Statistical Modeling Course

Multiple Linear Regression Assignment

The following problems involves the use of multiple linear regression on the `Auto` data set available in the `ISLR` package.

Problem 1

Use `GGpairs` in the `GGally` package to produce a scatterplot matrix which includes all of the variables in the data set and the pairwise corrlations. Set `progress = FALSE` so only the plot is printed. You will need to exclude the `name` variable, which is qualitative. Make sure to set the size to make it legible. You can change the font size by adding `theme(text=element_text(size=8))` to the plot.

Problem 2

Perform a multiple linear regression with `mpg` as the response and all other variables except `name` as the predictors. Print the results (including R^2 and p-values).

Problem 3

Comment on the output. Include the answers to the following questions: What fraction of the variance of `mpg` is explained by the model? Is there a relationship between the predictors and the response? Which predictors appear to have statistically significant relationship to the response? What does the coefficient for the `year` variable suggest?

Answer:.

Problem 4

Use the `plot()` function to produce diagnostic plots of the linear regression fit. Make sure your plots are visible in your pdf. Comment on any problems you see with the fit. Do the residual plots suggest any unusually large outliers? Does the leverage plot identify any observations with unusually high leverage?

Answer: