C P Patel & F H Shah Commerce (Autonomous) College, Anand
PROGRAMME STRUCTURE
Master of Science
M.Sc. Data Science
Semester: 2
(Syllabus with effect from the Academic Year June 2024)

Course Objectives - For M.Sc. Data Science Semester-II Programme	**After successful completion of this course, the student will be able to:** 1. The primary objective of the M.Sc. in Data Science program is to develop skilled professional workforce that is prepared to address the increasing needs in the rapidly expanding area of big data analytics. 2. The program aims to provide skills in quantitative data analysis, data mining, data modeling and prediction, data storage and management, machine learning, big data processing, data visualization, multimedia big data, programming and communication skills. Get familiarity with aspects of Business Environment. 3. Value Added Course/ training and a large number of practical case studies have been integrated in the program to boost the learner confidence and market acceptability.
Programme Specific Outcome (PSO) - For M.Sc. Data Science Semester - 2	**After successful completion of this semester, the student will be able to:** 1. Implement professional knowledge in setting up road map to be an entrepreneur and identify research areas. 2. Ability to develop skills to address and solve social and environmental problem with ethics and perform multidisciplinary projects with advance technologies and tools. 3. Data Science Post Graduates are able to become leaders in the society with the help of advanced knowledge and skill, which can empower them to analyses, design, develop and implement their learning to develop the society.

Course Type	Course Code	Course Title	T/P	Credit	Exam Duration Hours	Component of Marks		
						Internal	External	Total
Core Courses	PG02CMDS01	Programming in R	T	4	3	30/12	70/28	100/40
	PG02CMDS02	Advanced RDBMS	T	4	3	30/12	70/28	100/40
	PG02CMDS03	Big Data Analytics	T	4	3	30/12	70/28	100/40
	PG02CMDS04	Artificial Intelligence using ML	T	4	3	30/12	70/28	100/40
	PG02CMDS05	Computational Statistics	T	4	3	30/12	70/28	100/40
Practical Ability	PG02PMDS06	Practical's of R and Oracle RDBMS	P	5	3	--	100/40	100/40
		Total		25		150/60	450/180	600/240

3 (Lectures) + 1 (CSDS) + 1 (Assignments)

To Pass:	**1. At least 40% Marks in the University Examination in each paper and** **2. At least 40% Marks in the individual Head of passing or 30% marks in aggregate at the internal tests conducted by the PG Centers**

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE 'A+' UGC – MHRD, Govt of India – June 2022

Effective from the Academic Year 2024-2025

Master of Science (Data Science)
Semester –2

Course Code	PG02CMDS01	Title of the Course	Programming in R
Total Credits of the Course	04	Hours per Week	04

Course Objectives	• Understand the fundamentals, standards of Functions and capabilities of RLanguage. • Learning the basic R-Language Constructs • To demonstrate Simulation in R-Language, Math functions and files Processing. • To know the Principals of Graphics and R-Base Graphics Analyze

Course Outcomes	• Be comfortable using commercial and open source tools such as the R language and its associated libraries for data analytics and visualization. • Able to use basic R data structures in loading, cleaning the data and preprocessing the data.

Course Content		
Unit	**Description**	**Weight age (%)**
1.	**Introduction Of R-Language** • Introduction, How to run R, R Sessions and Functions, Basic Math, Variables, Data Types, Vectors, Advanced Data Structures, Data Frames, Lists, Matrices, Arrays, Classes. • R Programming Structures, Control Statements, Loops, - Looping Over Non-vector Sets, If-Else, Arithmetic and Boolean Operators	25%
2.	**R Functions And Matrices** • Functions - Creating User defined functions - Functions on Function Object - Scope of Variables - Accessing Global, Environment - Closures - Recursion. • Creating Matrices - Adding or removing rows/columns - Reshaping - Operations -Special functions on Matrices. Lists - Creating List – General List Operations - Special • Adding - Removing - Applying Special functions to Data Frames - Merging Data Frames Factors and Tables	25%

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND
(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)
AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR
An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE 'A$^+$' UGC – MHRD, Govt of India – June 2022
Effective from the Academic Year 2024-2025

3.	**Exploratory Data Analysis using R** • Data Preprocessing -Descriptive Statistics - Central Tendency - Variability - Mean - Median - Range - Variance -Summary - Handling Missing values and Outliers – Normalization • Linear Regression- Simple Linear Regression - Implementation in R - functions on lm() - predict() - plotting and fitting regression line.	**25%**
4.	**Data Visualization in R** • Types of visualizations - packages for visualizations – Basic Visualizations, Advanced Visualizations and Creating 3D plots. • Box plot, Histogram, Pareto charts, Pie graph, Line chart, Scatterplot, Developing graphs • Graphics, Creating Graphs, The Workhorse of R Base Graphics, the plot() Function Customizing Graphs, Saving Graphs to Files.	**25%**

Teaching-Learning Methodology	Material fort his course will be presented using multiple teaching approach: lecture and discussion, exploration and inquiry, cooperative groupwork, demonstrations, and presentations

EvaluationPattern		
Sr. No.	DetailsoftheEvaluation	Weightage
1.	Internal Written	**30%**
2.	Internal Continuous Assessment in the form of Quizzes, Seminars, Assignments, Attendance	
3.	External Examination	**70%**

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE 'A^+' UGC – MHRD, Govt of India – June 2022

Effective from the Academic Year 2024-2025

Suggested References:

S r .No.	References
1	R Cookbook, PaulTeetor, Oreilly: R Cookbook [R CKBK] [Paperback] R Cookbook [RKBK] [Paperback]Mar 31, 2011 by Paul Teetor
2	Nina Zumel, John Mount, "Practical Data Science with R", Manning Publications, 2014
3	R in Action,RobKabacoff, Manning: R in Action: Data Analysis and Graphics with RNov 5, 2018	Unabridged by Robert Kabacoff and Dale Ogden
4	Online Sources: https://www.geeksforgeeks.org/r-programming-language-introduction/ https://www.programiz.com/r

****

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)
AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE'A^+' UGC – MHRD, Govt of India – June 2022
Effective from the Academic Year 2024-2025

Master of Science (Data Science)
Semester –2

Course Code	PG02CMDS02	Title of the Course	Advanced RDBMS
Total Credits of the Course	04	Hours per Week	04

Course Objectives	- Learn about Database Life Cycle and administration. - Understand the concepts of Transactions, concurrency control, recovery, security and integrity - Understand various database system architectures - Learn to work with PL/SQL.Understand how to handle errors and use triggers

Course Outcomes	- Explain the Database Life Cycle and the process of database administration. - Use the concepts of Transactions, concurrency control, recovery, security and integrity.

	Course Content	
Unit	**Description**	**Weight age (%)**
1.	**Introductions of RDBMS** - History of Data base Systems. Data base System Applications, data base System VS file System. - The Database Life Cycle (DBLC): Initial Study of The Database, Database Design, Implementation and Loading, Testing And Evaluation, Operations and Maintenance. - Database Administration: The Role of Database Administrator, Modeling Enterprise Data, Planning for Databases, Managing Data Security, Backing Up Databases, Controlling Concurrent Access, Managing Data Quality, Data Dictionaries and Repositories.	25%
2.	**Data Models** ER Model – Relational Model – Other Models – Database Languages – DDL – DML. Introduction to the Relational Model – Integrity Constraint Over relations – Enforcing Integrity constraints – Querying relational data – Logical data base Design	25%

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE'A^+' UGC – MHRD, Govt of India – June 2022

Effective from the Academic Year 2024-2025

		• Query Processing and Optimization: Evaluation of Relational Algebra Expressions, Query Equivalence, Join strategies, Query Optimization Algorithms.	
3.	**PL/SQL**	• Introduction, the SQL execution environment, thePL/SQL syntax, block structure – declarative part, executablepart, exception handling part, variable declaration using%type, % rowtype, if statements, looping structures, oracletransactions, cursors & its types, cursor attributes, nesting ofcursors, parameterized cursors, error handling in SQL. • Concurrency control: Locks, implicit locking, levels oflocks, explicit locking, select for…update statement, usinglock table statement.	**25%**
4.	**Error handling,Stored Procedures and Functions, Triggers**	• Error handling: user named exception handlers for i/ovalidation and business rule validation. • Stored Procedures and Functions: creating a storedprocedure or function, syntax for declaration, execution andexception handling parts, advantages of using procedure orfunction. Deleting a procedure or function. • Database Triggers: Introduction, use, database triggers v/s procedures, database triggers v/s declarative integrity constraints, how to apply triggers. Types of triggers, Creating a trigger, deleting trigger.	**25%**

Teaching-Learning Methodology	Material for this course will be presented using multiple teaching approach: lecture and discussion, exploration and inquiry, cooperative group work, demonstrations, and presentations

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified

GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE'A^+' UGC – MHRD, Govt of India – June 2022

Effective from the Academic Year 2024-2025

EvaluationPattern

Sr. No.	Details of the Evaluation	Weightage
1.	Internal Written	**30%**
2.	Internal Continuous Assessment in the form of Quizzes, Seminars, Assignments, Attendance	
3.	External Examination	**70%**

Suggested References:

S r .No.	References
1	"Database System Concepts", Abraham Silberschatz, Henry Korth, S. Sudarshan, McGraw Hill
2	"Database Management System", Rajesh Narang, PHI
3	"SQL PL/SQL: The Programming Language of Oracle", Ivan Bayross, BPB Publications
4	Online Sources: https://www.javatpoint.com/what-is-rdbms https://www.tutorialspoint.com/sql/sql-rdbms-concepts.htm

****

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE'A^+' UGC – MHRD, Govt of India – June 2022
Effective from the Academic Year 2024-2025

Master of Science (Data Science)
Semester –2

Course Code	PG02CMDS03	Title of the Course	Big Data Analytics
Total Credits of the Course	04	Hours per Week	04

Course Objectives	• Understand the key issues in big data management and its associated applications in intelligent business and scientific computing. • Acquire fundamental enabling techniques and scalable algorithms like Hadoop,Map Reduce and NO SQL in big data analytics. • Interpret business models and scientific computing paradigms and apply software tools for big data analytics..

Course Outcomes	• Achieve adequate perspectives of big data analytics in various applications likerecommender systems and social media applications. • Evaluate and apply appropriate principles, techniques and theories to large-scaledata science problems using various databases withanalytics and visualizations.

	Course Content	
Unit	**Description**	**Weight age (%)**
1.	**Introduction to Big Data** • Definition – Big Data and Data Science • Big Data and its importance, Characteristics of Big Data, Limitation of Conventional Data Processing Approaches, Need of big data frameworks, Big data analytics, Limitations of Big Data and Challenges, Big data applications • Defining data science and big data, Recognizing the different types of data, Gaining insight into the data science process • Introducing the fields of data science and big data • Overview – Defining goals – Retrieving data – Data preparation – Data exploration – Data modeling – Presentation.	25%

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND
(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)
AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR
An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE'A^+' UGC – MHRD, Govt of India – June 2022
Effective from the Academic Year 2024-2025

2.	**Handling and Processing Big Data** • Types of Data Sources Sampling, Types of Data Elements ,Visual Data Exploration and Exploratory Statistical Analysis, MissingValues, Outlier Detection and Treatment, Standardizing Data,Categorization, Weights of Evidence Coding, Variable Selection,Segmentation • Problems when handling large data – General techniques for handling large data	**25%**

3.	**Overview of Big data Analytics** • Meaning and Characteristics of Big Data Analytics • Need of Big Data Analytics • Classification of Analytics • Importance of Big Data Analytics • Steps in big data analytics – Distributing data storage and processing with Frameworks • NO SQL in Big data Analytics	**25%**
4.	**Advanced Big data Analytics and data visualization** • Hadoop: Basic Concepts of Hadoop and its features -The Hadoop Distributed File System (HDFS)- Anatomy of a Hadoop Cluster - Hadoop cluster modes - Hadoop Architecture, Hadoop Storage • Methodological Challenges and Problems with Big data Analytics • Introduction to data visualization – Data visualization options – Filters – MapReduce –Dashboard development tools – Creating an interactive dashboard with dc.js	**25%**

Teaching-Learning Methodology	Material for this course will be presented using multiple teaching approach: lecture and discussion, exploration and inquiry, cooperative group work, demonstrations, and presentations

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE'A^+' UGC – MHRD, Govt of India – June 2022

Effective from the Academic Year 2024-2025

EvaluationPattern		
Sr. No.	Details of the Evaluation	Weightage
1.	Internal Written	**30%**
2.	Internal Continuous Assessment in the form of Quizzes, Seminars, Assignments, Attendance	
3.	External Examination	**70%**

Suggested References:	
S r .No.	References
1	Data Science from Scratch: First Principles with Python, Joel Grus, O'Reilly, 1st edition,2015
2	Boris lublinsky, Kevin t. Smith Alexey Yakubovich, "Professional Hadoop Solutions". Wiley, ISBN : 9788126551071, 2015.
3	Chris Eaton, Dirk Deroos et al., "Understanding Big Data", McGraw Hill , 2010.
4	Tom White, "HADOOP" : The definitive Guide", O Reilly 2012.
5	Online Sources: https://www.geeksforgeeks.org/what-is-big-data-analytics/ https://www.techtarget.com/searchbusinessanalytics/definition/big-data-analytics

****

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE 'A$^+$' UGC – MHRD, Govt of India – June 2022
Effective from the Academic Year 2024-2025

Master of Science (Data Science)
Semester –2

Course Code	PG02CMDS04	Title of the Course	Artificial Intelligence and ML
Total Credits of the Course	**04**	**Hours per Week**	**04**

Course Objectives	• To learn the basic concepts related to Artificial Intelligence and Knowledge Based Systems. • To understand the concept of fuzzy Logic and its applications. • To introduce the fundamental concepts of machine learning and its applications • To understand the deep learning architectures

Course Outcomes	• To learn the classification, clustering and regression based AI & machine learning algorithms. • To understand the methods of solving real life problems using the machine learning technique.

Course Content		
Unit	**Description**	**Weight age (%)**
1.	**Artificial Intelligence (AI) and Knowledge Based Systems (KBS)** • Artificial Intelligence, Applicationof AI, AI Problems, Problem Formulation, Intelligent Agents,Types of Agents, Agent Environments, PEAS representationfor an Agent, Architecture of Intelligent agents. Reasoning andLogic, Prepositional logic, first order logic, Using First-orderlogic, Inference in First-order logic, forward and backward Chaining. • KBS Structure, Components of KBS, Categories of KBS, • Knowledge-Based Shell, Advantages, Limitations and Applications • of KBS, Knowledge Acquisition, Knowledge Update • Factual and Procedural Knowledge Representations, Knowledge Based Systems Development Model	25%

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND
(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)
AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR
An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE 'A^+' UGC – MHRD, Govt of India – June 2022
Effective from the Academic Year 2024-2025

2.	**Fuzzy Logic, Connectionist Models** • Fuzzy Logic and Fuzzy Sets, Membership Functions, Fuzzification and Defuzzification, Operations on Fuzzy Sets,Fuzzy Functions and Linguistic Variables • Introduction to ANN, Biological Neuron and Artificial Neuron,Hopfield model of ANN, Parallel relaxation, Linearly Separable Problems, Single perceptron	**25%**
3.	**Machine Learning - I** • Machine Learning basics,Applications of ML,Data Mining Vs Machine Learning vs BigData Analytics. • Supervised Learning- Naïve Base Classifier, , Classifying withk-Nearest Neighbour classifier, Decision Tree classifier, Naïve Bayes classifier. • Unsupervised Learning - Grouping unlabeled items usingk-means clustering, Association analysis with the Apriorialgorithm Introduction to reinforcement learning	**25%**
4.	**Machine Learning - II** • Principal component analysis – Linear discriminant analysis - Independent component analysis. K-means clustering - fuzzy k-means clustering Expectation-maximization algorithm-Gaussian mixture models –auto associative neural network. • Perceptron and backpropagation neural network - k-nearest neighbor rule. Support vector machine: multicategory generalizations – Regression Decision trees: classification and regression tree – random forest.	**25%**

Teaching-Learning Methodology	Material for this course will be presented using multiple teaching approach: lecture and discussion, exploration and inquiry, cooperative group work, demonstrations, and presentations

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE 'A^+' UGC – MHRD, Govt of India – June 2022

Effective from the Academic Year 2024-2025

EvaluationPattern		
Sr. No.	Details of the Evaluation	Weightage
1.	Internal Written	
2.	Internal Continuous Assessment in the form of Quizzes, Seminars, Assignments, Attendance	**30%**
3.	External Examination	**70%**

Suggested References:	
S r .No.	References
1	EthemAlpaydin, "Introduction to Machine Learning", 3rd Edition, MIT Press, 2014.
2	C. M. Bishop, "Pattern Recognition and Machine Learning", Springer, 2006
3	Kevin P. Murphy, "Machine Learning: A Probabilistic Perspective", MIT Press, 2012.
4	Online Sources: https://www.javatpoint.com/difference-between-artificial-intelligence-and-machine-learning https://www.geeksforgeeks.org/difference-between-machine-learning-and-artificial-intelligence/

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE 'A+' UGC – MHRD, Govt of India – June 2022
Effective from the Academic Year 2024-2025

Master of Science (Data Science)
Semester –2

Course Code	PG02CMDS05	Title of the Course	Computational Statistics
Total Credits of the Course	**04**	**Hours per Week**	**04**

Course Objectives	- To enable the students to understand the fundamentals of statistics to apply descriptive measuresand probability for data analysis. - Understand the science of studying & analyzing numbers. - Identify and use various visualization tools for representing data. - Describe various statistical formulas. - Compute various statistical measures.

Course Outcomes	- To understand and produce rigorous statistical analysis including estimation, confidence intervals, hypothesis testing, regression, logistic regression. - Basic properties of optimization describe various statistical formulas.

	Course Content	
Unit	**Description**	**Weight age (%)**
1.	**Statistics and Probability, Collection of Data** - Introduction to Statistics – Origin of Statistics, Features of Statistics, Scope of Statistics, - Statistics vs. probability, sample vs population; - Summary statistics: Mean, SD, Median, IQR; - Graphical Summary: Pie Charts, Histograms, Box-plots - Sample space, event, probability, Conditional Probability, Bayes's Theorem, Independence	**25%**
2.	**Random variables and probability distributions, Commonly used distributions** - Random variables and probability distribution, Expected values and standard deviation, Probability density functions - Binomial distribution, Hypergeometric, negative bionomial, Poisson distributions, Normal distributions, - Normal approximations to data histograms, Exponential and Gammas distributions, Quantile-Quantile plot	**25%**

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND
(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)
AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR
An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE 'A+' UGC – MHRD, Govt of India – June 2022
Effective from the Academic Year 2024-2025

3.	**Joint Distributions and Random Samples** • Discrete joint distribution • Joint densities • Covariance and correlation • Multivariate random variables • Square root law • Central limit theorem • Concepts and Methods of Estimation.	**25%**
4.	**Hypothesis Testing** • Basic concept Hypothesis Testing • Test for population mean • t-test • Test for population proportion • Comparisons of two treatments • Inference based on two samples, Two-sample z-test, Two-sample t-test • Difference between two proportions, Analysis of paired data.	**25%**

Teaching-Learning Methodology	Material for this course will be presented using multiple teaching approach: lecture and discussion, exploration and inquiry, cooperative group work, demonstrations, and presentations

EvaluationPattern		
Sr. No.	Details of the Evaluation	Weightage
1.	Internal Written	**30%**
2.	Internal Continuous Assessment in the form of Quizzes, Seminars, Assignments, Attendance	
3.	External Examination	**70%**

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE'A^+' UGC – MHRD, Govt of India – June 2022

Effective from the Academic Year 2024-2025

Suggested References:	
S r .No.	**References**
1	Banfield J.(1999), Rweb: Web-based Statistical Analysis, Journal of Statistical Software
2	Bhattacharya,G.K. and Johnson, R.A.(19977), Statistical Concepts and Methods, New York, John Wiley & Sons.
3	Statistics and Data Analysis, A.Abebe, J. Daniels, J.W.Mckean, December 2000.
4	Introduction to Statistics, David M. Lane.
5	1. http://onlinestatbook.com/Online_Statistics_Education.pdf 2. https://textbookcorp.tn.gov.in/Books/12/Std12-Stat-EM.pdf 3. https://3lihandam69.files.wordpress.com/2015/10/introductorystatistics.pdf

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND
(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)
AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR
An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE'A^+' UGC – MHRD, Govt of India – June 2022
Effective from the Academic Year 2024-2025

Master of Science (Data Science)
Semester –2

Course Code	**PG02CMDS06**	Title of the Course	Practicals of R and RDBMS
Total Credits of the Course	**05**	**Hours per Week**	**10**

Course Objectives	- To provide practical experience of R programming - To provide practical experience off working with Database Tool. - To familiarize students with Data Science Libraries in R. - Database creation and operation in Database Tool. - Data science application in R.

Course Outcomes	- Gain knowledge of developing R programs and applications. - Create and manipulate Database Tool functionality for data analysis.

Course Content	
Description	**Weight age (%)**
Part-1 : Practical based on PG02CMDS01	**50%**
Part-2 : Practical based on PG02CMDS02	**50%**

Teaching-Learning Methodology	Blended learning approach incorporating traditional classroom teaching as well as online / ICT-based teaching practices

C P PATEL AND F H SHAH COMMERCE (AUTONOMOUS) COLLEGE, ANAND

(Managed by SARDAR PATEL EDUCATION TRUST,ANAND)

AFFILIATED TO SARDAR PATEL UNIVERSITY, V V NAGAR

An ISO 9001 2015 Certified / An ISO 14001-2015 Certified / An ISO 21001-2018 Certified
GUJARAT INSTITUTIONAL RATING FRAMEWORK (4 STAR)

AAA Reaccredited CGPA 3.56 – GRADE A^+ KCG-Dept of Edu. Got of Gujarat-April 2017

NAAC Reaccredited - CGPA 3.30 - GRADE 'A$^+$' UGC – MHRD, Govt of India – June 2022

Effective from the Academic Year 2024-2025

EvaluationPattern	
Details of the Evaluation	Weightage
Internal Written / Practical Examination (As per CBCS R.6.8.3)	**15%**
Internal Continuous Assessment in the form of Practical, Viva-voce, Quizzes, Seminars, Assignments, Attendance (As per CBCS R.6.8.3)	**15%**
External Examination	**70%**