

Data Analytics with R Training

COURSE CONTENT

GET IN TOUCH



Multisoft Systems
B - 125, Sector - 2, Noida



(+91) 9810-306-956



info@multisoftsystems.com



www.multisoftsystems.com

About Multisoft

Train yourself with the best and develop valuable in-demand skills with Multisoft Systems. A leading certification training provider, Multisoft collaborates with top technologies to bring world-class one-on-one and certification trainings. With the goal to empower professionals and business across the globe, we offer more than 1500 training courses, which are delivered by Multisoft's global subject matter experts. We offer tailored corporate training; project Based Training, comprehensive learning solution with lifetime e-learning access, after training support and globally recognized training certificates.

About Course

The Data Analytics with R Training offered by Multisoft Systems is an intensive course designed to equip participants with the essential skills needed to excel in the field of data science. This comprehensive program covers everything from basic data manipulation to advanced statistical analysis techniques using the R programming language.

Module 1: Introduction to Data Analytics

- ✓ Introduction to terms Business Intelligence, Business Analytics, Data, Information
- ✓ How information hierarchy can be improved/introduced
- ✓ Understanding Business Analytics and R
- ✓ Knowledge about the R language, its community, and its ecosystem
- ✓ Understand the use of 'R' in the industry
- ✓ Compare R with other software in analytics
- ✓ Install R and the packages useful for the course
- ✓ Perform basic operations in R using the command line
- ✓ Learn the use of IDE R Studio and Various GUI
- ✓ Use the 'R help' feature in R
- ✓ Knowledge about the worldwide R community collaboration

Module 2: Introduction to R Programming

- ✓ Various kinds of data types in R and their appropriate uses
- ✓ Built-in functions in R like seq(), cbind (), rbind(), merge()
- ✓ Knowledge of the various subsetting methods
- ✓ Summarize data by using functions like: str(), class(), length(), nrow(), ncol()
- ✓ Use of functions like head(), tail() for inspecting data
- ✓ Indulge in a class activity to summarize data
- ✓ dplyr package to perform SQL join in R

Module 3: Data Manipulation in R

- ✓ Steps involved in Data Cleaning
- ✓ Functions used in Data Inspection
- ✓ Tackling the problems faced during Data Cleaning
- ✓ Uses of the functions like grepl(), grep(), sub(), Coerce the data, uses of the apply() functions

Module 4: Data Import Techniques in R

- ✓ Import data from spreadsheets and text files into R
- ✓ Import data from other statistical formats like sas7bdat and spss, packages
- ✓ Installation used for database import
- ✓ Connect to RDBMS from R using ODBC
- ✓ Basic SQL queries in R
- ✓ Basics of Web Scraping

Module 5: Exploratory Data Analysis

- ✓ Exploratory Data Analysis(EDA)
- ✓ Implementation of EDA on various datasets
- ✓ Boxplots
- ✓ Understanding the cor() in R
- ✓ EDA functions like summarize(), llist()
- ✓ Multiple packages in R for data analysis
- ✓ Fancy plots like the Segment plot, and HC plot in R

Module 6: Data Visualization in R

- ✓ Data Visualization
- ✓ Graphical functions present in R
- ✓ Plot various graphs like tableplot, histogram, and Boxplot
- ✓ Customizing Graphical Parameters to improvise plots
- ✓ Understanding GUIs like Deducer and R Commander
- ✓ Introduction to Spatial Analysis

Module 7: Data Mining: Clustering Techniques

- ✓ Introduction to Data Mining
- ✓ Understanding Machine Learning
- ✓ Supervised and Unsupervised Machine Learning Algorithms

- ✓ K-means Clustering

Module 8: Data Mining: Association Rule Mining & Collaborative filtering

- ✓ Association Rule Mining
- ✓ User Based Collaborative Filtering (UBCF)
- ✓ Item Based Collaborative Filtering (IBCF)

Module 9: Linear and Logistic Regression

- ✓ Linear Regression
- ✓ Logistic Regression

Module 10: Anova and Sentiment Analysis

- ✓ Analysis of Variance (Anova) Technique
- ✓ Sentiment Analysis: fetch, extract and mine live data from Twitter

Module 11: Data Mining: Decision Trees and Random Forest

- ✓ Decision Tree
- ✓ Entropy
- ✓ Gini Index
- ✓ Pruning and Information Gain
- ✓ Algorithm for creating Decision Trees
- ✓ Bagging of Regression and Classification Trees
- ✓ Random Forest
- ✓ Working on Random Forest
- ✓ Features of Random Forest, among others

Module 12: Project Work

- ✓ Analyze census data to predict insights on the income of the people based on the factors like age, education, work class, and occupation using Decision Trees
- ✓ Logistic Regression and Random Forest
- ✓ Analyze the Sentiment of Twitter data, where the data to be analyzed is streamed live from Twitter, and sentiment analysis is performed on the same