

Parcial Práctico

BASES DE DATOS APLICADA

**Docente**: Rivas Sarquis Martin Miguel

**Alumnos:** Gomez, Lucas Mauricio

Lambezat, Juan Ignacio

Elebi, Yamil Matias(PM)

**Sede:** Lomas de Zamora

**Comisión:** 3-A **Turno:** Mañana

**Año:** 2024

# Índice

[**Índice 2**](#_9rbxuc9z7hzm)

[**Bitácora de modificaciones 2**](#_dzbwrc5858b3)

[**Introducción 3**](#_dhiawyl4g6ck)

[**Importación de BD 3**](#_4nz5v6t2isfp)

[**Creación DW y Diagrama Estrella 5**](#_zeew8okm4qu9)

[**Creación de Vistas 6**](#_piorve6eo9a2)

[**Servicios Utilizados 7**](#_4130c7vb1nc8)

[**Creación E.T.L. (Integration Service) 7**](#_23mlzdagyvkq)

[**Poblar DW 11**](#_lsva5zwxxn7w)

[**Creación del Cubo (Analysis Service) 11**](#_s1s8sv3chj01)

[**Uso de Power BI 12**](#_uaz1d6axtce1)

# Bitácora de modificaciones

| FECHA 18/10 | Creación de Diagrama Estrella, Creación de sakilaDW, Creación de E.T.L. y Ejecución. |
| --- | --- |
| FECHA 20/10  FECHA 21/10 | Creación del Cubo, errores en implementación (permisos). |
| FECHA 22/10 | Proyecto funcional completamente. |

# 

# 

# 

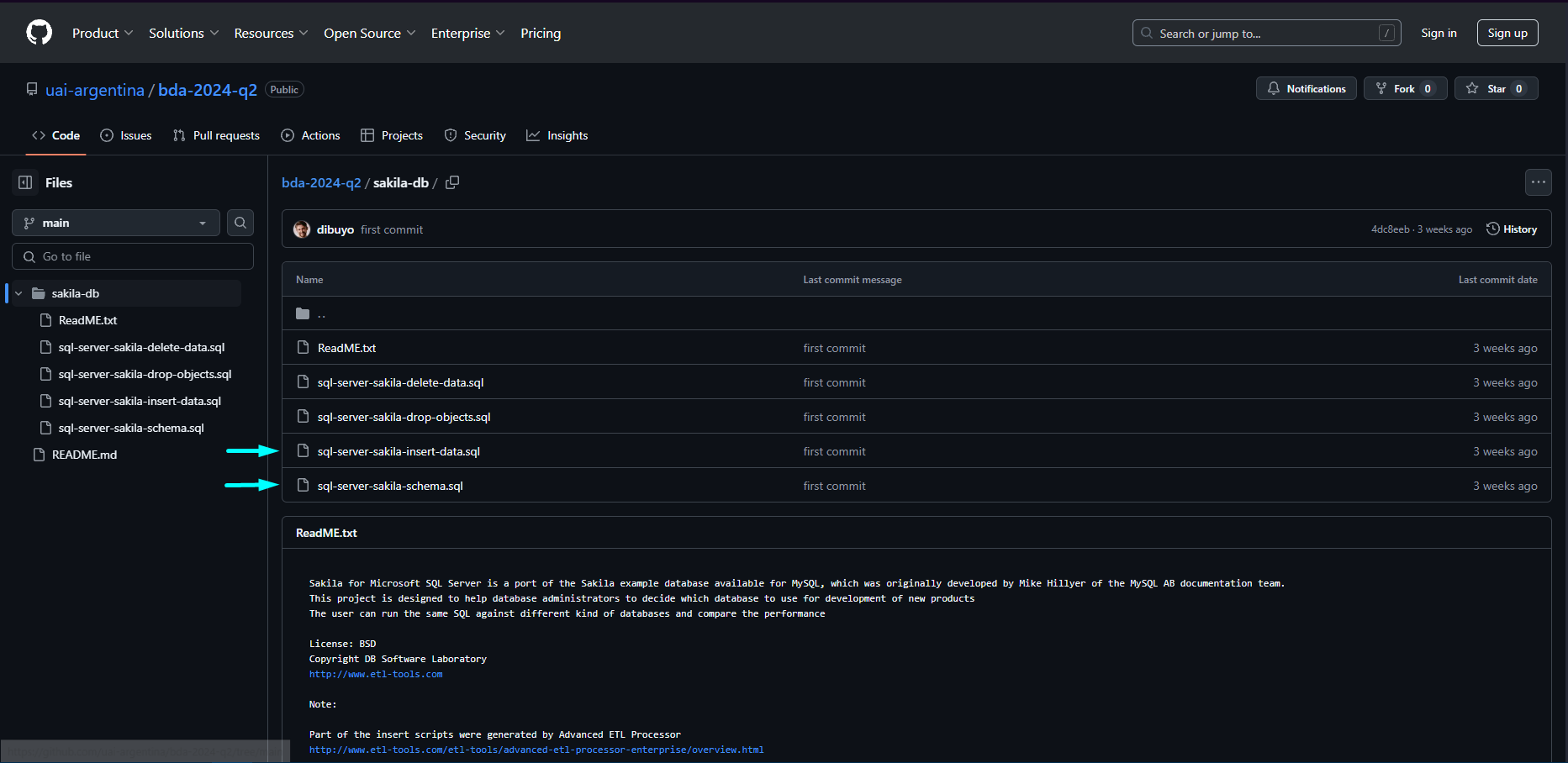
# Introducción

El documento es sobre la práctica realizada para el primer parcial domiciliario sobre la base de datos Sakila para la creación de un Data Warehouse, su población mediante E.T.L. y posterior análisis mediante un cubo y Power BI.

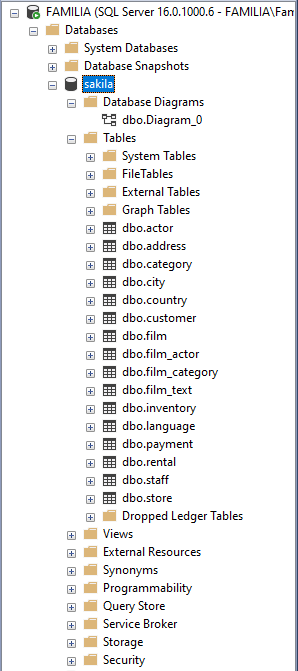
# Importación de BD

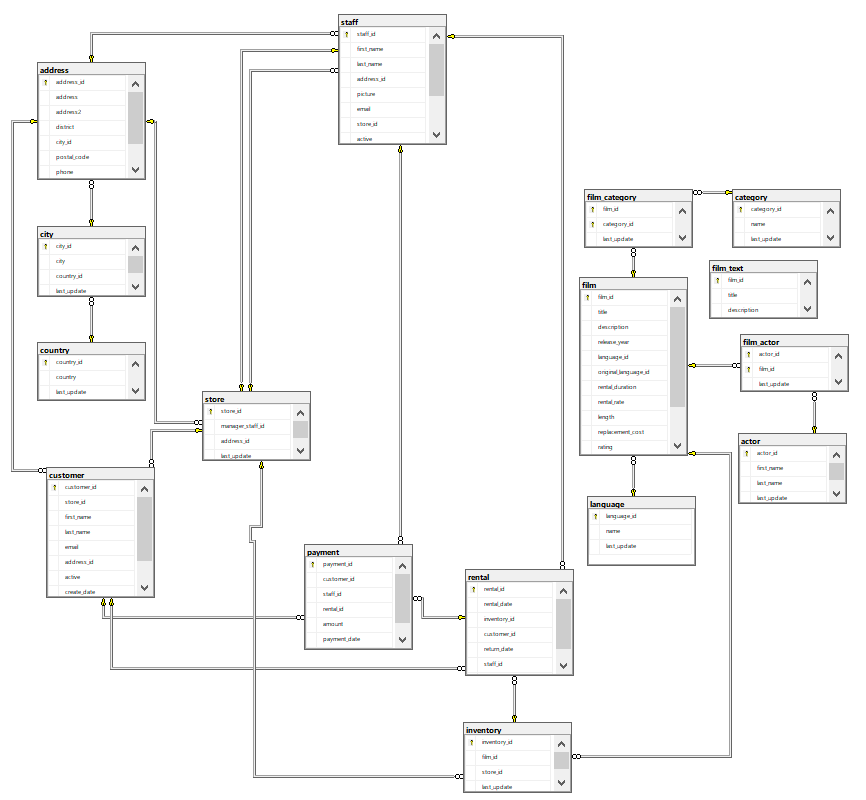
Ingresamos al siguiente link y descargamos los Script para SQL Server, usamos el de Schema que contiene toda la estructura de la base de datos Sakila y luego el de insert data para darle datos a todas las tablas de la base de datos.

<https://github.com/uai-argentina/bda-2024-q2/tree/main/sakila-db>



Al finalizar con los Scripts nos deja la BD Sakila con sus debidas tablas, con las que hacemos un diagrama automático para ver sus relaciones de manera más sencilla y será de utilidad para entender esta BD y poder crear un DW coherente.



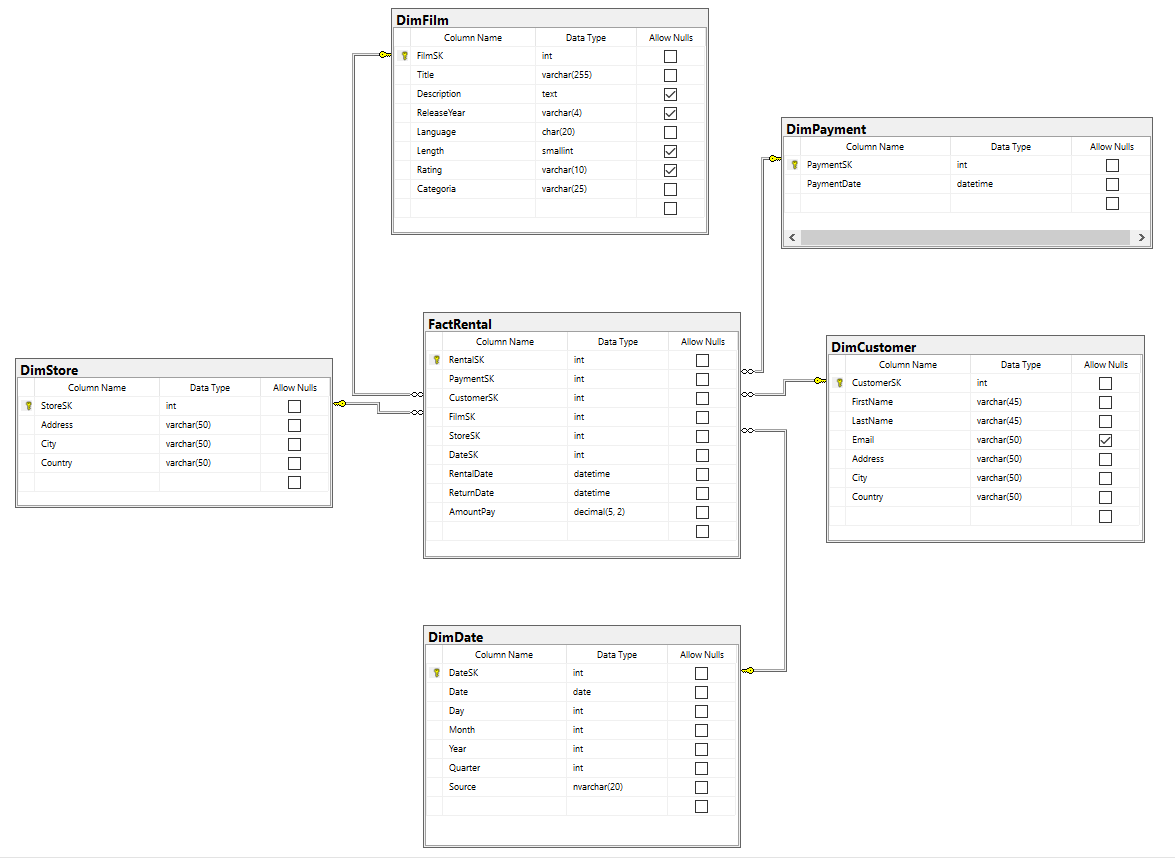


# Creación DW y Diagrama Estrella

Para la creación del diagrama creamos una nueva base de datos llamada “sakilaDW” que como su nombre indica será el data warehouse ya que este diagrama es el primer paso para el mismo.

Elegimos el diagrama estrella ya que proporciona simplicidad, velocidad y compatibilidad con herramientas BI, además de ser fácil de entender y escalar.

Nuestro diagrama se reparte en cinco dimensiones y una tabla de hechos. Siempre respetando los valores de los atributos que vendrán de “sakila” para evitar problemas a futuro.

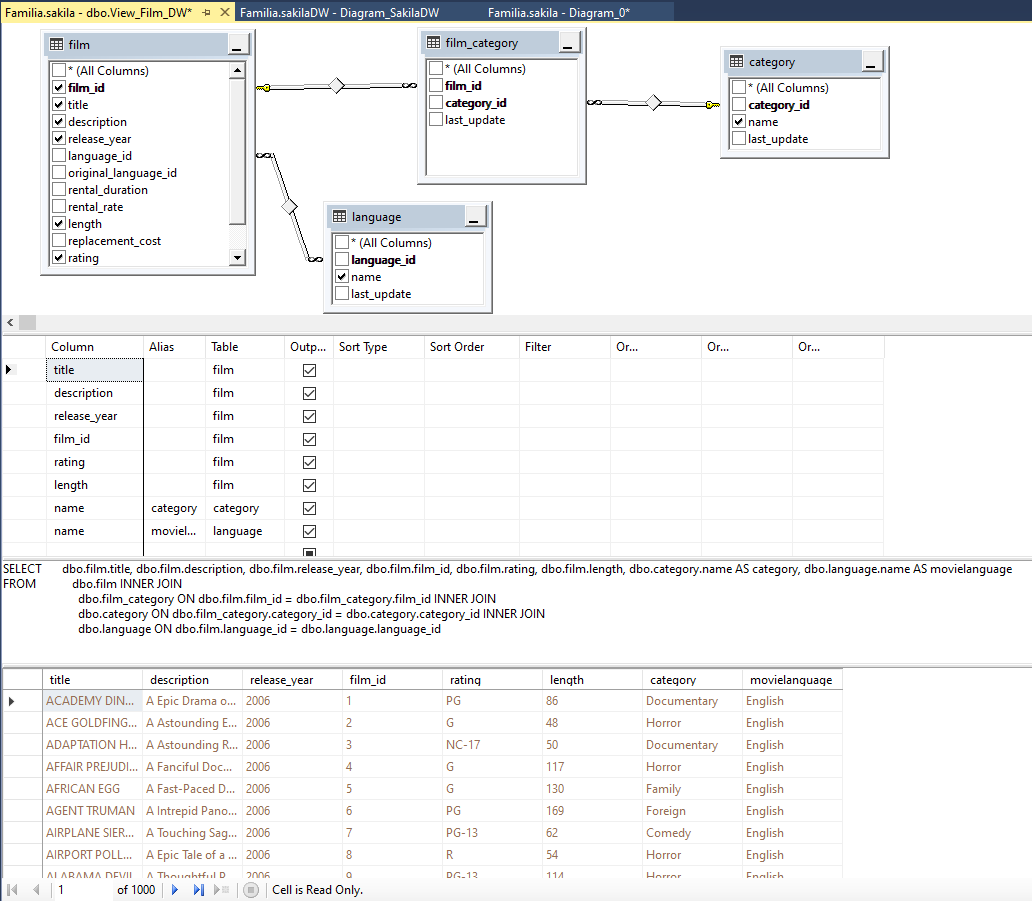


# Creación de Vistas

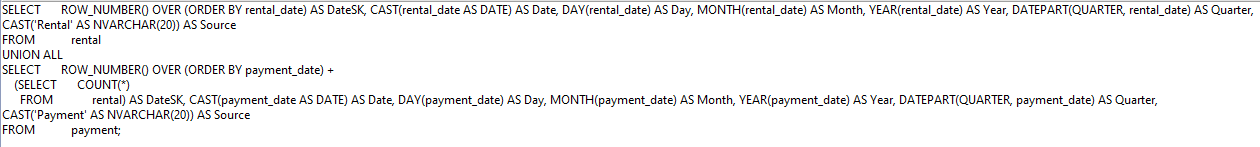
Las vistas fueron creadas en el origen de los datos, es decir, en “sakila” donde seleccionamos las tablas con los atributos que necesitamos para las dimensiones de nuestro data warehouse.

Fueron creadas mediante querys y el apartado de “New View”.

Ejemplo de vista creada con “New View” para la dimensión Film:



Ejemplo de vista creada con Query para la dimensión Time:



# Servicios Utilizados

Continuando con nuestro DW necesitamos nuevos servicios para poder poblar nuestro DW mediante ETL y Cubo para su posterior análisis además de Power BI para explotar los datos del Cubo.

En los siguientes links conseguimos dichos servicios.

**Analysis Service**

**Visual Studio 2022**

https://marketplace.visualstudio.com/items?itemName=ProBITools.MicrosoftAnalysisServicesModelingProjects2022

**Integration Service**

**Visual Studio 2022**

https://marketplace.visualstudio.com/items?itemName=SSIS.MicrosoftDataToolsIntegrationServices

**SQL Server Data Tools**

https://learn.microsoft.com/en-us/sql/ssdt/download-sql-server-data-tools-ssdt?view=sql-server-ver15

**Power BI Desktop**

https://www.microsoft.com/en-us/download/details.aspx?id=58494

# Creación E.T.L. (Integration Service)

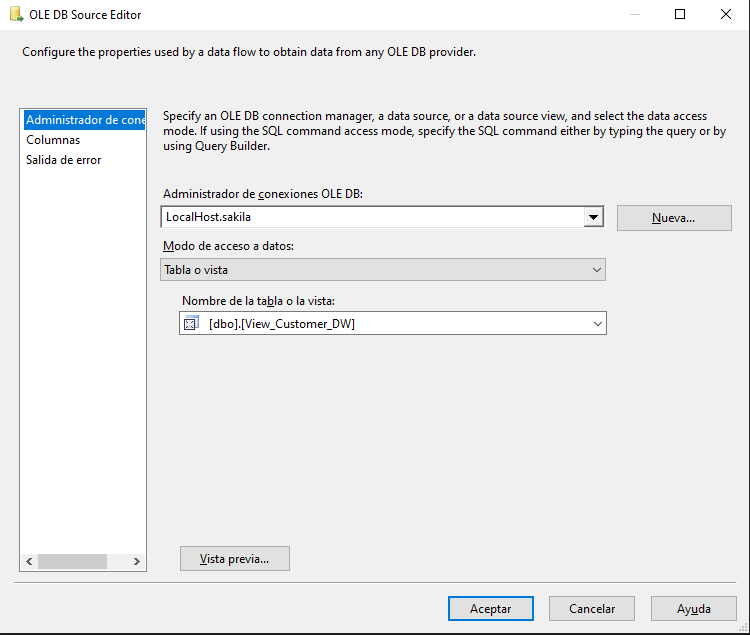
Ejecutamos Visual Studio 2022 y al momento de crear un nuevo proyecto buscamos la plantilla “Integration Services Project”.

Proceso de creación de ETL:

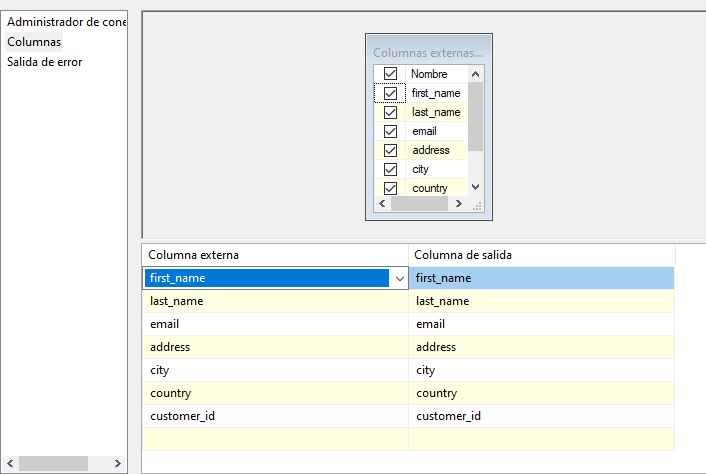
Seleccionamos en herramientas “Data Flow Task” en el apartado de Flujo de control e ingresamos con doble click en Data Flow Task, nos llevará al apartado de flujo de datos donde ingresamos un OLE DB Source y un OLE DB Destination conectandolos mediante la flecha azul.

Como indica su nombre uno será para la fuente origen de datos y otro para el destino de estos datos

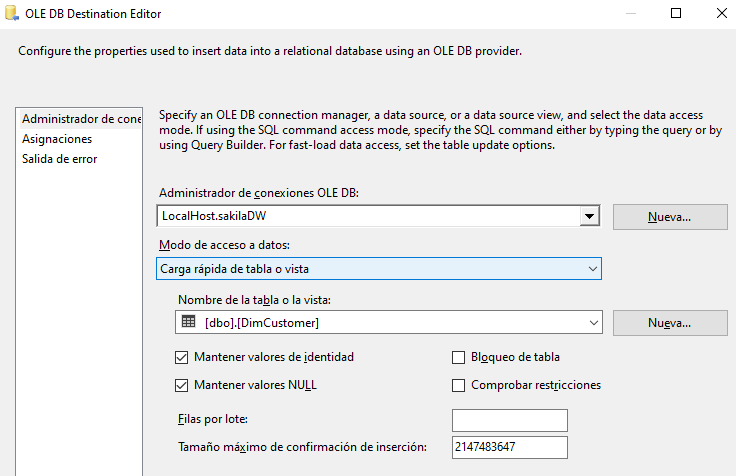
Este ejemplo con el origen de datos de Customer, .donde seleccionamos cual es el origen y la forma de acceder, en nuestro caso es con vistas y seleccionamos la vista a usar.



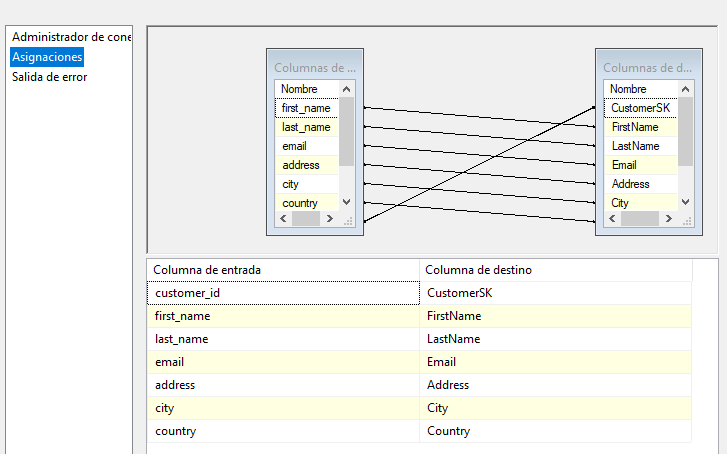
Luego en Columnas seleccionamos los campos necesarios, en caso de vistas serían todos porque es algo más personalizado, pero en caso de usar tablas debemos seleccionar las necesarias y asociarlas a los campos.



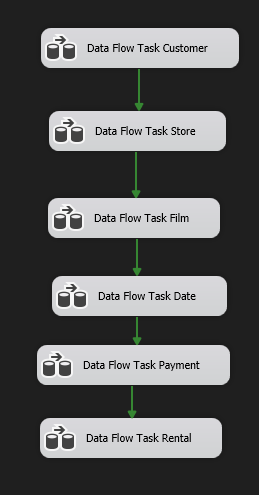
Continuamos con el Destination que reacciona de manera similar. Seleccionamos el lugar que recibirá estos datos previamente seleccionados donde usamos una carga rápida de tabla o vista y seleccionamos la dimensión a poblar. Aseguramos de condiciones como mantener valores NULL para evitar errores.



En asignaciones aseguramos que cada atributo sea correcto.



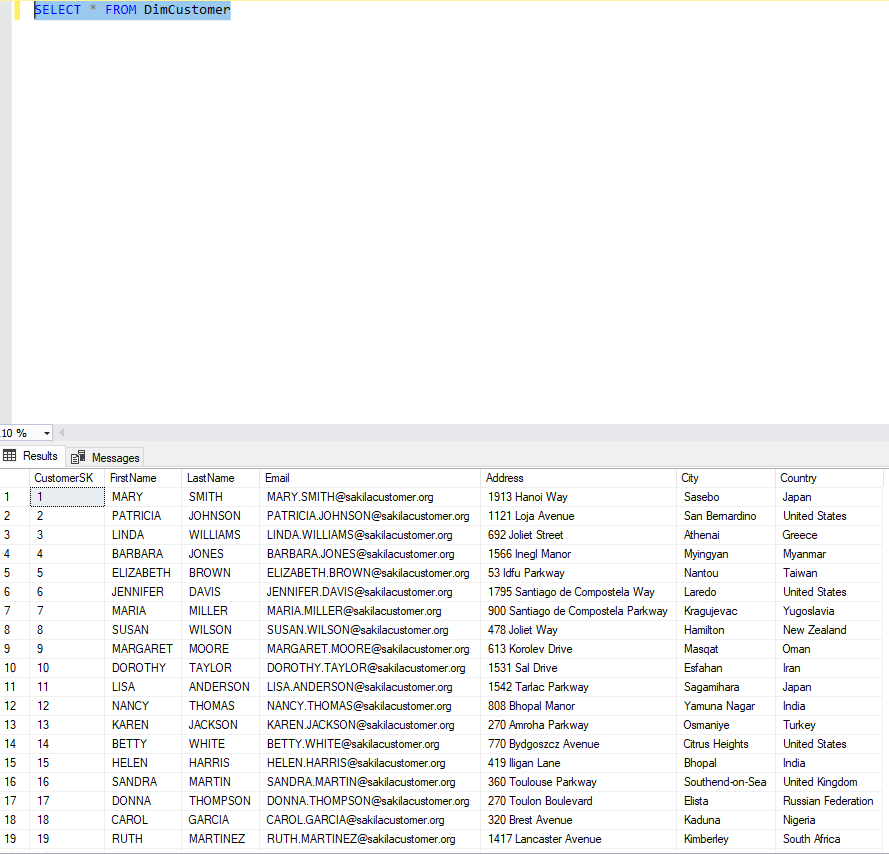
Una vez terminado el proceso con todas las dimensiones, relacionamos todos los data flow task para que se ejecuten en determinado orden.



# Poblar DW

Terminada la creación de los ETL, los ejecutamos e irán poblando las dimensiones del DW según lo planeado.

Con un Select general sobre la tabla veremos si los datos fueron correctamente poblados.



# Creación del Cubo (Analysis Service)

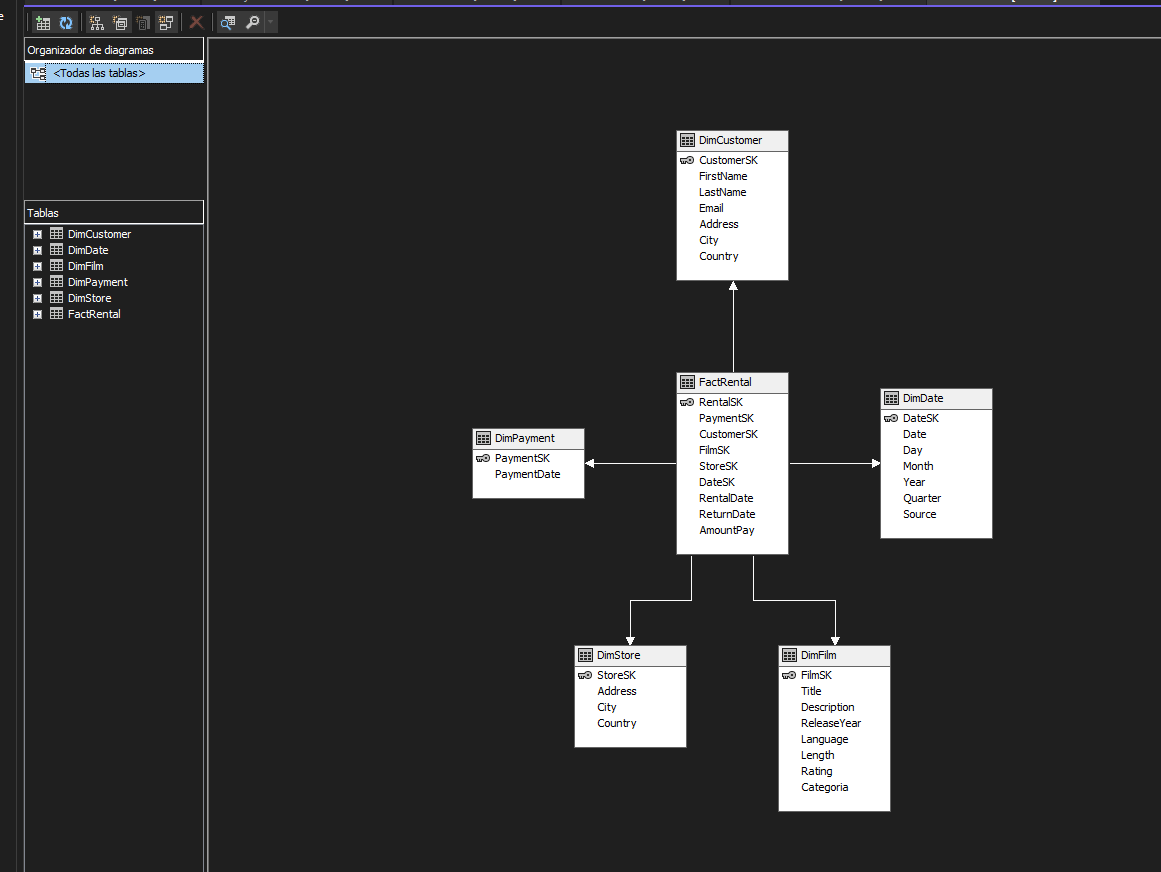
Al momento de crear un nuevo proyecto de visual studio 2022 usamos la plantilla de proyecto multidimensional de Analysis Services.

Tenemos tres etapas para concretar:

La primera es seleccionar el origen de los datos, es decir, el DW que creamos anteriormente. Hay que asegurarse de tener el servidor correcto y los permisos necesarios para evitar problemas al querer crear el cubo.

La segunda es crear una vista del origen de datos, resulta en el mismo diagrama que tenemos

Por último organizar los atributos de las dimensiones.



Si la ejecución es correcta dirá que la implementación fue exitosa. Por lo que nuestro cubo existe y puede ser ejecutado por Power BI.

# Uso de Power BI

Ingresamos al apartado de obtener datos de otro origen y buscamos SQL Server Analysis Services y seleccionamos el Servidor y BD que están en el modo Analysis Service de SQL Server. Asegurarse de dar los permisos de usuario para evitar errores.

Al ser correcta podremos utilizar el cubo creado y generar reportes / gráficos en base a ello.

Ejemplos:

