SQL | MASTER ACADEMY
RAISE AN EXPERT

COURSE: DP-601: Implementing a Lakehouse with Microsoft Fabric

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.

Summary

Duration:	*1 day (all day)*
Level:	*200*
Delivery method:	*Virtual Instructor-led class*
Language:	*English or Bulgarian*

** The difficulty level is consistent with the widely accepted scale of technical difficulty of training on Microsoft Corp*

AUDIENCE:

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.

Job role: Data Analyst, Data Engineer, Data Scientist.

PREREQUISITES:

You should be familiar with basic data concepts and terminology.

TOPICS:

Introduction to end-to-end analytics using Microsoft Fabric

Discover how Microsoft Fabric can meet your enterprise's analytics needs in one platform. Learn about Microsoft Fabric, how it works, and identify how you can use it for your analytics needs.

Get started with lakehouses in Microsoft Fabric

In this module, you'll learn how to:

- Describe core features and capabilities of lakehouses in Microsoft Fabric
- Create a lakehouse
- Ingest data into files and tables in a lakehouse
- Query lakehouse tables with SQL

Use Apache Spark in Microsoft Fabric

In this module, you'll learn how to:

- Configure Spark in a Microsoft Fabric workspace
- Identify suitable scenarios for Spark notebooks and Spark jobs
- Use Spark dataframes to analyze and transform data
- Use Spark SQL to query data in tables and views
- Visualize data in a Spark notebook

Work with Delta Lake tables in Microsoft Fabric

In this module, you'll learn how to:

- Understand Delta Lake and delta tables in Microsoft Fabric
- Create and manage delta tables using Spark
- Use Spark to query and transform data in delta tables
- Use delta tables with Spark structured streaming

Ingest Data with Dataflows Gen2 in Microsoft Fabric

In this module, you'll learn how to:

- Describe Dataflow (Gen2) capabilities in Microsoft Fabric
- Create Dataflow (Gen2) solutions to ingest and transform data
- Include a Dataflow (Gen2) in a pipeline