

Seat No./Enrollment No.: _____

GUJARAT TECHNOLOGICAL UNIVERSITY
BACHELOR OF ENGINEERING – SEMESTER 5 – EXAMINATION – SUMMER 2026

Subject Code: 3151608

Date: 02/06/2026

Subject Name: Data Science

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.**

- Q.1**
- (a) Define population and sample. **03**
 - (b) Differentiate Cross-sectional, Time Series and Panel Data. **04**
 - (c) Discuss how descriptive, predictive and prescriptive analytics techniques can be applied in healthcare applications. Explain with simple examples. **07**
- Q.2**
- (a) State Bayes' Theorem. **03**
 - (b) What is estimation? Differentiate point estimation and interval estimation. **04**
 - (c) Explain various measures of central tendency with suitable examples. **07**
- OR
- Q.2**
- (c) Explain various measures of shapes with suitable examples. **07**
- Q.3**
- (a) Explain how Maximum Likelihood Estimation can be used to estimate parameters. **03**
 - (b) Discuss CLT-Central Limit Theorem. **04**
 - (c) Why is the choice of measurement scale important in statistical analysis? Explain nominal, ordinal, interval, and ratio scales with examples. **07**
- OR
- Q.3**
- (a) Explain how sampling distribution can be used to make statistical inferences. **03**
 - (b) Discuss Method of Moments. **04**
 - (c) When percentile is preferred over percentage? Explain percentile, quartile and decile with examples. **07**
- Q.4**
- (a) Discuss applications of linear regression in data science. **03**
 - (b) Explain Simple Linear Regression model building steps. **04**
 - (c) Differentiate discrete and continuous probability distributions. **07**
- OR
- Q.4**
- (a) Discuss purpose of Ordinary Least Squares method in linear regression. **03**
 - (b) Explain steps for validation of Simple Linear Regression model. **04**
 - (c) Differentiate Chi-Square Distribution, Student's t-Distribution and F-Distribution. **07**
- Q.5**
- (a) Explain how logistic regression can be used in credit rating. **03**
 - (b) Differentiate Gain Chart and Lift Chart **04**

- (c) A retail company wants to predict whether a customer will buy a product based on features such as age, income, past purchase history, etc. Explain how decision tree method can be applied to build this model. **07**

OR

- Q.5** (a) Explain purpose of Optimal Cut-Off Probability in logistic regression. **03**
- (b) Differentiate Sensitivity and Specificity. **04**
- (c) A bank wants to predict whether a customer will repay a loan on time using features such as income, credit history, age, and employment status. Explain how bagging, random forests, and boosting can be applied to improve the accuracy of this prediction. **07**
