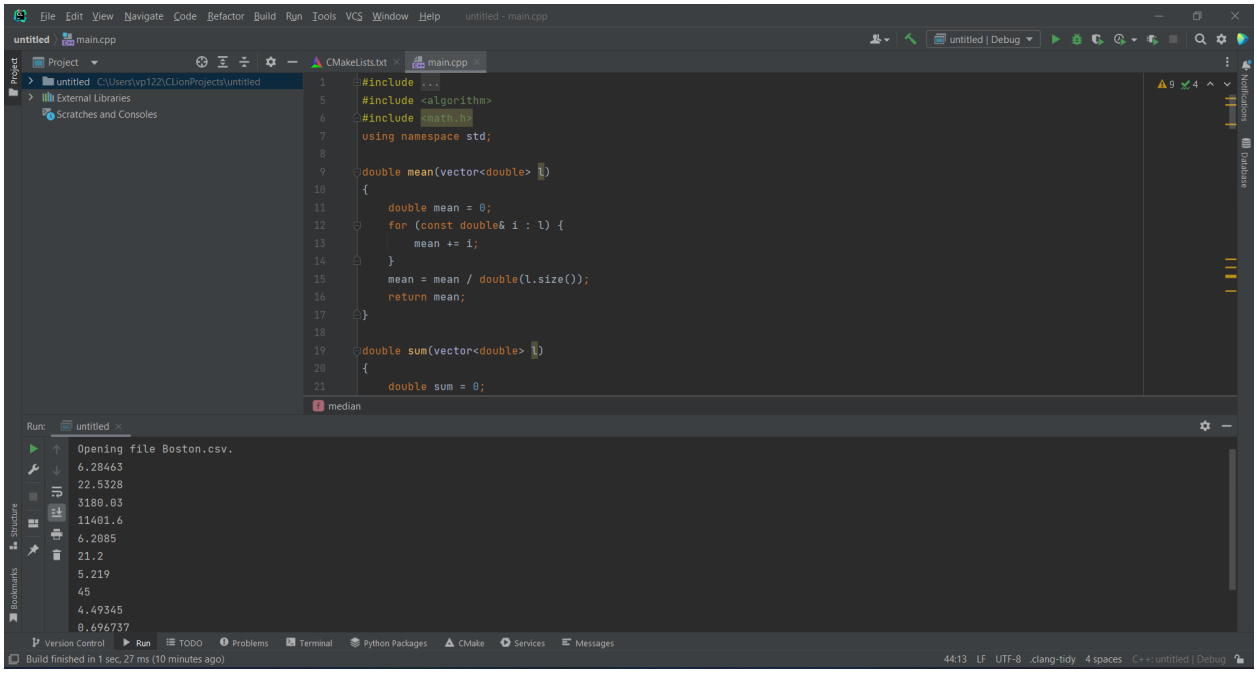
Vedant Prakash
Vxp200027
CS 4375



This is an output of my code. The first eight lines are the functions mean, sum, median , and range run on vectors rm and medv respectively and the last two lines are the result for the covariance and correlation functions respectively.

My experience with coding my own functions is that I gained a deeper understanding of how they worked. Before coding my own function, I did not know how covariance and correlation were calculated so this assignment gave me a deeper understanding of that. R makes it easy for me so this assignment for good for me to go a little deeper in the workings of the mathematical formulas.

Mean is the sum of all the values in a dataset divided by the number of values. Median is the middle data point of a dataset ordered from smallest to biggest or biggest to smallest. Range is the difference between the smallest and biggest values. These are useful in data exploration because these values can be used to compute more detailed statistics such as covariance and correlation that can be very useful for machine learning.

Covariance is how changes in one variable are associated with changes in a second variable. Correlation is the linear relationship between two variables. These statistics are useful in machine learning because if the model can know how two variables/features are related, it can use their relationship to increase its accuracy for predicting the target.