

GUJARAT TECHNOLOGICAL UNIVERSITY**BE(MINOR)- SEMESTER-I&II EXAMINATION – SUMMER 2025****Subject Code:115AG01****Date:03-06-2025****Subject Name: Introduction to Data Analytics****Time:02:30 PM TO 05:00 PM****Total Marks:70****Instructions:**

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Simple and non-programmable scientific calculators are allowed.

		Marks
Q.1	(a) Define the following terms: training dataset, testing dataset, validation dataset	03
	(b) What is data visualization? Explain data visualization using an example.	04
	(c) Explain the features of the python programming language. List out and explain any four libraries used in python for data analytics.	07
Q.2	(a) Differentiate between data science and data analytics.	03
	(b) Write a short note on exploratory data analysis.	04
	(c) Explain the data analytics process in detail using suitable example.	07
	OR	
	(c) What is data preprocessing? How should one treat missing/null values? How can we deal with them? Give an example.	07
Q.3	(a) Explain the need of data analytics.	03
	(b) Differentiate between classification and clustering applications.	04
	(c) What is machine learning? Explain supervised and unsupervised learning.	07
	OR	
Q.3	(a) Explain the future scope of data science.	03
	(b) How to select the right machine learning model for a given data for data analytics?	04
	(c) What do you mean by feature selection? List out and explain feature selection methods.	07
Q.4	(a) What do you mean by confusion matrix? Explain it in brief.	03
	(b) How can data analytics be applied for face detection application?	04
	(c) What are the different ways in which we can find outliers in the dataset? How to handle them?	07
	OR	
Q.4	(a) What is dimensionality reduction? Explain its significance.	03
	(b) Explain the role of the next generation data scientist.	04
	(c) How will you identify overfitting or under fitting models? How to deal with both the issues?	07
Q.5	(a) What do you mean by categorical variable? Give any one example.	03
	(b) List out various data collection methods used for data analytics. Explain any one in brief.	04
	(c) Write a python program to prepare scatter plot and box plot by considering suitable data.	07
	OR	
Q.5	(a) Give the example of tuple using python.	03
	(b) Explain security and ethical issues in data science.	04
	(c) Write a program to read data from a CSV file and apply data preprocessing using python.	07
