

# Microsoft Fabric

Fabric Analyst in a Day

Version: December 2023

# Agenda (times are approximate and will be fluid with the class)

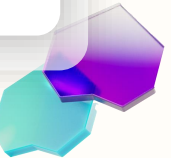
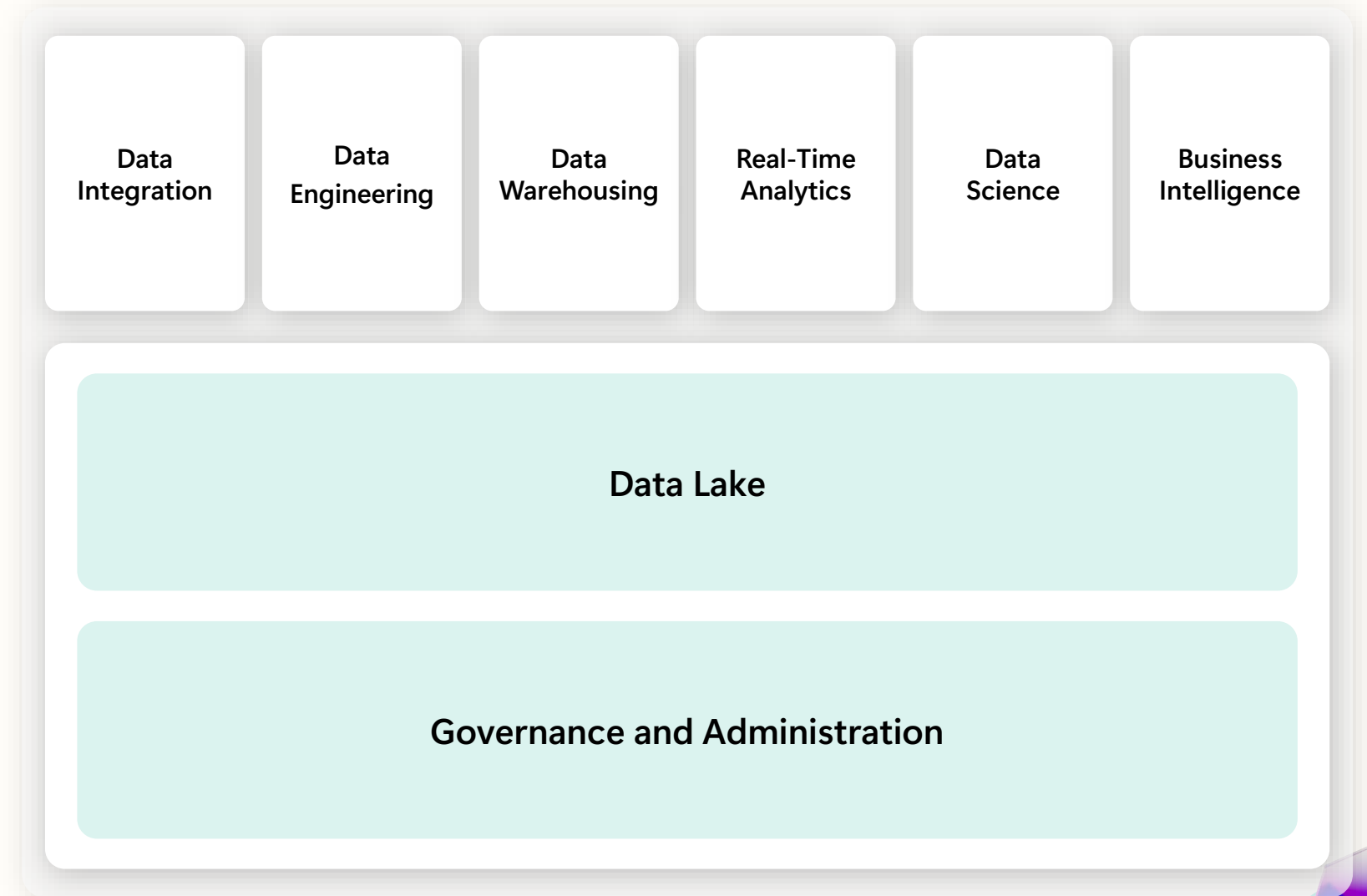
Morning	Presentation	Labs
09:00 AM – 10:00 AM	Introducing the new Microsoft Fabric solution	
10:00 AM – 10:30 AM	OneLake overview	Lab 1 – Problem Statement
<b>10:30 AM – 10:45 AM</b>	<b>Break</b>	
10:45 AM – 11:15 AM	Data Engineering overview	Lab 2 – Fabric Workspace
11:15 AM – 12:15 PM	Data Integration overview	Lab 3 – Data Integration Experience – Part 1
<b>12:15 PM – 01:00 PM</b>	<b>Break for Lunch</b>	
Afternoon	Presentation	Labs
01:00 AM – 02:00 PM		Lab 4 – Data Integration Experience – Part 2
02:00 PM – 03:00 PM	Data Warehousing overview	Lab 5 – Data Integration Experience – Part 3
<b>03:00 PM – 03:15 PM</b>	<b>Break</b>	
03:15 PM – 04:15 PM	Business Intelligence overview	Lab 6 – Data Engineering Experience
04:15 PM – 04:45 PM	Next steps and resources	Lab 7 – Business Intelligence Experience
04:45 PM – 05:00 PM		Lab 7 – Clean Up

Introducing the  
new Microsoft  
Fabric solution



# Analytics today

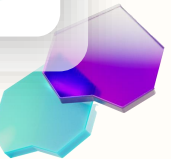
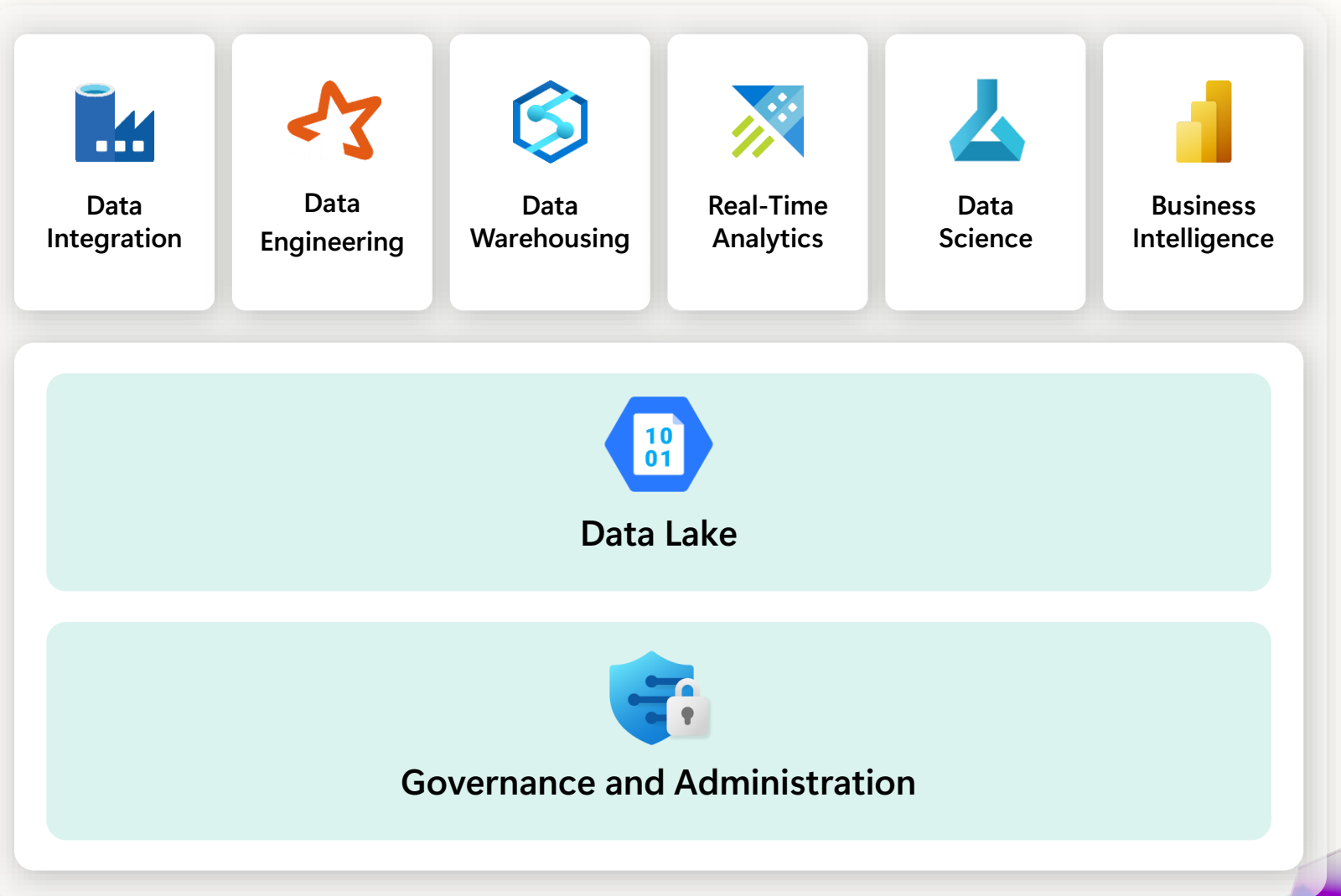
Analytics has very predictable patterns.



# Analytics today

Analytics has very predictable patterns.

Microsoft has all the products with the right scale needed to build a complete analytics system.



# Still far too complex

Many Products

Dedicated and Serverless

Steep Learning Curves

Different Experiences

PaaS and SaaS

Deep Expertise Needed

Proprietary and Open

Different Business Models

High Integration Effort



Purview



Power BI



Kusto



Data Factory



Azure AI



Synapse DW



Synapse Spark

# Introducing Microsoft Fabric—a unified analytics solution for the era of AI



OneLake



Data  
Factory



Synapse Data  
Engineering



Synapse Data  
Warehouse



Synapse Data  
Science



Synapse Real-  
Time Analytics



Power  
BI



Data  
Activator

Lake-centric  
and open

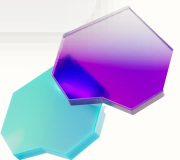
Role-specific  
tools

Empower every  
Microsoft 365 user

Persistent governance  
and security

## Single...

- Onboarding and trials
- Sign-on
- Navigation model
- UX model
- Workspace organization
- Collaboration experience
- Data Lake
- Storage format
- Data copy for all engines
- Security model
- CI/CD
- Monitoring hub
- Data Hub
- Governance & compliance





# Microsoft Fabric

The data platform for the era of AI

## Complete Analytics Platform

Everything, unified

---

SaaS-ified

---

Secured and governed

## Lake centric and open

OneLake

---

One Copy

---

Open at every tier

## Empower Every Business User

Familiar and intuitive

---

Built into Microsoft 365

---

Insight to action

## AI Powered

Copilot accelerated

---

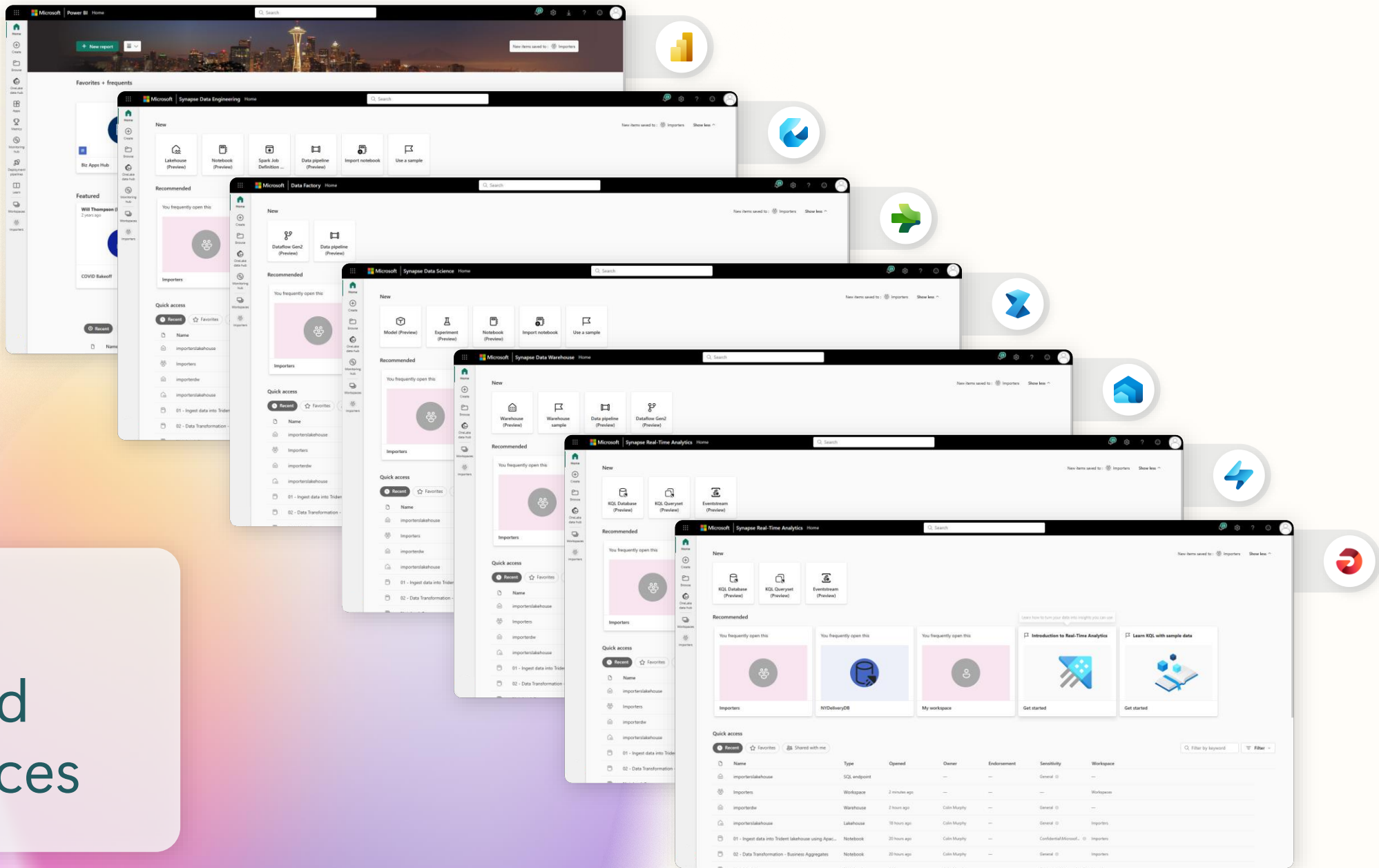
ChatGPT on your data

---

AI driven insights



# Persona optimized experiences



# Seven key experiences for end-to-end analytics

Experiences are designed to target specific personas and tasks, yet work together seamlessly in a unified platform via OneLake to enable creators to collaboratively do their best work:



Combines the ease of use of Power Query with the scale and power of Azure Data Factory to leverage 200+ native connectors to data sources on-premises and in the cloud



World-class Spark platform with great authoring experiences to empower data engineers to transform data at scale



Providing industry leading SQL performance and scale, fully separating compute from storage for independently scaling and natively storing data in open Delta Lake



Build, deploy, and operationalize machine learning models directly within Fabric to empower data scientists and analysts with predictive insights



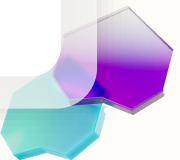
Best-in-class engine for observational data analytics to create actionable insights from real-time data



The world's leading business intelligence platform empowers users to quickly and intuitively to make better decisions with data



Automatically drive actions on your data, without writing code

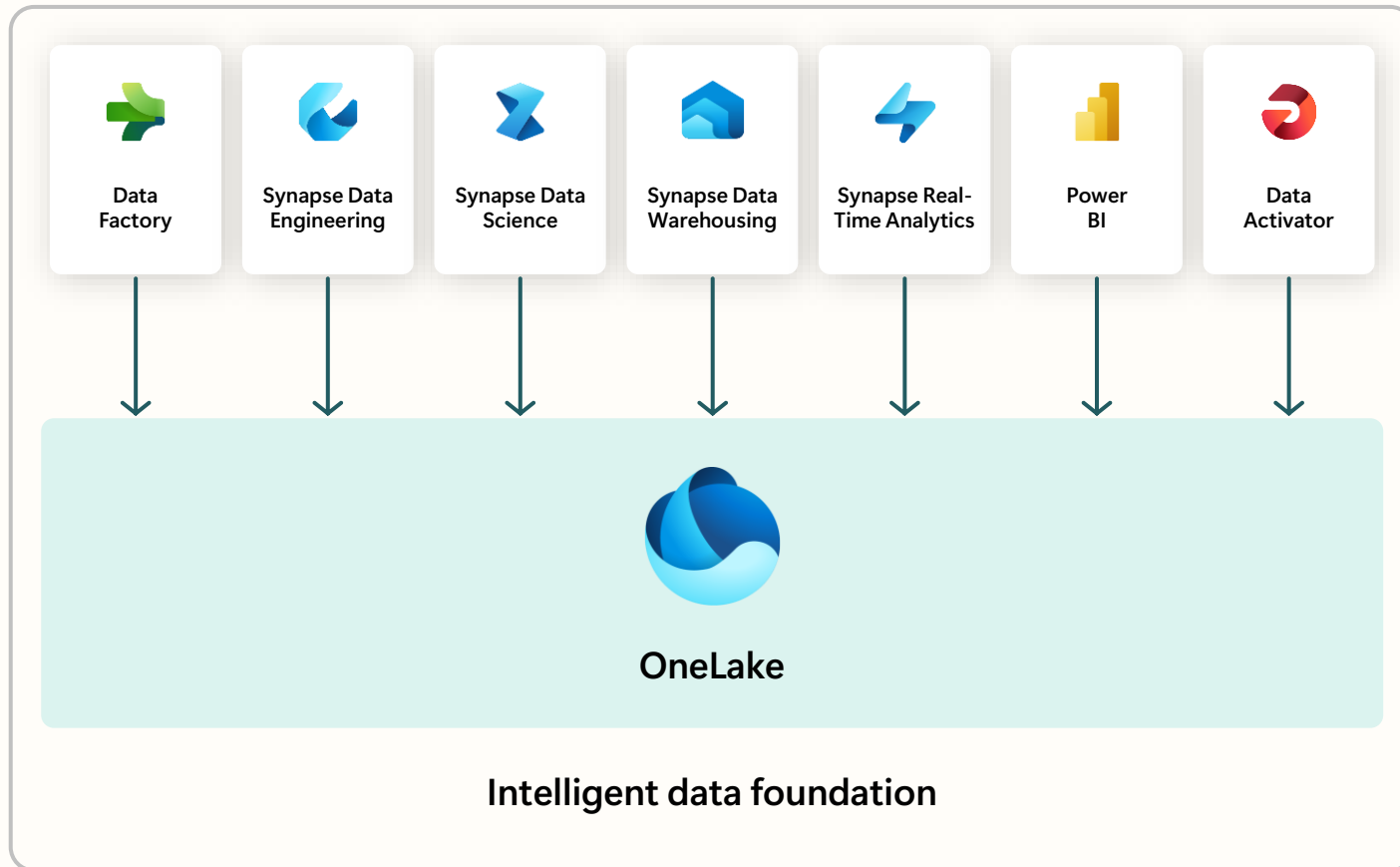


# OneLake overview



# OneLake for all data

“The OneDrive for data”



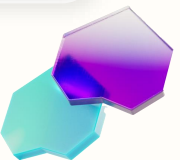
A single SaaS lake for the whole organization

Provisioned automatically with the tenant

All workloads automatically store their data in the OneLake workspace folders

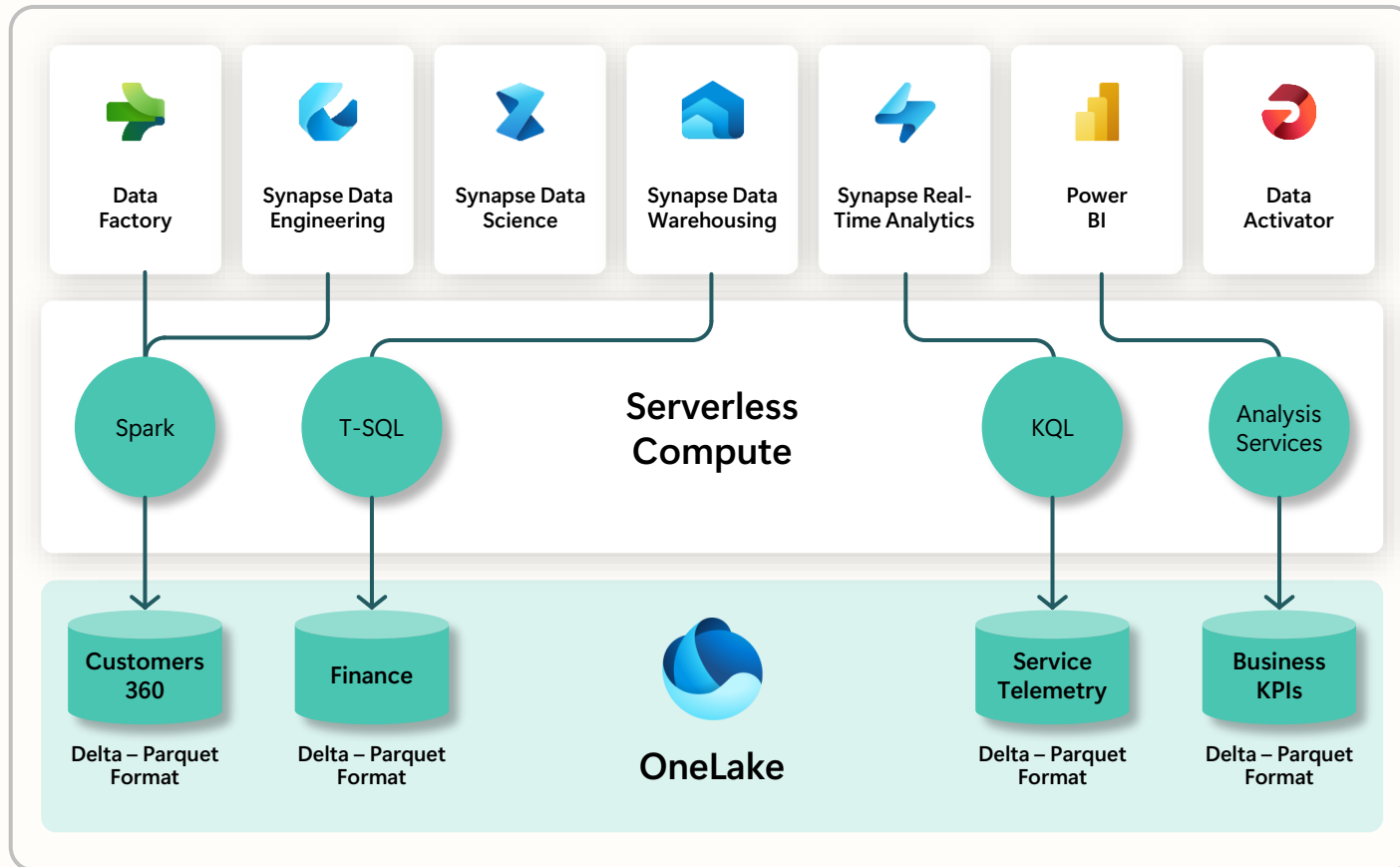
All the data is organized in an intuitive hierarchical namespace

The data in OneLake is automatically indexed for discovery, MIP labels, lineage, PII scans, sharing, governance, and compliance



# One Copy for all computes

Real separation of compute and storage



All the compute engines store their data automatically in OneLake

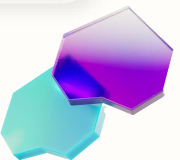
The data is stored in a single common format

[Delta - Parquet](#), an open standards format, is the storage format for all tabular data in Microsoft Fabric

Once data is stored in the lake, it is directly accessible by all the engines without needing any import / export

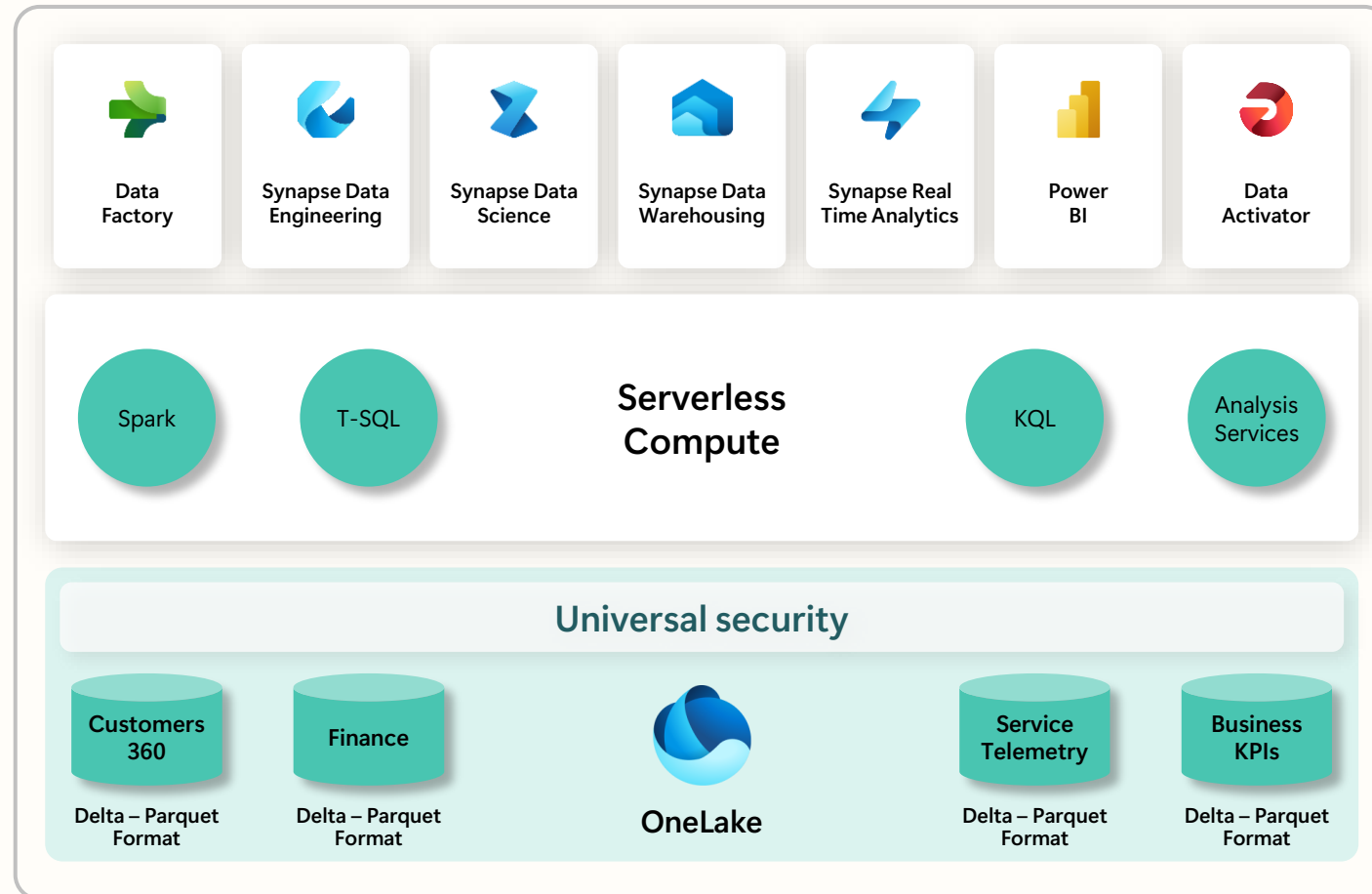
All the compute engines have been fully optimized to work with Delta Parquet as their native format

Shared universal security model is enforced across all the engines



# One Copy for all computes

Universal security makes it real



All the compute engines store their data automatically in OneLake

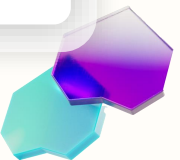
The data is stored in a single common format

[Delta - Parquet](#), an open standards format, is the storage format for all tabular data in Microsoft Fabric

Once data is stored in the lake, it is directly accessible by all the engines without needing any import / export

All the compute engines have been fully optimized to work with Delta Parquet as their native format

Shared universal security model is enforced across all the engines



# Shortcuts



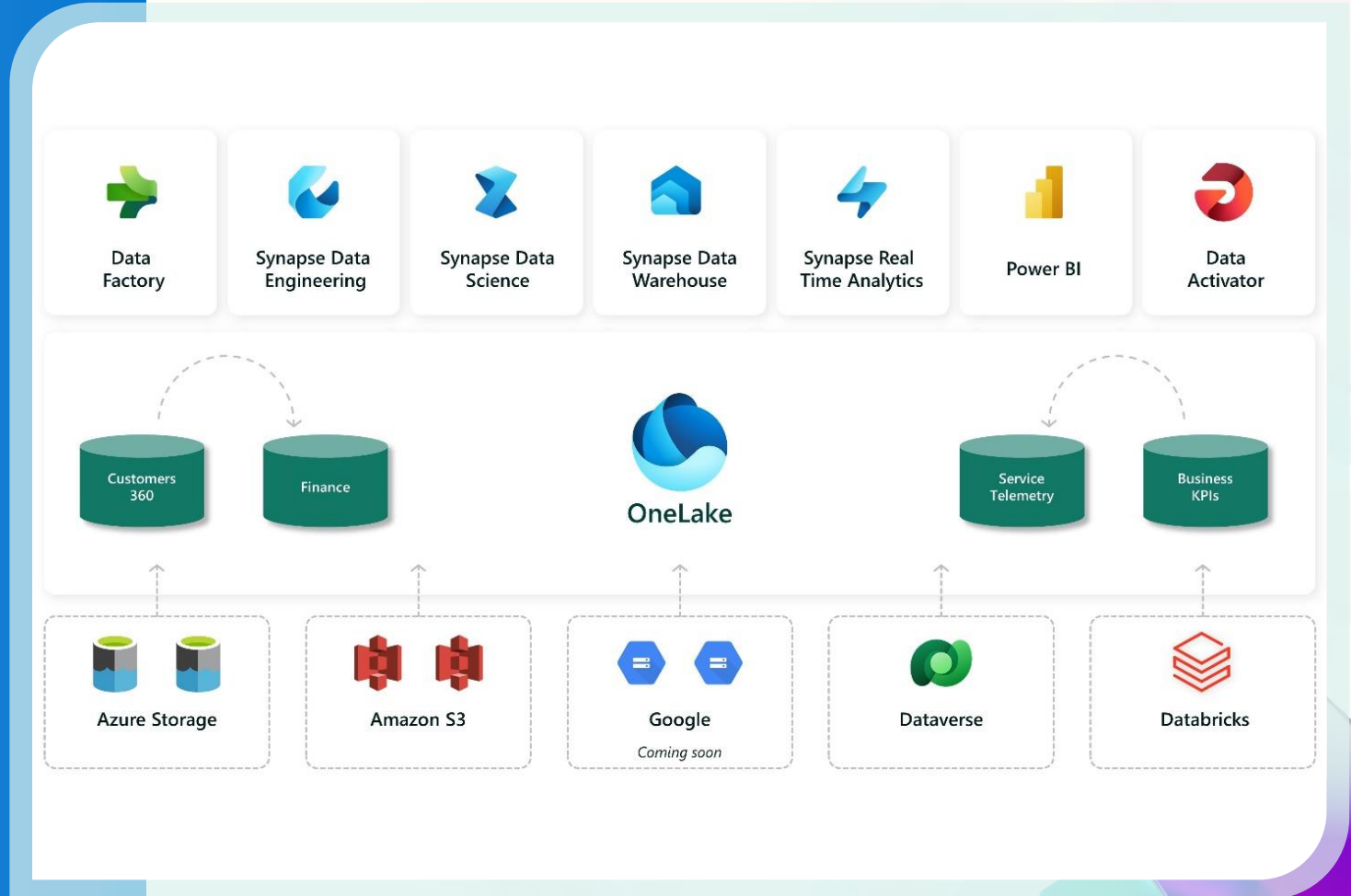
**Shortcuts** unify data without copying or moving existing data.

This means that data can be used multiple times without data duplication.



## Key Capabilities:

- Create shortcuts within Microsoft Fabric to consolidate data across artifacts or workspaces, without changing ownership of the data
- With shortcuts, data throughout OneLake can be composed together without any data movement
- Shortcuts also allow instant linking of data already existing in Azure and in other clouds, without any data duplication and movement, making OneLake the first multi-cloud data lake
- With support for industry standard APIs, OneLake data can be directly accessed by any application or service





# OneLake for all domains

## A true data mesh across organization domains

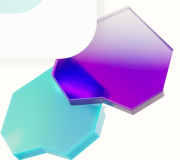
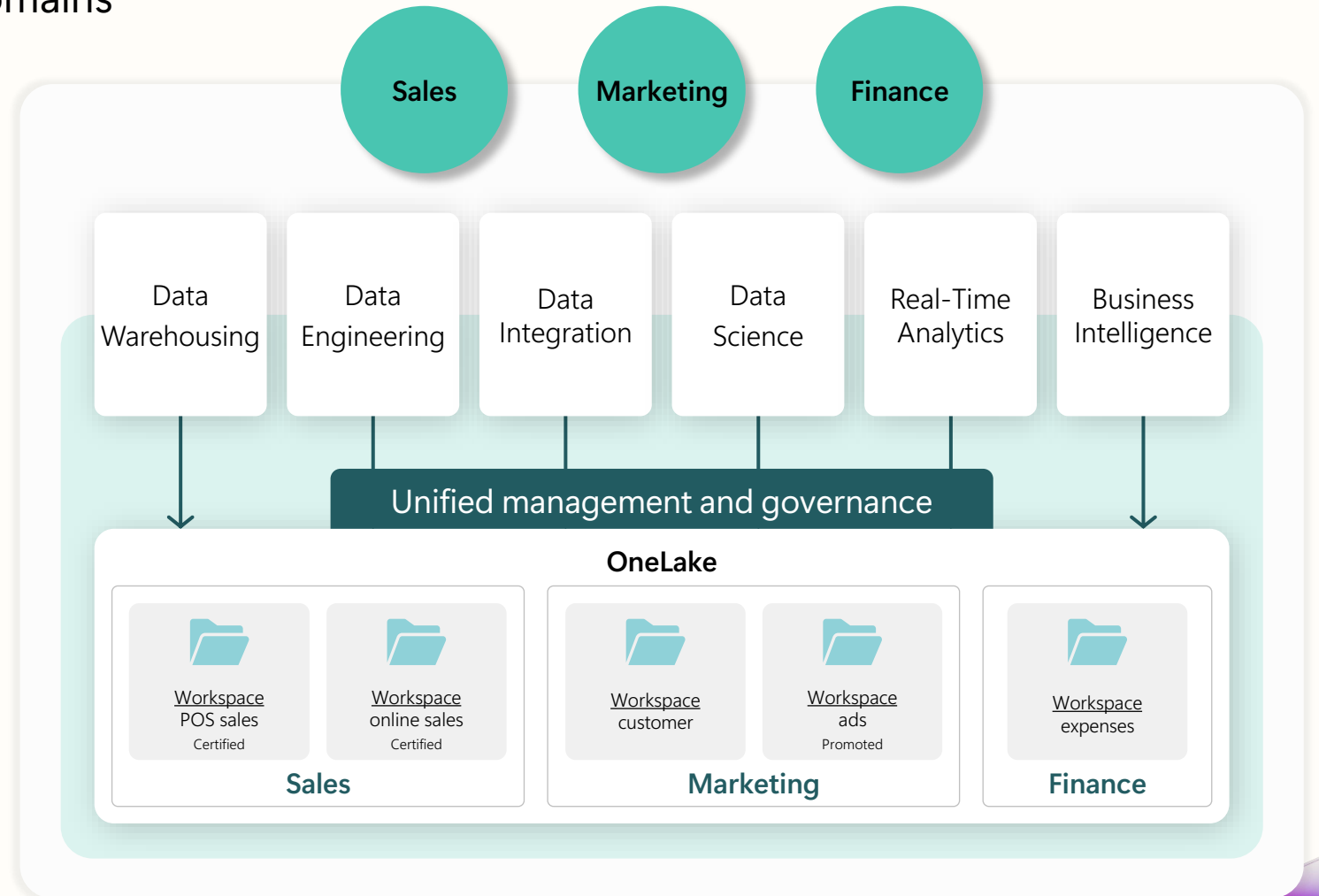
Introducing **domains** as an integral part of Fabric:  
A domain is a way to logically group **together** all the data in an organization relevant to an area or field, according to business needs

Domains are defined with **domain admins** and **contributors** who can **associate** workspaces and group them together under a relevant domain

**Federated governance** can be achieved by delegating settings to domain admins, thus allowing them to achieve more **granular control** over their business area

Domains simplify **discovery** and **consumption** of data across the organization, thus allowing business optimized consumption

Avoid data swamps by endorsing certain data as **certified** or **promoted**, thus encouraging **reuse**





# OneLake data hub

Discover, manage, and reuse data in one place

Name	Type	Owner	Refreshed	Location	Endorsement	Sensitivity
Business Dataset	Dataset	Arun Ulag	9/6/22, 5:20:24 PM	Sales Report - Contoso	-	Confidential GDPR
US Retail Sales	Dataset	hcl-admin	10/20/19, 10:42:15 PM	Sales HQ	-	-
Contoso Dataset	Dataset	Arun Ulag	3/17/22, 8:04:24 PM	Contoso Sales	-	General
Supplier-Quality-Analysis-Sampl...	Dataset	Arun Ulag	3/22/20, 3:12:13 AM	Sales HQ	-	General
Operating Report	Dataset	Arun Ulag	10/5/19, 1:20:04 AM	Sales HQ	-	Confidential/Intern...

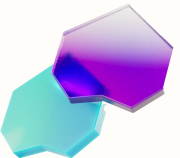
Central location within Fabric to discover, manage, and reuse data. Pervasive experience – available everywhere users discover data (create shortcuts, get data in dataflow, attach a notebook, create a semantic model, and more)

Data can be easily discovered by its domain (e.g., Finance) so users can see what matters to them

Efficient data discovery using search, filter and sort, or via browsing by folder (workspace) hierarchy

Data actions such as preview, exploration, tracking lineage, report creation, analyze in Excel, and export can be easily done, even by non-technical users

Available in Microsoft Teams, thus bringing OneLake Data hub to Office, enabling both technical and non-technical users with discovery, reuse, and exploration

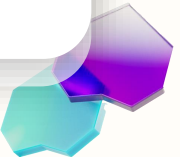


# Customer scenario

- Fabrikam, Inc. is a wholesale novelty goods distributor. As a wholesaler, Fabrikam's customers are mostly companies who resell to individuals. Fabrikam sells to retail customers across the United States including specialty stores, supermarkets, computing stores, and tourist attraction shops. Fabrikam also sells to other wholesalers via a network of agents who promote the products on Fabrikam's behalf. While all Fabrikam's customers are currently based in the United States, the company is intending to push for expansion into other countries/regions.
- You are a Data Analyst in the Sales team. You collect, clean, and interpret data sets to solve business problems. You also put together visualizations like charts and graphs, write reports, and present them to the decision-makers in the organization.
- To draw valuable insights from the data, you pull data from multiple systems, clean it, and mash it up together. You pull data from the following sources:
  - **Sales Data:** This data comes from the ERP system and is stored in an ADLS Gen2 database or Databricks. It gets updated at noon / 12 PM every day.
  - **Supplier Data:** This data comes from different suppliers and is stored in a Snowflake database. It gets updated at midnight / 12 AM every day.
  - **Customer Data:** This data comes from Customer Insights and is stored in Dataverse. The data is always up to date.
  - **Employee Data:** This data comes from the HR system; it is stored as an export file in a SharePoint folder. It gets updated every morning at 9 AM.

# Customer scenario continued

- You are currently building a semantic model on Power BI Premium that pulls the data from the above source systems order to satisfy your reporting needs as well as provide end users with the ability to self-serve. You use Power Query to update your semantic model.
- **You are facing the following challenges:**
  - You need to refresh your semantic model at least three times a day to accommodate the different update times for the different data sources.
  - Your refreshes take a long time as you need to do a full refresh every time to capture any updates that happened to the source systems.
  - Any errors in any of the data sources you are pulling from will result in your semantic model refresh breaking. A lot of times the employee file doesn't upload on time resulting in your semantic model refresh breaking.
  - It takes a very long time to make any changes to your data model as Power Query takes a long time to refresh your previews, given the large data sizes and complex transformations.
  - You need a Windows PC to use Power BI Desktop even though the corporate standard is Mac.
  - You heard about Microsoft Fabric, and decided to try to see if it will address your challenges.



# Lab 1

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



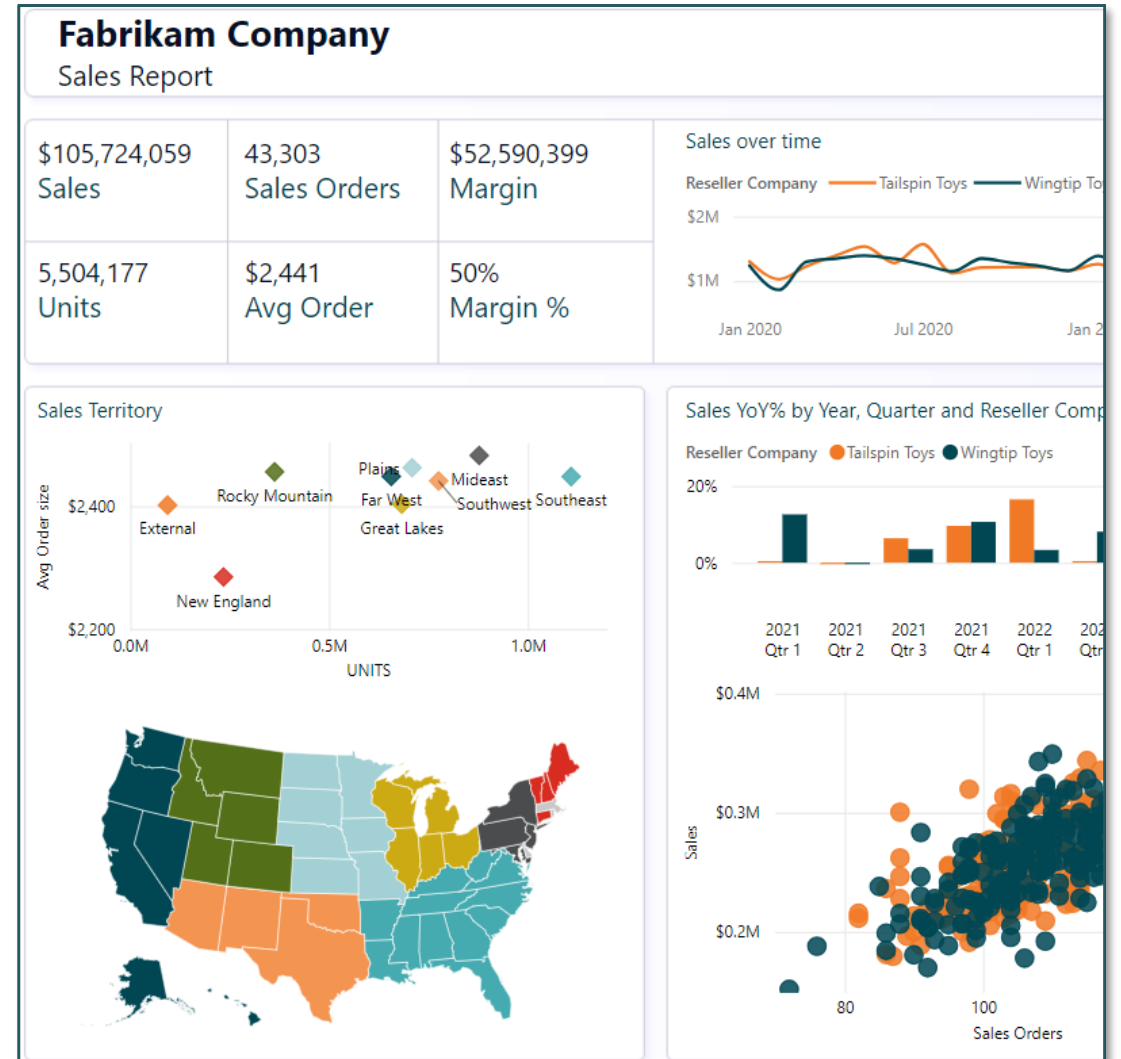
How to set up Power BI Desktop in lab environment



How to analyze Power BI Desktop Report



How to review Power Queries to understand the data sources



15 minute break



# Data Engineering overview





# Lakehouse | Overview



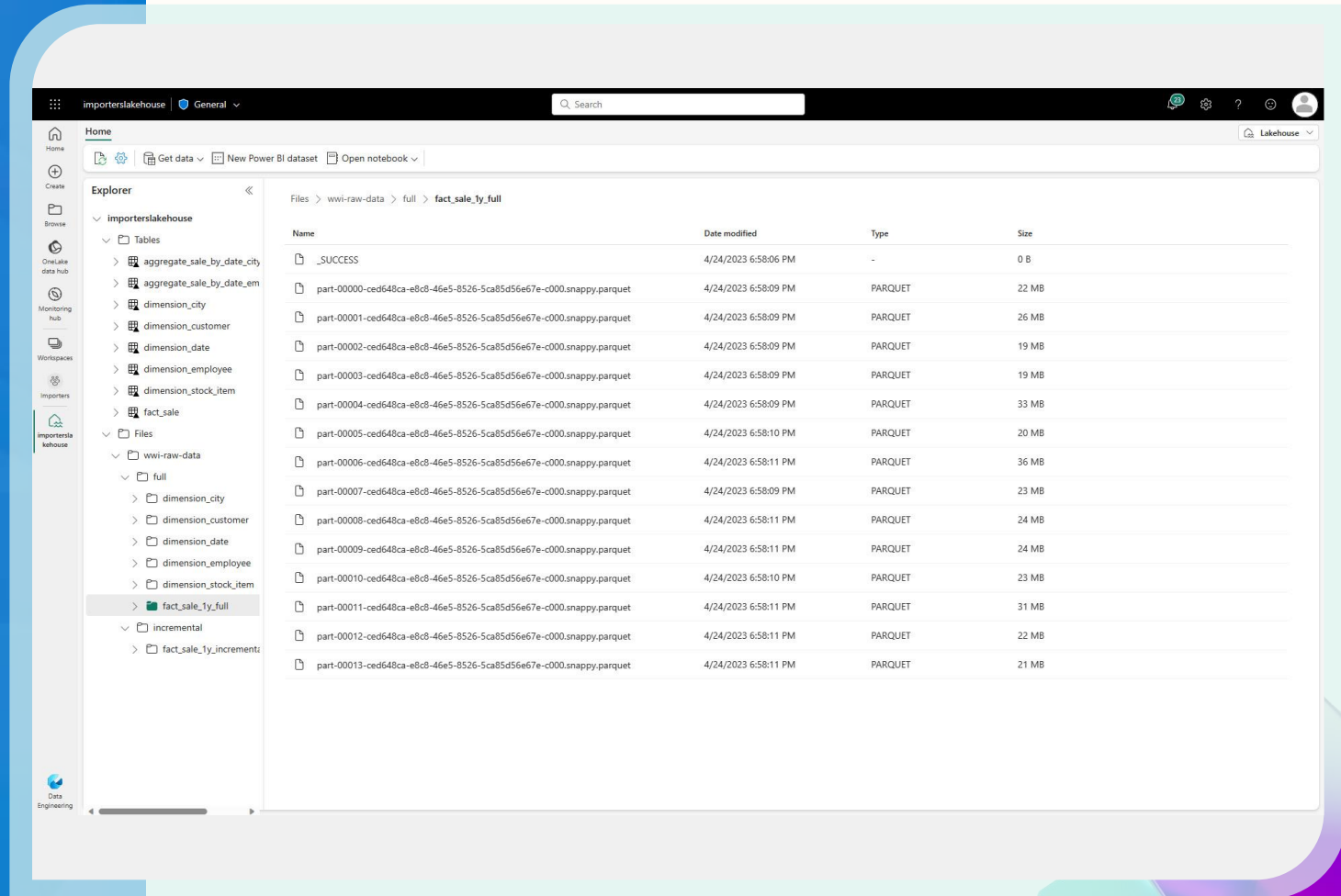
**Store, manage, and analyze** all your data in a single location and easily share across the entire enterprise.

Quickly and easily create a Lakehouse without having to provision and configure compute, storage, and networking.



## Key Capabilities:

- Flexible and scalable solution that enables organizations to handle large data volumes of all types and sizes
- Built-in SQL endpoint unlocks data warehouse capabilities on top of your Lakehouse with no data movement
- Use 'direct lake' mode to build reports in seconds directly on top of the data lake with blazing fast performance
- Easily ingest data into the Lakehouse through a variety of methods
- Share your Lakehouse as a data product with consumers



# Notebook | Overview

Immersive authoring experience for data developers.

Rich notebook capabilities including native Lakehouse integration, real-time collaboration with commenting support, auto-save support, lightweight scheduling, and pipeline integration.



## Key Capabilities:

- Manage your Python and R libraries through in-line installs using commands like %pip install
- Advanced notebook development support with ability to reference notebooks in notebooks, and snapshots for easy troubleshooting
- In context monitoring complete with real-time advice and error analysis
- Streamline data prep without giving up the power of reproducibility of Python

The screenshot displays a Databricks notebook interface. The top navigation bar includes 'Home', 'Edit', 'Run', 'Data', and 'View'. The notebook title is 'Visual 1: Distribution of trip duration(minutes) on linear and logarithmic scale'. The code cell contains the following Python code:

```
1 ## Compute trip duration(in minutes) on the sample using pandas
2 sampled_df['tripDuration'] = (sampled_df['tpepDropoffDateTime'] - sampled_df['tpepPickupDateTime']).astype('timedelta64[m]')
3 sampled_df = sampled_df[sampled_df['tripDuration'] > 0]
4
5 fig, axes = plt.subplots(1, 2, figsize=(18, 6))
6 sns.histplot(ax=axes[0], data=sampled_df,
7             x="tripDuration",
8             stat="count",
9             discrete=True).set(title='Distribution of trip duration(minutes)')
10 sns.histplot(ax=axes[1], data=sampled_df,
11             x="tripDuration",
12             stat="count",
13             log_scale=True).set(title='Distribution of trip duration(log scale)')
14 axes[1].xaxis.set_major_formatter(mticker.ScalarFormatter())
15 plt.show()
```

The output shows two histograms. The left histogram, titled 'Distribution of trip duration(minutes)', has a linear x-axis ranging from 0 to 1400 and a y-axis labeled 'Count' ranging from 0 to 3000. The right histogram, titled 'Distribution of trip duration(log scale)', has a logarithmic x-axis ranging from 1 to 1000 and a y-axis labeled 'Count' ranging from 0 to 3000. The notebook interface also shows a 'Lakehouse explorer' on the left with a tree view of tables and files.



# Lab 2

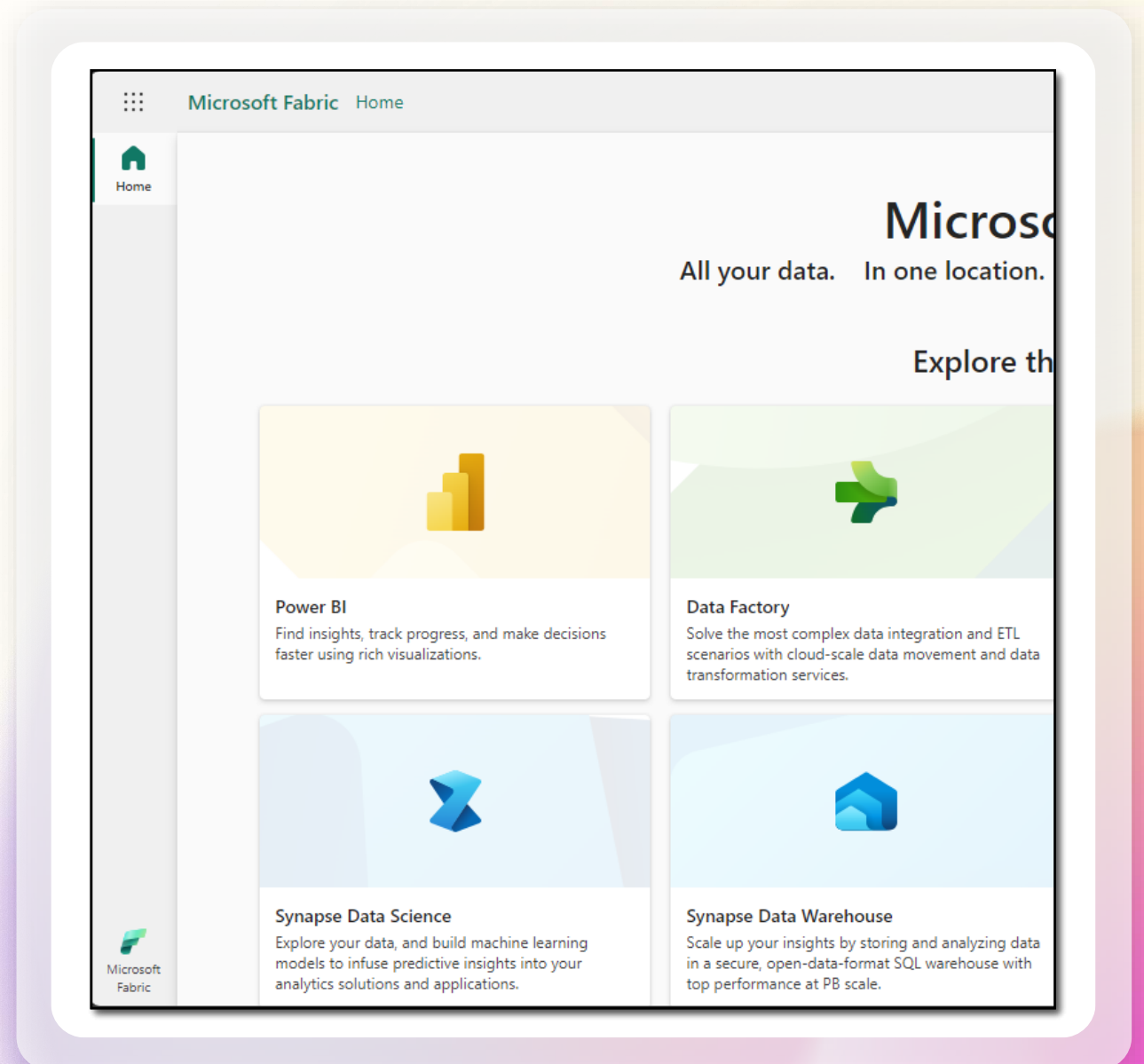
By the end of this lab, you will have learned:



How to create a Fabric workspace



How to create a Lakehouse



# Data Integration overview



# Data Factory



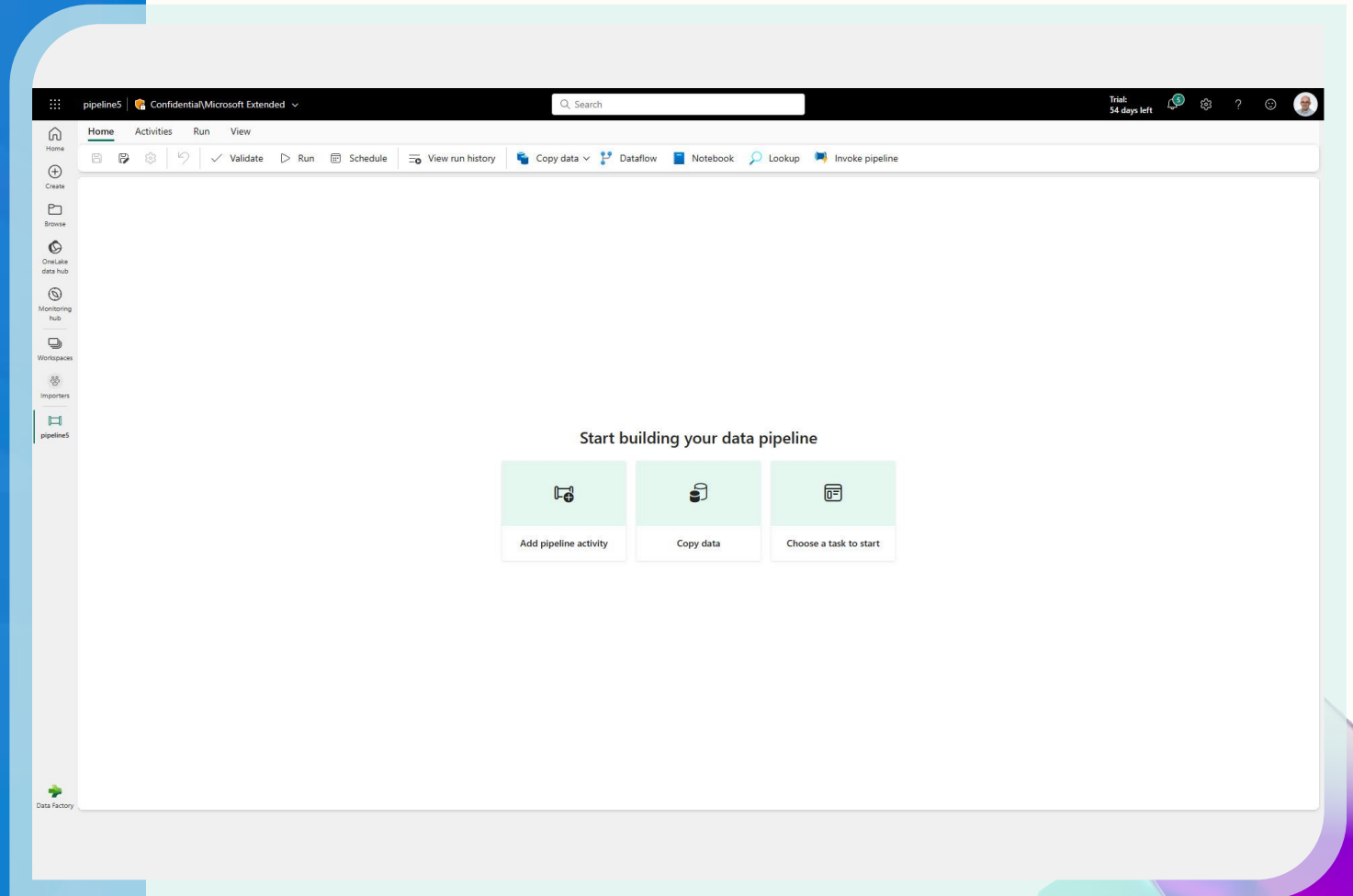
**Data Factory** in Microsoft Fabric provides cloud-scale data movement and data transformation services that allows you to solve the most complex ETL scenarios.

Core to Data Factory are Data Pipelines and Dataflows to give users the option to a low-code, collaborative, and enterprise scale approach for their ETL process.



## Latest capabilities:

- Output destination to Lakehouse
- 14 new connectors available in Pipeline copy activity
- Warehouse connector for Power BI Desktop
- Sample semantic models
- Pipeline Lakehouse copy assist
- Create data pipeline in Lakehouse portal
- Pipeline templates
- Pipeline support for Spark notebooks
- Service principle auth support



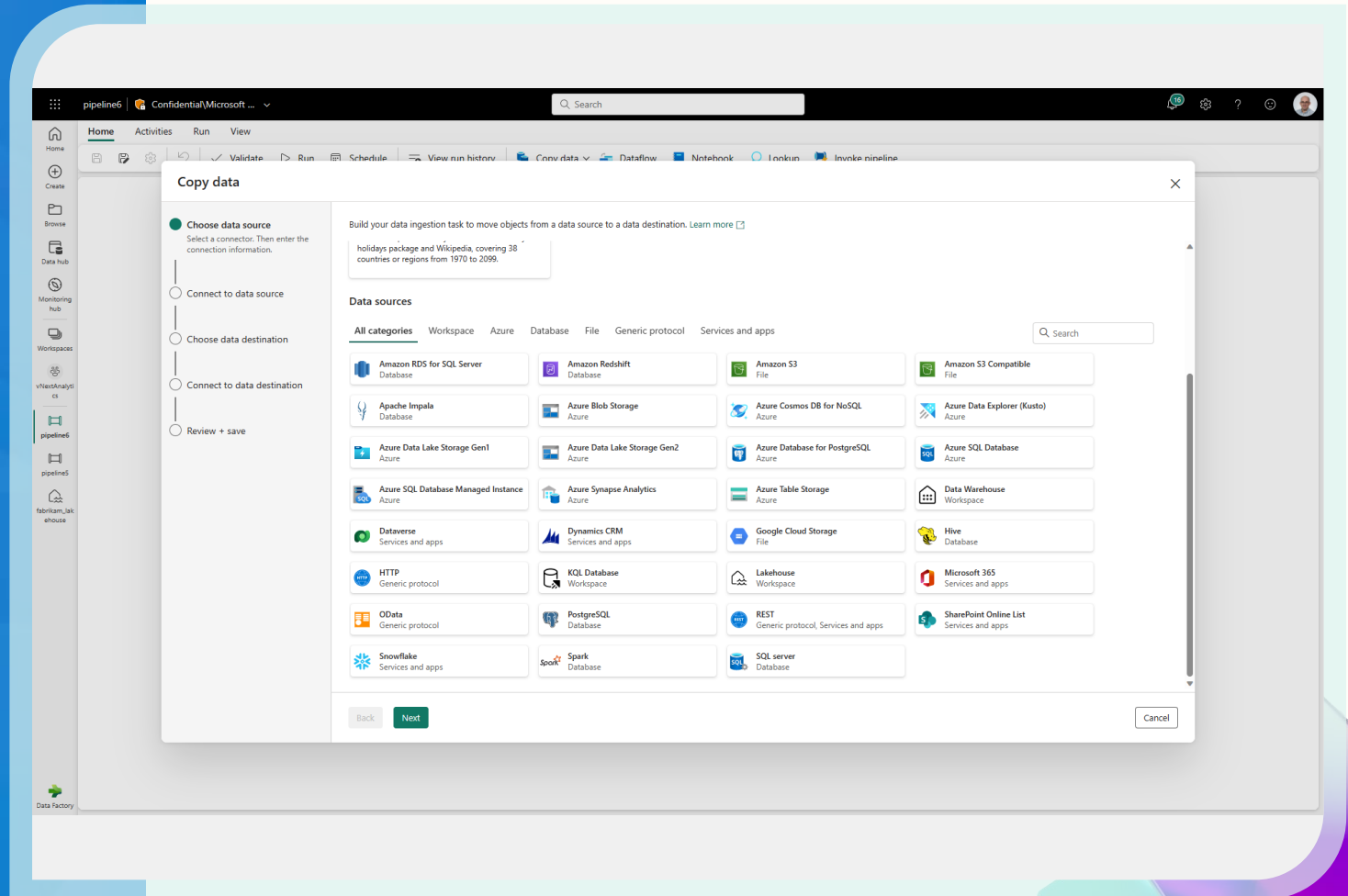
*Autonomous ETL can unlock operational efficiencies and help orchestrate, monitor, and manage pipeline performance.*

# Data Pipelines



**Data Pipelines** enable powerful workflow capabilities at cloud-scale like building complex workflows, moving PB-size data, and defining sophisticated control flow pipelines.

Data pipelines can be used to build complex ETL and data factory workflows that can perform a number of different tasks at scale. Additionally, control flow capabilities are built into pipelines so you can build workflow logic which provide loops and conditional.



# Data Pipelines | Connectors

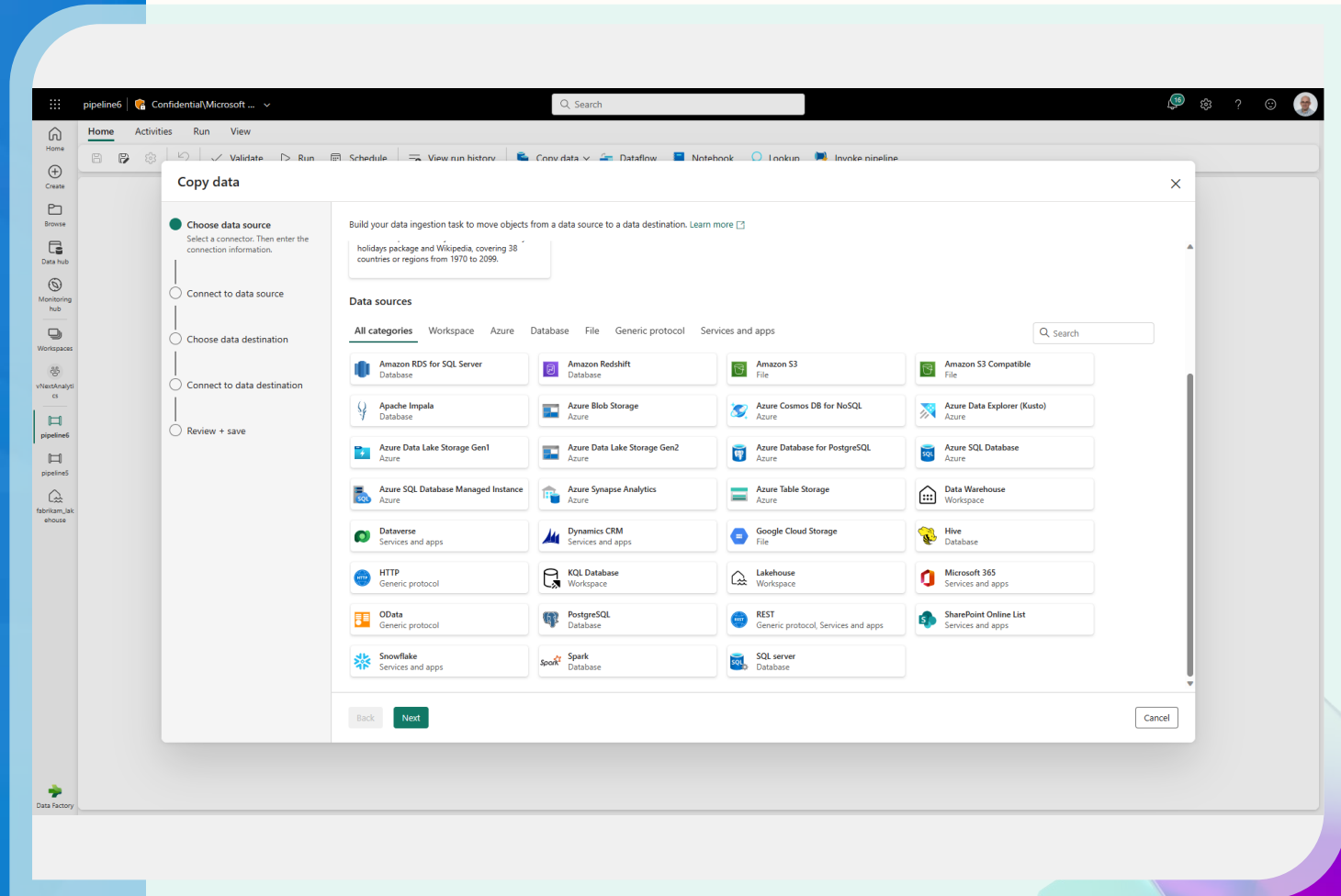


**New Connectors** provide a low-code interface for ingesting data from a variety of data sources.



## New Connectors:

- Warehouse connector; connect to existing Azure Synapse
- Lakehouse connector
- **14 new connectors in the copy activity:** Amazon S3, Azure Table, Amazon Redshift, OData, Google Cloud Storage, Apache Impala, Hive, SQL Server, Azure Synapse Analytics, HTTP, REST, PostgreSQL, Azure Database for PostgreSQL, and Azure Data Explorer



# Data Pipelines | Sample data

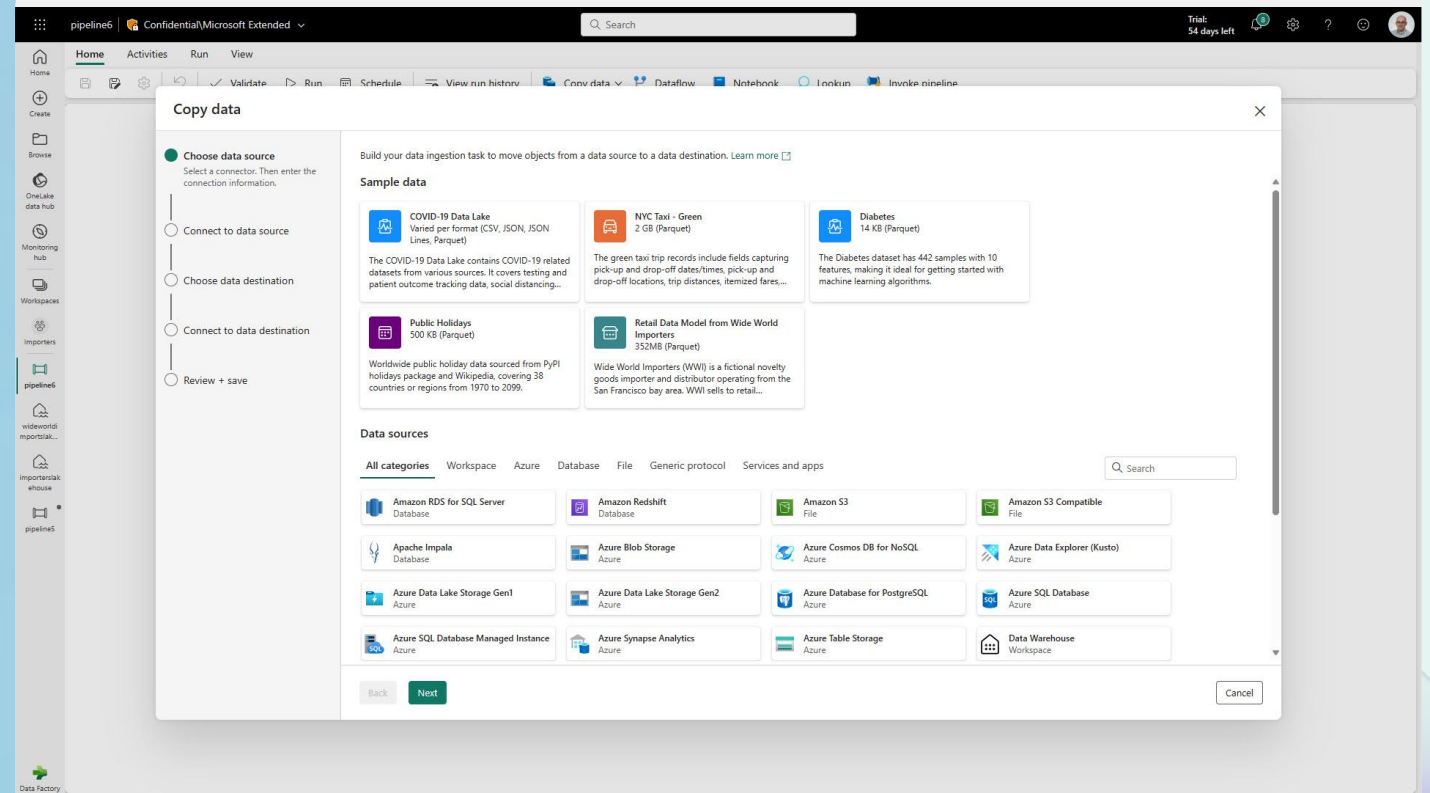


**Sample Semantic models** helps new users get started quickly, building out their ELT processes using Data Pipelines.



## Sample Semantic models:

- COVID-19 Data Lake (CSV, JSON, JSON Lines, Parquet)
- NYC Tax – Green (2GB Parquet)
- Diabetes (14K Parquet)
- Public Holidays (500KB Parquet)
- Retail Data Model from Wide World Importers (352MB Parquet)

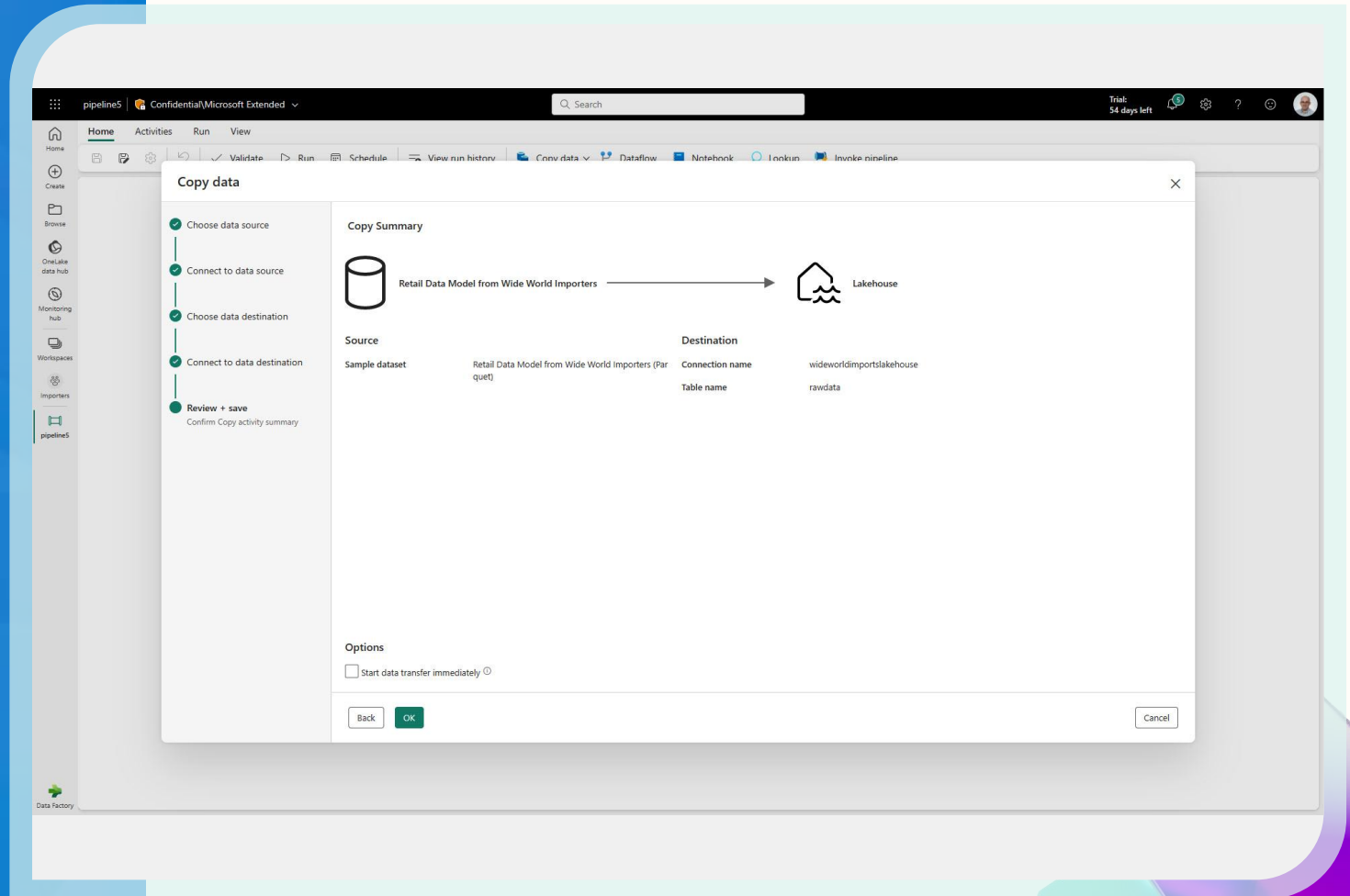




# Data Pipelines | Lakehouse copy assist

**Simply** copy data to a Lakehouse with copy assist capabilities within the Data Pipeline.

Additionally, user can create a Data Pipeline without having to leave the Lakehouse portal.



# Data Pipelines | Templates



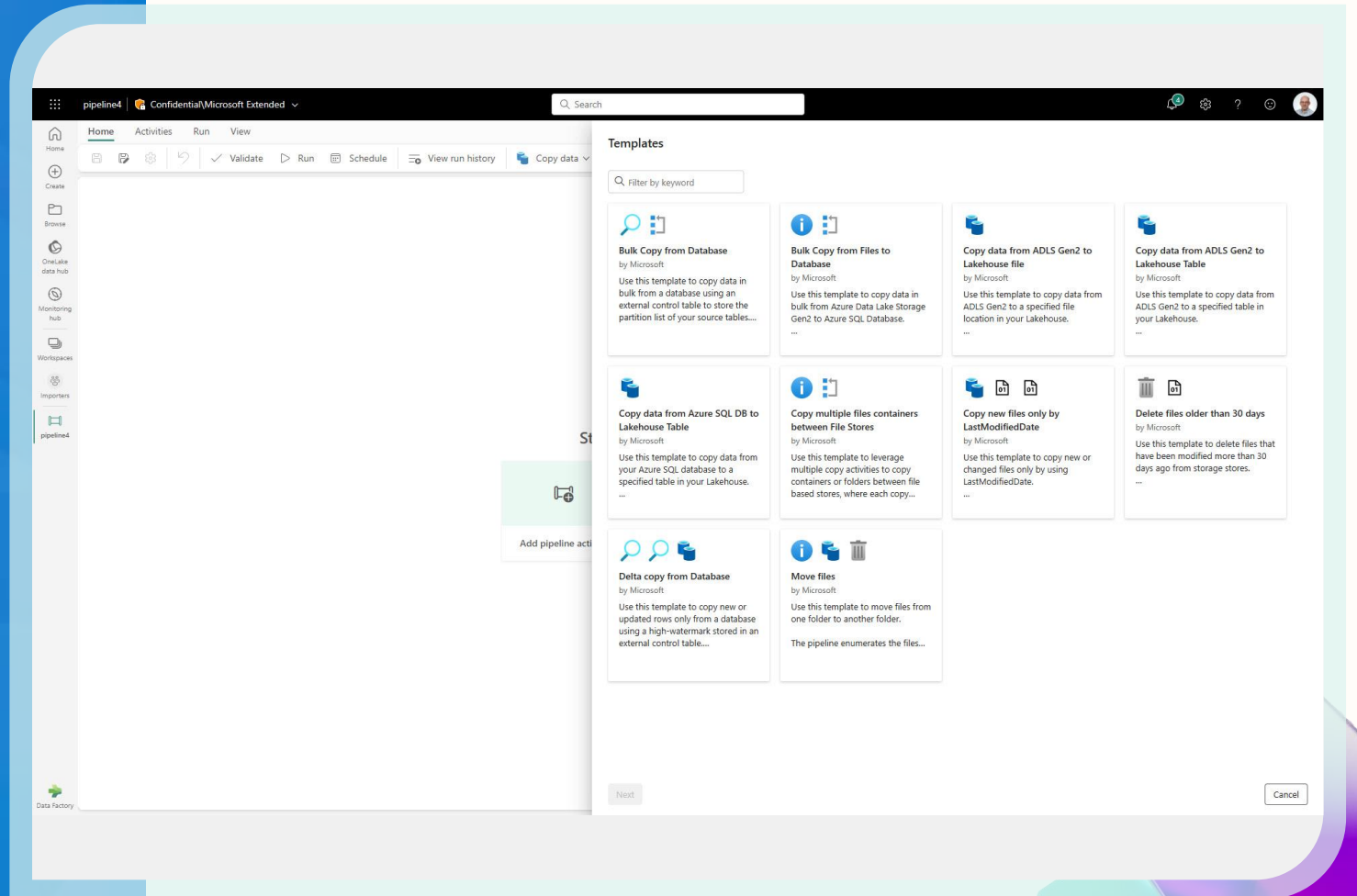
Quickly get started with data integration.

Templates help reduce development time by providing an easy way to create pipeline for common data integration scenarios.



## Available Data Pipeline Templates:

- Bulk copy from Database
- Bulk copy from File to Database
- Copy data from ADLS Gen2 to Lakehouse file
- Copy from ADLS Gen2 to Lakehouse Table
- Copy data from Azure AQL DB to Lakehouse Table
- Copy multiple files containers between File Stores
- Copy new files only by Last Modified Date
- Delete files older than 30 days
- Delta copy from Database
- Move files





# Dataflow

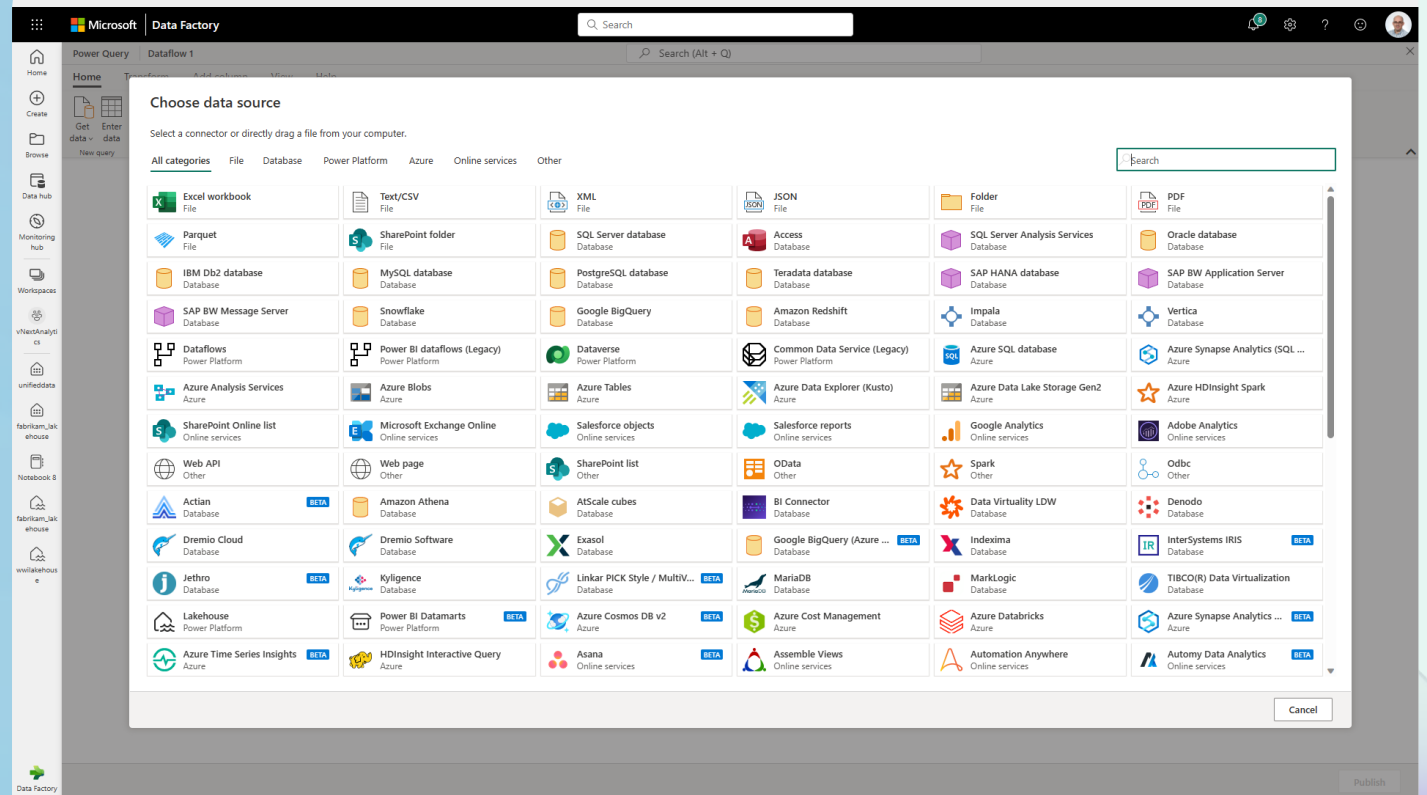
Dataflow provides a low-code interface for ingesting data from hundreds of data sources.

Dataflow quickly and easily unifies disparate data sources, establishes a more collaborative analytics approach, and promotes more informed, agile decision making.



## Key Capabilities:

- Accelerate data transformation with code-free data flows
- Standardize dataflows in Power Apps support for loading data
- Load results of data transformations into multiple destinations (Azure SQL Databases, Lakehouse, etc.)



# Dataflow | Gen1 vs Gen2



Dataflow Gen2 is the new generation of dataflows. The new generation of dataflows resides alongside the Power BI Dataflow (Gen1) and brings new features and improved experiences.

Feature	Dataflow Gen2	Dataflow Gen1
Author dataflows with Power Query	✓	✓
Shorter authoring flow	✓	
Auto-Save and background publishing	✓	
Data destinations	✓	
Improved monitoring and refresh history	✓	
Integration with data pipelines	✓	
High-scale compute	✓	
Get Data via Dataflows connector	✓	✓
Direct Query via Dataflows connector		✓
Incremental refresh		✓
AI Insights support		✓

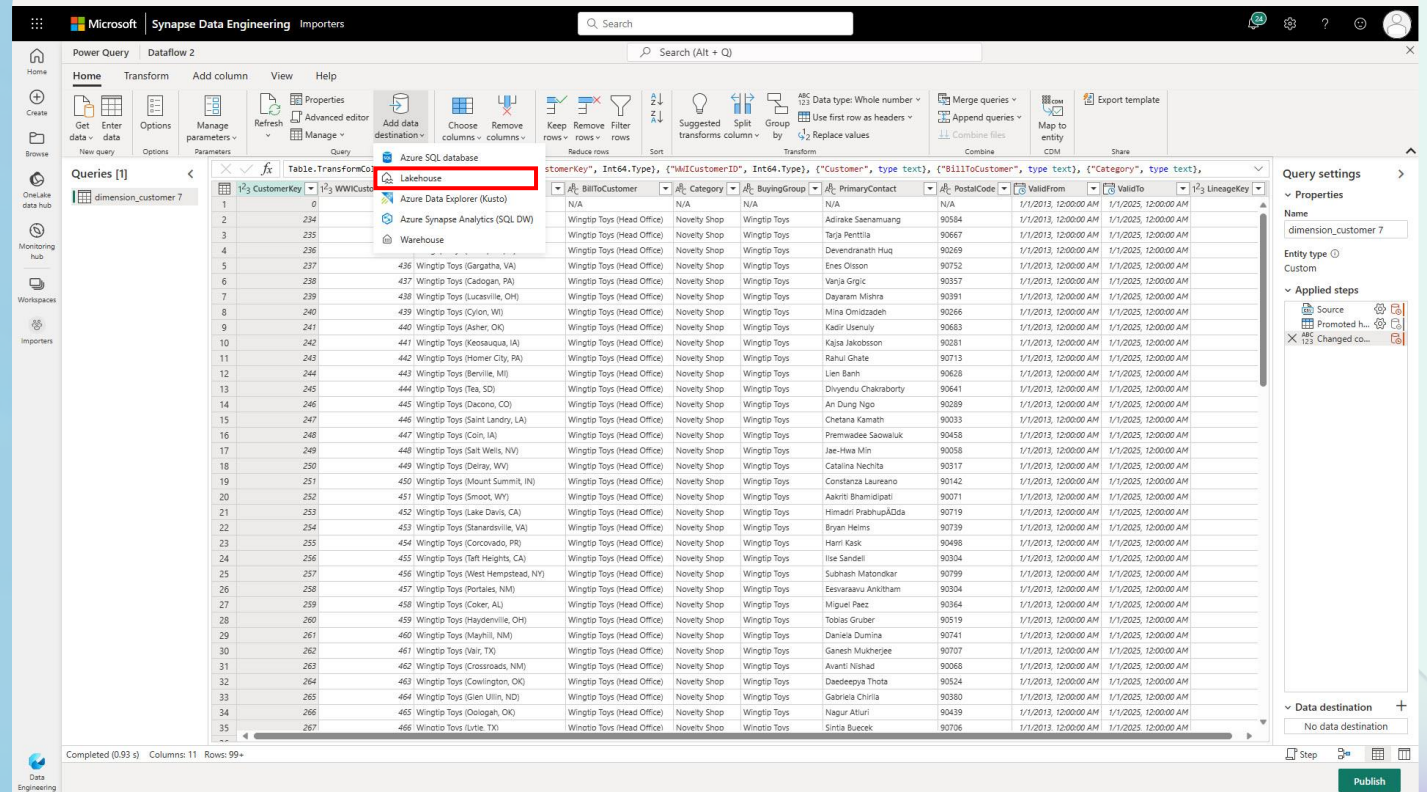
# Dataflow | Output to Lakehouse



Simply write into a Lakehouse from a Dataflow.

Users select the Lakehouse output destination from the list and configure the connection.

This requires the Lakehouse connector to be installed as a custom connector into your data gateway when loading data from on-premise.



# Lab 3

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



How to create Dataflow Gen2



How to connect to ADLS Gen2 using Dataflow Gen2 and transform data



How to ingest data into Lakehouse

The screenshot displays the Microsoft Fabric Dataflow Gen2 interface. On the left, a table shows the results of a query with columns: Unit Price, Tax Rate, and Tax Amount. The data includes rows with values such as 13, 30, 99, 50, 230, 2.7, 32, 13, 2.7, 2.4, 3.7, 13, 3.5, 32, 32, 15, and 25 for Unit Price; 15 for Tax Rate; and 1, 74, 6, 31, 38, 4, 1, 53, 17, 39, 2, 2, and 15 for Tax Amount. On the right, the 'Query settings' panel is visible, showing the name 'Sales', entity type 'Custom', and a list of applied steps including Source, Expanded Invoice, Removed Other Columns, Renamed Columns, Merged Queries, Added Custom, Changed Type, Removed Columns, and Renamed Columns1. The data destination is set to 'Lakehouse'. A 'Publish' button is highlighted with an orange box at the bottom right.

Unit Price	Tax Rate	Tax Amount
13	15	
13	15	
30	15	1
99	15	74
50	15	6
230	15	31
2.7	15	38
32	15	
13	15	
2.7	15	4
2.4	15	1
3.7	15	53
13	15	17
3.5	15	39
32	15	2
32	15	
15	15	2
25	15	

Lunch break



# Lab 4

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



How to connect to Snowflake using Dataflow Gen2 and ingest data into Lakehouse



How to connect to SharePoint using Dataflow Gen2 and ingest data into Lakehouse



How to connect to Dataverse using Dataflow Gen2 and ingest data into Lakehouse

### Dataflow 1

\* Required

Name

Description

Last edited on 10/26/2023 at 11:35:42 AM by ODL\_User 1111422



# Data Warehousing overview



# Data warehouse | Overview



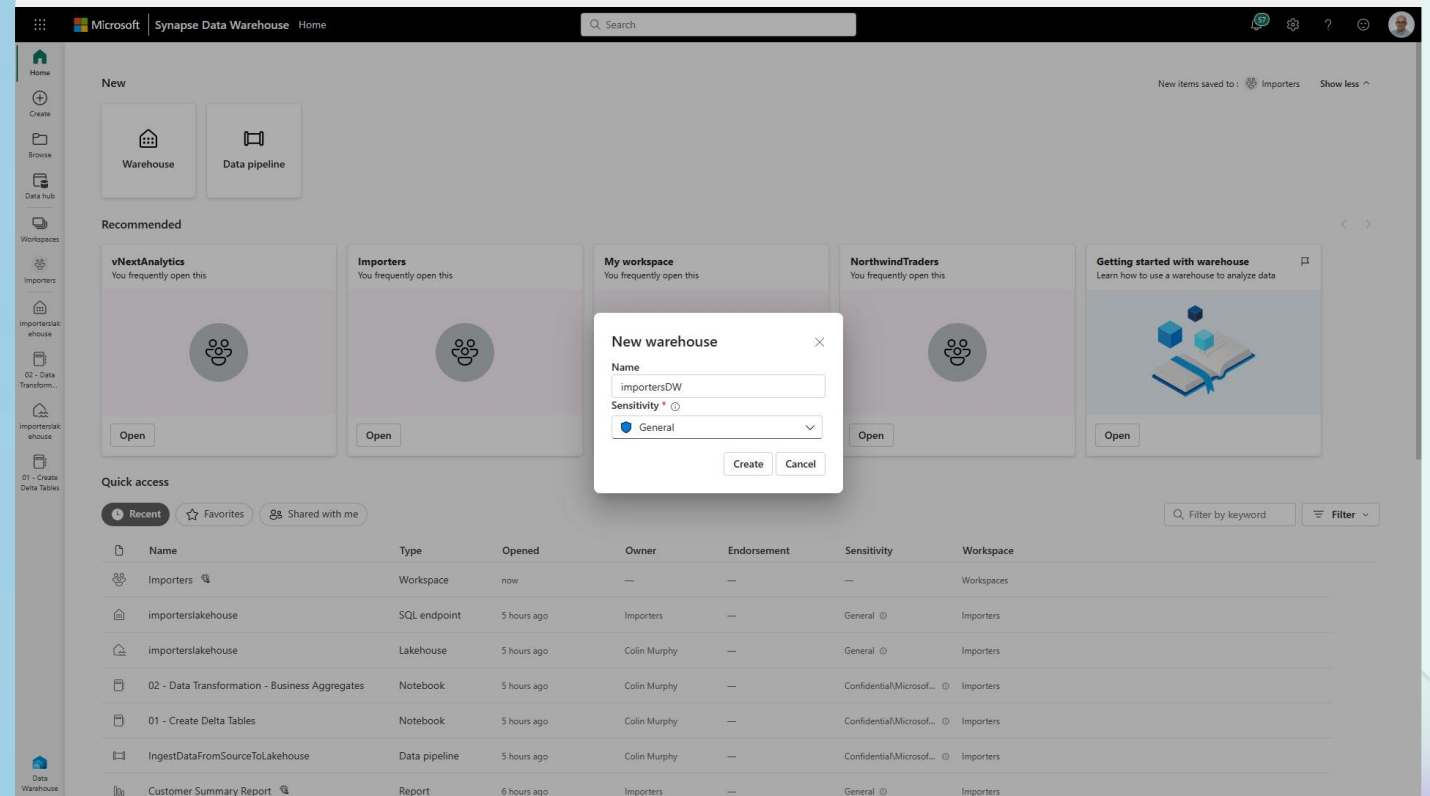
**Enterprise** scale data warehouse with open standard format.

No knobs performance with minimal set-up and deployment, no configuration of compute, or storage needed.



## Key Capabilities:

- Simple and intuitive warehouse experiences for both beginner and experienced data professionals
- Lake-centric warehouse stores data in OneLake in open Delta format with easy data recovery and management
- Fully integrated with all Fabric workloads out-of-the box
- Data loading and transforms at scale, with full multi-table transactional guarantees provided by the SQL engine
- Virtual warehouses with cross-database querying and a fully integrated semantic layer
- Enterprise-ready platform with end-to-end performance and usage visibility, with built-in governance and security
- Flexibility to build data warehouse or data mesh based on organizational needs and choice of no-code, low-code, or T-SQL for transformations





# Fully integrated Power BI semantic layer



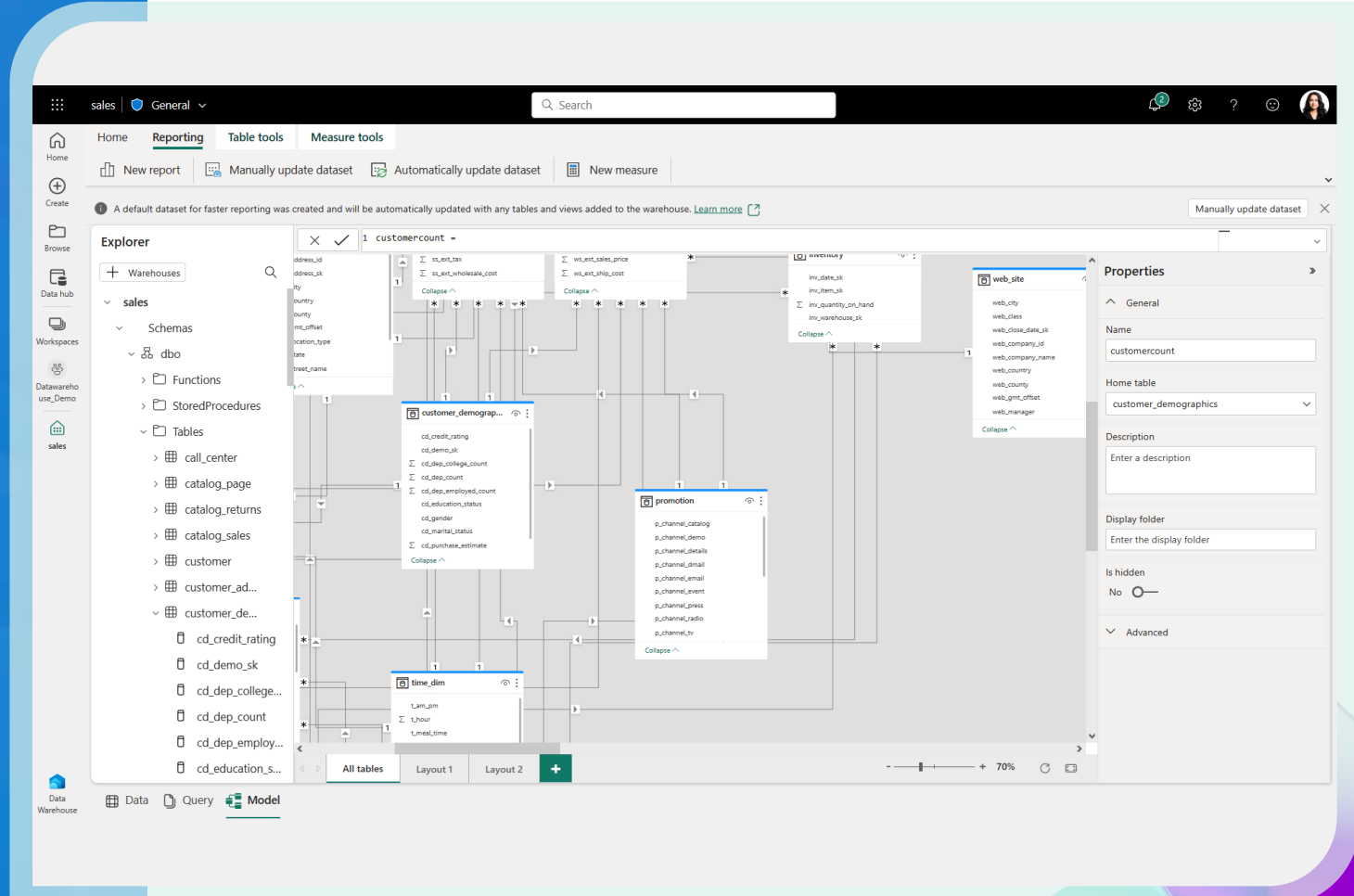
**Reduce integration** and gain insight from your data in seconds.

Built-in Power BI enables everyone to visualize their data in seconds.



## Key Capabilities:

- Auto-generated Power BI semantic model always in sync
- Default semantic model in Direct Lake mode, but automatically switches to Direct Query or Import mode as security or performance needs change
- Flexibility to add / remove tables to semantic model
- Full web authoring experience for creating measures



# Secure by default

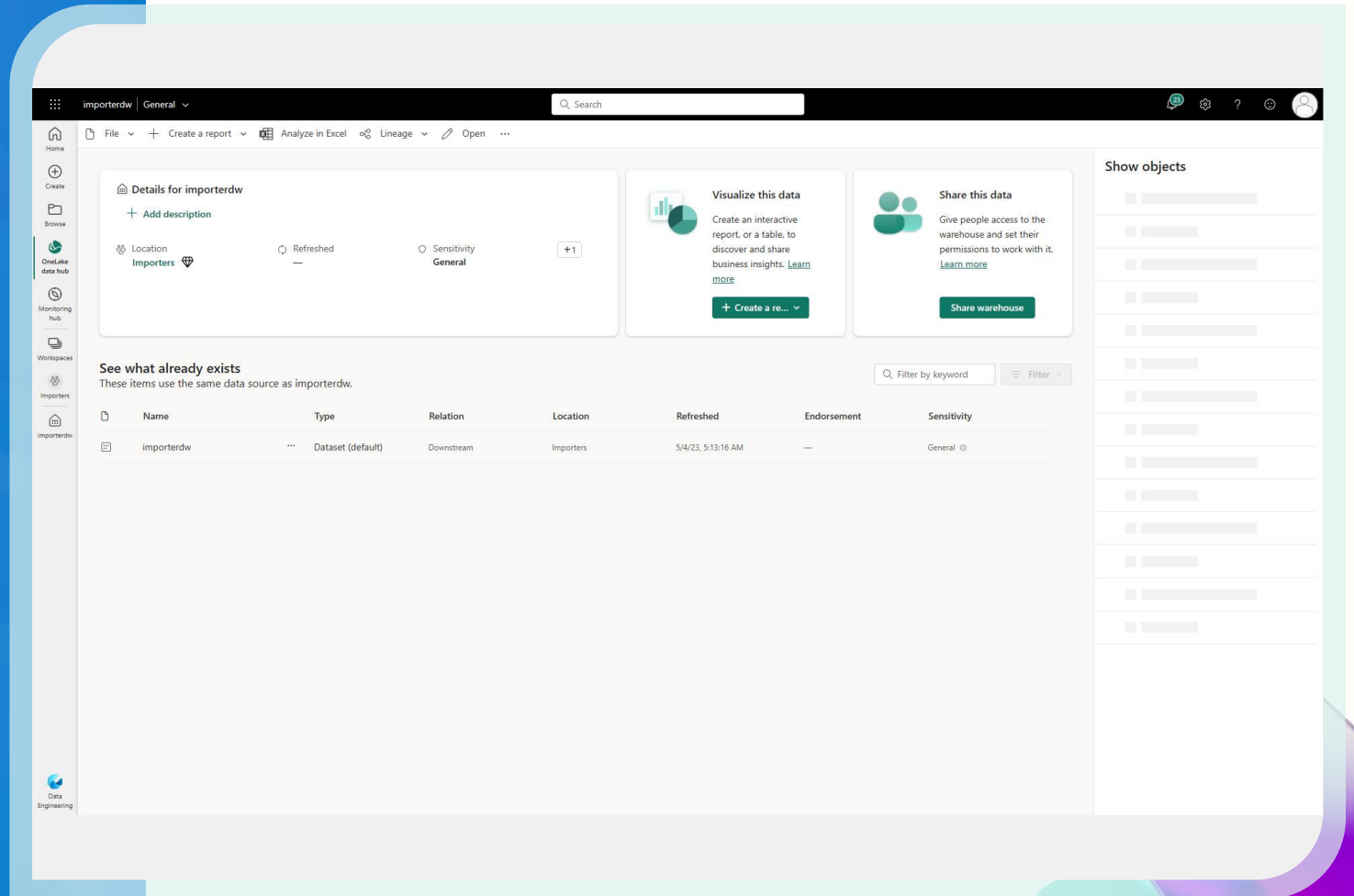
**Keep data secure** for any role accessing it and ensure peace of mind.

Customers can secure their data using familiar constructs and ensure data is only visible to those authorized to do so.



## Key Capabilities:

- **Workspace roles:** Workspace roles are used for collaboration with teams. Add users to workspace with role assignment of Admin, Member, Contributor, or Viewer
- **Artifact permissions:** Artifact permissions are used for sharing for consumption of Warehouses. Provide access and share individual Warehouses with specific permissions (coming soon)
- **Data security:** Use T-SQL, GRANT, REVOKE or DENY to secure any object within Warehouse. Users can be assigned to built-in custom roles
- **Sensitivity labels:** Apply sensitivity labels on your Warehouse to classify sensitive data



# Lab 5

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



How to configure a scheduled refresh of Dataflow Gen2



How to create a Data Pipeline



How to configure a scheduled refresh of a Data Pipeline

**pl\_Refresh\_People\_SharePoint\_Option2**  
Data pipeline

About

Endorsement

**Schedule**

The scheduled refresh is turned off

Schedule

Scheduled run

On  Off

Repeat

Daily

Time

09:00 AM

+ Add a time

Start date and time

11/05/2023

End date and time

02/29/2024

Time zone

(UTC-06:00) Central Time (US and Canada)

15 minute break



# Business Intelligence overview



# Power BI: The bridge between data and decisions



## Semantic model



Data Warehouse

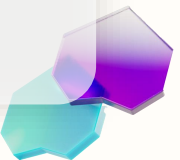


Lakehouse



## Consumption

- Web
- On-premises
- Mobile
- Embedded
- Teams
- PowerPoint
- SharePoint
- Excel
- Microsoft Power Platform
- "..."





# Fully integrated Power BI semantic layer



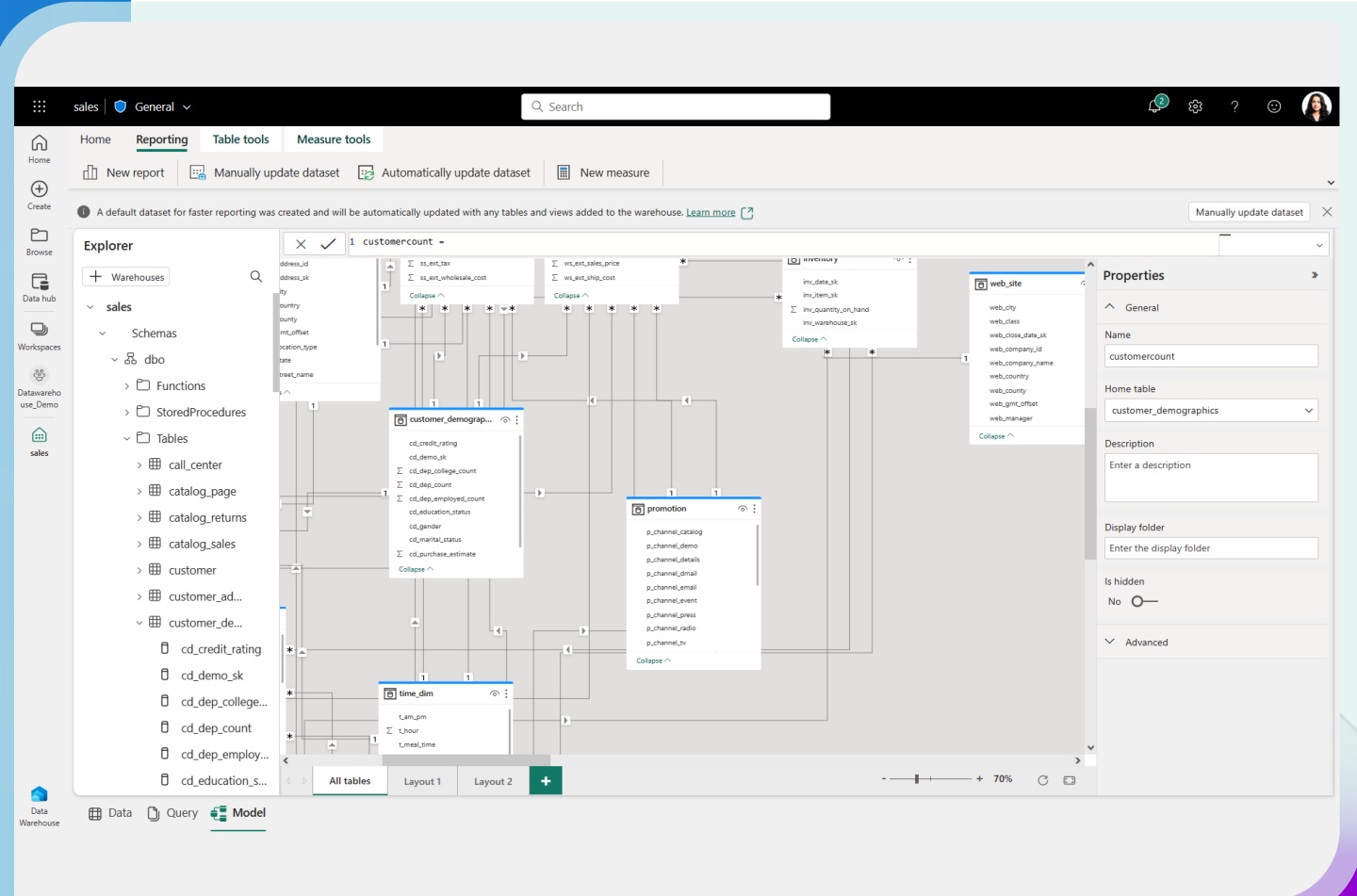
**Reduce integration** and gain insight from your data in seconds.

Built-in Power BI enables everyone to visualize their data in seconds.



## Key Capabilities:

- Auto-generated semantic model creating a Lakehouse / Warehouse
- Always in sync
- Default semantic model in Direct Lake mode but automatically switches to Direct Query or Import mode as security or performance needs change
- Flexibility to add / remove tables to semantic model
- Full web authoring experience for creating measures





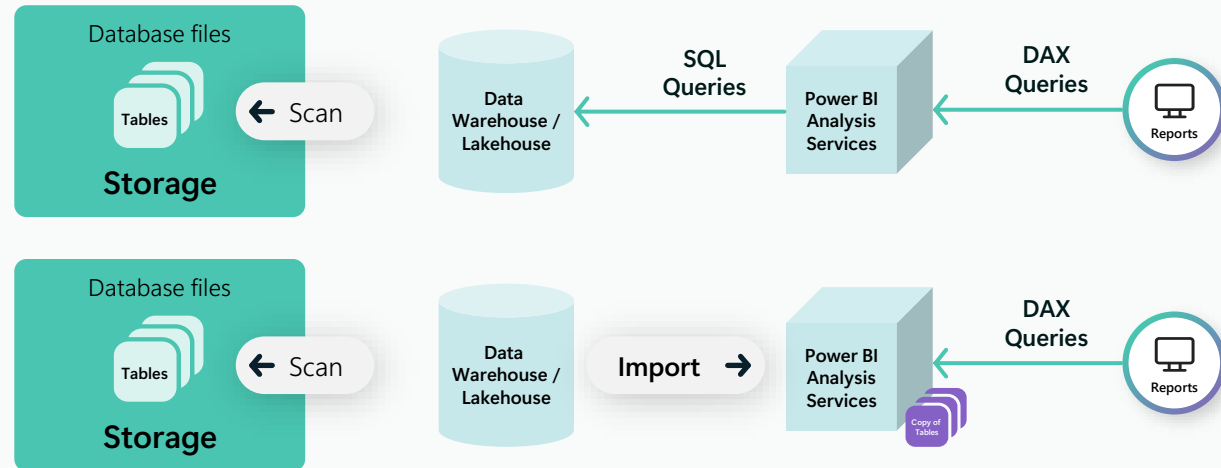
# Power BI | Direct Lake Mode



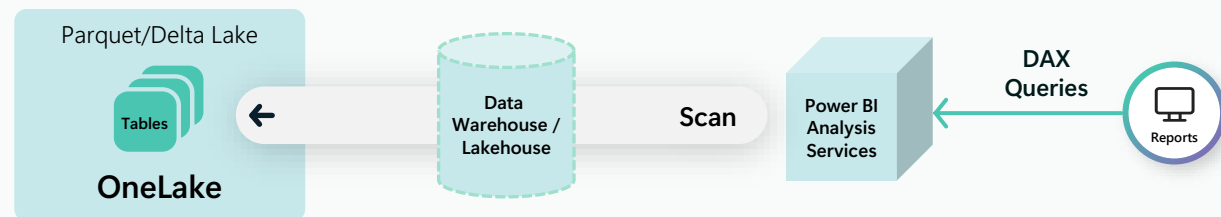
**Direct Lake** is a fast-path to load the data from the lake straight into the Power BI engine, ready for analysis.

Direct Lake is based on loading parquet-formatted files directly from a data lake without having to query a Lakehouse endpoint, and without having to import or duplicate data into a Power BI semantic model.

**Direct Query Mode.** Slow, but real time



**Direct Lake Mode.** Fast and real time



# AI-powered analytics

The most complete AI capabilities in a BI product



Information workers



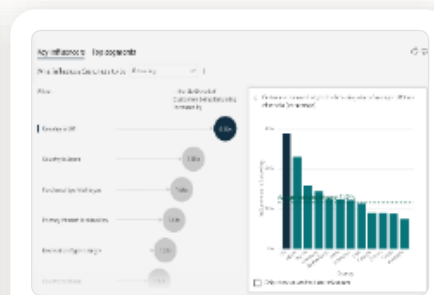
Business analysts



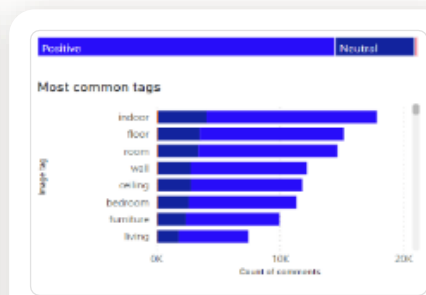
Data scientists



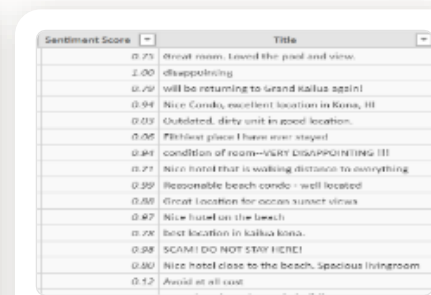
Q&A



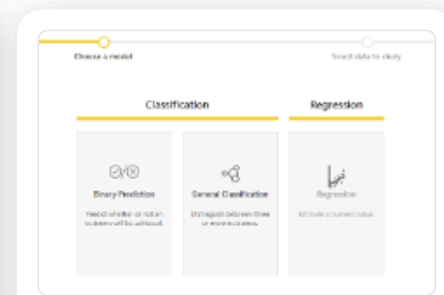
Key driver analysis



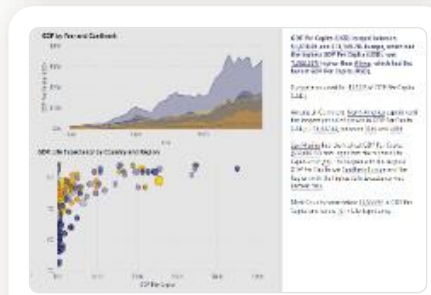
Key phrase extraction



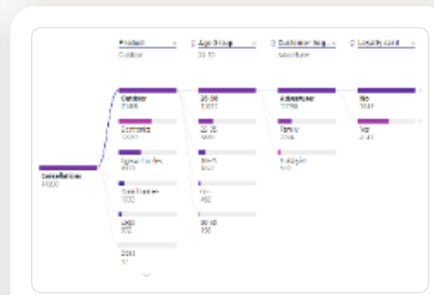
Sentiment analysis



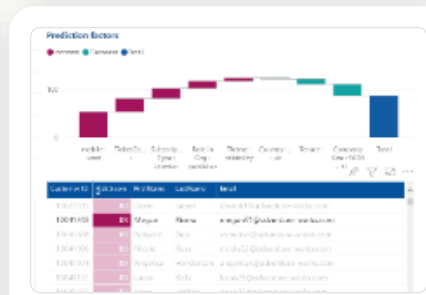
Create ML models



Smart narratives



Root cause analysis



Explore predictions



Python and R integration



Extend with Azure ML



# Track data lineage and impact analysis

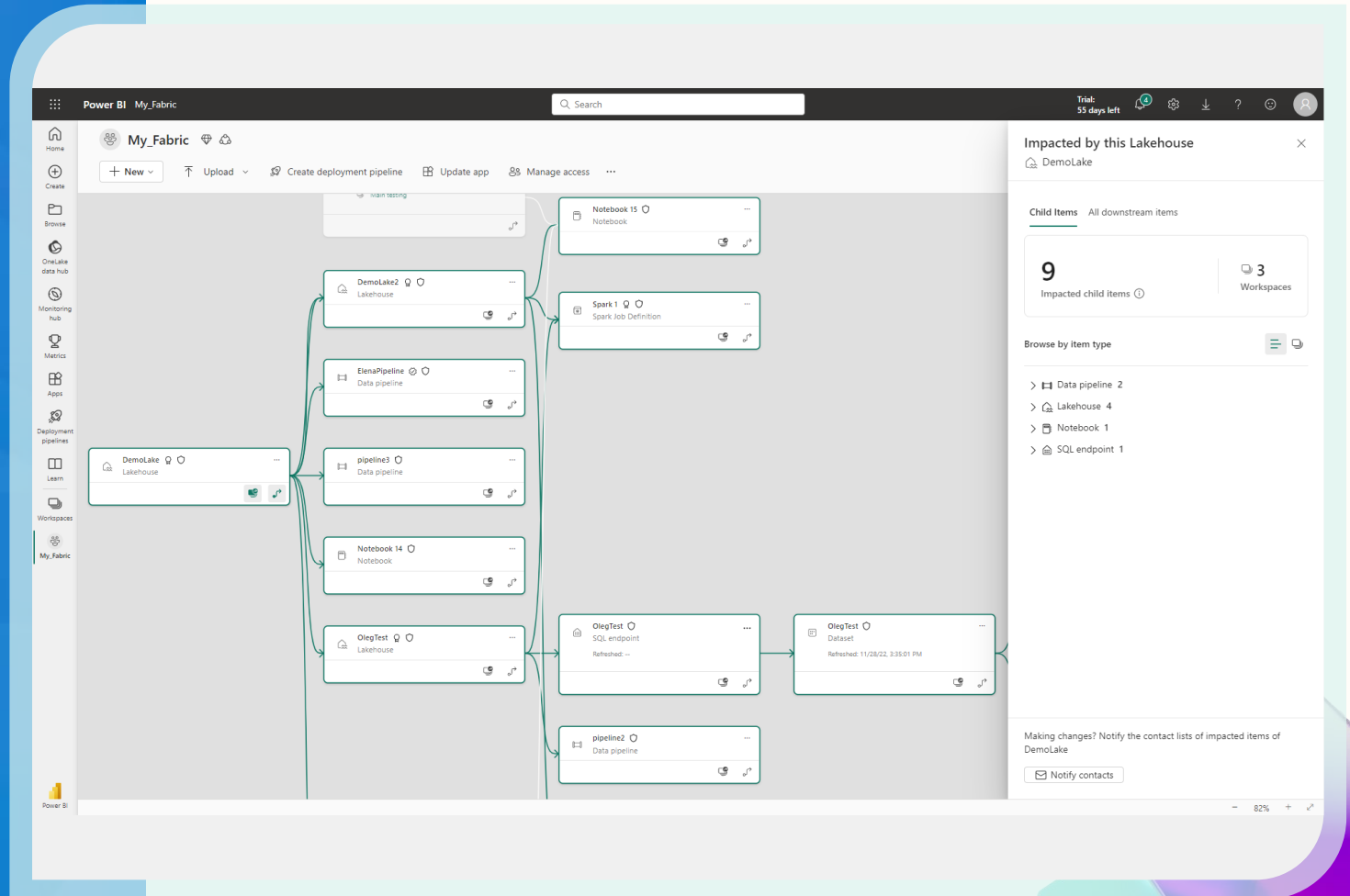


**Data owners** can gain insight and track data journey and impact analysis across their analytics projects.



## Key Capabilities:

- See lineage view of complex analytical projects
- Understand how the data flows from the source to the destination through all Fabric items, including Lakehouses, data warehouses, pipelines, semantic models, and reports
- Troubleshoot data refresh issues while tracing the data back from the destination to the source and monitor refresh status
- Gain trust in your data while understanding the source of the data
- Perform impact analysis to assess impact for upcoming changes



# Lab 6

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



How to explore Lakehouse



How to explore SQL view of Lakehouse



How to explore Data modeling in Lakehouse

The screenshot displays the Microsoft Fabric interface. At the top, a DAX formula is entered: `Order = DIVIDE([Sales], [Orders])`. Below the formula, a data model diagram is visible, showing two tables: **Reseller** and **Sales**. The **Reseller** table has columns: AccountOpenedDate, AlternateContactPersonID, BillToCustomerID, BuyingGroupID, CreditLimit, CustomerCategoryID, DeliveryAddressLine1, DeliveryAddressLine2, and DeliveryCityID. The **Sales** table has columns: ExtendedPrice, InvoiceDate, InvoiceID, InvoiceLineID, LineProfit, Quantity, ResellerID, Sales\_Amount, and SalespersonPersonID. A relationship line connects the **Reseller** table to the **Sales** table, with a '1' on the Reseller side and a '\*' on the Sales side, indicating a one-to-many relationship. The **Sales** table is highlighted with an orange border.

# Lab 7

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



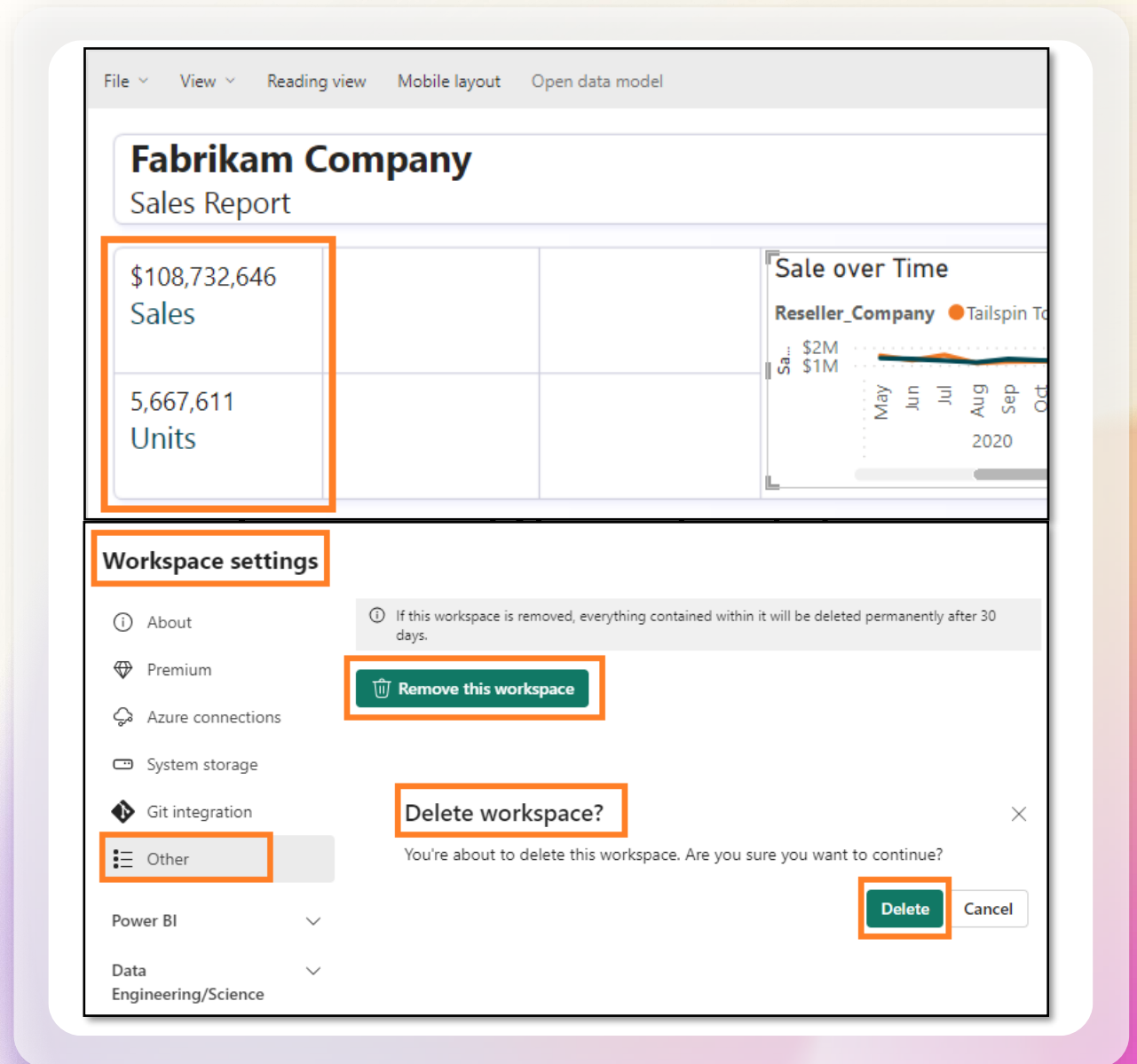
How to auto-create a report



How to build a report starting from a blank canvas



How to experience Direct Lake mode resulting in data automatically refreshing



# Next steps



Explore the product here:

<https://www.microsoft.com/en-us/microsoft-fabric>



Check out the Microsoft Mechanics show:

<https://youtu.be/oxMUJT3p9f4>



Explore the documentation:

<https://learn.microsoft.com/en-us/fabric/>



Read the free eBook: Getting started with Fabric

<https://aka.ms/fabric-get-started-ebook>

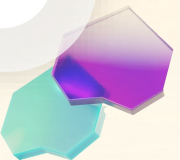
# Additional customer-facing links

## To learn more about Microsoft Fabric, consider these additional resources:

- See blog post to read the full [Microsoft Fabric GA announcement](#)
- Explore Fabric through the [Guided Tour](#)
- Sign up for the [Microsoft Fabric free trial](#)
- Visit the [Microsoft Fabric website](#)
- Learn new skills by exploring the [Fabric Learning modules](#)
- Explore the [Fabric technical documentation](#)
- Read the [free e-book on getting started with Fabric](#)
- Join the [Fabric community](#) to post your questions, share your feedback, and learn from others

## Read the more in-depth Fabric experience announcement blogs:

- [Data Factory experience in Fabric blog](#)
- [Synapse Data Engineering experience in Fabric blog](#)
- [Synapse Data Science experience in Fabric blog](#)
- [Synapse Data Warehousing experience in Fabric blog](#)
- [Synapse Real-Time Analytics experience in Fabric blog](#)
- [Power BI announcement blog](#)
- [Data Activator experience in Fabric blog](#)
- [Administration and governance in Fabric blog](#)
- [OneLake in Fabric blog](#)
- [Dataverse and Microsoft Fabric integration blog](#)





# End-to-end tutorials



## Lakehouse tutorial

<https://learn.microsoft.com/en-us/fabric/data-engineering/tutorial-lakehouse-introduction>

---



## Data Science tutorial

<https://learn.microsoft.com/en-us/fabric/data-science/tutorial-data-science-introduction>

---



## Real-Time Analytics tutorial

<https://learn.microsoft.com/en-us/fabric/real-time-analytics/tutorial-introduction>

---



## Data warehouse tutorial

<https://learn.microsoft.com/en-us/fabric/data-warehouse/tutorial-introduction>

---



## Power BI tutorial

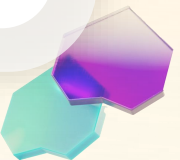
<https://learn.microsoft.com/en-us/power-bi/fundamentals/fabric-get-started>

---



## Data Factory tutorial

<https://learn.microsoft.com/en-us/fabric/data-factory/tutorial-end-to-end-introduction>



# Learn more about these experiences

Navigate to specific experiences and capabilities using the links below



## Microsoft Fabric get started

[What is Fabric?](#)  
[Microsoft Fabric licenses](#)  
[Enable Microsoft Fabric for your organization](#)  
[Navigate the Fabric portal](#)  
[Workspaces in Fabric](#)  
[See more](#)



## Power BI

[What is Power BI?](#)  
[What is a datamart?](#)  
[Azure and Power BI integration](#)  
[See more](#)



## OneLake

[What is OneLake?](#)  
[What are shortcuts?](#)  
[Create a lakehouse with OneLake](#)  
[OneLake and Azure Synapse Analytics integration](#)  
[See more](#)



## Data Factory

[What is Data Factory?](#)  
[Create your first pipeline](#)  
[Create your first dataflow](#)  
[Connectors](#)  
[See more](#)



## Synapse Data Engineering

[What is Data Engineering?](#)  
[Create a Lakehouse](#)  
[Create a Spark job definition](#)  
[See more](#)



## Synapse Data Science

[What is Data science?](#)  
[Machine learning experiment](#)  
[Use end-to-end AI samples](#)  
[See more](#)



## Synapse Data Warehouse

[What is Data Warehouse?](#)  
[Create a Warehouse](#)  
[Query using SQL query editor](#)  
[See more](#)



## Synapse Real-Time Analytics

[What is Real-Time Analytics?](#)  
[What is Event stream?](#)  
[Create a database](#)  
[See more](#)



## Data Activator

[What is Data Activator?](#)  
[Getting started](#)  
[Get data from Eventstreams](#)  
[See more](#)



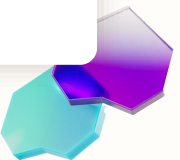
## Copilot in Fabric

[Introduction to copilot in Fabric](#) | [Enable copilot](#) | [Copilot for notebooks](#) | [See more](#)



## Security, Governance, and Administration

[Buy a Microsoft Fabric subscription](#) | [Fabric administration](#) | [Data governance and compliance](#) | [Security](#) | [See more](#)



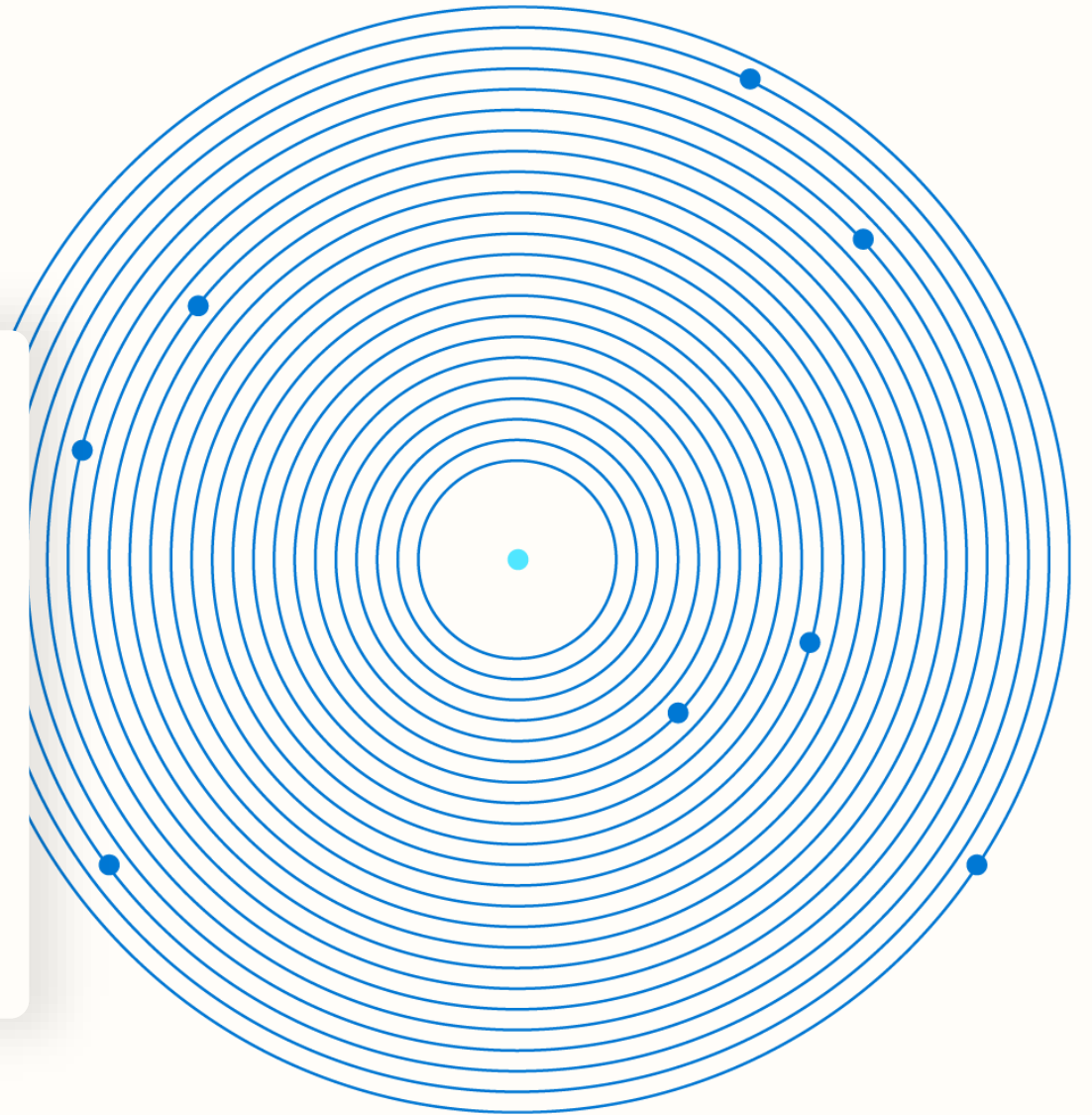
# Next steps

Get started with  
**Azure Migrate  
and Modernize &  
Azure Innovate!**

Leverage rich funding and resources available through partners via **Azure Migrate and Modernize & Azure Innovate** to simplify and accelerate your digital transformation.

[Ask us for more details](#)

[Learn more](#) 





# Appendix

# Data integration experience

Bring together low-code, AI-enabled experiences, multi-cloud connectivity, and persistent data security and governance to help solve complex scenarios for all developers

**Data Integration**  
Data Factory

**Data Engineering**  
Synapse

**Data Warehouse**  
Synapse

**Data Science**  
Synapse

**Real-Time Analytics**  
Synapse

**Business Intelligence**  
Power BI

**Data Activator**

**Serverless Compute**

**OneLake**

**Security, Governance and Administration**

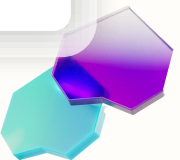
200+ native data source connectors

Cloud-scale data movement with Data Factory

Low-code interface for ingesting data from hundreds of data sources using Dataflows Gen2

Petabyte scale transformation

Powerful, enterprise-grade Data Factory experience with the best of ADF and Power Query together



# Data engineering experience

Build your data estate and empower data engineers with a world-class Spark platform, fully integrated with Data Factory, to transform and democratize data at scale

The image shows a grid of data engineering capabilities. The top row consists of seven tiles: 'Data Integration Data Factory', 'Data Engineering Synapse' (highlighted with a purple border), 'Data Warehouse Synapse', 'Data Science Synapse', 'Real-Time Analytics Synapse', 'Business Intelligence Power BI', and 'Data Activator'. Below this row are three larger horizontal tiles: 'Serverless Compute', 'OneLake', and 'Security, Governance and Administration'.

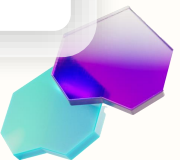
World-class Spark serverless compute

Pro and low-code authoring experience

Schedule and orchestrate data transformations with notebooks and Spark jobs

Launch clusters on demand and dynamically scale in, scale out, pause, and resume

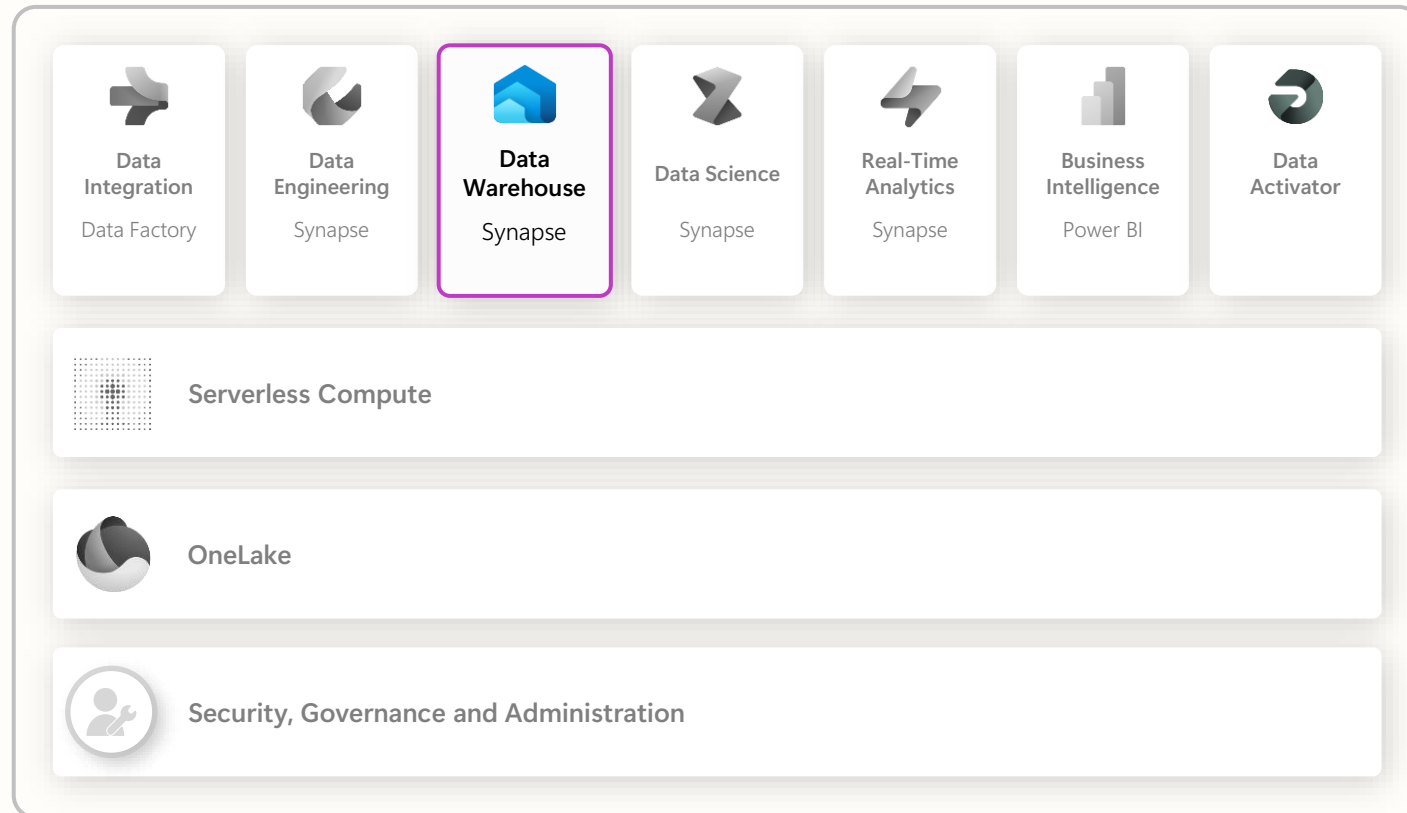
Perform code-free interactive data exploration and add to your data pipeline





# Data warehouse experience

Achieve data platform goals with ease and cost efficiency, while empowering your developers and engineers with accelerated reporting and insights



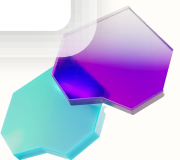
Industry-leading SQL performance and scale, including reporting tools using T-SQL and TDS end-points

Fully separate compute storage that can independently scale

Build relational layers on top of physical data in Lakehouse and natively store data in open Parquet/Delta Lake to remove need for data duplication

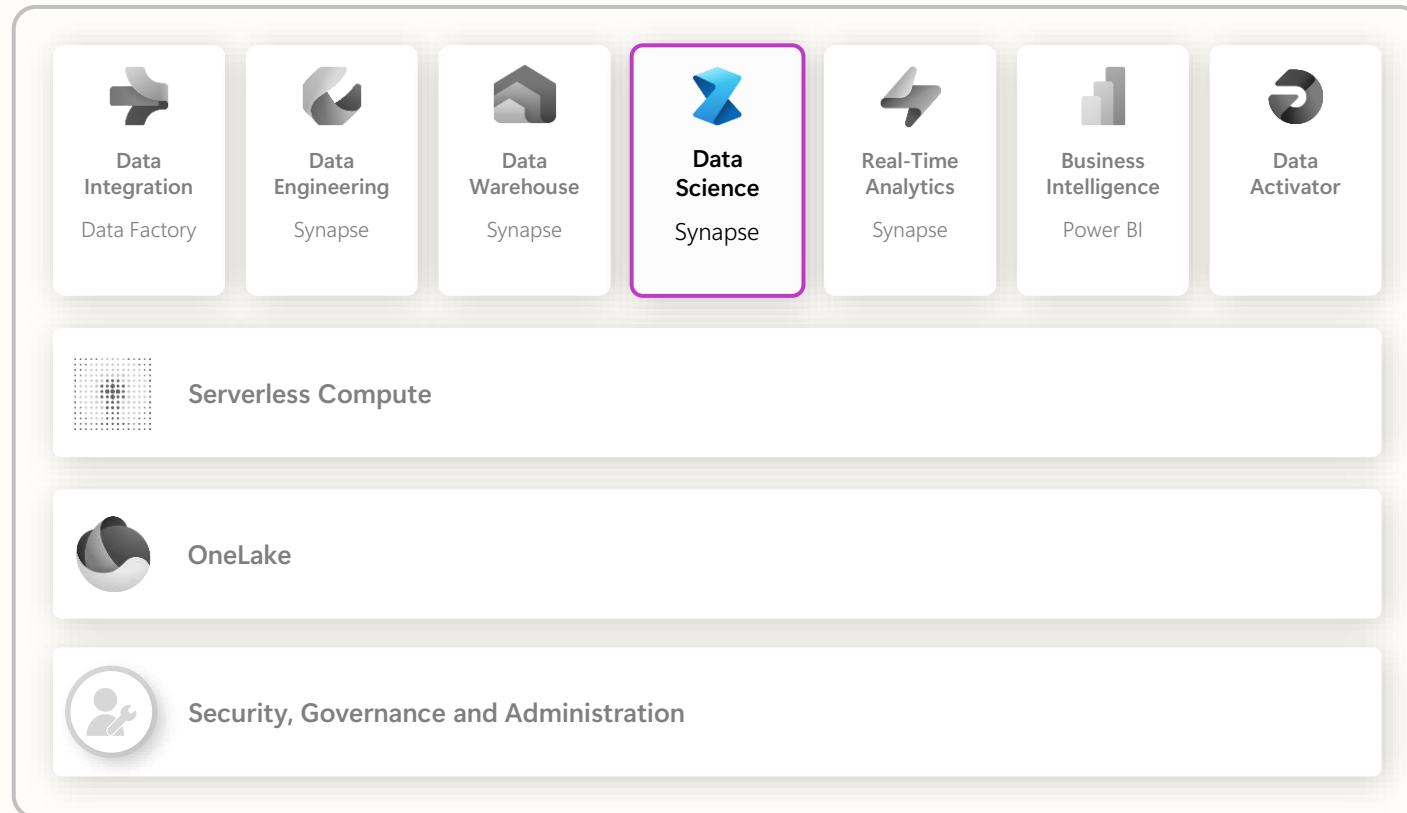
Consume data with Power BI for reporting and visualization

TLS 1.2 encryption protects all connections for granular security across your data platform



# Data science experience

Build, deploy, and operationalize machine learning models with boundaryless collaboration and seamless integration with Azure ML



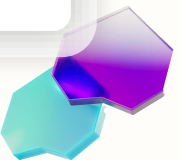
Access data from multiple sources and store data and insights in Lakehouse(s)

Leverage data science capabilities for model prediction at scale to gain and share business insights

Iterate, build, and track Machine Learning experiments using ML flow

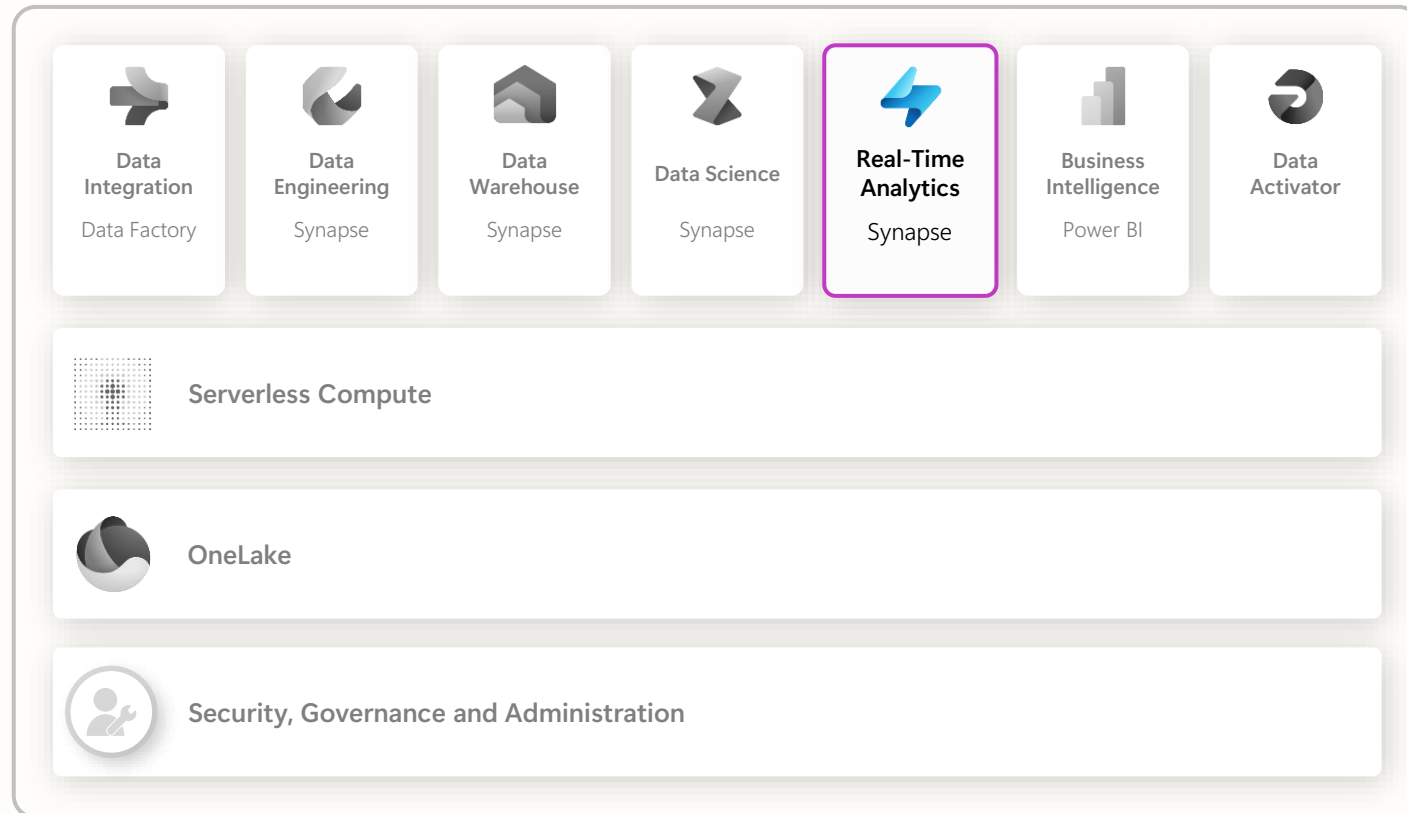
Perform transformation, exploration, and featurization by leveraging built-in experiences

Collaborate with others via Notebook, Power BI, and Lakehouses in real-time



# Real-time analytics experience

Unlock value and turn insights into actions with real-time analysis across telemetry data to better predict, optimize, and improve data applications

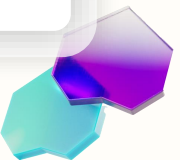


High velocity, low latency data analysis capable of indexing diverse data formats and structures up to several petabytes large

Leverage Kusto Query Language (KQL) to explore data and discover patterns, identify anomalies and outliers, and create statistical modeling

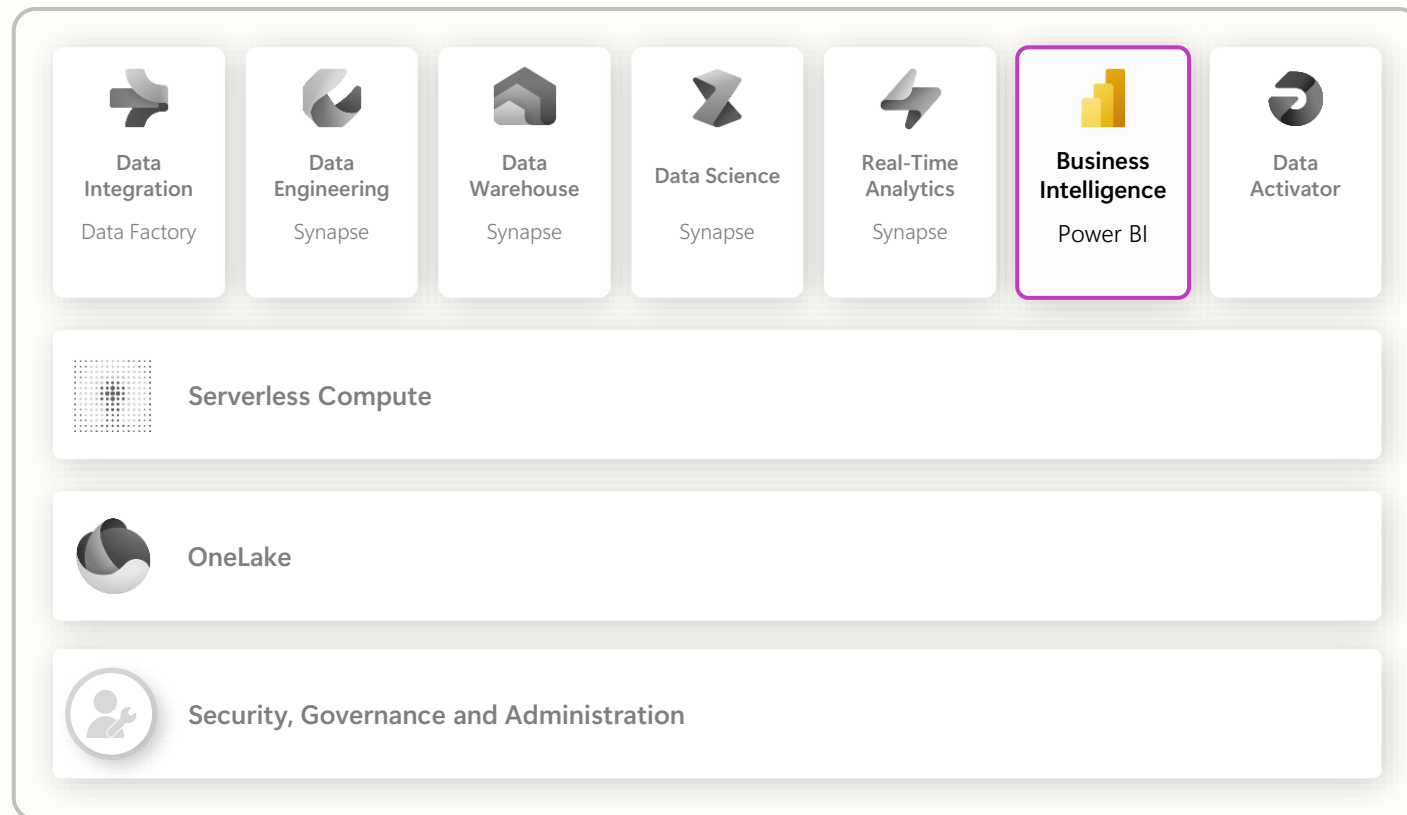
Easy ingestion of data from any source, in any data format, like applications, websites, and IoT devices

Democratize data responsibly with Microsoft Purview and use observability with Data Activator to activate timely actions



# Business intelligence experience

Uncover powerful insights with intelligence visuals, leverage data quickly and intuitively, and help achieve faster and better data-based decisions with the industry leading Power BI platform

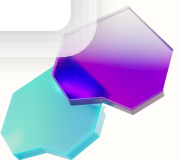


Native Integration with Microsoft 365

Powerful built-in AI capabilities and visuals illuminate hidden patterns, opportunities and anomalies in data with the click of a button

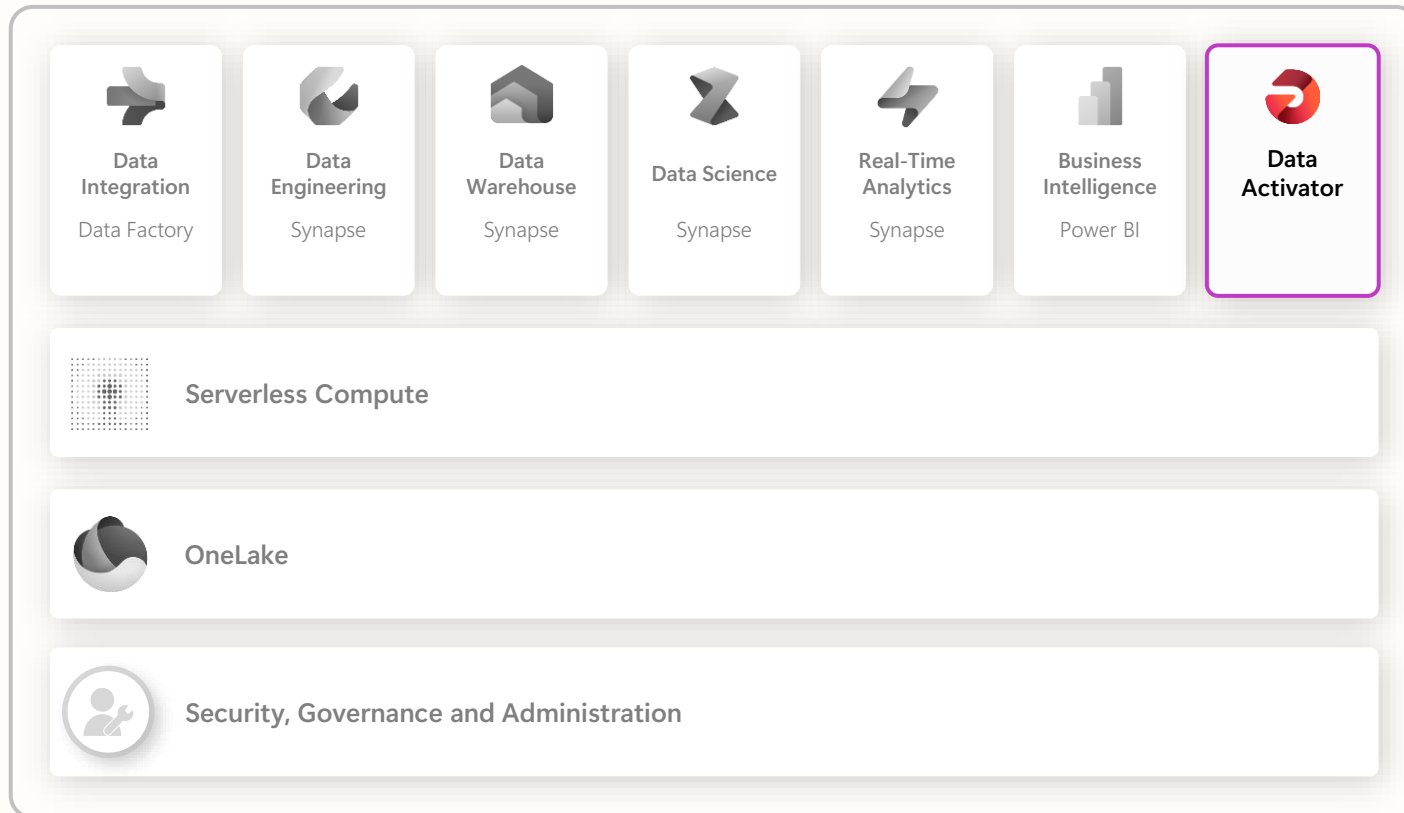
Connect to, index, and certify semantic models in the Power BI data hub

Build governed databases, like data models or data marts, in a trusted and secure hub



# Data Activator experience

Eliminate manual monitoring of operational dashboards. Automatically detect actionable conditions in your data, then drive alerts and actions, without writing code

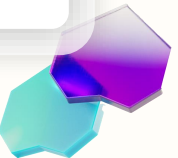


Automatically monitor Power BI semantic models and event streams.

Define no-code rules to detect actionable conditions in your data.

Send alerts in Microsoft Teams or email when actionable conditions occur.

Combine Data Activator with Power Automate to send alerts via 3<sup>rd</sup> party systems, or trigger actions in ticketing systems and other apps.



# Serverless compute

Run workloads without having to provision compute. Live compute pools are already provisioned with a workspace resulting in instant productivity.

Data Integration  
Data Factory

Data Engineering  
Synapse

Data Warehouse  
Synapse


Data Science  
Synapse

Real-Time Analytics  
Synapse


Business Intelligence  
Power BI

Data Activator


---

 **Serverless compute**

---

 OneLake

---

 Security, Governance and Administration

Limitless scale for the most demanding jobs

---

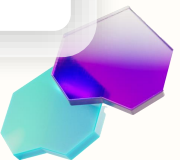
Automatically create with each workspace

---

With support for high concurrency mode there's no need to start a new session to run a notebook

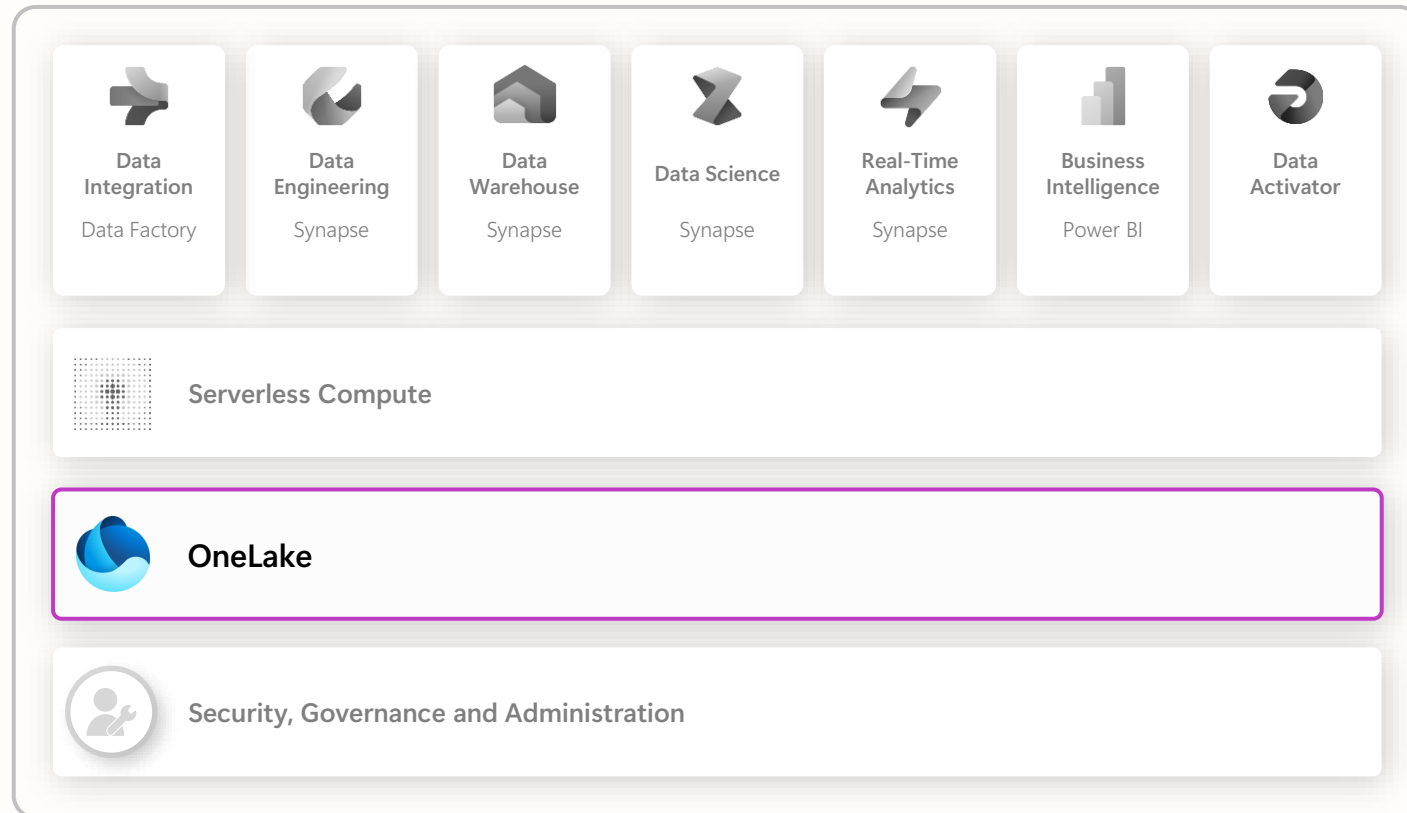
---

Memory optimized



# Unified data foundation with OneLake

Eliminate pervasive and chaotic data siloes to provide a unified, secure, and centralized storage system for developers with OneLake—the “OneDrive” for data



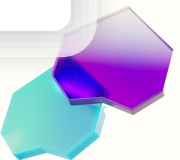
A single SaaS lake for the whole organization

Abstraction layer over Azure Storage and third-party storage services, provisioned automatically with the tenant

All workloads automatically store their data in OneLake workspace folders

All data is organized in an intuitive hierarchical namespace

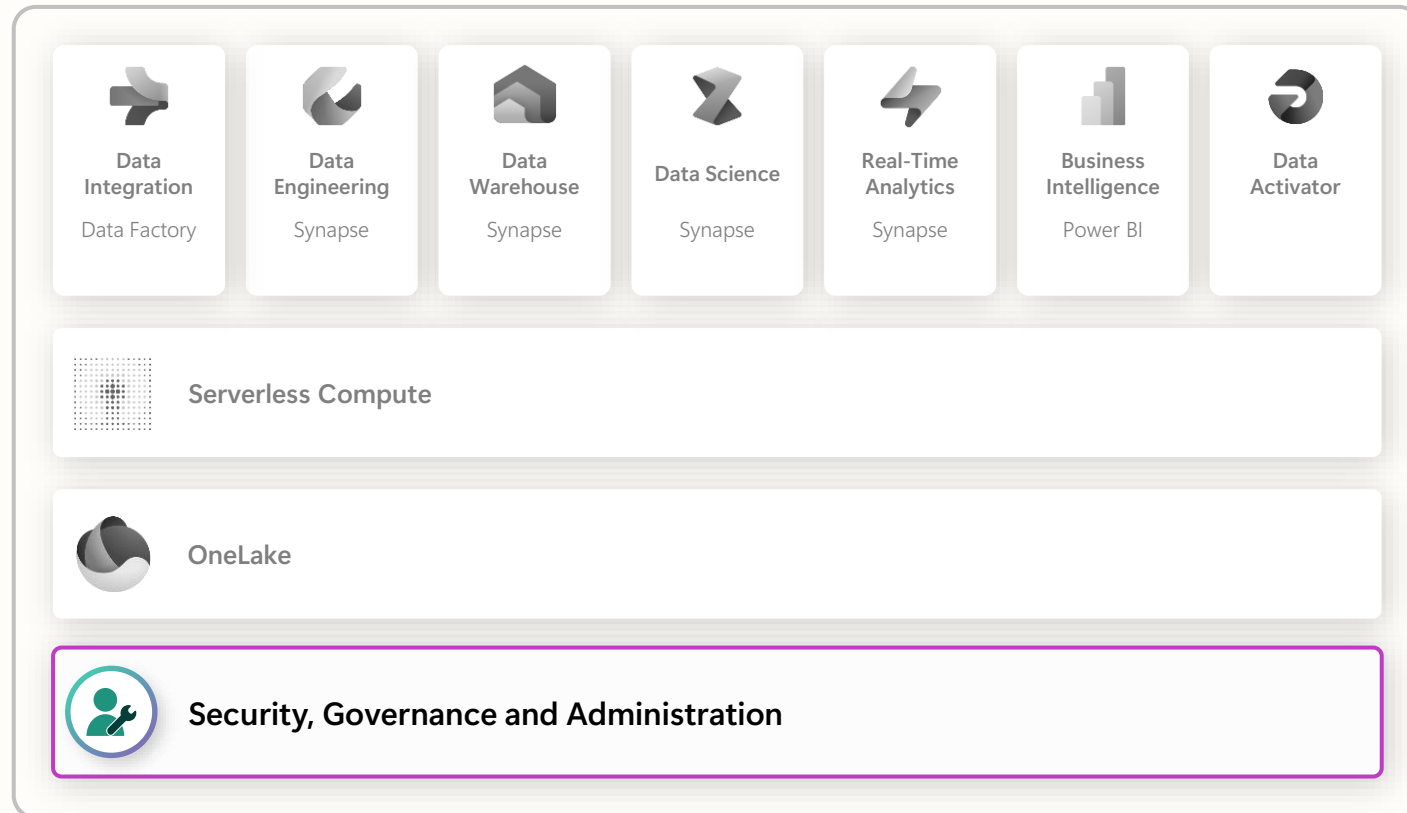
Data in OneLake is automatically indexed for discovery, sharing, governance, and compliance





# Persistent data governance and security

Built-in security and reliability features secure your data, and a set of governance and compliance capabilities help you manage, protect, and monitor sensitive information



Centralized Administration via the Admin Center

Governance powered natively by Microsoft Purview

Insights via Admin Monitoring, Capacity Insights, and more – enabling effective administration and governance

Security at Data, Workspace, and Network layers

Key enterprise promises like data residency, auditability, and much more

