

Roll No.																			
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



BENGALURU
School of School of Computer Science and Engineering
Mid - Term Examinations - November 2024

Semester: V

Date: 06-11-2024

Course Code: CSE3038

Time: 02.00pm to 03.30pm

Course Name: Applied Data Science

Max Marks: 50

Program: B. Tech

Weightage: 25%

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2 marks.

5Qx2M=10M

- | | | | | |
|----------|--|---------|----|-----|
| 1 | Define data and its types. | 2 Marks | L1 | C01 |
| 2 | List applications of data science in real world. | 2 Marks | L1 | C01 |
| 3 | List out differences between Data analysis and Data analytics. | 2 Marks | L1 | C01 |
| 4 | List out two functions to check null values in a data frame. | 2 Marks | L1 | C02 |
| 5 | Define interpolation. | 2 Marks | L1 | C02 |

Part B

Answer ALL Questions. Each question carries 10 marks.

4QX10M=40M

- 6** **a.** List any five data science job roles in current market. 5 Marks L1 C01
- b.** Explain any five data science applications in real world. 5 Marks L2 C01

or

- 7** **a.** Create a 2-D numpy array and check the type of array, dimension, shape, size, type of elements in the array (using python). 5 Marks L3 C01
- b.** Explain OSEM framework with diagram. 5 Marks L3 C01

- 8** **a.** Write a python program to explain how can you handle missing values using simpleimputer. 5 Marks L2 C01
- b.** Illustrate reshaping a 3*4 numpy array to 2*2*3 numpy array (using python). State the condition to reshape arrays from one array to another. 5 Marks L3 C01

or

- 9** **a.** Explain Data science life cycle with diagram. 5 Marks L2 C01
- b.** Create an empty series, series using numpy array, series using a list (using python). 5 Marks L3 C01

- 10** **a.** Explain any five types of Data quality assessment. 5 Marks L2 C02
- b.** Explain how can you handle missing values for time series data using python. 5 Marks L3 C02

or

- 11** **a.** Define feature aggregation. State one scenario where feature aggregation is applied. 3 Marks L3 C02

b. Explain imputation using K nearest neighbour with program.(using python) 7 Marks L3 C02

12 a. Explain how can we fill categorical missing values (using python). 5 Marks L3 C02

b. Demonstrate distribution measures on a sample data frame. (Using python) 5 Marks L3 C02

or

13 a. Explain how can you visualize null values using heat map. (Use python) 5 Marks L3 C02

b. What do you understand by imputation. Explain mean, median and mode imputation. (Using python) 5 Marks L3 C02