

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA  
FACULTAD DE INGENIERÍA  
ESCUELA DE CIENCIAS Y SISTEMAS  
SEMINARIO DE SISTEMAS 2  
ING. LUIS ALBERTO VETTORAZZI ESPANA  
AUX. ESCARLETH ANDREA VELASCO CAMPOS



# Proyecto 1

## SG-FOOD

Jose Fernando Alvarez Morales

201900528

### Descripción:

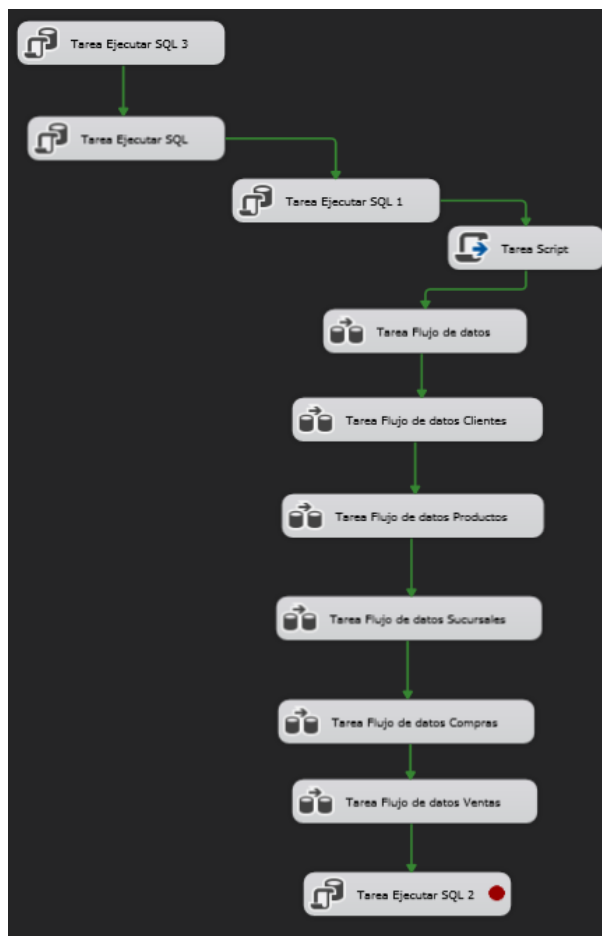
Se solicitan una solución que resuelva los problemas anteriormente planteados, esto como fin principal para la generación de reportes sobre las compras y las ventas de la empresa con la facilidad de generar estos en intervalos de tiempo según sean las necesidades que se presenten.

### Requerimientos del sistema para la creación:

- Windows 11 Home
- SQL Server 2019
- Visual Studio Code 2019
- SSIS 2019
- MySQL 8.0.32
- DataGrip 2022.3.3
- RAM 12GB
- Intel™ Core™ i7-8565U CPU @ 1.80GHz 1.99 GHz

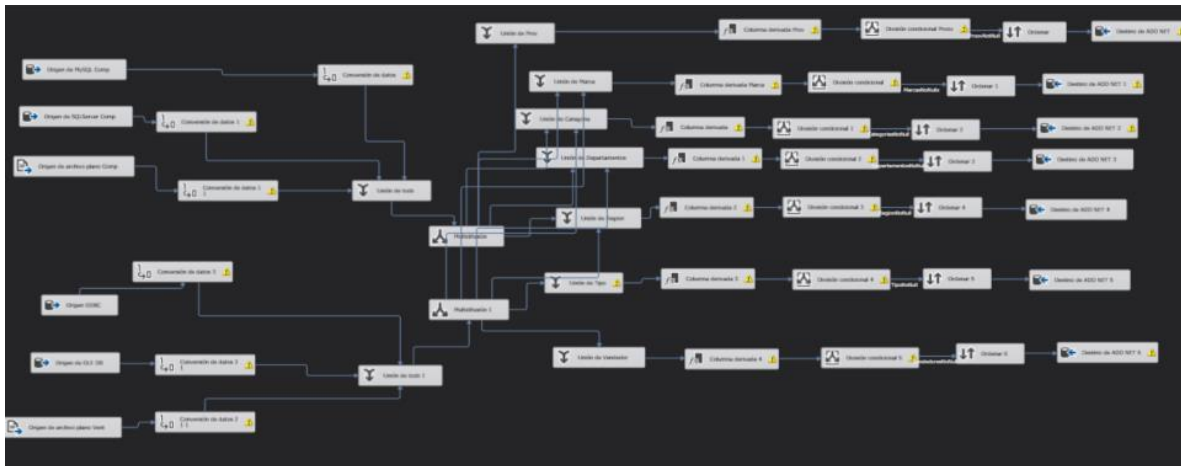
### Descripción de las fases de proceso ETL

Para empezar el modelo del Control Flow es este:

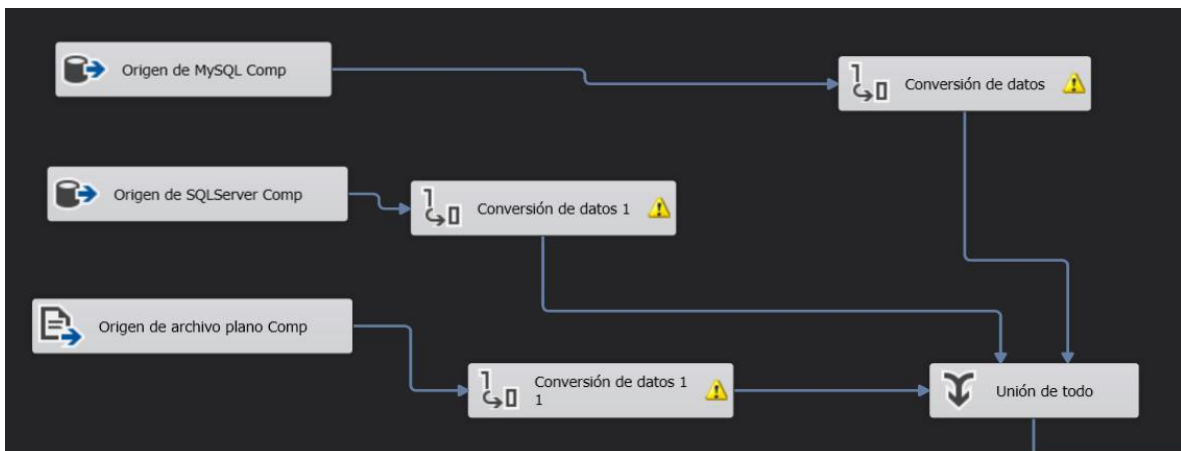


En las primeras dos instrucciones “Tarea Ejecutar SQL” lo que se realizó es una conexión con nuestras bases de datos pivotes, donde almacenaremos nuestra información temporalmente, y le haremos una limpieza de datos, incluso a nuestra datawarehouse para asegurarnos que estos no tengan ningún dato previamente.

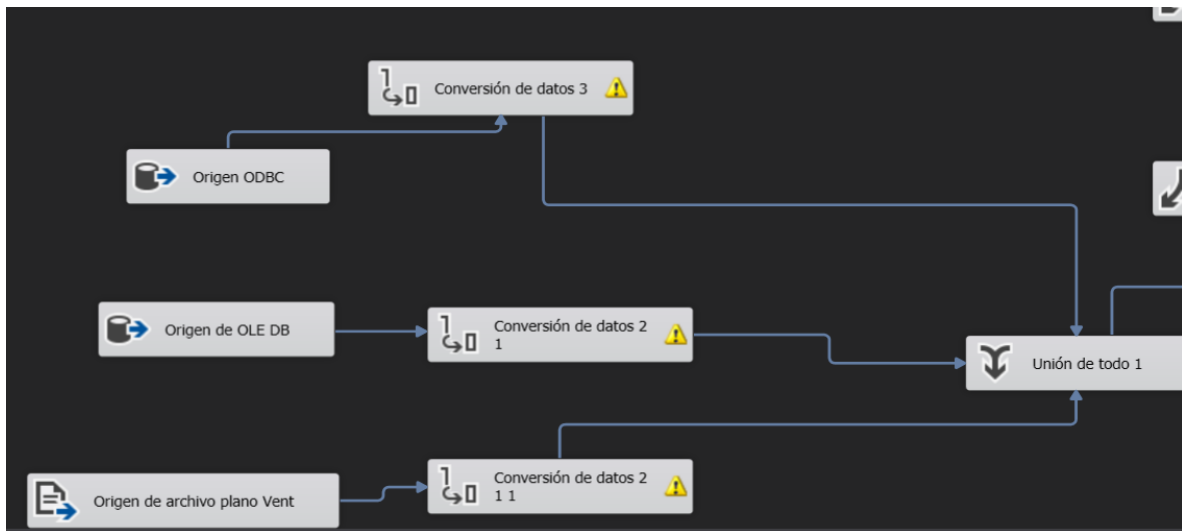
Luego en las siguientes son varios “DataFlow” donde se realizara varias acciones que se explicaran a continuación. Se realizó 1 “DataFlow” para la carga de datos en las tablas independiente, luego 1 “DataFlow” para poder llenar las tablas de Clientes, Productos, Sucursales, Compras y Ventas.



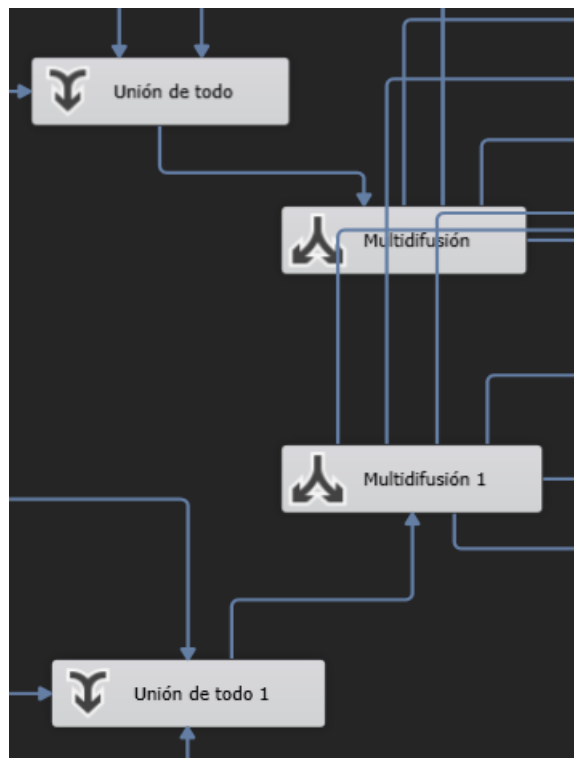
Este es el modelo que se encuentra en el primer “DataFlow” iremos explicando paso a paso.



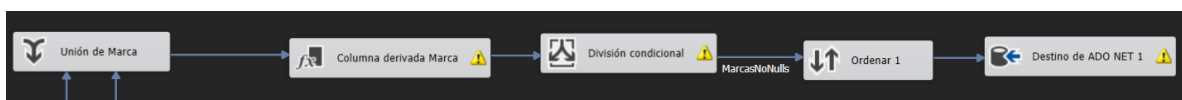
En esta primera parte se realiza la carga de datos que almacenamos en nuestras bases de datos pivote, ya que en dichas tablas aun no se encuentran con los correctos tipos los datos se deben convertir y para se utiliza la Conversion de Datos, luego se juntan todos los resultados, para poder manejarlas de una manera mas sencilla.



En el lado de la ventas se realiza de la misma manera, se obtienen los datos de las tablas pivote y del archivo plano para luego convertir los datos y luego los juntamos.

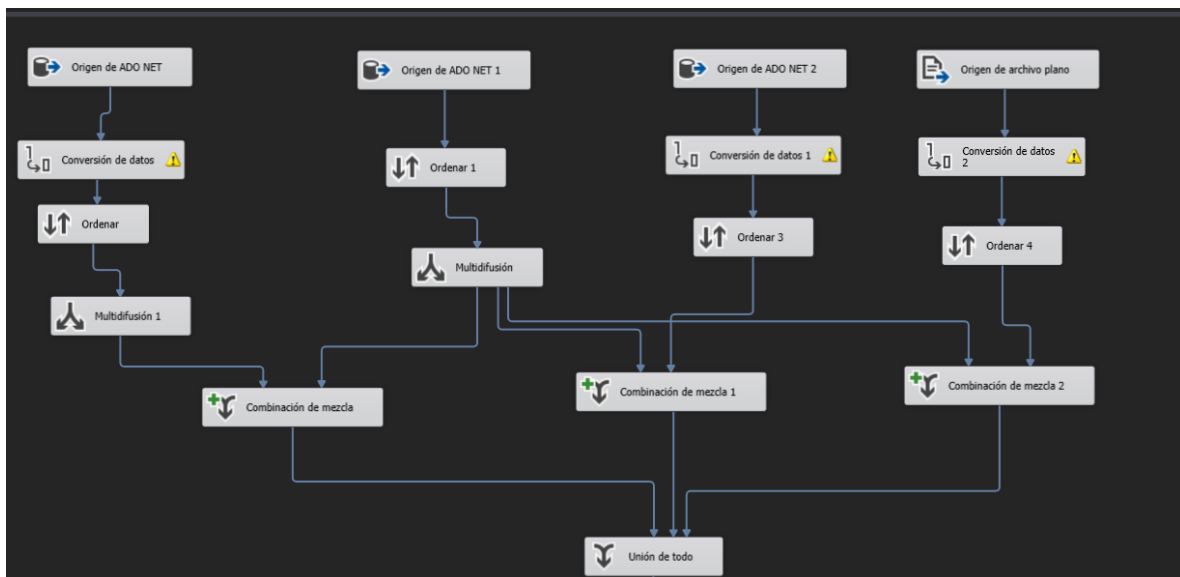


En estas instrucciones lo que se hace es dividir la información que se junto, ya que algunos datos que se encuentran los archivos de compra también se encuentran en los archivos de ventas, por lo que debemos de poder tener una forma de obtener los datos de ambos lados y para eso se utilizó la Multidifusión.



En esta serie de instrucciones como podemos observar, se envían los datos que se encuentran en el lado de ventas y de compras, luego se realizan validaciones para poder eliminar los espacios que se encuentran nulos, ya que esos datos no nos servirán. También en el ordenar del final se procede a eliminar los datos que estén repetidos y solamente nos aparecer 1 sola vez, para finalmente en el Destino de ADO NET poder guardar la información en la debida tabla de nuestro datawarehouse. Y así se hace con los demás tablas independientes.

Para los Data Flows donde se llenan las otras tablas se utiliza la misma idea de obtener los datos de nuestras tablas pivote y realizar la conversión de datos, pero ahora ya que debemos de realizar un inner join con una tabla que ya esta en nuestro datawarehouse, se obtiene dicha tabla, luego debemos de colocar un Ordenar, donde indicaremos que variable es la que se hará match entre las tablas par luego colocar su debido id, luego de realizar eso ya colocaremos la instrucción de Combinación de Mezcla que dicha tabla realizara el inner join y colocar los id correspondientes. Luego se juntan todos los resultados.



Y así mismo se realiza con las demás dataflows que tenemos en el control Flow.

De ultimo para poder obtener los counts de dichas tablas, debemos colocar otra “Tarea Ejecutar SQL” la cual pondremos las instrucciones sql para obtener los counts. Debemos de crear unas variables en donde se almacenarán dichos datos. Luego en la ventana de asignación le asignamos los datos para luego en un breakpoint poder observar dichos datos.



## Editor de la tarea Ejecutar SQL



Configure las propiedades requeridas para ejecutar instrucciones SQL y procedimientos almacenados mediante la conexión seleccionada.

General

Asignación de parámetro

Conjunto de resultados

Expresiones

Nombre de variable	Dirección	Tipo de ...	Nombre...	Tamaño ...
User::CountProveedor	Output	NVARC...	0	-1
User::CountCliente	Output	NVARC...	1	-1
User::CountVendedor	Output	NVARC...	2	-1
User::CountProducto	Output	NVARC...	3	-1
User::CountSucursal	Output	NVARC...	4	-1
User::CountCompra	Output	NVARC...	5	-1
User::CountVenta	Output	NVARC...	6	-1

Agregar

Quitar

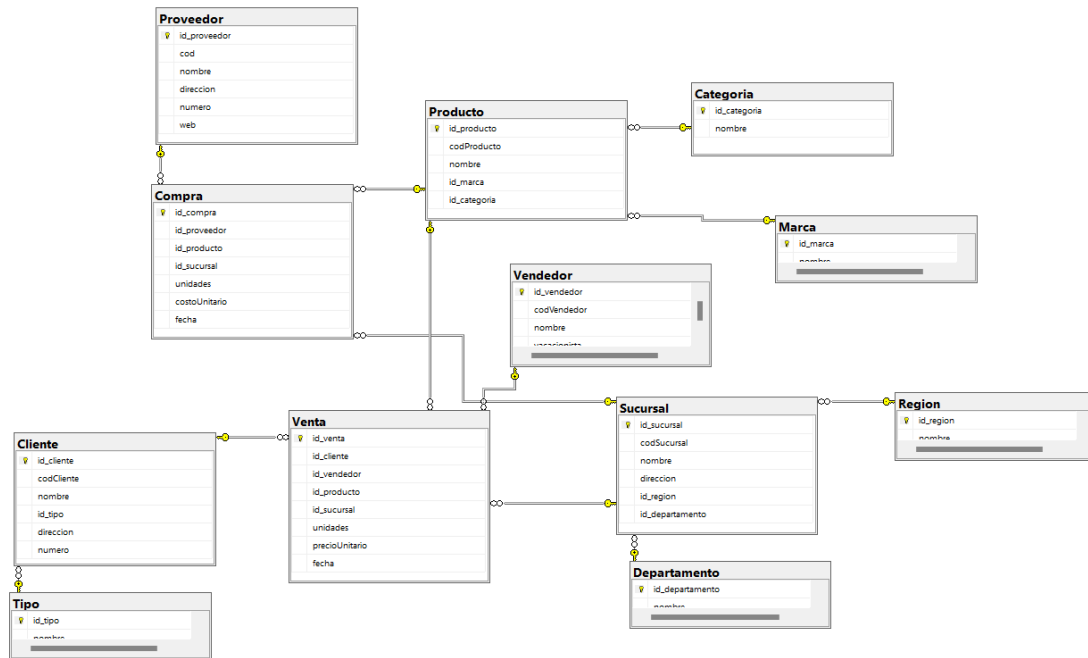
Aceptar

Cancelar

Ayuda

## Modelo DataWarehouse:

El modelo que se utilizó para dicho proyecto fue el siguiente:



Podemos observar que es un modelo Constelación, debido que hay mas de 1 tabla hecho donde se utilizan algunas dimensiones y en otra otras. Las tablas de hecho son las de Venta y Compra, mientras que las demás son dimensiones, dichas dimensiones tienen hijos las cuales las complementan.