

STATISTICS WORKSHEET-1

Answers

Q1- a

Q2- d

Q3- b

Q4- d

Q5- c

Q6- a

Q7- b

Q8- a

Q9- c

Answer Q10

The normal distribution is the most widely known and used of all distributions. Because the normal distribution

approximates many natural phenomena so well, it has developed into a standard of reference for many probability problems.

Answer Q12

A/B testing

A/B testing in its simplest sense is an experiment on two variants to see which performs better based on a given metric. Typically, two consumer groups are exposed to two different versions of the same thing to see if there is a significant difference in metrics like sessions, click-through rate, and/or conversions.

Answer Q13

Mean imputation is typically considered terrible practice since it ignores feature correlation.

Answer Q14

Linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable. For example, a modeler might want to relate the weights of individuals to their heights using a linear regression model.

Answer Q15

Statistics is a study of presentation, analysis, collection, interpretation and organization of data

There are two main branches of statistics

- Inferential Statistic.
- Descriptive Statistic.

* Inferential Statistic - Involves using a sample to draw conclusion about a population.

* Descriptive Statistic - Involves the organization, summarization, and display of data.

Answer Q11

We can use below method to handle Missing data

- 1- Deleting the record which has a missing value
- 2- Creating separate models to handle missing values
- 3- statistical methods like mean, median or mode