



# M.Sc. DATA ANALYTICS

## PROGRAMME SPECIFIC OUTCOME

- PSO1:** Learn Data analytics tools and techniques, SQL databases, the languages of R and Python, data visualizations, statistics and predictive analytics in a business environment.
- PSO2:** Explore basic programming skills and how to wrangle data from diverse sources and shape it to enable data-driven applications.
- PSO3:** Ability to pre-process and analyse various data, latest trends in technology development and thereby innovate new ideas and solutions to existing problems.
- PSO4:** Data analytics is the science of analyzing raw data in order to make conclusions about that information. Data analytics techniques can reveal trends and metrics that would otherwise be lost in the mass of information.
- PSO5:** Ability to apply mathematical / statistical implementation, manage Projects / Internship and function effectively as an individual, and as a member or leader in diverse teams.
- PSO6:** Apply knowledge of mathematics, statistics, science and computing/programming skill to model the software applications, configure software platform and analyze real time data in a heterogeneous domain.

# COURSE OUTCOME

## SJCSDC01 : PRINCIPLES OF DATA SCIENCE

- SJCSDC01.1 Understand Data sources, generations, data formats, Data Evolution, Data from various domains.
- SJCSDC01.2 Understand Big Data Characteristics What, Why, When, Limitation of traditional approaches and models. Map Big Vs to Data Domains.
- SJCSDC01.3 Understand Big Data Processing platform, frameworks, Hadoop, Spark, storage models – Hbase.
- SJCSDC01.4 Programming Model of Big Data MapReduce, Why MapReduce, Limitations of Traditional Models.
- SJCSDC01.5 Analyze various domains of Big Data Characteristics, Platform, Programming Model.
- SJCSDC01.6 Design Big Data framework ecosystem, and data processing framework of multidisciplinary domains.

## SJCSDC02: MATHEMATICS FOR COMPUTING

- SJCSDC02.1 Understand the principles of probability, frequency distribution measures.
- SJCSDC02.2 Understand the correlation and regression, hypothesis test, sampling techniques for specific applications.
- SJCSDC02.3 Apply probabilistic models and distribution models.
- SJCSDC02.4 Apply hypothesis testing and regression models for specific Domain.
- SJCSDC02.5 Illustrate statical model and infer.
- SJCSDC02.6 Design statistical models for specific domains.

## SJCSDC03 : DATA MINING

- SJCSDC03.1 Understand data mining tools and techniques and big data for various domains.
- SJCSDC03.2 Apply various data mining, text mining, web mining algorithms for real time applications.
- SJCSDC03.3 Analyze unsupervised and supervised algorithms for real world applications.
- SJCSDC03.4 Illustrate the mining techniques like association,

SJCSDC03.5	classifications and clustering on datasets.
SJCSDC03.6	Apply R programming packages for mining data. Compare various approaches of data mining algorithms.

## SJCSDC04: PYTHON PROGRAMMING

SJCSDC04.1	Understand the object-oriented concepts and control structures.
SJCSDC04.2	Understand the basic programming structure-list, dictionary, tuple.
SJCSDC04.3	Apply OOPs concept for designing software applications.
SJCSDC04.4	Understand the visualization methods, packages, statistical packages and other packages for building data models.
SJCSDC04.5	Design and analyze dataset applying statistical models, visualization and models using various tools.
SJCSDC04.6	Design data analytic model using the packages in python and provide inference for multidisciplinary domains.

## SJCSDC05: DESIGN AND ANALYSIS OF ALGORITHMS AND OBJECT

SJCSDC05.1	Identify classes and objects from the given problem description and create classes and objects using C++.
SJCSDC05.2	Code reusability and extensibility by means of Inheritance and Polymorphism.
SJCSDC05.3	Differentiate among various algorithmic approaches.
SJCSDC05.4	Design algorithms for problem solving by using the suitable algorithmic technique.
SJCSDC05.5	Analyze a given algorithm for its efficiency based on time and space it occupies.
SJCSDC05.6	Apply optimization techniques for improving the performance of algorithms.

## SJCSDC06: ADVANCE DATABASE MANAGEMENT

SJCSDC06.1	Explain the structure and model of the relational database system.
SJCSDC06.2	Design multiple tables and using group queries.
SJCSDC06.3	Design a database based on a data model normalization to a specified level.
SJCSDC06.4	Mongo DB and Operators.

## SJCSDC07: OPERATIONS RESEARCH

SJCSDC07.1	Identify the Assignment problem and to optimize in engineering fields
SJCSDC07.2	Solve linear programming techniques to optimization problems arising in all Computer fields.
SJCSDC07.3	Solve Integer linear programming techniques to optimization problems arising in all Computer fields.
SJCSDC07.4	Use Dynamic programming approach to real time problems.
SJCSDC07.5	Compare different Transportation algorithms
SJCSDC07.6	Write a case study using any Operations Research methods to get optimal solution for an organization

## SJCSDC08: R PROGRAMMING FOR DATA ANALYTICS

SJCSDC08.1	Understand the basic programming structure of R– Data frame, Matrix, List, Packages and Functions.
SJCSDC08.2	Understand various visualization models and their inference – Scatter plots, histogram, boxplot.
SJCSDC08.3	Apply statistical functions, models and their Inferences – Central tendency measure, Range, Variance, Standard Deviation.
SJCSDC08.4	Use data normalization for domain specific dataset.
SJCSDC08.5	Apply distribution models, Regression models and ANOVA.
SJCSDC08.6	Design data model, visualization and inference of dataset to gain insights.

## SJCSDC09: DATA VISUALIZATION

SJCSDC09.1	Understand the concepts of visualization.
SJCSDC09.2	Understand the methods for visualizing data in D3j, c3j, and Tableau.
SJCSDC09.3	Apply Visualization methods for different data domains.
SJCSDC09.4	Design Interactive Charts based on Data.
SJCSDC09.5	Distinguish and Suggest the appropriate data visualization tools for domain specific applications
SJCSDC09.6	Design an Interactive data visualization story board for data

## SJCSDC10: LINEAR ALGEBRA FOR MACHINE LEARNING

SJCSDC10.1	Understand the basics of Linear Programming constructs.
SJCSDC10.2	Apply vector spaces and their applications in Machine Learning
SJCSDC10.3	Understand the concepts of matrix and Gaussian Elimination.
SJCSDC10.4	Understand the concepts of differential equations.
SJCSDC10.5	Apply the concepts of Linear Algebra in Machine Learning Algorithms

## SJCSDC11: CLOUD COMPUTING

SJCSDC11.1	Understand Cloud Service Models and architectures.
SJCSDC11.2	Apply the concepts of Virtualization
SJCSDC11.3	Understand cloud programming models
SJCSDC11.4	Analyze various cloud services and features of cloud service providers.
SJCSDC11.5	Critically analyze case studies to derive the cloud models for developing and deploying cloud-based applications.
SJCSDC11.6	Understand the Risks and Management of Cloud environments.
SJCSDC11.7	Compare the cloud service models and providers to suit business needs.

## SJCSDC12: MAPREDUCE PROGRAMMING

SJCSDC12.1	Understand MapReduce Processing architectures.
SJCSDC12.2	Configure and setup MapReduce Processing architectures Ecosystem – Hadoop, Spark, Pig and Hive
SJCSDC12.3	Understand and write MapReduce program using Pig and Hive, spark.
SJCSDC12.4	Analyze dataset using Pig and Hive.
SJCSDC12.5	Critically analyze case studies for and suggest MapReduce Programming models based on domains and applications
SJCSDC11.6	Distinguish Hadoop and SPARK
SJCSDC11.7	Design and setup a Big Data Analytics Ecosystem for specific Business scenarios.

## SJCSDC13: MACHINE LEARNING

SJCSDC13.1	Understand the concepts of machine learning
SJCSDC13.2	Understand the theoretical concepts of probabilistic and linear methods
SJCSDC13.3	Distinguish Supervised, Unsupervised and semi supervised learning.
SJCSDC13.4	Understand and Apply the algorithms for a given specific problem in a specific tool.
SJCSDC13.5	Suggest Supervised, Unsupervised and semi supervised algorithms for specific application
SJCSDC13.5	Design a Machine Learning system for any specific domain.

## SJCSDL01: PYTHON AND C++ PRACTICAL

SJCSDL01.1	Implement Classes and Objects, Constructors and Destructors with array of Objects, Passing and returning parameters as objects by reference, String manipulation functions, different types of inheritances like Multiple, Multilevel and Hybrid..
SJCSDL01.2	Demonstrate Function Overloading, overload different operators – incr and decr operators with post and pre forms, use of Virtual Functions
SJCSDL01.3	Read, write, and execute simple Python programs, write simple Python programs for solving problems
SJCSDL01.4	Decompose a Python program into functions, lists etc
SJCSDL01.5	Read and write data from/to files in Python Programs.
SJCSDL01.6	Underline the use of package.

## SJCSDL02: R PROGRAMMING AND ADBMS PRACTICAL

- SJCSDL02.1 Configure R environment for development of application, Develop functional application in using r scripting.
- SJCSDL02.2 Develop application which process CSV, XML, JSON, XML, Develop application with visualization - bar Chart, line graph, box plot , Histogram and Scatter Plots.
- SJCSDL02.3 Apply basic statistical operation using R
- SJCSDL02.4 Use an SQL interface of a multi-user relational DBMS package to create, secure, populate, maintain, and query a database..
- SJCSDL02.5 Formulate query, using SQL, solutions to a broad range of query and data update problems
- SJCSDL02.6 Transform an information model into a relational database schema and to use a data definition language and/or utilities to implement the schema using a DBMS.

## SJCSDL03: MAPREDUCE AND MACHINE LEARNING PRACTICAL

- SJCSDL03.1 Implement a learning algorithm in Python, Predict classification or regression outcomes, with scikit-learn models in Python.
- SJCSDL03.2 Solve Non-linear problems using SVM. Apply machine learning algorithms to solve problems of moderate complexity.
- SJCSDL03.3 Apply and analyze dataset with machine learning algorithms
- SJCSDL03.4 Apply HIVE, Pig commands, data analysis and word count using PIG
- SJCSDL03.5 Spark RDD Commands, Data Analysis using Spark PySpark Commands  
Spark SQL commands.
- SJCSDL03.6 Word count using Spark, Word count using Map Reduce