Course Name: Statistical Data Analytics with R **Subject Code:** TMC 105

Program Name: Master of Computer Applications (MCA)

1 Contact Hours: 45 **L** 3 **T** 0 **P** 0

2 Examination Duration(Hrs): **Theory** 0 3 **Practical** 0 0

3 Relative Weightage: **CWE:** 25 **MTE:** 25 **ETE:** 50

4 Credits: 0 3

5 Semester: [*] [ ] [ ]
Autumn **Spring** **Both**

6 Pre-Requisite: Basic concepts of statistics and mathematics.

7 Subject Area: Computer Science

8 Objective: To familiarize students with the concept of statistics and provide a practical introduction to R programming language to understand the patterns in data for becoming a good data analyst.

9 Course Outcome:

A student who successfully fulfills the course requirements will be able to:

CO 1 Understand the concepts of statistics.
CO 2 Apply the probability distribution techniques in different applications.
CO 3 Learn how to import, pre-process, evaluate, manipulate and summarize data-sets in R.
CO 4 Apply the concepts of objects, data types and functions in R.
CO 5 Select appropriate statistical test for data analysis.
CO 6 Create visualizations with R in data Analytics.

10 Details of the Course:

Unit No.	CONTENT	CONTACT HOURS
1	**Statistics:** Introduction to Statistics- Descriptive Statistics, Summary Statistics Basic probability theory, Statistical Concepts (uni-variate and bi-variate sampling, distributions, re-sampling, statistical Inference, prediction error),	7
2	**Probability Distribution:** Introduction to Probability**,** Probability Distribution (Continuous and discrete- Normal, Bernoulli, Binomial, Negative Binomial, Geometric and Poisson distribution), Bayes' Theorem, Central Limit theorem, Data Exploration & preparation, Concepts of Correlation, Regression, Covariance, Outliers.	8
3	**Introduction to R and Data Preprocessing**: Introduction &	8

	Installation of R, R Basics, Finding Help, Code Editors for R, Command Packages, Manipulating and Processing Data in R, Reading and Getting Data into R, Exporting Data from R	
4	**Objects and Data Types**: Data Objects-Data Types & Data Structure. Viewing Named Objects, Structure of Data Items, Manipulating and Processing Data in R (Creating, Accessing, Sorting data frames, Extracting, Combining, Merging, reshaping data frames), Control Structures	14
5	**Functions:** Functions in R (numeric, character, statistical), working with objects, Viewing Objects within Objects, Constructing Data Objects, Building R Packages, Running and Manipulating Packages, Non parametric Tests- ANOVA, chi-Square, t-Test, U-Test, Introduction to Graphical Analysis, Using Plots(Box Plots, Scatter plot, Pie Charts, Bar charts, Line Chart), Plotting variables, Designing Special Plots, Simple Liner Regression,	8
	TOTAL	**45**

11 Suggested Books:

Sl. NO.	**NAME OF AUTHERS/BOOKS/PUBLISHERS**	**YEAR OF PUBLICATION**
1	Dr. Mark Gardener, Beginning R: "The Statistical Programming Language", John willey& Sons, 2012	2012
3	John M. Quick, "Statistical Analysis with R", Pckt Publishing.	2010