

## Data Science

Course Title	Data Science
Credit	2
Teaching per Week	3 hrs.
Minimum weeks per Semester	15 (Including Class work, examination, preparation etc.)
Purpose of Course	This course aims at introducing the students into the world of Data Science. Understand the basic concepts of Data Science, Life Cycle, Big Data, Advance Database, Data Warehouse
Course Objective	Provide fundamental knowledge about Data Science in real world.
Pre-requisite	Basic Knowledge of Database Management System.
Course Out come	After successful completion of the course a student will be <ul style="list-style-type: none"> <li>• Able to understand about need of Data Science</li> <li>• Able to understand about Big Data and Data Warehouse</li> <li>• Able to understand about Advance Database System</li> <li>• Differentiate Data Science and Data Analysis</li> </ul>

Unit	Content
<b>1</b>	<b>Introduction to Data Science</b> <ul style="list-style-type: none"> <li>• What is Data Science</li> <li>• Need of data Science</li> <li>• Business Intelligence v/s Data Science</li> <li>• Component of data science</li> <li>• Data Science Life Cycle</li> <li>• Tools for Data Science</li> </ul>
<b>2</b>	<b>Introduction to Big Data</b> <ul style="list-style-type: none"> <li>• Classification of Data</li> <li>• Definition and Evolution of Big Data</li> <li>• Challenges of Big Data</li> <li>• Characteristics of Big Data</li> <li>• Big Data Applications</li> <li>• Big Data Architecture.</li> </ul>
<b>3</b>	<b>Introduction to HADOOP</b> <ul style="list-style-type: none"> <li>• Apache Hadoop</li> <li>• Hadoop Architecture &amp; Hadoop Ecosystem</li> <li>• Hadoop ecosystem components</li> <li>• MapReduce</li> <li>• HDFS</li> <li>• YARN</li> </ul>
<b>4</b>	<b>Advance Database System</b> <ul style="list-style-type: none"> <li>• Types of Databases</li> </ul>

- Introduction of NoSQL
- Need of NoSQL
- Advantages of NoSQL
- SQL vs NoSQL
- Introduction to different type of NoSQL databases.

## **5 Data Analytics**

- What is Data analytics
- Use of Data Analytics
- Data Analytics Life Cycle
- Types of analysis
  - Predictive
  - Descriptive
  - Prescriptive
  - Diagnostic.

## **Reference Books**

- 1) Thomas Erl, Wajid Khattak, Paul Buhler, Big data Fundamentals Concepts, Driver & Techniques, Pearson
- 2) Tom White, "HADOOP: The definitive Guide", O Reilly
- 3) Dan Sullivan,"NoSQL for Mere Mortals", Pearson Education.
- 4) David Dietrich - Barry Hiller - "Data Science & Big Data Analytics" - EMC education services - Wiley publications