

UNIVERSITY OF MADRAS
M.Sc. DEGREE PROGRAMME IN COMPUTER SCIENCE
SYLLABUS WITH EFFECT FROM 2023-2024

Title of the Paper	Data Science and Analytics Practical		
Core–XII - Practical	II Year & III Semester	Credit:3	536C3D

Objectives:

- Understand Hadoop Distributed File System and examine MapReduce Programming.
- Explore Hadoop tools and manage Hadoop with Ambari.
- Appraise the role of Business intelligence and its applications across industries.
- Assess core data mining techniques for data analytics
- Identify various Text Mining techniques

Outcomes:

1.	To Describe the key issues in Big Data Management and experiment with the Hadoop framework.	K1, K2
2.	To Explain the structure and unstructured data by using NoSQL commands.	K2, K3
3.	To Apply scientific computing algorithms for finding similar items and clustering.	K3, K4
4.	To Test fundamental enabling techniques and scalable algorithms for data stream mining.	K4, K5
5.	To Develop problem solving and critical thinking skills in fundamental enable techniques like Hadoop & MapReduce	K6
K1-Remember;K2-Understand;K3-Apply;K4-Analyze;K5-Evaluate; K6-Create		

LIST OF PROGRAMS

1. Installation of Hadoop Framework, its components and study the HADOOP ecosystem
2. Write a program to implement word count program using MapReduce
3. Experiment on Hadoop Map-Reduce / PySpark: -Implementing simple algorithms in Map-Reduce: Matrix multiplication
4. Install and configure MongoDB/ Cassandra/ HBase/ Hypertable to execute NoSQL Commands.
5. Implementing DGIM algorithm using any Programming Language/ Implement Bloom Filter using any programming language
6. Implement and Perform Streaming Data Analysis using flume for data capture, PYSpark / HIVE for data analysis of twitter data, chat data, weblog analysis etc.
7. Implement any one Clustering algorithm (K-Means/CURE) using Map-Reduce.
8. Implement Page Rank Algorithm using Map-Reduce.
