SQL Saturday 1058 Portland/Vancouver

November 2nd 2024

Fabric Semantic Link Labs:

A Link to the Future

SQL Saturday Oregon & SW Washington

November 2024

Jason Romans

Senior BI Engineer
Builder of Models

The Dax Shepherd

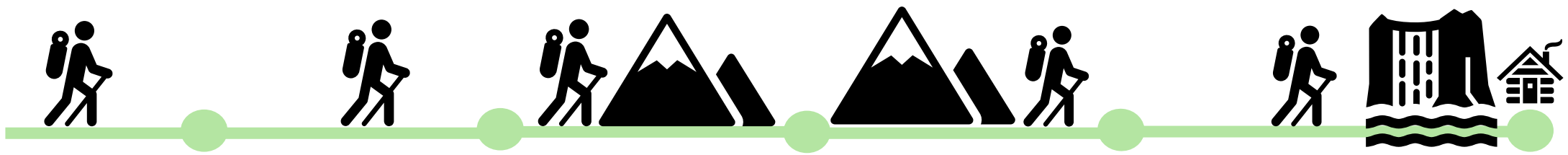
redgate COMMUNITY AMBASSADOR 2024

- Lives in Nashville, Tennessee, United States
- Started as SQL Server DBA
- Transitioned to the Microsoft BI Stack
- SSIS, SSAS, SQL Server Database, and Power BI
- Simple Talk Author at Redgate
- Favorite Data Model


SQL Saturday
2024
Sponsors
Microsoft
AMD
OMNIDATA
redgate
three tree tech
rev.io
NORTH BANK INNOVATIONS
INTERJECT
red9
PURESTORAGE®

Shoulders of Giants

Our Journey



1. Semantic Link
2. Semantic Link Labs
3. Direct Lake Migration
4. Additional Uses
5. Conclusion

Our Journey

1. Semantic Link

2. Semantic Link Labs

3. Direct Lake Migration

4. Additional Uses

5. Conclusion

Missing Link

Semantic link is a feature that allows you to establish a connection between semantic models and Synapse Data Science in Microsoft Fabric.

-- Reference: https://learn.microsoft.com/en-us/fabric/data-science/semantic-link-overview

All the Things

Semantic link is a feature that allows you to establish a connection between semantic models and Synapse Data Science in Microsoft Fabric.

semantic models

Models, Reports, Lakehouse, Workspaces and more

Synapse Data Science

Fabric Notebook – Apache Spark with Python and more

A Tale of Two Links

Both are Available Only in Microsoft Fabric

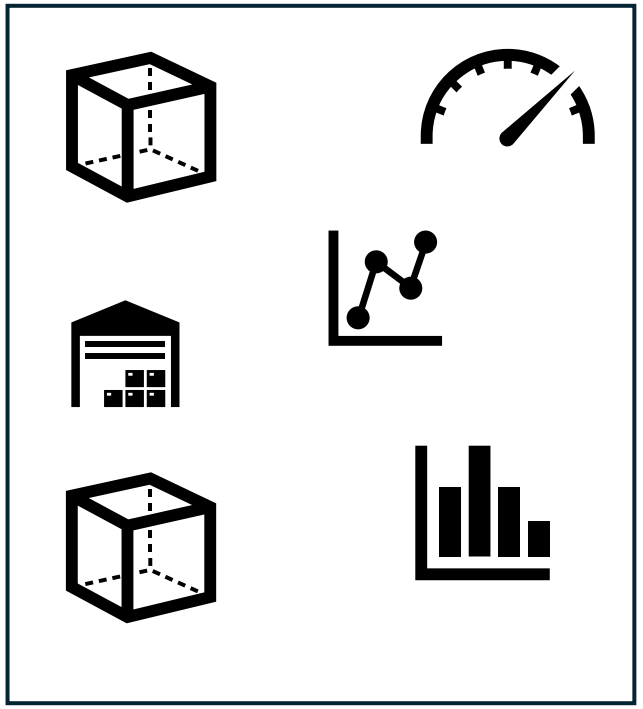
- Semantic Link
  - Base
  - Driver or API
  - Included in default runtime for the current version

A Tale of Two Links

- Semantic Link Labs
  - "Expansion Pack" -- Kurt Buhler
  - Uses Semantic Link
    - `import sempy.fabric as fabric`
  - Open source – GitHub Repository
  - Under Active Development

Semantic Link Labs

Microsoft Fabric



Semantic Link

List Tables

List Workspaces

List Models

List Reports

Semantic Link Labs

Migrate to Direct Lake

Best Practice Analyzer

Manage Reports

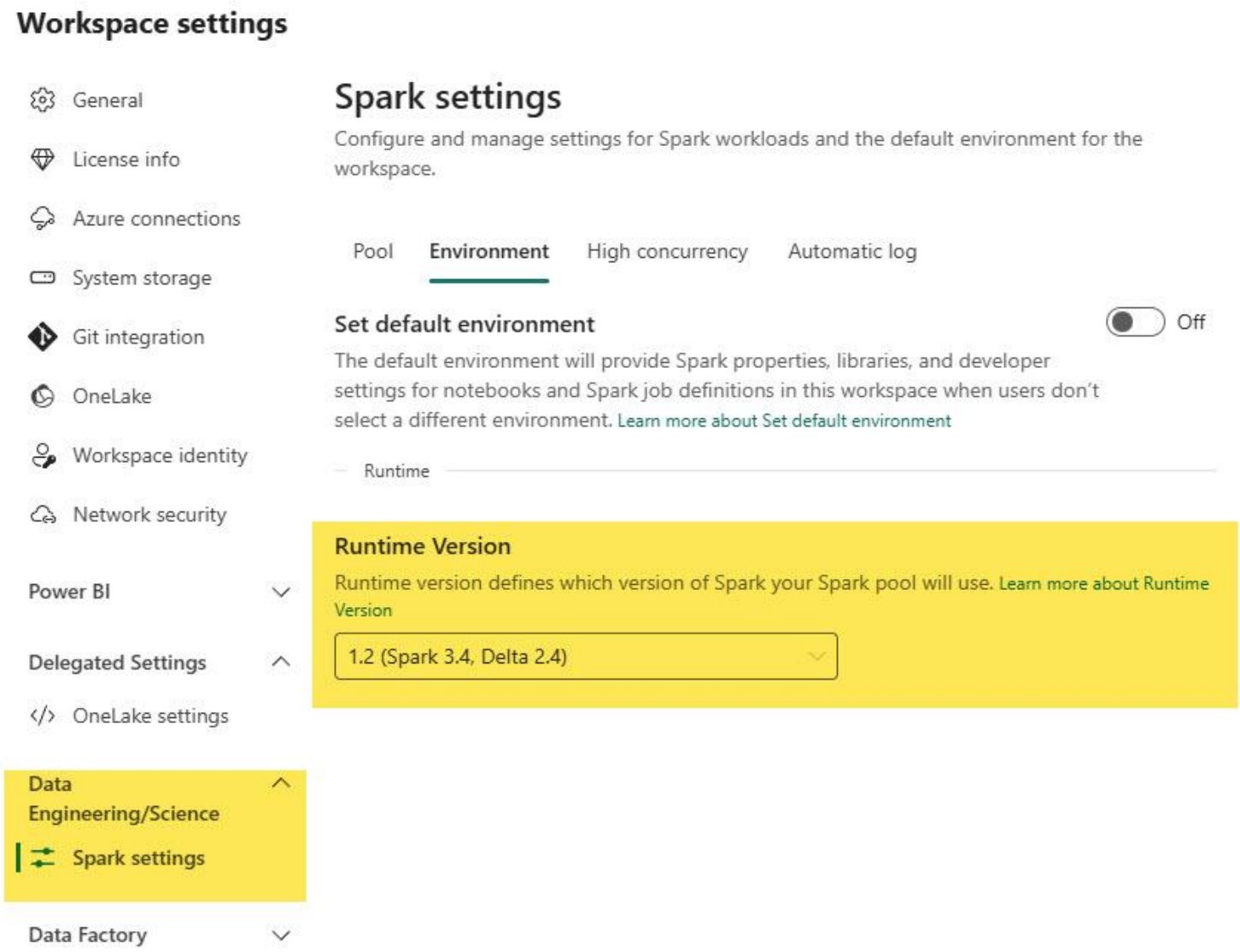
View Broken Reports

Rebind Reports

Installing

- Semantic Link
  - Spark 3.4 in default runtime
    - Update to newest version
      - %pip install –U semantic-link
  - Spark 3.3 or below – need to install
    - %pip install –U semantic-link

What Version of Spark



Spark Version Command

spark.version

```
1    spark.version
```

✓ <1 sec - Command executed in 260 ms by Jason Romans on 2:00:03 PM, 10/08/24

```
'3.4.3.5.3.20240904.5'
```

What Version of Semantic Link

- %pip show semantic-link

```
1  %pip show semantic-link
```

[23] ✓ 4 sec - Command executed in 3 sec 588 ms by Jason Romans on 11:49:10 PM, 10/19/24 — PySpark (Python)

```
Name: semantic-link
Version: 0.8.1
Summary: Semantic link for Microsoft Fabric
Home-page: https://learn.microsoft.com/en-us/fabric/data-science/semantic-link-overview
Author: Microsoft
Author-email: semanticdatascience@service.microsoft.com
License: proprietary and confidential
Location: /nfs4/pyenv-dd0ba783-2069-4fdc-8c80-bc03c74db705/lib/python3.11/site-packages
Requires: semantic-link-functions-geopandas, semantic-link-functions-holidays, semantic-link-functions-meteostat, semantic-link-functions-phonenumbers, semantic-link-functions-validators, semantic-link-sempy
Required-by:
Note: you may need to restart the kernel to use updated packages.
```

Import the Module

```
# Give it a friendly name – think Alias
import sempy.fabric as fabric
```

Done in a Notebook

Text Descriptions - Markdown

Code - Python

Install the latest .whl package

Check here to see the latest version.

```
1  %pip install semantic-link-labs
```

- Session ready in 8 sec 603 ms. Command executed in 26 sec 202 ms by Jason Romans on 10:29:04 AM, 10/23/24

Install Wheel from File

```
1  %pip install /lakehouse/default/Files/semantic_link_labs-0.8.3-py3-none-any.whl
```

- Command executed in 22 sec 745 ms by Jason Romans on 10:07:06 AM, 10/23/24

Show Semantic Link and Labs installed

```
1  %pip show semantic-link-sempy
2  print('\n')
3  %pip show semantic-link-labs
```

✓ - Command executed in 6 sec 382 ms by Jason Romans on 10:30:20 AM, 10/23/24

```
Name: semantic-link-sempy
Version: 0.8.1
Summary: Semantic link for Microsoft Fabric
Home-page: https://learn.microsoft.com/en-us/fabric/data-science/semantic-link-overview
Author: Microsoft
```

Uses Pandas Ecosystem (DataFrame)

Knowledge of working with Pandas DataFrame helpful

DataFrame – data table

Can convert Pandas -> Polars

List Semantic Models

```
import sempy.fabric as fabric
fabric.list_datasets()
```

Dataset = Semantic Model

	Dataset Name	Dataset ID	Created Timestamp	Last Update
0	Contoso10K	47c34560-ef4d-46c6-825e-20cb9f11ba9d	2023-05-04 14:36:12	NaT
1	FabSLL_Lakehouse	e6d18d3a-b407-4fb3-813a-418b76388b11	2021-02-12 23:00:58	NaT
2	DataflowsStagingLakehouse	7c76f16d-2364-4c33-89bb-6960ac29cb5d	2021-02-12 23:00:58	NaT
3	DataflowsStagingWarehouse	625cb1b9-de7e-425f-90b3-727ac87268db	2021-02-12 23:00:58	NaT
4	Contoso10K_DL	261101fb-fc2d-4511-be34-1def1b4530fe	2019-09-17 05:50:29	NaT
5	Contoso10K_DLL1	40b52877-7e62-43cd-9a38-7c09b62f9048	2019-09-17 05:50:29	NaT
6	Contoso10K_ABC	cc93c0f1-e112-4adc-a009-3fb3a7a12f3b	2019-09-17 05:50:29	NaT
7	Contoso10K_SLL	e45a6ce5-5ea5-4b87-ad88-136dc4cabc27	2019-09-17 05:50:29	NaT
8	Contoso10K_SL1	421040eb-ba74-4304-bb92-708c1667eb61	2019-09-17 05:50:29	NaT
9	Contoso10K_SL2	b84b790d-04a2-406d-ba17-2191186ebafa	2019-09-17 05:50:29	NaT
10	Contoso10K_SL9	5ca86b73-bef1-43d6-af09-5c12e358d391	2019-09-17 05:50:29	NaT

List Tables

```
tables = fabric.list_tables(workspace="SQLMAB", dataset = "SQLMab")
display(tables)
```

	Name	Description	Hidden	Data Category	Type
0	Customer		False		Table
1	Sales		False		Table
2	Date		False		Table
3	Store		False		Table
4	Product		False		Table

List Workspaces

```
1  fabric.list_workspaces()
```

✓ - Command executed in 806 ms by Jason Romans on 5:25:32 PM, 10/08/24 PySpark (Python)

	Id	Is Read Only	Is On Dedicated Capacity	Capacity Id	Default Dataset Storage Format	Type	Name
0	bfab8dff-bdfc-4943-9996-7dc97a4e4d38	False	False	NaN	NaN	Workspace	JAXSQL2023
1	41e69008-f6c3-42c0-8c0a-739f1a7a9a0a	False	True	3b9ac229-bcbb-4aa4-8543-72b6db25e330	Small	Workspace	ONSQL2023
2	10c0ad3f-7a0b-4c0e-8b7d-f4a6170c5dde	False	True	3b9ac229-bcbb-4aa4-8543-72b6db25e330	Small	Workspace	SOFLSQL2023
3	ee7a8d30-f109-4848-9db3-be711a0f24d4	False	False	NaN	NaN	AdminInsights	Admin monitoring
4	10c2bd0a-5d3d-4e77-883c-f16af027fcd5	False	True	3b9ac229-bcbb-4aa4-8543-72b6db25e330	Small	Workspace	COLSQLSAT2023
5	365a3880-83f7-4014-bad3-a006c34e2bb1	False	True	3b9ac229-bcbb-4aa4-8543-72b6db25e330	Small	Workspace	BRSQLSAT2023
6	e6e71dc2-2a8e-482a-8bc9-ff5c16ed0311	False	True	3b9ac229-bcbb-4aa4-8543-72b6db25e330	Small	Workspace	DENSQLSAT2023

List Workspaces on Dedicated Capacity

```
1   ws = fabric.list_workspaces()
2   dedicated = ws[ws["Is On Dedicated Capacity"] == True]
3   display(dedicated)
```

✓ - Command executed in 790 ms by Jason Romans on 5:37:07 PM, 10/08/24 PySpark (Python)

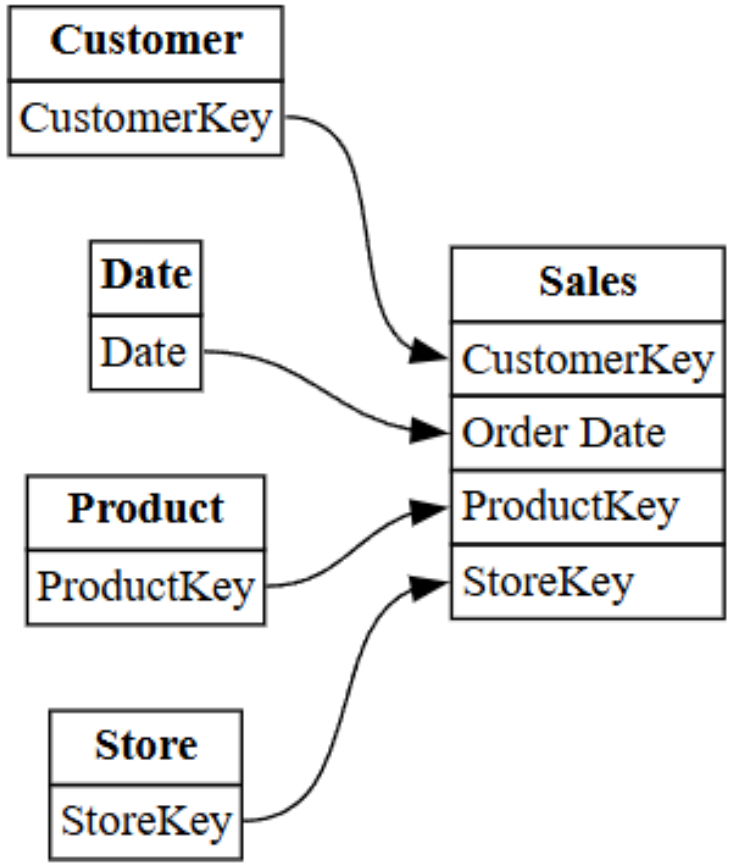
Table | Chart | Download | Showing rows 1 - 22 | Inspect | Search

	Id	Is Read Only	Is On Dedicated Capacity	Capacity Id	Default Dataset Storage Format	Type	Name
1	41e69008-f...	false	true	3b9ac229-bcbb...	Small	Workspace	ONSQL2023
2	10c0ad3f-7...	false	true	3b9ac229-bcbb...	Small	Workspace	SOFLSQL2023
3	10c2bd0a-...	false	true	3b9ac229-bcbb...	Small	Workspace	COLSQLSAT2023
4	365a3880-...	false	true	3b9ac229-bcbb...	Small	Workspace	BRSQLSAT2023
5	e6e71dc2-...	false	true	3b9ac229-bcbb...	Small	Workspace	DENSQLSAT2023
6	782b40a8-...	false	true	3b9ac229-bcbb...	Small	Workspace	SQLMAB
7	62e33fa7-f...	false	true	3b9ac229-bcbb...	Small	Workspace	SQLSAT_Denver2023
8	9a5d4f1a-b...	false	true	3b9ac229-bcbb...	Small	Workspace	MNSQLSAT2023
9	ba1e683d-...	false	true	3b9ac229-bcbb...	Small	Workspace	ORLANDOSQLSAT2023
10	be75ac57-...	false	true	3b9ac229-bcbb...	Small	Workspace	SV-SQLSat2023
11	4d139fc8-0...	false	true	3b9ac229-bcbb...	Small	Workspace	ORWA23_SQLSAT
12	23c86de6-...	false	true	3b9ac229-bcbb...	Large	Workspace	TE Training
13	ad0afb41-...	false	true	3b9ac229-bcbb...	Small	Workspace	PremTest
14	2e6ae09d-...	false	true	3b9ac229-bcbb...	Large	Workspace	SQLBits2024
15	5f6564c6-1...	false	true	27046857-fc30-...	Small	Workspace	MetaDriven-2

Selected Cell 14 of 21 ce

Relationships

```
from sempy.relationships import plot_relationship_metadata as prm
prm(fabric.list_relationships(workspace="SQLMAB", dataset = "SQLMab"))
```



Our Journey

1. Semantic Link
2. **Semantic Link Labs**
3. Direct Lake Migration
4. Additional Uses
5. Conclusion

In the Notebook:
Install the Expansion Pack (DLC)

Install the Package

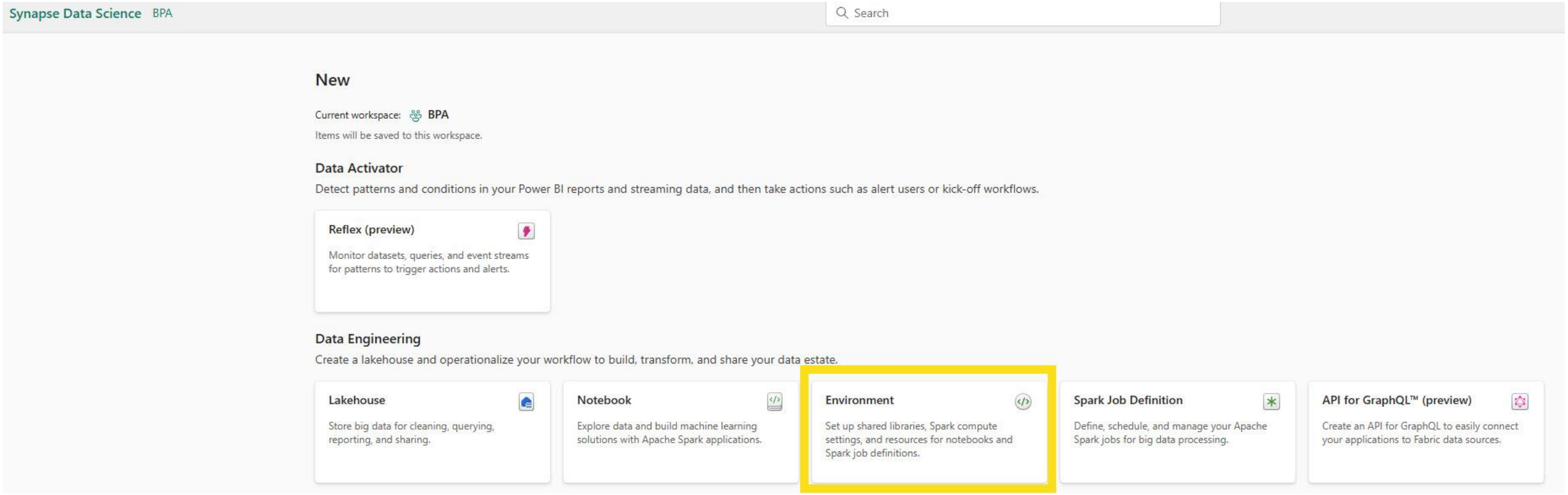
```
# Install Semantic Link Labs
%pip install semantic-link-labs


# Import module with a shorter name
# Easier to Type
import sempy_labs as labs
```

Expands Semantic Link

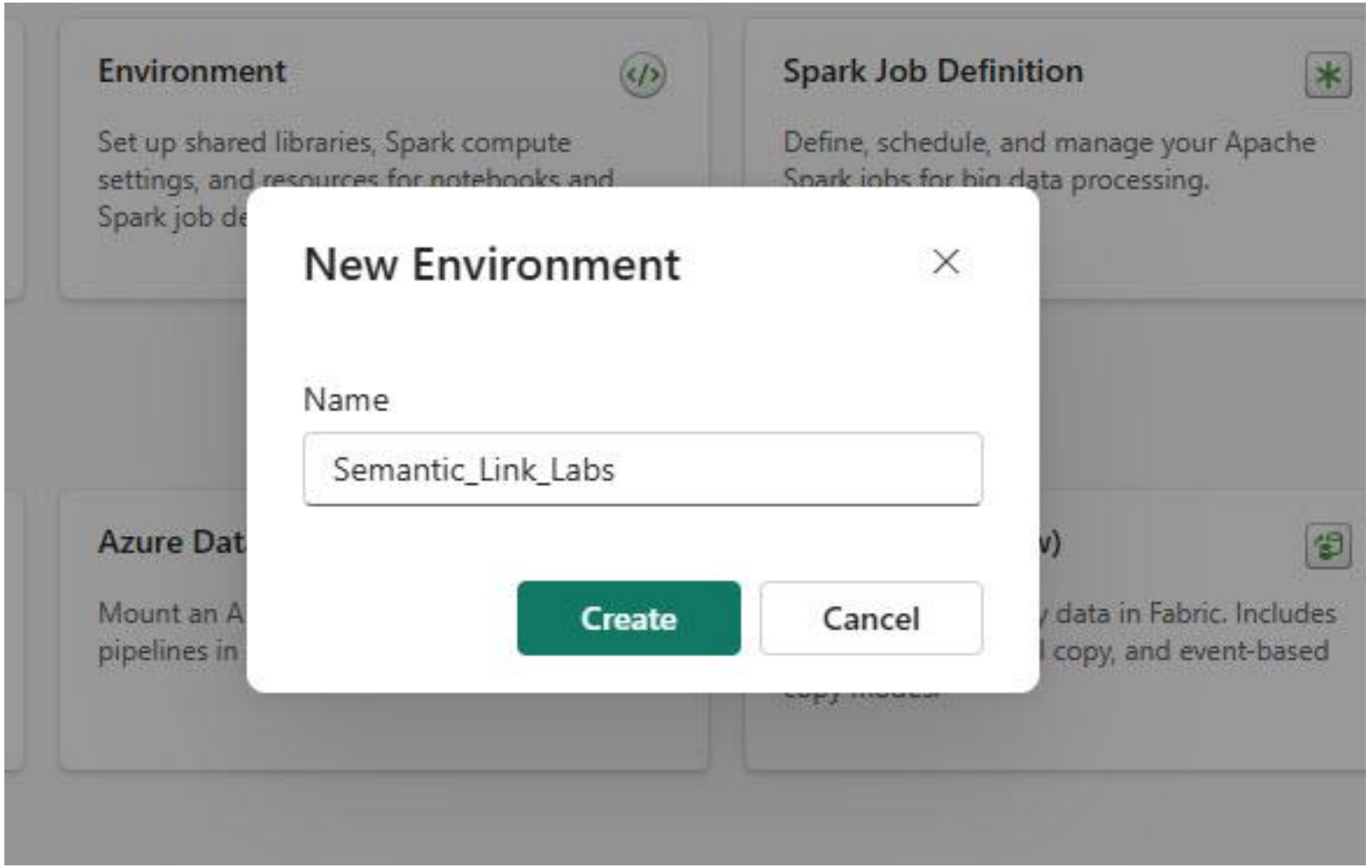
- Does not replace or overwrite
- Have both available
- Same as if imported Polars or other packages

In the Environment:
Install the Expansion Pack (DLC)

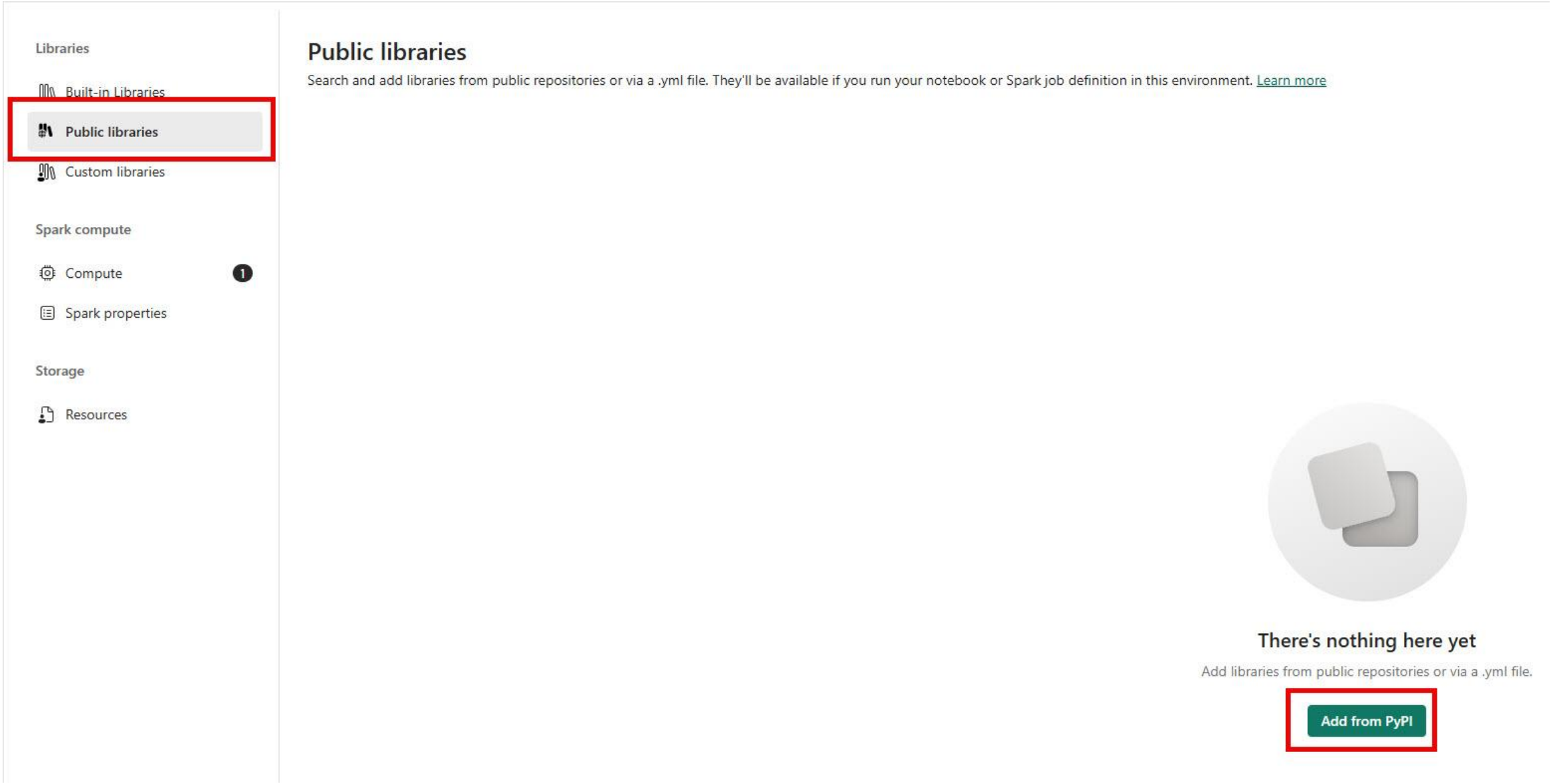
Create Environment (Workspace)



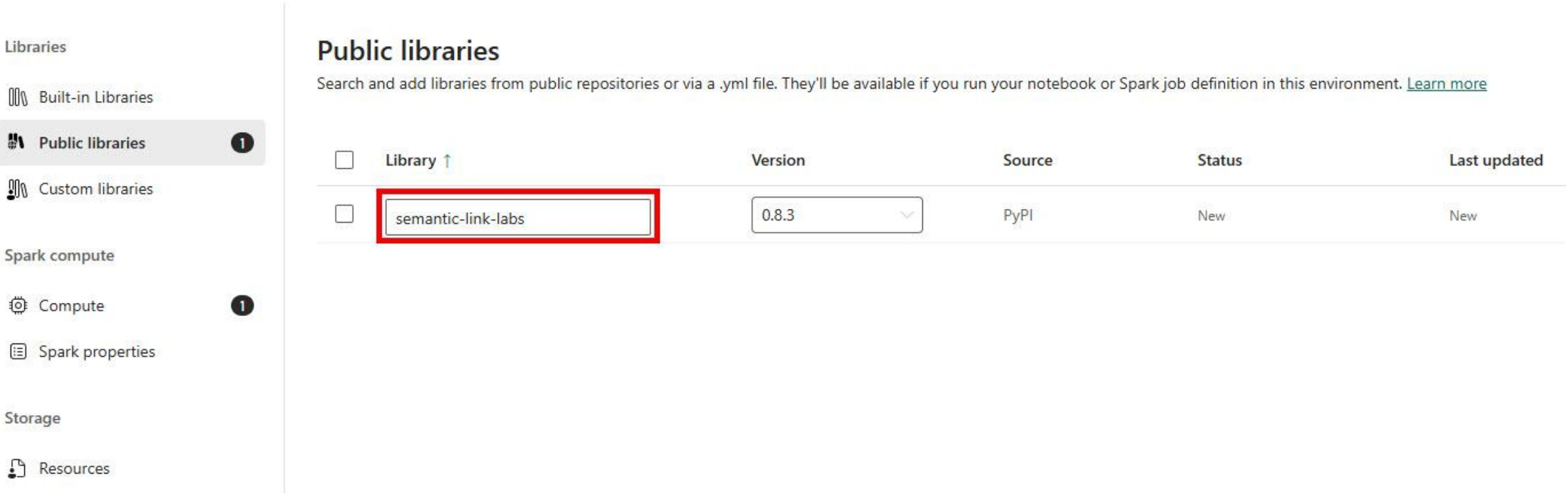
Create Environment



Create Environment



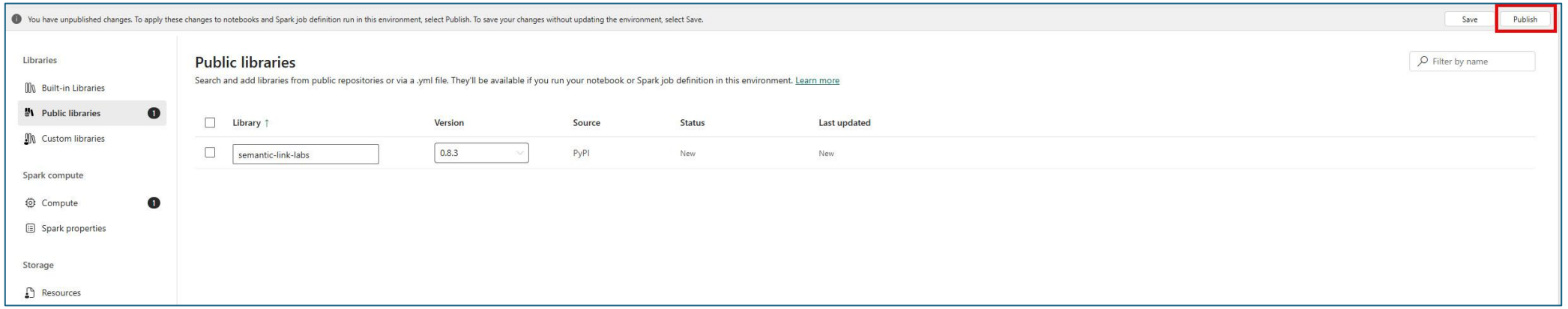
Create Environment



Specific Versions

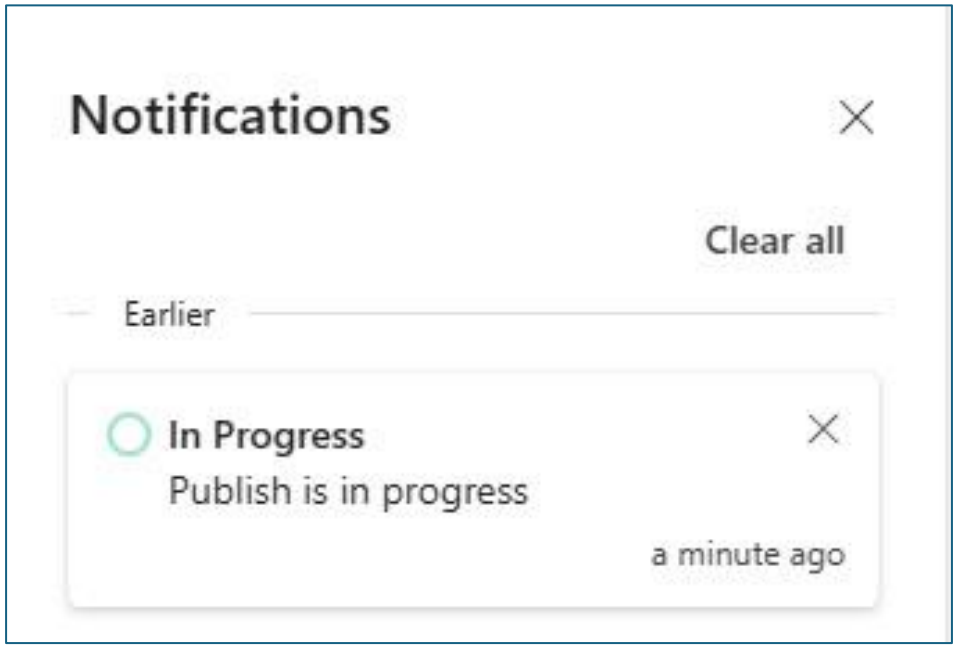
- Allows you to develop with specific versions
- Move to newer version when ready

Publish (Important)

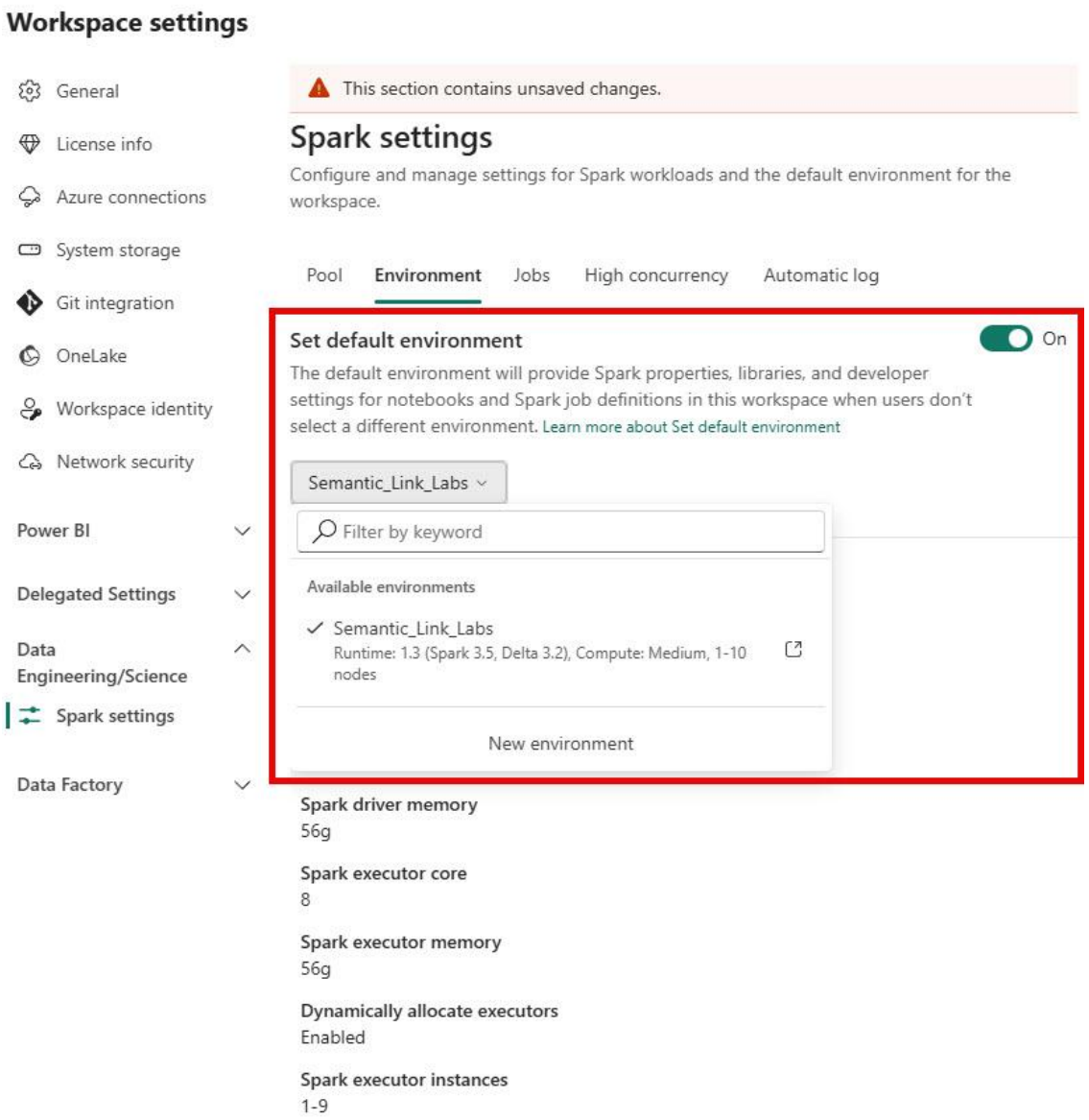


Environment takes time to publish

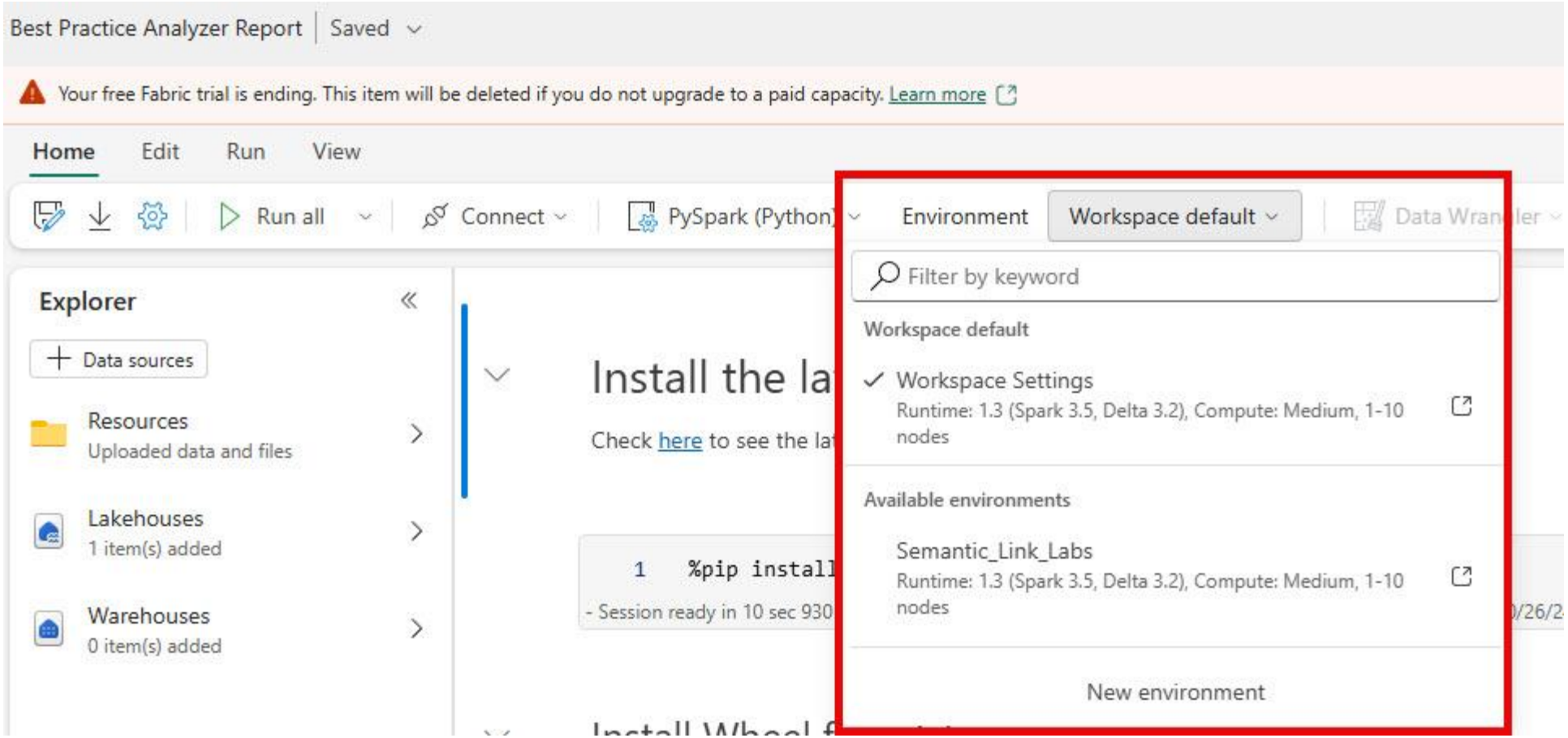
(Go watch the movie The Notebook while waiting)



Workspace Settings - Environment



Notebook Setting – Environment



Customize

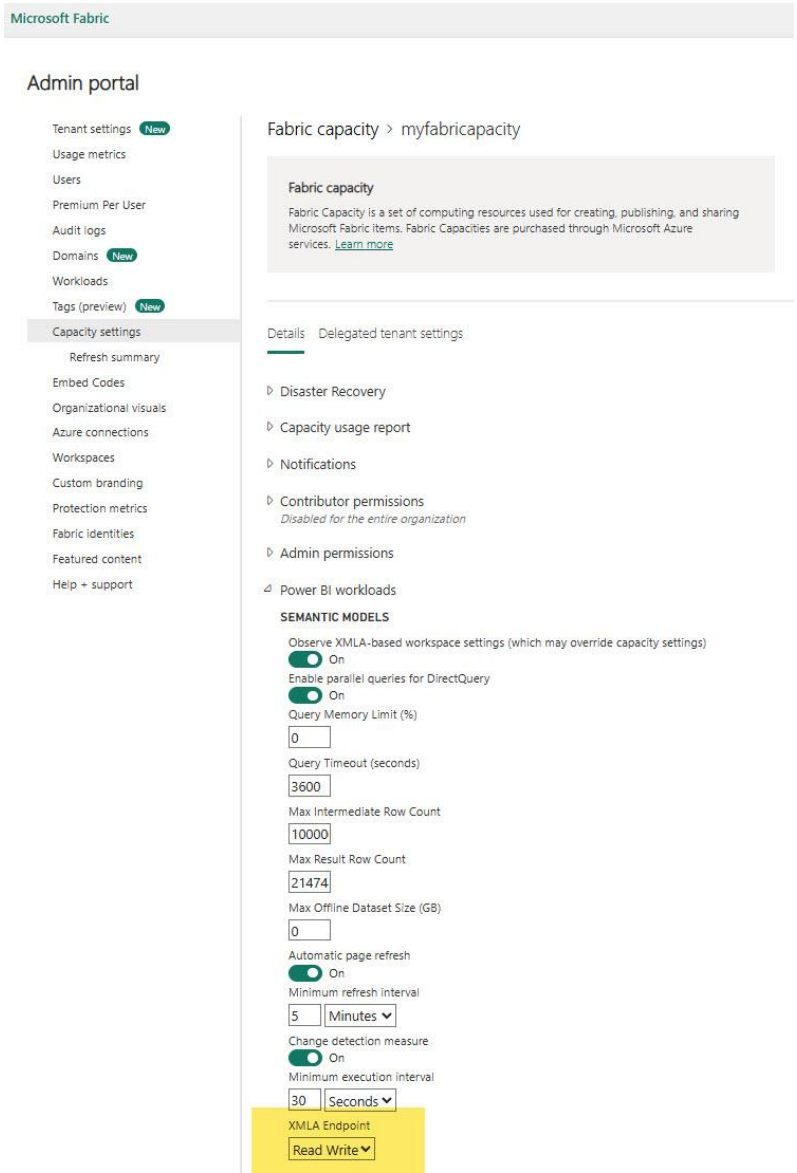
- Copy and paste Python code then modify it
- You can fork the repo (Great 80's Band Name)
  - Add Modifications
  - Build Wheel file
    - Either:
      - Install inline into Notebook
      - Install into Environment

Our Journey

1. Semantic Link
2. Semantic Link Labs
3. **Direct Lake Migration**
4. Additional Uses
5. Conclusion

Beyond the Scope

Enable XMLA


Microsoft Fabric
Admin portal
Tenant settings New
Usage metrics
Users
Premium Per User
Audit logs
Domains New
Workloads
Tags (preview) New
Capacity settings
Refresh summary
Embed Codes
Organizational visuals
Azure connections
Workspaces
Custom branding
Protection metrics
Fabric identities
Featured content
Help + support
Fabric capacity > myfabricapacity
Fabric capacity
Fabric Capacity is a set of computing resources used for creating, publishing, and sharing Microsoft Fabric items. Fabric Capacities are purchased through Microsoft Azure services. Learn more
Details
Delegated tenant settings
Disaster Recovery
Capacity usage report
Notifications
Contributor permissions
Disabled for the entire organization
Admin permissions
Power BI workloads
SEMANTIC MODELS
Observe XMLA-based workspace settings (which may override capacity settings)
On
Enable parallel queries for DirectQuery
On
Query Memory Limit (%)
0
Query Timeout (seconds)
3600
Max Intermediate Row Count
10000
Max Result Row Count
21474
Max Offline Dataset Size (GB)
0
Automatic page refresh
On
Minimum refresh interval
5
Minutes
Change detection measure
On
Minimum execution interval
30
Seconds
XMLA Endpoint
Read Write

Convert Import Mode to Direct Lake

1. Install the Package Semantic Link Labs
2. Import the Module
3. Need a LakeHouse
4. Set Parameters – Dataset, New Dataset, etc
5. Create Power Query Template file
6. Use Template file to create a Dataflows Gen2
7. Create Direct Lake Model based on the import/DQ model

Limitations – Direct Lake

- Calculated Tables are not supported…
- Calculated columns are not supported. Columns of binary data type are not supported.
- The columns used for the relationship cannot be of the datetime data type, and they must also be of the same data type.

Learn more about Direct Lake limitations here: https://learn.microsoft.com/power-bi/enterprise/directlake-overview#known-issues-and-limitations

Demo: Migrate to the Lake

Our Journey

1. Semantic Link
2. Semantic Link Labs
3. Direct Lake Migration
4. **Additional Uses**
5. Conclusion

Current Available Notebooks

- **Best Practice Analyzer Report**
  - Run BPA rules against one or more models
- **Capacity Migration**
  - Migrate P SKU -> F SKU
- **Migration to Direct Lake**
  - Import/Direct Query to Direct Lake
- **Model Optimization**
  - Vertipaq Analyzer

Current Available Notebooks

- **Query Scale Out**
  - Enable/Disable large semantic model format
- **SQL**
  - Run a SQL query (or queries) against a Fabric warehouse
- **Semantic Model Refresh**
  - Visualize the refresh of a semantic model
- **Tabular Object Model**
  - Add measure(s) to the semantic model

Best Practice Analyzer

- Collect stats for all semantic models within all accessible workspaces
  - labs.run_model_bpa_bulk(workspace='FabSLL’)
- Create a Direct Lake semantic model (called 'ModelBPA') for analyzing the Best Practice Analyzer results
  - labs.create_model_bpa_semantic_model()
- Create a Power BI report called 'ModelBPA' based semantic model created in the previous cell, which can be used to analyze the Best Practice Analyzer results
  - rep.create_model_bpa_report()

Automate Collection through Pipeline

- Either
  1. Use Environment with Semantic Link Labs
  2. Python inline installation
     - enable %pip install for pipeline, add "_inlineInstallationEnabled" as bool parameter equals True in the notebook activity parameters.

Reference:

- https://learn.microsoft.com/en-us/fabric/dataengineering/library-management

Multiple Runs

- RunId gets incremented
- Track improvements over time

Demo: Best Practice Analyzer

Our Journey

1. Semantic Link
2. Semantic Link Labs
3. Direct Lake Migration
4. Additional Uses
5. Conclusion

Conclusion

- Enhances Fabric
  - Development
  - Administration
  - Automation

Resources

Semantic Link

https://learn.microsoft.com/en-us/fabric/data-science/semantic-link-overview

Semantic Link Labs

https://github.com/microsoft/semantic-link-labs

Want to become a speaker or mentor?

NEW STARS OF DATA

newstarsofdata.com

5/16/2025

Call for Speakers & Helpers is open!

Thank you

The Dax Shepherd

Jason Romans

thedaxshepherd@gmail.com

www.thedaxshepherd.com