Microsoft Fabric in a Day Lab Manual – Lab 4

Prepared by: Will Crayger

# Lab 4: Working with Pipelines – Extracting Source Data

## Introduction:

For this lab, we will create two pipelines. Copying data from source to target is one of the most commonly used patterns for pipelines. As such, our first pipeline will be created to extract a single table from an Azure SQL Database. Our second pipeline will also be used to extract data, but we’ll create a pattern that loops through many tables using a lookup activity. Leveraging lookups and parameters are incredibly powerful and efficient, as you’ll see during the exercise.

## Part 1: Single Object Copy

1. From the Data Engineering Fabric Landing page, select New and chose Data Pipeline

A screenshot of a computer

Description automatically generated

1. Name the pipeline “Single Object Copy” and click Create

A screenshot of a computer

Description automatically generated

1. Select Copy Data from the pipeline landing page

A screenshot of a data

Description automatically generated

1. Notice the many available data sources from the list to extract data from
2. Select Azure from the top to filter the list

A screenshot of a computer

Description automatically generated

1. Select Azure SQL Database and click next

A screenshot of a computer

Description automatically generated

1. Ensure the “Existing connection” radial button is active, select “wwi-sample-sqldb” from the Connection dropdown list, and click Next

A screenshot of a computer

Description automatically generated

1. Select the “Sales.Customers” table from the list by checking the box and click Next

A white chair with a white frame

Description automatically generated

1. Select the “Workspace” tab from the top of the Destination menu, choose Lakehouse, and click Next

A screenshot of a computer

Description automatically generated

1. Ensure the “Existing Lakehouse” radial button is active, select the Lakehouse that was created in Lab 1 from the dropdown menu, and click Next

A screenshot of a computer

Description automatically generated

1. Review the settings on the next page, but leave them all default and click Next

A screenshot of a computer

Description automatically generated

1. Review the pipeline summary, uncheck the “Start data transfer immediately” box, and click OK

A screenshot of a computer

Description automatically generated

1. Select the copy data activity from the Canvas and change the name of the activity to Copy Single Object

A screenshot of a computer

Description automatically generated

1. Open the Source tab of the activity and review the configuration that was populated using the Copy Data Tool

A screenshot of a computer

Description automatically generated

1. Open the Destination tab of the activity and review the configuration

A screenshot of a computer

Description automatically generated

1. Click Advanced to open the advanced options

A screenshot of a computer

Description automatically generated

1. Notice that we can change the writer behavior to Append or Overwrite
2. Go back to the Home tab of the pipeline and click the Validate button in the activity bar of the pipeline

A screenshot of a computer

Description automatically generated

1. If everything was done correctly, you shouldn’t receive any errors and can close the Pipeline Validation Output

A white and grey logo

Description automatically generated

1. Click Run from the activity bar

A screenshot of a computer

Description automatically generated

1. Click Save and run

A screenshot of a computer error

Description automatically generated

1. The Output of the pipeline run will automatically come up
2. You can track the execution through the Output of the pipeline activity

A screenshot of a computer

Description automatically generated

1. You can also click the Activity Name to launch the Copy data details blade and track the progress from there as well as review various metrics about the run

A screenshot of a computer

Description automatically generated

1. Once the pipeline has completed, navigate to back to your Lakehouse to review the new table that has been created for Sales\_Customers.

A screenshot of a computer

Description automatically generated

1. You are now done with Part 1 of the lab

## Part 2: Batch Object Copy

1. From the Data Engineering Fabric Landing page, select New and chose Data Pipeline

A screenshot of a computer

Description automatically generated

1. Name the pipeline “Batch Object Copy” and click Create

A screenshot of a computer

Description automatically generated

1. Click “Add pipeline activity” from the pipeline landing page

A screenshot of a data pipeline

Description automatically generated

1. We’ll start by executing a Lookup to return a list of schema.table combos, so choose the Lookup activity from the list

A screenshot of a computer

Description automatically generated

1. Select the Lookup activity and name it “Lookup Schema and Table”

A screenshot of a computer

Description automatically generated

1. Navigate to the Settings tab and select the previously used “wwi-sample-sqldb” connection from the previous lab
2. Change the connection type to Azure SQL Database
3. Click the Query radial button

A screenshot of a computer

Description automatically generated

1. Find the “Lab 4 - Batch Load Source Query” file that was shared as part of the labs, copy the query from the file, and paste it into the Query box of the Lookup activity

A screen shot of a computer

Description automatically generated

1. Click Preview Data to see the output of the Lookup query

A screenshot of a computer

Description automatically generated

1. Uncheck the box for “First row only”

A screenshot of a computer

Description automatically generated

1. Open the Activities Tab from the top of the pipeline and click the ForEach activity to add to the pipeline canvas

A screenshot of a computer

Description automatically generated

1. Click and drag from the green checkmark on the Lookup activity to the ForEach activity to create a connection between them

A screenshot of a computer

Description automatically generated

1. Click on the ForEach activity and change the name to “Extraction Loop”

A screenshot of a computer

Description automatically generated

1. Click on the Settings tab
2. Click in the Items box and click “Add dynamic content” to edit the Items that will be looped through

A screenshot of a computer

Description automatically generated

1. Select “Lookup Schema and Table value array from the Activity outputs list to automatically populate the dynamic field and click OK

A screenshot of a computer

Description automatically generated

1. Click the Pencil icon on the ForEach Activity to edit the inside of the loop

A screenshot of a computer

Description automatically generated

1. From the Activities tab, select Copy Data and Add to canvas

A screenshot of a computer

Description automatically generated

1. Click the Copy Data task and change the name to “Batch Object Copy”

A screenshot of a computer

Description automatically generated

1. Under the Source tab, set the connection to “wwi-sample-sqldb”, connection type to Azure SQL Database, use query to Table, and check the Enter Manually box

A screenshot of a computer

Description automatically generated

1. Click the schema name box and select “Add dynamic content”

A screenshot of a computer

Description automatically generated

1. Click “Extraction Loop” from the ForEach iterator options to populate the first part of the schema name

A screenshot of a computer

Description automatically generated

1. Now, we must tell the pipeline which field from the Lookup activity we want to capture, which is “schema\_name”, and click OK

A screenshot of a computer

Description automatically generated

1. Repeat this process for the table name box

A screenshot of a computer

Description automatically generated

1. Navigate to the Destination tab and select the Lakehouse that was created in Lab 1

A screenshot of a computer

Description automatically generated

1. Click on the Table name box and click “Add dynamic content”

A screenshot of a computer

Description automatically generated

1. Paste the following into the expression box to create a concatenated “schema\_table” table name and click OK
   1. @concat(item().schema\_name, '\_', item().table\_name)

A screenshot of a computer

Description automatically generated

1. Open the Advanced menu and change the table action to Overwrite

A screenshot of a computer

Description automatically generated

1. Return to the Home tab and click Validate to ensure there are no errors in the pipeline
2. Click Run to execute the pipeline

A screenshot of a computer

Description automatically generated

1. Monitor the pipeline run in the Output tab
   1. Notice that there are multiple Copy activities happening simultaneously

A screenshot of a computer

Description automatically generated

1. Once the pipeline is complete, navigate back to your Lakehouse to confirm the delta tables have been created
   1. If the tables aren’t immediately visible, refresh your browser or right-click “Tables” and refresh the schema

A screenshot of a computer

Description automatically generated

1. You are now done with Part 2 this lab