

## DATA SCIENCE with R-PROGRAMMING & PYTHON

### Course contents

#### R - Programming

- ❖ Course Introduction
- ❖ How to get help in the course
- ❖ How to install the Software
- ❖ Development Environment Overview
- ❖ Introduction to R Basics
- ❖ Arithmetic in R
- ❖ Variables
- ❖ R Basic Data Types
- ❖ Vectors
  - Vector Basics
  - Vector Operations
  - Vector Indexing and Slicing
  - Vector Exercise
- ❖ Comparison Operators
- ❖ Matrices
  - Introduction to R matrices
  - Creating a Matrix
  - Matrix Arithmetic
  - Matrix Operations
  - Matrix Selection and Indexing
  - Factor and Categorical Matrices
  - Matrix Exercise
- ❖ Data Frames
  - Introduction to Data Frames
  - Data Frame Basics
  - Data Frame Indexing and Selection
  - Data Frame Operations
  - Data Frame Training Exercise
- ❖ Lists
- ❖ Data Input and Output with R
  - CSV Files
  - Excel Files
  - SQL with R
  - Web Scraping with R
- ❖ R Programming Basics
  - Logical Operators
  - If, else, and else if Statements
  - While Loops
  - For Loops
  - Functions
- ❖ Advanced R Programming

- Built-in R Features
- Apply Functions
- Math Functions with R
- Regular Expressions
- Dates and Timestamps
- ❖ Data Manipulation with R
  - Dplyr Functions
  - Pipe Operator
- ❖ Charting in R
- ❖ Data Visualization With R
  - Overview of ggplot2
  - Histograms
  - Scatterplots
  - Barplots
  - Boxplots
  - 2 Variable Plotting
  - Coordinates and Faceting
  - Themes
- ❖ Interactive Visualizations with Plotly
  - Overview of Plotly and Interactive Visualizations
  - Resources for Plotly and ggplot2
- ❖ Capstone Data Project

#### PYTHON

##### Introduction

- ❖ Python and its uses
- ❖ Installing Python and PyCharm
- ❖ Hello world Program in Python
- ❖ Some mathematical operations in Python
- ❖ Strings in Python
- ❖ Accepting Input from the user in Python
- ❖ Performing operations on a string in Python
- ❖ Variables in Python
- ❖ In place operators in Python
- ❖ Writing the first program in PyCharm

##### Control structures in Python

- ❖ If statement
- ❖ Elif statement
- ❖ Introduction to list in Python

### IT TRAINING & SERVICES

Regd. Office # 207, II floor, HUDA Maithrivanam, Ameerpet, Hyderabad 500 038.

Ph # +91 40 40310000, 23743392.

E-mail: enq@peerstech.com URL: www.peerstech.com

- ❖ List operations in Python
- ❖ List functions in Python
- ❖ Range function in Python
- ❖ Code reuse and functions in Python
- ❖ For loop in Python
- ❖ Boolean logic in Python
- ❖ While Loop in Python

### Functions and Modules in Python

- ❖ Passing arguments to functions
- ❖ Making function return value
- ❖ Passing functions as arguments
- ❖ Modules

### Exception Handling and File Handling in Python

- ❖ Errors and Exceptions
- ❖ Exception handling
- ❖ Finally Block
- ❖ File Handling
- ❖ Reading Data from file
- ❖ Adding data to file
- ❖ Appending to a file

### Some More Types in Python

- ❖ Dictionaries
- ❖ Dictionary Functions
- ❖ Tuples
- ❖ List Slicing
- ❖ List Comprehension
- ❖ String Formatting
- ❖ String functions
- ❖ Numeric Functions

### Functional Programming in Python

- ❖ Functional Programming
- ❖ Lambdas
- ❖ Map
- ❖ Filters
- ❖ Generators

## Business Statistics

### Introduction to Analytics

- ❖ Analytics Industry Overview
- ❖ Application of Analytics & Challenges to Analytics

### Data Understanding

- ❖ Data Types

- ❖ Summarizing Techniques
- ❖ Five Number Summary
- ❖ Histograms & Ogives
- ❖ Box Plots
- ❖ Scatter Diagram
- ❖ Frequency Tables and Distribution
- ❖ Cumulative Distributions

### Measure of Central Tendency, Dispersion and its importance

- ❖ Understanding Range
- ❖ Inter Quartile Range
- ❖ Variance
- ❖ Standard Deviation

### Probability and Probability Distribution

- ❖ Introduction to Probability
- ❖ Types of Probability
- ❖ Probability Rules
- ❖ Probability Distribution
- ❖ Random Variables
  - Discrete Random Variable
  - Continuous Random Variable
- ❖ Discrete Distributions
  - Binomial Distribution
  - Poisson Distribution
- ❖ Continuous Distribution
  - Normal Distribution
  - Standard Normal Distribution
  - Z scores

### Sampling and Sampling Distribution

- ❖ Introduction to Sampling
- ❖ Random Sampling & Non Random Sampling
- ❖ Sampling Techniques
  - Stratified Sampling Method
  - Cluster Sampling Method
- ❖ Sampling Distribution
- ❖ Central Limit Theorem
- ❖ Standard Error Concept

### Statistical Inference

- ❖ Estimation
  - Introduction
  - Point Estimates and Interval Estimates
  - Calculating Interval Estimates using 'Z' table
  - Introduction to 't' distribution

---

## IT TRAINING & SERVICES

Regd. Office # 207, II floor, HUDA Maithrivanam, Ameerpet, Hyderabad 500 038.

Ph # +91 40 40310000, 23743392.

E-mail: enq@peerstech.com URL: www.peerstech.com

- Degrees Of Freedom
- Calculating Interval Estimates using 't' table
- Confidence Intervals with t & z distributions
- Determining Sample Size in Estimation

### Hypotheses Testing

- ❖ Introduction
- ❖ One Sample Test & Two Sample Tests
  - Z test
  - t test
  - One Tail & Two Tail Test
  - Dependent & Independent Samples
- ❖ Concept of p-value
- ❖ ANOVA
  - Introduction
  - F distribution
  - One way ANOVA
- ❖ CHISQ Test
- ❖ Some Non Parametric Tests

### Simple Regression & Correlation

- ❖ Introduction
- ❖ Dependent and Independent Variables
- ❖ Correlation Analysis
- ❖ Estimation in Regression
- ❖ Least Squared Method
- ❖ Standard Error Of Line
- ❖ Finding Regression Equation
- ❖ Hypotheses Testing for estimates
- ❖ Limitations & Errors in Simple Regression Analysis
- ❖ Multiple Regression analysis: Introduction
- ❖ Multicollinearity
- ❖ Fitting the model
- ❖ Regression Assumptions
- ❖ Residual Analysis for Regression Assumptions

### Logistic Regression

- ❖ Understanding Logistics Regression
- ❖ Difference between linear and logistics regression
- ❖ Odds Ratio
- ❖ Logit Model
- ❖ Building Models
- ❖ ROC concept

- ❖ Model Fitting
- ❖ Evaluation of goodness of fit
- ❖ Model Suitability

## MACHINE LEARNING USING R-PROGRAMMING

- ❖ Introduction to Machine Learning
- ❖ Linear Regression
- ❖ Project : Linear Regression
- ❖ Logistic Regression
- ❖ Project : Logistic Regression
- ❖ K Nearest Neighbors
- ❖ Project : K Nearest Neighbors
- ❖ Decision Trees and Random Forests
- ❖ Project : Decision Trees and Random Forests
- ❖ Support Vector Machines
- ❖ Project : Support Vector Machines
- ❖ K – Means Clustering
- ❖ Project : K – Means Clustering
- ❖ Natural Language Processing
- ❖ Project : Natural Language Processing
- ❖ Neural Nets
- ❖ Project : Neural Nets

## MACHINE LEARNING USING PYTHON

- ❖ Introduction
- ❖ Environment Set-up
- ❖ Jupyter Overview
- ❖ Python for Data Analysis
  - Welcome to the Numpy
  - Numpy Arrays
  - Quick Note on Array Indexing
  - Numpy Array Indexing
  - Numpy Operations
- ❖ Numpy Exercises
- ❖ Python for Data Analysis
  - Welcome to Pandas
  - Series
  - Data Frames
  - Missing Data
  - Groupby
  - Merging, Joining and Concatenating
  - Operations
  - Data Input and Output
- ❖ Pandas Exercise

## IT TRAINING & SERVICES

Regd. Office # 207, II floor, HUDA Maithrivanam, Ameerpet, Hyderabad 500 038.

Ph # +91 40 40310000, 23743392.

E-mail: enq@peerstech.com URL: www.peerstech.com

- ❖ Python for Data Visualization
- ❖ Welcome to Matplotlib
  - Matplotlib
  - Matplotlib exercises
- ❖ Python for Data Visualization
  - Introduction to Seaborn
  - Distribution Plots
  - Categorical Plots
  - Matrix Plots
  - Grids
  - Regression Plots
  - Style and Color
- ❖ Seaborn Exercise
- ❖ Python for Data Visualization – Pandas Built-in Data Visualization
- ❖ Pandas Data Visualization Exercise
- ❖ Python for Data Visualization – Plotly and Cufflinks
- ❖ Python for Data Visualization – Geographical Plotting
- ❖ Introduction to Geographical Plotting
  - Choropleth Maps
  - Choropleth Exercises
- ❖ Capstone Project
- ❖ Introduction to Machine Learning
- ❖ Linear Regression
  - Theory
  - Model selection update for Scikit Learn 0.18
  - Linear Regression with Python
- ❖ Linear Regression Project
- ❖ Cross Validation and Bias-Variance Trade-off
- ❖ Logistic Regression
  - Theory
  - Logistic Regression with Python
- ❖ Logistic Regression Project
- ❖ K Nearest Neighbors
  - Theory
  - KNN with Python
- ❖ KNN Project
- ❖ Decision Trees and Random Forests
  - Introduction to Tree Methods
  - Decision Trees and Random Forest with Python
- ❖ Decision Trees and Random Forest Project
- ❖ Support Vector Machines
  - Theory
  - Support Vector Machines with Python
- ❖ SVM Project
- ❖ K Means Clustering
  - Theory
  - K Means with Python
- ❖ K Means Project
- ❖ Principal Component Analysis
  - Theory
  - PCA with Python
- ❖ Recommender Systems
  - Theory
  - Recommender Systems with Python
- ❖ Natural Language Processing
  - Theory
  - NLP with Python
- ❖ NLP Project
- ❖ Neural Nets and Deep Learning
  - Neural Network Theory
  - Welcome to Deep Learning
  - What is TensorFlow?
  - Changes with TensorFlow
  - TensorFlow Installation
  - TensorFlow Basics
  - MNIST with Multi-Layer Perceptron
  - TensorFlow with ContribLearn
- ❖ TensorFlow Project

## IT TRAINING & SERVICES

Regd. Office # 207, II floor, HUDA Maithrivanam, Ameerpet, Hyderabad 500 038.

Ph # +91 40 40310000, 23743392.

E-mail: enq@peerstech.com      URL: www.peerstech.com