1. Bernoulli random variables take (only) the values 1 and 0.
   a) True
2. Which of the following theorem states that the distribution of averages of iid variables, properly normalized, becomes that of a standard normal as the sample size increases?
   a) Central Limit Theorem
3. Which of the following is incorrect with respect to use of Poisson distribution?
   b) Modeling bounded count data
4. Point out the correct statement.
   d) All of the mentioned
5. _____ random variables are used to model rates.
   c) Poisson
6. 10. Usually replacing the standard error by its estimated value does change the CLT.
   b) False
7. Which of the following testing is concerned with making decisions using data?
   b) Hypothesis
8. Normalized data are centered at_____and have units equal to standard deviations of the original data.
   a) 0
9. Which of the following statement is incorrect with respect to outliers?
   c) Outliers cannot conform to the regression relationship
10. What do you understand by the term Normal Distribution?

    When we plot the distribution plot diagram for given data, we should be able to see the bell curve in the plot. There should not be any skewness in the data. That king of data are called as normally distributed dataset. It is symmetrical around its mean.
11. How do you handle missing data? What imputation techniques do you recommend?

    Depends on the dataset, if it makes sense for filling missing values, then we can start filling the missing values. Else we can drop those missing data and continue building the model with available data.

    If we decide to fill the missing values, we can follow any of the below methos.

    1. If the data is continuous, we can take the **mean of it and use fillna()** function to fill the missing values. If the data is categorical, we can fill the missing values using **fillna() by passing mode()** value.
    2. We can also **use simple imputer** also to fill the values by giving the required strategy as input (eg: mean,median,etc).Simple Imputer can be imported by using below code

       "form sklearn.impute import SimpleImputer"
    3. We can also use **knn imputer** to fill the missing values by passing the number of neighbors as hyper parameter. It can be imported using below code.

       "form sklearn.impute import KNNImputer"
12. What is A/B testing?

    A/B test is a **statistical hypothesis testing**. Hypothesis is between two data sets and those data sets are compared with each other to find is there a significant relationship between those data sets or not.

    Yes. Because filling with mean value will reduce the variance of those data points and will maintain the distribution of data normally and reduces the standard deviation.

13. Is mean imputation of missing data acceptable practice?

    Yes. Because filling with mean value will reduce the variance of those data points and will maintain the distribution of data normally and reduces the standard deviation.

14. What is linear regression in statistics?

    In statistics, **linear regression** is a linear approach for modelling the relationship between a target variable and features.

15. What are the various branches of statistics?

    1.data collection

    2.descriptive statistics

    3.inferential statistics