

Project 2 Theory answers

Q1 Types of Data Numerical & categorical

- Numerical Data is ① Discrete and ② Continuous Data
- Discrete means the countable number not in Decimal for eg:- Age, Mobile number
- Continuous means the number which can be in Decimal for eg:- Temperature, Score
- Categorical Data is Nominal and Ordinal
- Nominal Names/age, like male/female
- Ordinal means ordered category Low/Medium/High Ratings

Q-1 Types of Statistics: Descriptive and Inferential

- > Descriptive Data To Summarise Data
Using measures of central tendency and
measure of Dispersion, Shape of
Distribution and graphs
- > In Inferential we can Define Hypotheses
Testing Based on p-values, A/B Testing,
Null & Alternative Hypotheses

Q-2 What is Descriptive Statistics

- > Descriptive Statistics means from a large
Number of data (Population) we can
take a Sample group of data and
the we can summarise and visualise
Data

In the Central Tendency

- mean
 - median
 - mode
- we can use

Measure of Dispersion

- Range
- Variance
- Standard Deviation

Graphs like

- Bar graph, Pie chart, Line chart etc.

Shape of Distribution like

- Skewness
- Kurtosis ~~case~~

We can see

Can Difference between

→ Mean :- In mean we can take the Average of data or a sample data

Median :- In median we can find the middle value of data

Mode :- In mode we can find most repeated value

Range :- We can see how much data is spread

Variance:- In the we can find how far each value is from mean

Standard deviation:- It shows the typical deviation from the mean in original units

Q-5 • Gaussian Distribution

→ It is a bell shape curve normal distribution graph which shows by how far each data lies from mean

• Lognormal Distribution

→ when the ~~the~~ Data is right skewed and values are not negative is called Lognormal

• Percentiles

• Percentile Show by the which data is in which position and how much data falls at the point

$$P = \left(\frac{\text{rank}}{\text{total}} \right) \times 100$$

• Quartiles

- Quartiles means Q_1, Q_2, Q_3 where Q_1 is 25% of data, Q_2 is 50% of data and Q_3 is rest 25% of data. Q_2 is called median data & Q_1

• Five number summary

- minimum
- Q_1
- Q_2 Median
- Q_3
- maximum

• Skewness

→ This Describes the asymmetry of a data distribution around its mean

- Kurtosis is

- The measure the Curve / Tail of data is ~~heavy~~ ~~lighter~~ heavier or lighter from normal Distribution