**STATISTICS WORKSHEET - 1**

1. A) True
2. A) Central Limit Theorem
3. B) Modeling bounded count data
4. D) All of the Mentioned
5. C) Poisson
6. B) False
7. B) Hypothesis
8. A) 0
9. C) Outliers cannot confrom to the regression relationship
10. The Term Normal Distribution:

Normal Distribution is easy to work with the mathematically. And it is a bell shaped frequency distribution curve. For example height and intelligence are approximately normally distributed, The center of a normal distribution is located at its peak, and 50% of the data lies above the mean, while 50% lies below. It follows that the mean, median, and mode are all equal in a normal distribution.

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

* Fill Null Values (.fillna)
* Verify if NANs are filled
* Droppig unwanted cloumns
* Imputation

To fill the missing values in a column or feature are numerical, the values can be imputed by the **Linear Regression model.**

1. A/B testing :

It is a randomized control experiment. Need to make a hypothesis that is tentative insight then we should have two hypothesis that is The null hypothesis is the one that states that sample observations result purely from chance. he alternative hypothesis challenges the null hypothesis and is basically a hypothesis that the researcher believes to be true.

1. Mean imputation reduces the variance of the imputed variables.Mean imputation shrinks standard errors, which invalidates most hypothesis tests and the calculation of confidence interval.It does not preserve relationships between variables such as correlations.
2. What is linear regression in statistics?

Linear regression uses one independent variable to explain or predict the outcome of the dependent variable using a straight line. The simplest form of the regression equation with one dependent and one independent variable is defined by the formula y = c + b\*x, where y = estimated dependent variable score, c = constant, b = regression coefficient, and x = score on the independent variable.

1. What are the various branches of statistics?
2. Descriptive statistics : if data can be described without any statistical tools then it is called descriptive statistics . It describes the properties of sample and population data. example, marks in class , height of student.
3. Inferential statistics: if data is too big then then we use inferential statistics. It uses those properties to test hypotheses and draw conclusions.