

APRIL 2024

51432/SU45A/SE26B

---

Time : Three hours

Maximum : 75 marks

PART A — (10 × 2 = 20 marks)

Answer any TEN Questions each in 30 words.

1. What is data?
2. What is various stages in data science?
3. What are the benefits of data science?
4. What is modeling process?
5. Define data analysis.
6. What is machine learning?
7. Define frame work.
8. What is map reduce?
9. Define: BASE.
10. What is Hadoop.
11. What is research.
12. What is automation in data science?

PART B — (5 × 5 = 25 marks)

Answer any FIVE Questions each in 200 words.

13. Explain the goals of data science.
14. How to retrieve the data in data science process.
15. Why we need the machine learning in data science.
16. What are the applications of machine learning in data science?
17. Explain Hadoop concepts.
18. Discuss the data retrieval with examples.
19. Explain exploration and preparation of disease profiling with examples.

PART C — (3 × 10 = 30 marks)

Answer any THREE Questions each in 500 words.

20. Briefly Explain the various concepts of data science.
21. Discuss the data science process overview.
22. Briefly Explain supervised and unsupervised machine learning.

23. Discuss the following:

- (a) ACID
- (b) CAP
- (c) BASE

24. Briefly Explain the case study of data science.

---