

**Evidencia de aprendizaje 1. Modelo estrella de un Data Mart**

Cristian Andrei Rendón Alcaraz

Kevin Saldarriaga Vélez

Santiago Agudelo Escobar

**Ingeniería de Software y datos**

Antonio Jesús Valderrama Jaramillo

**Docente**

Bases de Datos II-PREICA2502B010064

**Curso**

**Institución Universitaria Digital de Antioquia**

**Antioquia**

**2025**

## **Evidencia de aprendizaje 1. Modelo estrella de un Data Mart**

### **Introducción**

El siguiente documento propone una explicación de la base de datos sector de jardinería el cual mostrara un análisis de ventas que permitirá responder 4 categorías en especial, optimizando e identificando a su vez el producto mas vendido, la categoría con mas productos y el año con más ventas, todo esto lo complementaremos con la correcta implementación de las tablas y sus relaciones, a su vez los campos necesarios, para la elaboración del modelo estrella, en la que se incluiría la tabla de hechos y dimensiones pertinentes.

### **Objetivos**

Los principales objetivos dentro la elaboración de la base de datos en la cual buscaremos la centralización de datos históricos de ventas, crear un diccionario vivo de productos, la normalización de atributos descriptivos, habilitar análisis multidimensional, facilitar reporte de rentabilidad por producto, con la correcta implementación a un solo clic se podrá acceder a las interrogantes, con sus índices optimizados para búsqueda rápida. La transformación de datos en estrategia, identificando las cifras mas significativas dentro de la jardinería.

## Planteamiento del problema

La base operacional actual presenta:

- Datos fragmentados en múltiples tablas (ej: Pedidos, Detalle\_Pedido, Productos)
- Complejidad para calcular agregaciones temporales (ventas anuales), dificultad de agregación.
- Falta de optimización para reporting analítico
- Requisitos del negocio (reporte diario de productos mas vendidos por categoría, análisis comparativo interanual de ventas, detección de patrones de compra por temporada)

## Análisis del problema

La compañía Jardinería conserva sus datos operativos en una base de datos normalizada, creada para transacciones cotidianas. A pesar de que esta estructura es eficaz para las operaciones, tiene limitaciones importantes en cuanto al análisis estratégico de la información.

- Complejidad en el análisis

Para conseguir información gerencial, se necesitan varias tablas vinculadas entre sí, lo que da lugar a procesos complicados en los que se requieren hasta cinco o cuatro tablas al mismo tiempo para responder preguntas elementales de negocio.

- Desempeño insuficiente

La toma de decisiones se ve afectada por la lentitud en los reportes gerenciales, que requiere procesar varias relaciones entre tablas.

### **Propuesta de la solución**

El modelo estrella diseñado para la base de datos de Jardinería tiene como propósito centralizar y simplificar el análisis de las ventas, permitiendo responder a las principales preguntas de negocio: identificar el producto más vendido, determinar la categoría con mayor número de productos y establecer el año con más ventas.

La tabla de hechos (FAC\_Pedido) constituye el núcleo del modelo, ya que concentra la información transaccional de las ventas. En ella se registran las métricas cuantitativas relevantes como la cantidad de productos vendidos, precio unitario y total de la venta. Además, almacena las claves foráneas que permiten enlazar con las distintas dimensiones de análisis.

En torno a esta tabla de hechos se estructuran las dimensiones, que contienen los atributos descriptivos que enriquecen el análisis:

- DIM\_Producto: incluye información detallada de los productos, como código, nombre, dimensiones, proveedor, descripción, cantidad en stock y precios (venta y proveedor). Esta dimensión es clave para identificar el producto más vendido y analizar la rentabilidad.
- DIM\_Categoria: agrupa los productos por tipo, permitiendo identificar qué categoría concentra más ventas o productos. Contiene el identificador, nombre, descripción en texto y HTML, así como una imagen representativa.
- DIM\_Tiempo: permite realizar análisis temporales de las ventas. Incluye la fecha exacta de cada pedido, así como atributos derivados (año, mes, día, día de la semana). Es fundamental para responder consultas como el año con más ventas o detectar patrones estacionales.
- DIM\_Cliente: enriquece el modelo con datos de los clientes como nombre, contacto, teléfono y ciudad, lo que permite segmentar ventas por ubicación geográfica o cliente recurrente.

Gracias a esta estructura en forma de estrella, el modelo asegura simplicidad en las consultas, integridad referencial entre hechos y dimensiones, y la flexibilidad necesaria para realizar análisis multidimensionales de manera ágil.

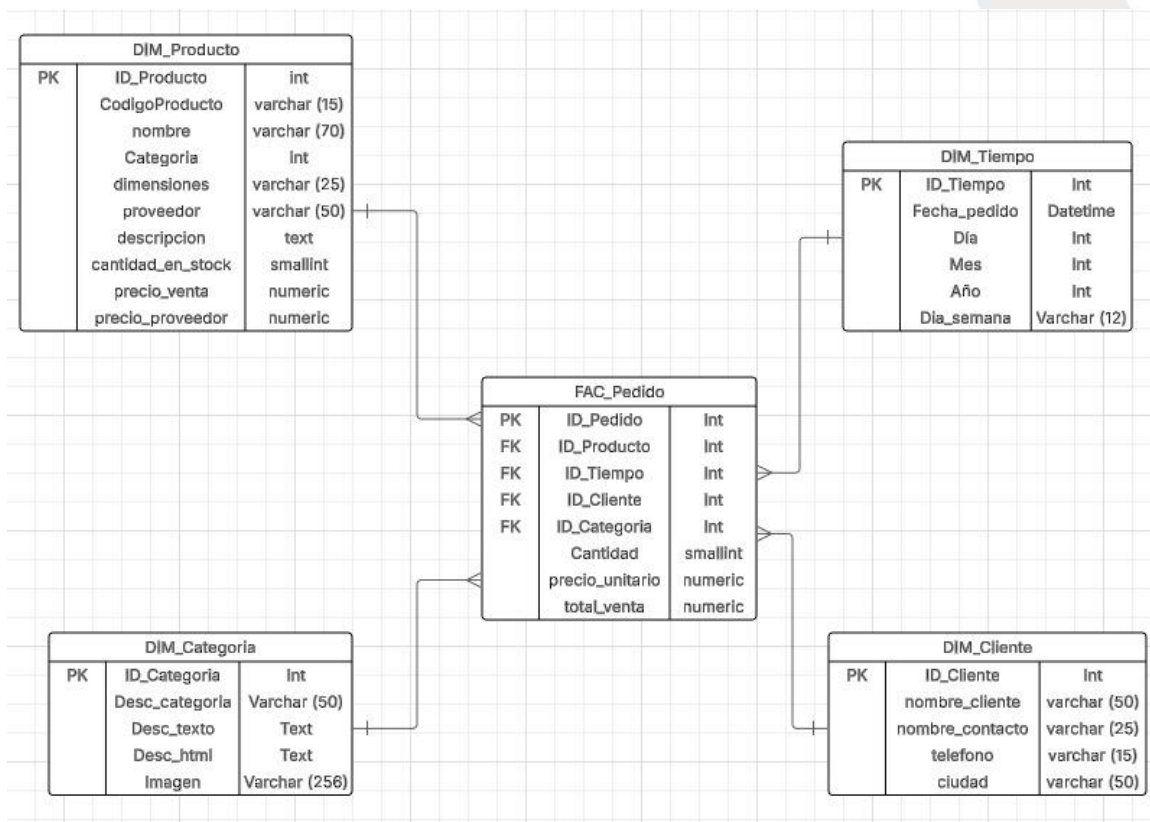


Figura 1.

Modelo estrella de la base de datos Jardinería

Nota. Elaboración propia (2025).

### Listas de dimensiones propuestas

DIM\_Producto - Dimensión de Análisis por Producto

¿Por qué esta dimensión?

Responde directamente al objetivo 1: "identificar el producto más vendido"

Permite desglosar y agrupar las ventas por producto individual.

¿Qué contiene?

Identificadores:

ID\_Producto (PK): Clave única para cada producto

CodigoProducto (varchar 15): Código interno del producto

Atributos descriptivos:

nombre (varchar 70): Nombre del producto - Esencial para mostrar "cuál es el producto más vendido"

dimensiones (varchar 25): Especificaciones físicas

proveedor (varchar 50): Información del proveedor

descripcion (text): Descripción detallada

Atributos comerciales:

cantidad\_en\_stock (smallint): Inventario disponible

precio\_venta (numeric): Precio de venta al público

precio\_proveedor (numeric): Costo de adquisición

DIM\_Categoria - Dimensión de Análisis por Categoría

¿Por qué esta dimensión?

Responde directamente al objetivo 2: "categoría con más productos"

Permite clasificar y agrupar las ventas por tipo de producto.

¿Qué contiene?

Identificador:

ID\_Categoria (PK): Clave única para cada categoría

Atributos descriptivos:

Desc\_categoria (varchar 50): Nombre de la categoría - Fundamental para mostrar "cuál categoría tiene más productos"

Desc\_texto (text): Descripción textual de la categoría

Desc\_html (text): Descripción en formato HTML para web

Imagen (varchar 256): Ruta de imagen representativa

#### DIM\_Tiempo - Dimensión de Análisis Temporal

¿Por qué esta dimensión?

Responde directamente al objetivo 3: "año con más ventas"

Permite analizar tendencias temporales y agrupar ventas por períodos.

¿Qué contiene?

Identificador:

ID\_Tiempo (PK): Clave única para cada fecha

Atributos temporales:

Fecha\_pedido (DateTime): Fecha exacta de la venta

Año (Int): Año de la venta - Crítico para identificar "el año con más ventas"

Mes (Int): Mes de la venta

Día (Int): Día de la venta

Día\_semana (varchar 12): Día de la semana

#### DIM\_Cliente - Dimensión Complementaria

¿Por qué esta dimensión?

Aunque no es requerida directamente por el objetivo, enriquece el análisis permitiendo segmentaciones adicionales por cliente.

¿Qué contiene?

Identificador:

ID\_Cliente (PK): Clave única para cada cliente



Información del cliente:

nombre\_cliente (varchar 50): Nombre de la empresa/cliente

nombre\_contacto (varchar 25): Persona de contacto

telefono (varchar 15): Teléfono de contacto

ciudad (varchar 50): Ubicación del cliente.

### Tabla de Hechos

FAC\_Pedido

¿Por qué esta tabla?

FAC\_Pedido es el corazón del análisis porque contiene las transacciones de venta, que son los datos cuantificables necesarios para responder las tres preguntas.

¿Qué contiene?

Claves de negocio:

ID\_Pedido (PK): Identificador único de cada transacción de venta

ID\_Producto (FK): Conecta con qué producto se vendió

ID\_Categoria (FK): Conecta con la categoría del producto vendido

ID\_Tiempo (FK): Conecta con cuándo se realizó la venta

ID\_Cliente (FK): Conecta con quién compró

Medidas cuantificables:

Cantidad (smallint): Unidades vendidas - Esencial para identificar el "producto más vendido"

precio\_unitario (numeric): Precio por unidad - Permite calcular valores monetarios

total\_venta (numeric): Valor total de la línea - Clave para identificar el "año con más ventas"

¿Por qué es fundamental?

Sin esta tabla no podríamos:

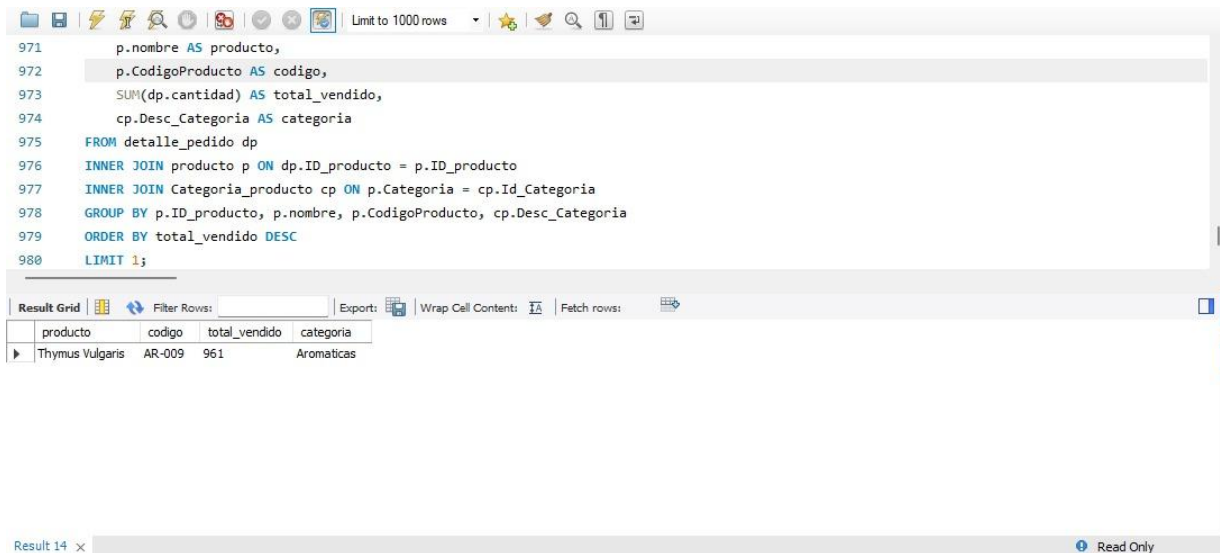
Contar cuántas unidades se vendieron de cada producto

Sumar las ventas totales por año

Relacionar productos con sus categorías en contexto de ventas

## Anexos

Consultas:



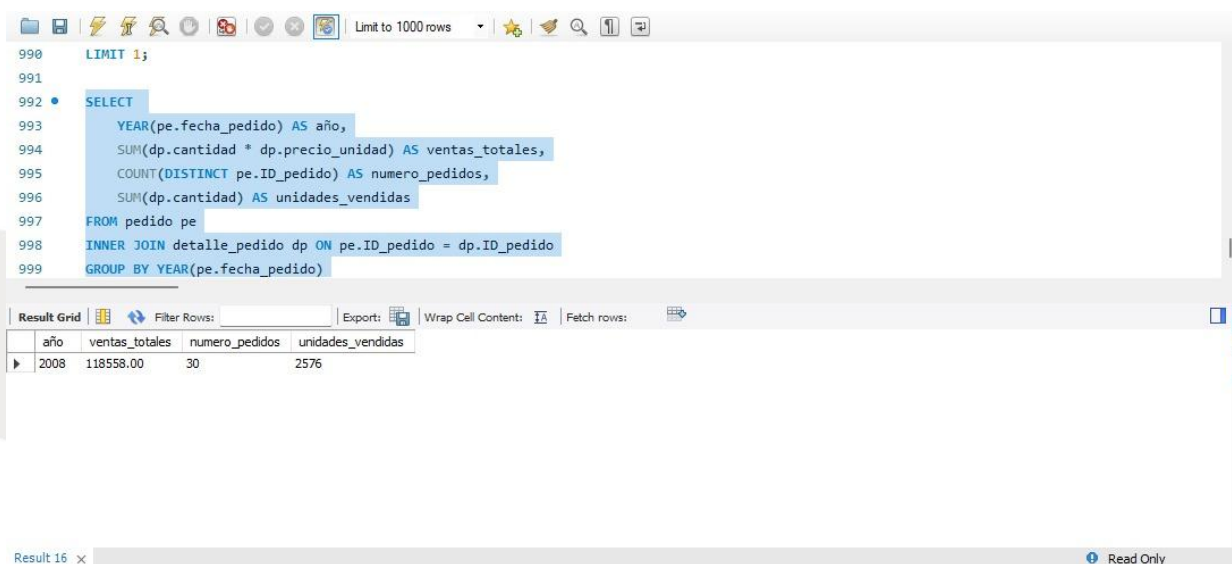
```

971 p.nombre AS producto,
972 p.CodigoProducto AS codigo,
973 SUM(dp.cantidad) AS total_vendido,
974 cp.Desc_Categoria AS categoria
975 FROM detalle_pedido dp
976 INNER JOIN producto p ON dp.ID_producto = p.ID_producto
977 INNER JOIN Categoria_producto cp ON p.Categoria = cp.Id_Categoria
978 GROUP BY p.ID_producto, p.nombre, p.CodigoProducto, cp.Desc_Categoria
979 ORDER BY total_vendido DESC
980 LIMIT 1;
  
```

Result Grid

producto	codigo	total_vendido	categoria
Thymus Vulgaris	AR-009	961	Aromaticas

Result 14 x Read Only



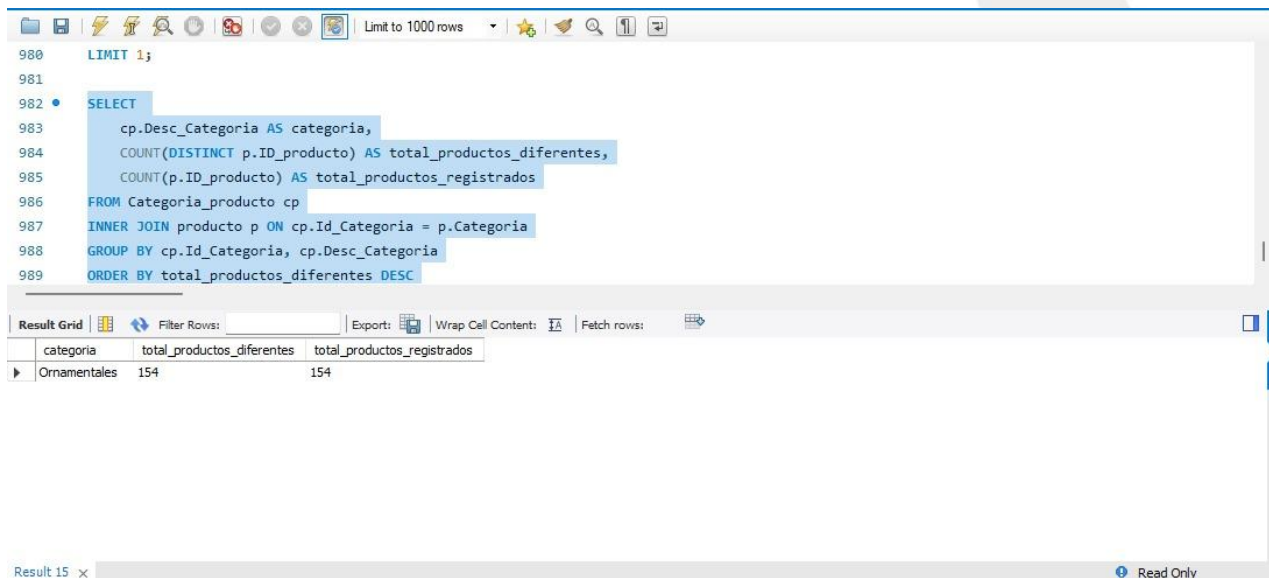
```

990 LIMIT 1;
991
992 SELECT
993     YEAR(pe.fecha_pedido) AS año,
994     SUM(dp.cantidad * dp.precio_unidad) AS ventas_totales,
995     COUNT(DISTINCT pe.ID_pedido) AS numero_pedidos,
996     SUM(dp.cantidad) AS unidades_vendidas
997 FROM pedido pe
998 INNER JOIN detalle_pedido dp ON pe.ID_pedido = dp.ID_pedido
999 GROUP BY YEAR(pe.fecha_pedido)
  
```

Result Grid

año	ventas_totales	numero_pedidos	unidades_vendidas
2008	118558.00	30	2576

Result 16 x Read Only



```

980 LIMIT 1;
981
982 SELECT
983     cp.Desc_Categoria AS categoria,
984     COUNT(DISTINCT p.ID_producto) AS total_productos_diferentes,
985     COUNT(p.ID_producto) AS total_productos_registrados
986 FROM Categoria_producto cp
987 INNER JOIN producto p ON cp.Id_Categoria = p.Categoria
988 GROUP BY cp.Id_Categoria, cp.Desc_Categoria
989 ORDER BY total_productos_diferentes DESC

```

categoria	total_productos_diferentes	total_productos_registrados
Ornamentales	154	154

Result 15 x Read Only

[https://lucid.app/lucidchart/df5602e3-2254-4c48-8c95-8859f34a419a/edit?viewport\\_loc=-230%2C-454%2C2686%2C1254%2C0\\_0&invitationId=inv\\_543b5d84-a536-4df8-acdf-dbdd94ee458a](https://lucid.app/lucidchart/df5602e3-2254-4c48-8c95-8859f34a419a/edit?viewport_loc=-230%2C-454%2C2686%2C1254%2C0_0&invitationId=inv_543b5d84-a536-4df8-acdf-dbdd94ee458a)

## Conclusión

El modelo estrella se desarrolló en un ejercicio integral de entendimiento teórico y aplicación práctica, lo que evidenció la posibilidad de convertir datos dispersos en información estratégica. Este proceso no solamente reforzó la comprensión conceptual acerca de las arquitecturas de data warehousing, sino que además otorgó experiencia concreta en la identificación, elección y organización de componentes esenciales en un entorno empresarial basado en datos.

El análisis exhaustivo de la base de datos original de jardinería reveló la complejidad inherente de los sistemas transaccionales, donde la información se encuentra distribuida en múltiples entidades relacionales. A través de un proceso meticuloso de ingeniería de datos, se procedió a:

- Auditoría Completa de la Base de Datos Fuente
- Diseño Arquitectónico del Modelo Dimensional

### Referencias

Peter-Myers. (s. f.). *Descripción de un esquema de estrella e importancia para Power BI - Power BI*. Microsoft Learn. <https://learn.microsoft.com/es-es/power-bi/guidance/star-schema>

El Dojo De La Informática. (2024, 14 febrero). *Sistemas de Bases de Datos 2 - Clase 01 - Conceptos iniciales de bases de datos* [Video]. YouTube. <https://www.youtube.com/watch?v=88axNwhDcdU>

José Gregorio Maestre. (2022, 2 agosto). *3 CVPBI Como crear Modelos Entidad Relacion como un PRO- modelo estrella con Mysql WorkBench* [Video]. YouTube. <https://www.youtube.com/watch?v=q8dXqrXz77M>