STATISTICS WORKSHEET-1

Q1 to Q9

1. A) True
2. A) Central Limit Theorem
3. C) Modeling contingency tables
4. D) All of the mentioned
5. C) Poisson
6. B) False (because estimated value doesn't change the CLT)
7. B) Hypothesis
8. A) 0
9. C) Outliers cannot conform to the regression relationship

Q10 to Q15 Subjective answers

10 - The term Normal Distribution refers to data which we want to analyze can have any distribution and the probability distribution graphs can take on very distinct and recognized shapes. It is the continuous probability distribution with a probability density function that gives a symmetrical bell curve.

11 - We handle missing data by using two primary methods to solve the error are imputation or the removal of data. It uses or guess for missing data. It's most useful when the percentage

of missing data is low. Imputation techniques is multiple imputation that we generate missing values from the dataset at many times.

12 - A/B testing is also known as Split testing, which means to a randomized experimentation process wherein two or more versions of a variable are shown to different segments of website visitors at same time to determine which versions leaves the maximum impact and drives business metrics. Creating two versions of a digital asset to see which one users respond to better. Examples of assets include a landing page, display ad, marketing email, and social post.

13 - True, imputing the mean preserves the mean of the observed data. So, if the data are missing at random, the estimate of the mean remains unbiased.

14 - Linear Regression is the simplest and most extensively used statistical a technique for predictive modelling analysis. It is a way to explain the relationship between a dependent variable(target) and one or more explanatory variables(predictors) using a straight line. There are two types of linear regression- Simple and Multiple.

15 - The various branches of Statistics:

- Mathematical or theoretical statistics: it helps in forming the experimental and statistical distribution.

- Statistical methods or functions: it helps in the collection, tabulation and interpretation of the data.
- Descriptive statistics: it helps in summarizing and organizing any data set characteristics.
- Inferential statistics: it helps in finding the conclusion regarding the population after analysis on the sample drawn from it.