



Lab 04: Dataflow Gen 2 in action

Introduction

In this lab, you will use Dataflow Gen2 to execute some advanced Data Preparation activities. You will first consume a complex JSON file with 2 methods, then you will extract and transform data from an Open Data environment based on REST API.

Objectives

After completing this lab, you will be better able to:

1. Import and Transform Data from a complex data structure
2. Understand how to implement advanced M code
3. Import and Transform Data from a REST API call

Estimated time to complete this lab

60 minutes

Contents

Lab 04: Dataflow Gen 2 in action.....	1
Introduction	1
Objectives.....	1
Task 1: Import and Transform Data from a complex data structure	3
Task 2: Execute advanced M code.....	14
Task 3: Extract data from OData.....	19

Lab Prerequisites

- Workspace: Fabric, Power Premium or Fabric trial
- Individual license: Power Pro or Premium Per User account

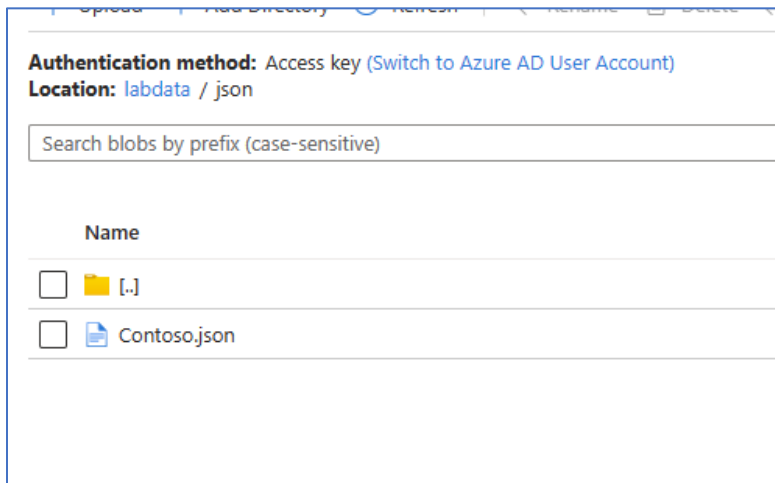
Information provided by your training provider

- Trial tenant (if applicable): login & password, workspace to use for the lab.
- Azure Data Lake Gen2 (containing data sources): account name & shared access signature.

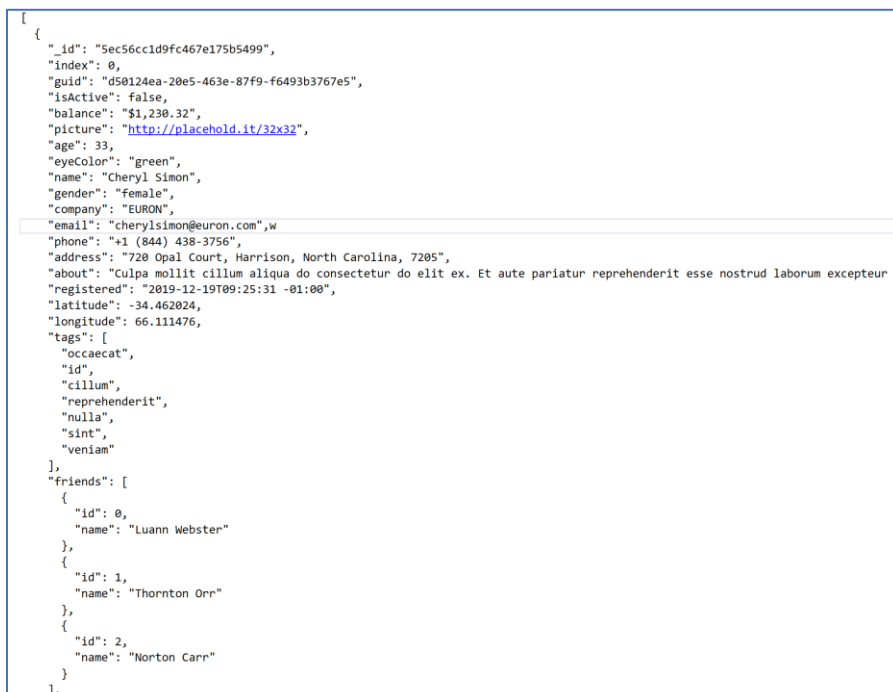
Task 1: Import and Transform Data from a complex data structure

In this task, you will import data from a JSON file then manipulate the structure before loading it to the Lakehouse.

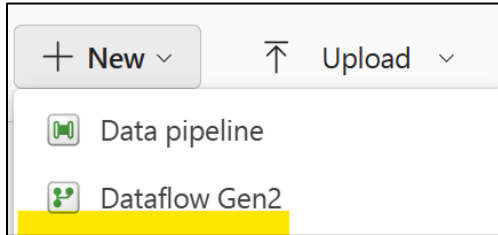
The file to be loaded is name **Contoso.json** and is stored in the labdate/json folder from the Azure Data Lake Store Gen 2 account used in the previous labs.



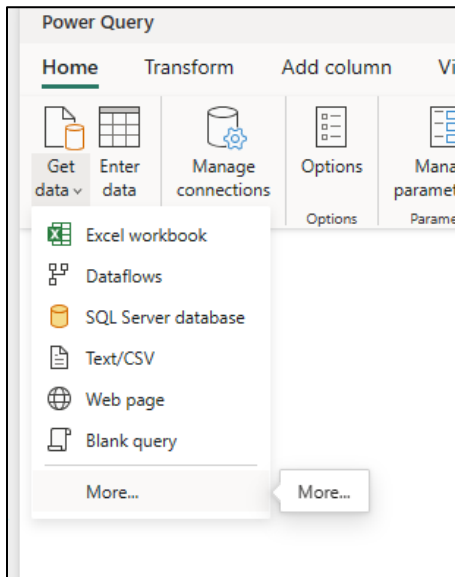
This file contains a list of contacts, with several attributes (age, gender...) and some multiple value attributes (tags, friends).



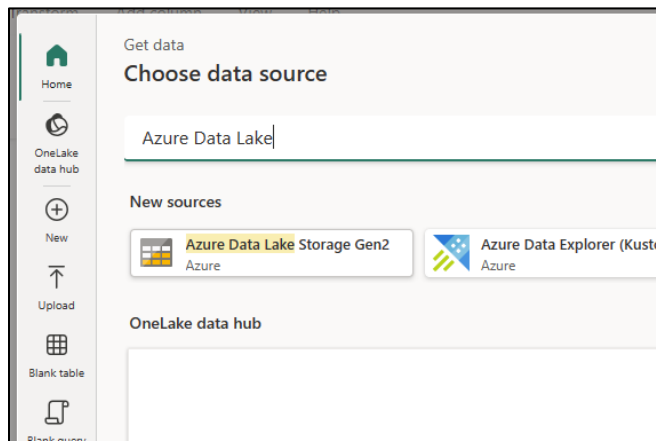
Create a **Dataflow Gen2** artifact in your workspace.



Go to the **Get Data** Menu and click on **More**.



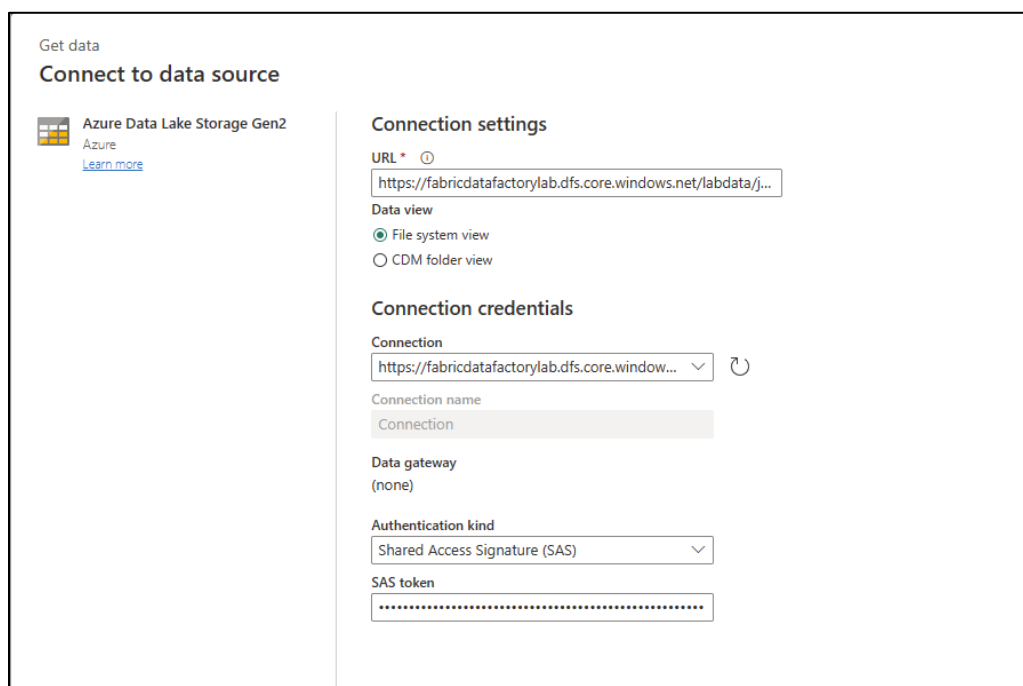
On the Choose data source window, type **Azure Data Lake** and select **Azure Data Lake Storage Gen2**.



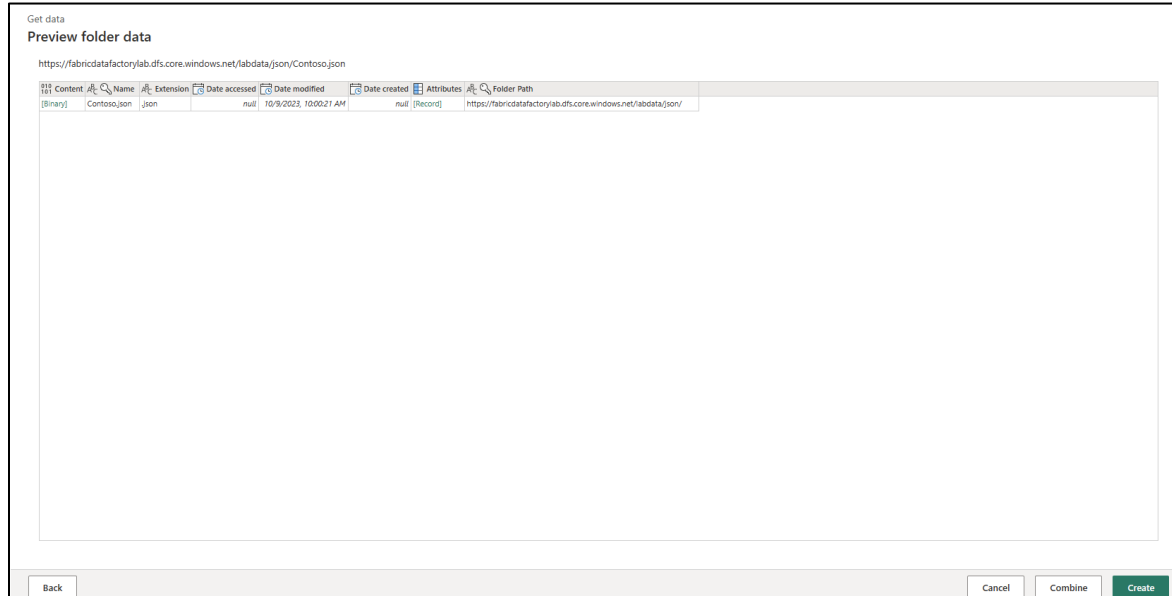
Configure the connection to the Contoso json file stored in the ADLS Gen2 account using a Shared Access Signature.

- **URL:** <https://fabricvbdadlspublic.dfs.core.windows.net/labdata/JSON/Contoso.json>
- **Authentication kind:** Shared Access Signature (SAS)

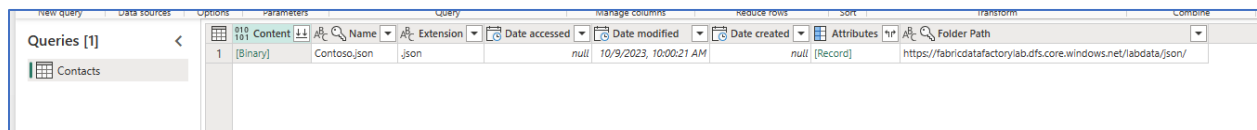
SAS Token: `sv=2022-11-02&ss=bf&srt=sco&sp=rlx&se=2024-05-01T02:07:21Z&st=2024-04-10T18:07:21Z&spr=https&sig=Rw%2BE248HSJssu2xV4%2FbLogGXUowVjVHQ98niqWaL0Uo%3D`



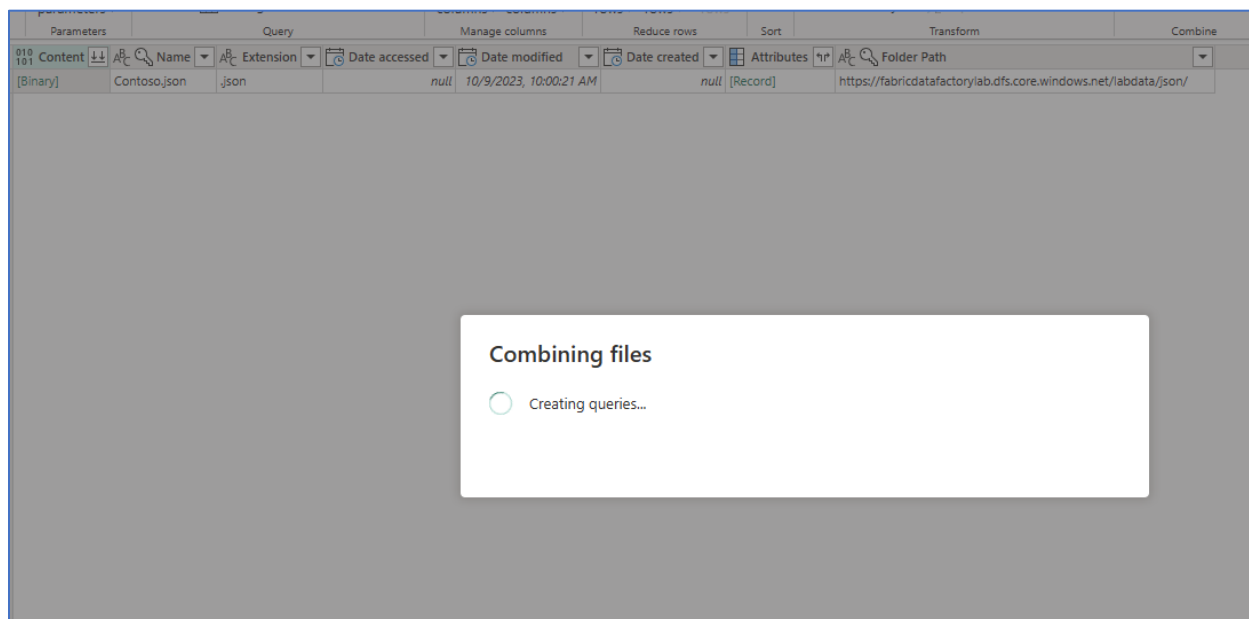
Click on **Create** to edit the query and extract data from the Contoso.json file.



Rename the query named "Query 1" as "Contacts" in the Queries panel, and click on the icon in the **Content** column header to ask the Dataflow Gen2 to explorer and extract the data from JSON file.



The Dataflow engine will automatically analyze the file structure and display the query result.

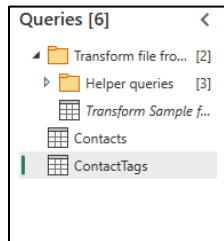


Upskilling on MS Fabric Data Factory

The query result is displayed on the Data Grid, and collection of resources have been automatically added.

	id	registered	latitude	longitude	tags	friends	greeting	favoritefruit
1	andent esse nostrud laborum excepteur fugiat nostrud. Lorem esse commodo ad reprehenderit est occaecat occaecat ad irure est. Ad...	12/19/2019, 9:25:31 AM -01:00	-34.462024	66.111476	[List]	[List]	Hello, Cheryl Simon! You have 8 unread messages.	strawberry
2	irt in reprehenderit adipiscing sint adipiscing esse dolor aliqua minim in irure fugiat. Panatur eiusmod proident eiusmod enim culpa ...	10/7/2017, 2:11:58 AM -02:00	-71.651883	46.040833	[List]	[List]	Hello, Rollins McIntyre! You have 8 unread messages.	apple
3	consequat. Exercitation irure non ad culpa ad aliquip non dolor nisi dolore et esse. Eu in velit ad consequat eu ex in commodo tempo...	9/4/2017, 4:29:37 AM -02:00	-74.224428	59.994114	[List]	[List]	Hello, Roxanne Carrillo! You have 1 unread messages.	apple
4	plit voluptate in ea eu culpa commodo velit. Cillum nisi aute sunt commodo. Proident anim aliqua et dolore proident eu sint qui moll...	8/15/2015, 9:04:40 AM -02:00	-51.757853	161.344879	[List]	[List]	Hello, Cornelia Stokes! You have 6 unread messages.	strawberry
5	s aliqua. Ea non proident veriam non aliquip eu non ullamco eiusmod ullamco voluptate do reprehenderit officia. Quis aute magna c...	12/5/2014, 4:10:19 AM -01:00	55.866653	-105.113761	[List]	[List]	Hello, Hayes Bender! You have 10 unread messages.	apple
6	erccitation quis reprehenderit officia nulla non quis consequat laboris labore pariatur dolor. Qui eiusmod do amet magna commodo l...	1/16/2017, 12:03:56 PM -01:00	52.43025	-65.747386	[List]	[List]	Hello, Annabelle Griffith! You have 4 unread messages.	banana
7	orem aute nostrud. Aliquip nisi sint cillum occaecat id sit. Minim proident ad sit ea aliqua eiusmod est laborum. Nostrud ut cupidatat ...	5/25/2016, 12:41:09 PM -02:00	39.025863	162.946564	[List]	[List]	Hello, Eva Gibson! You have 1 unread messages.	strawberry
8	at dolore aute et sit. Do id nulla proident commodo. Ex Lorem excepteur ipsum sit cillum est minim est dui est. Occaecat nostrud mi...	6/9/2016, 1:35:32 AM -02:00	-40.626388	-140.586797	[List]	[List]	Hello, Marcy Duffy! You have 9 unread messages.	banana
9	's. Minim voluptate quis ex ex labore tempor. Aute commodo sint amet et minim. Sit officia nulla sit aliqua sit labore do. Dolor eu offi...	8/11/2014, 11:45:34 AM -02:00	-27.539059	-45.977119	[List]	[List]	Hello, Lucy Berry! You have 4 unread messages.	apple
10	aliquip excepteur et. Deserunt enim nisi nisi aliqua amet ex incididunt. Ea fugiat laboris anim officia nulla nostrud consequat. Fugiat c...	10/24/2018, 10:54:40 AM -02:00	78.473938	71.181222	[List]	[List]	Hello, Leon McLaughlin! You have 2 unread messages.	apple
11	r fugiat dolore id. Aute adipiscing occaecat aliquip mollit proident tempor. Eiusmod velit quis labore esse. Veriam anim nulla in para...	9/7/2014, 9:12:24 AM -02:00	-56.207073	4.655699	[List]	[List]	Hello, Cochran Burks! You have 6 unread messages.	strawberry
12	ig ullamco non id. Est esse ex nisi dolore nostrud. Sit eu est cupidatat commodo incididunt sint tempor. Do irure dolor in non magna l...	6/17/2016, 7:41:12 AM -02:00	-63.79455	-21.520968	[List]	[List]	Hello, Suzette Gilmore! You have 5 unread messages.	apple

Select the **Contacts** query and clone it to create the new query **ContactTags** (to clone a query, right click, and select "Duplicate").



Select the **Contacts** query, and as indicated in the screenshot hereafter, select the following group of columns: age, eyeColor, name, gender, company, email and phone. Keep the selection active, right on the columns header and select **Remove other columns**.

	id	picture	age	eyeColor	name	gender	company	email	phone	ad
1	0.32	http://placeholder.it/32x32	33	green	Cheryl Simon	female	EURON	cheryl@euron.com	438-3756	720 C
2	4.09	http://placeholder.it/32x32	37	blue	Rollins McIntyre	male	GLUKGLUK	rollins@glukgluk.com	539-3799	335 C
3	0.22	http://placeholder.it/32x32	30	green	Roxanne Carrillo	female	KENGEN	roxanne@kengen.com	455-3405	724 K
4	2.79	http://placeholder.it/32x32	40	green	Cornelia Stokes	female	AEORA	cornelia@aeora.com	403-2507	356 D
5	9.68	http://placeholder.it/32x32	21	green	Hayes Bender	male	XIXAN	hayes@xixan.com	456-3196	248 G
6	7.91	http://placeholder.it/32x32	29	green	Annabelle Griffith	female	JETSILK	annabelle@jetsilk.com	524-3533	317 T
7	3.73	http://placeholder.it/32x32	40	blue	Eva Gibson	female	VITRICOMP	eva@vitricomp.com	484-3644	882 W
8	5.41	http://placeholder.it/32x32	28	blue	Marcy Duffy	female	FISHLAND	marcy@fishland.com	434-3257	299 D
9	8.84	http://placeholder.it/32x32	35	brown	Lucy Berry	female	GLUID	lucy@gluid.com	586-3546	757 L
10	9.77	http://placeholder.it/32x32	36	blue	Leon McLaughlin	male	CANOPOLY	leon@canopoly.com	569-2989	600 J
11	3.58	http://placeholder.it/32x32	27	brown	Cochran Burks	male	INVENTURE	cochranburks@inventure.com	+1 (842) 578-2777	938 M
12	5.52	http://placeholder.it/32x32	32	brown	Suzette Gilmore	female	ZEEDINA	suzette@zeedina.com	1 (852) 575-3852	525 B

The **Contacts** query should only contain the selected columns.

	age	eyeColor	name	gender	company	email	phone
1	33	green	Cheryl Simon	female	EURON	cherylsimon@euron.com	+1 (844) 438-3756
2	37	blue	Rollins McIntyre	male	GLUKGLUK	rollinsmcintyre@glukgluk.com	+1 (817) 539-3799
3	30	green	Roxanne Carrillo	female	KENGEN	roxannecarrillo@kengen.com	+1 (819) 455-3405
4	40	green	Cornelia Stokes	female	AEORA	corneliastokes@aeora.com	+1 (912) 403-2507
5	21	green	Hayes Bender	male	XIXAN	hayesbender@xixan.com	+1 (814) 456-3196
6	29	green	Annabelle Griffith	female	JETSILK	annabellegriffith@jetsilk.com	+1 (852) 524-3533
7	40	blue	Eva Gibson	female	VITRICOMP	evagibson@vitricomp.com	+1 (866) 484-3644
8	28	blue	Marcy Duffy	female	FISHLAND	marcyduffy@fishland.com	+1 (881) 434-3257
9	35	brown	Lucy Berry	female	GLUID	lucyberry@gluid.com	+1 (900) 586-3546
10	36	blue	Leon McLaughlin	male	CANOPOLY	leonmcLaughlin@canopoly.com	+1 (933) 569-2989
11	27	brown	Cochran Burks	male	INVENTURE	cochranburks@inventure.com	+1 (842) 578-2777
12	32	brown	Suzette Gilmore	female	ZERBINA	suzettegilmore@zerbina.com	+1 (862) 575-2853
13	39	green	Carol Velazquez	female	PROTODYNE	carovelazquez@protodyne.com	+1 (991) 483-3920
14	20	blue	Puckett Norris	male	ROOFORIA	puckettnorris@rooforia.com	+1 (881) 459-2737
15	40	blue	Barr Colon	male	BLUPLANET	barrcolon@bluplanet.com	+1 (928) 500-3883
16	35	brown	Sharron Bridges	female	EQUITOX	sharronbridges@equitox.com	+1 (891) 572-3608
17	24	brown	Russell Waters	male	BESTO	russellwaters@besto.com	+1 (942) 440-3056

Repeat the same operation for the **ContactTags** query, to keep the following columns: name and tags.

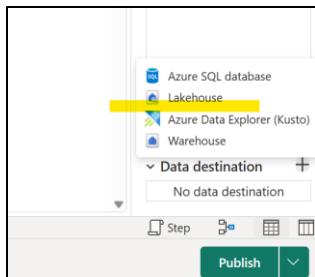
Then click on the small button located on the **tags** column header to expand the column values (which is a list).

	name	tags
1	Cheryl Simon	[List]
2	Rollins McIntyre	[List]
3	Roxanne Carrillo	[List]
4	Cornelia Stokes	[List]
5	Hayes Bender	[List]
6	Annabelle Griffith	[List]
7	Eva Gibson	[List]
8	Marcy Duffy	[List]
9	Lucy Berry	[List]
10	Leon McLaughlin	[List]
11	Cochran Burks	[List]
12	Suzette Gilmore	[List]

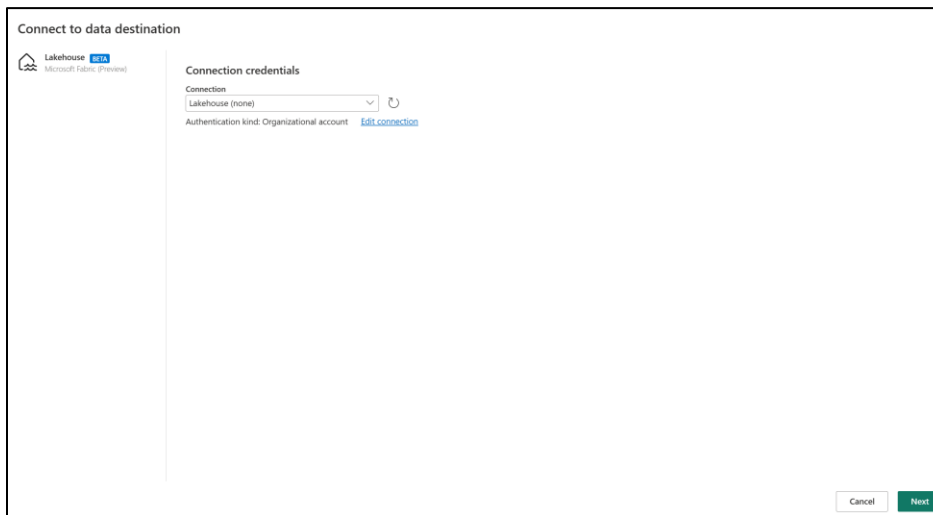
As the **ContactTags** query has been expanded, the same name appears who each corresponding tag value.

	name	tags
1	Cheryl Simon	occaecat
2	Cheryl Simon	id
3	Cheryl Simon	cillum
4	Cheryl Simon	reprehend...
5	Cheryl Simon	nulla
6	Cheryl Simon	sint
7	Cheryl Simon	veniam
8	Rollins McIntyre	ut
9	Rollins McIntyre	tempus

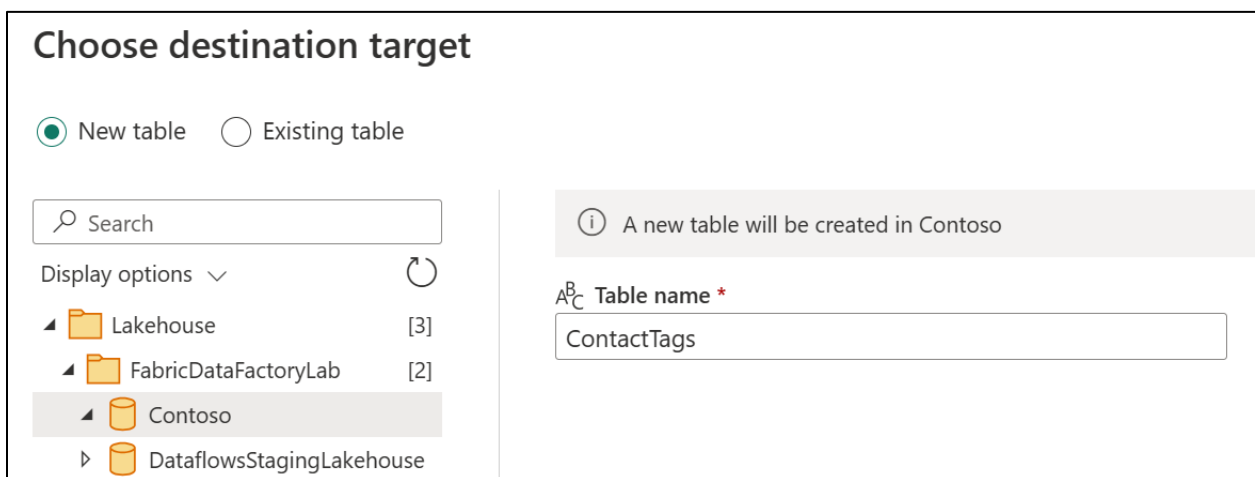
Click on the + icon in the **Data destination** area and select **Lakehouse**.



Keep the default configuration and click **Next**.



Select the existing **Contoso** lakehouse located in the same workspace, keep the destination table name as **ContactTags**, and click on **Next**.



Change Source type value as **Text** for the **tags** column and click on **Save settings**.

Choose destination settings

Update method

Existing data New data → Append Replace

Column mapping

Source	Source type	Destination	Destination type
name	Text	name	Text
tags	Text	tags	Text

Back Cancel Save settings

Repeat the same operation for the **Contacts** query to load data in the **Contoso** Lakehouse using **Contacts** as the destination table name.

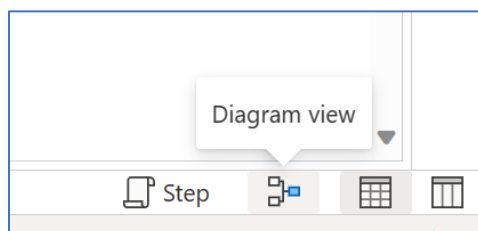
Display options

- Lakehouse [3]
 - FabricDataFactoryLab [2]
 - Contoso
 - DataflowsStagingLakehouse

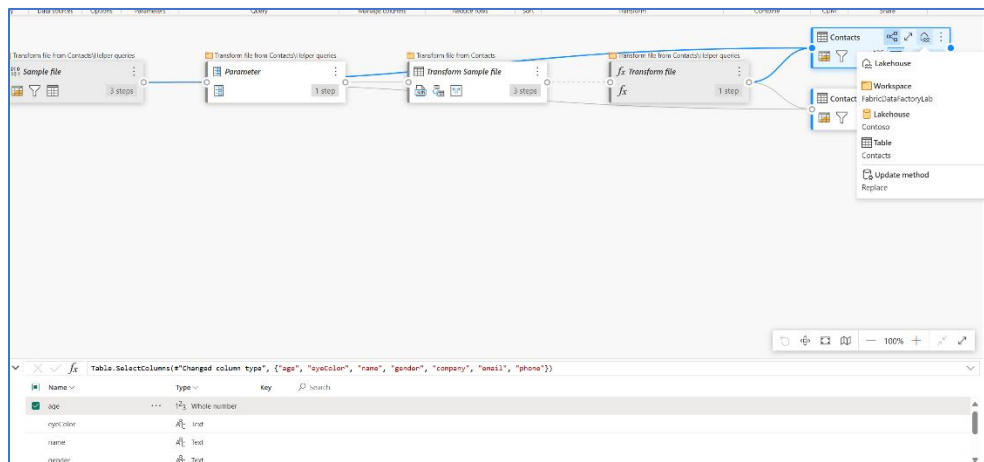
Table name *

Contacts

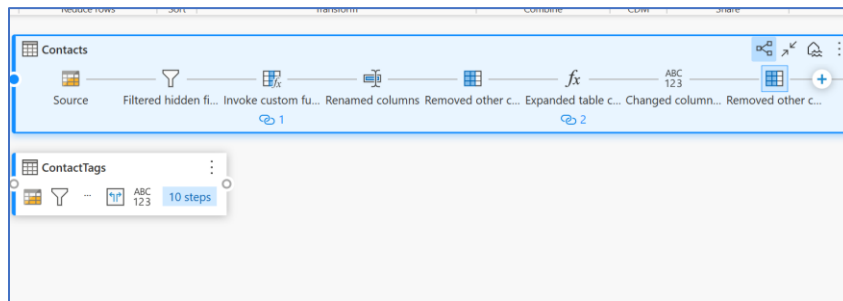
Click on **Diagram view** to see graphically how the 2 queries are structured.



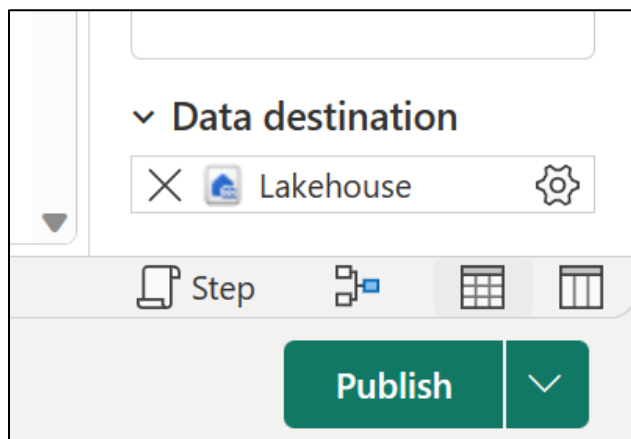
The **Contacts** query display the Lakehouse icon displaying where and how the data will be loaded.







Expand the **Contacts** query to display the sequence of transformation step.



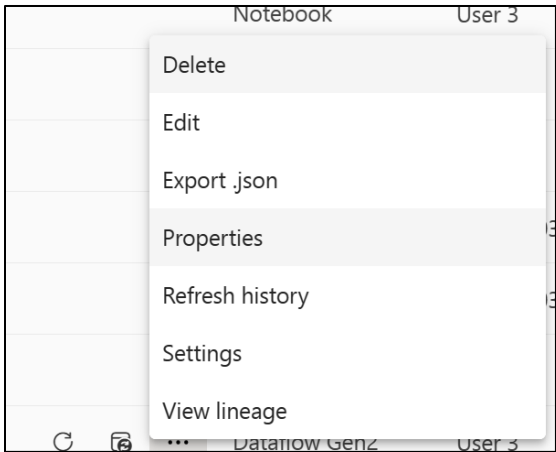
Click on **Publish** to save the Dataflow definition and start its first execution.



Wait until the end of the first Dataflow execution, by monitoring the column **Refreshed**.

	Contoso	 ...	Data pipeline	User 3	—
	Dataflow 1		Dataflow Gen2	User 3	1/15/24, 9:45:44 PM 

When the dataflow has been successfully executed, click on the Dataflow and select **Properties**.



Rename the dataflow as **ContactsLoading**.

Dataflow 2



* Required

Name

ContactsLoading

Description

Open the Contoso Lakehouse, and open the 2 tables Contacts and ContactTags, to make sure they contain data loaded from the CSV file.

✓ Contoso		age	eyeColor	name	gender	company	email	phone
✓	Tables							
>	 Contacts ...	1	20	brown	Landry Bro...	male	DOGTOWN	landrybrow... +1 (902) 48...
>	 ContactTags	2	20	brown	Munoz Boo...	male	MAZUDA	munozboo... +1 (870) 49...

Task 2: Execute advanced M code

In this task, you will process the same **Contoso.json** file using an advanced M (mashup) code snippet which:

- Directly extracts JSON structure based on a schema definition provided in the code
- Combines 2 instances of the same table object (with a single read of the json file)
- Expose detailed information about the main contact and the friend

```
let
//Read the JSON file
Source = AzureStorage.DataLake("https://<storage account>.dfs.core.windows.net/labdata/json/Contoso.json"),

JsonDoc = Json.Document(Table.Column(Source,"Content"){0}),

//Display the file content as a list
TableFromList= Table.FromList(
    JsonDoc,
    Splitter.SplitByNothing(),
    null,
    null,
    ExtraValues.Error
),

//Expand the list to get contact informations (Friends are not yet expanded)
ExpandedTable = Table.ExpandRecordColumn(TableFromList, "Column1",
    {"age", "eyeColor", "name", "gender", "company", "friends", "favoriteFruit"},
    {"age", "eyeColor", "name", "gender", "company", "friends", "favoriteFruit"}
),




//Remove duplicate contact (this table will be used 2 times : for contact information and friends informations)
TableRemoveDuplicates = Table.Distinct(ExpandedTable,{"name"}),


//Expand friends for each contact as a list
TableWithExpandedFriendsColumn = Table.ExpandListColumn(TableRemoveDuplicates, "friends"),

//Expand friends for each contact as records
PersonTable = Table.ExpandRecordColumn(TableWithExpandedFriendsColumn, "friends", {"name"}, {"friend"}),

//Prepare the table of friends by reusing the table "TableRemoveDuplicates"
//Rename columns with the "friend" prefix
FriendTable=Table.RenameColumns(
    //Keep a subset of columns (contact information) from the table "TableRemoveDuplicates"
    Table.SelectColumns(
        TableRemoveDuplicates,
```

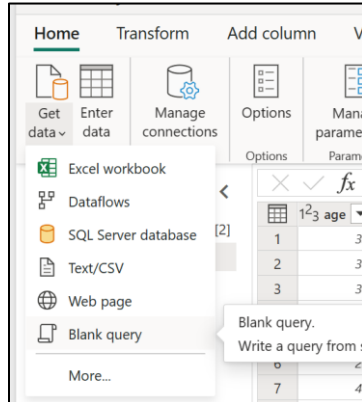
Edit the dataflow named **ContactsLoading** (right click and select **Edit**).

	Name	Type
	ContactsLoading	Dataflow Gen2
	Contoso	



Delete
Edit

Click on **GetData** and select **Blank query**



Paste the M code from the M_snippet.txt file, and replace <storage account> with the account name already used in the previous task.

Take the time to analyze the M code which is document with comments.

```
//Read the JSON file
Source = AzureStorage.DataLake("https://<storage account>.dfs.core.windows.net/labdata/json/Contoso.json"),

JsonDoc = Json.Document(Table.Column(Source, "Content"){0}),

//Display the file content as a list
```

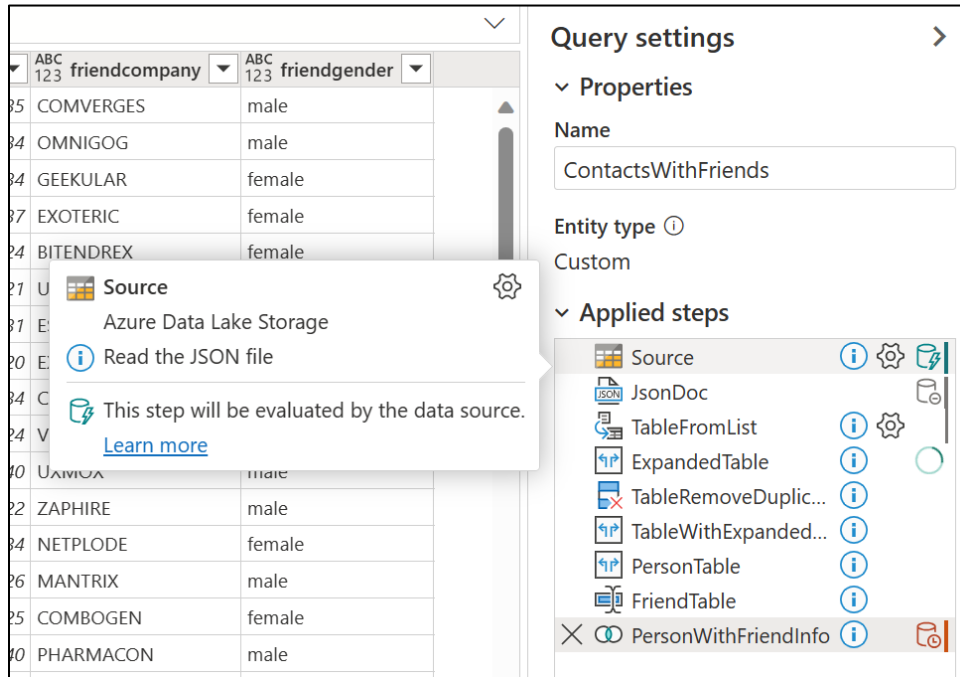
After few seconds, the query result should appear in the data grid.

Rename the query as **ContactsWithFriends**.

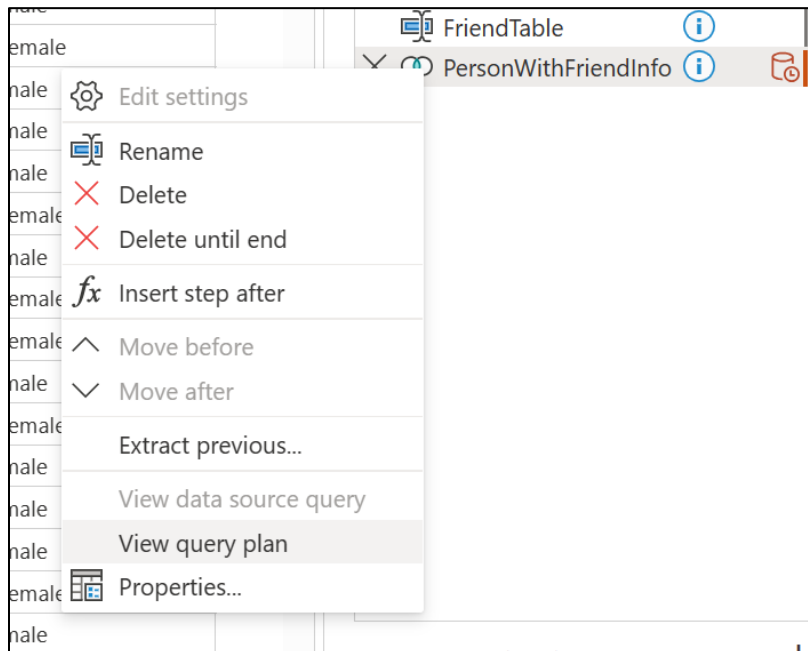
The screenshot shows the Microsoft Fabric Data Factory interface. On the left, the 'Queries [7]' pane lists 'ContactsWithFriends' as the selected query. The main area displays a table with 21 rows and 12 columns. The columns are: 'id', 'age', 'eyeColor', 'name', 'gender', 'company', 'friend', 'favoriteFruit', 'friendname', 'friendage', 'friendcompany', and 'friendgender'. The table contains data for various individuals and their friends. On the right, the 'Query settings' pane is open, showing the 'Properties' tab with the query name 'ContactsWithFriends' and the 'Entity type' set to 'Custom'. The 'Applied steps' pane shows a sequence of steps: 'Source', 'JsonDoc', 'TableFromList', 'ExpandedTable', 'TableRemoveDuplic...', 'TableWithExpanded...', 'PersonTable', 'FriendTable', and 'PersonWithFriendInfo'.

id	age	eyeColor	name	gender	company	friend	favoriteFruit	friendname	friendage	friendcompany	friendgender
1	20	green	Rosanne Carrillo	female	KENGEN	Buath Dominguez	apple	Buath Dominguez	35	CONVERGES	male
2	21	green	Hayes Bender	male	XXANH	Cooley House	apple	Cooley House	34	OMNIGOG	male
3	26	blue	Cecilia Bennett	female	CORPORANA	Nola Tucker	banana	Nola Tucker	34	GEKULAR	female
4	21	blue	Christensen Ellison	male	STREZZO	Lidia Nichols	banana	Lidia Nichols	37	EXOTIBIC	female
5	39	green	Sheena McLean	female	ESCHOR	Vera Black	banana	Vera Black	24	BITENDREX	female
6	25	green	Petra French	female	GEKKO	Letha Wilkinson	apple	Letha Wilkinson	21	UNWORLD	female
7	39	blue	Peters Barnes	male	SYNTAC	Glenda Curry	strawberry	Glenda Curry	31	EXCINTA	female
8	39	blue	Lester Holmes	male	UTARIAN	Hallie Rutledge	apple	Hallie Rutledge	20	EXOZENT	female
9	39	blue	Hawkins Ochoa	male	IMANT	Cohen Schneider	apple	Cohen Schneider	34	COMCLUBINE	male
10	23	brown	Jackson Miller	male	ZIGGLES	Irene Donaldson	banana	Irene Donaldson	24	VETRON	female
11	34	green	Madeline Ortiz	female	MAXEMIA	Jennings Farrell	banana	Jennings Farrell	40	URMOX	male
12	34	green	Rosetta Gay	female	EXOSIS	Holden Perez	apple	Holden Perez	22	ZAPHIRE	male
13	36	green	Jessica McGowan	female	ZAJ	Millicent Briggs	strawberry	Millicent Briggs	34	NETPLODE	female
14	31	blue	Walter Clark	male	RUGSTAR	Briggs Moran	apple	Briggs Moran	26	MANTRIX	male
15	39	green	Wilma Forbes	female	BLEEKO	Bethany Fulton	apple	Bethany Fulton	25	COMBOGEN	female
16	20	brown	Munoz Boone	male	MAZUDA	Sawyer Manning	banana	Sawyer Manning	40	PHARMACON	male
17	22	brown	Aurora Jones	female	MULTRON	Vang Rense	banana	Vang Rense	36	BEDLAM	male
18	23	blue	Murray Hickman	male	VISUALIX	Perry Ruiz	strawberry	Perry Ruiz	36	ZIPTOPE	male
19	20	green	Michelle Moon	female	GENEKOM	Lara Perkins	banana	Lara Perkins	38	CIPROMOX	female
20	24	green	Loretta Hill	female	FLOTONIC	Mcmillan Mitchell	apple	Mcmillan Mitchell	40	ZANITY	male
21	20	green	Solomon Townsend	male	LIQUICOM	Carry Goodwin	banana	Carry Goodwin	35	ROUGHES	female

In the Query settings panel, select the first step named **Source**, and look at the displayed properties : the lightning indicates that the activity is evaluated by the data source.

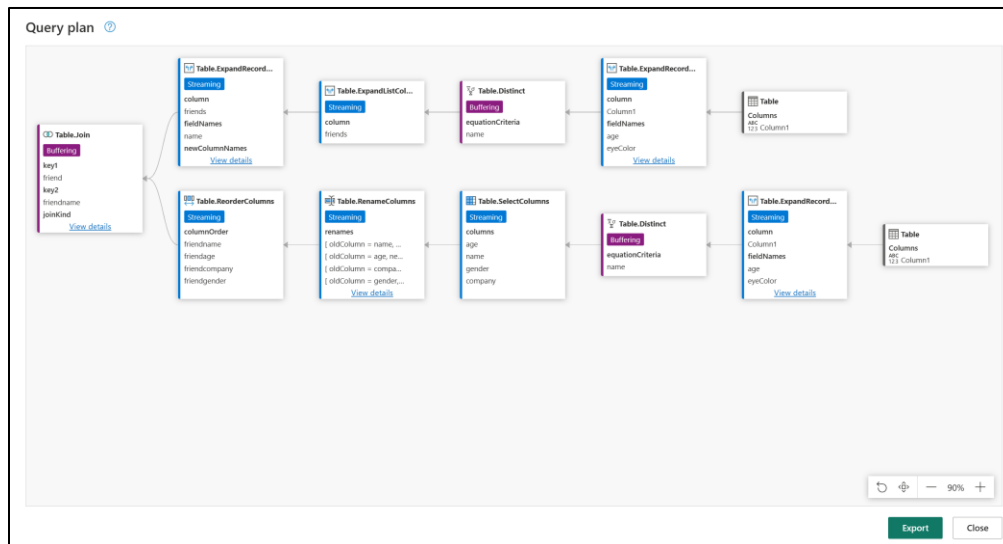


Select the last step **PersonWithFriendInfo**, right click and select **View query plan**.

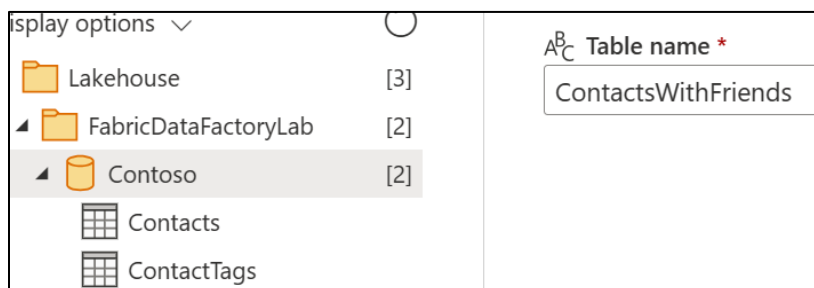


Review the Query Plan displayed with:

- The buffered tasks
- The streamed tasks



Close the query plan, and load the query to the Lakehouse using **ContactsWithFriends** as the table name.



Configure the source type as defined hereafter:

Choose destination settings

Update method

Existing data

New data

Append

Replace

Column mapping

Source	Source type	Destination	Destination type
<input checked="" type="checkbox"/> age	Whole number	age	Whole number
<input checked="" type="checkbox"/> eyeColor	Text	eyeColor	Text
<input checked="" type="checkbox"/> name	Text	name	Text
<input checked="" type="checkbox"/> gender	Text	gender	Text
<input checked="" type="checkbox"/> company	Text	company	Text
<input checked="" type="checkbox"/> friend	Text	friend	Text
<input checked="" type="checkbox"/> favoriteFruit	Text	favoriteFruit	Text
<input checked="" type="checkbox"/> friendname	Text	friendname	Text
<input checked="" type="checkbox"/> friendage	Text	friendage	Text
<input checked="" type="checkbox"/> friendcompany	Text	friendcompany	Text
<input checked="" type="checkbox"/> friendgender	Text	friendgender	Text

Back

Cancel

Save settings

Click on **Publish**

▼ Data destination

×

Lakehouse

⚙

Step

Publish

▼

Wait until the end of the dataflow execution.

	Name	Type	Owner	Refreshed
	ContactsLoading *	Dataflow Gen2	User1	13/10/23 13:33:52

Control the table content in the Lakehouse.

▼ Contoso	age	eyeColor	name	gender	company	friend	favoriteFruit	friendname	friendage	friendcom...	friendgender
▼ Tables	1	21	green	Hayes Bend...	male	XIXAN	Cooley Hou...	Cooley Hou...	34	OMNIGOG	male
> Contacts	2	36	green	Zimmerma...	male	FIREWAX	Gay Fry	Gay Fry	34	STOCKPOST	male
> ContactsWithFrie...	3	20	green	Boyer Mich...	male	NETERIA	Burnett Cas...	Burnett Cas...	24	QUILM	male
> ContactTags											

Task 3: Extract data from OData

In this task, you will extract data from an OData API endpoint:

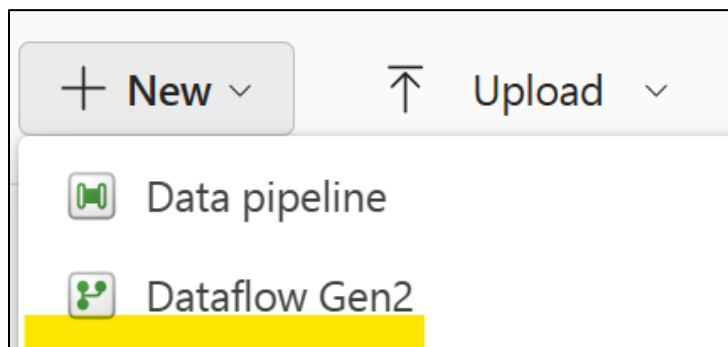
<https://services.odata.org/V4/Northwind/Northwind.svc/>

<https://services.odata.org/V3/Northwind/Northwind.svc/>

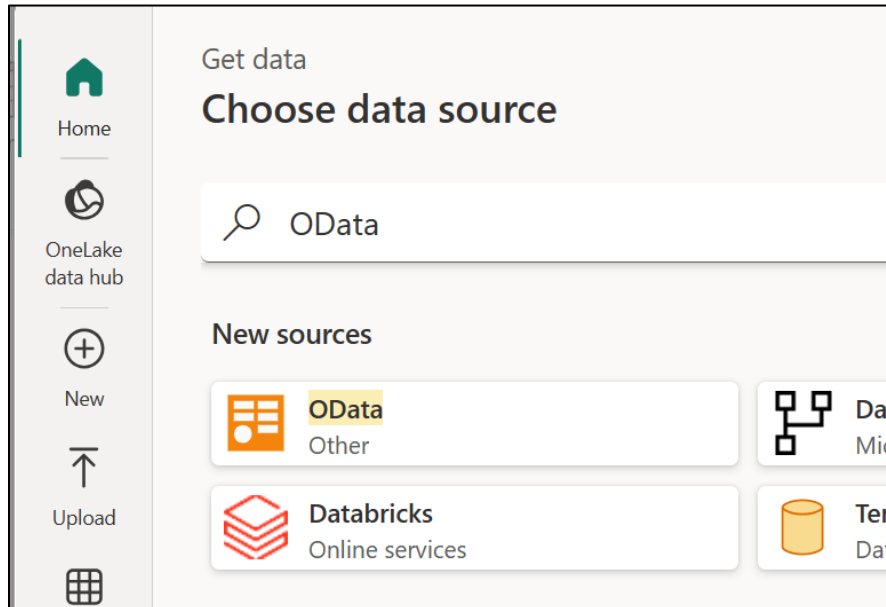
```
<?xml version='1.0' encoding='utf-8'>
<service xmlns="http://www.w3.org/2007/app" xmlns:atom="http://www.w3.org/2005/Atom" xmlns:m="http://docs.oasis-open.org/odata/ns/metadata" xml:base="https://services.odata.org/V4/Northwind/Northwind.svc/"
m:context="https://services.odata.org/V4/Northwind/Northwind.svc/$metadata">
  <workspace>
    <atom:title type="text">Default</atom:title>
    <collection href="Categories">
      <atom:title type="text">Categories</atom:title>
    </collection>
    <collection href="CustomerDemographics">
      <atom:title type="text">CustomerDemographics</atom:title>
    </collection>
    <collection href="Customers">
      <atom:title type="text">Customers</atom:title>
    </collection>
    <collection href="Employees">
      <atom:title type="text">Employees</atom:title>
    </collection>
    <collection href="Order_Details">
      <atom:title type="text">Order_Details</atom:title>
    </collection>
    <collection href="Orders">
      <atom:title type="text">Orders</atom:title>
    </collection>
    <collection href="Products">
      <atom:title type="text">Products</atom:title>
    </collection>
    <collection href="Regions">
      <atom:title type="text">Regions</atom:title>
    </collection>
    <collection href="Shippers">
      <atom:title type="text">Shippers</atom:title>
    </collection>
    <collection href="Suppliers">
      <atom:title type="text">Suppliers</atom:title>
    </collection>
    <collection href="Territories">
      <atom:title type="text">Territories</atom:title>
    </collection>
    <collection href="Alphabetical_list_of_products">
      <atom:title type="text">Alphabetical_list_of_products</atom:title>
    </collection>
    <collection href="Category_Sales_for_1997">
      <atom:title type="text">Category_Sales_for_1997</atom:title>
    </collection>
    <collection href="Current_Product_Lists">

```

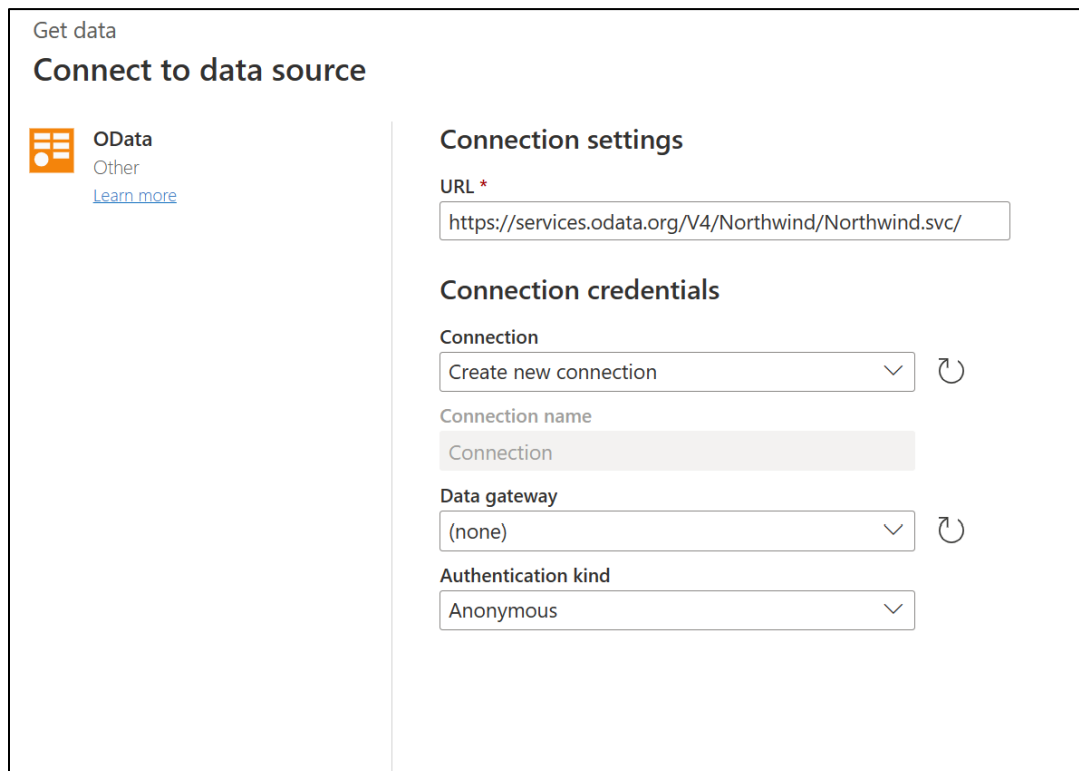
Create a new dataflow Gen2.



Choose the **OData** connector.



Define the URL and use the **Anonymous** authentication kind.



Select the table **Invoices** and **Products** and click on **Create**.

Get data

Choose data

Display options

Search

Alphabetical_list_of_products [26]

Categories

Category_Sales_for_1997

Current_Product_Lists

Customer_and_Suppliers_b...

CustomerDemographics

Customers

Employees

Invoices

Order_Details

Order_Details_Extendeds

Order_Subtotals

Orders

Orders_Queries

Product_Sales_for_1997

Products

Products_Above_Average...

Products_by_Categories

Regions

Sales_by_Categories

Sales_Totals_by_Amounts

Shippers

Select related tables

Back

Cancel

Create

ProductID	ProductName	SupplierID	CategoryID	QuantityPerUnit	UnitPrice	UnitsInStock	UnitsOnOrder	ReorderLevel	Discontinued	Category	Order_Details	Rel
1	Chai	1	1	10 boxes x 20 bags	18	39	0	10	FALSE	[Record]	[Table]	[Rel]
2	Chang	1	1	24 - 12 oz bottles	19	17	40	25	FALSE	[Record]	[Table]	[Rel]
3	Aniseed Syrup	1	2	12 - 550 ml bottles	10	13	70	25	FALSE	[Record]	[Table]	[Rel]
4	Chef Anton's Cajun Seasoning	2	2	48 - 6 oz jars	22	53	0	0	FALSE	[Record]	[Table]	[Rel]
5	Chef Anton's Gumbo Mix	2	2	36 boxes	21.25	0	0	0	TRUE	[Record]	[Table]	[Rel]
6	Grandma's Boysenberry Spread	3	2	12 - 8 oz jars	25	120	0	25	FALSE	[Record]	[Table]	[Rel]
7	Uncle Bob's Organic Dried Pears	3	7	12 - 1 lb pkgs.	30	15	0	10	FALSE	[Record]	[Table]	[Rel]
8	Northwoods Cranberry Sauce	2	2	12 - 12 oz jars	40	6	0	0	FALSE	[Record]	[Table]	[Rel]
9	Mishi Kobe Niku	4	6	18 - 500 g pkgs.	97	29	0	0	TRUE	[Record]	[Table]	[Rel]
10	Rua	4	8	12 - 200 ml jars	31	31	0	0	FALSE	[Record]	[Table]	[Rel]
11	Queso Cabrales	5	4	1 kg pkg.	21	22	30	30	FALSE	[Record]	[Table]	[Rel]
12	Queso Manchego La Pastora	5	4	10 - 500 g pkgs.	38	86	0	0	FALSE	[Record]	[Table]	[Rel]
13	Konbu	6	8	2 kg box	6	24	0	5	FALSE	[Record]	[Table]	[Rel]
14	Tofu	6	7	40 - 100 g pkgs.	23.25	35	0	0	FALSE	[Record]	[Table]	[Rel]
15	Garden of Eatin'	6	2	24 - 250 ml bottles	15.5	39	0	5	FALSE	[Record]	[Table]	[Rel]
16	Pavlova	7	3	32 - 500 g boxes	17.45	29	0	10	FALSE	[Record]	[Table]	[Rel]
17	Alice Mutton	7	6	20 - 1 kg tins	39	0	0	0	TRUE	[Record]	[Table]	[Rel]
18	Carnarvon Tigers	7	8	16 kg pkg.	62.5	42	0	0	FALSE	[Record]	[Table]	[Rel]
19	Teatime Chocolate Biscuits	8	3	10 boxes x 12 pieces	9.2	25	0	5	FALSE	[Record]	[Table]	[Rel]
20	Sir Rodney's Marmalade	8	3	30 gift boxes	81	40	0	0	FALSE	[Record]	[Table]	[Rel]
21	Sir Rodney's Scones	8	3	24 pkgs. x 4 pieces	10	3	40	5	FALSE	[Record]	[Table]	[Rel]
22	Gustaf's Knäckebröd	9	5	24 - 500 g pkgs.	21	104	0	25	FALSE	[Record]	[Table]	[Rel]
23	Turnip	9	5	12 - 250 g pkgs.	9	61	0	25	FALSE	[Record]	[Table]	[Rel]
24	Guaraná Fantástica	10	1	12 - 355 ml cans	4.5	26	0	0	TRUE	[Record]	[Table]	[Rel]
25	NuNuCa Nuß-Nougat-Creme	11	3	20 - 450 g glasses	14	76	0	30	FALSE	[Record]	[Table]	[Rel]
26	Gumbärchen	11	3	100 - 250 g bags	31.23	15	0	0	FALSE	[Record]	[Table]	[Rel]
27	Schoggi Schokolade	11	3	100 - 100 g pieces	43.9	49	0	30	FALSE	[Record]	[Table]	[Rel]
28	Rössle Sauerkraut	12	7	25 - 825 g cans	45.6	26	0	0	TRUE	[Record]	[Table]	[Rel]

Select the **Invoices** query, expand the **Data** table by selecting the **Data** column.

ABC 123	Name	ABC 123	Data	ABC 123	Signature
1	Alphabetical_list_of_products	[Table]	table		
2	Categories	[Table]	table		
3	Category_Sales_for_1997	[Table]	table		
4	Current_Product_Lists	[Table]	table		
5	Customer_and_Suppliers_by_Cities	[Table]	table		
6	CustomerDemographics	[Table]	table		
7	Customers	[Table]	table		
8	Employees	[Table]	table		
9	Invoices	[Table]	table		
10	Order_Details	[Table]	table		
11	Order_Details_Extendeds	[Table]	table		
12	Order_Subtotals	[Table]	table		
13	Orders	[Table]	table		

ProductID	ProductName	SupplierID	CategoryID	QuantityPerUnit	UnitPrice	UnitsInStock	UnitsOnOrder	ReorderLevel	Discontinued	CategoryName
1	Chai	1	1	10 boxes x 20 bags	18	39	0	10	FALSE	Beverages
2	Chang	1	1	24 - 12 oz bottles	19	17	40	25	FALSE	Beverages
3	Sasquatch Ale	16	1	24 - 12 oz bottles	14	111	0	15	FALSE	Beverages
4	Steeleye Stout	16	1	24 - 12 oz bottles	18	20	0	15	FALSE	Beverages
5	Côte de Blaye	18	1	12 - 75 cl bottles	263.5	17	0	15	FALSE	Beverages
6	Chartreuse verte	18	1	750 cc per bottle	18	69	0	5	FALSE	Beverages
7	Ipon Coffee	20	1	16 - 500 g tins	46	17	10	25	FALSE	Beverages
8	Laughing Lumberjack Lager	16	1	24 - 12 oz bottles	14	52	0	10	FALSE	Beverages
9	Outback Lager	7	1	24 - 355 ml bottles	15	15	10	30	FALSE	Beverages
10	Rhönbräu Klosterbier	12	1	24 - 0.5 l bottles	7.75	125	0	25	FALSE	Beverages

Select the **Products** query, expand the **Category** table by selecting the **CategoryName** column.

The screenshot shows the MS Fabric Data Factory interface. On the left, the 'Queries [2]' pane lists 'Invoices' and 'Products'. The 'Products' query is selected, and its data is displayed in a table. The table has columns: SupplierID, CategoryID, QuantityPerUnit, UnitPrice, UnitsInStock, UnitsOnOrder, ReorderLevel, Discontinued, and Category. The 'Category' column is expanded, showing a list of categories. A search dialog is open, showing the 'CategoryName' column selected. The dialog also has options for 'CategoryID', 'Description', 'Picture', and 'Products'. The 'Use original column name as prefix' checkbox is unchecked.

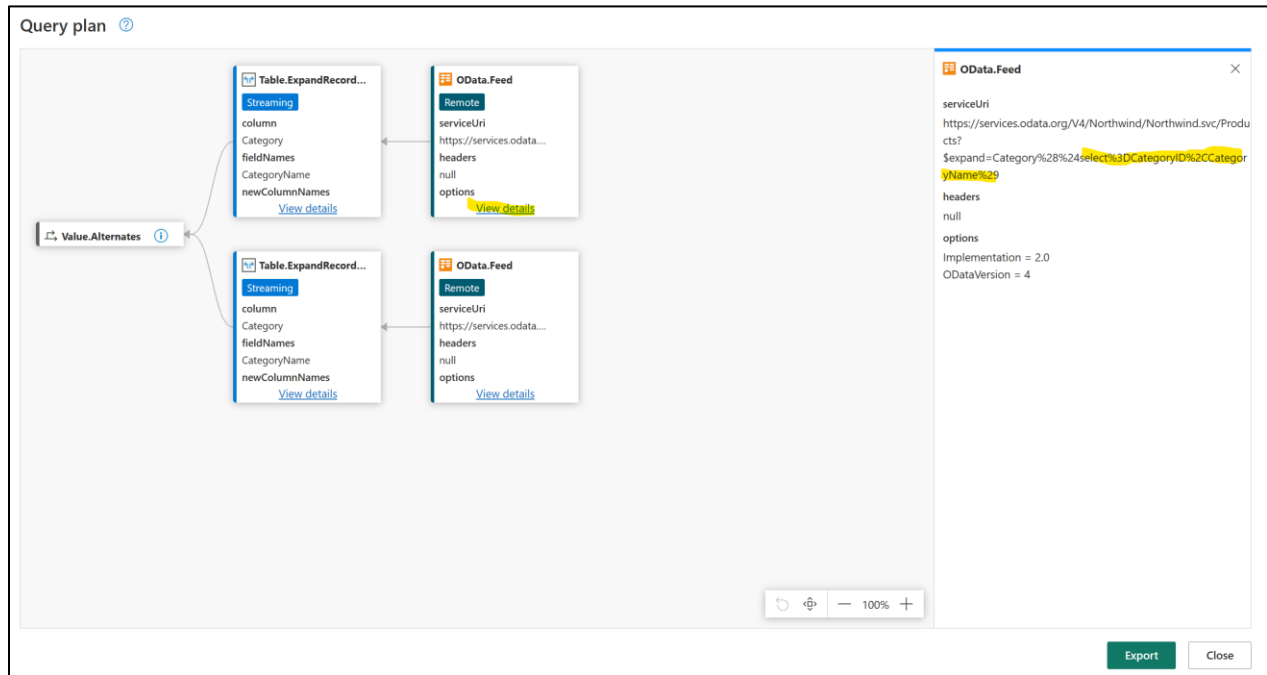
SupplierID	CategoryID	QuantityPerUnit	UnitPrice	UnitsInStock	UnitsOnOrder	ReorderLevel	Discontinued	Category
1	1	10 boxes x 20 bags	18	39				[Record]
2	1	24 - 12 oz bottles	19	17				[Record]
3	1	2 12 - 550 ml bottles	10	13				[Record]
4	2	2 48 - 6 oz jars	22	53				[Record]
5	2	2 36 boxes	21.35	0				[Record]
6	3	2 12 - 8 oz jars	25	120				[Record]
7	3	7 12 - 1 lb pkgs.	30	15				[Record]
8	3	2 12 - 12 oz jars	40	6				[Record]
9	4	6 18 - 500 g pkgs.	97	29				[Record]
10	4	8 12 - 200 ml jars	31	31				[Record]
11	5	4 1 kg pkg.	21	22				[Record]
12	5	4 10 - 500 g pkgs.	38	86				[Record]
13	6	8 2 kg box	6	24				[Record]
14	6	7 40 - 100 g pkgs.	23.25	35				[Record]
15	6	2 24 - 250 ml bottles	15.5	39				[Record]
16	7	3 32 - 500 g boxes	17.45	29	0	10	FALSE	[Record]
17	7	6 20 - 1 kg tins	39	0	0	0	TRUE	[Record]
18	7	8 16 kg pkg.	62.5	42	0	0	FALSE	[Record]
19	8	3 10 boxes x 12 pieces	9.2	25	0	5	FALSE	[Record]
20	8	3 30 gift boxes	81	40	0	0	FALSE	[Record]

Select the **Products** query, expand the **Products** table by selecting the **ProductID** column.

The screenshot shows the MS Fabric Data Factory interface. On the left, the 'Queries [2]' pane lists 'Invoices' and 'Products'. The 'Products' query is selected, and its data is displayed in a table. The table has columns: address, City, Region, PostalCode, Country, Phone, Fax, and HomePage. The 'Products' column is expanded, showing a list of products. A search dialog is open, showing the 'ProductID' column selected. The dialog also has options for 'ProductName', 'SupplierID', 'CategoryID', 'QuantityPerUnit', and 'UnitPrice'. The 'Use original column name as prefix' checkbox is unchecked.

address	City	Region	PostalCode	Country	Phone	Fax	HomePage
1 bert St.	London	null	EC1 4SD	UK	(171) 555-2222	null	
2 bert St.	London	null	EC1 4SD	UK	(171) 555-2222	null	
3 bert St.	London	null	EC1 4SD	UK	(171) 555-2222	null	
4 ox 78934	New Orleans	LA	70117	USA	(100) 555-4822	null	#CAJUN.HTM#
5 ox 78934	New Orleans	LA	70117	USA	(100) 555-4822	null	#CAJUN.HTM#
6 xford Rd.	Ann Arbor	MI	48104	USA	(313) 555-5735	(313) 555-3349	
7 xford Rd.	Ann Arbor	MI	48104	USA	(313) 555-5735	(313) 555-3349	
8 xford Rd.	Ann Arbor	MI	48104	USA	(313) 555-5735	(313) 555-3349	
9 ekimai Musashino-shi	Tokyo	null	100	Japan	(03) 3555-5011	null	
10 ekimai Musashino-shi	Tokyo	null	100	Japan	(03) 3555-5011	null	
11 del Rosal 4	Oviedo	Asturias	33007	Spain	(98) 598 76 54	null	
12 del Rosal 4	Oviedo	Asturias	33007	Spain	(98) 598 76 54	null	
13 tsuko Chuo-ku	Osaka	null	545	Japan	(06) 431-7877	null	Mayumi's (on the World Wide Web)#http://w
14 tsuko Chuo-ku	Osaka	null	545	Japan	(06) 431-7877	null	Mayumi's (on the World Wide Web)#http://w
15 tsuko Chuo-ku	Osaka	null	545	Japan	(06) 431-7877	null	Mayumi's (on the World Wide Web)#http://w
16 se St. Moonie Ponds	Melbourne	Victoria	3058	Australia	(03) 444-2343	(03) 444-6588	

Display the execution plan for the **Products** table which contains 2 possible execution plans, one containing the “Expand” predicate to extract the Category Name from the Products entity.



Select the **Invoices** query, then select the Merge queries from the Home ribbon.

<https://learn.microsoft.com/en-us/power-query/merge-queries-overview>

The screenshot shows the Power Query interface. The 'Home' ribbon is active, and the 'Merge queries' button is highlighted. The 'Queries [2]' list on the left shows 'Invoices' and 'Products'. The data table below shows columns for ShipName, ShipAddress, ShipCity, ShipRegion, ShipPostalCode, ShipCountry, CustomerID, CustomerName, Address, City, and Region. The data is as follows:

	ShipName	ShipAddress	ShipCity	ShipRegion	ShipPostalCode	ShipCountry	CustomerID	CustomerName	Address	City	Region
1	Alfred's Futterkiste	Obere Str. 57	Berlin	null	12209	Germany	ALFKI	Alfreds Futterkiste	Obere Str. 57	Berlin	
2	Alfred's Futterkiste	Obere Str. 57	Berlin	null	12209	Germany	ALFKI	Alfreds Futterkiste	Obere Str. 57	Berlin	
3	Alfred's Futterkiste	Obere Str. 57	Berlin	null	12209	Germany	ALFKI	Alfreds Futterkiste	Obere Str. 57	Berlin	
4	Alfred's Futterkiste	Obere Str. 57	Berlin	null	12209	Germany	ALFKI	Alfreds Futterkiste	Obere Str. 57	Berlin	
5	Alfred's Futterkiste	Obere Str. 57	Berlin	null	12209	Germany	ALFKI	Alfreds Futterkiste	Obere Str. 57	Berlin	
6	Alfred's Futterkiste	Obere Str. 57	Berlin	null	12209	Germany	ALFKI	Alfreds Futterkiste	Obere Str. 57	Berlin	

Upskilling on MS Fabric Data Factory

Power Query

Search (Alt + Q)

Home Transform Add column View Help

Get data Enter data Manage connections Options Manage parameters Refresh Advanced editor Add data destination Choose columns Remove columns Keep rows Remove rows Filter rows Sort Split column Group by Use first row as head Merge queries Map to entity Export template

Queries [2]

Invoices 2 steps

Products 2 steps

ProductID ProductName SupplierID CategoryID QuantityPerUnit UnitPrice UnitsInStock UnitsOnOrder ReorderLevel Discontinued CategoryName

1	Chai	1	1	10 boxes x 20 bags	18	39	0	10	FALSE	Beverages
2	Chang	1	1	24 - 12 oz bottles	19	17	40	25	FALSE	Beverages
3	Sauquatch Ale	16	1	24 - 12 oz bottles	14	111	0	15	FALSE	Beverages
4	Steele Stout	16	1	24 - 12 oz bottles	18	20	0	15	FALSE	Beverages
5	Côte de Blaye	18	1	12 - 75 cl bottles	263.5	17	0	15	FALSE	Beverages
6	Charlotte verte	18	1	750 cc per bottle	18	69	0	5	FALSE	Beverages
7	Ipoh Coffee	20	1	16 - 500 g tins	46	17	10	25	FALSE	Beverages
8	Laughing Lumberjack Lager	16	1	24 - 12 oz bottles	14	52	0	10	FALSE	Beverages
9	Outback Lager	7	1	24 - 355 ml bottles	15	15	10	30	FALSE	Beverages
10	Rhônebräu Klosterbier	12	1	24 - 0.5 l bottles	7.75	125	0	25	FALSE	Beverages
11	Lakkalikööri	23	1	500 ml	18	57	0	20	FALSE	Beverages

Define the **Merge** transformation using the **ProductID** for the **Inner** join and validate.

Merge

Select a table and matching columns to create a merged table.

Invoices

ShippedDate	ShipperName	ProductID	ProductName	QuantityPerUnit
/1997, 12:00:00 AM +00:00	United Package	63	Vegie-spread	
/1997, 12:00:00 AM +00:00	Speedy Express	3	Aniseed Syrup	
/1997, 12:00:00 AM +00:00	Speedy Express	76	Lakkalikööri	
/1998, 12:00:00 AM +00:00	Federal Shipping	59	Raclette Courdavault	

Right table for merge *

Products

ProductID	ProductName	SupplierID	CategoryID	QuantityPerUnit
1	Chai	1	1	10 boxes x 20 bags
2	Chang	1	1	24 - 12 oz bottles
3	Aniseed Syrup	1	2	12 - 550 ml bottles
4	Chef Anton's Cajun Seasoning	2	2	48 - 6 oz jars

Join kind *

Left outer

Right outer

Full outer

Inner

Left anti

Right anti

☐ Use fuzzy matching to perform the merge

> Fuzzy matching options

✓ The selection matches 328 rows from both the tables

OK

Cancel

Configure the Lakehouse as the query destination, using **Invoices** as table name.

Choose destination target

☒ New table ☐ Existing table

Display options ⌵ ↻

Lakehouse [3]

FabricDataFactoryLab [2]

Contoso [3]

Contacts

ContactsWithFriends

ContactTags

ⓘ A new table will be created in Contoso

Table name *

Publish the Dataflow.

Explore the **Invoices** table in the Lakehouse.

Home

Get data ▾ New Power BI dataset Open notebook ▾

A SQL endpoint for SQL querying and a default dataset for reporting were created and will be updated with any tables added to the lakehouse.

Explorer

Contoso

- Tables
 - Contacts
 - ContactsWithFriends
 - ContactTags
 - Invoices

Invoices

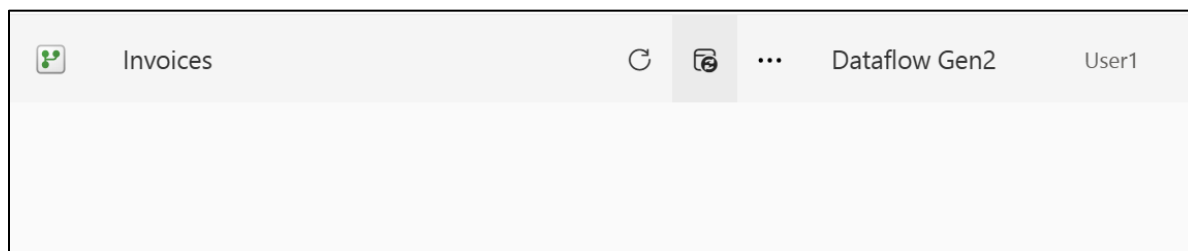
Showing 328 rows

	ShipName	ShipAddress	ShipCity	ShipRegion	ShipPostal...	ShipCountry	CustomerID	CustomerN...	Address	City	Region	PostalCode	Country	Salesperson
1	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Margaret P...
2	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Andrew Full...
3	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Robert King
4	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Laura Calla...
5	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Andrew Full...
6	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Nancy Dav...
7	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Nancy Dav...
8	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Janet Lever...
9	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Margaret P...
10	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Margaret P...
11	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Margaret P...
12	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Laura Calla...
13	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	QUICK	QUICK-Stop	Taucherstra...	Cunewalde		01307	Germany	Andrew Full...

Wait until the end of the dataflow execution and rename it as **Invoices**.

	Name	Type ↑
	Contoso	Data pipeline
	ContactsLoading	Dataflow Gen2
	Invoices	Dataflow Gen2

Open the **Refresh Schedule** option for the **Invoices** dataflow.



Select the **Refresh history** link.

Dataflows Gen2 is in preview


Settings for Invoices

This dataflow has been last modified by [User1@FabricDataFactoryLab.onmicrosoft.com](#)

Last refresh succeeded: Fri Oct 13 2023 23:28:00 GMT+0200 (heure d'été d'Europe centrale)
[Refresh history](#)

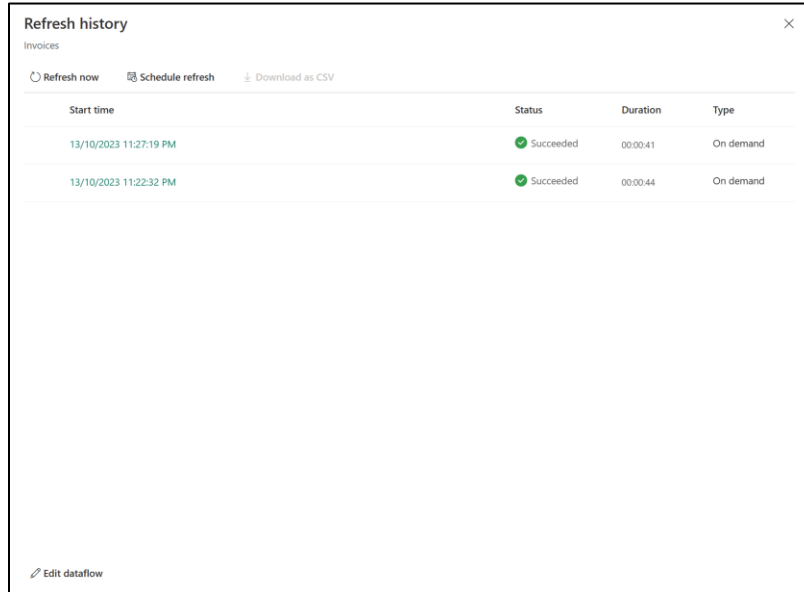
Gateway Connection

Dataflow on-premises gateways are currently editable through the Power Query Online experience. [Learn how to edit](#)

 Data source credentials

4 Refresh

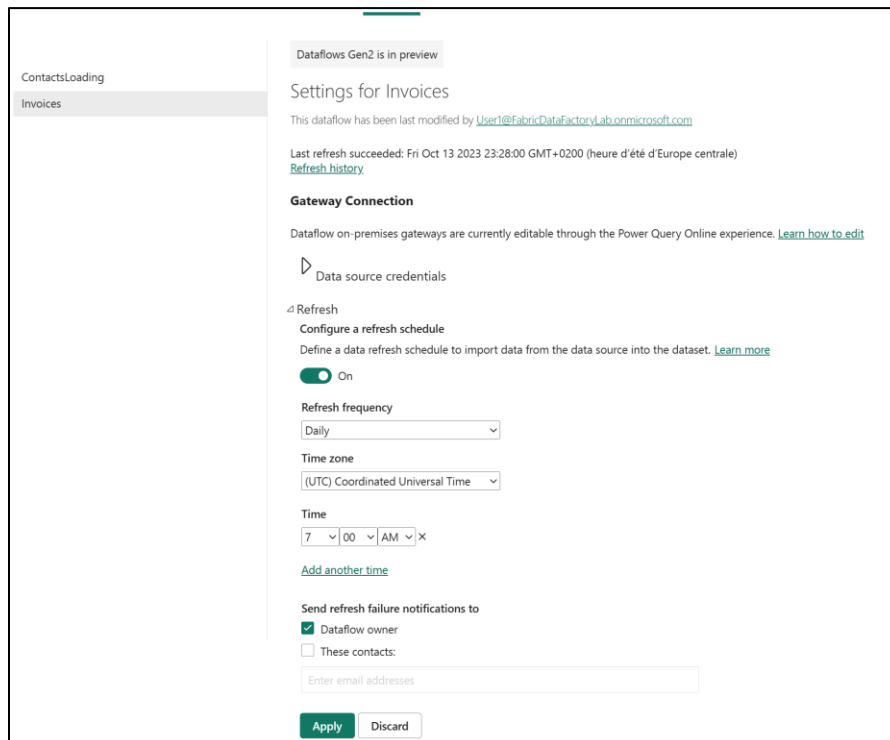
Refresh history gives details about the recent dataflow execution.



The screenshot shows a 'Refresh history' window for the 'Invoices' dataflow. It includes a table with columns for Start time, Status, Duration, and Type. Two successful refreshes are listed, both occurring on 13/10/2023. The interface also features buttons for 'Refresh now', 'Schedule refresh', 'Download as CSV', and 'Edit dataflow'.

Start time	Status	Duration	Type
13/10/2023 11:27:19 PM	Succeeded	00:00:41	On demand
13/10/2023 11:22:32 PM	Succeeded	00:00:44	On demand

Configure and **Daily refresh** and click on **Apply**.



The screenshot displays the 'Settings for Invoices' configuration page. It includes a sidebar with 'ContactsLoading' and 'Invoices' options. The main content area shows the 'Refresh' section, where the refresh schedule is configured. The 'Refresh frequency' is set to 'Daily', and the 'Time zone' is '(UTC) Coordinated Universal Time'. The 'Time' is set to 7:00 AM. There is also a section for 'Send refresh failure notifications to' with a checkbox for 'Dataflow owner' and a text input for 'Enter email addresses'. The 'Apply' button is highlighted in green.

Settings for Invoices

This dataflow has been last modified by [User1@FabricDataFactoryLab.onmicrosoft.com](#)

Last refresh succeeded: Fri Oct 13 2023 23:28:00 GMT+0200 (heure d'été d'Europe centrale)
[Refresh history](#)

Gateway Connection

Dataflow on-premises gateways are currently editable through the Power Query Online experience. [Learn how to edit](#)

▷ Data source credentials

Refresh

Configure a refresh schedule

Define a data refresh schedule to import data from the data source into the dataset. [Learn more](#)

☒ On

Refresh frequency

Daily

Time zone

(UTC) Coordinated Universal Time

Time

7:00 AM

[Add another time](#)

Send refresh failure notifications to

☒ Dataflow owner

☐ These contacts:

Enter email addresses

Apply **Discard**