

Descriptive Statistics with SPSS

Download the data file “**week1_data.csv**” from the moodle. It contains data for a sample of 30 employees, 15 male and 15 female, of a large organisation, between the ages of 20 and 50. Open the data file to view the variables and cases (observations).

Define the variables as folloes:

Refno is a 3-digit reference code, unique for each employee.

Gender is categorical Nominal variables i.e. male or female are names of unordered categories and are coded:

Gender: 1 = male, 2 = female

Agegroup and **Jobhappy** (Label: Are you happy in your job?) are categorical Ordinal variables i.e. its values are categories which can be ordered:

Agegroup: 1 = 20 - 29 years old, 2 = 30 - 39 years old, 3 = 40 - 50 years old.

Jobhappy: 1 = no, 2 = not sure, 3 = yes

Age, **Salary** (Label: Present salary (£000)) and **Prodbsal** (Label: Productivity-based salary (£000)) are quantitative Scale variables i.e. their values are measurements on a proper scale.

Salary is the actual salary in £1000's

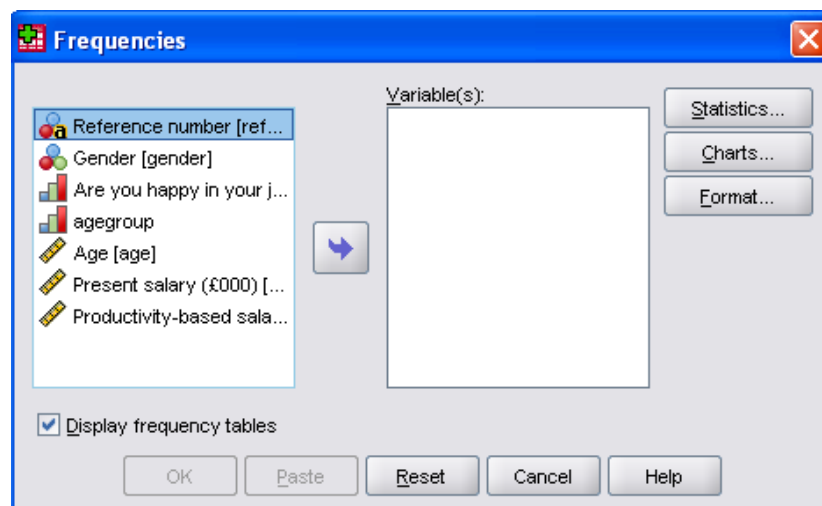
Prodbsal is the proposed salary in £1000's under a new productivity-based incentive scheme.

We will now consider a selection of ways of analysing the data and displaying the results. This first section does not involve any significance testing.

1) Frequency table and Percentages

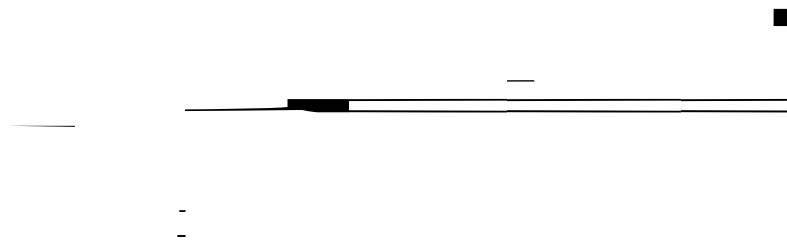
a) Obtain a frequency table for the agegroup variable.

Click on **Analyse**, **Descriptive Statistics** and **Frequencies** to give:



Click on **Agegroup** and the right-arrow .
Click on **Charts**, **Bar charts** and **Continue**, **OK**.
(Obtaining a bar-chart is normally optional)

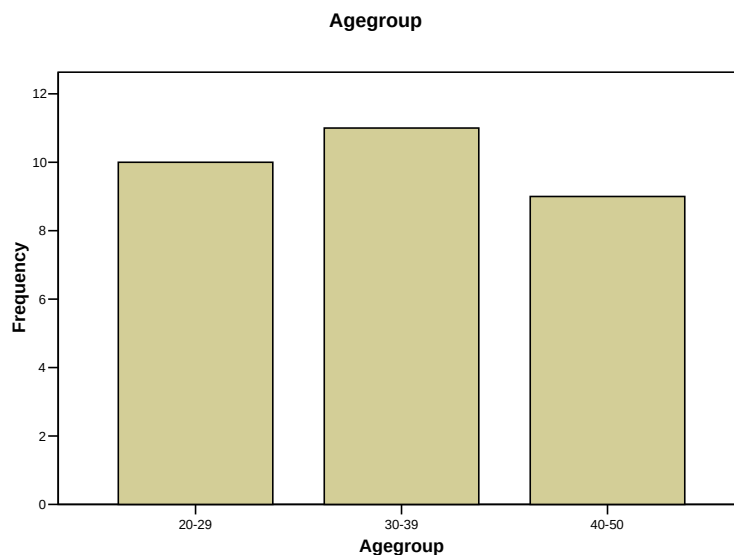
The following output is displayed in the output window:



- Note:
- i) The output shows value labels 20-29 etc. (if any).
 - ii) The valid percentage is the same as the percentage in this case. However if there was a blank cell i.e. a value missing in the data, the valid percentages would be calculated out of 29 not 30.
 - iii) The cumulative percentages show the percentage up to and including that value e.g. there are 70% up to 39 years old.

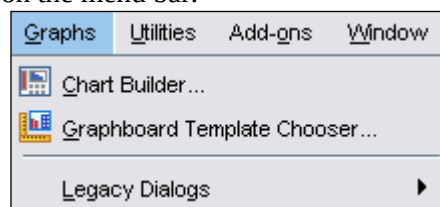
b) Bar-chart Option

The chart should appear as:



2) Charts (without obtaining frequencies)

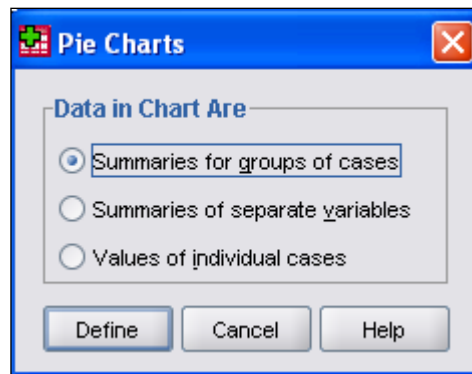
In SPSS v 17 there are a number of different ways to generate graphs which can be seen by clicking on the menu bar.



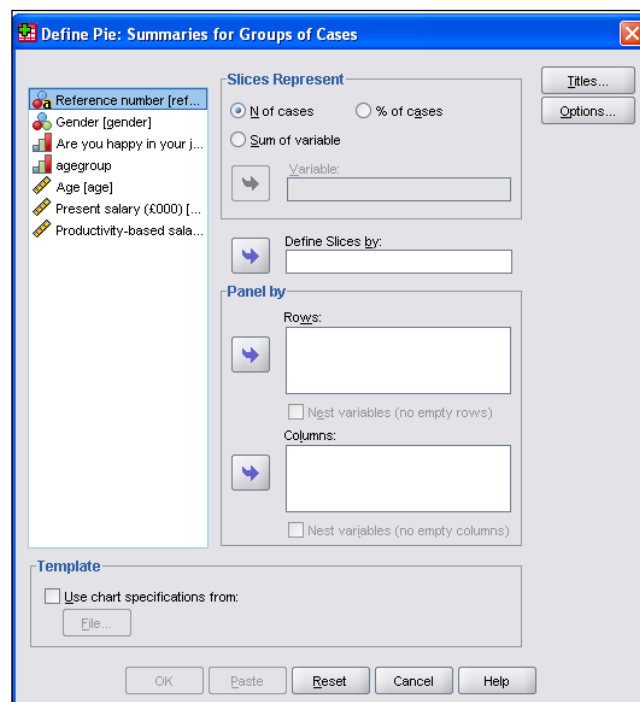
The option **Legacy Dialogs** is consistent with procedures available in earlier versions of SPSS, but more experienced users may choose to experiment with the alternative approaches to producing more interactive graphs.

a) **Obtain a pie-chart for the jobhappy variable**

Click on **Graphs**, **Legacy Dialogs** and **Pie**.



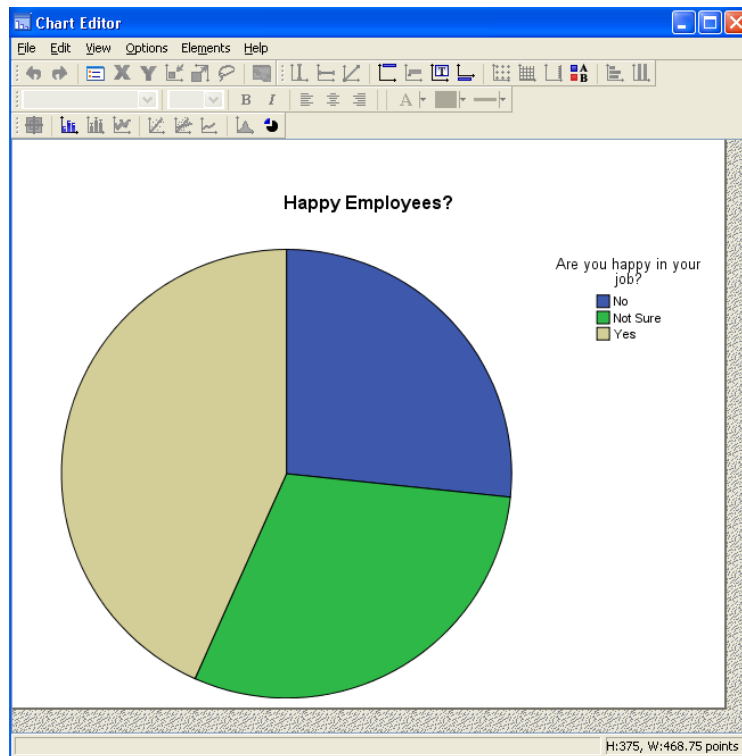
Select **Summaries for groups of cases** and click on **Define** to give:



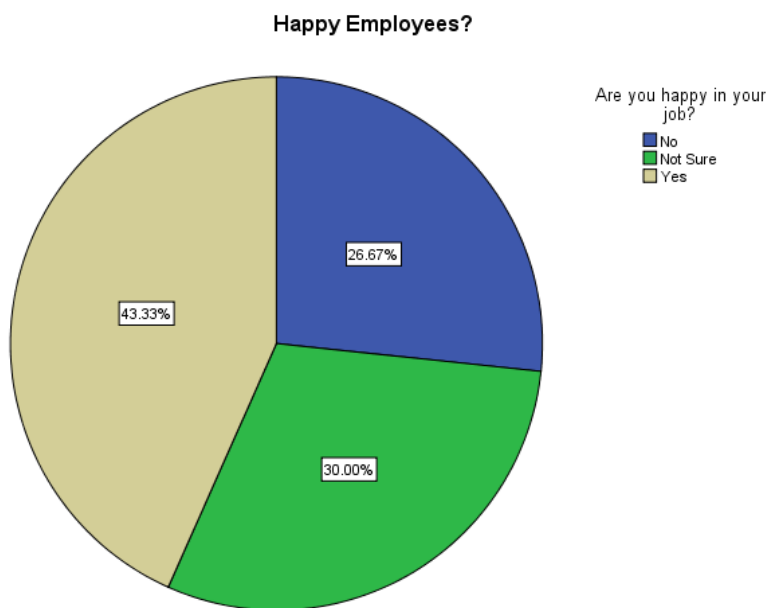
Click on **Jobhappy** and on the “Define Slices by” arrow  ..
N (umber) of cases should already be selected for Slices Represent.
Click on **Titles**, type in **Happy Employees?**, and click on **Continue, OK**.

Editing the Chart (Add labels)

Click and then double-click on the **Chart** and the SPSS Chart Editor window should appear.



Click on the pie chart and it should then be highlighted with a blue border. Click on **Elements, Show Data Labels** and the Properties window automatically appears for further editing. At this stage it is possible to change the labels, colours, titles or chart type. Close the Properties window to show the following chart.



Changing the chart type

Double click on the pie chart in the Chart Editor to highlight again with a blue border. In the Properties window, click on the **Variables** tab and change the Element type from **Pie** to **Bar** from the drop-down menu, **Apply**, **Close**.



Editing the Bar Chart

Change bar colour to white:

Double click on the bars and the Properties window should appear. Alternatively the Properties window can be opened with a right click over the selected bars.

Click on the **Fill & Border** tab, select each bar in turn, click on **Fill** and change the colour to **white** by clicking on the white rectangle, **Apply**.

Change bar style to drop shadow

Click on the **Depth & Angle** tab and click on **Shadow** under Effect, **Apply**, **Close**.

Remove the Legend

Double click on the legend and press delete on the keyboard

Change position of bar labels

Double click on the bar labels in the centre of the bars.

In the Properties window click on the **Data Value Labels** tab.

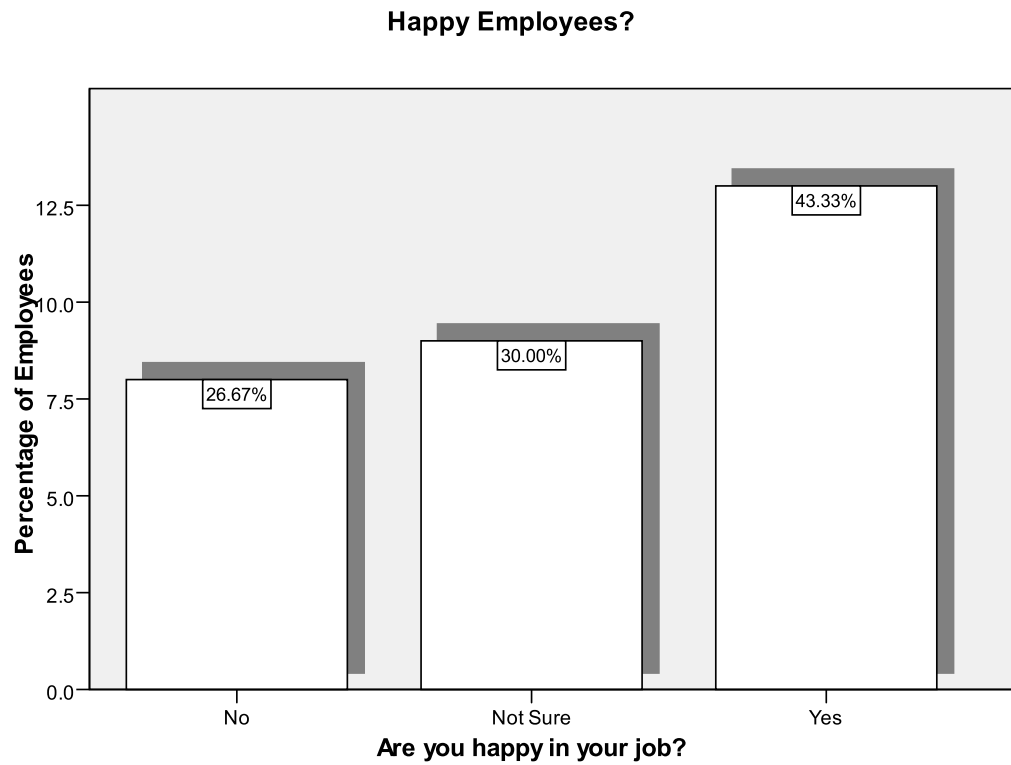
Click on **Custom** under Label Position and click on the **Above Center** icon.

Apply, **Close**.

Edit the Axis Label

Click on the vertical axis label and then click on it again to change it to **Percentage of Employees**.

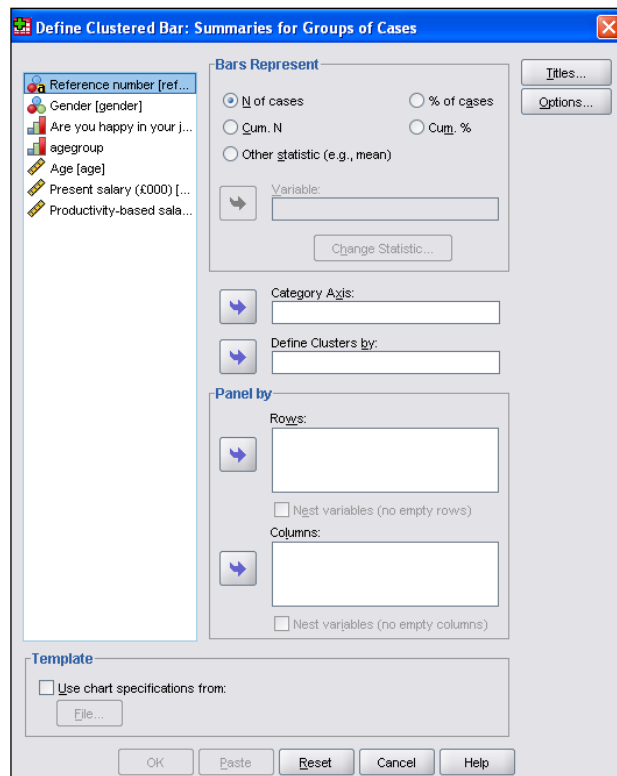
The edited chart should now appear as:



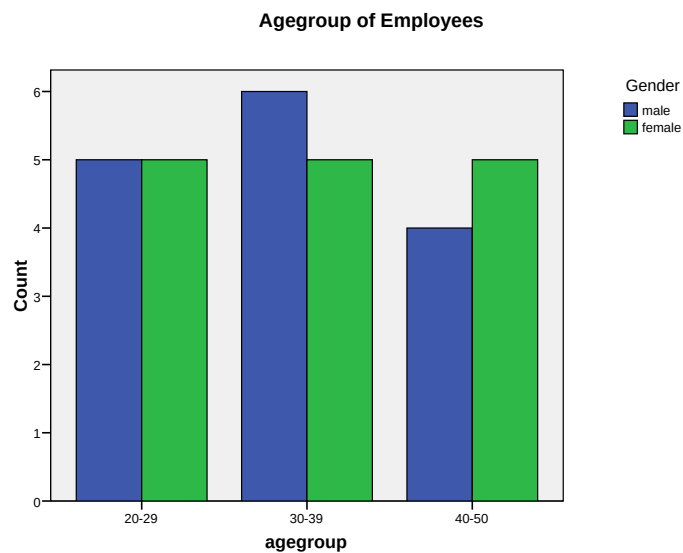
Close the Chart Editor window when editing has been completed

b) Cluster Bar Chart for gender and agegroup

Select **Graphs, Legacy Dialogs, Bar, Clustered, Define** to give:



Select **Agegroup** for the Category Axis, **Gender** for Define Clusters by. Add a Title, Continue, **OK**. The following chart appears, displaying a bar for each combination of Agegroup and Gender:



The frequencies and percentages represented by each bar can be displayed using cross-tabulations as shown in the next examples.

using cross-

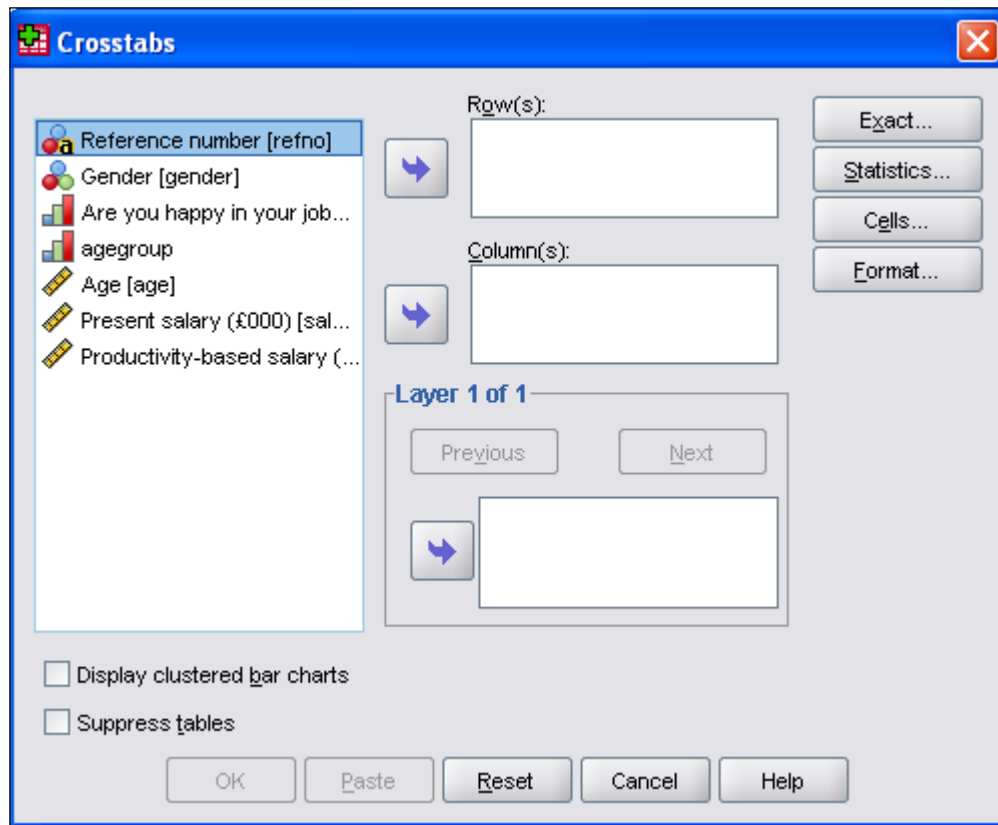
3) Cross-tabulations

a) Cross-tabulation of gender and agegroup

Produce a cross-tabulation between gender and agegroup to show the frequencies in each possible combination of the gender and agegroup values.

Click on **Analyse, Descriptive Statistics** and **Crosstabs**.

Select **Gender** for the row variable and **Agegroup** for the column variable, **OK**.



The output includes:

Gender * Agegroup Crosstabulation

Count		Agegroup			Total
		20-29	30-39	40-50	
Gender	male	5	6	4	15
	female	5	5	5	15
Total		10	11	9	30



Now click on the **Go to Data Editor** icon to return to the data.

b) Cross-tabulation of Gender and Jobhappy

Row, column or total percentages can be obtained together with frequencies.

Click on **Analyse, Descriptive Statistics** and **Crosstabs**.

Select **Gender** for the Row variable and **Jobhappy** for the Column variable.

Click on **Cells** and select **Row** under Percentages, **Continue** and **OK**, to give the following:

Gender * Are you happy in your job? Crosstabulation

			Are you happy in your job?			Total
			No	Not Sure	Yes	
Gender	male	Count	7	5	3	15
		% within Gender	46.7%	33.3%	20.0%	100.0%
	female	Count	1	4	10	15
		% within Gender	6.7%	26.7%	66.7%	100.0%
Total		Count	8	9	13	30
		% within Gender	26.7%	30.0%	43.3%	100.0%

This shows that 10 out of 15 or 66.7% of females are happy with their jobs, whereas 3 out of 15 or 20% of males are happy with their jobs.

4) Histograms

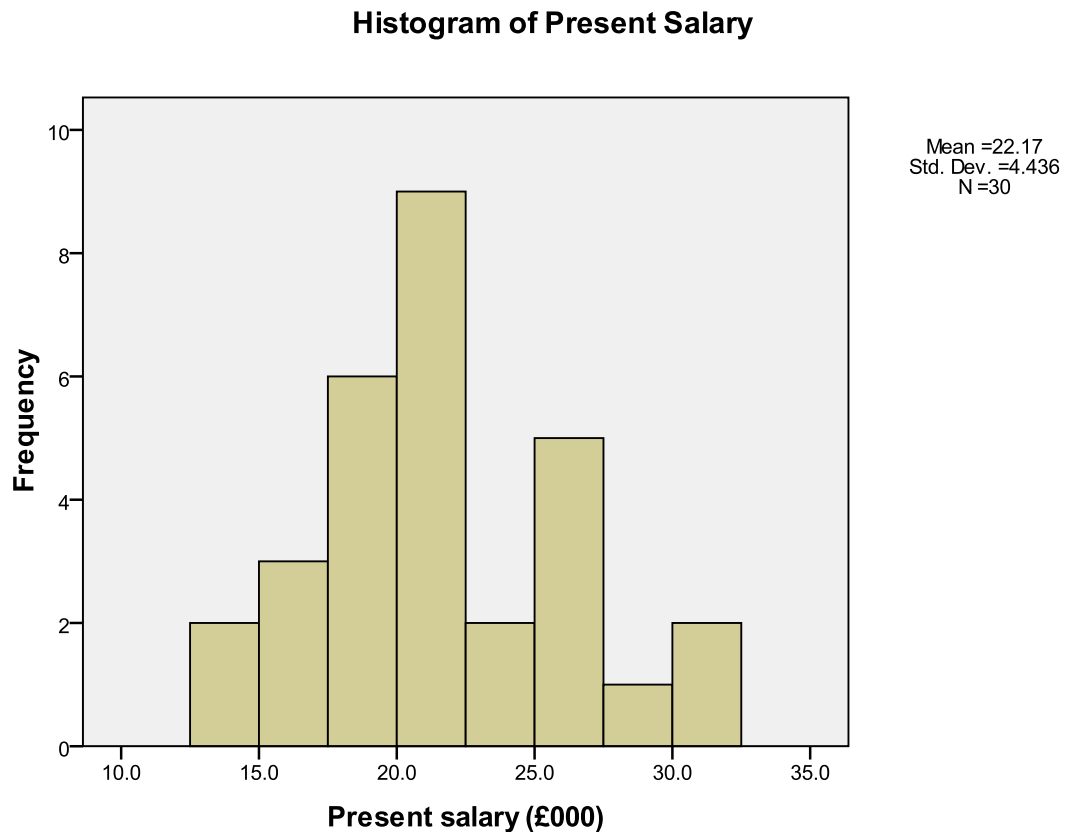
A histogram is used with quantitative variables to show the frequencies within a number of class intervals.

Histogram of salary

Click on **Graphs, Legacy Dialogs, Histogram**, select **Salary** for the variable.

Click on **Titles** and type **Histogram of Present Salary**

Click on **Continue, OK**, to give:



SPSS has used 8 intervals in the construction of the histogram.

The mean salary, standard deviation and number of values are also displayed.

Editing the Histogram

Double-click on the Histogram to view the SPSS Chart Editor.

To change the Bars to White

Double click on the bars and click on the Fill & Border tab in the Properties window. Click on **Fill** and then change the colour of the bars to white by clicking on the white rectangle, **Apply**, **Close**.

To improve the labelling

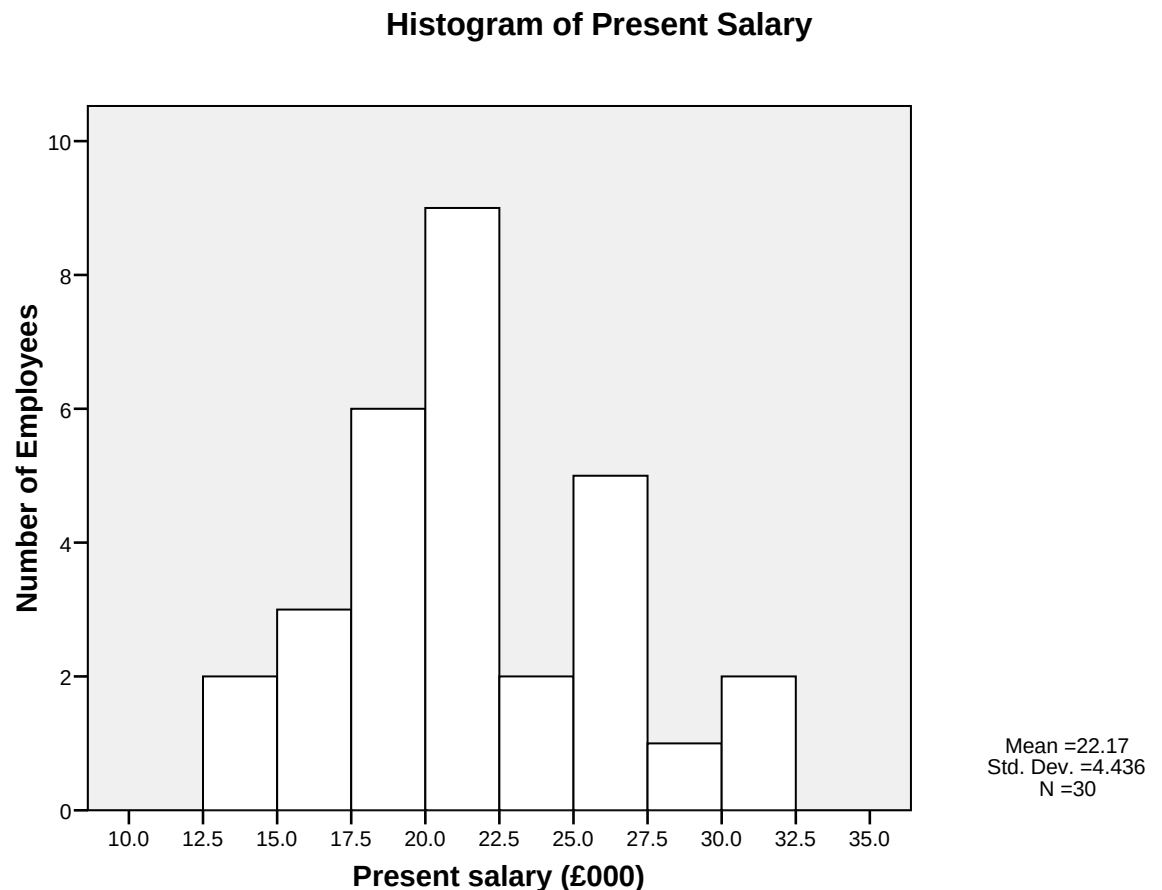
Double click on the Salary Axis to open the Properties window.

Click on the Scale tab.

For Major Increment deselect **Auto** and enter **2.5** under Custom, **Apply**, **Close**.

Click on the Frequency Axis Title, and then click again to edit the title. Change the title to **Number of Employees**.

Close the Chart Editor window.



Note that the boundary values of the intervals are labelled on the histogram.

In some cases SPSS labels the mid-points of the intervals.

This is dependent on the convenience of the data.

5) Averages and Measures of Spread

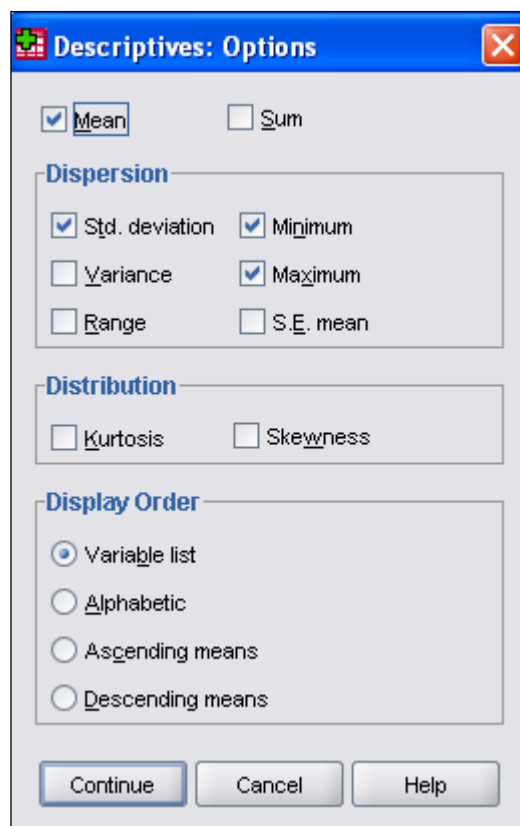
The mean, number of values and standard deviation are automatically displayed with a **Histogram**.

Frequencies (Analyse, Descriptive Statistics, Frequencies) has a Statistics option which will display such things as mean, median, mode, sum, standard deviation, maximum, minimum, range, variance, percentiles, etc.

Descriptives (Analyse, Descriptive Statistics, Descriptives) will give a number of measures of average and spread under Options. Although there are fewer options than with Frequencies, the Frequencies tool is not always appropriate.

Descriptives for salary and prodsal

Return to the data editor then select **Analyse, Descriptive Statistics, Descriptives**. Select both **Present salary** and **Productivity-based salary** for the variables and click on **Options** to give:



Select **Mean**, **Std. deviation**, **Maximum** and **Minimum**, **Continue** and **OK**. The output should be:

Descriptive Statistics

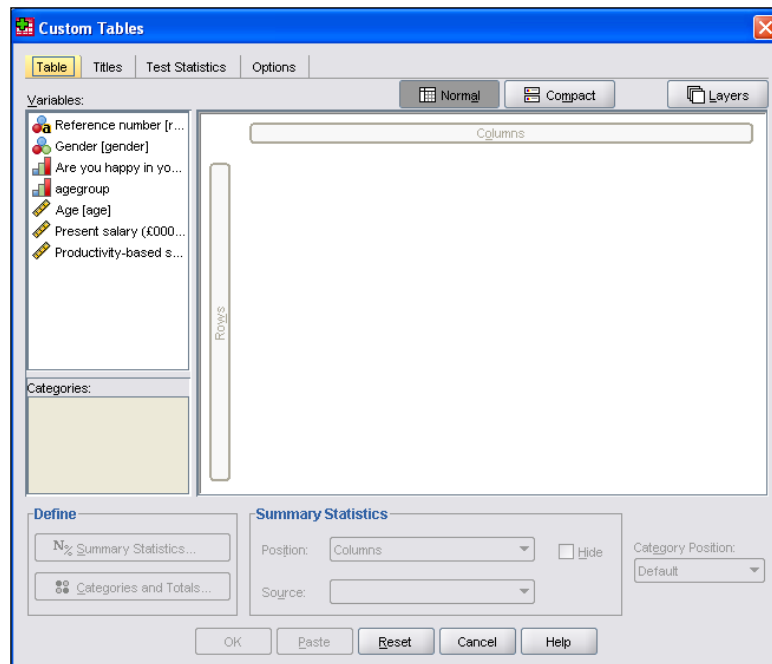
	N	Minimum	Maximum	Mean	Std. Deviation
Present salary (£000)	30	14.9	32.3	22.173	4.4357
Productivity-based salary (£000)	30	15.1	33.4	22.457	4.4264
Valid N (listwise)	30				

6) Comparison of Means

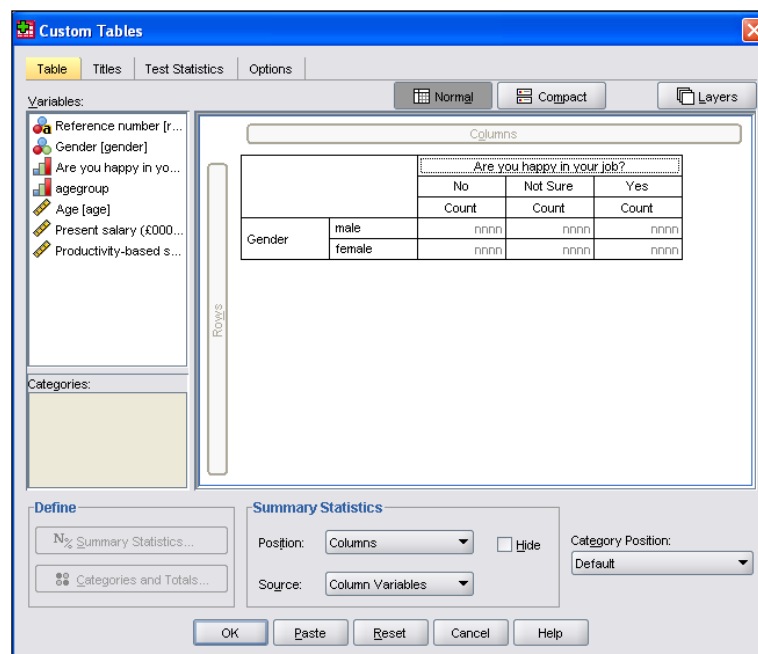
Summarising data for different groups of people can be displayed using a table of means and a bar chart of means.

a) Table of Mean Salaries for Gender and Jobhappy

Select **Analyze, Tables, Custom Tables**. Click **OK** to the dialogue box which provides advice on the use of variables within Custom Tables.



‘Click and drag’ the variables with **Jobhappy** for Columns and **Gender** for Rows. The following table will be generated.



Click and drag **Salary** to replace the **Count** for the cells of the table and the **Mean** will automatically be generated.

Click on the **Titles** tab and type **Mean Salaries of Employees** under Title. **Continue, OK.**

Mean Salary of Employees				
		Are you happy in your job?		
		No	Not Sure	Yes
		Present salary (£000)	Present salary (£000)	Present salary (£000)
		Mean	Mean	Mean
Gender	male	21.8	21.9	23.5
	female	19.8	24.6	21.4

b) Bar Chart of Mean Salaries for Gender and Jobhappy

Select **Graphs, Legacy Dialogs, Bar, Clustered, Define.**

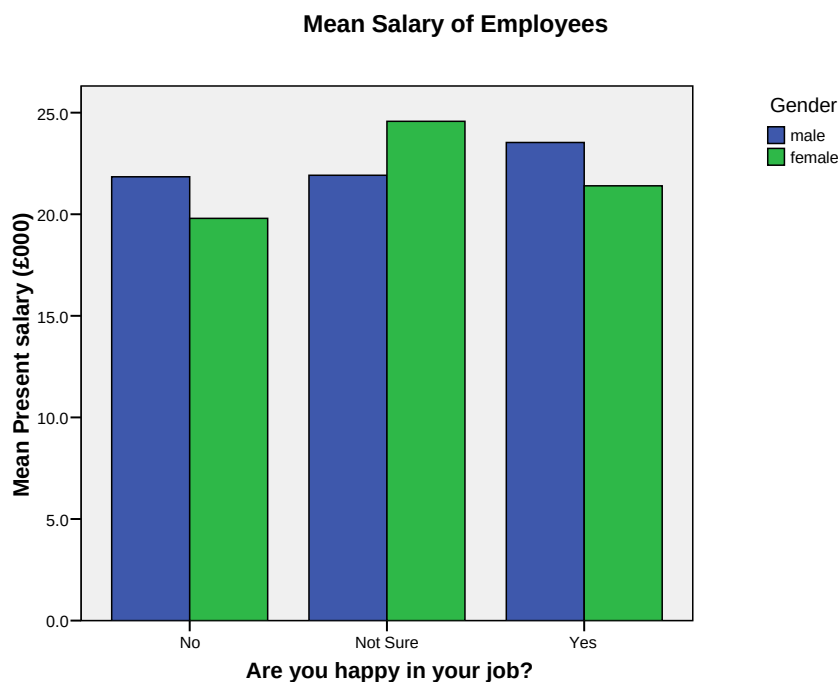
Click on **Other statistic** under Bars Represent.

Select **Salary** for Variable.

(Note that the mean score is selected as the default, but other statistics could be selected by clicking on **Change statistic**).

Select **Jobhappy** for the Category Axis and **Gender** for Define Clusters by.

Click on **Titles** and type **Mean Salaries of Employees, Continue, OK.**



7) Reporting a Statistical Output

It is useful to be able to insert relevant tables or charts into a word processed document as part of a report on statistical analysis.

To transfer SPSS output to a WORD document:

Use the output created from the Comparison of Means in the last analysis.

Click on the table of means from the output and this should now be surrounded by a border with a red arrow on the left-hand side.

Select **Edit, Copy**.

Minimise SPSS by clicking on the minus sign in the top right-hand corner of the screen.

Similarly, it may be necessary to minimise the data editor, by clicking on the minus sign in the top right-hand corner of the screen.

Load **WORD** and type in:

The results below show the analysis of the mean salaries for gender and jobhappy.

Press the return key twice on the keyboard.

Then select **Edit, Paste Special**, click on **Picture, OK**.

(Note that **Edit, Paste** does not always produce the best results for inserting SPSS output into WORD documents and hence **Paste Special** is a preferable alternative, especially for tables).

The selected output should now be pasted into the WORD document.

Press the return key twice on the keyboard before minimising WORD and returning to the SPSS output (this should be seen as an icon at the bottom of the screen).

Click on the Bar Chart of Mean Salaries of Employees and repeat the process:

Edit, Copy, minimise SPSS.

Return to WORD and select **Edit, Paste, OK**.

(Note that this time there is no **Picture** option in the **Paste Special**, and **Paste** gives the best results).

Notes

- (i) The size of the output table can be changed by clicking and dragging the small squares on the sides of the border of a selected table.
- (ii) The titles of the tables have to be copied and pasted separately, but it may be preferable to type these directly into the WORD document as an alternative.
- (iii) **Insert, New Text** can be used as an alternative to add text to the SPSS output.