Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous), Shivajinagar, Pune 5

(An Autonomous College Affiliated to Savitribai Phule Pune University)

Framework of Syllabus

For

B. Sc. (Data Science)

Choice Based Credit System (CBCS) Syllabus

Under National Education Policy (NEP)

To be implemented from Academic Year 2025-2026

Level:- 4.5 (First Year) **Sem:I**

Course Type	Course Code	Course Title	Credits		Teaching Scheme Hr/Week		Evaluation Scheme and Max Marks		
			TH	PR	TH	PR	CE	ESE	Total
Subject 1 T(2)+ (T/P)(2) or T(4)	24CsDscU1101	Problem Solving and Python Programming	2		2		20	30	50
	24CsDscU1102	Lab Course on 24CsDscU1101		2		4	20	30	50
Subject 2 T(2)+ (T/P) (2) or T(4)	24CsDscU1201	Descriptive Statistics	2		2		20	30	50
	24CsDscU1202	Lab Course on 24CsDscU1201		2		4	20	30	50
Subject 3 T(2)+ (T/P)(2) or T(4)	24CsDscU1301	Computational Mathematics	2		2		20	30	50
	24CsDscU1302	Lab Course on 24CsDscU1301		2		4	20	30	50
IKS T(2)	24CpCopU1901	Generic IKS	2		2		20	30	50
GE/OE (T/P) (2)	24CsCopU1401	Computer Fundamentals	2		2		20	30	50
SEC (P) (2)	24CsDscU1601	Software Engineering	2		2		20	30	50
AEC T(2)	24CpCopU1701 / 24CpCopU1702	MIL-I (Hindi) / MIL-I (Marathi)	2		2		20	30	50
VECT (2)	24CpCopU1801	Environmental Science	2		2		20	30	50
Total			16	06	16	12			550

Level:- 4.5 (First Year) **Sem:II**

Course Type	Course Code	Course Title	Credits		Teaching Scheme Hr/Week		Evaluation Scheme and Max Marks		
			TH	PR	TH	PR	CE	ESE	Total
Subject 1 T(2)+ T/P(2) or T(4)	24CsDscU2101	Advanced Python Programming	2		2		20	30	50
	24CsDscU2102	Lab Course on 24CsDscU2101		2		4	20	30	50
Subject 2 T(2)+ T/P(2) or T(4)	24CsDscU2201	Discrete Probability and Probability Distributions	2		2		20	30	50
	24CsDscU2202	Lab Course on 24CsDscU2201		2		4	20	30	50
Subject 3 T(2)+ T/P(2) or T(4)	24CsDscU2301	Graph Theory	2		2		20	30	50
	24CsDscU2302	Lab Course on 24CsDscU2301		2		4	20	30	50
GE/OE (T/P)(2)	24CsCopU2401	Digital Marketing	2		2		20	30	50
SEC T(2)	24CsStaU2601	Lab course on Excel		2		4	20	30	50
AEC T(2)	24CpCopU2703	English Communication Skills I	2		2		20	30	50
VEC T(2)	24CpCopU2801	Democracy, Election and Governance	2		2		20	30	50
CC(2)	24CpCopU2001/ 24CpCopU2011 / 24CpCopU2021 / 24CpCopU2031 / 24CpCopU2041 / 24CpCopU2051 / 24CpCopU2061 / 24CpCopU2071	Physical Education /Cultural Activities /NSS /NCC /Fine Arts /Applied Arts /Visual Arts Performing Arts	2		2		20	30	50
Total			14	08	14	16			550

Level:- 4.5 (Second Year) Sem:III

Course Type	Course Code	Course Title	Credi ts		Teaching Scheme Hr/Week		Evaluation Scheme and Max Marks		
			TH	PR	TH	PR	CE	EE	Total
Major Core T(2+2 or 4), (T/P)(2)	24CsDscU3101	Database Concepts	2		2		20	30	50
	24CsDscU3102	Web Technologies	2		2		20	30	50
	24CsDscU3103	Lab Course on 24CsDscU3101 and 24CsDscU3102		2		4	20	30	50
VSC P(2)	24CsDscU3501	Foundation of Data Science	2		2		20	30	50
IKS (T/P)(2)	24CsDscU3901	Vedic maths and Computer Fundamentals	2		2		20	30	50
FP P(2)	24CsDscU3002	Field Project I		2		4	20	30	50
Minor (T/P) (2+2 or 4)	24CsDscU3301	Project Management I (Object Oriented Software Engineering)	2		2		20	30	50
	24CsDscU3302	Project Management II (Lab Course on Programming in ‘C’)		2		4	20	30	50
GE/OE (T/P) (2)	24CsDscU3401	Cyber Security	2		2		20	30	50
AEC T(2)	24CpCopU3703	English Communication Skills II	2		2		20	30	50
CC T(2)	24CpCopU3001	Online Course on Yoga	2		2		20	30	50
Total			16	06	16	12			550

Level:- 4.5 (Second Year) Sem:IV

Course Type	Course Code	Course Title	Credits		Teaching Scheme Hr/Week		Evaluation Scheme and Max Marks		
			TH	PR	TH	PR	CE	EE	Total
Major Core T(2+2 or 4), (T/P)(2)	24CsDscU4101	Advanced Database Concepts	2		2		20	30	50
	24CsDscU4102	Data Structure	2		2		20	30	50
	24CsDscU4103	Lab course on 24CsDscU4101 and 24CsDscU4102		2		4	20	30	50
VSC P(2)	24CsDscU4501	Data Mining	2		2		20	30	50
CEP P(2)	24CsCopU4003	Community Engagement Project		2		4	20	30	50
Minor (T/P) (2+2 or 4)	24CsDscU4301	Project Management III (Software Project Management)	2		2		20	30	50
	24CsDscU4302	Project Management IV (Lab on Power BI)		2		4	20	30	50
GE/OE (T/P) (2)	24CsDscU4401	Web Designing	2		2		20	30	50
SEC T(2)	24CsDscU4601	Big Data Analytics	2		2		20	30	50
AEC T(2)	24CpCopU4701 / 24CpCopU4702	MIL-II (Hindi) / MIL-II (Marathi)	2		2		20	30	50
CC T(2)	24CpCopU4001	Health and Wellness	2		2		20	30	50
Total			16	06	16	12			550

Level:- 5.5 (Third Year) Sem:-V

Course Type	Course Code	Course Title	Credits		Teaching Scheme Hr/Week		Evaluation Scheme and Max Marks		
			TH	PR	TH	PR	CE	EE	Total
Major Core T(2+2+2+2or 4 + 2+2 or 4 + 4) P(2+2 or 4)	24CsDscU5101	NoSQL databases	2		2		20	30	50
	24CsDscU5102	Fundamentals of Block Chain Technology	2		2		20	30	50
	24CsDscU5103	Foundation of Artificial Intelligence	2		2		20	30	50
	24CsDscU5104	Lab Course on NOSQL		2		4	20	30	50
	24CsDscU5105	Lab Course on R Programming and Foundation of Artificial Intelligence		4		8	40	60	100
Major Elective (T/P) (2+2 or 4)	24CsBcaU5201	Business Analytics	2		2		20	30	50
	24CsDscU5202	Lab Course on Business Analytics		2		4	20	30	50
	OR								
	24CsDscU5203	Social Media Analytics	2		2		20	30	50
	24CsDscU5204	Lab Course on Social Media Analytics		2		4	20	30	50
VSC P(2)	24CsDscU5501	Lab Course on MATLAB		2		4	20	30	50
FP (2)	24CsDscU5001	Field Project II		2		4	20	30	50
Minor (T/P) (2)	24CsDscU5302	Project Management-V (Software Testing)	2		2		20	30	50
Total			10	12	14	24			550

Level:- 5.5 (Third Year) Sem:-VI

Course Type	Course Code	Course Title	Credits		Teaching Scheme Hr/Week		Evaluation Scheme and Max Marks		
			TH	PR	TH	PR	CE	EE	Total
Major Core T(2+2+2+2 or 4+2+2 or 4+4) P(2+2 or 4)	24CsDscU6101	Data Visualization and Modelling	2		2		20	30	50
	24CsDscU6102	Artificial Intelligence in Data Science	2		2		20	30	50
	24CsDscU6103	Data Security and Privacy	2		2		20	30	50
	24CsDscU6104	Machine Learning	2		2		20	30	50
	24CsDscU6105	Lab on Data Visualization and Modeling		2		4	20	30	50
	24CsDscU6106	Lab on Artificial Intelligence in Data Science and Machine Learning		2		4	20	30	50
Major Elective (T/P) (2+2 or 4)	24CsDscU6201	HR Analytics	2		2		20	30	50
	24CsDscU6202	Lab Course HR Analytics		2		4	20	30	50
	OR								
	24CsDscU6203	Financial Analytics	2		2		20	30	50
	24CsDscU6204	Lab on Financial Analytics		2		4	20	30	50
VSC P(2)	24CsDscU6501	Lab on Cloud Computing		2		4	20	30	50
OJT (2)	24CsDscU6004	On Job Training		4		8	40	60	100
Total			10	12	10	24			550

Semester III

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous),

Shivajinagar, Pune – 5

S.Y. B. Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU3101
Course Name: Database Concepts

Teaching Scheme: TH: 2 Hours/Week **Credit: 02**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Prerequisite Courses:

- Basic Knowledge of file system, storing data in file system and Operations on sets

Course Objectives:

- To explain the basic principles of files.
- A comprehensive introduction to the core concepts of database management systems,
- To study the basic concept of Entity relationship model.
- The process of data normalization and decomposition based on functional dependencies to create well-structured, efficient, and consistent database systems

Course Outcomes:
After successful completion of this course students will able to:

CO No	Course Outcomes (COs)	Blooms Cognitive level
CO 1	List and describe basic concepts of databases, data, information, and DBMS features.	1
CO 2	Demonstrate SQL commands (DDL, DML, constraints, joins, aggregate functions) for database operations.	2
CO 3	Apply SQL queries to create, modify, and retrieve data from relational databases.	3
CO 4	Analyze database schemas and relationships using ER models and normalization.	4
CO 5	Evaluate query results and integrity constraints for correctness and efficiency.	5
CO 6	Design normalized database schemas and implement them with SQL commands.	6

Course Contents:

Unit 1	Introduction to Database Concepts	06 lectures
	1.1 Data 1.2 Information 1.3 Database 1.4 DBMS 1.5 Data Dictionary 1.6 File system Vs. DBMS 1.7 Database Languages	

	1.8 Levels of Abstraction 1.9 Data Independence 1.10 Database Architecture	
Unit 2	**Structure of Relational Databases**	**04 lectures**
	2.1 Tuple, attribute & table 2.2 Types of attributes 2.3 Types of relationships 2.4 Concept of Primary Key and Foreign key	
Unit 3	**Structured Query Language**	**15 lectures**
	3.1 Introduction to SQL 3.2 Data types 3.3 DDL Commands with examples 3.4 DML Commands with examples 3.5 Basic structure of SQL queries 3.6 Integrity Constraints (primary key, referential integrity, not null constraint, unique constraint & check constraint) 3.7 Aggregate functions 3.8 Null Values 3.9 SQL mechanisms for joining relations	
Unit 4	**Conceptual Design (Entity Relationship model)**	**03 lectures**
	4.1 Overview of DB design 4.2 ER data model (Entities, Attributes, Entity sets, Relations, Relationship sets) 4.3 Entity Relationship Diagrams 4.4 Extended ER features 4.5 Case Studies	
Unit 5	**Relational Database Design**	**02 lectures**
	5.1 Concept of Normalization 5.2 Normalization forms (only definitions) 1NF, 2NF, 3NF and BCNF 5.3 Examples on Normalization 5.4 Decomposition 5.5 Desirable properties of Decomposition	

Reference Books:

1. Database System Concepts, Henry F. Korth, Abraham Silberschatz, S. Sudarshan, ISBN-13. 978- 9332901384, Tata McGraw-Hill Education
2. Database Management Systems, Raghu Ramakrishnan and Johannes Gehrke, ISBN-13. 978-0072465631 McGraw-Hill Science/Engineering/Math;
3. Database Systems, Shamkant B. Navathe, RamezElmasri, ISBN-10. 0805317554 Pearson Higher Education

Website:

1. https://www.tutorialspoint.com/postgresql/

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous),
Shivajinagar, Pune – 5

S.Y. B. Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU3102
Course Name: Web Technologies

Teaching Scheme: TH: 2 Hours/Week **Credit: 2**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Prerequisite Courses:

- Familiarity with fundamental computer operations,
- While not a prior skill, having access to a text editor or code editor like VS Code is necessary for practical application during the course.

Course Objectives:

- Understand the web ecosystem:
- Structure web pages with HTML
- Style web pages with CSS
- Add interactivity with JavaScript.
- Build functional web applications

Course Outcomes:
After successful completion of this course students will able to:

CO No	Course Outcomes (COs)	Blooms Cognitive Level
CO 1	Define and identify fundamental HTML elements, basic HTML tags ,CSS properties, and JavaScript syntax.	1
CO 2	Interpret the working of HTML5 forms and validation attributes. Explain CSS concepts such as display, position, and box model.	2
CO 3	Implement and use modern HTML5 semantic elements, CSS Transitions and Animations , JavaScript functions and loops to construct basic web pages.	3
CO 4	Analyze webpage layout, form validation methods, and JavaScript logic including DOM manipulation and event handling.	4
CO 5	Evaluate and optimize web pages by analyzing CSS layouts (box model, media queries, positioning). Evaluate the use of JavaScript Objects.	5
CO 6	Design and implement a complete interactive ,responsive web application integrating HTML5, CSS3, and advanced JavaScript concepts.	6

Course Contents:

Unit 1	**HTML Introduction**	**1 Lecture**
	1.1 HTML Elements and HTML Tags 1.2 HTML Basic Document and Structure 1.3 Interaction of HTML, CSS, and JavaScript to create modern web pages. 1.4 Static vs Dynamic websites 1.5 Introduction to HTML5	
Unit 2	**Basics of HTML5**	**5 Lectures**
	2.1 Structure of an HTML5 document, <! DOCTYPE html> declaration 2.2 Headings, Paragraphs, Lists 2.3 Links and Navigation 2.4 Images and Image Maps 2.5 SVG(Scalable Vector Graphics) 2.6 Tables 2.7 Semantic Elements: <header>, <footer>, <section>, <article>, <aside>, <nav>	
Unit 3	**HTML 5 Forms and Input Controls**	**6 Lectures**
	3.1 Input Types: Text, Email, Number, Range, Date, Time, Color, Radio, Checkbox, File 3.2 Input attributes: Value, Readonly, Disabled, Size, Maxlength, Min, Max, Required, Placeholder, Pattern, Autofocus 3.3 Validation in HTML5 3.4 Datalist, Textarea, Buttons and Select Elements 3.5 Fieldset and Legend 3.6 Progress and Meter	
Unit 4	**Introduction to CSS**	**7 Lectures**
	4.1 What is CSS? 4.2 Adding CSS (Inline, Internal, External) 4.3 CSS selectors 4.4 Colors, Fonts & Text Properties 4.5 Padding, Margin, Background, Border properties 4.6 CSS Lists, CSS links 4.7 Box Model, Display Property, Position property 4.8 CSS Transitions & Animations 4.9 Introduction to Responsive Web Design 4.10 Key Components in Responsive Web Design: Viewport <meta> tag, Flexible Layout (Grid and Flex), Mediaqueries 4.11 Introduction to Responsive Frameworks	
Unit 5	**JavaScript**	**6 Lectures**
	5.1 Introduction of JavaScript 5.2 What can JavaScript(JS) do? JS Statements, JS comments 5.3 Variables & Data Types 5.4 Operators, 5.5 Conditional Statements (If-Else, Switch) 5.6 Loops 5.7 Interaction: alert, prompt, confirm 5.8 Strings	

	5.9 Arrays & Array Methods 5.10 Functions Calling Functions, Function Parameters, Return Values, Arrow Functions	
Unit 6	**Advanced JavaScript**	**3 Lectures**
	6.1 JS Objects 6.2 JS HTML DOM 6.3 JS Events, Event Listeners 6.4 JS Form validation	
Unit 7	**Experiential Learning**	**2 Lectures**
	Project-based application to build interactive, dynamic web experiences with JS	

Reference Books:

1. HTML5 Cookbook, Christopher Schmitt and Kyle Simpson, O'Reilly Media, ISBN-13 : 978-1449396794
2. HTML, CSS, and JavaScript All in One: Covering HTML5, CSS3, and ES6, Sams Teach Yourself 3rd Edition, Julie Meloni, Jennifer Kyrnin , Pearson Publications,ISBN-13 : 978-0135167076
3. Pro HTML5 with CSS, JavaScript, and Multimedia, Collins, Apress publisher ISBN: 9781484252529
4. Mastering Html Css & Java Script Web Publishing By Lemay,BPB Publications, ISBN: 9788183335157

E-Books/MOOC links:

1. HTML, The Complete Reference
   http://www.htmlref.com/
2. https://www.w3schools.com/
3. http://www.tutorialspoint.com/css/
4. https://javascript.info/
5. https://developer.mozilla.org/en-US/docs/Web/HTML
6. http://digimat.in/nptel/courses/video/106105084/L14.html

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous)

Shivajinagar, Pune – 5

S.Y. B.Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU3103
Course Name: Lab Course on 24CsDscU3101 and 24CsDscU3102

Teaching Scheme:PR : 4 Hours/Week **Credit : 02**
Examination Scheme: CIA: 20 Marks **End-Sem : 30 Marks**

Prerequisite Courses:

- Basic knowledge of Database Management System concepts**.**
- Basic knowledge of HTML5, CSS, JavaScript
- Familiarity with text editors such as Notepad++, or Visual Studio Code.

Course Objectives:

- To provide hands-on experience in creating and managing databases using SQL.
- To enable students to design and implement database tables and relationships.
- To develop skills in executing SQL queries for data retrieval and manipulation.
- Understand the structure of an HTML5 document and create basic webpages
- Design and implement HTML5 forms using different input types, attributes, and validation techniques.
- Apply CSS for styling webpages,
- Create responsive web pages
- Use JavaScript to add interactivity to web pages, manipulate webpage elements using the JavaScript DOM and handle user events.
- Implement client-side form validation using HTML5 and JavaScript.
- Integrate HTML, CSS, and JavaScript to build simple interactive web applications.

Course Contents:

Sr.No	List of Assignments
	Lab Course on 24CsDscU3101 (Database Concepts Assignments)
1	Creation of Tables with Primary Key Constraint
2	Creation of Tables with Referential Integrity (Foreign Key)
3	Implementation of Integrity Constraints (Check, Unique, Not Null)
4	Table Manipulation using Insert, Update, Delete, Alter and Drop
5	Computation on Tables using Aggregate Functions and Clauses
6	Queries using Set Operations and Membership Operators
7	Queries using Nested Queries and Subqueries
Lab Course on 24CsDscU3101 (Web Technologies Assignments)	
1	Creating a simple web pages using HTML that includes Heading,Paragraph,Ordered and Unordered Lists, Image, Hyperlink, Text

	Formatting, Tables
2	HTML page that draws using SVG,HTML page that creates an image map
3	Adding CSS to web page e.g. Create a webpage and apply: Inline CSS,Internal CSS, External CSS and Style using headings, paragraphs, and background color
4	Use CSS Selectors and Text e. g Create a webpage and apply: Element selector, Class selector ,ID selector and Change :Font family, Font size, Text color, Text alignment.
5	Box Model: Create three boxes and apply: Border, Padding, Margin, Background color.
6	Create a webpage using HTML and CSS displaying Progress Bar showing course completion, Meter showing student performance.
7	Creating Registration Form using HTML5 Input Types ,Input Attributes and HTML5 Form Validation, Fieldset and Legend
8	CSS Lists and Links like Create a navigation menu using lists and style links with: Hover effect, Color change.
9	Responsive Web Page
10	JavaScript Basics: Conditional Statements,Loops,User Interaction, Functions
11	DOM Manipulation
12	Events, Event Listeners
13	JavaScript Form Validation

Progressive Education Society's
Modern College of Arts, Science and Commerce (Autonomous),
Shivajinagar, Pune – 5

S.Y. B. Sc (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU3501
Course Name: Foundation of Data Science

Teaching Scheme: TH: 2 Hours/ Week **Credit: 02**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Prerequisite Courses:
- Basic knowledge of mathematics, statistics and Python

Course Objectives:
- Introduction of fundamental concepts and techniques used in Data Science
- Develop an understanding of data pre-processing, analysis, and visualization techniques
- Introduce students to machine learning and data modelling
- Equip students with the ability to handle real-world data and derive knowledge and insights

Course Outcomes:
After successful completion of this course students will able to:

CO No	Course Outcomes (COs)	Blooms Cognitive level
CO 1	Understand the fundamental concepts of Data Science	1
CO 2	Analyze and describe data types, data collection methods, and data cleaning techniques	2
CO 3	Apply data preprocessing techniques, identify anomalies, Evaluate the uncover patterns	3
CO 4	Prepare a data model	4
CO 5	Apply various data visualization tools on different datasets	5
CO 6	Create Data Science solutions for Social Network Analysis	6

Course Contents:

Unit 1	Introduction to Data Science	04 lectures
	1.1 Definition and importance of Data Science 1.2 Role of a Data Scientist, Data Analyst, Data Engineer, Data Architect 1.3 Data science process life cycle 1.4 Overview of tools and technologies (Python, R, Power BI, Weka)	
Unit 2	**Data Collection and Data Types**	**04 lectures**

	2.1 Types of Data: Quantitative vs. Qualitative 2.2 Data Collection Methods: Structured, Unstructured and semi-structured Data 2.3 Data Sources: APIs, Web Scraping, Databases, CSV 2.4 Data Formats and Storage	
Unit 3	**Data Preprocessing using Python**	**05 lectures**
	3.1 Data Cleaning Techniques: Identifying and handling missing data. 3.2 Data Transformation: Normalization, Scaling, Encoding 3.2 Outlier Detection Methods 3.3 Outlier Treatment techniques 3.4 Data Integration 3.5 Data Reduction 3.6 Data Wrangling	
Unit 4	**Data Modelling**	**04 lectures**
	4.1 Data Modelling Process 4.2 Data Modelling : Overfitting & Underfitting 4.3 Avoid Overfitting & Underfitting 4.4 Training and Testing 4.5 Data Model Evaluation 4.6 Errors and Biases	
Unit 5	**Data Visualization using Python**	**05 lectures**
	5.1 Definition and Importance of Data Visualization 5.2 Principles of effective visualization 5.3 Types of graphs (E.g. Bar graph, Line graph, Pie Chart Histogram) 5.4 Usage of Libraries: Pandas, Matplotlib, Seaborn, Plotly (for Python) 5.5 Advanced Visualizations: Heatmaps, Scatter Plots, Box Plots	
Unit 6	**Introduction to data Analytics**	**08 lectures**
	6.1 Data driven decision Making 6.2 Types of data analytics 6.3 Data categories 6.4 Data Cycle 6.5 From small data to Big data 6.6 Levels of Data 6.7 Learning from Data 6.8 Analytical Thinking Models 6.9 Descriptive Analytics 6.10 Diagnostic Analysis 6.11 Predictive Analytics 6.12 Prescriptive Analytics	

Reference Books:

1. **Data Science for Business** by Foster Provost and Tom Fawcett **ISBN**: 9781449361327
2. **Python for Data Analysis** by Wes McKinney **ISBN**-13: 978-1491957660
3. **Introduction to Machine Learning with Python** by Andreas C. Müller and Sarah Guido **ISBN**: 9789352134571
4. **Practical Statistics for Data Scientists** by Peter Bruce and Andrew Bruce **ISBN**: 9781491952962
5. **Online Resources:** DataCamp, Coursera, and edX for hands-on tutorials and courses
6. **Fundamentals of Data Analytics(2024)** by Russell Dawson
7. **Data Science: A Beginners Guide(2023)** by C. Raju

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous),

Shivajinagar, Pune – 5

S.Y. B. Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU3301
Course Name: Project Management I (Object Oriented Software Engineering)

Teaching Scheme: TH: 2 Hours/Week **Credit: 02**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Prerequisite Courses:

- Knowledge of Software Engineering practices

Course Objectives:

- To get knowledge and understanding of Object Oriented Software Engineering disciplines, practices and pedagogy.
- To learn Analysis, Design principles for Object Oriented Software Project Development.

Course Outcomes:
After successful completion of this course students will able to:

CO No	Course Outcomes (COs)	Blooms Cognitive level
CO 1	Understand and describe the basic principles of Software Engineering and Object-Oriented Software Engineering.	1
CO 2	Compare Object-Oriented Software Engineering with traditional software engineering approaches and identify key differences.	2
CO 3	Apply structural UML diagrams (e.g., class diagrams, object diagrams) to model system components and relationships.	3
CO 4	Analyze system behavior using sequence, activity, and state diagrams to evaluate and predict system performance and flow.	4
CO 5	Evaluate system architecture using component and deployment UML diagrams to assess the effectiveness of hardware and software deployment.	5
CO 6	Create UML diagrams from real-world case studies to represent complex systems and their behavior accurately.	6

Course Contents:

Unit 1	Introduction to Object-Oriented Software Engineering	07 lectures
	1.1 Overview of Software Engineering 1.2 Object-Oriented Software Engineering vs. Traditional Software Engineering 1.3 Object-Oriented Concepts: Abstraction, Encapsulation, Inheritance, Polymorphism 1.4 Software Development Life Cycle (SDLC) in Object-Oriented Context	

	1.5 Introduction to Unified Modeling Language (UML): Concept and advantages	
Unit 2	**UML Diagrams – Structural Modeling**	**07 lectures**
	2.1 Classes, attributes, operations, relationships (association, aggregation, composition, dependency, generalization) 2.2 Advanced Classes 2.3 Advanced Relationship 2.4 Interface 2.5 Class Diagram - Multiplicity 2.6 Object Diagram – Object instances and their relationships 2.7 Case study of each (Minimum Two)	
Unit 3	**UML Diagrams – Behavioral Modeling**	**10 lectures**
	3.1 Sequence Diagram – Interaction between objects over time (lifelines, messages, activation bars) 3.2 Use Case Diagram – Actors, use cases, relationships (include, extend, generalization) 3.3 Interaction Diagram 3.4 Activity Diagram – Workflow representation, decision nodes, swimlanes 3.5 State Diagram – States, transitions, events, initial & final states 3.6 Case study of each (Minimum Two)	
Unit 4	**UML Diagrams – Architectural Modeling**	**06 lectures**
	4.1 Collaboration Diagram – Object interactions represented as links between objects 4.2 Component Diagram – Software components, dependencies, interfaces 4.3 Deployment Diagram – Hardware and software deployment, nodes, artifacts 4.4 Case study of each (Minimum Two)	

Reference Books:

1. The Unified Modeling Language, User Guide by Grady Booch, James Raumbaugh, Ivar Jacobson, ISBN: 0321267974.
2. Object Oriented Software Engineering by Ivar Jacobson, ISBN: 9780201544350
3. Software Engineering by Roger S. Pressman, ISBN: 9780078022128
4. Applying UML and Patterns by Craig Larman, ISBN-10: 8177589792

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous),

Shivajinagar, Pune – 5

S.Y. B. Sc. (Data Science)
(2024 Course under NEP)

Course Code : 24CsDscU3302

Course Name : Project Management II (Lab Course on Programming in 'C')

Teaching Scheme:PR : 4 Hours/Week **Credit : 02**

Examination Scheme: CIA: 20 Marks **End-Sem : 30 Marks**

Prerequisite Courses:

- Basic knowledge of computer fundamentals.
- Familiarity with algorithms and flowchart concepts.
- Basic logical thinking and problem-solving ability.
- Ability to use a computer system and basic software tools.

Course Contents:

Sr.No	List of Assignments
1	**Basics of C, Data Types and Operators:** • Study the structure of a C program (header files, main function, variable declaration). • Write programs using input/output statements such as printf() and scanf(). • Perform arithmetic operations using arithmetic operators. • Demonstrate relational, logical and assignment operators through simple programs.
2	**Conditional and Looping Statements:** • Write programs using decision making statements such as if, if–else, and nested if. • Implement programs using the switch statement. • Write programs using looping constructs such as for, while, and do–while. • Generate number series, multiplication tables, and factorial using loops.
3	**Arrays:** • Write programs to declare and initialize one-dimensional arrays. • Calculate the sum and average of array elements. • Find the maximum and minimum elements in an array. • Demonstrate traversal and searching of array elements using loops.
4	**Strings:** • Write programs to find the length of a string using strlen(). • Perform string concatenation using strcat().

	• Compare two strings using strcmp(). • Copy one string to another using strcpy().
5	**Functions:** • Write programs using user defined functions. • Implement a function to calculate factorial of a number. • Write a function to check whether a number is prime or not. • Demonstrate function declaration, definition and function calls.
6	**Pointers and Structures:** • Write programs to demonstrate pointer declaration and initialization. • Perform pointer arithmetic operations. • Access variables using pointers. • Create structures to store records such as student details. • Write programs to store and display structure data.

Semester IV

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous),

Shivajinagar, Pune – 5

S.Y. B. Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU4101
Course Name: Advanced Database Concepts

Teaching Scheme: TH: 2 Hours/Week **Credit: 02**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Prerequisite Courses:

- Student should have Basic Knowledge of DBMS is required.
- Basic Knowledge of ER model, relational design ,SQL Queries is required

Course Objectives:

- To learn fundamental concepts of RDBMS (PL/PgSQL).
- To learn database management operations.
- To understand the basic issues of transaction processing and concurrency Control
- To study data security and its importance

Course Outcomes:
After successful completion of this course students will able to:

CO No	Course Outcomes (COs)	Blooms Cognitive level
CO 1	List and describe advanced RDBMS concepts including PL/pgSQL, transactions, concurrency, and recovery.	1
CO 2	Demonstrate the use of PL/pgSQL constructs such as control structures, functions, cursors, and triggers.	2
CO 3	Apply advanced SQL queries and database programming techniques for data manipulation.	3
CO 4	Analyze transaction management issues, concurrency control mechanisms, and deadlock handling.	4
CO 5	Evaluate recovery techniques, database security methods, and efficiency of query execution.	5
CO 6	Design advanced database solutions incorporating distributed, object-based, and NoSQL concepts.	6

Course Contents:

Unit 1	PL/pgSQL Introduction	12 lectures
	1.1 Introduction to PL/pgSQL 1.2 pgAdmin - Introduction 1.3 Datatypes 1.4 Table creation in pgAdmin 1.5 Constraints, Viewing and Editing Data 1.6 Control Structure 1.7 Views 1.8 Stored Procedure and Function 1.9 Exception Handling and Errors 1.10 Cursors	

	1.11 Triggers	
Unit 2	**Transaction Concepts and Concurrency Control**	**12 lectures**
	2.1 Define Transaction, ACID properties, States of Transaction 2.2 Concurrent execution of transactions and conflicting operations 2.3 What is Schedule, Types of schedule 2.4 Concept of Serializability, conflict Serializability 2.5 Lock based protocols 2.6 Deadlock handling methods 2.7 Detection and Recovery (Wait for graph) 2.8 Prevention algorithms (Wound-wait, Wait- die)	
Unit 3	**Database Recovery**	**03 lectures**
	3.1 Transaction Failure classification 3.2 Recovery concepts 3.3 Recovery with concurrent transactions (Rollback, checkpoints, commit) 3.4 Log base recovery techniques (Deferred and Immediate update)	
Unit 4	**Database Security**	**01 lecture**
	4.1 Introduction to database security concepts 4.2 Methods for database security	
Unit 5	**Introduction to Other Databases**	**02 lectures**
	5.1 Parallel and Distributed Database 5.2 Object Based Databases 5.3 XML Databases 5.4 NoSQL Databases 5.5 Multimedia Databases	

Reference Books:

1. Database Management Systems – Database Management Systems, Raghu Ramakrishnan and Johannes Gehrke, McGraw-Hill Education, ISBN-13: 9780072465631, ISBN-10: 0072465638.
2. Fundamentals of Relational Database Management Systems – Fundamentals of Relational Database Management Systems, Springer, ISBN-13: 9783540483977 (Hardcover) / 9783642080128 (Softcover).
3. Database System Concepts – Database System Concepts, McGraw-Hill, ISBN-13: 9780073523323.
4. Practical PostgreSQL – Practical PostgreSQL, O'Reilly Media, ISBN-13: 9781565928466.
5. PostgreSQL – PostgreSQL, Sams Publishing, ISBN-13: 9780672327568.

Website:

2. https://www.tutorialspoint.com/postgresql/
3. https://www.pgadmin.org/docs/

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous),

Shivajinagar, Pune – 5

S.Y. B. Sc. (Data Science)
(2024 Course under NEP)

Course Code : 24CsDscU4102

Course Name : Data Structure

Teaching Scheme:PR : 2 Hours/Week **Credit : 02**

Examination Scheme: CIA: 20 Marks **End-Sem : 30 Marks**

Desirable Prerequisite:

- Students should be able to program in a standard programming language.
- Some mathematical maturity also will be expected; students should have some idea of what constitutes a mathematical proof and how to write one.

Course Objectives:

- To understand the concepts of ADTs
- To Learn linear data structures – lists, stacks, and queues
- To understand sorting, searching and hashing algorithms
- To apply Tree and Graph structures

Course Outcomes:

- **After successful completion of this course students will able to :**

CO No	Course Outcomes (COs)	Blooms Cognitive level
CO 1	List different types of data structures and identify appropriate algorithms by developing problem solving skills by analysing a problem.	1,3
CO 2	Demonstrate the usage of various structures in approaching the problem solution.	2
CO 3	Make use of Use effective and efficient data structures in solving various domain problems.	3
CO 4	Analyze the problems to apply suitable algorithm and data structure.	4
CO 5	Compare all algorithmic strategies for better efficiency	5
CO 6	Design the algorithms to solve the programming problems.	6

Course Contents

Unit No.	Title	Lectures
Unit 1	**Introduction to data structures**	**2**
	1.1. Data type and data object 1.2. Abstract Data Type (ADT) 1.3. Types of data structure 1.4. Algorithm analysis Space and Time complexity	
Unit 2	**Array as a data structure**	**4**
	2.1. Array representation- Row and column major 2.2. Sorting techniques 2.3. Searching techniques with time Complexity	
Unit 3	**Linked List**	**6**
	3.1. Introduction 3.2. Representation -Static &Dynamic 3.3. Types of linked lists- singly, doubly, circular 3.4. Operations-create, display, insert, delete, reverse, search	
Unit 4	**Stack and Queues**	**5**
	4.1. Introduction 4.2. Representation of stack -Static &Dynamic 4.3. Operations– push, pop 4.4. Representation of Queue -Static &Dynamic 4.5. Operations– insert, delete 4.6. Circular queue 4.7. Applications	
Unit 5	**Tree and Graph**	**8**
	6.1. Introduction 6.2. Tree terminologies- all definitions: root, leaf, level, height, depth 6.3. Binary tree 6.4. Types of Binary Tree 6.5. Binary Search Tree (BST) 6.6. Graph terminologies 6.7. Representation – Adjacency matrix, Adjacency list 6.8. Traversal– DFS, BFS	

Reference books:

1. Fundamentals of Data Structures in C by Ellis Horowitz, Sartaj Sahni and Susan Anderson-Freed, 2nd Edition, University Press, 2008, ISBN-13: 978-0929306407, ISBN-10: 0929306406
2. Introduction to Algorithms by Thomas H. Cormen, Charles E. Leiserson, Ronald L.Rivest, Clifford Stein, 3rd Edition, Mcgraw Hill, ISBN-13: 978-0262033848, ISBN-10: 9780262033848
3. Data Structures and Algorithms by Alfred V. Aho, Jeffrey D. Ullman, John E. Hopcroft,1st Edition , Pearson Education,1983. ISBN-13: 978-0201000238, ISBN-10: 0201000237
4. Programming in C by Stephen G. Kochan, 3rd edition, Pearson Education. ISBN-13: 978-0672326660,ISBN-10: 0672326663

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous)

Shivajinagar, Pune – 5

S.Y. B.Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU4103
Course Name: Lab on 24CsDscU4101 and 24CsDscU4102

Teaching Scheme:PR : 4 Hours/Week **Credit : 02**
Examination Scheme: CIA: 20 Marks **End-Sem : 30 Marks**

Prerequisite Courses:

- Students should be able to program in a standard programming language.
- Student should have Basic Knowledge of DBMS is required.

Course Objectives:

- To solve real world computational problems.
- To perform operations on relational database management systems.
- Understand the Fundamental Concepts of Non-Linear Data Structures
- Develop optimized algorithms for data science applications

Course Contents:

Sr.No	List of Assignments
	Lab on 24CsDscU4101 (Advanced Database Concepts Assignments)
1	**Assignment 1: Views** • To Create and Drop Views
2	**Assignment 2 : Stored Procedure and function** • Simple Stored Procedure • Stored Procedure with IN, OUT and IN/OUT parameter • Stored Function that returns
3	**Assignment 3 : Cursors** • Simple Cursor • Parameterize Cursor
4	**Assignment 4 : Exception Handling** • Simple Exception- Raise Debug Level Messages • Simple Exception- Raise Notice Level Messages • Simple Exception- Raise Exception Level Messages
5	**Assignment 5 : Triggers** • Before Triggers (insert, update, delete)

	• After Triggers (insert, update, delete)
Lab on 24CsDscU4102 (Data Structure Assignments)	
1	**Sorting Techniques** • Bubble Sort • Insertion Sort • Selection Sort • Quick Sort • Merge Sort
2	**Searching Techniques** • Linear search • Binary Search
3	**Linked List** • Singly Linked List • Singly Circular Linked List • Doubly Linked List • Doubly Circular Linked List
4	**Stack** • Static Stack Implementation • Dynamic Stack Implementation
5	**Queue** • Static Queue Implementation • Dynamic Queue Implementation

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous)

Shivajinagar, Pune – 5

S.Y. B.Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU4501
Course Name: Data Mining

Teaching Scheme: TH: 2 Hours/Week **Credit: 02**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Prerequisite Courses:

- Understand the concepts of Data science
- Revise the methodologies used for analysis of data
- Describe various techniques which enhance the data modelling.

Course Objectives:

- To understand Data mining techniques
- To understand data pre-processing techniques
- To study algorithms for finding hidden and interesting patterns in data
- To understand and apply various classification and clustering techniques using tools.

Course Outcomes:
After successful completion of this course, students will able to:

CO No	Course Outcomes (COs)	Blooms Cognitive level
CO 1	Define the functionality of the various data mining techniques	1
CO 2	Explain Various Data Preprocessing stages	2
CO 3	Apply appropriate classification, clustering techniques for data analysis	3
CO 4	Discover interesting patterns from large amounts of data to analyze for predictions and classification.	4
CO 5	Compare different approaches of data warehousing and data mining with various technologies.	5
CO 6	Design a Data warehouse system and perform business analysis with OLAP tools.	6

Course Contents:

Unit 1	Introduction to Data Mining	4 lectures
	1.1 Definition : Data mining , KDD 1.2 Stages of the Data Mining Process (KDD) 1.3 Data Mining Techniques/Tasks 1.4 Knowledge Representation Methods 1.5 Applications of Data mining 1.6 Data Mining issues	

Unit 2	**Data Preprocessing**	**5 lectures**
	2.1 Data Cleaning 2.2 Data Integration 2.3 Data Transformation 2.4 Data Reduction 2.5 Data Discretization 2.6 Handling Missing Values & Noisy Data	
Unit 3	**Data Warehousing & OLAP**	**5 lectures**
	3.1 Introduction to Data Warehouse , Data Mart 3.2 Data Warehouse Architecture and its components 3.3 Introduction to OLAP AND OLTP 3.4 Data Modeling with OLAP 3.4.1 Difference between OLTP and OLAP 3.4.2 OLAP Operations (Roll-up, Drill-down, Slice, Dice) 3.4.3 Fact Table, Dimension Table, OLAP cube 3.4.4 Introduction Star Schema, Snowflake Schema, Fact Constellation	
Unit 4	**Classification**	**7 lectures**
	4.1 Definition of Supervised Learning 4.2 Definition of Classification 4.3 General Approaches to solve a classification problem 4.4 Overview of Classification techniques 4.5 Classification Process & Workflow 4.6 Types of Classification Tasks 4.6.1 Introduction to Decision Trees 4.6.2 Decision Construction 4.6.3 Algorithm for Decision tree 4.6.4 Tree Pruning : Pre-Pruning , Post-Pruning Measures for Selecting the Best Split 4.7 Real-World Applications	
Unit 5	**Clustering**	**6 lectures**
	5.1 Definition of Unsupervised Learning 5.2 Cluster Analysis 5.2.1 Introduction to Unsupervised Learning 5.2.2 Requirements for Cluster Analysis. 5.2.3 Problems faced in clustering 5.3 Hierarchical Methods 5.3.1 Agglomerative Hierarchical Clustering 5.3.2 Divisive Hierarchical Clustering 5.4 Clustering Concepts 5.5 Real-World Applications	
Unit 6	**Mining Complex Data Types and Applications**	**3 lectures**

	6.1 Web Mining 6.2 Text Mining 6.3 Spatial & Temporal Data Mining 6.4 Multimedia Mining	

Reference Books:

1. Introduction to Data Mining - Pang-Ning-Tan | Vipin Kumar Michael Steinbach, paperback publication

2. Data Mining Concepts and Techniques - Jiawei Han and Micheline Kamber

3. Data Mining Introductory and Advanced Topics - Margaret H. Dunham S. Sridhar
4. S.C. Gupta -Fundamentals of Statistics, ISBN-13: 978-8183183390
5. D.N. Elhance -Fundamentals of Statistics, ISBN-13: 978-8122500332
6. Data Mining Concepts and Techniques - Jiawei Han | Micheline Kamber | Jian Pei
7. Data Mining - Pieter Adriaans , Dolf Zantinge
8. Data Warehousing in the Real World A Practical Guide for Building Decision Support Systems - Sam Anahory , Dennis Murray
9. The Data Warehouse Toolkit - Ralph Kimball, Margy Ross Amir D Aczel, Jayavel Sounderpandian -Complete Business statistics, ISBN-13: 978-0077108601

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous)

Shivajinagar, Pune – 5

S.Y. B.Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU4301

Course Name: Project Management-III (Software Project Management)

Teaching Scheme:PR : 2 Hours/Week **Credit : 02**

Examination Scheme: CIA: 20 Marks **End-Sem : 30 Marks**

Prerequisite Courses:

- Software Engineering

Course Objectives:

1. To learn process of software project management
2. To understand details of cost estimation
3. To Learn use of project Management tools
4. To understand configuration management
5. To learn user roles and software teams
6. To be familiar with the different methods and techniques used for project management.

Course Outcomes:

1. **Describe the principal tasks of software project managers, and basic concepts in software projects**
2. **Identify and describe the key phases of project management**
3. **Plan software projects, including risk and quality management**
4. **Explain basic concepts and principles of components of software**

CO.No	Course Outcome	Blooms cognitive Level
CO 1	Describe the principal tasks of software project managers, and basic concepts in software projects	1,4
CO 2	Identify and describe the key phases of project management	2,4
CO 3	Plan software projects, including risk and quality management	5,6,3
CO 4	Explain basic concepts and principles of components of software	1,2
CO 5	Be an active team member on an IT project.	3,6
CO 6	Apply project management concepts and	3,2

Course Contents:

Unit No.	Title	Lectures
Unit 1	**Introduction To Software Project Management**	**03 lectures**
	1.1. Project Definition 1.2. Contract Management 1.3. Activities covered by Software Project Management 1.4. Overview of Project Planning 1.5. Stepwise Project Planning.	
Unit 2	**Project Evaluation**	**06 lectures**
	2.1. Work Break Down for Project Estimation & setting 2.2. Milestones 2.3. Different methods of estimation 2.3.1. COCOMO model 2.3.2. Delphi cost estimation 2.3.3. Function point analysis	
Unit 3	**Activity Planning**	**04 lectures**
	3.1. Objectives 3.2. Project Schedule 3.3. Sequencing and Scheduling Activities 3.4. Network Planning Models 3.4.1. Forward Pass 3.4.2. Backward Pass	
Unit 4	**Risk Management**	**04 lectures**
	4.1. Nature Of Risk 4.2. Types of Risk 4.3. Managing Risk	
	4.4. Hazard Identification 4.5. Hazard Analysis 4.6. Risk Planning and Control 4.7. Risk Prioratization	
Unit 5	**Software Quality Management & Control**	**03 lectures**
	5.1. Quality Assurance & Standards 5.2. Six Sigma Model 5.3. The SEI Capability Maturity Model CMM 5.4. Concept of Software Quality 5.4.1. Software Quality Attributes, 5.4.2. Software Quality Metrics and Indicators, 5.5. Quality assurance & Validation plan (SQA Activities, reviews, walkthroughs, inspection, testing) Automation to improve Quality in testing 5.6. Defect Management	

Unit 6	**Configuration Management (CM)**	**03 lectures**
	6.1. Configuration management & Maintenance plan 6.2. Change Management 6.3. Version and Release Management 6.4. Configuration Management Tools 6.5. Library Support	
Unit 7	**Managing People and Organizing Teams**	**04 lectures**
	7.1. Introduction 7.2. Understanding Behaviour 7.3. Organizational Behaviour (Selecting The Right Person) 7.4. The Oldham Hackman Job Characteristics Model 7.5. Working In Groups 7.6. Becoming A Team 7.6.1. Decision Making 7.6.2. Leadership 7.6.3. Organizational Structures 7.7. Stress, Health and Safety 7.8. Case Studies	
Unit 8	**Project Management Tools**	**02 lectures**
	8.1. Project management tool like MS Project 8.2. Assignment based on the tool	
Unit 9	**EXPERIENTIAL LEARNING**	**01 lectures**
	Case study	

Reference books:

1. Software Project Management 5th Edition, McGraw. Hill. ISBN-10 : 0077122798, ISBN-13 : 978-0077122799
2. Effective Software Project Management 1st Edition Robert K. Wysocki ISBN-10 : 0764596365 , ISBN-13 : 978-0764596360
3. Information Technology Project Management by Kathy Schwalbe Thomson Publication. ISBN-101285847091
4. Software project management in practice, Pearson, 1st Edition, Pankaj Jalote ISBN- 13: 9780201737219
5. Software testing and quality assurance, Theory and practice, Wiley, 1st edition, Kshirsagar Naik ISBN-10 : 0471789119 ISBN-13 : 978-0471789116
6. Software project management, A Concise Study, S. A. Kelkar. ISBN-10 : 8120347021 ISBN-13 : 978-8120347021
7. Software Engineering-A Practitioner's Approach 7th or 8th edition, Roger Pressman, McGraw Hill Publication ISBN-10 : 0071184589 ISBN-13 : 978-0071184588
8. Reference website http://www.pmi.org

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous)

Shivajinagar, Pune – 5

S.Y. B.Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDscU4302

Course Name: Project Management-IV (Lab Course on Power BI)

Teaching Scheme:PR : 4 Practicals/Week **Credit : 02**

Examination Scheme: CIA: 20 Marks **End-Sem : 30 Marks**

Prerequisite Courses:

- Microsoft Excel

Course Objectives:

- To introduce students to Business Intelligence tools with a focus on Power BI and Excel for data driven decision-making.
- Enable data import, cleaning, and transformation from Excel and CSV files.
- Understand through simple database tasks.
- Develop skills in building interactive charts and dashboards.

Course Contents:

Sr.No	List of Assignments
1	**Data Import & Setup** • Install and set up Power BI Desktop. • Import an Excel/CSV sales dataset into Power BI Desktop. • Rename columns to meaningful names. • Change incorrect data types (date, number, text). • Remove blank and duplicate records.
2	**Data Cleaning using Power Query** • Remove null values from the Sales and Profit columns. • Split a "Date" column into Year, Month, and Day. • Merge two tables using a common column (Customer ID). • Replace incorrect values using Power Query.
3	**Data Modelling** • Create a relationship between Sales and Customer tables. • Set the correct cardinality for relationships. • Identify and fix incorrect relationships. • Hide unnecessary columns from the report view.
4	**Time-Based Analysis** • Create a Year-wise Sales measure. • Display Month-wise Sales trend. • Compare current year sales with previous year.
5	**Creating Visualizations** • Create a bar chart showing Category-wise Sales. • Create a pie chart showing Region-wise Sales. • Create a table showing Customer-wise Sales and Profit.

	• Add a KPI card for Total Revenue.
6	**Filters & Slicers** • Add a slicer for Year and Region. • Apply visual-level filters to show top 5 products by sales. • Apply page-level filters for a specific year. • Use report-level filters to limit data.
7	**Dashboard Design** • Design a one-page sales dashboard. • Align visuals properly using grid and formatting tools. • Apply consistent colour themes. • Add meaningful titles to visuals.
8	**Drill-Down & Interactions** • Enable drill-down on Category → Sub-Category. • Use drill-through to view detailed customer data. • Enable and test visual interactions. • Create tooltips showing additional insights.
9	**Publishing & Sharing** • Publish the report to Power BI Service. • Create a dashboard from the published report. • Share the dashboard with another user. • Set scheduled data refresh.
10	**Performance & Optimization** • Remove unused columns to reduce model size. • Identify slow visuals. • Optimize measures for better performance. • Explain steps taken to improve report speed.
11	**Mini Practical Project** • Create a complete sales analysis dashboard. • Show Total Sales, Profit, Top Products, and Regions. • Apply slicers for Date, Region, and Category. • Write 4–5 insights derived from the dashboard.

Note:- Students can use any open source dataset for assignments.

Progressive Education Society's

Modern College of Arts, Science and Commerce (Autonomous),

Shivajinagar, Pune – 5

S.Y. B. Sc. (Data Science)
(2024 Course under NEP)

Course Code: 24CsDScU4601
Course Name: Big Data Analytics

Teaching Scheme: TH: 2 Hours/Week **Credit: 2**
Examination Scheme: CIA: 20 Marks **End-Sem: 30 Marks**

Prerequisite Courses:

- Basic Programming (in Python, C, or Java)
- Database Management and SQL

Course Objectives:

- Explain the need and characteristics of Big Data.
- Understand Big Data architectures and ecosystems.
- Describe distributed storage and processing frameworks.
- Apply basic analytical concepts to large datasets.
- Identify real-world use cases of Big Data Analytics.

Course Outcomes:
After successful completion of this course students will able to:

CO No	Course Outcomes (COs)	Blooms Cognitive level
CO 1	Explain the evolution of data, the need for Big Data, and its key characteristics (5 Vs).	1
CO 2	Classify different types of Big Data and identify challenges associated with large-scale data processing.	2
CO 3	Describe Big Data architecture and major components of the Hadoop ecosystem.	3
CO 4	Explain the working of HDFS, including data storage, replication, and fault tolerance mechanisms.	4
CO 5	Illustrate distributed data processing using MapReduce and outline the fundamentals of Apache Spark.	5
CO 6	Identify and analyse real-world applications of Big Data Analytics across various domains.	6

Course Contents:

Unit 1	Introduction to Big Data	6 lectures
	1.1 Evolution of Data and Data Explosion 1.2 Traditional Data vs Big Data 1.3 Characteristics of Big Data (5 Vs)	

	1.4 Types of Big Data: Structured, Semi-structured, Unstructured 1.5 Challenges and Opportunities in Big Data 1.6 Big Data Use Cases	
Unit 2	**Big Data Architecture & Ecosystem**	**6 lectures**
	2.1 Big Data Architecture (Overview) 2.2 Components of Big Data Systems 2.3 Hadoop Architecture 2.4 Hadoop Ecosystem Overview • HDFS • YARN • MapReduce • Hive, Pig (Introduction) 2.5 Role of Big Data in Data Science	
Unit 3	**Distributed Storage – HDFS**	**6 lectures**
	3.1 Need for Distributed File Systems 3.2 HDFS Architecture 3.3 NameNode, DataNode, Secondary NameNode 3.4 HDFS Read and Write Operations 3.5 Data Replication and Fault Tolerance 3.6 Advantages and Limitations of HDFS	
Unit 4	**Big Data Processing Frameworks**	**6 lectures**
	4.1 Introduction to Distributed Processing 4.2 MapReduce Programming Model 4.3 MapReduce Phases 4.4 Limitations of MapReduce 4.5 Introduction to Apache Spark 4.6 Spark Architecture (Driver, Executor, Cluster Manager) 4.7 Spark vs MapReduce (Conceptual Comparison)	
Unit 5	**Big Data Analytics & Applications**	**6 lectures**
	5.1 Basics of Big Data Analytics 5.2 Types of Analytics: Descriptive, Diagnostic, Predictive, Prescriptive 5.3 Data Analytics Lifecycle 5.4 Tools for Big Data Analytics (Overview) 5.5 Applications of Big Data Analytics • Healthcare • Finance • E-commerce • Social Media • Smart Cities	

Reference Books:

1. Book Title: "Hadoop: The Definitive Guide" Author: Tom White Edition: 4th Edition (or the latest available edition) Publisher: O'Reilly Media
2. Hands-On Exploratory Data Analysis with Python" by Suresh Kumar Mukhiya

3. .DT Editorial Services,"Big Data, Black Book-Covers Hadoop2, MapReduce, Hive, YARN, Pig, R and Data Visualization" DreamtechPress,(2015).
4. Applied Big Data Analytics by Kiran Chaudhary and Pradeep Kumar