Semester-VI

Skill Enhancement Paper

SEC-4: Statistical Software: R

Total Marks: 100 (Theory: 38, Internal Assessment: 12, and Practical: 50)
Workload: 2 Lectures, 4 Practicals (per week) **Credits:** 4 (2+2)
Duration: 14 Weeks (28 Hrs. Theory + 56 Hrs. Practical) **Examination:** 2 Hrs.

Course Objectives: The purpose of this course is to help you begin using **R**, a powerful free software program for doing statistical computing and graphics. It can be used for exploring and plotting data, as well as performing statistical tests.

Course Learning Outcomes: This course will enable the students to:
  i)   Be familiar with **R** syntax and use **R** as a calculator.
 ii)   Understand the concepts of objects, vectors and data types.
iii)   Know about summary commands and summary table in **R**.
 iv)   Visualize distribution of data in **R** and learn about normality test.
  v)   Plot various graphs and charts using **R**.

Unit 1: Getting Started with R - The Statistical Programming Language
Introducing **R**, using **R** as a calculator; Explore data and relationships in **R**; Reading and getting data into **R:** combine and scan commands, viewing named objects and removing objects from **R**, Types and structures of data items with their properties, Working with history commands, Saving work in **R**; Manipulating vectors, Data frames, Matrices and lists; Viewing objects within objects, Constructing data objects and their conversions.

Unit 2: Descriptive Statistics and Tabulation
Summary commands**:** Summary statistics for vectors, Data frames, Matrices and lists; Summary tables.

Unit 3: Distribution of Data
Stem and leaf plot, Histograms, Density function and its plotting, The Shapiro−Wilk test for normality, The Kolmogorov−Smirnov test.

Unit 4: Graphical Analysis with R
Plotting in **R**: Box-whisker plots, Scatter plots, Pairs plots, Line charts, Pie charts, Cleveland dot charts, Bar charts; Copy and save graphics to other applications.

References:
1. Bindner, Donald & Erickson, Martin. (2011). *A Student's Guide to the Study, Practice, and Tools of Modern Mathematics*. CRC Press, Taylor & Francis Group, LLC.
2. Gardener, M. (2012). *Beginning* **R:** *The Statistical Programming Language*, Wiley Publications.

Additional Reading:
 i.   Verzani, John (2014). *Using R for Introductory Statistics* (2nd ed.). CRC Press, Taylor & Francis Group.

Practicals to be done in the Computer Lab using Statistical Software R:
[1] Chapter 14 (Exercises 1 to 3). [2] Relevant exercises of Chapters 2 to 5, and 7.

Note: The practical may be done on the database to be downloaded from https://data.gov.in/

Teaching Plan (Theory of SEC-4: Statistical Software: R):
Weeks 1 to 3: Introducing **R**, using **R** as a calculator; Explore data and relationships in **R**, Reading and getting data into **R**: Combine and scan commands, viewing named objects and removing objects from **R**, Types and structures of data items with their properties, Working with history commands, Saving work in **R**.
　　　[1] Chapter 14 (Sections 14.1 to 14.4).
　　　[2] Chapter 2.
Weeks 4 and 5: Manipulating vectors, Data frames, Matrices and lists; Viewing objects within objects, Constructing data objects and their conversions.
　　　[2] Chapter 3.
Weeks 6 to 8: Summary commands**:** Summary statistics for vectors, Data frames, Matrices and lists; Summary tables.
　　　[2] Chapter 4.
Weeks 9 to 11: Stem and leaf plot, Histograms, Density function and its plotting, The Shapiro–Wilk test for normality, The Kolmogorov-Smirnov test.
　　　[2] Chapter 5.
Weeks 12 to 14: Plotting in **R**: Box-whisker plots, Scatter plots, Pairs plots, Line charts, Pie charts, Cleveland dot charts, Bar charts; Copy and save graphics to other applications.
　　　[1] Chapter 14 (Section 14.7).
　　　[2] Chapter 7.

Facilitating the Achievement of Course Learning Outcomes

Unit No.	Course Learning Outcomes	Teaching and Learning Activity	Assessment Tasks
1.	Be familiar with **R** syntax and use **R** as a calculator. Understand the concepts of objects, vectors and data types.	(i) Topics to be explained with illustrations using **R** software. (ii) Students to be given homework/assignments. (iii) Students to be encouraged to look for new applications.	• Presentations and participation in discussions. • Assignments and class tests. • Mid-term examinations. • Practical examinations. • End-term examinations.
2.	Know about summary commands and summary table in **R**.		
3.	Visualize distribution of data in **R** and learn about normality test.		
4.	Plot various graphs and charts using **R**.		

Keywords: Objects, Vectors, Data types, Summary commands, Shapiro–Wilk test, Bar charts.