



35848

7E1824

Roll No.

Total No. of Pages: 2

7E1824

B. Tech. VII - Sem. (Main / Back) Exam., Nov-Dec 2025

Information Technology
7IT4-01 Big Data Analytics
AID, IT

Time: 3 Hours 156

Maximum Marks: 70

Instructions to Candidates:

Attempt all ten questions from Part A, five questions out of seven questions from Part B and three questions out of five questions from Part C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing may suitably be assumed and stated clearly. Units of quantities used/calculated must be stated clearly.

Use of following supporting material is permitted during examination. (Mentioned in form No. 205)

2. NIL**PART - A****[10×2=20]****(Answer should be given up to 25 words only)****All questions are compulsory**

- Q1 Define Big Data.
- Q2 What is Semi-structured data?
- Q3 Define Data Node.
- Q4 Mention any two important XML configuration files in Hadoop.
- Q5 What is MapReduce?
- Q6 What is the function of a Partitioner?
- Q7 Define Pig Latin.
- Q8 Write two Simple Pig Latin commands.
- Q9 List two advantages of using Hive.
- Q10 Define Hive QL.

[7E1824] / 428

Page 1 of 2

[1580]

PART - B

[5×4=20]

(Analytical/Problem solving questions)

Attempt any five questions

- Q.1. Explain the 3V's of Big data with suitable example.
- Q.2. Explain & differentiate Structured, Semi-structured and Unstructured data with examples.
- Q.3. Explain the role of Record Reader in Hadoop. Why it is important?
- Q.4. What is a Raw Comparator? How can Custom Comparators improve MapReduce performance? Explain.
- Q.5. Write and explain a Pig Latin Script to analyse student marks and calculate average scores.
- Q.6. Explain the interfaces available to run Pig Scripts.
- Q.7. Compare Hive QL with traditional SQL. How does Hive simplify Querying on Hadoop data?

PART - C

[3×10=30]

(Descriptive/Analytical/Problem Solving/Design Questions)

Attempt any three questions

- Q.1. Explain different sources of Big data. Also describe the problems faced by traditional large scale systems in handling Big data.
- Q.2. Explain MapReduce framework with a suitable example. Also write a MapReduce program for weather dataset analysis with the explanation of Driver, Mapper and Reducer code.
- Q.3. Explain the following Writable classes with example -
- (a) Text
 - (b) BytesWritable
 - (c) NullWritable
 - (d) ObjectWritable
 - (e) GenericWritable
- Q.4. Explain the Architecture of PIG with neat diagram. Also explain the application flow of PIG Latin Scripts with an example.
- Q.5. Explain the Hive architecture with diagram. Also describe the different Hive clients along with its roles.