

3718

**B. Tech. 8th Semester (Computer
Science & Information Technology)**

Examination – May, 2023

BIG DATA ANALYSIS

Paper : PCC-CSE-404-G/PCC-IT-403-G

Time : Three Hours] [Maximum Marks : 75

Before answering the questions, candidates should ensure that they have been supplied the correct and complete question paper. No complaint in this regard, will be entertained after examination.

Note : Attempt *five* questions in all, selecting at least *one* question from each Unit. Question No. 1 is *compulsory*. All questions carry equal marks.

1. Write short note on the following : **2.5 × 6 = 15**

- (i) Data Sciences
- (ii) ELT
- (iii) Sources of data using service bindings
- (iv) Data ingestion
- (v) MongoDB
- (vi) DFS

UNIT - I

2. (a) What is Big Data ? Explain various characteristics, challenges and applications of Big Data. 8
- (b) What is HDFS ? Explain its components. 7
3. (a) Explain six V's of Big Data in detail. 8
- (b) Write short note on foundation for big data system and programming. 7

UNIT - II

4. (a) Define Data Mart ? Explain different types of data marts with example. Also advantages and disadvantages of data marts. 8
- (b) Explain RDMS features and architecture in detail. 7
5. (a) What is NoSQL .Explain different types of NoSQL databases with example ? Differentiate SQL and NoSQL with example. 10
- (b) Write short note on different types of file formats used in big data. 5

UNIT - III

6. What is Modeling ? Explain various type of data models with example. 15
7. (a) Explain different types of big data management techniques. 10
- (b) Write short note on real life applications of big data. 5

UNIT - IV

8. (a) What is Hive ? Explain Hive features. Hive integration and workflow and architecture in detail. 10
- (b) Write short note on Hive Query Language (HQL). 5
9. (a) Write short note on Pig architecture and commands. 8
- (b) Explain MapReduce working ? WAP for Word Count using MapReduce. 7