Rachel Lindor
Data Science Bootcamp – HW 7

1. How do you assess the statistical significance of an insight?
By calculating the p-value and comparing it to a significance threshold (e.g., 0.05). If the p-value is less than the threshold, the result is statistically significant, meaning it is unlikely to have occurred by chance.

2. What is the Central Limit Theorem? Explain it. Why is it important?
The Central Limit Theorem (CLT) states that the sampling distribution of the sample mean approaches a normal distribution as the sample size increases, regardless of the population's original distribution.
Importance: It allows us to make inferences about population parameters using sample statistics and applies to hypothesis testing and confidence intervals.

3. What is the statistical power?
Statistical power is the probability of correctly rejecting the null hypothesis when it is false. It measures the sensitivity of a test to detect an effect, typically increased by larger sample sizes, higher effect sizes, or lower variability.

4. How do you control for biases?
Bias can be controlled by:

Using random sampling and random assignment.
Designing double-blind studies.
Standardizing procedures.
Applying statistical techniques like regression analysis to account for biases.
5. What are confounding variables?
Confounding variables are extraneous variables that are related to both the independent and dependent variables, potentially distorting the observed relationship between them.

6. What is A/B testing?
A/B testing is a statistical method to compare two versions (A and B) of a variable (e.g., web pages, features) to determine which performs better by measuring a predefined metric like click-through rate.

7. What are confidence intervals?
Confidence intervals are ranges of values, derived from sample data, that are likely to contain the true population parameter with a specified level of confidence (e.g., 95%).