Data Science using Python

60 Hours

(Basic + Intermediate +Advanced)

Programming & development

Python Basic:

An understanding of how to use the Python standard library to write programs, access various tools, and document and automate analytical processes.

- Types (strings, lists, dictionaries, and more)
- Control Flow (if-then statements, looping)
- Organizing code (functions, modules, packages)
- Reading and writing files
- Overview of Object-Oriented Programming (OOP)

NumPy & 2D Plotting Library:

Introduction to NumPy and 2D plotting. The NumPy package is presented as a tool for rapidly manipulating and processing large data sets. 2D plotting is introduced with matplotlib.

- Understanding the N-dimensional data structure
- Creating arrays
- Indexing arrays by slicing or more generally with indices or masks
- Basic operations and manipulations on N-dimensional arrays
- Plotting with matplotlib

Python Pandas & Data Analysis:

the Python Data Analysis Library (Pandas) is a powerful and convenient package

- Tabular Datasets
- Data Aggregation & Data Exploration
- Labelling data for each dimensionBasic operations and manipulations on N-dimensional arrays
- Dealing with missing values, and time series manipulations.

Accessing Data from & multiple sources:

- Reading and writing data from local files (.txt,.csv,.xls, json, etc.)
- Reading data from remote files
- Scraping tables from web pages (.html)
- Making the most of the powerful read table method

Data Preparation & Cleaning:

- Working with Pandas data structures: Series and Data Frames.
- Accessing your data: indexing, slicing, fancy indexing, Boolean indexing.
- Data wrangling, including dealing with dates and times and missing data's.
- Adding, dropping, selecting, creating, and combining rows and columns.

Data Access & Databases:

- Database access with DB-API2 and SQL Alchemy.
- Executing SQL commands from Pandas
- Loading database data into a Data Frame.
- Combining and manipulating Data Frames: merge, join, concatenate

Data Visualization:

- Understanding the structure of a Figure
- Data visualization: scatter plots, line plots, box plots, bar charts, and histograms with matplotlib
- Customizing plots: important attributes and arguments

Data Analysis:

- Split-apply-combine with Data Frames
- Data summarization and aggregation methods

- Pandas powerful group by method
- Reshaping, pivoting, and transforming your data
- Simple and rolling statistics

Real World Modelling & Problem Solving:

- Deep learning of the data analysis tools through lectures, Q&A, and hands-on exercises
- Develop transferable skills through application to authentic data sets
- Predict the future with time series analysis
- And more!

Python Data Science:

- Linear Regression
- SVM (Support Vector Machine)
- KNN (K-Nearest Neighbors)
- Logistic Regression
- Decision Tree
- K-Means
- Random Forest
- Naive Bayes
- Dimensional Reduction Algorithms
- Gradient Boosting Algorithms

Python Forecasting Modelling in Data Science:

- Autoregression (AR)
- Moving Average (MA)
- Autoregressive Moving Average (ARMA)
- Autoregressive Integrated Moving Average (ARIMA)
- Seasonal Autoregressive Integrated Moving-Average (SARIMA)
- Seasonal Autoregressive Integrated Moving-Average with Exogenous Regressors (SARIMAX)
- Vector Autoregression (VAR)
- Vector Autoregression Moving-Average (VARMA)
- Vector Autoregression Moving-Average with Exogenous Regressors (VARMAX)
- Simple Exponential Smoothing (SES)
- Holt Winter's Exponential Smoothing (HWES)