MA 542   S '17   Test 1 Practical

This test is open-book open-note. You may not use a cell phone. You must create and upload a report of your work in Canvas.

Part II: Practical

A finance company originates loans for individuals and businesses. The data set loans, found here and in Canvas under the name loans.txt, contains data on loan originations from 38 branches in the month of October, 2016. The variables are X, the number of loans originated, and Y, the time spent on those loans, in hours.

(a) **(13 points)** Examine the variables separately and jointly. Report what you see.

(b) **(12 points)** Fit the SLR $Y = \beta_0 + \beta_1 X + \epsilon$ to the data. Conduct a lack of fit test, and an analysis of the residuals and Studentized residuals, including tests for normality and White's test for heteroscedasticity. What do you conclude about the fit?

(c) **(5 points)** Find a Box-Cox transformation of the response variable. Round the choice of λ to the nearest integer.

(d) **(15 points)** Regress the Box-Cox transformed response on the predictor and proceed as in part (b).

(e) **(5 points)** Based on the fit in (d), how many extra hours would you estimate an additional loan application would take?

(f) **(5 points)** In November 2016, a branch office expended 214 hours in originating 10 loans. Is this significantly outside the norm represented by the data set? Use an appropriate level 0.99 interval based on the model in (d) to decide.

(g) **(15 points)** Regress the original response on the Box-Tidwell transformed predictor $X^{0.2}$ and proceed as in part (b).

(h) **(5 points)** Which model of the three would you choose? Why?