



Paper ID : 261301

Printed Page: 1

Subject Code:BCDS601

Roll No:

BTECH

(SEM. VI) THEORY EXAMINATION 2025-26

BIG DATA AND ANALYTICS

TIME: 3 HRS

M.MARKS: 70

Note: Attempt all Sections. In case of any missing data; choose suitably.

SECTION A

1. Attempt *all* questions in brief.

02 x 7 = 14

Q no.	Question	CO	Level
a.	Briefly explain 5Vs of Big Data.	1	K1
b.	What is the significance of the 'Heartbeat' signal in a Hadoop Cluster?	2	K2
c.	What is the purpose of 'Capped Collections' in MongoDB?	3	K2
d.	List the basic data types supported by the Scala programming language.	4	K1
e.	Define 'Resilient Distributed Datasets (RDD)' in Spark.	4	K2
f.	What is the function of the 'Hive Metastore'?	5	K2
g.	Mention the use of 'Zookeeper' in cluster monitoring.	5	K3

SECTION B

2. Attempt any *three* of the following:

07 x 3 = 21

Q no.	Question	CO	Level
a.	Discuss Big Data Privacy and Ethics. What are the major challenges in auditing big data?	1	K2
b.	Discuss the architecture of HDFS. Explain how it handles file abstraction and replication.	2	K2
c.	Differentiate between NoSQL and RDBMS. Discuss the key characteristics of NoSQL databases.	3	K4
d.	Describe the YARN Architecture. How does it support running MRv1 applications?	4	K3
e.	Explain Apache Pig's Data Model. Why is Pig Latin preferred for data transformation?	5	K2

SECTION C

3. Attempt any *one* part of the following:

07 x 1 = 07

Q no.	Question	CO	Level
a.	Explain the Big Data Technology Stack. Discuss the challenges of conventional systems.	1	K2
b.	Compare and contrast Analysis vs. Reporting. Discuss modern data analytic tools.	1	K4



Paper ID : 261301

Printed Page: 2

Subject Code:BCDS601

Roll No:

BTECH

(SEM. VI) THEORY EXAMINATION 2025-26

BIG DATA AND ANALYTICS

TIME: 3 HRS

M.MARKS: 70

4. Attempt any one part of the following:

07 x 1 = 07

Q no.	Question	CO	Level
a.	Describe the process of Data Ingestion using Flume and Sqoop.	2	K2
b.	Explain the HDFS Read and Write paths. How does the client interact with the NameNode?	2	K2

5. Attempt any one part of the following:

07 x 1 = 07

Q no.	Question	CO	Level
a.	Explain the querying mechanism in MongoDB. How does it handle complex data structures compared to traditional SQL queries?	3	K4
b.	Compare and contrast the Schema-less nature of NoSQL databases with the Schema-oriented nature of traditional RDBMS. Discuss the advantages of using NoSQL for Big Data.	3	K4

6. Attempt any one part of the following:

07 x 1 = 07

Q no.	Question	CO	Level
a.	Analyze the Resilient Distributed Dataset (RDD) model in Spark. How does Spark's "In-Memory" processing provide a performance advantage over the traditional MapReduce disk-based approach?	4	K4
b.	Discuss the importance of closures and higher-order functions in Scala. How do these features simplify the development of distributed processing logic for Spark?	4	K2

7. Attempt any one part of the following:

07 x 1 = 07

Q no.	Question	CO	Level
a.	Compare Local Mode and MapReduce Mode in Apache Pig. In what specific scenarios would a developer choose the Grunt shell over a pre-written Pig script?	5	K2
b.	Analyze why HBase is superior for random, real-time read/write access to Big Data compared to HDFS alone or a traditional RDBMS. Discuss its Schema Design and Advanced Indexing capabilities.	5	K4