GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Bachelor of Engineering
Level : UG
Subject Code: 3174604
Branch: Computer Science and Engineering (Data Science)
Semester: VII
Subject Name: R Programming for Data Analytics

Type of Course: Open Elective III
Prerequisite: knowledge in at least one programming language such C/Java/Python

Rationale:.

Teaching and Examination Scheme:

Teaching Scheme			Credit	Examination Marks				Total Marks
				Theory Marks		Practical Marks		
L	T	P	C	ESE (E)	PA (M)	ESE (V)	PA (I)	
2	0	2	3	70	30	30	20	150

Course Outcomes:

CO1: Utilize R Data types for developing programs.
CO2: Make use of different R Data Structures.
CO3: Develop programming logic using R Packages.
CO4: Analyze the datasets using R programming capabilities
CO5: Apply R programming for reading, cleaning, visualizing and analyzing data

Course Content:

Unit No.	Content	Total Lectures	weightage
1	Introduction to Algorithm and programming concepts. 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.	2	10

GUJARAT TECHNOLOGICAL UNIVERSITY

Program Name: Bachelor of Engineering

Level : UG

Subject Code: 3174604

Branch: Computer Science and Engineering (Data Science)

Semester: VII

Subject Name: R Programming for Data Analytics

2	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. Solving problems from Assignment sheet.	4	20
3	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 - R Vectors – Sequence vector, rep function, vector access, vector names, vector math, vector recycling, vector element sorting - R List - Creating a List, List Tags and Values, Add/Delete Element to or from a List, Size of List, Merging Lists, Converting List to Vector - R Matrices – Accessing Elements of a Matrix, Matrix Computations: Addition, subtraction, Multiplication and Division- R Arrays: Naming Columns and Rows, Accessing Array Elements, Manipulating Array Elements, Calculation Across Array Elements - R Factors –creating factors, generating factor levels gl().	4	20
4	String fnctions : grep(), nchar() , paste(), sprintf(), substr(), strsplit(), regex() gregexpr(), toupper(), tolower(), paste(),bitwOr(value1,value2), bitwXor(value1,value2), bitwNot(valoe), bitwAnd(value1,value2),bitwShiftL(value,shift), bitwShiftR(value,shift)	2	5
5	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(). Loading and handling Data in R: Getting and Setting the Working Directory – getwd(), setwd(), dir() File Handling in R language, -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.	6	20

6	Installing RMySQL Package, Creating Database, table under MYSQL, Inserting data in a table , Update and alter table, Display content of table.	2	5
7	Descriptive Statistics: Data Range, Frequencies, Mode, Mean and Median: Mean Applying Trim Option, Applying NA Option, Median - Mode - Standard Deviation – Correlation - Spotting Problems in Data with Visualization: visually Checking Distributions for a single Variable - R –Pie Charts: Pie Chart title and Colors – Slice Percentages and Chart Legend, 3D Pie Chart – R Histograms – Density Plot - R – Bar Charts: Bar Chart Labels, Title and Colors. Line Chart, Scatterplot, Developing graphs, Box Plot, Drawing line, circle, rectangle, triangle using R language .	6	20

Reference Books: :

1) Introduction to Statistics and Data Analysis With Exercises, Solutions and Applications in R Authors: Heumann, Christian, Schomaker, Michael, Shalabh, Publisher" Springer 2016
2) The R Software-Fundamentals of Programming and Statistical Analysis -Pierre Lafaye de Micheaux, Rémy Drouilhet, Benoit Liquet, Springer 2013
3) A Beginner's Guide to R (Use R) By Alain F. Zuur, Elena N. Ieno, Erik H.W.G. Meesters, Springer 2009
4) "The Book of R" by Tilman M. Davies, no starch press(San Francisco)
5) "The Art of R programming" by Norman Matloff, no starch press(San Francisco)

List of Laboratory/Learning Resources Required:

R Tutorial (https://www.w3schools.com/r/)
R Studio: RStudio Desktop - Posit (https://posit.co/download/rstudio-desktop/)

Other Resources/MOOCs:

NPTEL: Advanced R Programming for Data Analytics in Business
By Prof. Abhinava Tripathi | IIT Kanpur

Essentials of Data Science With R Software-1: Probability and Statistical Inference
By Prof. Shalabh | IIT Kanpur

Practical List:

1. Download and install R-Programming environment and install basic packages using install.packages() command in R.
2. Learn all the basics of R-Programming (Data types, Variables, Operators etc.
3. Implement R-Loops with different examples

4. Learn the basics of functions in R and implement with examples.
5. Implement data frames in R. Write a program to join columns and rows in a data frame using cbind() and rbind() in R
6. Implement different String Manipulation functions in R.
7. Implement different data structures in R (Vectors, Lists, Data Frames).
8. Write a program to read a csv file and analyze the data in the file in R.
9. Create pie charts and bar charts using R.
10. Create a data set and do statistical analysis on the data using R

****