

CBCS SCHEME

USN

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

21CS754

Seventh Semester B.E./B.Tech. Degree Examination, Dec.2024/Jan.2025 Introduction to Data Science

Time: 3 hrs.

Max. Marks: 100

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

Module-1

- 1 a. List the benefits of Data Science. (05 Marks)
- b. Explain the categories of data used in data science. (07 Marks)
- c. Write a note on NoSQL Databases. (08 Marks)

OR

- 2 a. Define Big data and Data science. List the characteristics of Big data. (04 Marks)
- b. Briefly explain the six steps of data science process. (12 Marks)
- c. Write a short note on Machine learning framework. (04 Marks)

Module-2

- 3 a. How do you set up a research goal and create a project charter. (08 Marks)
- b. Explain the process of Data Cleansing. (12 Marks)

OR

- 4 a. Explain the different ways of combining data. (08 Marks)
- b. How do you transform data suitable for data modeling? (08 Marks)
- c. Mention and show four different types of graphs used in exploratory analysis. (04 Marks)

Module-3

- 5 a. What is Machine learning? Mention its applications. (04 Marks)
- b. Briefly explain the four steps of Modeling Process. (08 Marks)
- c. Why is validation important? Explain three validation strategies. (08 Marks)

OR

- 6 a. List the python libraries to work with data in memory. (06 Marks)
- b. Mention the two common error measures in machine learning. Explain them with an example. (06 Marks)
- c. What do you mean by confusion matrix? Explain with an example. (08 Marks)

Module-4

- 7 a. Explain the different ways of delivering a new insight to end user by data scientists. (06 Marks)
- b. Explain the Javascript libraries available to work on visualization tools. (08 Marks)
- c. What are the reasons against developing your own application? (06 Marks)

OR

- 8 a. State the reasons for creating own custom reports instead of opting company tools. (06 Marks)
- b. Mention the ingredients of a dc.js application and three files created. (08 Marks)
- c. Explain three important reduce functions to write MapReduce logic. (06 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.

Module-5

- 9 a. Explain the components of Hadoop framework. (10 Marks)
b. Explain how Hadoop achieves parallelism with an example. (10 Marks)

OR

- 10 a. What is Spark? How does Spark solve the problem of MapReduce. (10 Marks)
b. Explain the components of Spark Framework. (10 Marks)

* * * * *