Q1. A

Q2 A

Q3 C

Q4 D

Q5 C

Q6 B

Q7 B

Q8. A

Q9 C

**Q10. What do you understand by Normal Distribution?**

Normal distribution is where the mean is equal to 0 & the standard deviation is plus or minus 1

**Q11. How do you handle missing data? What imputation techniques do you recommend?**

In case of missing numerical data we can choose to ignore that column or row which have missing data. But this is not the ideal process as we may risk losing valuable information. So instead of ignoring the column or row with missing data, we can use the mean or median of the columns where data is missing and incorporate the mean or median in the missing data. In case of missing string data we can use the mode method of the column in which data is missing and fill the missing data.

Q12 What is A/B Testing?

**Q13. Is mean imputation of missing data an acceptable practise?**

Yes

**Q14. What is linear regression in statistics?**

Linear Regression comes under supervised learning in ML where there will be a y output. The y output will be continuous values. It is used to predict house prices, product price, temperature, fees, electricity bill etc . Linear Regression will try to make the best fit line by joining as many points as possible.

Formula,

Y= a+bx+e

Where, y is the label/output/dependent variable, a=Intercept, b= coefficient of x, x= independent variable/input variable/explanatory variable and e= mean squared error.

**Q15. What are the various branches of statistics?**

There are 2 branches of statistics

1. Descriptive Statistics
2. Inferential statistics

Descriptive statistics is again divided into

1. Central Tendency-In this we have

* Mean
* Median
* Mode

1. Dispersion of data

* Standard deviation
* Range
* Percentile
* Skewness
* Variance

Inferential statistics is divided into

* Hypothesis Testing
* Zscore
* Chisquare
* Anova Test