Roll No.											

PRESIDENCY UNIVERSITY

BENGALURU

End - Term Examinations – December 2025

Date: 05-12-2025 **Time:** 01:00pm – 04:00pm

School: COMMON TO ALL SCHOOL	**Program:** COMMON TO ALL PROGRAM	
Course Code : MAT3041	**Course Name:** Statistical Reasoning & Data Analysis	
Semester: OE	**Max Marks**:100	**Weightage**: 50%

CO - Levels	**CO1**	**CO2**	**CO3**	**CO4**
Marks	**24%**	**26%**	**24%**	**26%**

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 2marks. **10Q x 2M=20M**

1.	The mean weight of a student in a group of 6 students is 119 lbs. The individual weights of five of them are 115,109, 129, 117 and 114 lbs. What is the weight of the sixth student?	**2 Marks**	**L2**	**CO1**
2.	Find the median for the data set: 1,5,8,10,12,15,18	**2 Marks**	**L2**	**CO1**
3.	When two fair dice are thrown, find the probability that the sum is exactly 7.	**2 Marks**	**L2**	**CO2**
4.	A coin throwing three times find the probability of getting at least one head?	**2 Marks**	**L2**	**CO2**
5.	Write the formula for the conditional probability of A given B, and B given A.	**2 Marks**	**L1**	**CO2**
6.	Define Exploratory Data Analysis (EDA).	**2 Marks**	**L1**	**CO3**
7.	Define unstructured Data with one example.	**2 Marks**	**L1**	**CO3**
8.	Write the formula for Karl Pearson's correlation coefficient.	**2 Marks**	**L1**	**CO4**
9.	State the different kinds of predictive models.	**2 Marks**	**L1**	**CO4**

10.	Consider the equation 3x-2y+11=0. Find the two regression coefficients and the correlation coefficient.	**2 Marks**	**L2**	**CO4**

Part B

Answer the Questions. **Total Marks 80M**

11.	**a.**	Consider the marks obtained by 10 students in statistics out of 50 as follows: 35, 28, 15, 45, 50, 23, 34, 36, 27, 24 (a) Calculate the mean, median, and mode for this dataset. (b) Calculate the Standard Deviation.	**10 Marks**	**L3**	**CO1**
		Or			
12.	**a.**	Consider the marks scored in MAT3041 by 10 students, each from two different sections, out of a total mark of 50 in Midterm examinations. Section A 44 30 28 10 2 48 21 15 50 20 Section B 28 32 25 49 20 24 30 26 22 25 Determine: (a) Which section students are better scorers? (b) Which section students are more consistent performers?	**10 Marks**	**L3**	**CO1**

13.	**a.**	Consider the marks obtained by 10 students in statistics out of 50 as follows: 35, 28, 15, 45, 50, 23, 34, 36, 27, 24 Determine the score that is: (a)Greater than the lower 25% scores of the students. (First Quartile) (b)Greater than the lower 75% scores of the students (Third Quartile)	**10 Marks**	**L3**	**CO1**
		Or			
14.	**a.**	A company wants to analyze the daily commute times (in minutes) for 20 of its employees. The data is as follows:	**10 Marks**	**L4**	**CO1**

		25, 30, 45, 15, 60, 35, 40, 50, 20, 25, 30, 35, 40, 25, 30, 55, 65, 70, 20, 10. Calculate the Interquartile Range (IQR) .			

15.	**a.**	An experiment involves rolling two fair dice and observing the sum of the numbers obtained. (a) Find the probability that the sum is 7. (b) Find the probability that the sum is greater than 10.	**10 Marks**	**L3**	**CO2**
		Or			
16.	**a.**	A lot of 100 semiconductor chips contain 20 that are defective. Two chips are selected at random, without replacement, from the lot. (a)What is the probability that the first one selected is defective? (b)What is the probability that the second one selected is defective given that the first one was defective? (c)What is the probability that both are defective?	**10 Marks**	**L3**	**CO2**

17.	**a.**	In a certain factory, three machines M1, M2 and M3 make 30%, 45% and 25% of the products respectively. It is known that 2%, 3% and 2% of the products made by each machine, respectively, are defective. Suppose that a finished product is randomly selected. If the randomly selected product is found to be defective, what is the probability that it is made by machine M1and M3?	**10 Marks**	**L4**	**CO2**
		Or			
18.	**a.**	A consulting firm rents cars from three agencies such that 20% from agency X, 30% from agency Y and 50% from agency Z. 90% of the cars from X, 80% of the cars from Y and 95% of the cars from Z are in good condition. If a car is in good condition, what is the probability that it has come from (a) agency X (b) agency Y?	**10 Marks**	**L3**	**CO2**

19.	**a.**	Explain the role of data visualization in EDA and describe how commonly used plots help in understanding data distribution and variable relationships.	**10 Marks**	**L3**	**CO3**
		Or			
20.	**a.**	Describe the key steps, tools, and insights involved in conducting EDA on a real-world dataset such as COVID-19 or financial market data.	**10 Marks**	**L3**	**CO3**

21.	**a.**	Discuss the importance of visualization in EDA and explain how different plots support data interpretation.	**10 Marks**	**L3**	**CO3**
		Or			
22.	**a.**	Explain how missing data and data quality issues are detected, handled, and prevented during EDA, with suitable examples.	**10 Marks**	**L3**	**CO3**

23.	**a.**	Calculate the Karl-Pearson's correlation co-efficient for the following data	**10 Marks**	**L4**	**CO4**

X	60	34	40	50	45	41	22	43	42	66	64	46
Y	75	32	35	40	45	33	12	30	36	72	41	57

Hence interpret the nature of correlation prevalent between the variables.

Or

24.	**a.**	Consider the marks scored in 2 courses History and Sociology for 10 students on a scale of 0 - 10.	**10 Marks**	**L3**	**CO4**

Students	1	2	3	4	5	6	7	8	9	10
History	9	4	7	8	5	6	7	9	8	2
Sociology	8	5	7	7	6	2	3	5	6	2

Calculate Karl Pearson's correlation coefficient for the given data and state whether the correlation is positive, negative, or negligible.

25.	**a.**	Estimate the value of X when Y = 30	**10 Marks**	**L3**	**CO4**

X	45	43	46	44	42	40	41
Y	41	39	40	36	38	35	37

Or

26.	**a.**	Calculate the two regression equations from the following data:	**10 Marks**	**L3**	**CO4**

Marks in Statistics	10	25	13	25	22	11	12	25	21	20
Marks in Mathematics	12	22	16	15	18	18	17	23	24	17