Problem 1 (45 points)

The dataset College contains a number of variables for 777 different universities and colleges in the US. The variables are

- Private : Public/private indicator
- Apps : Number of applications received
- Accept : Number of applicants accepted
- Enroll : Number of new students enrolled
- Top10perc : New students from top 10% of high school class
- Top25perc : New students from top 25% of high school class
- F.Undergrad : Number of full-time undergraduates
- P.Undergrad : Number of part-time undergraduates
- Outstate : Out-of-state tuition
- Room.Board : Room and board costs
- Books : Estimated book costs
- Personal : Estimated personal spending
- PhD : Percent of faculty with Ph.D.'s
- Terminal : Percent of faculty with terminal degree
- S.F.Ratio : Student/faculty ratio
- perc.alumni : Percent of alumni who donate
- Expend : Instructional expenditure per student
- Grad.Rate : Graduation rate

1. Use R functions to produce a scatterplot matrix of the first ten columns or variables of the data.
2. Use R functions to produce side-by-side boxplots of Outstate versus Private.
3. Use R functions to produce two different histograms with differing numbers of bins for the first two quantitative variables.
4. Create a new qualitative variable, called Elite, that is based on the Top10perc variable. The value of Elite is Yes if students coming from the top 10% of their high school classes exceeds 50%. Otherwise, the value is No

Problem 2 (35 points)

Consider the data set Auto. Make sure that the missing values have been removed from the data.

1. Which of the predictors are quantitative, and which are qualitative?
2. Use R functions to find the mean and standard deviation of each quantitative predictor.
3. Now remove the 10th through 85th observations. What are the mean and standard deviation of each predictor in the subset of the data that remains? Please use R functions to complete the task.