Faculty of Engineering & Technology
Master of Technology (M. Tech)
(W. E. F.: 2023-24)

Document ID: SUTEFETM-01

Name of Faculty	:	Faculty of Engineering & Technology
Name of Program	:	Master of Technology (M. Tech)
Course Code	:	2MSE06
Course Title	:	Data Science
Type of Course	:	VA
Year of Introduction	:	2023-24

Prerequisite	:	Mathematics & Statics and Programming & Computer Science
Course Objective	:	The course objectives of a data science course may vary depending on the institution or program offering the course. However, here are some common objectives that are often covered in data science courses: Introduction to Data Science, Data Acquisition and Preprocessing, Machine Learning Algorithms etc
Course Outcomes	:	At the end of this course, students will be able to:
	CO1	Understanding of Data Science Fundamentals.
	CO2	Proficiency in Programming: Data science courses often emphasize programming skills, particularly in languages such as Python or R.
	CO3	Data Acquisition and Cleaning: Students should learn techniques for acquiring data from various sources, such as databases, APIs, and web scraping.
	CO4	Exploratory Data Analysis: Students should be able to apply exploratory data analysis techniques to gain insights from raw data.
	CO5	Data scientists need to effectively communicate their findings and insights to various stakeholders.
	CO6	As data science often deals with large datasets, students should be introduced to big data technologies like Apache Hadoop, Apache Spark, and distributed computing frameworks.

Teaching and Examination Scheme

Teaching Scheme (Contact Hours)			Credits	Examination Marks				
				Theory Marks		Practical Marks		Total Marks
L	T	P	C	SEE	CIA	SEE	CIA	
3	0	2	4	70	30	30	20	150

Legends: ***L****-Lecture;* ***T****–Tutorial/Teacher Guided Theory Practice;* ***P*** *– Practical,* ***C*** *– Credit,* ***SEE*** *– Semester End Examination,* ***CIA*** *- Continuous Internal Assessment (It consists of Assignments/Seminars/Presentations/MCQ Tests, etc.))*

Course Content

Unit No.	Topics	Teaching Hours	Weightage	Mapping with CO
1	**An Introduction to core concepts & technologies:** Introduction, Terminology, data science process, data science toolkit, Types of data, Example applications	06	10%	CO1
2	**Data collection and management:** Introduction, Sources of data, Data collection and APIs, Exploring and fixing data, Data storage and management, Using multiple data sources	07	15%	CO2
3	**Data analysis:** Introduction, Terminology and concepts, Introduction to statistics, Central tendencies and distributions, Variance, Distribution properties and arithmetic, Samples/CLT, Basic machine learning algorithms, Linear regression, SVM, Naive Bayes.	10	25%	CO3
4	**Data visualisation:** Introduction, Types of data visualisation, Data for visualisation: Data types, Data encodings, Retinal variables, Mapping variables to encodings, Visual encodings	11	25%	CO4
5	Applications of Data Science, Technologies for visualisation, Bokeh (Python)	07	15%	CO5
6	Recent trends in various data collection and analysis techniques, various visualization techniques, application development methods of used in data science.	07	10%	CO6

Suggested Distribution of Theory Marks Using Bloom's Taxonomy						
Level	Remembrance	Understanding	Application	Analyse	Evaluate	Create
Weightage	**40**	**20**	**20**	**10**	**-**	**10**

NOTE: This specification table shall be treated as a general guideline for the students and the teachers. The actual distribution of marks in the question paper may vary slightly from above table.

Suggested List of Experiments/Tutorials

Sr. No.	Name of Experiment/Tutorial	Teaching Hours
1	Minimum 10 experiments based on the contents.	10
2	Mini Project in a group of max. 3 students	02
3	Writing a research paper on selected topic from content with latest research issues in that topic.	01

Reference Books

Sr. No.	Name of Reference Books
1	Data Mining Concepts and Techniques by Jiawei Han, Micheline Kamber and Jian Pei
2	Statistics and Data Analysis by A. Abebe