



Roll No:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

BTECH
(SEM VIII) THEORY EXAMINATION 2024-25
BIG DATA

TIME: 3 HRS

M.MARKS: 100

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

SECTION A

1. Attempt all questions in brief.

2 x 10 = 20

Q No.	Question
a.	What is Big Data?
b.	Explain characteristics of big data
c.	Explain four V's of big data
d.	Discuss application of big data
e.	What is big data analytics?
f.	Discuss challenges of big data.
g.	Differentiate between structured and non structured data
h.	Differentiate between HDFS and Hbase
i.	What is Zookeeper? List benefits of it
j.	Differentiate between Apache pig Vs Map Reduce

SECTION B

2. Attempt any three of the following:

10 x 3 = 30

Q No.	Question
a.	Explain how big data processing is different from distributed process
b.	Discuss Hadoop YARN in detail with failures in classic Map Reduce.
c.	Explain Map Reduce framework in details.
d.	What is name node and data node in Hadoop architecture
e.	Differentiate between NoSQL and SQL.

SECTION C

3. Attempt any one part of the following:

10 x 1 = 10

Q No.	Question
a.	What is map reduce? Explain working of various phase of map reduce with appropriate example and diagram.
b.	Explain working of Hive with proper steps and diagram

4. Attempt any one part of the following:

10 x 1 = 10

Q No.	Question
a.	Explain "Map Phase" and "Combiner Phase" in Map Reduce.
b.	What is data serialization? Make a note on how type of data affects data serialization.



Paper ID : 250138

Roll No:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

BTECH
(SEM VIII) THEORY EXAMINATION 2024-25
BIG DATA

TIME: 3 HRS

M.MARKS: 100

5. Attempt any one part of the following:

10 x 1 = 10

Q No.	Question
a.	What is Hadoop Ecosystem? Explain various components of Hadoop Ecosystem.
b.	Explain Spark components in details. Also list features of Spark.

6. Attempt any one part of the following:

10 x 1 = 10

Q No.	Question
a.	What is NoSQL database? List the differences between NoSQL and relational databases. Explain in brief various types of NoSQL databases in practice.
b.	Define HDFS. Discuss HDFS architecture and HDFS commands in brief.

7. Attempt any one part of the following:

10 x 1 = 10

Q No.	Question
a.	What do you mean by HiveQL data definition language? Explain any three HiveQL data definition language with their syntax and example.
b.	Explain job Scheduling in Map Reduce. How is it done in case of (i) fair scheduling (ii) capacity scheduling

QP25EP1_143 | 19-May-2025 9:06:19 AM | 182.71.247.82