

Course Title: Python + Statistics + Pyspark

Course Duration: 40 hours/20 Sessions

Introduction

Data analytics is the process of gathering, transforming, and organizing information in order to draw conclusions, make predictions, and make better decisions. The full stack data analytics course is meant to assist you in becoming a skilled data analyst. Learn how to deal with SQL databases, develop data visualizations, and apply predictive analytics and statistics in a corporate environment using the best analytics tools and methodologies.

Requirements

1. System with Internet Connection
2. Interest to learn
3. Dedication

What you will get

- Course Features
- Course material
- Course resources
- On demand recorded videos
- Practical exercises
- Quizzes
- Assignments for every topic
- Hands on Project
- Course completion certificate

Python Installation Link

<https://www.anaconda.com/products/distribution>

Google Colab Link

<https://colab.research.google.com/>

Step by step installation link

<https://docs.anaconda.com/anaconda/install/windows/>

Day 1.

- What is Python...?
- A Brief history of Python
- Why Should I learn Python...?
- Career opportunities after learning Python
- Installing Python
- Overview of Anaconda and different IDEs

Day 2.

- Write your first program
- Objects in Python
- Different data types
- Immutable and Mutable
- Object Referencing and Identity
- Operators in Python

Day 3.

- %Run %Load
- Strings and escape char
- String Formatting F string
- Formats for string formatting
- Input Statement
- Indexing, Slicing & Dicing

Day 4.

- List
- Tuple
- Dictionary
- Sets
- Control Flow Statement (If else conditions)

Day 5.

- For loop
- Enumerator for Iterations
- Break statement
- Continue statement
- Nested loops
- While loops
- Infinite loops

Day 6.

- List methods
- List Comprehension
- List Comprehension with If & Else
- Tuple Methods
- Tuple Comprehension

Day 7.

- Dictionary methods
- Dictionary Comprehension
- String methods
- Regular Expressions

Day 8.

- Function in python
- Lambda
- MAP
- Filter
- Reduce
- Zip
- Accumulate
- Decorators

Day 9.

- Modules
- Datetime
- Classes in Date time
- Iterable and Iterator
- Generator
- Generator vs. List

Day 10.

- What is Statistics?
- Population vs. Sample
- Different sampling techniques

- Different types of charts & Graphs
- Measure of Central Tendency

Day 11.

- Measure of Dispersion
- Normalisation vs. Standardisation (Z score)
- 5 number summary
- Probability
- Stats Practical with Math's, Stats & NumPy Library

Day 12.

- OOPS Concept
- Class & Objects
- `__init__` Method
- `__str__` Method
- Abstraction
- Inheritance
- Super() Method
- Encapsulation
- Polymorphism

Day 13.

- File Handling in Python
- File objects and Modes of file operations
- Reading, writing and use of 'with' keyword
- Exception Handling
- Try Except Block

Day 14.

- Logging and debugging
- Python Connection with SQL Server

Day 15.

- NumPy

Day 16.

- Pandas

Day 17.

- Pandas, Data analysis – Visualization using Pandas, Matplotlib, Plotly & Seaborn

Day 18.

- Python Project Discussion

Day 19.

- Introduction to Big Data
- Apache Spark Introduction
- PySpark Introduction
- PySpark Queries

Day 20.

- Introduction to Git Hub
- Version Controlling and useful Git Commands
- Introduction to API in Python



Institute for Data Science

Ph. +91 87500-28877

Web: www.thexlacademy.com

Email: info@thexlacademy.com

Thank You!!

The XL Academy