Reg No:..... K25FY2243
Name :.....

Second Semester FYUGP Computer Science Examination
APRIL 2025 (2024 Admission onwards)
KU2MDCCAP104 (INTRODUCTION TO DATA SCIENCE)
(DATE OF EXAM: 26-4-2025)

Time : 90 min Maximum Marks : 50

Part A (Answer any 6 questions. Each carries 2 marks)

1. Identify any four commonly used Python libraries in Data Science. 2
2. Explain the difference between a list and a tuple in Python. 2
3. What are measures of central tendency. 2
4. List out 3 types of machine learning. 2
5. Define the term 'Training Data' in Machine Learning. 2
6. What is Regression Analysis? 2
7. State two real-world applications of Classification. 2
8. Differentiate between hard clustering and soft clustering. 2

Part B (Answer any 4 questions. Each carries 6 marks)

9. Analyze how Data Science is transforming industries like healthcare, finance, and retail. Provide examples. 6
10. Compare mutable and immutable data structures in Python with examples. 6
11. Apply descriptive statistics to summarize a dataset containing student exam scores. Provide insights based on mean, median, standard deviation, and interquartile range. 6
12. Compare Supervised and Unsupervised Learning with real-life examples. 6
13. Compare and contrast Simple Linear Regression and Multiple Regression using a real-life dataset. 6
14. Analyze how hierarchical clustering works and compare the Agglomerative and Divisive approaches 6

Part C (Answer any 1 question(s). Each carries 14 marks)

15. Construct a robust missing data handling mechanism by comparing techniques like imputation(mean, median, mode), deletion, and predictive modeling. C2 14
16. Design a step-by-step approach to implementing K-Means clustering on a dataset. Explain how the number of clusters can be determined effectively. 14