Programme Structure and Course Syllabus

Sharda School of Basic Sciences & Research

Department of Mathematics

M.Sc. (Data Science & Analytics)

Programme Code: SBR0309

Batch: 2024-26

1.1 Vision, Mission, and Core Values of the University

Vision of the University

To serve the society by being a global University of higher learning in pursuit of academic excellence, innovation and nurturing entrepreneurship.

Mission of the University

1. **Transformative educational experience**
2. **Enrichment by educational initiatives that encourage global outlook**
3. **Develop research, support disruptive innovations and accelerate entrepreneurship**
4. **Seeking beyond boundaries**

Core Values

1. **Integrity**
2. **Leadership**
3. **Diversity**
4. **Community**

1.2 Vision and Mission of the School

Vision of the School

Achieving excellence in the realm of science to address the challenges of evolving society.

Mission of the School

1. Equip the students with knowledge and skills
2. Capacity building by providing academic flexibility to student and faculty members
3. To establish centre of excellence for innovative research
4. Address the deficiencies of the society pertaining to environment
5. To strengthen academic- industry collaboration for better employability
6. Developing a culture for continued betterment in all facets of life

Core Values

1.Integrity
2. Leadership
3. Diversity
4. Community

1.3 Vision and Mission Department of Mathematics

Vision of the Department

To become a globally recognized destination for education in applied mathematics and research.

Mission of the Department

1. To develop mathematical skills in students and make them employable across a wide range of professions and promote interest in research.
2. To develop entrepreneurial skills in students to serve the society at large.
3. To develop skills for the applications of mathematics in the various fields.

Core Values

1. Integrity
2. Leadership
3. Diversity
4. Community

1.4 Programme Educational Objectives (PEOs)

PEO1: The graduates will achieve deep subject knowledge in the courses of study to enable employed in industry, government, and entrepreneurial endeavors to have a successful professional career.
PEO2: The graduates will develop a positive attitude and skills to enable a multi-facet personality.
PEO3: The graduates will prepare to pursue higher education and research.
PEO4: The graduates will develop to contribute to society and human well-being by applying ethical principles.

1.4.1 Programme Outcomes (POs)

PO1: Data Science knowledge: Engage in continuous reflective learning in the context of technology and scientific advancement.
PO2: Modern software tool usage: Acquire the skills in handling data science programming tools for problem-solving and solution analysis for domain-specific problems.
PO3: Critical thinking: Ability to understand the abstract concepts that lead to various data science theories in Mathematics, Statistics, and Computer science.
PO4: Problem analysis: Problem analysis and design ability to identify analyze and design solutions for data science problems using fundamental principles of mathematics, Statistics, computing sciences, and relevant domain disciplines.
PO5: Innovation and Entrepreneurship: Produce innovative IT solutions and services based on global needs and trends.

1.4.2 Programme Specific Outcomes (PSOs)

PSO1: Utilize data science theories for societal and environmental concerns.
PSO2: Understand and commit to professional ethics and cyber regulations, responsibilities, and norms of professional computing practices.
PSO3: Use research-based knowledge and research methods including design of experiments, analysis, interpretation of data, and synthesis of the information to provide valid conclusions.
PSO4: Understand the role of statistical approaches and apply the same to solve real-life problems in the fields of data science and apply the research-based knowledge to analyze and solve advanced problems in data science.

Department of Mathematics
Sharda School of Basic Sciences and Research
M. Sc. (Data Science & Analytics) Batch: 2024-26
TERM: 2401 (Semester-I)

S. No.	COURSE CODE	Course Name	Teaching Load				CREDITS	PRE-REQUISITE/CO-REQUISITE	Type of Course: 1. CC 2. AECC 3. SEC 4. DSE
	THEORY								
			L	**T**	**P**	**TOTAL**			
1.	MDA101	Foundations of Data Science	4	0	0	4	4		CC
2.	MMT104	Statistical Methods	4	0	0	4	4		CC
3.	MDA102	Mathematics for Machine Learning	4	0	0	4	4		CC
4.	MDA103	Probability Theory and Distributions	4	0	0	4	4		CC
5.	MDA104	Next Generation Databases	4	0	0	4	4		AECC
	PRACTICALS								
6.	MDA151	Practical -I (Based on Paper MMT104, MDA102UsingExcel/SPSS/Mini-tab)	0	0	4	4	2		CC
7	MDA152	Practical -II (Based on Paper MMT104, MD104UsingR/Python)	0	0	4	4	2		AECC
8	RBL001	Research Based Learning-1	0	0	4	0	0		Project
TOTAL							**24**		

CC: Core Course, AECC: Ability Enhancement Compulsory Courses, SEC: Skill Enhancement Courses, DSE: Discipline Specific Courses

Department of Mathematics
Sharda School of Basic Sciences and Research
M. Sc. (Data Science & Analytics) Batch: 2024-26
TERM: 2402 (Semester-II)

S. No.	COURSE CODE	Course Name	Teaching Load				CREDITS	PRE-REQUISITE/CO-REQUISITE	Type of Course: 1. CC 2. AECC 3. SEC 4. DSE
	THEORY								
			L	T	P	TOTAL			
1.	MMT130	Numerical Analysis	4	0	0	4	4		CC
2.	MDA105	Regression Analysis and Predictive Models	4	0	0	4	4		CC
3.	MDA109	Survey Sampling	4	0	0	4	4		CC
4.	MDA107	Advanced Big Data and Text Analytics	4	0	0	4	4		CC
5.	MDA108	Data Mining & Artificial Intelligence	4	0	0	4	4		SEC
	PRACTICALS								
6.	NVI0101	Prompt Engineering	0	0	2	2	0		Audit
7.	CCU401	Community Connect	0	0	4	4	2		AECC
8.	MDA153	Practical-III (Based on Paper MDA105, 109, 107 Using R/Python/SAS/SPSS)	0	0	4	4	2		CC
9.	MDA154	Practical-IV(Based on Paper MDA108 using R/Python)	0	0	4	4	2		SEC
10	RBL002	Research Based Learning-2	0	0	4	0	0		Project
TOTAL							**26**		

Department of Mathematics
Sharda School of Basic Sciences and Research
M. Sc. (Data Science & Analytics) Batch: 2024-26
TERM: 2501 (Semester-III)

S. No.	COURSE CODE	Course Name	Teaching Load				CREDITS	PRE-REQUISITE/CO-REQUISITE	Type of Course: 1. CC 2. AECC 3. SEC 4. DSE
	THEORY								
			L	T	P	TOTAL			
1.	MDA201	Inferential Statistics	4	0	0	4	4		CC
2.	MDA202	Multivariate Data Analysis	4	0	0	4	4		CC
3.	MDA203	Soft Computing Techniques	4	0	0	4	4		AECC
4.	MDA215	Advances in Design of experiment	4	0	0	4	4		CC
5.	OPEXXX	Open elective (GE)	2	0	0	2	2		AECC
	PRACTICALS								
6.	MDA251	Practical -V (based on MDA201, MDA202, MDA215) (using SPSS/SAS/STRATA)	-	-	4	4	2		CC
7.	MDA252	Practical-VI (using based on MDA203)	-	-	4	4	2		AECC
		TOTAL					22		

Department of Mathematics
Sharda School of Basic Sciences and Research
M. Sc. (Data Science & Analytics) Batch: 2024-26
TERM: 2502 (Semester-IV)

S. No.	COURSE CODE	Course Name	HOURS				CREDITS	PRE-REQUISITE/CO-REQUISITE	Type of Course: 1. CC 2. AECC 3. SEC 4. DSE
	THEORY								
			L	**T**	**P**	**TOTAL**			
1.	MDAXXX	Elective-I(Online/Offline Courses)	4	0	0	4	4		DSE
2.	MDAXXX	Elective-II(Online/Offline Courses)	4	0	0	4	4		DSE
	DISSERTATION								
3.	MDA253	Capstone project (Based on fulltime training program/internship program in any government/private institute or industry during last semester)		-	20	6weeks (min. 30days)	10		AECC
		TOTAL					**18**		

List of Courses:

MDA212: Statistical Analysis Count Data and Survival Analysis, MDA213: Industrial Statistics, MDA214: Statistical Simulation, MDA222: Applied Econometrics and MDA: 229 Stochastic Processes.

School: SSBSR		**Batch: 2024-26**	
Program: M.Sc.		**Academic Year: 2024-25**	
Branch: Data Science & Analytics		**Semester: I**	
1	Course Code	**MDA101**	
2	Course Title	**Foundations of Data Science**	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	Compulsory	
5	Course Objective	The course is aimed at building the fundamentals of data science. Imparting design thinking capability to build big data and developing design skills of models for big data problems. Gaining practical experience in programming tools for data sciences and also empowering students with tools and techniques used in data science.	
6	Course Outcomes	CO1: Explain data evolution and application on the data. (K1, K2) CO2: Discuss the basic concepts of data science. (K2, K3) CO3: Apply Matrix decomposition techniques to perform data analysis.(K3, K4) CO4: Explain the concept of a real-life solution. (K3, K4) CO5: Apply and develop basic Machine Learning Algorithms. (K5, K6) CO6: Apply the statistical measures of R in a real-time environment.(K5, K6)	
7	Course Description	A PG-level course in the foundation of data science intended to versestudents in the techniques necessary to understand and carry out methods in the foundation of data science.	
8	**Outline syllabus**		**CO Mapping**
	Unit 1	**Introduction**	
	A	Introduction-What is Data Science?	CO1
	B	The steps in Doing Data Science-Skills needed to do DataScience storing data-combining bits into larger structures	CO1
	C	The steps in Doing Data Science-Skills needed to identify Data Problems.	CO1
	Unit 2	**EDA**	
	A	Big Data and Data Science - Big Data Analytics,Business intelligence vs big data, big data frameworks,	CO2
	B	Exploratory Data Analysis (EDA), statistical measures,	CO2
	C	Basic tools (plots, graphs, and summary statistics) of EDA, Data Analytics Lifecycle, Discovery	CO2
	Unit 3	**Data Pre-processing and Feature Selection**	
	A	Data cleaning - Data integration - Data Reduction - Data Transformation and Data Discretization.	CO3
	B	Feature Generation and Feature Selection, Feature Selection algorithms: Filters- Wrappers - Decision Trees -Random Forests	CO3
	C	Descriptive statistics-Using Histograms to understand a distribution-Normal Distribution.	CO3, CO6
	Unit 4	**Basic of R**	
	A	Getting Started with R-Installing R-Using R-Creating and Using Vectors-Follow the Data-Understanding existing.	CO4
	B	Data sources-Exploring Data Models-Rows and Columns-Creating Data Frames-Exploring.	CO4
	C	Importing Data Using R Studio-Accessing Excel data- Accessing Database-Comparing SQL and R for accessing a data set.	CO4, CO6
	Unit 5	**Basic Data Mining**	
	A	Data Mining Overview-Association Rule Mining-Text Mining-Supervised and Unsupervised Learning.	CO5

	B	Supervised Learning via Support Vector Machines- Support.			CO5
	C	Vector Machines in R-Creating Web Applications with R.			CO5, CO6
	Mode of examination	**Theory**			
	Weightage Distribution	CA	MTE	ETE	
		25%	25%	50%	
	Text book/s*	1. Jeffrey S. Saltz, Jeffre M. Stanton, "AnIntroduction to Data Science", Sage Publications.			
	Other References	1. Nina Zumal, John Mount (2014). Practical Data science in R, Managing Publication Company 2. Bernard Kolman, Robert C. Busby and Sharon Ross (2004). Discrete Mathematical Structures, New Delhi: Prentice Hall 3. V. Bhuvaneswari, T. Devi, (2016). Big Data Analytics: A Practitioner's Approach, Bharathiar University 4. V. Bhuvaneswari (2016). Data Analytics with R, Bharathiar University.			

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA101.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA101.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA101.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA101.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA101.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA101.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		**Batch: 2024-26**	
Program: M.Sc.		**Academic Year: 2024-25**	
Branch: Data Science & Analytics		**Semester: I**	
1	Course Code	MDA102	
2	Course Title	Mathematics for Machine Learning	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	Compulsory	
5	Course Objective	To enable the students to understand the concept of mathematics inmachine learning.	
6	Course Outcomes	CO1: Solve a system of Linear equations by applying the Gauss Elimination method. (K2, K3) CO2: Explain the basics of Vectors, Spaces, and Affine Spaces. (K2, K3) CO3: Apply different methods to evaluate the Inverse and Rank of aMatrix. (K1, K2, K3) CO4: Evaluate Eigen values and Eigen vectors using Linear transformation and power methods. (K3, K4) CO5: Evaluate Derivatives and Partial Derivatives using rules ofdifferentiation. (K4, K5) CO6: Apply optimization using gradient function. (K5, K6)	
7	Course Description	The course focuses on iterative techniques for solving large sparse linear systems of equations which typically stem from the Discretization of partial differential equations. In addition, the computation of eigenvalues, least square problems and error analysis will be discussed.	
8			CO Mapping
	Unit 1	**Matrices and Determinants**	
	A	Matrices – Determinant, Identity matrix, Inverse of amatrix.	CO1
	B	The rank of a matrix, Nullity, trace of a matrix.	CO1
	C	Eigen values, Eigen vectors, Matrix decompositions.	CO1
	Unit 2	**Basic Concept of Linear Algebra**	
	A	Linear Algebra-System of Linear equations, SolvingSystem of Linear equations.	CO2
	B	Linear Independence, Vectors, Scalars, Addition, Scalar multiplication.	CO2
	C	Dot product, vector projection, cosine similarity	CO2
	Unit 3	**Vector**	
	A	Orthogonal vectors, normal and Orthonormal vectors.	CO3
	B	Vector norm, vector space, linear combination.	CO3
	C	Basis of vectors, Affine spaces.	CO3
	Unit 4	**Derivatives**	
	A	Differentiation, rules of differentiation, Derivatives, Scalar derivatives.	CO4
	B	Partial derivatives, Principle Component analysis – Concepts and properties.	CO4

	C	Dimensionality reduction with PCA			CO4
	Unit 5	**Derivatives of Function**			
	A	Differentiation of univariate functions, Partial differentiation and gradients.			CO5
	B	Gradient of a vector-valued function. Gradient of matrices.			CO5
	C	Optimization using gradient functions, Constrained optimization, and Lagrange multipliers. Convex optimization.			CO6
	Mode of examination	**Theory**			
	Weightage Distribution	CA	MTE	ETE	
		25 %	25 %	50 %	
	Text book/s*	1. Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, "Mathematics for Machine Learning**", Cambridge University Press, 2020.**			
	Other References	1. Erwin Kreyszig, Advanced Engineering Mathematics, 10^{th} Edition., John Wiley & Sons, (2014). 2. B. S.Grewal, Higher Engineering Mathematics, 38th Edition. Khanna Publications, (2005).			

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA102.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA102.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA102.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA102.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA102.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA102.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		**Batch: 2024-26**	
Program: M.Sc.		**Academic Year: 2024-25**	
Branch: Data Science & Analytics		**Semester: I**	
1	Course Code	**MDA103**	
2	Course Title	**Probability Theory and Distributions**	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	Compulsory	
5	Course Objective	To incorporate the concepts of probability theory and its applications as the core material in building theoretical ideas along with real- lifedata.	
6	Course Outcomes	After completion of this course, students will be able to CO1: Develop problem-solving techniques needed to calculate probability and conditional probability. (K2, K3, K4) CO2: Formulate fundamental probability distribution and density functions, as well as functions of random variables, derive the probability density function of transformations. (K4, K5) CO3: Derive the expectation and conditional expectation, and describe their properties. (K4, K5) CO4: Discuss various types of generating functions used in statistics.(K3, K4) CO5:Apply sampling distributions to testing of hypotheses. (K4, K5) CO6: Illustrate and correlate the statistical problems into Statisticalanalysis. (K5, K6)	
7	Course Description	To integrate the intrinsic ideas of preliminary and advanced distributions to correlate with real-world scenarios.	
8			CO Mapping
	Unit 1	**Probability and Random variables**	
	A	Introduction to Random Experiments, Empirical basis of probability, Algebra of events, laws of probability; Conditional Probability, Independence, Bayes' law; Application of probability to business and economics.	CO1
	B	One-dimensional Random Variable-Discrete and Continuous; Distribution functions and their properties.	CO1
	C	Bivariate Random Variables- Joint Probability functions,marginal distributions, conditional distribution functions; The notion of Independence of Random variables.	CO1
	Unit 2	**Random Variables and Expectations**	
	A	Functions of random variables: introduction, distribution function technique, transformation technique: one variable, transformation technique: several variables, theory, and applications.	CO2
	B	Expectation, Variance, and Co-variance of random variables; Conditional expectation and conditional variance.	CO2

	C	Markov, Holder, Jensen, and Chebyshev's Inequality; Weak Law of Large numbers, Strong law of large numbers and Kolmogorov theorem; Central Limit Theorem.	CO2
	Unit 3	**Generating Functions and Discrete Distributions**	
	A	Probability generating function (p.g.f.), moment generating function (m.g.f.), characteristic function (c.f.).	CO3
	B	Properties and Applications. Probability distributions of functions of random variables: one and two dimensions.	CO3
	C	Bernoulli, Binomial, Poisson, Geometric, Hyper geometric, Negative Binomial, Multinomial, distributions and Discrete Uniform distribution - definition, properties and applications with numerical problems.	CO3, CO6
	Unit 4	**Continuous Distributions**	
	A	Uniform, Normal distribution function, Exponential distribution functions - definition, properties, and applications.	CO4
	B	Gamma, Beta distributions (First and Second kind), Weibull, Cauchy, and Laplace distribution functions-definition, properties, and applications.	CO4
	C	Lognormal, logistic, Pareto and Rayleigh distribution functions definition, properties and applications. Concept of truncated distributions.	CO4, CO6
	Unit 5	**Sampling Distributions**	
	A	Introduction, The sampling distribution of the Mean: Finite Populations, Sampling distribution of the proportion.	CO5
	B	t-distribution and F distribution, properties, applications, and procedure of hypothesis testing.	CO5
	C	Chi-square distribution and order statistics: properties, applications, and procedure of hypothesis testing.	CO5, CO6
	Mode of examination	**Theory**	
	Weightage Distribution	CA: 25 % ; MTE: 25 % ; ETE: 50 %	
	Text book/s*	1. Sheldon Ross; A First Course in Probability,Pearson, 2014. 2. Parimal Mukhopadhyay; An Introduction to theTheory of Probability, World scientific, 2012. 3. Irwin Miller, Marylee's Miller, John E. Freund's;Mathematical Statistics, Pearson, 2017	
	Other References	1. FetsjeBijma, Marianne Jonker and Aad van der Vaart; Introduction to Mathematical Statistics, Amsterdam University Press, 2018. 2. Krishnamoorthy, K., Handbook of Statistical Distributions with Applications, Chapman & Hall/CRC, 2006. 3. Rohatgi, V.K. and Ebsanes Saleh, A.K. Md., An introduction to Probability and Statistics, 2nd Ed.,John Wiley & Sons, 2002. 4. Shanmugam, R., Chattamvelli, R. Statistics forscientists and engineers, John Wiley, 2015.	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA103.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA103.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**

MDA103.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA103.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA103.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA103.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		**Batch: 2024-26**	
Program: M.Sc.		**Academic Year: 2024-25**	
Branch: Data Science & Analytics		**Semester: I**	
1	Course Code	**MDA104**	
2	Course Title	**Next Generation Databases**	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	Compulsory	
5	Course Objective	To explore the concepts of NoSQL Databases. To understand and use columnar and distributed database patterns.	
6	Course Outcomes	After completion of this course, students will be able to CO1: Develop and Explore the relationship between Big-Data andNoSQL databases. (K1, K2, K3) CO2: Formulate a fundamental relationship between Big-Data and NoSQL databases. (K2, K3) CO3: Describe various types of NoSQL databases to analyze the big data for useful business applications. (K3, K4) CO4: Derive and Work with NoSQL databases to analyze the big datafor useful business applications. (K4, K5) CO5: Discuss different data models to suit various data representations and storage needs. (K5, K6) CO6: Explain and correlate with different data models to suit variousdata representations and storage needs. (K5, K6)	
7	Course Description	To integrate the intrinsic ideas for the use of various Data models for a variety of databases.	
8			**CO Mapping**
	Unit 1		
	A	Database Revolutions- system Architecture-Relational Database. Database Design-Data Storage-Transaction Management.	CO1
	B	Data warehouse and Data Mining-Information Retrieval. Big-Data Revolution-CAP Theorem.	CO1
	C	Birth of NoSQL-Document Database—XML Databases. JSON Document Databases-Graph Databases.Probability and Random variables	CO1
	Unit 2		
	A	Big-Data Revolution-CAP Theorem.	CO2
	B	Birth of NoSQL-Document Database—XML Databases.	CO2
	C	JSON Document Databases-Graph Databases.	CO2
	Unit 3		

	A	ColumnDatabases-Data Warehousing Schemes- Columnar Alternative-Sybase IQ-C-Store.	CO3
	B	Vertica-Column Database Architectures-SSD and In-Memory Databases.	CO3
	C	In-Memory Databases-Berkeley Analytics Data Stack and Spark.	CO3, CO6
	Unit 4		
	A	Distributed Database Patterns-Distributed Relational Databases-Non- relational Distributed Databases.	CO4
	B	MongoDB Sharing and Replication-HBase-Cassandra- Consistency Models.	CO4
	C	Types of Consistency-Consistency MongoDB - HBase Consistency-Cassandra Consistency.	CO4, CO6
	Unit 5		
	A	Data Models and Storage-SQL-NoSQLAP Is-Return SQL-Advance Databases-Postgre SQL.	CO5
	B	Riak-CouchDB-NEO4J-Redis-Future, Databases- Revolution Revisited-Counter revolutionaries-Oracle HQ.	CO5
	C	Other Convergent Databases-Disruptive Database Technologies.	CO5, CO6
	Mode of examination	**Theory**	
	Weightage Distribution	CA 25 % / MTE 25 % / ETE 50 %	
	Text book/s*	1. Abraham Silberschatz, Henry F. Korth, S.Sudarshan, "Database System Concepts", Sixth Edition, McGraw Hill.	
	Other References	1. Guy Harrison, "Next Generation Databases",A Press, 2015. 2. Eric Redmond, Jim R Wilson, "Seven Databasesin Seven Weeks", LLC. 2012. 3. Dan Sullivan, "NoSQL for Mere Mortals",Addison-Wesley, 2015. 4. Adam Fowler, "NoSQL for Dummies", John	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA104.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA104.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA104.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA104.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA104.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA104.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		Batch: 2024-26	
Programme: M.Sc.		Academic Year: 2024-25	
Branch: Data Science & Analytics		Semester: I	
1	Course Code	**MDA151**	
2	Course Title	**Practical –I (Based on Paper MMT104, MDA102 Using Excel /SPSS /Minitab)**	
3	Credits	2	
4	Contact Hours (L-T-P)	0-0-4	
	Course Status	Compulsory	
5	Course Objective	Introduce basic concepts of Excel/SPSS/Minitab environment and provide students with a general understanding of Excel/SPSS/Minitab for solving the statistical-based problem. Equip students with the skills to apply Excel/SPSS/Minitab concepts and analytical tools to analyze statistical problems and handle real-world issues.	
6	Course Outcomes	CO1: Describe the overall process and particular steps in designing studies, collecting, analyzing data, and interpreting and presentingresults. (K1, K2, K3) CO2: Develop skills in presenting quantitative data using appropriate diagrams, tabulations, and summaries. (K2, K4) CO3: Test for various hypotheses of significance like means, proportions, independence of attributes, variance, etc. included in the theory. (K3, K4) CO4: Discuss and illustrate various discrete and continuousprobability distributions and study various real-life situations. (K4, K5) CO5: Identify the appropriate probability model that can be used. (K5, K6) CO6: Apply forecasting and data analysis techniques in the case of data sets. (K4, K5)	
7	Course Description	Introduce basic concepts of Excel/SPSS/Minitab environment and provide students with a general understanding of Excel/SPSS/Minitabfor solving the statistical-based problem. Equip students with the skills to apply Excel/SPSS/Minitab concepts and analytical tools to analyze statistical problems and handle real-world issues.	
8	Outline syllabus		CO Mapping
	Unit 1		
		Graphical representation of data by Histogram, Frequency polygons, frequency curves, and Ogives. Stem and Leaf Plot, Box Plot.	CO1
	Unit 2		
		Problems based on measures of central tendency. Problems based on measures of dispersion. Problems based on combined mean and variance and coefficient of variation. Problems based on moments, skewness, and kurtosis.	CO2
	Unit 3		
		Fitting of curves by the method of least squares. Determination of regression lines and calculation of correlation coefficient – grouped and ungrouped data. Calculation of multiple and partial correlation coefficients for three variables. Calculation of measures of association in contingency tables.	CO3
	Unit 4		
		Fitting of Binomial, Poisson, and Normal distributions toobserved data and testing of the goodness of fit.	CO4
	Unit 5		
		Analysis of variance in one-way and two-way classification (with and without interaction terms).	CO5, CO6
	Mode of examination	Practical	
	Weightage Distribution	CA: 25 % / CE: 25 % / ETE: 50 %	
	Text book	1. "Introductory Statistics with R" by Peter Dalgaard	

		2. Discovering Statistics Using IBM SPSS Statistics" by Andy Field	
	Other References		

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA151.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA151.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA151.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA151.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA151.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA151.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		Batch: 2024-26	
Programme: M.Sc.		Academic Year: 2024-25	
Branch: Data Science & Analytics		Semester: I	
1	Course Code	MDA152	
2	Course Title	**Practical –II (Based on Paper MMT104, MDA102, 103, 104UsingR/ Python)**	
3	Credits	2	
4	Contact Hours (L-T-P)	0-0-4	
	Course Status	Compulsory	
5	Course Objective	Introduce basic concepts of R/ Python environment and provide students with a genera skills to apply R/ Python concepts and analytical tools to analyze data analytics problem and handle real-world issues.	
6	Course Outcomes	CO1: Discuss and illustrate R/ Python environment. (K1,K2) CO2: Discuss and explain the importance of R/ Python workspace andworking directo CO3: Discuss, calculate and understands the Statistics and plot and interpret the graph CO4: Discuss probability distribution and testing of hypothesis through R / Python anc CO5: Discuss and Explain creating matrices and some simple matrix operations, Sub-r CO6: Develop a deeper understanding of the write R/ Python functions for Next Gene	
7	Course Description	Introduce basic concepts of R/ Python environment and provide students with a gener skills to apply R/ Python concepts and analytical tools to analyze data analytics problem and handle real-world issues.	
8	Outline syllabus		CO Mapping
	Unit 1		
		Use of basic R/ Python software commands c( ), scan( ), rep( ), seq ( ), min, max, sort, extract, data. frame, matrix, accessing resident data sets etc.	CO1
	Unit 2		
		Finding summary statistics using summary () and five num(). Calculate the arithmetic mean (AM),geometric mean (GM), harmonic mean (HM), median,mode, quantiles, range quartile deviation using R/Python.	CO2
	Unit 3		
		Computation of probabilities of negative binomial,multinomial, normal, exponential, gamma, , $\aleph^2$, using R/ Python.	CO3, CO6
	Unit 4		
		Creating matrices, some simple matrix operations,Sub-matrices also solve derivatives and some basic derivative function by using R/ Python.	CO4, CO6
	Unit 5		
		File operations, Reading Next Generation Databases,Data Structures.	CO5, CO6
	Mode of examination	Practical	
	Weightage Distribution	CA: 25 % · CE: 25 % · ETE: 50 %	
	Text book	1. "R for Data Science" by Hadley Wickham and Garrett Grolemu "Python for Data Analysis" by Wes McKinney	
	Other References	"The Elements of Statistical Learning: Data Mining, Inference, and Prec	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA152.1	3	2	2	3		3	3	2	1
MDA152.2	3	2	2	3		3	2	1	1
MDA152.3	3	2	2	3		2	2	1	1
MDA152.4	3	2	2	3		2	3	1	1
MDA152.5	3	2	2	3		3	2	2	1
MDA152.6	3	2	2	3		2	2	1	1

School: SSBSR		**Batch: 2024-26**	
Programme: M.Sc		**Academic Year: 2024-25**	
Branch: Data Science & Analytics		**Semester: I**	
1	Course Code	**RBL001**	
2	Course Title	**Research-Based Learning-1**	
3	Credits	0	
4	Contact Hours (L-T-P)	0-0-4	
	Course Status	Compulsory	
5	Course Objective	1. Deep knowledge of a specific area of specialization. 2. Develop communication skills, especially in project writing and oral presentation. Develop some time management skills.	
6	Course Outcomes	CO1: Explain the concept of research within the subject, as regards approaching a question, collecting and analyzing background material, and presenting research questions and conclusions. (K2, K4) CO2: Construct and develop a deeper interest in mathematics and a taste for research. (K5, K6) CO3: Select and recommend activities that support their professional goals. (K4, K6) CO4: Develop effective project organizational skills. (K5) CO5: Analyse the problem and summarize research findings. (K4, K5) CO6: Use research findings to develop education theory and practice. (K3, K6)	
7	Course Description	Maintain a core of mathematical and technical knowledge that is adaptable to changing technologies and provides a solid foundation for future learning.	
8	Outline syllabus		CO Achievement
	Unit 1	**Introduction**	**CO1**
	Unit 2	**Case study**	**CO1, CO2**
	Unit 3	**Conceptual**	**CO2, CO3**
	Unit 4	**Development**	**CO4, CO5**
	Unit 5	**Finalization**	**CO5, CO6**
	Mode of examination	Jury/Practical/Viva	
	Weightage Distribution	CA / ETE	
	Text book/s*	-	
	Other References		

School: SSBSR		Batch: 2024-26	
Programme: M.Sc.		**Academic Year: 2024-25**	
Branch: Mathematics		**Semester: II**	
1	Course Code	**MMT130**	
2	Course Title	**Numerical Analysis**	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	CC	
5	Course Objective	• To provide the student with numerical methods of solving the non-linear equations, interpolation, differentiation, and integration. • To improve the student's skills in numerical methods by using the MATLAB	
6	Course Outcomes	CO1: Estimate errors in numerical solution of a given problem. CO2: Find a root of transcendental equation. CO3: Solve a linear system of equations using iterative and factorization methods and discuss its convergence. CO4: Estimate numerical value of differentiation and integration using interpolation. CO5: Solve initial value problems numerically through single-step and multi-step methods. CO6: Apply finite difference technique for the solution of ordinary and partial differential equations.	
7	Course Description	This course is an introduction to the numerical analysis. The primary objective of the course is to develop the basic understanding of numerical algorithms and skills to implement algorithms to solve mathematical problems in MATLAB.	
8	Outline syllabus		CO Mapping
	Unit 1	**Error Analysis and solution of transcendental equations**	
	A	Definition and sources of errors, Propagation of errors, Sensitivity and conditioning, Stability and accuracy, Floating-point arithmetic and rounding errors.	CO1
	B	Intermediate value theorem, bisection method, method of false position, secant method, Newton Raphson method.	CO1, CO2
	C	Rate of convergence of iterative methods.	CO2
	Unit 2	**Solution of system of linear equations**	
	A	Iterative methods: Jacobi's method, Gauss-Seidal method	CO1, CO3
	B	Convergence criteria of iterative methods	CO3
	C	LU factorization methods: Crout, Choleski and Doolittle	CO3
	Unit 3	**Interpolation, differentiation and integration**	
	A	Finite difference operators, Newton Gregory forward and backward interpolation, Lagrange interpolation and Newton's divided difference interpolation	CO1, CO4
	B	Derivative formulae based on interpolating polynomial, Newton-Cotes quadrature formula	CO4
	C	Trapezoidal rule, Simpson's 1/3rd and 3/8th rules, Gauss quadrature formula.	CO1, CO4
	Unit 4	**Solution of ordinary differential equations**	
	A	Single-step methods: General definitions and Lipschitz condition, Derivations and stability analysis for Taylor series method	CO5
	B	Euler's method and its variants, Runge- Kutta second order and fourth order methods	CO1, CO5
	C	Solution of boundary value problems by finite difference technique.	CO1, CO6
	Unit 5	**Solution of Partial Differential Equations**	
	A	Finite difference approximations of partial derivatives	CO6

	B	Standard five-point and diagonal five-point formulae, solution of elliptic equations (Laplace and Poisson's equations) by Liebmann's iteration technique			CO1, CO6
	C	Solution of parabolic equation (one dimensional heat equation) by Bender-Schmidt and Crank Nicolson's methods, solution of hyperbolic equation (wave equation)			CO6
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25%	25%	50%	
	Text book/s*	1) M.K. Jain, S.R.K. Iyengar and R.K. Jain, Numerical Methods for Scientific and Engineering Computation, New Age International (P) Ltd., Publishers, 6 ed, 2012. 2) S.S. Sastry, Introductory Methods of Numerical Analysis, PHI Learning Pvt., Ltd., 5 ed, 2018. 3) C. F. Gerald and Patrick O. Wheatley, Applied Numerical Analysis, Pearson Education, 2006.			
	Other References	1) E. Kreyszig, Advanced Engineering Mathematics, Wiley Publications, 10 ed. 2) Steven C. Chapra and Raymond P. Canale, Numerical Methods for Engineers, Tata McGraw Hill Education Pvt., Ltd., 5 ed, 2007.			

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MMT130.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MMT130.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MMT130.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MMT130.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MMT130.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MMT130.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		**Batch: 2024-26**	
Program: M.Sc.		**Academic Year: 2024-25**	
Branch: Data Science & Analytics		**Semester: II**	
1	Course Code	**MDA105**	
2	Course Title	**Regression Analytics and Predictive Models**	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	Compulsory	
5	Course Objective	The main objective of this course is to demonstrate and intended to verse students in the techniques necessary to understand and carry out regression and predictive analysis.	
6	Course Outcomes	At the end of the course, the student should be able to CO1: Explain the concept of regression with two and multiple variables. CO2: Testing of the single and subset of the regression coefficient. CO3: Explain the concept of multicollinearity. CO4: Describe how to overcome the problem of heteroscedasticity and autocorrelation. CO5: Explain the concept of dummy variables. CO6: How to apply logistic regression on a dataset.	
7	Course Description	A PG-level course in regression analysis, intended to verse students in the techniques necessary to understand and carry out methods of research in serial analysis. Lectures study the large-sample properties of estimators based on one-sample, k-sample, and partial likelihood inference, with proofs based on the counting process and Martingale theory. The theory of competing risks is studied from several angles. Many extensions of the Cox model to more complex data structures are considered.	
8	Outline syllabus		CO Mapping
	Unit 1		
	A	Simple Linear Regression: Simple linear regression model. Least-squares estimation of parameters. Hypothesis testing on the slope and intercept. Interval estimation in simple linear regression.	CO1
	B	Prediction of new observations. Coefficient of determination. Estimation by maximum likelihood.	CO1
	C	Multiple linear regression: Multiple linear regression models. Estimation of the model parameters. Hypothesis testing in multiple linear regression. Confidence intervals in multiple regression. Coefficient of determination and Adjusted R2.	CO1
	Unit 2		
	A	Logistic Regression: Introduction, Linear predictor and link functions, logit, probit, odds ratio, the test of hypothesis. Discriminant Analysis.	CO2
	B	Model Adequacy: Checking of linearity between study and explanatory variable, Residual Analysis, Detection and treatment of outliers, Residual plots.	CO2
	C	The PRESS statistic. Outlier test based on Studentized Residual (R-student). Test for lack of fit of the regression model.	CO2
	Unit 3		
	A	Data Understanding and Preparation Introduction, Reading data from various sources, Data visualization, Distributions, and summary statistics, Relationships among variables	CO3
	B	The extent of Missing Data. Segmentation, Outlier detection, Automated Data Preparation	CO3

	C	Combining data files, Aggregate Data, Duplicate Removal, Sampling DATA, Data Caching, Partitioning data, and Missing Values.	CO3
	Unit 4		
	A	Model development & techniques Data Partitioning, Model selection, Model Development Techniques	CO4
	B	Neural networks, Decision trees, Logistic regression, Discriminant analysis, Support vector machine	CO4
	C	Bayesian Networks, Linear Regression, Cox Regression, and Association rules.	CO4
	Unit 5		
	A	Model Evaluation and Deployment Introduction, Model Validation, Rule Induction Using CHAID	CO5
	B	Automating Models for Categorical and Continuous targets, Comparing and Combining Models, and Evaluation Charts for Model Comparison	CO5
	C	Meta Level Modeling, Deploying Model, Assessing Model Performance, Updating a Model.	CO5, CO6
	Mode of examination	Theory	
	Weightage Distribution	CA: 25%, MTE: 25%, ETE: 50%	
	Text book/s*	1. Johnston, J. (1984). Econometric Methods, McGraw Hill Kogakusha Ltd.	
	Other References	1. Draper, N. R., and Smith, H. (1998). Applied Regression Analysis (John Wiley) Third edition.	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA105.1	3	2	2	3		3	3	2	1
MDA105.2	3	2	2	3		3	2	1	1
MDA105.3	3	2	2	3		2	2	1	1
MDA105.4	3	2	2	3		2	3	1	1
MDA105.5	3	2	2	3		3	2	2	1
MDA105.6	3	2	2	3		2	2	1	1

School: SSBSR		**Batch: 2024-26**
Programme: M.Sc. (Hons.)		**Academic Year: 2025-26**
Branch: Data Science & Analytics		**Semester: II**
1	Course Code	**MDA 118**
2	Course Title	**Survey Sampling**
3	Credits	4
4	Contact Hours (L-T-P)	4-0-0
	Course Status	DSE
5	Course Objective	To introduce the subject of Survey Sampling so that the students as data scientists are equipped to apply the theory and methods of survey sampling in practice for planning and conduct of sample surveys in various fields including market research, opinion poll, agriculture, business, and industry.
6	Course Outcome	**CO1**: Learn basic concepts of survey sampling and to learn how to draw a sample by SRSWR and SRSWOR and Estimation of Population parameters viz., total mean and proportion along with determination of sample size. (K1,K2,K3) **CO2**: Competent to draw sample by systematic and unequal probability sampling, corresponding estimation of population parameters and with standard errors along with comparison with SRSWR and SRSWOR. (K2,K3,K4) **CO3**: How to estimate population total / mean and proportion from stratified designs using SRSWR, SRSWOR, PPSWR Design and allocation of sample size to strata and also construction of strata under various allocations for different sampling designs. Comparison with unstratified sampling and also appraisal of gain due to stratification over unstratified sampling from stratified sample. (K3,K4,K5, K6) **CO4**: Use ratio and regression methods of estimation for estimation of population total / mean along with estimation if their bias, MSE etc., and their comparison with unbiased estimator for SRSWOR. (K3,K4,K5) **CO5**: Cluster sampling with equal and unequal sized clusters using various estimators and their comparison with mean per element. Learn two-stage sampling with equal and unequal sized FSUs using various sampling designs at the two-stages. Estimation of population parameters using different estimators and their comparison. (K3,K5, K6) **CO6**: Estimation in double sampling for stratification and for ratio-estimators. Sources of non-sampling errors and methods of their handling including randomized response technique. (K1,K5, K6)
7	Course Description	The course will provide first hand training in the design and analysis of sample surveys using various sampling designs beginning with simple random sampling with and without replacement designs (SRSWR & SRSWOR), probability proportional to size with and without replacement (PPSWR & PPSWOR) sampling designs, linear and circular systematic sampling (LSS & CSS) designs, stratified designs using SRSWR, SRSWOR and PPSWR Schemes, cluster, two-stage, and double sampling designs. Different estimation methods will be covered including Estimator based on distinct units in SRSWR, Ratio, Difference and Regression estimators in various sampling designs. Various practical aspects in the planning of sample surveys viz., determination of sample size, allocation of total sample size to strata, construction of strata, cluster size etc., will be given adequate coverage. Finally, students will be acquainted with various sources of non-sampling errors in surveys viz., types of non-response and item non-response

		along with methods of handling them including randomized response technique.	
8			
	Unit 1		
	A	**Basics of Survey Sampling** Types of data. Survey data. Complete Enumeration Survey (CES/Census) vs. Sample Survey. Need for sampling. Types of units. Sampling unit and sampling frame. Probability Sampling and alternatives. Sampling, non-sampling errors. Role of Sampling Theory. Principal steps in a sample survey. Sampling from finite population, unbiased and Consistent estimators. Measures of error-Mean Square Error (MSE), variance and standard error. Accuracy and precision. sampling and cost efficiency. Simple Random Sampling (SRS) with replacement (SRSWR) and without replacement (SRSWOR)-description and methods of sample selection.	CO1
	B	Estimation of population mean and total in SRSWR and SRSWOR. Sampling variance and comparison. Variance and standard error estimation. Estimation of population proportion in SRSWR and SRSWOR designs.	CO1
	C	Confidence intervals for mean, total and proportion. Estimator based on distinct units in SRSWR. Comparison with SRSWOR. Estimation of sample size.	CO2
	Unit 2		
	A	Estimates of the population mean, total, and proportion,	CO3
	B	Variances of these estimates	CO3
	C	Estimates of theses variances and sample size determination.	CO3
	Unit 3		
	A	Stratified random sampling, estimates of the population mean, and total variances of these estimates.	CO4
	B	Proportional and optimum allocations and their comparison with SRS.	CO4
	C	Systematic Sampling, estimates of the population mean and total, variances of these estimates.	CO4
	Unit 4		
	A	Ratio and regression methods of estimation, estimates of the population mean and total (for SRS of large size),	CO5
	B	Variances of these estimates and estimates of theses variances,	CO5
	C	Variances in terms of the correlation coefficient between X and Y for regression method and their comparison with SRS.	CO5
	Unit 5		
	A	Present official statistical system in India, Methods of collection of official statistics, their reliability and limitations.	CO6
	B	Principal publications containing data on the topics such as population, industry, and finance.	CO6
	C	Various official agencies are responsible for data collection and their main functions.	CO6
	Mode of examination	Theory	

	Weightage Distribution	CA:25%; MSE:25% ESE:50%	
	Text book/s*	1. Murthy M.N. (1977): Sampling Theory & Statistical Methods, Statistical Pub. Society, Calcutta 2. Cochran W.G (1984): Sampling Techniques (3rd Ed.), Wiley Eastern.	
	Other References	1. Sukhatme P.V., Sukhatme B.V, Sukhatme. S, Asok: Sampling Theory of Surveys with Applications, Publication: Indian Society of Agricultural Statistics, New Delhi. (3rd Ed.1984) 2. Raj Des & Chandhok Promod: Sample Survey Theory, Narosa Publishing House, New Delhi, 1999. 3. Mukhopadhyay P. (1998): Theory and Methods of Survey Sampling, Prentice Hall 4. Singh, D. and Chaudhari, F.S. (1986): Theory and Analysis of Sample Survey Designs. New Age International Publishers, New Delhi.	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA118.1	3	2	2	3		3	3	2	1
MDA118.2	3	2	2	3		3	2	1	1
MDA118.3	3	2	2	3		2	2	1	1
MDA118.4	3	2	2	3		2	3	1	1
MDA118.5	3	2	2	3		3	2	2	1
MDA118.6	3	2	2	3		2	2	1	1

School: SSBSR		**Batch: 2024-26**	
Program: M.Sc.		**Academic Year: 2024-25**	
Branch: Data Science & Analytics		**Semester: II**	
1	Course Code	**MDA107**	
2	Course Title	**Advanced Big Data and Text Analytics**	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	Compulsory	
5	Course Objective	This course aims to provide insight into the concepts of Natural Language Processing and its applications. This course helps the students to implement NLP applications using deep learning algorithms. This course helps to understand various word/text representation algorithms.	
6	Course Outcomes	At the end of the course, the student should be able to CO1: Learn about Big data techniques and their applications. CO2: Analyse various neural network problems. CO3: Use different word/text representation methods to see how words are related to each other. CO4: Model different NLP applications using Machine Learning/Deep learning algorithms CO5: Implement different deep learning models to solve real-time NLP problems CO6: Provide a body of concepts and techniques for designing intelligent systems.	
7	Course Description	A PG-level course in Soft Computing Techniques to Improve Big Data Analysis solutions is to strengthen the dialogue between the statistics and soft computing research communities.	
8	Outline syllabus		CO Mapping
	Unit 1		
	A	Introduction to Big Data: Introduction to Big Data, Big Data characteristics	CO1
	B	Types of Big Data, Structured Data, Unstructured Data, and semi Structured Data.	CO1
	C	Traditional vs. Big Data business approach, Case Study of Big Data Solutions.	CO1
	Unit 2		
	A	Mining Data Streams: The Stream Data Model: A Data Stream-Management System, Examples of Stream Sources, Stream Queries, Issues in Stream Processing.	CO2
	B	Sampling Data in a Stream: Obtaining a Representative Sample, The General Sampling Problem, Varying the Sample Size. Filtering Streams: The Bloom Filter Analysis.	CO2
	C	Counting Distinct Elements in a Stream: The Count-Distinct Problem, The Flajolet-Martin Algorithm, Combining Estimates, Space Requirements Counting Ones in a Window: The Cost of Exact Counts.	CO2
	Unit 3		
	A	The Big Data Analytics and Big Data Analytics Techniques: Big Data and its Importance, Drivers for Big data, Optimization techniques, Dimensionality Reduction techniques.	CO3
	B	Time series Forecasting, Social Media Mining, and Social Network Analysis, and its Application.	CO3
	C	Big Data analysis using Hadoop, Pig, Hive, MongoDB, Spark, and Mahout, Data analysis techniques like Discriminant Analysis and Cluster Analysis.	CO3
	Unit 4		
	A	Introduction to Natural Language Processing Words Regular Expressions N-grams Language modeling Part of Speech.	CO4
	B	Tagging Named Entity Recognition Syntactic and Semantic Parsing-Morphological Analysis	CO4
	C	Text Representation and Transformation-Vector space models Bag of Words Term Frequency Inverse Document Frequency Word Vector representations: Word2vec, GloVe, FastText, BERT-Topic Modelling	CO4
	Unit 5		

	A	Neural language models - Recurrent Neural Network - Long Short-Term Memory Networks	CO5
	B	Encoder decoder architecture - Attention Mechanism - Transformer networks	CO6
	C	Text classification-Sentiment Analysis-Neural Machine Translation - Question answering - Text summarization	CO6
	Mode of examination	Theory	
	Weightage Distribution	CA: 25%, MTE: 25%, ETE: 50%	
	Text book/s*	1. S.N. Sivanandam& S.N. Deepa, Principles of Soft Computing, Wiley Publications, 2nd Edition, 2011. 2.S, Rajasekaran& G.A. VijayalakshmiPai, Neural Networks, 3. Fuzzy Logic & Genetic Algorithms, Synthesis & applications, PHI Publication, 1st Edition, 2009.	
	Other References	1.N. K. Bose, Ping Liang, Neural Network fundamental with Graph, Algorithms & Applications, TMH, 1st Edition, 1998. 2. Rich E, Knight K, Artificial Intelligence, TMH, 3rd Edition, 2012. 3. Martin T Hagen, Neural Network Design, Nelson Candad, 2nd Edition, 2008.	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA107.1	3	2	2	3		3	3	2	1
MDA107.2	3	2	2	3		3	2	1	1
MDA107.3	3	2	2	3		2	2	1	1
MDA107.4	3	2	2	3		2	3	1	1
MDA107.5	3	2	2	3		3	2	2	1
MDA107.6	3	2	2	3		2	2	1	1

School: SSBSR		Batch: 2024-26	
Program: M.Sc.		**Academic Year: 2024-25**	
Branch: Data Science & Analytics		**Semester: II**	
1	Course Code	MDA108	
2	Course Title	Data Mining & Artificial Intelligence	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	Compulsory	
5	Course Objective	To introduce students to the applications, concepts, and techniques of data mining. To provide a strong foundation of fundamental concepts in Artificial Intelligence.	
6	Course Outcomes	CO1: Learn about the data mining pattern and functionalities CO2: Understand the basic concepts and classification of Data mining CO3: Explain the mining of frequency pattern CO4: Explain the correlation and cluster analysis with applications. CO5: Learn about the basic concept of AI CO6:Explain computable functions, predicates, forward and backward reasoning	
7	Course Description	The data mining process includes data selection and cleaning, machine learning techniques to ``learn" knowledge that is ``hidden" in data, and the reporting and visualization of the resulting knowledge. AI helps the students to understand various searching techniques, constraint satisfaction problems, and example problems- game playing techniques.	
8	Outline syllabus		CO Mapping
	Unit 1	Data Mining	
	A	Introduction, Data, Types of Data, Data Mining Functionalities,	CO1,
	B	Interestingness of Patterns, Classification of Data Mining Systems, Data Mining Task Primitives,	CO1,
	C	Integration of a Data Mining System with Data Warehouse Issues, Data Preprocessing	CO1,
	Unit 2	Mining Frequent Pattern	
	A	Mining Frequent Patterns, Associations, and Correlations, Mining Methods, Mining various Kinds of Association Rules,	CO2
	B	Correlation Analysis, Constraint-Based Association Mining Classification, and Prediction, Basic Concepts, Decision Tree Induction, Bayesian Classification, Rule Based Classification,	CO2
	C	Classification by Back propagation, Support Vector Machines, Associative Classification, Lazy Learners, Other Classification Methods, and Prediction.	CO3
	Unit 3	Cluster Analysis	
	A	Cluster Analysis, Types of Data, Categorization of Major Clustering Methods, K-means, Partitioning Methods, Hierarchical Methods,	CO4
	B	Density-Based Methods, Grid-Based Methods, Model-Based Clustering Methods, Clustering High Dimensional Data, Constraint, Based Cluster Analysis, and Outlier Analysis.	CO4
	C	Data Mining Applications. Apply data mining techniques and methods to large data sets, Use data mining tools, and Compare and contrast the various classifiers.	CO4
	Unit 4	Basic of AI	
	A	Defining Artificial Intelligence, Defining AI techniques,	CO5
	B	Defining problems such as State Space search, Production systems, and characteristics,	CO5
	C	Hill Climbing, Breadth first and depth first search, Best first search.	CO5
	Unit 5	Mapping in AI	
	A	Representations and Mappings, Approaches to knowledge representation, Representing simple facts in logic,	CO6
	B	Computable functions and predicates, Procedural vs Declarative knowledge, Logic Programming,	CO6
	C	Forward vs backward reasoning, Non-monotonic Reasoning, Logic for non-	CO6

		monotonic reasoning.			
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25%	25%	50%	
	Text book/s*	1. Alex Berson and Stephen J. Smith, "Data Warehousing, Data Mining and OLAP", Tata McGraw – Hill Edition, Thirteenth Reprint 2008. 2. Jiawei Han and Micheline Kamber, "Data Mining Concepts and Techniques", Third Edition, Elsevier, 2012. 3. Artificial Intelligence: A Modern Approach, Stuart Russel, Peter Norvig			
	Other References	1. Artificial Intelligence, 2nd Edition, Rich and Knight. 2. . K.P. Soman, ShyamDiwakar and V. Aja, "Insight into Data Mining Theory and Practice", Eastern Economy Edition, Prentice Hall of India, 2006.			

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA108.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA108.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA108.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA108.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA108.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA108.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		**Batch: 2024-26**			
Program: M.Sc.		**Academic Year: 2024-25**			
Branch: Data Science & Analytics		**Semester: II**			
1	Course Code	**MDA 153**			
2	Course Title	Practical -III (based on MDA 105, MDA 106 MDA 107 using R/SPSS/SAS/Python)			
3	Credits	2			
4	Contact Hours (L-T-P)	0-0-4			
	Course Status	Compulsory			
5	Course Objective	After studying these courses students will be able to understand how to calculate the power of the test, analyze the multivariate data and understand the characteristics of multivariate quantitative research, including strengths and weaknesses. It also discusses the principles and characteristics of the multivariate data analysis techniques.			
6	Course Outcomes	At the end of the course, the student should be able to CO1: Estimate the parameter by MLE CO2: Learn about how to calculate the Rao, Lehman, and Bhattacharya bounds CO3: Learn how to calculate the critical region, power of the test, unbiased test, and Neyman structure. CO4: Understand the basic concepts of multivariate normal distribution. CO5: Calculate Wishart distribution in the multivariate analysis also know how to find Mahalanobis D2 and HottelingT2. CO6: Apply the classification rule, PCA, and factor analysis.			
7	Course Description	In this course, students are concerned with making inferences based on relations found in the sample, to relations in the population. Also multivariate analysis of data deals with examining the interrelationship between three or more equally important variables or explaining variation in, usually one (or more than one) dependent variable(s) based on two or more independent (explaining) variables.			
8	Outline syllabus				CO Mapping
	Unit 1	Multiple regression analysis			
		Problem-based on Multiple regression analysis SPSS/SAS/STRATA/R/Python.			CO1 CO2
	Unit 2	Logistic regression analysis			
		Problem-based on Logistic regression analysis SPSS/SAS/STRATA/R/Python.			CO2, CO3
	Unit 3	Discriminant Analysis			
		Problem-based on Discriminant Analysis using SPSS/SAS/STRATA/R/Python.			CO3, CO4
	Unit 4	Principal Component Analysis			
		Problem-based on classification rule, PCA, and factor analysis using SPSS/SAS/STRATA/R/Python.			CO4,CO5
	Unit 5	Big Data Platform			
		Problem-based on Set up Hadoop Environment, Map Reduce Task using Hadoop			CO5, CO6
	Mode of examination	Practical			
	Weightage	CA	CE	ETE	

	Distribution	25%	25%	50%	
	Text book/s*				
	Other References				

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA153.1	3	2	2	3		3	3	2	1
MDA153.2	3	2	2	3		3	2	1	1
MDA153.3	3	2	2	3		2	2	1	1
MDA153.4	3	2	2	3		2	3	1	1
MDA153.5	3	2	2	3		3	2	2	1
MDA153.6	3	2	2	3		2	2	1	1

School: SSBSR		Batch: 2024-26			
Program: M.Sc.		**Academic Year: 2024-25**			
Branch: Data Science & Analytics		**Semester: II**			
1	Course Code	**MDA 154**			
2	Course Title	Practical-IV (using based on MDA 108, using R/ Python)			
3	Credits	2			
4	Contact Hours (L-T-P)	0-0-4			
	Course Status	Compulsory			
5	Course Objective	The objective of the course is to introduce basic fundamental concepts in Artificial Intelligence (AI), with a practical approach to understanding them. To visualize the scope of AI and its role in futuristic development.			
6	Course Outcomes	After the completion of this course, students will be able to: CO1: Relate the goals of Artificial Intelligence and AI and non-AI solutions. CO2: Analyze various AI uninformed and informed search algorithms. CO3: Extend knowledge representation, reasoning, and theorem proving techniques to real-world problems CO4: Make use: Machine learning algorithms in various application domains of AI. CO5: Select Artificial Intelligent based applications. CO6: Develop independent (or in a small group) research and communicate it effectively.			
7	Course Description	In this course, students will learn a basic introduction to Artificial Intelligence, problem-solving agents, reasoning, learning, and applications of artificial intelligence.			
8	Outline syllabus				CO Mapping
	Unit 1	Practical based on Data Mining			CO1
	A	Association Rule: Apriori Algorithm			
	B	Correlation Analysis			
	C	Practice on Real time dataset (Kaggle, Open Data)			
	Unit 2	Practical based on Packages			CO2
	A	Basic of Numpy and Pandas			
	B	Basic of Scikit Learn			
	C	Basic of Tensorflow/Keras			
	Unit 3	Practical based on Classification and Clustering			CO3, CO6
	A	Classification: Decision Tree, Baye's Classifier, KNN			
	B	Clustering: K Mean, SVM			
	C	Hybrid: Random Forest			
	Unit 4	Practical based on Pre Processing and Model Selection			CO4, CO6
	A	Pre Processing: Creating Pipeline			
	B	Standarization and Normalization			
	C	Model Building, Selection and Model Accuracy			
	Unit 5	Practical based on Neural Network			CO5, CO6
	A	CNN			
	B	RNN			
	C	Boosting Algorithm: XGBoost, AdaBoost			
	Mode of examination	Practical/CE			
	Weightage	CA	CE	ETE	

	Distribution	25%	25%	50%	
	Text book/s*	1. Rich E& Knight K, Artificial Intelligence, Tata McGraw Hill, Edition 3.			
	Other References	1. Russell S &Norvig P, *Artificial Intelligence: A Modern Approach*, Prentice Hall.			

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA154.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA154.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA154.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA154.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA154.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA154.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		**Batch: 2024-26**	
Programme: M.Sc.		**Academic Year: 2024-25**	
Branch: Data Science & Analytics		**Semester: II**	
1	Course Code	**RBL002**	
2	Course Title	**Research-Based Learning-2**	
3	Credits	0	
4	Contact Hours (L-T-P)	0-0-4	
	Course Status	Compulsory	
5	Course Objective	1. Deep knowledge of a specific area of specialization. 2. Develop communication skills, especially in project writing and oral presentation. Develop some time management skills.	
6	Course Outcomes	CO1: Explain the concept of research within the subject, as regards approaching a question, collecting and analyzing background material, and presenting research questions and conclusions. (K2, K4) CO2: Construct and develop a deeper interest in mathematics and a taste for research. (K5, K6) CO3: Select and recommend activities that support their professional goals. (K4, K6) CO4: Develop effective project organizational skills. (K5) CO5: Analyse the problem and summarize research findings. (K4, K5) CO6: Use research findings to develop education theory and practice. (K3, K6)	
7	Course Description	Maintain a core of mathematical and technical knowledge that is adaptable to changing technologies and provides a solid foundation for future learning.	
8	Outline syllabus		CO Achievement
	Unit 1	**Introduction**	CO1
	Unit 2	**Case study**	CO1,CO2
	Unit 3	**Conceptual**	CO2,CO3
	Unit 4	**Development**	CO4,CO5
	Unit 5	**Finalisation**	CO5,CO6
	Mode of examination	Jury/Practical/Viva	
	Weightage Distribution	CA	ETE	
	Text book/s*	-	
	Other References		

School:		**School of Basic Sciences & Research**	
Department		**Department of Mathematics**	
Program:		M. Sc. (Mathematics & Data Science	
Branch: 2024-28 onwards			
1	Course Code	**NV10101**	
2	Course Title	Prompt Engineering	
3	Credits	Audit Course (Zero Credits)	
4	Contact Hours (L-T-P)	0-0-4	
	Course Status	VAC	
5	Course Objective	To provide undergraduate and postgraduate students with a comprehensive introduction to the fundamental concepts and practical skills required for prompt engineering, covering essential techniques for crafting effective prompts, optimizing their performance, and understanding their applications across various AI domains.	
6	Course Outcomes	CO1. Demonstrate proficiency in understanding and crafting various types of prompts for AI applications. CO2. Apply techniques for creating contextually aware and adaptive prompts to enhance AI model performance. CO3. Utilize prompt engineering for data processing tasks such as extraction, summarization, and transformation. CO4. Implement fine-tuning and evaluation methods to optimize prompt performance and iteratively improve their effectiveness. CO5. Design and deploy prompts for specific applications like creative writing, customer support, and code generation. CO6. Analyze and discuss the ethical implications of prompt engineering, including bias detection and mitigation, and the responsible use of AI.	
7	Course Description	This course introduces students to the essential concepts and practical skills of prompt engineering, focusing on creating effective prompts for AI models. It covers the basics of prompt design, advanced techniques, practical applications, and ethical considerations. Through hands-on exercises and real-world examples, students will learn to craft, optimize, and deploy prompts across various AI domains, making prompt engineering accessible and applicable to their academic and professional pursuits.	
8	Outline syllabus		CO Mapping
	Unit 1	**Introduction to Prompt Engineering**	
	A	Overview of prompt engineering, significance and	CO1

		applications in AI			
	B	Basics of prompts: structure, components, and types			CO1
	C	Techniques for crafting clear and effective prompts			CO1, CO4
	Unit 2	**Advanced Prompting Techniques**			
	A	Contextual prompts: incorporating context to enhance performance			CO2
	B	Dynamic and adaptive prompts: creating flexible and responsive prompts			CO2
	C	Evaluation and iteration: methods for evaluating and improving prompts			CO4
	Unit 3	**Practical Applications and Ethical Considerations**			
	A	Using prompts in real-world scenarios: case studies and hands-on exercises			CO4
	B	Ethical considerations in prompt engineering: bias detection and mitigation			CO6
	C	Special applications: creative writing, customer support, and code generation			CO5
	Mode of examination	Practical			
	Weightage Distribution	CA		ETE	
		25%		75%	
	Text book/s*				
	Other References				

<table>
<tr><td colspan="2">SCHOOL:
School of Basic Sciences and Research</td><td>TEACHING DEPARTMENT:
Community Connect</td><td>Academic Year: 2024-25</td><td>FOR STUDENTS M.Sc.
Batch: 2024-26</td></tr>
<tr><td>1</td><td>Course Number</td><td colspan="3">Course Code: CCU401/ Course ID: 30804</td></tr>
<tr><td>2</td><td>Course Title</td><td colspan="3">Community Connect</td></tr>
<tr><td>3</td><td>Credits</td><td colspan="3">2</td></tr>
<tr><td>3.01</td><td>(L-T-P)</td><td colspan="3">(0-0-2)</td></tr>
<tr><td>4</td><td>Learning Hours</td><td colspan="3"><table><tr><td>Contact Hours</td><td>30</td></tr><tr><td>Project/Field Work</td><td>20</td></tr><tr><td>Assessment</td><td>00</td></tr><tr><td>Guided Study</td><td>10</td></tr><tr><td>Total hours</td><td>60</td></tr></table></td></tr>
<tr><td>5</td><td>Course Objectives</td><td colspan="3">1. To expose our students to different social issues faced by people in different sections of society.
2. To connect their classroom learning with problem-solving skills in real-life scenarios.</td></tr>
<tr><td>6</td><td>Course Outcomes</td><td colspan="3">After completion of this course, students will be able to:
CO1. Recognize social problems prevailing in different sections of society and find the solution sustainably.
CO2. Get practical exposure to all-round development which complements their classroom learning
CO3. These activities will add value to students, faculty members, the school, and the university.
CO4. Apply their knowledge via research, and training for community benefit.
CO5. Analyze work on socio-economic projects with teamwork and timely delivery.
CO6. A survey will help to identify the gaps and create a plan to further improve the situation related to social problems prevailing in different sections of society and find the solution sustainably.</td></tr>
<tr><td>7</td><td>Theme</td><td colspan="3">Major research themes:

1. Survey and self-learning: In this mode, students will make a survey, analyze data, and will extract results to correlate with their theoretical knowledge. E.g. Crops and animals, land holding, labor problems, medical problems of animals and humans, savage and sanitation situations, waste management, etc.
2. Survey and solution providing: In this mode, students will identify the common problems and will provide solutions/ educate the rural population. E.g. air and water pollution, need for treatment, use of renewable (mainly solar) energy, electricity saving devices, inefficiencies in the cropping systems, animal husbandry, poultry, pest control, irrigation, machining in agriculture, etc.
3. Survey and reporting: In this mode, students will educate villagers and survey the ground-level status of various government schemes meant for rural development. The analyzed results will be reported to concerned agencies which will help them for taking necessary/corrective measures. E.g. Pradhan Mantri Jan Dhan Yojana, Pradhan Mantri MUDRA Yojana, Pradhan Mantri Jeevan Jyoti Bima Yojana, Atal pension Yojana, Pradhan Mantri Awas Yojana, Pradhan Mantri FasalBima Yojana, Swachh Bharat Abhiyan, Soil Health Card Scheme, Digital India, Skill India Programme, Beti Bachao, Beti Padhao Yojana, DeenDayal Upadhyaya Gram Jyoti Yojana, Shyama Prasad Mukherjee Rurban Mission, UJWAL Discom Assurance Yojana, PAHAL, Pradhan Mantri</td></tr>
</table>

<table>
<tr><td></td><td></td><td>Awas Yojana-Gramin, Pradhan Mantri Yuva Yojana, Pradhan Mantri Jan Aushadhi Yojana, Pradhan Mantri KhanijKshetra Kalyan Yojana, Pradhan Mantri Suraksha Bima Yojana, UDAN scheme, DeenDayal Upadhyaya Grameen Kaushalya Yojana, Pradhan Mantri Sukanya Samriddhi Yojana, Sansad Adarsh Gram Yojana, Pradhan Mantri SurakshitMatritva Abhiyan, Pradhan Mantri RojgarProtsahan Yojana, Midday Meal Scheme, Pradhan Mantri Vaya Vandana Yojana, Pradhan Mantri Matritva Vandana Yojana, and Ayushman Bharat Yojana.</td></tr>
<tr><td>8.1</td><td>Guidelines for Faculty Members</td><td>It will be a group assignment.
There should be no more than 10 students in each group.
The faculty guide will guide the students and approve the project title and help the student in preparing the questionnaire and final report.
The questionnaire should be well-designed and it should carry at least 20 questions (Including demographic questions).
The faculty will guide the student to prepare the PPT.
The topic of the research should be related to social, economical, or environmental issues concerning the common man.
The report should contain 2,500 to 3,000 words and relevant charts, tables, and photographs.
The student should submit the report to CCC-Coordinator signed by the faculty guide by 15 April 2019.
The students have to send the hard copy of the report and PPT, and then only they will be allowed for ETE.</td></tr>
<tr><td>8.2</td><td>Role of CCC-Coordinator</td><td>The CCC Coordinator will supervise the whole process and assign students to faculty members.
1. PG-M.Sc.-Semester II – the students will be allocated to the faculty members (mentors/faculty members) in an even term.
2. UG- B.Sc.-Semester III - the students will be allocated to the faculty members (mentors/faculty members) in the odd terms.</td></tr>
<tr><td>8.3</td><td>The layout of the Report</td><td>Abstract (250 words)
a. Introduction
b. Literature review(optional)
c. The objective of the research
d. Research Methodology
e. Finding and discussion
f. Conclusion and recommendation
g. References
Note: The research report should base on primary data.</td></tr>
<tr><td>8.4</td><td>Guideline for Report Writing</td><td>Title Page: The following elements must be included:
• Title of the article;
• Name(s) and initial(s) of the author(s), preferably with first names spelled out;
• Affiliation(s) of author(s);
• Name of the faculty guide and Co-guide
Abstract: Each article is to be preceded by a succinct abstract, of up to 250 words, that highlights the objectives, methods, results, and conclusions of the paper.
Text: Manuscripts should be submitted in Word.</td></tr>
</table>

<table>
<tr><td></td><td></td><td>

Use a normal, plain font (e.g., 12-point Times Roman) for text.
Use italics for emphasis.
Use the automatic page numbering function to number the pages.
Save your file in Docx format (Word 2007 or higher) or doc format (older Word versions)

Reference list:
The list of references should only include works that are cited in the text and that have been published or accepted for publication.
The entries in the list should be in alphabetical order.
Journal article
Hamburger, C.: Quasimonotonicity, regularity, and duality for nonlinear systems of partial differential equations. Ann. Mat. Pura Appl. 169, 321–354 (1995)
Article by DOI
Sajti, C.L., Georgio, S., Khodorkovsky, V., Marine, W.: New nanohybrid materials for biophotonics. Appl. Phys. A (2007). doi:10.1007/s00339-007-4137-z
Book
Geddes, K.O., Czapor, S.R., Labahn, G.: Algorithms for Computer Algebra. Kluwer, Boston (1992)
Book chapter
Broy, M.: Software engineering — from auxiliary to key technologies. In: Broy, M., Denert, E. (eds.) Software Pioneers, pp. 10–13. Springer, Heidelberg (2002)
Online document
Cartwright, J.: Big stars have weather too. IOP Publishing PhysicsWeb. http://physicsweb.org/articles/news/11/6/16/1 (2007). Accessed 26 June 2007
Always use the standard abbreviation of a journal's name according to the ISSN List of Title Word Abbreviations, see
www.issn.org/2-22661-LTWA-online.php
For authors using EndNote, Springer provides an output style that supports the formatting of in-text citations and reference list.
EndNote style (zip, 2 kB)
Tables: All tables are to be numbered using Arabic numerals.
Figure Numbering: All figures are to be numbered using Arabic numerals.
The soft copy of the final report should be submitted by email to Dr. PialiHaldar(piali.haldar@sharda.ac.in) by 16th April 2019 along with a hard copy signed by the faculty guide.
</td></tr>
<tr><td>8.5</td><td><u>Format:</u></td><td>
The report should be Spiral/ hardbound
The Design of the Cover page to report will be given by the Coordinator- CCC
Cover page
Acknowledgment
Content
Project report
Appendices
</td></tr>
</table>

School: SSBSR		**Batch: 2024-26**
Program: M.Sc.		**Academic Year: 2025-26**
Branch: Data Science & Analytics		**Semester: III**
1	Course Code	**MDA201**
2	Course Title	**Inferential Statistics**
3	Credits	4
4	Contact Hours (L-T-P)	4-0-0
	Course Status	Compulsory
5	Course Objective	The course aims to understand the different properties of an estimator. After studying this course students will be able to understand the power of the test.
6	Course Outcomes	CO1: Learn about the properties of the estimator. CO2: Understand the concept of the best estimator with examples CO3: Learn about the Rao, Lehman, and Bhattacharya bounds CO4: Understand the properties of MLE CO5: Learn the concept of the critical region and the power of the test CO6: Understand the unbiased test and Neyman structure
7	Course Description	Inferential statistics are concerned with making inferences based on relations found in the sample, to relations in the population.

8	Outline syllabus				CO Mapping
	Unit 1	**Properties of Estimator**			
	A	Point estimator, Interval estimator, Unbiasedness, Consistency, Efficiency, Sufficiency, Neyman Fisher lemma, Sufficient Statistics, and completeness,			CO1, CO2
	B	UMVUE, Cramer Rao Inequality along with the underlying conditions,			CO1, CO2
	C	Modification and extension of CR inequality.			CO1, CO2
	Unit 2	**Blackwellization**			
	A	Rao Blackwell theorem			CO3
	B	Lehman Scheffe theorem,			CO3
	C	Introduction to Bhattacharya bounds, consistency of an estimator.			CO3
	Unit 3	**MLE**			
	A	Maximum Likelihood estimation			CO4
	B	Properties of MLE			CO4
	C	BAN, Pitman estimator, and its efficiency.			CO4
	Unit 4	**Critical Region**			
	A	Best critical region, Generalized Neyman Pearson lemma,			CO5
	B	UMP tests for distribution with MLR			CO5
	C	LR test and their properties.			CO5
	Unit 5	**Neyman Structure**			
	A	Unbiased tests,			CO6
	B	Locally most powerful tests,			CO6
	C	Similar regions and tests of Neyman structure.			CO6
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25%	25%	50%	
	Text book/s*	1. Mood, Graybill and Boes, An introduction to the theory of Statistics 3rd edition			
	Other References	1. Kendal & Stuart, The Advanced Theory of Statistics Vol II, Charles Griffin. 2. E. L. Lehman, Testing of Statistical Hypothesis, John Wiley & Wiley Eastern			

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA201.1	3	2	2	3		3	3	2	1
MDA201.2	3	2	2	3		3	2	1	1
MDA201.3	3	2	2	3		2	2	1	1
MDA201.4	3	2	2	3		2	3	1	1
MDA201.5	3	2	2	3		3	2	2	1
MDA201.6	3	2	2	3		2	2	1	1

School: SSBSR		**Batch: 2024-26**			
Program: M.Sc.		**Academic Year: 2025-26**			
Branch: Data Science & Analytics		**Semester: III**			
1	Course Code	MDA202			
2	Course Title	Multivariate Data Analysis			
3	Credits	4			
4	Contact Hours (L-T-P)	4-0-0			
	Course Status	Compulsory			
5	Course Objective	The course aims to analyze multivariate data and understand the characteristics of multivariate quantitative research, including strengths and weaknesses. It also discusses the principles and characteristics of multivariate data analysis techniques.			
6	Course Outcomes	CO1: Learn about the multivariate data; Evolution and understanding of the data. CO2: Understand the basic concepts of multivariate normal distribution. CO3: Utilize the Wishart distribution in multivariate analysis. CO4: Mahalanobis D^2 and Hotelling T^2 CO5:Apply the classification rule in decision theory CO6: Utilization of PCA and factor analysis.			
7	Course Description	A large amount of data is collected on many different variables across disciplines to understand the underlying process(es). The multivariate analysis of data deals with examining the interrelationship between three or more equally important variables or explaining variation in usually one (or more than one) dependent variable(s) based on two or more independent (explaining) variables.			
8	Outline syllabus				CO Mapping
	Unit 1	**Multivariate Normal Distribution**			
	A	Multivariate Normal Distribution			CO1, CO2
	B	Probability density function and other properties			CO1, CO2
	C	Marginal and condition distribution.			CO1, CO2
	Unit 2	**Wishart**			
	A	Wishart distribution			CO3
	B	Probability density and distribution function,			CO3
	C	Characteristic function and its properties.			CO3
	Unit 3	**Data Pre-processing and Feature Selection**			
	A	Hotelling T^2, Mahalanobis D^2,			CO4
	B	Properties and functional forms of T^2 and D^2			CO4
	C	Represent their relationship and application.			CO4
	Unit 4	**Basic of R**			
	A	Classification analysis,			CO5
	B	discrimination analysis,			CO5
	C	Bayesian classification and decision design.			CO5
	Unit 5	**Basic Data Mining**			
	A	Principal Component Analysis,			CO6
	B	Canonical Correlation and variables,			CO6
	C	Factor Analysis,			CO6
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25%	25%	50%	
	Text book/s*	1. T.W. Anderson, Multivariate Analysis, John Wiley & Wiley Eastern.			
	Other References	2. Johnson & Wichem, Applied Multivariate Analysis, Wiley & Wiley Eastern.			

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA202.1	3	2	2	3		3	3	2	1
MDA202.2	3	2	2	3		3	2	1	1
MDA202.3	3	2	2	3		2	2	1	1
MDA202.4	3	2	2	3		2	3	1	1
MDA202.5	3	2	2	3		3	2	2	1
MDA202.6	3	2	2	3		2	2	1	1

School: SSBSR	**Batch: 2024-26**	
Program: M.Sc.	**Academic Year: 2025-26**	
Branch: Data Science & Analytics	**Semester: III**	

1	Course Code	MDA203	
2	Course Title	Soft Computing Techniques	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	Compulsory	
5	Course Objective	The main objective of the Soft Computing Techniques to Improve Data Analysis Solutions is to strengthen the dialogue between the statistics and soft computing research communities to cross-pollinate both fields and generate mutual improvement activities.	
6	Course Outcomes	At the end of the course, the student should be able to CO1: Learn about soft computing techniques and their applications. CO2: Analyse various neural network architectures. CO3: Understand perceptrons and counter-propagation networks. CO4: Define the fuzzy systems. CO5: Analyse the genetic algorithms and their applications. CO6: Provide a body of concepts and techniques for designing intelligent systems.	
7	Course Description	A PG-level course in Soft Computing Techniques to Improve Data Analysis Solutions is to strengthen the dialogue between the statistics and soft computing research communities.	
8	Outline syllabus		CO Mapping
	Unit 1	Soft Computing & AI	
	A	Introduction to soft computing, soft computing vs. hard computing, various types of soft computing techniques, and applications of soft computing.	CO1
	B	Introduction, Various types of production systems, characteristics of production systems, breadth-first search, depth-first search techniques, other Search Techniques like hill Climbing, Best-first Search, A* algorithm, AO* Algorithms, and various types of control strategies.	CO1
	C	Knowledge representation issues,Prepositional and predicate logic, monotonic and non-monotonic reasoning, forward Reasoning, backward reasoning, Weak & Strong Slot & filler structures, NLP.	CO1
	Unit 2	Neural Network	
	A	Structure and Function of a single neuron.	CO2
	B	Biological neuron, artificial neuron, the definition of ANN, Taxonomy of the neural net, Difference b/w ANN and the human brain.	CO2
	C	Characteristics and applications of AssNN, single layer network.	CO2
	Unit 3	Perceptron & Counter propagation network	
	A	Perceptron training algorithm, Linear separability, Widrow & Hebb's learning rule/Delta rule, ADALINE, MADALINE, AI v/s ANN.	CO3
	B	Introduction of MLP, different activation functions, Error back propagation algorithm, derivation of BBPA, momentum, limitation, characteristics and application of EBPA.	CO3
	C	Architecture, functioning & characteristics of counter Propagation network, Hop field/ Recurrent network, configuration, stability constraints, associative memory, and characteristics, limitations, and applications. Hopfield v/s Boltzman machine. Adaptive Resonance Theory: Architecture, classifications, Implementation, and training. Associative Memory.	CO3
	Unit 4	Fuzzy Logic & Fuzzy rule base system	
	A	Fuzzy set theory, Fuzzy set versus crisp set, Crisp relation & fuzzy relations.	CO4
	B	Fuzzy systems: crisp logic, fuzzy logic, introduction & features of membership functions.	CO4
	C	Fuzzy propositions, formation, decomposition & aggregation of fuzzy Rules, fuzzy reasoning, fuzzy inference systems, fuzzy decision making & Applications of fuzzy logic.	CO4

	Unit 5	Genetic algorithm			
	A	Fundamental, basic concepts, working principle, encoding, fitness function, and reproduction.			CO5
	B	Genetic modeling: Inheritance operator, cross over, inversion & deletion, mutation operator, Bitwise operator, Generational Cycle, Convergence of GA.			CO6
	C	Applications & advances in GA, Differences & similarities between GA & other traditional methods.			CO6
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25%	25%	50%	
	Text book/s*	1. S.N. Sivanandam & S.N. Deepa, Principles of Soft Computing, Wiley Publications, 2nd Edition, 2011. 2. S, Rajasekaran & G.A. Vijayalakshmi Pai, Neural Networks, Fuzzy Logic & Genetic Algorithms, Synthesis & applications, PHI Publication, 1st Edition, 2009.			
	Other References	1. N. K. Bose, Ping Liang, Neural Network fundamental with Graph, Algorithms & Applications, TMH, 1st Edition, 1998. 2. Bart Kosko, Neural Network & Fuzzy System, PHI Publication, 1st Edition, 2009. 3. Rich E, Knight K, Artificial Intelligence, TMH, 3rd Edition, 2012. 4. George J Klir, Bo Yuan, Fuzzy sets & Fuzzy Logic, Theory & Applications, PHI Publication, 1st Edition, 2009. 5. Martin T Hagen, Neural Network Design, Nelson Candid, 2nd Edition, 2008.			

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA203.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA203.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA203.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA203.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA203.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA203.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		Batch: 2024-26			
Program: M.Sc.		**Academic Year: 2025-26**			
Branch: Data Science & Analytics		**Semester: III**			
1	Course Code	MDA204			
2	Course Title	Exploratory Data Analysis and Visualization			
3	Credits	4			
4	Contact Hours (L-T-P)	4-0-0			
	Course Status	Compulsory			
5	Course Objective	The main objective of the course is to introduce the methods for data preparation and data understanding. It covers essential exploratory techniques for understanding multivariate data by summarizing it through statistical methods and graphical methods.			
6	Course Outcomes	At the end of the course, the student should be able to CO1: Handle missing data in real-world data sets by choosing appropriate methods. CO2: Summarize the data using basic statistics. Visualize the data using basic graphs and plots. CO3: Identify the outliers if any in the data set. CO4: Choose appropriate feature selection and dimensionality reduction. CO5: Techniques for handling multi-dimensional data. CO6: Having problem-solving ability- solving social issues and problems related to data science.			
7	Course Description	A PG-level course in Exploratory Data Analysis and Visualization to support and summarize the insurer's use of predictive analytics, data science, and Data Visualization.			
8	Outline syllabus				CO Mapping
	Unit 1	Introduction To Exploratory Data Analysis			
	A	Data Analytics lifecycle, Exploratory Data Analysis (EDA).			CO1
	B	Definition, Motivation, Steps in data exploration.			CO1
	C	The basic data types Data Type Portability.			CO1
	Unit 2	Pre-processing-Traditional Methods, MLE and Bayesian Estimation			
	A	Introduction to Missing data, Traditional methods for dealing with missing data.			CO2
	B	Maximum Likelihood Estimation: Basics, Missing data handling, Improving the accuracy of the analysis.			CO2
	C	Introduction to Bayesian Estimation, Multiple Imputation-Imputation Phase, Analysis, and Pooling Phase, Practical Issues in Multiple Imputations, Models for Missing Notation Random Data.			CO2
	Unit 3	Data Summarization & Visualization			
	A	Statistical data elaboration, 1-D Statistical data analysis.			CO3
	B	Statistical data elaboration, 2-D Statistical data analysis.			CO3
	C	Statistical data elaboration, N-D Statistical data analysis.			CO3
	Unit 4	Outlier Analysis & Feature Subset Selection			
	A	Introduction, Extreme Value Analysis, Clustering based, Distance Based and Density Based outlier analysis, Outlier Detection in Categorical Data.			CO4
	B	Feature selection algorithms: filter methods, wrapper methods, embedded methods, Forward selection backward elimination.			CO4
	C	Relief, greedy selection, genetic algorithms for features election.			CO4
	Unit 5	Dimensionality Reduction & Contemporary issues			
	A	Introduction, Principal Component Analysis (PCA), Kernel PCA.			CO5
	B	Canonical Correlation Analysis, Factor Analysis, Multi-dimensional scaling, Correspondence Analysis.			CO6
	C	Recent Trends Problems.			CO6
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25%	25%	50%	

	Text book/s*	1. Charu C. Aggarwal, “Data Mining The Textbook”, Springer, 2015.	
	Other References	1. Craig K. Enders, “Applied Missing Data Analysis”, The Guilford Press, 2010. 2. Inge Koch, “Analysis of Multivariate and High dimensional data”, Cambridge University Press, 2014. 3. Michael Jambu, “Exploratory and multivariate data analysis”, Academic Press Inc., 1990. 4. Charu C. Aggarwal, “Data Classification Algorithms and Applications”, CRC Press, 2015	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA204.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA204.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA204.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA204.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA204.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA204.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR		**Batch: 2024-26**
Program: M.Sc.		**Academic Year: 2025-26**
Branch: Data Science & Analytics		**Semester: III**
1	Course Code	**MDA 251**
2	Course Title	**Practical-V (based on MDA 201, and MDA 202 using R/SPSS/SAS/STRATA/Python)**
3	Credits	2
4	Contact Hours (L-T-P)	0-0-4
	Course Status	Compulsory
5	Course Objective	After studying this course students will be able to understand how to calculate the power of the test, analyze the multivariate data and understand the characteristics of multivariate quantitative research, including strengths and weaknesses. It also discusses the principles and characteristics of multivariate data analysis techniques.
6	Course Outcomes	At the end of the course, the student should be able to CO1: Estimate the parameter by MLE CO2: Learn about how to calculate the Rao, Lehman, and Bhattacharya bounds CO3: Learn how to calculate the critical region, power of the test, unbiased test, and Neyman structure. CO4: Understand the basic concepts of multivariate normal distribution. CO5: Calculate Wishart distribution in the multivariate analysis also know how to find Mahalanobis D^2 and HottelingT^2. CO6: Apply the classification rule, PCA, and factor analysis.
7	Course Description	In this course, students are concerned with making inferences based on relations found in the sample, to relations in the population. Also multivariate analysis of data deals with examining the interrelationship between three or more equally important variables or explaining variation in, usually one (or more than one) dependent variable(s) based on two or more independent (explaining) variables.

8	Outline syllabus				CO Mapping
	Unit 1				
		Problem-based on the estimation of the parameter, Rao, Lehman, and Bhattacharya bounds using SPSS/SAS/STRATA/R/Python.			CO1, CO2
	Unit 2				
		Problem-based on critical region, power of the test, unbiased test, and Neyman structure using SPSS/SAS/STRATA/R/Python.			CO2, CO3
	Unit 3				
		Problem-based on multivariate normal distribution using SPSS/SAS/STRATA/R/Python.			CO3, CO4
	Unit 4				
		Problem-based on Wishart distribution, Mahalanobis D2, and HottelingT2using SPSS/SAS/STRATA/R/Python.			CO4, CO5
	Unit 5				
		Problem-based on classification rule, PCA, and factor analysis using SPSS/SAS/STRATA/R/Python.			CO5, CO6
	Mode of examination	Practical			
	Weightage Distribution	CA	CE	ETE	
		25%	25%	50%	
	Text book/s*				
	Other References				

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA251.1	3	2	2	3		3	3	2	1
MDA251.2	3	2	2	3		3	2	1	1
MDA251.3	3	2	2	3		2	2	1	1
MDA251.4	3	2	2	3		2	3	1	1
MDA251.5	3	2	2	3		3	2	2	1
MDA251.6	3	2	2	3		2	2	1	1

School: SSBSR		**Batch: 2024-26**	
Program: M.Sc.		**Academic Year: 2025-26**	
Branch: Data Science & Analytics		**Semester: III**	
1	Course Code	**MDA 252**	
2	Course Title	Practical -VI (using based on MDA 203, and MDA 204 using R/ Python)	
3	Credits	2	
4	Contact Hours (L-T-P)	0-0-4	
	Course Status	Compulsory	
5	Course Objective	The objective of this course is to analyze solutions to strengthen the dialogue between the statistics and soft computing research communities to cross-pollinate both fields and generate mutual improvement activities. It covers essential exploratory techniques for understanding multivariate data by summarizing it through statistical methods and graphical methods.	
6	Course Outcomes	At the end of the course, the student should be able to CO1: Learn about soft computing techniques and their applications, and analyze various neural network architectures. CO2: Understand perceptrons and counter propagation networks, Define the fuzzy systems. CO3: Analyze the genetic algorithms and their applications. CO4: Handle missing data in real-world data sets by choosing appropriate methods. CO5: Summarize the data using basic statistics. Visualize the data using basic graphs and plots. Identify the outliers if any in the data set. CO6: Choose appropriate feature selection and dimensionality reduction. Techniques for handling multi-dimensional data.	
7	Course Description	Using R/ Python try to solve the problem related to Soft Computing Techniques. Exploratory Data Analysis, Visualization, summarizes the insurer's use of predictive analytics, identifies the outliers, dimensionality reduction, and Data Visualization for multi-dimensional data.	
8	Outline syllabus		CO Mapping
	Unit 1		
		Create a perceptron with the appropriate no. of inputs and outputs. Train it using a fixed increment learning algorithm until no change in weights is required. Output the final weights. Create a simple ADALINE network with an appropriate no. of input and output nodes. Train it using the delta learning rule until no change in weights is required. Output the final weights.	CO1
	Unit 2		
		Train the autocorrelator by given patterns: A1=(-1,1,-1,1), A2=(1,1,1,-1), A3=(-1, -1, -1, 1). Test it using patterns: Ax=(-1,1,-1,1), Ay=(1,1,1,1), Az=(-1,-1,-1,-1). Train the hetrocorrelator using multiple training encoding strategies for given patterns: A1=(000111001) B1=(010000111), A2=(111001110) B2=(100000001), A3=(110110101) B3(101001010). Test it using pattern A2.	CO1, CO2
	Unit 3		
		Implement Union, Intersection, Complement, and Difference operations on fuzzy sets. Also, create fuzzy relation by the Cartesian product of any two fuzzy sets and perform max-min composition on any two fuzzy relations. Solve Greg Viot's fuzzy cruise controller using Python Fuzzy logic	CO2, CO3

		toolbox. Solve Air Conditioner Controller using Python Fuzzy logic toolbox. Implement TSP using GA.			
	Unit 4				
		Problem-based on Data Summarization, Visualization, Outlier Analysis, and Feature Subset Selection using R/python.			CO4,CO5
	Unit 5				
		Problem-based on PCA, Canonical Correlation Analysis, Factor Analysis, Multi-dimensional scaling, and Correspondence Analysis using R/python.			CO5, CO6
	Mode of examination	Practical			
	Weightage Distribution	CA	CE	ETE	
		25%	25%	50%	
	Text book/s*				
	Other References				

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA252.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA252.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA252.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA252.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA252.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA252.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

Branch: Data Science & Analytics		Semester: IV	
1	Course Code	MDA214	
2	Course Title	Statistical Simulation	
3	Credits	4	
4	Contact Hours (L-T-P)	4-0-0	
	Course Status	Elective	
5	Course Objective	To demonstrate and intended to verse students in the techniques necessary to understand and carry out methods of research in Statistical simulation.	
6	Course Outcomes	CO1: Explain the concept of statistical simulation. (K1, K2, K3). CO2: How to generate random numbers by the different methods (K1, K2, K4) CO3: Explain the concept of the MCMC technique. (K3, K4, K5). CO4: Recognize the concepts of probability and statistics that are relevant to modeling and simulation. (K3, K4, K5). CO5: Design and implement Bootstrapping; jackknife resampling. (K3, K4, K5). CO6: How simulation may be used to understand the behavior of real-world systems by utilizing mathematical models with an emphasis on simulation (K3, K4, K5).	
7	Course Description	A PG-level course in Statistics, intended to verse students in the techniques necessary to understand and carry out methods of research in Statistical simulation. Lectures study the various applications of the MCMC technique.	
8	Outline syllabus		CO Mapping
	Unit 1		
	A	Review of R/Python. Random number generation	CO1
	B	Inverse-transform; acceptance-rejection; transformations.	CO1, CO2
	C	Statistic simulations: generating random variables, and simulating normal, gamma, and beta random variables.	CO1, CO2
	Unit 2		
	A	Simulating multivariate distributions, MCMC methods.	CO3
	B	Gibbs sampler, simulating random fields, Simulating stochastic process.	CO3
	C	Variance reduction technique, importance sampling for integration, Control variate, and antithetic variables.	CO3
	Unit 3		
	A	Bootstrapping; jackknife resampling. Bootstrapping for estimation of the sampling distribution.	CO5
	B	Confidence intervals, variance stabilizing transformation.	CO5
	C	Bootstrapping in regression and sampling from finite populations.	CO5
	Unit 4		
	A	Simulating a non-homogeneous Poisson process.	CO4
	B	Optimization using Monte Carlo methods simulated annealing for optimization	CO4
	C	Solving differential equations by Monte Carlo methods	CO4
	Unit 5		
	A	Univariate density estimation, kernel smoothing multivariate density estimation	CO3, CO6
	B	Root finding: Numerical integration, numerical maximization/minimization, constrained and unconstrained optimization.	CO3, CO6
	C	EM algorithm, Simplex algorithm	CO3, CO6

	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25%	25%	50%	
	Text book/s*	Fishman, G.S. (1996). Monte Carlo: Concept, algorithm, and application. (Springer)			
	Other References	Rubinstien R.V. (1981). Simulation and Monte Carlo method. Reply, B. D. (1987). Stochastic Simulation. (Wiley).			

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA214.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA214.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA214.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA214.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA214.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA214.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

School: SSBSR	**Batch: 2024-26**
Program: M.Sc.	**Academic Year: 2025-26**
Branch: Data Science	**Semester: IV**

& Analytics					
1	Course Code	**MDA222**			
2	Course Title	**Applied Econometrics**			
3	Credits	4			
4	Contact Hours (L-T-P)	4-0-0			
	Course Status	Elective			
5	Course Objective	The objective of this course is to introduce regression analysis to students so that they can understand its applications in different fields of economics.			
6	Course Outcomes	CO1: Able to have a concise knowledge of basic regression analysis of economic data and interpret and critically evaluate outcomes of empirical analysis. (K1, K2, K3). CO2: Analyse the theoretical background for standard methods used in empirical analyses, like properties of least squares estimators and statistical testing of hypotheses. (K2, K3, K4). CO3: Able to apply for modern computer programs in regression analyses of empirical data, including statistical testing to investigate whether the classical assumptions in regression analysis are satisfied. (K2, K3, K4). CO4: Design and development of a real-life model based on econometric methods. (K4, K5, K6). CO5: Develop and apply advanced methods for the implementation of econometric techniques also various functions for economic analysis and future forecasting. (K5, K6). CO6: Able to use econometric models in their future work. (K4, K5).			
7	Course Description	The purpose of this course is to give students a solid foundation in econometric techniques, various functions for economic analysis, and future forecasting. Many of the methods introduced in this course are also useful in business, finance, and many other disciplines.			
8	Outline syllabus				CO Mapping
	Unit 1				
	A	Introduction to econometrics. A review of least squares and maximum likelihood estimation methods of parameters in the classical linear regression model and their properties.			CO1
	B	Generalized least squares estimation and prediction, construction of confidence regions.			CO1
	C	Tests of hypotheses, use of dummy variables, and seasonal adjustment.			CO1
	Unit 2				
	A	Regression analysis under linear restrictions, restricted least squares estimation method, and its properties.			CO2
	B	The problem of Multicollinearity, its implications, and tools for handling the problem.			CO2
	C	Ridge regression. Heteroscedasticity, consequences, and tests for it.			CO2
	Unit 3				
	A	Estimation procedures under heteroscedastic disturbances, Bartlett's test, Breusch Pagan test, and GoldfelfQuandt test.			CO3
	B	Autocorrelation, sources, and consequences.			CO3
	C	Autoregressive process tests for autocorrelation.			CO3
	Unit 4				
	A	Durbin Watson test. Asymptotic theory and regressors.			CO4
	B	Instrumental variable estimation, errors in variables.			CO4
	C	Simultaneous equations model, the problem of identification, a necessary and sufficient condition for the identifiability of parameters in a structural equation.			CO4
	Unit 5				
	A	Ordinary least squares, indirect least squares.			CO5
	B	Two-stage least square.			CO6
	C	Limited information maximum likelihood method.			CO6
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25%	25%	50%	

	Text book/s*	1. Gujarati, D.N. & Porter, D.C. (2017). Basic Econometrics, 6th Edition.McGraw Hill. 2. Maddala, G.S. &Lahiri, K. (2010). Introduction to Econometrics, 4th Edition.Wiley.	
	Other References	1. Greene, W.H. (2012). Econometric Analysis, 7th Edition. Pearson. 2. Studenmund, A.H. &Johnson, B.K. (2017). Using Econometrics: A Practical Guide, 7th Edition. Pearson.	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA222.1	**3**	**2**	**2**	**3**		**3**	**3**	**2**	**1**
MDA222.2	**3**	**2**	**2**	**3**		**3**	**2**	**1**	**1**
MDA222.3	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**
MDA222.4	**3**	**2**	**2**	**3**		**2**	**3**	**1**	**1**
MDA222.5	**3**	**2**	**2**	**3**		**3**	**2**	**2**	**1**
MDA222.6	**3**	**2**	**2**	**3**		**2**	**2**	**1**	**1**

Analytics					
1	Course Code	**MDA253**			
2	Course Title	**Capstone Project**			
3	Credits	10			
4	Contact Hours (L-T-P)	0-0-20			
	Course Status	Compulsory			
5	Course Objective	The course should be taught and implemented to develop the required course outcomes so that students will acquire the following competency needed by the industry: Plan innovative/creative solutions independently and/or collaboratively to integrate various competencies acquired during the semesters to solve/complete the identified problems/task/shortcomings faced by industry/user related to the concerned occupation.			
6	Course Outcomes	CO1: Plan a scientific project proposal with time duration (K2, K3). CO2: Select, collect, and use required information/knowledge to solve the problem/complete the task (K3, K4). CO3: Logically choose relevant possible solutions (K3, K4). CO4: Consider the ethical issues related to the project (if there are any) (K4, K5). CO5: Assess the impact of the project on society (if there is any) (K4, K5). CO6: Compile the entire project work to prepare a 'project report' with future scope. (K5, K6).			
7	Course Description	The course aims to give exposure to research in a real scenario to students. It caters to the needs of research designs, research methods, and various methodologies used. The course will further explain how to apply various data analysis tools to draw workable inferences for numerous problems and this course sharpens the student's analytical and decision-making skills.			
8	Outline syllabus				CO Mapping
	Unit 1				
	A	Feasibility studies, Design projects,			CO1
	B	Market surveys			CO1
	C	Prototype (design, make, test, and evaluate)			CO1
	Unit 2	.			
	A	Advanced work requires the development of existing work to be used and developed.			CO2
	B	Field works: This could include surveys			CO2
	C	Charting data and information from visual observation.			CO2
	Unit 3				
	A	Comparative Studies: Theoretical study of systems/mechanisms/ processes in detail and comparing them based on cost/energy conservation/impact on environment/technology used etc.			CO3
	B	Application of Emerging science/technology: Theoretical study of some emerging concepts,			CO3
	C	Feasibility of its application in some real-life situations in detail.			CO3
	Unit 4				
	A	Collection/combination of some concepts etc.			CO4
	B	Construction of some structure/concepts			CO4
	C	Development of software or use of software for solving some broad-based problem.			CO4
	Unit 5				
	A	Plan for a report must have the following contents: introduction, review of literature, and research gaps of the study.			CO5
	B	Significance of the study, research methodology: objectives of the study, hypotheses of the study.			CO6
	C	Data analysis and interpretation, findings and conclusion, recommendations and limitations, Bibliography Annexure- Questionnaire/Schedules if any.			CO6
	Mode of examination				
	Weightage Distribution	CA	CE	ETE	
		25%	25%	50%	

	Text book/s*	1. Rubin, Allen & Babbie, Earl (2009). Essential Research Methods for Social Work, Cengage Learning Inc., USA.	
	Other References	2. Neuman, W.L. (2008). Social research methods: Qualitative and quantitative approaches, Pearson Education. 3. Pawar, B.S. (2009). Theory building for hypothesis specification in organizational studies, Response Books, New Delhi.	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA253.1	**3**	**2**	**2**	**3**	**1**	**3**	**3**	**2**	**1**
MDA253.2	**3**	**2**	**2**	**3**	**1**	**3**	**2**	**1**	**1**
MDA253.3	**3**	**2**	**2**	**3**	**1**	**2**	**2**	**1**	**1**
MDA253.4	**3**	**2**	**2**	**3**	**1**	**2**	**3**	**1**	**1**
MDA253.5	**3**	**2**	**2**	**3**	**1**	**3**	**2**	**2**	**1**
MDA253.6	**3**	**2**	**2**	**3**	**1**	**2**	**2**	**1**	**1**