STATISTICS WORKSHEET-1

Q.1. **a**

Q.2. **a**

Q.3. **c**

Q.4. **a**

Q.5. **d**

Q.6. **a**

Q.7. **b**

Q.8. **a**

Q.9. **c**

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

Ans. The Normal distribution is a continuous probability distribution whose probability density curve is bell shaped and it's a perfect example of probability curve. The distribution has 0 skewness. The normal distribution is symmetrical around it's mean. Hence it has mean=Mode=Median. It is the most important probability distribution in Statistics because it accurately describes the distribution of values for many neural phenomena. Half of the data of Normal distribution is less than Mean and half is greater than Mean.

Q.11. How do you handle missing data? What imputation techniques do you recommend?

Ans :

There are several ways to deal with the missing data:

1. If your dataset is too large and having some missing value then you can delete the row with missing value and do your analysis over the leftover dataset.
2. You can use the average value of the response from the other participants to fill in the missing value in case of quantitative data.
3. In case of categorical data we can fill the missing value by the most common value which is also called as Mode.

Imputation Techniques:-

- .Mean/Median Imputation
- Arbitary value Imputation
- End of Tail Imputation
- Mode Imputation

Q. 12. What is A/B Testing?

Ans:

A/B testing is basically the type of Statistical Hypothesis testing in Statistical Inference. It is an analytical method for making decisions that estimates population parameter based on sample statistics.

Q.13. Is mean imputation of missing data is acceptable practice?

Ans:

In case of continuous data, mean imputation is acceptable. But in case of categorical data we can not use mean imputation.

Q.14. What is Linear Regression in Statistics?

Ans:

Linear regression is the most commonly used method of predictive analysis. It is used to predict the value of one variable based on the value of other variable. The variable which we have to predict is called as Dependent variable and the variable which is used to predict dependent variable is called independent variable.

Q.15. What are the various branches of Statistics?

Ans. There are two main branches of Statistics :

1. Descriptive Statistics : Descriptive statistics is used to get the brief summary of data either in Numerical form or in Graphical form.
2. Inferential Statistics : It is used to make inference and describe about the population. It is more useful when it is not easy or possible to examine each member of population.