Mid Semester Examination 2025

Name of the Program: MTECH | Semester: II

Name of the Course: Data Science & NLP | Course Code: MCS-251

Time: 1:30 Hours | **MM: 50**

Note:

(i) Answer all the questions by choosing any one of the sub questions.

(ii) Each question carries 10 marks

Q.1	**(10 Marks)**	**CO 1**
a.	1. Define **Data Science** and explain its significance in today's world. 2. Differentiate between **structured** and **unstructured** data with examples.	
	OR	
b.	Explain the difference between **quantitative** and **categorical** variables. Provide two real-world examples of each.	
Q.2	**(10 Marks)**	**CO 1**
a.	List the **five steps of a Data Science project** and briefly describe the role of each step.	
	OR	
b.	Write a program in python to compute, Mean, Median, Mode, Range, Average Deviation, Absolute Deviation, Squared Deviation, Standard Deviation, Total Sum of Squares for the following dataset. {18, 22, 33, 11, 9, 4}	
Q.3	**(10 Marks)**	**CO 2**
a.	What is **data preprocessing**, and why is it essential before performing analysis or modeling?	
	OR	
b.	Explain the difference between **supervised** and **unsupervised** machine learning with examples.	
Q.4	**(10 Marks)**	**CO 2**
a.	What are **Precision** and **Recall** in classification models? Why are they important in evaluating model performance?	
	OR	
b.	What is the difference between **population** and **sample** in data science? Why do we use sampling instead of analyzing the entire population?	
Q.5	**(10 Marks)**	**CO 1, CO 2**
a.	Explain the concept of **probability** in Data Science. How is probability useful in machine learning models?	
	OR	

b.	Consider the given data frame.
	# Sample data
	data = {
	'Name': ['Alice', 'Bob', 'Charlie', 'David', 'Emily'],
	'Age': [25, 30, 22, 28, 24],
	'Gender': ['Female', 'Male', 'Male', 'Male', 'Female']
	}
	1: Create a DataFrame from the data
	2: Summary statistics	

	3: Filtering rows based on a condition (Age>25) 4: Adding a new column('Age_Group' =age: 'Young' if age <= 25 else 'Old') 5: Grouping and aggregation 1. 6: Sorting by Age in descending order	