How to use the regression tool in Microsoft Excel
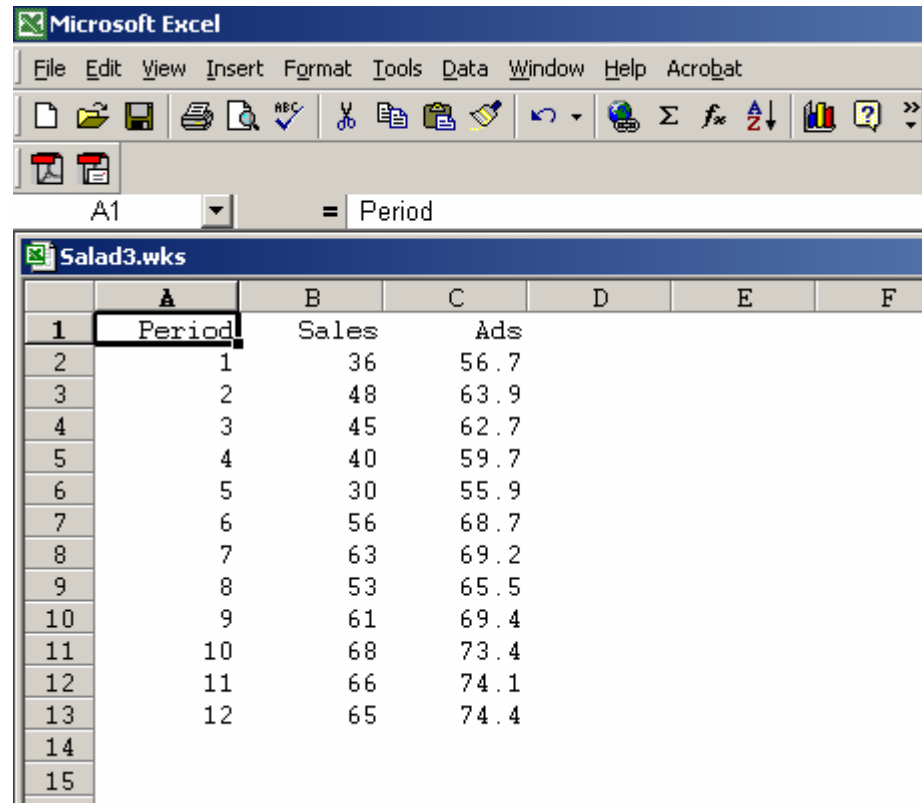
Consider the following simple (fictitious) example. A firm has recorded the sales and advertising figures for the last twelve months and set up a simple Excel spreadsheet as shown below in Figure 1.

	A	B	C	D	E	F
1	Period	Sales	Ads			
2	1	36	56.7			
3	2	48	63.9			
4	3	45	62.7			
5	4	40	59.7			
6	5	30	55.9			
7	6	56	68.7			
8	7	63	69.2			
9	8	53	65.5			
10	9	61	69.4			
11	10	68	73.4			
12	11	66	74.1			
13	12	65	74.4			
14						
15						

Figure 1   (Fictitious) Sales and Advertising data

It is believed that there may be a linear relationship of the form

Sales = a + b Ads + error

where a and b are constant parameters.

We expect b > 0, that is as advertising increases this will generate additional sales. We can't really say much about the size of the parameter a *a priori*. A negative value would indicate that until advertising reached at least some threshold level (-a/b) then Sales would remain zero.

Before running a regression lets first check that the assumption of a linear relationship is a reasonable one to work with by plotting an X-Y graph, or scatter plot of sales against ads (see Figure 2).
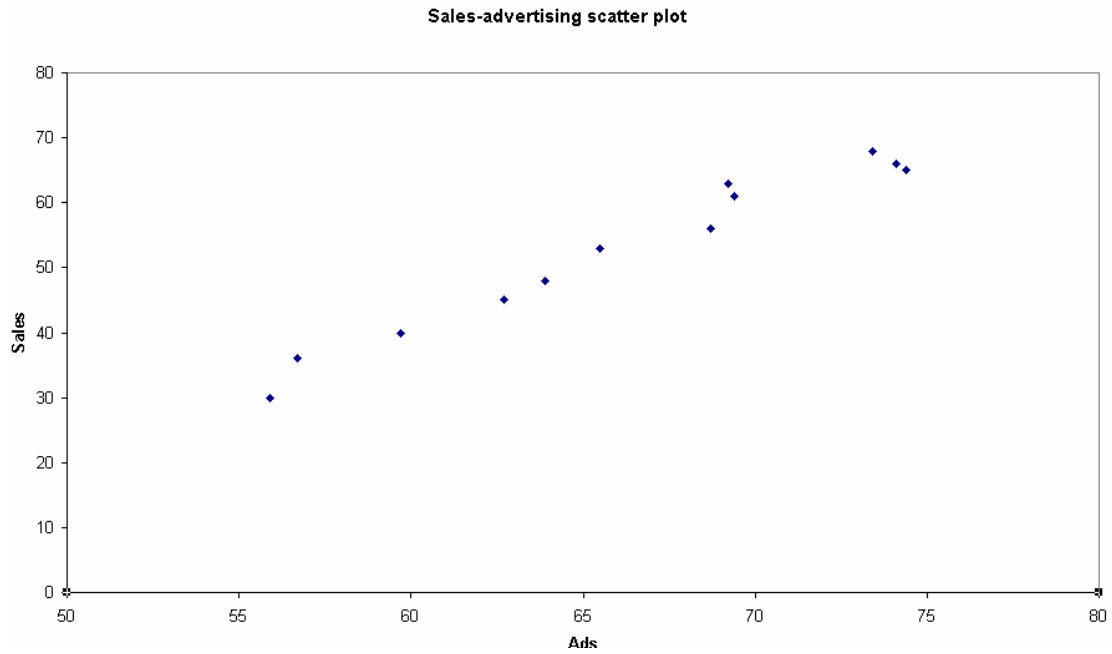
Figure 2 Sales-advertising scatter plot.

(Note: You can create a graph like this in Excel but the default will include grid lines and an X axis right back to the origin. I have edited the graph a bit to get it to look like this).

The scatter diagram would appear to confirm that we can proceed with a linear model. How can we do this automatically in Excel?

Microsoft Excel makes available a tool for running least squares regressions.

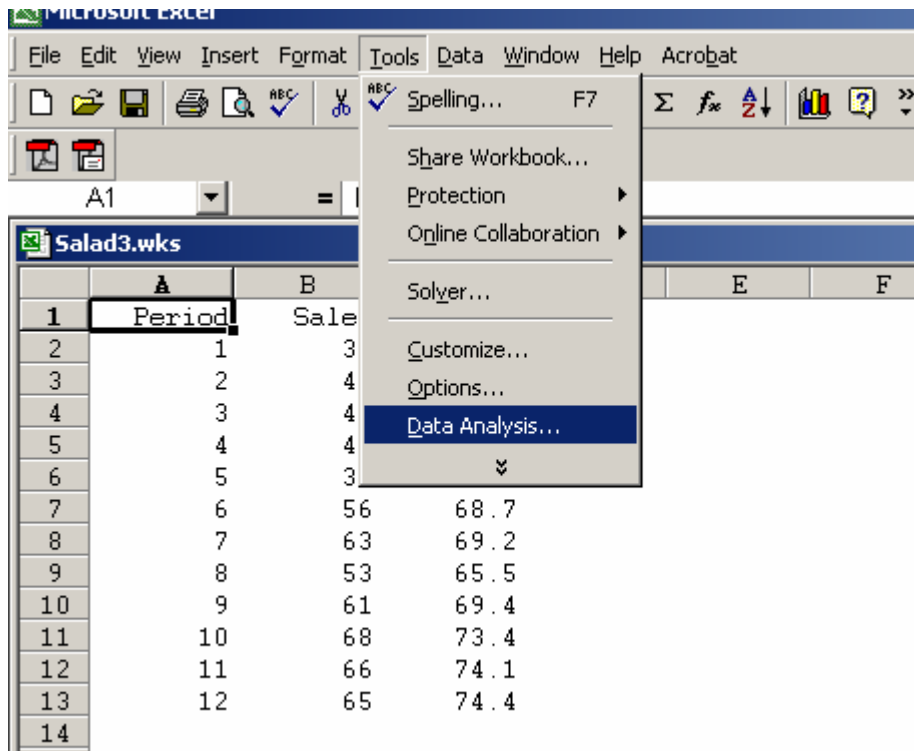Select Tools from the main menu and then Data analysis – see figure 3.

Figure 3. Where to find the Data Analysis tools in Excel.

If the Data Analysis tool does not appear it may be that you need to install it – see the Appendix.

Now choose Regression from the list of Analysis tools shown in the Data Analysis box (see Figure 4)
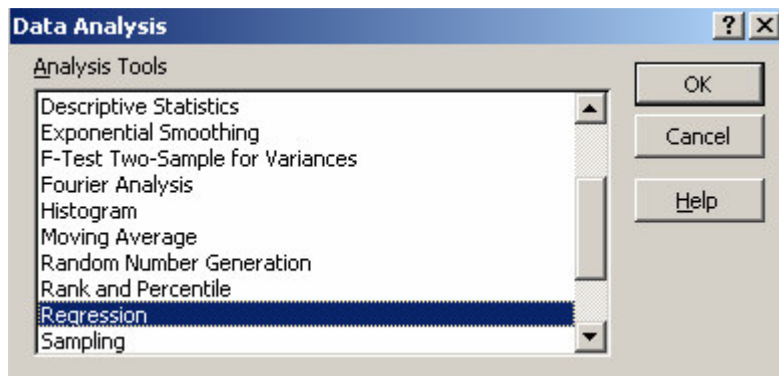

Figure 4 Data Analysis box

The Regression dialog box will open and you need to input the Y range (here B2 to B13) and the X range (here C2 to C13) – you can just use the mouse to select the block. (See Figure 5). You will notice that there are some other options that you could select here but as it is the first time you have used the regression tool just ignore them and click OK.

The standard regression results from Excel should then appear on a new worksheet (see Figure 6). Note: I have extended the width of some of the columns to make the output more readable.
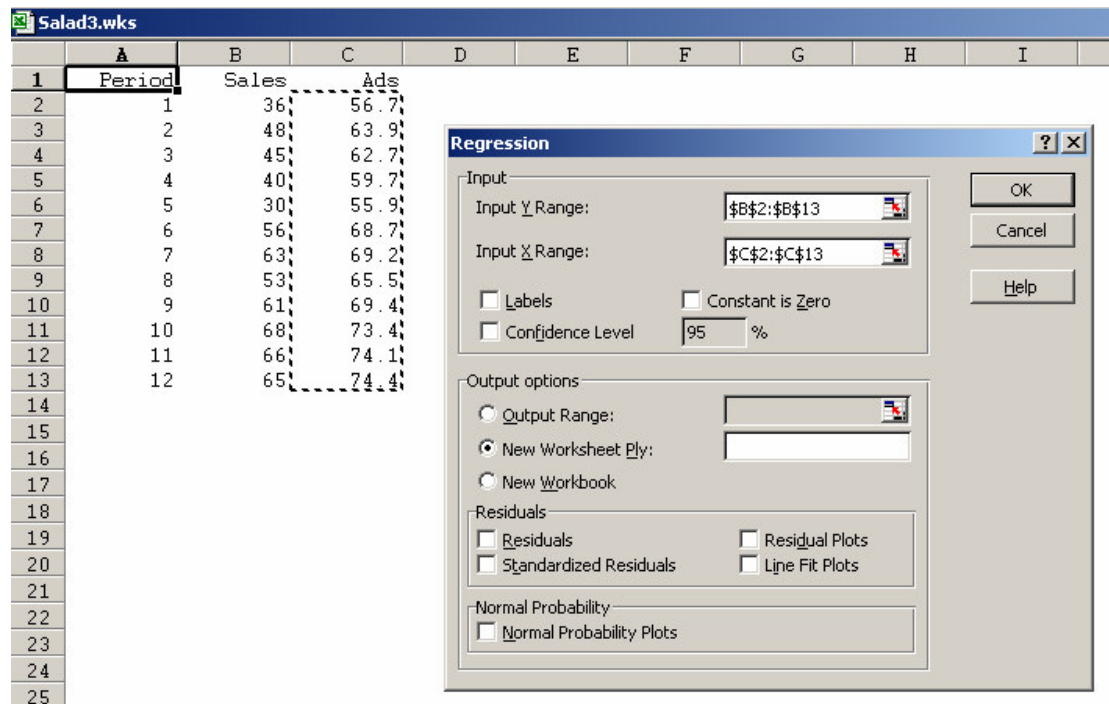

Figure 5 The Regression dialog box.

Figure 6 The basic regression results from Excel.

It is not our purpose here to interpret all the results provided – that will be dealt with on another handout. For the moment lets just pick out the values given for the Coefficients. The number for the Intercept coefficient (a) is close to –75, while the coefficient for the X Variable (b) is approximately 1.929.

So the estimated equation is

Sales = -75 + 1.929Ads + residual

(the residual is the difference between the actual value of sales in any period and the value predicted by the model).

Appendix. What to do if the Data Analysis tools don't appear.

If you can't see the Data Analysis option in the Tools menu, select Add-Ins from the menu. The dialog box below will open and you should check the Analysis ToolPak option. (see Figure 7). Note: Various Add-Ins are not automatically loaded each time with Excel to save the memoru overhead.
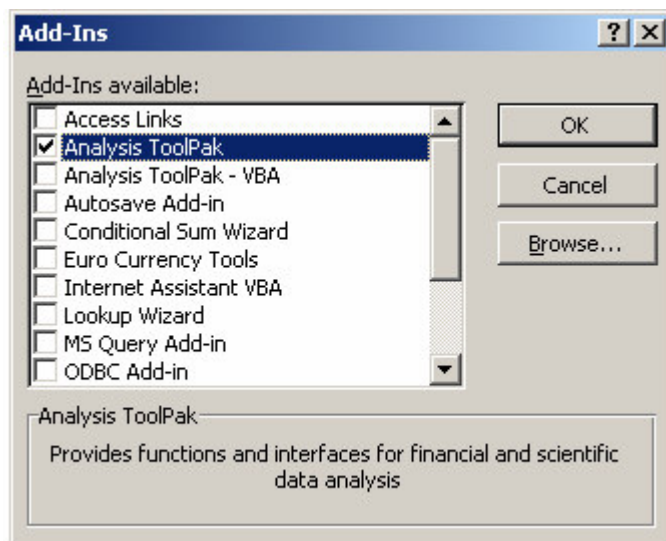
Figure 7 Enabling the Analysis ToolPak via the Add-Ins dialog box.

Guy Judge
February 2005