 How-to guide: regression analysis in Excel (Office 365)

Part 1 – making a scatterplot

1. Ensure your data is in two columns (or rows) in an Excel workbook. The data can be next to each other or separated by other columns.
	* Note: Ensure your independent variable is first (on the left). When Excel constructs a scatterplot, it will automatically put the first (left) variable on the horizontal axis.



1. Highlight your two rows of data.



1. Select ‘insert’, then ‘scatter’. Click on the first graph type.



1. The scatterplot will be produced.



1. You can now edit the chart. For example, you can add a chart title, axis labels, and the scale on each axis.



Part 2 – calculating Pearson’s correlation coefficient (r)

1. Ensure your data is in two columns (or rows) in an Excel workbook. The data can be next to each other or separated by other columns.



1. Click on the cell you want the correlation coefficient to be in. In this case, cell B1 has been selected.



1. To start a formula, type “=” then “correl(”



1. You will now have to select the data. Select one column, type “,”, then select the second column.



1. Close the formula with end brackets “)” before pressing enter to end the formula. Pearson’s correlation coefficient will now be in the selected cell (B1 in this case).



Part 3 – calculating the formula of and adding a least squares regression line

1. You will need to edit the Microsoft Excel 365 file in a standard version of Excel.



1. To add a trendline, right click the scatterplot and select ‘Add Trendline’.





1. To obtain the equation of the line, right click the line, select ‘Format Trendline’ and then select ‘Display Equation on Chart’.





Part 4 – forecasting with Excel

You may need to adjust the data range if you wish to forecast a value outside of the original data range using the graph. For example, in the ‘Foot Length and Height’ graph, we may want to predict the height of a person with a foot length of 40cm.

* To forecast, right click the line, select ‘Format Trendline’, then input the required number of periods. For example, a value of 10 will extend the line past 40 as its maximum value is currently 31.



