Homework Assignment 2

**Write an R script to perform following steps:**

1. load Auto.csv file in a variable called Auto.
2. What is the type and class of Auto variable
3. Use a R command/function to find the dimensions of the Auto variable
4. Use an R command to remove all NA’s from Auto variable. redo step (c) to report the new dimensions of Auto variable
5. Use an R command to report the names of all variables in Auto
6. Use attach() function to be able to directly access fields in Auto variable and then plot the cylinders by mpg. The axes of the plot should not have “Auto$.....” string
7. Convert the “cylinders” field in auto to a factor and use the plot function to generate a plot of cylinders by mpg. Make sure the results are draw with green color and the y and x axis of the graph has “CYLINDERS” and “MPG” respectively
8. Draw a histogram of mpg variable in auto, using green color to fill the bars. Make sure that the histogram only generate 10 bars
9. Use the pairs() function in R to generate a plot matrix, consisting of scatterplots for each variable-combination of “**mpg**”, “**displacement**”, “**horsepower**”, “**weight**”, and “**acceleration**” fields in Auto variable.

**Hint**: often in data analytics we are only interested in the correlations of a hand selected number of our variables. So, in such a case, function calls like e.g., pairs(Auto) will not provide the expected results because we are not interested in all fields of the Auto in this question. Also, please note that the Auto dataset contains non-numeric fields and pairs() function is unable to handle such variables and it will throw an error if it is used with such variable types

1. Report Descriptive Statistics for all fields of Auto variable

Please make sure to provide the R code and also the outcome of execution of each command in your report.