Implementing a Lakehouse with Microsoft Fabric

Duration: *1 Day(s)*

Course Overview

This course is designed to build your foundational skills in data engineering on Microsoft Fabric, focusing on the Lakehouse concept. This course will explore the powerful capabilities of Apache Spark for distributed data processing and the essential techniques for efficient data management, versioning, and reliability by working with Delta Lake tables. This course will also explore data ingestion and orchestration using Dataflows Gen2 and Data Factory pipelines. This course includes a combination of lectures and hands-on exercises that will prepare you to work with lakehouses in Microsoft Fabric.

Review this course online at https://www.alta3.com/courses/DP-601T00

Objectives

- Design robust lakehouses using Microsoft Fabric
- Implement data processing with Apache Spark
- Manage data with Delta Lake tables
- Optimize data ingestion using Dataflows Gen2 and pipelines

Who Should Attend

- Data Engineers
- Data Analysts
- Business Intelligence Professionals
- Data Architects

Prerequisites

Course Outline

Introduction to end-to-end analytics using Microsoft Fabric

1. Explore end-to-end analytics with Microsoft Fabric
2. Data teams and Microsoft Fabric
3. Enable and use Microsoft Fabric

Get started with lakehouses in Microsoft Fabric

4. Explore the Microsoft Fabric Lakehouse
5. Work with Microsoft Fabric Lakehouses
6. Explore and transform data in a lakehouse
7. Create and ingest data with a Microsoft Fabric Lakehouse

Use Apache Spark in Microsoft Fabric

8. Prepare to use Apache Spark
9. Run Spark code
10. Work with data in a Spark dataframe
11. Work with data using Spark SQL
12. Visualize data in a Spark notebook
13. Analyze data with Apache Spark

Work with Delta Lake tables in Microsoft Fabric

14. Understand Delta Lake
15. Create delta tables
16. Work with delta tables in Spark
17. Use delta tables with streaming data
18. Use delta tables in Apache Spark

Ingest Data with Dataflows Gen2 in Microsoft Fabric

19. Understand Dataflows (Gen2) in Microsoft Fabric
20. Explore Dataflows (Gen2) in Microsoft Fabric
21. Integrate Dataflows (Gen2) and Pipelines in Microsoft Fabric
22. Create and use a Dataflow (Gen2) in Microsoft Fabric

Use Data Factory pipelines in Microsoft Fabric

23. Understand pipelines
24. Use the Copy Data activity
25. Use pipeline templates
26. Run and monitor pipelines
27. Ingest data with a pipeline