Programming for Data Science

Data Science, also known as data-driven science, is an emerging field in the market. As per recent years' study, humungous amount of data is produced exponentially every year. Data Science is a process of using available data to understand and make decisions. To work on the available data, there are many tools such as Excel, SPSS, R, and so on. R is a General Public License (GNU) package, which is freely available and is the most emerging tool in present data science market. It is an open source programming language and software environment for statistical computation with graphics. The R language is widely used tool by statisticians and data miners for developing statistical software and data analysis.

In this course, you would learn to program in R and understand the data types and data structures supported in R. One would be able to demonstrate their ability to manipulate the R data objects and perform operations related to data science. This would be the foundation course for a seamless migration to statistical techniques and exploratory data analytics using R, which indeed are the foundation topics for machine learning using R.

Difficulty

Basic to Intermediate

Pedagogy

Online Learning + Live webinars

Duration

40 hours

Assessment

Topics Covered

- Data types and data structures
- Control structures
- Loops
- User defined functions
- File I/O operations

Learning Access

- Navigate the online content, study with video tutorials with examples
- Online reading material, assignments, and quizzes

Who Should Attend?

- Anyone interested in learning data science (keywords, R programming and massaging data)
- IT professionals and analysts interested in understanding about data science
- Business Intelligence (BI) professionals entering to data analytics projects
- Aspiring data scientists

Benefits to Learner

Programming for Data Science helps the learner to understand one of the most popular tool for data science (R). Upon completing this course, one would be able to seamlessly progress towards statistics and exploratory data analysis using R, which would lay the foundation for machine learning.

Learning Outcomes

1. Gain an insightful understanding of programming in R for data science
2. Understand data types, data structures and manipulate them using built-in functions in R
3. Ability to program using loops, control structures, and user-defined functions
4. Demonstrate the ability to perform file I/O operations

Course Offerings

Manipal Global Academy of Data Science (MGADS) is offering the PGP in Data Science through a comprehensive and modular set of learning content and interesting videos on a platform (EduNxt) that has been developed in-house.

Our objective is to create a talent pool of professionals skilled in the big data analytics or data science domain. We aim to train learners through an experiential model that incorporates several strategies including teaching through reading material, videos, practice questions, exercises, and illustrations explaining fundamental concepts.

Key Features

Students will get access to the following features:

Online faculty videos

Online instructor-led classes

Reading material

Assessments (quizzes)

Practice assignments in R

Course Curriculum

Unit No.	Unit Name	Topics Covered	Learning Objectives
1	Introduction to R programming	• Data types and data structures • Atomic types • Collection types • Operations on data structures	• Explain setting up of R-Studio • Describe the workspace • Explain the R data types and data structures • Manipulate the data structures using the built-in functions in R
2	Control structures	• Conditional statements (if-else blocks) • Switch statements • Loops	• Illustrate to program using if-else statement • Illustrate to program using nested conditional statements

			• Demonstrate the capability to use loops in R programming
3	Functions	• User-defined functions • Functions with default arguments • Functions with variable number of arguments • Functions returning values	• Demonstrate the programming ability to create user-defined functions
4	File I/O operations	• Reading data from a CSV file • Reading data from XML file • Saving the contents of data frame into a file	• Demonstrate the ability to load data into a data frame by reading from a file • Demonstrate the ability to save the contents of a data frame into files

Duration of the Course

Activity	Time Spent (hrs.)
Presentation materials	4
Online videos	3
Online quizzes/exercises	3
Online reading material	5
Instructor led webinars	10
Assignments / Hands-on exercises	15
Total	**40**