# Excel Practical 2: Using the Data Analysis package

Start by going to Tools > Data Analysis to have a look through the options, and explore the functionality.

Tasks:

1. Go the sheet with data set 1 in it and creating a set of bins (about 5 should be fine - see class 1 slides 29 and 30 if you’ve forgotten what these are), and put these in a line of cells. The bins should cover the range of the data. Now choose Tools > Data Analysis > Histogram, and input the data and the bins. Make sure to choose the ‘Chart Output’ option. Click OK and then tidy up the plot to your satisfaction.
   1. See if you can interpret the results. Can you guess at the sample mean from the histogram? Is there any skew
   2. Try changing the bins and looking at the effect on the histogram.
2. Now choose Tools > Data Analysis > Random Number Generation. Choose 1 variable and 10,000 random numbers. Pick the normal distribution option and leave the values as mean 0 and standard deviation 1. This will create 10,000 samples from a standard normal distribution. In class 2 we learnt that approximately 68.2% of samples from a standard normal will be between -1 and 1, and that 95.5% of samples will be between -2 and 2.
   1. Think how you might verify this with your list of 10,000 samples.
   2. Try repeating the exercise with 100 random samples, or 100,000 random samples and seeing how close you get to 68.2% and 95.5%
3. (Harder) Go to data set 3. Use the function Z.TEST to see if the mean of run 1, fortification level 1, is equal to 0.91 or not across replicates. (Note that this function returns only the p-value). See if you can re-create this p-value by running the test by hand (i.e. not using Z.TEST) by following the instructions in Class 3 slides 21 and 22. It also helps to read the help files on the Z.TEST function (we want the two-tailed version of the p-value).