## PRESIDENCY UNIVERSITY

## BENGALURU

## SCHOOL OF INFORMATION SCIENCE END TERM EXAMINATION - JAN 2023

Semester: Semester I-2022
Course Code : CSA1003
Course Name : Sem I-CSA1003 - Fundamentals of Data Science Program : BCA

Date : 12-JAN-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.

## PART A

## ANSWER ALL THE FOLLOWING QUESTIONS

$10 \times 2=20 \mathrm{M}$

1. List the Pieces of Data Science Puzzle.
(CO1) [Knowledge]
2. Explaine the VLOOKUP function with syntax.
(CO1) [Knowledge]
3. Write the syntax for following functions:
a. SUMIF
(CO1) [Knowledge]
4. 

| 4 | A | $B$ | C | D |  | E |  | F |  |  | H |  | I |  | J |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Month | Sales |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | Jan | 186,983 | Sales by Month |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | Feb | 179,009 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | Mar | 193,422 | 250,000 |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | Apr | 135,956 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 | May | 196,025 | 200,000 |  |  |  |  |  |  |  |  |  |  |  |  |
| 7 | Jun | 140,323 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 8 | Jul | 139,302 | 150,000 |  |  |  |  |  |  |  |  |  |  |  |  |
| 9 | Aug | 134,922 | 100,000 |  |  |  |  |  |  |  |  |  |  |  |  |
| 10 | sep | 184,773 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 11 | Oct | 201,233 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 12 | Nov | 213,039 | 50,000 |  |  |  |  |  |  |  |  |  |  |  |  |
| 13 | Dec | 240,093 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 14 |  |  | 0 |  |  |  |  |  |  |  |  |  |  |  |  |
| 15 |  |  |  | Ian | Fcb | Mar | Apr | May | Jun | Ju | A4] | Sep | Oct | Nov | Dec |
| 16 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Analyse the above chart and draw few insights out of it.
(CO2) [Knowledge]
5. Write the difference between data visualization and data cleaning.
(CO2) [Knowledge]
6. What are the three types of data varieties? Give an example for each.
(CO2) [Knowledge]
7. Write all the steps which is used to clean data.
(CO3) [Knowledge]
8. Explain Descriptive Analysis \& Diagnostic Analysis.
(CO3) [Knowledge]
9. Explain Mean, Mode \& Median with example.
(CO3) [Knowledge]
10. How to select most appropiate graph for data visualization?
(CO4) [Knowledge]

## PART B

ANSWER ALL THE FOLLOWING QUESTIONS
$5 \times 10=50 M$
11. Explain types of feature selection techniques in machine learning (ML).
(CO4) [Comprehension]
12. Explain Knowledge Based System and its component with suitable diagram.
(CO4) [Comprehension]
13. Name and explin the disciplines that contributed ideas, viewpoints and techniques to foundations of Artificial Intelligence.
(CO4) [Comprehension]
14. Explain Machine Learning (ML) and its lifecycle.
(CO4) [Comprehension]
15. Draw a Semantic Network using following rules: Every Human, animal and birds is living things who breath and eat.All Birds can fly.All man and woman are humans who have two legs.Cat is an animal and has fur.All animals have skin and can move.Giraffe is an animal who is tall and has long legs.Parrot is a bird and is green in color.
(CO4) [Comprehension]

## PART C

## ANSWER ALL THE FOLLOWING QUESTIONS <br> $2 \times 15=30 M$

16. Sam has 20 rose bushes, but only counted the flowers on 6 of them!. The "population" is all 20 rose bushes, and the "sample" is the 6 bushes that Sam counted the flowers of. Let us say Sam's flower counts are: 9, 2, 5, 4, 12, 7 find out the Mean, Median and Standard Deviation.
(CO4) [Application]
17. Consider the following data points in ( $X, Y$ ) plane. $(3,30),(8,57),(9,64),(13,72),(6,43),(11,59),(21,90)$, $(1,20),(16,83)$. Find the regression equation and then value of $y$ when $x=10$.
(CO4) [Application]
