

Microsoft Fabric in a Day Lab Manual – **Lab 1**

Author: Will Crayger

Email: wcrayger@lucidbi.co

LinkedIn: [LinkedIn.com/in/willcrayger/](https://www.linkedin.com/in/willcrayger/)

Course Material: [GitHub.com/Lucid-Will/FabCon-EU-Zero-To-Hero-with-Fabric](https://github.com/Lucid-Will/FabCon-EU-Zero-To-Hero-with-Fabric)

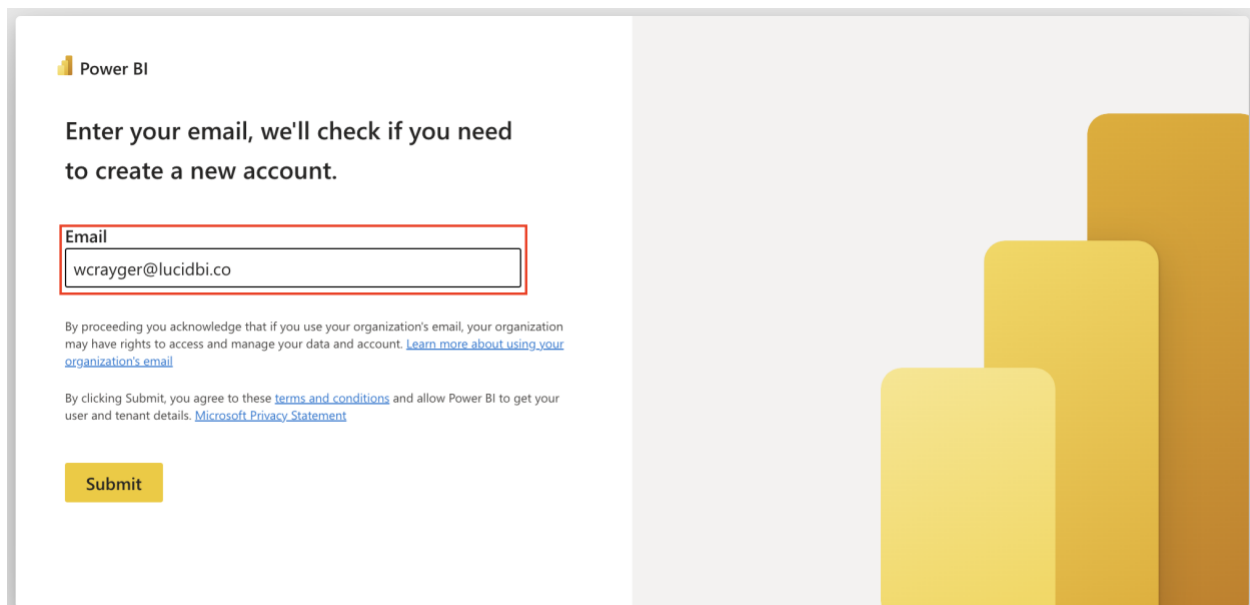
Working with Fabric Lakehouse – Creating your Lakehouse

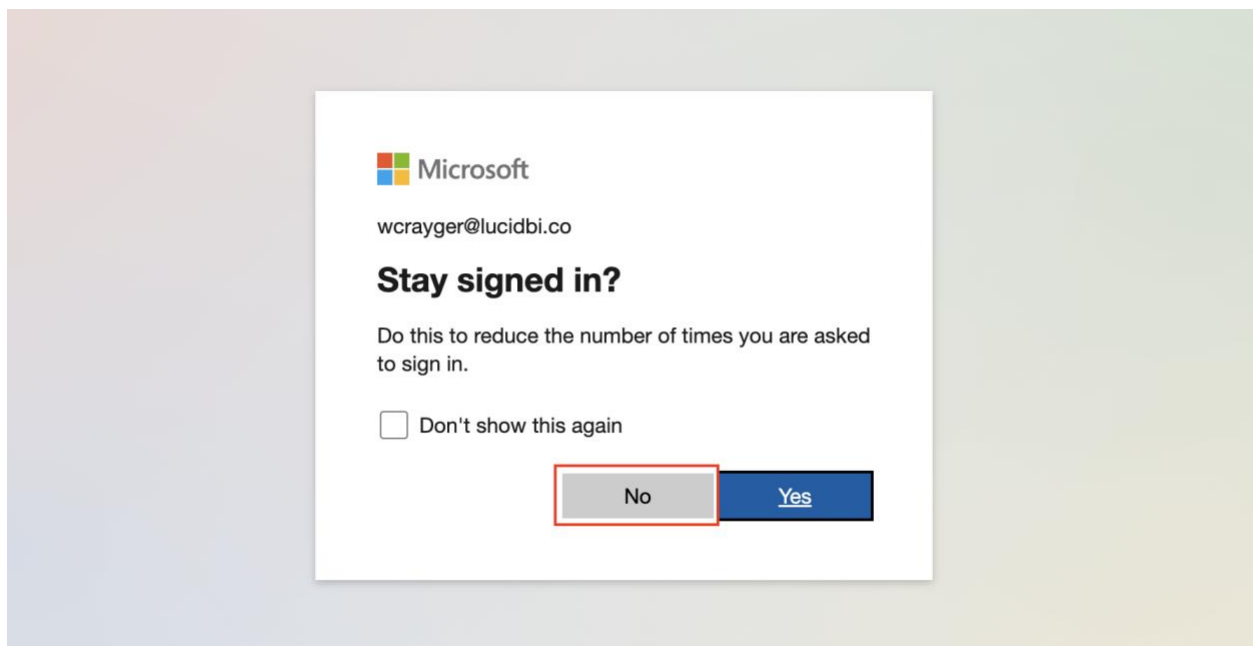
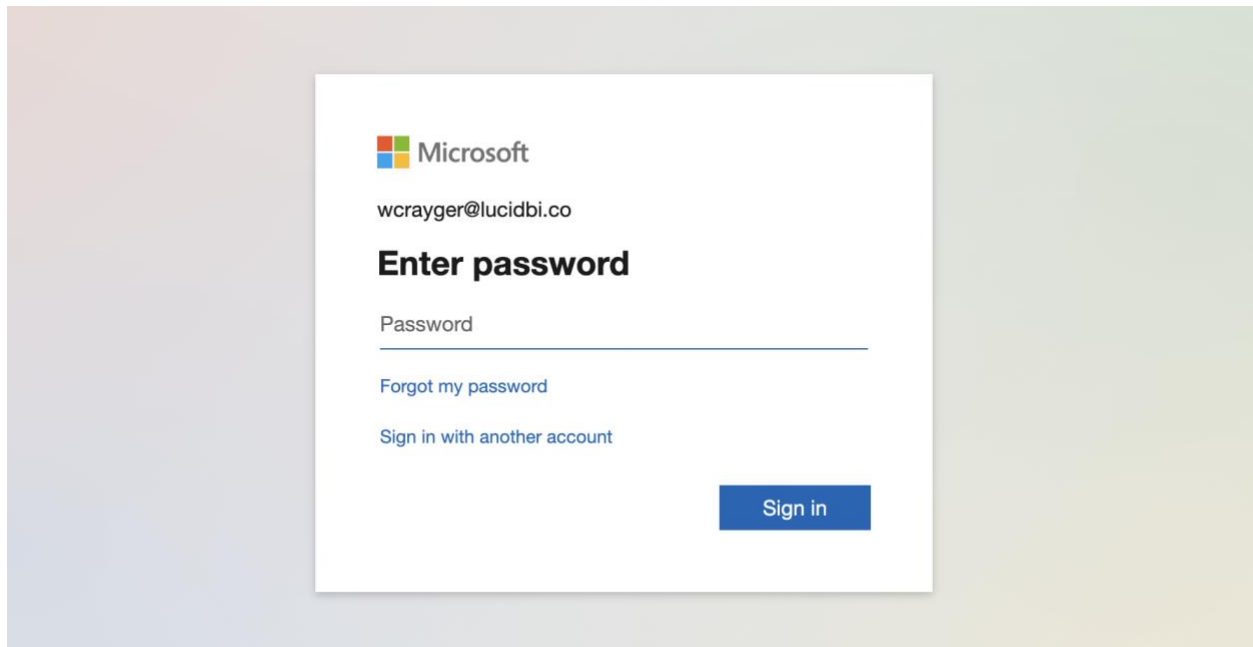
Introduction:

In this lab, you will create a Lakehouse in the Fabric workspace. A Lakehouse is essential because it serves as a centralized repository for all types of data, structured or unstructured. It enables efficient data management and analysis, forming the backbone of any data-driven operation. Go to home screen of your Fabric Workspace

Logging into Fabric/Power BI:

Authenticate into Fabric / Power BI: Navigate to app.powerbi.com and authenticate in with the credentials provided to you as part of the FIAD program.


The image shows the Power BI login interface. On the left, there is a white panel with the Power BI logo at the top. Below the logo, the text reads "Enter your email, we'll check if you need to create a new account." There is an email input field with the text "wcrayger@lucidbi.co" inside. Below the input field, there is a small text block stating: "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 this, another small text block states: "By clicking Submit, you agree to these [terms and conditions](#) and allow Power BI to get your user and tenant details. [Microsoft Privacy Statement](#)". At the bottom of the white panel is a yellow "Submit" button. The right side of the screen has a light gray background with three overlapping yellow and orange rounded rectangles of increasing size, creating a sense of depth.




Access the Fabric Workspace Home Screen: This is your starting point for building the Lakehouse. The home screen is where you'll navigate through different aspects of your Fabric coursework.


Creating Your Fabric Workspace:


Creating your training Workspace: Select **Workspaces** from the left-side navigation blade and choose **New Workspace**. Create a new Workspace using **fiad_<your_initials>** (e.g. **fiad_wtc**) as the naming pattern.


**Power BI**


My workspace


 Search



Home



Create



Browse



OneLake data hub


Apps



Workspaces




My workspace






Power BI



Workspaces



 Admin monitoring



 **My workspace** 



All



 Development Sandbox 


 Fabric in a Day - Zero to H... 


 Fabric Training Developme... 

 Grid Testing 

 Lucid 

 Lucid Capacity Monitor 

 Deployment pipelines



Power BI My workspace

Search

Trial: 45 days left

Home

Create

Browse

OneLake data hub

Apps

Workspaces

My workspace

My workspace

+ New item

Create a workspace

Name *

fiad_wtc

✓ This name is available

Description

Describe this workspace

Domain ⓘ

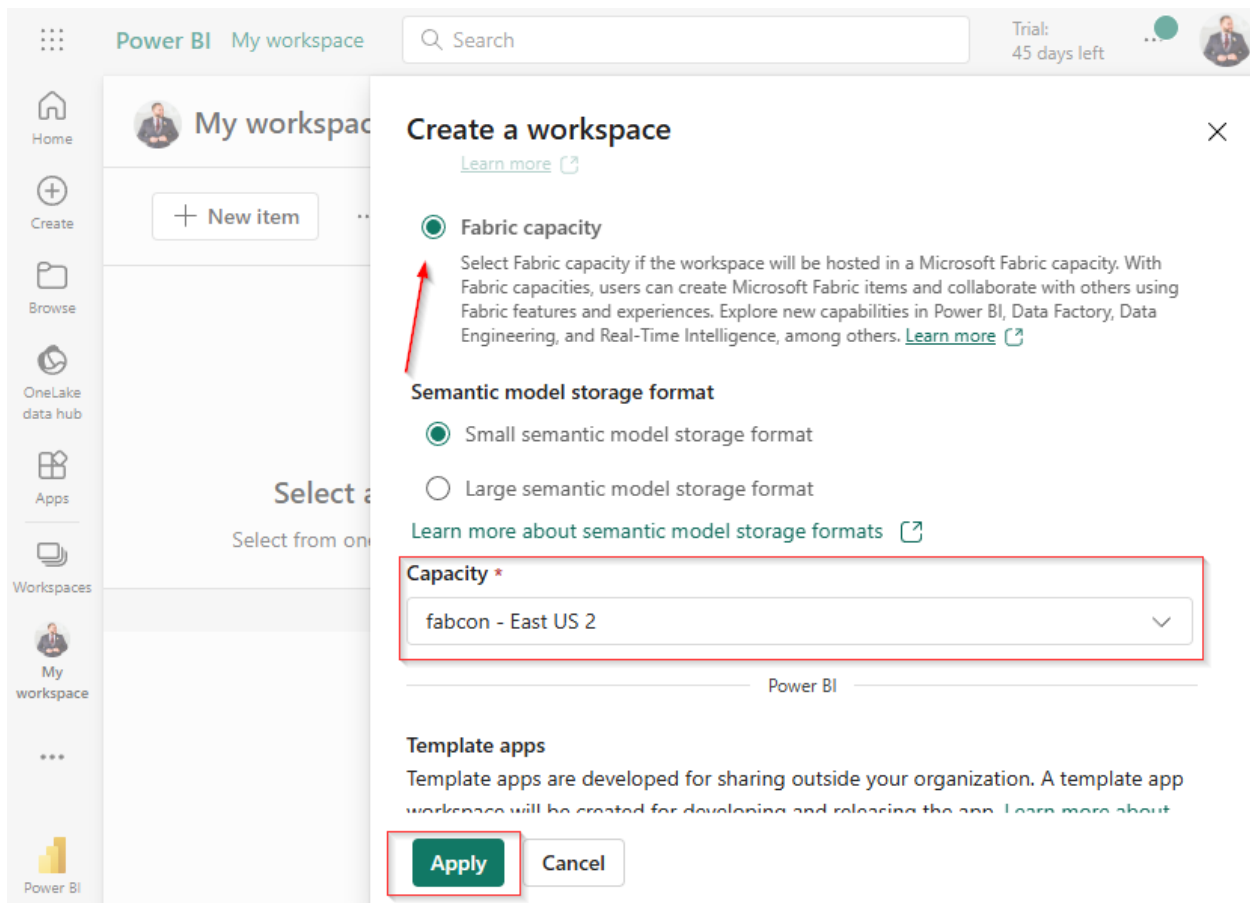
Assign to a domain (optional)

[Learn more about workspace settings](#)

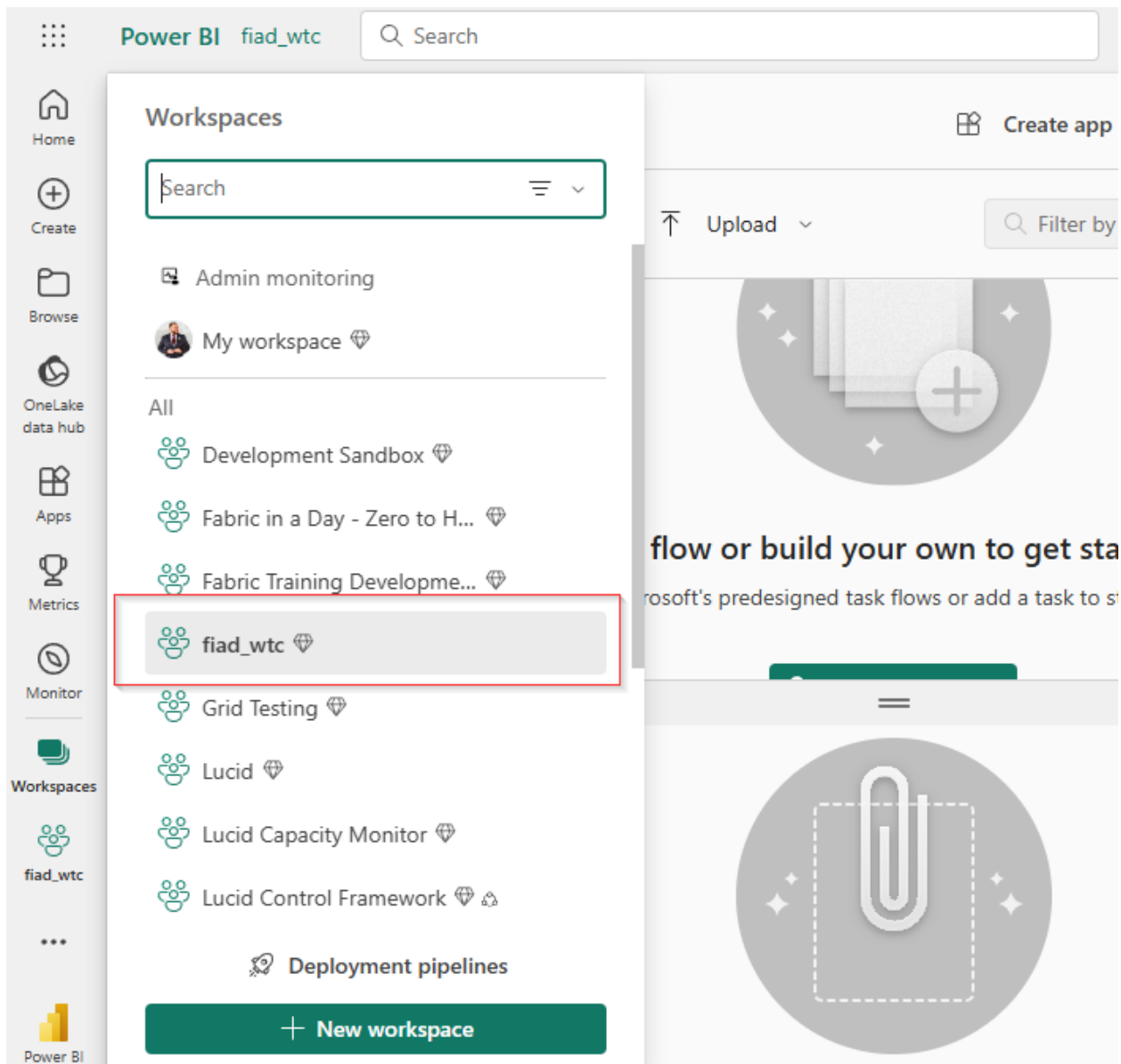
Workspace image

Upload

Assign your Workspace to Fabric Capacity: Expand the **Advanced settings** in the **Create a workspace** blade. Select the **Fabric capacity** radial button and choose the available capacity from the dropdown then click **Apply**.

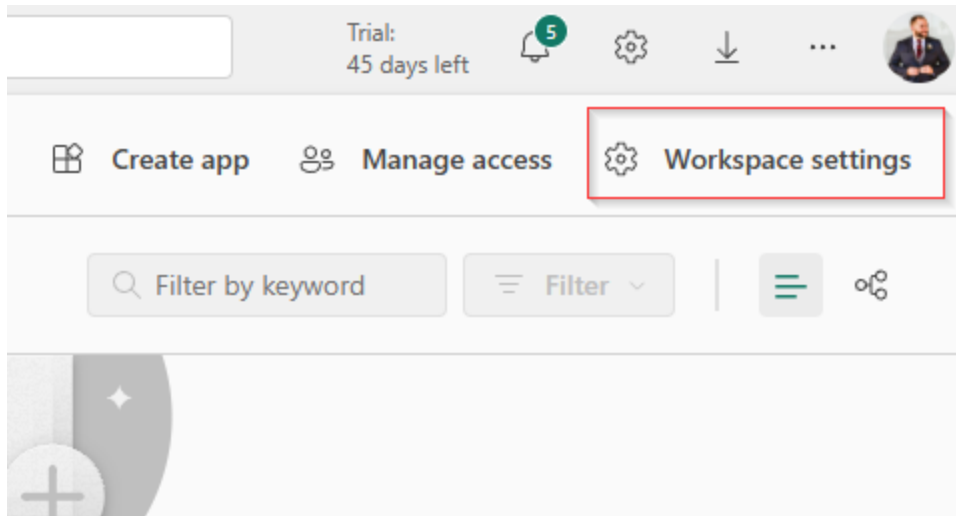


Selecting the Correct Workspace: Ensure that you're working in the correct Fabric Workspace. From the left navigate pane, click the **Workspaces** button and select the workspace created in the previous step. **Do not use the "My Workspace" environment.**



Configuring Workspace Settings:

Create a new Spark pool: Enter the **Workspace settings** from the top-right and open the **Data Engineering/Science** menu.



Workspace settings

General

- License info
- Azure connections
- System storage
- Git integration
- OneLake
- Workspace identity
- Network security

Power BI

Data
Engineering/Science

Spark settings

Data Factory

General

About

Workspace image



Upload

Delete

Name *

fiad_wtc

Description

Describe this workspace (Optional)

Domain

Assign this workspace to a relevant domain to help people discover the content inside it. Each workspace can be assigned to one domain.

Configure default spark pool: Select the **Default pool for workspace** dropdown and select **New pool**. Name your new pool **single_node**. Set the **node size** to **small**, disable **autoscale** and **dynamically allocate executors**, and reduce the number of nodes to 1 then click **Create**. Ensure the **single_node** pool is selected as the **Default**.

Workspace settings

- General
- License info
- Azure connections
- System storage
- Git integration
- OneLake
- Workspace identity
- Network security

Power BI

Data Engineering/Science

Spark settings

Spark settings

Configure and manage settings for Spark workloads and the default environment for the workspace.

Pool Environment High concurrency Automatic log

Default pool for workspace

Use the automatically created starter pool or create custom pools for workspaces and items in the capacity. If the setting Customize compute configurations for items is turned off, this pool will be used for all environments in this workspace.

StarterPool

Starter pool

✓ StarterPool
Node family: Memory optimized; Node size: Medium

New pool

Number of nodes
1 - 10

Customize compute configurations for items

On

Workspace settings

- General
- License info
- Azure connections
- System storage
- Git integration
- OneLake
- Workspace identity
- Network security

Power BI

Data Engineering/Science

Spark settings

Data Factory

Create new pool

Spark pool name *

single_node

Node family

Memory optimized

Node size

Small

Autoscale

If enabled, your Apache Spark pool will automatically scale up and down based on the amount of activity.

☐ Enable autoscale

1

Dynamically allocate executors


☐ Enable dynamic allocation

Create

Cancel

Workspace settings

- General
- License info
- Azure connections
- System storage
- Git integration
- OneLake
- Workspace identity
- Network security
- Power BI
- Data Engineering/Science
- Spark settings

 This section contains unsaved changes.

Spark settings

Configure and manage settings for Spark workloads and the default environment for the workspace.

Pool

Environment



High concurrency

Automatic log

Default pool for workspace

Use the automatically created starter pool or create custom pools for workspaces and items in the capacity. If the setting Customize compute configurations for items is turned off, this pool will be used for all environments in this workspace.

single_node

Pool details				
Node family	Node size	Number of nodes		
Memory optimized	Small	1		

Disable High Concurrency mode: Under the **High Concurrency** tab, toggle off **For notebooks** and click **Save**.

Workspace settings



- General
- License info
- Azure connections
- System storage
- Git integration
- OneLake
- Workspace identity
- Network security

- Power BI
- Data Engineering/Science
- Spark settings**
- Data Factory

⚠ This section contains unsaved changes.

Spark settings

Configure and manage settings for Spark workloads and the default environment for the workspace.

Pool Environment **High concurrency** Automatic log

For notebooks

When high concurrency for notebooks is on, multiple notebooks can use the same Spark application to reduce the start time for each session. [Learn more about running notebooks in high concurrency mode](#)

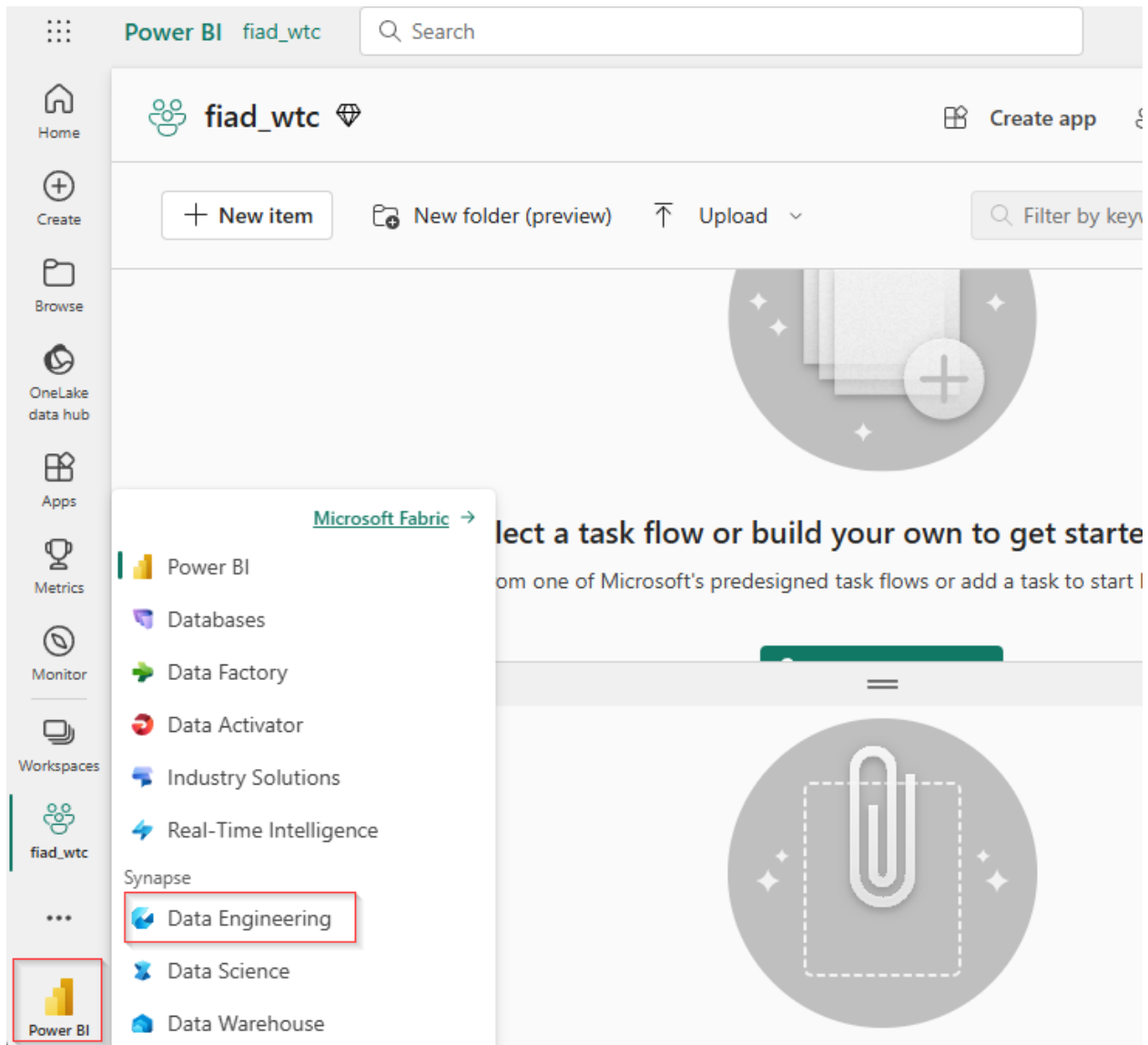
☐ Off

Save

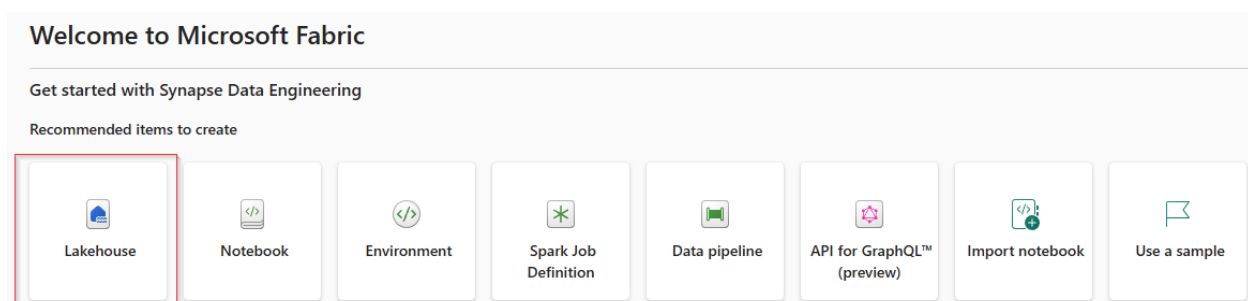
Discard

Create Bronze Lakehouse:

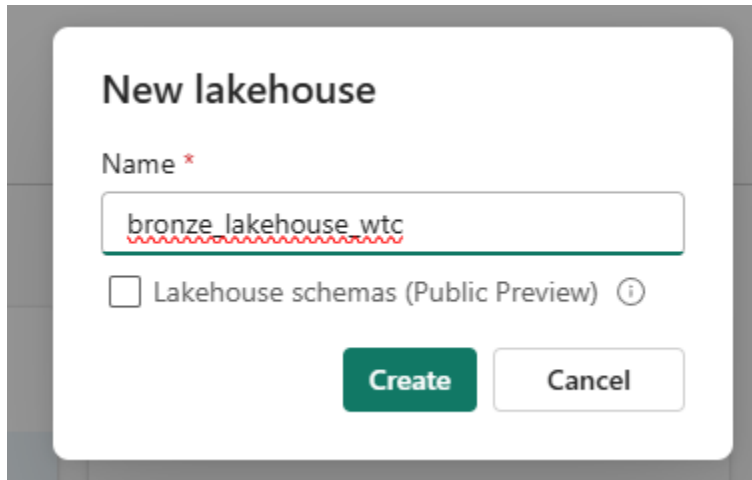
Navigate to the Data Engineering Experience: Use the experience toggle in the bottom left corner to access the data engineering tools. The experience toggle can be used to access persona specific menus.



Select Lakehouse: By choosing **Lakehouse**, you initiate the process of setting up your data repository. This is an important step in establishing a central location for data storage and analysis.



Name Your Lakehouse: Personalize your workspace by naming your Lakehouse **bronze_lakehouse_{your initials}**. Leave Lakehouse schemas disabled and click **Create**.



New lakehouse

Name *

bronze_lakehouse_wtc

☐ Lakehouse schemas (Public Preview) ⓘ

Create **Cancel**

Uploading a CSV File – Creating Your First Delta Table

Introduction:

Now that our Lakehouse has been created, we will hydrate the lake with our first batch of data. To do so, we will be using Fabric's Lakehouse ability of uploading a CSV file and converting it to a delta table in the lake.

Load CSV and Create Delta Table:

Load the sample file to Lakehouse File Container: Navigate to your Lakehouse and right-click the **Files** section. Hover over **Upload** and select **Upload files**. Click the folder icon in the **Upload files** blade that opens on the right side of the window. Navigate to the location of the saved **package_types.csv** file from the course material, select the file, and click **Open**. Click **Upload** and wait for the file to be uploaded to the Lakehouse.

Home

Get data ▾ New semantic model Open notebook ▾ Manage OneLake data access (preview)

A SQL analytics endpoint for SQL querying and a default Power BI semantic model for reporting were created with this item.

Explorer

- bronze_lakehouse_wtc
 - Tables
 - Files
 - New shortcut
 - New subfolder
 - Upload
 - Upload files
 - Upload folder
 - Properties
 - Refresh

Workspaces

- fiad_wtc
- bronze_lakehouse_wtc

Real-Time hub

OneLake data hub

Monitor

Create

Home

Browser

Up arrow icon

← → ▾ ↑ Training and Education > Fabric > Zero-To-Hero-with-Fabric > Labs > Lab 1 and Lab 2 - Creating a Lakehouse

Organize ▾ New folder

	Name	Status	Date modified
> Apps			
> Attachments			
> Documents			
> Documents - Lucid Asset Management			
> Documents - Lucid Consulting Group			
> Documents - Lucid Holdings			
	Fabric in a Day Lab Instructions - Lab 1 a...	🔄	9/8/2024 3:07 PM
	package_types	✅	9/7/2024 10:08 PM

Upload files



Files/

package_types.csv

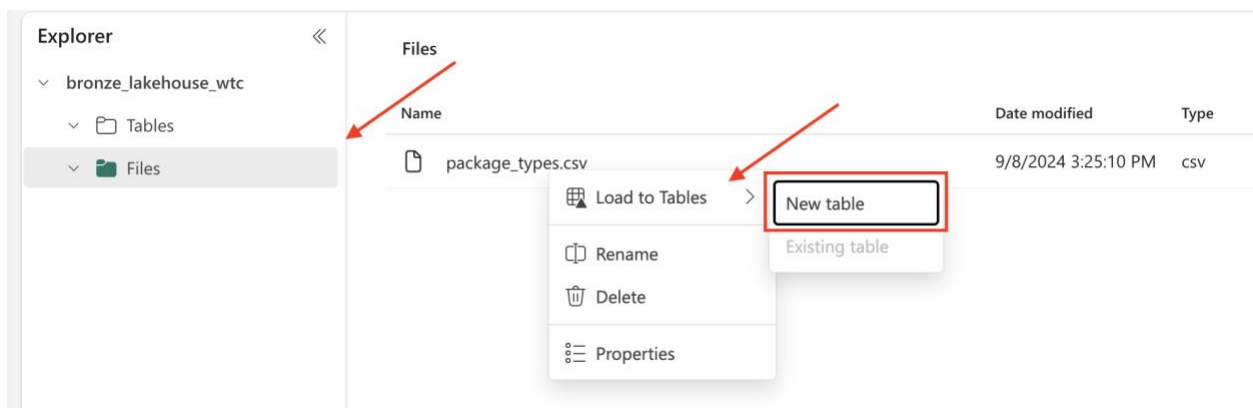


☐ Overwrite if files already exist

Upload

Create a Delta Table:

Load CSV to Lakehouse Delta Table: Navigate back to the **Files** section of the Lakehouse explorer. Right-click the **package_types.csv** file, select **Load to Tables**, then **New Table**. Name the new table **package_types**, specify file has column headers, and it's a **“,”** separated file, and click **Load**.



Load file to new table

All fields marked with * are required

New table name *

package_types

Names can include letters, numbers, spaces, and underscores (_).

Column header ⓘ

☒ Use header for column names

Separator ⓘ

,

Separators cannot use the following characters: () [] { } ' " "

Load **Cancel**

Verify Table Creation: Navigate back to the Table section of the Lakehouse and confirm the table was created. It may take a few minutes for the table to create as a spark session is being created to perform the operation.

Note: It may take 60-90 seconds for the table to load due to background processes. Please be patient and do not try to create the table multiple times.

Explorer

«

▼ bronze_lakehouse_wtc

▼ Tables

▶ package_types ...

▼ Files

package_types

	123	PackageTypeID	ABC PackageTypeName	123 LastEditedBy	ValidFrom	ValidTo
1	1		Bag	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
2	2		Block	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
3	3		Bottle	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
4	4		Box	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
5	5		Can	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
6	6		Carton	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
7	7		Each	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
8	8		Kg	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
9	9		Packet	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
10	10		Pair	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
11	11		Pallet	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
12	12		Tray	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
13	13		Tub	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM
14	14		Tube	1	1/1/2013 12:00:00 AM	12/31/9999 11:59:59 PM

Note: If at any point you see “Undefined” instead of a table name in your Lakehouse please refresh your browser. The Fabric UI has not registered the Delta table being created yet and a refresh should resolve the issue.

This lab is now complete.