**R Programming**

**Course Overview**

The **R Programming** course is designed for individuals who want to explore **data analysis, statistics, and visualization using R**, a language widely used by statisticians, data analysts, and researchers. This beginner-friendly yet powerful course introduces learners to R syntax, data structures, statistical computing, and graphing techniques.

The course is especially beneficial for students, professionals, and data enthusiasts looking to enter fields like data science, bioinformatics, econometrics, academic research, and machine learning. With its easy syntax and extensive libraries, R is the go-to tool for turning raw data into meaningful insights.

**Course Objectives**

By the end of this course, learners will:

* Understand the fundamentals of R programming language.
* Work efficiently with RStudio and R environment.
* Learn data structures such as vectors, matrices, lists, and data frames.
* Perform data cleaning, manipulation, and transformation using built-in and external packages.
* Apply statistical techniques and generate descriptive summaries.
* Visualize data using base R and ggplot2.
* Build basic models and interpret statistical results.
* Handle real-world datasets for analysis and reporting.

**Syllabus Breakdown**

**Module 1: Introduction to R and RStudio**

* What is R and why use it?
* Installing R and RStudio
* Understanding the RStudio interface
* Writing and executing R scripts
* Basic syntax: variables, comments, operators

**Module 2: R Data Types and Structures**

* Data types: numeric, integer, character, logical, factor
* Vectors and vectorized operations
* Lists and their manipulation
* Matrices: creation, indexing, and operations
* Arrays and multi-dimensional data
* Data frames: structure, indexing, subsetting

**Module 3: Control Structures and Functions**

* Conditional statements: if, else, ifelse()
* Loops: for, while, repeat
* Writing custom functions
* Using built-in functions
* Scope and environments

**Module 4: Data Import, Export, and Cleaning**

* Reading data: read.csv(), read.table(), readxl, and data.table
* Writing data to files
* Data exploration: str(), summary(), head(), tail()
* Handling missing values and NA
* Data transformation and cleaning techniques

**Module 5: Data Manipulation with dplyr and tidyr**

* Selecting, filtering, and mutating data
* Grouping and summarizing
* Joins and combining datasets
* Data reshaping: gather(), spread(), pivot\_longer(), pivot\_wider()

**Module 6: Data Visualization**

* Base R plotting system
* Customizing plots: titles, legends, colors
* Introduction to ggplot2
* Creating bar charts, histograms, line graphs, scatter plots, boxplots
* Combining plots using patchwork or cowplot

**Module 7: Basic Statistics and Analysis**

* Descriptive statistics: mean, median, mode, variance, standard deviation
* Probability distributions
* Hypothesis testing: t-tests, chi-square, ANOVA
* Correlation and regression analysis
* Interpreting statistical results

**Module 8: Working with Real Datasets**

* Importing datasets from the internet
* Exploratory Data Analysis (EDA) project
* Case Study: COVID-19 Dataset or Iris Dataset
* Creating summary reports and visual dashboards

**Career Opportunities**

R programming is in high demand in analytics-driven roles and industries. After completing this course, learners can pursue careers in:

* **Data Analyst**
* **Statistician**
* **Business Intelligence Analyst**
* **Market Research Analyst**
* **Data Scientist (with further study)**
* **Research Assistant in Universities/Think Tanks**
* **Bioinformatics Specialist**
* **Healthcare Data Analyst**

R is particularly popular in sectors like healthcare, academia, research organizations, public policy, and market research.