1. **How do you assess the statistical significance of an insight?**

   We can use P-value to determine statistical significance

   If P-value is less than alpha, we can reject the null hypothesis. If the P-value is higher than alpha, we fail to reject null hypothesis.

   (If P-value is less than alpha, reject Ho for Ha. If P-value is greater than alpha, fail to reject Ho for Ha)

2. **What is the Central Limit Theorem? Explain it. Why is it important?**

   If we have a population with mean μ and standard deviation σ and take sufficiently large random samples from the population, then the distribution of the sample means will be approximately normally distributed.

   The central limit theorem is a valuable tool for analyzing large datasets, as it enables us to make the reasonable assumption that, in most cases, the sampling distribution of the mean will follow a normal distribution.

3. **What is statistical power?**

   Statistical power refers to the likelihood of observing a statistically significant result at significance level alpha (α) when a true effect of a specific magnitude exists. It enables the detection of a genuine difference between test variations when such a difference is indeed present.

4. **How do you control for biases?**

   There are many ways to control for biases in statistics.

   Random Sampling:

   Use random sampling techniques to select a representative sample from the population of interest. Random sampling reduces selection bias by giving each element in the population an equal chance of being included in the sample.

   Large Sample Sizes: Larger sample sizes can help reduce sampling bias and improve the accuracy of your estimates and results.

5. **What are confounding variables?**

   A confounding variable impacts both the dependent variable and the independent variable, leading to a misleading or spurious association.

6. **What is A/B testing?**

   A/B testing is a randomized controlled experiment that involves comparing two versions of a variable to determine their relative performance in a controlled environment.

7. **What are confidence intervals?**

   A confidence interval represents the likelihood that a parameter will lie within a specified range around the mean. Confidence intervals quantify the level of confidence or uncertainty associated with a sampling method.