A close-up of a person smiling

Description automatically generated

# Lab 01: Getting Started with Fabric Data Factory

### 

### Introduction

In this lab, you will initiate your lab environment in Fabric by provisioning a Lakehouse, a Data Factory, and copying a set of CSV files from an Azure Blob Storage container.

### Objectives

After completing this lab, you will be better able to:

1. Provision a Fabric Lakehouse and a Fabric Data Factory
2. Configure a Data Pipeline to copy a set of CSV files
3. Load Lakehouse tables using the Data Factory

**Estimated time to complete this lab**

60 minutes

Contents

[Lab 01: Getting Started with Fabric Data Factory 1](#_Toc148303619)

[Introduction 1](#_Toc148303620)

[Objectives 1](#_Toc148303621)

[Task 1: Provision the Lakehouse and the Data Factory 3](#_Toc148303622)

[Task 2: Use the Copy Data wizard 5](#_Toc148303623)

**Lab Prerequisites**

* Workspace: Fabric, Power Premium or Fabric trial
* Individual license: Power Pro or Premium Per User account

**Information provided by your training provider**

* Trial tenant (if applicable): login & password, workspace to use for the lab.
* Azure Data Lake Gen2 (containing data sources): account name & shared access signature.

### Task 1: Provision the Lakehouse and the Data Factory

In this task, you will provide the Lakehouse and the Data Factory environment to be used in this training.

Connect to the Microsoft Fabric environment and go to your assigned workspace, priorly indicated by your trainer. If you don’t already have an assigned workspace, create a new one using either:

* Fabric Capacity (F sku)
* Fabric Capacity Trial
* PBI Premium Capacity (P sku)

Using the Menu on the bottom left corner, switch to the **Data Engineering** mode.

A screenshot of a computer

Description automatically generated

Click on **New Item**

A screenshot of a computer

Description automatically generated

Create a new **Lakehouse** et define the name as **Contoso**.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Using the Menu on the bottom left corner, switch to the **Data Factory** mode.

A screenshot of a computer

Description automatically generated

Create a new **Data pipeline** et define the name as **Contoso.**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

Please note that you can create any kind of Fabric artifact using the Create button on the tool bar on the left side of the UI.

**A screenshot of a computer

Description automatically generated**

Your workspace should contain the following artifacts:

**A screenshot of a computer

Description automatically generated**

### Task 2: Use the Copy Data wizard

In this task, you will copy a set of CSV files from an Azure Data Lake Gen 2 account to the “Bronze” zone of your Lakehouse using the Copy Data wizard.

These files are stored in the following path:

A screenshot of a computer

Description automatically generated

Open the **Data Pipeline** named Contoso created in the previous lab task.

**A screenshot of a computer

Description automatically generated**

Click on **Copy data assistant** to start the Copy data wizard.

A screenshot of a computer

Description automatically generated

Select **New**, then **Azure** connection type and click on **Azure Data lake Storage Gen 2**

A screenshot of a computer

Description automatically generated

Create a new connection based on SAS token.

Your trainer will give you the URL and the SAS token to use.

A screenshot of a computer

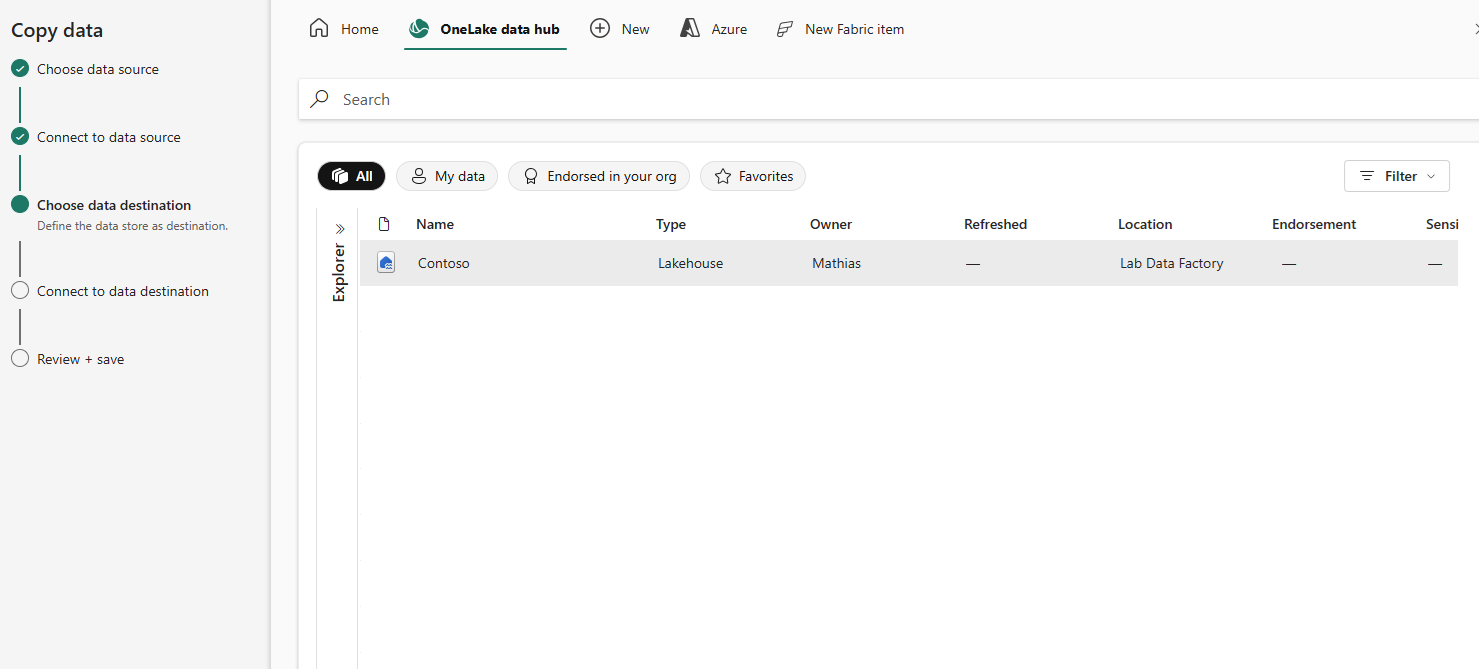
Description automatically generated

Select the **Dimensions** folder located in the **CSV** and select the **DelimitedText** as File Format. Click on **Next**.

A screenshot of a computer

Description automatically generated

Go to the **OneLake data hub** section and select the previously create **Lakehouse** as the destination. Click on **Next**.



Sign in to enable the connection and click on **Connect.**

A screenshot of a computer

Description automatically generated

Type **raw** as the folder path for the destination, keep blank the File Name textbox and select **Preserve hierarchy** as the Copy behavior. Click on **Next**.

A screenshot of a computer

Description automatically generated

Select the **Delimited** file format as configured hereafter. Click on **Next**.

A screenshot of a computer

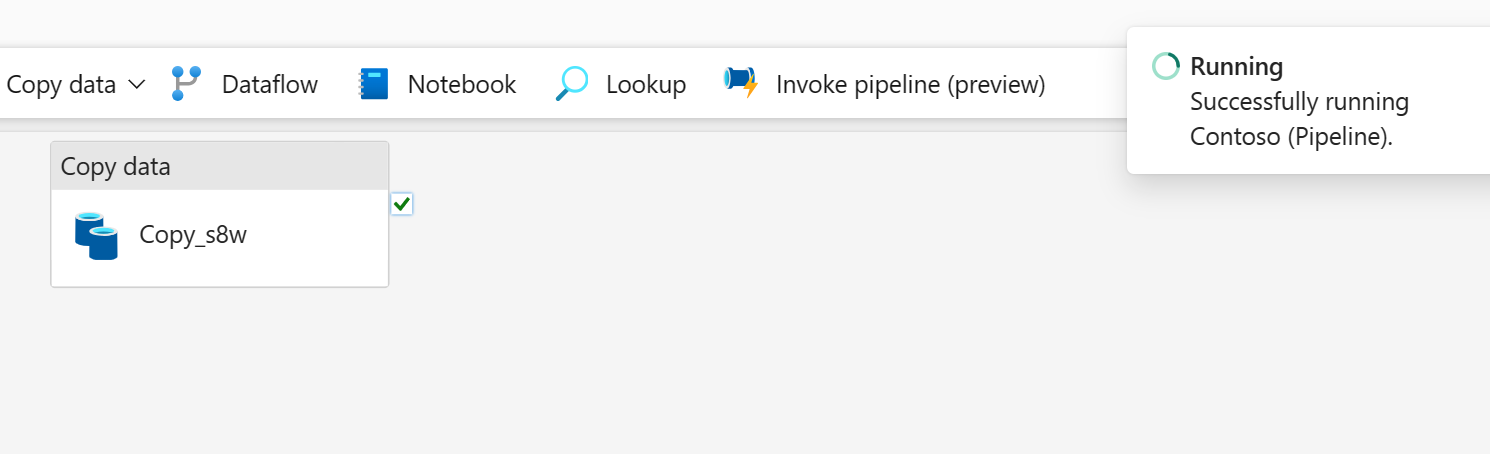
Description automatically generated

Click on **Save + Run**.

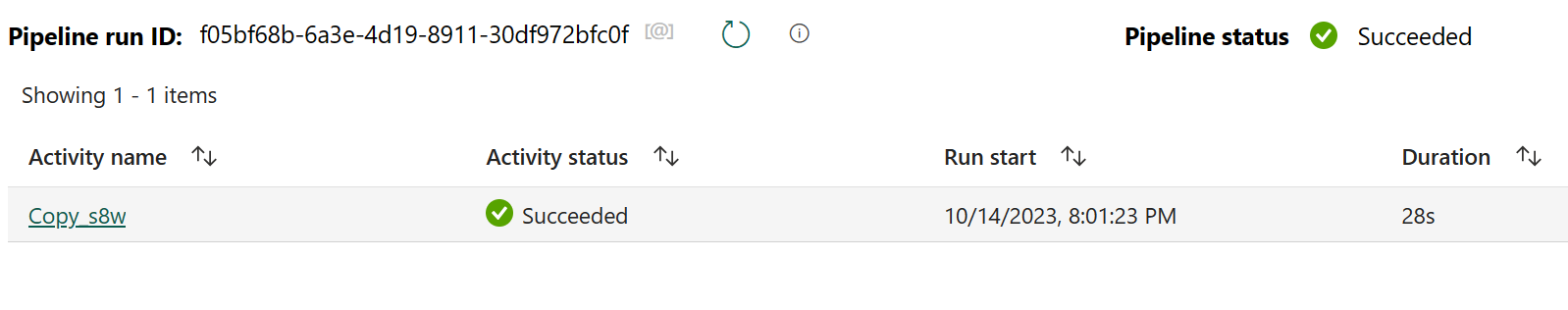
A screenshot of a computer

Description automatically generated

The **Copy data** component will be executed automatically.



A new activity appears, representing the current pipeline execution. Wait until having the end of the execution with the **Succeeded** Activity status.



Click on the activity name to see the execution details, including the duration, the number of files processed, and the quantity of data read and written.

A screenshot of a computer

Description automatically generated

Click on **Duration breakdown** to see the detailed sequence of the activity and the level of parallelism.

A blue rectangles with black text

Description automatically generated

Rename the **Copy data** activity as **Copy\_CSV**.

A screenshot of a computer

Description automatically generated

Save the pipeline and click on the Workspace in the toolbar to see the list of the workspace artifacts. Click on the **Contoso Lakehouse** created in the same workspace.

A screenshot of a computer

Description automatically generated

On the Lakehouse explorer, verify the 11 CSV files copied in the **raw** folder.

A screenshot of a computer

Description automatically generated