

Roll No.										
----------	--	--	--	--	--	--	--	--	--	--

BCADS Second Sessional Tests 2023-24  
Fourth Semester

BCADS1401: Big Data Fundamentals

Max. Marks: 60

Time: 3 Hours

(Section A)

(1\*8)

**Q1. Attempt All Questions**

- What is YARN's Primary function in Hadoop?
- List down the 5 V's of Big Data.
- What is the purpose of mapper in Map Reduce?
- What is the default block size and replication count in Hadoop?
- Define the use of Ls, cat, gedit, and touch commands?
- Write down the list of libraries of apache spark
- Explain the benefits of using sqoop in Hadoop
- What are the 4 different forms of data in NoSQL databases?

**Section B**

(6\*2)

**Q2. Attempt any Two.**

- List the segments of HDP and also name 2 components from each
- What is big data, what are its scope and challenges?
- Explain the architecture of HBASE in detail
- Explain the working of 5 components of hdfs and mapreduce in Hadoop

(5\*2)

**Q3 Attempt Any Two.**

- List down the differences between sql and NoSQL Databases
- Discuss the working and benefits of Horizontal scaling
- Explain the use of ambari in Hadoop

(5\*2)

**Q4 Attempt Any Two.**

- List and explain different types of analysis with examples
- Write down the procedure of static and dynamic partitioning In HIVE
- Explain the architecture of hive with neat diagram and also Explain Its use cases

(5\*2)

**Q5 Attempt Any Two.**

- a) Describe the use of zookeeper and IBM fluid query in HDP
- b) Why is PIG used in Hadoop, explain the components of PIG
- c) Describe the Pillars of security in the big data ecosystem

(5\*2)

**Q6 Attempt Any Two**

- a) What is Map Reduce, explain its working with example of Word count diagram
- b) Describe Apache Spark and its working in brief
- c) How is YARN helping the growth of Hadoop?