S.No. 7255 **P 20 MCA 15**

(For candidates admitted from 2020–2021 onwards)

M.C.A. DEGREE EXAMINATION, NOVEMBER 2023.

Computer Applications

PRINCIPLES OF DATA SCIENCE

Time : Three hours Maximum : 75 marks

PART A — (10 × 2 = 20)

Answer ALL questions.

1. Define Probabilistic model.
2. What are Discrete and Continuous data?
3. Define Random variables.
4. What is meant by set theory?
5. Define sampling bias.
6. What is coefficient of variation?
7. What are Scatter plots?
8. Define Regression.
9. Define cluster.
10. How will you calculate A Silhouette Coefficient?

PART B — (5 × 5 = 25)

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

11. (a) Explain Nominal level and ordinal level of data with examples.

Or

(b) Explain the concepts of Filtering in pandas.

12. (a) Write short notes on vectors and Matrices.

Or

(b) Explain the Frequentist approach with examples.

13. (a) Explain Measures of center in statistics.

Or

(b) Explain the Empirical rule.

14. (a) Differences between correlation and causation.

Or

(b) Write short notes on Histogram.

15. (a) How bias/variance play into error functions?

Or

(b) Write short notes on Grid Searching.

PART C — (3 × 10 = 30)

Answer any THREE questions.

16. Explain about case study – Marketing dollars in Data science with examples in detail.

17. Explain any two applications of Bayes theorem.

18. Briefly explain the Hypothesis tests.

19. Describe the concepts of unsupervised learning with examples.

20. Discuss about Decision trees with examples.

———————