Analyzing Big Data with Microsoft R

Modality: On Demand

Duration: 16 Hours

About this Course:

This intermediate-level training program is specifically designed for IT Professionals and Developers working within Big Data Environment. This course presents a great learning opportunity for students and candidates striving to learn the fundamentals of working with Big Data. Microsoft R Programming Language and its functionalities are comprehensively elaborated in this course and professionals get practical experience of analyzing and working with big datasets.

This course majorly focuses on the core features and functionalities of RevoScaleR Package which is the most essential component of Microsoft R Server. Professionals get to develop familiarity with these functionalities and apply these features to process and analyze large datasets seamlessly. Professionals also get to learn the art of deploying Microsoft R Server Functionalities on SQL Server Database and Spark Cluster. In this way, candidates will get to learn the basics of big data problem solving with the help of Microsoft R Server (MRS).

Course Objective:

The core objective of this course is to help professionals gain a better knowledge and understanding of the following key concepts:

- Data Interpretation from Flat Files into Microsoft R Data Frame
- Dataset Structure Investigation and Resolutions
- Prepared Datasets Storage and Applications
- Data Transformation and Preparation
- Essential Summary Statistics Calculation and Crosstabulation’
- Creating Summary Functions and ggplot2 Data Visualization
- Data Deployment and Scaling
- Predictive Model Development and New Data Prediction Generation
- Model Comparison and Evaluation

Audience:

This course is specifically tailored for Data Scientists and Data Analysts. Professionals liable for R Analysis and Solution Integration can also greatly benefit from the teachings of this course.

Prerequisites:

Professionals planning to enroll in the Analyzing Big Data with Microsoft R Course must comply with the following prerequisites:

- Familiarity with the Core Features and Functionalities of Microsoft R.

- Fundamental Knowledge of Microsoft R Data Structures
- Working Experience with Third-Party Packages is also highly Recommended

Course Outline:

1. Introduction

- Introduction

2.Reading and Preparing Data

- Reading the Data
- Preparing the Data
- LAB

3. Examining and Visualizing Data

- Examining the Data
- Visualizing the Data
- LAB

4. Clustering and Modeling

- Clustering
- Predictive Modelling
- LAB

5. Deploying and Scaling

- Deploying and Scaling

Final Exam and Wrap-up

- Final Exam
- Exam Wrap-up!