R Programming

Beginner Level

Duration: 1 Day (9am – 4pm)

Course Pre-Requisites: None

Course Overview

Dive into the dynamic world of R programming with our beginner's course designed to provide a comprehensive introduction to one of the most powerful statistical programming languages. Starting with the foundational elements of R and its premier integrated development environment, RStudio, participants will be equipped with the skills to perform basic mathematical operations, manage and manipulate diverse data structures, and visualise complex datasets. The course emphasises hands-on learning, ensuring attendees are well-versed in importing and exporting data, utilising functions effectively, and navigating potential package conflicts. By the end of this course, participants will have a solid grasp of the essentials of R programming, setting the stage for more advanced explorations in data analysis and visualisation.

1. Introduction

In the introduction to R, participants will be acquainted with the foundational elements of the R programming language. This module delves into the distinction between base R and the plethora of contributed packages available. It provides a step-by-step guide to downloading and installing the base R software, ensuring attendees are set up to commence their R journey. Additionally, guidance on installing essential R packages is provided.

2. The R Studio IDE

RStudio, the premier integrated development environment (IDE) for R, is thoroughly explored in this section. Attendees will be guided through the process of downloading and installing RStudio. An exhaustive overview of the RStudio environment is provided, emphasising its main panes, the significance of the working directory, the utility of the workspace, and the concept of projects, including creating one.

3. Using R as a Calculator

This section capitalises on R's robust computational capabilities. Participants will be trained on executing various commands both from the command line and the source pane. The fundamentals of arithmetic, relational, and logical operators in R are covered, ensuring attendees can perform basic mathematical operations.

4. Creating Objects

Delving deeper into R's object-oriented nature, this module emphasises the process of creating and managing objects. From understanding the different types of objects in R to the nuances of assignment operators and naming rules, participants will gain a thorough understanding. The basics of R syntax are covered, with hands-on demonstrations on creating and viewing objects both in RStudio and the console.

5. Data Types and Classes

A deep dive into R's data types and classes, this section offers insights into basic data types and the myriad data structures R supports, from vectors and factors to complex tibbles. Attendees will learn about implicit and explicit coercion methods, how to name different parts of data objects, and the art of accessing, replacing, adding to, and removing data from various R data structures.

6. Changing Data Types

In R, there are moments when data types might not align with the operations or analyses, we wish to perform. This section will delve into the mechanisms of Implicit Coercion where R automatically converts data types based on the context. However, such automatic conversions might not always yield the desired results. Hence, we also explore Explicit Coercion, where users manually specify the data type conversion, offering more control and precision.

7. Naming Parts of Data Objects

A well-structured dataset is as much about the data it holds as it is about the labels that define it. This module illuminates the process of naming various components of data objects. Whether it's Column Names for distinguishing different variables, Row Names for identifying individual observations, or Dimension Names for multi-dimensional arrays, clear naming conventions enhance data clarity. The section also touches upon the importance and methods of general naming for various objects in R, encapsulated under Names.

8. Accessing Data within Data Structures

Data stored in R structures are of no use unless we can efficiently access them. This section introduces participants to the art of data retrieval. Whether it's Referring to Data by Position, which is crucial for indexed data structures like matrices, or Referring to Data by Name,

which is vital for data frames with column names, this module ensures attendees can swiftly and efficiently navigate through their data.

9. Replace Parts of an Object

As datasets evolve, there's often a need to modify them. This module equips participants with skills to Replace Names of data structures and Replace Values within them. Such techniques are essential for data cleaning and transformation, ensuring datasets remain relevant and accurate.

10. Add to a Data Object

Data accumulation is a constant process. Whether you're appending new observations or adding new variables, this section has you covered. Learn to Add Elements to Vectors for simple data structures or Add Rows or Columns to more complex ones like matrices and data frames. Whether you're adding by Position or by Name, this module ensures you can grow your datasets seamlessly.

11. Removing Data from a Data Object

Not all data remains relevant, and there's often a need to prune unnecessary or redundant information. This section delves into techniques to Remove Elements from Vectors, ensuring simple data structures remain concise. For more complex structures, learn to Remove Rows or Columns from Matrices and Data Frames, ensuring your datasets are always streamlined and relevant.

12. Evaluation in R

This module delves into R's evaluation mechanisms. Participants will explore vector arithmetic, understand the order of operations in R, and learn about vector recycling. The session emphasises the importance of vectorised operations and introduces methods to apply functions to elements within data structures.

13. Using Functions

Functions are the building blocks of R programming. In this section, participants will learn what constitutes a function in R, the correct syntax for using them, and the different arguments they can accept. Essential to this module is the guidance on seeking help with functions and understanding the help documentation in R. The section also touches upon the basics of statistical summary functions and addresses potential conflicts when multiple packages offer functions with the same name.

14. Importing Data

Effective data analysis begins with efficient data import. This module emphasises the tools and techniques to import data into RStudio. A significant focus is given to importing data from text files, specifically CSVs, ensuring participants can seamlessly integrate data into their R environment.

15. Exporting Data

Data exportation is equally crucial as its import counterpart. In this section, participants will learn the best practices for exporting data to text files, ensuring their analyses can be shared and utilised outside the R environment.

16. Basic Data Visualisation

Visualisation aids in comprehending complex data narratives. This module introduces participants to R's plot function, a versatile tool for data visualisation. Attendees will learn how to enhance their plots with reference lines, text, and legends. The module concludes with instructions on exporting these visual representations for external use.