Seat No: _____ **Enrollment No: _____**

PARUL UNIVERSITY

FACULTY OF ENGINEERING & TECHNOLOGY

B. Tech Summer 2024 - 25 Examination

Semester: 5 **Date: 12-04-2025**
Subject Code: 303105314 **Time: 10.30am to 1.00pm**
Subject Name: Data Analytics and Data Visualization **Total Marks: 60**

Instructions:

1. This question paper comprises of two sections. Write answer of both the sections in separate answer books.
2. From Section A, **Q.1 is compulsory,** From Section B, **Q.1 is compulsory.**
3. Figures to the right indicate full marks.
4. Draw neat and clean drawings & make suitable assumptions wherever necessary.
5. Start new question on new page.
6. BT- Blooms Taxonomy Levels – Remember-1, Understand -2, Apply-3, Analyse-4, Evaluate-5, Create-6

	SECTION-A	Marks	CO	BT
Q.1	**Answer the following questions.**			
	A. i) Differentiate between data analysis and data analytics. ii) Define Data Visualization. iii) Justify the need for Decision Trees in classification tasks.	**[06]**	**1** **5** **4**	**2** **1** **5**
	B. i) State the importance of Python in Data Analytics. ii) Define Central Tendency and its types. iii) Define Hypothesis Testing and its significance	**[06]**	**2** **2** **3**	**2** **1** **1**
Q.2	**A.** List the four types of data analytics with brief explanations.	**[04]**	**1**	**1**
	B. Differentiate between Linear and Logistic Regression.	**[05]**	**4**	**4**
	OR			
	B. Compare the roles and responsibilities of a data analyst and a data scientist.	**[05]**	**1**	**4**
Q.3	**A.** What are the different types of data used in data analytics?	**[04]**	**1**	**2**
	B. What is an ANOVA test? Explain it with simple example.	**[05]**	**2**	**2**
	OR			
	B. What is Hierarchical clustering, and how does it work?	**[05]**	**4**	**2**
	SECTION-B			
Q.1	**Answer the following questions.**			
	A. i) Define Dendrogram. ii) List three measures of Dispersion. iii) Illustrate the process of data transformation in Power BI.	**[06]**	**4** **3** **5**	**2** **1** **3**
	B. i) Define Linear Regression ii) State what Power BI is and its primary purpose. iii) Illustrate how the Chi-Square test is applied in real-world scenarios.	**[06]**	**4** **5** **3**	**1** **1** **3**
Q.2	**A**. Evaluate the benefits of using Power BI for business analytics.	**[04]**	**5**	**6**
	B. A company wants to estimate the average monthly salary of its employees. A random sample of 50 employees is taken, and their average salary is found to be ₹45,000 with a standard deviation of ₹5,000. Calculate the 95% confidence interval for the true average salary of all employees.	**[05]**	**3**	**3**
	OR			
	B. How does a decision tree handle continuous and categorical data?	**[05]**	**4**	**2**
Q.3	**A.** What are different types of data visualizations available in Power BI? Explain any three.	**[04]**	**5**	**2**

	B. A spam detection model classifies emails as either Spam (Positive) or Not Spam (Negative). The model produces the following confusion matrix: Calculate Accuracy, Precision, Recall and F1-Score.	**[05]**	**4**	**3**

Actual\Predicted	Spam (Positive)	Not Spam (Negative)
Spam	50	10
Not Spam	5	100

	OR			
	B. A dataset contains the following four points in a 1D space: X= {3,5,12,14}. If we apply K-Means clustering with K=2 and choose 3 and 14 as the initial centroids, which cluster will each point belong to after the first iteration? Also, find the new centroids.	**[05]**	**4**	**3**