

## **Assignment - 1**

<b>Subject</b>	<b>Issue Date</b>	<b>Due Date</b>
Big Data Analytics (01CE0719)	10/07/2025	23/07/2025

**Note:**

- **Write Assignment in the file pages.**
- **Submit assignment in the file.**
- **Assignment will be checked in the respective lab sessions only.**
- **Mandatory to submit your checked assignment in the Google Classroom.**

<b>Sr. No.</b>	<b>Question</b>	<b>Cos</b>	<b>Bloom Taxonomy</b>
1	How 4 V's are important in Big Data Analytics.	CO1	Analyse
2	Illustrate the elements of Big Data Architecture with suitable diagram.	CO1	Apply
3.	Compare: RDBMS vs Hadoop	CO2	Analyse
4.	Break down the architecture of HDFS with a suitable diagram and describe the role of its components in managing data replication and fault tolerance.	CO2	Analyse
5.	Apply Map Reduce algorithm on following sentences. Dear Bear River Car Car River Deer Car Bear	CO2	Apply
6.	Break down the architecture of YARN with a labelled diagram and discuss how its components coordinate to manage resource allocation and job execution in a Hadoop cluster.	CO2	Analyse
7.	Explain the concept of Sharding and demonstrate how different types of sharding can be applied to improve the	CO1	Apply

---

	performance of a distributed database in a real-world scenario.		
8	Illustrate the architecture and working of Hive, and explain how its components collaborate to translate SQL-like queries into MapReduce jobs within the Hadoop ecosystem.	CO3	Analyse
9	Using practical examples, explain how Data Definition and Data Manipulation Language in Hive are applied to manage and query large datasets stored in HDFS.	CO3	Apply
10	Write HIVE queries to find the salary whose age is greater than 25.	CO3	Apply

Subject Faculty  
Prof. Urvi Dhamecha

Prof. (Dr.) Krunal Vaghela  
Head of Department  
Dept. of Computer Engineering