

NIRMA UNIVERSITY

Institute:	Institute of Technology
Name of Programme:	B.Tech. All (Other than CSE)
Course Code:	XXXX
Course Title:	Data Analysis and Visualisation
Course Type:	Inter-disciplinary Minor-Core
Year of Introduction:	2024-25

L	T	Practical Component				C
		LPW	PW	W	S	
3	0	2	-	-	-	4

Course Learning Outcomes (CLO):

At the end of the course, students will be able to –

1. demonstrate data characteristics using visualisation tools (BL2)
2. identify common data types and corresponding analysis approaches (BL3)
3. analyse the data using various statistical tools (BL4)
4. build data visualisation systems for interdisciplinary problems (BL6)

Unit	Contents	Teaching Hours (Total 45)
Unit-I	Introduction: Data Understanding, types of data, information and uncertainty, classes and attributes, interactions among attributes, relative distributions, summary statistics. Data Quality: inaccurate data, sparse data, missing data, insufficient data, imbalanced data	10
Unit-II	Definition, Purpose, Usage, Business Data Visualization: Features of Business Data, Different Visualization fields. Forms of Business Data Visualization. Social Challenges: Data ownership, data security, ethics and privacy	10
Unit-III	The Data: Data Examination, Data Visualization Patterns, the Categories of Data Visualization. Data Visualization using different tools: Refine data and create, edit, alter, and display their visualizations (x-y graph, bar chart, pie chart, cube etc)	10
Unit-IV	Data Reduction and Feature Enhancement: Standardizing data, sampling data, using principal components to eliminate attributes, limitations and pitfalls of principal component analysis (PCA), curse of dimensionality	10
Unit-V	Showing Complex Data: Organizational Models, Preattentive Variables, Sorting and Rearranging, Searching and Filtering, Datatips, Data Spotlight, Dynamic Queries, Data Brushing, Local Zooming, Sortable Table, Radial Table, Muti-Y Graphs, Treemap, Small Multiples	05

Self-Study: The self-study contents will be declared at the commencement of semester. Around 10% of the questions will be asked from self-study contents

Suggested Readings/
References:

1. Jack G. Zheng, Data Visualization for Business Intelligence, Taylor and Fransis
2. Jiahei Han & Micheline Kamber, Data Mining Concepts and Techniques, Morgan Kaufmann
3. Jenifer Tidwell, Designing Interfaces, O'Reilly Media, inc.
4. Edward Tufte, The Visual Display of Quantitative Information, Graphics Press LLC.
5. Ben Fry, Visualizing Data, O'Reilly Media inc.
6. Noab Iliinsky, Julie Steele, Designing Data Visualization, O'reilly Media inc.
7. Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, Pearson
8. Wes McKinney, Python for Data Analysis, Oreilly
9. S. Nagabhushana, Data Warehousing OLAP and Data Mining, New Age publishers

Suggested List of Experiments:

Sr. No.	Title	Hours
1	(a) Aim: Data Domain selection and Identification of Characteristics of selected Dataset of different formats. 1. What data domain you have selected? 2. What are the information dataset contains? 3. Identify the characteristics of various fields of the dataset. (The distribution, inference etc.) 4. What are the insight (knowledge) we can generate for the selected dataset? 5. What are the pattern available in the dataset? (b) For selected dataset generate Five Number Summary using Python. Also generate mode and midrange, outlier detection using concept of Quartile method and other. Compare the results.	06
2	Case Study for Data visualization using Tableau. Use the dataset selected in practical -1 and design an interactive Dashboard for analysing data for selected KPI.	04
3	Data Preprocessing (Data Quality): Aim: Implement data smoothing and data normalization methods. Redundancy analysis using Pearson correlation and Chi- Square. Discretization by Intuitive Partitioning.	04
4	Data Reduction and Feature Enhancement: Aim: Implement Dimensionality reduction and Feature selection technique with selected dataset.	04
5	Data Analysis and Visualization Aim: Apply Classification Techniques for selected datasets visualize the results.	04

- | | | |
|---|---|----|
| 6 | Data Analysis and Visualization | 04 |
| | Aim: Apply Clustering Techniques for selected datasets and visualize the results. | |
| 7 | Pattern Analysis and Visualization | 04 |
| | Aim: Implementation of various classification and regression techniques and visualize the result with selected visualization tool like Tableau. | |

Suggested Case
List:

-NA-

