STAT 210
Applied Statistics and Data Analysis:
Homework 4

Due on Nov. 12/2020

Question 1

For this question we will use the data set `cars`.

 (i) Plot `dist` as a function of `speed`. Fit a simple linear regression model of `dist` as a function of `speed`. Add the regression line to the previous plot. Obtain the summary for this regression. Obtain an estimator for the error variance. Observe the value for the intercept and comment.

 (ii) Based on your comments to the previous section, fit a model without an intersect. Draw a scatterplot and add the two regression lines. Obtain a summary for the new regression and comment on the differences with the previous model, including the estimated error variance.

 (iii) We want to compare the predictive power of these two models. Using the same procedure as in exercise 1 of the list for week 9, compare the predictive power of both models and comment on your results.

Question 2

For this question use the data set `data1`.

 (i) Read the data and draw a scatterplot of `var1` as a function of `var2`. Fit a simple linear regression and add the line to the plot. Comment. Obtain a summary of the regression.

 (ii) Draw the diagnostic plots. Do you identify any point as an outlier? If you do, which point is this? Can you identify this point in the initial scatterplot?

 (iii) Fit a new regression model excluding the outlier(s) that you identified in the previous section. Draw a scatterplot with both regression lines. Compare the summary tables. Draw the diagnostic plots and comment.

Question 3

For this question use the data set `data2`.

 (i) Read `data2` and plot `yval` as a function of `xval`. Fit a simple linear regression and add the regression line to the plot. Comment. Obtain a summary for the regression and draw the diagnostic plots. Comment on the results

 (ii) Use the function `boxcox` on the package `MASS` with the argument set to the model you fitted in (i).

 (iii) If the confidence interval in the graph includes zero, use a logarithmic transformation for `yval` and fit a new model. Obtain a summary of the new regression and compare with the previous one. Draw the diagnostic plots and compare with the previous results.

(iv) Write down the final model in terms of the original variables. Draw a scatterplot of `yval` against `xval` and add the regression line for the first model and the curve you obtained with the second regression.

Question 4

For this question use the data set `data3`.

(i) Read `data3` and plot `vary` as a function of `varx`. Fit a simple linear regression and add the regression line to the plot. Comment. Obtain a summary for the regression and draw the diagnostic plots. Comment on the results.

(ii) Use the function `residualPlots` in package `car` and interpret the test produced by the function. What is your conclusion?

(iii) Add a quadratic term to the regression model and obtain a summary, draw the diagnostic plots and comment. Draw a scatterplot of the data and add the lines/curves for both models. Write down you final model.