MATU9D2 : PRACTICAL STATISTICS

Spring 2017

PRACTICAL SESSION 1

In this Practical, which introduces Minitab, you will:

1. Import data from a spreadsheet file.
2. Enter data directly into Minitab.
3. Draw some simple graphs: Pie Charts, Bar Graphs
4. Plot more than one set of data for comparison
5. Change the scale of plots.
6. Perform arithmetic calculations using Minitab
7. Use graphical and numerical summaries to formulate your subjective impressions of data.
8. Save all the data and graphs to a file.

- Handout 2 of 2

A O'Hare
Computing Science & Maths
School of Natural Sciences
University of Stirling

EXERCISES

1. Open Minitab (***Section 2***).

2. The data on Social Class and Voting Preference which is summarised in the table below is in an Excel file on Succeed under '**Learning Content/Practicals' in a file called MATU9D2 Practical 1 data.**

   Copy the data into your Minitab file (***Section 6.2***).

3. Save the file as a Minitab Project File (***Section 7.1***).

4. Check that it is the same data by re-creating the table below using the menus in ***Section 10.2*** with Class as the Row Variable and Voting as the Column Variable.

		Political Preference			
		Tory	Labour	LibDem	Other
Social	A	110	11	37	2
Class	B	420	99	237	8
	C1	321	119	143	12
	C2	353	363	205	9
	D	183	286	134	6

Exercises 5-9 Make sure that you label the graphs

5. Draw a Pie Chart for 'Voting' (***Section 10.1.1***).

6. Draw two Bar Charts for 'Voting' (i) Counts and (ii) Percentages (***Section 10.1.2***).

7. Draw a Side by Side Bar graph for using both Social Class and Voting (***Section 10.1.2***). Are percentages or counts more appropriate?

8. Draw a Stacked Bar graph for using both Social Class and Voting (***Section 10.1.2***). Are percentages or counts more appropriate?

9. Draw a Side by Side Bar graph for only Tory and Labour Voters (***Section 10.1.2***).
   i.e. Choose 'Data Options' then 'Rows that match' and enter the Condition
   Voting= "Conservative" Or Voting="Labour"

10. Enter the data from Question 1 of the By Hand Exercises from Practical 1 (see below) into a single column. (***Section 5.1***)

The following yields (kg) were obtained from plots of a fixed size in a field of potatoes growing under the same fertiliser treatments.

28	21	14	17	24	19	22	21	16	26
20	24	21	19	17	15	18	22	23	20
26	18	24	20	19	23	22	18	20	22

(i) Draw a Stem and Leaf plot. (***Section 11.1***).
(ii) Calculate the Five Number Summary of the data (***Section 12***).
(iii) Draw a Box and Whisker plot. (***Section 11.2***).

11. The rainfall was measured on nine days and the results (in mm) were:

11.7 12.2 10.9 11.4 11.3 12.0 11.1 10.7 11.6

Is the mean rainfall different from 12.1 mm (the average measured over the same period last year)? (***Sections 11 & 12***).

12. The birth weights (kgs) of 36 babies born after normal pregnancies of 40 weeks were:

3.5	4.1	2.8	3.2	2.8	3.1	3.4	3.0	2.3
3.8	2.7	3.7	3.9	2.6	2.7	3.1	2.2	2.9
3.2	3.7	3.3	4.3	3.4	3.5	4.6	3.1	3.4
3.5	3.5	3.8	2.4	3.0	3.6	4.0	2.9	3.3

Is the mean weight of babies born after a normal pregnancy 3.6kgs? (***Sections 11 & 12***).

13. Drug levels (in ng/ml) in blood samples from two groups of subjects gave values of:

Group 1:	3.3	3.7	3.5	4.1	3.4	3.5	4.0	3.8	3.2	3.7
Group 2:	3.2	3.6	3.1	3.4	3.0	3.4	2.8	3.1	3.3	3.6

Is there a difference in the means? (***Sections 11 & 12***).

14. Save the file as a Minitab Project File (note that this saves all data, all results and all graphs)

15. Leave Minitab (***Section 11***)

16. Logoff (***Section 12***)

MINITAB is a statistical package which will perform calculations, exploratory data analysis, plotting and statistical tests on data. Data is stored in a worksheet with columns labelled C1, C2, ... and rows numbers 1,2, ... (This worksheet looks like an Excel spreadsheet but does not operate like one. It is only for the data!!!) Single numbers can also be stored as constants labelled K1, K2, ...

Data can be either numeric or alpha-numeric and be entered directly from the keyboard or imported from files stored by other software packages.

This series of Workshops cannot cover all the facilities available in Minitab. In fact, they will only scratch the surface. You can investigate other facilities using the Tutorial and Help options.

1. Starting Up

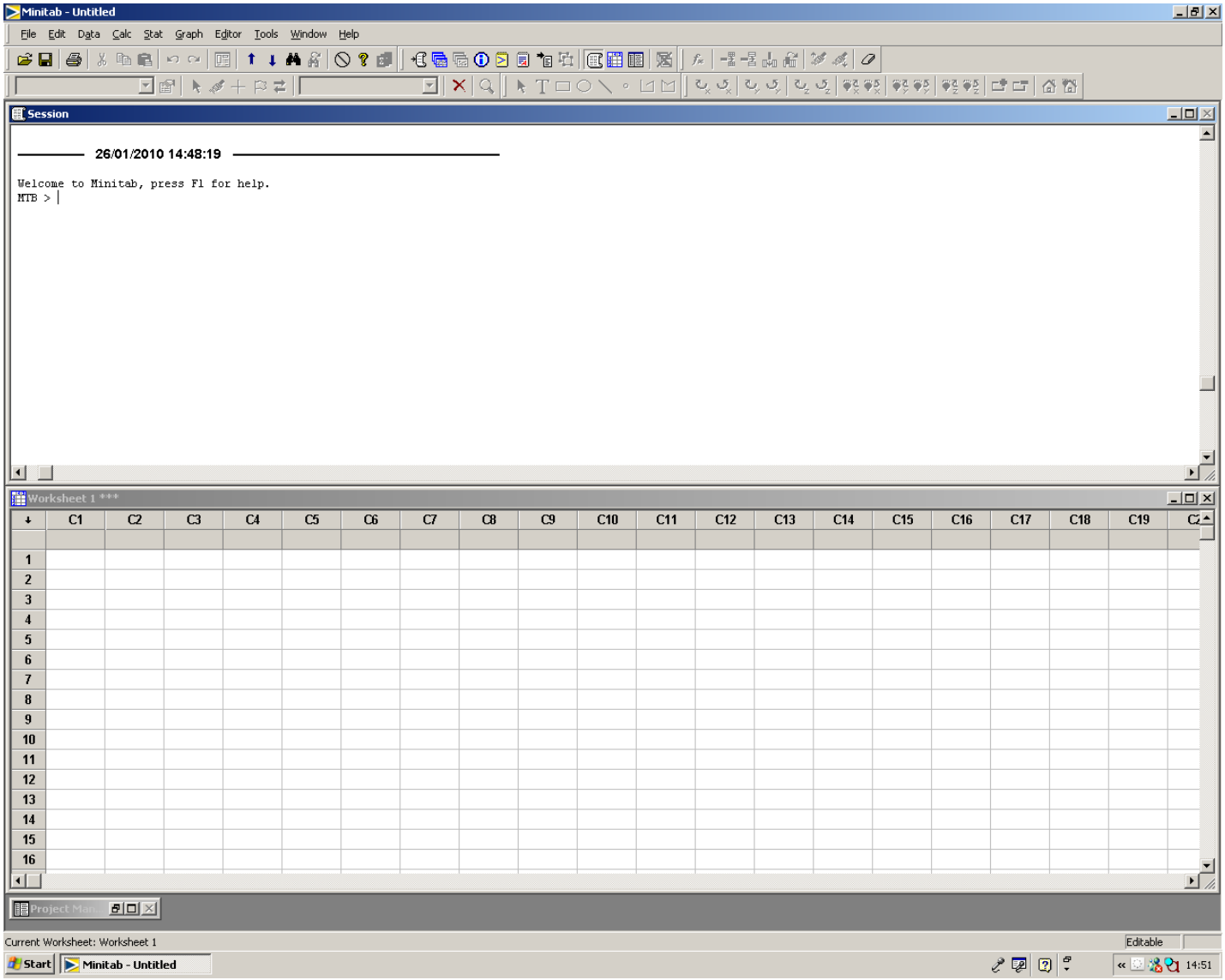
- Switch on the Computer
- When prompted enter your logon name and password. (Use the TAB key to move between the name and password positions)

  Follow the instructions on the screen.
- The screen should then include a number of icons on the left hand side.

2. Accessing MINITAB

- Click on Start at the bottom left of Screen
  then on All Programs
  then on Minitab
  then on Minitab 16 Statistical Software

The Menu Bar (containing the titles of 'pull-down' menus: e.g. File, Edit, Calc, Stat, Graph), title bar will appear, including the MINITAB Data Window (at the bottom) and Session Window (at the top). See illustration below.



3. The Windows in MINITAB

There are 4 different windows accessed in MINITAB -

Session	:	contains the commands and output from any commands resulting in non high resolution output
Data	:	shows the data rows and columns
History	:	lists the previous commands performed
Graphs	:	Windows containing high resolution plots

The History can be accessed via the Project Manager via the Window drop-down menu – see above.

- To access a new window. Click once on the WINDOW menu then on the title of the Window of your choice.
  Or click once on an already visible window or double click on an icon.

In addition, high resolution graphs are created they will be stored in 'graph windows'.

4. Tasks in MINITAB

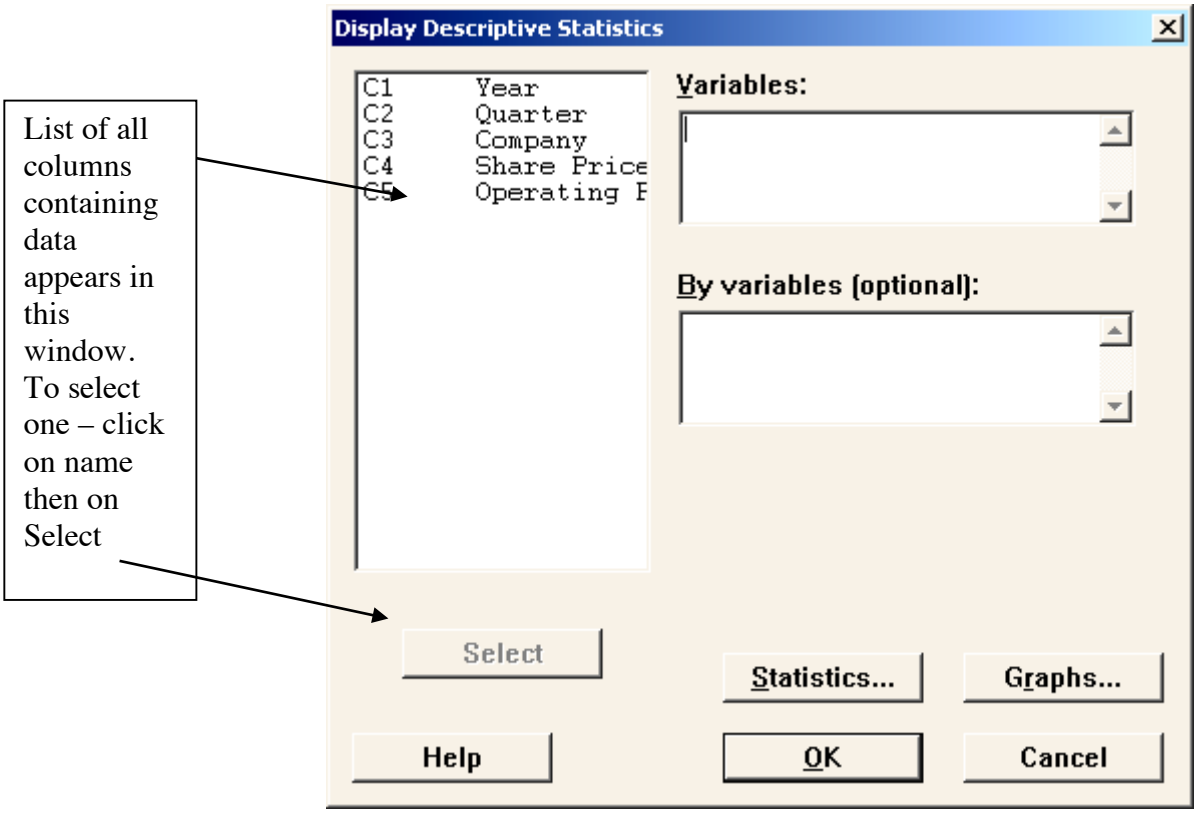
Tasks can usually be undertaken in at least two ways. It is a matter of personal choice as to which way you choose. Most new users will use the 'Drop-Down' Menus.

(i) Using the 'Drop-Down' Menus

- Move the pointer over the menu of choice, click once, then move the pointer to the option of choice then click once.

  You will have opened a 'dialogue box'. This prompts you for answers to all possible options related to the task.

  Example of a Dialogue Box.



To leave a dialogue box and perform the task - click on OK. To leave a dialogue box abandoning the task - click on CANCEL.

N.B. [?] clicking once over this button gives help information
N.B. In those cases where columns are required either type c1 c2 etc. in the box or highlight the columns in the data section by pointing and either double clicking or click on select.

(ii) Entering the Line Commands

Ensure that the Session Window is 'active' i.e. click once on the window, then the prompt MTB > will be displayed.

If this does not appear then - go to Editor and click on Enable Commands.

Then enter the appropriate command. Note that when you use the pull down menus the corresponding line command, and sub-command if applicable, will appear in the Session and History windows.

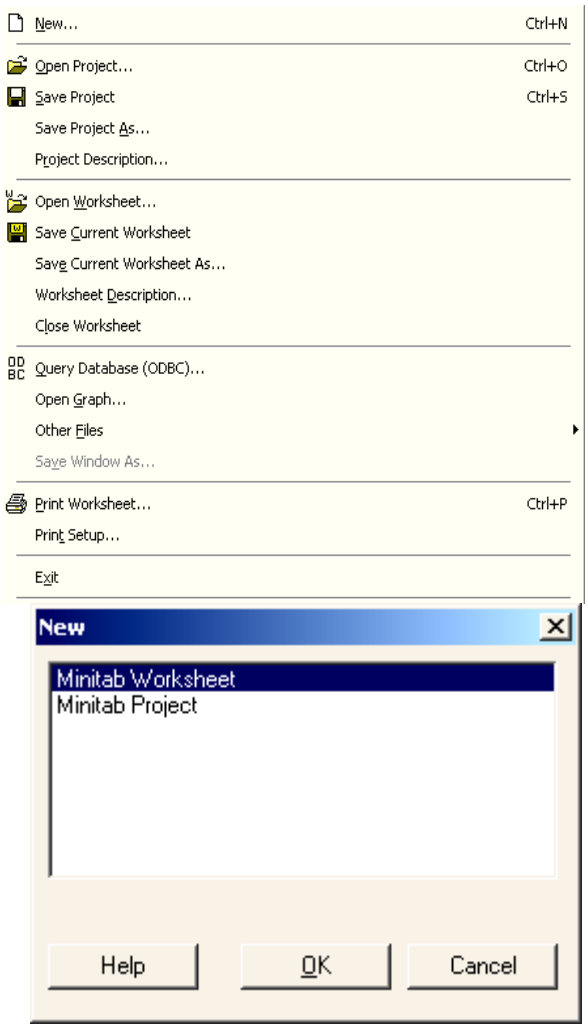
5. Opening a New Project or Worksheet

5.1 Opening a New Worksheet

If you are accessing a Minitab Project and you want a new worksheet within the project Access File Menu -> New -> Worksheet -> OK

5.2 Opening a New Project

If you have completed work within one project file and wish to open a new blank project Access File Menu -> New -> Project -> OK



6. Entering Data

6.1 Entering Data from the Keyboard

Ensure that the Data Window is active, i.e. click once on the window, enter the data in columns from row 1 downwards.

Moving from cell to cell can be done either using the mouse and clicking once on the appropriate cell or by using the arrow keys.

To name a column, move to the 'cell' immediately under the column number.

	c1	c2	c3	c4	c5
	Name1	**Name2**	**Name3**	**Name4**	**Name5**
1					
2					

If you enter a number incorrectly, (a) if you have not either pressed enter or moved to the next cell then use the backspace to delete the characters or number then repeat correctly, or (b) move back to the cell and re-enter.

If you require to delete cells or insert cells, then

- Access Edit -> Choose appropriate option

If you require to insert or delete rows or columns or edit cell contents, then

- Access Editor -> Choose appropriate option

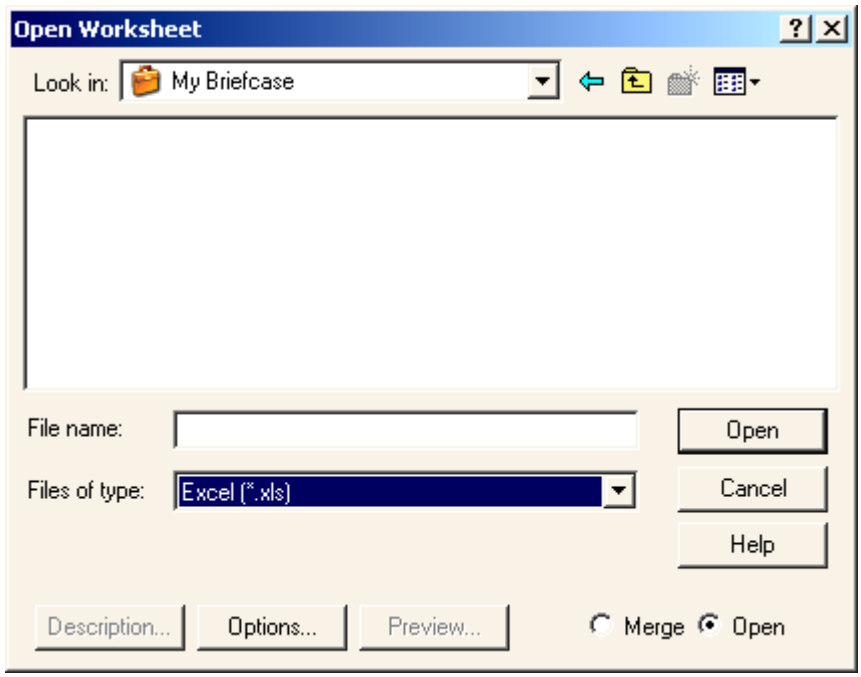
6.2 Copying Data from an Excel File

- Create a Worksheet in Excel containing variable names in row 1 and the data in rows 2 onwards. i.e. no graphs or tables. Row 1 need not contain the labels but usually would.
- You could copy and paste the data and names

OR

- Access the File Menu
- Then Open Worksheet
- Look in the correct Folder

- Change File of Type to Excel (*.xls)



- Choose the File you require then Click on Open

  If this is the correct Worksheet in the Workbook and you have variable names in first row. This command opens the first Worksheet in a Workbook.

7. Saving to a File

When you choose the **File Menu** the following option choice appears. The options listed in this section are marked.

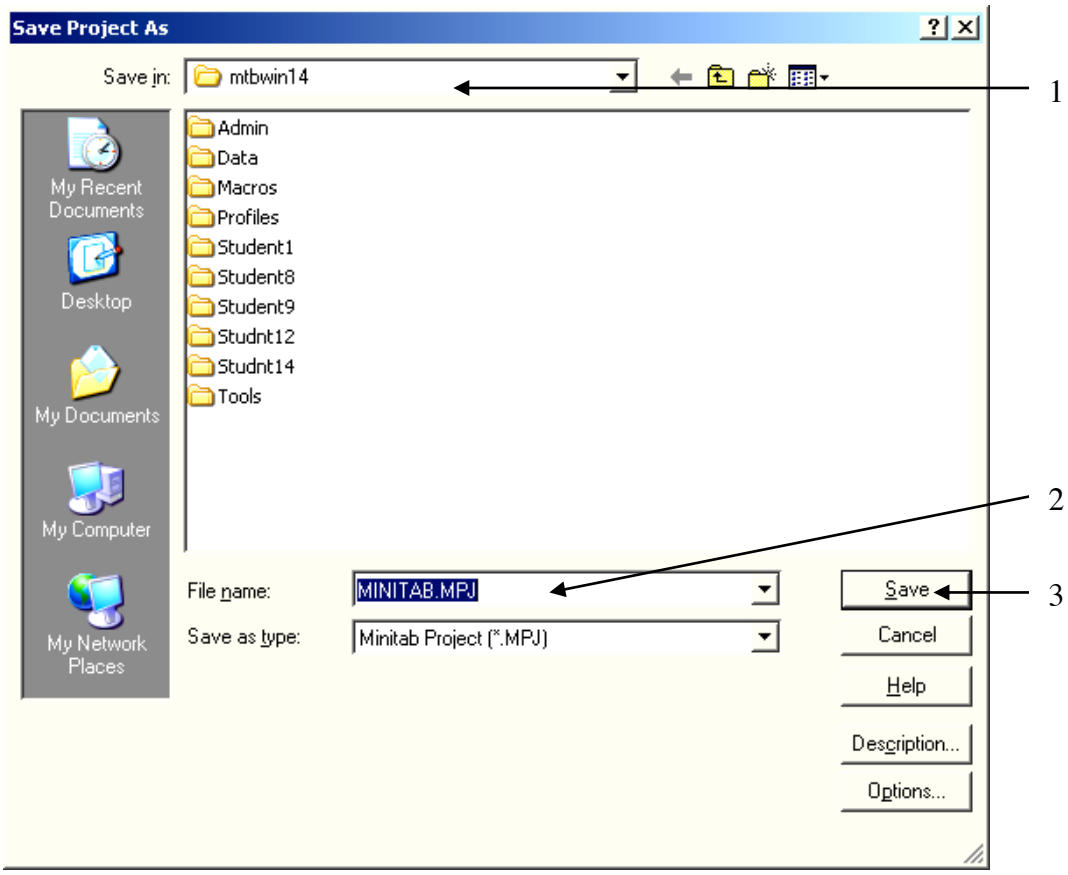
7.1 All Windows to a File

This saves all the data, any output in the Session Window and any High Resolution Graphs to a 'Project'

- Access the File Menu -> Save Project As

  You will have accessed the following dialogue box-
  1. You may need to alter the drive and directory specifications.
  2. Enter the file name under which you want to save the data.
  3. Then click Save.

N.B. Minitab automatically gives a file of this type a **MPJ extension**.



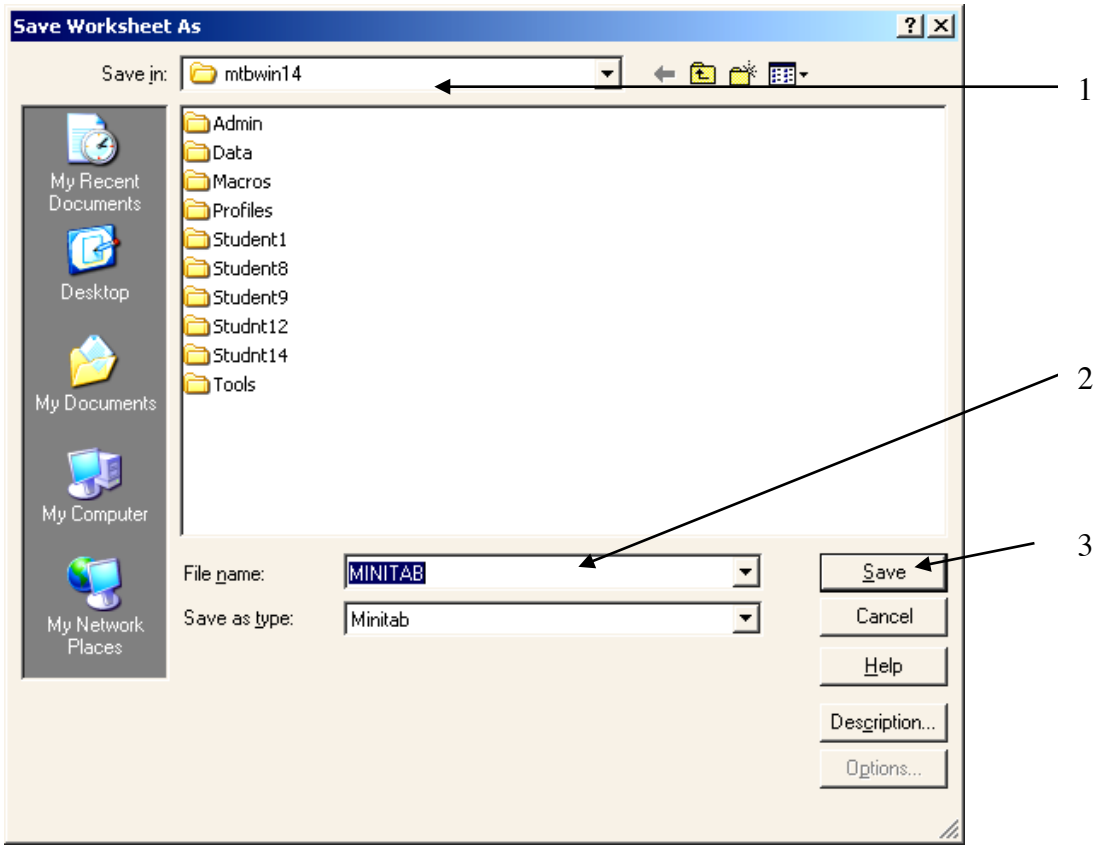
7.2 Data to a Worksheet

This saves all the data stored in all **constants, matrices and columns as well as column names.**

- Access the File Menu -> Save Current Worksheet As

  You will have accessed the following dialogue box-
  1. You may need to alter the drive and directory specifications.
  2. Enter the file name under which you want to save the data.
  3. Then click Save.

N.B. Minitab automatically gives a file of this type a **MTW extension**.



7.3 Data to an Excel Workbook

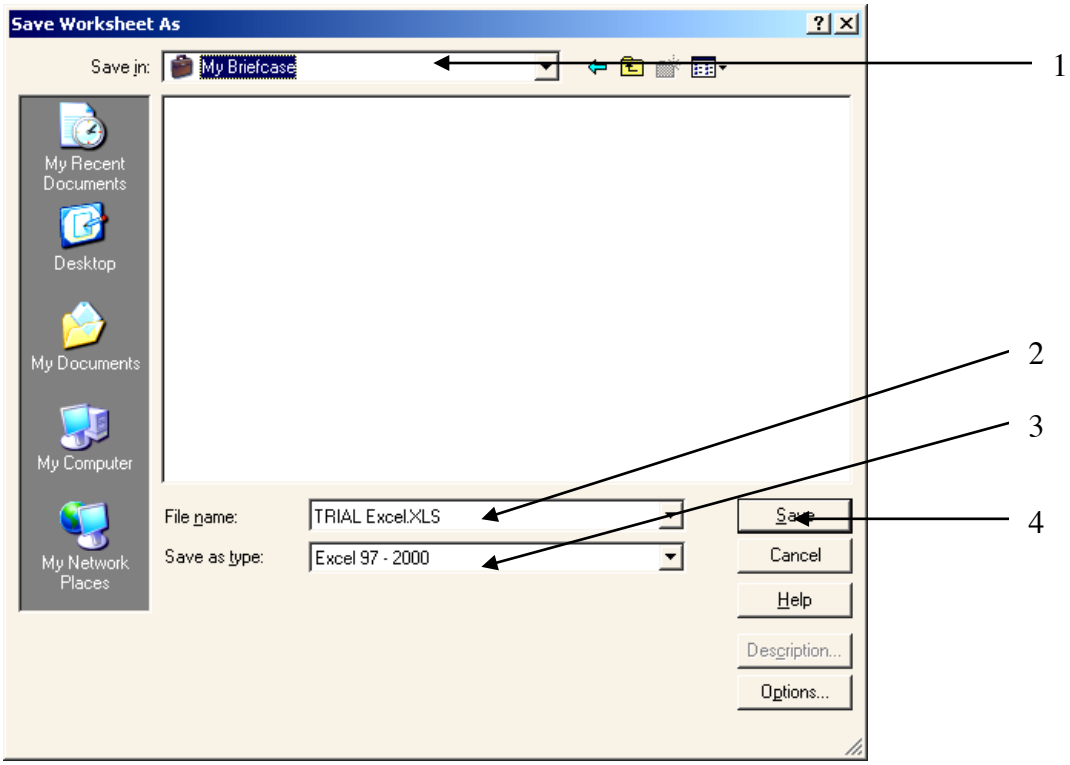
This saves the **data and column names** to an Excel Workbook.

- Access the File Menu -> Save Current Worksheet As

You will have accessed the following dialogue box-

1. You may need to alter the drive and directory specifications.
2. Enter the file name under which you want to save the data.
3. Change 'Save as type' to the correct version of Excel
4. Then click Save.

N.B. Minitab automatically gives a file of this type a **XLS extension**.



7.4 Information in the Session Window to a File

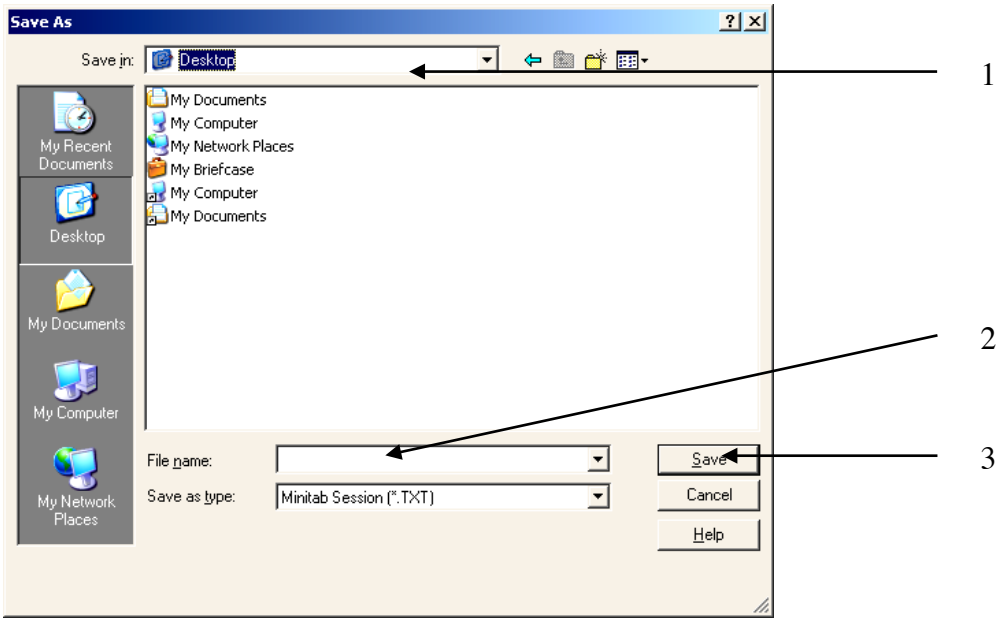
Information from the session window can be saved to an appropriate file by selecting the window to be 'active' then

- Access File Menu -> Save Session Window As -> Name of File

You will have accessed the following dialogue box-

1. You may need to alter the drive and directory specifications.
2. Enter the file name under which you want to save the data.
3. Then click Save.

N.B. Minitab automatically gives a file of this type a **TXT extension**.



8. Retrieving a File

When you choose the **File Menu** the following option choice appears. The options listed in this section are marked.

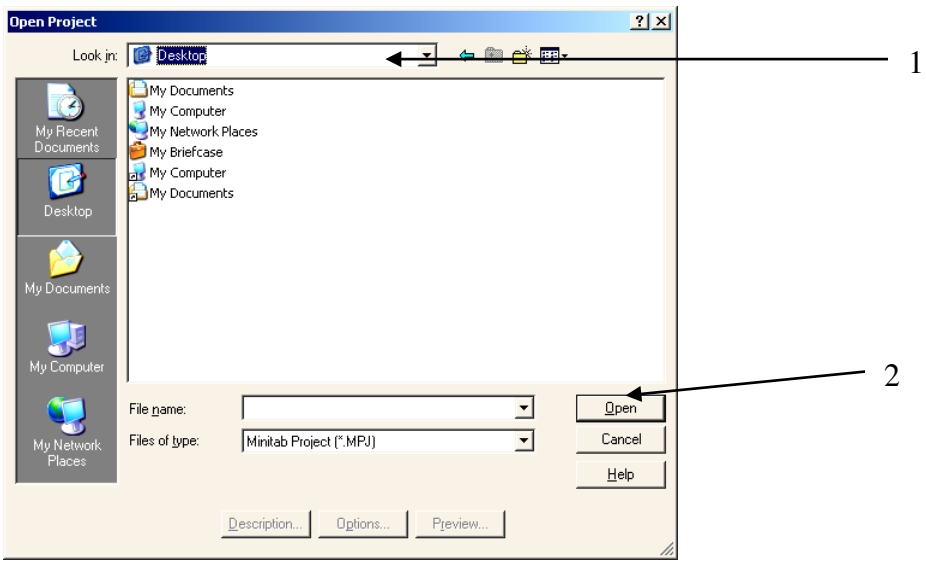
8.1 Information from a Project File

This retrieves all the information (data, results and graphs) from a project file.

- Access the File Menu -> Open Project

Once again you have accessed a dialogue box:

1. Select the drive, directory and file from which you want to import the data.
   N.B. Minitab automatically lists files with a **MPJ extension,** if the file does not have a MPJ extension you will have change the file type.
2. Click Open



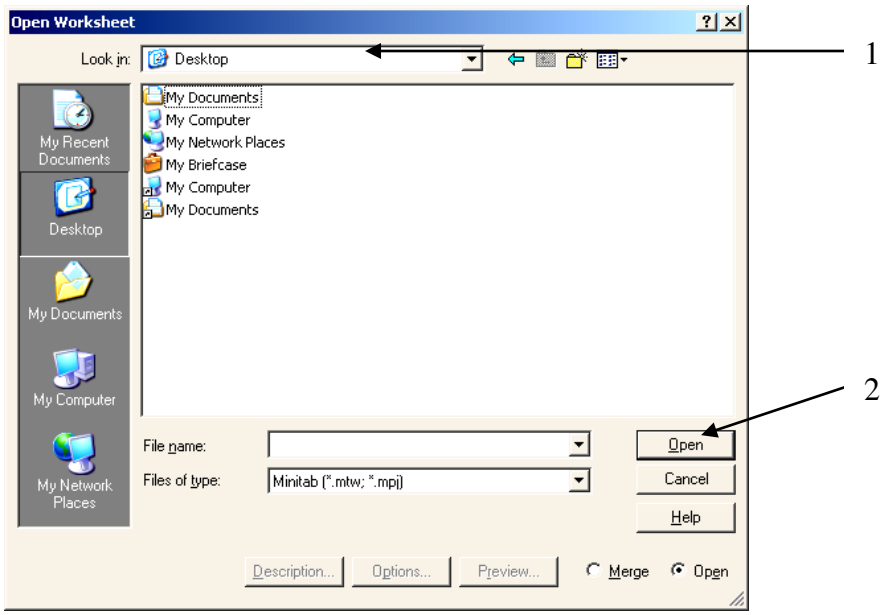
8.2 Data from a Worksheet

This retrieves all the data stored in all **constants, matrices and columns as well as column names.**

- Access the File Menu -> Open Worksheet

Once again you have accessed a dialogue box:

1. Select the drive, directory and file from which you want to import the data.
   N.B. Minitab automatically lists files with a **MTW extension,** if the file does not have a MTW extension you will have change the file type.
2. Click Open



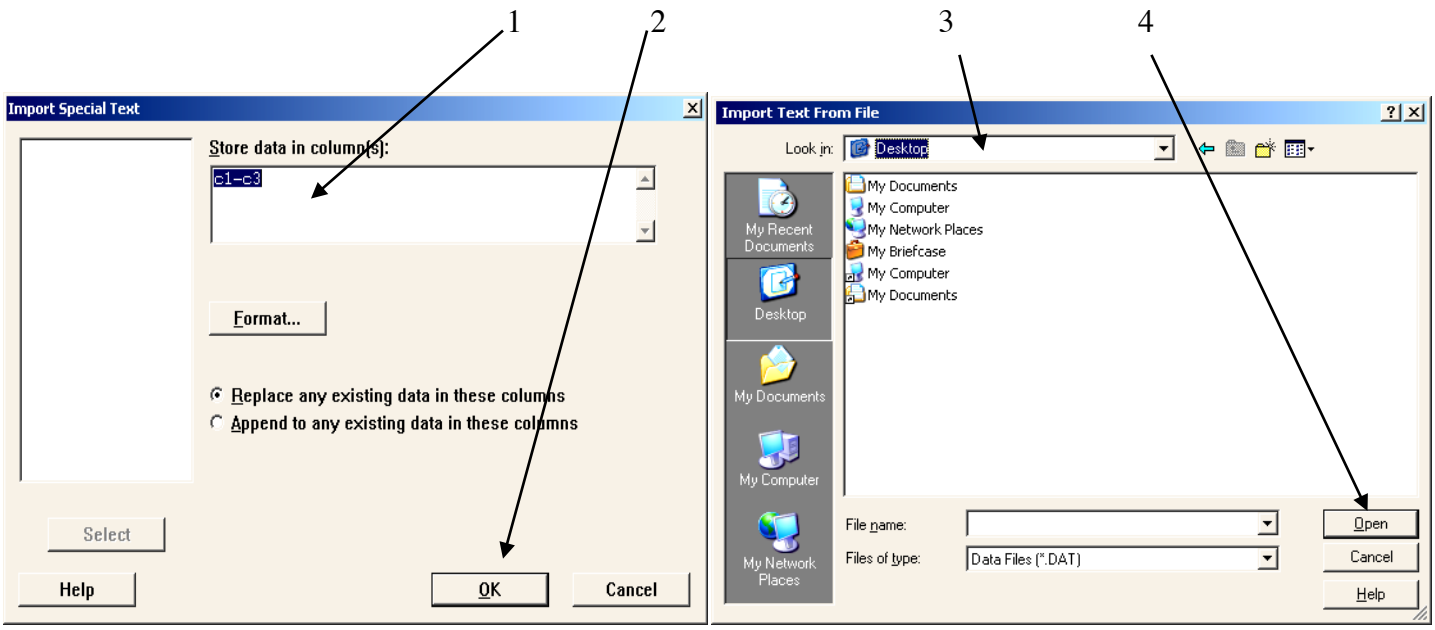
8.3 Data from a Text File

This only retrieves **data into a specific set of columns** to a file.

- Access the File Menu -> Other Files -> Import Special Text

Once again you have accessed a dialogue box:

1. Enter column numbers into which you want the data to go
2. Click OK
3. In the next dialogue box select the drive, directory and file from which you want to import the data.
   N.B. Minitab automatically lists files with a **DAT extension,** if the file does not have a DAT extension you will have change the file type.
4. Click Open.



9. Printing Information

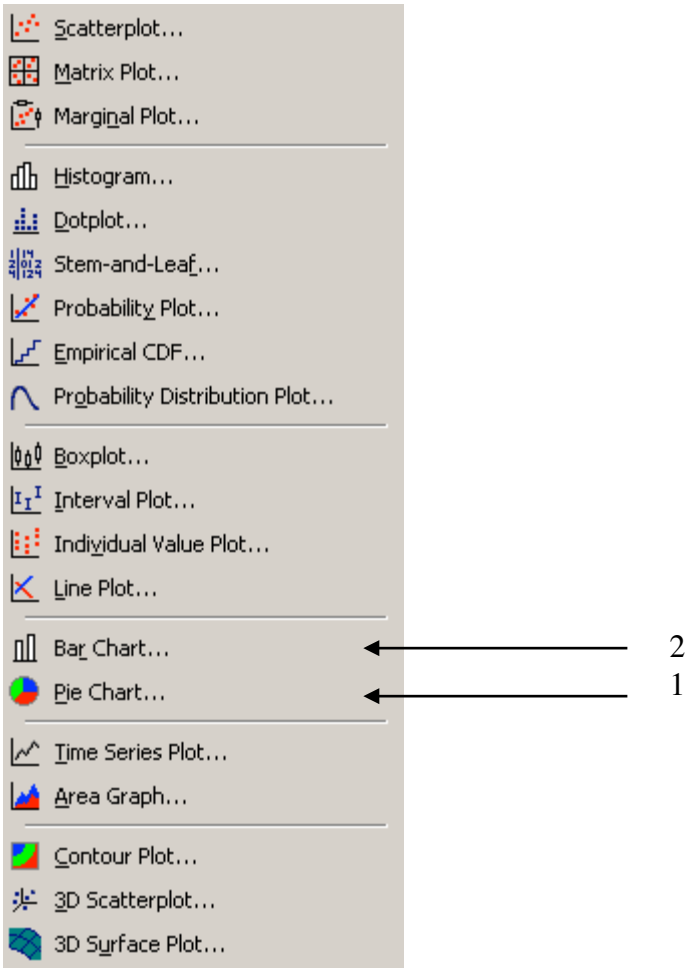
The information in any of the four windows (plus any High Resolution Graph) can be printed.

- Ensure that the appropriate Window is 'Active' i.e. click over it. Then access the File Menu -> Print Session Window or Print Worksheet -> OK

10. Exploratory Data Analysis

10.1 Graphical Summaries

When you choose the Graph Menu, the following list of options appears. The options we will use during this workshop are marked:



10.1.1 Plots – Pie Chart

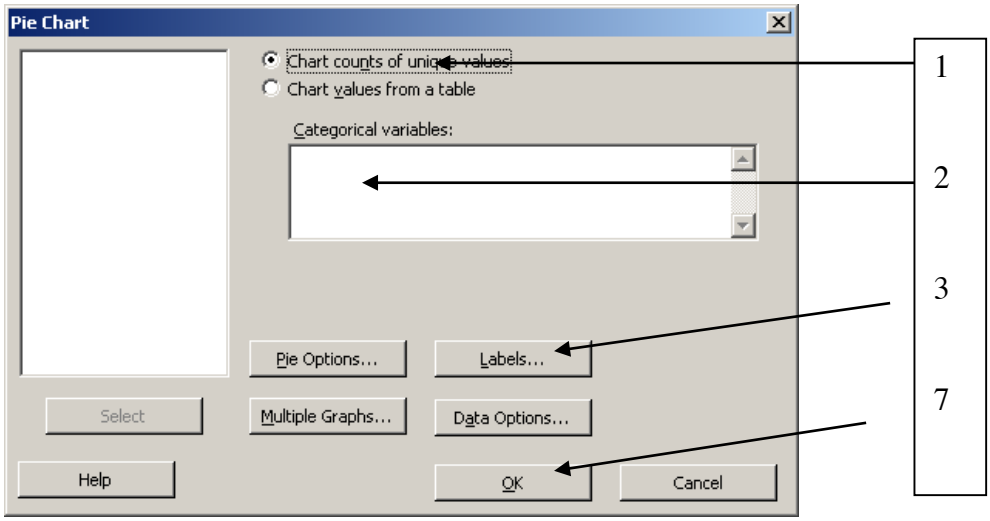
Obtain a Pie Chart of the data directly from the **full data** set by...

Access Graph -> Pie Chart

1. Choose Chart counts of unique values (dialogue box is shown on the next page)

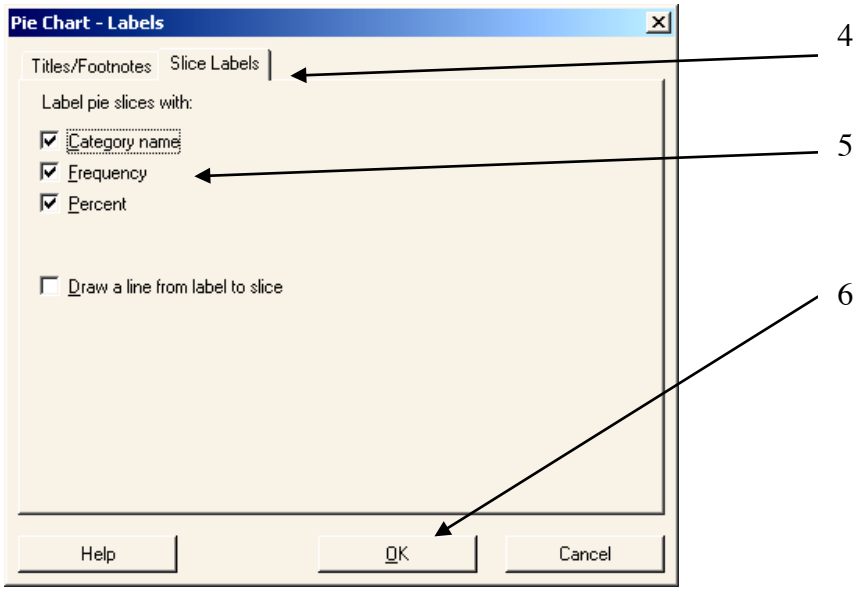
2. Then in the dialogue box select the column for which you want the plot (Categorical Variables).

3. Click Labels in order to label the plot appropriate (dialogue box is shown on the next page)



4. ‘Toggle’ between Titles/Footnotes and Slice Labels by clicking on the tabs

5. Good practice to label with both the Frequency and Percent.

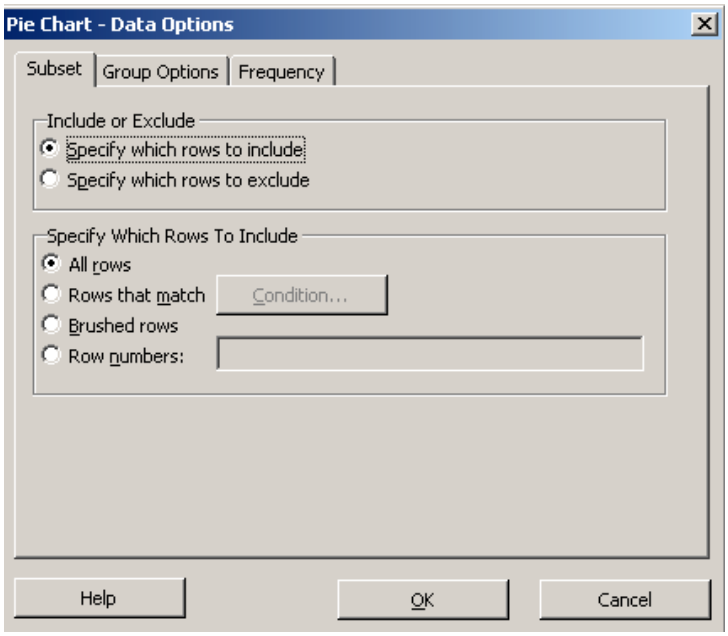


6. Click OK

   If you want to select a Subset of the Data, see over the page, if not go to 7 below.

7. Click OK

To **Select a Subset of Data**, choose Data Options in the Pie Chart dialogue box, the following options will be displayed:



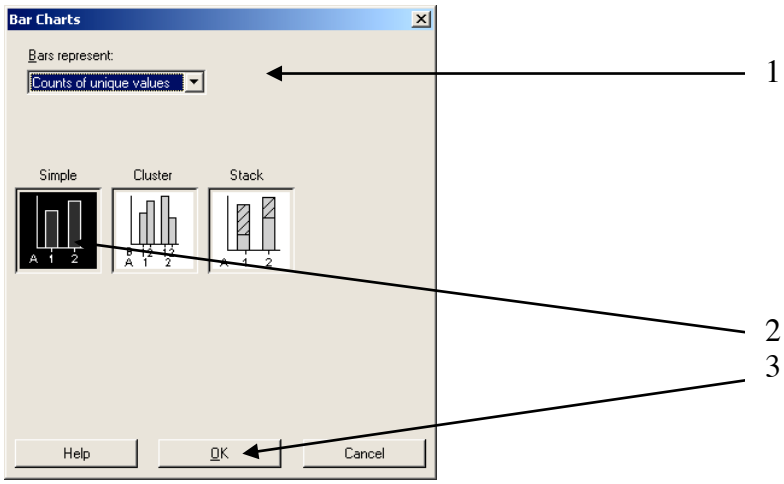
As you can see there is a lot of choice!!

10.1.2 Plots – Bar Graph

Obtain a Bar Graph of the data directly from the full data set by...

- Access Graph -> Bar Chart

1. Choose Bars represent Counts of Unique Values
2. Choose Simple Chart. (or for Side-by-Side bar chart for two or more columns of data then choose Cluster or Stack)
3. Click OK.



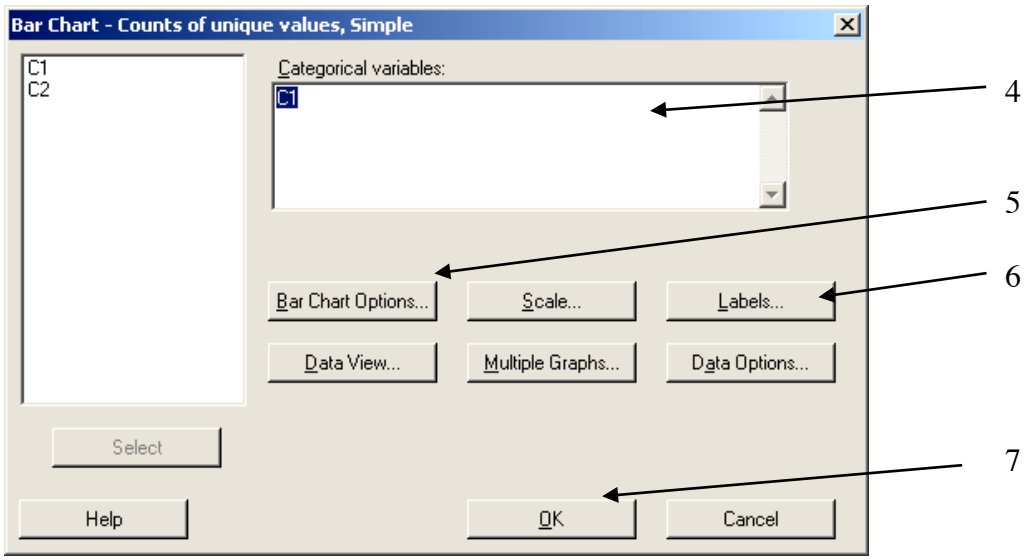
Then in the next dialog box (shown overleaf):

4. Select the column (or columns) for which you want the plot (Categorical variables).

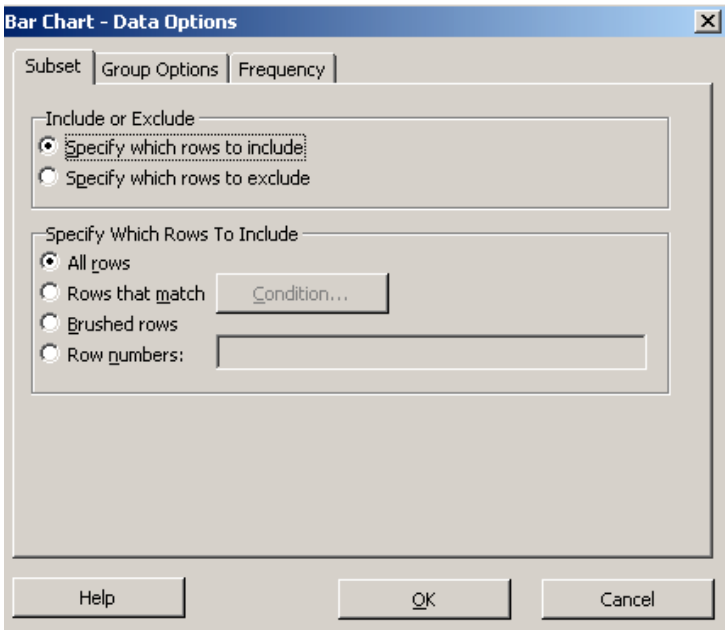
5. If you want to change on Y axis to **Percentages** from **Counts** then Click on **Bar Chart Options** the Percent and Accumulate then Show Y as Percent, If **Cluster or Stack Graph** then you should use Percentages on Y axis so also choose under Take Percent and/or Accumulate and the Within Categories at level 1 option. Click OK.

6. To label the plot then Click on Labels and enter Titles etc.

7. Click OK.



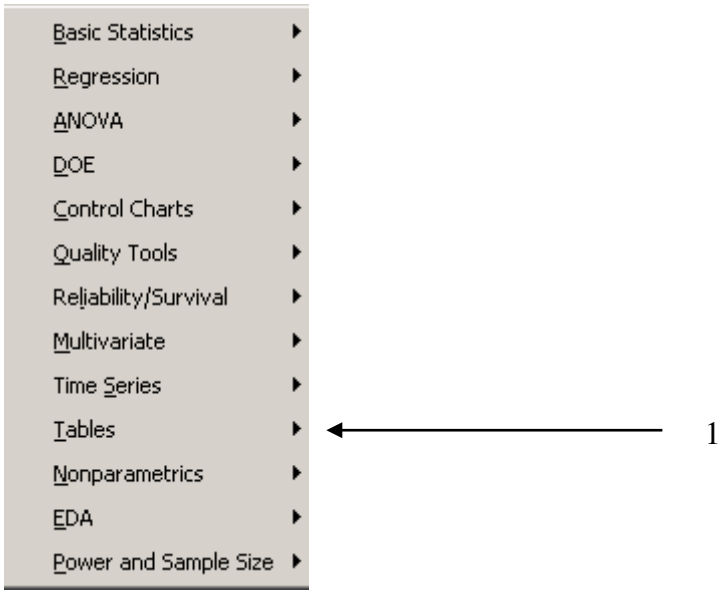
To **Select a Subset of Data**, choose Data Options in the Pie Chart dialogue box, the following options will be displayed:



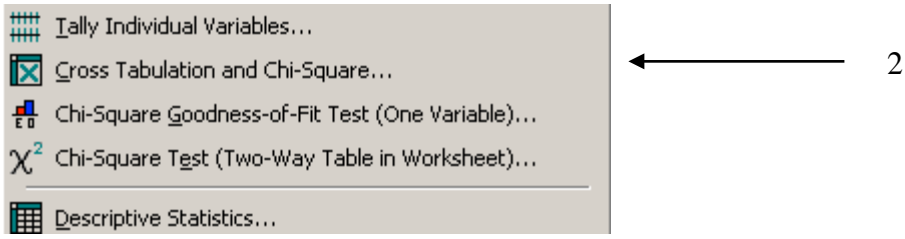
As you can see there is a lot of choice!!

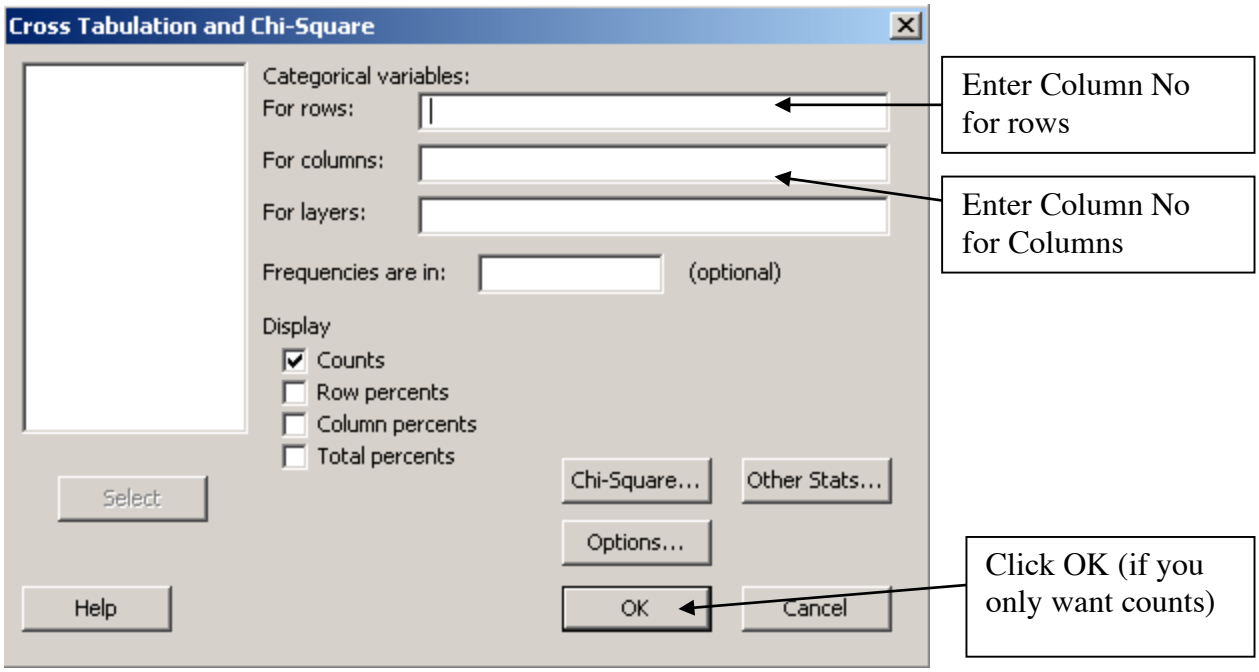
10.2 Numerical Summaries for Categorical Data

When you choose the **Stat Menu**, the following list of options appears. The following options are given, choose **Tables**



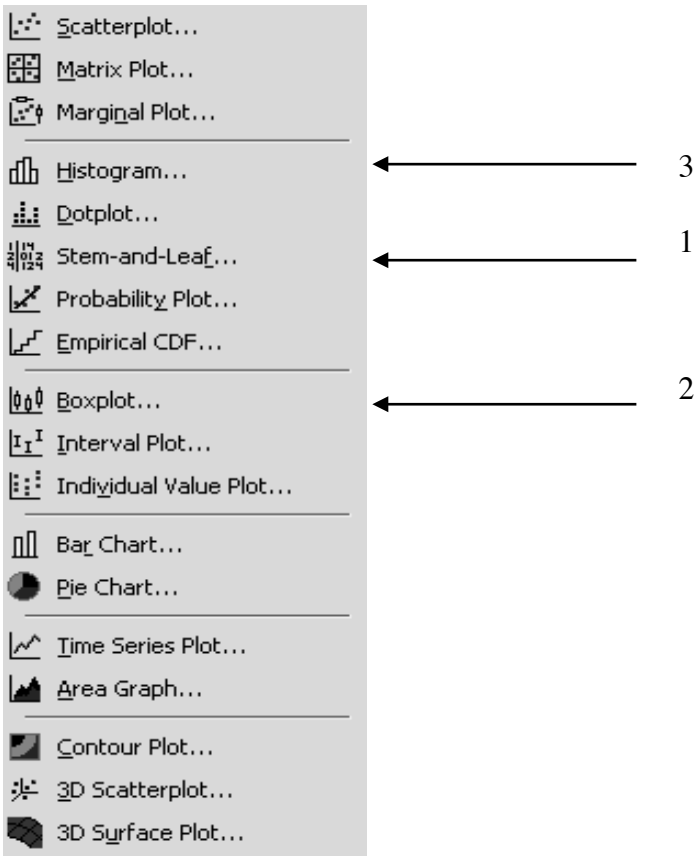
The following options are given under **Tables,** choose **Cross Tabulation and Chisquare**





11. Graphical Summaries for Quantitative Data

When you choose the Graph Menu the following list of options appears. The options we will use during this workshop are marked:



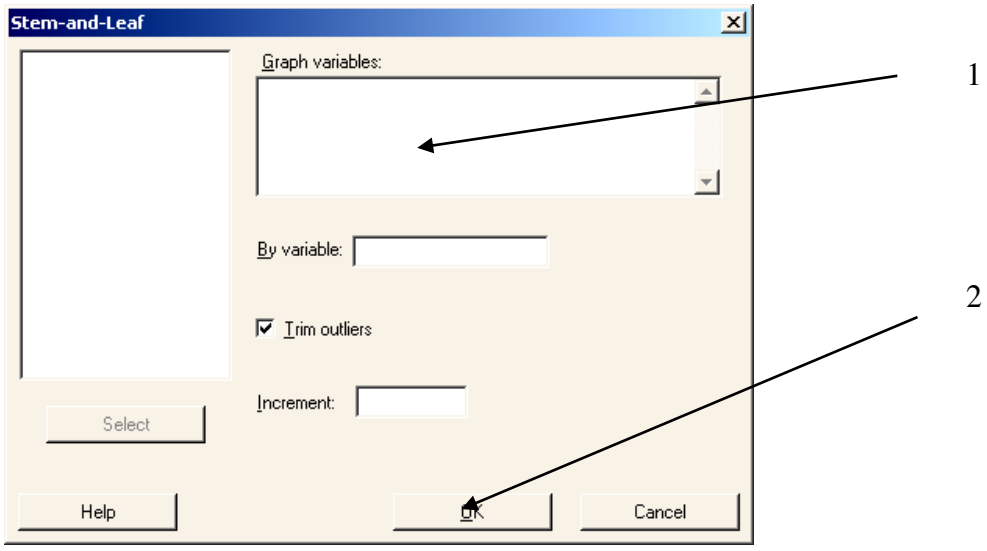
11.1 Plots - Stem and Leaf for 1 Group of Data

Obtain a stem and leaf plot of the data by...

Access Graph -> Stem and Leaf

Then in the dialog box:

1. Select the columns for which you want the plot.
2. Click OK

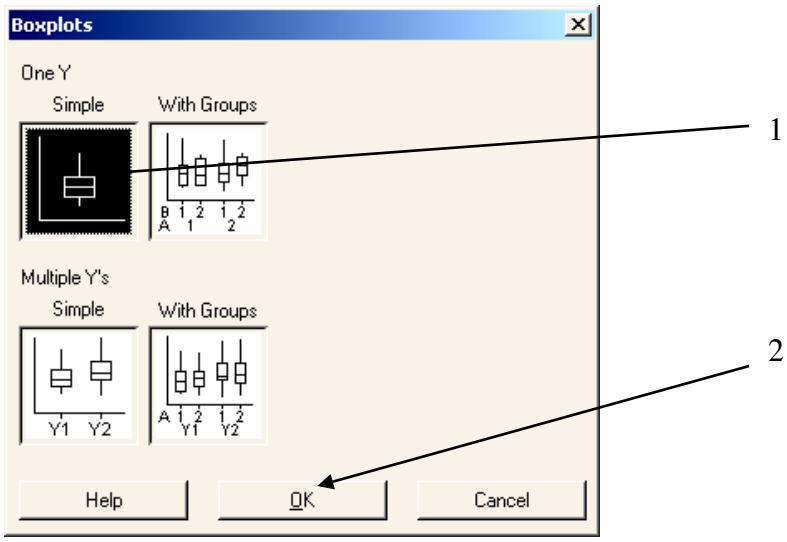


11.2 Plots - Box and Whisker Plot for 1 Group of Data

Obtain a simple box and whisker plot for one set of data by....

Access Graph -> Boxplot ->

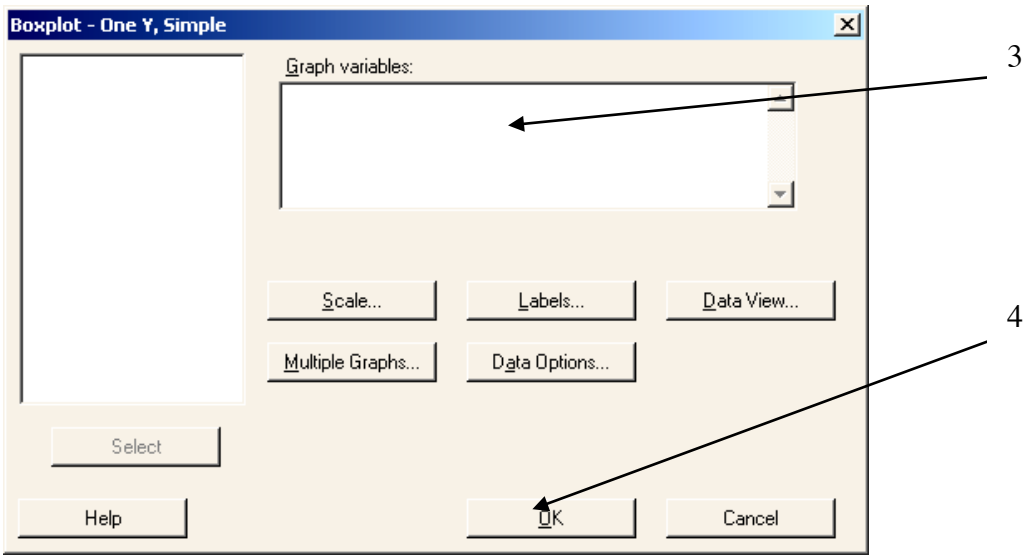
1. Choose the type you want. i.e. in this case click One Y - Simple.
2. Click OK



Then in the dialogue box :
(If you want to label the plot then Click on Labels before leaving the dialogue box.)

3. Select the columns for which you want the plot.

4. Click OK



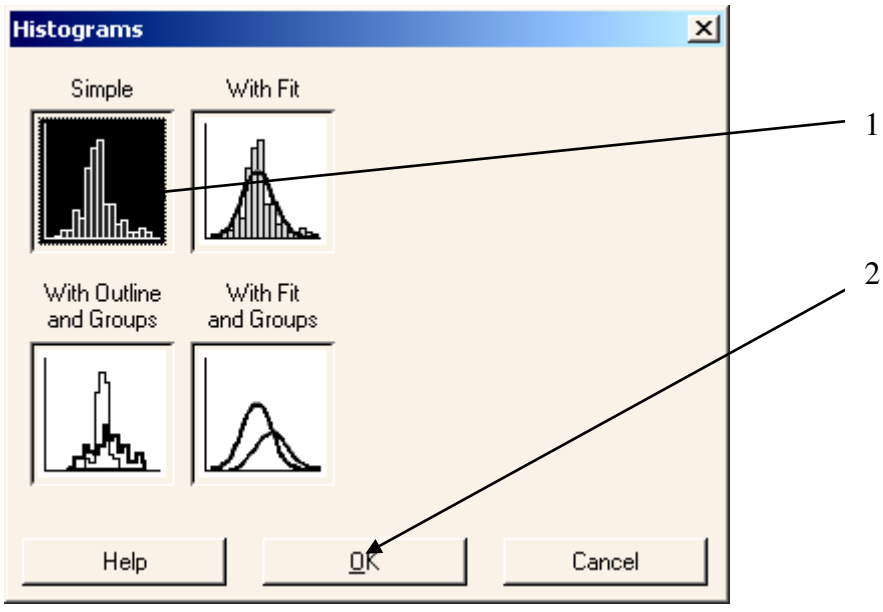
11.3 Plots – Histogram for 1 Group of Data

Obtain a histogram of the data in any column by..

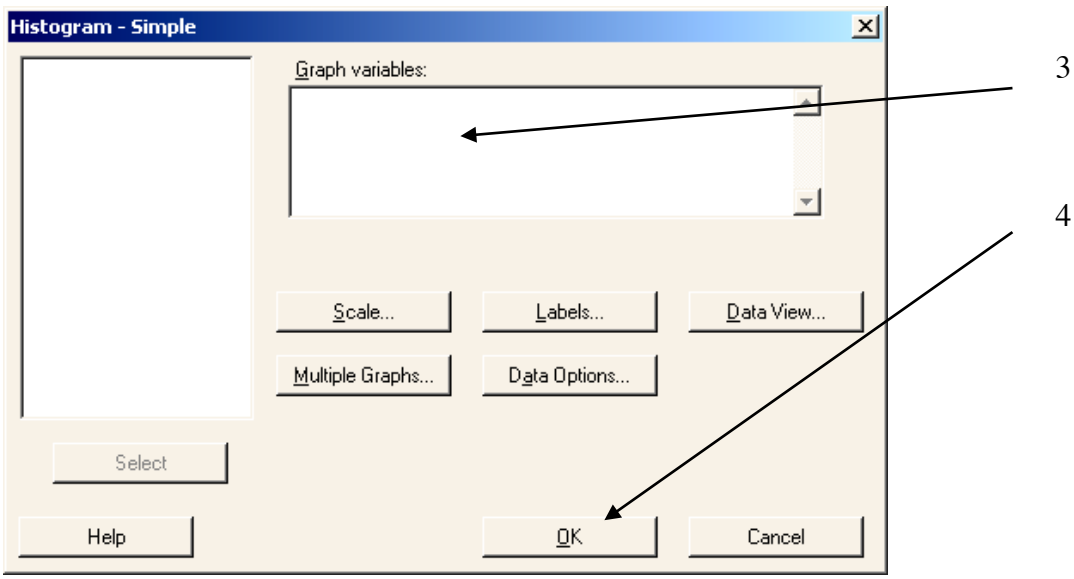
Access Graph -> Histogram

In the dialogue box,

1. Select the type of histogram you want. For example, Simple.
2. Click OK.



3. Select the data you want to plot then any options you want!!
4. Then Click OK



11.4 Plots for More Than One Set of Data

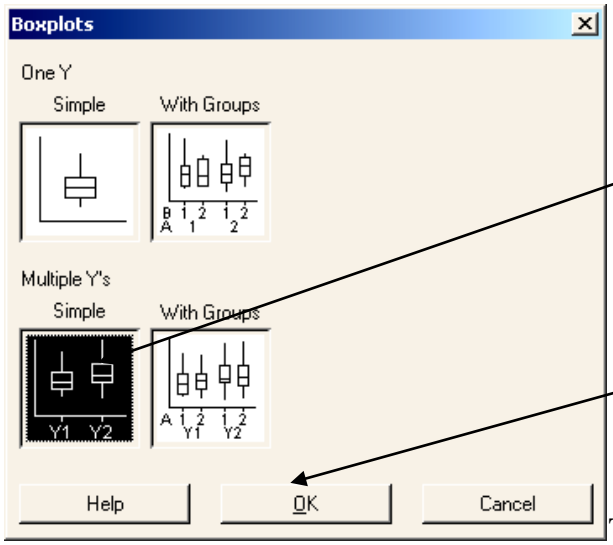
If we want to plot more than one set so as to make direct comparisons then….

11.4.1 Box and Whisker Plots

11.4.1.1 Data in Separate Columns

Access Graph -> Boxplot ->

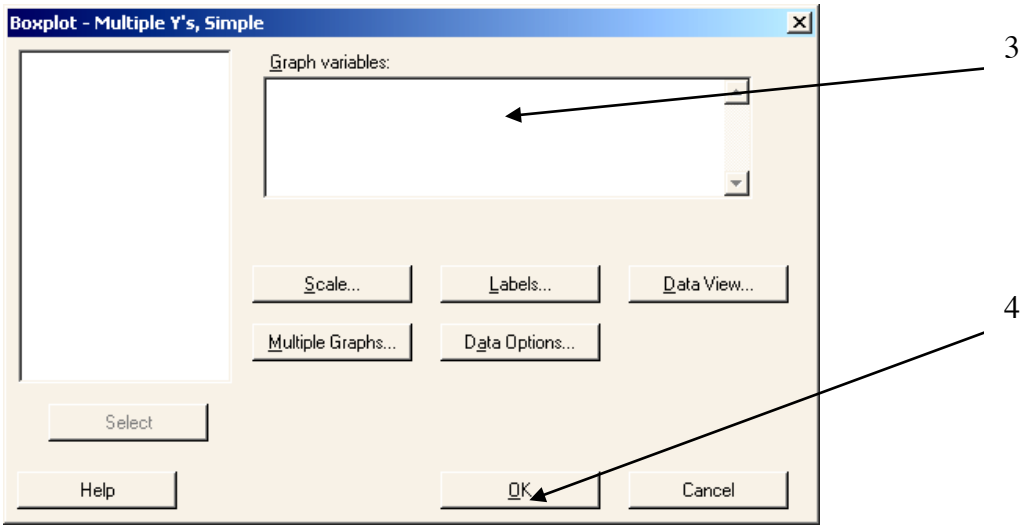
1. Choose the type you want. i.e. in this case click Multiple Y's Simple (if your two or more sets of data are in separate columns.
2. Click OK



Then in the dialogue box :

3. Select the columns for which you want the plot.

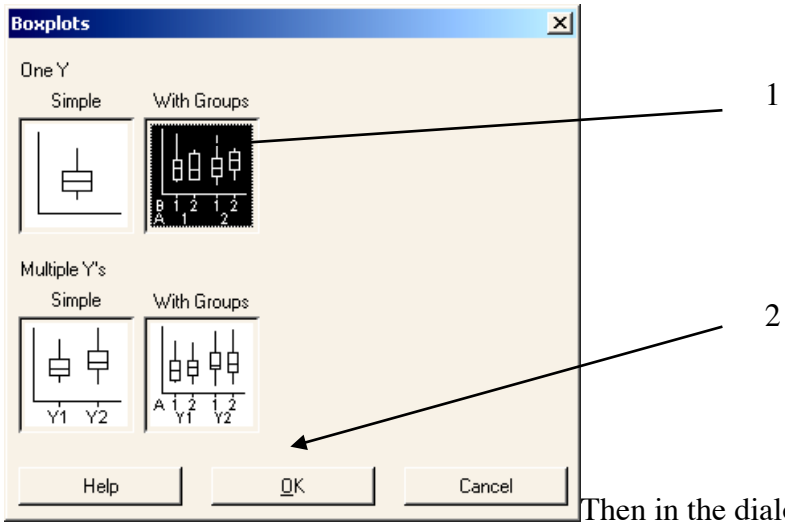
4. Click OK



11.4.1.2 Data in Single Column

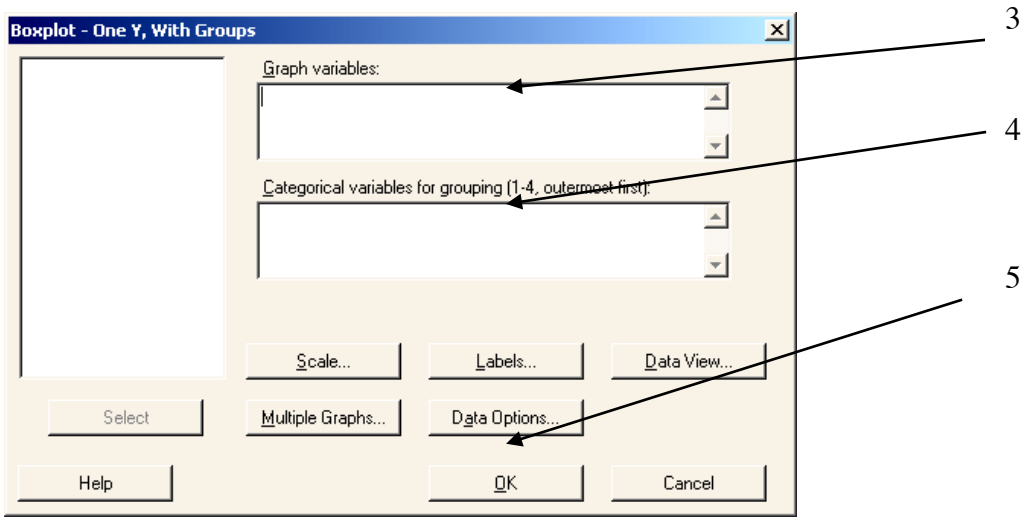
Access Graph -> Boxplot ->

1. Choose the type you want. i.e. in this case click One Y – With Groups (if your two or more sets of data are in a single column with group labels in another column..
2. Click OK



Then in the dialogue box :

3. Select the column containing the data as the Graph variables
4. Select the column containing the group codes as the Categorical variables for grouping
5. Click OK



11.4.2 Histograms

You can plot more than one histogram on a single graph but this is often not informative as they will be superimposed on each other.

So alternatively you can draw more than one histogram but choose the same scale for the graphs. You can then plot them on the same 'page'!

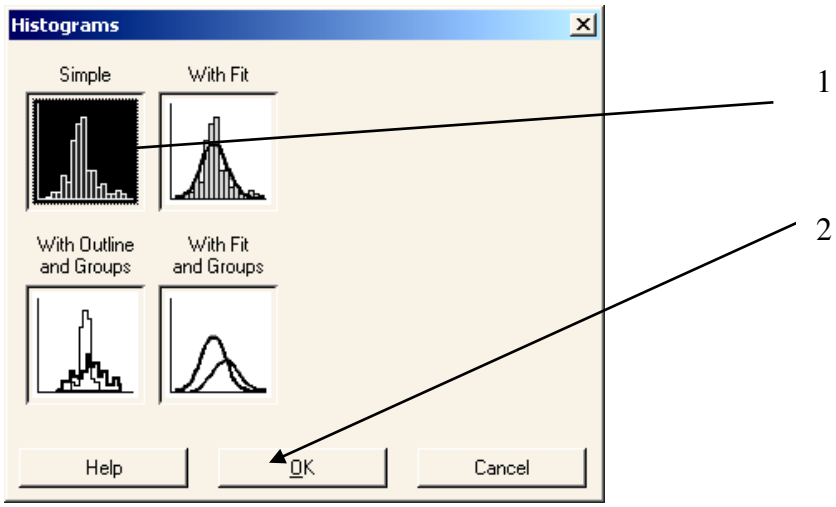
Note also that if comparing groups of different sizes examining percentages may be more appropriate. See section 2.4 Relative Frequency Histograms.

11.4.2.1 Histogram – Data in Separate Columns

Access Graph -> Histogram

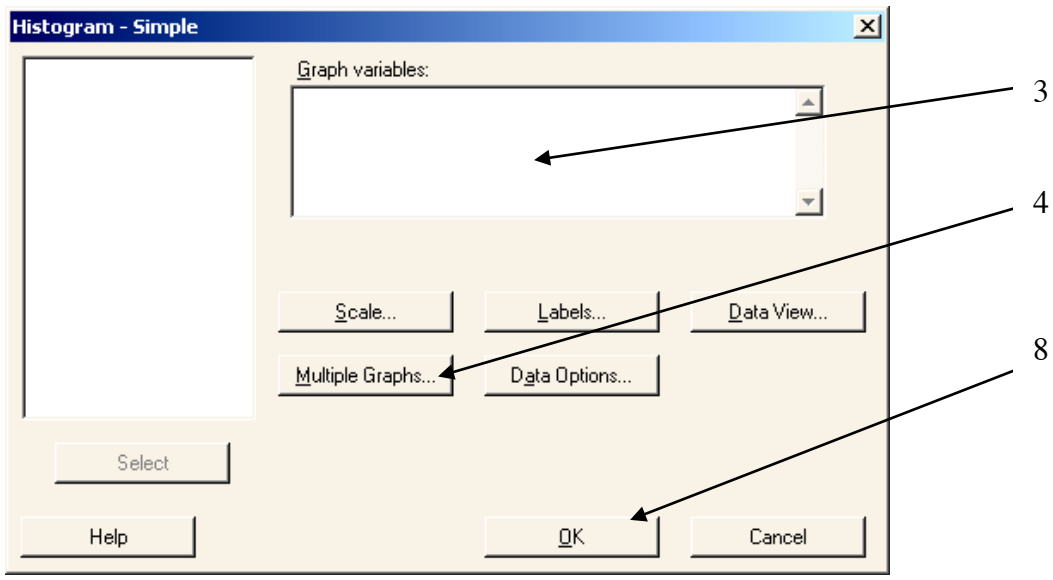
In the dialogue box,

1. Select the type of histogram you want - Simple.
2. Click OK.

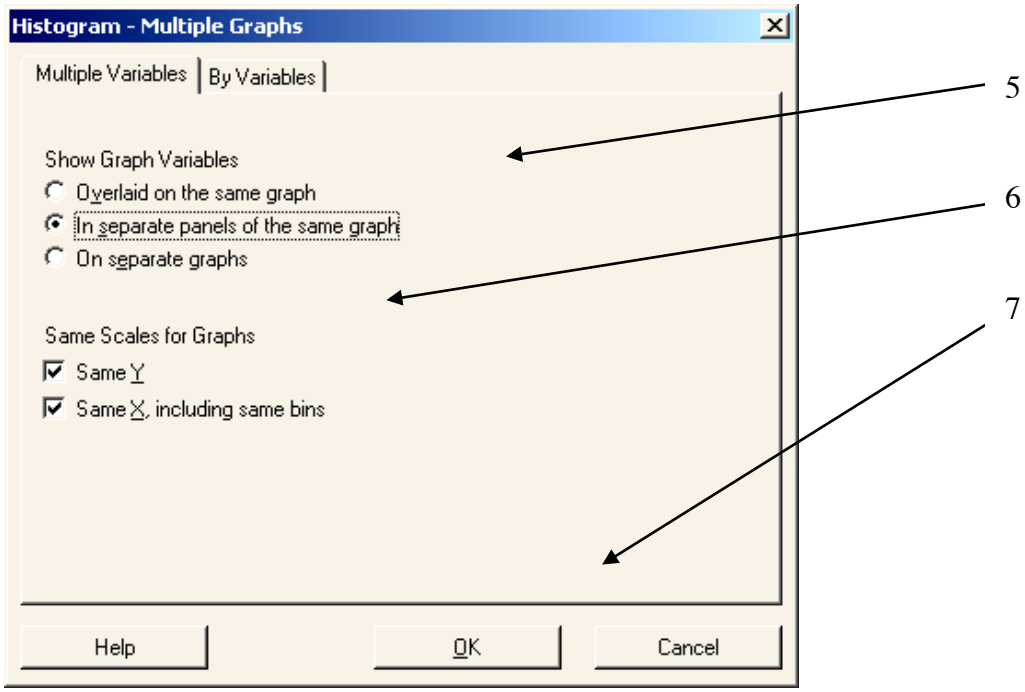


3. Select the data you want to plot then any options you want!!

4. Click Multiple Graphs



5. Choose 'In separate panels of the same graph'
6. Choose Same Y and Same X
7. Click OK
8. Click OK



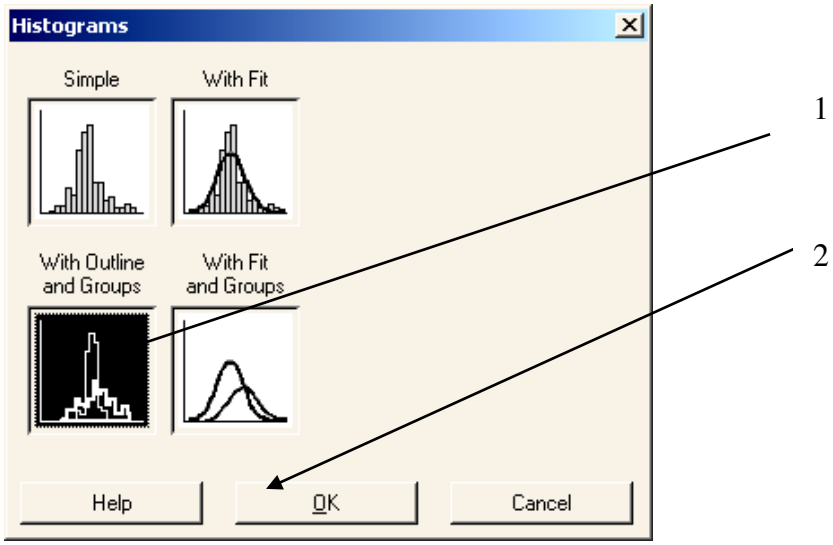
11.4.2.2 Histogram – Data in a Single Column

Access Graph -> Histogram

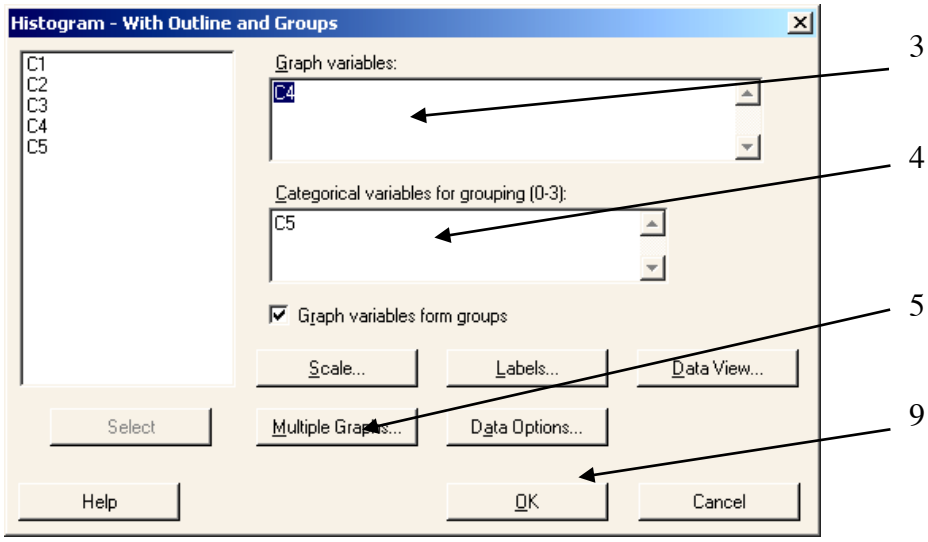
In the dialogue box,

1. Select the type of histogram you want – With Outline and Groups

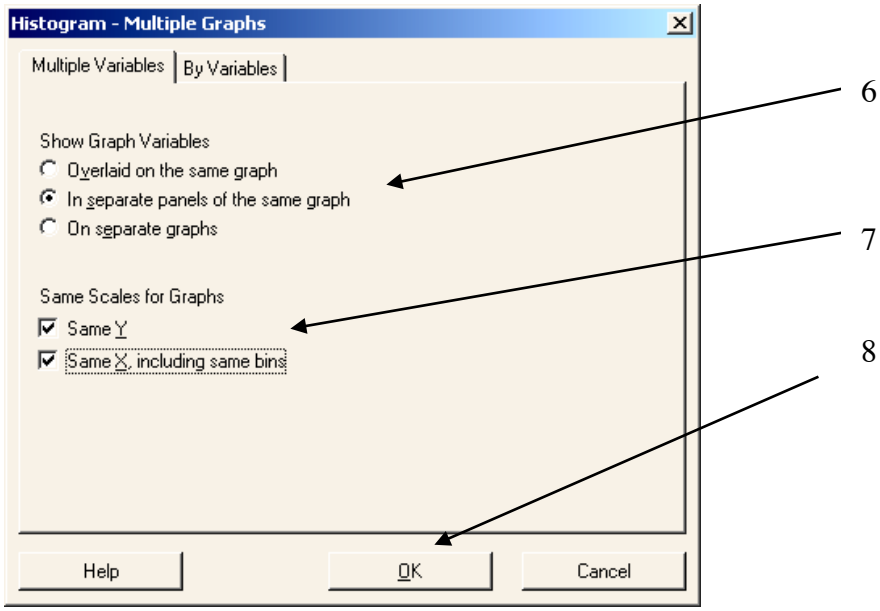
2. Click OK.



3. Select the column containing the data as the Graph variables
4. Select the column containing the group codes as the Categorical variables for grouping
5. Click Multiple Graphs



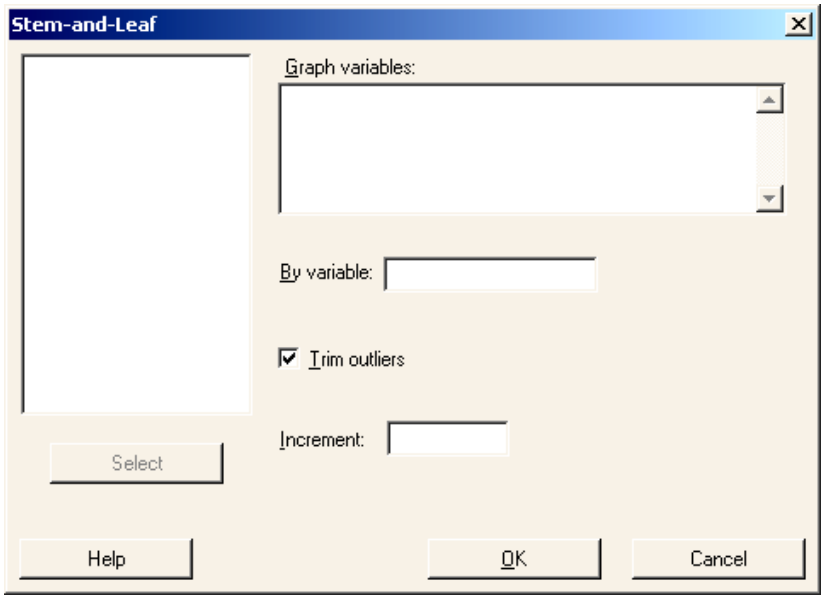
6. Choose ‘In separate panels of the same graph’
7. Choose Same Y and Same X
8. Click OK
9. Click OK



11.5 Change the Scale on Plots

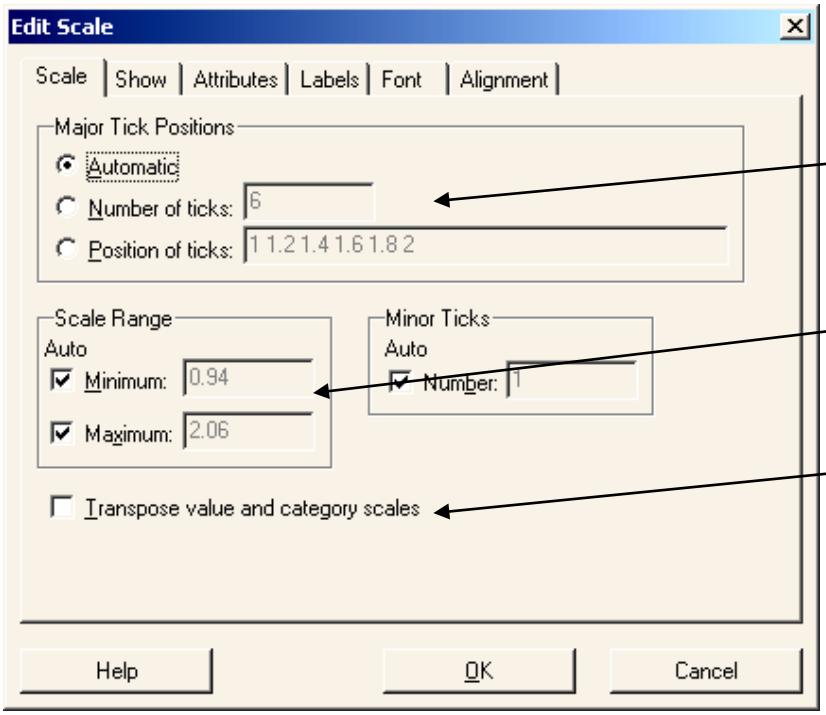
11.5.1 Stem and Leaf Plot

You can change the scale by changing the increment in the following dialogue box which appears after choosing Stem and Leaf Plot (see Workshop 1):



11.5.2 Box and Whisker Plot

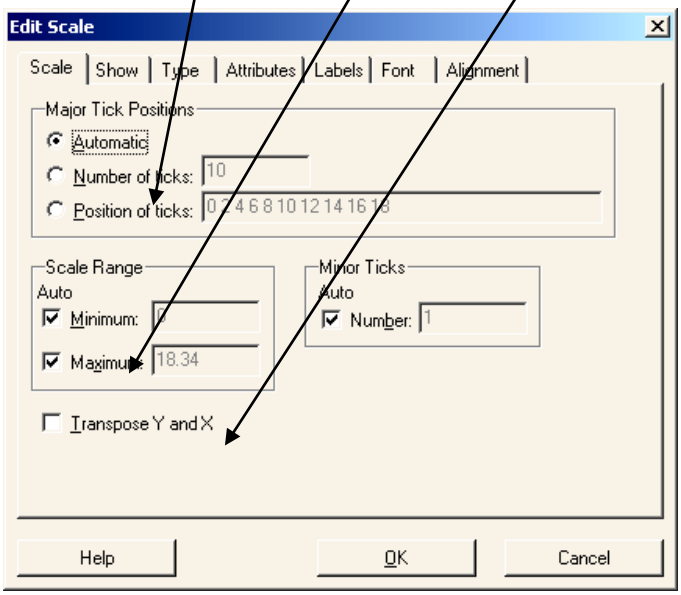
After you have created the plot- click on the Y-axis and the following dialogue box appears:



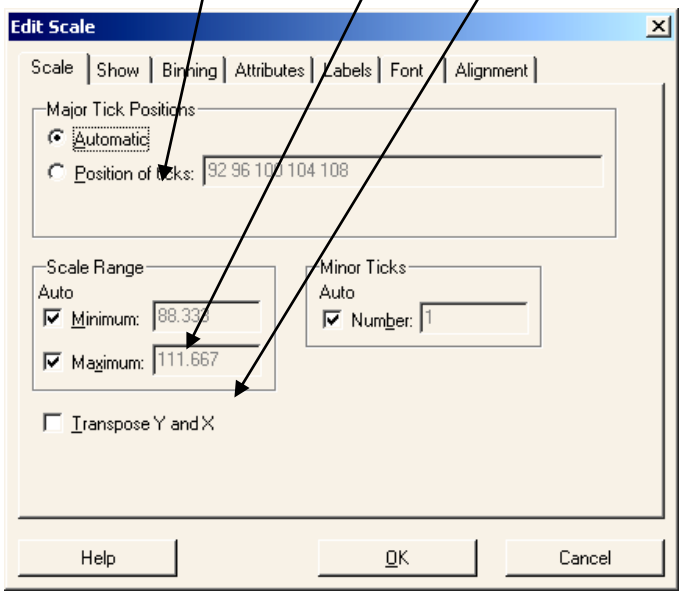
11.5.3 Histogram

After you have created the plot- click on the axes and the following dialogue boxes appear:

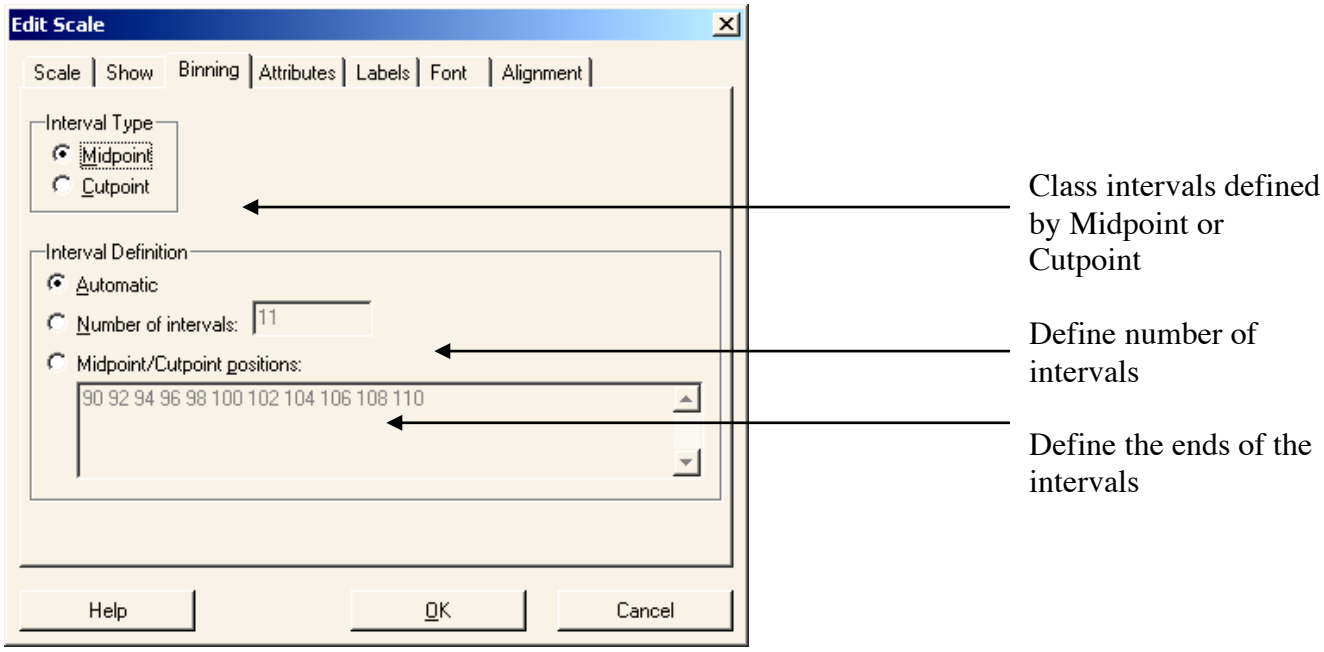
Click on Y Axis
Change intervals, min and max, direction



Click on X Axis
Change intervals, min and max, direction



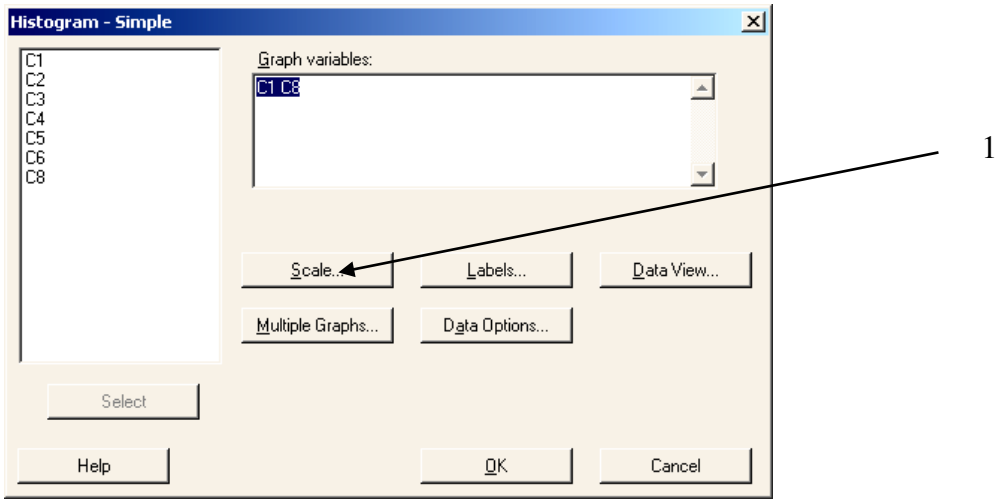
Click on Binning to change the 'Class Intervals'



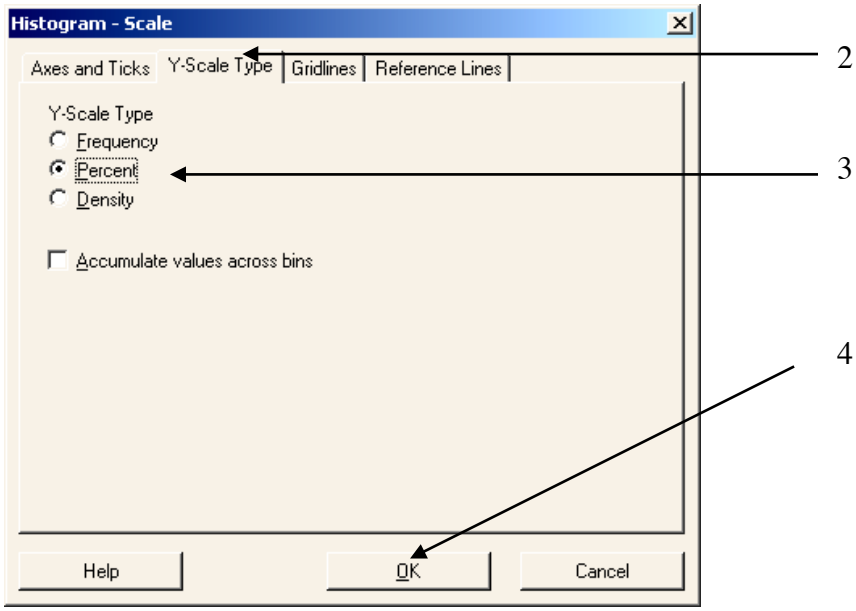
11.5.4 Relative Frequency Histogram.

If comparing two of more groups of different sizes then may be more appropriate to use percentages instead of counts when drawing the histogram.

1. Choose Scale from the dialogue box which appears having selected the type of histogram (example below)



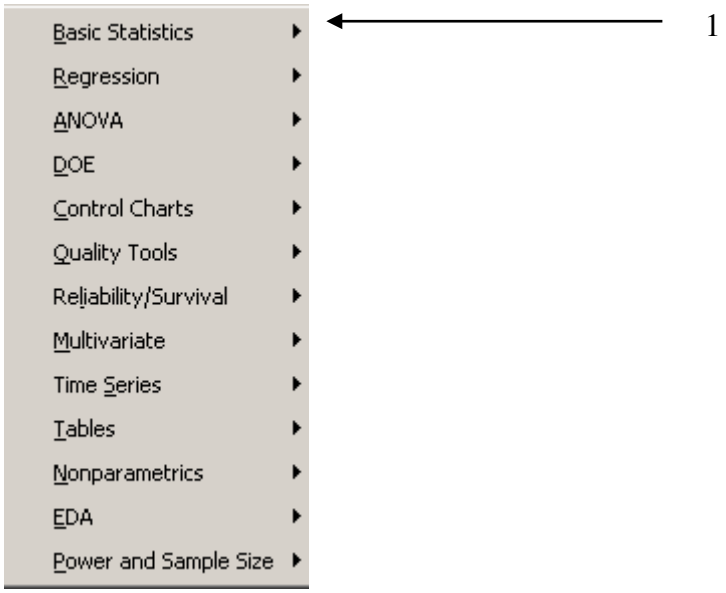
2. Choose Y-Scale Type from the dialogue box that then appears by clicking on tab
3. Select Percent
4. Click OK



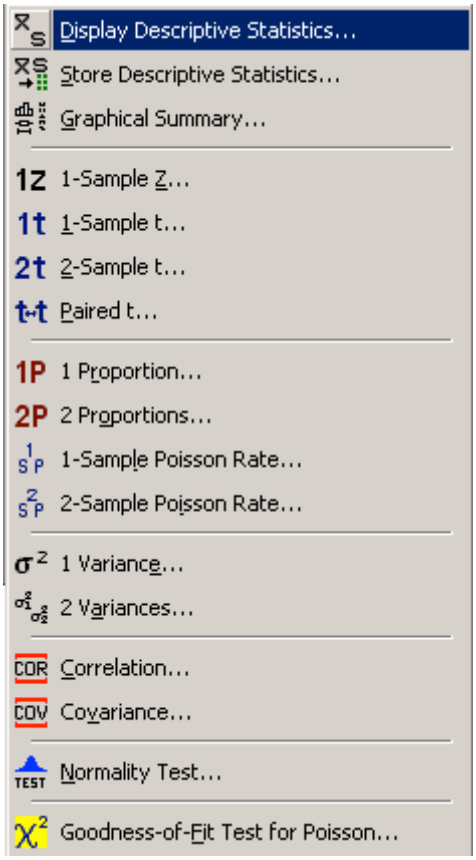
12 Numerical Summaries

i.e. Find the mean, median, trimmed mean, standard deviation, standard error, range and lower and upper quartiles for data in any number of columns.

When you choose the Stat Menu, the following list of options appears. The options we will use during this workshop are marked:



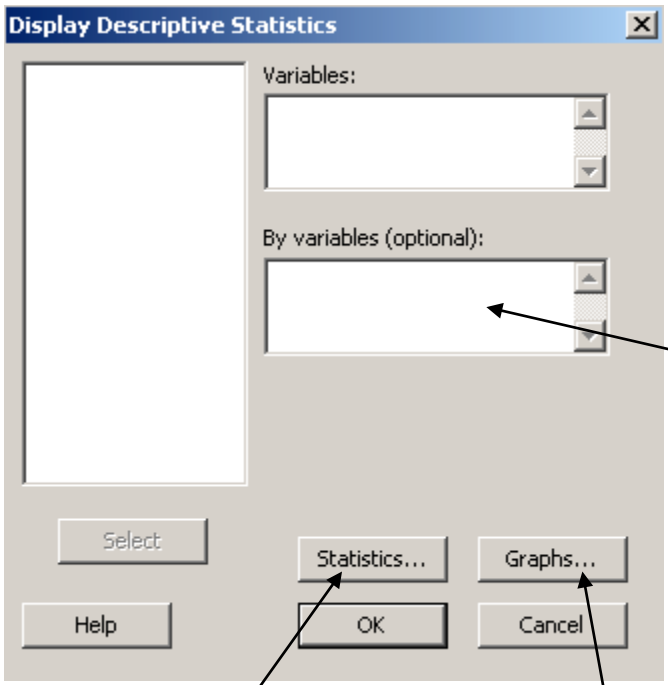
The following menu choice is given, choose Display Descriptive Statistics



2

Try this too!!!

In the dialog box, choose the columns (in the Variables box) for which you want summary statistics then click on Select, then Click on OK.



Separate summaries for different groups obtained by selecting the label column as the 'By Variable'

Choose what you want!!

You can also get a Boxplot of your data by using the Graphs option before leaving the dialogue box.

13. **Leaving MINITAB**

- Access the File Menu -> Exit (see Section 5.2)

14. **Finishing the Session**

- Click on **Start** (See Section 2)
- Click on **Logoff**