

Azure Data Factory Copy files of last 7 days

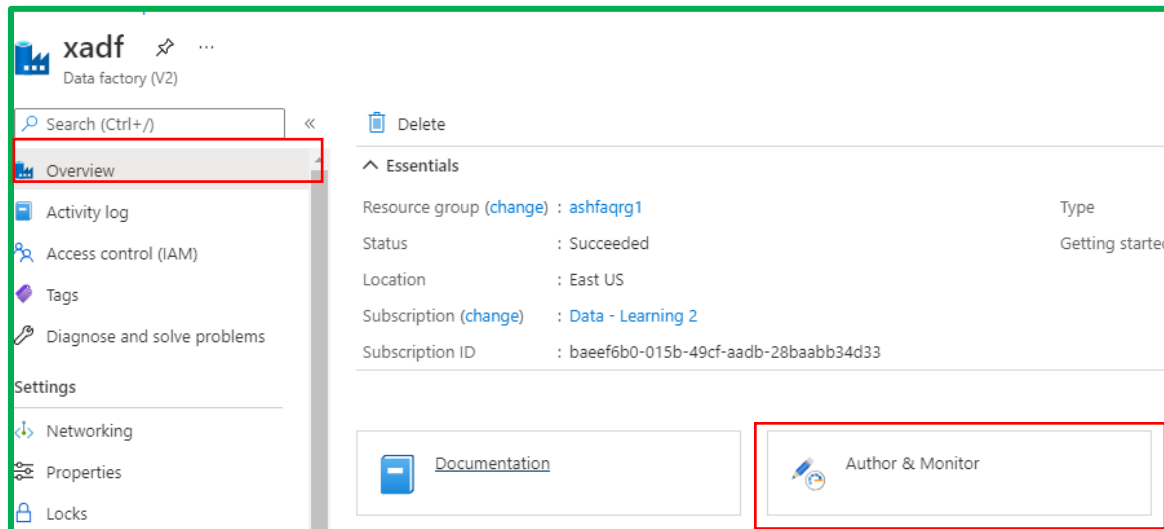
Statement: Copy files from a one container to another container based on X days [ex: copy last 7 days data] using ADF pipeline.

Pre-requisites

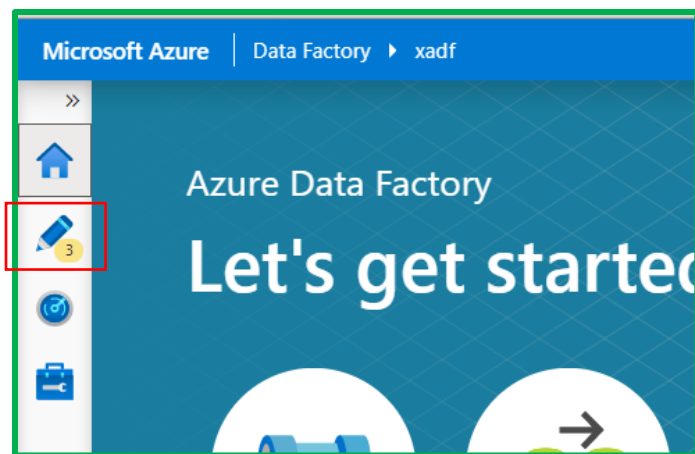
- **Need to have an azure subscription.**
- **Need a storage account create 2 containers with name.**
container1 name = 'input' adds some files
container2 name = 'output'
- **Create a ADF Resource using azure portal.**

Following are the Steps.

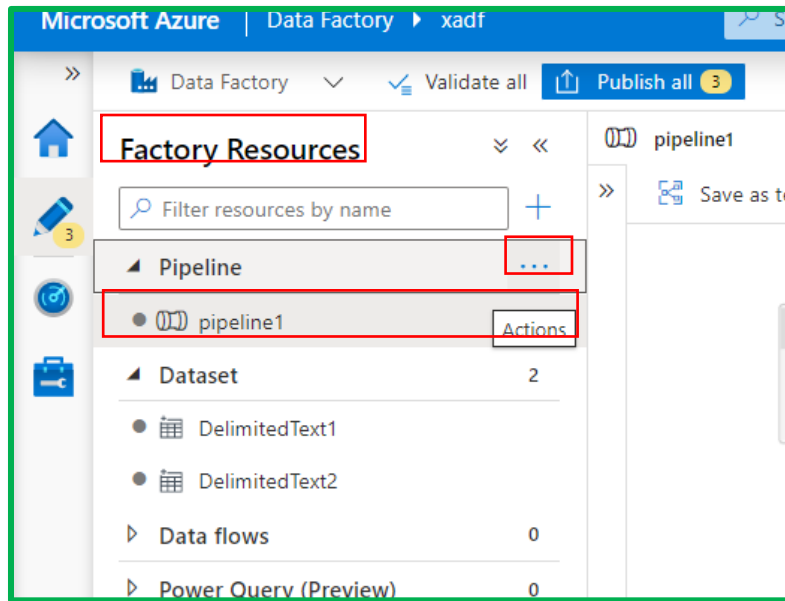
Step 1 Open Data factory Select Author & Monitor from Overview! Click on Author & Monitor .



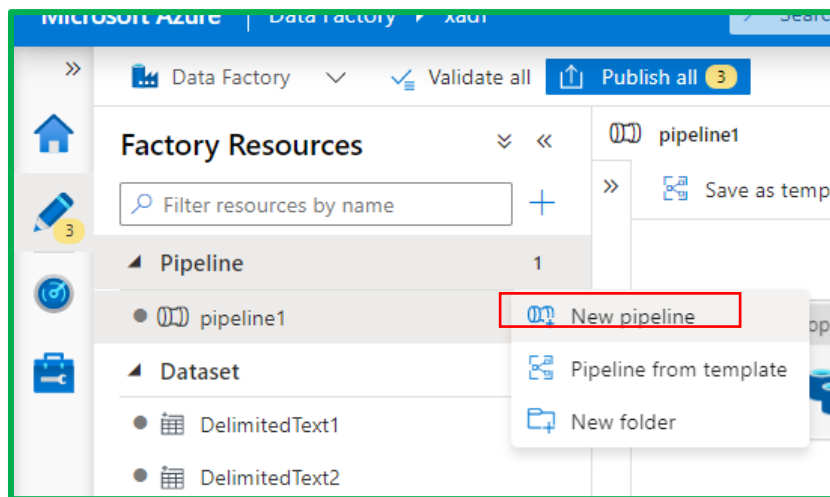
Step 2: Click on Author (Pencil icon).



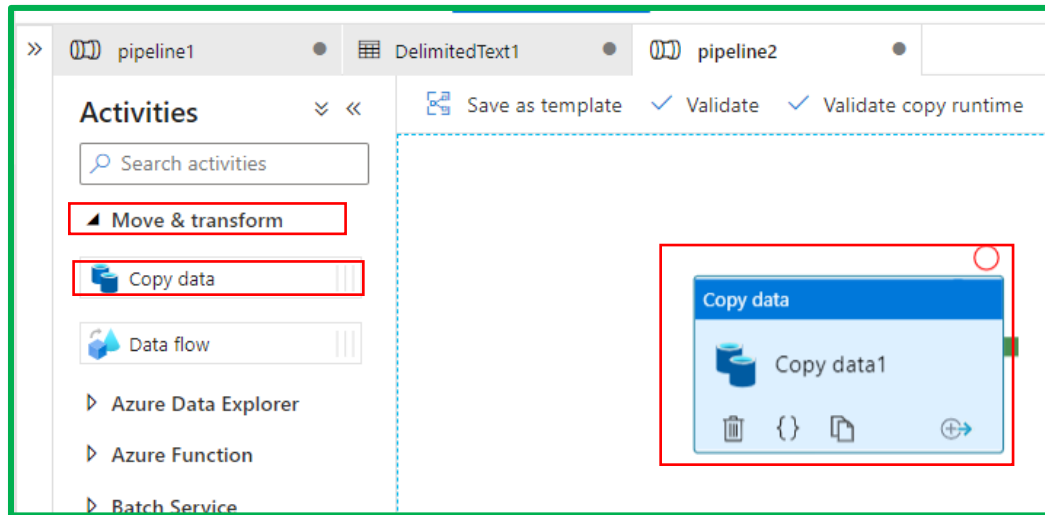
Step 3: Go to Factory Resources > Pipeline > Click on 3 dot .



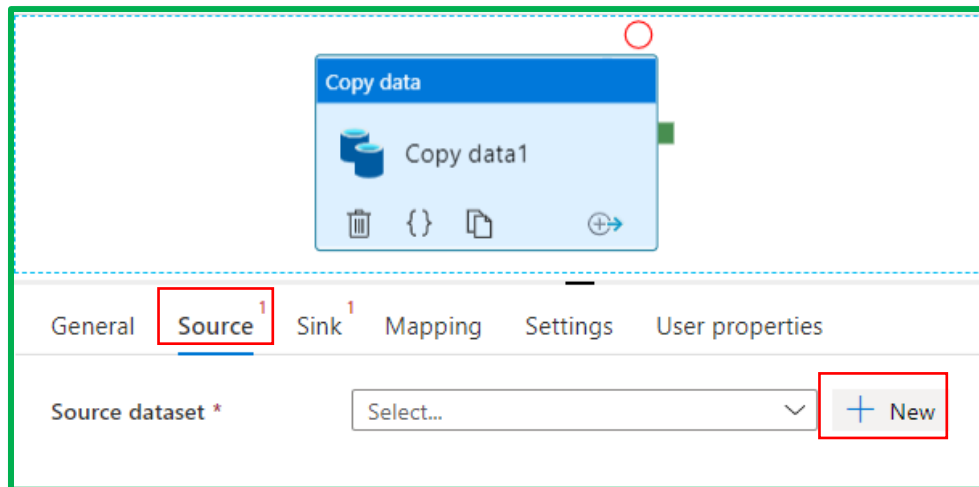
3.1 Create a new pipeline.



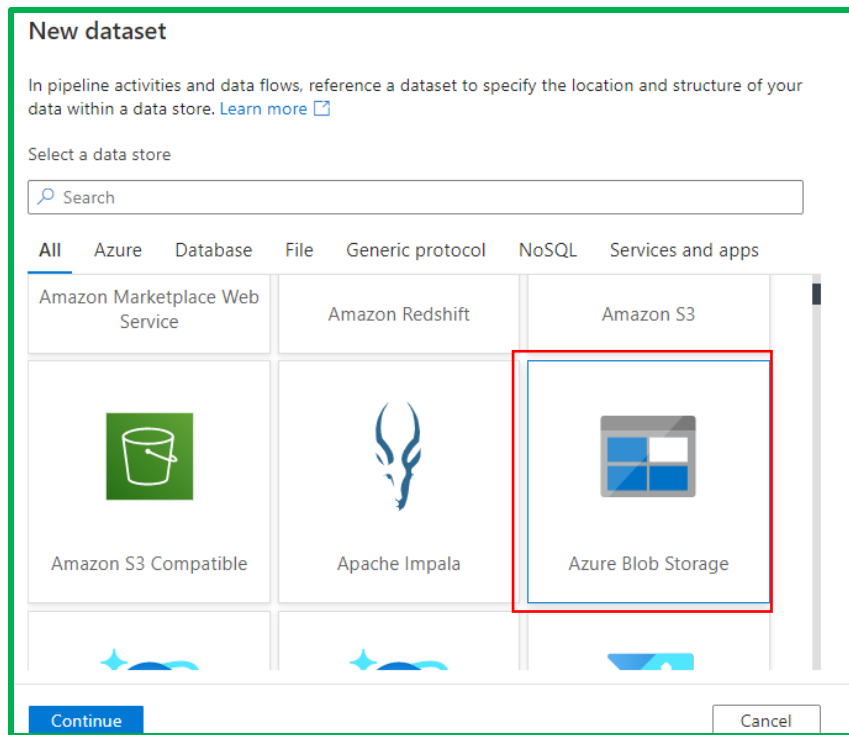
Step 4: Pipeline go to Activities > move & transform > drag and drop copy Activity.



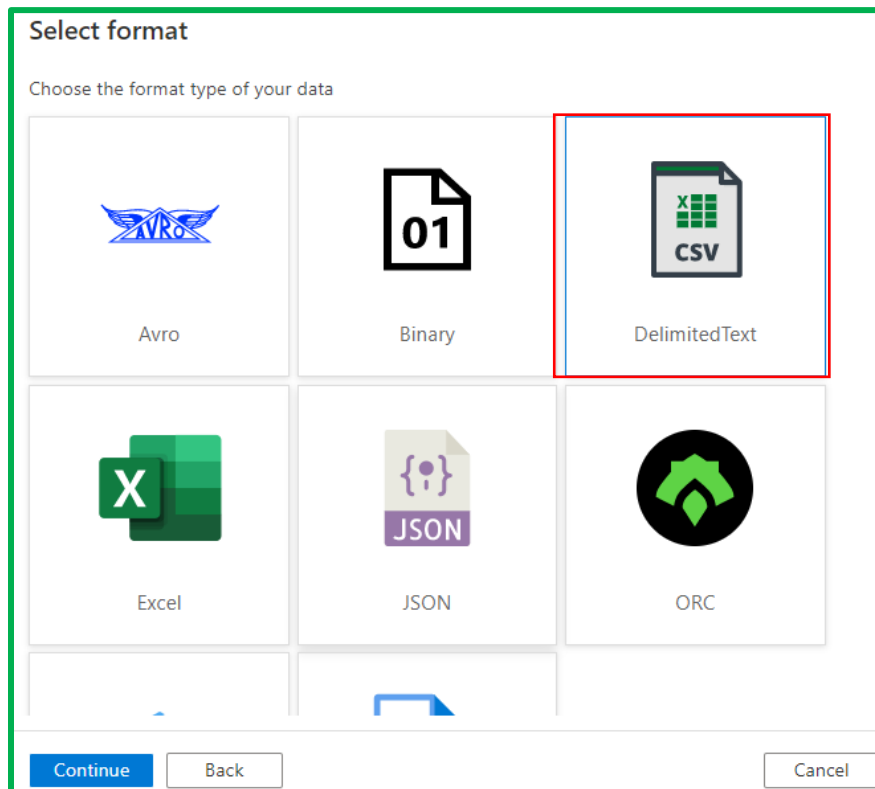
Step 5: Click on **Source** in that add new source data click on **New** option.



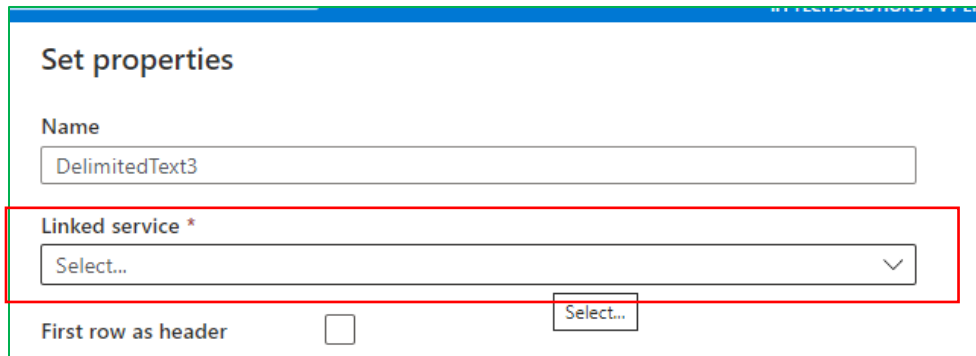
Step 6: Select Azure Blob Storage.



Step 7: Select **DelimitedText**.



Step 8: Add a **linked service** add a new service.



Set properties

Name

DelimitedText3

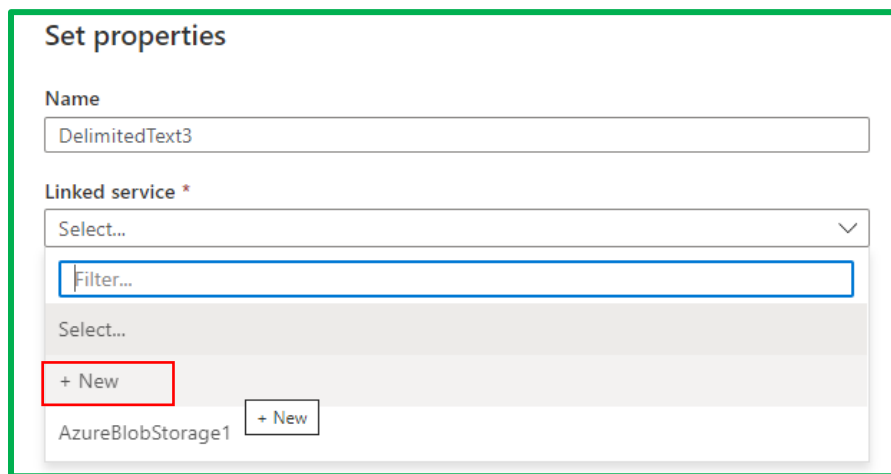
Linked service *

Select...

First row as header ☐

Select...

8.1 Click on **NEW**.



Set properties

Name

DelimitedText3

Linked service *

Select...

Filter...

Select...

+ New

AzureBlobStorage1 + New

Step 9: In this step only, add.

- Azure subscription = add your azure subscription in which you have your storage account
- Storage account name = Storage account we have created using azure portal in a resource group
- Click on Create Button at bottom of screen.

New linked service (Azure Blob Storage)

Authentication method
Account key

Connection string Azure Key Vault

Account selection method ⓘ
☒ From Azure subscription ☐ Enter manually

Azure subscription ⓘ
Data - Learning 2 (baef6b0-015b-49cf-aadb-28baabb34d33)

Storage account name *
sdrq01

Additional connection properties
+ New

Test connection ⓘ
☒ To linked service ☐ To file path

Annotations

Create Test connection Cancel

Step 10: Select Container.

Set properties

Name
DelimitedText3

Linked service *
AzureBlobStorage1

File path
Container / Directory / File

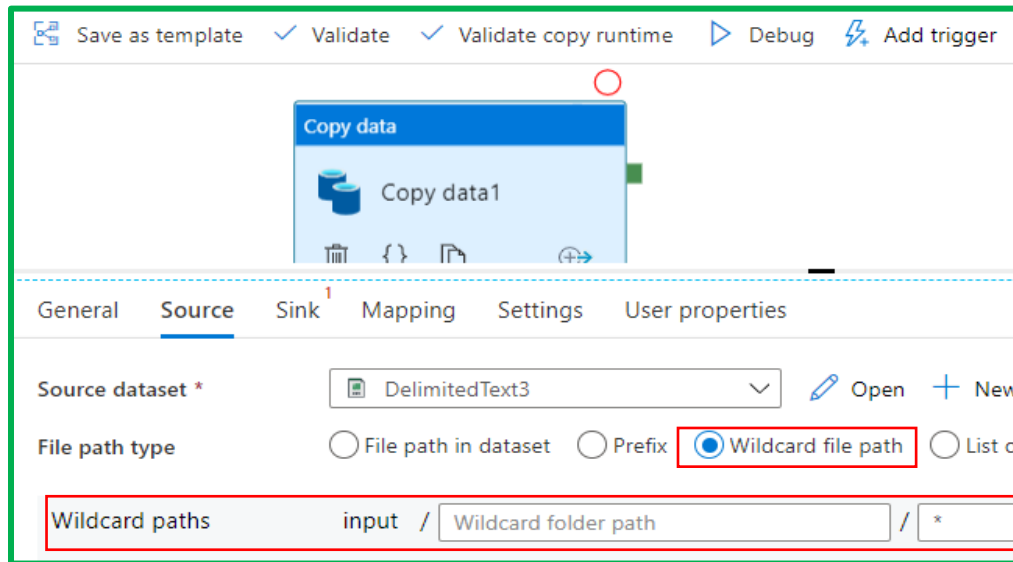
First row as header ☐

Import schema
☐ From connection/store ☐ From sample file ☒ None

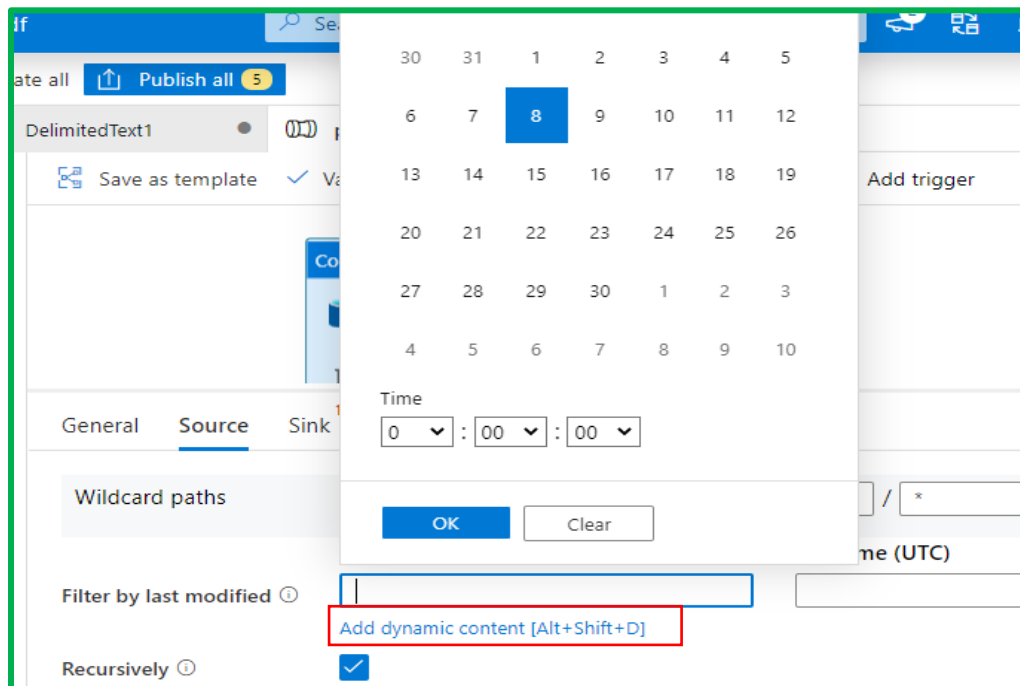
Advanced

Step 11: Now Configure Source of Copy activity

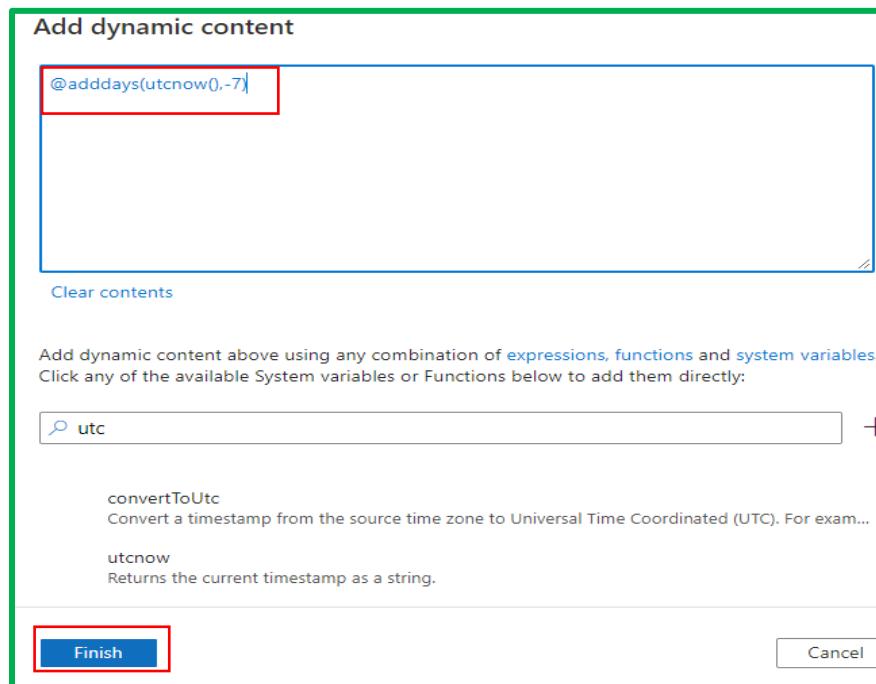
Select Wildcard path from file path type add extension (*) This will take all the files in the selected Container.



11.1 Now Configure Filter by last modified. Click on [Add dynamic content \[Alt+Shift+D\]](#).



11.2 Add this syntax to the Content. `@adddays(utcnow(),-7)` Click Finish.



Add dynamic content

`@adddays(utcnow(),-7)`

[Clear contents](#)

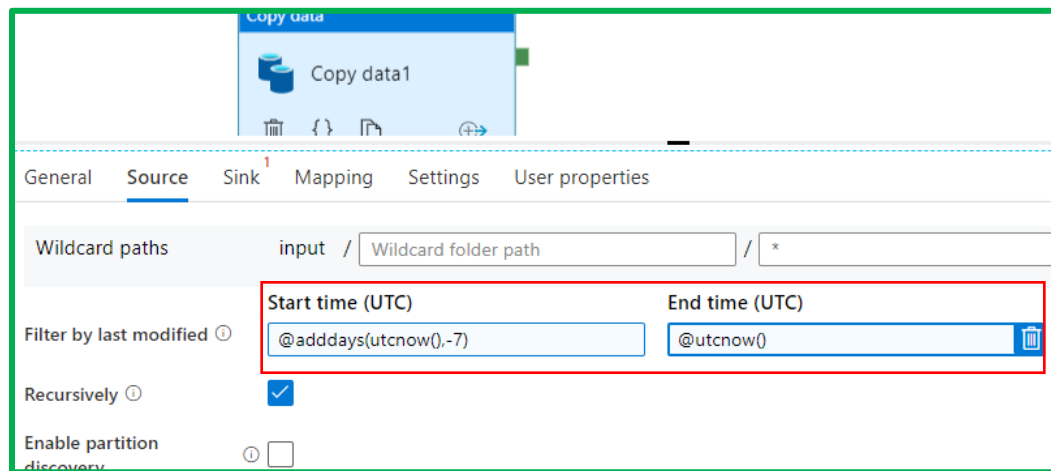
Add dynamic content above using any combination of [expressions](#), [functions](#) and [system variables](#). Click any of the available System variables or Functions below to add them directly:

Search: +

- convertToUtc**
Convert a timestamp from the source time zone to Universal Time Coordinated (UTC). For exam...
- utcnow**
Returns the current timestamp as a string.

Finish Cancel

11.3 For the End Time add `@utcnow()`.



Copy data

Copy data1


General **Source** Sink Mapping Settings User properties

Wildcard paths: input / / *

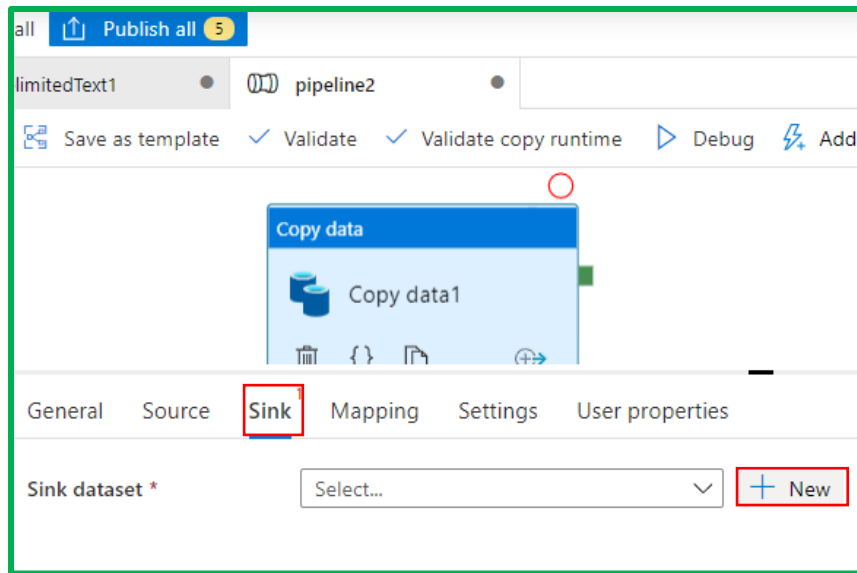
Filter by last modified ⓘ

Recursively ⓘ ☒

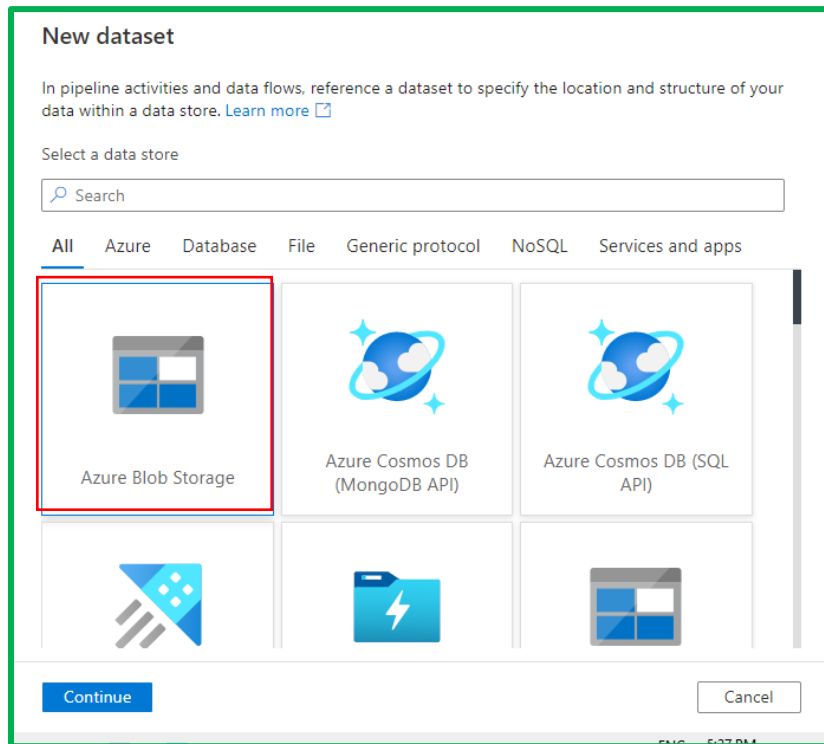
Enable partition discovery ⓘ ☐

Start time (UTC) **End time (UTC)** 

Step 12: Configure the Sink for copy activity.









12.1 Click on New Button from Sink dataset Select Azure Blob Storage.



12.2 Select Delimited Text Click on Continue.

Select format

Choose the format type of your data

 Avro	 Binary	 DelimitedText
 JSON	 ORC	 Parquet


Continue Back Cancel


12.3 Select Output option and click on ok button.

Browse

Select a file or folder.

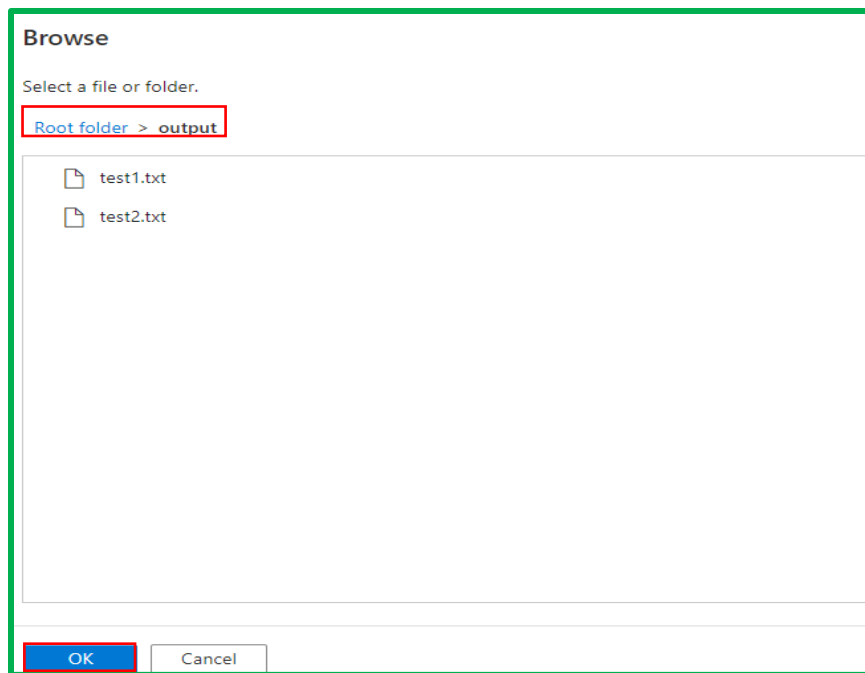
Root folder

 input

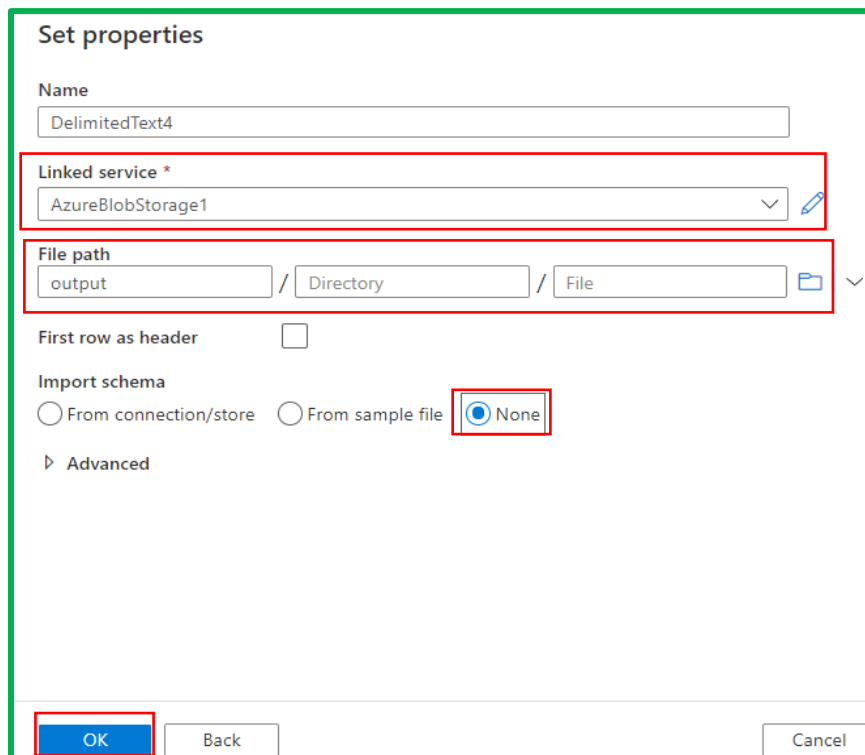
 output

OK Cancel

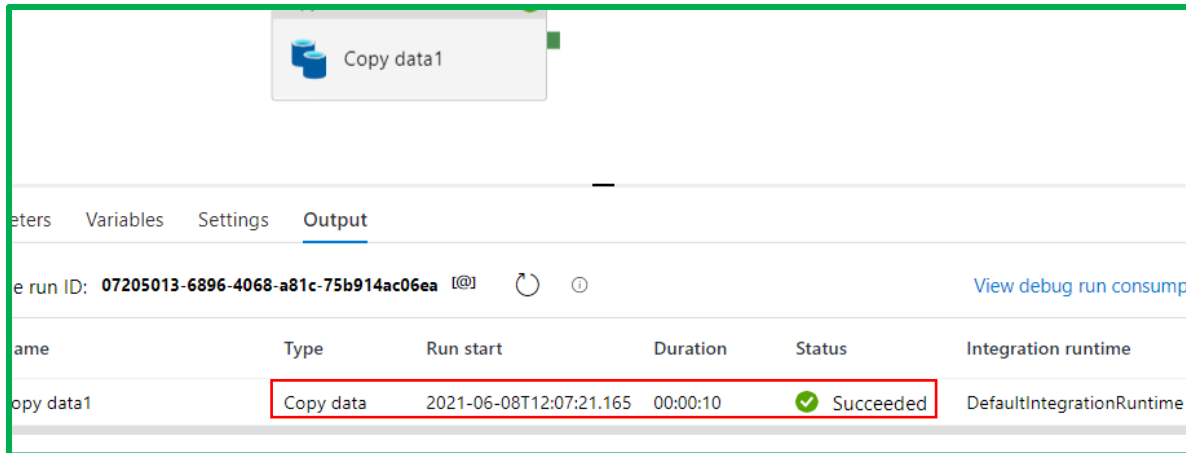
12.4 Open the container and click ok button.



12.5 Select None option from import schema.



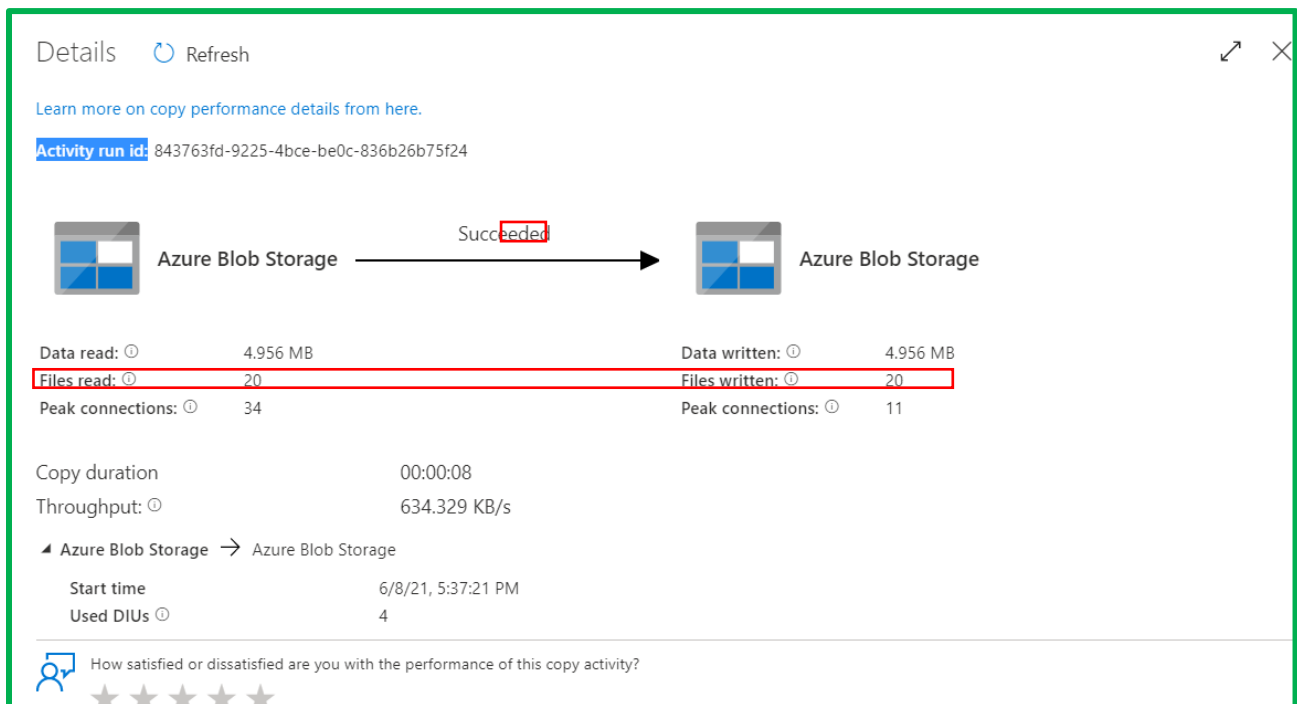
Step 13: Debug and check the operation status.



Name	Type	Run start	Duration	Status	Integration runtime
Copy data1	Copy data	2021-06-08T12:07:21.165	00:00:10	✓ Succeeded	DefaultIntegrationRuntime

Output:

As we can see in the below image 20 files have been copied from input container to output container.



Details Refresh

Learn more on copy performance details from here.

Activity run id: 843763fd-9225-4bce-be0c-836b26b75f24

Azure Blob Storage → Succeeded → Azure Blob Storage

Data read: ⓘ	4.956 MB	Data written: ⓘ	4.956 MB
Files read: ⓘ	20	Files written: ⓘ	20
Peak connections: ⓘ	34	Peak connections: ⓘ	11

Copy duration: 00:00:08
Throughput: ⓘ 634.329 KB/s

▲ Azure Blob Storage → Azure Blob Storage

Start time: 6/8/21, 5:37:21 PM
Used DIUs ⓘ: 4

How satisfied or dissatisfied are you with the performance of this copy activity?

★ ★ ★ ★ ★