

## BASIC MODULE

### R Basics

16.5 Hours

#### Introduction to R

R-Studio, Packages in R, Installing Packages, Setting Directories

#### Basic Operations in R

Programming Language Basics, Scalars, Vectors, Simple Calculations Data Structure, Data Frames

#### Data Manipulation

Data Acquisition, Sub-setting Observations & Variables, Transforming Variables, Conditional Processing, Merging & Concatenating Datasets, Using SQL in R

### Statistics Using R

28.5 Hours

#### Basic Statistics in R

Descriptive Statistics, Cross-Tabs, T-Tests, ANOVA, n-way ANOVA, Correlations, Sampling, Bootstrapping, Random Number Generation

#### Graphics in R

Line Plots, Bar Charts, Pie Charts, Histograms, Scatter Plots, 3-D, Parallel Coordinates

#### Linear Models in R

Simple Regression Analysis, Multiple Regression

#### Logistics Regression in R

Logic Function, Maximum Likelihood Optimization, Odds Ratio

#### Model Selection in R

Selection Criteria, Stepwise Methods

Writing Functions in R: Creating User Defined Functions

## ADVANCED MODULE

### Predictive Analytics

10.5 Hours

#### Review of Basic Statistics in R

Basic Statistics (Mean, Median, S.D.), ANOVA (CRD, RBD), Correlations, Outliers, Statistical Graphs

#### Predictive Modelling using Regression Analysis

Linear Regression: Modelling Assumptions

#### Predictive Modelling using Regression Analysis

Logistic Regression: Modelling Assumptions, Model Selection, Scoring

#### Predictive Modelling using Decision Trees Analysis & Random Forest

Recursive Partitioning, CART, Splitting Criterion, Pruning

#### Other Predictive Modelling Tools

Support Vector Machine, K-Nearest Neighbor Algorithm

### Text Analytics

8 Hours

#### Basics of Text Mining

Preparing and Loading a Corpus, Performing Semming, Creating & Exploring Term-Document matrix, N-Gram Analysis

#### Visualizations in Text Mining

Creating and modifying a WordCloud, Reshaping, Resizing and Recoloring, Word Association Analysis

#### Advances in Text Analytics

Text Clustering & Categorization Sentiment Analysis

### Multivariate Analysis

12 Hours

#### Clustering

Hierarchical Clustering, K-means Clustering

#### Variable Reduction

Factor Analysis, Principal Component Analysis

#### Association Mining

Market Basket Analysis

### Optimization/ Forecasting

16.5 Hours

#### Linear Programming & Goal programming in R

Objective Functions, Constraints and Solving LP & GP Problems

#### Integer Programming & Mixed Integer Linear Programming

Objective Functions, Constraints and Solving ILP and MILP Problems

#### Business Forecasting in R

Basic Time Series, Exponential Smoothing Models, ARIMA Modelling in R.

### Experiential Learning and Training

#### Case-study

Real life Industry case studies

#### Career Assistance

Placement Workshops  
Mock Interviews  
Personal Development

#### Industry Mentorship

Evaluation of case studies & project work.

## CASE STUDIES



### TITANIC

An interesting project where students are given the task to predict the survival of passengers on the Titanic based on various factors such as gender, age, passenger class and port of embarkation.



### COLONIAL BROADCASTING

A major American television network, must determine whether fact-based or fiction-based television movies garner higher ratings. Furthermore, the network must decide whether to accept a fixed fee or a sliding-scale contract.



### CLASS EXAMINATION

Here we look at a class with 22 students and how they performed in a given exam. Using analytical tools, you need to identify how the overall class performed and provide a report on the performance of every student in this examination.