



MICROSOFT TRAINING

Azure Databricks Online

Master Azure Databricks for scalable data engineering and machine learning workflows in the cloud.

DURATION

16 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 Online" course empowers participants with the essential skills to leverage Microsoft Azure's powerful analytics platform for big data and artificial intelligence workloads. You will gain hands-on experience with Azure Databricks, a unified analytics platform built on Apache Spark, designed for data engineering, data science, and machine learning. The curriculum delves into key functionalities such as data ingestion, transformation, and analysis using Spark SQL, Python, and Scala within the Databricks environment. You will also explore the integration of Databricks with other Azure services for a comprehensive data solution.

By completing this 16-hour online course, you will be proficient in building scalable data pipelines, performing advanced analytics, and developing machine learning models on vast datasets. The course emphasizes practical application, ensuring you can effectively utilize core technologies like Delta Lake for reliable data lakes and MLflow for managing the machine learning lifecycle. This expertise will significantly enhance your ability to contribute to data-driven decision-making and innovation within any organization, preparing you for high-demand roles in the rapidly evolving cloud analytics landscape.

Duration	16 hours
Level	Intermediate
Vendor	Microsoft

COURSE CURRICULUM

01 Module 1: Introduction to Azure Databricks and Apache Spark

- Overview of Azure Databricks architecture and components
- Understanding Apache Spark fundamentals and its ecosystem
- Databricks workspace navigation and cluster management
- Introduction to notebooks, languages (Python, SQL, Scala, R), and commands
- Connecting Azure Databricks to Azure Data Lake Storage and Blob Storage

02 Module 2: Data Engineering with Delta Lake on Databricks

- Ingesting data into Databricks using various sources (streaming, batch)
- Introduction to Delta Lake: ACID transactions, schema enforcement, time travel
- Building reliable and performant data lakes with Delta Lake
- Implementing ETL/ELT processes using Spark SQL and DataFrames
- Optimizing Delta Lake tables for query performance

03 Module 3: Advanced Data Processing and Analytics

- Advanced Spark DataFrame operations for data manipulation and transformation
- Working with structured, semi-structured, and unstructured data
- Implementing user-defined functions (UDFs) in Spark
- Performing complex analytics with Spark SQL and PySpark
- Introduction to structured streaming for real-time data processing

04 Module 4: Machine Learning on Azure Databricks with MLflow

- Overview of machine learning concepts and lifecycle
- Preparing data for machine learning using Databricks
- Building and training machine learning models using MLlib and popular libraries (Scikit-learn, TensorFlow, PyTorch)
- Managing the machine learning lifecycle with MLflow: tracking experiments, models, and artifacts

- Deploying and serving machine learning models with Databricks and Azure services

05 Module 5: Performance Tuning, Security, and Best Practices

- Optimizing Spark jobs for performance and cost efficiency
- Troubleshooting common Databricks issues
- Implementing security best practices within Azure Databricks
- Monitoring Databricks workloads and clusters
- Integrating Databricks with Azure DevOps for CI/CD

CAREER PATHWAYS

Graduates of this program are prepared for the following roles:

- > **Azure Data Engineer specializing in scalable ETL/ELT pipelines**
- > **Machine Learning Engineer building and deploying models on Databricks**
- > **Big Data Architect designing unified analytics solutions on Azure**

FREQUENTLY ASKED QUESTIONS

Q: What are the recommended prerequisites for effectively engaging with this Azure Databricks Online course?

Participants should have a foundational understanding of cloud computing concepts, preferably with some exposure to Azure services. Basic programming knowledge in Python or SQL is highly recommended, as the course involves hands-on coding for data manipulation and analysis within the Databricks environment. Familiarity with big data concepts or Apache Spark is beneficial but not strictly required, as key concepts will be introduced.

Q: How will this course specifically prepare me for building real-world data engineering and machine learning solutions on Azure?

This course is designed with a practical, hands-on approach, focusing on real-world scenarios. You will learn to ingest and transform diverse datasets using Delta Lake, build scalable data pipelines, and develop and manage machine learning models using MLflow. By working through practical labs and exercises, you'll gain the direct experience needed to design, implement, and optimize data engineering workflows and deploy robust ML solutions on the Azure Databricks platform, making you job-ready for relevant industry roles.

Q: Will this course cover advanced features like Delta Lake optimization techniques or MLflow model deployment strategies?

Yes, this intermediate course delves into crucial advanced features. You will learn about Delta Lake optimization techniques, including Z-ordering and compaction, to enhance query performance and storage efficiency. Furthermore, the course covers MLflow's capabilities beyond basic experiment tracking, exploring model versioning, registry, and strategies for deploying registered models for inference, providing a comprehensive understanding of the ML lifecycle 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