Faysal Aziz

Beginning R Programming

Data preparation, visualisation and introductory statistical analysis

FA-0776 · 2 days

Data & Analytics · Software Development

Introduction

Begin programming in R for statistics and data analysis. Work through guided examples of data structures, functions, data preparation and visualisation, then interpret a simple linear regression. The two-day course introduces these techniques rather than providing advanced statistical training.

Learning outcomes

- Use R objects, indexing, operators, conditions, loops and user-defined functions.
- Import, export, clean, transform and tidy data frames.
- Explore data using ggplot2 charts and summary statistics.
- Fit a guided linear regression and interpret its output and R-squared.
- Use external R packages in the course environment.

Prerequisites

No prior R programming knowledge is required. Use a compatible R and RStudio environment, a browser and access to the course data sources. A computer with approximately 8 GB RAM and sufficient storage for the prepared datasets is suitable; a second screen is optional.

Training outline

1. Day 1 — Environment, objects and programming

- Environment
- Workspace
- Data Source

Programming fundamentals

- Data types
- Operations
- Coercion
- Functions

Composite objects

- Vectors
- Vector operations
- Recycling
- Dimensions
- Matrices
- Indexing
- Slicing
- Arithmetic and operations
- Lists

Check vector lengths and understand recycling rather than assuming all recycled operations are appropriate.

Control flow and functions

- Operators
- Conditions
- Loops
- UDFs

2. Day 2 — Data preparation and visualisation

- Data Frames
- Import from data source
- Export to data destination
- Functions and operations with Data Frames
- Data cleaning
- Transformations
- Sampling
- Tidying
- Use ggplot2 to create histograms, box-and-whisker plots, bar charts and scatter plots.

3. Day 2 — Exploratory statistics and regression

Exploratory analytics

- Population vs. Sample
- Mean, median and mode
- Skewness
- Variance, standard deviation and coefficient of variation
- Covariance and correlation
- The linear regression model
- Correlation vs regression
- Geometrical representation
- Practical regression in R
- How to interpret the regression table
- Decomposition of variability: SST, SSR, SSE
- R-squared

- Assessment

Statistical and regression topics use simple guided examples; interpretation includes the limits of correlation and model fit.

© 2026 Faysal Aziz · Kuala Lumpur