

Mahatma Education Society's
Pillai HOC College of Arts, Science & Commerce (Autonomous)
Rasayani

Affiliated to University of Mumbai
NAAC Accredited with "A+" Grade in cycle II
ISO 9001:2015 Certified



SYLLABUS

Masters of Science (M. Sc.) in Data Science
F. Y. M. Sc. Data Science

As per National Education Policy 2020
Academic Year 2026-27



Mahatma Education Society's

College Code: 870

PILLAI HOC COLLEGE OF ARTS, SCIENCE & COMMERCE

Pillai HOCL Educational Campus, HOC Colony, Rasayani, Via. Panvel, Dist. Raigad. Pin 410207

Tel: 02192 - 669000 / 01 / 02 / 03 / 04 / 05 / 06 / 07 / 08 / 09

Website : www.phcasc.ac.in Email : phcasc@mes.ac.in

(NAAC Accredited 'A+' Grade , CGPA – 3.26 in Cycle 2 & ISO 9001:2015 Certified)

Affiliated to the University of Mumbai, Approved by Government of Maharashtra





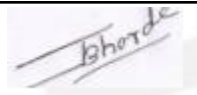
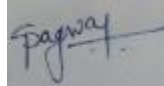
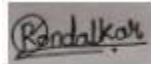
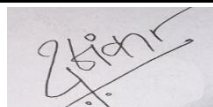


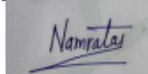



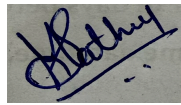
(AUTONOMOUS COLLEGE)

Ref.No: PHCASC/BOS Meeting/Attendance/2026-27

Date: 01.04.2026

Board of Studies Mathematics, Statistics and Computer Application

Sr. No	Name	Signature
1.	Dr. Swapna Kadam Vice Chancellor Nominee	
2.	Dr. Annie Rajan, Subject Experts	
3.	Dr. Homraj Patelpaik Subject Experts	
4.	Mr. Swapnil H. Patil Industry Expert	
5.	Mr. Akash Ghadge Meritorious Alumnus	
6.	Dr. Rinkoo Shantnu Principal	
7.	Ms. Priyanka Sorte Chairperson (Programme Co-ordinator BSc. CS)	
8.	Ms. Priya Prakash Member (Mathematics, Statistics)	
9.	Ms. Harshita Singh Member (Mathematics, Statistics)	
10.	Ms. Mrunal Wanjale Member (Mathematics, Statistics)	
11.	Ms. Arpita Kante Member (Programme Co-ordinator BSc. DS)	
12.	Ms. Anita Mhatre Member (Programme Co-ordinator MSc. DS)	

13.	Ms. Sangeeta Menon Member (Programme Co-ordinator BSc.AIML)	
14.	Ms. Aarti Wani Member	
15.	Ms. Rutuja Madane Member	
16.	Ms. Kranti Vartak Member	
17.	Ms. Jyoti Borade Member	
18.	Ms. Sonali Dagwar Member	
19.	Ms. Rutuja Kondalkar Member	
20.	Mr. Yash Karkhanis Member	
21.	Ms. Aishwarya Mokal Member	
22.	Ms. Vinaya Bhoir Member (Programme Co-ordinator BSc. Cyber Security)	
23.	Ms. Namrata Kedari Member	
24.	Ms. Ashwini Patil Member	
25.	Ms. Ashwini Lad Member (Programme Co-ordinator MSc.IT)	
26.	Ms. Priyanka Sonawane Member (Programme Co-ordinator BSc.IT)	
27.	Mr. Shinoj Mathew Member	

INTRODUCTION

An M.Sc. Data Science programme is designed to empower students with practical skills to analyze data, build intelligent models, and make informed decisions. It focuses on developing strong foundations in data handling, programming, and statistical thinking, which are essential for solving real-world problems.

As students progress, they gain hands-on experience with modern tools and techniques in machine learning, artificial intelligence, and big data technologies. The programme emphasizes learning by doing through projects, case studies, and research work, enabling students to apply concepts in real-life scenarios

It also nurtures critical thinking, problem-solving abilities, and effective communication skills, helping students present insights clearly to different audiences. Alongside technical knowledge, the programme highlights ethical data usage and responsible decision-making. Overall, the course prepares students to become skilled data professionals who can confidently work across industries and contribute to data-driven innovation.

Programme Outcomes (POs)

PO. No.	PO Title	PO's in Brief
PO1	Fundamental Knowledge Acquisition	Graduates will demonstrate a comprehensive and foundational knowledge of their chosen discipline along with an awareness of interdisciplinary connections.
PO2	Critical Thinking and Analytical Reasoning	Graduates will be able to analyse complex problems, synthesize data from multiple sources (qualitative and quantitative), and employ logical reasoning to formulate well-supported conclusions and arguments.
PO3	Effective Communication	Graduates will exhibit proficiency in both written and oral communication, articulating ideas clearly, persuasively, and ethically to diverse audiences
PO4	Problem Solving	Graduates will possess the ability to identify, formulate, and design solutions for real-world problems in their professional or social contexts, applying relevant theoretical knowledge and practical skills.
PO5	Information and Digital Literacy	Graduates will demonstrate the capability to locate, evaluate, and effectively use information from various sources, and utilize modern tools and Information and Communication Technology (ICT) for professional and academic tasks.
PO6	Research Skills and Scientific Temperament	Graduates will develop a sense of inquiry and research methodology, including the ability to design experiments (where applicable), collect and analyse data, and interpret results while maintaining scientific rigor and intellectual honesty.
PO7	Ethical Reasoning and Professional Integrity	Graduates will recognize ethical dilemmas, commit to professional and academic ethics, and demonstrate an understanding of moral and social responsibilities in their personal and professional conduct.
PO8	Employability and Professional Skills	Graduates will acquire the necessary job-ready skills, managerial competencies, and professional values to secure gainful employment or pursue advanced education in their respective fields.
PO9	Environmental and Sustainability Consciousness	Graduates will understand the importance of environmental conservation and sustainable development, displaying responsibility toward ecological challenges and advocating for healthy environmental practices.

PO10	Life-Long Learning	Graduates will develop the capacity for independent and self-directed learning to continuously upgrade their knowledge and skills, enabling them to adapt to rapid technological and societal changes.
PO11	Civic and Social Responsibility	Graduates will act as responsible citizens with an informed awareness of constitutional values, engaging proactively in community development and addressing social needs.
PO12	Empathy and Social Intelligence	Graduates will be able to cultivate and demonstrate affective, interpersonal, social and emotional intelligence.

Programme Specific Outcomes (PSOs)

PSOs. No.	PSO Title	PSOs in brief
PSO1	Advanced Analytical and Modeling Skills	Students will be able to apply statistical methods, machine learning, and AI techniques to analyze data and develop effective solutions for real-world problems.
PSO2	Big Data and System Competency	Students will gain the ability to design and manage data systems, including handling large-scale data using modern databases, big data tools, and data processing techniques.
PSO3	Research and Ethical Data Practices	Students will develop research-oriented thinking, apply appropriate methodologies, and understand ethical issues related to data privacy, security, and responsible use of data.
PSO4	Communication and Practical Application	Students will effectively communicate insights using data visualization tools and demonstrate adaptability by applying data science skills in diverse domains through projects and teamwork.

Evaluation Pattern

Marking Code	Marking Scheme
A	50 Marks Semester End Exam, 50 Marks Continuous Assessment (distributed within 15 Marks Class Test, 15 Marks Presentation & Assignment, 10 Marks Online Quiz, 10 Marks Attendance & Class Participation)
B	50 Marks Semester End Exam

C	100 marks Continuous Assessment (distributed within 30 Marks Class Test, 30 Marks Presentation & Assignment, 30 Marks Online Quiz, 10 Attendance & Class Participation)
D	50 Marks of Continuous Assessment (distributed within 15 Marks Class Test, 15 Marks Presentation & Assignment, 10 Marks Online Quiz, 10 Marks Attendance & Class Participation)
E	50 Marks Practical Examination (distributed within 30 Marks Practical Module 1 & 2, 10 Marks Journal, 10 Marks Viva)

Course Structure

SEMESTER I							
Course Code	Course Type	Course Title	Theory/ Practical	Marks	Credits	Lectures /Week	Evaluation Pattern
HPSDS101	Major	Foundation of Data Science	Theory	100	4	4	A
HPSDS101P	Major-Practical	Practical (HPSDS101)	Practical	50	2	2	E
HPSDS102	Major	Advance Bigdata Analytics	Theory	100	4	4	A
HPSDS102P	Major-Practical	Practical (HPSDS102)	Practical	50	2	2	E
HPSDS103P	Major	Data Visualization with Power BI	Practical	50	2	2	E
HPSDS104	Minor	Numerical Methods with Programming	Theory	100	3	3	A
HPSDS104P	Minor-Practical	Practical (HPSDS104)	Practical	50	1	1	E
HPSDS105	Major	Research Methodology for Data Science	Theory	100	4	4	A
		Total Credits		600	22		

BOS	Mathematics, Statistics and Computer Application				
Course	Foundation of Data Science				
Course Code	HPSDS101	Level	6.0		
		Type	Theory	Practical	Total
Semester	I	Credits	4	2	6
Type	Major	No of Teaching Hours	60	60	120
Evaluation Pattern	Total Marks	Semester End	Continuous		Practical
	150	50	50		50

Learning Objectives	
1	Understand the role of a data scientist and the complete process of transforming raw data into actionable insights.
2	Apply core Python programming concepts including data types, control structures, functions, and file handling for data-driven tasks.
3	Perform data cleaning and preprocessing by handling missing values, outliers, duplicates, and performing data type conversions.
4	Implement machine learning algorithms such as regression, classification, clustering, and dimensionality reduction using Scikit-learn.
5	Understand the role of a data scientist and the complete process of transforming raw data into actionable insights.

Course Outcomes	
CO1	Understand the fundamental concepts of data science and identify the role of a data scientist.
CO2	Use NumPy and Pandas to analyze, manipulate, and clean structured and unstructured data.
CO3	Analyze and visualize datasets to identify trends, patterns, and relationships.
CO4	Interpret machine learning results and visualize insights effectively.

Modules at Glance

Module No.	Content	No. of Hours	CO Mapping
1	Introduction to Data Science	15	CO1, CO2
2	Data Manipulation and Cleaning	15	CO2, CO3
3	Introduction to Machine Learning	15	CO3, CO4
4	Machine Learning Applications	15	CO4

Syllabus

Module No.	Content	No. of Lectures
1	Trails and Tribulations of Data Scientist: Data? Science? Data Science? The data Scientist, Data Science Tools, From Data to insight, Python: For something Completely Different, Introduction to Python for Data Science: Python basics: Variables, data types, control structures, functions, and modules, File I/O: Reading/writing files (CSV, JSON, Excel) using built-in functions and open(), Introduction to key libraries: NumPy for numerical computing (arrays, operations), Pandas for data structures (Series, DataFrames).	15
2	TData Manipulation and Cleaning: Basic data loading and manipulation: Importing data, indexing, slicing, and basic operations, Data cleaning: Handling missing values, outliers, duplicates; data type conversions, Time series data: Working with dates, resampling, and rolling windows using Pandas.	15
3	Introduction to Machine Learning: Regression models: Linear, multiple regression with Scikit-learn, Classification basics: Logistic regression, k-Nearest Neighbors (kNN), Model evaluation: Metrics (accuracy, precision, recall, R ²), cross-validation.	15
4	Machine Learning and Applications: Supervised learning: Decision trees, random forests, Unsupervised learning: Clustering (k-Means), Principal Component Analysis and dimensionality reduction, Ethical considerations and case studies in data science.	15
Case Study:		
1	<p>Problem: A telecom company wants to predict which customers are likely to leave (churn).</p> <p>Approach:</p> <ul style="list-style-type: none"> ● Load customer dataset (CSV) using Pandas ● Perform data cleaning (handle missing values, remove duplicates) ● Convert categorical data into numerical form ● Apply Logistic Regression / k-NN models ● Evaluate using accuracy, precision, recall 	
2	<p>Problem: An e-commerce company wants to group customers based on purchasing behavior.</p> <p>Approach:</p> <ul style="list-style-type: none"> ● Load and preprocess sales data 	

	<ul style="list-style-type: none"> • Perform data manipulation (filtering, aggregation) • Apply K-Means clustering for segmentation • Use PCA to reduce dimensions and visualize clusters
--	---

Books and References:

1. Python for Data Analysis by Wes McKinney
2. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow by Aurélien Géron
3. Introduction to Machine Learning with Python” – Andreas C. Müller & Sarah Guido
4. Data Science from Scratch” – Joel Grus
5. Practical Statistics for Data Scientists” – Peter Bruce & Andrew Bruce

Semester End Evaluation (50 Marks)

Time : 2 Hours

Paper Pattern

Question No.	Questions	Total Marks : 50
Q1	Attempt 3 out of 5	15
Q2	Attempt 3 out of 5	15
Q3	Attempt 3 out of 5	15
Q4	Case Study	05

Practical Syllabus

Sr. No	List of Practical	No. of Lectures	CO Mapping
1	Install and set up the Python environment for data science using VS Code and required libraries like NumPy, Pandas, Matplotlib, and Scikit-learn.	3	CO1
2	Perform basic Python operations using variables, data types, control structures, functions, and file I/O for reading and writing CSV files.	3	CO1
3	Explore and manipulate data using Pandas — creating Series and DataFrames, performing indexing, slicing, and basic statistical operations.	3	CO2
4	Clean and preprocess data by handling missing values, duplicates, and outliers using appropriate Pandas functions.	3	CO2

Semester End Practical Evaluation

Time: 2 Hours

Question No.	Questions	Total Marks
Q.1	Program	30
Q.2	Journal	10
Q.3	Viva & Attendance	10

Course	Advance Big Data Analytics				
Course Code	HPSDS102	Level	6.0		
		Type	Theory	Practical	Total
Semester	I	Credits	4	2	6
Type	Major	No of Teaching Hours	60	60	120
Evaluation Pattern	Total Marks	Semester End	Continuous	Practical	
	150	50	50	50	
BOS	Mathematics, Statistics and Computer Application				

Learning Objectives	
1	Understand the role and importance of Big Data
2	Discuss the use and applications of Big data in various industries
3	Identify the major technologies associated with Big data
4	Explain the roles of various components of Hadoop Ecosystems

Course Outcomes	
After successful completion of this course, students would be able to: -	
CO1	Understand the key issues in big data management and its associated applications in intelligent business and scientific computing.
CO2	Acquire fundamental enabling techniques and scalable algorithms like Hadoop, Map Reduce and Big SQL in big data analytics.
CO3	Interpret business models and scientific computing paradigms, and apply software tools for big data analytics.
CO4	Categorize between the Hadoop components and tools for processing various types of data to fulfil business requirement.

Modules at Glance

Module No.	Content	No. of Hours	CO Mapping
1	Introduction to Big Data	15	CO1, CO3
2	HDFS (Hadoop Distributed File System)	15	CO2, CO4
3	MapReduce Framework	15	CO2, CO3
4	Hadoop Ecosystem Tools	15	CO3, CO4

Syllabus

Module No.	Content	No. of Lectures
1	Introduction to big data and Hadoop: Types of Digital Data, Introduction to Big Data, Big Data Analytics, History of Hadoop, Apache Hadoop, Analysing Data with Unix tools, Analysing Data with Hadoop, Hadoop Streaming, Hadoop Echo System, IBM Big Data Strategy, Introduction to Info sphere Big Insights and Big Sheets	15
2	HDFS (Hadoop Distributed File System): The Design of HDFS, HDFS Concepts, Command Line Interface, Hadoop file system interfaces, Data flow, Data Ingest with Flume and Scoop and Hadoop archives, Hadoop I/O: Compression, Serialization, Avro and FileBased Data structures.	15
3	Map Reduce Anatomy of a Map Reduce Job Run, Failures, Job Scheduling, Shuffle and Sort, Task Execution, Map Reduce Types and Formats, Map Reduce Features. Hadoop Eco System Pig: Introduction to PIG, Execution Modes of Pig, Comparison of Pig with Databases, Grunt, Pig Latin, User Defined Functions, Data Processing operators	15
4	Hive: Hive Shell, Hive Services, Hive Metastore, Comparison with Traditional Databases, HiveQL, Tables, Querying Data and User Defined Functions. Hbase: HBasics, Concepts, Clients, Example, Hbase Versus RDBMS.	15
Case Study Scenario		
M1	<p>Problem: A large e-commerce company generates massive server log data daily. The company wants to analyze user behavior, most visited pages, and error patterns.</p> <p>Approach:</p> <ul style="list-style-type: none"> ● Store log data in HDFS ● Use Sqoop/Flume to ingest real-time and batch data ● Process data using MapReduce to count page visits and identify error logs ● Use Hive to query structured insights (e.g., top pages, peak traffic time) 	

M2	<p>Problem: A retail company wants to analyze large volumes of customer purchase data to improve marketing strategies.</p> <p>Approach:</p> <ul style="list-style-type: none"> ● Load structured and semi-structured data into HDFS ● Use Pig Latin scripts to clean and transform raw data ● Use HiveQL to perform queries like customer segmentation and sales trends ● Store processed data in HBase for fast access and real-time querying
-----------	--

Reference Books:

1. Fundamentals of Computer Graphics, Steve Marschner, Peter Shirley, CRC Press, 4th, 2016
2. The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, Ralph Kimball, Margy Ross, John Wiley & Sons, 3rd, 2013
3. Data Mining: Concepts and Techniques, Jiawei Han, Micheline Kamber, Jian Pei, Morgan Kaufmann (Elsevier), 3rd, 2011
4. Mastering Data Warehouse Design: Relational and Dimensional Techniques, Claudia Imhoff, Nicholas Gallemmo, Jonathan G. Geiger, John Wiley & Sons, 1st, 2003
5. Data Warehousing for Dummies, Alan Simon, Wiley Publishing, 1st, 2010

Semester End Evaluation (50 Marks)

Time : 2 Hours

Paper Pattern

Question No.	Questions	Total Marks : 50
Q1	Attempt 3 out of 5	15
Q2	Attempt 3 out of 5	15
Q3	Attempt 3 out of 5	15
Q4	Case Study	05

Practical Syllabus

Sr. No	List of Practical	No. of Lectures	CO Mapping
1	a. Perform installation of Hadoop in its three operating modes: Standalone, PseudoDistributed and Fully Distributed b. Use web based tools to monitor your Hadoop Setup	3	CO2
2	Implement the following file management tasks in Hadoop: a. Adding files and directories b. Retrieving files c. Deleting files Hint: A typical Hadoop workflow creates data files (such as log files) elsewhere and copies them into HDFS using one of the above command line utilities.	3	CO2, CO4
3	Run a basic word count Map Reduce program to understand Map Reduce Paradigm.	3	CO2
4	Write a Map Reduce program that mines weather data. Weather sensors collecting data every hour at many locations across the globe gather a large volume of log data, which is a good candidate for analysis with Map Reduce, since it is semi structured and record – oriented	3	CO2
5	Implement Matrix Multiplication with Hadoop Map Reduce	3	CO3
6	Install and Run Pig then write Pig Latin scripts to sort, group, join, project, and filter your data.	3	CO3
7	Install and Run Hive then use Hive to create, alter and drop databases, tables, Views, functions and indexes.	3	CO4

Semester End Practical Evaluation

Time: 2 Hours

Question No.	Questions	Total Marks
Q.1	Program	30
Q.2	Journal	10
Q.3	Viva & Attendance	10

Course	Data Visualization with Power BI				
Course Code	HPSDS103P	Level	6.0		
		Type	Theory	Practical	Total
Semester	I	Credits	-	2	6
Type	Major	No of Teaching Hours	-	60	120
Evaluation Pattern	Total Marks	Semester End	Continuous		Practical
	50	-	-		50
BOS	Mathematics, Statistics and Computer Application				

Learning Objectives	
1	Understand the fundamentals and importance of data visualization.
2	Connect, transform, and model data using Power BI tools.
3	Create interactive dashboards and reports for data analysis.
4	Use DAX (Data Analysis Expressions) for calculated fields and measures.

Course Outcomes	
After successful completion of this course, students would be able to: -	
CO1	Explain the Power BI interface and describe the business intelligence workflow.
CO2	Import, clean, and transform datasets from multiple sources using Power BI.
CO3	Develop data models and apply DAX functions for analytical calculations.
CO4	Analyze data and design effective visualizations and interactive dashboards for decision-making.

Practical Syllabus

Sr. No	List of Practical	No. of Lectures	CO Mapping
1	Explore the Power BI interface and import data from Excel and CSV files.	3	CO1
2	To clean and transform data using Power Query Editor.	3	CO2
3	Create relationships between multiple tables and build a data model.	3	CO3
4	Create calculated columns and measures using DAX functions.	3	CO3
5	Visualize data using bar, line, and pie charts in Power BI.	3	CO4
6	Add slicers, filters, and hierarchies for interactive dashboards.	3	CO4
7	Design KPI cards and gauge visuals for performance tracking.	3	CO4
8	Perform time-series analysis using DAX and date	3	CO4

	hierarchies.		
9	Publish and share Power BI reports using Power BI Service.	3	CO1, CO4
10	Create a final dashboard project combining multiple visuals and insights.	3	CO4

Semester End Practical Evaluation

Time: 2 Hours

Question No.	Questions	Total Marks
Q.1	Program	30
Q.2	Journal	10
Q.3	Viva & Attendance	10

BOS	Mathematics, Statistics and Computer Application				
Course	Numerical Methods with Programming				
Course Code	HPSDS104	Level	6.0		
		Type	Theory	Practical	Total
Semester	I	Credits	3	1	4
Type	MINOR	No of Teaching Hours	45	30	75
Evaluation Pattern	Total Marks	Semester End	Continuous	Practical	
	150	50	50	50	

Learning Objectives	
1	Identify types of computational errors and analyze their propagation.
2	Implement numerical methods for solving equations and linear systems in Python.
3	Perform interpolation, curve fitting, differentiation, and integration using Python.
4	Solve initial and boundary value problems using numerical ODE solvers and visualize results.
5	Identify types of computational errors and analyze their propagation.

Course Outcomes	
After successful completion of this course, students would be able to: -	
CO1	Understand the fundamentals of numerical methods and computational errors.
CO2	Solve nonlinear equations and systems of linear equations using numerical techniques.
CO3	Apply interpolation, curve fitting, differentiation, and integration for data analysis.
CO4	Solve ordinary differential equations (ODEs) numerically and apply them to real-world problems.

Modules at Glance

Module No.	Content	No. of Hours	CO Mapping
1	Errors and Root Finding	15	CO1
2	Linear Algebraic Methods:	15	CO2
3	a) Interpolation and Curve Fitting b) Numerical Calculus & Differential Equations	15	CO3, CO4

Syllabus

Module No.	Content	No. of Lectures
1	<p>Errors and Root Finding:</p> <p>Errors: Floating point representation, Errors: absolute, relative, truncation, round-off, Error propagation and stability</p> <p>Root Finding : Bisection, False Position, Newton-Raphson, Secant methods, Convergence analysis</p>	15
2	<p>Linear Algebraic Methods:</p> <p>Solution of linear equations: Gaussian elimination, Gauss-Jordan, LU decomposition, Iterative methods: Jacobi, Gauss-Seidel, Eigenvalues and eigenvectors (basic idea)</p>	15
3	<p>Interpolation and Curve Fitting: Lagrange interpolation, Newton forward & backward methods, Cubic spline interpolation, least squares method, Linear & polynomial regression</p> <p>Numerical Calculus & Differential Equations : Numerical differentiation (finite differences), Numerical integration: Trapezoidal, Simpson's rules, IVP: Euler, Modified Euler, Runge-Kutta , BVP: Finite difference method, Applications in data science</p>	15
Case Study Scenario		
M1	<p>Estimating Pollution Levels in a City:</p> <p>A city monitoring system collects air pollution data (PM2.5 levels) at irregular time intervals. Some data points are missing or noisy. The goal is to estimate missing values, analyze trends, and predict future pollution levels.</p>	
M2	<p>Modeling Temperature Distribution in a Rod:</p> <p>A metal rod is heated at one end, and temperature is measured at different points. The aim is to model temperature distribution and find steady-state conditions.</p>	

Recommended Reference Books:

1. S. S. Sastry — *Introductory Methods of Numerical Analysis*, PHI Learning.
2. **M. K. Jain, S. R. K. Iyengar, R. K. Jain** — *Numerical Methods for Scientific and Engineering Computation*, New Age International.
3. Steven C. Chapra, Raymond P. Canale — *Numerical Methods for Engineers*, McGraw-Hill.
4. K. E. Atkinson — *An Introduction to Numerical Analysis*, Wiley.
5. Chapra & Canale – *Numerical Methods for Engineers*
6. Burden & Faires – *Numerical Analysis*
7. Kiusalaas – *Numerical Methods in Engineering*

Semester End Evaluation (50 Marks)

Time : 2 Hours

Paper Pattern

Question No.	Questions	Total Marks : 50
Q1	Attempt 3 out of 5	15
Q2	Attempt 3 out of 5	15
Q3	Attempt 3 out of 5	15
Q4	Case Study	05

Practical Syllabus

Sr. No	List of Practical(Conducted in Python)	No. of Lectures	CO Mapping
1	Error Analysis and Stability: To understand different types of errors (absolute, relative, truncation, round-off) and study error propagation in numerical computations.	3	CO1
2	Root Finding Methods: To implement and compare basic numerical methods for solving nonlinear equations.	3	CO1
3	Gaussian Elimination and Gauss–Jordan Method: To solve a system of linear equations using direct methods.	3	CO2
4	LU Decomposition: To decompose a matrix and solve linear equations efficiently.	3	CO2
5	Iterative Methods (Jacobi & Gauss–Seidel): To solve linear equations using iterative techniques.	3	CO2
6	Eigenvalues and Eigenvectors (Basic Idea): To compute eigenvalues and eigenvectors of a matrix.	3	CO2
7	Interpolation Methods: To estimate unknown values using interpolation techniques	3	CO3, CO4
8	Curve Fitting and Regression: To fit curves to data using least squares and regression techniques.	3	CO3, CO4
9	Numerical Differentiation and Integration: To compute derivatives and integrals numerically.	3	CO3, CO4
10	Numerical Solution of Differential Equations: To solve differential equations numerically	3	CO3, CO4

Semester End Practical Evaluation**Time: 2 Hours**

Question No.	Questions	Total Marks
Q.1	Program/Problems	30
Q.2	Journal	10
Q.3	Viva & Attendance	10

BOS	Mathematics, Statistics and Computer Application				
Course	Research Methodology for Data Science				
Course Code	HPSDS105	Level	6.0		
		Type	Theory	Practical	Total
Semester	I	Credits	4	-	4
Type	MAJOR	No of Teaching Hours	60	-	60
Evaluation Pattern	Total Marks	Semester End	Continuous	Practical	
	100	50	50	-	

Learning Objectives	
1	Understand the fundamentals of research
2	Develop research problems, objectives, and hypotheses.
3	Apply data collection, sampling, and analysis techniques for interpretation.
4	Recognize the importance of research ethics.
5	Understand the fundamentals of research

Course Outcomes	
After successful completion of this course, students would be able to: -	
CO1	Explain the design and formulation of a research problem and the basics of intellectual property rights.
CO2	Analyze research-related information and apply statistical methods in research.
CO3	Conduct independent research using scientific methods to solve a defined problem.
CO4	Evaluate patent information, copyright, trademarks, and IPR tools for effective protection of innovations.

Modules at Glance

Module No.	Content	No. of Hours	CO Mapping
1	Fundamentals of Research	15	CO1
2	Research Methods & Modeling	15	CO2
3	Data Analysis & Statistical Techniques	15	CO3
4	Research Reporting, Ethics & IPR	15	CO4

Syllabus

Module No.	Content	No. of Lectures
1	Introduction to research, Definitions and characteristics of research, Types of Research, Research Process, Problem definition, Objectives of Research, Research Questions, Research design, Quantitative vs. Qualitative Approach, Building and Validating Theoretical Models, Exploratory vs. Confirmatory Research, Experimental vs. Theoretical Research, Importance of reasoning in research	15
2	Problem Formulation, Understanding Modeling & Simulation, Literature Review, Referencing, Information Sources, Information Retrieval, Indexing and abstracting services, Citation indexes, Development of Hypothesis, Measurement Systems Analysis, Error Propagation, Validity of experiments, Statistical Design of Experiments, Data/Variable Types & Classification, Data collection, Numerical and Graphical Data Analysis: Sampling, Observation, Interpretation of Results.	15
3	Statistics: Probability & Sampling distribution, Estimation, Measures of central Tendency, Arithmetic mean, Median, Mode, Standard deviation, Co efficient of variation (Discrete serious and continuous serious), Hypothesis testing & application, Correlation & regression analysis, Orthogonal array, ANOVA, Standard error, Concept of point and interval estimation, Level of significance, Degree of freedom, Analysis of variance, One way and two way classified data, 'F' test.	15
4	Preparation of Dissertation and Research Papers, Tables and illustrations, Guidelines for writing the abstract, introduction, methodology, results and discussion, conclusion sections of a manuscript. References, Citation and listing system of documents. Intellectual property rights (IPR) patents copyrights Trademarks Industrial design geographical indication. Ethics of Research Scientific Misconduct Forms of Scientific Misconduct. Plagiarism, Unscientific practices in thesis work, Ethics in science.	15
Case Study Scenario		
M1	A study was conducted to analyze the effect of social media usage on students' academic performance. Data was collected from 100 students, considering time spent on social media and their exam scores. Statistical tools like mean, correlation, and regression were used. The results showed a negative correlation, indicating that higher social media usage leads to lower academic performance. The study concluded that excessive use of social media adversely affects students' results.	
M2	An institution compared the effectiveness of online and traditional classroom learning using two groups of 50 students each. Exam scores were analyzed using mean, ANOVA, and F-test. The findings revealed that classroom learning	

	resulted in slightly better performance, with a significant difference between the two methods. The study concluded that traditional teaching is more effective in this case.
--	---

Recommended Reference Books:

1. K. S. Bordens, and B. B. Abbott, , “Research Design and Methods – A Process Approach”, McGraw Hill
2. C. R. Kothari, “Research Methodology – Methods and Techniques”, New Age International Publishers
3. Douglas C. Montgomery & George C. Runger, Applied Statistics & probability for Engineers, Wiley
4. Robert P. Merges, Peter S. Menell, Mark A. Lemley, “Intellectual Property in New Techn
5. Fundamentals of Intellectual Property – Anita Bhatt

Semester End Evaluation (50 Marks)

Time : 2 Hours

Paper Pattern

Question No.	Questions	Total Marks : 50
Q1	Attempt 3 out of 5	15
Q2	Attempt 3 out of 5	15
Q3	Attempt 3 out of 5	15
Q4	Case Study	05