



MICROSOFT TRAINING

Azure Databricks Course

Master Azure Databricks to build and manage scalable data engineering and machine learning solutions for big data analytics.

DURATION

40 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 Azure Databricks course is meticulously designed for data professionals seeking to master the unified data analytics platform on Microsoft Azure. You will dive deep into leveraging Apache Spark's powerful processing capabilities within the Databricks environment, exploring core technologies like Delta Lake for reliable data lakes and MLflow for end-to-end machine learning lifecycle management. The curriculum is structured to provide a comprehensive understanding of how to build robust data pipelines, perform advanced data transformations, and implement scalable machine learning solutions.

Upon completion, participants will possess highly sought-after skills crucial for modern data engineering and data science roles. This course not only covers the theoretical underpinnings but also emphasizes practical, hands-on application, enabling you to confidently tackle complex big data challenges. You will learn to optimize performance, ensure data governance, and integrate Databricks with other Azure services, ultimately accelerating data initiatives and driving data-driven decisions within your organization.

Professional benefits include enhanced career prospects in rapidly evolving data domains, a deeper understanding of cloud-native analytics, and the ability to contribute significantly to projects involving large-scale data processing and machine learning operations. It's an ideal pathway for those looking to solidify their expertise and become pivotal players in the big data ecosystem.

Duration	40 hours
Level	Intermediate
Vendor	Microsoft

COURSE CURRICULUM

01 Module 1: Introduction to Azure Databricks and Apache Spark Fundamentals

- Understanding the Azure Databricks Platform and Architecture
- Databricks Workspace, Clusters, and Notebooks
- Introduction to Apache Spark: Core Concepts and Architecture
- Spark DataFrames and Spark SQL for Data Manipulation
- Working with Databricks Utilities and Connectors

02 Module 2: Data Engineering with Delta Lake on Databricks

- Data Ingestion Strategies into Databricks (ADLS Gen2, Blob Storage)
- Introduction to Delta Lake: ACID Properties, Schema Enforcement, Time Travel
- Building Bronze, Silver, and Gold Layer Data Lakes with Delta Lake
- Optimizing Delta Lake Tables for Performance (Z-Ordering, Compaction)
- Implementing Change Data Capture (CDC) with Delta Lake

03 Module 3: Advanced Data Processing and Structured Streaming

- Complex Data Transformations with Spark (Joins, Aggregations, Window Functions)
- Introduction to Structured Streaming for Real-time Data Processing
- Implementing Real-time ETL Pipelines with Databricks Structured Streaming
- Error Handling and Monitoring in Streaming Workloads
- Integrating Databricks with Azure Data Factory for Orchestration

04 Module 4: Machine Learning Lifecycle with Databricks and MLflow

- Preparing Data for Machine Learning on Databricks
- Model Training with Spark MLlib, scikit-learn, and Deep Learning Frameworks
- Introduction to MLflow: Experiment Tracking, Model Management, and Registries
- Hyperparameter Tuning and Model Evaluation on Databricks
- Model Deployment Strategies and MLOps Concepts with Databricks

05 Module 5: Performance Optimization, Security, and Governance

- Optimizing Spark Applications for Performance and Resource Utilization
- Troubleshooting Common Databricks Performance Issues
- Security Best Practices in Azure Databricks (Identity, Access Management)
- Data Governance and Compliance Considerations on Databricks
- Monitoring and Alerting for Databricks Workloads with Azure Monitor

CAREER PATHWAYS

Graduates of this program are prepared for the following roles:

- > **Azure Databricks Data Engineer**
- > **Cloud Machine Learning Engineer (Azure Focus)**
- > **Big Data Solutions Architect (with Databricks Specialization)**

FREQUENTLY ASKED QUESTIONS

Q: What specific hands-on data engineering skills will I acquire that are directly transferable to my job role using Azure Databricks?

This course focuses on practical application. You will gain expertise in building robust ETL/ELT pipelines using Delta Lake, performing advanced data transformations with Spark SQL and DataFrames, implementing real-time data ingestion with Structured Streaming, and optimizing Spark job performance. These skills are critical for data engineers tasked with managing and processing large datasets on Azure.

Q: What are the recommended prerequisites for this intermediate Azure Databricks course, especially regarding prior experience with Azure or Apache Spark?

While a foundational understanding of cloud concepts (preferably Azure), SQL, and at least one programming language (Python or Scala) is highly recommended, prior extensive Apache Spark experience is not strictly required. The course will cover Spark fundamentals in depth, but participants will benefit from basic familiarity with data processing concepts and scripting.

Q: How will this course help me in building and deploying machine learning models, and does it cover MLOps practices within the Azure ecosystem?

Yes, a dedicated module focuses on leveraging Databricks for the complete machine learning lifecycle. You will learn to prepare data for ML, train models using various libraries, and critically, utilize MLflow for experiment tracking, model versioning, and model registry. The course introduces key MLOps concepts, demonstrating how Databricks facilitates repeatable and manageable model deployment and governance within the Azure environment.

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