

# Data Analytics with Python and R Programming Languages

## Module-1: Total Duration of Python Programming Course=24 hrs

**M-1.1:** Introduction to algorithm to solve a numerical problem. Downloading latest free Python package from <https://www.python.org/downloads/> under c-drive, Description of different components of Python, writing simple Python Code, basic syntax, editing, saving and running a script, data types, variables, assignments, arithmetic operators and expressions, writing comments, error messages, conditions, Boolean logic, logical operators, ranges, control statements: if-else, loops(for, while). Solving simple problems using Python language.  
**[3 hrs]**

**M-1.2:** Introduction to Lists, tuples and dictionaries, basic list operators, replacing, inserting, removing an element; searching and sorting lists; dictionary literals, adding and removing keys, accessing and replacing values, traversing dictionaries., Solving problems from assignment sheet. **[3 hrs]**

**M-1.3:** Introduction to String manipulations, subscript operator, indexing, slicing a string, string and number system, converting strings to numbers and vice versa, Solving problems .  
**[ 2hrs]**

**M-1.4:** Introduction to user defined functions, arguments and return values, formal vs actual arguments, named arguments, recursive call, applications., Solving problems from Assignment sheet.  
**[3 hrs]**

**M-1.5:** Introduction to Data Files, Strings and text files in Python,

manipulating files and directories, reading/writing text and numbers from/to file, creating and reading a formatted file(csv or tab-separated), Solving problems from assignment sheet. [ 3 hrs]

**M-1.6:** Introduction to Bit-wise operators : Bit wise or(|), Bit wise and (&), Bit wise XOR(^), Bit wise Complement(~), Bit wise Left Shift(<<), Bit wise right shift(>>), to convert number/character to bits, To convert data file to bits and vice versa, Solving problems from assignment sheet. [ 2hrs]

**M-1.7:** Introduction to Computer Graphics using Python, Using simple in built functions : GraphWin(), Point(), Circle(), Line(), Rectangle(), Oval(), setFill(), setWidth(), draw(), getMouse(). Drawing simple diagrams using in-built functions in Python. [2 hrs]

Free download Anaconda Software from <https://www.anaconda.com/products/distribution/start-coding-immediately>.

**M-1.8:** Introduction to numpy, importing numpy in Python code, converting lists to array, array manipulation, problem solving using Numpy. [2 hrs]

**M-1.9:** Introduction to Pandas, importing pandas in Python code, Application of Pandas for data manipulation, solving problems from assignment sheet. Accessing csv files from Kaggle. Introduction to Matplotlib, importing matplotlib in python code, reading standard database from Kaggle, Converting database to csv file, drawing different types of graphs using matplotlib, Solving problems from assignment sheet. [4 hrs]

Online Examination , Duration of Examination : 2 hrs(outside theory Class)

## **References:**

1. Fundamentals of Python : First Programs, Author: Kenneth A. Lambert  
Publisher: Course Technology, Cengage Learning , 2012, ISBN-13: 978-1-111-82270-5.
2. Python Notes for Professionals , e-book : goalkicker.com
3. Class notes will be supplied to all participants after every class through individual's email-id.

## **Module-2: “R Programming Language (Duration: 24 hrs)”**

### **M-2.1: Introduction to R :**

Introduction to R: What is R? – Why R? – Advantages of R over Other Programming Languages - R Studio: R command Prompt, R script file, comments – Handling Packages in R: Installing a R Package, Few commands to get started: installed.packages(), packageDescription(), help(), find.package(), library() - Input and Output – Entering Data from keyboard – Printing fewer digits or more digits – Special Values functions : NA, Inf and -inf. [ 1 hr]

### **M-2.2: R Data Types:**

R Data Types: Vectors, Lists, Matrices, Arrays, Factors, Data Frame – R - Variables: Variable assignment, Data types of Variable, Finding Variable ls(), Deleting Variables - R Operators: Arithmetic Operators, Relational Operators, Logical Operator, Assignment Operators, Miscellaneous Operators - R Decision Making: if statement, if – else statement, if – else if statement, switch statement – R Loops: repeat loop, while loop, for loop - Loop control statement: break statement, next statement. [5 hrs]

### **M-2.3: Functions in R-Language:**

R-Function : function definition, Built in functions: mean(), paste(), sum(), min(), max(), seq(), user-defined function, calling a function, calling a function without an argument, calling a function with argument values - R-Strings – Manipulating Text in Data: substr(), strsplit(), paste(), grep(), toupper(), tolower() - R Vectors – Sequence vector, rep function, vector access, vector names, vector math, vector recycling, vector element sorting - R List - Creating a List, List Tags and Values, Add/Delete Element to or from a List, Size of List, Merging Lists, Converting List to Vector - R Matrices – Accessing Elements of a Matrix, Matrix Computations: Addition, subtraction, Multiplication and Division- R Arrays: Naming Columns and Rows, Accessing Array Elements, Manipulating Array Elements, Calculation Across Array Elements - R Factors –creating factors, generating factor levels.

[6 hrs]

#### **M-2.4: String Manipulation in R language:**

String functions : grep(), nchar() , paste(), sprintf(), substring(), sub(),strsplit(), r

[ 2 hrs]

#### **M-2.5: Bit-wise operators using R:**

bitwOr(value1,value2), bitwXor(value1,value2), bitwNot(value),  
bitwAnd(value1,value2),bitwShiftL(value,shift), bitwShiftR(value,shift),  
Applications

[2 hrs]

#### **M-2.6: File Creation and visualization of Data:**

Data Frames –Create Data Frame, Data Frame Access, Understanding Data in Data Frames: dim(), nrow(), ncol(), str(), Summary(), names(), head(), tail(), edit() functions - Extract Data from Data Frame, Expand Data Frame: Add Column, Add Row - Joining columns and rows in a Data frame rbind() and cbind() – Merging Data frames merge() – Melting and Casting data melt(), cast(). Loading and handling Data in R: Getting and Setting the Working Directory – getwd(), setwd(), dir() File Handling in R language, -CSV Files - Input as a CSV file, Reading a CSV File, Analyzing the CSV File: summary(), min(), max(), range(), mean(), median(), apply() - Writing into a CSV File – R -Excel File – Reading the Excel file.

[3 hrs]

#### **M-2.7: Descriptive Statistics using R:**

Descriptive Statistics: Data Range, Frequencies, Mode, Mean and Median: Mean Applying Trim Option, Applying NA Option, Median - Mode - Standard Deviation – Correlation - Spotting Problems in Data with Visualization: visually Checking Distributions for a single Variable - R –Pie Charts: Pie Chart title and Colors – Slice Percentages and Chart Legend, 3D Pie Chart – R Histograms – Density Plot - R – Bar Charts: Bar Chart Labels, Title and Colors. Line Chart, Scatterplot, Developing graphs, Box Plot  
[ 3 hrs]

1. Drawing line, circle, rectangle, triangle using R language . [2 hrs]

**Exam : 2 hrs(It is outside theory class theory timing)**

**REFERENCES:**

1. “The Book of R” by Tilman M. Davies, no starch press(San Francisco)
2. “The Art of R programming” by Norman Matloff, no starch press(San Francisco)