Q1 to Q9 have only one correct answer. Choose the correct option to answer your question.

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?
d) All of the mentioned

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. 1. Which of the following testing is concerned with making decisions using data?
b) Hypothesis

8. 4. 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

Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.

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

Normal distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean.
That means Half of the data will fall to the left of the mean; half will fall to the right.

11. How do you handle missing data? What imputation techniques do you recommend?
There are various ways to handle missing data based on the data

1. Deleting the rows with missing data
2. Filling the missing data with a value – Imputation
3. Imputation with an additional column
4. Filling with a Regression Model

We can use Mean / Median method as it will not give accurate data but close to it and eassy method to follows

12. What is A/B testing?

An AB test is an example of **statistical hypothesis testing**, a process whereby a hypothesis is made about the relationship between two data sets and those data sets are then compared against each other to determine if there is a statistically significant relationship or not

13. Is mean imputation of missing data acceptable practice?
I Guess no, as it is not give accurate data is just minimize the variance of data in big picture.
So mostly for top approach with min variance we can use it but not for accuracy.

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 the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

15. What are the various branches of statistics?
There are three real branches of statistics: data collection, descriptive statistics and inferential statistics