Course Code:DS8002   Course Title: Data Anlaytics and Visulization

Course Credits: [Lecture: 02, Tutorial: 00, Practical: 01]

Prerequisites: A basic understanding of Python programming and knwoldege of dataset.

Objectives: To equip students with foundational and practical knowledge of Data Analytics, covering data handling, transformation, analysis, and visualization.

Course Outcomes: Upon completion of the course, students shall be able to

CO1:	To understand the fundamentals and importance of Data Analytics, its roles across domains, and distinctions between Analytics, Data Science, and Machine Learning
CO2:	To observe, explore, and interpret patterns and relationships in datasets through variable analysis and outlier detection techniques.
CO3:	To perform data preparation and transformation tasks including handling missing values, text cleaning, binning, normalization, encoding, and reshaping datasets.
CO4:	To apply statistical thinking, sampling, hypothesis testing, and evaluate models using concepts like bias, variance, and overfitting.
CO5:	To design impactful data stories and apply visualization techniques such as bar plots, line charts, heatmaps, and box plots using appropriate styling.
CO6:	To examine ethical aspects of data collection and analysis while understanding current trends including fairness in algorithms and real-time analytics.

Course Objective and Course Outcomes Mapping:

- To gain a comprehensive understanding of introduction and visualization of data science CO1, CO2, CO3
- To explore predictive moodel using advance analytical coneptsCO4,C05, C06

Programme Outcomes:The student will have

PO1-Computational Knowledge: Apply knowledge of computing fundamentals, computing specialisation, mathematics, and domain knowledge appropriate for the computing specialisation to the abstraction and conceptualisation of computing models from defined problems and requirements.

PO2-Problem Analysis: Identify, formulate, research literature, and solve complex computing problems reaching substantiated conclusions using fundamental principles of mathematics, computing sciences, and relevant domain disciplines.

PO3-Design /Development of Solutions: Design and evaluate solutions for complex computing problems, and design and evaluate systems, components, or processes that meet specified needs with appropriate consideration for public health and safety, cultural, societal, and environmental considerations.

PO4-Conduct Investigations of Complex Computing Problems: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of t h e information to provide valid conclusions.

PO5-Modern Tool Usage: Create, select, adapt and apply appropriate techniques, resources, and modern computing tools to complex computing activities, with an understanding of the limitations.

PO6-Professional Ethics: Understand and commit to professional ethics and cyber regulations, responsibilities, and norms of professional computing practices.

PO7-Life-long Learning: Recognise the need, and have the ability, to engage in independent learning for continual development as a computing professional.

PO8-Project management and finance: Demonstrate knowledge and understanding of the computing and management principles

and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

PO9-Communication Efficacy: Communicate effectively with the computing community, and with society at large, about complex computing activities by being able to comprehend and write effective reports, design documentation, make effective presentations, and give and understand clear instructions.

PO10-Societal and Environmental Concern: Understand and assess societal, environmental, health, safety, legal, and cultural issues within local and global contexts, and the consequential responsibilities relevant to professional computing practices.

PO11-Individual and Team Work: Function effectively as an individual and as a member or leader in diverse teams and in multidisciplinary environments.

PO12-Innovation and Entrepreneurship: Identify a timely opportunity and using innovation to pursue that opportunity to create value and wealth for the betterment of the individual and society at large.

Programme Outcomes and Course Outcomes mapping:

Course Outcomes	Programme Outcomes											
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	P10	P11	P12
CO1	3	2	0	0	2	0	2	0	0	0	0	0
CO2	3	3	0	0	2	0	2	0	0	0	0	0
CO3	3	2	0	0	2	0	2	0	0	0	0	0
CO4	3	2	0	2	2	0	2	0	0	0	0	0
CO5	3	2	0	2	2	0	2	0	0	0	0	0
CO6	3	2	0	2	2	2	3	3	3	2	3	3

Curriculum Unit Titles

Units	Unit Description	Weightage
1	Foundations of Data Analytics	[10%]

1.1.   Basics and Importance of Data Analytics in Today's World

1.2.   Roles and Applications of Analytics in Different Domains

1.3.   Key Differences between Data Analytics, Data Science, and Machine Learning

1.4.   Overview of Python Libraries and Tools for Data Handling

2	Data Observation and Exploration	[20%]

2.1.   Techniques to Observe and Understand Patterns in Data

2.2.   Procedures for Exploring Relationships Between Variables

2.3.   Understanding Shape and Spread of Data

2.4.   Summary to Outlier Identification Approaches

3	Data Preparation and Transformation	[20%]

3.1.   Handling Missing Data using Imputation and Deletion Techniques

3.2.   Cleaning and Standardizing Text Data with Regular Expressions

3.3.   Feature Transformation: Binning, Normalization, and Encoding

3.4.   Data Integration: Merging, Concatenation, and Reshapingand Restructuring Datasets for Better Use

4	Statistical Thinking and Analysis	[20%]

4.1.   Sampling Methods and Estimation Techniques

4.2.   Confidence Intervals and Margin of Error Interpretation

4.3.   Hypothesis Testing: t-Test, ANOVA, and Chi-Square

4.4. Model Evaluation Concepts: Bias, Variance, Misinterpretation and Overfitting in Analysis

5 Data Storytelling and Visualization [15%]

5.1. Best Practices for Designing Visual Narratives

5.2. Visualization Techniques: Bar, Line, Heatmap, and Box Plots

5.3. Customizing Visuals with Labels, Legends, and Styles

6 Ethics and Emerging Trends in Analytics [15%]

6.1. Principles of Ethical Data Collection and Use

6.2. Addressing Algorithmic Bias and Promoting Fairness

6.3. Overview of Trends: Real-Time Analytics

Course Units and Course Outcomes Mapping:

Unit No.	Unit	Course Outcomes					
		CO1	CO2	CO3	CO4	CO5	CO6
1	Foundations of Data Analytics	✓					
2	Data Observation and Exploration		✓				
3	Data Preparation and Transformation			✓			
4	Statistical Thinking and Analysis				✓		
5	Data Storytelling and Visualization					✓	
6	Ethics and Emerging Trends in Analytics						✓

Computing Environment:

A student must have the following computing environment in the laboratory and/or on his/her laptop.

- CE#1 Python for all the 6 units with Numpy, Pandas and Matplotlib

Text Books

1. Wes McKinney, *Python for Data Analysis*, 2nd Edition, O'Reilly Media..

2. Jake VanderPlas, *Python Data Science Handbook*, O'Reilly Media.

References:

1. Marek Gagolewski, *Minimalist Data Wrangling with Python*.

2. Joel Grus, *Data Science from Scratch: First Principles with Python*, 2nd Edition, O'Reilly Media.