

## **Lab – Build a Direct Lake dataset**

# Content in Lab

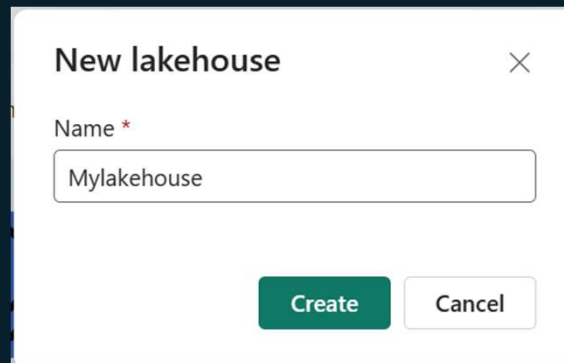
1. Build lakehouse
2. Ingest data
3. Build data model
4. Create report

# Prerequisites

- Fabric workspace is created – i.e. workspace is in Trial, Premium capacity or Fabric capacity

# Build lakehouse

1. In workspace, select New > Show all and select Lakehouse under Data Engineering,  
or  
In bottom left-hand corner, select Data Engineering and select Lakehouse on this screen
2. Provide a name for your lakehouse

A screenshot of a 'New lakehouse' dialog box. The dialog has a title bar with 'New lakehouse' and a close button (X). Below the title bar, there is a label 'Name \*' followed by a text input field containing the text 'Mylakehouse'. At the bottom of the dialog, there are two buttons: a green 'Create' button and a white 'Cancel' button with a grey border.

New lakehouse

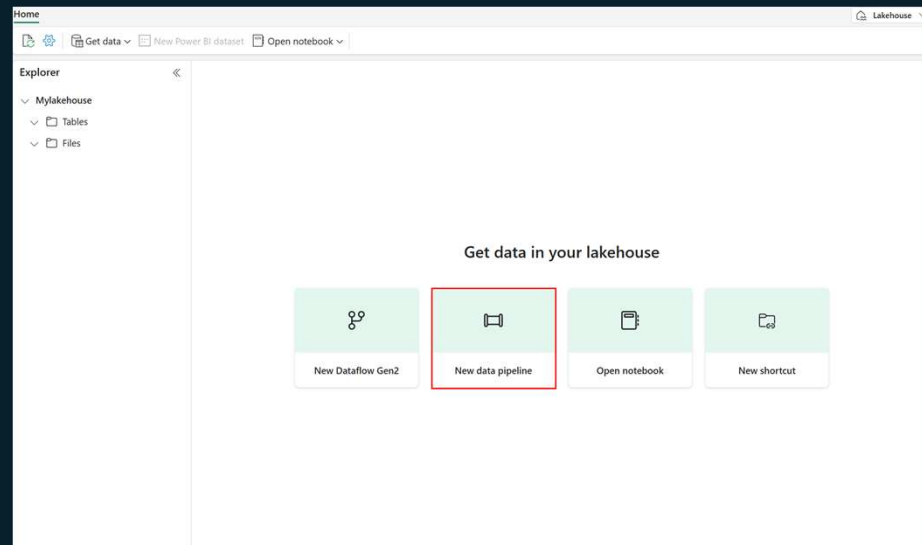
Name \*

Mylakehouse

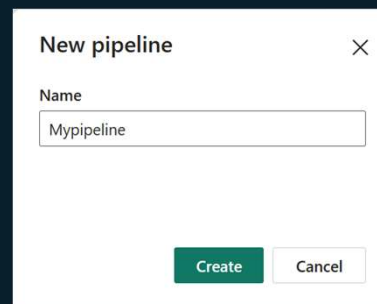
Create Cancel

# Ingest data

1. In the lakehouse, select New data pipeline

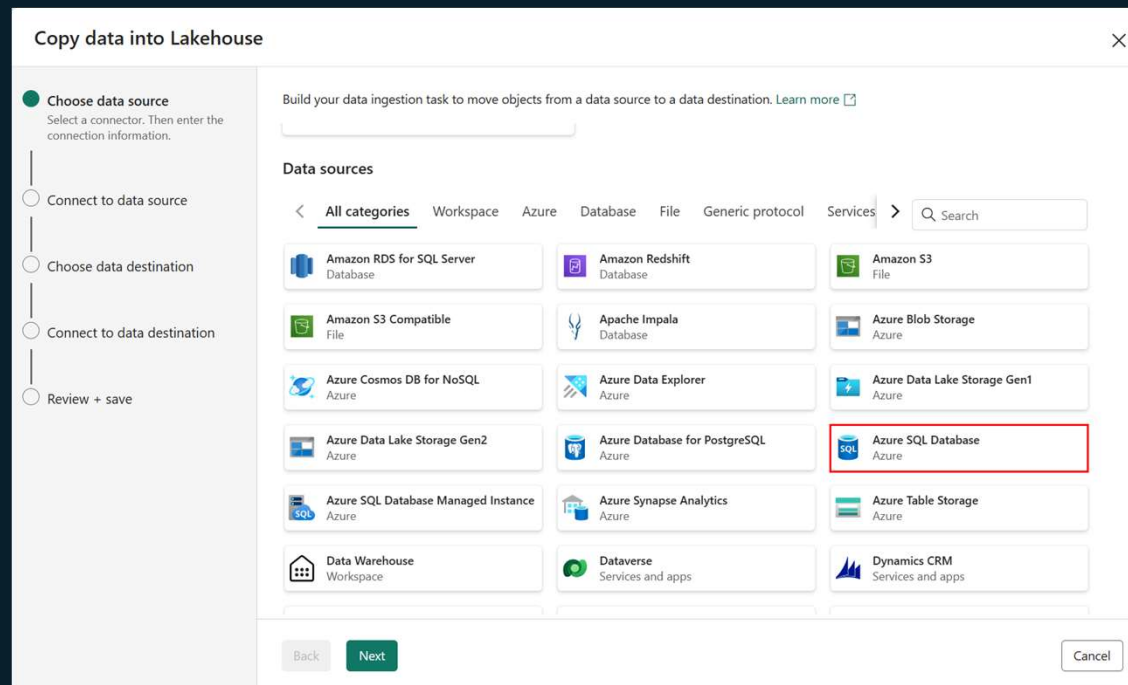


2. Provide a name for your pipeline

A screenshot of a 'New pipeline' dialog box. It has a title bar with 'New pipeline' and a close button. Below the title is a 'Name' label and a text input field containing the text 'Mypipeline'. At the bottom of the dialog are two buttons: 'Create' (in green) and 'Cancel'.





# Ingest data

3. In the data pipeline experience, "Copy data" activity has been pre-selected.  
Scroll to "Data sources" and select "Azure SQL Database"



# Ingest data

4. Use following connection settings:  
Server: laander.database.windows.net  
Database: DemoDW  
Authentication kind: Basic  
Username: fabric5sep  
Password: fabric.12345
5. Select all 4 tables

<input checked="" type="checkbox"/>	Select all
<input checked="" type="checkbox"/>	 dbo.DimCustomerGeo
<input checked="" type="checkbox"/>	 dbo.DimDate
<input checked="" type="checkbox"/>	 dbo.DimProduct
<input checked="" type="checkbox"/>	 dbo.FactInternetSales

### Connection settings

Server \* ⓘ  
laander.database.windows.net

Database  
DemoDW

### Connection credentials

Connection  
Create new connection ▼ ↺

Connection name  
laander.database.windows.net;DemoDW

Authentication kind  
Basic ▼

Username  
fabric5sep

Password  
.....

☒ Use encrypted connection

# Ingest data

6. Make sure your newly created lakehouse is set as your destination

The screenshot shows the 'Copy data into Lakehouse' wizard at step 3, 'Choose data destination'. The left sidebar shows a progress bar with three steps: 'Choose data source' (completed), 'Connect to data source' (completed), and 'Choose data destination' (current step). The main area has a 'Lakehouse' icon and 'Learn more' link. On the right, there are two radio buttons: 'Existing Lakehouse' (selected) and 'Create new Lakehouse'. Below, there is a dropdown menu labeled 'Lakehouse \*' with 'Mylakehouse' selected, and a 'Refresh' button.

7. Make sure all 4 tables are in the list of "Connect to data destination" preferably set "Table action" to Overwrite for each table, if you need to re-run

The screenshot shows the 'Copy data into Lakehouse' wizard at step 4, 'Connect to data destination'. The left sidebar shows a progress bar with four steps: 'Choose data source' (completed), 'Connect to data source' (completed), 'Choose data destination' (completed), and 'Connect to data destination' (current step). The main area has a 'Root folder' section with 'Tables' selected. Below is a 'Source Table' section with a search bar and a list of tables: 'dbo.DimCustomerGeo' (highlighted), 'dbo.DimDate', 'dbo.DimProduct', and 'dbo.FactInternetSales'. On the right, there is a 'Destination table name \*' field with 'dbo\_DimCustomerGeo' entered. Below that, there is a 'Table action' section with 'Append' and 'Overwrite' radio buttons, where 'Overwrite' is selected. There is also an 'Enable partition' checkbox.



# Ingest data

## 7. Save + Run your pipeline

### Copy data into Lakehouse

✓ Choose data source

✓ Connect to data source


✓ Choose data destination

✓ Connect to data destination


● Review + save

Confirm Copy activity summary

#### Copy Summary

 Azure SQL Database

→

 Lakehouse

Source		Destination	
Connection name	laander.database.windows.net;DemoDW laander	Connection name	Mylakehouse
Number of tables	4	Number of tables	4

#### Options

☒ Start data transfer immediately ⓘ

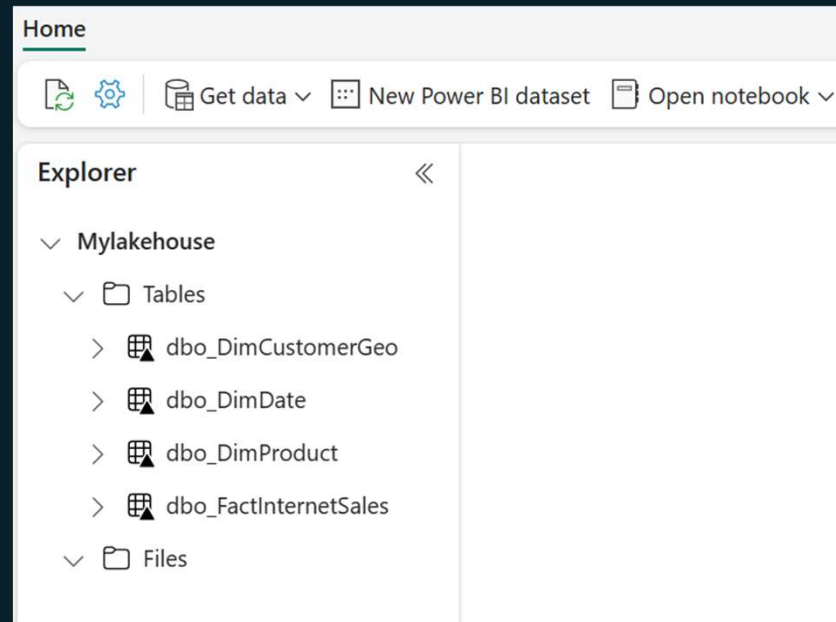
Back

Save + Run

Cancel

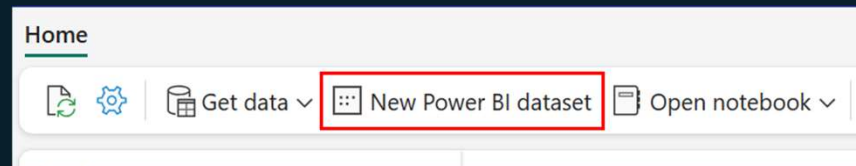
# Ingest data

8. Now your lakehouse has 4 tables

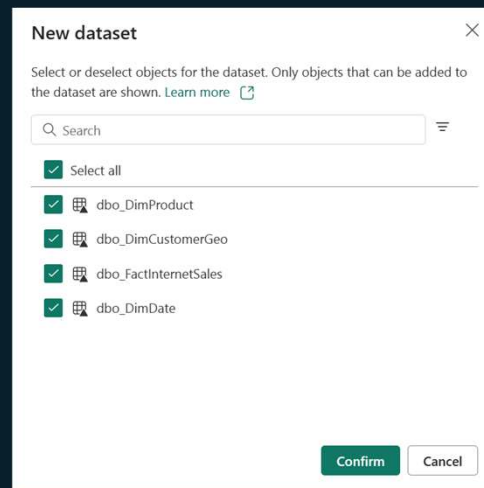


# Build data model

1. In lakehouse view, click on "New Power BI dataset" to build the Direct Lake model



2. Include all 4 tables



# Build data model

3. You are now in the web modeling experience and need to configure relationships between fact and dimension tables

The screenshot displays the Microsoft Power BI web modeling experience interface. At the top, there is a navigation bar with 'Home' and 'Help' tabs. Below this, a ribbon contains icons for 'New measure', 'New column', 'New table', 'Manage roles', 'New report', and 'Edit tables', each with a corresponding label and category (Calculations, Security, Reporting, Modeling). A yellow status bar below the ribbon states: 'Keep in mind your changes will be permanent and automatically saved.' with a 'Learn more' link.

The main workspace shows three dimension tables: **dbo\_DimCustomerGeo**, **dbo\_DimDate**, and **dbo\_DimProduct**. Each table is represented by a card with a list of fields. **dbo\_DimCustomerGeo** includes BirthDate, City, Country, CustomerKey (marked with a sigma symbol), FirstName, Gender, LastName, MaritalStatus, and Occupation. **dbo\_DimDate** includes CalendarQuarter, CalendarYear, DateKey (marked with a sigma symbol), DayNameOfWeek, DayNumberOfMonth, DayNumberOfWeek, DayNumberOfYear, FullDateAlternateKey, and MonthName. **dbo\_DimProduct** includes CategoryName, ProductKey (marked with a sigma symbol), ProductName, and SubcategoryName. Each card has a 'Collapse' button at the bottom.

On the right side, there are two panels: 'Properties' and 'Data'. The 'Properties' panel has a 'Cards' section with two settings: 'Show the database in the header when applicable' (set to 'No') and 'Show related fields when card is collapsed' (set to 'Yes'). The 'Data' panel has a search bar and a list of tables: **dbo\_DimCustomerGeo**, **dbo\_DimDate**, **dbo\_DimProduct**, and **dbo\_FactInternetSales**.

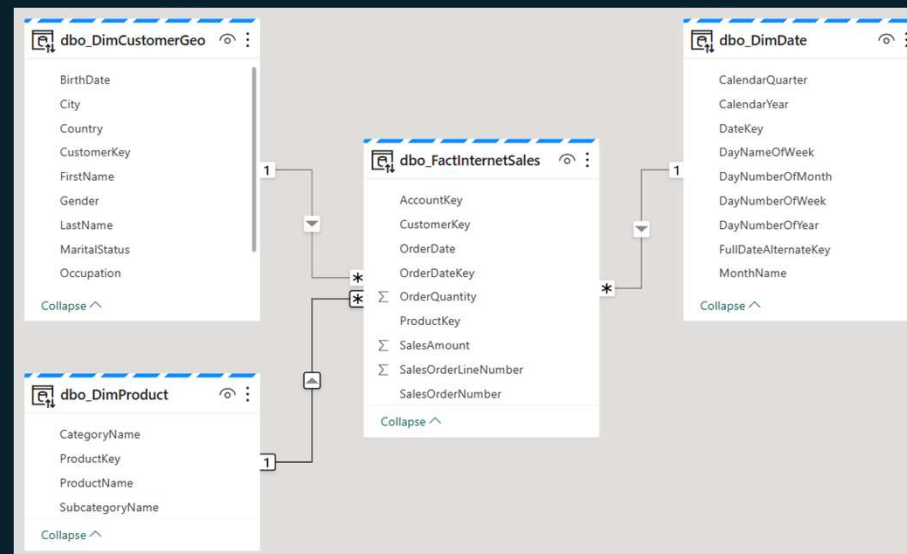
# Build data model

4. Create the following relationships – drag column from/to dimension to/from fact:

`dbo_DimCustomerGeo.CustomerKey = dbo_FactInternetSales.CustomerKey`

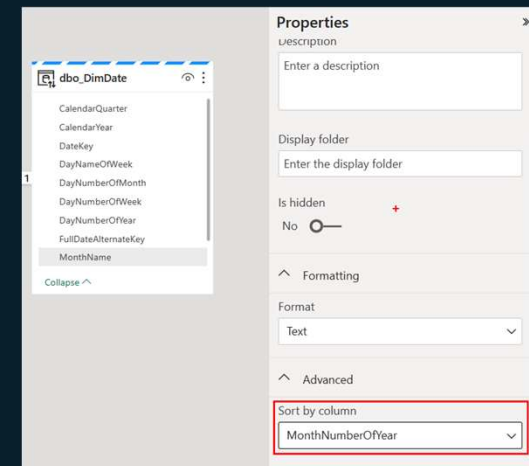
`dbo_DimProduct.ProductKey = dbo_FactInternetSales.ProductKey`

`dbo_DimDate.DateKey = dbo_FactInternetSales.OrderDateKey`



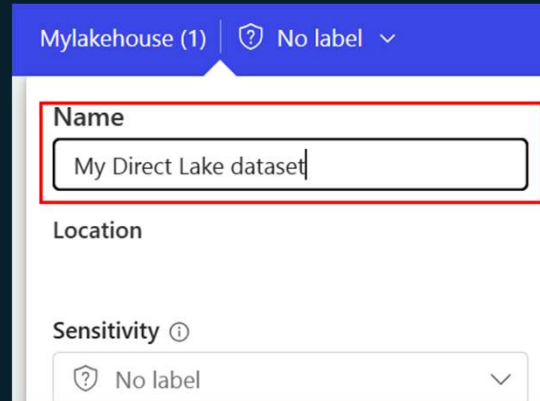
# Build data model

5. Refine the model (*just a little bit, in real-world scenarios you will do more 😊*)
  - > Hide all columns in dbo\_FactInternetSales
  - > Create two measures in dbo\_FactInternetSales
    - Total Sales = SUM(dbo\_FactInternetSales[SalesAmount])
    - Total Orders = SUM(dbo\_FactInternetSales[OrderQuantity])
  - > In dbo\_DimDate select the column MonthName. In Properties > Advanced, change "Sort by column" to MonthNumberOfYear



# Build data model

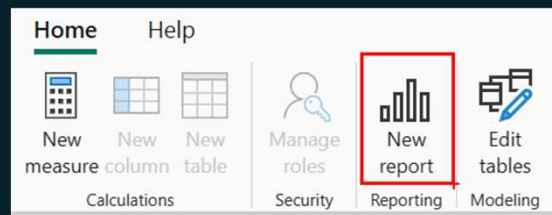
6. Optionally, rename your Direct Lake dataset – click name in upper left corner



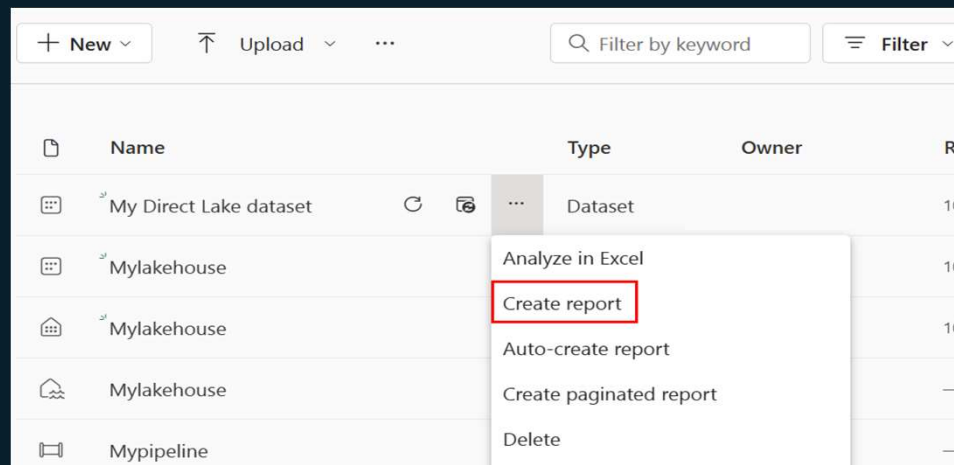
The screenshot shows a user interface for managing a dataset in Microsoft Fabric. At the top, there is a blue header bar with the text 'Mylakehouse (1)' and a dropdown menu showing 'No label'. Below this, the 'Name' field is highlighted with a red rectangle and contains the text 'My Direct Lake dataset'. The 'Location' field is empty. The 'Sensitivity' field is set to 'No label' with a dropdown arrow.

# Create report

1. If you are in the web modeling experience, select “New report”



Else navigate to workspace and select “Create report”





# Create report

- Look at the data model and build your own report, it could look something like this



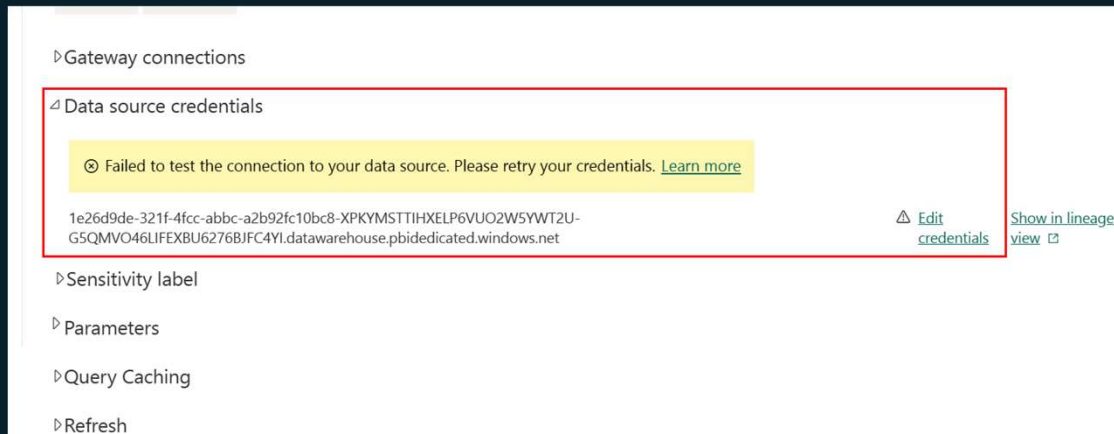
# Extra – Refreshing your Direct Lake dataset

1. You can configure refresh on your Direct Lake dataset by going to the setting of the dataset

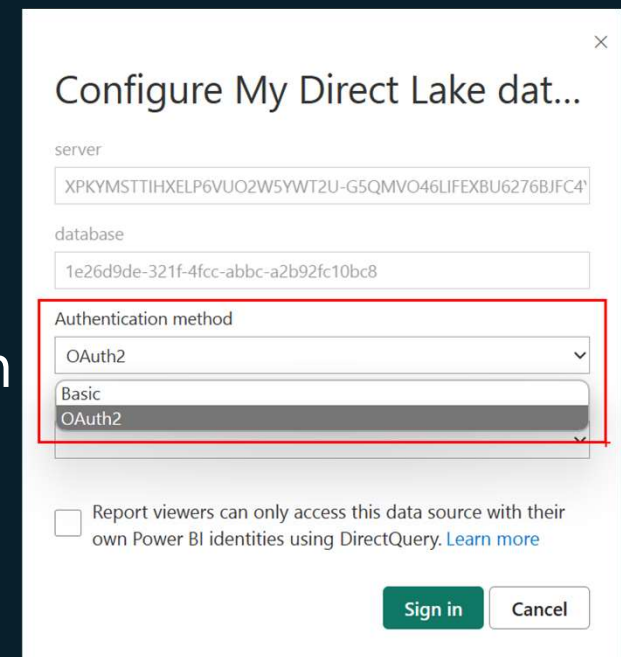
Name	Type	Owner
My Direct Lake dataset	Dataset	Lars202308
Mylakehouse	Analyze in Excel	
Mylakehouse	Create report	
Mylakehouse	Auto-create report	
Mylakehouse	Create paginated report	
Mypipeline	Delete	
Report on Direct Lake dataset	Rename	
	Open data model	
	Settings	
	Manage permissions	
	View lineage	

# Extra – Refreshing your Direct Lake dataset

- Initially, you must enter credentials for your dataset. Select “Edit credentials” under “Data source credentials”

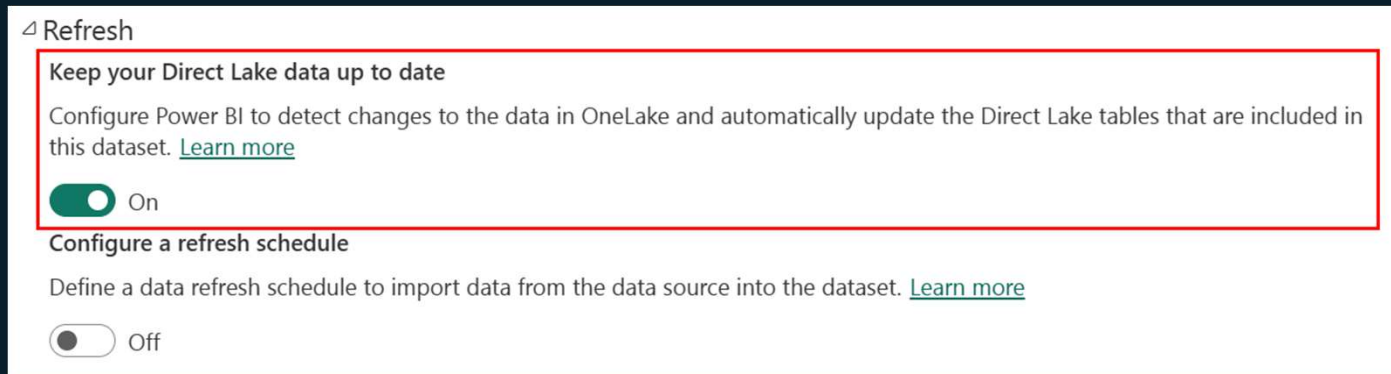


Use “Authentication method” OAuth2 and sign in with your Fabric credentials  
Notice: this is credentials for your lakehouse and not the Azure SQL database



## Extra – Refreshing your Direct Lake dataset

2. Under Refresh a Direct Lake dataset has a new setting “Keep your Direct Lake data up to date” which is turned on by default



Having this setting turned on mean that whenever new data is added to your table in lakehouse, it will automatically be available in your Direct Lake dataset. Turning the setting off mean that you must control the refresh of your dataset either by “Configure a refresh schedule” or trigger refresh using API

# Direct Lake documentation

[Learn about Direct Lake in Power BI and Microsoft Fabric](https://learn.microsoft.com/en-us/power-bi/enterprise/directlake-overview)

(<https://learn.microsoft.com/en-us/power-bi/enterprise/directlake-overview>)