INTERNSHIP Task
STATISTICS WORKSHEET-1

1) a) True

2) a) Central Limit Theorem

3) b) Modeling bounded count data

4) a) The exponent of a normally distributed random variables follows what is called the log- normal distribution

5) c) Poisson

6) b) False

7) b) Hypothesis

8) a) 0

9) d) None of the mentioned

10) What do you understand by the term Normal Distribution?
Ans: A normal distribution is a type of continuous probability distribution in which most data points cluster toward the middle of the range, while the rest taper off symmetrically toward either extreme. The middle of the range is also known as the mean of the distribution.

11) How do you handle missing data? What imputation techniques do you recommend?
Ans: 1. Missing data can be dealt with in a variety of ways. I believe the most common reaction is to ignore it. ...

2. Another common strategy among those who pay attention is imputation. ...

3. Mean imputation. ...

4. Substitution. ...

5. Hot deck imputation. ...

6. Cold deck imputation. ...

7. Regression imputation. ...

8. Stochastic regression imputation.

12) What is A/B testing?
Ans: It is a methodology for comparing two versions of a webpage or app against each other to determine which one performs better.

13) Is mean imputation of missing data acceptable practice?
Ans: Mean imputation is typically considered terrible practice since it ignores feature correlation.

Consider the following scenario: we have a table with age and fitness scores, and an eight-year-old has a missing fitness score.

14) What is linear regression in statistics?
Ans: Linear regression is a data analysis technique that predicts the value of unknown data by using another related and known data value. It mathematically models the unknown or dependent variable and the known or independent variable as a linear equation.

15) What are the various branches of statistics?
Ans: The two major areas of statistics are known as descriptive statistics, which describes the properties of sample and population data, and inferential statistics, which uses those properties to test hypotheses and draw conclusions.