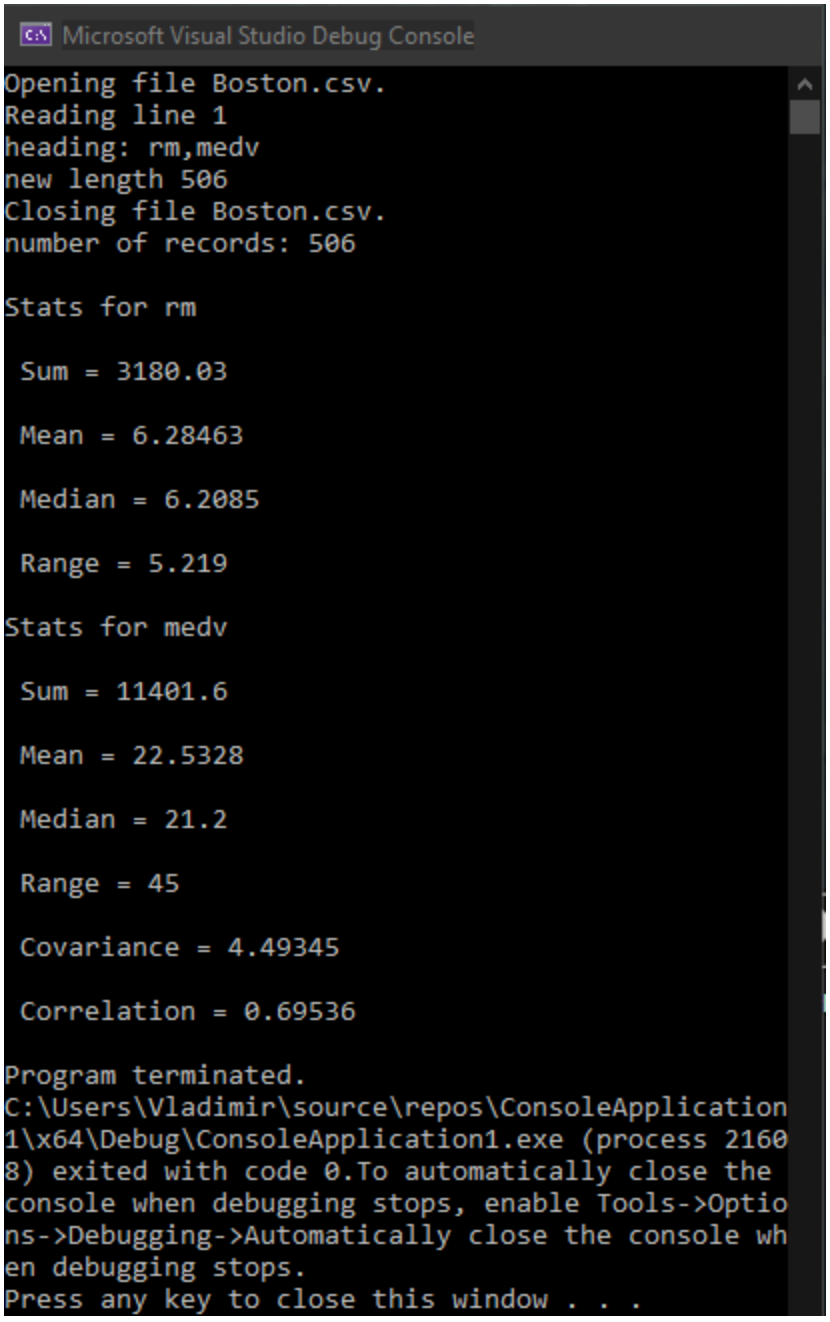

a. Describe your experience using built-in functions in R versus coding your own functions in C++

Using built-in functions in R are much simpler than coding our own functions. There is much more room for error but there is more room for flexibility when writing your own functions. The provided code for reading in the file was very helpful.

b. Describe the descriptive statistical measures mean, median, and range, and how these values might be useful in data exploration prior to machine learning

Mean is the sum divided by the number of values and provides the starting point for analyzing the data. Median is the middle value in a sorted data set and a large difference from the mean could indicate outliers or extreme values. Range is the largest value minus the smallest value and describes the spread of data points.

c. Describe the covariance and correlation statistics, and what information they give about two attributes. How might this information be useful in machine learning?

Covariance describes how strongly related two attributes are with large positive and negative numbers meaning strongly related while numbers close to zero meaning weakly related. Correlation is covariance adjusted to [-1,1] where -1 and 1 mean perfectly correlated. Machine learning is all about finding patterns in data to figure out which attributes are related and how which uses correlation.