## 1. Analysis of cars dataset (Auto)

This exercise involves the Auto data set studied in the lab.

**(a)** Which of the predictors are quantitative, and which are qualitative?

**(b)** What is the range of each quantitative predictor? You can answer this using the range() function.

**(c)** What is the mean and standard deviation of each quantitative predictor?

**(d)** Now remove the 10th through 85th observations. What is the range, mean, and standard deviation of each predictor in the subset of the data that remains?

**(e)** Using the full data set, investigate the predictors graphically, using scatterplots or other tools of your choice. Create some plots highlighting the relationships among the predictors.  
Comment on your findings.

**(f)** Suppose that we wish to predict gas mileage (mpg) on the basis of the other variables. Do your plots suggest that any of the other variables might be useful in predicting mpg? Justify your answer.

## 2. Analysis of Boston Housing dataset (Boston)

This exercise involves a dataset that stores looks at the characteristics of houses in the different suburbs (districts) around Boston in the USA.

**(a)** To begin, load in the Boston data set. This is part of the MASS library in R.

> library(MASS)

Now the data set is contained in the object Boston.

> Boston

Read about the data set:

> ?Boston

How many rows are in this data set? How many columns? What do the rows and columns represent?

**(b)** Make some pairwise scatterplots of the predictors (columns) in this data set. Describe your findings.

**(c)** Are any of the predictors associated with per capita crime rate? If so, explain the relationship.

**(d)** Do any of the suburbs of Boston appear to have particularly high crime rates? Tax rates? Pupil-teacher ratios? Comment on the range of each predictor.

**(e)** How many of the suburbs in this data set bound the Charles river?

**(f)** What is the median pupil-teacher ratio among the towns in this data set?

**(g)** Which suburb of Boston has lowest median value of owner occupied homes? What are the values of the other predictors for that suburb, and how do those values compare to the overall ranges for those predictors? Comment on your findings.

**(h)** In this data set, how many of the suburbs average more than seven rooms per dwelling? More than eight rooms per dwelling? Comment on the characteristics of the suburbs that average more than eight rooms per dwelling.