EOSC 510/410 Assignment 1:

Note: Please do **not** submit your code; only submit your assignment as a PDF with the figures/results/tables embedded inside the document. **Include your name(s) in the document name**, e.g. Assignment1_Anderson.pdf

Please submit the assignment to the TA (Geena Littel): glittel@eoas.ubc.ca

Problem 1 *[total of 7 points]:*

From the given data file (corr_data.mat file in Matlab format or corr.csv as a text file), compare the Pearson correlation with the Spearman rank correlation for the time series x and y (each with 40 observations). [*1 point for Pearson correlation coefficients, 1 point for Spearman correlation coefficients*]

Repeat the comparison for the time series x2 and y2 (from the same data file as above), where x2 and y2 are the same as x and y, except that the 5th data point in y is replaced by an outlier in y2.

Repeat the comparison for the time series x3 and y3, where x3 and y3 are the same as x and y, except that the 5th data point in x and y is replaced by an outlier in x3 and y3.

Make scatterplots of the data points in the x-y space, the x2-y2 space and the x3-y3 space. Also plot the linear regression line in the scatterplots. [*1 point for each plot (data and regression line), so 3 points in total*]

Briefly discuss the results (e.g. which correlation coefficient is more resistant to outliers and why?; in potential cases with different data but same correlation coefficients: why are the coefficients unchanged while the data is altered?) [*max 2 points for the discussion*]

Problem 2 *[total of 5 points]:*

From the given data file (MLR.mat file in Matlab format or MLR.csv as a text file), perform multiple linear regression and stepwise regression with predictors x1, x2, x3, x4, x5 and x6, and the response variable y. Rank the importance of the individual predictors in their influence on y. Briefly discuss the results.

[*1 point for the regression coefficients for MLR, and 1 point for regression coefficients for stepwise. 1 point for ranking the importance of predictors, 1 point for discussing the difference between MLR and stepwise results, and 1 point if correct standardization of data is applied*]