



Institución Universitaria Digital de Antioquia

Facultad de Ingeniería

Ingeniería de software y datos

BASE DE DATOS II
S25-CONSTRUCCIÓN DE MODELO
ESTRELLA

Autor:

Robinson Zapata Villada

CC:1017242373

Medellín – Colombia

02 de sep. de 2025

Introducción

El presente documento detalla la construcción de un modelo de datos tipo modelo estrella a partir de la base de datos Jardinería. Este modelo está diseñado para optimizar el análisis de información clave del negocio, como el rendimiento de productos, categorías y ventas a lo largo del tiempo. Al transformar los datos de un esquema relacional a uno dimensional, se busca facilitar la consulta de métricas de negocio de manera ágil y eficiente.

Objetivos

- **General:** Construir un modelo de datos tipo estrella para un data mart que permita un análisis eficiente de las ventas relacionadas en la BD Jardinería.
- **Específicos:**
 - Identificar las tablas y campos relevantes en la base de datos Jardinería para el análisis de ventas.
 - Diseñar y documentar la estructura de una tabla de hechos y sus dimensiones asociadas.
 - Crear un modelo que permita responder a las siguientes preguntas de negocio:
 - ¿Cuál es el producto más vendido?
 - ¿Cuál es la categoría con más productos?
 - ¿Cuál es el año con más ventas?

Planteamiento del Problema

La base de datos presenta un esquema relacional diseñado para la gestión de transacciones diarias, lo cual es óptimo para las operaciones, pero bastante ineficiente para la toma de decisiones. Analizar métricas como las ventas por año, producto o categoría requiere de múltiples joins entre tablas, lo que ralentiza las consultas y complica el proceso de análisis de datos. Por lo tanto, se necesita una solución que estructure los datos de forma que las consultas analíticas sean directas y de alto rendimiento.

Análisis del Problema

La base de datos Jardinería presenta la siguiente estructura relacional:

1. **detalle_pedido**: Tabla central que registra cada línea de venta, conteniendo:
 - ID_detalle_pedido, ID_pedido, ID_producto
 - cantidad, precio_unidad, numero_linea
2. **pedido**: Registra las órdenes de compra con:
 - ID_pedido, fecha_pedido, fecha_esperada, fecha_entrega
 - estado, comentarios, ID_cliente
3. **producto**: Catálogo de productos con:
 - ID_producto, CódigoProducto, nombre, Categoría
 - dimensiones, proveedor, descripción
 - cantidad_en_stock, precio_venta, precio_proveedor
4. **Categoría_producto**: Clasificación de productos con:
 - Id_Categoría, Desc_Categoría, descripción_texto
 - descripción_html, imagen
5. **cliente**: Información de clientes con:
 - ID_cliente, nombre_cliente, datos de contacto
 - ubicación geográfica, ID_empleado_rep_ventas, límite_credito
6. **empleado**: Datos del personal con:
 - ID_empleado, nombre, apellidos, email
 - ID_oficina, ID_jefe, puesto
7. **oficina**: Ubicaciones de la empresa con:
 - ID_oficina, Descripción, ciudad, país, región
 - código_postal, teléfono, direcciones
8. **pago**: Registros de pagos con:
 - ID_pago, ID_cliente, forma_pago
 - id_transacción, fecha_pago, total

Para responder a las preguntas de negocio planteadas, las tablas más relevantes son:

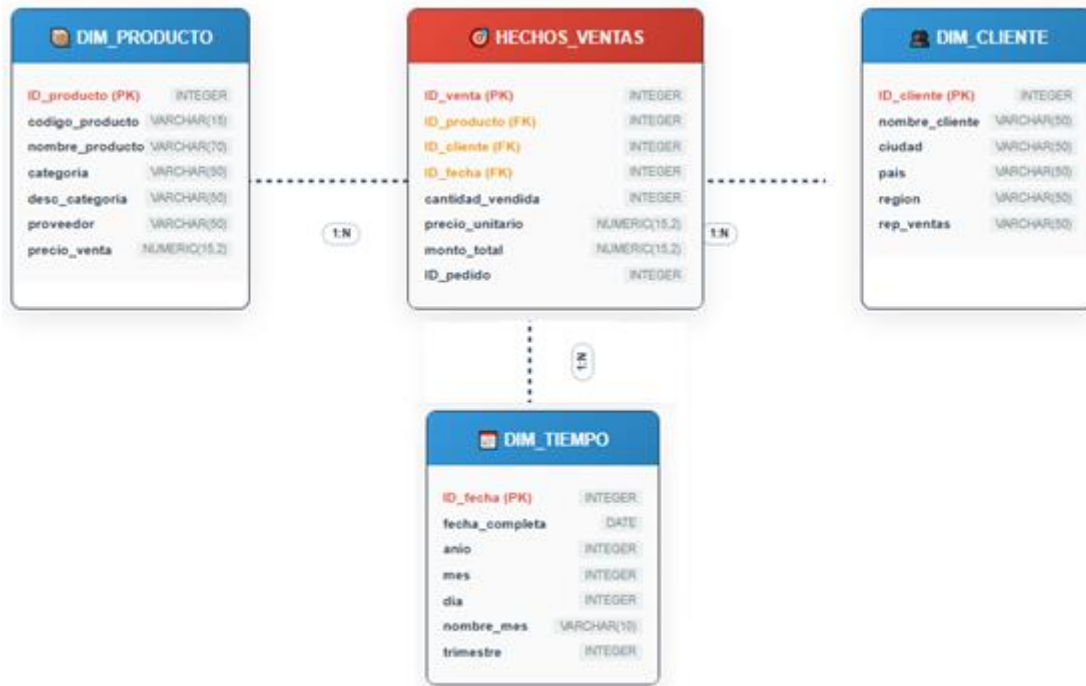
- **detalle_pedido**: Fuente principal para la tabla de hechos (transacciones de venta)
- **pedido**: Proporciona información temporal de las ventas
- **producto**: Contiene información de productos y su categorización
- **Categoría_producto**: Descripción de categorías para análisis
- **cliente**: Información dimensional de clientes

El problema radica en que para hacer cualquier consulta de manera objetiva y eficiente es necesario realizar muchos JOINS lo cual hace que sea demasiado ineficiente y poco escalable. Por lo cual, es necesario aplicar una solución del tipo modelo estrella para facilitar las consultas.

Solución

Se propone un modelo estrella con una tabla de hechos central y tres dimensiones que facilitan el análisis requerido.

El modelo estrella se compone de una tabla de hechos central, Hechos_Ventas, rodeada por tres tablas de dimensiones: Dim_Producto, Dim_Cliente y Dim_Tiempo. Las dimensiones actúan como filtros o agrupaciones para analizar los datos cuantitativos de la tabla de hechos. Este diseño permite responder de forma directa y rápida a las preguntas de negocio planteadas.



Lista de Dimensiones

1. Dim_Producto

Propósito: Almacenar información detallada sobre productos y sus categorías para análisis de rendimiento por producto y categoría.

Campos y Tipos de Datos:

- ID_producto (INTEGER): Clave primaria, identificador único del producto
- codigo_producto (VARCHAR(15)): Código interno del producto
- nombre_producto (VARCHAR(70)): Nombre descriptivo del producto
- categoria (VARCHAR(50)): Descripción de la categoría del producto
- desc_categoria (VARCHAR(50)): Descripción detallada de la categoría
- proveedor (VARCHAR(50)): Nombre del proveedor del producto
- precio_venta (NUMERIC(15,2)): Precio de venta actual del producto

2. Dim_Cliente

Propósito: Contener información demográfica y de contacto de clientes para análisis por segmentación geográfica y representante de ventas.

Campos y Tipos de Datos:

- ID_cliente (INTEGER): Clave primaria, identificador único del cliente
- nombre_cliente (VARCHAR(50)): Nombre completo o razón social del cliente
- ciudad (VARCHAR(50)): Ciudad donde se ubica el cliente
- pais (VARCHAR(50)): País del cliente
- region (VARCHAR(50)): Región geográfica
- rep_ventas (VARCHAR(50)): Nombre del representante de ventas asignado

3. Dim_Tiempo

Propósito: Facilitar análisis temporal de ventas con diferentes niveles de granularidad (día, mes, año, trimestre).

Campos y Tipos de Datos:

- ID_fecha (INTEGER): Clave primaria en formato YYYYMMDD
- fecha_completa (DATE): Fecha completa de la transacción
- anio (INTEGER): Año de la venta
- mes (INTEGER): Mes de la venta (1-12)
- dia (INTEGER): Día del mes (1-31)
- nombre_mes (VARCHAR(10)): Nombre del mes (Enero, Febrero, etc.)
- trimestre (INTEGER): Trimestre del año (1-4)

Tabla de Hechos: Hechos_Ventas

Campos y Tipos de Datos:

- ID_venta (INTEGER): Clave primaria, identificador único de la transacción
- ID_producto (INTEGER): Clave foránea que referencia Dim_Producto
- ID_cliente (INTEGER): Clave foránea que referencia Dim_Cliente
- ID_fecha (INTEGER): Clave foránea que referencia Dim_Tiempo
- cantidad_vendida (INTEGER): Cantidad de unidades vendidas en la transacción
- precio_unitario (NUMERIC(15,2)): Precio de venta por unidad al momento de la transacción
- monto_total (NUMERIC(15,2)): Valor total de la línea de venta (cantidad * precio_unitario)
- ID_pedido (INTEGER): Referencia al pedido original para trazabilidad

Relaciones Definidas:

1. Dim_Producto → Hechos_Ventas: Un producto puede aparecer en múltiples ventas
 - Dim_Producto.ID_producto ← Hechos_Ventas.ID_producto
2. Dim_Cliente → Hechos_Ventas: Un cliente puede realizar múltiples compras
 - Dim_Cliente.ID_cliente ← Hechos_Ventas.ID_cliente
3. Dim_Tiempo → Hechos_Ventas: Una fecha puede tener múltiples transacciones
 - Dim_Tiempo.ID_fecha ← Hechos_Ventas.ID_fecha

Conclusiones

La implementación del modelo estrella propuesto para la base de datos Jardinería representa una solución integral que transforma la estructura relacional normalizada en un esquema dimensional optimizado para análisis de negocio

Ventajas Técnicas:

- Rendimiento optimizado: Las consultas analíticas se ejecutan de manera más eficiente al reducir la cantidad de JOINS necesarios
- Simplicidad conceptual: La estructura estrella es intuitiva y fácil de entender para usuarios de negocio
- Escalabilidad: El modelo puede extenderse fácilmente agregando nuevas dimensiones o métricas

Anexos:

Scripts para crear las dimensiones y la tabla de hechos:

-- Dimensión Producto

```
CREATE TABLE Dim_Producto (  
    ID_producto INTEGER PRIMARY KEY,  
    codigo_producto VARCHAR(15) NOT NULL,  
    nombre_producto VARCHAR(70) NOT NULL,  
    categoria VARCHAR(50) NOT NULL,  
    desc_categoria VARCHAR(50),  
    proveedor VARCHAR(50),  
    precio_venta NUMERIC(15,2)  
);
```

-- Dimensión Cliente

```
CREATE TABLE Dim_Cliente (  
    ID_cliente INTEGER PRIMARY KEY,  
    nombre_cliente VARCHAR(50) NOT NULL,  
    ciudad VARCHAR(50),  
    pais VARCHAR(50),  
    region VARCHAR(50),  
    rep_ventas VARCHAR(50)  
);
```

-- Dimensión Tiempo

```
CREATE TABLE Dim_Tiempo (  
    ID_fecha INTEGER PRIMARY KEY,  
    fecha_completa DATE NOT NULL,  
    anio INTEGER NOT NULL,  
    mes INTEGER NOT NULL,  
    dia INTEGER NOT NULL,  
    nombre_mes VARCHAR(10),  
    trimestre INTEGER  
);
```

-- Tabla de Hechos

```
CREATE TABLE Hechos_Ventas (  
    ID_venta INTEGER IDENTITY(1,1) PRIMARY KEY,  
    ID_producto INTEGER NOT NULL,  
    ID_cliente INTEGER NOT NULL,  
    ID_fecha INTEGER NOT NULL,  
    cantidad_vendida INTEGER NOT NULL,  
    precio_unitario NUMERIC(15,2) NOT NULL,  
    monto_total NUMERIC(15,2) NOT NULL,  
    ID_pedido INTEGER,  
  
    FOREIGN KEY (ID_producto) REFERENCES Dim_Producto(ID_producto),  
    FOREIGN KEY (ID_cliente) REFERENCES Dim_Cliente(ID_cliente),  
    FOREIGN KEY (ID_fecha) REFERENCES Dim_Tiempo(ID_fecha)  
);
```