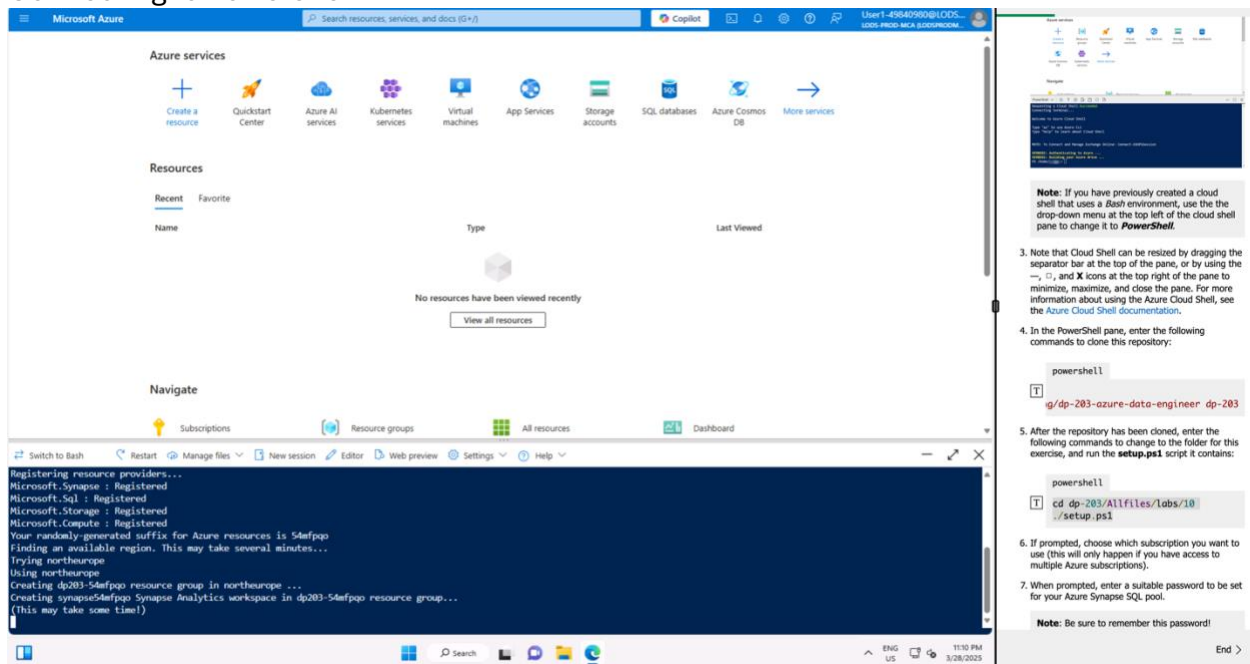


Lab 10 Synapse Pipeline

BUAN 6390.001 – Analytics Practicum

Name: Swastik Bhatnagar
Net ID: Sxb220210

Connecting to Powershell:



Microsoft Azure

Azure services

Resources

Navigate

Switch to Bash

Registering resource providers...

Microsoft.Synapse : Registered

Microsoft.Sql : Registered

Microsoft.Storage : Registered

Microsoft.Compute : Registered

Your randomly-generated suffix for Azure resources is 54mfpp

Finding an available region. This may take several minutes...

Trying northeurope

Using northeurope

Creating dp203-54mfpp resource group in northeurope ...

Creating synapse54mfpp Synapse Analytics workspace in dp203-54mfpp resource group...

(This may take some time)

Note: If you have previously created a cloud shell that uses a Bash environment, use the drop-down menu at the top left of the cloud shell pane to change it to **PowerShell**.

3. Note that Cloud Shell can be resized by dragging the separator bar at the top of the pane, or by using the **+**, **[-]**, and **X** icons at the top right of the pane to minimize, maximize, and close the pane. For more information about using the Azure Cloud Shell, see the [Azure Cloud Shell documentation](#).

4. In the PowerShell pane, enter the following commands to clone this repository:

```
powershell
git clone https://github.com/AzureDataEngineering/dp-203-azure-data-engineer dp-203
```

5. After the repository has been cloned, enter the following commands to change to the folder for this exercise, and run the **setup.ps1** script it contains:

```
powershell
cd dp-203/Allfiles/Labs/10
./setup.ps1
```

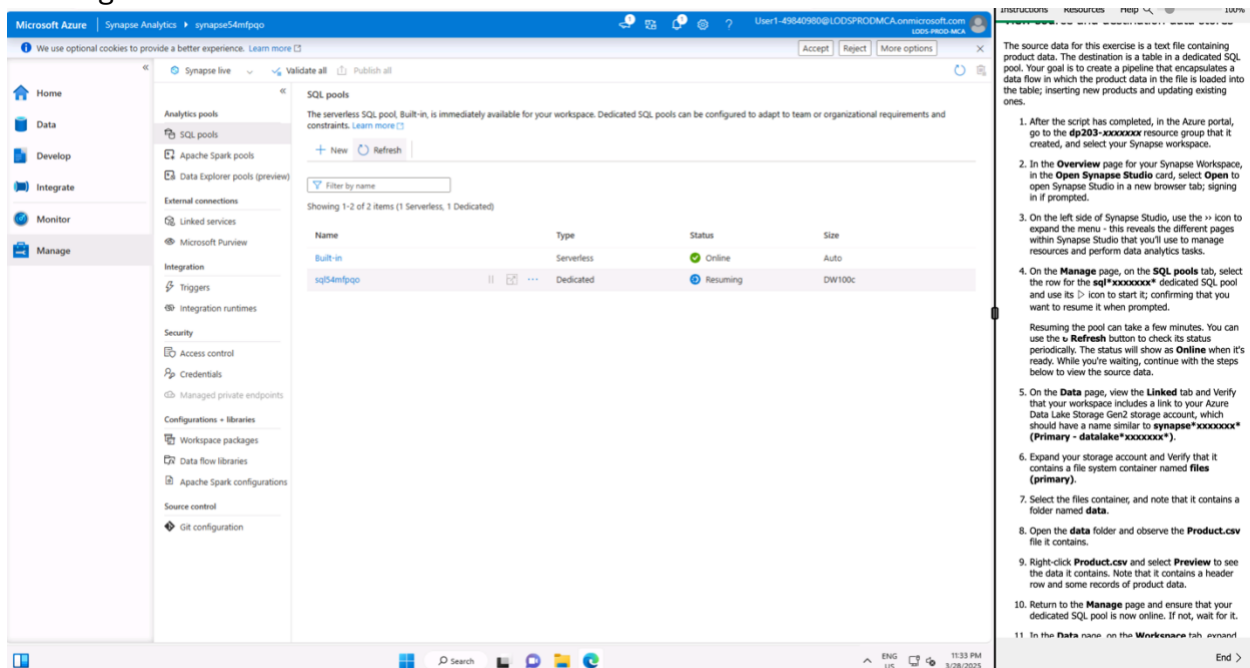
6. If prompted, choose which subscription you want to use (this will only happen if you have access to multiple Azure subscriptions).

7. When prompted, enter a suitable password to be set for your Azure Synapse SQL pool.

Note: Be sure to remember this password!

End >

Viewing Source and Destination data stores:



Microsoft Azure | Synapse Analytics | synapse54mfpp

SQL pools

The serverless SQL pool, built-in, is immediately available for your workspace. Dedicated SQL pools can be configured to adapt to team or organizational requirements and constraints. [Learn more](#)

+ New Refresh

Filter by name

Showing 1-2 of 2 items (1 Serverless, 1 Dedicated)

Name	Type	Status	Size
Built-in	Serverless	Online	Auto
sql54mfpp	Dedicated	Resuming	DW100c

The source data for this exercise is a text file containing product data. The destination is a table in a dedicated SQL pool. Your goal is to create a pipeline that encapsulates a data flow in which the product data in the file is loaded into the table; inserting new products and updating existing ones.

1. After the script has completed, in the Azure portal, go to the **dp203-xxxxxx** resource group that it created, and select your Synapse workspace.

2. In the **Overview** page for your Synapse Workspace, in the **Open Synapse Studio** card, select **Open** to open Synapse Studio in a new browser tab; signing in if prompted.

3. On the left side of Synapse Studio, use the **>>** icon to expand the menu - this reveals the different pages within Synapse Studio that you'll use to manage resources and perform data analytics tasks.

4. On the **Manage** page, on the **SQL pools** tab, select the row for the **sql54mfpp** dedicated SQL pool and use its **>** icon to start it; confirming that you want to resume it when prompted.

Resuming the pool can take a few minutes. You can use the **Refresh** button to check its status periodically. The status will show as **Online** when it's ready. While you're waiting, continue with the steps below to view the source data.

5. On the **Data** page, view the **Linked** tab and verify that your workspace includes a link to your Azure Data Lake Storage Gen2 storage account, which should have a name similar to **synapsexxxxxx** (**Primary - datalakexxxxxx**).

6. Expand your storage account and verify that it contains a file system container named **files** (**primary**).

7. Select the **files** container, and note that it contains a folder named **data**.

8. Open the **data** folder and observe the **Product.csv** file it contains.

9. Right-click **Product.csv** and select **Preview** to see the data it contains. Note that it contains a header row and some records of product data.

10. Return to the **Manage** page and ensure that your dedicated SQL pool is now online. If not, wait for it.

11. In the **Data** pane, on the **Workspace** tab, expand

End >

Product.csv

PRODUCTID	PRODUCTNAME	COLOR	SIZE
AR-5381	Adjustable Race	Red	10.00
BA-8327	Bearing Ball	NA	10.00
BE-2349	Ball Bearing	NA	10.00
BE-2908	Headset Ball B...	NA	10.00
BL-2036	Blade	NA	10.00
CA-5965	LL Crankarm	Silver	10.00
CA-6738	ML Crankarm	Black	10.00
CA-7457	HL Crankarm	Silver	10.00
CB-2903	Chaining Bolts	Silver	10.00
CN-6137	Chaining Nut	Silver	10.00
CR-7833	Chaining	Black	10.00

Showing 1 to 1 of 1 cached items

Instructions

- On the left side of Synapse Studio, use the **+** icon to expand the menu - this reveals the different pages within Synapse Studio that you'll use to manage resources and perform data analytics tasks.
- On the **Manage** page, on the **SQL pools** tab, select the row for the **sql*xxxxxxx*** dedicated SQL pool and use its **D+** icon to start it; confirming that you want to resume it when prompted.
- Resuming the pool can take a few minutes. You can use the **Refresh** button to check its status periodically. The status will show as **Online** when it's ready. While you're waiting, continue with the steps below to view the source data.
- On the **Data** page, view the **Linked** tab and Verify that your workspace includes a link to your Azure Data Lake Storage Gen2 storage account, which should have a name similar to **synapse*xxxxxxx*** (**Primary - data lake*xxxxxxx***).
- Expand your storage account and Verify that it contains a file system container named **files** (**primary**).
- Select the files container, and note that it contains a folder named **data**.
- Open the **data** folder and observe the **Product.csv** file it contains.
- Right-click **Product.csv** and select **Preview** to see the data it contains. Note that it contains a header row and some records of product data.
- Return to the **Manage** page and ensure that your dedicated SQL pool is now online. If not, wait for it.
- In the **Data** page, on the **Workspace** tab, expand **SQL database**, your **sql*xxxxxxx*** (**SQL**) database, and its **Tables**.
- Select the **dbo.DimProduct** table. Then in its **...** menu, select **New SQL script** > **Select TOP 100 rows**; which will run a query that returns the product data from the table - there should be a single row.

Implement a pipeline

To load the data in the text file into the database table, you will implement an Azure Synapse Analytics pipeline that contains a dataflow encapsulating the logic to ingest the data from the text file. In this example, **LoadProducts**.

Implementing a Pipeline:

Creating a pipeline with a data flow activity:

Integrate

Filter resources by name

Pipelines

- Load Product Data

Activities

- Synapse
- Move and transform
- Copy data
- Data flow
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning

Data flow

LoadProducts

Settings

General

Compute size: Small

Advanced

Logging level: Verbose

Sink properties

Staging linked service: synapse54mfpgo-WorkspaceDefault...

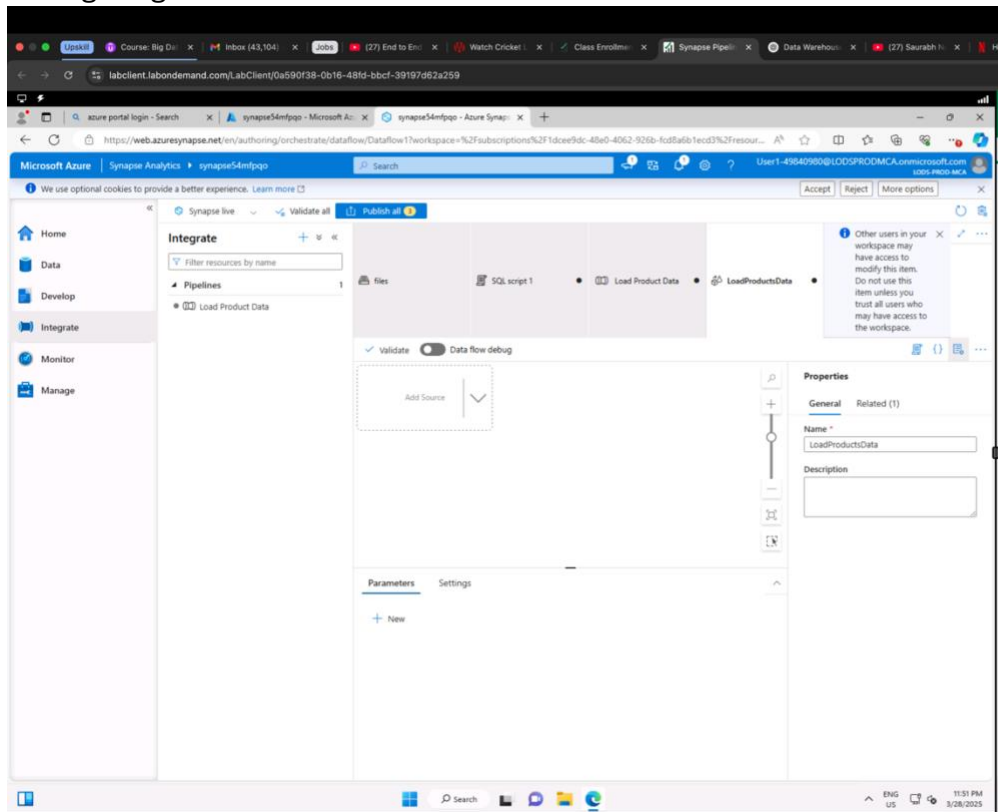
Integration runtime: AutoResolveIntegrationRuntime

Staging storage folder: files

Configure the data flow

- At the top of the **Settings** tab for the **LoadProducts** data flow, for the **Data flow** property, select **New**.
- In the **Properties** pane for the new data flow design surface that opens, set the **Name** to **LoadProductsData** and then hide the **Properties** pane. The data flow designer should look like this:

Configuring the data flow:



Microsoft Azure | Synapse Analytics | synapse54mfpqo

Integrate

Filter resources by name

Pipelines

Load Product Data

Validate

Data flow debug

Add Source

Properties

General

Related (1)

Name *

LoadProductsData

Description

Source type

Integration dataset

Dataset: Add a New dataset with the following properties:

- Type: Azure Datalake Storage Gen2
- Format: Delimited text
- Name: Products_Csv
- Linked service: synapse*xxxxxxxx*-WorkspaceDefaultStorage
- File path: files/data/Product.csv
- First row as header: Selected
- Import schema: From connection/store

Other users in your workspace may have access to modify this item. Do not use this item unless you trust all users who may have access to the workspace.

End >

Synapse Pipeline

1 Hr 13 Min Remaining

Instructions Resources Help

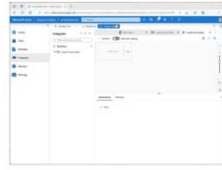
100%

workspaceuserstorage inbox service.

Staging storage folder: Set container to files and Directory to stage_products.

Configure the data flow

- At the top of the **Settings** tab for the **LoadProducts** data flow, for the **Data flow** property, select **+ New**.
- In the **Properties** pane for the new data flow design surface that opens, set the **Name** to **LoadProductsData** and then hide the **Properties** pane. The data flow designer should look like this:

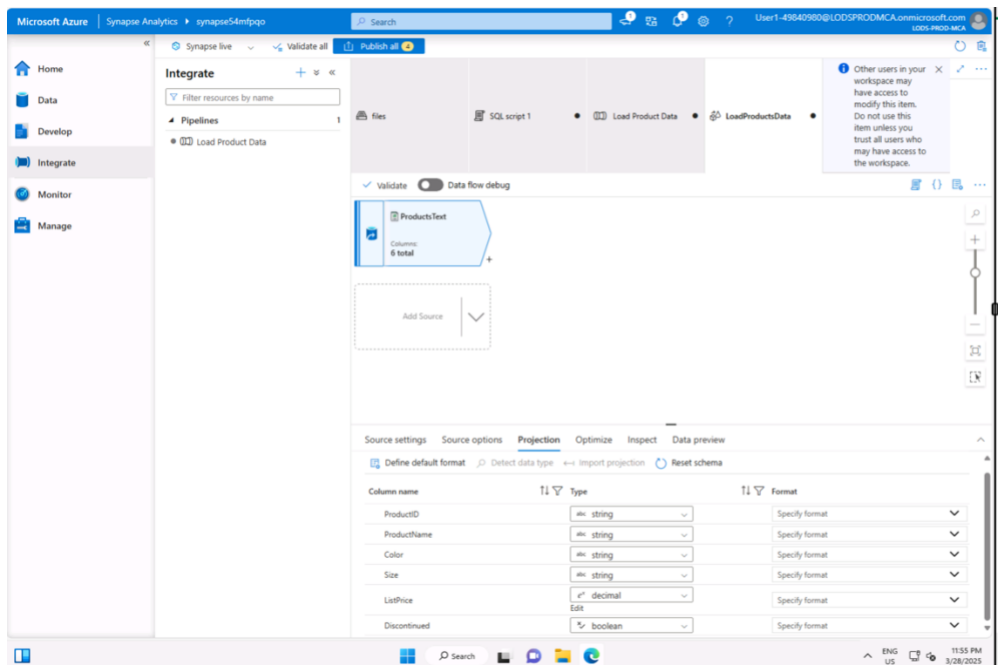


Add sources

- In the data flow design surface, in the **Add Source** drop-down list, select **Add Source**. Then configure the source settings as follows:
 - Output stream name: ProductsText
 - Description: Products text data
 - Source type: Integration dataset
 - Dataset: Add a New dataset with the following properties:
 - Type: Azure Datalake Storage Gen2
 - Format: Delimited text
 - Name: Products_Csv
 - Linked service: synapse*xxxxxxxx*-WorkspaceDefaultStorage
 - File path: files/data/Product.csv
 - First row as header: Selected
 - Import schema: From connection/store
 - Allow schema drift: Selected

End >

Now Adding Sources:



Microsoft Azure | Synapse Analytics | synapse54mfpqo

Integrate

Filter resources by name

Pipelines

Load Product Data

Validate

Data flow debug

Add Source

Properties

General

Related (1)

Name *

LoadProductsData

Description

Source type

Integration dataset

Dataset: Add a New dataset with the following properties:

- Type: Azure Datalake Storage Gen2
- Format: Delimited text
- Name: Products_Csv
- Linked service: synapse*xxxxxxxx*-WorkspaceDefaultStorage
- File path: files/data/Product.csv
- First row as header: Selected
- Import schema: From connection/store

Other users in your workspace may have access to modify this item. Do not use this item unless you trust all users who may have access to the workspace.

End >

Synapse Pipeline

1 Hr 13 Min Remaining

Instructions Resources Help

100%

workspaceuserstorage inbox service.

Staging storage folder: Set container to files and Directory to stage_products.

Add sources

- In the data flow design surface, in the **Add Source** drop-down list, select **Add Source**. Then configure the source settings as follows:
 - Output stream name: ProductsText
 - Description: Products text data
 - Source type: Integration dataset
 - Dataset: Add a New dataset with the following properties:
 - Type: Azure Datalake Storage Gen2
 - Format: Delimited text
 - Name: Products_Csv
 - Linked service: synapse*xxxxxxxx*-WorkspaceDefaultStorage
 - File path: files/data/Product.csv
 - First row as header: Selected
 - Import schema: From connection/store
 - Allow schema drift: Selected
- On the **Projection** tab for the new **ProductsText** source, set the following data types:
 - ProductID: string
 - ProductName: string
 - Color: string
 - Size: string
 - ListPrice: decimal
 - Discontinued: boolean
- Add a second source with the following properties:
 - Output stream name: ProductTable
 - Description: Product table
 - Source type: Integration dataset
 - Dataset: Add a New dataset with the following properties:
 - Type: Azure Synapse Analytics

End >

Microsoft Azure | Synapse Analytics | synapse54mfppg

Integrate

Filter resources by name

Pipelines

Load Product Data

Files

SQL script 1

Load Product Data

LoadProductsData

ProductsText

ProductTable

Columns: 7 total

Add Source

Source settings

Source options

Projection

Optimize

Inspect

Data preview

Column name	Type
ProductKey	integer
ProductAltKey	string
ProductName	string
Color	string
Size	string
ListPrice	decimal
Discontinued	boolean

Synapse Pipeline

1 Hr 3 Min Remaining

Instructions

Resources

Help

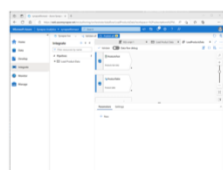
100%

- Server name: synapse*xxxxxx* (Synapse workspace)
- Database name: sqlxxxxxx
- SQL pool: sqlxxxxxx
- Authentication type: System Assigned Managed Identity
- Table name: dbo.DimProduct
- Import schema: From connection/store
- Allow schema drift: Selected

4. On the **Projection** tab for the new **ProductTable** source, Verify that the following data types are set:

- ProductKey: integer
- ProductAltKey: string
- ProductName: string
- Color: string
- Size: string
- ListPrice: decimal
- Discontinued: boolean

5. Verify that your data flow contains two sources, as shown here:



Add a Lookup

- Select the **+** icon at the bottom right of the **ProductsText** source and select **Lookup**.
- Configure the Lookup settings as follows:

- Output stream name: MatchedProducts
- Description: Matched product data

End >

Microsoft Azure | Synapse Analytics | synapse54mfppg

Integrate

Filter resources by name

Pipelines

Load Product Data

Files

SQL script 1

Load Product Data

LoadProductsData

ProductsText

ProductTable

Columns: 13 total

Add Source

Lookup settings

Match on *

Sort conditions *

Lookup conditions *

Synapse Pipeline

59 Minutes Remaining

Instructions

Resources

Help

100%

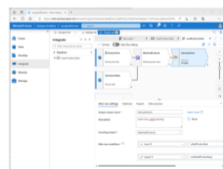
The lookup returns a set of columns from both sources, essentially forming an outer join that matches the **ProductID** column in the text file to the **ProductAltKey** column in the data warehouse table. When a product with the alternate key already exists in the table, the dataset will include the values from both sources. When the product does not already exist in the data warehouse, the dataset will contain NULL values for the table columns.

Add an Alter Row

- Select the **+** icon at the bottom right of the **MatchedProducts** Lookup and select **Alter Row**.
- Configure the alter row settings as follows:

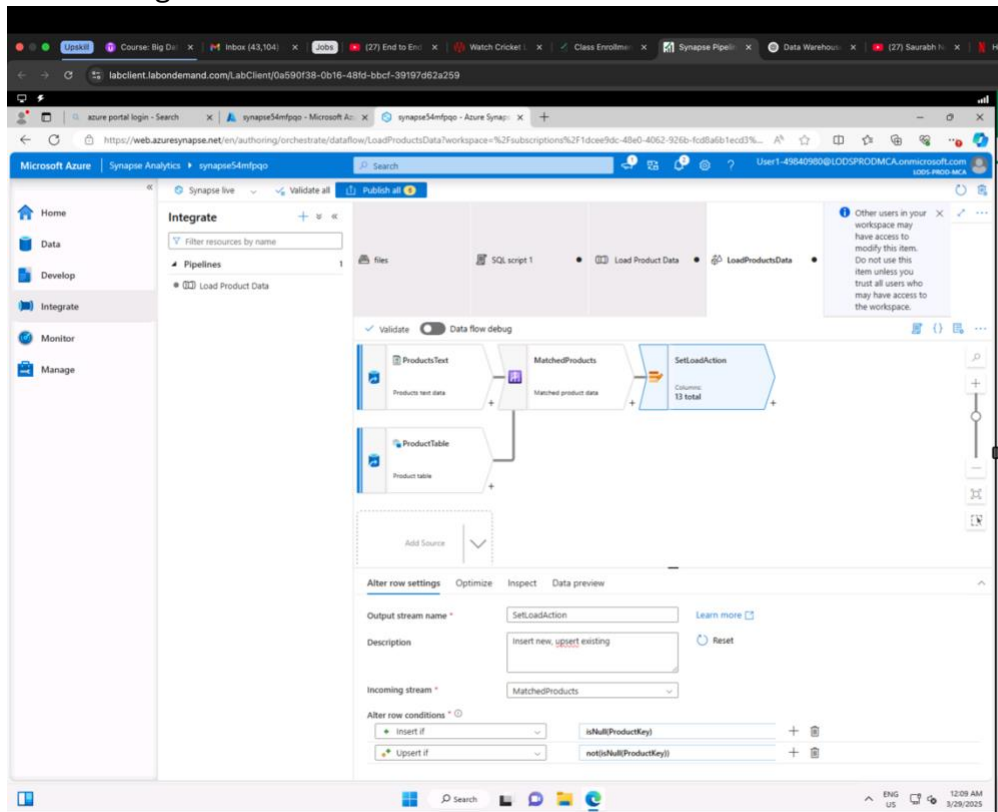
- Output stream name: SetLoadAction
- Description: Insert new, upsert existing
- Incoming stream: MatchedProducts
- Alter row conditions: Edit the existing condition and use the **+** button to add a second condition as follows (note that the expressions are case-sensitive):
 - InsertIf: `isNotNull(ProductKey)`
 - UpsertIf: `not(isNotNull(ProductKey))`

3. Verify that the data flow looks like this:



End >

Now adding an Alter row:



The screenshot shows the Microsoft Azure Synapse Analytics interface. The 'Integrate' tab is selected, and the 'Load Product Data' pipeline is being edited. The 'SetLoadAction' step is highlighted, and the 'Alter row settings' tab is active. The 'Output stream name' is 'SetLoadAction', the 'Description' is 'Insert new, upsert existing', and the 'Incoming stream' is 'MatchedProducts'. The 'Alter row conditions' section shows 'Insert if' and 'Upsert if' conditions.

Alter row settings

Output stream name: SetLoadAction

Description: Insert new, upsert existing

Incoming stream: MatchedProducts

Alter row conditions:

- Insert if: isNull(ProductKey)
- Upsert if: not(isNull(ProductKey))

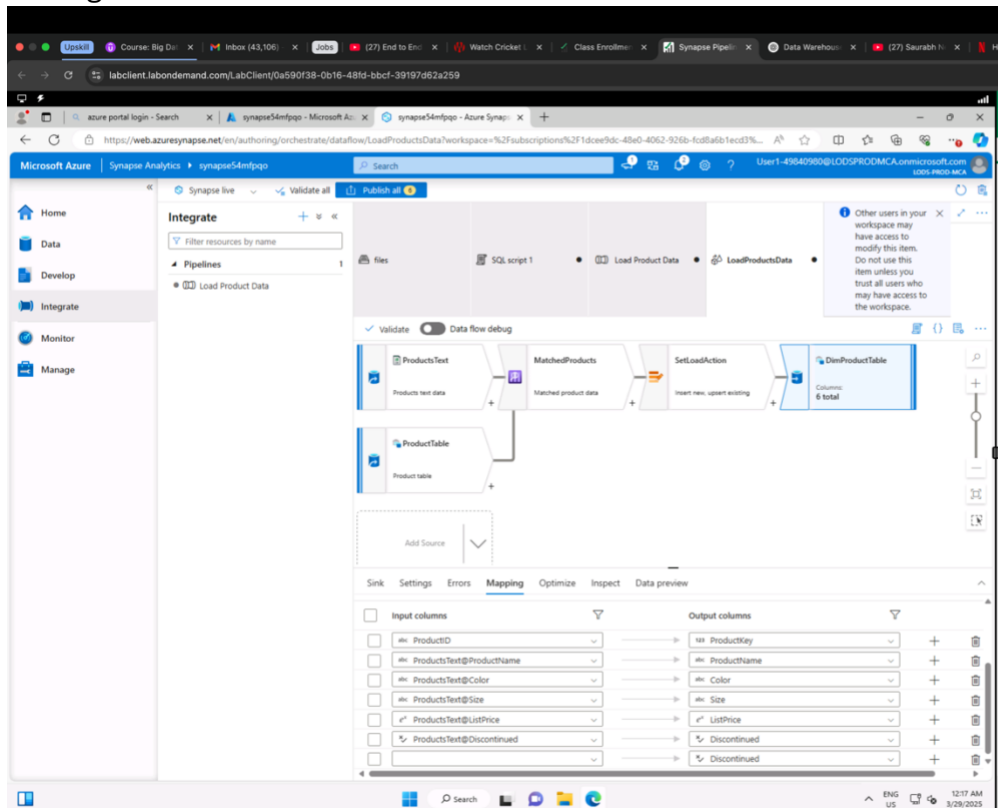
Instructions

The alter row step configures the kind of load action to perform for each row. Where there's no existing row in the table (the **ProductKey** is NULL), the row from the text file will be inserted. Where there's already a row for the product, an upsert will be performed to update the existing row. This configuration essentially applies a type 1 slowly changing dimension update.

Add a sink

- Select the + icon at the bottom right of the **SetLoadAction** alter row step and select **Sink**.
- Configure the **Sink** properties as follows:
 - Output stream name:** DimProductTable
 - Description:** Load DimProduct table
 - Incoming stream:** SetLoadAction
 - Sink type:** Integration dataset
 - Dataset:** DimProduct
 - Allow schema drift:** Selected
- On the **Settings** tab for the new **DimProductTable** sink, specify the following settings:
 - Update method:** Select **Allow insert** and **Allow Upsert**.
 - Key columns:** Select **List of columns**, and then select the **ProductID** column.
- On the **Mappings** tab for the new **DimProductTable** sink, clear the **Auto mapping** checkbox and specify the following column mappings:
 - ProductID: ProductID

Adding a sink:



The screenshot shows the Microsoft Azure Synapse Analytics interface. The 'Integrate' tab is selected, and the 'Load Product Data' pipeline is being edited. The 'SetLoadAction' step is highlighted, and the 'DimProductTable' sink is added. The 'Mapping' tab is active, showing the column mappings for the 'DimProductTable' sink.

Instructions

4. On the **Mappings** tab for the new **DimProductTable** sink, clear the **Auto mapping** checkbox and specify the following column mappings:

- ProductID: ProductID
- ProductsText@ProductName: ProductName
- ProductsText@Color: Color
- ProductsText@Size: Size
- ProductsText@ListPrice: ListPrice
- ProductsText@Discontinued: Discontinued

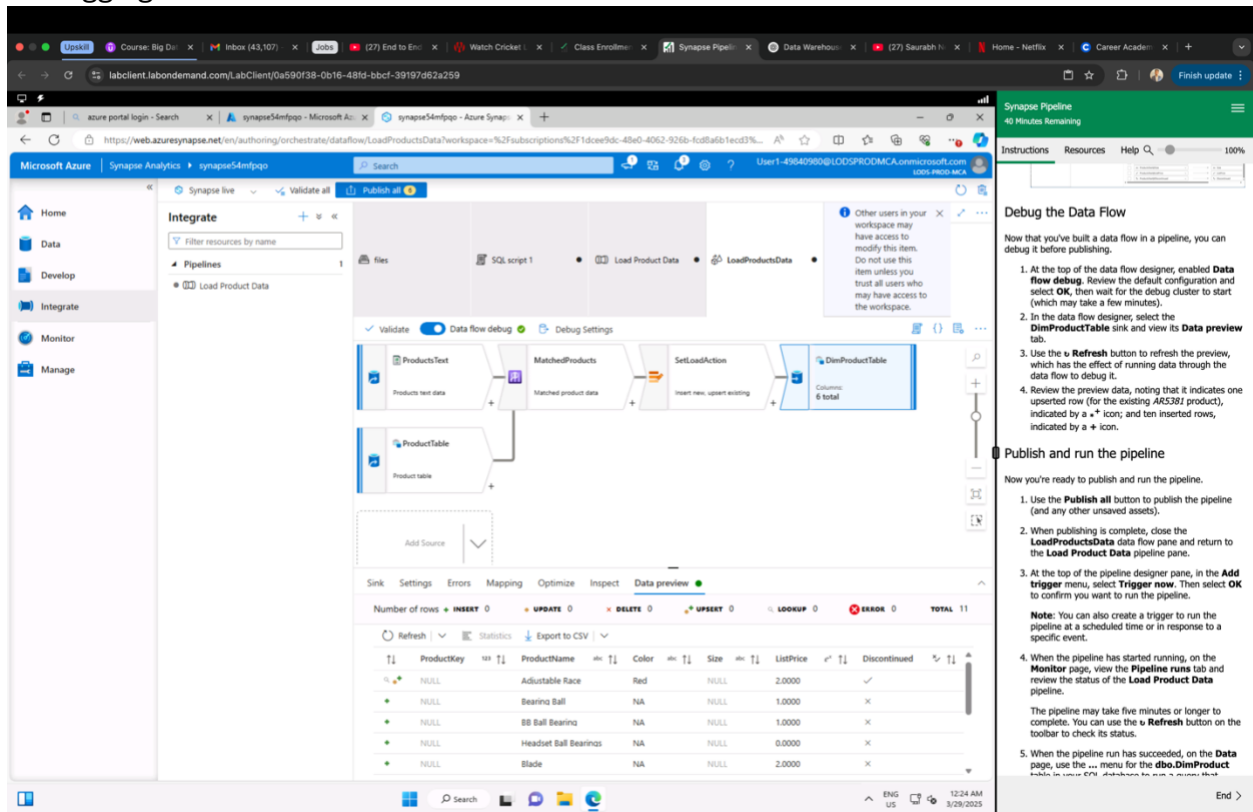
5. Verify that your data flow looks like this:

Debug the Data Flow

Now that you've built a data flow in a pipeline, you can debug it before publishing.

- At the top of the data flow designer, enabled **Data flow debug**. Review the default configuration and select **OK**, then wait for the debug cluster to start (which may take a few minutes).
- In the data flow designer, select the **DimProductTable** sink and view its **Data preview** tab.
- Use the **Refresh** button to refresh the preview, which has the effect of running the data through the data flow to debug it.
- Review the preview data, noting that it indicates one

Debugging the data flow:



Debug the Data Flow

Now that you've built a data flow in a pipeline, you can debug it before publishing.

- At the top of the data flow designer, enable **Data flow debug**. Review the default configuration and select **OK**, then wait for the debug cluster to start (which may take a few minutes).
- In the data flow designer, select the **DimProductTable** sink and view its **Data preview** tab.
- Use the **Refresh** button to refresh the preview, which has the effect of running data through the data flow to debug it.
- Review the preview data, noting that it indicates one upserted row (for the existing AR5381 product), indicated by a **+** icon; and ten inserted rows, indicated by a **+** icon.

Publish and run the pipeline

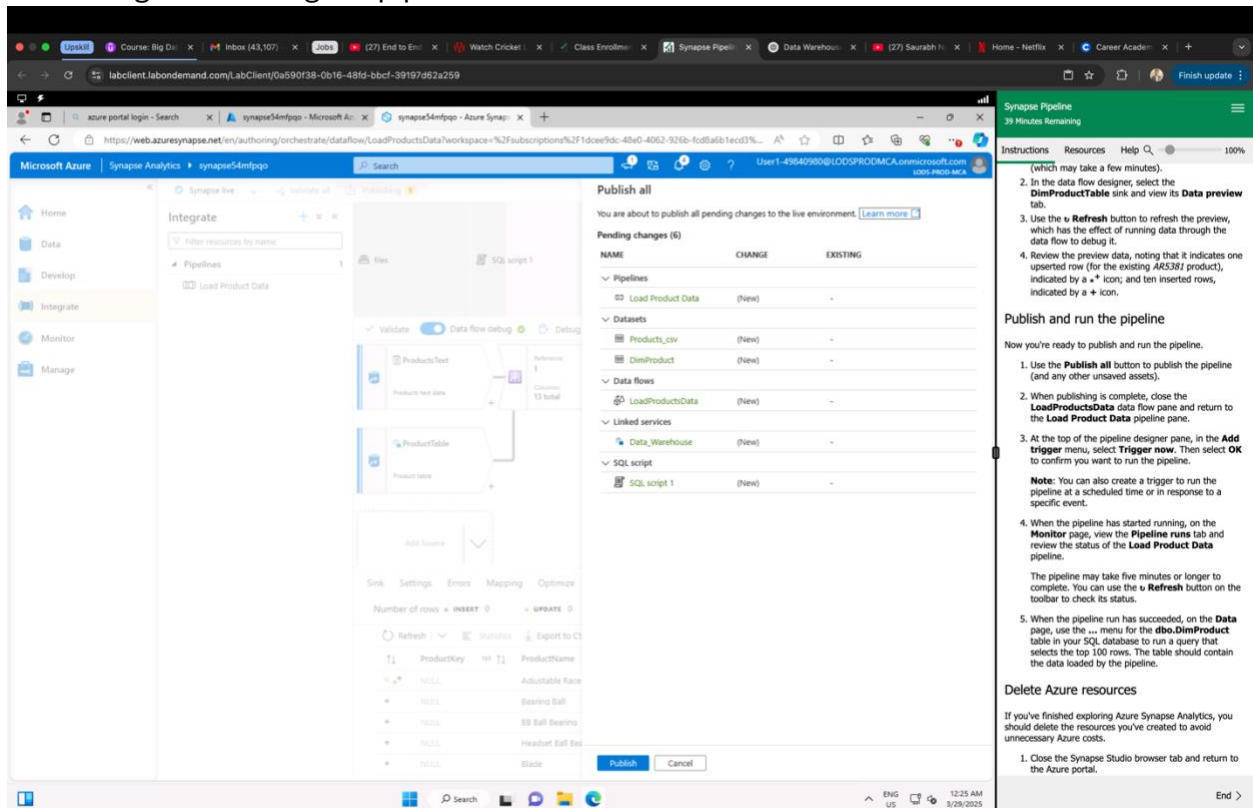
Now you're ready to publish and run the pipeline.

- Use the **Publish all** button to publish the pipeline (and any other unsaved assets).
- When publishing is complete, close the **LoadProductsData** data flow pane and return to the **Load Product Data** pipeline pane.
- At the top of the pipeline designer pane, in the **Add trigger** menu, select **Trigger now**. Then select **OK** to confirm you want to run the pipeline.
- Note: You can also create a trigger to run the pipeline at a scheduled time or in response to a specific event.
- When the pipeline has started running, on the **Monitor** page, view the **Pipeline runs** tab and review the status of the **Load Product Data** pipeline.

The pipeline may take five minutes or longer to complete. You can use the **Refresh** button on the toolbar to check its status.

5. When the pipeline run has succeeded, on the **Data** page, use the **...** menu for the **dbo.DimProductTable** in your SQL database to run a query that

Publishing and running the pipeline:



Publish all

You are about to publish all pending changes to the live environment. [Learn more](#)

Pending changes (6)

NAME	CHANGE	EXISTING
Pipelines		
Load Product Data	(New)	-
Datasets		
Products_csv	(New)	-
DimProduct	(New)	-
Data flows		
LoadProductsData	(New)	-
Linked services		
Data_Warehouse	(New)	-
SQL script		
SQL script 1	(New)	-

Publish and run the pipeline

Now you're ready to publish and run the pipeline.

- Use the **Publish all** button to publish the pipeline (and any other unsaved assets).
- When publishing is complete, close the **LoadProductsData** data flow pane and return to the **Load Product Data** pipeline pane.
- At the top of the pipeline designer pane, in the **Add trigger** menu, select **Trigger now**. Then select **OK** to confirm you want to run the pipeline.
- Note: You can also create a trigger to run the pipeline at a scheduled time or in response to a specific event.
- When the pipeline has started running, on the **Monitor** page, view the **Pipeline runs** tab and review the status of the **Load Product Data** pipeline.

The pipeline may take five minutes or longer to complete. You can use the **Refresh** button on the toolbar to check its status.

5. When the pipeline run has succeeded, on the **Data** page, use the **...** menu for the **dbo.DimProductTable** in your SQL database to run a query that selects the top 100 rows. The table should contain the data loaded by the pipeline.

Delete Azure resources

If you've finished exploring Azure Synapse Analytics, you should delete the resources you've created to avoid unnecessary Azure costs.

- Close the Synapse Studio browser tab and return to the Azure portal.

Microsoft Azure | Synapse Analytics | synapse54mfppo

Home

Data

Develop

Integrate

Monitor

Manage

Analytics pools

SQL pools

Apache Spark pools

Data Explorer pools (preview)

Activities

SQL requests

KQL requests

Apache Spark applications

Data flow debug

Integration

Pipeline runs

Trigger runs

Integration runtimes

Link connections

Pipeline runs

Triggered Debug Rerun Cancel options Refresh Edit columns List Gantt

Filter by run ID or name Pacific Time (US & C...) Last 24 hours Pipeline name: All Status: All Copy filters Export to CSV

Runs: Latest runs Triggered by: All Add filter

Showing 1 - 1 items Last refreshed 0 minutes ago

Pipeline name	Run start	Run end	Duration	Triggered by	Status
Load Product Data	3/29/2025, 12:26:50 AM	--	1m 5s	Manual trigger	In progress

Synapse Pipeline

36 Minutes Remaining

Instructions Resources Help 100%

which has the effect of running data through the data flow to debug it.

4. Review the preview data, noting that it indicates one upserted row (for the existing AR5381 product), indicated by a + icon; and ten inserted rows, indicated by a + icon.

Publish and run the pipeline

Now you're ready to publish and run the pipeline.

1. Use the **Publish all** button to publish the pipeline (and any other unsaved assets).

2. When publishing is complete, close the **LoadProductsData** data flow pane and return to the **Load Product Data** pipeline pane.

3. At the top of the pipeline designer pane, in the **Add trigger** menu, select **Trigger now**. Then select **OK** to confirm you want to run the pipeline.

Note: You can also create a trigger to run the pipeline at a scheduled time or in response to a specific event.

4. When the pipeline has started running, on the **Monitor** page, view the **Pipeline runs** tab and review the status of the **Load Product Data** pipeline.

The pipeline may take five minutes or longer to complete. You can use the **Refresh** button on the toolbar to check its status.

5. When the pipeline run has succeeded, on the **Data** page, use the ... menu for the **dbo.DimProduct** table in your SQL database to run a query that selects the top 100 rows. The table should contain the data loaded by the pipeline.

Delete Azure resources

If you've finished exploring Azure Synapse Analytics, you should delete the resources you've created to avoid unnecessary Azure costs.

1. Close the Synapse Studio browser tab and return to the Azure portal.

2. On the Azure portal, on the **Home** page, select **Resource groups**.

3. Select the **dp203-xxxxxxx** resource group for your Synapse Analytics workspace (not the managed resource group), and verify that it contains the Synapse workspace, storage account, and dedicated SQL pool for your workspace.

End >

Deleting the resources:

Microsoft Azure | Resource groups | dp203-54mfppo

Resource groups

dp203-54mfppo

Overview

Activity log

Access control (IAM)

Tags

Resource visualizer

Events

Settings

Cost Management

Monitoring

Automation

Help

Essentials

Resources

Recommendations

Filter for any field... Type equals all

Showing 1 to 3 of 3 records. Show hidden types

Name
datalake54mfppo
sql54mfppo (synapse54mfppo/sql54mfppo)
synapse54mfppo

Delete a resource group

The following resource group and all its dependent resources will be permanently deleted.

Resource group to be deleted

dp203-54mfppo

Dependent resources to be deleted (3)

All dependent resources, including hidden types, are shown.

datalake54mfppo

sql54mfppo

synapse54mfppo

Delete confirmation

Deleting this resource group and its dependent resources is a permanent action and cannot be undone.

Delete Go back

Enter resource group name to confirm deletion *

dp203-54mfppo

Delete Cancel

PS[cmd]Time

PS[JobType]Name

Output

Error

Progress

Verbose

Debug

Warning

Information

State

Script completed at 03/29/2025 06:29:55

PS /home/user1-49840980/dp-203/AllFiles/labn/10>

Synapse Pipeline

25 Minutes Remaining

Instructions Resources Help 100%

3. At the top of the pipeline designer pane, in the **Add trigger** menu, select **Trigger now**. Then select **OK** to confirm you want to run the pipeline.

Note: You can also create a trigger to run the pipeline at a scheduled time or in response to a specific event.

4. When the pipeline has started running, on the **Monitor** page, view the **Pipeline runs** tab and review the status of the **Load Product Data** pipeline.

The pipeline may take five minutes or longer to complete. You can use the **Refresh** button on the toolbar to check its status.

5. When the pipeline run has succeeded, on the **Data** page, use the ... menu for the **dbo.DimProduct** table in your SQL database to run a query that selects the top 100 rows. The table should contain the data loaded by the pipeline.

Delete Azure resources

If you've finished exploring Azure Synapse Analytics, you should delete the resources you've created to avoid unnecessary Azure costs.

1. Close the Synapse Studio browser tab and return to the Azure portal.

2. On the Azure portal, on the **Home** page, select **Resource groups**.

3. Select the **dp203-xxxxxxx** resource group for your Synapse Analytics workspace (not the managed resource group), and verify that it contains the Synapse workspace, storage account, and dedicated SQL pool for your workspace.

4. At the top of the **Overview** page for your resource group, select **Delete resource group**.

5. Enter the **dp203-xxxxxxx** resource group name to confirm you want to delete it, and select **Delete**.

After a few minutes, your Azure Synapse workspace resource group and the managed workspace resource group associated with it will be deleted.

End the lab

Please be sure to end the lab.

End >

Conclusion:

This lab provided valuable experience in building and managing a Synapse Pipeline within Azure Synapse Analytics. I learned how to create and configure a pipeline with data flow activities, set up source and destination data stores, and apply transformations such as adding sources, altering rows, and defining sinks. Additionally, I gained practical skills in debugging, publishing, and executing pipelines to ensure seamless data integration. This exercise enhanced my understanding of orchestrating data workflows and optimizing data processing within a cloud-based environment.