

Course Title: Data Science Starter

Duration: 9 hours

Module – 1: Big data analytics overview

- Data and its changing nature
- Big data overview
- Hadoop Overview
- NoSQL overview
- Data Engineer v/s Data Scientist
- Data science tools
- Data science on Hadoop

Module – 2: Overview of R

- R Graphical User Interface
- Importing and Exporting Data
- Attributes and Data Types
- Descriptive Statistics

Module – 3: Exploratory Data Analysis

- Visualization Before Analysis
- Dirty Data
- Visualizing a Single Variable
- Examining Multiple Variables
- Data Exploration vs. Presentation

Module – 4: Statistical Methods for Evaluation

- Hypothesis Testing
- Difference of Means
- Wilcoxon Rank-Sum Test
- Type I and Type II Errors
- Power and Sample Size
- ANOVA

Module – 5: Case Studies

- Customer Propensity Models
- Customer Churn Models
- Customer Segmentation Models
- Predictive maintenance models

- Recommender models
- Natural language Processing