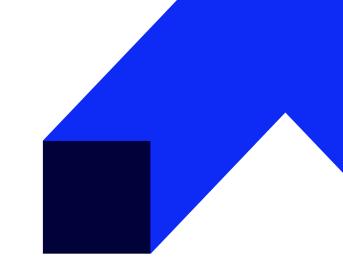# Implementing a Lakehouse with Microsoft Fabric

The Implementing a Lakehouse with Microsoft Fabric training introduces participants to the concept of the Lakehouse – a modern approach to data management and analytics that combines the flexibility of data lakes with the capabilities of a traditional data warehouse.
During the course, participants learn about the capabilities of Microsoft Fabric for acquiring, processing, and organizing data using Apache Spark, Delta Lake, Dataflows Gen2, and Data Factory. The training focuses on practical aspects of data engineering, such as versioning, dataflow orchestration, and building scalable analytical environments.

Training recipients

The training is intended for data professionals who want to explore the capabilities of Microsoft Fabric in data processing and integration within the Lakehouse architecture.
It is particularly dedicated to:

- data engineers,
- data analysts working with large datasets,
- Power BI specialists who want to expand their skills with data engineering capabilities,
- data and analytics architects,
- members of teams designing Lakehouse environments in the Microsoft cloud.

Benefits

- Understanding the Lakehouse concept – you will learn the structure and use cases of the Lakehouse architecture in the Microsoft Fabric environment.
- Using Apache Spark – you will learn how to process, transform, and analyze distributed data using

Spark DataFrames and Spark SQL.

- Working with Delta Lake – you will learn how to create and manage Delta Lake tables, use data versioning, and work with the "time travel" feature.
- Data ingestion and integration – you will master the use of Dataflows Gen2 and Data Factory pipelines for ingesting and processing data.
- Building modern analytical solutions – you will learn how to create integrated, high-performance data environments in Microsoft Fabric, ready for further analysis in Power BI or other tools.

Training program

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

- Introduction to Microsoft Fabric
- Data teams and Fabric
- Enable and use Microsoft Fabric

Module 2: Get started with lakehouses in Microsoft Fabric

- What is a Lakehouse
- Work with a Fabric Lakehouse
- Explore, transform and visualize data in the Lakehouse

Module 3: Use Apache Spark in Microsoft Fabric

- Prepare to use Apache Spark
- Run Spark in Fabric
- Load data in a Spark DataFrame
- Transform data in a Spark DataFrame
- Partition the output file
- Work with data using Spark SQL
- Query Data using Spark SQL API
- Visualize Data

Module 4: Work with Delta Lake tables in Microsoft Fabric

- Understand Delta Lake
- Create delta tables using code in Spark
- Managed vs External Tables
- Work with delta tables in Spark
- Data versioning and Time Travel
- Use delta tables with Streaming data

Module 5: Ingest Data with Dataflows Gen2 in Microsoft Fabric

- Understand Dataflows (Gen2)
- Dataflow (Gen2) benefits and limitations
- Explore Dataflows (Gen2) in Microsoft Fabric
- Integrate Dataflows (Gen2) and Pipelines in Microsoft Fabric

Module 6: Use Data Factory pipelines in Microsoft Fabric

- Pipelines in Microsoft Fabric
- Common Activities – Copy Data
- Common Activities – pipeline templates
- Run and monitor pipelines

Expected preparation of the participant

You should be familiar with basic data concepts and terminology. It is suggested that you familiarize yourself with the AZ-900 and DP-900 training materials in advance, or attend these trainings.

To increase the comfort of work and training's effectiveness we suggest using an additional monitor. The lack of additional monitor does not exclude participation in the training, however, it significantly influences the comfort of work during classes.

Training Includes

- manual in electronic form available on the platform:
- https://learn.microsoft.com/pl-pl/training/
- access to Altkom Akademia's student portal

This course includes a combination of lectures and hands-on exercises that will prepare you to work with lakehouses in Microsoft Fabric.

Duration

1 days / 7 hours

Language

- **Training**: English
- **Materials**: English