



No:-

Date:

MC460305:

Data Preparation and Visualization

L-T-P-Cr: 3-0-2-3

Prerequisite – None

Course Objectives:

1. To understand the basics of data preprocessing techniques.
2. To understand the basics of data visualization.
3. To Learn various techniques for visualizing the data.
4. To implement data visualization tools.
5. To learn about methods of visualizing distributions, associations among two or more quantitative variables, and different types of data.

Course Outcomes:

After completion of the course, students will be able to:

1. Understand basic types of data and data preprocessing techniques.
2. Understand and analyze the visualization for data and distributions.
3. Identify and develop the different tools for visualizing the data.
4. Understand how to visualize complex datasets.
5. Understand the different Image file formats.

Sl. No	Course Outcome (CO)	Mapping to PO
1	Understand basic types of data and data preprocessing techniques.	PO1, PO2
2	Understand and analyse the visualization for data and distributions.	PO1, PO2, PO6, PO8
3	Identify and develop the different types of tools for visualizing the data.	PO1, PO3, PO6, PO8
4	Understand how to visualize the complex datasets.	PO1, PO2, PO6, PO8
5	Understand the different Image file formats.	PO1, PO6

UNIT I

Lectures: 8

Introduction to Data: The Basic Data Types – Nominal, Ordinal, interval, Ratio scaled.

Basic Statistical Descriptions of Data: Measuring the Central Tendency, Measuring the Dispersion of Data, Statistical Descriptions of Data.

Non-dependency-oriented data: Quantitative, multidimensional data, categorical and mixed attribute data, binary and set data, and text data.

Dependency-Oriented Data: Time-Series Data, Discrete Sequences and Strings, Spatial Data, Network and Graph Data.

UNIT-II

Lectures10

Data Preprocessing: An Overview, Data Cleaning: Missing Values, Noisy Data Data Cleaning as a Process; Data Integration: Entity Identification Problem, Redundancy and Correlation Analysis, Tuple Duplication; Overview of Data Reduction Strategies; Data Transformation and Data Discretization

UNIT-III

Lectures:8

Why visualize Data? History of Visualization, The Visualization Process. Visualizing Data: Mapping Data onto Aesthetics, Coordinate Systems and Axes, Colour scales, Directory of Visualizations,

UNIT-V

Lectures:8

Visualizing Amounts, Visualizing Distributions and visualizing many Distributions at once, Visualizing Associations among Two or More Quantitative Variables, Visualizing Time series, and other functions of an independent variable.

UNIT-VI

Lectures:8

Visualize Trends, Geospatial data, and uncertainty, Visualization of Networks and Trees, visualize multidimensional data, Understand the most commonly used Image file formats, and choose the right visualization software.

Text Books:

1. Jiawei Han and Micheline Kamber, Data Mining- Concepts and Techniques-Morgan Kaufmann Publishers, Elsevier, 2nd Edition, 2006.
2. Claus O. Wilke, Fundamentals of Data Visualization, O'Reilly publication, first Release Edition.
3. Tamara Munzner. Visualization Analysis and Design. A K Peters Visualization Series, CRC Press, 2014
4. Kieran Healy, Data Visualization: A Practical Introduction 1st Edition, Princeton University Press.