TechyEdz Solutions

Training | Consulting | Developement | Outsourcing

Python and Data Science

9032803832

contact@techyedz.com

9032803832

www.techyedz.com

Python + Data Science Combo Course

Course Overview:

Python is the most popular programming language for Data Science as on Today. Python is powerful , easy to learn and flexible tool for coding Data Science and Machine Learning algorithms. In recent years, Python has evolved immensely with respect to Data Science sphere, with a huge community around Python creating quite a few power data science and analytics packages such as Pandas, Numpy, Scikit Learn, Scipy and more. As a result, analyzing data, modeling machine learning algorithms with Python has never been easier.

This course "Python for Data Science", is designed for candidates with or without programming skills, with basics of Data importing, Data mugging and coding Machine Learning algorithms along with effective programming techniques. This also includes Python Data Science challenges kit, enabling the candidates to not only understand Python core concepts but also gain practical mastery over Python for Data Science, which is very much in demand in Today's Data Science job opportunities.

Course Outline:

Python

1. Getting Started

- History
- A Python Q&A Session
- How Python Runs Programs
- How You Run Programs

2. Introduction to Python:

- What is Python?
- Why Python?
- Python Applications in real life
- Brief history of Python
- Versions of Python

- Installing Python
- Using IDLE
- First Python Program
- Getting help from Python Docs

3. Types and Operations

- Introducing Python Object Types
- Numeric Types
- The Dynamic Typing Interlude
- Strings
- Lists and Dictionaries
- Tuples, Files and Everything Else

4. Variables Data types

- Intro to dynamic typing
- Variables in Python
- Naming conventions
- Basic Data types (representation of strings, integer, floats)

5. Basic Syntax

- Basic syntax
- Commenting
- Indentation
- Python keywords
- Strings
- String values
- String Operations
- String slicing
- Built in string methods
- Formatted printing
- Simple Input and Output handling

6. **Language Building blocks**

- Control statements, the if, elif, else
- True and False
- Arithmetic Operators
- Relational Operators
- Logical Operators
- Bitwise Operators
- While loop
- Usage of pass, break and continue
- For each loop

7. **Collections**

- Lists
- Tuples
- Sets
- Dictionaries
- Sorting collections
- Operations on collections
- Discussion on real life application of above collections

8. **Functions**

- Introduction to functions
- Built in functions
- User defined functions
- Function parameters
- Variable arguments ,args and kwargs
- Positional and named arguments
- Discussion scope of variables with respect to functions and namespace
- Passing function to another function

9. **Project**

10. File Handling

11. Modules

- Introduction to modules
- Introduction to standard modules
- OS module
- path module
- Sys module
- sub process module
- Argument parsing using argparse module
- .csv file parsing using csv module
- .jason file paring using Jason module
- Xml file parsing using xml module
- Introduction to logging module

12. Project 2: Building log parser and reporting the results

13. Object Oriented Programming

- Introduction to Classes and Objects
- Principles of OOP
- Instance methods
- Special methods
- Encapsulation
- Inheritance
- Polymorphism

14. Regular Expressions

- Introduction to regular exceptions
- Introduction to re module
- Simple character matches
- Match function

- Searching function
- Regular expression patterns
- Patterns in Regex
- Search And Replace

15. Optional I(for testers)

- Introduction to testing using Python
- Introduction to test automation
- Introduction to Selenium web deriver
- Web testing using selenium

16. Option II(developers) Advance topics:

- Generators
- Decorators
- Iterators and iterator protocol
- Debugging using PDB

17. Options III(Web programming)

- Introduction to web programming using Python
- Introduction to Django/Flask
- Introduction to Restful API's using Python

18. Option IV(Data science)

- Introduction to data science using python
- Introduction to pandas module
- Introduction to data visualization using matplotlib
- Introduction to numpy
- Introduction to scipy

Data Science

Introduction to Data Science

- What is Data Science?
- What is Machine Learning?
- What is Deep Learning?
- What is AI?
- Data Analytics & it's types

Python Programming Fundamentals

- Programming Basics
- Python Data Types
- Structures and conditional statements
- Python core packages
- Introduction to Jupyter Notebook

Data Science Essentials

- Data Science Introduction
- Data Science work flow
- Machine Learning Overview

Data Mugging with Numpy and Pandas

- Introduction to Numpy and Pandas
- Data filtering and selecting
- Find duplicates and treating missing values
- Concatenate and transform data

Basic Statistics

1. Central Tendency

- Mean
- Median
- Mode

- Skewness
- Normal Distribution

2. Probability Basics

- What does it mean by probability?
- Types of Probability
- ODDS Ratio?

3. Standard Deviation

- Data deviation & distribution
- Variance

4. Bias variance Tradeoff

- Underfitting
- Overfitting

5. Distance metrics

- Euclidean Distance
- Manhattan Distance

6. Outlier analysis

- What is an Outlier?
- Inter Quartile Range
- Box & whisker plot
- Upper Whisker
- Lower Whisker
- catter plot
- Cook's Distance

7. Missing Value treatments

- What is an NA?
- Central Imputation
- KNN imputation
- Dummification

8. Correlation

- Pearson correlation
- Positive & Negative correlation

9. Error Metrics Duration

- Classification
- Confusion Matrix
- Precision
- Recall
- Specificity
- F1 Score

10. Regression

- MSE
- RMSE
- MAPE

11. Visualization, web scraping

- Creating basic charts
- Statistical Charts
- Web Scrapping tools

Introduction to Machine Learning

- Overview of Supervised and Unsupervised Machine Learning
- Linear Regression
- Clustering with K-means
- Naive Bayes Classification
- Introduction to Neural Networks

Supervised Learning

Linear Regression

- Linear Equation
- Slope
- Intercept
- R square value

Logistic regression

- ODDS ratio
- Probability of success
- Probability of failure
- ROC curve
- Bias Variance Tradeoff

Unsupervised Learning

- K-Means
- K-Means ++
- Hierarchical Clustering

Other Machine Learning algorithms

- K – Nearest Neighbour
- Naïve Bayes Classifier
- Decision Tree – CART
- Decision Tree – C50
- Random Forest

Prerequisites:

- Basic Programming is recommended
- Basic Statistics knowledge is recommended

Who Can attend:

- Candidates wanted to pursue Data Science career, with basic or no programming skills
- Seasoned conventional programmer aiming to gain basic machine learning coding skills
- Job seekers, pursuing a career as Data Science Developer
- Professionals, whose job involves Data Science and Python.

Number of Hours: 70hrs

Key Features:

- One to One Training
- Online Training

- Fastrack & Normal Track
- Resume Modification
- Mock Interviews
- Video Tutorials
- Materials
- Real Time Projects
- Virtual Live Experience
- Preparing for Certification