

# Day8\_HarshThakur

Friday, 8 September 2023

22:19

The screenshot displays the Microsoft Azure Data Factory portal interface. The left sidebar shows the 'Factory Resources' tree with 'Pipelines' expanded, listing IDA\_PL\_01, IDA\_PL\_02, and IDA\_PL\_03. The 'Activities' pane on the right shows a 'Copy data' activity. The main canvas displays a 'Copy data' activity with a green checkmark, indicating success. Below the canvas, the 'Output' tab shows the 'Pipeline run ID' as 2db0b939-af34-46a1-85f8-2194d052d0ec and the 'Pipeline status' as 'Succeeded'. A table below lists the activity details:

Activity name	Activity status	Run start	Duration	Integration runtime	User properties	Activity run ID
Copy data1	Succeeded	9/7/2023, 3:21:03 PM	9s	AutoResolveIntegration		b75ad1c7-a971-4204-b411-dc14113c

The screenshot displays the Microsoft Azure Data Factory portal interface. The left sidebar shows the 'Factory Resources' tree with 'Pipelines' expanded, listing IDA\_PL\_01 and IDA\_PL\_02. The 'Activities' pane on the right shows a 'Copy data' activity. The main canvas displays a 'Copy data' activity with a green checkmark, indicating success. Below the canvas, the 'Publishing' tab shows the 'Pipeline run ID' as 2db0b939-af34-46a1-85f8-2194d052d0ec and the 'Pipeline status' as 'Publishing'. A table below lists the activity details:

Activity name	Activity status	Run start	Duration	Integration runtime	User properties	Activity run ID
Copy data1	Publishing	9/7/2023, 3:21:03 PM	9s	AutoResolveIntegration		b75ad1c7-a971-4204-b411-dc14113c



part-00000-d0b3c8ea-c692-454e-ad6b-6631bde578e6-c000.csv

part-00000-d0b3c8ea-c692-454e-ad6b-6631bde578e6-c000.csv

Overview Versions Snapshots Edit Generate SAS

Properties

URL <https://harshisblob...>

LAST MODIFIED 9/7/2023, 12:12:41 PM

CREATION TIME 9/7/2023, 12:12:41 PM

VERSION ID -

TYPE Block blob

SIZE 12.74 KiB

ACCESS TIER Hot (Inferred)

ACCESS TIER LAST MODIFIED N/A

ARCHIVE STATUS -

REHYDRATE PRIORITY -

SERVER ENCRYPTED true

ETAG 0x0B8AF6DA29137D1

VERSION-LEVEL IMMUTABILITY POLICY Disabled

CACHE-CONTROL

CONTENT-TYPE application/octet-stream

CONTENT-MD5

CONTENT-ENCODING

CONTENT-LANGUAGE

CONTENT-DISPOSITION

LEASE STATUS Unlocked

LEASE STATE Available

LEASE DURATION -

COPY STATUS Success

COPY COMPLETION TIME 9/7/2023, 12:12:41 PM

COPY SOURCE <https://harshisblob...>

Undo

Metadata

Key	Value
hd_permission	("owner":"trusted-service-user","group":"supergroup...")

output - Microsoft Azure

output

Overview Versions Snapshots Edit Generate SAS

Authentication method: Access key (Switch to Azure AD User Account)

Location: output

Search blobs by prefix (case-sensitive)

Show deleted blobs

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
<input type="checkbox"/> _SUCCESS	9/7/2023, 12:12:41 PM	Hot (Inferred)		Block blob	0 B	Available
<input type="checkbox"/> part-00000-618741fb-b17b-40e3-b764-77b29bc60b7f-c000.csv	9/7/2023, 11:27:20 AM	Hot (Inferred)		Block blob	115.47 KiB	Available
<input type="checkbox"/> part-00000-d0b3c8ea-c692-454e-ad6b-6631bde578e6-c000.csv	9/7/2023, 12:12:41 PM	Hot (Inferred)		Block blob	12.74 KiB	Available



output - Microsoft Azure

harshisdatafactory - Azure Data Factory

adf.azure.com/authoring/dataflow/dataflow?factory=...&blob=...

Microsoft Azure | Data Factory | harshiddatafactory

Search factory and documentation

Preview experience Off

Factory Resources

- Pipelines: 1
  - IDA\_PL\_01
- Datasets: 3
  - ds\_destination
  - ds\_orderdetails
  - ds\_orders
- Data flows: 1
  - dataflow1
- Power Query: 0

dataflow1 x IDA\_PL\_01

Validate Data flow debug Debug Settings

Visual Studio Code interface showing a data flow pipeline:

- orders (Import data from ds\_orders)
- orderdetails (Import data from ds\_orderdetails)
- join1 (Right outer join on 'orders' and 'orderdetails')
- select1 (Renaming join1 to select1 with columns: Order ID, Order Date, CustomerName, State, City, Amount, Profit...)
- derived1 (Creating/Updating the columns: Order ID, Order Date, CustomerName, State, City, Amount, Profit...)
- filter1 (Filtering rows using expressions on columns: 'Profit' between [integer(Profit), 100, 500])
- sink1 (Export data to ds\_destination)

Filter settings

Output stream name: filter1

Description: Filtering rows using expressions on columns: 'Profit'

Incoming stream: derived1

Filter on: between [integer(Profit), 100, 500]

harshidblob - Microsoft Azure | harshiddatafactory - Azure Di... | Lumen | SalesDataset - Google Drive

adfs.azure.com/en/authoring/dataflow/dataflow1?factory=%2Fsubscriptions%2F66fb922-d220-4e51-bf86-0cccc0bbe535%2FresourceGroups%2Fharshidresource%2Fproviders%2FMicrosoft.DataFactory%2Ffactories%2Fharshiddatafactory

Microsoft Azure | Data Factory | harshiddatafactory

Search factory and documentation

Preview experience Off

Factory validation output

Factory Resources

- Pipelines: 1
  - IDA\_PL\_01
- Datasets: 3
  - ds\_destination
  - ds\_orderdetails
  - ds\_orders
- Data flows: 1
  - dataflow1
- Power Query: 0

dataflow1 x IDA\_PL\_01

Validate Data flow debug Debug Settings

Visual Studio Code interface showing a data flow pipeline:

- orders (Import data from ds\_orders)
- orderdetails (Import data from ds\_orderdetails)
- join1 (Right outer join on 'orders' and 'orderdetails')
- select1 (Renaming join1 to select1 with columns: Order ID, Order Date, CustomerName, State, City, Amount, Profit...)
- sink1 (Export data to ds\_destination)

Sink Settings

Output stream name: sink1

Description: Export data to ds\_destination

Incoming stream: select1

Sink type: Dataset

Dataset: ds\_destination

Skip line count:

Options: ☒ Allow schema drift ☐ Validate schema

Your factory has been validated. No errors were found.

Close

harshidblob - Microsoft Azure | harshiddatafactory - Azure Di... | Lumen | SalesDataset - Google Drive

adfs.azure.com/en/authoring/pipeline/IDA\_PL\_01?factory=%2Fsubscriptions%2F66fb922-d220-4e51-bf86-0cccc0bbe535%2FresourceGroups%2Fharshidresource%2Fproviders%2FMicrosoft.DataFactory%2Ffactories%2Fharshiddatafactory

Microsoft Azure | Data Factory | harshiddatafactory

Search factory and documentation

Preview experience Off

Factory validation output

Factory Resources

- Pipelines: 1
  - IDA\_PL\_01
- Datasets: 3
  - ds\_destination
  - ds\_orderdetails
  - ds\_orders

Activities

- Move and transform
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service

dataflow1 x IDA\_PL\_01

Validate Cancel options Add trigger Data flow debug

Visual Studio Code interface showing a data flow pipeline:

- dataflow1 (Data flow)

ds\_orders

▲ Data flows

1

dataflow1

0

Power Query

0

- > Databricks
- > Data Lake Analytics
- > General
- > HDInsight
- > Iteration & conditionals
- > Machine Learning
- > Power Query



Your factory has been validated.  
No errors were found.

Parameters Variables Settings **Output**

Pipeline run ID: 274d1e5b-fa8f-46e4-a29d-dbe191f72e5d

🔄

🔍

Pipeline status 🔵 In progress

[Monitor in Azure Metrics](#) [Export to CSV](#)

All status ▾

Showing 1 - 1 of 1 items

Activity name	Activity status	Run start	Duration	Integration runtime	User properties	Activity run ID
dataflow	<span>🟢</span> Succeeded	9/7/2023, 11:27:04 AM	1m 1s	debugpool-8Cores-Gei		0dc32276-2915-4

Close