

DP-700: Ingeniero de datos de Microsoft Fabric



Introducción a los Data Warehouses en Microsoft Fabric



Objetivos de aprendizaje



- Descripción de los Data Warehouses en Fabric.
- Descripción de un Data Warehouse frente a un almacén de lago de datos.
- Trabajo con los Data Warehouses en Fabric.
- Creación y administración de tablas de hechos y dimensiones dentro de un Data Warehouse.

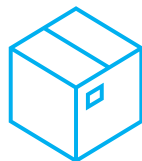
Conceptos básicos del Data Warehouse

Ingesta de datos



Mueve los datos de sistemas de origen a un Data Warehouse.

Data Warehouse



Almacena los datos en un formato optimizado para el análisis.

Procesamiento de datos



Transforma los datos en un formato listo para su consumo con las herramientas de análisis.

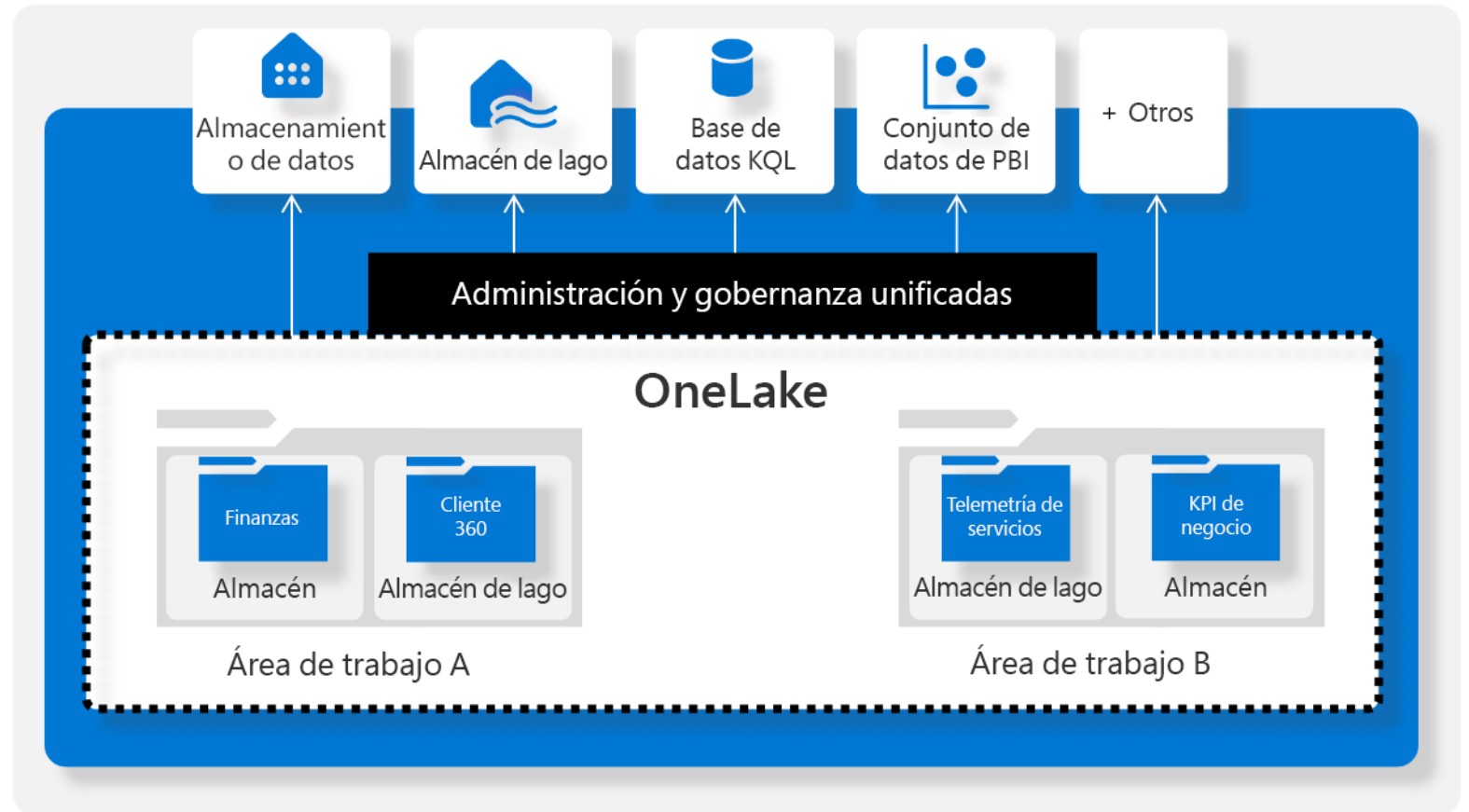
Análisis y entrega de datos



Analiza los datos para obtener información y proporcionarla a la empresa.

Descripción de los almacenes de Fabric

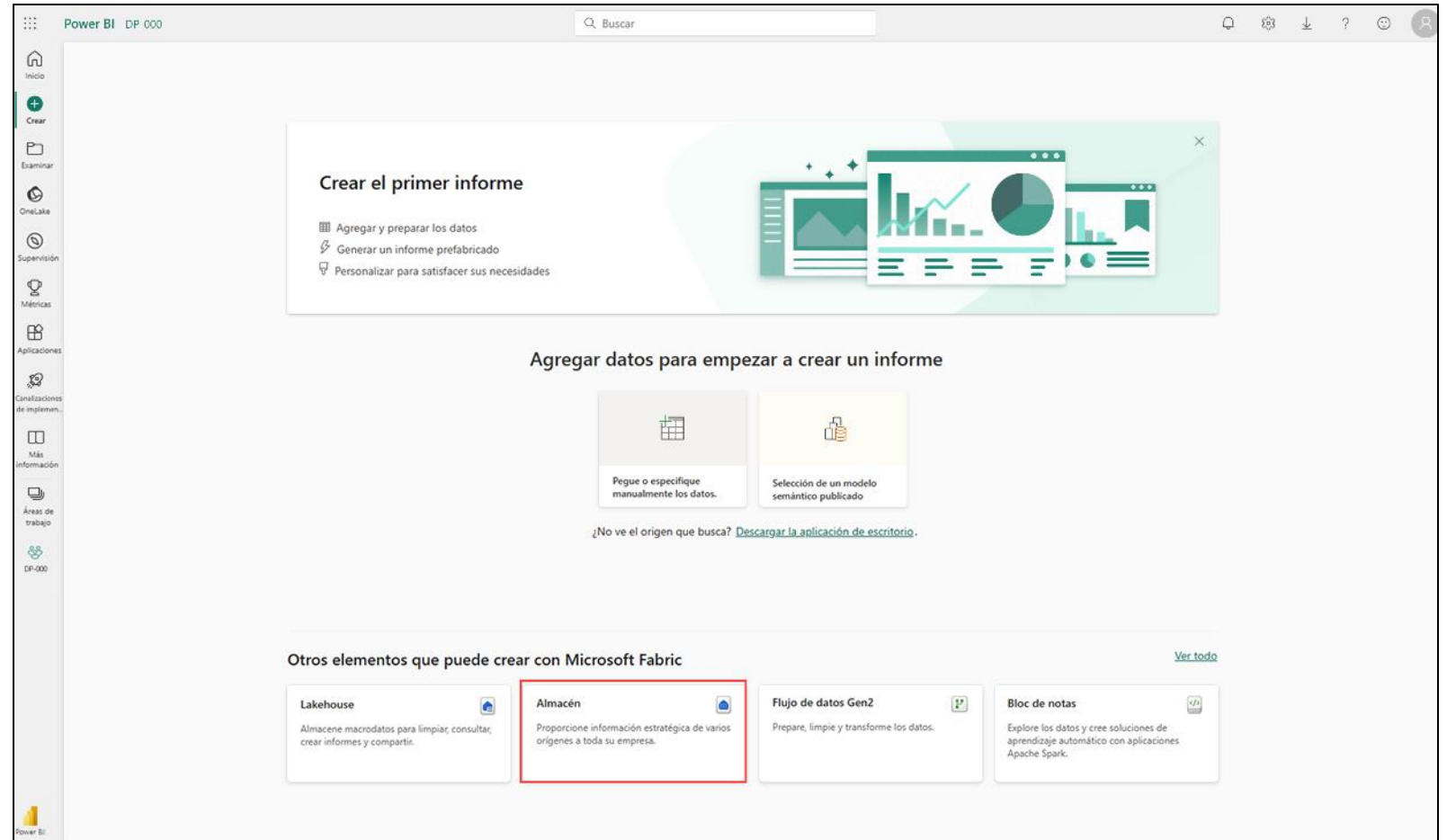
- Centrados en un único lago de datos
- Con tecnología de Synapse Analytics
- Totalmente compatibles con T-SQL
- Formato de archivo Parquet



Creación de un Data Warehouse en Fabric

Dentro del portal de servicios de Fabric

- Navegar al área de trabajo de Fabric
- Seleccionar nuevo elemento > Almacén
- Nombrar almacén



Diseño de un Data Warehouse

Al igual que todas las bases de datos relacionales, el Data Warehouse de Fabric contiene tablas para almacenar los datos para el análisis.

Tablas de hechos:

- Las tablas de hechos contienen los datos numéricos que quieres analizar.
- Las tablas de hechos suelen tener un gran número de filas y son la principal fuente de datos para el análisis.

Tablas de dimensiones:

- Contienen información descriptiva sobre los datos de las tablas de hechos.
- Las tablas de dimensiones suelen tener un pequeño número de filas y se usan para proporcionar contexto para los datos de las tablas de hechos.
- Es habitual que una tabla de dimensiones incluya dos columnas de claves: una ***clave suplente*** y una ***clave alternativa***

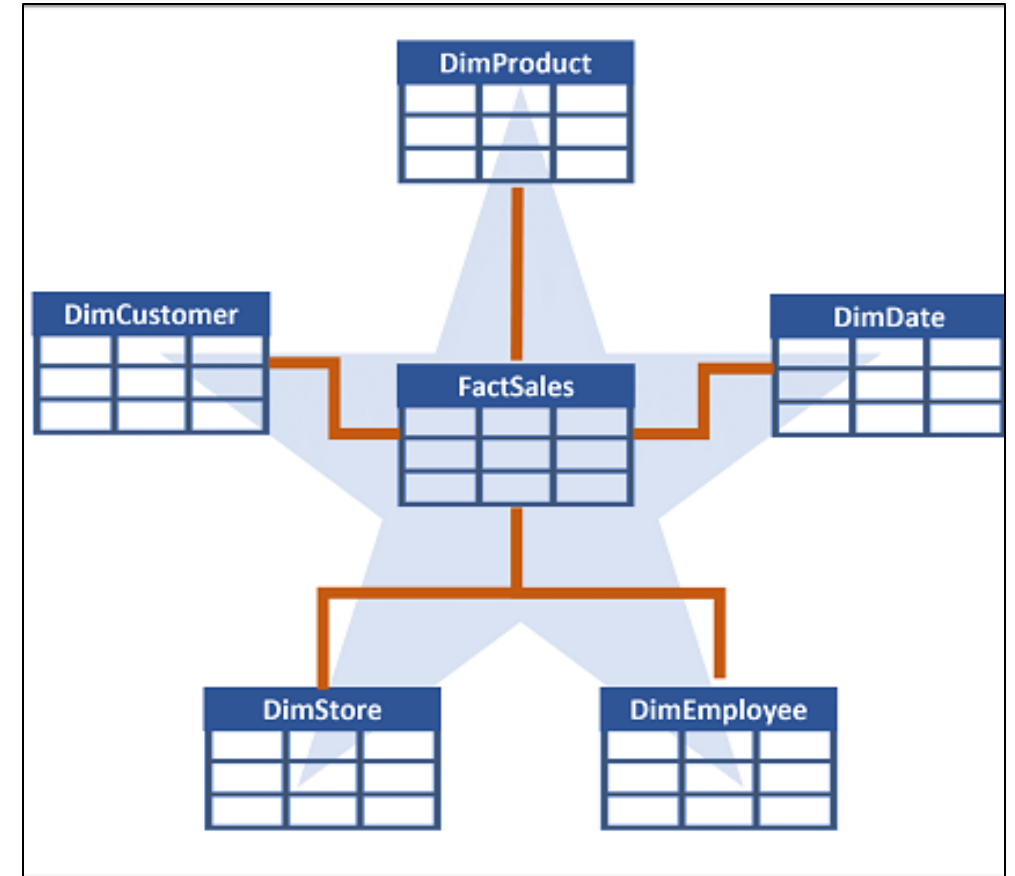
Tipos especiales de tablas de dimensiones

Las ***dimensiones de tiempo*** proporcionan información sobre el período de tiempo en el que se produjo un evento.

Las dimensiones de variación lenta son tablas de dimensiones que hacen el seguimiento de los cambios en los atributos de dimensión a lo largo del tiempo, como los cambios en la dirección de un cliente o el precio de un producto.

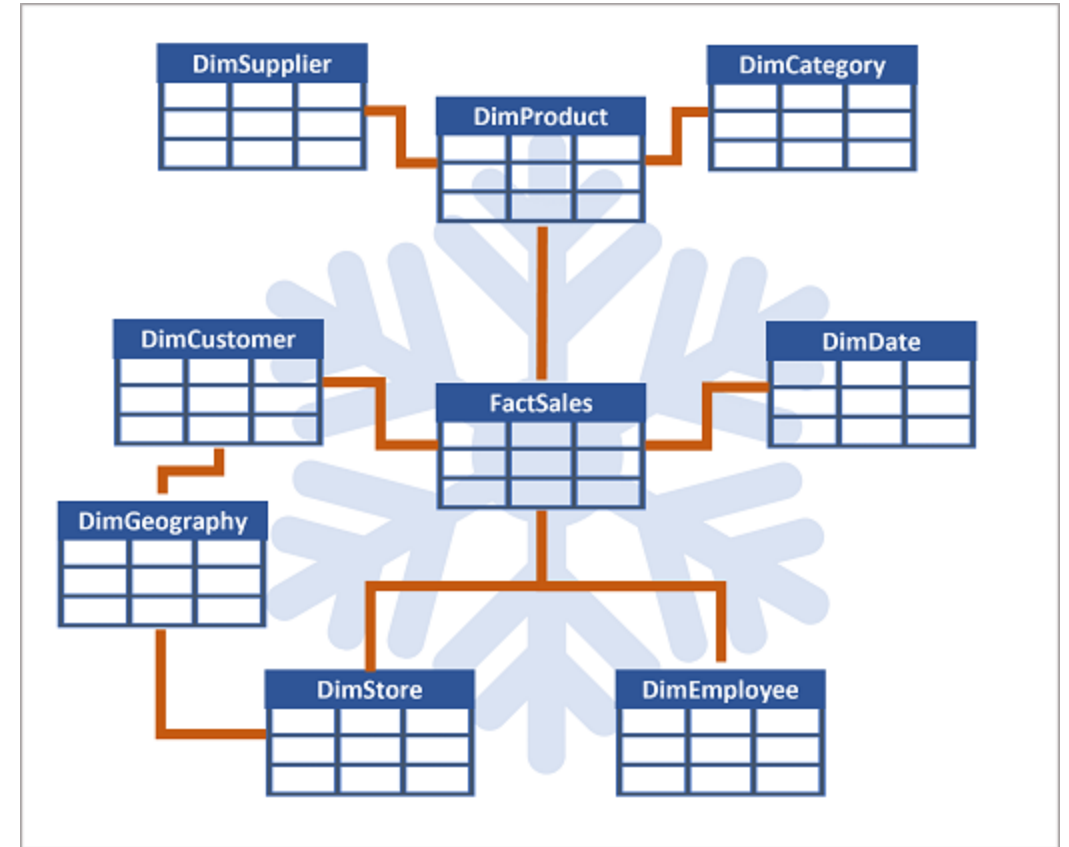
Diseños de esquemas de Data Warehouse (1/2)

- Un Data Warehouse se organiza como un esquema en estrella, en el que una tabla de hechos está directamente relacionada con las tablas de dimensiones.



Diseños de esquemas de Data Warehouse (2/2)

- Si hay muchos niveles o se comparte información por cosas diferentes, podría tener sentido usar un *esquema de copo de nieve* en su lugar.



Ingesta de datos en un Data Warehouse

Canalizaciones
de Fabric

Flujos de
datos de
Fabric Gen2

Consultas
entre bases de
datos

Comando
COPY INTO

Clonación de tabla

Réplicas de tablas creadas al copiar los metadatos mientras sigue haciendo referencia a los mismos archivos de datos en OneLake.

Desarrollo y
pruebas

Recuperación
de datos

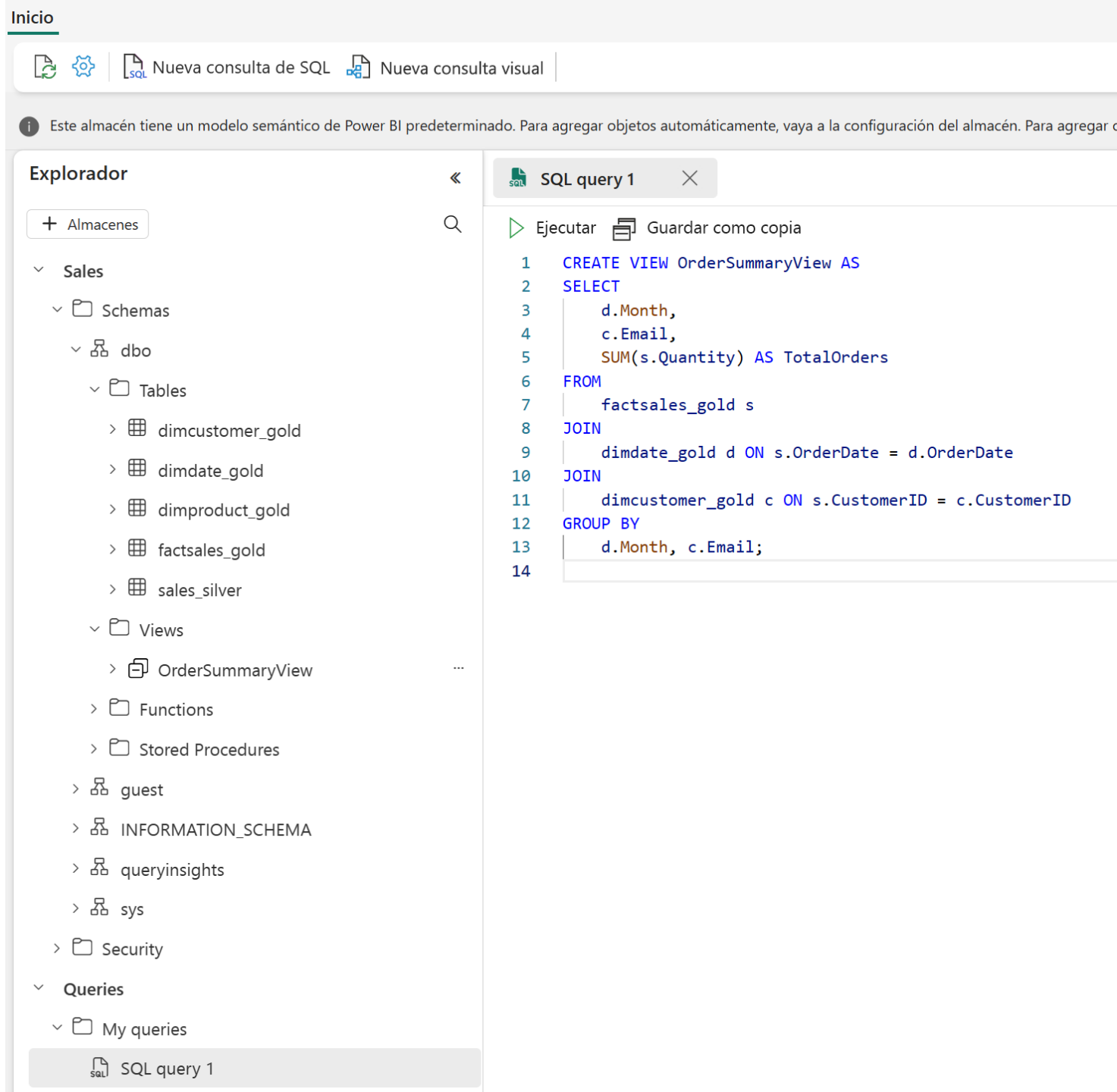
Informes
históricos

Consultar datos

Uso del editor de consultas SQL

Como usar SQL Server Management Studio (SSMS) o Azure Data Studio (ADS).

- Intellisense
- Finalización de código
- Resaltado de sintaxis
- Análisis del lado del cliente
- Validation



Visualización de consultas

Creación de una consulta visual

Arrastrar la tabla al lienzo y, después, transformarla según sea necesario:

- Administrar columnas
- Filtrar filas
- Reemplazo de valores
- Agrupar por
- Y mucho más

Nueva consulta visual

ántico de Power BI predeterminado. Para agregar objetos automáticamente, vaya a la configuración del almacén. Para agregar objetos manualmente, use Administrar modelo semántico predeterminado. [Más información](#)

SQL query 1 SQL query 2 SQL query 3 Visual query 1

Administrar columnas Reducir filas Ordenar Transformación Combinar Guardar como copia Vista SQL

DimProduct FactSalesOrder

Source Database Table Consultas combinadas

Buscar comandos

- Administrar columnas
 - Elegir columnas
 - Quitar columnas
 - Quitar otras columnas
- Reducir filas
 - Conservar filas superiores
 - Conservar duplicados
 - Filtrar filas
- Ordenar
 - Orden ascendente
 - Orden descendente
- Combinar
 - Combinar consultas
 - Combinar consultas como nuevas
 - Anexar consultas
 - Anexar consultas como nuevas
- Transformar tabla
 - Agrupar por

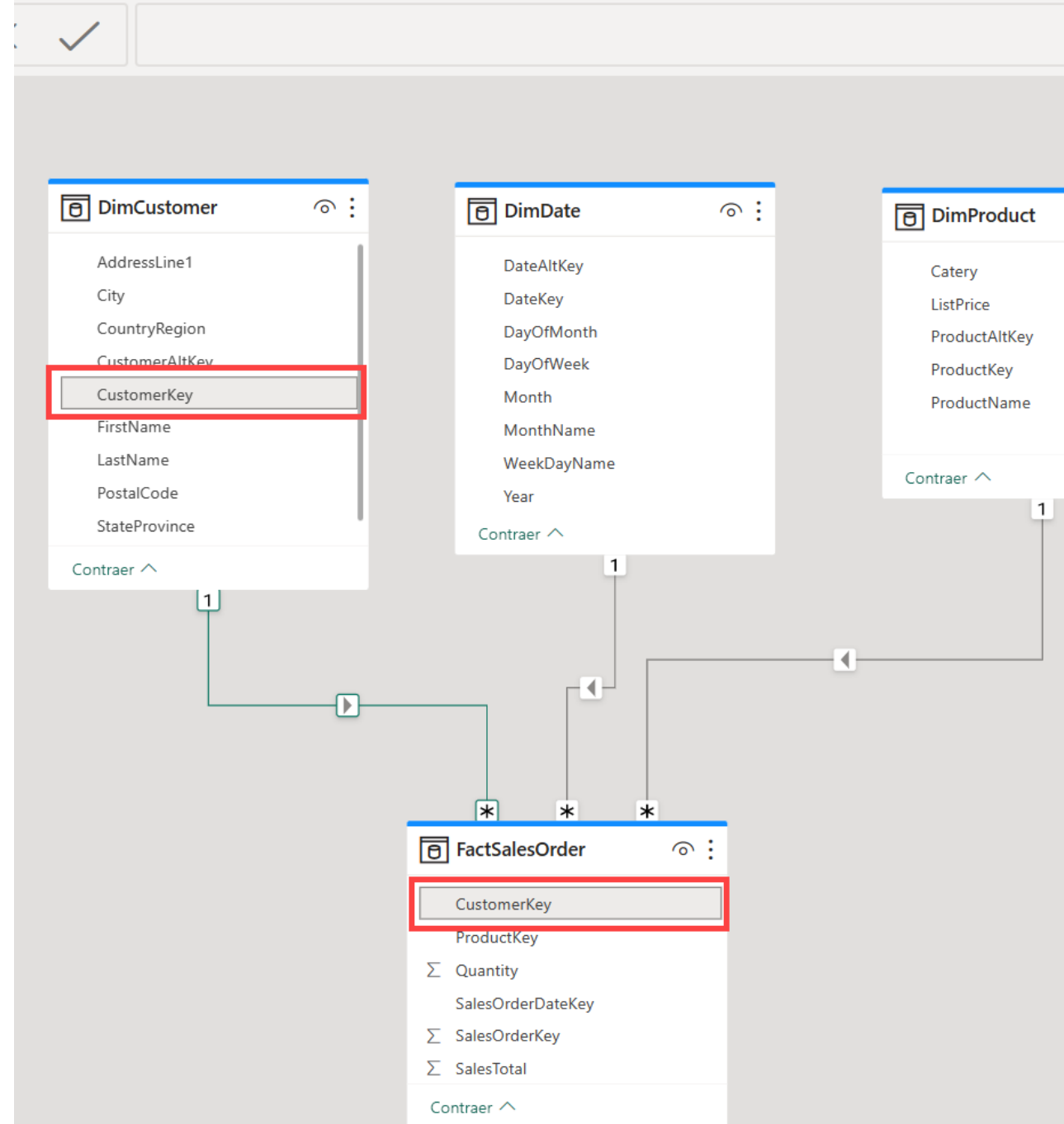
Descargar archivo de Excel Visualización de resultados

	1 ² SalesOrderKey	1 ² SalesOrderDateKey	1 ² ProductKey	1 ² CustomerKey	1 ² Quantity	1.2 SalesTotal	1 ² ProductName
1	5306	20221121	843	480	2	50	Cable Lock
2	4685	20220819	843	52	5	125	Cable Lock
3	4690	20220821	843	29580	9	225	Cable Lock
4	5007	20221011	843	29583	3	75	Cable Lock
5	3734	20220317	843	400	11	275	Cable Lock
6	3917	20220416	843	29778	8	200	Cable Lock
7	3922	20220418	843	1	8	200	Cable Lock
8	2932	20211024	843	665	5	125	Cable Lock
9	2976	20211102	843	29915	2	50	Cable Lock
10	3199	20211207	843	564	11	275	Cable Lock
11	1319	20210118	843	29911	8	200	Cable Lock
12	1536	20210222	843	131	5	125	Cable Lock
13	1558	20210224	843	214	8	200	Cable Lock

Construir relaciones

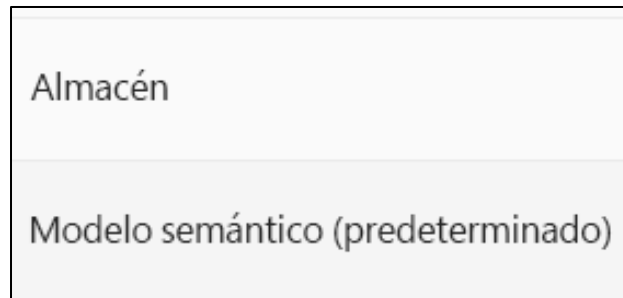
Conexión de tablas de hechos y dimensiones

- Usar la vista **Modelo** en el almacén del portal de Fabric.
- Identificar la clave adecuada entre tablas.
- Arrastrar columnas de una tabla encima de la columna relacionada de la otra tabla.



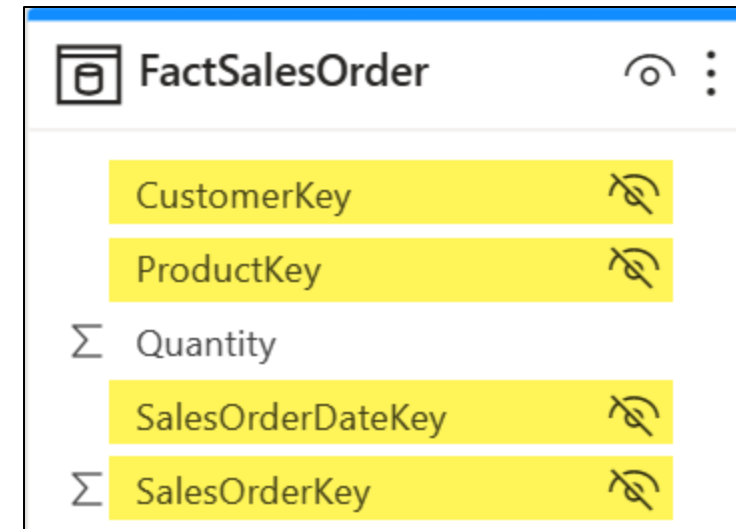
Descripción del modelo semántico predeterminado

Fabric crea automáticamente un modelo semántico predeterminado para que los usuarios de Power BI lo usen para la creación de informes.



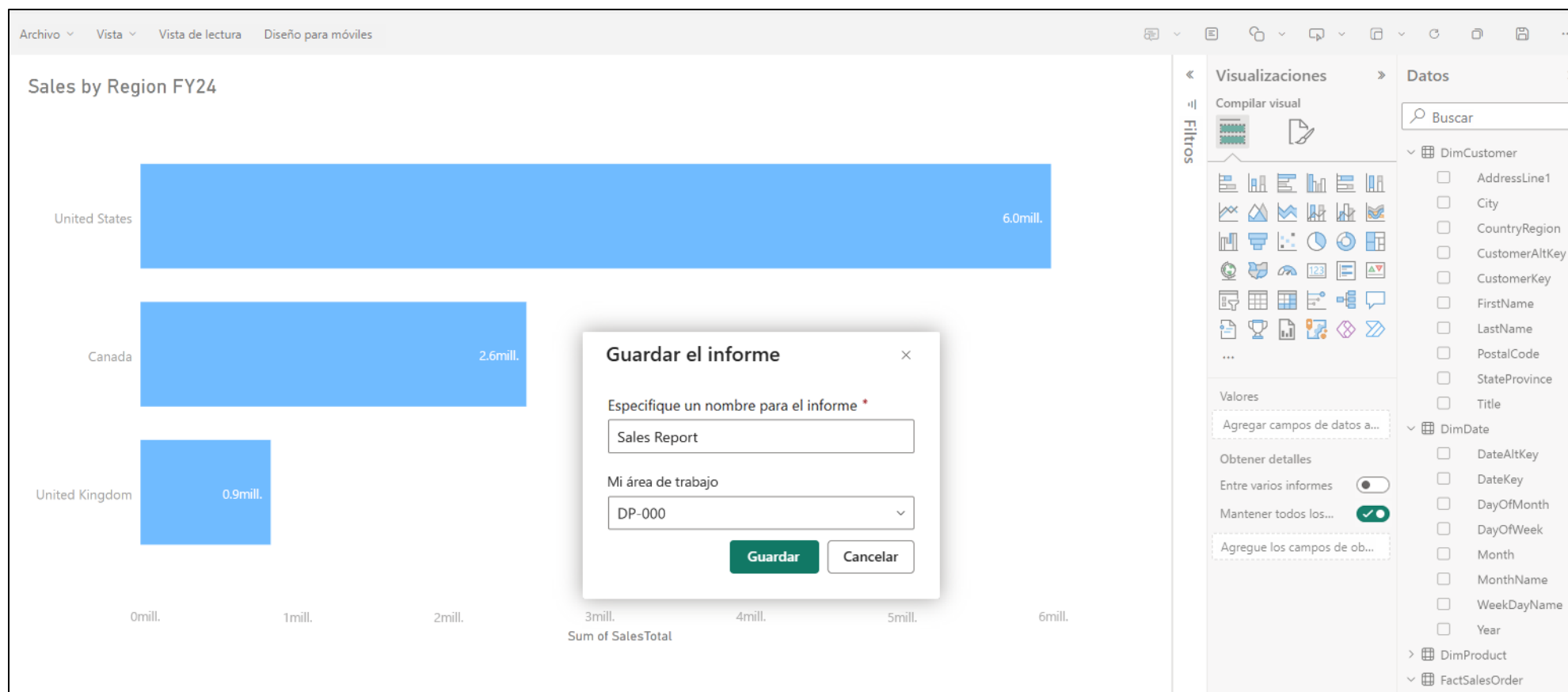
Considera la posibilidad de dar forma al modelo semántico para reducir las transformaciones de nivel inferior.

- Nombres de columna descriptivos
- Ocultar columnas o tablas de la vista



Visualización de datos

Aunque esta tarea normalmente puede ser responsabilidad de los analistas de datos, es útil saber que puedes crear informes de Power BI directamente desde el almacén de Fabric.



Introducción a la seguridad

- Control de acceso basado en rol (RBAC)
- Cifrado SSL
- Cifrado del servicio Azure Storage
- Azure Monitor y Azure Log Analytics
- Autenticación multifactor (MFA)
- Integración de Microsoft Entra ID

Permisos de área de trabajo y elementos

Áreas de trabajo

- Contenedores, almacenar elementos relacionados
- Los roles conceden acceso al área de trabajo
 - Administración
 - Colaborador
 - Miembro
 - Visor

Productos

- Almacenes de lagos de datos, almacenes, modelos semánticos, informes, paneles, etc.
- Acceso de nivel de elemento en el área de trabajo
- Acceso de punto de conexión de análisis SQL
 - Leer
 - ReadData
 - ReadAll

Ejercicio

30 minutos



Análisis de los datos en un Data Warehouse

Prueba de conocimientos



1 ¿Qué tipo de tabla debe usar una compañía de seguros para almacenar los detalles del atributo de proveedor para agregar reclamaciones?

- ☐ Tabla de hechos.
- ☒ Tabla de dimensiones.
- ☐ Tabla de almacenamiento provisional.

2 ¿Qué es un conjunto de datos en la experiencia del Data Warehouse?

- ☒ Los conjuntos de datos son un modelo semántico que se usan para crear informes.
- ☐ Los conjuntos de datos son una colección de archivos que contienen datos no organizados.
- ☐ Los conjuntos de datos son un conjunto de variables y observaciones que no se han limpiado ni preprocesado.

3 ¿Cuál es la finalidad de los permisos de elemento en un área de trabajo?

- ☐ Conceder acceso a todos los elementos de un área de trabajo.
- ☐ Conceder acceso a columnas específicas dentro de una tabla.
- ☒ Conceder acceso a almacenamientos individuales para el consumo descendente.

Resumen

En esta sección, se ha descrito lo siguiente:

- Los conceptos de Data Warehouse y modelado dimensional
- La creación de almacenes y la ingesta de datos
- La consulta y visualización de datos
- Los modelos semánticos y su propósito para los informes derivados
- La protección de un Data Warehouse

Información adicional

Introducción a los Data Warehouses en Microsoft Fabric


<https://aka.ms/fabric-warehouse-spa>



Cargar datos en un Data Warehouse de Microsoft Fabric

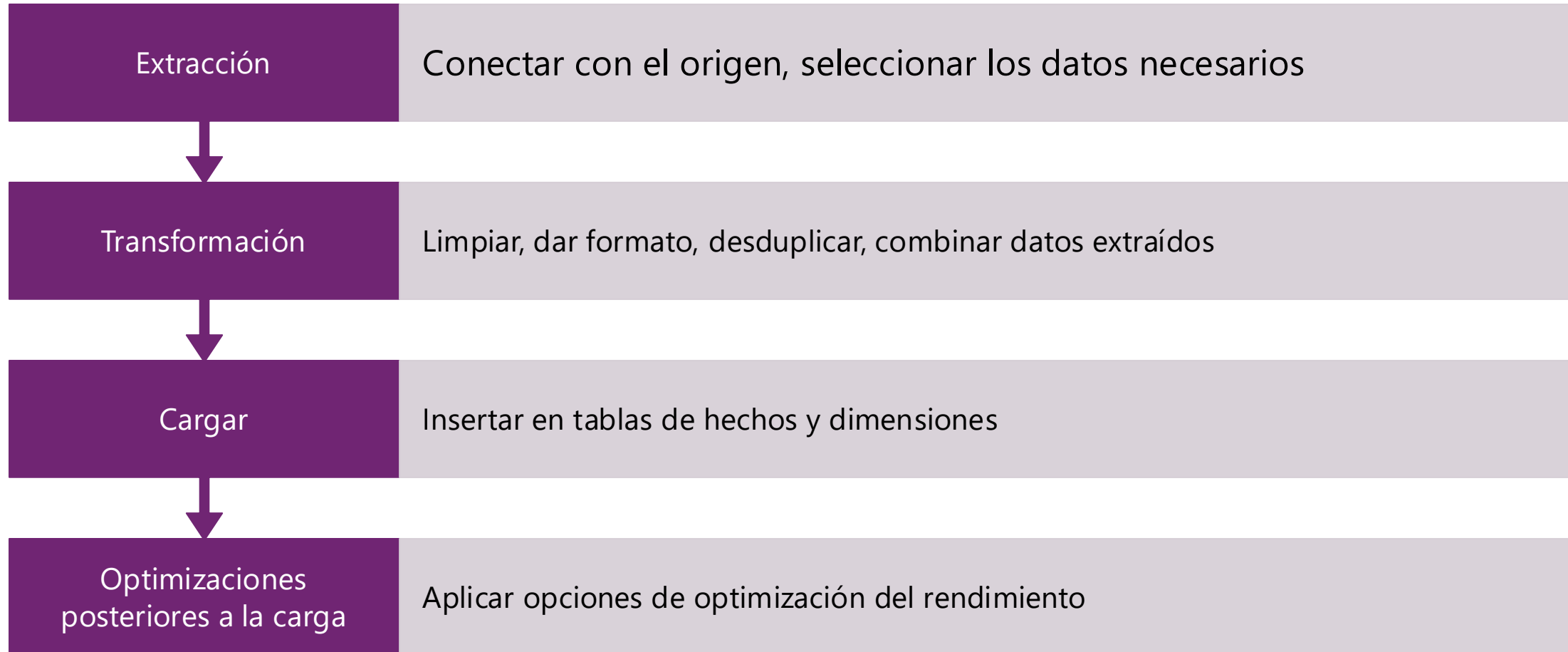


Objetivos de aprendizaje



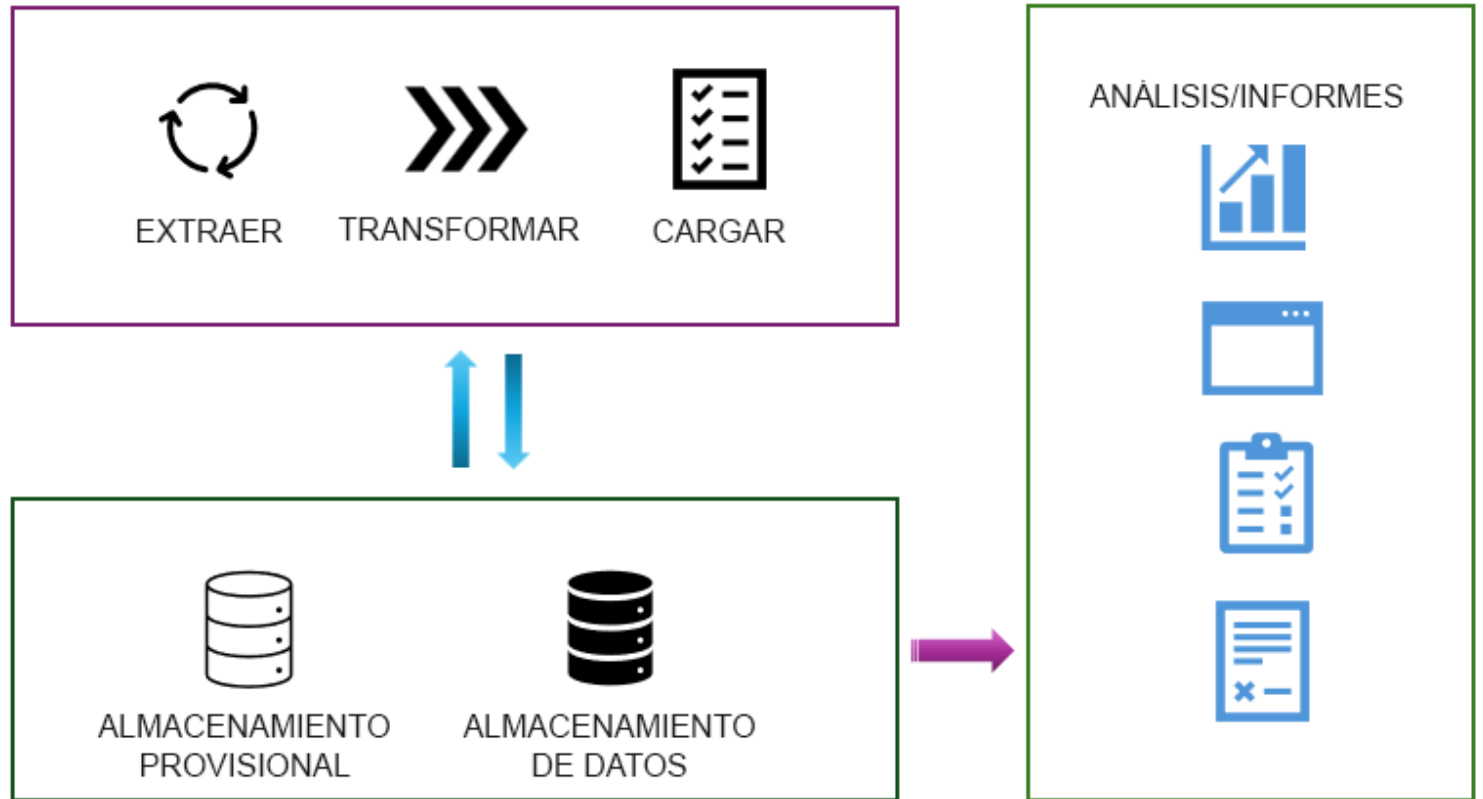
- Estrategias para cargar datos en un Data Warehouse en Microsoft Fabric.
- Creación de una canalización de datos para cargar un almacén en Microsoft Fabric.
- Carga de datos en un almacén mediante T-SQL.
- Carga y transformación de datos con flujo de datos (Gen 2).

Descripción de ETL (extracción, transformación y carga)



Preparación de los datos

- Almacenamiento y transformación temporales
- Simplifica las operaciones de carga



Diferentes tipos de carga de datos

Carga completa (inicial)

- Trunca la tabla y carga los datos
- Tiempo de carga más largo
- Ningún historial almacenado
- Es mejor para la carga inicial o la actualización

Carga incremental

- Anexar datos a tablas
- Actualizaciones más rápidas
- Conserva el historial (marca de tiempo)
- Es mejor para las actualizaciones frecuentes

Claves de dimensión

Clave suplente

- Identifica de forma única una instancia de una entidad de dimensión (es decir, una fila).
- Normalmente, un valor entero simple.
- Debe ser única en la tabla de dimensiones.

Clave alternativa

- Identifica una entidad en el sistema de origen operativo.
- A menudo, una clave *empresarial* (por ejemplo, un código de producto o un identificador de cliente) o una clave *natural* (p. ej., un valor datetime en una dimensión de tiempo).
- Se puede duplicar en la tabla de dimensiones para representar la misma entidad en distintos momentos en el tiempo.

CustomerKey	CustomerAltKey	Name	Email	Street	City	PostalCode	CountryRegion
123	I-543*	Navin Jones	navin1@contoso.com	1 Main St.	Seattle	90000	United States
124	R-589	Mary Smith	mary2@contoso.com	234 190th Ave	Buffalo	50001	United States
125	I-321	Antoine Dubois	antoine1@contoso.com	2 Rue Jolie	Paris	20098	France
126	I-543*	Navin Jones	navin1@contoso.com	24 125th Ave.	New York	50000	United States
...

*Este cliente se trasladó de Seattle a Nueva York, por lo que se agregó un nuevo registro con la misma clave alternativa, pero con una nueva clave suplente.

Creación de tablas con T-SQL

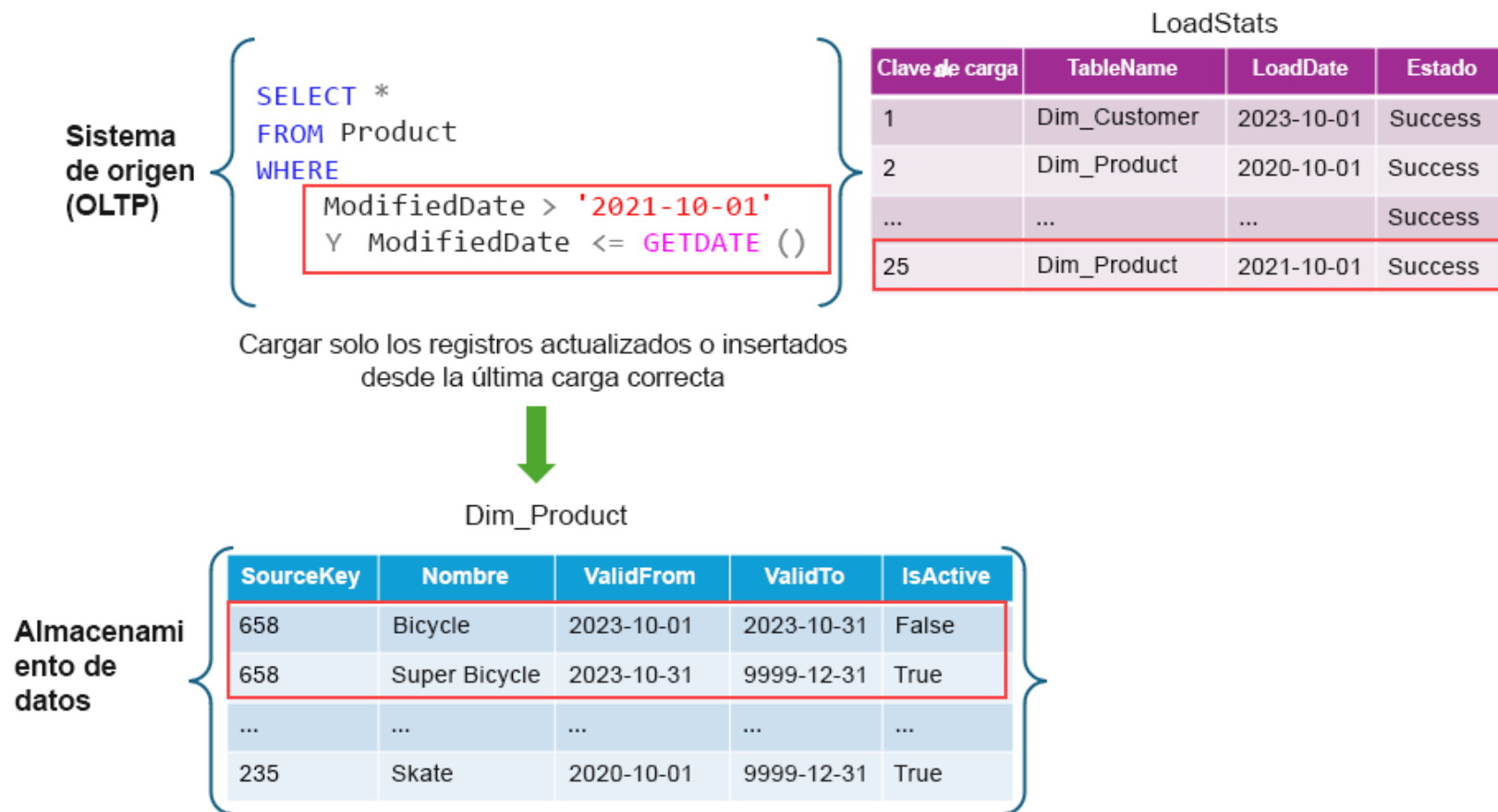
- SQL Server Management Studio
- Portal de servicio de Microsoft Fabric

```
CREATE TABLE DimProduct (  
    ProductID INT PRIMARY KEY,  
    ProductName NVARCHAR(100),  
    CategoryID INT,  
    Price DECIMAL(10, 2)  
);
```

Carga de tablas de dimensiones

Tipos de dimensión que cambian lentamente

- **0**: nunca cambia
- **1**: se sobrescribe, sin historial
- **2**: anexa, realiza un seguimiento de los cambios
- **3**: agrega columna de historial
- **4**: agrega nueva dimensión
- **5**: como el tipo 2 para cambios grandes
- **6**: combinación de los tipos 2 y 3



Ejemplo

Actualiza la tabla *Dim_Products*,
diseñada como **SCD de tipo 2**.

```
IF EXISTS (SELECT 1 FROM Dim_Products WHERE SourceKey =
@ProductID AND IsActive = 'True')
BEGIN
    -- Existing product record
    UPDATE Dim_Products
    SET ValidTo = GETDATE(), IsActive = 'False'
    WHERE SourceKey = @ProductID AND IsActive = 'True';
END
ELSE
BEGIN
    -- New product record
    INSERT INTO Dim_Products (SourceKey, ProductName,
StartDate, EndDate, IsActive)
    VALUES (@ProductID, @ProductName, GETDATE(), '9999-
12-31', 'True');
END
```


Cargar tablas de hechos

- Usa una instrucción INSERT
- Busca claves suplentes en las tablas de dimensiones en función de la clave alternativa:

Caso simple: obtén la instancia de dimensión cargada más recientemente (clave suplente de incremento máximo).

O bien, usa un campo de marca *IsCurrent*.

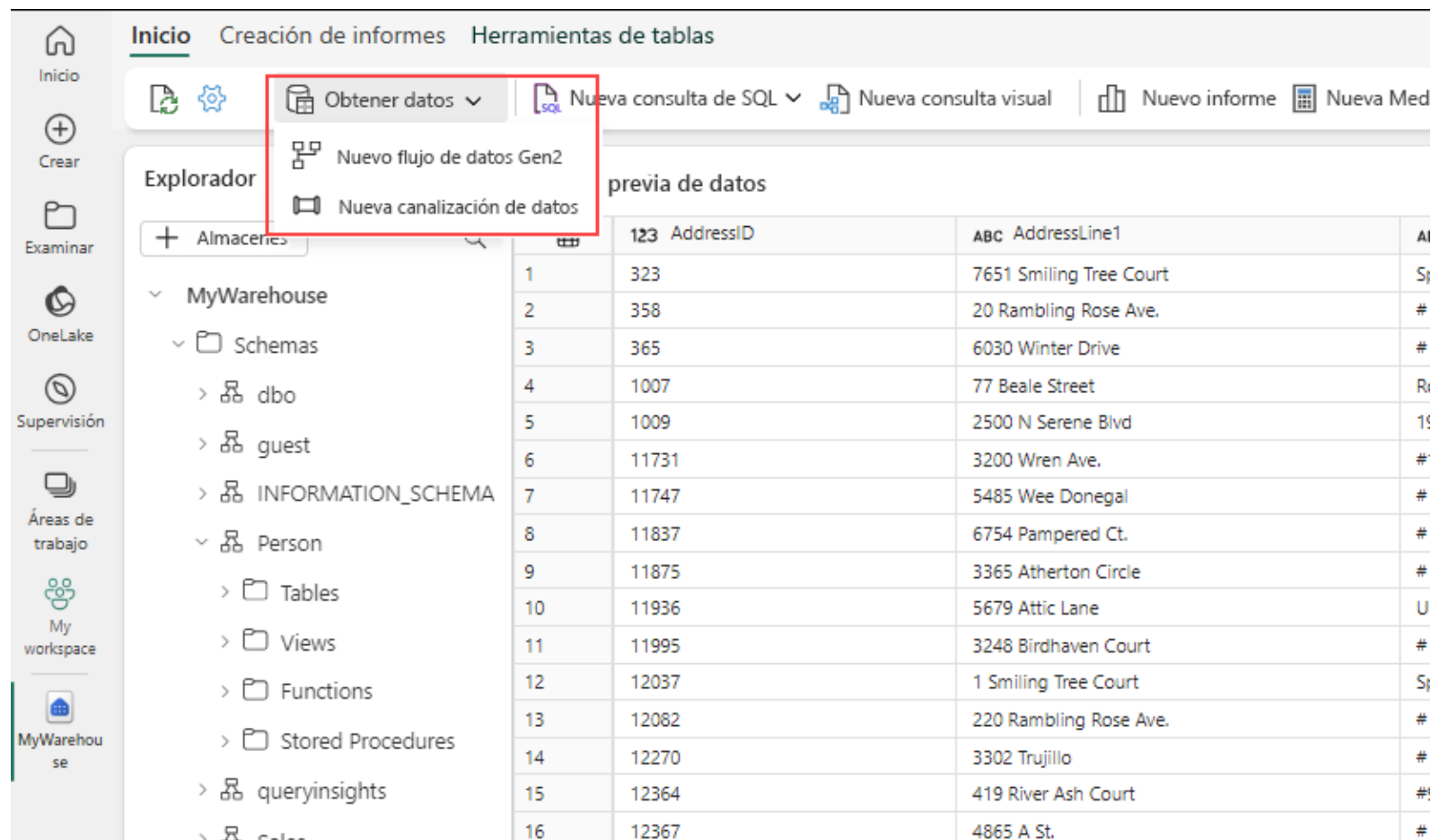
O bien, usa las fechas de *inicio* y *finalización* para encontrar la instancia adecuada para la fecha de los hechos.

```
INSERT INTO dbo.FactSales
SELECT (SELECT MAX(DateKey)
        FROM dbo.DimDate
        WHERE FullDateAlternateKey = stg.OrderDate) AS
OrderDateKey,
       (SELECT MAX(CustomerKey)
        FROM dbo.DimCustomer
        WHERE CustomerAlternateKey = stg.CustNo) AS
CustomerKey,
       (SELECT MAX(ProductKey)
        FROM dbo.DimProduct
        WHERE ProductAlternateKey = stg.ProductID) AS
ProductKey,
       (SELECT MAX(StoreKey)
        FROM dbo.DimStore
        WHERE StoreAlternateKey = stg.StoreID) AS StoreKey,
OrderNumber,
OrderLineItem,
OrderQuantity,
UnitPrice,
Discount,
Tax,
SalesAmount
FROM dbo.StageSales AS stg
```

Carga de datos con canalizaciones de Fabric

Organizar las actividades

- Copiar datos
- Programación
- Parámetros
- Eliminar datos

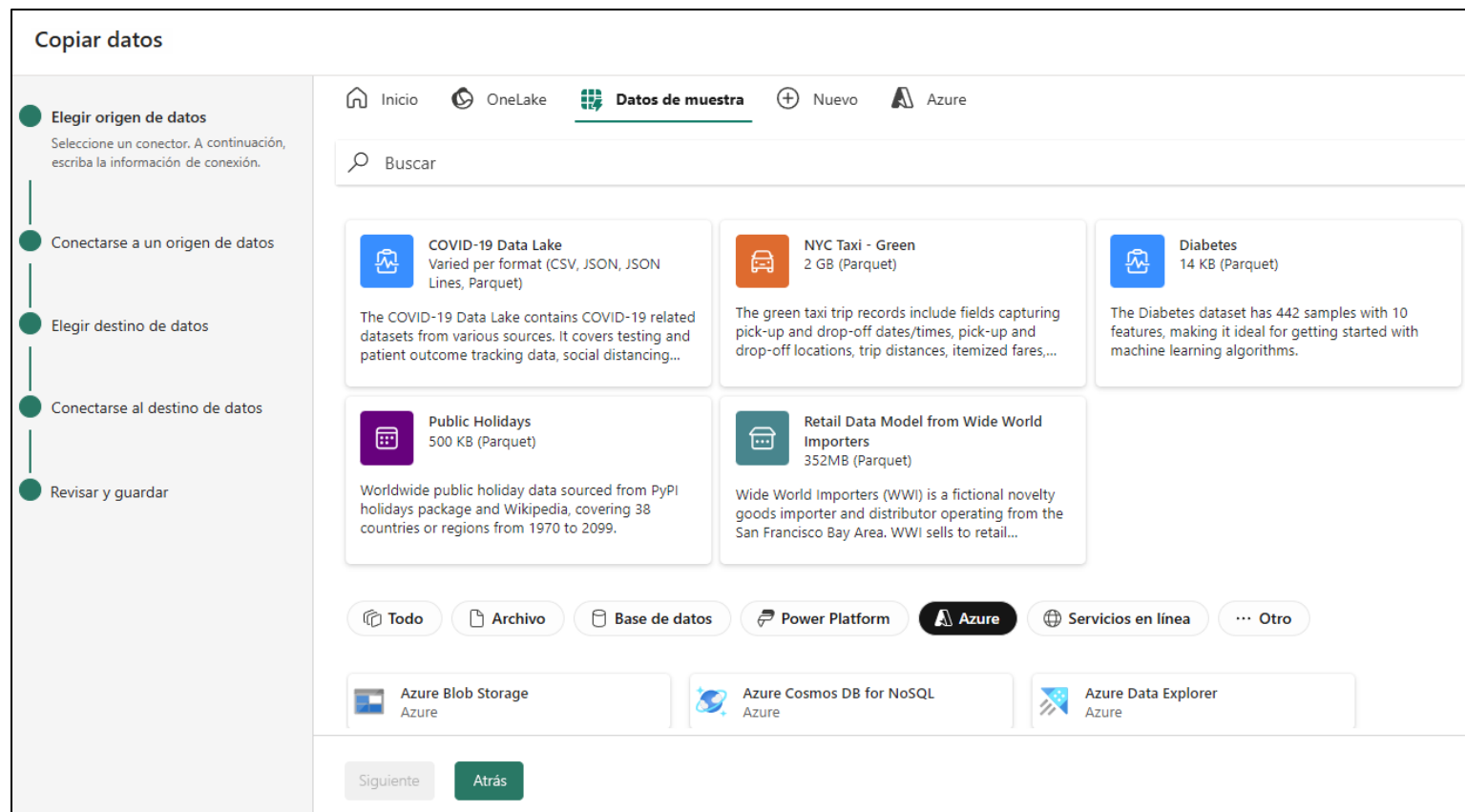


The screenshot displays the Microsoft Fabric user interface. On the left is a vertical navigation pane with icons for Inicio, Crear, Examinar, OneLake, Supervisión, Áreas de trabajo, My workspace, and MyWarehouse. The main area is divided into three sections: a top ribbon with 'Inicio', 'Creación de informes', and 'Herramientas de tablas'; a central 'Explorador' (Explorer) pane showing a tree view of 'MyWarehouse' (Schemas, dbo, guest, INFORMATION_SCHEMA, Person, Tables, Views, Functions, Stored Procedures, queryinsights); and a right pane titled 'previa de datos' (Data preview) showing a table of address data. A red rectangle highlights the 'Obtener datos' (Get data) dropdown menu, which contains options for 'Nuevo flujo de datos Gen2' and 'Nueva canalización de datos'.

	123 AddressID	ABC AddressLine1	AI
1	323	7651 Smiling Tree Court	S
2	358	20 Rambling Rose Ave.	#
3	365	6030 Winter Drive	#
4	1007	77 Beale Street	R
5	1009	2500 N Serene Blvd	1
6	11731	3200 Wren Ave.	#
7	11747	5485 Wee Donegal	#
8	11837	6754 Pampered Ct.	#
9	11875	3365 Atherton Circle	#
10	11936	5679 Attic Lane	U
11	11995	3248 Birdhaven Court	#
12	12037	1 Smiling Tree Court	S
13	12082	220 Rambling Rose Ave.	#
14	12270	3302 Trujillo	#
15	12364	419 River Ash Court	#
16	12367	4865 A St.	#

Configurar el asistente de datos de copia

- Elegir origen de datos
- Conectarse a un origen de datos de datos
- Elegir destino de datos
- Conectarse al destino de datos
- Configuración



Carga de datos con T-SQL

Uso de la instrucción COPY

- Almacenamiento externo de Azure
- Especificar el formato de archivo (PARQUET /CSV)
- Control de errores
- Varios archivos

```
-- Load as CSV
```

```
COPY INTO my_table
```

```
FROM
```

```
'https://myaccount.blob.core.windows.net/myblobcontainer/folder0/*.csv',  
https://myaccount.blob.core.windows.net/myblobcontainer/folder1/ '
```

```
WITH (FILE_TYPE = 'CSV',
```

```
      CREDENTIAL=(IDENTITY= 'Shared Access Signature',
```

```
SECRET='<Your_SAS_Token>'),
```

```
      FIELDTERMINATOR = '|' )
```

```
-- Load as PARQUET
```

```
COPY INTO test_parquet
```

```
FROM
```

```
'https://myaccount.blob.core.windows.net/myblobcontainer/folder1/*.parquet'
```

```
WITH (CREDENTIAL=(IDENTITY= 'Shared Access Signature',
```

```
SECRET='<Your_SAS_Token>'))
```

Carga desde otros recursos

Combinación en un almacén

- Requiere nomenclatura de tres partes
- CREATE TABLE AS SELECT
- INSERT...SELECT

```
CREATE TABLE
[analysis_warehouse].[dbo].[combined_data]
AS
SELECT
FROM [sales_warehouse].[dbo].[sales_data] sales
INNER JOIN [social_lakehouse].[dbo].[social_data]
social
ON sales.[product_id] = social.[product_id];
```

```
INSERT INTO
[analysis_warehouse].[dbo].[combined_data]
SELECT
    sales.product_id
FROM [sales_warehouse].[dbo].[sales_data] sales
INNER JOIN [social_lakehouse].[dbo].[social_data]
social
ON sales.product_id = social.product_id;
```

Carga de datos con Dataflows Gen2

Ingesta y transformación con Power Query

- Conexión a orígenes de datos
- Anexar o reemplazar
- Agregar destino de datos
 - Lakehouse
 - Almacén de lago
 - Azure SQL Database
 - Azure Data Explorer (Kusto)
 - Azure Synapse Analytics (SQL DW)
- Publicar

Configuración de consulta >

▼ Propiedades

Nombre

customer-churn 1

Tipo de entidad ⓘ

Personalizado

▼ Pasos aplicados



Origen



Encabezados promovidos



ABC
123

Tipo de columna cambiado



Personalizado

▼ Destino de datos +

Sin destino de datos

Transformación de datos con Copilot

Ejemplo:

- Supongamos que tenemos una columna Gender (Género) que contiene "Male" (Hombre) y "Female" (Mujer) y queremos transformarla en entero.

The screenshot displays the Microsoft Power Query interface. The main window shows a table with columns: CustomerID, Gender, years_customer, and total_day_calls. The 'Gender' column is highlighted, and a red box encloses the 'Personalizado' (Custom) step in the 'Pasos aplicados' (Applied Steps) pane. The 'Copilot' sidebar on the right is also highlighted with a red box, showing a preview of the transformation step: 'Transform the Gender column. If Male 0, if Female 1. Then convert it to integer.'

Power Query

Buscar (Alt + Q)

Inicio Transformación Agregar columna Ver Ayuda

Obtener datos Introducir... Administrar conexiones Opciones Administrar parámetros Actualizar Propiedades Editor avanzado Agregar destino Elegir columnas Quitar columnas Conservar filas Transform... Combinar Asignar a entidad Copilot Exportar pla

Consultas [1] customer-churn 1

123 CustomerID 123 Gender 123 years_customer 123 total_day_calls

1 1716863 1 0 9

2 1282735 1 8 31

3 1135563 1 7 8

4 1782208 1 9 4

5 1136268 1 1 22

6 1945781 1 0 29

7 1786395 1 0 7

8 1590819 1 0 5

9 1402701 1 8 8

10 1223908 1 1 7

11 1202065 1 1 17

12 1980226 1 3 21

13 1484165 0 5 14

14 1233874 1 7 5

15 1311291 1 0 6

16 1621154 1 3 1

17

Configuración de... Propiedades Nombre: customer-churn 1 Tipo de entidad: Personalizado Pasos aplicados: Origen, Encabezad..., Tipo de columna, Personalizado Destino de datos: Sin destino de datos

Copilot Preview

includes sample data

- Add a step to an existing query
- Describe an existing query or steps

6:09 PM

Transform the Gender column. If Male 0, if Female 1. Then convert it to integer.

6:09 PM

Describe what you'd like to do

AI-generated content can have mistakes. Make sure it's accurate and appropriate before using it. [Review terms](#)

Se ha completado en (1.94 s) Columnas: 11 Filas: 99+

Paso

Publicar

Ejercicio

30 minutos



Cargar datos en un almacén de datos de Microsoft Fabric

Prueba de conocimientos



- 1 ¿Cuáles son las cuatro opciones de ingesta de datos disponibles en Microsoft Fabric para cargar datos en un Data Warehouse?**
 - ☒ Instrucción COPY (Transact-SQL), canalizaciones de datos, flujos de datos y almacenamiento cruzado.
 - ☐ Instrucción COPY (Transact-SQL), canalizaciones de datos, flujos de datos y Limpieza y transformación de datos.
 - ☐ Instrucción COPY (Transact-SQL), canalizaciones de datos, flujos de datos e ingesta multiplataforma.
- 2 ¿Cuáles son los orígenes de datos y los formatos de archivo admitidos para la instrucción COPY (Transact-SQL) en el almacenamiento?**
 - ☒ Azure Data Lake Storage (ADLS) Gen2 y Azure Blob Storage, con formatos de archivo PARQUET y CSV.
 - ☐ Azure Data Lake Storage (ADLS) Gen1 y Azure Blob Storage, con formatos de archivo PARQUET y CSV.
 - ☐ Azure Data Lake Storage (ADLS) Gen2 y Azure Blob Storage, con formatos de archivo ORC y CSV.
- 3 ¿Cuál es el tamaño mínimo de archivo recomendado al trabajar con datos externos en archivos de Microsoft Fabric?**
 - ☒ Al menos 4 MB.
 - ☐ Al menos 1 MB.
 - ☐ Al menos 10 MB.

Resumen

En esta sección, se ha descrito lo siguiente:

- Diferentes estrategias para cargar datos en un Data Warehouse
- Cómo crear una canalización de datos para cargar un almacén
- Cómo cargar datos en un almacén mediante T-SQL.
- Cómo cargar y transformar datos con flujos de datos Gen2.

Información adicional

Cargar datos en un Data Warehouse de Microsoft Fabric


<https://aka.ms/fabric-loaddw-spa>



Supervisión de un Data Warehouse de Microsoft Fabric



Objetivos de aprendizaje

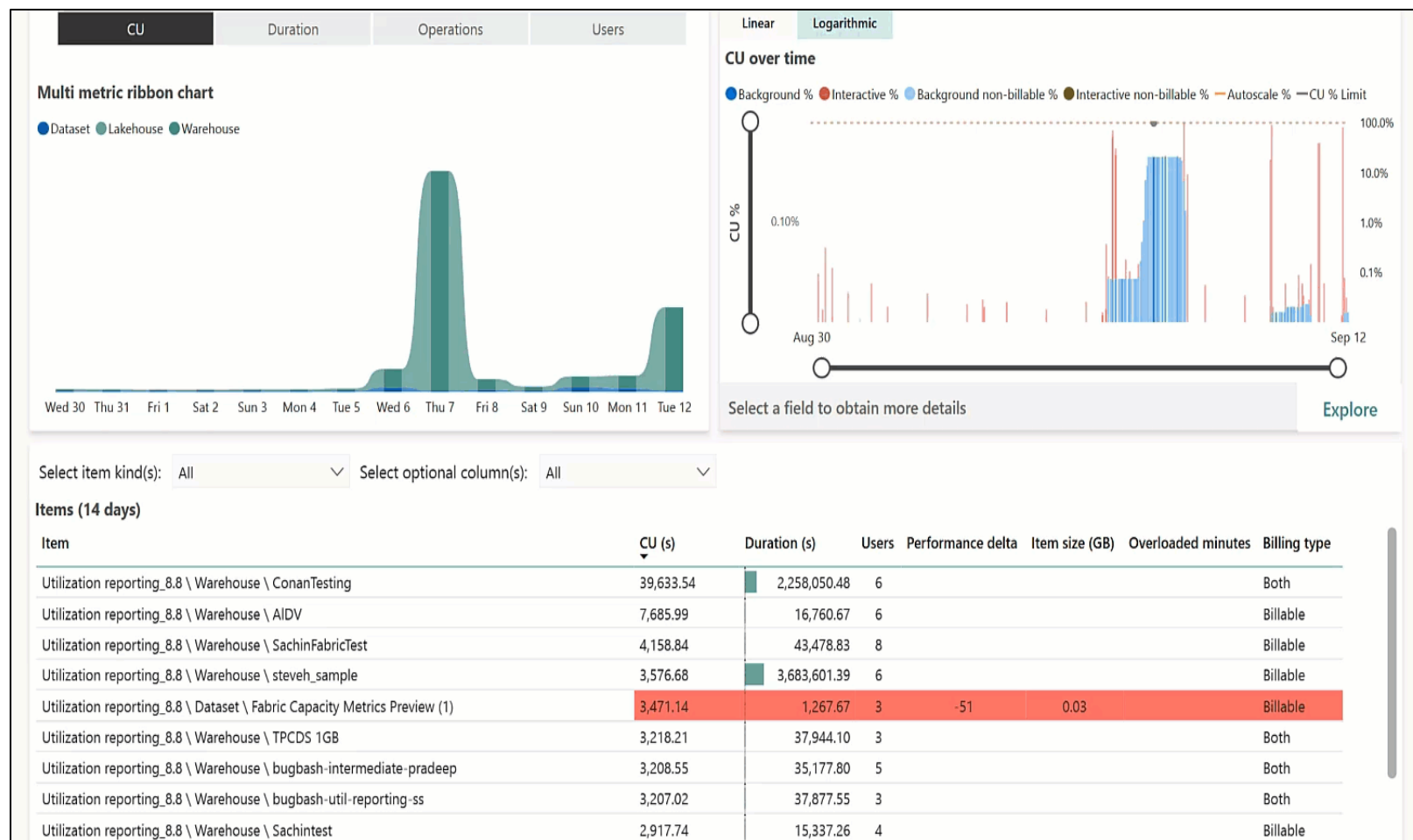


- Supervise el uso de la unidad de capacidad con la aplicación Métricas de capacidad de Microsoft Fabric.
- Supervise la actividad actual en el Data Warehouse con vistas de administración dinámica.
- Supervisión de las tendencias de consulta con vistas de información de consultas.

Supervisión de métricas de capacidad

Aplicación de métricas de capacidad de Fabric

- Supervisar y administrar costes de carga de trabajo
- Requiere instalación del administrador
- Filtrar por tipo de actividad
- Realiza un seguimiento de las unidades de capacidad (CU)
- Principalmente lecturas y escrituras



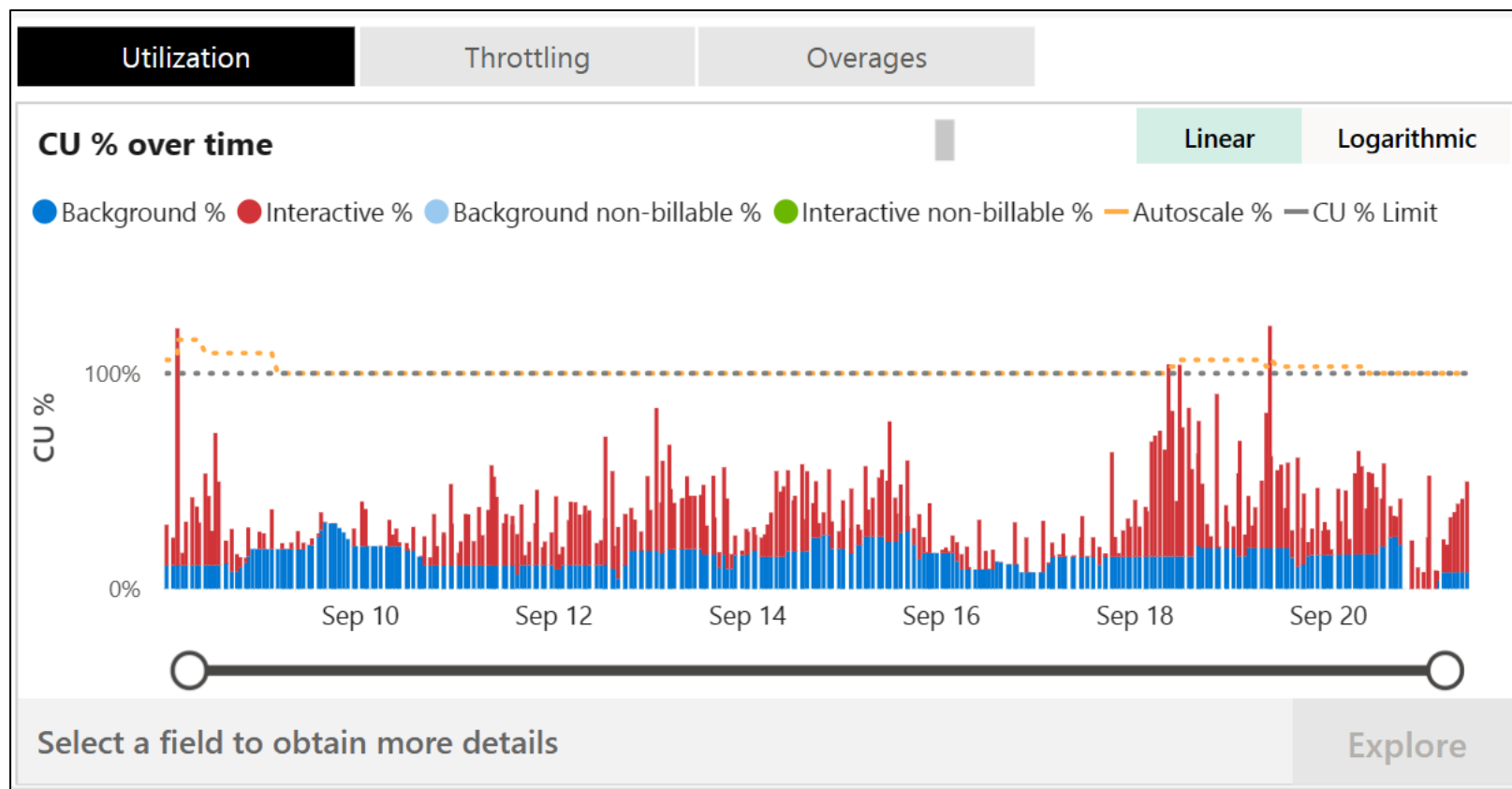
Categorías de operaciones de almacenamiento

- Consulta de almacenamiento
- Consulta de punto de conexión de análisis SQL
- Proceso de OneLake

Operation name	CU (s) ▼	Duration (s)	Users	Billing type
Warehouse Query	2,187.77	1,360.42	5	Billable
OneLake Compute	0.01	11,880.00	1	Billable
Total	2,187.78	13,240.42	6	

Operation name	CU (s) ▼	Duration (s)	Users	Billing type
SQL Endpoint Query	4,086.87	1,503.81	6	Both
Total	4,086.87	1,503.81	6	

Gráfico de exploración de puntos de tiempo



Consideraciones de uso

- Informes entre bases de datos
- Consultas de vistas de administración dinámica y de catálogo del sistema
- El campo "Duraciones" en la aplicación de métricas de capacidad de Fabric

Supervisión de la actividad actual

Vistas de administración dinámica (DMV) con T-SQL

- **sys.dm_exec_connections:**
Conexiones entre el almacén y el motor
- **sys.dm_exec_sessions:**
Sesiones entre elemento y motor
- **sys.dm_exec_requests:**
Solicitudes activas en una sesión

```
-- Identify long-running queries
SELECT request_id, session_id, start_time,
total_elapsed_time
FROM sys.dm_exec_requests
WHERE status = 'running'
ORDER BY total_elapsed_time DESC;

-- Find session_id (user)
SELECT login_name
FROM sys.dm_exec_sessions
WHERE 'session_id' = 'SESSION_ID WITH LONG-
RUNNING QUERY';

-- Terminate session (ADMIN only)
KILL 'SESSION_ID WITH LONG-RUNNING QUERY';
```

Vistas de información de consulta

- **queryinsights.exec_requests_history:**
consultas SQL completadas
- **queryinsights.long_running_queries:**
tiempos de ejecución de la consulta
- **queryinsights.frequently_run_queries:**
consultas que se ejecutan con frecuencia

--- Queries from previous hour

```
SELECT start_time, login_name, command
FROM queryinsights.exec_requests_history
WHERE start_time >= DATEADD(MINUTE, -60,
GETUTCDATE())
```

--- Long-running queries, multiple runs

```
SELECT last_run_command, number_of_runs,
median_total_elapsed_time_ms,
last_run_start_time
FROM queryinsights.long_running_queries
WHERE number_of_runs > 1
ORDER BY median_total_elapsed_time_ms DESC;
```

-- Frequently run queries

```
SELECT last_run_command, number_of_runs,
number_of_successful_runs,
number_of_failed_runs
FROM queryinsights.frequently_run_queries
ORDER BY number_of_runs DESC;
```

Ejercicio

30 minutos



Supervisión de Data Warehouses en Microsoft Fabric

Prueba de conocimientos



1 Quiere supervisar el consumo de unidades de capacidad en el Data Warehouse de Fabric. ¿Qué herramientas debe usar?

- ☐ Microsoft Azure Monitor
- ☒ Aplicación Métricas de capacidad de Microsoft Fabric
- ☐ Microsoft Azure Data Studio.

2 ¿Qué vista de administración dinámica proporciona detalles sobre los comandos SQL que se ejecutan en el Data Warehouse?

- ☒ sys.dm_exec_requests
- ☐ sys.dm_exec_connections
- ☐ sys.dm_exec_sessions

3 ¿Qué vista debes usar para identificar los comandos ejecutados habitualmente en el Data Warehouse?

- ☐ queryinsights.exec_requests_history
- ☐ queryinsights.long_running_queries
- ☒ queryinsights.frequently_run_queries

Resumen

En esta sección, se ha descrito lo siguiente:

- La aplicación Métricas de capacidad de Microsoft Fabric.
- Las vistas de administración dinámica.
- Las vistas de información de consulta.

Información adicional

Supervisión de un Data Warehouse de Microsoft Fabric

<https://aka.ms/fabric-monitordw-spa>



Protección de un Data Warehouse de Microsoft Fabric



Objetivos de aprendizaje



- Obtención de información sobre los conceptos de protección de un Data Warehouse en Microsoft Fabric.
- Implementación del enmascaramiento dinámico de datos, la seguridad a nivel de fila y a nivel de columna.
- Configuración de permisos granulares mediante T-SQL

Enmascaramiento dinámico de datos

- Enmascara los datos para evitar que los usuarios sin privilegios vean información confidenciales
- Los administradores especifican qué datos se van a revelar
- Se aplica cuando se devuelven los resultados de la consulta: sin cambios en los datos
- Hay varias funciones de enmascaramiento disponibles para diferentes categorías de información confidencial

Table.CreditCardNo
4465-6571-7868-5796
4468-7746-3848-1978
4484-5434-6858-6550

*Enmascaramiento de datos
en tiempo real*

CreditCardNo
XXXX-XXXX-XXXX-5796
XXXX-XXXX-XXXX-1978

Reglas de enmascaramiento

Tipo de enmascaramiento	Descripción	Caso de uso	Limitaciones	Regla de enmascaramiento
Valor predeterminado	Enmascaramiento completo de acuerdo con los tipos de datos de los campos designados.	Resulta útil cuando desea ocultar completamente los datos reales.	Enmascara completamente los datos. No hay información visible.	Default()
Correo electrónico	Expone la primera letra de una dirección de correo electrónico y el sufijo constante ".com"	Resulta útil cuando desea mostrar que el campo de datos contiene un correo electrónico sin revelar el correo electrónico real.	Solo se aplica a los campos de correo electrónico	Email()
Texto personalizado	Expone el primer y el último carácter "n" y agrega una cadena de relleno personalizada en el medio.	Resulta útil cuando desea ocultar parcialmente los datos reales.	No es adecuado para los tipos de datos numéricos, de fecha y hora.	Partial(prefix_padding,padding_string,suffix_padding)
Aleatorio	Reemplaza cualquier valor numérico o binario por un número aleatorio dentro de un intervalo especificado.	Resulta útil cuando desea ocultar los datos numéricos o binarios reales.	Solo se aplica a los tipos de datos numéricos y binarios.	Random(low, high)

Configuración del enmascaramiento de datos dinámicos

1. CREATE o ALTER TABLE con ADD MASK
2. GRANT UNMASK para usuarios con privilegios
3. Los usuarios sin privilegios ven los datos enmascarados
4. Los usuarios o roles administrativos y con privilegios ven datos sin máscara

Agregación de máscaras mediante T-SQL

```
ALTER TABLE Employee ALTER COLUMN Email  
ADD MASKED WITH (FUNCTION = 'email()')  
  
ALTER TABLE Employee ALTER COLUMN Salary  
ADD MASKED WITH (FUNCTION = 'random(1,200000)')  
  
GRANT UNMASK ON Employee  
TO [kim2@adventure-works.com]
```

```
SELECT Email, Salary  
FROM Employee
```

El inicio de sesión de kim2@adventure-works.com ve:

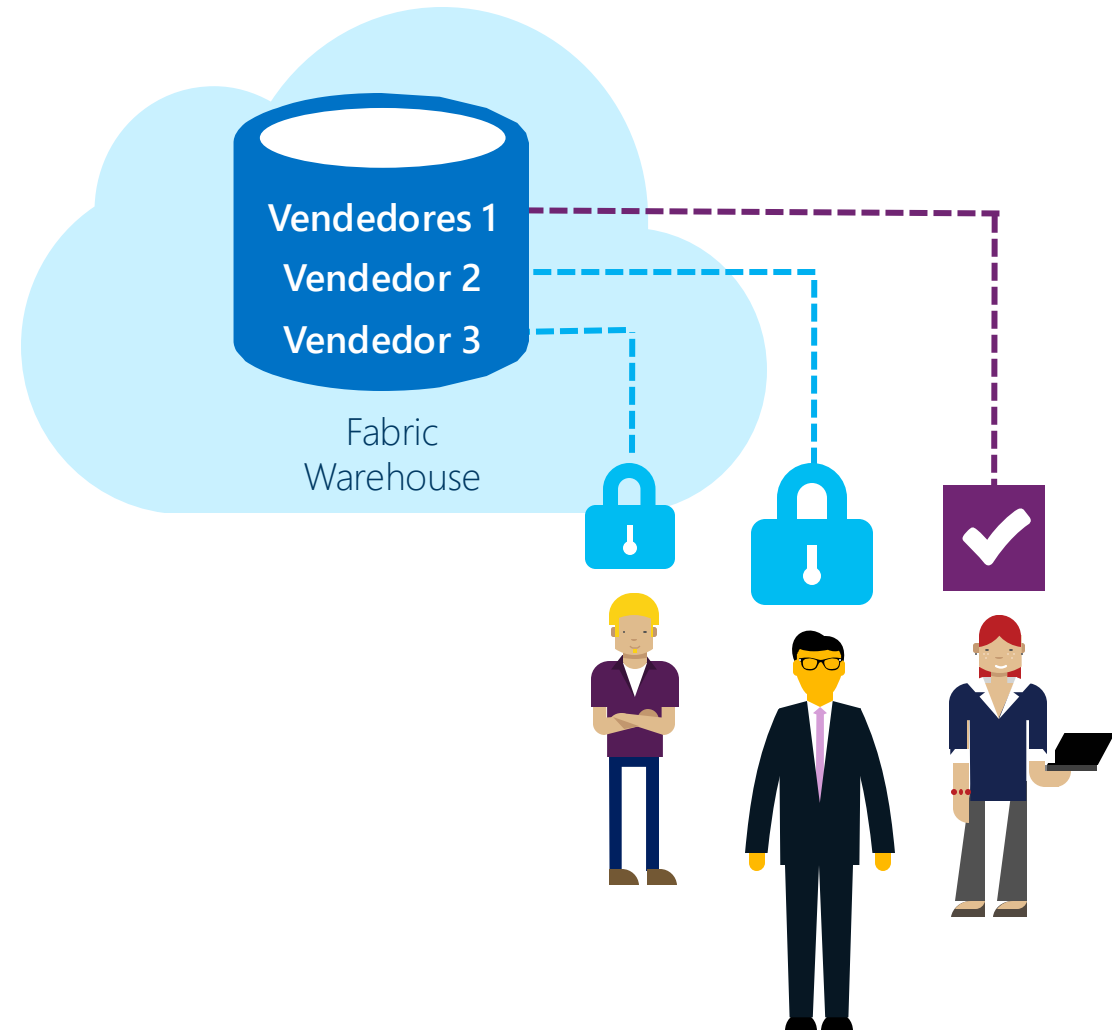
	ABC Email ↑	123 Salary
1	catherine0@adventure-works.com	10000
2	frances0@adventure-works.com	20000

Otros inicios de sesión ven:

	ABC Email	123 Salary
1	c)00X(@)000X.com	13156
2	f00X(@)000X.com	198291

Seguridad de nivel de fila

- Te permite controlar qué usuarios pueden acceder a filas de datos específicas
- El acceso a datos está restringido en función de la identidad del usuario
- Las filas no se devuelven a los usuarios que no tienen permiso para verlas
- No cambia los datos del almacén
- Caso de uso de ejemplo: los vendedores solo deben ver sus pedidos, no los pedidos de todos los vendedores



Implementación de la seguridad de nivel de fila

1. Determinación de roles y criterios para el acceso a datos

2. Creación de una **función de predicado**

Determina los criterios para el acceso a filas de usuario

Contiene lógica de seguridad y **predicado de seguridad**

3. Creación de una **directiva de seguridad**

Invoca la función predicado

Enlaza la función de predicado a una tabla

Ejemplo de seguridad de nivel de fila

Escenario: los representantes de ventas solo pueden ver las filas asociadas a sus pedidos en una tabla Ventas. No pueden ver las filas colocadas por otros representantes de ventas.

Función

Evalúa quién está consultando la tabla. Devuelve 1 (true) cuando un valor de la columna SalesRep de una fila es el mismo que el usuario que ejecuta una consulta.

```
CREATE FUNCTION rls.fn_securitypredicate(@SalesRep AS VARCHAR(60))
    RETURNS TABLE
WITH SCHEMABINDING
AS
    RETURN SELECT 1 AS fn_securitypredicate_result
WHERE @SalesRep = USER_NAME();
```

Directiva de seguridad

Asocia la función a la tabla Ventas e invoca la función cuando se consulta la tabla Ventas. Filtran en modo silencioso las filas disponibles para leer operaciones (SELECT, UPDATE y DELETE).

```
CREATE SECURITY POLICY SalesFilter
ADD FILTER PREDICATE rls.fn_securitypredicate(SalesRep)
ON dbo.Sales
WITH (STATE = ON);
```

Seguridad a nivel de columna

- Restringe el acceso a columnas en una tabla
- Se implementa mediante los permisos GRANT SELECT y DENY SELECT

Ejemplo

Permitir que [mary@contoso.com] vea todas las columnas de una tabla Pedidos, excepto la columna CreditCard.

```
GRANT SELECT ON Orders TO [mary@contoso.com];  
DENY SELECT ON Orders (CreditCard) TO [mary@contoso.com];
```

```
SELECT *  
FROM Orders;
```

mary@contoso.com ve:

Error: se ha denegado el permiso SELECT en la columna "CreditCard" del objeto "Orders", la base de datos "Sales", el esquema "dbo"

Los usuarios no restringidos ven:

	123 OrderID	123 CustomerID ↑	ABC CreditCard
1	1234	5678	1111111111111111
2	2341	6785	2222222222222222
3	3412	7856	3333333333333333

Permisos granulares mediante T-SQL

- Control de acceso específico a los elementos protegibles del almacén
Tablas, vistas, procedimientos
- Usa comandos del lenguaje de control de datos de SQL
GRANT, DENY, REVOKE
- Se usa cuando los roles de espacio de trabajo de Fabric y los permisos de elemento no son lo suficientemente específicos

```
DENYSELECT on dbo.Parts to [RoleName];
```

```
GRANTEXECUTE on dbo.sp_UpdateParts to [RoleName];
```

Ejercicio

30 minutos



Asegura un almacenamiento en Microsoft Fabric

Prueba de conocimientos



- 1** ¿Cuál es la principal ventaja del enmascaramiento dinámico de datos (DDM)?
 - ☒ Limita la exposición de datos al ocultar información confidencial en tiempo real.
 - ☐ Cambia los datos reales de la base de datos
 - ☐ Requiere codificación compleja para implementar.

- 2** ¿Cuál es el propósito de una función de predicado de seguridad en seguridad de nivel de fila (RLS)?
 - ☒ Determina si una fila es accesible para un usuario en función de determinados criterios.
 - ☐ Habilita las conversiones de tipos en funciones de predicado.
 - ☐ Permite a los usuarios ejecutar consultas ad hoc

- 3** ¿Qué ocurre cuando se concede un permiso a un usuario y después se le deniega el mismo permiso de un almacén?
 - ☐ La instrucción GRANT reemplaza a DENY y el usuario tendrá acceso al objeto específico.
 - ☒ DENY siempre reemplaza a GRANT y se deniega el acceso del usuario al objeto específico.
 - ☐ El usuario tiene ambos permisos y provoca un conflicto.

Resumen

En esta sección, se ha descrito lo siguiente:

- El enmascaramiento dinámico de datos (DDM).
- La seguridad a nivel de fila (RLS).
- La seguridad a nivel de columna (CLS).
- Los permisos granulares de SQL.

Información adicional

Protección de un Data Warehouse de Microsoft Fabric

<https://aka.ms/fabric-securedw-spa>

