

UNIVERSIDAD DE SAN CARLOS DE
GUATEMALA
SEMINARIO DE SISTEMAS 2
SECCIÓN A
ING. LUIS VETTORAZI
SEGUNDO SEMESTRE 2023



FIUSAC
FACULTAD DE INGENIERÍA
UNIVERSIDAD DE SAN CARLOS DE GUATEMALA

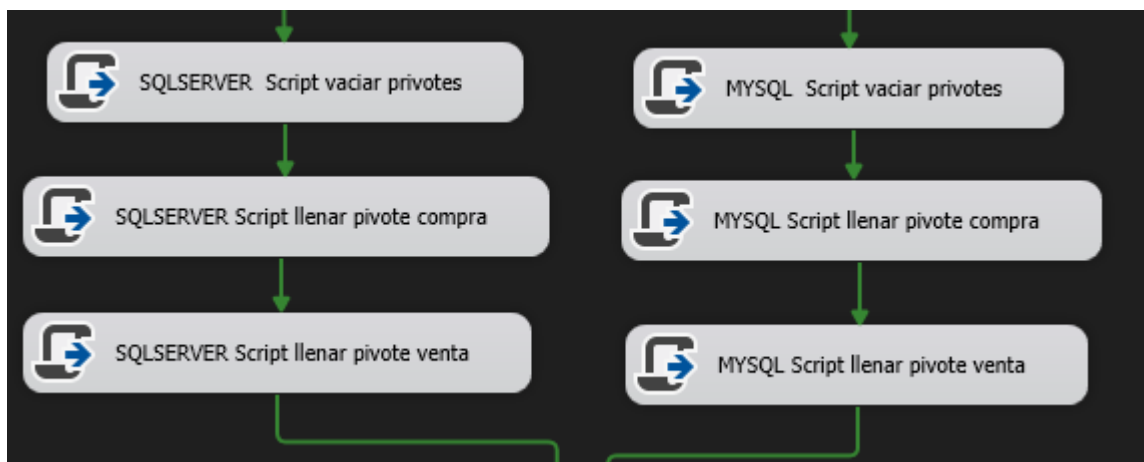
DOCUMENTACIÓN PROYECTO

201905554
MARVIN EDUARDO CATALAN VELIZ

NOMBRE

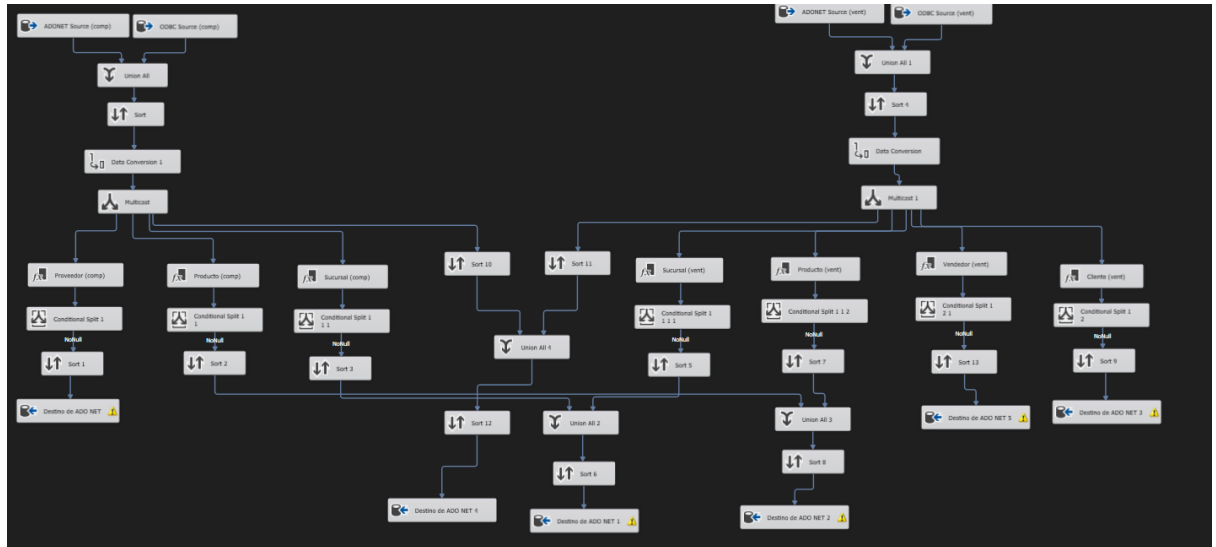
El nombre de este proyecto es **ETL FoodVital**, El ETL (Extracción, Transformación y Carga) tiene como propósito fundamental facilitar la obtención, procesamiento y almacenamiento eficiente de datos en el contexto de Business Intelligence (BI). En el caso específico de FoodVital, el ETL se utiliza para:

1. Extracción de Datos: Recopilar información relevante de archivos de texto delimitados por pipe, pleca o barra vertical (.comp y .vent) y de dos bases de datos de distintos fabricantes (SQL Server y MySQL).

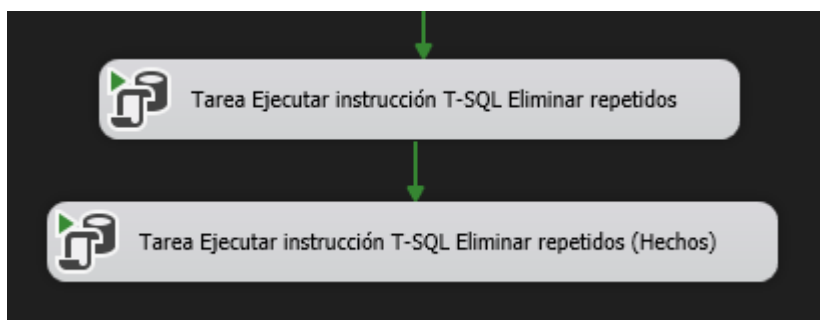


2. Transformación de Datos: Realizar procesos de limpieza y corrección en los archivos de entrada, abordando posibles errores como campos en blanco, números negativos o mal escritos. Además, se busca integrar datos de distintas fuentes y formatos en una estructura coherente y compatible con el modelo de DataWarehouse.

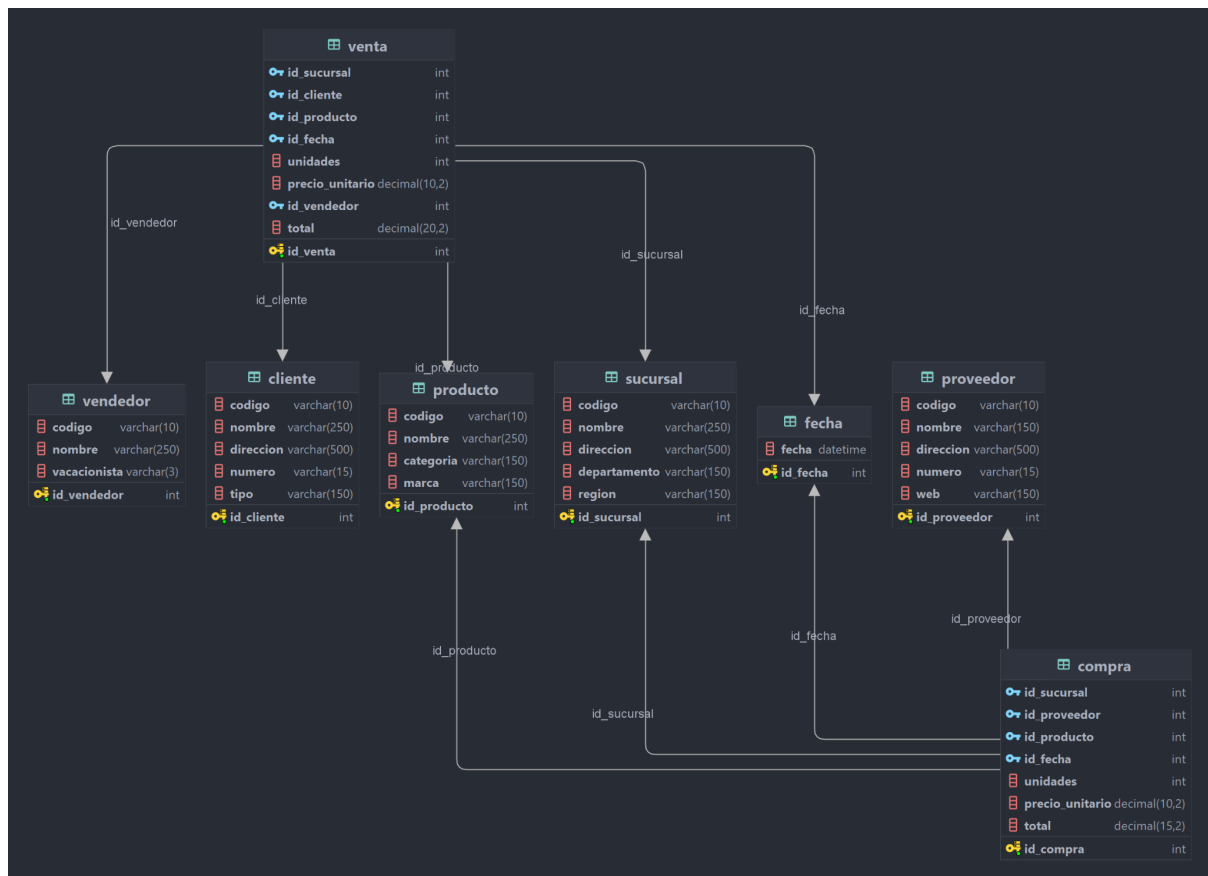
UNIVERSIDAD DE SAN CARLOS DE
GUATEMALA
SEMINARIO DE SISTEMAS 2
SECCIÓN A
ING. LUIS VETTORAZI
SEGUNDO SEMESTRE 2023



3. Carga de Datos: Insertar los datos procesados en un DataWarehouse, que actúa como un repositorio centralizado optimizado para la generación de informes y análisis. Este DataWarehouse se construyó en SQL Server y permitirá a FoodVital realizar consultas eficientes sobre sus ventas y compras.



MODELO IMPLEMENTADO



Se trata de un diseño de modelo de constelación. En un modelo de constelación, las tablas de dimensiones se normalizan adicionalmente, dividiéndose en subdimensiones para reducir la redundancia de datos. Veamos cómo se clasifican las tablas en este modelo:

1. Tablas de Hechos:

- compra: Representa las transacciones de compra con detalles específicos, referencias a las tablas de dimensiones y medidas como la cantidad de unidades, precio unitario y total.
- venta: Similar a compra, pero para las transacciones de venta.



UNIVERSIDAD DE SAN CARLOS DE
GUATEMALA
SEMINARIO DE SISTEMAS 2
SECCIÓN A
ING. LUIS VETTORAZI
SEGUNDO SEMESTRE 2023

2. Tablas de Dimensiones:
 - cliente, fecha, producto, proveedor, sucursal, vendedor: Estas tablas almacenan información detallada sobre dimensiones específicas.
3. Subdimensiones (normalizadas):
 - Las tablas de dimensiones mencionadas anteriormente podrían dividirse en subdimensiones adicionales para normalizar aún más los datos y reducir redundancias.

Justificación del Modelo de Constelación:

- Normalización Adicional: La normalización en subdimensiones permite reducir la redundancia de datos al dividir las tablas de dimensiones en partes más pequeñas y normalizadas.
- Ahorro de Espacio: Al normalizar las tablas de dimensiones, se puede ahorrar espacio de almacenamiento al eliminar la repetición de datos en múltiples filas.
- Integridad de Datos: La normalización ayuda a mantener la integridad de los datos al reducir la probabilidad de anomalías de actualización, inserción o eliminación.
- Flexibilidad: Aunque puede haber una mayor complejidad en la estructura, el esquema de constelación sigue siendo bastante flexible y puede adaptarse a cambios en las necesidades del negocio.