Savitribai Phule Pune University

(Formerly University of Pune)

Bachelors Degree in Data Science

(Faculty of Science and Technology)

Syllabi for
B.Sc. (Data Science)-Second Year
Sem-III and IV

(For Colleges Affiliated to Savitribai Phule Pune University)

Choice Based Credit System (CBCS) Syllabus
Under National Education Policy (NEP)

To be implemented from Academic Year 2025-26

Savitribai Phule Pune University

Syllabus Structure as per NEP Guidelines

B.Sc. (Data Science) from 2025-26

SY (Level 5.0) SEMESTER III

Course Type	Course code	Course Name	Credits		Teaching Scheme Hrs/Week		Examination Scheme and Marks		
			TH	PR	TH	PR	CE	EE	Total
Major Core	DS-201-MJ-T	Database Management System	2	-	2	-	15	35	50
	DS-202-MJ-T	Data Structure-I	2	-	2	-	15	35	50
	DS-203-MJ-P	Lab Course on DS-201-MJ-T and DS-202-MJ-T	-	2	-	4	15	35	50
VSC	DS-221-VSC-T	Foundations of Data Science	2	-	2	-	15	35	50
FP/ OJT/ CEP	DS-231-FP	Mini Project	-	2	-	4	15	35	50
Minor	DS-241-MN-T	Probability Distribution and Modelling	2	-	2	-	15	35	50
	DS-242-MN-P	Lab Course on DS-241-MN-T	-	2	-	4	15	35	50
IKS	DS-201-IKS	Indian Knowledge System in Computing	2	-	2	-	15	35	50
GE/OE *	OE-201-DS-T* OE-202-DS-T* OE-203-DS-T* OE-204-DS-T*	• E commerce-I / • Web Design-I/ • Digital Marketing-I/ • AI for Everyone-I	2	-	2	-	15	35	50
AEC	AEC-201-T	**From University Basket**	2	-	2	-	15	35	50
CC	CC-201-T	**From University Basket**	2	-	2	-	15	35	50
		Total	**16**	**6**	**16**	**12**			**550**

**** These subjects are offered to other faculty students under GE/OE vertical. The students of B.Sc. (Data Science) will opt the subjects offered by other faculty given in University Basket.***

Savitribai Phule Pune University

Syllabus Structure as per NEP Guidelines

B.Sc. (Data Science) from 2025-26

SY (Level 5.0) SEMESTER IV

Course Type	Course code	Course Name	Credits		Teaching Scheme Hrs/Week		Examination Scheme and Marks		
			TH	PR	TH	PR	CE	EE	Total
Major Core	DS-251-MJ-T	Relational Database Management System	2	-	2	-	15	35	50
	DS-252-MJ-T	Data Structure-II	2	-	2	-	15	35	50
	DS-253-MJ-P	Lab Course on DS-251-MJ-T and DS-252-MJ-T	-	2	-	4	15	35	50
VSC	DS-271-VSC-P	Data Analytics	-	2	-	4	15	35	50
FP/ OJT/ CEP	DS-281-FP	Mini Project	-	2	-	4	15	35	50
Minor	DS-291-MN-T	Testing of Hypothesis and Sampling Distributions	2	-	2	-	15	35	50
	DS-292-MN-P	Lab Course on DS-291-MN-T	-	2	-	4	15	35	50
SEC	SEC-251-DS-T	Software Engineering	2	-	2	-	15	35	50
GE/OE *	OE-251-DS-T* OE-252-DS-T* OE-253-DS-T* OE-254-DS-T*	• E commerce-II / • Web Design-II / • Digital Marketing-II/ • AI for Everyone-II	2	-	2	-	15	35	50
AEC	AEC-251-T	**From University Basket**	2	-	2	-	15	35	50
CC	CC-251-T	**From University Basket**	2	-	2	-	15	35	50
		Total	**14**	**8**	**14**	**16**			**550**

**** These subjects are offered to other faculty students under GE/OE vertical. The students of B.Sc. (Data Science) will opt the subjects offered by other faculty given in University Basket.***

Exit option: Award of UG Diploma in B.Sc. (Data Science) with 88 credits and an additional 04 credits as per University guidelines or else continue with Major and Minor.

Continue option: Third year will be continued with Data Science as Major and Statistics as a Minor subject.

Detail Syllabus
B.Sc. (Data Science)
Semester-III

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-III
DS-201-MJ-T : Database Management System

No. of Credits: 2	Teaching Scheme Theory: 2 Hrs/Week	Examination Scheme Continuous Evaluation: 15 Marks End Semester : 35 Marks

Prerequisites

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

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

On Completion of this course, student will be able to -

CO1: Understand basic concepts of file operations and organization.

CO2: Basic concepts of DBMS, Data Models and Relational Data Model terminologies.

CO3: Construct an Entity-Relationship (E-R) model from specifications.

CO4: Formulate Queries using SQL and Relational Formal Query Languages, working with tables, applying and modifying constraints, join queries.

CO5: Design Normalize database.

Unit No.	Name of Unit	Teaching Hours	CO Targeted
1	**Introduction of DBMS**	**5**	**CO1, CO2**

Introduction, Definition of file, Basic file Operation, Types of file (Logical & Physical)
Basic Concept and Definitions (Data, Information, Database), Definition of DBMS, File System Vs. DBMS, Application of DBMS, Purpose of Database Systems, Levels of abstraction, data independence, Structure of DBMS, DBMS users, Advantages and disadvantages of DBMS

2	**Entity-Relationship Model**	**9**	**CO2, CO3**

Introduction to Data Models (E-R Models, Relational Model, Network Model, Hierarchical Model), Basic Concepts (entities, attributes, entity sets, relations, relationship sets), Constraints (Mapping Constraints, key constraints, referential Constraints, Unique, NULL, Check constraint), Strong and weak entities, Extended ER Features (generalization, specialization, Aggregation), Structure of Relational Databases (concepts of a table),
Examples on E-R Model (Case studies)

3	**Structured Query Language (SQL)**	**10**	**CO4**
Introduction, Basic Structure, DDL Commands, DML Commands, Form a Basic SQL Query Set Operations (union / intersection / except), Aggregate & String Function, NULL Values, Aggregate operators (Group by, having), Nested Queries (Set Membership & Set Comparison operator), SQL mechanisms for joining relations (inner joins, outer joins and their types) case study on SQL			
4.	**Relational Database Design**	**6**	**CO5**
Introduction to Schema Refinement, Problems Caused by Redundancy, Use of Decompositions, Problems Related to Decomposition, Functional Dependencies (Closure of a Set of FDs, Attribute Closure), Normalization forms (1NF, 2NF, 3NF, BCNF), Decomposition (lossless join, dependency preserving property)			
Reference Books			
1. Book1 Database System Concepts by Henry Korth and A. Silberschatz 2. Book2 Introduction to SQL by Reck F. van der Lans by Pearson 3. Book 3 Database Management System by Raghu Ramakrishnan			

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-III

DS-202-MJ-T : Data Structure-I

No. of Credits:	Teaching Scheme	Examination Scheme
2	Theory: 2 Hrs./Week	Continuous Evaluation: 15 Marks End Semester : 35 Marks

Prerequisites

- Basic knowledge of algorithms and problem solving
- Knowledge of Python Programming Language

Objectives

- To Know the Basic Concept of Data Structures and Algorithm
- To Understand different Sorting and Searching Methods with their Time Complexity
- To learn various Application of Array, Linked List,Stack,Queue
- To know the good choice of Data Structure for Data Science Application

Course Outcomes

On Completion of this course, student will be able to -

CO1: Understanding basic concept of Data Structure

CO2: To use well-organized Data Structures in solving various problems.

CO3: Develop Algorithms Using Stacks and Queues for Real-Time Applications

CO4: To efficiently implement the different Data Structures

CO5:To apply linear data structures

Unit No.	Name of Unit	Teaching Hours	CO Targeted
1	**Introduction to Data Structure and Algorithm Analysis**	4 Hrs.	CO1

Overview of Data Structures

Need of Data Structure, Definitions: Data and Information, Data Type, Data Object, ADT, Data Structure, Types of Data Structures: Linear vs. Non-linear, Static vs. Dynamic

Complexity Analysis

Space and Time Complexity, Asymptotic Notation (Big O, Omega ,Theta), Best, Worst, and Average Case Analysis

2	**Array as a Data Structure**	7 Hrs.	CO2

Introduction to Arrays

Representation of Arrays in memory (Contiguous Memory Allocation), Types of Array, Operations: Access, Insertion, Deletion, Traversing

Applications of Array

Searching (Linear Search, Binary Search), Comparison of searching Methods

Sorting Terminology: Internal ,External , Stable, In-Place Sorting, Types of Sorting

Study of Comparison Based Sorting-Efficient Bubble Sort, Insertion Sort, Selection Sort

Divide and Conquer Strategy – Merge Sort, Quick Sort			
3	**Linked List**	**7 Hrs.**	**CO2,CO5**
Introduction to Linked Lists List as a Data Structure, Comparison of Array and Linked list Static and dynamic Implementation of Linked List Types of Linked list -Singly Linked List, Doubly Linked List, Circular Linked List Node structure and pointer-based representation **Operations on Linked Lists** Create, Traversing, Insertion, Deletion, Searching, Reverse, Concatenate, Merge, Time Complexity of Operations **Application of linked List** Online Shopping Cart (E-Commerce) Managing Large Datasets in Big Data Processing LRU Cache Implementation (Optimized Data Retrieval) Memory-Efficient Data Structures for NLP			
4	**Stack**	**6 Hrs.**	**CO3,CO4**
Introduction Stack Operations – init(), Push(), isEmpty(), isFull(), Peek(), Pop(), Time Complexity of operations, Implementation –Static and Dynamic with Comparison **Applications of Stack** Data Preprocessing (Undo/Redo in Data Cleaning), Function Call Stack in Recursive Algorithms (Tree Traversal), Expression Evaluation in Data Analysis (Parsing Expressions) Reversing a Dataset (Time Series Data)			
5	**Queue**	**6 Hrs.**	**CO3,CO4**
Introduction Queue Operation-init(), enqueue(), dequeue(), IsEmpty(), isFull, Peek(), Time Complexity of Operations, Comparison with Stack Implementation –Static and Dynamic with Comparison Types of Queue : Linear Queue, Circular Queue, Priority Queue Dequeue **Applications of Queue** **Data Processing Pipelines (Stream Processing),** Task Scheduling in Machine Learning, Web Scraping, **Chatbots, Cloud Computing**			
Reference Books			
1. Book1: **Data Structures and Algorithms in Python** by Michael T. Goodrich 2. Book2: Python Data Structures and Algorithms by Benjamin Baka 3. Book3: Algorithms in Python" by Robert Sedgwick and Kevin Wayne			

Savitribai Phule Pune University

B.Sc. Data Science (2024 Pattern)

Sem-III

DS-203-MJ-P : Lab Course on DS-201-MJ-T and DS-202-MJ-T (Database Management System and Data Structure-I)

No. of Credits	Teaching Scheme	Examination Scheme
2	Practical 4 Hrs. / Week	Continuous Evaluation: 15 Marks End Semester : 35 Marks

Prerequisites

- Knowledge of Database Management System, storing data in file system
- Basic knowledge of algorithms and problem solving
- Knowledge of Python Programming Language

Course Objectives

- Understand basic database management operations.
- Design E-R Model for given requirements and convert the same into database tables.
- Understand Basic Concept of Searching and Sorting Techniques
- Understand Concept of Linear Data structure(Linked list, Stack, Queue) and their application

Course Outcomes:-

On completion of this course, students will be able to :

CO1: Create database tables in postgreSQL.

CO2: Write and execute simple, nested queries.

CO3: Implement Sorting and Searching Techniques

CO4: Implement Linear Data Structure dynamically with their application

Operating Environment:

For DBMS:

- Operating System: Linux or any relevant Operating System
- DBMS: PostgreSQL

For Data Structures:

Python is cross-platform and can run on:

- **Windows** (Windows 10/11)
- **Linux** (Ubuntu, Fedora, Debian, etc.)
- **macOS** (Latest versions)

To write and execute Python programs efficiently, use the following:

- **VS Code** (Lightweight, supports extensions)
- **PyCharm** (Best for large projects, auto-suggestions)
- **Jupyter Notebook** (Great for data visualization & step-by-step execution)
- **IDLE** (Comes pre-installed with Python, good for beginners)

- **Spyder** (Best for scientific computing)

Database Management System Assignment

Assignment 1

To create simple tables with only the primary key (PK) constraint (as a table level constraint & as a field level constraint) (include all data types)

Assignment 2

To create more than one table, with referential integrity constraint, FK constraint.

Assignment 3

To create one or more tables with following constraints, in addition to the first two constraints (PK & FK)

- Check constraint
- Unique constraint
- Not null constraint

Assignment 4

To drop a table, alter schema of a table, insert / update / delete records using tables created in previous Assignments. (use simple forms of insert / update / delete statements)

Assignment 5

Computation on table using , aggregate functions ,string functions, special clauses (order by, group by, Having)

Assignment 6

To query table, using set operations (union, intersect) set membership operator (in ,not in)

Assignment 7.

To query tables using nested queries (use of 'Except', exists, not exists, all clauses)

Data Structure-I Assignments

Assignment 1: Searching Algorithms

Implementation of searching algorithms to search an element using:

Linear Search, Search, Binary Search (with time complexity)

Assignment 2: Sorting Algorithms - I

Implementation of sorting algorithms:

Bubble Sort, Insertion Sort, Selection Sort

Assignment 3: Sorting Algorithms - II

Implementation of sorting algorithms:

Quick Sort, Merge Sort
Assignment 4: Singly Linked List
1.Dynamic implementation of Singly Linked List to perform following operations: Create, Insert, Delete, Display, Search, Reverse 2. Create a list in the sorted order.
Assignment 5: Doubly Linked List
Dynamic implementation of Doubly circular Linked List to perform following operations: Create, Insert, Delete, Display, Search
Assignment 6: Linked List Applications
1. Merge two sorted lists. 2. Create Simple Music Playlist using Linked List
Assignment 7: Stack
Static and Dynamic implementation of Stack to perform following operations: Init,Push, Pop, Peek, Isempty, Isfull
Assignment 8: Applications of Stack
Data Preprocessing(Undo/Redo in Data Cleaning) Expression Evaluation in Data Analysis Tokenization in Natural Language Processing
Assignment 9: Linear Queue and Circular Queue
1.Static and Dynamic implementation of linear Queue to perform following operations: Init, enqueue, dequeue Peek, IsEmpty, IsFull. 2. Implementation of circular queue
Assignment 10 Application of Queue
1. Web Scraping (Handling Multiple request)
2. Customer Support Chatbots (Message Queues)
Reference Books Book 1 : Database System Concepts by Henry Korth and A. Silberschatz Book 2 : Introduction to SQL by Reck F. van der Lans by Pearson Book 3 : Database Management System by Raghu Ramakrishnan Book 4: Data Structures and Algorithms in Python by Michael T. Goodrich Book 5: Python Data Structures and Algorithms by Benjamin Baka Book 6: Algorithms in Python" by Robert Sedgwick and Kevin Wayne

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-III

DS-221-VSC-T : Foundations of Data Science

No. of Credits:	Teaching Scheme	Examination Scheme
2	Theory: 2 Hrs/Week	Continuous Evaluation:15 Marks End Semester : 35 Marks

Prerequisites

- Knowledge of basic programming concepts
- Knowledge of basic concepts in Mathematics and statistics
- Knowledge of basic concepts in databases

Objectives

- To gain a comprehensive understanding of the core concepts, tools, and Methodologies used in large dataset.
- To know how to Collect, pre-process, and explore datasets to extract meaningful insights.
- To study statistical techniques and models to analyze data and draw conclusions.
- To Understand Data visualization techniques to effectively communicate insights from complex data sets through charts and graphs.

Course Outcomes

On Completion of this course, student will be able to -

CO1: Understand the basic principles and concepts of data science and its relevance to data analytics

CO2: Apply statistical techniques and models to analyze data and draw conclusions.

CO3: Explain how data is collected, managed and stored for data science.

CO4: Build, and prepare data for use with a variety of statistical methods and models.

CO5: Analyze Data using various Visualization techniques.

Unit No.	Name of Unit	Teaching Hours	CO Targeted
1	**Introduction to Data Science**	**6**	**CO1**

Data Science: - Definition, History, Benefits, scope and Applications of Data Science.
The 3 V's: Volume, Velocity, Variety
The Data Science Lifecycle :- Overview – Defining research goals – Retrieving data – Data preparation - Exploratory Data analysis – build the model– presenting findings and building applications
Why learn Data Science?
Real life usage of Data Science Systems
More about Data Science: Data Analysis Vs. Data Analytic, Qualitative Analysis Vs. Quantitative Analysis
Data Mining - Data Warehousing

2	**Statistical Analysis in data science**	**8**	**CO2**

What is Statistical Analysis?
Types of Data in statistics:-Qualitative , Quantitative :-Discrete, continuous
Types of Statistical Analysis

Descriptive Statistical Analysis: Measuring the Frequency
Measuring the Central Tendency: Mean, Median, and Mode
Measuring the Dispersion: Range, Standard deviation, Variance, Interquartile Range
Measure of Relationship:-Covariance, Correlation
Inferential statistics:-Parameter estimation for Statistical Inference
Hypothesis:-Types of hypothesis, Hypothesis Testing
Statistical Tests:-Parametric Tests-Z-test ,t-test ,F-test ,ANOVA ,Chi-square test,
Non-parametric Tests
Regression analysis

3	**Data Collection and Preprocessing**	8	**CO3,CO4**

Types of Data: Structured, semi-structured, Unstructured Data
Data sources:-Internal Data, External data:-Open Data ,Social media data , Standard Dataset Data Formats
Data preprocessing:-Steps in data preprocessing
Data Collection and Import, Data Exploration ,Data Cleaning:- Handling missing data, Removing duplicates, Correcting errors, Dealing with outliers
Data Transformation:- Normalization or standardization, Encoding categorical variable, Feature engineering, Handling skewed data
Data Reduction:- Dimensionality reduction, Numerosity Reduction, Data Compression
Data Validation, Data Discretization,
Imbalanced Dataset, Uses of Data Preprocessing, Advantages and Disadvantages of Data preprocessing.

4	**Visualization tools and software**	8	**CO5**

Introduction to Data visualization
Why Data visualization important?
Data Scientist's Toolbox:- Python ,R-Programming ,SAS, Tableau public ,Microsoft Excel ,Rapid Miner, Tensor flow
Types of visualization**:-** Chart, Graph, Geospatial, Info graphic , Dashboards ,Temporal data, Hierarchical data ,network data ,multi-dimensional data ,Univariate data, multivariate data
Data visualization Software:- Tableau ,Looker studio , Zoho Analytics , Sisense , IBM Watson Analytics , Qlik Sense , Domo , Microsoft Power BI, Klipfolio , Dundas BI , Infogram, Chart Blocks
Data visualization liabraries: - Matplotlib library, seaborn library, ggplot library, Bokeh library, plotly library, pygal library, geoplotlib library, Altair library, Holoviews libraries.
Graphs and Chart for Data Visualization:-Bar Graph , Line Graph,
Histogram, scatter plot, Area chart, Pie charts, Bubble chart, Boxplot, Donut chart, Heat map, Radar Charts, Treemap Charts, Waterfall Charts

Reference Books

1.Data Science Fundamentals and Practical Approaches-Dr.Gypsi Nandi,Dr.Rupam Kumar Sharma

2. Data science concept and practice-Vijay kotu and bala Deshpande

3. Introducing Data Science Big Data, Machine Learning, And More, Using Python Tools-Davy Cielen ,Arno D. B. Meysman, Mohamed Ali

4. Hand-On Introduction to Data Science-Chirag Shah

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-III

DS-231-FP : Mini Project (Using Excel & Advanced Excel)

No. of Credits: 2	Teaching Scheme Practical: 4 Hrs/Week	Examination Scheme Continuous Evaluation: 15 Marks End Semester : 35 Marks

Prerequisites

- Knowledge of Excel & Advanced Excel

Objectives

- Develop hands-on experience in using Excel and Advanced Excel for real-world applications.
- Improve proficiency in data entry, formatting, and worksheet management.
- Learn how to use functions and formulas in excel.
- Learn data analysis and visualization using charts and pivot tables.
- Learn use of macros for repetitive tasks.
- Learn use of Excel for **financial modeling, budgeting, and forecasting**.

Course Outcomes

On Completion of this course, student will be able to -

CO1: Develop Excel & Advanced Excel Skills.

CO2: Format and manage worksheet.

CO3: Use functions and formulas in excel.

CO4: Use macros for repetitive tasks.

CO5: Analyze & visualize data using charts and pivot tables

CO6: Understand use of Excel for **financial modeling, budgeting, and forecasting**.

Building an **Excel project** involves planning, structuring, and implementing various Excel features to meet specific requirements. The expectations for an Excel project depend on its complexity and purpose.

Students can find data for analysis from variety of sources including online platforms like Data.gov, Kaggle and Google Dataset Search as well as collect data from Government agencies, research institutions, concern organizations and industries.

Sr. No.	Guidelines for Mini Project using Excel & Advanced Excel
1	**Define the Purpose** • What problem will the project solve? • Is it for data analysis, automation, financial modeling, or reporting? • Who will use the file (yourself, a team, or a client)?
2	**Data Structure & Formatting** • Organize data using proper headers and columns. • Use tables to manage structured data.

	• Apply consistent formatting for readability. • Use data validation to prevent incorrect entries.
3	**Formulas & Functions** • Implement basic formulas (SUM, AVERAGE, IF) Use advanced functions (FILTER, COUNTIF, SUMIF, TEXT, DATE functions).
4	**Automation Using Macros** • Automate repetitive tasks with macros.
5	**Data Visualization & Dashboards** • Use charts (bar, pie, line, pivot charts) for visualization. • Create interactive dashboards using slicers and pivot tables. • Apply conditional formatting to highlight key insights.
6	**Security & Protection** • Protect sensitive sheets and cells. • Use passwords to restrict access if needed. • Prevent accidental modifications using data validation and locked cells.
7	**Documentation & User Guide** • Provide clear instructions for users. • Add comments or notes to explain complex formulas. • Include a help sheet if required.

Examples of Mini Project using Excel & Advanced Excel

1. Personal Budget Tracker
2. Customer Feedback Analysis
3. Meal Planner & Shopping List
4. Grade Book for Teachers
5. Event Planning Organizer
6. Inventory Management System
7. Sales Report and Forecasting Tool
8. Task List with Priority Levels

And many other

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-III

DS-241-MN-T: Probability Distribution and Modelling

No. of Credits: 02	Teaching Scheme **Theory:** 2 Hours/Week	Examination Scheme Continuous Evaluation:15 Marks End Semester : 35 Marks

Prerequisites

- Basics of Probability, discrete and continuous random variable

Objectives

- To study probability distribution of (univariate and bivariate) continuous random variables, expectation and moments of probability distribution
- To understand the concept of discrete and continuous probability distributions
- To calculate the probabilities of discrete and continuous probability distributions

Course Outcomes

On Completion of this course, student will be able to –

CO1: Identify the given function is probability density function (pdf) or not

CO2: Calculate cumulative distribution function (cdf), mean, variance, Moment generating function (MGF), Cumulant generating function (CGF), skewness and kurtosis for univariate continuous random variable

CO3: Calculate cdf, mean and variance of (XY), joint moments, MGF, CGF, skewness and kurtosis for bivariate continuous random variable

CO4: Obtain marginal and conditional expectation and distribution

CO5: Identify the distribution

CO6: Calculate the probabilities of discrete and continuous probability distributions

CO7: Fit distribution to the particular data

CO8: Interpret the results after fitting

Unit No.	Name of Unit	Teaching Hours	CO Targeted
1	**Continuous Univariate Distributions**	**04**	**CO1, CO2**

Continuous sample space: Definition, illustrations.

Continuous random variable: Definition, pdf, cdf, properties of cdf (without proof), probabilities of events related to random variable.

Expectation of continuous r.v., expectation of function of r.v. $E[g(X)]$, mean, variance, geometric mean, harmonic mean, raw and central moments, skewness, kurtosis, mean deviation about mean.

MGF: Definition, properties (without proof). CGF, Mode, partition values

Probability distribution of function of r. v. : $Y = g(X)$ using i) Jacobian of transformation for $g(.)$ monotonic function and one-to-one, on to functions, ii) Distribution function for $Y = X^2$, $Y = |X|$ etc., iii) M.G.F. of $g(X)$.

2	**Continuous Bivariate Distributions**	**04**	**CO3, CO4**

Continuous bivariate random vector or variable (X, Y): Joint p. d. f., joint c. d. f,
properties (without proof), probabilities of events related to random variables. Marginal and conditional distributions. Expectation of r.v. (X,Y), expectation of function of r.v. $E[g(X,Y)]$, joint moments,
$Cov\ (X,Y)$, $Corr\ (X,Y)$, conditional mean, conditional variance, $E[E(X|Y = y)] = E(X)$ & $E[E(Y/X = x) = E(Y)$. Theorems on expectation (without proof):
i) $E(X + Y) = E(X) + E(Y)$,
ii) $E(XY) = E(X)\,E(Y)$, $if\ X\ and\ Y$ are independent, generalization to k variables.
iii) $E(aX + bY + c), Var\ (aX + bY + c)$.
Independence of random variables $X\ and\ Y$ and also its extension to k random variables.
MGF for bivariate distribution and its properties without proof.
Probability distribution of transformation of bivariate r. v. $U = \phi_1\ (X,Y)$, $V = \phi_2(X,Y)$.

3	**Standard Discrete Distributions**	**10**	**CO5, CO6, CO7, CO8**

Uniform Distribution: Pmf of uniform distribution, mean, variance, moments, skewness, kurtosis, standard results and theorems, examples on calculation of uniform probabilities, drawing a random sample from uniform distribution
Bernoulli Distribution: Pmf of Bernoulli distribution, mean, variance, moments, skewness, kurtosis, standard results and theorems, examples on calculation of Bernoulli probabilities
Binomial Distribution: Pmf of binomial distribution, mean, variance, moments, skewness, kurtosis, standard results and theorems, examples on calculation of binomial probabilities, drawing a random sample from binomial distribution, fitting of binomial distribution
Poisson Distribution: Pmf of poisson distribution, mean, variance, moments, skewness, kurtosis, standard results and theorems, examples on calculation of poisson probabilities, drawing a random sample from poisson distribution, fitting of poisson distribution
Geometric Distribution: Pmf of geometric distribution, mean, variance, moments, skewness, kurtosis, standard results and theorems, examples on calculation of geometric probabilities

4	**Standard Univariate Continuous Distribution**	**12**	**CO5, CO6, CO7, CO8**

Uniform or Rectangular Distribution: Pdf of uniform distribution, mean, variance, moments, skewness, kurtosis, standard results and theorems, examples on calculation of uniform probabilities
Normal Distribution: Pdf of normal distribution, mean, variance, moments, skewness, kurtosis, standard results and theorems, examples on calculation of normal probabilities, fitting of normal distribution
Exponential Distribution: Pdf of exponential distribution, mean, variance, moments, skewness, kurtosis, standard results and theorems, examples on calculation of normal probabilities, drawing a random sample from exponential distribution
Gamma Distribution: Pdf of gamma distribution, mean, variance, moments, skewness, kurtosis, standard results and theorems, examples on transformation of random variable

Reference Books

1. Goon, A. M., Gupta, M. K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1, Sixth Revised Edition, The World Press Pvt. Ltd., Calcutta.
2. Gupta, S. C. and Kapoor, V. K. (2002), Fundamentals of Mathematical Statistics,

(Eleventh Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi, 110002.
3. Gupta, S. P. (2002), Statistical Methods (Thirty First Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi 110002.
4. Hogg, R. V. and Craig, A. T. , Mckean J. W. (2012), Introduction to Mathematical Statistics (Tenth Impression), Pearson Prentice Hall.
5. Neil, A. Weiss, (2016). Introductory Statistics, Tenth Edition, Pearson.
6. Purohit, S. G., Gore S. D., Deshmukh S. R. (2008). Statistics Using R, Narosa Publishing House, New Delhi.
7. Sarma, K. V. S. (2001). Statistics Made it Simple: Do it yourself on PC. Prentce Hall of India, New Delhi.
8. W. and Cochran W. G. (1989). Statistical Methods, Eighth Ed. East West Press.
9. Mood, A. M. and Graybill, F. A. and Boes D.C. (1974). Introduction to the Theory of Statistics, Ed. 3, McGraw Hill Book Company.

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-III

DS-242-MN-P : Lab Course on DS-241-MN-T

(Probability Distribution and Modelling)

No. of Credits: 2	**Teaching Scheme** Practical: 4 Hrs/Week	**Examination Scheme** Continuous Evaluation: 15 Marks End Semester : 35 Marks

Prerequisites

- Basics of Probability, discrete and continuous random variable

Objectives

Practical Implementation of

- Probability distribution of (univariate and bivariate) continuous random variables, expectation and moments of probability distribution
- Discrete and continuous probability distributions

Course Outcomes

On Completion of this course, student will be able to -

CO1: Fit suitable discrete and continuous probability distributions to the data.

CO2: Identify the suitable probability model for the population.

CO3: Generate random samples from the discrete and continuous probability distributions.

List of Assignments

Sr. No.	**Assignment**
1	Fitting of Binomial distribution and computation of expected frequencies.
2	Fitting of Poisson distribution and computation of expected frequencies.
3	Applications of Binomial distribution
4	Application of Poisson and geometric distribution.
5	Model sampling from Poisson and Binomial distributions.
6	Fitting of normal distribution and computation of expected frequencies.
7	Model sampling from exponential distribution using distribution function
8	Generating random samples from normal distribution using (i) distribution function (ii) Box-Muller transformation.
9	Applications of normal and exponential distribution.
10	Problems on Bivariate Probability distributions.

Reference Books

10. Goon, A. M., Gupta, M. K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1, Sixth Revised Edition, The World Press Pvt. Ltd., Calcutta.

11. Gupta, S. C. and Kapoor, V. K. (2002), Fundamentals of Mathematical Statistics, (Eleventh Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi, 110002.
12. Gupta, S. P. (2002), Statistical Methods (Thirty First Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi 110002.
13. Hogg, R. V. and Craig, A. T. , Mckean J. W. (2012), Introduction to Mathematical Statistics (Tenth Impression), Pearson Prentice Hall.
14. Neil, A. Weiss, (2016). Introductory Statistics, Tenth Edition, Pearson.
15. Purohit, S. G., Gore S. D., Deshmukh S. R. (2008). Statistics Using R, Narosa Publishing House, New Delhi.
16. Sarma, K. V. S. (2001). Statistics Made it Simple: Do it yourself on PC. Prentce Hall of India, New Delhi.
17. W. and Cochran W. G. (1989). Statistical Methods, Eighth Ed. East West Press.
18. Mood, A. M. and Graybill, F. A. and Boes D.C. (1974). Introduction to the Theory of Statistics, Ed. 3, McGraw Hill Book Company.

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-III

DS-201-IKS: Indian Knowledge System in Computing

No. of Credits: 02	Teaching Scheme **Theory:** 2 Hours/Week	Examination Scheme Continuous Evaluation:15 Marks End Semester : 35 Marks

Objectives

- To introduce Vedic mathematical techniques and their relevance to modern computational methods.
- To understand Nyaya's logical framework and its application in reasoning and AI.
- To explore the algorithmic structure of Panini's grammar and Chandasastra's binary system in computational linguistics and mathematics.
- To explore real-world applications of IKS concepts in computational sciences.

Course Outcomes

On Completion of this course, student will be able to –

CO1: Understand the computational foundations of Indian Knowledge Systems by applying Vedic mathematical techniques in problem-solving.

CO2: Use Nyaya's logical reasoning in AI and decision-making.

CO3: Explore the connection between Panini's grammar and NLP technologies.

CO4: Recognize the applications of IKS in modern computing fields.

Unit No.	Name of Unit	Teaching Hours	CO Targeted
1	**Vedic Mathematics & Computational Thinking**	**8**	**CO1**
1.1 Introduction to Vedic Mathematics: Origins and importance in ancient India, Sutras and their logical foundation **1.2** Basic Arithmetic using Vedic Methods: Addition, subtraction, multiplication, and division tricks **1.3** Algebraic Applications of Vedic Mathematics: Squaring, square roots, cube roots, and factorization			
2	**Introduction to Nyaya (Indian Logic)**	**8**	**CO2**
2.1 Introduction to Nyaya Philosophy: Introduction to Nyaya (Indian Logic), Overview of Indian philosophical schools, Importance of Nyaya in logical reasoning, Types of reasoning (Anumana, Pramana, etc.) **2.2** Nyaya's Four Sources of Knowledge (Pramaṇa): Perception, inference, comparison, verbal testimony			

2.3 Types of Argumentations in Nyaya
Vada (truth-based), Jalpa (debate-focused), Vitanda (criticism)

2.4 Applications in AI & Machine Learning**:** Logical reasoning models, expert systems, and rule-based AI

3	**Panini's Astadhyayi & Chandasāstra**	**8**	**CO3**

3.1 Introduction to Panini's Astadhyayi: Historical background and linguistic importance

3.2 Rule-Based System of Sanskrit Grammar: Sutras, meta-rules, recursion, and transformations

3.3 Chandasastra's Binary logic and combinatorial techniques

4	**Applications of IKS in Computer Science**	**6**	**CO4**

4.1 Mind and cognition in Samkhya and Yoga: AI insights

4.2 Machine Learning and Indian philosophies: Understanding of human cognition in Indian philosophical schools (Advaita, Samkhya and Yoga)

4.3 Cryptography and Security: Ancient cryptographic methods in Kautilya's Arthashastra, protecting information: analogies from Indian traditions

Reference Books

1. Vedic Mathematics, Jagadguru Swami Bharati Krishna Tirtha, Motilal Banarsidass Publishing House, New Delhi.
2. "The Power of Vedic Maths" – Atul Gupta, JAICO publishing
3. Nyaya Theory of Knowledge" – S.C. Vidyabhusana
4. "A Primer of Indian Logic" – Kuppuswami Sastri, Hassell Street Press.2021
5. "Indian Logic: A Reader" – Jonardon Ganeri
6. "Aṣṭādhyāyī of Pāṇini" (Volumes 1 & 2) – Rama Nath Sharma, Munshirm Manoharlal publication
7. "Panini: His Work and Its Traditions" – George Cardona, Motilal Banarsidass Publishing House
8. "The Mathematics of Metre" – Satyanarayana Das
9. "Samkhya and Science" – Debabrata Sen Sharma
10. Explores the cognitive science aspects of Samkhya and Yoga in AI research.
11. "AI and Indian Philosophy" – Sangeet Kedia
12. "Kautilya's Arthashastra" – R. Shamasastry (Translation)
13. "History of Indian Cryptography" – Subhash Kak
14. Discusses coded messages, steganography, and security concepts in ancient India.
15. Saubhagya Vardhan, AI in Land of Vedas, Notion Press, 2023

Savitribai Phule Pune University

B.Sc. Data Science (2024 Pattern)

Sem-III

Course Type: GE/OE Course Code: OE-201-DS-T

Course Title: Ecommerce I

(To be offered to faculty other than Science and Technology)

Teaching Scheme 2 Hours /Week	No. of Credits: 2	**Examination Scheme** Continuous Evaluation:15 Marks End Semester: 35 Marks

Prerequisites

- Basic Computer and Internet Knowledge
- Fundamentals of Business Commerce and Digital Marketing Basics
- Financial and Payment Systems

Course Objectives

- To understand basic concepts about e-Commerce.
- To understand the applications of e-Commerce.
- To learn business model knowledge.
- To enable knowledge about E-payment system.
- To get a general idea of M-commerce

Course Outcomes

On completion of the course, student will be able to–

CO1: Implementation of basic concepts and application of e-Commerce

CO2: Use of electronic payment system.

CO3: Implementation of Business model knowledge

CO4:M-commerce technology use

Course Contents

Chapter 1	**E- Commerce and Business Model Concepts**	**7 Hours**

1.1. Main Activities of E Commerce
1.2. Definition
1.3. Goals
1.4. Technical Components
1.5. Functions
1.6. Status
1.7. Prospects
1.8. Significance
1.9. Advantages
1.10. Disadvantages E-Commerce Business Models
1.11. Major Business to Consumer (B2C)Business Model Portal, E-tailor
1.12. Major Business to Business (B2B) Business Model
1.13. E Distributor, E-Procurement, Exchanges
1.14. Business models in Emerging E-Commerce Areas - C2C, P2P, and B2G.

Chapter 2	**E-Marketing and E- Commerce Application**	**7 Hours**

2.1. Identifying Goals
2.2. Browsing Behaviour Model

2.3. Online Marketing
2.4. e-Commerce and retailing
2.5. e-Commerce and banking,
2.4. E Advertising
2.5. Internet Marketing Trends
2.6. Target Markets
2.7. E-Branding
2.8. Marketing Strategies
2.9. Consumer Online: The Internet Audience and Consumer Behaviors
2.10. E-cycle of Internet Marketing

Chapter 3	**E-commerce Payment Method**	**6 Hours**

3.1. The requirements of an electronic payment system,
3.2. The constraints of a traditional payment system, and an overview of electronic payment technology.
3.3. Electronic payment gateways
3.4. B2B electronic payments.
3.5. Third-party payment processing electronic or digital currency, its characteristics, and its operation.
3.6. System for online credit card payments and smart cards.

Chapter 4	**Title E-Commerce Laws & Regulations**	**5 Hours**

4.1.Introduction to E-Commerce Laws
4.2.Information Technology (IT) Act, 2000 (India)
4.3.GDPR (General Data Protection Regulation) – EU
4.4.Other Global E-Commerce Laws.

Chapter 5	**Title Future Trends in E-Commerce**	**5 Hours**

5.1. AI & Chatbots in E-Commerce
5.2. AR/VR Shopping Experiences
5.3. Personalization & Data Analytics
5.4. The Role of IoT in E-Commerce

Reference Books:

1. Kenneth C. Laudon, E-Commerce : Business, Technology, Society, 4th Edition, Pearson
2. S. J. Joseph, E-Commerce: an Indian perspective, PHI
3 E-Commerce Law: National and Transnational Topics– Alan Davidson
4. Artificial Intelligence in E-Commerce– Richard Boire

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-III

Course Type: GE/OE Course Code : OE-202-DS-T

Course Title: Web Design I

(To be offered to faculty other than Science and Technology)

Teaching Scheme	No. of Credits	**Examination Scheme**
02 Hrs/ week	2	Continuous Evaluation:15 Marks End Semester: 35 Marks

Prerequisites

- Basic computer knowledge and the ability to work with files.
- Knowledge of other programming languages will enhance your abilities with HTML and web development.
- Knowledge and understanding of Internet.

Course Objectives

- To learn HTML tags and programming concepts and techniques.
- To develop the ability to logically plan and develop web pages.
- To learn to write, test, and debug web pages using HTML.
- To learn to design Style sheets.
- To know to link and publish Web pages.

Course Outcomes

On completion of the course, student will be able to–

- Learn and use the HTML Tags.
- Understand and resolves written HTML codes.
- Design and develop the page using HTML codes.

Course Contents

Chapter 1	**Introduction to Web Design**	**8**

1.1 Introduction
1.2 Working of the Internet
1.3 Role of Web Servers, Clients(Communication)
1.4 Web Browsers
1.5 Working of the Internet , Intranet and WWW
1.6 E-Mail Servers and Protocols
1.7 E-mail Clients and Web Based Mail Access using Browser
1.8 Messenger Services and Clients(Chat)
1.9 Advantages and Disadvantages of Internet
1.10 Concept of effectiveWeb Design (Web site, classification of website, Advantages and Disadvantages. Of website)
1.11 Fundamental Principles of Web page design and issues

Chapter 2	Getting Started with HTML	6

2.1 Introduction to scripting Languages
2.2 HTML Editing Tools
2.3 WYSISYG Authoring Tools
2.4 My First HTML Script
2.5 Basic HTML Document Structure
2.6 Common HTML Tags and it's attributes
2.7 Design HTML Tags
2.8 Text Formatting and Styles
2.9 Images and Graphics
2.10 Button, Formatting and Style
2.11 Lists
2.12 Hyperlinks
2.13 Multimedia
2.14 Frames
2.15 HTML Forms
2.16 Linking Web pages
2.17 Publishing Web Pages

Chapter 3	Tables	6

3.1 The summary of Table tags
3.2 Introduction to tables
3.3 The Basic table structure
3.4 Affecting table appearance
3.5 Table troubleshooting
3.6 Tips and tricks
3.7 Standard table templates
3.8 Multipart images in tables

Chapter 4	Frame / Forms	6

4.1.Introduction to frames
4.2.The basic frameset structure
4.3.The frame function and appearance and Targeting frames
4.4.The Inline (Floating) frames and Frame design tips and tricks
4.5.Introduction to Forms and The basic form (FORM)
4.6. The FORM elements and FORM attributes
4.7.Unconventional use of FORM elements
4.8.Demystifying CGI
4.9.Retrieving the parameter value using getParameter () method

Case Studies	4

Case study 1: Creation of forms, small case study to create HTML pages using all the above learnt techniques.

Case study 2: Creation of Forms layout designing by using div element with CSS property

Case study 3: Create Multiple Web pages link them to publish a small website.

Reference Books:

1. Computer Programming For Beginners:Learn The Basics Of HTML5-Joseph Connor
2. The Complete Reference HTML & CSS-Fifth Edition-Thomas A.Powell
3. Learning Web Design: A beginner's Guide To HTML, CSS, Javascript, and Web Graphics - Jennifer Robbins
4. HTML5: The Missing Manual - Matthew MacDonald.

Reference Link :

1. HTML reference - HTML: HyperText Markup Language | MDN
2. HTML Standard

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-III

Course Type: GE/OE. Course Code: OE-203-DS-T
Course Title: Digital Marketing I
(To be offered to faculty other than Science and Technology)

Teaching Scheme 02 Hours /Week	No. of Credits 2	**Examination Scheme** Continuous Evaluation:15 Marks End Semester: 35 Marks

Prerequisites

- Creative & Logical thinking ability,
- Digital devices operational skills and Knowledge

Course Objectives

- To understand Digital Marketing as the most powerful marketing tool.
- Learn to create digital marketing artworks.
- Learn how to use email campaigns, blogging to produce worthwhile, pertinent material that draws in and engages a target audience.

Course Outcomes

On completion of the course, student will be able to–

CO1: Students will have improved their visualization power to explore new ideas.

CO2: Students will learn to develop their Marketing skills.

CO3: Student will making sure that tactics are in line with corporate objectives and produce quantifiable outcomes from e-payment mechanisms.

CO4: Student will be in line with target demographics and company objectives, increasing brand awareness and boosting conversions.

Course Contents

Chapter1	**History of Digital Marketing**	**8**

1.1.What is Digital Marketing ?

1.2.Offline digital marketing - Electronic billboards, Radio marketing, T.V. marketing, Phone marketing

1.3.Online marketing - Search Engine Optimization (SEO), Social media marketing, E-mail marketing

1.4.Difference between conventional marketing and online marketing.

Chapter2	**Internet Marketing**	**8**

2.1. Structure of Website - Team

2.2. Types of website - Static Website, Dynamic website, Personal, Commercial, Governmental, Non- profit organisation

2.3. Web Portals - Type of Portals

Chapter3	**Classification of e-Commerce**	**8**

3.1. Business to Business (B2B) Model

3.2. Business to Consumer(B2C) Model

3.3. Consumer to Consumer(C2C) Model 3.4. Consumer to Business (B2B) Model	
Case Study	**6**
Case Study 1 Case Study 2 Case Study 3	
Reference Books:	
1. SEO 2025: Learn Search Engine Optimization with Smart Internet Marketing Strategies" by Adam Clarke 2. Digital Marketing : Nitin Kamat, ChinmayKamat (Himalaya Publishing House) 3. "Made to Stick: Why Some Ideas Survive and Others Die" by Chip Heath and Dan Heath 4. "Digital Marketing: Strategy, Implementation, and Practice" by Dave Chaffey and Fiona Ellis-Chadwick	
Reference Links: 1. https://www.coursera.org/browse/business/digital-marketing 2. https://learndigital.withgoogle.com/digitalgarage	

Savitribai Phule Pune University

B.Sc. Data Science (2024 Pattern)

Sem-III

Course Type: GE/OE. Course Code: OE-204-DS-T

Course Title: AI for Everyone-I

(To be offered to faculty other than Science and Technology)

Teaching Scheme: 2 hours / week	No. of Credits: **2**	**Examination Scheme** Continuous Evaluation:15 Marks End Semester: 35 Marks

Course Objectives: -

1. Understand the basics of artificial intelligence and its subfields.
2. Explore real-world applications of AI across different industries.
3. Gain insights into the ethical, social, and economic implications of AI.
4. Develop an appreciation for the potential of AI to drive innovation and transformation.

Course Outcomes: -On completion of the course, student will be able to–

CO1: Define and explain the fundamental concepts and subfields of AI.
CO2: To understand the potential of AI to drive innovation and transformation in different domains.
CO3: Identify real-world applications of AI across various industries.
CO4: Analyse the ethical, social, and economic implications of AI.

Course Contents

Unit 1	**Introduction to Artificial Intelligence**	**8 hours**	**CO1**

1.1 Definition and scope of AI
1.2 Historical overview and key milestones
1.3 Differentiating AI from human intelligence
1.4 Types of AI tools: Text, image, audio, video, coding, and automation.
1.5 Where to find free AI tools? (Google AI, OpenAI, Hugging Face, etc.)

Unit 2	**AI Subfields**	**6 hours**	**CO2**

2.1 Machine learning: Supervised, unsupervised, and reinforcement learning
2.2 Deep learning and neural networks
2.3 Natural language processing (NLP) and computer vision

Unit 3	**Applications of AI**	**8 hours**	**CO3**

3.1 AI in healthcare: Diagnosis, treatment, and medical imaging
3.2 AI in finance: Fraud detection, algorithmic trading, and risk assessment
3.3 AI in transportation: Autonomous vehicles and traffic optimization
3.4 AI in customer service and chatbots
3.5 AI in education: Personalized learning and intelligent tutoring systems

Unit4	Ethical and Social Implications of AI	8 hours	CO4
4.1 Bias and fairness in AI systems. 4.2 Privacy and data protection concerns 4.3 Impact of AI on employment and the workforce 4.4 AI and social inequality			
Reference Books:			
1. Artificial Intelligence: A Guide for Thinking Humans" – Melanie Mitchell 2. The AI Revolution in Medicine: GPT-4 and Beyond" – Peter Lee, Carey Goldberg, Isaac Kohane 3. AI 2041: Ten Visions for Our Future" – Kai-Fu Lee, Chen Qiufan 4. The Business of AI: AI Technologies and How to Leverage Them for Business Success" – Anirudh Koul 5. AI-Powered Marketing: Harness the Future of Marketing with AI" – Peter Gentsch 6. The AI Marketing Handbook" – Ryan McKenzie			

Detail Syllabus
B.Sc. (Data Science)
Semester-IV

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

DS-251-MJ-T: Relational Database Management System

No. of Credits: 2	Teaching Scheme Theory: 2 Hrs/Week	Examination Scheme Continuous Evaluation: 15 Marks End Semester : 35 Marks

Prerequisites

- Basic Knowledge of DBMS, Knowledge of SQL Queries, Basics of relational design

Objectives

- To teach the essential concepts of RDBMS with PL/PgSQL
- Understand the need of transaction processing and learn techniques for controlling the consequences of concurrent data access.
- Learn how different recovery techniques restore the database to a consistent state after failures

Course Outcomes

On Completion of this course, student will be able to -

CO1: Understand the basic syntax and structure of PL/SQL language.

CO2: Analyze database transactions and can control them by applying ACID properties.

CO3: Demonstrate the basic concepts of transaction processing and concurrency control.

CO4: Understand concurrent transactions affect the recovery process.

CO5: Analyze various databases.

Unit No.	Name of Unit	Teaching Hours	CO Targeted
1	**Introduction to PL/SQL**	**10**	**CO1**
PL/SQL: Structure, Elements, Data types, Control Structure, Iterative Control,Views, Cursors - Stored Procedure and Function, Exceptional Handling and Errors, Triggers.			
2	**Transaction Management**	**5**	**CO2**
Define Transaction, Properties of Transaction, States of Transaction , The problems associated with concurrently executing transactions What is Schedule Types of Schedule Concept of Serializability, Precedence graph for serializability			
3	**Concurrency Control Techniques**	**7**	**CO3**
Ensuring serializability by locks, Lock Based Protocols ,Different locks modes, 2PL & it's Variation, Timestamp Based Protocol, Thomas Writes rule, Multiple Granularity Phantom problem, Deadlock & Deadlock handling technique.			

4.	**Database Recovery Technique**	**6**	**CO4**
Recovery concepts, Log based Recovery technique (Deferred and Immediate), Checkpoints, Shadow paging, Aries Algorithm, Recovery with concurrent Transactions. Database backup and Recovery from Catastrophics failures.			
5	**Advanced Topics in Database**	**2**	**CO5**
Parallel Database, Distributed Database, Multimedia Database , Mobile Database , Web Database , Multidimensional Database, Data Warehouse , OLTP Vs OLAP, Bigdata, NoSQL Databases			

Reference Books

1. **Book1** - Database Management Systems, Raghu Ramakrishnan and Johannes Gehrke
2. **Book2** - "Fundamentals of Relational Database Management System", Sumathi, S. Esakkirajan
3. **Book 3 -** Database System Concepts by Henry Korth and A. Silberschatz
4. **Book 4** - Practical Postgresql , By Joshua D. Drake, John C Worsley (O'Reillypublications)
5. **Book 5** - PostgreSQL, Korry Douglas, ISBN:9780672327568, Sams

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

DS-252-MJ-T : Data Structure II

No. of Credits:	Teaching Scheme	Examination Scheme
2	Theory: 2 Hrs./Week	Continuous Evaluation: 15 Marks End Semester : 35 Marks

Prerequisites

- Basic knowledge of Python Programming.
- Fundamental Knowledge of Array, Linked List, Stack, Queue Data Structure
- Familiarity with basic Searching and Sorting techniques

Objectives

- To understand advanced data structures and their applications in Data Science.
- To analyze and implement complex algorithms using Python.
- To improve problem-solving skills by applying data structures to real-world problems.
- To introduce students to efficient searching, sorting, and graph-based techniques.
- To enhance the ability to optimize performance in large-scale data-driven applications

Course Outcomes

On Completion of this course, student will be able to -

CO1: Understand the Fundamental Concepts of Non-Linear Data Structures

CO2: Develop optimized algorithms for data science applications

CO3: Apply graph algorithms for solving real-world problems.

CO4: Evaluate the efficiency of algorithms using Time and Space complexity.

Unit No.	Name of Unit	Teaching Hours	CO Targeted
1	**Tree**	10	CO1

Concept and Terminologies

Types of Binary Tree-Binary Tree, Skewed tree, Strict Binary Tree, Full Binary Tree, Complete

Binary Tree, Expression tree, Binary Search Tree, Heap

Representation

Static and Dynamic Representation of tree

Implementation and Operations on Binary Search Tree

Create, Insert, Delete, Search

Tree Traversal-preorder, inorder, postorder ,Level order traversal using queue

Counting Leaf node, Non leaf nodes, Copy, Mirror

Application of Trees

Heap Sort

Introduction to Greedy Strategy, Huffman encoding (Implementation Using Priority Queue)

Applications of tree in Data Science

Decision tree in machine learning

Hierarchical clustering in data analysis			
2	**Efficient Search Tree**	**4**	**CO1,CO2**
Terminology- Balanced trees-AVL Trees, Red Black tree, Splay Tree, Lexical Search tree-Trie, Multiway search Trees-B and B+ tree **AVL Tree-** Concept and Rotations			
3	**Graph**	**10**	**CO1,CO2,CO3, CO4**
Concept and Terminology Graph representation-Adjacency Matrix, Adjacency List Graph Traversals-Breadth First Search, Depth First Search (with implementation) **Applications of Graph -** Topological Sorting Use of Greedy Strategy in Minimal Spanning Trees(Prim's and Kruskal Algorithm) Single Source Shortest Path-Dijkstra's Algorithm Dynamic Programming Strategy, All Pair shortest Path(Floyd Warshall Algorithm) **Applications of Graph in Data Science** Use of Graph in Social Network Graph based Machine Learning Transportation & Route Optimization			
4	**Hash table**	**6**	**CO1, CO2**
Concept of Hashing Terminologies-Hash Table, Hash Function, Bucket, Hash address, Collision, Synonym, Overflow Properties of good Hash function Hash Function-Division function, MID Square, Folding Method **Collision Resolution Techniques** Open Addressing-Linear Probing, Quadratic Probing, Rehashing Chaining: Coalesced, Separate Chaining **Applications of Hash Table in Data Science** Efficient Lookup in Datasets : Searching specific values or attributes in large datasets Counting Frequencies : Counting frequencies of words in large datasets Caching Results : Storing computed results of expensive operations for reuse			
Reference Books			
5. Data Structures and Algorithms in Python Authored by Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser 6. Data Structures Using Python by Shriram K. Vasudevan, 7. Data Structures and Algorithm Analysis in Python by Mark Allen Weiss 8. Python Data Structures and Algorithms by Benjamin Baka			

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

DS-253-MJ-P : Lab Course on DS-251-MJ-T and DS-252-MJ-T
(Relational Database Management System and Data Structure-II)

No. of Credits:	Teaching Scheme	Examination Scheme
2	Practical: 4 Hrs./Week	Continuous Evaluation: 15 Marks End Semester : 35 Marks

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 Outcomes

On completion of this course students will able to:

CO1: To use PostgreSQL (PL/SQL.)
CO2: To perform advanced database operations (View ,Function, cursor, Trigger)
CO3: Implement Fundamental Concepts of Non-Linear Data Structures
CO4: Implement nonlinear Data Structure dynamically with their application

Operating Environment:

For DBMS:

- Operating System: Linux or any relevant Operating System
- DBMS: PostgreSQL

For Data Structures:

- Python is cross-platform and can run on:
  - **Windows** (Windows 10/11)
  - **Linux** (Ubuntu, Fedora, Debian, etc.)
  - **macOS** (Latest versions)

To write and execute Python programs efficiently, you can use:

- **VS Code** (Lightweight, supports extensions)
- **PyCharm** (Best for large projects, auto-suggestions)
- **Jupyter Notebook** (Great for data visualization & step-by-step execution)
- **IDLE** (Comes pre-installed with Python, good for beginners)
- **Spyder** (Best for scientific computing)

Course Contents:

Relational Database Management Systems Assignments

Assignment 1: Views

1) To Create and Drop Views
Assignment 2 : Stored Procedure
1) Simple Stored Procedure 2) Stored Procedure with IN, OUT and IN/OUT parameter
Assignment 3: Stored Function
1) Simple Stored Function 2) Stored Function that returns
Assignment 4 : Cursors
1) Simple Cursor 2) Parameterize Cursor
Assignment 5 : Exception Handling
1) Simple Exception- Raise Debug Level Messages 2) Simple Exception- Raise Notice Level Messages 3) Simple Exception- Raise Exception Level Messages
Assignment 6 : Triggers
1) Before Triggers (insert, update, delete) 2) After Triggers (insert, update, delete)
Assignments for Data Structures – II
Assignment 1: Binary Search Tree and Traversals
1. Implement Binary Search Tree (BST) to perform following operations on BST– Create, Recursive Traversals - Inorder, Preorder, Postorder **2.** Perform following operations: insert, delete
Assignment 2: Binary Search Tree Operations
1. Implement Binary Search Tree (BST) to perform following operations on BST– Copy and mirror image of BST, counting leaf, non-leaf and total nodes. 2. Level-order traversal of binary search tree using queue.
Assignment 3: Applications of Binary Tree
1. Sort set of elements using Heap sort 2. Encode a set of characters using Huffman encoding
Assignment 4: Graph implementation
1 Implement Graph as adjacency matrix and adjacency list 2. Calculate indegree and out degree of vertices 3. Graph traversals: BFS and DFS.
Assignment 5: Graph Applications - I
1. Implementation of Topological sorting 2. Implementation of Prims/Kruskals Minimum spanning tree algorithm
Assignment 6: Graph Applications - II
1. Implementation of Dijkstra's shortest path algorithm for finding Shortest Path from a given source vertex using adjacency cost matrix. 2. Implementation of Floyd Warshall algorithm for all pairs shortest path.
Assignment 7: Hash Table
1. Implementation of static hash table with Linear Probing. 2. Implementation of static hash table with chaining

Books

1. Database Management Systems, Raghu Ramakrishnan and Johannes Gehrke
2. “Fundamentals of Relational Database Management System”, Sumathi, S. Esakkirajan
3. Database System Concepts by Henry Korth and A. Silberschatz
4. Practical Postgresql , By Joshua D. Drake, John C Worsley (O’Reillypublications)
5. PostgreSQL, Korry Douglas, ISBN:9780672327568, Sams
6. Data Structures and Algorithms in Python Authored by Michael T. Goodrich, Roberto Tamassia, and Michael H. Goldwasser,
7. Data Structures Using Python by Shriram K. Vasudevan,
8. Data Structures and Algorithm Analysis in Python by Mark Allen Weiss
9. Python Data Structures and Algorithms by Benjamin Baka

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

DS-271-VSC-P: Data Analytics

No. of Credits: 02	Teaching Scheme Practical: 4 Hours/Week	Examination Scheme Continuous Evaluation: 15 Marks End Semester: 35 Marks

Prerequisites

- Basic knowledge of logic and Python programming concepts
- Knowledge of problem-solving tools like algorithms, flowcharts and pseudo codes will be an added advantage

Objectives

- To explore the fundamental concepts of data analytics.
- To understand the various methods and tools for data analytics.
- Design a simple database.
- Query a database using SQL in Python

Course Outcomes

On Completion of this course, student will be able to -

CO1: Use different data analytics functions to solve a problem.
CO2: Apply different data cleaning functions on a given data.
CO3: Explain basic excel formulas and functions and apply them on a data set.
CO4: Design a simple database with DDL and DML commands.
CO5: Write sub queries and join operations for retrieving data from various tables.
CO6: Use the aggregations and group operations for data analysis in python.

Guidelines:

Operating Environment: Excel, PostgreSQL, Python

List of Assignments:

Excel:

Assignment 1 : Getting Started with Excel

- Creation of spread sheets, Insertion of rows and columns, Drag & Fill, use of Aggregate functons.

Assignment 2 : Working with Data

- Importing data, Data Entry & Manipulation, Sorting & Filtering, Data Validation, Pivot Tables & Pivot Charts

Assignment 3 : Data Analysis Process

- Conditional Formatting, What-If Analysis, Data Tables, Charts & graphs

Assignment 4 : Cleaning Data with Text Functions and Cleaning Data Containing Date and Time Values

- Use of UPPER and LOWER, TRIM function, Concatenate, use of DATEVALUE function, DATEADD and DATEDIF, TIMEVALUE functions.

Assignment 5 : Conditional Formatting

- Formatting, parsing, and highlighting data in spreadsheets during data analysis.

Assignment 6 : Working with Multiple Sheets

- Work with multiple sheets within a workbook is crucial for organizing and managing data, perform complex calculations and create comprehensive reports

Assignment 7 : Generation of report & presentation using Autofilter & macro

SQL (Using Python)

Connect to PostgreSQL database using Python and Write SQL queries.

- Filtering and sorting data (SELECT, WHERE, ORDER BY), Aggregation (SUM(), AVG(), COUNT()), Joining multiple tables (JOIN), Fetch results and analyze them in Pandas.

Python:

Assignment 1: Load a dataset and perform some primary operations.

- info(), describe(), shape, size, loc, sort_values, value_counts() used to describe the dataframe, Handle missing values, Remove duplicate records, Standardize column names
- Getting introduced to essential packages like NumPy, SciPy, and pandas from the Data Analytics point of view.

Assignment 2 : Basic statistical operations

- Apply basic statistical operations on a dataset- Compute the mean, median, mode, range, quartiles, and variance for one or more attributes.

Assignment 3 : Data pre-processing

Apply data pre-processing techniques

- Describing data set
- Shape the data set
- Displaying specific rows from the data set
- Partition them into appropriate number of bins by equal-frequency as well as equal-width partitioning.

Assignment 4 : Data Visualization

- View the data using various 2-D, 3-D plots and charts to detect and handle outliers

Reference Books

1. How to solve it by Computer, R.G. Dromey, Pearson Education.
2. Programming Python, O`Reilly, 4th Edition, 2010

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

DS-281-FP : Mini Project
(Software Engineering Project)

No. of Credits: 2	Teaching Scheme Practical: 4 Hrs/Week	Examination Scheme Continuous Evaluation: 15 Marks End Semester : 35 Marks

Prerequisites

- Fundamental Knowledge of Software Engineering, RDBMS

Objectives

- Apply software engineering principles to design, develop a small-scale software project.
- Demonstrate understanding of software development life cycles, methodologies, and tools.
- Develop teamwork, communication and problem-solving skills.

Course Outcomes

On Completion of this course, student will be able to -

CO1: Understand software engineering discipline.

CO2: The project should have a clear set of requirements, design documents.

CO3: The project can be developed using any process model.

Unit No.	Name of Unit
1	**Project Proposal**
• Project title and description • Project goals and objectives • Technical specifications and requirements	
2	**Problem Definition and Scope**
• Problem Description • Study of Existing system(Manual or computerized) • Drawbacks of Existing system • Scope of the Proposed System	
3	**Feasibility Study**
• Technical Feasibility • Economical Feasibility • Operational Feasibility	
4	**Fact Finding Techniques**
• Interview • Questionary • Survey • Observations	

5	**Entity Relationship Diagram**
• Identify entities & attributes • Identify relations in entities	
6	**UML Diagrams**
• Class Diagram • Use case Diagram • Sequence Diagram • Activity Diagram • Component Diagram • Deployment Diagram	
7	**Data Dictionary**
• Designing the Database • Normalized to 3NF	
8	**Designing queries related to Functional requirements**
• Simple Queries • Nested Queries(using Aggregate Functions) • Views • Functions • Cursors • Trigger	
Reference Books	
1. Book1:Software Engineering", by Roger Pressman, 8th Edition 2. Book2: Practical PostgreSQL", O'Reilly Publications	

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

DS-291-MN-T : Testing of Hypothesis and Sampling Distributions

No. of Credits: 02	Teaching Scheme **Theory:** 2 Hours/Week	Examination Scheme Continuous Evaluation:15 Marks End Semester : 35 Marks

Prerequisites

- Basics of Distribution

Objectives

- To study the hypothesis
- To perform the large sample test
- To overview the sampling distribution
- To understand the difference between small sample and large sample test
- To study hypothesis testing based on sampling distributions

Course Outcomes

On Completion of this course, student will be able to –

CO1: Identify the difference between statistic, estimate and parameter.

CO2: Formulate the null and alternative hypotheses

CO3: Calculate type-I and type-II error, p-value

CO4: Apply small, large sample tests in real life problems

CO5: Derive probability distribution function of chi-square, t, F distribution

CO6: Explains interrelation between the above distributions and their properties.

CO7: Get familiar with statistical tests of hypothesis and are able to apply in real life situations in various fields

Unit No.	Name of Unit	Teaching Hours	CO Targeted
1	**Testing of Hypothesis**	**08**	**CO1, CO2, CO3**

Statistic, estimate and parameter. Sampling distribution of a statistic, standard error of a statistic with illustrations. Statistical Inference: Introduction to problem of Estimation and testing of hypothesis. Difference between estimator and estimate. Concept of unbiased estimator. Point and interval estimation. Statistical hypothesis, null and alternative hypothesis, simple and composite hypothesis, one sided and two-sided alternative hypothesis, critical region, type-I and type-II error, level of significance, p-value. Two-sided confidence interval. Tests of hypotheses using i) critical region approach, ii) p-value approach and iii) confidence interval approach.

Tests for population means (large sample / approximate tests): Testing of population mean (μ) with a specified value (μ_0) for the known variance. Testing equality of two population means $(\mu_1 = \mu_2)$ for the known variance. Construction of two-sided confidence interval for population mean μ and for the mean difference $\mu_1 - \mu_2$.

Testing of population proportion (P) with a specified value (P_0). Testing equality of two population proportions $(P_1 = P_2)$. Construction of two-sided confidence interval for $P\ an\ P_1 - P_2$.

2	**Chi-square Distribution**	**08**	**CO5**

Definition of chi-square random variable as a sum of squares of independent and identical standard normal variables. Derivation of the p.d.f. of Chi-square variable with n degrees of freedom (d.f.) using MGF. Mean, variance, MGF, CGF, central moments skewness, kurtosis, mode, additive property. Use of chi-square tables for calculations of probabilities. Normal approximation: $\frac{\chi_n^2 - n}{\sqrt{2n}}$ (statement only) Distribution of $\underline{X}$ and $\frac{nS^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n} \left(X_i - \underline{X}\right)^2$ for a random sample from a normal distribution using orthogonal transformation, independence of $\underline{X}$ and S^2.

3	**t-Distribution**	**04**	**CO5**

Definition of r.v. with n d.f. in the form of $= \frac{U}{\sqrt{\frac{V}{n}}}$, where $U \sim N(0,1)$ and V is chi-square with n d.f., where U and V are independent random variables. Derivation of the p.d.f of distribution, nature of probability curve, mean, variance, moments, mode. Use of t-tables for calculations of probabilities, Normal approximation (statement only).

4	**Snedecore's F-Distribution**	**05**	**CO5**

Definition of r.v. with n_1 and n_2 d.f. as $F_{n_1,n_2} = \frac{X_1 n_1}{X_2 n_2}$ where & are independent chi-square variables with n_1 and n_2 d.f. Derivation of the p.d.f, nature of probability curve, mean, variance, moments, mode. Distribution of $\frac{1}{F_{n_1,n_2}}$, use of tables for calculation of probabilities. Interrelationship between Chi-square, t and F-distributions.

5	**Tests based on Sampling Distributions**	**05**	**CO6, CO7**

Tests based on chi-square distribution: Test for independence of two attributes, Test for goodness of fit, Test for variance for known and unknown mean.

Tests based on t-distribution: Tests for population means, Paired t-test for one-sided and two-sided alternatives.

Test based on F-distribution: Test for $H_0: \sigma_1^2 = \sigma_2^2$ against one-sided and two-sided alternatives when i) means are known and ii) means are unknown. Take $F = \frac{S_1^2}{S_2^2}$.

Reference Books

1. Goon, A. M., Gupta, M. K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1, Sixth Revised Edition, The World Press Pvt. Ltd., Calcutta.
2. Gupta, S. C. and Kapoor, V. K. (2002), Fundamentals of Mathematical Statistics, (Eleventh Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi, 110002.
3. Kulkarni, M. B., Ghatpande, S. B. and Gore, S. D. (1999), Common Statistical Tests, Satyajeet Prakashan, Pune 411029
4. Medhi, J., Statistical Methods, Wiley Eastern Ltd., 4835/24, Ansari Road, Daryaganj, New Delhi – 110002.
5. Meyer, P. L., Introductory Probability and Statistical Applications, Oxford and IBH Publishing Co. New Delhi.
6. Mood, A. M. and Graybill, F. A. and Boes D.C. (1974). Introduction to the Theory of Statistics, Ed. 3, McGraw Hill Book Company.

7. Mood, A. M., Graybill F. A. and Bose, F. A. (1974), Introduction to Theory of Statistics (Third Edition, Chapters II, IV, V, VI), McGraw - Hill Series G A 276
8. Neil, A. Weiss, (2016). Introductory Statistics, Tenth Edition, Pearson.
9. Purohit, S. G., Gore S. D., Deshmukh S. R. (2008). Statistics Using R, Narosa Publishing House, New Delhi.
10. W. and Cochran W. G. (1989). Statistical Methods, Eighth Ed. East West Press.

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

DS-292-MN-P : Lab Course on DS-291-MN-T

(Testing of Hypothesis and Sampling Distributions)

No. of Credits: 2	**Teaching Scheme** Practical: 4 Hrs/Week	**Examination Scheme** Continuous Evaluation: 15 Marks End Semester : 35 Marks

Prerequisites

- Basics of Distribution

Objectives

Practical Implementation of

- Hypothesis
- Performing the large sample test
- The sampling distribution
- The difference between small sample and large sample test
- Hypothesis testing based on sampling distributions

Course Outcomes:

At the end of this course, students will be able to:
CO1: To draw conclusions and determining probability of making errors in hypothesis tests using critical values.
CO2: To conduct various tests of significance like averages, population proportions, independence of attributes, variance etc.

List of Assignments

Sr. No.	Assignment
1	Test for proportions (one sample problem and two sample problem)
2	Test for means (large sample test).
3	Test for means (one sample problem) (small sample test).
4	Test for means (two sample problem) (small sample test).
5	Paired t-test.
6	Test for independence of attributes
7	χ^2 goodness of fit test.
8	Test based on χ^2 distribution : $H_0: \sigma^2 = \sigma_0^2$ for μ unknown
9	Tests for population variance (two sample problem only).

Reference Books

19. Goon, A. M., Gupta, M. K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1, Sixth Revised Edition, The World Press Pvt. Ltd., Calcutta.

20. Gupta, S. C. and Kapoor, V. K. (2002), Fundamentals of Mathematical Statistics, (Eleventh Edition), Sultan Chand and Sons, 23, Daryaganj, New Delhi, 110002.
21. Kulkarni, M. B., Ghatpande, S. B. and Gore, S. D. (1999), Common Statistical Tests, Satyajeet Prakashan, Pune 411029
22. Medhi, J., Statistical Methods, Wiley Eastern Ltd., 4835/24, Ansari Road, Daryaganj, New Delhi – 110002.
23. Meyer, P. L., Introductory Probability and Statistical Applications, Oxford and IBH Publishing Co. New Delhi.
24. Mood, A. M. and Graybill, F. A. and Boes D.C. (1974). Introduction to the Theory of Statistics, Ed. 3, McGraw Hill Book Company.
25. Mood, A. M., Graybill F. A. and Bose, F. A. (1974), Introduction to Theory of Statistics (Third Edition, Chapters II, IV, V, VI), McGraw - Hill Series G A 276
26. Neil, A. Weiss, (2016). Introductory Statistics, Tenth Edition, Pearson.
27. Purohit, S. G., Gore S. D., Deshmukh S. R. (2008). Statistics Using R, Narosa Publishing House, New Delhi.
28. W. and Cochran W. G. (1989). Statistical Methods, Eighth Ed. East West Press.

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

SEC-251-DS-T: Software Engineering

No. of Credits: 2	**Teaching Scheme** Theory: 2 Hrs/Week	**Examination Scheme** Continuous Evaluation: 15 Marks End Semester: 35 Marks

Prerequisites

Fundamentals of programming language, Databases, ER Modeling

Objectives

1. To get knowledge and understanding of software engineering discipline.
2. To learn analysis and design principles for software project development.
3. To become familiar with software development life cycle models and their applications.
4. To understand software testing strategies and techniques.

Course Outcomes

On Completion of this course, student will be able to -

CO1: Understand software engineering discipline.

CO2: Compare and choose a process model for a software project development.

CO3: Identify requirements analyze, prepare models and Agile Models.

CO4: Apply design concepts and metrics in software development.

CO5: Implement testing strategies to ensure software quality.

Unit No.	Name of Unit	Teaching Hours	CO Targeted
1	**Software Engineering Fundamentals**	**3**	**CO1**
1.1 Definition of Software and the evolving role of software 1.3 Changing nature of software 1.4 Software myths. 1.5 Software Process 1.5.1 The Process Framework 1.5.2 Umbrella Activities 1.5.3 Process Adaptation			
2	**Software Development Life Cycle (SDLC)**	**5**	**CO2**
2.1 Generic Process Model 2.2 Prescriptive Process Models 2.2.1 The Waterfall Model 2.2.2 Incremental Process Models 2.2.3 Evolutionary Process Models 2.2.4 Concurrent Models 2.2.5 The Unified Process 2.6 Reverse Engineering			
3	**Agile Development**	**6**	**CO3**
3.1 Agility and Agile Process			

3.2 Agility Principles
3.3 The Politics of Agile Development
3.4 Human Factors
3.5 Extreme Programming (XP)
3.5.1 XP Values
3.5.2 XP Process
3.6 Adaptive Software Development (ASD)
3.7 Scrum
3.8 Dynamic System Development Model (DSDM)
3.9 Agile Unified Process (AUP)

4	**Software Requirement Analysis and Design Using UML**	**12**	**CO3,CO4**

4.1 Requirement Elicitation
4.2 Software requirement specification (SRS)
4.3 Building the Analysis Model
4.4 Negotiating Requirements
4.5 Validating Requirements
4.6 Introduction to UML
4.7 Structural Modeling
4.7.1 Use case model
4.7.2 Class model
4.8 Behavioral Modeling
4.8.1 Sequence model
4.8.2 Activity model
4.8.3 Collaboration model
4.9 Architectural Modeling
4.9.1 Component model
4.9.2 Deployment model

5	**Software Testing**	**4**	**CO5**

5.1 Introduction to Software Testing
5.2 Software Testing Life Cycle (STLC)
5.3 Software Testing Strategies & Techniques
5.3.1 Black box and White box Testing
5.4 Level of Testing
5.4.1 Unit Testing
5.4.2 Integration Testing
5.4.3 System Testing

Reference Books

1. Software Engineering : A Practitioner's Approach - Roger S. Pressman, McGraw hill(Eighth Edition) ISBN-13: 978-0-07-802212-8, ISBN-10: 0-07-802212-6
2. The Unified Modeling Language Reference Manual - James Rambaugh, Ivar Jacobson, Grady Booch ISBN 0-201-30998-X
3. Effective Methods of Software Testing, William E Perry, 3rd Edition, Wiley Publishing Inc

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

Course Type: GE/OE. Course Code: OE-251-DS-T
Course Title: Ecommerce II
(To be offered to faculty other than Science and Technology)

Teaching Scheme 02 Hours /Week	No. of Credits 2	**Examination Scheme** Continuous Evaluation: 15 Marks End Semester: 35 Marks

Prerequisites

- Knowledge of basic e-commerce concepts, business models and payment systems.
- Understanding of digital marketing, financial systems, and internet technologies

Course Objectives

- To understand the technical and security aspects of e-commerce.
- To explore data-driven decision-making and analytics in e-commerce.
- To study supply chain and logistics management in e-commerce.
- To gain insights into global e-commerce trends and challenges.
- To learn about the integration of AI, Blockchain, and Cloud Computing in e-commerce.

Course Outcomes

On completion of the course, student will be able to–

CO1: Implement secure e-commerce transactions and protect user data.

CO2: Apply analytics tools to track and enhance e-commerce performance.

CO3: Manage e-commerce logistics and understand global trends.

CO4: Use emerging technologies such as AI, Blockchain, and Cloud for e-commerce applications.

Course Contents

Chapter 1	**E-Commerce Logistics and Supply Chain Management**	**6 Hours**

1.1. E-Commerce Supply Chain Overview
1.2. Inventory Management in E-Commerce
1.3. Role of Warehousing and Fulfillment Centers
1.4. Last-Mile Delivery Challenges
1.5. Reverse Logistics and Return Policies
1.6. Green and Sustainable Supply Chains

Chapter 2	**Data Analytics and Consumer Behavior in E-Commerce**	**7 Hours**

2.1. Importance of Data Analytics in E-Commerce
2.2. Key Performance Indicators (KPIs) and Metrics

2.3. Web Analytics (Google Analytics, Heatmaps, A/B Testing) 2.4. Predictive Analytics & Customer Insights 2.5. Recommendation Engines & Personalization 2.6. Conversion Rate Optimization (CRO) 2.7. Fraud Detection Using AI in E-Commerce		
Chapter 3	**E-Commerce Security and Privacy**	**6 Hours**
3.1. Importance of Security in E-Commerce 3.2. Threats to E-Commerce (Phishing, Fraud, Cyber Attacks) 3.3. Cryptography & Secure Transactions (SSL/TLS, Encryption) 3.4. Digital Signatures & Certificates 3.5. Firewalls & Intrusion Detection Systems		
Chapter 4	**Advanced E-Commerce Technologies**	**5 Hours**
4.1. Machine Learning for Product Recommendations 4.2. Blockchain for Secure Transactions and Smart Contracts 4.3. Cloud Computing and SaaS Platforms for E-Commerce 4.4. The Role of 5G in E-Commerce Growth		
Chapter 5	**Global Trends of E-Commerce**	**6 Hours**
5.1. Cross-Border E-Commerce and Global Expansion 5.2. Mobile Commerce (M-Commerce) Innovations 5.3. Subscription-Based E-Commerce Models 5.4. Social Commerce (Instagram, Facebook Shops, TikTok Commerce) 5.5. Ethical and Sustainable E-Commerce Practices 5.6. Future Challenges in E-Commerce		
Reference Books:		
1. E-Commerce Analytics: Analyze and Improve the Impact of Your Digital Strategy, Judah Phillips,Pearson 2. Cloud Computing: Principles and Paradigms, Rajkumar Buyya, James Broberg, Andrzej Goscinski, Wiley 3. Global E-Commerce: Theory and Case Studies, Jie Lin, Fei Gao, Springer 4. Logistics and Supply Chain Management, Martin Christopher, Pearson		

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

Course Type: GE/OE Course Code: OE-252-DS-T

Course Title: Web Design II

(To be offered to faculty other than Science and Technology)

Teaching Scheme 02 Hrs/ week	No. of Credits 2	**Examination Scheme** Continuous Evaluation: 15 Marks End Semester: 35 Marks

Prerequisites

- Knowledge of other programming languages will enhance your abilities with HTML and web development.
- Knowledge and understanding of Internet.

Course Objectives

- To learn to design Style sheets.
- To know to link and publish Web pages.
- To prepare the learners with the fundamentals of HTML programming and scripting languages.
- Learners should know how to create HTML documents, applying styles using CSS.

Course Outcomes

On completion of the course, student will be able to–

- Learn and use the CSS to design Web Pages.
- Link and publish Web pages.
- Develop and create a webpage by planning, designing, and implementing everything themselves.
- Understand and develop a dynamic web pages using JavaScript (client side programming) .

Course Contents

Chapter 1	**CSS**	**8**

1.1. Introduction of CSS and its Syntax
1.2. Ways to Insert CSS and Background image handling
1.3. Background colour management using CSS
1.4. Text and Font management using CSS
1.5. Managing Hyperlinks and List using CSS
1.6. Designing Borders and Outline
1.7. Setting Page Margin using CSS

Chapter 2	**XML**	**8**

2.1. XML Namespaces and Infoset and Document Type
2.2.Definitions (DTDs)

2.3.XML Schemas and XML-Parser 2.4.Data Modeling , Document and Object Model (DOM) 2.5.Displaying XML with XSLT		
Chapter 3	**Introduction to JavaScript**	**8**
3.1. Concept of Script, Types of Scripts 3.2. Introduction to JavaScript. 3.3. Variables, identifier and Operator, Control structure. 3.4. Examples on JavaScript Operators. 3.5. Functions 3.6. Event Handling in JavaScript with examples.		
Case Study		**6**
Case study 1: Creation of forms, small case study to create HTML pages using all the above learnt techniques. **Case study 2:** Redesigning the Website of a Small Business. **Case study 3:**Create a Styled Web Page for a Coffee Shop.		
Reference Books:		
1. Learning Web Design: A beginner's Guide To HTML, CSS, Javascript, and Web Graphics - Jennifer Robbins 2. HTML5: The Missing Manual - Matthew MacDonald 3. HTML and JavaScript – Ivan Bayross 4. Mastering HTML, CSS &Javascript Web Publishing		
Reference Link: **1.** JavaScript Guide - JavaScript | MDN **2.** JavaScript reference - JavaScript | MDN **3.** XML.com		

Savitribai Phule Pune University

B.Sc. Data Science (2024 Pattern)

Sem-IV

Course Type: GE/OE Course Code: OE-253-DS-T

Course Title: Digital Marketing- II

(To be offered to faculty other than Science and Technology)

Teaching Scheme	No. of Credits 2	**Examination Scheme**
02 Hours /Week		Continuous Evaluation: 15 Marks End Semester: 35 Marks

Prerequisites

- Digital marketing requires creativity and problem-solving abilities.
- Experience with social media platforms (Facebook, Instagram, Twitter, LinkedIn, etc.) is beneficial, as digital marketing

Course Objectives

- To understand Digital Marketing as the most powerful marketing tool.
- To Learn to create digital marketing artworks.
- To use social media sites like Facebook, Instagram, Twitter, LinkedIn, and others to raise sales, engage customers, and establish your brand.

Course Outcomes

On completion of the course, student will be able to–

CO1: Students will be able to communicate marketing strategies and results effectively to stakeholders.

CO2: Student will be able to assess and enhance digital marketing campaigns' return on Investment.

CO3: Students will gain practical experience with industry-standard digital marketing tools.

CO4: Student will be adept at using a variety of social media channels to create and interact with communities, raise awareness of a brand.

Course Contents

Chapter 1	**Online Consumer Behaviour Analysis**	**8**

1.1 Consumer Behavior
1.2 Segmentation and targeting online customers
1.3 Psychological Responses
1.4 Social Trends

Chapter 2	**Social Media Marketing**	**8**

2.1. Social Media Sites
2.2. Influence of Social Media Marketing
2.3. Power of Social Media
2.4. Monetization through Social Media

Chapter 3	**Future of Digital Marketing**	**8**

3.1. Use of Artificial Intelligence (AI) in Digital Marketing.
3.2. Common use of household gadgets for online marketing.

3.3. Digital Marketing strategies.	
Case Study	**6**
Case Study 1 : Experiential Learning : Creating a website. Case Study 2 : Online Consumer Behavior Analysis for an E-Commerce Fashion Brand Case Study 3	
Reference Books:	
1. Digital Marketing : Nitin Kamat, Chinmay Kamat (Himalaya Publishing House) 2. "Digital Marketing for Dummies" by Ryan Deiss and Russ Henneberry 3. "Influence: The Psychology of Persuasion" by Robert B. Cialdini 4. "Social Media Marketing Workbook: How to Use Social Media for Business" by Jason McDonald	
Reference Links: 1.https://www.socialmediaexaminer.com/ 2. https://www.marketingprofs.com/	

Savitribai Phule Pune University
B.Sc. Data Science (2024 Pattern)
Sem-IV

Course Type: GE/OE Course Code: OE-254-DS-T

Course Title: AI for Everyone- II

(To be offered to faculty other than Science and Technology)

Teaching Scheme: 2hours / week	No. of Credits: **2**	**Examination Scheme** Continuous Evaluation: 15 Marks End Semester: 35 Marks

Course Objectives: -

1. Understand the basics of artificial intelligence and its subfields.
2. Explore real-world applications of AI across different industries.
3. Gain insights into the ethical, social, and economic implications of AI.
4. Develop an appreciation for the potential of AI to drive innovation and transformation.

Course Outcomes: -On completion of the course, student will be able to–

CO1: To understand different types of AI Models
CO2: To understand content optimization using AI.
CO3: To understand Animations and motions in AI
CO4: To Understand uses of AI tools.

Course Contents

Unit 1	**Advanced AI Fundamentals**	**6 hours**	**CO1**

1.1 Deep Dive into AI, Machine Learning & Deep Learning
1.2 Types of AI Models: Generative AI, NLP, Computer Vision, Reinforcement Learning
1.3 Latest AI Trends: AGI, Large Language Models (LLMs), and multimodal AI
1.4 Exploring AI Frameworks & APIs: Open AI, Hugging Face, Google AI

Unit 2	**AI for Advanced Text & Content Creation**	**8 hours**	**CO2**

2.1 AI for Long-form Writing & Reports
2.2 Automating Research & Citation Management
2.3 AI for SEO & Content Optimization
2.4 Using AI for Professional Emails & Business Writing

Unit 3	**Advanced AI for Image & Video Processing**	**8 hours**	**CO3**

3.1 AI Image Generation Beyond Basics
3.2 Deepfake Technology & Ethical Concerns
3.3 AI Video Editing & Creation

3.4 AI Animation & Motion Capture			
Unit 4	**AI Tools**	**8 hours**	**CO4**
4.1 Chat GPT (Open AI) 4.2 Google Gemini (Bard AI) 4.3 Canva, Beautiful. AI, Gamma, Slides AI 4.4 Rytr, Grammarly			

References:

1. Artificial Intelligence: A Modern Approach – Stuart Russell & Peter Norvig.
2. Practical AI for Business Leaders – Anand S. Rao
3. AI-Powered Automation Handbook – Will Kelly
4. AI for Content Creators: How to Use AI Tools for Writing and Marketing – Rob Lennon
5. Human Compatible: Artificial Intelligence and the Problem of Control – Stuart Russell
6. https://www.grammarly.com/ai/ai-writing-tools/article-writer