

Universidad de San Carlos de Guatemala  
Facultad de Ingeniería  
Seminario de sistemas 2  
Aux. Sergio Gonzales  
Ing. Luis Vettorazzi  
Sección A

## **Proyecto SG-Food**

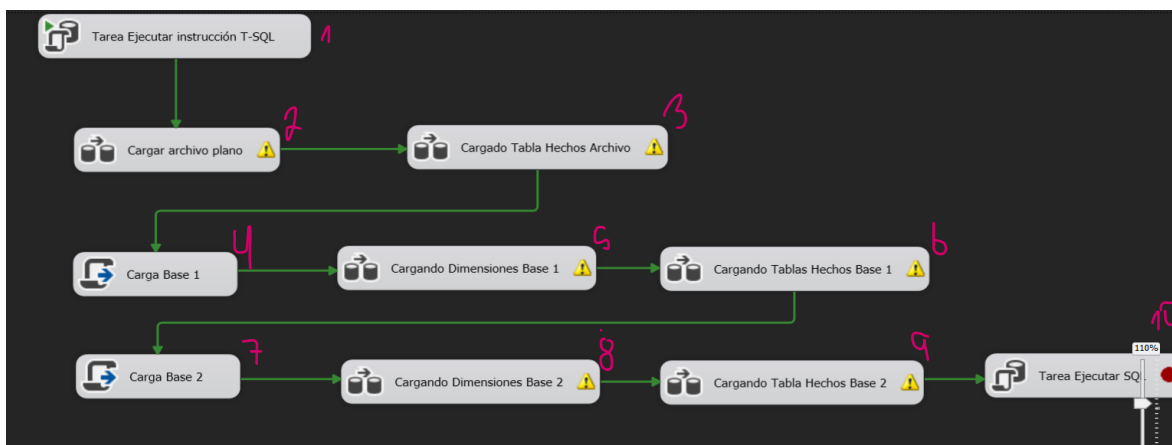
*Fase 1*

Mynor Alison Isai Saban Che

201800516

Guatemala, 20 de septiembre de 2022

## Flujo del sistema

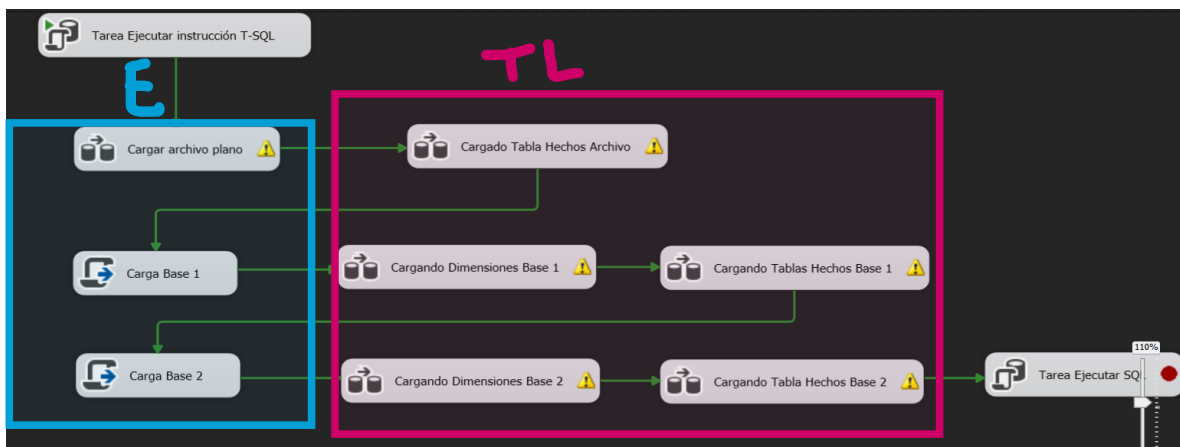


El flujo del sistema es el mostrado en la imagen anterior, el cual cuenta con 10 pasos lo cual se explica a continuación:

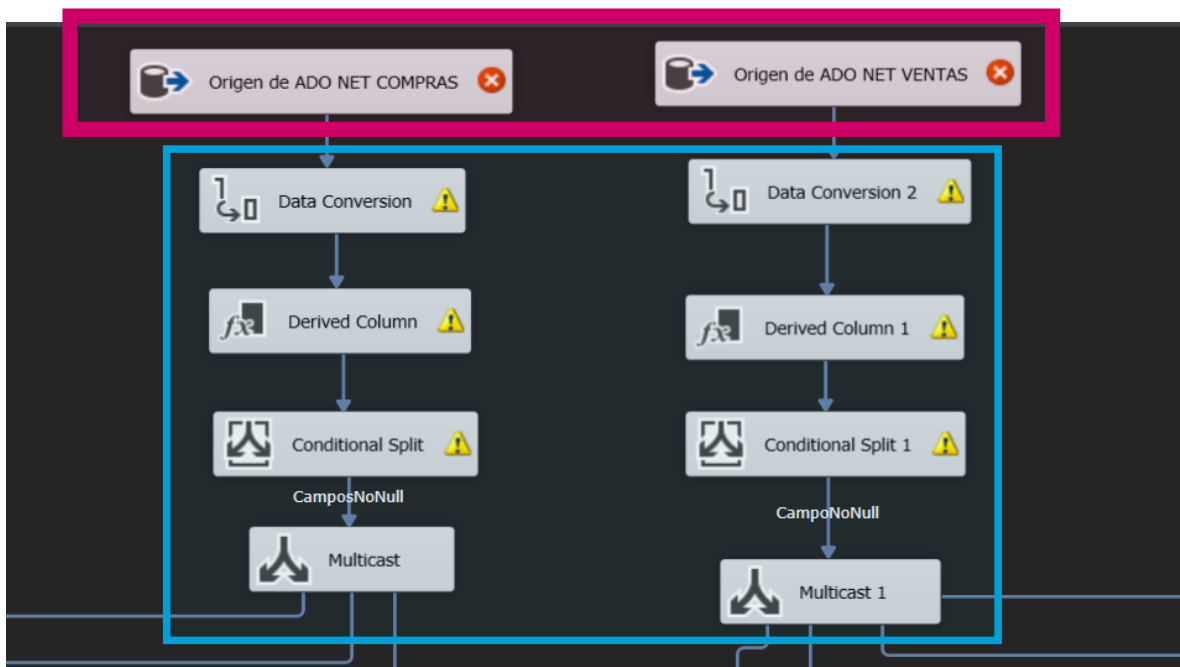
1. Tarea T-SQL: aquí se ejecuta los comandos de borrado de datos en la base de datos que contiene nuestra Datawarehouse, lo cual permitirá que se encuentre vacía la base de datos.
2. Cargar archivo plano: este lee el archivo el cual contiene datos y es transformado a hacia la base de Datawarehouse, el cual se transforma en las tablas necesarias para lograr el llenado de las tablas.
3. Cargar tablas Hechos Archivo: aquí se hace la transformación para realizar el llenado en la tabla de hechos de siempre consultando las tablas de dimensión.
4. Carga base 1: Se carga la información de los archivos en la base de datos 1 que en este caso se uso es MySQL, para el llenado de las tablas temporales (pivote), las cuales son dos tablas de compras y la tabla de ventas.
5. Cargando Dimensiones de base 1: este es el proceso de ETL el cual es el proceso de Extracción (Extrae los datos de la base de datos que se encuentra en MySQL), el proceso de Transformación (Se realiza el proceso de transformación de los datos, de cadenas de texto a numérico, el cual se adapta a la base de datos con el modelo Datawarehouse), y por último el proceso de Load (Se realiza el proceso de la carga en el modelo final, de la data que está en el proceso de transformación).
6. Cargando Tabla de hechos de base 1: es el proceso de ETL para el llenado de la tabla de hechos la cual es algo fundamental para el respectivo análisis que se realizara más adelante, este se hace el llenado con la base de datos 1 el cual se encuentra en MySQL.
7. Carga base 2: Se carga la información de los archivos en la base de datos 2 que en este caso se uso es SQL Server, para el llenado de las tablas temporales (pivote), las cuales son dos tablas de compras y la tabla de ventas.
8. Cargando Dimensiones de base 2: este es el proceso de ETL el cual es el proceso de Extracción (Extrae los datos de la base de datos que se encuentra en SQL Server), el

proceso de Transformación (Se realiza el proceso de transformación de los datos, de cadenas de texto a numérico, el cual se adapta a la base de datos con el modelo Datawarehouse), y por último el proceso de Load (Se realiza el proceso de la carga en el modelo final, de la data que está en el proceso de transformación).

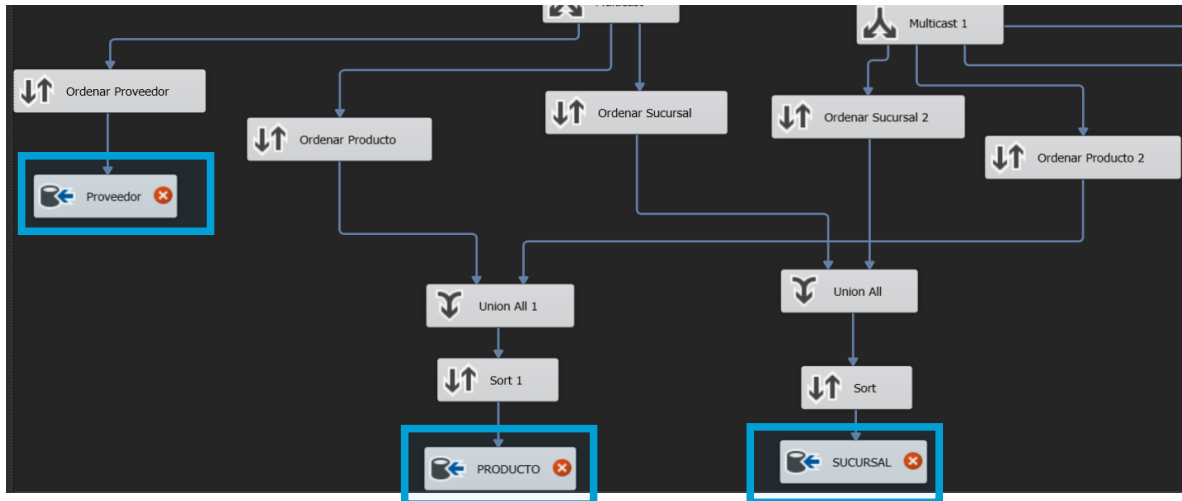
9. Cargando Tabla de hechos de base 2: es el proceso de ETL para el llenado de la tabla de hechos la cual es algo fundamental para el respectivo análisis que se realizara más adelante, este se hace el llenado con la base de datos 2 el cual se encuentra en SQL Server.
10. Tarea de ejecutar SQL: En esta tarea se realiza los count de cada tabla, para verificar que la tabla de hechos y las tablas de dimensiones se llenó correctamente



## Proceso de carga en tabla dimensión.

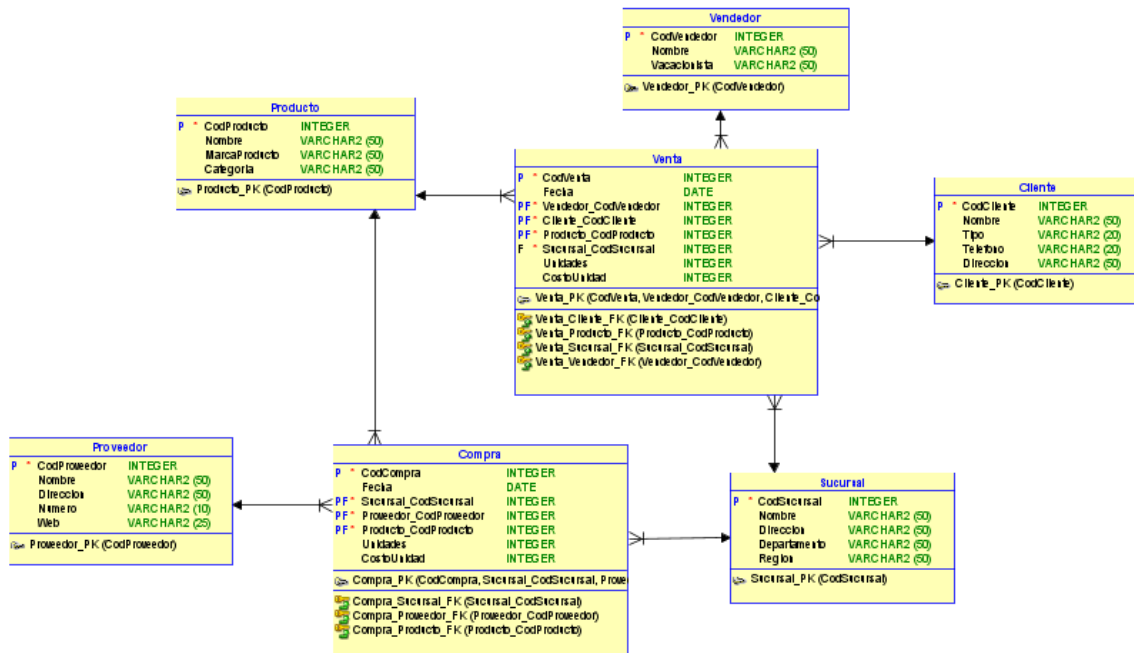


Se muestra la extracción de los datos lo cuales se realiza desde la base de datos 1 (recuadro en rojo), y a la vez pasa por la transformación, la cual se realiza una conversión de datos (string → entero), además quita los campos que están vacíos además solo envía los datos correctos a la fase de carga.



Ahora nos encontramos en la fase de carga, ya habiendo realizado el proceso de formación esta carga la data a sus respectivas tablas de dimensiones las cuales se antes de su ingreso se debe de realizar un filtro con el fin de que no exista data repetida en cada dimensión, ya que esto causaría problemas en el análisis de la información en la Datawarehouse.

## Entidad Relación de la Datawarehouse



## Explicación del modelo

### Dimensiones

#### Proveedor

En esta tabla se tienen todos los datos necesarios del proveedor para lograr su registro en Datawarehouse, esta dimensión esta compuesta por los siguientes campos:

- Código del proveedor
- Nombre.
- Dirección.
- Número telefónico.
- Web.

#### Producto

En esta tabla se tienen todos los datos necesarios del producto para lograr su registro en Datawarehouse, esta dimensión está compuesta por los siguientes campos:

- Código del producto.
- Nombre.
- Categoría.
- Marca del producto.

## **Sucursal**

En esta tabla se tienen todos los datos necesarios de la sucursal para lograr su registro en Datawarehouse, esta dimensión está compuesta por los siguientes campos:

- Código de la sucursal
- Nombre
- Dirección
- Departamento
- Región

## **Vendedor**

En esta tabla se tienen todos los datos necesarios del vendedor para lograr su registro en Datawarehouse, esta dimensión está compuesta por los siguientes campos:

- Código del vendedor.
- Nombre.
- Vacacionista (si es o no vacacionista).

## **Cliente**

En esta tabla se tienen todos los datos necesarios del cliente para lograr su registro en Datawarehouse, esta dimensión está compuesta por los siguientes campos:

- Código del cliente.
- Nombre.
- Tipo
- Teléfono.
- Dirección.

## **Tabla hechos**

### **Compra**

En esta tabla se tienen todos los datos necesarios de la compra para lograr su registro en Datawarehouse, esta dimensión está compuesta por los siguientes campos:

- Código de la compra.
- Fecha.
- Id de la sucursal.
- Id del producto.
- Id del proveedor.
- Cantidad comprada.
- Precio del producto.

## **Venta**

En esta tabla se tienen todos los datos necesarios de la venta para lograr su registro en Datawarehouse, esta dimensión está compuesta por los siguientes campos:

- Código de la venta.
- Fecha.
- Id de la sucursal.
- Id del producto.
- Id del cliente.
- Cantidad vendida.
- Precio del producto.