Implement a data engineering solution with Azure Databricks (DP-3027)

Description

Master data processing with Azure Databricks

Azure Databricks is becoming a key pillar in the world of modern data engineering. Thanks to its power, flexibility, and integrated tools, this platform allows you to process large volumes of data efficiently. This DP-3027 training offers complete mastery of cloud data pipelines through hands-on cases and progressive modules.

A practical course for today's data engineers

This course has been designed to meet the real needs of data professionals. You will learn how to implement robust engineering solutions, automate workflows, and ensure data quality in production. The modules cover the entire pipeline lifecycle, from real-time processing with Spark Structured Streaming to implementing CI/CD workflows, including automation through Azure Data Factory.

Course Content
Module 1: Perform incremental processing with Spark Structured Streaming

- Set up real-time data sources for incremental processing
- Optimize Delta Lake for incremental processing in Azure Databricks
- Handle late data and out-of-order events in incremental processing
- Monitoring and performance tuning strategies for incremental processing in Azure Databricks

Module 2: Implement streaming architecture patterns with Delta Live Tables

- Event driven architectures with Delta Live Tables
- Ingest data with structured streaming
- Maintain data consistency and reliability with structured streaming
- Scale streaming workloads with Delta Live Tables

Module 3: Optimize performance with Spark and Delta Live Tables

- Optimize performance with Spark and Delta Live Tables
- Perform cost-based optimization and query tuning
- Use change data capture (CDC)
- Use enhanced autoscaling
- Implement observability and data quality metrics

Module 4: Implement CI/CD workflows in Azure Databricks

- Implement version control and Git integration
- Perform unit testing and integration testing
- Manage and configure your environment
- Implement rollback and roll-forward strategies

Module 5: Automate workloads with Azure Databricks Jobs

- Implement job scheduling and automation
- Optimize workflows with parameters
- Handle dependency management
- Implement error handling and retry mechanisms
- Explore best practices and guidelines

Module 6: Manage data privacy and governance with Azure Databricks

- Implement data encryption techniques in Azure Databricks
- Manage access controls in Azure Databricks
- Implement data masking and anonymization in Azure Databricks
- Use compliance frameworks and secure data sharing in Azure Databricks
- Use data lineage and metadata management
- Implement governance automation in Azure Databricks

Module 7: Use SQL Warehouses in Azure Databricks

- Get started with SQL Warehouses
- Create databases and tables
- Create queries and dashboards

Module 8: Run Azure Databricks Notebooks with Azure Data Factory

- Understand Azure Databricks notebooks and pipelines
- Create a linked service for Azure Databricks
- Use a Notebook activity in a pipeline
- Use parameters in a notebook

Lab / Exercises

- This course provides you with exclusive access to the official Microsoft lab, enabling you to practice your skills in a professional environment.

Documentation

- Access to Microsoft Learn, Microsoft's online learning platform, offering interactive resources and educational content to deepen your knowledge and develop your technical skills.

Participant profiles

- Data Engineer
- Big Data Developer
- Cloud Architect
- Data Consultant
- Azure Administrator
- BI Manager

Prerequisites

- Know how to use a computer
- Understand the basics of cloud computing

Objectives

- Create real-time data streams
- Build architectures with Delta Live Tables
- Improve data processing performance
- Deploy CI/CD workflows in Databricks
- Automate tasks with Azure Data Factory
- Secure and govern data
- Use SQL warehouses in Databricks
- Run notebooks from Azure Data Factory

Description
Implement a data engineering solution with Azure Databricks (DP-3027)
Niveau
Intermédiaire
Classroom Registration Price (CHF)
900
Virtual Classroom Registration Price (CHF)
850
Duration (in Days)
1
Reference
DP-3027