SNS COLLEGE OF ENGINEERING

Kurumbapalayam(Po), Coimbatore – 641 107
Accredited by NAAC-UGC with 'A' Grade
Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

Department of Information Technology

Course Name – 23ADT202 Fundamental of Data science and Analytics

II Year / IV Semester

Unit 2 – Descriptive Analytics

Interpreting Distributions

Introduction to Distributions

Distributions show how data points are spread across values. They allow us to understand frequency, patterns, and anomalies. Analyzing distributions helps us draw insights from data. This is key in descriptive analytics.

What is a Distribution?

A distribution represents how data points spread across values.
It shows how often each value or range appears in the dataset.
Example: Test scores [85, 90, 75, 85, 80, 90, 95, 75, 80, 85].
A histogram visualizes the frequency of these values.

Key Properties of Distributions

1. Central Tendency: Measures data's center (Mean, Median, Mode).
2. Spread (Variability): Describes data variability (Range, SD).
3. Shape: Symmetry or skewness of data.
4. Outliers: Extreme values outside general patterns.

Central Tendency

Central tendency shows where data is centered.
Mean, median, and mode are the common measures.
Example: In [85, 85, 90, 95, 100], mean = 91.
It helps us understand the general "average" of data.

Spread (Variability)

Spread describes how much the data varies.
Common metrics include range, variance, and standard deviation.
Example: In [85, 85, 90, 95, 100], range = 15.
It gives insights into the data’s consistency or variability.

Types of Distributions

Normal: Bell-shaped curve, symmetric around the mean.
Uniform: All values occur with equal frequency.
Skewed: Right or left-skewed based on tails.
Bimodal: Two peaks indicating distinct groups.

Visualizing Distributions

Histograms show frequency of data points in intervals.
Boxplots summarize distributions and highlight outliers.
Density plots show probability density of data.
Visuals help in interpreting the underlying distribution patterns.

Practical Applications

Business: Analyze sales distributions for trends.
Education: Understand grade distributions for exam assessment.
Healthcare: Analyze patient data for treatment planning.
Distributions provide insights for various industries.

THANK YOU