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
   a) The exponent of a normally distributed random variables follows what is called the log-normal distribution
5. _____ random variables are used to model rates.
   b) Poisson
6. 10. Usually replacing the standard error by its estimated value does change the CLT
   c) False
7. 1. Which of the following testing is concerned with making decisions using data?
   c) Hypothesis
8. 4. Normalized data are centered at_____and have units equal to standard deviations of the original dat
   a) 0
9. Which of the following statement is incorrect with respect to outliers?
   b) Outliers cannot conform to the regression relationship

10. . What do you understand by the term Normal Distribution?

    The normal distribution, is one of the relevant distributions in statistics for independent variables, and it can be identified with his bell shape curve in graph analysis. And this curve is centered around the mean. It only has 3 significant probability values standard deviation from -3 -2 -1 U +1 +2 +3 which is below and above the mean. The normal distribution has a Standard deviation between -1 to +1 which representing 68% of the data, between -2 to +2 is 95% of the data, between -3 to +3 is 99.7% of the data. It is one of the most popular distribution simply because is applicable to real life application.

11. How do you handle missing data? What imputation techniques do you recommend?
    First before imputing any techniques, I will identify my data type. Let's my features contains categorical data and there is a missing value, I will prefer to use (. dropna) now. The reason is because is better to loos some data than to input wrong information. As for .mode it will depend of the categorical data and the problem statement.
    Secondly, if my features are continuous that is float or integer data I will use (.mean).
12. What is A/B testing

13. Is mean imputation of missing data acceptable practice?

    Yes, mean imputation for missing data is accepted when the features are continuous data.

14. What is linear regression in statistics?

Regression in statistic is the process of predicting a label (or dependent variable) based on the features (independent variable at hand. Regression analysis is an important tool for analyzing the data

15. What are the various branches of statistics?

There are four various branches of statistics

- Descriptive Statistics: Is talking about the numbers of of rows, columns, means, outliers, 1st quantile, 2nd quantile, 3rd quantile, how it can be spread, how much deviation it has etc. (when you get the basic description of the data
- Inferential Statistics: is establishing the relationship between data (columns) structured data in the form of table and format, e.g. Excel -worksheet, sql-table, python- DataFrame
- Predictive Statistics: is the machine learning part, where you will make predictions of the target variables. Let's say you are predicting salary target column base on the name, age, gender, experience, department, age, ethnicity, country, base on this you can predict any employee.
- Prescriptive statistics: Base on the data, will try to prescribe or suggest something base on the analysis or information you have in the data.