

**Year: B. Tech III (Semester VI)**

**Subject Name:** Computational Data Analytics  
**Type of course:** Minor (Group: Data Science)  
**Teaching and Examination Scheme:**

**Subject Code:** BTEA19651

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 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.	07
2.	Data visualization, Data import and visualization, Introduction to various plots	09
3.	Frequentist Hypothesis Testing, Z-Tests, Power Analysis	10
4.	Linear regression, diagnostics, visualization, Likelihood Inference, Fitting a line with Likelihood, Model Selection with one predictor	10
5.	Bayesian Inference, Fitting a line with Bayesian techniques, Multiple Regression and Interaction Effects. Information Theoretic Approaches	09

**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	Practical Data Science with R	Nina Zumel	John Wiley & Sons.

2	Experimental Designs in Data Science with Least Resources	N. C. Das	Shroff Publisher Publisher
3	R for Data Science	Hadley Wickham, Garret Golemund	Shroff Publisher/O'Reilly Publisher Publisher
4	Ecological Models and Data in R	Benjamin M. Bolker	University Press
5	An R Companion to Applied Regression	John Fox and Sanford Weisberg	Sage Publications

**Note:** Students should refer to the latest editions of books

**Course Outcomes (CO):**

Sr. No.	CO statements	Marks % weightage
CO-1	Explain how data is collected, managed and stored for data science	30%
CO-2	Select and apply appropriate Machine learning model to real world problems	40%
CO-3	Implement ML algorithms using R computing	30%

**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, Forecats.
3. Simulation and Frequentist Hypothesis testing, Simulation and Power.
4. Bayesian computation in R, Fitting a line with Bayesian techniques.

