

UNIVERSITY OF MADRAS  
MASTER OF COMPUTER APPLICATIONS (MCA) DEGREE PROGRAMME  
SYLLABUS WITH EFFECT FROM 2023-2024

Title of the Paper	Big Data Analytics		
Core - VI Theory	I Year & II Semester	Credit:4	435C2B

**Course Objectives**

- To introduce big data tools & Information Standard formats.
- To understand the basic concepts of big data.
- To learn Hadoop, HDFS and Map Reduce concepts.
- To teach the importance of NoSQL.
- To explore the big data tools such as Hive, HBase and Pig.

**UNIT I:**Big Data and Analytics: Classification of Digital Data: Structured Data- Semi Structured Data and Unstructured Data.

Introduction to Big Data: Characteristics – Evolution – Definition - Challenges with Big Data - Other Characteristics of Data - Big Data - Traditional Business Intelligence versus Big Data - Data Warehouse and Hadoop.

Environment Big Data Analytics: Classification of Analytics – Challenges - Big Data Analytics important - Data Science - Data Scientist - Terminologies used in Big Data Environments – Basically Available Soft State Eventual Consistency - Top Analytics Tools

**UNIT II:** Technology Landscape:NoSQL, Comparison of SQL and NoSQL, Hadoop -RDBMS Versus Hadoop - Distributed Computing Challenges – Hadoop Overview - Hadoop Distributed File System - Processing Data with Hadoop - Managing Resources and Applications with Hadoop YARN - Interacting with Hadoop Ecosystem

**UNIT III:** MongoDB and Mapreduce Programming:MongoDB: Mongo DB - Terms used in RDBMS and Mongo DB - Data Types - MongoDB Query Language.

MapReduce: Mapper – Reducer – Combiner – Partitioner – Searching – Sorting – Compression

**UNIT IV:** Hive: Introduction – Architecture - Data Types - File Formats - Hive Query Language Statements – Partitions – Bucketing – Views - Sub- Query – Joins – Aggregations - Group by and Having – RCFile - Implementation - Hive User Defined Function - Serialization and Deserialization.

**UNIT V:** Pig: Introduction - Anatomy – Features – Philosophy - Use Case for Pig - Pig Latin Overview - Pig Primitive Data Types - Running Pig - Execution Modes of Pig - HDFS Commands - Relational Operators - Eval Function - Complex Data Types - Piggy Bank - User-Defined Functions - Parameter Substitution – Diagnostic Operator - Word Count Example using Pig - Pig at Yahoo! - Pig Versus Hive

**Text Book:**

1. Seema Acharya, Subhashini Chellappan, “Big Data and Analytics”, Wiley Publications, First Edition,2015

# UNIVERSITY OF MADRAS

## MASTER OF COMPUTER APPLICATIONS (MCA) DEGREE PROGRAMME SYLLABUS WITH EFFECT FROM 2023-2024

### Reference Book:

1. Judith Huruwitz, Alan Nugent, Fern Halper, Marcia Kaufman, “Big data for dummies”, John Wiley & Sons, Inc. (2013)
2. Tom White, “Hadoop The Definitive Guide”, O’Reilly Publications, Fourth Edition, 2015
3. Dirk Deroos, Paul C.Zikopoulos, Roman B.Melnky, Bruce Brown, Rafael Coss, “Hadoop For Dummies”, Wiley Publications, 2014
4. Robert D.Schneider, “Hadoop For Dummies”, John Wiley & Sons, Inc. (2012)
5. Paul Zikopoulos, “Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data, McGraw Hill, 2012 Chuck Lam, “Hadoop In Action”, Dreamtech Publications, 2010

### Course Outcomes

On the successful completion of the course, students will be able to

CO1	To understand, illustrate and evaluate the concepts and techniques of Data Science, Big Data Analytics and its tools.	K1-K6
CO2	To collaborate, apply and review the computing for big data in Hadoop, and NoSQL environment.	K1-K6
CO3	To comprehend, implement and review the concepts of data science and big data analytics projects using MapReduce, and MongoDB.	K1-K6
CO4	To understand, use and analyze the concepts of big data analytics projects using HIVE database.	K1-K6
CO5	To illustrate, develop and review the concepts of PIG database in Hadoop environment.	K1-K6

K1- Remember, K2- Understand, K3- Apply, K4- Analyze, K5 Evaluate, K6- Create

### Mapping with Programme Outcomes:

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10
CO1	S	-	M	-	M	L	-	-	-	-
CO2	S	-	S	-	S	L	-	-	-	S
CO3	S	-	S	-	S	L	-	-	-	S
CO4	S	-	S	-	S	L	-	-	-	S
CO5	S	-	S	-	S	L	-	-	-	S

S- Strong; M-Medium; L-Low