Master of Science

Data Science &Analytics

AY: 2021- 22

Program Structure

School of Basic Science and Research

Department of Mathematics

M.Sc. (Data Science& Analytics)

Batch 2021-23

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 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 (PEO's)

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 positive attitude and skills to enable a multi facet personality.
PEO3: The graduates will prepare for pursue higher education and research.
PEO4: The graduates will develop for contribute to the society and human well-being by applying ethical principles.

1.4.1 Program Outcomes (PO's)

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 towards 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 (PSO's)

PSO1 : Utilize the 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 and 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 the real life problems in the fields of data science and apply the research-based knowledge to analyse and solve advanced problems in data science.

1.4.2 Mapping of PEOs with Mission Statements:

PEO Statements	School Mission 1	School Mission 2	School Mission 3	School Mission 4	School Mission 5	School Mission 6
PEO1:	3	2	3	1	2	3
PEO2:	3	2	3	2	2	3
PEO3:	3	3	3	3	3	3
PEO4:	3	2	3	2	3	3

1.4.3 Mapping of Program Outcome (PO's)Vs Program Educational Objectives (PEO's)

	PEO1	PEO2	PEO3	PEO4
PO1	3	3	3	2
PO2	3	3	3	2
PO3	3	3	3	2
PO4	3	2	3	2
PO5	2	3	2	3
PSO1	2	2	3	2
PSO2	3	2	2	3
PSO3	3	3	2	3
PSO4	3	2	3	3

1. Slight (Low) ***2. Moderate (Medium)*** ***3. Substantial (High)***

1.3.5 Program Outcome (PO's)Vs Courses Mapping Table:

1.3.5.1 COURSE ARTICULATION MATRIX

Co's	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
MDA101	3	2	2	3	2	3	3	2	2
MMT104	3	2	2	3	3	3	2	2	2
MDA102	3	2	2	3	2	3	3	2	2
MDA103	3	2	2	3	2	3	3	2	2
MDA104	3	2	2	3	2	3	3	2	2
MDA151	3	3	2	3	3	3	3	3	3
MDA152	2	3	2	3	3	2	3	3	3
MMT123	3	2	2	3	2	3	3	2	2
MDA105	3	2	3	3	2	3	3	2	2
MDA106	3	2	2	3	2	3	3	2	2
MDA107	3	2	3	3	3	3	3	2	2
MDA108	2	1	1	2	1	1	2	1	1
CCU401	3	2	3	3	3	3	3	2	2

MDA153	3	3	2	2	3	3	3	3	3
MDA154	3	2	3	3	2	3	3	2	3
MDA201	3	2	2	3	3	3	2	2	2
MDA202	3	2	2	3	2	3	3	2	2
MDA203	3	3	3	3	2	3	2	2	2
MDA204	3	2	2	3	3	3	3	2	2
OPE XXX	3	2	2	3	2	3	2	2	2
MDA251	3	3	2	3	2	3	2	3	3
MDA252	3	2	3	3	2	3	3	2	3
MDAXXX	3	3	2	2	3	3	3	2	3
MDA XXX	2	3	2	2	3	3	3	3	2
MDA 253	3	2	2	3	3	3	3	2	2

1-Slight (Low)
2-Moderate (Medium)
3-Substantial (High)

Department of Mathematics
School of Basic Sciences and Research
M. Sc. (Data Science & Analytics)
Batch: 2021-23
TERM: I

S. No.	SUBJECT CODE	Title of Paper	Teaching Load				CREDITS	PRE-REQUISITE/CO-REQUISITE	Type of Course[1]: 1. CC 2. AECC 3. SEC 4. DSE
	THEORY								
			L	**T**	**P**	**TOTAL**			
1.	MDA101	Foundations of Data Science	4	0	-	4	4		CC
2.	MMT104	Statistical Methods	4	0	-	4	4		CC
3.	MDA102	Mathematics for Machine Learning	4	0	-	4	4		CC
4.	MDA103	Probability Theory and Distributions	4	0	-	4	4		CC
5.	MDA104	Next Generation Databases	4	0	-	4	4		AECC
	PRACTICALS								
6.	MDA151	Practical -I (Based on Paper MMT104, MDA102 Using Excel/SPSS/Mini-tab)	-	-	3	3	2		AECC
7	MDA152	Practical -II (Based on Paper MMT104, MDA102, 103, 104 UsingR/ Python)	-	-	3	3	2		AECC
TOTAL			**15**	**-**	**6**	**26**	**24**		

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

Department of Mathematics
School of Basic Sciences and Research
M. Sc. (Data Science & Analytics)
Batch: 2020-22
TERM: II

S. No.	SUBJECT CODE	Title of Paper	Teaching Load				CREDITS	PRE-REQUISITE/CO-REQUISITE	Type of Course2: 1. CC 2. AECC 3. SEC 4. DSE
	THEORY								
			L	**T**	**P**	**TOTAL**			
1.	MMT123	Numerical Methods with Programming	4	0	0	4	4		CC
2.	MDA105	Regression Analysis and Predictive Models	4	0	0	4	4		CC
3.	MDA106	Statistical Data Preparation& Analytics	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
6.	CCU401	Community Connect	-	-	2	2	2		SEC
	PRACTICALS								
7.	MDA153	Practical -III (Based on Paper MDA105, 106, 107UsingR/ Python/SAS/SPSS)	-	-	3	3	2		AECC
8.	MDA154	Practical -IV (Based on Paper MDA108UsingR/Python)	-	-	3	3	2		AECC
TOTAL			**15**	**-**	**13**	**28**	**26**		

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

Department of Mathematics
School of Basic Sciences and Research
M. Sc. (Data Science & Analytics)
Batch: 2021-23
TERM: III

S. No.	SUBJECT CODE	Title of Paper	Teaching Load				CREDITS	PRE-REQUISITE/ CO-REQUISITE	Type of Course3: 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	MDA204	Exploratory Data Analysis and Visualization	4	0	0	4	4		
5	OPE XXX	Open elective (GE)	-	-	-	2	2		AECC
	PRACTICALS								
6.	MDA251	Practical -V (based on MDA 201 , MDA 202) (using SPSS/SAS/STRATA)	-	-	3	3	2		AECC
7.	MDA252	Practical -VI (using based on MDA 203 , MDA 204)	-	-	3	3	2		
TOTAL			**12**	**-**	**10**	**22**	**22**		

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

Department of Mathematics
School of Basic Sciences and Research
M. Sc. (Data Science & Analytics)
Batch: 2021-223
TERM: IV

S. No.	SUBJECT CODE	Title of Paper	HOURS				CREDITS	PRE-REQUISITE/ CO-REQUISITE	Type of Course4: 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		DSC
2.	MDA XXX	Elective-II (Online/Offline Courses)	4	0	0	4	4		DSC
	DISSERTATION								
	MDA253	Capstone project (Based on full time training program/internship program in any government/private institute or industry during last semester)	-	-	30	6 weeks (min. 30 days)	10		AECC
TOTAL			**6**	**2**	**30**	**38**	**18**		

TOTAL CREDIT: 90

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

SYLLABUS

M. Sc. (Data Science & Analytics)

School: SBSR		**Batch :2021-23**	
Program: M.Sc.		**Current Academic Year: 2021-22**	
Branch: M. Sc. 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 applying 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 for real life solution.(K3,K4) CO5:Apply and develop basic Machine Learning Algorithms. (K5,K6) CO6: Apply the statistical measures of R in real time environment. (K5,K6)	
7	Course Description	A PG-level course in foundation of data science, intended to verse students in the techniques necessary to understand and carry out methods in 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 Data Science storing data-combining bits into larger structures	CO1
	C	The steps in Doing Data Science-Skills needed to-Identifying 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
	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
	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.	CO6
	C	Vector Machines in R-Creating Web Applications With R.	CO6
	Mode of examination	Theory	
	Weightage Distribution	CA / MTE / ETE 25 Marks / 25 Marks / 50 Marks	
	Text book/s*	1. Jeffrey S. Saltz, Jeffre M. Stanton, "An Introduction 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
C101.1	3	3	3	3	3	3	3	2	2
C101.2	3	2	3	3	2	3	2	1	1
C101.3	2	2	2	2	2	2	2	1	2
C101.4	2	2	1	2	2	2	3	1	1
C101.5	3	2	2	3	2	3	2	2	2
C101.6	3	2	1	3	2	2	2	1	2

School: SBSR		Batch :2021-23	
Program: M.Sc.		Current Academic Year: 2021-22	
Branch: M. Sc. 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 in machine learning.	
6	Course Outcomes	CO1: Solve system of Linear equations by applying 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 a Matrix. (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 of differentiation. (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, computation of eigenvalues, least square problems and error analysis will be discussed.	
8	Outline syllabus		CO Mapping
	Unit 1	**Basic Concept of Linear Algebra**	
	A	Linear Algebra – System of Linear equations, Solving System of Linear equations.	CO1
	B	Linear Independence, Vectors, Scalars, Addition, Scalar multiplication.	CO1
	C	Dot product, vector projection, cosine similarity.	CO1
	Unit 2	**Vector**	
	A	Orthogonal vectors, normal and Orthonormal vectors.	CO2
	B	Vector norm, vector space, linear combination.	CO2
	C	Basis of vectors, Affine spaces.	CO2
	Unit 3	**Matrices and Determinants**	
	A	Matrices – Determinant, Identity matrix, Inverse of a matrix.	CO3
	B	Rank of a matrix, Nullity, trace of a matrix.	CO3

	C	Eigen values, Eigen vectors, Matrix decompositions.			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 vector valued function. Gradient of matrices.			CO6
	C	Optimization using gradient functions, Constrained optimization and Lagrange multipliers. Convex optimization.			CO6
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25 Marks	25 Marks	50 Marks	
	Text book/s*	1. Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, "Mathematics for Machine Learning", **Cambridge University Press, 2020.** 2.			
	Other References	1. Erwin Kreyszig, Advanced Engineering Mathematics, 10th 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
C102.1	**3**	**3**	**3**	**3**	**3**	**3**	**3**	**2**	**1**
C102.2	**3**	**2**	**3**	**3**	**2**	**3**	**2**	**1**	**2**
C102.3	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**
C102.4	**2**	**2**	**1**	**2**	**2**	**2**	**3**	**1**	**2**
C102.5	**3**	**2**	**2**	**3**	**2**	**3**	**2**	**2**	**1**
C102.6	**3**	**2**	**1**	**3**	**3**	**2**	**2**	**1**	**2**

School: SBSR		**Batch :2021-23**	
Program: M.Sc.		**Current Academic Year: 2021-22**	
Branch: M. Sc. 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 the real life data.	
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 Statistical analysis. (K5,K6)	
7	Course Description	To integrate the intrinsic ideas of preliminary and advanced distributions to correlate with the real-world scenarios.	
8	Outline syllabus		CO Mapping
	Unit 1	**Probability and Random variables**	
	A	Introduction – 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 its properties.	CO1
	C	Bivariate Random Variables- Joint Probability functions, marginal distributions, conditional distribution functions; Notion of Independence of Random variables.	CO1
	Unit 2	**Random Variables and Expectation**	

	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
	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
	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, and applications, procedure of hypothesis testing.			CO6
	C	Chi-square distribution and order statistics: properties, and applications, procedure of hypothesis testing.			CO6
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25 Marks	25 Marks	50 Marks	
	Text book/s*	1. Sheldon Ross; A First Course in Probability, Pearson, 2014.			

		2. Parimal Mukhopadhyay; An Introduction to the Theory of Probability, World scientific, 2012. 3. Irwin Miller, Marylees 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 for scientists and engineers, John Wiley, 2015.	

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
C103.1	3	3	3	3	3	3	3	2	1
C103.2	3	3	3	3	2	3	2	1	2
C103.3	2	3	2	2	2	2	2	2	2
C103.4	2	2	1	2	2	2	3	1	1
C103.5	3	2	2	3	2	3	2	2	2
C103.6	3	2	2	3	2	2	2	1	2

School: SBSR		Batch :2021-23	
Program: M.Sc.		Current Academic Year: 2021-22	
Branch: M. Sc. 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 data base patterns.	
6	Course Outcomes	After completion of this course, students will be able to CO1: Develop and Explore the relationship between Big-Data and NoSQL databases. (K1,K2,K3) CO2: Formulate 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-data for useful business applications. (K4,K5) CO5: Discuss different data models to suit various data representation and storage needs. (K5,K6) CO6: Explain and correlate with different data models to suit various data representation and storage needs. (K5,K6)	
7	Course Description	To integrate the intrinsic ideas for use various Data models for a variety of databases.	
8	Outline syllabus		CO Mapping
	Unit 1		
	A	Database Revolutions- system Architecture-Relational Database.	CO1
	B	Database Design-Data Storage-Transaction Management.	CO1
	C	Data warehouse and Data Mining-Information Retrieval.	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	Column Databases—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

	Unit 4				
	A	DistributedDatabasePatterns—DistributedRelationalDatabases-Non-relationalDistributed Databases.			CO4
	B	MongoDB - Sharingand Replication-HBase-Cassandra-Consistency Models.			CO4
	C	Types of Consistency-Consistency MongoDB - HBase Consistency-Cassandra Consistency.			CO4
	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.			CO6
	C	Other Convergent Databases-Disruptive Database Technologies.			CO6
	Mode of examination	Theory			
	Weightage Distribution	CA	MTE	ETE	
		25 Marks	25 Marks	50 Marks	
	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”, Apress, 2015. 2. Eric Redmond, Jim R Wilson, “Seven Databases in 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
C104.1	3	3	3	3	3	3	3	2	1
C104.2	3	2	3	3	2	3	2	1	2
C104.3	2	2	2	2	2	2	2	2	1
C104.4	2	2	1	2	2	2	3	1	1
C104.5	3	2	2	3	2	3	2	2	2
C104.6	3	2	1	3	2	2	2	1	2

Practical-I (MDA151)

School: SBSR		**Batch: 2021-23**
Program: M.Sc.		**Current Academic Year: 2020-21**
Branch: Mathematics		**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)	0-0-3
	Course Status	Compulsory
5	Course Objective	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/Minitabconcepts and analytical tools to analyze statistical problem and handle real-world issues.
6	Course	CO1: Describe the overall process and particular steps in designing

	Outcomes	studies, collecting and analyzing data, interpreting and presenting results. (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 theory. (K3,K4) CO4: Discuss and illustrate various discrete and continuous probability distributions and to 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 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 problem 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.	CO1, CO2
	Unit 3		
		Fitting of curves by 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.	CO2, CO3
	Unit 4		

		Fitting of Binomial, Poisson and Normal distributions to observed data and testing of goodness of fit.			CO4,CO5
	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	Viva	ETE	
		25 Marks	25 Marks	50 Marks	
	Text book/s*				
	Other References				

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
C151.1	3	3	2	2	2	3	3	3	3
C151.2	2	3	3	2	3	2	3	2	2
C151.3	2	3	2	2	3	1	3	2	3
C151.4	2	3	2	2	1	2	1	2	2
C151.5	3	3	2	2	2	2	3	3	3
C151.6	3	3	2	3	2	2	3	3	3

Practical-II (MDA152)

School: SBSR		**Batch: 2021-23**
Program: M.Sc.		**Current Academic Year: 2020-21**
Branch: Mathematics		**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	0-0-3

	(L-T-P)		
	Course Status	Compulsory	
5	Course Objective	Introduce basic concepts of R/ Python environment and provide students with a general understanding of R/ Python for solving the data analytics based problem. Equip students with the 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 and working directory. (K2,K3) CO3: Discuss, calculate and understands the Statistics and plot and interpret the graph in R/ Python in R/ Python. (K2,K3,K4) CO4: Discuss probability distribution and testing of hypothesis through R / Python and explain R/ Python programming language for it. (K3,K4) CO5: Discuss and Explain creating matrices and some simple matrix operations, Sub-matrices in R/ Python. (K4,K5) CO6: Develop a deeper understanding of the write R/ Python functions for Next Generation Databases. (K4,K5)	
7	Course Description	Introduce basic concepts of R/ Python environment and provide students with a general understanding of R/ Python for solving the data analytics based problem. Equip students with the 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 fivenum (). Calculate arithmetic mean (AM), geometric mean (GM), harmonic mean (HM), median, mode, quantiles, range, quartile deviation (QD), variance, coefficient of variation (CV) using R/ Python.	CO1, CO2

	Unit 3				
		Computation of probabilities of negative binomial, multinomial, normal, exponential, gamma, $t, \aleph^2, F$ using R/ Python.			CO2, CO3
	Unit 4				
		Creating matrices, some simple matrix operations, Sub-matrices also solve derivatives and some basic derivative function by using R/ Python.			CO4,CO5
	Unit 5				
		File operations, Reading Next Generation Databases, Data Structures.			CO5, CO6
	Mode of examination	Practical			
	Weightage Distribution	CA	Viva	ETE	
		25 Marks	25 Marks	50 Marks	
	Text book/s*				
	Other References				

COURSE OUTCOMES – PROGRAMME OUTCOMES MAPPING TABLE

PO / CO	PO1	PO2	PO3	PO4	PO5	PSO1	PSO2	PSO3	PSO4
C152.1	3	3	2	3	2	2	3	1	2
C152.2	2	3	3	2	3	2	3	2	2
C152.3	2	3	1	2	3	1	3	2	3
C152.4	2	3	2	2	1	2	3	2	2
C152.5	3	3	1	2	2	2	3	3	2
C152.6	3	3	2	3	2	2	3	2	3