

CBGS SCHEME

USN

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

21CS754

Seventh Semester B.E./B.Tech. Degree Examination, June/July 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. Explain main categories of data. (10 Marks)
- b. Explain NoSQL database in big data technologies. (10 Marks)

OR

- 2 a. Explain data science process. (10 Marks)
- b. Write the benefits and uses of data science and big data. (10 Marks)

Module-2

- 3 a. Write an overview of techniques to handle missing data. (10 Marks)
- b. Why correct errors as early as possible in data science? (10 Marks)

OR

- 4 a. Write a note on retrieving data. (10 Marks)
- b. Explain presentation and automation. (10 Marks)

Module-3

- 5 a. What is machine learning? Applications for machine learning in data science. (10 Marks)
- b. Explain packages for working with data in memory. (10 Marks)

OR

- 6 a. Explain modeling process. (10 Marks)
- b. Explain types of machine learning. (10 Marks)

Module-4

- 7 a. Explain dashboard development tools. (10 Marks)
- b. Design a code to generate total stock overtime graph. (10 Marks)

OR

- 8 a. What are uses of data visualization options? (10 Marks)
- b. Write a code to generate average stock per medicine graph. (10 Marks)

Module-5

- 9 a. Explain Map Reduce flow for counting the colors in input texts. (10 Marks)
- b. What is spark? How does spark solve the problems of Map Reduce? (10 Marks)

OR

- 10 a. Explain Data preparation in spark. (10 Marks)
- b. List of common hadoop file system commands. (10 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.