

Using Microsoft Excel - an introduction

by Erling Strandberg and Birgitta Malmfors

Dept. of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden

This brief guide focuses on some basics in handling data and doing calculations and matrix operations in Microsoft Excel. It also includes making graphs. The advice given is based on the authors' own experience of using this software. The software versions referred to are mainly MS Excel 2002 and 2003, but much of the advice is applicable also for earlier versions.

Enter data or text into cells

Mark the cell you want to write in (click on it) and write the text. Leave the cell by pressing Enter or the down or right arrow, or by clicking in another cell.

Refer to data in cells

The cells are usually referred to by their column and row identifiers, e.g., the first cell (upper left corner) is called A1. To refer to several rows within a column, write e.g., A3:A20. To refer to a set of rows and columns, give the cell in the top left corner and the one in the bottom right corner, e.g., A3:E20. To refer in Sheet 1 to cell B4 in Sheet 2 of your Excel workbook, write: =sheet2!B4. If you have named sheet 2, use that name instead.

Sort data

Mark the data you want to sort, open the Data menu and click Sort; decide what column(s) to sort by, and in what order (ascending, descending). Make sure you sort the whole area you want sorted!

Name cells

Sometimes it's useful to give cells meaningful (but short) names so you can use the name instead of the column/row identifiers when referring to the cell in calculations. The simplest way to name a cell is to click on the cell, and thereafter click the Name Box that you find just above the A-column (where it says A1 if you are in cell A1). Write in the box the name you give the cell and press Enter. Another way to do this: Insert menu – Name – Define – Write the name – Click OK

You can also name a field consisting of many cells by marking them with the mouse and then writing the name you want for it in the name box just as you did for one cell.

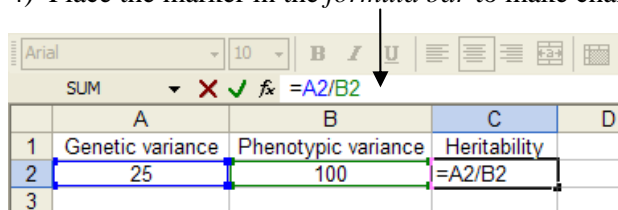
Note! Excel does not differ between lowercase and uppercase letters, so you can use whatever type you like, but you cannot name one variable as "name" and another one as "NAME". They will be identified as the same variable.

To *delete* a name entered: Insert menu – Name – Define – Mark the name – Click Delete.

Enter equations, to do calculations

You can programme Excel to do a lot of calculations for you. Just give the equation for what you want the program to do.

- 1) Start every equation by writing an equals sign: =
- 2) Write the equation (e.g. A2/B2)
- 3) Finish by pressing Enter (Return)
- 4) Place the marker in the *formula bar* to make changes in an equation.



	A	B	C	D
1	Genetic variance	Phenotypic variance	Heritability	
2	25	100	=A2/B2	
3				

Addition: +
Subtraction: -
Multiplication: *
Division: /
Raise value to a power: ^ (e.g.: ^2, ^3, ^4)

To sum values in rows 3 to 15 within column B, for example, click on the cell where you want the sum and write: =SUM(B3:B15) and press Enter. You can also mark the cells and click the Σ icon.

If you want to divide the value in cell B3 with the value in B17, and get the answer rounded with two decimals, write: =ROUND(B3/B17,2). Note! Some versions of Excel use *semicolon* here instead of comma. You can also do it this way: write =B3/B17 and press Enter. Mark the cell, click Format menu - Cells – Number - choose number of decimals. If you want to raise a value in cell D6 to the power of 3, click the cell where you want the result and write: =D6^3.

You can refer to *cell names* instead of cell positions. Assume that you have given a value for additive genetic variance in A2 and a value for phenotypic variance in B2. Now, if you name A2 as varA, and B2 as varP (and if you want, C2 as herit) you can calculate the heritability in cell C2 by writing =varA/varP. Doing so means that you don't need to keep track of the cell numbers when you write the equations; the equations are also more informative. Note, you cannot give a name "h2" because that is in conflict with the cell identifier "h2" (and "H2").

	A	B	C
1	Genetic variance	Phenotypic variance	Heritability
2	25	100	0.25
3			

Functions

The feature "Function" in Excel uses predefined formulas that perform calculations. There are numerous mathematical, statistical and other functions available. You can see all of them if you click the Function icon [f_x]. You can also go to the Insert menu – Function.

To get the square-root of a value, for example: write the value, click on the cell where you want the result, click on the f_x icon and select SQRT from the menu of the category *Math & Trig*, give the cell identification for your value, e.g. A7, and click OK. Instead of using the f_x icon you could have written the equation yourself as: =SQRT(A7), which is the faster way. If you would want to get the standard deviation of values in rows B4 to B75, click where you want the result, click the f_x icon, select STDEV from the menu of the category *Statistical*, write B4:B75 in the upper box (or mark those cells), click OK. Use Excel **Help** for further advice on using the Function feature.

Note! It can be useful to look at the function names shown in the Insert Function window, and also to test some of the functions. You then get ideas on how to express things when you write your own equations. You see the formula resulting from using a specific function command if you look in the formula bar to the right of the f_x icon.

Cut, copy and paste

Very similar to a word processing program - mark the cells you want to cut/copy, click on the appropriate icon (or do Ctrl+X for cut and Ctrl+C for copy) and then go to the cell where you want to paste the cell(s) and click on the paste icon (or Ctrl+V).

If you want to copy a value in a cell to many cells in the same column or in the same row, there is a simpler way. Click on the cell and then move the mouse pointer to the lower right corner of that cell (it then changes from a thick white cross to a thinner black cross). Then hold down the left mouse button while dragging downwards (or to the right) and the value will be copied to all cells you drag through. This is called "filling down" or "filling right".

Create a series

Often one wants to create a series of values, e.g. 1, 2, 3, ... 100. Writing them all is tedious and you are bound to make mistakes. A very simple way to do this is to write the first two values 1 and 2 in the first two cells, mark these cells and place the mouse pointer in the lower right corner (it changes to a black cross). Thereafter, drag downwards as many cells as you want (filling down). Excel will automatically fill in the appropriate values in the series. You can create any series; writing 1980 and 1985 in the first two cells, for example, will give the series 1980, 1985, 1990, 1995, etc.

Using "fill down" with functions and fixed references

Often one may want to create a "table" with values in the first column and a function of these values in the following columns. Let's say that you have various heritabilities in the first column and want to know how large the genetic variance is for each heritability (h^2), given a certain phenotypic variance. Then you could create the table as follows:

First create the series of h^2 -values you want, say write 0.1 and 0.2 in A5 and A6, and fill down to make a series to 0.9. In B4 you can write the value for the phenotypic variance. In B5 (corresponding to h^2 of 0.1) you write the equation = A5*B4 (which means that you multiply h^2 with σ_p^2 to get σ_A^2). Now, in order to be able to fill down and copy this equation to all cells below, but still always be referring to B2, you need to *fix* the row pointer of that reference so it always points to row 2. To do this you put a dollar sign (\$) in front of the "4" in B4 when you write the equation; it then will be: =A5*B\$4. Now you can fill down from cell B5 and get the correct values in all cells. (If you would not have fixed the row pointer to B4, the following multiplications would have been done when you were filling down: A6*B5, A7*B6 etc, which does not give the series you wanted).

If you want another column where you start from another phenotypic variance you just write that value in C4, fix the column reference to A5 in cell B5 (i.e. change from =A5*B\$4 to =\$A5*B\$4), fill right from B5 to C5, mark the cell C5 and then fill down from it.

These tricks with fixing references with "\$" may sound complicated when written in text, but if you just try it a couple of times you will figure out which part to fix in what situation. Just remember that you write the \$ sign in front of what you want to be fixed, i.e. the column letter and/or the row number. If you want to fix both column and row (e.g. \$A\$5), a quick way of getting that when you write the equation is to write A5 (or click on the cell) and then press F4. Pressing F4 again will give A\$5, and once more will give \$A5.

Filling down or filling right can be very useful if you want an equation to be valid for data in a number of columns or rows. In this situation you will not fix any columns or rows. Say that you have entered data for different traits in columns B, C, D and E, and that you in two separate rows in column B give the equation for calculating the mean and the standard deviation. Thereafter you just need to mark the two cells in those rows, place the mouse pointer in the bottom right cross and pull to the right. This will give you the calculations also for columns C, D and E.

Matrix functions

You can also handle matrices in Excel:

Start by writing the individual values of your matrix in a set of columns and rows.

To refer to the matrix: mark the area and name the matrix, or refer to its position, e.g. A3:D6.

Note! When writing a matrix function: always start with = and finish by pressing **Ctrl+Shift+Enter** (i.e. press the three keys simultaneously)

Transpose

If you want the transpose of a matrix: mark the cells where you want the transpose (make sure to mark the correct number of rows and columns, and in the correct direction), write =TRANSPOSE(A), where A is the reference to the matrix you want the transpose of (i.e. the name you have given that matrix, or a description of its location, e.g. C1:D2). Press Ctrl+Shift+Enter.

Matrix multiplication

To multiply two matrices with each other, say A with B: mark the correct number of cells where you want the result, and write =MMULT(A;B) and then press Ctrl+Shift+Enter.

Note! Instead of semicolon, you here use a comma in some versions of Excel.

Matrix inversion

Mark the cells where you want the result and write MINVERSE(A), where A is the name of the matrix. Press Ctrl+Shift+Enter.

Borders and Highlighting

You can make *borders* in your Excel sheet; it might be lines or rectangles, | — □ to separate different parts of the sheet or indicating a matrix.

To do this: Mark the cells – click the down arrow at the Border icon – choose border.

You might also want to fill some cells with *colour*. You then do like this:

Mark the cells – click the down arrow at the Fill Colour icon – choose colour.

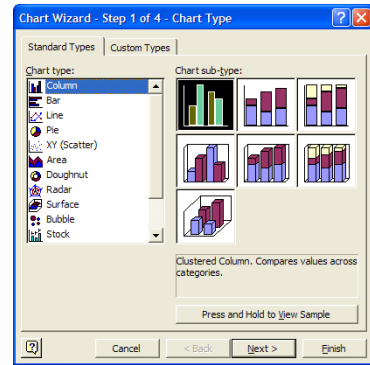
If you don't find this icon among the other icons in the Formatting toolbar at the top of the screen, click the drop down menu to the right of that toolbar and tick the Fill Colour option.

Graphs

You can create different types of graphs in MS Excel, e.g. Column and Bar charts, Line and Scatter charts, Stock charts, and Pie charts. You see all chart types available if you open the Chart Wizard.

Create a graph

In the Excel worksheet, *organize your data* in columns or in rows, including also column/row headings. Note, it can be useful to have the variable/category that is to be on the X-axis in the first column (or row), because for some types of charts that is assumed. For more details on arranging the data, search “create a graph” in Excel Help.



To create a graph from the data you can do like this:

Mark the data you want to be included in the graph (including column/row headings).

If you don't want all columns (or rows) of the data to be included:

Hold down the Ctrl key and mark the parts of the data that should be included.

Click the *Chart Wizard* icon (or click Insert menu, and then Chart)¹.

Thereafter four steps follow:

1. Chart type

Select *Chart type* and *Sub-Type*. Read the explanatory text that appears when you click a sub-type chart. If the sub-chart text only explains whether you get lines with markers, for example, click the sub-chart shown in the upper left corner to read about the *function* of the chart type.

Note that the Line chart is used to show trends in data over *categories* or *time* (assuming that there are equal distances between the time events); the X-axis is here categorical. The Scatter chart, on the other hand, is used to show relationships among *numeric values* in two (or more) data series; the x-axis is here continuous. The scatter chart displays pairs of x- and y-values, either as single data points, or as lines connecting the points.

Note also, it is much easier to see the height of a column (and read its value in the y-axis) in a plain column chart than in one with three-dimensional visual effects. The situation is the same for a bar chart and a pie chart. We strongly advise against using graphs with 3-D visual effects, unless your data would be three-dimensional.

You can read about more about chart types in Excel Help.

2. Chart Source Data

Data Range (the tab to the left): Sometimes Excel is “too smart”, it tries to predict what you want to have as a series, and where your categories or continuous x-axis values are. If it guesses wrong, it's usually helps if you in the Data Range tab indicate (tick) whether your data (series) are organized in columns or in rows. See the result for the graph shown in the window. If the graph seems to be correct, click Next to go to step 3. If the graph is not correct, have a look at the tab Series (see below).

Series (the tab to the right): If the graph or the labels are not correct then you can try the following: Check in the box “Series” if possibly the variable for the X-axis appears there, and if it does, mark it and click the Remove button. Have a look at the graph.

If the labels for the X-axis are not correct (e.g. in a line chart or a column chart), place the marker to the left in the box “Category X-axis labels” – Go to your data set – Mark the X labels (don't mark the column/row heading). The correct labels for the X-axis should now appear in the diagram.

If you don't manage to get a graph that looks correct, check that your data were organized the way they should be. Or, try another type of graph; maybe you happened to choose Line chart where you should have chosen Scatter chart, or vice versa.

¹ If you have Excel 4 the pointer here changes to a small cross with a diagram attached to it. Draw a square where

you want the graph to appear (hold the left mouse button down while doing this)

3. Chart Options

There are several tabs in Chart Options. The most common thing to do here is to click the tab Titles and write the titles for the X- and Y-axes. Remember to give the measurement unit for the Y-axis, and when relevant also for the X-axis. You might also give the chart a title, but you normally don't do that if the graph is to appear in a text and have a figure caption. If you don't want gridlines in the graph, click the Gridlines tab and delete the tick for gridlines; you can also delete the gridlines later if you want to see the graph before you decide on gridlines.

For some graphs you might want to use the options for Data Labels. For example, in a pie Chart it can choose to outside each "pie-piece" show its name and value or proportion (%) of the total. That is achieved by ticking "Category name" and "Percentage" in the Data labels tab. When you use this option, remember to un-tick "Show legend" in the tab Legend.

4. Chart Location

Decide (tick) where in your Excel Workbook you want to place the graph. Click Finish.

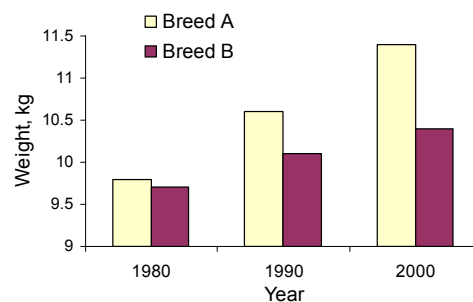
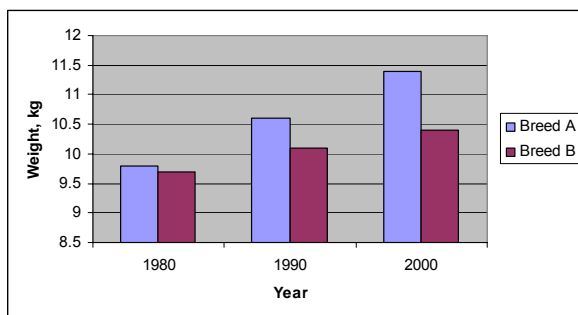
Format a Graph

The graph you get after having done steps 1-4 above usually gains from some additional formatting.

The principle for formatting a chart is to:

Right-click a specific part of the graph – Shortcut menu opens – Click an item in the menu – Formatting window opens – Make desired changes. Right-click the next item you want to format ...

Below you see two figures showing the same column chart. The figure to the left shows the chart you get after the four steps described above to create a chart (and keeping the tick for gridlines and legend). The figure to the right shows the column chart after doing some formatting of the layout.



To format the left figure to get the figure to the right you do like this,

right-click:

- the white chart area: click Format Chart Area in shortcut menu; in tab Patterns, tick Border none
- the grey plot area: Format Plot Area; tab Patterns, tick Border none and make Area Colour white (or Area none)
- a gridline: click Clear
(if you want to retain the gridlines but make them less dominant: Right-click a gridline and click Format Gridlines; tab Patterns, choose Style (e.g. dotted line) and Colour (e.g. grey))
- the first pile: Format data series; tab Patterns, change Area Colour (to better see the difference between the breed columns when the graph is printed in black and white)
- the y-axis title: Format Axis Title; tab Font, select font 12, and change bold to regular
Do the same for the X-axis title
- a year in the X-axis: Format Axis; tab Font, select font 11
- a value in the Y-axis: Format Axis; tab Scale, change minimum to 9 and maximum to 11.5
- the legend: Format legend; tab Patterns, tick Border none; tab Font, select font 12

Furthermore:

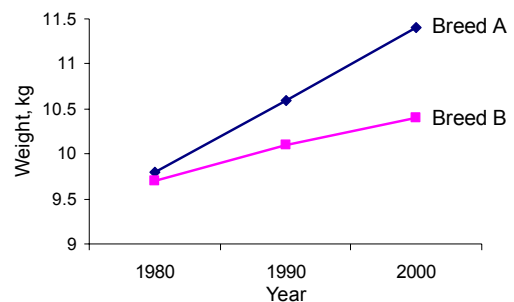
- Place the marker at the legend – hold down left mouse button and *drag* the legend into the graph
- Place the marker at the X-axis title – hold down left mouse button, drag the title upwards a little.

If you want to get back to one of the steps 2-4 that you did when you created the chart:
Right-click the chart area – Shortcut menu opens – click the name of the step you want to go to.

You might think it seems laborious to format a graph; actually, it is done quite quickly once you know how to do it.

You can also change the graph you produced to another type of graph (you do it in the Excel sheet):
Right-click the chart area – Shortcut menu opens – click Chart Type – select chart type.

We have performed this operation with the column chart that was formatted (but copied it first to maintain also the column chart). We chose Line chart² with markers displayed and got a figure like the one shown here, but some further formatting was also done:



- The thickness (weight) of the lines was increased:
Right-click line; Format Data Series; tab Patterns, increase Weight (marker size is then adjusted, but you might reduce it so markers don't dominate).
- The figure legend was deleted and breed definitions were written in text boxes at the end of each line.

To use a Text Box in Excel:

Mark the chart – click the Text Box icon in the Drawing toolbar³ – go to the graph and click where the text box should appear approximately – write the text – click the frame of the text box – hold down the left mouse bottom – drag the text box to where you want it (for better precision, press the Alt key while you drag the text box).

Copy a graph into MS Word or MS Power Point

Copy the graph in Excel – Go to your MS Word or MS PowerPoint file – Paste the picture into the file using one of the following two alternatives:

Alt 1. Edit – Paste Special – Picture (Enhanced Metafile), or Picture. The graph is entered as an object that can be formatted, i.e. you can change size, brightness and contrast, and you can crop it. When entered into PowerPoint you can also recolor the graph, and ungroup it to modify individual parts. In Word you can choose style for text wrapping (e.g. “Square” or “Top and Bottom”) and also set distance to the text.

Pictures entered as described in this alternative display quickly, and are usually easy to handle. If you would want to edit the data behind the graph, you do that in the Excel file where you produced it (assuming you saved the file), and make a new copy to replace the previous graph.

Alt 2. Edit – Paste Special – MS Excel Chart Object. A graph entered like this can be formatted (as described above); it can also be edited as if you were still in your Excel file (the full Excel workbook is copied into the Word or PowerPoint file). However, the picture might not be so easy to handle, and it will take more disc space (KB) compared to Alt 1. Note: if you use Paste (instead of Paste Special) the graph will be entered as when using Alt 2.

Help



You will learn a lot more about MS Excel if you use the Help function. Search for a specific item or explore the contents and all features that are available.

² Note: here a Line chart is OK because we have equal distance between the years, if we e.g. had years 1980, 1981, 1990, we would choose scatter plot instead, otherwise the graph will give the wrong perception.

³ If the Drawing Toolbar is not displayed in your Excel sheet, click View menu, select Toolbars and click Drawing. Usually this toolbar is displayed at the bottom of the Excel sheet.