

## Data Science with R Course Content

### Course Description:

Data Science using R gives the best statistical computing and design. Data science with R offers data exploration, visualization, predictive, and descriptive analytics techniques with R language. R is the most flexible and powerful analytic tool to develop statistical software and data analysis. R is trending now and has immense demand in the future. R is open-source, good data handling capability, graphical capabilities, advancements, and cost-efficient as compared to SAS statistical tool. R is also very simple and easy to learn.

Hachion Data Science with R online training provided by industry experts from end to end. Data Science using R course tutorial includes data science tools and techniques including cleaning, exploring, visualizing data using the R language. Through our Data science R language online course, you will learn about packages, structures, statistical concepts, cluster analysis, import/export data, and forecasting. By enhancing theoretical and practical knowledge one can easily appear for the Data Science certification exam.

### Course Content:

#### Introduction to R

- What is R?
- Why R?
- Installing R
- R environment
- How to get help in R
- R Studio Overview

#### Understanding R data structure

- Variables in R
- Scalars
- Vectors
- Matrices
- List
- Data frames
- Cbind, Rbind, attach and detach functions in R
- Factors
- Getting a subset of Data
- Missing values
- Converting between vector types

#### Importing data

- Reading Tabular Data files
- Reading CSV files

- Importing data from excel
- Loading and storing data with clipboard
- Accessing database
- Saving in R data
- Loading R data objects
- Writing data to file
- Writing text and output from analyses to file

## **Manipulating Data**

- Selecting rows/observations
- Rounding Number
- Creating string from variable
- Search and Replace a string or Number
- Selecting columns/fields
- Merging data
- Relabeling the column names
- Data sorting
- Data aggregation
- Finding and removing duplicate records

## **Using functions in R**

- Apply Function Family
- Commonly used Mathematical Functions
- Commonly used Summary Functions
- Commonly used String Functions
- User defined functions
- Local and global variable
- Working with dates

## **Charts and Plots**

- Box plot
- Histogram
- Pie graph
- Line chart
- Scatterplot
- Developing graphs
- Cover all the current trending packages for Graphs

## **Machine Learning Algorithm**

- Sentiment analysis with Machine learning
- C 5.0
- Support vector Machines

- K Means
- Random Forest
- Naïve Bayes algorithm

## Statistics

- Correlation
- Linear Regression
- Non-Linear Regression
- Predictive time series forecasting
- K means clustering
- P value
- Find outlier
- Neural Network
- Error Measure

## Leading Topics

- Overture of R Shiny
- What is Hadoop
- Integration of Hadoop in R
- Data Mining using R
- Clinical research preface in R
- API in R (Twitter and Facebook)
- Word Cloud in R