

Data Science & Machine Learning with Python Curriculum

Training Provider

Company: Scenario Academy

Office: LKJ House, 3/5 Adeyemo Alajika, Victoria Island, Lagos

Office Hours: 9:00am to 5:00pm

Office Telephone: +234(0)8076774912

E-mail: training@scenarioacademy.com

Website: www.scenarioacademy.com

Course Details

Duration

- **Weekday:** Ten (10) days (Monday–Friday)
- **Weekend:** Eight (8) days (Saturday–Sunday)

Description

Data Science is about extraction, preparation, analysis, visualization, and maintenance of information. It is a cross-disciplinary field, which uses scientific methods and processes to draw insights from data.

With the emergence of new technologies, there has been an exponential increase in data. This has created an opportunity to analyze and derive meaningful insights from data. It requires special expertise of a 'Data Scientist' who can use various statistical and machine learning tools to understand and analyze data. A Data Scientist, specializing in Data Science, not only analyzes the data but also uses machine-learning algorithms to predict future occurrences of an event.

Therefore, we can understand Data Science as a field that deals with data processing, analysis, and extraction of insights from the data using various statistical methods and computer algorithms.

Benefits

- A workshop certificate issued after the training
- Post-training employability support and
- Mentored Learning Approach

Course Outline

Introduction to Data Science

Module 1- Introduction- Getting Started with Python

- Brief review of what Data science is (Differences between a data analyst and a data scientist)
- Environment Set-up
- Overview of different Integrated Development Environment (Jupyter Notebook, Spyder, Pycharm)
- Exercise

Module 2- Handling Program Flow in Python

- Exploring Data types- Numbers, Strings, Print Formatting, Lists, Dictionaries, Booleans, Tuples and Set
- Comparison Operators- IF, ELIF and ELSE statements, For loops, range(), List comprehension, Functions, Lambda Expression
- Take home work

Module 3- Handling Data with NumPy

- Introduction to Numpy- Numerical Python
- Exploring one-dimensional array and two-dimensional arrays
- Numpy methods
- Numpy array indexing
- Numpy Operations

Module 4- Data Wrangling with Pandas

- Introduction to Pandas
- Series and DataFrames
- GroupBy and Missing Data
- Merging, Joining and Concatenation
- Data Input and Output (Extract-Transform-Load)
- Pandas Operations

Module 5- Python for Data Visualization- Matplotlib & Seaborn

- Introduction to Matplotlib
- Visualization with Matplotlib
- Introduction to seaborn, Visualizations with Seaborn

FOUNDATIONAL MACHINE LEARNING

Module 6- Introduction to Machine Learning

- What is machine learning?
- Types of machine learning- Supervised and Unsupervised

Module 7- Summarizing Data with Statistics

- Data preparation with pandas
- Filling missing data

Module 8- Making inference from Data

- Matplotlib and seaborn for inference

Module 9- Build your first model – Linear Regression

- What is Linear regression
- Univariate Linear regression
- Multivariate linear regression
- EDA and Data Pre-processing
- Feature Engineering
- Improving Model with Feature Selection Technique
- Linear Regression

Other Machine Learning Algorithms

- Logistic Regression
- Random Forests/Decision Trees
- SVM
- Naïve Bayes
- KNN

FINAL CAPSTONE PROJECT/ HACKATHON