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

The Covariance Formula
The formula is: Cov(X,Y) = Σ E((X – μ) E(Y – v)) / n-1 where: X is a random variable. E(X) = μ is the expected value (the mean) of the random variable X and. E(Y) = v is the expected value (the mean) of the random variable Y.

2. What makes Correlations better than Covariance?

Both correlation and covariance measures are also unaffected by the change in location. However, when it comes to making a choice between covariance vs correlation to measure relationship between variables, correlation is preferred over covariance because it does not get affected by the change in scale.

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

Pearson correlation: Pearson correlation evaluates the linear relationship between two continuous variables. Spearman correlation: Spearman correlation evaluates the monotonic relationship. The Spearman correlation coefficient is based on the ranked values for each variable rather than the raw data.

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

Also, does not estimate a natural population parameter (unlike Pearson's which estimates ). An advantage of the Spearman rank correlation coefficient is that the X and Y values can be continuous or ordinal, and approximate normal distributions for X and Y are not required.

5. Describe the Central Limit Theorem.

The central limit theorem states that if you have a population with mean μ and standard deviation σ and take sufficiently large random samples from the population with replacement , then the distribution of the sample means will be approximately normally distributed.