

Roll No



**PRESIDENCY UNIVERSITY
BENGALURU**

**SCHOOL OF ENGINEERING
MID TERM EXAMINATION - APR 2023**

Semester : Semester VI - 2020

Course Code : CSE2027

Course Name : Sem VI - CSE2027 - Fundamentals of Data Analytics

Program : CAI,CEI,CST

Date : 13-APR-2023

Time : 2:00PM - 3:30PM

Max Marks : 60

Weightage : 30%

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. Define Data analysis. (CO1) [Knowledge]
2. Find the median of the following dataset:
3, 4, 7, 2, 3, 7,4, 2, 4, 7, 4. (CO1) [Knowledge]
3. What is Prescriptive Analytics. (CO1) [Knowledge]
4. 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]
5. Find Mean of the following data set :
3,10,6,2,1,7,10,8,7,5,2,4,3,5,7,8,6 (CO1) [Knowledge]
6. What is skewness of data. (CO2) [Knowledge]
7. What is standard deviation? Write its formula. (CO2) [Knowledge]
8. For any given dataset ; Define median. (CO2) [Knowledge]
9. Define Sampling. List two types of sampling. (CO2) [Knowledge]

10. Define stratified sampling?

(CO2) [Knowledge]

PART B

ANSWER ALL THE QUESTIONS

(4 X 5 = 20M)

11. Mr.Roshan is required to identify the different set of analysis types available. His customer comes with the requirement that, he want to know the analysis of sale going to happen. Suggest him which is best possible analysis type he can adopt to solve his requirement and expalin in detail the analysis to be adopted.

(CO1) [Comprehension]

12. For business analysis lot of data is collected from all social media platforms.The collected data for business analysis will be of what type? Differentiate between structured data and unstructured data.

(CO1) [Comprehension]

13. A family of five members recorded their weekly expense continuously for a period of 12 weeks in order to cut down on unnecessary expenses. based on the data recorded by them, help them to calculate the variance and standard deviation.

Week	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

(CO2) [Comprehension]

14. A class Teacher announces a class committee meeting. She instructs few students to attend the meeting without fail. While selecting the students she picks few students based on some criteria in each cluster. Identify which sampling technique she has adopted and elaborate in detail.

(CO2) [Comprehension]

PART C

ANSWER ALL THE QUESTIONS

(2 X 10 = 20M)

15. Describe the many methods of data analysis and give thorough justifications for each. Select what kinds of analyses will be utilised to help a businessman decide how many things to produce. There are several types of Data Analysis techniques that exist based on business and technology. However, the major Data Analysis methods are: Need to explain with an example for each in detail.

(CO1) [Application]

16. The table shows a random sample of 100 hikers and the areas of hiking they prefer.

Hiking Area Preference.

Gender	Coastline	Lakes and Streams	Mountain Peaks	Total
Female	18	16	--	45
Male	--	---	14	55
Total	---	41	--	--

Hiking Area Preference.

- (i) Complete the contingency table
- (ii) Are the events 'being female' and 'preferring coastline' independent events ?
- (iii) Find the probability that hiker is a male given that the hiker prefers hiking near lakes and streams
- (iv) Find the probability that hiker is a female OR prefers hiking near mountain peaks
- (v) Are the events 'being male' and 'preferring lakes and streams' independent events ?

(CO2) [Application]