

USN: _____

Department of Artificial Intelligence and Machine Learning

Course Code: 21AI41

Sem: IV

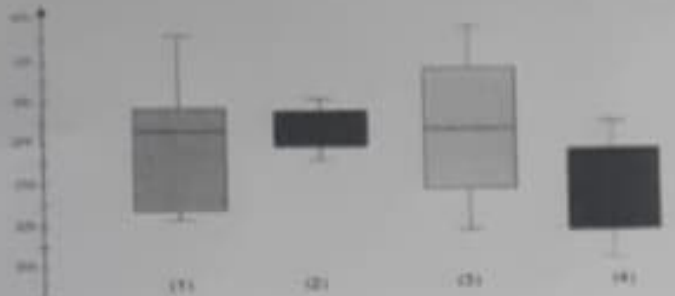
Date: 05-July-2023

Duration: 90 Minutes


CIE-I

Statistics for Data Science

Answer all the Questions

SL. No	Questions	M	BT	CO
1	In an online examination for the course on ML for Sports, the average score was 20 on 50. List out a maximum of 5 possible inferences that you can make about the class performance from this information.	05	04	01
2	There are three ants on a triangle, one at each corner. At a given point of time they all set off for a different corner at random. What is the probability that they don't collide?	05	03	01
3	Given a set of data, we tend to compute the variance and then obtain the standard deviation as the square-root of variance. Why is it that we prefer standard deviation and not variance?	05	01	02
4	 <p>The above boxplot gives the response of students to 4 different questions related to their well-being in college/school. The y-axis is the well-being scale.</p> <p>(a) Which of the 4 questions have a very high level of agreement among students? Justify your answer.</p> <p>(b) Some questions seem to have high levels of disagreement among students, indicating different opinions. Identify the boxplot(s) which indicate this scenario. Justify your answer.</p>	05+05	04	02
5	Let X be the data as given: $X = 2, 3, 4, 0, 1, 2, 3, 4, 2, 9$	10	03	01
	Recall from our class lecture, the geometrical interpretation of mean as a projection of data onto the vector with all its components equal to 1. Use this interpretation to obtain the deviation vector and subsequently, the length of the deviation vector for the data given above.			

Department of Artificial Intelligence and Machine Learning

9	An integer is randomly selected between 1 and 50, inclusively. (a) Find the probability that the number is not divisible by 8. (b) Find the probability that the number is divisible by 9. (c) Find the probability that the number is neither divisible by 8 nor by 9.	01+01 =02	02	02
10	Assess or Reject each of the following statements with proper justification. (a) Statement 1: The histogram given below depicts the average salary offered to students in a college through campus placement programme. (x-axis depicts the salary band)  (b) Statement 2: Disjoint events are necessarily independent events.	01+01	02	02

Course Outcome	
CO1	Apply the knowledge of statistics in providing solutions to some common business problems.
CO2	Perform statistical inferencing on some real-world scientific/business use cases and present the analysis results.
CO3	Investigate the need for distributions, statistical experiments, and significance testing in solving some problems of industry and society.
CO4	Use statistical tools to illustrate the principles of data distribution, data sampling, and data visualization.
CO5	Appraise the knowledge of statistics in data science to build a successful career as an AIML engineer, work in teams, and communicate their ideas effectively.

L-Marks, BT-Blooms Taxonomy Levels, CO-Course Outcomes

Marks Distribution	Particulars	CO1	CO2	CO3	CO4	CO5	L1	L2	L3	L4	L5
	Max Marks	20	30	—	—	—	—	15	20	15	—