| A picture containing text  Description automatically generated | **MAN809** | **Data Analytics with R** | L | T | P | J | S | C |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 2 |  | 2 |  |  | 3 |
| Course Owner | Department of Business Analytics | Syllabus version | | | | 1.0 | |
| Course Pre-requisite(s) |  | Contact hours | | | | 45 | |
| Course Co-requisite(s) |  | Date Approved | | | |  | |
| Alternate Exposure |  | | | | | | |

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 business applications.

**Course Objectives**

1. Understand the programming concepts of R
2. To handle data on R platform
3. To be able to Descriptive Analytics using R
4. To be able to do Predictive Analytics using R
5. To Build Machine Learning models using R

| **UNIT - I** | **Elements of R** | **No of Hours:9** | |
| --- | --- | --- | --- |
| Concept of GNU/R, IDE of R, Mathematical Operators and Vectors, Assigning Variables, Special Numbers, Logical Vectors, Classes, Different types of numbers, Changing classes, Examining Variables, The workplace, Elements in R – Vectors, Matrices and Arrays, Lists, Conversion between vectors and lists, Combining lists, Data Frames | | | |
|  | | | |
| **Learning Outcomes:** | | | |
| After completion of this unit, the student will be able to | | | |
| Install the R and R studio | | | L1 |
| Understand the mathematical operators | | | L2 |
| Understand the Different types of numbers | | | L4 |
| Change the class of objects | | | L3 |
| Deal with matrices and arrays | | | L2 |
|  | | | |
| **Pedagogy tools**: Blended learning,Case let, video lectures, self-reading | | | |
|  | | | |
| **UNIT - II** | **Functions, Strings and Factors and Flow Controls** | | **No of Hours:9** |
| Environments, Functions, Strings, Factors, 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 | | | |
|  | | | |
| **Learning Outcomes:** | | | |
| After completion of this unit, the student will be able to | | | |
| Create list | | L3 | |
| Convert the lists into other types of data | | L3 | |
| Do the basic data manipulation | | L3 | |
| Create basic functions | | L3 | |
| Construct the strings | | L4 | |
|  | | | |
| **Pedagogy tools**: Blended learning,Case let, video lectures, self-reading | | | |
|  | | | |
| **UNIT - III** | **Packages and Visualization** | **No of Hours:9** | |
| 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 | | | |
|  | | | | |
| **Learning Outcomes:** | | | | |
| After completion of this unit, the student will be able to | | | | |
| Implement if and else statement | | L3 | | |
| Do multiple selection on data frames | | L3 | | |
| Implement different looping techniques | | L4 | | |
| Do the instant vectorization | | L4 | | |
| Can split the data frame | | L4 | | |
|  | | | | |
| **Pedagogy tools**: Blended learning,Case let, video lectures, self-reading | | | | |
|  | | | | |
| **UNIT - IV** | **Computing Statistics and Exploratory Data Analysis with R** | **No of Hours:9** | | |
| 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, Variations, Missing Values, Covariation, Patterns and Models | | | | |
|  | | | | |
| **Learning Outcomes:** | | | | |
| After completion of this unit, the student will be able to | | | | |
| Summarize the data | | L3 | | |
| Create contingency tables | | L3 | | |
| Normalize the dataset | | L4 | | |
| Testing on sample data | | L4 | | |
| Do the regression analysis | | L4 | | |
|  | | | | |
| **Pedagogy tools**: Blended learning,Case let, video lectures, self-reading | | | | |
|  | | | | |
| **UNIT - V** | **Machine Learning and Model Building with R** | **No of Hours:9** | | |
| Types of machine learning algorithm, supervised learning algorithms – Linear regression in R, Logistic Regression in R Unsupervised Learning in R -Clustering with R, Recommendation Algorithms, Steps to generate recommendations in R, Model Building: Model basics, Type of Models, Visualizing models – Predictions, Residuals, Model Building, Communicating results – Basics of R Markdown | | | | |
|  | | | | |
| **Learning Outcomes:** | | | | |
| After completion of this unit, the student will be able to | | | | |
| Deal with R packages | | L2 | | |
| Visualize the given dataset | | L2 | | |
| Create different plots | | L2 | | |
| Deal with packages meant for visualization | | L2 | | |
|  | | | | |
| **Pedagogy tools**: Blended learning,Case let, video lectures, self-reading | | | | |
| **Course Outcomes:**   * Differentiate different programming elements of R * Write programs in R with flow control * Work with R packages * Perform statistical analysis using R * Build machine learning model using R | | | | |
| **Textbook(s):** | | | | |
| Abraham Silberschatz, Henry F Korth, Database System Concepts, McGraw Hill Education | | | | |
| **Additional Reading** | | | | |
| **Reference Book(s):** | | | | |
| A. Hoffer Jeffrey, V. Ramesh, Topi Heikki, Modern Database Management, Pearson | | | | | |
| Andrew Couch, Microsoft Access Plain & Simple | | | | | |
| **Journal(s):** | | | | | |
| Journal of Business Analytics, Taylor & Francis | | | | | |
| **Website(s):** | | | | | |
| 1. | | | | | |

***Practical Experiments***

| **Topics** | | | | | | | | | | **Type (Experiment, Project, Exercise) Choose an item.** | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Creating data frames in R | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Creating matrices and arrays | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Creating and Calling Functions | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Constructing and printing strings | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Looping over lists and arrays | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Tabulating Factors and creating contingency tables | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Normalizing the data | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Testing the normality | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Classification algorithms with R | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Clustering algorithms with R | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Dealing with R packages | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| Data visualization | | | | | | | | | | Lab Experiment | | | | | | | | | | | | |
| **Text Books** | | | | | | | | | | | | | | | **Topics** | | | | | | | |
| **1: Gardener, M (2013), Beginning R, New Delhi: Wiley India.** | | | | | | | | | | | | | | | **All** | | | | | | | |
| **Additional Reading** | | | | | | | | | | | | | | | | | | | | |  | |
| **Reference Book(s):** | | | | | | | | **Topics** | | | | | | | | | | | | | | |
| **1**: Teetor, P. (2014), R Cookbook, Mumbai: O’ Reilly India / Shroff Publishers.  2.Cotton, R. (2014), Learning R, Mumbai : O’ Reilly India / ShroffPublishers. | | | | | | | | **All** | | | | | | | | | | | | | | |
|  | **Programme Objectives (POs)** | | | | | | | | | | | | | | | | **PSOs** | | | |  | |
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | | **8** | | **9** | **10** | **11** | **12** | | **1** | | **2** | **3** |  | |
| **CO1** | **0** | **2** | **2** | **1** | **1** | **3** | **1** | | **1** | |  |  |  |  | |  | |  |  |  | |
| **CO2** | **0** | **2** | **2** | **1** | **1** | **3** | **2** | | **1** | |  |  |  |  | |  | |  |  |  | |
| **CO3** | **0** | **2** | **2** | **1** | **1** | **3** | **2** | | **3** | |  |  |  |  | |  | |  |  |  | |
| **CO4** | **0** | **2** | **2** | **1** | **1** | **3** | **2** | | **3** | |  |  |  |  | |  | |  |  |  | |
| **CO5** | **0** | **2** | **2** | **1** | **1** | **3** | **2** | | **3** | |  |  |  |  | |  | |  |  |  | |

1-Low, 2- Medium and 3- High Correlation