



# GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Minor/Hons. Program

Level: UG

Branch: Minor/Hons. -Artificial Intelligence & Machine Learning

Subject Code : BE050AG011

Subject Name : Introduction to Data Analytics

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

<b>Prerequisite:</b>	Basic Programming
<b>Rationale:</b>	Introduction to Data Analytics equips students with the ability to collect, clean, analyze, and interpret data for informed decision-making. It builds a foundation for understanding trends, patterns, and insights across engineering, business, and research domains. The course prepares learners for advanced analytics, AI, and data-driven problem solving in real-world applications.

### Course Outcomes:

Sr. No.	CO statement	Marks% weightage
1	Explain how data is collected, managed and stored for data science;	20
2	Understand the key concepts in data science, including their real-world applications and the toolkit used by data scientists;	30
3	Implement data collection and management scripts using MongoDB	30
4	Design end to end data pipeline for given domain.	20

### Teaching and Examination Scheme:

Teaching / Learning Scheme (in Hours per semester)					Total Credits = TH/30	Assessment Pattern and Marks					Total Marks
L	T	P	PBL*	TH		Theory		Tutorial / Practical			
						ESE (E)	PA (M)	PA (I)	PBL (I)	ESE (V)	
45	00	30	45	120	4	70	00	00	30	50	150

\* **Problem Based Learning (PBL)** aims to accommodate learning beyond syllabus as per clause 9.4 of NBA manual.

Where L = Lecture, T= Tutorial, P= Practical, TW/SL = Term-Work / Self-Learning, TH = Total Hours, PA = Progressive Assessment, ESE = End-Semester Examination



# GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Minor/Hons. Program

Level: UG

Branch: Minor/Hons. -Artificial Intelligence & Machine Learning

Subject Code : BE050AG011

Subject Name : Introduction to Data Analytics

## Content:

Sr. No.	Content	Total Hrs
1.	Introduction to Data Science, Different Sectors using Data science, Purpose and Components of Python in Data Science.	06
2.	Data Analytics Process, Knowledge Check, Exploratory Data Analysis (EDA), EDA- Quantitative technique, EDA- Graphical Technique, Data Analytics Conclusion and Predictions.	09
3.	Feature Generation and Feature Selection (Extracting Meaning from Data)- Motivating application: user (customer) retention- Feature Generation (brainstorming, role of domain expertise, and place for imagination)- Feature Selection algorithms.	10
4.	Data Visualization- Basic principles, ideas and tools for data visualization, Examples of inspiring (industry) projects- Exercise: create your own visualization of a complex dataset.	10
5.	Applications of Data Science, Data Science and Ethical Issues- Discussions on privacy, security, ethics- A look back at Data Science- Next-generation data scientists.	10
<b>TOTAL</b>		<b>45</b>

## Suggested Specification table with Marks (Theory):

Distribution of Theory Marks					
R Level	U Level	A Level	N Level	E Level	C Level
20	30	20	20	10	-

**R: Remembrance; U: Understanding; A: Application, N: Analyze and E: Evaluate C: Create and above Levels (Revised Bloom's Taxonomy)**

## Reference Books:

1. Joel Grus, Data Science from Scratch, Shroff Publisher Publisher /O'Reilly Publisher Media
2. Annalyn Ng, Kenneth Soo, Numsense! Data Science for the Layman, Shroff Publisher Publisher
3. Cathy O'Neil and Rachel Schutt. Doing Data Science, Straight Talk from The Frontline. O'Reilly Publisher Media.
4. Jure Leskovek, Anand Rajaraman and Jeffrey Ullman. Mining of Massive Datasets. v2.1, Cambridge University Press.
5. Jake VanderPlas, Python Data Science Handbook, Shroff Publisher Publisher /O'Reilly Publisher Media
6. Philipp Janert, Data Analysis with Open Source Tools, Shroff Publisher Publisher /O'Reilly Publisher Media.



# GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Minor/Hons. Program

Level: UG

Branch: Minor/Hons. -Artificial Intelligence & Machine Learning

Subject Code : BE050AG011

Subject Name : Introduction to Data Analytics

## List of Practical:

1. Python Environment setup and Essentials.
2. Mathematical computing with Python (NumPy).
3. Scientific Computing with Python (SciPy).
4. Data Manipulation with Pandas.
5. Prediction using Scikit-Learn
6. Data Visualization in python using matplotlib

## List of learning website:

- 1) SWAYAM/NPTEL courses

## List of suggested activities for Problem Based Learning:

No.	Description	No. of hours	Total hours
1.	Students will have to undergo industrial visit related to any one of the contents mentioned in syllabus and submit the report for the same	Visit = 5h, Report preparation = 5h Total = 10h	10
2.	Students will have to undergo technical video based learning related to subject and have to present it or prepare a report of that learning	Duration of video = 5 hrs Report Preparation = 5hrs	10
3.	Students will have to submit assignment work assigned to them.	5 assignments of 2h each.	10
4.	Discussion on research paper based on relevant subject. Summarize research paper and evaluation critical parameters	3 research papers can be included. Each one has 5h	15
5.	Student have to undergo an on-line course	Minimum duration of the course should be 10h.	10
6.	Students will have to prepare poster, chart or PowerPoint presentation on technical topic related to subject content	Poster/Chart/Power point preparation and presentation	10
7.	Students will have to prepare Working/non-working model on technical topics	For each model 10 h duration	10
8.	Real world case studies-based learning	Duration of data collection/study = 5h Report preparation = 5h	10
<b>Maximum Hours to be allotted</b>			<b>45</b>