Assignment 3 Answers

- **Statistics Assignment 3 Answers**

Answer 1 - B) Total Variation = Residual Variation + Regression Variation

Answer 2 - B) Binomial

Answer 3 – A) 2

Answer 4 – A) Type 1 Error

Answer 5 – B) Size of the test

Answer 6 – A) Decrease

Answer 7 – B) Hypothesis

Answer 8 – D) All of the mentioned

Answer 9 – A) 0

Question 10 - What Is Bayes' Theorem?

Answer 10 - Bayes' theorem describes the probability of occurrence of an event related to any condition. It is also considered for the case of conditional probability. Bayes theorem is also known as the formula for the probability of "causes".

Question 11 - What is z-score?

Answer 11 - A Z-score is a numerical measurement used in statistics of a value's relationship to the mean (average) of a group of values, measured in terms of standard deviations from the mean. If a Z-score is 0, it indicates that the data point's score is identical to the mean score.

Question 12 - What is t-test?

Answer 12 - A t-test is a type of inferential statistic used to determine the significant difference between the means of two groups, which may be related to certain features. A t-test is used as a hypothesis testing tool, which allows testing an assumption applicable to a population. A t-test looks at the t-statistic, the t-distribution values, and the degrees of freedom to determine the statistical significance. A t-test allows us to compare the two data sets' average values and determine if they came from the same population.

For example, if we take a sample of students from class A and another sample of students from class B, we don't get the same mean and standard deviation. Similarly, samples taken from the placebo-fed control group and those taken from the drug prescribed group should have a slightly different mean and standard deviation.

Question 13- What is percentile?

Answer 13 – A percentile is a term that describes how a score compares to other scores from the same set. While there is no universal definition of percentile, it is commonly expressed as the percentage of values in a set of data scores that fall below a given value

Question 14 - What is ANOVA?

Answer 14 - Analysis of variance (ANOVA) is a collection of statistical models. It is one of the significant aspects of statistics. The statistics students should be aware of the analysis of variance. But most of the statistics students find it challenging to understand the analysis of variance.

Question 15 - How can ANOVA help?

Answer 15 -

- **Hypothesis Testing:** Enables the comparison of independent and dependent variables.
- **Understanding Data Sets:** An analyst or statistician can best determine inconsistencies in data sets.
- **Group Comparisons:** Allows multiple groups to be compared at the same time to uncover relationships between data.
- **Sales and Marketing Improvement:** Businesses can answer customer and product research questions to improve advertising and marketing for better sales.
- **Project Management:** Leadership, such as project management, can better align their goals and strategies with business and departmental cost objectives.
- **Industry-Wide Approach:** ANOVA is effective for a wide variety of uses across different industries, including financial services, eCommerce, industrial, R&D, and more.

- **Product Development:** Organizations can better pinpoint and understand what product features to improve or adapt for the best results.

• SQL Assignment 3 Answers

1. create database customers;

use customers;

create table customers (custumerNumber int, custumerName char(20), custumerFirstName char(20), custumerLastName char(20),phone int(), addressLine1 char(50), addressLine2 char(50),city char(20),state char(20),postalCode int (6),country char (20),salesRepEmployeeNumber int(10),creditLimit int);

2. create database orders;

use orders;

create table orders (orderNumber int, orderDate int, requiredDate int, shippedDate int, status char(20),comments char(1000),custsomerNumber int);

3. SELECT orderNumber, orderDate, requiredDate, shippedDate, status, comments, custsomerNumber FROM orders;

4.SELECT comments From orders;

6. SELECT employeNumber, lastName, firstName From employees;

Machine learning Answers

1. **(D) All of above**
2. **(D) None**
3. **(C) Reinforcement learning and Unsupervised learning**
4. **(B) The tree representing how close the data points are to each other**
5. **(A) A distance metric**
6. **(C) k-nearest neighbour is same as k-means**
7. **(D) 1,2 and 3**
8. **(B) 2nd only**
9. **(A) 2**
10. **(D) Given historical weather records, predict if tomorrow's weather will be sunny or rainy**
11. **A**
12. **A**
13.

• Having clustering methods helps in restarting the local search procedure and remove the inefficiency. In addition, clustering helps to determine the internal structure of the data.

• This clustering analysis has been used for model analysis, vector region of attraction.

• Clustering quality depends on the methods and the identification of hidden patterns.

Answer 14 -