SARDAR PATEL UNIVERSITY

Vallabh Vidyanagar

NAAC 'A' Grade (10-01-2023 To 09-01-2028)

NEP-2020 aligned Curriculum with effect from Academic Year 2026-27

Master of Science Applied Statistics and Predictive Data Science Semester-I

Course Type	Course Code	Course Title	Teaching-Learning Scheme L-P-T	Total Notional Hours	Course credits
DSC	**P2S01NCASP01**	Data Wrangling and Exploratory Analysis	4-0-0	120	04

- **Course Learning Outcomes (CLOs)**

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

CLO1: Demonstrate understanding of data types, data sources and data quality issues in real-world datasets.

CLO2: Apply appropriate data cleaning, transformation and integration techniques to prepare datasets for analysis.

CLO3: Analyze datasets using descriptive statistical methods and exploratory data analysis techniques.

CLO4: Evaluate patterns, trends and relationships in data using visualization and statistical measures.

CLO5: Develop and demonstrate data analysis workflows using software tools such as Python or R for effective decision-making.

Unit	Course Content	Learning Pedagogies*	CLO(s)
I	Data Wrangling Fundamentals • Data Sources and Collection: Types of data (structured, semi-structured, unstructured), sources of data, file formats, data acquisition methods, introduction to big data sources, data quality issues, ethical considerations in data collection. • Data Cleaning Techniques: Types of data errors (missing values, duplicates), handling missing data, outlier detection, data correction and validation, data type conversion, removing duplicates, noise filtering.	CL CBL	1 2
II	Data Transformation and Integration • Data Transformation: Data normalization (min-max, standardization), encoding categorical variables, feature engineering basics, aggregation and summarization, discretization (binning), scaling techniques, derived variables creation. • Data merging and joining, concatenation, pivot tables, reshaping data, handling multi-dimensional data, data reduction techniques, tidy data principles.	PBL MP	2 5
III	Exploratory Data Analysis • Descriptive Statistics: Measures of central tendency, measures of dispersion, skewness and kurtosis, frequency distributions, cross-tabulation, summary statistics using	CBL MP ICTEL	3 4 5

	software, interpretation. • Data Visualization Techniques: Principles of visualization, univariate analysis, bivariate analysis, multivariate visualization, graphical representation, visualization tools (Python/R).		
IV	Advanced Exploratory Data Analysis • Multivariate and Correlation Analysis: Correlation (Pearson, Spearman, Kendall), covariance, multivariate analysis, dimensionality reduction (PCA introduction), association analysis, feature selection basics. • Pattern Identification and Reporting: Trend analysis, anomaly detection, data storytelling, dashboard basics, report writing, communication of insights.	CBL ICTEL PBL	3 4 5

Learning Pedagogies/Methods

(a) Classroom Lecture (CL)
(b) Case-Based Learning (CBL)
(c) Micro-Projects (MP)
(d) Problem-Based Learning (PBL)
(e) ICT-Enabled Learning through Digital Resources/ Virtual Labs/Webinars (ICT EL)

- **Assessment Methodologies**

(A) Internal Assessment

a. Internal Formative assessment (25 Marks)

i. MCQ and Objective question Quiz: Students select or give the correct answer from the list of options four /none for a given questions.

ii. Assignment: A practical task or setup questions given to measure a student's knowledge, skill and ability in subject or field.

iii. Seminar: A student-led academic activity designed to evaluate student's ability to research, understand, and present a specific topic.

iv. Group Discussion: An interactive method to asses a student's personality and inter- personal skills in team setting. It typically involves 5 to 15 participants discussing a topic for 15 to 30 minutes in presence of an observer.

b. Internal Summative Assessment (25 Marks)

i. Mid-term tests
A test conducted roughly halfway through an academic term or semester.

ii. Laboratory performance
Students demonstrate their laboratory hands-on performance skills, knowledge and ability to apply concepts in subjects.

iii. Viva-voce
An oral examination conducted to evaluate a student's in-depth knowledge, understanding, and critical thinking regarding a subject or project.

Master of Science Applied Statistics and Predictive Data Science Semester-I

(B) Weightage of Learning Efforts for External Assessment

Unit	Aligned CLOs	Total Learning Hours	Approximate weightage (Marks) to Learning levels (BT)			Total Marks
			Remember (R)	Understanding (U)	Application/ Analyse & above (A)	
I	1, 2	30	4	5	3	12
II	2, 5	30	4	5	4	13
III	3, 4, 5	30	4	5	4	13
IV	3, 4, 5	30	3	5	4	12
		120	**15**	**20**	**15**	**50**

- **Assessment and Evaluation**

Sr. No.	Assessment/Evaluation	Component	Weightage (%)
1	Continuous Internal Evaluation	FA: Seminars, Assignments, Quizzes, Group Discussion	25%
		SA: Mid Term Test	25%
2	End-Semester Examination	Written Exam	50%

(C) CLOs – PLOs Matrix

CLO	PLO										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
CLO1	**3**	**-**	**-**	**2**	**-**	**2**	**-**	**-**	**-**	**-**	**-**
CLO2	**3**	**2**	**2**	**3**	**-**	**2**	**-**	**-**	**-**	**2**	**-**
CLO3	**3**	**2**	**3**	**3**	**2**	**-**	**2**	**-**	**-**	**2**	**-**
CLO4	**3**	**2**	**3**	**-**	**2**	**2**	**2**	**2**	**-**	**-**	**-**
CLO5	**3**	**3**	**3**	**3**	**-**	**2**	**2**	**3**	**2**	**2**	**2**

CLO – PLO correlation	Value
Strong	3
Moderate	2
Low	1
No correlation	-

- **Suggested Learning Materials Books:**

Sr.No.	Title	Author(s)	Edition/Year	Publisher
1	R for Data Science: Import, Tidy, Transform, Visualize, and Model Data	Hadley Wickham, Garrett Grolemund	1st Edition/2017	O'Reilly Media
2	Minimalist Data Wrangling with Python	Marek Gagolewski	1st Edition /2022	Self-Published
3	Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython	Wes McKinney	2nd Edition/2017	O'Reilly Media

Master of Science Applied Statistics and Predictive Data Science Semester-I

4	Exploratory Data Analysis with R	Roger D. Peng	1st Edition/2016	Leanpub
5	OpenIntro Statistics	David Diez, Christopher Barr, Mine Çetinkaya-Rundel	4th Edition/2019	OpenIntro Inc

- **Online Resources (Open Source)**

Sr. No.	Description of Resource(s)	Weblink
1	Prof. Shalabh, Indian Institute of Technology, Kanpur Exploratory Data Analysis for Data Science with R Software	https://onlinecourses.nptel.ac.in/noc26_cs07/preview

SARDAR PATEL UNIVERSITY

Vallabh Vidyanagar

NAAC 'A' Grade (10-01-2023 To 09-01-2028)

NEP-2020 aligned Curriculum with effect from Academic Year 2026-27

Master of Science Applied Statistics and Predictive Data Science Semester-I

Course Type	Course Code	Course Title	Teaching-Learning Scheme	Total Notional Hours	Course credits
			L-P-T		
DSC	**P2S01NCASP02**	Probability and Distribution Theory	4-0-0	120	04

- **Course Learning Outcomes (CLOs)**

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

CLO1: Demonstrate knowledge of axiomatic probability, conditional probability, Bayes' theorem, and advanced concepts like convergence in probability and distribution.

CLO2: Analyze discrete and continuous random variables, joint, marginal, and conditional distributions, and calculate moments, Cumulative Distribution Functions (CDF).and probability models using Characteristic Functions

CLO3: Apply the identified probability distributions (e.g., Cauchy, Beta, Gamma, Weibull, Pareto, Log-Normal) for data modelling.

CLO4: Evaluate convergence criteria and apply the Weak/Strong Laws of Large Numbers (LLN) and Central Limit Theorems (CLT) to analyse large sample behaviours.

CLO5: Derive properties of sampling distributions (t, F, Chi-square, min, max), multivariate distribution and understand the mathematical foundations of statistical inference (estimation theory) towards modelling complex data scenarios

Unit	Course Content	Learning Pedagogies*	CLO(s)
I	Measurement of Events • Probability and random Variables: Classes of Events and Data, Limit of sequence of events, Probability, Axiomatic approach, Random Variables and vectors, Decomposition of Distribution Function • Expectation and Moment's Inequalities: Expectation including Conditional expectation, Characteristic function, Holder's and Cauchy-Schwartz, Basic, Markov and Chebyshev, Jensen's and Liaponouv's inequalities, Multiplication Theorem of expectation	CL PBL ICT EL	1 2 4
II	Measurement of Sequence of Events • Modes of Convergences: Four modes of Convergences, Inversion theorem of characteristic function • Convergence Theorems: Weak and Strong Laws of large numbers, Central Limit Theorems for iid and independent sequence of random variables	CL PBL ICT EL	1 4
III	Statistical Modelling of Events and Data • Probability Distributions: Transformation of random variables, Generating Functions, Discrete & Continuous distributions (including advanced types), distribution function of sums and product of random variables • Sampling distributions: central and non-central t-, chi-square, F- distributions and their MGF and CGF	CL ICT EL	3 4

IV	Statistical Modelling of Joint Events • Joint Probability Distributions: Joint distributions, marginal distributions, conditional distributions, mixture distributions for bivariate normal distribution. • Multivariate Distributions: distribution of order statistics, multinomial and multivariate distribution	CL CBL MP	3 4 5

- **Learning Pedagogies/Methods**
  (a) Classroom Lecture (CL)
  (b) Case-Based Learning (CBL)
  (c) Micro-Projects (MP)
  (d) Problem-Based Learning (PBL)
  (e) ICT-Enabled Learning through Digital Resources/ Virtual Labs/Webinars (ICT EL)
- **Assessment Methodologies**

(A) Internal Assessment

a. Internal Formative assessment

i. MCQ and Objective question Quiz: Students select or give the correct answer from the list of options four /none for a given questions.

ii. Assignment: A practical task or setup questions given to measure a student's knowledge, skill and ability in subject or field.

iii. Seminar: A student-led academic activity designed to evaluate student's ability to research, understand, and present a specific topic.

iv. Group Discussion: An interactive method to asses a student's personality and inter- personal skills in team setting. It typically involves 5 to 15 participants discussing a topic for 15 to 30 minutes in presence of an observer.

b. Internal Summative Assessment

i. Mid-term tests
A test conducted roughly halfway through an academic term or semester.

ii. Laboratory performance
Students demonstrate their laboratory hands-on performance skills, knowledge and ability to apply concepts in subjects.

iii. Viva-voce
An oral examination conducted to evaluate a student's in-depth knowledge, understanding, and critical thinking regarding a subject or project.

(B) Weightage of Learning Efforts for External Assessment

Unit	Aligned CLOs	Total Learning Hours	Approximate weightage (Marks) to Learning levels (BT)			Total Marks
			Remember (R)	Understanding (U)	Application/ Analyse & above (A)	
I	1, 2, 4	30	4	5	3	12
II	1,4	30	4	5	4	13
III	3,4	30	4	5	4	13
IV	3,4, 5	30	3	5	4	12
		120	**15**	**20**	**15**	**50**

Master of Science Applied Statistics and Predictive Data Science Semester-I

- **Assessment and Evaluation**

Sr. No.	Assessment/Evaluation	Component	Weightage (%)
1	Continuous Internal Evaluation	FA: Seminars, Assignments, Quizzes, Group Discussion	25%
		SA: Mid Term Test	25%
2	End-Semester Examination	Written Exam	50%

(C) CLOs – PLOs Matrix

CLO	PLO														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CLO1	-	-	3	2	1	1	1	2	1	1	-	-	2	3	1
CLO2	1	-	-	2	1	-	-	1	1	-	1	-	2	3	-
CLO3	1	-	2	-	1	1	-	1	-	3	-	-	2	3	-
CLO4	-	-	1	2	1	1	-	2	1	-	1	-	2	3	-
CLO5	1	1	2	3	1	3	2	2	1	-	1	-	2	3	-

CLO – PLO correlation	Value
Strong	3
Moderate	2
Low	1
No correlation	-

- **Suggested Learning Materials Books:**

Sr.No.	Title	Author(s)	Edition/Year	Publisher
1	Probability Theory: An Introduction Using R	Shailaja R. Deshmukh, Akanksha S. Kashikar	First/2024	Taylor and Francis
2	Modern Probability Theory, An introductory Textbook	Bhat, B. R.	Third/1999	New Age International (P) Ltd, India
3	A First Course in Probability	Ross, S.	Ninth/2014	Pearson Edu. Dorling Kindersley (India) Pvt Ltd.
4	Probability Through Problems	Capinski, M. and Zastawniak T.	First/2003	Springer
5	Mathematical Statistics	Mukhopadhyay P.	Third/ 2006	Books and Allied Pvt. Ltd.

- **Online Resources (Open Source)**

Sr. No.	Description of Resource(s)	Weblink
1	Probability Theory: An Introduction Using R Book by S R Deshmukh and A. S. Kashikar	https://doi.org/10.1201/9781032619057

Master of Science Applied Statistics and Predictive Data Science Semester-I

Course Type	Course Code	Course Title	Teaching-Learning Scheme	Total Notional Hours	Course credits
			L-P-T		
DSE	**P2S01NEASP01**	Operations Research	4-0-0	120	04

- **Course Learning Outcomes (CLOs)**

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

CLO1: Apply optimization techniques such as Linear Programming, Transportation, and Assignment models to solve real-world decision-making problems.

CLO2: Demonstrate the ability to formulate and solve mathematical models using methods like Simplex, Integer Programming techniques.

CLO3: Evaluate project scheduling problems and inventory control models to improve operational efficiency.

CLO4: Apply queueing theory models to analyze service systems and evaluate system performance measures such as waiting time and utilization.

CLO5: Demonstrate analytical and problem-solving skills by interpreting results of Operations Research models and applying them in practical scenarios.

Unit	Course Content	Learning Pedagogies*	CLO(s)
I	Introduction to Operations Research & Linear Programming • Optimization Techniques: Overview of the optimization techniques, Applications and scope of Operations Research. • Linear Programming Problem (LPP): Introduction to Linear Programming Problems (LPP), General Structure of LPP, Advantages and Limitations of LPP, Application of LPP, Mathematical formulation of the problem, Linear Programming: Simplex method, Integer Programming problem, Gomory's Cut method, Branch and Bound method.	CL PBL	1 2
II	Transportation and Assignment Problems • Transportation Problem (TP): Introduction, Area of application, mathematical model of TP maximization and minimization problems, Degeneracy, Balanced and unbalanced TP. North-West Corner Method (NWCM), Least Cost Method (LCM), Vogel's Approximation Method (VAM), Modified Distribution Method (MODI). • Assignment Problem (AP):Introduction, Area of application, mathematical model of AP-maximization and minimization problems, Hungarian Method, Multiple Optimal Solutions.	CBL CL PBL	1 2 5
III	**Network Analysis and Inventory Control Models** • Network Analysis: Definition and formulation, critical path method, Project Evaluation and Review Technique (PERT).	CL MP CBL	3 5

	• An Inventory control problem, reasons for carrying inventory, Deterministic inventory control models with and without shortages, Problem of EOQ with quantity discounts, Inventory problems with uncertain demand: single period problem without set-up cost and single period model with set-up cost (s, S)) policy.		
IV	Queueing Theory and Queueing Models • Basics of Queueing Theory: Introduction, structure of Queuing system, performance measures of a queuing system. • Queueing Models and Performance Measures: steady state solution of M/M/C/∞/FIFO and M/M/C/N/FIFO with associated distributions of queue length and waiting time. (C=1 as particular case).	PBL MP ICT EL	4 5

- **Learning Pedagogies/Methods**
  (f) Classroom Lecture (CL)
  (g) Case-Based Learning (CBL)
  (h) Micro-Projects (MP)
  (i) Problem-Based Learning (PBL)
  (j) ICT-Enabled Learning through Digital Resources/ Virtual Labs/Webinars (ICT EL)
- **Assessment Methodologies**

(D) Internal Assessment

a. Internal Formative assessment (25 Marks)

v. MCQ and Objective question Quiz: Students select or give the correct answer from the list of options four /none for a given questions.

vi. Assignment: A practical task or setup questions given to measure a student's knowledge, skill and ability in subject or field.

vii. Seminar: A student-led academic activity designed to evaluate student's ability to research, understand, and present a specific topic.

viii. Group Discussion: An interactive method to asses a student's personality and inter- personal skills in team setting. It typically involves 5 to 15 participants discussing a topic for 15 to 30 minutes in presence of an observer.

b. Internal Summative Assessment (25 Marks)

i. Mid-term tests
A test conducted roughly halfway through an academic term or semester.

ii. Laboratory performance
Students demonstrate their laboratory hands-on performance skills, knowledge and ability to apply concepts in subjects.

iii. Viva-voce
An oral examination conducted to evaluate a student's in-depth knowledge, understanding, and critical thinking regarding a subject or project.

Master of Science Applied Statistics and Predictive Data Science Semester-I

(E) Weightage of Learning Efforts for External Assessment

Unit	Aligned CLOs	Total Learning Hours	Approximate weightage (Marks) to Learning levels (BT)			Total Marks
			Remember (R)	Understanding (U)	Application/ Analyse & above (A)	
I	1, 2	30	4	5	3	12
II	1, 2, 5	30	4	5	4	13
III	3, 5	30	4	5	4	13
IV	4, 5	30	3	5	4	12
		120	**15**	**20**	**15**	**50**

- **Assessment and Evaluation**

Sr. No.	Assessment/Evaluation	Component	Weightage (%)
1	Continuous Internal Evaluation	FA: Seminars, Assignments, Quizzes, Group Discussion	25%
		SA: Mid Term Test	25%
2	End-Semester Examination	Written Exam	50%

(F) CLOs – PLOs Matrix

CLO	PLO										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
CLO1	**3**	**3**	**2**	**1**	**2**	**-**	**-**	**2**	**1**	**2**	**1**
CLO2	**3**	**2**	**2**	**2**	**2**	**-**	**-**	**2**	**1**	**2**	**1**
CLO3	**3**	**3**	**2**	**1**	**3**	**-**	**1**	**2**	**1**	**2**	**1**
CLO4	**3**	**2**	**3**	**1**	**2**	**-**	**1**	**2**	**1**	**2**	**1**
CLO5	**3**	**3**	**2**	**2**	**2**	**2**	**2**	**3**	**2**	**3**	**2**

CLO – PLO correlation	Value
Strong	3
Moderate	2
Low	1
No correlation	-

- **Suggested Learning Materials Books:**

Sr. No.	Title	Author(s)	Edition/Year	Publisher
1	Operations Research: An Introduction	Hamdy A. Taha	10th Edition/2017	Pearson
2	Operations Research	Kanti Swarup, P. K. Gupta, Man Mohan	20th Edition/2022	Sultan Chand & Sons
3	Operations Research: A Model-Based Approach	H. A. Eiselt, Carl-Louis Sandblom	3rd Edition/2022	Springer

Master of Science Applied Statistics and Predictive Data Science Semester-I

4	Introduction to Operations Research	Hillier, Frederick S., and Gerald J. Lieberman	11th Edition/2018	McGraw-Hill
5	Operations Research: Problems and Solutions	Sharma, J. K.	3rd Edition/2009	Macmillan India Ltd.

- **Online Resources (Open Source)**

Sr. No.	Description of Resource(s)	Weblink
1	Operations Research By Prof. Rajendra Singh | IIT Kharagpur	https://onlinecourses.nptel.ac.in/noc25_ag15/preview

Master of Science Applied Statistics and Predictive Data Science Semester-I

Course Type	Course Code	Course Title	Teaching-Learning Scheme	Total Notional Hours	Course credits
			L-P-T		
DSE	**P2S01NEASP02**	Reinforcement Learning and Optimization	4-0-0	120	04

- **Course Learning Outcomes (CLOs)**

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

CLO1: Demonstrate reinforcement learning theory and algorithms and their application to sequential decision-making.

CLO2: Formulate RL problems using Markov Decision Processes (MDPs), defining state spaces, actions, and reward signals.

CLO3: Implement Deep Q-Networks (DQN) and policy-based methods Actor-Critic.

CLO4: Evaluate and compare optimization algorithms.

CLO5: Apply RL and optimization in clinical, trading and other professional environments.

Unit	Course Content	Learning Pedagogies*	CLO(s)
I	Fundamentals of RL • MDP framework: Markov Decision Process; states, actions, rewards, transitions; Policy and value functions; Bellman equations; Monte Carlo methods. • Temporal Difference learning: TD(0), TD(λ); SARSA; On-policy vs off-policy learning. • Exploration vs exploitation: ε-greedy, UCB, Thompson sampling; Tabular RL.	CL PBL ICT EL	1 2
II	Q-Learning and Deep RL • DQN: experience replay, target networks; Double DQN; Prioritized experience replay; Dueling DQN; Rainbow DQN. • Actor-Critic methods: A2C, A3C; Proximal Policy Optimization (PPO); Trust Region Policy Optimization (TRPO); SAC for continuous actions.	CL CBL ICT EL	1 3 5
III	Advanced RL and Applications • Model-based RL: Dyna architecture, World Models; Meta-RL • RL in Profession: Clinical treatment optimization; Trading and Finance; Robotic control; NLP; Game playing • Simulation environments: OpenAI Gym	CL ICT EL	1 5
IV	Metaheuristic Optimization • Integer Programming: Branch and bound; Genetic algorithm; simulated annealing • Ant colony optimization; Particle Swarm optimization	CL MP	4 5

- **Learning Pedagogies/Methods**
  (k) Classroom Lecture (CL)
  (l) Case-Based Learning (CBL)
  (m) Micro-Projects (MP)
  (n) Problem-Based Learning (PBL)
  (o) ICT-Enabled Learning through Digital Resources/ Virtual Labs/Webinars (ICT EL)

- **Assessment Methodologies**

(G) Internal Assessment

a. Internal Formative assessment

ix. MCQ and Objective question Quiz: Students select or give the correct answer from the list of options four /none for a given questions.

x. Assignment: A practical task or setup questions given to measure a student's knowledge, skill and ability in subject or field.

xi. Seminar: A student-led academic activity designed to evaluate student's ability to research, understand, and present a specific topic.

xii. Group Discussion: An interactive method to asses a student's personality and inter- personal skills in team setting. It typically involves 5 to 15 participants discussing a topic for 15 to 30 minutes in presence of an observer.

b. Internal Summative Assessment

iv. Mid-term tests
A test conducted roughly halfway through an academic term or semester.

v. Laboratory performance
Students demonstrate their laboratory hands-on performance skills, knowledge and ability to apply concepts in subjects.

vi. Viva-voce
An oral examination conducted to evaluate a student's in-depth knowledge, understanding, and critical thinking regarding a subject or project.

(H) Weightage of Learning Efforts for External Assessment

Unit	Aligned CLOs	Total Learning Hours	Approximate weightage (Marks) to Learning levels (BT)			Total Marks
			Remember (R)	Understanding (U)	Application/ Analyse & above (A)	
I	1, 2	30	4	5	3	12
II	1, 3, 5	30	4	5	4	13
III	1, 5	30	4	5	4	13
IV	4, 5	30	3	5	4	12
		120	**15**	**20**	**15**	**50**

Master of Science Applied Statistics and Predictive Data Science Semester-I

- **Assessment and Evaluation**

Sr. No.	Assessment/Evaluation	Component	Weightage (%)
1	Continuous Internal Evaluation	FA: Seminars, Assignments, Quizzes, Group Discussion	25%
		SA: Mid Term Test	25%
2	End-Semester Examination	Written Exam	50%

(I) CLOs – PLOs Matrix

CLO	PLO										
	1	2	3	4	5	6	7	8	9	10	11
CLO1	2	-	-	--	-	-	2	2	1	3	2
CLO2	2	-	-	2	-	2	2	-	1	3	2
CLO3	2	3	-	-	-	-	-	-	2	3	2
CLO4	-	-	-	2	-	-	-	2	1	3	2
CLO5	2	3	-	-	-	-	2	2	2	3	2

CLO – PLO correlation	Value
Strong	3
Moderate	2
Low	1
No correlation	-

- **Suggested Learning Materials Books:**

Sr. No.	Title	Author(s)	Edition/Year	Publisher
1	Deep Reinforcement Learning Hands-On	Lapan, M.	Second/2020	Packt
2	Algorithms for Reinforcement Learning	Szepesvári, C.	First/2010	Morgan & Claypool
3	Reinforcement Learning: An Introduction	Sutton, R.S. & Barto, A.G.	Second/2020	MIT Press

Master of Science Applied Statistics and Predictive Data Science Semester-I

Course Type	Course Code	Course Title	Teaching-Learning Scheme L-P-T	Total Notional Hours	Course credits
DSC	**P2S01NCASP03**	Database Management and SQL Analytics	0-4-0	120	04

- **Course Learning Outcomes (CLOs)**

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

CLO1: Design database schemas with appropriate constraints and data structures, and implement data manipulation and access control operations for efficient data management.

CLO2: Analyze datasets using SQL queries involving filtering, aggregation, joins, and subqueries to extract meaningful insights for decision-making.

CLO3: Develop PL/SQL programs using block structures, conditional and looping statements to automate data processing tasks.

CLO4: Evaluate database operations and query performance by applying advanced SQL techniques and structured programming for optimized data handling.

CLO5: Construct data-driven solutions by integrating SQL and PL/SQL techniques to manage, process, and analyze structured datasets in real-world scenarios with ethical considerations.

Unit	Course Content	Learning Pedagogies*	CLO(s)
I	Database Fundamentals and Basic SQL Operations • Database Creation and Connection • Table Creation using DDL Commands • Data Manipulation using DML Commands • Access Control using DCL Commands • Table Design with Constraints	ICT EL	1 3 5
II	Data Retrieval Techniques • Data Retrieval using SELECT Queries • Aggregate Functions with GROUP BY and HAVING • Use of SQL Operators	ICT EL	2 4 5
III	Advanced SQL Querying • Correlated Subqueries • Modifying Constraints using ALTER • User Creation and Privilege Management	ICT EL	1 2 4
IV	PL/SQL Programming • Basic PL/SQL Programming using Block Structure • PL/SQL Conditional Statements • PL/SQL Looping Statements • Exception Handling in PL/SQL	ICT EL	3 4 5

- **Learning Pedagogies/Methods**

(p) Classroom Lecture (CL)
(q) Case-Based Learning (CBL)
(r) Micro-Projects (MP)
(s) Problem-Based Learning (PBL)
(t) ICT-Enabled Learning through Digital Resources/ Virtual Labs/Webinars (ICT EL)

- **Assessment Methodologies**

(J) Internal Assessment

a. Internal Formative assessment

xiii. **MCQ and Objective question Quiz:** Students select or give the correct answer from the list of options four /none for a given questions.

xiv. **Assignment:** A practical task or setup questions given to measure a student's knowledge, skill and ability in subject or field.

xv. **Seminar:** A student-led academic activity designed to evaluate student's ability to research, understand, and present a specific topic.

xvi. **Group Discussion:** An interactive method to asses a student's personality and inter- personal skills in team setting. It typically involves 5 to 15 participants discussing a topic for 15 to 30 minutes in presence of an observer.

b. Internal Summative Assessment

vii. **Mid-term tests**
A test conducted roughly halfway through an academic term or semester.

viii. **Laboratory performance**
Students demonstrate their laboratory hands-on performance skills, knowledge and ability to apply concepts in subjects.

ix. **Viva-voce**
An oral examination conducted to evaluate a student's in-depth knowledge, understanding, and critical thinking regarding a subject or project.

(K) Weightage of Learning Efforts for External Assessment

Unit	Aligned CLOs	Total Learning Hours	Approximate weightage (Marks) to Learning levels (BT)			Total Marks
			Remember (R)	Understanding (U)	Application/ Analyse & above (A)	
I	1,3,5	30	4	5	3	12
II	2,4,5	30	4	5	4	13
III	1,2,4	30	4	5	4	13
IV	3,4,5	30	3	5	4	12
		120	**15**	**20**	**15**	**50**

- **Assessment and Evaluation**

Sr. No.	Assessment/Evaluation	Component	Weightage (%)
1	Continuous Internal Evaluation	FA: Seminars, Assignments, Quizzes, Group Discussion	25%
		SA: Mid Term Test	25%
2	End-Semester Examination	Written Exam	50%

Master of Science Applied Statistics and Predictive Data Science Semester-I

(L) CLOs – PLOs Matrix

CLO	PLO										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
CLO1	-	-	-	3	-	2	1	-	-	-	-
CLO2	-	-	-	3	-	2	2	-	-	-	-
CLO3	-	-	-	3	-	2	1	-	-	-	-
CLO4	-	-	-	3	-	2	1	1	-	-	-
CLO5	-	-	-	3	-	3	3	2	3	1	2

CLO – PLO correlation	Value
Strong	3
Moderate	2
Low	1
No correlation	-

- **Suggested Learning Materials Books:**

Sr. No.	Title	Author(s)	Edition/Year	Publisher
1	Database Management System: Oracle SQL and PL/SQL	Pranab Kumar Das Gupta & P. Radha Krishna	2nd Edition/ 2013	PHI Learning Pvt. Ltd. (Prentice Hall India)
2	SQL, PL/SQL: The Programming Language of Oracle	Ivan Bayross	2009	BPB Publications, New Delhi
3	Fundamentals of Database Systems	Ramez Elmasri & Shamkant B. Navathe	5th Edition/2008	Dorling Kindersley (India) Pvt. Ltd., New Delhi

- **Online Resources (Open Source)**

Sr. No.	Description of Resource(s)	Weblink
1	PL/SQL Complete Tutorial	**https://www.plsql.co/procedure.html**

Master of Science Applied Statistics and Predictive Data Science Semester-I

Course Type	Course Code	Course Title	Teaching-Learning Scheme	Total Notional Hours	Course credits
			L-P-T		
DSC	**P2S01NCASP04**	Statistical Computing through Python	0-4-0	120	04

- **Course Learning Outcomes (CLOs)**

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

CLO1: Demonstrate basic Python concepts such as variables, data types and control structures for statistical problems.

CLO2: Apply functions, recursion and modules to perform statistical computations.

CLO3: Apply object-oriented programming concepts like classes, inheritance and polymorphism for data handling.

CLO4: Apply Python libraries such as NumPy, Pandas, Matplotlib for data analysis and visualization.

CLO5: Evaluate statistical programs using debugging, testing and interpretation of results.

Unit	Course Content	Learning Pedagogies*	CLO(s)
I	Basics of Python for Statistics • Python Basics: Variables and Data Types • Descriptive Statistics using Python • Conditional Statements in Statistical Problems • Loops and Data Structures in Python	CL PBL ICT EL	1
II	Functions and Modules • User-Defined Functions for Statistical Measures • Recursive Functions in Python • Lambda Functions and Exception Handling • Using Python Modules for Mathematical Operations	CL CBL PBL	2
III	Object-Oriented Programming • Classes and Objects for Data Handling • Inheritance in Python • Polymorphism in Python • Debugging and Testing of Programs	CL CBL MP	3
IV	Python for Data Analysis • Array Operations using NumPy • Data Analysis using Pandas • Data Visualization using Matplotlib • Mathematical Computations using SciPy/SymPy	CL CBL MP ICT EL	4 5

- **Learning Pedagogies/Methods**

(a) Classroom Lecture (CL)
(b) Case-Based Learning (CBL)
(c) Micro-Projects (MP)
(d) Problem-Based Learning (PBL)
(e) ICT-Enabled Learning through Digital Resources/ Virtual Labs/Webinars (ICT EL)

- **Assessment Methodologies**

(M) Internal Assessment

a. Internal Formative assessment

xvii. Group Discussion: An interactive method to asses a student's personality and inter- personal skills in team setting. It typically involves 5 to 15 participants discussing a topic for 15 to 30 minutes in presence of an observer.

b. Internal Summative Assessment

iv. Mid-term tests
A test conducted roughly halfway through an academic term or semester.

v. Laboratory performance
Students demonstrate their laboratory hands-on performance skills, knowledge and ability to apply concepts in subjects.

(N) Weightage of Learning Efforts for External Assessment

Unit	Aligned CLOs	Total Learning Hours	Approximate weightage (Marks) to Learning levels (BT)			Total Marks
			Remember (R)	Understanding (U)	Application/ Analyse & above (A)	
I	1	30	4	5	3	12
II	2	30	4	5	4	13
III	3	30	4	5	4	13
IV	4, 5	30	3	5	4	12
		120	**15**	**20**	**15**	**50**

- **Assessment and Evaluation**

Sr. No.	Assessment/Evaluation	Component	Weightage (%)
1	Continuous Internal Evaluation	FA: Seminars, Assignments, Quizzes, Group Discussion	25
		SA: Mid Term Test	25
2	End-Semester Examination	Written Exam	50

(O) CLOs – PLOs Matrix

	PLO										
CLO	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
CLO1	**3**	**2**	**-**	**2**	**-**	**-**	**-**	**1**	**-**	**-**	**-**
CLO2	**3**	**2**	**2**	**3**	**-**	**-**	**-**	**1**	**-**	**-**	**-**
CLO3	**2**	**2**	**2**	**3**	**-**	**-**	**-**	**1**	**1**	**-**	**-**
CLO4	**3**	**3**	**3**	**3**	**-**	**-**	**2**	**2**	**1**	**1**	**2**
CLO5	**2**	**2**	**2**	**3**	**1**	**2**	**3**	**2**	**-**	**1**	**2**

Master of Science Applied Statistics and Predictive Data Science Semester-I

CLO – PLO correlation	Value
Strong	3
Moderate	2
Low	1
No correlation	-

- **Suggested Learning Materials Books:**

Sr. No.	Title	Author(s)	Edition/Year	Publisher
1	An Introduction to Statistics with Python	Thomas Haslwanter	1st Edition/2016	Springer
2	Introduction to Computation and Programming Using Python	John V Guttag	3rd Edition /2021	Prentice Hall of India
3	Fundamentals of Python: First Programs	Kenneth A. Lambert	3rd Edition/2023	Cenagepublication
4	Data Structures and Algorithms in Python	Michael T. Goodrich, Roberto Tamassia, Michael H. Goldwasser	1st Edition/2013	Wiley

- **Online Resources (Open Source)**

Sr. No.	Description of Resource(s)	Weblink
1	IBM, Statistics for Data Science with Python	https://www.coursera.org/learn/statistics-for-data-science-python
2	100 Days of Python Bundle by Sarath Srinivas Sannyboyina	https://swayam-plus.swayam2.ac.in/courses/course-details?id=P_GUVI_01

Master of Science Applied Statistics and Predictive Data Science Semester-I

Course Type	Course Code	Course Title	Teaching-Learning Scheme	Total Notional Hours	Course credits
			L-P-T		
DSC	**P2S01NCASP05**	Communication Skills	0-2-0	60	02

- **Course Learning Outcomes (CLOs)**

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

CLO1: Demonstrate clear, concise, and courteous verbal and non-verbal communication skills in interpersonal, academic, and professional contexts through self-introductions, role plays, discussions, and presentations.

CLO2: Apply active listening techniques and identify communication barriers to achieve complete, coherent, and effective communication in individual and group interactions.

CLO3: Compose and evaluate emails, business letters, reports, and other professional documents that are correct, concise, complete, and coherent, following appropriate formats and communication etiquette.

CLO4: Communicate ideas and information in a clear, concrete, and coherent manner through group discussions, oral presentations, debates, and interview simulations while demonstrating teamwork and leadership skills.

CLO5: Analyze and adapt communication strategies in digital and cross-cultural environments to ensure correct, courteous, complete, and culturally sensitive communication while adhering to netiquette and ethical practices.

Unit	Course Content	Learning Pedagogies*	CLO(s)
I	Icebreaker & Self-Introduction Skills • Students prepare a 2-minute self-introduction • Include education, strengths, hobbies, goals Active Listening Drill • One student speaks/read aloud from a book for 2 minutes • Listener take dictation and give feedback Communication Barriers Role Play • Groups perform scenarios showing barriers (noise, misunderstanding, language issues) • Other group identifies barriers Non-Verbal Communication Awareness • Silent role-play (no speaking allowed) • Express emotions using only body language Email Writing Workshop • Write formal and informal emails (complaint, request, inquiry) • Peer review correction session Business Letter Writing Exercise • Draft a formal business letter (e.g., leave request, complaint letter)	CBL ICT EL	1 2 3

II	Report Writing Task • Write a short report on an event (seminar, workshop, survey) Group Discussion Simulation • Topics: "Social media impact", "Online education vs classroom learning", and others. • Evaluate participation, clarity, leadership Mock Interview Session (Online / Offline Mode) • Students act as interviewer and candidate • Questions on education, skills, goals Oral Presentation Practice (Story Telling Format, Auditorium) • 5-minute presentation with PPT • Topics: climate change, technology, education Netiquette & Digital Communication Activity • Identify mistakes in sample emails/chats • Rewrite corrected versions Cross-Cultural Communication Debate • Debate on cultural communication differences • Example: "Work culture differences across countries"	CBL ICT EL	1 3 4 5

- **Learning Pedagogies/Methods**
  (u) Classroom Lecture (CL)
  (v) Case-Based Learning (CBL)
  (w) Micro-Projects (MP)
  (x) Problem-Based Learning (PBL)
  (y) ICT-Enabled Learning through Digital Resources/ Virtual Labs/Webinars (ICT EL)
- **Assessment Methodologies**

(P) Internal Assessment

a. Internal Formative assessment

xviii. Active Listening: An interactive method to asses a student's ability to listen, grasp, take dictation accurately and give feedback of lower to higher level.

xix. Group Discussion: An interactive method to asses a student's personality and inter- personal skills in team setting. It typically involves 5 to 15 participants discussing a topic for 15 to 30 minutes in presence of an observer.

xx. Cross-Cultural Communication Debate: An interactive method to assess a student's communication, argumentation, critical thinking, and cross-cultural awareness skills through structured debates on cultural and workplace communication issues. It typically involves team-based discussions in the presence of an evaluator and focuses on clarity, coherence, confidence, teamwork, and respectful exchange of ideas.

b. Internal Summative Assessment

x. Oral Performance: An interactive method to assess a student's verbal communication, presentation, storytelling, and public speaking skills. It typically involves a 5-minute oral presentation with or without visual aids (PPT) on a given topic in an auditorium or classroom setting. The assessment evaluates clarity, confidence, organization of ideas,

pronunciation, body language, audience engagement, and the ability to communicate information effectively.

xi. **Mock Interview Performance:** An interactive method to assess a student's professional communication, interpersonal skills, confidence, and interview preparedness. It typically involves a simulated interview in which students act as candidates and respond to questions related to their education, skills, achievements, career goals, and problem-solving abilities in the presence of an evaluator. The assessment focuses on content of responses, communication effectiveness, confidence, professional etiquette, and overall interview performance.

- **Assessment and Evaluation**

Sr. No.	Assessment/Evaluation	Component	Weightage (%)
1	Continuous Internal Evaluation	FA: Active Listening, Group Discussion, Cross-Cultural Communication Debate, Group Discussion	25%
		SA: Oral Performance, Mock Interview Performance	25%

(Q)CLOs – PLOs Matrix

CLO	PLO										
	1	2	3	4	5	6	7	8	9	10	11
CLO1	-	-	-	-	-	1	-	3	2	-	1
CLO2	-	-	-	-	-	1	1	2	2	-	1
CLO3	-	1	-	-	-	2	1	3	1	2	-
CLO4	-	1	-	-	-	1	-	3	3	1	1
CLO5	-	-	-	-	-	3	2	2	2	-	2

CLO – PLO correlation	Value
Strong	3
Moderate	2
Low	1
No correlation	-

- **Suggested Learning Materials Books:**

Sr.No.	Title	Author(s)	Edition/Year	Publisher
1	Indian Folk Tales	Prithvishwar Gayen	1999	A CBT Publication
2	Amar Chitra Katha Series: A Jataka Stories	Pratap Mulick	1976	Amar Chitra Katha Pvt Ltd
3	The Art of Statistics Learning from Data	David Spiegelhalter	2019	Penguin Random House UK

- **Online Resources (Open Source)**

Sr. No.	Description of Resource(s)	Weblink
	NPTEL	**https://nptel.ac.in/courses/109104031**

Master of Science Applied Statistics and Predictive Data Science Semester-I