STATISTICS

----

1. ***A***
2. ***A***
3. ***B - It is used for modelling unbounded count data***
4. ***D***
5. ***C***
6. ***B***
7. ***B***
8. ***A***
9. ***C***
10. **A Normal Distribution is an arrangement of dataset which prioritises majority of the values in the middle of the given limit and in which rest of the other values are scattered off symetrically scattered in the graph. Ex: For ex a dataset having columns Grade1 and Grade2 shows the distribution of its values in each graph, and the distribution is denoted by a Bell-Curve and usally showing the distribution of the dataset.**
11. **We can handle missing values on the dataset by diferent methods based on its characteristics (1) Deleting Rows with Missing Values (2) Impute missing values for continus variable (3) Impute missing values for categorical variable (4) Prediction of missing values. The code is : df.isull().sum().any() is used to find any empty rows on the given column or the dataset. Understanding the nature of missing data is crucial in confirming**

what technique is to be used to treat the lack of data. The following ways can be used to treat missing values or Null values: (1) Missing Completely At Random (MCAR), (2) Missing At Random (MAR), (3) Not Missing At Random (NMAR).

12. **A/B is a basic randominized controlled experiment. It checks for the best model for the given target variable and to denote the best enviroment for it. For ex assume that you have made a product X and its A part would remain the same and since it is fixed variable, the B part could be tweaked for experiment. So from this we would use B part to try and get the bestfit packing on the product X.**

13. **Mean Imputation of missing data remains unbiased since Imputed values don't really make sense.**

14. **A basic Linear Regression is a model that estimates the relationship between one independant variable and one dependant variable using a straight line Diagonally.**

15. **The FOur most important branches in Statistics in accordance to its flow is: (1) Collecting Data, (2) Choosing the best parameters on best model using HPT (Hyper-parameter-tuning), (3) Statistical Analysis, (4) Descriptive Analysis.**