SNS COLLEGE OF TECHNOLOGY

An Autonomous Institution

Accredited by NBA – AICTE and Accredited by NAAC – UGC with 'A' Grade
Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai

Department of Computer Science and Engineering

Course Code & Title : 23AD0201 - Data Science Fundamentals

III YEAR / VI SEMESTER - EEE
Unit 2 - DESCRIPTIVE ANALYTICS

Topic : FREQUENCY DISTRIBUTIONS

Mr.K.Karthikeyan AP /CSE

Empathy in Understanding Frequency Distributions

Understanding Frequency Distributions

Applications
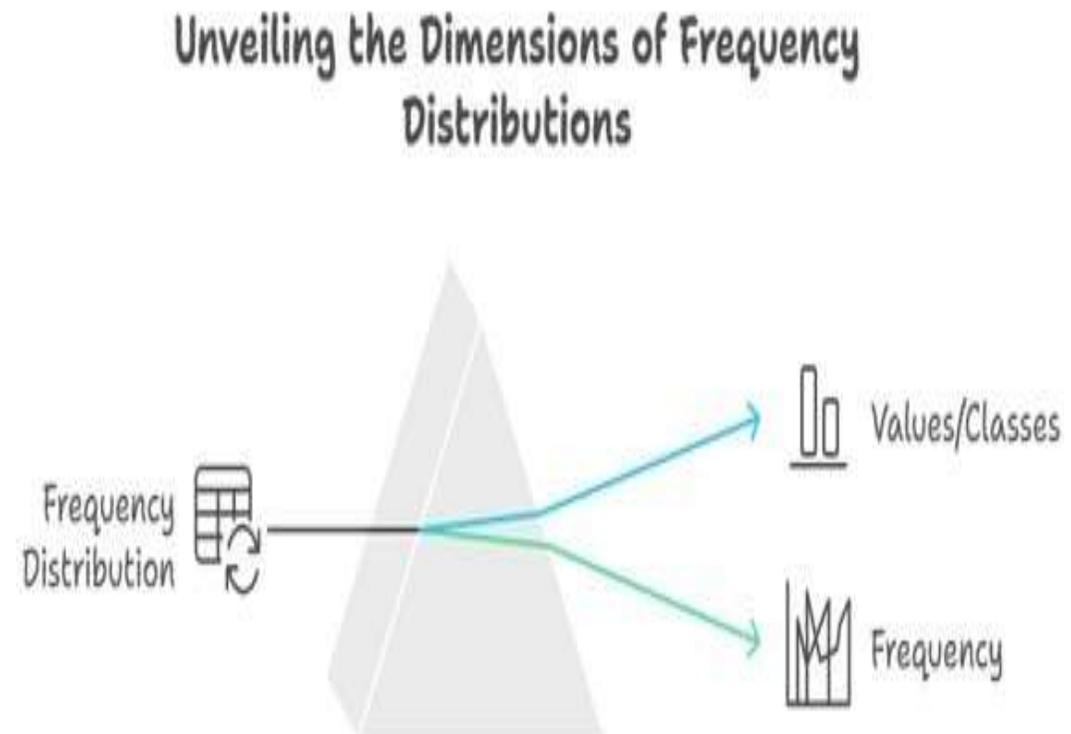How frequency distributions are used in analysis

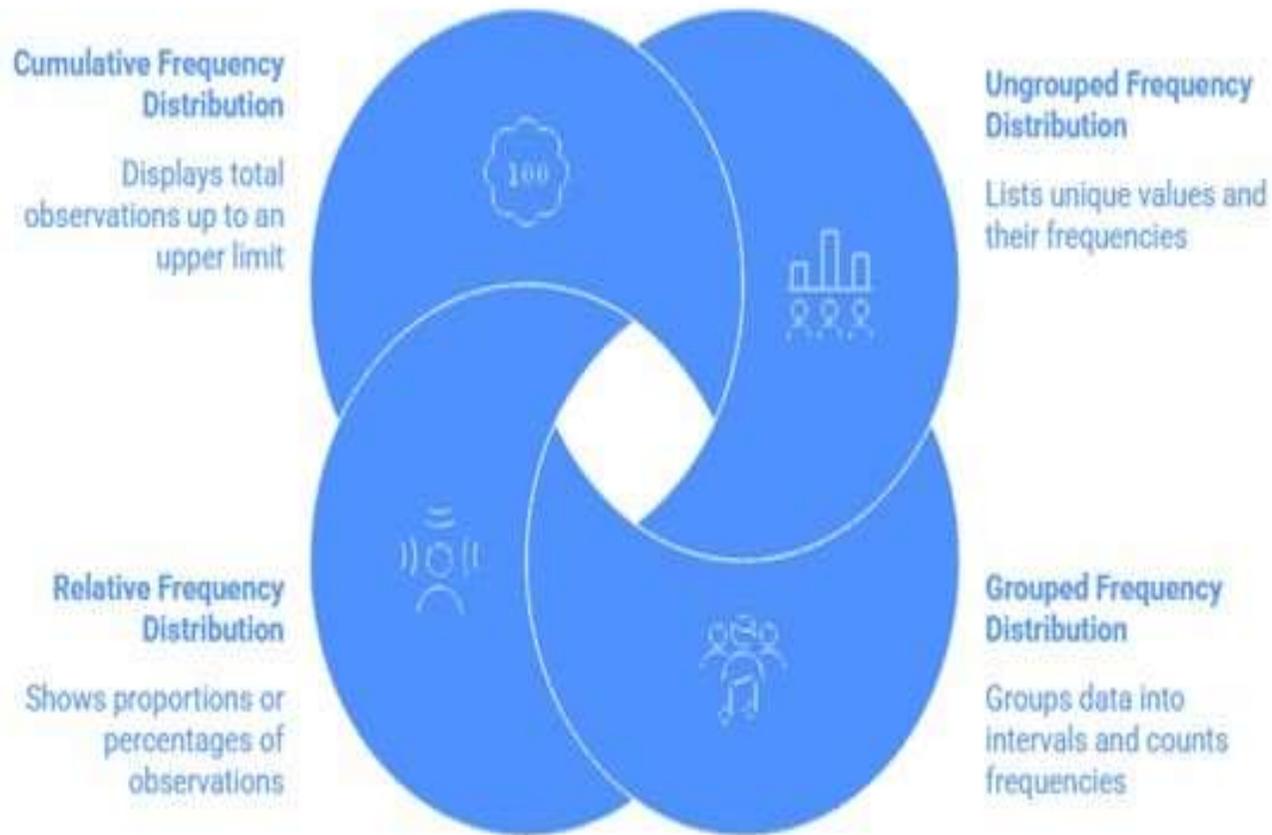**Definition**
A way to organize and summarize data

Construction
Steps to create a frequency distribution

Types
Different forms of frequency distributions

- A **frequency distribution** is a table or chart that summarizes the values and their frequencies for a set of data. It shows how often each value or range of values occurs in the dataset.
- In simpler terms, it's a way of organizing data to see the pattern of how frequently different values appear.

Unveiling the Dimensions of Frequency Distributions

Frequency Distribution

Values/Classes

Frequency

Overview of Frequency Distributions

Cumulative Frequency Distribution

Displays total observations up to an upper limit

Ungrouped Frequency Distribution

Lists unique values and their frequencies

Relative Frequency Distribution

Shows proportions or percentages of observations

Grouped Frequency Distribution

Groups data into intervals and counts frequencies

Example Sum : Constructing a Grouped Frequency Distribution

Let's say we have the following dataset representing the scores of 30 students on a test:

65, 70, 72, 75, 78, 80, 82, 85, 88, 90, 92, 95, 68, 73, 76, 79, 81, 83, 86, 89, 91, 93, 96, 67, 71, 74, 77, 84, 87, 94

1.Range: 96 - 65 = 31

2.Number of Classes: Let's choose 6 classes.

3.Class Width: 31 / 6 ≈ 5.17. We'll round this up to 6.

4.Class Boundaries:

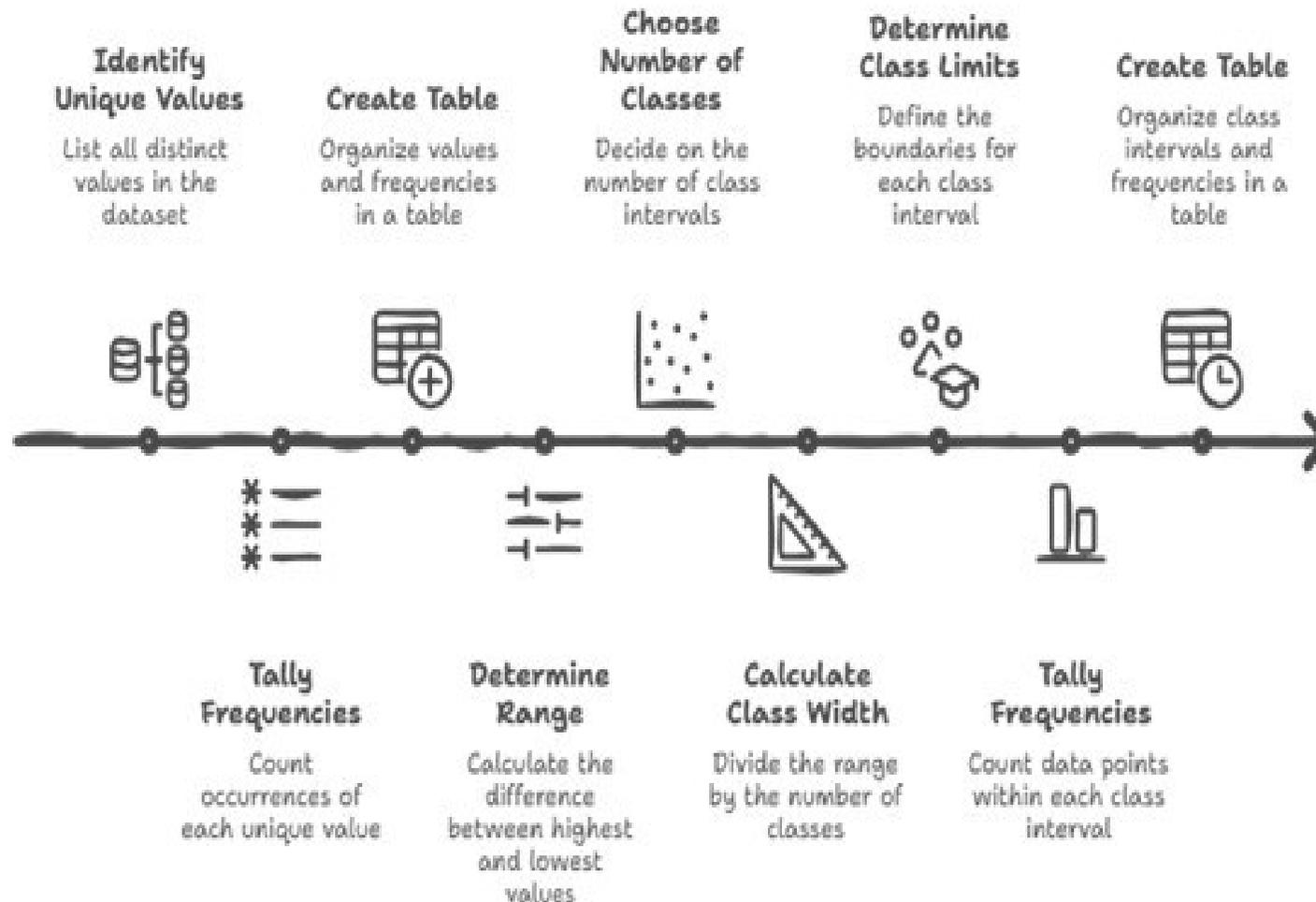1. 65-70
2. 71-76
3. 77-82
4. 83-88
5. 89-94
6. 95-100

5.Tally Frequencies:

* 65-70: 4
* 71-76: 5
* 77-82: 6
* 83-88: 5
* 89-94: 6
* 95-100: 4

6.Frequency Distribution Table:

Class	Frequency	Relative Frequency	Cumulative Frequency
65-70	4	0.133	4
71-76	5	0.167	9
77-82	6	0.200	15
83-88	5	0.167	20
89-94	6	0.200	26
95-100	4	0.133	30

Constructing a Frequency Distribution

Identify Unique Values
List all distinct values in the dataset

Create Table
Organize values and frequencies in a table

Choose Number of Classes
Decide on the number of class intervals

Determine Class Limits
Define the boundaries for each class interval

Create Table
Organize class intervals and frequencies in a table

Tally Frequencies
Count occurrences of each unique value

Determine Range
Calculate the difference between highest and lowest values

Calculate Class Width
Divide the range by the number of classes

Tally Frequencies
Count data points within each class interval

FREQUENCY DISTRIBUTIONS – FUN ACTIVITIES

1. Collect & Classify 📋

Time: 10–15 minutes

Mode: Individual → Group

Steps:

•Ask students to collect one variable (e.g., marks, heights, mobile usage hours).

•Write raw data on the board.

•Students:

- Decide class intervals

- Create a frequency table

✅ **Outcome:** Understanding raw data → frequency distribution

Visualizing Frequency Distributions

Histogram

Bar graph with adjacent bars representing class frequencies

Frequency Polygon

Line graph connecting midpoints of histogram bars

Bar Chart

Separated bars for categorical data
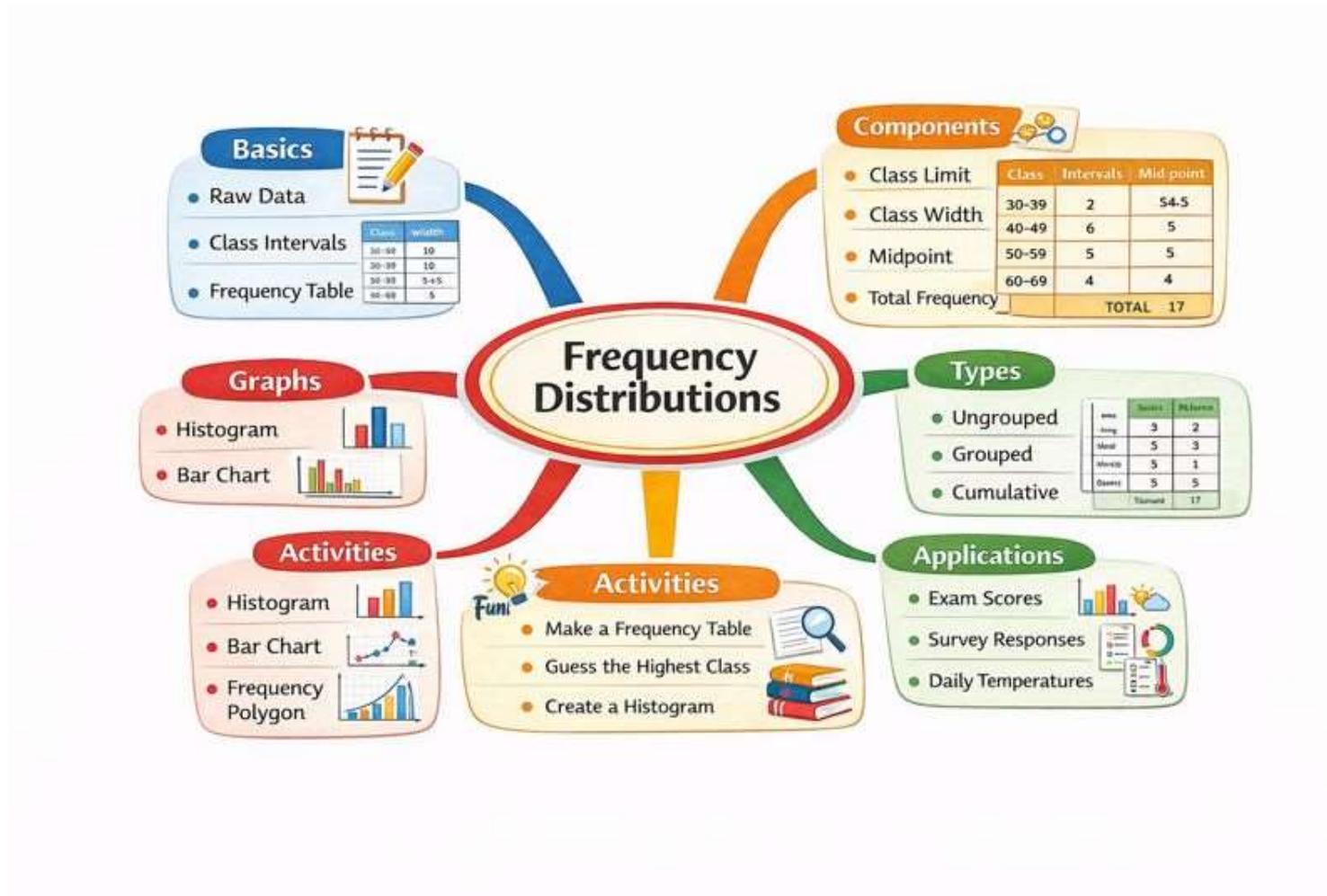
Pie Chart

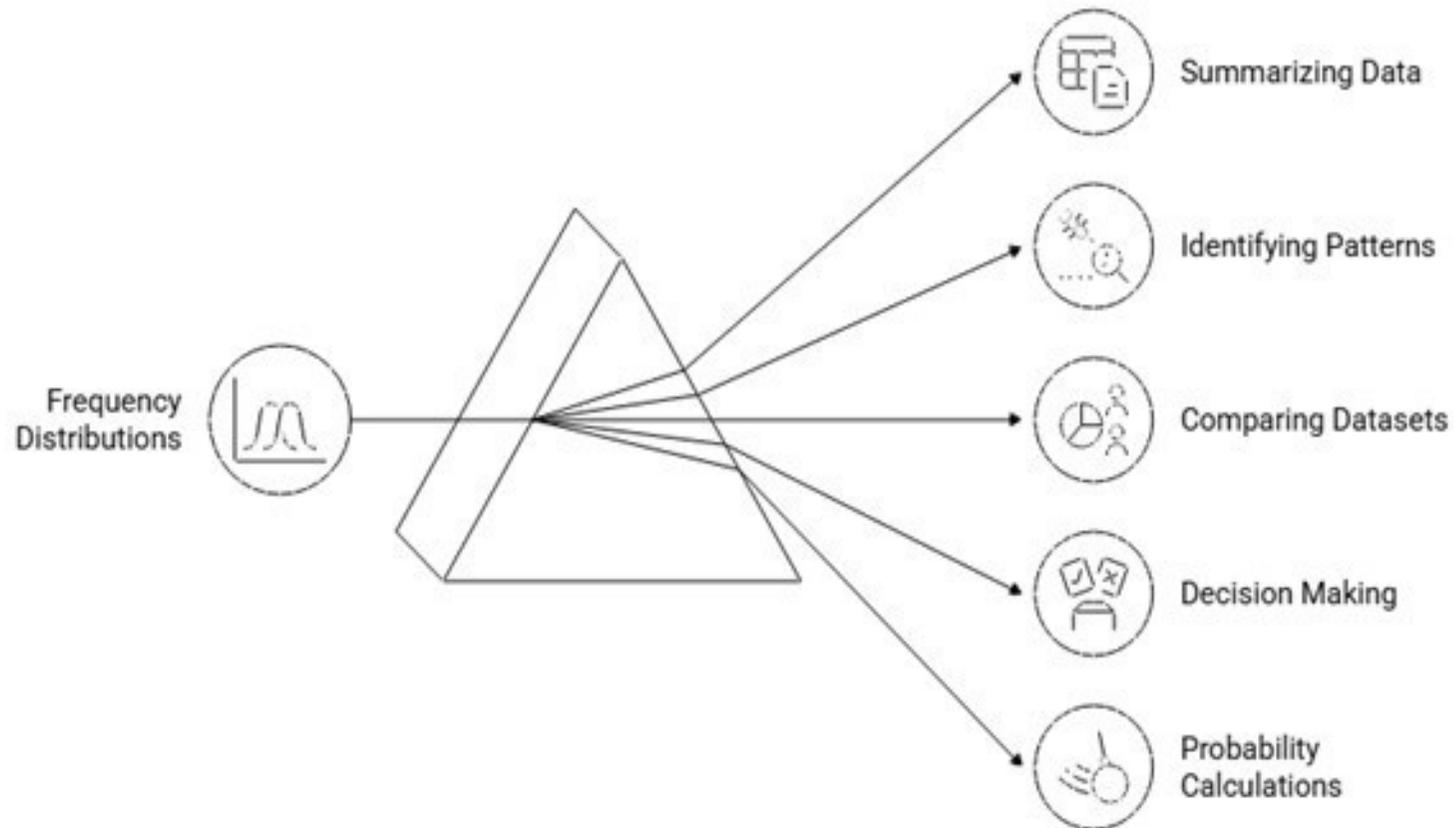Circular chart with slices for relative frequencies

Ogive

Line graph of cumulative frequency against upper limits

MIND MAP

FREQUENCY DISTRIBUTIONS / 23ADO201 - DATA SCIENCE FUNDAMENTALS / Mr.K.KARTHIKEYAN/CSE/SNSCT

Unveiling the Multifaceted Role of Frequency Distributions

Frequency Distributions

- Summarizing Data
- Identifying Patterns
- Comparing Datasets
- Decision Making
- Probability Calculations

ASSESSMENT

FREQUENCY DISTRIBUTIONS – MCQs

1. Basics

1. A frequency distribution is:

A. A graphical representation of data

B. A table showing how often values occur

C. A summary using averages only

D. A random arrangement of data

✅ **Answer:** B

2. Raw data refers to:

A. Grouped data

B. Processed data

C. Original collected data

D. Summarized data

✅ **Answer:** C

REFERENCE BOOKS

1.Allen B. Downey, "Think Stats: Exploratory Data Analysis in Python", Green Tea Press, 2014.
2.Sanjeev J. Wagh, Manisha S. Bhende, Anuradha D. Thakare, "Fundamentals of Data Science", CRC Press, 2022.
3.Chirag Shah, "A Hands-On Introduction to Data Science", Cambridge University Press, 2020.
4.Vineet Raina, Srinath Krishnamurthy, "Building an Effective Data Science Practice: A Framework to Bootstrap and Manage a Successful Data Science Practice", A press, 2021.