R. P. Gogate College of Arts & Science and

R. V. Jogalekar College of Commerce (Autonomous),

Ratnagiri

Master of Science (Data Science) (M.Sc.)
Two Year Integrated Programme
Four Semesters

Course: Data Science

Course Structure

Syllabus for MSc Semester I & II

Under Choice Based Credit System (CBCS)
To be implemented from Academic Year 2025-2026
Progressively

Preamble

The Master of Science (M.Sc.) in Data Science is a graduate-level program offered by the autonomous college, focusing on the study and application of data analysis, data mining, machine learning, and statistical modelling techniques to extract valuable insights and knowledge from large and complex datasets. This program is meticulously designed to equip students with the necessary skills and knowledge to tackle the challenges and opportunities in the rapidly evolving field of data science.

The curriculum of the M.Sc. in Data Science program typically includes a blend of core courses and elective courses. Core courses provide a solid foundation in fundamental concepts and techniques of data science, such as programming languages (e.g., Python, R), essential technologies for data science, data visualization, statistical analysis, natural language processing, soft computing, time series analysis and forecasting, and deep neural networks. Elective courses allow students to specialize in specific areas of interest, such as healthcare analytics, sports analytics, legal analytics, human resource and retail marketing analytics, block chain technologies, image and video analytics, and social media analytics.

Additionally, the M.Sc. in Data Science program incorporates On Job Training (OJT) and hands-on projects. This intensive OJT training, equivalent to a full course, provides invaluable exposure to real-world scenarios within IT or IT-related organizations. By applying their theoretical knowledge in practical settings, students gain first-hand experience and develop the necessary skills to thrive in the professional world. The projects help students gain practical experience in applying their knowledge to solve complex data challenges and provide valuable insights to organizations. The inclusion of a Research Methodology Course helps students develop a strong research attitude, enabling them to contribute meaningfully to the advancement of Data Science.

Career prospects for graduates of the M.Sc. in Data Science program are highly promising, as data science skills are in high demand across various industries. Graduates can pursue careers as data scientists, data analysts, and machine learning engineers, data engineers, or data consultants in sectors such as finance, healthcare, e-commerce, marketing, and technology. They may also find opportunities in research institutions or pursue further studies at the doctoral level.

Overall, the M.Sc. in Data Science program offers a comprehensive and rigorous education in the field of data science, equipping students with the skills and knowledge necessary to succeed in a data-driven world. The curriculum's continuous refinement has been made possible through the valuable inputs, suggestions, and observations of colleagues, experts from premier institutions, and industry professionals. The autonomous college extends its gratitude to all those who have directly or indirectly contributed to the development of this program.

Name of Programme	**Masters of Data Science**
Level	PG
No of Semesters	04
Year of Implementation	**2025-26**
Programme Specific Outcomes (PSO)	At the end of the Programme, Learner will be able to 1. Develop a solid foundation in fundamental concepts, theories, and methodologies of Data Science. 2. Offer opportunities for specialization in a chosen area of Data Science. 3. Foster a research-oriented mindset and contribute to the advancement of Data Science. 4. Prepare learners for lifelong learning, adapting to emerging technologies and industry requirements. 5. Inculcate professional attitudes, leadership qualities and social responsibility. 6. Equip students with industry-relevant skills and experiences for successful careers. 7. Enhance critical thinking and innovative problem solving abilities.
Relevance of PSOs to the local, regional, national, and global developmental needs	1. **Provide In-depth Knowledge:** The program aims to provide students with a comprehensive understanding of the key concepts, theories, and methodologies in Data Science. It covers a range of topics including Data Analysis and Visualization, Statistical Methods for Data Science, data visualization, and SPARK Technologies, Soft Computing ,etc. enabling students to develop a deep knowledge base in these areas. 2. **Develop Programming Skills:** The program aims to equip students with strong programming skills by providing hands-on experience with different tools and technologies. Students will gain proficiency in designing front-end and back-end solutions, enhancing their ability to develop robust and scalable applications. 3. **Foster Problem-solving Abilities:** The program aims to enhance students' problem-solving abilities by training them to approach real-world data challenges critically and creatively. Students will learn to identify problems, design appropriate data analysis strategies, and develop innovative solutions using their knowledge of Computer Science. 4. **Encourage Collaboration and Teamwork:** The program aims to foster collaboration and teamwork skills among students, recognizing that computer science projects often require interdisciplinary collaboration. Students will learn to effectively communicate, collaborate, and contribute as part of a team, preparing them for collaborative work environments. 5. **Foster Industry Relevance:** The program aims to stay up-to-date with industry trends and technologies to ensure graduates are well-prepared for the demands of the Data Science job market. Through industry partnerships and

	internships, students will have the opportunity to gain practical experience and stay in touch with the latest advancements in the field. 6. **Professional Development:** The program aims to prepare students for successful careers in the field of Data Science. In addition to technical skills, students will develop professional skills such as teamwork, project management, and leadership. The program may also provide networking opportunities, internships, or collaborations with industry partners to enhance students' industry readiness and employability. 7. **Cultivate Research Skills**: The program aims to cultivate research skills among students by providing a Research Methodology Course and encouraging participation in research projects. Students will learn to conduct literature reviews, design experiments, analyze data, and present their findings, fostering a research oriented mindset and contributing to the advancement of Data Science.

Master of Science (Data Science) Programme
Under Choice Based Credit System (CBCS)
Course Structure (Autonomous)
Department of Data Science
(To be implemented from Academic Year 2025-26)

Course Code	Semester I	Credits	Course Code	Semester II	Credits
	Major Mandatory			***Major Mandatory***	
25_PSDSM101	Essential Technologies for Data Science (T)	**04**	25_PSDSM201	Soft Computing (T)	**04**
25_PSDSM102	Essential Technologies for Data Science Practical (P)	**02**	25_PSDSM202	Soft Computing Practical (P)	**02**
25_PSDSM103	Data Analysis and Visualization (T)	**04**	25_PSDSM203	Time Series Analysis and Forecasting (T)	**04**
25_PSDSM104	Data Analysis and Visualization Practical (P)	**02**	25_PSDSM204	Time Series Analysis and Forecasting Practical (P)	**02**
25_PSDSM105	Statistical Methods for Data Science (T)	**02**	25_PSDSM205	Ethical Issues in Data Science (T)	**02**
	Major Electives (Any 1)			***Major Electives (Any 1)***	
25_PSDSE106	SPARK Technologies (T)	**02**	25_PSDSE206	Human Resource Analytics (T)	**02**
25_PSDSE107	SPARK Technologies Practical (P)	**02**	25_PSDSE207	Human Resource Analytics Practical (P)	**02**
OR			OR		
25_PSDSE108	Retail Marketing Analytics (T)	**02**	25_PSDSE208	Public Health care Analytics (T)	**02**
25_PSDSE109	Retail Marketing Analytics Practical (P)	**02**	25_PSDSE209	Public Health care Analytics Practical (P)	**02**
OR			OR		
25_PSDSE110	Sports Data Analytics (T)	**02**	25_PSDSE210	Social Media Analytics (T)	**02**
25_PSDSE111	Sports Data Analytics Practical (P)	**02**	25_PSDSE211	Social Media Analytics Practical (P)	**02**
	Research Methodology (RM)			***On Job Training(OJT)***	
25_PSDSR112	Research Methodology (T)	**04**	25_PSDSJ212	On Job Training (P)	**04**
Total Credits		**22**	**Total Credits**		**22**

Nomenclature of the Course	Essential technologies for Data science
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSM101
No. of Credits	4
Nature	Theory
Type	Major Mandatory
Highlight revision specific to employability/ entrepreneurship/ skill development	The course *Essential Technologies for Data Science* is revised to enhance employability, entrepreneurship, and skill development. It now includes hands-on training in industry-relevant tools like Python, R, Git, and Tableau, along with cloud platforms such as Google Colab and AWS. Students will build portfolios through real-world projects and GitHub repositories. Entrepreneurship is encouraged via modules on data-driven business models, MVP development, and startup case studies. Communication and visualization skills are strengthened through dashboard creation and data storytelling. These updates ensure learners gain practical, job-ready skills while fostering innovation and self-reliance in the evolving data science industry.

Course Outcomes:

The learner will be able to :

CO1: To understand the core concepts of programming before starting to write new programs. Students should be able to develop logic for Problem Solving.

CO2: To manage data by relying on data structures such as strings, arrays, files, lists, and Dictionaries. Exemplify the various levels of decision making on a program and implement a mix of loops, functions, and control flow to extract information from a data structure.

CO3: Students should be able to learn different programming techniques and tools related to data science.

C**O4 :** Students should able to implement basic Data analysis techniques on datasets

Syllabus:

Unit No.	Unit Title	Sub Titles (Learning Points)
1	**Introduction to Data science and Python**	Introduction to Data Science, data science life cycle, Applications, and advantages of Python over other programming languages, What is Python? Why Should I Learn Python? Installing Python How to execute Python program Writing your first program. Basic programming elements of Python-variables and constants, identifiers, Typecasting or Type Conversion in Python, indentation, comments, rules of writing identifiers, primitive data types, writing command line programs in python, Operators in Python: Arithmetic operators,

		relational operators, Logical operators, Membership operators, Taking user input.
2	**Data structures and control flow**	Collection data structures in Python- List, tuples, dictionary, sets and strings, Control flow- Sequential, Branching or Conditional, Iteration or Repetition, Modular or Subroutines Conditional and iteration statements: if elif else statements, loops, for loop and while loops User defined functions in Python- No Value Pass and No Return, Value Pass and No Return, Value Pass and Return, Function with default arguments, Function with variable arguments, Higher order functions, list comprehension
3	**Statistics for Data Analysts**	Permutations and combinations, probability, Descriptive statistics (mean, median, mode), point estimation, quartiles and boxplot, methods of dispersion, random variables and probability distribution Measures of shape- skewness, kurtosis, outlier detection, transformation (log, square root) Introduction to NumPy, creating NumPy arrays, indexing and slicing, vectorization, Boolean indexing, transformation, inferential statistics using NumPy.
4	**Data wrangling using Pandas.**	Data wrangling using Pandas - Creating Series, Creating Data frame from dictionary, attributes, and method description of a data frame. Drop columns, add columns, add rows, iloc , loc, indexing and slicing data frames, selection with condition, group by summary operation, sorting Operations. Introduction to R IDE- components of R IDE, Basic data types in R, Data structures in R, data coercion, importing files, visualisation using ggplot2. Basic visualisation using matplotlib- Components of a chart, line chart, scatter chart, pie chart, sub plots

Prescribed Text/s (If any):

1. Data Analysis with Pandas and Python by Boris Paskhaver, Manning Publications. Available at: https://www.perlego.com/book/2881120/pandas-in-action-pd
2. Practical Statistics for Data Scientists: 50 Essential Concepts by Peter Bruce, Andrew.
Bruce, Peter Gedeck, O'Reilly Media, 2017 ISBN-10: 1491952962 ISBN-13: 978-1491952962
3. Foundations of Statistics for Data Scientists With R and Python By Alan Agresti, Maria Kateri, CRC Press Taylor and Francis group.

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**Introduction to Data science and Python**	Classroom Teaching and ICT	15
2	**Data structures and control flow**	Classroom Teaching and ICT	15
3	**Statistics for Data Analysts**	Classroom Teaching and ICT	15
4	**Data wrangling using Pandas.**	Classroom Teaching and ICT	15

Nomenclature of the Course	Essential technologies for Data Science Practical
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSM102
No. of Credits	2
Nature	Practical
Type	Major Mandatory
Highlight revision specific to employability/ entrepreneurship/ skill development	The practical component of Essential Technologies for Data Science has been revised to focus on employability, entrepreneurship, and skill development. Students will work on real-time datasets using tools like Python, Jupiter Notebooks to develop hands-on expertise. These enhancements ensure students gain technical proficiency, critical thinking, and innovation skills essential for modern data careers.

Course Outcomes:

The learner will be able to :

CO1: To write basic data handling programs in Python.

CO2: To use Python data structures and able to use conditional and iterative control flow.

CO3: To demonstrate descriptive, diagnostic, and inferential statistics using, Python, R or Excel (use Data analysis tool pack in Excel or Data analyser tool in Microsoft office 365)

CO4: Perform basic data wrangling using R or Pandas

Note: Following practical's will be implemented using any version of Python.

Syllabus:

Sr. No.	Aim of the Practical
1	Write a Python program to accept inputs from users and perform arithmetic operations.
2	Write a program to demonstrate relational and logical operators in Python.
3	Write a Python program to demonstrate usage of loops. Use both for and while loops to distinguish between them. [e.g., Reversing the digits of a number without converting to String]
4	Demonstrate the use of data structures list, sets, and dictionaries.
5	Import a dataset and perform univariate analysis on the numeric columns to analyse the shape of the data. Write inference of the output. [Python or R or Excel]
6	Demonstrate Hypothesis testing, and ANOVA using a dataset [use Python, R or Excel]
7	Demonstrate correlation analysis. Use heatmap for visualization. Write inferences.
8	Import a csv or Excel dataset and demonstrate data wrangling, view shape, dimension, column names of the dataset, ways to select data using column number, column names, simple and compound conditional selection, update and modify dataset.
9	Demonstrate group by summary operations and sorting techniques.

10	Perform univariate, bivariate and multivariate analysis using visualization techniques in Python, R or Excel
11	Build a simple linear regression model predicting a dependent variable based on an independent variable.
12	Import a dataset and demonstrate: identifying missing values, Handling missing values (dropna, fillna with mean/median/mode), Detecting and handling outliers using IQR and Z-Score methods.

Teaching Plan:

Practicals	**Unit Title**	**Teaching Methods**	**No. of Lectures**
All	All Practical's	Lab Session with ICT	60

Nomenclature of the Course	Data Analysis and Visualization
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSM103
No. of Credits	4
Nature	Theory
Type	Major Mandatory
Highlight revision specific to employability/ entrepreneurship/ skill development	This course builds strong employability and entrepreneurship skills by training learners in data analysis, visualization, and storytelling using industry-standard tools like Tableau and Power BI, preparing them for roles such as Data Analyst, Business Intelligence Developer, and Data Consultant.

Course Outcomes:
The learner will be able to : **CO1:** Understand fundamental data analysis concepts, sources, and processes in data science. **CO2:** Understand data wrangling techniques, including data gathering, cleaning, transformation, and integration from various sources. **CO3:** Develop skills in data visualization using Tableau, including creating, customizing, and publishing interactive dashboards. **CO4:** Demonstrate data storytelling using Power BI, including report creation, visualization, interactivity, and publishing.

Syllabus:		
Unit No.	**Unit Title**	**Sub Titles (Learning Points)**
1	**Introduction to Data Analysis**	Data Analysis, Exploratory Data Analysis, Data Science Process, Responsibilities of a Data Analyst, Data Analytics vs. Data Analysis. Understanding Different Types of File Formats, Sources of Data, and Languages for Data Professionals. Overview of Data Repositories, Data Marts, Data Lakes, ETL, and Data Pipelines, Foundations of Big Data, Identifying Data for Analysis.
2	**Data Wrangling**	How to gather and Import Data, Data Loading, Storage of Data, Reading and Writing Data in Text Format, Web Scraping, Binary Data Formats, interacting with Web APIs, Interacting with Databases. Data Wrangling, Hierarchical Indexing, Combining and Merging Data Sets Reshaping and Pivoting Tools for Data Wrangling. Data Cleaning and Preparation, Handling Missing Data, DataTransformation, String Manipulation.
3	**Data Visualization**	Introduction to Data Visualization and Dashboarding Software, Visualization Tools, Getting started with Tableau Desktop, Connecting to the dataset. Creating charts, Creating common visualizations, Filtering and sorting data, Adding Titles, Labels, and descriptions. Publish your work to Tableau Cloud, Interactivity with text and visual tooltips, Interactivity with actions, Assembling dashboards from multiple charts.
4	**Story Telling**	Introduction to Power BI, Understanding Desktop, Understanding Power BI Report Designer, Report Canvas, Report Pages Creation, Report Visuals. Fields and UI Options, Experimenting Visual Interactions, Advantages, Reports with Multiple Pages and Advantages, Pages with Multiple Visualizations. Publish Options and Report Verification in Cloud, Adding Report Titles, Report Format Options, Introduction to data storytelling, Creating a data story.

Reference Books:

1. Python for Data Analysis: Data Wrangling with Pandas, NumPy and IPython by McKinney, W., 2nd edition. O‟Reilly Media, 2017
2. Doing Data Science: Straight Talk from the Frontline by O‟Neil, C., & Schutt, R, O‟Reilly Media, 2013
3. The Big Book of Dashboards by Steve Wexler, Jeffrey Shaffer, Andy Cotgreave, John Wiley & Sons, 2017
4. Practical Tableau by Ryan Sleeper, O‟Reilly Media, 2018
5. Power BI. Book-1, Business Intelligence Clinic: Create and Learn by Roger F Silva, 2018

6. Introducing Microsoft Power BI by Alberto Ferrari and Marco Russo, Microsoft Press, Washington, 2016			
Teaching Plan:			
Unit No.	**Unit Title**	**Teaching Methods**	**No. of Lectures**
1	**Introduction to Data Analysis**	Classroom Teaching and ICT	15
2	**Data Wrangling**	Classroom Teaching and ICT	15
3	**Data Visualization**	Classroom Teaching and ICT	15
4	**Story Telling**	Classroom Teaching and ICT	15

Nomenclature of the Course	Data Analysis and Visualization Practical
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSM104
No. of Credits	2
Nature	Practical
Type	Major Mandatory
Highlight revision specific to employability/ entrepreneurship/ skill development	This practical course builds hands-on skills in data handling, cleaning, visualization, forecasting, and real-time dashboard creation using Tableau and Power BI, preparing learners for job roles and entrepreneurial projects in data analytics and business intelligence.

Course Outcomes:

The learner will be able to :

CO1: Handle missing data and manage data wrangling and manipulation
CO2: Create data visualization and report making using various software tools
CO3: Demonstrate the visualizations and make interpretations
CO4: Create a data story using various software tools.

Note: The following set of practical's could be implemented in Python/ R/ Power BI/ Tableau or any other suitable software.

Syllabus:

Sr. No.	Aim of the Practical
1	Implement Data Loading, Storage and File Formats. Read data and store them in text format.
2	Implement the code to interact with Web APIs and to perform web scrapping
3	Demonstrate Data Cleaning and Preparation.
4	Implement Data wrangling on a data set.
5	Demonstrate the handling of missing data and string manipulation.
6	Create common charts with title, labels and descriptions using Tableau.
7	Perform sorting and filtering using tableau, create visualizations and publish it on Tableau Cloud.
8	Perform data visualization using Power BI.
9	Create reports using Power BI.
10	Create a data story in Tableau or power BI

11	Load time-series data and visualize trends over time using line charts and Apply basic forecasting techniques (e.g., ARIMA or Prophet) in Power BI or Tableau.
12	Build a real-time data dashboard in Power BI or Tableau using live data sources (e.g., API connections). Monitor and visualize real-time data streams (e.g., stock prices, social media mentions).

Teaching Plan:

Practical's	Unit Title	Teaching Methods	No. of Lectures
All	All Practical's	Lab Session with ICT	60

Nomenclature of the Course	Statistical Methods for Data Science
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSM105
No. of Credits	2
Nature	Theory
Type	Major Mandatory
Highlight revision specific to Employability/ entrepreneurship/ skill development	This course will develop key employability skills in quantitative reasoning, data literacy and analytical thinking enabling students to interpret and apply statistical methods like t-tests, ANOVA, and regression. These skills are valuable across sectors such as research, healthcare, finance, and technology.

Course Outcomes:

The learner will be able to :

CO1: Evaluate sample size and power, validate models, and interpret results, analyse a variety of datasets using statistical inference and significance testing.

CO2: Understand and apply advanced statistical techniques, including various regression models to analyze complex datasets.

Syllabus:

Unit No.	Unit Title	Subtitles (Learning Points)
1.	**Introduction to Applied Statistics**	The Nature of Statistics and Inference: What is "Big Data?", Statistical Modelling, Statistical Significance Testing and Error Rates, Simple Example of Inference Using a Coin, Point Estimates and Confidence Intervals, Variable Types, Sample Size, Statistical Power and Statistical Significance, The Verdict on Significance Testing, Training versus Test Data, Mean, Correlations, Counts Drawing Inferences: Computing z and Related Scores, Statistical Tests, Plotting Normal Distributions, Correlation Coefficients. Tests of Mean Differences: t-Tests for One Sample, Two Sample t-Test, Paired-Samples t- Test Categorical Data: Binomial Test, Categorical Data Having More Than Two Possibilities. Power Analysis and Sample Size Estimation: Power for t-Tests, Power for One-Way ANOVA, Power for Correlations. Analysis of Variance: Fixed Effects, Random Effects, Mixed Models, Introducing the Analysis of Variance (ANOVA), Performing the ANOVA, Random Effects ANOVA and Mixed Models, One-Way Random Effects ANOVA

2	**Multivariate Techniques**	Regression: Simple and Multiple Linear Regression, Hierarchical Regression, Logistic Regression and the Generalized Linear Model, Predicting Probabilities, Training Error Rate Versus Test Error Rate. Multivariate Analysis of Variance (MANOVA)and Discriminant Analysis: Multivariate Tests of Significance, Example of MANOVA, Outliers, Homogeneity of Covariance Matrices, Linear Discriminant Function Analysis, Theory of Discriminant Analysis, Predicting Group Membership, Visualizing Separation. Principal Component Analysis: Principal Component Analysis Versus Factor Analysis, Properties of Principal Components, Component Scores, Exploratory Factor Analysis, Common Factor Analysis Model, Factor Analysis Versus Principal Component Analysis on the Same. Cluster Analysis: k-Means Cluster Analysis, Minimizing Criteria, Example of k-Means Clustering, Hierarchical Cluster Analysis

Prescribed Text/s (If any)

1. Fundamentals of Mathematical Statistics; Tenth Edition.Sultan Chand &Sons. (2000), Gupta S. C., Kapoor V. K
2. Applied Multivariate Statistical Analysis, Prentice-Hall, NewJersey, 2002, Johnson, R.A., Wichern, D.W.

Other Learning Resources recommended:

1. Applied Regression Analysis (John Wiley), Third Edition. Draper, N. R. and Smith, H. (1998)
1. Statistics using R, second edition. Narosa Publishing House, New Delhi.Purohit, S. G. Gore, S.D. and Deshmukh, S.R. (2015).
2. Applied Non-Parametric Statistics, First edition Boston-Houghton Mifflin Company, Daniel W. W.

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**Introduction to Applied Statistics**	Classroom Teaching and ICT	15
2	**Multivariate Techniques**	Classroom Teaching and ICT	15

Nomenclature of the Course	SPARK Technologies
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSE106
No. of Credits	2
Nature	Theory
Type	Major Elective
Highlight revision specific to employability/ entrepreneurship/ skill development	Mastering Apache Spark enhances employability by equipping learners with in-demand big data processing skills crucial for roles in data engineering, analytics, and AI development. Proficiency in RDDs, Spark applications, and performance tuning builds practical, job-ready skills.

Course Outcomes:

The learner will be able to :

CO1: Understand the concept of SPARK Technologies and its implementation with the RDD concept.

CO2: Understand the implementation of SPARK SQL, GraphX, and Performance Tuning.

Syllabus:

Unit No.	Unit Title	Subtitles (Learning Points)
1	**Introduction to SPARK Technologies**	Components of the Spark unified stack, Features of Spark, Spark Web UI, An introduction to RDDs - Resilient Distributed Datasets, Launching and using Spark's Scala and Python shell, Spark Context, Spark Ecosystem, In-Memory data – Spark, Creating, Loading and Saving RDD, Transformations in RDD, Actions in RDD, Key-Value Pair RDD, Map Reduce and Pair RDD operations RDD Partitions.
2	**Implementation of SPARK Technologies**	Spark Applications vs. Spark Shell, Creating Spark Context, Building a Spark Application, Spark and Hadoop Integration-HDFS, Handling Sequence File, Spark RDD,RDD Lineage, RDD Persistence Overview, Distributed Persistence. Spark Streaming, ML library for Spark, Working with Statistics, SPARK SQL, GraphX, Performance Tuning.

Prescribed Text/s (If any):

1. Learning Spark: Lightning-Fast Data Analytics 2nd Edition, by Jules S. Damji,Brooke Wenig , Tathagata Das, Denny Lee, O‟Reilly , 2020
2. Apache Spark Machine Learning Blueprints 1st Edition, Kindle Edition by Alex Liu, Packt Publishing, 2016
3. Apache Spark 2.x Cookbook: Cloud-ready recipes for analytics and data science 2nd Edition, by Rishi Yadav, Packt Publishing, 2017

Teaching Plan:			
Unit No.	**Unit Title**	**Teaching Methods**	**No. of Lectures**
1	**Introduction to SPARK Technologies**	Classroom Teaching and ICT	15
2	**Implementation of SPARK Technologies**	Classroom Teaching and ICT	15

Nomenclature of the Course	SPARK Technologies Practical
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSE107
No. of Credits	2
Nature	Practical
Type	Major Electives
Highlight revision specific to employability/ entrepreneurship/ skill development	This practical course builds industry-ready skills in big data processing, real-time analytics, and machine learning using Apache Spark. It enhances employability and entrepreneurship by enabling learners to design scalable data solutions like recommendation engines and streaming dashboards.

Course Outcomes:

The learner will be able to :
CO1: Understand the concept of SPARK Technologies and its implementation
CO2: Understand the concept of RDD
CO3: Understand the implementation of SPARK SQL, GraphX Performance Tuning.
CO4: Use Spark ML to Produce Movie Recommendations.

Note:The following practical's will be implemented using Python.

Syllabus:

Sr. No.	Aim of the Practical
1	Demonstrate installation of Apache Spark.
2	Introduction of Spark Basics and RDD interface.
3	Filtering RDDs, and the Minimum Temperature by Location Example.
4	Demonstrate use of Counting Word Occurrences using flatmap().
5	Executing SQL commands and SQL-style functions on a DataFrame.
6	Implement Total Spent by Customer with DataFrames.
7	Demonstrate use of Broadcast Variables to Display Movie Names Instead of ID Numbers.
8	Create Similar Movies from One Million Rating.
9	Demonstrate use of Spark ML to Produce Movie Recommendations.
10	Demonstrate use of Windows with Structured Streaming to Track Most-Viewed URLs (Spark Streaming).
11	Implement a code for Handling Missing Data and Data Cleaning in Spark
12	Demonstrate how to join two large DataFrames (inner join, left join) in Spark

Teaching Plan:

Practicals	Unit Title	Teaching Methods	No. of Lectures
All	All Practical's	Lab Session with ICT	60

Nomenclature of the Course	Retail Marketing Analytics
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSE108
No. of Credits	2
Nature	Theory
Type	Major Electives
Highlight revision specific to employability/ entrepreneurship/ skill development	This course equips learners with essential marketing analytics and data visualization skills, enhancing their ability to make data-driven decisions. It supports employability and entrepreneurship by teaching practical tools like R, hypothesis testing, and CLV prediction—crucial for roles in digital marketing, product strategy, and customer analytics.

Course Outcomes:

The learner will be able to :

CO1: Understand the importance of marketing analytics for forward looking and systematic allocation of marketing resources.

CO2: Understand how to use marketing analytics to develop predictive marketing dashboard for organization.

Syllabus:

Unit No.	Unit Title	Sub Titles (Learning Points)
1	**Introduction to Marketing Analytics and Exploratory Data analytics using R**	Marketing Analytics Overview, how analytics can assist marketing decision-making, the framework of marketing optimization .Tabulate and Summarize data, cleaned data, simple histogram plot, use histogram and boxplot to inform data distribution. Visualize data, elements of data visualization, histogram, scatter plot, line plot, bar charts, line fits with the ggplot() function
2	**Marketing Campaigns - Experiment Design, Customer Lifetime Value (CLV) and Cohort Analysis**	Design and Conduct Experiments , Design experiments, examples , randomization/sample splitting , conduct experiments , Assess Experiment Outcome Using Hypothesis Testing , hypothesis testing for experiment outcomes , terminologies for hypothesis testing, how does hypothesis testing work , power calculation , conduct hypothesis testing in R, Calculate and Predict CLV , calculate CLV, typical frameworks in predicting CLV, using linear regression and logistic regression to predict CLV , CLV Analysis and Cohort Analysis Introduction to Experiment.

Prescribed Text/s (If any):

1. Hands-on Data Science for Marketing by Yoon Hyup Hwang, Packt Publishing, 2019
2. Retail Analytics: The Secret Weapon by Emmett Cox,1st edition , Weily , 2011
3. Cutting Edge Marketing Analytics: Real World Cases and Data Sets for Hands on

Learning by Venkatesan Rajkumar, Farris Paul and Ronald Wilcox, Pearson FT Press, 2014 4. Marketing Analytics: A Practical Guide to Real Marketing Science by Grigsby Mike,Kogan Page, 2015			
Teaching Plan:			
Unit No.	**Unit Title**	**Teaching Methods**	**No. of Lectures**
1	**Introduction to Marketing Analytics and Exploratory Data analytics using R**	Classroom Teaching and ICT	15
2	**Marketing Campaigns - Experiment Design, Customer Lifetime Value (CLV) and Cohort Analysis**	Classroom Teaching and ICT	15

Nomenclature of the Course	Retail Marketing Analytics Practical
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSE109
No. of Credits	2
Nature	Practical
Type	Major Elective
Highlight revision specific to employability/ entrepreneurship/ skill development	This practical course builds job-ready skills in marketing data analysis, visualization, and experimentation using R. It enhances employability by training learners in real-world techniques like CLV prediction, A/B testing, and data-driven campaign design—key for roles in digital marketing and business analytics.

Course Outcomes:

The learner will be able to :

CO1: Learn working and analyzing marketing data.
CO2: Develop predictive marketing dashboard for organization.
CO3: Understand the concept of hypothesis testing and its role in assessing experiment. outcomes.
CO4: Understand the concept Time Series Forecasting.

Note:The following practical's will be implemented using RStudio and MS Excel

Syllabus:

Sr. No.	Aim of the Practical
1	Learn how to tabulate and summarize marketing data using R. ● Clean and preprocess the marketing data. ● Generate a simple histogram plot to visualize data distribution. ● Use tabulation and summary functions to gain insights from the data. ● Interpret the findings and discuss the implications for marketing analysis.
2	Gain proficiency in visualizing marketing data using R. ● Understand the key elements of data visualization. ● Create various visualizations such as histograms, scatter plots, line plots, and bar charts using the ggplot() function in R. ● Apply appropriate visualization techniques to effectively communicate marketing insights.
3	Design and conduct experiments for marketing campaigns. ● Learn about experimental design and its application in marketing. ● Design experiments using examples from marketing scenarios. ● Implement randomization and sample splitting techniques. ● Conduct the experiments and collect relevant data for analysis.
4	Understand the concept of hypothesis testing and its role in assessing experiment outcomes. ● Explore the purpose of hypothesis testing in analyzing experiment results. ● Familiarize with key terminologies related to hypothesis testing. ● Learn the process of hypothesis testing and power calculation. ● Conduct hypothesis testing using R to evaluate experiment outcomes.
5	Calculate and predict Customer Lifetime Value (CLV).

	● Calculate CLV using different approaches and frameworks. ● Explore predictive modeling techniques such as linear regression and logistic regression for CLV prediction. ● Assess the accuracy and reliability of CLV predictions.
6	Apply CLV analysis and cohort analysis in marketing analytics. ● Analyze CLV data and identify patterns and trends. ● Perform cohort analysis to segment customers based on their behavior or characteristics. ● Interpret the results of CLV analysis and cohort analysis to derive actionable insights for marketing strategies.
7	Extract data from social media platforms and perform analysis to gain insights into customer behavior and preferences. ● Utilize Python libraries like Beautiful Soup and requests to scrape data from social media platforms. ● Clean and preprocess the scraped data. ● Analyze the data to identify trends, sentiment analysis, or customer engagement metrics. ● Visualize the findings using appropriate charts or graphs.
8	Analyze customer purchasing patterns and build a recommender system based on market basket analysis. ● Use transactional data to identify frequently occurring item sets using association rule mining algorithms. ● Calculate support, confidence, and lift for the identified item sets. ● Build a recommendation engine using collaborative filtering techniques. ● Evaluate the performance of the recommender system and make recommendations based on customer preferences.
9	Segment customers based on their recency, frequency, and monetary value (RFM) to better target marketing efforts. ● Analyze customer transaction data to calculate RFM scores. ● Segment customers into different groups using clustering algorithms such as k-means or hierarchical clustering. ● Perform descriptive analysis on each customer segment to understand their characteristics. ● Develop targeted marketing strategies for each segment based on their RFM profiles.
10	Conduct A/B testing to evaluate the impact of different marketing strategies and make data-driven decisions. ● Design and implement A/B tests for marketing campaigns using randomized assignments. ● Collect relevant data and perform statistical analysis to compare the performance of different strategies. ● Calculate key metrics such as conversion rates, clickthrough rates, or revenue. ● Interpret the results and provide recommendations for optimizing marketing campaigns based on the findings.
11	Time Series Forecasting for Marketing Trends. ● Use historical marketing/sales data. ● Apply time series forecasting methods (ARIMA, Exponential Smoothing). ● Forecast future trends and demand. ● Visualize forecasts and interpret business implications.
12	Churn Analysis and Retention Modeling ● Identify churners based on customer transaction data. ● Build churn prediction models using classification algorithms (Logistic Regression , Decision Trees).

	● Analyze important factors contributing to churn. ● Recommend strategies to improve customer retention.

Teaching Plan:

Practicals	Unit Title	Teaching Methods	No. of Lectures
All	All Practical's	Lab Session with ICT	60

Nomenclature of the Course	Sports Data Analytics
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSE110
No. of Credits	2
Nature	Theory
Type	Major Elective
Highlight revision specific to employability/ entrepreneurship/ skill development	This course builds employability by developing practical skills in data-driven decision-making for sports, using tools like machine learning, predictive modeling, and visualization. It fosters entrepreneurship by enabling the creation of innovative solutions in sports performance, injury prevention, and fan engagement—key areas in the growing sports tech industry.

Course Outcomes:

The learner will be able to :

CO1: Understand the role and importance of data analytics in sports

CO2: Develop skills in collecting, cleaning, and managing sports data

Syllabus:

Unit No.	Unit Title	Sub Titles (Learning Points)
1	**Fundamentals of Sports Data Analytics**	Overview of sports data analytics , Evolution of analytics in sports , Importance and applications of sports data analytics , Data Collection and Preprocessing, Sources of sports data , Data collection methods , Data cleaning and preprocessing techniques ,Database management for sports analytics, Data Visualization for Sports Analytics, Principles of data visualization ,Visualization tools and libraries ,Creating effective visualizations for sports data ,Interactive dashboards for sports analytics , Statistical Analysis in Sports , Descriptive statistics for sports data , Hypothesis testing in sports analytics , Regression analysis in sports, Analysis of variance (ANOVA) in sport.
2	**Advanced Techniques in Sports Data Analytics**	Introduction to predictive modelling , Feature selection and engineering for sports data , Linear regression models for sports prediction , Classification models for sports outcomes , Machine Learning in Sports Analytics ,Overview of machine learning algorithms , Decision trees and random forests in sports analytics, Support vector machines for sports prediction , Neural networks and deep learning in sports analytics, Advanced Topics in Sports Data Analytics , Sports performance analysis , Player tracking and motion analytics , Sports marketing and fan engagement analytics , Sports injury prediction and prevention , Sports Business Analytics , Revenue generation and marketing in sports , Fan engagement and customer analytics.

Prescribed Text/s (If any):

1. Sports Analytics: A Guide for Coaches, Managers, and Other Decision Makers by Benjamin C. Alamar, Columbia university press, 2013
2. Sports Analytics and Data Science: Winning the Game with Methods and Models by Thomas Miller, 1st edition, Pearson FT Press, 2015
3. Sports Analytics: Analysis, Visualisation and Decision Making in Sports Performance by Daniel Memmert, Tim McGarry, and Tony Reilly, 2018
4. Cricket Analytics: Analytics and Data Science in Cricket by Tapan Bagchi and S. Raghunathan
5. Machine Learning using Python by Manaranjan Pradhan and U. Dinesh Kumar, Weily , 2020

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**Fundamentals of Sports Data Analytics**	Classroom Teaching and ICT	15
2	**Advanced Techniques in Sports Data Analytics**	Classroom Teaching and ICT	15

Nomenclature of the Course	Sports Data Analytics Practical
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSE111
No. of Credits	2
Nature	Practical
Type	Major Elective
Highlight revision specific to employability/ entrepreneurship/ skill development	This practical course enhances employability by building hands-on skills in sports data analysis using real cricket datasets, focusing on performance metrics and statistical comparisons. It supports entrepreneurship by enabling data-driven insights for talent scouting, team strategy, and fan engagement platforms within the sports analytics domain.

Course Outcomes:

The learner will be able to :

CO1: Understand performance metrics using histograms, box plots, or scatter plots.
CO2: Understand Data Manipulation and Exploratory Data Analysis
CO3: Understand Player Performance/ Team Performance Analysis
CO4: Develop a regression model and classification mode

Note:The following practical's will be implemented using n Python/ R/ Power BI/ Tableau or any other suitable software

Syllabus:

Sr. No.	Aim of the Practical
1	Exploratory Data Analysis ● Perform exploratory data analysis on a cricket dataset, analyzing variables such as number of matches, runs, not outs, wickets, etc. ● Visualize the distribution of player performance metrics using histograms, box plots, or scatter plots. Investigate the relationship between player age and performance metrics using correlation analysis.
2	Batting Performance Analysis ● Analyze batting performance in a Cricket dataset, calculating metrics such as batting average, strike rate, and runs scored by players. ● Identify top-performing batsmen based on performance metrics and compare their performance against different opponents or in specific conditions.
3	Bowling Performance Analysis ● Analyze bowling performance in the Cricket dataset, calculating metrics such as bowling average, economy rate, and wickets taken by players. ● Identify top-performing bowlers based on performance metrics and analyze their performance against different teams or in various match situations
4	Performance Comparison ● Compare the scoring averages of top-performing batsman in different seasons. ● Analyze the runs scoring (strike rate of Batting) of players from various teams in a specific league
5	Player Position Analysis ● Calculate position-specific performance metrics and compare players within each

	position.
6	Injury Analysis ● Investigate the relationship between player injuries and their subsequent performance using historical injury and performance data. ● Identify patterns and trends in the data to determine the impact of injuries on player performance and team success.
7	Team Analysis ● Analyze the impact of toss on a team's overall scoring and winning percentage. ● Study the relationship between batting averages of players and their team's win-loss record
8	Sports Revenue Analysis ● Analyze revenue generation in sports organizations by examining factors such as ticket sales, merchandise sales, and sponsorship deals. ● Identify key drivers of revenue and provide recommendations for maximizing financial performance.
9	Predictive Modeling ● Build a regression model to predict the number of runs scored by players based on their historical performance data. ● Develop a classification model to predict the outcome of match based on team‟s statistics.
10	Fantasy League Points Prediction ● Create a model to predict fantasy cricket points based on player stats (runs, wickets, catches, strike rate, economy rate). ● Help optimize fantasy team selections using predictive analytics.
11	Sentiment Analysis on Sports Commentary or Tweets ● Scrape live tweets or commentary about matches. ● Perform sentiment analysis to measure fan sentiment toward players/teams. ● Visualize sentiments using pie charts, bar graphs, or word clouds.
12	Visualization and Reporting: (Mini-Project) ● Prepare a comprehensive report summarizing the findings of the analysis and providing actionable insights for sports teams or organizations

Teaching Plan:

Practical's	Unit Title	Teaching Methods	No. of Lectures
All	All Practical's	Lab Session with ICT	60

Nomenclature of the Course	Research Methodology
Class	M.Sc. Data Science
Semester	I
Course Code	25_PSDSR112
No. of Credits	4
Nature	Theory
Type	Research Methodology
Highlight revision specific to employability/ entrepreneurship/ skill development	This course aims at employability by building a strong foundation in research principles, which fosters critical thinking and problem-solving skills essential across various industries. The paper sharpens analytical thinking and project planning skills, data driven roles in research, business, and technology sectors.

Course Outcomes:

The learner will be able to :

CO1: Understand fundamentals of research methods.
CO2: Get knowledge of scientific report writing.
CO3: Learn design and measurement concepts of research.
CO4: Know data collection and analysis tools.

Syllabus:

Unit No.	Unit Title	Sub Titles (Learning Points)
1	**Fundamentals of Research Methods**	Definition of research, Role and objectives of research, importance of research, Applications and types of research, Creativity and innovation, Critical thinking, Research process and steps in it, Collecting and reviewing the literature, Conceptualization and Formulation of: research problem, identifying variables, constructing hypothesis and Synopsis, Interpretation of results and discussion. Selecting and defining a research problem, Need for research design, Features of a good research design, Different research designs, Digital sources: Web sources, E-journals, Journal access, Citation Index, Impact factor, H- index, E-consortium, UGC info net, eBooks, Internet discussion groups and communities, Blogs, preprint servers, Search engines, Scirus, Google Scholar, Scopus

2	**Documentation, scientific writing and Academic Integrity**	Documentation and scientific writing: Structure and Components of Research Report, Types of Report: research papers, thesis, Research Project Reports, Pictures and Graphs, writing a review of paper, Preparation of manuscript for Publication of Research paper, citation styles, Results and Conclusions, Bibliography, Presenting a paper in scientific seminar, publications of scientific work, Thesis Writing Research and Academic Integrity: Intellectual property rights, Plagiarism, Copyright issues, Ethics in research.
3	**Research Design and Measurement Concepts and Literature Searching**	Scales of measurements, Nominal, Ordinal, Internal and ratio scales, Errors in measurements, Validity and Reliability in measurement, Scale Construction Techniques. Optimization of Algorithm: Introduction, Constraints, Critical points, condition of local minima, First order methods- Gradient Descent, Momentum, Nesterov momentum, Adagrad, RMSProp, Adadelta, Adam
4	**Research Testing**	Statistical Analysis for Data Sciences: Inferential statistics for hypothesis testing, Parametric and non parametric test, Regression analysis and correlation, Statistical software tools Recent trends in research: Machine Learning Predictions: Machine learning concepts for predictive modeling, model selection and evaluation Big Data Analytics: Understanding big data and its challenges, Tools and frameworks for big data processing (e.g., Hadoop, Spark), Analyzing and extracting insights from large datasets

Prescribed Text/s (If any):

A] Books and Textbooks:

1. Research Methodology, Methods and Techniques‖, Kothari C. R., Second revised edition, New Age International Publication, 2004.
2. Research Methodology, Saravanavel P., Kitab Mahal, Sixteenth edition, 2007.
3. Research Methodology, a step-by-step guide for beginners, Ranjit Kumar, Second edition, Pearson education Australia, 2005.
4. Research Methods for Business Students, Mark Saunders, Philip Lewis, Adrain Thornhiu, Seventh edition, Pearson Education ltd, 2016

B] Academic Journals and Publications:

1. Thesis & Assignment Writing, J Anderson, B.H.Dursten & M.Poole, Wiley Eastern, 1977
2. A Hand Book of Methodology of Research , P. Rajammal and P. Devadoss, R. M. M. Vidya Press,1976.
3. The Craft of Scientific Writing by Michael Alley, (Springer).

4. Research Methodology, R. Panneerselvam, PHI, New Delhi 2005
5. Research Methodology- A step by step Guide for Beginners, Kumar Ranjit, Second Edition, Pearson Education, 2005.
6. How to write and publish by Robert A. Day and Barbara Gastel, Cambridge University Press.
7. Research Methodology and Statistical techniques, S. Gupta, Deep and Deep Publications (P) Ltd. New Delhi, India, 2005.
8. Research Methodology, R. Kothari, New Age International, New Delhi, India, 2008.
9. Standard /Reputed Journal authors' instructions.

C] Web Resourecs:

1. www.sciencedirect.com for journal references,
2. www.aip.org and www.aps.org for reference styles.
3. www.nature.com, www.sciencemag.org,
4. www.springer.com, www.pnas.org, www.tandf.co.uk,
5. www.opticsinfobase.org for research updates.

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**Fundamentals of Research Methods**	Classroom Teaching and ICT	15
2	**Research Design and Measurement Concepts and Literature Searching**	Classroom Teaching and ICT	15
3	**Documentation, scientific writing and Academic Integrity**	Classroom Teaching and ICT	15
4	**Research Testing**	Classroom Teaching and ICT	15

Nomenclature of the Course	Soft Computing
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSM201
No. of Credits	4
Nature	Theory
Type	*Major Mandatory*
Highlight revision specific to employability/ entrepreneurship/ skill development	The revised Soft Computing for Data Science course boosts employability and entrepreneurship through hands-on learning with fuzzy logic, neural networks, and genetic algorithms using Python tools. Real-world case studies, collaborative projects, and innovative applications in finance, healthcare, and logistics enhance critical thinking, problem-solving, and technical communication skills.

Course Outcomes:

The learner will be able to :

CO1 :-Understand the concepts and techniques of soft computing.

CO2 :- Apply fuzzy logic, neural networks, and evolutionary computing to real-world data science problems.

CO3: Explore hybrid soft computing techniques for optimization and prediction tasks.

CO4: Develop the ability to design intelligent systems that solve complex, real-life problems.

Syllabus:

Unit No.	Unit Title	Sub Titles (Learning Points)
1	**Introduction to Soft Computing**	Soft Computing :Hard vs. Soft Computing, Characteristics of Soft Computing, Components of Soft Computing: Fuzzy Logic, Neural Networks, Genetic Algorithms, Swarm Intelligence, Probabilistic Reasoning, Applications in Data Science Fuzzy Logic and Systems: Introduction to Fuzzy Sets, Operations on Fuzzy Sets, Fuzzy Relations and Fuzzy Inference Systems, Mamdani and Sugeno Fuzzy Models, Fuzzification and Defuzzification, Applications in Control Systems and Data Classification.
2	**Artificial Neural Networks (ANN)**	Basic Concepts of Neurons: Biological Neurons and Artificial Neurons, ANN Architectures: Feedforward, Feedback, Recurrent Networks, Learning Rules: Hebbian, Delta, Perceptron Learning Rule, Multilayer Perceptrons (MLP) and Backpropagation, Introduction to Deep Learning Concepts, Applications: Classification, Regression, Feature Extraction.

3	**Evolutionary Computation**	Genetic Algorithms : Basics of Genetic Algorithms (GAs), Representation, Selection, Crossover, Mutation Operators, Fitness Functions ,Hybridization with Neural Networks and Fuzzy Systems. Applications in Optimization Problems, Holland classifier systems, genetic programming, advantages and limitations and applications of genetic algorithm.
4	**Intelligent Systems**	Swarm Intelligence Techniques: Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Applications in Clustering and Optimization Hybrid Soft Computing Systems:- Neuro-Fuzzy Systems, Genetic-Fuzzy Systems, Hybrid Models in Real-World Applications (Predictive Analytics, Decision Support Systems) Recent Trends in Soft Computing: Neuro-symbolic AI, Explainable AI (XAI) and Soft Computing, Quantum-Inspired Soft Computing

Prescribed Text/s (If any):

1. Artificial Intelligence and Soft Computing by Anandita Das Bhattacharya, SPD, 3rd edition 2018 .
2. Principles of Soft computing by S.N.Sivanandam S.N.Deepa , Wiley, 3rd edition, 2019
3. Neuro-Fuzzy and Soft Computing by J.S.R.Jang, C.T.Sun and E.Mizutani, Prentice Hall of India, 1st edition, 2004.
4. Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis & Applications by S.Rajasekaran, G. A. Vijayalakshami , Prentice Hall of India, 1st edition, 2004
5. Fuzzy Logic with Engineering Applications, Timothy J.Ross, McGrawHill 1st edition, 1997 .
6. Genetic Algorithms: Search, Optimization and Machine Learning by Davis E.Goldberg, Addison Wesley, 1st edition, 1989.
7. Introduction to AI and Expert System by Dan W. Patterson, Prentice Hall of India, 2nd edition, 2009.

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**Introduction to Soft Computing**	Classroom Teaching and ICT	15
2	**Artificial Neural Networks (ANN)**	Classroom Teaching and ICT	15
3	**Evolutionary Computation**	Classroom Teaching and ICT	15
4	**Intelligent Systems**	Classroom Teaching and ICT	15

Nomenclature of the Course	Soft Computing Practical
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSM202
No. of Credits	2
Nature	Practical
Type	*Major Mandatory*
Highlight revision specific to employability/ entrepreneurship/ skill development	These practical tasks are revised to boost employability, entrepreneurship, and skill development by focusing on real-world problem-solving using fuzzy logic, neural networks, and evolutionary algorithms. Students gain hands-on experience with industry-relevant tools and datasets, build AI-based solutions, and optimize models for applications like stock prediction and credit risk. This fosters critical thinking, innovation, and readiness for data-driven careers or startups.

Course Outcomes:

The learner will be able to :

CO1: Apply fuzzy logic principles and fuzzy inference systems to model uncertainty and approximate reasoning in real-world data science problems**.**

CO2: Design and implement various types of artificial neural networks (ANNs) for solving classification, regression, and prediction tasks using real datasets..

CO3: Develop and evaluate optimization algorithms like Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) for solving complex computational problems.

CO4: Integrate and apply hybrid soft computing techniques (e.g., Neuro-Fuzzy Systems, GA-ANN models) for advanced data analytics and decision-making applications.

Note: The following Practical's will be implemented using Python.
Link :- https://www.python.org/ftp/python/3.13.3/python-3.13.3-amd64.exe
Packages: -scikit-fuzzy, TensorFlow/Keras or PyTorch, DEAP, pyswarm,anfis
Package Versions : Any

Syllabus:

Sr. No.	Aim of the Practical
1	Implement fuzzy set operations like union, intersection, complement.
2	Design a fuzzy controller for a temperature control system.
3	Map continuous input to fuzzy sets and perform centroid defuzzification.
4	Train a single-layer perceptron for classifying linearly separable data (like AND, OR gate).
5	Implement a three-layer network to classify the XOR problem.
6	Train an MLP on the Iris Dataset or MNIST digits classification.
7	Find the maximum/minimum of a mathematical function using a genetic algorithm.
8	Use Genetic Algorithm to select optimal features from the Titanic dataset for survival prediction.
9	Cluster the Iris dataset using PSO and compare it with K-Means.
10	Implement Ant Colony Optimization to find the shortest path in TSP.
11	Apply an Adaptive Neuro-Fuzzy Inference System (ANFIS) for stock price prediction.

12	Use Genetic Algorithm to optimize the weights of an ANN for credit risk prediction.

Teaching Plan:

Practicals	Unit Title	Teaching Methods	No. of Lectures
All	All Practical's	Lab Session with ICT	60

Nomenclature of the Course	Time Series Analysis and Forecasting
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSM203
No. of Credits	4
Nature	Theory
Type	*Major Mandatory*
Highlight revision specific to employability/ entrepreneurship/ skill development	Enhances employability through hands-on skills in time series forecasting and data interpretation. Supports entrepreneurship by enabling data-driven planning and strategic decision-making capabilities. Fosters skill development in modern analytical tools, aligning with industry and startup ecosystem demands.

Course Outcomes:

The learner will be able to :

CO1 : Forecast the trend pattern exhibited by the given data by using various methods.
CO2 : Run and interpret time series models and regression models for time series.
CO3 : Use the Box-Jenkins approach to model and forecast time series data empirically.
CO4 : Analyze and estimate the cyclic components using special processes.

Syllabus:

Unit No.	Unit Title	Sub Titles (Learning Points)
1	**Introduction to Trends**	Introduction to times series data, application of time series from various fields, Components of a time series, Decomposition of time series. Trend: Estimation of trend by free hand curve method, method of semi averages, fitting a various mathematical curve, and growth curves.
2	**Trend and Seasonal component**	Method of moving averages. Detrending. Effect of elimination of trend on other components of the time series. Seasonal Component: Estimation of seasonal component by Method of simple averages, Ratio to Trend, Ratio to moving average and Link relatives.
3	**Forecasting variate component method**	Stationary Time series: Weak stationary, autocorrelation function and correlogram of moving average. Forecasting: Exponential smoothing methods, Short term forecasting methods: Brown's discounted regression, Box-Jenkins Method.

4	**Cyclic Component Deseasonalization**	Deseasonalization. Cyclic Component: Harmonic Analysis. Some Special Processes: Moving-average (MA) process and Autoregressive (AR) process of orders one and two, Estimation of the parameters of AR (1) and AR (2) – Yule-Walker equations.

Prescribed Text/s (If any):

1. Time Series , Kendall, M. , 2nd Edition , 1976.
2. The Analysis of Time Series – An Introduction , Chatfield, C. , 6th Edition , 1980.
3. Applied Statistics , Mukhopadhyay, P. , 2nd Edition (Revised Reprint) , 2011.
4. Time Series Analysis and Its Applications: With R Examples , Shumway, R. H., and Stoffer, D. S. , 2nd Edition , 2006.
5. Time Series Analysis: Forecasting and Control , Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. , (Not mentioned, typically 3rd/4th Edition), 1994.
6. Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS , Yaffee, R., and McGee, M. , (Edition not mentioned) , 2000.

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**Introduction to Trend**	Classroom Teaching and ICT	15
2	**Trend and Seasonal component**	Classroom Teaching and ICT	15
3	**Forecasting Variate component method**	Classroom Teaching and ICT	15
4	**Cyclic Component Deseasonalization**	Classroom Teaching and ICT	15

Nomenclature of the Course	Time Series Analysis and Forecasting Practical
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSM204
No. of Credits	2
Nature	Practical
Type	*Major Mandatory*
Highlight revision specific to employability/ entrepreneurship/ skill development	Enhances skill development through practical application of forecasting methods and machine learning tools like Prophet and LSTM. Promotes employability and entrepreneurship by training students in data-driven decision-making using real-world time series data.

Course Outcomes:

The learner will be able to :

CO1: Fit various growth curves and trend models, and measure seasonal indices in time series data.
CO2: Apply different forecasting methods to analyze and predict future trends effectively.
CO3: Calculate and interpret the variance of the random (irregular) component in time series models.
CO4: Visualize time series components and fitted models using appropriate plotting techniques to aid in data-driven decision making**.**

Note: The following set of practicals can be implemented using platforms such as Time Series Lab, Tableau,any statistical software, or any suitable programming platform.

Syllabus:

Sr. No.	Aim of the Practical
1	Fitting and plotting of modified exponential curves.
2	Fitting and plotting of Gompertz curve.
3	Fitting and plotting of logistic curves.
4	Fitting of trend by Moving Average Method.
5	Measurement of Seasonal indices Ratio-to-Trend method
6	Measurement of Seasonal indices Ratio-to-Moving Average method.
7	Measurement of seasonal indices Link Relative method.
8	Calculation of variance of random component by variate difference method
9	Forecasting by exponential smoothing.
10	Forecasting by short term forecasting methods.
11	Time Series Decomposition using STL and X-13ARIMA-SEATS
12	Machine Learning Forecasting with Prophet and LSTM

Teaching Plan:

Practicals	Unit Title	Teaching Methods	No. of Lectures
All	All Practical's	Lab Session with ICT	60

Nomenclature of the Course	Ethical Issues in Data Science
Class	M.Sc. I
Semester	II
Course Code	25_PSDSM205
No. of Credits	2
Nature	Theory
Type	Major Mandatory
Highlight revision specific to Employability/ entrepreneurship/ skill development	This course prepares students for careers in entrepreneurship and data science by teaching them to handle data responsibly and make ethical decisions. It helps them promote accountability and create socially responsible innovations in business.

Course Outcomes:

The learner will be able to :

CO1: Apply ethical principles to data science, focussing particular attention to accountability, transparency, governance, and bias in automated decision-making and data processing.

CO2: Examine ethical issues in data science, including research, privacy, AI, and visualisation to guarantee ethical data practices.

Syllabus:

Unit No.	Unit Title	Subtitles (Learning Points)
1	**Ethical foundation in Data Science**	Foundational Concepts: Ethics in data science, Introduction to ethical frameworks, principles, challenges and importance. Ethical considerations in data handling: Data collection methods, storage, sharing and its ethical considerations, Types of bias in data, Importance of transparency in data science. Ethics in Automated Systems and Governance: Ethical considerations in automated decision-making, Data governance frameworks and practices, Ensuring accountability in data science projects.
2	**Emerging Ethical issues in Data Science**	Ethical Dimensions of Data Science Techniques and Applications: Ethical Issues in data visualization, Ethical challenges in emerging technologies e.g., AI, IoT, biometrics, blockchain, Ethical Challenges in Data Science: Ethical challenges in data science research, Ethical considerations in collaborative data science environments, Ethical issues in using the internet, privacy and security in the context of data science.

Prescribed Text/s (If any):

1. Data Science Ethics, David Martens ISBN: 9780192847263 Oxford University Press 2023
2. Ethics of Big Data: Balancing Risk and Innovation" by Kord Davis and Doug Patterson O′Reilly 2012.

Other Learning Resources recommended:

1. Data Science Ethics Resources - Concepts, Techniques, and Cautionary Talesby David Martens-Oxford University Press-2022.
2. Data Science Association, Data Science Code of Professional Conduct.- https://www.datascienceassn.org/code-of-conduct.html

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**Ethical foundation in Data Science**	Classroom Teaching and ICT	15
2	**Emerging Ethical issues in Data Science**	Classroom Teaching and ICT	15

Nomenclature of the Course	Human Resource Analytics
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSE206
No. of Credits	2
Nature	Theory
Type	*Major Elective*
Highlight revision specific to Employability/ entrepreneurship/ skill development	HR Analytics empowers better talent acquisition, upskilling, and career pathing through data-driven insights, enhancing employability and workforce planning. Tools like HRIS and competency mapping support entrepreneurial decision-making and targeted skill development aligned with business goals.

Course Outcomes:

The learner will be able to :

CO1: Understand and evaluate the significance, concepts, and frameworks of HR analytics, measurement, and HRIS applications.

CO2: Evaluate the application of HR analytics across key HR systems and frameworks, along with global and local challenges in HR valuation and assessment.

Syllabus:

Unit No.	Unit Title	Sub Titles (Learning Points)
1	**HR Measurement**	Introduction to HR measurement:Need for HR Measurement, Significance and concept of HR Analytics, HR Analytics and business linkages, Prerequisites of HR Analytics, Models and frameworks of HR Analytics, Measuring intellectual capital, Research Gap,approaches and Metho in HR Accounting. HRIS for HR Analytics: What is Human Resource Information System, Role of HRIS in analytics, HRIS development and Implementation, the development process need analysis, systems design, structure and culture, HRIS Applications-making HRIS work.
2	**Analytics for HR subsystems**	HR Analytics for Staffing, Training & Development: Performance Management Systems, Career Planning Systems, Rewards and Compensation Management, Employee Relations Systems. Analytics for HR system: HR performance frameworks and measurement systems, Measuring HR Climate and People Management Capabilities, Competency Management Frameworks & Competency Mapping, Integration of competency-based HR System. Measuring HR Effectiveness, The HR Scorecard Trends and Future Challenges:Technology and changes in HR Analytics, Role of social media, Big Data and Predictive Analytics in HR, Assessing the effectiveness of HR

		Analytics, Post analysis steps, Review and monitoring, Issues in HR valuation and measurement

Prescribed Text/s (If any):

1. Ulrich, D. & Brockbank, W., The HR Value Proposition. Harvard Business School Press 2016 .
2. How to measure HRM by Jac Fitz-enz 2002 .
3. Predictive Analytics for Human Resources by Jac Fitz-enz, John Mattox II, Wiley 2014 .
4. Making Human Capital Analytics Work: Measuring the ROI of Human Capital Processes and Outcomes. By by Jack Phillips,Patricia Pulliam Phillips- 2014 .

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**HR Measurement**	Classroom Teaching and ICT	15
2	**Analytics for HR sub-systems**	Classroom Teaching and ICT	15

Nomenclature of the Course	Human Resource Analytics Practical
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSE207
No. of Credits	2
Nature	Practical
Type	*Major Elective*
Highlight revision specific to Employability/ entrepreneurship/ skill development	Monitoring and analysing HR metrics using tools like dashboards, statistical software, and HRIS enhances data literacy and enables evidence-based decision-making—key skills for employability in modern HR roles.

Course Outcomes:

The learner will be able to :
CO1: Analyse different data to monitor and analyse HR metrics.
CO2: Use and understand tools require for HR analytics
CO3: Identify the application and uses of HR analytics in various HR sub-systems.

Note: It is expected to R for all assignments and practical's throughout this course.
Download R from http://cran.r-project.org/
Download R Studio fromhttp://www.rstudio.com/products/rstudio/download/

Syllabus:

Sr. No.	Aim of the Practical
1	Analyze employee turnover rates and identify factors contributing to attrition ● Collect historical employee data, including tenure, performance ratings, salary, and job satisfaction. ● Calculate employee turnover rates for different departments and job roles. ● Conduct statistical analysis to identify correlations between turnover and variables such as salary, job satisfaction, and performance. ● Generate visualizations (e.g., charts, graphs) to present the findings and propose recommendations to reduce turnover.
2	Develop a user-friendly HRIS dashboard for monitoring and analysing HR metrics ● Identify key HR metrics to be displayed on the dashboard (e.g., headcount, recruitment pipeline, training hours). ● Design the layout and interface of the HRIS dashboard using appropriate programming languages and tools. ● Integrate data from various HR systems and databases to populate the dashboard in real-time. ● Implement interactive features, such as drill-down capabilities and data filters, to facilitate data exploration and analysis
3	Analyse training effectiveness and identify skill gaps in the organization ● Collect training data, including participant demographics, training modules, pre/post-assessment scores, and performance metrics.

	● Perform statistical analysis to evaluate the impact of training on employee performance. ● Identify areas of improvement and recommend targeted training programs based on identified skill gaps. ● Develop a visualization or report summarizing the training needs analysis results.
4	Develop an HR scorecard to measure HR effectiveness and align HR strategies with organizational goals ● Identify key HR performance indicators aligned with the organization's strategic objectives. ● Collect relevant data for each HR indicator, such as employee satisfaction surveys, training investment data, and performance metrics. ● Calculate HR metrics and indicators, such as turnover rate, training ROI, and employee engagement index. ● Design a dashboard or report to present the HR scorecard and analyze trends over time.
5	Use predictive analytics to forecast employee attrition and develop retention strategies ● Gather historical HR data, including employee demographics, performance metrics, compensation, and employee exit data. ● Build a predictive model (e.g., logistic regression, decision tree) to predict employee attrition. ● Validate the model's accuracy and evaluate its performance using appropriate evaluation metrics. ● Generate actionable insights and recommendations to proactively address potential attrition risks.
6	Use predictive analytics to forecast employee attrition and develop retention strategies ● Gather historical HR data, including employee demographics, performance metrics, compensation, and employee exit data. ● Build a predictive model (e.g., logistic regression, decision tree) to predict employee attrition. ● Validate the model's accuracy and evaluate its performance using appropriate evaluation metrics. ● Generate actionable insights and recommendations to proactively address potential attrition risks.
7	Measure and analyze employee engagement levels within the organization ● Collect employee engagement survey data, including responses to survey questions related to job satisfaction, work environment, and organizational culture. ● Calculate engagement scores and identify key drivers of engagement. ● Conduct a sentiment analysis on employee feedback to understand areas of improvement. ● Present the findings and propose strategies to enhance employee engagement based on the analysis.
8	Develop a program to automate repetitive HR processes, such as leave management or performance appraisal ● Identify the HR process to be automated and define the required functionalities. ● Design and implement a web-based application or script to streamline the process using appropriate programming languages and frameworks. ● Integrate the application with relevant HR systems and databases to ensure data accuracy and consistency.

	● Test and validate the automated process, considering different scenarios and user inputs.
9	Analyse the effectiveness of the organization's performance management system and provide insights for improvement. ● Collect performance evaluation data, including performance ratings, goal chievement metrics, and feedback. ● Analyze the distribution of performance ratings across different departments or job roles. ● Identify trends and patterns in performance data and assess the fairness and consistency of the evaluation process. ● Propose recommendations for enhancing the performance management system based on the analysis results.
10	Analyse the organization's compensation structure and compare it to industry benchmarks. ● Gather salary data for different job roles and levels within the organization. ● Perform a salary analysis, including measures like average salary, salary distribution, and salary competitiveness. ● Conduct benchmarking by comparing the organization's salary data with industry standards or competitor data.
11	Recruitment Funnel Analysis and Optimization ● Collect data across hiring stages (applications, interviews, offers, hires). ● Calculate funnel conversion rates. ● Identify bottlenecks in recruitment processes. ● Recommend improvements to optimize hiring efficiency.
12	Succession Planning Analytics ● Identify high-potential employees (HiPos) from performance and competency data. ● Analyze leadership pipeline gaps. ● Build dashboards to track readiness of internal candidates for critical roles.

Teaching Plan:

Practicals	Unit Title	Teaching Methods	No. of Lectures
All	All Practical's	Lab Session with ICT	60

Nomenclature of the Course	Public Health Care Analytics
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSE208
No. of Credits	2
Nature	Theory
Type	*Major Elective*
Highlight revision specific to Employability/ entrepreneurship/ skill development	Healthcare data management and analysis open vast employability opportunities in roles such as data scientists, health informatics specialists, bioinformaticians, and clinical analysts, while enabling entrepreneurship in areas like EHR solutions, predictive health apps, and privacy-preserving analytics platforms.

Course Outcomes:

The learner will be able to :

CO1: Analyse the principles and challenges of health data management to enhance decision-making and improve healthcare outcomes.

CO2: Apply advanced data analytics techniques while addressing privacy and data governance challenges in healthcare.

Syllabus:

Unit No.	Unit Title	Sub Titles (Learning Points)
1	**Healthcare data management**	Healthcare data management:What is Health Data Management?, Benefits and challenges of health data management, how to store all that data,Electronic Health Records,Components of EHR,Coding Systems,Benefits of EHR,Barrier to Adopting EHR Challenges, Phenotyping Algorithms. Statistical analysis of healthcare data:Measures of Central Tendency and Dispersion, Confidence Limits and Hypothesis Testing, Statistical Tests for Categorical Data, T-Tests for Related and Unrelated Data, Analysis of Variance, Data Quality and Governance
2	**Healthcare data Analysis**	Healthcare data Analysis:Biomedical Image and Signal Analysis, Genomic Data Analysis for Personalized Medicine ,Natural Language Processing and Data Mining for Clinical Text, Mining the Biomedical- Social Media Analytics for Healthcare. Predictive Models for Integrating Clinical and Genomic Data, Privacy-Preserving Data, Publishing Methods in Healthcare . Mobile Imaging and Analytics for Biomedical Data, Data Analytics for Pharmaceutical Discoveries- Clinical Decision Support Systems .

Prescribed Text/s (If any):

1. Healthcare Business Intelligence + Website - A Guide to Empowering Successful Data Reporting and Analytics Hardcover by LB Madsen (Author),2012
2. Practical Text Analytics: Interpreting Text and Unstructured Data for Business Intelligence (Marketing Science) 1st Edition by Dr. Steven Struhl,2016

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**Healthcare data management**	Classroom Teaching and ICT	15
2	**Healthcare data Analysis**	Classroom Teaching and ICT	15

Nomenclature of the Course	Public Health Care Analytics Practical
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSE209
No. of Credits	2
Nature	Practical
Type	*Major Elective*
Highlight revision specific to Employability/ entrepreneurship/ skill development	Employability can be enhanced through hands-on experience with EHR platforms (like Epic or Cerner), statistical software (R, Python), and real-world healthcare datasets for tasks like patient outcome prediction, image analysis, or genomic pattern detection.

Course Outcomes:

The learner will be able to :
CO1: Understand, visualize and infer healthcare data.
CO2: Use linear regression and forecasting methods for predicting growth rates, expenditure or any other numeric variable related to the medical field.
CO3: Demonstrate descriptive, diagnostic, and inferential statistics using Python, R or Excel.
CO4: Perform predictive analysis using machine learning algorithms and deep learning.

Prerequisite: Basic computer skills and Basic understanding of elementary Math.
Note: The following practicals will be implemented using Python.

Syllabus:

Sr. No.	Aim of the Practical
1	EHR Data Modeling, Data Mining, and Exploratory Data Analysis (EDA) ● Model Electronic Health Record (EHR) data. ● Perform EDA using Tableau or Power BI. ● Visualize patient demographics, diagnoses, treatments, and outcomes.
2	Medical Expenditure Prediction ● Build regression models to predict patient medical expenses based on demographics, insurance, diagnoses, and treatment history. ● Evaluate model performance (e.g., RMSE, R^2).
3	Twitter Healthcare Data Extraction, Processing, and Sentiment Analysis ● Scrape tweets using healthcare-related keywords (e.g., COVID-19, vaccine). ● Preprocess tweets: clean, tokenize, remove noise. ● Perform sentiment analysis (positive/negative/neutral). ● Visualize findings using word clouds and sentiment graphs.
4	Using Classification Algorithms for Disease Prediction ● Apply algorithms (Logistic Regression, Random Forest, SVM) to predict diseases like diabetes, cancer, or heart disease. ● Evaluate performance using accuracy, precision, recall, and ROC curve.
5	Clinical Text Data Categorization and Feature Extraction ● Process clinical notes (text) using NLP. ● Perform feature extraction (TF-IDF, Bag-of-Words). ● Categorize text into diagnosis, treatment, prescription, etc.

6	Medical Image Classification (X-Rays) ● Classify medical images (e.g., pneumonia detection in chest X-rays) using CNNs (Convolutional Neural Networks). ● Evaluate using confusion matrices and accuracy scores.
7	Image Object Detection using Deep Learning ● Detect tumors or anomalies in medical imaging (MRI, CT scans). ● Implement object detection models (YOLO, Faster R-CNN).
8	Health Forecasting Using Time-Series Analysis ● Use time-series forecasting (ARIMA, Prophet) to predict health-related metrics (e.g., hospital admissions, disease outbreak rates).
9	Heart Sound Classification Using Signal Processing ● Analyze heart sound signals (phonocardiograms). ● Apply signal processing techniques (FFT, MFCC) and classify normal vs. abnormal heart sounds.
10	AI Conversational Chatbot for Primary Healthcare Diagnosis ● Develop a rule-based or ML-based chatbot. ● Enable it to diagnose common symptoms and suggest primary care actions.
11	Predictive Modeling of Hospital Readmission Rates ● Use patient data (age, diagnosis, treatment, length of stay) to predict likelihood of hospital readmission within 30 days. ● Apply classification models and evaluate their effectiveness.
12	Wearable Device Data Analysis for Health Monitoring ● Analyze data collected from wearable devices (e.g., Fitbit, Apple Watch). ● Monitor vital signs (heart rate, activity, sleep). ● Detect abnormal patterns and suggest health interventions.

Teaching Plan:

Practicals	**Unit Title**	**Teaching Methods**	**No. of Lectures**
All	All Practical's	Lab Session with ICT	60

Nomenclature of the Course	Social Media Analytics
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSE210
No. of Credits	2
Nature	Theory
Type	*Major Elective*
Highlight revision specific to Employability/ entrepreneurship/ skill development	Entrepreneurship and skill development can focus on creating data-driven social media consulting services, building recommendation engines, or developing location-based marketing solutions, supported by expertise in social network theory, strategy formulation, and risk management.

Course Outcomes:

The learner will be able to :

CO1: Analyze the core characteristics and types of social media to manage risks and evaluate online behavior through real-world case studies.

CO2: Evaluate the structure and features of social media network analytics.

Syllabus:

Unit No.	Unit Title	Sub Titles (Learning Points)
1	**Social Media Analytics: An Overview**	Core Characteristics of social media, Types of social media, social media landscape, Need for Social Media Analytics (SMA), Seven Layers of Social Media Analytics, Types of Social Media Analytics, Social Media Analytics Cycle, Location Analytics ,Sources of Location Data, Categories of Location Analytics.. Social Information Filtering: Social Sharing and filtering, Automated Recommendation systems, Traditional v/s social Recommendation Systems, understanding social media and Business Alignment, social media KPI, formulating a Social Media Strategy, Managing Social Media Risks
2	**Social Network Structure, Measures & Visualization**	Basics of Social Network Structure, Nodes, Edges & Tie Describing the Networks Measures,Degree Distribution, Density, Connectivity, Centralization, Tie Strength & Trust. Network Visualization, Graph Layout, Visualizing Network features, Scale Issues.Capturing Correlations: Triangles, Clustering, and Assortativity. **S**ocial Media Network Analytics ,Common Network Terms, Common Social Media Network Types, Types of Networks, Common Network Terminologies, Network Analytics Tools. Case Study:On any Social media.

Prescribed Text/s (If any):

1. Seven Layers of Social Media Analytics_ Mining Business Insights from Social Media Text, Actions, Networks, Hyperlinks, Apps, Search Engine, and Location Data, Gohar F. Khan, 2015
2. Analyzing the Social Web 1st Edition by Jennifer Golbeck, 2013
3. Mining the Social Web_ Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites, Matthew A Russell, O,,Reilly, 2019
4. Social Network Data Analytics, Springer, Charu Aggarwal (ed.),2011

Teaching Plan:

Unit No.	Unit Title	Teaching Methods	No. of Lectures
1	**Social Media Analytics: An Overview**	Classroom Teaching and ICT	15
2	**Social Network Structure, Measures & Visualization**	Classroom Teaching and ICT	15

Syllabus of Courses of **M.Sc. Data Science**
Programme at Semester II
with Effect from the Academic Year 2025-2026

Nomenclature of the Course	Social Media Analytics Practical
Class	M.Sc. Data Science
Semester	II
Course Code	25_PSDSE211
No. of Credits	2
Nature	Practical
Type	*Major Elective*
Highlight revision specific to Employability/ entrepreneurship/ skill development	Social media analytics and network analysis offer employability in roles such as social media strategist, data analyst, or network scientist, requiring hands-on skills in sentiment analysis, location analytics and tools like Gephi, Power BI, and Python libraries .

Course Outcomes:
The learner will be able to :
CO1: To understand the fundamental concepts of social media networks.
CO2: To Collect, monitor, store and track social media data
CO3: To analyse and visualize social media data
CO4: To design and develop social media analytics models.

Note: - The following set of practicals should be implemented in Scrape, python and python related tools.

Syllabus:

Sr. No.	Aim of the Practical
1	Study Various ● Social Media platforms (Facebook, twitter, YouTube etc) ● Social Media analytics tools (Facebook insights, google analytics netlyticetc) ● Social Media Analytics techniques and engagement metrics (page level, post level, member level) using Gephi Tool
2	Scrape an online Social Media Site for Data. Use python to scrape information from twitter. Exploratory Data Analysis and visualization of Social Media Data
3	Create sociograms for the persons-by-persons network and the community-by-community network for a given relevant problem. Create a one-mode network and two-node network for the same. Datasets: les-Misérables, Airlines, Internet Core Routers.
4	Develop Content (text, emoticons, image, audio, video) based social media analytics model for business. (e.g., Content Based Analysis: Topic, Issue, Trend, sentiment/opinion analysis, audio, video, image analytics)
5	Develop Structure based social media analytics model for any business. (e.g., Structure Based Models -community detection, influence analysis)
6	Develop a dashboard and reporting tool based on real time social media data Using Power BI
7	Use Google Visualization Charts to analyze social media data.
8	Analyze social media data Network Analysis with Orange Software

9	Use Graph Neural Networks on the datasets (Planetoid Cora Dataset)/ Jazz Musicians Network.
10	Analyze Twitter conversations to identify the most active and influential users using Machine Learning Algorithms with Gephi Tool.
11	Predict trending topics using supervised ML techniques (e.g., Naive Bayes, SVM).
12	Develop a mzachine learning model to rank influencers based on engagement, reach, and influence scores extracted from social media datasets.

Teaching Plan:

Practicals	Unit Title	Teaching Methods	No. of Lectures
All	All Practical's	Lab Session with ICT	60

Guidelines and Evaluation pattern for On Job Training **(100 Marks)**

Name of the Course	On Job Training
Course Code	25_PSDSJ212
Class	PG
Semester	II
No of Credits	4
Nature	Practical
Type	On Job Training
Relevance with Employability/ Entrepreneurship/ Skill development	On the job training provides learner with the opportunity to acquire hands on experience and practical skills required for specific job roles. It bridges the gap between theoretical knowledge and the practical requirements of the job. Learner can gain valuable insights into the industry practice, company culture, this experience makes them confident and competent candidate when applying for the position increasing the employability prospects. OJT is instrumental in skill development as it focuses on practical job specific competencies like technical skills, soft skills. Overall OJT enhances employability, foster entrepreneurship by providing valuable exposure in various field.

Aim:

The aim of on-job training during an internship is to provide students with practical experience in a professional setting, allowing them to apply classroom knowledge to real-world tasks. It helps students develop job-specific skills, improve their problem-solving abilities, and gain a deeper understanding of industry practices, preparing them for future career opportunities.

Introduction:

The main objective of inclusion of On Job Training and Field Project is to inculcate the ability to interpret particular aspects of the study in his/ her own words.

Guidelines for On Job Training:

On-the-Job Training: Students will be required to undertake a designated project or tasks in an organization or industry relevant to their field of study. The course aims to provide students with practical exposure and hands-on experience in a professional work environment related to their field of study.

Course Objectives:

By the end of the course, students should be able to:

1. Gain exposure to real-world insights and apply theoretical knowledge to practical situations
2. Enhance his/her skills regarding problem-solving, decision-making, and communication skills.
3. Understand organizational dynamics and work culture.
4. Build industry connections and networking opportunities.

Course Outcomes:

1. Apply theoretical knowledge and concepts acquired during the academic program to real-world work scenarios.
2. Develop practical skills and competencies necessary for successful professional engagement.
3. Demonstrate effective problem-solving, decision-making, and critical thinking abilities in a work environment.
4. Adapt to and navigate organizational dynamics and work culture in the chosen industry.
5. Prepare a comprehensive report documenting the training/project experience, findings, and recommendations.

Course Duration:

Minimum 1 months / 120 hours of On Job Training with an Organization/ Startup/ Charitable Organization/ Private firm/ Private Limited Company.

- The theme of the On Job Training should be based on any study area of the Major/Minor/VSC course
- Experience certificate, Evaluation report, working module and attendance report is Mandatory

Evaluation Scheme:

OJT Evaluation (100 Marks)				
Internal (40 Marks)		**External (60 Marks)**		
Mid –Term Report	**Presentation**	**End –Term Report**	**Working Module**	**Presentation**
30	10	30	20	10

Report Structure:

Mid-Term Report:

Mid-Term Report will be utilized for internal evaluation. The presentation of work done so far will be presented up to 10 minutes in the form of power point presentation which will include only introduction slide and working module/work done/skills earned so far. Mid-Term report will be submitted at the time of presentation. Please find the format of Mid-Term Report in Annexure A.

End-Term Report:

The students will be required to submit a comprehensive report at the end of the On-the-Job Training. A project report has to be brief in content and must include the following aspects:

a) Title Page:

Mentioning the title of the report, name of the student, program, institution, and the period of training/project. (Refer **Annexure B)**

b) Certificate of Completion:

1. A certificate issued by the organization or supervisor confirming the successful completion of the training/project. (Refer Annexure C)
2. A certificate issued by the institution confirming the successful completion of the training/project (duly signed by internal guide and HOD). (Refer Annexure D)

c) Professional Evaluation of intern:

Mentioning the behavior and punctuality of learner in the organization during On Job Training. (Refer Annexure E)

d) Declaration:

A statement by the student declaring that the report is their original work and acknowledging any assistance or references used. (Refer Annexure F)

e) Acknowledgments:

Recognizing individuals or organizations that provided support, guidance, or resources during the training/project. (Refer Annexure G)

f) Table of Contents:

Providing a clear outline of the report's sections and page numbers. (Refer Annexure H)

g) Introduction of the Company:

A Concise representation of company/ organization defining its scope, products / service. (Refer Annexure I)

h) Your Role in the Organization during the on-Job Training:

The key aspects handled, the department under which you were deployed and brief Summary report duly acknowledged by the reporting head. (Refer Annexure J)

Annexure A

(Proforma for Mid Term Report)

1. Name of the Trainee: _____
2. Academic Roll No.: _____
3. Position (If Any): _____
4. Name of the Company in which OJT is performed: _____

_____

5. Name of Guide from the Company: _____
6. No of Weeks/Hours for which mid-term Report is submitted: _____
7. Duration: From _____/____/_____to ____/____/_____
8. Submission date: _____

Signature

Signature Student **Internal Guide**

Internship Letter (If Given):

(Proforma for the certificate for internship in official letter head)

This is to certify that Mr/Ms_____ of_____College/Institution worked as an intern as part of his/her M.Sc. course in Computer Science of University of Mumbai. The particulars of internship are given below:

Internship starting date: _____

Internship ending date: _____

Actual number of days worked: _____

Tentative number of hours worked:_____Hours

Broad area of work: _____

A small description of work done by the intern during the period:

_____

_____

Signature: _____

Name: _____

Designation: _____

Contact number: _____

Email: _____

Tasks & Actions Taken (So-far)

1. Assigned Task:

2. Work performed so far (Min 100 words):

3. Any new skills learned(Min 100 words):

4. Action taken on assigned task(Min 100 words):
   Note: Mention in points. Do not mention module / working script

Annexure B

(Proforma for the Title Page of OJT End Term Report)

On-the-Job Training Report (OJT Report)

Student Information

Full Name:
Course/Program Name: MSc (Data Science)
Exam Seat Number:
Contact Information (Mobile Number):

OJT Details

Company Name: Company
Address:
OJT Department/Division:
OJT Supervisor Name and Position: Duration of
OJT:

Academic Information

Institution Name: Department/Faculty
Name: Instructor/Advisor Name: Course
Code:
Date of Submission (Month, Year):

Annexure C

(Proforma for the certificate for OJT End Term Report on official letter head of Company)

Certificate

This is to certify that Mr/Ms_____ of _____College/Institution worked as an intern as part of his/her M.Sc course in Data Science of R. P. Gogate College of arts an Science & R. V. Jogalekar College of Commerce(Autonomous), Ratnagiri . The particulars of internship are given below:

On Job Training starting date: _____

On Job Training ending date: _____

Actual number of days worked: _____

Tentative number of hours worked:_____Hours

Broad area of work: _____

A small description of work done by the Student during the period:

_____

_____

Signature: _____

Name: _____

Designation: _____

Contact number: _____

Email: _____

(Seal of the organization)

Annexure D

(Proforma for the certificate issued by institution)

R. E. Society's

R. P. Gogate College of Arts and Science and R. V. Jogalekar College of Commerce (Autonomous), Ratnagiri

Department of Computer Science

CERTIFICATE

This is to certify that Mr./Ms. _____ of M.Sc. (Sem II) class bearing examination seat no._____has satisfactorily completed On Job Training in _____ , as laid by the Board of Studies of Data Science for the year 20 -- in Data Science as Major. His/ Her bonafide work was completed under the guidance of Mr./Mrs.___.

Signature of Guide **Examiner** **Head**
Department Data Science

Date:
Place:

Annexure E

(Proforma for Professional Evaluation of Intern)

Professional Evaluation of Intern

Name of intern: _____

College/institution: _____

[Note: Give a score in the 1 to 5 scale by putting √ in the respective cells]

No	Particular	Excellent	Very Good	Good	Moderate	Satisfactory
1	Attendance & Punctuality					
2	Ability to work in a team					
3	Written and oral communication skills					
4	Problem solving skills					
5	Ability to grasp new concepts					
6	Technical skill in terms of technology,programming etc					
7	Ability to complete the task					
8	Quality of overall work done					

Comments:

_____

_____

Signature:_____

Name:_____

Designation:_____

Contact number: _____

Email: _____

(Seal of the organization)

Annexure F

(Proforma for the Declaration in OJT End Term Report)

DECLARATION BY STUDENT

I, **[Full Name]**, hereby declare that this On-the-Job Training (OJT) report titled **"[Title of the Report]"** is my own work and has been written and prepared in compliance with the guidelines and requirements set by **[Institution Name]**. All information and references from external sources have been properly cited and acknowledged.

This report has not been submitted for any other academic or professional purpose, and no part of it has been plagiarized or copied from other sources without appropriate citations. I understand the consequences of academic dishonesty, and I assure the authenticity of the content presented in this report.

I further declare that I have completed the OJT at **[Company Name]** during the period from **[Start Date]** to **[End Date]**, under the supervision of **[OJT Supervisor's Name]**, and the activities and experiences discussed in this report accurately reflect my involvement during the training.

Signed:
[Full Name of Student]
Date: [Month, Year]

Annexure G

(Proforma for the acknowledgments in OJT End Term Report)

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to all the individuals who contributed to the successful completion of my On-the-Job Training (OJT) and this report.

First and foremost, I would like to extend my deepest thanks to **[OJT Supervisor's Name]**, **[Position]** at **[Company Name]**, for his/her invaluable guidance, mentorship, and support throughout my training. His/her expertise and encouragement significantly enhanced my learning experience.

I would also like to express my appreciation to the **Head of the Department (HOD)**, **[HOD's Name]**, for his/her continuous support and for providing me with the opportunity to undergo this OJT program. His/her leadership and direction have played a significant role in shaping my academic and professional development.

My sincere thanks go to **[Vice Principal's Name]**, Vice Principal of Science Faculty for his/her support and for facilitating the training opportunities that enriched my learning experience. I am grateful for the opportunities provided under his/her leadership.

I would also like to express my heartfelt appreciation to **[Principal's Name]**, Principal of **[Institution Name]**, for his/her encouragement, and for ensuring that the institution maintains strong ties with industry partners, allowing students like me to gain practical experience in the field.

I am equally grateful to my academic advisor, **[Instructor/Advisor's Name]**, for providing me with the academic guidance and knowledge that prepared me for the challenges and opportunities during my OJT.

Special thanks are due to the entire team at **[Department/Division Name]** at **[Company Name]**, particularly **[Names of colleagues or team members, if applicable]**, for their collaboration and support, and for creating a positive and learning-focused environment during my training.

Lastly, I would like to thank my family and friends for their unwavering support and encouragement, which motivated me to give my best during this experience.

Sign [Your Full Name]

Date: [Month, Year]

Annexure H

(Proforma for the table of contents of OJT End Term Report)

TABLE OF CONTENTS

Annexure I

(Proforma for Company Introduction of OJT End Term Report)

Company Introduction

- **Company Name**: [Insert the full name of the company]
- **Industry**: [What industry does the company belong to? For example, manufacturing, marketing, IT, healthcare, etc.]
- **Location**: [Where is the company based? Include main offices or branches, if relevant.]
- **Year Established**: [When was the company founded?]
- **Founders**: [If applicable, mention who founded the company.]
- **Mission Statement (If Any):** [Include the company's mission or vision statement if available. This provides insight into the company's core values and objectives.]
- **Company Structure:** Explain the organizational structure of the company, particularly highlighting the departments or divisions where you worked or interacted.
- **Products/Services Offered**: [List the key products or services offered by the company. Focus on what was relevant to your training.]
- **Target Market**: [Who are the company's primary customers? For example, businesses, consumers, governments, etc.]
- **Clients**: [If relevant, mention some key clients or partners the company works with.]

Annexure J

(Proforma for role in organization during OJT of OJT End Term Report)

Role in the Organization during On-the-Job Training

Position/Title:

- Position/Title: [Your official position or title during the OJT, e.g., OJT Trainee, Marketing Assistant, IT Intern, etc.]
- Department/Division: [The department where you were assigned, e.g., Marketing, HR, IT, Production, etc.]
- Supervisor: [Name of your OJT supervisor, position, and department]

Primary Responsibilities and Tasks:

Provide a detailed description of the key tasks and responsibilities you were given during your OJT. Mention any specific projects or activities you worked on, and describe how these contributed to the organization's objectives.

- Task 1: [Description of the first key responsibility or task. Explain what you did, how you did it, and why it was important.]
- Task 2: [Description of the second responsibility, and so on.]

Skills and Knowledge Applied:

Explain the specific skills and knowledge you applied during your OJT, and how your academic background helped you in your role. This shows the connection between theory and practice.

- Skills Applied: [Mention the technical, professional, and soft skills you utilized. For example, communication skills, data analysis, project management, technical skills, etc.]
- Knowledge Applied: [Describe the theoretical knowledge you applied, such as principles from your coursework in marketing, engineering, business management, etc.]

Challenges and Problem-Solving:

Briefly mention any challenges or problems you faced in your role and how you addressed or overcame them. This demonstrates your ability to adapt and problem-solve in a professional environment.

Contribution to the Organization:

Highlight how your work and efforts contributed to the success of the organization during your OJT. This could include improvements in processes, successful projects, or other positive outcomes as a result of your involvement.

Bridge Course: Foundations of Data Science (2 Credits)

Course Objectives

1. Provide fundamental knowledge of programming (Python) and essential mathematical concepts.
2. Introduce basic data handling and statistical techniques required for data science.
3. Prepare students for advanced coursework in the M.Sc. Data Science program.

Prerequisites

- Basic understanding of mathematics
- No prior coding experience required

Course Structure (Total: 30 hours)

Module 1: Mathematics for Data Science (15 hours)

- Linear Algebra: Vectors, Matrices, and their operations
- Probability Basics: Conditional probability, Bayes' theorem
- Descriptive Statistics: Mean, Median, Mode, Variance, Standard Deviation
- Introduction to Calculus: Differentiation and Integration concepts in machine learning

Module 2: Introduction to Data Science Concepts (15 hours)

- What is Data Science? Overview of ML, AI, and Big Data
- Data Preprocessing: Handling missing data, normalizing, encoding categorical variables
- Introduction to Python and Jupyter Notebooks
- Data structures (Lists, Tuples, Dictionaries, and Sets)
- Functions and loops in Python
- Libraries: NumPy, Pandas, and Matplotlib
- Basic data manipulation and visualization

Evaluation Pattern:

- 50 Marks Online MCQ based Examination will be taken on LMS.
- Minimum 40% Marks required to qualify Bridge Course.
- List of Qualified Students will be displayed on Notice Board.
- In case of failure, ONE more valid attempt will be provided during the same academic year.

Evaluation Scheme

For Theory Course:

The allocation of marks for the Internal Assessment and Semester End Examinations are as follows-

Evaluation for the course will be of 100 marks conducted in a 60: 40 pattern. 60 marks will be

for semester end examination and 40 marks will be for internal evaluation.

A) Internal Assessment for 4 Credit Course: 40 % (40 Marks)

	Particulars	
	Mid-Term Class Test It should be conducted using any **learning management system** such as **Moodle**(Modular object-oriented dynamic learning environment) Question Paper Pattern for Periodical Class Test/ Online Examination Duration: 40 Minutes Match the Column / Fill in the Blanks / Multiple Choice Questions/ True or False / Answer in One or Two Lines (Concept based Questions) (1 Marks each)	
	Assignment/ Case study/ Presentations	
	Attendance	
	Behaviour	

B) Semester End Examination for 4 Credit Course: 60% (60 Marks)

Question Paper Pattern

Maximum Marks: 60

Questions to be set: 05

Duration: 2 Hours

Question no	Unit	Options	Marks
Q1	Unit 1	Any 2 out of 4	12
Q2	Unit 2	Any 2 out of 4	12
Q3	Unit 3	Any 2 out of 4	12
Q4	Unit 4	Any 2 out of 4	12
Q5	Unit 5	MCQ and Fill in the Blanks	12

C) Internal Assessment for 2 Credit Course: 40 % (20 Marks)

Sr. No.	Particulars	Marks
01	**Mid-Term Class Test** It should be conducted using any **learning management system** such as **Moodle**(Modular object-oriented dynamic learning environment) Question Paper Pattern for Periodical Class Test/ Online Examination Duration: 40 Minutes Match the Column / Fill in the Blanks / Multiple Choice Questions/ True or False / Answer in One or Two Lines (Concept based Questions) (1 Marks each)	**10**
02	**Assignment/ Case study/ Presentations**	**05**
03	**Attendance & Behavior**	**05**

D) Semester End Examination for 2 Credit Course: 60% (30 Marks)

Question Paper Pattern

Maximum Marks: 30

Questions to be set: 03

Duration: 1 Hour

Question No	Unit	Options	Marks
Q1	Unit 1	Any 2 out of 4	10
Q2	Unit 2	Any 2 out of 4	10
Q3	Unit 3	Any 2 out of 4	10

For Practical Course:

The allocation of marks for the Internal Assessment and Semester End Examinations are as follows-

Evaluation for the course will be of 100 marks conducted in a 60: 40 pattern. 60 marks will be for semester end examination and 40 marks will be for internal evaluation.

a. Internal Examination: 20 Marks

Sr. No	Criterion	Marks
1	Journals containing minimum 20 practical's which are timely completed with desired output	**10**
2	Attendance & Practical Performance	**10**

b. Semester End Examination: 30 Marks Duration: 2 hrs

Sr. No	Criterion	Marks
1	One Practical Question **OR** Combination of Practical Questions **OR** Combination of Practical Question and Theory Question	**25**
2	Viva	**05**

Standard of Passing

The learner to pass a course shall have to obtain a minimum of 40% marks in aggregate for each course where the course consists of Internal Assessment & Semester End Examination. The learner shall obtain minimum of 40% marks (i.e. 16 out of 40) in the Internal Assessment and 40% marks in Semester End Examination (i.e. 24 out of 60) separately, to pass the course and minimum of Letter Grade "P" in the project component, wherever applicable to pass a particular semester. A learner will be said to have passed the course if the learner passes the Internal Assessment & Semester End Examination together.

Performance Grading:

Letter Grades and Grade Points

Semester GPA/ Program CGPA Semester/Program	**% of Marks**	**Alpha-Sign / Letter Grade Result**
9.00-10.00	90.0 -100	0 (Outstanding)
8.00 ≤ 9.00	80.0 ≤ 90.0	A+ (Excellent)
7.00 ≤ 8.00	70.0 ≤ 80.0	A (Very Good)
6.00 ≤ 7.00	60.0 ≤ 70.0	B+ (Good)
5.50 ≤ 6.00	55.0 ≤ 60.0	B (Above Average)
5.00 ≤ 5.50	50.0 ≤ 55.0	C (Average)
4.00 ≤ 5.00	40.0 ≤ 50.0	P (Pass)
Below 4.00	Below 40	F (Fail)
Ab (Absent)	-	Absent

Date: 09/05/2025
Place:- Ratnagiri

The Chairperson
BoS of Data Science