

S.No. 6338

P 16 CSE 5 A/P 16 ITE 3 A

(For candidates admitted from 2016–2021 Batch)

M.Sc. DEGREE EXAMINATION, NOVEMBER 2023.

Computer Science/ Information Technology — Elective

BIG DATA ANALYTICS

Time : Three hours

Maximum : 75 marks

PART A — (10 × 2 = 20)

Answer ALL questions.

1. What is Semi-Structured Data?
2. Define big data
3. List out the need of big data Analytics
4. What is mean by Data Science?
5. Write about NewSQL.
6. List out the characteristics of MongoDB.
7. Define Hadoop.
8. List the components of Hadoop.

9. Define YARN.
10. What is meant by Data Serialization?

PART B — (5 × 5 = 25)

Answer ALL the questions, choosing either (a) or (b)

11. (a) Write about Unstructured data in digital data.

Or

- (b) Describe about characteristics of Big data.

12. (a) Discuss about technology expertise in Data Science.

Or

- (b) Explain about classification of analytics in big data.

13. (a) Write short on NoSQL and its Types.

Or

- (b) Explain the concepts of Apache Cassandra.

14. (a) What are the key components available in Hadoop? Explain.

Or

- (b) Discuss about concept of HDFS in Hadoop.

15. (a) Explain about Map reduce application with neat diagram.

Or

- (b) How to working with common serialization formats? Explain.

PART C — (3 × 10 = 30)

Answer any THREE questions.

16. Differentiate data environment and big data environment.

17. Discuss in detail about basic terminologies in data environment

18. Explain about features of Cassandra in detail.

19. Illustrate the concepts of Hadoop ecosystem

20. Describe about storage formats in Big data serialization.