



Institución Universitaria Digital de Antioquia

Facultad de Ingeniería

Ingeniería de software y datos

## BASE DE DATOS II

S30 - Evidencia de aprendizaje 3. Proceso de transformación de datos y carga en el data mart final

Autor:

**Robinson Zapata Villada**

**CC:1017242373**

**Medellín – Colombia**

**30 de sep. de 2025**

## Introducción

El objetivo de este proyecto es implementar un proceso ETL completo para la base de datos transaccional (OLTP) "Jardinería". Este proceso consiste en extraer los datos de origen, cargarlos en una base de datos de Staging, transformarlos para adaptarlos a un modelo analítico y, finalmente, cargarlos en un Data Mart con un modelo de Estrella.

El propósito final del Data Mart es facilitar consultas analíticas complejas para la toma de decisiones, como identificar los productos más vendidos, analizar ventas por región o entender el comportamiento de los clientes a lo largo del tiempo.

## Objetivos

- **General:** Desarrollar un proceso completo de ETL (Extract, Transform, Load) para migrar datos desde la base de datos transaccional de Jardinería hacia un Data Mart dimensional optimizado para análisis de ventas y toma de decisiones empresariales.
- **Específicos:** Estos objetivos cubren las tres fases del proceso (Extracción, Transformación y Carga) y la validación final, asegurando que se cumpla el objetivo general.
  - **Extracción y Preparación (E - Staging)**
    - **Confirmar de Staging:** Verificar la disponibilidad y consistencia de la base de datos staging\_jardineria y sus tablas (stg\_\*) para asegurar que contenga los datos extraídos previamente de la BD jardinería.
    - **Validar la Integridad de la Extracción:** Asegurar la integridad de los datos extraídos en la Staging, confirmando que no existen valores huérfanos ni inconsistencias críticas que impidan el proceso de transformación posterior.
  - **Transformación (T)**
    - **Diseñar el Modelo Dimensional:** Definir y documentar el Modelo Estrella del Data Mart, especificando las tablas de Hechos (Hechos\_Ventas) y las Dimensiones (Dim\_Tiempo, Dim\_Producto, Dim\_Cliente, Dim\_Empleado, Dim\_Oficina) y sus relaciones.
    - **Implementar Transformaciones Analíticas:** Desarrollar y aplicar scripts SQL para realizar transformaciones clave en los datos
  - **Carga y Validación (L)**
    - **Construir el Esquema del Data Mart:** Crear la base de datos data\_mart\_jardineria e implementar todas las tablas del Modelo Estrella, incluyendo sus claves primarias y foráneas.
    - **Ejecutar la Carga Dimensional:** Cargar los datos transformados desde la base de datos de Staging hacia las tablas de **Dimensión** del Data Mart.
    - **Ejecutar la Carga de Hechos:** Cargar los datos transformados en la tabla de Hechos (Hechos\_Ventas), estableciendo las relaciones con las dimensiones mediante las claves subrogadas.
    - **Realizar Pruebas de Calidad:** Ejecutar consultas de validación en el Data Mart final para confirmar que la cantidad de registros sea correcta y que las relaciones dimensionales se hayan establecido adecuadamente.

## Análisis del Problema

- Se debe realizar una arquitectura de 3 capas para su desarrollo

Capa de origen:  
-BD jardineria  
- 9 tablas relacionales

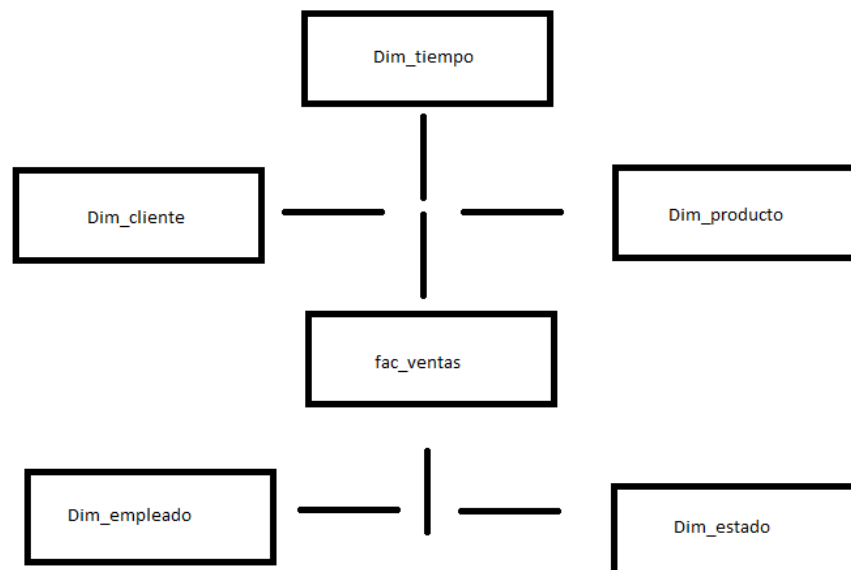
•Extraccion

Capa Staging:  
-BD Staging\_jardineria  
- 8 tablas temporales

•Transformación

Capa Data Mart  
-BD datamart\_jardineria  
-5 dimensiones + 1 tabla de hechos

- Flujo de datos: ORIGEN → STAGING → DATA MART → ANÁLISIS
- El Data Mart implementa un modelo estrella que centraliza las métricas de ventas y las relaciona con dimensiones descriptivas para facilitar el análisis multidimensional.



**Dimensiones del modelo**

**Dimensión Tiempo (dim\_tiempo)**

Campo	Tipo	Descripción
sk_tiempo	INT (PK)	Clave surrogate
fecha	DATE	Fecha completa
anio	INT	Año (2006-2009)
mes	INT	Mes (1-12)
trimestre	INT	Trimestre (1-4)
nombre_mes	VARCHAR(20)	Nombre del mes

### Dimensión Cliente (dim\_cliente)

Campo	Tipo	Descripción
sk_cliente	INT (PK)	Clave surrogate
nombre_cliente	VARCHAR(50)	Nombre o razón social
ciudad	VARCHAR(50)	Ciudad del cliente
pais	VARCHAR(50)	País del cliente
limite_credito	NUMERIC(15,2)	Límite de crédito asignado
segmento_credito	VARCHAR(20)	Sin Crédito / Bajo / Medio / Alto

### Dimensión Producto (dim\_producto)

Campo	Tipo	Descripción
sk_producto	INT (PK)	Clave surrogate
codigo_producto	VARCHAR(15)	Código único del producto
nombre_producto	VARCHAR(70)	Nombre descriptivo
categoria	VARCHAR(50)	Categoría del producto
rango_precio	VARCHAR(20)	Económico / Medio / Premium / Lujo

Tabla de Hechos (fact\_ventas)

Granularidad: Línea de detalle de pedido

Campo	Tipo	Descripción
sk_venta	INT (PK)	Clave primaria surrogate
sk_tiempo	INT (FK)	Referencia a dim_tiempo
sk_cliente	INT (FK)	Referencia a dim_cliente
sk_producto	INT (FK)	Referencia a dim_producto
cantidad	INT	Métrica: Unidades vendidas
precio_unitario	NUMERIC(15,2)	Métrica: Precio por unidad
monto_linea	NUMERIC(15,2)	Métrica: Total línea (cant x precio)

SQLQuery1.sql - RO...binson Zapata (58))

```
WHERE de.tipo_estado = 'Exitoso'
GROUP BY dt.anio, dt.mes, dt.nombre_mes
)
SELECT
  anio,
  mes,
  nombre_mes,
  ingresos,
  LAG(ingresos) OVER (ORDER BY anio, mes) AS ingresos_mes_anterior,
  ingresos - LAG(ingresos) OVER (ORDER BY anio, mes) AS diferencia,
  CASE
    WHEN LAG(ingresos) OVER (ORDER BY anio, mes) IS NOT NULL
    THEN CAST((ingresos - LAG(ingresos) OVER (ORDER BY anio, mes)) * 100.0 /
      LAG(ingresos) OVER (ORDER BY anio, mes) AS DECIMAL(10,2))
    ELSE NULL
  END AS porcentaje_crecimiento
FROM ventas_mensuales
ORDER BY anio, mes;
GO

PRINT '';
PRINT 'Todas las consultas analíticas ejecutadas exitosamente';
GO
```

100 %

Results

Messages

	codigo_producto	nombre_producto	categoria	rango_precio	total_unidades_vendidas	numero_pedidos	ingresos_totales	precio_promedio
1	AR-009	Thymus Vulgaris	Aromaticas	Económico	961	6	986.00	1.166666
2	FR-17	Rosal bajo 1Aº-En maceta-inicio brotación	Frutales	Económico	423	1	846.00	2.000000
3	OR-214	Brahea Armata	Omamentales	Medio	212	1	2120.00	10.000000
4	OR-247	Trachycarpus Fortunei	Omamentales	Lujo	198	4	72408.00	348.500000
5	OR-208	Tuja orientalis \"Aurea nana\"	Omamentales	Económico	191	5	764.00	4.000000
6	OR-227	Chamaerops Humilis	Omamentales	Premium	185	5	11840.00	64.000000

Query executed successfully.

ROBINSON-PC\SQLEXPRESS (16... | ROBINSON-PC\Robinson Z... | datamart\_jardineria | 00:00:00 | 103 rows

