

## **Introduction to R**

- Installing R and R Studio
- What is CRAN?
- Getting Started
- Data Structure in R
- Input and output
- Working with various data structures
- Data Manipulations and exploration
- Control Structures in Programming
- Loop functions - apply/lapply/split etc
- Creating user defined functions
- Generating Random numbers
- Data Splitting: training and test
- Data Summary and dealing with missing data
- Data Connection different sources
- Working with Data Frame dplyr

## **Exploratory Data Analysis (EDA) with Excel, R, SAS**

- Understanding various statistics of data
- Dealing with graphs (continuous & categorical data)
- Working with ggplot2 package in R
- Factor Analysis
- Various methods/concepts related to factors
- Dimension Reduction

## **Basic Statistics Probability with Excel, R, SAS**

- Conditional Probability
- Expected value
- Variance and Standard Error
- Binomial, Normal and Poisson distribution
- Central Limit Theorem
- Hypothesis Testing
- P-Value, T-Value, Z-score

## **Regression Modeling with R, SAS**

- What is regression?
- Model fitting
- Diagnostics plots
- Model comparison
- Cross validation
- Variable selection
- Outlier detection
- Influential Observations
- Heteroscedasticity
- Non normality
- Variable transformation
- Skewness, Kurtosis
- ANOVA/MANOVA
- Least square method
- R-Square and Adj R-Square
- Interpreting Coefficients
- Prediction using linear model
- Multivariate regression
- Multicollinearity
- Model selection

## **Logistic Regression with R, SAS**

- What is Logistic Regression?
- When is this needed?
- Null Deviance and residual deviance
- Model performance: confusion matrix
- ROC curve -AUC
- Prediction with Logit model
- Multinomial logistic model
- Working with count data: Poisson regression

## **Time series modeling with R, SAS**

- Model of Time Series
- ARIMA model
- MA model
- Understanding ACF plots
- Seasonality and trend in data
- Handling seasonality and trends in model
- Forecasting using model

## **Text Analytics with R, SAS**

- Read the Text Data in R
- User Defined Functions
- Text Preprocessing - Cleaning
- Create Document Term Matrix
- Basic Text Analysis
  - Using Term frequency (TF)
  - Using Term Frequency Inverse Document Frequency(TF-IDF)
  - Sentiment Analysis
- Create Positive Words word cloud
- Create Negative Words word cloud
- Positive words v/s Negative Words plot
- Twitter API
- Extract Reviews from Website
- Topic Analysis
- Overview of NLP
- Case Study –Sentiment Analysis for IPL Teams

## **Data Mining with Excel, R, SAS**

- Basic of machine learning
- Supervised/Unsupervised learning
- Conjoint analysis
- CHAID Analysis
- Factor analysis
- Cluster analysis
- Clustering Methods
- K-Means clustering algorithm
- Hierarchical Clustering
- Association Rules (Market basket Analysis)
- Overview of KNN/Neural Net/SVM/Ensemble
- Overview of Reinforcement Learning