Data analysis with Python

Duration: 5 Days

Overview:
This course is intended to convert a professional into a job-ready data analyst. To achieve that, we will cover the fundamental topics extensively. In other words, we are not going to teach you how to analyse data before you know how to gather and clean it. The course provides complete preparation for someone who wants to become a data. We believe that this course will significantly boost your chances of landing a job, as it will prepare you for practical tasks and concepts that are frequently included in interviews.

Course Objectives:
In this course, you will learn:
- Acquire a big picture understanding of the data analyst role
- Be able to work with text files
- Complete a data cleaning exercise
- Learn how to create pie, bar, line, area, histogram, scatter, regression, and combo charts
- Python programming, NumPy, pandas, data preparation - data collection, data cleaning, data preprocessing
- Learn beginner and advanced Python
- Understand different data types and their memory usage
- Master data visualization

Course Content:

Lesson 1: Introduction to Data Analytics
- Introduction to the World of Business and Data
- Key terms
- Data Analyst vs. Other Data Jobs
- What is Programming
- Getting started with Jupyter

Lesson 2: Python Basics with Exercises
- Python Variables
- Types of Data - Numbers and Boolean Values
- Types of Data – Strings
- Arithmetic Operators
- The Double Equality Sign
- Reassign Values
- Add Comments
- Line Continuation
- Indentation
- Comparison Operators
- Logical and Identity Operators

- Conditional Statements –
    - The IF Statement
    - The ELSE Statement
    - The ELIF Statement
    - A Note on Boolean Values
- Functions –
    - Defining a Function in Python
    - Creating a Function with a Parameter
    - Another Way to Define a Function
    - Using a Function in Another Function
    - Combining Conditional Statements and Functions
    - Creating Functions That Contain a Few Arguments
    - Notable Built-in Functions in Python
- Sequences-
    - Lists
    - Methods
    - List Slicing
    - Tuples
    - Dictionaries
- Iterations-
    - For Loops
    - While Loops and Incrementing
    - Create Lists with the range () Function
    - Use Conditional Statements and Loops Together
    - Conditional Statements, Functions, and Loops
    - Iterating over Dictionaries

Lesson 3: Fundamentals for coding in Python
- Object-Oriented Programming
- Modules, Packages, and the Python Standard Library
- Importing Modules
- Introduction to Using NumPy and pandas
- What is Software Documentation?
- The Python Documentation

Lesson 4: Mathematics for Python
- What is Matrix
- Scalars and Vectors
- Linear Algebra and Geometry
- Arrays in Python
- What Is a Tensor?
- Adding and Subtracting Matrices
- Errors When Adding Matrices
- Transpose

- Dot Product of Vectors
- Dot Product of Matrices

Lesson 5: NumPy Basics
- Importance of NumPy Package
- Installing/ Upgrading NumPy
- Nd array

Lesson 6: Pandas Basics
- Introduction to the pandas Library
- Installing and Running pandas
- Introduction to pandas Series
- Working with Attributes in Python
- Using an Index in pandas
- Label-based vs Position-based Indexing
- Working with Indices in Python
- Using Methods
- Parameters vs Arguments
- Introduction to pandas DataFrames
- Creating DataFrames from Scratch

Lesson 7: Working with Text Files
- Working with Files in Python
- File vs File Object, Read vs Parse
- Structured vs Semi-Structured and Unstructured Data
- Data Connectivity through Text Files
- Principles of Importing Data in Python
- More on Text Files (*.txt vs *.csv)
- Fixed-width Files
- Common Naming Conventions Used in Programming
- Importing Data –
    - Text Files in Python (open())
    - Importing *.csv Files with pandas
    - "index_col" Parameter
    - With NumPy - .loadtxt() vs genfromtxt()
    - With NumPy - Partial Cleaning While Importing
    - *.json Files
- Working with Excel Data
- Importing Data with the pandas' "Squeeze" Method
- Saving Your Data with pandas and NumPy

Lesson 8: Working with Text Data
- Working with Text Data and Argument Specifiers
- Manipulating Python Strings

- Using Various Python String Methods
- String Accessors
- Using the .format() Method

Lesson 9: Python Tools
- Iterating Over Range Objects
- Nested For Loops
    - Introduction
    - Triple Nested For Loops
- List Comprehensions
- Anonymous (Lambda) Functions

Lesson 10: Data Processing
- Data Gathering and collection
- Cleaning and Pre-processing

Lesson 11: Pandas
- Running Pandas
- .unique(), .nunique()
- Converting Series into Arrays
- .sort_values()
- Attribute and Method Chaining
- .sort_index()
- Recap to Pandas DataFrames
- Common Attributes for Working with DataFrames
- Data Selection in pandas DataFrames

Lesson 12: NumPy Fundamentals
- Indexing in NumPy
- Assigning Values in NumPy
- Elementwise Properties of Arrays
- Types of Data Supported by NumPy
- Characteristics of NumPy Functions
- Ndarrays
- Arrays vs Lists
- Strings vs Object vs Number

Lesson 13: Working with Arrays
- Basic Slicing in NumPy
- Stepwise Slicing in NumPy
- Conditional Slicing in NumPy
- Dimensions and the Squeeze Function

Lesson 14: Generating Data with NumPy
- Arrays of 0s and 1s
- "_like" functions in NumPy
- A Non-Random Sequence of Numbers
- Random Generators and Seeds
- Basic Random Functions in NumPy
- Probability Distributions in NumPy
- Applications of Random Data in NumPy

Lesson 15: Statistics with NumPy
- Using Statistical Functions in NumPy
- Minimal and Maximal Values in NumPy
- Statistical Order Functions in NumPy
- Averages and Variance in NumPy
- Covariance and Correlation in NumPy
- Histograms in NumPy

Lesson 16: NumPy – Preprocessing
- Checking for Missing Values in Ndarrays
- Substituting Missing Values in Ndarrays
- Reshaping Ndarrays
- Removing Values from Ndarrays
- Sorting Ndarrays
- Argument Sort in NumPy
- Argument Where in NumPy
- Shuffling Ndarrays
- Casting Ndarrays
- Striping Values from Ndarrays
- Stacking Ndarrays
- Concatenating Ndarrays

Lesson 17: Data Visualization
- What Is Data Visualization and Why Is It Important?
- Why Learn Data Visualization?
- Choosing the Right Visualization
- Introduction into Colors and Color Theory
- Bar Chart
- Pie Chart
- Stacked Area Chart
- Line Chart
- Histogram
- Scatter Plot
- Regression Plot
- Combination Charts

Practical Case studies

- A Loan Data Example with NumPy
- The "Absenteeism"