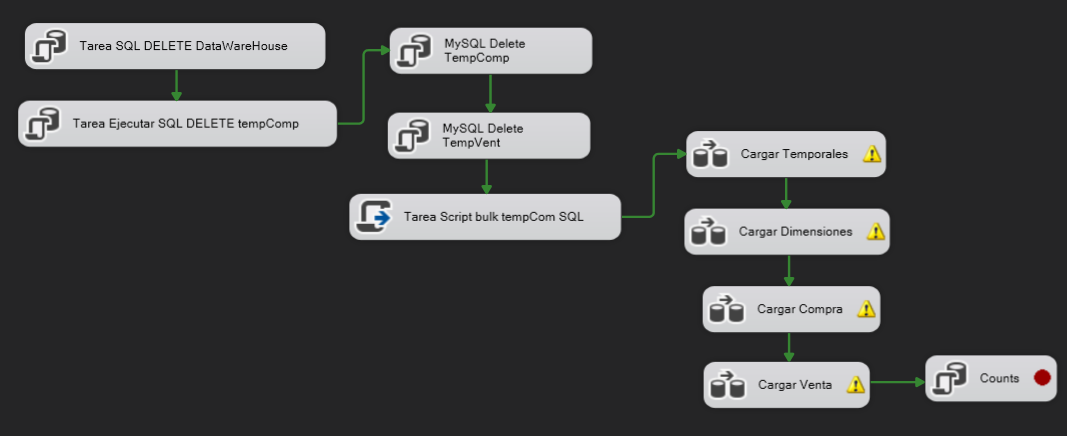
|  |  |
| --- | --- |
|  |  |
| Proyecto Fase 1 |  |
|  |  |
|  | 12/02/2023Seminario de Sistemas 2 - A |
|  | Anggelo Santiago Son Mux201709502 |

## Nombre del proyecto

Proyecto 1 Fase1 - SG-Food

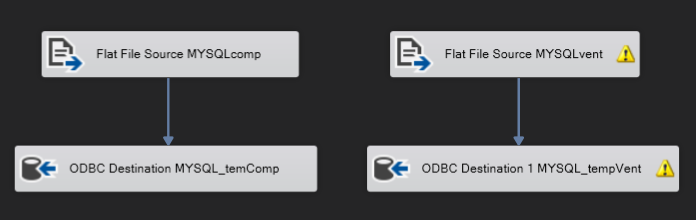
## Fases del proceso ETL



Para el control del flujo se utilizaron las herramientas:

* Tarea Ejecutar SQL: Ejecuta instrucciones SQL o procedimientos almacenados en una base de datos relacional.
* Tarea Script: Realiza funciones que no proporcionan las tareas estándar de integración service, se utilizo para la lectura de archivos y almacenamiento de la información obtenida en las tablas temporales ubicadas en la base de datos de SQL Server.
* Data Flow Task: Mueve los datos entre fuentes y destinos mientras se transforman y limpian.

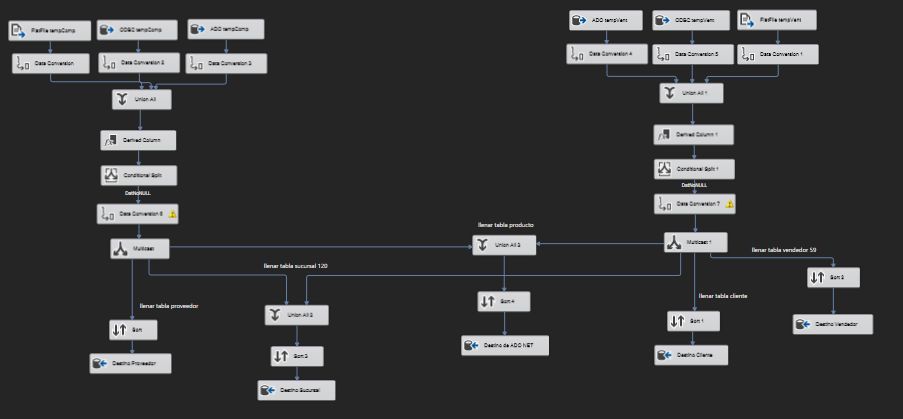
#### Carga de tablas temporales



Para la carga de las tablas temporales se utilizo la herramienta:

* Flat File Source: permite la lectura de archivos de texto y utilizar delimitadores para la identificación de columnas.
* ODBC Destination: se utilizo para cargar los datos obtenidos por el Flat File en las tablas temporales de la base de datos de MySQL.

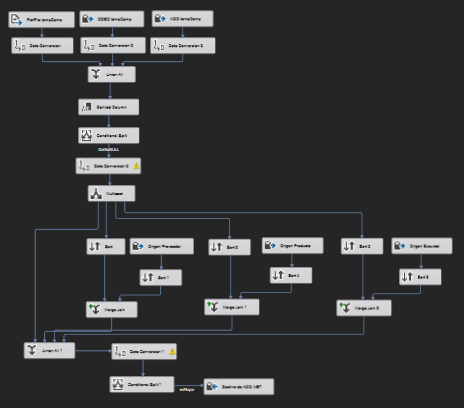
#### Carga de las tablas dimensiones



Para cargar los datos en las tablas dimensiones se utilizaron las herramientas:

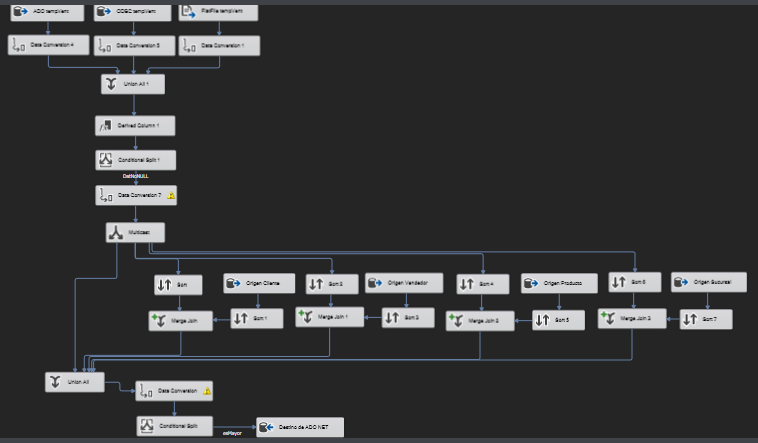
* Flat File Destination: permite la lectura de archivos de texto y utilizar delimitadores para la identificación de columnas.
* ODBC Source: extrae la información de las tablas temporales localizadas en la base de datos MySQL.
* Origen de ADO Net: Consume datos de SQL Server, se utilizo para extraer los datos almacenados en las tablas temporales localizadas en la base de datos SQL Server.
* Union All: combina la información de múltiples orígenes.
* Derived Column: Crea nuevas columnas y valores por la aplicación de expresiones para la transformación de los datos.
* Conditional Split: Permite dirigir los datos hacia diferentes entradas dependiendo de la condición especificada.
* Data Convertion: Convierte el tipo de un dato a otro tipo.
* Sort: Clasifica los datos de entrada en ascendente o descendente cuando no son clasificados por su fuente.

#### Carga de la tabla de hechos Compra



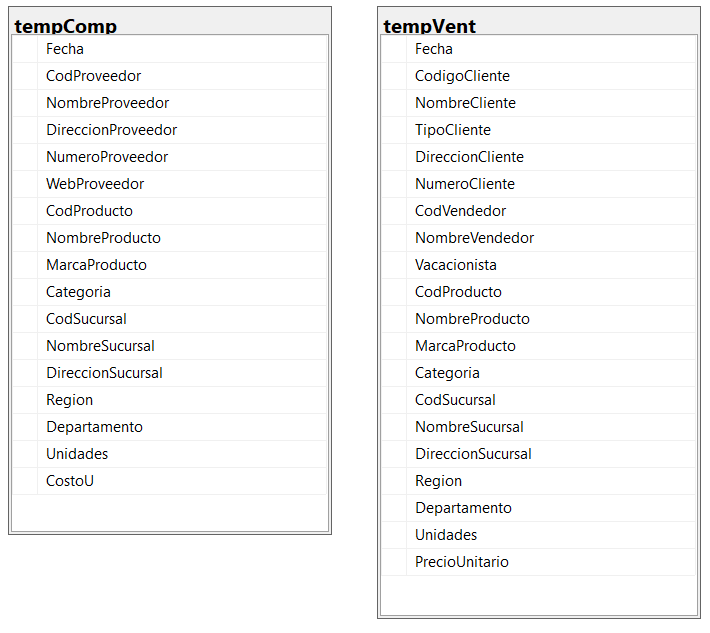
* Flat File Destination: permite la lectura de archivos de texto y utilizar delimitadores para la identificación de columnas.
* ODBC Source: extrae la información de las tablas temporales localizadas en la base de datos MySQL.
* Origen de ADO Net: Consume datos de SQL Server, se utilizo para extraer los datos almacenados en las tablas temporales localizadas en la base de datos SQL Server.
* Union All: combina la información de múltiples orígenes.
* Derived Column: Crea nuevas columnas y valores por la aplicación de expresiones para la transformación de los datos.
* Conditional Split: Permite dirigir los datos hacia diferentes entradas dependiendo de la condición especificada.
* Data Convertion: Convierte el tipo de un dato a otro tipo.
* Sort: Clasifica los datos de entrada en ascendente o descendente cuando no son clasificados por su fuente.
* Merge Join: Combina dos flujos de datos clasificados en uno usando inner join, left join o full join.

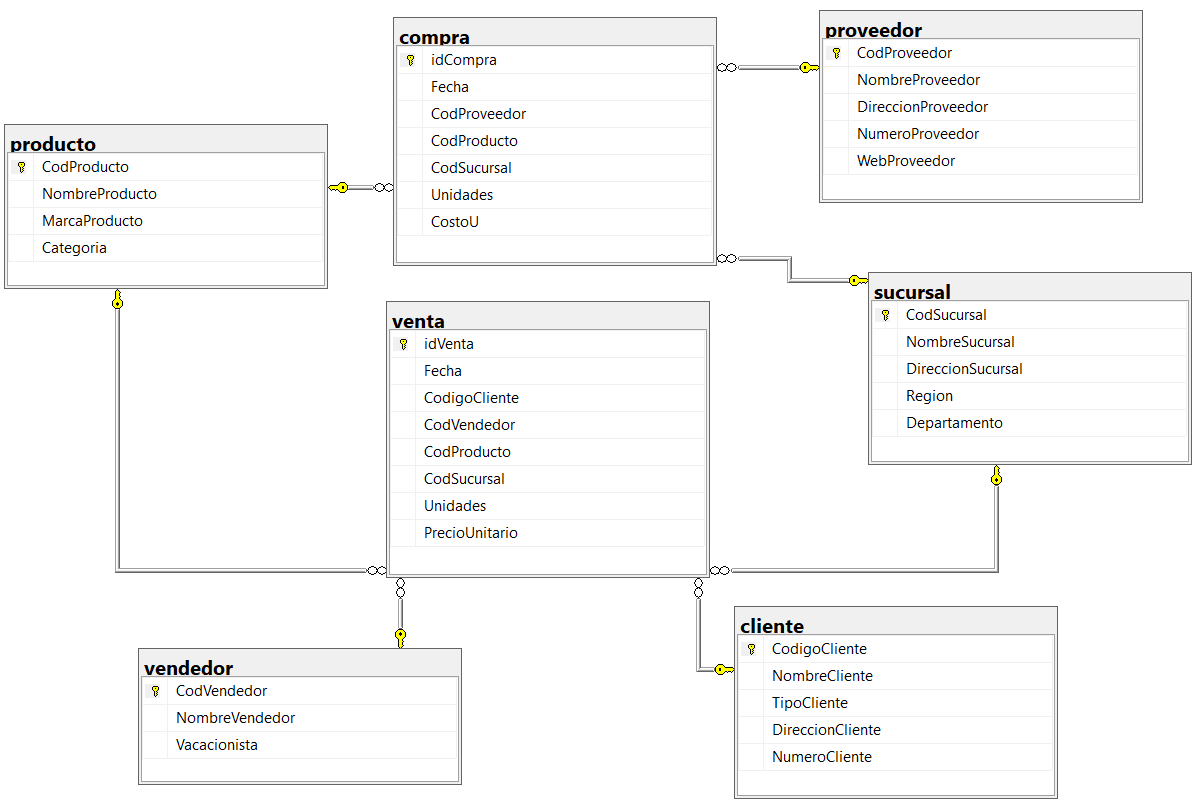
#### Carga de la tabla de hechos Venta



* Flat File Destination: permite la lectura de archivos de texto y utilizar delimitadores para la identificación de columnas.
* ODBC Source: extrae la información de las tablas temporales localizadas en la base de datos MySQL.
* Origen de ADO Net: Consume datos de SQL Server, se utilizo para extraer los datos almacenados en las tablas temporales localizadas en la base de datos SQL Server.
* Union All: combina la información de múltiples orígenes.
* Derived Column: Crea nuevas columnas y valores por la aplicación de expresiones para la transformación de los datos.
* Conditional Split: Permite dirigir los datos hacia diferentes entradas dependiendo de la condición especificada.
* Data Convertion: Convierte el tipo de un dato a otro tipo.
* Sort: Clasifica los datos de entrada en ascendente o descendente cuando no son clasificados por su fuente.
* Merge Join: Combina dos flujos de datos clasificados en uno usando inner join, left join o full join.

## Modelo del DataWarehouse





Para la implementación del DataWarehouse se hizo utilización de 5 dimensiones y 2 hechos optando por la utilización de un modelo constelación.

Dimensiones: producto, proveedor, sucursal, cliente, vendedor

Hechos: compra, venta