CHARTS AND GRAPHS

INTRODUCTION

SPSS and Excel each contain a number of options for producing what are sometimes known as ‘business’ graphics - i.e. statistical charts and diagrams. This handout explores some of the options available in each application and also looks at how the graphics may be copied into Word documents.

Run Cag08.bat to download some sample data to work with.

USING SPSS TO DRAW GRAPHS

SPSS provides a large number of options for producing charts and diagrams. The graphics options in SPSS are available on the Graphs menu option. Each option on the Graphs menu contains a number of sub-options, some of which contain further sub-options. The first two options are a bit different to the others. The Graphs|Gallery option provides on-line help about each of the major types of graph available. It also provides step by step instructions about how to draw the graph. The Graphs|Interactive option (introduced in version 8.0) may be used to draw Interactive Graphs (see below). Each of the other options on the Graphs menu may be used to draw specific types of non-interactive graphs. The procedures used within each of these options are somewhat similar.

When you draw a graph, the graph is initially displayed in the Viewer window. Graphs may be moved within the Viewer window, hidden or displayed, and deleted in the same way as tables output from the Analyze options.

SPSS GRAPH OPTIONS

The main non-interactive options available from Graphs option include:

- **Bar** allows you to generate a simple, clustered, or stacked bar chart from your data.
- **Line** allows you to generate a simple or multiple line chart from your data.
- **Area** allows you to generate a simple or stacked area chart from your data.
- **Pie** allows you to generate a simple pie chart or a composite bar chart from your data.
- **High-Low** allows you to plot pairs or triples of values, for example high, low, and closing prices.
- **Pareto** creates Pareto charts - bar charts with a line superimposed showing the cumulative sum.
- **Control** produces the most commonly-used process-control charts.
- **Boxplot** allows you to generate boxplots showing the median, interquartile range, outliers, and extreme cases of individual variables.
- **Error Bar** allows you to generate boxplots showing the mean, standard deviation, standard error and confidence intervals of individual variables.
- **Scatter** allows you to generate a simple or overlay scatterplot, a scatterplot matrix, or a 3-D scatterplot from your data.
- **Histogram** allows you to generate a histogram showing the distribution of an individual variable.
- **Normal P-P** plots the cumulative proportions of a variable's distribution against the cumulative proportions of the normal distribution.
- **Normal Q-Q** plots the quantiles of a variable's distribution against the quantiles of the normal distribution.
- **Sequence** produces a plot of one or more variables by order in the file, suitable for examining time-series data.
- **ROC Curves** may be used to test the efficiency of a two category classification.
- **Time Series: Autocorrelations** calculates and plots the autocorrelation function (ACF) and partial autocorrelation function of one or more series to any specified number of lags, displaying the Box-Ljung statistic at each lag to test the overall hypothesis that the ACF is zero at all lags.
- **Time Series: Cross-correlations** calculates and plots the cross-correlation function of two or more series for positive, negative, and zero lags.
Time Series: Spectral calculates and plots univariate or bivariate periodograms and spectral density functions, which express variation in a time series (or covariation in two time series) as the sum of a series of sinusoidal components. It can optionally save various components of the frequency analysis as new series.

The best way to get a feel for the options available is to explore the Graphs|Gallery menu option and then to try some of the options.

We will use the data file c:\temp\spss.sav to explore some of the options. Start SPSS and load the file c:\temp\spss.sav using the File|Open option.

Select the Graphs|Bar option. Click on the Simple button, click the radio button beside ‘Summaries for groups of cases’, then click on the Define button. Click the radio button beside ‘N of cases’ on the second dialogue box, highlight age in the list of variables and then click the arrow beside Category Axis so that age is loaded into the box. At this point we can click the OK button to draw our first graph. It will be noted that the graph displays a bar indicating the frequency (i.e. number of cases) of every age value in the data file.

Select Graphs|Bar again, leave the settings on the first dialogue box unchanged (i.e. click the Define button), and click the radio button beside ‘Cum. N of cases’ on the second dialogue box, and then OK. The histogram this time will display the cumulative number of cases.

Repeat a third time, but this time click the radio button beside ‘Other summary function’ on the second dialogue box. The OK button will become dimmed, indicating that something further is required. Select educ in the list of variables, and click the arrow beside the box marked Variable. MEAN(educ) should appear in the box. When you click OK, a graph will be drawn showing the mean education level for each age. It will be seen that the number of years of education generally decreases with increasing age, although there is a curious cyclical effect.

The ‘% of cases’ and ‘Cum. % of cases’ are identical to their non-percentage equivalents, except that the Y-axis is expressed as a percentage of the total valid cases.

Exercise 1.

Draw a graph showing the mean number of years of education for each racial group. (N.B. To empty a box already containing a variable name, click on the box, then click on the arrow button beside the box. Double clicking on the name will also remove it from the box).

Exercise 2.

Adopting a similar approach to the above, see if you can find out what the Clustered and Stacked options do on the first dialogue box. Likewise see if you can establish what the ‘Summaries of separate variables’ and ‘Values of individual cases’ options do. Check out the Help option to see how informative it is. (N.B. You may find it easier to use variables with a small number of classes - e.g. race and sex - for these experiments).

Exercise 3.

Use the Graphs|Histogram option to draw a histogram of age. How does this differ from a bar chart of age? Which option would be better for data on a continuous scale (e.g. temperatures) where no one value is likely to occur more than once?

Exercise 4.

Explore the other options on the Graphs menu, paying special attention to Line, Area, Pie, Boxplot, and Scatter. (N.B. You may need to play around with different variables before some options will do anything constructive - e.g. the time series options obviously require time series data).
EDITING SPSS GRAPHS

If you wish, you can edit a chart to modify its appearance. To edit a chart, double click on the chart within the Viewer. This will open the chart in a Chart Editor window. The chart may be edited (a) by selecting options from the Chart or Format menus, (b) by double clicking on the object you wish to change, or (c) by selecting an object with a single click and clicking on an icon on the toolbar. The toolbar icons fulfil the following functions (from left to right):

- **Point Id** allows you to identify the case represented by a point in scatterplots etc.
- **Fill pattern** allows you to change the fill pattern of a selected object.
- **Colour** allows you to change the colour of a selected object.
- **Markers** allows you to change the markers on a chart or plot.
- **Line Style** allows you to change line styles.
- **Bar Style** allows you to change bar styles for charts that contain bars.
- **Bar Label Style** allows you to label bars with the numeric values that they represent.
- **Interpolation** allows you to select an interpolation method for a line or scatterplot series.
- **Text** allows you to change the font and point size of text within a chart or plot.
- **3-D Rotation** rotates 3-D scatterplots, using a dialog box that shows only the axes.
- **Swap Axes** interchanges the vertical and horizontal axes of a chart.
- **Explode Slice** visually removes or "explodes" the selected slice from a pie chart.
- **Break Line at Missing** breaks a line chart where missing values should be.
- **Chart Options** allows you to change options for certain types of chart.
- **Spin Mode** rotates 3-D scatterplots directly, hiding only the labels. It is slower than 3-D Rotation when many data points are present.

When the graph looks the way you want it to, close the Chart Editor window. The modified chart should appear in the Viewer window.

INTERACTIVE CHARTS

Version 8.0 (and upwards) of SPSS includes options for Interactive Graphs. These are initiated from the Graphs|Interactive menu option. This provides a number of sub-options (Bar, Dot, Line, Ribbon, Drop-Line, Area, Pie, BoxPlot, Error Bar, Histogram and Scatter Plot), most of which are similar to the non-interactive graphs discussed above.

The Interactive Graphs options operate differently from the other graph options. One difference is that the variables to be graphed are selected by dragging and dropping operations - i.e. you drag the name of the variable (or the category within a variable) that you wish to graph and drop it in the appropriate box. For example, open the file c:\temp\spss.sav if it is not already open, and select the Graphs|Interactive|Bar option. A list of variables or a list of categories will be displayed in a panel on the left of the dialogue box that opens. To determine which is currently displayed, right click anywhere in the panel to display a pop-up menu. Make sure the ‘Display Variable Names’ option is selected. Drag the variable ‘Age’ from the panel on the left to box on the right representing the horizontal axis of the graph. Click the OK button to draw the graph.

The bar chart will be displayed in the Viewer window in the normal manner. An Interactive Graph can be edited in a similar manner to ordinary graphs by double clicking on the graph in the Viewer. However, double clicking on an Interactive Graph gives a number of additional options. For example, you can change the variables displayed, add additional variables, change the graph’s appearance (e.g. from 2-D to 3-D), and so on. Most of these options are initiated either from the icons which are displayed in the graph window, double clicking on the object you wish to change, or right clicking on an object.

Interactive Graphs are extremely versatile, fairly intuitive to use, but difficult to describe in a condensed manner. The best way to find out how they work is by trial and error. Further information is available from the Help|Topics and Help|Tutorial menu options.
Exercise 5.

Using the file c:\temp\spss.sav, experiment with different types of Interactive Graphs. Edit the graphs you produce to change their appearance. Experiment with the interactive properties to change the variables and/or categories graphed. Consult the Help|Topics and Help|Tutorial menu options to find out what sort of help is available.

COPYING SPSS GRAPHS TO WORD DOCUMENTS

Graphs and Interactive Graphs created within SPSS can be exported to Word (and other applications). To copy a SPSS graph into a Word document, create the graph as described above, select the graph to be copied (i.e. click on it - a box will appear around the graph to indicate it has been selected) and then select the Edit|Copy menu option. Open the Word document that you want to copy the graph to, locate the cursor at the point where the graph is to be inserted, and select the Edit|Paste menu option. The graph should appear in your Word document. (N.B. If the Edit|Paste menu options does not produce a graph as expected, try the Edit|Paste Special option. This will give you a choice between a number of formats. Trial and error will establish which one works best with your application).

If you click on the graph in the Word document, little squares (known as sizing handles) will appear around the edges of the graph to indicate that it is selected. When the graph is selected you can delete the graph by pressing the <Delete> key on the keyboard, or you can re-size it by dragging any of the sizing handles. It is probably advisable to drag only the corner handles as these will resize the graph but retain the original proportions. However, if you want your graph to look taller or thinner, you can drag one of the handles at the middle of one of the sides. The graph may also be centred, or right- or left-aligned when selected.

If you double click on the graph in Word, you should go into editing mode, which allows you to edit the appearance of the graph. However, the options are very limited, so it is generally better to edit the graph within SPSS before copying it to Word. If you do not like the appearance of the graph in Word, the simplest solution is to delete the graph within Word, edit it within SPSS (as described above), and then copy a new version to Word.

CREATING GRAPHS USING EXCEL

Excel provides a number of graph options, including most of the more useful types available in SPSS - e.g. bar graphs, line graphs, area graphs, pie diagrams and scattergrams. However, SPSS provides a greater range of options, so you may find it preferable to use SPSS. For example, although Excel provides options for drawing histograms and trend lines (e.g. regression lines) through a scatter of points, these options tend to be more cumbersome than their equivalents in SPSS. On the other hand, Excel provides better options for setting up ‘live links’ between applications (see below).

Graphs created in Excel may be located either on a separate worksheet referred to as a chart sheet, or as an embedded chart on the same page as the data from which it is constructed. The latter option is generally the more useful.

Irrespective of which option you select, the graph will automatically be redrawn if any of the data values are altered. This is a very useful feature, especially if you are producing reports based on data that change at frequent intervals. (N.B this option is not available in SPSS. If your data values are changed when using SPSS, you will need to explicitly re-draw any graphs that use the data).

Creating an embedded chart in Excel is reasonably simple. To create a chart you should select the data that you wish to graph, then click on the Chart Wizard icon. The Chart Wizard will initiate a series of dialogue boxes. Simply select the options you wish at each step and click the Next button. An impression of what your graph will look like will be displayed at step 2. If this is not what you want, you can back-pedal by clicking on the Back button and then select alternative options. (This, incidentally, provides a handy way of exploring what each of the different options does). When you get to step 4, click the Finish button and your graph will appear on your worksheet. The sizing handles may be used to re-size the graph if required. The graph may also be dragged to a different location (and may even be placed on top of some of the data in the worksheet).
The procedure is somewhat similar if you wish to create a graph on a separate chart sheet, except that you specify a chart sheet on step 4.

**Exercise 6.**

Open an Excel spreadsheet that you have created in previous exercises (or, if none are available, open c:\temp\excel.xls). Select some data, and explore the various graph options available. Create both embedded charts and graphs on separate chart sheets.

**EDITING AND DELETING EXCEL GRAPHS**

The appearance of a graph may be edited. Select the feature you wish to change (e.g. line, bar, axis text, etc.) by double clicking on it. A dialogue box containing all the available options will be displayed. Select whatever changes you wish to make, and click on the **OK** button.

If you wish to delete an embedded chart, click once on the chart to select it, then press the `<Delete>` key on the keyboard. To delete a chart sheet, right click on its tab, and select Delete from the pop-up menu. (The sheet may also be renamed to something other than Chart1 etc. from the same pop-up menu).

**Exercise 7.**

Use the editing options to change the colour of bars, the colour of the background, the font used for different items of text, the style of lines etc. until you get a feel for the types of options available.

**COPYING EXCEL GRAPHS TO WORD DOCUMENTS**

Graphs created within Excel may be exported to Word (and other applications) using a similar procedure to that used for exporting SPSS graphs - i.e. select the graph in Excel, click the **Copy** button in Excel (or select the **Edit|Copy** option) to copy the graph into the clipboard, move to the position you where you want to insert the graph in the Word document, and then click the **Paste** button. (You may also achieve the same ends either by selecting the **Edit|Paste** menu option, or the **Edit|Paste Special** menu option whilst making sure to select the Microsoft Excel Chart Object).

At first glance the outcome will appear to be exactly the same as for a SPSS graph. However, there are some very important differences behind the scenes. Excel supports OLE (Object Linking and Embedding). When you copy and paste from SPSS, the graph is copied as a picture. However, when you do a simple copy and paste from Excel, the graph is copied as an **embedded object** - i.e. instead of copying a picture of the graph, the copy and paste operation copies all the information required to create the graph. The difference can be seen if you double click on the graph in Word. A miniature spreadsheet will be opened for editing. If you were to change some of the data values upon which the graph is based from Word, then the graph will be adjusted accordingly in Word, but the data values and graph in the original spreadsheet will remain unaffected.

It is also possible to **link** the graph in Word with the original spreadsheet. To do this you should select the graph within Excel and click the **Copy** button, move to the position where you want to insert the graph within the Word document, and then select the **Edit|Paste Special** option. Click on the radio button beside **Paste Link**, make sure Microsoft Excel Chart Object is selected, and then click OK. The graph will appear as before. However, if you double click the graph in Word, the original spreadsheet will be opened in Excel. Any changes you make at this point will be reflected not only in the graph in Word but also in the original spreadsheet. More usefully, if at any later stage you change the data in the Excel spreadsheet, then the graph in Word will automatically be updated to take account of the new information.

Linking is an extremely useful feature, especially if you need to write reports where the data is continuously changing. Linking also results in smaller Word files (due to the fact that they simply contain a link to the source information in Excel, rather than a copy of it), although it presumes that the Excel source file is still available (i.e. has not been deleted or moved). However, embedding is obviously more useful if you do not wish to have your Word file automatically updated each time the data in the spreadsheet change. Note that if you have a
linked graph and you accidentally delete your original spreadsheet, Word will still display a ‘linked’ graph but it will no longer be updated by the original spreadsheet.

N.B. SPSS provides some OLE options, but they require the target application to support Active-X objects. It is also necessary to have some programming experience in the SPSS scripting language. Consult the ‘OLE’ and ‘ActiveX’ topics in the SPSS Help for further details).

Homework

Start Excel. Create a column containing the numbers 1 to 10 in a new worksheet. Use this column to create a bar graph. Select the graph and press the Copy button. Open a new Word document. Press Paste to create an embedded copy of the graph, then use Edit|Paste Special to create a link to the graph immediately below it. Save the spreadsheet as test.xls. Insert text before each of the bar graphs to indicate which is the embedded object and which is linked to the spreadsheet.

Return to Excel and change the data value for 5 to 10. What happened to each of the graphs in Word? Why? (Type your answers to these and all subsequent questions into the Word document).

Double click on the embedded graph in Word. Change the data value for 5 to 2. What happened to each of the graphs in Word? Why?

Double click on the linked graph in Word. Change the data value for 6 to 10. Close the Excel spreadsheet without saving it. What happened to each of the graphs in Word? Why?

Save the word document as Graph**.doc (where ** is your initials) and then close Word. Start Word again and open Graph**.doc. What happened to the lower graph? Why?

Close Graph**.doc and email it to dennis.tringle@may.ie.