

# CERTIFICATION IN DATA SCIENCE WITH R



## Course Code : OCIT0011

R is a programming language and software environment for statistical analysis, graphics representation and reporting. R is a programming language and environment commonly used in statistical computing, data analytics and scientific research. R allows us to integrate with other languages (C, C++). Thus, you can easily interact with many data sources and statistical packages. As a result, R programming language has large growing community of users.

## Curriculum

### Module 1: Advance Time Series Modeling

- Box - Jenkins Methodology, Introduction to Auto Regression and Moving Averages, ACF, PACF, Detecting order of ARIMA processes, Seasonal ARIMA Models (P,D,Q)(p,d,q), Introduction to Multivariate Time series Analysis, Using built in R datasets

### Module 2: Stock market prediction

- Live example/ live project, Using client given stock prices/taking stock price data

### Module 3: Pharmaceuticals

- Case Study with the Data, Based on open set data

### Module 4: Market Research

- Case Study with the Data, Based on open set data

### Module 5: Machine Learning

- Supervised Learning Techniques, Conceptual Overview

### Module 6: Machine Learning

- Unsupervised Learning Techniques, Association Rule Mining Segmentation, Conceptual Overview

### Module 7: Fraud Analytics

- Fraud Identification Process in Parts procuring, Sample data from online

### Module 8: Text Analytics

- Text Analytics, Sample text from online

### Module 9: Social Media Analytics

- Social Media Analytics, Sample text from online

## Learning Outcomes

- Understand the core programming concepts of R language.
- Know the Looping and condition statements in R Programming
- Understand the different options in I/O operations in R programming.
- Understand the importance of statistical functions in R programming language
- Understand the basic concepts of integrating other object oriented programming with R.