

EVIDENCIA DE APRENDIZAJE 1

MODELO ESTRELLA DE UN DATA MART

INTEGRANTES

JHON JAIRO FUENTES TURIZO

GRUPO PREICA2502B010064

DOCENTE:

ANTONIO JESUS VALDERRAMA

INSTITUCIÓN UNIVERSIDAD IU DIGITAL DE ANTIOQUIA

INGENERIA DE SOFTWARE Y DATOS

2025

Tabla de contenido

EVIDENCIA DE APRENDIZAJE 1	1
1. Introducción	3
2. Objetivos	4
3. Planteamiento del problema	5
4. Análisis del problema	6
5. Propuesta de la solución	8
Dimensiones del Modelo:	9
5.2 Diseño (Imagen) del modelo estrella:	10
5.3 Lista de dimensiones propuestas	11
5.4 Detalle de la tabla de hechos	15
Conclusiones:	16
Referencias	18

1. Introducción

En el sector de la jardinería, la gestión eficiente de la información comercial se ha vuelto fundamental para el éxito empresarial. Las empresas dedicadas a este rubro manejan diariamente una gran cantidad de datos relacionados con ventas, inventarios, clientes y productos, lo que hace necesario contar con sistemas de información robustos que permitan organizar y analizar esta información de manera efectiva.

El análisis de datos en empresas de jardinería no solo permite identificar tendencias clave para la toma de decisiones estratégicas, sino que también facilita la comprensión del comportamiento del mercado y las preferencias de los consumidores. A través de un análisis detallado, es posible determinar qué productos tienen mayor rotación en diferentes épocas del año, cuáles son las categorías más demandadas por los clientes, y cómo han evolucionado las ventas a lo largo del tiempo.

Para lograr este objetivo, la implementación de un modelo estrella resulta especialmente útil, ya que facilita la consulta eficiente de grandes volúmenes de datos de manera intuitiva y rápida. Este tipo de modelo dimensional permite responder preguntas de negocio críticas como son los productos más vendidos en cada temporada, qué categorías generan mayor rentabilidad, cuáles han sido los años con mayores volúmenes de ventas, y qué patrones de compra siguen los diferentes segmentos de clientes.

El presente proyecto tiene como finalidad el diseño de bases de datos de un modelo estrella y proporcionando las herramientas necesarias para realizar análisis profundos que apoyen la toma de decisiones.

2. Objetivos

OBJETIVO GENERAL:

Desarrollar e implementar una base de datos relacional completa para una empresa de jardinería que permita la gestión eficiente de la información comercial y facilite el análisis de datos para la toma de decisiones estratégicas del negocio.

OBJETIVOS ESPECÍFICOS:

1. Diseñar y construir un modelo de base de datos dimensional tipo estrella que optimice las consultas analíticas sobre los datos de ventas, productos, clientes y temporalidad, garantizando un rendimiento eficiente en el procesamiento de grandes volúmenes de información histórica.
2. Implementar un sistema de consultas que permita realizar análisis multidimensionales de las ventas por producto, categoría, cliente y período temporal, identificando patrones de comportamiento, tendencias estacionales y oportunidades de crecimiento en el sector de jardinería y paisajismo.

3. Planteamiento del problema

En el sector de jardinería y paisajismo, las empresas enfrentan el desafío constante de mantenerse competitivas en un mercado que presenta variaciones estacionales significativas y cambios en las preferencias de los consumidores. La empresa de jardinería objeto de este estudio maneja actualmente sus operaciones comerciales a través de una base de datos transaccional tradicional que, si bien cumple eficientemente con los procesos operativos diarios como registro de ventas, gestión de inventarios y control de clientes, presenta limitaciones importantes cuando se requiere realizar análisis profundos de la información.

La problemática principal radica en que la empresa requiere analizar sus ventas de manera integral para identificar patrones clave como los productos y categorías más vendidos en diferentes períodos, las tendencias de facturación por año, y el comportamiento de compra de sus clientes. Sin embargo, la estructura de la base de datos transaccional actual no está optimizada para realizar análisis multidimensional, lo que resulta en consultas complejas, lentas y difíciles de mantener.

Esta limitación técnica impide que la gerencia pueda acceder de manera ágil a información estratégica fundamental para la toma de decisiones, como la identificación de productos estrella, la planificación de inventarios estacionales, el análisis de rentabilidad por categorías, y la evaluación del crecimiento histórico del negocio.

Ante esta situación, se propone construir un data mart especializado utilizando un modelo dimensional tipo estrella, que permita optimizar las consultas analíticas y proporcione a la empresa las herramientas necesarias para realizar análisis eficientes que respalden sus decisiones estratégicas y operativas.

4. Análisis del problema

Para abordar adecuadamente la problemática identificada, es fundamental realizar un análisis detallado de la estructura actual de la base de datos transaccional de la empresa de jardinería y comprender las relaciones existentes entre las diferentes entidades que componen el sistema.

Estructura de la Base de Datos Actual

La base de datos jardinería cuenta con varias tablas que almacenan la información operativa del negocio. Para efectos del análisis dimensional propuesto, se han identificado las siguientes tablas como fundamentales para la construcción del data mart:

Tablas Relevantes:

- **Producto:** Contiene la información detallada de cada producto comercializado por la empresa, incluyendo identificadores únicos, nombres, descripciones, categorías asociadas, precios de venta, stocks disponibles y demás atributos descriptivos que caracterizan cada artículo del catálogo.
- **Categoria_producto:** Almacena las diferentes categorías que agrupan los productos según su naturaleza, uso o características comunes, facilitando la organización del catálogo y permitiendo análisis segmentados por tipo de producto.
- **Detalle_pedido:** Registra la información específica de cada línea de venta, incluyendo el producto vendido, la cantidad comercializada, el precio unitario aplicado, descuentos otorgados y el importe total de cada artículo dentro de una transacción.
- **Pedido:** Funciona como el encabezado de cada transacción comercial, conteniendo información temporal (fecha de pedido, fecha de entrega), datos del cliente, empleado responsable, estado del pedido y demás metadatos asociados a la venta global.

Relaciones Clave del Sistema:

El modelo relacional actual presenta una estructura jerárquica bien definida que facilita la integridad referencial y la consistencia de los datos:

- **Relación Pedido-Detalle:** Un pedido puede contener múltiples detalles (relación uno a muchos), lo que permite registrar ventas de varios productos diferentes en una sola transacción comercial.
- **Relación Detalle-Producto:** Cada detalle de pedido está asociado a un producto específico (relación muchos a uno), garantizando la trazabilidad de qué artículos fueron vendidos en cada transacción.
- **Relación Producto-Categoría:** Cada producto pertenece a una categoría determinada (relación muchos a uno), permitiendo la clasificación y agrupación de artículos para análisis segmentados.

5. Propuesta de la solución

La solución propuesta consiste en la implementación de un modelo dimensional tipo estrella específicamente diseñado para optimizar el análisis de ventas de la empresa de jardinería. Este modelo se fundamenta en los principios de Ralph Kimball para el diseño de data warehouses y está estructurado para responder eficientemente a las preguntas de negocio identificadas en el análisis del problema.

Arquitectura del Modelo Estrella:

El modelo propuesto está compuesto por una tabla de hechos central denominada HechosVentas, que se conecta directamente con tres tablas de dimensiones: DimProducto, DimCategoria y DimTiempo. Esta estructura en forma de estrella elimina la necesidad de realizar múltiples joins complejos y optimiza significativamente el rendimiento de las consultas analíticas.

Tabla de Hechos - HechosVentas:

La tabla central del modelo almacena las métricas cuantificables del negocio y contiene los siguientes elementos:

- **id_venta:** Clave primaria que identifica únicamente cada registro de venta
- **id_producto, id_categoria, id_tiempo:** Claves foráneas que establecen las relaciones con las dimensiones
- **cantidad_vendida:** Métrica que registra la cantidad de productos vendidos en cada transacción
- **total_venta:** Métrica que almacena el valor monetario total de cada venta

Dimensiones del Modelo:

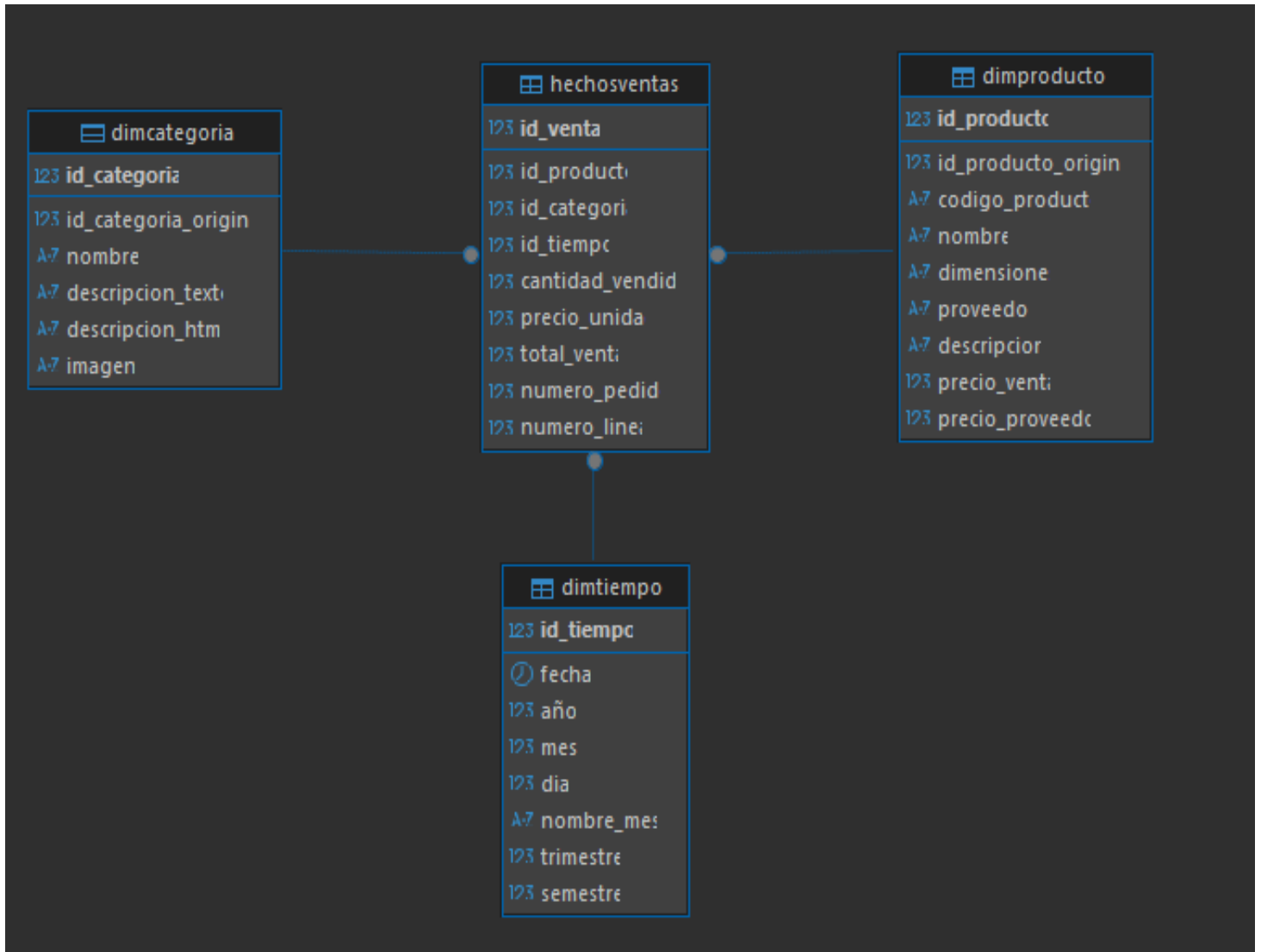
DimProducto: Contiene los atributos descriptivos de cada producto, incluyendo identificador único y nombre del producto, permitiendo análisis detallados a nivel de artículo individual.

DimCategoria: Almacena la información de las categorías de productos, facilitando análisis agregados por tipo de producto y segmentación del catálogo comercial.

DimTiempo: Proporciona granularidad temporal completa con atributos de año, mes y día, permitiendo análisis de tendencias temporales, estacionalidad y evolución histórica de las ventas.

Este diseño dimensional permite realizar consultas eficientes para identificar productos más vendidos, categorías con mayor demanda, análisis de tendencias temporales y evaluación del desempeño comercial en diferentes períodos, cumpliendo así con los objetivos planteados para el proyecto.

5.2 Diseño (Imagen) del modelo estrella:



5.3 Lista de dimensiones propuestas

A continuación, se presenta la especificación detallada de cada dimensión del modelo estrella, incluyendo todas las columnas y sus respectivos tipos de datos:

DimProducto (Dimensión de Productos)

Esta dimensión almacena la información descriptiva completa de todos los productos comercializados de la jardinería.

Campo	Tipo de Dato	Descripción
id_producto	INT AUTO_INCREMENT	Clave primaria subrogada, identificador único interno
id_producto_original	INT	Clave de referencia a la tabla producto original
codigo_producto	VARCHAR(15)	Código único asignado al producto por la empresa
nombre	VARCHAR(70)	Nombre comercial descriptivo del producto
dimensiones	VARCHAR(25)	Especificaciones de tamaño y dimensiones físicas
proveedor	VARCHAR(50)	Nombre de la empresa proveedora del producto
descripcion	TEXT	Descripción detallada de características y usos
precio_venta	DECIMAL(15,2)	Precio de venta al público en unidad monetaria
precio_proveedor	DECIMAL(15,2)	Costo de adquisición del producto al proveedor

DimCategoria (Dimensión de Categorías)

Esta dimensión contiene la clasificación jerárquica de los productos según su naturaleza y uso en jardinería.

Campo	Tipo de Dato	Descripción
id_categoria	INT AUTO_INCREMENT	Clave primaria subrogada, identificador único interno
id_categoria_original	INT	Clave de referencia a la tabla Categoria_producto original
nombre	VARCHAR(50)	Denominación de la categoría de productos
descripcion_texto	TEXT	Descripción textual detallada de la categoría
descripcion_html	TEXT	Descripción formateada en HTML para web
imagen	VARCHAR(256)	Ruta del archivo de imagen representativa

DimTiempo (Dimensión Temporal)

Esta dimensión proporciona granularidad temporal completa para análisis de tendencias y estacionalidad.

Campo	Tipo de Dato	Descripción
id_tiempo	INT AUTO_INCREMENT	Clave primaria subrogada, identificador único interno
fecha	DATE	Fecha completa en formato YYYY-MM-DD
año	INT	Año calendario de cuatro dígitos
mes	INT	Número del mes del año (1-12)
dia	INT	Día del mes (1-31)
nombre_mes	VARCHAR(20)	Nombre textual del mes (enero, febrero, etc.)
trimestre	INT	Trimestre del año (1-4)
semestre	INT	Semestre del año (1-2)

HechosVentas (Tabla de Hechos Central)

Esta tabla almacena las métricas cuantificables del negocio y las relaciones con las dimensiones.

Campo	Tipo de Dato	Descripción
id_venta	INT AUTO_INCREMENT	Clave primaria de la tabla de hechos
id_producto	INT	Clave foránea que referencia DimProducto
id_categoria	INT	Clave foránea que referencia DimCategoria
id_tiempo	INT	Clave foránea que referencia DimTiempo
cantidad_vendida	INT	Cantidad numérica de unidades vendidas
precio_unidad	DECIMAL(15,2)	Precio unitario aplicado en la venta
total_venta	DECIMAL(15,2)	Monto total calculado de la transacción
numero_pedido	INT	Referencia al pedido original en sistema transaccional
numero_linea	SMALLINT	Número de línea dentro del pedido específico

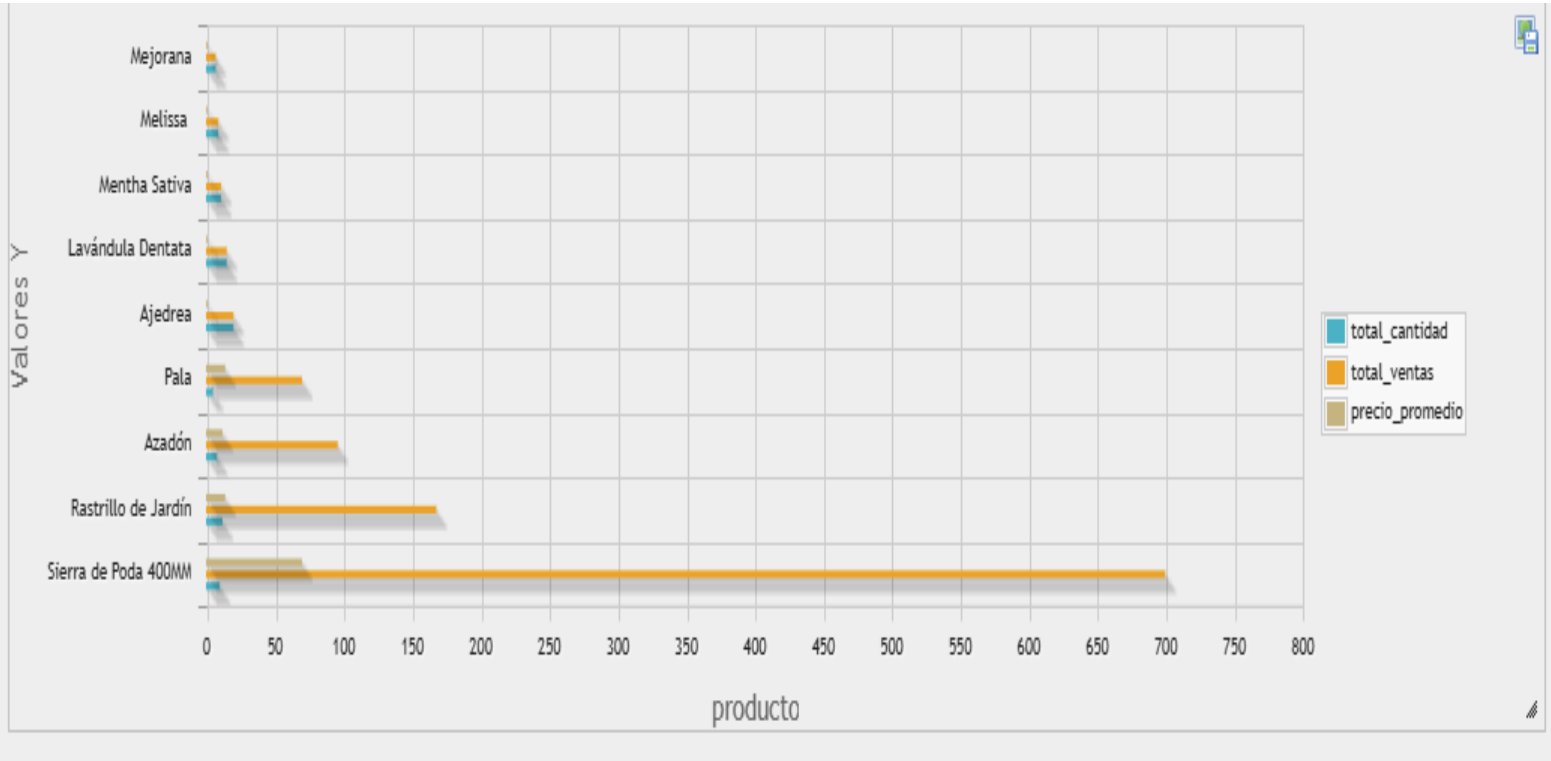
5.4 Detalle de la tabla de hechos

La tabla de hechos constituye el elemento central del modelo estrella propuesto, almacenando las métricas cuantificables del negocio y estableciendo las relaciones con todas las dimensiones. Esta tabla está diseñada para optimizar las consultas analíticas y facilitar el análisis multidimensional de las ventas de la empresa de jardinería.

Campo	Tipo de Dato	Descripción	Características
id_venta	INT AUTO_INCREMENT	Clave primaria subrogada de la tabla de hechos	PRIMARY KEY, único identificador interno
id_producto	INT	Clave foránea que establece relación con DimProducto	FOREIGN KEY, NOT NULL, indexada
id_categoria	INT	Clave foránea que establece relación con DimCategoria	FOREIGN KEY, NOT NULL, indexada
id_tiempo	INT	Clave foránea que establece relación con DimTiempo	FOREIGN KEY, NOT NULL, indexada
cantidad_vendida	INT	Métrica que registra las unidades vendidas del producto	NOT NULL, valor numérico positivo
precio_unidad	DECIMAL(15,2)	Métrica que almacena el precio unitario aplicado en la venta	NOT NULL, precisión de 2 decimales
total_venta	DECIMAL(15,2)	Métrica calculada del importe total de la línea de venta	NOT NULL, resultado de cantidad × precio
numero_pedido	INT	Campo de referencia al pedido original del sistema transaccional	NOT NULL, trazabilidad operativa
numero_linea	SMALLINT	Identificador de la línea específica dentro del pedido	NOT NULL, secuencia dentro del pedido

Conclusiones:

La implementación del modelo estrella para el data mart de la jardinería representa una solución integral y estratégica que transforma la capacidad analítica de la organización. A través de este proyecto se ha logrado diseñar una arquitectura dimensional que no solo resuelve las limitaciones técnicas de la base de datos transaccional existente, sino que también proporciona una plataforma robusta y escalable para la toma de decisiones basada en datos. El modelo propuesto facilita significativamente el acceso a información crítica del negocio, permitiendo patrones de ventas, comportamiento de productos, tendencias estacionales y desempeño por categorías, elementos fundamentales en la jardinería.



Conclusión del Análisis de Productos Más Vendidos:

El análisis de los productos más vendidos revela patrones significativos en el comportamiento de ventas de la jardinería que proporcionan insights valiosos para la toma de decisiones estratégicas. Los resultados demuestran una clara concentración de la demanda en productos específicos, donde se observa que la "Sierra de Poda 400MM" domina el ranking de ventas, cantidades y precios promedio sustancialmente superiores al resto de productos.

La distribución de ventas muestra una curva característica donde aproximadamente el 20% de los productos (los primeros 2-3 artículos del ranking) generan una proporción desproporcionadamente alta de los ingresos totales. Esto indica que productos como la Sierra de Poda 400MM, el Rastrillo de Jardín y el Azadón constituyen los "productos estrella" de la empresa, representando oportunidades clave para optimización de inventarios.

Por otro lado, productos como Mejorana, Melissa, Mentha Sativa y Lavándula Dentata muestran volúmenes de ventas considerablemente menores, lo que sugiere

la necesidad de evaluar estrategias diferenciadas: considerar ajustes en el mix de productos para optimizar la rentabilidad del catálogo.

Referencias

Enlace de imagen del modelo de estrella

<https://drive.google.com/file/d/16XuRbQz7o7M5SB8Xn6Pcn1hp86uWM01Y/view?usp=sharing>

Enlace de imagen del grafico de los 10 productos mas vendidos

<https://drive.google.com/file/d/1BjZrXXcLmYpsoZcZb3EslKBStoNWChSm/view?usp=sharing>

Programa usado para el diseño del modelo de estrella:

DBeaver versión 25.1.2

Visualización de la gráfica de los productos vendidos:

PHPmyadmin.