Certificate Course Offered by Department of Information Technology

Name of the course: Basic R Programming

Course Code: 401004/IT100A

Objectives

In this course students will learn how to program in R and how to use R for effective data analysis and visualization. They would learn the foundations – how to install R and load data into it – and continue with data manipulation, visualization, and implementation of standard statistical functions. Students will be able to understand Probability and Sampling Distributions and learn the creative applications of Linear Regression in multivariate context for predictive purposes.

Course Outcomes

Sl. No.	Description
CO 1	Possess an in-depth understanding of the R programming environment and of the requirements for, and programming implications of, writing code using basic R
CO 2	Extract data from files and other sources and perform various data manipulation tasks on them.
CO 3	Acquire the skills needed to successfully develop general-purpose programming applications in the R environment.
CO4	Learn the basic statistical concepts, followed by application of these concepts using R.
CO5	Use R Graphics and Tables to visualize results of various statistical operations on data .
CO6	Implement Probability and Probability Distributions to solve a wide variety of problems.
CO7	Apply the knowledge of R gained to data Analytics for real life applications.

Syllabus

1. Introduction to R and RStudio

Overview of R - Introduction to R and RStudio, Installing R and RStudio, Command Line and Script, Functions and Packages, Management of Code and Other Files.

2. Basic Object Types and Operations in R

Basic Data Types and Data Structures in R - Strings, Factors, Vectors and Simple Manipulations, Arrays and Matrices, Lists and Data Frames, Operations in Data Frames.

3. Control Structures and Functions

Decision making and Loops in R - Sample programs, loop functions and the debugging tools in R, Introduction to functions, Examples of functions.

4. Descriptive Statistics using R

Measurement of Central Tendency – Mean, Median and Mode (Using R). Measurement of Variation - Range, IQR and Standard Deviation (Using R), Descriptive Statistics Using psych Package.

5. Data Visualization using R

Base-R and ggplot2 Graphics, Creating graphs - scatterplot, histogram, boxplot, time series plot.

6. Probability Distribution

R Functions for Normal Distribution - rnorm, pnorm, qnorm and dnorm, R Functions for Binomial Distribution - rbinom, pbinom, qbinom and dbinom, R Functions for Poisson Distribution - rpois, ppois, qpois and dpois. Basic Statistics- Correlation and Covariance, T-Tests, ANOVA.

7. Linear Regression

Linear Models, Simple Linear Regression, -Multiple Regression Generalized Linear Models, Logistic Regression, - Poisson Regression- other Generalized Linear Models-Survival Analysis, Nonlinear Models, Splines- Decision- Random Forests.

Practical Sessions

1. Installing R and RStudio (Windows),
2. Data Import and Export. Creating and Saving the R Script File.

3. Practice basic R commands and create simple R programs using mathematical operations.
4. Simple programs using Control structures Using debugging tools
5. Examples of functions
6. Creating vectors, arrays and matrices using R commands
7. Sample programs to implement matrix addition and subtraction
8. Sample programs to implement matrix multiplication, transpose and inverse matrix functions.
9. Creating Lists, Data frames and Factors.
10. Simple programs using Control structures Using debugging tools
11. Examples of functions.
12. Measurement of Central Tendency – Mean, Median and Mode (Using R functions)
13. Measurement of Variation - Range, IQR and Standard Deviation (Using R)
14. Descriptive Statistics Using psych Package
15. Dirty Data handling: remove null values from a dataset.
16. Plot scatterplot, histogram, boxplot, and time series plot.
17. Plotting Normal Distribution Using R Functions
18. Plotting Binomial Distribution Using R Functions.
19. Plotting Poisson Distribution Using R Functions.
20. Project on statistical data modeling

References

1. SandipRakshit, "Statistics with R Programming", McGraw Hill Education, 2018.
2. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, "AN Introduction to Statistical Learning: with Applications in R", Springer Texts in Statistics, 2017.
3. Joseph Schmuller, "Statistical Analysis with R for Dummies", Wiley, 2017.
4. K G Srinivasa, G M Siddesh, ChetanShetty, Sowmya B J, "Statistical Programming in R", Oxford Higher Education, 2017.

Resource Persons

1. **Dr. Neeba E A**
   Associate Professor & HoD
   Department of Information Technology
   Email: neebaea@rajagiritech.edu.in

2. **Dr. Ranju S Kartha**
   Assistant Professor
   Department of Information Technology
   Email: ranjusk@rajagiritech.edu.in

Online Division of Teaching & Learning

	Topic	Resource Person	Total Hours
1	Introduction to R and RStudio	Dr. Neeba E A	4
2	Basic Object Types and Operations in R	Dr. Neeba E A	8
3	Control Structures and Functions	Dr. Neeba E A	8
4	Descriptive Statistics using R	Dr. Ranju S Kartha	8
5	Data Visualization using R	Dr. Ranju S Kartha	6
6	Probability Distribution	Dr. Ranju S Kartha	8
7	Linear Regression	Dr. Ranju S Kartha	8

Mode of Delivery: Online/ Offline

Duration: 22 Hrs. of Theory & 28 Hrs. of Practical Session

Fee Structure

The registration fee for inhouse candidates - **Rs 500/-**

The registration fee for external candidates - **Rs 1000/-**

Eligibility Criteria

This certificate course is mainly for the students pursuing B.Tech. in Computer Science, Information Technology, Electronics and Communication, Applied Electronics and Instrumentation, B.Sc. Computer Science & Electronics. Those who are completed plus two are also eligible for this course.

Evaluation scheme

Assignment/Quiz: Total Marks: 4**0**

Assignments will be provided after the completion of each module.

Exam (Marks: 6**0**)

The exam will be conducted after the completion of the entire course.

Cut off mark

Those students who acquire a minimum of **60** marks from both the assignments and the exam will be eligible to get the certificate

Project

The interested inhouse students will get a chance to carry out a project after the successful completion of the course.