Synopsis

Data Science is emerging as a hot new profession and academic discipline. Harvard Business Review says Data Scientist is the Sexiest Job of the 21st Century. But demand for data scientists is racing ahead of supply.

People with the necessary skills are scarce, primarily because the discipline is so new. This course is designed to give a start and introduction to this new discipline. This course is spread across 2 days and will have a plenty of hands on exercises on real world data.

Pre-requisites:

- A basic understanding of data and programming is required.
- Programming knowledge using Python is essential

Hardware & Software:

- A desktop or notebook with 64 bit OS (Windows/Mac)
- 8 GB RAM
- High speed Internet connection

Course Structure

Introduction to Data Science and Setting up data analysis environment
Introduction to Data Science
Setting up Python Environment for Data Analysis
Overview of Data Analysis Software Stack - Numpy, Pandas, Matplotlib, Scipy and Scikit-learn
Hands On Exercise
Working with Numbers
Introduction to Numpy array
Overview of Array and operations
N-Dimensional array and manipulations
Accessing and preparing data with Pandas
Loading data from varieties of sources: CSV, Databases
Data manipulation - Filtering, Grouping, Ordering of data
Dealing with missing Data
Dealing with Continuous and categorical variables
Normalizing and transforming data
Data Exploration and Visualization using Scipy and matplotlib
Basic Statistical analysis using Scipy
Univariate Analysis – Statistics Summary, Hypothesis tests
Drawing Histograms, Bar charts, Density Plots, Box Plots
Drawing Density plots and understating data distributions
Regression and Classification Algorithms
Understand Regression Techniques
Simple Linear Regression & Multiple Linear Regressions

Measuring accuracy of the models
Regression Diagnostics - Validating Models
Making Predictions using Classification algorithms - Logistic Regression
Clustering
Understanding k-means clustering and creating Segments
Creating clustering plots
Creating customer segments using Clustering Algorithms
Text Analytics
Handling Text and unstructured Data
Accessing Social Data - Integrate with Twitter
Trend Analysis & Topic Modelling

Benefits

In this special two-day course, you will gain practical data science experience under the guidance of an industry expert. You will learn how to structure a data science project which will significantly increase your chances of producing valuable work.
At the end of the two days, you will have the ability to talk about data science with confidence and begin to work on data science projects.