

MS-DP601TOO: IMPLEMENTING A LAKEHOUSE WITH MICROSOFT FABRIC



DURATION	LEVEL	TECHNOLOGY	DELIVERY METHOD	TRAINING CREDITS
1 Day	Intermediate	Azure Data & AI	Instructor-led	NA

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.

AUDIENCE PROFILE

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

You should be familiar with basic data concepts and terminology.

COURSE OBJECTIVES

After completing this course, students will be able to:

- Describe end-to-end analytics in Microsoft Fabric
- Describe core features and capabilities of lakehouses in Microsoft Fabric
- Identify suitable scenarios for Spark notebooks and Spark jobs
- Understand Delta Lake and delta tables in Microsoft Fabric
- Describe pipeline capabilities in Microsoft Fabric

COURSE CONTENT

Module 1: 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.

Lessons

- Introduction
- Explore end-to-end analytics with Microsoft Fabric
- Data teams and Microsoft Fabric
- Enable and use Microsoft Fabric
- Knowledge Check
- Summary

Learning objectives

In this module, you'll learn how to:

- Describe end-to-end analytics in Microsoft Fabric

Module 2: Get started with lakehouses in Microsoft Fabric

Lakehouses merge data lake storage flexibility with data warehouse analytics. Microsoft Fabric offers a lakehouse solution for comprehensive analytics on a single SaaS platform.

Lessons

- Introduction
- 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
- Knowledge check
- Summary

Learning objectives

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

Module 3: Use Apache Spark in Microsoft Fabric

Apache Spark is a core technology for large-scale data analytics. Microsoft Fabric provides support for Spark clusters, enabling you to analyze and process data in a Lakehouse at scale.

Lessons

- Introduction
- Prepare to use Apache Spark
- Run Spark code
- Connect to data sources and ingest data
- 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
- Knowledge check
- Summary

Learning objectives

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

Module 4: Work with Delta Lake tables in Microsoft Fabric

Tables in a Microsoft Fabric lakehouse are based on the Delta Lake storage format commonly used in Apache Spark. By using the enhanced capabilities of delta tables, you can create advanced analytics solutions.

Lessons

- Introduction
- Understand Delta Lake
- Create delta tables
- Create delta tables
- Work with delta tables in Spark
- Use delta tables with streaming data
- Exercise - Use delta tables in Apache Spark
- Knowledge check
- Summary

Learning objectives

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

Module 5: Ingest Data with Dataflows Gen2 in Microsoft Fabric

Data ingestion is crucial in analytics. Microsoft Fabric's Data Factory offers Dataflows (Gen2) for visually creating multi-step data ingestion and transformation using Power Query Online.

Lessons

- Introduction
- 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
- Knowledge check
- Summary

Learning objectives

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

Module 6: Use Data Factory pipelines in Microsoft Fabric

Microsoft Fabric includes Data Factory capabilities, including the ability to create pipelines that orchestrate data ingestion and transformation tasks.

Lessons

- Introduction
- Understand pipelines
- Use the Copy Data activity
- Use pipeline templates
- Run and monitor pipelines
- Exercise - Ingest data with a pipeline
- Knowledge check
- Summary

Learning objectives

In this module, you'll learn how to:

- Describe pipeline capabilities in Microsoft Fabric
- Use the Copy Data activity in a pipeline
- Create pipelines based on predefined templates
- Run and monitor pipelines

Module 7: Organize a Fabric lakehouse using medallion architecture design

Explore the potential of the medallion architecture design in Microsoft Fabric. Organize and transform your data across bronze, silver, and gold layers of a lakehouse for optimized analytics.

Lessons

- Introduction
- Describe medallion architecture
- Implement a medallion architecture in Fabric
- Query and report on data in your Fabric lakehouse
- Considerations for managing your lakehouse
- Exercise - Organize your Fabric lakehouse using a medallion architecture
- Knowledge check
- Summary

Learning objectives

- Describe the principles of using the medallion architecture in data management.
- Apply the medallion architecture framework within the Microsoft Fabric environment.
- Analyze data stored in the governance of data stored in the medallion architecture.

ASSOCIATED CERTIFICATIONS & EXAM

There is no Associated certification & Exam for this course, however, there is an assessment to achieve your Applied Skills credential. ([Assessment Link](#))