1. Use the given link Data Set.

Answer the below questions:

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

Answer:

**Covariance** is a statistical term, defined as a systematic relationship between a pair of random variables wherein a change in one variable reciprocated by an equivalent change in another variable.

**Correlation** is described as a measure in statistics, which determines the degree to which two or more random variables move in tandem. During the study of two variables, if it has been observed that the movement in one variable, is reciprocated by an equivalent movement another variable, in some way or the other, then the variables are said to be correlated.

Difference between Correlation and Covariance are:

1. 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.
2. Covariance is nothing but a measure of correlation. On the contrary, correlation refers to the scaled form of covariance.
3. The value of correlation takes place between -1 and +1. Conversely, the value of covariance lies between -∞ and +∞.
4. 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.
5. 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.