Roll No										

PRESIDENCY UNIVERSITY
BENGALURU

SET A

SCHOOL OF INFORMATION SCIENCE
END TERM EXAMINATION - JAN 2024

Semester : Semester I - 2023
Course Code : CSA1003
Course Name : Fundamentals of Data Science
Program : B.Sc. Data Science

Date : 10-JAN-2024
Time : 1:00 PM - 4:00 PM
Max Marks : 100
Weightage : 50%

Instructions:
(i) Read all questions carefully and answer accordingly.
(ii) Question paper consists of 3 parts.
(iii) Scientific and non-programmable calculator are permitted.
(iv) Do not write any information on the question paper other than Roll Number.

PART A

ANSWER ALL THE QUESTIONS **5 X 2M = 10M**

1. **What is the use of Data Warehousing?**

(CO1) [Knowledge]

2. Define Data Science.

(CO1) [Knowledge]

3. What is the use of Z Test?

(CO2) [Knowledge]

4. Define Correlation Coefficient. Write the formulae to find r.

(CO3) [Knowledge]

5. Write the formulae to find Euclidean Distance.

(CO4) [Knowledge]

PART B

ANSWER ALL THE QUESTIONS **5 X 10M = 50M**

6. Obtain mean, variance and standard deviation for the following observation X and Y. Analyze which is best

x	y
10	45
15	22
42	70
26	16
40	74
34	11
53	24
19	14
20	30
21	17

(CO1) [Comprehension]

7. Explain Median, Median, Mode, Range, Percentile with Example.

(CO1) [Comprehension]

8. Explain Different types of Charts with neat Diagram.

(CO2) [Comprehension]

9.

Solve the prediction to obtain least square method for the following observation	
x	y
30	40
60	50
90	70
85	61
72	87
63	79
45	12
22	14
13.5	16
14	18

(CO3) [Comprehension]

(CO4) [Comprehension]

10. Illustrate KNN Algoritthm.

PART C

ANSWER ALL THE QUESTIONS **2 X 20M = 40M**

11. a) From a statistics standpoint, the standard deviation of a dataset is a measure of the magnitude of deviations between the values of the observations contained in the dataset. From a financial standpoint, the standard deviation can help investors quantify how risky an investment is and determine their minimum required return on the investment. Solve the standard deviation for the following 12 weeks data set. [15 Marks]

Weeks	Expenditure
1	$48.50
2	$87.40
3	$19.98
4	$59.74
5	$40.87
6	$105.51
7	$40.80
8	$23.10
9	$98.10
10	$60.54
11	$64.81
12	$48.01

b) A garden contains 39 plants.
The following plants were chosen at random, and their heights were recorded in cm: 38, 51, 46, 79, and 57. Calculate their heights' standard deviation. [5 Marks] (CO2) [Application]

12. Consider the following given data and find student s1(8.5,7,7) will get placed or not by using KNN, where K=3 ?

10th CGPA	12th CGPA	BTech CGPA	Status
9	8	7	NOT PLACED
7	7.5	8	PLACED
8	8.5	9	PLACED
9.5	9	7	NOT PLACED
6.8	8	7	PLACED
7.7	8	8	PLACED
8.5	8	9	PLACED

(CO4) [Application]