# **Creating a Tumbling Window Trigger**

**Table of Contents**

[ADF or Synapse Pipelines 2](#_Toc180495578)

[I. Create the Pipeline 2](#_Toc180495579)

[II. Alter Names 9](#_Toc180495580)

[III. Alter Copy data Defaults and Wildcard path 10](#_Toc180495581)

[IV. Get any old files 11](#_Toc180495582)

[V. Monitor Pipeline Run 12](#_Toc180495583)

[VI. View Data 12](#_Toc180495584)

[VII. Pause the Trigger 13](#_Toc180495585)

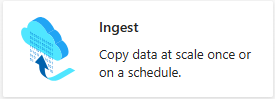
[Fabric Pipelines 15](#_Toc180495586)

[I. Create the Pipeline 15](#_Toc180495587)

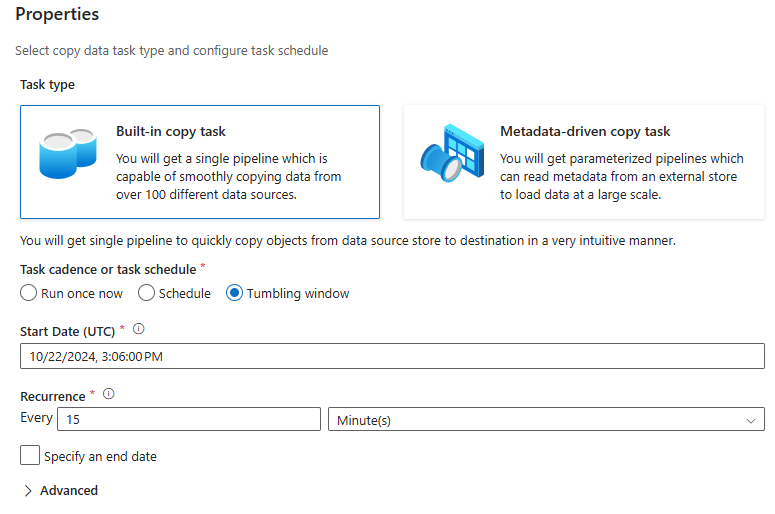
[II. View Data 19](#_Toc180495588)

## **ADF or Synapse Pipelines**

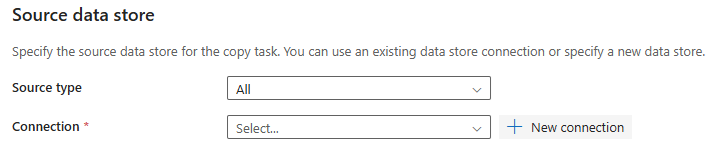
### **Create the Pipeline**



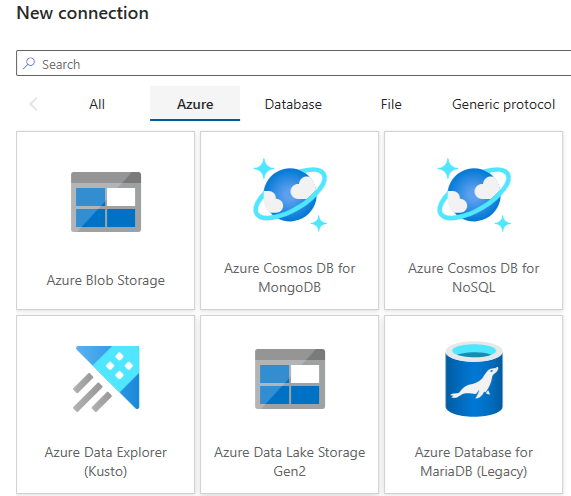
1. Go to the Home page of ADF
2. Click on Copy Data tool



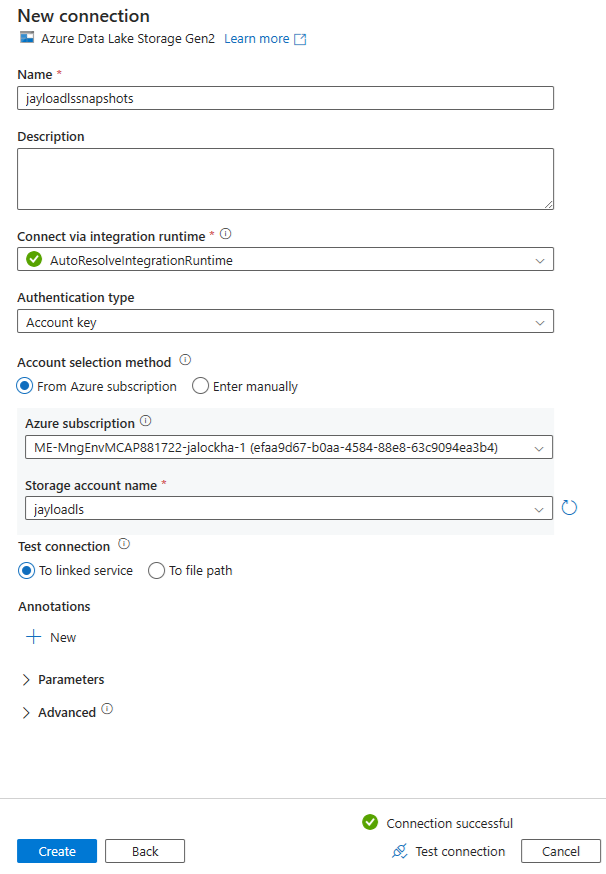
1. Select **Tumbling window**
2. Set a Start time that is 15-30 in the future (UTC time)
3. Set the recurrence as needed
4. Click **Next >**



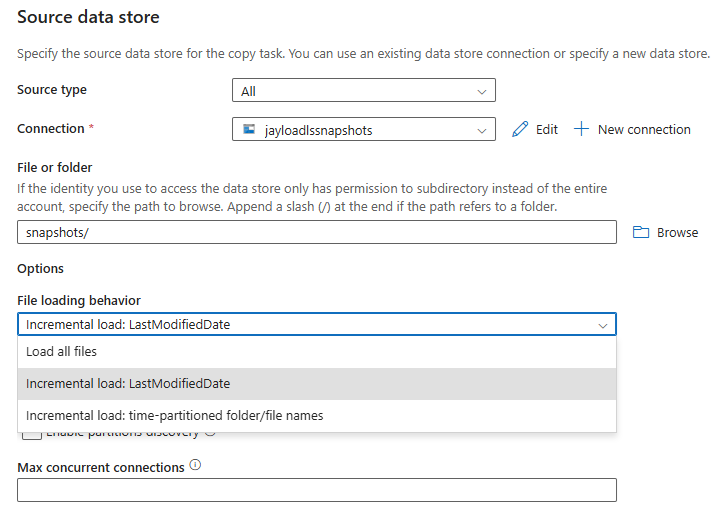
1. Select a Connection or click **+ New connection** to create a new one



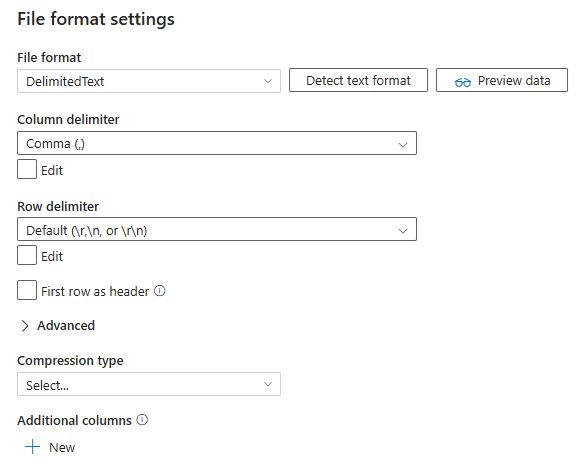
1. Select **Azure**, then **Azure Data Lake Storage Gen2** and click **Continue**



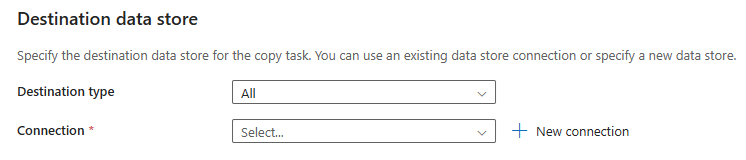
1. Give the connection a name
2. Leave the **Integration Runtime** as is
3. Choose your **Authentication type** and enter or select the appropriate settings
4. Leave Test connection as **To linked service**
5. Click **Test Connection**
6. If connection is successful, click **Create**



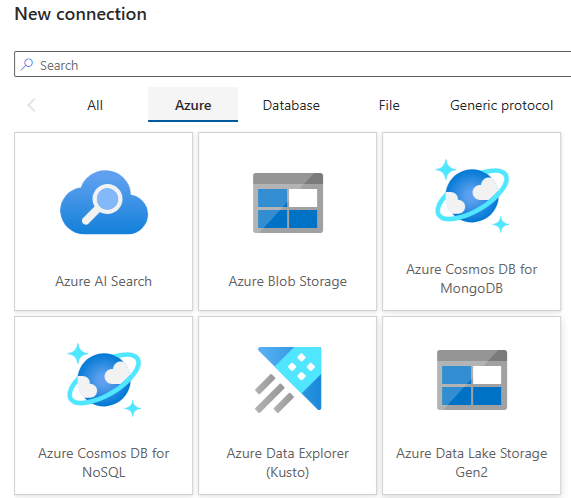
1. Click **Browse** and select the Container where your files will be located
2. Select **Incremental load: LastModifiedDate** as the File loading behavior
3. Leave other options as is
4. Click **Next >**



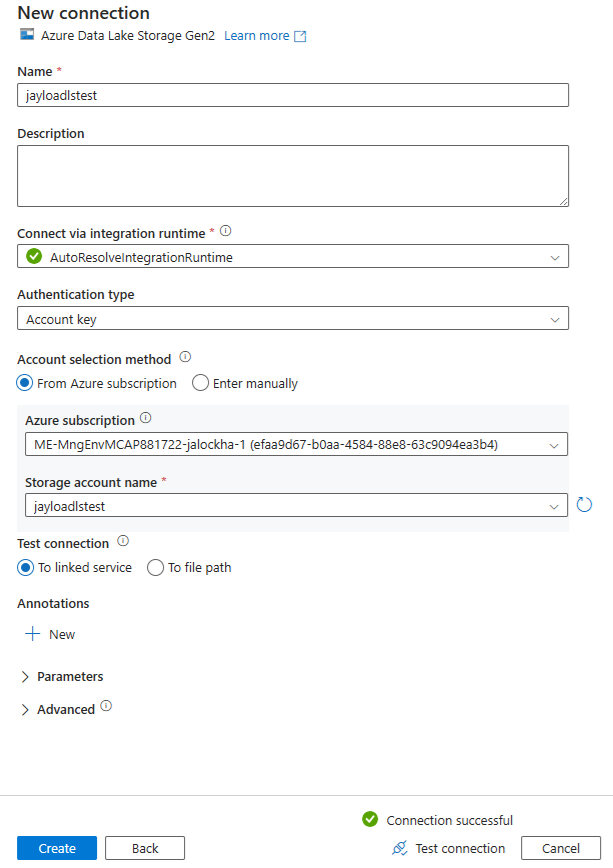
1. Choosethe appropriate **File format settings**
2. Click **Next >**



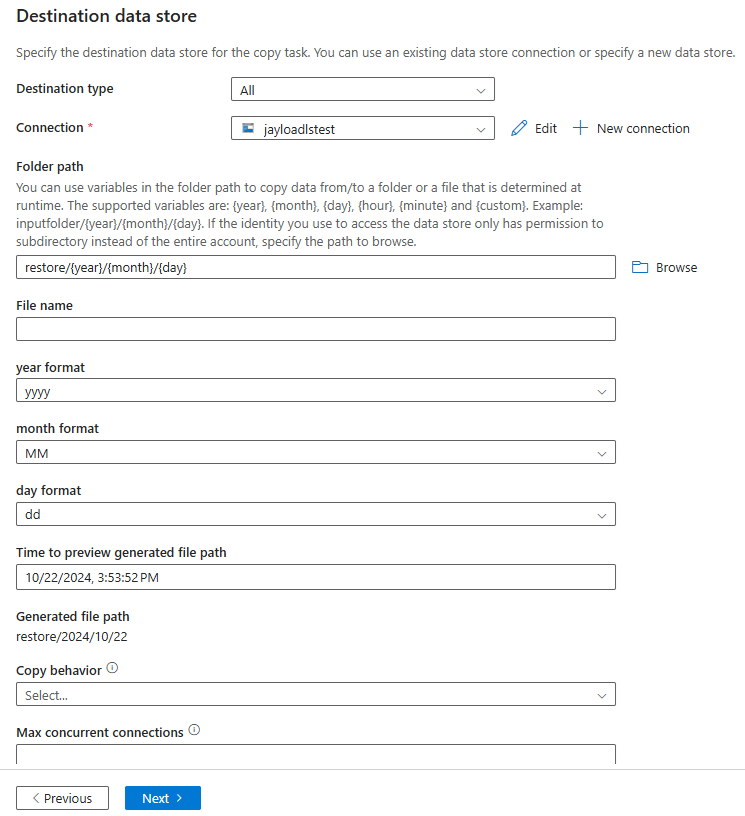
1. Select a Connection or click **+ New connection** to create a new one



1. Select **Azure**, then **Azure Data Lake Storage Gen2** and click **Continue**



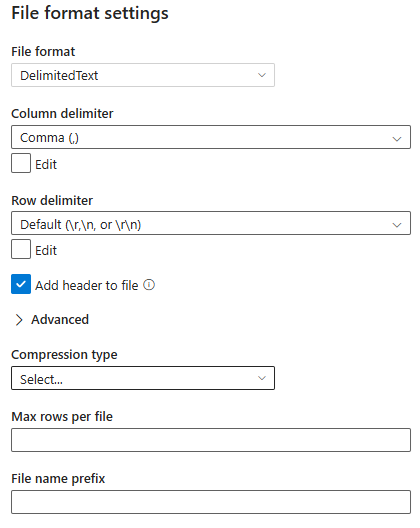
1. Give the connection a name
2. Leave the **Integration Runtime** as is
3. Choose your **Authentication type** and enter or select the appropriate settings
4. Leave Test connection as **To linked service**
5. Click **Test Connection**
6. If connection is successful, click **Create**



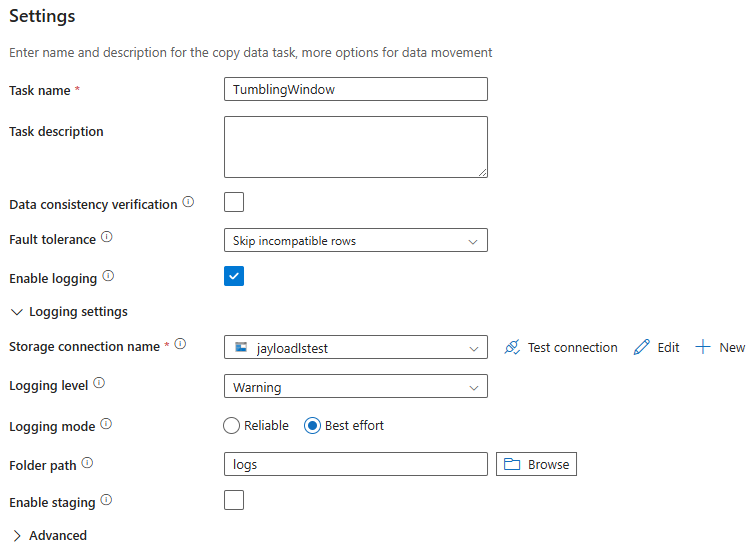
1. For the Output Folder, enter a folder name with the year, month and day parameter

*IE: restore/{year}/{month}/{day}*

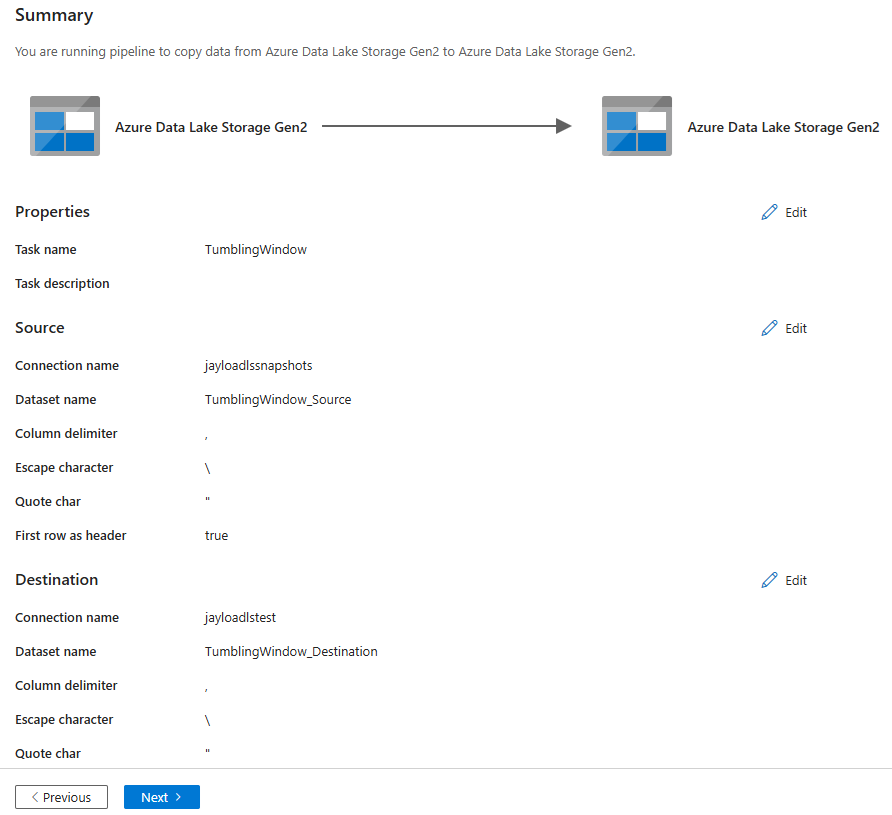
1. Change year format to **yyyy**
2. Leave everything else as is
3. Click **Next >**



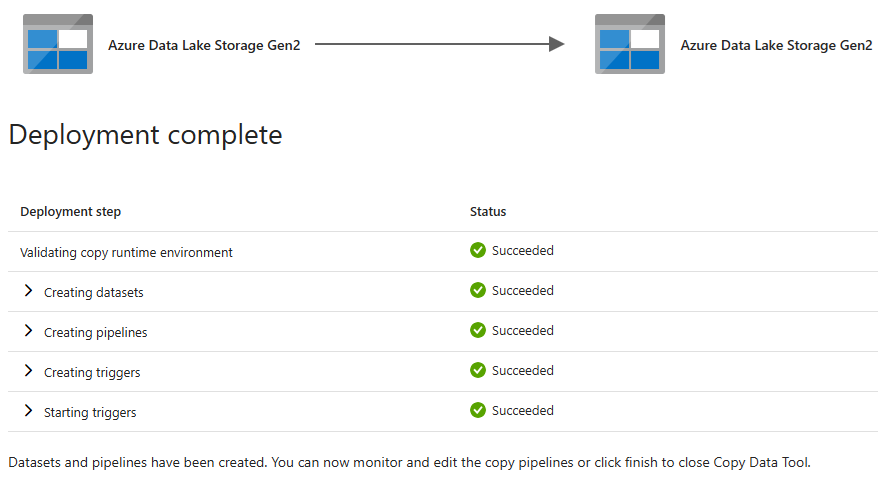
1. Leave file format info as is
2. Click **Next >**



1. Change the Task name if need be
2. Select **Skip incompatible rows** for Fault tolerance
3. Check **Enable logging**
4. Select your Data Lake Storage as the Storage connection name
5. Select **logs** as the Folder path
6. Click **Next >**



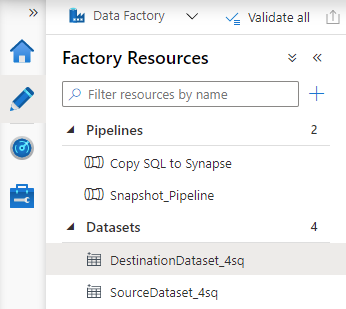
1. Change the source and destination Connections name if need be
2. Click **Next >**



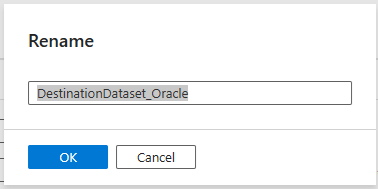
1. Wait for the deployment to complete, click **Edit pipeline**

## **Alter Names**

1. Click on the **Pencil**

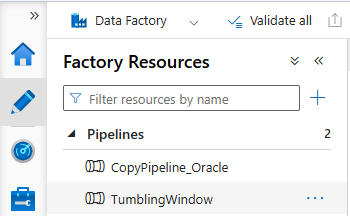


1. You can change the name of the Datasets to something more meaningful if you’d like

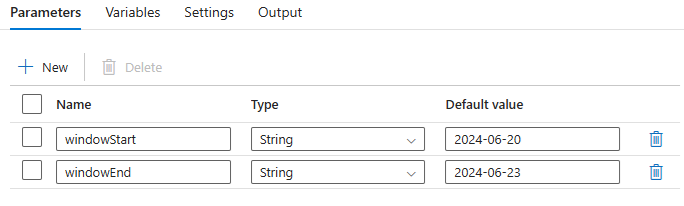


1. Clicking on the Dataset name and select the **…** and select **Rename**
2. Rename the dataset then click **OK**

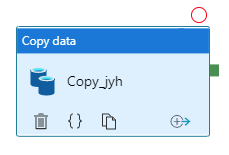
## **Alter Copy data Defaults and Wildcard path**



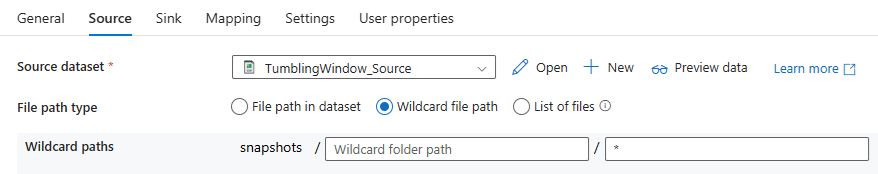
1. Click on the **Pencil**, then click the pipeline to edit it
2. Click anywhere in the white area around the Copy data activity



1. Change the **Default values** for the **windowStart** and **window**End



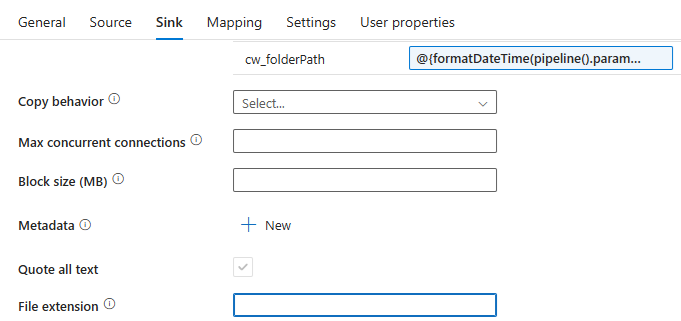
1. You can change the Wildcard paths if desired



1. Click **Source**, and change the **Wildcard path** to be an appropriate filter for the input files

*IE:* ***\*.csv***

1. Click on the **Sink** tab

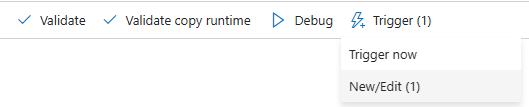


1. Remove the **.txt** in the File extension field

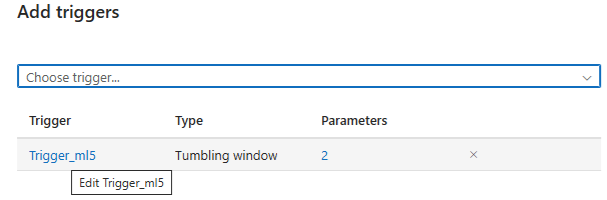


1. Click **Publish all**
2. Click **Publish**

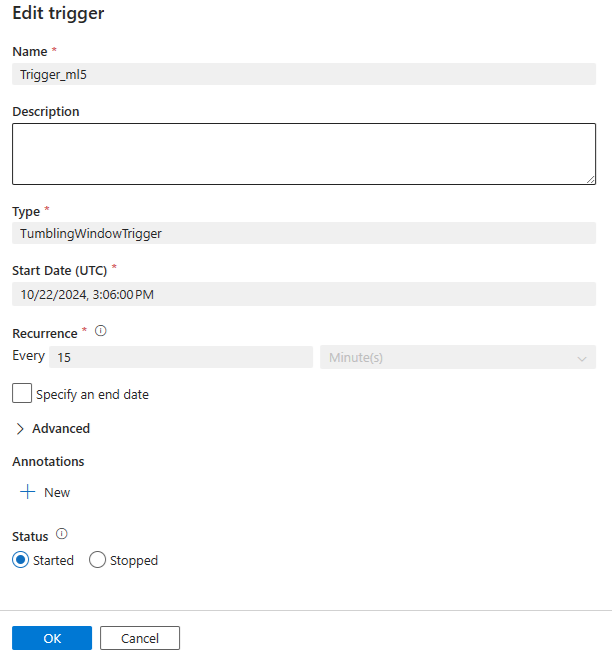
## **Get any old files**



1. Click **Trigger**, then **New/Edit (1)**



1. Click the trigger name

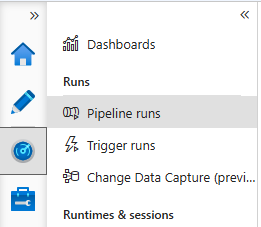


1. Change anything parameters

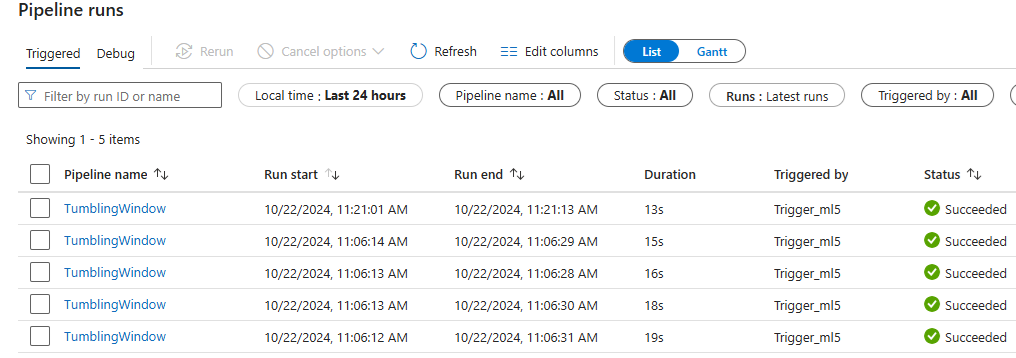
*FYI: if you need a different Type, Start Date or Recurrence…you will have to create a new Trigger*

1. Click **Ok**
2. Click **Close**
3. If you changed anything…click **Publish All** then **Publish**

## **Monitor Pipeline Run**



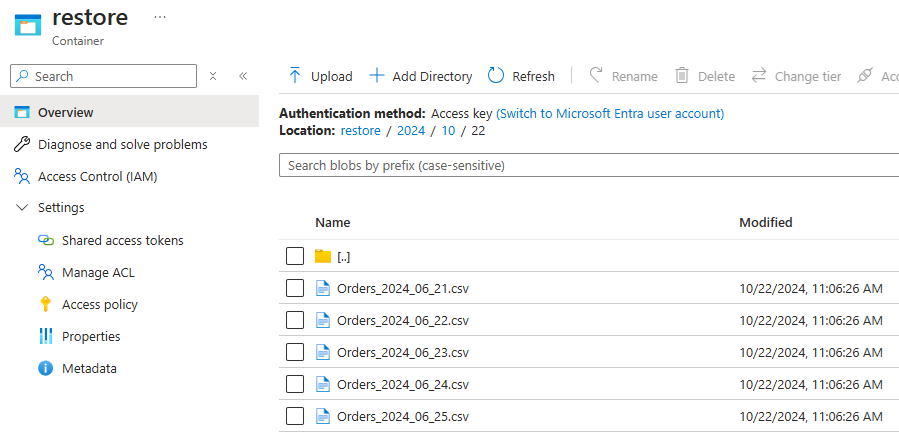
1. Click on the **Speedomator** then click **Pipeline runs**



1. You should see that your Pipeline has run several times successfully

## **View Data**

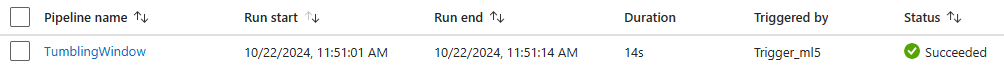
1. In the Azure Portal, go to your destination Data Lake Store



1. Click on the destination container and drill down to the files
2. Notice that folders were created for **Year**, **Month** and **Day**



1. Drop a new file into the source container
2. Go back to ADF and check the Pipeline runs. Then wait up to 15 minutes for the new run



1. Go back to the destination container under your storage account and drill down to the files



1. You should see that the new file has been copied

## **Pause the Trigger**

A screenshot of a computer

Description automatically generated

1. Click on the **briefcase**
2. Select **Triggers** under Author
3. Click **Stop**

A screenshot of a computer

Description automatically generated

1. Click **Publish all**

A screenshot of a computer

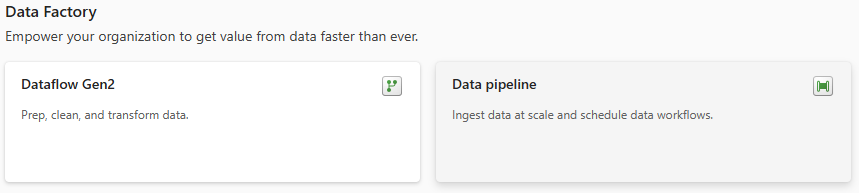
Description automatically generated

1. Click **Publish**

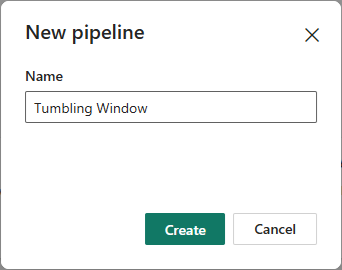
## **Fabric Pipelines**

### **Create the Pipeline**

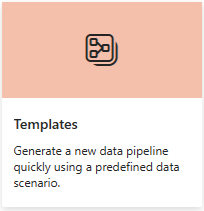
1. In **Fabric** open a Workspace, then click **+ New** and select **More options**



1. Under **Data Factory** select **Data pipeline**



1. Give your new Pipeline a name, then click **Create**

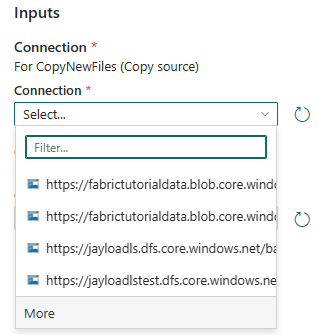


1. Click on **Templates**

A screenshot of a computer

Description automatically generated

1. Enter **lastMod** in the search field then select **Copy new files only by LastModifiedDate**



1. Under Connection select the source Connection or click **More** to create a new one

*FYI: Skip to* ***Step 15*** *if you selected one from the list*

A screenshot of a computer

Description automatically generated

1. Type **data lake** in the search bar and select **Azure Data Lake Storage Gen2**

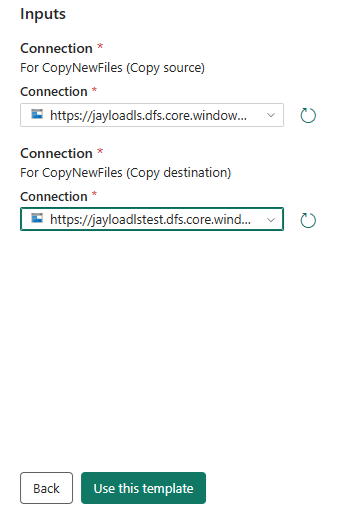
A screenshot of a computer

Description automatically generated

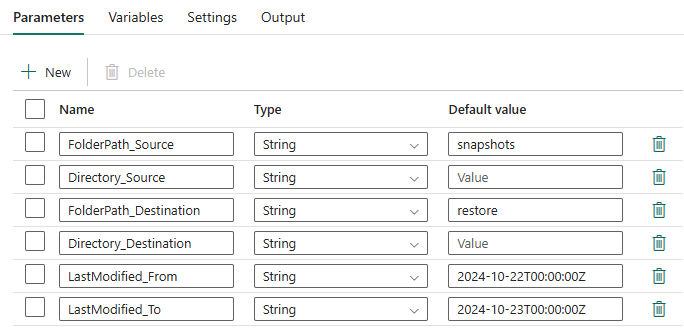
1. Enter the URL for your Data Lake Gen 2 container to copy from

*Found under* ***Properties*** *of the container*

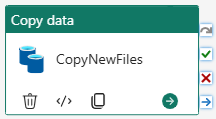
1. Change the work ‘**blob’** in the URL to ‘**dfs’**
2. Leave Connection as **Create new connection**
3. Change the Connection name if needed
4. Leave Data gateway as **(none)**
5. Select **Account Key** as the **Authentication kind**
6. Copy and paste **key 1** found under the Storage Account **Access keys** setting
7. Click **Next**
8. Repeat the above steps for the destination ADLS Gen2 container.



1. Once the **Copy source** and **Copy Destination** are selected, click **Use this template**



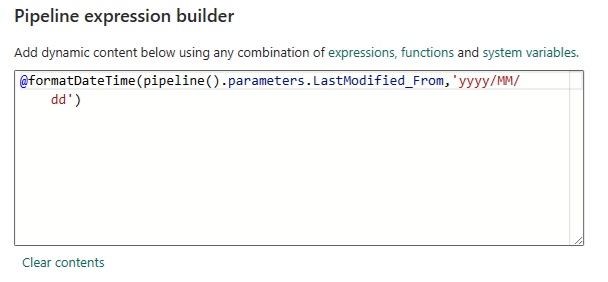
1. Change the **Parameters** to match your ADLS Gen2 folder structure
2. For Directory\_Source and Directory\_Destination change the default value to nothing



1. Click on the **Copy data** in activity in the middle of the screen



1. Click on the **Destination** tab
2. Double click the middle File path to open the Pileline expression builder



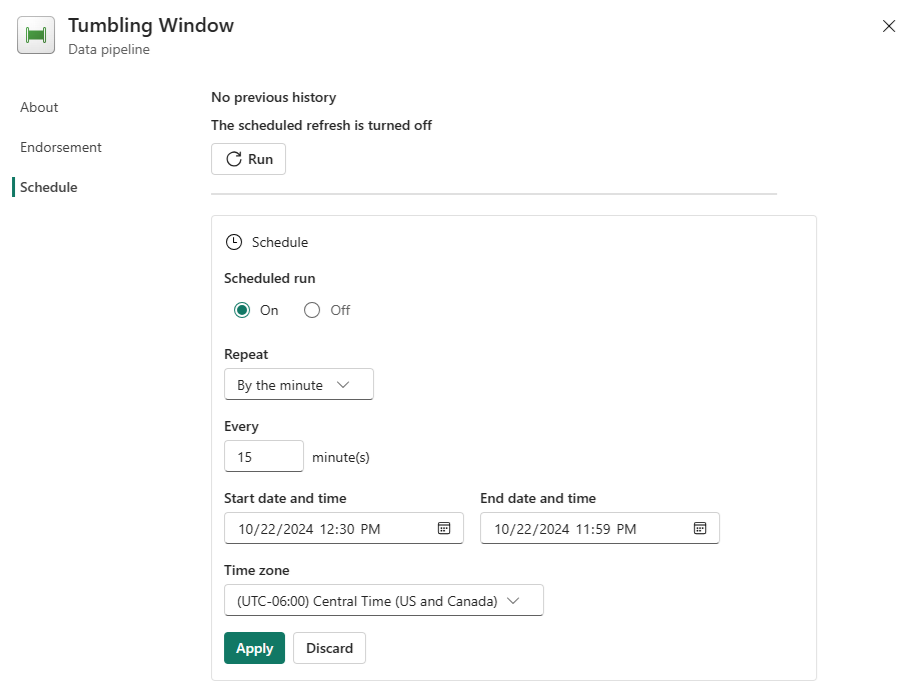
1. Change the content to the following:

@formatDateTime(pipeline().parameters.LastModified\_From,'yyyy/MM/dd')

1. Click **OK**
2. Click **Save**



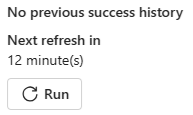
1. Click on **Add trigger** at the top



1. Select **On** for **Scheduled run**
2. Select the **Repeat frequency**
3. Select the **Every amount**
4. Enter or select the **Start date and time** and **End date and time**
5. Select the **Time zone**
6. Click **Apply**



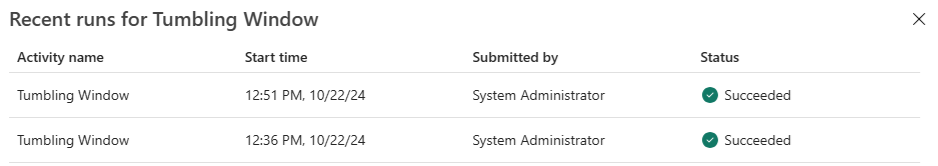
1. Click the **X** to close the Tumbling Window



1. Click **Add Trigger** again to see how long until the next run
2. Click **Save**
3. Go ahead and drop a new file into the source container so that the pipeline can pick it up
4. Wait until the time has passed



1. Click **View run history**



1. Click **Go to Monitor** to see more details

### **View Data**

1. In the Azure Portal, go to your destination Data Lake Store

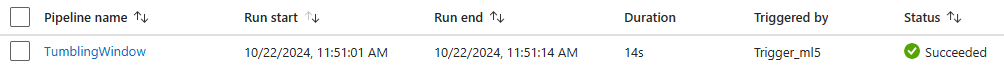
A screenshot of a computer

Description automatically generated

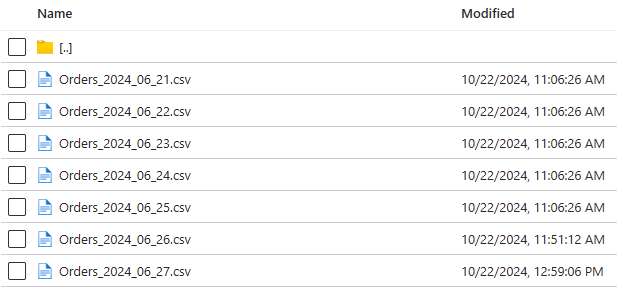
1. Click on the destination container and drill down to the files
2. Notice that folders were created for **Year**, **Month** and **Day**



1. Drop a new file into the source container
2. Go back to ADF and check the Pipeline runs. Then wait up to 15 minutes for the new run



1. Go back to the destination container under your storage account and drill down to the files



1. You should see that the new file has been copied