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

Frequency distributions

What is a Frequency Distribution?

A frequency distribution organizes data into categories or intervals and shows how often each category occurs. It is used to summarize large datasets, making it easier to identify patterns, trends, and outliers in the data.

Example of a Frequency Distribution

For a dataset of student grades: [85, 90, 75, 85, 80, 90, 95, 75, 80, 85], the frequency distribution is as follows:

- Grade 75: 2 occurrences
- Grade 80: 2 occurrences
- Grade 85: 3 occurrences
- Grade 90: 2 occurrences
- Grade 95: 1 occurrence

Types of Frequency Distributions

1. **Absolute Frequency**: The number of times a value appears in the dataset.
2.Relative Frequency: The proportion of the total occurrences represented by a specific value.
3.Cumulative Frequency: A running total of frequencies as you progress through the dataset.

Calculating Absolute Frequency

The absolute frequency shows how often each value occurs in the dataset. For instance, in the dataset [85, 90, 75, 85, 80, 90, 95, 75, 80, 85], grade 85 appears 3 times. This value is listed in the frequency distribution table.

Relative Frequency Calculation

Relative frequency shows the proportion of occurrences of each value relative to the total number of observations.

Example: For grade 85: Relative Frequency = 3 occurrences / 10 total observations = 0.3 or 30%.

Cumulative Frequency Calculation

Cumulative frequency is the sum of frequencies up to a certain point. For grade 85, the cumulative frequency is calculated by adding all previous frequencies.
Example: Cumulative frequency for grade 85 = 2 (for 75) + 2 (for 80) + 3 (for 85) = 7.

Steps to Construct a Frequency Distribution

Organize the data in ascending order.
Define categories or intervals (for continuous data).
Count the occurrences of each value or interval.
Create a table to display frequencies.
Example: Sorted dataset = [75, 75, 80, 80, 85, 85, 85, 90, 90, 95].

Visualization of Frequency Distributions

Bar Charts: Ideal for categorical data.

Histograms: Used for continuous data, displaying intervals or bins.

These visualizations help to clearly communicate the frequency distribution and its patterns.

Practical Applications

Business: Analyze sales data to identify trends.
Education: Assess grade distributions to evaluate student performance.
Healthcare: Track patient recovery times and outcomes.
Frequency distributions help in decision-making across various sectors.

THANK YOU