Total No. of Printed Pages: 2

PRN No.	

PAPER CODE	V124-315

May 2024 (ENDSEM) EXAM

F.Y.B. TECH. (SEMESTER - II)

COURSE NAME: Fundamentals of Data Science Branch: AI & DS COURSE CODE: ADUA12236

(PATTERN 2023)

Time: [1Hr. 30 Min] **[Max. Marks: 40]**

(*) Instructions to candidates:

1) **Figures to the right indicate full marks.**
2) **Use of scientific calculator is allowed**
3) **Use suitable data wherever required**
4) **All questions are compulsory. Solve any one sub question from each Question 1 and 2 and any three sub questions each from Questions 3 and 4.**

Q. No.	Question Description	Max. Marks	CO mapped	BT Level
Q.1	a) List the role of each component in the data science workflow and how they contribute to achieving project objectives.	[5]	CO1	Remember
	b) Define the key reasons why data science is important in today's digital age.	[5]	CO1	Remember
Q2	a) Summarize the potential causes of missing data and outliers in datasets and their implications for analysis.	[5]	CO2	Understand
	b) Explain the characteristics and limitations of various data types, such as structured, unstructured, and semi-structured data.	[5]	CO2	Understand
Q.3	a) Calculate the following descriptive statistics for the ages of the students: Mean , Median, Mode, Range Student : 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 Age: 15, 18, 19, 18, 20, 16, 15, 19, 20, 19	[5]	CO3	Apply
	b) Apply common techniques used in exploratory data analysis, such as scatter plots, heatmaps, and correlation matrices using suitable data.	[5]	CO3	Apply
	c) Illustrate the purpose of measure of central tendency in summarizing and interpreting data using a suitable example.	[5]	CO3	Apply
	d)Demonstrate the process to build statistical models to make predictions using an example.	[5]	CO3	Apply

Q.4	a) Illustrate how machine learning adds value to businesses and organizations. Give specific examples where machine learning has led to tangible benefits or competitive advantages.	[5]	CO4	Apply
	b) Apply a supervised learning algorithm to predict whether a customer will make a purchase in the future using a dataset (assume a suitable dataset). Outline the steps involved in preprocessing the data and selecting an appropriate algorithm for this task.	[5]	CO4	Apply
	c) Illustrate the importance of data visualization to present complex datasets and communicating insights effectively take appropriate data.	[5]	CO4	Apply
	d) Apply appropriate unsupervised learning algorithm for a real application, give justification and the process of model training.	[5]	CO4	Apply