PY200: *Data Analytics with Python*

Course Duration: 2 days; Instructor-led

WHAT YOU WILL LEARN

Learn how to analyze data using Python. This course will take you from the basics of Python to exploring many different types of data. You will learn how to prepare data for analysis, perform simple statistical analyses, create meaningful data visualizations, predict future trends from data, and more.

AUDIENCE

Anyone who wants to use Python programming language to do Data Analysis

PREREQUISITES

Python programming knowledge is preferred.

METHODOLOGY

This program will be conducted with interactive lectures, PowerPoint presentation, discussion and practical exercise.

COURSE OBJECTIVES

The course will introduce data manipulation and cleaning techniques using the popular python pandas data science library and introduce the abstraction of the Series and Data Frame as the central data structures for data analysis, along with tutorials on how to use functions such as group by, merge, and pivot tables effectively. At the end of this course, participants will be able to take tabular data, clean it, manipulate it, and run basic inferential statistical analyses.

OUTLINE

Module 1: Python Programming Refresher

- Control Flows
- Data Types
- Functions and Parameters
- Useful Collections

Module 2: Importing libraries

- Creating module for reuse
- Importin modules

Module 3: File Operations (Reading/Writing)

- TXT
- CSV
- MS EXCEL

Module 4: Database Operations: MySQL

- The DB API
- Connecting to a server
- Creating and executing a cursor
- Fetching data
- Parameterized statements

Module 5: Introduction to Pandas

- What is Pandas?
- Creating data sets and data frames

Module 6: Data Manipulation, simple statistics and data plotting using Pandas

- Finding Maximum & Minimums
- Calculating and identifying outliers
- Indexing and Selecting Data
- Reshaping and Pivot Tables
- Time Series / Date functionality
- Categorical Data
- Simple Statistical Analysis

Module 7: Visualization

- Basic Plotting
- Plot Customization