PASW (SPSS)/Excel Workshop 1 – Semester Two, 2010

In Assignment 1 of STATS 101/108 you will need to use:

- PASW (SPSS) to generate descriptive statistics and plots in **Question 4**;
- Excel to create two appropriate tables of information in **Question 5 (a)** to **Question 5 (e)**; and
- Excel or PASW (SPSS) to appropriately display data in **Question 5 (h)**.

Instructions from your assignment sheet read:

| Question guide |
| --- |
| • Questions 4 and 5 will require use of PASW and/or Excel. Hand in the required computer output. |

To learn skills needed for these computing components of Assignment 1, we will be using a number of files from www.stat.auckland.ac.nz/~leila.
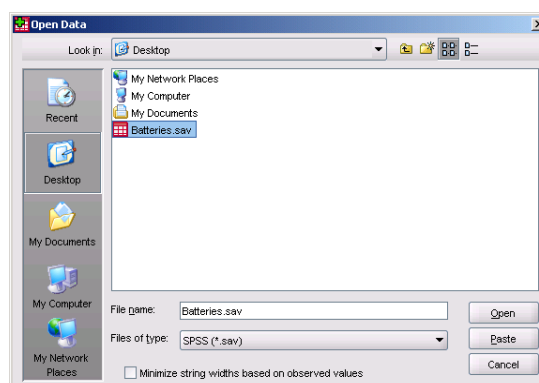
<u>PASW (SPSS) Basics</u>

In STATS 101/108 you will be using PASW (SPSS) to plot data and perform simple calculations.

Opening PASW (SPSS) files

1. Data file: click **File → Open → Data**.
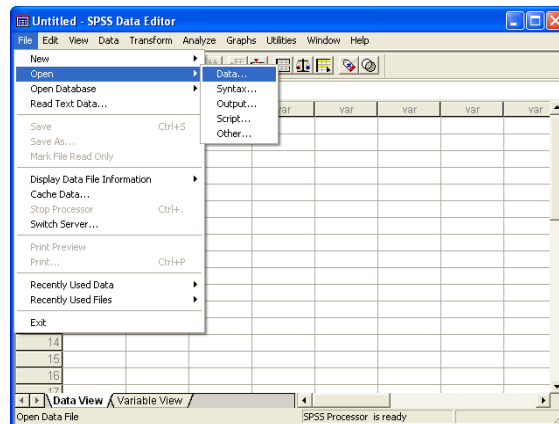   Output file: click **File → Open → Output**.
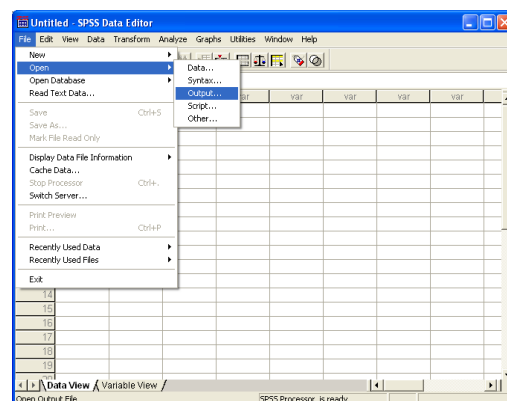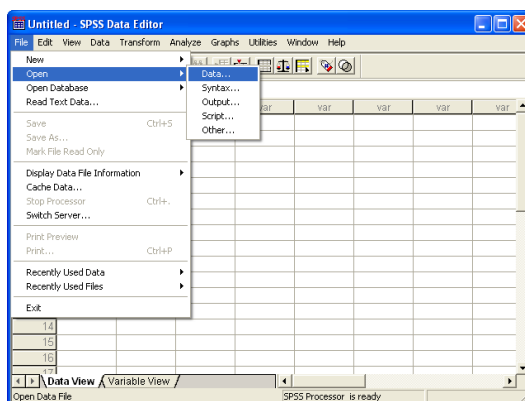
2. Choose the file.
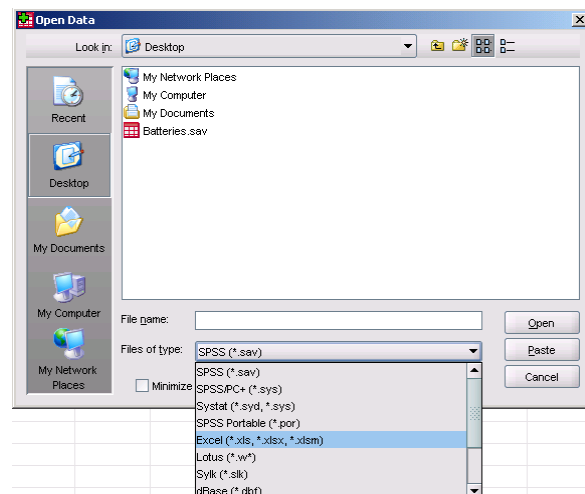   Find and click the required file and click **Open**.
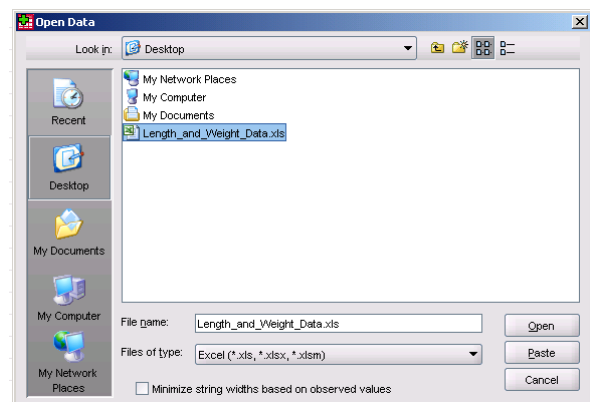
Importing data from an *Excel* file

1. Click **File → Open → Data**.
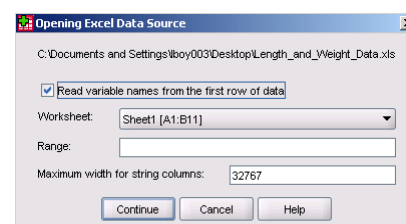
2. Choose the file type.
   In the **Files of Type** box,
   choose **Excel (*.xls)**.

3. Choose the file.
   Find and click the required file
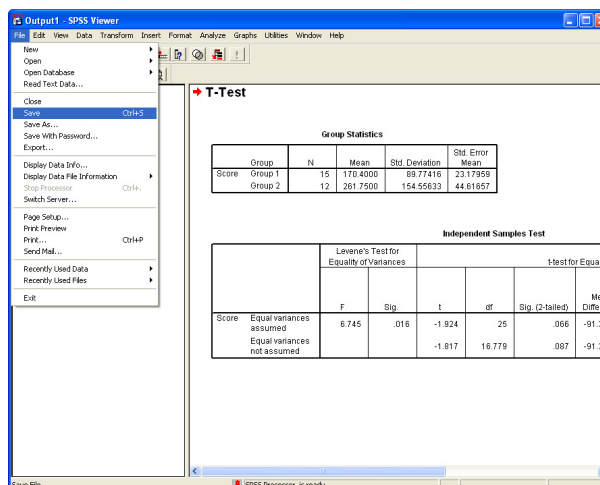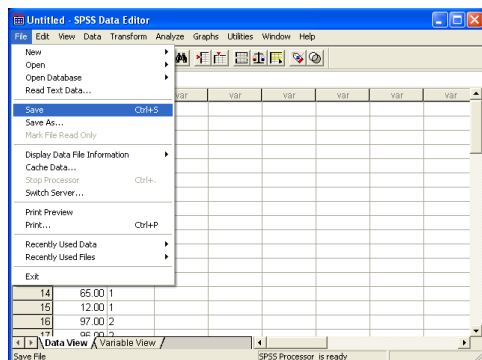   and click **Open**.

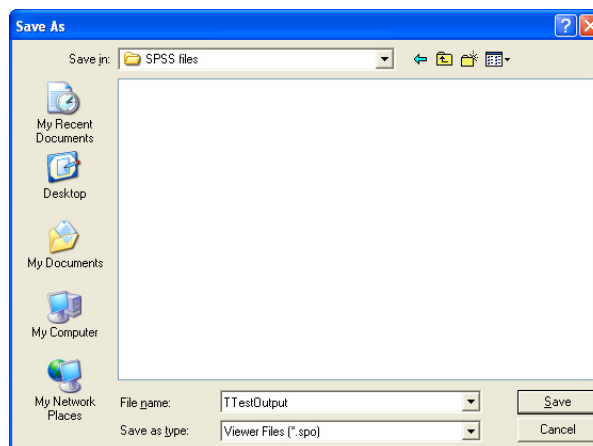4. Click **Continue/OK**.

Saving

In PASW (SPSS), data and output are saved separately.

1. Click **File → Save**.

2. Enter the file name.

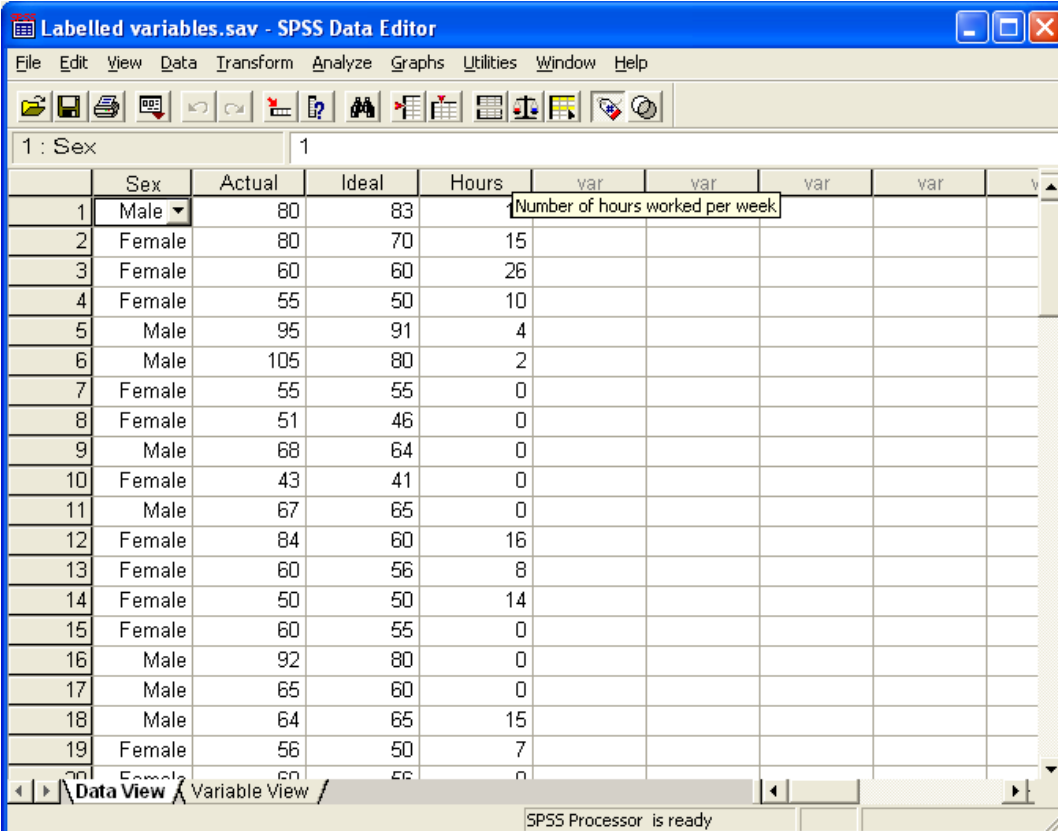   Type a name for the file in the **File name** box and click **Save**.

Data View

To view the data view, click the **Data View** tab at the bottom of the window.

All data is entered in the data view. Each row corresponds to one case (an individual, an experimental unit, etc.), and each column is a variable.

You can also view labels in the data view.

To view value labels, click the **Value Labels** button ⬦. The labels will now appear instead of the values.

To view variable labels, hold the mouse over the variable names at the top of each column. The variable label will appear.

	Sex	Actual	Ideal	Hours	var	var	var	var	v
1	Male ▾	80	83	Number of hours worked per week					
2	Female	80	70	15					
3	Female	60	60	26					
4	Female	55	50	10					
5	Male	95	91	4					
6	Male	105	80	2					
7	Female	55	55	0					
8	Female	51	46	0					
9	Male	68	64	0					
10	Female	43	41	0					
11	Male	67	65	0					
12	Female	84	60	16					
13	Female	60	56	8					
14	Female	50	50	14					
15	Female	60	55	0					
16	Male	92	80	0					
17	Male	65	60	0					
18	Male	64	65	15					
19	Female	56	50	7					

Data View / Variable View

SPSS Processor is ready

Variable View

To view the variable view, click the **Variable View** tab at the bottom of the window.

Variables are created in the **Variable View** window. Each row corresponds to one variable, and the columns contain information about each variable.
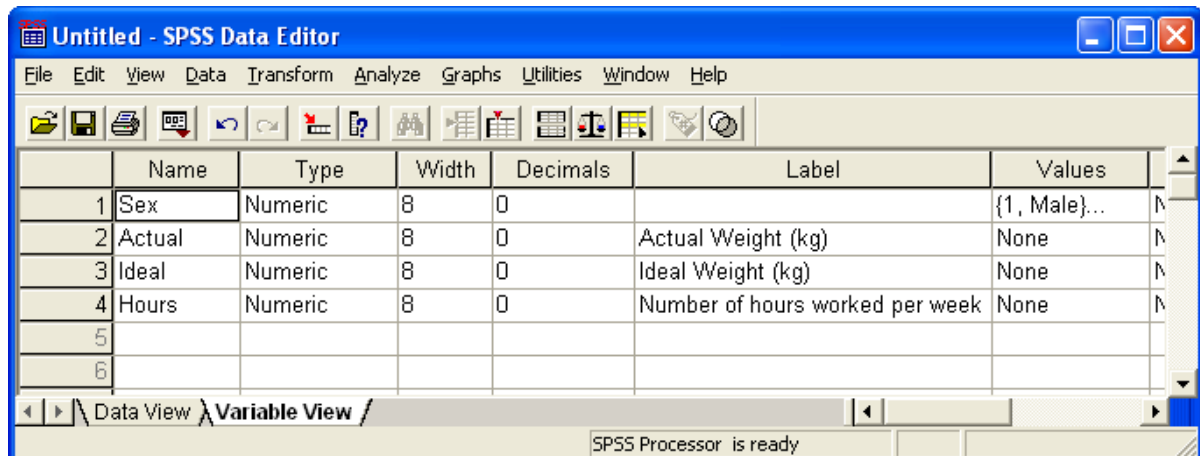
Name

This contains the name of the variable. Variable names are abbreviations of what the variable is. There is a 64-character limit, the name must start with an upper or lowercase letter and may include any combination of letters, numbers and underscores but no other characters are allowed.

To enter the name, click in the cell and type the variable name.

Type

This contains the type of the variable – usually Numeric. When entering qualitative data, you should always use the Numeric type. You can then use numbers (e.g. 1, 2, 3) when entering data and value labels to make them more meaningful.

If you want to change the type, click in the cell and then click ⊡.

Choose the type and click **OK**.

Width

If the variable is of type String, this represents the maximum number of characters the string can contain. You should make sure this is high enough to allow all of your data to be entered correctly.
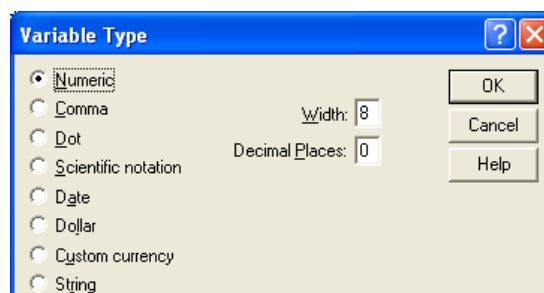
To enter the width of the string, click in the cell and type the width.

Decimals

If the variable is of type Numeric, this represents the number of decimal places the variable can contain.

To enter the number of decimal places, click in the cell and type the number.

Label

Variable labels are used to give a more meaningful description of the variable than the variable name. Understanding the computer output is made easier by using variable labels. They can be of any length, and spaces are allowed.

To enter a label, click in the cell and type the label.

Values

Value labels are used to give more meaningful descriptions of the numerical values used for qualitative data. Understanding the computer output is made easier by using value labels.

They can be of any length, and spaces are allowed.

To enter value labels, click in the cell and then click ⊡.

In the **Value** box, type one of the values that your variable takes.

In the **Value Label** box, type the label that you want that value to have.

Click **Add**.

Repeat this process for all values. Then click **OK**.
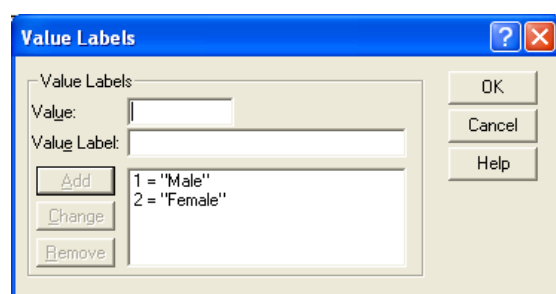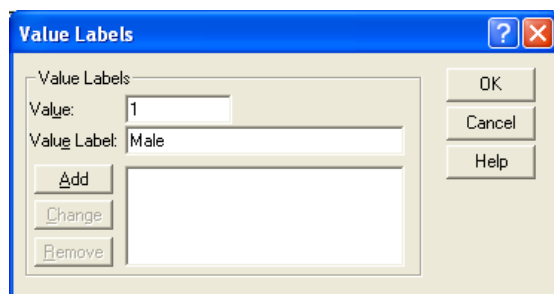
Entering data

1.   Set up the variables in the **Variable View**.

Click **Variable View**. Enter each variable that you will be using.

2.   Type in the data in the **Data View**.

Click **Data View**. Enter all the data.

Generating Descriptive Statistics and Creating a Stem and Leaf Plot and a Box Plot in PASW (SPSS)

Example: Generate descriptive statistics and create a stem-and-leaf plot and a box plot for the breaking strengths of gear teeth in certain positions of a gear.

1. Enter the data into PASW (SPSS) or open the GearTeeth.sav file.

   Label **strength** as **Breaking Strengths of Gear Teeth**.
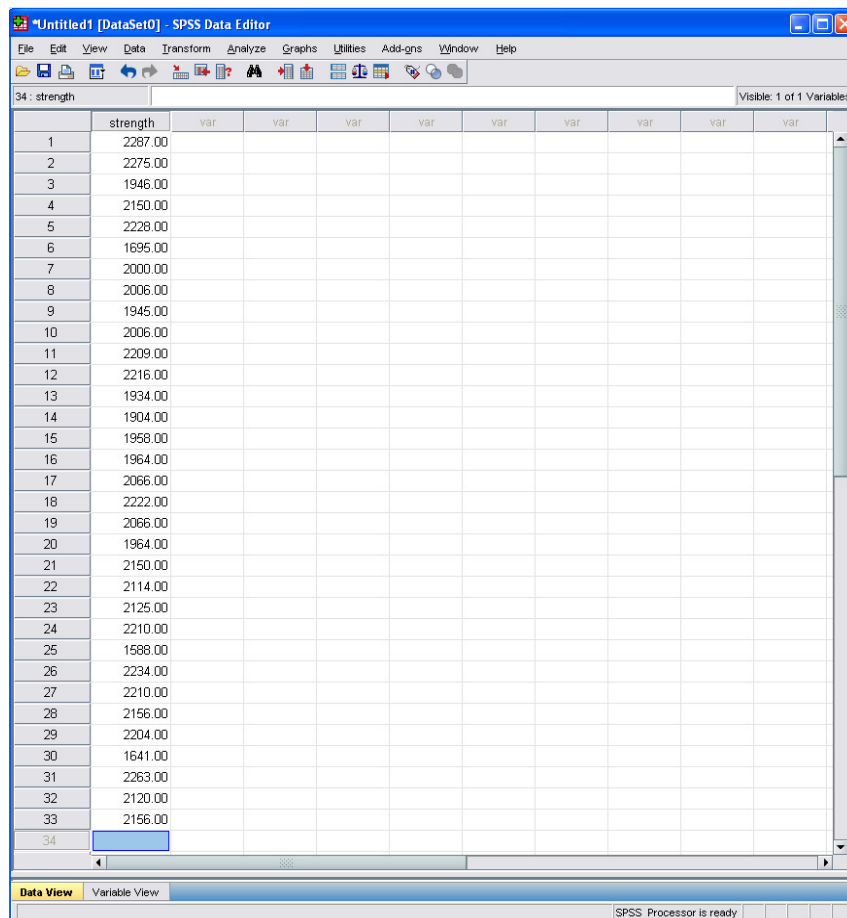
2. Choose the analysis tool: **Explore.**

   Click **Analyze → Descriptive Statistics → Explore**.
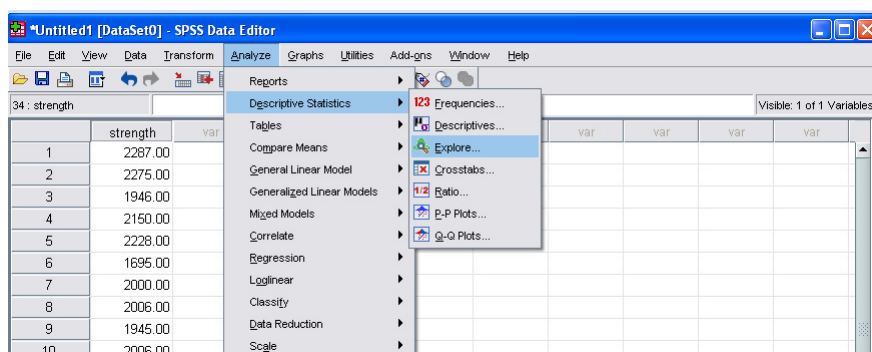
3. Select the relevant variable(s).

   Quantitative variable(s) → **Dependent List** box.

   Click **Breaking Strengths of Gear Teeth [strength]**. Click the first . Then click **OK**.

4. The results appear in the Output Window.

Explore

Case Processing Summary

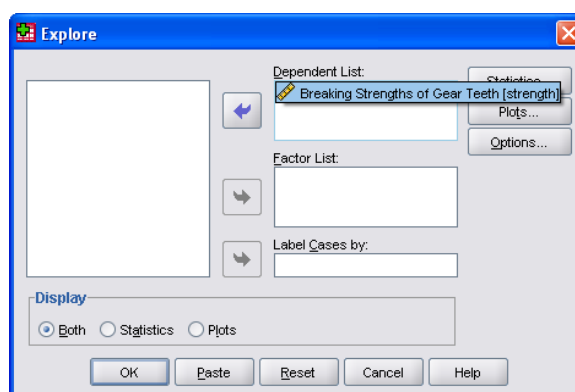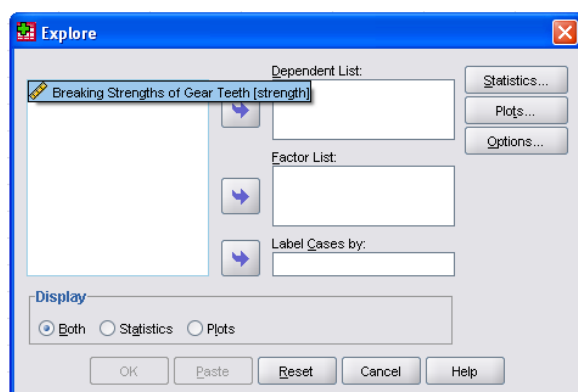|	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Breaking Strengths of Gear Teeth	33	100.0%	0	.0%	33	100.0%

Descriptives

|			Statistic	Std. Error
Breaking Strengths of Gear Teeth	Mean		2067.0303	31.01907
	95% Confidence Interval for Mean	Lower Bound	2003.8465	
		Upper Bound	2130.2141	
	5% Trimmed Mean		2080.5253	
	Median		2120.0000	
	Variance		31752.030	
	Std. Deviation		178.19099	
	Minimum		1588.00	
	Maximum		2287.00	
	Range		699.00	
	Interquartile Range		249.00	
	Skewness		-1.152	.409
	Kurtosis		1.060	.798

Breaking Strengths of Gear Teeth

```
Breaking Strengths of Gear Teeth Stem-and-Leaf Plot

 Frequency    Stem &  Leaf

     1.00 Extremes    (=<1588)
     2.00      16 .  49
      .00      17 .
      .00      18 .
     7.00      19 .  0344566
     5.00      20 .  00066
     7.00      21 .  1225555
    11.00      22 .  00111223678

 Stem width:     100.00
 Each leaf:       1 case(s)
```
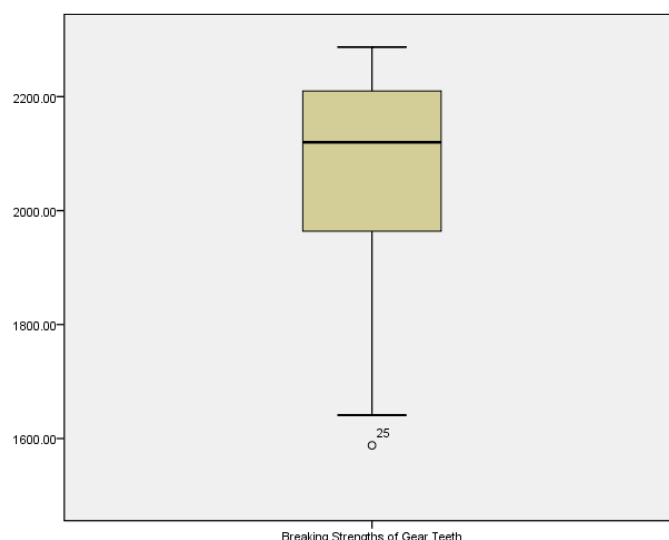
Creating a Scatter Plot in PASW (SPSS)

Example: Create a scatter-plot of the female coyote length and weight data.

1. Enter the data into PASW (SPSS)
    OR import <u>Length_and_Weight_Data.xls</u> into PASW (SPSS)
    OR copy the numerical values from *Excel* to PASW (SPSS) and enter
    the variable names **Length** and **Weight** in the **Variable View**.
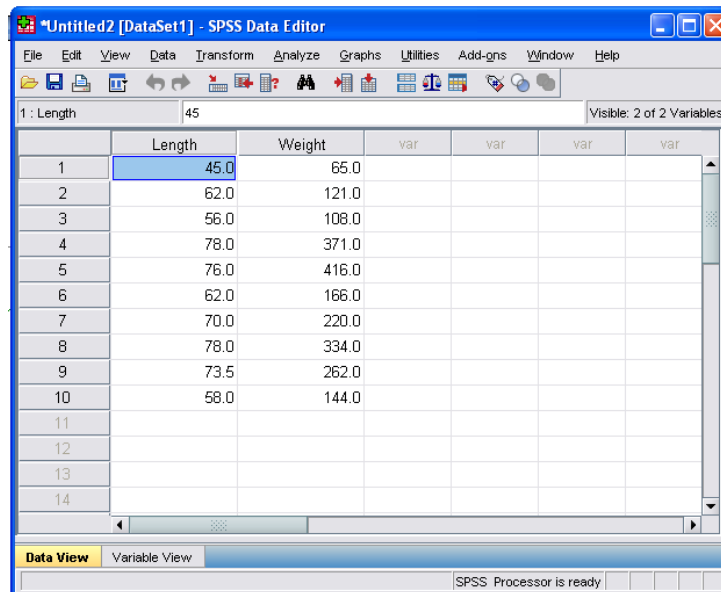
2. Choose the graph type.

    Click **Graphs** → **Legacy Dialogs** → **Scatter/Dot**.
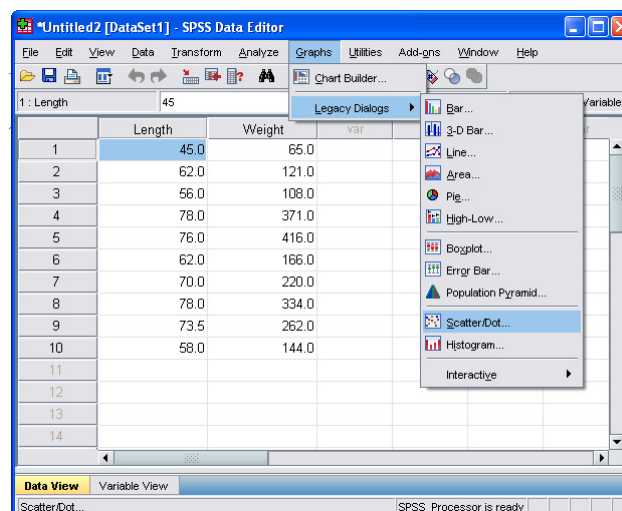
    Click **Simple Scatter**.
    Click **Define**.
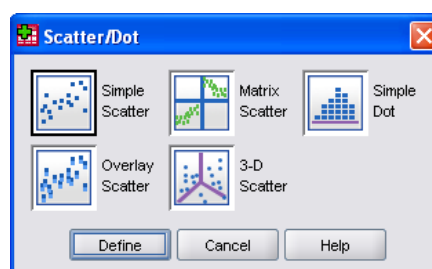
3.   Select the relevant variables.

        Dependent variable → **Y Axis** box.
        Click **Length**. Click the first ⬛.

        Independent variable → **X Axis** box.
        Click **Weight**. Click the second ⬛.
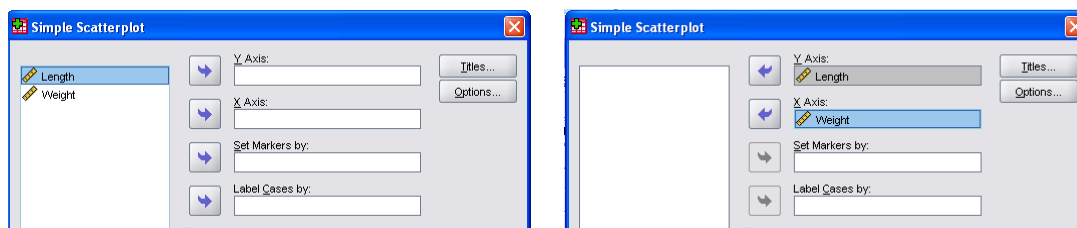
4.   Enter titles.

        Click **Titles**.
        Type **Scatterplot of Length versus Weight** in the **Title Line 1** box.
        Click **Continue**. Then click **OK**.
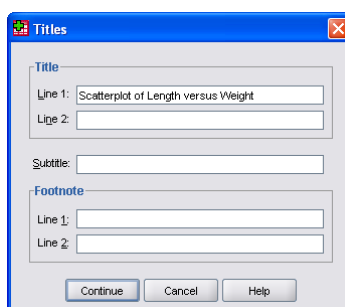
5.   The graph appears in the Output Window.
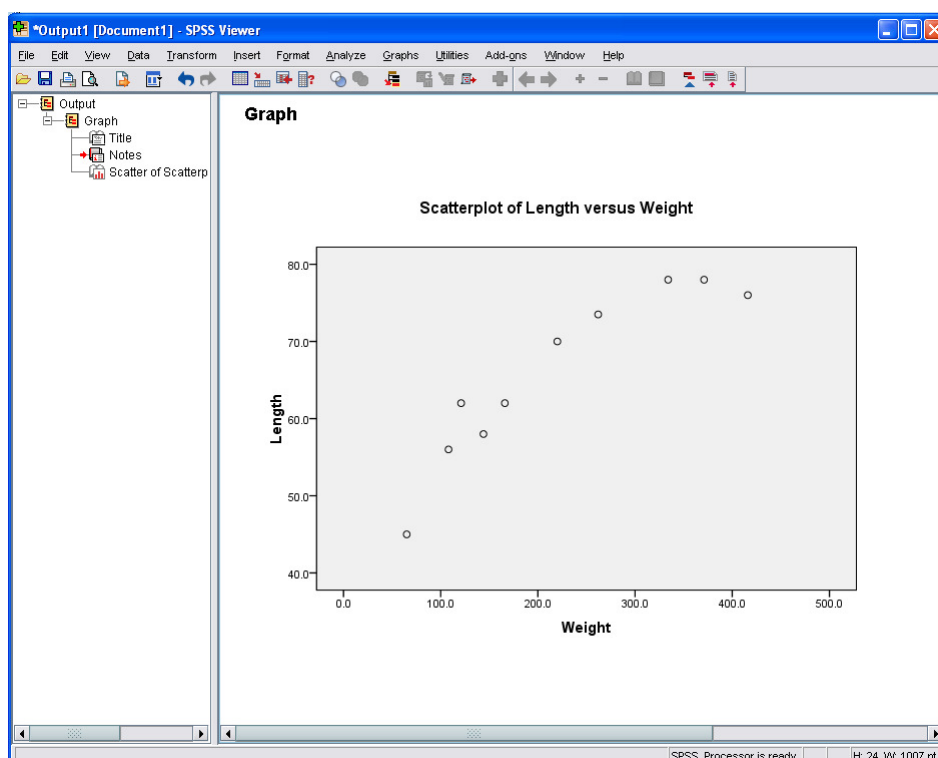
Excel Basics

In STATS 101/108 you will be using Excel to present tables of data and perform simple calculations, as well as calculating probabilities and plotting data.

The following exercise will introduce you to basic concepts in Excel that will help you to get started and should be useful for what you are required to do in STATS 101/108 Assignment 1.

Presentation of Data in Tables

In all aspects of your university study and professional career, good, clear presentation of data and information is essential to the success with which your audience receives your information. Simple features within Excel can really enhance how your data looks and improve comparability of data sets.

After entering the following data set into Excel we will use features, such as, *Column width, Merging cells, Bold, Underline, Wrapping text, Autosum, Copying, Centering, Borders, Sorting, Function Wizard (Average), Decimals,* to improve the presentation of the data.

From this…

	A	B	C	D	E	F
1	Japanese Car Data					
2	1991-1993 models			1994-1996 models		
3	Make	Trouble f	Had Prob	Trouble f	Had Problems	
4	Mazda	44	41	46	33	
5	Toyota	212	196	123		
6	Mitsubish	110	134	89		
7	Nissan	88	120	80		
8	Subaru	37	36	22		
9	Honda	82	70	80		
10						

...to this...

	A	B	C	D	E
1	Japanese Car Data				
2		1991-1993 models		1994-1996 models	
3	Make	Trouble free	Had Problems	Trouble free	Had Problems
4	Mazda	44	41	46	33
5	Toyota	212	196	123	87
6	Mitsubishi	110	134	89	84
7	Nissan	88	120	80	74
8	Subaru	37	36	22	13
9	Honda	82	70	80	68
10					

...to this...

Japanese Car Data							
	1991-1993 models				**1994-1996 models**		
Make	**Trouble free**	**Had Problems**	**Total**	**Trouble free (%)**	**Trouble free**	**Had Problems**	**Total**	**Trouble free (%)**
Subaru	37	36	73	51%	22	13	35	63%
Toyota	212	196	408	52%	123	87	210	59%
Mazda	44	41	85	52%	46	33	79	58%
Honda	82	70	152	54%	80	68	148	54%
Nissan	88	120	208	42%	80	74	154	52%
Mitsubishi	110	134	244	45%	89	84	173	51%
Total	573	597	1170		440	359	799	
Average	95.5	99.5	195	49%	73.3	59.8	133.2	56%

To open Excel

EITHER:

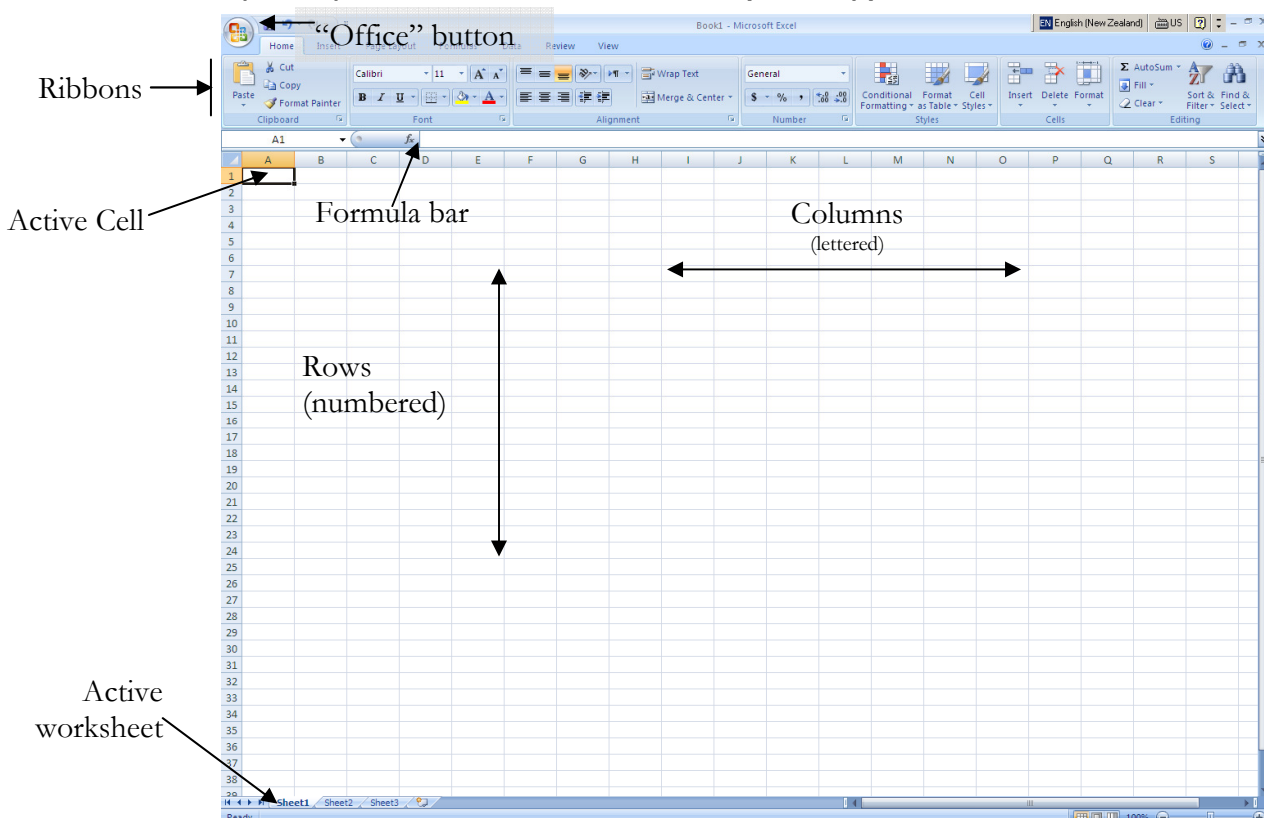- Click on the ***Start*** button, click on "New Office Document" and then select "Blank Workbook"

OR:

- Click on the ***Start*** button, click on "Programs" and then select "Microsoft Office Excel 2007" (it may be in a folder labelled "Microsoft Office")

OR:

- Double click on the Microsoft Excel icon on the desktop (if there is one)

Excel will open up a new **workbook** with (usually) 3 blank **worksheets**:

Entering Data

When you open a workbook, by default, Excel selects cell A1 as the **active cell**.

1. Begin typing your data into the active cell.

2. Press the **Enter** key to move the active cell down (e.g. A2)

   OR

   Press the **Tab** key to move the active cell to the right (e.g. B1)

3. To activate a random cell, move the cursor to that cell with the mouse and click on the mouse button.

4. To select a group of adjacent cells, click on the first cell and drag the cursor across the adjoining cells.

5. To select a group of random cells, click the first cell, hold down the CTRL key and click the additional cells.

	A	B	C	D	E	F
1	Japanese Car Data					
2	1991-1993 models			1994-1996 models		
3	Make	Trouble f	Had Prob	Trouble f	Had Problems	
4	Mazda	44	41	46	33	
5	Toyota	212	196	123	87	
6	Mitsubish	110	134	89	84	
7	Nissan	88	120	80	74	
8	Subaru	37	36	22	13	
9	Honda	82	70	80	68	
10						

In order to save time in this workshop, please feel free to download the spreadsheet Car_Data.xls from Leila's webpage: www.stat.auckland.ac.nz/~leila which is the same as the table above with just a few numbers missing.
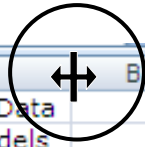
Column Widths / Row Widths

To change the column width to display all the data clearly,

1. Move the cursor with the mouse to the right of the column heading of the column you want to widen (i.e. A B C D, etc). The cursor will turn into a **cross** symbol.

2. Double click the mouse and the column will automatically widen to the width of the contents.

   OR

   Click and drag the column to widen manually.

	A	B	C
1	Japanese Car Data		
2	1991-1993 models		
3	Make	Trouble f	Had Pro
4	Mazda	44	4
5	Toyota	212	19
6	Mitsubishi	110	13
7	Nissan	88	12
8	Subaru	37	3
9	Honda	82	7
10			

3. Row widths are adjusted the same way, except you place the cursor between the row headings.

To insert Columns or Row

To insert a column within a data set:

1. Select the column to the right of where you want to position the new column, by clicking on the column heading (i.e. A B C D, etc).

2. Right click your mouse while your cursor is over the column heading.
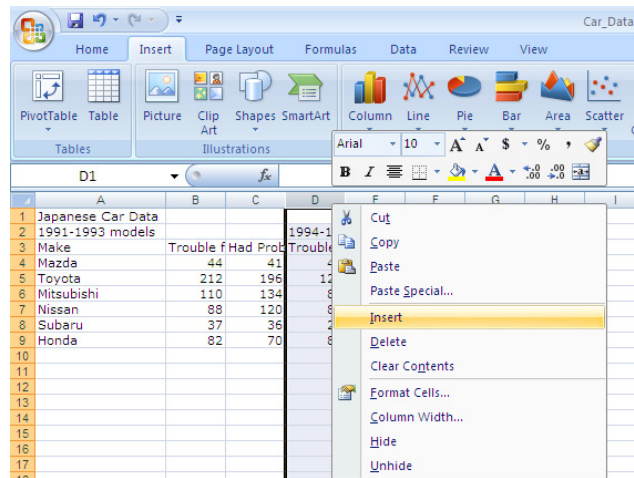
3. Click on **Insert**.

To insert a row within a data set:

1. Select the row below where you want to position the new row, by clicking on the row heading (i.e. 1 2 3 4, etc).

2. Right click your mouse while your cursor is over the row heading.

3. Click on **Insert**.

Simple Formatting - **Bold**, <u>Underline</u>, *Italics*

To **bold,** *italicise* or <u>underline</u> text

1. Select the cell(s) you want to format

2. Click on either the **B** to **bold**, *I* to *italicise* or <u>U</u> to <u>underline</u> buttons on the toolbar

   OR
   Press Ctrl + B to **bold**, Ctrl + I to *italicise* or Ctrl + U to <u>underline</u>.

3. To turn off the formatting, repeat one of the options in step 2 above.

Merging Cells

To centre a heading across the width of the data set

1. Select the cell that contains the heading and drag the cursor across the columns that represent the data set.
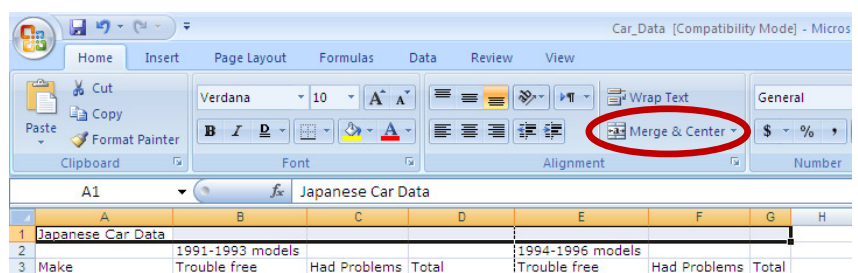
2. Click on the **Merge and Center** button on the toolbar.

AutoSum Function

To automatically total a row or column of numbers

For a row of numbers, select the empty cell directly to the right of the numbers, then:

1. Click the **AutoSum** button on the right hand side of the **Home** ribbon

2. The SUM formula including the range of numbers selected will appear in the formula bar

3. Check you have the correct range and then press the **Enter** button.

   OR

For a column of numbers, select the empty cell directly below the numbers, then continue with steps 1 to 3 above.
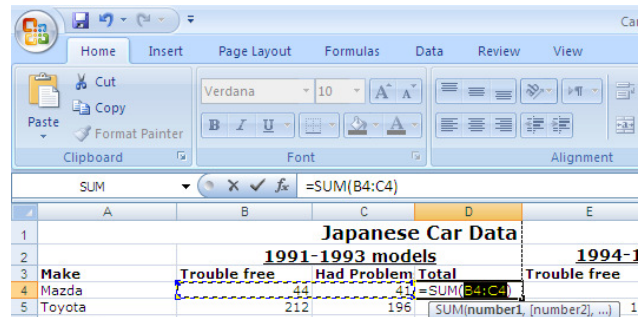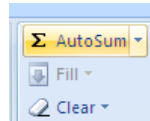
Copying & Pasting formula

To copy the same formula relative to a different group of cells

1. Select the cell that contains the formula you want to copy

2. Click on the **Copy** button on the **Home** ribbon **OR** press CTRL + C

3. Select the cell or group of cells where you want to paste the formula

4. Click on the **Paste** button on the **Home** ribbon **OR** press CTRL + V
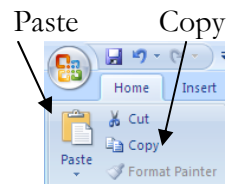
If the group of cells is adjacent to the original cell containing the formula you want to copy you can drag and paste the formula by clicking on the little square box in the bottom right hand corner of the active cell containing the formula you want to copy and dragging the cursor down (for rows) or across (for columns) to paste the formula.
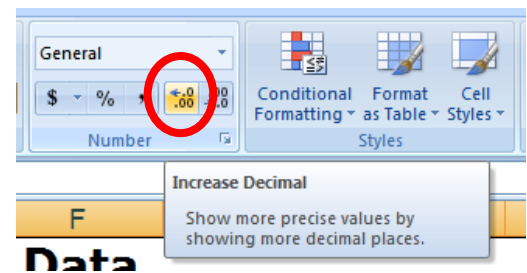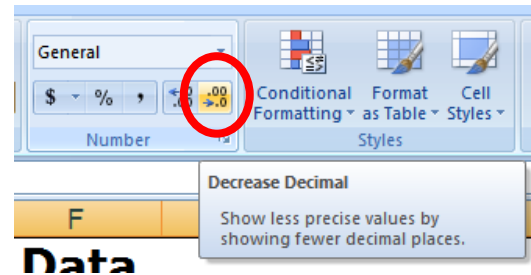
Rounding

To round a decimal value to a whole number or simply increase or decrease the number of decimal places a value has:

1. Select the cell/s containing the value/s you want to alter

2. Click on the **Decrease Decimal** button to remove decimal places

OR

3. Click on the **Increase Decimal** button to add decimal places

Simple calculations

To carry out simple calculations:

1. Think about how you would do the problem in your head or on your calculator.

2. Think about how Excel may carry this out and get Excel to do it by clicking on the relevant cell/s and using the appropriate combination of:

addition	+
subtraction	-
multiplication	*
division	/
brackets	()
to the power of	^

3. If the same formula is to be applied more than once, simply copy and paste it as per **Copying & Pasting formula** instructions on page 17.

4. Round appropriately using the "Decrease Decimal" button or the "Increase Decimal" button as per **Rounding** instructions above.

Percentages

To create a percentage from a proportion (decimal), simply use the **Percent Style** button. This will automatically display the proportion (decimal) to the nearest whole percent, so:
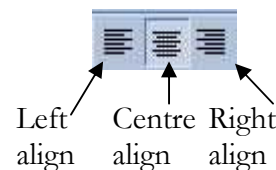
0.003	will become	0%
0.03	will become	3%
0.3	will become	30%
3	will become	300%
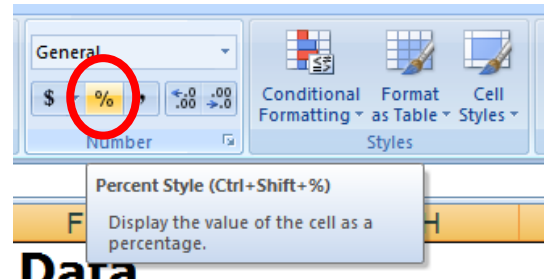
1. Select the cell/s containing the number/s you want to alter

2. Click on the **Percent Style** button to round to the nearest whole percentage

3. Click on the **Increase Decimal** button to add decimal places if appropriate.

Align Text

To align text in a cell to the left, centre or right

1. Select the cell(s) to be aligned

2. Click on either the **Left align**, **Centre align** or **Right align** buttons.

Borders

Borders can be applied individually to each of the 4 sides of a cell or applied to a group of cells. Borders allow you to separate data groups and highlight specific sections of data. To add a border to a cell or group of cells

1. Select the cell(s) that you want to place a border around

2. Click on the down arrow ▾ of the **Borders** button on the toolbar. A range of different border buttons will be displayed.

3. Click on the **Borders** button you want for the selected cells.
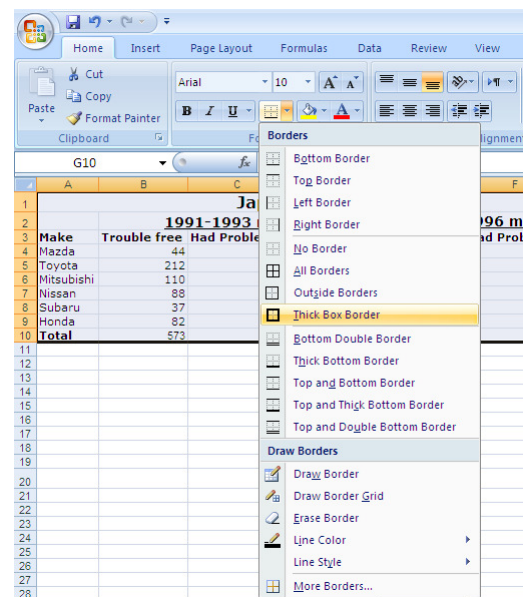
Sorting Data

To sort data numerically or alphabetically

1. Select the rows of data you want to sort

2. Click on the **Sort and Filter** button on the **Home** ribbon.

3. Click on **Custom Sort**.

4. Next to the **Sort By** heading, click on the down arrow ▾ to select the column on which you want to sort the data

5. Select either **Smallest to Largest** or **Largest to Smallest**.

6. Click **OK**.
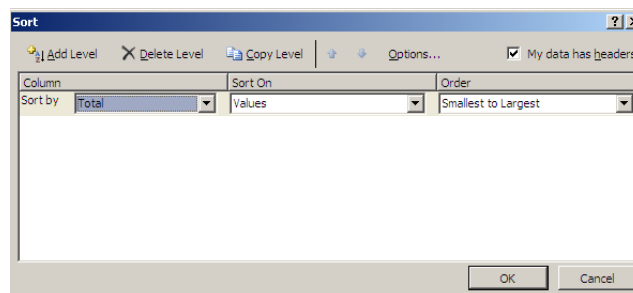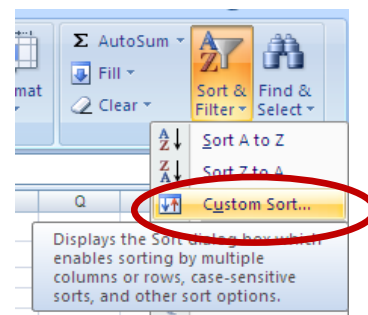
Print Preview

Print preview allows you to check whether all the columns and rows of your table will print on the paper size you have chosen.

1. Click on the "Office" button, then position your mouse over "Print" and then click on "Print Preview". Alternatively, use the "Ctrl + F2" shortcut.

2. Check that all your columns and rows appear on the one page.

3. If there are columns missing, you may have to print the worksheet in "Landscape" orientation (use the "Page Setup" button to change from "Portrait" orientation) and/or adjust the "Margins" (check the box next to "Show Margins" and adjust the page margins as appropriate). Shrink the worksheet to fit on one page. Click on the Close button (of the Print Preview screen) and adjust the width of the columns to make them smaller. Remember that you still need to see the data in each column.
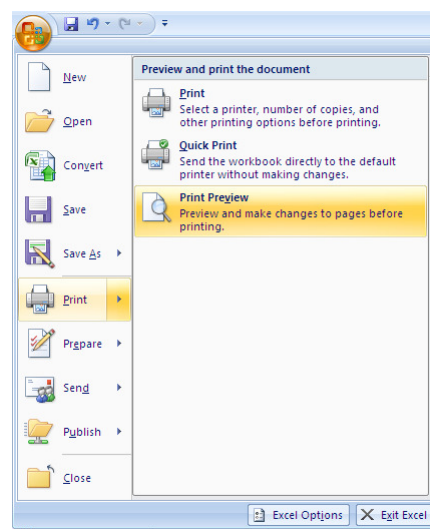
Excel Help

You can find more information in "Excel Help" which you can access by pressing the F1 key.