



Proceso de transformación de datos y carga en el data mart final

Grupo #5

Jhon Neider Cuervo Quintero

Yuri Marcela Cuervo Quintero

Formación

ING software y datos

Base De Datos II - PREICA2502B010064

Docente

Antonio Jesús Valderrama

30/09/2025

Introducción

El presente informe describe el desarrollo de un proceso de Extracción, Transformación y Carga (ETL) utilizando la base de datos Jardinería. El objetivo del proyecto fue construir un Data Mart basado en un modelo estrella, que permita realizar análisis relevantes como la identificación del producto más vendido y las ventas por cliente, empleado y tiempo.

Preparación

Se revisó el modelo estrella definido para comprender la estructura de las tablas de hechos y dimensiones. También se verificó la consistencia y disponibilidad de la base de datos staging previamente creada, asegurando que los datos estuvieran listos para su transformación y carga en el Data Mart.

Extracción de datos

La extracción se realizó desde la base de datos original 'jardineria', copiando la información de las tablas clientes, productos, pedidos, empleados y detalle_pedido hacia la base de datos staging. Esta etapa permitió centralizar la información en un entorno intermedio antes de aplicar transformaciones.

Transformación de datos

Se aplicaron transformaciones para limpiar y estandarizar los datos, incluyendo:

- Conversión de valores nulos a valores válidos (ejemplo: límite de crédito).

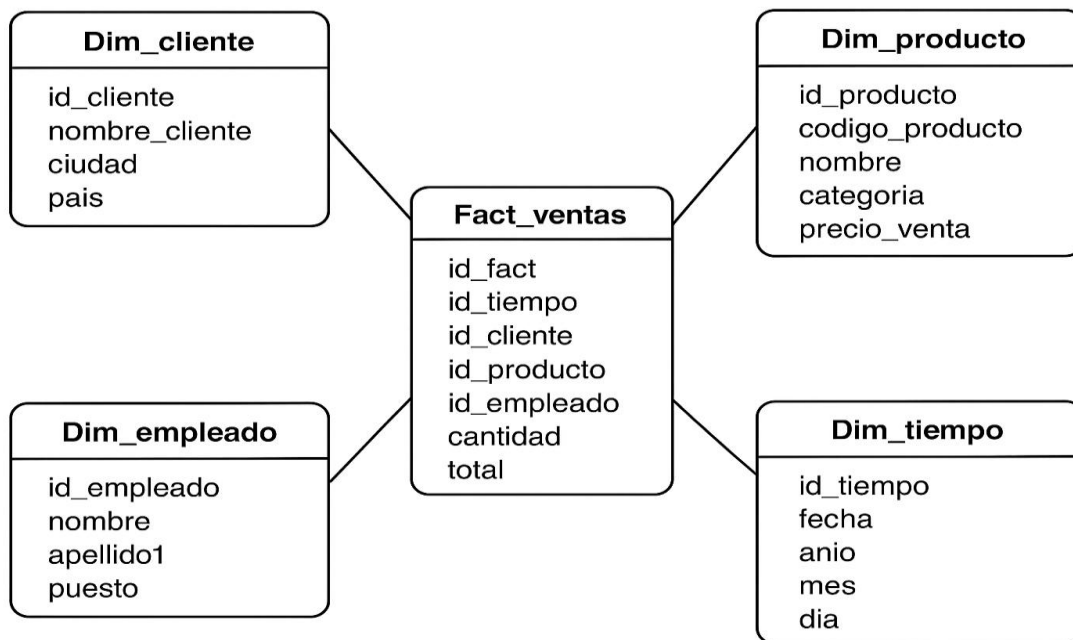
- Corrección de valores negativos en precios.
- Creación de campos derivados como año, mes, día y trimestre en la dimensión tiempo.
- Normalización de datos de clientes y productos.

Carga en el data mart

Los datos transformados fueron cargados en el Data Mart 'jardineria_datamart'. Se poblaron las dimensiones (dim_cliente, dim_producto, dim_empleado, dim_tiempo) y la tabla de hechos (fact_ventas). Esto permitió consolidar la información en un modelo estrella optimizado para el análisis.

Modelo estrella

A continuación, se muestra el diagrama del modelo estrella implementado, en el cual la tabla de hechos fact_ventas se relaciona con las dimensiones cliente, producto, empleado y tiempo.



Consultas analíticas

Algunas de las consultas realizadas sobre el Data Mart fueron:

- Total de ventas por cliente.
- Productos más vendidos.
- Ventas realizadas por cada empleado.
- Ventas agrupadas por año y mes.

Estas consultas permiten apoyar la toma de decisiones estratégicas en la empresa.

Conclusiones

El proceso ETL permitió consolidar la información del sistema Jardinería en un Data Mart con estructura estrella. Gracias a este modelo, es posible realizar análisis eficientes y relevantes para la gestión del negocio, identificando tendencias de ventas, clientes más importantes y desempeño de empleados.