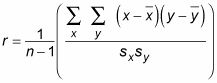
The formula for the correlation (r) is



Where:

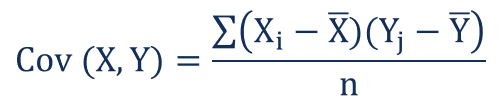
N is the number of pairs of data;

1590722223(1) and 1590722258(1) are the sample means of all the x-values and all the y-values, respectively;

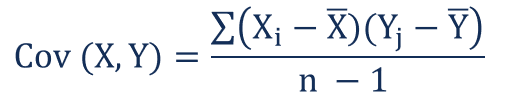
Sx and Sy are the sample standard deviations of all the x- and y-values, respectively.

### **Formula for Covariance**

The covariance formula is similar to the formula for correlation and deals with the calculation of data points from the average value in a dataset. For example, the covariance between two random variables X and Y can be calculated using the following formula (for population):



For a sample covariance, the formula is slightly adjusted:



Where:

Xi – the values of the X-variable

Yj – the values of the Y-variable

X̄ – the mean (average) of the X-variable

Ȳ – the mean (average) of the Y-variable

n – the number of data points

1. ****In what ways they similar? In what ways they differ?****

**Similar:**

Covariance and correlation both primarily assess the relationship between variables. The closest analogy to the relationship between them is the relationship between the variance and [standard deviation](https://corporatefinanceinstitute.com/resources/knowledge/standard-deviation/" \o ").

**Differ:**

Covariance measures the total variation of two random variables from their expected values. Using covariance, we can only gauge the direction of the relationship (whether the variables tend to move in tandem or show an inverse relationship). However, it does not indicate the strength of the relationship, nor the dependency between the variables.

1. ****What is each metrix actually measuring?****

Correlation measures the strength of the relationship between variables. Correlation is the scaled measure of covariance. It is dimensionless. In other words, the correlation coefficient is always a pure value and not measured in any units.

1. ****Outline a scenario in which the covariance is preferable than correlation.****

In this scenario:

As we know, people’s every meal consisted by 3 parts: carbohydrate, fats and protein.

we want to know how if people having more percentage of carbohydrate in their meal will help or resist people to gain weight. So the two variables here are:

P: Percentage of calories that carbohydrate in people’s meal.

W: People’s weight gained in a observation time.

So , in this scenario, we just want to find the two variables have a positive or negative influence to each other or not, but we do not need to find how strong they influence each other, we will prefer to use covariance than correlation.