1. A
2. A
3. B
4. D
5. C
6. B
7. B
8. A
9. C
10. Normal distribution is one of the most important probability distributions of independent random variable. Normal distribution is measured by the probability density function for the continues random variable. Normal distribution shapes like bell, and it's also called as Gaussian Distribution. The parameters like Mean, Median and mode will the same in Normal Distribution. The distribution of Probability density will be symmetrical to both left and right side of the distribution curve. Normal distribution cannot have only one mode.
11. Missing Data handling is one of the important in the preprocessing of a dataset. Most of the cases Missing data are handled by some imputation techniques. Imputation is a technique used for substituting or replacing any missing vales in a dataset. It will not be accurate in prediction and also the model will be tempted to biasing if we remove all the missing feature/label records from the Dataset. There are three types of imputation techniques widely used, **Simple Imputer iterative Imputer and KNN imputer. Simple Imputer** will treat the null values by filling the mean. KNN imputer imputes null values based on the neighborhood values, we need to define how many numbers of Neighbours needs to be considered. Also, **KNN imputer** will try to find the relationship with other continuous features which are having values and impute the value accordingly. The Average of nearest neighbor's continuous feature value are imputed based on the number of neighbors mentioned. In **Iterative imputer,** each feature value is modelled against other features. We need to pass the list of features which Iterative imputer needs to learn. This is a strategy for imputing missing values by modeling each feature with missing values as a function of other features in a round-robin fashion.
12. A/B testing is a randomized control experiment which compares two versions of a variable in a controlled environment based on Hypothesis testing. Example An organization conducting a test for usage of a particular product by making two versions. Both the version of the products is launched to random sampled users for usage. The Null and alternative hypothesis are defined. Say Null hypothesis H0, the change in the product will not make any impact to users and Alternate hypothesis(Ha), the version 2 of the product has some impacts by analyzing the sample study from users. This way null hypothesis can be rejected A/B testing.
13. No, Mean imputation is a bad practice in general, If we are estimating the mean the mean will only be calculated based on the observed, If we impute the mean this cause the model to behave biased also this leads to underestimate the standard deviation. Also this will distort the correlation between variables, other imputation technique like knn and iterative imputer will assess the correlation between features.

14. Linear Regression is one of the most Fundamental and widely known machine learning Algorithm. Building blocks of Linear Regression models are Continuos and Descrete Independent Variables, a Best fit line representing the regression, Continuous Independent Variable. The Linear regression predicts the dependent variable by using the best fit line and the independent variables. The Linear Regression can be mathematically expressed as y=a+bx+e, which a is the intercept and x is the slope, and e is the error term used.
15. There are three real branches of statistics: data collection, descriptive statistics and inferential statistics.