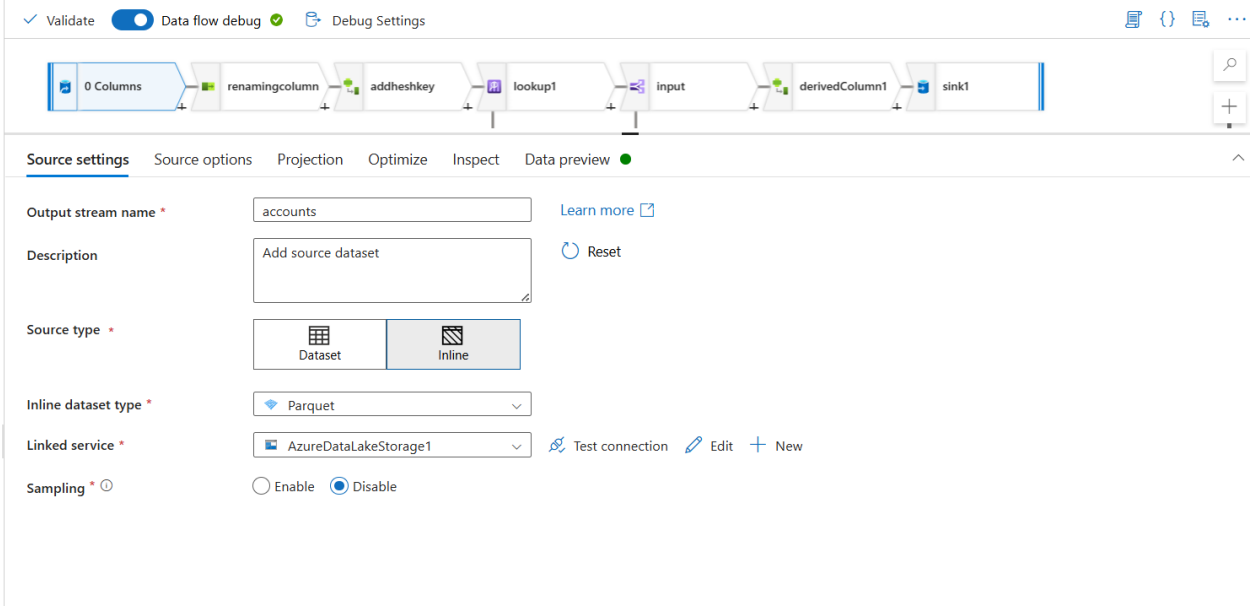
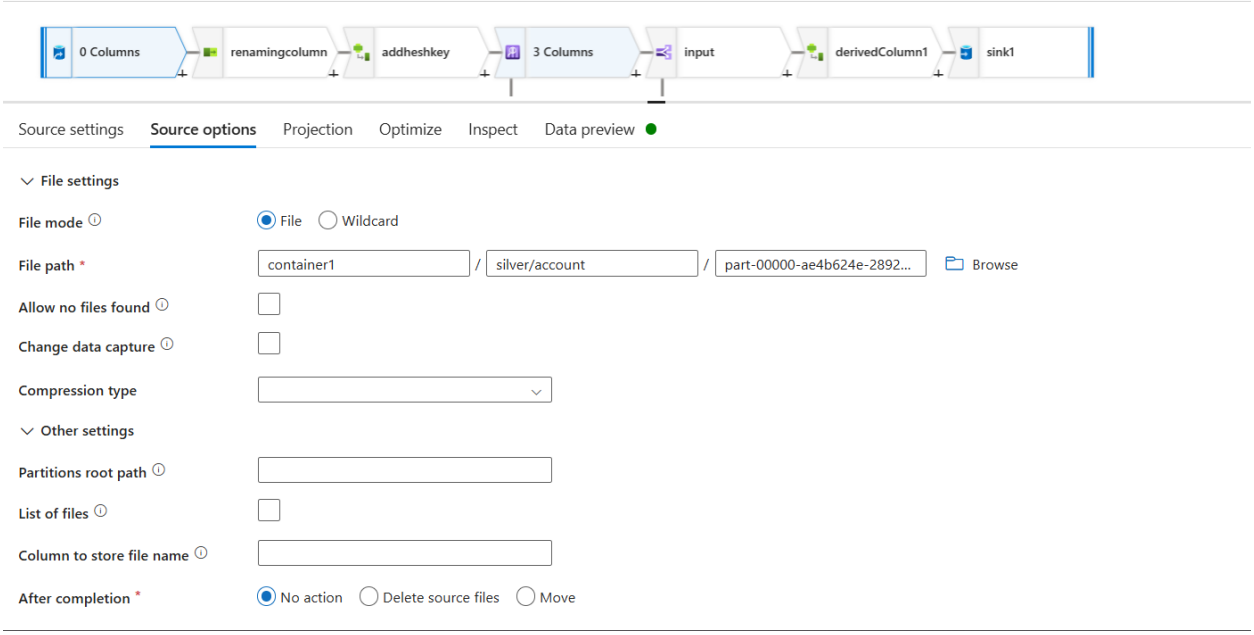
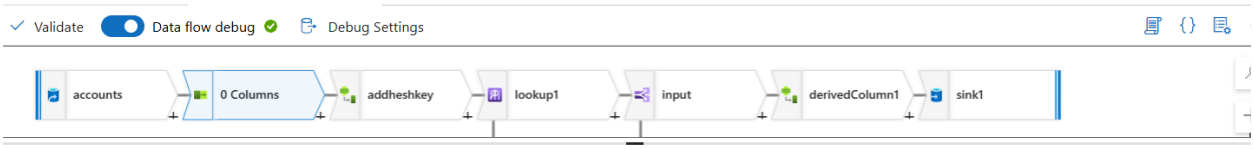
SCD Type -1 & 2 Pipeline Details

Accounts :- SCD-1







Select settings | Optimize | Inspect | Data preview

Output stream name * renamingcolumn — Learn more

Description: Add columns to rename — Reset

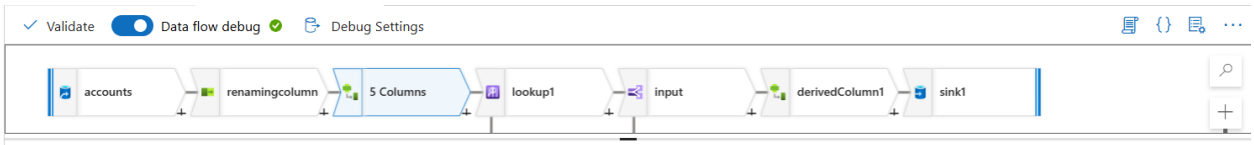
Incoming stream * accounts

Options: Skip duplicate input columns; Skip duplicate output columns

Input columns *

Auto mapping | Reset | Add mapping | Delete | 1 mappings: All inputs mapped

accounts's column	Name as
1==1	concat('scr_',$$)



Derived column's settings | Optimize | Inspect | Data preview

Output stream name * addheshkey — Learn more

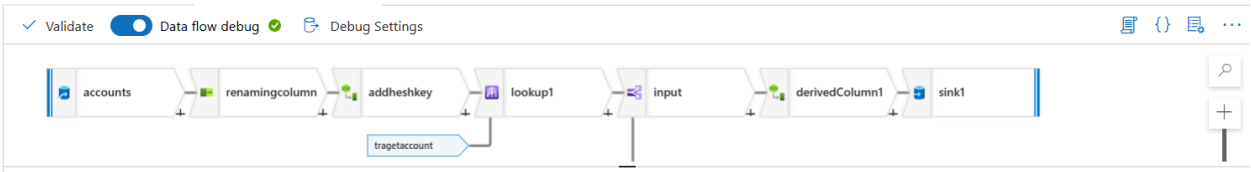
Description: Creating/updating the columns 'scr_account_id, scr_customer_id, scr_account_type, scr_balance, — Reset

Incoming stream * renamingcolumn

Add | Clone | Delete | Open expression builder

Columns *

Column	Expression
src_hashkey	crc32(concat(toString(scr_account_id),toString(sc...



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

Output stream name * tragetaccount — Learn more

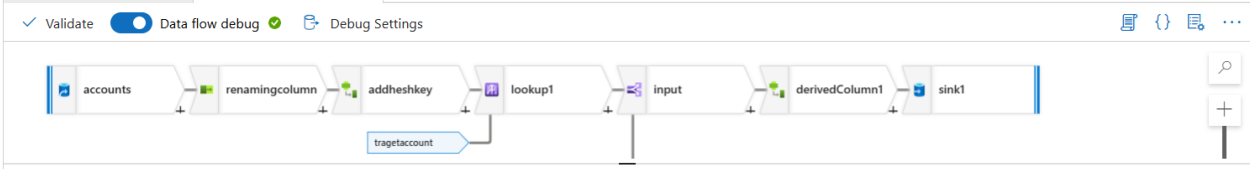
Description: Import data from AzureSqlDatabase1 — Reset

Source type *: Dataset | Inline

Inline dataset type *: Azure SQL Database

Linked service *: AzureSqlDatabase1 — Test connection | Edit | New

Sampling * ⓘ: Enable | Disable



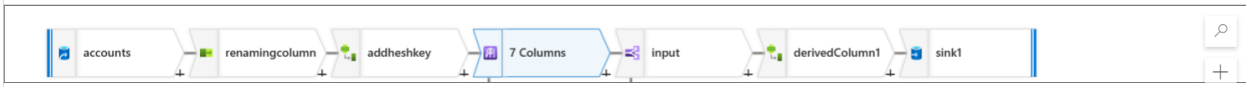
Source settings | Source options | Projection | Optimize | Inspect | Data preview

Input: Table | Query | Stored procedure

Query * ⓘ:
```
select AccountID ,HashKey from [dbo].[Account]
```

Incremental column ⓘ

Isolation level ⓘ: Read uncommitted



Lookup settings | Optimize | Inspect | Data preview

Output stream name * lookup1 Learn more

Description Lookup on 'addheshkey' from 'tragetaccount' Reset

Primary stream * addheshkey

Lookup stream * tragetaccount

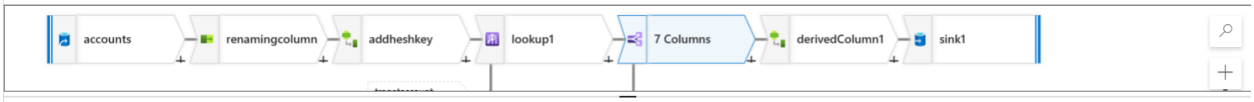
Match multiple rows

Match on * Any row

Lookup conditions *

Left: addheshkey's column		Right: tragetaccount's column
123 scr_account_id	==	123 AccountID

Validate | Data flow debug



Conditional split settings | Optimize | Inspect | Data preview

Output stream name * split1 Learn more

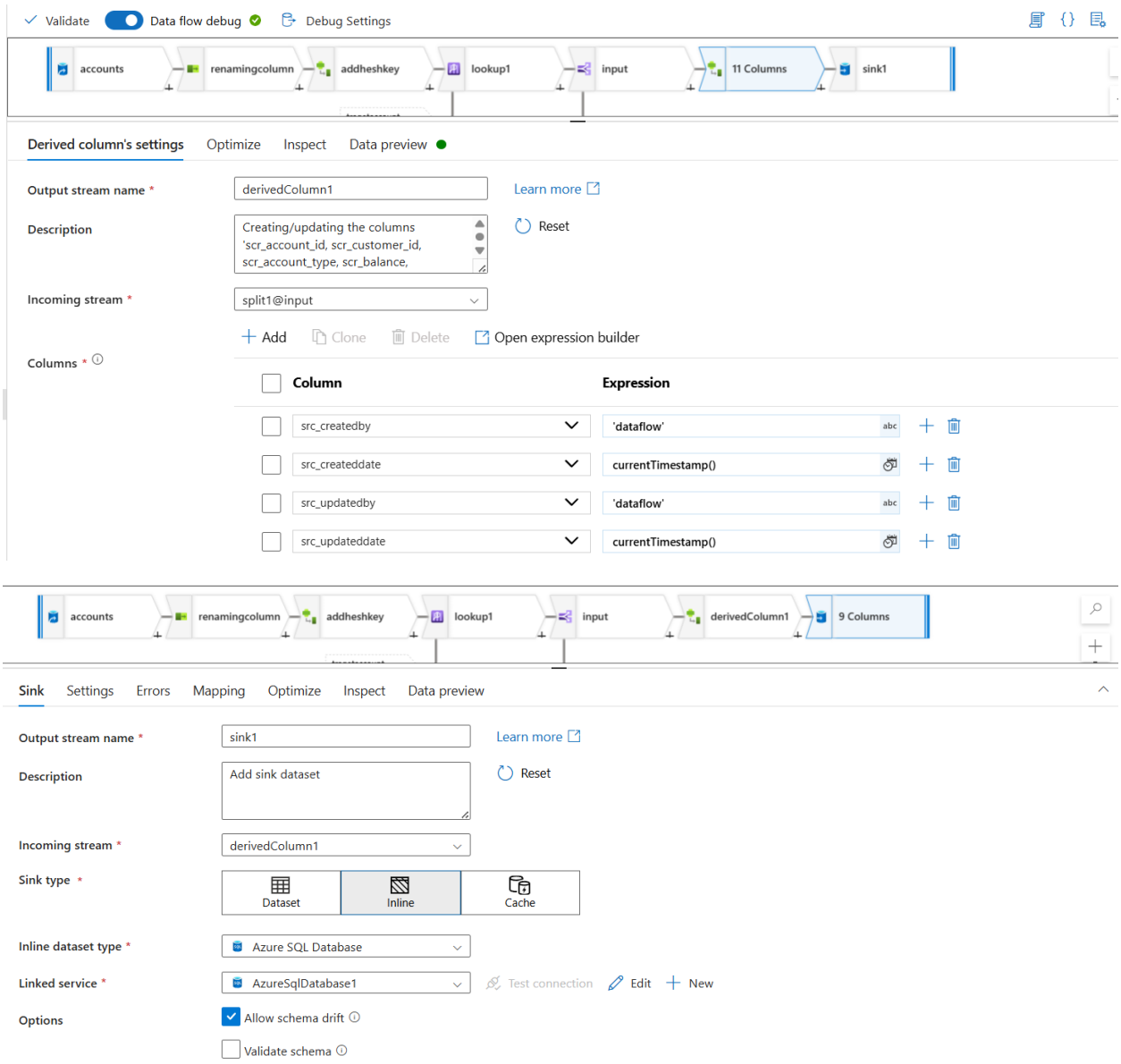
Description Conditionally distributing the data in AccountID, scr_account_id, AccountID, src_hashkey, HashKey groups, based on Reset

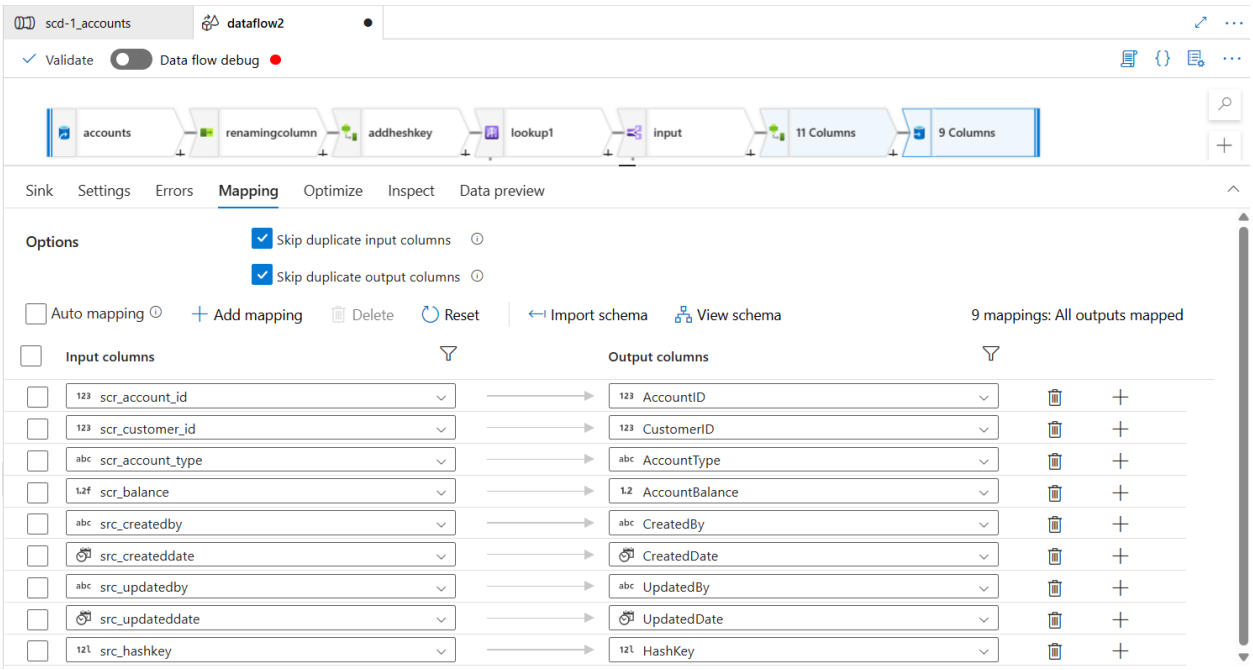
Incoming stream * lookup1

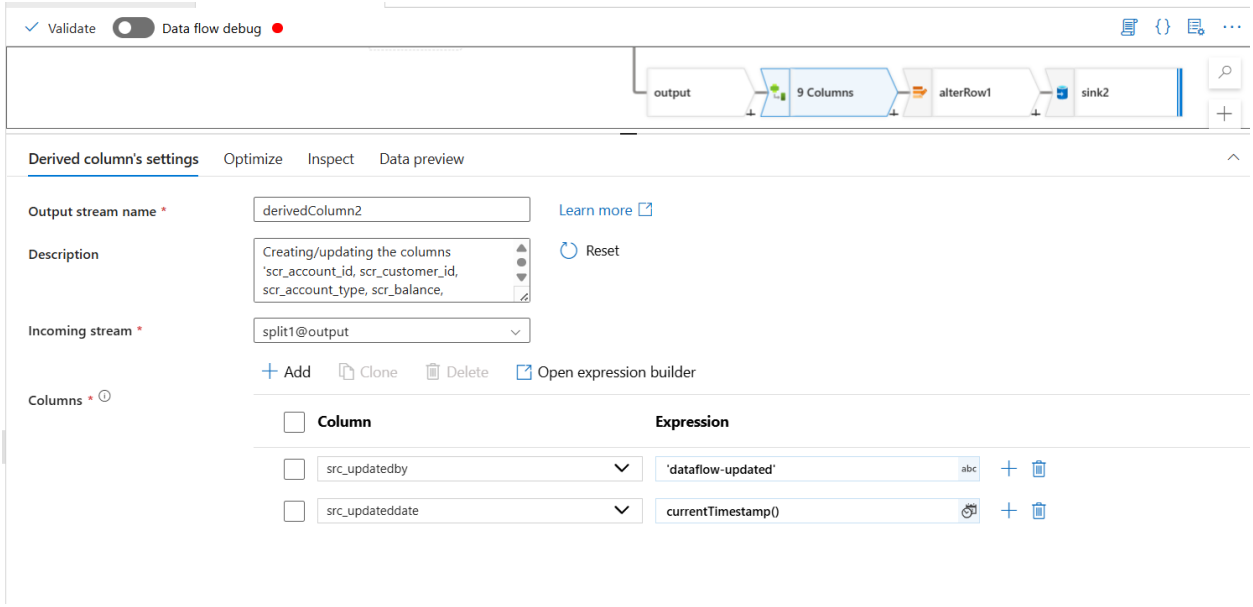
Split on (•) First matching condition ( ) All matching conditions

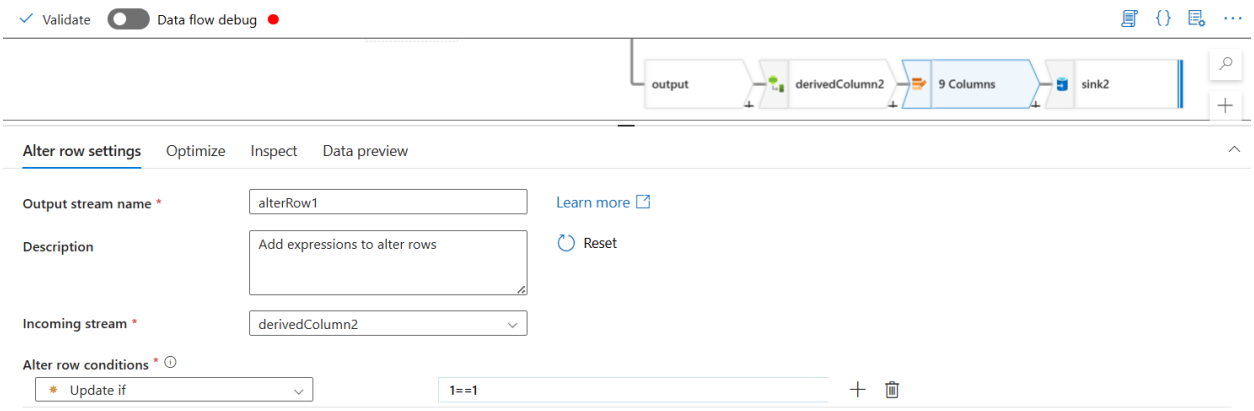
Split condition

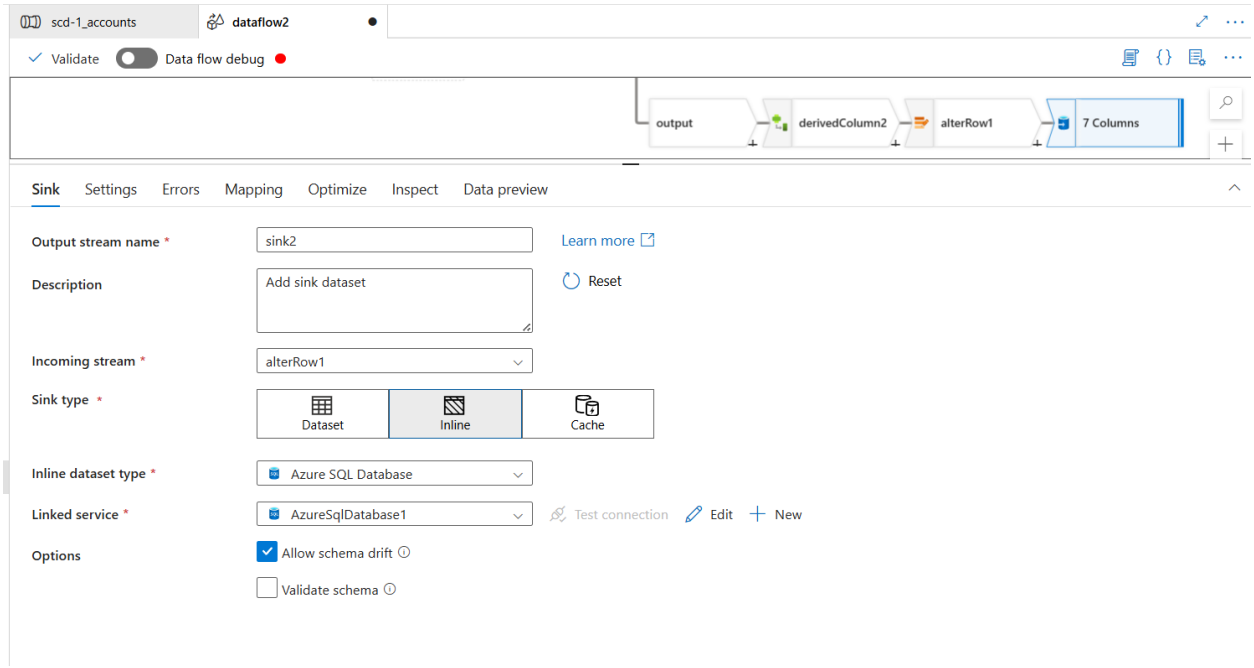
Stream names	Condition
input	isNull(AccountID)
output	scr_account_id==AccountID &&src_hashkey!=HashKey

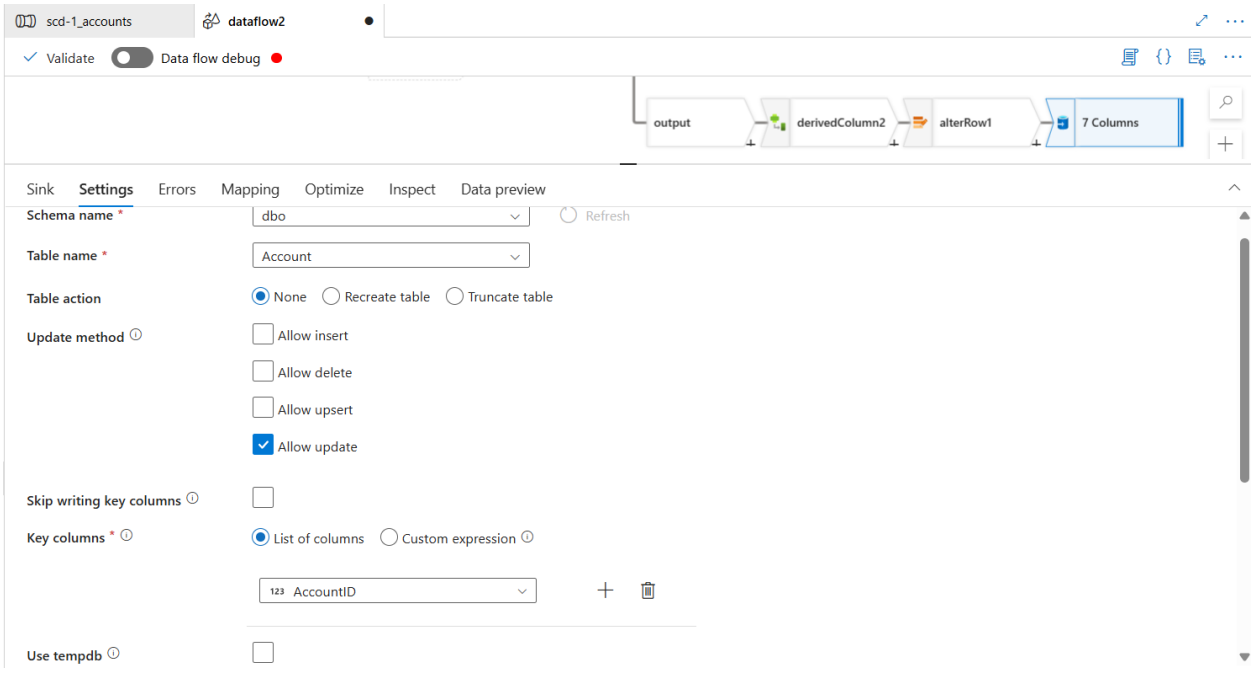

Validate
Data flow debug
Debug Settings
accounts
renamingcolumn
addheshkey
lookup1
input
11 Columns
sink1
Derived column's settings
Optimize
Inspect
Data preview
Output stream name *
derivedColumn1
Learn more
Description
Creating/updating the columns 'scr_account_id, scr_customer_id, scr_account_type, scr_balance,
Reset
Incoming stream *
split1@input
Add
Clone
Delete
Open expression builder
Columns *
Column
Expression
src_createdby
'dataflow'
src_createddate
currentTimestamp()
src_updatedby
'dataflow'
src_updateddate
currentTimestamp()
accounts
renamingcolumn
addheshkey
lookup1
input
derivedColumn1
9 Columns
Sink
Settings
Errors
Mapping
Optimize
Inspect
Data preview
Output stream name *
sink1
Learn more
Description
Add sink dataset
Reset
Incoming stream *
derivedColumn1
Sink type *
Dataset
Inline
Cache
Inline dataset type *
Azure SQL Database
Linked service *
AzureSqlDatabase1
Test connection
Edit
New
Options
Allow schema drift
Validate schema

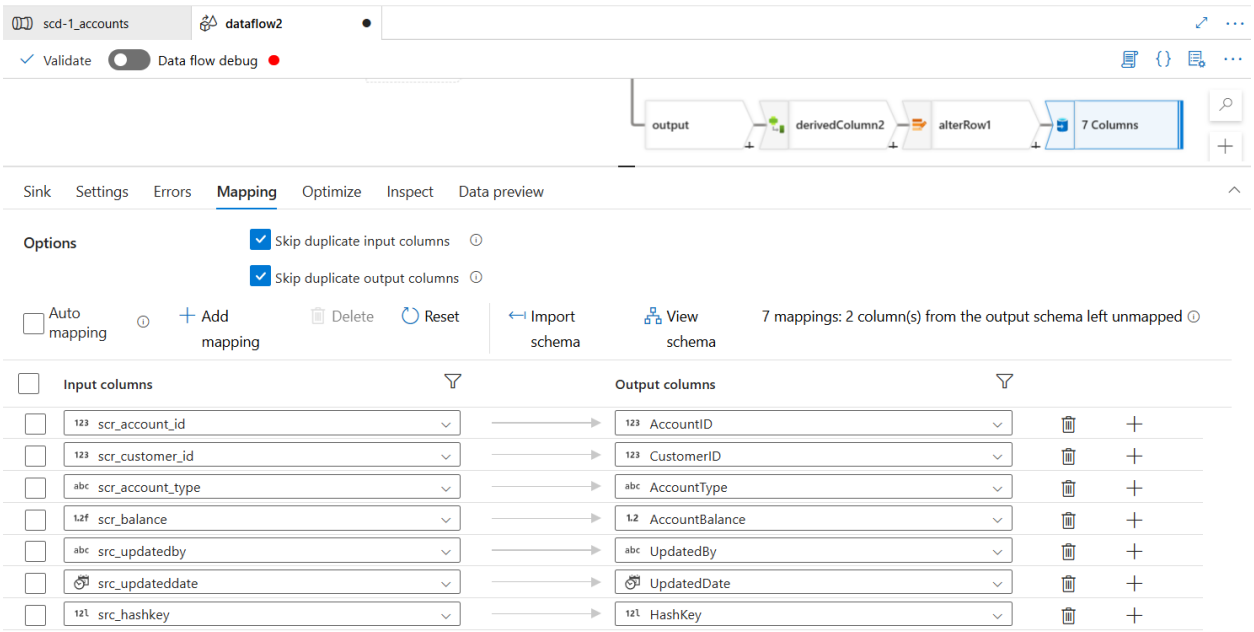

scd-1_accounts
dataflow2
Validate
Data flow debug
accounts
renamingcolumn
addheshkey
lookup1
input
11 Columns
9 Columns
Sink
Settings
Errors
Mapping
Optimize
Inspect
Data preview
Options
Skip duplicate input columns
Skip duplicate output columns
Auto mapping
Add mapping
Delete
Reset
Import schema
View schema
9 mappings: All outputs mapped
Input columns
Output columns
scr_account_id
AccountID
scr_customer_id
CustomerID
scr_account_type
AccountType
scr_balance
AccountBalance
src_createdby
CreatedBy
src_createddate
CreatedDate
src_updatedby
UpdatedBy
src_updateddate
UpdatedDate
src_hashkey
HashKey


Validate
Data flow debug
output
9 Columns
alterRow1
sink2
Derived column's settings
Optimize
Inspect
Data preview
Output stream name *
derivedColumn2
Learn more
Description
Creating/updating the columns 'scr_account_id, scr_customer_id, scr_account_type, scr_balance,
Reset
Incoming stream *
split1@output
Add
Clone
Delete
Open expression builder
Columns *
Column
Expression
src_updatedby
'dataflow-updated'
src_updateddate
currentTimestamp()

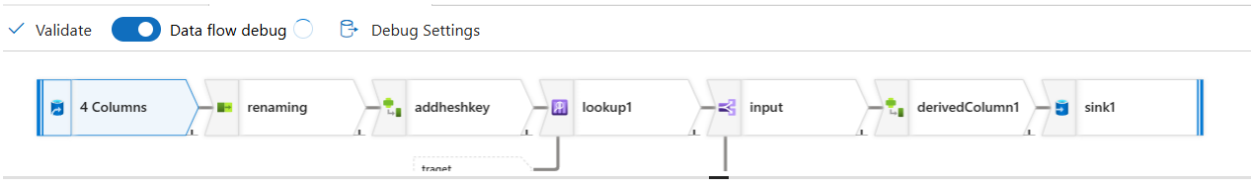

Validate
Data flow debug
output
derivedColumn2
9 Columns
sink2
Alter row settings
Optimize
Inspect
Data preview
Output stream name *
alterRow1
Learn more
Description
Add expressions to alter rows
Reset
Incoming stream *
derivedColumn2
Alter row conditions *
Update if
1==1

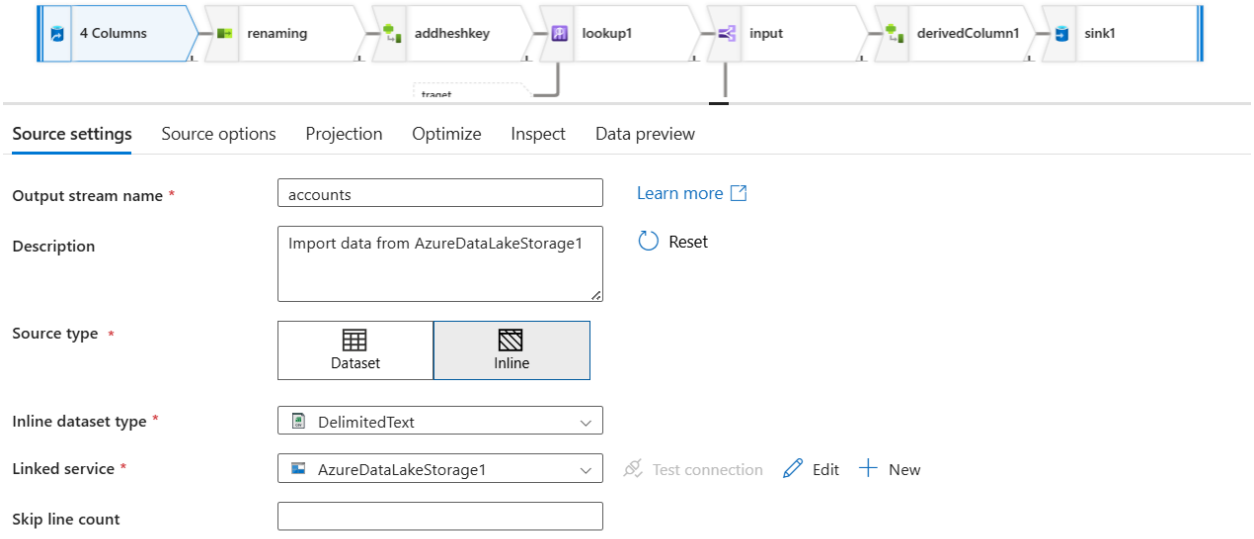

scd-1_accounts
dataflow2
Validate
Data flow debug
output
derivedColumn2
alterRow1
7 Columns
Sink
Settings
Errors
Mapping
Optimize
Inspect
Data preview
Output stream name *
sink2
Learn more
Description
Add sink dataset
Reset
Incoming stream *
alterRow1
Sink type *
Dataset
Inline
Cache
Inline dataset type *
Azure SQL Database
Linked service *
AzureSqlDatabase1
Test connection
Edit
New
Options
Allow schema drift
Validate schema

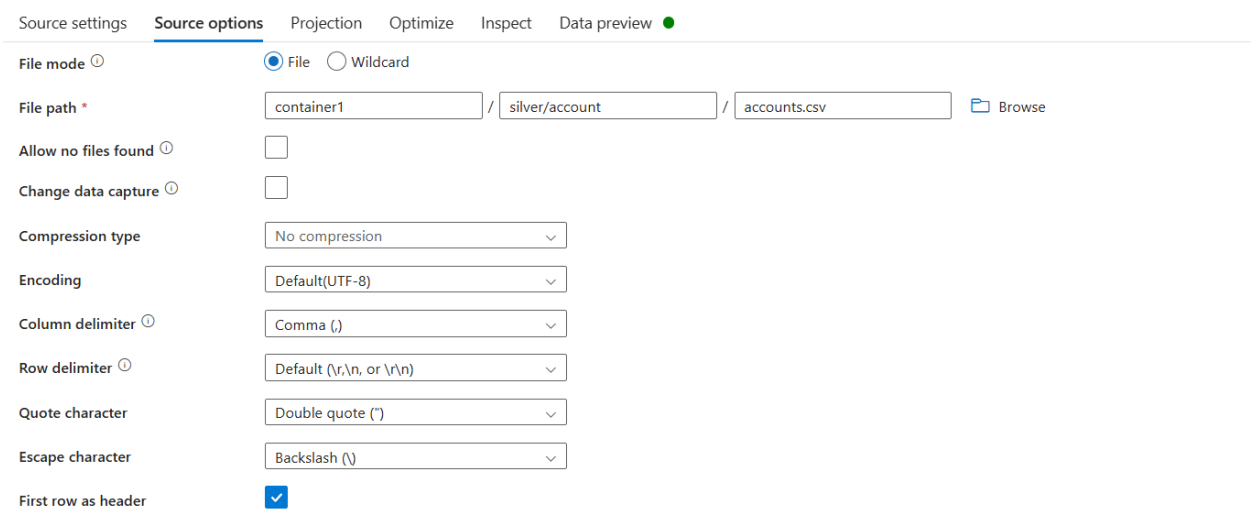




After adding new file which have updated data and new data

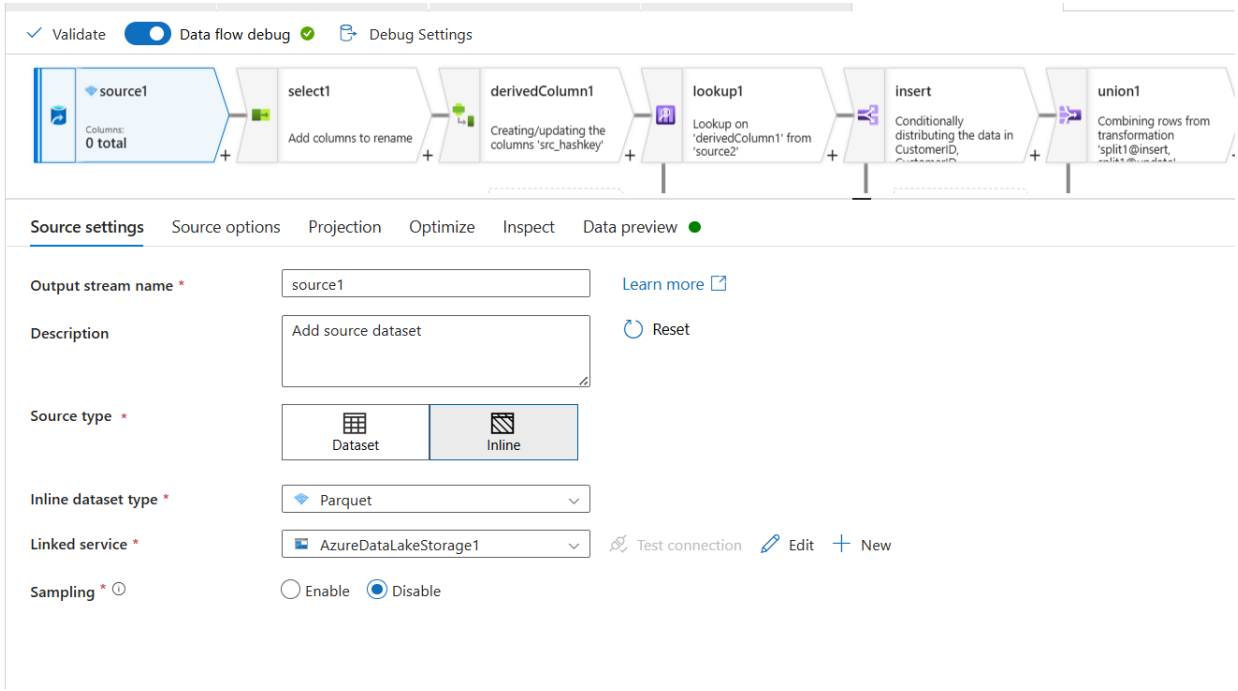


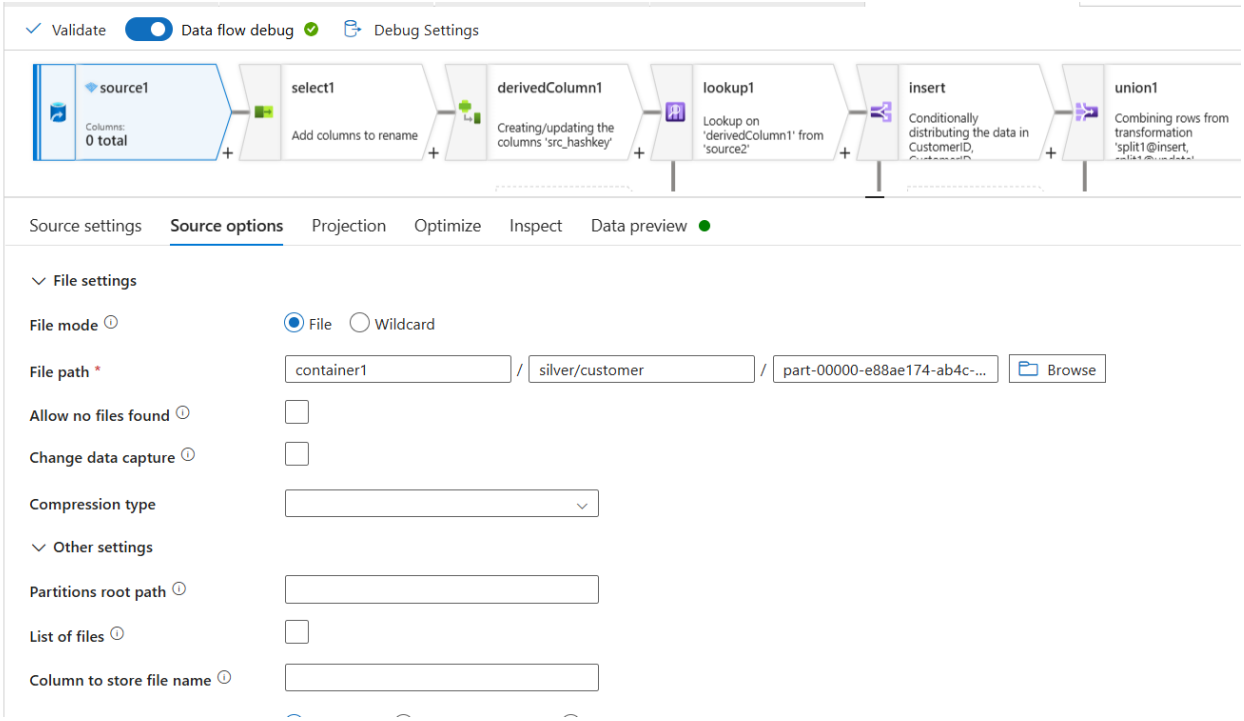




Similar Logic for Transaction data as well.

Customer:- SCDType-2





Select settings | Optimize | Inspect | Data preview

Output stream name * select1 Learn more

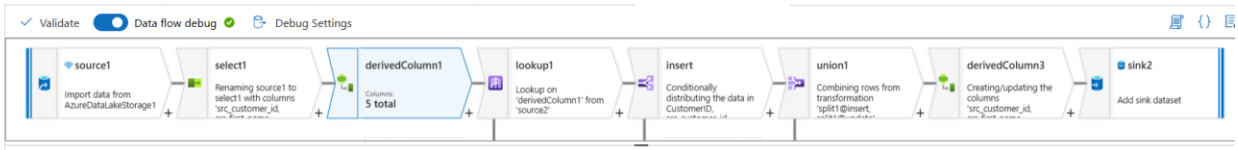
Description: Renaming source1 to select1 with columns 'src_customer_id, src_first_name, src_last_name, src_full_address' Reset

Incoming stream * source1

Options: Skip duplicate input columns; Skip duplicate output columns

Input columns * Auto mapping Reset Add mapping Delete 1 mappings: All inputs ma

source1's column	Name as
1==1	concat('src_',$$)



Derived column's settings | Optimize | Inspect | Data preview

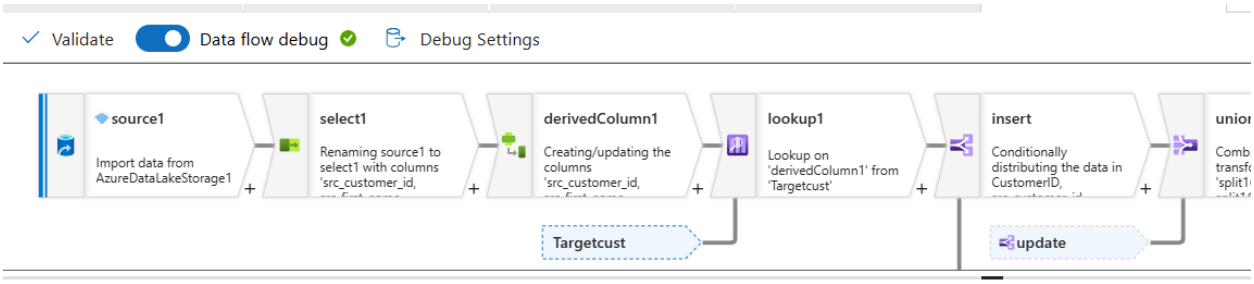
Output stream name * derivedColumn1 Learn more

Description: Creating/updating the columns 'src_customer_id, src_first_name, src_last_name, src_full_address, Reset

Incoming stream * select1

Columns * Add Clone Delete Open expression builder

Column	Expression
src_hashkey	crc32(concat(toString(src_customer_id),src_first_...



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

Output stream name * Targetcust — Learn more

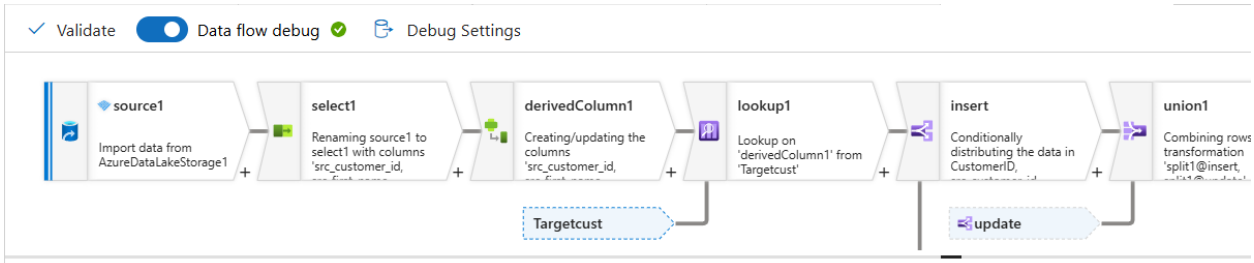
Description Import data from AzureSqlDatabase1 — Reset

Source type * Dataset | Inline

Inline dataset type * Azure SQL Database

Linked service * AzureSqlDatabase1 — Test connection | Edit | New

Sampling * Enable | Disable



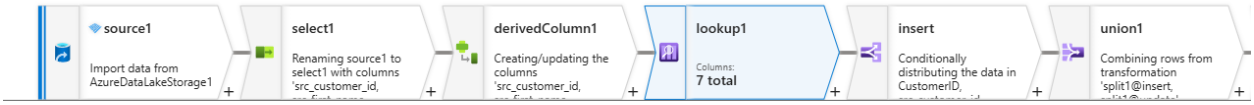
Source settings | Source options | Projection | Optimize | Inspect | Data preview

Input Table | Query | Stored procedure

Query * select CustomerID, HashKey from Customers where IsActive=1

Incremental column

Isolation level Read uncommitted



Lookup settings | Optimize | Inspect | Data preview

Output stream name * lookup1 — Learn more

Description: Lookup on 'derivedColumn1' from 'Targetcust' — Reset

Primary stream * derivedColumn1

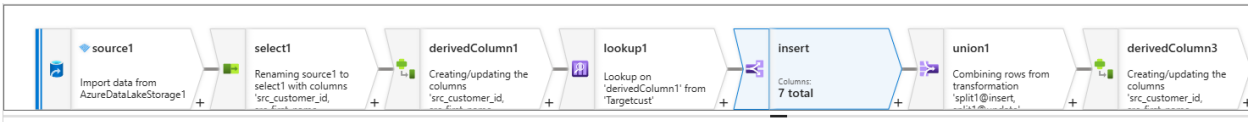
Lookup stream * Targetcust

Match multiple rows

Match on * Any row

Lookup conditions *

Left: derivedColumn1's column		Right: Targetcust's column
src_customer_id	==	CustomerID



Conditional split settings | Optimize | Inspect | Data preview

Output stream name * split1 — Learn more

Description: Conditionally distributing the data in CustomerID, src_customer_id, CustomerID, src_hashkey, HashKey — Reset

Incoming stream * lookup1

Split on: (•) First matching condition ( ) All matching conditions

Split condition

Stream names	Condition
insert	isNull(CustomerID)
update	src_customer_id==CustomerID&&src_hashkey!=HashKey

Derived column's settings | Optimize | Inspect | Data preview

Output stream name * derivedColumn2 Learn more

Description Creating/updating the columns 'src_customer_id, src_first_name, src_last_name, src_full_address, Reset

Incoming stream * split1@update

Add | Clone | Delete | Open expression builder

Columns *

Column	Expression
src_updatedby	'dataflow-updated'
src_updateddate	currentTimestamp()
src_isactive	0

Alter row settings | Optimize | Inspect | Data preview

Output stream name * alterRow1 Learn more

Description Add expressions to alter rows Reset

Incoming stream * derivedColumn2

Alter row conditions *

Update if 1==1

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

Output stream name * : sink1 — Learn more

Description : Add sink dataset — Reset

Incoming stream * : alterRow1

Sink type * : Dataset | Inline | Cache

Inline dataset type * : Azure SQL Database

Linked service * : AzureSqlDatabase1 — Test connection | Edit | New

Options : Allow schema drift ⓘ

Validate schema ⓘ

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

Schema name * : dbo — Refresh

Table name * : Customers

Table action : None | Recreate table | Truncate table

Update method ⓘ : Allow insert

Allow delete

Allow upsert

Allow update

Skip writing key columns ⓘ

Key columns * ⓘ : List of columns | Custom expression ⓘ

123 CustomerID

123 HashKey

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

Options : Skip duplicate input columns ⓘ

Skip duplicate output columns ⓘ

Auto mapping ⓘ | Add mapping | Delete | Reset | Import schema | View schema

5 mappings: 4 column(s) from the output schema left unmapped

Input columns	Output columns
CustomerID	CustomerID
src_updatedby	UpdatedBy
src_updateddate	UpdatedDate
HashKey	HashKey
src_isactive	IsActive

Union settings | Optimize | Inspect | Data preview

Output stream name * union1 Learn more

Description Combining rows from transformation 'split1@insert, split1@update' Reset

Incoming stream * split1@insert

Union by * Name Position

Union with * split1@update

Derived column's settings | Optimize | Inspect | Data preview

Output stream name * derivedColumn3 Learn more

Description Creating/updating the columns 'src_customer_id, src_first_name, src_last_name, src_full_address, Reset

Incoming stream * union1

Add | Clone | Delete | Open expression builder

Columns *

Column	Expression
src_createdby	'dataflow'
src_createddate	currentTimestamp()
src_updatedby	'dataflow'
src_updateddate	currentTimestamp()
src_isactive	1

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

Output stream name * sink2 Learn more

Description Add sink dataset Reset

Incoming stream * derivedColumn3

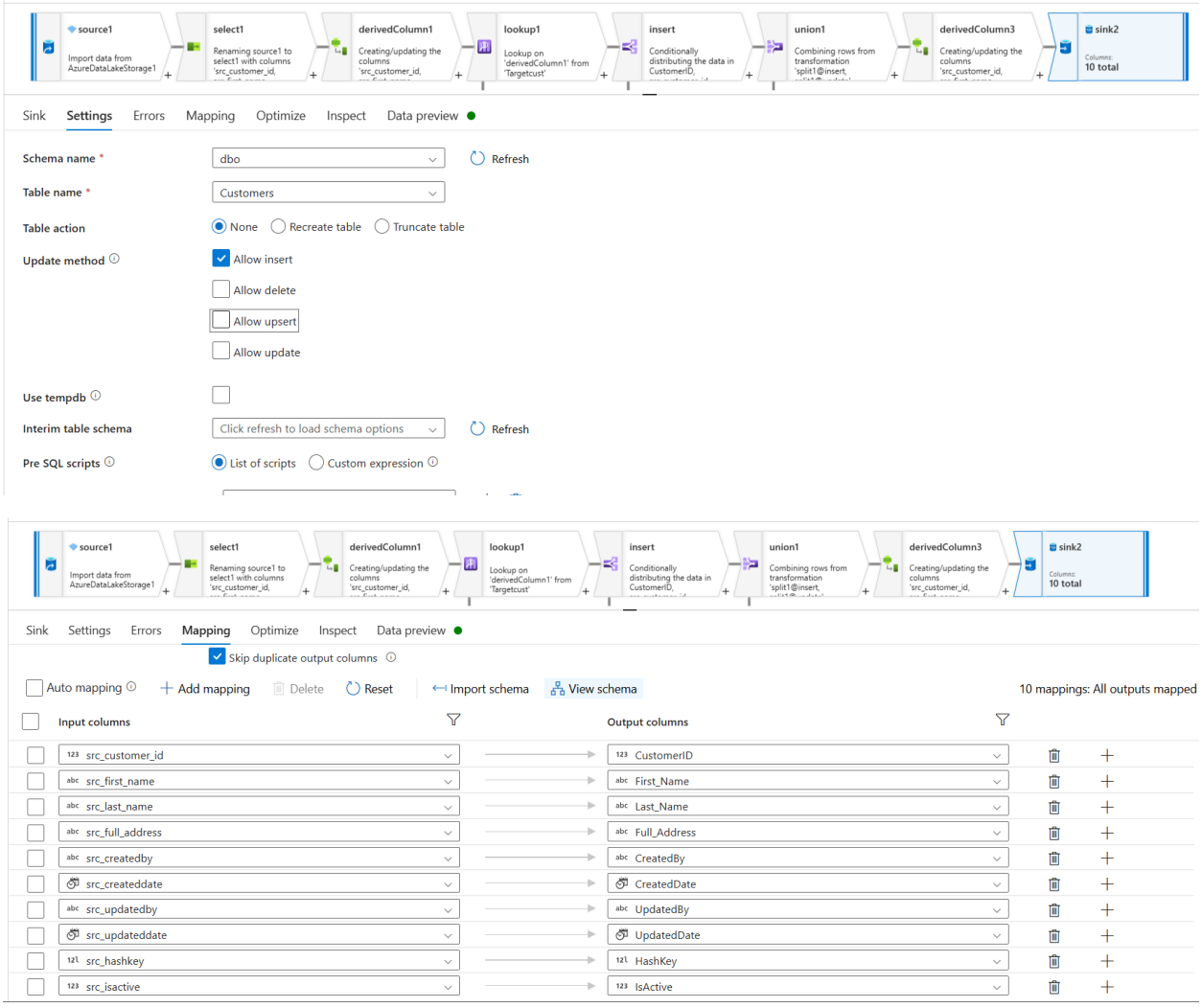
Sink type * Dataset | Inline | Cache

Inline dataset type * Azure SQL Database

Linked service * AzureSqlDatabase1 Test connection Edit New

Options Allow schema drift

Validate schema



Similar logics for other files.