

**Implementing a Lakehouse with Microsoft Fabric (DP-601)**

**Course Number:** MOC-DP-601
**Duration:** 1 day

**Overview**

This Microsoft official course, Implementing a Lakehouse with Microsoft Fabric training (DP-601), teaches attendees how to leverage Apache Spark for distributed data processing and efficient data management, versioning, and reliability by working with Delta Lake tables. This course explores data ingestion and orchestration using Dataflows Gen2 and Data Factory pipelines.

**Prerequisites**

All students should be familiar with basic data concepts and terminology.

**Materials**

All Microsoft training students receive Microsoft official courseware.

For all Microsoft Official Courses taught in their entirety that have a corresponding certification exam, an exam voucher is included for each participant.

**Software Needed on Each Student PC**

Attendees will not need to install any software on their computer for this class. The class will be conducted in a remote environment that Accelebrate will provide; students will only need a local computer with a web browser and a stable Internet connection. Any recent version of Microsoft Edge, Mozilla Firefox, or Google Chrome will be fine.

**Objectives**

* Describe end-to-end analytics in Microsoft Fabric
* 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
* 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
* 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
* Describe Dataflow (Gen2) capabilities in Microsoft Fabric
* Create Dataflow (Gen2) solutions to ingest and transform data
* Include a Dataflow (Gen2) in a pipeline
* Describe pipeline capabilities in Microsoft Fabric
* Use the Copy Data activity in a pipeline
* Create pipelines based on predefined templates
* Run and monitor pipelines

**Outline**

* Introduction to end-to-end analytics using Microsoft Fabric
	+ Explore end-to-end analytics with Microsoft Fabric
	+ Data teams and Microsoft Fabric
	+ Enable and use Microsoft Fabric
* Get started with lakehouses in Microsoft Fabric
	+ Explore the Microsoft Fabric Lakehouse
	+ Work with Microsoft Fabric Lakehouses
	+ Explore and transform data in a lakehouse
	+ Exercise - Create and ingest data with a Microsoft Fabric Lakehouse
* 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
	+ Exercise - Analyze data with Apache Spark
* Work with Delta Lake tables in Microsoft Fabric
	+ Understand Delta Lake
	+ Create delta tables
	+ Work with delta tables in Spark
	+ Use delta tables with streaming data
	+ Exercise - Use delta tables in Apache Spark
* 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
	+ Exercise - Create and use a Dataflow (Gen2) in Microsoft Fabric
* Use Data Factory pipelines in Microsoft Fabric
	+ Understand pipelines
	+ Use the Copy Data activity
	+ Use pipeline templates
	+ Run and monitor pipelines
	+ Exercise - Ingest data with a pipeline
* Conclusion