

- COURSE :Data Science with R
- DURATION : 2 MONTH
- FEES : 13500/-
- GOVERNMENT APPROVED CERTIFICATE

DATA SCIENCE WITH R

Module 1- Introduction to Data Analytics (Duration: 04:00:00 hrs)

Objectives

- This module introduces you to some of the important keywords in R like Business Intelligence, Business Analytics, Data and Information.
- You can also learn how R can play an important role in solving complex analytical problems.
- This module tells you what is R and how it is used by the giants like Google, Facebook, etc.
- Also, you will learn use of 'R' in the industry, this module also helps you compare R with other software in analytics, install R and its packages.

Topics

- Business Analytics, Data, Information
- Understanding Business Analytics and R
- Compare R with other software in analytics
- Install R
- Perform basic operations in R using command line
- Learn the use of IDE R Studio
- Use the 'R help' feature in R

Module 2- Introduction to R programming (Duration : 03:00:00 hrs)

Objectives

- This module starts from the basics of R programming like datatypes and functions.
- In this module, we present a scenario and let you think about the options to resolve it, such as which datatype should one to store the variable or which R function that can help you in this scenario.
- You will also learn how to apply the 'join' function in SQL.

Topics

- Variables in R
- Scalars
- Vectors

SHAH INSTITUTE OF COMPUTER SCIENCE

- Matrices
- List
- Data frames
- Using c, Cbind, Rbind, attach and detach functions in R
- Factors

Module 3- Data Manipulation in R (Duration : 04:00:00 hrs)

Objectives

- In this module, we start with a sample of a dirty data set and perform Data Cleaning on it, resulting in a data set, which is ready for any analysis.
- Thus using and exploring the popular functions required to clean data in R.

Topics

- Data sorting
- Find and remove duplicates record
- Cleaning data
- Recoding data
- Merging data
- Slicing of Data
- Merging Data
- Apply functions

Module 4- Data Import techniques in R (Duration : 04:00:00 hrs)

Objectives:

- This module tells you about the versatility and robustness of R which can take-up data in a variety of formats, be it from a csv file to the data scraped from a website.
- This module teaches you various data importing techniques in R.

Topics

- Reading Data
- Writing Data
- Basic SQL queries in R
- Web Scraping

Module 5- Exploratory Data Analysis (Duration : 04:00:00 hrs)

Objectives

- In this module, you will learn that exploratory data analysis is an important step in the analysis.
- EDA is for seeing what the data can tell us beyond the formal modeling or hypothesis. You will also learn about the various tasks involved in a typical EDA process.

Topics

- Box plot
- Histogram
- Pareto charts
- Pie graph
- Line chart
- Scatterplot
- Developing Graphs

Module 6- Basics of Statistics & Linear & Logistic Regression (Duration : 05:00:00 hrs)

Objectives

- This module touches the base of Descriptive and Inferential Statistics and Probabilities & 'Regression Techniques'.
- Linear and logistic regression is explained from the basics with the examples and it is implemented in R using two case studies dedicated to each type of Regression discussed.

Topics

- Basics of Statistics
- Inferential statistics
- Probability
- Hypothesis
- Standard deviation
- Outliers
- Correlation
- Linear & Logistic Regression

Module 7- Data Mining: Clustering techniques, Regression & Classification (Duration : 04:00:00 hrs)

Objectives

- Linear and logistic regression is explained from the basics with the examples and it is implemented in R using two case studies dedicated to each type of Regression discussed.
- The two Machine Learning types are Supervised Learning and Unsupervised Learning and the difference between the two types.
- We will also discuss the process involved in 'K-means Clustering', the various statistical measures you need to know to implement it in this module.

Topics

- Introduction to Data Mining
- Understanding Machine Learning
- Supervised and Unsupervised Machine Learning Algorithms
- K- means clustering

Module 8- Project work (Duration : 08:00:00 hrs)

- 2 Real-time projects