Year: B. Tech III (Semester V)

Subject Name: Introduction to Data Science
Type of course: Minor (Group: Data Science)

Subject Code: BTEA19551

Teaching and Examination Scheme:

Teaching Scheme				Theory Marks			Practical Marks		Total
L	T	P	C	TEE	CA1	CA2	TEP	CA3	
3	0	2	4	60	25	15	30	20	150

CA1: Continuous Assessment (assignments / projects / open book tests / closed book tests) CA2: Sincerity in attending classes / class tests / timely submissions of assignments / self-learning attitude / solving advanced problems TEE: Term End Examination TEP: Term End Practical Exam (Performance and viva on practical skills learned in course) CA3: Regular submission of Lab work / Quality of work submitted / Active participation in lab sessions / viva on practical skills learned in course.

Contents:

Sr. No.	Contents	Total Hours
1.	Introduction to Data Science, Different Sectors using Data science, Purpose and Components of Python in Data Science.	08
2.	Data Analytics Process, Knowledge Check, Exploratory Data Analysis (EDA), EDA- Quantitative technique, EDA- Graphical Technique, Data Analytics Conclusion and Predictions.	08
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.	11
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.	08

Suggested Specification table with Marks (Theory): (For B. Tech only)

Distribution of Theory Marks					
R Level	U Level	A Level	N Level	E Level	C Level
15	20	20	5	-	-

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

Reference Books:

Sr. No.	Title of book /article	Author(s)	Publisher and details like ISBN
1	Business Analytics: The Science of Data - Driven Decision Making	U Dinesh Kumar	John Wiley & Sons
2	Introducing Data Science: Big Data, Machine Learning, and More, Using Python Tools	Davy Cielen	John Wiley & Sons
3	Data Science from Scratch	Joel Grus	O'Reilly Publisher Media
4	Numsense! Data Science for the Layman	Annalyn Ng, Kenneth Soo	Shroff Publisher
5	Doing Data Science, Straight Talk from The Frontline	Cathy O'Neil and Rachel Schutt	O'Reilly Publisher

Note: Students should refer to the latest editions of books

Course Outcomes (CO):

Sr. No.	CO statements	Marks % weightage
CO-1	Describe how data is collected, managed and stored for data science.	25%
CO-2	Explain the key concepts in data science, including their real-world applications and the toolkit used by data scientists.	40%
CO-3	Demonstrate different data visualization methods on complex dataset.	20%
CO-4	Discuss ethical and privacy issues in data science applications.	15%

Lab Work:

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