

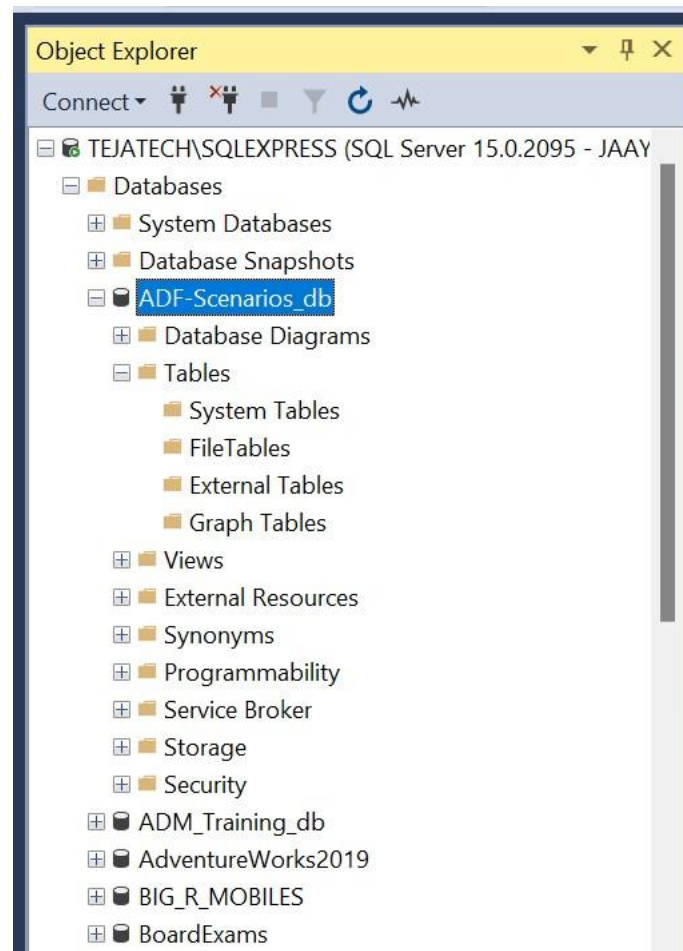
# ADF-Copy Data Scenario

## Problem statement:

Copy data from Azure data lake storage gen 2 to on premise MS SQL Server 2019.

## Solution:

SSMS- MS SQL Data base: SQL server authentication



## ADF Copy Data Pipeline - General

The screenshot shows the Microsoft Azure Data Factory console for a pipeline named 'Csv2SQLTrasfer'. The 'Activities' pane on the left lists 'Move & transform' as the selected category, with 'Copy data' and 'Data flow' as sub-options. The main pane displays the 'Copy data' activity 'CopyCSV2SQL' in the 'General' tab. The activity is configured with the following properties:

- Name:** CopyCSV2SQL
- Description:** (Empty text box)
- Timeout:** 0.12:00:00
- Retry:** 0
- Retry interval (sec):** 30

Navigation tabs at the bottom include 'General', 'Source', 'Sink', 'Mapping', 'Settings', and 'User properties'.

## ADF Copy Data Pipeline -Source

The screenshot shows the same Microsoft Azure Data Factory console, but with the 'Source' tab selected for the 'CopyCSV2SQL' activity. The configuration is as follows:

- Source dataset:** LogFileCSV
- File path type:** ☒ File path in dataset, ☐ Wildcard file path, ☐ List of files
- Filter by last modified:**  Start time (UTC),  End time (UTC)

Buttons for 'Open', 'New', 'Preview data', and 'Learn more' are visible below the source dataset selection.

## ADF Copy Data Pipeline -Preview data

Preview data

Linked service: ADLS\_ADF\_Scenario

Object: 15092022\_log\_csv.csv

	Prop_0	Prop_1	Prop_2	Prop_3	Prop_4
1	EMPLOYEE ID	FIRST NAME	LAST NAME	OPERATION NAME	TIME STAMP
2	1	Rosy	Ava	List Storage Account Keys	26-06-2021 14:24
3	2	Mary	Amelia	Delete SqlPools	25-06-2021 14:22
4	3	Jack	Haland	Pause SQL Analytics pools.	25-06-2021 14:22
5	4	Raees	khan	Pause SQL Analytics pools.	23-06-2021 07:23
6	5	Ruby	Jasline	List Storage Account Keys	20-06-2021 07:23
7	6	Romeo	Lukaku	Pause a Datawarehouse database.	19-06-2021 07:23
8	7	Mathew	Hayden	Create Pipeline Run	17-06-2021 07:23
9	8	Mark	Alexander	Create Pipeline Run	15-06-2021 14:22

## ADF Copy Data Pipeline -Sink

Microsoft Azure | ADF-Scenarios-Instance

Copy data

CopyCSV2SQL

Sink dataset \*

Write behavior ☒ Insert ☐ Upsert ☐ Stored procedure

Bulk insert table lock ☐ Yes ☒ No

Table option ☐ None ☒ Auto create table

Pre-copy script

## ADF Copy Data Pipeline -Mapping

Copy data

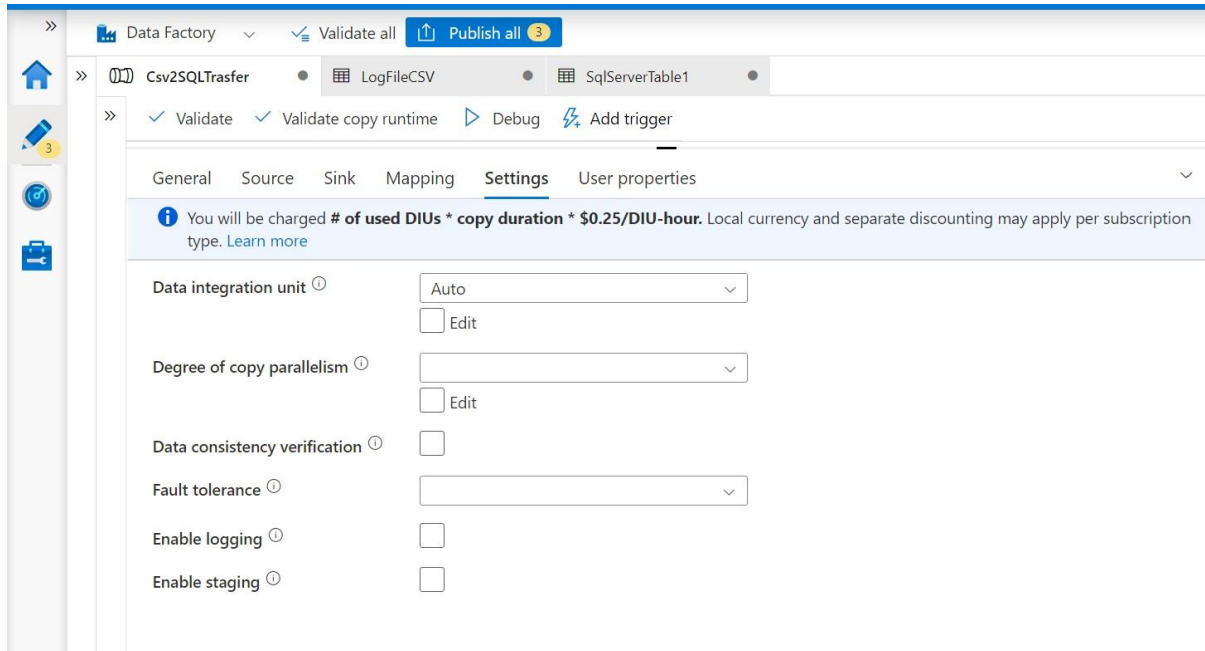
CopyCSV2SQL

Mapping

Source (Column number)	Type	Destination	Type
<input type="checkbox"/> 1	abc String	EmployeeID	String
<input type="checkbox"/> 2	abc String	FIRSTNAME	String
<input type="checkbox"/> 4	abc String	OPERATION NAME	String
<input checked="" type="checkbox"/> 5	abc String	TimeStamp	String

Add dynamic content [Alt+Shift+D]

## ADF Copy Data Pipeline -Settings



The screenshot shows the 'Settings' tab of an ADF Copy Data Pipeline. The pipeline is named 'Csv2SQLTrasfer' and is currently in the 'Settings' tab. The 'Data integration unit' is set to 'Auto'. The 'Degree of copy parallelism' is set to '1'. The 'Data consistency verification' checkbox is unchecked. The 'Fault tolerance' is set to 'Continue'. The 'Enable logging' checkbox is unchecked. The 'Enable staging' checkbox is unchecked. A warning message at the top states: 'You will be charged # of used DIUs \* copy duration \* \$0.25/DIU-hour. Local currency and separate discounting may apply per subscription type. Learn more'.

General Source Sink Mapping **Settings** User properties

**Settings**

**Data integration unit**

**Degree of copy parallelism**

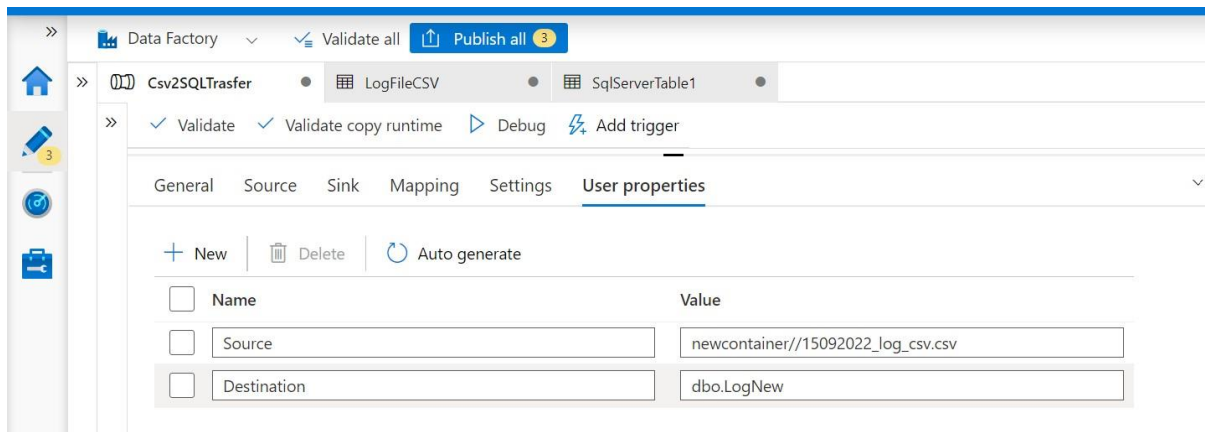
**Data consistency verification** ☐

**Fault tolerance**

**Enable logging** ☐

**Enable staging** ☐

## ADF Copy Data Pipeline-User Properties



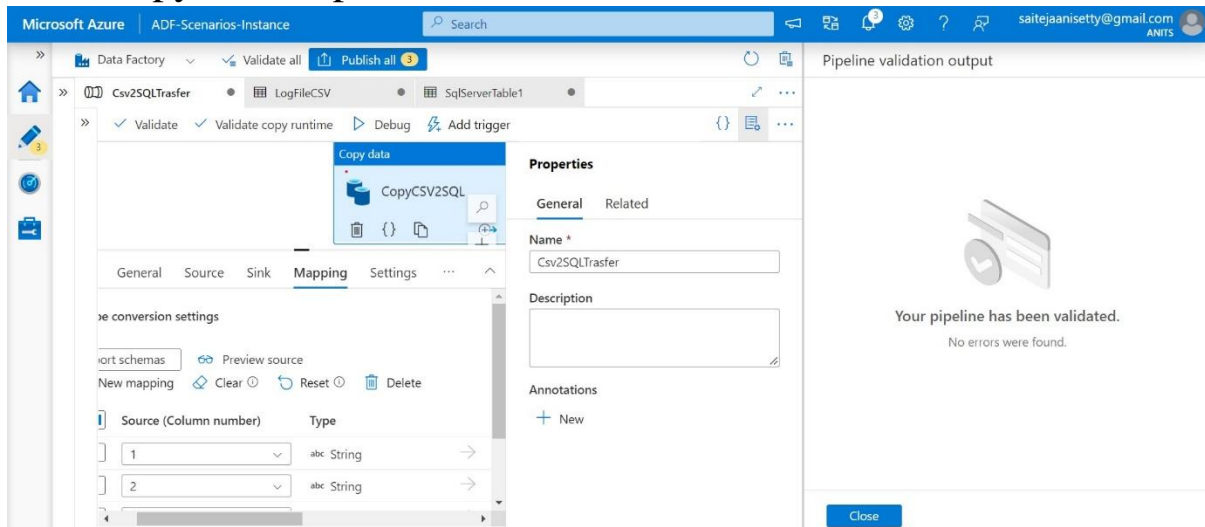
The screenshot shows the 'User properties' tab of an ADF Copy Data Pipeline. The pipeline is named 'Csv2SQLTrasfer' and is currently in the 'User properties' tab. The 'User properties' section contains a table with columns 'Name' and 'Value'. The table has two rows: 'Source' with value 'newcontainer//15092022\_log\_csv.csv' and 'Destination' with value 'dbo.LogNew'.

General Source Sink Mapping Settings **User properties**

**User properties**

Name	Value
Source	newcontainer//15092022_log_csv.csv
Destination	dbo.LogNew

## ADF Copy Data Pipeline -Validation



The screenshot shows the 'Validation' tab of an ADF Copy Data Pipeline. The pipeline is named 'Csv2SQLTrasfer' and is currently in the 'Validation' tab. The 'Validation' section contains a table with columns 'Source (Column number)' and 'Type'. The table has two rows: '1' with type 'abc String' and '2' with type 'abc String'. The 'Properties' section contains a table with columns 'Name' and 'Value'. The table has one row: 'Name' with value 'Csv2SQLTrasfer'. The 'Description' section contains a text area with the text 'Your pipeline has been validated. No errors were found.'.

General Source Sink Mapping Settings **Validation**

**Validation**

**Source (Column number)** **Type**

1	abc String
2	abc String

**Properties**

**General** **Related**

**Name \***

**Description**

**Annotations**

# ADF-Factory Validation

Factory validation output

Your factory has been validated.  
No errors were found.

Close

# ADF-Publishing

Publish all

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

Pending changes (3)

NAME	CHANGE	EXISTING
Pipelines		
Csv2SQLTrasfer	(New)	-
Datasets		
LogFileCSV	(New)	-
SqlServerTable1	(New)	-

Publish Cancel

# ADF Copy Data Pipeline-Trigger

Csv2SQLTrasfer

List Gantt

Rerun Rerun from activity Rerun from failed activity Refresh Update pipeline

Activity runs

Pipeline run ID 25a2db3e-cd66-4bc4-a18c-ae03b3e78736

All status

Showing 1 - 1 items

Activity name	Activity type	Run start	Duration	Status	Error
CopyCSV2SQL	Copy data	Sep 17, 2022, 5:28:10 pm	00:00:01	Queued	

## ADF Copy Data Pipeline-Run pipeline

Pipeline runs

Triggered Debug Rerun Cancel Refresh Edit columns List Gantt

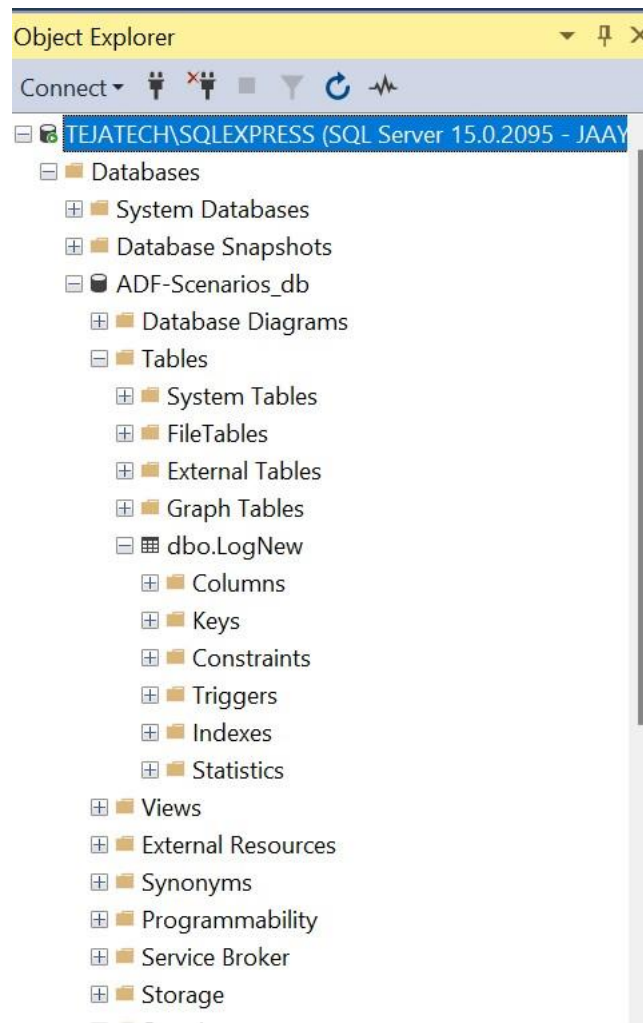
Filter by run ID or name Chennai, Kolkata, Mu... : Last 24 hours Pipeline name : All Status : All

Runs : Latest runs Triggered by : All Add filter X

Showing 1 - 1 items Last refreshed 0 minutes ago

<input type="checkbox"/> Pipeline name ↑↓	Run start ↑↓	Run end ↑↓	Duration	Triggered by	Status ↑↓	Error	Run
<input type="checkbox"/> Csv2SQLTrasfer	Sep 17, 2022, 5:28:08 pm	Sep 17, 2022, 5:28:33 pm	00:00:24	Manual trigger	✔ Succeeded		Original

## SSMS- MS SQL Server Table created [dbo.LogNew]





## SSMS-Output Table

SQLQuery1.sql - TEJ...rios_db (JAAY (63))*				
SELECT * FROM dbo.LogNew				
100 %				
Results Messages				
	EmployeeID	FIRSTNAME	OPERATION NAME	TimeStamp
1	EMPLOYEE ID	FIRST NAME	OPERATTION NAME	TIME STAMP
2	1	Rosy	List Storage Account Keys	26-06-2021 14:24
3	2	Mary	Delete SqlPools	25-06-2021 14:22
4	3	Jack	Pause SQL Analytics pools.	25-06-2021 14:22
5	4	Raees	Pause SQL Analytics pools.	23-06-2021 07:23
6	5	Ruby	List Storage Account Keys	20-06-2021 07:23
7	6	Romeo	Pause a Datawarehouse database.	19-06-2021 07:23
8	7	Mathew	Create Pipeline Run	17-06-2021 07:23
9	8	Mark	Create Pipeline Run	15-06-2021 14:22
10	9	Merlin	Create or Update any Pipeline.	13-06-2021 07:23
11	10	David	Create or Update any Pipeline.	12-06-2021 07:23
12	11	William	Validate Deployment	10-06-2021 07:23
13	12	Mitchel	Update SQL server firewall rules	06-06-2021 07:23
14	13	Mike	'auditIfNotExists' Policy action.	04-06-2021 07:23
15	14	Rachie	'auditIfNotExists' Policy action.	02-06-2021 07:23
16	15	Shane	Resume a Datawarehouse datab...	01-06-2021 07:23
17	NULL	NULL	NULL	NULL