A) Copy/Pasted output from code:

```
Opening file Boston.csv
Reading line 1
heading: rm,medv
new length 506
Closing file Boston.csv
Number of records: 506

Stats for rm
Sum: 3180.03
Mean: 6.28463
Median: 6.209
Range: 5.219

Stats for medv
Sum: 11401.6
Mean: 22.5328
Median: 21.2
Range: 45

 Covariance = 4.49345

 Correlation = 0.696737
```

B) Using Built in Functions in R vs Creating functions in C++.

Using the built in functions in R is much easier and more time efficient than having to code our own functions in C++. R has a lot of functions that you can use so every function we had to code in C++ was already pre coded in R. Overall, R's built in functions are convenient but for what we did in this assignment the functions were pretty simple to code.

C)

Mean is the average of a group. The way to calculate the mean is to add up all the values and then divide by the number of elements. Median is the value of the element at the middle of a sorted group. If there are two middle elements then take the mean of those two elements. Range is the difference between the maximum and minimum value of a set. I believe, given a set you could take these values and it would give you a summary of what the set was like. Also, I think you could make some predictions given this information.

D) Covariance and Correlation

Covariance is a representation on how similar the data is between two sets. If the data is similar you will get a positive covariance value, if they are not similar, then you will get a negative covariance value. Correlation refers to the linear correlation of two sets. I believe both covariance and correlation are useful in machine learning. In class we went over linear regression. You can use the correlation to help the machine since correlation

deals strictly with how related two sets are based on their linearity. I think covariance is important in this example since, when finding the correlation of two sets you must first know the covariance of these sets.