Prepare and visualize data with Microsoft Power BI

Course DP-601 1 Days Instructor-led, Hands-on

Introduction

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.

Who Should Attend?

The primary audience for this course is data professionals who are familiar with data modeling, extraction, and analytics. It is designed for professionals who are interested in gaining knowledge about Lakehouse architecture, the Microsoft Fabric platform, and how to enable end-to-end analytics using these technologies.

Prerequisites

Students should be familiar with basic data concepts and terminology.

Outline

Module 1: Introduction to end-to-end analytics using Microsoft Fabric

- Explore end-to-end analytics with Microsoft Fabric
- Explore data teams and Microsoft Fabric
- Enable and use Microsoft Fabric
- Module assessment

Module 2: Get started with lakehouses in Microsoft Fabric

- Describe lakehouse features and capabilities
- Ingest and transform data in a lakehouse
- Query and analyze lakehouse data
- Module assessment

Module 3: Use Apache Spark in Microsoft Fabric

- Prepare to use Apache Spark
- Run Spark code
- Work with data in a Spark dataframe
- Work with data using Spark SQL
- Visualize data in a Spark notebook
- Module assessment

Module 4: Work with Delta Lake tables in Microsoft Fabric

- Understand Delta Lake
- Create delta tables
- Optimize delta tables
- Work with delta tables in Spark
- Use delta tables with streaming data
- Module assessment

Module 5: Ingest Data with Dataflows Gen2 in Microsoft Fabric

- Understand Dataflows Gen2 in Microsoft Fabric
- Explore Dataflows Gen2 in Microsoft Fabric
- Integrate Dataflows Gen2 and Pipelines in Microsoft Fabric
- Module assessment

Module 6: Orchestrate processes and data movement with Microsoft Fabric

- Understand pipelines
- Use the Copy Data activity
- Use pipeline templates
- Run and monitor pipelines
- Module assessment

Module 7: Organize a Fabric lakehouse using medallion architecture design

- Describe medallion architecture
- Plan a medallion architecture in Fabric
- Query and report on data in your Fabric lakehouse
- Secure and govern your medallion lakehouse
- Module assessment