

Dimension Tables:

In Blob: deniesdata

1. Categories.xlsx
2. SalesRep.xlsx
3. SubCategories.xlsx

In MS_SQL(On-Premise)

1. Product

In Cosmos MongoDB

1. Geography

Fact Table:

In ADLS: deniessalesdata / Sales

1. Sales 2014.csv
2. Sales 2015.csv
3. Sales 2016.csv
4. Sales 2017.csv

Blob data

Dashboard > ksrdatasa | Containers >

deniesdata ...

Container

Search

Upload Change access level Refresh Delete Change tier Acquire lease Break lease View snapshots Create snapshot Give feedback

Diagnose and solve problems

Access Control (IAM)

Authentication method: Access key (Switch to Microsoft Entra user account)

Location: deniesdata

Search blobs by prefix (case-sensitive)

Show deleted blobs

Add filter

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
Categories.xlsx	28/04/2024, 20:09:12	Hot (Inferred)		Block blob	17.03 KiB	Available
SalesRep.xlsx	28/04/2024, 20:09:13	Hot (Inferred)		Block blob	17.62 KiB	Available
SubCategories.xlsx	28/04/2024, 20:09:13	Hot (Inferred)		Block blob	18.09 KiB	Available

On Prem data

Object Explorer

Connect ...

NAGARAJA (SQL Server 15.0.2110.4 - NAGARAJA\akula)

- Databases
 - System Databases
 - AdventureWorks2012
 - AdventureWorks2014
 - AdventureWorks2016
 - AdventureWorks2017
 - AdventureWorks2019
 - AdventureWorksDW2012
 - AdventureWorksDW2014
 - AdventureWorksDW2016
 - AdventureWorksDW2017
 - AdventureWorksDW2019
 - ADVV_2019_STAGING
 - Agri_Culture
 - demo
 - Duplicates
 - Empinfo
 - HR_Data
 - Indian_Bank
 - KSR_DATA
 - Database Diagrams
 - Tables
 - System Tables
 - FileTables
 - External Tables
 - Graph Tables
 - dbo.Product
 - Views
 - External Resources
 - Synonyms
 - Programmability
 - Server Roles

SQLQuery1.sql - N...GARAJA\akula (69)*

```
1 | use [KSR_DATA]
2 |
3 | Select * FROM [dbo].[Product]
```

Results

ProductID	Sub_Category_Key	Color	ProductName	RetailPrice	StandardCost	
1	3	Red	Alder	23.9500007629395	7.55000019073486	
2	2	Blue	Linder	23.9500007629395	7.55000019073486	
3	3	Green	Magnum	23.9500007629395	7.55000019073486	
4	4	Red	Quad	43.9500007629395	13.75	
5	5	Blue	Black Monk	43.9500007629395	13.75	
6	6	Green	Quad	43.9500007629395	13.75	
7	7	Red	Bing	26.9500007629395	8.25	
8	8	Blue	VanHelen	26.9500007629395	8.25	
9	9	Green	Magnum	26.9500007629395	8.25	
10	10	1	Florentine Pink	Carlo	29.9500007629395	9.14999961853027
11	11	4	Florentine Blue	Carlo	29.9500007629395	9.14999961853027
12	7	1	Red	Bing	26.9500007629395	8.25
13	10	1	Florentine Pink	Carlo	29.9500007629395	9.14999961853027

MongoDB data

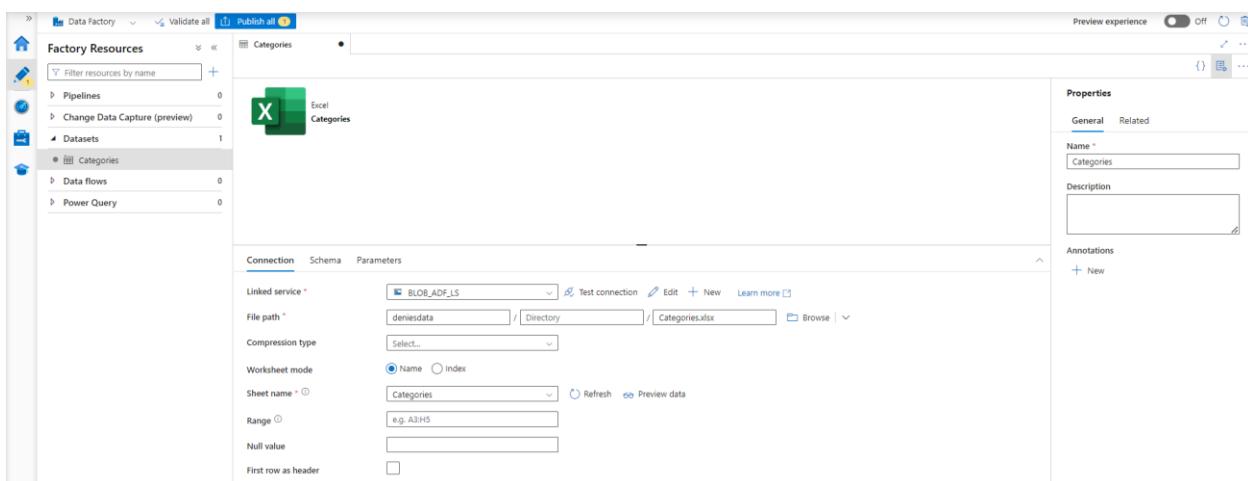
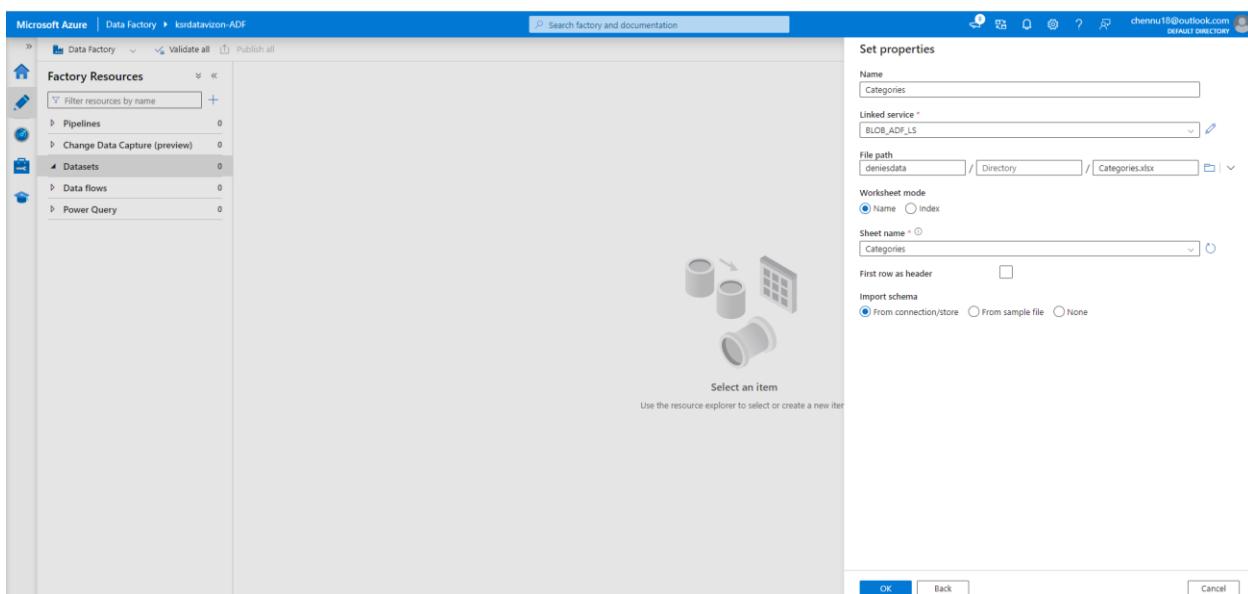
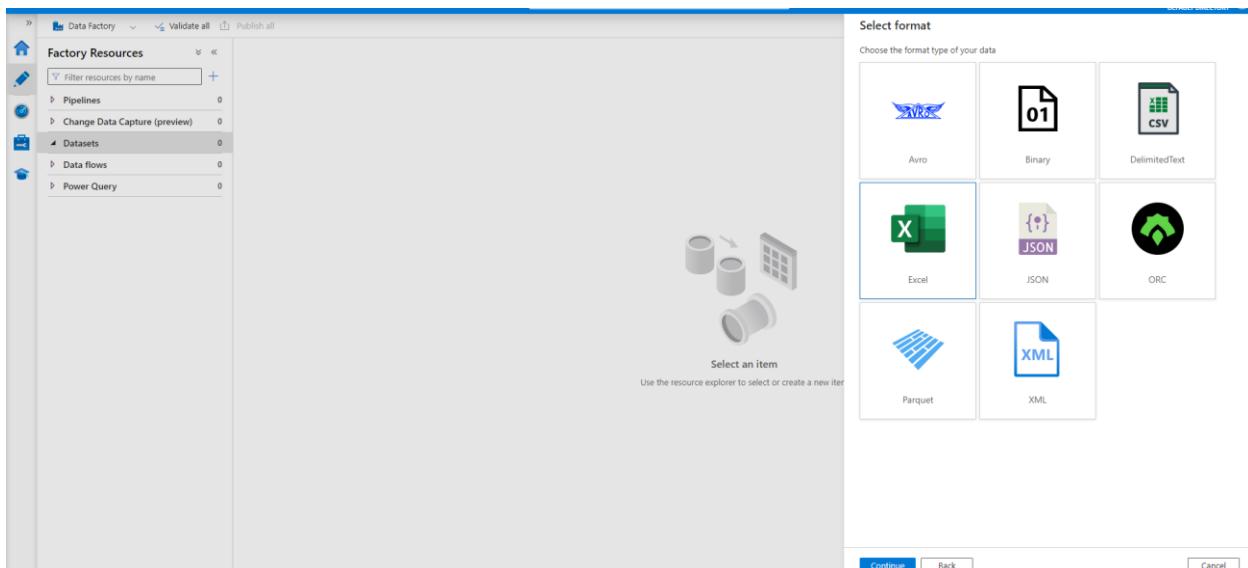
The screenshot shows the Azure Cosmos DB MongoDB API portal. On the left, there's a sidebar with 'CONNECTIONS' and 'SERVERS' sections. The main area has tabs for 'PROBLEMS', 'OUTPUT' (selected), 'TERMINAL', and 'TASKS'. A mongo shell session is running, displaying a query for geographical data. Below the terminal, there are two buttons: 'New Collection' and 'Import Sample Data'. On the right, a 'NOTIFICATIONS' panel shows a message: 'Successfully inserted 7 documents (took 0s)'.

Data lake data:

The screenshot shows the Azure Data Lake Storage Gen2 container named 'deniessalesdata'. The left sidebar includes 'Overview', 'Diagnose and solve problems', 'Access Control (IAM)', 'Settings', 'Shared access tokens', 'Manage ACL', 'Access policy', 'Properties', and 'Metadata'. The main area shows a table of blobs with columns: Name, Modified, Access tier, Archive status, Blob type, Size, and Lease state. The blobs listed are 'Sales 2014.csv', 'sales 2015.csv', 'sales 2016.csv', and 'sales 2017.csv'.

Datasets:

The screenshot shows the Azure Data Factory 'Factory Resources' blade. The left sidebar lists 'Factory Resources' like Pipelines, Datasets, Data flows, and Power Query. The 'Datasets' item is selected. The main area shows a central workspace with a 'Select an item' placeholder and a 'New dataset' dialog on the right. The dialog lists various Azure data stores under the 'Azure' tab, including Azure AI Search, Azure Blob Storage, Azure Cosmos DB for MongoDB, Azure Cosmos DB for NoSQL, Azure Data Explorer (Kusto), Azure Data Lake Storage Gen2, Azure Database for MariaDB, Azure Database for MySQL, and Azure Database for PostgreSQL. Buttons at the bottom are 'Continue' and 'Cancel'.



Data Factory Validate all Publish all

Factory Resources

- Pipelines 0
- Change Data Capture (preview) 0
- Datasets** 0
- Data flows 0
- Power Query 0

Filter resources by name +

Select an item
Use the resource explorer to select or create a new item

New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more [\[?\]](#)

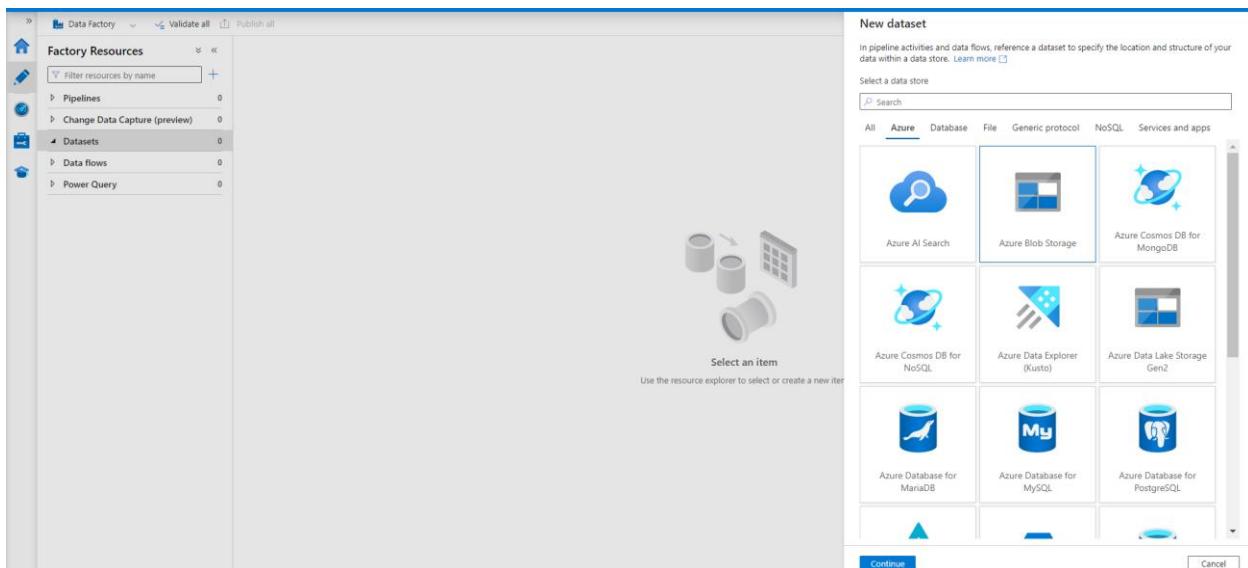
Select a data store

Search

All Azure Database File Generic protocol NoSQL Services and apps

Azure AI Search	Azure Blob Storage	Azure Cosmos DB for MongoDB
Azure Cosmos DB for NoSQL	Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2
Azure Database for MariaDB	Azure Database for MySQL	Azure Database for PostgreSQL

Continue Cancel



Data Factory Validate all Publish all

Factory Resources

- Pipelines 0
- Change Data Capture (preview) 0
- Datasets** 0
- Data flows 0
- Power Query 0

Filter resources by name +

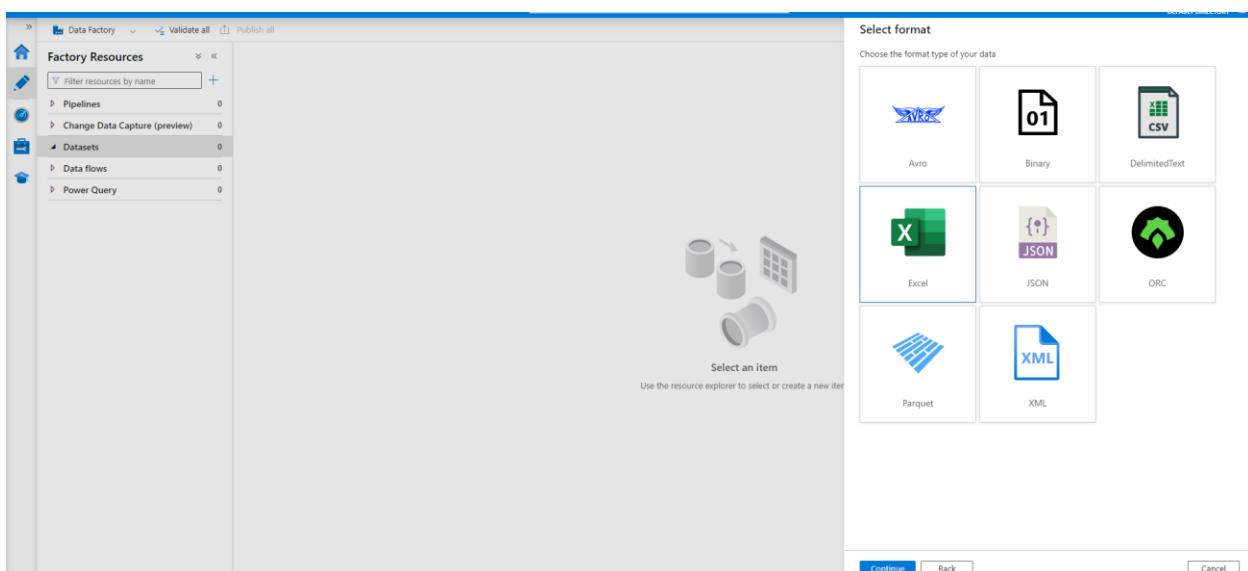
Select an item
Use the resource explorer to select or create a new item

Select format

Choose the format type of your data

Avro	Binary	DelimitedText
Excel	JSON	ORC
Parquet	XML	

Continue Back Cancel



Set properties

Name: SalesRep_DS

Linked service: BLOB_ADF_LS

File path: deniedata / [Directory] / SalesRep.xlsx

Worksheet mode: Name

Sheet name: Sales rep

First row as header:

Import schema: From connection/store

Connection: BLOB_ADF_LS

Schema: Categories

Parameters:

- File path: deniedata / [Directory] / SalesRep.xlsx
- Compression type: Select...
- Worksheet mode: Name
- Sheet name: Categories
- Range: e.g. A3:H5
- Null value:
- First row as header:

OK **Back** **Cancel**

SalesRep_DS

Properties

General: Name: SalesRep_DS

Description:

Annotations: + New

Connection: BLOB_ADF_LS

Schema: Sales rep

Parameters:

- File path: deniedata / [Directory] / SalesRep.xlsx
- Compression type: Select...
- Worksheet mode: Name
- Sheet name: Sales rep
- Range: e.g. A3:H5
- Null value:
- First row as header:

New dataset

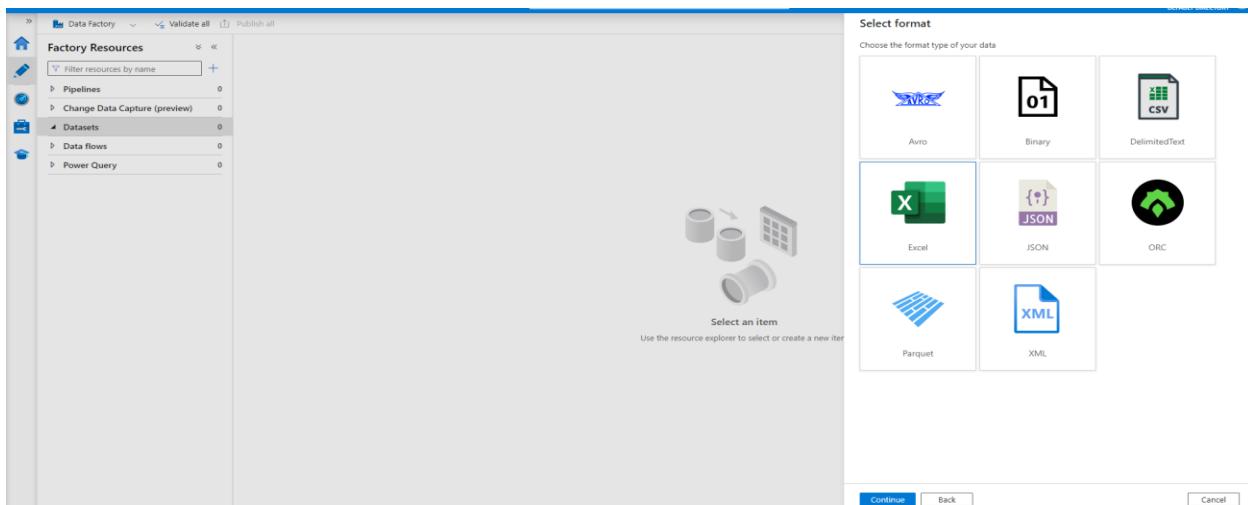
In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more

Select a data store

All Azure Database File Generic protocol NoSQL Services and apps

Azure AI Search	Azure Blob Storage	Azure Cosmos DB for MongoDB
Azure Cosmos DB for NoSQL	Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2
Azure Database for MariaDB	Azure Database for MySQL	Azure Database for PostgreSQL

Continue **Cancel**



Factory Resources

Categories_DS

Set properties

Name: SubCategories_DS

Linked service: BLOB_ADF_LS

File path: deniedata / Directory / Categories.xlsx

Compression type: Select...

Worksheet mode: Name (radio button selected)

Sheet name: Categories

Range: e.g. A3:H5

Null value:

First row as header:

Import schema: From connection/store (radio button selected)

OK **Back** **Cancel**

Factory Resources

Categories_DS **SalesRep_DS** **SubCategories_DS**

Properties

Name: SubCategories_DS

Description:

Annotations: + New

Connection **Schema** **Parameters**

Linked service: BLOB_ADF_LS

File path: deniedata / Directory / SubCategories.xlsx

Compression type: Select...

Worksheet mode: Name (radio button selected)

Sheet name: Sub categories

Range: e.g. A3:H5

Null value:

First row as header:

Data Factory

Factory Resources

- Pipelines 0
- Change Data Capture (preview) 0
- Datasets 3
 - Categories_DS
 - SalesRep_DS
 - SubCategories_DS
- Data flows 0
- Power Query 0

Categories_DS **SalesRep_DS** **SubCategories_DS**

Excel SubCategories_DS

New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more [\[?\]](#)

Select a data store

Search

All	Azure	Database	File	Generic protocol	NoSQL	Services and apps
PostgreSQL (Legacy)	Presto	SAP BW Open Hub				
SAP BW	SAP HANA	SAP CDC				
SAP Table	SQL server	Snowflake				
Spark	Sybase					

Continue **Cancel**

Data Factory

Factory Resources

- Pipelines 0
- Change Data Capture (preview) 0
- Datasets 3
 - Categories_DS
 - SalesRep_DS
 - SubCategories_DS
- Data flows 0
- Power Query 0

Categories_DS **SalesRep_DS** **SubCategories_DS**

Excel SubCategories_DS

Set properties

Name

Linked service *

Connect via integration runtime *

Table name Enter manually

Import schema From connection/store None

OK **Back** **Cancel**

Set properties

Name: Product_DS

Linked service: SHIR_MSSQL_ADF_LS

Connect via integration runtime: SelfHostedIR

Table name: dbo.Product

Import schema: From connection/store

Properties

General: Name: Product_DS

Description:

Annotations: + New

New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more

Select a data store: cos

All: Azure Database File Generic protocol NoSQL Services and apps

Azure Cosmos DB for MongoDB and Azure Cosmos DB for NoSQL

Set properties

- Name: Geography_DS
- Linked service: CosmosDbMongoDB_ADF_LS
- Collection name: Geography
- Enter manually

Connection

- Linked service: SHIR_MSSQL_ADF_LS
- Integration runtime: Self-hosted
- Table: dbo.Product

Parameters

Properties

- General
- Related

Name: Geography_DS

Description

Annotations

Preview experience: Off

Connection

- Linked service: CosmosDbMongoDB_ADF_LS
- Collection name: Geography

Parameters

From On-Prem to Blob

Activities

- Copy data
- Copy data1

Properties

- General
- Related

Name: On_Prem-MSSQL_to_Blob

Source dataset: Product_DS

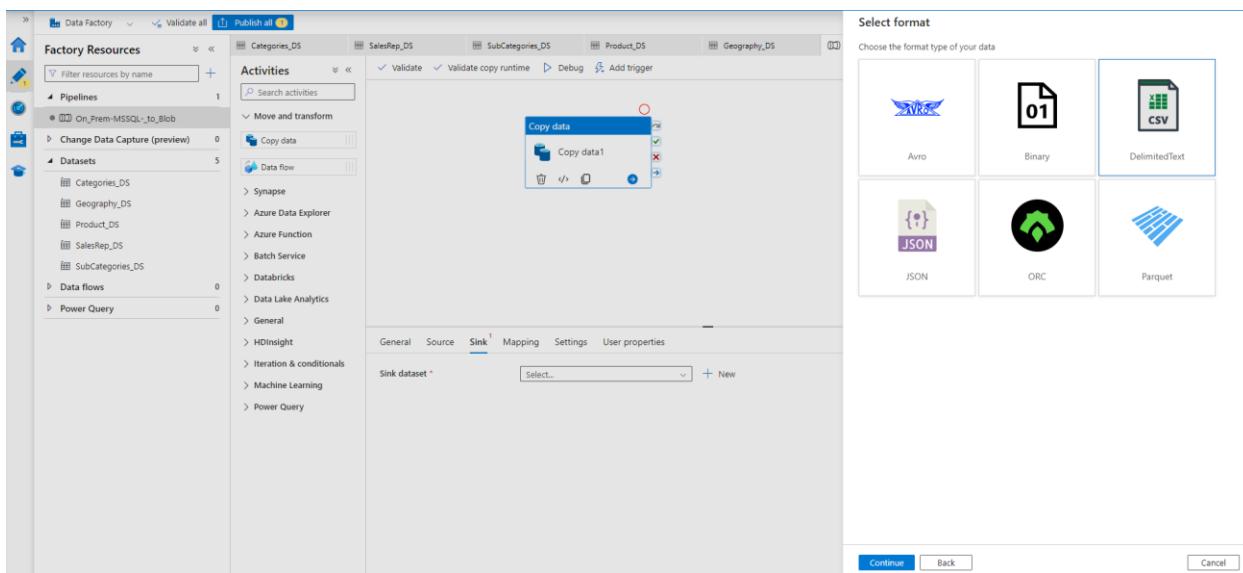
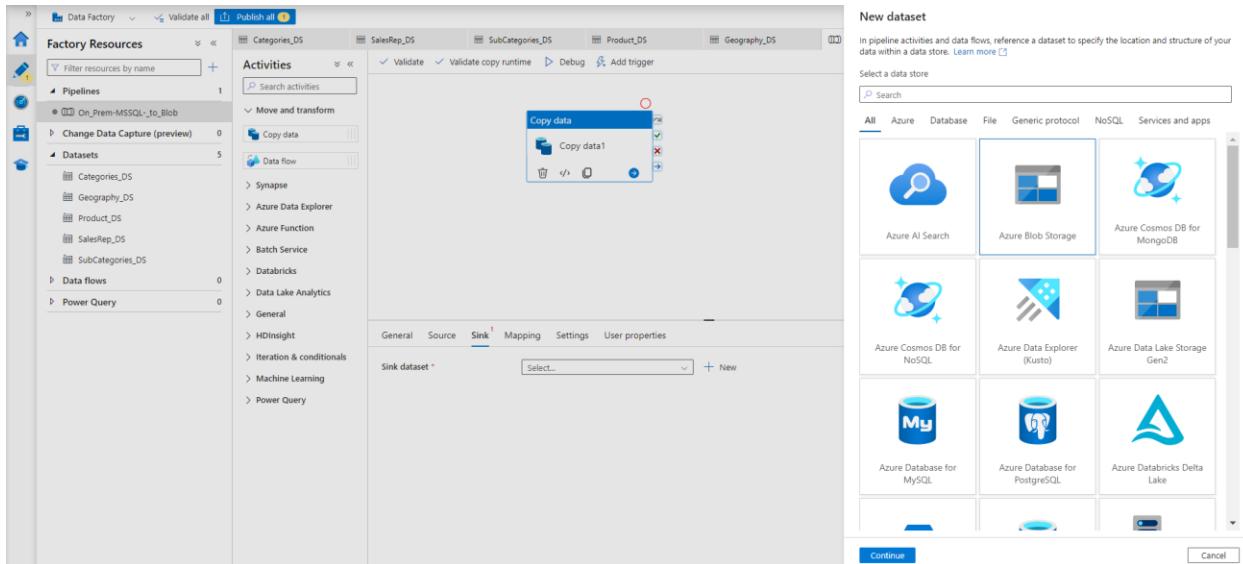
Use query: Table

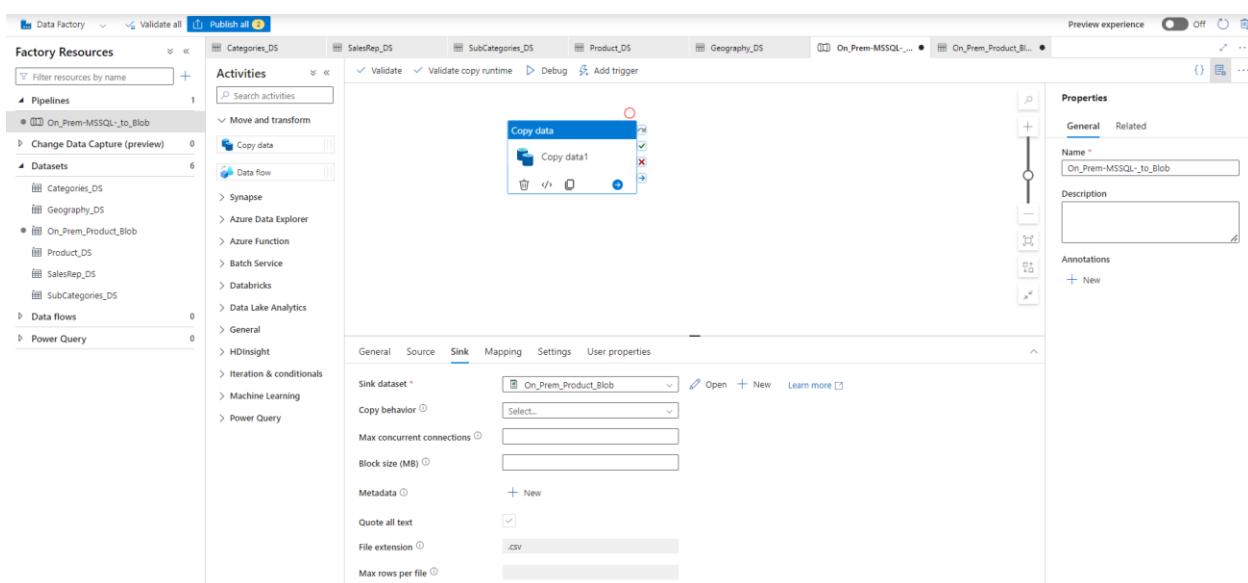
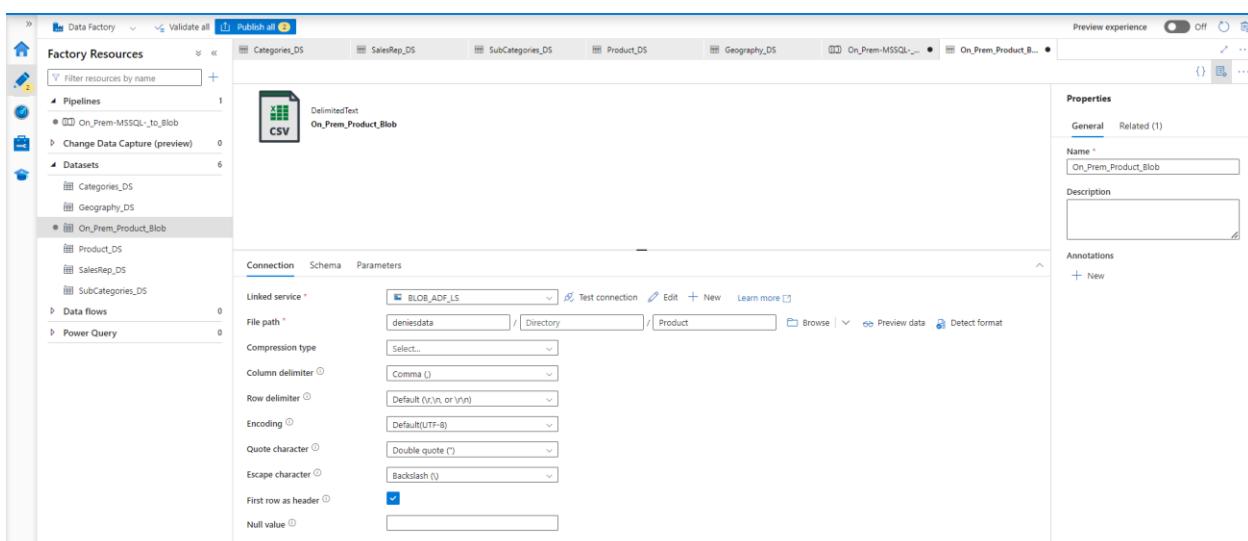
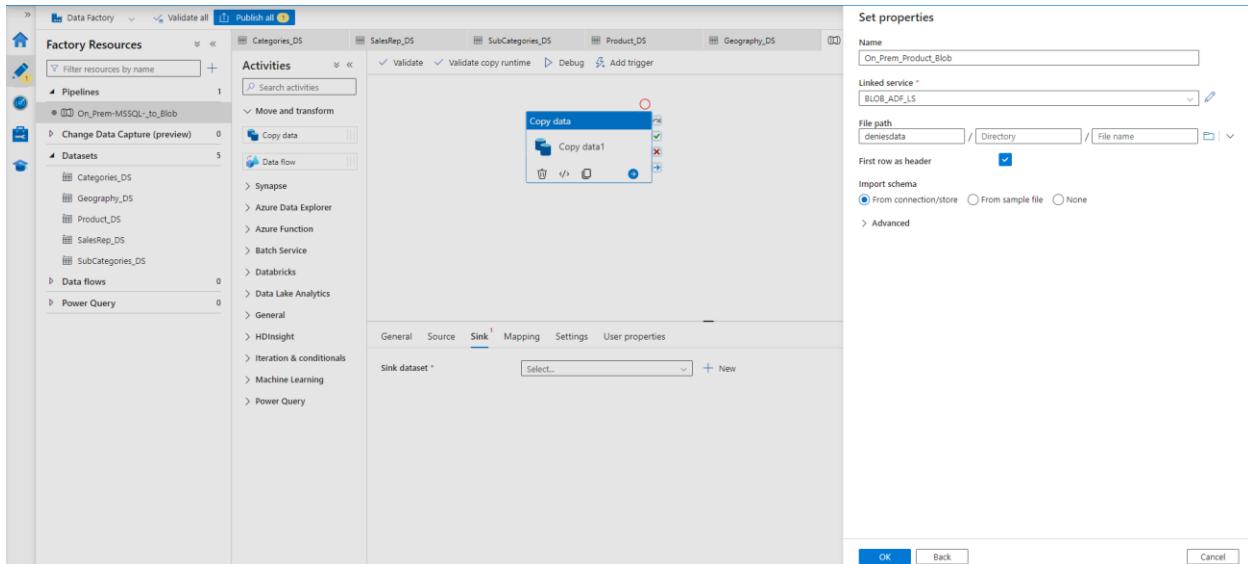
Query timeout (minutes): 120

Isolation level: Select...

Partition option: None

Additional columns: + New





The screenshot shows the Azure Data Factory Pipeline Editor. On the left, the 'Factory Resources' sidebar lists 'Pipelines' (1), 'Datasets' (6), and other resources. In the center, a 'Copy data' activity is selected under 'Activities'. The 'Output' tab is active, showing a single pipeline run entry:

Activity name	Activity status	Activity type	Run start	Duration	Integration runtime	User properties
Copy data from Onprem to Bl...	Succeeded	Copy data	4/28/2024, 9:14:05 PM	12s	SelfHostedIR	

The screenshot shows the Azure Storage Blob container 'deniesdata'. The 'Overview' tab is selected. The container has an 'Authentication method' of 'Access key' and a 'Location' of 'deniesdata'. A search bar at the top right allows filtering by prefix. The table below lists the blobs:

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
Categories.xlsx	28/04/2024, 20:09:12	Hot (Inferred)		Block blob	17.03 KB	Available
Product.csv	28/04/2024, 21:15:54	Hot (Inferred)		Block blob	778 B	Available
SalesRep.xlsx	28/04/2024, 20:09:13	Hot (Inferred)		Block blob	17.62 KB	Available
SubCategories.xlsx	28/04/2024, 20:09:13	Hot (Inferred)		Block blob	18.09 KB	Available

The screenshot shows the 'Product.csv' file within the 'deniesdata' container. The 'Edit' tab is selected. The file content is displayed as follows:

```

1 ProductID,Sub_Catgeory_Key,Color,ProductName,RetailPrice,StandardCost
2 1,3,"Red","Alder",23.950000762939453,7.5500001907348633
3 2,2,"Blue","Linder",23.950000762939453,7.5500001907348633
4 3,2,"Green","Magnum",23.950000762939453,7.5500001907348633
5 4,1,"Red","Quad",43.950000762939453,13.75
6 5,1,"Blue","Black Monk",43.950000762939453,13.75
7 6,4,"Green","Quad",43.950000762939453,13.75
8 7,1,"Red","Bing",26.950000762939453,8.25
9 8,3,"Blue","VanHelen",26.950000762939453,8.25
10 9,1,"Green","Magnum",26.950000762939453,8.25
11 10,5,"Fluorescent Pink","Carlota",29.950000762939453,9.1499996185302734
12 11,4,"Fluorescent Blue","Carlota",29.950000762939453,9.1499996185302734
13 7,1,"Red","Bing",26.950000762939453,8.25
14 10,1,"Fluorescent Pink","Carlota",29.950000762939453,9.1499996185302734
15

```

At the bottom, there are 'Csv' and 'Preview' buttons.

From MongoDB to Blob

The screenshot shows the Azure Data Factory pipeline editor interface. A pipeline named "MongoDB_Geography_Blob" is being configured. The pipeline consists of two main stages: "Copy data" and "Copy data from mongoDB to blob".

Properties Panel:

- Name: MongoDB_Geography_Blob
- Description: (empty)
- Annotations: + New

Copy data Activity (Source Tab):

- Source dataset: Geography_DS
- Filter: (empty)
- Batch size: 100
- Additional columns: + New

Copy data Activity (Sink Tab):

- Sink dataset: Select...

New dataset Dialog:

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more [\[link\]](#)

Select a data store: Search

Categories:

- All
- Azure
- Database
- File
- Generic protocol
- NoSQL
- Services and apps

Azure AI Search	Azure Blob Storage	Azure Cosmos DB for MongoDB
Azure Cosmos DB for NoSQL	Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2
Azure Database for MySQL	Azure Database for PostgreSQL	Azure Databricks Delta Lake

Buttons: Continue, Cancel

Select format

Choose the format type of your data

 Avro	 Binary	 CSV
 JSON	 ORC	 Parquet

Continue **Back** **Cancel**

Set properties

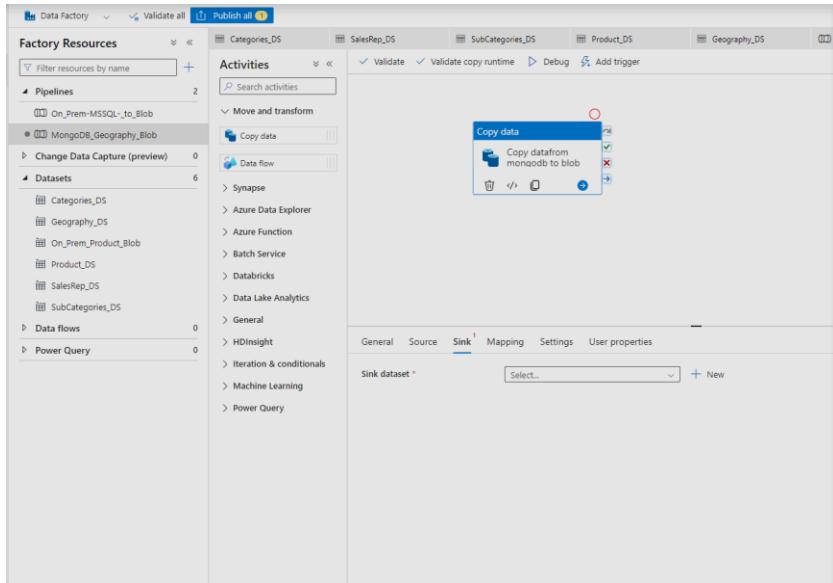
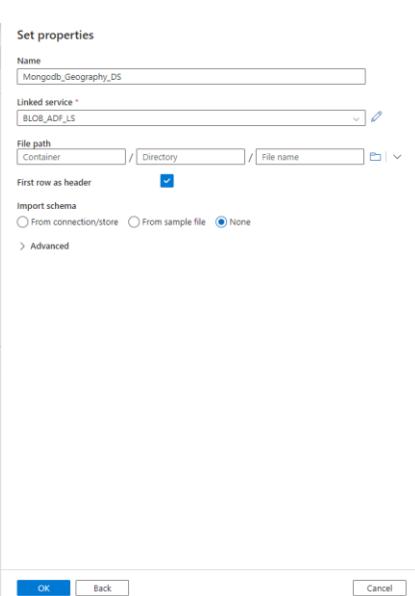
Name: MongoDB_Geography_DS
Linked service: BLOB_ADF_LS

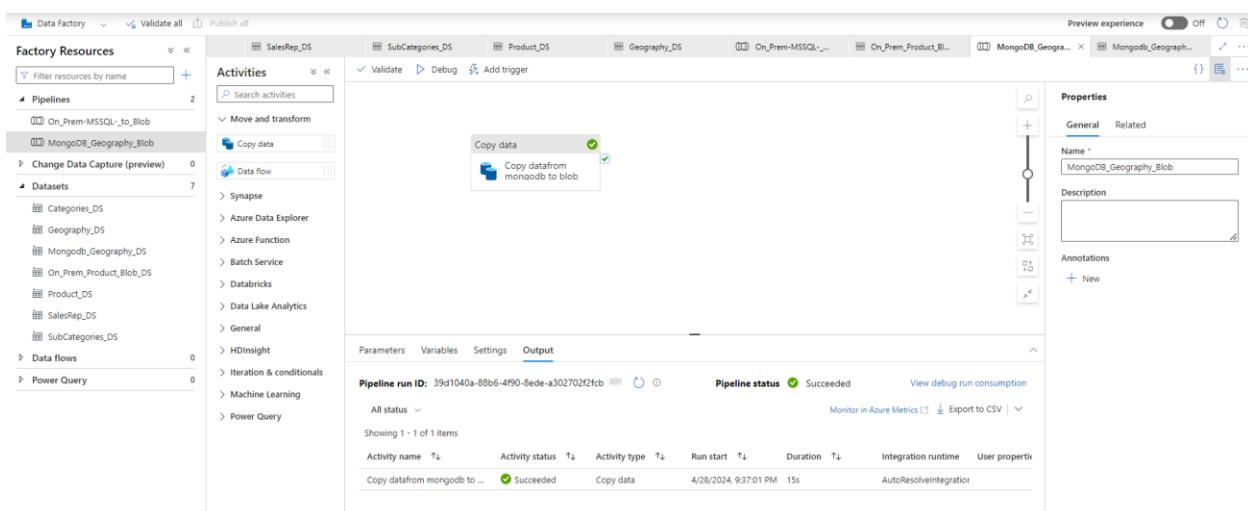
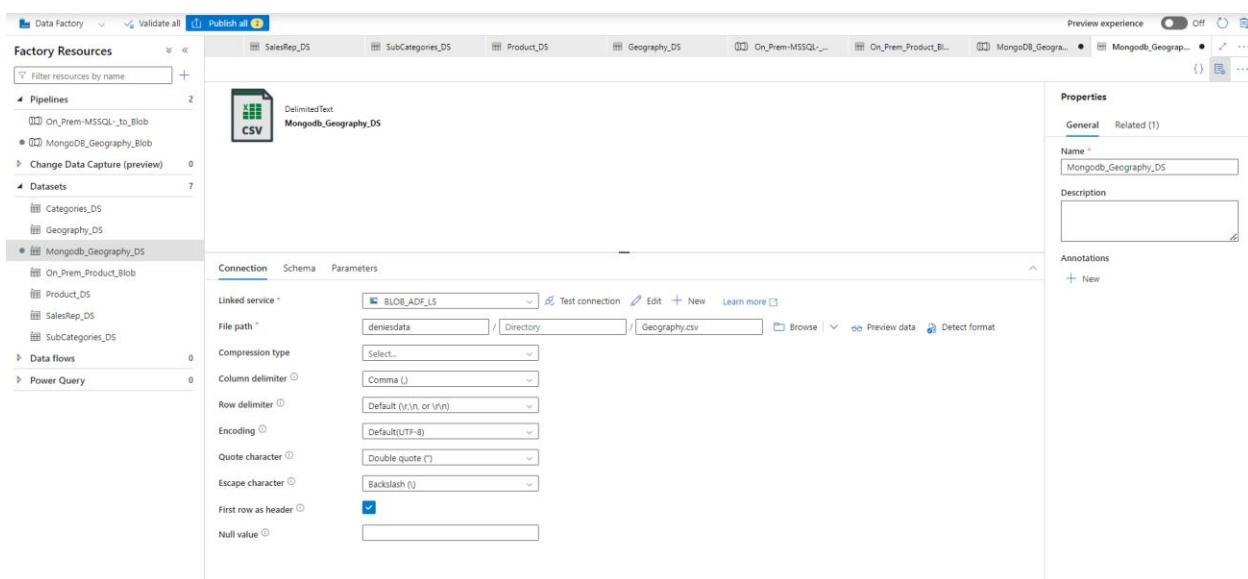
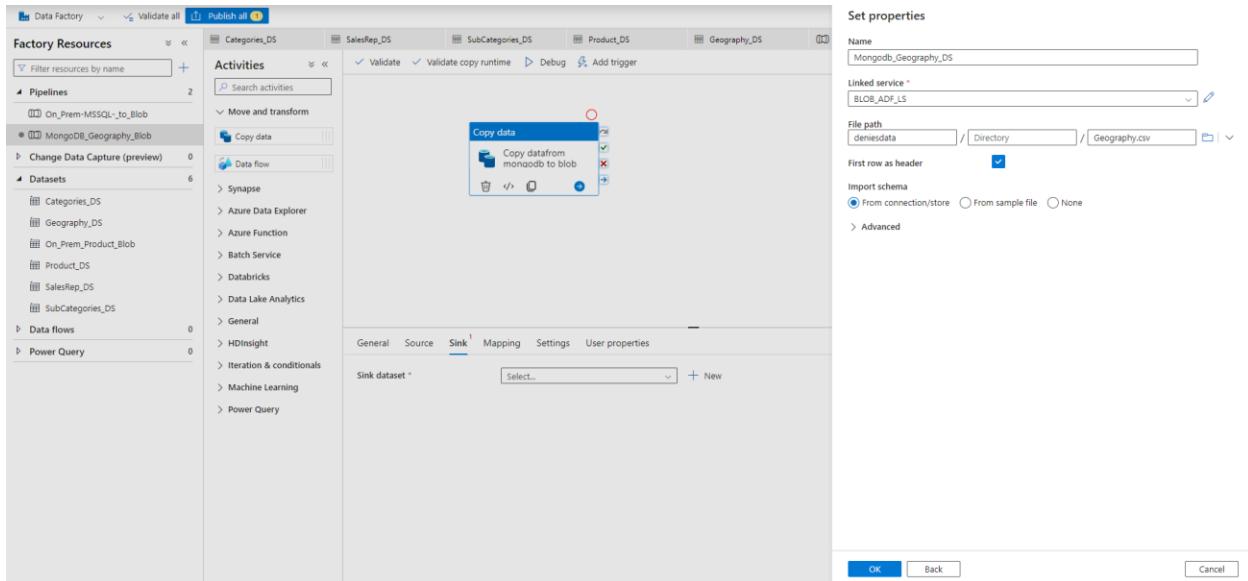
File path: Container / Directory / File name

First row as header:

Import schema: From connection/store From sample file None

OK **Back** **Cancel**



Dashboard > ksrdatas | Containers >

deniesdata

Container

Overview Diagnose and solve problems Access Control (IAM)

Settings Shared access tokens Access policy Properties Metadata

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
Categories.xlsx	28/04/2024, 20:09:12	Hot (Inferred)		Block blob	17.03 KB	Available
Geography.csv	28/04/2024, 21:37:15	Hot (Inferred)		Block blob	472 B	Available
Product.csv	28/04/2024, 21:15:54	Hot (Inferred)		Block blob	778 B	Available
SalesRep.xlsx	28/04/2024, 20:09:13	Hot (Inferred)		Block blob	17.62 KB	Available
SubCategories.xlsx	28/04/2024, 20:09:13	Hot (Inferred)		Block blob	18.09 KB	Available

Show deleted blobs

Dashboard > ksrdatas | Containers > deniesdata >

deniesdata

Container

Overview Diagnose and solve problems Access Control (IAM)

Settings Shared access tokens Access policy Properties Metadata

Add filter

Name
Categories.xlsx
Geography.csv
Product.csv
SalesRep.xlsx
SubCategories.xlsx

Geography.csv ...

Blob

```

1 Country,Town,Wikipedia
2 "Czech republic","Prague","https://en.wikipedia.org/wiki/Prague"
3 "Denmark","Copenhagen","https://en.wikipedia.org/wiki/Copenhagen"
4 "Germany","Frankfurt","https://en.wikipedia.org/wiki/Frankfurt"
5 "Germany","Berlin","https://en.wikipedia.org/wiki/Berlin"
6 "Germany","Dresden","https://en.wikipedia.org/wiki/Dresden"
7 "Czech republic","Brno","https://en.wikipedia.org/wiki/Brno"
8 "Czech republic","Ostrava","https://en.wikipedia.org/wiki/Ostrava"
9

```

Create Data flows:

Data Factory

Factory Resources

- Pipelines: On_Prem_MSSQL_to_Blob, MongoDB_Geography_Blob
- Change Data Capture (preview): 0
- Datasets: 7
- Data flows: 1
- Power Query: 0

Categories_DF

Categories_DF

Data flow debug

CategoriesData

Columns: 2 total

Add Source

Properties

General Related

Name: Categories_DF

Description

Source settings

Output stream name: CategoriesData

Description: Import data from Categories_DS

Source type: Dataset Inline

Dataset: Categories_DS

Options: Allow schema drift Infer drifted column types Validate schema

Sampling: Enable Disable

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
- Power Query

Data Flow Editor - Categories_DF

Data Flow Editor - Categories_DF

Data Flow Editor - Categories_DF

Data Factory Pipeline Editor

Pipeline Details: Categories_DF

Stages:

- Import data from Categories_DS:** Categoriesdataset
- aggregate1:** Aggregating data by CategoryKey, Category, CCategories
- select1:** Renaming aggregate1 to select1 with columns 'CategoryKey, Category, CCategories'

Output Stream: select1

Description: Renaming aggregate1 to select1 with columns 'CategoryKey, Category, CCategories'

Incoming Stream: aggregate1

Options:

- Skip duplicate input columns
- Skip duplicate output columns

Input Columns:

aggregate1's column	Name as
121 CategoryKey	CategoryKey
abc Category	Category
121 CCategories	CCategories

Data Factory Pipeline Editor

Pipeline Details: Categories_DF

Stages:

- Import data from Categories_DS:** Categoriesdataset
- aggregate1:** Aggregating data by CategoryKey, Category, CCategories
- select1:** Renaming aggregate1 to select1 with columns 'CategoryKey, Category'

Output Stream: select1

Description: Renaming aggregate1 to select1 with columns 'CategoryKey, Category'

Incoming Stream: aggregate1

Options:

- Skip duplicate input columns
- Skip duplicate output columns

Input Columns:

aggregate1's column	Name as
121 CategoryKey	CategoryKey
abc Category	Category

Note: 2 mappings: 1 column(s) from the inputs left unmapped.

Data Factory Pipeline Editor

Pipeline Details: Categories_DF

Stages:

- Import data from Categories_DS:** Categoriesdataset
- aggregate1:** Aggregating data by CategoryKey, Category, CCategories
- select1:** Renaming aggregate1 to select1 with columns 'CategoryKey, Category'
- sink1:** Column: 2 total

Sink:

Output Stream: sink1

Description: Add sink dataset

Incoming Stream: select1

Sink type: Dataset

Dataset: Select...

Options:

- Allow schema drift
- Validate schema

New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more

Select a data store

All	Azure	Database	File	Generic protocol	NoSQL	Services and apps
Amazon S3	Azure Blob Storage	Azure Cosmos DB for NoSQL				
Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2	Azure Database for MySQL				
Azure Database for PostgreSQL	Azure SQL Database	Azure SQL Database Managed Instance				

Buttons: Continue, Cancel

Data Factory Validate all Publish all

Factory Resources

- Pipelines
- Datasets
- Data flows
- Power Query

Categories_DF

Validate Data flow debug Debug Settings

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name: sink1 Learn more

Description: Add sink dataset Reset

Incoming stream: select1

Sink type: Dataset Inline Cache

Dataset: Select... New

Options: Allow schema drift (checked) Validate schema (unchecked)

Select format

Choose the format type of your data

Avro	DelimitedText	JSON
ORC	Parquet	Binary

Continue Back Cancel

Data Factory Validate all Publish all

Factory Resources

- Pipelines
- Datasets
- Data flows
- Power Query

Categories_DF

Validate Data flow debug Debug Settings

Sink Settings Errors Mapping Optimize Inspect Data preview

Output stream name: sink1 Learn more

Description: Add sink dataset Reset

Incoming stream: select1

Sink type: Dataset Inline Cache

Dataset: Select... New

Options: Allow schema drift (checked) Validate schema (unchecked)

Set properties

Name: Stg_Categories_DS

Linked service: BLOB_ADF_LS

File path: stagingdata / Directory / File name

First row as header:

Import schema: From connection/store From sample file None

> Advanced

OK Back Cancel

Data Factory Validate all Publish all

Factory Resources

- Pipelines
- Datasets
- Data flows
- Power Query

Stg_Categories_DS

DelimitedText

CSV

Connection Schema Parameters

Linked service: BLOB_ADF_LS Test connection Edit New Learn more

File path: stagingdata / Directory / Stg_Categories.csv Browse Preview data Detect format

Compression type: Select...

Column delimiter: Comma (,)

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

Encoding: Default(UTF-8)

Quote character: Double quote (")

Escape character: Backslash (\)

First row as header:

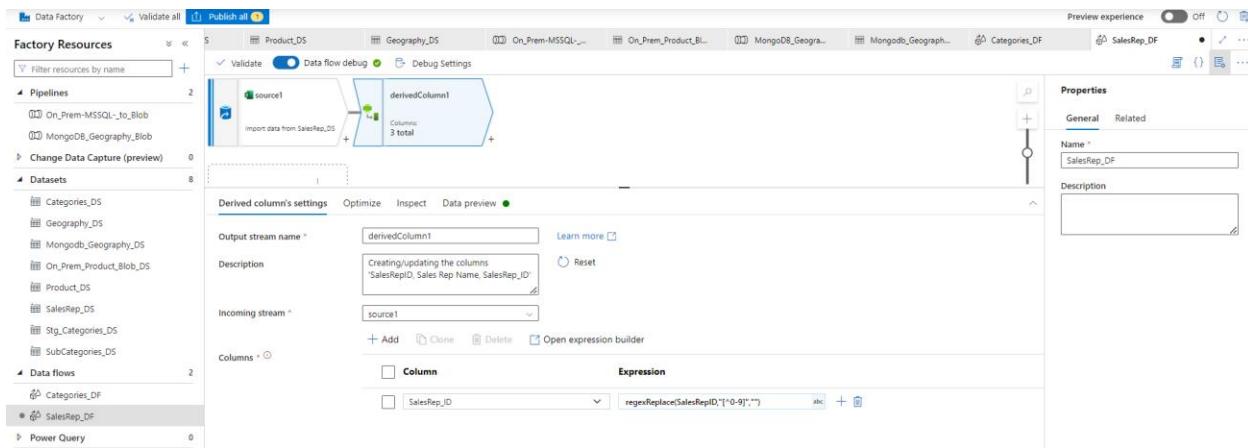
Null value:

Preview experience: Off

The screenshot shows the Azure Data Factory Data Flow interface. On the left, the 'Factory Resources' sidebar lists Pipelines, Datasets, and Data flows. Under 'Data flows', 'SalesRep_DF' is selected. The main workspace displays a 'source1' component with two columns: SalesRepID and Sales Rep Name. The 'Properties' pane on the right shows the 'Name' field set to 'SalesRep_DF'.

The screenshot shows the 'Projection' tab for the 'source1' component. It lists two columns: SalesRepID and Sales Rep Name, both mapped to 'string' type. A success message 'Successfully imported' is displayed in the top right corner.

The screenshot shows the 'Data preview' tab for the 'source1' component. It displays 7 rows of data from the SalesRep table, showing columns SalesRepID and Sales Rep Name. The data includes rows for Jan Novotny, John White, Ellen Woody, Mark Spancer, Ellie Gill, Bill Murray, and El Bob.



Dataflow expression builder

derivedColumn1

Derived Columns

+ Create new

SalesRep_ID

Expression

`regexReplace(SalesRepID,"[^0-9]","",)`

Expression reference documentation [Expression reference documentation](#)

Expression elements

Expression values
reg

All

Functions

Input schema

Parameters

Cached lookup

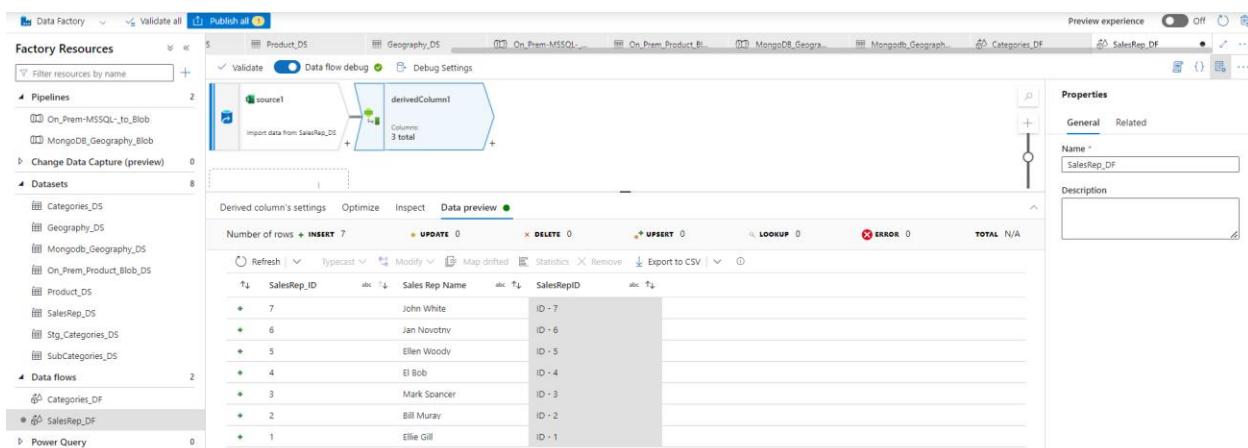
Data flow library functions

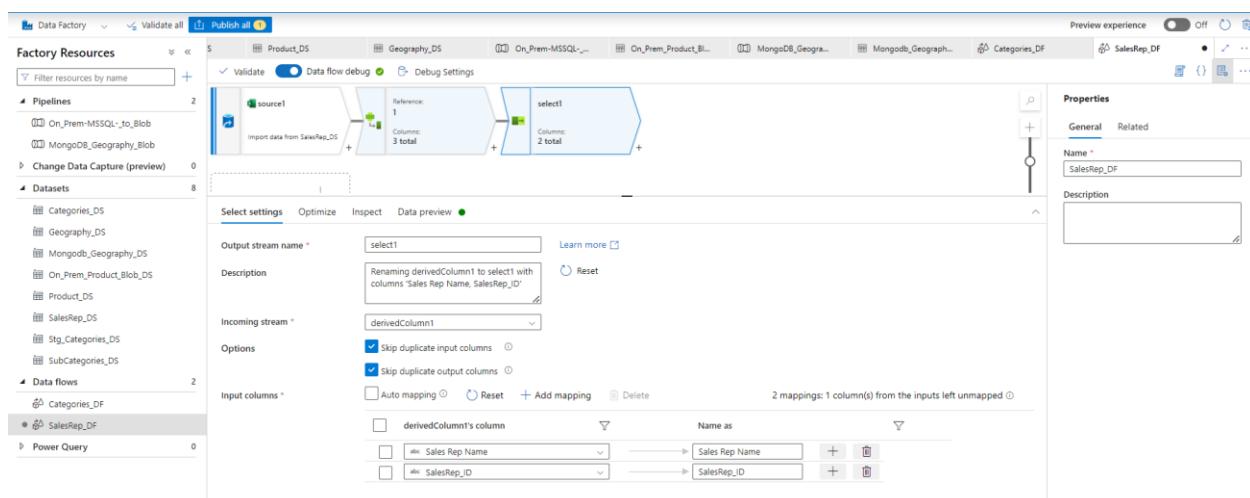
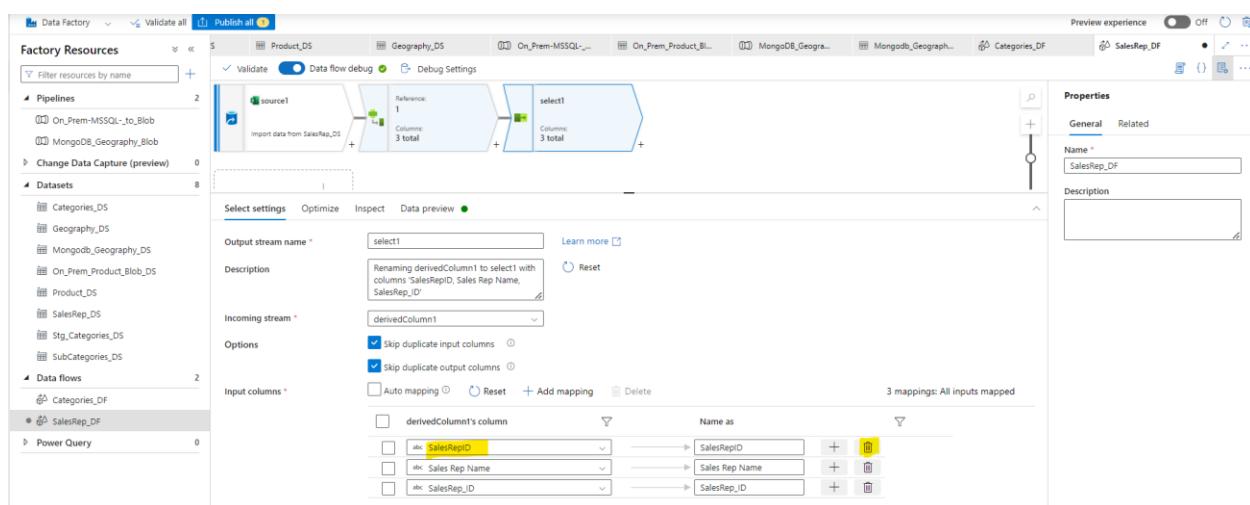
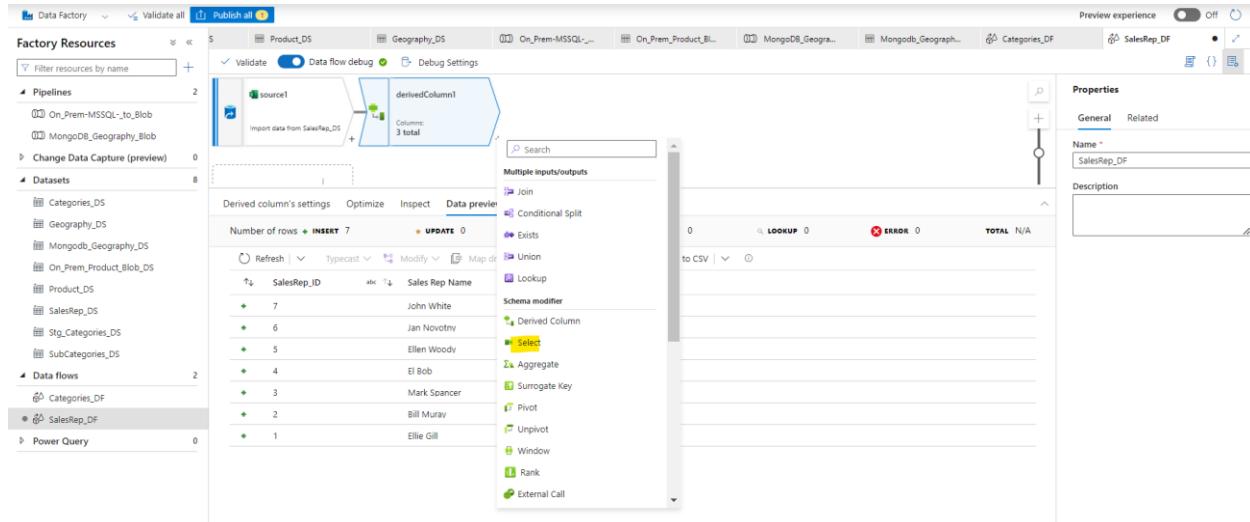
Locals

Save

Data preview Refresh

Save and finish Cancel Clear contents





Data Flow Editor - SalesRep_DF

New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more [\[\]](#)

Select a data store

All Azure Database File Generic protocol NoSQL Services and apps

Amazon S3	Azure Blob Storage	Azure Cosmos DB for NoSQL
Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2	Azure Database for MySQL
Azure Database for PostgreSQL	Azure SQL Database	Azure SQL Database Managed Instance
Continue	Cancel	

Data Factory Pipeline Editor

Pipeline Details: Product_DS

Stages:

- Stage 1:** Import data from SalesRep_DS → derivedColumn1 → select1 → sink1
- Stage 2:** (Placeholder)

Sink: On_Prem_Product_Blob

Output stream name: sink1

Description: Add sink dataset

Incoming stream: select1

Sink type: Dataset

Dataset: Select...

Options: Allow schema drift (checked), Validate schema (unchecked)

Select format: Choose the format type of your data

Avro	DelimitedText	JSON
ORC	Parquet	Binary

Buttons: Continue, Back, Cancel

Data Factory Pipeline Editor

Pipeline Details: Product_DS

Stages:

- Stage 1:** Import data from SalesRep_DS → derivedColumn1 → select1 → sink1
- Stage 2:** (Placeholder)

Sink: On_Prem_Product_Blob

Output stream name: sink1

Description: Add sink dataset

Incoming stream: select1

Sink type: Dataset

Dataset: Select...

Options: Allow schema drift (checked), Validate schema (unchecked)

Set properties:

- Name:** Stg_SalesRep_DS
- Linked service:** BLOB_ADF_LS
- File path:** stagingdata / Directory / File name
- First row as header:** checked
- Import schema:** From connection/store (radio button selected)

Buttons: Continue, Back, Cancel

Data Factory Pipeline Editor

Pipeline Details: Product_DS

Stages:

- Stage 1:** Import data from SalesRep_DS → derivedColumn1 → select1 → sink1
- Stage 2:** (Placeholder)

Sink: DelimitedText

Format: CSV

Properties:

- Name:** Stg_SalesRep_DS
- Description:** (empty)
- Annotations:** + New

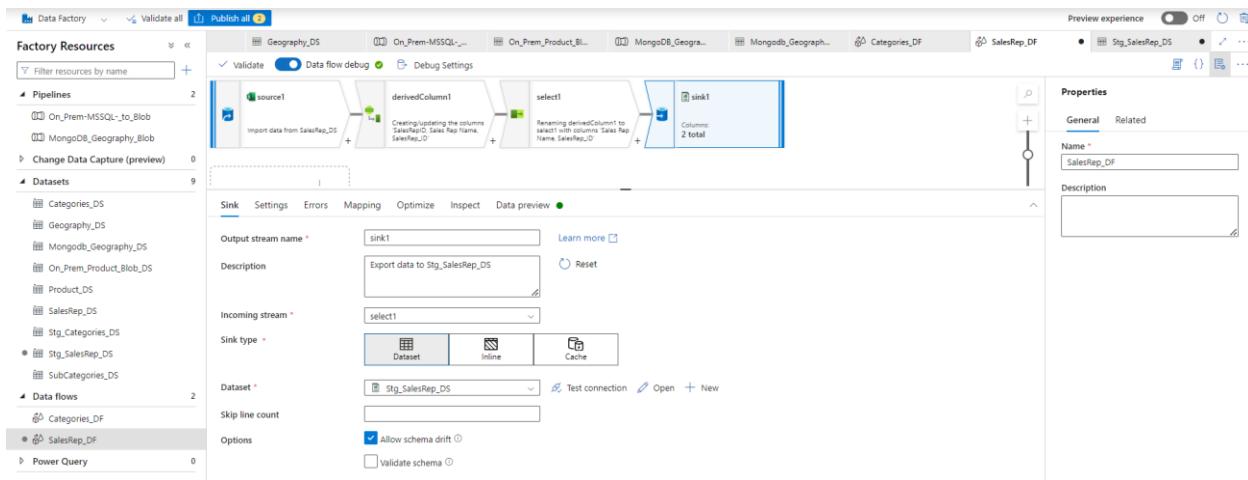
Connection:

- Linked service:** BLOB_ADF_LS
- File path:** stagingdata / Directory / Stg_SalesRep.csv
- Test connection:** (button)
- Preview data:** (button)
- Detect format:** (button)

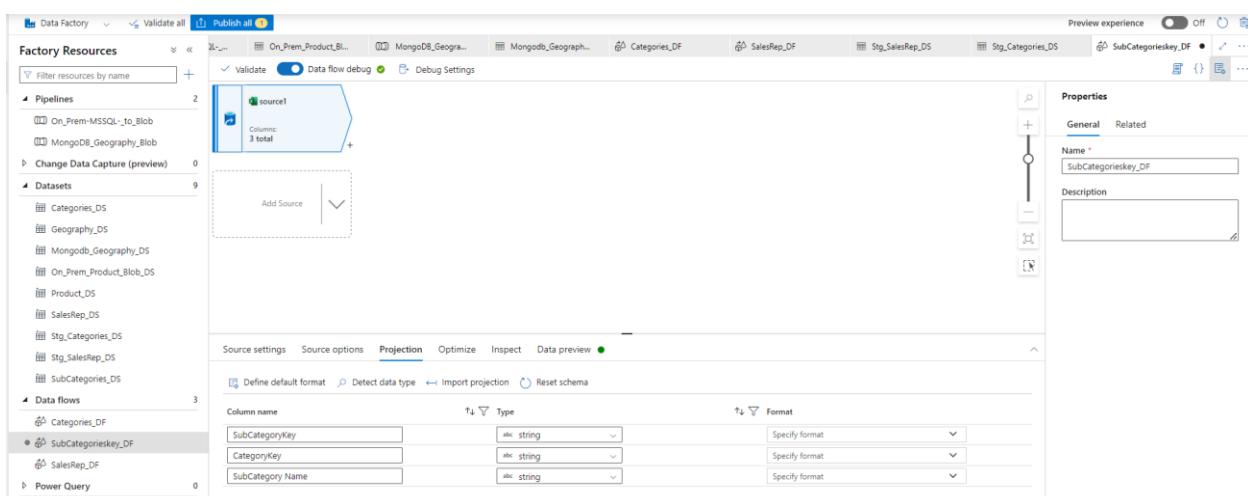
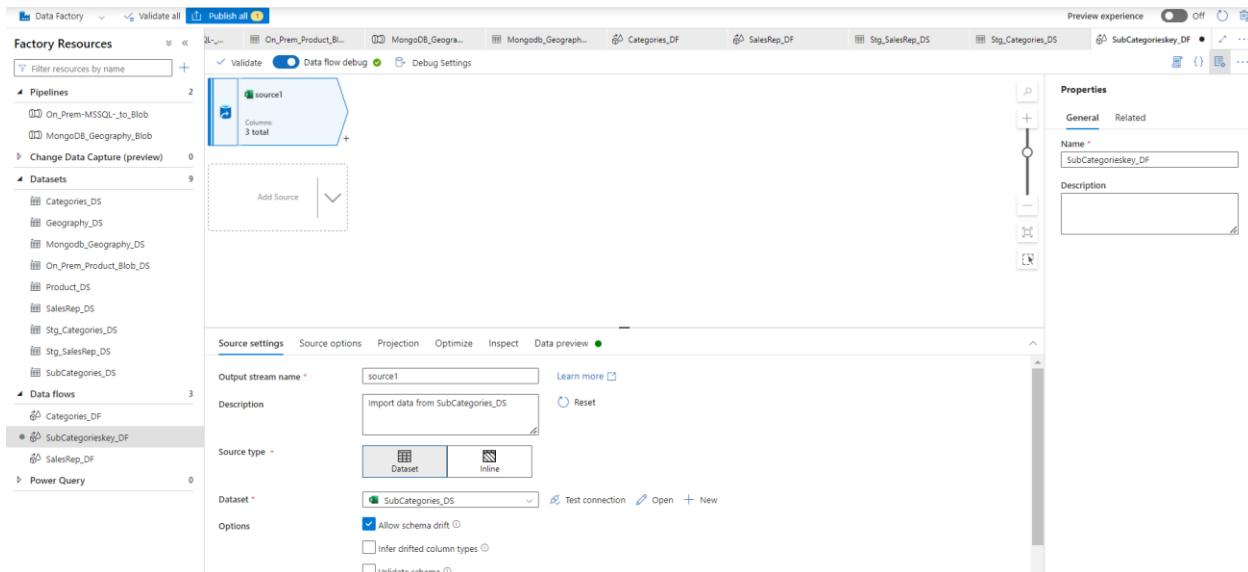
Schema:

- Column delimiter:** Comma (,)
- Row delimiter:** Default (\r\n or \n)
- Encoding:** Default(UTF-8)
- Quote character:** Double quote (")
- Escape character:** Backslash (\)
- First row as header:** checked
- Null value:** (empty)

Buttons: Preview experience (switch), Off, ...



SubCategories Key:



Factory Resources

- Pipelines
 - On_Prem_MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_SalesRep_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - SubCategorieskey_DF
 - SalesRep_DF
- Power Query

source1

Validate all

Properties

Name: SubCategorieskey_DF

Description:

Source settings Source options Projection Optimize Inspect Data preview

Number of rows: INSERT 4 UPDATE 0 DELETE 0 UPSERT 0 LOOKUP 0 ERROR 0 TOTAL 4

Refresh Typecast Modify Map drifted Statistics Export to CSV

SubCategoryKey	CategoryKey	SubCategory Name
1	ID - 1	Extra
2	ID - 2	Regular
3	ID - 1	Micro
4	ID - 2	Super

Factory Resources

- Pipelines
 - On_Prem_MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_SalesRep_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - SubCategorieskey_DF
 - SalesRep_DF
- Power Query

source1

Validate all

Properties

Name: SubCategorieskey_DF

Description:

Source settings Source opt: Select Aggregate Surrogate Key

Number of rows: INSERT 4 UPDATE 0 DELETE 0 UPSERT 0 LOOKUP 0 ERROR 0 TOTAL N/A

Refresh Typecast Pivot Unpivot Window Rank External Call

Multiple inputs/outputs

- Join
- Conditional Split
- Exists
- Union
- Lookup
- Schema modifier
 - Derived Column

Data preview

DELETE 0 UPSERT 0 LOOKUP 0 ERROR 0 TOTAL N/A

Statistics Remove Export to CSV

SubCategory Name
Extra
Regular
Micro
Super

Dataflow expression builder

derivedColumn1

Derived Columns

+ Create new ▾

Category_Key

Column name *
Category_Key

Expression
`regexReplace(CategoryKey, "[^0-9]", "")`

Expression elements

All
Functions
Input schema
Parameters
Cached lookup
Data flow library functions
Locals

Expression values

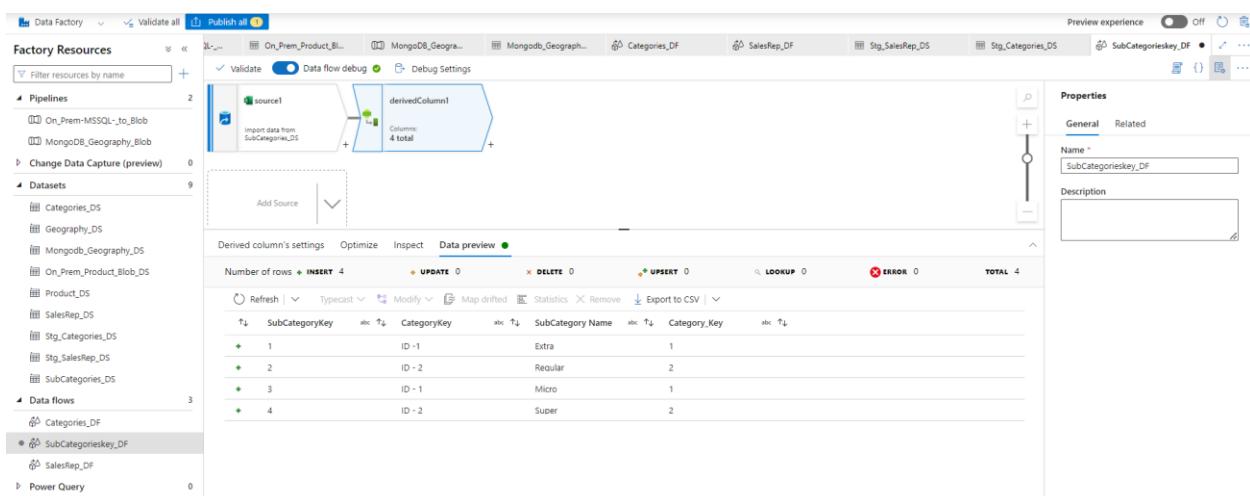
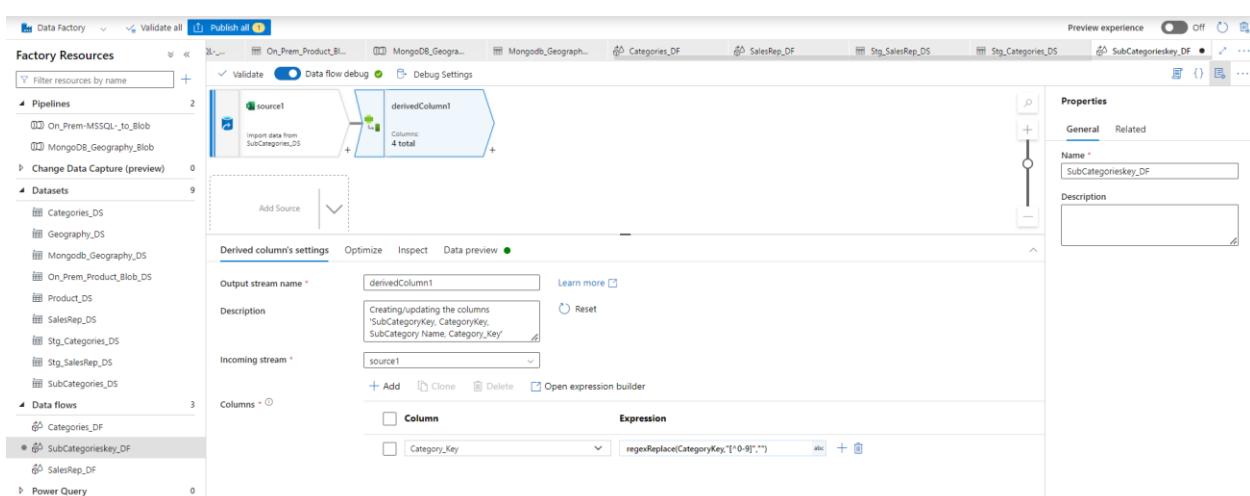
rep

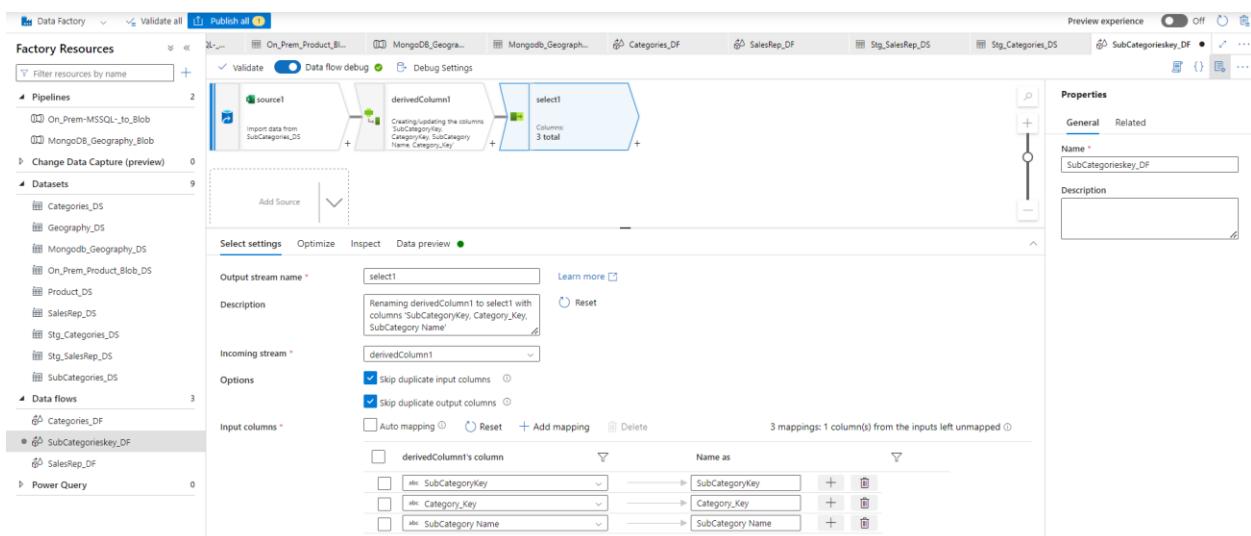
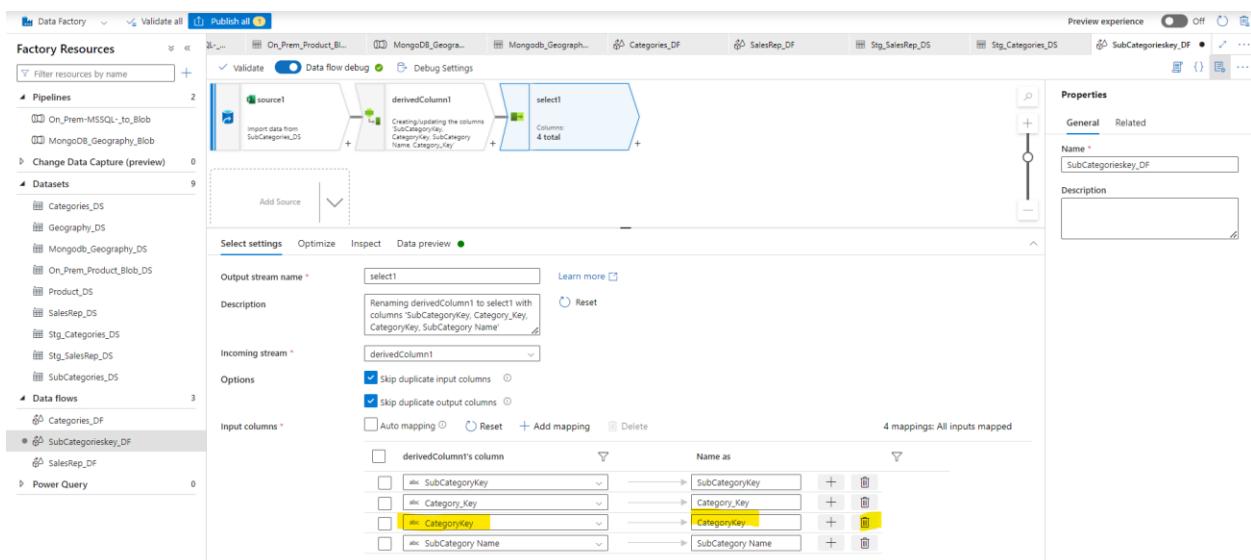
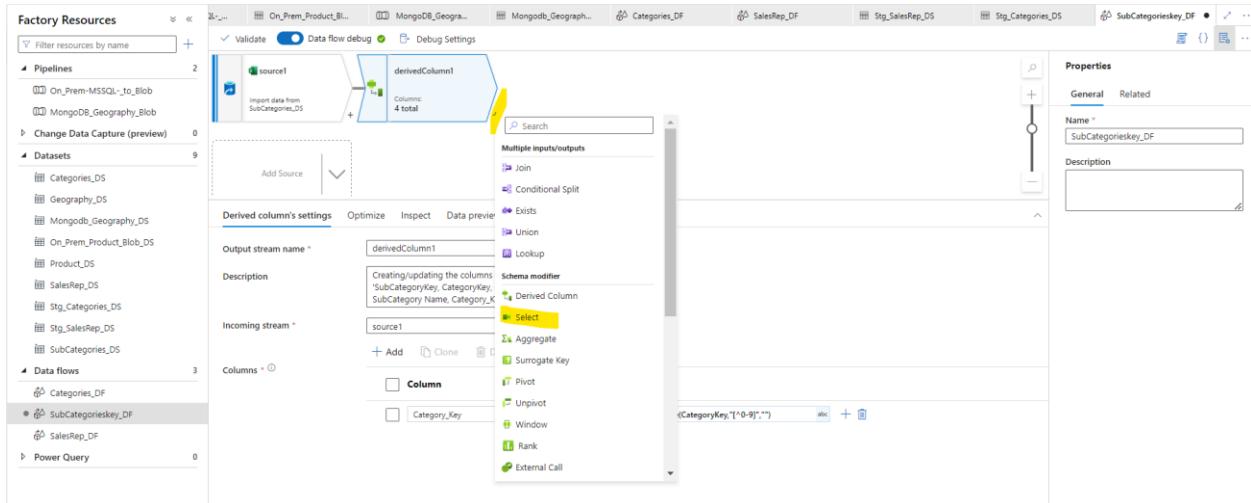
+ Create new ▾

char(rep Input number)
expr(rep Literal expression)
hex(rep Binary value)
regexReplace(rep String, rep Regex to find, rep Substring to replace)
replace(rep String, rep Substring to find, rep Substring to replace)
toBinary(rep Expression)
translate(rep String to translate, rep Lookup characters, rep Replace characters)

Data preview Refresh

Save and finish Cancel Clear contents





Data Flow 1: This flow imports data from SubCategories_DS and performs a derived column operation to rename columns. It then uses a select operation to filter rows. The final step is an aggregate operation.

```

graph LR
    source1[Import data from SubCategories_DS] --> derivedColumn1[derivedColumn1]
    derivedColumn1 --> select1[select1]
    select1 --> aggregate1[aggregate1]
  
```

Data Flow 2: This flow imports data from SubCategories_DS and performs a derived column operation to rename columns. It then uses a select operation to filter rows. The final step is an aggregate operation.

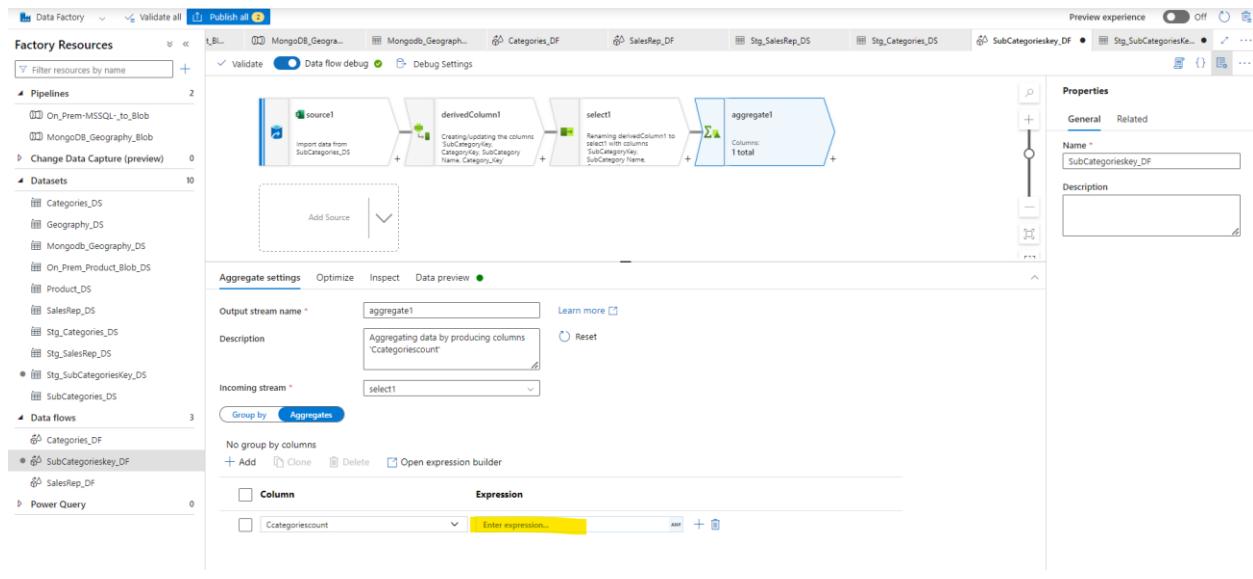
```

graph LR
    source1[Import data from SubCategories_DS] --> derivedColumn1[derivedColumn1]
    derivedColumn1 --> select1[select1]
    select1 --> aggregate1[aggregate1]
  
```

Data Flow 3: This flow imports data from SubCategories_DS and performs a derived column operation to rename columns. It then uses a select operation to filter rows. The final step is an aggregate operation.

```

graph LR
    source1[Import data from SubCategories_DS] --> derivedColumn1[derivedColumn1]
    derivedColumn1 --> select1[select1]
    select1 --> aggregate1[aggregate1]
  
```



Dataflow expression builder

aggregate1

Aggregate Columns

+ Create new

ANY Categoriescount

Column name *
Categoriescount

Expression

```
count(SubCategoryKey)
```

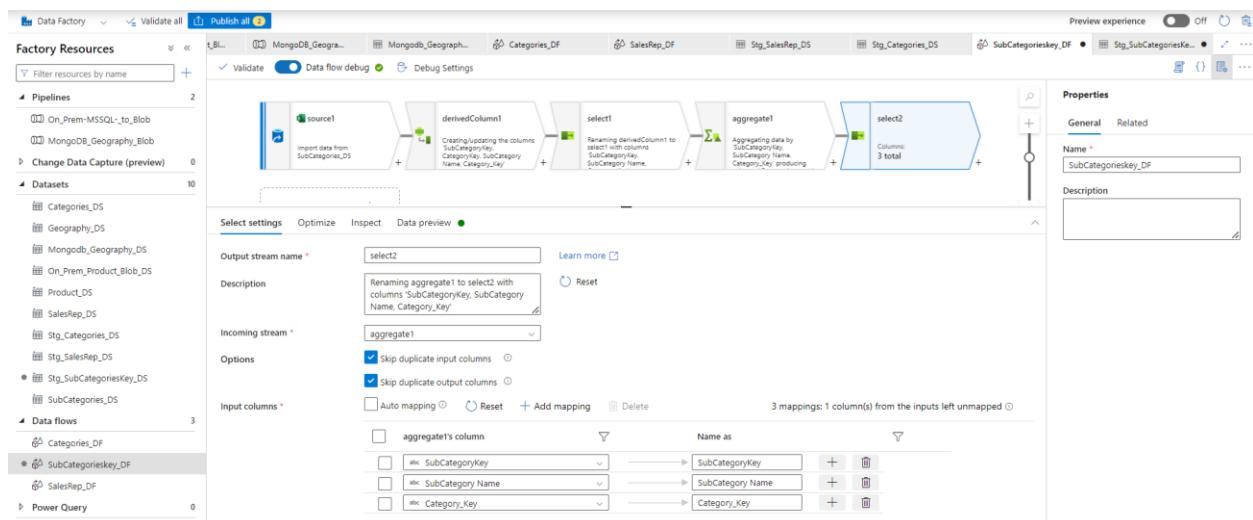
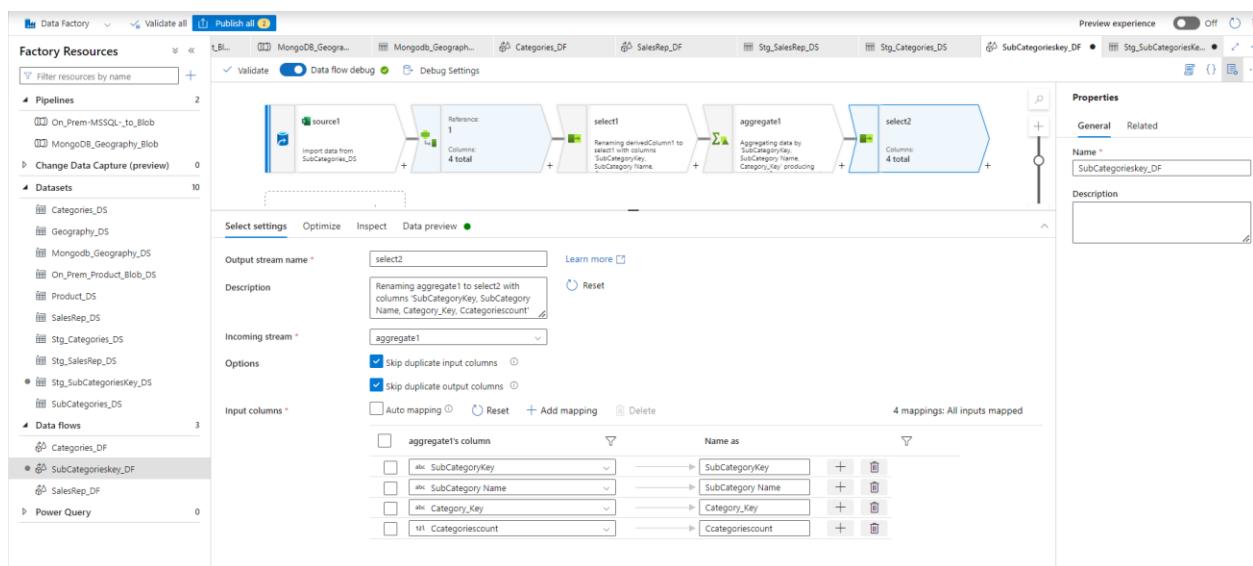
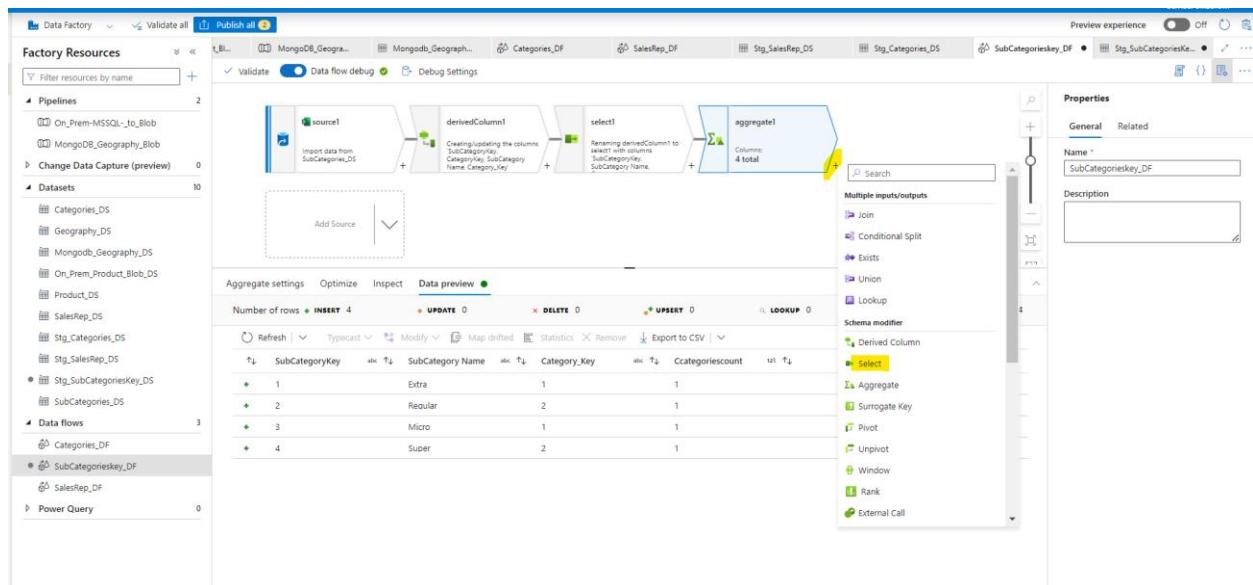
Expression elements

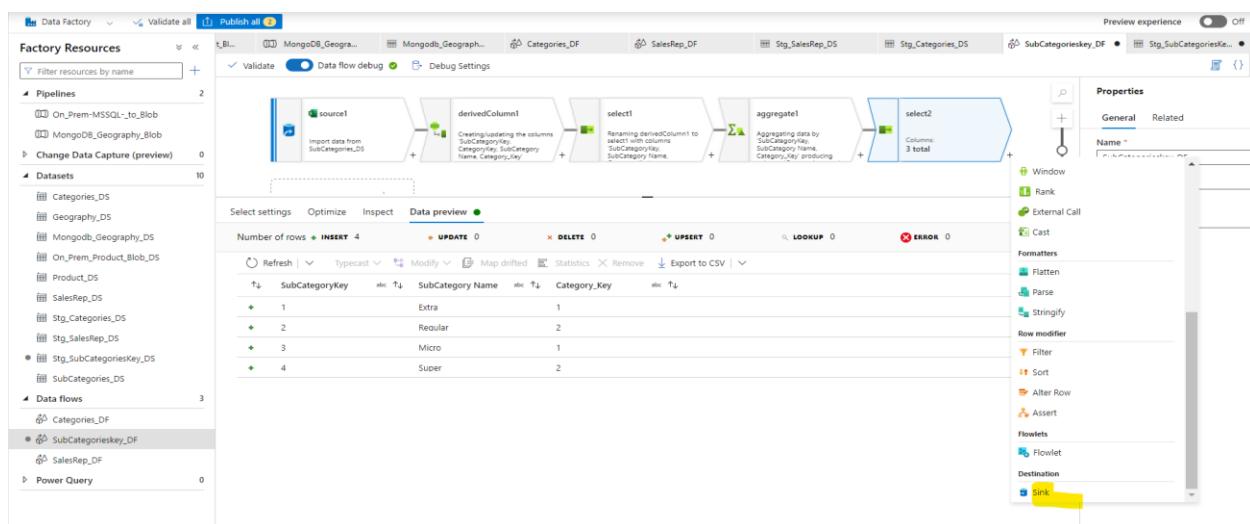
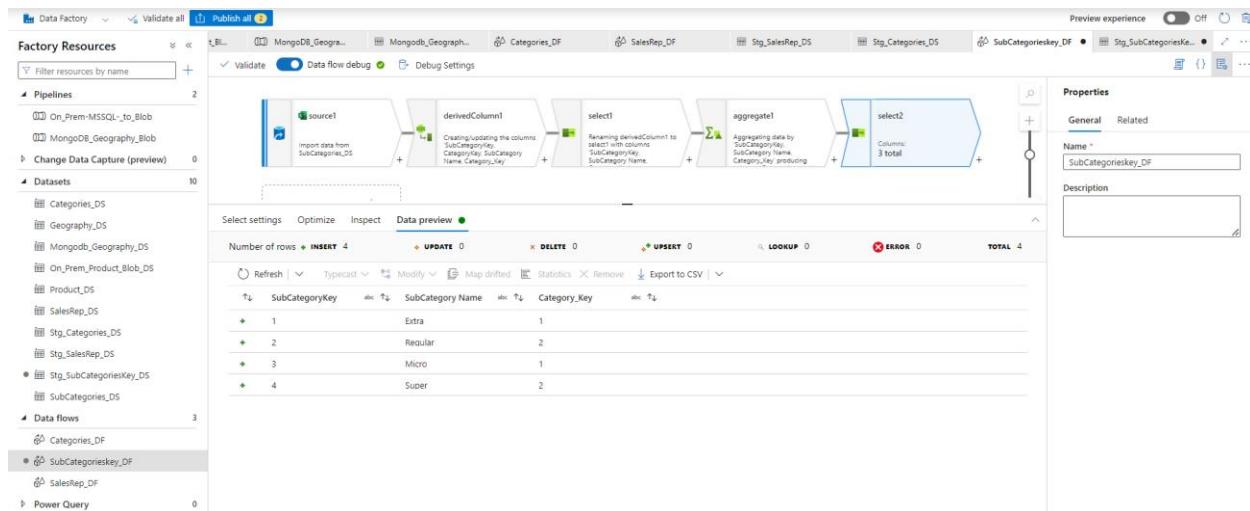
- All
- Functions
- Input schema
- Parameters
- Cached lookup
- Data flow library functions
- Locals

Expression values

- + Create new
- abc SubCategoryKey
- abc SubCategory Name
- abc Category_Key
- 123 abs(123 numeric_value)
- 123 acos(123 numeric_value)
- ANY add(ANY first_expression, ANY second_expression)
- ANY addDays(ANY date/timestamp, ANY days to add)

Save and finish **Cancel** **Clear contents**





New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more ↗

Select a data store

Search:

All Azure Database File Generic protocol NoSQL Services and apps

Amazon S3	Azure Blob Storage	Azure Cosmos DB for NoSQL
Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2	Azure Database for MySQL
Azure Database for PostgreSQL	Azure SQL Database	Azure SQL Database Managed Instance

Sink

Output stream name: sink1

Description: Add sink dataset

Incoming stream: select2

Sink type: Dataset

Dataset: Select...

Options:

- Allow schema drift
- Validate schema

Continue Cancel

Data Factory Pipeline Editor

Pipeline Details: SubCategorieskey_DF

Source: Import data from SubCategories_DS

Operations:

- derivedColumn1: Creating/updating the column 'SubCategoryKey'.
- select1: Renaming derivedColumn1 to select1 with columns 'SubCategoryKey'.
- aggregate1: Aggregating data by 'SubCategoryKey', 'SubCategory Name'.
- select2: Renaming aggregate1 to select2 with columns 'SubCategoryKey'.

Sink: sink1

Description: Add sink dataset

Incoming stream: select2

Sink type: Dataset

Dataset: Select...

Options:
 Allow schema drift
 Validate schema

Format Selection: Choose the format type of your data

Avro	DelimitedText	JSON
ORC	Parquet	Binary

Buttons: Continue, Back, Cancel

Data Factory Pipeline Editor

Pipeline Details: SubCategorieskey_DF

Source: Import data from SubCategories_DS

Operations:

- derivedColumn1: Creating/updating the column 'SubCategoryKey'.
- select1: Renaming derivedColumn1 to select1 with columns 'SubCategoryKey'.
- aggregate1: Aggregating data by 'SubCategoryKey', 'SubCategory Name'.
- select2: Renaming aggregate1 to select2 with columns 'SubCategoryKey'.

Sink: sink1

Description: Add sink dataset

Incoming stream: select2

Sink type: Dataset

Dataset: Select...

Options:
 Allow schema drift
 Validate schema

Set Properties:

Name: Stg_SubCategorieskey_DS1

Linked service: BLOB_ADF_LS

File path: stagingdata / Directory / File name

Import schema:
 From connection/store
 From sample file
 None

Advanced:

First row as header:

Buttons: OK, Back, Cancel

Data Factory Pipeline:

Stg_SubCategoriesKey_DS (DelimitedText CSV) -> **BLOB_ADF_LS** (File path: stagingdata / Stg_SubCategories.csv)

Properties:

- Name: Stg_SubCategoriesKey_DS
- Description:
- Annotations: + New

Data Flow:

```

graph LR
    source1((source1)) -- "Import data from SubCategories_DS" --> derivedColumn1[derivedColumn1]
    derivedColumn1 -- "Creating using expression: SubCategoryKey, CategoryName, SubCategoryName" --> select1[select1]
    select1 -- "Reusing derivedColumn1 to select with columns SubCategoryName" --> aggregate[aggregate]
    aggregate -- "Aggregating data by SubCategoryKey, CategoryName, SubCategoryName" --> select2[select2]
    select2 -- "Reusing aggregated to select2 with columns SubCategoryKey, SubCategoryName" --> sink1[sink1]
    sink1 -- "3 total" --> end(( ))
  
```

Properties:

- Name: SubCategorieskey_DF
- Description:

MongoDB Data flow:

Data Flow:

```

graph LR
    source1((source1)) -- "Columns: 6 total" --> Geography_DF[Geography_DF]
  
```

Properties:

- Name: Geography_DF
- Description:

Data Factory Pipeline Editor

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

source1

Properties

- General
- Related
- Name: Geography_DF
- Description

Source settings

Source type: Dataset

Output stream name: source1

Schema modifier: Derived Column

Sampling: 0

Source options

Projection

Optimize

Inspect

Data preview

Test connection

General

Source settings

Source type: Dataset

Output stream name: source1

Schema modifier: Derived Column

Sampling: 0

Source options

Projection

Optimize

Inspect

Data preview

Test connection

General

Source settings

Source type: Dataset

Output stream name: source1

Schema modifier: Derived Column

Sampling: 0

Source options

Projection

Optimize

Inspect

Data preview

Test connection

Data Factory Pipeline Editor

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

source1

Properties

- General
- Related
- Name: Geography_DF
- Description

Source settings

Source type: Dataset

Output stream name: source1

Schema modifier: Derived Column

Sampling: 0

Source options

Projection

Optimize

Inspect

Data preview

Test connection

General

Source settings

Source type: Dataset

Output stream name: source1

Schema modifier: Derived Column

Sampling: 0

Source options

Projection

Optimize

Inspect

Data preview

Test connection

General

Source settings

Source type: Dataset

Output stream name: source1

Schema modifier: Derived Column

Sampling: 0

Source options

Projection

Optimize

Inspect

Data preview

Test connection

Data Factory Pipeline Editor

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

source1

Properties

- General
- Related
- Name: Geography_DF
- Description

Source settings

Source type: Dataset

Output stream name: source1

Schema modifier: Derived Column

Sampling: 0

Source options

Projection

Optimize

Inspect

Data preview

Define default format

Column name

Column name	Type	Format
Country	string	Specify format
Town	string	Specify format
Wikipedia	string	Specify format

General

Source settings

Source type: Dataset

Output stream name: source1

Schema modifier: Derived Column

Sampling: 0

Source options

Projection

Optimize

Inspect

Data preview

Define default format

Column name

Column name	Type	Format
Country	string	Specify format
Town	string	Specify format
Wikipedia	string	Specify format

Data Factory Pipeline Overview

Pipeline Details: Geography_DF

Source: Import data from Mongodb_Geography_DS

Transformation: select1 (Columns: 3 total)

Properties:

- Name: Geography_DF
- Description:

Input Mappings:

Source Column	Output Column
abc_Country	Country
abc_Town	Town
abc_Wikipedia	Wikipedia

Data Factory Pipeline Overview

Pipeline Details: Geography_DF

Source: Import data from Mongodb_Geography_DS

Transformation: select1 (Columns: 2 total)

Properties:

- Name: Geography_DF
- Description:

Input Mappings:

Source Column	Output Column
abc_Country	Country
abc_Town	Town

Data Factory Pipeline Overview

Pipeline Details: Geography_DF

Source: Import data from Mongodb_Geography_DS

Transformation: select1 (Columns: 2 total)

Properties:

- Name: Geography_DF
- Description:

Input Mappings:

Source Column	Output Column
abc_Country	Country
abc_Town	Town

Schema Modifier Options:

- Join
- Conditional Split
- Exists
- Union
- Lookup
- Derived Column
- Select
- Aggregate
- Surrogate key
- Pivot
- Unpivot
- Window
- Rank
- External Call

Data Flow Editor - Geography_DF

```

graph LR
    source1[Import data from Mongodb_Geography_DS] --> select1[Renaming source1 to select1 with columns "Country", "Town"]
    select1 --> AddingGeographyID[Adding new key Geography_ID starting from 1 with step 1]

```

Data Preview - Geography_DF

Country	Town	Geography_ID
Czech republic	Prague	1
Denmark	Copenhagen	2
Germany	Frankfurt	3
Germany	Berlin	4
Czech republic	Dresden	5
Czech republic	Brno	6
Czech republic	Ostrava	7

Data Flow Editor - Geography_DF

```

graph LR
    source1[Import data from Mongodb_Geography_DS] --> select1[Renaming source1 to select1 with columns "Country", "Town"]
    select1 --> AddingGeographyID[Adding new key Geography_ID starting from 1 with step 1]
    AddingGeographyID --> select2[Adding Geography_ID's column to select2 with columns "Geography_ID", "Country", "Town"]

```

New dataset

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more ▾

Select a data store

All Azure Database File Generic protocol NoSQL Services and apps

Continue Cancel

Choose the format type of your data

Avro CSV JSON

ORC Parquet Binary

Continue Back Cancel

Name: Stg_Geography_DS

Linked service: Blob_ADF_L5

File path: stagingdata / Directory / File name

First row as header:

Import schema: From connection/store From sample file None

Advanced

OK Back Cancel

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_Geography_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

Properties

General

Name: Stg_Geography_DS

Description:

Annotations

+ New

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_Geography_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

Properties

General

Name: Geography_DS

Description:

Number of rows	INSERT	UPDATE	DELETE	UPSERT	LOOKUP	ERROR	TOTAL
7							7
Data preview							
Geography_ID Country Town 1 Czech republic Prague 2 Denmark Copenhagen 3 Germany Frankfurt 4 Germany Berlin 5 Germany Dresden 6 Czech republic Brno 7 Czech republic Ostrava							

PRODUCT DATA FLOW:

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_Geography_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - Product_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

Properties

General

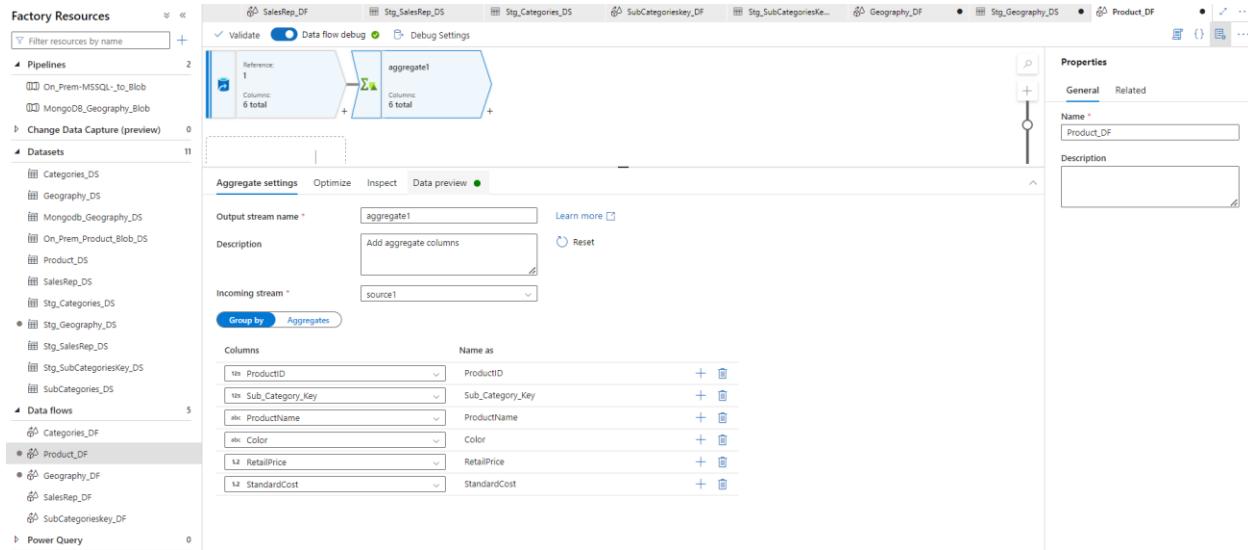
Name: Product_DF

Description:

Data Flow Editor (Top):

Data Preview (Bottom):

ProductID	Sub_Category_Key	Color	ProductName	RetailPrice	StandardCost
1	3	Red	Alder	23.950007629394...	7.550000190734863
2	2	Blue	Linder	23.950007629394...	7.550000190734863
3	2	Green	Magnum	23.950007629394...	7.550000190734863
4	1	Red	Quad	43.9500076293945	13.75
5	1	Blue	Black Monk	43.9500076293945	13.75
6	4	Green	Quad	43.9500076293945	13.75
7	1	Red	Bind	26.950007629394...	8.25
8	3	Blue	VanHelen	26.950007629394...	8.25
9	1	Green	Magnum	26.950007629394...	8.25
10	1	Florescent Pink	Carlota	29.950007629394...	9.149999618530273
11	4	Florescent Blue	Carlota	29.950007629394...	9.149999618530273
7	1	Red	Bing	26.950007629394...	8.25
10	1	Florescent Pink	Carlota	29.950007629394...	9.149999618530273



Dataflow expression builder

aggregate1

Aggregate Columns

Create new CProductID

Column name: CProductID

Expression

```
count(ProductID)
```

Save

Expression elements

Expression values
All
Functions
Input schema
Parameters
Cached lookup
Data flow library functions
Locals
ProductID
Sub_Catgeory_Key
Color
ProductName
RetailPrice
StandardCost
abs(numeric_value)

Factory Resources

- Pipelines
- Datasets
- Data flows
- Power Query

SalesRep_DF Data flow debug Debug Settings

Properties

- General
- Related
- Name: Product_DF
- Description

Aggregate settings

Output stream name: aggregate1

Description: Aggregating data by 'ProductID, Sub_Catgeory_Key, ProductName, Color, RetailPrice, StandardCost'

Incoming stream: source1

Group by Aggregates

Column Expression

```
count(ProductID)
```

Grouped by: ProductID, Sub_Catgeory_Key, ProductName, Color, RetailPrice, StandardCost

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_Geography_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - Product_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

SalesRep_DF

Validate Data flow debug Debug Settings

Properties

General Related

Name: Product_DF

Description:

Data preview

Number of rows: 11

ProductID	Sub_Categor...	ProductName	Color	RetailPrice	StandardCost	CP...
1	3	Alder	Red	23.95000762...	7.550001907...	1
2	2	Linder	Blue	23.95000762...	7.550001907...	1
3	2	Magnum	Green	23.95000762...	7.550001907...	1
4	1	Quad	Red	43.95000762...	13.75	1
5	1	Black Monk	Blue	43.95000762...	13.75	1
6	4	Quad	Green	43.95000762...	13.75	1
7	1	Bing	Red	26.95000762...	8.25	2
8	3	VanHelen	Blue	26.95000762...	8.25	1
9	1	Magnum	Green	26.95000762...	8.25	1
10	1	Carlota	Florescent Pink	29.95000762...	9.149996185...	2
11	4	Carlota	Florescent Blue	29.95000762...	9.149996185...	1

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_Geography_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - Product_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

SalesRep_DF

Validate Data flow debug Debug Settings

Properties

General Related

Name: Product_DF

Description:

Select settings

Output stream name: select1

Description: Renaming aggregate1 to select1 with columns ProductID, Sub_Categor..., ProductName.

Incoming stream: aggregate1

Options: Skip duplicate input columns, Skip duplicate output columns

Input columns: Auto mapping, Reset, Add mapping, Delete

Data preview

Number of rows: 11

ProductID	Sub_Categor...	ProductName	Color	RetailPrice	StandardCost	CP...
1	3	Alder	Red	23.95000762...	7.550001907...	1
2	2	Linder	Blue	23.95000762...	7.550001907...	1
3	2	Magnum	Green	23.95000762...	7.550001907...	1
4	1	Quad	Red	43.95000762...	13.75	1
5	1	Black Monk	Blue	43.95000762...	13.75	1
6	4	Quad	Green	43.95000762...	13.75	1
7	1	Bing	Red	26.95000762...	8.25	2
8	3	VanHelen	Blue	26.95000762...	8.25	1
9	1	Magnum	Green	26.95000762...	8.25	1
10	1	Carlota	Florescent Pink	29.95000762...	9.149996185...	2
11	4	Carlota	Florescent Blue	29.95000762...	9.149996185...	1

Factory Resources

- Pipelines
 - On_Prem-MSSQL_to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_Geography_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data flows
 - Categories_DF
 - Product_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

SalesRep_DF

Validate Data flow debug Debug Settings

Properties

General Related

Name: Product_DF

Description:

Select settings

Output stream name: select1

Description: Renaming aggregate1 to select1 with columns ProductID, Sub_Categor..., ProductName.

Incoming stream: aggregate1

Options: Skip duplicate input columns, Skip duplicate output columns

Input columns: Auto mapping, Reset, Add mapping, Delete

Data preview

Number of rows: 11

ProductID	Sub_Categor...	ProductName	Color	RetailPrice	StandardCost
1	3	Alder	Red	23.95000762...	7.550001907...
2	2	Linder	Blue	23.95000762...	7.550001907...
3	2	Magnum	Green	23.95000762...	7.550001907...
4	1	Quad	Red	43.95000762...	13.75
5	1	Black Monk	Blue	43.95000762...	13.75
6	4	Quad	Green	43.95000762...	13.75
7	1	Bing	Red	26.95000762...	8.25
8	3	VanHelen	Blue	26.95000762...	8.25
9	1	Magnum	Green	26.95000762...	8.25
10	1	Carlota	Florescent Pink	29.95000762...	9.149996185...
11	4	Carlota	Florescent Blue	29.95000762...	9.149996185...

Factory Resources

- Pipelines
 - On_Prem-MSSQL-to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_Geography_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data Flows
 - Categories_DF
 - Product_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

SalesRep_DF

Data preview

ProductID	Sub_Category_Key	ProductName	Color	RetailPrice	StandardCost
1	3	Alder	Red	23.9500007629394...	7.550000190734863
2	2	Linder	Blue	23.9500007629394...	7.550000190734863
3	2	Maonum	Green	23.9500007629394...	7.550000190734863
4	1	Quad	Red	43.95000076293945	13.75
5	1	Black Monk	Blue	43.95000076293945	13.75
6	4	Quad	Green	43.95000076293945	13.75
7	1	Bino	Red	26.9500007629394...	8.25
8	3	VanHelen	Blue	26.9500007629394...	8.25
9	1	Maonum	Green	26.9500007629394...	8.25
10	1	Carlota	Florescent Pink	29.9500007629394...	9.149999618530273
11	4	Carlota	Florescent Blue	29.9500007629394...	9.149999618530273

Properties

- General
- Related

Name: Product_DF

Description:

Factory Resources

- Pipelines
 - On_Prem-MSSQL-to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_Geography_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data Flows
 - Categories_DF
 - Product_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

SalesRep_DF

Sink

Output stream name: sink1

Description: Add sink dataset

Incoming stream: select1

Sink type: Dataset

Dataset: Select...

Options: Allow schema drift (checked), Validate schema

Properties

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more [\[\]](#)

Select a data store

All Azure Database File Generic protocol NoSQL Services and apps

Amazon S3	Azure Blob Storage	Azure Cosmos DB for NoSQL
Azure Data Explorer (Kusto)	Azure Data Lake Storage Gen2	Azure Database for MySQL
Azure Database for PostgreSQL	Azure SQL Database	Azure SQL Database Managed Instance

Continue Cancel

Factory Resources

- Pipelines
 - On_Prem-MSSQL-to_Blob
 - MongoDB_Geography_Blob
- Change Data Capture (preview)
- Datasets
 - Categories_DS
 - Geography_DS
 - Mongodb_Geography_DS
 - On_Prem_Product_Blob_DS
 - Product_DS
 - SalesRep_DS
 - Stg_Categories_DS
 - Stg_Geography_DS
 - Stg_SalesRep_DS
 - Stg_SubCategoriesKey_DS
 - SubCategories_DS
- Data Flows
 - Categories_DF
 - Product_DF
 - Geography_DF
 - SalesRep_DF
 - SubCategorieskey_DF
- Power Query

SalesRep_DF

Sink

Output stream name: sink1

Description: Add sink dataset

Incoming stream: select1

Sink type: Dataset

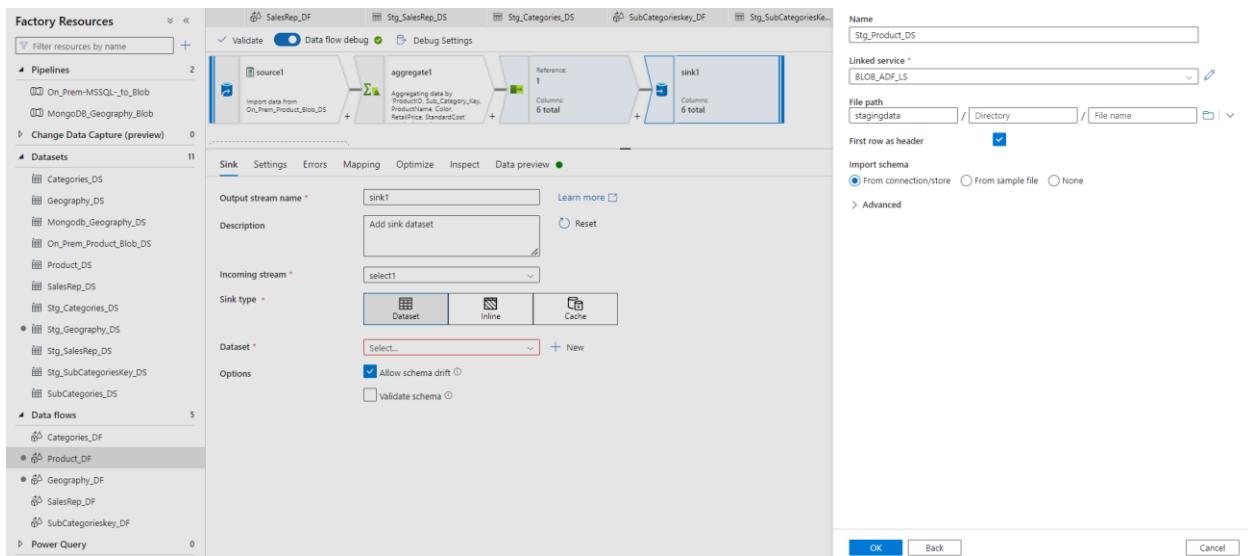
Dataset: Select...

Options: Allow schema drift (checked), Validate schema

Choose the format type of your data

Avro	DelimitedText	JSON
ORC	Parquet	Binary

Continue Back Cancel



Stg_Product_DS

Properties
General
Name: Stg_Product_DS
Description:
Annotations: + New

SalesRep_DF

Properties
General
Name: Product_DF
Description:

Number of rows	INSERT N/A	UPDATE N/A	DELETE N/A	UPSERT N/A	LOOKUP N/A	ERROR N/A	TOTAL 11				
ProductID	asc ↑	Sub_Categroy_Key	asc ↑	ProductName	asc ↑	Color	asc ↑	RetailPrice	asc ↑	StandardCost	asc ↑
1		3		Alder		Red		23.950000762999453		7.550000190734863	
2		2		Linder		Blue		23.950000762999453		7.550000190734863	
3		2		Maonum		Green		23.950000762999453		7.550000190734863	
4		1		Quad		Red		43.95000076299945		13.75	
5		1		Black Monk		Blue		43.95000076299945		13.75	
6		4		Quad		Green		43.95000076299945		13.75	
7		1		Bing		Red		26.950000762999453		8.25	
8		3		VanHelen		Blue		26.950000762999453		8.25	
9		1		Maonum		Green		26.950000762999453		8.25	
10		1		Carlota		Florescent Pink		29.950000762999453		9.149999618530273	
11		4		Carlota		Florescent Blue		29.950000762999453		9.149999618530273	

Create Pipelines:

The screenshot shows three separate Azure Data Factory pipeline configurations, each with its own properties and activities.

Pipeline 1: Categories_PL

- Properties:**
 - Name: Categories_PL
 - Description: (empty)
 - Activity state: Activated
 - Timeout: 0:12:00:00
 - Retry: 0
 - Retry interval (sec): 30
 - Secure output: (checkbox)
 - Secure input: (checkbox)
- Activities:**
 - Categories data flow (Data Flow activity)

Pipeline 2: Categories_DF

- Properties:**
 - Name: Categories_DF
 - Description: (empty)
 - Run on (Azure IR): AutoResolveIntegrationRuntime
 - Compute size: Small
 - Logging level: Verbose
- Activities:**
 - Categories data flow (Data Flow activity)

Pipeline 3: SubCategoriesKey_PL

- Properties:**
 - Name: SubCategoriesKey_PL
 - Description: (empty)
- Activities:**
 - SubCategoriesKey data Flow (Data Flow activity)

Factory Resources

- Pipelines
 - Categories_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - SubCategoriesKey_PL
 - pipeline2
- Change Data Capture (preview) 0
- Datasets 12
- Data flows 5
- Power Query 0

Activities

- Move and transform
 - Copy data
 - Data flow
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Properties

General

Name: SubCategoriesKey_PL

Description:

Annotations

+ New

Settings

Data flow: SubCategorieskey_DF

Run on (Azure IR): AutoResolveIntegrationRuntime

Compute size: Small

Logging level: Verbose

> Sink properties

> Staging

Factory Resources

- Pipelines
 - Categories_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - SubCategoriesKey_PL
 - Product_PL
- Change Data Capture (preview) 0
- Datasets 12
- Data flows 5
- Power Query 0

Activities

- Move and transform
 - Copy data
 - Data flow
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Properties

General

Name: Product_PL

Description:

Annotations

+ New

Settings

Name: Product Data flow

Description:

Activity state: Activated

Timeout: 0:12:00:00

Retry: 0

Retry interval (sec): 30

Secure output:

Secure input:

Factory Resources

- Pipelines
 - Categories_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - SubCategoriesKey_PL
 - Product_PL
- Change Data Capture (preview) 0
- Datasets 12
- Data flows 5
- Power Query 0

Activities

- Move and transform
 - Copy data
 - Data flow
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Properties

General

Name: Product_PL

Description:

Annotations

+ New

Settings

Data flow: Product_DF

Run on (Azure IR): AutoResolveIntegrationRuntime

Compute size: Small

Logging level: Verbose

> Sink properties

> Staging

Factory Resources

- Pipelines
 - Categories_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - SalesRep_PL
 - pipeline2
 - Product_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 12
- Data flows 5
- Power Query 0

Activities

- Move and transform
 - Copy data
 - Data flow
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Data flow

Properties

General

Name: Sales Rep Data flow

Description:

Activity state: Activated

Timeout: 0:12:00:00

Retry: 0

Retry interval (sec): 30

Secure output:

Secure input:

Factory Resources

- Pipelines
 - Categories_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - SalesRep_PL
 - pipeline2
 - Product_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 12
- Data flows 5
- Power Query 0

Activities

- Move and transform
 - Copy data
 - Data flow
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Data flow

Properties

General

Name: SalesRep_DF

Run on (Azure IR): AutoResolveIntegrationRuntime

Compute size: Small

Logging level: Verbose

Advanced

Sink properties

Staging

Factory Resources

- Pipelines
 - Categories_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - Geography_PL
 - SalesRep_PL
 - pipeline2
 - Product_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 12
- Data flows 5
- Power Query 0

Activities

- Move and transform
 - Copy data
 - Data flow
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Data flow

Properties

General

Name: Geography PL

Description:

Activity state: Activated

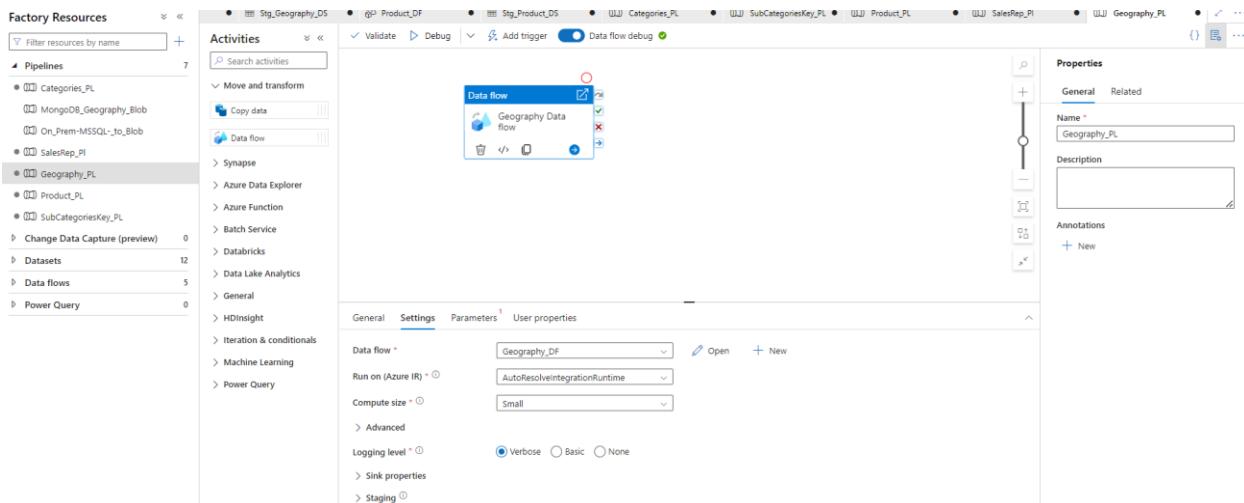
Timeout: 0:12:00:00

Retry: 0

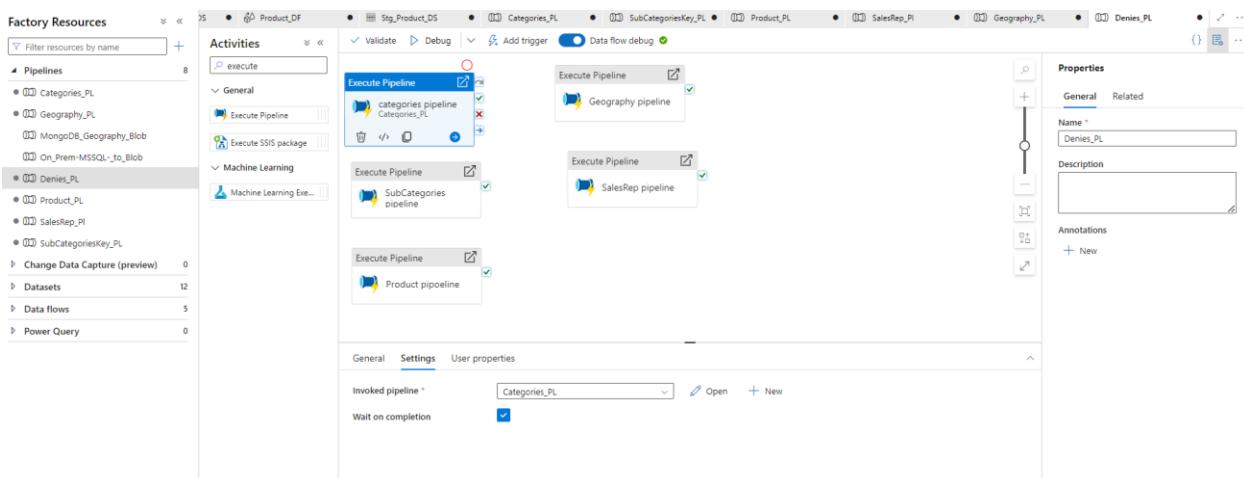
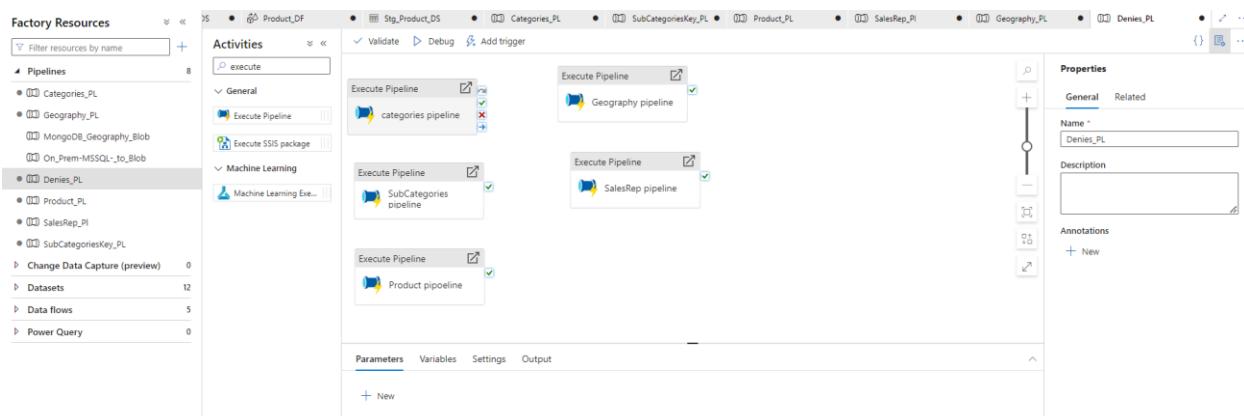
Retry interval (sec): 30

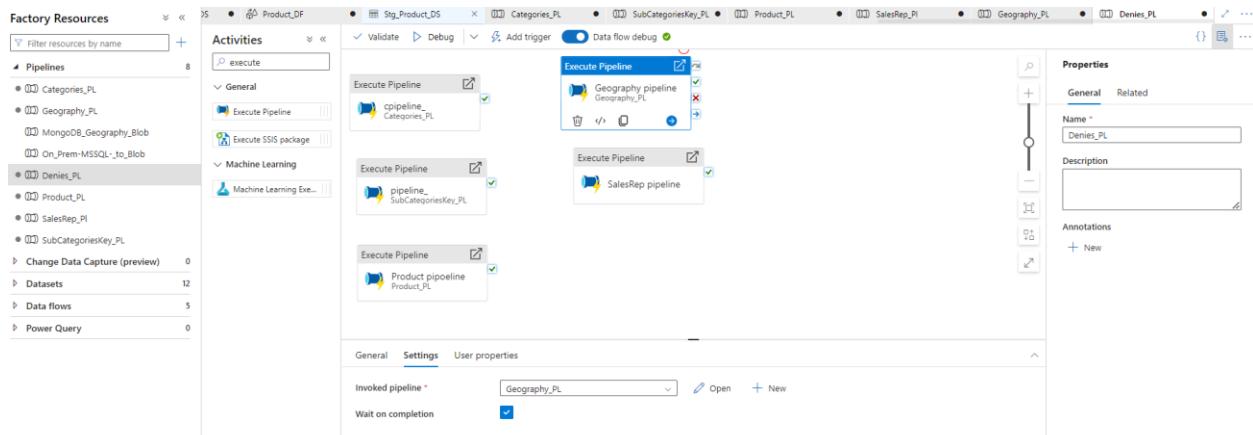
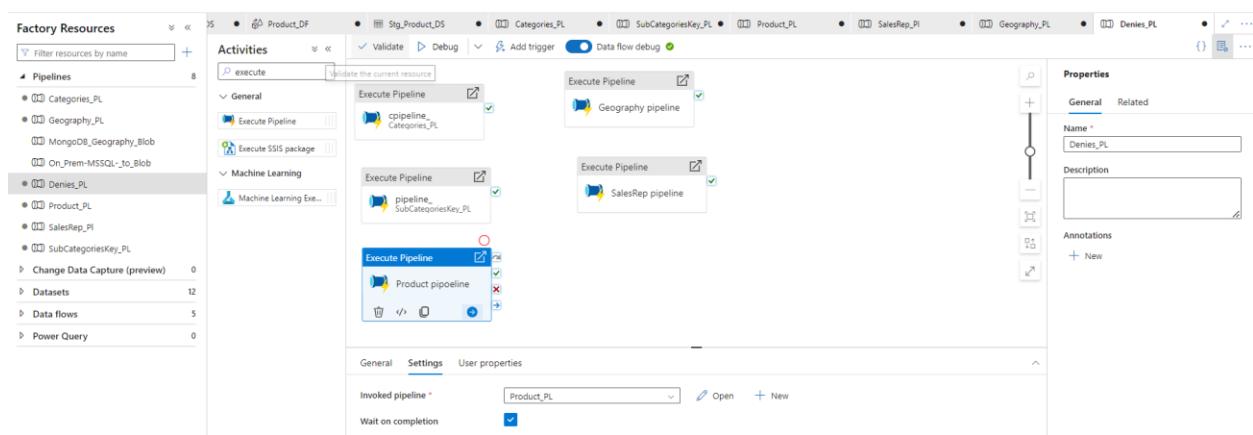
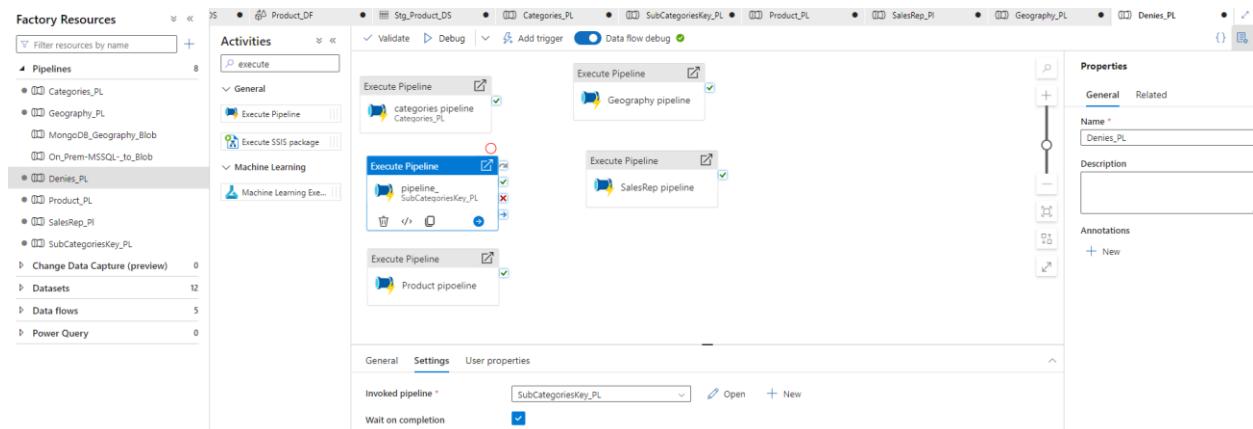
Secure output:

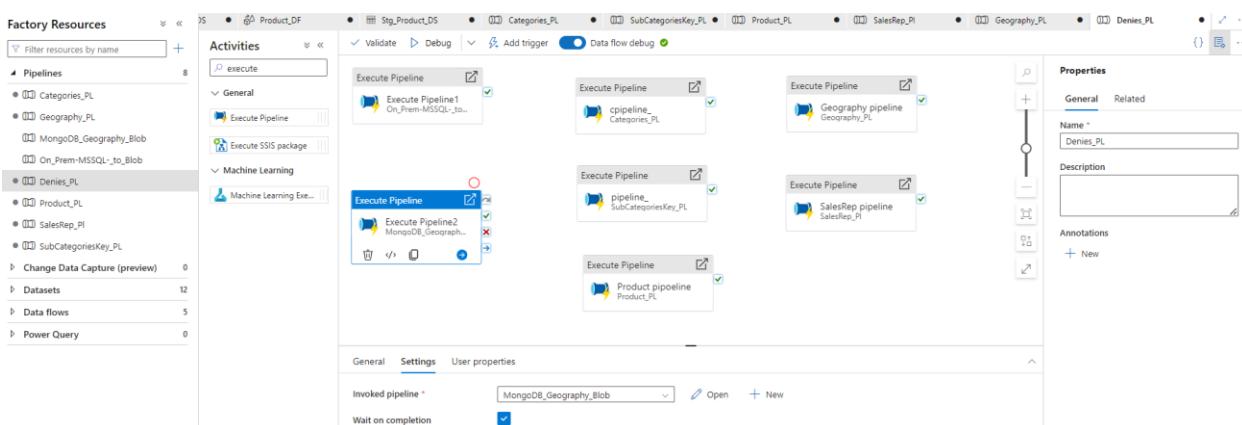
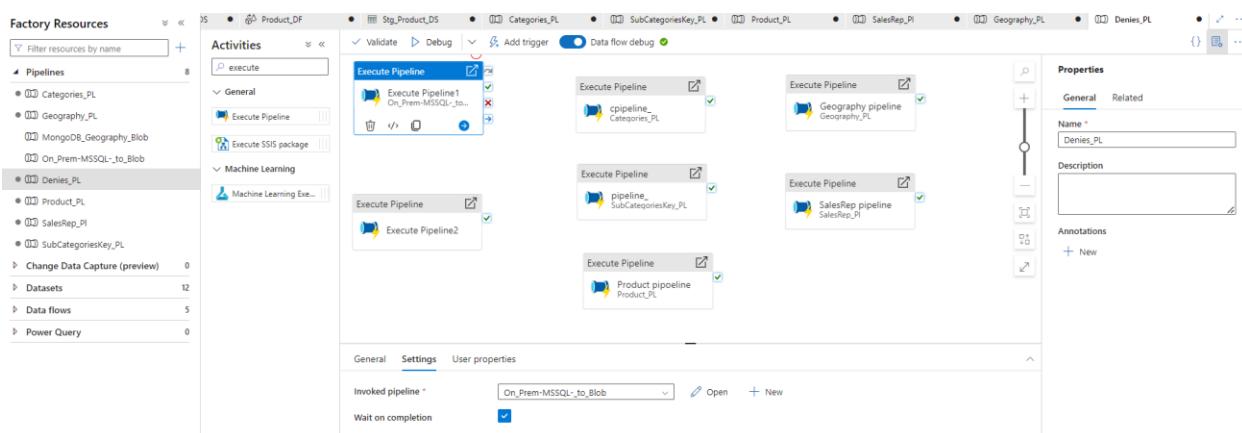
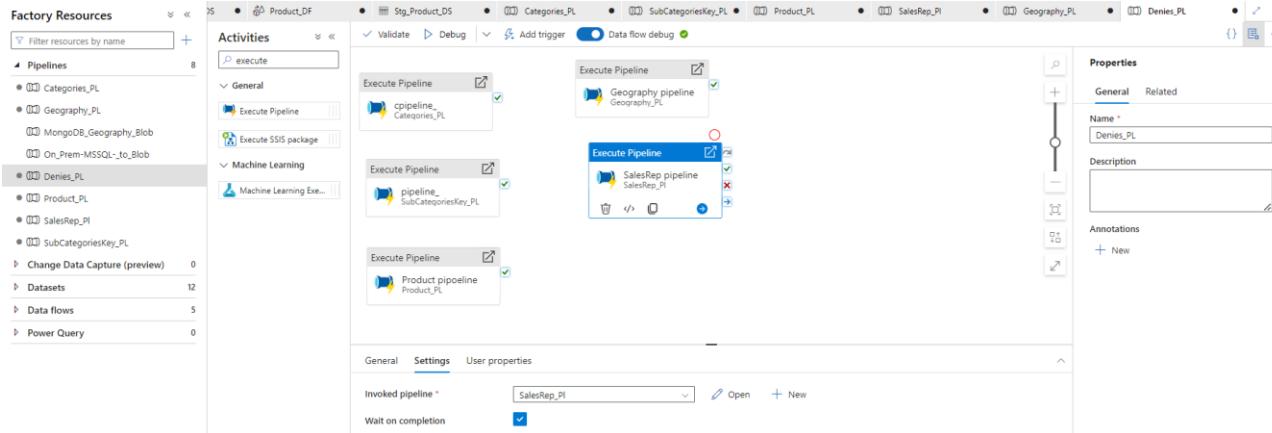
Secure input:

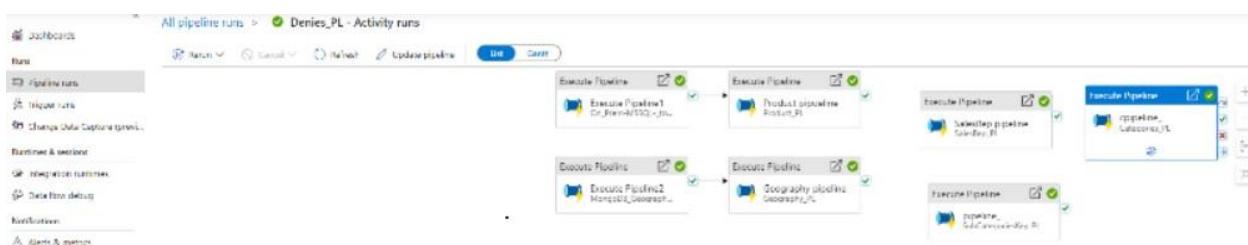
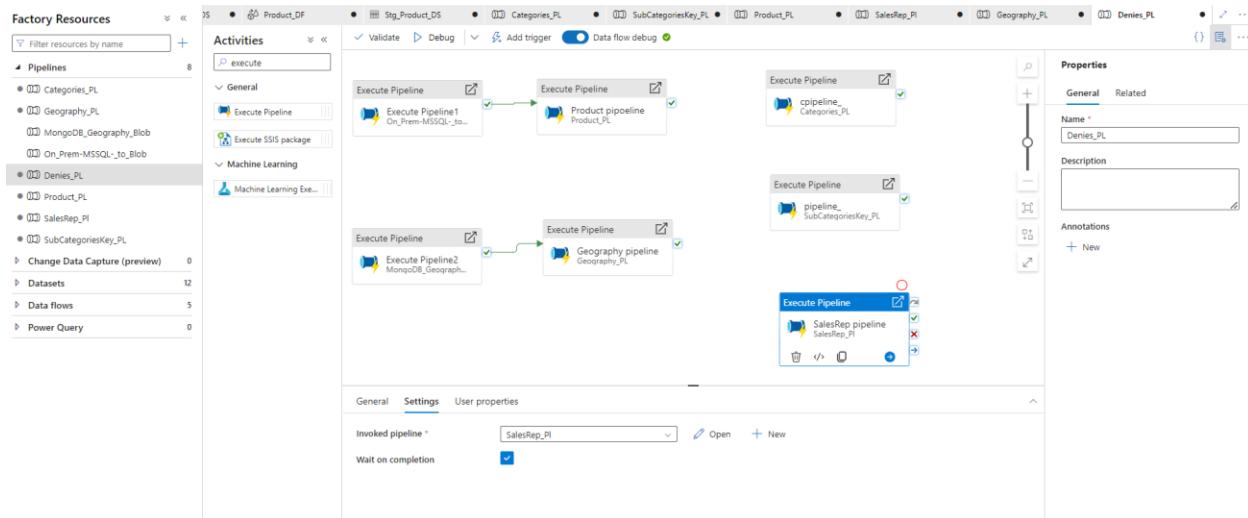


Denies pipeline Creation:









Activity runs

Pipeline run ID: 02446220-1224-4a12-9f5b-22c0d

All status

Showing 1 - 7 items

Activity name	Activity status	Activity type	Run start	Duration	Integration runtime	User properties	Activity run ID	Log
Product pipeline	Successful	Compose Pipeline	4/26/2024, 11:57:07 PM	2m 33s			5fb6c15b-43f5-4e43-81ca-595d1154b538	
Geography pipeline	Successful	Compose Pipeline	4/26/2024, 11:57:24 PM	3m 14s			23beec5-40fd-46f3-8053-375c80057668	
pipeline_	Successful	Compose Pipeline	4/26/2024, 11:57:27 PM	3m 27s			07c251b-77b4-4939-b6b9-f4f514c956	
Compose Pipeline2	Successful	Compose Pipeline	4/26/2024, 11:57:27 PM	1s			ee550024-4059-4f31-90f9-4e59ec2853ff	
SalesRep pipeline	Successful	Compose Pipeline	4/26/2024, 11:57:27 PM	3m 29s			8844f59b-23f7-4a21-9f1e-1e7710d1c62	
Compose Pipeline1	Successful	Compose Pipeline	4/26/2024, 11:57:27 PM	2s			2223816-308c-4853-9d21-028800ef110	
pipeline_	Successful	Compose Pipeline	4/26/2024, 11:57:27 PM	3m 20s			c165254b-07b4-479c-8c3b-545f9cc42f5	

stagingdata Container

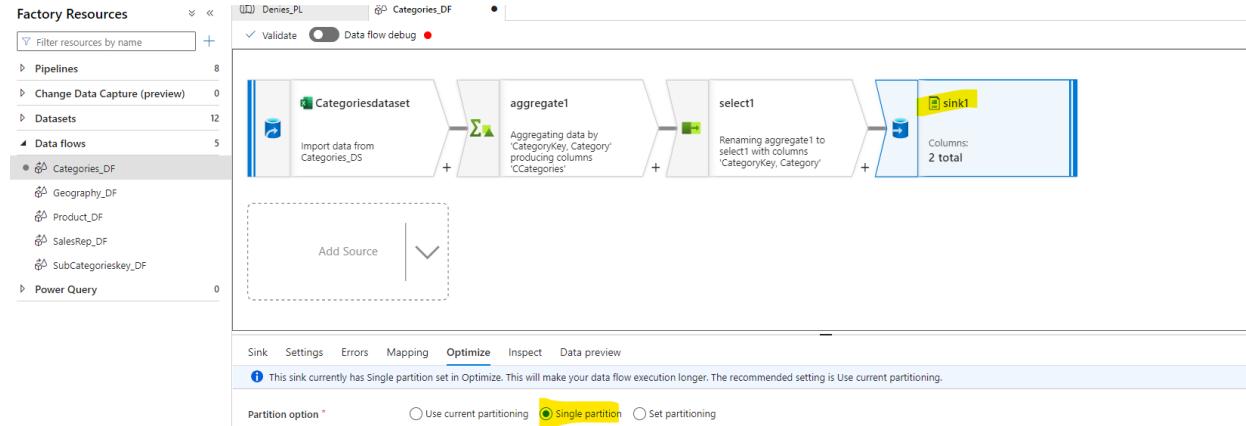
Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
_SUCCESS	29/04/2024, 00:01:04	Hot (Inferred)		Block blob	0 B	Available
part-00000-5be493bc-c27b-4fda-8f08-924fb2...	29/04/2024, 00:00:59	Hot (Inferred)		Block blob	118 B	Available
part-00000-61adfcc6-8945-4f68-9781-07b5bb2b...	29/04/2024, 00:01:03	Hot (Inferred)		Block blob	173 B	Available
part-00000-8db0f76e-df96-455a-ac86-0e9e329e...	29/04/2024, 00:00:43	Hot (Inferred)		Block blob	70 B	Available
part-00000-9a3944ae-9daf-4fe4-9eea-fda79456a...	29/04/2024, 00:00:52	Hot (Inferred)		Block blob	45 B	Available
part-00000-9cbc7a7d-84be-406d-bd7e-5d4708e...	29/04/2024, 00:00:48	Hot (Inferred)		Block blob	21 B	Available
part-00002-8db0f76e-df96-455a-ac86-0e9e329e...	29/04/2024, 00:00:42	Hot (Inferred)		Block blob	107 B	Available
part-00013-9a3944ae-9daf-4fe4-9eea-fda79456a...	29/04/2024, 00:00:51	Hot (Inferred)		Block blob	55 B	Available
part-00015-8db0f76e-df96-455a-ac86-0e9e329e...	29/04/2024, 00:00:42	Hot (Inferred)		Block blob	136 B	Available
part-00058-8db0f76e-df96-455a-ac86-0e9e329e...	29/04/2024, 00:00:42	Hot (Inferred)		Block blob	124 B	Available
part-00060-8db0f76e-df96-455a-ac86-0e9e329e...	29/04/2024, 00:00:42	Hot (Inferred)		Block blob	112 B	Available
part-00091-9a3944ae-9daf-4fe4-9eea-fda79456a...	29/04/2024, 00:00:51	Hot (Inferred)		Block blob	55 B	Available
part-00100-8db0f76e-df96-455a-ac86-0e9e329e...	29/04/2024, 00:00:42	Hot (Inferred)		Block blob	111 B	Available
part-00129-8db0f76e-df96-455a-ac86-0e9e329e...	29/04/2024, 00:00:42	Hot (Inferred)		Block blob	153 B	Available
part-00130-9cbc7a7d-84be-406d-bd7e-5d4708e...	29/04/2024, 00:00:48	Hot (Inferred)		Block blob	31 B	Available
part-00136-9a3944ae-9daf-4fe4-9eea-fda79456a...	29/04/2024, 00:00:51	Hot (Inferred)		Block blob	55 B	Available
part-00140-8db0f76e-df96-455a-ac86-0e9e329e...	29/04/2024, 00:00:42	Hot (Inferred)		Block blob	121 B	Available

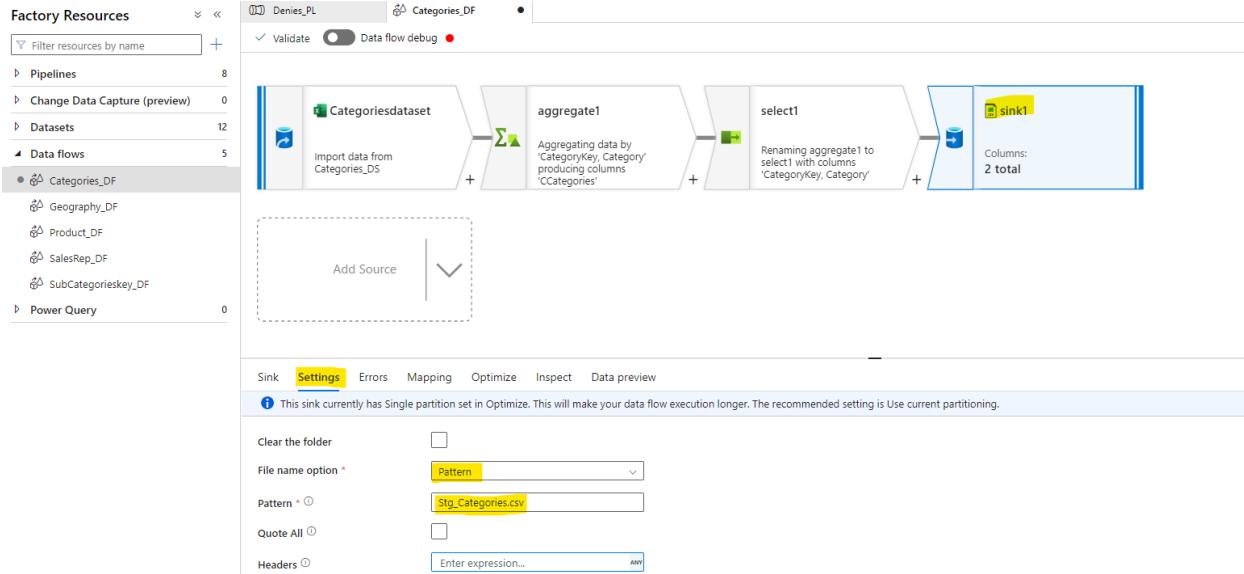
Pipeline runs

Run ID	Triggered	Run start	Run end	Duration	Triggered by	Status	Run	Parameters	Annotations	Last refreshed
0f318c05-d5d7-40bf-b951-f0c48cb4f81	Triggered	4/28/2024, 11:57:57 PM	4/29/2024, 12:00:51 AM	2m 55s	f4bcb69b-43d5-4c43...	Succeeded	Original			0 minutes ago
ae1b1a60-1150-4e39-99a2-c4250a0343	Triggered	4/28/2024, 11:57:54 PM	4/29/2024, 12:01:07 AM	3m 13s	aabeeaca-d30d-469...	Succeeded	Original			0 minutes ago
9fb5457a-9950-4270-a2a3-e8e685109f5	Triggered	4/28/2024, 11:57:37 PM	4/29/2024, 12:01:11 AM	3m 34s	6fa6251b-77bf-489d...	Succeeded	Original			0 minutes ago
eb1ffea-68c9-441c-b004-2ec2c480ad8	Triggered	4/28/2024, 11:57:37 PM	4/28/2024, 11:57:53 PM	16s	ee3552c8-8a59-405...	Succeeded	Original			0 minutes ago
7c27bc53-3be5-4224-a1a4-c3d2669012	Triggered	4/28/2024, 11:57:37 PM	4/29/2024, 12:01:04 AM	3m 28s	8b94f21b-22d7-4ab9...	Succeeded	Original			0 minutes ago
ebd7e656-5ba0-46b3-a2c9-b7791c0e07	Triggered	4/28/2024, 11:57:37 PM	4/28/2024, 11:57:56 PM	20s	aaa31d3-263d-462...	Succeeded	Original			0 minutes ago
24e7ea49-e822-431f-9718-7fd20abb5c	Triggered	4/28/2024, 11:57:37 PM	4/29/2024, 12:00:56 AM	3m 20s	e1d0a643-d7fa-476c...	Succeeded	Original			0 minutes ago
654ab2a0-cd3e-4e7c-9ce4-dfbf6756ca6	Triggered	4/28/2024, 11:57:34 PM	4/29/2024, 12:01:13 AM	3m 39s	Manual trigger	Succeeded	Original			0 minutes ago

In data flow we have to do small changes

Categories dataflow





Geography_Datflow:

Factory Resources

- Pipelines: 8
- Change Data Capture (preview): 0
- Datasets: 12
- Data flows: 5
- Categories_DF
- Geography_DF
- Product_DF
- SalesRep_DF
- SubCategorieskey_DF
- Power Query: 0

Denies_PL Categories_DF

Validate Data flow debug ●

```

graph LR
    A[source1] --> B[select1]
    B --> C[AddingGeographyID]
    C --> D[select2]
    D --> E[sink1]
    
```

Import data from Mongodo_Geography_DS
Renaming source1 to select1 with columns 'Country, Town'
Adding new key Geography_ID starting from 1 with step 1
Renaming AddingGeographyID to select2 with columns Geography_ID, Country, Town
Renaming select2 with columns Geography_ID, Country, Town
sink1
Columns: 3 total

Stg_Categories_DS Geography_DF

Sink Settings Errors Mapping Optimize Inspect Data preview

This sink currently has Single partition set in Optimize. This will make your data flow execution longer. The recommended setting is Use current partitioning.

Partition option: Use current partitioning Single partition Set partitioning

Denies_PL Geography_DF

Validate Data flow debug ●

```

graph LR
    A[source1] --> B[select1]
    B --> C[AddingGeographyID]
    C --> D[select2]
    D --> E[sink1]
    
```

Import data from Mongodo_Geography_DS
Renaming source1 to select1 with columns 'Country, Town'
Adding new key Geography_ID starting from 1 with step 1
Renaming AddingGeographyID to select2 with columns Geography_ID, Country, Town
Renaming select2 with columns Geography_ID, Country, Town
sink1
Columns: 3 total

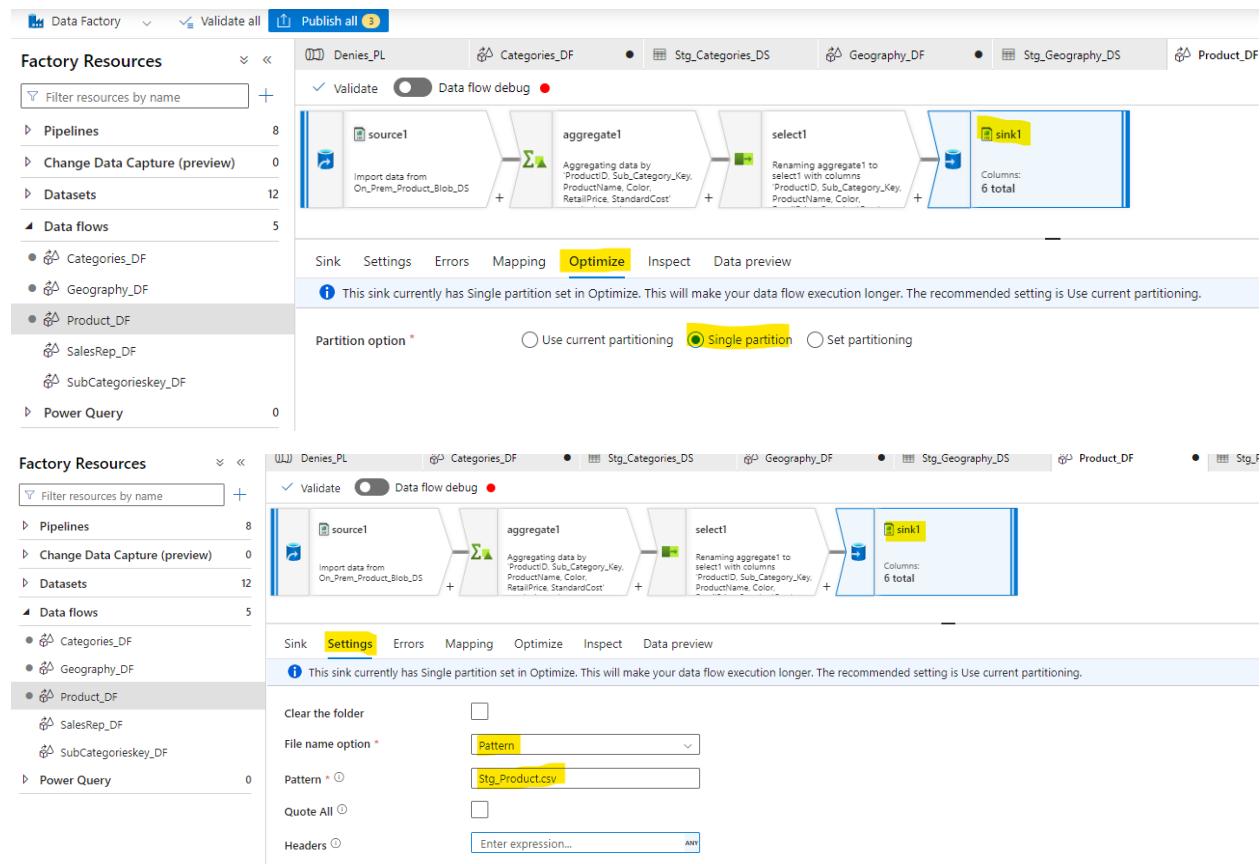
Stg_Geography_DS

Sink Settings Errors Mapping Optimize Inspect Data preview

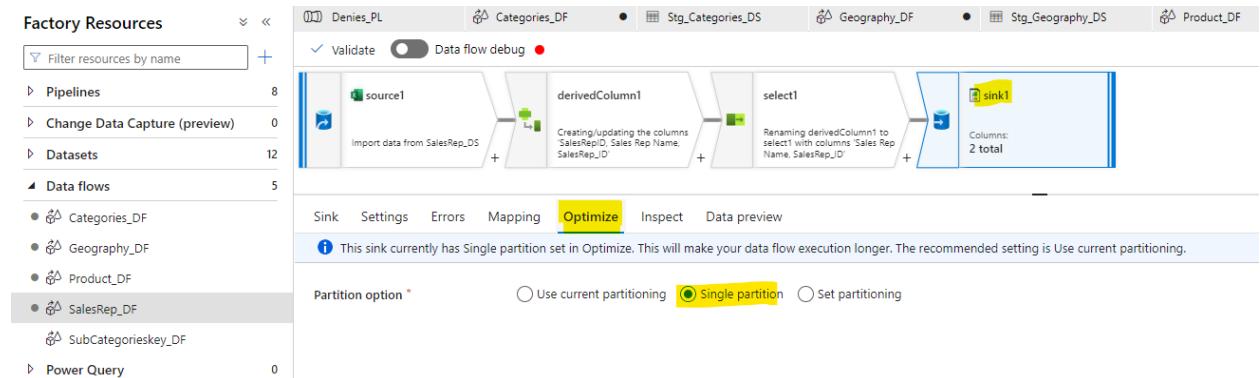
This sink currently has Single partition set in Optimize. This will make your data flow execution longer. The recommended setting is Use current partitioning.

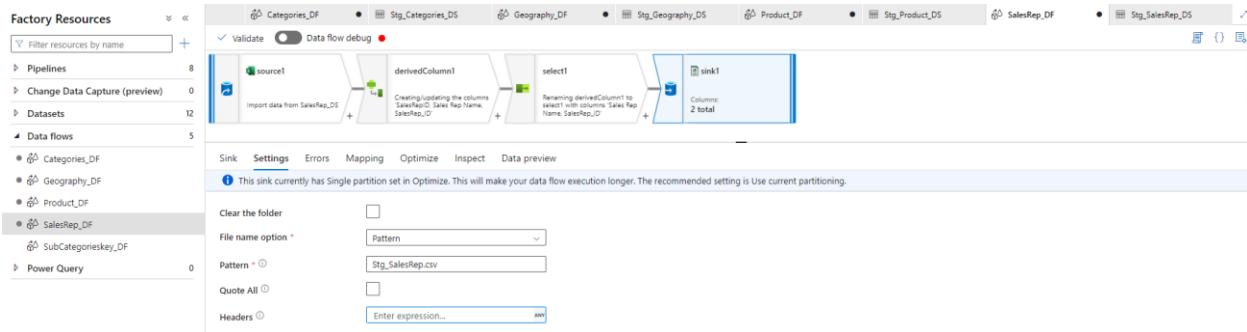
Clear the folder
File name option * Pattern: Stg_Geography.csv
Quote All
Headers Enter expression...

Product Data Flow:

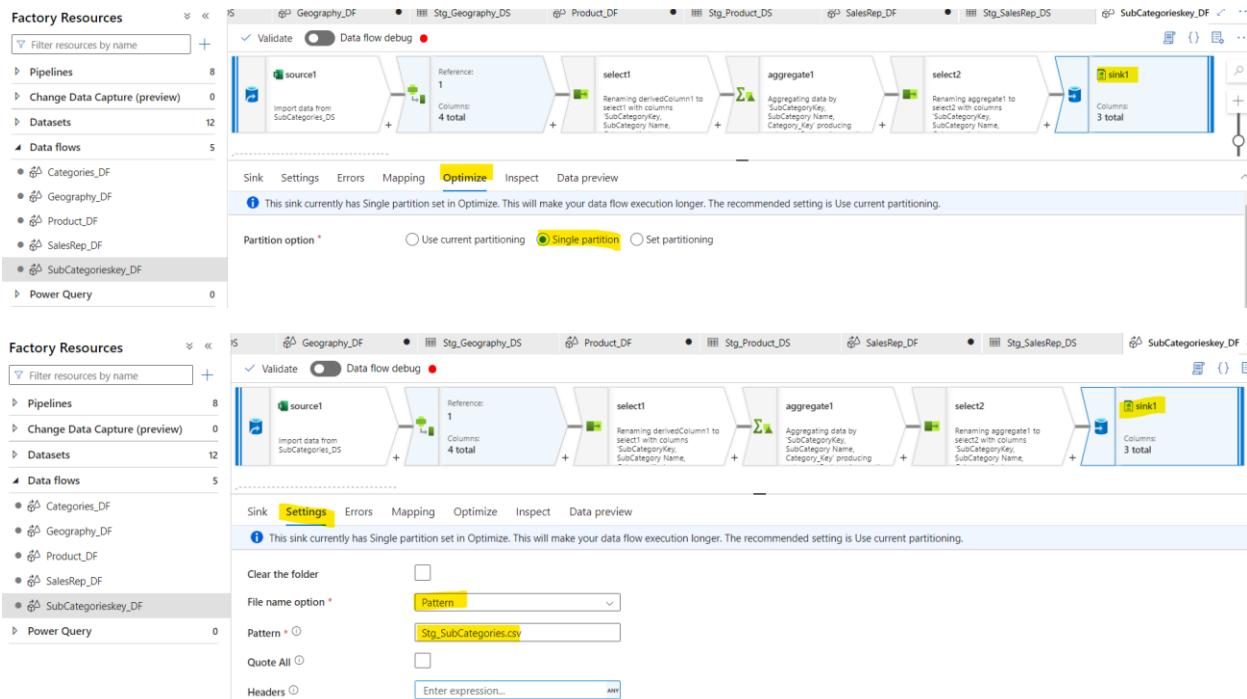


SalesRep Data Flow:





SubCategories Data_Flow:



Delete Previously uploaded files in ksrdatasa → re run the pipeline(Denis_PL)

Getting mails using logic apps:

Dashboard > datavizLogicApp

datavizLogicApp | Logic app designer

Add a trigger

Runtime: Select a runtime Action Type: Triggers

Group by Connector:

- HTTP** In App
- HTTP** In App Trigger
- HTTP Webhook** In App Trigger
- HTTP + Swagger** In App Trigger
- Request** In App
- When a HTTP request is received** In App Trigger

Pipeline expression builder

```
{
  "Status": "Successful",
  "PLName": "@{pipeline().Pipeline}",
  "ExecutionTime": "@{pipeline().TriggerTime}"
}
```

Activity outputs Parameters System variables Functions Variables

OK Cancel

The screenshot shows the Azure Logic App designer interface. On the left, there's a sidebar with navigation links like Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Development Tools (selected), Logic app designer, Logic app code view, Run History, Versions, API connections, Quick start guides, Settings, Workflow settings, Authorization, Access keys, and Identity. The main area has tabs for Data Factory, Validate all, Publish at, Pipelines, Categories_PL, Denies_PL, Geography_PL, MongoDB_Geography_Blob, On_Prem_MSSQL_to_Blob, Product_PL, SalesRep_PL, SubCategoriesKey_PL, Change Data Capture (preview), Datasets (12), Data flows (5), and Power Query (0). A central workspace shows a flow with three Execute Pipeline activities connected sequentially. The top one is triggered by a When a HTTP request is received trigger. The middle one is triggered by Execute Pipeline1 (On_Prem-MSSQL_to_Blob). The bottom one is triggered by Execute Pipeline2 (MongoDB_Geography...). Below the workspace is a Pipeline expression builder with a JSON template. At the bottom is an Enter or paste a sample JSON payload input field containing the same JSON template.

Enter or paste a sample JSON payload.

```

1 {
2   "Status": "Successful",
3   "PLName": "@{pipeline().Pipeline}",
4   "ExecutionTime": "@{pipeline().TriggerTime}"
5 }
6

```

Dashboard > datavizonLogicApp

datavizonLogicApp | Logic app designer

The Logic App Designer interface shows a complex workflow. It starts with a trigger "When a HTTP request is received". This triggers an "Execute Pipeline" action, which then branches into three parallel paths. Each path contains an "Execute Pipeline" action, followed by a "Web" action (either "Web1" or "Web2"). The "Settings" tab for the main trigger is open, showing the URL: <https://prod-51.eastus.logic.azure.com:443/workflows/fc5d5a8df9ec4357beccc47717d37828/triggers/manual/Run>. The Request Body JSON Schema is defined as:

```
{
  "type": "object",
  "properties": {
    "Status": {
      "type": "string"
    },
    "PLName": {
      "type": "string"
    },
    "ExecutionTime": {
      "type": "string"
    }
  }
}
```

Factory Resources

- Denies_PL
- Categories_PL
- Product_PL
- SalesRep_PL
- SubCategoriesKey_PL

Activities

- General
 - Web
 - WebHook

Triggers

- When a HTTP request is received

Actions

- Execute Pipeline
- Execute Pipeline1 On_Prem-MSSQL_to_Blob
- Execute Pipeline2 MongoDB_Geography
- Geography pipeline Geography_PL
- Product pipeline product_PL
- Web
- Web1
- Web2

Dashboard > datavizonLogicApp

datavizonLogicApp | Logic app designer

The Logic App Designer interface shows a simple workflow. It starts with a trigger "When a HTTP request is received", which then triggers an "Send an email (V2)" action. The "Send an email (V2)" action is connected to a "Create Connection" section on the right.

Actions

- Send an email (V2)

Create Connection

Send an email (V2)

Create a new connection

Sign in to create a connection to Outlook.com.

Sign in

Sign in to your Microsoft account - Google Chrome

login.live.com/oauth20_authorize.srf?client_id=44da0bc6-ecea-41e8-a74c-532e754b...



Pick an account



Chennu Akula
chennu18@outlook.com
Signed in

⋮



Use another account

Legal
tio
(V2)
tion

Dashboard > datavizonLogicApp

datavizonLogicApp | Logic app designer

Logic app

Search

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Development Tools

- Logic app designer
- Logic app code view
- Run History
- Versions
- API connections
- Quick start guides

Settings

- Workflow settings
- Authorization
- Access keys

When a HTTP request is received

+ Send an email (V2)

To: chennu18@outlook.com

Subject: Product Pipelins status is: Status

Body:

Normal Arial 15px

Here are the Pipe line Details:
Pipeline: PLName Execution time: ExecutionTime

Thanks and regards
KSR Dev Team

Home View Help

New mail Delete Archive Report Sweep Move to Reply all Read / Unread ...

Favorites

- Inbox 6
- Sent Items
- Drafts
- Add favorite

Folders

- Inbox** 6
- Junk Email 1
- Drafts
- Sent Items
- Deleted Items 7
- Archive
- Notes
- Conversation History
- Create new folder

Product Pipelins status is: Successful

Chennu Akula Product Pipelins status is: 11:33 Here are the Pipe line Details: Pip...

Microsoft Azure Optimize your virtual ma... 01:35 Improve security, monitor perfor...

This month

Microsoft Azure Improve data perfor... Sat 20:57 Learn about database migration i...

Microsoft Azure Take action to keep yo... Fri 20:55 Learn what to do to keep your Az...

Microsoft Azure Take two important ... Wed 24/04 Ensure your databases run at pea...

Chennu Akula The Pipeline: Bolb_to... Tue 23/04 The Pipeline encountered an issu...

Chennu Akula

CA Chennu Akula To: You Mon 29/04/2024 11:33

Here are the Pipe line Details:
Pipeline: Denies_PL
Execution time: 2024-04-29T05:58:45.7721739Z

Thanks and regards
KSR Dev Team

Reply Forward

www.amazon.com/S. Buy sportswear sets men - Amazon.com Official Site Enjoy low prices a get fast, free delivery with Prim on millions of products. Try eligil items in sportswear sets men using Prime Try Before Y Buy. Shop now!

Learn more >

Pipeline expression builder

Add dynamic content below using any combination of expressions, functions and system variables.

```
{
  "Status": "Failed",
  "PLName": "@{pipeline().Pipeline}",
  "ExecutionTime": "@{pipeline().TriggerTime}",
  "Error": "@{activity('Product pipeline').output.errors[0].Message}"
}
```

Activity outputs Parameters System variables Functions Variables

Search

Execute Pipeline1
Execute Pipeline1 activity output

Execute Pipeline1
Execute Pipeline1 pipeline return value

Product pipeline
Product pipeline activity output

Product pipeline
Product pipeline pipeline return value

OK **Cancel**

The screenshot shows the Pipeline Expression Builder dialog. It contains a JSON template for failed pipeline status, which includes pipeline name, execution time, and error message. Below the JSON, there are sections for Activity outputs, Parameters, System variables, Functions, and Variables. A search bar is also present. The background shows the Data Factory pipeline editor with two Execute Pipeline activities connected to a Web activity.

Fact table(Sales):

The screenshot shows the Azure Data Factory pipeline editor. On the left, the 'Factory Resources' sidebar lists various pipelines, datasets, and data flows. In the center, the 'Activities' pane shows a sequence of steps: 'Execute Pipeline' followed by 'Execute' (with a sub-step 'Execute Pipeline1'). A tooltip for 'Execute Pipeline1' indicates it's an 'On_Prem-MSSQL_to_Blob' activity. Below these steps is a 'Parameters' tab with a 'New' button.

To the right, a 'New dataset' dialog is open. It displays a grid of data store icons. The 'Azure Data Lake Storage Gen2' icon is highlighted with a yellow box. The dialog also includes tabs for 'All', 'Azure', 'Database', 'File', 'Generic protocol', 'NoSQL', and 'Services and apps'. At the bottom right of the dialog are 'Save' and 'Cancel' buttons.

This screenshot continues from the previous one. The 'New dataset' dialog has been closed, and the pipeline editor now shows the 'Output' tab selected. The 'Select format' section contains a grid of file format icons: Avro, Binary, DelimitedText (highlighted with a yellow box), Excel, JSON, ORC, Parquet, and XML. Below the grid are 'Continue' and 'Back' buttons.

The screenshot shows the 'Properties' page for the 'Sales_DS' dataset. The 'General' tab is selected. The dataset name is 'Sales_DS' and its description is empty. The 'Annotations' section has a '+ New' button. On the left, the 'Factory Resources' sidebar shows the 'Datasets' section with 'Sales_DS' selected. The 'Connection' tab of the dataset properties page shows a connection named 'ADLS_ADF_LS' and various parameters like 'File name', 'Separator', and 'Quote character'. The 'Schema' and 'Parameters' tabs are also visible.

Factory Resources

- Categories_DS
- Geography_DS
- Mongodb_Geography_DS
- On_Prem_Product_Blob_DS
- Product_DS
- Sales_DS
- SalesRep_DS
- Stg_Categories_DS
- Stg_Geography_DS
- Stg_Product_DS
- Stg_SalesRep_DS
- Stg_SubCategoriesKey_DS
- SubCategories_DS

Data flows (6)

- Sales_DF**
- Categories_DF
- Geography_DF
- Product_DF
- SalesRep_DF
- SubCategorieskey_DF

source1

Validate Data flow debug

Source settings

Output stream name: source1

Description: Import data from Sales_DS

Source type: Dataset

Dataset: Sales_DS

Options

- Allow schema drift
- Infer drifted column types
- Validate schema

Skip line count:

Sampling: Enable Disable

Properties

Name: Sales_DF

Description:

Factory Resources

- Categories_DS
- Geography_DS
- Mongodb_Geography_DS
- On_Prem_Product_Blob_DS
- Product_DS
- Sales_DS
- SalesRep_DS
- Stg_Categories_DS
- Stg_Geography_DS
- Stg_Product_DS
- Stg_SalesRep_DS
- Stg_SubCategoriesKey_DS
- SubCategories_DS

Data flows (6)

- Sales_DF**
- Categories_DF
- Geography_DF
- Product_DF
- SalesRep_DF
- SubCategorieskey_DF

source1

Validate Data flow debug

Source settings

Projection

Partition option:

- Use current partitioning
- Single partition
- Set partitioning

Properties

Name: Sales_DF

Description:

The screenshot shows two views of the Azure Data Factory Data Flow interface, illustrating the configuration of a data pipeline.

Top View: This view shows the Data preview pane. A data source named "source1" is connected to a table named "Sales_DS". The table has 8 columns: fSalesPr..., ProductID, SalesRepID, Location, Date, Units, and PercentO... . The data preview shows 100 rows of sales data from various locations and dates.

fSalesPr...	ProductID	SalesRepID	Location	Date	Units	PercentO...
36943	10	7	Germany...	1.10.2017	60	0.99
36944	8	2	Denmark...	26.4.2017	71	0.972
36945	9	6	Germany...	28.9.2017	68	0.988
36946	2	5	Germany...	16.4.2017	157	0.97
36947	8	2	Denmark...	13.2.2017	103	0.965
36948	8	1	Czech rep...	3.12.2017	48	0.997
36949	4	2	Denmark...	26.8.2017	30	0.984
36950	4	7	Germany...	16.4.2017	57	0.97
36951	11	4	Germany...	3.11.2017	145	0.994
36952	5	5	Germany...	1.1.2017	31	0.954
36953	9	6	Germany...	4.3.2017	87	0.964

Bottom View: This view shows the Source settings pane for the "source1" component. The "Schema modifier" dropdown is set to "Derived Column". Other settings include "Source type" (Dataset), "Sampling" (0%), and "Skip line count" (0).

Dataflow expression builder

Adding columns

Derived Columns

Create new

Country

Column name *

Country

Expression

```
split(Location, ";"){1}
```

Expression elements

- All
- Functions
- Input schema
- Parameters
- Cached lookup
- Data flow library functions
- Locals

Expression values

sp

- Create new
- regexSplit(string to split, regex expression)
- trim(string to trim, trim characters)
- split(string to split, split characters)
- toDecimal(value, precision, scale, format, locale)
- toString(value, number format/date format, timezone option, date locale)
- trim(string to trim, trim characters)

Data preview Refresh

Save and finish Cancel Clear contents

Dataflow expression builder

Adding columns

Derived Columns

Create new

Country

City

DimDate

Column name *

DimDate

Expression

```
toTimestamp(Date, 'YYYY-MM-DD')
```

Expression elements

- All
- Functions
- Input schema
- Parameters
- Cached lookup
- Data flow library functions
- Locals

Expression values

Filter by keyword

- Create new
- fSalesPrimarykey
- ProductID
- SalesRepID
- Location
- Date
- Unit

Save

Data preview Refresh

Dataflow expression builder

Adding columns

Derived Columns

Create new

Country

City

DimDate

Column name *

DimDate

Expression

```
toTimestamp(Date, 'dd.mm.yyyy')
```

Expression elements

- All
- Functions
- Input schema
- Parameters
- Cached lookup
- Data flow library functions
- Locals

Expression values

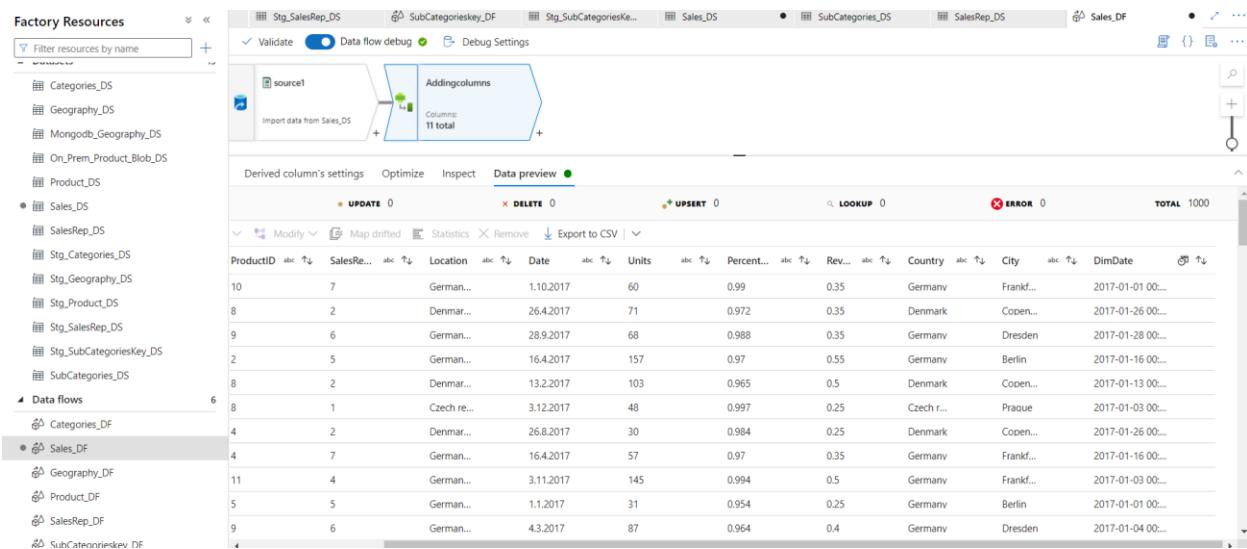
Filter by keyword

- Create new
- fSalesPrimarykey
- ProductID
- SalesRepID
- Location
- Date
- Unit

Save

Data preview Refresh

Save and finish Cancel Clear contents



Dataflow expression builder

Addingcolumns

Derived Columns

Create new ▾

abc Country
abc City
DimDate

Column name * DimDate

Expression

```
toDate(toTimestamp(Date,'dd.mm.yyyy'))
```

Save

Expression elements

All
Functions
Input schema
Parameters
Cached lookup
Data flow library functions
Locals

Expression values

Filter by keyword

Create new ▾
fSalesPrimaryKey
ProductID
SalesRepID
Location
Date

Data preview Refresh

Save and finish Cancel Clear contents

Factory Resources

Data flows

- Categories_DF
- Sales_DS
- SalesRep_DS
- Stg_Categories_DS
- Stg_Geography_DS
- Stg_Product_DS
- Stg_SalesRep_DS
- SubCategoriesKey_DS
- SubCategories_DS
- Sales_DF**
- Categories_DF
- Geography_DF
- Product_DF
- SalesRep_DF
- SubCategoriesKey_DS

Factory Resources

Data flows

- Categories_DF
- Sales_DS
- SalesRep_DS
- Stg_Categories_DS
- Stg_Geography_DS
- Stg_Product_DS
- Stg_SalesRep_DS
- SubCategoriesKey_DS
- SubCategories_DS
- Sales_DF**
- Categories_DF
- Geography_DF
- Product_DF
- SalesRep_DF
- SubCategoriesKey_DS

Factory Resources

Data flows

- Categories_DF
- Sales_DS
- SalesRep_DS
- Stg_Categories_DS
- Stg_Geography_DS
- Stg_Product_DS
- Stg_SalesRep_DS
- SubCategoriesKey_DS
- SubCategories_DS
- Sales_DF**
- Categories_DF
- Geography_DF
- Product_DF
- SalesRep_DF
- SubCategoriesKey_DS

Stg_SalesRep_DS

SubCategorieskey_DF

Stg_SubCategoriesKe...

Sales_DS

SubCategories_DS

SalesRep_DS

Sales_DF

Source1

Addingcolumns

Select1

Columns:

Select settings

Optimize

Inspect

Data preview

Description

Renaming Addingcolumns to select1 with columns : fSalesPrimaryKey, ProductID, SalesRepID, Country, City.

Incoming stream *

Options

Input columns *

Auto mapping

Reset

Add mapping

Delete

9 mappings: 2 column(s) from the inputs left unmapped

Name as	Addingcolumns's column
fSalesPrimaryKey	abc fSalesPrimaryKey
ProductID	abc ProductID
SalesRepID	abc SalesRepID
Country	abc Country
City	abc City
Date	abc DimDate
Units	abc Units
PercentOfStandardCost	abc PercentOfStandardCost
RevenueDiscount	abc RevenueDiscount

Stg_SalesRep_DS

SubCategorieskey_DF

Stg_SubCategoriesKe...

Sales_DS

SubCategories_DS

SalesRep_DS

Sales_DF

Source1

Addingcolumns

Columns: 11 total

Derived column's settings

Optimize

Inspect

Data preview

UPDATE 0

DELETE 0

UPSERT 0

LOOKUP 0

ERROR 0

TOTAL 1000

ProductID

SalesRepID

Location

Date

Units

Percent...

Revenue...

Country

City

DimDate

City	DimDate
Berlin	2017-01-16
Berlin	2017-01-01
Berlin	2017-01-07
Berlin	2017-01-29
Berlin	2017-01-13
Berlin	2017-01-21
Berlin	2017-01-09
Berlin	2017-01-22
Berlin	2017-01-12
Berlin	2017-01-08
Berlin	2017-01-30

Preview experience

Off

Source1

Reference: 1

select1

Columns:

Select settings

Optimize

Inspect

Data preview

Options

Input columns *

Auto mapping

Reset

Add mapping

Delete

11 mappings: All inputs mapped

Name as	Addingcolumns's column
fSalesPrimaryKey	abc fSalesPrimaryKey
ProductID	abc ProductID
SalesRepID	abc SalesRepID
Location	abc Location
Date	abc Date
Units	abc Units
PercentOfStandardCost	abc PercentOfStandardCost
RevenueDiscount	abc RevenueDiscount
Country	abc Country
City	abc City
DimDate	abc DimDate

Factory Resources

Data flows

Stg_SalesRep_DS

Sink

Output stream name: sink1

Description: Add sink dataset

Incoming stream: select1

Sink type: Dataset

Dataset: Select...

Options: Allow schema drift (checked), Validate schema (unchecked)

Select a data store

All, **Azure**, **Database**, **File**, **Generic protocol**, **NoSQL**, **Services and apps**

Continue | **Cancel**

Factory Resources

Data flows

Stg_SalesRep_DS

Sink

Output stream name: sink1

Description: Add sink dataset

Incoming stream: select1

Sink type: Dataset

Dataset: Select...

Options: Allow schema drift (checked), Validate schema (unchecked)

Choose the format type of your data

Continue | **Back** | **Cancel**

Data Factory | **Validate all** | **Publish all**

Factory Resources

Data flows

Stg_SalesRep_DS

Sink

Output stream name: sink1

Description: Add sink dataset

Incoming stream: select1

Sink type: Dataset

Dataset: Select...

Options: Allow schema drift (checked), Validate schema (unchecked)

Set properties

Name: Stg_Sales_DS

Linked service: ADLS_ADF_LS

File path: stagesadata / Sales / File name

First row as header: checked

Import schema: From connection/store (radio button selected), From sample file, None

Advanced

OK | **Back** | **Cancel**

Data Factory < Validate all Publish all

Factory Resources

- Categories_DS
- Geography_DS
- Mongodb_Geography_DS
- On_Prem_Product_Blob_DS
- Product_DS
- Sales_DS**
- SalesRep_DS
- Stg_Categories_DS
- Stg_Geography_DS
- Stg_Product_DS
- Stg_Sales_DS**
- Stg_SalesRep_DS
- Stg_SubCategoriesKey_DS
- SubCategories_DS

Data flows 6

- Categories_DF
- Sales_DF**
- Geography_DF
- Product_DF

Preview experience Data flow debug Debug Settings

Data Factory < Validate all Publish all

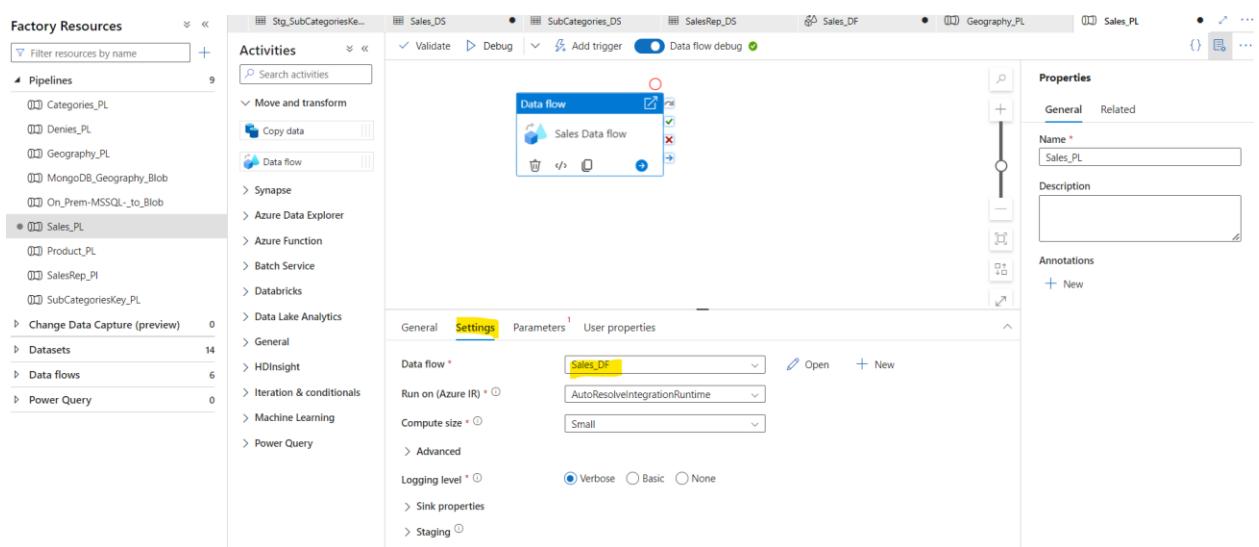
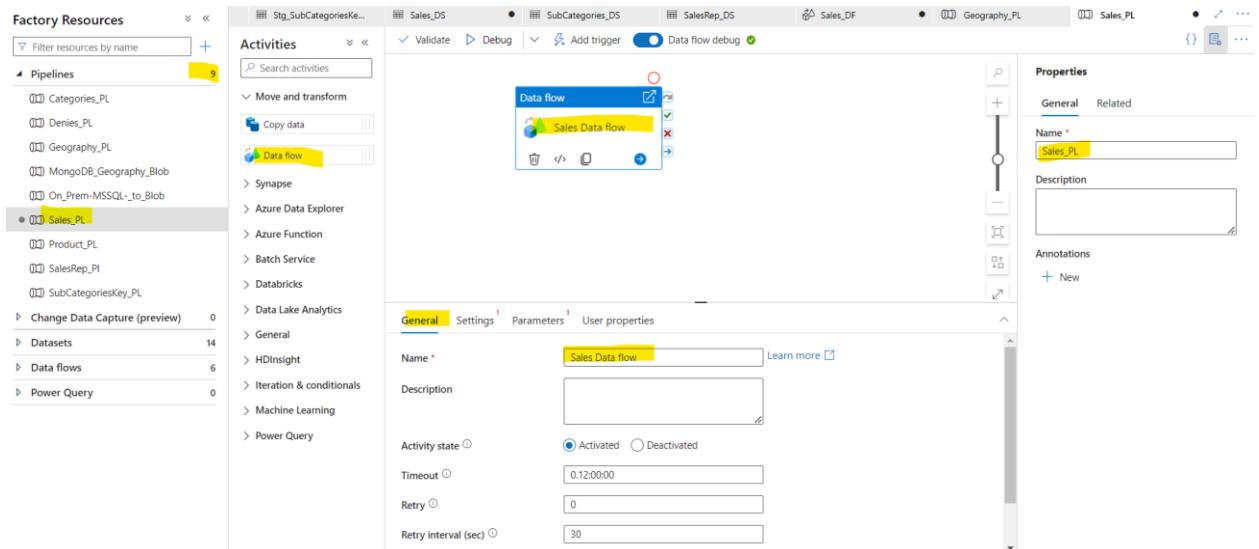
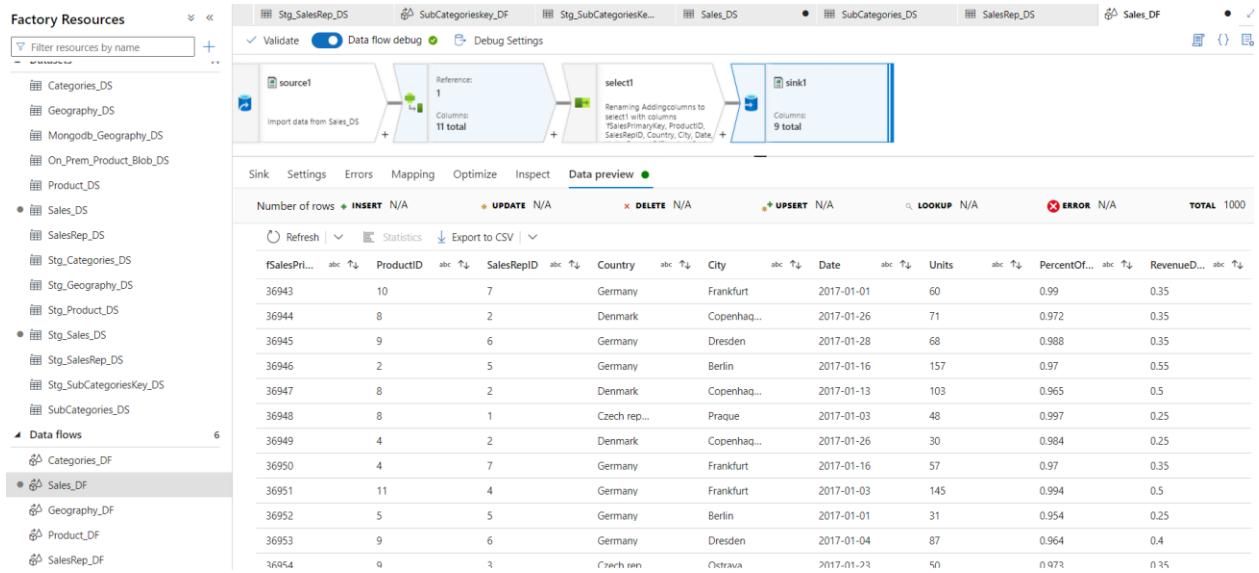
Factory Resources

- Categories_DS
- Geography_DS
- Mongodb_Geography_DS
- On_Prem_Product_Blob_DS
- Product_DS
- Sales_DS**
- SalesRep_DS
- Stg_Categories_DS
- Stg_Geography_DS
- Stg_Product_DS
- Stg_Sales_DS**
- Stg_SalesRep_DS
- Stg_SubCategoriesKey_DS
- SubCategories_DS

Data flows 6

- Categories_DF
- Sales_DF**
- Geography_DF
- Product_DF
- SalesRep_DF

Preview experience Off Data flow debug Debug Settings



Factory Resources

- Pipelines
 - Categories_PL
 - Denies_PL
 - Geography_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - Sales_PL
 - Product_PL
 - SalesRep_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 14
- Data flows 6
- Power Query 0

Activities

- General
 - Execute Pipeline
 - Execute SSIS package
- Machine Learning
 - Machine Learning Batch...
 - Machine Learning Execute...

Execute Pipeline

- SalesRep pipeline SalesRep_PL

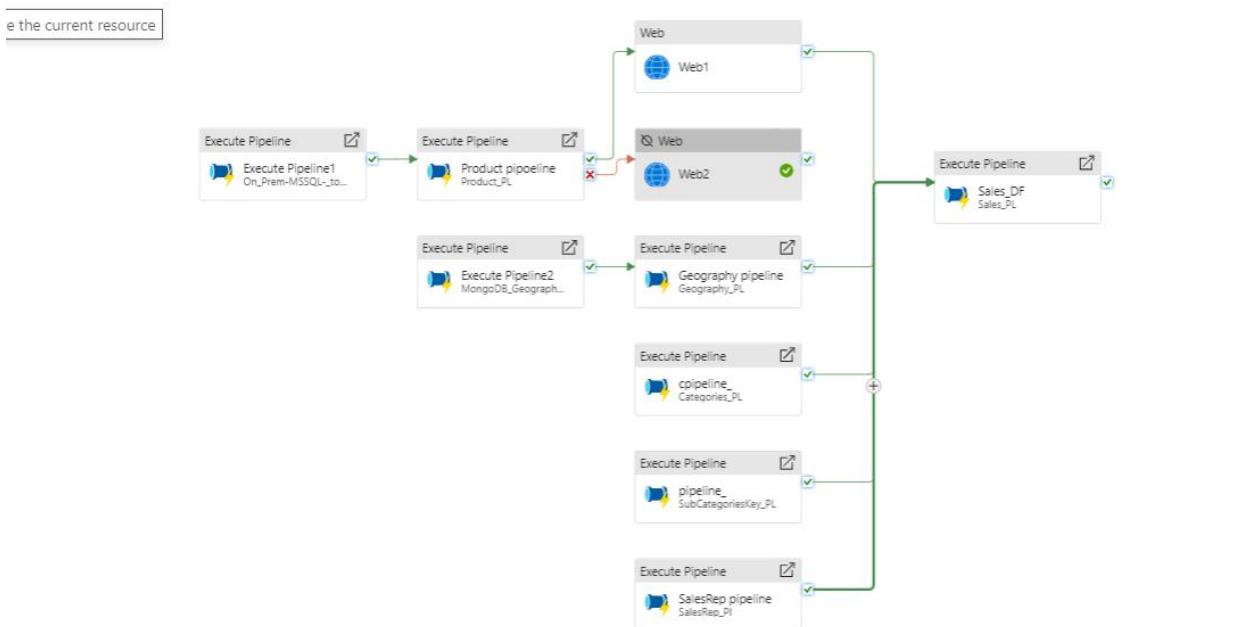
Execute Pipeline

- Sales_DF Sales_PL

General Settings User properties

Invoked pipeline * Sales_PL

Wait on completion



All pipeline runs > Denies_PL - Activity runs

Run Succeeded
Successfully ran Denies_PL (Pipeline). View pipeline run

Activity name	Activity status	Activity type	Run start	Duration	Integration runtime	User properties	Activity run ID	Log
Web1	Succeeded	Web	4/29/2024, 12:54:30 PM	2s	AutoResolveIntegratio		534412d3-22ce-46ce-a7b4-570d60e549a4	
Geography pipeline	Succeeded	Execute Pipeline	4/29/2024, 12:54:12 PM	2m 55s			57745fb2-6e1b-46df-b407-47c3cae5b6ce	
Product pipeline	Succeeded	Execute Pipeline	4/29/2024, 12:54:02 PM	3m 28s			0f626906-778b-45e7-a169-0a02a23a85e	
cpipeline_	Succeeded	Execute Pipeline	4/29/2024, 12:54:42 PM	3m 48s			0050f940-9371-4074-b20d-fed914b5b4ea	
Execute Pipeline2	Succeeded	Execute Pipeline	4/29/2024, 12:54:42 PM	30s			aaa1d136-2457-4ebc-a369-c884ae060a8	
Execute Pipeline1	Succeeded	Execute Pipeline	4/29/2024, 12:54:42 PM	18s			a8656418-e73f-491d-a0b3-e286ea2babf2	
SalesRep pipeline	Succeeded	Execute Pipeline	4/29/2024, 12:54:42 PM	3m 35s			f24de5c4-3e50-4ee8-b3be-d1fc1d157e3ae	
pipeline_	Succeeded	Execute Pipeline	4/29/2024, 12:54:42 PM	3m 34s			8d9e29e7-a242-42d4-a20a-8e371bd53636	

Home > **stagingdata** Container

Authentication method: Access key ([Switch to Microsoft Entra user account](#))
Location: stagingdata

Search blobs by prefix (case-sensitive) Show deleted blobs

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

- Shared access tokens
- Access policy
- Properties
- Metadata

Add filter

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
Stg_Categories.csv	29/04/2024, 12:54:24	Hot (Inferred)		Block blob	41 B	Available
Stg_Geography.csv	29/04/2024, 12:54:02	Hot (Inferred)		Block blob	173 B	Available
Stg_Product.csv	29/04/2024, 12:54:24	Hot (Inferred)		Block blob	600 B	Available
Stg_SalesRep.csv	29/04/2024, 12:54:14	Hot (Inferred)		Block blob	118 B	Available
Stg_SubCategories.csv	29/04/2024, 12:54:12	Hot (Inferred)		Block blob	87 B	Available

Home > **ksrdatadisa** | Containers > **stagesalesdata** Container

Authentication method: Access key ([Switch to Microsoft Entra user account](#))
Location: stagesalesdata / Sales

Search blobs by prefix (case-sensitive) Show deleted objects

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

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

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
Stg_Sales.csv	29/04/2024, 12:54:47	Cool (Inferred)		Block blob	2.73 MiB	Available

Denies Data Flow:

Factory Resources

- Pipelines
- Datasets
- Data flows
- Change Data Capture (preview)
- Denies
- Geography
- MongoDB, Geography, Blob
- On_Prem-MSSQL-to_Blob
- DeniesFinal
- Product
- Sales
- SalesRep
- SubCategoriesKey

Properties

General

Name * Denies_DF

Description

Source settings

Output stream name * StgCategoriesData

Description Import data from Stg_Categories_DS

Source type *

- Dataset
- Inline

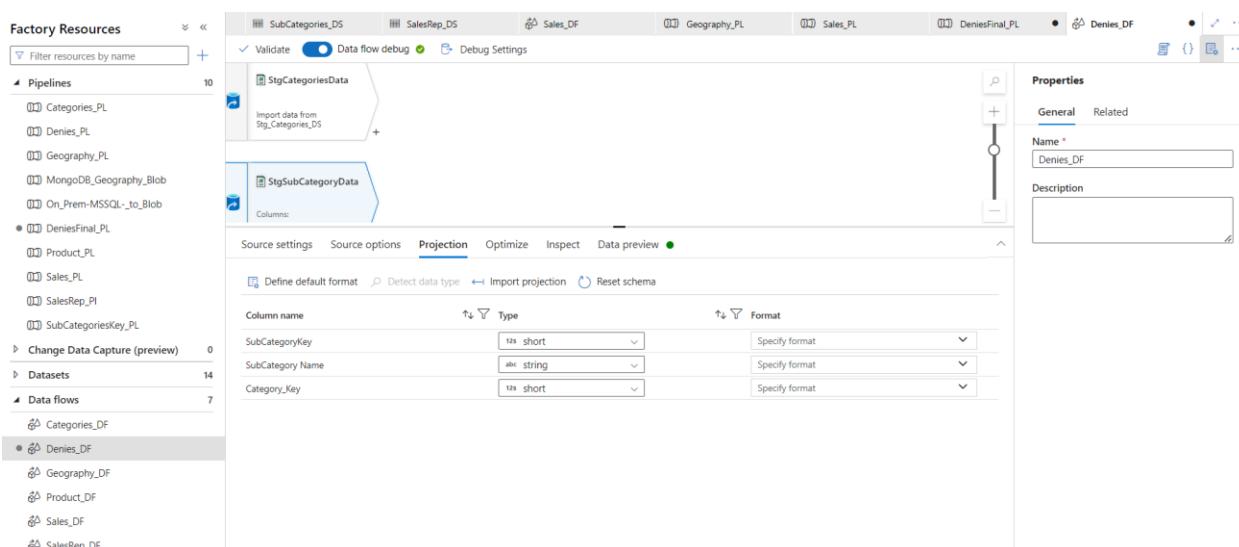
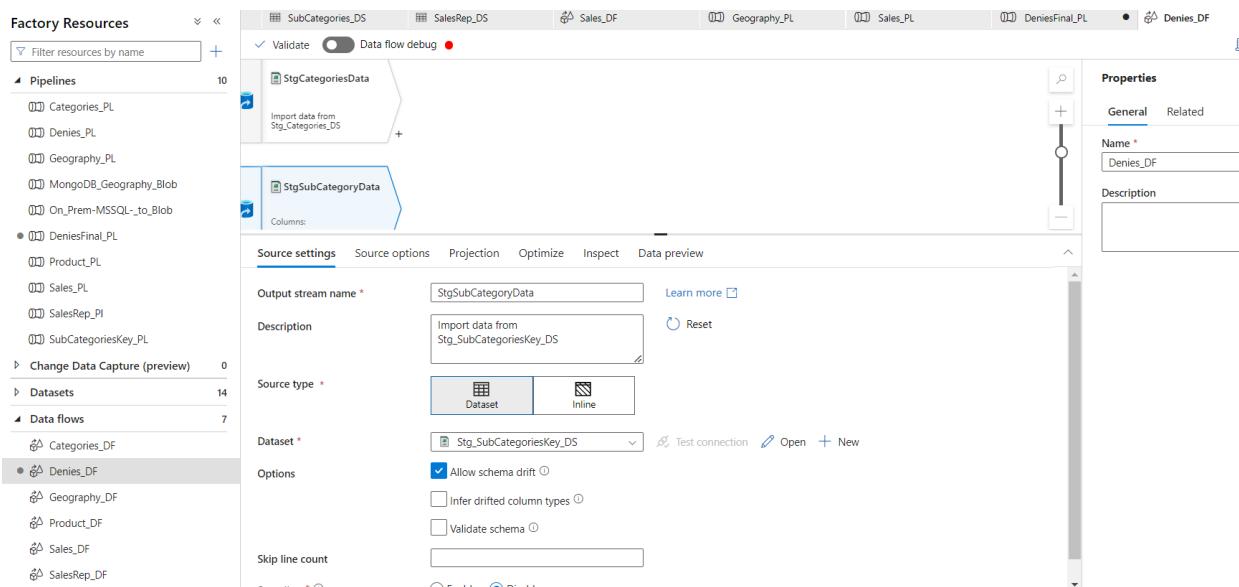
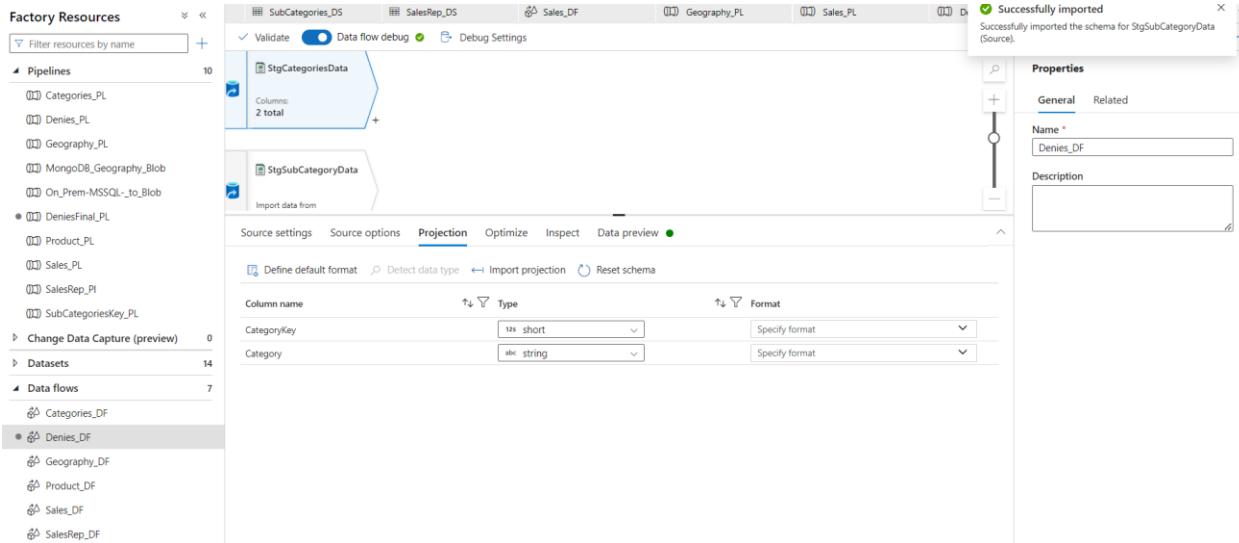
Dataset * Stg_Categories_DS

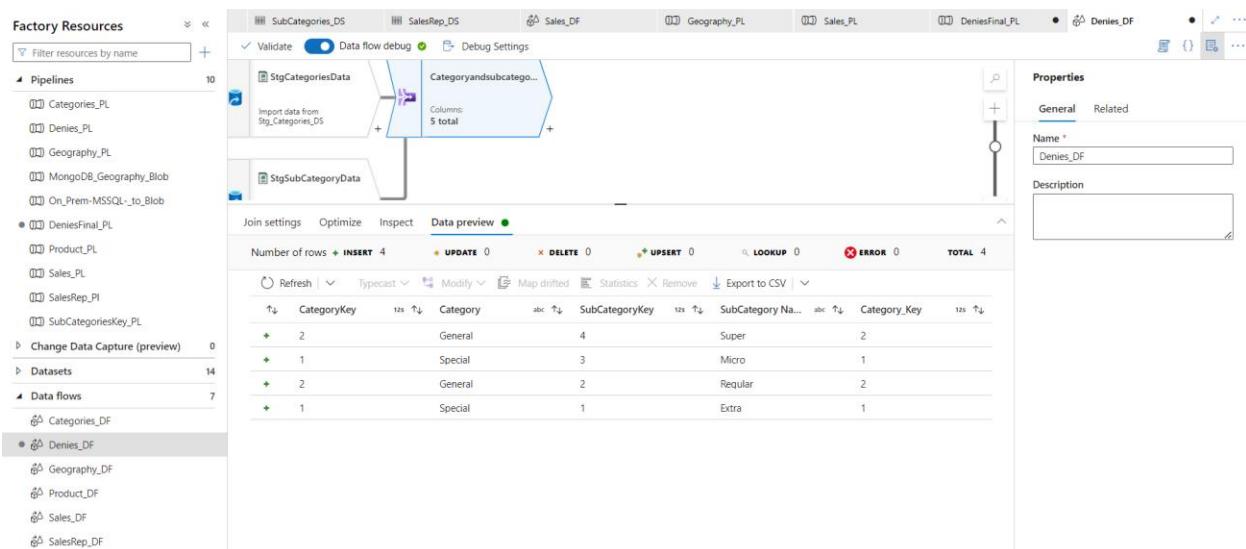
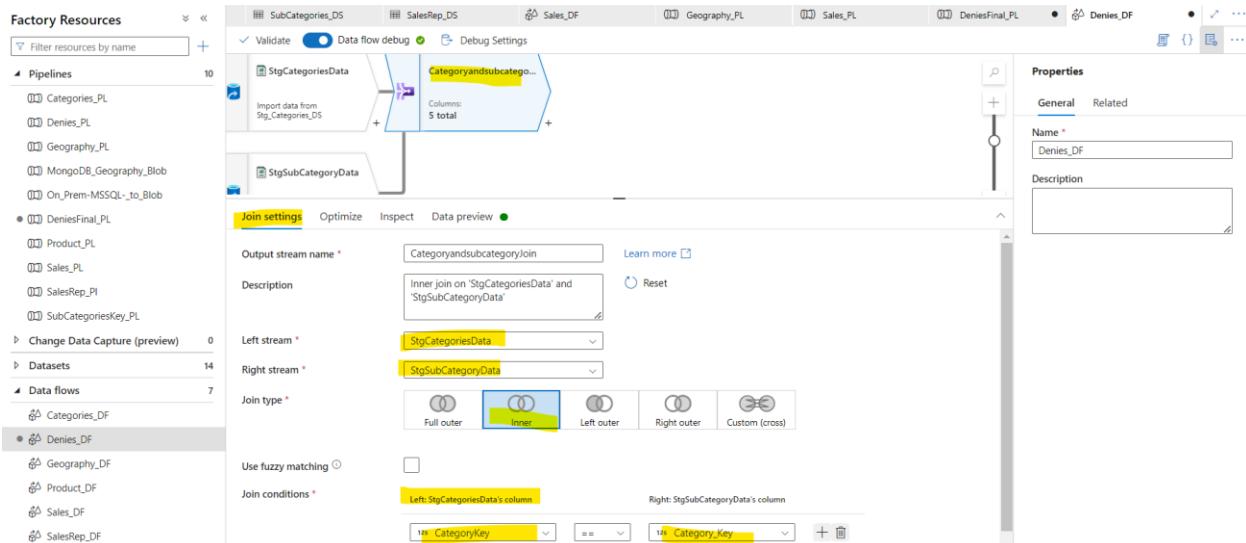
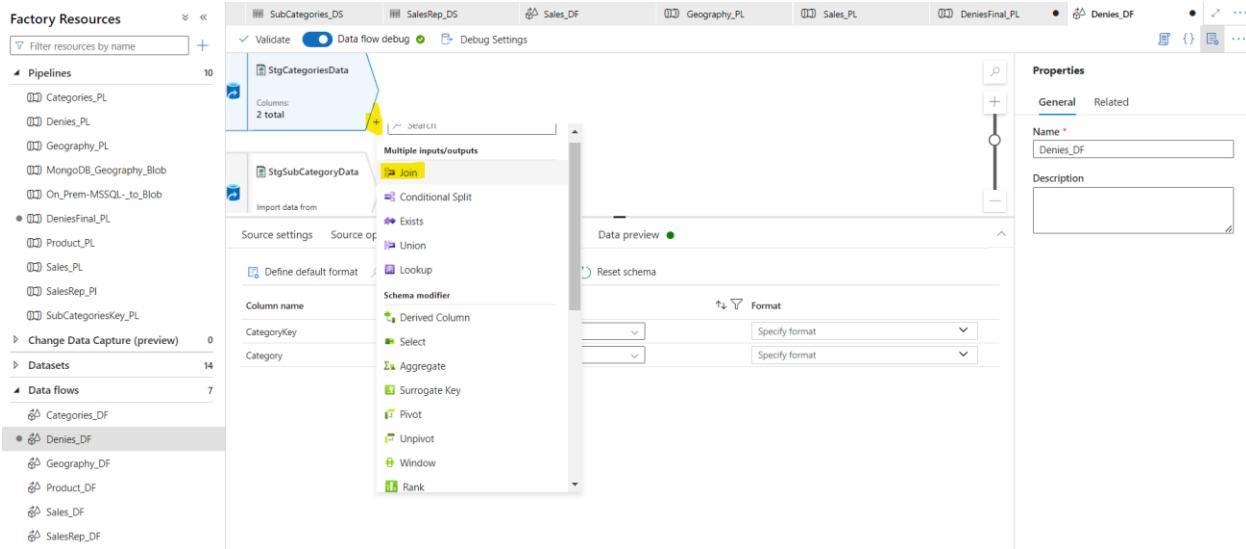
Options

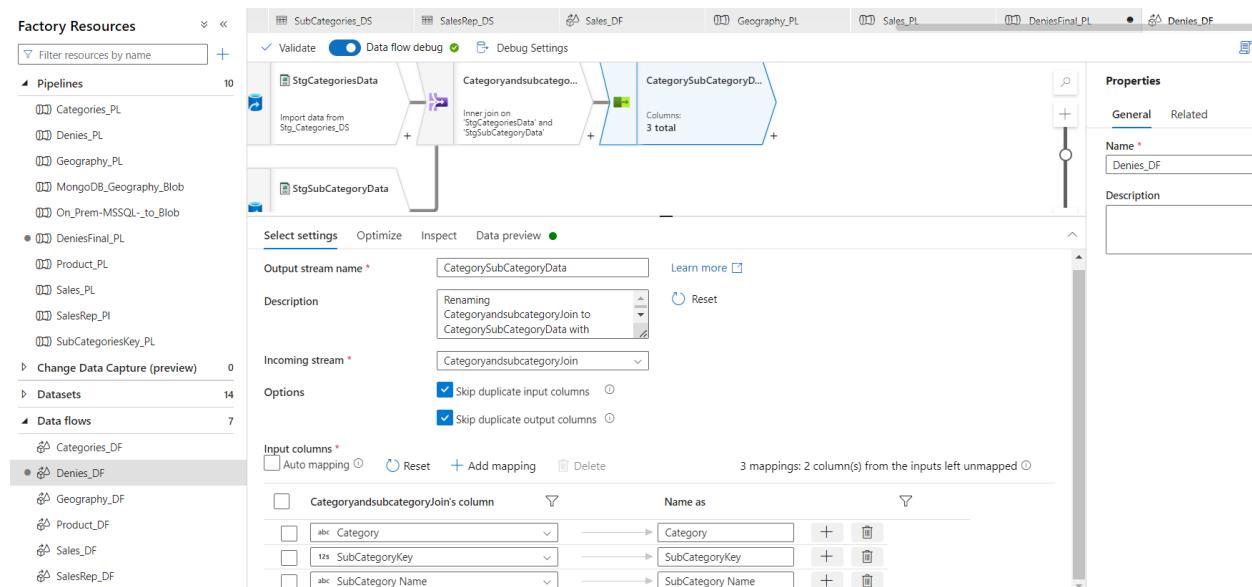
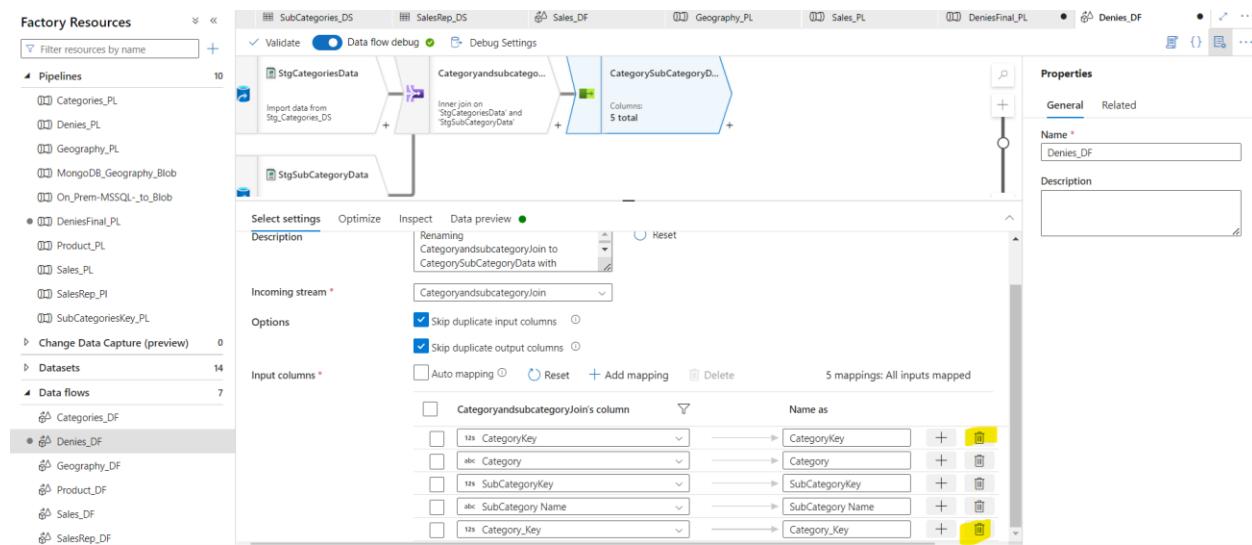
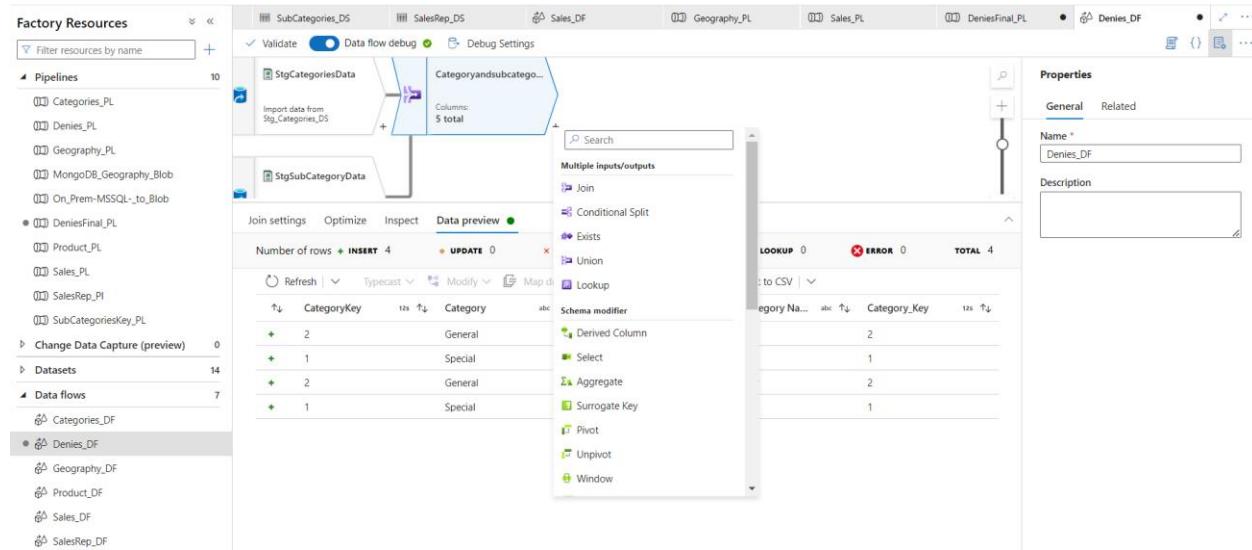
- Allow schema drift
- Infer drifted column types
- Validate schema

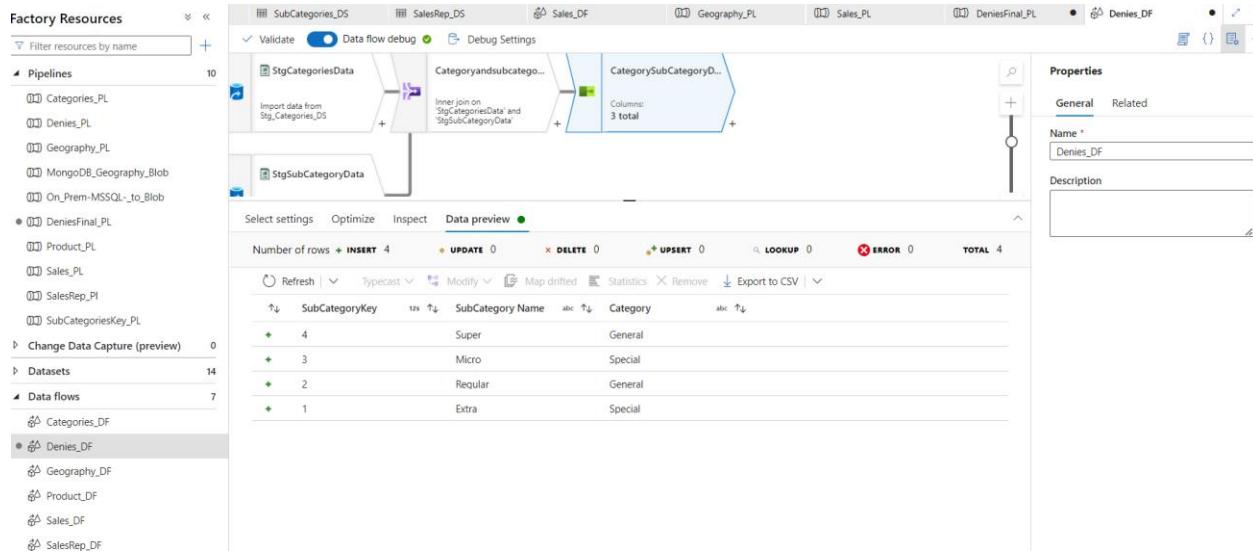
Skip line count

Sampling * Enable Disable

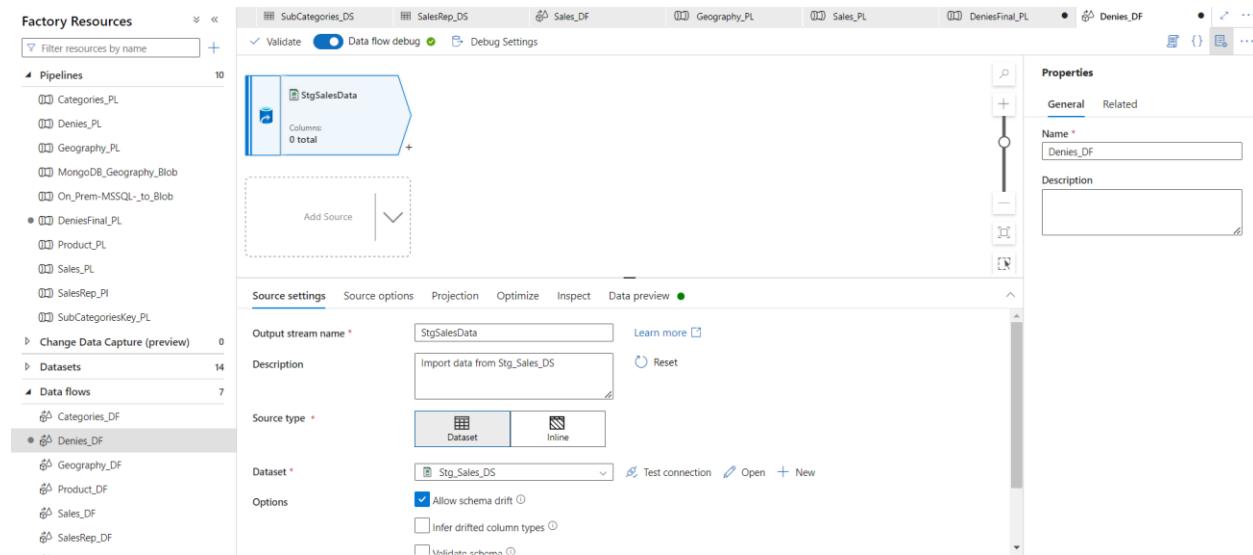








Another join condition:



Factory Resources

- Pipelines (10)
 - Categories_PL
 - Denies_PL
 - Geography_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - DeniesFinal_PL
 - Product_PL
 - Sales_PL
 - SalesRep_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) (0)
- Datasets (14)
 - Categories_DS
 - Denies_DF
 - Geography_DS
 - Product_DS
 - Sales_DS
 - SalesRep_DS
- Data flows (7)
 - Categories_DF
 - Denies_DF
 - Geography_DF
 - Product_DF
 - Sales_DF
 - SalesRep_DF

Properties

General

Name * Denies_DF

Description

Projection

Column name	Type	Format
fSalesPrimaryKey	integer	Specify format
ProductID	short	Specify format
SalesRepID	short	Specify format
Country	string	Specify format
City	string	Specify format
Date	date	Specify format
Units	short	Specify format

Factory Resources

- Pipelines (10)
 - Categories_PL
 - Denies_PL
 - Geography_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - DeniesFinal_PL
 - Product_PL
 - Sales_PL
 - SalesRep_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) (0)
- Datasets (14)
- Data flows (7)
 - Categories_DF
 - Denies_DF
 - Geography_DF
 - Product_DF
 - Sales_DF
 - SalesRep_DF

Properties

General

Name * Denies_DF

Description

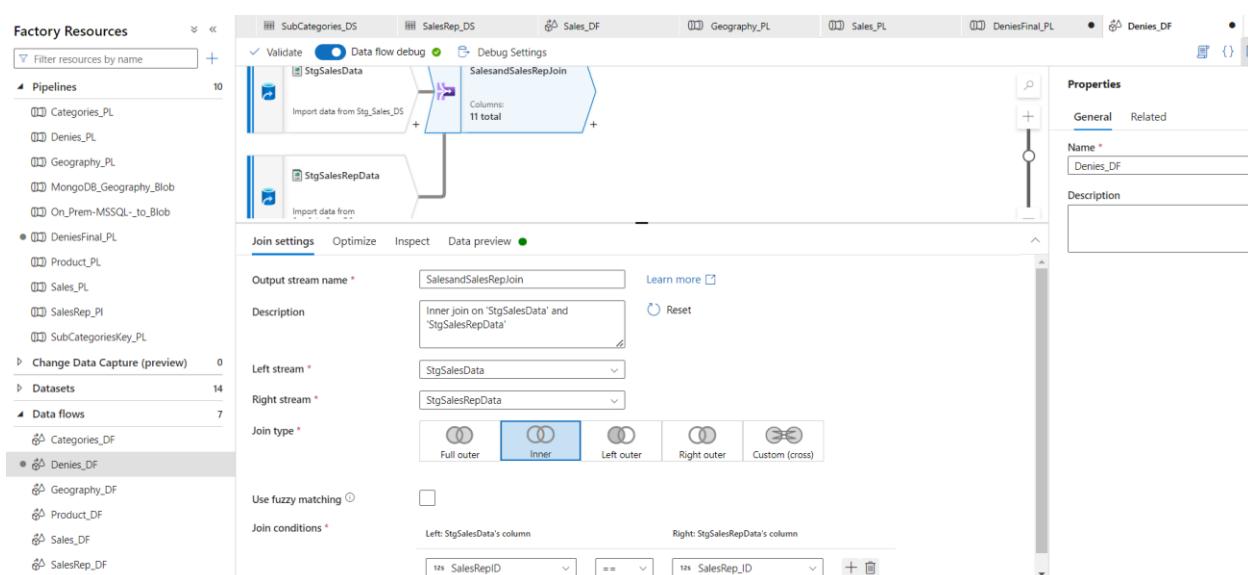
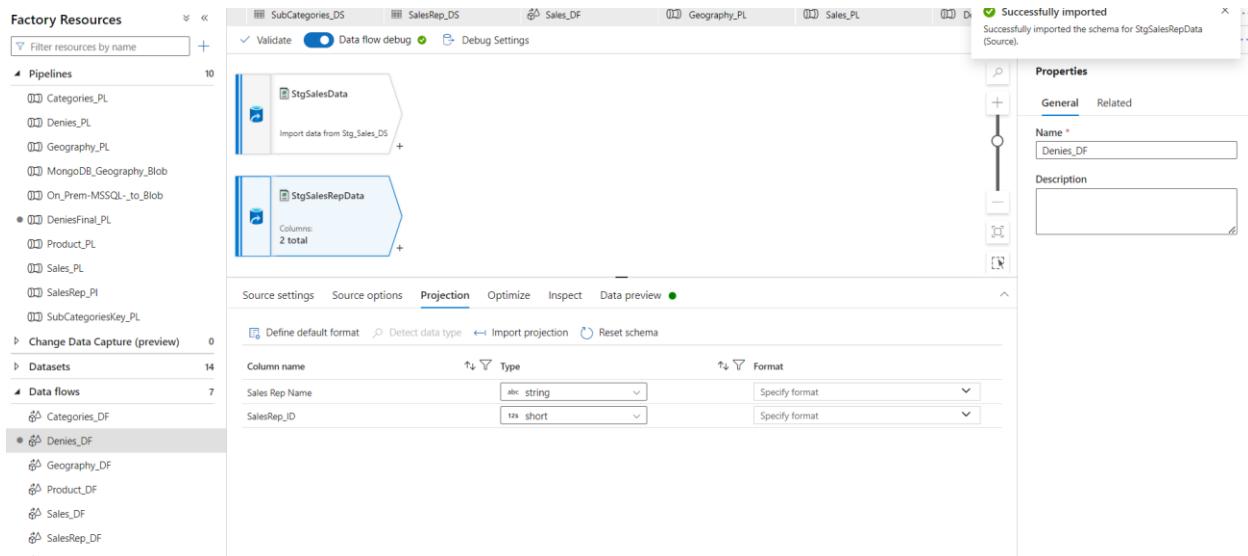
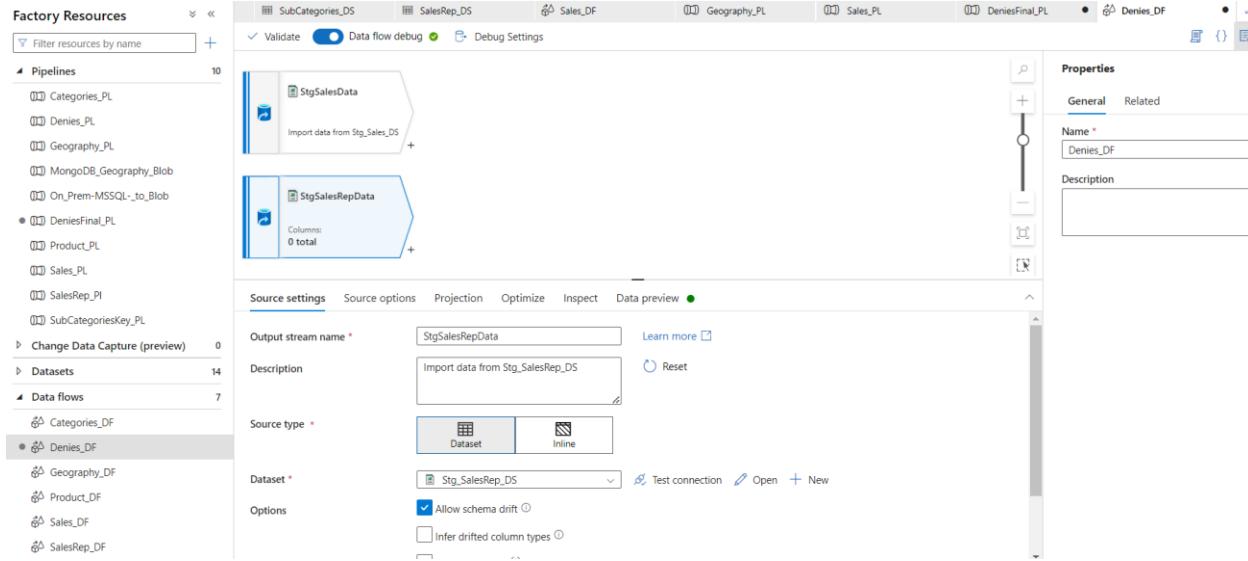
Source settings

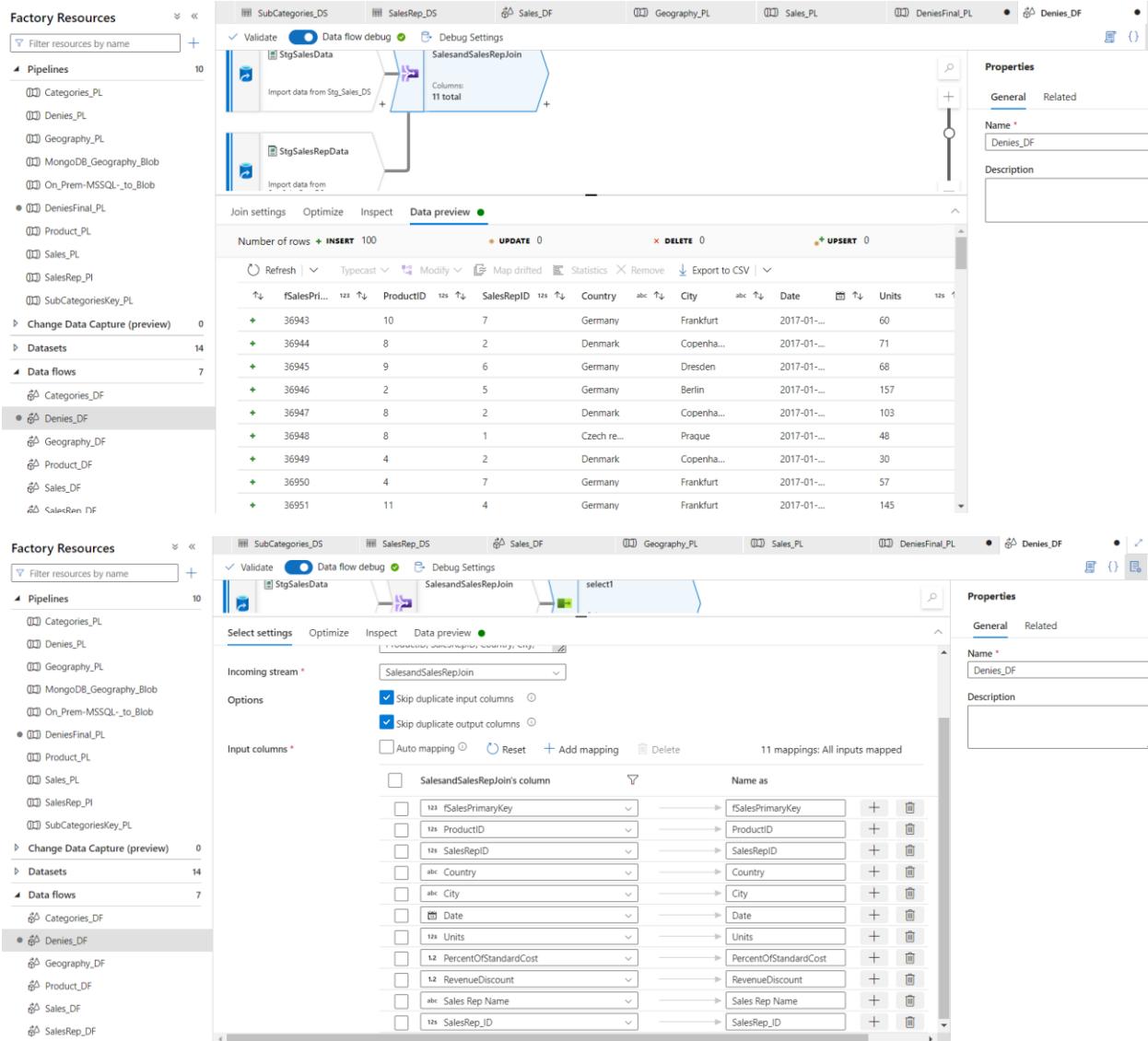
Projection

Data preview

Number of rows: 100

	INSERT	UPDATE	DELETE	UPSERT	LOOKUP		
Refresh	▼	Typecast	▼	Map drifted	▼		
+	36943	10	7	Germanv	Frankfurt	2017-01...	60
+	36944	8	2	Denmark	Copenha...	2017-01...	71
+	36945	9	6	Germanv	Dresden	2017-01...	68
+	36946	2	5	Germanv	Berlin	2017-01...	157
+	36947	8	2	Denmark	Copenha...	2017-01...	103
+	36948	8	1	Czech re...	Praque	2017-01...	48





Data Factory Validate all Publish all

Factory Resources

- Pipelines
 - Categories_PL
 - Denies_PL
 - Geography_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - DeniesFinal_PL
 - Product_PL
 - Sales_PL
 - SalesRep_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 14
- Data flows 7
 - Categories_DF
 - Denies_DF
 - Geography_DF
 - Product_DF
 - Sales_DF
 - SalesRep_DF

SubCategories_DS SalesRep_DS Sales_DF Geography_PL Sales_PL DeniesFinal_PL

Validate Data flow debug Debug Settings

StgSalesData SalesandSalesRepJoin select1

Select settings Optimize Inspect Data preview

Description Renaming SalesandSalesRepJoin to select1 with columns 'ISalesPrimaryKey', 'ProductID', 'Sales Rep Name', 'Country', 'City', 'Date', 'Units', 'PercentOfStandardCost', 'RevenueDiscount'

Incoming stream SalesandSalesRepJoin

Options Skip duplicate input columns Skip duplicate output columns

Input columns Auto mapping Reset Add mapping Delete 9 mappings: 2 column(s) from the inputs left unmapped

SalesandSalesRepJoin's column	Name as
ISalesPrimaryKey	ISalesPrimaryKey
ProductID	ProductID
Sales Rep Name	Sales Rep Name
Country	Country
City	City
Date	Date
Units	Units
PercentOfStandardCost	PercentOfStandardCost
RevenueDiscount	RevenueDiscount

Properties

General Related

Name * Denies_DF

Description

Data Factory Validate all Publish all

Factory Resources

- Pipelines
 - Categories_PL
 - Denies_PL
 - Geography_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - DeniesFinal_PL
 - Product_PL
 - Sales_PL
 - SalesRep_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 14
- Data flows 7
 - Categories_DF
 - Denies_DF
 - Geography_DF
 - Product_DF
 - Sales_DF
 - SalesRep_DF

SubCategories_DS SalesRep_DS Sales_DF Geography_PL Sales_PL DeniesFinal_PL

Validate Data flow debug Debug Settings

StgSalesData select1

Select settings Optimize Inspect Data preview

Number of rows + INSERT 100 UPDATE 0 DELETE 0 UPSERT 0 LOOKUP 0

Refresh Typecast Modify Map drifted Statistics Remove Export to CSV

ISalesPri...	ProductID	Sales Re...	Country	City	Date	Units
36943	10	John White	Germany	Frankfurt	2017-01-...	60
36944	8	Bill Muray	Denmark	Copenhagen	2017-01-...	71
36945	9	Jan Novo...	Germany	Dresden	2017-01-...	68
36946	2	Ellen Wo...	Germany	Berlin	2017-01-...	157
36947	8	Bill Muray	Denmark	Copenhagen	2017-01-...	103
36948	8	Ellie Gill	Czech re...	Prague	2017-01-...	48
36949	4	Bill Muray	Denmark	Copenhagen	2017-01-...	30
36950	4	John White	Germany	Frankfurt	2017-01-...	57
36951	11	El Bob	Germany	Frankfurt	2017-01-...	145
36952	5	Ellen Wo...	Germany	Berlin	2017-01-...	31
36953	9	Jan Novo...	Germany	Dresden	2017-01-...	87
36954	9	Mark Spa...	Czech re...	Ostrava	2017-01-...	50
36955	3	Ellie Gill	Czech re...	Prague	2017-01-...	72
36956	1	Franz Ma...	Germany	Berlin	2017-01-...	76

Properties

General Related

Name * Denies_DF

Description

Stg Product data

Data Factory Validate all Publish all

Factory Resources

- Pipelines
 - Categories_PL
 - Denies_PL
 - Geography_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - DeniesFinal_PL
 - Product_PL
 - Sales_PL
 - SalesRep_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 14
- Data flows 7
 - Categories_DF
 - Denies_DF
 - Geography_DF
 - Product_DF
 - Sales_DF
 - SalesRep_DF

SubCategories_DS SalesRep_DS Sales_DF Geography_PL Sales_PL DeniesFinal_PL

Validate Data flow debug Debug Settings

Import data from Stg_SalesRep_DS StgProductData

Source settings Source options Projection Optimize Inspect Data preview

Output stream * StgProductData Learn more

Description Import data from Stg_Product_DS

Source type Dataset Inline

Dataset Stg_Product_DS Test connection Open New

Options Allow schema drift Infer drifted column types Validate schema

Skip line count

Properties

General Related

Name * Denies_DF

Description

Data Flow Editor - Denies_DF

```

graph TD
    StgProductData[StgProductData] --> DeniesDF[Denies_DF]

```

Column name	Type	Format
ProductID	short	Specify format
Sub_Category_Key	short	Specify format
ProductName	string	Specify format
Color	string	Specify format
RetailPrice	double	Specify format
StandardCost	double	Specify format

Data Flow Editor - DeniesFinal_PL

```

graph TD
    StgProductData[StgProductData] --> DeniesFinalPL[DeniesFinal_PL]

```

Column name	Type	Format
ProductID	short	Specify format
Sub_Category_Key	short	Specify format
ProductName	string	Specify format
Color	string	Specify format
RetailPrice	double	Specify format
StandardCost	double	Specify format

Data Flow Editor - Denies_DF

```

graph TD
    StgProductData[StgProductData] --> DeniesDF[Denies_DF]

```

Column name	Type	Format
ProductID	short	Specify format
Sub_Category_Key	short	Specify format
ProductName	string	Specify format
Color	string	Specify format
RetailPrice	double	Specify format
StandardCost	double	Specify format

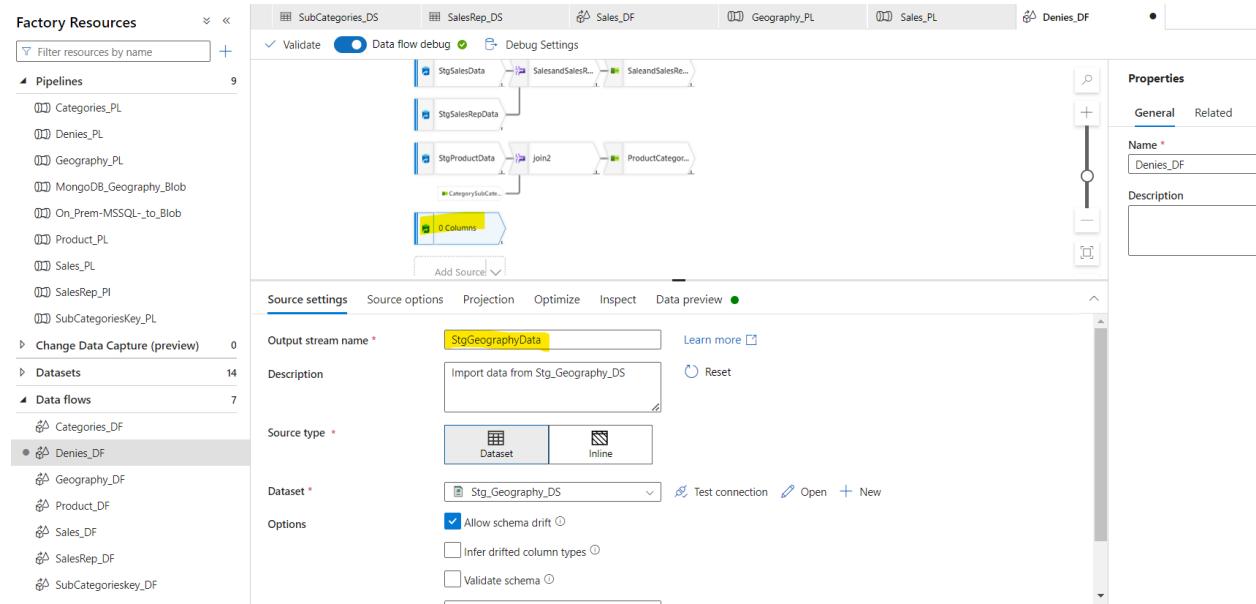
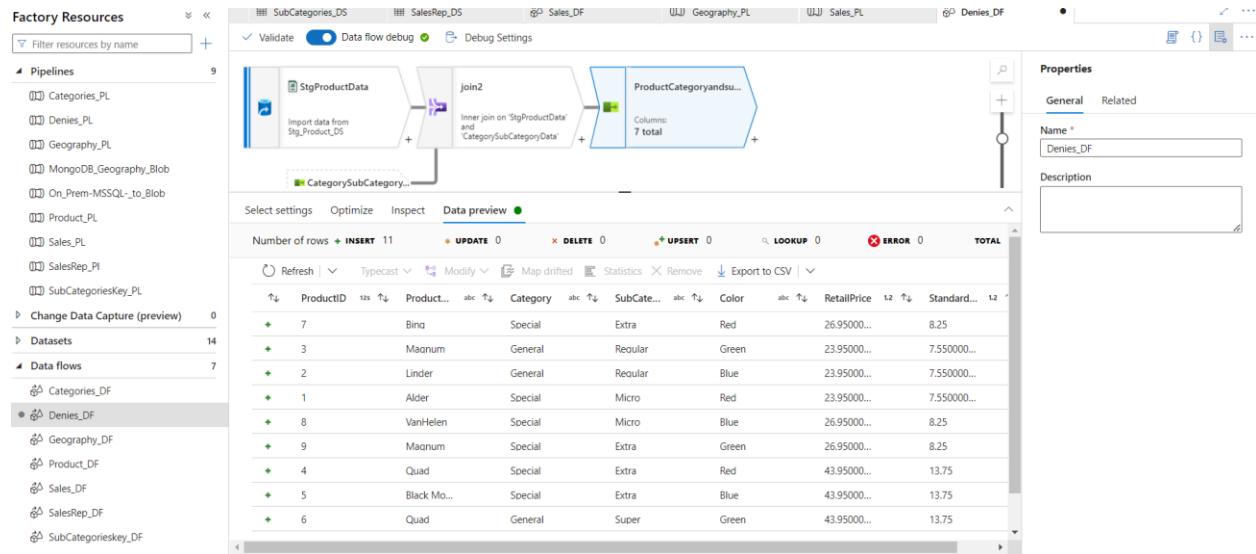
Data Flow Editor - DeniesFinal_PL

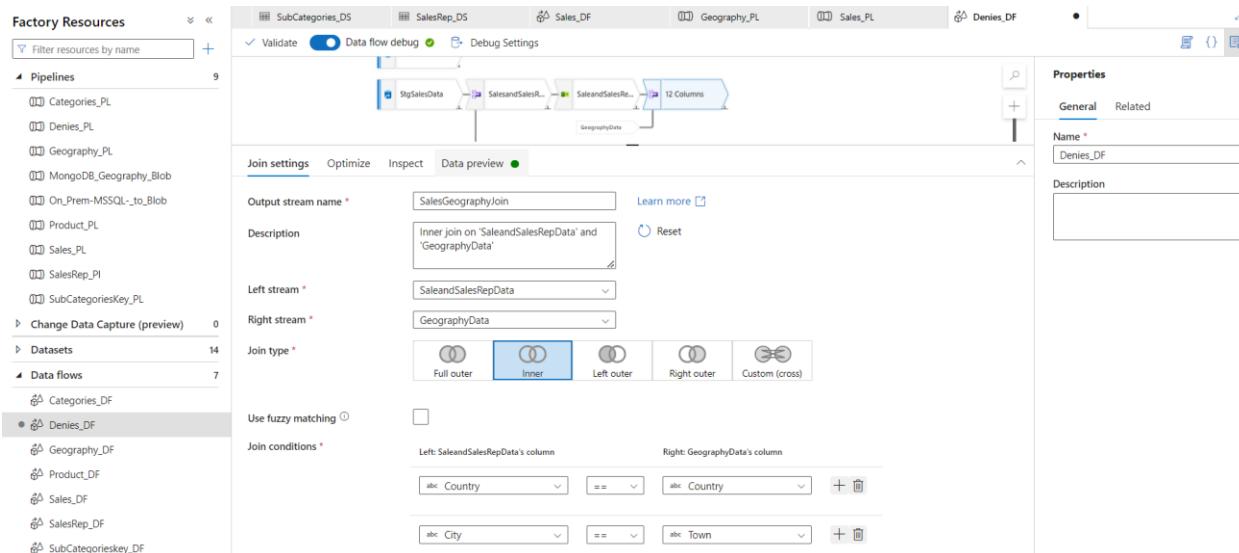
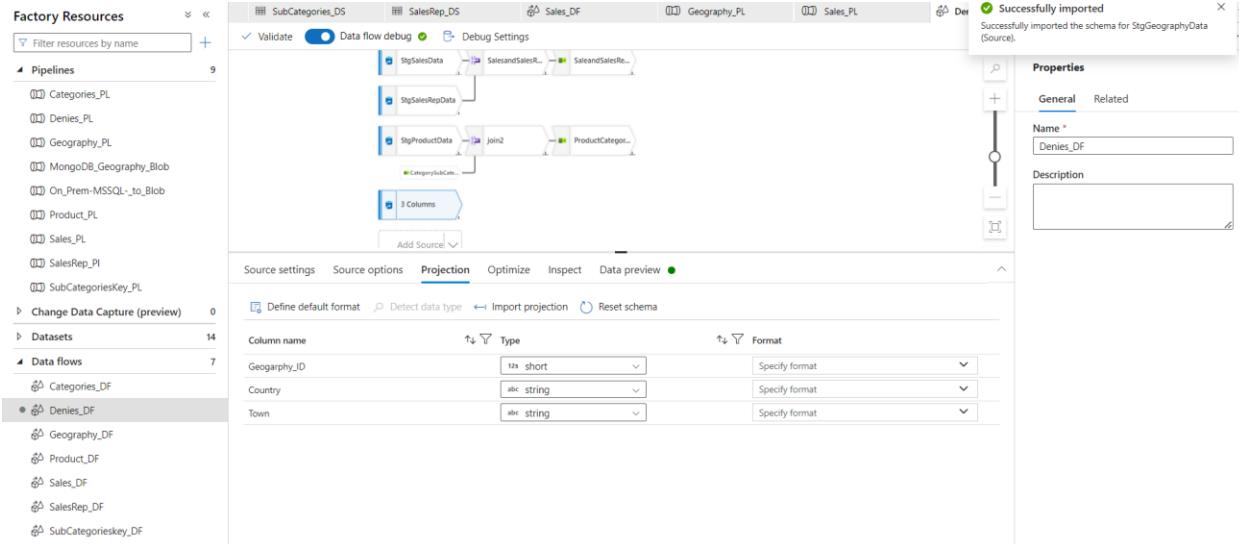
```

graph TD
    StgProductData[StgProductData] --> DeniesFinalPL[DeniesFinal_PL]

```

Column name	Type	Format
ProductID	short	Specify format
Sub_Category_Key	short	Specify format
ProductName	string	Specify format
Color	string	Specify format
RetailPrice	double	Specify format
StandardCost	double	Specify format





Factory Resources

- Pipelines
 - Categories_PL
 - Denies_PL
 - Geography_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - Product_PL
 - Sales_PL
 - SalesRep_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 14
- Data flows 7
 - Categories_DF
 - * Denies_DF
 - Geography_DF
 - Product_DF
 - Sales_DF
 - SalesRep_DF
 - SubCategorieskey_DF

Data flow debug **Validate** **Debug Settings**

Properties

General

Name * Denies_DF

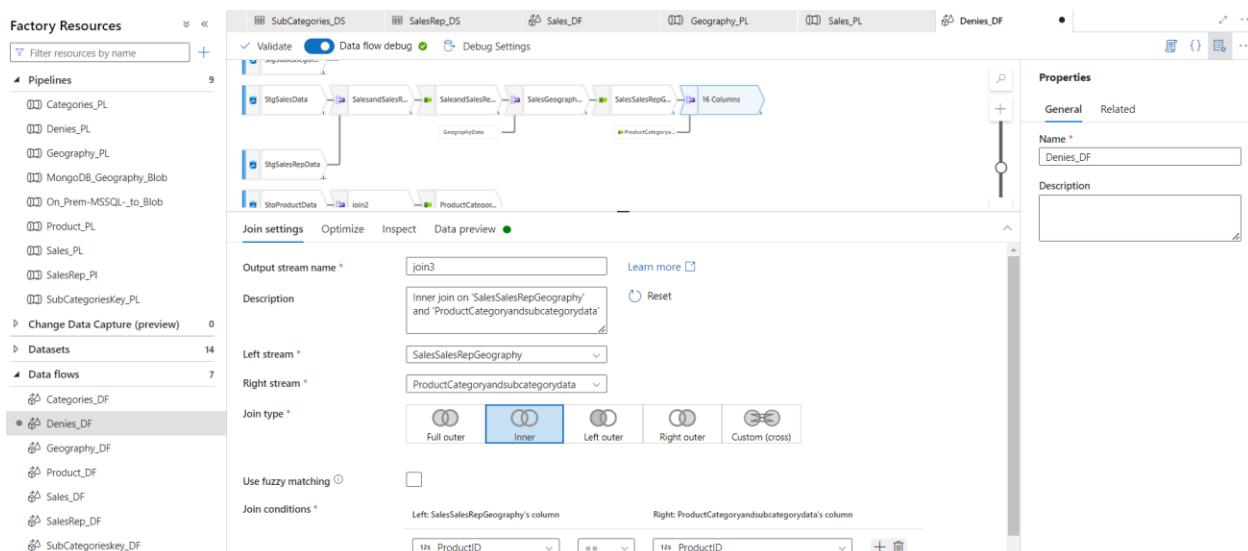
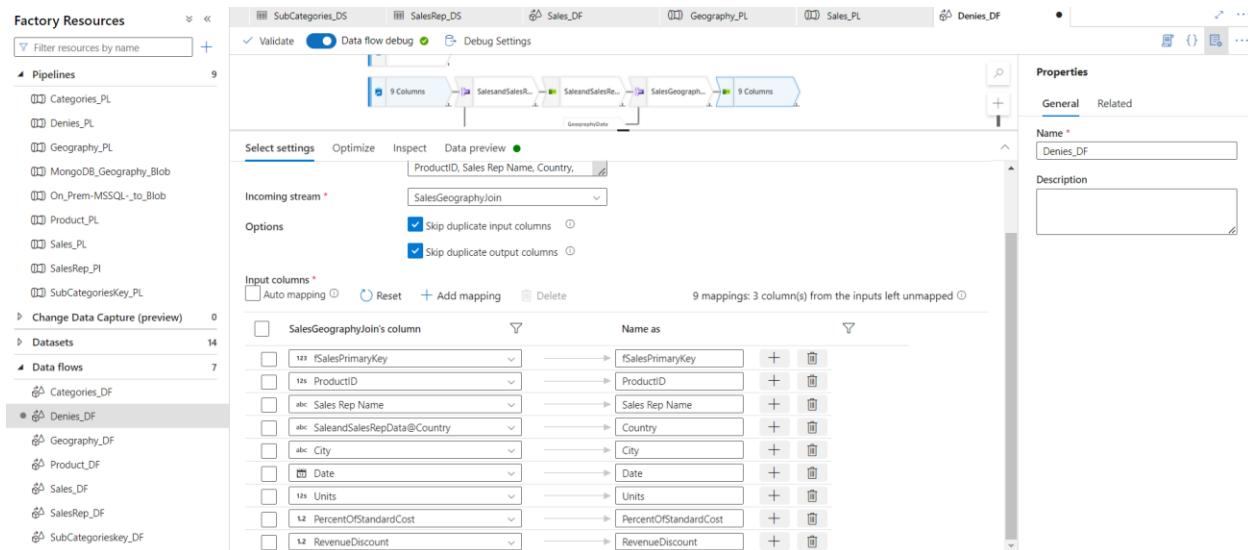
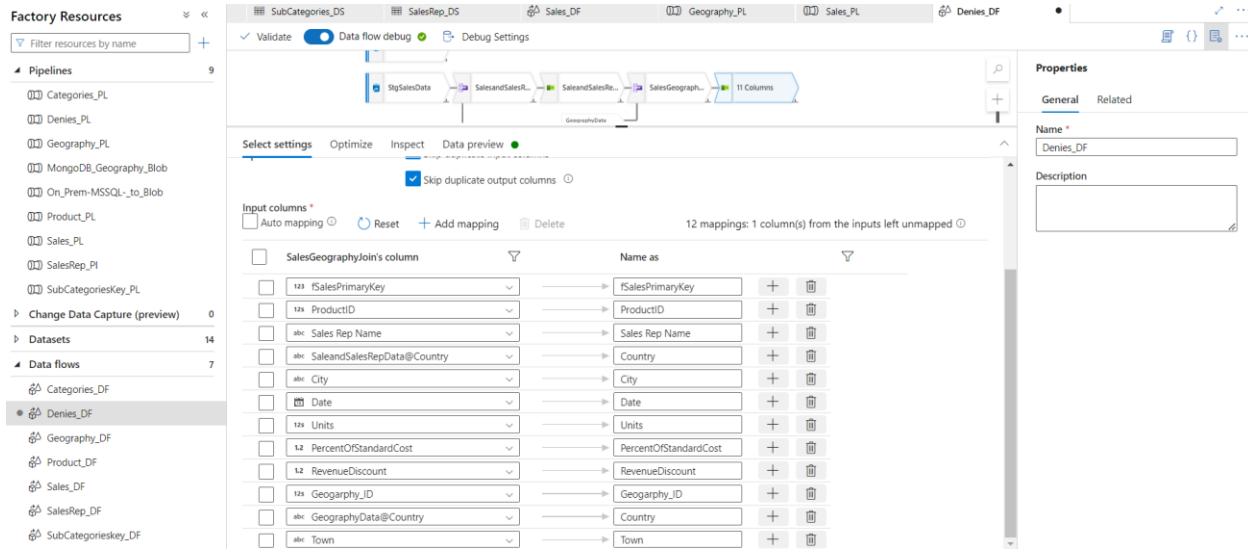
Description

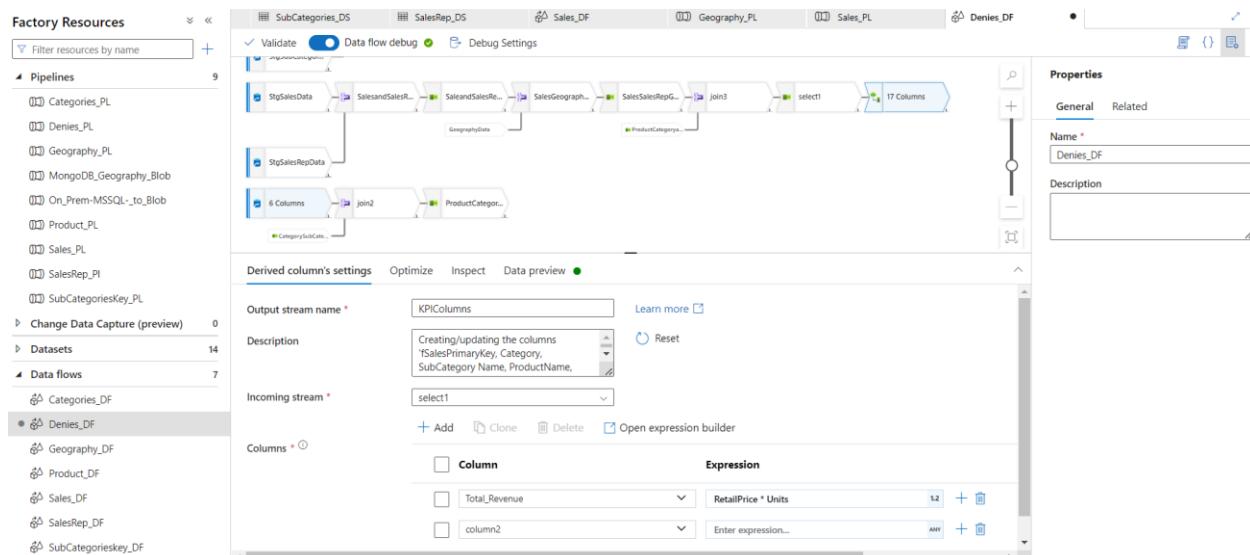
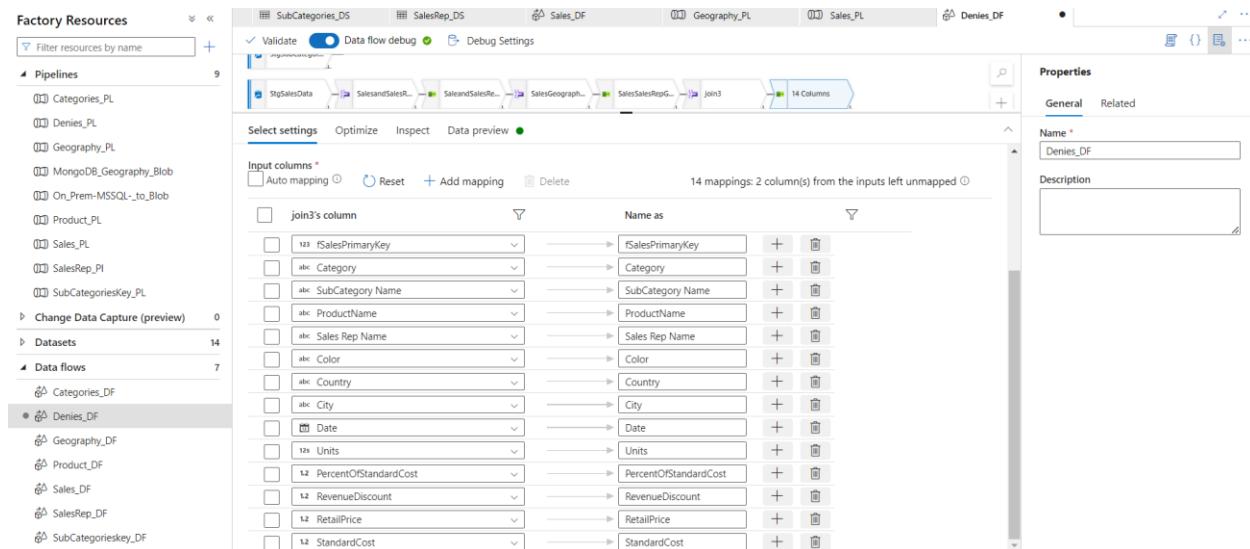
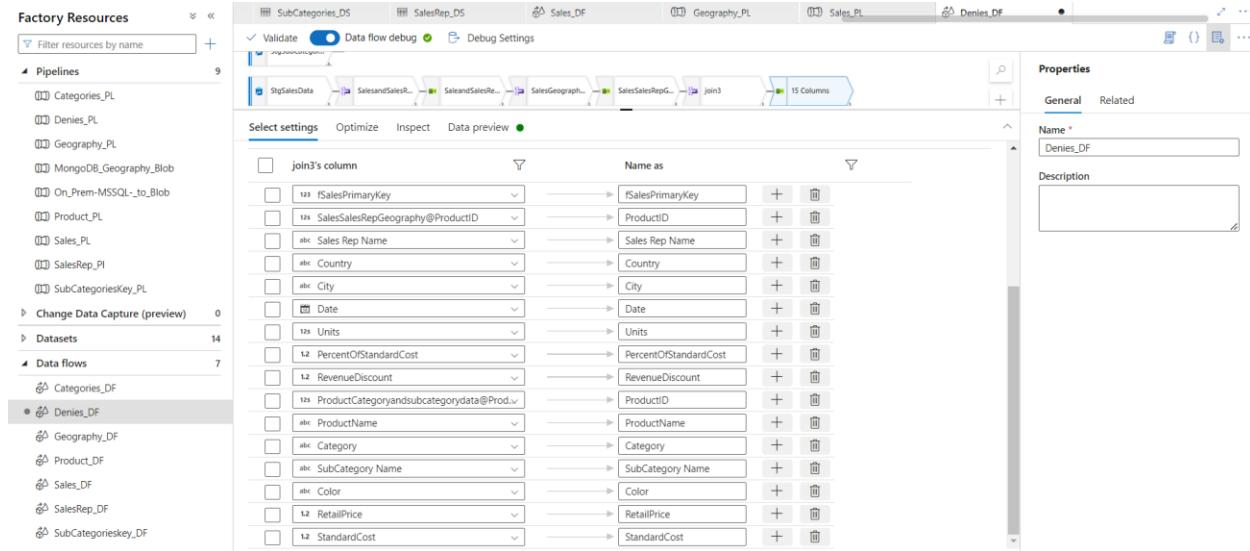
Data preview

Number of rows + INSERT 100 UPDATE 0 DELETE 0 UPSERT 0

FSalesPri...	ProductID	Sales Re...	Country	Country	City	Date
36943	10	John Wh...	Germany	Germany	Frankfurt	2017-01-...
36944	8	Bill Muray	Denmark	Denmark	Copenhagen...	2017-01-...
36945	9	Jan Nov...	Germany	Germany	Dresden	2017-01-...
36946	2	Ellen Wo...	Germany	Germany	Berlin	2017-01-...
36947	8	Bill Muray	Denmark	Denmark	Copenhagen...	2017-01-...
36948	8	Ellie Gill	Czech re...	Czech re...	Praque	2017-01-...
36949	4	Bill Muray	Denmark	Denmark	Copenhagen...	2017-01-...
36950	4	John Wh...	Germany	Germany	Frankfurt	2017-01-...
36951	11	Ei Bob	Germany	Germany	Frankfurt	2017-01-...
36952	5	Ellen Wo...	Germany	Germany	Berlin	2017-01-...
36953	9	Jan Nov...	Germany	Germany	Dresden	2017-01-...
36954	0	Mark Sn...	Czech re...	Czech re...	Ostrava	2017-01...

Refresh **Typecast** **Modify** **Map drifted** **Statistics** **Remove** **Export to CSV**





Dataflow expression builder

Expression reference documentation ▾

KPIColumns

Derived Columns

+ Create new ▾

All Total_Revenue
column2
column1

Column name *
Total_Revenue

Expression
RetailPrice * Units

Save

Expression elements

All Functions Input schema Parameters Cached lookup Data flow library functions Locals

Expression values

Filter by keyword

Create new

123 ISalesPrimaryKey
abc Category
abc SubCategory Name
abc ProductName
abc Sales Rep Name

Color

Data preview Refresh

Save and finish Cancel Clear contents

Factory Resources

Filter resources by name +

Pipelines

- Categories_PL
- Denies_PL
- Geography_PL
- MongoDB_Geography_Blob
- Cn_Prem-MSSQL-.to_Blob
- Product_PL
- Sales_PL
- SalesRep_PL
- SubCategoriesKey_PL

Change Data Capture (preview) 0

Datasets 14

Data flows 7

- Categories_DF
- Denies_DF
- Geography_DF
- Product_DF
- Sales_DF
- SalesRep_DF

StgSalesData → SalesandSalesRep → SalesGeography → SalesSalesRepG → select1 → 17 Columns

StgSalesRepData → join2 → ProductCategory → CategorySubCategory

Properties

Name * Denies_DF

Description

Derived column's settings

Output stream name * KPIColumns

Description Creating/updating the columns ISalesPrimaryKey, Category, SubCategory Name, ProductName.

Incoming stream * select1

Add Clone Delete Open expression builder

Columns

Column	Expression
Total_Revenue	RetailPrice * Units
Total_Cost	StandardCost * Units

Learn more ▾

Optimize Inspect Data preview

Validate Data flow debug Debug Settings

Dataflow expression builder

Expression reference documentation ▾

KPIColumns

Derived Columns

+ Create new ▾

All Total_Revenue
Total_Cost
column1

Column name *
Total_Cost

Expression
StandardCost * Units

Save

Expression elements

All Functions Input schema Parameters Cached lookup Data flow library functions Locals

Expression values

Filter by keyword

Create new

123 ISalesPrimaryKey
abc Category
abc SubCategory Name
abc ProductName
abc Sales Rep Name

Color

Data preview Refresh

Save and finish Cancel Clear contents

Factory Resources

Filter resources by name +

Pipelines

- Categories_PL
- Denies_PL
- Geography_PL
- MongoDB_Geography_Blob
- Cn_Prem-MSSQL-.to_Blob
- Product_PL
- Sales_PL
- SalesRep_PL
- SubCategoriesKey_PL

Change Data Capture (preview) 0

Datasets 14

Data flows 7

- Categories_DF
- Denies_DF
- Geography_DF
- Product_DF
- Sales_DF
- SalesRep_DF

StgSalesData → SalesandSalesRep → SalesGeography → SalesSalesRepG → select1 → 17 Columns

StgSalesRepData → join2 → ProductCategory → CategorySubCategory

Properties

Name * Denies_DF

Description

Derived column's settings

Output stream name * KPIColumns

Description Creating/updating the columns ISalesPrimaryKey, Category, SubCategory Name, ProductName.

Incoming stream * select1

Add Clone Delete Open expression builder

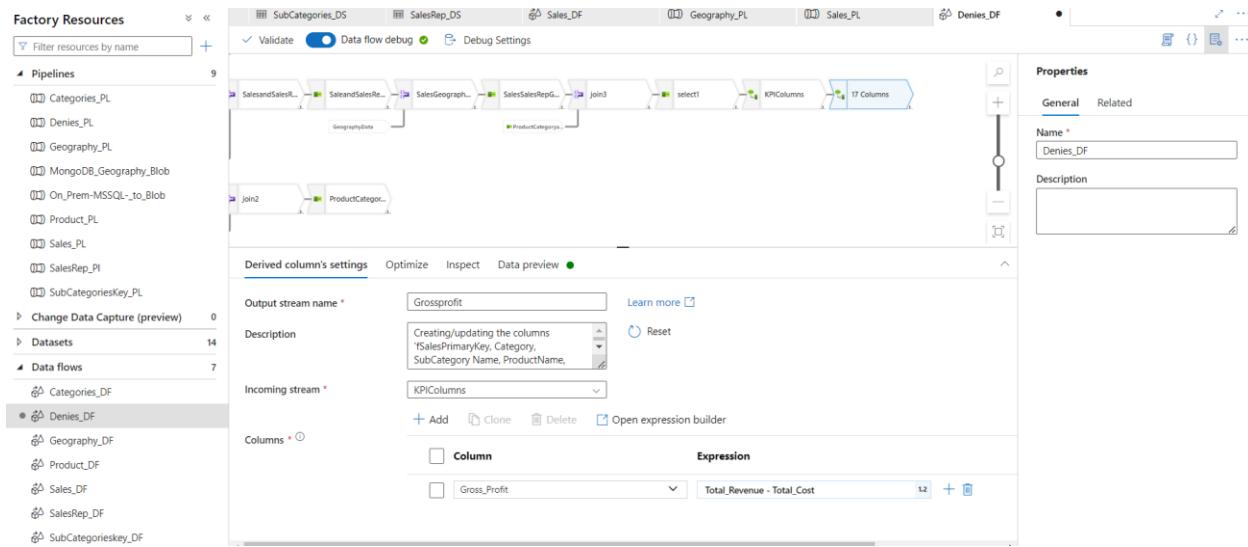
Columns

Column	Expression
Total_Revenue	RetailPrice * Units
Total_Cost	StandardCost * Units

Learn more ▾

Optimize Inspect Data preview

Validate Data flow debug Debug Settings



Dataflow expression builder

Grossprofit

Derived Columns

Create new

Gross_Profit

Column name: Gross_Profit

Expression: Total_Revenue - Total_Cost

Save

Expression elements

All

Functions

Input schema

Parameters

Cached lookup

Data flow library functions

Locals

Expression values

Filter by keyword

Create new

fSalesPrimaryKey

Category

SubCategory Name

ProductName

Sales Rep Name

Save

Data preview Refresh

Save and finish Cancel Clear contents

Factory Resources

- Pipelines
- Datasets
- Data flows
- Categories_DF
- Denies_DF**
- Geography_DF
- Product_DF
- Sales_DF
- SalesRep_DF
- SubCategorieskey_DF

SubCategories_DS SalesRep_DS Sales_DF Geography_PL Denies_DF

Validate Data flow debug Debug Settings

Sink

Output stream name: DeniesFinal

Description: Add sink dataset

Incoming stream: Grossprofit

Sink type: Dataset

Dataset: Select... New

Options: Allow schema drift (checked), Validate schema

In pipeline activities and data flows, reference a dataset to specify the location and structure of your data within a data store. Learn more

Select a data store

Search

All Azure Database File Generic protocol NoSQL Services and apps

Azure Data Lake Storage Gen2	Azure Database for MySQL	Azure Database for PostgreSQL
Azure SQL Database Managed Instance	Azure Synapse Analytics	

Continue Cancel

Factory Resources

- Pipelines
 - Categories_PL
 - Denies_PL
 - Geography_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - Product_PL
 - Sales_PL
 - SalesRep_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 14
- Data flows 7
 - Categories_DF
 - Denies_DF
 - Geography_DF
 - Product_DF
 - Sales_DF
 - SalesRep_DF
 - SubCategorieskey_DF

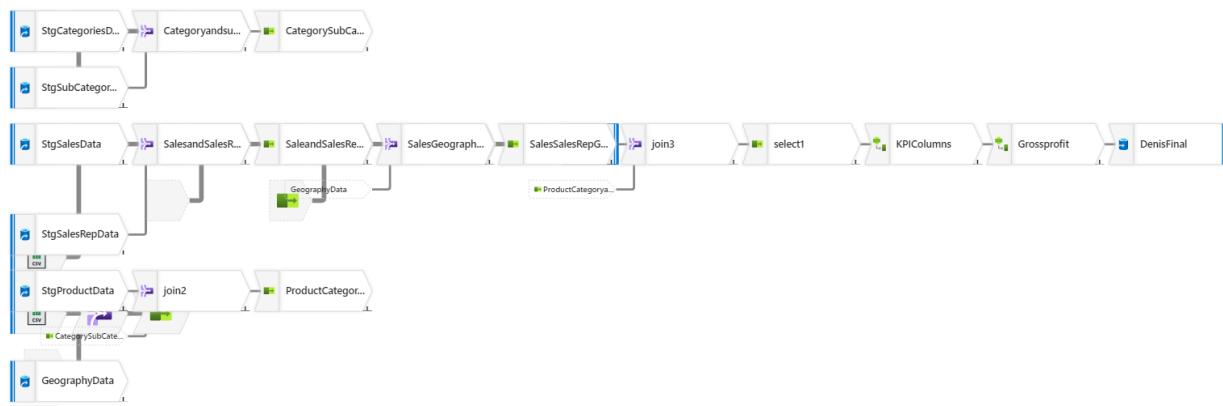
Denies_DF Data Flow Configuration

General

- Name: Denies_Data_DS
- Linked service: ADF_AZSQL_LS
- Schema and table name: Schema name: Denies_Data_TB
- Output stream name: DenisFinal
- Description: Add sink dataset
- Incoming stream: Grossprofit
- Sink type: Dataset
- Dataset: Select...
- Options: Allow schema drift (checked), Validate schema (unchecked)

Advanced

OK Back Cancel



Factory Resources

- Pipelines 10
 - Categories_PL
 - Denies_PL
 - Geography_PL
 - MongoDB_Geography_Blob
 - On_Prem-MSSQL_to_Blob
 - DeniesFinal_PL (selected)
 - Product_PL
 - Sales_PL
 - SalesRep_PL
 - SubCategoriesKey_PL
- Change Data Capture (preview) 0
- Datasets 15
- Data flows 7
- Power Query 0

Activities

- Move and transform
 - Copy data
 - Data flow (selected)
- Synapse
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning
- Power Query

Properties

General

- Name: DeniesFinal_PL
- Description:
- Annotations: + New

Settings

- Data flow: Denies_DF
- Run on (Azure IR): AutoResolveIntegrationRuntime
- Compute size: Small
- Logging level: Verbose (checked)
- Sink properties
- Staging

