

Home

Author

Monitor

Manage

Learning Center

Data Factory

Validate all

Publish all

General

Factory settings

Connections

Linked services

Integration runtimes

Microsoft Purview

Source control

Git configuration

ARM template

Author

Triggers

Global parameters

Data flow libraries

Security

Credentials

Customer managed key

Outbound rules

Managed private endpoi...

## Linked services

Linked service defines the connection information to a data store or compute service.

+ New

Filter by name Annotations : Any

### New linked service

Azure Blob Storage [Learn more](#)

Name \* ls\_azblob\_ercovidreportingsa

Description

Connect via integration runtime \* ⓘ AutoResolveIntegrationRuntime

Authentication type Account key

Connection string Azure Key Vault

Account selection method ⓘ

From Azure subscription  Enter manually

Azure subscription ⓘ

Azure for Students (50381b31-77bb-4cc2-a6f1-156c0ef3f834)

Storage account name \* ercovidreportingsa

Connection successful

Test connection Cancel

If you expected to see result

Home

Author

Monitor

Manage

Learning Center

Data Factory

Validate all

Publish all

## Linked services

Linked service defines the connection information to a data store or compute

New

Filter by name

Annotations : Any

Showing 1 - 1 of 1 items

Name ↑	Type ↑
ls_azblob_ercovidreportingsa	Azure Blob Storage

### New linked service

Azure Data Lake Storage Gen2 [Learn more](#)

Name \*  
ls\_azdl\_ercovidreportingdatalake

Description

Connect via integration runtime \* ⓘ  
AutoResolveIntegrationRuntime

Authentication type  
Account key

Account selection method ⓘ  
 From Azure subscription  Enter manually

Azure subscription ⓘ  
Azure for Students (50381b31-77bb-4cc2-a6f1-156c0ef3f834)

Storage account name \*  
ercovidreportingdatalake

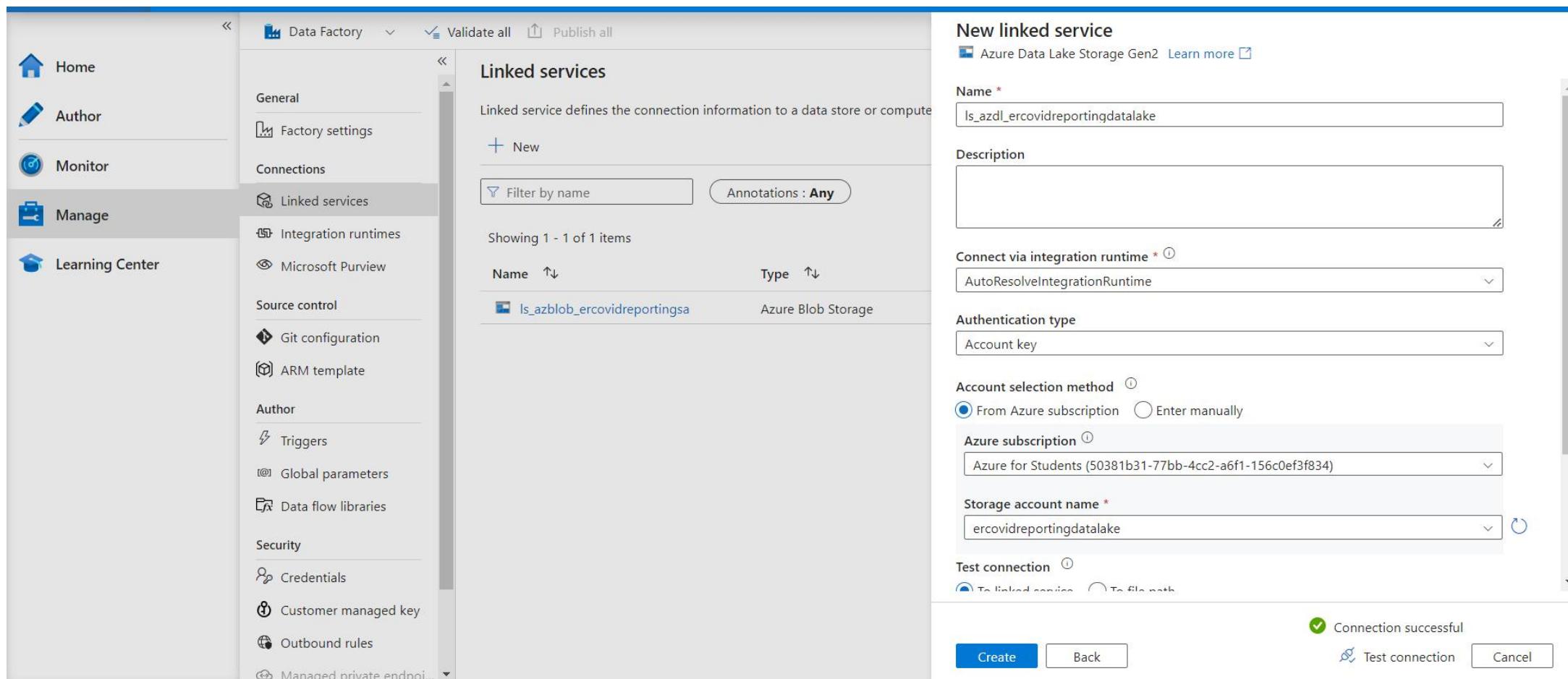
Test connection ⓘ

To linked service  To file path

Connection successful

Create Back

Test connection Cancel



Home    Author    Monitor    Manage    Learning Center

Data Factory    Validate all    Publish all

Factory Resources

Filter resources by name

Pipelines    0  
Change Data Capture (preview)    0  
**Datasets**    0  
Data flows    0  
Power Query    0

Set properties

Name: ds\_populationdata\_gz

Linked service \*: ls\_azblob\_ercovidreportingsa

File path: population / Directory / population\_by\_age.tsv.gz

First row as header:

Import schema:

From connection/store    From sample file    None

OK   Back   Cancel

Home Author Monitor Manage Learning Center

Data Factory Validate all Publish all 1

Factory Resources << ds\_populationdata\_gz <<

Filter resources by name +

Pipelines 0 Change Data Capture (preview) 0 Datasets 1 ds\_populationdata\_gz Data flows 0 Power Query 0

**ds\_populationdata\_gz**

DelimitedText CSV

Connection Schema Parameters

Compression level Optimal

Column delimiter ⓘ Tab (\t)

Row delimiter ⓘ Default (\r,\n, or \r\n)

Encoding ⓘ Default(UTF-8)

Quote character ⓘ Double quote ("")

Escape character ⓘ Backslash (\)

First row as header ⓘ

Null value ⓘ

Set properties

Name ds\_populationdata\_tsv

Linked service \* ls\_azdl\_ercovidreportingdatalake

File path raw / Directory / population\_by\_age.tsv

First row as header

Import schema

From connection/store  From sample file  None

OK Back Cancel

Home    Author    Monitor    Manage    Learning Center

Data Factory    Validate all    Publish all (2)

Preview experience    Off

Factory Resources

Activities

Pipelines: pl\_ingest\_population\_data (1)

Datasets: ds\_populationdata\_gz, ds\_populationdata\_tsv (2)

Data flows: Change Data Capture (preview) (0)

Power Query: (0)

Move & transform: Copy data, Data flow

Synapse, Azure Data Explorer, Azure Function, Batch Service, Databricks, Data Lake Analytics, General, HDInsight, Iteration & conditionals, Machine Learning, Power Query

Copy data: Copy Population data

Properties

General    Related

Name: pl\_ingest\_population\_data

Description:

Annotations

Parameters, Variables, Settings, Output

Pipeline run ID: 2f84ceba-d70f-4961-8cf5-b87cab5919bb

All status, Export to CSV

Showing 1 - 1 of 1 items

Activity name	Status	Activity type
Copy Population data	Succeeded	Copy data

<adf.azure.com/en/authoring/pipeline/pipeline1?factory=%2Fsubscriptions%2F50381b31-77bb-4cc2-a6f1-156c0ef3f834%2FresourceGroups%2Fcovid-reporting-rg%2Fproviders%2FMicrosoft.DataF...>

## Microsoft Azure | Data Factory > er-covid-reporting

Details    Refresh

Learn more on copy performance details from here.

Activity run id: b377adcd-2828-4231-a19b-9bf39dcd83a0

```

graph LR
    A[Azure Blob Storage  
Region: East US] -- "Succeeded" --> B[Azure Data Lake Storage Gen2  
Region: East US]
  
```

**Azure Blob Storage**  
Region: East US

**Azure Data Lake Storage Gen2**  
Region: East US

Data read:	5.986 KB	Data written:	26.07 KB
Files read:	1	Files written:	1
Peak connections:	1	Peak connections:	1

Copy duration: 00:00:05  
Throughput: 5.986 KB/s

**Copy details:**

- From: Azure Blob Storage → Azure Data Lake Storage Gen2
- Start time: 6/29/2023, 9:06:11 PM
- Used DIUs: 4
- Used parallel copies: 1
- Duration: 00:00:05

Details	Working duration	Total duration
Queue	00:00:00	00:00:03

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

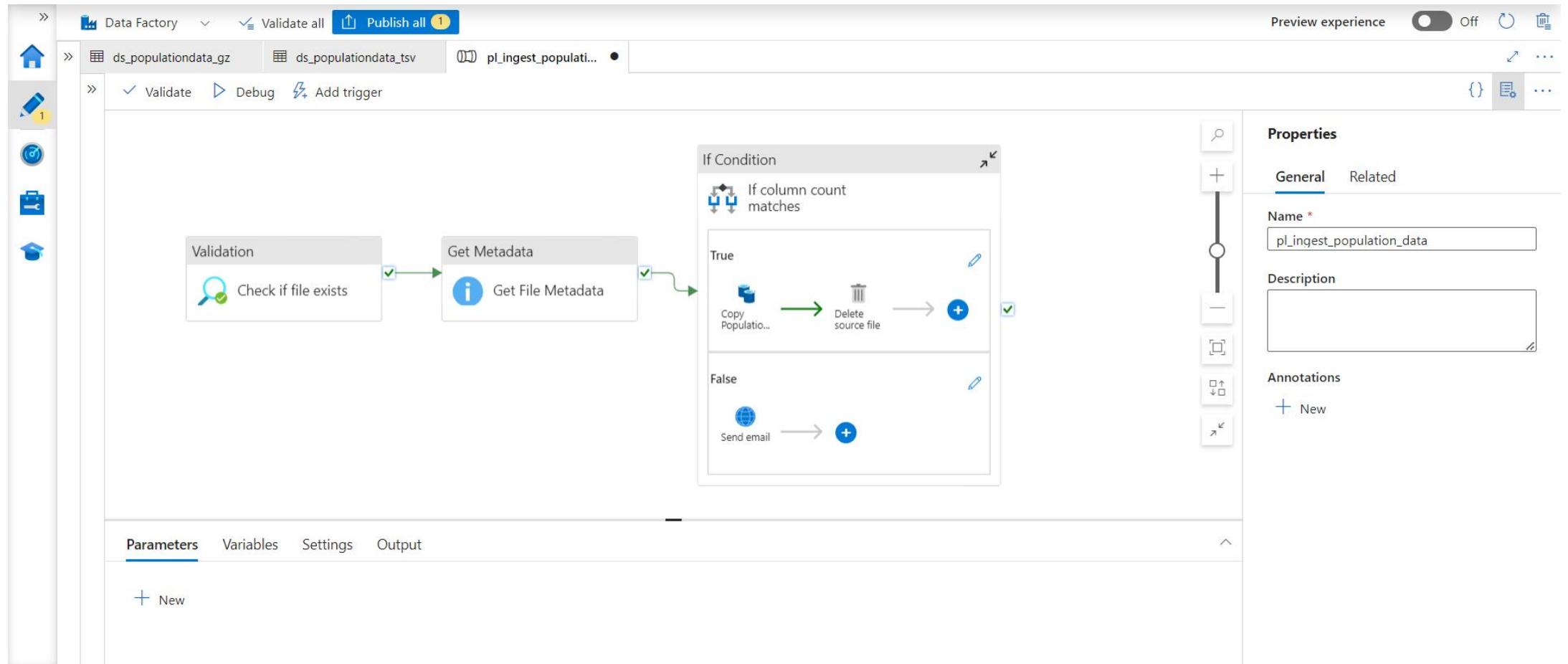
5 stars

population\_by\_age.tsv population\_by\_a....tsv.gz

Show all

Type here to search

9:08 PM 6/29/2023



Data Factory   Validate all   Publish all 1   Preview experience   Off

ds\_populationdata\_gz   ds\_populationdata\_tsv   pl\_ ingest\_population\_data

Validate   Debug   Add trigger

Validation   Get Metadata   If Condition

Check if file exists   Get File Metadata   Copy Population data   Delete source file

True   False

Properties

General   Related

Name \* pl\_ ingest\_population\_data

Description

Annotations

Pipeline run ID: 0131b4cf-5c11-492d-9e07-50db60f217d1   View debug run consumption

All status   List   Export to CSV

Showing 1 - 5 of 5 items

Activity name	Status	Activity type	Run start	Duration	Log	Integration runtime
Delete source file	Succeeded	Delete	6/29/2023, 10:27:52 PM	3s		AutoResolveIntegrationRuntime
Copy Population data	Succeeded	Copy data	6/29/2023, 10:27:43 PM	7s		AutoResolveIntegrationRuntime
If column count matches	Succeeded	If Condition	6/29/2023, 10:27:43 PM	15s		
Get File Metadata	Succeeded	Get Metadata	6/29/2023, 10:27:40 PM	3s		AutoResolveIntegrationRuntime
Check if file exists	Succeeded	Validation	6/29/2023, 10:27:36 PM	4s		

Data Factory < Validate all Publish all

Triggers

To execute a pipeline set the trigger. Triggers represent a unit of processing that determines when a pipeline execution needs to be kicked off.

+ New

Filter by name Annotations : Any

No triggers to show

If you expected to see results, try changing your filters or create a new trigger

Create trigger

General

Factory settings

Connections

Linked services

Integration runtimes

Microsoft Purview

Source control

Git configuration

ARM template

Author

Triggers

Global parameters

Data flow libraries

Security

Credentials

Customer managed key

Outbound rules

Managed private endpoints

Workflow orchestration manager

Apache Airflow

### New trigger

Name \* blobarrivaltrigger

Description

Type \* Storage events

Account selection method \*  From Azure subscription  Enter manually

Azure subscription  Azure for Students (50381b31-77bb-4cc2-a6f1-156c0ef3f834)

Storage account name \* ercovidreportingsa

Container name \* population

Blob path begins with \* population\_by\_age.tsv.gz

Blob path ends with \*

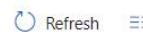
Event \*  Blob created  Blob deleted

Ignore empty blobs \*  Yes  No

Continue Cancel

## Trigger runs

All Schedule Tumbling window Storage events Custom events



Refresh



Edit columns

Local time : **Last 24 hours**

Trigger name : **All**

Status : **All**

Export to CSV |



No results to show

If you expected to see results, try changing your filters.

EXPLORER

Search for resources

Collapse all Refresh all

- Quick Access
- Emulator & Attached
- Storage Accounts
- Azure for Students (AvinashChinta@my.unt.edu)
  - Storage Accounts
    - cs71003200143e43ad6
    - dbstoragegeocb4jhwww7hy (ADLS Gen2)
    - dispractise (ADLS Gen2)
    - ercovidreportingdatalake (ADLS Gen2)
      - Blob Containers
        - \$logs
        - raw
      - File Shares
      - Queues
      - Tables
    - ercovidreportingsa
      - Blob Containers
        - \$logs
        - population
      - File Shares
      - Queues
      - Tables
  - Disks
    - cloud-shell-storage-southcentralus
    - covid-reporting-rg
    - databricks-rg-dbs-hands-on-bj47lln3b
    - NetworkWatcherRG
    - rg-avinash-hands-on

population raw

Upload Download

← → ↕ Active blobs

Name Access Tier

Selected files:

No data available.

population\_by\_age.tsv.gz

Blob type:

Block Blob

Upload .vhdx files as page blobs (recommended)

Target Access Tier:

Default

Destination directory:

/

Tags:

key value

Activities

Clear completed Clear success

- Deletion of 'population\_... Started at: 6/29/2023 10:48 PM, Duration: 4 seconds
- Transfer of 'C:\Users\Dehe... Started at: 6/29/2023 10:48 PM, Duration: 4 seconds
- Deletion of 'population\_... Started at: 6/29/2023 10:48 PM, Duration: 4 seconds
- Transfer of 'C:\Users\Dehe... complete: 1 item transferred (used SAS, discovery completed)

Upload Cancel

Manage History

Folder Statistics

Refresh

Status	Remaining Days	Deleted Time	Lease State	Disk Name	VM Name	Dis

Type here to search

10:52 PM  
6/29/2023

Trigger runs

All Schedule Tumbling window Storage events Custom events Refresh Edit columns

Local time : Last 24 hours Trigger name : All Status : All X Export to CSV |

Showing 1 - 1 items

Trigger name ↑↓	Trigger time ↑↓	Status ↑↓	Pipelines	Trigger file name	Trigger file size	Message	Properties
blob_ingestpopulationdata_tr...	6/29/2023, 10:52:44 PM	✓ Succeeded	1	/blobServices/default/contai	5986 B		⚙️

Alerts & metrics

## Microsoft Azure Storage Explorer

File Edit View Help

EXPLORER

population raw

Search for resources

Collapse all Refresh all

Quick Access

Emulator & Attached

Storage Accounts

Azure for Students (AvinashChinta@my.unt.edu)

- Storage Accounts
  - cs71003200143e43ad6
  - dbstorageocb4jhwww7hy (ADLS Gen2)
  - dispractise (ADLS Gen2)
  - ercovidreportingdatalake (ADLS Gen2)
    - Blob Containers
      - \$logs
      - raw
    - File Shares
    - Queues
    - Tables
  - ercovidreportingsa
    - Blob Containers
      - \$logs
      - population
    - File Shares
    - Queues
    - Tables
- Disk

No data available.

Name Access Tier Access Tier Last Modified Last Modified Blob Type Content Type Size Status Remaining Days Deleted Time Lease State Disk Name VM Name Dis

Showing 0 to 0 of 0 entries

Activities

Clear completed Clear successful

- Transfer of 'C:\Users\Deheeraji Bajjuri\Downloads\population\_by\_age.tsv.gz' to 'population/' complete: 1 item transferred (used SAS, discovery completed) Started at: 6/29/2023 10:26 PM, Duration: 4 seconds Copy AzCopy Command to Clipboard
- Deletion of 'population\_by\_age.tsv' from 'raw/population/' completed: 1 completed (used name and key) Started at: 6/29/2023 10:26 PM, Duration: 4 seconds Copy AzCopy Command to Clipboard
- Deletion of 'population\_by\_age.tsv.gz' from 'population/' completed: 1 completed (used SAS, discovery completed) Started at: 6/29/2023 10:12 PM, Duration: 7 seconds Copy AzCopy Command to Clipboard
- Successfully deleted blob container 'population'

Actions Properties

Type here to search

10:28 PM 6/29/2023

## Microsoft Azure Storage Explorer

File Edit View Help

EXPLORER

population raw

Search for resources

Collapse all Refresh all

Quick Access

Emulator & Attached

Storage Accounts

Azure for Students (AvinashChinta@my.unt.edu)

- Storage Accounts
  - cs71003200143e43ad6
  - dbstorageocb4jhwww7hy (ADLS Gen2)
  - dispractise (ADLS Gen2)
  - ercovidreportingdatalake (ADLS Gen2)
    - Blob Containers
      - \$logs
      - raw
- ercovidreportingsa
  - Blob Containers
    - \$logs
    - population
- File Shares
- Queues
- Tables

Disks

- cloud-shell-storage-southcentralus
- covid-reporting-rg
- databricks-rg-dbs-hands-on-bj47ln3b
- NetworkWatcherRG
- rg-avinash-hanson

Showing 1 to 1 of 1 cached items

Activities

Clear completed Clear successful

- Transfer of 'C:\Users\Deheeraji Bajjuri\Downloads\population\_by\_age.tsv.gz' to 'population/' complete: 1 item transferred (used SAS, discovery completed) Started at: 6/29/2023 10:26 PM, Duration: 4 seconds Copy AzCopy Command to Clipboard
- Deletion of 'population\_by\_age.tsv' from 'raw/population/' completed: 1 completed (used name and key) Started at: 6/29/2023 10:26 PM, Duration: 4 seconds Copy AzCopy Command to Clipboard
- Deletion of 'population\_by\_age.tsv.gz' from 'population/' completed: 1 completed (used SAS, discovery completed) Started at: 6/29/2023 10:12 PM, Duration: 7 seconds Copy AzCopy Command to Clipboard
- Successfully deleted blob container 'population'

Actions Properties

Type here to search

10:29 PM 6/29/2023

» Data Factory    Validate all    Publish all 1

Linked services

Linked service defines the connection information to a data store or compute. [Learn more](#)

+ New

Filter by name Annotations : Any

Showing 1 - 3 of 3 items

Name	Type	Related
ls_azblob_ercovidreportingsa	Azure Blob Storage	1
ls_azdl_ercovidreportingdatalake	Azure Data Lake Storage Gen2	1
ls_http_ecdc_data	HTTP	0

Edit linked service [HTTP Learn more](#)

Name \* ls\_http\_ecdc\_data

Description

Connect via integration runtime \* ⓘ AutoResolveIntegrationRuntime

Base URL \* https://raw.githubusercontent.com

⚠ Information will be sent to the URL specified. Please ensure you trust the URL entered.

Server Certificate Validation ⓘ  Enable  Disable

Authentication type \* ⓘ Anonymous

Auth headers ⓘ + New

Annotations + New

> Parameters

Save Cancel ⚙ Test connection

Data Factory   Validate all   Publish all (4)

Preview experience   Off

Factory Resources

Filter resources by name

Pipelines (2)

- pl\_ingest\_cases\_deaths\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) (0)

Datasets (4)

- ds\_cases\_deaths\_data\_dl
- ds\_http\_cases\_deaths\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows (0)

Power Query (0)

**ds\_cases\_deaths\_data\_dl**

DelimitedText

CSV

Connection Schema Parameters

Linked service \* ls\_azdl\_ecovidreportingdatalake Test connection Edit New Learn more

File path \* raw / ecdc / cases\_deaths.csv Browse Preview

Compression type None

Column delimiter ⓘ Comma (,)

Row delimiter ⓘ Default (\r,\n, or \r\n)

Encoding ⓘ Default(UTF-8)

Quote character ⓘ Double quote ("")

Escape character ⓘ Backslash (\)

First row as header ⓘ

Properties

General Related (1)

Name \* ds\_cases\_deaths\_data\_dl

Description

Annotations

+ New

Expand Data Factory Validate all Publish all 4 Preview experience Off

Factory Resources Filter resources by name +

Pipelines 2

- pl\_ingest\_cases\_deaths\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 4

- ds\_cases\_deaths\_data\_dl
- ds\_http\_cases\_deaths\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows 0

Power Query 0

**Properties**

General Related (1)

Name \* ds\_http\_cases\_deaths\_data

Description

Annotations + New

**Connection** Schema Parameters

Linked service \* ls\_http\_ecdc\_data Test connection Edit New Learn more

Base URL https://raw.githubusercontent.com/Eshwarreddyt/AzureDataFactory\_Project-... Detect format

Relative URL Eshwarreddyt/AzureDataFactory\_Project-... Preview data

Compression type None

Column delimiter Comma (,)

Row delimiter Default (\r,\n, or \r\n)

Encoding Default(UTF-8)

Quote character Double quote ("")

Escape character Backslash (\)

The screenshot shows the Azure Data Factory interface with the 'Properties' blade open for a dataset named 'ds\_http\_cases\_deaths\_data'. The 'General' tab is selected. The dataset is defined as a 'DelimitedText' type with a 'CSV' icon. The 'Connection' tab shows it is connected to a linked service named 'ls\_http\_ecdc\_data'. The 'Schema' and 'Parameters' tabs are also visible. The 'Properties' section on the right contains fields for 'Name' (set to 'ds\_http\_cases\_deaths\_data'), 'Description' (empty), and 'Annotations' (with a '+ New' button). The 'Annotations' section is currently collapsed.

Screenshot of the Azure Data Factory studio interface showing the configuration of a pipeline activity.

**Left Sidebar:**

- Factory Resources**:
  - Pipelines: pl\_ ingest\_cases\_deaths\_data (selected), pl\_ ingest\_population\_data
  - Change Data Capture (preview): 0
  - Datasets: ds\_cases\_deaths\_data\_dl, ds\_http\_cases\_deaths\_data, ds\_populationdata\_gz, ds\_populationdata\_tsv
  - Data flows: 0
  - Power Query: 0

**Top Bar:**

- Expand Data Factory, Validate all, Publish all (4)
- Preview experience: Off
- ...

**Activities Panel:**

- Search activities: Copy data, Move & transform, Synapse, Azure Data Explorer, Azure Function, Batch Service, Databricks, Data Lake Analytics, General, HDInsight, Iteration & conditionals, Machine Learning, Power Query

**Properties Panel (Copy data activity):**

**General Tab:**

- Name \***: Copy Cases and Deaths data
- Description**: (empty)
- Activity state (preview)**: Active (radio button selected)
- Timeout**: 0.12:00:00
- Retry**: 0
- Retry interval (sec)**: 30

**Related Tab:**

- General, Related

**Annotations:**

- + New

Expand Data Factory   Validate all   Publish all (4)

Preview experience: Off

Factory Resources

- Pipelines (2)
  - pl\_ ingest\_cases\_deaths\_data
  - pl\_ ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (4)
  - ds\_cases\_deaths\_data\_dl
  - ds\_http\_cases\_deaths\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
- Data flows (0)
- Power Query (0)

Activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Copy data

Copy Cases and Deaths data

Source dataset: ds\_http\_cases\_deaths\_data

Request method: GET

Additional headers:

Request body:

Properties

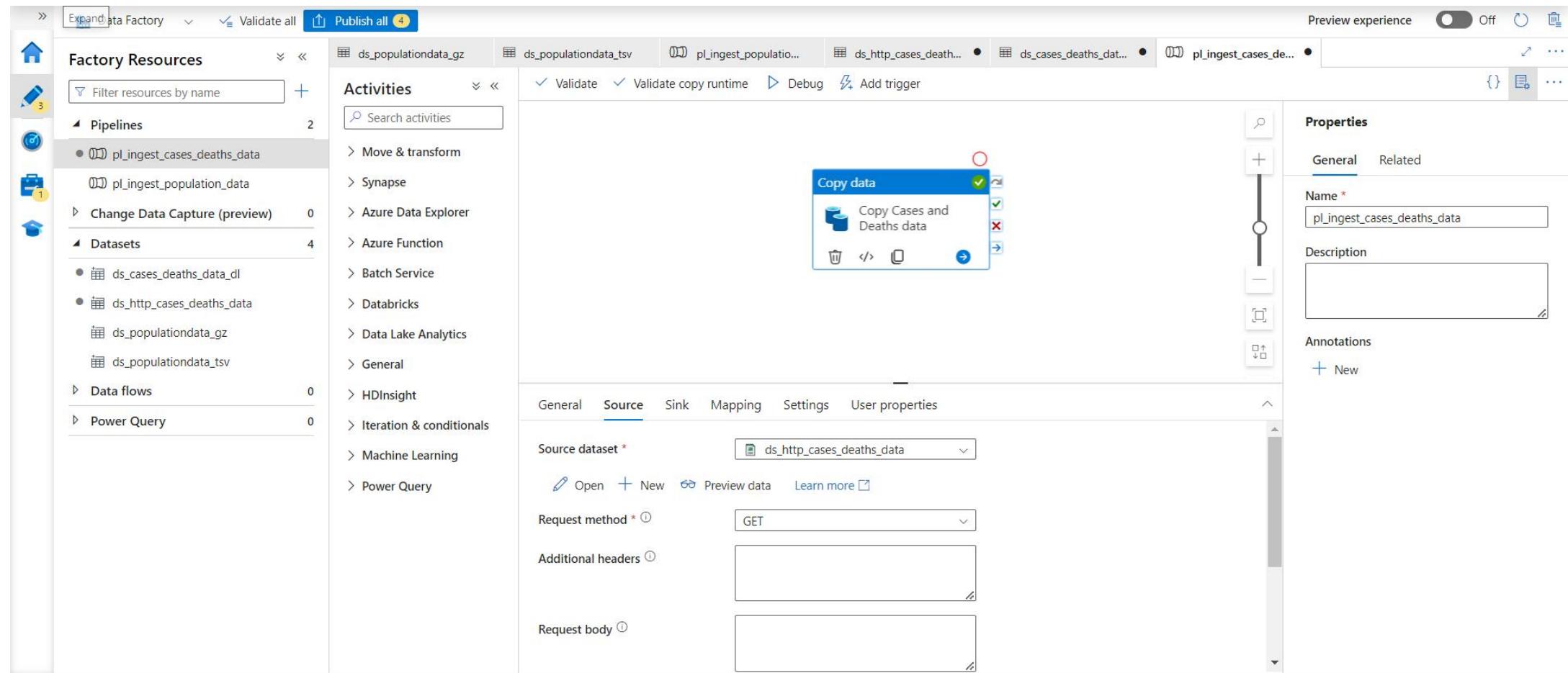
General

Name: pl\_ ingest\_cases\_deaths\_data

Description:

Annotations

New



Validate all Publish all 4

Preview experience Off

Factory Resources

Pipelines 2

- pl\_ingest\_cases\_deaths\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 4

- ds\_cases\_deaths\_data\_dl
- ds\_http\_cases\_deaths\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows 0

Power Query 0

Activities

Move & transform

Synapse

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Power Query

Copy data

Copy Cases and Deaths data

Properties

General Related

Name \* pl\_ingest\_cases\_deaths\_data

Description

Annotations + New

General Source Sink Mapping Settings User properties

Sink dataset \* ds\_cases\_deaths\_data\_dl Open + New Learn more

Copy behavior None

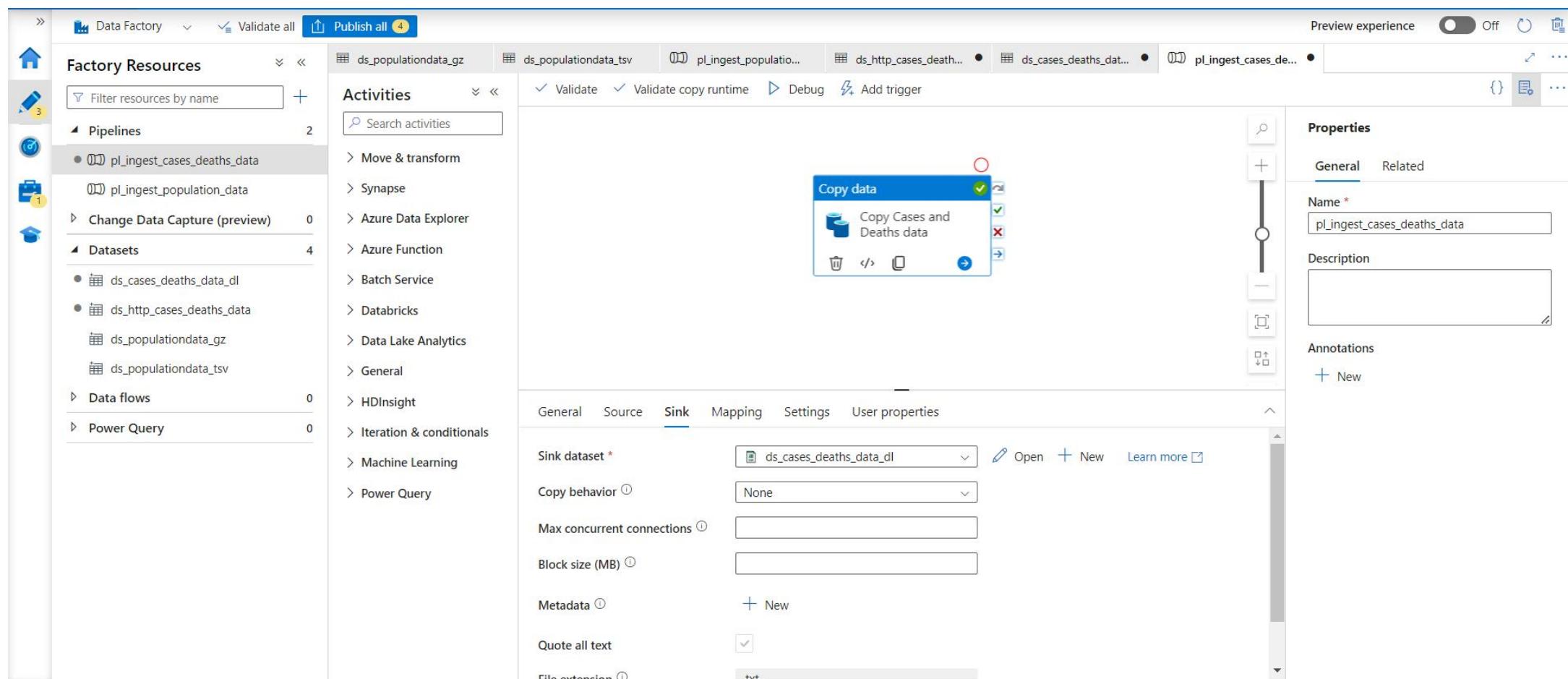
Max concurrent connections

Block size (MB)

Metadata + New

Quote all text

File extensions



Home Data Factory Validate all Publish all 4 Preview experience Off

Factory Resources

Filter resources by name

Pipelines 2

- pl\_ ingest\_cases\_deaths\_data
- pl\_ ingest\_population\_data

Change Data Capture (preview) 0

Datasets 4

- ds\_cases\_deaths\_dl
- ds\_http\_cases\_deaths\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows 0

Power Query 0

Activities

Search activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDI insight
- Iteration & conditionals
- Machine Learning
- Power Query

Copy data

Copy Cases and Deaths data

Properties

General Related

Name \* pl\_ ingest\_cases\_deaths\_data

Description

Annotations + New

Pipeline run ID: b8bd946d-8bea-431d-9675-da5d03207ffe

View debug run consumption

All status

Export to CSV

Showing 1 - 1 of 1 items

Activity name	Status	Activity type	Run start	Duration	Lc
Copy Cases and Deaths data	<span style="color: green;">Succeeded</span>	Copy data	6/30/2023, 12:35:31 PM	8s	

## Microsoft Azure Storage Explorer

File Edit View Help



## EXPLORER

Search for resources	
Collapse all	Refresh all
Quick Access	
Emulator & Attached	
> Storage Accounts	
✓ Azure for Students (AvinashChinta@my.unt.edu)	
✓ Storage Accounts	
> cs71003200143e43ad6	
> dbstorageoch4zhwww7hy (ADLS Gen2)	
> dlspractise (ADLS Gen2)	
✓ ercovidreportingdatalake (ADLS Gen2)	
✓ Blob Containers	
> \$logs	
> raw	
> File Shares	
> Queues	
> Tables	
> ercovidreportingsa	
✓ Disks	
> cloud-shell-storage-southcentralus	
> covid-reporting-rg	
> databricks-rg-dbs-hanson-bj47ln3b	
> NetworkWatcherRG	
> rg-avinash-hanson	

raw population

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undelete Folder Statistics Refresh

← → ↑ Active blobs (default) raw > ecdc Filter by prefix (case-sensitive)

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted Time	Lease
cases_deaths.csv	Hot (inferred)		6/30/2023 12:35 PM	Block Blob	application/octet-stream	13.78 MiB	Active				

Showing 1 to 1 of 1 cached items

## Activities

Clear completed Clear successful

- ✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\population\_by\_age.tsv.gz' to 'population/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/29/2023 10:52 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- ✓ Deletion of 'population\_by\_age.tsv.gz' from 'population/' completed: 1 completed (used SAS, discovery completed)  
Started at: 6/29/2023 10:51 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- ✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\population\_by\_age.tsv.gz' to 'population/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/29/2023 10:48 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- ✓ Deletion of 'population\_by\_age.tsv' from 'raw/population/' completed: 1 completed (used name and key) [Copy AzCopy Command to Clipboard](#)

Actions Properties



Type here to search

12:40 PM

6/30/2023



Data Factory   Validate all   Publish all (3)

Preview experience   Off

Factory Resources

Filter resources by name

Pipelines (2)

- pl\_ingest\_cases\_deaths\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) (0)

Datasets (4)

- ds\_cases\_deaths\_data\_dl
- ds\_http\_cases\_deaths\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows (0)

Power Query (0)

**ds\_http\_cases\_deaths\_data**

DelimitedText

CSV

Connection Schema Parameters

+ New   Delete

Name	Type	Default value
RelativeURL	String	Value

Properties

General   Related (1)

Name \* ds\_http\_cases\_deaths\_data

Description

Annotations

+ New

Validate all Publish all 3

Preview experience Off

Factory Resources

Pipelines 2

- pl\_ingest\_cases\_deaths\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 4

- ds\_cases\_deaths\_data\_dl
- ds\_http\_cases\_deaths\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows 0

Power Query 0

**ds\_http\_cases\_deaths\_data**

DelimitedText CSV

Connection Schema Parameters

Linked service \* ls\_http\_ecdc\_data Test connection Edit New Learn more

Base URL https://raw.githubusercontent.com Detect format

Relative URL @dataset().RelativeURL Preview data

Compression type None

Column delimiter Comma (,)

Row delimiter Default (\r\n or \r\n)

Encoding Default(UTF-8)

Quote character Double quote ("")

Escape character Backslash (\)

Properties

General Related (1)

Name \* ds\_http\_cases\_deaths\_data

Description

Annotations

+ New

Data Factory   Validate all   Publish all (3)   Preview experience   Off

Factory Resources   Filter resources by name

Pipelines: pl\_ingest\_cases\_deaths\_data, pl\_ingest\_population\_data

Change Data Capture (preview): 0

Datasets: ds\_cases\_deaths\_data\_dl (selected), ds\_http\_cases\_deaths\_data, ds\_populationdata\_gz, ds\_populationdata\_tsv

Data flows: 0

Power Query: 0

**ds\_cases\_deaths\_data\_dl**

DelimitedText

**Properties**

General   Related (1)

Name \*: ds\_cases\_deaths\_data\_dl

Description:

Annotations

**Parameters**

New   Delete

Name	Type	Default value
filename	String	Value

Data Factory   Validate all   Publish all (3)

Preview experience   Off

Factory Resources   <<

Filter resources by name

Pipelines   2

- pl\_ingest\_cases\_deaths\_data
- pl\_ingest\_population\_data

Change Data Capture (preview)   0

Datasets   4

- ds\_cases\_deaths\_data\_dl
- ds\_http\_cases\_deaths\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows   0

Power Query   0

**Properties**

General   Related (1)

Name \* ds\_cases\_deaths\_data\_dl

Description

Annotations

**ds\_cases\_deaths\_data\_dl**

DelimitedText

CSV

Connection   Schema   Parameters

Linked service \* ls\_azdl\_ecovidreportingdatalake   Test connection   Edit   New   Learn more

File path \* raw / ecdc / @dataset().filename   Browse

Compression type None

Column delimiter Comma (,)

Row delimiter Default (\r,\n, or \r\n)

Encoding Default(UTF-8)

Quote character Double quote ("")

Escape character Backslash (\)

First row as header

Expand Data Factory   Validate all   Publish all (3)

Preview experience   Off

Factory Resources   ds\_populationdata\_gz   ds\_populationdata\_tsv   pl\_ingest\_population...   ds\_http\_cases\_death...   ds\_cases\_deaths\_dat...   pl\_ingest\_cases\_de...

Activities   Validate   Debug   Add trigger

Filter resources by name

Pipelines (2)

- pl\_ingest\_cases\_deaths\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) (0)

Datasets (4)

- ds\_cases\_deaths\_data\_dl
- ds\_http\_cases\_deaths\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows (0)

Power Query (0)

Activities

Search activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Copy data

Copy Cases and Deaths data

Properties

General   Related

Name \* pl\_ingest\_cases\_deaths\_data

Description

Annotations

New

Parameters   Variables   Settings   Output

+ New   Delete

Name	Type	Default value
sourceRelativeURL	String	Eshwarreddyt/AzureDataFactor
sinkFileName	String	hospital_admissions.csv

Expand Data Factory Validate all Publish all 3 Preview experience Off

### Factory Resources

- Pipelines (2)
  - pl\_ ingest\_cases\_deaths\_data
  - pl\_ ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (4)
  - ds\_cases\_deaths\_data\_dl
  - ds\_http\_cases\_deaths\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
- Data flows (0)
- Power Query (0)

### Activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Copy data

Copy Cases and Deaths data

Source dataset \* ds\_http\_cases\_deaths\_data

Request method \* GET

Additional headers

Name Value Type

RelativeURL @variables('sourceRelativeURL') string

Properties

General Related

Name \* pl\_ ingest\_cases\_deaths\_data

Description

Annotations New

This screenshot shows the Azure Data Factory studio interface. On the left, the 'Factory Resources' sidebar lists various resources like Pipelines, Datasets, and Activities. The 'Activities' section is expanded, showing options like Move & transform, Synapse, and Azure Function. In the center, a 'Copy data' activity is selected, which is used to copy data from 'ds\_http\_cases\_deaths\_data' to another dataset. The 'Source' tab of the activity's properties is active, showing the source dataset as 'ds\_http\_cases\_deaths\_data'. Below this, 'Dataset properties' are defined, including 'RelativeURL' set to '@variables('sourceRelativeURL')' of type 'string'. The 'Request method' is set to 'GET'. On the right, the 'Properties' panel shows the general settings for the 'pl\_ ingest\_cases\_deaths\_data' pipeline, including its name and a description field.

» Data Factory < Validate all Publish all 3

Preview experience Off

Factory Resources <

Filter resources by name +

Pipelines 2

- pl\_ ingest\_cases\_deaths\_data
- pl\_ ingest\_population\_data

Change Data Capture (preview) 0

Datasets 4

- ds\_cases\_deaths\_data\_dl
- ds\_http\_cases\_deaths\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows 0

Power Query 0

Activities <

Validate Validate copy runtime Debug Add trigger

Search activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Copy data

Copy Cases and Deaths data

General Source Sink Mapping Settings User properties

Sink dataset \* ds\_cases\_deaths\_data\_dl Open New Learn more

Dataset properties

Name	Value	Type
filename	@variables('sinkFileName')	string

Copy behavior None

Max concurrent connections

Block size (MB)

Properties

General Related

Name \* pl\_ ingest\_cases\_deaths\_data

Description

Annotations

+ New

The screenshot shows the Azure Data Factory pipeline editor interface. On the left, the 'Factory Resources' sidebar lists various datasets and pipelines. In the center, a 'Copy data' activity is selected, which is part of a pipeline named 'pl\_ ingest\_cases\_deaths\_data'. The 'Sink' tab is active in the configuration pane. The sink dataset is set to 'ds\_cases\_deaths\_data\_dl'. A dataset property 'filename' is defined with the value '@variables('sinkFileName')'. The 'Copy behavior' is set to 'None'. The right side of the screen displays the 'Properties' panel for the selected pipeline, showing its general settings like name and description.

Preview experience  Off

Data Factory Validate all Publish all

### Factory Resources

Filter resources by name

- Pipelines 2
  - pl\_ingest\_cases\_deaths\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 4
  - ds\_cases\_deaths\_data\_dl
  - ds\_http\_cases\_deaths\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
- Data flows 0
- Power Query 0

### Activities

Validate Debug Add trigger

Search activities

- > Move & transform
- > Synapse
- > Azure Data Explorer
- > Azure Function
- > Batch Service
- > Databricks
- > Data Lake Analytics
- > General
- > HDInsight
- > Iteration & conditionals
- > Machine Learning
- > Power Query

Copy data

Copy Cases and Deaths data

Properties

General Related

Name \* pl\_ingest\_cases\_deaths\_data

Description

Annotations + New

Pipeline run ID: 4d035e8c-71e9-4762-b449-304e81251cfb

View debug run consumption

All status

Showing 1 - of 1 items

Activity name	Status	Activity type	Run start	Duration	Loc
Copy Cases and Deaths data	Succeeded	Copy data	6/30/2023, 2:01:47 PM	11s	

## Microsoft Azure Storage Explorer

File Edit View Help



EXPLORER

Search for resources

raw population

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undo Refresh

← → ↕ Active blobs (default) raw > ecdc Filter by prefix (case-sensitive)

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted Time
cases_deaths.csv	Hot (inferred)		6/30/2023 12:35 PM	Block Blob	application/octet-stream	13.78 MiB	Active			
hospital_admissions.csv	Hot (inferred)		6/30/2023 2:01 PM	Block Blob	application/octet-stream	1.01 MiB	Active			

Showing 1 to 2 of 2 cached items

Activities

Clear completed Clear successful

- Transfer of 'C:\Users\AvinashChinta@my.unt.edu\Downloads\population\_by\_age.tsv.gz' to 'population/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 12:35 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Deletion of 'population\_by\_age.tsv.gz' from 'population/' completed: 1 completed (used SAS, discovery completed)  
Started at: 6/30/2023 12:35 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Transfer of 'C:\Users\AvinashChinta@my.unt.edu\Downloads\population\_by\_age.tsv.gz' to 'population/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 12:35 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Deletion of 'population\_by\_age.tsv' from 'raw/population/' completed: 1 completed (used name and key) [Copy AzCopy Command to Clipboard](#)

Actions Properties

Screenshot of the Azure Data Factory studio interface showing the configuration of a pipeline activity.

**Top Bar:** Contains "Expand Data Factory", "Validate all", "Publish all (1)", "Preview experience" toggle (Off), and other navigation icons.

**Left Sidebar (Factory Resources):**

- Pipelines: 2 items (pl\_ingest\_cases\_deaths\_data, pl\_ingest\_population\_data)
- Change Data Capture (preview): 0 items
- Datasets: 4 items (ds\_cases\_deaths\_data\_dl, ds\_http\_cases\_deaths\_data, ds\_populationdata\_gz, ds\_populationdata\_tsv)
- Data flows: 0 items
- Power Query: 0 items

**Middle Panel (Activities):**

- Search activities: Copy data, Move & transform, Synapse, Azure Data Explorer, Azure Function, Batch Service, Databricks, Data Lake Analytics, General, HDInsight, Iteration & conditionals, Machine Learning, Power Query.

**Selected Activity (Copy data):**

**Properties Panel (General tab):**

- Name: **pl\_ingest\_cases\_deaths\_data**
- Description: (empty)
- Annotations: + New

**Variables Tab (Variables section):**

Name	Type	Default value
sourceRelativeURL	String	Eshwarreddyt/AzureDataFactory
sinkFileName	String	hospital_admissions.csv

Preview experience  Off

Data Factory Validate all Publish all 1

### Factory Resources

- Pipelines (2)
  - pl\_ingest\_cases\_deaths\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (4)
  - ds\_cases\_deaths\_data\_dl
  - ds\_http\_cases\_deaths\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
- Data flows (0)
- Power Query (0)

### Activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Copy data

Copy Cases and Deaths data

Source dataset \* ds\_http\_cases\_deaths\_data

Request method \* GET

Additional headers

Name Value Type

RelativeURL @pipeline().parameters.sourceRelativ... string

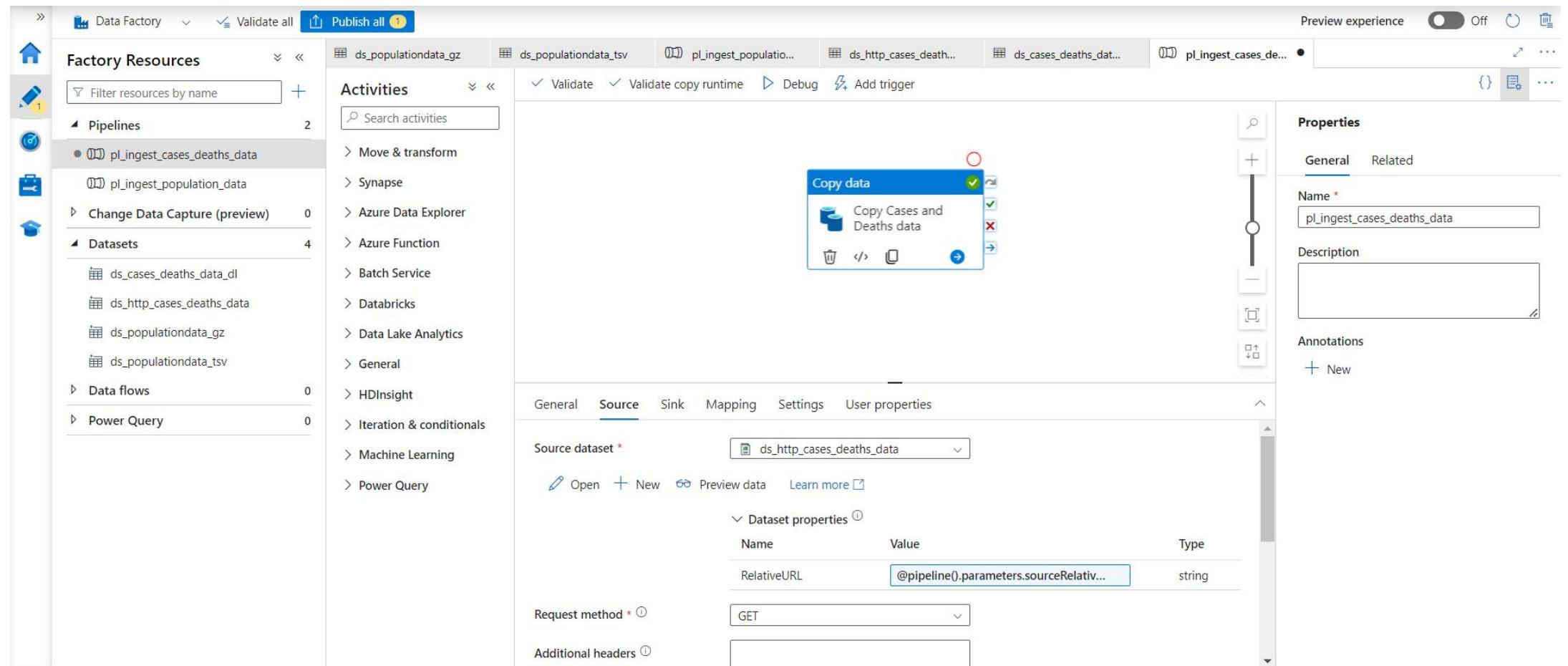
Properties

General Related

Name \* pl\_ingest\_cases\_deaths\_data

Description

Annotations + New



Screenshot of the Azure Data Factory studio interface showing the configuration of a pipeline activity.

**Factory Resources** sidebar:

- Pipelines: 2 items (pl\_ ingest\_cases\_deaths\_data, pl\_ ingest\_population\_data)
- Datasets: 4 items (ds\_cases\_deaths\_data\_dl, ds\_http\_cases\_deaths\_data, ds\_populationdata\_gz, ds\_populationdata\_tsv)
- Data flows: 0 items
- Power Query: 0 items

**Activities** list:

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

**Copy data** activity configuration (selected in the pipeline):

**Properties** pane (General tab):

- Name: pl\_ ingest\_cases\_deaths\_data
- Description: (empty)

**Annotations** pane: + New

**Sink** tab configuration:

- Sink dataset: ds\_cases\_deaths\_data\_dl
- Dataset properties:

Name	Value	Type
filename	@pipeline().parameters.sinkFileName	string

- Copy behavior: None
- Max concurrent connections: (empty)
- Block size (MB): (empty)

Data Factory    Validate all    Publish all (3)

### Factory Resources

- Pipelines (2)
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (4)
  - ds\_ecdc\_data\_dl
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
- Data flows (0)
- Power Query (0)

### Activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

✓ Validating... ✓ Validate copy runtime ▶ Debug ⚡ Add trigger

**Copy data** ✓

Copy Cases and Deaths data

General   Source   **Sink**   Mapping   Settings   User properties

**Sink dataset \*** ds\_ecdc\_data\_dl

Dataset properties

Name	Value
filename	@pipeline0.para

Copy behavior None

Max concurrent connections

Block size (MB)

OK   Cancel

### Pipeline run

Parameters

Name	Type	Value
sourceRelativeURL	string	Eshwarreddy/AzureDataFact...
sinkFileName	string	testing.csv

Preview experience  Off

Data Factory Validate all Publish all 3

### Factory Resources

- Pipelines 2
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 4
  - ds\_ecdc\_data\_dl
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
- Data flows 0
- Power Query 0

### Activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Validate Debug Add trigger

Copy data ✓

Copy ecdc data

Properties

General Related

Name \* pl\_ingest\_ecdc\_data

Description

Annotations + New

Pipeline run ID: 3d50fdd3-23e0-4ce8-aff2-07b1cba7ad0d [View debug run consumption](#)

All status [Export to CSV](#)

Showing 1 - 1 of 1 items

Activity name	Status	Activity type	Run start	Duration	L
Copy ecdc data	<span style="color: green;">✓ Succeeded</span>	Copy data	6/30/2023, 2:26:31 PM	8s	

## Microsoft Azure Storage Explorer

File Edit View Help



EXPLORER

Search for resources

raw population

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undo Refresh

← → ↑ Active blobs (default) raw &gt; ecdc

Filter by prefix (case-sensitive)

- Collapse all Refresh all
- Quick Access
  - Emulator & Attached
  - Storage Accounts
  - Azure for Students (AvinashChinta@my.unt.edu)
    - Storage Accounts
      - cs71003200143e43ad6
      - dbstorageocb4jzhwww7hy (ADLS Gen2)
      - dlspractise (ADLS Gen2)
    - ercovidreportingdatalake (ADLS Gen2)
      - Blob Containers
        - \$logs
        - raw
      - File Shares
      - Queues
      - Tables
    - ercovidreportingsa
  - Disks
    - cloud-shell-storage-southcentralus
    - covid-reporting-rg
    - databricks-rg-dbs-hanson-bj47ln3b
    - NetworkWatcherRG
    - rg-avinash-hanson

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted Time
cases_deaths.csv	Hot (inferred)		6/30/2023 12:35 PM	Block Blob	application/octet-stream	13.78 MiB	Active			
hospital_admissions.csv	Hot (inferred)		6/30/2023 2:01 PM	Block Blob	application/octet-stream	1.01 MiB	Active			
testing.csv	Hot (inferred)		6/30/2023 2:26 PM	Block Blob	application/octet-stream	83.87 KiB	Active			

Showing 1 to 3 of 3 cached items

Activities

Clear completed Clear successful

✓ Deletion of 'testing.csv' from 'raw/ecdc/' completed: 1 completed (used name and key)  
Started at: 6/30/2023 2:25 PM, Duration: 5 seconds

[Copy AzCopy Command to Clipboard](#)

✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\population\_by\_age.tsv.gz' to 'population/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/29/2023 10:52 PM, Duration: 4 seconds

[Copy AzCopy Command to Clipboard](#)

✓ Deletion of 'population\_by\_age.tsv.gz' from 'population/' completed: 1 completed (used SAS, discovery completed)  
Started at: 6/29/2023 10:51 PM, Duration: 4 seconds

[Copy AzCopy Command to Clipboard](#)

✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\population\_by\_age.tsv.gz' to 'population/' complete: 1 item transferred (used SAS, discovery completed)

[Copy AzCopy Command to Clipboard](#)

Actions Properties



Data Factory    Validate all    Publish all (4)    Preview experience Off

### Factory Resources

- Pipelines (2)
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (5)
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
- Data flows (0)
- Power Query (0)

### Activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

✓ Validate    ▶ Debug    ⚙ Add trigger

```
graph LR; Lookup((Lookup)) --> ForEach((ForEach)); subgraph ForEach [execute copy for every record]; Copy[Copy ecdc data]; end
```

The screenshot shows the Azure Data Factory pipeline editor. On the left, the 'Factory Resources' sidebar lists various datasets and activities. The main workspace displays a pipeline structure: a 'Lookup' activity is connected to an 'ForEach' loop. Inside the 'ForEach' loop is a single 'Activities' box containing a 'Copy ecdc data' activity. Below the pipeline, the 'Settings' tab of the 'Lookup' activity configuration pane is visible, showing settings for the source dataset (ds\_ecdc\_filelist), file path type (File path in dataset), and other options like 'First row only' and 'Recursively'. The pipeline itself consists of a 'Lookup' activity followed by an 'ForEach' loop. The 'ForEach' loop has a green checkmark indicating it is active. Inside the loop is a single activity labeled 'Activities' which contains a 'Copy ecdc data' activity.

Validate all Publish all 4

Preview experience Off

Factory Resources

Pipelines

- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview)

Datasets

- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows

Power Query

Activities

Move & transform

Synapse

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Power Query

Lookup

Lookup\_ecdc\_filelist

ForEach

execute copy for every record

Activities

Copy ecdc data

General Settings Activities (1) User properties

Sequential

Batch count

Items

```
@activity('Lookup_ecdc_filelist').outp...
```

```
graph LR; Lookup[Lookup] --> ForEach[ForEach]; subgraph ForEach [ForEach]; direction TB; subgraph Activities [Activities]; direction TB; Copy[Copy ecdc data]; end; end;
```

Preview experience Off

Data Factory Validate all Publish all 4

Factory Resources

Filter resources by name

Pipelines 2

- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 5

- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows 0

Power Query 0

Activities

Validate Debug Add trigger

Search activities

Move & transform

Synapse

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Power Query

Lookup

Lookup\_ecdc\_filelist

ForEach

execute copy for every record

Activities

Copy ecdc data

General Source Sink Mapping Settings User properties

Source dataset \* ds\_http\_ecdc\_data Open New Preview data Learn more

Dataset properties

Name	Type
RelativeURL	string
baseURL	string

Request method \* GET

Additional headers

```
graph LR; Lookup[Lookup] --> ForEach[ForEach]; subgraph ForEach [ForEach]; direction TB; subgraph Activities [Activities]; direction TB; Copy[Copy ecdc data] --> End1(( )); end; end;
```

Factory Resources

Pipelines

- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview)

Datasets

- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows

Power Query

Activities

Move & transform

Synapse

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Power Query

Lookup

Lookup\_ecdc\_filelist

ForEach

execute copy for every record

Activities

Copy ecdc data

General

Source

Sink

Mapping

Settings

User properties

Sink dataset \*

ds\_ecdc\_data\_dl

Open

New

Learn more

Dataset properties

Name	Type
filename	string

None

Max concurrent connections

Block size (MB)

```
graph LR; Lookup((Lookup)) --> ForEach[ForEach]; subgraph ForEach [ForEach]; direction TB; subgraph Activities [Activities]; direction TB; Copy[Copy ecdc data] --> End[End]; end; end;
```

Home

Data Factory

Validate all

Publish all 4

Preview experience Off

Factory Resources

Filter resources by name

Activities

Search activities

Pipelines 2

- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 5

- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv

Data flows 0

Power Query 0

Activities

Validate Debug Add trigger

Lookup

Lookup\_ecdc\_filelist

ForEach

execute copy for every record

Activities

Copy ecdc data

Parameters Variables Settings Output

Showing 1 - 6 of 6 items

Activity name	Status	Activity type	Run start	Duration	Log	Integration runtime	Use
Copy ecdc data	Succeeded	Copy data	6/30/2023, 4:53:50 PM	7s		AutoResolveIntegration	
Copy ecdc data	Succeeded	Copy data	6/30/2023, 4:53:50 PM	7s		AutoResolveIntegration	
Copy ecdc data	Succeeded	Copy data	6/30/2023, 4:53:50 PM	8s		AutoResolveIntegration	
Copy ecdc data	Succeeded	Copy data	6/30/2023, 4:53:50 PM	8s		AutoResolveIntegration	
execute copy for every record	Succeeded	ForEach	6/30/2023, 4:53:49 PM	10s			
Lookup_ecdc_filelist	Succeeded	Lookup	6/30/2023, 4:53:46 PM	3s		AutoResolveIntegration	

## Microsoft Azure Storage Explorer

File Edit View Help

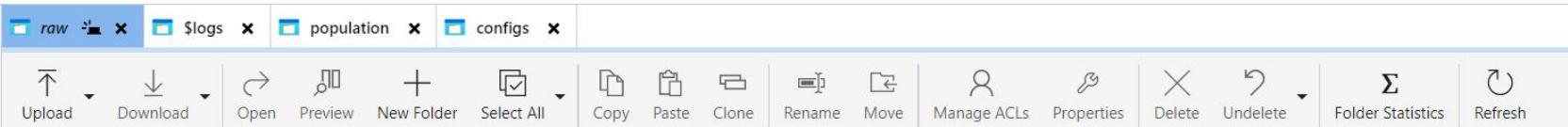


EXPLORER

Search for resources

Collapse all Refresh all

- Quick Access
- Emulator & Attached
- Storage Accounts
  - Storage Accounts
    - cs71003200143e43ad6
    - dbstoragegeocb4jzhwww7hy (ADLS Gen2)
    - dlspractise (ADLS Gen2)
    - ercovidreportingdatalake (ADLS Gen2)
      - Blob Containers
        - \$logs
        - raw
      - File Shares
      - Queues
      - Tables
    - ercovidreportingsa
      - Blob Containers
        - \$logs
        - configs
        - population
      - File Shares
      - Queues
      - Tables
  - Disks
    - cloud-shell-storage-southcentralus
    - covid-reporting-rg
    - databricks-rg-dbs-hands-on-bj47lln3b
    - NetworkWatcherRG
    - rg-avinash-hands-on



← → ▲ Active blobs (default) ▼ raw > ecdc Filter by prefix (case-sensitive) ✖

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted Time
cases_deaths.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	13.78 MiB	Active			
country_response.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	46.20 KiB	Active			
hospital_admissions.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	1.01 MiB	Active			
testing.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	83.87 KiB	Active			

Showing 1 to 4 of 4 cached items

## Activities

Clear completed Clear successful

- Transfer of 'C:\Users\Dheeraj Bajjuri\Downloads\ecdc\_file\_list.json' to 'configs/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 4:22 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Transfer of 'C:\Users\Dheeraj Bajjuri\Downloads\githubusercontent\_ecdc\_file\_list.json' to 'configs/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 3:57 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Deletion of 'githubusercontent\_ecdc\_file\_list.json' from 'configs/' completed: 1 completed (used SAS, discovery completed)  
Started at: 6/30/2023 3:57 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Deletion of 'ecdc' from 'raw/' completed: 1 completed (used name and key) [Copy AzCopy Command to Clipboard](#)

Actions Properties

## Microsoft Azure Storage Explorer

File Edit View Help



## EXPLORER

Search for resources	
Collapse all	Refresh all
Quick Access	
Emulator & Attached	
Storage Accounts	
Azure for Students (AvinashChinta@my.unt.edu)	
Storage Accounts	
cs71003200143e43ad6	
dbstoragegeocb4jzhwww7hy (ADLS Gen1)	
dlspractise (ADLS Gen2)	
ercovidreportingdatalake (ADLS Gen2)	
Blob Containers	
\$logs	
lookup	
raw	
File Shares	
Queues	
Tables	
ercovidreportingsa	
Blob Containers	
\$logs	
configs	
population	
File Shares	
Queues	
Tables	
Disks	
cloud-shell-storage-southcentralus	
covid-reporting-rg	
databricks-rg-dbs-hanson-bj47lln3b	
NetworkWatcherRG	
rg-avinash-hanson	

raw \$logs population configs lookup



← → ↕ Active blobs (default) ↕ lookup Filter by prefix (case-sensitive) 🔎

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted Time	Lease
country_lookup.csv	Hot (inferred)		6/30/2023 6:37 PM	Block Blob	application/vnd.ms-excel	6.86 KiB	Active				

Showing 1 to 1 of 1 cached items

## Activities

Clear completed Clear successful

- Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\case\_deaths\_uk\_ind\_only.csv' to 'raw/ecdc/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\country\_lookup.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 5 seconds [Copy AzCopy Command to Clipboard](#)
- Successfully created blob container 'lookup'

EXPLORER

raw \$logs population configs lookup

Search for resources

Collapse all Refresh all

Quick Access Emulator & Attached Storage Accounts

Storage Accounts

Azure for Students (AvinashChinta@my.unt.edu)

Storage Accounts

cs71003200143e43ad6 dbstorageocb4jzhwww7hy (ADLS Gen1) dispractise (ADLS Gen2) ercovidreportingdatalake (ADLS Gen2)

Blob Containers \$logs lookup raw

File Shares Queues Tables

ercovidreportingsa

Blob Containers \$logs configs population

File Shares Queues Tables

Disks

cloud-shell-storage-southcentralus covid-reporting-rg databricks-rg-dbs-hands-on-bj47lln3b NetworkWatcherRG rg-avinash-hands-on

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undelete Folder Statistics Refresh

Active blobs (default) raw > ecdc

Name Access Tier Access Tier Last Modified Last Modified Blob Type Content Type Size Status Deletion ID Remaining Days Deleted T

case_deaths_uk_ind_only.csv	Hot (inferred)		6/30/2023 6:37 PM	Block Blob	application/vnd.ms-excel	131.48 KiB	Active				
cases_deaths.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	13.78 MiB	Active				
country_response.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	46.20 KiB	Active				
hospital_admissions.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	1.01 MiB	Active				
testing.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	83.87 KiB	Active				

Showing 1 to 5 of 5 cached items

Activities

Clear completed Clear successful

Transfer of 'C:\Users\Deepraj Bajuri\Downloads\case\_deaths\_uk\_ind\_only.csv' to 'raw/ecdc/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)

Transfer of 'C:\Users\Deepraj Bajuri\Downloads\country\_lookup.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 5 seconds [Copy AzCopy Command to Clipboard](#)

Successfully created blob container 'lookup'

Turn on data flow debug

Integration runtime  
AutoResolveIntegrationRuntime

Region AutoResolve

Compute size Small

Debug time to live 1 hour

Data Factory Validate all Publish all 1

Factory Resources Filter resources by name +

Pipelines 2

- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 5

Data flows 1

- dataflow1

Power Query 0

Add Source

Parameters Settings

New

OK Cancel

Data Factory   Validate all   Publish all 2   Preview experience Off

Factory Resources   Filter resources by name

Pipelines: 2  
pl\_ingest\_ecdc\_data, pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets: 6

Data flows: 1  
df\_transform\_cases\_deaths

Power Query: 0

Validate   Data flow debug   Debug Settings

casesanddeathsSource  
Columns: 9 total

Add Source

Properties

General   Related

Name \* df\_transform\_cases\_deaths

Description

Source settings   Source options   Projection   Optimize   Inspect   Data preview

Output stream name \* casesanddeathsSource   Learn more

Description Import data from df\_raw\_cases\_deaths   Reset

Source type \* Dataset   Inline

Dataset \* df\_raw\_cases\_deaths   Test connection   Open   New

Options   Allow schema drift

Data Factory   Validate all   Publish all 2

Preview experience Off

Factory Resources

Pipelines: 2

- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview): 0

Datasets: 6

Data flows: 1

df\_transform\_cases\_deaths

Power Query: 0

Filter resources by name

Validate Data flow debug Debug Settings

**Properties**

General Related

Name \* df\_transform\_cases\_deaths

Description

**filterEuropeOnly**

Import data from df\_raw\_cases\_deaths

Columns: 9 total

Add Source

Filter settings

Optimize Inspect Data preview ●

Output stream name \* filterEuropeOnly

Description Filtering rows using expressions on columns 'continent, country\_code'

Incoming stream \* casesanddeathsSource

Filter on \* continent == 'Europe' && not(isNull(country\_code))

```
graph LR; casesanddeathsSource[casesanddeathsSource] --> filterEuropeOnly[filterEuropeOnly]; filterEuropeOnly --> output[Output stream name: filterEuropeOnly]
```



Data Factory   Validate all   Publish all 2   Preview experience Off

### Factory Resources

Filter resources by name

- Pipelines
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview)
- Datasets
- Data flows
  - df\_transform\_cases\_deaths
- Power Query

✓ Validate   Data flow debug   Debug Settings

```
graph LR; A[casesanddeathsSource] --> B[filterEuropeOnly]; B --> C[selectOnlyReqFields];
```

**Properties**

**General**

Name \* df\_transform\_cases\_deaths

Description

**Select settings**

Output stream name \* selectOnlyReqFields

Description 'country, country\_code, population, indicator, daily\_count, source, reported\_date'

Incoming stream \* filterEuropeOnly

Options

Skip duplicate input columns

Skip duplicate output columns

**Input columns \***

Auto mapping   Reset   Add mapping   Delete

7 mappings: 2 column(s) from the inputs left unmapped

filterEuropeOnly's column	Name as
abc country	country
abc country_code	country_code
123 population	population
abc indicator	indicator
123 daily_count	daily_count

Data Factory   Validate all   Publish all 2

Home   Factory Resources   Filter resources by name +

Pipelines   2

- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview)   0

Datasets   6

Data flows   1

- df\_transform\_cases\_deaths

Power Query   0

Validate   Data flow debug   Debug Settings

**Properties**

General   Related

Name \* df\_transform\_cases\_deaths

Description

**DF Flow Details**

**Input Stream**

casesanddeathsSource   Import data from df\_raw\_cases\_deaths

filterEuropeOnly   Filtering rows using expressions on columns 'continent, country\_code'

selectOnlyReqFields   Columns: 7 total

**Select settings**

Optimize   Inspect   Data preview ●

reported\_date\*

Incoming stream \* filterEuropeOnly

Options

Skip duplicate input columns ⓘ

Skip duplicate output columns ⓘ

**Input columns \***

Auto mapping ⓘ   Reset   Add mapping   Delete   7 mappings: 2 column(s) from the inputs left unmapped ⓘ

filterEuropeOnly's column	Name as
abc country	country
abc country_code	country_code
123 population	population
abc indicator	indicator
123 daily_count	daily_count
abc source	source
name == 'date'	'reported' + '_date'
date	reported_date

Validate all | Publish all (2)

Preview experience Off

### Factory Resources

- Pipelines (2)
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (6)
- Data flows (1)
  - df\_transform\_cases\_deaths
- Power Query (0)

Filter resources by name

Validate Data flow debug Debug Settings

casesanddeathsSource filterEuropeOnly selectOnlyReqFields

Import data from df\_raw\_cases\_deaths Filtering rows using expressions on columns 'continent, country\_code' Columns: 7 total

Select settings Optimize Inspect Data preview

Number of rows: 100 INSERT 100 UPDATE 0 DELETE 0 UPSERT 0 LOOKUP 0 ERROR 0 TOTAL 596

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

country	country...	populat...	indicator	daily_co...	source	reporter...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...
United ...	GBR	66647112	confirm...	0	Epidemi...	2020-01...

Properties

General Related

Name: df\_transform\_cases\_deaths

Description

Data Factory   Validate all   Publish all (2)

Home   Factory Resources   Filter resources by name

Pipelines: pl\_ingest\_ecdc\_data, pl\_ingest\_population\_data

Change Data Capture (preview): 0

Datasets: 6

Data flows: df\_transform\_cases\_deaths

Power Query: 0

Validate   Data flow debug   Debug Settings

**df\_transform\_cases\_deaths**

```
graph LR; casesanddeathsSource[casesanddeathsSource  
Import data from df_raw_cases_deaths] --> filterEuropeOnly[filterEuropeOnly  
Filtering rows using expressions on columns 'continent, country_code']; filterEuropeOnly --> selectOnlyReqFields[selectOnlyReqFields  
Renaming filterEuropeOnly to selectOnlyReqFields with columns 'country, country_code, population, abc']; selectOnlyReqFields --> pivotCounts[pivotCounts  
Columns: 7 total]
```

Add Source

**Properties**

General   Related

Name \*: df\_transform\_cases\_deaths

Description:

**Pivot settings**

Output stream name \*: pivotCounts

Description:

Incoming stream \*: selectOnlyReqFields

1. Group by   2. Pivot key   3. Pivoted columns

Columns	Name as
abc country	country

Preview experience: Off

```
graph LR; casesanddeathsSource[casesanddeathsSource  
Import data from df_raw_cases_deaths] --> filterEuropeOnly[filterEuropeOnly  
Filtering rows using expressions on columns 'continent, country_code']; filterEuropeOnly --> selectOnlyReqFields[selectOnlyReqFields  
Renaming filterEuropeOnly to selectOnlyReqFields with columns 'country, country_code, population, abc']; selectOnlyReqFields --> pivotCounts[pivotCounts  
Columns: 7 total]; pivotCounts --> df_transform_cases_deaths[Properties: General, Name: df_transform_cases_deaths, Description: ]
```

Data Factory   Validate all   Publish all 2   Preview experience Off

### Factory Resources

- Pipelines: pl\_ingest\_ecdc\_data, pl\_ingest\_population\_data
- Change Data Capture (preview): 0
- Datasets: 6
- Data flows: df\_transform\_cases\_deaths
- Power Query: 0

Filter resources by name

✓ Validate   Data flow debug   Debug Settings

```
graph LR; A[casesanddeathsSource] --> B[filterEuropeOnly]; B --> C[selectOnlyReqFields]; C --> D[pivotCounts];
```

**Properties**

**General**

Name \* df\_transform\_cases\_deaths

Description

**Pivot settings**

Optimize   Inspect   Data preview

Columns	Name as
abc_country	country
abc_country_code	country_code
123_population	population
abc_source	source
reported_date	reported_date

Data Factory < Validate all > Publish all (2)

Factory Resources <>

Filter resources by name

Pipelines 2

- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 6

Data flows 1

- df\_transform\_cases\_deaths

Power Query 0

Validate Data flow debug Debug Settings

Preview experience Off

Properties

General Related

Name \* df\_transform\_cases\_deaths

Description

Diagram:

```
graph LR; A[casesanddeathsSource  
Import data from df_raw_cases_deaths] --> B[filterEuropeOnly  
Filtering rows using expressions on columns 'continent, country_code']; B --> C[selectOnlyReqFields  
Renaming filterEuropeOnly to selectOnlyReqFields with columns 'country, country_code, population,']; C --> D[pivotCounts  
Columns: 7 total];
```

Add Source

Pivot settings Optimize Inspect Data preview

Incoming stream \* selectOnlyReqFields

1. Group by 2. Pivot key 3. Pivoted columns

Pivot key \* abc indicator

Value

- confirmed cases
- deaths

Null value

Data Factory   Validate all   Publish all (2)

Factory Resources   Filter resources by name

Pipelines

- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 6

Data flows 1

- df\_transform\_cases\_deaths

Power Query 0

Preview experience Off

Validate Data flow debug Debug Settings

**Properties**

General   Related

Name \* df\_transform\_cases\_deaths

Description

**Diagram**

```
graph LR; A[casesanddeathsSource] --> B[filterEuropeOnly]; B --> C[selectOnlyReqFields]; C --> D[pivotCounts];
```

**Pivot settings**

Incoming schema: selectonlyreqfields

1. Group by   2. Pivot key   3. Pivoted columns

Column name pattern \* prefix(Pivot key value)middle(expression prefix)suffix

Prefix   -   Suffix

Column arrangement \*

Normal   Lateral (selected)

sum(daily\_count)   count

Add Source

The screenshot shows the Microsoft Data Factory Data Flow editor interface. On the left, there's a sidebar with navigation links like Home, Pipelines, Datasets, Data flows, and Power Query. The main area displays a pipeline named 'df\_transform\_cases\_deaths' with four stages: 'casesanddeathsSource', 'filterEuropeOnly', 'selectOnlyReqFields', and 'pivotCounts'. The 'pivotCounts' stage is highlighted. To the right of the stages is a 'Properties' panel with tabs for General and Related, and fields for Name (df\_transform\_cases\_deaths) and Description. Below the stages is a 'Pivot settings' section with tabs for Incoming schema, 1. Group by, 2. Pivot key, and 3. Pivoted columns. Under 'Pivoted columns', there's a column name pattern input ('prefix(Pivot key value)middle(expression prefix)suffix') and a 'Column arrangement' section with 'Normal' and 'Lateral' options, where 'Lateral' is selected. At the bottom, there are buttons for sum(daily\_count) and count, along with a '+' and '-' button for adding or removing columns.

Data Factory < Validate all Publish all 2

### Factory Resources

- Pipelines
  - pl\_ ingest\_ecdc\_data
  - pl\_ ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets
  - ds\_raw\_cases\_deaths
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
- Data flows
  - df\_transform\_cases\_deaths
- Power Query 0

Filter resources by name +

✓ Validate  Data flow debug [Debug Settings](#)

Add Source

Parameters Settings

+ New

### Set properties

Name

Linked service \*

File path  /  /

First row as header

Import schema

From connection/store  From sample file  None

OK Back Cancel

Data Factory   Validate all   Publish all (3)

Factory Resources   Filter resources by name

Pipelines: pl\_ingest\_ecdc\_data, pl\_ingest\_population\_data

Change Data Capture (preview): 0

Datasets: ds\_country\_lookup, ds\_ecdc\_data\_dl, ds\_ecdc\_filelist, ds\_http\_ecdc\_data, ds\_populationdata\_gz, ds\_populationdata\_ts, ds\_raw\_cases\_deaths

Data flows: df\_transform\_cases\_deaths

Power Query: 0

Validate   Data flow debug   Debug Settings

**df\_transform\_cases\_deaths**

```
graph LR; casesanddeathsSource[casesanddeathsSource] --> filterEuropeOnly[filterEuropeOnly]; filterEuropeOnly --> selectOnlyReqFields[selectOnlyReqFields]; selectOnlyReqFields --> pivotCounts[pivotCounts]; countryLookup[countryLookup] --> AddSource[Add Source];
```

**Properties**

**General**

Name \*: df\_transform\_cases\_deaths

Description:

**Source settings**

Output stream name \*: countryLookup

Description: Import data from ds\_country\_lookup

Reset

Source type \*: Dataset

Dataset \*: ds\_country\_lookup

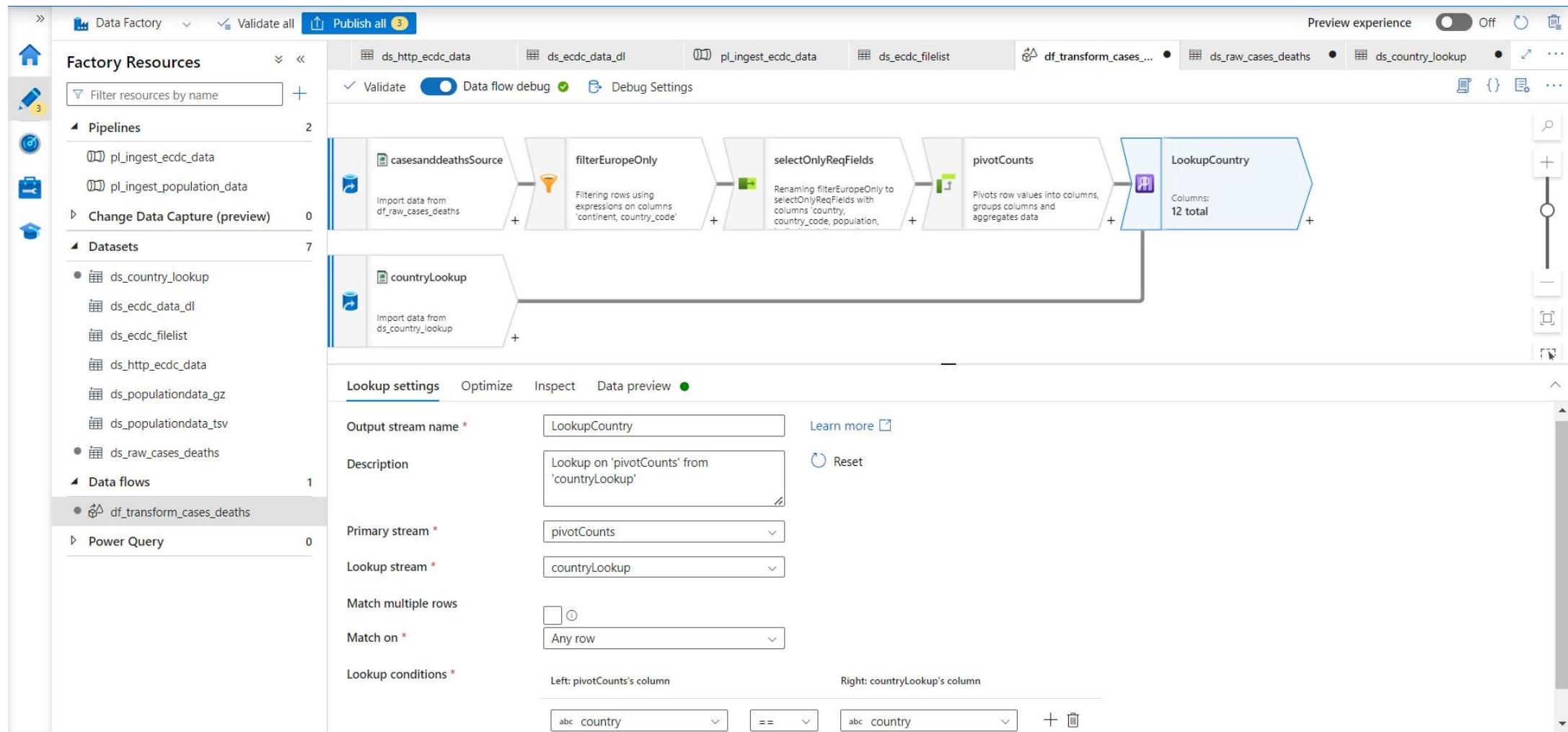
Test connection   Open   New

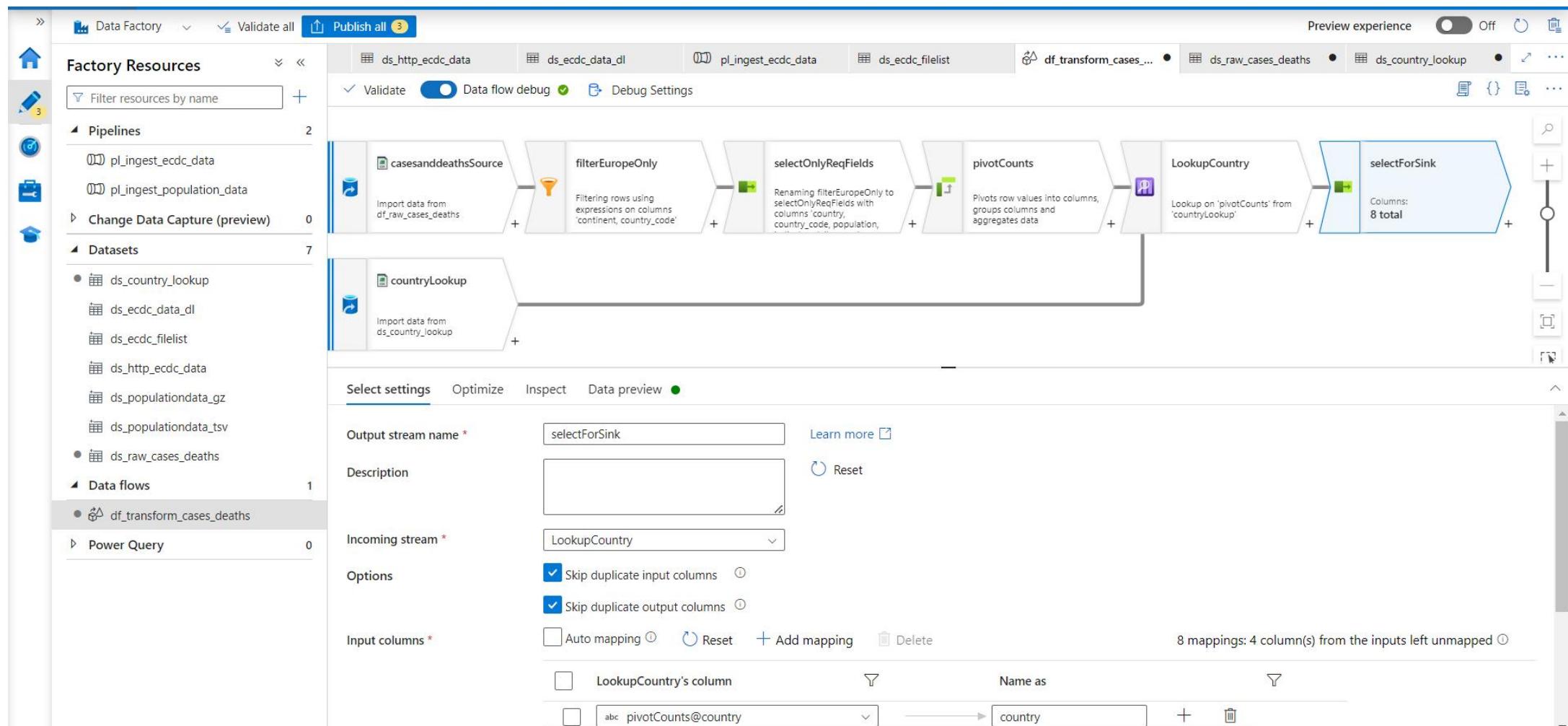
Options

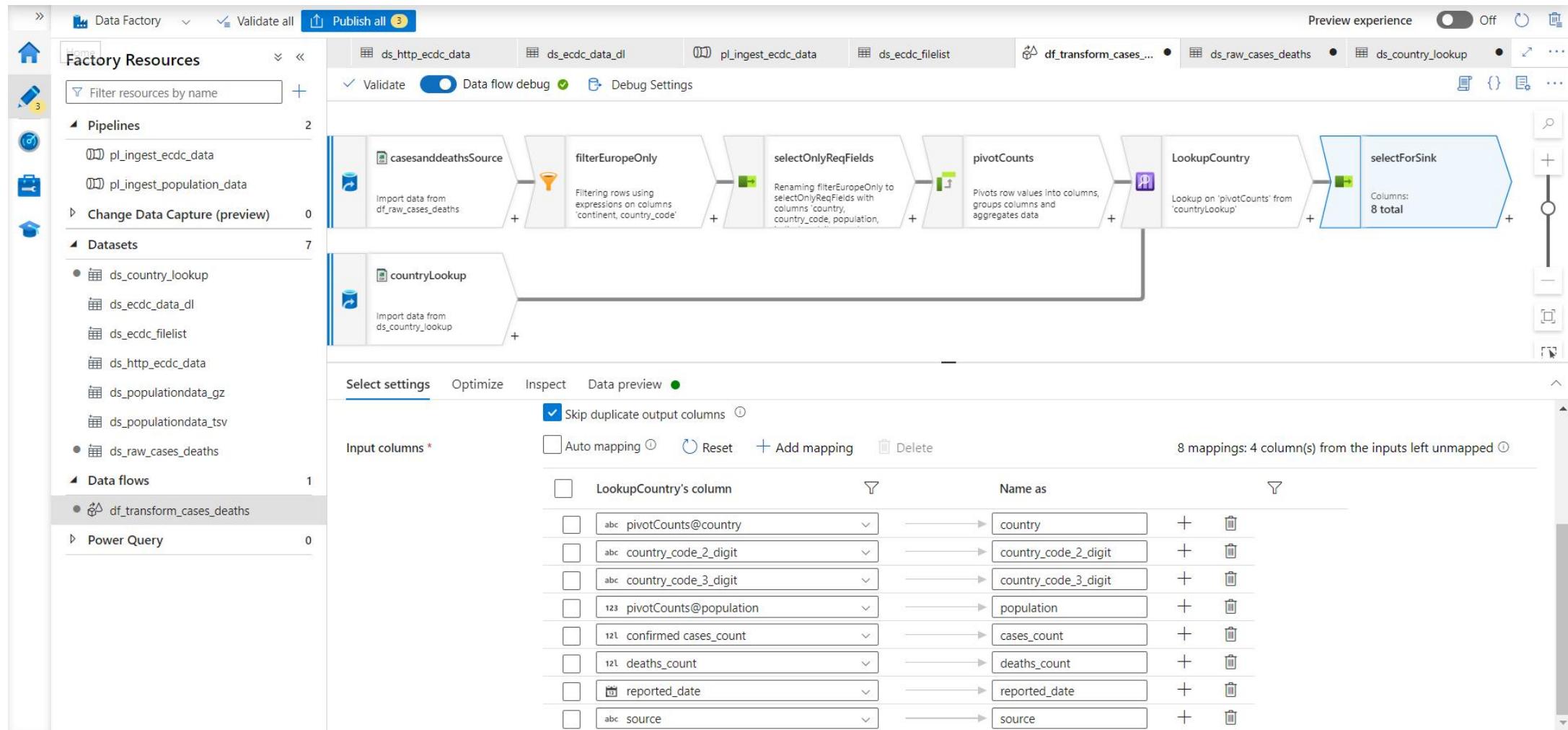
Allow schema drift

Infer drifted column types

Preview experience: Off







Data Factory < Validate all Publish all

### Factory Resources

- Pipelines: pl\_ingest\_ecdc\_data, pl\_ingest\_population\_data
- Change Data Capture (preview): 0
- Datasets: ds\_country\_lookup, ds\_ecdc\_data\_dl, ds\_ecdc\_filelist, ds\_http\_ecdc\_data, ds\_populationdata\_gz, ds\_populationdata\_tsv, ds\_raw\_cases\_deaths
- Data flows: df\_transform\_cases\_deaths
- Power Query: 0

**Set properties**

Name: ds\_processed\_cases\_deaths

Linked service: ls\_azdl\_ecovidreportingdatalake

File path: processed / ecdc/cases\_deaths / File name

First row as header:

Import schema:  From connection/store  From sample file  None

> Advanced

**Sink**

Output stream name: CasesandDeathsSink

Description: Add sink dataset

Incoming stream: selectForSink

Sink type: Dataset

Dataset: Select...

Options:  Allow schema drift  Validate schema

OK Back Cancel

The screenshot shows the Microsoft Azure Data Factory interface. On the left, the 'Factory Resources' sidebar lists various resources: Pipelines (pl\_ingest\_ecdc\_data, pl\_ingest\_population\_data), Change Data Capture (0), Datasets (ds\_country\_lookup, ds\_ecdc\_data\_dl, ds\_ecdc\_filelist, ds\_http\_ecdc\_data, ds\_populationdata\_gz, ds\_populationdata\_tsv, ds\_raw\_cases\_deaths), Data flows (df\_transform\_cases\_deaths), and Power Query (0). The 'df\_transform\_cases\_deaths' data flow is currently selected.

The main workspace displays a data flow pipeline with three stages: 'selectOnlyReqFields', 'pivotCounts', and 'LookupCountry'. The 'selectOnlyReqFields' stage has a note: 'Renaming filterEuropeOnly to selectOnlyReqFields with columns 'country, country\_code, population,''. The 'pivotCounts' stage has a note: 'Pivots row values into columns, groups columns and aggregates data'. The 'LookupCountry' stage has a note: 'Lookup on 'pivotCounts' from 'countryLookup''. Below the pipeline, the 'Sink' tab is selected in the configuration pane.

The 'Sink' configuration pane includes the following fields:

- Output stream name: CasesandDeathsSink
- Description: Add sink dataset
- Incoming stream: selectForSink
- Sink type: Dataset (selected)
- Dataset: Select... (dropdown menu)
- Options:
  - Allow schema drift
  - Validate schema

On the right side of the configuration pane, there is a 'Set properties' section with fields for Name (ds\_processed\_cases\_deaths), Linked service (ls\_azdl\_ecovidreportingdatalake), File path (processed / ecdc/cases\_deaths / File name), First row as header (checked), Import schema (None selected), and an 'Advanced' section.

Preview experience  Off

Data Factory Validate all Publish all 1

Factory Resources < <

Filter resources by name +

Pipelines 3

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 8

- ds\_country\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths

Data flows 1

- df\_transform\_cases\_deaths

Power Query 0

Activities < <

Validate Cancel options Add trigger  Data flow debug

Search activities

- > Move & transform
- > Synapse
- > Azure Data Explorer
- > Azure Function
- > Batch Service
- > Databricks
- > Data Lake Analytics
- > General
- > HDInsight
- > Iteration & conditionals
- > Machine Learning
- > Power Query

Data flow df\_cases\_deaths

General Settings Parameters User properties

Data flow \* df\_transform\_cases\_deaths

Run on (Azure IR) \* AutoResolveIntegrationRuntime

Compute size \* Small

Advanced

Logging level \*  Verbose  Basic  None

Sink properties

Staging

Properties

General Related

Name \* pl\_cases\_deaths\_data\_processed

Description

Annotations

» Data Factory Validate all Publish all 1 Preview experience Off

Factory Resources

- Pipelines 3
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 8
  - ds\_country\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
- Data flows 1
  - df\_transform\_cases\_deaths
- Power Query 0

Activities

- Move & transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

✓ Validate ⏺ Cancel options ⏹ Add trigger ⏹ Data flow debug ✓

Data flow df\_cases\_deaths

Properties

General Related

Name \* pl\_cases\_deaths\_data\_processed

Description

Annotations + New

Pipeline run ID: 4fc59733-09a4-4e19-850f-430872f885c1

All status Export to CSV

Showing 1 - 1 of 1 items

Activity name	Status	Activity type	Run start	Duration	Lc
df_cases_deaths	Succeeded	Data flow	6/30/2023, 8:35:54 PM	1m 22s	

## Microsoft Azure Storage Explorer

File Edit View Help



EXPLORER

Search for resources

raw x \$logs x population x configs x lookup x processed x

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undelete Folder Statistics Refresh

← → ↑ Active blobs (default) processed > ecdc > cases\_deaths Filter by prefix (case-sensitive)

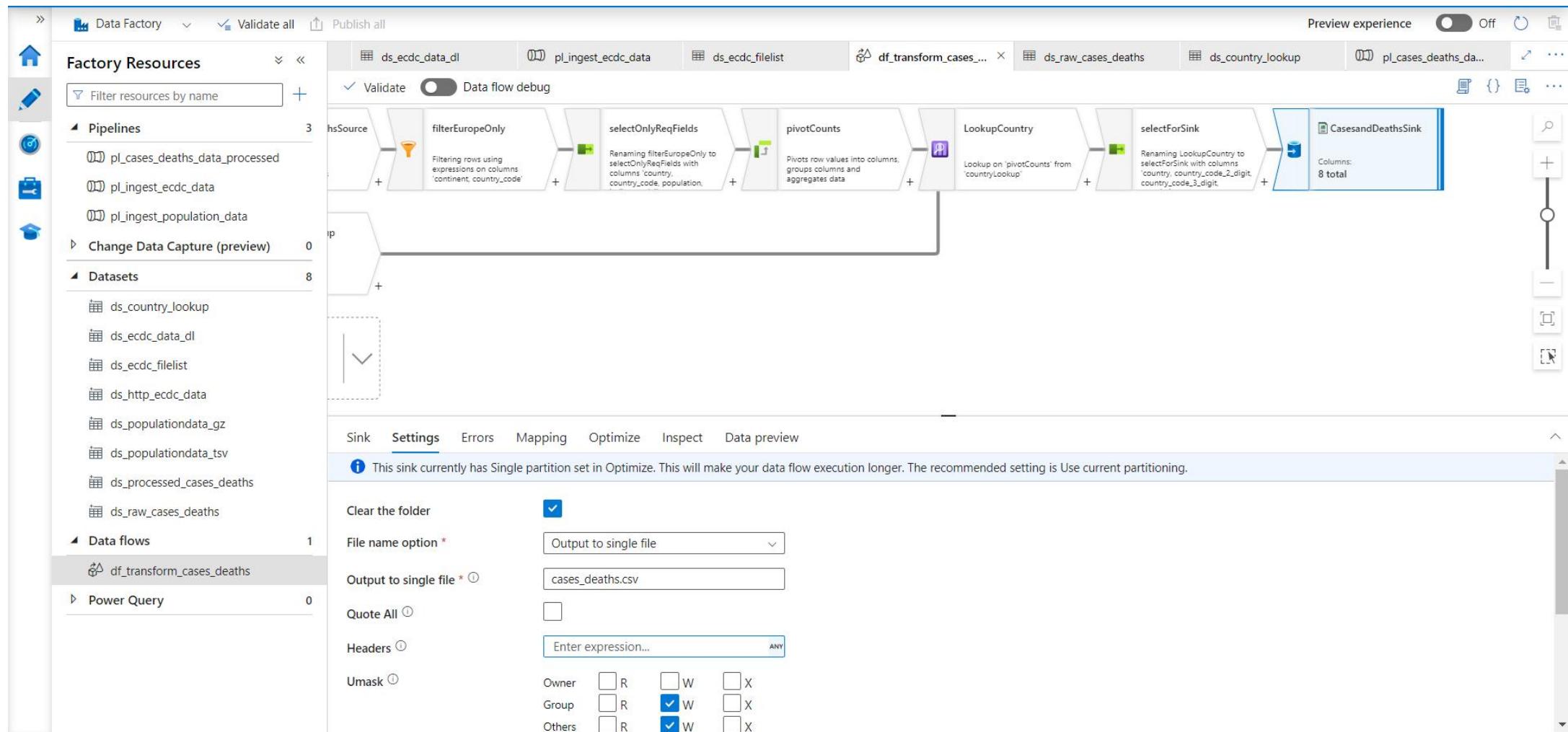
Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID
_SUCCESS	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	0 B	Active	
part-00000-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	6.81 KiB	Active	
part-00002-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	7.09 KiB	Active	
part-00003-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	13.10 KiB	Active	
part-00004-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	7.14 KiB	Active	
part-00005-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	7.76 KiB	Active	
part-00006-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	7.87 KiB	Active	
part-00007-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	6.29 KiB	Active	
part-00010-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	14.72 KiB	Active	
part-00011-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	13.83 KiB	Active	
part-00012-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	18.91 KiB	Active	
part-00013-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	6.63 KiB	Active	
part-00014-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	14.05 KiB	Active	
part-00018-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	6.49 KiB	Active	
part-00019-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	13.06 KiB	Active	
part-00021-c8f26247-79da-4338-9bd5-a37a5c0259d6-c000.csv	Hot (inferred)		6/30/2023 8:36 PM	Block Blob	application/octet-stream	21.02 KiB	Active	

Showing 1 to 100 of 129 cached items

Activities

Clear completed Clear successful

- Successfully created blob container 'processed'
- Transfer of 'C:\Users\Deheraji Bajjuri\Downloads\case\_deaths\_uk\_ind\_only.csv' to 'raw/ecdc/' complete: 1 item transferred (used SAS, discovery completed) Started at: 6/30/2023 6:37 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Transfer of 'C:\Users\Deheraji Bajjuri\Downloads\country\_lookup.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed) Started at: 6/30/2023 6:37 PM, Duration: 5 seconds [Copy AzCopy Command to Clipboard](#)
- Successfully created blob container 'lookup'
- Transfer of 'C:\Users\Deheraji Bajjuri\Downloads\code\_file\_list.json' to 'configs/' complete: 1 item transferred (used SAS, discovery completed) [Copy AzCopy Command to Clipboard](#)



## Pipeline runs

Triggered   Debug



Rerun   Cancel options



Refresh   Edit columns

List

Gantt

Filter by run ID or name

Local time : Last 24 hours

Pipeline name : All

Status : All

Runs : Latest runs

Triggered by : All

Add filter

Copy filters   Export to CSV



Showing 1 - 2 items

Last refreshed 0 minutes ago

<input type="checkbox"/> Pipeline name ↑↓	Run start ↑↓	Run end ↑↓	Duration	Triggered by	Status ↑↓	Run	Parameters	Annotations	Run
<input type="checkbox"/> pl_cases_deaths_data_process...	6/30/2023, 8:44:35 PM	6/30/2023, 8:48:53 PM	4m 18s	Manual trigger	Succeeded	Original			e4
<input type="checkbox"/> pl_ingest_population_data	6/29/2023, 10:52:45 PM	6/29/2023, 10:53:09 PM	24s	blob_ingestpopulati...	Succeeded	Original			9d

## Notifications

Alerts & metrics

## Microsoft Azure Storage Explorer

File Edit View Help



EXPLORER

Search for resources

raw x \$logs x population x configs x lookup x processed x

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undo Refresh

← → ↑ Active blobs (default) processed &gt; ecdc &gt; cases\_deaths Filter by prefix (case-sensitive)

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted Time	Lease St
cases_deaths.csv	Hot (inferred)		6/30/2023 8:48 PM	Block Blob	application/octet-stream	1.33 MiB	Active				

- Collapse all Refresh all
- Quick Access
- Emulator & Attached
- Storage Accounts
- Azure for Students (AvinashChinta@my.unt.edu)
  - Storage Accounts
    - cs71003200143e43ad6
    - dbstoragegeocb4jzhwww7hy (ADLS Gen2)
    - dlspractise (ADLS Gen2)
    - ercovidreportingdatalake (ADLS Gen2)
      - Blob Containers
        - \$logs
        - lookup
        - processed
        - raw
      - File Shares
      - Queues
      - Tables
    - ercovidreportingsa
      - Blob Containers
        - \$logs
        - configs
        - population
      - File Shares
      - Queues
      - Tables
    - Disk
      - cloud-shell-storage-southcentralus
      - covid-reporting-rg
      - databricks-rg-dbs-hands-on-bj47lln3b
      - NetworkWatcherRG
      - rg-avinash-hands-on

Showing 1 to 1 of 1 cached items

## Activities

[Clear completed](#) [Clear successful](#)

- ✓ Successfully created blob container 'processed'
- ✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\case\_deaths\_uk\_ind\_only.csv' to 'raw/ecdc/' complete: 1 item transferred (used SAS, discovery completed) Started at: 6/30/2023 6:37 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- ✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\country\_lookup.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed) Started at: 6/30/2023 6:37 PM, Duration: 5 seconds [Copy AzCopy Command to Clipboard](#)

Validate all Publish all 1

### Factory Resources

Pipelines

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview)

Datasets

- ds\_country\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths

Data flows

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query

HospitalAdmissions

Columns: 0 total

Add Source

Source settings

Source type \*

- Dataset
- Inline

Dataset \*

Select... + New

Options

- Allow schema drift
- Infer drifted column types
- Validate schema

Sampling \*

- Enable
- Disable

Set properties

Name

ds\_raw\_hospital\_admissions

Linked service \*

ls\_azdl\_ercovidreportingdatalake

File path

raw / ecdc / hospital\_admissions.csv

First row as header

Import schema

- From connection/store
- From sample file
- None

Advanced

OK Back Cancel

Data Factory   Validate all   Publish all (2)

Preview experience   Off

### Factory Resources

- Pipelines
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 9
  - ds\_country\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

Filter resources by name +

Validate Data flow debug

HospitalAdmissions  
Columns: 7 total

Add Source

Properties

General   Related

Name \* df\_transform\_hospital\_admissions

Description

Source settings   Source options   **Projection**   Optimize   Inspect   Data preview

Column name	Type	Format
country	string	Specify format
indicator	string	Specify format
date	date	Specify format
year_week	string	Specify format
value	double	Specify format
source	string	Specify format
url	string	Specify format

Preview experience Off

Data Factory Data Factory Validate all Publish all 2

Factory Resources Filter resources by name +

Pipelines 3

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 9

- ds\_country\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows 2

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query 0

Properties General Related

Name \* df\_transform\_hospital\_admissions

Description

Validate Data flow debug ●

HospitalAdmissions Import data from ds\_raw\_hospital\_admissions + selectOnlyReqFields Columns: 6 total +

Add Source

Select settings Optimize Inspect Data preview

Output stream name \* selectOnlyReqFields Learn more □

Description Renaming HospitalAdmissions to selectOnlyReqFields with columns 'country, indicator, reported\_date, ...' Reset

Incoming stream \* HospitalAdmissions

Options  Skip duplicate input columns ⓘ  Skip duplicate output columns ⓘ

Input columns \*

```
graph LR; SA[Import data from ds_raw_hospital_admissions] --> S[selectOnlyReqFields]; S --> DS[ds_raw_cases_deaths]
```

Properties General Related

Name \* df\_transform\_hospital\_admissions

Description

Validate Data flow debug ●

HospitalAdmissions Import data from ds\_raw\_hospital\_admissions + selectOnlyReqFields Columns: 6 total +

Add Source

Select settings Optimize Inspect Data preview

Output stream name \* selectOnlyReqFields Learn more □

Description Renaming HospitalAdmissions to selectOnlyReqFields with columns 'country, indicator, reported\_date, ...' Reset

Incoming stream \* HospitalAdmissions

Options  Skip duplicate input columns ⓘ  Skip duplicate output columns ⓘ

Input columns \*

Factory Resources < >

Filter resources by name +

Pipelines 3

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 9

- ds\_country\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows 2

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query 0

Validate all Publish all 2

Preview experience Off

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

HospitalAdmissions Import data from ds\_raw\_hospital\_admissions

selectOnlyReqFields Columns: 6 total

Add Source

Select settings Optimize Inspect Data preview

Input columns\* Auto mapping ⚡ Reset Add mapping Delete 6 mappings: 1 column(s) from the inputs left unmapped ⚡

HospitalAdmissions's column	Name as
abc country	country
abc indicator	indicator
date	reported_date
abc year_week	reported_year_week
1.2 value	value
abc source	source

Factory Resources < Filter resources by name +

Pipelines 3

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 9

- ds\_country\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows 2

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query 0

Validate all Publish all 2

Preview experience Off

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

Diagram:

```
graph LR; HospitalAdmissions[HospitalAdmissions] --> selectOnlyReqFields[selectOnlyReqFields]; selectOnlyReqFields --> countrySource[countrySource];
```

Source settings

Output stream name \* countrySource

Description Import data from ds\_country\_lookup

Source type \*

Dataset

ds\_country\_lookup

Options Allow schema drift

Data Factory    Validate all    Publish all (2)

Preview experience    Off

### Factory Resources

- Pipelines
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 9
  - ds\_country\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

Filter resources by name +

✓ Validate    Data flow debug ●

```
graph LR; A[HospitalAdmissions] --> B[selectOnlyReqFields]; B --> C[lookupCountry]; D[countrySource] --> C;
```

Properties

General    Related

Name \* df\_transform\_hospital\_admissions

Description

Lookup settings

Output stream name \* lookupCountry

Description Lookup on 'selectOnlyReqFields' from 'countrySource'

Primary stream \* selectOnlyReqFields

Lookup stream \* countrySource

Match multiple rows

Match on \* Any row

Lookup conditions \*

Left: selectOnlyReqFields's column    Right: countrySource's column

abc country    ==    abc country

Validate all Publish all 2

Preview experience Off

### Factory Resources

Filter resources by name

- Pipelines
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets
  - ds\_country\_lookup
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

Validate Data flow debug

```
graph LR; A[HospitalAdmissions] --> B[selectOnlyReqFields]; B --> C[lookupCountry]; C --> D[selectReqFields2];
```

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

Select settings Optimize Inspect Data preview

Output stream name \* selectReqFields2 Learn more

Description Renaming lookupCountry to selectReqFields2 with columns 'country', 'indicator', 'reported\_date',

Incoming stream \* lookupCountry

Options

- Skip duplicate input columns
- Skip duplicate output columns

Input columns \* Auto mapping Reset + Add mapping Delete 10 mappings: 2 column(s) from the inputs left unmapped

lookupCountry's column	Name as
abc selectOnlyReqFields@country	country
abc indicator	indicator

Validate all ↗ Publish all 2

Preview experience Off

**Factory Resources**

- Pipelines
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 9
  - ds\_country\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

**Properties**

General Related

Name \* df\_transform\_hospital\_admissions

Description

```

graph LR
    A[HospitalAdmissions] --> B[selectOnlyReqFields]
    B --> C[lookupCountry]
    C --> D[selectReqFields2]
  
```

**Select settings**

Input columns \*

lookupCountry's column	Name as
abc selectOnlyReqFields@country	country
abc indicator	indicator
abc reported_date	reported_date
abc reported_year_week	reported_year_week
1.2 value	value
abc source	source
abc countrySource@country	country
abc country_code_2_digit	country_code_2_digit
abc country_code_3_digit	country_code_3_digit
abc population	population

10 mappings: 2 column(s) from the inputs left unmapped

Factory Resources < >

Validate all Publish all 2

Preview experience Off

Validate Data flow debug

**Pipelines**

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

**Change Data Capture (preview)** 0

**Datasets** 9

- ds\_country\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

**Data flows** 2

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

**Power Query** 0

A data flow diagram illustrating a pipeline. It starts with an 'Import data from ds\_raw\_hospital\_admissions' source, followed by a 'selectOnlyReqFields' transformation (Renaming HospitalAdmissions to selectOnlyReqFields with columns 'country', 'indicator', 'reported\_date'). This is then followed by a 'lookupCountry' transformation (Lookup on 'selectOnlyReqFields' from 'countrySource'), which is connected to a 'selectReqFields2' transformation (Columns: 9 total). The final output stage is shown on the right.

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

Select settings Optimize Inspect Data preview

Schema

Input Output

Number of columns Updated* 0		Dropped 2	Unchanged 9	Total 9
Order ↑	Column ↑	Type ↑↓	Updated ↑↓	Input column ↑↓
1	country	abc string		selectOnlyReqFields@country
2	indicator	abc string		indicator
3	reported_date	date		reported_date
4	reported_year_week	abc string		reported_year_week
5	value	1.2 double		value
6	source	abc string		source
7	country_code_2_digit	abc string		country_code_2_digit
8	country_code_3_digit	abc string		country_code_3_digit
9	population	abc string		population

Expand Data Factory   Validate all   Publish all (2)

Preview experience: Off

### Factory Resources

- Pipelines: pl\_cases\_deaths\_data\_processed, pl\_ingest\_ecdc\_data, pl\_ingest\_population\_data
- Datasets: ds\_country\_lookup, ds\_ecdc\_data\_dl, ds\_ecdc\_filelist, ds\_http\_ecdc\_data, ds\_populationdata\_gz, ds\_populationdata\_tsv, ds\_processed\_cases\_deaths, ds\_raw\_cases\_deaths, ds\_raw\_hospital\_admissions
- Data flows: df\_transform\_hospital\_admissions, df\_transform\_cases\_deaths
- Power Query: 0

**Properties**

**General**

Name: df\_transform\_hospital\_admissions

Description:

```

graph LR
    ReqFields[ReqFields] --> lookupCountry[lookupCountry]
    lookupCountry --> selectReqFields2[selectReqFields2]
    selectReqFields2 --> Weekly[Weekly]
    selectReqFields2 --> Daily[Daily]
    
```

**Conditional split settings**

Output stream name: SplitDailyfromWeekly

Description: Conditionally distributing the data in indicator groups, based on columns '{1}'

Incoming stream: selectReqFields2

Split on: First matching condition

Split condition	Stream names	Condition
	Weekly	indicator == 'Weekly new hospital admissions per 100k'    indicator ...

»



## Dataflow expression builder

[Expression reference documentation](#)[SplitDailyfromWeekly](#)

### Expression

```
indicator == 'Weekly new hospital admissions per 100k' || indicator == 'Weekly new ICU admissions per 100k'
```

+ - \* / || && ! ^ == === <=> != > < >= <= [ ]

#### Expression elements

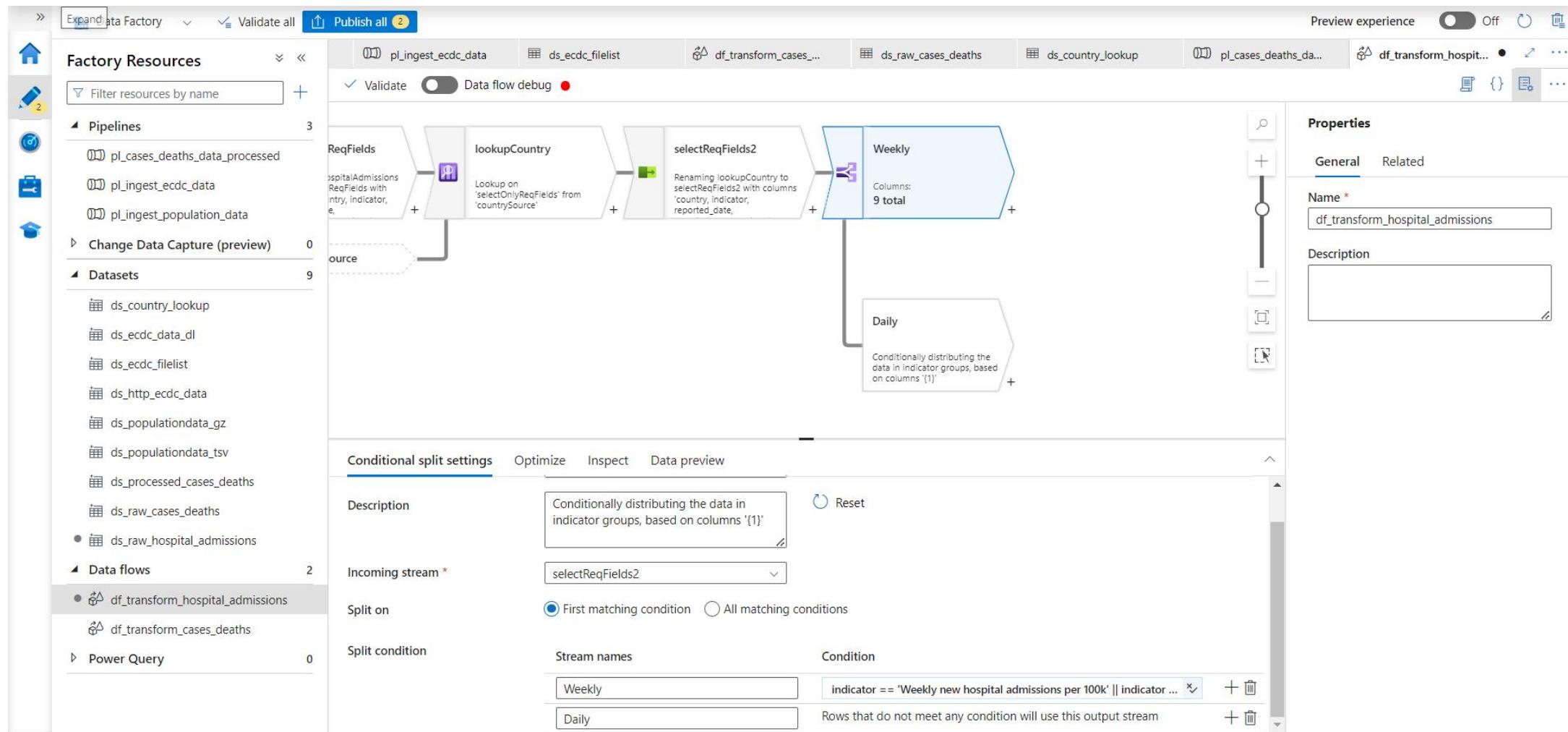
All  
Functions  
Input schema  
Parameters  
Cached lookup  
Data flow library functions

#### Expression values

Filter by keyword  
+ Create new ▾  
abc country  
abc indicator  
abc reported\_date  
abc reported\_year\_week  
123 value  
abc source

#### Data preview

[Save and finish](#)[Cancel](#)[Clear contents](#)



Microsoft Azure Storage Explorer

File Edit View Help

EXPLORER

Search for resources

Collapse all Refresh all

Quick Access

Emulator & Attached

Storage Accounts

Azure for Students (AvinashChinta@my.unt.edu)

- Storage Accounts
  - cs71003200143e43ad6
  - dbstoragegeocb4jzhwww7hy (ADLS Gen2)
  - dlspractise (ADLS Gen2)
- ercovidreportingdatalake (ADLS Gen2)
  - Blob Containers
    - \$logs
    - lookup (selected)
    - processed
    - raw
  - File Shares
  - Queues
  - Tables
- ercovidreportingsa
  - Blob Containers
    - \$logs
    - configs
    - population
  - File Shares
  - Queues
  - Tables

Disks

- cloud-shell-storage-southcentralus
- covid-reporting-rg
- databricks-rg-dbs-hanson-bj47lln3b
- NetworkWatcherRG
- rg-avinash-hanson

raw \$logs population configs lookup processed

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undelete Folder Statistics Refresh

← → ↑ Active blobs (default) lookup Filter by prefix (case-sensitive)

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted Time	Le
country_lookup.csv	Hot (inferred)		6/30/2023 6:37 PM	Block Blob	application/vnd.ms-excel	6.86 KiB	Active				
dim_date.csv	Hot (inferred)		6/30/2023 9:35 PM	Block Blob	application/vnd.ms-excel	73.67 KiB	Active				

Showing 1 to 2 of 2 cached items

Activities

Clear completed Clear successful

- Transfer of 'C:\Users\AvinashChinta\Downloads\dim\_date.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 9:35 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Successfully created blob container 'processed'
- Transfer of 'C:\Users\AvinashChinta\Downloads\country\_lookup.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)
- Transfer of 'C:\Users\AvinashChinta\Downloads\dim\_date.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)

» Data Factory Validate all Publish all (3)

Preview experience Off

**Factory Resources** Filter resources by name ...

**Pipelines** 3

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

**Change Data Capture (preview)** 0

**Datasets** 10

- ds\_country\_lookup
- ds\_dimdate\_lookup**
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

**Data flows** 2

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

**Power Query** 0

**Properties**

**General** **Related (1)**

**Name \***

**Description**

**Annotations** + New

**DelimitedText**  
**CSV**

**ds\_dimdate\_lookup**

**Connection** **Schema** **Parameters**

**Linked service \***  Test connection Edit New Learn more

**File path \***  /  /  Browse Preview

**Compression type**

**Column delimiter**

**Row delimiter**

**Encoding**

**Quote character**

**Escape character**

**First row as header**

**Null value**

Data Factory    Validate all    Publish all

Factory Resources    Filter resources by name

Pipelines

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview)

Datasets

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query

Preview experience Off

Properties

General    Related

Name \* df\_transform\_hospital\_admissions

Description

Diagram:

```
graph LR; countrySource[Import data from ds_country_lookup] --> DeriveECDCYearWeek[DeriveECDCYearWeek]; dimdataSource[Import data from ds_dimdate_lookup] --> DeriveECDCYearWeek;
```

Derived column's settings

Output stream name \* DeriveECDCYearWeek

Description Creating/updating the columns 'date\_key, date, year, month, day, day\_name, day\_of\_year,'

Incoming stream \* dimdataSource

Add Clone Delete Open expression builder

Columns \* ①

Column	Expression
ecdc_year_week	year + '-W' + lpad(week_of_year,2,'0')

Data Factory   Validate all   Publish all (3)

Preview experience   Off

### Factory Resources

- Pipelines
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

Filter resources by name

Validate   Data flow debug

Properties

General   Related

Name \* df\_transform\_hospital\_admissions

Description

Aggregate settings

Output stream name \* AggDimdate

Description Aggregating data by 'ecdc\_year\_week' producing columns 'week\_start\_date', 'week\_end\_date'

Incoming stream \* DeriveECDCYearWeek

Group by   Aggregates

Grouped by: ecdc\_year\_week

Add   Clone   Delete   Open expression builder

Column	Expression
week_start_date	min(date)
week_end_date	max(date)

Data Factory    Validate all    Publish all (3)    Preview experience Off

### Factory Resources

- Pipelines
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

Validate Data flow debug

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

Aggregate settings

Output stream name \* AggDimdate

Description Aggregating data by 'ecdc\_year\_week' producing columns 'week\_start\_date', 'week\_end\_date'

Incoming stream \* DeriveECDCYearWeek

Group by Aggregates

Columns Name as

year + '-W' + Ipad(week\_of\_year,2,'0') ecdc\_year\_week

Published Data Flow: df\_transform\_hospital\_admissions

**Pipeline:** pl\_cases\_deaths\_data\_processed

**Data Flow Components:**

- selectOnlyReqFields:** Renaming HospitalAdmissions to selectOnlyReqFields with columns 'country', 'indicator'.
- lookupCountry:** Lookup on 'selectOnlyReqFields' from 'countrySource'.
- selectReqFields2:** Renaming lookupCountry to selectReqFields2 with columns 'country', 'indicator', 'reported\_date'.
- Weekly:** Conditionally distributing the data in indicator groups, based on columns '(1)'.
- JoinWithDate:** Columns: 12 total.
- AggDimdate:** Daily

**Properties:**

- Name:** df\_transform\_hospital\_admissions
- Description:** Inner join on 'SplitDailyfromWeekly@Weekly' and 'AggDimdate'

**Join settings:**

- Output stream name:** JoinWithDate
- Left stream:** SplitDailyfromWeekly@Weekly
- Right stream:** AggDimdate
- Join type:** Inner (selected)
- Use fuzzy matching:**
- Join conditions:**
  - Left: SplitDailyfromWeekly@Weekly's column: reported\_year\_week
  - Right: AggDimdate's column: ecdc\_year\_week

Home Data Factory Validate all Publish all 3 Preview experience Off

### Factory Resources

- Pipelines
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 10
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions**
  - df\_transform\_cases\_deaths
- Power Query 0

Validate Data flow debug

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

Join settings	Optimize	Inspect	Data preview	
Order ↑	Column ↑	Type ↑	Fed by ↑	Used in Join ↑
1	country	abc string	SplitDailyfromWeekly	
2	indicator	abc string	SplitDailyfromWeekly	
3	reported_date	abc date	SplitDailyfromWeekly	
4	reported_year_week	abc string	SplitDailyfromWeekly	✓
5	value	1.2 double	SplitDailyfromWeekly	
6	source	abc string	SplitDailyfromWeekly	
7	country_code_2_digit	abc string	SplitDailyfromWeekly	
8	country_code_3_digit	abc string	SplitDailyfromWeekly	
9	population	abc string	SplitDailyfromWeekly	
10	ecdc_year_week	abc string	AggDimdate	✓
11	week_start_date	abc string	AggDimdate	
12	week_end_date	abc string	AggDimdate	

Data Factory < Validate all Publish all 3

Factory Resources < Filter resources by name +

Pipelines 3

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 10

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows 2

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query 0

Validate Data flow debug

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

Pivot settings Optimize Inspect Data preview

Output stream name \* pivotWeekly

Description

Pivots row values into columns, groups columns and aggregates data

Incoming stream \* JoinWithDate

1. Group by 2. Pivot key 3. Pivoted columns

Columns	Name as
abc_country	country
abc_country_code_2_digit	country_code_2_digit
abc_country_code_3_digit	country_code_3_digit
abc_population	population
abc_reported_year_week	reported_year_week
abc_source	source

```
graph LR; lookupCountry[lookupCountry] --> selectReqFields2[selectReqFields2]; selectReqFields2 --> Weekly[Weekly]; Weekly --> JoinWithDate[JoinWithDate]; JoinWithDate --> pivotWeekly[pivotWeekly];
```

Validate all | Publish all (3)

Preview experience Off

### Factory Resources

Filter resources by name

- Pipelines (3)
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (10)
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows (2)
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query (0)

✓ Validate Data flow debug ●

The data flow consists of the following stages:

- lookupCountry**: Lookup on 'selectOnlyReqFields' from 'countrySource'.
- selectReqFields2**: Renaming 'lookupCountry' to 'selectReqFields2' with columns 'country', 'indicator', and 'reported date'.
- Weekly**: Conditionally distributing the data in 'selectReqFields2' based on indicator groups ('11').
- JoinWithDate**: Inner join on 'SplitDailyfromWeekly@Weekly' and 'AggDimdate'.
- pivotWeekly**: Pivots row values into columns, groups columns and aggregates data. It has 8 total columns.

**Pivot settings**

Output stream name \*: pivotWeekly

Description: Pivots row values into columns, groups columns and aggregates data

Incoming stream \*: JoinWithDate

1. Group by    2. Pivot key    3. Pivoted columns

Pivot key \*: abc indicator

Value:  
Weekly new hospital admissions per 100k  
Weekly new ICU admissions per 100k

Null value

**Properties**

General Related

Name \*: df\_transform\_hospital\_admissions

Description:

» Data Factory Validate all Publish all 3 Preview experience Off

### Factory Resources

Filter resources by name +

- Pipelines 3
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 10
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

✓ Validate Data flow debug ●

```
graph LR; A[lookupCountry] --> B[selectReqFields2]; B --> C[Weekly]; C --> D[JoinWithDate]; D --> E[pivotWeekly];
```

**Pivot settings** Optimize Inspect Data preview

Output stream name \* pivotWeekly ? Help Learn more

Description  
Pivots row values into columns, groups columns and aggregates data

Incoming stream \* JoinWithDate

1. Group by    2. Pivot key    3. Pivoted columns

Column name pattern \* prefix{Pivot key value}middle{expression prefix}suffix

Column arrangement \* Normal Lateral

sum(value) 1.2 count + -

**Properties**

General Related

Name \* df\_transform\_hospital\_admissions

Description

Expand Data Factory   Validate all   Publish all (3)

Preview experience: Off

**Factory Resources**

- Pipelines (3)
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (10)
  - ds\_country\_lookup
  - ds\_dimdate\_lookup (●)
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions (●)
- Data flows (2)
  - df\_transform\_hospital\_admissions (●)
  - df\_transform\_cases\_deaths
- Power Query (0)

**Properties**

General   Related

Name \* df\_transform\_hospital\_admissions

Description

The data flow consists of five main stages connected sequentially:

- lookupCountry**: A Lookup transformation that looks up 'selectOnlyReqFields' from 'countrySource'.
- selectReqFields2**: A Rename transformation that renames the columns from the lookup result to 'country', 'indicator', and 'reported date'.
- Weekly**: A Conditional distribution transformation that groups the data by indicator groups based on the 'Indicator' column.
- JoinWithDate**: An Inner join transformation that joins the weekly data with a date dimension table ('DimDate') on 'SplitDailyfromWeekly@Week' and 'AggDimDate'.
- pivotWeekly**: A Pivot transformation that creates two new columns: 'Weekly new hospital admissions per 100k\_count' and 'Weekly new ICU admissions per 100k\_count'.

**Inspect** tab selected.

**Schema**

Input		Output	
Number of columns New* 2		Dropped 6	
		Unchanged 6	
		Total 8	
Order ↑↓	Column ↑↓	Type ↑↓	Aggregated as ↑↓
1	country	abc string	Group by
2	country_code_2_digit	abc string	Group by
3	country_code_3_digit	abc string	Group by
4	population	abc string	Group by
5	reported_year_week	abc string	Group by
6	source	abc string	Group by
7	Weekly new hospital admissions per 100k_count	1..2 double	Pivot
8	Weekly new ICU admissions per 100k_count	1..2 double	Pivot

Data Factory   Validate all   Publish all (3)

Preview experience   Off

### Factory Resources

Filter resources by name

- Pipelines
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 10
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

Validate   Data flow debug ●

The data flow diagram illustrates a sequence of steps:

- An incoming stream labeled "SplitDailyfromWeekly@Daily" enters the "AggDimdate" step.
- The output of "AggDimdate" splits into two parallel paths, each leading to a "Daily" step.
- The outputs from the "Daily" steps merge into a single path leading to the "pivotDaily" step.
- The "pivotDaily" step has 8 total columns.

**Properties**

General   Related

Name \* df\_transform\_hospital\_admissions

Description

Pivot settings

Incoming stream \* SplitDailyfromWeekly@Daily

1. Group by   2. Pivot key   3. Pivoted columns

Columns	Name as
abc_country	country
abc_country_code_2_digit	country_code_2_digit
abc_country_code_3_digit	country_code_3_digit
abc_population	population
reported_date	reported_date
abc_source	source

Expand Data Factory   Validate all   Publish all (3)   Preview experience   Off

Factory Resources   Filter resources by name

Pipelines

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query 0

Validate   Data flow debug

reported\_date (1)

AggDimdate

Daily

Conditionally distributing the data in indicator groups, based on columns '(1)'.

pivotDaily

Columns: 8 total

Properties

General   Related

Name \* df\_transform\_hospital\_admissions

Description

Pivot settings   Optimize   Inspect   Data preview

Output stream name \* pivotDaily

Description

Pivots row values into columns, groups columns and aggregates data

Incoming stream \* SplitDailyfromWeekly@Daily

1. Group by   2. Pivot key   3. Pivoted columns

Pivot key \* abc indicator

Value

Daily hospital occupancy

Daily ICU occupancy

Null value

Validate all Publish all

Factory Resources

Pipelines

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview)

Datasets

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query

Preview experience Off

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

Output stream name \* pivotDaily

Description Pivots row values into columns, groups columns and aggregates data

Incoming stream \* SplitDailyfromWeekly@Daily

1. Group by 2. Pivot key 3. Pivoted columns

Column name pattern \* prefix{Pivot key value}middle{expression prefix}suffix

Prefix \_ Suffix

Column arrangement \* Normal Lateral

sum(value) 1.2 count

```
graph LR; A[reported_date] --> B((AggDimdate)); B --> C[Daily]; C --> D[pivotDaily]; D --> E[sum(value)]; D --> F[count]
```

Expand Data Factory   Validate all   Publish all (3)

Preview experience Off

### Factory Resources

- Pipelines (3)
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (10)
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows (2)
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query (0)

Filter resources by name +

Validate Data flow debug ●

**Properties**

General Related

Name \* df\_transform\_hospital\_admissions

Description

Pivot settings Optimize **Inspect** Data preview

Schema **Input** Output

Number of columns <b>New</b> 2		Dropped 3	Unchanged 6	Total 8
Order ↑	Column ↑	Type	Aggregated as ↑	Based on ↑
1	country	abc string	Group by	country
2	country_code_2_digit	abc string	Group by	country_code_2_digit
3	country_code_3_digit	abc string	Group by	country_code_3_digit
4	population	abc string	Group by	population
5	reported_date	date	Group by	reported_date
6	source	abc string	Group by	source
7	Daily hospital occupancy_count	1..2 double	Pivot	value
8	Daily ICU occupancy_count	1..2 double	Pivot	value

Validate all | Publish all (3)

Preview experience Off

### Factory Resources

- Pipelines (3)
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (10)
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows (2)
  - df\_transform\_hospital\_admissions**
  - df\_transform\_cases\_deaths
- Power Query (0)

Filter resources by name

Validate Data flow debug

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

Sort settings

Output stream name \* sortWeekly

Description Sorting rows on columns 'reported\_year\_week, country'

Incoming stream \* pivotWeekly

Options \*
 Case insensitive
 Sort only within partition

Sort conditions \*

pivotWeekly's column	Order	Nulls first
abc reported_year_week	Descending	<input checked="" type="checkbox"/>
abc country	Ascending	<input checked="" type="checkbox"/>

Expand Data Factory   Validate all   Publish all (3)

Preview experience: Off

### Factory Resources

Filter resources by name: +

- Pipelines:
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 10
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

✓ Validate   Data flow debug: ●

```
graph LR; Daily[Daily] --> sortDaily[sortDaily];
```

The data flow consists of two main components: "Daily" and "sortDaily". The "Daily" component is described as "Conditionally distributing the data in indicator groups, based on columns '11'". It has one output stream labeled "pivotDaily". This stream is then processed by the "sortDaily" component, which is described as "Sorting rows on columns 'reported\_date, country'". The "sortDaily" component has an output stream labeled "sortDaily".

**Properties**

General   Related

Name \* df\_transform\_hospital\_admissions

Description

**Sort settings**

Output stream name \* sortDaily

Description Sorting rows on columns 'reported\_date, country'

Incoming stream \* pivotDaily

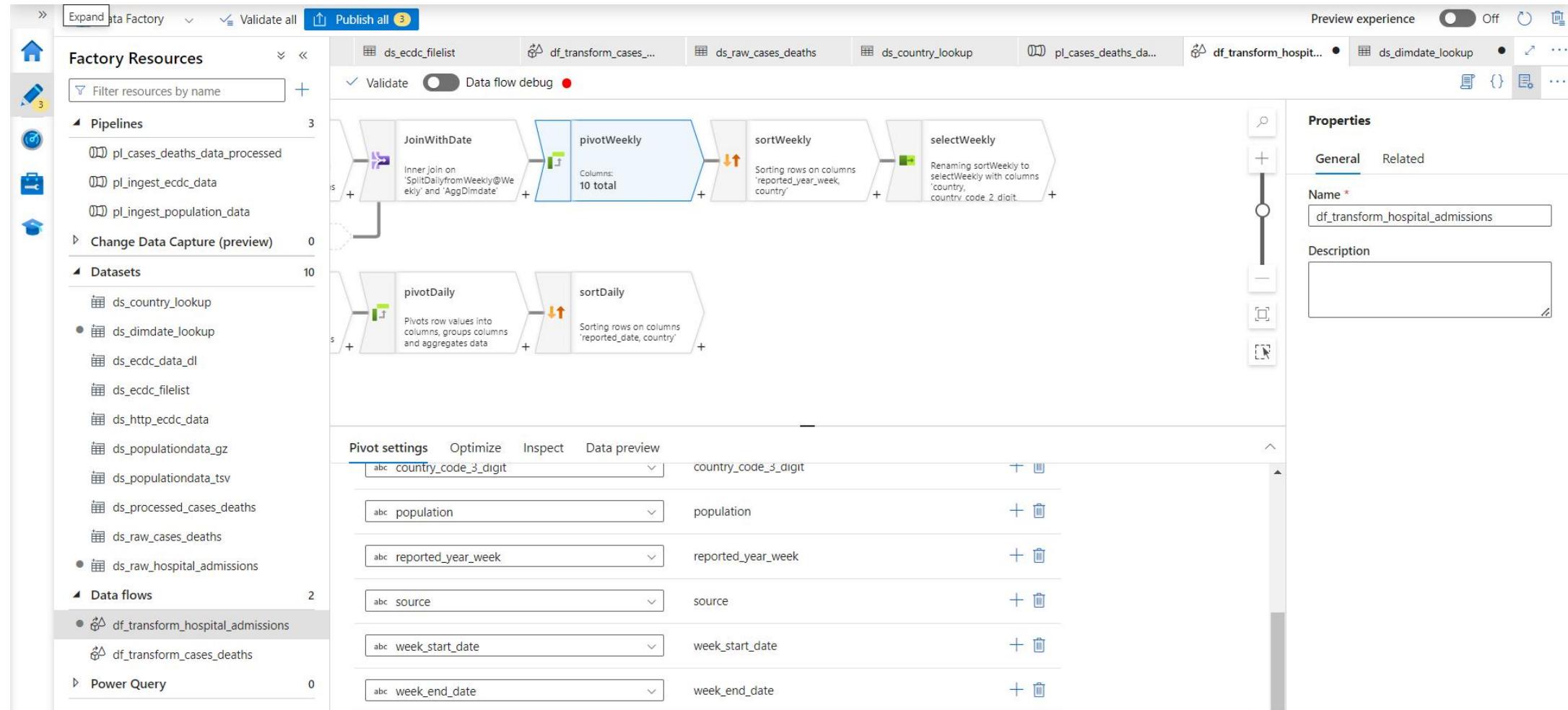
Options \*

Case insensitive

Sort only within partition

**Sort conditions \***

pivotDaily's column	Order	Nulls first
reported_date	Descending	<input checked="" type="checkbox"/>
country	Ascending	<input checked="" type="checkbox"/>



Expand Data Factory   Validate all   Publish all (3)   Preview experience Off

### Factory Resources

- Pipelines (3)
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (10)
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows (2)
  - df\_transform\_hospital\_admissions**
  - df\_transform\_cases\_deaths
- Power Query (0)

**Properties**

**General**

Name \* df\_transform\_hospital\_admissions

Description

**df\_transform\_hospital\_admissions**

**Data flow debug** (On)

**Flow Diagram**

```

graph LR
    A[JoinWithDate] --> B[pivotWeekly]
    B --> C[sortWeekly]
    C --> D[selectWeekly]
    
```

**Select settings**

with columns 'country, country\_code\_2\_digit,'

**Incoming stream \***

sortWeekly

**Options**

Skip duplicate input columns

Skip duplicate output columns

**Input columns \***

10 mappings: All inputs mapped

sortWeekly's column	Name as
abc_country	country
abc_country_code_2_digit	country_code_2_digit
abc_country_code_3_digit	country_code_3_digit
abc_population	population
abc_reported_year_week	reported_year_week
abc_week_start_date	reported_week_start_date
abc_week_end_date	reported_week_end_date
1.2 Weekly new hospital admissions per 100k_c...	new_hospital_occupancy...
1.2 Weekly new ICU admissions per 100k_count	new_icu_occupancy_count
abc_source	source

Data Factory    Validate all    Publish all (3)

Preview experience    Off

### Factory Resources

- Pipelines
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 10
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions**
  - df\_transform\_cases\_deaths
- Power Query 0

✓ Validate    Data flow debug ●

```

graph LR
    A[JoinWithDate] --> B[pivotWeekly]
    B --> C[sortWeekly]
    C --> D[selectWeekly]
  
```

Properties

General    Related

Name \* df\_transform\_hospital\_admissions

Description

Schema

Input		Output	
Number of columns	Updated* 4	Dropped 0	Unchanged 6
Total 10			
Order ↑	Column ↑	Type ↑	Updated ↑
1	country	abc string	country
2	country_code_2_digit	abc string	country_code_2_digit
3	country_code_3_digit	abc string	country_code_3_digit
4	population	abc string	population
5	reported_year_week	abc string	reported_year_week
6	reported_week_start_date	abc string	* week_start_date
7	reported_week_end_date	abc string	* week_end_date
8	new_hospital_occupancy_count	1..2 double	* Weekly new hospital admissions per 100k_count
9	new_icu_occupancy_count	1..2 double	* Weekly new ICU admissions per 100k_count
10	source	abc string	source

Factory Resources < Filter resources by name +

Pipelines 3

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 10

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows 2

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query 0

Preview experience Off

Validate all Publish all 3

✓ Validate Data flow debug ●

pivotDaily sortDaily selectDaily

Pivots row values into columns, groups columns and aggregates data  
Sorting rows on columns 'reported\_date, country'  
Columns: 8 total

Select settings Optimize Inspect Data preview

Description Renaming sortDaily to selectDaily with columns 'country, country\_code\_2\_digit,'

Incoming stream \* sortDaily

Options  Skip duplicate input columns ⓘ  Skip duplicate output columns ⓘ

Input columns \*  Auto mapping ⓘ Reset Add mapping Delete 8 mappings: All inputs mapped

sortDaily's column	Name as	+	–
abc_country	country	+	–
abc_country_code_2_digit	country_code_2_digit	+	–
abc_country_code_3_digit	country_code_3_digit	+	–
abc_population	population	+	–
reported_date	reported_date	+	–
1.2 Daily hospital occupancy_count	hospital_occupancy_cou...	+	–
1.2 Daily ICU occupancy_count	ICU_occupancy_count	+	–
abc_source	source	+	–

Properties

General Related

Name \* df\_transform\_hospital\_admissions

Description

☰ Data Factory ▾ ✓ Validate all Publish all 3

### Factory Resources

- Filter resources by name
- Pipelines (3)
  - pl\_cases\_deaths\_data\_processed
  - pl\_ingest\_ecdc\_data
  - pl\_ingest\_population\_data
- Change Data Capture (preview) (0)
- Datasets (10)
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows (2)
  - df\_transform\_hospital\_admissions**
  - df\_transform\_cases\_deaths
- Power Query (0)

✓ Validate Data flow debug ●

```

graph LR
    A[pivotDaily] --> B[sortDaily]
    B --> C[selectDaily]
    C -- "Columns: 8 total" --> D[...]
  
```

pivotDaily  
Pivots row values into columns, groups columns and aggregates data

sortDaily  
Sorting rows on columns 'reported\_date, country'

selectDaily  
Columns: 8 total

Select settings Optimize Inspect Data preview

**Properties**

General Related

Name \* df\_transform\_hospital\_admissions

Description

Schema		Input	Output	
Number of columns	Updated* 2	Dropped 0	Unchanged 6	Total 8
Order ↑↓	Column ↑↓	Type ↑↓	Updated ↑↓	Input column ↑↓
1	country	abc string		country
2	country_code_2_digit	abc string		country_code_2_digit
3	country_code_3_digit	abc string		country_code_3_digit
4	population	abc string		population
5	reported_date	date		reported_date
6	hospital_occupancy_count	1.2 double	*	Daily hospital occupancy_count
7	ICU_occupancy_count	1.2 double	*	Daily ICU occupancy_count
8	source	abc string		source

Data Factory   Validate all   Publish all (4)   Preview experience   Off

Factory Resources   Filter resources by name

Pipelines

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview)

Datasets

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_processed\_hospitaladmissions\_w...
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query

**Properties**

**General**   Related (1)

Name \* ds\_processed\_hospitaladmissions\_weekly

Description

Annotations

**Connection**   Schema   Parameters

Linked service \* ls\_azdl\_ecovidreportingdatalake   Test connection   Edit   New   Learn more

File path \* processed / ecdc/hospitaladmissions\_we... / File name   Browse   Preview

Compression type None

Column delimiter Comma (,)

Row delimiter Default (\r\n or \n\r)

Encoding Default(UTF-8)

Quote character Double quote ("")

Escape character Backslash (\)

First row as header

Null value

Data Factory   Validate all   Publish all 4

Factory Resources   Filter resources by name

Pipelines: pl\_cases\_deaths\_data\_processed, pl\_ingest\_ecdc\_data, pl\_ingest\_population\_data

Change Data Capture (preview): 0

Datasets: ds\_country\_lookup, ds\_dimdate\_lookup, ds\_ecdc\_data\_dl, ds\_ecdc\_filelist, ds\_http\_ecdc\_data, ds\_populationdata\_gz, ds\_populationdata\_tsv, ds\_processed\_cases\_deaths, ds\_processed\_hospitaladmissions\_w..., ds\_raw\_cases\_deaths

Data flows: df\_transform\_hospital\_admissions, df\_transform\_cases\_deaths

Power Query: 0

Preview experience: Off

Properties: General, Related

Name: df\_transform\_hospital\_admissions

Description:

Sink tab:

- Output stream name: Weeklysink
- Description: Export data to ds\_processed\_hospitaladmissions\_weekly
- Reset
- Incoming stream: selectWeekly
- Sink type: Dataset
- Dataset: ds\_processed\_hospitaladmissions\_w...
- Test connection, Open, New
- Skip line count:
- Options:
  - Allow schema drift (checked)
  - Validate schema

Diagram:

```
graph LR; JoinWithDate[JoinWithDate] --> pivotWeekly[pivotWeekly]; pivotWeekly --> sortWeekly[sortWeekly]; sortWeekly --> selectWeekly[selectWeekly]; selectWeekly --> Weeklysink[Weeklysink];  
pivotDaily[pivotDaily] --> sortDaily[sortDaily]; sortDaily --> selectDaily[selectDaily]; selectDaily --> Weeklysink;
```

The diagram shows a data flow pipeline. It starts with a 'JoinWithDate' component, which performs an inner join between 'SplitDailyfromWeekly@Weekly' and 'AggDimdate'. This is followed by a 'pivotWeekly' component, which pivots row values into columns, groups columns, and aggregates data. Next is a 'sortWeekly' component, which sorts rows on columns 'reported\_year\_week' and 'country'. Then comes a 'selectWeekly' component, which renames 'selectWeekly' to 'selectWeekly with columns country, countryv\_code\_2\_dinit'. Finally, the data is written to a 'Weeklysink' sink, which has 10 total columns.

Factory Resources < Filter resources by name +

Pipelines 3

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 11

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_processed\_hospitaladmissions\_w...
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows 2

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Power Query 0

Validate all Publish all 4

✓ Validate Data flow debug ●

Set properties

Name ds\_processed\_hospitaladmissions\_daily

Linked service \* ls\_azdl\_ecovidreportingdatalake

File path processed / ecdc/hospitaladmission... / File name

First row as header

Import schema

From connection/store  From sample file  None

> Advanced

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name \* Dailysink Learn more □

Description Add sink dataset Reset

Incoming stream \* selectDaily

Sink type \*

Dataset   +

Options  Allow schema drift ⓘ  Validate schema ⓘ

OK Back Cancel

Data Factory   Validate all   Publish all (5)

Factory Resources   Filter resources by name

Pipelines

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 12

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_processed\_hospitaladmissions\_da...
- ds\_processed\_hospitaladmissions\_w...
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows 2

- df\_transform\_hospital\_admissions
- df\_transform\_cases\_deaths

Validate   Data flow debug

Properties

General   Related

Name \* df\_transform\_hospital\_admissions

Description

Sink   Settings   Errors   Mapping   Optimize   Inspect   Data preview

Output stream name \* Dailysink   Learn more

Description Export data to ds\_processed\_hospitaladmissions\_daily   Reset

Incoming stream \* selectDaily

Sink type \* Dataset   Inline   Cache

Dataset \* ds\_processed\_hospitaladmissions\_da...   Test connection   Open   New

Skip line count

Options

Allow schema drift    Validate schema

Data Factory   Validate all   Publish all 1

Preview experience Off

Factory Resources

Filter resources by name

Pipelines (4)

- pl\_processed\_hospitaladmissions\_data
- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) (0)

Datasets (12)

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_ts
- ds\_processed\_cases\_deaths
- ds\_processed\_hospitaladmissions\_da...
- ds\_processed\_hospitaladmissions\_w...
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows (2)

- df\_transform\_hospitaladmissions

Activities

Validate   Debug   Add trigger   Data flow debug

Search activities

Move & transform

- Copy data
- Data flow

Synapse

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Power Query

Data flow

df\_transform\_hospitaladmissions

Properties

General   Related

Name \* pl\_processed\_hospitaladmissions\_data

Description

Annotations

+ New

General Settings Parameters User properties

Name \* df\_transform\_hospitaladmissions

Description

Activity state (preview)  Active  Inactive

Timeout

Retry

Retry interval (sec)

```
graph LR; subgraph Pipeline [pl_processed_hospitaladmissions_data]; direction TB; A[Data flow df_transform_hospitaladmissions]; end; subgraph Properties [Properties]; direction TB; G[General]; R[Related]; end; G --> PN[Name * pl_processed_hospitaladmissions_data]; G --> DS[Description]; G --> AS[Activity state (preview)]; AS --> A1((Active)); AS --> A2((Inactive)); G --> T[Timeout 0.12:00:00]; G --> R1[Retry 0]; G --> RI[Retry interval (sec) 30];
```

Preview experience  Off

Data Factory Validate all Publish all 1

Factory Resources

Filter resources by name  +

Pipelines 4

- pl\_processed\_hospitaladmissions\_data
- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 12

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_processed\_hospitaladmissions\_da...
- ds\_processed\_hospitaladmissions\_w...
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

Data flows 2

- df\_transform\_hospital\_admissions

Activities

Validate Debug Add trigger  Data flow debug

Search activities

Move & transform

- Copy data
- Data flow

Synapse

Azure Data Explorer

Azure Function

Batch Service

Databricks

Data Lake Analytics

General

HDInsight

Iteration & conditionals

Machine Learning

Power Query

Data flow df\_transform\_hospital\_admissions

Properties

General Related

Name \* pl\_processed\_hospitaladmissions\_data

Description

Annotations

General Settings Parameters User properties

Data flow \* df\_transform\_hospital\_admissions

Run on (Azure IR) \* AutoResolveIntegrationRuntime

Compute size \* Small

Advanced

Logging level \* Verbose Basic None

Sink properties

Staging

## Pipeline runs

Triggered Debug

Rerun

Cancel options

Refresh

Edit columns

List

Gantt

Filter by run ID or name

Local time : Last 24 hours

Pipeline name : All

Status : All

Runs : Latest runs

Triggered by : All

Add filter

Copy filters

Export to CSV

Showing 1 - 2 items

Last refreshed 0 minutes ago

<input type="checkbox"/> Pipeline name ↑↓	Run start ↑↓	Run end ↑↓	Duration	Triggered by	Status ↑↓	Run	Parameters	Annotations	RU
<input type="checkbox"/> pl_processed_hospitaladmissi...	6/30/2023, 11:16:53 PM	6/30/2023, 11:21:02 PM	4m 9s	Manual trigger	<span style="color: green;">✓ Succeeded</span>	Original			67
<input type="checkbox"/> pl_cases_deaths_data_process...	6/30/2023, 8:44:35 PM	6/30/2023, 8:48:53 PM	4m 18s	Manual trigger	<span style="color: green;">✓ Succeeded</span>	Original			e4

## Microsoft Azure Storage Explorer

File Edit View Help



EXPLORER

Search for resources

 processed ✖  
 raw ✖  
 \$logs ✖  
 population ✖  
 configs ✖  
 lookup ✖

 Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undelete Σ Folder Statistics Refresh

Active blobs (default) Filter by prefix (case-sensitive)

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID
_SUCCESS	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	0 B	Active	
part-00000-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.66 KiB	Active	
part-00001-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.60 KiB	Active	
part-00002-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.56 KiB	Active	
part-00003-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.58 KiB	Active	
part-00004-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.58 KiB	Active	
part-00005-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.58 KiB	Active	
part-00006-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.57 KiB	Active	
part-00007-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.58 KiB	Active	
part-00008-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.57 KiB	Active	
part-00009-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.58 KiB	Active	
part-00010-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.56 KiB	Active	
part-00011-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.57 KiB	Active	
part-00012-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.57 KiB	Active	
part-00013-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.58 KiB	Active	
part-00014-b6c2531e-a544-438b-9e94-546eda41728b-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	1.57 KiB	Active	

Showing 1 to 100 of 201 cached items

Activities

[Clear completed](#) [Clear successful](#)

✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\dim\_date.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
 Started at: 6/30/2023 9:35 PM, Duration: 4 seconds

[Copy AzCopy Command to Clipboard](#)

✓ Successfully created blob container 'processed'

✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\case\_deaths\_uk\_ind\_only.csv' to 'raw/ecdc/' complete: 1 item transferred (used SAS, discovery completed)  
 Started at: 6/30/2023 6:37 PM, Duration: 4 seconds

[Copy AzCopy Command to Clipboard](#)

✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\country\_lookup.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)

[Copy AzCopy Command to Clipboard](#)

## Microsoft Azure Storage Explorer

File Edit View Help



EXPLORER

Search for resources

Collapse all Refresh all

- Quick Access
- Emulator & Attached
- Storage Accounts
- Azure for Students (AvinashChinta@my.unt.edu)
  - Storage Accounts
    - cs71003200143e43ad6
    - dbstorageocb4jzhwww7hy (ADLS Gen2)
    - dlspractise (ADLS Gen2)
    - ercovidreportingdatalake (ADLS Gen2)
      - Blob Containers
        - \$logs
        - lookup
        - processed
        - raw
  - ercovidreportingsa
  - Disk
    - cloud-shell-storage-southcentralus
    - covid-reporting-rg
    - databricks-rg-dbs-hanson-bj47lln3b
    - NetworkWatcherRG
    - rg-avinash-hanson

processed raw \$logs population configs lookup

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undelete Σ Folder Statistics Refresh

Active blobs (default) processed > ecdc > hospitaladmissions\_weekly

Filter by prefix (case-sensitive)

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID
_SUCCESS	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	0 B	Active	
part-00000-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	613 B	Active	
part-00001-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	573 B	Active	
part-00002-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	508 B	Active	
part-00003-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	650 B	Active	
part-00004-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	455 B	Active	
part-00005-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	587 B	Active	
part-00006-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	484 B	Active	
part-00007-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	626 B	Active	
part-00008-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	493 B	Active	
part-00009-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	613 B	Active	
part-00010-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	563 B	Active	
part-00011-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	612 B	Active	
part-00012-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	483 B	Active	
part-00013-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	598 B	Active	
part-00014-c6c11802-c8dd-43e8-bc82-c167a0a716ec-c000.csv	Hot (inferred)		6/30/2023 11:20 PM	Block Blob	application/octet-stream	522 B	Active	

Showing 1 to 100 of 201 cached items

&lt;&lt; &lt; 1 2 3 &gt; &gt;&gt;

Activities

Clear completed Clear successful

✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\dim\_date.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 9:35 PM, Duration: 4 seconds

[Copy AzCopy Command to Clipboard](#)

✓ Successfully created blob container 'processed'

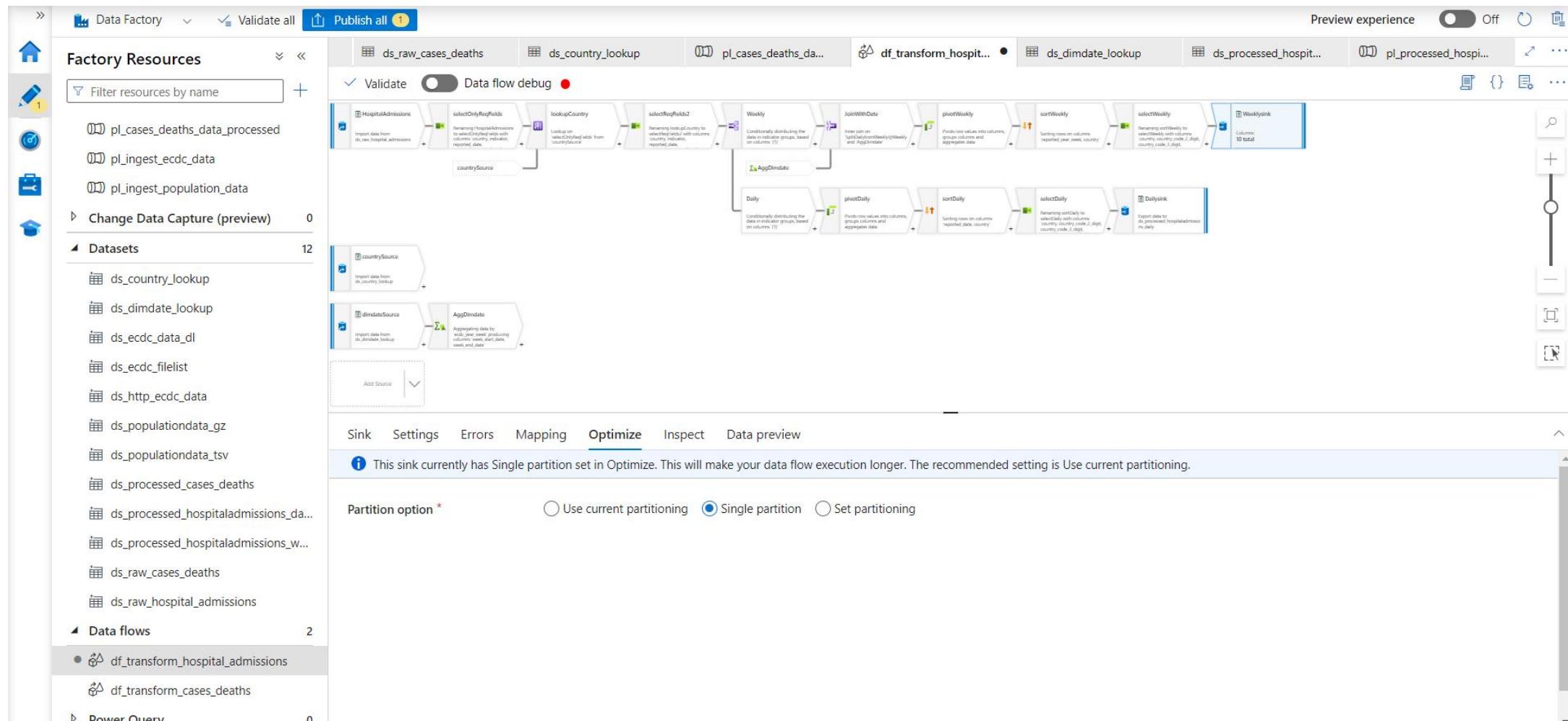
✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\case\_deaths\_uk\_ind\_only.csv' to 'raw/ecdc/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 4 seconds

[Copy AzCopy Command to Clipboard](#)

✓ Transfer of 'C:\Users\Dheeraji Bajjuri\Downloads\country\_lookup.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 5 seconds

[Copy AzCopy Command to Clipboard](#)

Actions Properties



Factory Resources < >

Filter resources by name +

- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data
- Change Data Capture (preview) 0
- Datasets 12
  - ds\_country\_lookup
  - ds\_dimdate\_lookup
  - ds\_ecdc\_data\_dl
  - ds\_ecdc\_filelist
  - ds\_http\_ecdc\_data
  - ds\_populationdata\_gz
  - ds\_populationdata\_tsv
  - ds\_processed\_cases\_deaths
  - ds\_processed\_hospitaladmissions\_da...
  - ds\_processed\_hospitaladmissions\_w...
  - ds\_raw\_cases\_deaths
  - ds\_raw\_hospital\_admissions
- Data flows 2
  - df\_transform\_hospital\_admissions
  - df\_transform\_cases\_deaths
- Power Query 0

Publish all 1 | Validate all | Preview experience Off

Validate Data flow debug ●

Sink Settings Errors Mapping Optimize Inspect Data preview

This sink currently has Single partition set in Optimize. This will make your data flow execution longer. The recommended setting is Use current partitioning.

Clear the folder

File name option \*

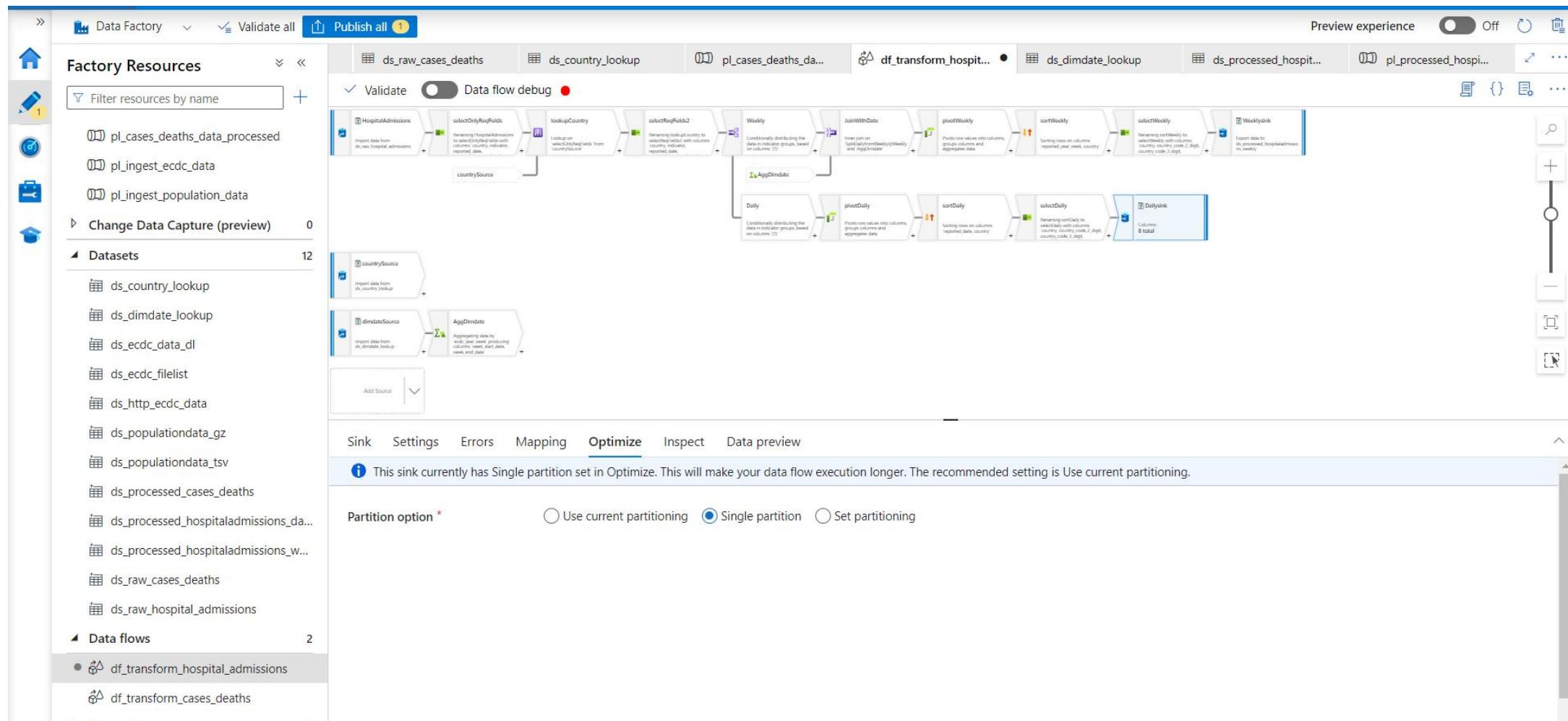
Output to single file \*

Quote All

Headers

Umask
 

Owner	<input type="checkbox"/> R	<input type="checkbox"/> W	<input type="checkbox"/> X
Group	<input type="checkbox"/> R	<input checked="" type="checkbox"/> W	<input type="checkbox"/> X
Others	<input type="checkbox"/> R	<input checked="" type="checkbox"/> W	<input type="checkbox"/> X



Factory Resources < Filter resources by name +

✓ Validate Data flow debug

Sink **Settings** Errors Mapping Optimize Inspect Data preview

**Settings**

This sink currently has Single partition set in Optimize. This will make your data flow execution longer. The recommended setting is Use current partitioning.

**Clear the folder**

**File name option \*** Output to single file

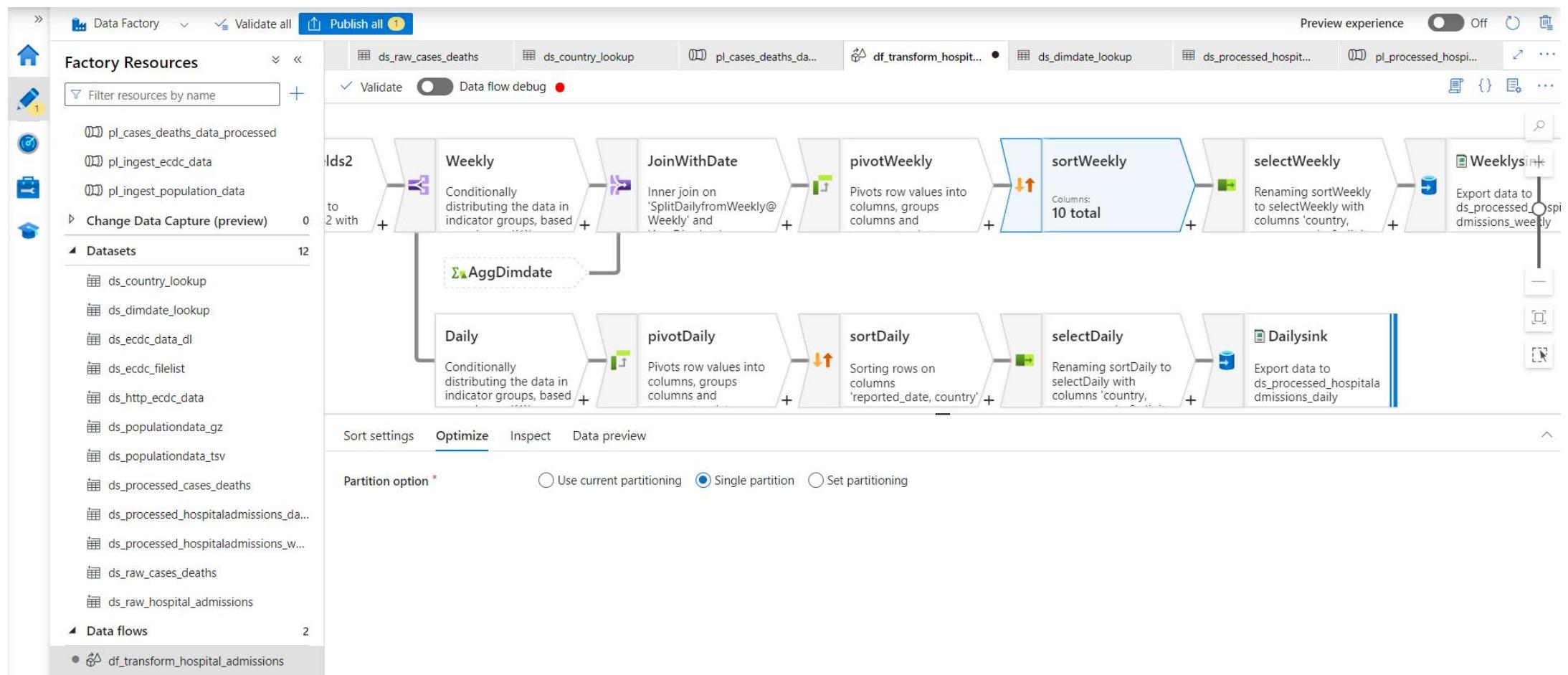
**Output to single file \*** hospitaladmissions\_daily.csv

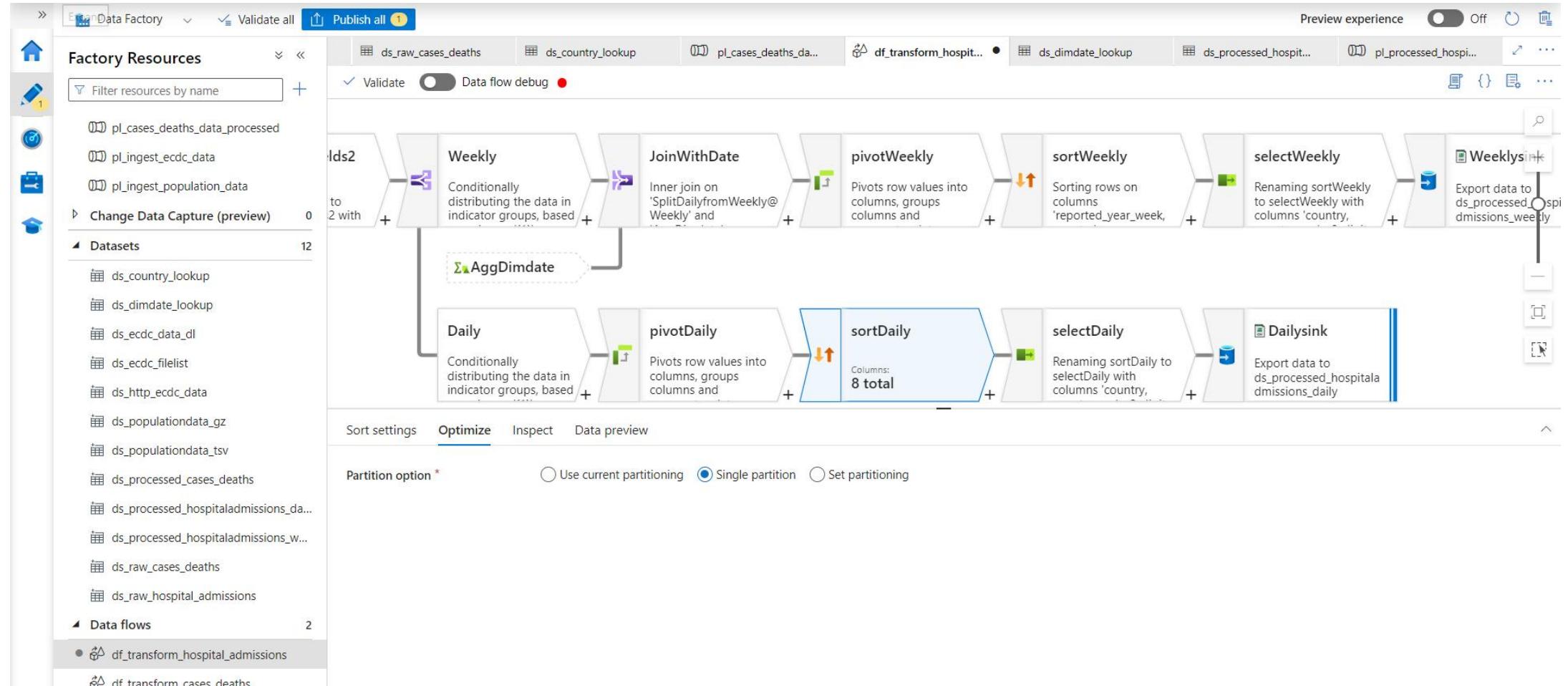
**Quote All**

**Headers** Enter expression... ANY

**Umask**

Owner	<input type="checkbox"/> R	<input type="checkbox"/> W	<input type="checkbox"/> X
Group	<input type="checkbox"/> R	<input checked="" type="checkbox"/> W	<input type="checkbox"/> X
Others	<input type="checkbox"/> R	<input checked="" type="checkbox"/> W	<input type="checkbox"/> X





» Expand

Dashboards

Runs

Pipeline runs

Trigger runs

Change Data Capture (previ...)

Runtimes & sessions

Integration runtimes

Data flow debug

Notifications

Alerts & metrics

## Pipeline runs

Triggered Debug Rerun Cancel options Refresh Edit columns List Gantt

Filter by run ID or name Local time : Last 24 hours Pipeline name : All Status : All Runs : Latest runs Triggered by : All Add filter Copy filters Export to CSV

Showing 1 - 3 items Last refreshed 0 minutes ago

Pipeline name ↑	Run start ↓	Run end ↓	Duration	Triggered by	Status ↑	Run	Parameters	Annotations	Run
pl_processed_hospitaladmissi...	6/30/2023, 11:31:50 PM	6/30/2023, 11:36:07 PM	4m 17s	Manual trigger	Succeeded	Original			ab
pl_processed_hospitaladmissi...	6/30/2023, 11:16:53 PM	6/30/2023, 11:21:02 PM	4m 9s	Manual trigger	Succeeded	Original			67
pl_cases_deaths_data_process...	6/30/2023, 8:44:35 PM	6/30/2023, 8:48:53 PM	4m 18s	Manual trigger	Succeeded	Original			e4

Microsoft Azure Storage Explorer

File Edit View Help

EXPLORER    processed raw \$logs population configs lookup

Upload Download Open Preview New Folder Select All Copy Paste Clone Rename Move Manage ACLs Properties Delete Undo Folder Statistics Refresh

Collapse all Refresh all

Quick Access

Emulator & Attached

Storage Accounts

Azure for Students (AvinashChinta@my.unt.edu)

- Storage Accounts
  - cs71003200143e43ad6
  - dbstoragegeocb4jzhwww7hy (ADLS Gen2)
  - dlspractise (ADLS Gen2)
  - ercovidreportingdatalake (ADLS Gen2)
    - Blob Containers
      - \$logs
      - lookup
      - processed
      - raw
    - File Shares
    - Queues
    - Tables
  - ercovidreportingsa
- Disks
  - cloud-shell-storage-southcentralus
  - covid-reporting-rg
  - databricks-rg-dbs-hanson-bj47lln3b
  - NetworkWatcherRG
  - rg-avinash-hanson

Active blobs (default) processed > ecdc > hospitaladmissions\_daily

Filter by prefix (case-sensitive)

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted
hospitaladmissions_daily.csv	Hot (inferred)		6/30/2023 11:35 PM	Block Blob	application/octet-stream	285.18 KiB	Active			

Showing 1 to 1 of 1 cached items

Activities

Clear completed Clear successful

Transfer of 'C:\Users\AvinashChinta\Downloads\dim\_date.csv' to 'processed/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 9:35 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)

Successfully created blob container 'processed'

Transfer of 'C:\Users\AvinashChinta\Downloads\case\_deaths\_uk\_ind\_only.csv' to 'raw/ecdc/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)

Transfer of 'C:\Users\AvinashChinta\Downloads\country\_lookup.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 4 seconds [Copy AzCopy Command to Clipboard](#)

Actions Properties

## Microsoft Azure Storage Explorer

File Edit View Help



EXPLORER

Search for resources

processed ✖raw ✖\$logs ✖population ✖configs ✖lookup ✖↑ ↓ ↶ ↷ +/- ☒ 📄 📁 📅 🕒 👤 🔗 ✖ ↶ ↷ Σ ⟳

Collapse all Refresh all

Quick Access

Emulator &amp; Attached

Storage Accounts

Azure for Students (AvinashChinta@my.unt.edu)

Storage Accounts

cs71003200143e43ad6

dbstoragegeocb4jzhwww7hy (ADLS Gen2)

dlspractise (ADLS Gen2)

ercovidreportingdatalake (ADLS Gen2)

Blob Containers

\$logs

lookup

processed

raw

File Shares

Queues

Tables

ercovidreportingsa

Disks

cloud-shell-storage-southcentralus

covid-reporting-rg

databricks-rg-dbs-hands-on-bj47lin3b

NetworkWatcherRG

rg-avinash-hanson

← → ↕ Active blobs (default) processed > ecdc > hospitaladmissions\_weekly Filter by prefix (case-sensitive) ✖

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted
hospitaladmissions_weekly.csv	Hot (inferred)		6/30/2023 11:36 PM	Block Blob	application/octet-stream	70.68 KiB	Active			

Showing 1 to 1 of 1 cached items

Activities

Clear completed Clear successful

✓ Transfer of 'C:\Users\AvinashChinta\Downloads\dim\_date.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 9:35 PM, Duration: 4 seconds

[Copy AzCopy Command to Clipboard](#)

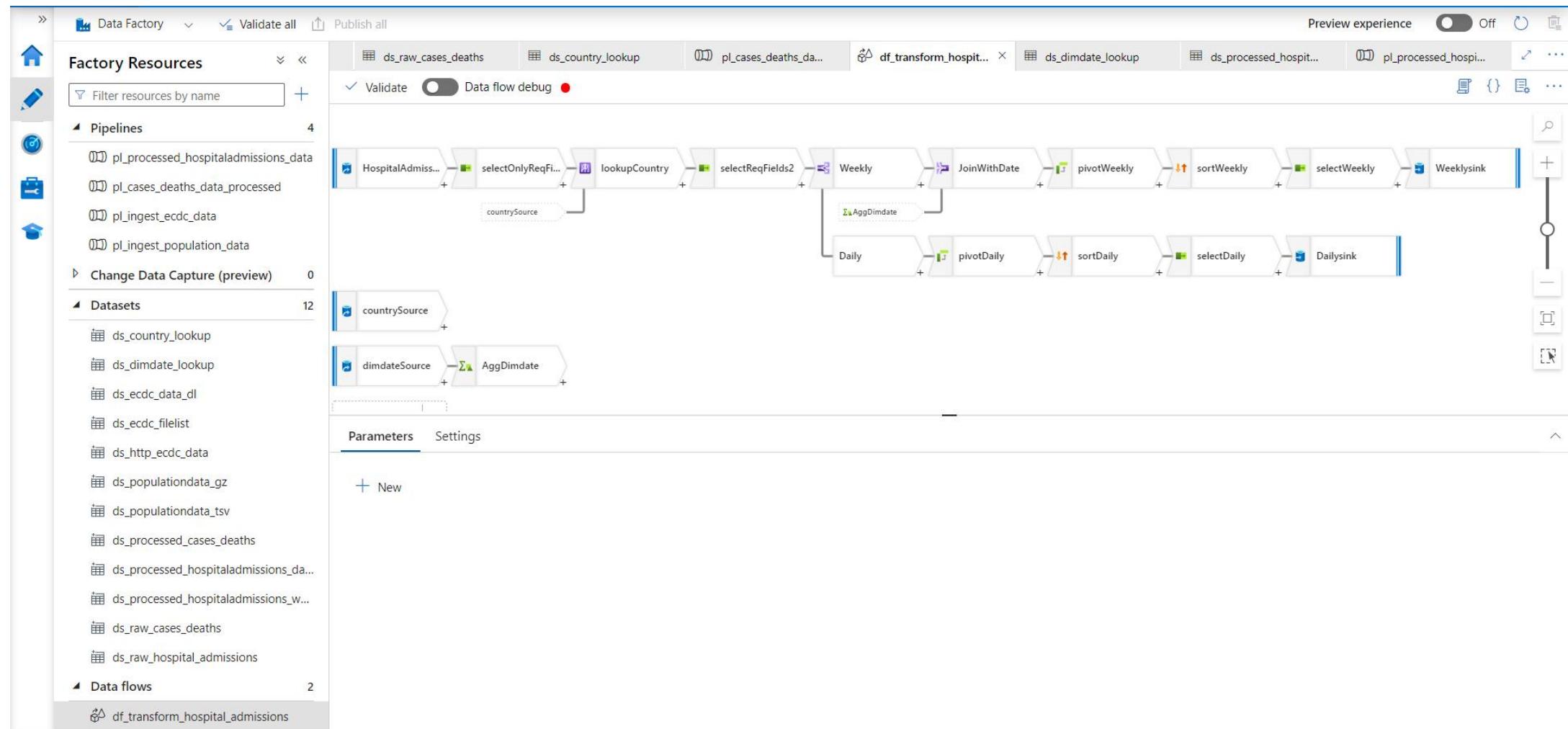
✓ Successfully created blob container 'processed'

✓ Transfer of 'C:\Users\AvinashChinta\Downloads\case\_deaths\_uk\_ind\_only.csv' to 'raw/ecdc/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 6:37 PM, Duration: 4 seconds

[Copy AzCopy Command to Clipboard](#)

✓ Transfer of 'C:\Users\AvinashChinta\Downloads\countryv\_lookup.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)

[Copy AzCopy Command to Clipboard](#)



## Microsoft Azure Storage Explorer

File Edit View Help



EXPLORER

Search for resources

processed raw \$logs population configs lookup

Upload Download Open Preview New Folder Select All Copy Paste Clone Delete Undelete Manage History Folder Statistics Refresh

Collapse all

Refresh all

- Quick Access
- Emulator & Attached
- Storage Accounts
  - Azure for Students (AvinashChinta@my.unt.edu)
    - Storage Accounts
      - cs71003200143e43ad6
      - dbstoragegeocb4jzhwww7hy (ADLS Gen 1)
      - dlspractise (ADLS Gen 2)
      - ercovidreportingdatalake (ADLS Gen 2)
        - Blob Containers
          - \$logs
          - lookup
          - processed
          - raw
        - File Shares
        - Queues
        - Tables
    - ercovidreportingsa
      - Blob Containers
        - \$logs
        - configs
        - population
      - File Shares
      - Queues
      - Tables
  - Disks
    - cloud-shell-storage-southcentralus
    - covid-reporting-rg
    - databricks-rg-dbs-hanson-bj47lln3b
    - NetworkWatcherRG
    - rg-avinash-hanson

← → ↑ Active blobs (default)

configs

Name	Access Tier	Access Tier
ecdc_file_list_new.json	Hot (inferred)	
githubusercontent_ecdc_file_list.json	Hot (inferred)	

ecdc\_file\_list (1).json - Notepad

File Edit Format View Help

```
[  
 {  
 "sourceBaseUrl": "https://github.com",  
 "sourceRelativeURL": "cloubboxacademy/covid19/raw/main/ecdc_dat  
 "sinkFileName": "cases_deaths/cases_deaths.csv"  
 },  
 {  
 "sourceBaseUrl": "https://github.com",  
 "sourceRelativeURL": "cloubboxacademy/covid19/raw/main/ecdc_dat  
 "sinkFileName": "hospital_admissions/hospital_admissions.csv"  
 },  
 {  
 "sourceBaseUrl": "https://github.com",  
 "sourceRelativeURL": "cloubboxacademy/covid19/raw/main/ecdc_dat  
 "sinkFileName": "testing/testing.csv"  
 },  
 {  
 "sourceBaseUrl": "https://github.com",  
 "sourceRelativeURL": "cloubboxacademy/covid19/raw/main/ecdc_dat  
 "sinkFileName": "country_response/country_response.csv"  
 }]
```

Showing 1 to 2 of 2 cached items

Activities

Clear completed Clear successful

- Transfer of 'C:\Users\Deheraji Bajuri\Downloads\ecdc\_file\_list\_new.json' to 'configs/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 7/1/2023 10:49 AM, Duration: 4 seconds
- Deletion of 'ecdc\_file\_list.json' from 'configs/' completed: 1 completed (used SAS, discovery completed)  
Started at: 7/1/2023 10:49 AM, Duration: 6 seconds
- Transfer of 'C:\Users\Deheraji Bajuri\Downloads\dim\_date.csv' to 'lookup/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 6/30/2023 9:35 PM, Duration: 4 seconds
- Successfully created blob container 'processed'

Copy Command to Clipboard

Copy AzCopy Command to Clipboard

Copy AzCopy Command to Clipboard

Actions Properties

» Data Factory Data Factory Validate all Publish all 1 Preview experience Off

Factory Resources << Filter resources by name +

Pipelines 4

- pl\_processed\_hospitaladmissions\_data
- pl\_cases\_deaths\_data\_processed
- pl\_ingest\_ecdc\_data
- pl\_ingest\_population\_data

Change Data Capture (preview) 0

Datasets 12

- ds\_country\_lookup
- ds\_dimdate\_lookup
- ds\_ecdc\_data\_dl
- ds\_ecdc\_filelist
- ds\_http\_ecdc\_data
- ds\_populationdata\_gz
- ds\_populationdata\_tsv
- ds\_processed\_cases\_deaths
- ds\_processed\_hospitaladmissions\_da...
- ds\_processed\_hospitaladmissions\_w...
- ds\_raw\_cases\_deaths
- ds\_raw\_hospital\_admissions

**ds\_ecdc\_filelist** JSON

Connection Schema Parameters

Linked service \* ls\_azblob\_ercovidreportingsa Test connection Edit New Learn more

File path \* configs / Directory / ecdc\_file\_list\_new.json Browse Preview data

Compression type None

Encoding Default(UTF-8)

The screenshot shows the Azure Data Factory interface with the 'ds\_ecdc\_filelist' dataset selected. The dataset is defined as a JSON file located in a blob storage linked service. The configuration includes the linked service, file path, compression type (None), and encoding (Default(UTF-8)). The interface also displays other datasets and pipelines in the factory.

## Pipeline runs

Triggered Debug

Rerun

Cancel options

Refresh

Edit columns

List

Gantt

Succeeded

Successfully ran pl\_ingest\_ecdc\_data (Pipeline).

[View pipeline run](#)

[Copy filters](#) [Export to CSV](#)

Filter by run ID or name

Last 24 hours

Pipeline name : All

Status : All

Runs : Latest runs

Triggered by : All

Ac

Showing 1 - 6 items

Last refreshed 0 minutes ago

	Pipeline name ↑↓	Run start ↑↓	Run end ↑↓	Duration	Triggered by	Status ↑↓	Run	Parameters	Annotations	Run
	<a href="#">pl_ingest_ecdc_data</a>	7/1/2023, 11:10:50 AM	7/1/2023, 11:11:04 AM	15s	Manual trigger	<span>✓ Succeeded</span>	Original			48
	<a href="#">pl_ingest_ecdc_data</a>	7/1/2023, 11:07:22 AM	7/1/2023, 11:07:26 AM	5s	Manual trigger	<span>✗ Failed</span>	Original			00
	<a href="#">pl_ingest_ecdc_data</a>	7/1/2023, 11:04:03 AM	7/1/2023, 11:04:09 AM	6s	Manual trigger	<span>✗ Failed</span>	Original			c1
	<a href="#">pl_processed_hospitaladmissi...</a>	6/30/2023, 11:31:50 PM	6/30/2023, 11:36:07 PM	4m 17s	Manual trigger	<span>✓ Succeeded</span>	Original			ab
	<a href="#">pl_processed_hospitaladmissi...</a>	6/30/2023, 11:16:53 PM	6/30/2023, 11:21:02 PM	4m 9s	Manual trigger	<span>✓ Succeeded</span>	Original			67
	<a href="#">pl_cases_deaths_data_process...</a>	6/30/2023, 8:44:35 PM	6/30/2023, 8:48:53 PM	4m 18s	Manual trigger	<span>✓ Succeeded</span>	Original			e4

## Microsoft Azure Storage Explorer

File Edit View Help



EXPLORER

Search for resources

Collapse all

Refresh all

Quick Access

Emulator &amp; Attached

Storage Accounts

Azure for Students (AvinashChinta@my.unt.edu)

Storage Accounts

cs71003200143e43ad6

dbstoragegeocb4jzhwww7hy (ADLS Gen2)

dlspractise (ADLS Gen2)

ercovidreportingdatalake (ADLS Gen2)

Blob Containers

\$logs

lookup

processed

raw

File Shares

Queues

Tables

ercovidreportingsa

Blob Containers

\$logs

configs

population

File Shares

Queues

Tables

Disks

cloud-shell-storage-southcentralus

covid-reporting-rg

databricks-rg-dbs-hanson-bj47lln3b

NetworkWatcherRG

rg-avinash-hanson

[processed](#) [raw](#) [\\$logs](#) [population](#) [configs](#) [lookup](#)[Upload](#) [Download](#) [Open](#) [Preview](#) [New Folder](#) [Select All](#) [Copy](#) [Paste](#) [Clone](#) [Rename](#) [Move](#) [Manage ACLs](#) [Properties](#) [Delete](#) [Undelete](#) [Folder Statistics](#) [Refresh](#)[Active blobs \(default\)](#) [raw > ecdc](#) Filter by prefix (case-sensitive)

Name	Access Tier	Access Tier Last Modified	Last Modified	Blob Type	Content Type	Size	Status	Deletion ID	Remaining Days	Deleted T
cases_deaths			7/1/2023 11:11 AM		Folder		Active			
country_response			7/1/2023 11:11 AM		Folder		Active			
hospital_admissions			7/1/2023 11:11 AM		Folder		Active			
testing			7/1/2023 11:11 AM		Folder		Active			
case_deaths_uk_ind_only.csv	Hot (inferred)		6/30/2023 6:37 PM	Block Blob	application/vnd.ms-excel	131.48 KiB	Active			
cases_deaths.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	13.78 MiB	Active			
country_response.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	46.20 KiB	Active			
hospital_admissions.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	1.01 MiB	Active			
testing.csv	Hot (inferred)		6/30/2023 4:53 PM	Block Blob	application/octet-stream	83.87 KiB	Active			

Showing 1 to 9 of 9 cached items

Activities

[Clear completed](#) [Clear successful](#)

✓ Transfer of 'C:\Users\AvinashChinta\Downloads\ecdc\_file\_list.json' to 'configs/' complete: 1 item transferred (used SAS, discovery completed)  
Started at: 7/1/2023 11:09 AM, Duration: 4 seconds

[Copy AzCopy Command to Clipboard](#)

✓ Deletion of 'ecdc\_file\_list\_new.json' from 'configs/' completed: 1 completed (used SAS, discovery completed)  
Started at: 7/1/2023 11:09 AM, Duration: 5 seconds

[Copy AzCopy Command to Clipboard](#)

● Transfer of 'C:\Users\AvinashChinta\Downloads\ecdc\_file\_list.json' to 'configs/' in progress: 1 item transferred (used SAS, discovery completed)

[Copy AzCopy Command to Clipboard](#)