



**MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT  
(AUTONOMOUS)**

**2250546: DATA SCIENCE**

**III Year B.Tech. DS I – Sem.**

<b>L</b>	<b>T</b>	<b>P</b>	<b>C</b>
<b>3</b>	<b>0</b>	<b>0</b>	<b>3</b>

**Prerequisites:** Data Base Management Systems

**Course Objectives:**

- To use different functions in R, how to read data into R, accessing R packages, writing R functions, debugging, and organizing data using R functions.
- To learn basics of statistical data analysis with examples.

**Course Outcomes:** The students should be able to

- Understand basic terms what Statistical Inference means.
- Identify probability distributions commonly used as foundations for statistical modelling. Fit a model to data
- Describe the data using various statistical measures
- Utilize R elements for data handling
- Perform data reduction and apply visualization techniques

**UNIT – I**

**Introduction** to Data Science, Data Analytics, Data Manipulation, Data Import Techniques, Exploratory Data Analysis, Data Visualization.

**Introduction to R:** What is R? – Why R? – Advantages of R over Other Programming Languages - R Studio: R command Prompt, R script file, comments – Handling Packages in R: Installing a R Package, Few commands to get started: installed.packages(), packageDescription(), help(), find.package(), library() - Input and Output – Entering Data from keyboard – Printing fewer digits or more digits – Special Values functions : NA, Inf and–inf.

**UNIT –II**

**R Data Types:** Vectors, Lists, Matrices, Arrays, Factors, Data Frame – **R - Variables:** Variable assignment, Data types of Variable, Finding Variable ls(), Deleting Variables - **R Operators:** Arithmetic Operators, Relational Operators, Logical Operator, Assignment Operators, Miscellaneous Operators - **R Decision Making:** if statement, if – else statement, if– else if statement, switch statement – **R Loops:** repeat loop, while loop, for loop - Loop control statement: break statement, next statement.

**UNIT –III**

**R-Function** :function definition, Built in functions: mean(), paste(), sum(), min(), max(), seq(), user-defined function, calling a function, calling a function without an argument, calling a function with argument values - R-Strings – Manipulating Text in Data: substr(), strsplit(), paste(), grep(), toupper(), tolower() .

**UNIT – IV**

**Data Frames** –Create Data Frame, Data Frame Access, Understanding Data in Data Frames: dim(), nrow(), ncol(), str(), Summary(), names(), head(), tail(), edit() functions - Extract Data from Data Frame, Expand Data Frame: Add Column, Add Row - Joining columns and rows in a Data frame rbind() and cbind() – Merging Data frames merge() – Melting and Casting data melt(), cast().



MARRI LAXMAN REDDY INSTITUTE OF TECHNOLOGY AND MANAGEMENT  
(AUTONOMOUS)

**UNIT – V****Data Visualization in R**

Understanding on Data Visualization, graphical functions present in R, plot various graphs like tableplot, histogram, Boxplot, customizing Graphical Parameters to improvise plots, understanding GUIs like Deducer and R Commander, introduction to Spatial Analysis.

**Loading and handling Data in R: Getting and Setting the Working Directory** – getwd(), setwd(), dir() - R-CSV Files - Input as a CSV file, Reading a CSV File, Analyzing the CSV File: summary(), min(), max(), range(), mean(), median(), apply() - Writing into a CSV File – R -Excel File – Reading the Excel file.

**TEXT BOOK:**

1. Sandip Rakshit, R Programming for Beginners, McGraw Hill Education (India),2017, ISBN :978-93-5260-455-5.

**REFERENCES:**

1. Seema Acharya, Data Analytics using R, McGrawHill Education (India), 2018, ISBN: 978-93-5260-524-8.
2. Andrie de Vries, Joris Meys, R for Dummies A Wiley Brand, 2nd Edition, John Wiley and Sons, Inc, 2015, ISBN: 978-1-119-05580-8