

PART A – THEORY (Short Questions)

1. Define Mean, Median, and Mode with an example (Salary)

Mean:

Mean is the average of all values.

$$\text{Mean} = \frac{\text{Sum of all values}}{\text{Number of values}}$$

Example:

If employee salaries are ₹30,000, ₹40,000, ₹50,000

$$\text{Mean} = (30000 + 40000 + 50000) / 3 = ₹40,000$$

Median:

Median is the middle value when data is arranged in ascending order.

Example:

Salaries: ₹30,000, ₹40,000, ₹50,000

$$\text{Median} = ₹40,000$$

Mode:

Mode is the value that occurs most frequently.

Example:

Salaries: ₹30,000, ₹40,000, ₹40,000, ₹50,000

$$\text{Mode} = ₹40,000$$

2. Difference between Range and Variance

Range	Variance
Difference between maximum and minimum value	Average of squared deviations from mean
Simple measure of spread	Detailed measure of dispersion

Range	Variance
Affected by extreme values	Uses all data points

Formula:

- Range = Max – Min
 - Variance = $\frac{\sum(x-\bar{x})^2}{n}$
-

3. Difference between Normal Distribution and Poisson Distribution

Normal Distribution Poisson Distribution

Continuous Discrete

Bell-shaped curve Right-skewed

Used for salary, height Used for number of events

Defined by mean & SD Defined by λ (rate)

4. Explain Skewness with workplace example

Skewness measures asymmetry of a distribution.

- **Positive Skewness:** Long tail on right
Example: Few employees earn very high salaries.
 - **Negative Skewness:** Long tail on left
Example: Most employees earn high salaries, few earn low.
 - **Zero Skewness:** Symmetrical distribution
-

5. Define Conditional Probability and explain with example

Conditional probability is the probability of an event occurring **given that another event has occurred.**

Formula:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

Example:

Probability that an employee is promoted **given** performance score > 80.

6. What is Bayes' Theorem? Explain its use

Bayes' Theorem helps update probabilities based on new information.

Formula:

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

Use in real life:

Used in promotion decisions, spam detection, medical diagnosis, and risk assessment.

7. Write a short note on PCA (Principal Component Analysis)

PCA is a dimensionality reduction technique used to convert large datasets into fewer variables while retaining maximum information.

Advantages:

- Reduces complexity
- Removes multicollinearity
- Improves model performance

Used in data analysis and machine learning.