

✓ Validate

Data flow debug

✓

Debug Settings

accounts

0 Columns

addheshkey

lookup1

input

derivedColumn1

sink1

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

1==1

Name as

concat('scr_',\$\$)

✓ Validate

Data flow debug

✓

Debug Settings

accounts

renamingcolumn

5 Columns

lookup1

input

derivedColumn1

sink1

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

Columns *

 ⓘ

+ Add

Clone

Delete

Open expression builder

Column

Expression

src_hashkey

crc32(concat(toString(scr_account_id),toString(sc... 121)

✓ Validate

Data flow debug

Debug Settings

accounts

renamingcolumn

addheshkey

lookup1

input

derivedColumn1

sink1

tragetaccount

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

✓ Validate

Data flow debug

Debug Settings

accounts

renamingcolumn

addheshkey

lookup1

input

derivedColumn1

sink1

tragetaccount

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

Output stream name * [Learn more](#)

Description [Reset](#)

Primary stream *

Lookup stream *

Match multiple rows ☐ [?](#)

Match on *

Lookup conditions *

Left: addheshkey's column		Right: tragetaccount's column	
<input type="text" value="123 scr_account_id"/>	<input type="text" value="=="/>	<input type="text" value="123 AccountID"/>	+ -

✓ Validate ☐ Data flow debug



Conditional split settings

Output stream name * [Learn more](#)

Description [Reset](#)

Incoming stream *

Split on ☒ First matching condition ☐ All matching conditions

Split condition

Stream names	Condition	
<input type="text" value="input"/>	<input type="text" value="isNull(AccountID)"/>	+ -
<input type="text" value="output"/>	<input type="text" value="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 * ①

<input type="checkbox"/>	Column	Expression		
<input type="checkbox"/>	src_createdby	'dataflow'	abc	+ 🗑
<input type="checkbox"/>	src_createddate	currentTimestamp()		+ 🗑
<input type="checkbox"/>	src_updatedby	'dataflow'	abc	+ 🗑
<input type="checkbox"/>	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 ①

☐ Validate schema ①



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		
<input type="checkbox"/> 123 scr_account_id	→	<input type="checkbox"/> 123 AccountID	Delete	+
<input type="checkbox"/> 123 scr_customer_id	→	<input type="checkbox"/> 123 CustomerID	Delete	+
<input type="checkbox"/> abc scr_account_type	→	<input type="checkbox"/> abc AccountType	Delete	+
<input type="checkbox"/> 12f scr_balance	→	<input type="checkbox"/> 1.2 AccountBalance	Delete	+
<input type="checkbox"/> abc src_createdby	→	<input type="checkbox"/> abc CreatedBy	Delete	+
<input type="checkbox"/> src_createddate	→	<input type="checkbox"/> CreatedDate	Delete	+
<input type="checkbox"/> abc src_updatedby	→	<input type="checkbox"/> abc UpdatedBy	Delete	+
<input type="checkbox"/> src_updateddate	→	<input type="checkbox"/> UpdatedDate	Delete	+
<input type="checkbox"/> 12l src_hashkey	→	<input type="checkbox"/> 12l HashKey	Delete	+



Derived column's settings Optimize Inspect Data preview

Output stream name * [Learn more](#)

Description [Reset](#)

Incoming stream *

[+ Add](#) [Clone](#) [Delete](#) [Open expression builder](#)

Column	Expression
<input type="checkbox"/> src_updatedby	<input type="text" value="'dataflow-updated'"/> abc + Delete
<input type="checkbox"/> src_updateddate	<input type="text" value="currentTimestamp()"/> + Delete

✓ 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 ⓘ

scd-1_accounts
dataflow2

Validate
Data flow debug

output
derivedColumn2
alterRow1
7 Columns

Sink
Settings
Errors
Mapping
Optimize
Inspect
Data preview

Schema name *
dbo
Refresh

Table name *
Account

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 AccountID

Use tempdb ⓘ

scd-1_accounts
dataflow2

Validate
Data flow debug

output
derivedColumn2
alterRow1
7 Columns

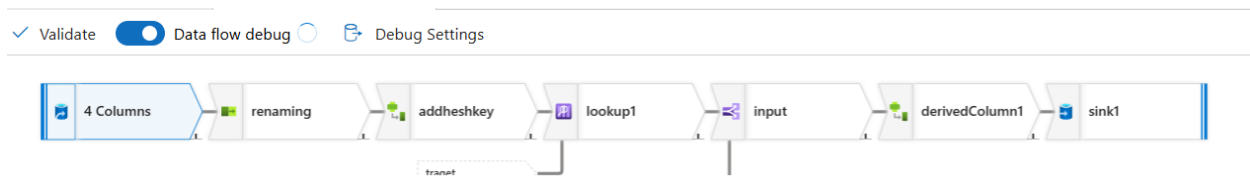
Sink
Settings
Errors
Mapping
Optimize
Inspect
Data preview

Options
Skip duplicate input columns ⓘ
Skip duplicate output columns ⓘ

Auto mapping ⓘ
Add mapping
Delete
Reset
Import schema
View schema
7 mappings: 2 column(s) from the output schema left unmapped ⓘ

Input columns	Output columns
123 scr_account_id	123 AccountID
123 scr_customer_id	123 CustomerID
abc scr_account_type	abc AccountType
12f scr_balance	1.2 AccountBalance
abc src_updatedby	abc UpdatedBy
src_updateddate	UpdatedDate
12l src_hashkey	12l HashKey

After adding new file which have updated data and new data



Source settings Source options Projection Optimize Inspect Data preview

Output stream name * [Learn more](#)

Description [Reset](#)

Source type * ☐ Dataset ☒ Inline

Inline dataset type *

Linked service * [Test connection](#) [Edit](#) [New](#)

Skip line count

Sampling * ☐ Enable ☒ Disable



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

File mode ☒ File ☐ Wildcard

File path * / / [Browse](#)

Allow no files found ☐

Change data capture ☐

Compression type

Encoding

Column delimiter

Row delimiter

Quote character

Escape character

First row as header ☒

Similar Logic for Transaction data as well.

Customer:- SCDType-2

✓ Validate ☒ Data flow debug ✓ Debug Settings

source1
Columns:
0 total

select1
Add columns to rename

derivedColumn1
Creating/updating the columns 'src_hashkey'

lookup1
Lookup on 'derivedColumn1' from 'source2'

insert
Conditionally distributing the data in CustomerID, CustomerName

union1
Combining rows from transformation 'split1@insert, split1@insert'

Source settings

Source options

Projection

Optimize

Inspect

Data preview

Output stream name *

source1

Learn more

Description

Add source dataset

Reset

Source type *

Dataset

Inline

Inline dataset type *

Parquet

Linked service *

AzureDataLakeStorage1

Test connection

Edit

New

Sampling *

Enable

Disable

✓ Validate ☒ Data flow debug ✓ Debug Settings

source1
Columns:
0 total

select1
Add columns to rename

derivedColumn1
Creating/updating the columns 'src_hashkey'

lookup1
Lookup on 'derivedColumn1' from 'source2'

insert
Conditionally distributing the data in CustomerID, CustomerName

union1
Combining rows from transformation 'split1@insert, split1@insert'

Source settings

Source options

Projection

Optimize

Inspect

Data preview

File settings

File mode

☒ File ☐ Wildcard

File path *

container1

/

silver/customer

/

part-00000-e88ae174-ab4c-...

Browse

Allow no files found

☐

Change data capture

☐

Compression type

Other settings

Partitions root path

List of files

☐

Column to store file name

source1

Import data from AzureDataLakeStorage1

+

select1

Columns: 4 total

+

derivedColumn1

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

+

lookup1

Lookup on 'derivedColumn1' from 'source2'

+

insert

Conditionally distributing the data in CustomerID, src_customer_id

+

union1

Combining rows from transformation 'split1@insert, split1@insert'

+

derivedColumn3

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

+

sink2

Add sink dataset

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 mapped

☐ source1's column

☐ 1=1

Filter

Name as

concat('src',\$\$)

abc

+

Delete

...

Validate

Data flow debug

Debug Settings

source1

Import data from AzureDataLakeStorage1

+

select1

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

+

derivedColumn1

Columns: 5 total

+

lookup1

Lookup on 'derivedColumn1' from 'source2'

+

insert

Conditionally distributing the data in CustomerID, src_customer_id

+

union1

Combining rows from transformation 'split1@insert, split1@insert'

+

derivedColumn3

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

+

sink2

Add sink dataset

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

src_hashkey

▼

Expression

crc32(concat(toString(src_customer_id),src_first_name,src_last_name,src_full_address))

123

+

Delete

✓ Validate

Data flow debug

✓

Debug Settings

source1

Import data from AzureDataLakeStorage1

+

select1

Renaming source1 to select1 with columns 'src_customer_id,'

+

derivedColumn1

Creating/updating the columns 'src_customer_id,'

+

lookup1

Lookup on 'derivedColumn1' from 'Targetcust'

+

insert

Conditionally distributing the data in CustomerID,

+

union

Comb transfo 'split1'

Targetcust

update

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

✓ Validate

Data flow debug

✓

Debug Settings

source1

Import data from AzureDataLakeStorage1

+

select1

Renaming source1 to select1 with columns 'src_customer_id,'

+

derivedColumn1

Creating/updating the columns 'src_customer_id,'

+

lookup1

Lookup on 'derivedColumn1' from 'Targetcust'

+

insert

Conditionally distributing the data in CustomerID,

+

union1

Combining rows transformation 'split1@insert,'

Targetcust

update

Source settings

Source options

Projection

Optimize

Inspect

Data preview

Input

Table

Query

Stored procedure

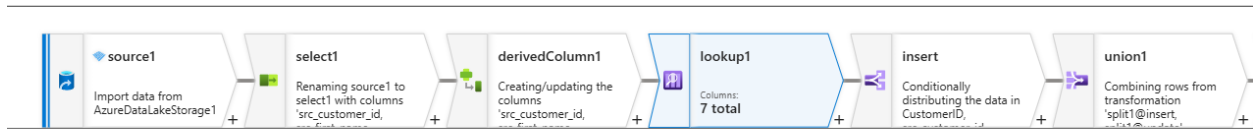
Query * ⓘ

select CustomerID, HashKey from Customers where IsActive=1

Incremental column ⓘ

Isolation level ⓘ

Read uncommitted



Lookup settings Optimize Inspect Data preview ●

Output stream name * [Learn more](#)

Description [Reset](#)

Primary stream *

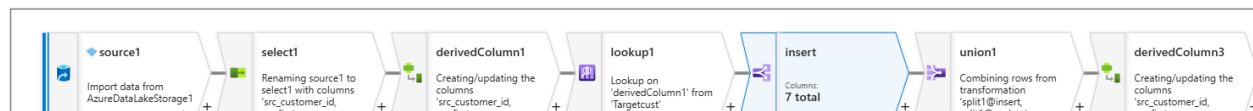
Lookup stream *

Match multiple rows ☐

Match on *

Lookup conditions *

Left: derivedColumn1's column		Right: Targetcust's column
<input type="text" value="123 src_customer_id"/>	<input type="text" value="=="/>	<input type="text" value="123 CustomerID"/>



Conditional split settings Optimize Inspect Data preview ●

Output stream name * [Learn more](#)

Description [Reset](#)

Incoming stream *

Split on ☒ First matching condition ☐ All matching conditions

Split condition

Stream names	Condition
<input type="text" value="insert"/>	<input type="text" value="isNull(CustomerID)"/>
<input type="text" value="update"/>	<input type="text" value="src_customer_id!=CustomerID&src_hashkey!=HashKey"/>

✓ Validate

Data flow debug

🔗 Debug Settings

update

Conditionally distributing the data in CustomerID

+

derivedColumn2

Columns: 10 total

+

alterRow1

Add expressions to alter rows

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

Columns *

+ Add

📄 Clone

🗑 Delete

🔗 Open expression builder

<input type="checkbox"/>	Column	Expression
<input type="checkbox"/>	src_updatedby	'dataflow-updated'
<input type="checkbox"/>	src_updateddate	currentTimestamp()
<input type="checkbox"/>	src_isactive	0

update

Conditionally distributing the data in CustomerID

+

derivedColumn2

Creating/updating the columns 'src_customer_id,

+

alterRow1

Columns: 10 total

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

update
Conditionally distributing the data in CustomerID...

derivedColumn2
Creating/updating the columns 'src_customer_id'...

alterRow1
Add expressions to alter rows

sink1
Columns: 5 total

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

Reference: 2

derivedColumn2
Creating/updating the

alterRow1

sink1

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

+

Reference: 2

derivedColumn2
Creating/updating the

alterRow1

sink1

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

123 CustomerID

abc src_updatedby

src_updateddate

121 HashKey

123 src_isactive

Output columns

123 CustomerID

abc UpdatedBy

UpdatedDate

121 HashKey

123 IsActive



Union settings Optimize Inspect Data preview ●

Output stream name * [Learn more](#)

Description [Reset](#)

Incoming stream *

Union by * ☒ Name ☐ Position

Union with * [+](#) [-](#)



Derived column's settings Optimize Inspect Data preview ●

Output stream name * [Learn more](#)

Description [Reset](#)

Incoming stream *

[+](#) Add [Clone](#) [Delete](#) [Open expression builder](#)

<input type="checkbox"/>	Column	Expression		
<input type="checkbox"/>	src_createdby	'dataflow'	abc	+ -
<input type="checkbox"/>	src_createddate	currentTimestamp0		+ -
<input type="checkbox"/>	src_updatedby	'dataflow'	abc	+ -
<input type="checkbox"/>	src_updateddate	currentTimestamp0		+ -
<input type="checkbox"/>	src_isactive	1	123	+ -



Sink Settings Errors Mapping Optimize Inspect Data preview ●

Output stream name * [Learn more](#)

Description [Reset](#)

Incoming stream *

Sink type * ☐ Dataset ☒ Inline ☐ Cache

Inline dataset type *

Linked service * [Test connection](#) [Edit](#) [+](#) [New](#)

Options ☒ Allow schema drift ☐ Validate schema

source1 select1 derivedColumn1 lookup1 insert union1 derivedColumn3 sink2

Import data from AzureDataLakeStorage1 Renaming source1 to select1 with columns 'src_customer_id' Creating/updating the columns 'src_customer_id' Lookup on 'derivedColumn1' from 'Targetcust' Conditionally distributing the data in CustomerID Combining rows from transformation 'split1@insert' Creating/updating the columns 'src_customer_id' Columns: 10 total

Sink Settings Errors Mapping Optimize Inspect Data preview ●

Schema name * Refresh

Table name * Refresh

Table action ☒ None ☐ Recreate table ☐ Truncate table

Update method ⓘ ☒ Allow insert ☐ Allow delete ☐ Allow upsert ☐ Allow update

Use tempdb ⓘ ☐

Interim table schema Refresh

Pre SQL scripts ⓘ ☒ List of scripts ☐ Custom expression ⓘ

source1 select1 derivedColumn1 lookup1 insert union1 derivedColumn3 sink2

Import data from AzureDataLakeStorage1 Renaming source1 to select1 with columns 'src_customer_id' Creating/updating the columns 'src_customer_id' Lookup on 'derivedColumn1' from 'Targetcust' Conditionally distributing the data in CustomerID Combining rows from transformation 'split1@insert' Creating/updating the columns 'src_customer_id' Columns: 10 total

Sink Settings Errors Mapping Optimize Inspect Data preview ●

☒ Skip duplicate output columns ⓘ

☐ Auto mapping ⓘ + Add mapping Delete Reset Import schema View schema 10 mappings: All outputs mapped

Input columns	Output columns
<input type="checkbox"/> 123 src_customer_id	<input type="checkbox"/> 123 CustomerID
<input type="checkbox"/> abc src_first_name	<input type="checkbox"/> abc First_Name
<input type="checkbox"/> abc src_last_name	<input type="checkbox"/> abc Last_Name
<input type="checkbox"/> abc src_full_address	<input type="checkbox"/> abc Full_Address
<input type="checkbox"/> abc src_createdby	<input type="checkbox"/> abc CreatedBy
<input type="checkbox"/> src_createddate	<input type="checkbox"/> CreatedDate
<input type="checkbox"/> abc src_updatedby	<input type="checkbox"/> abc UpdatedBy
<input type="checkbox"/> src_updateddate	<input type="checkbox"/> UpdatedDate
<input type="checkbox"/> 121 src_hashkey	<input type="checkbox"/> 121 HashKey
<input type="checkbox"/> 123 src_isactive	<input type="checkbox"/> 123 IsActive

Similar logics for other files.