STAT511 HW#11

Reading: CH11 in O&L
See Canvas Calendar for Due Date.

36 points total, **2** points per problem part unless otherwise noted.

1. Review problem 11.22 from Ott & Longnecker regarding treadmill "time to exhaustion" (X) and 10km race times (Y).
   A. Regress 10.K (Y) on Treadmill (X) and include the "summary" information in your assignment.
   B. Create a scatterplot of 10-K vs Treadmill with fitted regression line overlaid.
   C. Give the estimate, 95% confidence interval and interpretation of the **slope**. (**4 pts**)
   D. Give the R^2 value and interpretation in terms of this scenario.
   E. Give the predicted 10.K time for a runner with Treadmill = 11. Also provide a corresponding prediction interval.
   F. Create the plots of (1) residuals vs fitted values and (2) qqplot of residuals
   G. Based on the plots above, subject 13 appears to be a bit of an outlier. Run a formal outlier test for this observation. Provide the p-value and make a conclusion. Note that since we identified this observation after looking at the data, a Bonferonni adjustment is appropriate.

2. Data on age in coating Thickness (X) and Strength (Y) from an experiment involving steel are available from Canvas as Steel.csv.
   A. Regress Strength (Y) against Thick (X) and look at (1) the plot of Strength versus Thick (2) residuals versus predicted values and (3) qqplot of residuals. Include these plots in your assignment. Do the regression assumptions appear to be met? Discuss. (**4 pts**)
   B. Perform an F-test for "lack of fit". Give your p-value and make a conclusion. (**4 pts**)
   C. Now perform a quadratic regression and create a scatterplot with the fitted curve overlaid. Include the "summary" table and plot in your assignment. This can be done with code like the following. (**4 pts**)

```
QFit <- lm(Strength ~ Thick + I(Thick^2), data =
Steel)
summary(QFit)
plot(Strength ~ Thick, data = Steel)
curve(b0 + b1*x + b2*x^2, add = TRUE)
```

   Note that b0,b1,b2 need to be replaced with estimates from the quadratic regression.

3. Review problem 11.50 from Ott & Longnecker regarding SAT Scores.
   A. Create pairwise scatterplots for all 4 variables (Male.Verbal, Female.Verbal, Male.Math, Female.Math)
   B. Calculate pairwise (Pearson) correlations for all 4 variables. Which pair of variables has the strongest correlation? (**4 pts**)
   C. Provide a test of the correlation for Female.Verbal vs Female.Math. Give the p-value and conclusion in your assignment.