# Implement data engineering solutions using Azure Databricks (DP-750T00)

Duration: 4 days
Audience: Analyste de données

## Who is this training for?

The target audience is data engineers who have fundamental knowledge of data analytics concepts, a basic understanding of cloud storage, and familiarity with data organization principles. They should be comfortable working with SQL and have experience using Python, including notebooks, for data engineering tasks. Learners are expected to have a good understanding of Azure Databricks workspaces and Unity Catalog, along with familiarity with data access patterns and core data engineering and data warehouse concepts. In addition, they should have foundational knowledge of Azure security, including Microsoft Entra ID, and be familiar with Git version control fundamentals.

## Training objectives

By the end of this course, students will be able to:

- Provision and configure an Azure Databricks workspace and understand its architectural components.
- Select and configure appropriate compute resources for different workload types.
- Create and organize Unity Catalog objects including catalogs, schemas, tables, views, and volumes.
- Implement security controls using access control lists, row filtering, column masking, and Azure Key Vault.
- Apply governance practices including data lineage, audit logging, tags, and Delta Sharing.
- Design and implement data models using Delta Lake with appropriate partitioning and clustering strategies.
- Ingest data using multiple patterns: batch, streaming, Auto Loader, COPY INTO, and Lakeflow Connect.
- Cleanse and transform raw data using SQL and PySpark, including joins, pivots, and merge operations.

Copyright © 2026 AFI by Edgenda inc.

- Implement data quality constraints using schema enforcement and Lakeflow Spark Declarative Pipeline expectations.
- Design and implement multi task data pipelines with error handling and retry logic. Create and schedule Lakeflow Jobs with triggers, alerts, and automated restarts.
- Apply development lifecycle practices using Git folders and Declarative Automation Bundles.
- Monitor, troubleshoot, and optimize Spark workloads using the Spark UI and Azure Log Analytics.

## Summary

DP-750T00: Implement Data Engineering Solutions Using Azure Databricks is a 4-day instructor led training (ILT) course that prepares data engineers to design, build, and maintain scalable data engineering solutions on Azure Databricks. The course covers the full lifecycle of a data engineering workload—from setting up a workspace and configuring compute, to ingesting and transforming data, enforcing governance, and deploying production pipelines.

## Course outline

**Learning Path 1 — Set up and configure an Azure Databricks environment**

- Module 1 — Explore Azure Databricks
- Module 2 — Select and configure compute in Azure Databricks
- Module 3 — Create and organize objects in Unity Catalog

**Learning Path 2 — Secure and govern Unity Catalog objects in Azure Databricks**

- Module 1 — Secure Unity Catalog objects
- Module 2 — Govern Unity Catalog objects

**Learning Path 3 — Prepare and process data with Azure Databricks**

- Module 1 — Design and implement data modeling with Azure Databricks
-

Copyright © 2026 AFI by Edgenda inc.

Module 2 — Ingest data into Unity Catalog

- Module 3 — Cleanse, transform, and load data into Unity Catalog
- Module 4 — Implement and manage data quality constraints with Azure Databricks

**Learning Path 4 — Deploy and maintain data pipelines and workloads with Azure Databricks**

- Module 1 — Design and implement data pipelines with Azure Databricks
- Module 2 — Implement Lakeflow Jobs with Azure Databricks
- Module 3 — Implement development lifecycle processes in Azure Databricks
- Module 4 — Monitor, troubleshoot, and optimize workloads in Azure Databricks

## Approach and methodology

Practical and structured approach combining focused theory with guided hands on labs. Participants progressively build data engineering solutions using Azure Databricks through real world, scenario based exercises that promote immediate application of learning.

They learn how to provision and configure Azure Databricks workspaces, organize and secure data with Unity Catalog, design Delta Lake data models, ingest and transform data using SQL and PySpark, and implement data quality constraints. The course emphasizes modern data engineering practices including Lakeflow pipelines, Lakeflow Jobs, Git based development workflows, and performance optimization using the Spark UI and Azure Log Analytics.

Led by a Microsoft certified trainer (MCT), the course encourages interactivity and the development of directly transferable technical skills that enable participants to confidently design, build, secure, and maintain production grade data engineering workloads on Azure Databricks.

Copyright © 2026 AFI by Edgenda inc.

## Prerequisites

Students should have the following knowledge and experience before attending this course:

- Fundamental knowledge of data analytics and data warehousing concepts.
- Basic understanding of cloud storage and Azure resource management.
- Familiarity with SQL for querying and managing data.
- Basic understanding of Python programming (used in notebooks and PySpark).
- Understanding of Git and version control fundamentals.
- Knowledge of Microsoft Entra ID and Azure security basics (recommended for LP 2).

Copyright © 2026 AFI by Edgenda inc.