R Homework for Chapter 18

Motorcycles

More than one million motorcycles are sold annually (www.webbikeworld.com). Off-road motorcycles (often called "dirt bikes") are a market segment (about 18%) that is highly specialized and offers great variation in features. This makes it a good segment to study to learn about which features account for the cost (manufacturer's suggested retail price, MSRP) of a dirt bike. Researchers collected data on 2005 model dirt bikes based on a randomized experiment.

You can find the data file on Blackboard. Download it and put it in the **same folder** as your `R` program file. Then, use the following command to read in the data

```
motor <- read.table('Motorcycles.txt', sep = '\t',  header = TRUE)
```

and answer the questions below.

1. Let's take $MSRP$ ($) as response variable and consider $Wheelbase$ (in), $Displacement$ ($cu\ in$), $Bore$ (in) and $Clearance$ (in) as potential predictors. Use scatterplots to see which variables can be appropriately used as predictors in simple linear regression.

2. Build a multiple regression model for $MSRP$ using $Displacement$ and $Bore$ as predictors. Write down the fitted model. Report R^2 and adjusted R^2. Interpret the coefficients for $Displacement$ and $Bore$.

3. Check the model you fitted in the previous question to see if it satisfies the assumptions as required in multiple regression.

4. Conduct a test to see if the fitted multiple regression model is statistically useful. If useful, find the predictors that make significant contributions to the $MSRP$ in the model. Explain.

5. Suppose we are not satisfied with the R^2 given by the current model. Please propose a new multiple regression model in order to improve R^2. Compare the new model to the current one with respect to their R^2, coefficient estimates and hypothesis tests. Don't forget to check assumptions of the new model for its validity. (Hint: We have two potential predictors $Wheelbase$ and $Clearance$ in the pool. Think about how to use them to improve the model.)