GUJARAT TECHNOLOGICAL UNIVERSITY

BACHELOR OF ENGINEERING SYLLABUS

Minor Degree : DATA SCIENCE

Subject Code : N116AT01

Subject Name : Computational Data Analytics

WEF Academic Year :	2025 - 26
Semester :	6
Category of the Course :	Compulsory

Course Objective :

- To learn how to think about your study system and research question of interest in a systematic way in order to design an efficient sampling and experimental research program.
- To understand how to analyze collected data to derive the most information possible about your research questions.

Course Scheme :

Teaching Scheme			Total Credits	Assessment Pattern and Marks				Total Marks
				Theory		Practical		
L	T	PR	C	ESE (E)	PA(M)	ESE (V)	PA (I)	
3	0	2	4	70	0	30	0	100

Course Content :

Sr. No.	Course Content	No. of Hours
1	**Module 1 :** Introduction to R Computing language. Best practices in executing Reproducible Research in data science, Sampling and Simulation. Descriptive statistics, and the creation of good observational sampling designs.	6
2	**Module 2 :** Data visualization, Data import and visualization, Introduction to various plots.	8
3	**Module 3 :** Frequentist Hypothesis Testing, Z-Tests, Power Analysis.	10
4	**Module 4 :** Linear regression, diagnostics, visualization, Likelihoodist Inference, Fitting a line with Likelihood, Model Selection with one predictor.	10
5	**Module 5 :** Bayesian Inference, Fitting a line with Bayesian techniques, Multiple Regression and Interaction Effects, Information Theoretic Approaches.	8
Total		**42**

References / Text Book :

1. Practical Data Science with R, Nina Zumel, John Wiley & Sons.
2. N. C. Das, Experimental Designs in Data Science with Least Resources, Shroff Publisher.
3. Hadley Wickham, Garret Grolemund, R for Data Science, Shroff Publisher/O'Reilly Publisher.
4. Benjamin M. Bolker. Ecological Models and Data in R. Princeton University Press, 2008. ISBN 978-0-691-12522-0.
5. John Fox and Sanford Weisberg. An R Companion to Applied Regression. Sage Publications, Thousand Oaks, CA, USA, second edition, 2011. ISBN 978-1-4129-7514- 8.

Lab Work :

1. To give a basic insight of R and its various libraries.
2. Libraries in R. R as a Data Importing Tool, Dplyr. Forcats.
3. Simulation and Frequentist Hypothesis testing, Simulation and Power.
4. Bayesian computation in R, Fitting a line with Bayesian techniques.

Course Outcome :

After Completion of the Course, Student will able to :

No.	Course Outcomes
01	Explain how data is collected, managed and stored for data science.
02	When to use which type of Machine learning model.
03	Implement various ML algorithms on data models.

* * * * * * *