

INDIRA UNIVERSITY, PUNE

SCHOOL OF INFORMATION TECHNOLOGY-BSC AIML

Backlog Examination (2025 Pattern) April/ May – 2026 - Semester – I

Subject Name: Notion of Statistical Data Analysis
Subject Code: 25AML102T

Max. Marks: 25
Time: 1:30 Hrs.

Instructions

- All Questions are Compulsory.
- Use of Scientific calculator is allowed.

CO #	Cognitive Ability	Course Outcome
CO1	Remember	Recall and define fundamental statistical concepts.
CO2	Understand	Organize and manage data effectively for analysis.
CO3	Apply	Use statistical thinking to explore and interpret data.

Q1.	Attempt any 5 out of 7. (1 mark each) a) Define primary data. b) Name three common measures of central tendency. c) If $\mu'_1 = 2$, $\mu'_2 = 10$, find μ_2. d) Write the formula for Bowley's coefficient of skewness (S_B). e) Which diagram is used to study correlation analysis? f) If $r = 0.8$, find coefficient of determination. g) Name the two types of ogive curves.	(5 Marks)	CO1
Q2.	Attempt any 2 out of 4. (5 mark each) a) The following data represents the marks obtained by 15 students in a test: 23, 27, 11, 35, 32, 29, 41, 38, 36, 25, 33, 30, 22, 38, 40 Construct a stem-and-leaf chart for the given data. Also answer the following questions: i) Identify the smallest and largest values. ii) Find the median using the stem-and-leaf chart.	(10 Marks)	CO2

	b) The distribution of daily wages is as follows: <table><tr><td>Wages in Rs.</td><td>0-100</td><td>100-200</td><td>200-300</td><td>300-400</td><td>400-500</td></tr><tr><td>No. of workers</td><td>2</td><td>5</td><td>8</td><td>10</td><td>5</td></tr></table> Compute Median and Mean of the above data. c) For the data: 2, 4, 6, 8, 10, calculate the first four raw moments about the origin. d) Explain the concept of Kurtosis with its types. State the measure of kurtosis with its formula. Also interpret its values for identifying the types of kurtosis.	Wages in Rs.	0-100	100-200	200-300	300-400	400-500	No. of workers	2	5	8	10	5			
Wages in Rs.	0-100	100-200	200-300	300-400	400-500											
No. of workers	2	5	8	10	5											
Q.3.	Attempt all questions: (10 Marks) (A) i) In a study of 6 students, study hours (X) and test marks (Y) were recorded: <table><tr><td>X</td><td>1</td><td>2</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>Y</td><td>15</td><td>20</td><td>25</td><td>28</td><td>30</td><td>35</td></tr></table> Find the correlation coefficient between X and Y and interpret it. ii) Apply the regression concept to a real-life situation. Interpret the meaning of regression coefficients. Also describe any three key properties of regression coefficients. OR	X	1	2	2	3	4	5	Y	15	20	25	28	30	35	CO3
X	1	2	2	3	4	5										
Y	15	20	25	28	30	35										
	(B) The monthly incomes (Rs.) of two groups of workers are: <table><tr><td>Group A</td><td>10,000</td><td>12,000</td><td>13,000</td><td>14,000</td><td>15,000</td></tr><tr><td>Group B</td><td>8,000</td><td>10,000</td><td>14,000</td><td>18,000</td><td>20,000</td></tr></table> (i) Compute mean of income of Group A and mean of income of Group B. (ii) Which group of workers has better monthly income? Justify (iii) Compute variance, standard deviation and coefficient of variation for both the groups A and B. (iv) Identify which group has greater income variability with the help of coefficient of variation with proper justification.	Group A	10,000	12,000	13,000	14,000	15,000	Group B	8,000	10,000	14,000	18,000	20,000	CO3		
Group A	10,000	12,000	13,000	14,000	15,000											
Group B	8,000	10,000	14,000	18,000	20,000											
