Evidencia de aprendizaje 4.Proyecto integrador-Repositorio de todas las actividades

Elaborado por:

Damián Pérez Salvatierra

Catalina Vinasco Arredondo

Yilma Valeria Espitia Sanabria

Docente:

Víctor Hugo Mercado

Curso:

Base de Datos II

Grupo: 76

Ingeniería de Software y Datos

Facultad de Ingeniería y de Ciencias Agropecuarias

Institución Universitaria Digital de Antioquia

2024

**Carga de Datos en Data Mart**

Una vez realizada la transformación de los datos, procedimos a realizar la carga de esta transformación al data mart. Primeramente creamos una nueva base de datos llamada ‘DataMartJardineria076’. Luego, procedimos a realizar la carga de cada una de las cuatro dimensiones que planteamos en la primera evidencia de la creación del modelo estrella. Que lo podemos visualizar a continuación:

Dimensión Geografía:

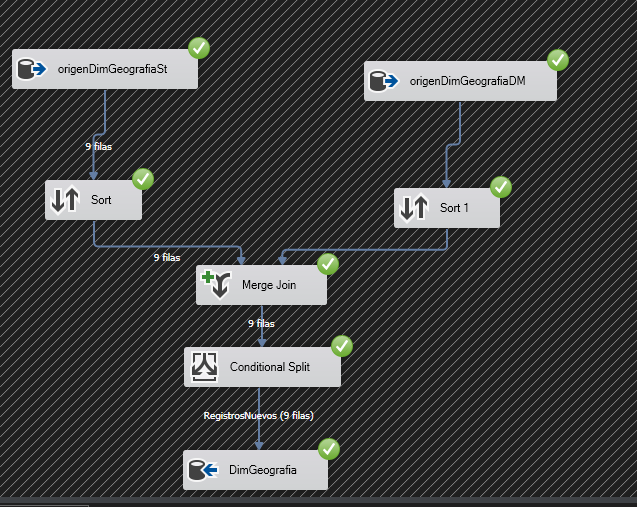


Imagen 1. Flujo de Datos exitoso sobre la dimensión de Geografía

Resultado final de la dimensión:

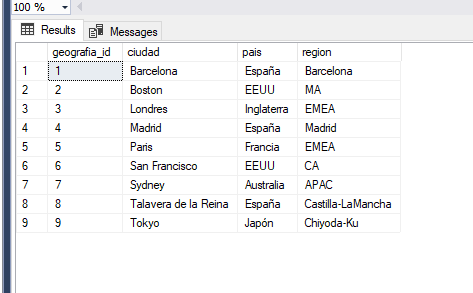


Imagen 2. Valores en la tabla de dimensión de Geografía

Dimension Empleado

Flujo de Datos en Visual

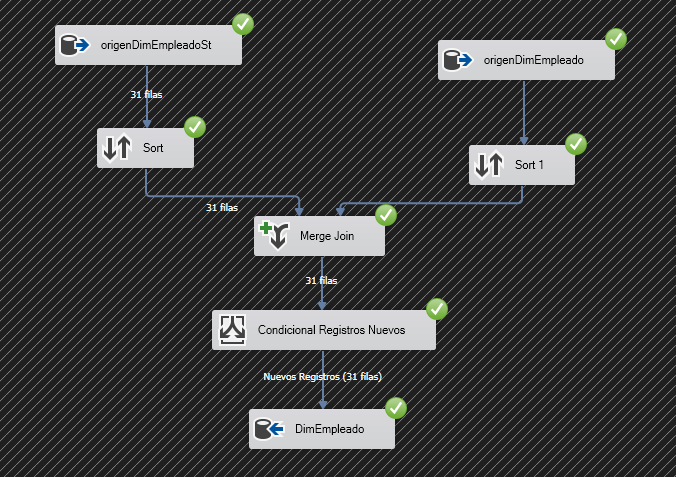


Imagen 3. Creación de la dimensión de Empleado

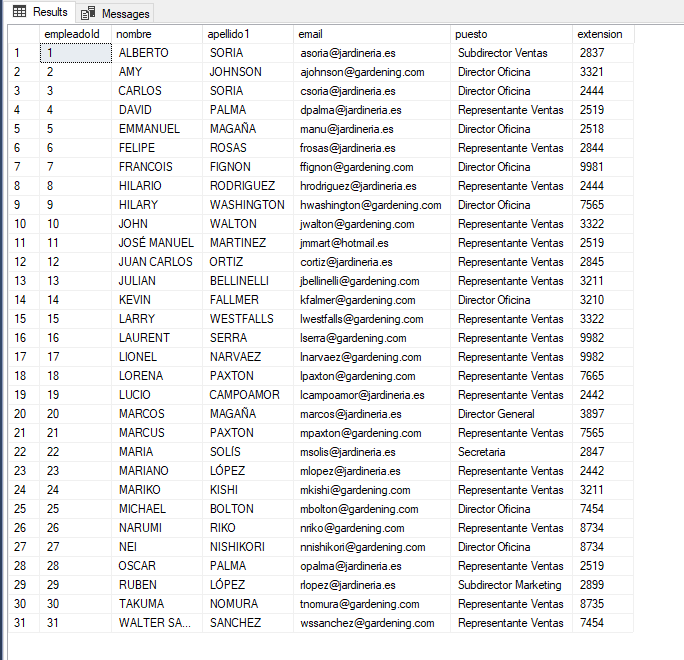
Resultados en la Tabla:

Imagen 4. Resultado de los valores de la dimensión de empleado

Dimension de Cliente:

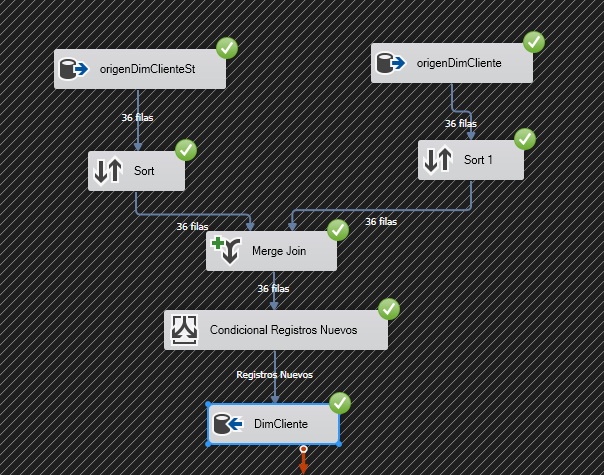


Imagen 5. Creación de la dimensión de Cliente

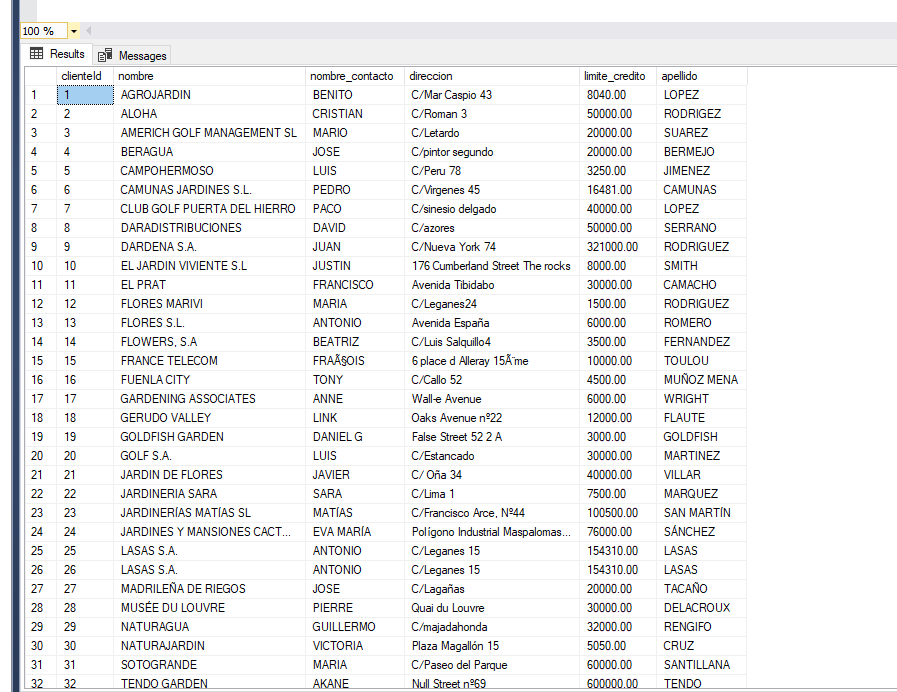
Resultados en la Tabla:

Imagen 6. Valores de la tabla de dimensión de Cliente

Dimensión de Producto:

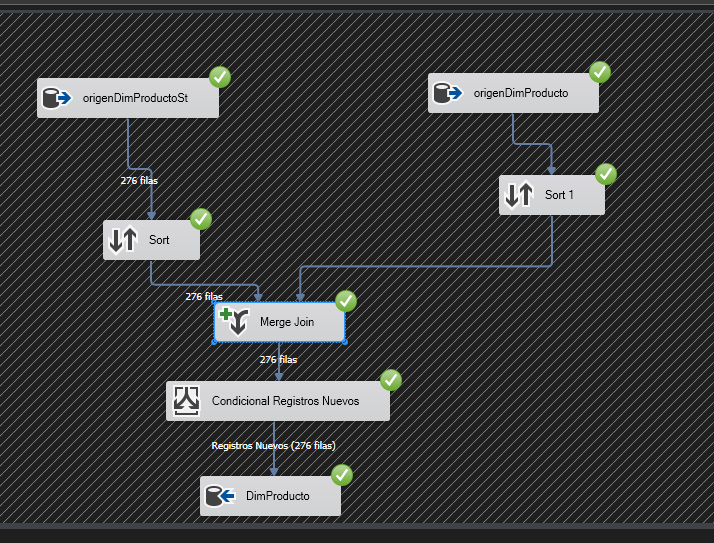


Imagen 7. Creación de Dimensión de Producto

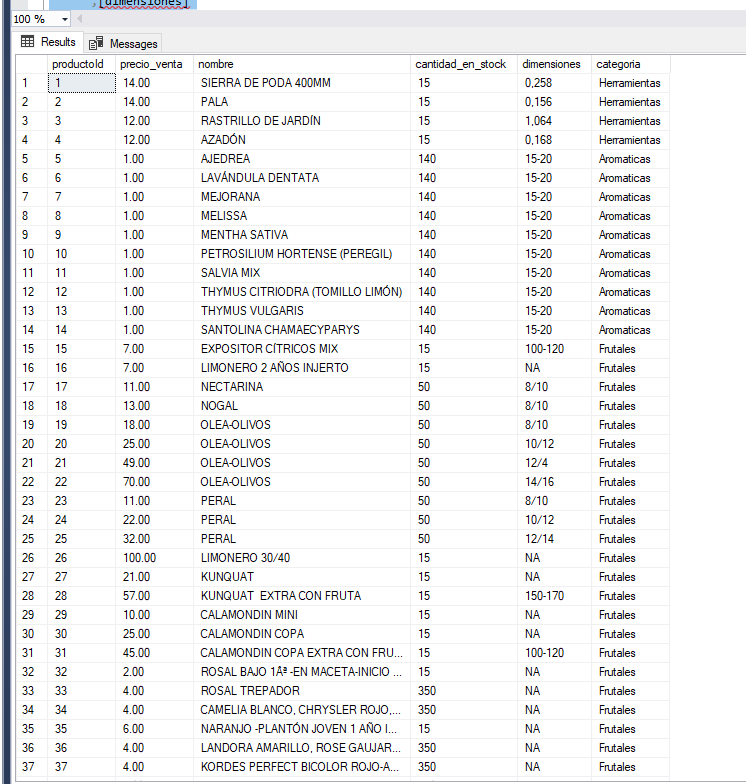
Resultados en la Tabla:

Imagen 8. Valores de la tabla de dimensión de Cliente

Creación Tabla de Hechos Factura de Venta

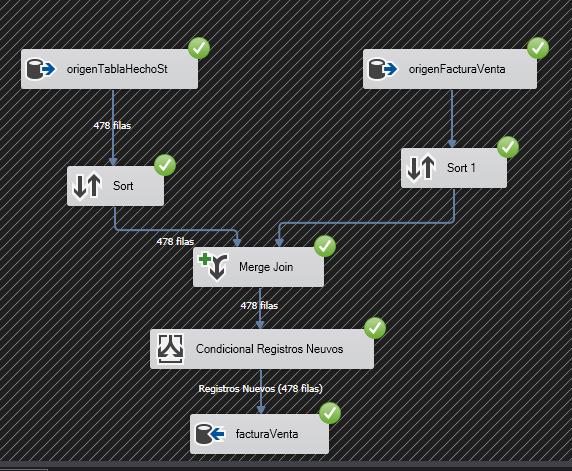


Imagen 9. Creación de Tabla de Hechos de Factura de Venta

Resultados en la Tabla:

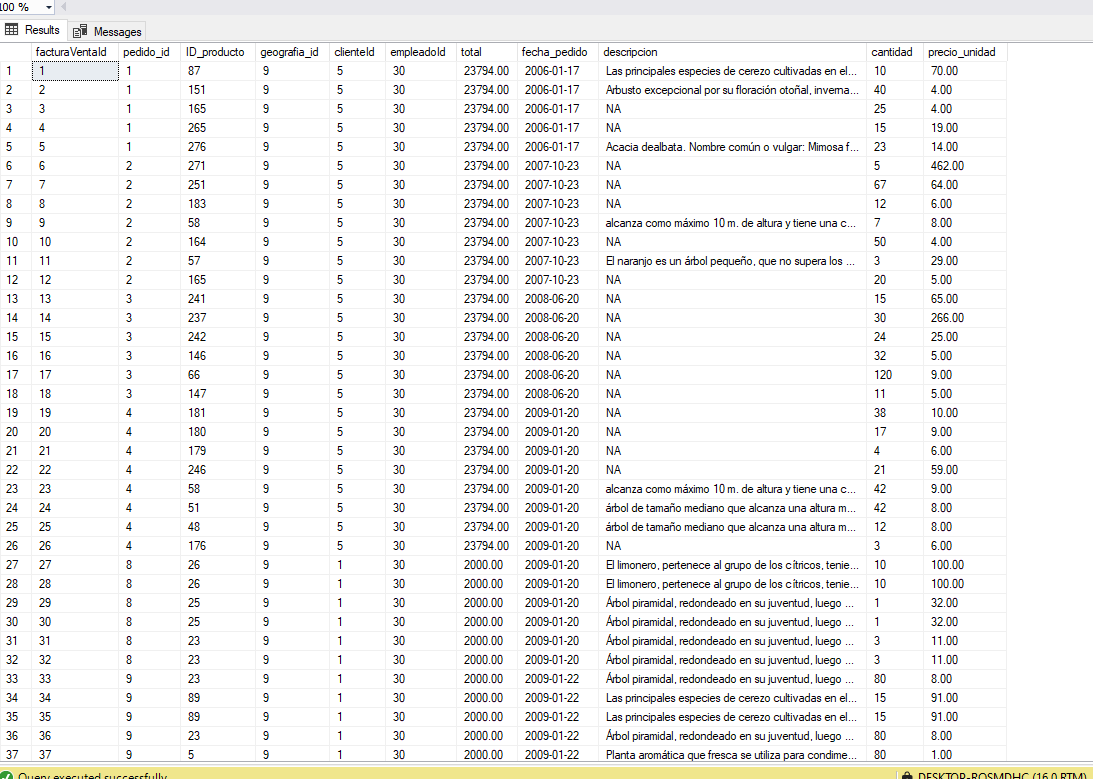


Imagen 10. Resultados de la tabla de hecho de Factura de Ventas

Después de realizar exitosamente la carga de datos en el Data Mart, nos encargamos de realizar las respectivas relaciones entre la tabla de hechos de factura de ventas y las distintas dimensiones como podemos ver a continuación:

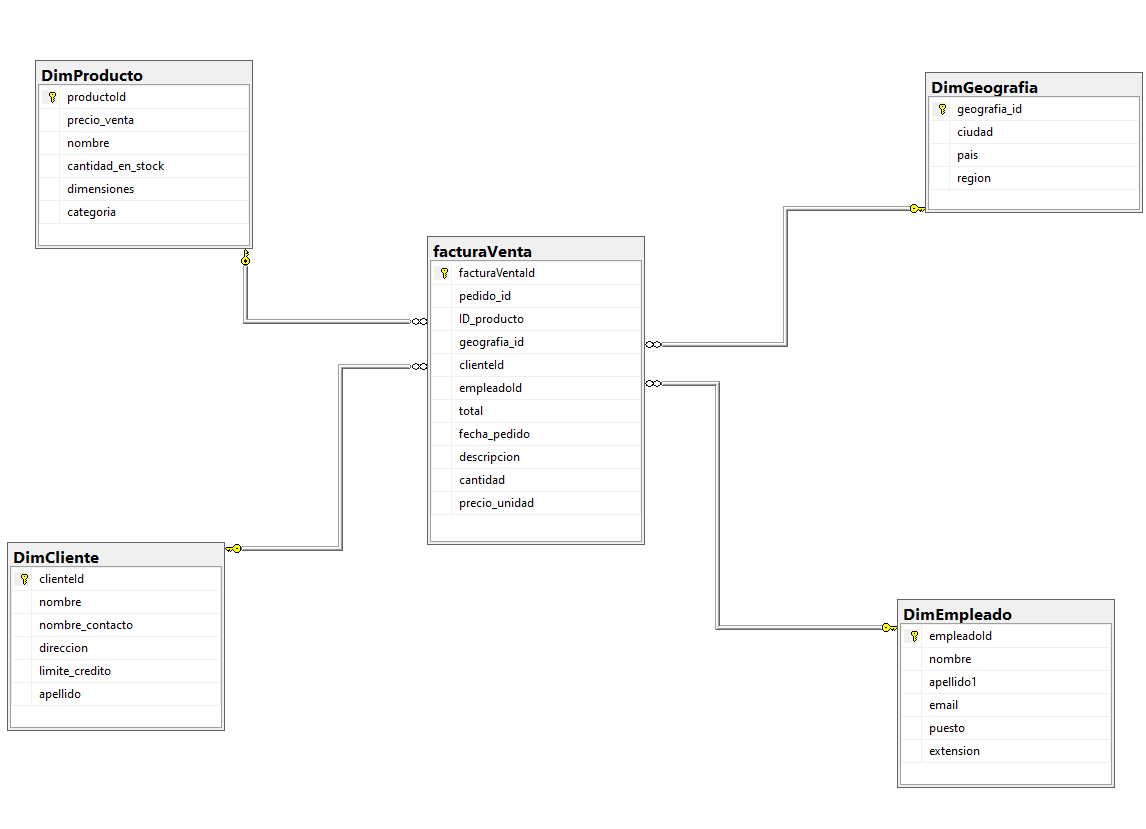
**Diagrama del Data Mart**

Imagen 10. Diagrama Estrella del data Mart.

Finalmente podemos observar como el proceso de ETL se llevó a cabo exitosamente en visual Studio 2022:

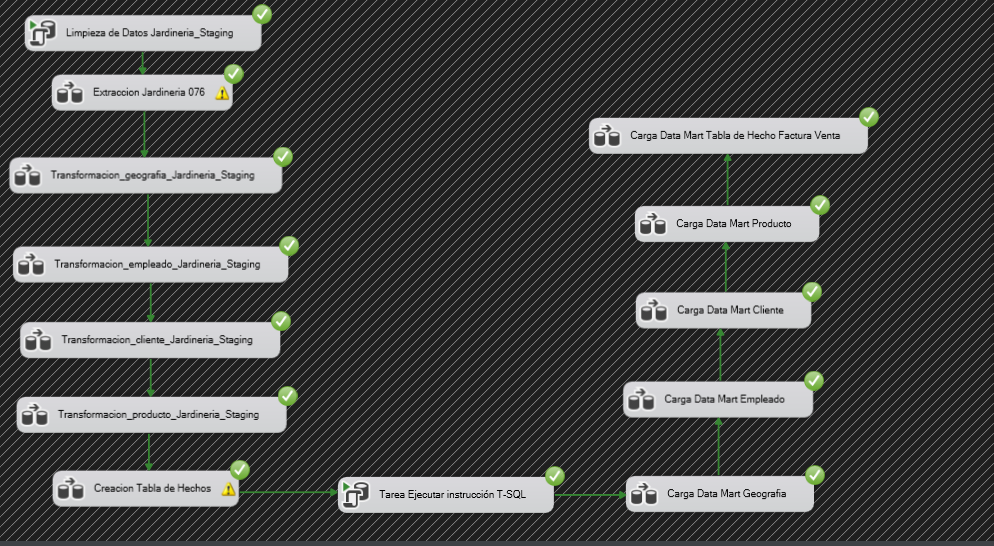


Imagen 11. Flujo de Datos exitoso en Visual Studio

**Verificación de calidad de datos**

Ya con el modelo estrella creado podemos realizar un control de la calidad de los datos

Ejemplo 1: Ver todas las compras que se realizaron en la ciudad de “Tokyo”, de aquí queremos obtener, quien hizo las compras, el empleado que asesoró en cada compra, el producto, total, fecha de pedido, descripción y cantidad:

SELECT FV.facturaVentaId, CONCAT(C.nombre\_contacto,' ', C.apellido) AS nombre\_completo, PR.nombre as producto, CONCAT(EM.nombre,' ', EM.apellido1) as nombre\_empleado,

FV.total, FV.fecha\_pedido, FV.descripcion, FV.cantidad

FROM facturaVenta FV

INNER JOIN DimGeografia G on FV.geografia\_id = G.geografia\_id

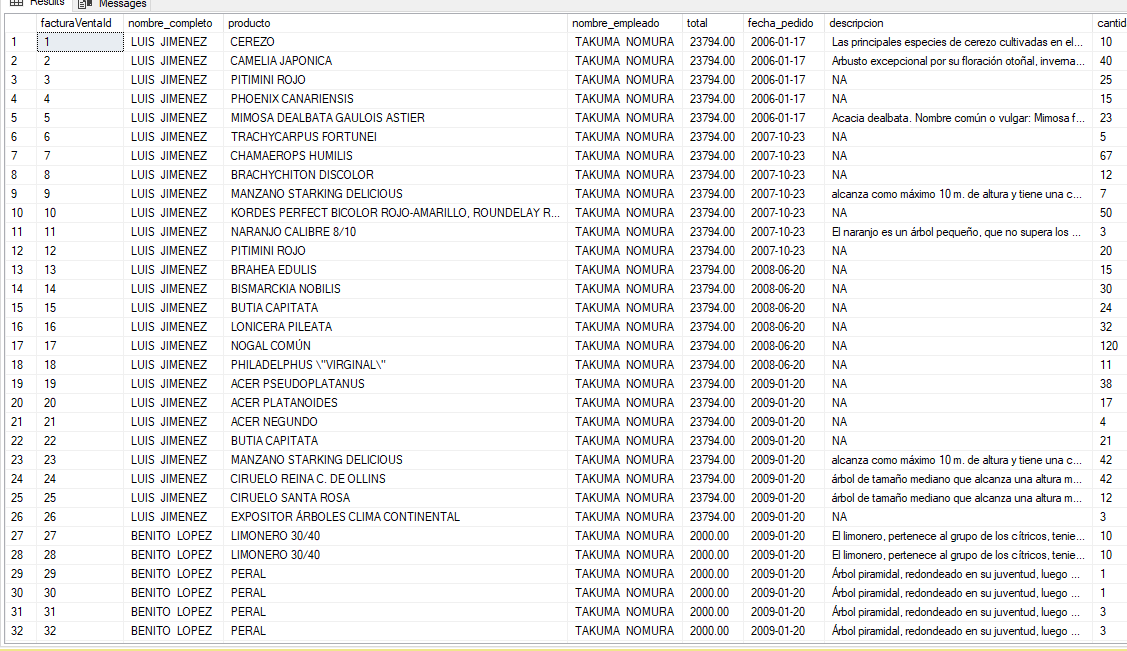
INNER JOIN DimCliente C on FV.clienteId= C.clienteId

INNER JOIN DimProducto PR on FV.ID\_producto= PR.productoId

INNER JOIN DimEmpleado EM on FV.empleadoId= EM.empleadoId

WHERE G.ciudad= 'Tokyo'

Resultado:



**Conclusión**

El proceso ETL (Extract, Transform, Load) en bases de datos se compone de tres actividades fundamentales que son vitales para la gestión efectiva de datos en entornos empresariales y/o inteligencia de negocios:

Extracción (Extract): Durante esta etapa, los datos se obtienen de una variedad de fuentes, incluyendo bases de datos, archivos planos y servicios web. Esta actividad es crucial para reunir datos crudos de múltiples fuentes y prepararlos para su procesamiento posterior.

Transformación (Transform): La transformación de datos implica una serie de operaciones destinadas a limpiar, estructurar y enriquecer los datos extraídos. Esta etapa es fundamental para garantizar la coherencia y la calidad de los datos, así como para prepararlos para su análisis y uso en aplicaciones empresariales.

Carga (Load): Finalmente, los datos transformados se cargan en el destino final que es la base de datos llamada ’DataMartJardineria076’, generalmente una base de datos o un almacén de datos. Esta actividad facilita el acceso y la consulta de datos por parte de usuarios y aplicaciones, permitiendo la generación de conocimiento y la toma de decisiones informadas.