

Project of Theory answers

(Q1) Types of Data Numerical & Categorical

→ Numerical Data is Discrete and Continuous Data

→ Discrete means the countable numbers not in decimal for e.g.: age, mobile number

→ Continuous means the number which can be in decimal for e.g.: Temperature, Score

→ Categorical Data is Nominal and Ordinal

→ Nominal Numbers like male/female

→ Ordinal means orders category Low/medium/High/Ratings

(Q-1) Types of Statistics: Descriptive and
Inferential

-> Descriptive Data To summarize Data
using means or central tendency and
measures of dispersion, shape of
Distribution and graphs

-> In inferential we can Define hypothesis
Testing Based of p-values, A/B Testing,
Null & Alternative Hypotheses

(Q-2) What is Descriptive Statistics

-> Descriptive statistics means from a large
numbers of data (Population) we can
take a sample group of data and
we can summarize and visualize
Data

In the Central Tendency

- mean
- median
- mode

we can use

Measure of Dispersion

- Range
- Variance
- Standard Deviation

Graphs like

Box graph, pie chart, line chart etc

Shape of Distribution like

- Skewness
- Kurtosis

We consider

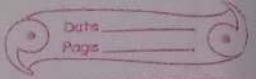
One difference between

→ 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 ~~Data~~ 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

- Examples: mean (\bar{x}), Q_1 , Q_2 , Q_3 where Q_1 is 25% of data, Q_2 is 50% of data and Q_3 is next 25% of data. Q_2 is called median data as well.

- Five Number Summary

- minimum
- Q_1
- Q_2 Median
- Q_3
- maximum

- Skewness

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

- There is 3

- The mean, the curve / tail of data is ~~too steep~~ ~~less frequent~~ lighter than heavy and lighter than normal Distribution