Microsoft Fabric in a Day Lab Manual – Lab 5

Prepared by: Will Crayger

# Lab 5: Data Engineering – Data Transformation and Engineering in Fabric

## Introduction:

In this lab, you’re going to create a “silver” Lakehouse and begin loading data from the Lakehouse that was created in Lab 1.

**Note: As you begin querying or transforming data, please note that you’re working in a case-sensitive environment.**

## Part 1: Creating the Silver Lakehouse

The reason for creating a Silver Lakehouse is to simulate the medallion development pattern. The Silver Lakehouse will give you separation from the Bronze layer, allowing you to begin cleansing and shaping the data.

1. Navigate to the data engineering landing page of Fabric

A screenshot of a computer

Description automatically generated

1. Chose Lakehouse from the tiles across the top

A screenshot of a computer

Description automatically generated

1. Name your lakehouse “Silver\_Lakehouse\_{your\_initials}” and click Create

A screenshot of a computer

Description automatically generated

1. Notice that because we’re creating a new lakehouse the tables and files sections from the Lakehouse Explorer are empty

A screenshot of a computer

Description automatically generated

1. You’re now done with Part 1 of the lab

## Part 2: Data Engineering via Fabric Notebook

In the Notebook section of lab 5, the notebook will serve as the primary documentation source and have been commented for visibility, clarity, and are complete with code examples that can be referenced to complete each code exercise.

The notebook that will be used for these exercises is the “Lab 5 - Data Engineering and Modeling Notebook” that was provided as part of the program. To set up the environment and launch the notebook, complete the following steps.

1. Navigate to the data engineering landing page of Fabric

A screenshot of a computer

Description automatically generated

1. Click Import Notebook from the actions across the top

A screenshot of a computer

Description automatically generated

1. Click upload, navigate to the location of the course files that were saved, select the Data Engineering and Modeling Notebook, and click Open

A screenshot of a computer

Description automatically generated

1. Navigate back to the workspace home page and you should see the notebook as part of the workspace artifacts
2. Click the notebook to open

A screenshot of a computer

Description automatically generated

1. Click Add in the Add Lakehouse blade

A screenshot of a computer

Description automatically generated

1. Select Existing Lakehouse and click Add

A screenshot of a computer

Description automatically generated

1. Choose the Lakehouse that was created as part of Lab 1 as well as the Silver Lakehouse from Part 1 of this lab from the list and click Add

A screenshot of a computer

Description automatically generated

1. You’re now ready to begin executing code from your notebook

**Note: Notebooks are structured in a way that reads top to bottom and should be executed in the same manner. There are two types of cells in a notebook, Markdown and Code cells. Markdown cells are not executable and are intended for comments.**

**In the markdown cells, you’ll find overviews of what each subsequent cell (or group of cells) does as well as a code snippet that can be used for reference to complete the code cells. I encourage you to write some of the code to familiarize yourself with the different approaches, but you can copy/paste the code as well.**

**Code cells are intended to be completed by you and give you a hands-on experience to data engineering in a notebook.**

**To execute a code cell in the notebook (after adding code), click the “play” icon next to it to “Run Cell”.**

**Don’t forget to change the Lakehouse reference if you do choose the copy/paste approach.**

**A screenshot of a computer program

Description automatically generated**

**When you’ve completed the exercises, you will have 2 new tables in your Silver\_Lakehouse visible from the Lakehouse Explorer as shown below.**

**A screenshot of a computer

Description automatically generated**

**Note: If you encounter error indicators in your notebook as shown below, disregard them. The code cell will still execute.**

A screenshot of a computer program

Description automatically generated

## Part 3: Data Engineering via Lakehouse SQL Endpoint

We can also leverage the Lakehouse SQL Analytics Endpoint to transform and prepare our data. In this section, you’ll use cross-database joins to query data from the bronze Lakehouse and create a view in the silver Lakehouse.

1. Navigate to the Silver Lakehouse SQL Analytics Endpoint from the workspace landing page

A screenshot of a computer

Description automatically generated

1. Click the “+ Warehouses” button at the top of the Explorer

A screenshot of a computer

Description automatically generated

1. Check the box for your Lab 1 Lakehouse and click Connect

A screenshot of a computer

Description automatically generated

1. You will now see the schema from the Lab 1 Lakehouse available in the Explorer

A screenshot of a computer

Description automatically generated

1. Select New SQL Query from the action ribbon at the top

A screenshot of a computer

Description automatically generated

1. Begin building the query needed to retrieve data from the Lab 1 (bronze) Lakehouse by simply writing SELECT FROM and dragging “Application\_People” from the Lakehouse\_WC schema to your query

A screenshot of a computer

Description automatically generated

1. Update the SELECT statement to include only the PersonID, FullName, PreferredName, and LogonName columns by either dragging the column names to the query window or by typing them in directly
   1. As you begin typing, notice the intellisense working to identify the schema and aid you in writing your query

A screenshot of a computer

Description automatically generated

1. Your final query should look similar to the image below

A screen shot of a computer

Description automatically generated

1. Execute the query by clicking Run in the query window and review the output

A screenshot of a computer

Description automatically generated

1. Now that we have our query complete, copy the query from the query window and click “Save as view”

A blue and black text on a white background

Description automatically generated

1. Name the view salesperson and click OK

A screenshot of a computer

Description automatically generated

1. You will now have a view created in your Silver\_Lakehouse called salesperson

A screenshot of a computer

Description automatically generated

1. You are now done with Part 3 of this lab