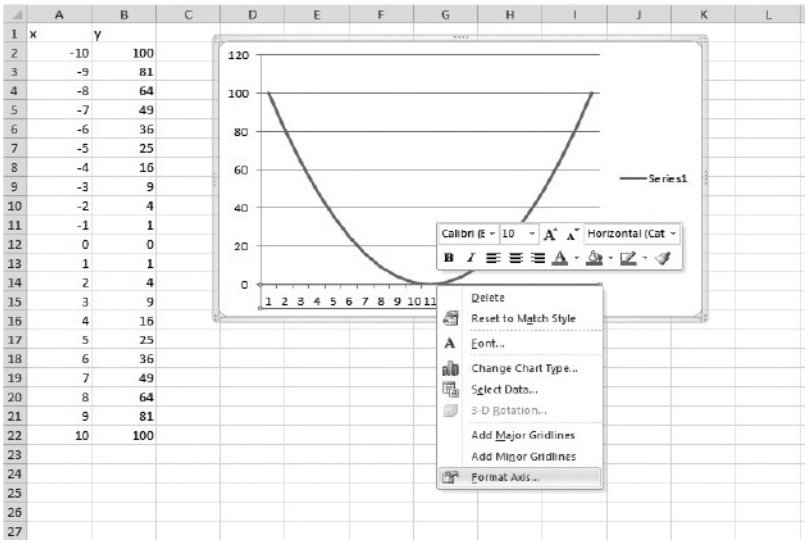
3

THE SOFTWARE ENVIRONMENT

QUESTION 1.

Using Excel create a spreadsheet that will plot the function $y = x^2$ (you may remember this equation is classed as a quadratic equation).



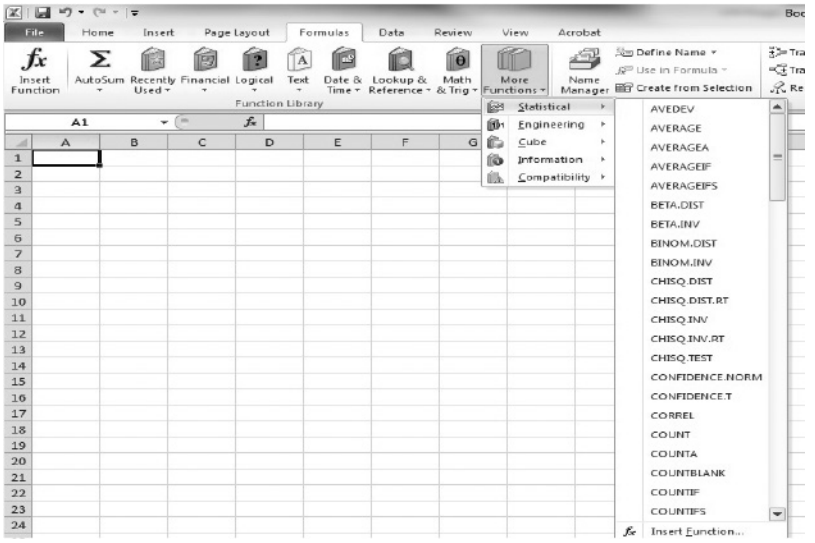
The diagram shows you my solution, which is untidy. It would be nice to have the vertical axis crossing the x axis at 0, the 'scale' or 'grid'? marks' aligned with the numbers, plus titles and axis labels. In the diagram you can see a dialogue box which is shown when you place your cursor near the horizontal axis.

Click on 'Format Axis' and experiment with the various options until you produce a neat and tidy graph.

QUESTION 2.

In the diagram below you will see the drop down menu associated with the formula tab. You will be using the functions in this drop down menu a lot and I mean a LOT!

Find and select 'average'. In the box below describe what Excel does when this function is used. You will have to find out what 'arguments' are (mathematics arguments not the ones you have with friends or family!), what the 'arithmetic mean' is and so on.



Excel calulates the arithmetic mean when the average function is used.
An argument can be a number or a name, an array or references that contain numbers.
In statistics we mainly use numbers.

QUESTION 3.

In question 2 you were asked to explore the average function. Following the same process find the 4 functions for the standard deviation and explain when they would be used and the differences between them.

STDEV.P is used to calculate the standard deviation based on the entire population.
STDEV.S is used to calculate the standard deviation based on a sample of the population.
STDEVA and STDEVPA are similar to STDEV.P and STDEV.S but will also accept text and logical values.

QUESTION 4.

Create an instance of SPSS then create a variable called 'scores' and enter the data shown below into SPSS:

53	80	43	57	57	53	80	43	57	57

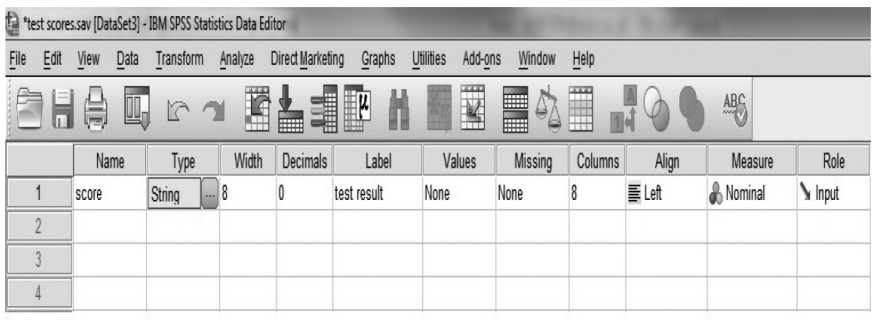
QUESTION 5.

The diagram below shows the results of an exam. Enter the data into SPSS and in order to get an idea of the range of the data, sort the data so that the lowest value is first.

score		score		score
53.00		54.00		57.00
80.00		68.00		54.00
43.00		79.00		54.00
57.00		54.00		76.00
57.00		54.00		66.00
53.00		63.00		64.00
80.00		65.00		68.00
43.00		63.00		69.00
57.00		65.00		
57.00		65.00		
39.00		70.00		
58.00		58.00		
55.00		41.00		
54.00		58.00		
80.00		45.00		
52.00		52.00		
55.00		55.00		
62.00		17.00		
63.00		53.00		
74.00		50.00		
100.00		61.00		

QUESTION 6.

The diagram below shows the variable view for the scores data as entered by a colleague, who you know to be a bit careless. Check the values of each of the attributes to make sure they are correct and make any necessary changes.



Type should be numeric, measure scale.

QUESTION 7.

In many cases you will get duplicated data. In other words, many people will select the same option from a questionnaire. It is sometimes useful to be able to identify the duplicated responses. The diagram below shows a screenshot from SPSS's output viewer. Identify the command that was used under the data tab and give an explanation of the information contained in the screenshot.

Frequencies

Statistics

Indicator of first matching case as Primary

N	Valid	50
	Missing	0

Indicator of each first matching case as Primary

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Duplicate Case	24	48.0	48.0	48.0
	Primary Case	26	52.0	52.0	100.0
	Total	50	100.0	100.0	

There were 24 same responses and 26 dierent responses.

QUESTION 8.

Create an Excel worksheet and enter the following data:

ID	1	2	3	4	5	6	7	8	9	10
Glasses	0	3	7	15	2	1	0	0	10	5

Once you have created the Excel spreadsheet, create an instance of SPSS and import the data.

MINI PROJECT

You have been asked to demonstrate to a group of students what random numbers are. In order to do this you decide to use Excel to create a spreadsheet with a column of random numbers. In anticipation of the question 'How do you know they are random?' you decide to investigate the functions Excel offers and see if any of them can be used to demonstrate randomness. You might also consider plotting the data.

And finally...

There is an acronym that has been around for a long time in the computer world, GIGO (Garbage In Garbage Out), take heed of it; if you enter rubbish data, you will get rubbish analysis!