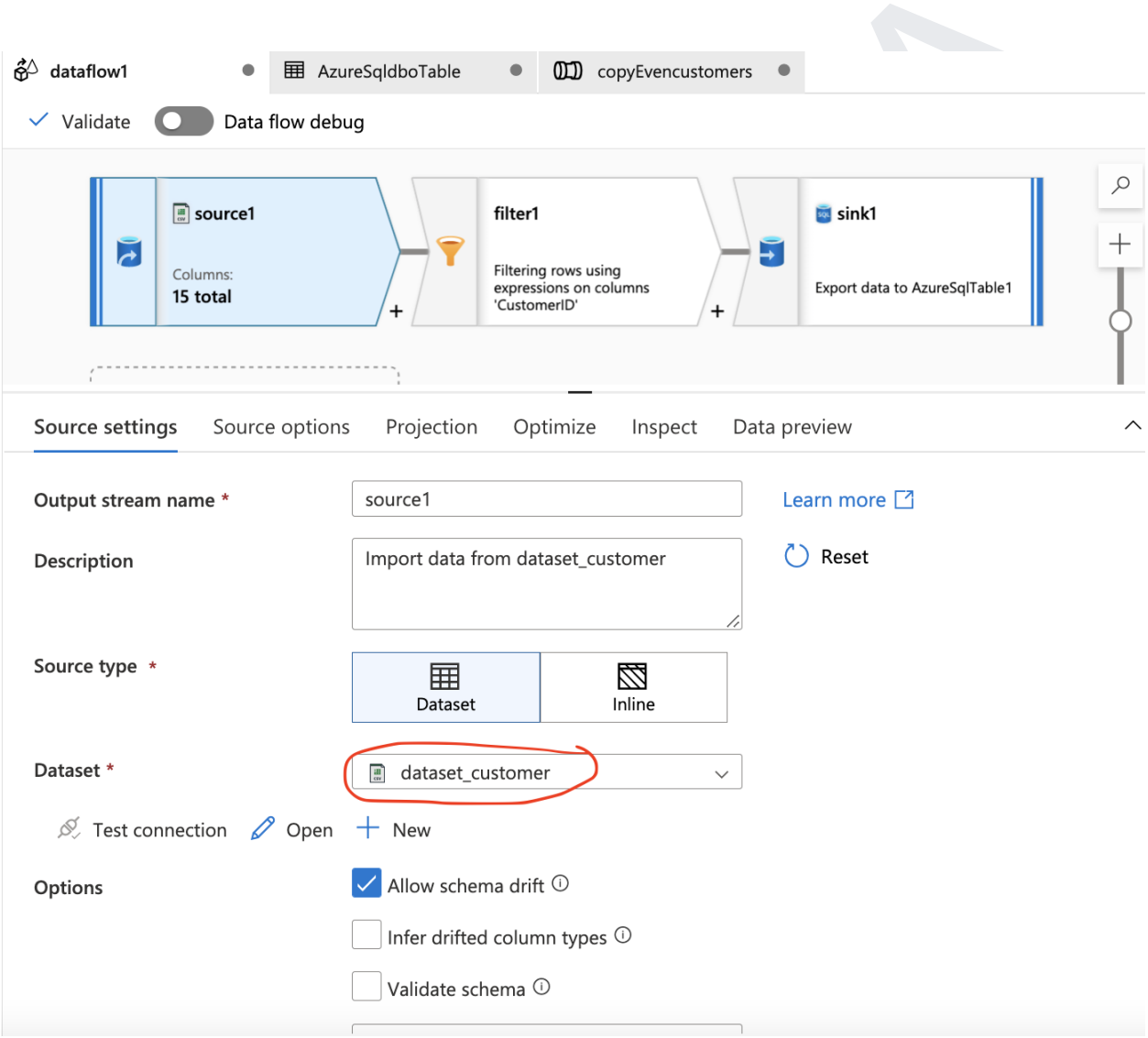
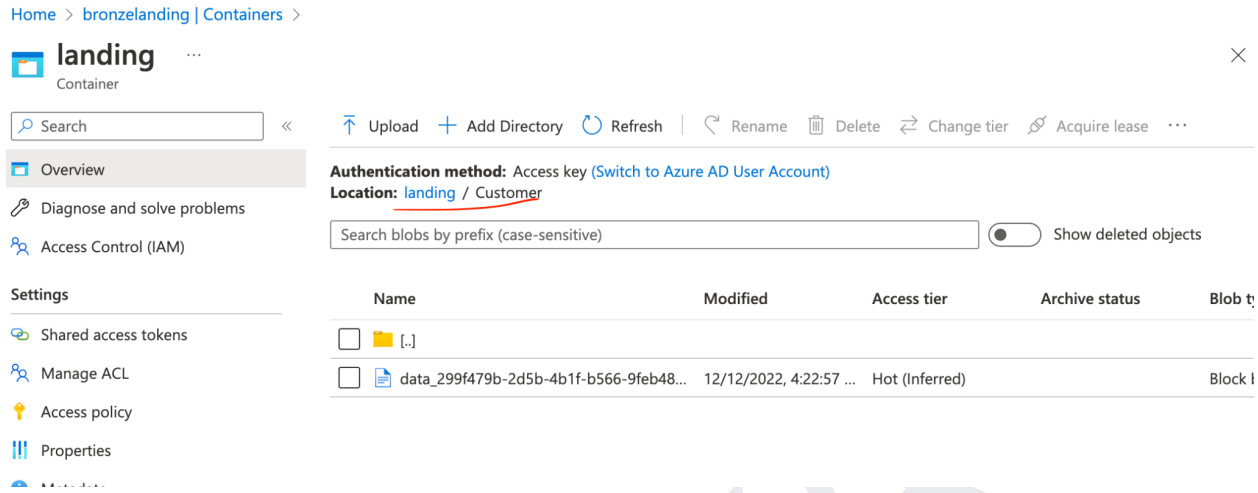
1. **Create a pipeline to copy the customer data from csv file to SQL where the customer id is an even number.**

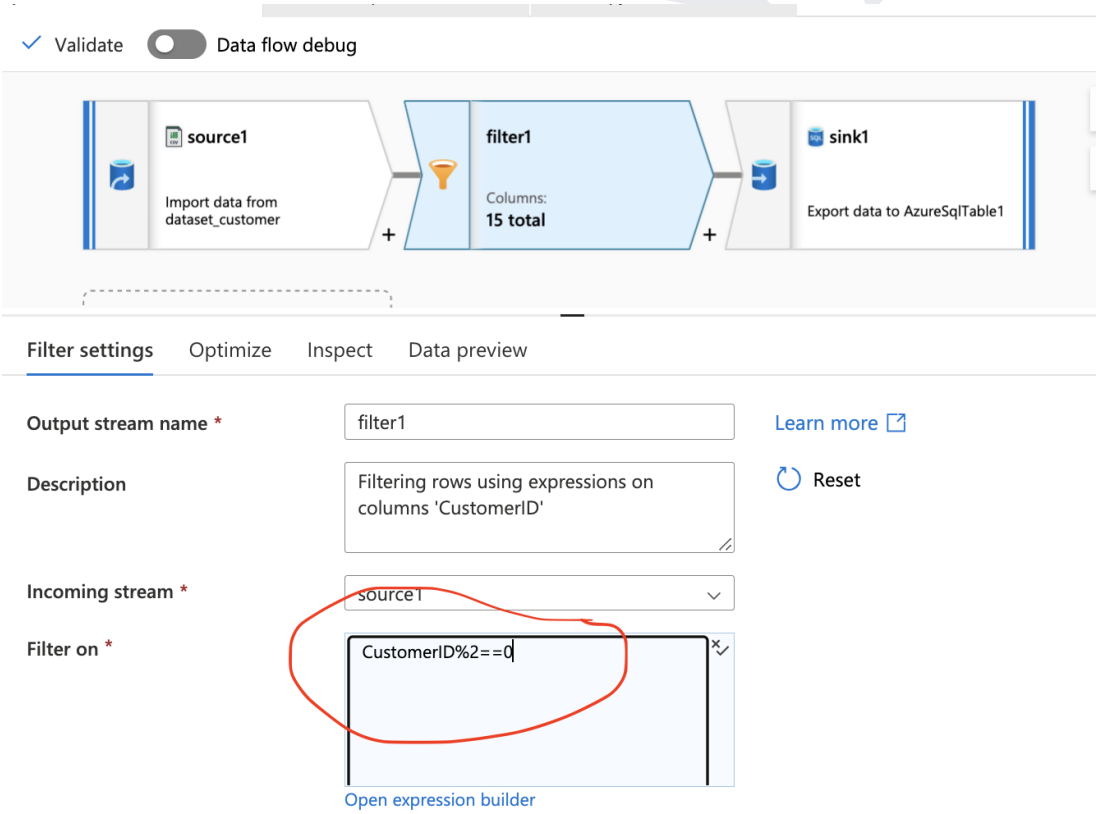
Click on dataflow and select source1 as customer file located in Azure datalake file storage



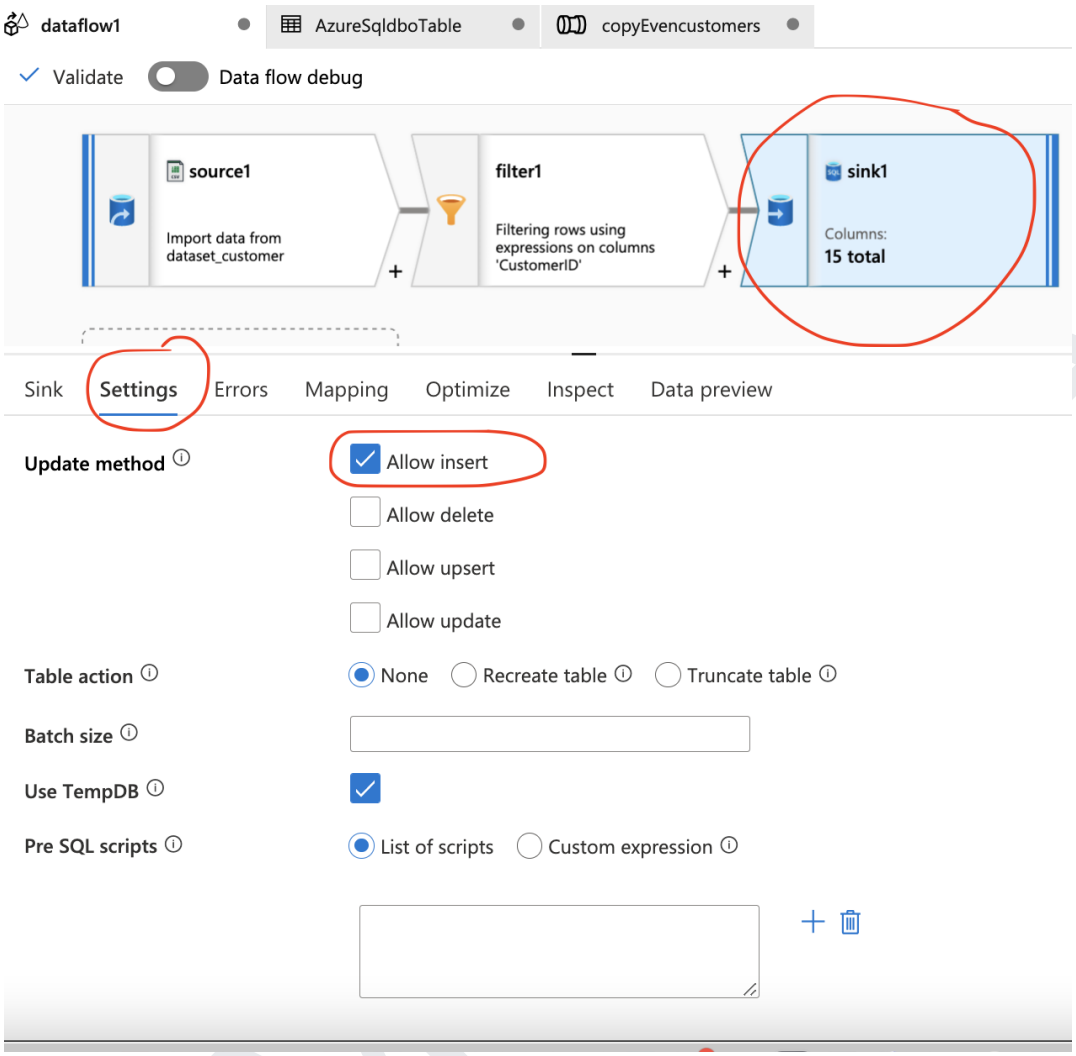
Customer File location:



Use Filter to get even number customers:



Sink into sqldataDB with new table as customerClone.



Set properties

Name

AzureSqlTable2

Linked service *

AzureSqlDatabase

Connect via integration runtime *

AutoResolveIntegrationRuntime (Managed Virtual Network)

Interactive authoring enabled

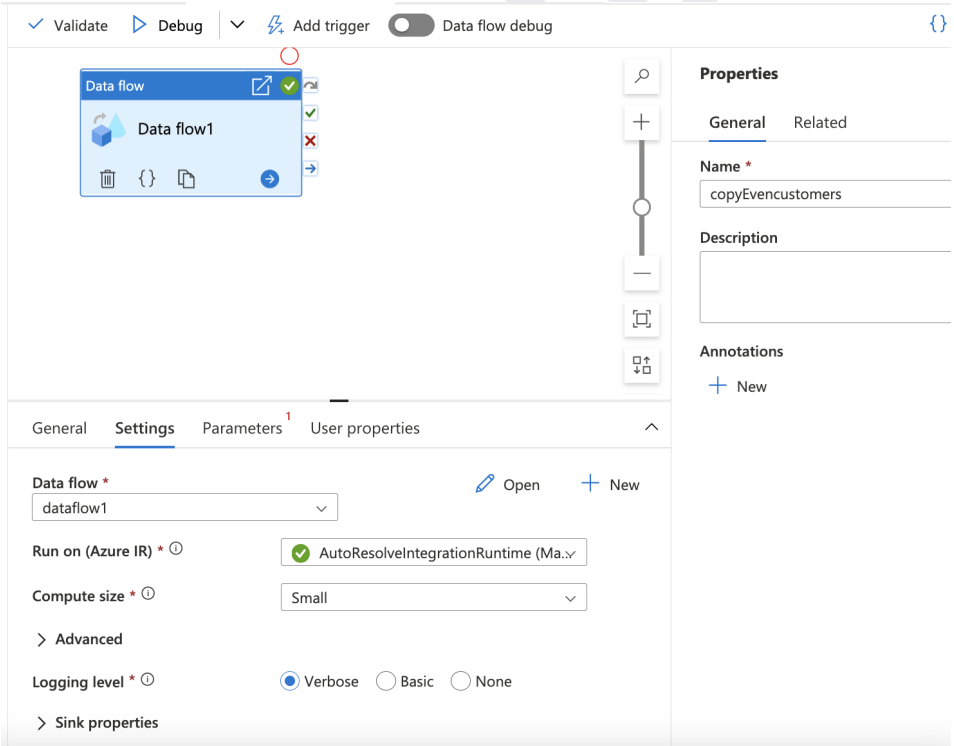
Select from existing table | Create new table

Schema and table name

Schema name . Table name

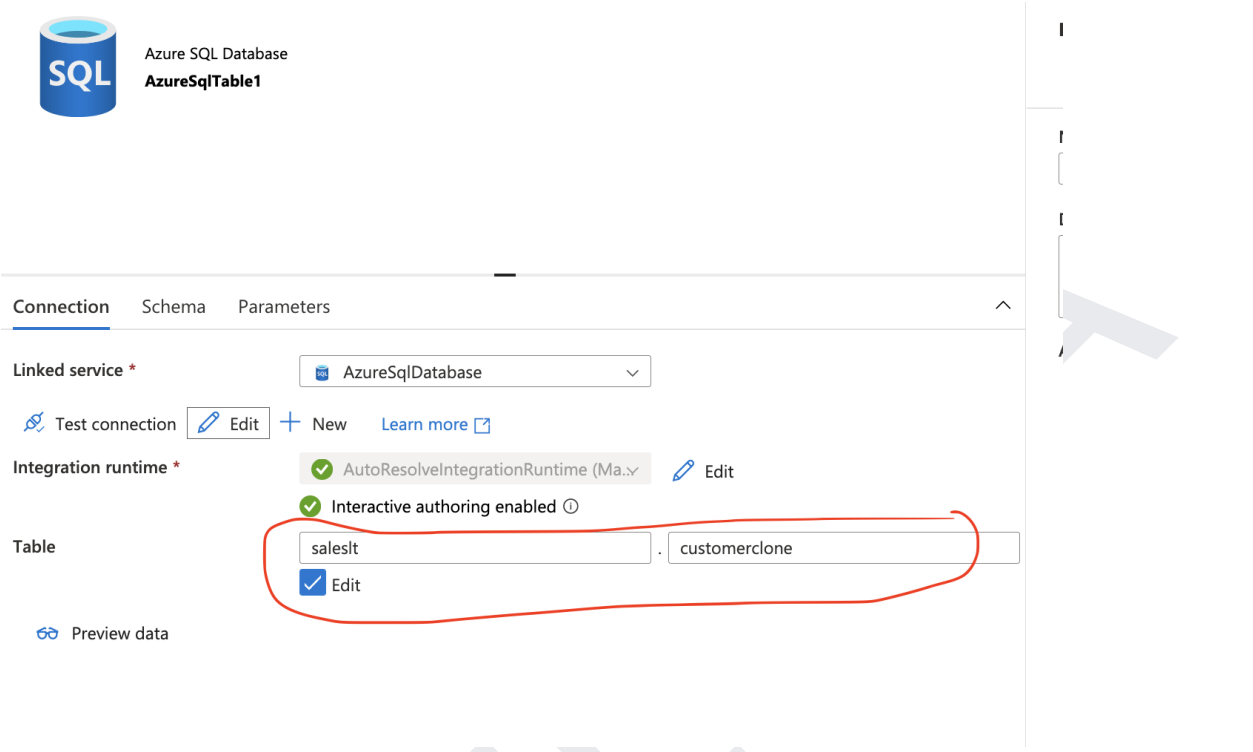
Advanced

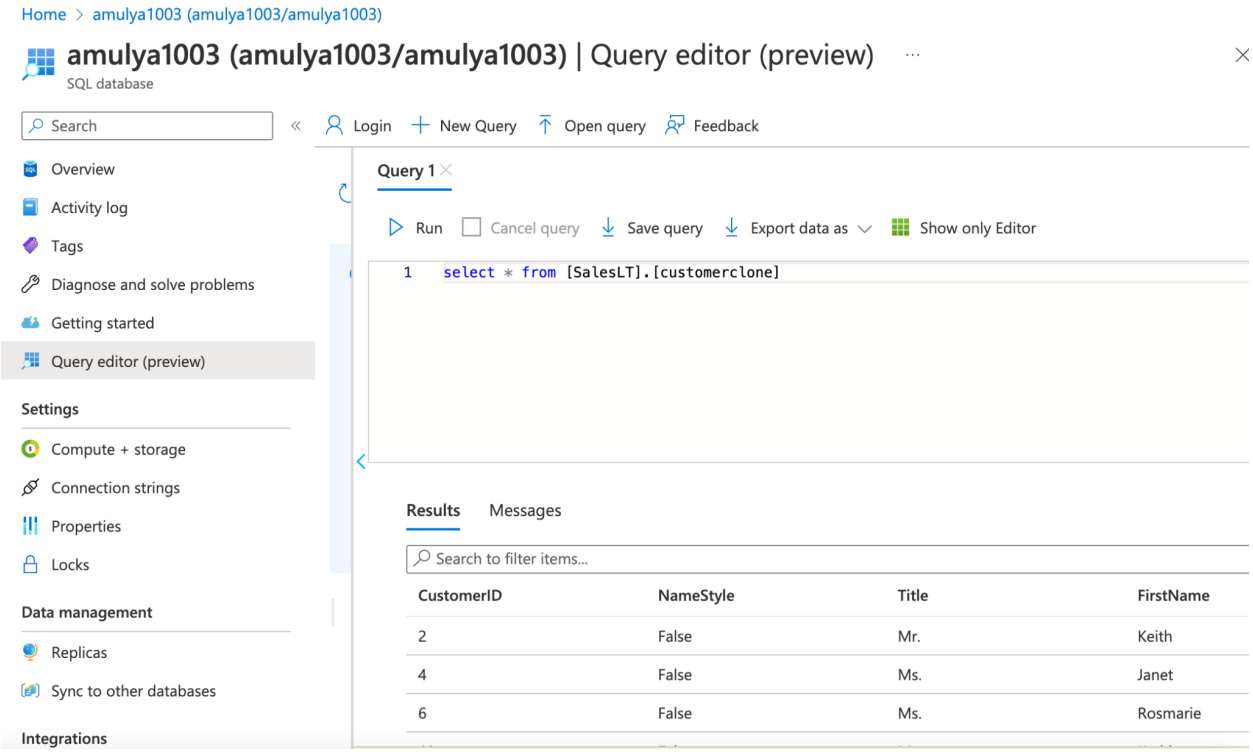
Create a pipeline and add dataflow to execute :



Publish and debug

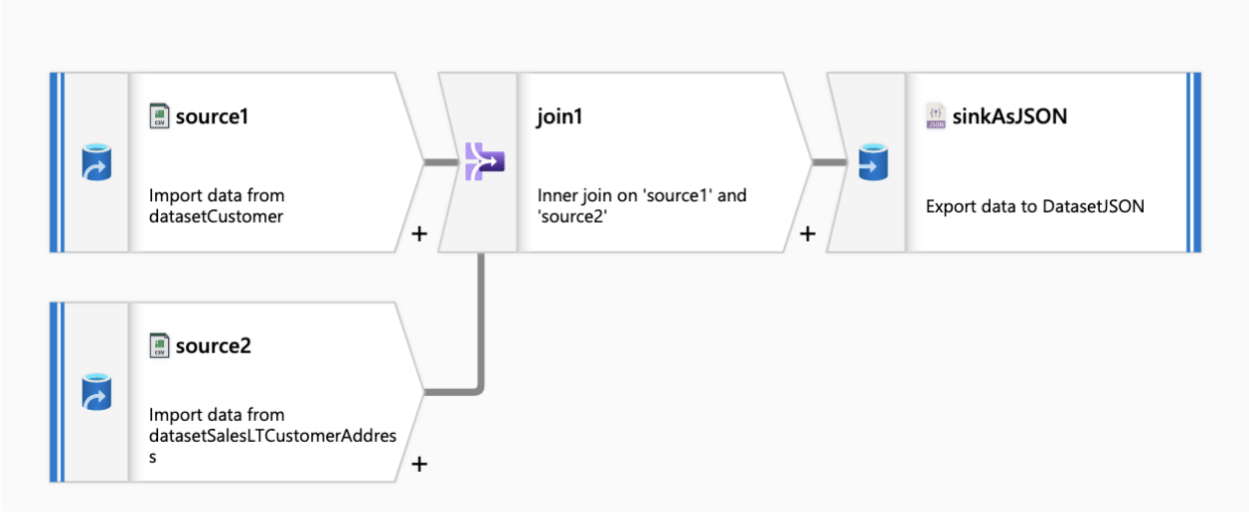
Verfying table created in SQLDB:





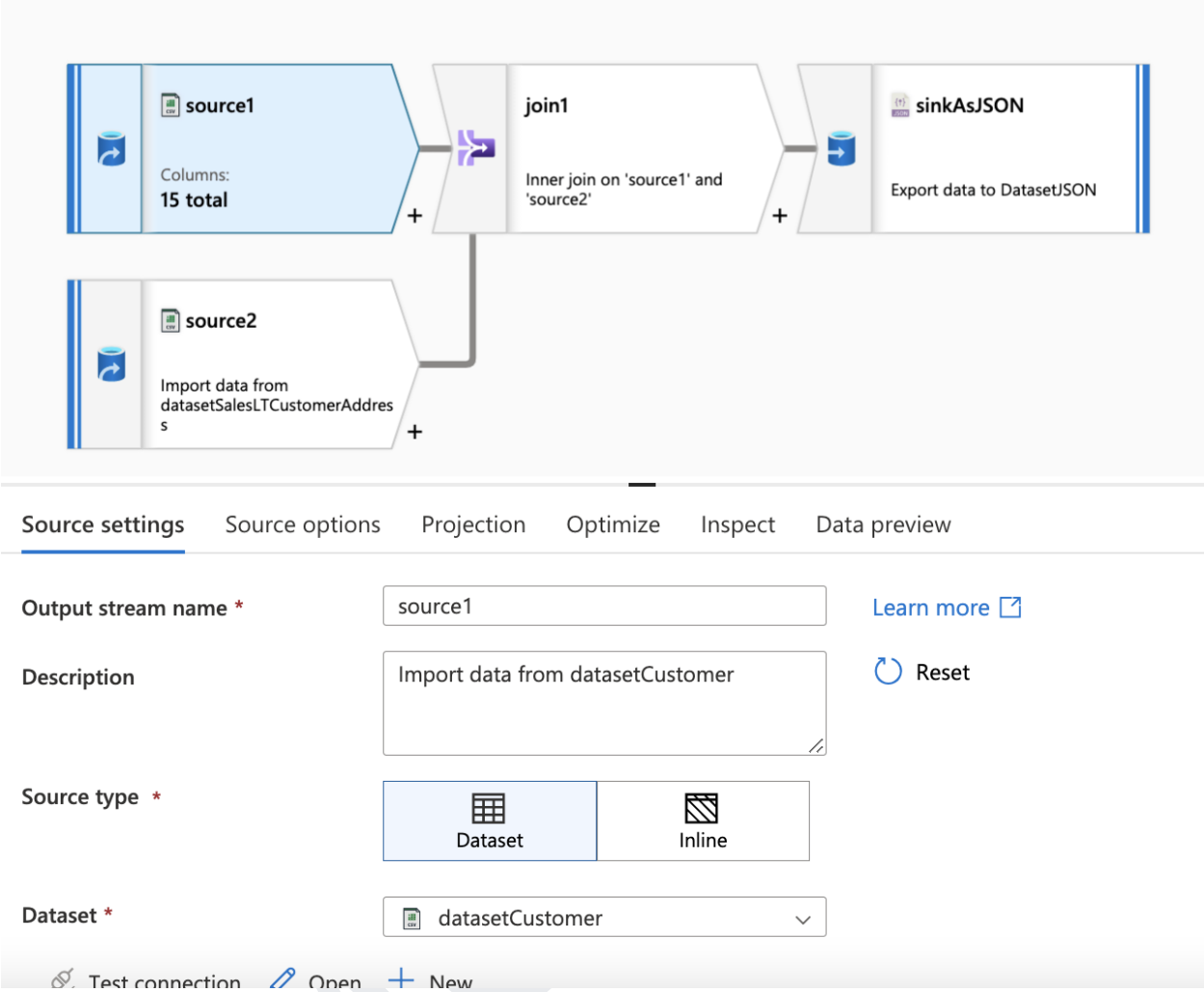
2. **Create a pipeline to join the two files (Customer, Customer Address) based on customer id and save the result as a JSON file.**

Import Customer data from datalake:

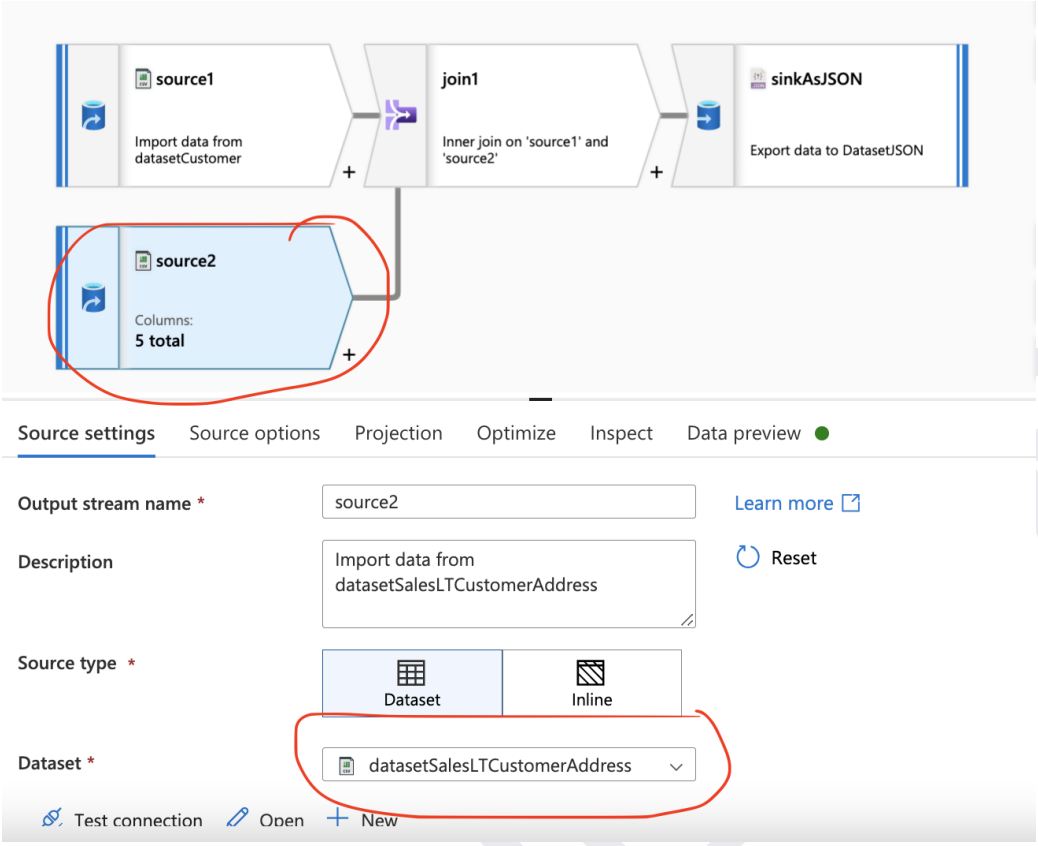


Parameters Settings

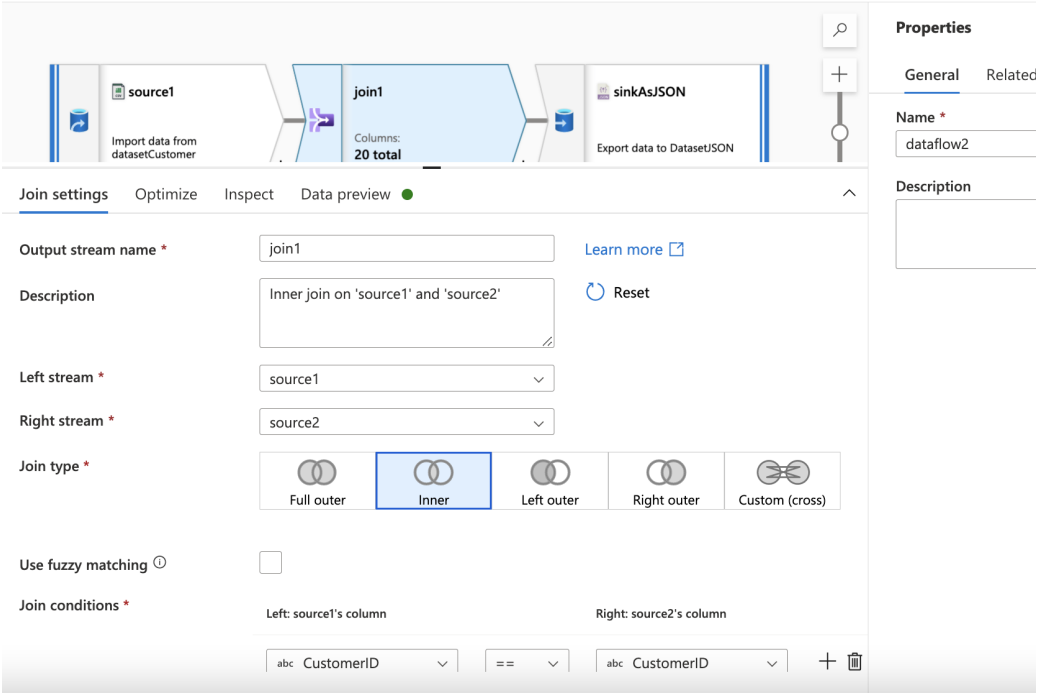
Select Dataset Pointing to customer.txt file:



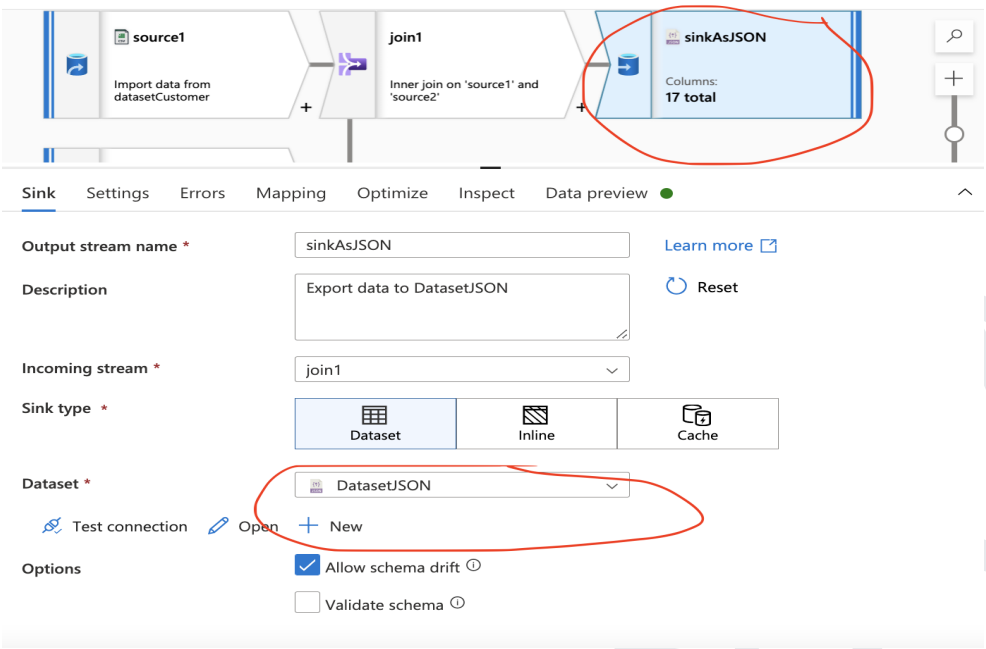
Import CustomerAddress data from datalake:



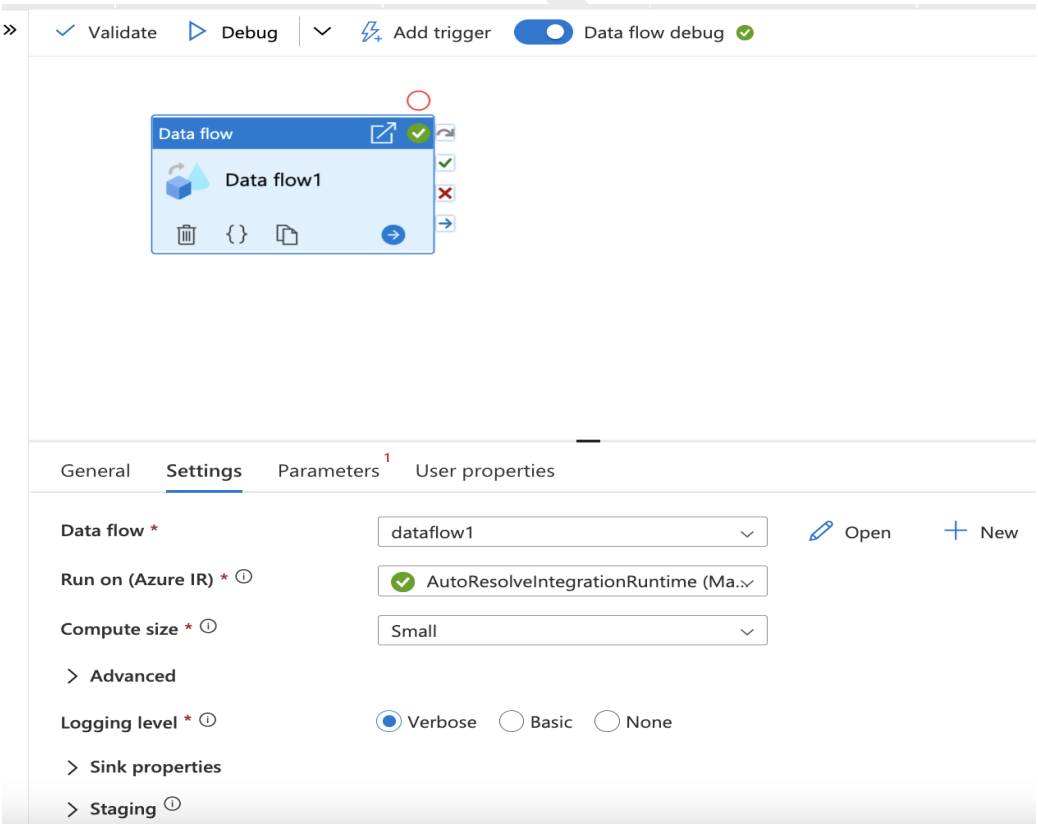
Join both tables based on customer ID



Select dataset -JSON to save:



Create a pipeline to execute the dataflow:



Home > bronzelanding | Containers >

landing

Container

Search

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Manage ACL

Access policy

Properties

Metadata

Upload | Add Directory | Refresh | Rename | Delete | Change

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

Search blobs by prefix (case-sensitive)

Name
Customer
_SUCCESS
part-00000-0f33fcca-af83-40db-a6cb-317e010d43fe-c000.json
SalesLt.Customer.txt
SalesLt.CustomerAddress.txt
SalesLt.Product.txt

!!!! Successfully created JSON file!!!

Home > bronzelanding | Containers > landing >

part-00000-0f33fcca-af83-40db-a6cb-317e010d43fe-c000.json

Blob

Save Discard Download Refresh Delete

Overview Versions Edit Generate SAS


```
1  {"CustomerID":"29485","NameStyle":"False","Title":"Ms.","FirstName":"Catherine","MiddleName":"R.","LastName":"Abel","CompanyName":"Pr
2  {"CustomerID":"29486","NameStyle":"False","Title":"Ms.","FirstName":"Kim","LastName":"Abercrombie","CompanyName":"Riders Company","Sa
3  {"CustomerID":"29489","NameStyle":"False","Title":"Ms.","FirstName":"Frances","MiddleName":"B.","LastName":"Adams","CompanyName":"Are
4  {"CustomerID":"29490","NameStyle":"False","Title":"Ms.","FirstName":"Margaret","MiddleName":"J.","LastName":"Smith","CompanyName":"Bi
5  {"CustomerID":"29492","NameStyle":"False","Title":"Mr.","FirstName":"Jay","LastName":"Adams","CompanyName":"Valley Bicycle Specialist
6  {"CustomerID":"29494","NameStyle":"False","Title":"Mr.","FirstName":"Samuel","MiddleName":"N.","LastName":"Agcaoili","CompanyName":"V
7  {"CustomerID":"29496","NameStyle":"False","Title":"Mr.","FirstName":"Robert","MiddleName":"E.","LastName":"Ahlering","CompanyName":"F
8  {"CustomerID":"29497","NameStyle":"False","Title":"Mr.","FirstName":"François","LastName":"Ferrier","CompanyName":"Great Bikes ","Sal
9  {"CustomerID":"29499","NameStyle":"False","Title":"Ms.","FirstName":"Amy","MiddleName":"E.","LastName":"Alberts","CompanyName":"Valle
10 {"CustomerID":"29502","NameStyle":"False","Title":"Mr.","FirstName":"Paul","MiddleName":"L.","LastName":"Alcorn","CompanyName":"Major
11 {"CustomerID":"29503","NameStyle":"False","Title":"Mr.","FirstName":"Gregory","MiddleName":"F.","LastName":"Alderson","CompanyName":"
12 {"CustomerID":"29503","NameStyle":"False","Title":"Mr.","FirstName":"Gregory","MiddleName":"F.","LastName":"Alderson","CompanyName":"
13 {"CustomerID":"29505","NameStyle":"False","Title":"Ms.","FirstName":"Michelle","LastName":"Alexander","CompanyName":"Imported and Dom
14 {"CustomerID":"29506","NameStyle":"False","Title":"Mr.","FirstName":"Sean","MiddleName":"P.","LastName":"Jacobson","CompanyName":"Sys
15 {"CustomerID":"29508","NameStyle":"False","Title":"Mr.","FirstName":"Marvin","MiddleName":"N.","LastName":"Allen","CompanyName":"Moun
16 {"CustomerID":"29510","NameStyle":"False","Title":"Mr.","FirstName":"Cecil","MiddleName":"J.","LastName":"Allison","CompanyName":"Des
17 {"CustomerID":"29511","NameStyle":"False","Title":"Mr.","FirstName":"Oscar","MiddleName":"L.","LastName":"Alpuerto","CompanyName":"Bo
18 {"CustomerID":"29515","NameStyle":"False","Title":"Mr.","FirstName":"Maxwell","MiddleName":"J.","LastName":"Amland","CompanyName":"Se
19 {"CustomerID":"29517","NameStyle":"False","Title":"Ms.","FirstName":"Ramona","MiddleName":"J.","LastName":"Antrim","CompanyName":"Nov
```

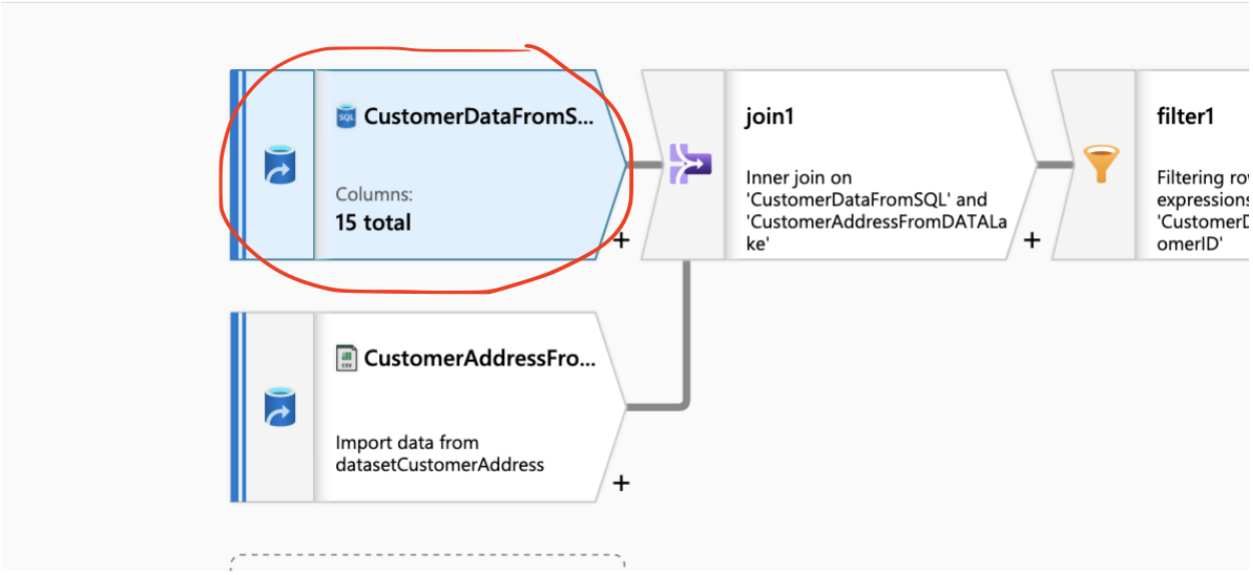

Json Preview

3.Create a pipeline to read the Customer table data from SQL and CustomerAddress data from CSV, join both of them ,and then save the result where customer id> 1000 & Customer id <2000 in ascending order as a Parquet file.

Select CustomerData from SQLDB:

AMULYA

Validate Data flow debug Debug Settings



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

Output stream name * CustomerDataFromSQL Learn more

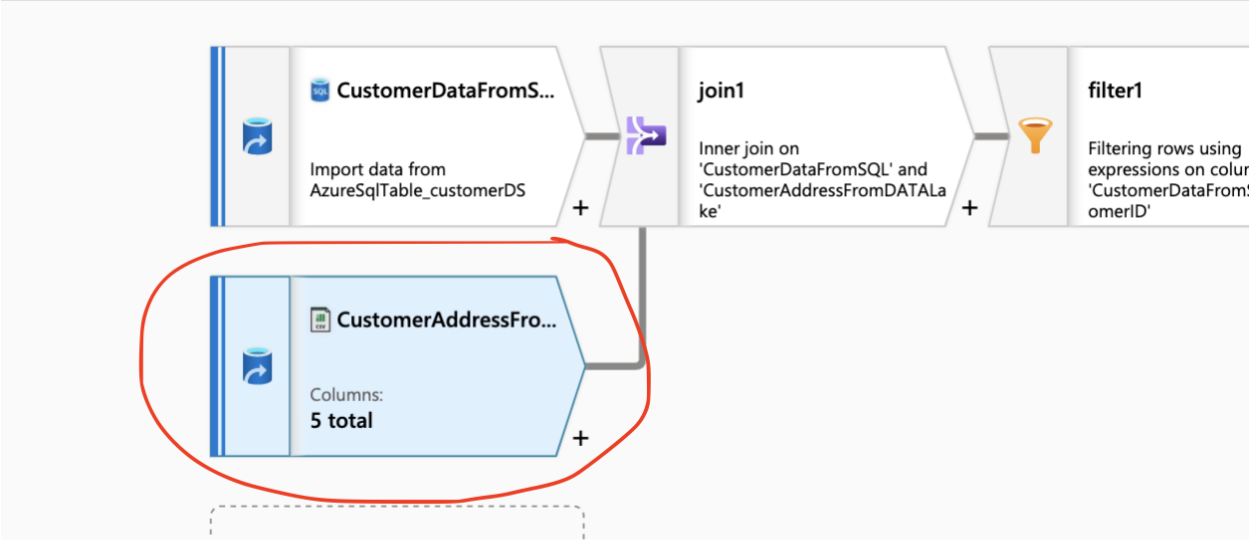
Description Import data from AzureSqlTable_customerDS Reset

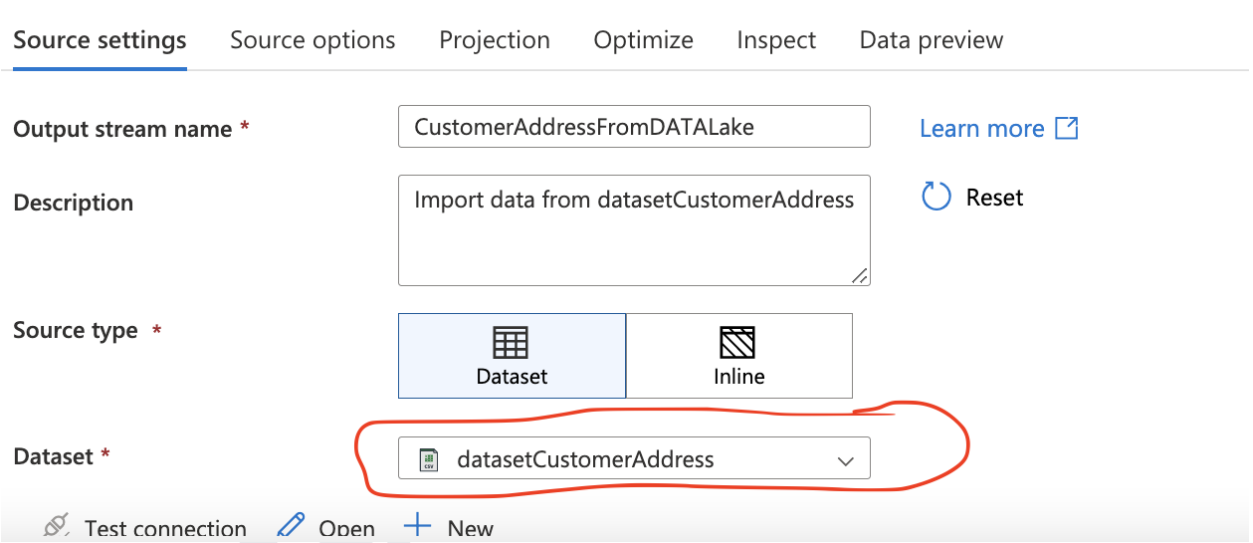
Source type * Dataset | Inline

Dataset * AzureSqlTable_customerDS

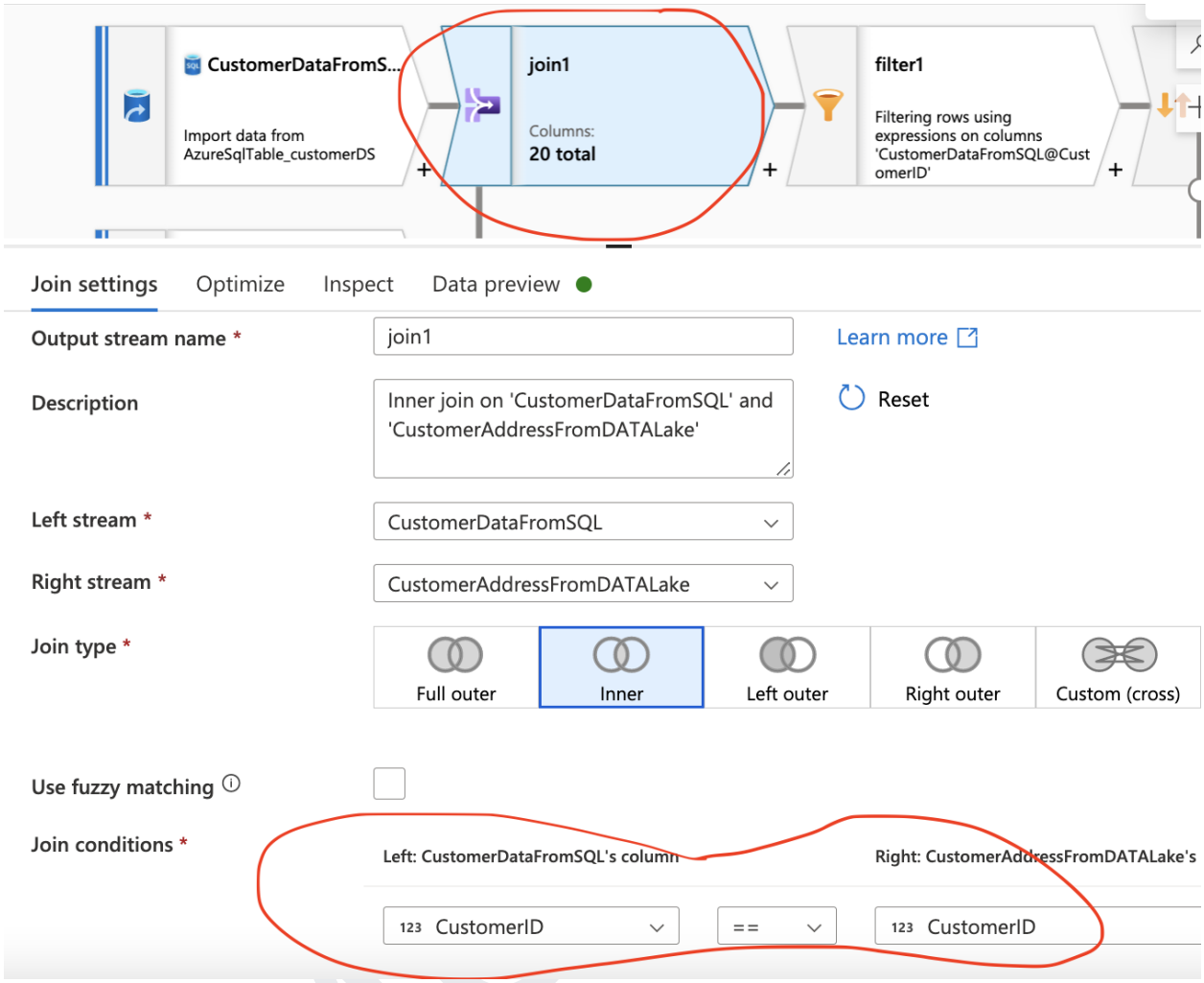
Test connection Open New

Select CustomerAddress data from dataLake:

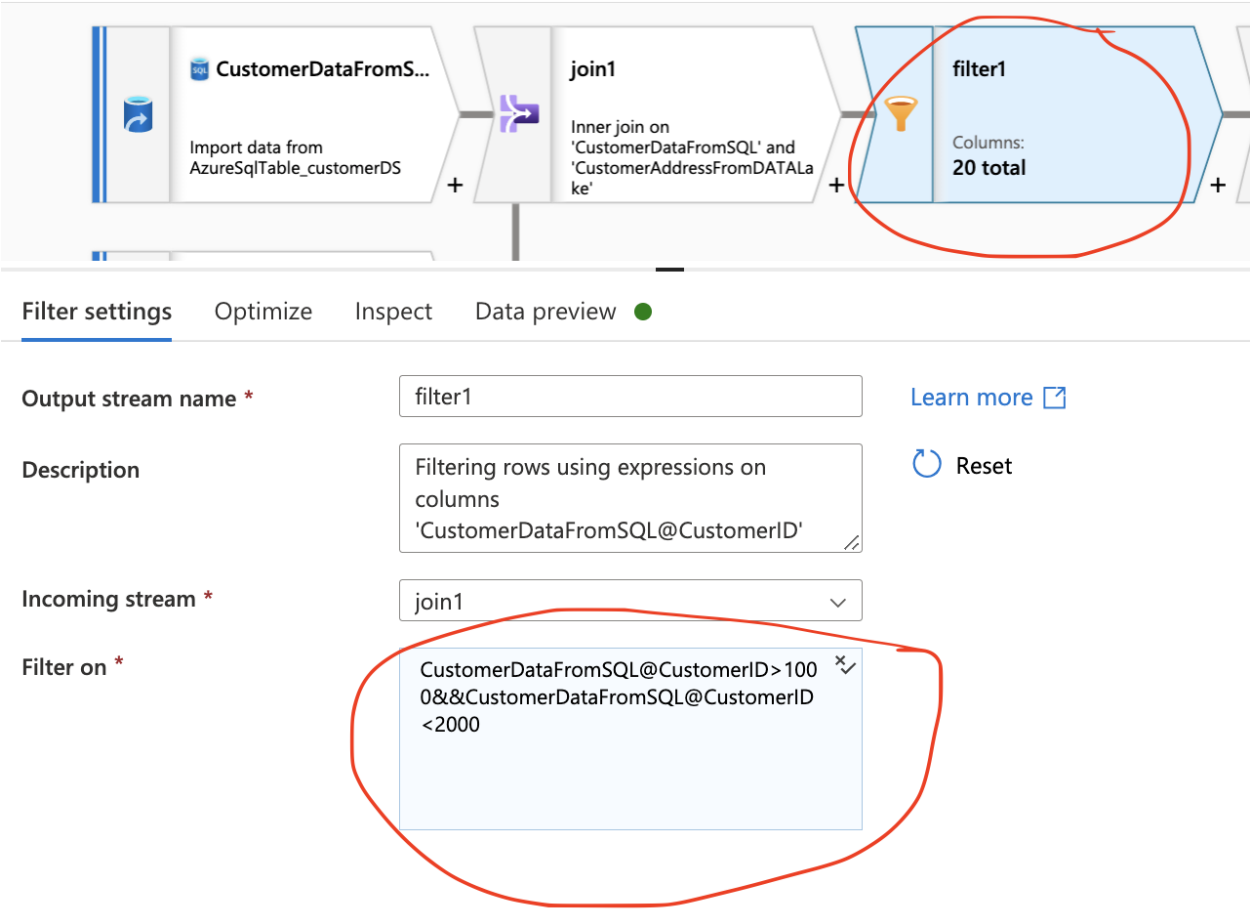




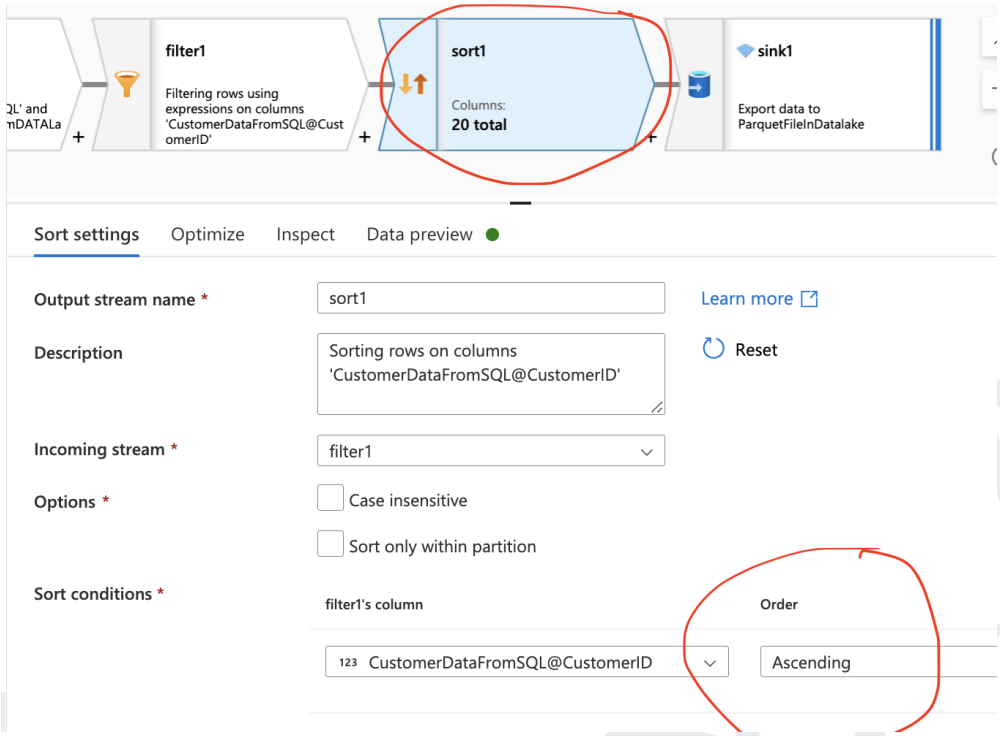
Join both tables based on customerID



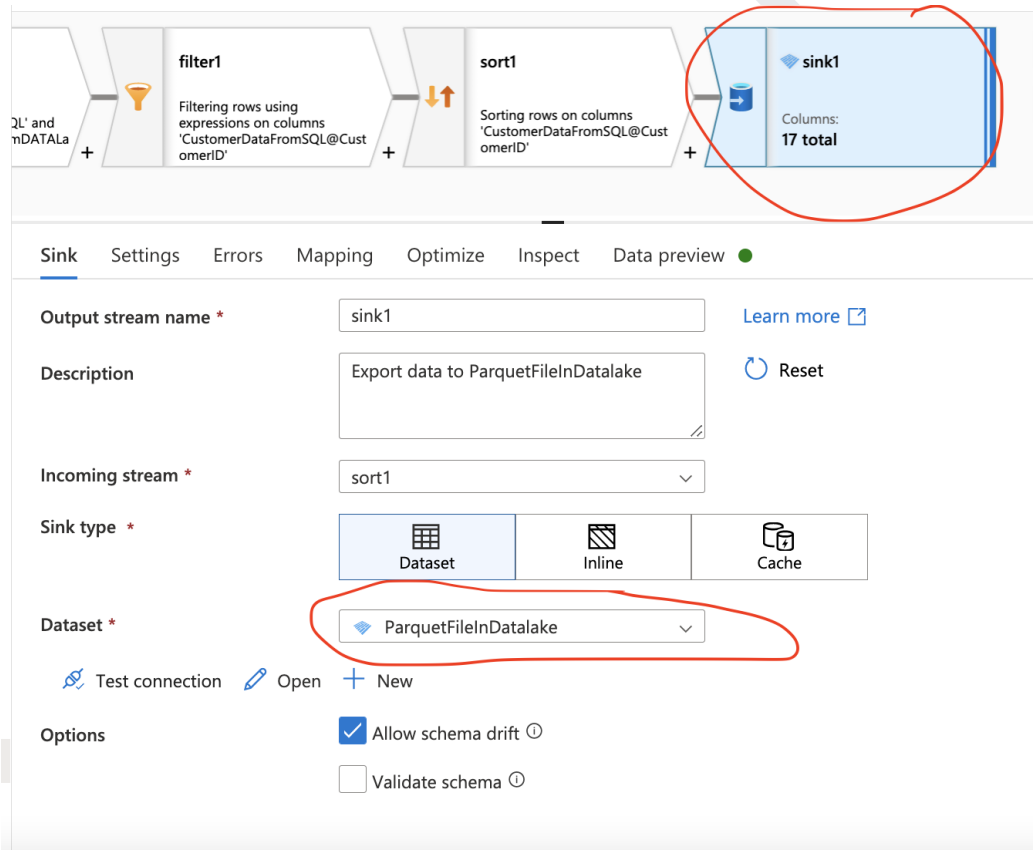
Filter customer id > 1000 && Customer id <2000



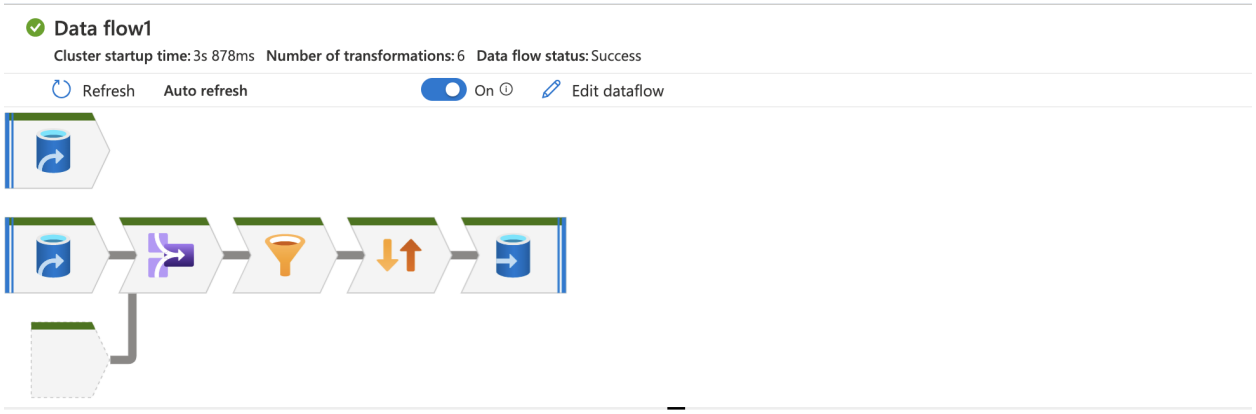
Sort CustomerID in ascending order:



Sink as Parquet File:

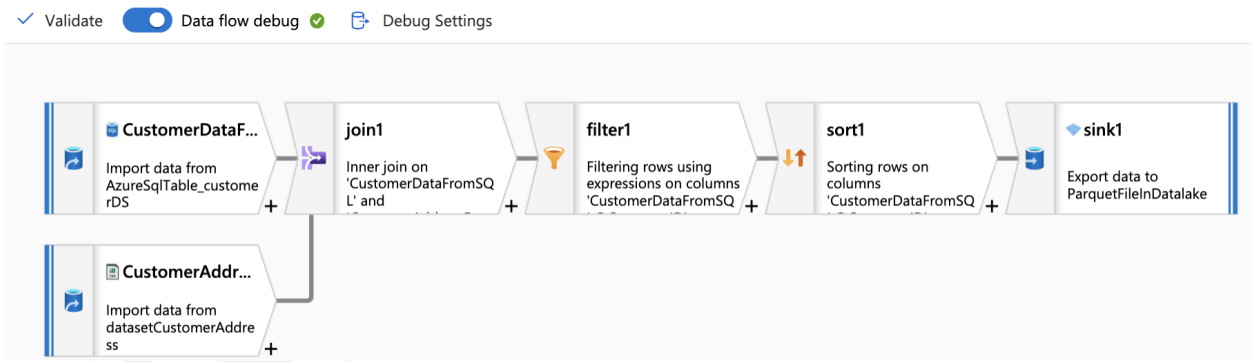


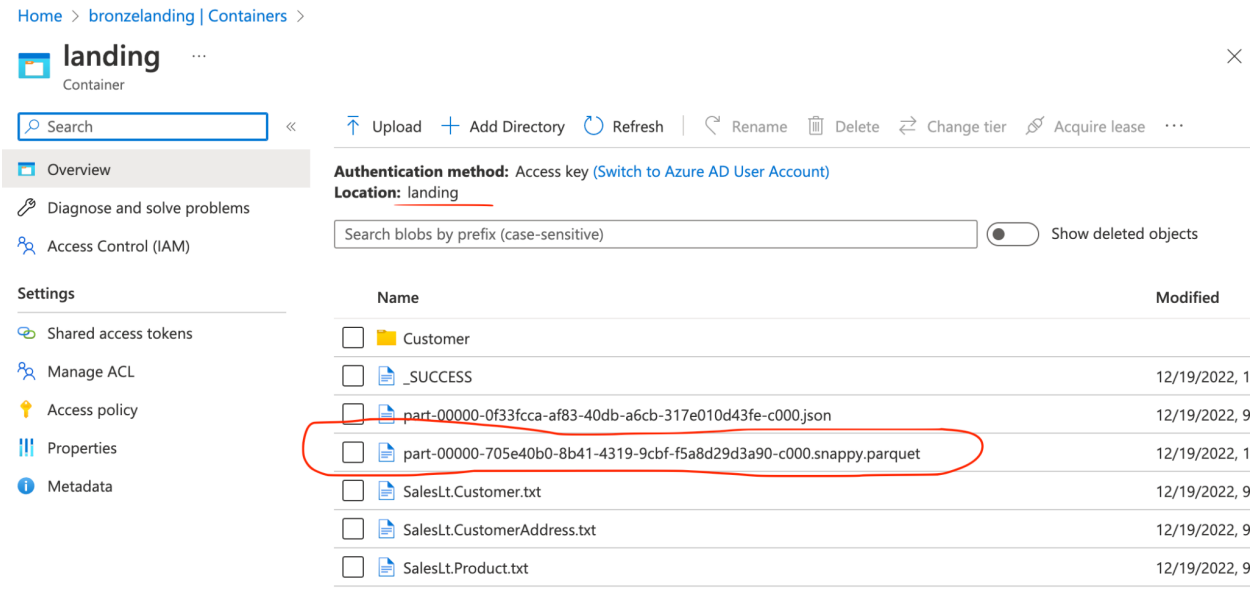
Output Flow:



Sinks All streams

Sink	Status	Processing time	Highest processing time	Rows written	Stages	Lineage
sink1	Succeeded	13s	4s 314ms	0		



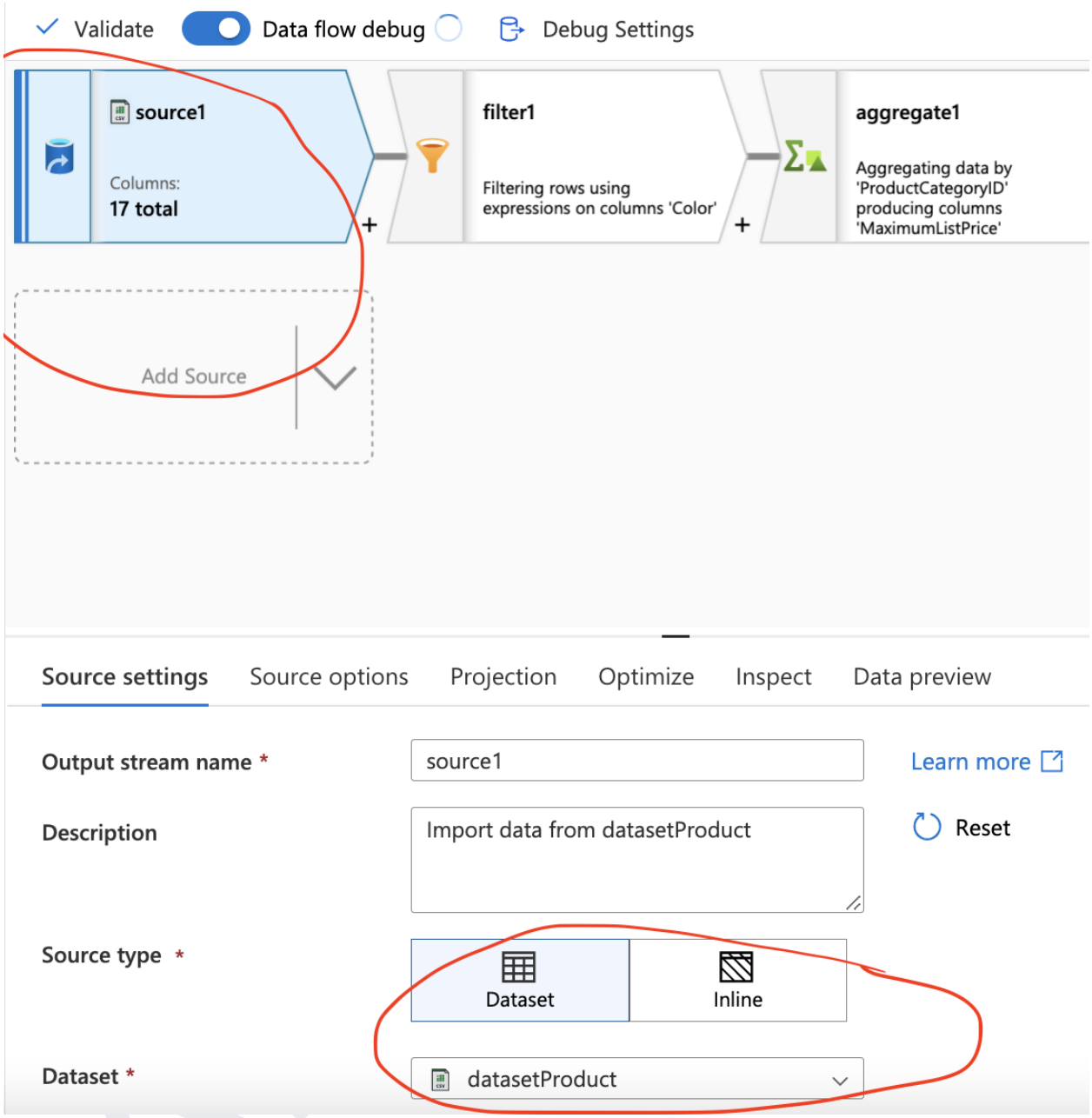


!!!!! Successfully Parquet file created !!!!

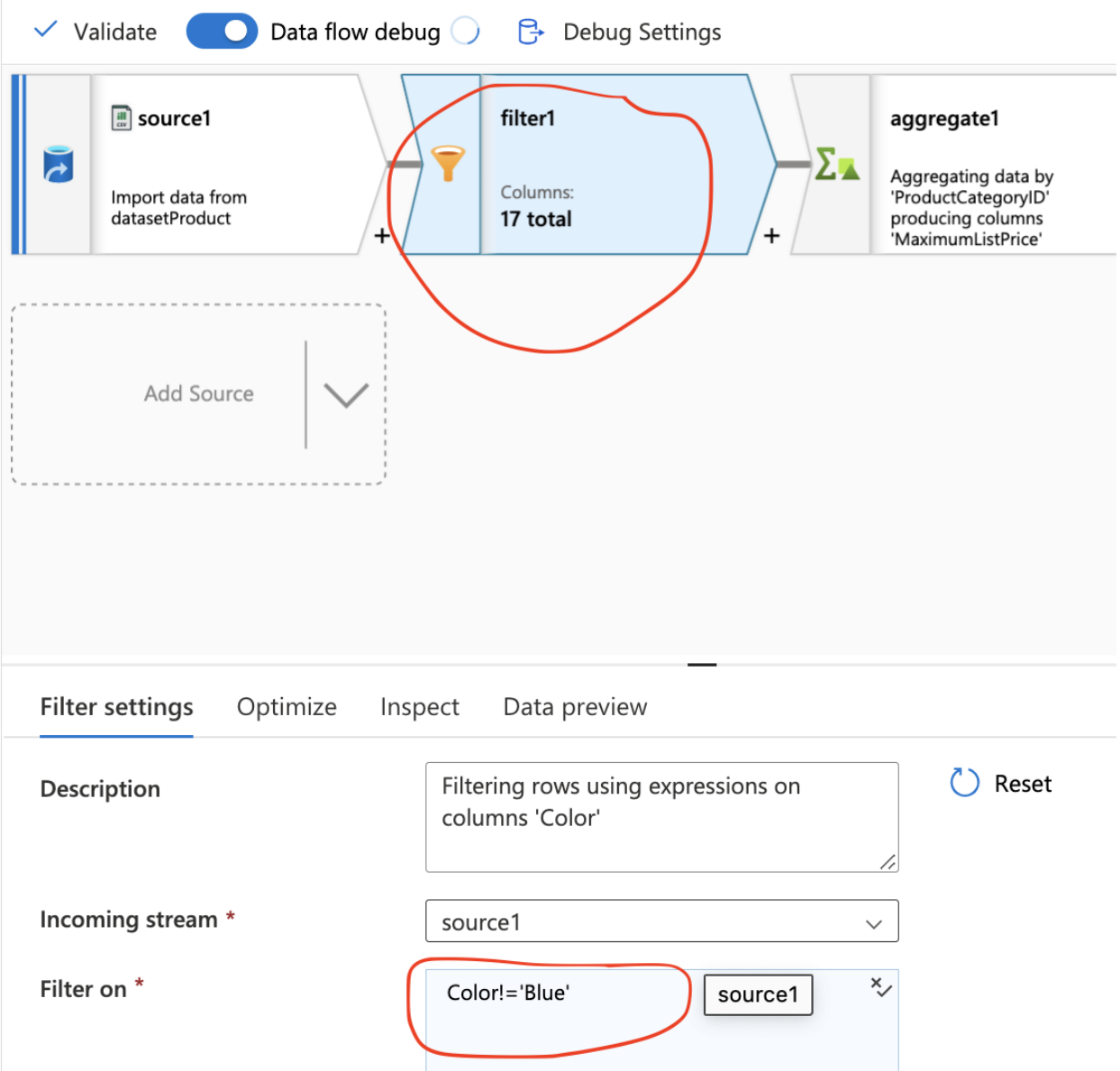
4. create a pipeline to read the Product CSV file, and calculate the highest listPrice of any product under each productcategory. Ensure that product shouldn't be of blue in color and save the result as CSV file inside ProductResult folder.

Select Product table from datalake

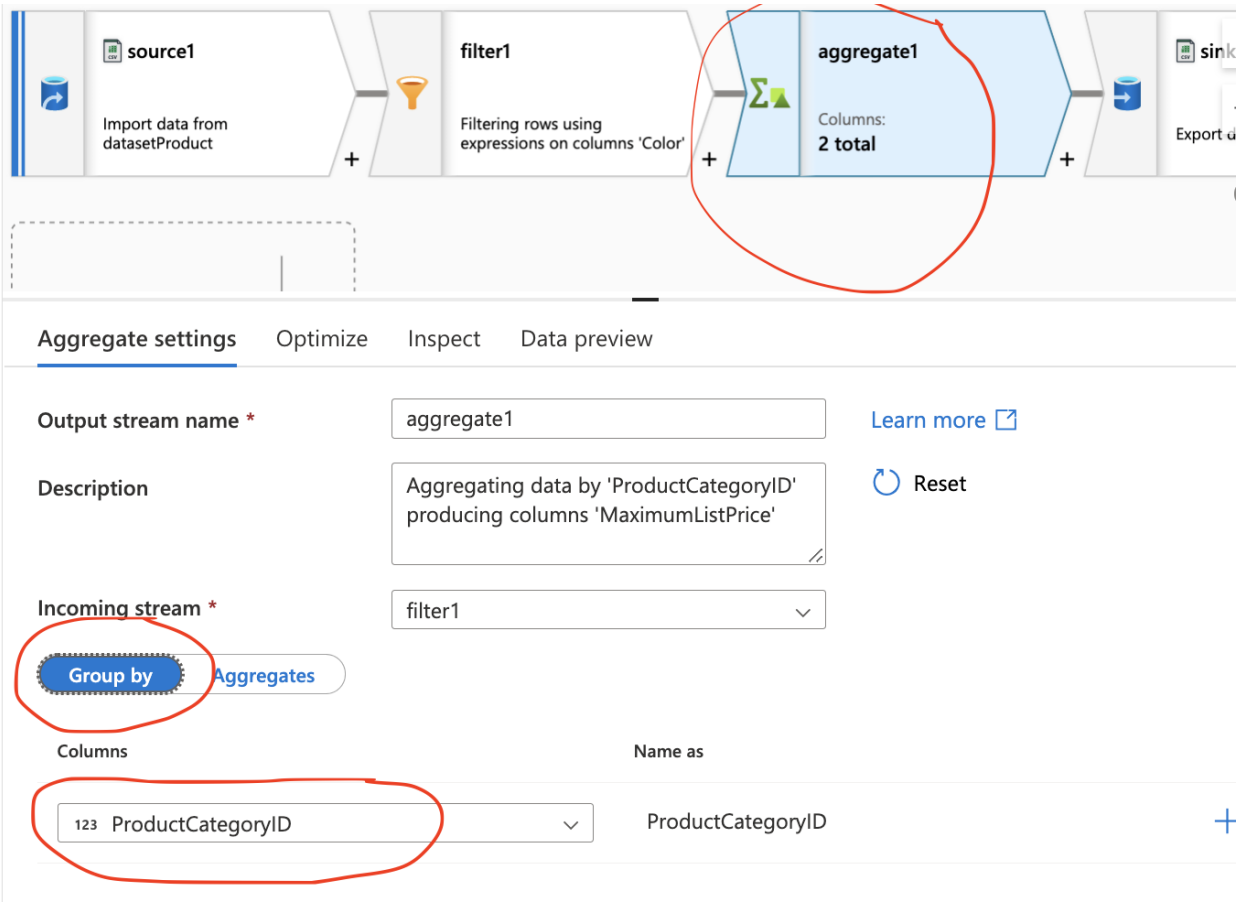
AMULY



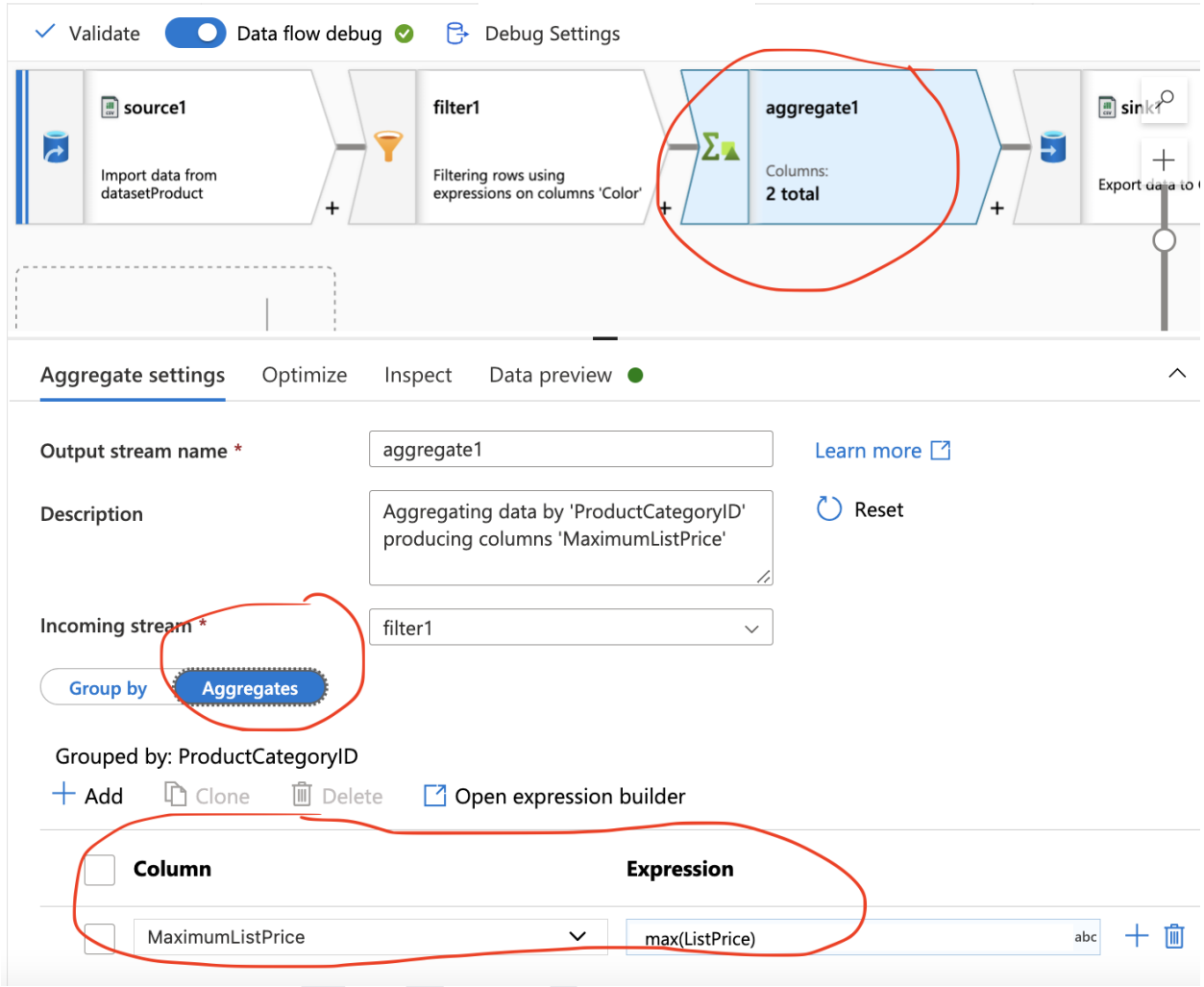
Color Should not be in Blue Color



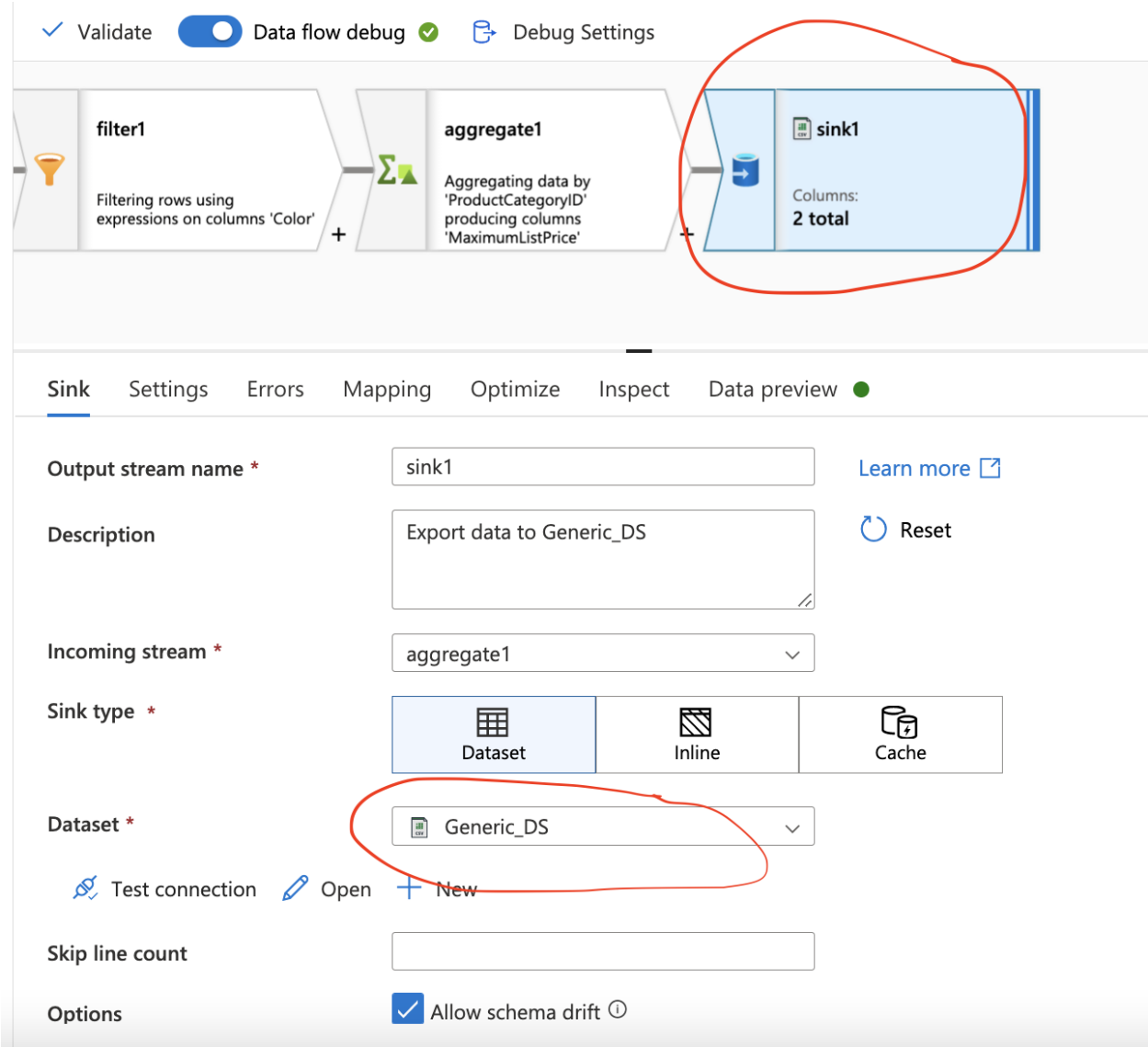
maximum of listprice in ProductID



SELECT aggregate to get MAX(listPrice)



Sink in productResult folder in datalake


Validate
Data flow debug
Debug Settings
filter1
Filtering rows using expressions on columns 'Color'
aggregate1
Aggregating data by 'ProductCategoryID' producing columns 'MaximumListPrice'
sink1
Columns:
2 total
Sink
Settings
Errors
Mapping
Optimize
Inspect
Data preview
Output stream name *
sink1
Learn more
Description
Export data to Generic_DS
Reset
Incoming stream *
aggregate1
Sink type *
Dataset
Inline
Cache
Dataset *
Generic_DS
Test connection
Open
New
Skip line count
Options
Allow schema drift

DelimitedText
Generic_DS

Connection Schema **Parameters**

New Delete

Name	Type	Default value
FolderPath	String	ProductResult

DelimitedText
Generic_DS

Connection Schema Parameters

Linked service * AzureDataLakeStorage_LS

Test connection Edit New Learn more

Integration runtime * AutoResolveIntegrationRuntime (Ma... Edit

Interactive authoring enabled

File path * landing / @dataset().FolderPath / File name

Compression type None

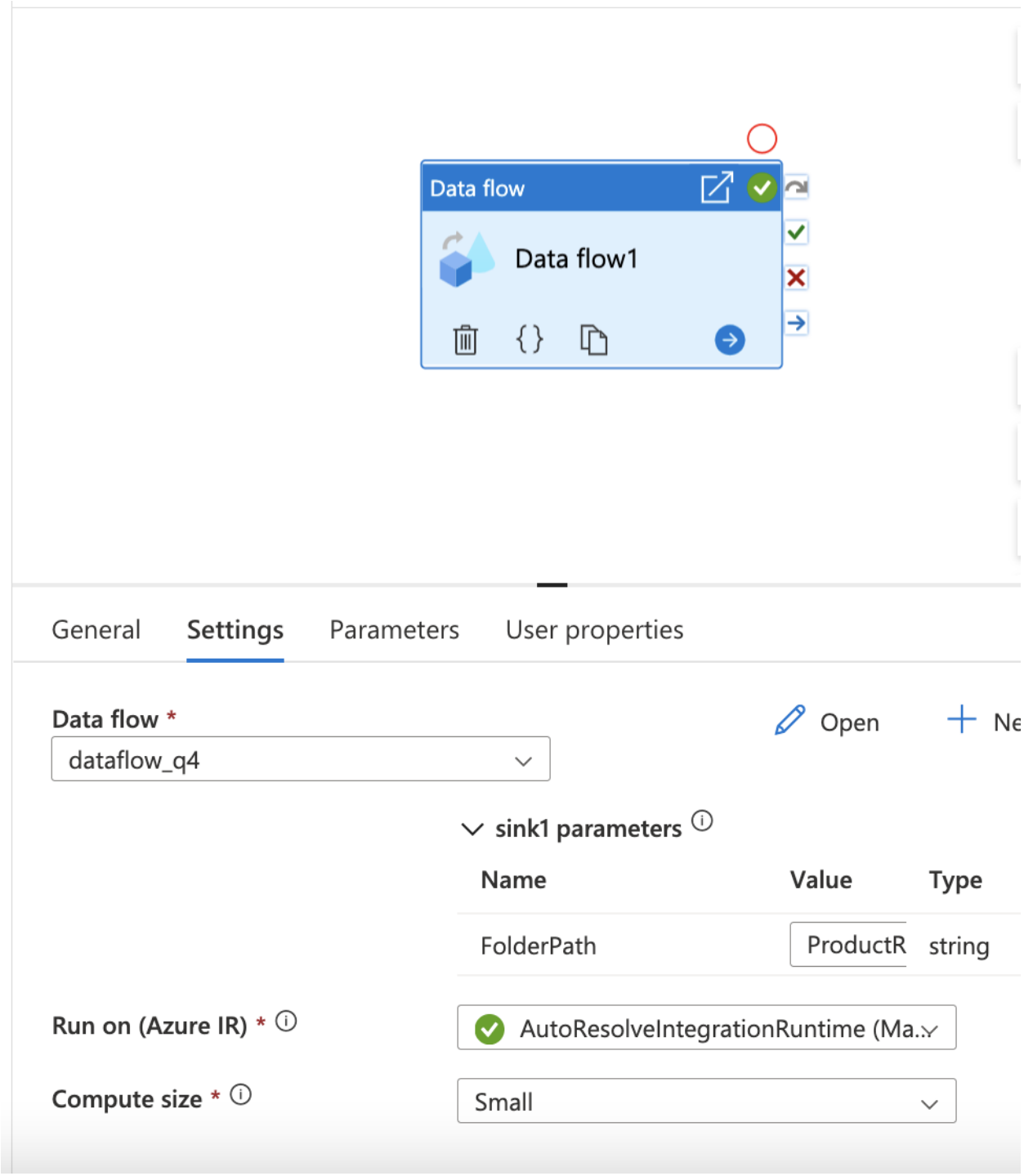
Column delimiter Comma (,)

Edit

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

Edit

Add dataflow to pipeline in order to execute:



Pipeline was modified after this run. The current pipeline configuration is shown.

Data flow1

Cluster startup time: 1s 282ms **Number of transformations:** 4 **Data flow status:** Success

Refresh **Auto refresh** On Edit dataflow

Sinks All streams

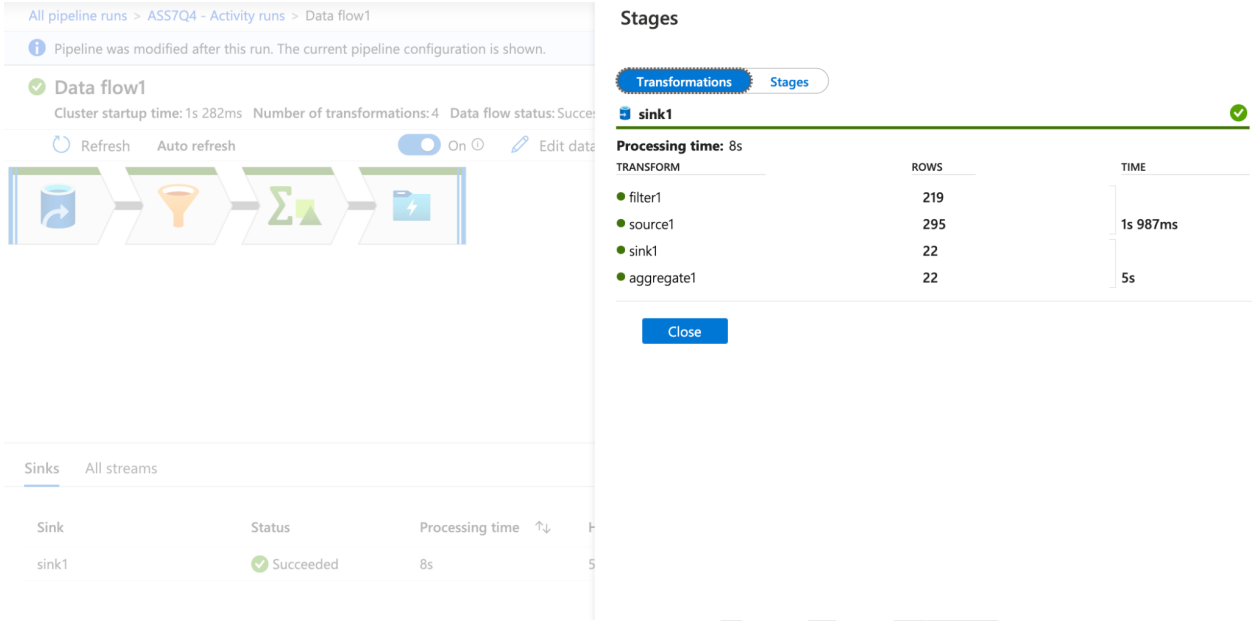
Sink	Status	Processing time	Highest processing time	Rows written	Stages	Lineage
sink1	Succeeded	8s	5s	22		

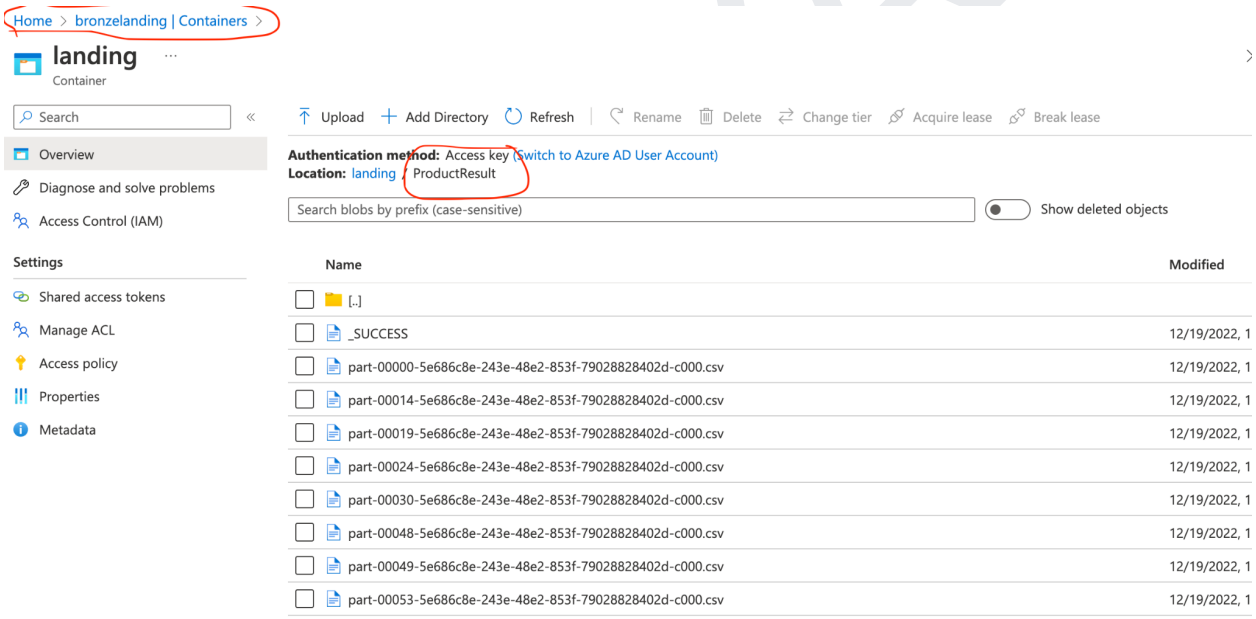
Lineage

sink1

COLUMN	METHOD	ORIGINAL SOURCE
MaximumListPrice	Calculated	source1
		ListPrice
ProductCategoryID	Mapped	source1
		ProductCategoryID
-	Used	source1
		Color

Close

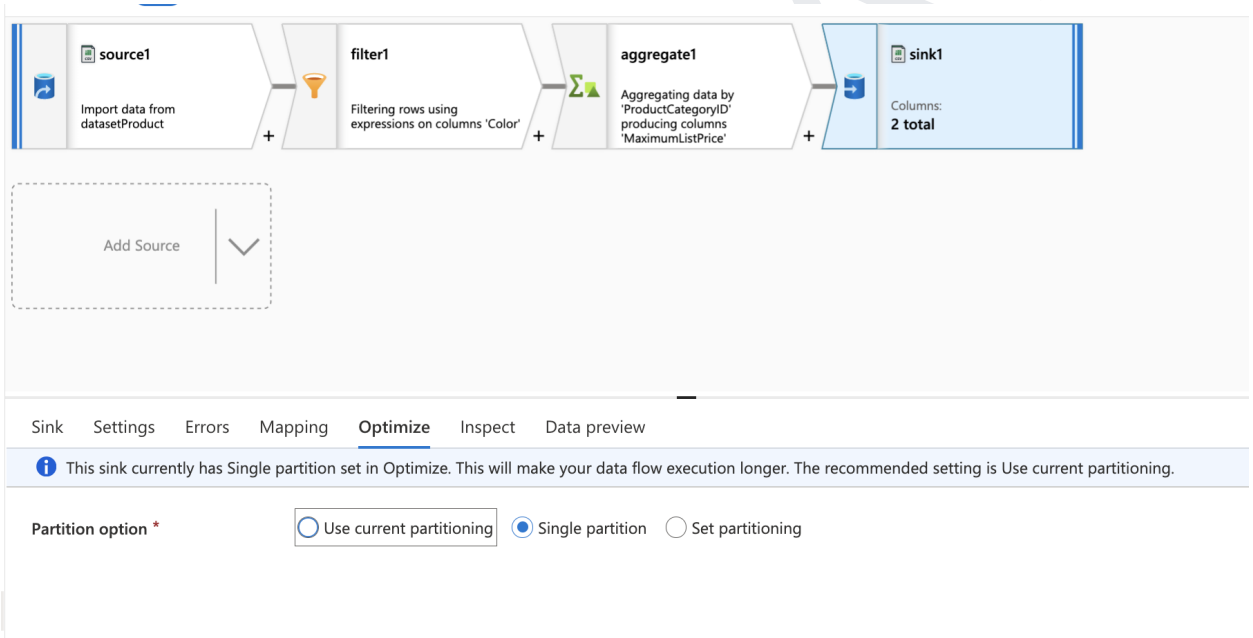




Files created in partitions

create a pipeline to read the Product CSV file, and calculate the highest listPrice of any product under each productcategory. Ensure that product shouldn't be of blue in color and save the result as a SINGLE CSV file inside ProductSingleResult folder.

A:It is the same as above question but instead of saving in multiple partitions save as single partition.



Home > bronzelanding | Containers >

landing

Container

Search

- Overview
- Diagnose and solve problems
- Access Control (IAM)

Settings

- Shared access tokens
- Manage ACL
- Access policy
- Properties
- Metadata

Upload | Add Directory | Refresh | Rename | Delete | Change tier | Acquire lease | Break l

part-00048-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00049-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00053-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00066-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00089-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00095-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00105-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00107-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00122-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00132-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00135-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00143-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:21...	Hot (Inferred)
part-00163-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:22...	Hot (Inferred)
part-00192-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:22...	Hot (Inferred)
part-00198-5e686c8e-243e-48e2-853f-79028828402d-c000.csv	12/19/2022, 10:49:22...	Hot (Inferred)
part-merged.csv	12/19/2022, 11:01:50...	Hot (Inferred)

AMUL

Home > bronzelanding | Containers > landing >

ProductResult/part-merged.csv

Blob

Save Discard Download Refresh Delete

Overview Versions **Edit** Generate SAS

```
ProductCategoryID,MaximumListPrice
6,782.9900
16,364.0900
23,8.9900
39,125.0000
26,69.9900
27,9.5000
12,404.9900
20,333.4200
5,769.4900
17,80.9900
35,34.9900
7,742.3500
25,53.9900
10,106.5000
18,594.8300
28,74.9900
22,89.9900
13,91.4900
```

Csv Preview

For more updates Follow me on LinkedIN: LinkedIn.com/in/amulya1003

AMULYAA