

# Microsoft Fabric

Chat with your Data in a Day  
Lab #1

Version: September 2025



Document Structure.....	3
Scenario / Problem Statement.....	3
Introduction.....	3
Clean data with dataflows gen2.....	3
Task 1: Setting up a workspace.....	3
Task 2: Creating the dataflow gen2 .....	7
Task 3: Uploading data .....	11
Task 4: transforming data with Copilot .....	13
Step 1: Format table.....	14
Step 2: Clean data .....	14
Step 3: remove redundant data .....	16
Step 4: Update to correct data types .....	17
Appendix .....	18
Answers for what need to be cleaned/transformed in the data.....	18
Troubleshooting.....	19
Failed to create the dataflow.....	19
Failed to get the existing dataflow names.....	22

## Document Structure

The lab includes steps for the user to follow along with associated screenshots that provide visual aid. In each screenshot, sections are highlighted with red boxes to indicate the area(s) user should focus on.

## Scenario / Problem Statement

You are getting data ready to create a Power BI report. The CSV you're working with does not have clean data. You'll use Fabric dataflow Gen2 to clean and transform the data with natural language

### Current Challenges

- Data is not consistent within fields
- Null data
- Duplicate data

## Introduction

Cleaning and transforming data with Microsoft Fabric is easier than ever. Instead of pointing and clicking to do the transformations, you can work with Copilot to get the desired results and clean and transform your data.

By the end of this lab, you will have learned:

- How to access Copilot in a dataflow Gen2
- Examples of natural language questions Copilot supports to transform
- How to use Copilot to clean and transform your data

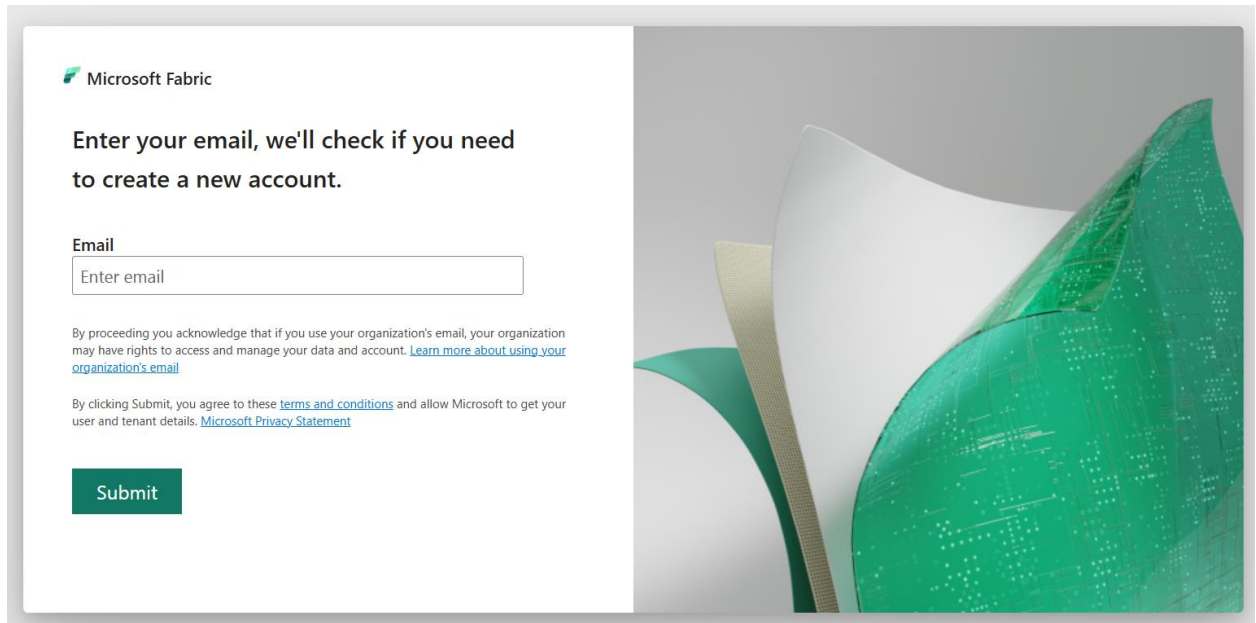
## Clean data with dataflows gen2

### Task 1: Setting up a workspace

In this section we will get your Fabric workspace created so that you can upload your data to clean and transform it.

1. Navigate to <https://app.fabric.microsoft.com>

## 2. Enter the email provided on your printed credentials



The image shows the Microsoft Fabric account creation page. It has a white background with a green Microsoft Fabric logo at the top left. The main heading says "Enter your email, we'll check if you need to create a new account." Below this is an "Email" label and a text input field with the placeholder "Enter email". Under the input field, there is a paragraph of text: "By proceeding you acknowledge that if you use your organization's email, your organization may have rights to access and manage your data and account. [Learn more about using your organization's email](#)". Below that is another paragraph: "By clicking Submit, you agree to these [terms and conditions](#) and allow Microsoft to get your user and tenant details. [Microsoft Privacy Statement](#)". At the bottom left is a green "Submit" button. The right side of the page features a large, abstract green and white graphic.

Microsoft Fabric

Enter your email, we'll check if you need to create a new account.

Email

Enter email

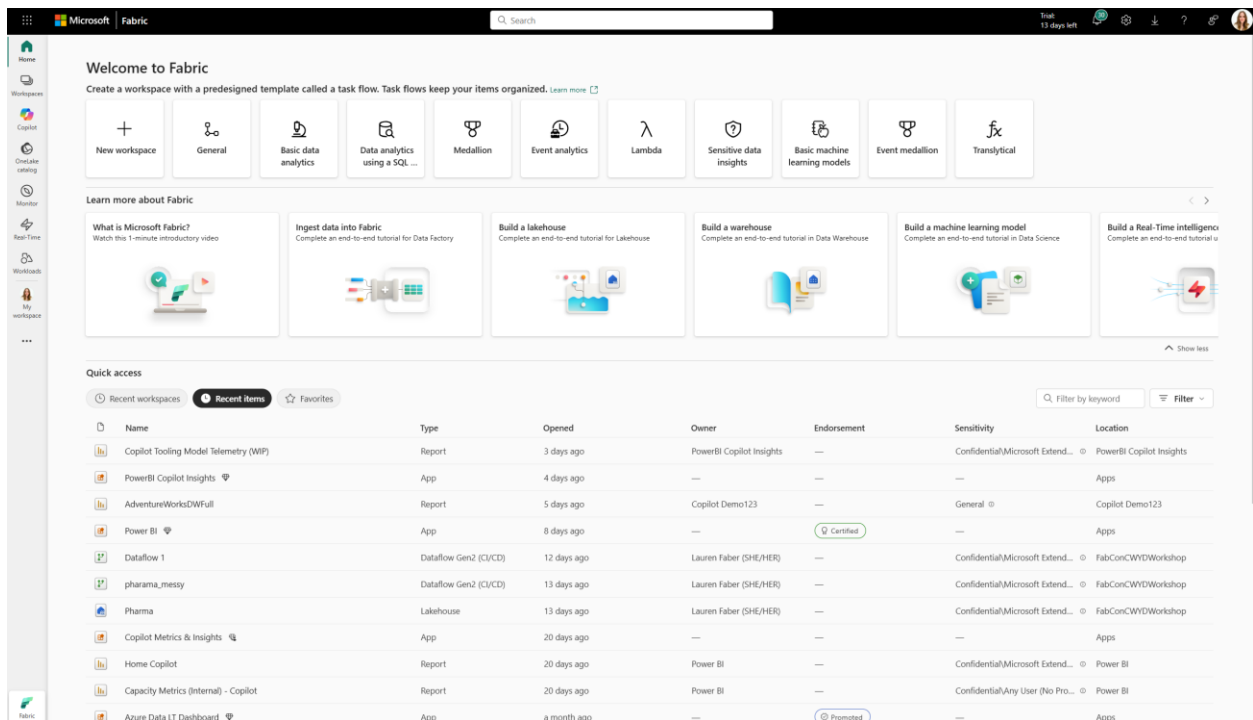
By proceeding you acknowledge that if you use your organization's email, your organization may have rights to access and manage your data and account. [Learn more about using your organization's email](#)

By clicking Submit, you agree to these [terms and conditions](#) and allow Microsoft to get your user and tenant details. [Microsoft Privacy Statement](#)

Submit

## 3. Enter the password provided on your printed credentials

## 4. Land on the Fabric homepage



The image shows the Microsoft Fabric homepage. It has a dark header with the Microsoft logo and "Fabric" text. A search bar is in the center, and a "Trial 11 days left" badge is on the right. The left sidebar contains navigation icons for Home, Workspaces, Copilot, OneLake catalog, Monitor, Real-Time, Workloads, and My workspace. The main content area starts with "Welcome to Fabric" and a description of task flows. Below this is a row of 11 tiles for different capabilities: New workspace, General, Basic data analytics, Data analytics using a SQL..., Medallion, Event analytics, Lambda, Sensitive data insights, Basic machine learning models, Event medallion, and Translytical. Underneath is a "Learn more about Fabric" section with six cards: "What is Microsoft Fabric?", "Ingest data into Fabric", "Build a lakehouse", "Build a warehouse", "Build a machine learning model", and "Build a Real-Time intelligence". The bottom section is "Quick access", which includes tabs for "Recent workspaces", "Recent items", and "Favorites". It contains a table of recent workspaces with columns for Name, Type, Opened, Owner, Endorsement, Sensitivity, and Location.

Microsoft Fabric

Welcome to Fabric

Create a workspace with a predesigned template called a task flow. Task flows keep your items organized. [Learn more](#)

Learn more about Fabric

What is Microsoft Fabric? Watch this 1-minute introductory video

Ingest data into Fabric Complete an end-to-end tutorial for Data Factory

Build a lakehouse Complete an end-to-end tutorial for Lakehouse

Build a warehouse Complete an end-to-end tutorial in Data Warehouse

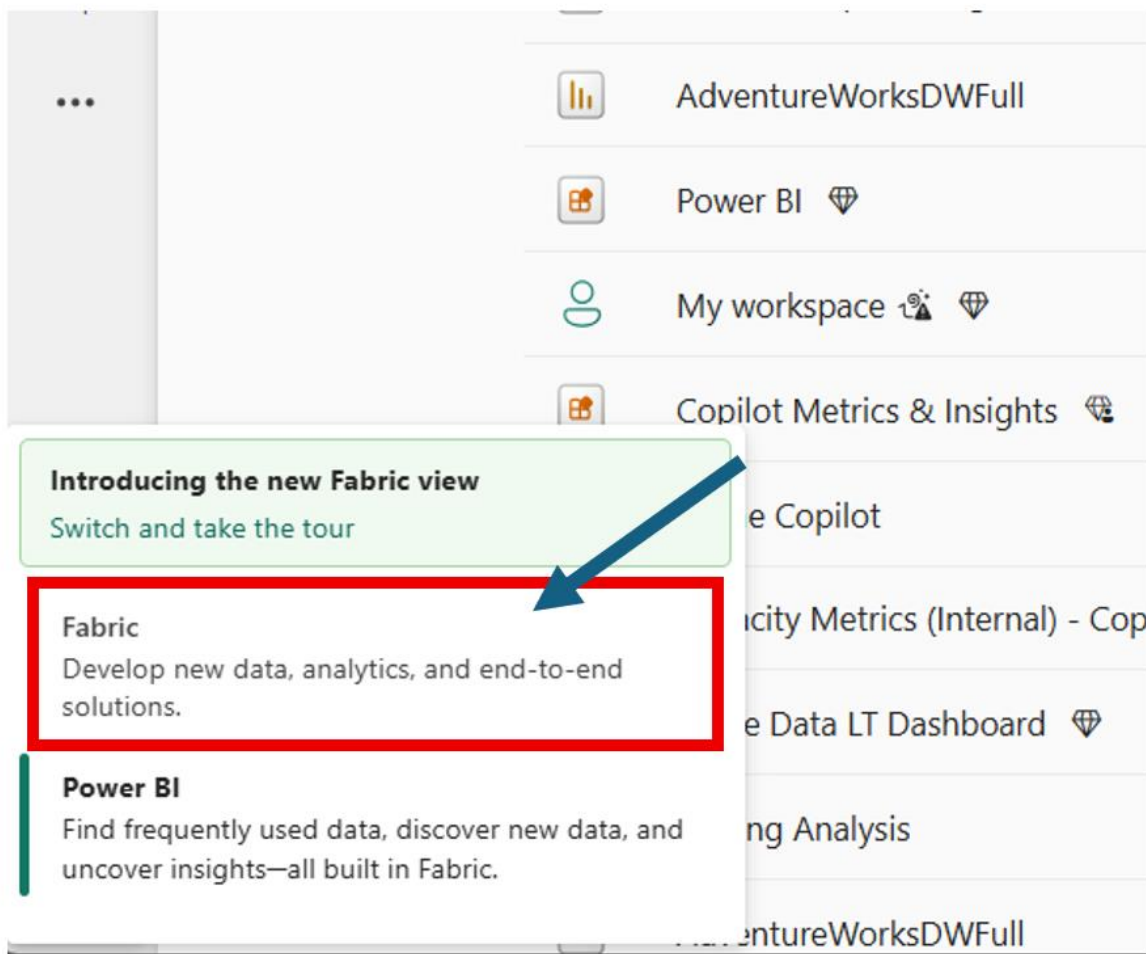
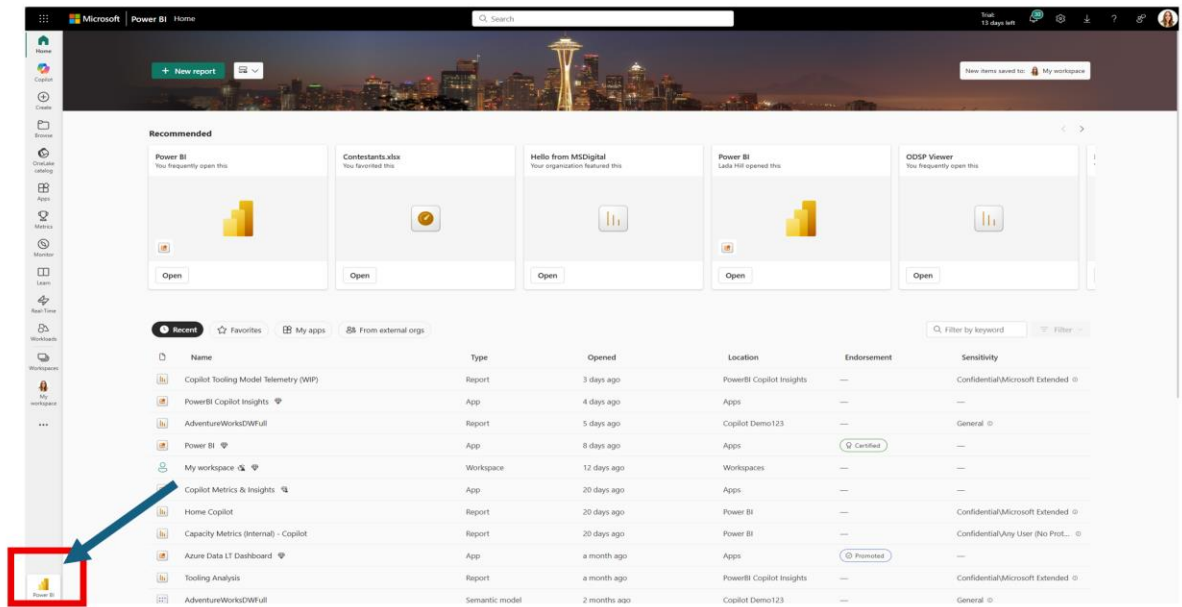
Build a machine learning model Complete an end-to-end tutorial in Data Science

Build a Real-Time intelligence Complete an end-to-end tutorial in

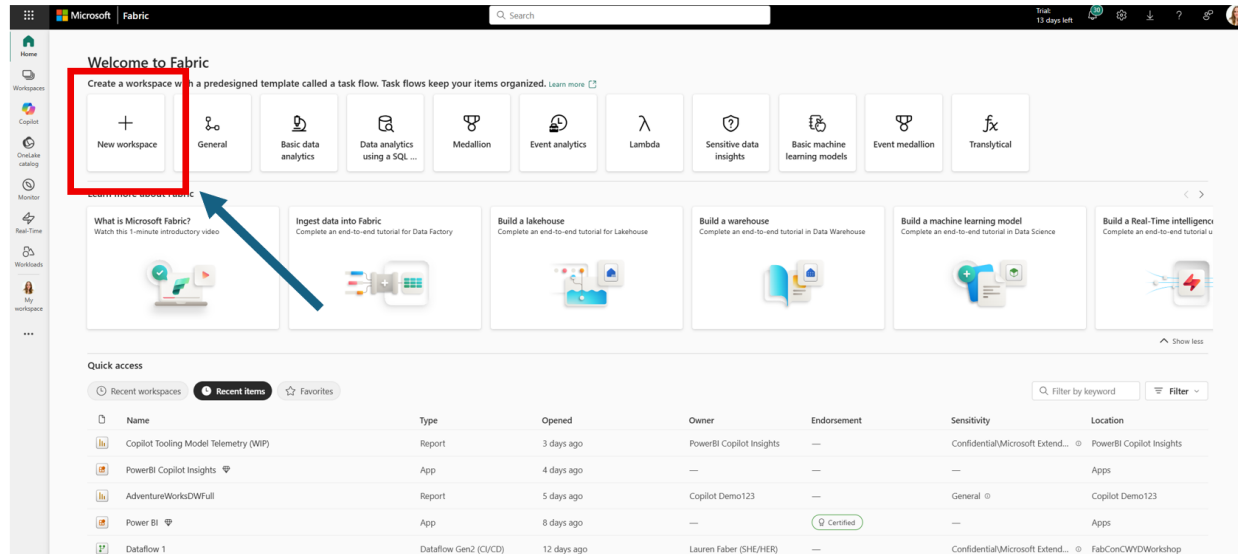
Quick access

Name	Type	Opened	Owner	Endorsement	Sensitivity	Location
Copilot Tooling Model Telemetry (WIP)	Report	3 days ago	PowerBI Copilot Insights	—	Confidential/Microsoft Extend...	PowerBI Copilot Insights
PowerBI Copilot Insights	App	4 days ago	—	—	—	Apps
AdventureWorksDWFull	Report	5 days ago	Copilot Demo123	—	General	Copilot Demo123
Power BI	App	8 days ago	—	Q Certified	—	Apps
Dataflow 1	Dataflow Gen2 (CU/CD)	12 days ago	Lauren Faber (SHE/HER)	—	Confidential/Microsoft Extend...	FabConCWIDWorkshop
pharama_mesny	Dataflow Gen2 (CU/CD)	13 days ago	Lauren Faber (SHE/HER)	—	Confidential/Microsoft Extend...	FabConCWIDWorkshop
Pharma	Lakehouse	13 days ago	Lauren Faber (SHE/HER)	—	Confidential/Microsoft Extend...	FabConCWIDWorkshop
Copilot Metrics & Insights	App	20 days ago	—	—	—	Apps
Home Copilot	Report	20 days ago	Power BI	—	Confidential/Microsoft Extend...	Power BI
Capacity Metrics (Internal) - Copilot	Report	20 days ago	Power BI	—	Confidential/Any User (No Pro...	Power BI
Azure Data LT Dashboard	App	a month ago	—	Promoted	—	Apps

## 5. If you see the Power BI homepage instead of fabric, navigate to the Fabric home page by clicking on the Power BI button in the bottom left corner of the browser. Choose "Fabric"



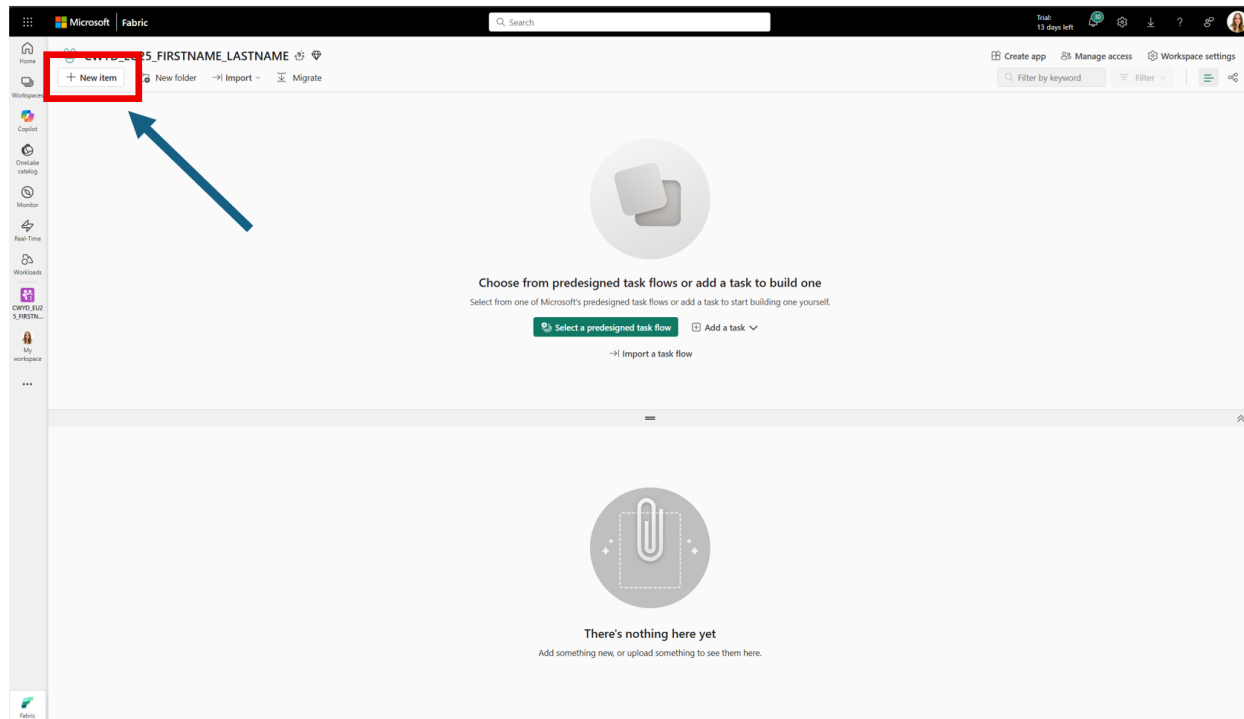
6. Select “New workspace” in the top left.



7. Name the workspace your CWYD\_EU25\_FIRSTNAME\_LASTNAME

a. For example I would name my workspace CWYD\_EU25\_LAUREN\_FABER





## 2. Select Dataflow Gen2

- a. Select the dataflow that has the description of Prep, clean, and transform data
- b. If you cannot scroll to find the item, search for dataflow in the search bar
- c. Select dataflow Gen2. You do not want to select Dataflow Gen1.
- d. It does not matter if you select the Dataflow Gen2 under the “Get data” or “Prepare data” headers.



**New item** ×

☆ Favorites ☒ All items

**Get data**  
Ingest batch and real-time data into a single location within your Fabric workspace.

**Copy job**  
Makes it easy to copy data in Fabric. Includes full copy, incremental copy, and event-based copy modes.

**Data pipeline**  
Ingest data at scale and schedule data workflows.

**Dataflow Gen1**  
Prep, clean, and transform data.

**Dataflow Gen2**  
Prep, clean, and transform data.

**Eventstream**  
Capture, transform, and route real-time event stream to various destinations in desired format with no-code experience.

**Mirrored Azure Cosmos DB (pre...**  
Easily replicate data from an existing source into an analytics-friendly format.

**Mirrored Azure Database for Po...**  
Easily replicate data from an existing source into an analytics-friendly format.

**Mirrored Azure Databricks catalog**  
Explore Unity Catalog Tables

**Mirrored Azure SQL Database**  
Easily replicate data from an existing source into an analytics-friendly format.

**Mirrored Azure SQL Managed In...**  
Easily replicate data from an existing source into an analytics-friendly format.

**Mirrored database**  
Easily replicate data from an existing source into an analytics-friendly format.

**Mirrored Dataverse (preview)**  
Easily access your enterprise data from Dynamics 365 and PowerApps in Dataverse.

**Mirrored Google Big Query (pre...**  
Easily replicate data from an existing source into an analytics-friendly format.

**Mirrored Oracle (preview)**  
Easily replicate data from an existing source into an analytics-friendly format.

**Mirrored Snowflake**  
Easily replicate data from an existing source into an analytics-friendly format.

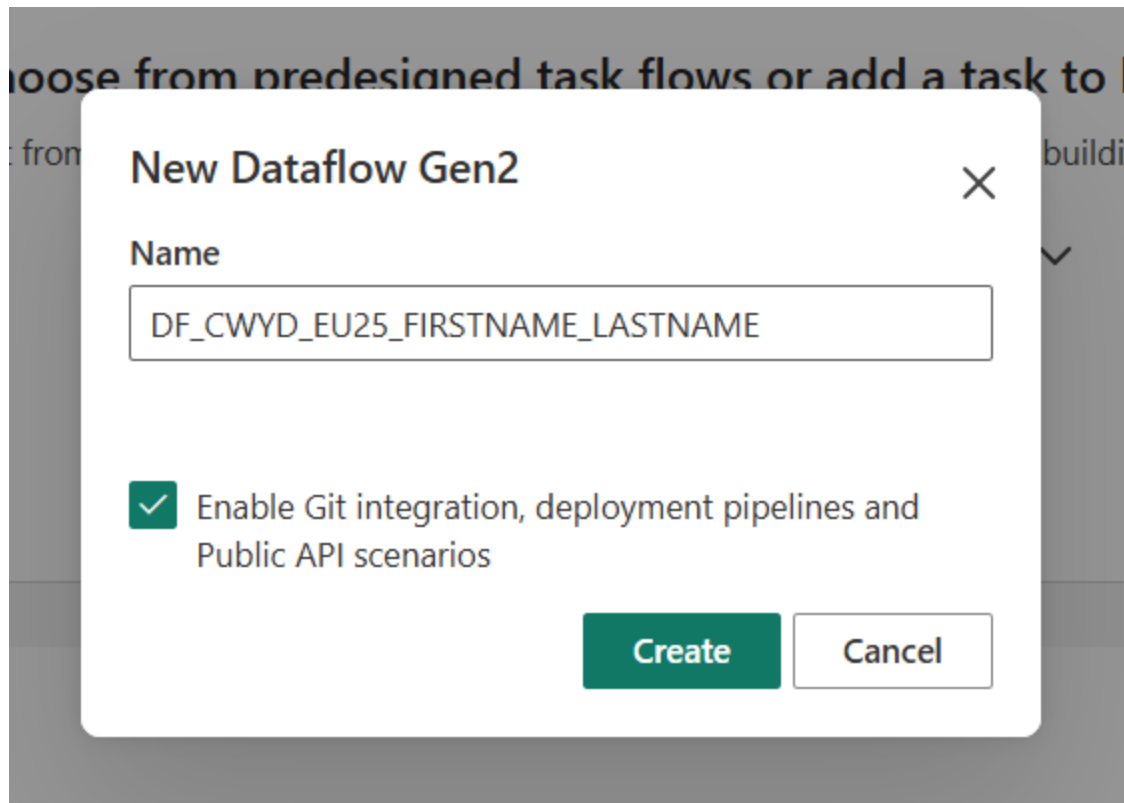
**Mirrored SQL Server (preview)**  
Easily replicate data from an existing source into an analytics-friendly format.

**Notebook**  
Explore, analyze, and visualize data and build ML models. Supports Apache Spark, Python, T-SQL, and more.

**Spark Job Definition**  
Define, schedule, and manage your Apache Spark jobs for big data processing.

3. Name the dataflow

- Name the dataflow as DF\_CWYD\_EU25\_FIRSTNAME\_LASTNAME
- For example I would name my dataflow DF\_CWYD\_EU25\_LAUREN\_FABER
- Name the dataflow

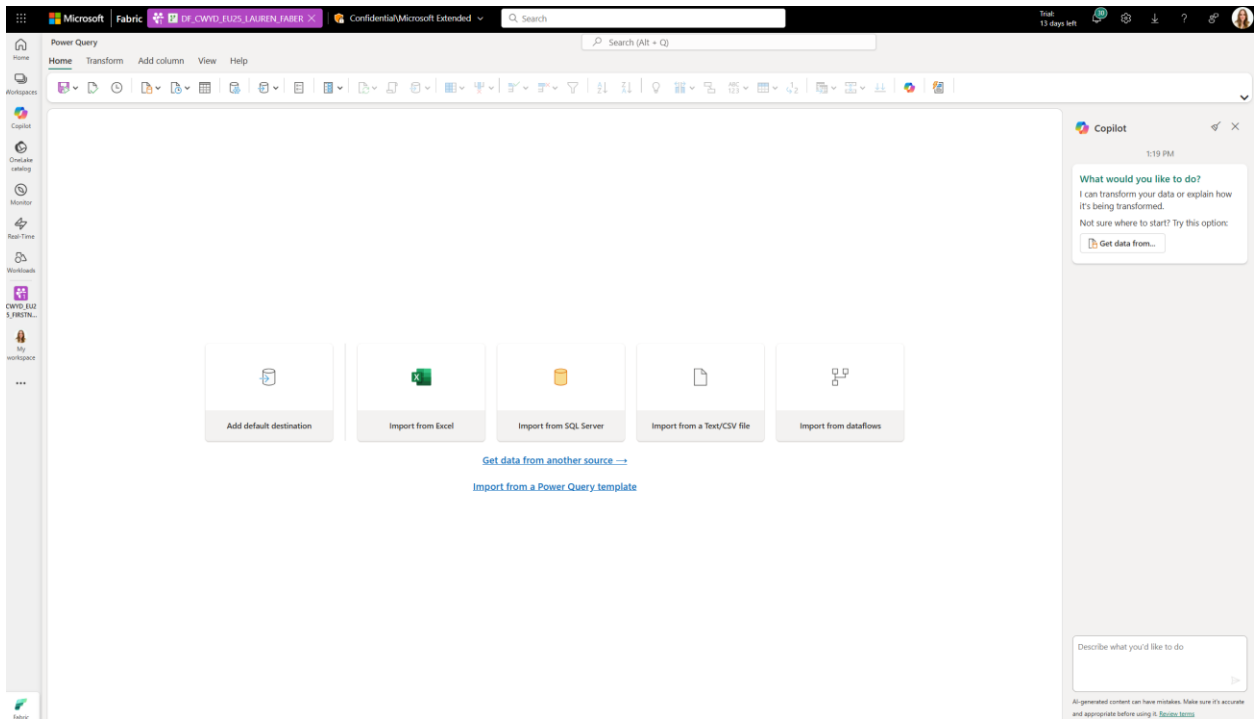


\*You can leave the Enable Git integration box checked, or choose to uncheck it. For this workshop it does not matter.

4. Click Create

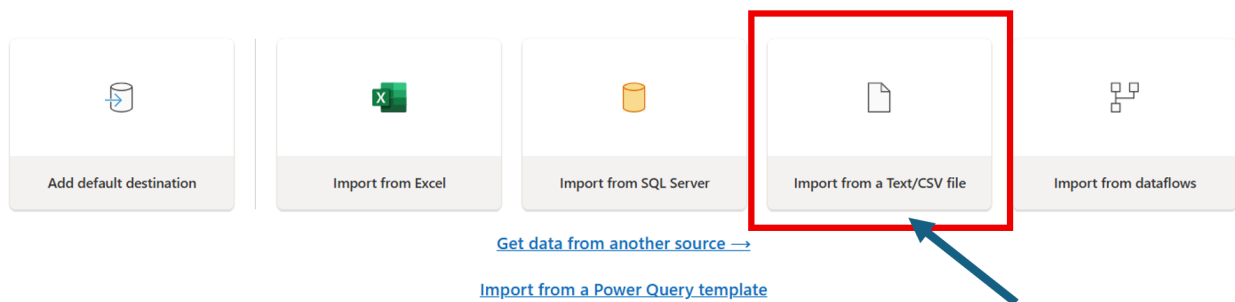
(If you run into the [Failed to create the dataflow](#) error or [failed to get the existing dataflow name](#) troubleshooting guidelines [here](#).)

5. You should now see the dataflow created



### Task 3: Uploading data

1. Select “Import from a Text/CSV file”




2. By default it will be set to “Link to file.” Paste the following link into the “File path or URL” input box:

<https://raw.githubusercontent.com/microsoft/FabConEU25CWYDWorkshop/refs/heads/main/Lab%201%20-%20CWYDIAD%20dirty%20customer%20CSV.csv>

Get data

## Connect to data source

 **Text/CSV**  
File  
[Learn more](#)

### Connection settings

☒ Link to file ☐ Upload file 

File path or URL \*

[Browse OneDrive...](#)

### Connection credentials

Connection



Connection name

Data gateway



Authentication kind

Privacy Level

☐ This connection can be used with on-premises data gateways and VNet data gateways.

### 3. Click “Next”. You will see a screen like the following

Get data

#### Preview file data

URL: <https://raw.githubusercontent.com/microsoft/FabConEU25CWYDWorkshop/refs/heads/main/Lab%201%20-%20CWYDIAD%20dirty%20customer%20CSV.csv>

File origin

65001: Unicode (UTF-8)

Delimiter

Comma

Data type detection

Based on first 200 rows

Column1	Column2	Column3	Column4	Column5	Column6	Column7	Column8	Column9	Column10	Column11
Customer ID	Address	Age	Brand Affinity	City	Country	Full Name	Generation	Interest Affinity 1	Interest Affinity 2	Reseller ID
129b68be0497ce59797e68492f6...	708 Nienow Annex Unit 19,Columbia,New Jersey	36	Ferrari	Columbia	USA	Mabel Guskowski	Gen-Y	Women's Empower...	Environmental Activism	585
9b4b795687e39931867dfe3a11a...	599 Runte Underpass,Columbia,New Jersey	Thirty-six	Ferrari	Columbia	USA	Jerde, Ramona	Gen-Y	Women's Empower...	Environmental Activism	585
9b4b795687e39931867dfe3a11a...	599 Runte Underpass,Columbia,New Jersey	36	Ferrari	Columbia	USA	Ramona Jerde	Gen-Y	Women's Empower...	Environmental Activism	585
ca0708a20d004f38ee3a9e20fd34...	218 Pollich Passage Suite 10,Columbia,New Jersey	36	Ferrari	Columbia	USA	Goyette, Joan	Gen-Y	Women's Empower...	Environmental Activism	585
c25fd60747bac90d310247f6a52...	238 Gleichner Tunnel Apt 32,Columbia,New Jersey	36	Ferrari	Columbia	USA	Freda Waters	Gen-Y	Women's Empower...	Environmental Activism	#585
916211a5e4841af6dc97be8703...	241 Balistreri Corner,Columbia,New Jersey	36	Ferrari	Columbia	USA	Patsy Zemlak	Gen-Y	Women's Empower...	Environmental Activism	585
700c62e8a47dfbd328eebd49ce...	579 Corkery Street Apt 89,Columbia,New Jersey	36	Ferrari	Columbia	USA	Minnie Powlowski	Gen-Y	Women's Empower...	Environmental Activism	585
67595214c5c2f6d5a715cc787ead...	525 Stanton Lane,Columbia,New Jersey	36	Ferrari	Columbia	USA	Boyle, Veronica	Gen-Y	Women's Empower...	Environmental Activism	585
f9b4f1edacbe310a4849b1c53a77...	122 Rosenbaum Forge Suite 83,Columbia,New Jersey	36	Ferrari	Columbia	USA	Della Kiehn	Gen-Y	Women's Empower...	Environmental Activism	585
f9b4f1edacbe310a4849b1c53a77...	122 Rosenbaum Forge Suite 83,Columbia,New Jersey	36	Ferrari	Columbia	USA	Della Kiehn	Gen-Y	Women's Empower...	Environmental Activism	585
5527b16ce30b5182e18f5f822e43...	790 Dickinson Frontage road Unit 22,Columbia,New Jersey	Thirty-six	Ferrari	Columbia	USA	Carolyn Beahan	Gen-Y	Women's Empower...	Environmental Activism	585
5527b16ce30b5182e18f5f822e43...	790 Dickinson Frontage road Unit 22,Columbia,New Jersey	36	Ferrari	Columbia	USA	Carolyn Beahan	Gen-Y	Women's Empower...	Environmental Activism	585
0e2c6e92a4a32f521c0301e05d3...	1036 Gutkowski Knoll Apt 21,Columbia,New Jersey	36	Ferrari	Columbia	USA	Janet Willms	Gen-Y	Women's Empower...	Environmental Activism	585
67595214c5c2f6d5a715cc787ead...	525 Stanton Lane,Columbia,New Jersey	36	Ferrari	Columbia	USA	Veronica Boyle	Gen-Y	Women's Empower...	Environmental Activism	585
a7f88031cfa03041f13b2652bb1...	968 Hill Green,Columbia,New Jersey	Thirty-six	Ferrari	Columbia	USA	Rochelle Hermiston	Gen-Y	Women's Empower...	Environmental Activism	#585
9090fc749b4d9064d9a5a950fca...	471 Klocko Alley,Columbia,New Jersey	36	Ferrari	Columbia	USA	Candace Vandervort	Gen-Y	Women's Empower...	Environmental Activism	585
17edf50b64062314ab878a511cd...	448 Batz Grove,Columbia,New Jersey	36	Ferrari	Columbia	USA	Erma Kling	Gen-Y	Women's Empower...	Environmental Activism	585
59aca4b0a5ecad22fbaa97a88aa4...	417 Schaefer Lodge Unit 63,Columbia,New Jersey	36	Ferrari	Columbia	USA	Hilli, Desiree	Gen-Y	Women's Empower...	Environmental Activism	585
f7bec3f875135a75c1f500036b64...	158 Rowe Corner,Columbia,New Jersey	36	Ferrari	Columbia	USA	Shella Ruecker	Gen-Y	Women's Empower...	Environmental Activism	585
dbcbfd969dd4c5dbd77a8d40afa...	576 Haley Mission Unit 22,Columbia,New Jersey	36	Ferrari	Columbia	USA	Angela Goodwin	Gen-Y	Women's Empower...	Environmental Activism	585
22cadb6975b533bd1d8ff6d223e4...	456 Weimann Trafficway Suite 42,Columbia,New Jersey	Thirty-six	Ferrari	Columbia	USA	Suzanne Leffler	Gen-Y	Women's Empower...	Environmental Activism	585
57df25c5e424d9c099d6d55ce8bd...	781 Wintheiser Heights,Columbia,New Jersey	36	Ferrari	Columbia	USA	Cruickshank, Isabel	Gen-Y	Women's Empower...	Environmental Activism	585
bc90d01afc4c4c059d002feef5481...	206 Skiles Park,Columbia,New Jersey	36	Ferrari	Columbia	USA	Josephine Weimann	Gen-Y	Women's Empower...	Environmental Activism	585
63617ab19337c6b521c8783bacb...			Ferrari	Columbia	USA		Gen-Y	Women's Empower...	Environmental Activism	
7669eddc7ecb047a1571d8b6b6...	863 Pfannerstill Course Suite 83,Columbia,New Jersey	36	Ferrari	Columbia	USA	Winifred Gutmann	Gen-Y	Women's Empower...	Environmental Activism	585
22c8c8ba9d6281d9d5fd3bd65e...	845 Skiles Throughway Suite 62,Columbia,New Jersey	36	Ferrari	Columbia	USA	Candace Conn	Gen-Y	Women's Empower...	Environmental Activism	585

Back

Cancel

Add table using examples

Create

### 4. Click “Create”

You should now see all your data loaded into the dataflow!

## Task 4: transforming data with Copilot

Now we will use Copilot to clean the messy data in our CSV. Before jumping into the rest of the lab see if you can identify problems with the data. There are at least 8 things that need to be fixed. How many can you find? (See answers at [in index](#))

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	Customer ID	Address	Age	Brand Affinity	City	Country	Full Name	Generation	Interest Affinity 1	Interest Affinity 2	Reseller ID	State	Zip Code	
1	129b68be0497ce59797e6849216de86	708 Nienow Annex Unit 19,Columbia,New Jersey	36		Columbia	USA	Mabel Guskowski	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
2	9b4b795687e39931867dfe3a11a62e36	599 Runte Underpass,Columbia,New Jersey	Thirty-six		Columbia	USA	Jerde, Ramona	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
3	9b4b795687e39931867dfe3a11a62e36	599 Runte Underpass,Columbia,New Jersey	36		Columbia	USA	Ramona Jerde	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
4	ca0708a20d004f38ee3a92e0fd34a140	218 Pollich Passage Suite 10,Columbia,New Jersey	36		Columbia	USA	Goyette, Joan	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
5	c25f60747bac90d310247f6a52e5b0	238 Gleichner Tunnel Apt 32,Columbia,New Jersey	36		Columbia	USA	Freda Waters	Gen-Y	Women's Empowerment	Environmental Activism	#585 NJ		7832	
6	916211a5e48481a6dc097be870307f1	241 Balistrieri Corner,Columbia,New Jersey	36		Columbia	USA	Patsy Zemlak	Gen-Y	Women's Empowerment	Environmental Activism	#585 NJ		7832	
7	700c62e8a47dfb8d32eebd49ced7ce1	579 Corkery Street Apt 89,Columbia,New Jersey	36		Columbia	USA	Minnie Powlowski	Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
8	67595214c5c2605a715cc787ead1b3e	525 Stanton Lane,Columbia,New Jersey	36		Columbia	USA	Boyle, Veronica	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
9	19b4f1edacbe310a4849b1c53a77775e	122 Rosenbaum Forge Suite 83,Columbia,New Jersey	36		Columbia	USA	Della Kiehn	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
10	19b4f1edacbe310a4849b1c53a77775e	122 Rosenbaum Forge Suite 83,Columbia,New Jersey	36		Columbia	USA	Della Kiehn	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
11	5527b16ce30b5182e18f5822e43f342	790 Dickinson Frontage road Unit 22,Columbia,New Jersey	Thirty-six		Columbia	USA	Carolyn Beahan	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
12	5527b16ce30b5182e18f5822e43f342	790 Dickinson Frontage road Unit 22,Columbia,New Jersey	36		Columbia	USA	Carolyn Beahan	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
13	0e2c6e92a432f521c0f01a05d33d4f8	1036 Gutkowski Knoll Apt 21,Columbia,New Jersey	36		Columbia	USA	Janet Willms	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		#7832	
14	67595214c5c2605a715cc787ead1b3e	525 Stanton Lane,Columbia,New Jersey	36		Columbia	USA	Veronica Boyle	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
15	7f88031cfa0304113b2652bb13c6	968 Hill Green,Columbia,New Jersey	Thirty-six		Columbia	USA	Rochelle Hermiston	Gen-Y	Women's Empowerment	Environmental Activism	#585 NJ		7832	
16	909bfc749b4d9064d9a5a950fca94a8	471 Klocko Alley,Columbia,New Jersey	36		Columbia	USA	Candace Vandervort	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
17	f7ef50b64062314ab878a511cd0b19d	448 Batz Grove,Columbia,New Jersey	36		Columbia	USA	Erma Kling	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
18	59aca4b0a5ecad22fbaa97a888a4a5ed	417 Schaefer Lodge Unit 63,Columbia,New Jersey	36		Columbia	USA	Hilli, Desiree	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
19	78ec3f875135a75c1f500036b6437e1	158 Rowe Corner,Columbia,New Jersey	36		Columbia	USA	Shelia Ruecker	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
20	dbcfb969dd4c5dbd77ab404fa25b75	576 Haley Mission Unit 22,Columbia,New Jersey	36		Columbia	USA	Angela Goodwin	Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
21	22cad6b97f5b5338d1d8f6d23e492d0	456 Weimann Trafficway Suite 42,Columbia,New Jersey	Thirty-six		Columbia	USA	Suzanne Leffler	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
22	57df25c4e24d9c099d6d55c8dcad02e	781 Wintheiser Heights,Columbia,New Jersey	36		Columbia	USA	Cruickshank, Isabel	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
23	bc30d01afc4c059a6022eef5481d2a	206 Skiles Park,Columbia,New Jersey	36		Columbia	USA	Josephine Weimann	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		#7832	
24	63617ab19337c6b521c8783bacbdc...	863 Planerstill Course Suite 83,Columbia,New Jersey	36		Columbia	USA	Winifred Gutmann	Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
25	7669eddc7ecb047a1571d8b6b613cbb	845 Skiles Thoroughway Suite 62,Columbia,New Jersey	36		Columbia	USA	Candace Conn	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
26	22c8bca99d6281d9d5d3dbd65e129b8	339 Eichmann Meadow,Columbia,New Jersey	Thirty-six		Columbia	USA	Ashley Heidenreich	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
27	69173418d10b17986d446eba5988cd7b	465 Rodriguez Path,Columbia,New Jersey	36		Columbia	USA	Kay Champlin	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
28	0e56f98c8e80b0a024212ac517908301	197 McClure Passage Apt 18,Columbia,New Jersey	36		Columbia	USA	Elena Tremblay	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
29	9b4b795687e39931867dfe3a11a62e36	599 Runte Underpass,Columbia,New Jersey	36		Columbia	USA	Ramona Jerde	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
30	0b316106166206a33f672e0daf51534	656 Barrows Ramp,Columbia,New Jersey	36		Columbia	USA	Geraldine Wisoky	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
31	5459fd266134e7c7f654e0b5d4b79f				Columbia	USA			Women's Empowerment	Environmental Activism				
32	90de846ad50feaa042879411ebf590	497 Rosenbaum Plaza,Columbia,New Jersey	36		Columbia	USA	Elsie Sipes	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
33	eed4d65a3989e753aaa38d2c132a6e5	686 Murray Boulevard Apt 41,Columbia,New Jersey	36		Columbia	USA	Ashley Herzog	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	
34	5a291c520425dca11b9564863e84042	440 Huel Skyway Unit 23,Columbia,New Jersey	36		Columbia	USA	Heather Klocko	Gen-Y	Women's Empowerment	Environmental Activism	585 NJ		7832	

**\*\*Troubleshooting tip.** If Copilot starts to get overwhelmed, save your progress, exit the dataflow, re-enter, and try again.

### 1. Open the Copilot pane

The screenshot shows the Power Query interface with the Copilot pane open on the right. The main area displays a data table with columns: Customer ID, Address, Age, Brand Affinity, City, Country, Full Name, Generation, Interest Affinity 1, Interest Affinity 2, Reseller ID, State, Zip Code. The Copilot pane on the right has a search bar and a 'What would you like to do?' prompt with a 'Get data from...' button. A red box highlights the Copilot icon in the top right corner of the Power Query ribbon.

- c. "In Full\_Name switch order of names if there is a comma, then remove the comma"

Before

Full_Name	
Mabel Gusikowski	C
Jerde, Ramona	C
Ramona Jerde	C
Joan Goyette	C
Freda Waters	C
Patsy Zemlak	C
Minnie Powlowski	Y
Boyle, Veronica	C
Della Kiehn	C
Della Kiehn	C
Carolyn Beahan	C
Carolyn Beahan	C
Janet Willms	C
Veronica Boyle	C
Rochelle Hermiston	C
Candace Vandervort	C
Erma Kling	C
Hilll, Desiree	C
Shelia Ruecker	C
Angela Goodwin	Y

After

Full_Name	
Mabel Gusikowski	Gr
Ramona Jerde	Gr
Ramona Jerde	Gr
Joan Goyette	Gr
Freda Waters	Gr
Patsy Zemlak	Gr
Minnie Powlowski	Y
Veronica Boyle	Gr
Della Kiehn	Gr
Della Kiehn	Gr
Carolyn Beahan	Gr
Carolyn Beahan	Gr
Janet Willms	Gr
Veronica Boyle	Gr
Rochelle Hermiston	Gr
Candace Vandervort	Gr
Erma Kling	Gr
Desiree Hilll	Gr
Shelia Ruecker	Gr

d. "Remove #" anywhere in the data"

Before

After

ABC Reseller_ID	ABC
585	N
585	N
585	N
585	N
585	N
585	N
#585	N
585	N
585	N
585	N
585	N
585	N
585	N
585	N
585	N
585	N
#585	N
585	N

ABC Reseller_ID	ABC
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑
585	↑

Step 3: remove redundant data

Now the data looks uniform, but we still have some problems. We have empty of null values, and rows that are duplicates. Let’s fix that!

- e. “Remove rows with null or empty values”

Before

rmcColumns({# Custom 3 , { State , each IT \_ = New York , Lien NY

ABC Address	ABC Age	ABC Brai
576 Haley Mission Unit 22,Columbia,New Jersey	36	
456 Weimann Trafficway Suite 42,Columbia,New Jersey	36	
781 Wintheiser Heights,Columbia,New Jersey	36	
206 Skiles Park,Columbia,New Jersey	36	
863 Pfannerstill Course Suite 83,Columbia,New Jersey	36	
845 Skiles Throughway Suite 62,Columbia,New Jersey	36	
339 Eichmann Meadow,Columbia,New Jersey	36	
465 Rodriguez Path,Columbia,New Jersey	36	
197 McClure Passage Apt 18,Columbia,New Jersey	36	
599 Runte Underpass,Columbia,New Jersey	36	
656 Barrows Ramp,Columbia,New Jersey	36	
497 Rosenbaum Plaza,Columbia,New Jersey	36	

After



11	cecb7ef824f42e921a8c4a62bbac...	282 Krajcik Green Apt 27,Miami,Florida	18	Ferrari	Miami	t
12	787b5a23a2b12f1838eaf76ad6c0...	109 Davis Heights Unit 93,Miami,Florida	18	Ferrari	Miami	t
13	3a6c0c78f953a18bb9add670641...	900 Smitham Quay Unit 16,Miami,Florida	18	Ferrari	Miami	t
14	7f9a53dd7db05197572131dc431...	517 Metz Bay Suite 94,Miami,Florida	23	Ferrari	Miami	t
15	f420b7d3e35cb3cf639c69542610...	1079 Erdman Lock Unit 81,Miami,Florida	23	Ferrari	Miami	t
16	511ad80d9b20e81e644a78a0daa...	667 Schowalter Skyway Suite 26,Miami,Florida	23	Ferrari	Miami	t
17	66dacbbe43dae888b1f4fce505f0...	481 Nolan Rest Suite 50,Miami,Florida	23	Ferrari	Miami	t
18	dd4a02296060f4786ebf8e39f0e0...	150 Durgan Station Apt 5,Miami,Florida	22	Ferrari	Miami	t
19	341ef7ef735ce882da2da53973c1...	841 Kuvalis Throughway,Miami,Florida	22	Ferrari	Miami	t
20	4b7ef26a042a3ef4b4488c22d03c...	1058 Lueilwitz Bay,Miami,Florida	22	Ferrari	Miami	t
21	47628210fe3b9bc5216ef64d4009...	961 Schultz Gate Unit 57,Miami,Florida	27	Ferrari	Miami	t
22	0bc121a4f35a3fb5c4b09b099027...	516 Heller Landing,Miami,Florida	27	Ferrari	Miami	t
23	18ae8fa85845d12bb807178c452...	201 Zulauf Stravenue,Miami,Florida	27	Ferrari	Miami	t
24	183e0e9a0cfa59175583cb9156b6...	897 Watsica Vista Unit 23,Miami,Florida	27	Ferrari	Miami	t

#### f. "Remove duplicate rows"

Before

1	129b68be0497ce59797e68492f6...	708 Nienow Annex Unit 19,Columbia,New Jersey	36		Columbia	USA	Rai
2	9b4b795687e39931867dfe3a11a...	599 Runte Underpass,Columbia,New Jersey	36		Columbia	USA	Rai
3	9b4b795687e39931867dfe3a11a...	599 Runte Underpass,Columbia,New Jersey	36		Colum	USA	Rai
4	ca0708a20d004f38ee3a92e0fd34...	218 Pollich Passage Suite 10,Columbia,New Jersey	36		Columbia	USA	Ins

After

1	129b68be0497ce59797e68492f6...	708 Nienow Annex Unit 19,Columbia,New Jersey	36				
2	9b4b795687e39931867dfe3a11a...	599 Runte Underpass,Columbia,New Jersey	36				
3	ca0708a20d004f38ee3a92e0fd34...	218 Pollich Passage Suite 10,Columbia,New Jersey	36				
4	c25fcf60747bac90d310247f6a52...	238 Gleichner Tunnel Apt 32,Columbia,New Jersey	36				

#### Step 4: Update to correct data types

Lastly, we want to ensure our data is ready to be put into a semantic model, let's update all the field types to be the right type.

#### g. "Update field type to whole number instead of text where applicable"

Before

	ABC 123	Customer_ID	ABC 123	Address	ABC 123	Age	ABC 123	E
1		129b68be0497ce59797e68492f6...		708 Nienow Annex Unit 19,Columbia,New Jersey		36		
2		9b4b795687e39931867dfe3a11a...		599 Runte Underpass,Columbia,New Jersey		36		

After

	ABC 123	Customer_ID	ABC 123	Address	123	Age	ABC 123	Brand_Affinity	ABC 123	City
1		129b68be0497ce59797e68492f6...		708 Nienow Annex Unit 19,Columbia,New Jersey		36				Columbia
2		9b4b795687e39931867dfe3a11a...		599 Runte Underpass,Columbia,New Jersey		36				Columbia
3		ca0708a20d004f38ee3a92e0fd34...		218 Pollich Passage Suite 10,Columbia,New Jersey		36				Columbia

You did it! You successfully used Copilot to clean your data!

Now hit the save and run button in the top left corner

The screenshot shows the Microsoft Power Query interface. At the top, the title bar reads "DF\_CWYD\_EU25\_SID\_VEMURI". Below it, the "Power Query" tab is active, and a green checkmark indicates "Draft saved". The ribbon includes "Home", "Transform", "Add column", "View", and "Help". The "Home" ribbon is expanded, showing the "Save & run" button (a purple icon with a green play button) and a tooltip that reads: "Save & run. Save your changes to this dataflow and run the dataflow." Below the ribbon, a data table is visible with the following columns: "ID", "Name", "Generation", and "Interest\_Affinity\_1". The table contains 10 rows of data, with IDs ranging from 447 to 457. The "Name" column shows names like "Mante", "McKenzie", "s Kiehn", "o Zulauf", "erry", "s Willms", "Smitham", "l Hammes", "Marks", and "n Miller". The "Generation" column is consistently "Gen-Y". The "Interest\_Affinity\_1" column lists various interests like "Women's Empowerment" and "Auto Racing".

ID	Name	Generation	Interest_Affinity_1
447	Mante	Gen-Y	Women's Empowerment
448	McKenzie	Gen-Y	Auto Racing
449	s Kiehn	Gen-Y	Women's Empowerment
450	o Zulauf	Gen-Y	Auto Racing
451	erry	Gen-Y	Auto Racing
452	s Willms	Gen-Y	Auto Racing
453	Smitham	Gen-Y	Auto Racing
454	l Hammes	Gen-Y	Women's Empowerment
455	Marks	Gen-Y	Women's Empowerment
456	y Hudson	Gen-Y	Women's Empowerment
457	n Miller	Gen-Y	Women's Empowerment

## Appendix

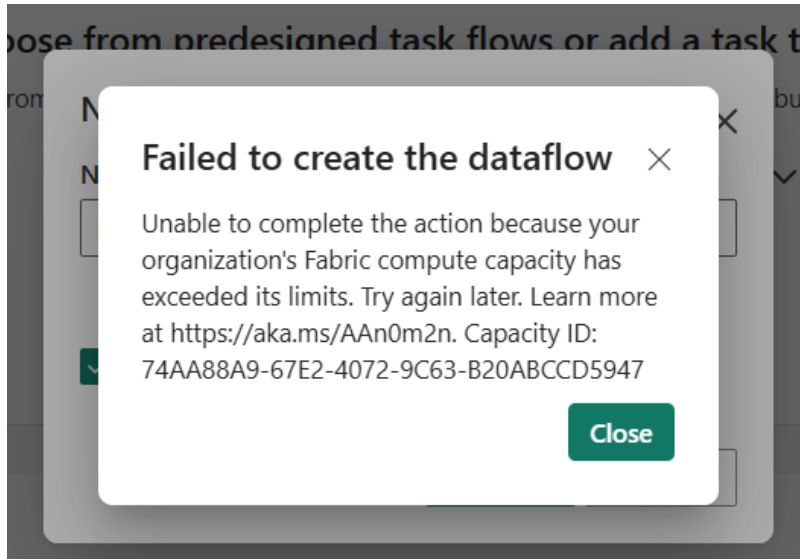
Answers for what need to be cleaned/transformed in the data

1. First row needs to be header
2. Age column has a mix of word and numbers
3. Generation is inconsistent
4. Reseller ID and Zip Code has random “#”s
5. State is inconsistent
6. Full Name has inconsistent format (some are First name Last name and others are Last Name, First name)
7. Null values
8. Duplicate items

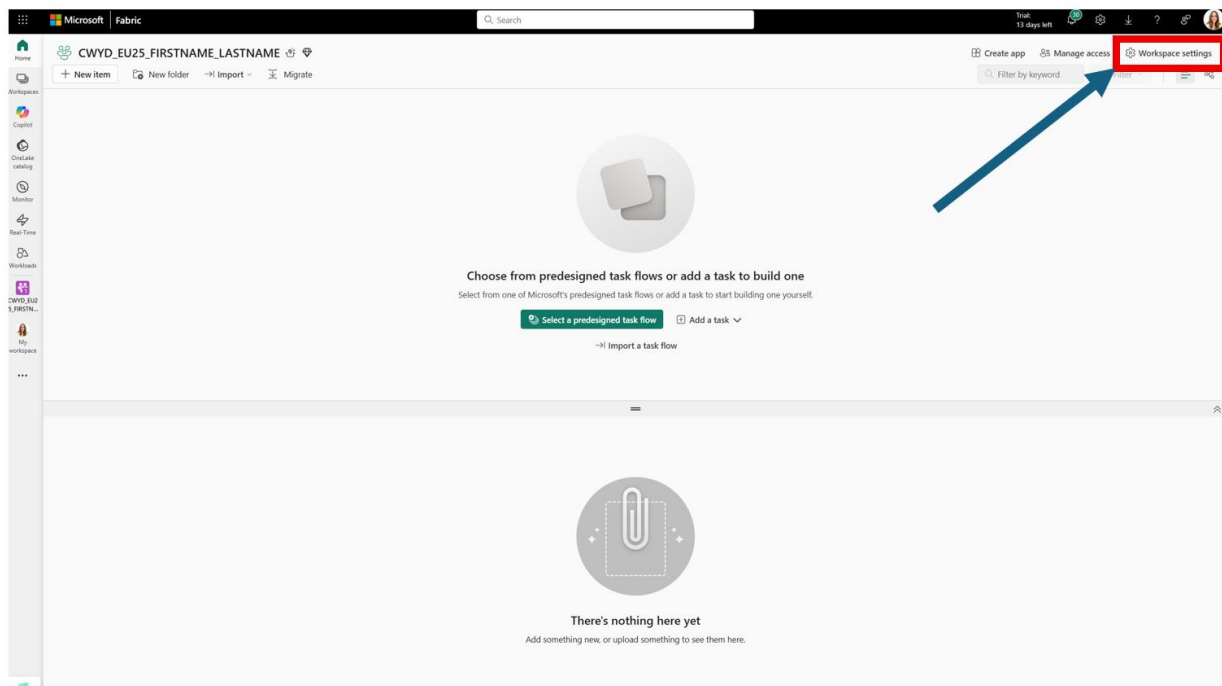
## Troubleshooting

### Failed to create the dataflow

- If you run into the error failed to create the dataflow, follow these steps



- Click on the workspace settings in the top right of the workspace



- Select “License info” and click the Edit icon

**Workspace settings**

Search

**License info**

Choose a license for this workspace.

**License Configuration**

Current license  
**Trial**

License capacity  
**Trial-20250228T133033Z-ALgjmjUZK0alWnSaw0IPVg**  
SKU: FT1, Region: West Central US  
Capacity ID: 74AA88A9-67E2-4072-9C63-B20ABCCD5947

Semantic model storage format  
**Small semantic model storage format**

**Connection link**

Use this link to connect third-party software to the workspace. Copy the link and add it to your third-party software.

powerbi://api.powerbi.com/v1.0/myorg/CWYD\_EU25\_FIRSTNAME\_LASTNAME

- Under license capacity, select the dropdown. Choose any new capacity and try again. If you still get the error try a few more capacities. If the error persists please raise your hand.

(preview)


J

ngs


ience

e

### ☐ Fabric capacity

Select Fabric capacity if the workspace will be hosted in a Microsoft Fabric capacity. With Fabric capacities, users can create Microsoft Fabric items and collaborate with others using Fabric features and experiences. Explore new capabilities in Power BI, Data Factory, Data Engineering, and Real-Time Intelligence, among others. [Learn more](#) 

### ☒ Trial

Select Trial to assign this workspace to a Fabric trial capacity. A Microsoft Fabric trial capacity allows user to explore the capabilities of Microsoft Fabric like Data Factory, Data Engineering and Real-Time Intelligence among others. [Learn more](#) 

### License capacity

Please specify the remote computing resources your items will use with this workspace.

Trial-20250227T174539Z-4wkTZyrezke4cZPFmFt9QQ - West Central US

Trial-20250227T192529Z-FGTvbYB9gUu2gQLXdsjDdg - West Central US

Trial-20250227T180136Z-HejdwS651kOxGwQOn0FEZQ - West Central US

Trial-20250227T175825Z-P\_OPWXvo90WPdXETv6EN2g - West Central US

Trial-20250227T175148Z-JLza8YgWs0GAo5rW83EoZg - West Central US

Trial-20250227T174539Z-4wkTZyrezke4cZPFmFt9QQ - West Central US

Trial-20250227T165659Z-SwDMC4cvWUuoC6qJr3yU5g - West Central US

Trial-20250227T163654Z-fAMLUfjhYkeSGpjrkhUBDw - West Central US

Trial-20250227T162703Z-O0iZiGKwxEuvAEK9BxEKcg - West Central US

Trial-20250227T162036Z-D2-eV-EGsUCfQs9Erj9TXA - West Central US

Trial-20250227T161633Z-eOmGSLUuEmI0KeeC-OLKa - West Central US

### Conn

Use th

your t

powerbi://api.powerbi.com/v1.0/myorg/CWYD\_EU25\_FIRSTNAME\_LASTNAME

- Click “Select license”
- Try creating a dataflow Gen2 again

## Failed to get the existing dataflow names

If you run into a “failed to get the existing dataflow names” error. Try creating the dataflow in a different workspace (like MyWorkspace)

