

CBCS SCHEME

USN

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

18CS72

Seventh Semester B.E. Degree Examination, June/July 2024 Big Data Analytics

Time: 3 hrs.

Max. Marks: 100

Note: Answer any FIVE full questions, choosing ONE full question from each module.

Module-1

- 1 a. What is data? How do you classify data as structured, semistructured, multistructured and unstructured data? Give suitable examples. (10 Marks)
- b. Define big data. Describe characteristics and types of Big data. (10 Marks)

OR

- 2 a. Explain the role of big data in medicine and healthcare system. (10 Marks)
- b. Explain the functions of each of the five layers in big data architecture design with a diagram. (10 Marks)

Module-2

- 3 a. Give the purpose of Hadoop, describe the core components and features. (10 Marks)
- b. List and explain the HDFS commands. (10 Marks)

OR

- 4 a. Explain in detail about Hadoop MapReduce framework with programming model. (12 Marks)
- b. Describe Hadoop YARN and Hadoop execution model. (08 Marks)

Module-3

- 5 a. What is NOSQL? Explain the characteristics and features of NOSQL. (08 Marks)
- b. Explain in detail about NOSQL data architecture pattern. (12 Marks)

OR

- 6 a. Explain BASE and CAP theorems in detail. (10 Marks)
- b. Explain in detail the MongoDB along with its characteristics. (10 Marks)

Module-4

- 7 a. Explain relation algebra operations with an example. (08 Marks)
- b. Define Hive. List out characteristics, limitations and explain hive architecture in detail. (12 Marks)

OR

- 8 a. Explain in detail about map-tasks, key value, pair grouping by key. (10 Marks)
- b. Explain the pig architecture in detail. (10 Marks)

Module-5

- 9 a. Explain the text mining process architecture and applications. (10 Marks)
- b. Explain the 3 types of web mining. (10 Marks)

OR

- 10 a. Explain the working of simple linear regression, along with a real-world application illustration. (12 Marks)
- b. Compare Text Mining and Data Mining. (08 Marks)

* * * * *

Important Note : 1. On completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages.
2. Any revealing of identification, appeal to evaluator and /or equations written eg, 42+8 = 50, will be treated as malpractice.