Ganpat University
॥ विद्यया समाजोत्कर्षः ॥

Faculty of Computer Applications

Programme	B.Sc. IT Honours (Data Science)	**Branch**	Computer Applications
Semester	VI	**Version**	1.0.0.0
Effective from Academic Year	2026-27	**Effective for the batch Admitted in**	June 2024
Subject code	U76A2IRP2	**Subject Name**	INTRODUCTION TO R PROGRAMMING - II

Teaching scheme						**Examination scheme (Marks)**			
(Per week)	**Lecture (DT)**		**Practical (Lab.)**		**Total**		**CCE**	**SEE**	**Total**
	L	TU	P	TW					
Credit	2	-	2	-	4	Theory	50	50	100
Hours	2	-	4	-	6				

Objective:

To enable students to perform advanced data analysis in R using modern tools for data wrangling, statistical modeling, visualization, and reporting and to build interactive dashboards and reports suitable for real-world data science applications

Pre-requisites:

Basics of R Programming, Statistics

Learning Outcome:

Name of CO	Description
CO1	Perform data cleaning, transformation, and summarization using tidyverse tools
CO2	Generate advanced data visualizations using ggplot2 and interpret patterns and insight
CO3	Apply statistical tests and regression models to analyze data
CO4	Analyze and interpret time series data
CO5	Build reports with R Markdown and develop basic Shiny dashboards

Mapping of CO and PO:

Cos	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	2	2	3	1	2	2	2	1	2	2
CO2	3	3	3	2	3	1	2	2	3	2	2	2
CO3	3	3	2	3	3	1	2	2	2	1	1	2
CO4	3	3	2	3	3	1	2	2	2	2	1	1
CO5	3	3	3	2	3	1	2	3	3	2	3	3

Content:

Unit	Content	Hrs.
1	**Data Manipulation** Introduction to Tidyverse, Filtering and selecting data, Mutating and summarizing, Grouped operations, Tidying data with gather(), spread(), pivot_longer(), Joining datasets	06
2	**Advanced Data Visualization** ggplot2: layers, aesthetics, color, themes, labels, facets, plot saving, interactive plots using plotly	06

3	**Statistical Analysis** Hypothesis testing, Correlation and Covariance, Linear regression , ANOVA , Residual analysis and model fit, Use of lm() and interpretation of results	06
4	**Introduction to Time Series** Date and Time handling (lubridate package), Time series data basics and plotting, Moving average and smoothing, Decomposing time series	06
5	**R Markdown, Shiny, and Reporting** Introduction to R Markdown, Creating HTML and PDF reports, Embedding plots and tables in reports, Introduction to Shiny for web apps, Building a basic interactive dashboard, Deploying and sharing reports/dashboards	06

Practical Content:	
List of practical specified by subject teacher based on above mentioned topics	
Reference Books:	
1	Advanced R (2nd Ed.) by Hadley Wickham CRC Press, 2019.
2	Practical Statistics for Data Scientists by Peter Bruce, Andrew Bruce & Peter Gedeck,, O'Reilly Media, 2020.
3	R Graphics Cookbook by Winston Chang O'Reilly Media, 2013.
Web Reference:	
1	https://adv-r.hadley.nz/
2	https://www.datacamp.com/doc/r/category/r-documentation
3	https://www.geeksforgeeks.org/r-language/r-tutorial/
MOOC/Certificate Course:	
1	www.coursera.org/learn/advanced-r
2	www.coursera.org/specializations/r
3	www.alison.com/course/introduction-to-data-structures?show_modal=true
Question Paper Scheme:	
	End Semester Examination Duration: (2 Hours Theory Examination) **Note for Examiner: -** Q-1 Any Five out of Seven (25 Marks) Q-2 Any Two out of Three (06 Marks) Q-3 Mandatory question (05 Marks) Q-4 Any Two out of Three (08 Marks) Q-5 Any Two out of Three(06 Marks) *The question paper must comprehensively address all Course Outcomes (COs), align with Bloom's Taxonomy levels, and ensure complete syllabus coverage