

Orchestrate processes and data movement with Microsoft Fabric

1. Introduction

<https://learn.microsoft.com/en-us/training/modules/use-data-factory-pipelines-fabric/1-introduction>

Introduction

Completed

Data pipelines define a sequence of activities that orchestrate an overall process, usually by extracting data from one or more sources and loading it into a destination; often transforming it along the way. Pipelines are commonly used to automate *extract*, *transform*, and *load* (ETL) processes that ingest transactional data from operational data stores into an analytical data store, such as a lakehouse, data warehouse, or SQL database.

If you're already familiar with Azure Data Factory, then data pipelines in Microsoft Fabric will be immediately familiar. They use the same architecture of connected activities to define a process that can include multiple kinds of data processing tasks and control flow logic. You can run pipelines interactively in the Microsoft Fabric user interface, or schedule them to run automatically.

2. Understand pipelines

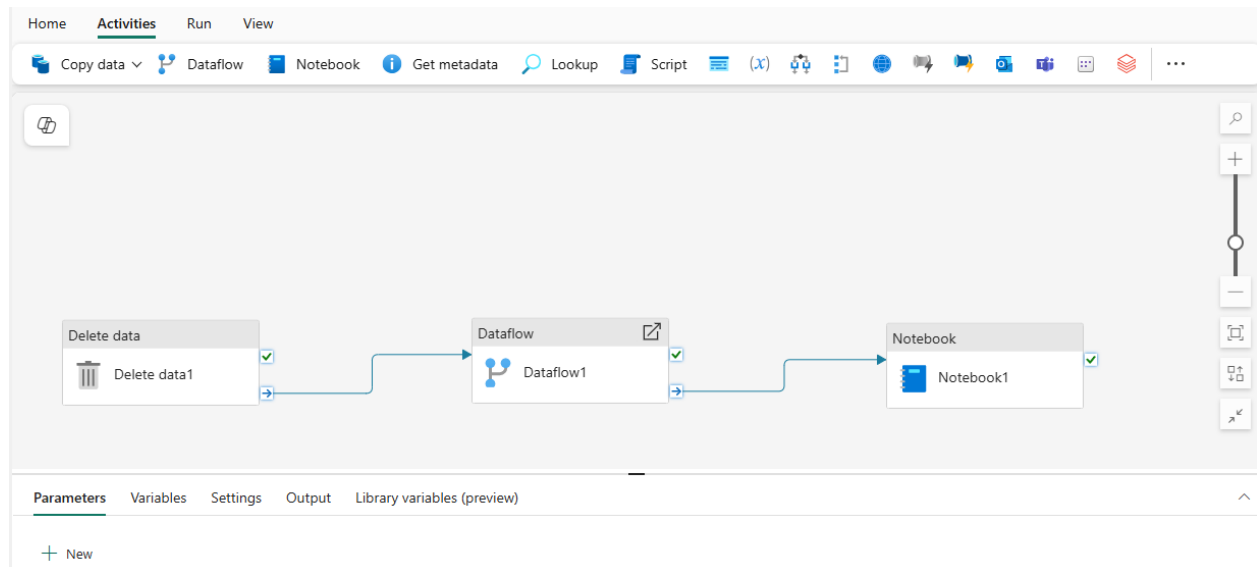
<https://learn.microsoft.com/en-us/training/modules/use-data-factory-pipelines-fabric/2-understand-fabric-pipeline>

Understand pipelines

Completed

Pipelines in Microsoft Fabric encapsulate a sequence of *activities* that perform data movement and processing tasks. You can use a pipeline to define data transfer and transformation activities, and

orchestrate these activities through control flow activities that manage branching, looping, and other typical processing logic. The graphical pipeline *canvas* in the Fabric user interface enables you to build complex pipelines with minimal or no coding required.



Core pipeline concepts

Before building pipelines in Microsoft Fabric, you should understand a few core concepts.

Activities

Activities are the executable tasks in a pipeline. You can define a flow of activities by connecting them in a sequence. The outcome of a particular activity (success, failure, or completion) can be used to direct the flow to the next activity in the sequence.

There are two broad categories of activity in a pipeline.

- **Data transformation activities** - activities that encapsulate data transfer operations, including simple **Copy Data** activities that extract data from a source and load it to a destination, and more complex **Data Flow** activities that encapsulate dataflows (Gen2) that apply transformations to the data as it is transferred. Other data transformation activities include **Notebook** activities to run a Spark notebook, **Stored procedure** activities to run SQL code, **Delete data** activities to delete existing data, and others. In OneLake, you can configure the destination to a lakehouse, warehouse, SQL database, or other options.
- **Control flow activities** - activities that you can use to implement loops, conditional branching, or manage variable and parameter values. The wide range of control flow activities enables you to implement complex pipeline logic to orchestrate data ingestion and transformation flow.

Tip

For details about the complete set of pipeline activities available in Microsoft Fabric, see [Activity overview](#) in the Microsoft Fabric documentation.

Parameters

Pipelines can be parameterized, enabling you to provide specific values to be used each time a pipeline is run. For example, you might want to use a pipeline to save ingested data in a folder, but have the flexibility to specify a folder name each time the pipeline is run.

Using parameters increases the reusability of your pipelines, enabling you to create flexible data ingestion and transformation processes.

Pipeline runs

Each time a pipeline is executed, a *data pipeline run* is initiated. Runs can be initiated on-demand in the Fabric user interface or scheduled to start at a specific frequency. Use the unique run ID to review run details to confirm they completed successfully and investigate the specific settings used for each execution.

3. Use the Copy Data activity

<https://learn.microsoft.com/en-us/training/modules/use-data-factory-pipelines-fabric/3-copy-data>

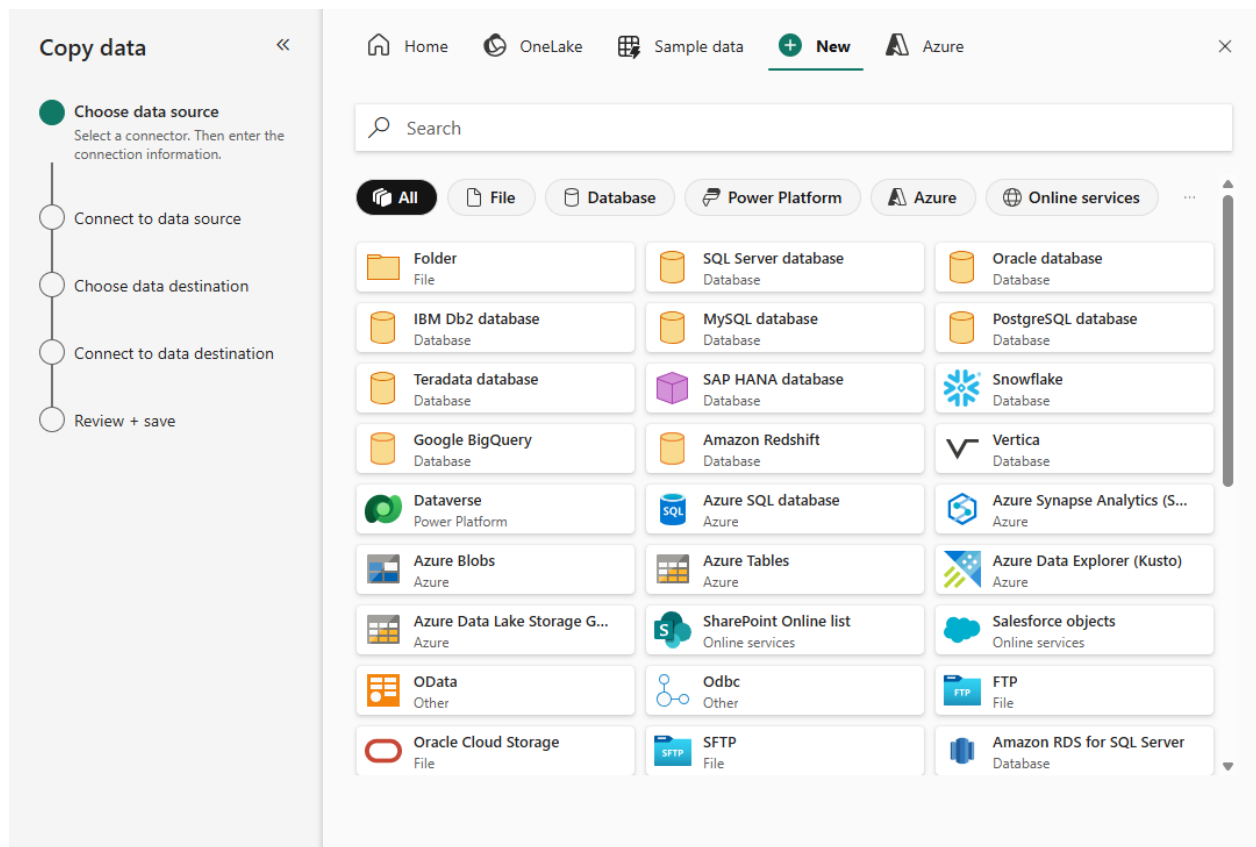
Use the Copy Data activity

Completed

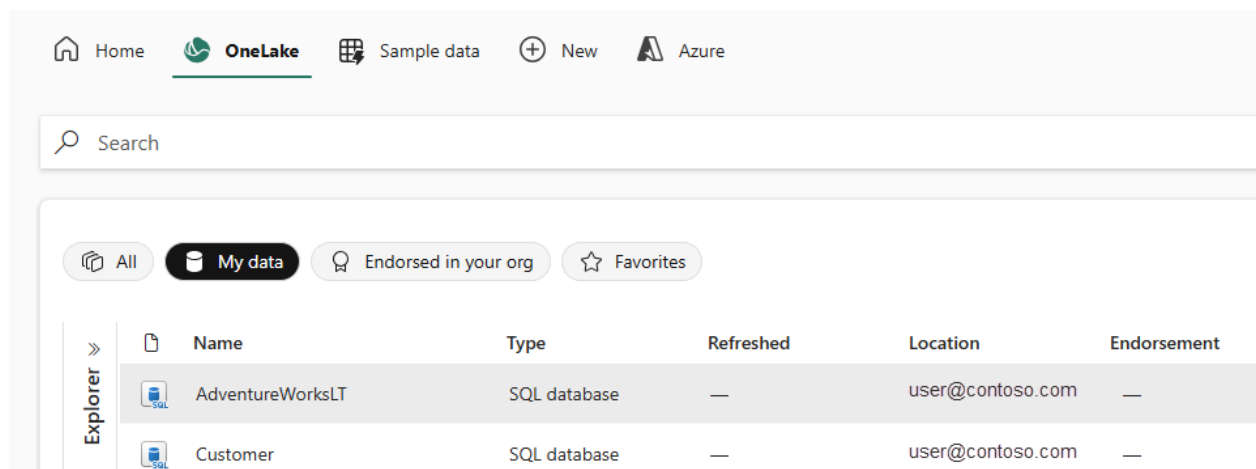
The **Copy Data** activity is one of the most common uses of a data pipeline. Many pipelines consist of a single **Copy Data** activity that is used to ingest data from an external source into a lakehouse file or table.

You can also combine the **Copy Data** activity with other activities to create a repeatable data ingestion process - for example by using a **Delete data** activity to remove existing data, a **Copy Data** activity to replace the deleted data with a file containing data from an external source, and a **Notebook** activity to run Spark code that transforms the data in the file and loads it into a table.

The Copy Data tool

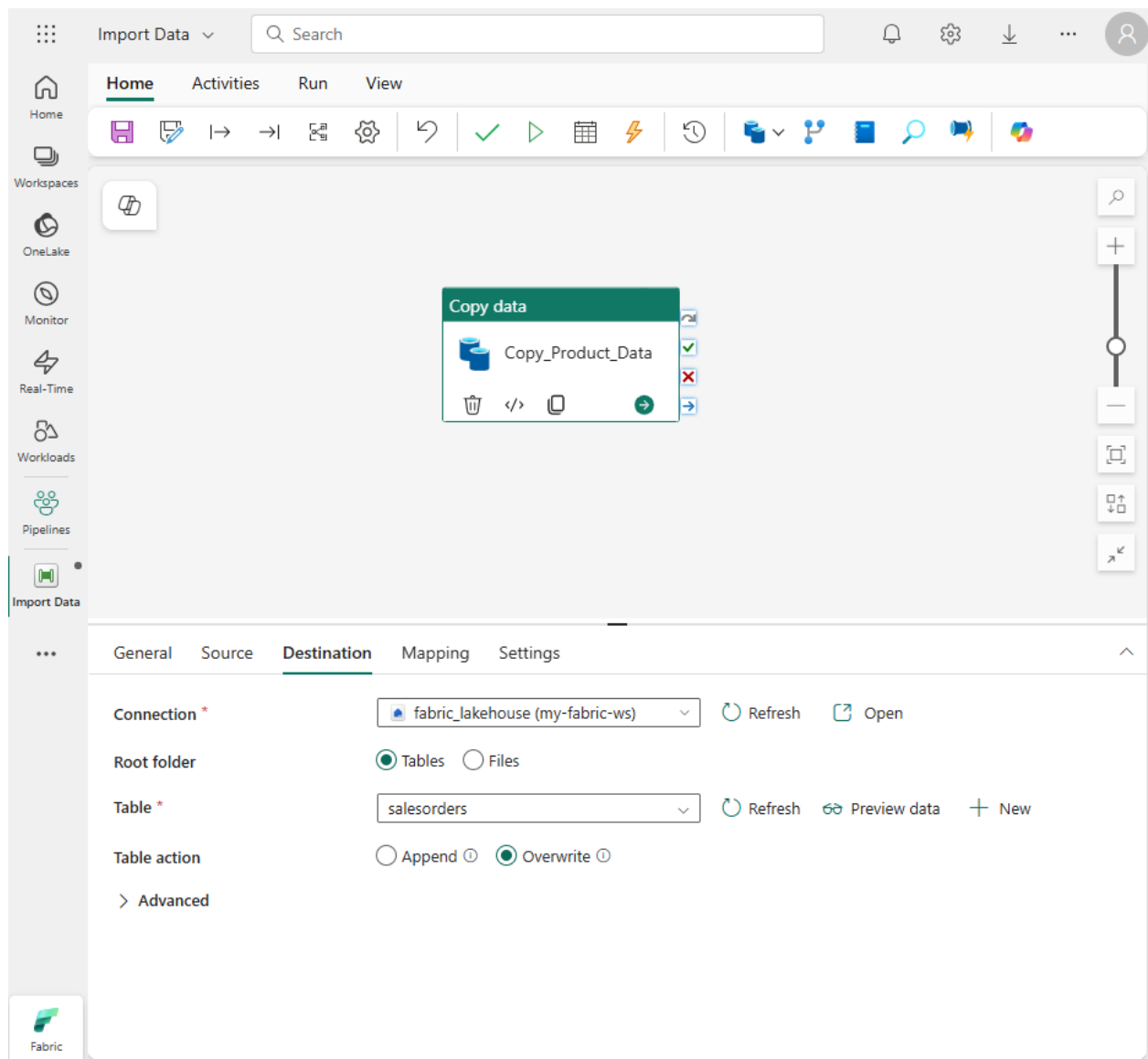


When you add a **Copy Data** activity to a pipeline, a graphical tool takes you through the steps required to configure the data source and destination for the copy operation. A wide range of source connections is supported, making it possible to ingest data from most common sources. In OneLake, this includes support for lakehouse, warehouse, SQL Database, and others.



Copy Data activity settings

After you've added a **Copy Data** activity to a pipeline, you can select it in the pipeline canvas and edit its settings in the pane underneath.



When to use the Copy Data activity

Use the **Copy Data** activity when you need to copy data directly between a supported source and destination without applying any transformations, or when you want to import the raw data and apply transformations in later pipeline activities.

If you need to apply transformations to the data as it is ingested, or merge data from multiple sources, consider using a **Data Flow** activity to run a dataflow (Gen2). You can use the Power Query user interface to define a dataflow (Gen2) that includes multiple transformation steps, and include it in a pipeline.

Tip

To learn more about Dataflow (Gen2) in Microsoft Fabric to ingest data, consider completing the [Ingest Data with Dataflows Gen2 in Microsoft Fabric](#) module.

4. Use pipeline templates

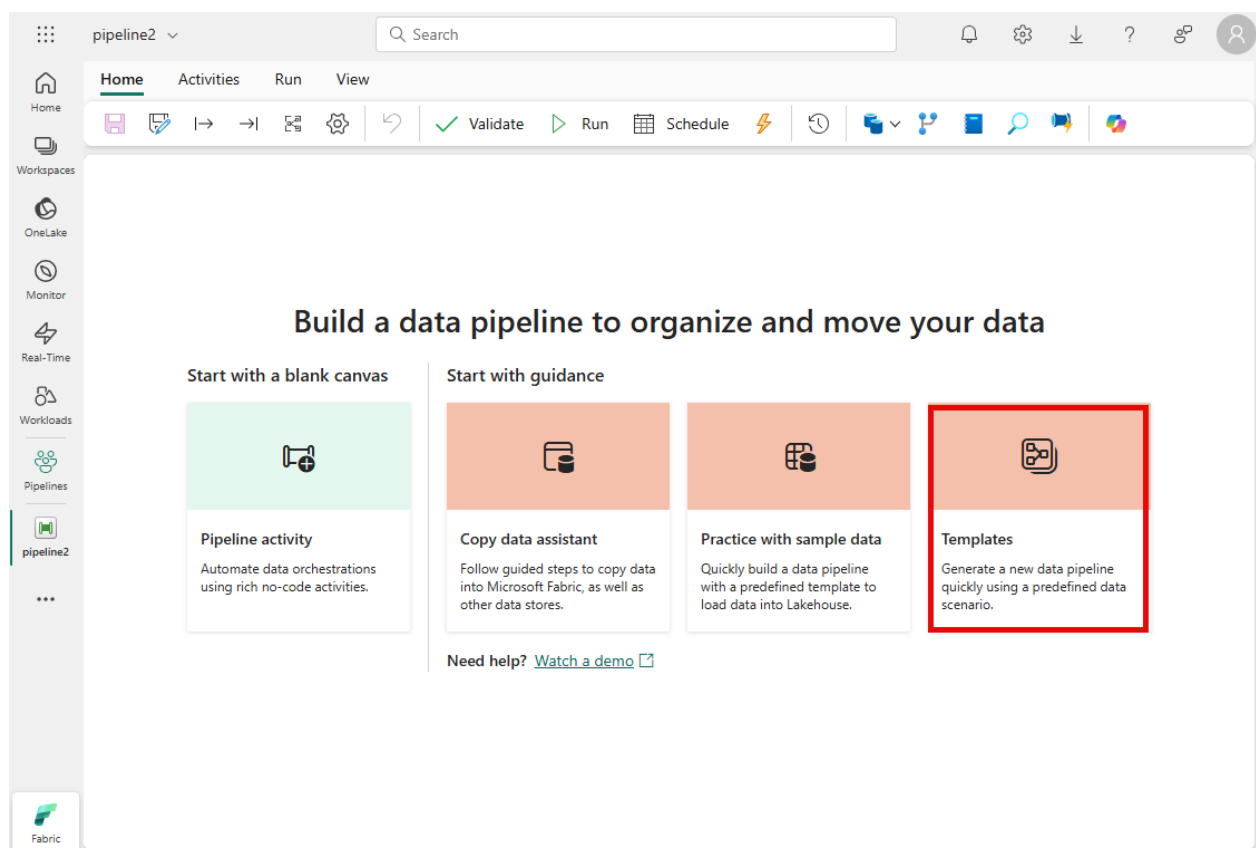
<https://learn.microsoft.com/en-us/training/modules/use-data-factory-pipelines-fabric/4-pipeline-templates>

Use pipeline templates

Completed

You can define pipelines from any combination of activities you choose, enabling to create custom data ingestion and transformation processes to meet your specific needs. However, there are many common pipeline scenarios for which Microsoft Fabric includes predefined pipeline templates that you can use and customize as required.

To create a pipeline based on a template, select the **Templates** tile in a new pipeline as shown here.



Selecting this option displays a selection of pipeline templates, as shown here.

 AI-Develop RAG pipeline using SQL database in Fabric by Microsoft Use this Retrieval Augmented Generation (RAG) data pipeline template to get your data ready in SQL Database in fabric for buildin...	 Bulk Copy from Database by Microsoft Use this template to copy data in bulk from a database using an external control table to store the partition list of your source tables...	 Bulk Copy from Files to Database by Microsoft Use this template to copy data in bulk from Azure Data Lake Storage Gen2 to Azure SQL Database.	 Copy data from ADLS Gen2 to Lakehouse file by Microsoft Use this template to copy data from ADLS Gen2 to a specified file location in your Lakehouse.
 Copy data from ADLS Gen2 to Lakehouse Table by Microsoft Use this template to copy data from ADLS Gen2 to a specified table in your Lakehouse.	 Copy data from Azure SQL DB to Lakehouse Table by Microsoft Use this template to copy data from your Azure SQL database to a specified table in your Lakehouse.	 Copy data from sample data to Lakehouse file by Microsoft Use this template to copy data from sample data (NYC Taxi - Green) to Lakehouse file.	 Copy data from sample data to Warehouse by Microsoft Use this template to copy data from sample data (Retail Data Model from Wide World Importers) to your Fabric Warehouse....

You can select the most appropriate template for your needs, and then edit the pipeline in the pipeline canvas to customize it to your needs.

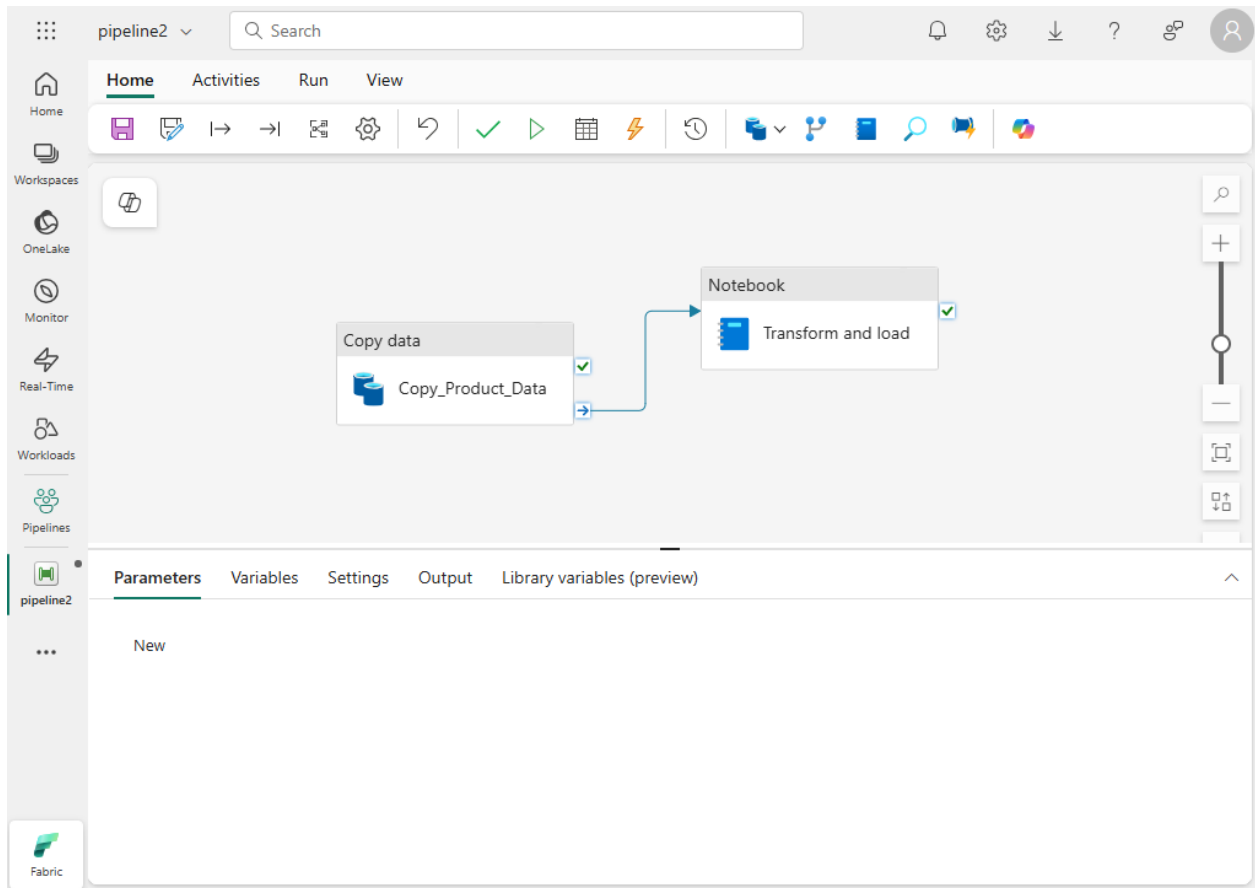
5. Run and monitor pipelines

<https://learn.microsoft.com/en-us/training/modules/use-data-factory-pipelines-fabric/5-run-monitor-pipelines>

Run and monitor pipelines

Completed

When you have completed a pipeline, you can use the **Validate** option to check that the configuration is valid, and then either run it interactively or specify a schedule.



View run history

You can view the run history for a pipeline to see details of each run, either from the pipeline canvas or from the pipeline item listed in the page for the workspace.

Home Activities Run View

Validate Run Schedule Add trigger View run history

Copy data

Copy data to Lakehouse

General Source Destination Mapping Settings

Name * Copy data to Lakehouse [Learn more](#)

Description

Activity state ☒ Activated ☐ Deactivated

Timeout 0.12:00:00

Retry 0

> Advanced

Recent runs for pipeline3

Activity name	Start time	Submitted by	Status
pipeline3	05/23/2025, 3:24 PM	System Administrator	Succeeded
pipeline3	05/23/2025, 3:19 PM	System Administrator	Succeeded
pipeline3	05/23/2025, 3:16 PM	System Administrator	Succeeded
pipeline3	05/23/2025, 3:15 PM	System Administrator	Succeeded

6. Exercise - Ingest data with a pipeline

<https://learn.microsoft.com/en-us/training/modules/use-data-factory-pipelines-fabric/6-exercise-pipelines>

Exercise - Ingest data with a pipeline

Completed

Now it's your chance to implement a pipeline in Microsoft Fabric. In this exercise, you create a pipeline that copies data from an external source into a lakehouse. Then enhance the pipeline by adding activities to transform the ingested data.

Note

You need a Microsoft Fabric trial license with the Fabric preview enabled in your tenant. See [Getting started with Fabric](#) to enable your Fabric trial license.

Launch the exercise and follow the instructions.

[Launch Exercise](#)

7. Module assessment

<https://learn.microsoft.com/en-us/training/modules/use-data-factory-pipelines-fabric/7-knowledge-check>

Module assessment

Completed

8. Summary

Summary

Completed

With Microsoft Fabric, you can create pipelines that encapsulate complex data ingestion and transformation processes. Pipelines provide an effective way to orchestrate data processing tasks that can be run on-demand or at scheduled intervals.

Tip

For more information about pipelines in Microsoft Fabric, see [Data pipelines](#) in the Microsoft Fabric documentation.