Q1-Which of the following can be considered as random variable
Answer-d) All of the mentioned

Q2-Which of the following random variable that take on only a countable number of possibilities?
Answer-a) Discrete

Q3-Which of the following function is associated with a continuous random variable?
Answer-a) pdf

Q4-The expected value or _____ of a random variable is the center of its distribution
Answer-c) mean

Q5- Which of the following of a random variable is not a measure of spread?
Answer- c) empirical mean

Q6-The _____ of the Chi-squared distribution is twice the degrees of freedom
Answer-b) standard deviation

Q7- The beta distribution is the default prior for parameters between _____
Answer-c) 0 and 1

Q8-Which of the following tool is used for constructing confidence intervals and calculating standard errors for difficult statistics?
Answer-b) bootstrap

Q9-Data that summarize all observations in a category are called _____ data.
Answer-b) summarized

Q10and Q15 are subjective answer type questions, Answer them in your own words briefly.
Q10-What is the difference between a boxplot and histogram?
Answer-While boxplots and histograms are visualizations used to show the distribution of the data, they communicate information differently.

Histograms are bar charts that show the frequency of a numerical variable's values and are used to approximate the probability distribution of the given variable. It allows you to quickly understand the shape of the distribution, the variation, and potential outliers.

Boxplots communicate different aspects of the distribution of data. While you can't see the shape of the distribution through a box plot, you can gather other information like the quartiles, the range, and outliers. Boxplots are especially useful when you want to compare multiple charts at the same time because they take up less space than histograms.

Q11- How to select metrics?
Answer-Metrics are measures of quantitative assessment commonly used for comparing, and tracking performance or production. Metrics can be used in a variety of scenarios. They are used to track specific processes, such as the conversion rate of a marketing or sales initiative. Metrics help transform the vague requirements that a customer gives into a series of numbers that can be used to accurately map the process for its efficiency. Metrics tell us whether a process is good enough to meet the customer's requirements or whether it needs to be better.

Q12-How do you assess the statistical significance of an insight?

Answer-You would perform hypothesis testing to determine statistical significance. First, you would state the null hypothesis and alternative hypothesis. Second, you would calculate the p-value, the probability of obtaining the observed results of a test assuming that the null hypothesis is true. Last, you would set the level of the significance (alpha) and if the p-value is less than the alpha, you would reject the null - in other words, the result is statistically significant.

Q13-Give examples of data that does not have a Gaussian distribution, nor log-normal?

Answer-**Any type of categorical data won't** have a gaussian distribution or log-normal distribution.

Exponential distributions - eg**. the amount of time that a car battery lasts or the amount of time until an earthquake occurs**

Q14- Give an example where the median is a better measure than the mean.

Answer- Income is the classic example of when to use the median instead of the mean because its distribution tends to be skewed. The median is generally used for skewed distributions.The mean is typically better when the data follow a symmetric distribution. When the data are skewed, the median is more useful because the mean will be distorted by outliers.

Q15-What is the Likelihood**?**

Answer-Likelihood function is a fundamental concept in statistical inference. It indicates how likely a particular population is to produce an observed sample.

The likelihood is the probability that a particular outcome is observed when the true value of the parameter is , equivalent to the probability mass on ; it is not a probability density over the parameter . The likelihood, should not be confused with , which is the posterior probability of given the data . Likelihood is a strange concept in that it is not a probability but is proportional to a probability. The likelihood of a hypothesis (H) given some data (D) is the probability of obtaining D given that H is true multiplied by an arbitrary positive constant

$$\mathbf{K: L(H) = K \times P(D|H).}$$

In most cases, a hypothesis represents a value of a parameter in a statistical model, such as the mean of a normal distribution. Because a likelihood is not actually a probability, it does not obey various rules of probability; for example, likelihoods need not sum to 1
