MAY 2022 **P/ID 17621/PCAE5**

Time : Three hours Maximum : 80 marks

PART A — (10 × 2 = 20 marks)

Answer any TEN questions.

1. What are outliers? Mention the types of outlier with example.

2. List out the various job profiles involved in Data Analytics.

3. What is meant by segmentation?

4. What do you mean by Standardization of Data?

5. List out the advantages of statistical methods for predictive analysis.

6. Define the terms Support and Confidence.

7. Mention any two applications of Association rules.

8. Explain Ego nets.

9. What are Self-organizing maps?

10. What do you mean by collective Inferencing?

11. Define Web Data Analytics.

12. List out the various dimensions of data quality.

PART B — (5 × 6 = 30 marks)

Answer any FIVE questions.

13. Discuss on different types of data elements with seperate examples.

14. Explain the various methods to deal with missing values.

15. Discuss on the Analytics process model with a neat diagram.

16. Write short notes on the various types of clusters.

17. Explain Gibb's sampling method of collective inferencing.

18. Discuss on the basic principles to safeguard privacy as defined by OECD.

19. Explain the various Methods used for web Data Collection.

PART C — (3 × 10 = 30 marks)

Answer any THREE questions.

20. Write short notes on Decision trees.

21. Discuss briefly on the performance measures for classification and Regression models.

22. What is Association rule mining? Explain the steps involved in association rule mining with example.

23. Write short notes on the various approaches used to build a Recommender system.

24. Discuss briefly on the centrality measures in a Social Network.

———————