QP Code: D 122750	Total Pages: 1	Name:
		Register No.

SECOND SEMESTER (CUFYUGP) DEGREE EXAMINATION APRIL 2025

COMPUTER SCIENCE

CSC2MN102 – INTRODUCTION TO DATA SCIENCE

2024 Admission onwards

Maximum Time: 2 Hours | **Maximum Marks: 70**

Section A

All Questions can be answered. Each Question carries 3 marks (Ceiling: 24 Marks)

1	What are the key challenges in data science?
2	Differentiate between descriptive and predictive data analysis.
3	What is kurtosis? Explain its importance in data analysis.
4	Describe different types of bar charts used in data visualization.
5	What is data cleaning? Why is it necessary?
6	Explain the importance of dimensionality reduction in data science.
7	Define model training and its role in machine learning.
8	Differentiate between classification and regression in machine learning.
9	Explain k-means clustering with a simple example.
10	What is the role of probability distributions in data science?

Section B

All Questions can be answered. Each Question carries 6 marks (Ceiling: 36 Marks)

11	Describe the different types of data analysis used in data science.
12	Discuss different statistical measures used for data analysis.
13	Explain various data transformation techniques with examples.
14	What is a heatmap? How is it useful in data analysis?
15	Explain the importance of feature scaling in machine learning models.
16	What are the differences between underfitting and overfitting in machine learning?
17	Describe the role of machine learning in real-world applications.
18	Explain the concept of Singular Value Decomposition (SVD).

Section C

Answer any ONE. Each Question carries 10 marks (1x10=10 Marks)

19	Discuss various exploratory data analysis (EDA) techniques with examples.
20	Explain different types of machine learning algorithms with case studies.