

## Excel Practical 4: correlation, linear regression, and Control Charts

---

Note: if you want to go back and cover more of the previous practical sheet, please do so

Tasks:

1. Go to Data Set 2 and re-create your scatter plot all the way back in Practical 1. There are two ways to calculate the correlation in Excel. One is to use the CORREL function, and the other is to choose Data Analysis > Correlation. Use both and check that they come up with the same answer. Try to interpret the correlation value. Does it match the relationship you can see from the plot?
2. Run a linear regression with Breeding Value as the response (Y) variable and Expression as the explanatory (X) variable. Make sure to tick the Confidence level box, and the residuals box. Try and work your way through the output. You should be able to find:
  - a. The correlation between the two variables
  - b. The residual standard error (i.e. how variable the residuals are)
  - c. The best value of the slope and the intercept (hint: look at the coefficients column)
  - d. Confidence intervals and t-tests with associated p-values for whether the intercept and slope are significantly different from 0 or not

See if you can interpret all these values. Much of the output is pretty useless so can be ignored but there's some useful stuff in there

3. Open the other\_data\_sets.xlsx file and look at the control 1 sheet.
  - a. Re-create the x-bar chart in slides 31-32 in Class 6
  - b. Re-create the R-bar chart in slides 35-36

For each try adjusting the data to see what kinds of values make the data go 'out of control'