

Organising data and creating figures (charts and graphs) in Excel



What is in this guide

- Familiarisation with Excel (for beginners)
- Setting up data sheets
- Creating a chart (graph)
- Formatting the chart
- Creating error bars
- Follow-up learning activity
- Other related guides
- Links to screencasts on using Excel

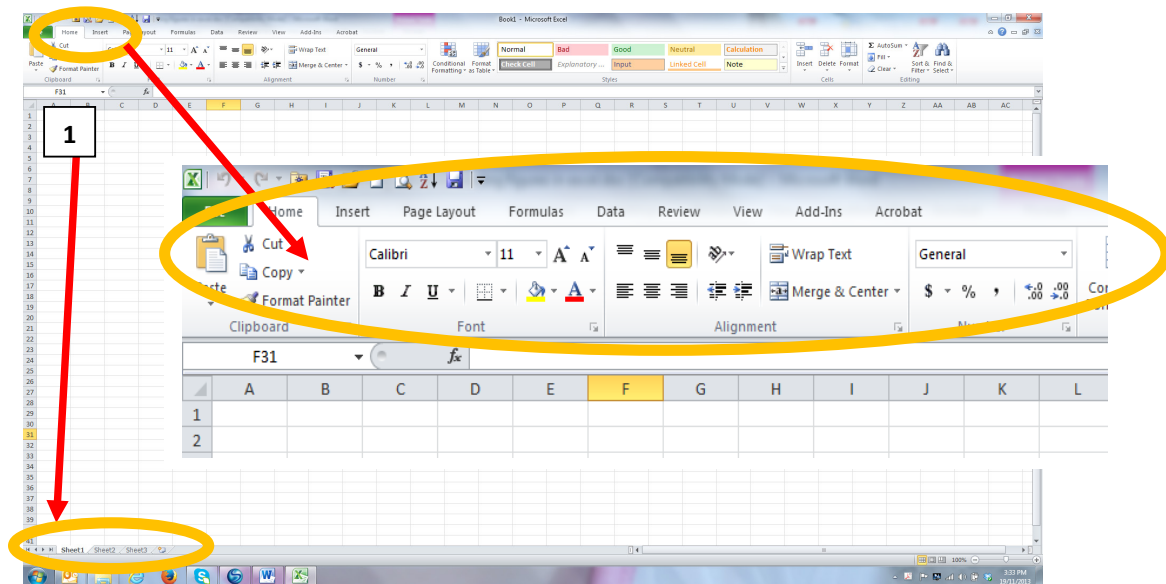
Studying in the Sciences

This guide has been designed for students to follow the steps being explained using **Microsoft Excel 2010**. For different versions of Windows, while how you perform the functions may be slightly different to what is described here; these other versions will still have basically the same functionality. Mac versions (e.g. Numbers) will operate quite differently, although the same functions should be able to be performed.

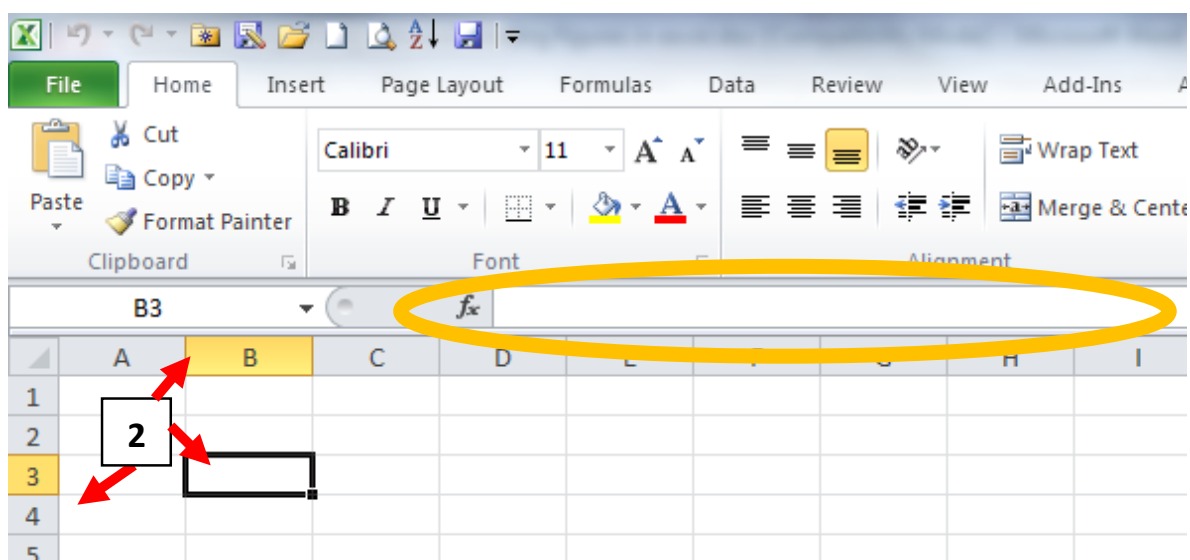
Familiarisation with Excel (for beginners)

Open Excel and view the following features:

1. Click on the **Home** tab. What you see is known as an Excel workbook and you will notice tabs down the bottom saying sheet 1; sheet 2 etc. Each of these is known as an Excel worksheet. Explore the features that are situated under the home tab. Look at the other tab functions from the menu across the top as well

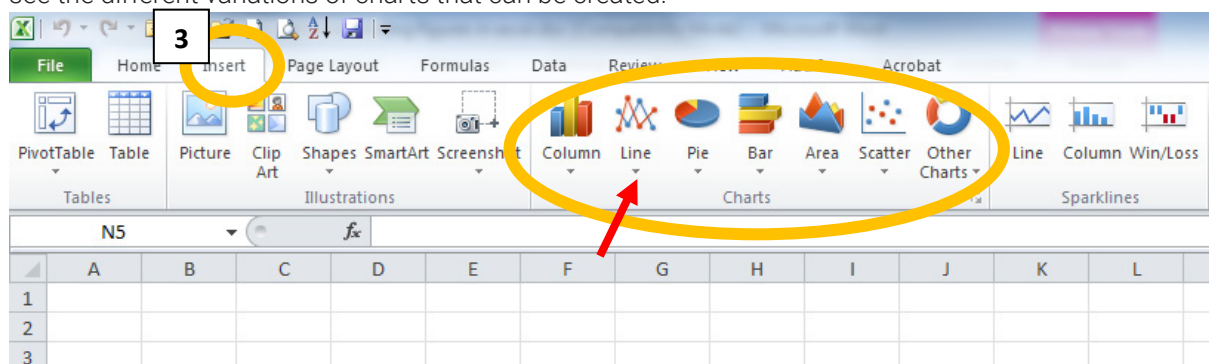


2. Identify the columns (labelled across the top as A; B; C etc.) and rows (labelled down the side as 1; 2; 3 etc). Click on the cell that is the intersection of column B and row 3. You will notice the B and 3 will be highlighted and this cell has a black border with a little square in the bottom right corner indicating the cell is active.



This cell would be referred to as B3. To enter text (or other information e.g. a formula) into the cell you need to click on the cell and type in the cell directly or, with the cell active, put your cursor in the function bar (fx) and type here. Have a go; enter the year you were born. Your year of birth should appear in cell B3.

3. Click on the **Insert tab**. In the **Charts** section, you will see symbols of various types of charts that can be created e.g., line, column, pie, scatter, etc. Click on the down arrowhead under each of these chart types to see the different variations of charts that can be created.



To be able to use these functions, the first thing to do is organise the data you want to plot in the worksheet (data sheet).

Video 1

Screencast link: [Video1_familiarisation.mp4](https://www.youtube.com/watch?v=Video1_familiarisation.mp4)

Setting up data sheets

You need to organise the data in a meaningful way. Use the Excel worksheet to put your data into a table format. This organisation may vary according to the nature of the data you are entering but generally:

- Use column and row labels to identify the data.
- If you have more than one set of data the you want to plot on the same graph.
- To enter the data:
 - Click on the cell where you want the data to go.
 - Type your data into the cell.
 - Press the ENTER key on the keyboard or click on another cell with the mouse.

There are many tools that can be used in Excel to help sort and organise data for more help with this refer to the Excel help menus.

Enter some data

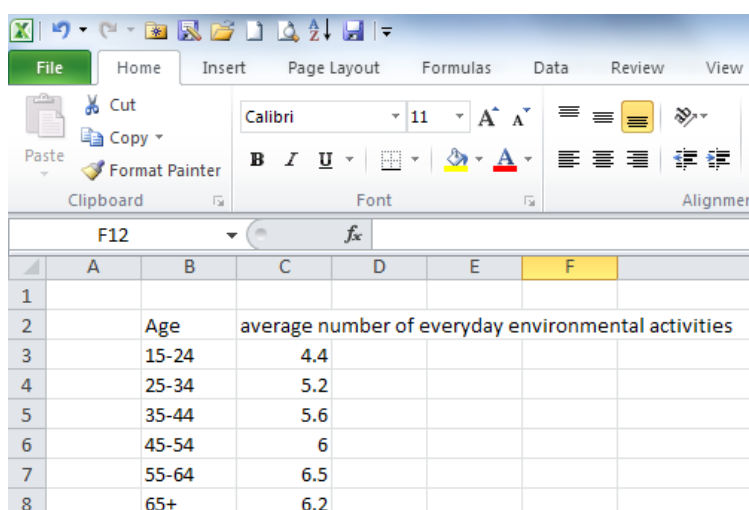
The following data has been extracted from the report – Who Cares About the Environment in 2012? (NSW Office of Environment and Heritage, 2013) and looks at the average number of ‘everyday’ environmental activities carried out by NSW residents, according to age, in 2012.

Age (years)	Average number of ‘everyday’ environmental activities
15–24	4.4
25–34	5.2
35–44	5.6
45–54	6
55–64	6.5
65+	6.2

NSW Office of Environment and Heritage (2013, p. 8)

Activity

Enter the data from the above table into a spread sheet. Remember to enter the data in a systematic manner keeping the paired (independent and dependent) variables together.



Video 2

Screencast link: [Video2_entering data.mp4](#)

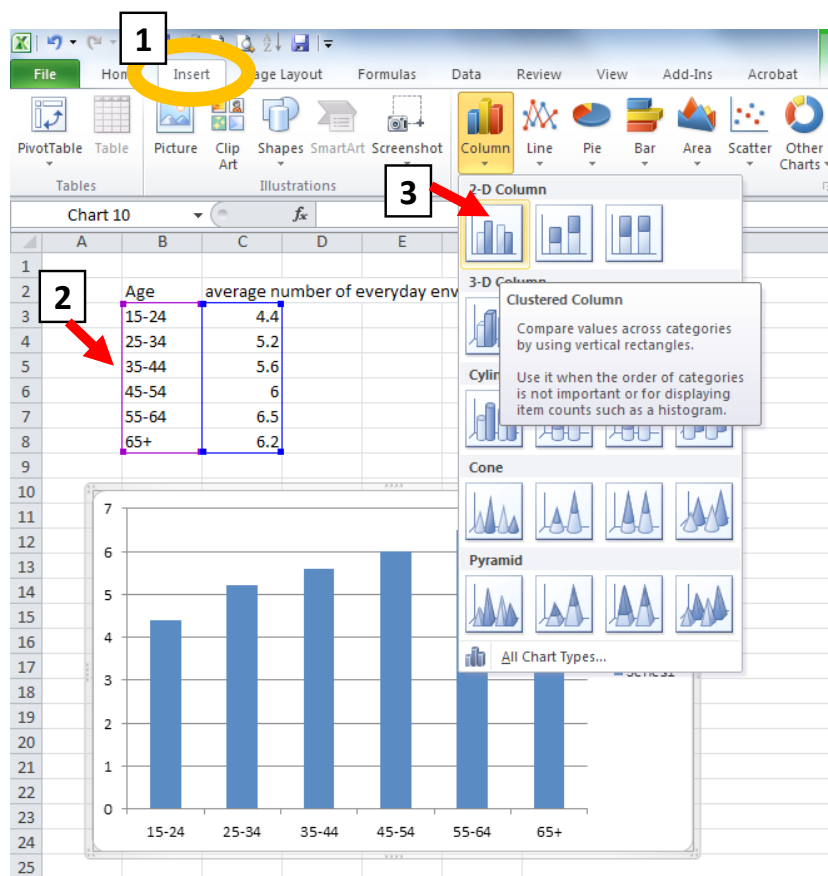
Creating a chart

From the data entered above, the next thing is to create a graph that represents this information in a meaningful way. The data represents variations in the number of environmental activities carried out every day (independent variable) according to the category of age (dependant variable). A **histogram** (a bar graph where columns touch), or a **column graph** (where columns are separate) could be used to show this information. Usually, when the categories are discrete numeric values a histogram would be used, where as if the categories are a nominal group (like gender or geographical location) a column chart is used. The purpose of this exercise however, is to become familiar with using Excel for creating figures. So let's now focus on this.

For information on **what type of graphs to use to display different types of information** see Australian Bureau of Statistics (n.d.) *What graph to display to use when* by clicking this [link](#).

Activity

Using your Excel spreadsheet with the entered data, follow the steps outlined below to produce the graph (chart).



1. Select the **insert** tab at the top of the spread sheet.
2. Select the **data** you want to plot, by clicking and dragging the mouse cursor over the required data to activate the cells.
3. Choose the chart type to display this type of data. A histogram is a good choice as it compares values (number of environmental activities) across numerical categories (age). To produce a histogram, click on the **2-D clustered column** option.

The **gaps between the columns** can then be removed (to form a histogram) or decreased in width – see section on [Formatting the chart](#) (3)

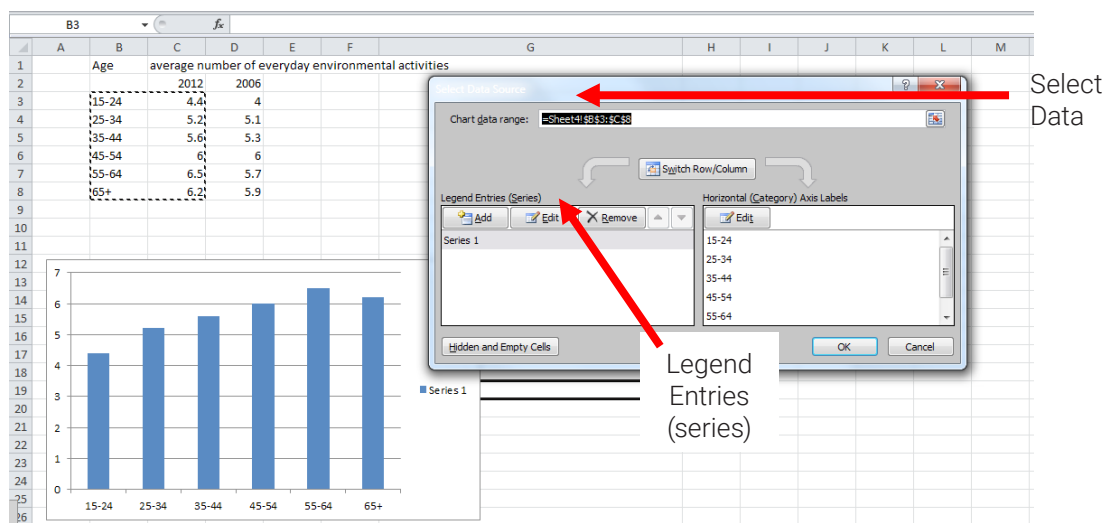
You can use this method to choose other chart types by selecting the type of chart you want to plot (e.g. line, scatter, bar etc.). If you have more than one set of data that you want to display, for example, if you wanted to add a new data set for 2006 to the graph in the above example for 2012, you can do this by adding a second data series. The new data set should have the same variables as the original data set: the same age categories showing the average number of everyday environmental activities conducted. You would enter this data in a new column next to the original 2012 data set.

Video 3

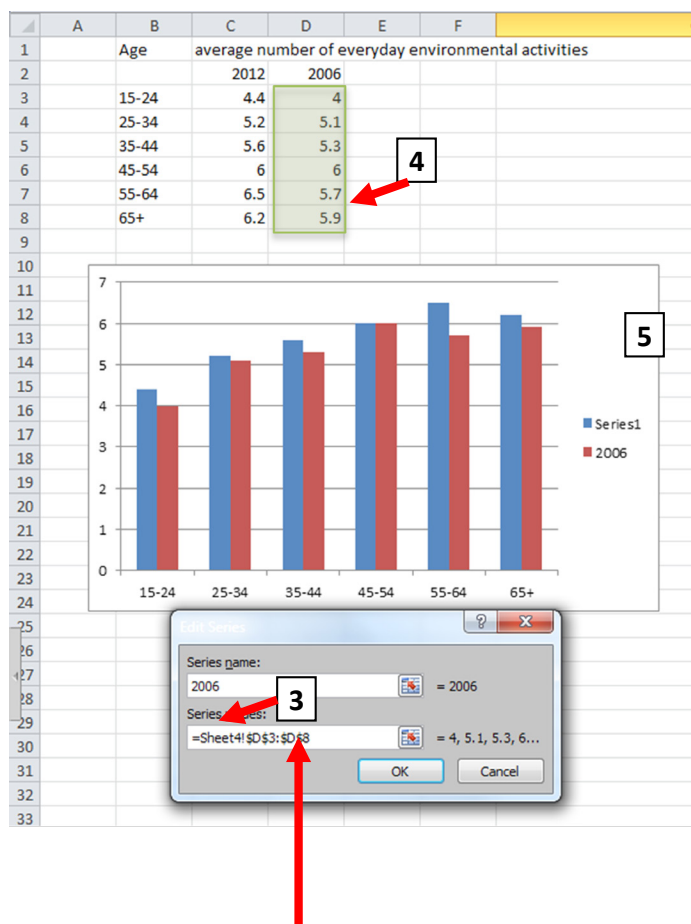
Screencast link: [Video3_creating a chart.mp4](#)

Adding another data set (series)

1. Activate the chart area by clicking in the chart area. A thicker rectangle will appear around the chart.
2. With your cursor on the activated chart area boarder, right click your mouse to get an options menu and select the **data source** option. The **Select Data Source** box will appear. Your column chart so far displays Series 1– this is indicated under the **Legend Entries (series)**. You now want to enter the second data set, Series 2 in other words, the data for the year 2006.



3. Select the **Add** option underneath **Legend Entry (series)**. Another box will appear. Here, in the **Series name** text box, **type** the year '**2006**'. This will label Series 2 as 2006. (see below)
4. You then need to select the series values; this is the data or the variables you want plotted. Put the cursor in the **Series value box** and delete anything that is in there. With your cursor active in this box, **click and drag over the data** you want to be plotted, in this case the variables for the hypothetical 2006 data set. (see next page)



The series values are automatically entered into the box in a code. In the example here:

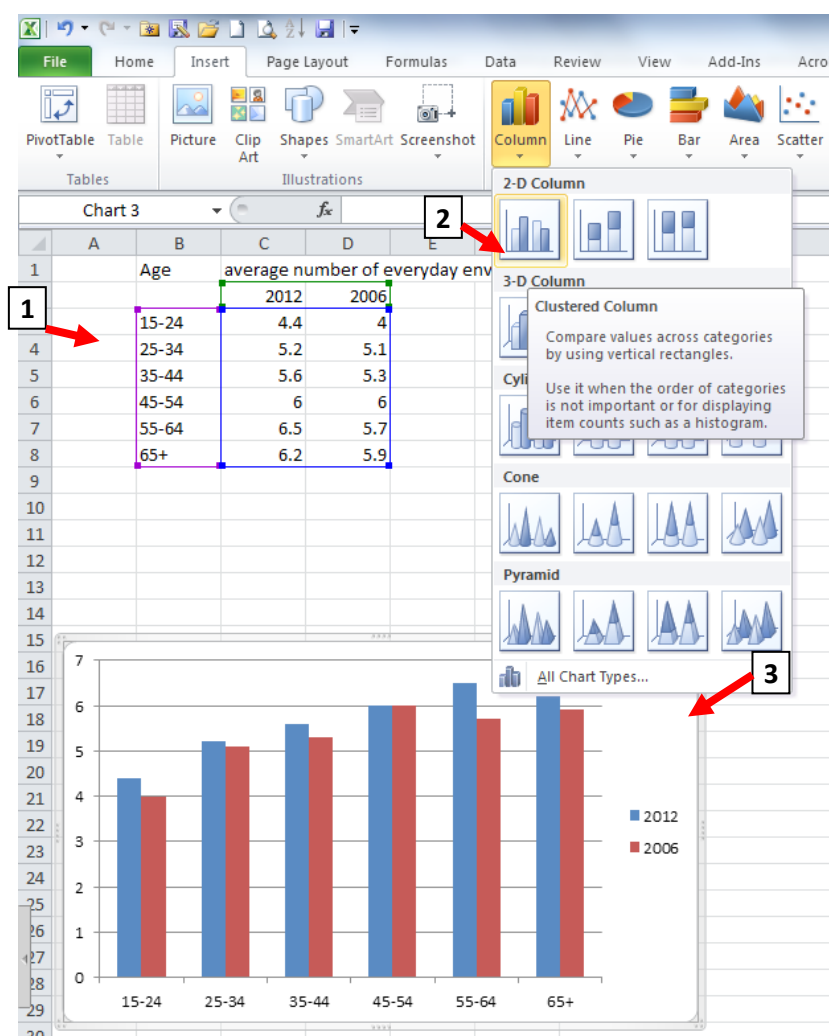
=Sheet4!\$D\$3:\$D\$8. This code tells us that the series values are the values contained on worksheet 4 from cells D3 to D8. (see above)

- Once the series values are entered then click **OK** and the **Select Data Source** box will reappear. (While you are here you can change the series 1 legend entry label to 2012. To do this, **click on the text Series 1**, then underneath **Legend Entries (series)** select **edit**, then in the **Series name** text box that appears type **2012** then select **OK**). Click **OK** in the **Select Data Source** box and the chart will now display the second hypothetical data set – 2006. (see above)

Quick tip

Alternatively, if you know you want to plot two different data sets (series) from the start, arrange your data in columns and label the two data sets as per [1] on the screenshot below, then:

- Select and activate both data sets and data set labels (2012, 2006).
- Choose from the **chart options** the type of chart you want: **Column> 2-D Column > Clustered Column**. Click this option.
- The chart will be generated with both data sets and, as you have included the data set labels, a legend is provided also showing what data is for 2012 and 2006.

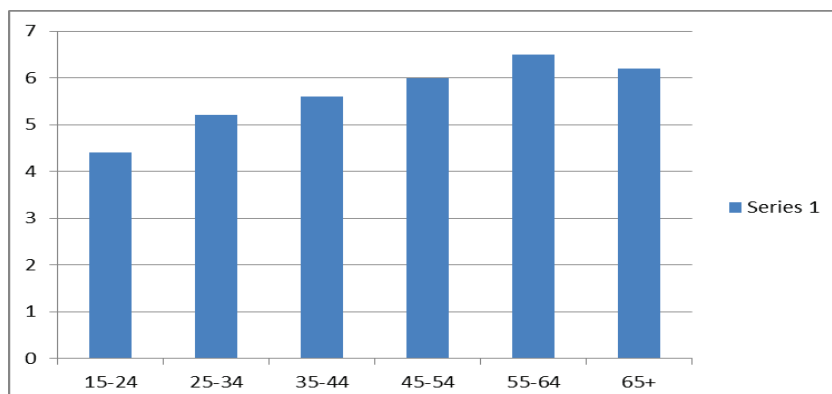


Video 4

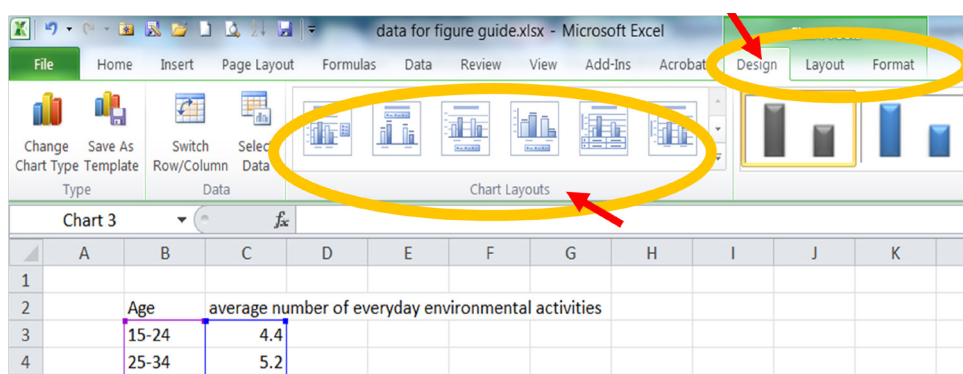
Screencast link: [Video4_adding another data series.mp4](#)

Formatting the chart

The next thing to do is to make sure that your chart's appearance is clear and has all the required features such as axes labels. Let's go back to the plot of the 2012 data on everyday environmental activities carried out. The basic chart generated looked like this:

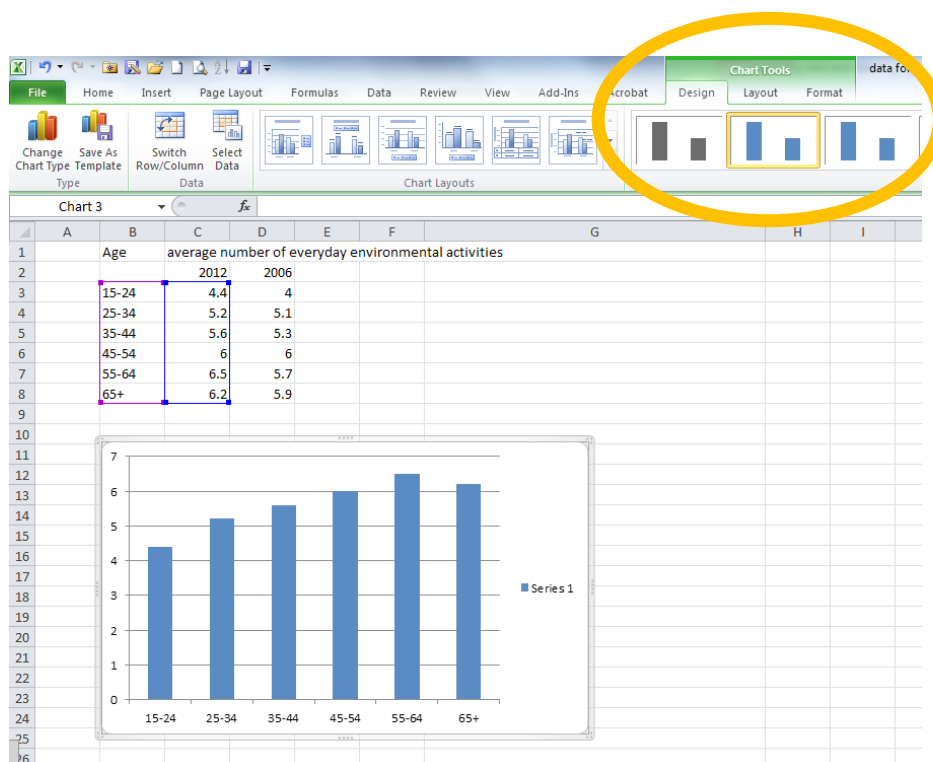


With the chart active (click in the chart area) you can select **Chart tools > Design > Chart Layouts** from the menu tabs at the top. The Chart Layout options, under Design, allow you to select a variety of layout designs that will hopefully match the layout you require. Activate your chart area and look at what the various Chart Layout option produce.

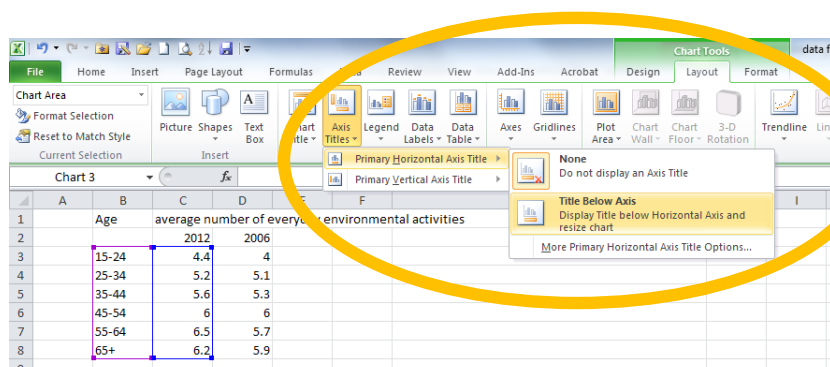


You may need to be more selective and navigate through different menus to get the layout to display the information you need, you may also want to change its appearance or format. For example, get rid of the horizontal grid lines, and change the size. Let's work through some of the other ways to format the chart and/or change the layout:

1. Activate the chart area by clicking in the chart area. A **Chart Tools** tab will appear at the top of the page with three options: **Design, Layout and Format**. Click through these three to get an idea of what functions they provide.



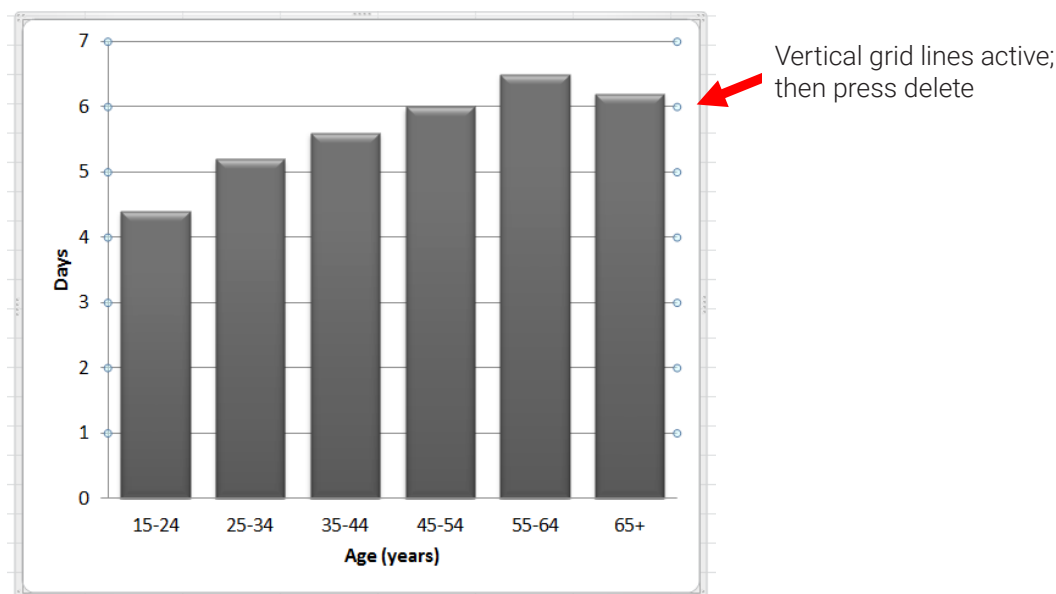
2. To add axis labels: With the chart area active, select – **Layout > Axes titles > Primary Horizontal Axis > Title Below Axis**



A **text box** will appear on the chart under the horizontal axis. Type the axis title – Age (years) – here.

3. Repeat, this time choosing **Primary Vertical Axis > Rotated Title**, and label the axis 'Days'.
4. **Convert column chart into a histogram (change the width of gaps between columns).** **RIGHT click** anywhere on a column to activate a menu box. Select **Format data series** at the bottom of the menu. The **Series Options** menu (see below) will appear. Use the **Gap width** slide to change the width of the columns and gaps. For a histogram there will be no gap between columns. In this case, by selecting **Border Color** from the **LEFT** hand menu and opting for a colour different to the fill, each column will be more clearly defined.
5. **Change the colour/ style of the chart:** With the chart area active, select – **Design** then under the **Chart Styles** section, select the chart colour/ style desired
6. **Remove or format the legend:** To remove the legend 'Series 1' click on 'Series 1' in the chart area and it will activate the legend. Press **delete** to remove. If you want to keep it you can drag and drop it in another position. Alternatively, formatting options for the legend are under **Layout > Legend**. To change the label (Series 1), follow the steps above ([click here](#)).

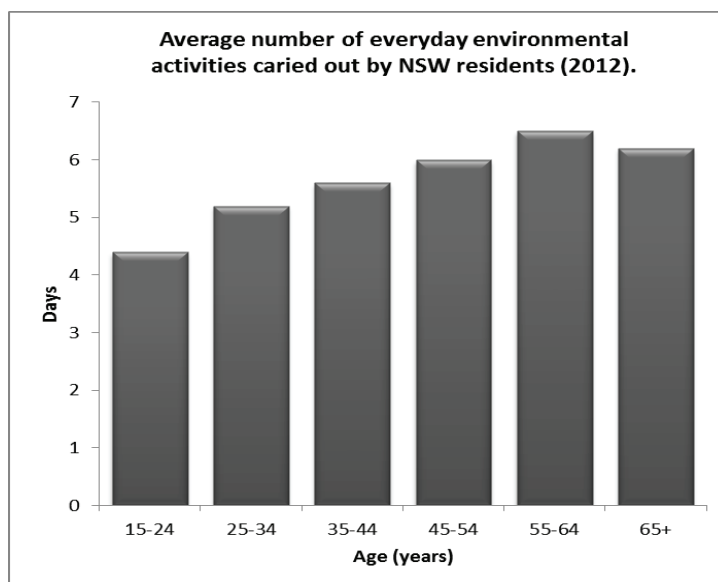
7. **Remove the gridlines:** Click on a vertical grid line to activate all the gridlines in the plot area, then press **delete**. The gridlines will be removed.



8. **Create a title:** You may want to add a title to the chart. If you are pasting the chart into another document as a figure it is probably better not to provide a title here but rather a figure caption in the document where you are pasting it. If you want to create a chart title here, then follow these steps. Activate the chart area, then select; **Layout > Chart Title > Above Chart**. This will put a text box at the top of the chart where you can insert text to create a title. You can change the font style, size etc. by highlighting the text and using the **font tools** under the **home tab**.

9. **Resize the chart:** With the chart area activated, drag the corner of the chart area.

Below, formatting changes have been made with the end result looking like this:



For other formatting options, explore the functions under **Chart tools** and use the Excel **help menu** for assistance.

Video 5

Screencast link: [Video5_formatting the chart.mp4](#)

Creating error bars

You may have been asked to add error bars to your chart. This is more advanced than some of the earlier concepts and it is something you may not need to know. The following example shows **how to add error bars to a chart**. The data being plotted are mean scores and standard deviations for items on a test used in speech therapy (Hill et al., 2009, p. 739). There are two data sets: scores for the test administered face to face, and scores for the test administered via tele-rehabilitation. The mean values have been plotted using a column chart as per the method used earlier in this guide. The next step is to add the error bars, in this case standard deviations (SD). Each mean value entered into the column has an associated standard deviation; thereby the SDs need to be entered into a column with so that the corresponding SDs and mean values are linked, that is they occur in the same row/ next to one and other.

1. Activate the chart area. Select one of the two data sets (series) by clicking on one of the columns.
2. Select **Chart tools tab > layout > error bars**. From the drop down menu in error bars select **more error bar options**.
3. Select **Vertical error bars** and click the option you want. Here we are going for error bars in the **Plus** direction and the End style is Cap. Because we want the error bars to represent specific values, we have to designate where those values are so select **Custom** underneath the **Error Amount** option, then **Specify Value**. A **Custom Error Bar** box will appear.
4. Click on the **Positive Error Value** function box, delete any content then, on the spread sheet, click and drag over the cells containing the standard deviation values you want plotted. (Tip. As each standard deviation to plot is specific to a certain mean value, it is very important that the data is arranged so that the mean and related standard deviations are in the same row). The code for the cells contain this data should appear in the **Positive Error Value** function box then click OK and the error bars should appear. Close the **Format Error Bars** box
5. You then need to repeat these steps for the other data series.

Video 6

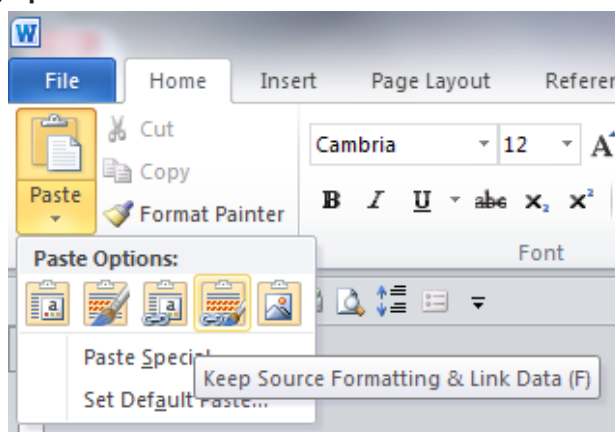
Screencast link: [Video6_adding error bars.mp4](#)

Inserting the chart into a document or PowerPoint presentation

Once you have created your chart (figure) using Excel you can copy and paste it into another document.

1. Copy the chart by activating the chart menu by right clicking on its boarder then selecting **copy** on the **control tab** or using the **CTRL+C** short cut
2. Paste the chart where you want it in the new document or presentation by putting your cursor where you want it then selecting the **paste** options under the **Home** tab (see screen shot below). You can choose the option **keep source formatting and link data** – this allows you to open the source data file and edit the chart directly from the pasted figure, or you can paste the object as a picture file where you can crop and resize it (you won't be able to edit the figure itself). If you are placing it in a presentation ensure that colours are appropriate for easy viewing and font sizes can be read. If inserting the figure into a presentation you can play around with the paste options to get the best appearance for the figure. It is preferable to use the **embedded work books or source formatting options** (off the **paste menu**) as this will allow you to change things so the presentation style is clear. The **use destination them** option will tend to select colours that are consistent with the presentation design and that can be seen on the background theme of the presentation.
NB. You can simply choose **paste or CTRL+C**. This will paste the figure using the default paste options set up in your system.

3. Once you have inserted the chart you can resize it by clicking and drag on the corner boarder. You can align the chart on the page by clicking on the figure then aligning it centre right or left by choosing these options under the **Paragraph** tab under the **Home** tab.



4. Resize the plot area so that it fills most of the chart window. Click inside the plot area and use the mouse to drag on the little black boxes around the plot.

Follow-up learning activity

The following data has been extracted from the NSW Office of Environment and Heritage (2013, p. 14) showing response rates for the two biggest environmental concerns identified by NSW residents over the past 15 years. The data shown for this activity is a selection of three of fourteen categories identified from the source data set: namely, water supply/conservation, climate change and mining.

Table1 Percentage of surveyed NSW residents identifying the selected environmental issues (water supply/conservation, climate change, mining) as either the most or second most important issue (1997–2012).

	Year					
	1997 (%)	2000 (%)	2003 (%)	2006 (%)	2009 (%)	2012 (%)
Water Supply/ Conservation	3	4	27	58	43	18
Climate Change	3	5	5	13	23	12
Mining	1	1	2	0	3	17

Task

Have a go:

1. Organise the data from the above table in an Excel (or equivalent) spread sheet.
2. Create a line graph to show the trends in perception of NSW residents for these three environmental issues from 1997 to 2012.
3. Check your work by viewing the [feedback](#) provided below.

Other related guides (available on the CTL site)

- Studying in the sciences resources: Using figures to present information in science.
- Numeracy resources: Statistics: Topic 6: Correlation and Regression (for how to create a linear graph showing correlation/regression)

Other helpful links

Australian Bureau of Statistics (n.d.) What graph to display to use when. Retrieved from http://agencysearch.australia.gov.au/search/click.cgi?rank=1&collection=agencies&url=http%3A%2F%2Fwww.abs.gov.au%2Fwebsitedbs%2FCaSHome.nsf%2Fhome%2Fdownloadable%2Bfiles.es%2F%24FILE%2FWhat_Graph_or_Display_to_use_When.docx&index_url=http%3A%2F%2Fwww.abs.gov.au%2Fwebsitedbs%2FCaSHome.nsf%2Fhome%2Fdownloadable%2Bfiles.es%2F%24FILE%2FWhat_Graph_or_Display_to_use_When.docx&auth=C2xqci83fYTCzdHkaCwzbA&query=what%20graph%20to%20display%20when&profile=abs

Learning Development, University of Leicester. (2012). Presenting numerical data. Retrieved from <http://www2.le.ac.uk/offices/ld/resources/numeracy/numerical-data>

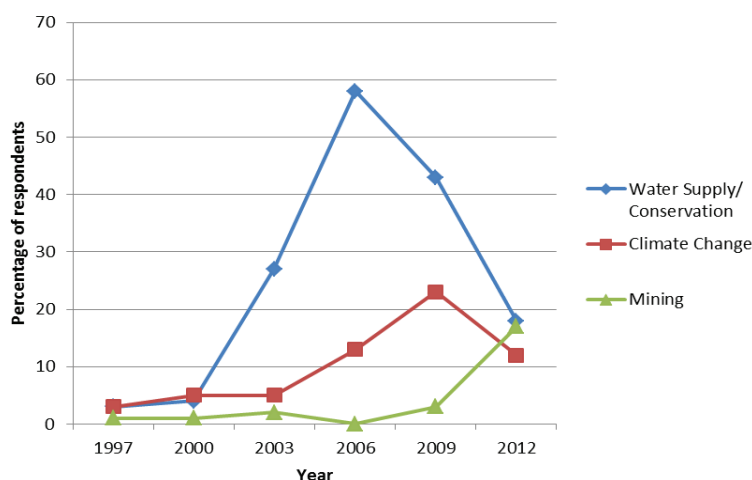
References

Hill, A.J., Theodoros, D., Russell, T., & Ward, E. (2009). Using telerehabilitation to assess apraxia of speech in adults. *International Journal of Language & Communication Disorders*, 44(5), 731–47.

NSW Office of Environment and Heritage. (2013). Who cares about the environment in 2012? At a glance. Retrieved from <http://www2.le.ac.uk/offices/ld/resources/numeracy/numerical-data>

Feedback on learning activity

Your chart should look something like this.



Here is a summary of the steps used to construct this chart in Excel:

1. Organise data think about what you want to show in a line graph on trends, that is what happens to the data over time for each of the three categories (environmental issues). As such time (or years) should be represented on the horizontal axis (independent variable) and the percentage of respondents (dependant variable) on the vertical axis. Organise your data like this (see screen shot below) in the spread sheet with t dependant variable running down the column and the independent variable that will appear on the horizontal axis running along the top row. Type each of the 3 data categories for type of environmental issue as row headings. This will save you entering titles for each of the data categories (series) in Excel later.
2. On the insert tab at the top of the Excel spread sheet to get the Chart options up.
3. Highlight the data you want to use to create the line graph. Selecting the data and columns (years) and row (categories of environment). This way you will plot the data and have a legend for the 3 categories and the years will be labelled on the horizontal axis. Once you have selected the data then click on the line chart icon and select 2-D line chart with markers.
4. Format the chart to add axes titles. Under **chart tools tab select > layout > Axes titles > Primary horizontal (or vertical) axis title > Title below axes**. This will give you a text box to type a label along the chosen axis. Repeat for the other axis.
5. There are other formatting and layout options you could complete but this gives you a pretty good chart. Don't forget to the label and number your figure and provide a caption/title when you insert it into a document (see studying in the sciences resource: Using figures to present information in science for further information).