1. Bernoulli random variables take (only) the values 1 and 0.
   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?
   Central Limit Theorem
3. Which of the following is incorrect with respect to use of Poisson distribution?
   Modelling event/time data.
4. Point out the correct statement.
   All of the mentioned
5. _____ random variables are used to model rates.
   Poisson
6. Usually replacing the standard error by its estimated value does change the CLT.
   False
7. Which of the following testing is concerned with making decisions using data?
   Hypothesis
8. Normalized data are centered at_____and have units equal to standard deviations of the original data.
   0
9. Which of the following statement is incorrect with respect to outliers?
   Outliers cannot conform to the regression relationship

10. What do you understand by the term Normal Distribution?
    Normal Distribution is a probability distribution that is symmetric about a mean. The data near the mean are more frequent in the occurrence than the data far from the mean. The graph of a Normal Distribution will be like Bell Shape Curve. In Normal Distribution Mean is Zero and the Standard Deviation is 1 and it has Zero Skew. Normal Distribution is always a symmetrical distribution.

11. How do you handle missing data? What imputation techniques do you recommend?
    Missing Data is defined as the values or data that is not stored (or not present) in the given data set. In Pandas it is represented by Nan. It is important to handle this missing data appropriately. The process of filling/updating these missing values (Nan) is known as imputation. There are many imputations techniques - Replacing it with an arbitrary value, Replacing with Mean, Replacing with Mode, Replacing with Median, replacing with next value (Backward Fill), Replacing with the previous value (Forward Fill), Interpolation etc. Handling Missing Values is one of the challenges of data analysis. Understanding the different categories of missing data helps in handling the missing data efficiently. It is very important to understand your data well and why it's missing, talk to the experts, if possible, to figure out what's going on with the data before blindly following any of the above methods.

12. What is A/B testing?
    A/B testing in simple terms is an experiment done on two variants to see which performs better in each metrics. Typically, two consumers are exposed to two different version of same thing to see the significant difference in the metrics. The testing is not limited to two variants, their can be n number of variants in the testing environment.

13. Is mean imputation of missing data acceptable practice?
    The process of replacing NULL values in a data collection with the data's mean value is know as mean imputation. Mean imputation is typically considered a terrible practice since it ignores features corelation. Second mean imputation decreases the variance of our data while increasing the bias. As a result of this the model is less accurate.

14. What is linear regression in statistics?
    Linear Regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called dependant variable and the variable that you are using to predict is independent variable. The technique uses statistical calculation to plot a trend line in a set of data point. The equation of an linear regression can be written as mentioned below.
    Y = C + B*X
    Where Y = Estimated Dependant Variable
    C = Constant
    B = Regression Coefficient
    X = Independent Variable.

15. What are the various branches of statistics?
    Statistic is basically concern with collection, analysis, interpretation and presentation of data. Two main branches of statistic are descriptive statistic and inferential statistic. Descriptive Statistic summarizes data from an sample using indexes such as mean or standard deviation. It just describes and summarises the data but do not draa any conclusion. Whereas Inferential Statistic draws conclusion and make prediction from data that are subject to random variations.