

Branch Name:	Artificial Intelligence And Data Science
Program Code:	ET108
Semester/Year:	Semester V / Third Year
Course Name:	R Programming for Data Science
Course Code:	2ET1080502T
Pre-requisite:	Knowledge of statistical theory in mathematics. Understanding of different types of graphs for data representation and most important is that we should have prior knowledge of any programming.

Course Objective:

1. Understand the use of the R and R Studio interactive environment
2. Expand R by installing R packages
3. Utilize and R Data types for developing programs.
4. Understand the different data structures in R

Teaching and Examination Scheme:

Credit	Teaching Scheme Hours per week				Evaluation Scheme (Marks)								Total (Marks)
					Theory (Marks)				Practical (Marks)				
	L	T	P	Total	SEE		CCE		SEE		CCE		
					Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.	
3	3	-	-	3	24	60	16	40	-	-	-	-	100

Subject Contents

Sr. No.	Topic	Total Hours	Weightage %
1	Introduction to R Programming: What is R? Installing R and RStudio, RStudio Overview, Working in the Console, Arithmetic Operators, Logical Operations, Using Functions, Getting Help in R and Quitting RStudio	06	14
2	R-Data Types, Operators and Packages: Data types, Variables: Variable assignment, Finding Variable ls(), Deleting Variables, Operators, Decision Making, R Loops, Vectors, Lists, Matrices, Arrays, Factors, Data Frame, R Packages	07	17
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 Text in Data: substr(), strsplit(), paste(), grep(), toupper(), tolower() - 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().	07	17

4	R Studio IDE: Environment setup with R Studio, R Commands, R command Prompt, R script file, comments – Handling Packages in R: Installing R Packages, 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.	08	19
5	Data querying: SQL and R Writing SQL statements in R, Using the Select, From, Where, Is, Like, Order By, Limit, Max, Min SQL functions	08	19
6	R Data Visualization: R Pie Charts: Pie Chart title and Colors – Slice Percentages and Chart Legend, 3D Pie Chart, R Histograms, RBoxplots, R Bar Charts, R Line Graphs, R Scatterplots	06	14

List of Reference Books:

1. Jared P Lander, R for everyone: advanced analytics and graphics, Pearson Education, 2013
2. Dunlop, Dorothy D., and Ajit C. Tamhane. Statistics and data analysis: from elementary to intermediate. Prentice Hall, 2000.
3. G Casella and R.L. Berger, Statistical Inference, Thomson Learning 2002.
4. P. Dalgaard. Introductory Statistics with R, 2nd Edition. (Springer 2008)
5. Michael Berthold, David J. Hand, Intelligent Data Analysis, Springer
6. Hastie, Trevor, et al. The elements of statistical learning. Vol. 2. No. 1. New York: springer, 2009.
7. Montgomery, Douglas C., and George C. Runger. Applied statistics and probability for engineers. John Wiley & Sons, 2010
8. Joseph F Hair, William C Black et al , “Multivariate Data Analysis” , Pearson Education, 7th edition, 2013.
9. Mark Gardener, “Beginning R - The Statistical Programming Language”, John Wiley & Sons, Inc., 2012.
10. W. N. Venables, D. M. Smith and the R Core Team, “An Introduction to R”, 2013.

Course Learning Outcomes (CLO): On completion of this course, the students will be able to:

CLO	Description	Bloom's Taxonomy Level
CLO1	Show the installation of R Programming Environment	1 Remembering 2 Understanding
CLO2	Utilize and R Data types for developing programs.	1 Remembering 2 Understanding
CLO3	Make use of different R Data Structures.	4 Analyze
CLO4	Develop programming logic using R Packages.	3 Applying
CLO5	Analyze the datasets using R programming capabilities	3 Applying

Mapping of CLOs with POs & PSOs

Course Learning Outcomes	Program Outcomes (POs)												Program Specific Outcomes (PSOs)			
	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO 10	PO 11	PO 12	PSO 1	PSO 2	PSO 3	PSO 4
CLO1	H	M			M	L	H		M	L		H	H	M	L	
CLO2	H	H			M			L		M		M		H	H	
CLO3	H	L	M		M	L		L		M	L			H		L
CLO4	H	M	H	M	M			M			M	L	L		M	H
CLO5	H	M	H	H	L	M			M		L	L		M	H	

H: High, M: Medium, L: Low