1. What is the definition of covariance? Create the formula for it.

Ans: In statistics, covariance is a measure of joint variability of two random variables. For two jointly distributed random variables, covariance is defined as the expected value of the product of their deviations from their individual expected values.

1. What makes correlations better than covariance?

Ans: The main reason why correlations are better than covariance that it does not get affected by the change in scale of data.

1. Explain the process as well as Pearson and Spearman Correlation.

Ans: **Pearson correlation**, is the measure of linear correlation between two sets of data. It is the ratio of the covariance of the two variables and the product of their standard deviations. Thus, it is a normalized measurement of covariance, such that the result always has value between -1 and +1.

**Spearman correlation**: While Pearson corelation only assess the linear relationships, spearman can assess the monotonic relationships (whether linear or not). If there are no repeated data values, a perfect Spearman correlation of +1 or −1 occurs when each of the variables is a perfect monotone function of the other.

1. What are the advantages of Spearman Correlation over Pearson Correlation?

Ans: **Spearman correlation**: While Pearson corelation only assess the linear relationships, spearman can assess the monotonic relationships (whether linear or not). If there are no repeated data values, a perfect Spearman correlation of +1 or −1 occurs when each of the variables is a perfect monotone function of the other.

1. Describe the Central Limit Theorem?

Ans: The central limit theorem states that, in many situations when independent random variables are summed up, their properly normalized sum tends towards a normal distribution even of the original variables themselves are not normally distributed.