



PYTHON TRAINING

Python Data Engineering Machine Learning Course

Master Python for data engineering and machine learning. Develop advanced ML solutions and optimize data workflows.

DURATION

2 hours

LEVEL

Intermediate

FORMAT

Instructor-Led

- Globally Recognized Certification

- Expert-Led Hands-On Training

- Flexible Scheduling

Computer Pride | 1st Floor, ICEA Building, Kenyatta Avenue, Nairobi | www.computer-pride.co.ke

COURSE OVERVIEW

This intermediate-level Python Data Engineering Machine Learning Course is meticulously designed for professionals aiming to bridge the gap between robust data infrastructure and intelligent machine learning applications. Over 2 intensive hours, you will dive deep into leveraging Python for critical data engineering tasks, including data acquisition, cleaning, transformation, and building efficient data pipelines. The course then seamlessly transitions into the practical implementation of machine learning models, focusing on how engineered data fuels powerful predictive analytics and AI solutions.

Participants will gain hands-on experience with industry-standard Python libraries such as Pandas for data manipulation, NumPy for numerical operations, and Scikit-learn for foundational machine learning algorithms. We'll also touch upon concepts vital for operationalizing ML models, understanding the lifecycle from raw data to deployed intelligent systems. The professional benefits are immense: you'll enhance your ability to design, build, and maintain scalable data solutions that directly support and empower advanced machine learning initiatives, making you an invaluable asset in data-driven organizations. This course is your gateway to developing full-stack data capabilities, enabling you to contribute to cutting-edge AI projects with confidence and expertise.

Duration	2 hours
Level	Intermediate
Vendor	Python

COURSE CURRICULUM

01 Module 1: Python Fundamentals for Data & ML Readiness

- Review of Python programming essentials (data structures, functions, OOP concepts)
- Introduction to Anaconda and Jupyter Notebooks
- Efficient data handling with NumPy arrays
- Data manipulation and analysis with Pandas DataFrames

02 Module 2: Core Data Engineering Principles with Python

- Understanding ETL (Extract, Transform, Load) processes
- Data cleaning and preprocessing techniques (missing values, outliers, data types)
- Building basic data pipelines for structured and semi-structured data
- Data persistence: Working with CSV, JSON, and simple database connections

03 Module 3: Introduction to Machine Learning Concepts & Scikit-learn

- Overview of supervised vs. unsupervised learning
- Key ML algorithms: Linear Regression, Logistic Regression, Decision Trees
- Model training, evaluation metrics (accuracy, precision, recall, F1-score)
- Feature engineering and selection basics using Python

04 Module 4: Integrating Data Engineering with Machine Learning

- Preparing engineered data for ML model input
- Workflow for data transformation and model training integration
- Introduction to MLOps concepts: Model versioning and basic deployment considerations
- Building a simple end-to-end data engineering to ML pipeline project

CAREER PATHWAYS

Graduates of this program are prepared for the following roles:

- > **Junior Data Engineer specializing in Machine Learning Pipelines**
- > **Machine Learning Operations (MLOps) Associate**
- > **Data Scientist (Focus: ML Model Deployment & Integration)**

FREQUENTLY ASKED QUESTIONS

Q: What specific Python prerequisites are essential to succeed in this intermediate course?

To excel in this course, participants should have a foundational understanding of Python programming, including familiarity with basic data types, control structures (loops, conditionals), functions, and ideally, some exposure to object-oriented programming concepts. While we will briefly review data fundamentals, prior experience with libraries like Pandas and NumPy is beneficial but not strictly required, as key concepts will be covered.

Q: How will this course prepare me to build and maintain data pipelines specifically for machine learning applications?

This course focuses on teaching you to engineer data effectively to serve machine learning models. You will learn methodologies for data extraction, robust cleaning, transformation, and loading (ETL) using Python. Crucially, we cover how to structure these processes into reliable data pipelines that provide clean, preprocessed data ready for model training and inference, ensuring your ML projects are built on a solid data foundation.

Q: Beyond training models, does this course cover any aspects of deploying or operationalizing machine learning solutions in a practical setting?

While the primary focus is on data engineering and machine learning model development, this course introduces essential MLOps (Machine Learning Operations) concepts crucial for practical deployment. You will gain an understanding of how engineered data facilitates model operationalization, including discussions around preparing models for deployment, basic versioning, and the challenges of integrating models into real-world applications. This provides a critical stepping stone towards a deeper dive into MLOps.

ABOUT COMPUTER PRIDE



East Africa's Premier IT Training Center

With over 15 years of excellence, Computer Pride delivers globally recognized certification programs from leading technology vendors. Located on Kenyatta Avenue in Nairobi's CBD, our state-of-the-art facilities and expert instructors have empowered thousands of professionals across East Africa.

15+

Years of Excellence

10,000+

Professionals Trained

50+

Vendor Certifications

READY TO ENROL?

Take the next step in your technology career.

Location

1st Flr, ICEA Building
Kenyatta Ave, Nairobi

Email

info@computer-pride.co.ke

Phone

+254 705 900 900

www.computer-pride.co.ke