



DATABRICKS TRAINING

Databricks Data Engineer

Become a Databricks Data Engineer. Learn to build and manage data pipelines on the Databricks Lakehouse Platform.

DURATION

Flexible

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 Databricks Data Engineer course is meticulously designed for data professionals looking to master the art of building and maintaining robust, scalable data pipelines on the Databricks Lakehouse Platform. You'll dive deep into core technologies such as Apache Spark for distributed data processing, Delta Lake for reliable data storage and ACID transactions, and Structured Streaming for real-time data ingestion. The curriculum emphasizes hands-on experience with SQL and Python (or Scala), equipping you with the practical skills to transform raw data into actionable insights efficiently and effectively.

By completing this course, you will be proficient in designing, developing, and deploying high-performance data engineering solutions within the Databricks ecosystem. You'll learn best practices for data ingestion, ETL/ELT, data quality, and optimization, crucial for handling large datasets and complex analytical workloads. This certification not only validates your expertise in Databricks data engineering but also significantly enhances your marketability, paving the way for advanced roles in the rapidly evolving field of data.

Level	Intermediate
Vendor	Databricks

COURSE CURRICULUM

01 Module 1: Introduction to Databricks and the Lakehouse Platform

- Understanding the Data Lakehouse Architecture
- Databricks Workspace Navigation and Components
- Introduction to Apache Spark and its role in Databricks
- Setting up Databricks Clusters and Notebooks
- Fundamentals of Delta Lake

02 Module 2: Apache Spark for Data Engineering

- Spark Core Concepts: RDDs, DataFrames, Datasets
- Working with Spark SQL for Data Manipulation
- Spark Data Structures and Common Transformations
- Optimizing Spark Jobs for Performance
- Introduction to Spark APIs (Python/Scala)

03 Module 3: Delta Lake Fundamentals and Advanced Features

- Delta Lake as a Storage Layer for the Lakehouse
- ACID Transactions, Schema Enforcement, and Evolution
- Time Travel and Data Versioning
- Upserts and Change Data Capture (CDC) with Delta Lake
- Optimizing Delta Tables (Z-Ordering, Vacuum)

04 Module 4: Data Ingestion and ETL/ELT Pipelines

- Ingesting data from various sources (cloud storage, databases)
- Building multi-hop architecture (Bronze, Silver, Gold layers)
- Designing and implementing ETL/ELT workflows with Databricks
- Data Quality and Validation Techniques
- Using Databricks Autoloader for incremental data ingestion

Module 5: Real-time Data Processing with Structured Streaming

- 05**
- Introduction to Spark Structured Streaming
 - Processing real-time data from Kafka, Event Hubs, and other sources
 - Stateful and Stateless Stream Processing
 - Monitoring and troubleshooting streaming applications
 - Integrating Structured Streaming with Delta Lake

06 **Module 6: Deployment, Monitoring, and Governance**

- Scheduling and orchestrating Databricks jobs (Databricks Workflows, external orchestrators)
- Monitoring performance and cost with Databricks UI and external tools
- Securing data and workspaces (Unity Catalog, ACLs)
- Collaboration and CI/CD best practices for Databricks projects
- Cost management and optimization strategies

CAREER PATHWAYS

Graduates of this program are prepared for the following roles:

- > **Senior Databricks Data Engineer**
- > **Cloud Data Architect (Specializing in Azure/AWS Databricks)**
- > **ETL Developer (Lakehouse Platform Specialist)**

FREQUENTLY ASKED QUESTIONS

Q: What are the recommended technical prerequisites for enrolling in the Databricks Data Engineer course?

While there are no strict official prerequisites, candidates will benefit greatly from foundational knowledge in SQL, basic programming concepts (preferably Python or Scala), and familiarity with cloud computing basics (e.g., AWS S3, Azure Data Lake Storage). Prior experience with data warehousing or ETL concepts would also be advantageous, though not mandatory.

Q: How does this Databricks Data Engineer course prepare me for real-world data lakehouse projects compared to generic Spark or ETL training?

This course is specifically tailored to the Databricks Lakehouse Platform, integrating Spark, Delta Lake, and Structured Streaming into a unified architecture. Unlike generic training, it focuses on Databricks-specific tools and best practices for building end-to-end data pipelines, managing data quality, and optimizing performance within the Databricks ecosystem, which is critical for modern data lakehouse implementations.

Q: What specific Databricks-related skills will I gain that are most valuable for a hiring manager?

You will gain hands-on proficiency in designing and implementing medallion architectures (Bronze, Silver, Gold), utilizing Delta Live Tables for declarative pipeline development, optimizing Spark workloads on Databricks, performing advanced data transformations with Spark SQL and Python/Scala, and managing data governance and security through Unity Catalog. These are highly sought-after skills for roles requiring expertise in scalable, reliable data solutions on Databricks.

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