GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Industry Led Minor/Hons.

Level: UG

Branch: Data Science

Course / Subject Code : BE05IAR011

Course / Subject Name: Demystifying Data Analytics

w. e. f. Academic Year:	2026-2027
Semester:	5
Category of the Course:	Core Course

Prerequisite:	• Basic programming knowledge (preferably in Python). • Basic understanding of mathematics, including probability and statistics. • Familiarity with data handling and basic visualization techniques.
Rationale:	This course aims to equip students with essential data science skills, focusing on data analysis, statistical inference, and machine learning. The rationale behind this syllabus is to provide hands-on experience with real-world data, enabling students to make data-driven decisions and understand key concepts in probability, hypothesis testing, and statistical modeling. These skills are foundational for careers in data science, analytics, and related fields.

Course Outcome:

After Completion of the Course, Student will be able to:

No	Course Outcomes	RBT Level
01	Understand Python libraries' role in data science	U
02	Derive insights from real datasets	N
03	Visualize data effectively	A
04	Build predictive machine learning models	C
05	Apply probability concepts including Bayes theorem	A
06	Compute statistical measures and interpret result	N
07	Understand confidence intervals and hypothesis testing	U
08	Compare treatments using statistical techniques	E

**Revised Bloom's Taxonomy (RBT)*

Teaching and Examination Scheme:

Teaching Scheme (in Hours/week)			Total Credits	Assessment Pattern and Marks				Total Marks
				Theory		Tutorial / Practical		
L	T	PR	C	ESE (E)	PA / CA (M)	PA/CA (I)	ESE (V)	
3	0	2	4	100	0	0	0	100

*Total Lecture Hrs. (L) =45	Total Practical Hrs. (PR) =30.	Total Hours =75 Hrs

GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Industry Led Minor/Hons.

Level: UG

Branch: Data Science

Course / Subject Code : BE05IAR011

Course / Subject Name: Demystifying Data Analytics

Course Content:

Unit No.	Content	No. of Hours	% of Weightage
1.	**Python for Data Science:** What is Machine Learning? Linear Regression Single Variable, Linear Regression Multiple Variables, Gradient Descent and Cost Function, Save Model Using Joblib And Pickle, Dummy Variables & One Hot Encoding, Training and Testing Data, Logistic Regression (Binary Classification), Logistic Regression (Multiclass Classification), Decision Tree, Support Vector Machine (SVM), Random Forest, K Fold Cross Validation, K Means Clustering	15	20
2.	Introduction to Numpy, Operations on Numpy arrays, Image as a Numpy matrix, Pandas - Introduction to Pandas, Introduction to Pandas objects, working with datasets, Operations in Pandas, Pandas Plots, Matplotlib - Introduction to Data Visualization, Introduction to Matplotlib, Object Oriented Interface in Matplotlib, Types of plots, Scikit-learn, Machine learning using skitlearn	15	20
3.	**Probability and Statistics using Python:** Introduction to Probability, Understanding a Chance, The term Probability, Uses of Probability, Probabilistic Model - Introduction, Set Theory, Venn Diagrams, Counting Techniques, Introduction - Counting Techniques, Product Rule, Sum Rule, Combination, Permutation, Probability, Introduction - Probability, Axioms, Elementary properties, Computing Probability, Extending Conditional Probability, Introduction to Statistics, Types of statistics, Measures of central tendency, Arithmetic mean, Weighted mean, Geometric mean, Harmonic mean, Median, Mode, Introduction - Dispersion, Range, Variance, Standard deviation, z-score, Quartiles, Five number summary, Interquartile range, Introduction - Measures of shape, Skewness, Kurtosis,	15	20
4.	**Statistical Inference using Python** - Introduction to Statistical Inference, Case Study, Definition, Scope of Statistical Inference, Sampling and Sampling Methods, Introduction to Sampling, Determining Sample Size, Sampling Techniques, Sampling Methods - Exercise, Quiz, Parameter Estimation - Introduction to Parameter Estimation, Point Estimator of Population Mean, Sampling Distribution of Means, Confidence Interval of Population Mean, Point Estimate of Population Variance, Sampling Distribution of Sample Variance, Confidence Interval of Population Variance, Sampling Distribution of Variance - Exercise, Hypothesis Testing, Why hypothesis testing ?, Steps Involved in Hypothesis Testing, Performing Hypothesis Test to Validate the Estimates, Comparing Two Populations, Estimating	15	20

	Difference Between Two Population Means, Sampling Distribution of Difference in Sample Means, Confidence Interval of Difference in Two Population Means, Comparing Two Populations by Conducting Two Sample Test		
5.	**Comparing More Than Two Populations** - One-Way ANOVA, Two-Way ANOVA, ANOVA, Case Study.	15	20
	Total	**75**	**100**

Suggested Specification Table with Marks : Given here tentative, which may vary as per Author and Course.

Distribution of Marks (in %)					
R Level	**U Level**	**A Level**	**N Level**	**E Level**	**C Level**
10%	**25%**	**25%**	**10%**	**5%**	**15%**

Where R: Remember; U: Understanding; A: Application, N: Analyze and E: Evaluate C: Create (as per Revised Bloom's Taxonomy)

Skill & Practical Activities to be carried out during Semester						
Important Note:- Please keep only applicable categories relevant to your offerings in this table and delete not applicable categories"						
Sr. No.	**Category of Engagement**	**Describe the activities to be carried out by students in brief**	**Expected Frequency & Duration**	**Mode of Delivery (Online / Offline / Hybrid)**	**Tools / Platforms / Equipment / Machinery to be Used**	**Expected major Learning Outcomes in 2 or 3 bullet points**
1	Tutorials	Complete the lectures	Twice a week	Online	Infosys Springboard platform/app	• Apply machine learning algorithms to

						real-world datasets. • Perform data preprocessing, visualization, and statistical analysis using Python.
1	Master Classes / Expert Lectures by Industry Professionals	Attend the live class, ask doubt and complete practical	5 Days Master class twice in a semester (20 Hours)	Online	Infosys Springboard/VS Code/Jupiter notebook	Students will learn hands on machine learning and Python
2	Case Study	• Build ML models (Regression, Classification, Clustering). • Perform data cleaning, analysis, and visualization. • Apply probability, statistics, and	15 Hour	Guided on platform	Infosys Springboard/Vs Code/Jupiter notebook	• Apply ML algorithms on real datasets. • Perform data preprocessing & visualization. • Interpret statistical results for decision-making.

		hypothesis testing. • Conduct ANOVA and inference on datasets				
3	Quizzes	Complete quizzes on portal	After each module	online	Infosys springboard platform/app	Better understanding about topic
4	Hands-on Training / Lab Exercises / Tool-Based Learning	Complete exercises	2–3-hour / week	Can be taken online on platform/offline in own system	Infosys springboard platform/app	Model building • Performance evaluation

* * * * * * *