Statistical Learning, Exercise 5 : Multiple linear regression with R

Question 1

Load the education data set from the file Education5.txt. Use the lm() function to create a multiple linear model on this data set to predict wage and interpret the result you obtain from R.

Question 2

Load the Computer data set from the file ComputerData.txt. The variable PRP is the performance if the system.

1. Remove the variable that you cannot use to predict performances;

2. Try to build the best multiple linear regression model to predict performances, you can use all the variables;

3. Does you model explain something?

4. Explain the strategy you applied.

5. Interpret the important values of the output of the final model you obtained with R;

Question 3

Plot your model graphically on a graph with all the data. Is there outliers or predictions that are hard to predict?

Question 4

Do question 2 and 3 with the car data set from the file Cars2Data.txt. Here you have to predict the variable mpg (miles per gallon).