a. Visualize the correlation between all variables in a meaningful and clear way of representing. Find out top 3 reasons for having more crime in a city.

b. What is the difference between co-variance and correlation? Take an example from this dataset and show the differences if any?

Covariance and Correlation are two mathematical concepts which are quite commonly used in statistics. Both of these two determine the relationship and measures the dependency between two random variables. Despite, some similarities between these two mathematical terms, they are different from each other. Correlation is when the change in one item may result in the change in the another item. On the other hand, covariance is when two items vary together. Read the given article to know the differences between covariance and correlation.

Key Differences Between Covariance and Correlation

The following points are noteworthy so far as the difference between covariance and correlation is concerned:

A measure used to indicate the extent to which two random variables change in tandem is known as covariance. A measure used to represent how strongly two random variables are related known as correlation.

Covariance is nothing but a measure of correlation. On the contrary, correlation refers to the scaled form of covariance.

The value of correlation takes place between -1 and +1. Conversely, the value of covariance lies between -∞ and +∞.

Covariance is affected by the change in scale, i.e. if all the value of one variable is multiplied by a constant and all the value of another variable are multiplied, by a similar or different constant, then the covariance is changed. As against this, correlation is not influenced by the change in scale.

Correlation is dimensionless, i.e. it is a unit-free measure of the relationship between variables. Unlike covariance, where the value is obtained by the product of the units of the two variables.

Conclusion

Both measures only linear relationship between two variables, i.e. when the correlation coefficient is zero, covariance is also zero. Further, the two measures are unaffected by the change in location.

Correlation is a special case of covariance which can be obtained when the data is standardized. Now, when it comes to making a choice, which is a better measure of the relationship between two variables, correlation is preferred over covariance, because it remains unaffected by the change in location and scale, and can also be used to make a comparison between two pairs of variables.