Marwadi University
Faculty of Diploma Studies
Information and Communication Technology

# Subject Code: 09CT0401
# Subject Name: Foundation of Data science Tools
# Diploma Year – II ( Semester IV)

**Diploma Branches in which this subject is offered:** Information and Communication Technology

**Objective:**
As the amount of data suddenly rises with exponential way, there is need to extract meaningful insights from this data. This course is designed to introduce you to a range of topics and concepts related to the data analytics process. It will cover tools used in data collection, processing, analysis and visualisation.

**Credits Earned:** 02 Credits

**Course Outcomes:** After completion of this course, student will be able to

- Understand process involved in data Science
- Apply various functionality of excel for data analysis
- Understand the fundamental syntax of R
- Apply critical programming language concepts such as data types, vectors, lists, matrices, arrays, factors, data frames, variables
- Analyse a data set in R and present findings
- Visualize data attributes using pie chart, bar chart, histograms line graphs, scatterplots

**Pre-requisite of course:** Elementary knowledge of excel and programming

**Teaching and Examination Scheme**

| Teaching Scheme (Hours) | | | Credits | Theory Marks | | | Tutorial/ Practical Marks | | Total Marks |
|---|---|---|---|---|---|---|---|---|---|
| Theory | Tutorial | Practical | | ESE | IA | CSE | Viva | Term work | |
| 0 | 0 | 4 | 2 | 0 | 30 | 20 | 25 | 25 | 100 |

**Contents:**

| Unit | Topics | Contact hours |
|---|---|---|
| 1 | **Introduction to data analytics**<br>Types of Data Analysis, Data Mining, Business Intelligence, Statistical Analysis, Predictive Analytics, Text Analytics, Data Analysis Process, Data Analysis Process, Data Collection, processing, cleaning and analysis | 6 |

| | | |
|---|---|---|
| **2** | **Data analytics with Excel**<br>Introduction to Excel, reading data, manipulating data, Basic excel operations and functions, Tables, Sort, Filter, Conditional Formatting, Charts, Pivot Tables, Tables, What-If Analysis, Solver, Analysis ToolPack | **12** |
| **3** | **Introduction to R and R Studio**<br>Evolution of R, Features of R, R environment set up, installation of R Studio, Introduction to R Studio | **6** |
| **4** | **Basics of R language**<br>Basic Syntax of R, R command Prompt, R script file, comments, R-data types, vectors, Lists, Matrices, Arrays, Factors, Data Frames, Variables | **6** |
| **5** | **Operators in R**<br>Arithmetic Operators, Logical Operators, Relational Operators, Assignment Operators, Miscellaneous Operators | **6** |
| **6** | **R-Decision and Control Loop Statements**<br>if condition, if else condition, switch condition, repeat loop, while loop, for loop, break statement, Next statement | **6** |
| **7** | **Flavors in R**<br>Functions in R, R-Strings, String Manipulations, R-vectors, R-Lists, R-Matrices, R-arrays, R-factors, R-data frames, R-Packages, R-Data Reshaping, R-Excel Files, Mean, median and Mode | **8** |
| **8** | **Data Representation using R**<br>R-Pie Chart, R-Bar Chart, R-Histograms R-Line Graphs, R-Scatterplots | **6** |
| **TOTAL HOURS** | | **56** |

## Suggested Theory distribution:

The suggested theory distribution as per Bloom's taxonomy is as per follows. This distribution serves as guidelines for teachers and students to achieve effective teaching-learning process.

| **Distribution of Theory for course delivery and evaluation** | | | | | |
|---|---|---|---|---|---|
| Remember | Understand | Apply | Analyse | Evaluate | Create |
| 15% | 20% | 30% | 30% | 5% | 0% |

## Instructional Method:

- The course delivery method will depend upon the requirement of content and need of students. The teacher in addition to conventional teaching method by black board, may also use any of tools such as demonstration, role play, Quiz, brainstorming, MOOCs etc.
- The internal evaluation will be done on the basis of continuous evaluation of students in the laboratory.
- Students will use supplementary resources such as online videos and e-courses.

**Reference Books:**

- Excel 2016 Bible, 1st Edition by bohnWalkenbach
- Excel Data Analysis For Dummies by Paul McFedries, John Wiley & Sons
- Data Analysis Using Microsoft Excel by Ash Narayan Sah, Excel Books India
- R Cookbook, Paul Teetor, Pub: Penram International.
- The Art of R Programming: A Tour of Statistical Software Design, Norman Matloff
- R for Data Science, Garrett Grolemund and Hadley Wickham
- Hands-On Programming with R: Write Your Own Functions and Simulations, Garrett Grolemund

**Supplementary Resources / Open Source Software:**

- https://www.coursera.org/learn/excel-data-analysis
- https://www.coursera.org/learn/analytics-excel
- https://www.udemy.com/r-basics/learn/v4/overview
- https://intellipaat.com/r-programming-certification-training/#course-content