

Covariance:

Covariance is a measure of joint variability of 2 random variables. It basically quantifies the relationship between 2 random variables.

Formula

$$cov_{x,y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N - 1}$$

$cov_{x,y}$ = covariance between variable a and y

x_i = data value of x

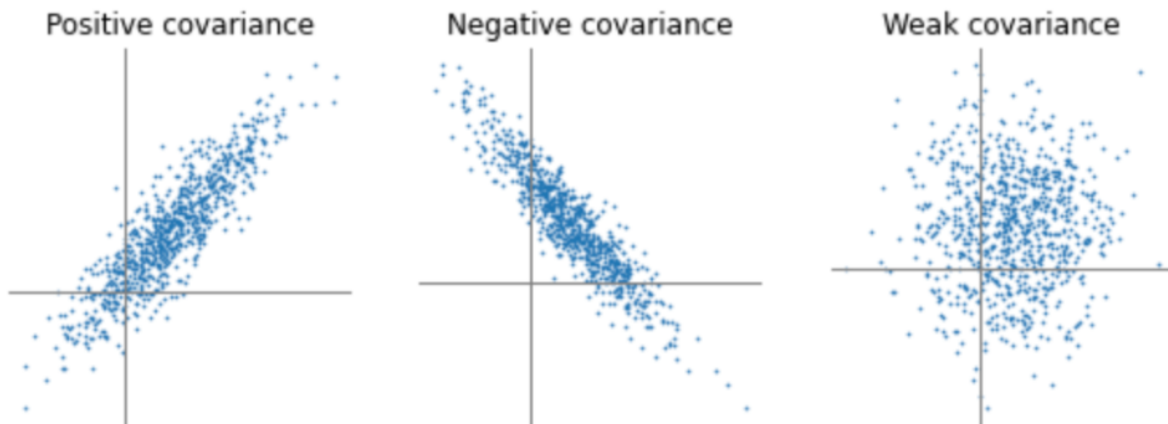
y_i = data value of y

\bar{x} = mean of x

\bar{y} = mean of y

N = number of data values

If greater values of one random variable correspond with the greater values of other and same followed for lesser values, then it is positive covariance. For negative covariance, greater values of one variable correspond with the lesser values of other variable, and, vice-versa.



Hence, covariance can be used to analyze the tendency in linear relationship between 2 variables. But the magnitude of covariance is not easy to interpret, as it has no limit and is NOT normalized. To get around this problem we use Pearson Correlation.

Pearson Correlation ⇒ https://en.wikipedia.org/wiki/Pearson_correlation_coefficient