Q. 1). Define Covariance and explain how it differs from Correlation in terms of scale and interpretation.

Answer,

Covariance is a statistical measure that indicates the direction of the linear relationship between two random variables.
Covariance shows direction, while correlation shows direction and strength on a standardized scale.

Q 2). What does a positive, negative, and zero covariance indicate about the relationship between two variables?

Answer,

Positive covariance: Indicates that the two variables tend to move in the same direction.
Negative covariance: Indicates that the two variables tend to move in opposite directions.
Zero covariance: Indicates no linear relationship between the two variables. Changes in one variable do not systematically correspond to changes in the other (though a nonlinear relationship may still exist).

Q 3). Discuss the limitations of covariance as a measure of relationship between two variables. Why is correlation preferred in many cases?

answer,

Covariance is useful for identifying the direction of a relationship, but due to its scale dependence and poor interpretability, it is limited.
Correlation is preferred in most applications because it is standardized, easier to interpret, and more meaningful for comparing relationships between variables.

Q 4). Explain the difference between Pearson's correlation coefficient and Spearman's rank correlation coefficient. When would you prefer to use Spearman's correlation?

Answer,

Pearson's Correlation Coefficient
Measures the strength and direction of a linear relationship between two variables.
Based on the actual numerical values of the data.

Use Spearman's correlation when data are ranked, non-normal, contain outliers, or when the relationship is monotonic rather than strictly linear.

Q 5). If the correlation coefficient between two variables X and Y is 0.85, interpret this value in context. Can you infer causation from this value? Why or why not?

Answer,

A correlation coefficient of 0.85 suggests a strong positive relationship between X and Y, but it does not establish cause-and-effect.
Determining causation requires controlled experiments, temporal evidence, or additional statistical and theoretical justification.

Q. 6). Using the dataset below, calculate the covariance between X and Y.

x	2	4	6	8
Y	3	7	5	10

answer,

Sample covariance between X and Y ≈ 6.33
Population covariance between X and Y = 4.75
The positive value indicates a positive relationship between X and Y.

Q. 7). Compute the Pearson correlation coefficient between variables A and B:

A	10	20	30	40	50
B	8	14	18	24	28

answer,

This value indicates an extremely strong positive linear relationship between variables A and B.

r = 0.998

Q 8). The following table shows heights (in cm) and weights (in kg) of 5 students. Find the correlation coefficient between Height and Weight.

Height	150	160	165	170	180
waight	50	55	58	62	70

answer,

There is an extremely strong positive linear relationship between height and weight for these students.
Taller students tend to weigh more.

r=0.993

Q 9). Given the dataset below, determine whether there is a positive or negative correlation between X and Y. (No need for exact calculation, just reasoning.)

x	1	2	3	4	5
Y	15	12	9	7	3

answer,

There is a negative correlation between X and Y.

As X increases from 1 to 5, Y consistently decreases from 15 to 3.
This indicates that the two variables move in opposite directions.

Q. 10).Two investment portfolios have the following returns (%) over 5 years. Compute the covariance and correlation coefficient, and interpret whether the portfolios move together.

year	portfolio A	portfolio B
1	8	6
2	10	9

	3	12	11
	4	9	8
	5	11	10

Answer,

Covariance = 3
Correlation coefficient ≈ 0.99

The positive covariance indicates that both portfolios tend to move in the same direction.

The very high positive correlation (≈ 0.99) shows a strong linear relationship between the returns.