Faculty of Computer Science & Applications
Master of Computer Application with Cyber Security
(W. E. F.: 2023-24)

Document ID: SUTEEB-01

Name of Faculty	:	Faculty of Computer Science & Applications
Name of Program	:	Master of Computer Application with Cyber Security
Course Code	:	2MCY01
Course Title	:	Statistical computing using R
Type of Course	:	Professional core
Year of Introduction	:	2023-24

Prerequisite	:	Programming Languages, statistics and mathematics, Graph and most importantly zeal to learn
Course Objective	1	Students exercise the fundamentals of statistical analysis in R environment.
	2	Students can analysis data for the purpose of exploration using descriptive and Inferential Statistics.
	3	Students will understand Probability and Sampling distributions and learn the creative application of Linear Regression in multivariate context for predictive purpose.
Course Outcomes	:	After learning the course the students will be able to:
	CO1	To understand the basic programming concepts of R programming language.
	CO2	To understand the data structures in R Statistical computing programming language.
	CO3	To understand the important packages and functions in R statistical computing programming language.
	CO4	To understand the importance of R in statistical analysis and customizing the analysis.
	CO5	To understand the impact of R in the current analytical organization over proprietary statistical software.

Teaching and Examination Scheme

Teaching Scheme (Contact Hours)			Credits	Examination Marks				
				Theory Marks		Practical Marks		Total Marks
L	T	P	C	SEE	CIA	SEE	CIA	
3	0	2	4	70	30	30	20	150

*Legends: **L**-Lecture; **T**–Tutorial/Teacher Guided Theory Practice; **P**– Practical,**C** – Credit, **SEE** – Semester End Examination, **CIA** - Continuous Internal Assessment (It consists of Assignments/Seminars/Presentations /MCQ Tests, etc.))*

Course Content

Unit No.	Topics	Teaching Hours	Weightage	Mapping with CO
1	**Introduction to R Programming:** R and R Studio, Logical Arguments, Missing Values, Characters, Factors and Numeric, Help in R, Vector to Matrix, Matrix Access, Data Frames, Data Frame Access, Basic Data Manipulation Techniques, Usage of various apply functions – apply, lapply, sapply and tapply, Outliers treatment	07	20%	CO1
2	**Descriptive Statistics:** Types of Data, Nominal, Ordinal, Scale and Ratio, Measures of Central Tendency, Mean, Mode and Median, Bar Chart, Pie Chart and Box Plot, Measures of Variability, Range, Inter-QuartileRange, Standard Deviation, Skewness and Kurtosis, Histogram, Stem and Leaf Diagram, Standard Error of Mean and Confidence Intervals	07	20%	CO2 CO3
3	**Probability, Probability& Sampling Distribution:** Experiment, Sample Space and Events, Classical Probability, General Rules Of Addition, Conditional Probability, General Rules For Multiplication, Independent Events, Bayes' Theorem, Discrete Probability Distributions: Binomial, Poisson, Continuous Probability Distribution, Normal Distribution & t-distribution, Sampling Distribution and Central Limit Theorem	07	20%	CO4
4	**Statistical Inference and Hypothesis Testing:** Population and Sample, Null and Alternate Hypothesis, Level of Significance, Type I and Type II Errors, One Sample t Test, Confidence Intervals, One Sample Proportion Test, Paired Sample t Test, Independent Samples t Test, Two Sample Proportion Tests, One Way Analysis of Variance and Chi Square Test	08	20%	CO2
5	**Correlation and Regression:** Analysis of Relationship, Positive and Negative Correlation, Perfect Correlation, Correlation Matrix, Scatter Plots, Simple Linear Regression, R Square, Adjusted R Square, Testing of Slope, Standard Error of Estimate, Overall Model Fitness, Assumptions of Linear Regression, Multiple Regression, Coefficients of Partial Determination, Durbin Watson Statistics, Variance Inflation Factor.	07	20%	CO5

Suggested Distribution of Theory Marks Using Bloom's Taxonomy						
Level	Remembrance	Understanding	Application	Analyse	Evaluate	Create
Weightage	**20**	**30**	**30**	**20**	**0**	**0**

NOTE: This specification table shall be treated as a general guideline for the students and the teachers. The actual distribution of marks in the question paper may vary slightly from above table.

Suggested List of Experiments/Tutorials

Sr. No.	Name of Experiment/Tutorial	Teaching Hours
1	Install and configure R, set working directory.	01
2	Install Packages and calling installed packages.	01
3	R studio environment and functionalities of R studio.	01
4	Implement basic R operations (data input, missing values, importing data into R using different formats: xlsx, CSV, Text files).	02
5	Use R as a calculator.	01
6	Explore various functionalities of data frames.	01
7	Create data set using data frames, list and tables.	01
8	Calculate the interest earned after 5 years on an investment of $2000,	02
9	Assuming an interest rate of 3% compounded annually.	01
10	Calculate the remainder after dividing 31079 into 170166719	01

Major Equipment/ Instruments and Software Required

Sr. No.	Name of Major Equipment/ Instruments and Software
1	RStudio is a free and open-source IDE (integrated development environment) for programming in R.
2	I3/ I5 processor; 8GB RAM; 250GB HDD

Suggested Learning Websites

Sr. No.	Name of Website
1	https://www.w3schools.com/r/
2	https://www.javatpoint.com/r-tutorial
3	https://www.programiz.com/r

Reference Books

Sr. No.	Name of Reference Books
1	Introduction to Scientific Programming and Simulation using R – Owen Jones, Robert Maillardet and Andrew Robinson, CRC Press
2	Advanced R – Hadley Wickham, CRC Press.
3	The art of R programming – Norman Matloff, no starch Press, San Francisco.