



About Big Data Analytics with R

The course on R gives students a systematic overview of how to solve programming problems using the R language. It exposes students to basic concepts in R, followed by the use of various libraries in R. In addition, this course covers basic concepts in statistics using R Commander, as well as implementing algorithms for data visualization in R. In addition, students are exposed to advanced topics like time series analysis.

Course Objectives

- To get a basic understanding of R and the various ways to create scripts and programs in R
- To understand some of the key constructs in R for data handling
- To understand how to use R in interacting with the external world
- To understand and appreciate how to summarize large volumes of data effectively by appropriate use of charts of different types
- Understand how to use R for real-life applications, such as regression and forecasting

Curriculum

Module 1: Introduction to Data Science, Machine Learning

Module 2: Introduction to R, datatypes in R, Importing Exporting Data in R

Module 3: Using R for statistics, R Apply functions

Module 4: Data Visualization in R

Module 5: Unsupervised Learning with R

Module 6: Linear, Logistic Regression with R

Module 7: Decision Tree, Random Forest in R

Module 8: Project Discussion, Working on Case study

Module 9: Time-Series Analysis in R

Module 10: Text/Sentiment Analysis in R



Instructor

Instructors are handpicked from a selected group of industry experts and mentors and trained to deliver the best online learning experience. All training instructors have at least ten years of industry experience and extensive functional expertise in the field they train.

Certification

- The assessment will be done on the basis of an online test and Project Evaluation at the end of the course.
- On completion of the Assessment (Project + Test) with a minimum of 70% marks, the candidate will be issued a certificate of successful completion from NIIT.

Pre-requisites

Learner should have basic knowledge of statistics and computer programming. Preferably a reasonable level of proficiency in any data handling tool like MS Excel

FAQs

What is the Case Study based Project which I need to develop as part of this program?

Project Specifications

Background: A finance company deals in all home loans. They have presence across all urban, semi urban and rural areas. Customer first applies for home loan after that company validates the customer eligibility for loan.

Problem: Company wants to automate the loan eligibility process (real time) based on customer detail provided while filling online application form. These details are Gender, Marital Status, Education, Number of Dependents, Income, Loan Amount, Credit History and others. To automate this process, they have given a problem to identify the customers segments; those are eligible for loan amount so that they can specifically target these customers. You are required to use your “R” skills to deliver a model for predicting loan eligibility of a customer. In other words, you will tell the company by your solution that whether customer applied for loan will default or not in future!