

Microsoft Fabric

Fabric for the Data Analysts

Platinum
partners

creates.

In Summa

Goud
partners



Zilver
partners



Brons
partners



macaw

iqbs



illionx

valcon



Community
partners



ĐashData.



easypdash

Session Instructors



Benni De Jagere

Senior Program Manager



Daan Humblé

Cloud Solution Architect

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

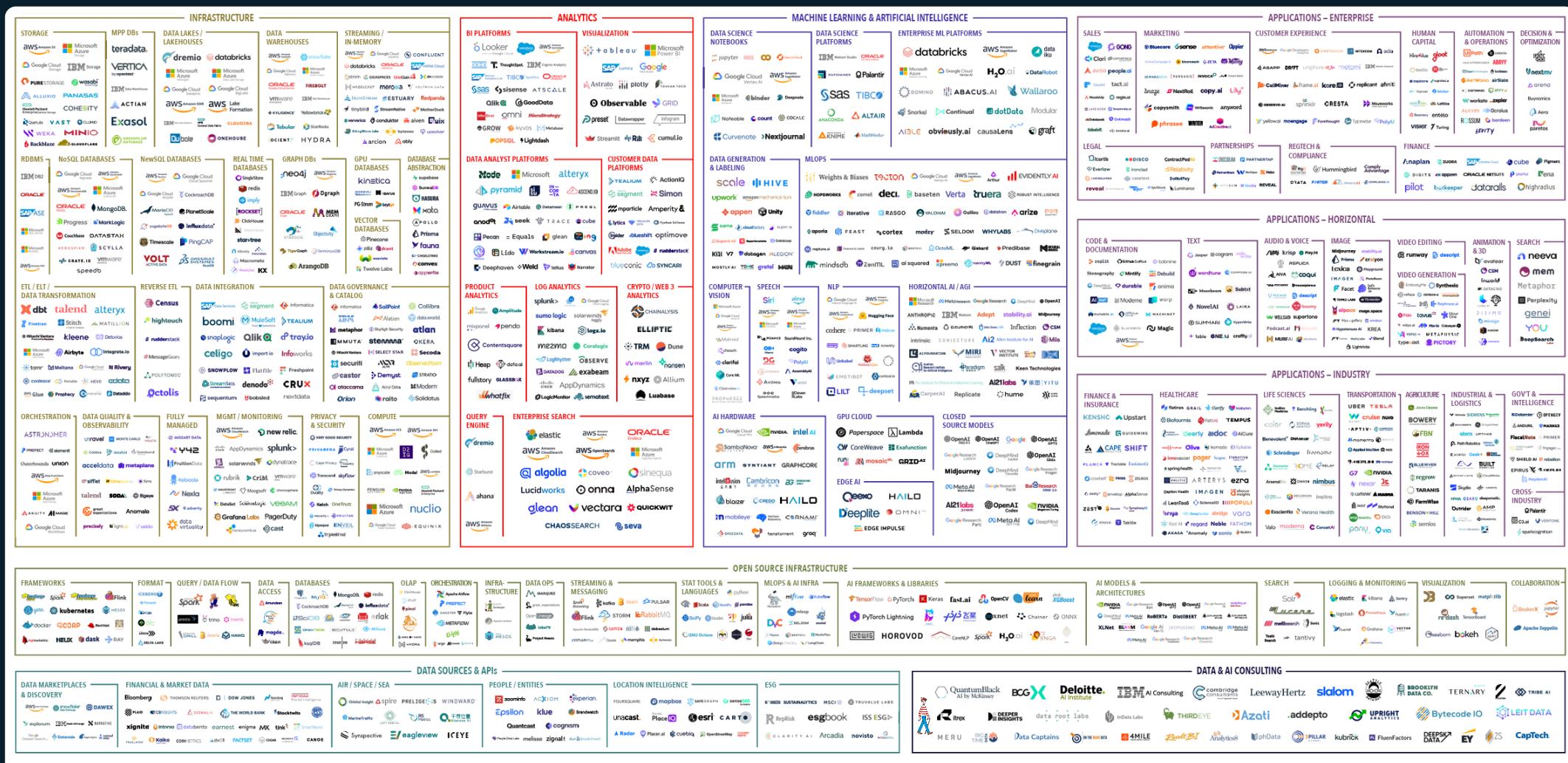
	Time	Presentation
Morning	09:00 – 9:30	Welcome, Intro, Setting the Scene
	09:30 – 10:30	Ingest to Lakehouse: Shortcuts and Data Factory
	10:30 – 11:00	Break
	11:00 – 12:00	Explore and prepare with Dataflow Gen2 and SQL
	12:00 – 13:00	Break for Lunch
Afternoon	13:00 – 13:30	Notebooks & Data Wrangler
	13:30 – 14:15	Serve with Semantic models
	14:15 – 15:00	Copilot for Microsoft Fabric
	15:00 – 15:30	Break
	15:30 – 16:00	Visualizations
	16:00 – 17:00	Wrap up, Kahoot & Open Q&A

Introducing the
new Microsoft
Fabric solution





The 2023 Data and AI landscape



Version 1.0 - Feb 2023

© Matt Turck (@mattturck), Kevin Zhang (@kevinzhang) & FirstMark (@firstmarkcap)

Blog post: mattturck.com/MAD2023

Interactive version: MADfirstmarkcap.com

Comments? Email MAD2023@firstmarkcap.com

FIRSTMARK ■
EARLY STAGE VENTURE CAPITAL

<https://mad.firstmark.com/>

Analytics today

Analytics has very predictable patterns.

Data Integration

Data Engineering

Data Warehousing

Real-Time Analytics

Data Science

Business Intelligence

Data Lake

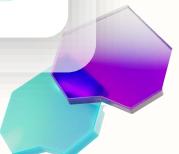
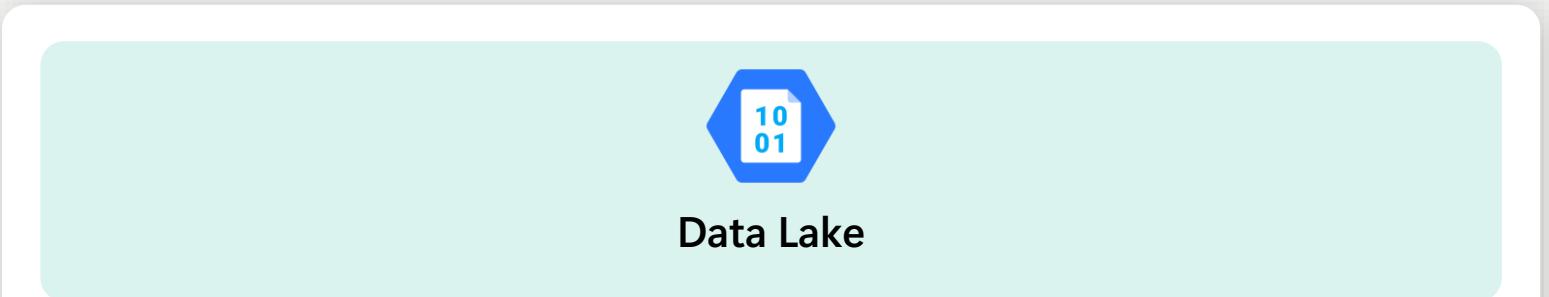
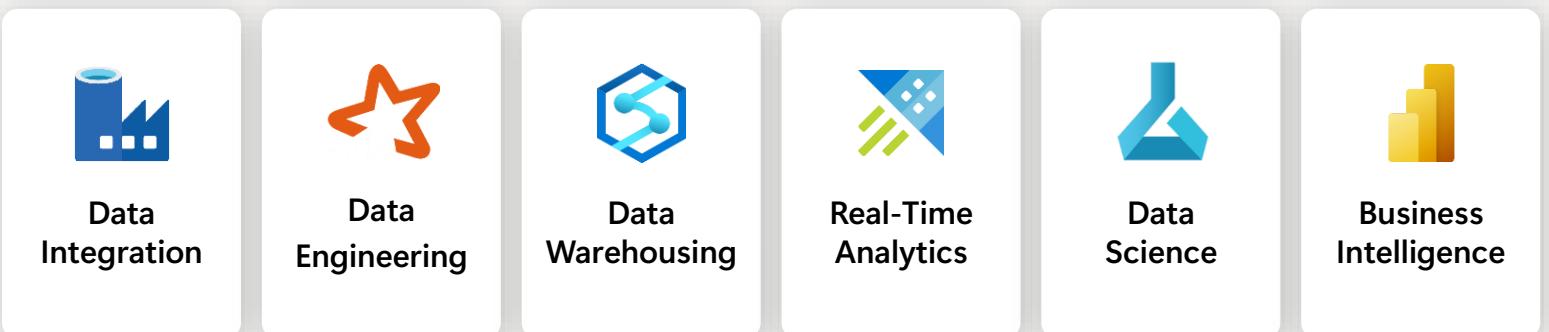
Governance and Administration



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

Different Experiences

Proprietary and Open

Dedicated and Serverless

PaaS and SaaS

Different Business Models

Steep Learning Curves

Deep Expertise Needed

High Integration Effort



Purview



Power BI



Kusto



Data Factory



Azure AI

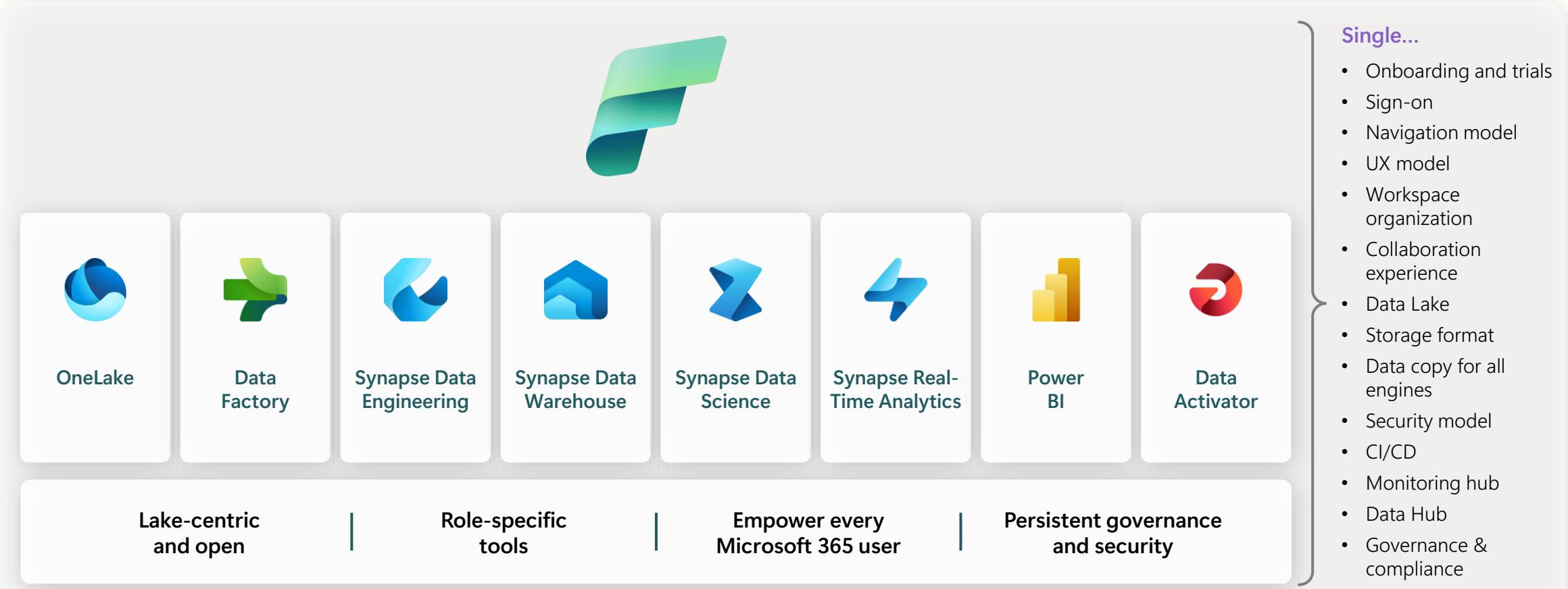


Synapse DW



Synapse Spark

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





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

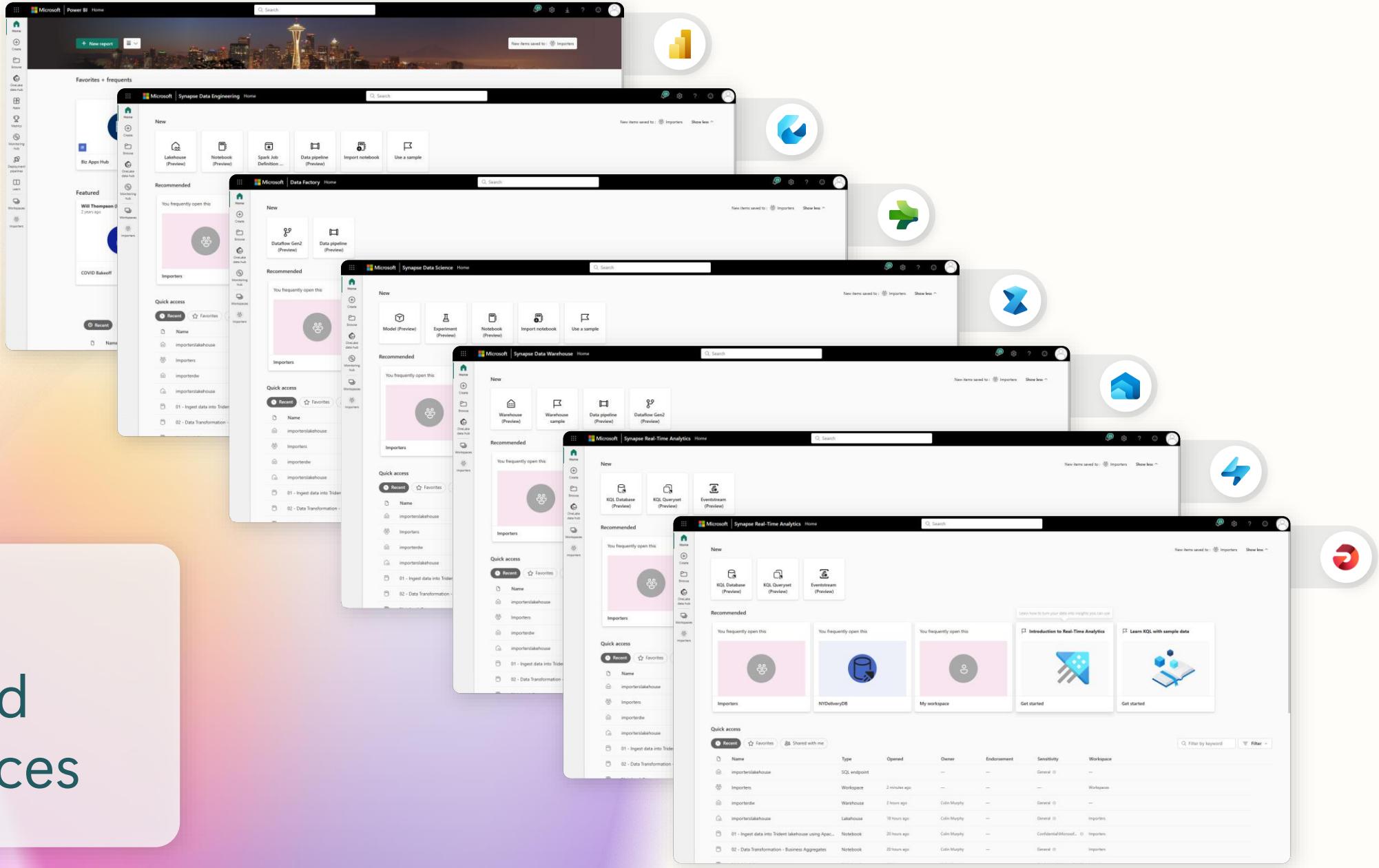
AI Powered

Copilot accelerated

ChatGPT on your data

AI driven insights

Persona optimized experiences



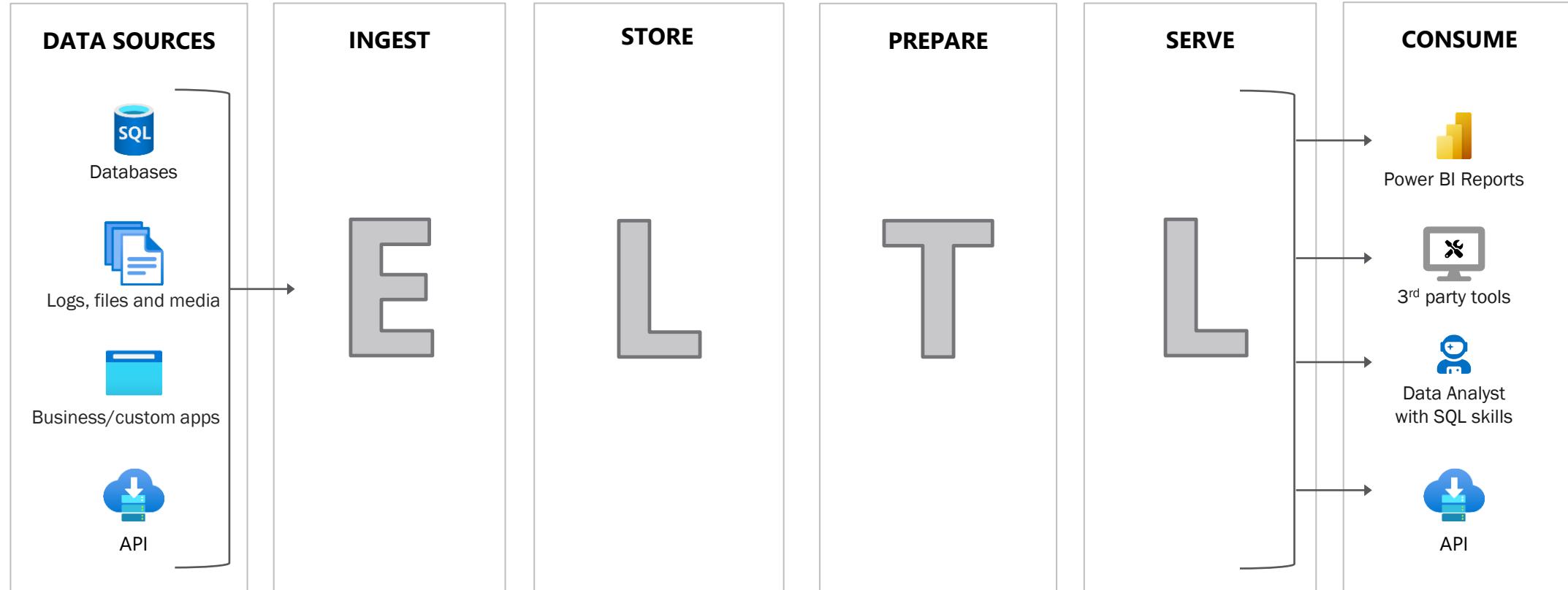


Separating Development

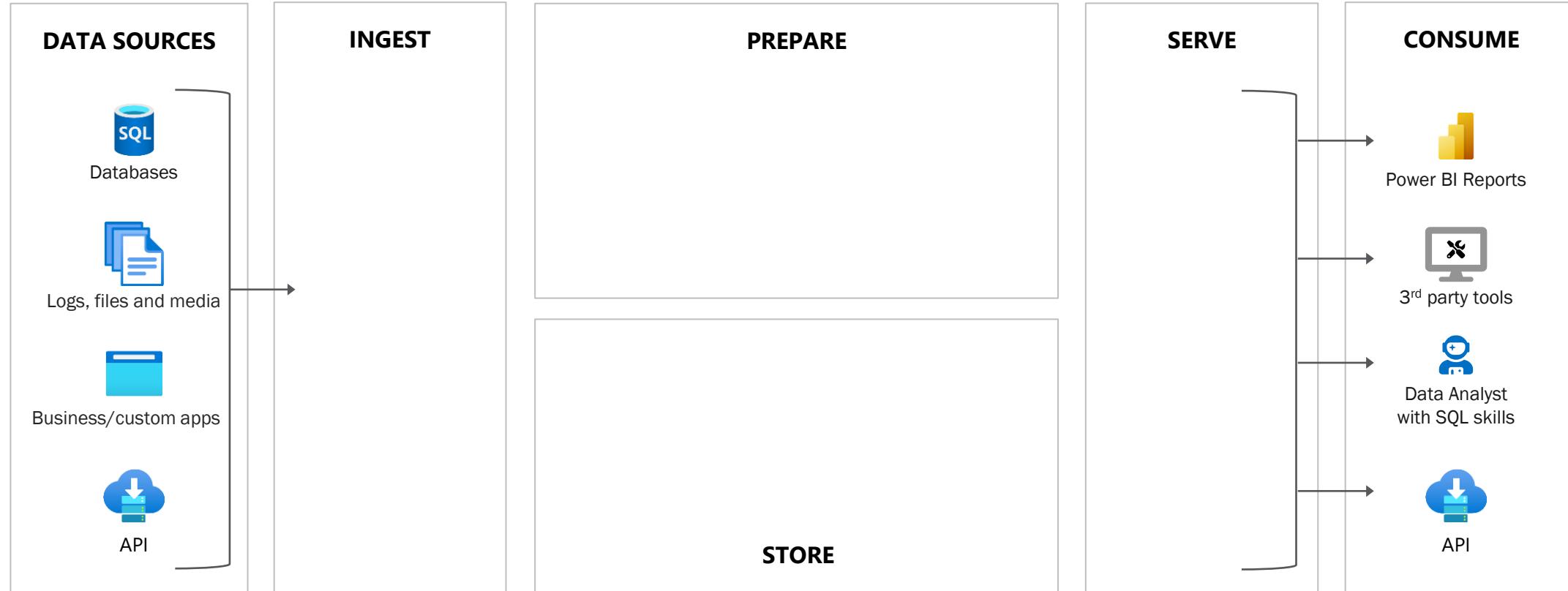
For enterprise scale deployments, it's recommended to separate the data preparation, semantic model, and report development from a single file solution.

This technique allows different solution developers to work independently and collaboratively by sharing components.

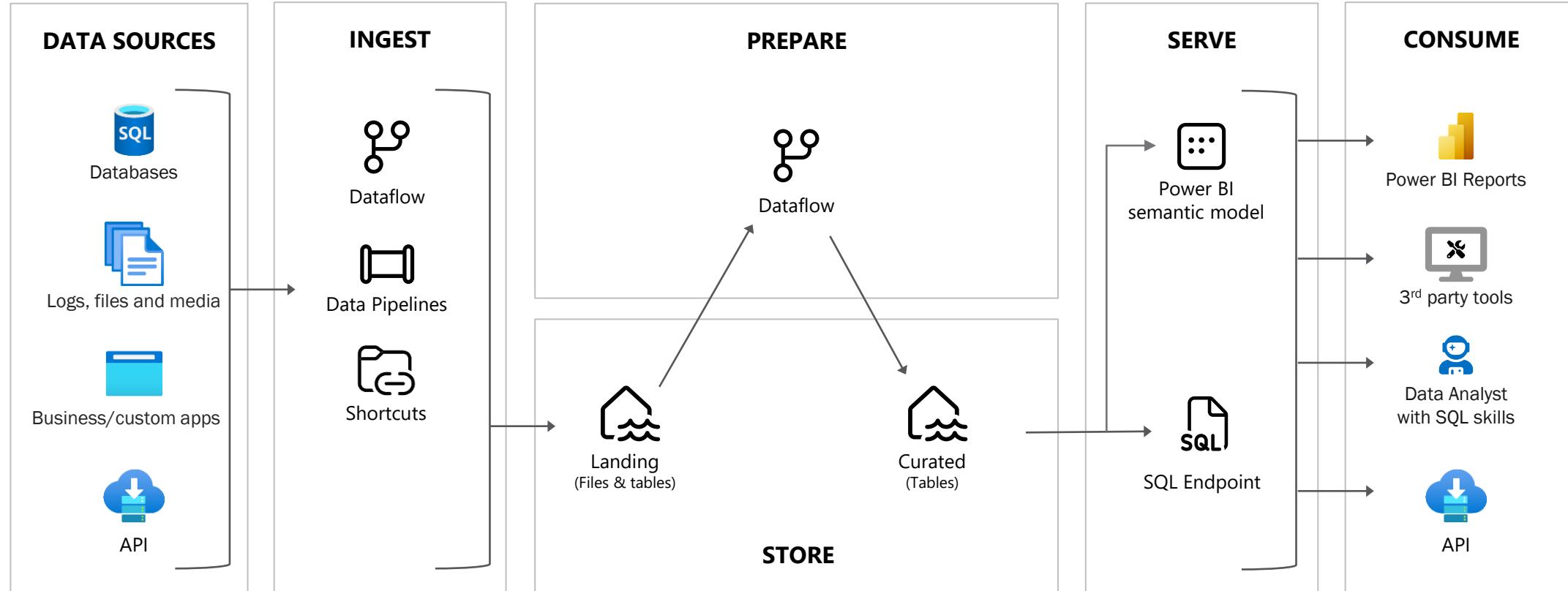
Dataflow in a Modern Data Warehouse



Dataflow in a Data Lakehouse in Fabric



Dataflow in a Data Lakehouse in Fabric



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

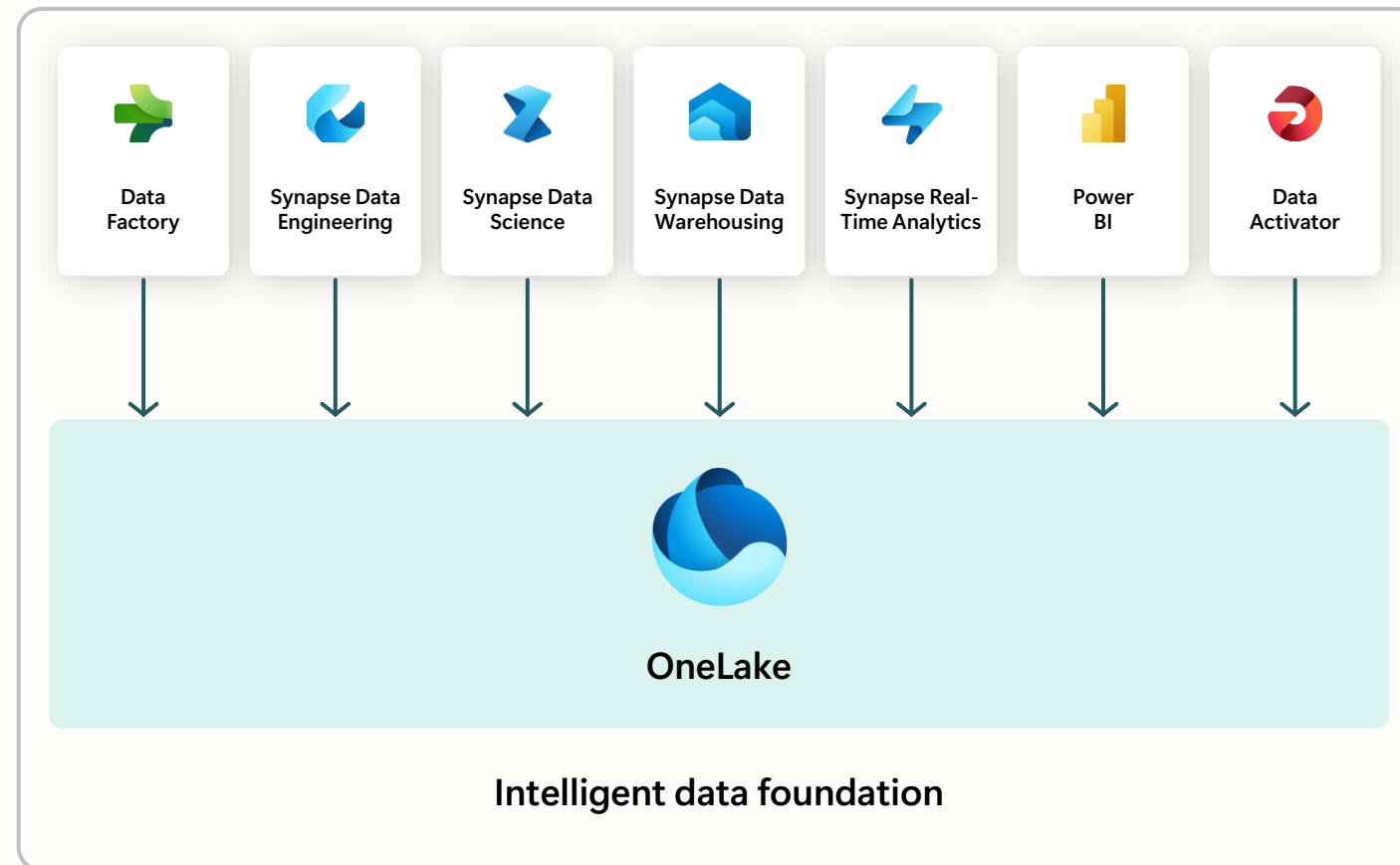


OneLake overview



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

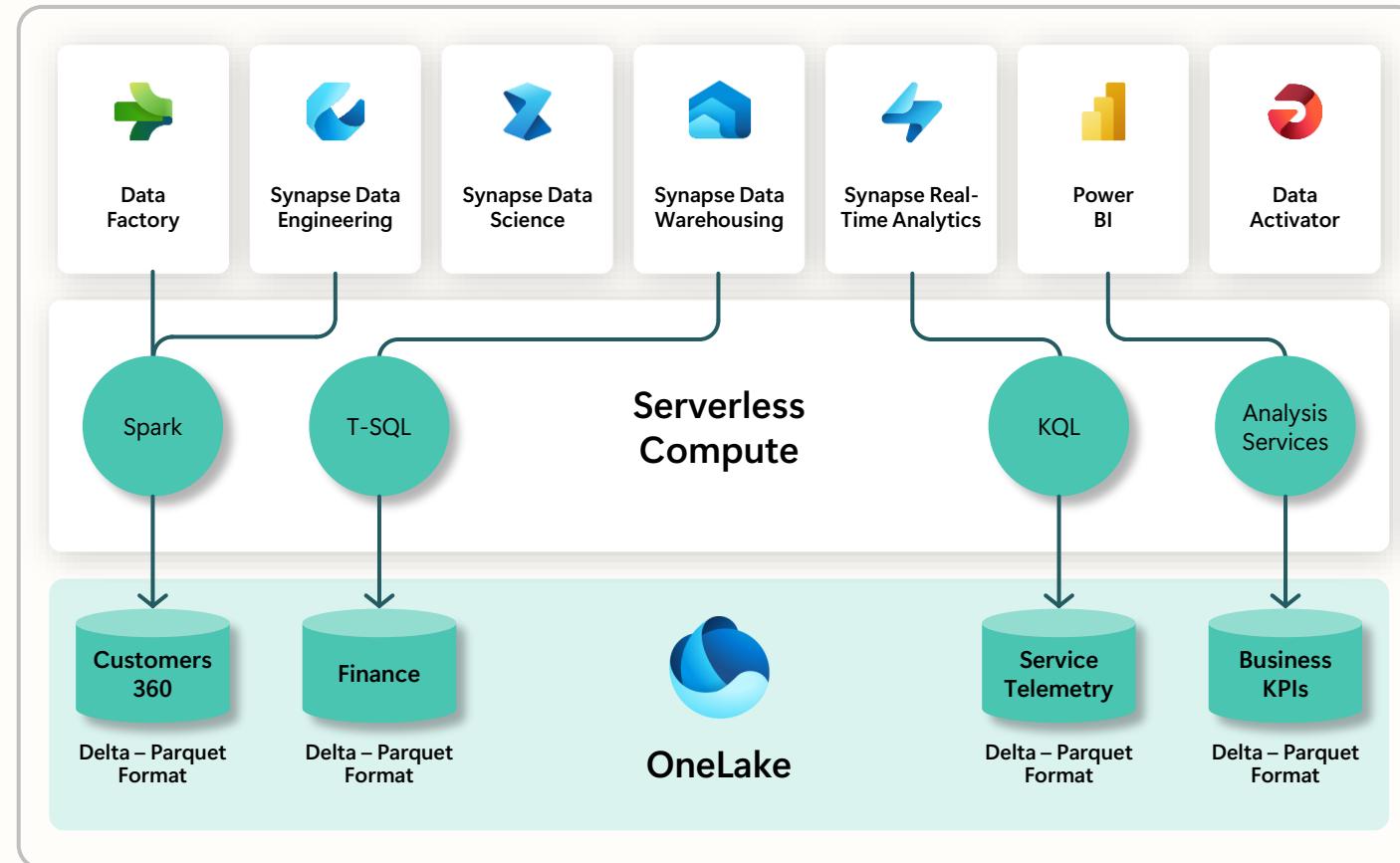
All the data is organized in an intuitive hierarchical namespace

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



One Copy for all computers

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

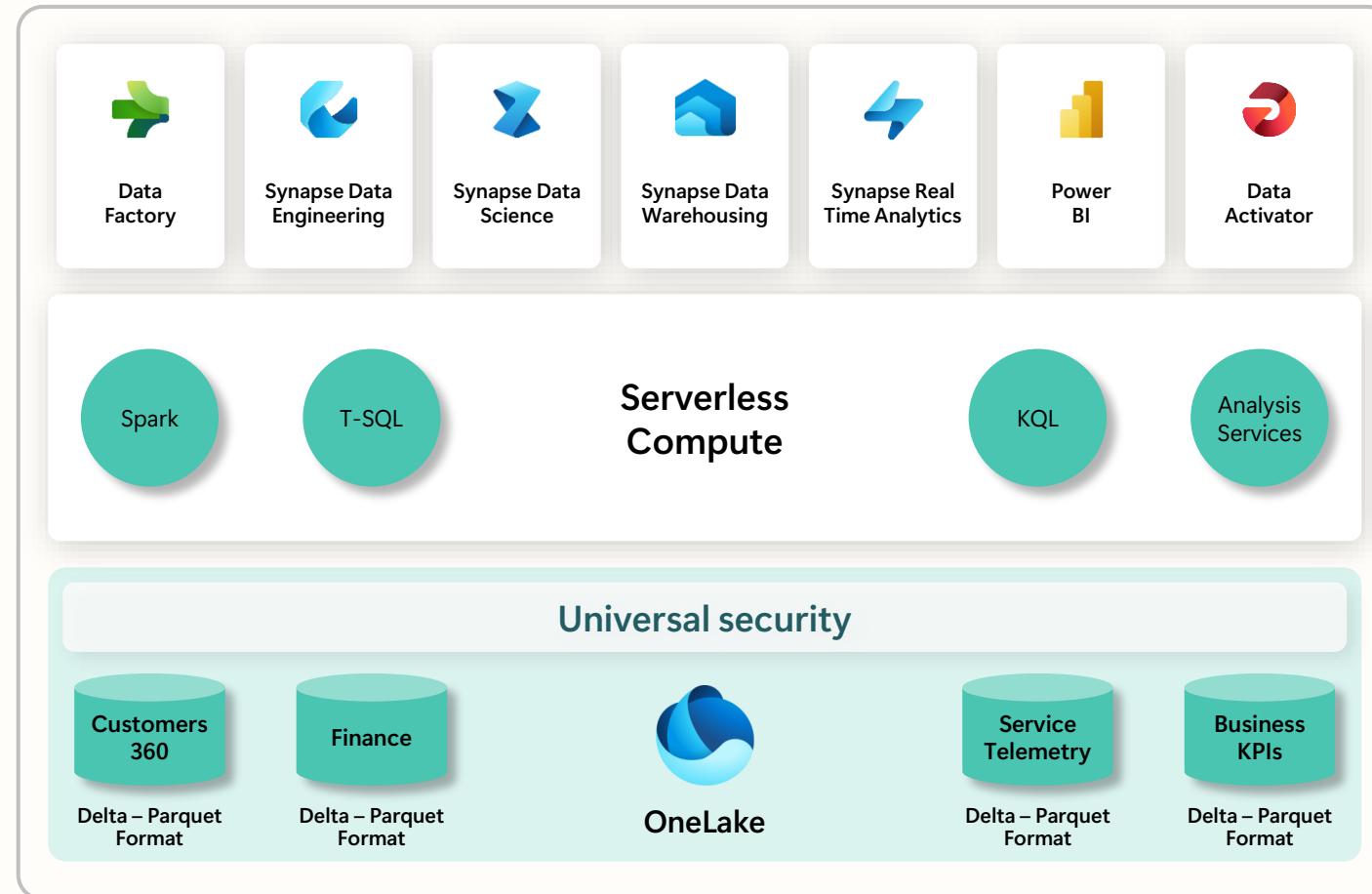
Once data is stored in the lake, it is directly accessible by all the engines without needing any import / export

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 computers

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

OneLake for all domains

A true data mesh across organization domains

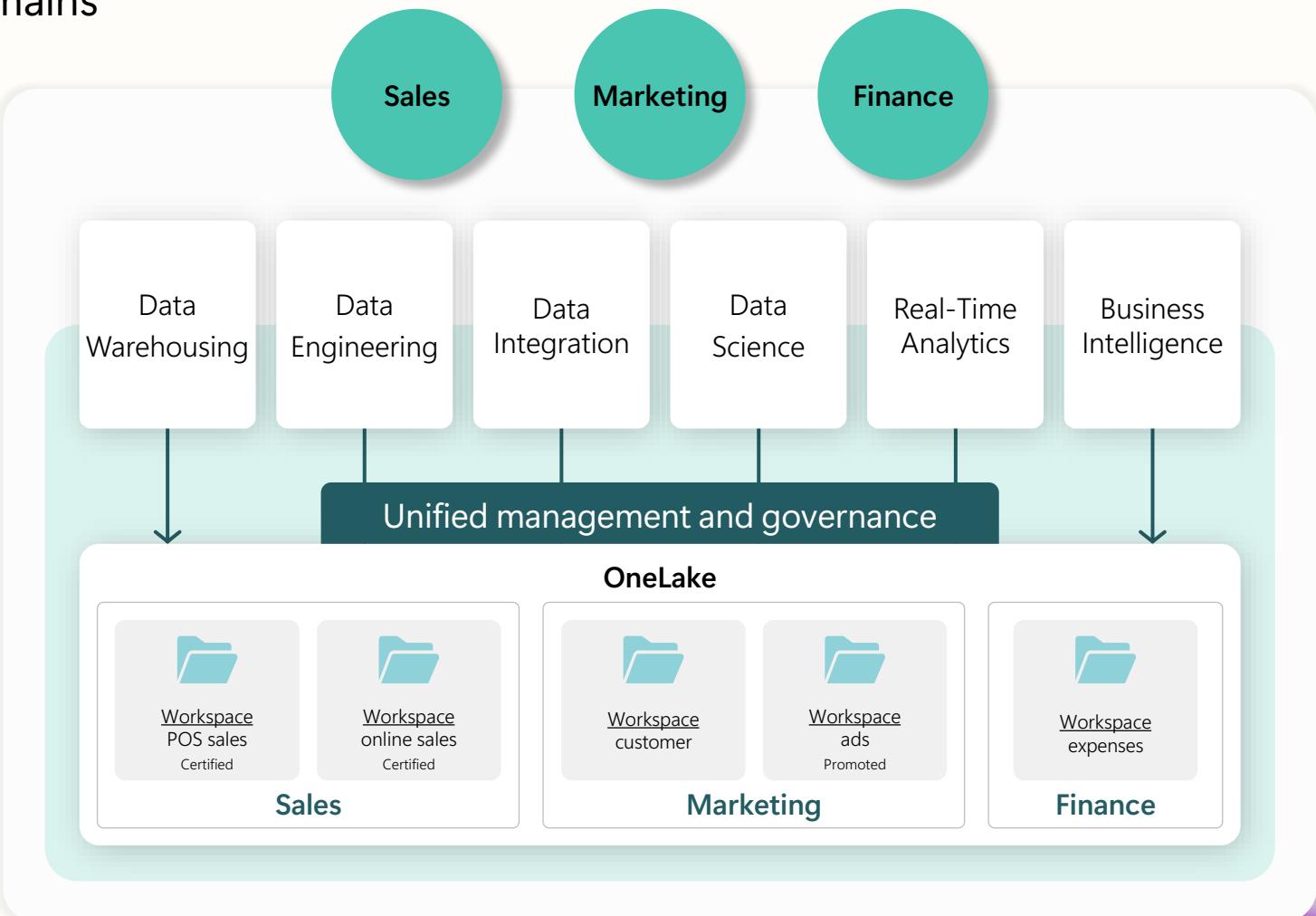
Introducing **domains** as an integral part of Fabric:
A domain is a way to logically group **together** all
the data in an organization relevant to an area or
field, according to business needs

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

The screenshot shows the Microsoft Fabric OneLake data hub interface. On the left, a sidebar lists various fabric components: Home, Create, Browse, OneLake data hub (selected), Apps, Metrics, Monitoring hub, Deployment pipelines, Learn, Workspaces, and Contoso CRM... Below this is a Power BI icon. The main area has a search bar at the top. A banner at the top says "Discover, manage, and use data from across your org. Learn more about OneLake data hub". A "Fabric Trial 59 days left" message is also present. The interface is divided into two main sections: "Recommended" datasets and an "Explorer" table.

Recommended Datasets:

- Business Dataset
- US Retail Sales
- Contoso Dataset
- Supplier-Quality-Analysis-Sample...
- Operating Report

Explorer Table:

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-Sample...	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...

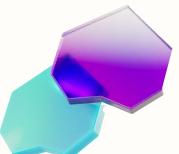
Central location within Fabric to discover, manage, and reuse data. Pervasive experience – available everywhere users discover data (create shortcuts, get data in dataflow, attach a notebook, create a semantic model, and more)

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

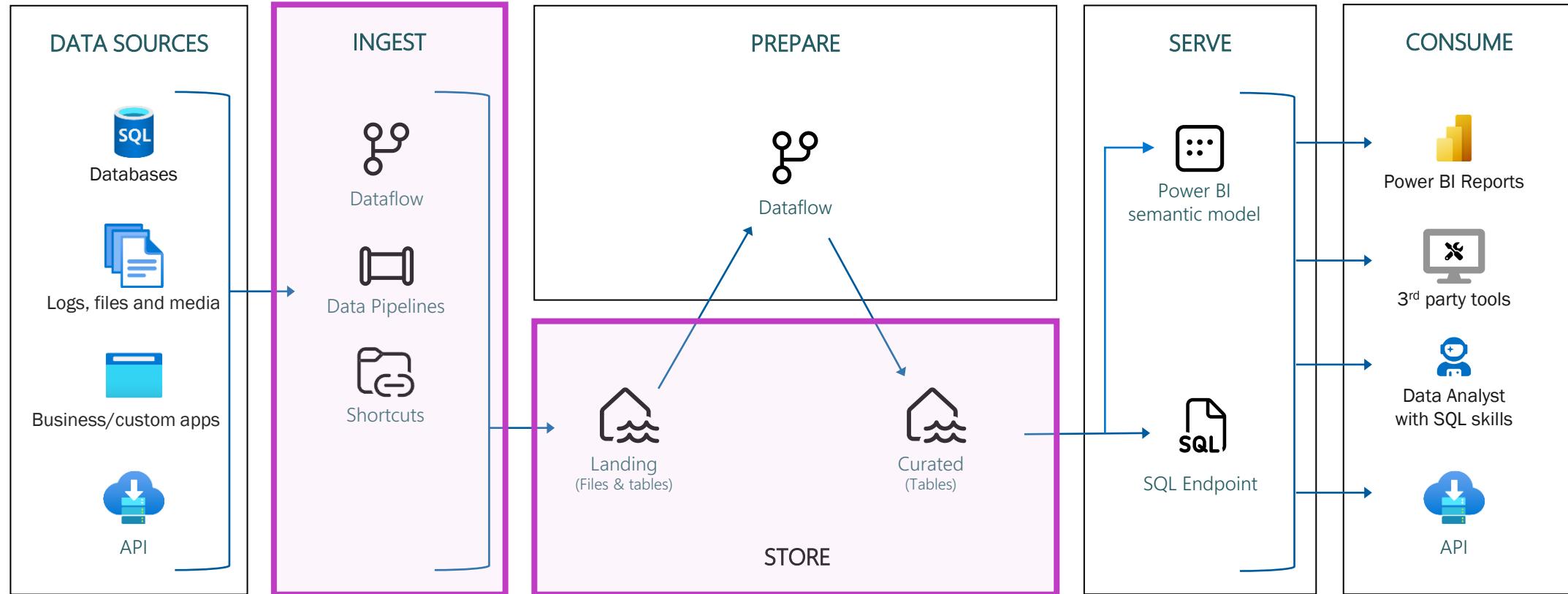
Available in Microsoft Teams, thus bringing OneLake Data hub to Office, enabling both technical and non-technical users with discovery, reuse, and exploration



Data Ingestion

Lakehouse, Data Factory,
and Shortcuts

Ingest Data in Fabric



Lakehouse



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
- Easily ingest data into the Lakehouse through a variety of methods
- Share your Lakehouse as a data product with consumers
- Use 'direct lake' mode to build reports in seconds directly on top of the data lake with blazing fast performance

The screenshot shows the Microsoft Fabric Data Explorer interface. On the left, the sidebar navigation includes Home, Create, Browse, OneLake data hub, Monitoring hub, Workspaces, Importers, and Importers Lakehouse. The Importers Lakehouse section is currently selected. The main area displays a file tree under 'importerslakehouse'. The 'fact_sale_ly_full' folder is expanded, showing its contents. A list of Parquet files is displayed in a table format:

Name	Date modified	Type	Size
_SUCCESS	4/24/2023 6:58:06 PM	-	0 B
part-00000-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:09 PM	PARQUET	22 MB
part-00001-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:09 PM	PARQUET	26 MB
part-00002-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:09 PM	PARQUET	19 MB
part-00003-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:09 PM	PARQUET	19 MB
part-00004-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:09 PM	PARQUET	33 MB
part-00005-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:10 PM	PARQUET	20 MB
part-00006-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:11 PM	PARQUET	36 MB
part-00007-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:09 PM	PARQUET	23 MB
part-00008-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:11 PM	PARQUET	24 MB
part-00009-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:11 PM	PARQUET	24 MB
part-00010-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:10 PM	PARQUET	23 MB
part-00011-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:11 PM	PARQUET	31 MB
part-00012-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:11 PM	PARQUET	22 MB
part-00013-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:11 PM	PARQUET	21 MB

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

The screenshot shows the Microsoft Data Factory user interface. At the top, there's a navigation bar with tabs like Home, Activities, Run, View, and specific pipeline details (pipeline5, Confidential/Microsoft Extended). Below the navigation is a toolbar with icons for Validate, Run, Schedule, View run history, Copy data, Dataflow, Notebook, Lookup, and Invoke pipeline. On the left, a sidebar lists various components: Home, Create, Browse, OneLake data hub, Monitoring hub, Workspaces, Importers, and pipeline5 (which is selected). The main area is titled "Start building your data pipeline" and contains three buttons: "Add pipeline activity" (with a gear icon), "Copy data" (with a clipboard icon), and "Choose a task to start" (with a square icon). The bottom right corner features a large, stylized purple and blue geometric shape.

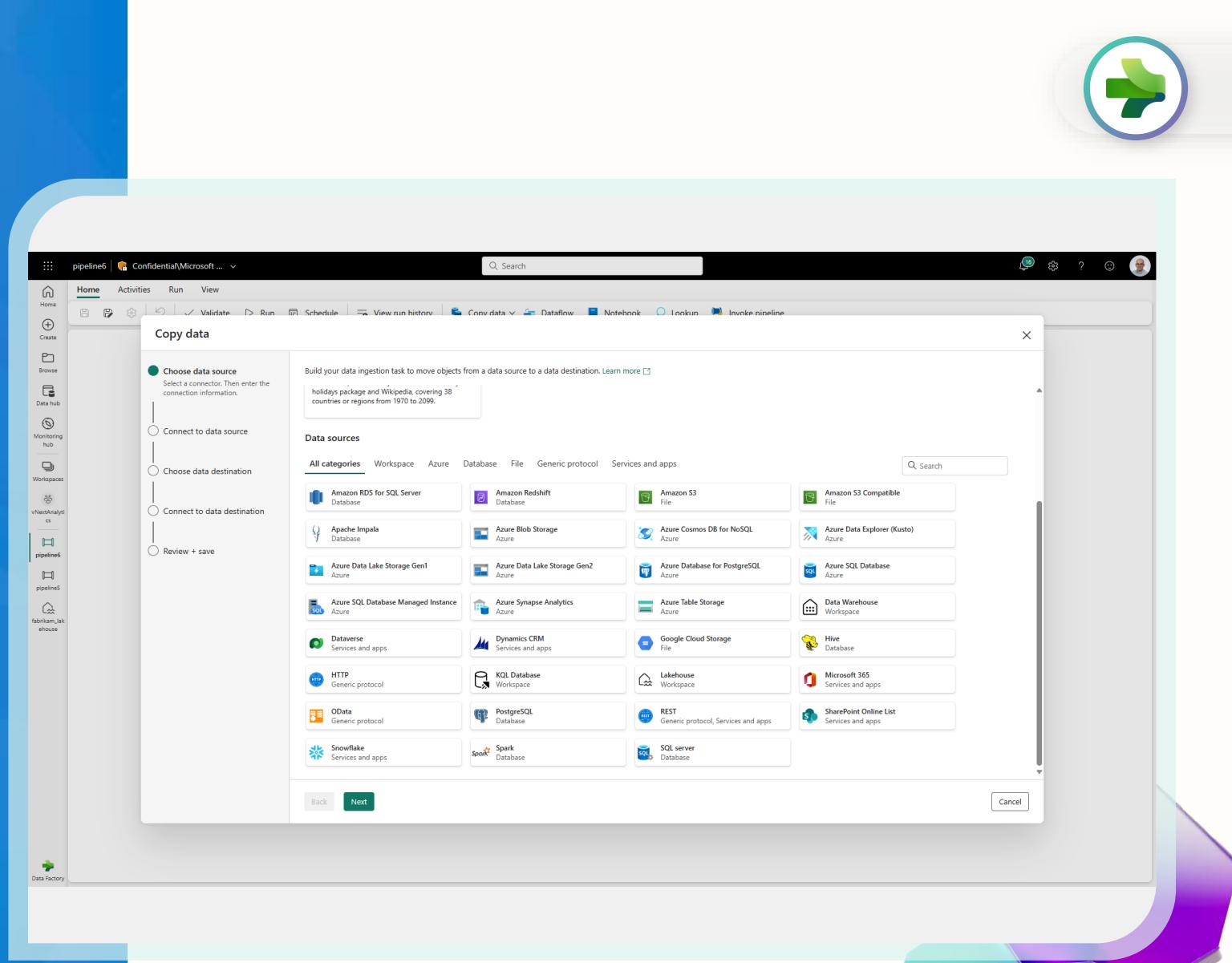
Autonomous ETL can unlock operational efficiencies and help orchestrate, monitor, and manage pipeline performance.

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.



The screenshot shows the Microsoft Fabric Data Factory interface with a 'Copy data' wizard open. The left sidebar lists various pipeline components: Home, Create, Browse, Data hub, Monitoring hub, Workspaces, Visual Analytics, and two pipeline items named 'pipeline6' and 'fabrikam_lakehouse'. The main area has a header 'Copy data' with sub-tasks: 'Choose data source', 'Choose data destination', and 'Review + save'. A note says: 'Build your data ingestion task to move objects from a data source to a data destination. Learn more' with a link to 'holidays package and Wikipedia, covering 38 countries or regions from 1970 to 2099'. Below this is a 'Data sources' section with tabs for 'All categories', 'Workspace', 'Azure', 'Database', 'File', 'Generic protocol', and 'Services and apps'. The 'All categories' tab is selected, showing a grid of 24 data source icons. At the bottom of the wizard are 'Back', 'Next', and 'Cancel' buttons.

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

The screenshot shows the Microsoft Fabric Data Factory interface. A modal window titled "Copy data" is open, specifically the "Choose data source" step. The window includes a descriptive text block about moving objects from a data source to a destination, a search bar, and a grid of data source icons categorized by type. The categories at the top of the grid are "All categories", "Workspace", "Azure", "Database", "File", "Generic protocol", and "Services and apps". The "All categories" tab is selected. The grid contains numerous icons representing different data sources, such as Amazon RDS for SQL Server, Amazon Redshift, Amazon S3, Azure Blob Storage, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2, Azure Synapse Analytics, Azure Table Storage, DataVerse, Dynamics CRM, Google Cloud Storage, HTTP, OData, PostgreSQL, REST, Snowflake, Spark, and SQL server. At the bottom of the modal are "Back", "Next", and "Cancel" buttons.

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)

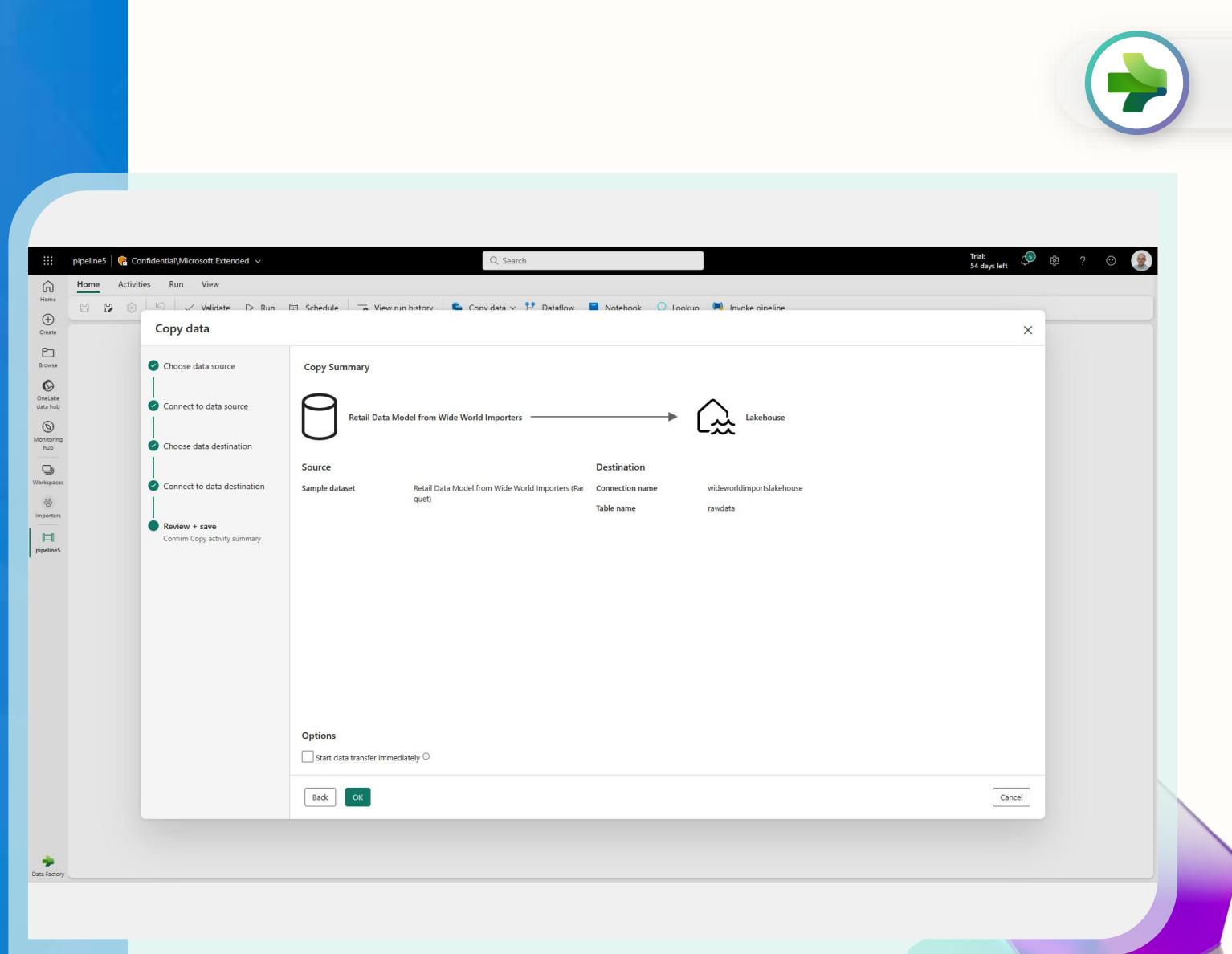
The screenshot shows the Microsoft Fabric Data Pipelines interface. On the left, a sidebar lists various data assets and pipelines. The main area is titled "Copy data" and is the first step in a pipeline named "pipeline6". It prompts the user to "Choose data source" and lists several options: "COVID-19 Data Lake", "NYC Taxi - Green", "Diabetes", "Public Holidays", and "Retail Data Model from Wide World Importers". Below these, a "Data sources" section lists various cloud services. At the bottom of the dialog are "Back", "Next", and "Cancel" buttons.



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

The screenshot shows the Microsoft Fabric Data Factory interface. On the left, a sidebar lists various workspace options like Home, Browse, OneLake data hub, Monitoring hub, Workspaces, Importers, and the current pipeline named 'pipeline4'. The main area is titled 'Templates' and displays a grid of 12 data transfer and management templates, each with a preview icon, title, and brief description. A modal window titled 'Select template' is open in the center, showing a preview of the 'Bulk Copy from Database' template, which includes fields for 'Source database' (selected as 'AdventureworksDW'), 'Destination database' (selected as 'AdventureworksLT'), and 'External control table' (selected as 'AdventureworksLT.dbo.dimCustomer'). Below the preview are 'Next' and 'Cancel' buttons. The overall interface has a light blue and white color scheme with a large green checkmark icon in the bottom right corner.

Template	Description
Bulk Copy from Database	Use this template to copy data in bulk from a database using an external control table to store the partition list of your source tables...
Bulk Copy from Files to Database	Use this template to copy data in bulk from Azure Data Lake Storage Gen2 to Azure SQL Database...
Copy data from ADLS Gen2 to Lakehouse file	Use this template to copy data from ADLS Gen2 to a specified file location in your Lakehouse...
Copy data from ADLS Gen2 to Lakehouse Table	Use this template to copy data from ADLS Gen2 to a specified table in your Lakehouse...
Copy data from Azure SQL DB to Lakehouse Table	Use this template to copy data from your Azure SQL database to a specified table in your Lakehouse...
Copy multiple files containers between File Stores	Use this template to leverage copy activities to copy containers or folders between file based stores, where each copy...
Copy new files only by LastModifiedDate	Use this template to copy new or changed files only by using LastModifiedDate...
Delete files older than 30 days	Use this template to delete files that have been modified more than 30 days ago from storage stores...
Delta copy from Database	Use this template to copy new or updated rows only from a database using a high-watermark stored in an external control table...
Move files	Use this template to move files from one folder to another folder. The pipeline enumerates the files...

Shortcuts



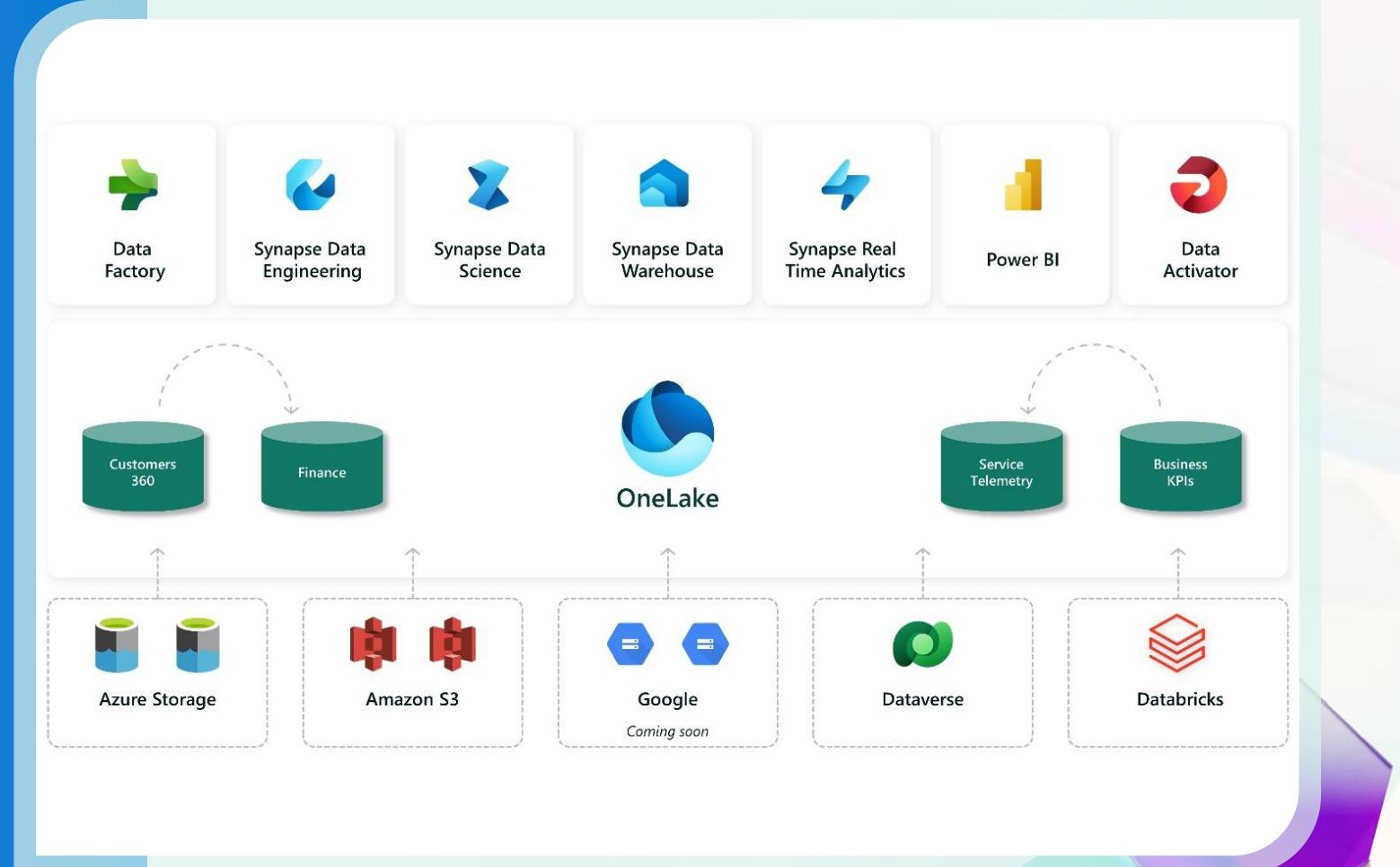
Shortcuts unify data without copying or moving existing data.

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



Data Warehouse

On SQL Endpoint

Lakehouse SQL Endpoint

Use data warehouse capabilities on top of your lakehouse with no data movement



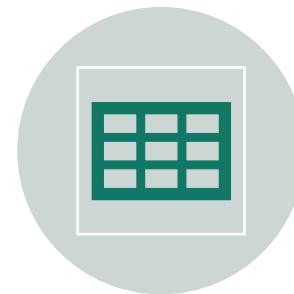
Run and store SQL queries
using the SQL visual editor
or authoring T-SQL



Define functions, views and
custom SQL security



Connect through external
tools like Azure Data Studio
and SSMS



Cross-table and DB querying
(across warehouse artifacts
and lakehouse SQL
endpoints)

Power BI on top of Lakehouse

Build semantic datasets and reports directly on top of your lakehouse



Built-in semantic modelling capabilities enabling users to add measures and relationships inside the lakehouse experience



Blazing fast query performance with Vertipaq compression and direct lake connectivity



Generate reports in seconds directly from the lakehouse

Data Ingestion
Lakehouse, Shortcuts and
Data Factory

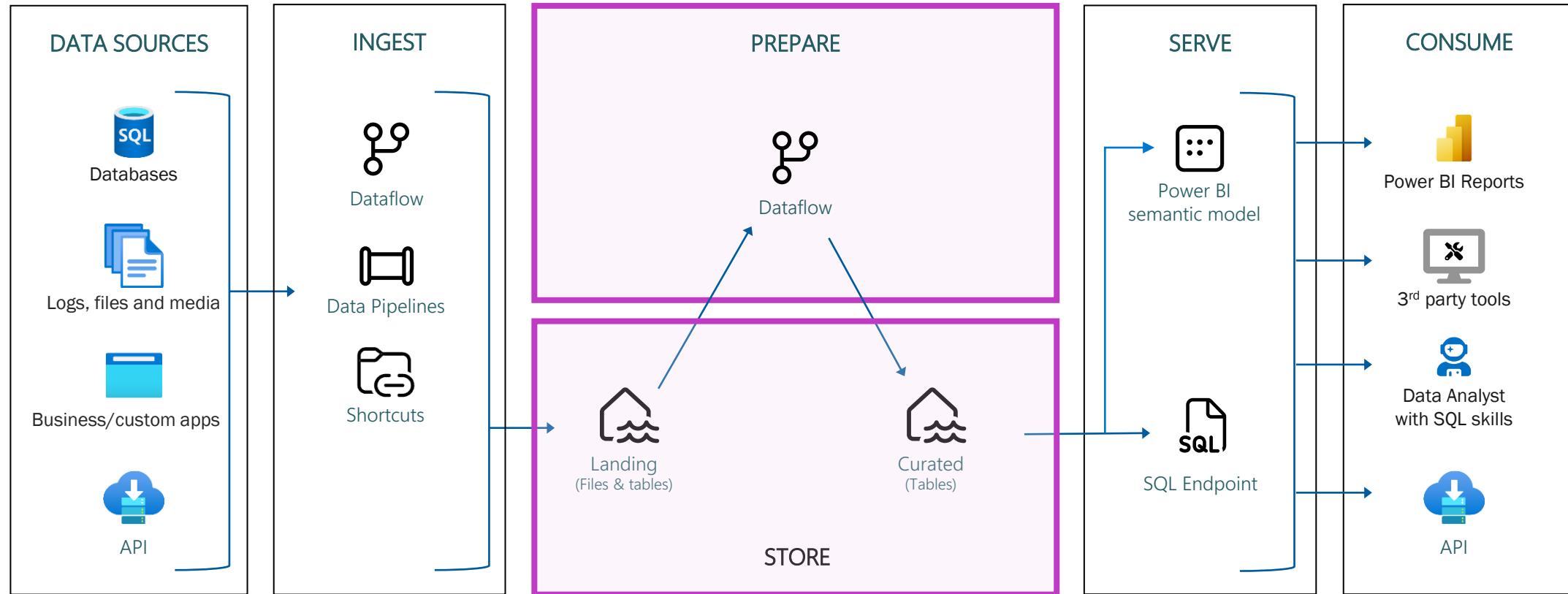


DEMO

Data preparation

Dataflows Gen2

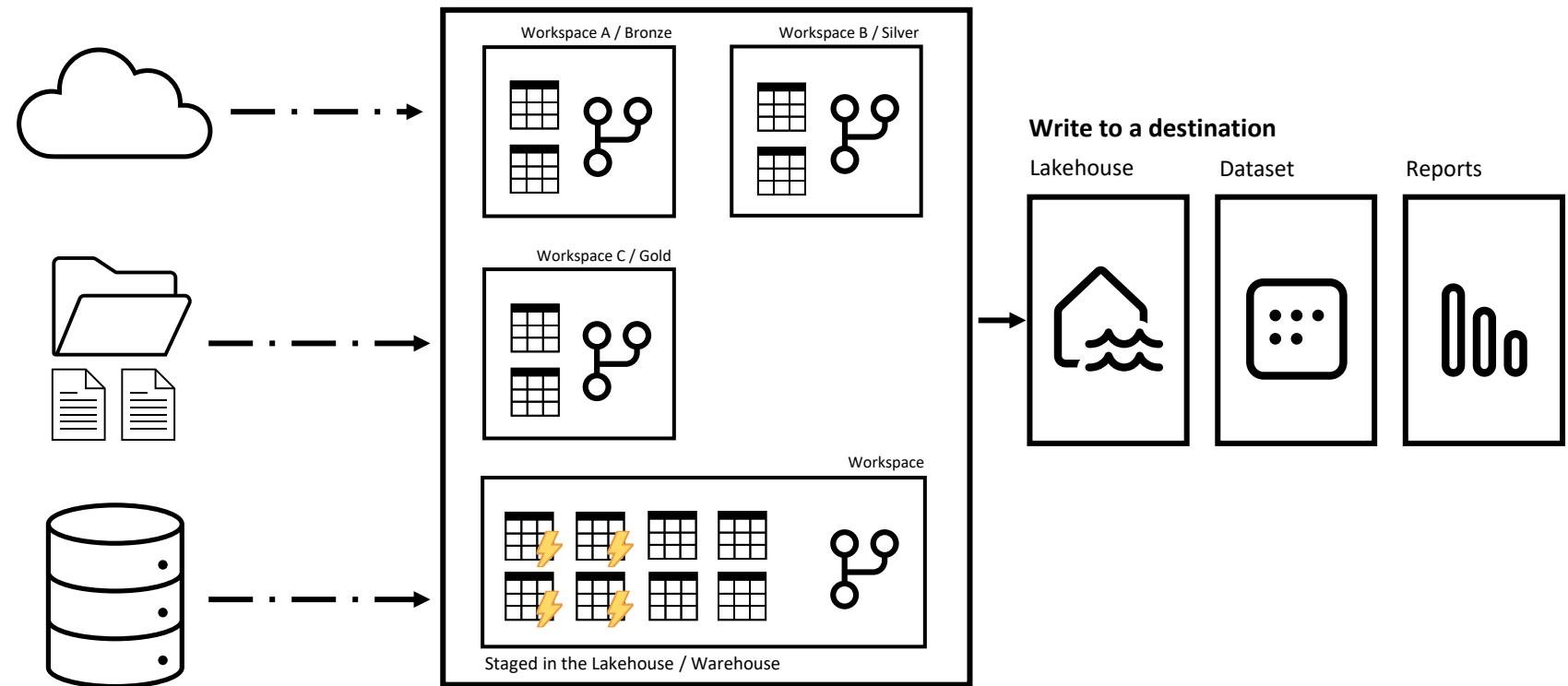
Data preparation in Fabric



Dataflows Gen2

Dataflows are a self-service, cloud-based, data preparation technology

Dataflows are available in Power BI, Power Apps and Dynamics 365 Customer Insights



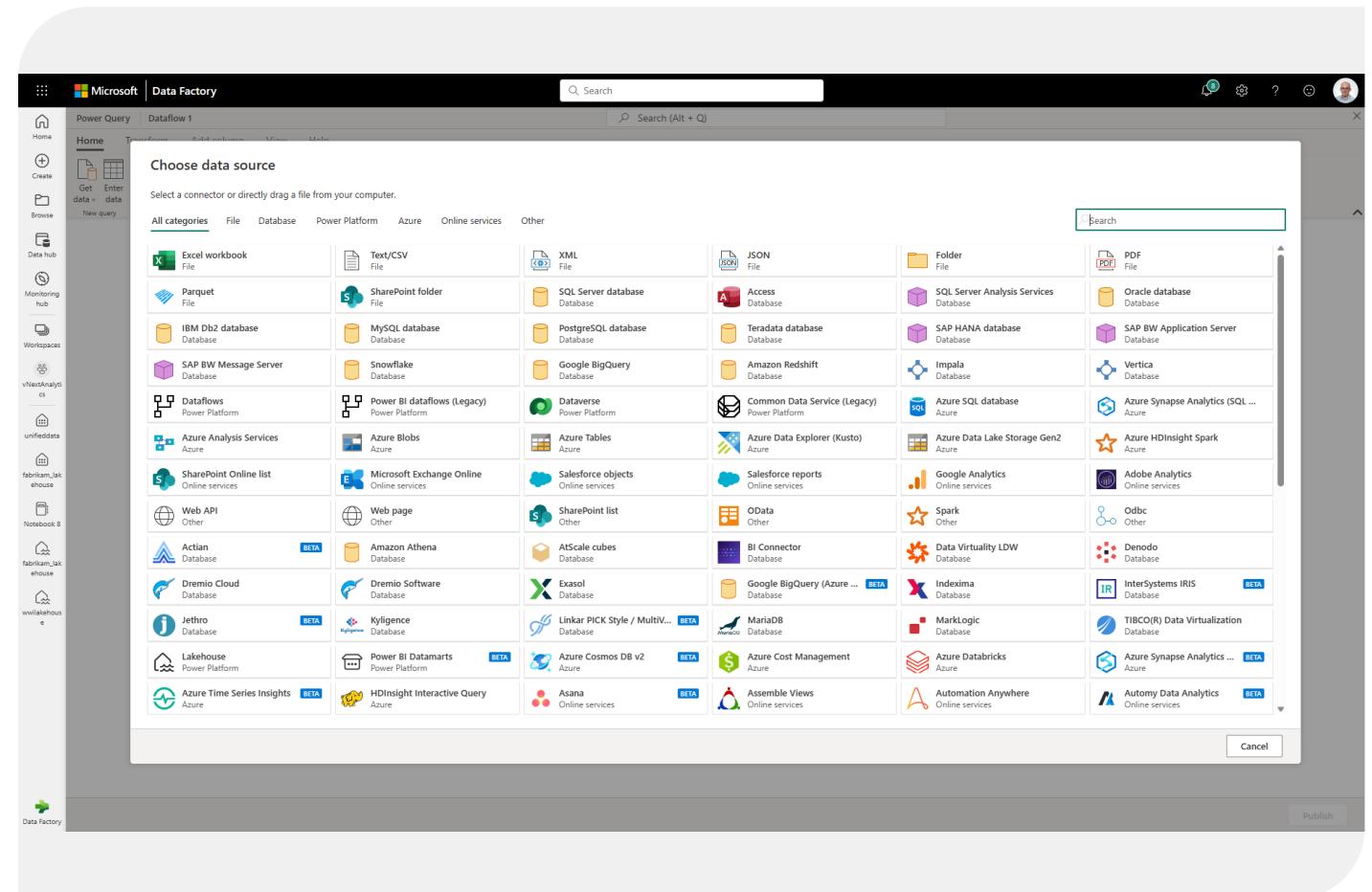
Dataflows Gen2

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.)



Dataflows Gen2 vs Gen1

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		✓

Dataflows Output to Lakehouse

Simply write into a Lakehouse from a Dataflow.

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

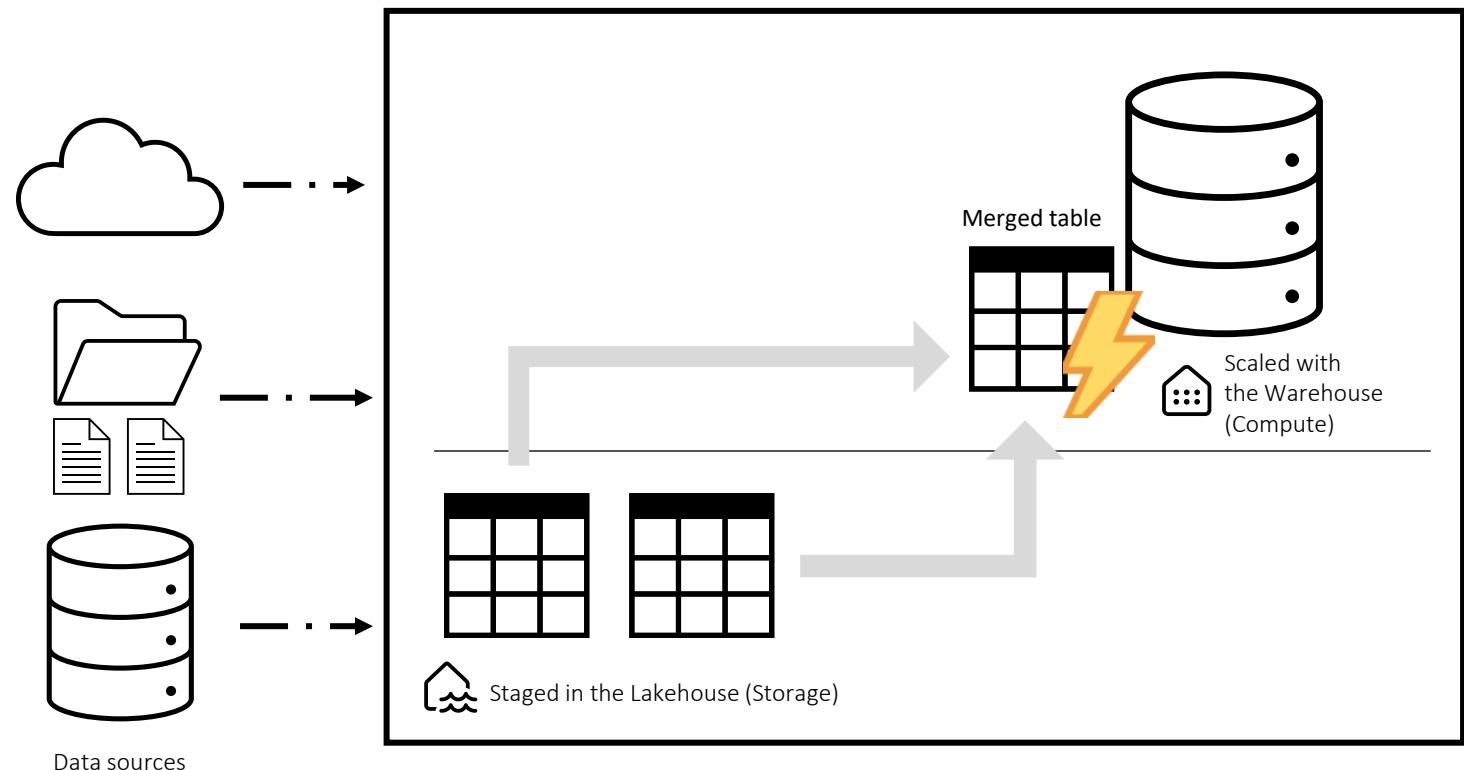
This requires the Lakehouse connector to be installed as a custom connector into your data gateway when loading data from on-premise.

The screenshot shows the Microsoft Synapse Data Engineering Importers interface. On the left, the 'Importers' sidebar is visible with options like Home, Create, Browse, OneLake data hub, Monitoring hub, Workspaces, and Importers. The main area shows a 'Power Query | Dataflow 2' window. In the center, there's a 'Queries [1]' pane with a table named 'Table.TransformCol1'. The table has columns: 'CustomerKey' (Int64), 'WWICustomerID' (Int64), 'Type' (Text), 'Customer' (Text), 'BillToCustomer' (Text), 'Category' (Text), 'BuyingGroup' (Text), 'PrimaryContact' (Text), 'PostalCode' (Text), 'ValidFrom' (Text), 'ValidTo' (Text), and 'LineageKey' (Text). The data consists of 99 rows, mostly from 'Wingtip Toys' head offices. To the right of the table, there's a 'Transform' ribbon bar with various data manipulation tools. On the far right, the 'Query settings' pane is open, showing 'Properties' (Name: dimension_customer ?, Entity type: Custom) and 'Applied steps' (Source, Promoted h., Changed co...). At the bottom, the status bar says 'Completed (0.93 s) Columns: 11 Rows: 99+' and 'Data Engineering'.

Computed Tables

Power Query Online enables Premium subscribers to use their capacity to optimize the use of dataflows with Fabric compute.

1. Connect to your data and copy it into the Lakehouse using *“Enable staging” (**on by default!**)
2. Create a reference query in a new query.
3. Apply transformation steps to the computed table for complex ETL operations such as join, distinct, filter and group by – leveraging the Warehouse for compute



*Previously titled “Enable load”

Computed Tables

Power Query Online enables Premium subscribers to use their capacity to optimize the use of dataflows with Fabric compute.

1. Connect to your data and copy it into the Lakehouse using *“Enable staging” (**on by default!**)
2. Create a reference query in a new query.
3. Apply transformation steps to the computed table for complex ETL operations such as join, distinct, filter and group by – leveraging the Warehouse for compute

*Previously titled “Enable load”

The screenshot shows the Power Query Online interface with a dataflow diagram. The dataflow consists of several stages: Data staging (DimProductCategory_raw, DimProductSubcategory_raw, DimCustomer_raw, DimGeography_raw), Data load (DimDate, DimEmployee, DimStore), and Data transformation (DimProduct, DimCustomer). A context menu is open over a step in the DimProduct transformation stage, with the 'Enable staging' option highlighted. The main pane displays a table with columns: ProductKey, ProductLabel, ProductName, and ProductDescription. The table has 99+ rows and was completed in 1.17 seconds. A status bar at the bottom indicates column profiling based on top 1,000 rows.

ProductKey	ProductLabel	ProductName	ProductDescription
101001	Contoso 512MB MP3 Player E51 Silver	512MB USB driver plays MP3 and WMA	
101002	Contoso 512MB MP3 Player E51 Blue	512MB USB driver plays MP3 and WMA	
101003	Contoso 1G MP3 Player E100 White	1GB flash memory and USB driver plays MP3 and WMA	
101004	Contoso 2G MP3 Player E200 Silver	2GB flash memory, LCD display, plays MP3 and WMA	
101005	Contoso 2G MP3 Player E200 Red	2GB flash memory, LCD display, plays MP3 and WMA	
101006	Contoso 2G MP3 Player E200 Black	2GB flash memory, LCD display, plays MP3 and WMA	

Closing Thoughts on Data Preparation

1. "As far upstream as possible, as far downstream as necessary."
2. Spending hours, to save seconds on refresh times is a cost-benefit analysis of your time.
3. Always default to the UI first and the code last.

Data preparation
Dataflows Gen2



DEMO

Notebooks And Data Wrangler



Data Wrangler

```
1 # Welcome to your new notebook
2 # Type here in the cell editor to add code!
3
4
5 # With Spark SQL, Please run the query onto the lakehouse which is from the same workspace as the current default lakehouse.
6
7 df = spark.sql("SELECT * FROM ingest1.Customer LIMIT 1000")
8 display(df)
```

[10] -Command executed in 2 sec 597 ms by Daan on 4:18:42 PM, 3/06/24

Output is hidden

```
1 # Code generated by Data Wrangler for pandas sample
2
3 def clean_data(pandas_df):
4     # Select columns: 'CustomerKey', 'Gender' and 10 other columns
5     pandas_df = pandas_df.loc[:, ['CustomerKey', 'Gender', 'GivenName', 'Surname']]
6     # Derive column 'Name' from columns: 'GivenName', 'Surname'
7     # Transform based on the following examples:
8     #   GivenName Surname   Output
9     # 1: "Pasquale" "Gallo" => "Pasquale Gallo"
10    pandas_df.insert(4, "Name", pandas_df["GivenName"].str.strip() + " " + pandas_df["Surname"])
11    # Drop columns: 'Surname', 'GivenName'
12    pandas_df = pandas_df.drop(columns=['Surname', 'GivenName'])
13    # Scale column 'Age' between 0 and 1
14    new_min, new_max = 0, 1
15    old_min, old_max = pandas_df['Age'].min(), pandas_df['Age'].max()
16    pandas_df['Age'] = (pandas_df['Age'] - old_min) / (old_max - old_min) * (new_max - new_min)
17    return pandas_df
18
19 # Loaded variable 'df' from kernel state
20 pandas_df = df.limit(5000).toPandas()
21
22 pandas_df_clean = clean_data(pandas_df.copy())
23 pandas_df_clean.head()
```

[11] -Command executed in 754 ms by Daan on 4:18:43 PM, 3/06/24

CustomerKey	Gender	Name	City	State	StateFull	Country	CountryFull	Birthday	0.87	
0	717000	Male	Pasquale Gallo	Valtina	BZ	Bolzano	IT	Italy	1943-11-08	0.87
1	726955	Male	Ludovico Bianchi	Senales	BZ	Bolzano	IT	Italy	1977-09-12	0.36
2	727649	Male	Giacinto Greece	Monte Isola	BS	Brescia	IT	Italy	1955-05-16	0.69
3	732772	Male	Giraldo Manfrin	Gaibola	BO	Bologna	IT	Italy	1990-01-10	0.18
4	754897	Male	Urielle Endrizzi	Racines	BZ	Bolzano	IT	Italy	1966-10-21	0.53

← Data Wrangler: df

+ Add code to notebook Copy code to clipboard Save as CSV Views

Your Spark DataFrame has been converted into a pandas sample for performance, but all the code generated by Data Wrangler will be converted to PySpark when you add it to your notebook.

# index	# CustomerKey	Missing: Distinct:	Gender	Title	GivenName	MiddleInitial	Summary
0	717000	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	711 Distinct values	A: G: M: Other: 14% 9% 8% 70%	Data shape 1,000 rows x 21 columns
1	726955	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Pasquale	D	Columns 21
2	727649	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Ludovico	L	Rows (1,000)
3	732772	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Giacinto	A	Rows with missing values 0 (0.0%)
4	754897	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Giraldo	D	Duplicate rows 0 (0.0%)
5	758863	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Urielle	A	Missing values (0)
6	761495	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Ulderico	F	No missing values
7	768089	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Aladino	P	
8	785344	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Prisco	L	
9	785381	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Salvo	A	
10	789061	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Amalio	C	
11	791808	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Eugenio	M	
12	715258	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Calogero	R	
13	718173	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Miranda	A	
14	722794	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Gaetana	E	
15	730651	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Italia	G	
16	737216	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Iva	T	
17	747984	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Serena	A	
18	761297	1000 (100%)	Female: Male:	64% Dr. 36% Mrs. 29% Ms. 4% Other:	Iole	D	
					Susanna	E	

Cleaning steps

1 Load data from variable

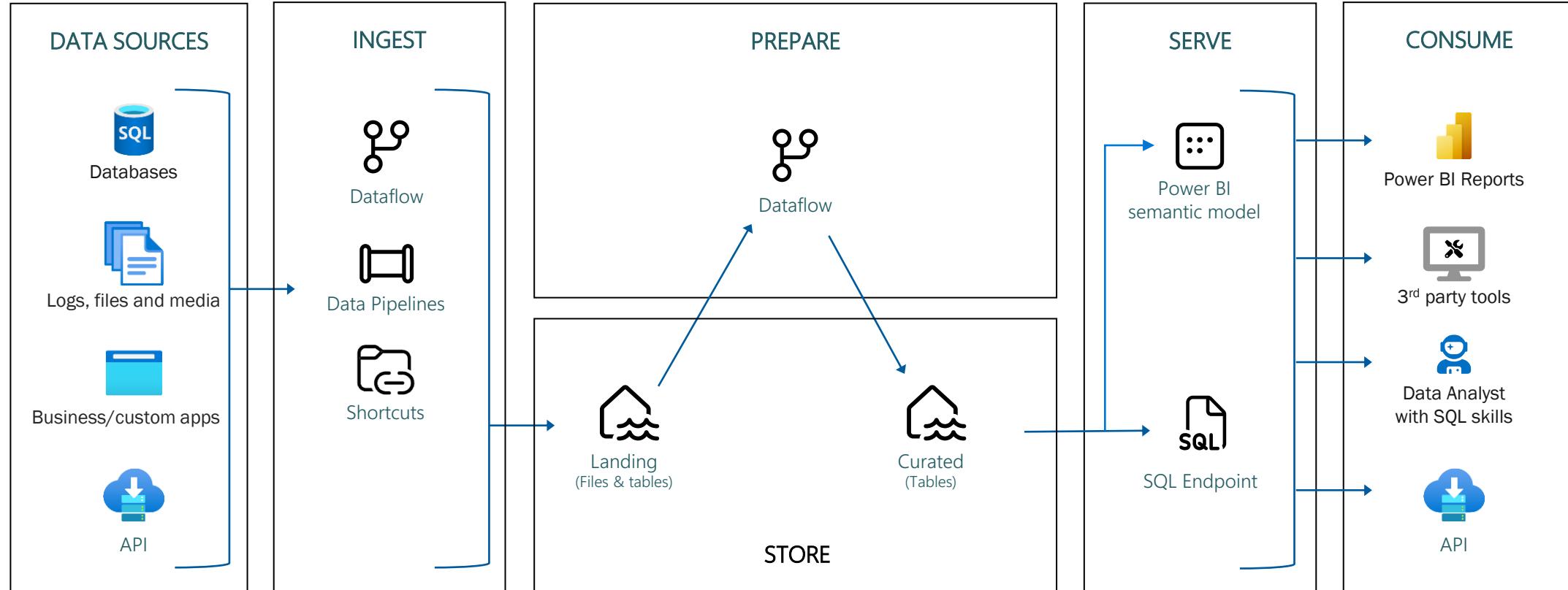
2 New operation

Choose an operation to get started

Preview code for all steps

Apply Discard

Ingest Data in Fabric



Data modelling

Some considerations

the

FLATFILE

the

vs. STAR

which is better?...

Normalized (Flat File)

1. Advantages

- Single table to manage
- Limited maintenance

2. Disadvantages

- Slow refresh
- No organization
- Complicated DAX (possibly)
- Time Intelligence can be challenging

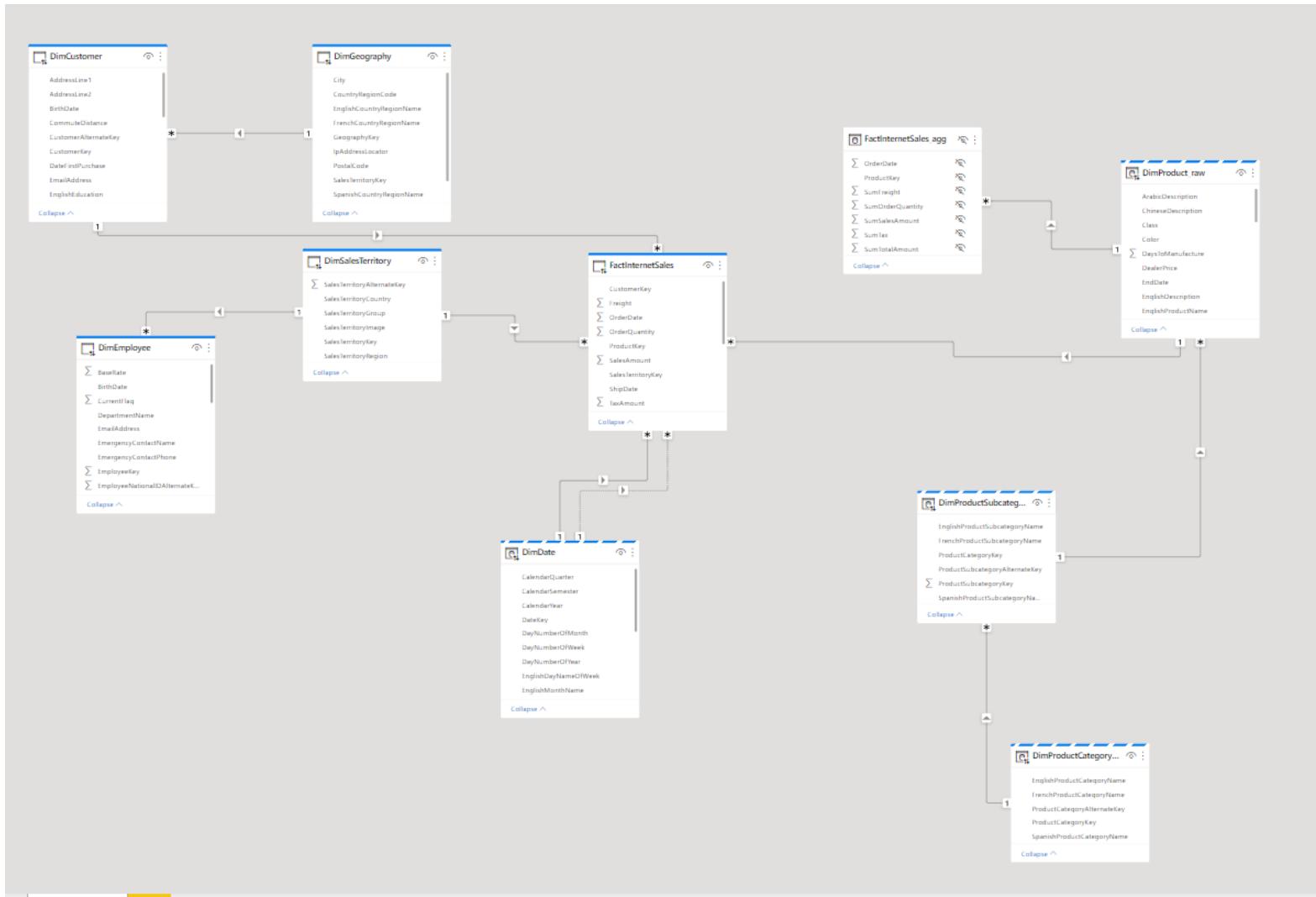
Relationships

1. Advantages:

- Filtering
- Aggregating
- Simpler DAX

2. Disadvantages:

- Only one column
- Can be costly
- Be careful with
 - Many-to-many
 - Bi-directional



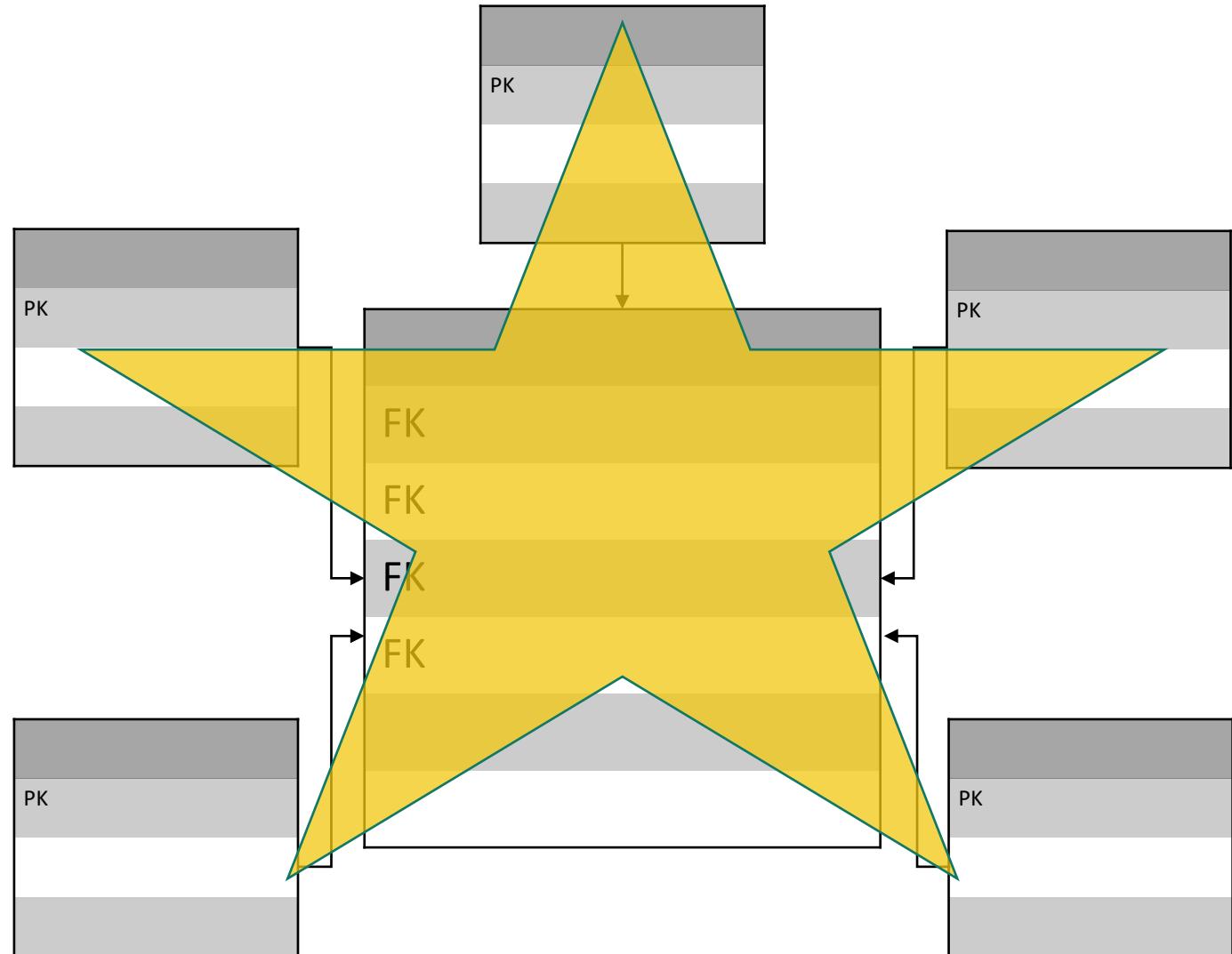
Denormalized (Star Schema)

1. Advantages

- Simple DAX (possibly)
- Organized
- Faster Refresh
- Consistency

2. Disadvantages

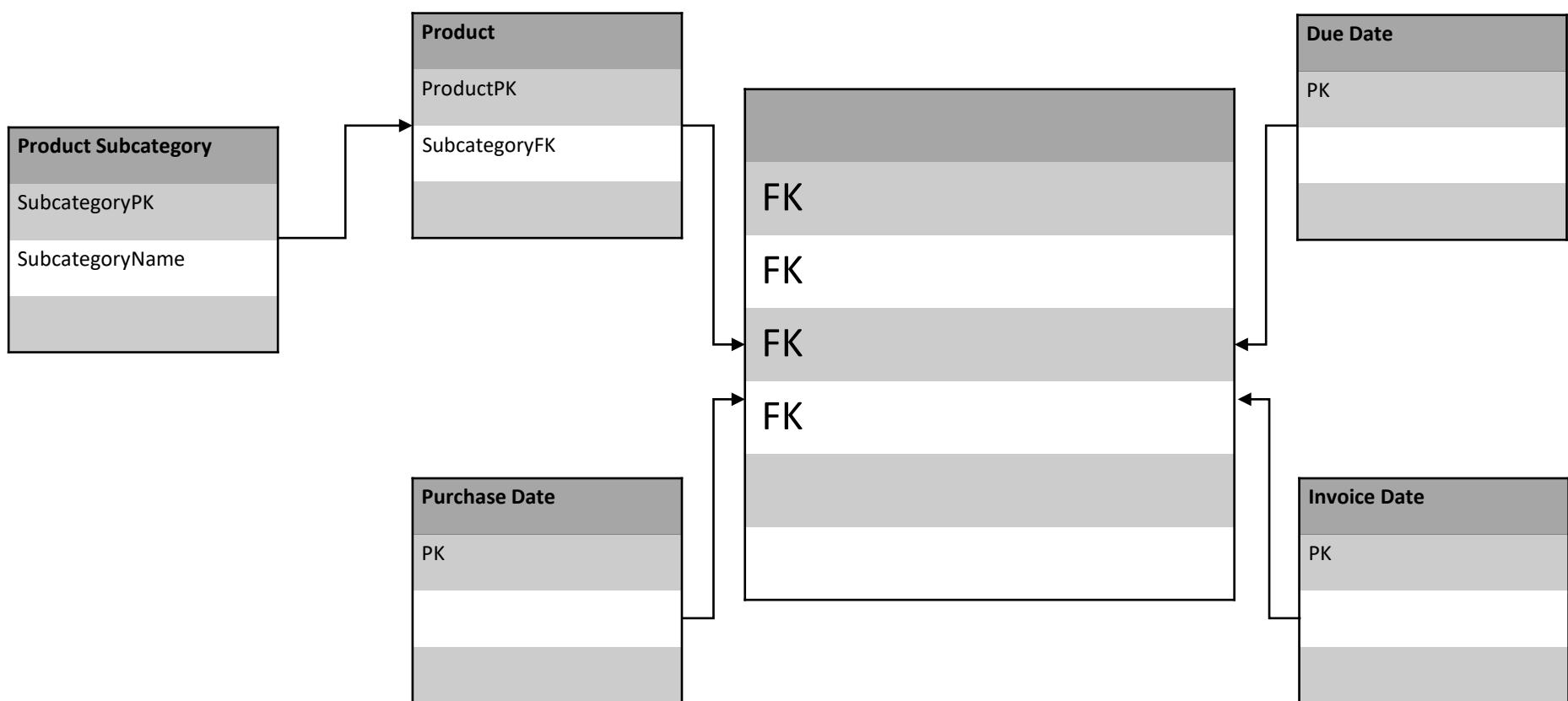
- Subject Matter Expert
- Maintenance
- Cost



Snowflake Schema

Design decisions

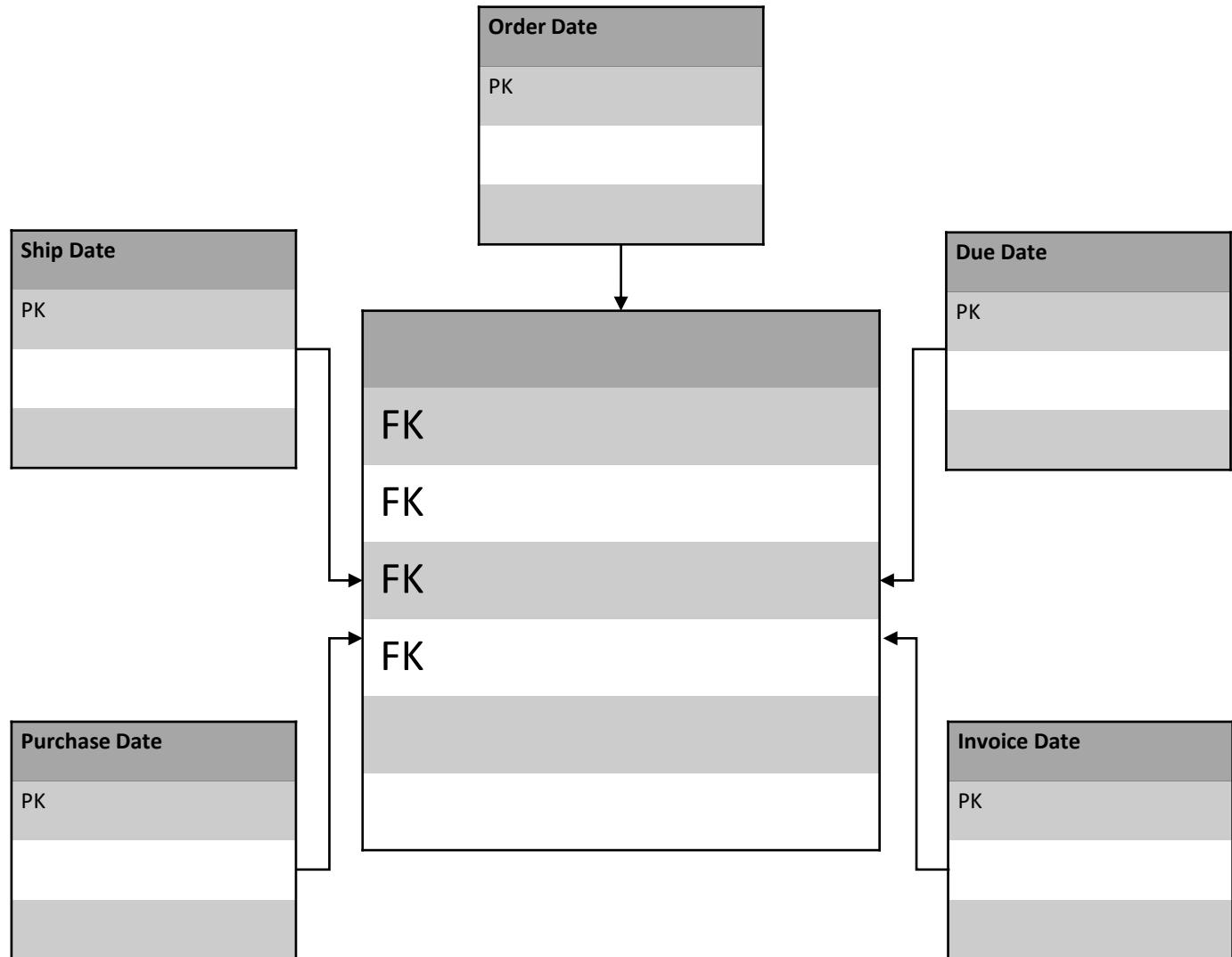
- More Tables
- Possibly more DAX
- Relationship size
- Possible query performance



Role-Playing Dimensions

Design decisions

- Multiple Tables
- Single-Table
- UseRelationship()



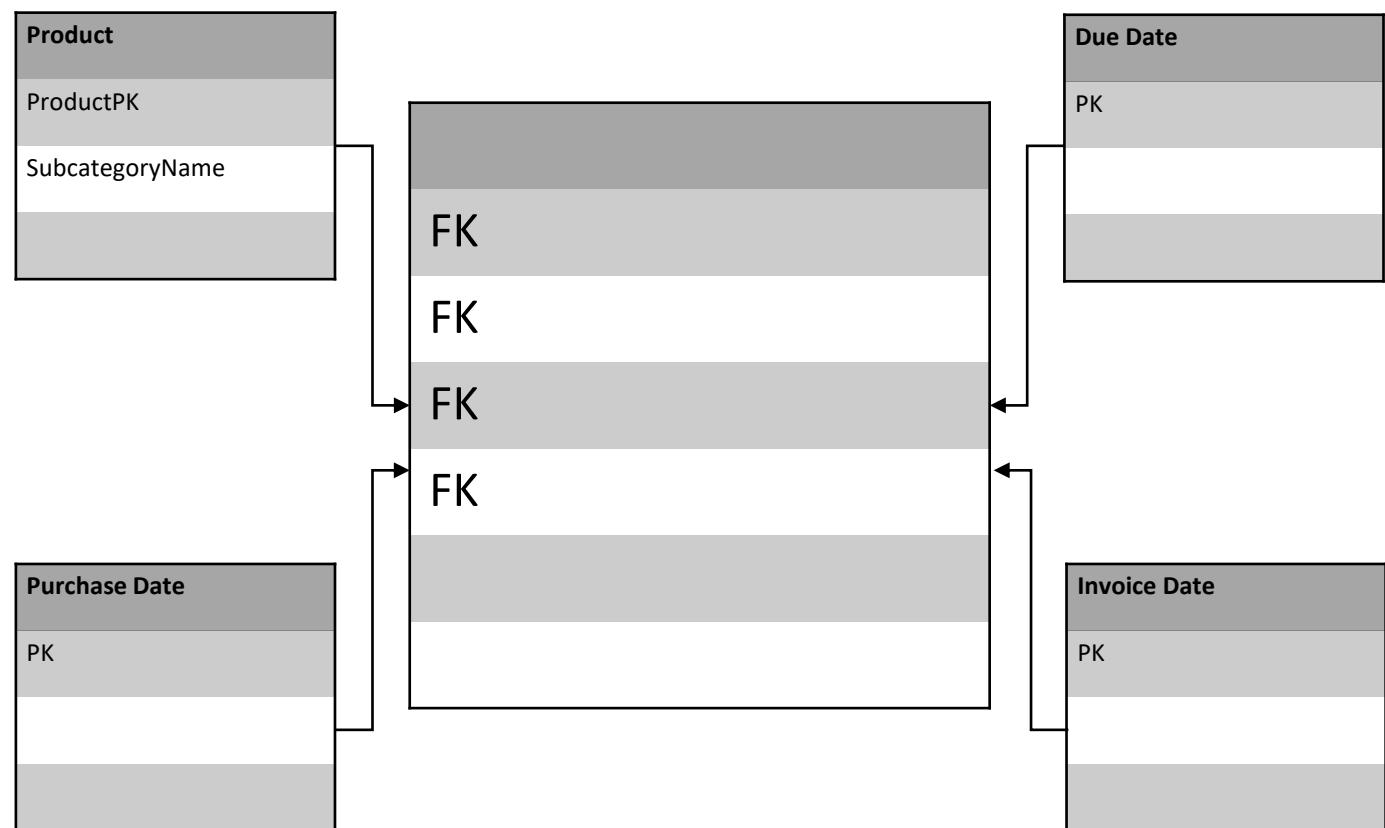
One Dataset...

1. Advantages

- Endorsement
 - Promotion
 - Certification
- Consistency
- Extensible

2. Disadvantages

- Maintenance
- Discipline
- Complication
- Scalability



Closing Thoughts on Data Modeling

1. "As far upstream as possible, as far downstream as necessary."
2. Less is Always More. Only Bring in What You Need.
3. Keep it Simple. Keep it Clean.
4. There is no such thing as "**Come Back To It**".
5. Test. Test. Test.

Lunch break



Semantic Models overview



Power BI: The bridge between data and decisions



Semantic model



Data
Warehouse



Lakehouse



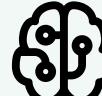
Power BI
service



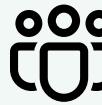
Interactive reports,
real-time dashboards,
and paginated reports



Custom visualizations



AI-infused insights and
intuitive data exploration
using natural language query



Sharing and collaboration

Consumption



Web



On-premises



Mobile



Embedded



Teams



PowerPoint



SharePoint



Excel



Microsoft
Power Platform

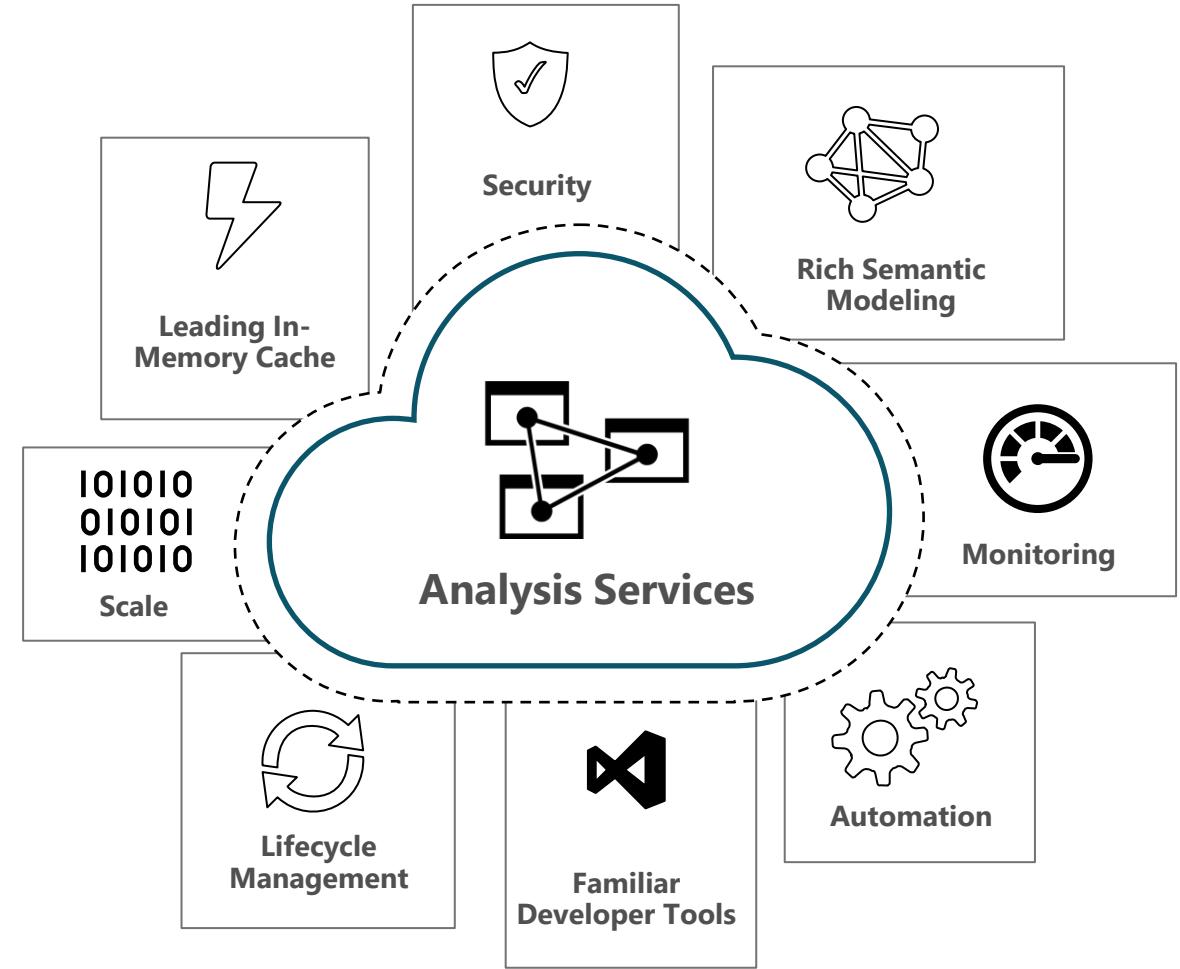


..."

Analysis Services Engine

Most Used, Most Battle-Tested Analytics Engine Technology

- Enterprise grade analytics **engine** as a service
- Leading query performance over massive data volumes
- Rich semantic modeling
- Interactive Exploration over even **TRILLIONS** of rows of data, by intelligent proactive caching of aggregates

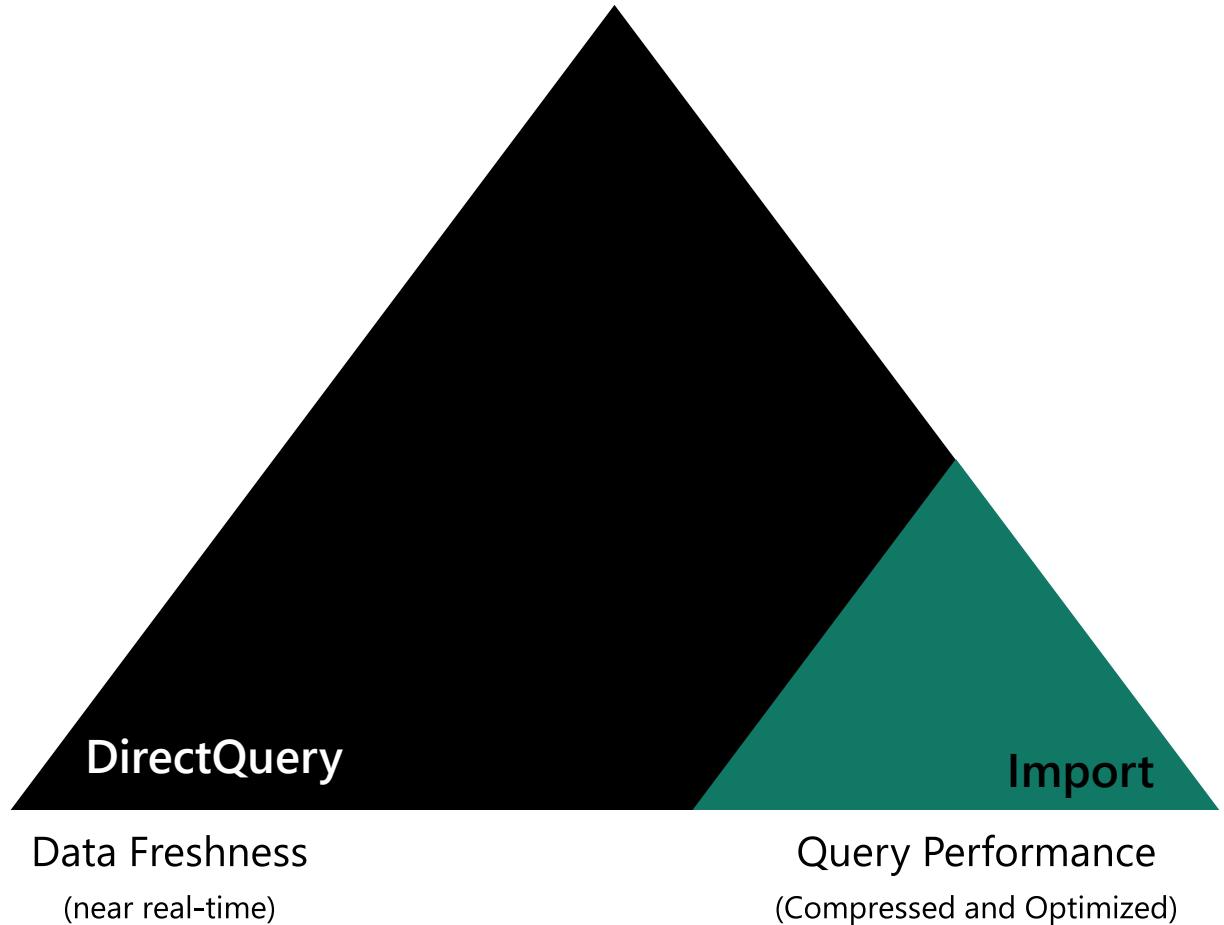


Storage Modes

Import: Caches data into memory to deliver extremely fast performance using the **analysis services** database engine. The default mode when creating a new Power BI Desktop solution along with providing Data Modelers the most design flexibility.

DirectQuery: Does not import the data into memory, consists only of the metadata defining the structure. When the model is queried, native queries are used to retrieve data from the underlying data source.

Changing the **Storage mode** of a table to **Import** is an irreversible operation. Once set, this property can't later be changed using Power BI Desktop.



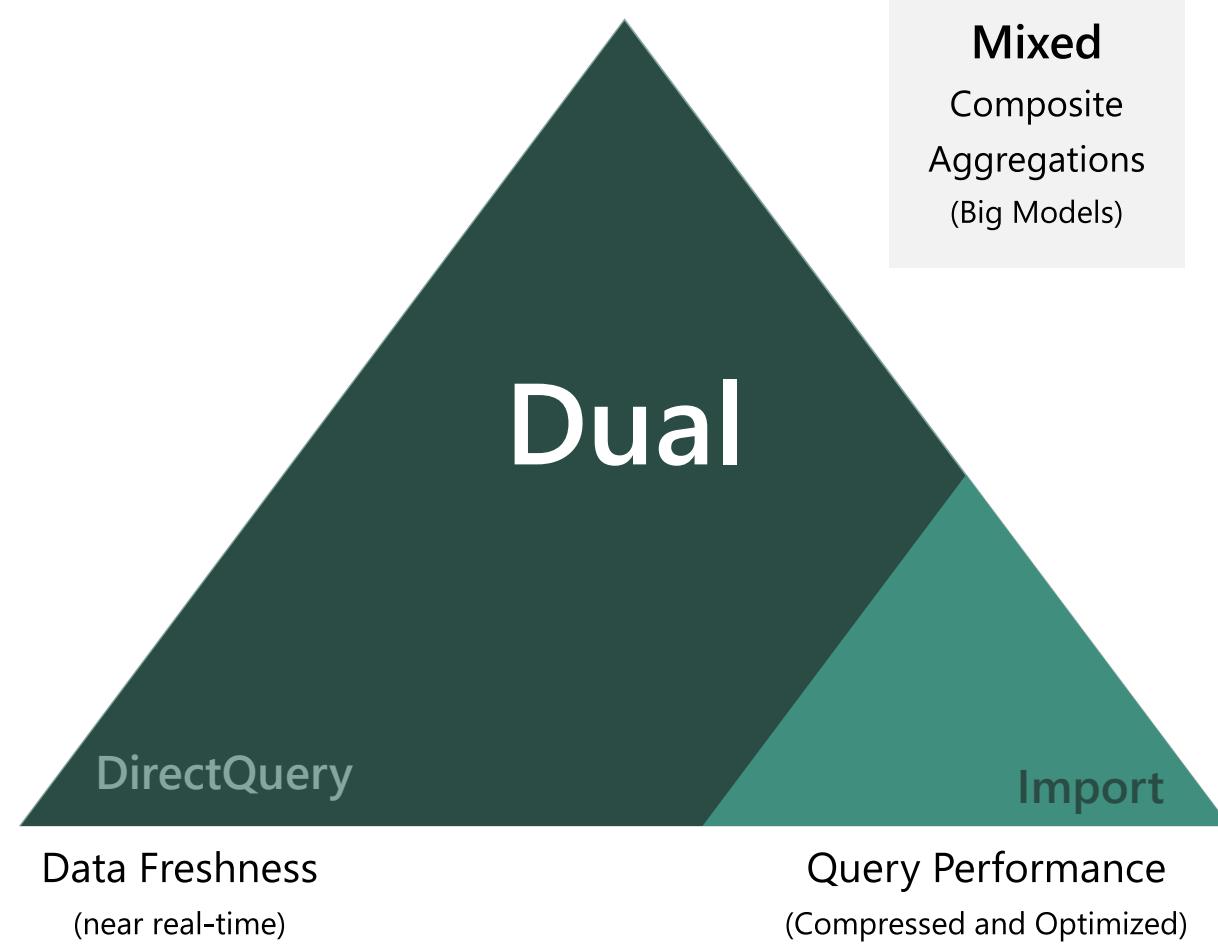
Storage Modes

Import: Caches data into memory to deliver extremely fast performance using the **analysis services** database engine. The default mode when creating a new Power BI Desktop solution along with providing Data Modelers the most design flexibility.

DirectQuery: Does not import the data into memory, consists only of the metadata defining the structure. When the model is queried, native queries are used to retrieve data from the underlying data source.

Dual: Can act as either cached or not cached, depending on the context of the query that's submitted to the Power BI dataset. In some cases, you fulfill queries from cached data. In other cases, you fulfill queries by executing an on-demand query to the underlying data source.

Changing the **Storage mode** of a table to **Import** is an irreversible operation. Once set, this property can't later be changed to either **DirectQuery** or **Dual** using Power BI Desktop.



Storage Modes

Direct Lake: A groundbreaking new dataset capability for analyzing very large data volumes. 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 dataset. Direct Lake is a fast-path to load the data from the lake straight into the Power BI engine, ready for analysis and yielding performance similar to import mode.

DirectQuery (fallback): Automatically switches modes—either due to current limitations or based on factors such as available memory in the capacity.



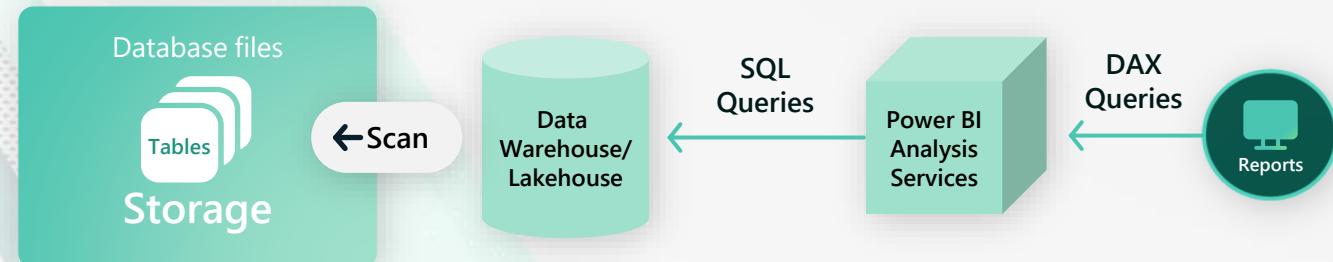


Power BI | Direct Lake Mode

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 dataset

DirectQuery Mode. Slow, but real time



Import Mode. Fast, but latent and duplicative



Direct Lake Mode. Fast and real time



Data Model Properties

With the **Properties** pane in the Data modeling view, dataset developers can enrich their Tabular Object Model (TOM) with additional rich metadata...

- Description
- Row label
- Key column
- Formatting
- Data category
- and more...

The screenshot shows two side-by-side 'Properties' panes for a dataset named 'Online Sales'. The left pane is for the dataset itself, and the right pane is for a specific table or query within it.

Dataset Properties (Left Pane):

- Name:** Online Sales
- Description:** Contains all online sales transactions including quantities and amounts.
- Synonyms:** online sale, sale
- Row label:** Select a row label
- Key column:** OnlineSalesKey
- Is hidden:** No (checkbox)
- Is featured table:** No (checkbox)
- Storage mode:** Import

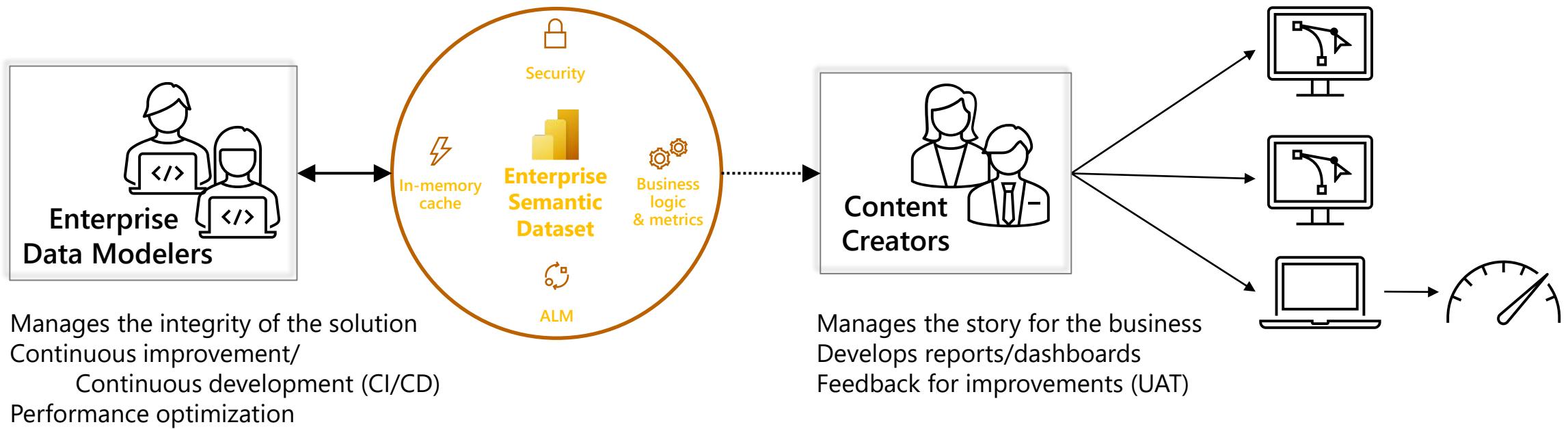
Table Properties (Right Pane):

- Description:** Enter a description (placeholder: Enter a description)
- Synonyms:** PoP % change
- Display folder:** Enter the display folder
- Is hidden:** No (checkbox)
- Formatting:**

 - Format:** General
 - Percentage format:** No (checkbox)
 - Thousands separator:** No (checkbox)
 - Decimal places:** Auto

Shared Dataset

Live Connection: You can establish a connection to a shared dataset in the Power BI service, and create many different reports from the same dataset. Then you and others can create multiple different reports in separate files and save them to different workspaces.



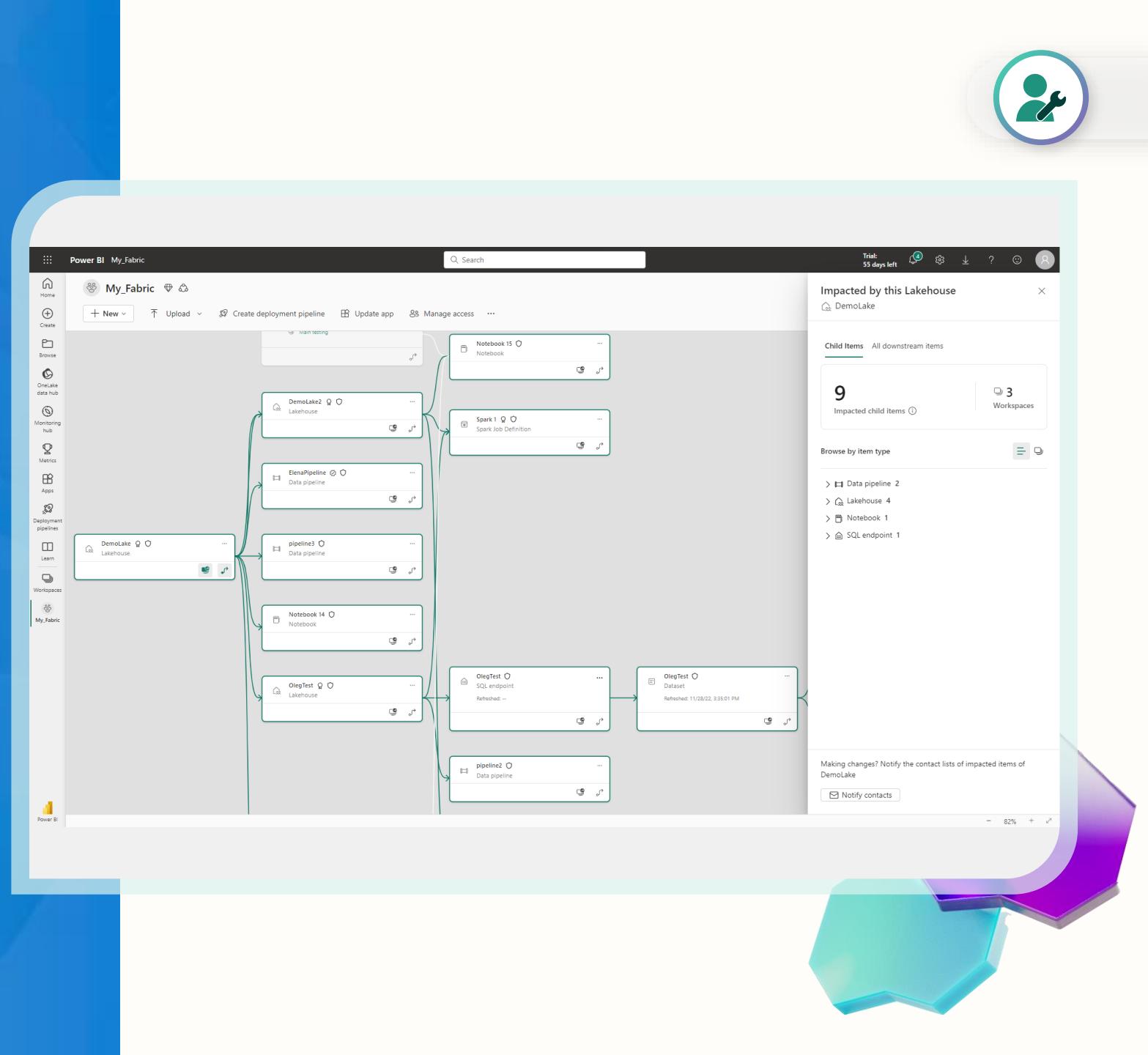
Track data lineage and impact analysis

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



Report Types

Executive

Description: Provides very quick access to most important metrics, leaving little room for misinterpretation.

Analytical

Description: Supports the strategic functions of senior management. Often a balance between simplicity and utility through key performance indicators and business measures displayed in a Dashboard or Scorecard report style. Often built to add depth to an executive report

Operational

Description: Supports an organization's day-to-day functions. Often a highly-formatted and consistent output displayed in a Tabular or List report style.

Ad hoc

Description: Supports the flexibility of a one-time report, created in real-time for a particular purpose or business necessity. Often summarized in a PivotTable report style.



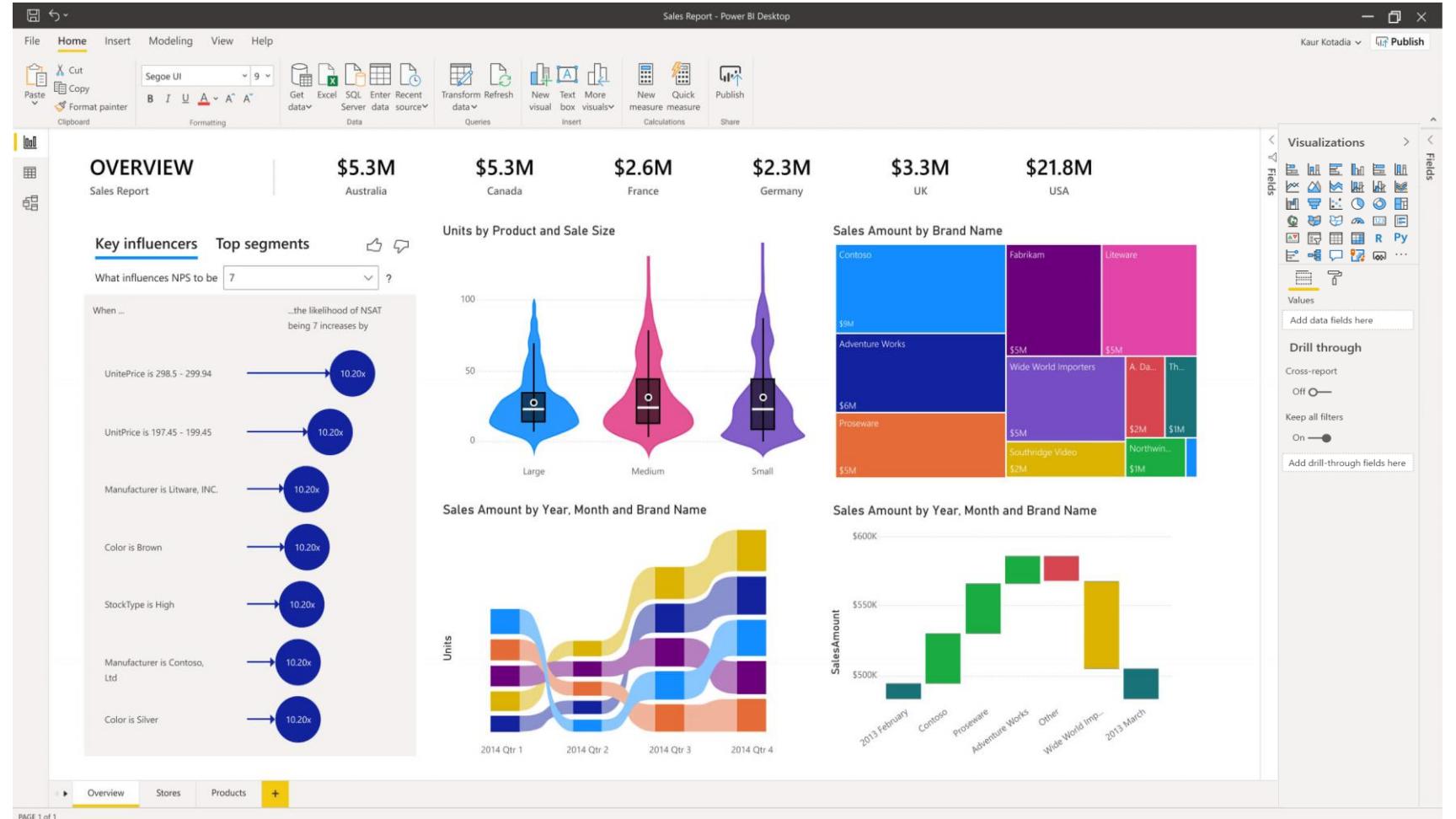
How does Power BI do it?





Power BI Desktop

- Familiar O365 like experience, the “**PowerPoint for data**”
- Create **interactive, immersive** reports that provide **actionable insights** and drive business results





Power BI Report Builder

- Create and distribute **highly formatted, multi-page, pixel-perfect** reports
- Use **parameters** to filter data for customized views

The screenshot shows the Power BI Report Builder application. The interface includes a ribbon with tabs like File, Home, Insert, and View. The Insert tab is selected, showing various data visualization options such as Table, Matrix, List, Chart, Gauge, Map, Data, Sparkline, Indicator, Text Box, Image, Line, Rectangle, Subreport, Header, and Footer. A 'Report Data' pane on the left lists items like Built-in Fields, Parameters, Images, Data Sources, and Datasets. A 'Parameters' pane shows a dropdown set to 'Product' with a value of 'NULL'. The main workspace displays a report titled 'SALES & RETURNS' with a table containing columns: Category, Product, Date, and Price. The table has four rows of sample data. Below the table, the 'Row Groups' and 'Column' sections are visible. To the right, a 'Properties' pane provides detailed settings for the report, including General, Localization, Other, Page, and References sections. A preview window shows the report's layout and navigation controls. The bottom of the screen features a zoom control at 100%.

Category	Product	Date	Price
[Category]	[Product]	[Date]	[Price]
			[&Execute]
Office 365	Excel	1/27/2019 12:00:00 AM	\$550.00
Office 365	Excel	1/8/2019 12:00:00 AM	\$550.00
Office 365	Excel	1/13/2019 12:00:00 AM	\$550.00
Office 365	Excel	1/20/2019 12:00:00 AM	\$550.00
Office 365	Excel	2/17/2019 12:00:00 AM	\$550.00
Office 365	Excel	2/24/2019 12:00:00 AM	\$550.00
Office 365	Excel	2/10/2019 12:00:00 AM	\$550.00
Office 365	Excel	2/3/2019 12:00:00 AM	\$550.00
Office 365	Excel	3/24/2019 12:00:00 AM	\$550.00
Office 365	Excel	3/3/2019 12:00:00 AM	\$550.00
Office 365	Excel	3/31/2019 12:00:00 AM	\$550.00
Office 365	Excel	3/10/2019 12:00:00 AM	\$550.00
Office 365	Excel	3/17/2019 12:00:00 AM	\$550.00
Office 365	Excel	4/28/2019 12:00:00 AM	\$550.00
Office 365	Excel	4/14/2019 12:00:00 AM	\$550.00
Office 365	Excel	4/21/2019 12:00:00 AM	\$550.00
Office 365	Excel	4/7/2019 12:00:00 AM	\$550.00
Office 365	Excel	5/12/2019 12:00:00 AM	\$550.00



Microsoft Excel

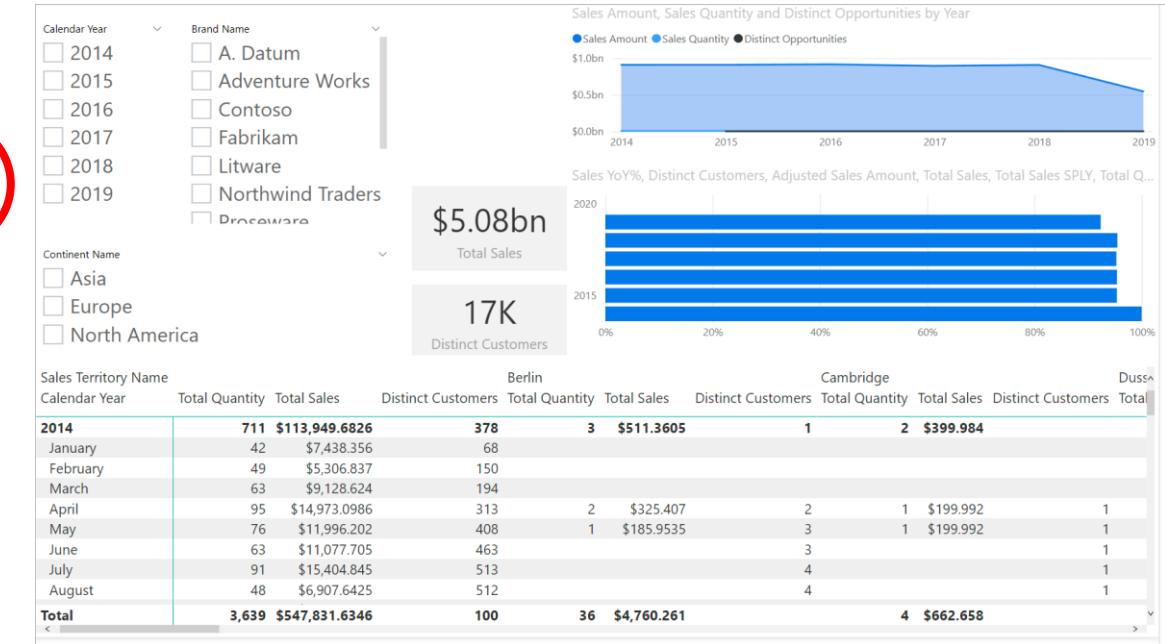
- Power BI users can analyze their data in Excel by using a **live, connected** experience
 - With a PivotTable you can quickly summarize your data, pivot the data to **view it from different angles** and analyze in specific detail

What does a *good
Report look like?



We know it when we see it...

Sales Territory Name	Calendar Year	Total Quantity	Total Sales	Berlin	Cambridge	Dussa
		Distinct Customers	Total Quantity	Total Sales	Distinct Customers	Total
2014		711	\$113,949.6826	378	3	\$511.3605
January		42	\$7,438.356	68		
February		49	\$5,306.837	150		
March		63	\$9,128.624	194		
April		95	\$14,973.0986	313	2	\$325.407
May		76	\$11,996.202	408	1	\$185.9535
June		63	\$11,077.705	463		
July		91	\$15,404.845	513		
August		48	\$6,907.6425	512		
September		22	\$4,662.0475	490		
October		88	\$13,230.3875	432		
November		31	\$6,996.25	373		
December		43	\$6,827.6875	378		
2015		653	\$105,848.6646	367	6	\$862.774
January		57	\$8,235.2625	370		
February		41	\$6,414.9715	358	2	\$325.407
March		61	\$11,425.9515	397		
April		68	\$10,689.5925	402		
May		47	\$6,997.4397	428		
June		36	\$6,497.7525	425		
July		61	\$8,872.836	395		
August		54	\$7,126.87	402		
September		41	\$8,657.625	355		
October		46	\$7,346.9817	328		
November		58	\$8,984.923	315	1	\$119.96
December		63	\$14,598.4587	367	3	\$417.407
Total		3,639	\$547,831.6346	100	36	\$4,760.261



We know it when we see it...

Contoso

SALES

Sales Overview

Revenue won \$11.43M	Qualified Pipeline \$19.90M	Revenue goal \$23M	Forecast 136%
--------------------------------	---------------------------------------	------------------------------	-------------------------

Revenue Open by Sales Stage

Sales Stage	Revenue
1-Qualify	\$7,912.02K
2-Develop	\$8,170.42K
3-Propose	\$7,264.68K
4-Close	\$4,465.27K

Revenue Won and Revenue In Pipeline by Product LOB

Product Category	Revenue Won	Qualified P...	Forecast %
Accessories	\$4,485.19K	\$7,035.43K	128%
Devices	\$3,672.98K	\$6,605.71K	128%
Warranties	\$3,271.66K	\$6,259.22K	136%

WHAT IF the qualified forecast was adjusted by 0 % ?

By Team + User By Product

Product Category Revenue Won Qualified P... Forecast %

Product Category	Revenue Won	Qualified P...	Forecast %
Accessories	\$4,485.19K	\$7,035.43K	128%
Devices	\$3,672.98K	\$6,605.71K	128%
Warranties	\$3,271.66K	\$6,259.22K	136%

Forecast by Territory

Territory	Revenue Won	In Pipeline	Forecast %
US-SOUTH	\$4,520,554.00	\$7,269.60K	131%
US-WEST	\$3,041,107.00	\$6,061.89K	130%
US-MIDWEST	\$2,686,629.00	\$4,367.79K	141%
US-NORTHEAST	\$1,181,536.00	\$2,201.09K	113%
Total	\$11,429,826.00	\$19,900.36K	136%

Forecast by Location

Product Categories

- Accessories: Black cover 7", E-reader pen, Black cover 6"
- Devices: E-reader Plati..., E-reader Dia..., E-reader Plati..., E-reader Dia..., E-reader Stan...
- Warranties: 1 Year Warranty

We know it when we see it...



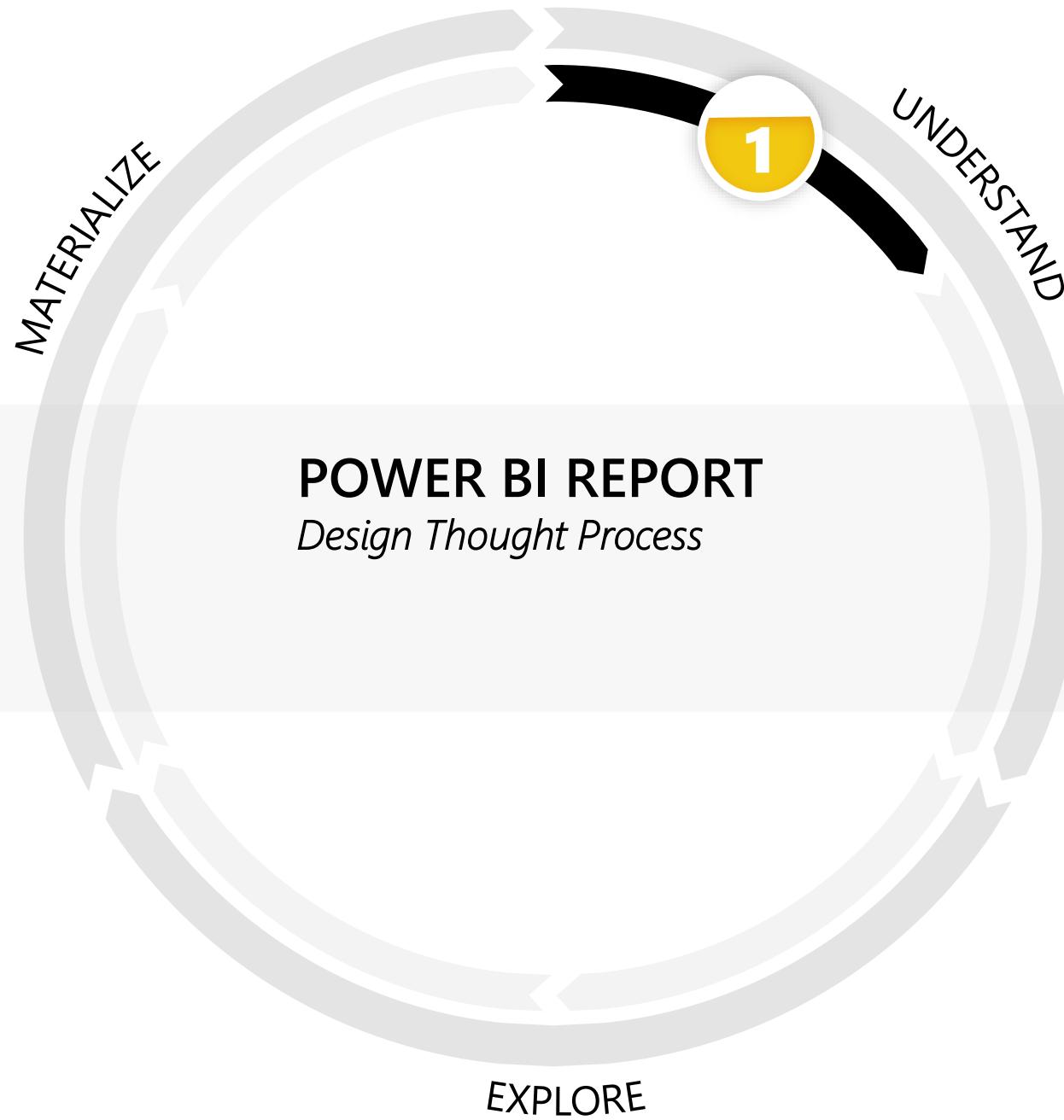
Report Design

Process

Design Tools

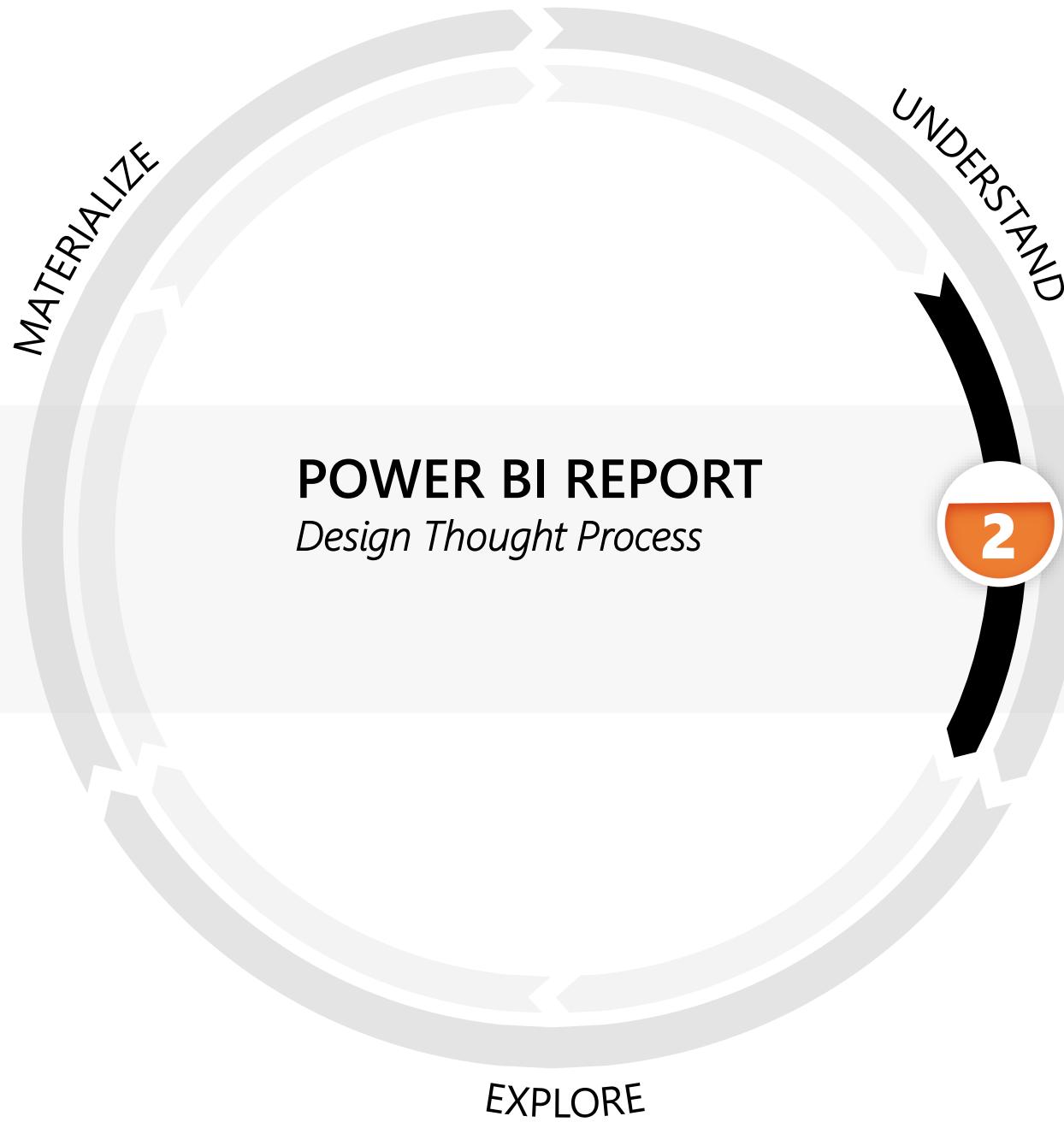
Testing



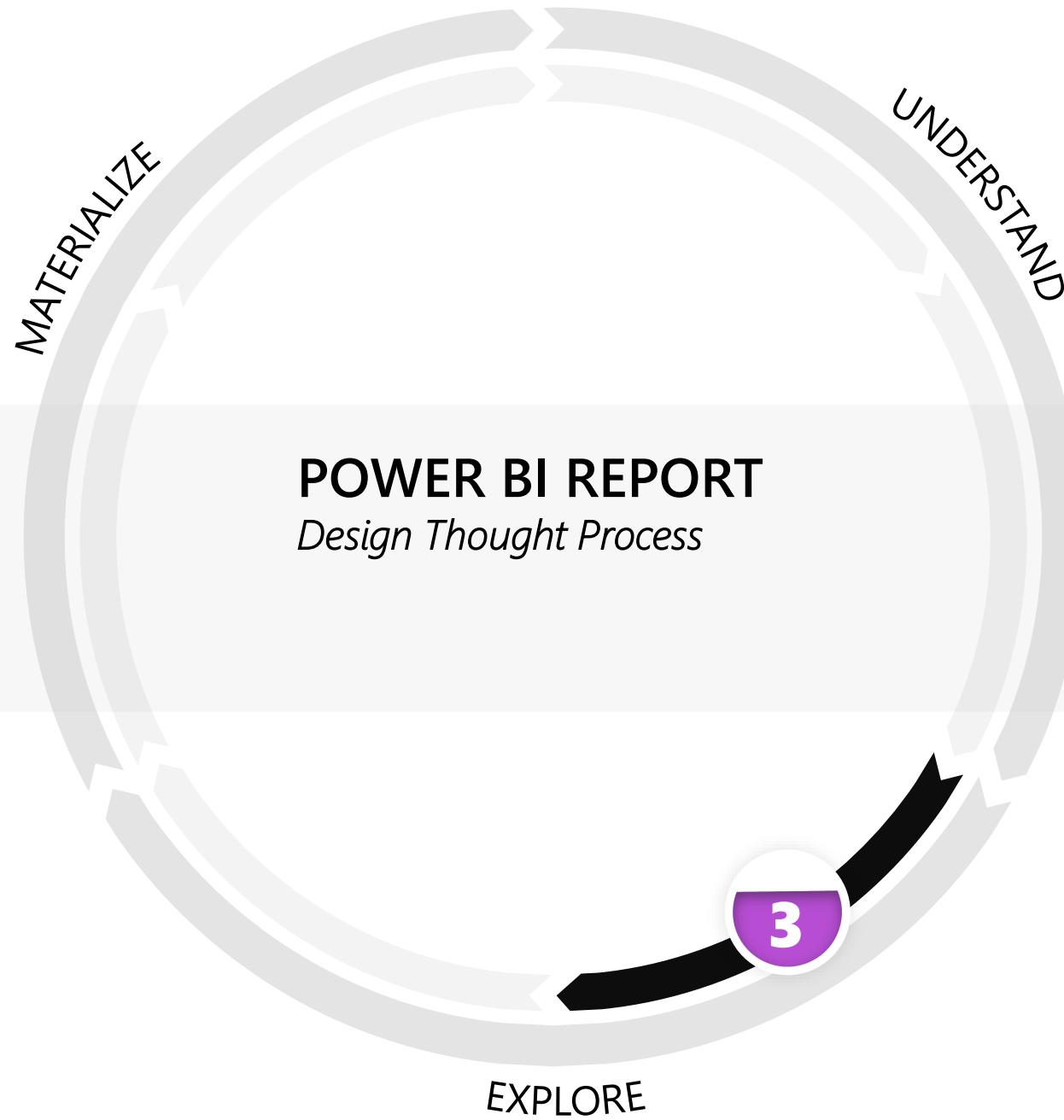


**EMPATHIZE
WITH YOUR AUDIENCE**

Conduct research to develop an understanding of your users

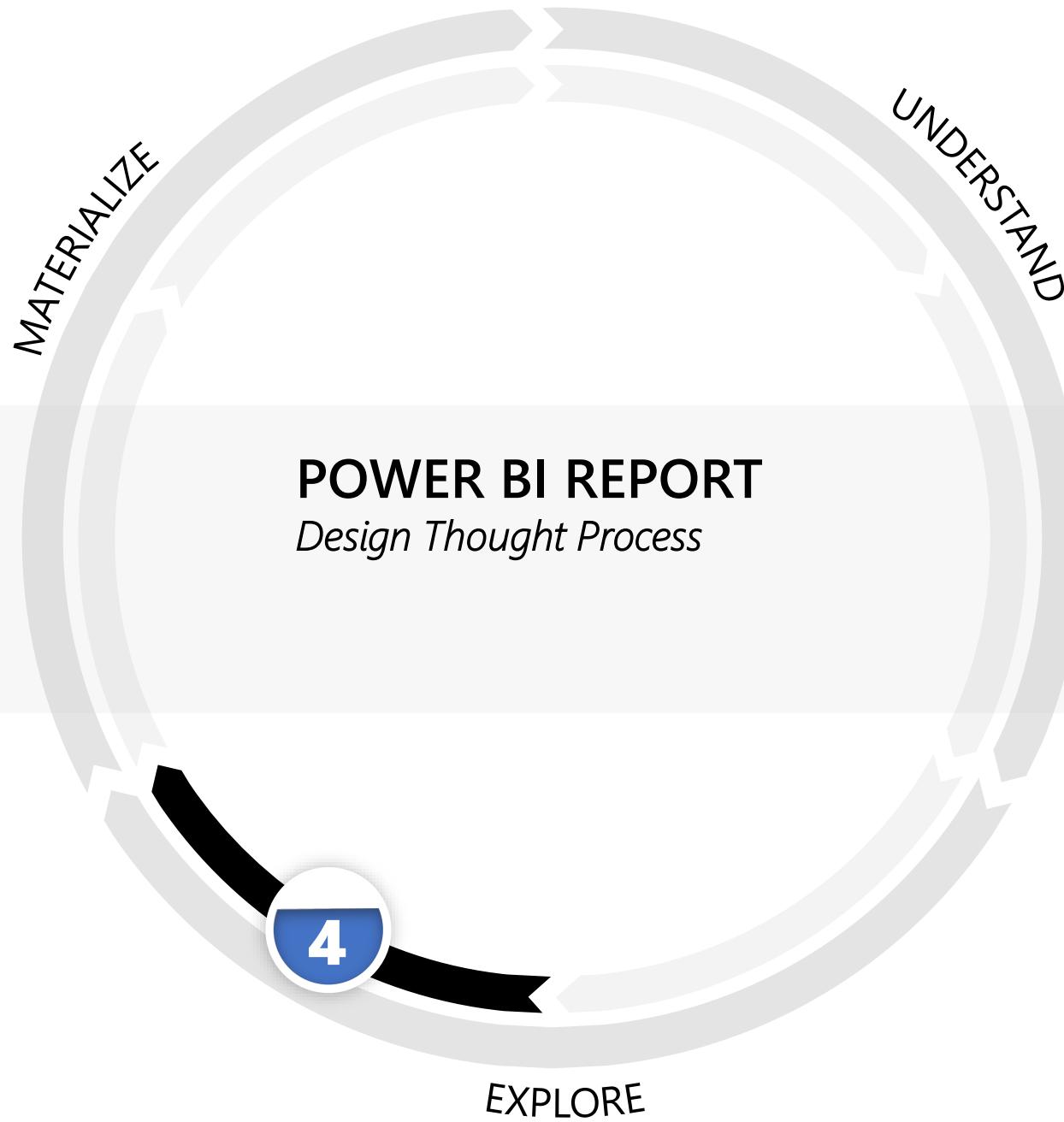


DEFINE
THE PLATFORM & ENVIRONMENT
Combine all research and observe where your users' problems & needs will exist



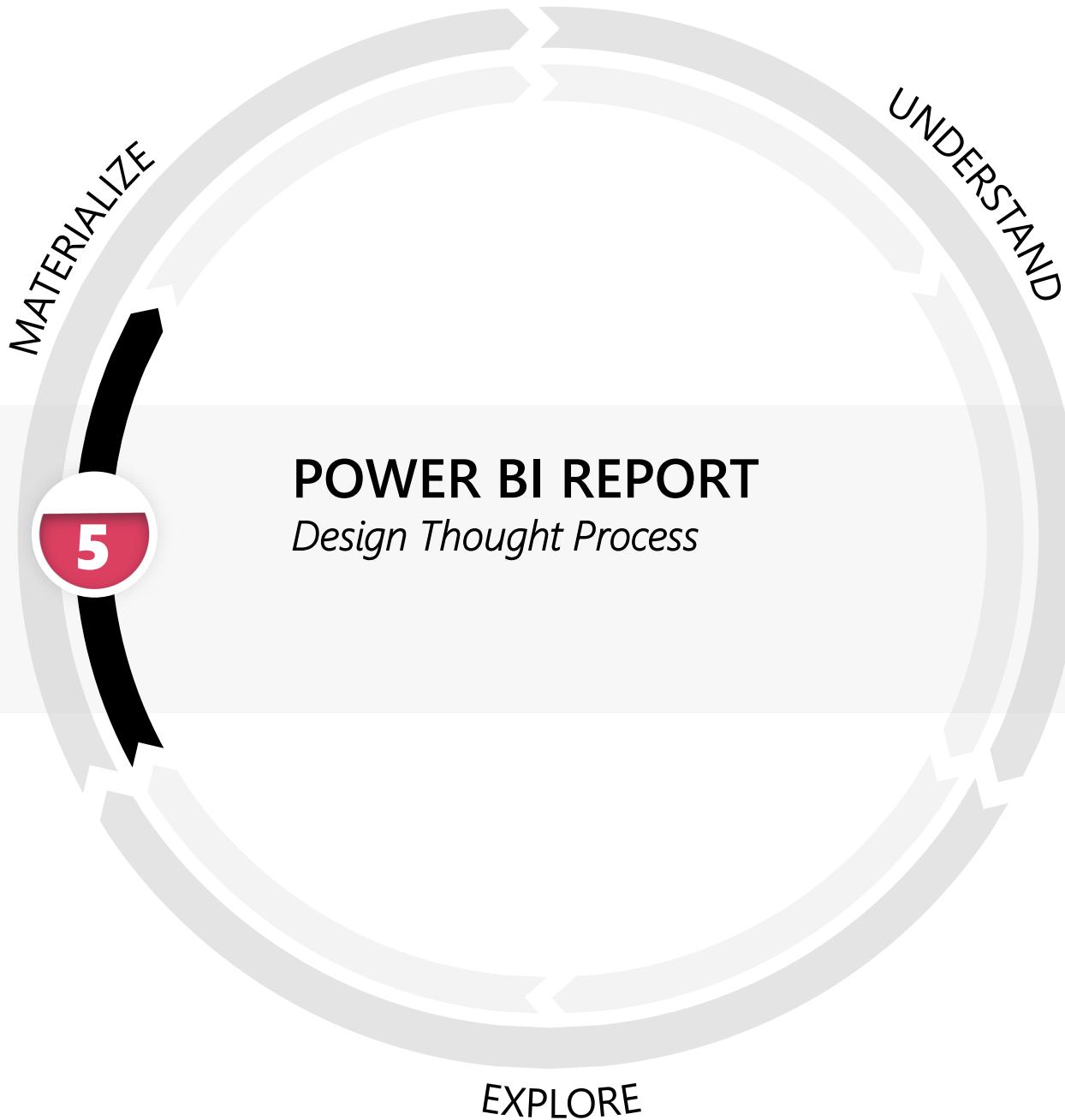
ENCOMPASS THE USER EXPERIENCE

Create easy, efficient, relevant, and all-around pleasant experiences for the user.

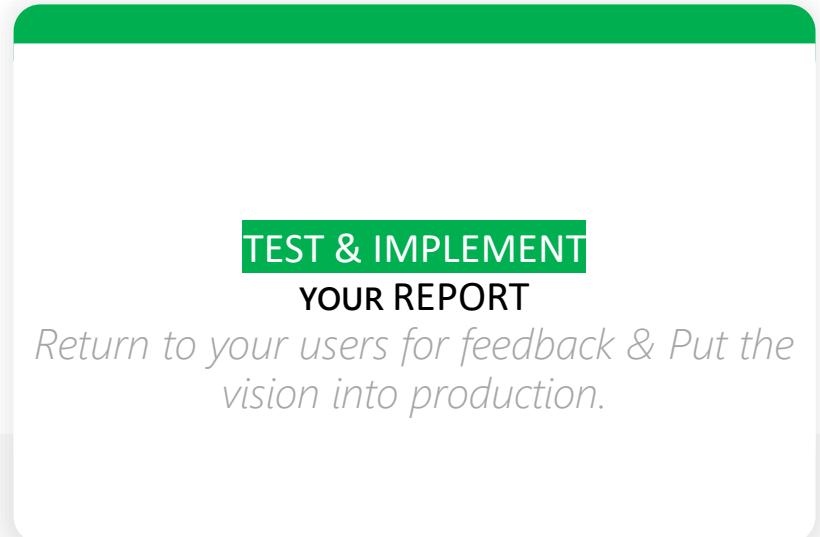
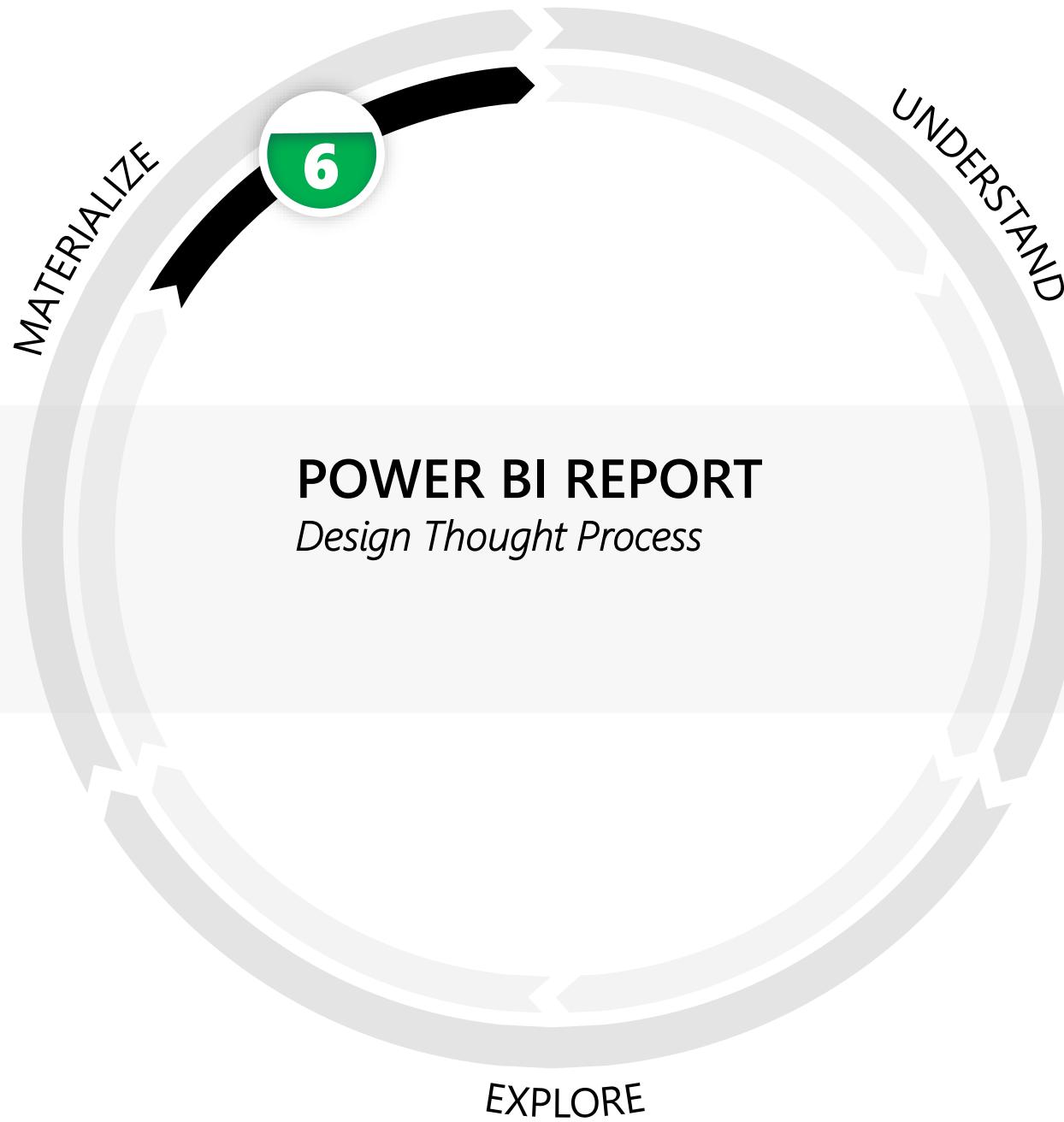


VISUALIZE YOUR DATA

Generate a range of alternative, creative visualizations to tell the stories your users want to know



DESIGN
THE USER INTERFACE
Build real, tactile representations for a range of your ideas



Effective design elements for practical application

USER EXPERIENCE



Navigation



Depth



Context

VISUALS



Visual selection



Formatting options



Conditional format



Dynamic

USER INTERFACE



Margins



Distribution



Alignment



Branding



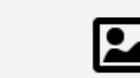
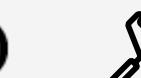
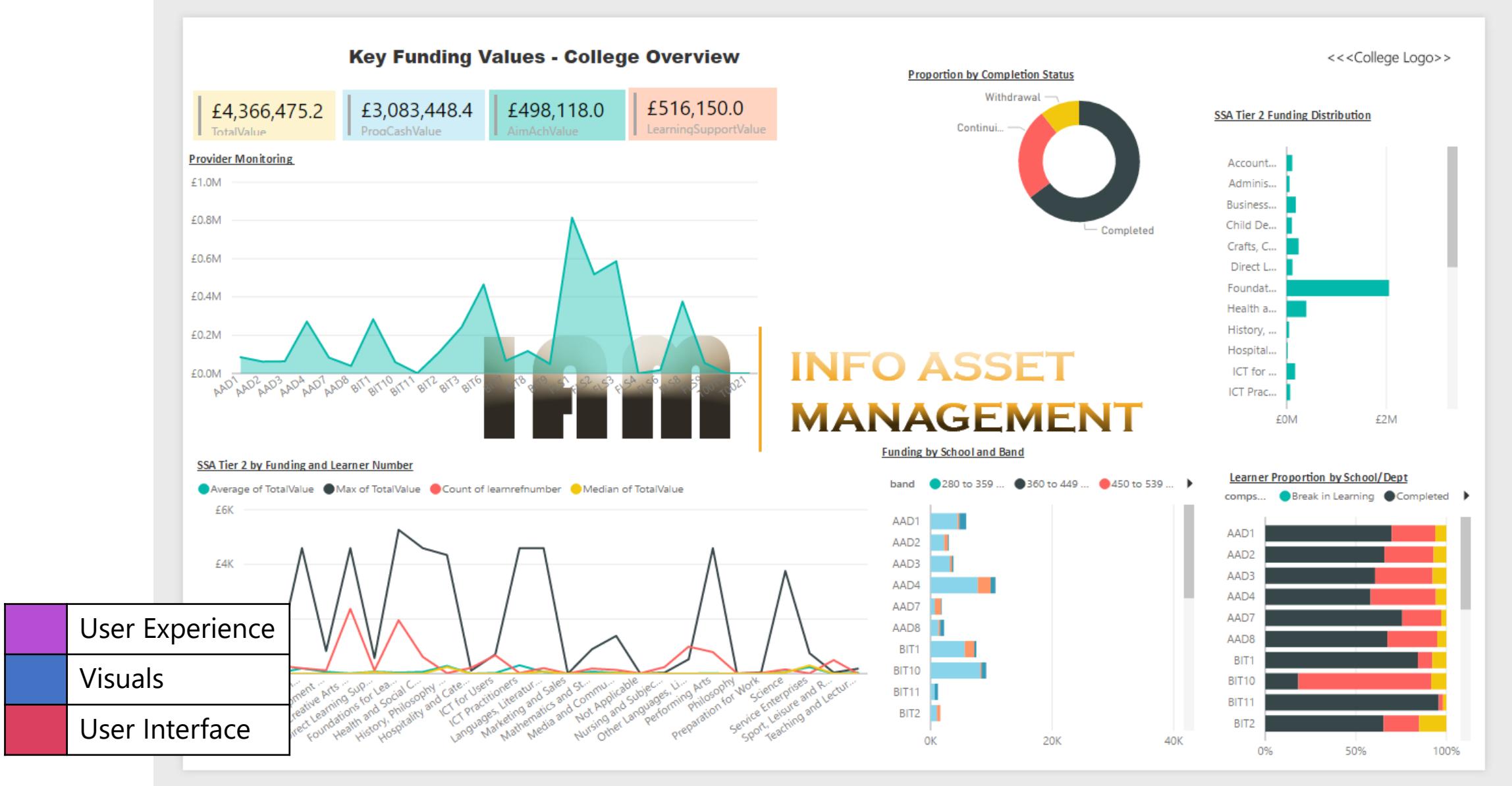
Color



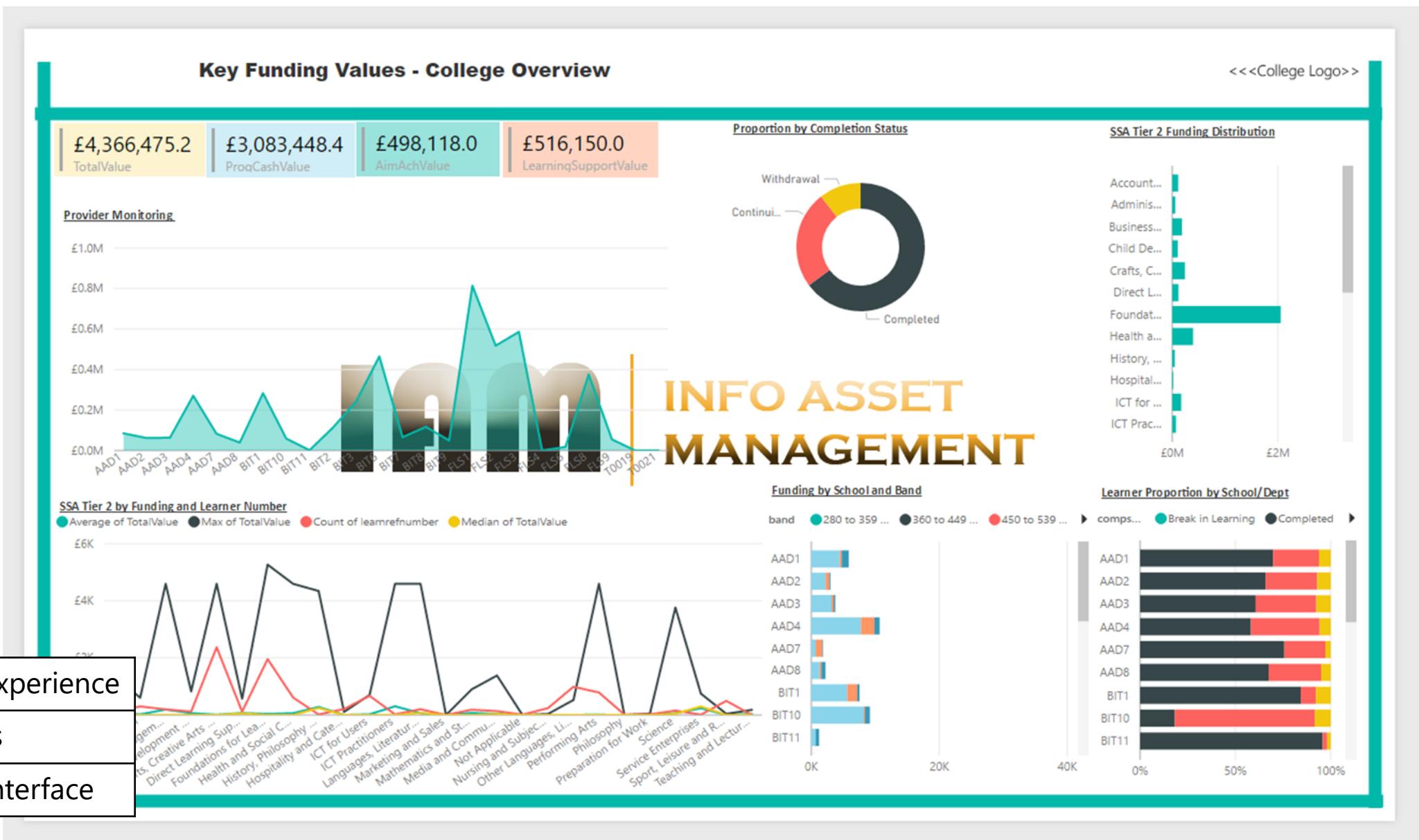
Background



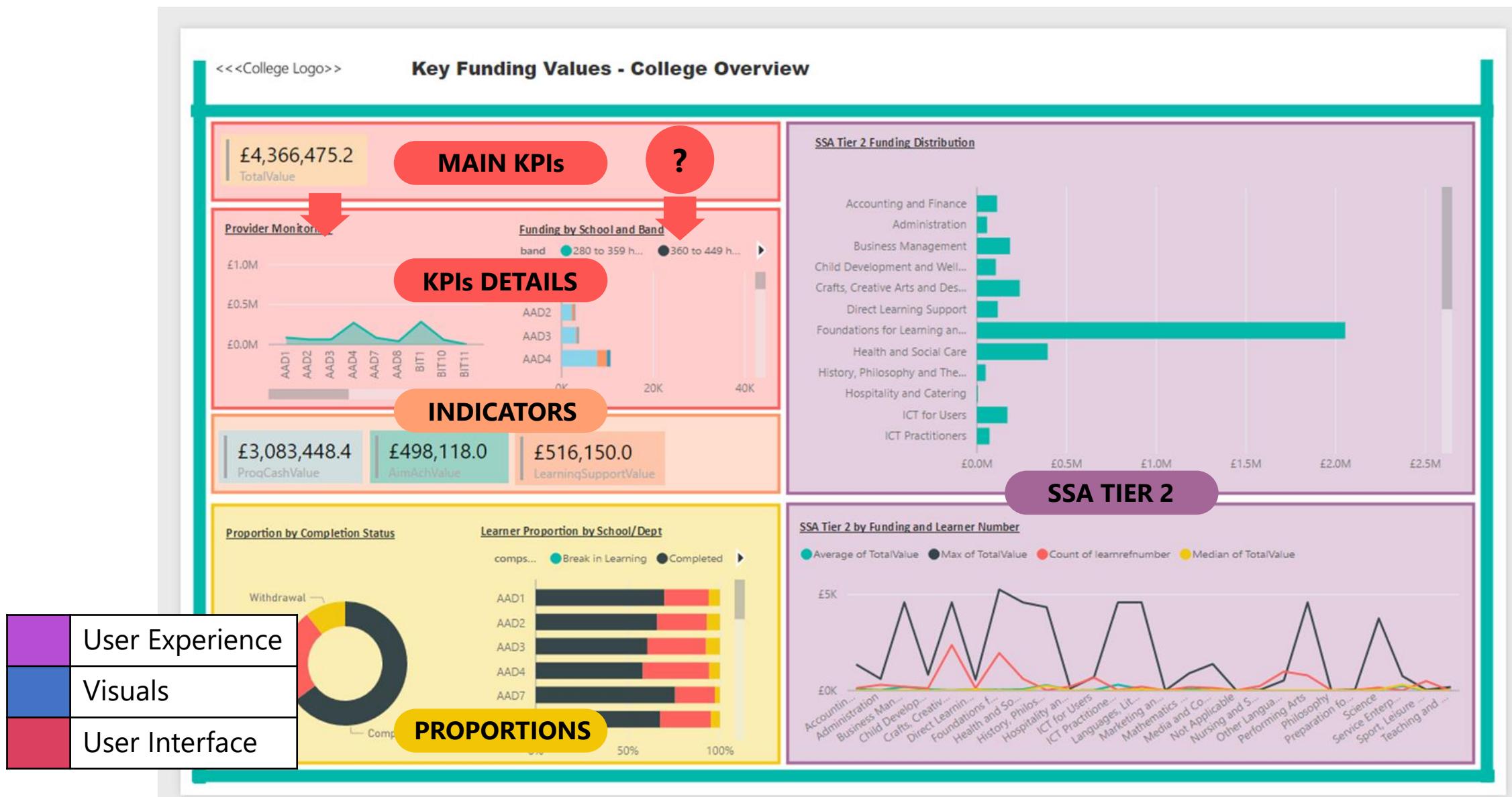
Text consistency



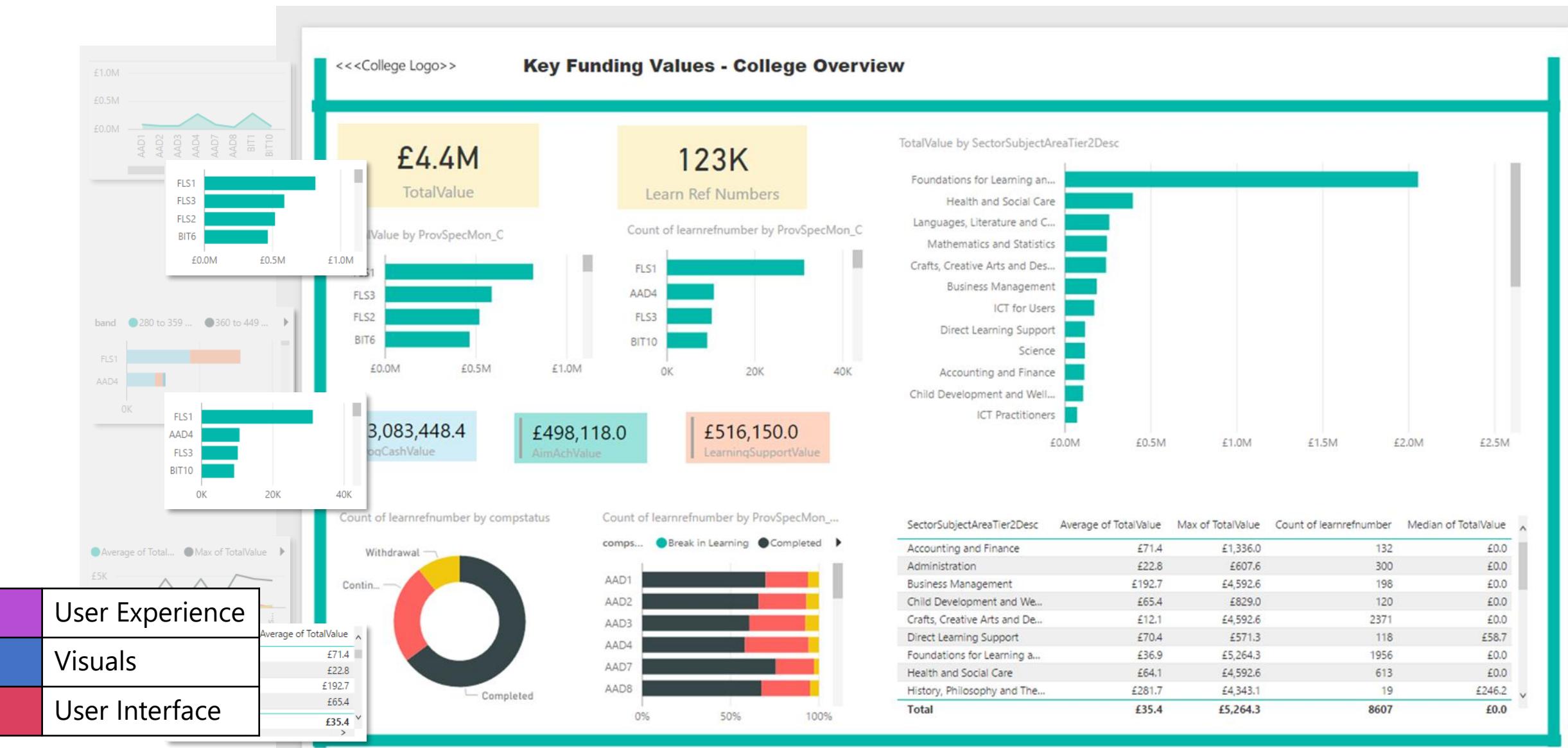
A-Z



A-Z



A-Z



A-Z

ALIGNMENT

SSA TIER 2

MAIN KPIs

KPIs DETAILS

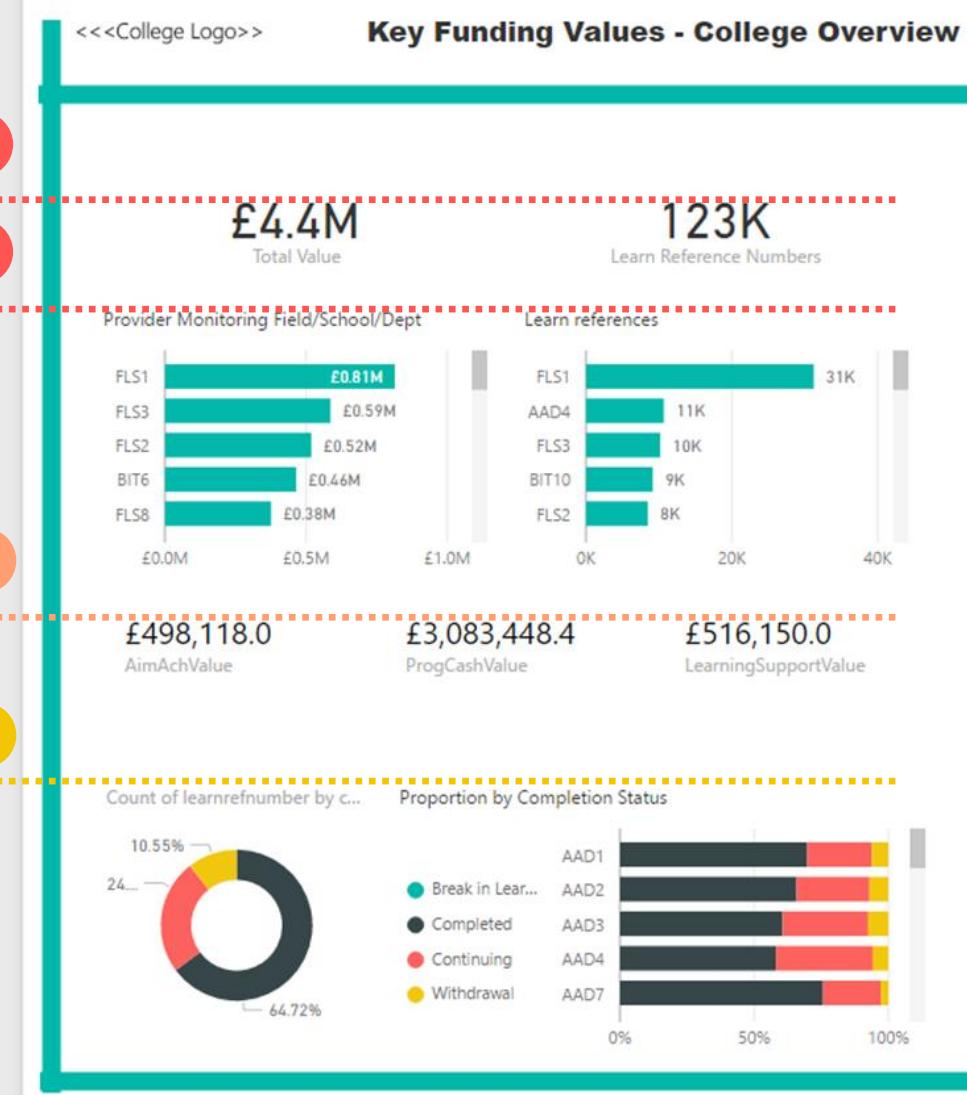
INDICATORS

PROPORTIONS

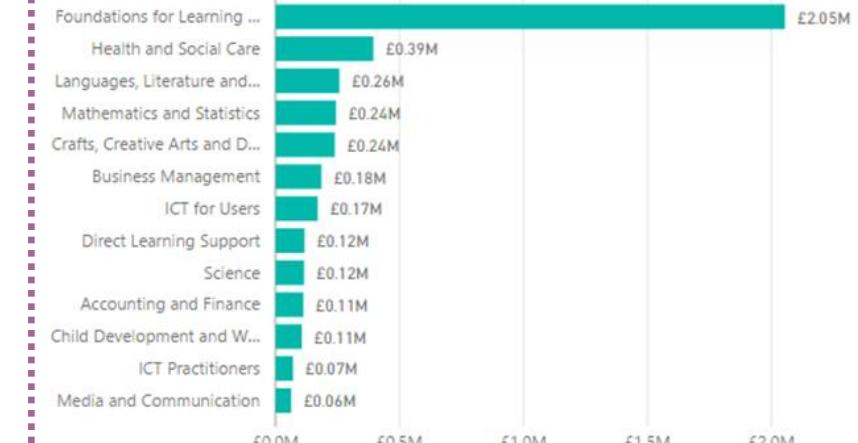
User Experience

Visuals

User Interface



TotalValue by SectorSubjectAreaTier2Desc



SectorSubjectAreaTier2Desc	Average of TotalValue	Max of TotalValue	Count of learnrefnumber	Median of TotalValue
ICT Practitioners	£305.4	£4,592.6	22	£0.0
History, Philosophy and Theology	£281.7	£4,343.1	19	£246.2
Service Enterprises	£239.1	£749.3	12	£299.7
Business Management	£192.7	£4,592.6	198	£0.0
Accounting and Finance	£71.4	£1,336.0	132	£0.0
Mathematics and Statistics	£71.3	£894.9	186	£0.0
Direct Learning Support	£70.4	£571.3	118	£58.7
Total	£35.4	£5,264.3	8607	£0.0



A-Z

BRANDING/TARGETING

FLEXIBLE LOGO



infoassetmanagement.com/products.html

Facebook Gmail Pinterest File sharing and Archivos - Dropbox Inicie sesión en su ImageShack - All Merging States -

IAM | INFO ASSET MANAGEMENT

HOME ABOUT US CONTACT US PRODUCTS SERVICES PRIVACY [f](#) [g](#) [t](#)

College Information Asset Manager

The amount of information within colleges is constantly expanding, and to stay competitive, your data-insight strategy has to keep pace with business change. With multiple sources of data, it can be difficult for your users to get access to the information they need, when they need it. College Information Asset Manager (CIAM) is not just a dashboard but your strategic information asset manager to help you address these challenges here.

Try CIAM

A screenshot of a web browser displaying the IAM website. The header features the IAM logo and navigation links for Home, About Us, Contact Us, Products, Services, and Privacy. Below the header, there's a section for social media sharing with icons for Facebook, Twitter, Print, Email, and Plus. A main text area discusses the challenges of managing information in colleges and introduces CIAM as a solution. To the right, there's a thumbnail image of a laptop screen showing a dashboard interface. A call-to-action button labeled 'Try CIAM' is visible. The overall design uses a dark blue background with white and light blue text.

User Experience
Visuals
User Interface

MAIN COLOR



COLOR

INFO ASSET MANAGEMENT

Key Funding Values
College Overview

band
All

£4.4M
Total Value

123K
Learn Reference Numbers

Provider Monitoring Field/School/Dept

FLS1	£0.81M
FLS3	£0.59M
FLS2	£0.52M
BIT6	£0.46M
FLS8	£0.38M
BIT1	£0.28M

Learn references

FLS1	31K
AAD4	11K
FLS3	10K
BIT10	9K
FLS2	8K
BIT1	8K

£498,118.0
AimAchValue

£3,083,448.4
ProgCashValue

£516,150.0
LearningSupportValue

Proportion by Completion Status

SSA Tier 2 Funding Distribution

SectorSubjectAreaTier2Desc	Average of TotalValue	Max of TotalValue	Count of learnrefnumber	Median of TotalValue
ICT Practitioners	£305.4	£4,592.6	22	£0.0
History, Philosophy and Theology	£281.7	£4,343.1	19	£246.2
Service Enterprises	£239.1	£749.3	12	£299.7
Business Management	£192.7	£4,592.6	198	£0.0
Accounting and Finance	£71.4	£1,336.0	132	£0.0
Mathematics and Statistics	£71.3	£894.9	186	£0.0
Direct Learning Support	£70.4	£571.3	118	£58.7
Child Development and Well Being	£65.4	£829.0	120	£0.0
Languages, Literature and Culture of...	£65.1	£4,592.6	203	£72.7
Health and Social Care	£64.1	£4,592.6	613	£0.0
Science	£62.1	£3,757.6	156	£43.7
Total	£35.4	£5,264.3	8607	£0.0

User Experience

Visuals

User Interface



A-Z

BACKGROUND

INFO ASSET MANAGEMENT

Key Funding Values
College Overview

KEY METRICS

- £4.4M** Total Value
- 123K** Learn Reference Numbers

Provider Monitoring Field/School/Dept

Category	Value
FLS1	£0.81M
FLS3	£0.59M
FLS2	£0.52M
BIT6	£0.46M
FLS8	£0.38M
BIT1	£0.28M

Learn references

Category	Value
FLS1	31K
AAD4	11K
FLS3	10K
BIT10	9K
FLS2	8K
BIT1	8K

£498,118.0 AimAchValue

£3,083,448.4 ProgCashValue

£516,150.0 LearningSupportValue

Proportion by Completion Status

Status	Percentage
Break in Lear...	10.55%
Completed	64.72%
Continuing	23.73%
Withdrawal	0.70%

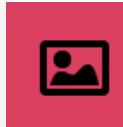
SSA Tier 2 Funding Distribution

SectorSubjectAreaTier2Desc	Average of TotalValue	Max of TotalValue	Count of learnrefnumber	Median of TotalValue
ICT Practitioners	£305.4	£4,592.6	22	£0.0
History, Philosophy and Theology	£281.7	£4,343.1	19	£246.2
Service Enterprises	£239.1	£749.3	12	£299.7
Business Management	£192.7	£4,592.6	198	£0.0
Accounting and Finance	£71.4	£1,336.0	132	£0.0
Mathematics and Statistics	£71.3	£894.9	186	£0.0
Direct Learning Support	£70.4	£571.3	118	£58.7
Child Development and Well Being	£65.4	£829.0	120	£0.0
Languages, Literature and Culture of ...	£65.1	£4,592.6	203	£72.7
Health and Social Care	£64.1	£4,592.6	613	£0.0
Science	£62.1	£3,757.6	156	£43.7
Total	£35.4	£5,264.3	8607	£0.0

User Experience

Visuals

User Interface



A-Z

CONTEXT AND DEPTH

INFO ASSET MANAGEMENT | **Key Funding Values** | **College Overview**

KEY METRICS

- £4.4M** Total Value
- 123K** Learn Reference Numbers

Provider Monitoring Field/School/Dept

Field/School/Dept	Total Value
FLS1	£0.81M
FLS3	£0.59M
FLS2	£0.52M
BIT6	£0.46M
FLS8	£0.38M
BIT1	£0.28M

Learn references Drill Through →

Learn reference	Count
FLS1	31K
AAD4	11K
FLS3	10K
BIT10	9K
FLS2	8K
BIT1	8K

£498,118.0 AimAchValue

£3,083,448.4 ProgCashValue

£516,150.0 LearningSupportValue

Proportion by Completion Status

Completion Status	Percentage
Break in Lear...	10.55%
Completed	64.72%
Continuing	24.73%
Withdrawal	0.00%

SSA Tier 2 Funding Distribution

band: All

Total value Trend

SectorSubjectAreaTier2Desc	Average of TotalValue	Max of TotalValue	Count of learnrefnumber	Median of TotalValue
ICT Practitioners	£305.4	£4,592.6	22	£0.0
History, Philosophy and Theology	£281.7	£4,343.1	19	£246.2
Service Enterprises	£239.1	£749.3	12	£299.7
Business Management	£192.7	£4,592.6	198	£0.0
Accounting and Finance	£71.4	£1,336.0	132	£0.0
Mathematics and Statistics	£71.3	£894.9	186	£0.0
Direct Learning Support	£70.4	£571.3	118	£58.7
Child Development and Well Being	£65.4	£829.0	120	£0.0
Languages, Literature and Culture of ...	£65.1	£4,592.6	203	£72.7
Health and Social Care	£64.1	£4,592.6	613	£0.0
Science	£62.1	£3,757.6	156	£43.7
Total	£35.4	£5,264.3	8607	£0.0

User Experience
Visuals
User Interface



TEXT

Font
Segoe UI

Sizes

8
9
14
30

Types
Normal
Bold

	User Experience
	Visuals
	User Interface

Key Funding Values
College Overview

band
All

KEY METRICS

£4.4M Total Value 123K Learn Reference Numbers

Provider Monitoring Field/School/Dept Learn references

Category	Value
FLS1	£0.81M
FLS3	£0.59M
FLS2	£0.52M
BIT6	£0.46M
FLS8	£0.38M
BIT1	£0.28M

31K 11K 10K 9K 8K 8K

£498,118.0 AimAchValue £3,083,448.4 ProgCashValue £516,150.0 LearningSupportValue

Proportion by Completion Status

Completion Status	Percentage
Break in Learner	10.55%
Completed	64.72%
Continuing	24.73%
Withdrawal	0%

SSA Tier 2 Funding Distribution

Funding Category	Value (£M)
Foundations for Learning ...	£2.05M
Health and Social Care	£0.39M
Languages, Literature and...	£0.26M
Mathematics and Statistics	£0.24M
Crafts, Creative Arts and D...	£0.24M
Business Management	£0.18M
ICT for Users	£0.17M
Direct Learning Support	£0.12M
Science	£0.12M
Accounting and Finance	£0.11M
Child Development and W...	£0.11M
ICT Practitioners	£0.07M
Media and Communication	£0.06M
Other Languages, Literatu...	£0.06M

SectorSubjectAreaTier2Desc Average of TotalValue Max of TotalValue Count of learnrefnumber Median of TotalValue

SectorSubjectAreaTier2Desc	Average of TotalValue	Max of TotalValue	Count of learnrefnumber	Median of TotalValue
ICT Practitioners	£305.4	£4,592.6	22	£0.0
History, Philosophy and Theology	£281.7	£4,343.1	19	£246.2
Service Enterprises	£239.1	£4,592.6	12	£299.7
Business Management	£192.7	£4,592.6	198	£0.0
Accounting and Finance	£71.4	£1,336.0	132	£0.0
Mathematics and Statistics	£71.3	£894.9	186	£0.0
Direct Learning Support	£70.4	£571.3	118	£58.7
Child Development and Well Being	£65.4	£829.0	120	£0.0
Languages, Literature and Culture of ...	£65.1	£4,592.6	203	£72.7
Health and Social Care	£64.1	£4,592.6	613	£0.0
Science	£62.1	£3,757.6	156	£43.7
Total	£35.4	£5,264.3	8607	£0.0



A-Z

FINAL REPORT

IAM | INFO ASSET MANAGEMENT

Key Funding Values
College Overview

KEY METRICS

£4.4M Total Value	123K Learn Reference Numbers	
Provider Monitoring Field/School/Dept	Learn references	
<ul style="list-style-type: none"> FLS1 £0.81M FLS3 £0.59M FLS2 £0.52M BIT6 £0.46M FLS8 £0.38M BIT1 £0.28M 	<ul style="list-style-type: none"> FLS1 31K AAD4 11K FLS3 10K BIT10 9K FLS2 8K BIT1 8K 	
£498,118.0 AimAchValue	£3,083,448.4 ProgCashValue	£516,150.0 LearningSupportValue
<p>Proportion by Completion Status</p> <ul style="list-style-type: none"> Break in Lear... Completed Continuing Withdrawal 		

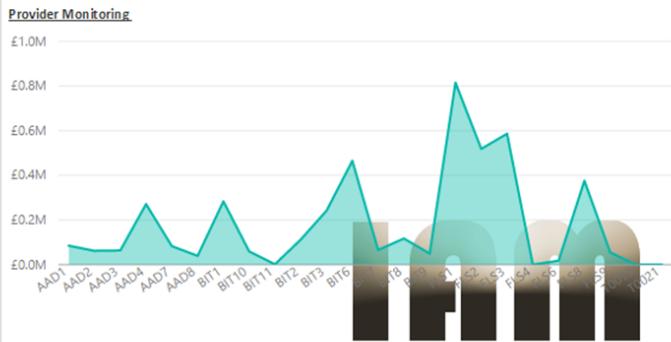
SSA Tier 2 Funding Distribution

Sector	Subject Area	Tier 2 Description	Average of Total Value	Max of Total Value	Count of LearnRefNumber	Median of Total Value
ICT Practitioners			£305.4	£4,592.6	22	£0.0
History, Philosophy and Theology			£281.7	£4,343.1	19	£246.2
Service Enterprises			£239.1	£749.3	12	£299.7
Business Management			£192.7	£4,592.6	198	£0.0
Accounting and Finance			£71.4	£1,336.0	132	£0.0
Mathematics and Statistics			£71.3	£894.9	186	£0.0
Direct Learning Support			£70.4	£571.3	118	£58.7
Child Development and Well Being			£65.4	£829.0	120	£0.0
Languages, Literature and Culture of ...			£65.1	£4,592.6	203	£72.7
Health and Social Care			£64.1	£4,592.6	613	£0.0
Science			£62.1	£3,757.6	156	£43.7
Total			£35.4	£5,264.3	8607	£0.0

band
All

Key Funding Values - College Overview

£4,366,475.2 TotalValue
£3,083,448.4 ProgCashValue
£498,118.0 AimAchValue
£516,150.0 LearningSupportValue

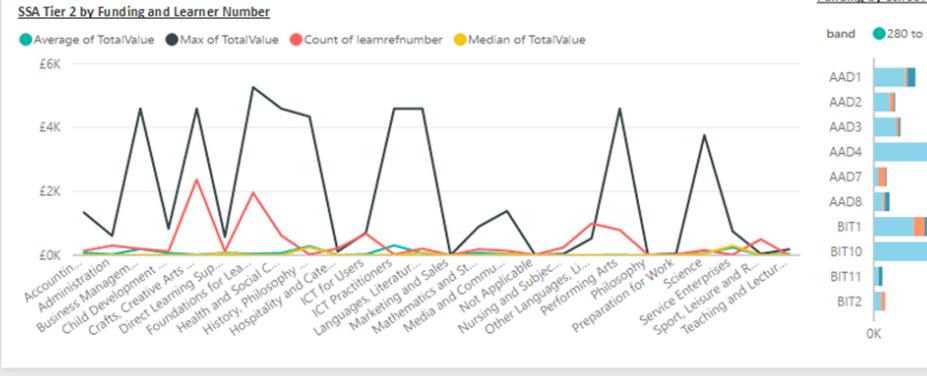


<<<College Logo>>

SSA Tier 2 Funding Distribution



INFO ASSET MANAGEMENT



Funding by School and Dept



INFO ASSET MANAGEMENT

Key Funding Values
College Overview

KEY METRICS

£4.4M

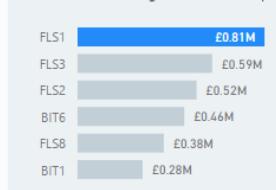
Total Value

123K

Learn Reference Numbers

Drill Through →

Provider Monitoring Field/School/Dept



£498,118.0

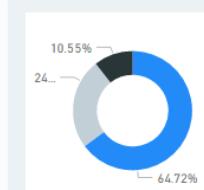
AimAchValue

£3,083,448.4

ProgCashValue

£516,150.0

LearningSupportValue



10.55%
24.72%
64.72%

Completed
Continuing
Withdrawal

Break in Learner Pathway

100%

SSA Tier 2 Funding Distribution



SectorSubjectAreaTier2Desc	Average of TotalValue	Max of TotalValue	Count of learnerrefnumber	Median of TotalValue
ICT Practitioners	£305.4	£4,592.6	22	£0.0
History, Philosophy and Theology	£281.7	£4,343.1	19	£246.2
Service Enterprises	£239.1	£4,592.6	12	£299.7
Business Management	£192.7	£4,592.6	198	£0.0
Accounting and Finance	£71.4	£1,336.0	132	£0.0
Mathematics and Statistics	£71.3	£894.9	186	£0.0
Direct Learning Support	£70.4	£571.3	118	£58.7
Child Development and Well Being	£65.4	£289.0	120	£0.0
Languages, Literature and Culture of ...	£65.1	£4,592.6	203	£72.7
Health and Social Care	£64.1	£4,592.6	613	£0.0
Science	£62.1	£3,757.6	156	£43.7
Total	£35.4	£2,644.3	8607	£0.0

Report Building Steps

1. Create visuals
2. Add margins *
3. Redistribute to logical sections *
4. Modify visual selection
5. Formatting & alignment
6. Background shapes *
7. Labeling *

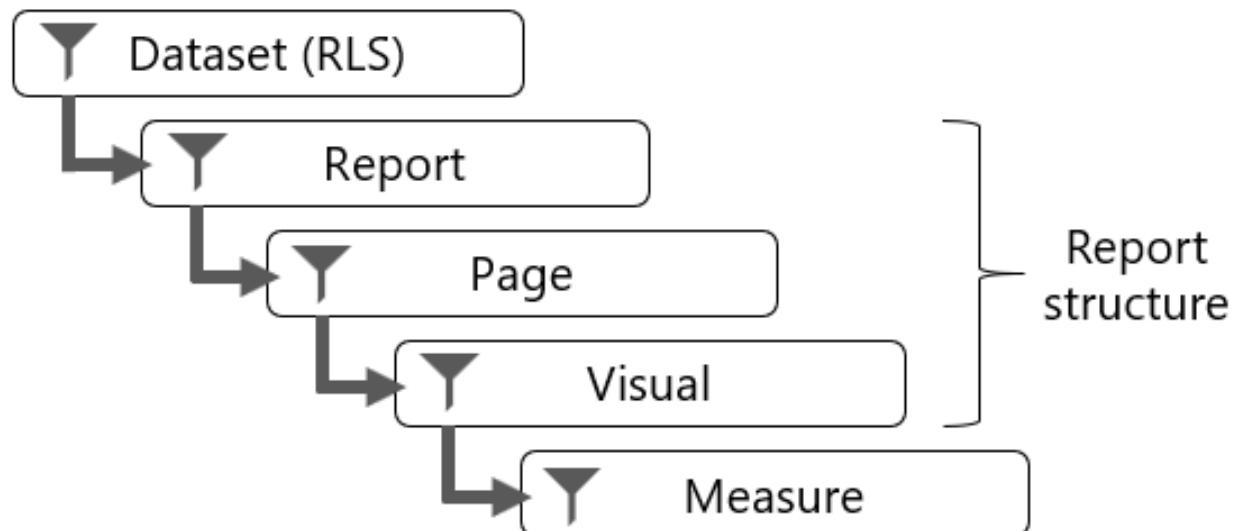
*Can partially be handled using a background image

Report Design Tools

- Format commands (Format ribbon, help to arrange objects)
- Selection pane (Layer order, Tab order)
- Report layout options (View ribbon, gridlines, snap to grid, lock objects)
- Page view (View ribbon, fit to page, fit to width, show actual size)
- Undo/redo (CTRL+Z or CTRL +Y)

Filtering

- You can filter at five different levels
 - Dataset (RLS)
 - Report
 - Page
 - Visual
 - Measure (using DAX/Report-level measure)
- Filter using
 - Filter pane
 - Slicers



Filter Pane or Slicer

Filter pane

- Advantages; same location, advanced filtering, better performance
- Disadvantages; less design flexibility, more difficult to see which filters are applied

Slicers

- Advantages; intuitive layout on report canvas, hierarchical slicer, filter context visible, support edit visual interaction, synced slicers
- Disadvantages; can impact performance, take up space in report canvas

Filtering Tips

- Use either filters or slicers.
 - Avoid using both filter techniques because it can create confusion.
- Use the clear all slicers button to reset to default values
- When a requirement is in place to lay out many slicers, consider creating a page or fly out pane that is dedicated to showing all slicers.

Design For Accessibility

- **Styling**
 - Larger fonts, color blindness. E.g., built-in theme Color blind safe or High contrast
- **Alt text**
 - Can describe the appearance and function of report objects to screen reader users
- **Tab order**
 - Build a logical sequence
- **Conditional formatting**
 - Use icons and colors

Before Publishing

Testing

- User acceptance test, test on different interfaces (Service, mobile app, embedded), Performance!

Prepare for publication

- Set intended initial experience, e.g., define first page, reset filters and slicers, visual drill state, sort orders, and button state

Provide support

- Training, built-in assistance, documentation

Manage change

- Often, it's better to publish a new report that includes many updates than release smaller incremental changes, consider using Deployment pipelines

And...

Validate The Performance

If it doesn't perform well in Power BI Desktop, don't expect it to fix itself in the Power BI service!

Performance analyzer

Start recording Refresh visuals Stop

Clear Export

Name	Duration (ms)
Recording started (9/18/2022 10:41:33 AM)	-
Changed page	-
Really really big Matrix	3958
Sales YoY%, Distinct Customers, Adjusted Sale...	1939
Sales Amount, Sales Quantity and Distinct Opp...	1905
Distinct Customers	1477
Continent Slicer	1748
Brand Slicer	1691
Year Slicer	1722
Total Sales Card	1520

Session evaluation



Event evaluation





Slides



https://github.com/BenniDeJagere/Presentations/{Year}/{YYYYMMDD}_{Event}



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](#)



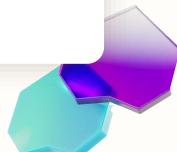
Copilot in Fabric

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



Security, Governance, and Administration

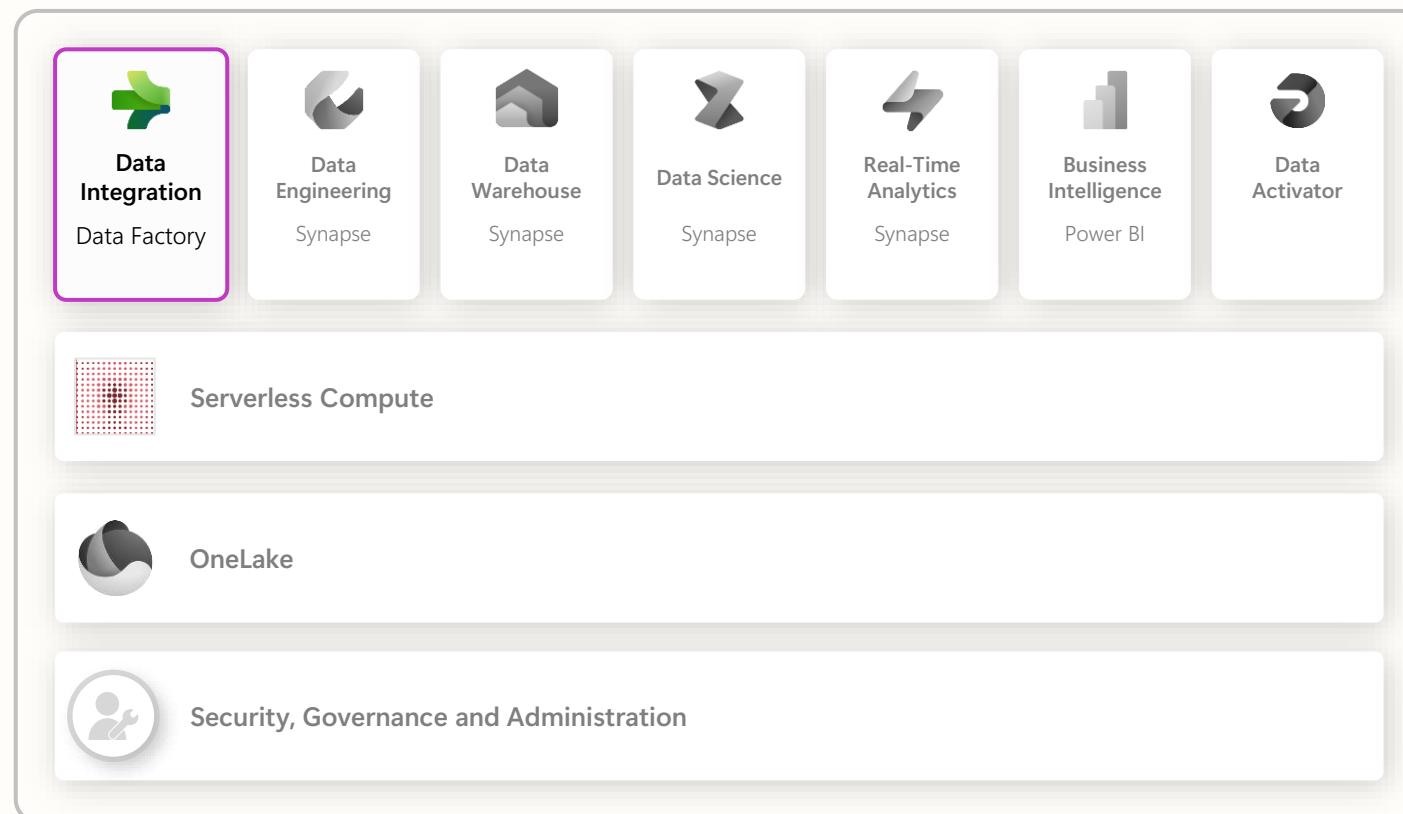
[Buy a Microsoft Fabric subscription](#) | [Fabric administration](#) | [Data governance and compliance](#) | [Security](#) | [See more](#)



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



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

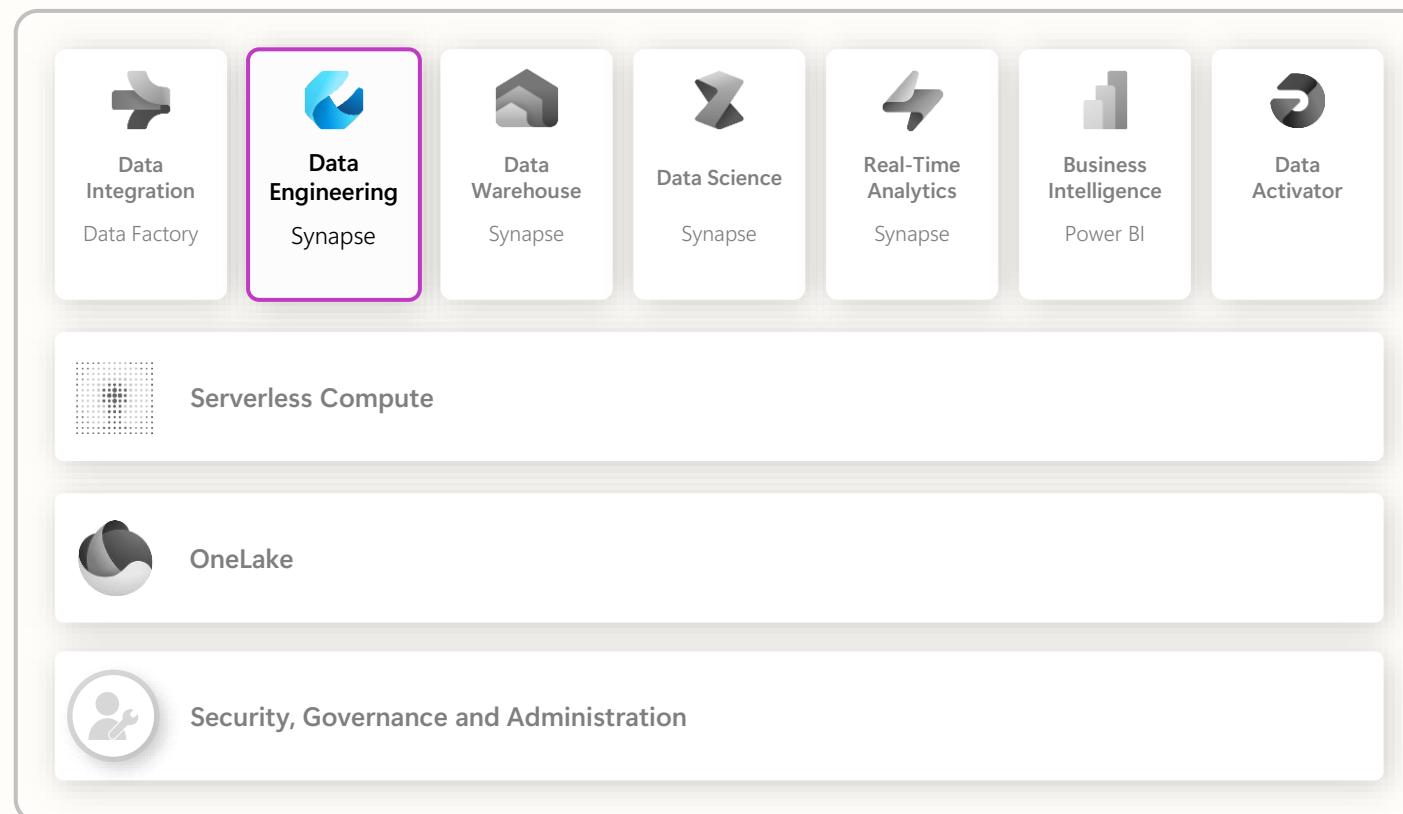
Petabyte scale transformation

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



World-class Spark serverless compute

Pro and low-code authoring experience

Schedule and orchestrate data transformations with notebooks and Spark jobs

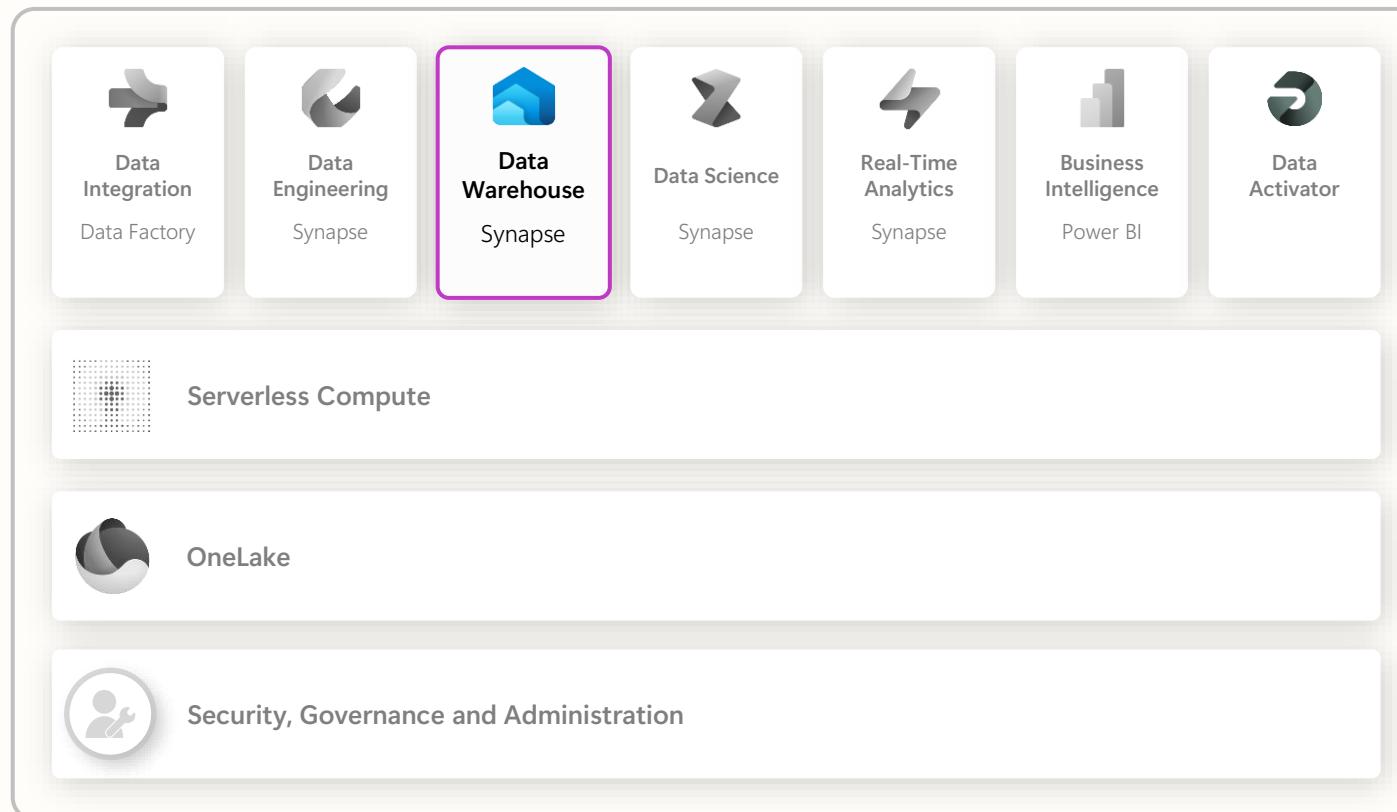
Launch clusters on demand and dynamically scale in, scale out, pause, and resume

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 duplicity

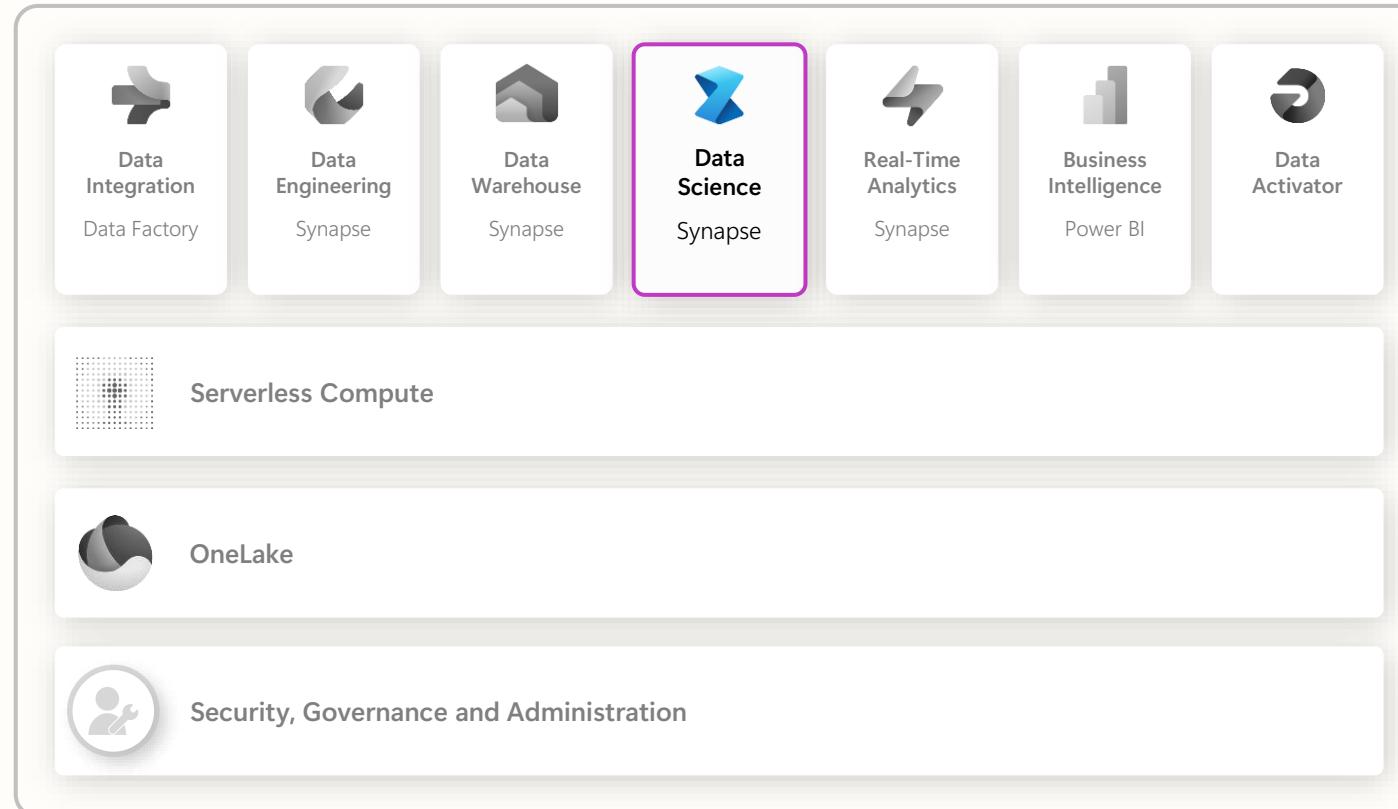
Consume data with Power BI for reporting and visualization

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

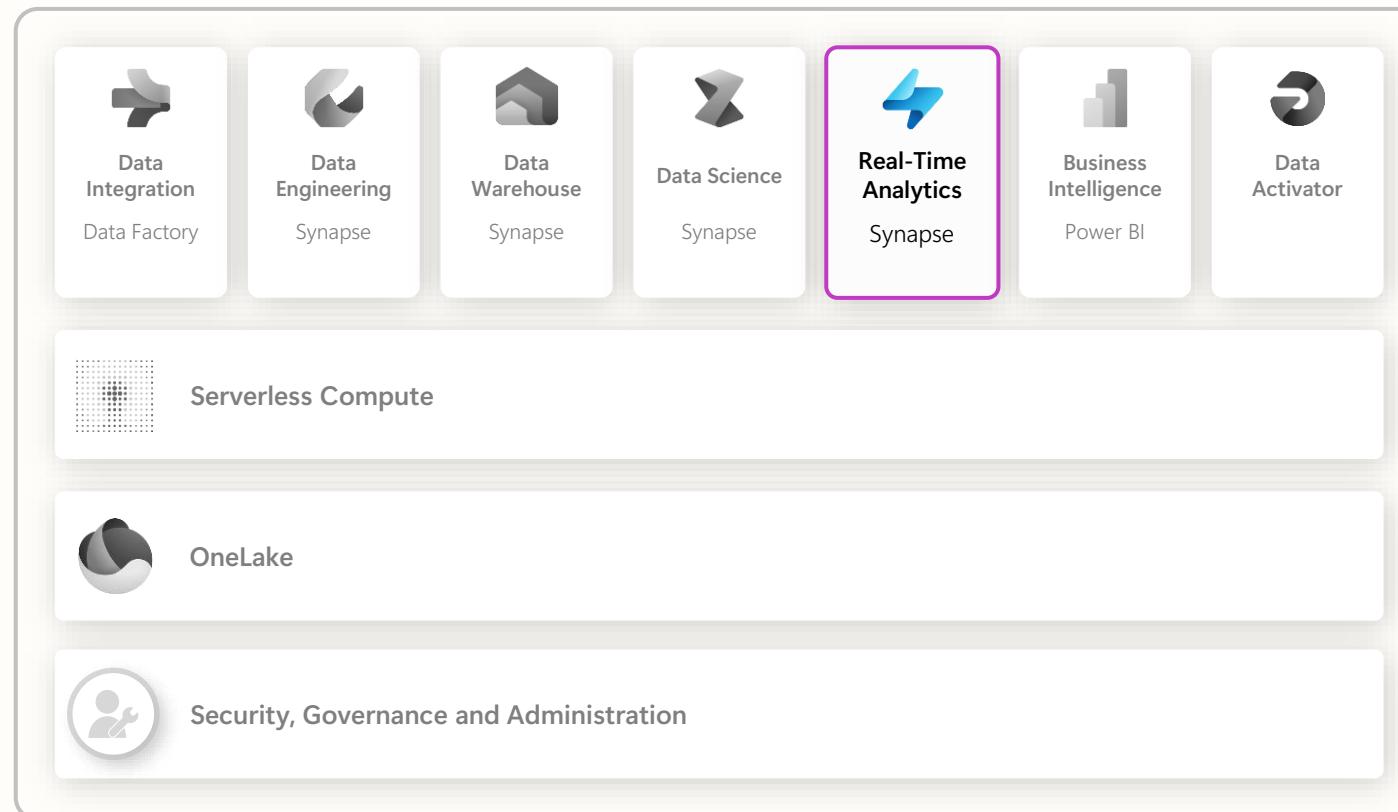
Perform transformation, exploration, and featurization by leveraging built-in experiences

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

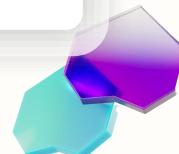


High velocity, low latency data analysis capable of indexing diverse data formats and structures up to several petabytes large

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

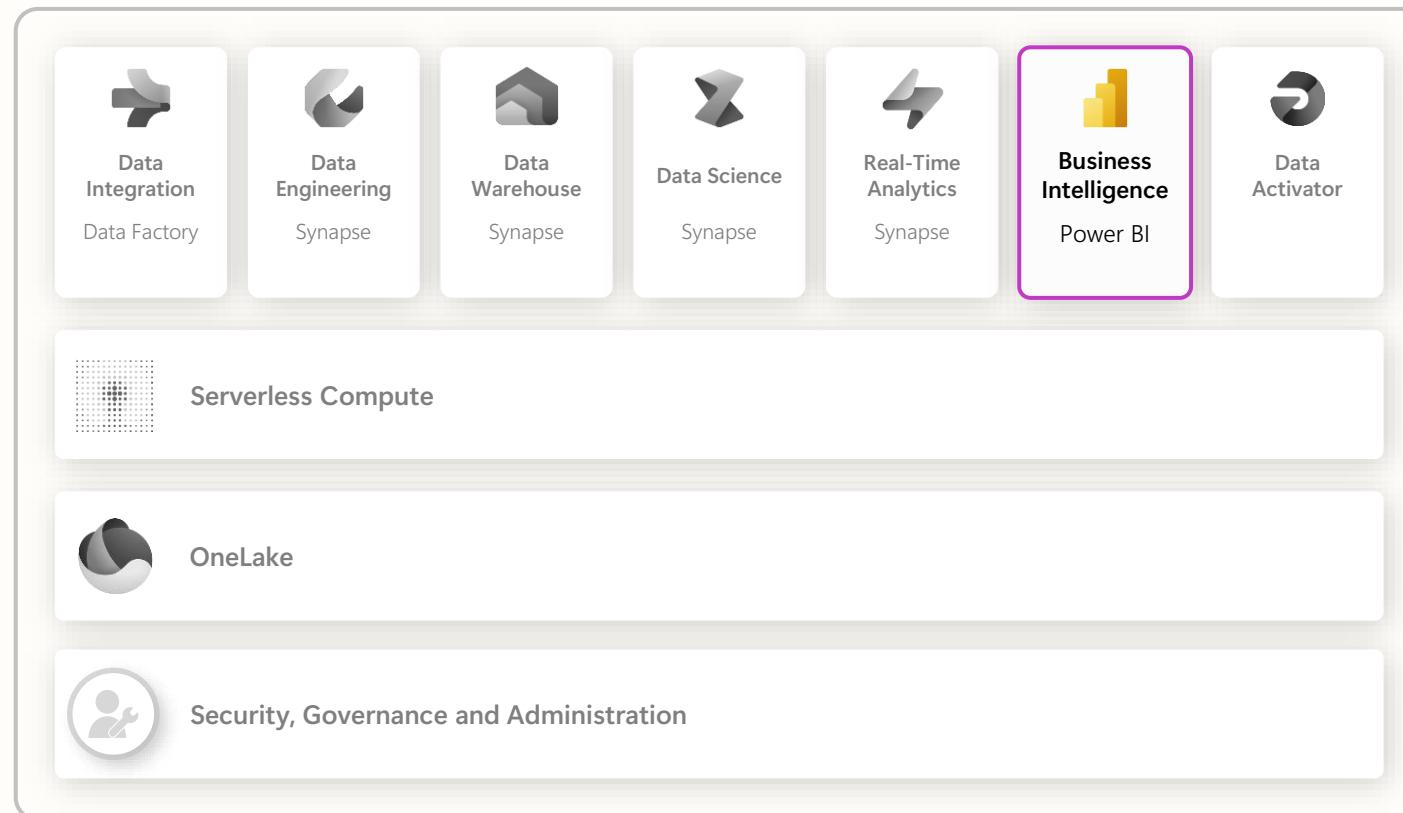
Easy ingestion of data from any source, in any data format, like applications, websites, and IoT devices

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

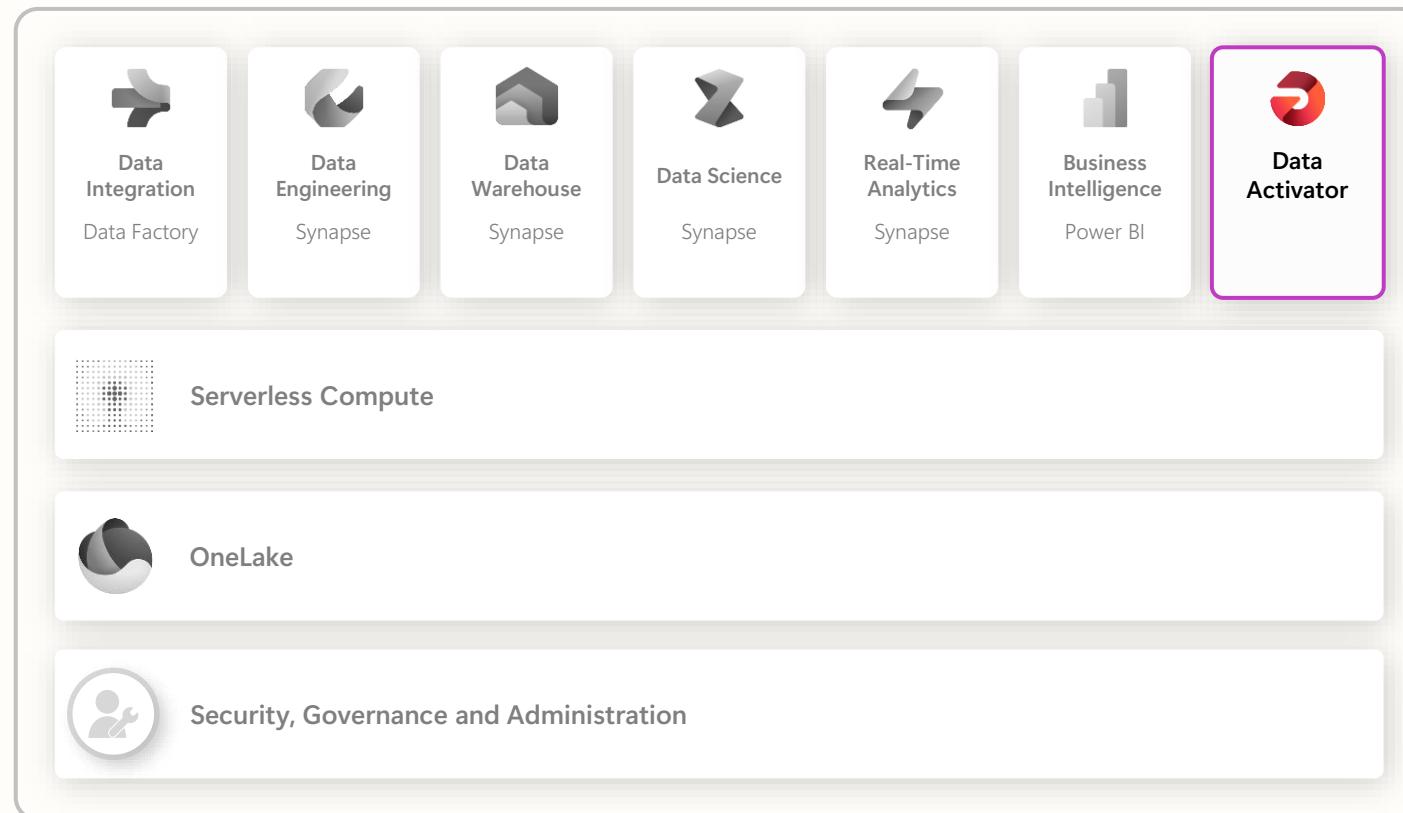
Connect to, index, and certify semantic models in the Power BI data hub

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

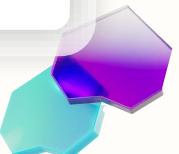


Automatically monitor Power BI semantic models and event streams.

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

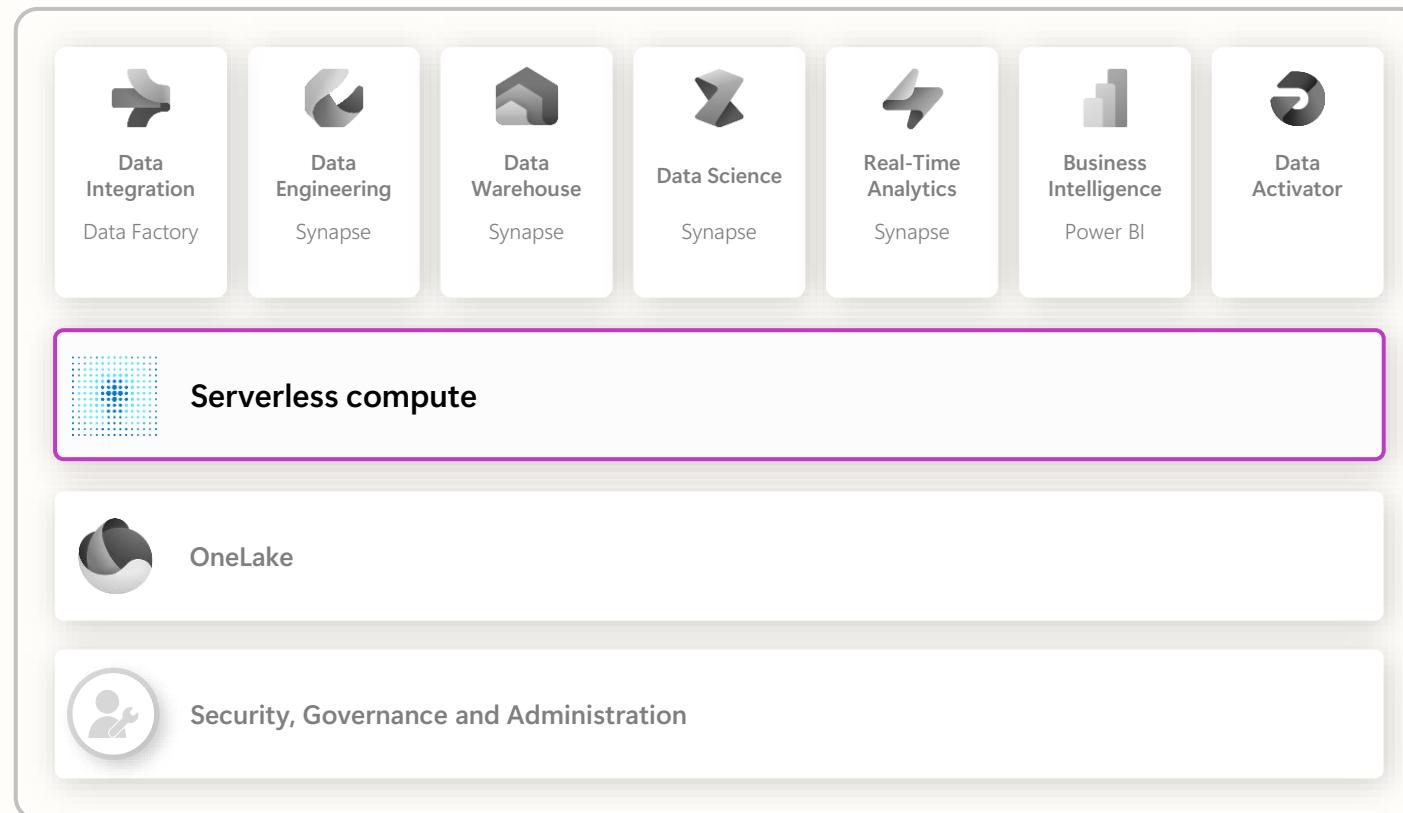
Send alerts in Microsoft Teams or email when actionable conditions occur.

Combine Data Activator with Power Automate to send alerts via 3rd 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.



Limitless scale for the most demanding jobs

Automatically create with each workspace

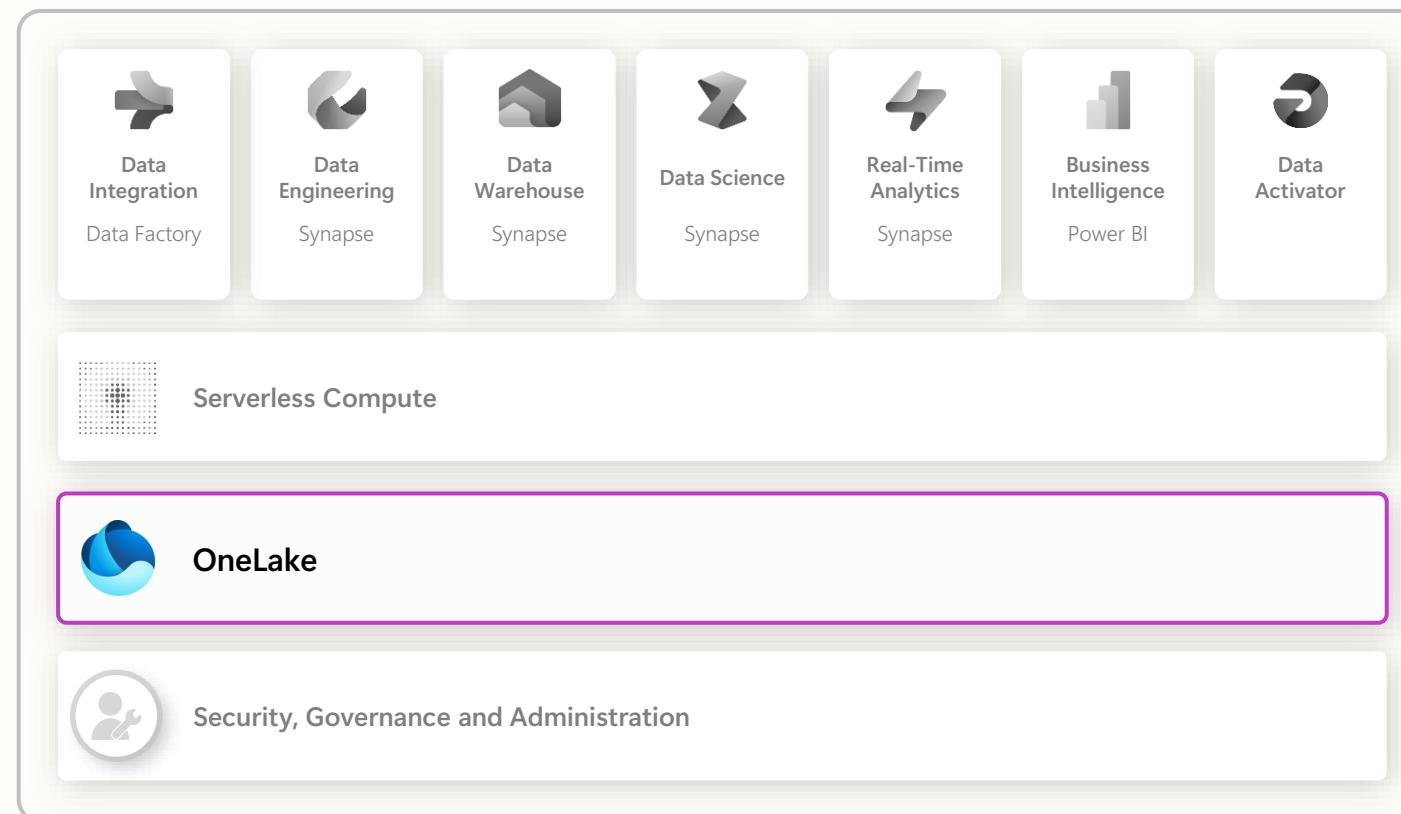
With support for high concurrency mode there's no need to start a new session to run a notebook

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



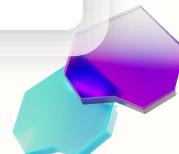
A single SaaS lake for the whole organization

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

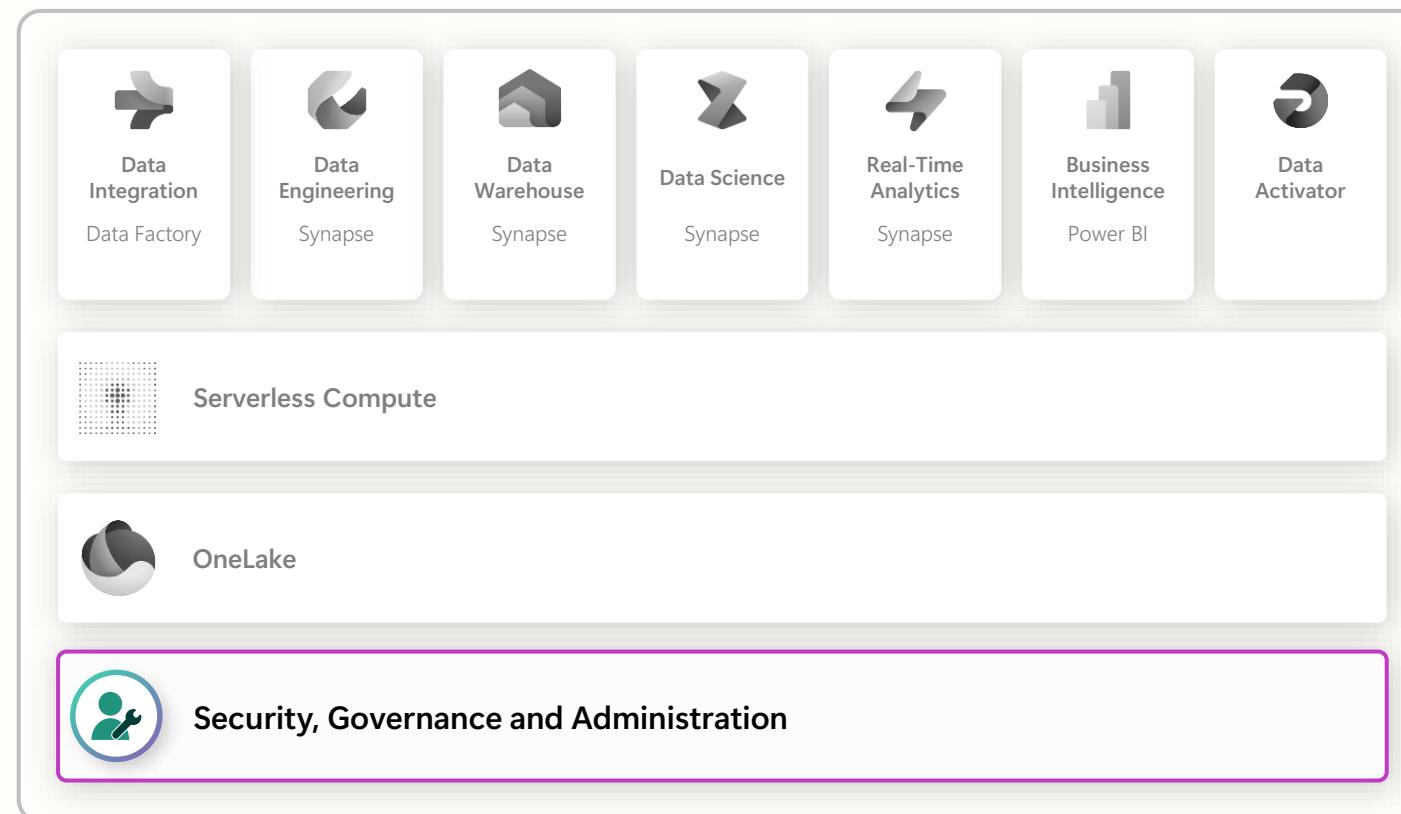
All data is organized in an intuitive hierarchical namespace

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

