CIND 123, Lab 5

You are out in the market to buy a new car and you have found a list[1] of all of the cars in the market. First you want to get a feeling of the data. So you decide to look at the outputs of the summary function. When you run the data, you see that the horse power column is classified as character.

- Use the as.numeric() function to correct this column.

Now that all columns are numeric, you want to plot graphs to start your exploratory analysis.

- Create a scatter plot using car weights in the x axis and acceleration in the y axis. (Use plot(x,y)).

After your initial findings, you want to tackle your main issue; What car should you buy? You do not want to spend too much time for this pursuit. So, you have decided to introduce a rule based approach to narrow down the list to two cars. These rules are:

- Pick the cars that have the median acceleration and find the lightest (in terms of weight) one.
- Pick the car whose weight is closest to the mean.

Guide for the rules:

1. You will have to use the which() function to access the indices of the cars with median acceleration. Then use which.min() to find the lightest car from the initial list. You will need to store two indices to access the car that match the rule's requirements.
2. For the second rule, derive the difference between each car and the mean weight of the data set by using R's built-in mean function. Add these observations to the cars dataset.
3. Then access the smallest value in the newly created column by using the which.min() function (You might need to use the absolute function before using the which.min() function).

[1] http://archive.ics.uci.edu/ml/datasets/Auto+MPG
Access the data by clicking the Data Folder, and then clicking auto-mpg.data. Select all the data and put it in a text file. Save the file at a location you can easily access. To understand what each column means please skim the auto-mpg.names file. Use the read.table() function to load the data and then convert the raw data to a data frame.