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

An insight is statistically significant if it's linked to causation instead of correlation. Any possibilities of occurring factors from which the insight is derived due to correlation are "weeded" out via randomization and maintaining a large sample size.

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

The sample data follows a normal distribution under three conditions: random sampling, independent sampling, large sample size. Knowing the distribution enables one to make predictions on the data, specifically determining how far the sample mean is from the true population mean.

3. What is the statistical power?

Statistical power is the probability of observing the sample data given that the null hypothesis is false / the hypothesis is true.

4. How do you control for biases?

Null hypothesis significance testing controls for confirmation bias by seeing if a hypothesis can be proven false instead of proven true. The support for the hypothesis is strengthened if there is failure in proving the hypothesis false.

5. What are confounding variables?

Confounding variables are factors that contribute to the observation of both the independent variable and the dependent variable, making it difficult to determine whether the independent variable correlates to the dependent variable or causes the dependent variable.

6. What is A/B testing?

A/B testing defines guidelines for forming a hypothesis where the null hypothesis represents a scenario reflecting the present (thus also representing the control group) and the hypothesis represents a scenario that is singularly modified in some respect to the null hypothesis.

7. What are confidence intervals?

It defines the range of data that reflects a particular percentage of the true population falls under. For example, a 95% confidence interval of "x" from the sample mean "u" (given normal distribution) means that 95% of the true population have a value in the range [u - x, u + x]