STATISTICS WORKSHEET-1

Q1. Answer – A

Q2. Answer – A

Q3. Answer – C

Q4. Answer – C

Q5. Answer – C

Q6. Answer – B

Q7. Answer – B

Q8. Answer – A

Q9. Answer – C

Q10. What is Normal Distribution?

Answer - In a normal distribution, the data is symmetrically distributed around the mean, with the highest frequency of observations occurring at the mean. The distribution is bell-shaped, with tails that extend infinitely in both directions.

Q11. - How do you handle missing data? What imputation techniques do you recommend?

Answer – There are several ways to handle missing data:

a. Delete rows with missing data

b. Mean/Median/Mode imputation

c. Assigning a unique value

d. Predicting the missing values

e. Using an algorithm which supports missing values, like random forests.

The best method is to delete rows with missing data as it ensures that no bias or variance is added or Removed, and ultimately results in a robust and accurate model. However, this is only recommended If there's a lot of data to start with and the percentage of missing values is low.

Q12. - What is A/B testing?

Answer - A/B testing is a form of hypothesis testing and two-sample hypothesis testing to compare Two versions. The control and variant, of a single variable. It is commonly used to improve and Optimize user experience and marketing.

Q13. Is mean imputation of missing data acceptable practice?

Answer - Mean imputation is generally bad practice because it doesn't take into account feature Correlation. For example, imagine we have a table showing age and fitness score and imagine that an Eighty-year-old has a missing fitness score. If we took the average fitness score from an age range of 15 to 80, then the eighty-year old will appear to have a much higher fitness score that he actually

Should.

Q14. What is linear regression in statistics?

Answer - Linear regression is one of the statistical techniques used in predictive analysis, in this

Technique will identify the strength of the impact that the independent variables show on deepened

Variables.

Q15. What are the various branches of statistics?

Answer - a. Descriptive Statistics: This usually summarizes the data from the sample by making use of an index

Like mean or standard deviation. The methods which are used in the descriptive statistics are

Displaying, organizing, and describing the data.

b. Inferential Statistics: These conclude from data which are subject to random variations like

Observation mistakes and other sample variations.