## Question 3 Statistical Data Mining HW – 1

### Part a

Make pairwise scatterplots of the predictors, and describe your findings

Comment: Some variables are definitely correlated.

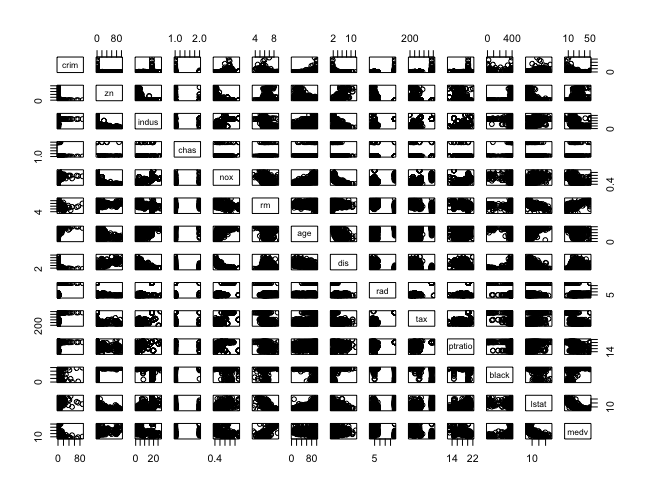
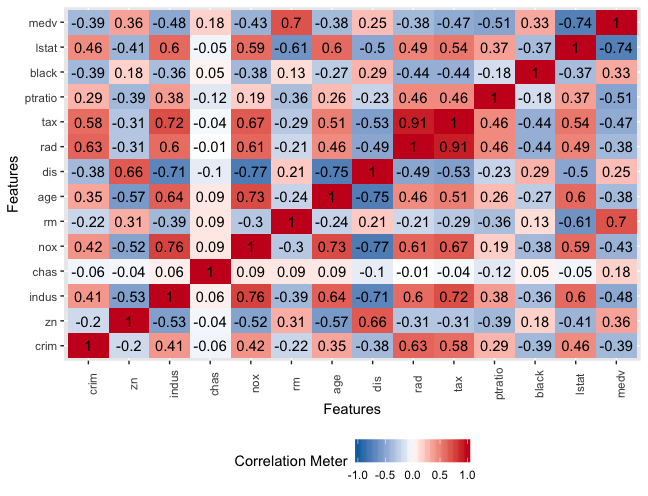


Fig: Correlation Matrix

Observe the above correlation matrix this does helps a lot

From the above figure, it is difficult to extract good insights

Let’s see co-relation heat map:



To be more specific a heat map is the best tool.

### Part b

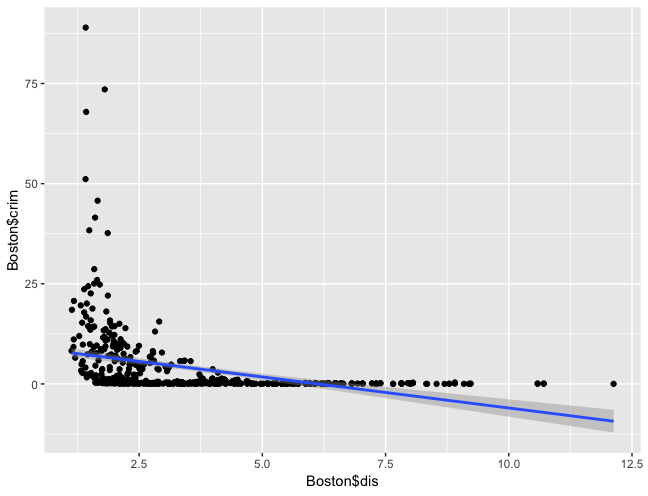
Based on the heat map we can be certain that there is some good association between per capita crime rate and various factors.

For Example:

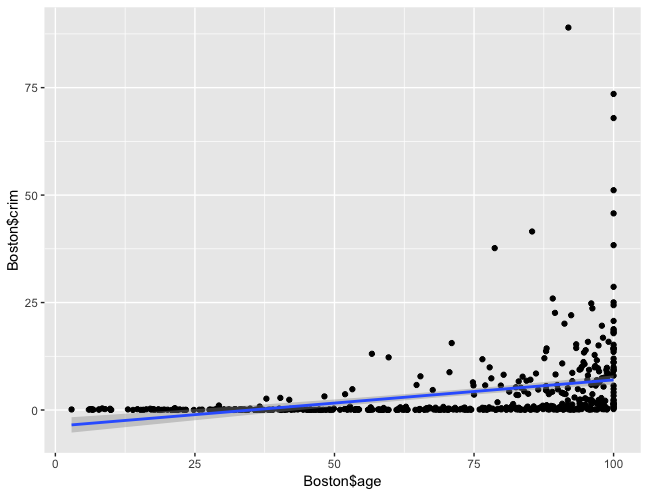
Observe the positive correlation between Features Crim and rad, Crim and tax.

A negative correlation between crim and black.

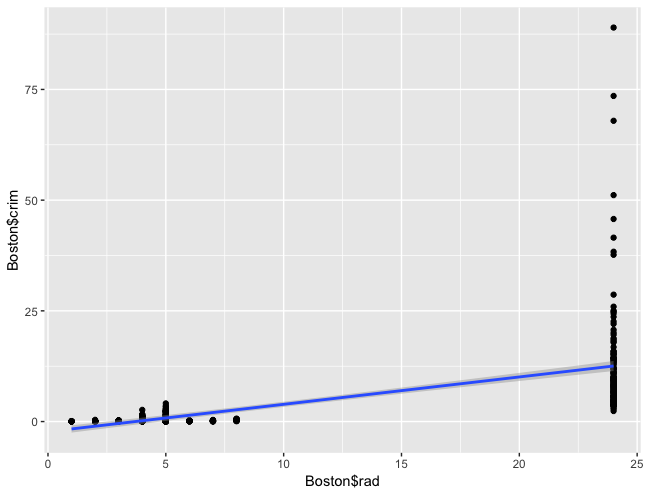
A detailed approach to elaborate the above correlations in the following figures:



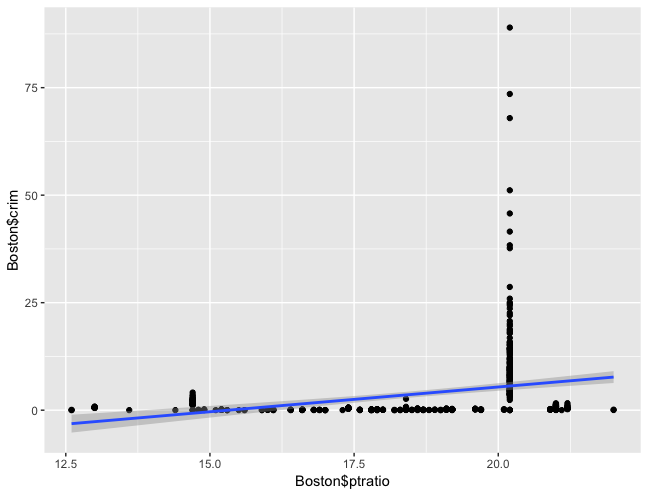
* Crime is high in the areas near Employment Centers



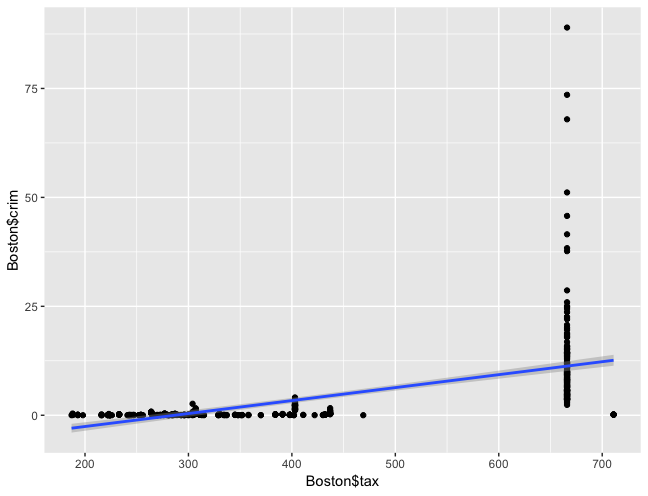
* There is a significant skew in the areas near older homes



* A surprisingly positive correlation between distance to highways and crime rate.



* Not much of the correlation between crime and pupil-teacher ratio by town.

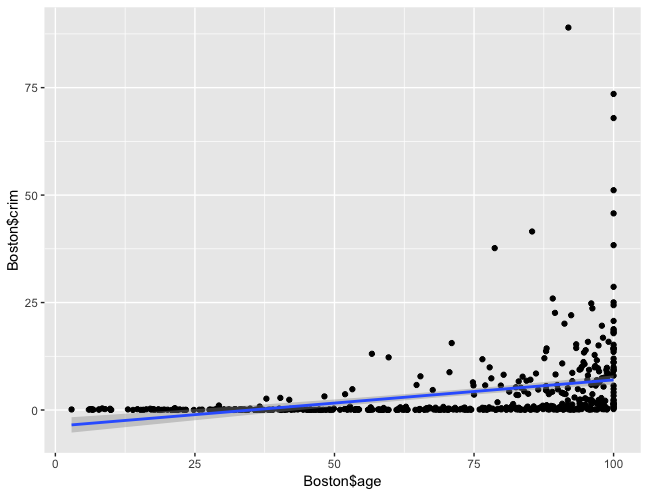


* A positive skew in the crime rate near the high tax paying areas

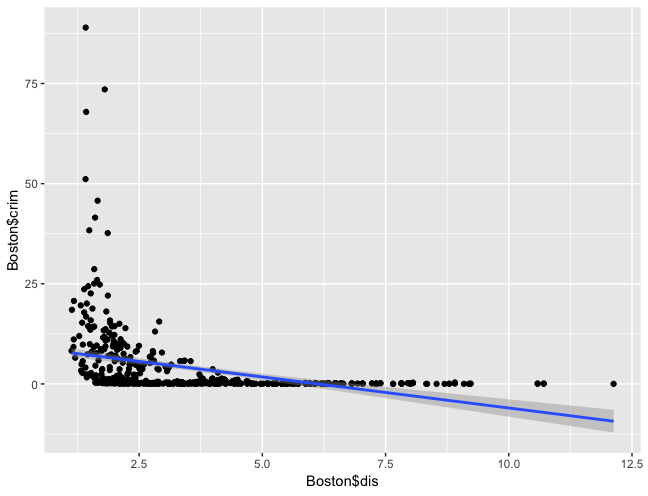
### Part c

Most of the Boston is safe except 18 Suburbs

there are indeed some neighborhoods where the crime rate is alarmingly high and the count is 18

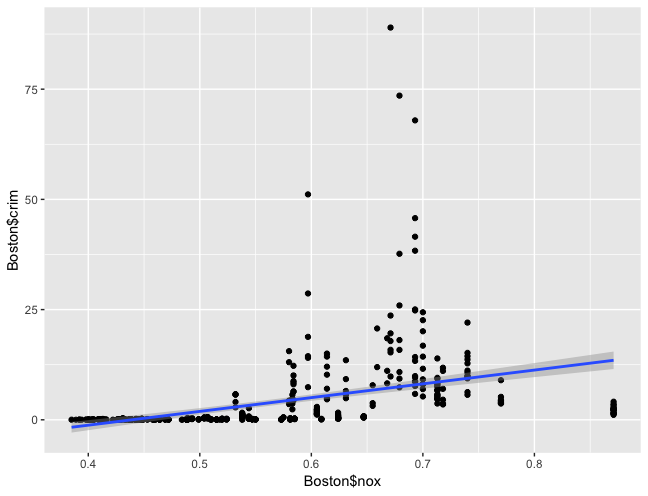


* areas near the older homes appear to have high crime rates

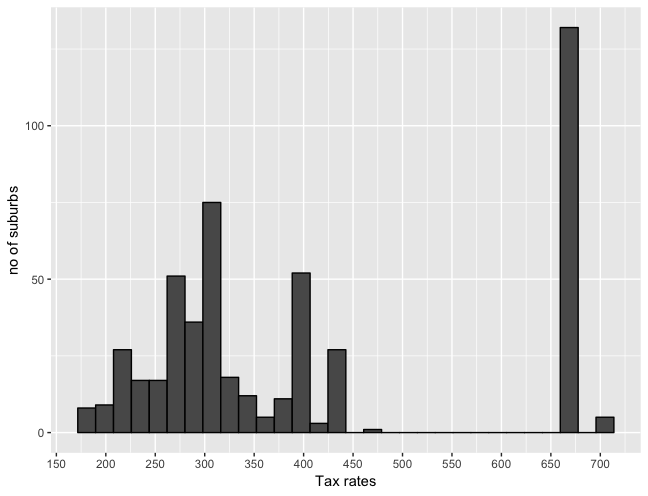


* High crime rate in the areas near Employment centers

crime rate values are respectively about 0.26% and 89%, there are indeed some neighborhoods where the crime rate is alarmingly high



practically unrelated features like nitrogen oxide concentration does show positive correlation to high crime rates



* A few of the suburbs pay the high tax rates

### Part d

how many of the suburbs average more than Seven rooms per

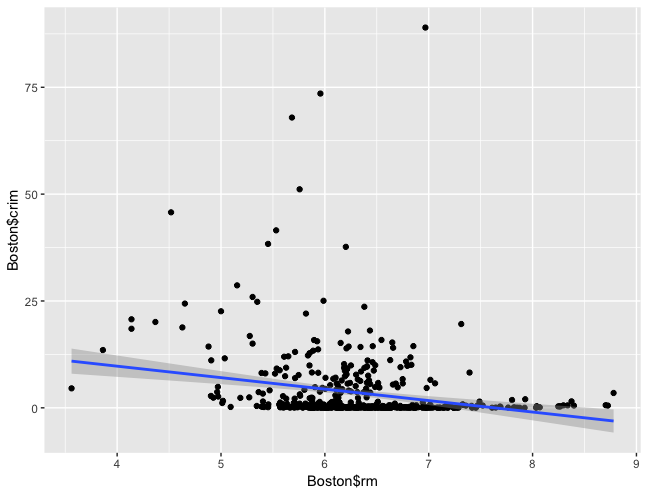
dwelling?

64

More than eight rooms per dwelling?

13

Comment on the suburbs that average more than eight rooms per dwelling.



* We can definitely say there is low crime in the dwellings that average more than 8 rooms