ASSIGNMENT 6
STATISTICS

1. Which of the following can be considered as random variable?
Ans: Option D; all of the above

2. Which of the following random variable that take on only a countable number of possibilities?
Ans: Option A; discrete

3. Which of the following function is associated with a continuous random variable?
Ans: Option A; pdf

4. The expected value or _____ of a random variable is the center of its distribution.
Ans: Option C; mean

5. Which of the following of a random variable is not a measure of spread?
Ans: Option C; empirical mean

6. The _____ of the Chi-squared distribution is twice the degrees of freedom.
Ans: Option A; variance

7. The beta distribution is the default prior for parameters between _____
Ans: Option C; 0 and 1

8. Which of the following tool is used for constructing confidence intervals and calculating standard errors for difficult statistics?
Ans: Option B; boostrap

9. Data that summarize all observations in a category are called _____ data.
Ans: Option b; summarized

10. What is the difference between a boxplot and histogram?
Ans: Histograms and boxplots are kind of visaalization plots. Histogram gives frequency/ counts of data presemt in specified column. Incase of boxplots they are better used to check for outliers as the data above and below the quamtiles of boxplots shows that the outliers present.

11. How to select metrics?
Ans: Metrics are measures of quantitative assessment commonly used for assessing, comparing, and tracking performance or production. Generally, a group of metrics willtypically be used to build a dashboard that management or analysts review on a regular
basis to maintain performance assessments, opinions, and business strategies.

12. How do you assess the statistical significance of an insight?
Ans: To assess statistical significance, you would use hypothesis testing. The null hypothesis and alternate hypothesis would be stated first. Second, you'd calculate the p-value, which is the likelihood of getting the test's observed findings if the null hypothesis is true. Finally, you would select the threshold of significance (alpha) and reject the null hypothesis if the p-value is smaller than the alpha — in other words, the result is statistically significant.

13. Give examples of data that doesnot have a Gaussian distribution, nor log-normal.
Ans:distributions of income; distributions of house prices; distributions of bets placed on a sporting event.
14. Give an example where the median is a better measure than the mean.
Ans: Income is the classic example of when to use the median instead of the mean because its distribution tends to be skewed.

15. What is the Likelihood?
Ans: The likelihood is the probability that a particular outcome is observed when the true value of the parameter is equivalent to the probability mass.