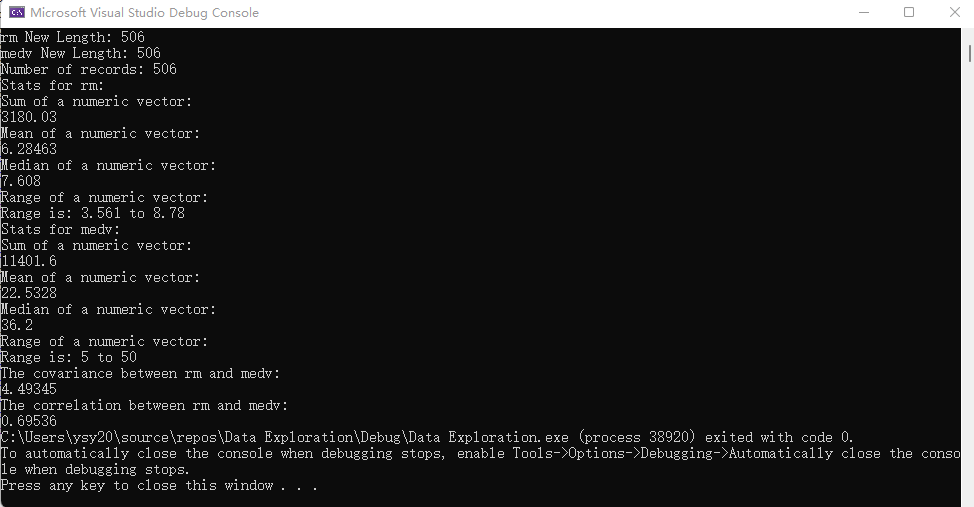
C++ Data Exploration

1. Copy/paste runs of your code showing the output.



1. Describing your experience using built-in functions in R versus coding your own functions in C++.

In R versus, it is easy to get answers for sum, average, median, correlation, and covariance from function like median(), sum(), and so on. However, in C++, I cannot get answers directly from function. I need to calculate each part of the formula and combine them to get the final answer.

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

Mean in descriptive statistical is adding all figures within the data set and divide the number of figures. Median in descriptive statistical is the figure in the middle of the data set. Range in descriptive statistical is the lowest to the highest value in the data set. They are considered as different parameters in the machine learning.

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

Covariance is to measure of how much random variables vary together. It can tell us what’s the relationship between variables. The large covariance means strong relationship, otherwise means weak relationship. Correlation is to indicates how strongly two variables are related. Range is -1 to 1. The figure of correlation is closer to the 1, it means that two variables have strong relevance. These information can be used in Image Recognition.