PRESIDENCY UNIVERSITY
BENGALURU

SCHOOL OF ENGINEERING
END TERM EXAMINATION - JUN 2023

Semester : Semester VI - 2020
Course Code : CSE2027
Course Name : Sem VI - CSE2027 - Fundamentals of Data Analytics
Program : CAI

Date : 16-JUN-2023
Time : 9.30AM - 12.30PM
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 (10 X 2 = 20M)

1. Interquartile range is the measure of how the data is dispersed. Write the formula for IQR(Inter Quartile Range).

   (CO2) [Knowledge]

2. Write the formula for Recall and Precision with respect to Confusion matrix.

   (CO5) [Knowledge]

3. In which technique do we divide the entire population into subgroups, wherein, each of those subgroups has similar characteristics?

   (CO2) [Knowledge]

4. What is a secondary data?Give any two sources of secondary data.

   (CO3) [Knowledge]

5. What is regression in context of machine learning?

   (CO5) [Knowledge]

6. Which chart is suitable to measure the trend for various categories over a period of time.

   (CO4) [Knowledge]

7. Mention any two advantages of primary data.

   (CO3) [Knowledge]

8. Following data depicts the number of calls received for 7 consecutive days:
   11,13,17,19,13,23,25
   Find the mode of the calls received.

   (CO1) [Knowledge]

9. What are the different Data measurement scales used during analysis?

(CO1) [Knowledge]

10. List the uses of Dashboard as a outcome of data anlysis.

(CO4) [Knowledge]

PART B

ANSWER ALL THE QUESTIONS   (5 X 10 = 50M)

11. Explain essential characteristis of a good questionnarie.

(CO3) [Comprehension]

12. Explain five major characterstics of big data.

(CO1) [Comprehension]

13. How do you visualise the data effectively? Explain with an example using, i. Column chart ii. Bar chart iii. Pie chart.

(CO4) [Comprehension]

14. Explain Simple Linear Regression and Multiple Linear Regression. How can a regression line be calculated, explain with equations and graphs.

(CO5) [Comprehension]

15. Sales of a company (in Lakhs per month) are given below:
69.47, 41.22, 93.1, 70.45, 59.74, 36.9, 72.5.
Find the fluctuation in the sales per month.

(CO2) [Comprehension]

PART C

ANSWER ALL THE QUESTIONS   (2 X 15 = 30M)

16. An 'ABC' automobile company is planning to launch two new variants of a four wheeler. Both the variants are tested for speeds at two different locations and following data is collected:
Vehicle-1(Km/H) 110, 119, 123, 132, 125, 135
Vehicle-2(Miles/H) 104, 87, 92, 103, 86, 93
As a data analyst, you are required to determine which vehicle has the highest variation in speed.

(CO2) [Application]

17. Discuss different types of data visualisations. Support your answer with graphs and specific use case for each type.

(CO4) [Application]