasing contact between people from diverse backgrounds, prejudice can be reduced and positive attitudes towards 'out-groups' can be fostered.

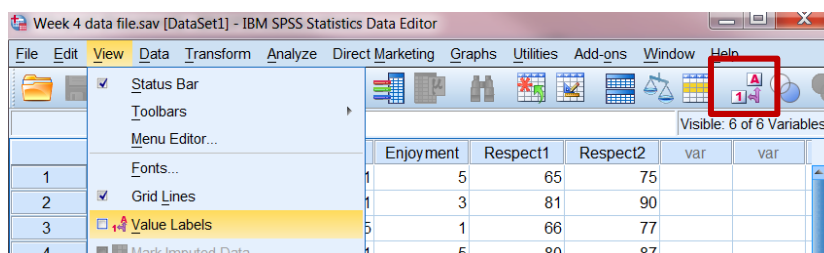The data can be found in the SPSS file: 'Week 4 data file.sav' and looks like this:

For an in-depth explanation of what each of the variables represent, revisit the 'Descriptive Statistics' tutorial.

When deciding what type of graph to produce, you first need to think about (1) the type of data you have collected: interval, ordinal or nominal data. And (2) what information are you trying to convey: means, medians, modes or frequency data.

Let's start by exploring our *nominal* (or categorical) variables. SPSS codes different categories by assigning each one a number (which is displayed in the datasheet above). To see how this is done, revisit the **Adding Variables** tutorial.

To see what the different coding options are in any dataset, you can **CLICK** on the **View** menu and select the **Value Labels** option *or* use the **Value Labels** shortcut button on the task bar (outlined in red below).

Doing this will display the variable labels for each of your conditions or groups.

So how might you best summarise this kind of data? What information might you want to show?

Let's explore this using one of the variables above: **Ethnicity**. One thing you might want to consider when exploring the effectiveness of the Schools Linking Network is how diverse the students who took part were. In other words, you would want a way to represent the proportion of participants who fell into each of the different ethnic groups defined here. A good way to represent proportions is using a **pie chart.**

Pie Charts

Just to confuse you, SPSS has multiple ways of producing charts and graphs... but this tutorial is going to focus on the method you are likely to use the more: using the chart builder.
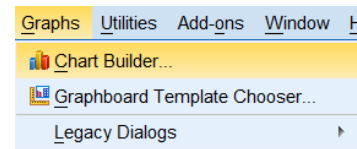

To produce a pie chart you first need to **CLICK** on the **Graphs** menu and select the **Chart Builder** option.

The first thing you will see is a pop-up box asking you to define your level of measurement for each variable (i.e. tell SPSS whether it is interval, ordinal or nominal). This has already been done for this example (to see how, revisit the **Adding Variables** tutorial).

To stop this pop-up from displaying again, tick the 'Don't show this dialog again' box.

CLICK on **OK** to continue.
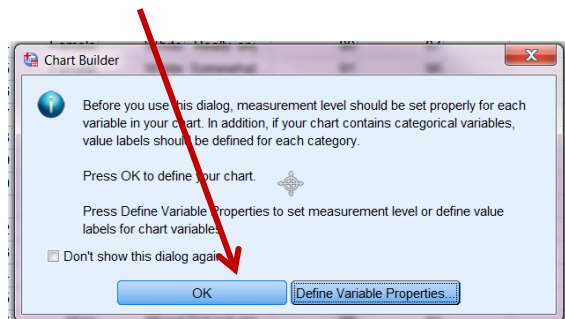
This will open up the chart builder for you.

The chart builder is an interactive drag-and-drop dialog box where you can define exactly what type of graph you would like.

The **Variables** list displays the different variables you have in your dataset.

The **Gallery** displays each type of graph you can create and its variants.

The **Chart Preview** box is the canvas, where you can drag and drop information to build and preview your graph or chart.

To create a pie chart, **CLICK** on **Pie/Polar** from the Gallery options, and **SELECT** the image of the pie chart and **drag and drop** it to the Chart Preview window.

You may have noticed two things happen when you do this:

1. The first is that the Chart preview window now displays a preview of a standard pie chart, which provides drag-and-drop boxes for you to define which information you want it to display.

2. The second is the opening of the **Element Properties** dialogue box, which allows you to define what numerical aspects of your data will be displayed.

So, you now need to define the information you want the pie chart to display.
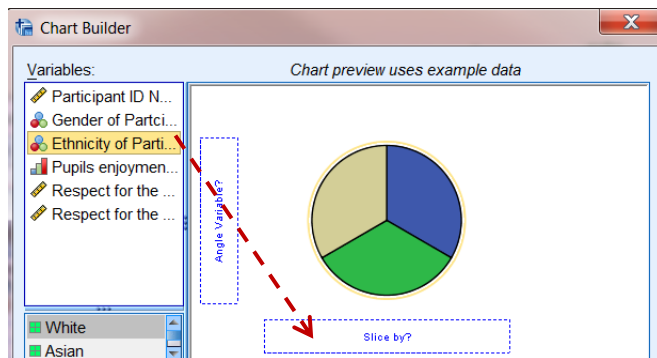
In this case, you want the **Slices** of the pie to represent the ethnicity of your participants. So, **SELECT** the Ethnicity variable from the Variables List and **drag and drop** it to box marked 'Slice by?' in the Chart Preview window.
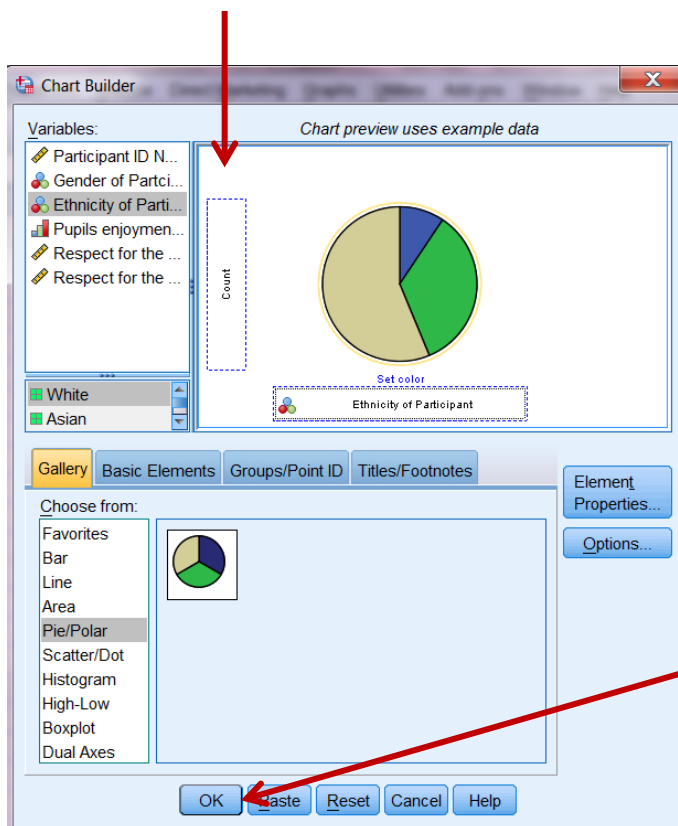

You may have noticed that in doing this, your **Angle Variable** box has automatically filled itself to '**Count**'. This means the angles of the slices will be determined by how many participants are in each group.

If we wanted to change this, we could do so using the **Element Properties** box.

However, as this is what we want to display, simply click on the **OK** button in the **Chart Builder** to produce your chart.

In the **SPSS Output** viewer, you should now find you have produced the following pie chart:

From this we can easily see that the vast majority of pupils taking part in the Linking Schools Network were white, making up almost three-quarters of the sample.
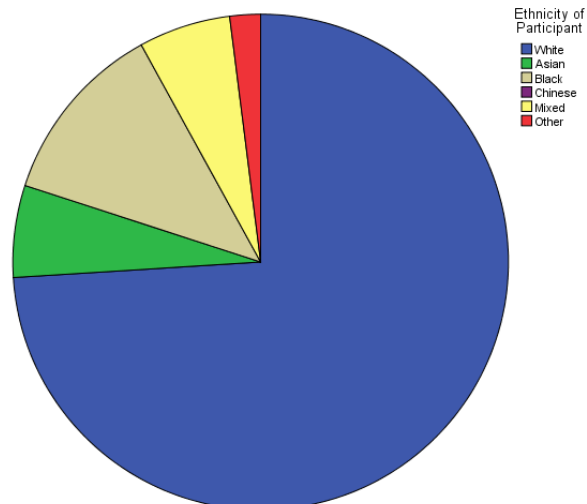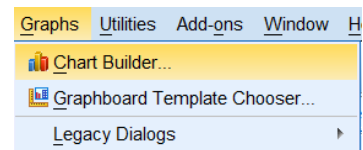
Histograms

Now you know how to produce a simple pie chart, let's try to produce a different type of graph. This time, let's investigate how much the participants in the Linking Schools Project enjoyed meeting people through the program.

The variable **Enjoyment** recorded how much pupils enjoyed this on a scale of 1-5 (where 1 = did not enjoy at all and 5 = really enjoyed). As this is a single scale with 5 ordered categories to choose from, it's an *ordinal* variable. In this case, you might be interested in the frequency with which participants chose each of these five response options. A good way to investigate a frequency distribution like this would be with a **histogram**.

Again, open up the **Chart Builder** by selecting this option through the **Graphs** menu.

The Chart Builder will remember your settings from the last graph you produced (unless you have closed SPSS in between). Don't worry about that – you just need to replace the options.

First, tell SPSS what graph you want by selecting **Histogram** from the **Gallery** window.

This time, there are a number of graph options that you can choose from. You are only ever likely to need two of these options:

Simple Histogram: when you want to look at the frequency distribution for a single variable (as we do now).

Stacked Histogram: where you want to compare the frequency distribution for a single variable across different groups.
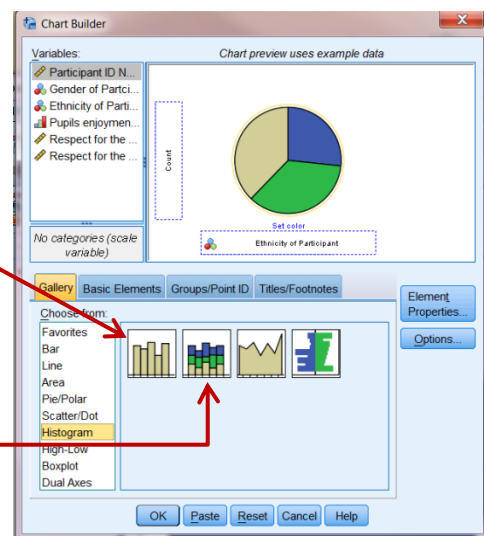
As we only have one variable that we are interested in here, **CLICK** on **Histogram** from the Gallery options, and **SELECT** the image of the simple histogram and **drag and drop** it to the Chart Preview window.
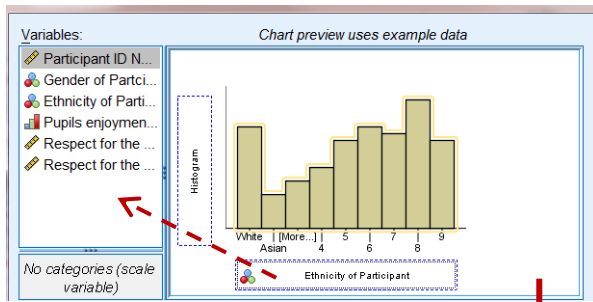
The Chart Preview window now displays a histogram, but it still contains the Ethnicity variable. To replace this, simply click on the box containing the variable and drag it back to the **Variables** list.
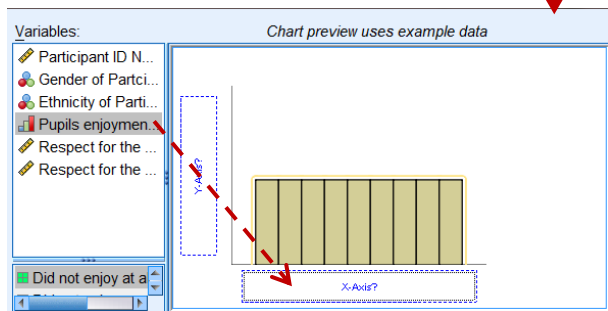
This gives you a blank histogram template. You now need to define what information will be displayed on the **X-axis** and the **Y-axis** of the graph.

In this case, you want each bar of the histogram to represent a score on the Enjoyment scale.

So, **SELECT** the Enjoyment variable from the Variables List and **drag and drop** it to box marked X-Axis' in the Chart Preview window.
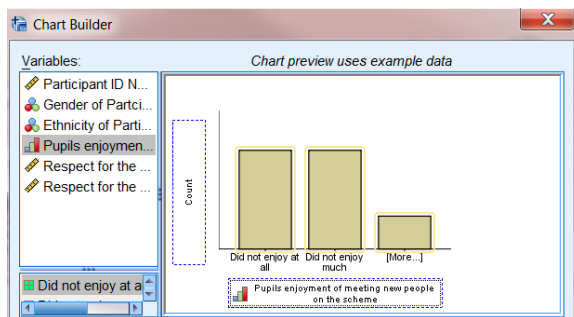

As with the pie chart, doing this automatically fills the other box (Y-Axis,) with '**Count**'. This means the height of each of the bars will be determined by how many participants are in each group. While this is absolutely fine for a histogram, for this example let's change the display to something else.

We can change what the Y-Axis represents using the **Element Properties** dialog box.

In this case, we want to change the display from Count to **Percentage** (which will display the percentage of participants who selected each of the response options).

To do this, **CLICK** on the drop-down **Statistic** menu in the **Statistics** box, and **SELECT** 'Percentage ()'.

Once this is selected and displays in the **Statistic** box, **CLICK** on **Apply** to change the histogram display.
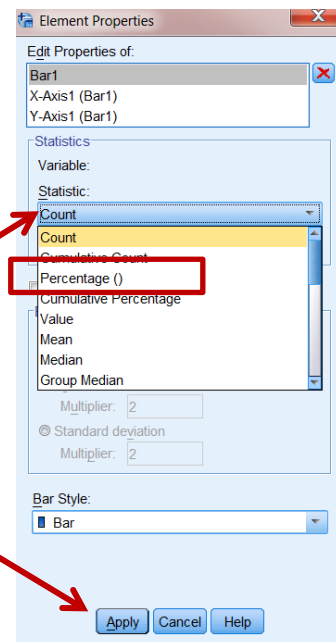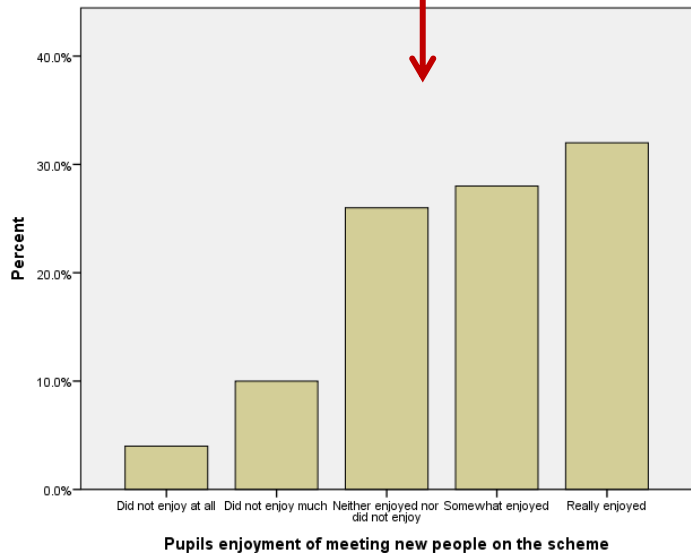
When the changes have been applied in the Chart Preview display, **CLICK** on **OK** in the **Chart Builder** window to produce your histogram.

Your histogram should now appear in the **SPSS Output** window:
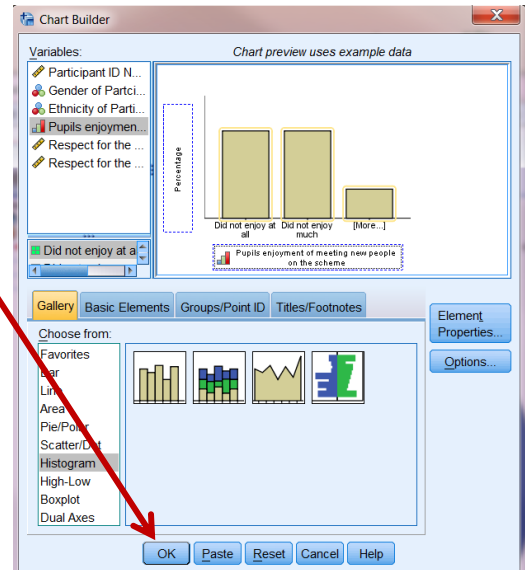
From the histogram you can see that the majority of pupils enjoyed meeting new people through the Linking Schools Network.

Bar Charts

So far we have looked at producing charts that display the frequency distributions of nominal or ordinal variables. But what about situations where we want to compare mean scores on one variable across different conditions or groups. For example:

- You might want to compare the mean Respect scores that boys and girls had before taking part in the Linking Schools Network. In this case, the groups being compared are **independent** from one another, as participants can only fall into one of the two grouping categories (they are either male or female).

- Alternatively, you might want to investigate the effectiveness of the Linking Schools Network by comparing the mean Respect scores participants had both before and after taking part in the program. In this case, the groups being compared are **related**, as the same participants would give data at both of the two time points.

When comparing group means, the best type of graph to use is a **bar chart.**
But whether the comparison you want to make is Independent (between participants) or Related (within-participants) affects the way you would produce the bar chart.

Regardless of the example you choose, you would start off in the same way. First, let's imagine we want to compare mean Respect1 scores across gender.
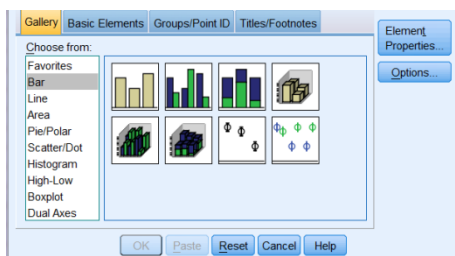
In this case you would open up the Chart Builder (as above) and select **Bar** from the Gallery window.

Again, several options are displayed, but you are only ever likely to use two:

Simple Bar: when you are investigating mean scores on one variable (e.g. respect scores) across the different groups or conditions of another single variable (e.g. across gender, or across time points).

Clustered Bar: when you are investigating differences in mean scores on one variable across the different groups or conditions of multiple other variables (e.g. across gender and ethnicity simultaneously).

In both of our examples here, we would use a **Simple Bar**. As such, __SELECT__ the image of the simple bar and __drag and drop__ it to the Chart Preview window.

Now, empty the chart preview of any existing variables (as before).

Bar Charts Comparing Independent Group Means

Let's take our first example, displaying mean respect scores for boys and girls in the program. In this example, we would want two bars; one to each represent gender. Their height would represent their Respect scores.
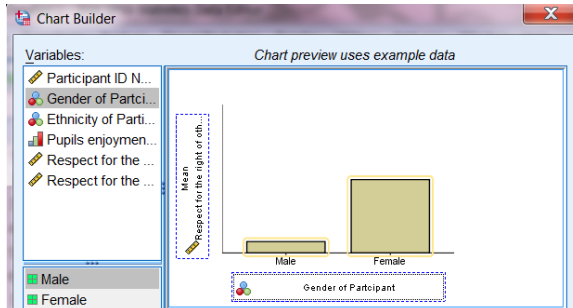

As such, **SELECT** the first **Respect** variable from the Variables List and **drag and drop** it to the box marked **Y-Axis** in the Chart Preview window.

Next, **SELECT** the **Gender** variable and **drag and drop** it to the box marked **X-Axis**.

You may have noticed that in doing this the **Statistics** box has detected the variable on the Y-Axis and has automatically recommended displaying the **Mean** value here. As this is what we want to display there is no need to change this.

But when producing a bar chart it is always good to include error bars, as this gives an idea of the dispersion in your data and the representativeness of your sample.

To do this, **SELECT** the **Display error bars** box and choose an option from the 'Error Bars Represent' box. In this case, choose the '**Standard error**' option with a Multiple of 2. This measure is important, as you will learn when it comes to inferential statistics later on.
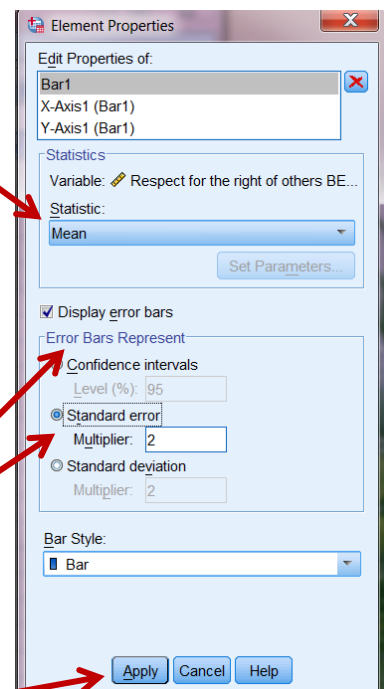
CLICK on **Apply** to save the changes.

When the changes have been applied in the Chart Preview display, **CLICK** on **OK** in the **Chart Builder** window to produce the following bar chart in the **Output** window:

As you can see from the graph, Respect score appear very similar for the two genders (something you can tell from the overlap of the error bars).
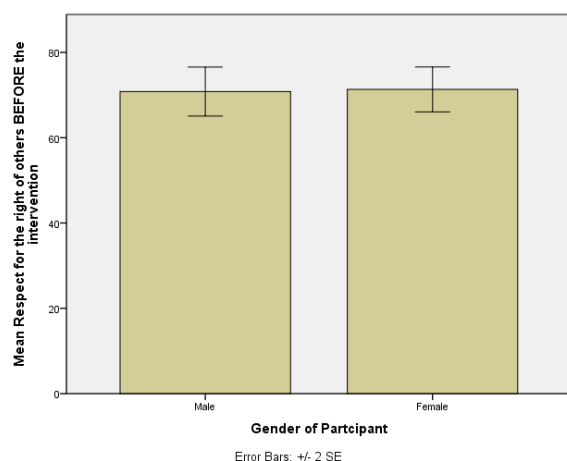
Bar Charts Comparing Related Group Means

As our final example, let's go back to our suggested example investigating the effectiveness of the Linking Schools Network. In this case, participants' mean Respect scores need to be compared both before and after taking part in the program. The means are related, because the same participants give data at each time point.

Following the steps you have gone through before, open up the Chart Builder and empty the Chart Preview window of any existing variables.

In this example, we want the bars to represent each time point (before and after taking part in the program) and their height to show their Respect scores.

To do this, you need to hold down the Ctrl key on your keyboard and **SELECT** *both* Respect variables simultaneously.

Once they are both highlighted, **drag and drop** them to the box marked **Y-Axis** in the Preview window.
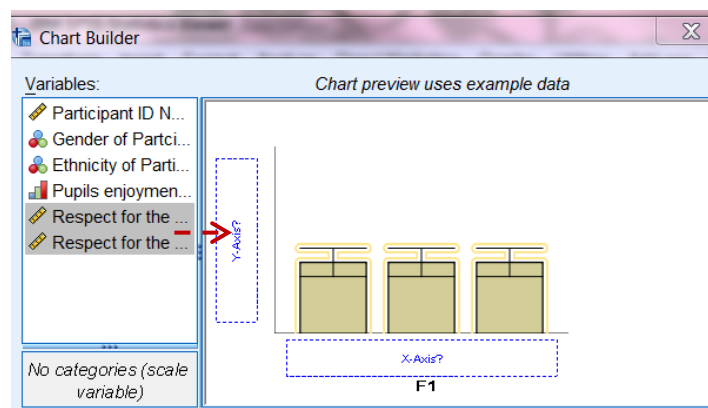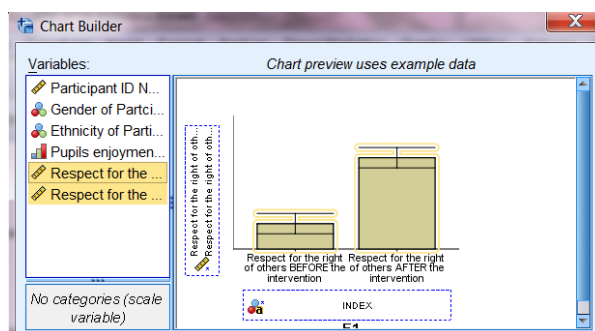

This brings up the following dialog box: Essentially this box tells us that it is creating two new variables:

Summary: which is the variable we are measuring directly (i.e. Respect scores)

Index: which represents the grouping categories (i.e. Before and After taking part in the program)
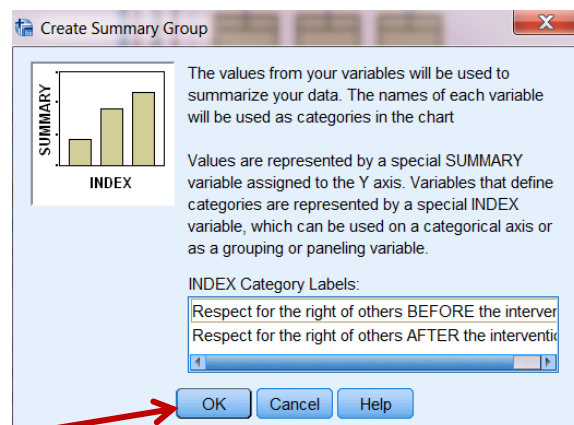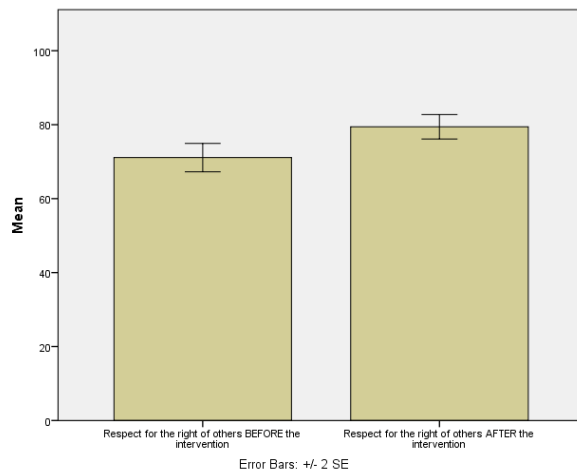
Just **CLICK** on **OK** to accept this.

As SPSS remembers all of our previous settings (i.e. displaying Means and +/- 2SE error bars) the Chart Preview window should be correct.

CLICK on **OK** in the **Chart Builder** window to produce the bar chart.

The SPSS Output window should now display the following bar chart:

In this case it appears that Respect scores after the intervention were higher than beforehand.

You have now produced several different types of graphs. Why don't you take the time to explore the other different types of charts you can produce by downloading the data file used in this tutorial and playing with the different graph options.

For example, see if you can produce:

- a pie chart for the different genders
- a box plot for one of the variables
- a bar chart looking at respect scores before and after the intervention, clustered by gender

The more you play with SPSS, the more familiar with the program you will become.