

Data Analysis using R Programming

What is Big Data & Why Hadoop?

- Big Data Characteristics, Challenges with traditional system

Introduction to R

- History of R
- An Insight into R
- Data Structure and Data Type

Data Management and Data Cleaning

- Missing Value Treatment
- Outlier Treatment
- Sorting Datasets
- Merging Datasets
- Creating new variables
- Binning variables
- Reading datasets from other environments into R (importing)
- Writing datasets from R environment to other environments (exporting)

Data Visualization in R

- Bar Chart
- Dot Plot
- Scatter Plot (3D)
- Spinning Scatter Plots
- Pie Chart
- Histogram (3D) [including colourful ones]
- Overlapping Histograms
- Boxplot
- Plotting with Base and Lattice Graphics
- Plotting and Colouring
- Geo Charts
- Motion Charts

Using functions in R

- Sapply function
- Apply function
- Tapply function

Basic statistics in R

- Summarizing data
- Measures of Central Tendency
- Measures of Variability and Distributions
- Hypothesis Testing
- Performing Anova (One – Way and Two – Way)