

GYANMANJARI INNOVATIVE UNIVERSITY

GYANMANJARI INSTITUTE OF TECHNOLOGY

B.Tech.-Mid Semester Examination (MSE)- S2026

Enrollment No.: 529

Subject Code: BETIT16327

Date: 23/03/2026

Subject Name: Data Analysis and Visualization

Semester: 6

Time: 02:30PM to 04:30PM

Total Marks: 60

Instructions:

1. Question No. 1 is compulsory.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.

		Marks	
Q.1	(a) Explain the types of data with suitable examples	05	
	(b) Explain different data sources such as databases, APIs, and online repositories with suitable examples.	05	
	(c) Explain normalization and standardization techniques and illustrate when each method should be applied in data preprocessing.	10	
Q.2	(a) Define data cleaning and data wrangling in data preprocessing.	05	
	(b) Define outliers and list different outlier detection methods used in data analysis.	05	
	OR		
Q.2	(b) Differentiate between correlation and covariance with suitable examples.	05	
	(c) Explain how histograms, scatter plots, box plots, heatmaps, and pair plots help in Exploratory Data Analysis (EDA).	10	
	OR		
Q.3	(c) Define Exploratory Data Analysis (EDA), descriptive statistics, and measures of central tendency and dispersion.	10	
	Q.3	(a) Define hypothesis testing and chi-square test.	05
		(b) Differentiate between Matplotlib and Seaborn with suitable examples.	05
(c) Define K-Nearest Neighbor (KNN) classification and K-Means clustering algorithms.		10	
OR			
Q.3	(a) Define Linear Regression and Logistic Regression.	05	
	(b) Define Big Data, NoSQL databases, and components of the Hadoop ecosystem (HDFS, MapReduce, Hive, Pig).	05	
	(c) List and define different chart types such as bar, histogram, line, pie, scatter, heatmap, and violin plots.	10	