

LEARINGO SOFTWARE AND EDUCATION PVT LTD

PYTHON DATA SCIENCE

MODULE -1

- Introduction to Data Science
- Data Science Era
- Data Science involvement in industries
- Business Intelligence vs. Data Science
- Data Science Life Cycle
- Tools of Data Science
- Introduction to Python
- Introduction to Machine Learning

MODULE -2

- Introduction to Python Programming
- Introduction to Python
- Basic Operations in Python
- Variable Assignment
- Functions in-built functions, user defined functions
- Condition if,if-else,nested if-else, else-if
- Pre Reads (Attachment for students)
- Assignment Solution

MODULE -3

- Data Structure – Introduction
- List: Different Data Types in a List, List in a List
- Operations on a list:Slicing,Splicing,Sub-setting
- Condition(true/false) on a List
- Applying functions on a list
- Dictionary Index ,Value
- Operation on a Dictionary Slicing, Splicing, Sub-setting
- Condition (true/false) on a Dictionary
- Applying functions on a Dictionary
- Modules and Packages
- Regex operation's
- Pre Reads
- Assignment Solution

MODULE -4

- Introduction to SQL(Structured Query Language)
- Basic SQL statement
- Advanced SQL(Searching sorting, grouping)
- Accessing databases using python

MODULE -5

- Data Types in an Array, Dimensions of an Array
- Operations on Array : Indexing ,Slicing, Splicing, Sub-setting
- Conditional(T/F) on an Array
- Loops:For,While
- Shorthand for For
- Control statements
- Shape Manipulation
- Linear Algebra

MODULE -6

- Python Pandas – Home
- Python Pandas – Introduction
- Python Pandas – Environment Setup
- Introduction to Data Structures
- Python Pandas – Series
- Python Pandas – Data Frame
- Python Pandas – Panel
- Python Pandas – Basic Functionality
- Function Application
- Python Pandas – Reindexing
- Python Pandas – Iteration
- Python Pandas – Sorting
- Working with Text Data
- Options & Customization
- Indexing & Selecting Data
- Python Pandas – Missing Data
- Python Pandas – Group By
- Python Pandas – Merging / Joining
- Python Pandas – Concatenation
- Python Pandas – Date Functionality
- Python Pandas – Categorical Data
- Python Pandas – Visualization
- Pre Reads (Attachment for students)
- Assignment (For Student)
- Assignment Solution

MODULE -7

- Intro to Statistics
- Statistical Inference
- Terminologies of Statistics, Descriptive statistics
- Statistical functions Measures of Centres
- Mean
- Median
- Mode Measures of Spread
- Variance Standard Deviation
- Histogram Probability
- Normal Distribution
- Binary Distribution
- Poisson Distribution
- Skewness
- Bell curve
- Hypothesis Building and Testing
- Chi-Square Test
- Correlation Matrix

MODULE -9

- Data Analysis Pipeline
- What is Data Extraction
- Types of Data (Raw & Processed Data)
- Data Wrangling Exploratory Data Analysis Data Visualization Matplotlib
- Bar Plot
- Histograms Plot
- Box Plot
- Area Plot
- Scatter Plot
- Pie Plot
- Sea born
- Pre Reads (Attachment for student)

MODULE -11

- Data Pre-processing
- Data preparation
- Intro to Scikit Learn

MODULE -13

- Classification K-nearest neighbours
- Metrics
- Confusion Matrix
- Classification report
- Support Vector Machines
- Kernel

MODULE -8

- Scientific computing with python
- SciPY and its Characteristics
- SciPy sub-packages
- SciPy sub- packages – Integration
- SciPy sub – packages – Optimize
- Linear Algebra
- SciPy sub-packages – Statistics

MODULE -10

- Introduction to Machine Learning
- Machine Learning Use-Cases
- Machine Learning Process Flow
- Machine Learning Categories

MODULE -12

- Regression
- Types
- Algorithms
- Linear Regression
- RMSE
- R2 score
- Logistic Regression
- Introduction to Dimensionality
- Why Dimensionality Reduction
- PCA
- Factor Analysis
- Scaling Dimensional model
- Encoding
- Implementation with Case Studies
- Intro to Kaggle and UCI repository
- Pre Reads (Attachment for student)

MODULE -14

- Unsupervised learning
- Clustering Algorithms
- K-Means Clustering
- Hierarchical Clustering
- Implementation with Case Studies
- Pre Reads (Attachment for students)

MODULE -15

- Recommendation Engine
- Collaborative filtering
- The course will be covering 12 + Industrial real time case studies

- Working of SVM
- Native Bayes
- Hyper parameter Optimization
- Decision Tree Classifier
- Random Forest Classifier
- Ensemble Techniques and SVM tuning
- Underfitting & Overfitting
- Entropy
- AUC-ROC Curve
- Implementation with Case Studies
- Cross – Validation
- Pre Reads (Attachment for students)

MODULE-16

Specialization Course: Students can opt any one Specialization

Spec-1: Deep Learning with Computer Vision

- Deep Learning
- Computer Vision OCR/OCV
- Advanced SQL
- Tableau
- Cloud Computing

Spec-2: Deep Learning with NLP

- Deep Learning
- NLP
- Advanced SQL
- Deep NLP
- Tableau
- Cloud Computing

Spec-3: Business Analytics with R

- R and R Shiny
- Advanced SQL
- Power BI
- Advanced Excel
- Cloud Computing

Spec-4: Business Analytics with Tableau

- Tableau
- Advance Excel
- Google Data Studio
- Power BI
- Advanced SQL
- Cloud Computing

Spec-5: Data Engineering

- Big Data & Linux os
- Hadoop
- Apache Spark
- Advanced SQL
- Google Data Studio
- Cloud Computing

Spec-6: Business Analytics with SAS

- Power BI
- Advanced Excel
- Google Data Studio
- SAS
- Advanced SQL
- Cloud Computing