Lab 1 - Practice Questions

Q1. Load in the Boston data set. The Boston data set is part of the MASS library in R. How many rows are in this data set? How many columns? What do the rows and columns represent? Also, read about the dataset to see what do various columns represent.

Q2. Make some plots of the predictors 'nox –'dis', 'dis' –'medv' ,'crim'- 'medv' in this data set. Describe your findings.

Q3. Are any of the predictors associated with per capita crime rate? You can use the cor() function to do so and explain your observations. [The cor() function prints the correlation coefficient between the two input variables. The correlation coefficient is a measure of the strength of the relationship between two variables. The values range between -1 and 1. A correlation of -1 shows a perfect negative correlation, while a correlation of 1 shows a perfect positive correlation. A correlation of 0 shows no linear relationship between the movement of the two variables.]

Q4. Do any of the suburbs of Boston appear to have particularly high crime rates? Tax rates? Comment on the range of the crime rate.

Q5. How many of the suburbs in this data set bound the Charles river?

Q6. What is the median pupil-teacher ratio among the towns in this data set?