Microsoft Excel 2003™ - An Intermediate Guide

http://www.reading.ac.uk/its/documents/training/excel/inter/

Contents:

See Document Conventions for information on the meaning of the text formatting used below.

Introduction

This intermediate level guide deals with some of the more scientific facilities available in Excel. It also covers database facilities (sorting and selection) plus customisation of charts. The guide assumes that the user is familiar with the topics covered in *Microsoft Excel - A Beginners' Guide*. More advanced notes (*Advanced Spreadsheet Topics Using Microsoft Excel*) are also available.

Opening the Example File

Open: 

Log into an IT Services PC as usual and start up Excel.

1. Click on the **[Open]** button, or use **Open...** from the **File** menu
2. Change the default directory by clicking on the *list arrow* attached to the *Look In:* box - choose **User (D:)** then *double click* on the folder called **Training**
3. Click on **phoenix.xls** and press **<Enter>** for **[Open]**

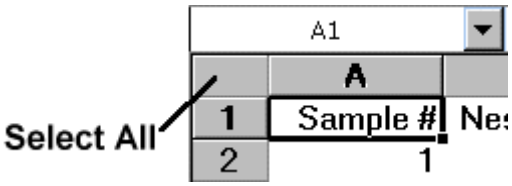
Note: For those using these notes on a computer not run by IT Services, the example file can be downloaded from the link provided on the web page at step 3 above.

The set of data which appears is information on the size, weight and colour of 50 birds eggs recorded from nests at a variety of altitudes. The initials of the collectors are also included.

Selecting Cells

Excel has various shortcuts for selecting a range of cells (besides the usual method of dragging through the required cells). These include:

- *entire rows/columns*: click on the row/column heading number/letter
- *a block of cells*: click in the top left corner of the block then, holding down **<Shift>**, press ***<arrow>*** keys to extend the selection in the directions of the arrows. Alternatively, select the top left cell then hold down **<Shift>** and click in the bottom right cell
- *a row/column range*: click on the starting row/column heading then either hold down **<Shift>** and use the ***<arrow>*** keys or **<Shift>** click on the end row/column heading
- *non-contiguous cells or blocks*: hold down **<Ctrl>** as you make your selection
- *part rows/columns*: click on the cell at the start of the selection then hold down both **<Ctrl>** and **<Shift>** and press the ***<arrow>*** key in the direction required. Alternatively, hold down **<Shift>** and *double click* on the active cell border in the direction required. Repeat this, using either method, to extend the new selection to a block of cells
- *contiguous cells*: hold down **<Ctrl>** and press the **<*>** key (or use **<Ctrl Shift 8>**)
- *the entire worksheet*: click on the *Select All* button - the grey rectangle in the top left corner of the worksheet (where the row and column headings meet) or use **<Ctrl a>**



You can also select a range of cells using a command - open the **Edit** menu and choose **Go To...**. In the *Reference:* box, type in the cell or range of cells you wish to select. Note that when using this method, the selection is added to the *Go to:* list so that you can easily make the same selection again. Named ranges also appear in this list. **Go To...** is also activated by **<Ctrl g>** and **<F5>**.

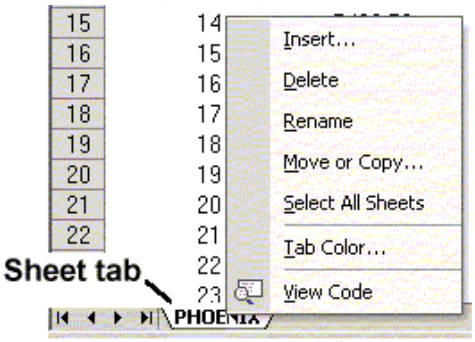
Try using some of the above methods to select different areas of the worksheet.

Data Manipulation

Using Multiple Sheets

In the *Beginners' Guide*, everything was done on a single worksheet, however an Excel workbook can contain several sheets, each with different sets of data, charts etc. Indeed, graphs are usually stored on separate *chart sheets*.

The *sheet tab* (at the foot of the screen) shows the worksheet name. New sheets can be created, old ones renamed, copied or deleted. In this first exercise, you will create a copy of the *phoenix* sheet (this is quite a good idea with any worksheet as it allows you to play around with the data while preserving a copy of the original). You will also see how to insert, rename and delete a sheet.



To make a copy of a worksheet:

1. Position the cursor over the *PHOENIX* tab, press the *right* mouse button then, from the shortcut menu, select **Move or Copy...** (or, from the **Edit** menu, use **Move or Copy Sheet...**)
2. In the *Move or Copy* window, click on the *Create a copy* check box in the bottom left corner and set *Before sheet* to **(move to end)** - note that you can also copy or move a worksheet to an entirely new workbook or one already open under *To book:*
3. Press **<Enter>** or click on **[OK]** and a copy, *PHOENIX (2)*, will be created on the right

To create a new worksheet:

4. Repeat step **1** but select **Insert...** from the shortcut menu
5. In the *Insert* window, check **Worksheet** is selected then press **<Enter>** or click on **[OK]** (or, from the **Insert** menu, use **Worksheet**)

To rename a worksheet:

6. *Right click* on the new *Sheet1* tab and select **Rename** from the shortcut menu (or, from the **Format** menu, select **Sheet** then **Rename**) - *Sheet1* is now highlighted
7. Type in the new name (call it `Red-Brown` - you'll be using this later) then press **<Enter>**

Note: you can also rename a worksheet by double clicking on the sheet tab name.

To delete a worksheet (eg the copy of the data):

8. *Right click* on the *PHOENIX (2)* tab then select **Delete** from the shortcut menu (or, from the **Edit** menu, use **Delete Sheet**)
9. You need all the sheets here, so click on **[Cancel]** - normally you would press **<Enter>** for **[Delete]** to confirm the deletion

The order of the sheets can be altered simply by dragging and dropping the sheet tabs.

10. Position the mouse cursor over the *Red-Brown* sheet tab
11. Hold down the mouse button and drag the sheet to the right hand side
12. Release the mouse button to drop the sheet tab in the new position

Cell References when Using Multiple Sheets

If you want to use a cell on another sheet in a formula, then you have to include the sheet name in the cell reference. The full reference is of the form *Sheet*!*Cell* (you may have already noticed such references, eg when you have used the *Chart Wizard*). Usually it's easiest to fill in the reference by moving to the sheet concerned and clicking on the required cell:

1. Move to the empty sheet by clicking on the *Red-Brown* tab
2. In cell *A1* type `=10*`
3. Click on the *PHOENIX* tab then on cell *D2* - note how the cell reference is added to the formula
4. Press <**Enter**> and cell *A1* on the *Red-Brown* sheet is filled in
5. Move up to *A1*, note the cell reference in the formula again, then press <**Delete**> to empty the cell contents (they are not needed here)

Sorting the Data

Sort Ascending:  Sort Descending: 

It's often useful to have data displayed in some order - alphabetical or increasing/decreasing numbers. To sort the data based on a particular column, the active cell must be in that column. Here, you are going to sort the data in order of increasing nest altitude:

1. Move to the original copy of the data by clicking on the *PHOENIX* tab
2. Move the active cell into column *B* - click on any cell containing data
3. Click on the **[Sort Ascending]** button

The eggs should now be ranked in order of nest altitude, with the lowest at the top. Note how the sample numbers in column A have changed.

Tip: Occasionally you may want to sort only part of the data, leaving the surrounding data exactly as it is. To do this you must first select the data before you click on a sort button. Try this out if you like, but make sure you reverse any sort before you continue the course - use the **[Undo]** button to undo any changes. For more complex sorts, you have to use **Sort...** from the **Data** menu.

To learn more about sorting, see the *Sorting, Subtotals and Outlines in Microsoft Excel* advanced notes

Creating Data Subsets

Excel offers limited database facilities whereby you can extract a subset of the data to work on. As an example, store the red-brown eggs on the sheet created earlier.

1. Make sure you are using the original copy of the data - *PHOENIX* - and that the active cell is somewhere within the data
2. Open the **Data** menu, select **Filter** then **AutoFilter** - a *filter arrow* is added to each column heading

3. Click on the *filter arrow* in cell *F1* and choose *Red-Brown* from the list which appears - only red-brown eggs are now shown (it tells you *21 of 50 records found* at the bottom left)
4. Select all the red-brown records - press **<Ctrl *>**
5. Press **<Ctrl c>** or click on the **[Copy]** button (or choose **Copy** from the **Edit** menu)
6. Move to the *Red-Brown* worksheet by clicking on its tab
7. Make sure you are in cell *A1* then just press **<Enter>** for **[Paste]**
8. To resize (*autofit*) the columns, drag across the column headings *A* to *G* then *double click* on any of the dividing lines separating one column heading from its neighbour
9. Move back to the original data by clicking on the *PHOENIX* tab

For practice, you could repeat the whole procedure for the blue-green eggs, storing them on a new sheet labelled *Blue-Green*.

Excel allows you to set selection criteria on more than one column - for example, you might want eggs that are both red-brown and which have a diameter of more than 8.5cm. To do this:

10. Click on the *filter arrow* in cell *D1* and this time choose **(Custom...)**
11. In the new *Custom AutoFilter* window which appears, use the *list arrow* attached to the left-hand box to select **is greater than**, in the right-hand box, type a value of **`8.5`**
12. Press **<Enter>** for **[OK]** - only red-brown eggs over 8.5cm in diameter are now displayed

Finally, turn off both selection criteria and redisplay all the original data as follows:

13. From the **Data** menu choose **Filter** followed by **Show All**

Note: you can turn off one criteria at a time by using the list arrow at the top of each column and choosing **All**

14. To remove the filter arrows, open the **Data** menu and choose **Filter** followed by **AutoFilter**

More complicated selections can be made using the **Advanced Filter...** command. To learn more about filters, see the *Using Filters in Microsoft Excel* advanced notes.

Text Rotation and Wrapping

One feature which you may find useful, especially if you have a column heading which is much wider than the data beneath, is text rotation.

1. Mark row *1* by clicking on the row number **1**
2. Open the **Format** menu and choose **Cells...**
3. Click on the **[Alignment]** tab

Under *Orientation* on the right-hand side, you can either have the text reading downwards or at any angle to the horizontal. Text can be rotated in any cell or group of cells using this feature.

4. Alter the *Orientation* (eg to 90°) then press **<Enter>** for **[OK]** to see the effect
5. Click on the **[Undo]** button to reset to normal

Also on this tab, **Wrap text**, under *Text control*, allows text to wrap onto several lines within a cell. You can also force text onto a new line by holding down **<Alt>** and pressing **<Enter>** - Excel automatically adjusts the row height for you.

6. Repeat step **2** then, on the **[Alignment]** tab, turn on the **Wrap text** check box

7. Under *Text alignment* at the top, set *Vertical:* to **Center** then press **<Enter>** for **[OK]**
8. *Double click* on cell *D1*, move the insertion point to immediately before *(cm)* then hold down **<Alt>** and press **<Enter>** - you can *autofit* the column width, if you want
9. Repeat step **8** on cell *E1*
10. Decrease the width of column *B* so that the contents wraps round onto 2 or 3 lines

Tip: Double clicking on a cell lets you edit the cell contents without having to use the *Formula Bar*.

Further Graphics

XY Graphs

A quick examination of the sorted data should show you that a clear relationship exists between nest altitude and the other variables - the eggs get bigger the higher the altitude. Try plotting this on a chart next. With scientific data, both the independent and dependent variables usually vary over a *continuous* range. This contrasts with the business data used in the *Beginners' Guide*, where the dependent variables (income, costs, etc) applied to equal time periods (years). To activate the x-axis, an *X-Y* chart must be used.

1. Select the data needed for the graph *B1* to *E51* - starting in cell *B1* hold down **<Ctrl>** and **<Shift>** and press ***<down arrow>***, then use **<Shift>** and ***<right arrow>*** *three times* to stretch the selection across to column *E*
2. Click on the **[Chart Wizard]** button
3. Select an **XY (Scatter)** chart (use the default sub-type) - press **<Enter>** or click on **[Next>]**
4. Check the *Data range:* is correct (**B1:E51** - the extra $ signs will be explained later) and that *Series in:* is set to **Columns** then press **<Enter>** or click on **[Next>]**
5. Type in a *Chart title:* of **`Phoenix Eggs`** (press **<Tab>**), a *Value (X) Axis:* label of **`Metres`** (**<Tab>**) and a *Value (Y) axis:* label of **`Centimetres`** - press **<Enter>** or click on **[Next>]**
6. Change *Place chart:* to **As new sheet:**, call it **`EggPlot`**, then press **<Enter>** for **[Finish]**

The chart is drawn on a separate sheet, *EggPlot*, as indicated at the foot of the screen, and the *Chart Toolbar* appears. If the toolbar doesn't appear (or if it disappears later) open the **View** menu, choose **Toolbars** and make sure **Chart** is ticked on.

Customising your Graphs

When you draw a graph on a separate chart sheet, the menus along the top of the screen are set ready for you to customise it. If you draw a graph on a worksheet then you must first click on it to change the top menus (handles appear on the graph border and the *Chart* toolbar is displayed). Excel now offers various facilities, including adding trendlines; adding or changing text (titles, axis labels etc) or the legend; modifying the axes (range, tick marks etc); and changing line and shading patterns. You can also select different font, foreground, background and infill colours.

Secondary Axes

Though all three measurements correlate positively with nest altitude, the length and diameter are on a much smaller scale than the mass and as such look a bit silly on the same graph. To overcome this, you can add a second y-axis.

1. *Double click* on one of the (blue) egg mass points - alternatively *right click* and choose **Format Data Series...**
2. In the *Format Data Series* window, click on the *Axis* tab
3. Set *Plot series on* to the **Secondary axis**

4. Move to the *Patterns* tab and set *Line* first to **Custom** then to **None**
5. Press **<Enter>** for **[OK]**

To complete the graph you need to add a label to the new axis.

6. Open the **Chart** menu (or *right click*), select **Chart Options...** then click on the *Titles* tab
7. Click in the **Second value (Y) axis** box, type **`Grams`** then press **<Enter>** for **[OK]**

The graph is now much clearer. Note that once a secondary Y-axis has been turned on, you can also create a secondary X-axis from the *Axis* tab in **Chart Options...**.

Trendlines

Excel gives you the opportunity of adding best-fit lines to your plotted data. To do this:

1. *Right click* once on one of the data points (or select a data series by clicking normally then open the **Chart** menu) and choose **Add Trendline...**

You now have a variety of *Trend/Regression types* to choose from, while the *Options* tab offers several important extras. Here, keep the type of line as the default - ie linear.

2. Click on the *Options* tab and set on **Display equation on chart** and **Display R-squared value on chart** by clicking in the check boxes
3. Set the *Forecast* to extend the line *Backward:* by **`1158`** units (to meet the y-axis)
4. Press **<Enter>** or click on **[OK]**
5. Repeat from step 1 for the other data series, if you want to fully annotate your graph

Changing Text

To modify the main title, for example, *double click* on the words *Phoenix Eggs* on the graph. A *Format Chart Title* window will appear in which you can change:

- **Patterns** allows you to box the title or have a different coloured background
- **Font** lets you change the lettering. You could select a different *Font:* (use the scroll bar to move up and down and click on a font to select it), change the *Size:* and also reset the *Color:*
- **Alignment** allows you to change the text orientation and alignment within the text box

Reset some of the available options, if you like (eg make the font size bigger). Press **<Enter>** or click on **[OK]** and the title is changed appropriately. For practice, reset some of the characteristics of some of the other text on the graph.

Changing the Legend

To change the legend *double click* anywhere inside the box. The *Format Legend* window appears - here you can again change the **Patterns** and **Font**, but also:

- **Placement** the position of the legend on the page. Note that you can also drag the legend around to any position you like, as you can any other text box (eg the regression equations)

Press **<Enter>** or click on **[OK]** to close the *Format Legend* window.

Click on the already-selected legend then *double click* inside it again and a different window appears - *Format Legend Entry*. Here you can change the characteristics of the legend font.

Modifying the Axes

To change the settings on an axis simply double click on it. As an example, have a look at the y- or value (vertical) axis. *Double click* anywhere on the (left) axis line itself or on the labels to its left. A *Format Axis* window will open, allowing you to modify:

- **Patterns** allows you to change the axis width, style and colour and tick mark settings
- **Scale** lets you set the range (max/min values) and tick mark intervals. You can also select a logarithmic scale or reverse the order of the data. Note that the options available for the x-axis differ (unless you have an X-Y graph)
- **Font** and **Alignment** give you the same options as before
- **Number** sets a format for the numbers used on the axis - the default is *Linked to source*

Press **<Enter>** or click on **[OK]** to return to the graph.

Changing Data Markers and Line Style

To change the point markers or line style, simply *double click* on one of the points. The *Format Data Series* window will appear in which you can change:

- **Patterns** allows you to modify the line style and point markers. You can also smooth the line joining up data points. On a chart with shaded areas you can change the shading colour and, under *Fill Effects*, add patterns, texture, two-colours and shading gradients
- **Axis** lets you plot the data on a secondary axis - it creates one if necessary
- **Y Error Bars** lets you add error bars - for X-Y graphs there is also an **X Error Bars** tab
- **Data Labels** gives you the opportunity of attaching data values or labels to your data points
- **Series Order** lets you change the order of the series in the legend
- **Options** other options can sometimes be set which vary with the chart type

Press **<Enter>** or click on **[OK]** to get back to your graph. You can now repeat the process with the other series if you want. Note that trendlines have to be formatted independently - double click on the line and set the options required in the *Format Trendline* window.

You can also format the gridlines - double click on any gridline and set the options required in the *Format Gridlines* window.

Warning: *Excel also allows you to format a single data point in a series. However, you must be very careful when doing so or you can change your original data. Select the data series, click on the required point then open the* ***Format*** *menu and choose* ***Selected Data Point****. You could now use a different marker for that point or display its value. Take care you don't move the point slightly when selecting it as this alters the original data value on the worksheet!*

Changing Chart and Plot Area

The Chart and Plot Areas form the background to the chart itself. By default, a grey *Plot Area* is placed on a white *Chart Area*. When printing out on a black and white printer, this can cause problems - it's better to make the *Plot Area* white also.

Double click on the grey background to the Chart - a *Format Plot Area* window appears

- Under **Patterns**, set *Area* either to **None** or select white from the colours below. Note that a wide range of *Fill Effects* is also provided. Click on **[OK]** to get back to your graph

If you *double click* on the white *Chart Area*, a *Format Chart Area* window appears. This also has the *Patterns* tab and one in which you can set the default *Font* for all the lettering on the chart.

Further Data Manipulation

Naming Ranges

In the *Beginners' Guide* you saw how to name a single cell; you can apply the same procedure to a cell range and then use this in your formulae. A range of cells is also known as an array.

1. On the *PHOENIX* sheet, select cells *C2* to *C51* - click on *C2* then hold down **<Ctrl>** and **<Shift>** and press ***<down arrow>***
2. Click on the active cell name (*C2*) in the *Name Box* (just above the heading to column A) - the name will be highlighted
3. Type in the new name for the range, call it **`mass`**, and press **<Enter>**
4. Note that cells *D2:D51* and *E2:E51* have already been named **diameter** and **length** for you

Note that cells *D2:D51* and *E2:E51* have already been named **diameter** and **length** for you.

The next section demonstrates the use of named ranges in formulae and functions.

Using Ranges in Formulae and Functions

A wide range of functions is available within Excel. These can either be typed in directly, or can be invoked by clicking on the **[Insert Function]** button or by opening the **Insert** menu and choosing **Function...**. Examples include MAX and MIN (the largest and smallest values); SQRT; LOG/LN (logarithms); PI (ie 3.14159); SIN, COS and TAN (the trigonometric functions); and RAND (a random number). Note that all functions use brackets - PI and RAND need dummy brackets ().

1. Click on cell *H2*
2. Type in the formula **`=length*diameter`** and press **<Ctrl Enter>** (the result should be 37.35)
3. Check you are in cell *H2* (pressing **<Ctrl Enter>** keeps you in the same cell)
4. Copy the formula down the column using *autofill* (ie dragging the cell *handle* - the mouse cursor changes shape to a + when over the handle)

Tip: A quick method of filling down a column is to *double click* on the top cell's *handle* .

Note how, for each row, the value for the length is multiplied by that of the diameter. As an example of using named ranges in a function, work out average values for egg mass, diameter and length at the base of each column.

5. Move down to cell *A53* (use the *scroll bars*, **<Page Down>** or press **<F5>** and type in **`A53`**)
6. Enter the label **`Average`** then press ***<right arrow>*** *twice* to move to cell *C53*
7. In *C53* type **`=AVERAGE(mass)`** and press ***<right arrow>*** (the result should be 51.79)
8. In cells *D53* and *E53*, type **`=AVERAGE(diameter)`** and **`=AVERAGE(length)`** respectively (the values should be 7.84 and 11.21)

When using ranges in complex formulae, it is sometimes necessary to enter the formula as an *array formula*. You do this by holding down **<Shift>** and **<Ctrl>** as you press **<Enter>** to enter the formula. To illustrate this:

9. Move to cell *H53*
10. Type the formula **`=AVERAGE(length*diameter)`**

11. Press <**Enter**> and you will find Excel cannot do the calculation - a *#VALUE!* error message appears showing you that you haven't entered the formula as an array
12. Move back to cell *H53*, click at the end of the formula on the *Formula Bar* (the named ranges are shown in different colours), hold down <**Ctrl**> and <**Shift**> then press <**Enter**> again

The resultant value of 89.99 represents the average of (D2*E2) + (D3*E3) +... + (D51*E51). Note how Excel has surrounded the formula by { and } on the command line - you *cannot* type these in directly yourself.

Although using named ranges seems a user-friendly way of typing in formulae, they do not change when formulae are copied. Whenever you need a formula to change according to its position, the range should be specified explicitly. To demonstrate this, carry out the following:

13. Move to cell *C54* and type in the formula **`=AVERAGE(C2:C51)`**
14. Copy the formula into *D54* and *E54* simply by dragging C54's cell *handle* across the cells

When you release the mouse button, the formula will be copied across (giving the same values as in row 53). Move between the cells and note how the formula has been modified for each column; compare these formulae with those in row 53.

Tracing Links and Errors

As a speadsheet is developed, it becomes more and more difficult to see how the values are built up. Excel provides a very useful facility that clearly shows which cells are dependent on which and the source of an error.

First, *standardise* some of the data. Standardisation is a statistical technique that involves dividing each value by the average value for that set of data. It is often used to compare data which varies widely in scale - here, for example, mass ranges up to over 100, while the maximum length is just over 15. Dividing by the respective averages (51.79 and 7.84), produces values between 0 and around 2. These could then be plotted on a graph using a single y-axis.

1. Move to cell *I2* (press <**F5**>, type in **I2** then press <**Enter**>)
2. Enter the formula **`=C2/C53`** then *double click* on the *cell handle* to fill down the column

You will find that an error **(#DIV/0!)** occurs in all but the top two cells. To trace the error:

3. Move to any cell reporting the error (eg *I4*)
4. Open the **Tools** menu and choose **Formula Auditing** followed by **Trace Error**

The arrows show that *I4* is calculated from the values in *C4* and *C55* (which is empty - hence the *division by zero* error). Each of the values should have been divided by the average (in *C53*). You will see how to fix a value in a formula in a minute.

Other error messages which might appear include **#NAME?** (unrecognised text in a formula), **#NUM!** (invalid numeric values in a formula or function), **#REF!** (invalid cell reference) and **#VALUE!** (wrong type of argument or operand used).

Even where you haven't got an error, you can trace how a value is calculated by using **Trace Precedents** (another *Auditing* tool). Where you have built up a series of calculations, you can use this repeatedly to see how the formula in each cell is itself calculated. Another tool, **Trace Dependents**, shows you which other cells depend on a particular cell. Try this out on cell *E51*.

5. Move to cell *E51* (press **<F5>**, type in **E51** then press **<Enter>**)
6. Open the **Tools** menu and choose **Formula Auditing** followed by **Trace Dependents**

To turn off the tracing:

7. Open the **Tools** menu, choose **Formula Auditing** followed by **Remove All Arrows**

Fixing a Row/Column in a Formula (*Absolute References*)

As has just been seen, when entering a formula it is sometimes necessary to fix either the row or the column or both. In Excel, this is achieved either by using a named cell or by using a *$ notation* (the data range was fixed like this in the Chart Wizard).

1. Move to cell *I2* (press **<F5>**, type in **I2** then press **<Enter>**)
2. Correct the formula to read **`=C2/C$53`**

The $ notation fixes row 53 - had you also wanted to fix column C you would have used *C53*. When both the row and column are fixed, this is known as a *fully absolute reference*.

Tip: You can also fix a cell reference by pressing **<F4>** immediately after typing the reference when entering a formula. This circles round the possible notations each time the key is pressed.

3. *Double click* on the *cell handle* to copy the amended formula down the column - the error messages should disappear
4. Using the handle again, drag the formula in cell *I2* across to *J2* and *K2*
5. Complete the standardized data by *double clicking* on the *cell handles* in *J2* and *K2*

Move around the new cells and note how the formulae have automatically been adjusted for their relative position - the column reference varies while the row reference remains fixed

Freezing the Column Headings

When you have data which stretches down more than one screen, you lose the column headings as you scroll down. You can permanently display these heading by freezing them:

1. Move to cell *A2* then open the **Window** menu and choose **Freeze Panes**
2. Try using the scroll bars to move down the data - the column headings do not disappear
3. To return to normal working, open the **Window** menu and choose **Unfreeze Panes**

Freeze Panes can also be used to permanently display row headings - it freezes rows and/or columns immediately above and to the left of the active cell. To learn more about this, see the *Splitting Windows and Fixing Panes in Microsoft Excel* advanced notes.

Data Analysis

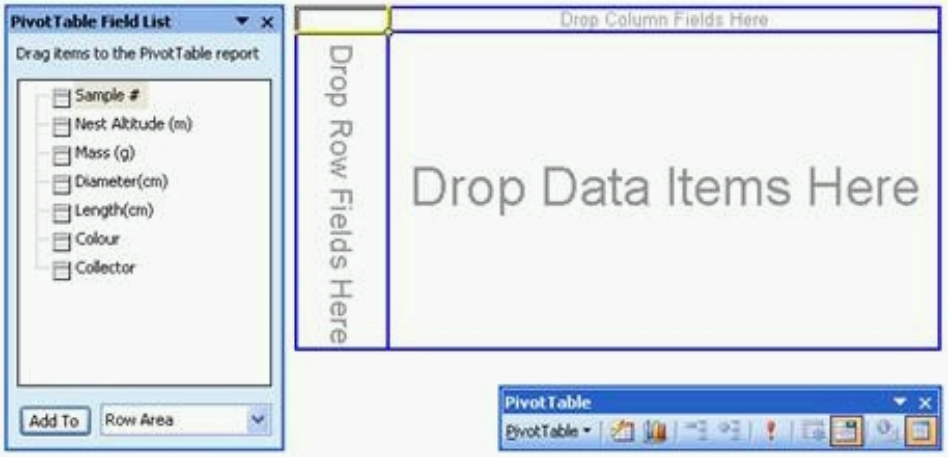
Although Excel is not a comprehensive statistical package, it does provide some rudimentary analysis tools - including Anova, Fourier Analysis, Regression and T-tests. For more thorough analyses of your data, a statistical package such as SAS, SPSS, Genstat or Minitab should be used. Consult the Applied Statistics Advisory Service for help with this. They can advise you on the suitability of Excel for your work and its limitations.

Pivot Tables

Pivot tables allow you to create interactive summary tables of your data. Here is a simple example:

1. Check you are using the *PHOENIX* data
2. Open the **Data** menu and choose **PivotTable and PivotChart Report...** - a *PivotTable and PivotChart Wizard* now runs
3. Step 1 of the wizard asks about the types of data source and report required. The defaults, **Microsoft Office Excel list or database** and **PivotTable**, are correct here so press **<Enter>** for **[Next>]**
4. Step 2 defines the data source - accept the named area *Database* (or type **`A1:G51`**) then press **<Enter>** for **[Next>]**
5. Step 3 asks where you want the Pivot Table to be placed - place it on the **Existing worksheet** in cell **`M2`** then press **<Enter>** for **[Finish]**

A skeleton Pivot Table is now drawn and the PivotTable toolbar and Field List window appears:



You now have to select which data series you are summarizing by dragging the field buttons from the *Field List* window and dropping them onto the Pivot Table skeleton. For this particular set of data, it's easy to summarize by collector and egg colour.

6. Drag and drop the **Colour** field button into the area marked *Drop Column Fields Here*
7. Drag and drop the **Collector** field button into the area marked *Drop Row Fields Here*
8. Drag and drop the **Collector** field button again, this time into the *Drop Data Items Here* area

The end result should look like this:

Count of Collector	Colour		
Collector	Blue-Green	Red-Brown	Grand Total
BN	1	8	9
BRV	6	3	9
CDS	9	1	10
FLC	1	9	10
RFA	4	8	12
Grand Total	21	29	50

Because you used a text field for the summary data, the default is to count the number of cells for each collector. Had you used a numeric field, the default would have been to sum the values. Various other measures are available, including average, maximum and minimum. You can have several of these showing by dragging further field buttons into the data area.

9. *Double click* on **Count of Collector** in cell *M2* - the following window appears:

Explore (but don't change) the list of measures in the *Summarize by:* box. Further options are available by clicking on **[Options>>]**. Use the *list arrow* attached to *Show data as:* to reveal these.

10. Click on **[Cancel]** to leave **Count** as the chosen option
11. Next select **Mass (g)** from the *Pivot Table Field List*
12. Using the *list arrow* to the right of **[Add To]** (currently showing *Row Area*) select **Data Area** then click on **[Add To]** itself - this is an alternative way of setting up your *PivotTable*

You now have both a count of the eggs for each collector and the sum of their mass. The results might look clearer if the two sets of figures are separated out:

13. Drag the **Data** row field heading in *N3* and drop it to the left of **Collector** in *M3*

If you change any of your data, the Pivot Table is not updated automatically. Instead, you have to use the **[Refresh Data]** button:

Refresh Data: 

14. Move to cell *G2* and change the collector from CDS to **`BN`**
15. Click on the pivot table (to activate it) then on the **[Refresh Data]** button - note how the values are updated (collector BN now has two blue-green eggs)
16. Reset the value in *G2* to **`CDS`** and update the pivot table again

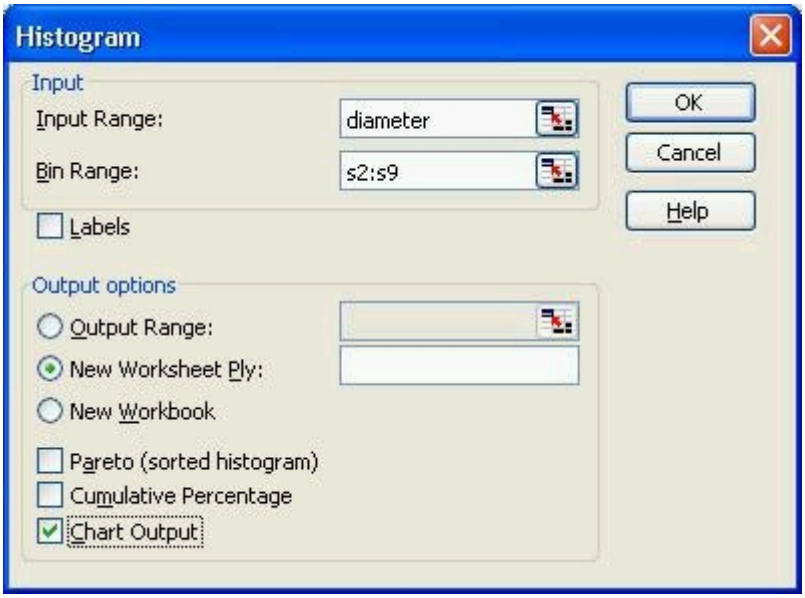
For more about pivot tables, see the *Using Pivot Tables in Microsoft Excel* advanced notes.

Histograms

Before running the histogram tool, you should first set up *bin values* (to define the divisions for each histogram range). As you will see, Excel can also draw a chart for you, if requested.

1. Starting in *S2*, type the number **`5`** - press **<Ctrl Enter>**
2. Drag the *cell handle* down to cell *S9* then click on the **[Auto Fill Options]** button which appears next to the bottom of the selection and choose **Fill Series**
3. Check you have the rest of the bin numbers: **`6`**, **`7`**, **`8`**, **`9`**, **`10`**, **`11`** and **`12`** in *S3:S9*
4. From the **Tools** menu choose **Data Analysis...** - if this isn't showing, click on **Add-Ins...** in the same menu and select **Analysis ToolPak** then click on **[OK]**
5. From the list of analysis tools, choose **Histogram** - press **<Enter>** or click on **[OK]**
6. In the *Histogram* window, set the *Input Range:* to **`diameter`** and *Bin Range:* to **`S2:S9`**
7. Under *Output options*, keep the default of output to a **New Worksheet Ply:** but name the sheet **`Histogram`**
8. Click on the **Chart Output** check box for a graph

The *Histogram* window should look as follows:



9. Press <**Enter**> or click on **[OK]** and watch what happens
10. Enlarge the chart by clicking on it (to activate it) then dragging the bottom handle downwards

The Histogram tool works very well but it can be annoying to have a *More* category. To remove this from the chart (and frequency table):

11. Select cells *A10* and *B10*
12. Open the **Edit** menu and choose **Delete...**
13. Leave *Delete* as **Shift cells left** - just press <**Enter**> or click on **[OK]**

The labeling on the X-axis could be improved further by editing the bin values in cells *A2* to *A9*:

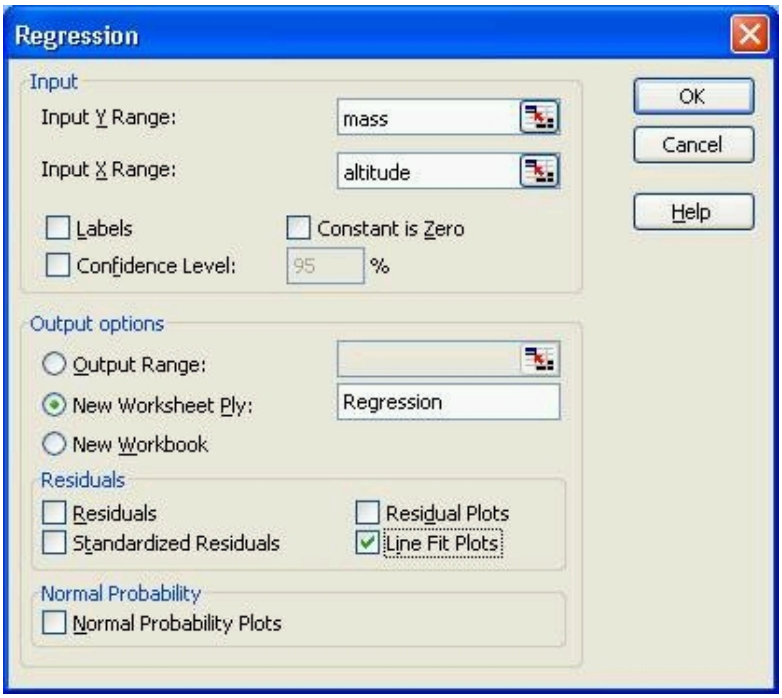
14. Move to cell *A2* and type **`<5`** then press <**Enter**>
15. In cell *A3* type **`'5-6`** (the single quote is vital - without it you get a date) and press <**Enter**>
16. Continue typing **`'6-7`**, **`'7-8`** ... **`'10-11`**, **`>11`** into cells *A4* to *A9*
17. Finally, move the chart onto its own sheet - click on it then open the **Chart** menu and choose **Location...**
18. Choose *As new sheet:* and name it **`Histogram Plot`** - press <**Enter**> for **[OK]**

Regression Analysis

Regression is a commonly-used statistical tool in the scientific world. If you are not familiar with this statistical technique, ignore this next exercise. The hypothesis you'll be testing out is that egg mass is dependent on the nest altitude.

1. Click on the *PHOENIX* tab to return to your data
2. Open up the **Tools** menu and select **Data Analysis...**
3. Select **Regression** from the list of *Analysis Tools* - press <**Enter**> or click on **[OK]**
4. In the *Regression* window, set **Input Y Range:** to **`mass`** (or type **`C2:C51`**)
5. Press <**Tab**> then in **Input X Range:** type **`altitude`** (or type **`B2:B51`**)
6. Set *Output options* to **New Worksheet Ply:** and name the new sheet **`Regression`**
7. To get a graph of the results, turn on **Line Fit Plots** under *Residuals*

The *Regression* window should look as follows:



8. Press **<Enter>** or click on **[OK]** and watch what happens

The regression results are displayed on a new worksheet, with a graph plotted at the side.

9. Move the chart onto its own sheet - click on it then open the **Chart** menu and choose **Location...**
10. Choose *As new sheet:* and name it **`Regression Plot`** - press **<Enter>** for **[OK]**

To add a best-fit line to your graph:

11. *Double click* on one of the (pink) *Predicted Y* values then, on the *Patterns* tab in the *Format Data Series* window, set *Line* to **Custom** and *Marker* to **None**
12. Press **<Enter>** or click on **[OK]**
13. Click away from the *Predicted Y* series and the line will be properly displayed, without the individual values showing

Note that Excel does allow rudimentary multiple regression - set up the *Input X Range* to a block covering more than one adjacent column. However, your data may not be suitable for this so always seek guidance from the Applied Statistics Advisory Service.

Pasting Charts and Worksheets into other Documents

If you want to incorporate information from Excel into Microsoft Word (or other software), the procedure is as follows:

a) *Worksheets* can be copied across simply by using **[Copy]** and **[Paste]** - in Excel, select and copy the cells you require then switch to Word, move the insert position if required and paste in the data. In Word, the data will be displayed as a table with any formatting preserved. Such tables can be edited as normal text. Note, however, that with certain packages some Excel formats cause problems and have to be cleared before the copy/paste is carried out.

If **Paste Special...** is used instead of **Paste**, then the data can be moved as *unformatted text*. This does not appear as a table but is instead *tab separated*. Using **Paste Special...** you can also copy a worksheet as a *picture*. Note that the data in a picture of the worksheet cannot be edited.

If you want to preserve a link between the Excel worksheet and other document (such that changes made to one file are matched in the other) then turn on **Paste link** in the *Paste Special* window.

Worksheets can also be saved in various formats readable by many other applications including space/tab/comma separated values. You can also save as an earlier version of Excel. The format type can be set in the *Save As* window.

b) *Charts* can also be copied to a Word document (or other software, such as PowerPoint or Access) by using **[Copy]** and **[Paste]**. When pasting into Word, it is advisable to use **Paste Special** from the **Edit** menu, as this gives you the options of pasting the chart as a picture or as a link (so that changes to the chart made in Excel are automatically reflected in the Word document).

Changes can be made to a pasted chart by *double clicking* on it (Excel will automatically be loaded). Note that if your original chart was not on a separate chart sheet, the copy in Word will be put on one. By moving to the worksheet, you can change the data to alter the graph. Move back to the chart sheet before returning to your document.

For details about copying Excel charts to Word see the document *Graphics in Microsoft Word*.

Leaving Excel

To quit from Excel:

1. Choose **Exit** from the **File** menu
2. Click on **[No]** when you are asked whether you want to save the changes to *phoenix.xls* - there's no need to keep the changes

Finally, close down your session as usual.

™ Trademark owned by Microsoft Corporation.
© Screen shot(s) reprinted by permission from Microsoft Corporation.

Page last updated Friday, 17 August 2007
Email: its-help@reading.ac.uk
© The University of Reading