GITAM INSTITUTE OF MANAGEMENT (GIM)
Gandhi Institute of Technology and Management (GITAM)
(Declared as Deemed to be University u/s 3 of UGC Act. 1956)
Visakhapatnam – 45.

Course Code: MAN 227	**Course Title: Data Analysis with R**	
Semester: III	**Course Type: Core**	**Credits: 3**
Home Programme(s): BBA (Business Analytics)		**Batch: 2020 - 2023**
Course Leader: Dr. M. Kamakshaiah		

Course description and learning outcomes

R is an open source programming language for statistical computing and graphics. Being open source, it has found huge acceptance among data scientists and is one of the popular tool for data science and machine learning.

Learning Objectives

- Understand the programming concepts of R (L1)
- Gain hands on experience in working with R (L2)

On successful completion of this course, students will be able to:

CO	Course Outcome	Assessment
CO1	Understand the elements of R programming	A1,
CO2	Write basic programs in R language	A1, A2, A3
CO3	Write programs in R using control structures	A1, A2, A3
CO4	Perform statistical analysis in R language	A2, A3, A4
CO5	Perform data visualization with R	A2, A3, A4

Course outline and indicative content

UNIT–I (CO1, L1)

Introduction to R ; Concept of R, Installing R, IDE of R, Getting help from R, Mathematical Operators and Vectors, Assigning Variables, Special Numbers, Logical Vectors, Classes, Different types of numbers, Changing classes, Examining Variables, The workplace, Vectors – Sequences, Lengths, Names, Indexing Vectors, Vector Recycling and Repetition, Matrices and Arrays – Creating Arrays and Matrices, Rows, Columns, Dimensions, Indexing Arrays, Combining Matrices, Array Arithmetic,

UNIT–II (CO2, L1)

Lists, Functions, Strings and Factors : Lists – Creating lists, Automatic and recursive variables, List dimensions and arithmetic, indexing lists, Conversion between vectors and lists, Combining lists, NULL, Pairlists, Data Frames – Creating Data Frames, Indexing Data Frames, Basic Data Frame Manipulation, Environments, Functions – Creating and Calling Functions, Passing functions, variable scope, Strings – Constructing and printing strings, Formatting numbers, Special characters, Changing case, Extracting Substrings, Splitting Strings, File paths, Factors – Creating, factor levels, ordered factors, conversion of variables

UNIT–III (CO3, L2)

Flow Controls ; Conditional – if and else, Vectorized if, Multiple Selection, Loops – repeat loops, while loops, for loops, Advanced looping – replication, looping over lists, looping over arrays, Multiple – Input Apply, Instant vectorization, Split-Apply-Combine

UNIT–IV (CO4, L2)

Statistics with R : Summarizing data, Calculating relative frequencies, Tabulating Factors and creating contingency tables, Testing categorical variables for independence, Calculating Quantiles of a dataset, Converting data into z-scores, t-test, testing sample proportions, testing normality, comparing means of two samples, testing correlation for significance, Linear regression in R, Logistic Regression in R Clustering with R

UNIT–V (CO5, L2)

Packages and Visualization : Loading packages, search path, libraries and installed packages, installing packages, maintaining packages, Visualization – The three plotting systems, Scatterplots – base graphics, lattice graphics, ggplots, Line Plots, Histograms, Box Plots, Bar Charts, Other plotting packages and systems

Assessment methods

Task	Task type	Task mode	Weightage (%)
A1	Quiz (2)	Individual	40
A2	Assignments / Lab Tasks / Written Test/Coursera Groups* or Individual (3)	Individual / Group	30
A3	Project	Individual	20
A4	Presentation	Individual	10

Mapping Cos – Blooms Levels – Assessment Tools

Knowledge dimension / Cognitive dimension	L1. Remember	L2. Understand	L3. Apply	L4. Analyze	L5. Evaluate	L6. Create
Factual Knowledge	CO1 (A1, A2)					
Conceptual Knowledge	CO1 (A1, A2)	CO1, CO2 (A1, A2, A3)	CO2, CO3, CO4 (A1, A2, A3)	CO3 (A3)	CO3, CO4(A3, A4)	
Procedural Knowledge		CO1, CO2, CO3 (A1, A2, A3)	CO2 (A3, A4)	CO4 (A3)	CO3, CO4 (A3, A4)	CO3, CO4 (A3, A4)
Meta Cognitive Knowledge						CO3, CO4(A3, A4)

Learning and teaching activities

Classroom Lectures, Problem solving exercises, Demonstration, Lab Sessions

Teaching and learning resources

Computer Lab, R Package, Textbooks, Ebooks, Reference Materials, Web resources

CO PO Mapping

This is to map the level of relevance of the Course Outcome (CO) with Programme Outcome (PO). 0= No Relevance; 1= Low Relevance; 2= Medium Relevance; 3= High Relevance

CO PO Mapping							
	PO1	PO2	PO3	PO4	PO5	PO6	Sum
CO1	0	0	2	0	0	0	6
CO2	0	1	3	3	3	3	14
CO3	0	3	3	3	2	3	14
CO4	0	2	2	3	3	3	13
CO5	0	1	3	1	2	3	10
Target Level Max.	0	9	12	11	10	15	58

Program Outcomes

1	Ability to understand the business problem with their knowledge in different functional areas of management
2	Ability to work with structured, semi – structured and unstructured data.
3	Ability to use tools such as Microsoft Excel, SPSS, R, Weka and Tableau to solve business analytics problem
4	Ability to apply analytics techniques to analyze and interpret the data.
5	Ability to perform descriptive, predictive and prescriptive analytics.
6	Have necessary skills and understanding to take up advanced topics in the area of analytics and thus enhance their career.