STATISTICS WORKSHEET ASSIGNMENT-1

(1) a) True
(2) a) Central Limit Theorem
(3) b) Modelling bounded count data
(4) d) All of the mentioned
(5) c) Poisson
(6) b) False
(7) b) Hypothesis
(8) c) 1
(9) c) Outliers cannot conform to the regression relationship
(10) What do you understand by the term Normal Distribution?
    a. ANS:- The normal distribution is the probability distribution of the values of a random variable equally in the left and right side of the central tendency (Mean) in the graph, so that the graph looks like the bell shaped curve. In this type all the data points lying within the standard deviation 1 in the both sides of the mean.
(11) How do you handle missing data? What imputation techniques do you recommend?
    a. ANS:- There are many ways to handle the missing data, as per preferences (1) Imputation of the missing data with the very previous data value or with the very next data value (2) Regression Imputation by the information of other variables to predict the missing values by using regression model, (3) Hot-Deck Imputation by replacing missing values with the similar records randomly and (4) Imputation with the Mean value of the respective column data if outliers are not present in the dataset.
(12) What is A/B testing?
    a. ANS:- The A/B testing is the method to compare two versions of features or products by splitting them in two groups, that why it is also known as Split testing. In this method we split the feature data into two groups “A” as the control group and “B” as the experimental group to check the further user responses of which version is better.
(13) Is mean imputation of missing data acceptable practice?
    a. ANS:- The mean imputation of missing data is acceptable if outliers are not present in the dataset, otherwise it would convey the wrong predictions. So we have to first deal with the Outliers before using of the mean imputation of the missing data.
(14) What is linear regression in statistics?
    a. ANS:- If we are having two types of variables (1) Independent variables and (2) Dependent variable, dependent on the mentioned (1) Independent variables. In such type of scenarios, when we predict the values of the dependent variable on based on independent variable is called Linear Regression in Statistics. So Linear Regression is the Variable based prediction of another Variable.
(15) What are the various branches of statistics?
    a. ANS:- There are two main branches of Statistics (1) Descriptive Statistics and (2) Inferential Statistics. Wherein the Descriptive Statistics as its name shows is the method to consider and describe the properties of the population data and its sample data for further statistical procedures and whereas on the other hand the Inferential Statistics uses these data for further testing and drawing the conclusions.