T.K.M. COLLEGE LIBRARY
MAVELIKARA
PIN: 690110
KERALA

(Pages : 3) **S – 6516**

**Reg. No. : .....**

**Name : .....**

**Second Semester M.A. Degree Examination, January 2024**

**Behavioural Economics and Data Science**

**BEDS 523 : FOUNDATIONS OF DATA SCIENCE**

**(2020 Admission Onwards)**

Time : 3 Hours

Max. Marks : 75

SECTION – A

Answer **all** questions. Each question carries **1** mark.

1. Data cleaning.
2. Descriptive statistics.
3. Clulstering.
4. Student t test.
5. Statistic.
6. EDA
7. Probability distribution.
8. Linear regression.
9. Machines learning.
10. Random forest.

**(10 × 1 = 10 Marks)**

P.T.O.

## SECTION – B

Answer any **seven** questions. Each question carries **5** marks.

11. Describe different perspectives of Data Science.

12. Briefly explain the steps involved in testing of hypothesis.

13. Write a short note on hierarchical agglomerative clustering.

14. Discuss the practical applications of Data Science in real-world scenarios.

15. Explain the concept of data discretization and its applications in data processing.

16. Write a detailed note on Machine Learning algorithm.

17. What is data pre-processing, and why is it essential in the field of Data Science?

18. Discuss the role of Student's t-tests and ROC curves in evaluating classification methods.

19. Explain the concept of linear regression and its applications in predictive modeling.

20. Describe Lloyd's algorithm and how it partitions data into clusters.

**(7 × 5 = 35 Marks)**

## SECTION – C

Answer any **three** questions. Each question carries **10** marks.

21. Explain the concept of matrix decomposition and its relevance in Data Science, focusing on SVD and PCA.

22. Discuss the concept of clustering quality and the various metrics used to assess the quality of clustering results.

23. Describe logistic regressing classifiers and discuss their role in binary classification problems.

24. Describe the steps involved in the Data Science process, from data processing to generating insights.

25. Discuss the challenges and considerations involved in visualizing big data and complex datasets.

**(3 × 10 = 30 Marks)**

---