**Correlation :**

Correlation refers to the statistical relationship between the two entities. It measures the extent to which two variables are linearly related[1].

The correlation coefficient is calculated by determining the covariance of the variables and dividing that number by the product of those variables' standard deviations.

Steps to calculate :

1.Calculate the mean (average) of each variable.

2.Calculate the difference between each variable value and its corresponding mean.

3.Multiply the differences for each variable.

4.Sum up the multiplied differences for each variable.

5.Calculate the standard deviation of each variable.

6.Multiply the standard deviations of each variable.

7.Divide the sum of the multiplied differences by the product of the standard deviations.

**Covariance:**

Covariance is a statistical term that refers to a systematic relationship between two random variables in which a change in the other reflects a change in one variable.

Assuming the two variables A and B

The covariance between two variables A and B is calculated as follows:

1.Calculate the average of A and B.

2.Calculate :X= A -mean(A).

3.Calculate :Y= B-mean(B).

4.Multiply the differences:.X\*Y

5.Sum up the multiplied differences.

6.Divide the sum by the total number of data points

**Similarities:**

Both of them shows the relationship between the variables , a positive correlation or covariance indicates positive relationship and negative values represent the inverse relationships.

**Differences:**

Range of correlation is -1 to +1 and range of the covariance is -infinity to infinity.

Correlation measures the strength of the variables under comparison and covariance measures the extent of the change with respect to change in another.

Correlation is independent of units while covariance depends on the units.

**What is each metric actually measuring?**

Correlation measures the strength and direction of the linear relationship between two variables.

Covariance is a statistical measure that quantifies the relationship between two variables. It measures how much the variables change together and provides information about the direction and magnitude of their relationship.

**Outline a scenario in which the use of a covariance metric would  
prove to be preferable to the use of a correlation metric.**

If we want to analyze the relationship between the height and weight of individuals of a population. Of a diverse groups -adults, teenagers, and children covariance metric would be preferable to a correlation metric. Correlation is a standardized measure and it assumes a linear relationship between variables . But this case, the relationship between height and weight may not strictly follow a linear pattern.By calculating the covariance between height and weight, you can uncover the general trends in their co-variation, regardless of the linearity assumption.

References :

[1]. <https://www.simplilearn.com/tutorials/statistics-tutorial/what-is-correlation-in-statistics#:~:text=Correlation%20refers%20to%20the%20statistical,two%20variables%20are%20linearly%20related>

[2]. ChatGpt

[3]. <https://machinelearninginterview.com/topics/machine-learning/covariance-and-correlation/>

[4]. https://www.excelr.com/blog/data-science/statistics-for-data-scientist/correlation-vs-covariance