

GLOBAL POWER PLATFORM BOOTCAMP MADRID 2024

Trabajando con Datos en
Microsoft Fabric de 0 a 100



Power Platform
Bootcamp

Power Platform Madrid



#GPPB2024



#PowerPlatformMadrid



Power Platform
Bootcamp



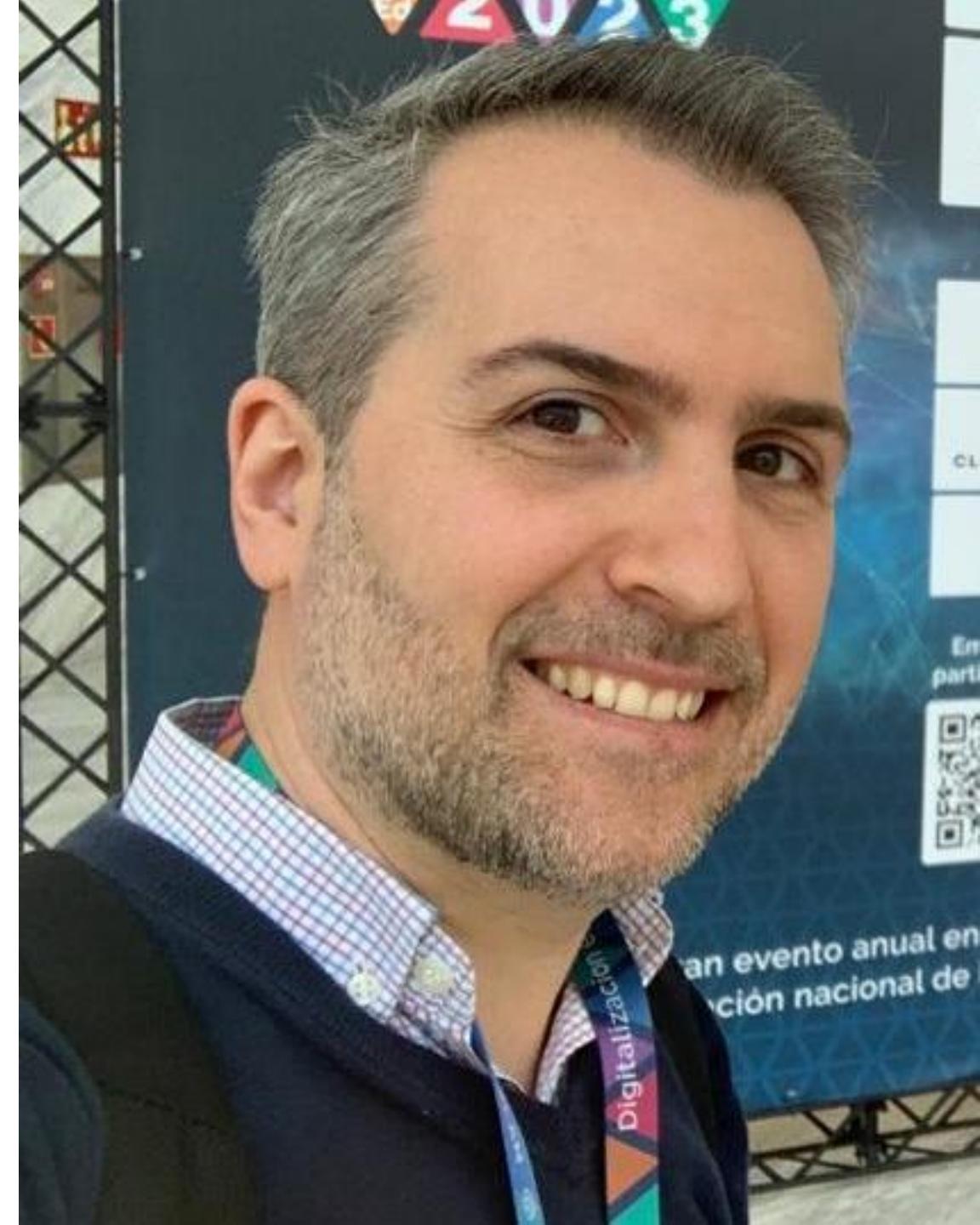
Power Platform Madrid



Mariano Kovo

- Microsoft MVP Data Platform
- Co-fundador Grupo Usuarios SQLArgentina
- Co-Fundador Precision IT (España / USA / Argentina)
- Trabajando con SQL 20+ años
- Blog: <https://precisionit.com.ar/blogs/marianokovo/>
- Email: mkovo@precisionit.tech
- LinkedIn: [/in/marianokovo/](https://in/marianokovo/)

 <https://www.linkedin.com/in/marianokovo/>
 @Mariano_Kovo
 www.PrecisionIT.Tech



Evento posible gracias a nuestros patrocinadores

Power Platform Madrid



#GPPB2024



#PowerPlatformMadrid

Agenda



#GPPB2024



#PowerPlatformMadrid

Agenda

Presentation	Labs
Introducción a Microsoft Fabric	
Overview de OneLake	Lab 1 – Fabric Workspace
Data Integration	Lab 2 – Experiencia Data Integration – Part 1
	Lab 3 – Experiencia Data Integration – Part 2
Break	
Data Engineering	Lab 4 – Trabajando con Delta Tables
Data Warehousing	Lab 5 – Experiencia Data Engineering
Capacidades y Licenciamiento	
Próximos pasos y recursos	Lab 6 – Limpieza de Entorno



#GPPB2024



#PowerPlatformMadrid

Introducción a Microsoft Fabric



#GPPB2024



#PowerPlatformMadrid

Analytics hasta hoy

El mundo de la analítica de datos tiene varios patrones predecibles.

Data Integration

Data Engineering

Data Warehousing

Real-Time Analytics

Data Science

Business Intelligence

Data Lake

Governance and Administration



#GPPB2024

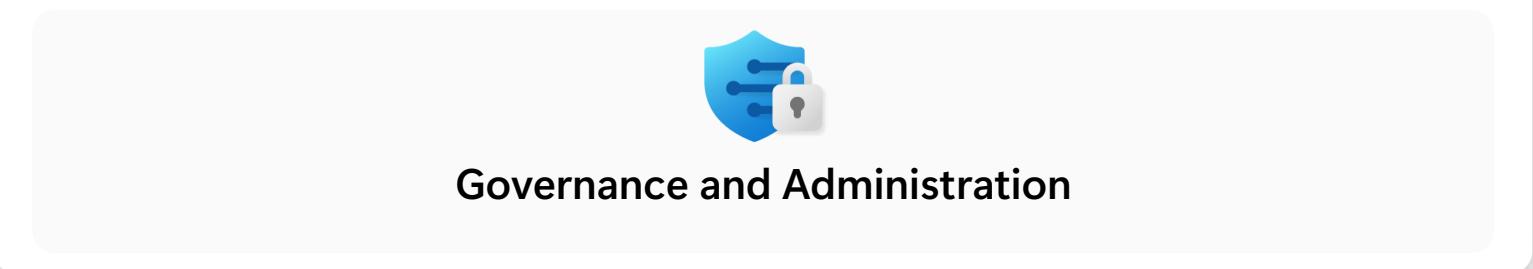
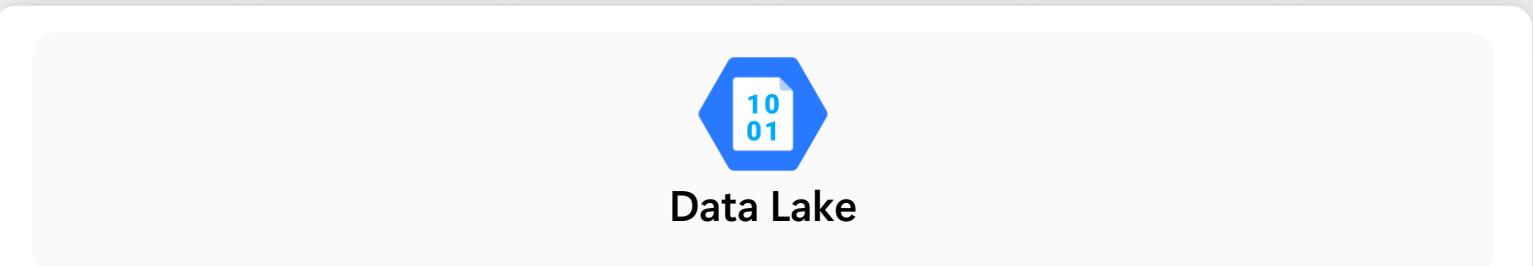
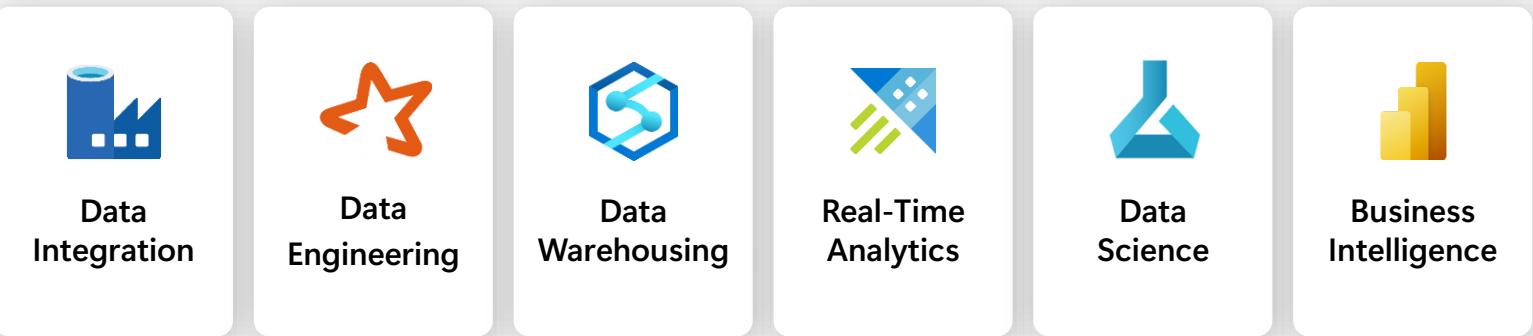


#PowerPlatformMadrid

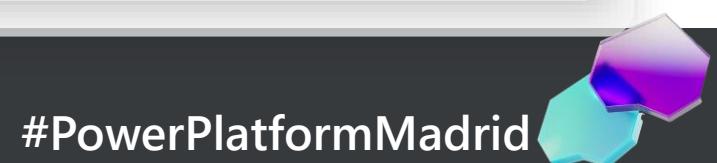
Analytics hasta hoy

El mundo de la analítica de datos tiene varios patrones predecibles.

Microsoft tiene todos los productos con la escala adecuada necesaria para crear un sistema de análisis completo.



#GPPB2024



#PowerPlatformMadrid

Aun sigue siendo complejo

Muchos productos

Experiencias diferentes

Propietario y abierto

Dedicado y sin servidor

PaaS y SaaS

Diferentes modelos de negocio

Curvas de aprendizaje pronunciadas

Se necesita una amplia experiencia

Alto esfuerzo de integración



Purview



Power BI



Kusto



Data Factory



Azure AI



Synapse DW



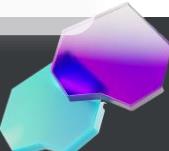
Synapse Spark



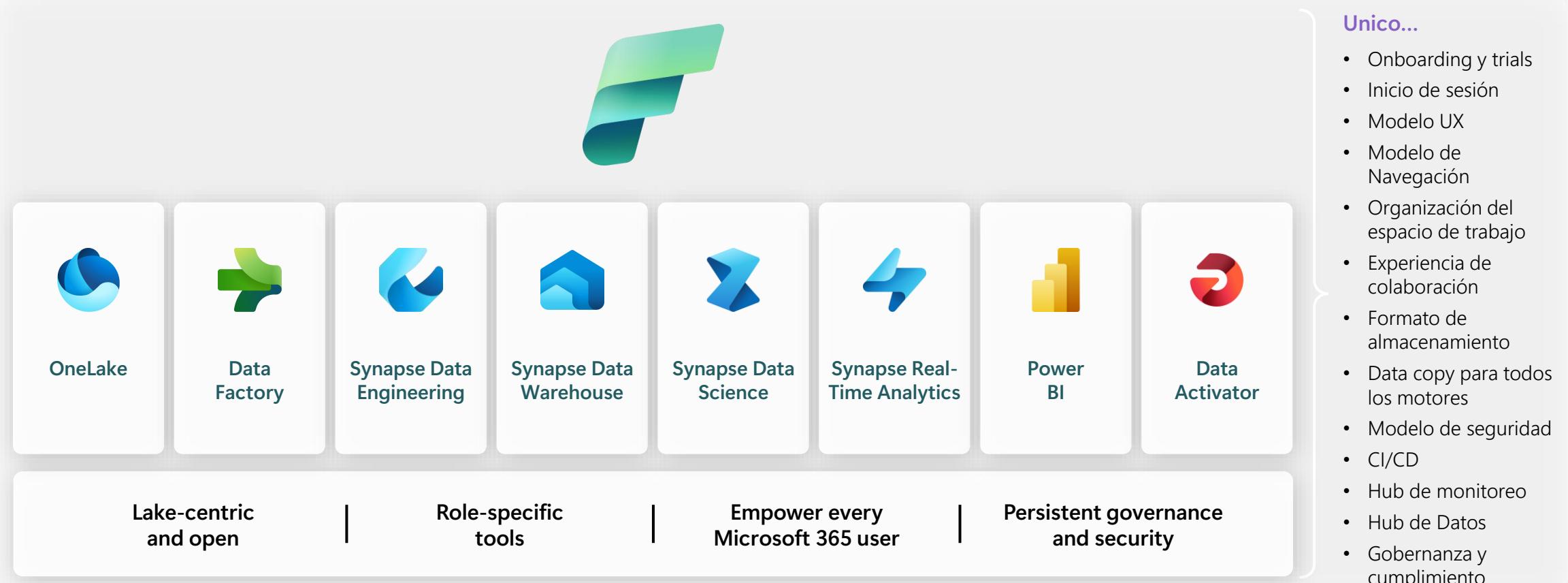
#GPPB2024



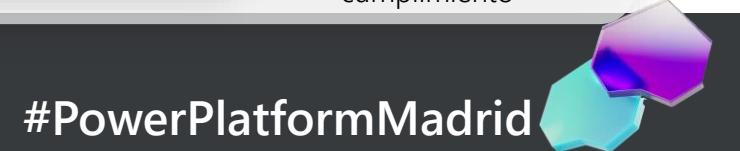
#PowerPlatformMadrid



Microsoft Fabric, una solución de análisis unificada para la era de la IA



#GPPB2024



#PowerPlatformMadrid



Microsoft Fabric

La plataforma de datos para la era de la IA

Plataforma de Analítica de datos Integral

Todo, unificado

SaaS-ificado

Asegurado y gobernado

Centrado en Datalake y abierto

OneLake

Copia Única

Abierto en todos los niveles

Potencia para todos los usuarios

Familiar e intuitivo

Integrado en Microsoft 365

De la visión a la acción

Integrado con IA

Acelerado por Copilot

ChatGPT en tus datos

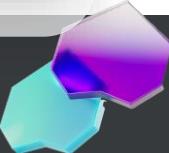
Información impulsada por IA



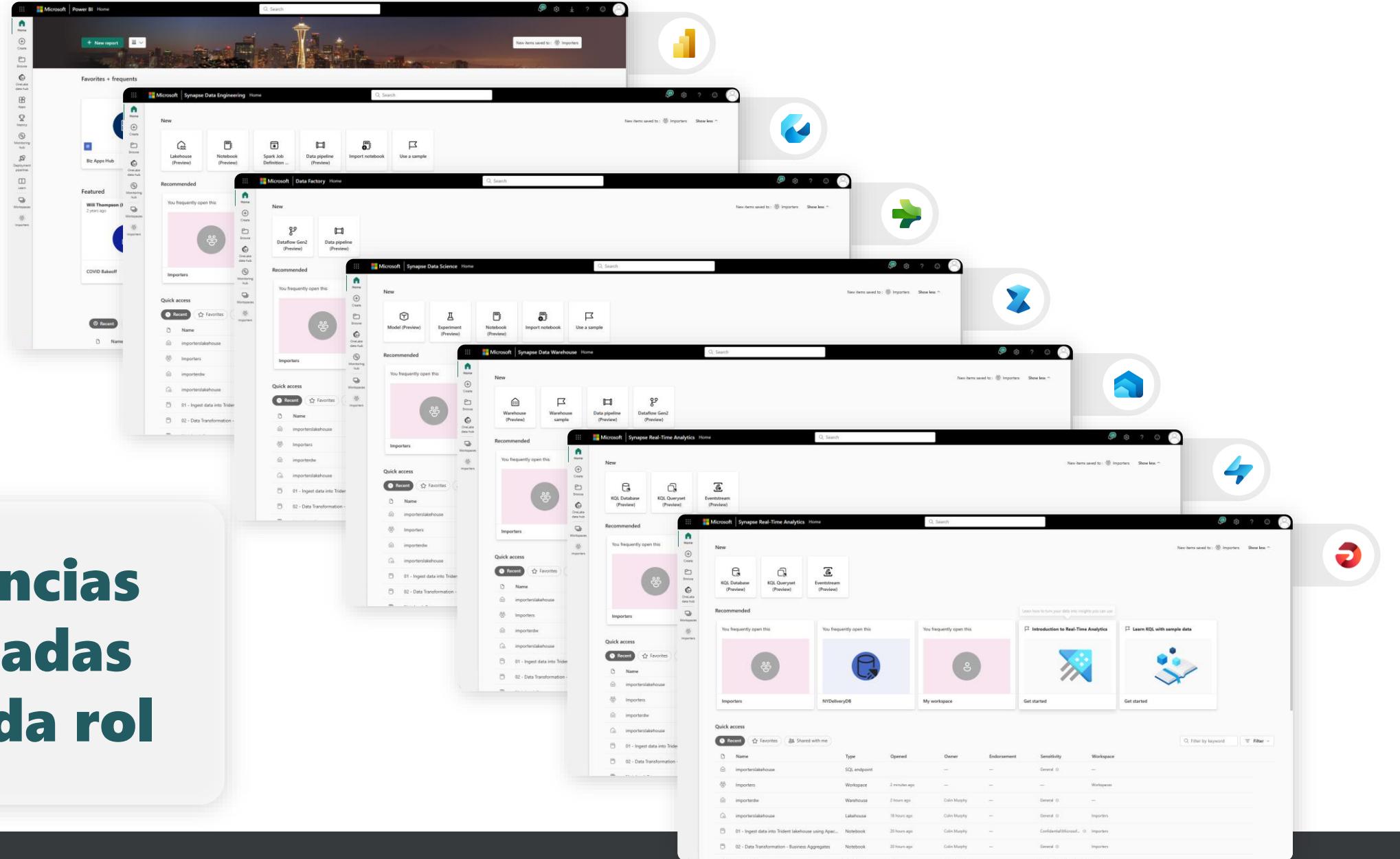
#GPPB2024



#PowerPlatformMadrid



Experiencias optimizadas para cada rol



#GPPB2024



#PowerPlatformMadrid

Siete experiencias clave basadas en Roles y Personas

Las experiencias están diseñadas para dirigirse a personas y tareas específicas, pero funcionan juntas sin problemas en una plataforma unificada a través de OneLake para permitir que los usuarios hagan su mejor trabajo de forma colaborativa:



Combina la facilidad de uso de Power Query con la escala y la potencia de Azure Data Factory para aprovechar 200+ conectores nativos a orígenes de datos locales y en la nube.



Plataforma Spark de clase mundial con excelentes experiencias de creación para capacitar a los ingenieros de datos para transformar los datos a escala



Proporciona un rendimiento y escalabilidad de SQL líderes en la industria, separando completamente la computación del almacenamiento para escalar de forma independiente y almacenar datos de forma nativa en Delta Lake abierto



Cree, implemente y ponga en funcionamiento modelos de aprendizaje automático directamente en Fabric para capacitar a los científicos y analistas de datos con información predictiva



El mejor motor de su clase para el análisis de datos observacionales para crear información procesable a partir de datos en tiempo real



La plataforma de inteligencia empresarial líder en el mundo permite a los usuarios tomar mejores decisiones con los datos de forma rápida e intuitiva



Impulse automáticamente acciones en sus datos, sin necesidad de escribir código



#GPPB2024



#PowerPlatformMadrid



Overview de OneLake



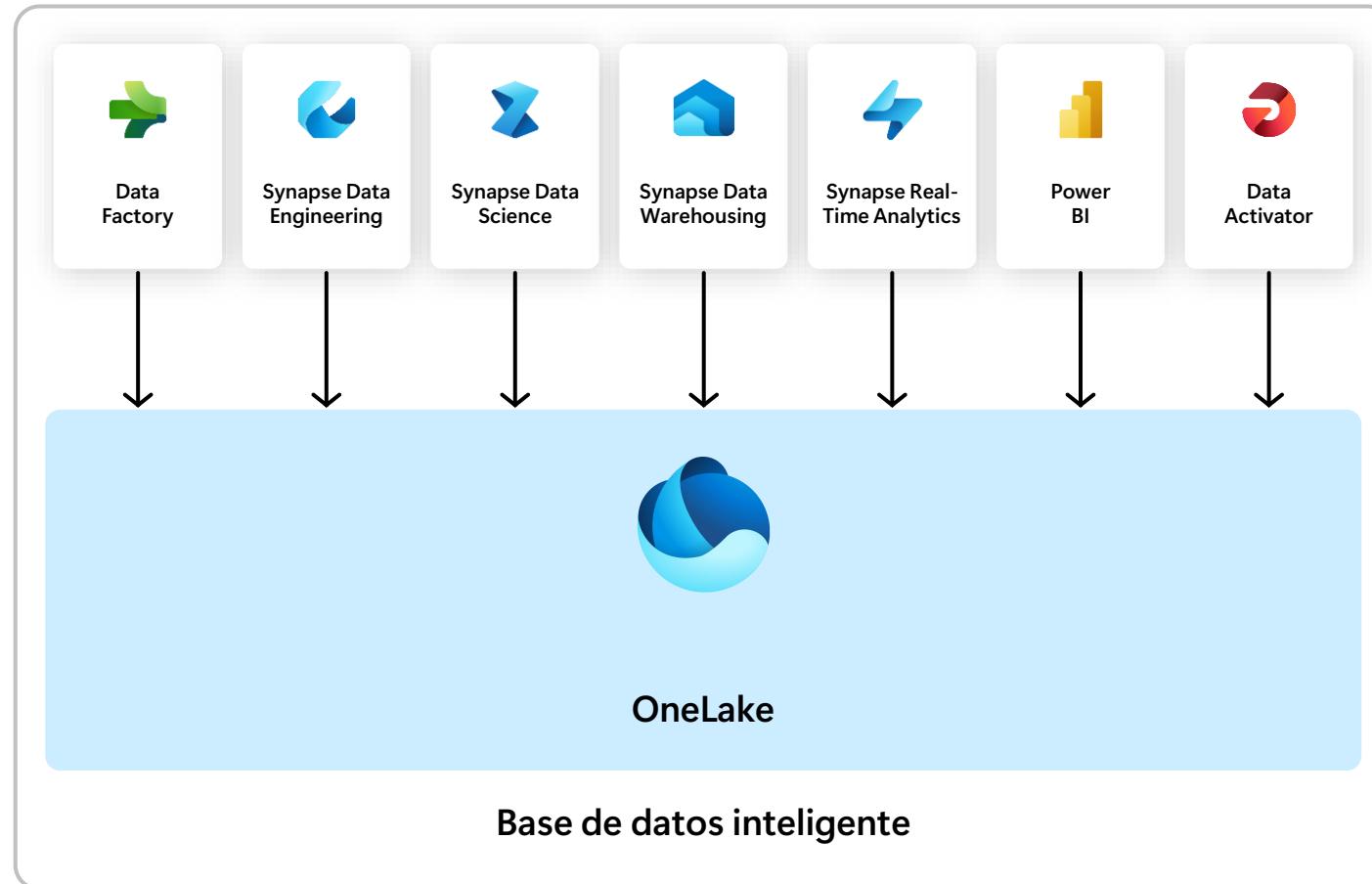
#GPPB2024



#PowerPlatformMadrid

OneLake para todos los datos

"El OneDrive para los datos"



Un únicodatalake SaaS para toda la organización

Aprovisionamiento automático con el tenant

Todas las cargas de trabajo almacenan automáticamente sus datos en las carpetas del área de trabajo de OneLake

Todos los datos están organizados en un espacio de nombres jerárquico intuitivo

Los datos de OneLake se indexan automáticamente para la detección, las etiquetas MIP, el linaje, los análisis de PII, el uso compartido, la gobernanza y el cumplimiento



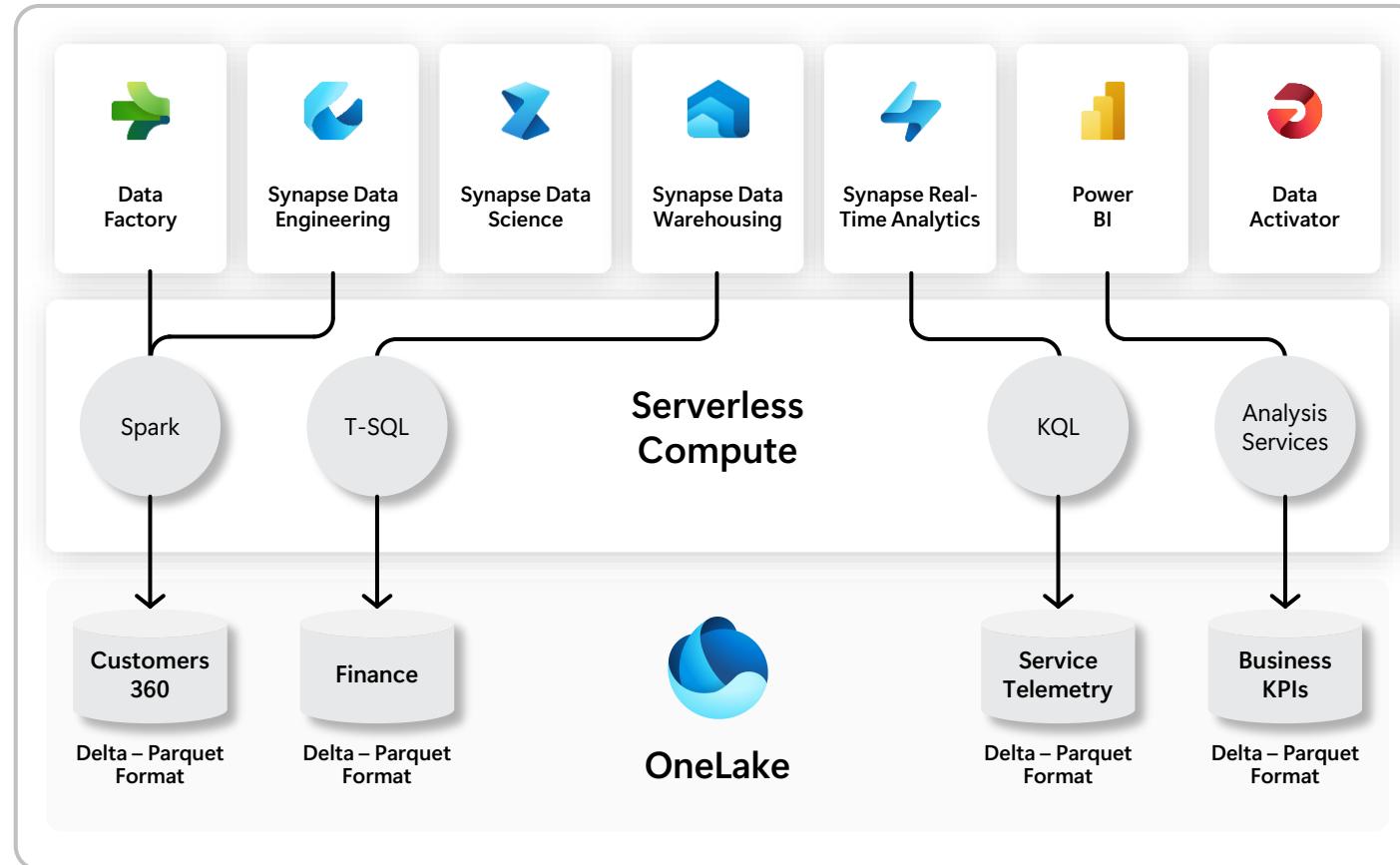
#GPPB2024



#PowerPlatformMadrid

Una copia única para todo

Separación real del cómputo y el almacenamiento



Todos los motores de cómputo almacenan sus datos automáticamente en OneLake

Los datos se almacenan en un único formato común

Delta – Parquet, un formato de estándares abiertos, es el formato de almacenamiento de todos los datos tabulares de Microsoft Fabric

Una vez que los datos se almacenan en el lago, se pueden acceder directamente a ellos sin necesidad de cualquier importación / exportación

Todos los motores de cómputo se han optimizado por completo para que funcionen con Delta Parquet como su formato nativo

El modelo de seguridad universal compartido se aplica en todos los motores de cómputo



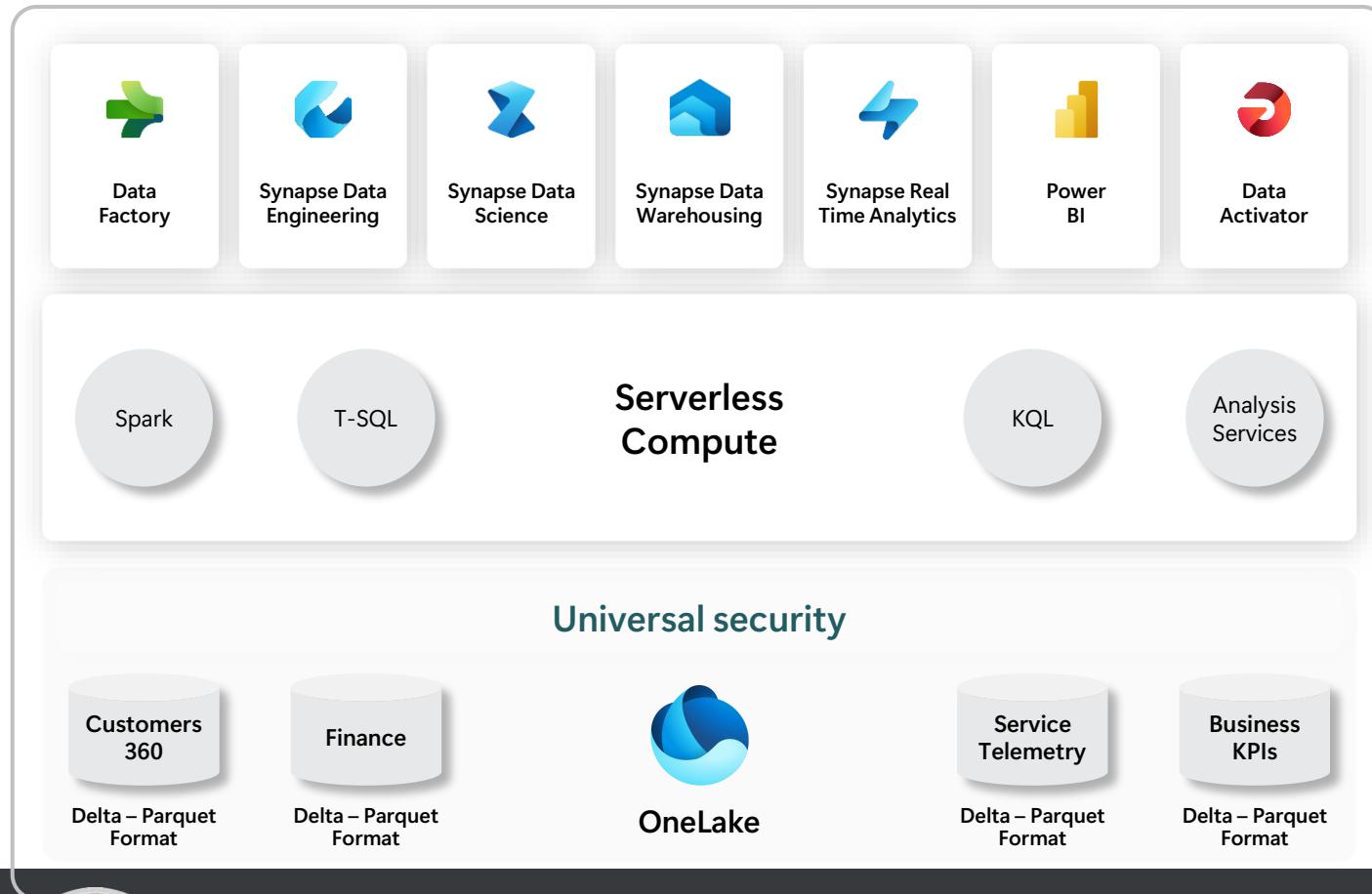
#GPPB2024



#PowerPlatformMadrid

Una copia única para todo

Separación real del cómputo y el almacenamiento



Todos los motores de cómputo almacenan sus datos automáticamente en OneLake

Los datos se almacenan en un único formato común

Delta – Parquet, un formato de estándares abiertos, es el formato de almacenamiento de todos los datos tabulares de Microsoft Fabric

Una vez que los datos se almacenan en el lago, se pueden acceder directamente a ellos sin necesidad de cualquier importación / exportación

Todos los motores de cómputo se han optimizado por completo para que funcionen con Delta Parquet como su formato nativo

El modelo de seguridad universal compartido se aplica en todos los motores de cómputo



#GPPB2024



#PowerPlatformMadrid



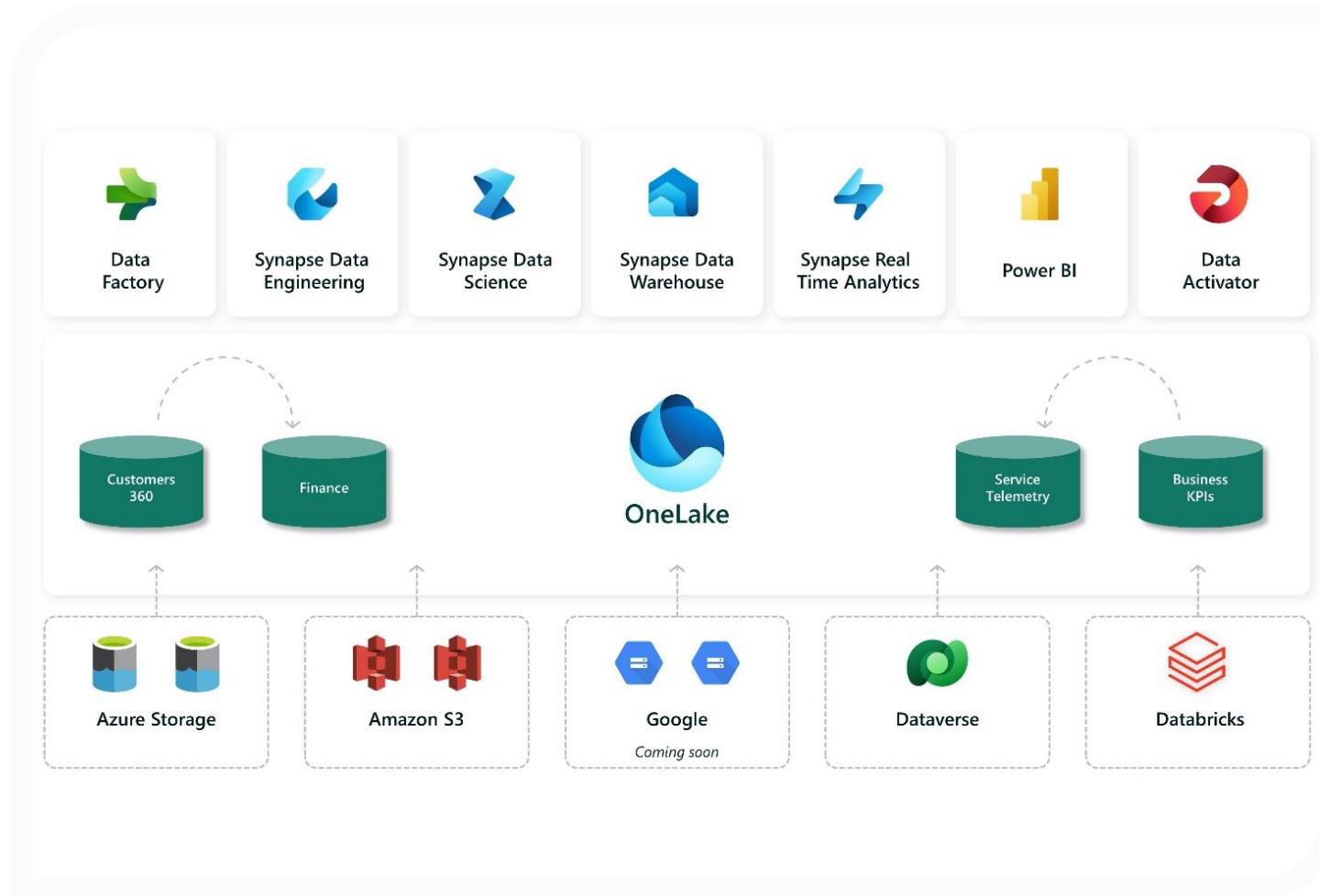
Shortcuts

Los **Shortcuts** unifican los datos sin copiar ni mover los datos existentes.

Esto significa que los datos se pueden utilizar varias veces sin duplicación de datos.

Capacidades clave:

- Cree accesos directos dentro de Microsoft Fabric para consolidar datos en artefactos o áreas de trabajo, sin cambiar la propiedad de los datos
- Los accesos directos también permiten la vinculación instantánea de datos ya existentes en Azure y en otras nubes, sin duplicación ni movimiento de datos, lo que convierte a OneLake en el primer lago de datos multinube
- Con soporte para APIs estándar de la industria, cualquier aplicación o servicio puede acceder directamente a los datos de OneLake



#GPPB2024



#PowerPlatformMadrid

OneLake para todos los dominios

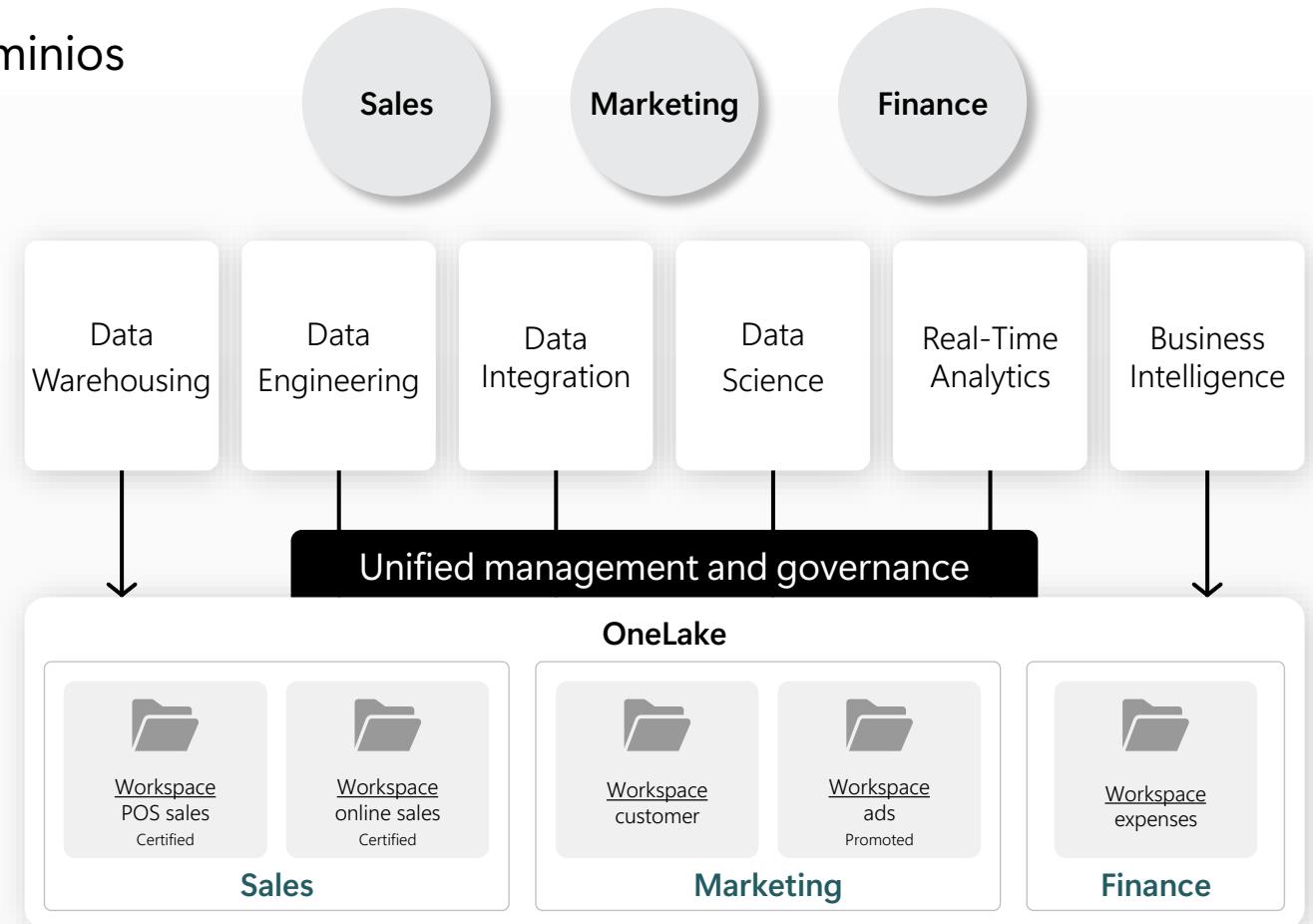
Una verdadera malla de datos en todos los dominios de la organización

Un dominio es una forma de **agrupar** lógicamente todos los datos de una organización relevantes para un área o campo, de acuerdo con las necesidades del negocio

Los dominios se definen con **administradores** y **colaboradores** que pueden asociar espacios de trabajo y agruparlos en un dominio relevante

La **gobernanza federada** se logra delegando la configuración a los administradores de dominio, otorgando un **control más granular** sobre su área de negocio

Evite los pantanos de datos respaldando ciertos datos como **certificados** o **promocionados**, fomentando así la **reutilización**



#GPPB2024



#PowerPlatformMadrid

OneLake data hub

Descubra, administre y reutilice datos en un solo lugar

The screenshot shows the Power BI OneLake data hub interface. On the left is a navigation sidebar with icons for Home, Create, Browse, OneLake data hub, App, Metrics, Monitoring hub, Deployment pipelines, Learn, Workspaces, and Connect CRM. The main area has a search bar at the top. Below it, a banner says "Discover, manage, and use data from across your org." and "Learn more about OneLake data hub". A "Fabric Trial: 59 days left" message is also present. The "Sales" category is selected. A "Recommended" section displays five cards: "Business Dataset" (Owner: Arun Ulag), "US Retail Sales" (Owner: hd-admin), "Contoso Dataset" (Owner: Arun Ulag), "Supplier-Quality-Analysis-Sample..." (Owner: Arun Ulag), and "Operating Report" (Owner: Arun Ulag). At the bottom, there's an "Explorer" section with a search bar and a 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	hd-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/Internal

Ubicación central dentro de Fabric para detectar, administrar y reutilizar datos. Crear accesos directos, Obtener datos en Dataflow, adjuntar un bloc de notas, crear un modelo semántico y mucho más

Los datos pueden ser descubiertos fácilmente por su dominio (ej. Finanzas) para que los usuarios puedan ver lo que les importa

Detección eficiente de datos mediante la búsqueda, el filtrado y orden, o la exploración por jerarquía de carpetas (espacio de trabajo)

Las acciones de datos, como la vista previa, la exploración, el seguimiento del linaje, la creación de informes, el análisis en Excel y la exportación, se pueden realizar fácilmente.

Disponible en Microsoft Teams, lo que lleva el centro de datos de OneLake a Office, lo que permite a los usuarios técnicos y no técnicos la detección, la reutilización y la exploración



#GPPB2024



#PowerPlatformMadrid

Pero que es un Delta Lake?

Delta Lake es una capa de almacenamiento open source que agrega semántica de base de datos relacional al procesamiento de un datalake basado en Spark.

The image displays two side-by-side screenshots of the Microsoft Fabric Data Explorer interface, illustrating the state of a table named "salesorders".

Screenshot 1 (Left): Shows the "salesorders" table in its relational form. The table has columns: SalesOrderID, SalesOrderStatus, OrderDate, CustomerName, Email, Item, and Quantity. The data shows 12 rows of purchase details, such as item descriptions like "Patch Kit/8" and "Half-Finger ...".

	SalesOrderID	SalesOrderStatus	OrderDate	CustomerName	Email	Item	Quantity
1	S051555	7	6/23/2021 ...	Chloe Garcia	chloe27@...	Patch Kit/8 ...	1
2	S054042	7	8/9/2021 1...	Logan Collins	logan29@...	Half-Finger ...	1
3	S054784	7	8/22/2021 ...	Autumn Li	autumn3@...	All-Purpose...	1
4	S058572	7	10/25/2021...	Cesar Sara	cesar9@ad...	Short-Sleev...	1
5	S058845	7	10/30/2021...	Peter She	peter8@ad...	Sport-100 ...	1
6	S058845	8	10/30/2021...	Peter She	peter8@ad...	Long-Sleev...	1
7	S060233	7	11/16/2021...	Jason Mitch...	jason40@...	Sport-100 ...	1
8	S061412	7	12/3/2021 ...	Nathaniel C...	nathaniel9@...	Short-Sleev...	1
9	S062984	7	12/29/2021...	Miguel San...	miguel72@...	Racing Soc...	1
10	S051555	6	6/23/2021 ...	Chloe Garcia	chloe27@...	Mountain B...	1
11	S052058	6	7/4/2021 1...	Elijah Ross	elijah7@ad...	Short-Sleev...	1
12	S052255	6	7/28/2021 ...	Edward Tard	edward21@...	Short-Sleev...	1

Screenshot 2 (Right): Shows the same "salesorders" table in its Delta Lake file-based view. The table is now represented as a folder structure within a "Files" folder. It contains four Parquet files (part-00000 through part-00002) and a log file (_delta_log). The log file was last modified on 5/5/2023 at 11:00:08 AM.

Name	Date modified	Type	Size
_delta_log	5/5/2023 11:00:08 AM	Folder	2 items
part-00000-270f72f4-dd...	5/5/2023 11:00:09 AM	PARQUET	413 KB
part-00001-c433570f-28...	5/5/2023 11:00:09 AM	PARQUET	80 KB
part-00002-7c7b8c78-a8...	5/5/2023 11:00:09 AM	PARQUET	37 KB



#GPPB2024



#PowerPlatformMadrid

INSIDE THE APACHE PARQUET FORMAT IT'S OPEN SOURCE!

IN TRADITIONAL DATABASES (POSTGRES, SQL SERVER*) AND FILE TYPES (CSV, JSON), DATA IS STORED AS ROWS...

FIRST_NAME	LAST_NAME	STREET	ISPREMIUMC USTOMER	NEXT MONTH BUYING PROBABILITY
ABIGAIL	HAYES	4009 9TH STREET	1	1
AMANDA	CLARK	87 5TH STREET	0	2
MELISSA	CLARK	12650 5TH STREET NORTH	1	0.5438465585898219
SARA	BENNETT	203 STATE STREET	0	0.15856312140709694

Customers.csv

```
first_name,last_name,street,isPremiumCustomer,nextMonthBuyProb
Abigail,Hayes,4009 9th Street,1,0.26340968441351104
Amanda,Clark,87 5th Street,0,0.17177893540625921
Melissa,Clark,12650 5th Street North,1,0.5438465585898219
Sara,Bennett,203 State Street,0,0.15856312140709694
Isaac,Henderson,732 Magnolia Drive,1,0.19389081059172666
Emma,Stewart,47 Route 32,1,0.8833057932776623
Sean,Morris,254 Lincoln Avenue,0,0.415633382834478
Jason,Sanchez,26050 Madison Avenue,0,0.5862717219123177
Alexander,Walker,54051 Valley Road,1,0.5941814808482648
Megan,Morris,31 Delaware Avenue,1,0.6289626353900487
```

... WHICH IS NOT OPTIMAL FOR ANALYTICAL NEEDS:

- ☒ FOR SOME FILE TYPES (LIKE CSV), THERE IS NO EMBEDDED SCHEMA
- ☒ COMPRESSION IS NOT GOOD (YOU'RE KINDA LIMITED TO TEXT COMPRESSION)
- ☒ YOU NEED TO SCAN(READ) ALL THE ROWS

*UNLESS YOU'RE USING COLUMNSTORE TABLES :)

IN PARQUET, DATA IS STORED AS COLUMNS

Customers.parquet

```
first_name Abigail Amanda Melissa Sara Isaac Emma Sean Jason
last_name Hayes Clark Clark Bennett Henderson Stewart Morris Sanchez
isPremiumCustomer 1 0 1 0 1 x2 0 x2
Category 0.263409684 0.171778935 0.543846559 0.158563121 0.193890811
0.883305793 0.415633338 0.586271722
Col 1 STRING Col 2 STRING Col 3 BOOLEAN Col 4 UINT_32
```

FILE STRUCTURE IS SIMPLIFIED, OMITTING DETAILS LIKE ROW GROUPS.



PARQUET EMBEDS THE DATA SCHEMA

WANT THE NUMBER OF PREMIUM CUSTOMERS? YOU CAN COMPUTE THAT AGGREGATE WITHOUT READING ANY OTHER UNNECESSARY DATA!



WE CAN EFFICIENTLY COMPRESS REPETITIVE DATA LIKE THIS ONE

WITH FABRIC, YOU CAN EASILY READ PARQUET FILES STORED IN ONE LAKE OR IN OTHER CLOUDS

DENTRO DEL FORMATO DELTA PARQUET

EL FORMATO DELTA AMPLÍA LA FUNCIONALIDAD DE PARQUET CON CARACTERÍSTICAS COMO ELIMINACIONES, ACTUALIZACIONES Y FUSIONES, TRANSACCIONES ACID, CONTROL DE VERSIONES Y VIAJES EN EL TIEMPO, EVOLUCIÓN DE ESQUEMAS, OPTIMIZACIONES DE LECTURA Y MÁS.



UNA TABLA DELTA ES SOLO UNA
CARPETA

EL REGISTRO DE TRANSACCIONES HABILITA LAS
CARACTERÍSTICAS DELTA.
CADA ARCHIVO REPRESENTA UNA CONFIRMACIÓN Y
CONTIENE PUNTEROS A ARCHIVOS PARQUET Y
METADATOS

Líneas añadidas. Contenido en part_0001.parquet
Líneas modificadas. Contenido en part_0002.parquet

LOS ARCHIVOS PARQUET SE PUEDEN
ALMACENAR EN LA CARPETA RAÍZ U
ORGANIZAR POR PARTICIONES

Lab 1



Cómo crear un espacio de trabajo de Fabric



Cómo crear un Lakehouse

The screenshot shows the Microsoft Fabric Home interface. At the top right, it says "Microsoft Fabric" and "All your data. In one location." Below that, there's a heading "Explore the Fabric ecosystem". The main area displays four service cards:

- Power BI**: Find insights, track progress, and make decisions faster using rich visualizations.
- Data Factory**: Solve the most complex data integration and ETL scenarios with cloud-scale data movement and data transformation services.
- Synapse Data Science**: Explore your data, and build machine learning models to infuse predictive insights into your analytics solutions and applications.
- Synapse Data Warehouse**: Scale up your insights by storing and analyzing data in a secure, open-data-format SQL warehouse with top performance at PB scale.



#GPPB2024



#PowerPlatformMadrid

Data Integration



#GPPB2024



#PowerPlatformMadrid

Data Factory



Data Factory en Microsoft Fabric proporciona servicios de transformación de datos y movimiento de datos a escala de nube que le permiten resolver los escenarios de ETL más complejos.

El núcleo de Data Factory son las canalizaciones de datos y los flujos de datos para ofrecer a los usuarios la opción de un enfoque de bajo código, colaborativo y a escala empresarial para su proceso ETL.



Nuevas capacidades:

- Destino de salida a Lakehouse
- 14 nuevos conectores disponibles en la actividad de copia de canalización
- Conector de almacén para Power BI Desktop
- Asistencia de copia de Pipeline Lakehouse
- Plantillas de canalización
- Compatibilidad con canalizaciones para cuadernos de Spark
- Soporte de autenticación Service Principals



#GPPB2024



El ETL autónomo puede desbloquear eficiencias operativas y ayudar a orquestar, supervisar y gestionar el rendimiento de los pipelines.

The screenshot shows the Microsoft Data Factory interface. At the top, there's a navigation bar with tabs like Home, Activities, Run, View, and pipeline5. 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 OneLake data hub, Monitoring hub, Workspaces, Importers, and pipeline5 (which is selected). The main area has a heading "Start building your data pipeline". It features three green buttons: "Add pipeline activity" (with a gear icon), "Copy data" (with a copy icon), and "Choose a task to start" (with a square icon). A small "Data Factory" logo is at the bottom left of the main area.

#PowerPlatformMadrid

Data Pipelines



Data Pipelines habilita potentes capacidades de flujo de trabajo a escala de la nube, como la creación de flujos de trabajo complejos, el traslado de datos de tamaño PB y la definición de canalizaciones de flujo de control sofisticadas.

Las canalizaciones de datos se pueden usar para crear flujos de trabajo complejos de ETL y factoría de datos que pueden realizar una serie de tareas diferentes a escala.

Además, las capacidades de flujo de control están integradas en las canalizaciones para que pueda crear una lógica de flujo de trabajo que proporcione bucles y condicionales.

The screenshot shows the Microsoft Data Factory interface. A modal window titled "Copy data" is open, specifically the "Choose data source" step. The left side of the modal has four options: "Connect to data source", "Choose data destination", "Connect to data destination", and "Review + save". The right side displays a grid of data sources categorized under "All categories". The categories include Workspace, Azure, Database, File, Generic protocol, and Services and apps. Examples of data sources listed include 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, Azure Data Explorer (Kusto), Azure SQL Database, Data Warehouse, Hive, Microsoft 365, SharePoint Online List, and many others.

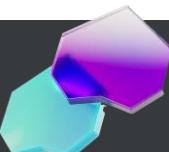


#GPPB2024

Microsoft Fabric



#PowerPlatformMadrid



Data Pipelines | Conectores



Los Nuevos Conectores proporcionan una interfaz con poco código para ingerir datos de una variedad de orígenes de datos.

Nuevos conectores:

- Conector de almacén; conexión a Azure Synapse existente
 - Conector Lakehouse
 - 14 nuevos conectores en la actividad de copia: 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 y Azure Data Explorer

The screenshot shows the 'Copy data' wizard in the Azure Data Factory portal. The left sidebar lists various pipeline components like Home, Activities, Run, View, and pipeline6. The main area has a title 'Copy data' and a sub-section 'Choose data source'. It includes a search bar and a note: 'Build your data ingestion task to move objects from a data source to a data destination. Learn more'. Below this is a list of steps: 'Connect to data source', 'Choose data destination', 'Connect to data destination', and 'Review + save'. A preview pane shows 'holidays package and Wikipedia, covering 38 countries or regions from 1970 to 2099'. The 'Data sources' section is open, displaying a grid of connectors categorized by type. The categories are All categories, Workspace, Azure, Database, File, Generic protocol, and Services and apps. The 'All categories' tab is selected. The grid contains 24 items:

All categories	Workspace	Azure	Database	File	Generic protocol	Services and apps
Amazon RDS for SQL Server	Amazon Redshift	Amazon S3	Amazon S3 Compatible			
Apache Impala	Azure Blob Storage	Azure Cosmos DB for NoSQL	Azure Data Explorer (Kusto)			
Azure Data Lake Storage Gen1	Azure Data Lake Storage Gen2	Azure Database for PostgreSQL	Azure SQL Database			
Azure SQL Database Managed Instance	Azure Synapse Analytics	Azure Table Storage	Data Warehouse			
Dataverse	Dynamics CRM	Google Cloud Storage	Hive			
HTTP	KQL Database	Lakehouse	Microsoft 365			
OData	PostgreSQL	REST	SharePoint Online List			
Snowflake	Spark	SQL server				

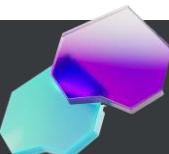
At the bottom are 'Back' and 'Next' buttons, and a 'Cancel' button on the right.



#GPPB2024



#PowerPlatformMadrid



Data Pipelines | Datos de ejemplo

Los Ejemplos de modelos semánticos ayudan a los nuevos usuarios a comenzar rápidamente, creando sus procesos ELT mediante canalizaciones de datos.



Ejemplos de modelos semánticos:

- COVID-19 Data Lake (CSV, JSON, JSON Lines, Parquet)
- Impuestos de la ciudad de Nueva York - Verde (parquet de 2 GB)
- Diabetes (14K Parquet)
- Días festivos (parquet de 500 KB)
- Modelo de datos de venta al por menor de Wide World Importers (parquet de 352 MB)

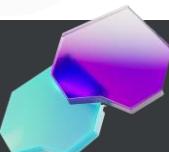
The screenshot shows the Microsoft Power BI Data Factory interface. A modal window titled 'Copy data' is open, guiding the user through the process of moving objects from a data source to a destination. The first step, 'Choose data source', is selected. Below it, there are four options: 'Connect to data source', 'Choose data destination', 'Connect to data destination', and 'Review + save'. To the right of the steps, a section titled 'Sample data' lists several datasets: COVID-19 Data Lake (2 GB, Parquet), NYC Taxi - Green (2 GB, Parquet), Diabetes (14 KB, Parquet), Public Holidays (500 KB, Parquet), and Retail Data Model from Wide World Importers (352 MB, Parquet). Below this, a 'Data sources' section displays a grid of various Azure and AWS services, each with a small icon and a brief description. The categories listed are All categories, Workspace, Azure, Database, File, Generic protocol, Services and apps. Some visible service icons include Amazon RDS for SQL Server, Amazon Redshift, Amazon S3, Apache Impala, Azure Blob Storage, Azure Data Lake Storage Gen1, Azure Data Lake Storage Gen2, Azure SQL Database Managed Instance, Azure Synapse Analytics, Azure Cosmos DB for NoSQL, Azure Data Explorer (Kusto), Azure Database for PostgreSQL, Azure SQL Database, Azure Table Storage, and Data Warehouse.



#GPPB2024



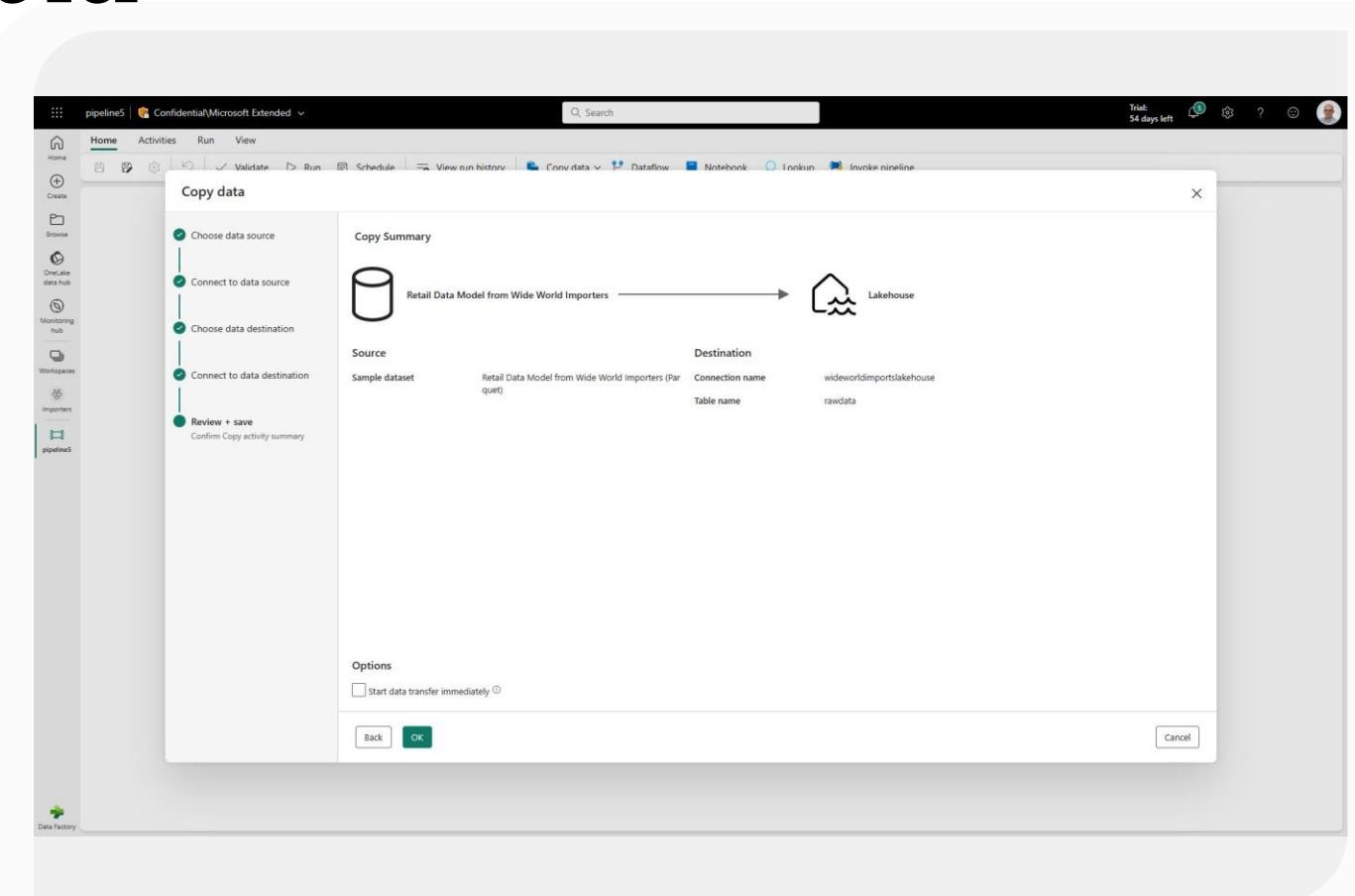
#PowerPlatformMadrid



Data Pipelines | Asistente de copia de Lakehouse

Simplemente copie los datos en un Lakehouse con capacidades de asistencia de copia dentro de la canalización de datos.

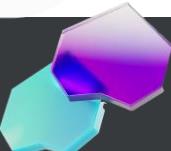
Además, el usuario puede crear un archivo de datos Pipeline sin tener que salir del portal de Lakehouse.



#GPPB2024



#PowerPlatformMadrid



Data Pipelines | Templates

Las plantillas ayudan a reducir el tiempo de desarrollo al proporcionar una manera fácil de crear Pipelines en escenarios comunes de integración de datos.



Plantillas de Pipelines disponibles:

- Copia masiva desde bases de datos y archivos
- Copia de datos de ADLS Gen2 al Lakehouse
- Copiar de ADLS Gen2 a tabla de Lakehouse
- Copia de Azure AQL DB a tabla de Lakehouse
- Copiar varios contenedores de archivos entre almacenes de archivos
- Copiar archivos nuevos solo por fecha de última modificación
- Eliminar archivos con más de 30 días de antigüedad
- Copia delta de la base de datos
- Mover archivos

The screenshot shows the Microsoft Data Factory pipeline template gallery. The interface includes a top navigation bar with 'pipeline4' and 'Confidential/Microsoft Extended'. Below the navigation is a toolbar with 'Home', 'Activities', 'Run', 'View', 'Validate', 'Run', 'Schedule', 'View run history', and 'Copy data'. On the left, a sidebar lists 'Home', 'Create', 'Browse', 'OneLake data hub', 'Monitoring hub', 'Workspaces', 'Importers', and the current 'pipeline4'. The main area is titled 'Templates' with a search bar 'Search'. It displays a grid of 12 templates, each with an icon, title, and brief description. The templates are:

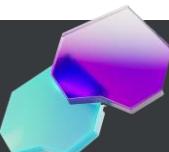
- Bulk Copy from Database** by Microsoft: Use this template to copy data in bulk from a database using an external control table to store the partition list of your source tables...
- Bulk Copy from Files to Database** by Microsoft: Use this template to copy data in bulk from Azure Data Lake Storage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



#GPPB2024



#PowerPlatformMadrid



Dataflow



Dataflow proporciona una interfaz low-code para ingerir datos de cientos de orígenes de datos.

Dataflow unifica rápida y fácilmente fuentes de datos dispares, establece un enfoque de análisis más colaborativo y promueve una toma de decisiones más informada y ágil.



Capacidades clave:

- Acelere la transformación de datos con flujos de datos sin código
- Estandarizar los flujos de datos en la compatibilidad con Power Apps para cargar datos
- Cargue los resultados de las transformaciones de datos en varios destinos (Azure SQL Databases, Lakehouse, etc.)

The screenshot shows the Microsoft Data Factory interface with a search bar at the top. Below it, a large grid titled 'Choose data source' displays numerous data connectors. The connectors are organized into categories: All categories, File, Database, Power Platform, Azure, Online services, and Other. Each category contains several icons representing different data sources, such as Excel workbook, Parquet, SharePoint folder, MySQL database, PostgreSQL database, Access Database, SAP HANA database, and many others. Some icons have a 'BETA' label. At the bottom right of the grid, there is a 'Cancel' button.

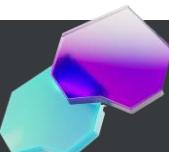


#GPPB2024



LA NAVE

#PowerPlatformMadrid





Dataflow | Gen1 vs Gen2

Dataflow Gen2 es la nueva generación de flujos de datos. La nueva generación de flujos de datos reside junto con el flujo de datos de Power BI (Gen1) y aporta nuevas características y experiencias mejoradas.

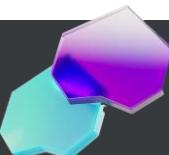
Feature	Dataflow Gen2	Dataflow Gen1
Author dataflows with Power Query	✓	✓
Shorter authoring flow	✓	
Auto-Save and background publishing	✓	
Data destinations	✓	
Improved monitoring and refresh history	✓	
Integration with data pipelines	✓	
High-scale compute	✓	
Get Data via Dataflows connector	✓	✓
Direct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



#GPPB2024



#PowerPlatformMadrid



Dataflow | Salida al Lakehouse



Simplemente se puede escribir en un Lakehouse desde un flujo de datos.

Los usuarios seleccionan el destino de salida de Lakehouse de la lista y configuran la conexión.

Esto requiere que el conector de Lakehouse se instale como un conector personalizado en la puerta de enlace de datos al cargar datos on-premises.

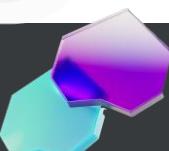
The screenshot shows the Microsoft Synapse Data Engineering interface. A Dataflow job named "Dataflow 2" is running. In the "Data destination" section, "Azure SQL database" is selected. The main pane displays a preview of data from a "dimension_customer" table, showing columns like CustomerKey, BirthDateCustomer, Category, BuyingGroup, PrimaryContact, PostCode, and Valueto. The bottom right corner shows a "Publish" button.



#GPPB2024



#PowerPlatformMadrid



Lab 2



Cómo crear Dataflow Gen2



Cómo conectarse a ADLS Gen2 con Dataflow Gen2 y transformar datos



Cómo ingerir datos en Lakehouse

The screenshot shows the Microsoft Dataflow Gen2 interface. On the left is a data preview grid with three columns: Unit Price, Tax Rate, and Tax Amount. The data consists of 15 rows of numerical values. To the right of the preview is the 'Query settings' pane. It includes sections for 'Properties' (Name: Sales, Entity type: Custom), 'Applied steps' (listing Source, Expanded Invoice, Removed Other Columns, Renamed Columns, Merged Queries, Added Custom, Changed Type, and Removed Columns, with Renamed Columns1 highlighted in red), and 'Data destination' (Lakehouse). At the bottom right of the pane is a green 'Publish' button, which is outlined in orange.



#GPPB2024

Lab 3



Cómo configurar una actualización programada de Dataflow Gen2



Cómo crear un Data Pipeline



Cómo configurar una actualización programada de un Data Pipeline

The screenshot shows the configuration interface for a data pipeline named "pl_Refresh_People_SharePoint_Option2". The "Schedule" tab is selected. A large orange rectangle highlights the "Scheduled run" section, which includes a radio button for "On" (selected) and "Off", a dropdown for "Repeat" set to "Daily", and a "Time" field set to "09:00 AM". Below this are fields for "Start date and time" (11/05/2023) and "End date and time" (02/29/2024), and a "Time zone" dropdown set to "(UTC-06:00) Central Time (US and Canada)". At the bottom are "Apply" and "Discard" buttons.



#GPPB2024

Break



#GPPB2024



#PowerPlatformMadrid

Data Engineering



#GPPB2024



#PowerPlatformMadrid

Lakehouse

Almacene, administre y analice todos sus datos en una sola ubicación y compártalos fácilmente en toda la empresa.

Cree rápida y fácilmente un Lakehouse sin tener que aprovisionar y configurar la computación, el almacenamiento y las redes.



Capacidades clave:

- Solución flexible y escalable que permite a las organizaciones manejar grandes volúmenes de datos de todos los tipos y tamaños
- El punto de conexión SQL integrado desbloquea las capacidades de almacenamiento de datos en la parte superior de su Lakehouse sin movimiento de datos
- Utilice el modo "lago directo" para crear informes en segundos directamente sobre el lago de datos con un rendimiento ultrarrápido
- Ingiera fácilmente datos en Lakehouse a través de una variedad de métodos
- Comparta su Lakehouse como un producto de datos con los consumidores

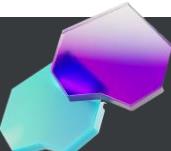
Name	Date modified	Type	Size
_SUCCESS	4/24/2023 6:58:06 PM	-	0 B
part-00000-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:09 PM	PARQUET	22 MB
part-00001-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	4/24/2023 6:58:09 PM	PARQUET	26 MB
part-00002-ced648ca-e8c8-46e5-8526-5ca85d56e67e-c000.snappy.parquet	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



#GPPB2024



#PowerPlatformMadrid



Notebooks



Experiencia de creación inmersiva para desarrolladores de datos.

Capacidades de bloc de notas enriquecidas, incluida la integración nativa de Lakehouse, la colaboración en tiempo real con soporte de comentarios, el soporte de guardado automático, la programación ligera y la integración de canalizaciones.



Capacidades clave:

- Administre sus bibliotecas de Python y R a través de instalaciones en línea mediante comandos como %pip install
- Soporte avanzado de desarrollo de blocs de notas con capacidad para hacer referencia a blocs de notas en blocs de notas e instantáneas para facilitar la resolución de problemas
- Monitoreo en contexto completo con asesoramiento en tiempo real y análisis de errores
- Optimice la preparación de datos sin renunciar al poder de reproducibilidad de Python

The screenshot shows the Microsoft Power BI Notebook interface. On the left, there's a sidebar with navigation links like Home, Create, Browse, OneLake data hub, Monitoring hub, Workspaces, Importers, and Data Science. Below that is a detailed file tree under '02 - Explore and...'. The main area has a title 'Visual 1: Distribution of trip duration(minutes) on linear and logarithmic scale'. It contains two plots: one on a linear scale from 0 to 1400 minutes with a count up to 3000, and another on a logarithmic scale from 1 to 1000 minutes with a count up to 3000. Above the plots is a code block in PySpark (Python) that generates the histograms using pandas and sns.histplot. The code includes comments explaining the steps, such as computing trip duration in minutes and setting different parameters for each plot.

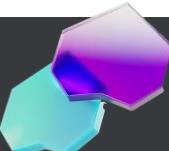
```
## Compute trip duration(in minutes) on the sample using pandas
1 sampled_df['tripDuration'] = (sampled_df['tpepDropoffDateTime'] - sampled_df['tpepPickupDateTime']).astype('timedelta64[m]')
2 sampled_df = sampled_df[sampled_df['tripDuration'] > 0]
3
4 fig, axes = plt.subplots(1, 2, figsize=(18,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



#GPPB2024



#PowerPlatformMadrid



Creación de tablas delta

Al crear una tabla en un lakehouse de Microsoft Fabric, se define una tabla delta en el metastore del lakehouse y los datos de la tabla se almacenan en los archivos Parquet subyacentes de la tabla.



#GPPB2024



#PowerPlatformMadrid

TABLAS: MANAGED VS EXTERNAL

MANAGED TABLES

EXTERNAL TABLES

AMBOS ALMACENAN DATOS EN EL FORMATO OPEN DELTA PARQUET

LA ADM DE DATOS ES MANEJADA POR FABRIC SPARK

ADMINISTRADO POR FABRIC SPARK, EN DIRECTORIO *Tables*

LOS DATOS SE BORRAN CUANDO LA TABLE SE
ELIMINA

PERFECTO DENTRO DE FABRIC.
POSIBLE SI LAS TOOLS ACCEDEN AL LAKEHOUSE

USE CASE

LOCATION

DATA RETENTION

INTEGRATION WITH OTHER
TOOLS

USO DE CONJUNTOS DE DATOS EXISTENTES,
ADMINISTRADOS FUERA DE FABRIC

ESPECIFICADO POR USUARIO, EN EL LAKEHOUSE

LOS DATOS SE CONSERVAN CUANDO SE ELIMINA
LA TABLA

GENIAL DENTRO DE FABRIC.
FÁCIL, YA QUE LOS DATOS SE PUEDEN
ALMACENAR

SOPORTADO EN AMBOS CASOS: TRANSACCIÓN ACID, BATCH & STREAMING, MANEJO DE METADATOS, ...

RENDIMIENTO: SIMILAR PARA AMBOS TIPOS DE TABLAS (CUANDO SE CREAN DENTRO DE FABRIC CON LAS OPCIONES ADECUADAS)

CREACIÓN DE TABLAS DELTA CON SPARK

MANAGED TABLE

EXTERNAL TABLE

CREACIÓN DE TABLAS DELTA CON SPARK

MANAGED TABLE

FROM DATA



CREACIÓN DE TABLAS DELTA CON SPARK

MANAGED TABLE

FROM DATA

```
PYTHON  
df.write.format("delta").  
saveAsTable("mytable")
```

SPARK SQL

```
CREATE TABLE PastYearSales USING DELTA  
AS SELECT * FROM SalesOrders WHERE SalesYear = 2023
```

FROM SCHEMA

```
PYTHON  
DeltaTable.create(spark).tableName("products")  
.addColumn("Productid", "INT").execute()
```

SPARK SQL

```
CREATE TABLE salesorders ( Orderid  
INT NOT NULL, OrderDate TIMESTAMP NOT NULL,  
CustomerName STRING, SalesTotal FLOAT NOT  
NULL ) USING DELTA
```

EXTERNAL TABLE

```
PYTHON  
df.write.format("delta").saveAsTable("myext  
ernaltable", path="Files/myexternalTable")
```

SPARK SQL

```
CREATE TABLE PastYearSales USING DELTA  
LOCATION 'Files/PastYearSales'  
AS SELECT * FROM SalesOrders WHERE SalesYear = 2023
```

PYTHON

```
from delta import *  
DeltaTable.create(spark).tableName("myextern  
altable").addColumn("c1", dataType="INT",  
nullable=False).execute()
```

SPARK SQL

```
CREATE TABLE salesorders ( Orderid INT NOT  
NULL, OrderDate TIMESTAMP NOT NULL, CustomerName  
STRING, SalesTotal FLOAT NOT NULL ) USING  
DELTA LOCATION 'Files/sales_orders_ext'
```

Lab 4



Cómo crear Tablas con Spark Delta API



Cómo crear Tablas Delta con Spark SQL



Usar Time Travel con Tablas Delta



#GPPB2024



Python

```
from delta.tables import *
from pyspark.sql.functions import *

# Create a DeltaTable object
delta_path = "Files/mytable"
deltaTable = DeltaTable.forPath(spark,
delta_path)
```

SQL

```
%%sql
```

```
UPDATE products
SET Price = 2.49 WHERE ProductId = 1;
```

#PowerPlatformMadrid

Data Warehousing



#GPPB2024



#PowerPlatformMadrid

Data warehouse



Data Warehouse de escala empresarial con un formato abierto.

Configuración e implementación mínimas sin complejidad, no se necesita configuración de cómputo o almacenamiento.



Capacidades clave:

- Experiencias sencillas e intuitivas tanto para profesionales de datos principiantes como experimentados
- Centrado en Lakehouse almacena datos en OneLake en formato Delta abierto con una fácil recuperación y gestión de datos
- Totalmente integrado con todas las cargas de trabajo de Fabric
- Carga de datos y transformaciones a escala, con garantías transaccionales completas de varias tablas proporcionadas por el motor SQL
- Almacenes virtuales con consultas entre bases de datos y una capa semántica totalmente integrada
- Vsibilidad de uso de extremo a extremo, con gobernanza y seguridad integradas

The screenshot shows the Microsoft Synapse Data Warehouse Home interface. In the center, a modal dialog box titled "New warehouse" is open, prompting the user to enter a name ("importersDW") and sensitivity ("General"). The background displays a grid of workspace cards, including "vNextAnalytics", "Importers" (which is highlighted), "My workspace", "NorthwindTraders", and "Getting started with warehouse". On the left, a sidebar lists various workspace categories like Home, Create, Browse, Data hub, Workspaces, Importers, Importers lakehouse, 02 - Data Transformation, and 01 - Create Delta Tables. At the bottom, a table provides a quick access overview of existing workspaces, including columns for Name, Type, Opened, Owner, Endorsement, Sensitivity, and Workspace.

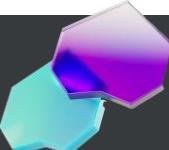
Name	Type	Opened	Owner	Endorsement	Sensitivity	Workspace
Importers	Workspace	now	—	—	—	Workspaces
importerslakehouse	SQL endpoint	5 hours ago	Importers	—	General	Importers
importerslakehouse	Lakehouse	5 hours ago	Colin Murphy	—	General	Importers
02 - Data Transformation - Business Aggregates	Notebook	5 hours ago	Colin Murphy	—	Confidential/Microsoft...	Importers
01 - Create Delta Tables	Notebook	5 hours ago	Colin Murphy	—	Confidential/Microsoft...	Importers
IngestDataFromSourceToLakehouse	Data pipeline	5 hours ago	Colin Murphy	—	Confidential/Microsoft...	Importers
Customer Summary Report	Report	6 hours ago	Importers	—	General	Importers



#GPPB2024



#PowerPlatformMadrid



Capa semántica de Power BI totalmente integrada

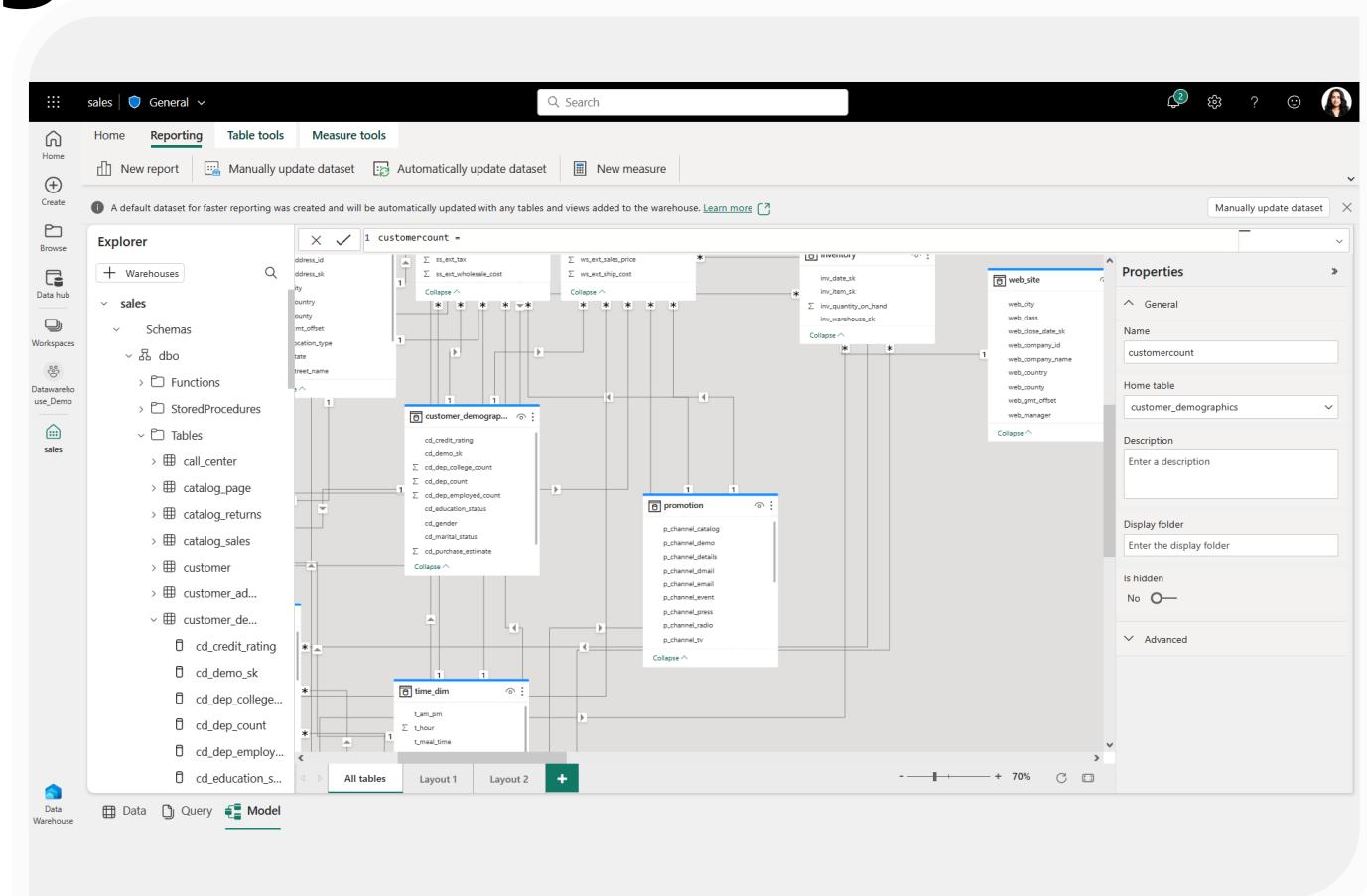
Reduzca la integración y obtenga información de sus datos en segundos.

La integración de Power BI permite a todos los usuarios visualizar sus datos en cuestión de segundos.



Capacidades clave:

- Modelo semántico de Power BI generado automáticamente siempre sincronizado
- Modelo semántico predeterminado en el modo Direct Lake, pero cambia automáticamente al modo Direct Query o Importación a medida que cambian las necesidades de seguridad o rendimiento
- Flexibilidad para agregar/eliminar tablas al modelo semántico
- Experiencia completa en creación web para crear medidas



#GPPB2024

Microsoft Fabric



#PowerPlatformMadrid



Seguro por defecto



Mantenga los datos **seguros** para cualquier rol que acceda a ellos y garantice la tranquilidad.



Capacidades clave:

- **Roles de Workspace :** se utilizan para la colaboración con equipos. Agregar usuarios al espacio de trabajo con la asignación de roles de Administrador, Miembro, Colaborador o Lector
- **Permisos de Artefactos:** se utilizan para compartir para el consumo de los almacenes. Proporcionar acceso y compartir almacenes individuales con permisos específicos (próximamente)
- **Seguridad de Datos :** Use T-SQL, GRANT, REVOKE o DENY para proteger cualquier objeto dentro de Warehouse. Los usuarios se pueden asignar a un rol personalizado integrado
- **Etiquetas de confidencialidad:** Aplique etiquetas de confidencialidad en su almacén para clasificar los datos confidenciales

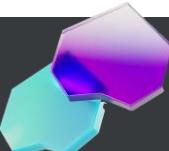
The screenshot shows the Microsoft Power BI Data Studio interface. On the left, there's a sidebar with navigation links like Home, Create, OneLake data hub, Monitoring hub, Workspaces, Importers, and Importerdw. The main area displays the 'importerdw | General' view. It includes sections for 'Details for importerdw' (Location: Importers, Refreshed: —, Sensitivity: General), 'Visualize this data' (with a 'Create a report' button), and 'Share this data' (with a 'Share warehouse' button). Below these are sections for 'See what already exists' and a table listing datasets. The table has columns for Name, Type, Relation, Location, Refreshed, Endorsement, and Sensitivity. One row is shown: 'importerdw' (Dataset (default), Downstream, Importers, 5/4/23, 5:13:16 AM, —, General). On the right, there's a 'Show objects' section with a list of items and a 'Filter by keyword' search bar.



#GPPB2024



#PowerPlatformMadrid



Lab 5



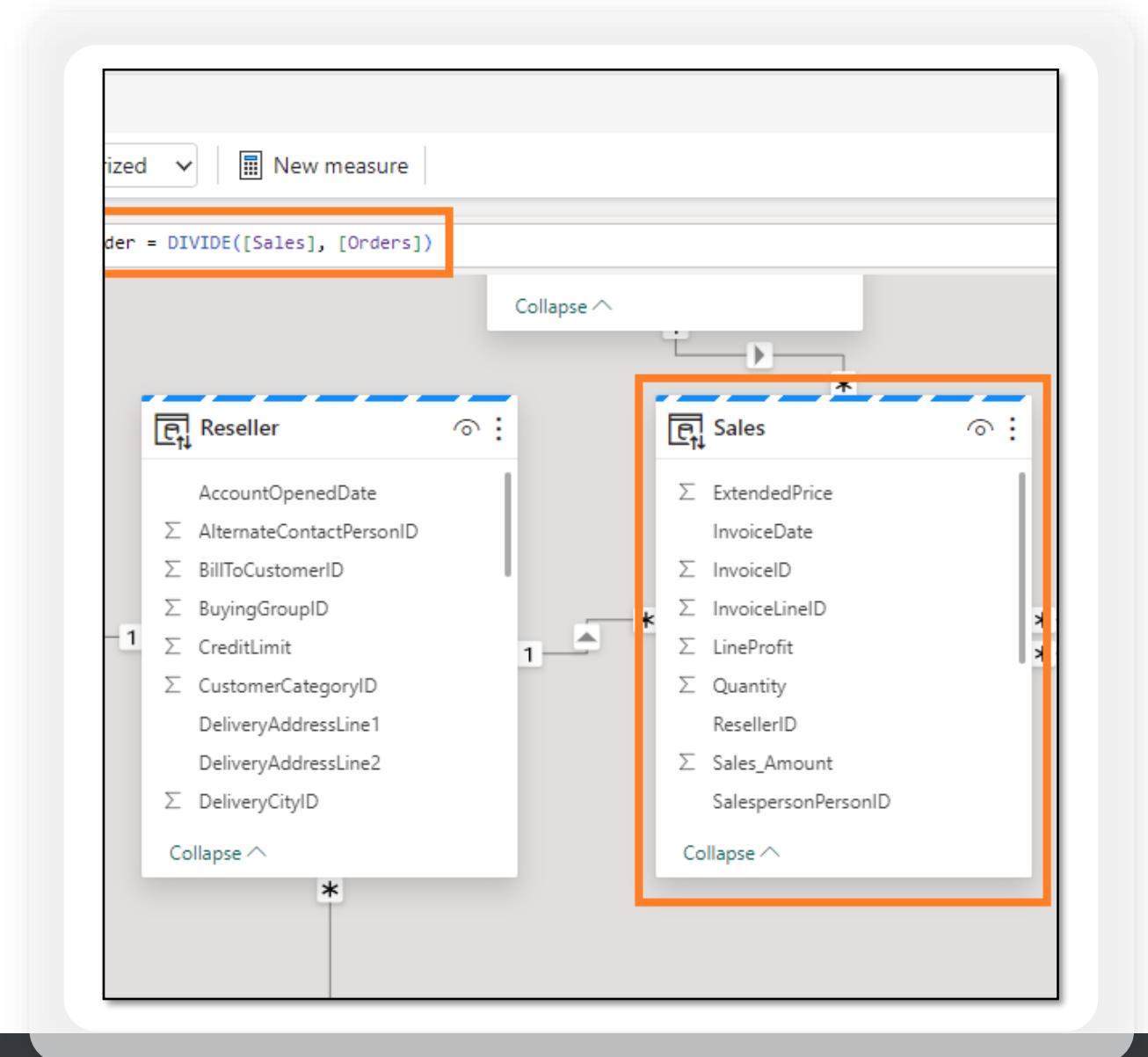
Explorar un Lakehouse



Explorar una Vista SQL de un
Lakehouse



Como trabajar con Data modeling



#GPPB2024



#PowerPlatformMadrid

Capacidades y Licenciamiento



#GPPB2024



#PowerPlatformMadrid

Licenciamiento

Dentro de una capacidad de Fabric, en Spark, los usuarios obtienen dos Spark VCores por cada unidad de capacidad que reservan como parte de su SKU.

Una Capacity Unit (CU) = dos Spark VCores

Los VCores de Spark se comparten entre todos los elementos basados en Spark:

- Notebooks
- Definiciones de trabajos de Spark
- Lakehouses creados en estas áreas de trabajo.

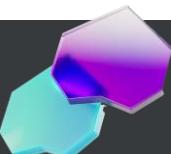
Fabric capacity SKU	Equivalent Power BI SKU	Spark VCores	Max Spark VCores with Burst Factor	Queue limit
F2	-	4	20	4
F4	-	8	24	4
F8	-	16	48	8
F16	-	32	96	16
F32	-	64	192	32
F64	P1	128	384	64
F128	P2	256	768	128
F256	P3	512	1536	256
F512	P4	1024	3072	512
F1024	-	2048	6144	1024
F2048	-	4096	12288	2048
Trial Capacity	P1	128	128	NA



#GPPB2024



#PowerPlatformMadrid



FABRIC LICENSING

AS OF TODAY, CAN CHANGE BETWEEN NOW AND GENERAL AVAILABILITY.

CAPACITY

SHARED ACROSS ALL ANALYTICS WORKLOADS
CAN BE RESIZED UP, DOWN AND EVEN PAUSED
ATTACHED TO A SPECIFIC REGION
SOLD AS CAPACITY UNIT (CU), FROM 2 TO 2048

\$0.18/CU/Hour, MINIMUM 2 CU (ONLY \$9/DAY!)

STORAGE

YOU PAY FOR THE STORAGE YOU CONSUME.
PRICE VARIES FROM REGION TO REGION
\$0.023/GB/MONTH FOR US WEST 2

NETWORK

CROSS-REGION DATA TRANSFER CHARGES,
EGRESS CHARGES FOR MULTI-CLOUD SHORTCUTS

INFRASTRUCTURE



INDIVIDUAL LICENSE

EACH FABRIC USER NEEDS A LICENCE. THERE ARE
3 TYPES OF LICENCE: FREE, PRO AND PREMIUM.

FREE

ACCESS FABRIC WEB APP
CREATE FABRIC WORKSPACES
CREATE NON-POWER BI ITEMS IN FABRIC WORKSPACES
SHARE NON-POWER BI ITEMS IN FABRIC WORKSPACES
VIEW SHARED POWER BI ITEMS IF WORKSPACE HAS A
FABRIC CAPACITY => 64CU

PRO

\$10/USER
/MONTH

VIEW SHARED CONTENT
CREATE POWER BI ITEMS IN ALL WORKSPACES

PREMIUM

ENABLE POWER BI PREMIUM FEATURES
\$20/USER/MONTH

Lab 6



Limpieza de Entorno



#GPPB2024

Microsoft Fabric



#PowerPlatformMadrid

Workspace settings



About



Premium



Azure connections



System storage



Git integration



OneLake



Other

Delete this workspace and all data and items in it?

This workspace will be permanently deleted after a retention period of your Fabric administrator. Contact your Fabric administrator if you need to workspace and recover the items in it during this retention period.



Remove this workspace

Power BI



Data

Engineering/Science



Próximos Pasos



- Explore el producto aquí:
- <https://www.microsoft.com/en-us/microsoft-fabric>



- Vea el programa de Microsoft Mechanics:
- <https://youtu.be/oxMUT3p9f4>



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



- Lea el E-Book gratuito: Introducción a Fabric
- <https://aka.ms/fabric-get-started-ebook>



#GPPB2024



#PowerPlatformMadrid

Links de Referencia

Links para obtener más información sobre Microsoft Fabric:

- Consulte la publicación del blog para leer la información completa [Microsoft Fabric GA announcement](#)
- Explora Fabric a través del [Guided Tour](#)
- Regístrate al [Microsoft Fabric free trial](#)
- Visiti [Microsoft Fabric website](#)
- Aprenda nuevas habilidades explorando los [Fabric Learning modules](#)
- Explore [Fabric technical documentation](#)
- Lea [free e-book on getting started with Fabric](#)
- Unase a la [Fabric community](#) para publicar sus preguntas, compartir sus comentarios y aprender de los demás

Lea los blogs más detallados sobre el anuncio de la experiencia de Fabric:

- [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](#)



#GPPB2024



#PowerPlatformMadrid

Tutoriales

 **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>



#GPPB2024



#PowerPlatformMadrid