David Teran
February 4, 2023
CS 4375.004

Snapshot of code output

```
Loaded '/usr/lib/liboah.dylib'. Symbols loaded.
Opening file: Boston.csv for data exploration.
Headings: rm,medv

Records of data: 506
     rm
Sum: 3180.03
Mean: 6.28463
Median: 6.209
Range: 5.219

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

Covariance of rm and medv: 4.49345
Correlation of rm and medv: 0.69536

Program Terminated! Have a great day!
The program '/Users/davidt/Documents/SE4375_Project1/data_exploration_Component1' has exited with code 0 (0x00
000000).
```

Experience Using R Built-in Functions vs. Coding Functions in C++:
A good refresher in programming using C++, but also not too bad to program. Using the built-in functions in R is much easier and convenient than programming the functions themselves in C++. It was still a good way to get back into remembering how to program in C++ after using Java for a while.

Statistical Measures:
The mean, or the average of values in a set of data, is useful in data exploration to determine patterns in sets of data, whether the average is rising or falling. The range, or the difference between the minimum and maximum values in a data set, can also help in determining how far apart the data values. The median, or the middle values of a set of data, can also determine the average values as well. All three descriptive statistics can be useful in data exploration for determining how varied the values are in a data set, how apart they are, and can help determine a pattern or relationship between data sets.

Covariance and Correlation:
The covariance helps determine how changes in one variable can be associated with the changes of a second or other variable, and how these variables are related to each other. Correlation determines the extent of how two variables or data sets are related to each other. The information from both covariance and correlation can be used to predict the future behavior of data sets or variables, which is useful in machine learning.