Microsoft Fabric in a Day Lab Manual – **Lab 2**

**Author**: Will Crayger

**Email**: [[wcrayger@lucidbi.co](mailto:wcrayger@lucidbi.co)](mailto:wcrayger@lucidbi.co)

**LinkedIn**: [LinkedIn.com/in/willcrayger/](https://www.linkedin.com/in/willcrayger/)

**Course Material**: [GitHub.com/Lucid-Will/FabCon-EU-Zero-To-Hero-with-Fabric](https://github.com/Lucid-Will/FabCon-EU-Zero-To-Hero-With-Fabric)

# Lab 2: Dataflows Generation 2 (Gen2) – Creating a Location Table

## Introduction

This lab will explore Fabric's low/no-code data integration workload, Dataflows gen2. The steps below will walk through creating a Dataflow, connecting it to a sample database, filtering by the necessary schema, and integrating base tables that will be used in the following labs.

## Creating the Dataflow Gen2

**Creating the Dataflow Gen2:** Select **Workspaces** from the left-side navigation blade and click your workspace name **fiad\_****{your\_initials}**, e.g. fiad\_wtc. This is your central hub for creating and managing all Fabric workloads. Select the **New item** dropdown and select the **Dataflow Gen2** workload from the menu that appears. Rename your dataflow by selecting Dataflow 1 in the top-left corner, type **locations**, and pressing Enter.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

## Connecting to Source Tables

**Connect to a Source**: begin by selecting **Get Data** from the top left of your screen. This will open a window allowing you to explore available data sources. From there, choose **View More** under the **New Sources** section to expand the available options.

Next, select **Azure** from the slicer. Once filtered, choose **Azure SQL database** from the list of options presented. You will then be prompted to input connection settings. Using to the “Source Connection Details” file, complete the **Server** and **Database** fields, and then set the **Authentication Kind** to **Basic** before completing the **Username** and **Password** fields. Click **Next**.

**Add Source Tables:** Tick checkboxes for the **Application.Cities** and **Application.StateProvinces** tables. Click **Create**.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

## Transforming Application Cities

Now that we’ve connected to the sample database and created queries for the tables we’ll be working with, we can start the data transformation process. We’ll begin by working with the **Application Cities** table before moving on to joining the **Application StateProvinces** table.

**Apply transformations to Cities:** Start by selecting **Application Cities** from the queries pane on the left. Navigate to the **Home** tab in the top ribbon and select **Choose Columns**. In the dialog box that appears, uncheck the **Select All** box. Then, check the boxes for the following fields: **CityID**, **CityName**, and **StateProvinceID**. Click **OK**. Reorganize the columns of the table to shuffle the Id fields to the beginning by clicking and dragging **StateProvinceId** to the second position.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

To enhance the readability of the transformation steps, edit the names of the **Applied steps**. Right-click on each step and select **Rename** to give meaningful descriptions:

* Rename **Navigation 1** to **Application Cities Table**
* Rename **Choose Columns** to **Removed Unused Application Cities Fields**
* Rename **Reordered Columns** to **Moved State Province ID**

A screenshot of a computer

Description automatically generated

After completing these steps, open the **Advanced Editor** to review the **Mashup (M code)** that has been written as the steps were applied. Pay close attention to the naming conventions we’ve applied, noting how they improve code navigation and readability.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Disable dataflow staging:** Right click Application Cities from the Queries blade and disable **Enable Staging**.

A screenshot of a computer

Description automatically generated

## Transforming Application StateProvinces

Now that we’ve had a detailed walkthrough of how to transform the Application Cities table, we’re going to repeat the process for Application StateProvinces with abbreviated instructions.

**Apply transformations to StateProvinces:** Select **Application StateProvinces** from the queries pane on the left then use **Choose columns** to keep only **StateProvinceID, StateProvinceCode, StateProvinceName** and **SalesTerritory**. Click **OK**.

A screenshot of a computer

Description automatically generated

**Apply meaningful names to applied steps:**

**A screenshot of a computer

Description automatically generated**

**Disable Dataflow staging:**

A screenshot of a computer

Description automatically generated

## Merge Tables

Now that our tables are prepared, we’re ready to join them together to create a single **locations** table. There are several ways to accomplish this. We’ll use the diagram view to merge Application Cities and Application StateProvinces tables.

**Merge tables using Diagram View:** To merge the tables, start by navigating to the **View** tab in the top ribbon. From there, select the **Diagram View** toggle. In this view, you’ll see the two tables represented in a visual interface.

Next, click on the ellipsis (three dots) in the top right corner of the **Application StateProvinces** table. From the options, select **Merge queries as new**. Select **StateProvinceID** from the **Application StateProvinces** table. In the second dropdown menu, select **Application Cities** and choose **StateProvinceID** as the join field. For the join type, select **Inner** and then click **OK**.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Review the newly created table:** At this point, your diagram view should show the two **Application** tables merging into the **Merge** table. Right-click this new table, select Rename, type **locations** and press enter. In the table preview at the bottom, expand the contents of the **Application Cities** column by clicking , unchecking the **StateProvinceID** box (as this would result in a duplicate column), and pressing OK. Reorganize the columns by moving the **CityID** field to the beginning.

A screenshot of a computer

Description automatically generated

**Disable dataflow staging:** To finalize, expand the **Queries** pane on the left and disable staging for the **locations** table.

A screenshot of a computer

Description automatically generated

## Writing to a Destination

With the transformation process complete, we now have the final version of the **locations** table. The next step is to set our Lakehouse as the destination to the dataflow and load the **locations** table.

**Add dataflow destination:** To begin adding a destination to the dataflow, expand the **Query settings** pane to the right, and click the “+” button in the bottom right corner. From the available options, select **Lakehouse**. Leave the default settings as they are and click **Next**.

**Note: You may need to re-authenticate as the cached credentials may have expired. Use the organizational account credentials assigned to you to re-authenticate.**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

Navigate to your **Fabric in a Day Workspace** and expand the folder. Select the Lakehouse that you created earlier, labeled as **bronze\_lakehouse\_{your initials}**, and click **Next**.

A screenshot of a computer

Description automatically generated

Untoggle **Use automatic settings** at the top, ensure **Replace** is selected, and click **Save Settings**. To complete the development of the dataflow, click the **Publish** button down the bottom-right, which will initiate loading the **locations** table into your Lakehouse.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Once the publishing process begins, you'll see two different loading symbols. The first "swirl" should be quick, and it indicates that the metadata of the dataflow is being saved. You may also notice other artifacts appear in your workspace during this process, which are temporary and will automatically disappear upon completion.

A screenshot of a computer

Description automatically generated

The second "swirl" represents the execution of the dataflow. After it finishes, navigate to the **Lakehouse**. You should now see your **locations** table under the **Tables** section of the **Lakehouse Explorer**. Click on the table to preview the data.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Note: If the locations table is missing or you see “Undefined” in your Lakehouse, please refresh your browser.**

You have now successfully completed this lab.