Q1-Q9
Answers
1. a) True
2. a)
3. c)
4. d)
5. c)
6. b)
7. b)
8. a)
9. c)

Q10- Q15
Answers

Q10.
Normal Distribution
Is known by Gaussian distribution. It is a probability distribution which is symmetrical about the mean. So this corresponds to each half being a mirror image to the other half.
It is a continuous distribution.
The total area under the normal curve is equal to 1.
It is also asymptotic to the horizontal axis and is unimodal.
It is beneficial for Model Building.
It consists of different Visualization Techniques: Histograms, kdeplot.

Q11.
Handling Missing Data: The concept of missing data is that it is not stored for some variables in the given dataset. The dataset is missing for possible reasons stated below:
• Past data might get corrupted due to improper maintenance.
• Observations are not recorded for certain fields due to some reasons.(Failure can also be due to human error.
• The user has not provided the values intentionally.
The Imputations of Missing Values:
• Univariate Approach
• Multivariate Approach
• Nearest Neighbours Imputations

Q12.

A/B testing is an optimisation technique often used to understand how an altered variable affects audience or user engagement. It can be a useful when optimising machine learning algorithm, whether testing new models or gaining insight into training data. It can also be used to automate the A/B testing process, making experimentation more efficient over manual processes.

Its type of a split testing and is commonly used to drive improvements to a specific variables or elements by measuring user or audience engagement.

The approach is commonly used to optimise marketing campaigns or digital assets like websites.

Q13

With every method we use there comes a drawback and a plus point. There are pros and cons to each of the methods and resolutions we come across, Mean Imputation has its own, as listed below:

It is easy.

It can be painful to lose a large part of data, so its necessary to be mindful while using mean imputation.

It does not preserve the relationships among variables.

It leads to an underestimate of Standard errors.

So mean imputation comes with its own ways. It is on the persons aim when utilising Mean imputations. As less than 5% of values are missing then its acceptable to ignore them.

Q14

Linear Regression in Statistics

Its a type of Predictive analysis, where the idea of regression analysis is to examine two things : 1) to check whether the predictor variables do a good job. In predicting an outcome variable.

2) which variables in particular are significant predictors of the outcome variable, and in what way do they give the magnitude, etc.

These regression estimates are used to explain the relationships between one dependent variables and one or more independent variables.

Q15

Branches of Statistics

The two main branches of statistics are:

Descriptive statistics and Inferential statistics

Descriptive Statistics:

Descriptive statisticsdeals with the presentation and collection of data. This needs for the statistician to be sure of designing experiments, choosing the right focus group and avoid biases.

Inferential Statistics:Inferential Statistics, it involves drawing the right conclusions from the statistical analysis that has been performed using descriptive statistics. Most predictions of the future and generalisations about a population by studying a smaller sample come under the purview of inferential statistics.