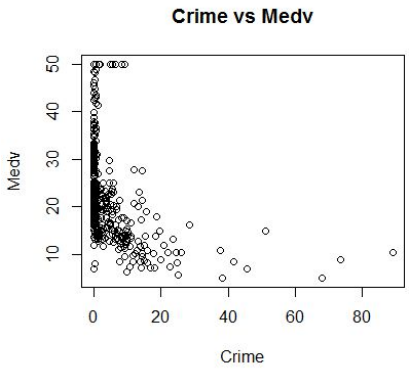
****Don't forget to append your first initial + last name to the end of the file when submitting!

11. In the Boston Housing data set, what is the relationship between crime and housing prices? Please support your claims with exploratory analysis conducted in R. Does this relationship make sense? Justify your answer. IE: What are some reasons this relationship makes sense or does not make sense?

The scatterplot between crime and the median value of owner-occupied homes does not show a clear linear pattern and the correlation value is also moderately negative hence there is an indication that as the crime rate increases the housing price gets negatively impacted and decreases, but further analysis is needed.

```
> cor(BostonHousing$medv,BostonHousing$crim)
[1] -0.3883046
```



12. Based on your analysis of the Boston Housing data set, please provide an interpretation for the top 3 strongest absolute correlations. Offer some hypothesis as to why these correlations may be present.

The strongest absolute correlation found between tax and rad at 0.9102282 suggest that they are interrelated to each which makes sense practically because as the accessibility to radial highways we can expect higher property tax rates.

The second strongest absolute correlation is noted between dis and nox which is -0.7692301 which makes complete sense since as the distances to employment centres increase the nitrogen oxides concentration decreases.

The third strongest absolute correlation noted is between nox and indus at 0.7636514 proportion of non-retail business acres which can be hypothesized as if there was increase in the non-retail business acres we expect the nitrogen oxides concentration also increases in that area.

13. Based on your analysis of the 90th%, 92.5th%, 95th%, 97.5th% and 99th% confidence intervals for the mean of passing yards, you should have noticed that the bounds of the confidence intervals increase as the percentile (90th%, 92.th%, etc) increases. In your own words, explain why this relationship exists.

The bounds of the confidence intervals tend to increase as the percentile of the confidence increases because the scope for the range of predictions widens, it happens because with the increase in the percentile the error needs to be reduced for which we widen the interval.

14. Based on your anova of the food calorie counts, please state the null and alternative hypothesis and interpret the results of your anova.

Null hypothesis: The mean calorie counts of food1,food2,food3 and food4 are all the same.

Alternative Hypothesis: The means are not the same

As the Pr(>F) value standing at 0.00688 seems very significant we can reject the null hypothesis, and state that the means are not equal.