Registration Number:

Date & session:

ST JOSEPH'S UNIVERSITY, BENGALURU -27
B.C.A (DATA ANALYTICS) – II SEMESTER
SEMESTER EXAMINATION: APRIL 2024
(Examination conducted in May / June 2024)
BCADA 2121: PRINCIPLES OF DATA SCIENCE

(For current batch students only)

Time: 2 Hours **Max Marks: 60**

This paper contains TWO printed pages and THREE parts

PART - A

ANSWER ALL THE QUESTIONS **5 x 2 = 10**

1. Describe the types of data with an example.
2. Define two strategies used to handle missing data.
3. List the differences between a discrete variable and a continuous variable with an example.
4. Define correlation coefficient.
5. What is the use of scatter plot?

PART - B

ANSWER ANY FIVE QUESTIONS **5 x 4 = 20**

6. Explain the different facets of data with an example.
7. Suppose a hospital tested the age and body fat data for randomly selected adults with the following result:

Age	**23**	**27**	**39**	**49**	**50**	**52**	**54**	**56**	**57**	**58**	**60**
% fat	**9.5**	**17.8**	**31.4**	**27.2**	**31.2**	**34.6**	**42.5**	**33.4**	**30.2**	**34.1**	**41**

   Draw the boxplots for age.
8. Explain the need for Correlation and Regression analysis in Data analysis with an example.
9. Write down the types of machine learning and explain the same.
10. Explain the steps involved in modelling process for machine learning.
11. Discuss in detail about Text mining Applications.
12. How to Value different aspects of privacy in data science? Justify.

PART - C

ANSWER ANY THREE QUESTIONS **3 x 10 = 30**

13. Elaborate about the steps in the data science process with a diagram.
14. Write the procedure of solving two-way ANOVA and discuss how it is used as a feature selection technique.
15. Describe the concept of Decision Tree Classification in detail.
16. Explain about the 5 C's of data science ethics.