



IU Digital
de Antioquia



Evidencia de Aprendizaje : Modelo estrella de un data mart

Curso : Bases de datos II

Estudiantes

Juan David Ramírez García

Liz Karen Espinosa Cantor

Oscar Luis Marquez Arrieta

Docente:

Antonio Jesús Valderrama

IU digital de Antioquia

Medellín

31 de Agosto del 2023

ÍNDICE

Introducción.....	4
Objetivos.....	4
Planteamiento del problema.....	4
Análisis del problema.....	5
Propuesta de la solución.....	5
Respondiendo a las preguntas que se plantearon en los objetivos del proyecto.....	9
Conclusiones.....	10
Anexos.....	11
Bibliografía.....	13

Introducción

La idea de este trabajo es aprender a aplicar y a entender la técnica de modelado de datos llamada modelo estrella, sus muchas ventajas a la hora de analizar y recuperar información desde una base de datos, a través de un proyecto práctico en el cual se verá la utilidad de esta técnica. Para ello, partiremos de una base de datos llamada Jardinería, que corresponde a un modelo de tablas-relaciones, las cuales analizaremos y posteriormente obtendremos los componentes principales de un modelo estrella, como su tabla de hechos y tablas de dimensiones, las cuales nos facilitarán organizar los datos de la base de datos Jardinería de manera práctica y eficiente.

La metodología de desarrollo del proyecto corresponde al paso a paso propuesto de la actividad de aprendizaje, en el cual se implementará la construcción del modelo estrella de la siguiente manera: definición de los objetivos, planteamiento y análisis del problema, propuesta de solución que incluye descripción del modelo, listas de dimensiones propuestas y explicada detalladamente, y unas conclusiones que permitan dar respuesta a la idea del trabajo, que es el aprendizaje de implementación de modelo estrella.

Objetivos

- **Objetivo general** :aprender a implementar un modelo estrella para el análisis y evaluación de información de manera más práctica y ágil, a partir de la base de datos de Jardinería.
- **Objetivos específicos:**
 1. Identificar el producto más vendido, la categoría con más productos y el año con más ventas mediante la implementación del modelo estrella.
 2. Construir las dimensiones y la tabla de hechos utilizando métricas, campos y tablas de la base de datos Jardinería.

Planteamiento del problema

El modelo estrella es una estrategia para poder manipular mayores volúmenes de datos distribuidos en múltiples tablas de manera sencilla, rápida y entendible. La base de datos Jardinería hace referencia a ese escenario donde analizar los datos puede ser complejo y lento, debido al volumen de datos, así como esta necesitaría consultas muy extensas con muchos scripts y el uso de muchos join para establecer las relaciones entre las diferentes tablas como cliente, detalle_pedido, producto, categoría_producto, etc. Por lo tanto, ante este contexto vamos a usar el modelado estrella para facilitar el análisis de la información y responder preguntas como: ¿cuál es el producto más vendido?, ¿cuál es la categoría con más productos? ¿cuál es el año con más ventas ?

Análisis del problema

La base de datos de Jardinería contiene información sobre operaciones, la gestión y ventas de productos de jardinería por categorías, como pedidos, detalles de pedido y clientes. Sin embargo, la forma en que está distribuida la información resulta difícil de entender a primera vista, ya que existen muchas tablas con diferentes relaciones entre los datos que tienen en común y para acceder a información en específico es necesario establecer consultas complejas o largas que unan varias tablas a través de sus llaves foráneas y primarias.

Esto último dificulta la respuesta a preguntas sencillas de manera fácil y rápida, como cuál fue el producto más vendido o qué año tuvo más ventas.

Por lo cual, se plantea construir un modelo estrella, que organiza los datos de manera más clara y permite un análisis ágil y fácil de entender.

Propuesta de la solución

Descripción del modelo estrella propuesto:

El modelo estrella que se construyó fue conformado a partir de las tablas y campos de la base de datos jardinería. Para su elaboración se usaron los campos más relevantes que permitieran responder a las preguntas: el producto más vendido, la categoría con más productos y el año con más ventas.

Las tablas y campos usados son:

- De la tabla “*detalle_pedido*” se tomaron los campos “*cantidad*” y “*precio_unidad*” para calcular el total de ventas.
- En la tabla de “*pedido*” se tomó el campo “*fecha_pedido*” para tomar el año y poder establecer el año con más ventas junto a los campos “*cantidad*” y “*precio_unidad*”

mencionados anteriormente.

- De la tabla *producto* se tomaron el “ID_producto”, “nombre”, “dimension”, “descripción” y *categoría* que está relacionada con “Desc_categoria” de la tabla “*Categoría_producto*”, los cuales describen los productos con sus respectivas categorías.
- De la tabla “*categoría*” se tomaron el “ID_categoria”, la “desc_categoria” y la “descripción_texto”, porque ayudan a responder la pregunta de qué categoría tiene más productos.

Diseño modelo estrella

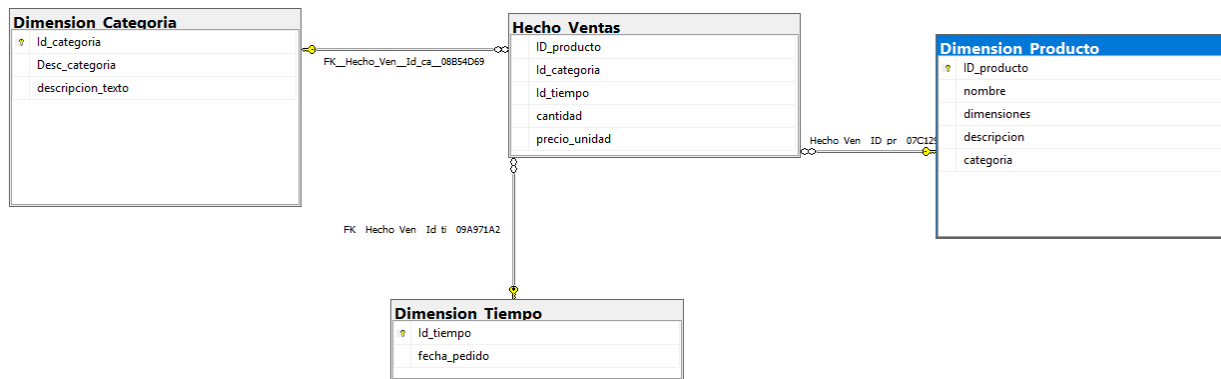


Imagen 1
Realizada en SQL server MANAGAMENT Studio

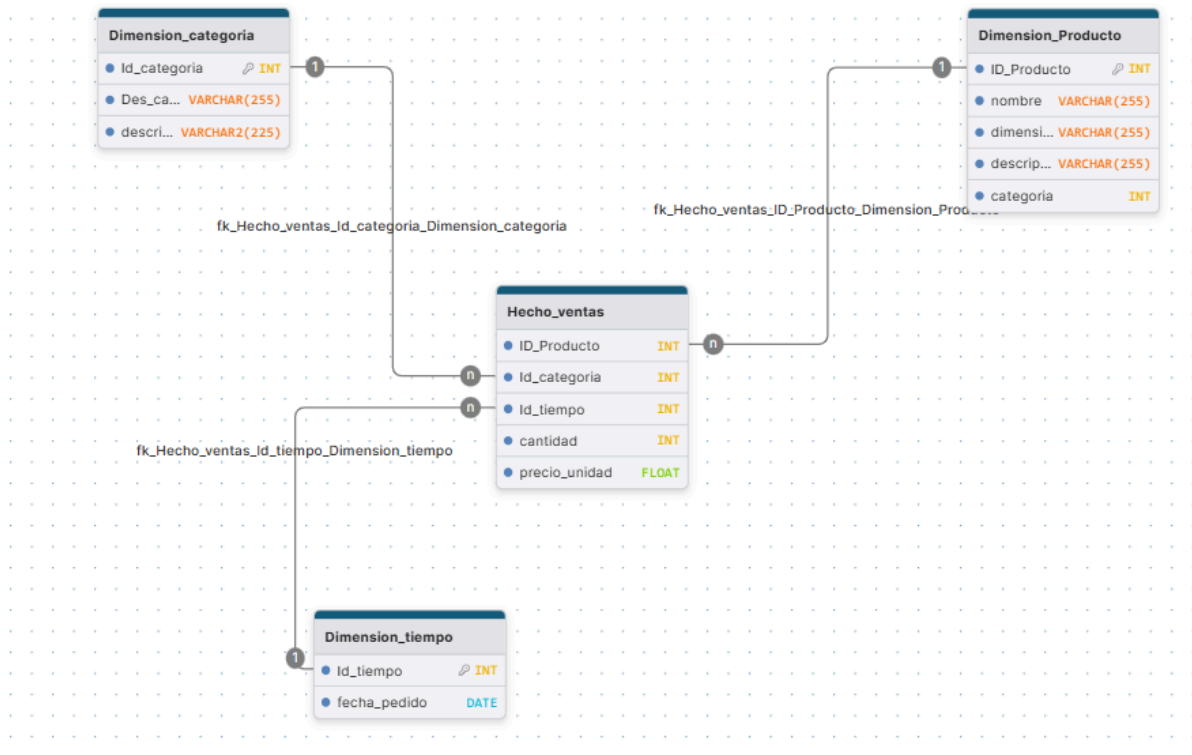


Imagen 2
realizada en drawDB

```
-- Tabla Dimensión Producto
CREATE TABLE Dimension_Producto (
    ID_producto INT PRIMARY KEY,
    nombre NVARCHAR(100),
    dimensiones NVARCHAR(100),
    descripcion NVARCHAR(255),
    categoria INT
);

-- Tabla Dimensión Categoría
CREATE TABLE Dimension_Categoria (
    Id_categoria INT PRIMARY KEY,
    Desc_categoria NVARCHAR(100),
    descripcion_texto NVARCHAR(255)
);

-- Tabla Dimensión Tiempo
CREATE TABLE Dimension_Tiempo (
    Id_tiempo INT PRIMARY KEY,
    fecha_pedido DATE
);

-- Tabla Hechos Ventas
CREATE TABLE Hecho_Ventas (
    ID_producto INT,
    Id_categoria INT,
    Id_tiempo INT,
    cantidad INT,
    precio_unidad DECIMAL(10,2),
    FOREIGN KEY (ID_producto) REFERENCES Dimension_Producto(ID_producto),
    FOREIGN KEY (Id_categoria) REFERENCES Dimension_Categoria(Id_categoria),
    FOREIGN KEY (Id_tiempo) REFERENCES Dimension_Tiempo(Id_tiempo)
);
```

Imagen 3

Script para crear la tablas de hechos y de dimensión del modelo estrella

Lista de dimensiones propuestas.

Dimensión_Producto:

- “ID_producto”: Tipo entero(int) y es PK el identificador único del producto.
- “nombre”: Tipo varchar y contiene los nombres de los productos en este caso, productos de jardinería como pala, rastrillo de jardín, etc.
- “dimensiones”: Tipo varchar y contiene las medidas de cada producto.
- “descripcion”: Tipo texto y contiene la descripción de cada producto.
- “categoria”: Tipo entero que hace referencia a la categoria del producto el cual se describe en la tabla “Dimensión_categoria”

Dimensión_Categoria:

- “Id_categoria”: Tipo entero (int) y es la clave primaria, el identificador único de la categoría.
- “desc_categoria”: Tipo varchar y son los nombres de las categorías como herbáceas y herramientas.
- “descripcion_texto”: Tipo texto y contiene la descripción de cada categoría.

Dimensión_tiempo:

- “Id_tiempo”: Es tipo date y es la llave primaria de la tabla, es el identificador único de cada fecha.
- “fecha_pedido”: Tipo date es la fecha de pedidos de cada producto y permite calcular periodos de ventas.

Hecho_ventas:

- “ID_producto”: Tipo entero (int) y es la llave foránea con la tabla Dimensión_Producto
- “Id_categoria” : Tipo entero (int) y es la llave foránea con la tabla Dimensión_Categoría y permite identificar producto por categoría.
- “Id_tiempo”: Tipo entero (int) y es la llave foránea de la tabla Dimensión_Tiempo y permite identificar pedido de productos por fecha.
- “cantidad”: Tipo entero(int) y establece el número de los diferentes productos
- precio_unidad: Tipo float o decimal y contiene el precio de las unidades que se vendieron.

Respondiendo a las preguntas que se plantearon en los objetivos del proyecto

¿Cuál es el producto más vendido?

El producto más vendido fue el producto de nombre Thymus vulgaris con 2124 unidades vendidas. En la siguiente imagen se comparte la consulta(query) que se realizó para obtener el dato con los resultados que arrojó. Entre ellos podemos destacar el segundo producto más vendido, Tomillo limón, con 905 unidades.


```

select dp.nombre, SUM(hv.cantidad) as total_producto from Dimension_Producto as dp
JOIN Hecho_Ventas as hv on dp.ID_producto=hv.ID_producto
Group by dp.nombre
Order by total_producto Desc;

```

31 %

Results Messages

	nombre	total_producto
1	Thymus Vulgaris	2124
2	Thymus Citriodra (Tomillo limón)	905
3	Rosal bajo 1Aº -En maceta-inicio brotación	846
4	Cerezo	752
5	Robinia Pseudoacacia Casque Rouge	750
6	Chamaerops Humilis	677
7	Petrosilium Hortense (Peregil)	611
8	Trachycarpus Fortunei	599
9	Acer Pseudoplatanus	529
10	Beucarnea Recurvata	450
11	Brahea Armata	424
12	Azadón	423
13	Landora Amarillo, Rose Gaujard bicolor blanco-rojo	400
14	Lagunaria patersonii calibre 8/10	365
15	Mimosa Semilla Bayleyana	347
16	Tuja orientalis "Aurea nana"	320
17	Peral	312
18	Ajedrea	291
19	Lavándula Dentata	278

¿Categoría con más productos?

La categoría con más productos es ornamentales con 46.252 productos.

En la siguiente imagen podemos apreciar la consulta (query) para obtener este resultado. De igual manera destacamos otras categorías como frutales y aromáticas con 18.897 y 623 productos.

```

Select dc.Desc_categoria, SUM(hv.ID_producto) as numero_productos From Dimension_categoria as dc
join HECHO_VENTAS as hv on dc.Id_categoria=hv.Id_categoria
Group by dc.Desc_categoria
Order by numero_productos desc;

```

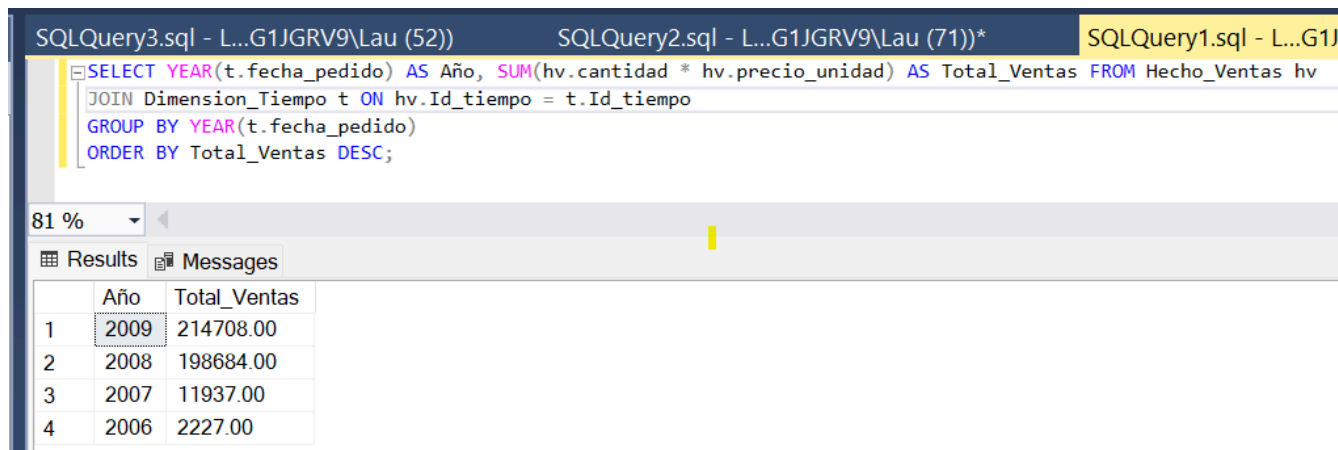
81 %

Results Messages

	Desc_categoria	numero_productos
1	Ornamentales	46252
2	Frutales	18897
3	Aromaticas	623
4	Herramientas	131

¿Cuál es el año con más ventas?

El año con más ventas fue 2009 con un total de 214708 ventas; el segundo año fue 2008 con 198.864 ventas. En la siguiente imagen se comparte la consulta (query) con los datos obtenidos.



The screenshot shows the SQL Server Enterprise Manager interface. At the top, there are three tabs for SQL queries: 'SQLQuery3.sql - L...G1JGRV9\Lau (52))', 'SQLQuery2.sql - L...G1JGRV9\Lau (71))*', and 'SQLQuery1.sql - L...G1J'. The active query is 'SQLQuery1.sql'. The query text is as follows:

```
SELECT YEAR(t.fecha_pedido) AS Año, SUM(hv.cantidad * hv.precio_unidad) AS Total_Ventas FROM Hecho_Ventas hv
JOIN Dimension_Tiempo t ON hv.Id_tiempo = t.Id_tiempo
GROUP BY YEAR(t.fecha_pedido)
ORDER BY Total_Ventas DESC;
```

Below the query, there is a progress bar showing '81 %'. Underneath, there are two tabs: 'Results' and 'Messages'. The 'Results' tab is active, displaying a table with the following data:

	Año	Total_Ventas
1	2009	214708.00
2	2008	198684.00
3	2007	11937.00
4	2006	2227.00

Conclusiones

- La construcción del modelo estrella permitió evidenciar el aprendizaje y entendimiento del mismo de manera práctica. Desde la base de datos de jardinería se identificaron las tablas y campos relevantes para la creación de la tabla de hechos y de dimensiones que componen el modelo, logrando que la información quede más clara, estructurada y accesible para su análisis. De esta forma, se facilita la obtención de respuestas a preguntas clave de los datos.
- Se pudo dar respuesta a las preguntas que se propusieron después de construir el modelo estrella, las cuales son: producto más vendido, la categoría con más productos y el año con más ventas. Esto permitió responder de manera organizada y estructurada las preguntas planteadas, mostrando cómo el modelo facilita el análisis y hace más accesible la información para la toma de decisiones.
- Se cumplió el objetivo de construir las dimensiones y la tabla de hechos a partir de métricas, campos y tablas de la base de datos Jardinería. Con ello, se logró transformar la información en una estructura clara y coherente, lo que evidencia el valor del modelo estrella para organizar los datos y potenciar su análisis.

Anexos

- Enlace donde se construyó el modelo estrella en el editor drawDB:
<https://www.drawdb.app/editor?shareId=b76f0111abeabab29da250060285432b>
- Imágenes de cómo se insertaron los datos en las tablas de dimensión y hechos. Motor de base de datos usado: **microsoft sql server management studio**

SQLQuery3.sql - L...G1JGRV9\Lau (72))* SQLQuery2.sql - L...G1JGRV9\Lau (53))*

```

INSERT INTO Dimension_Categoria (Id_categoria, Desc_categoria, descripcion_texto)
SELECT Id_categoria, Desc_categoria, descripcion_texto FROM categoria_producto;

```

select * from Dimension_Categoria;

82 %

Results Messages

	Id_categoria	Desc_categoria	descripcion_texto
1	1	Herbaceas	Plantas para jardín decorativas
2	2	Herramientas	2 para todo tipo de acción
3	3	Aromaticas	Plantas aromáticas
4	4	Frutales	Árboles pequeños de producción frutal
5	5	Ornamentales	Plantas vistosas para la decoración del jardín

Inserción de datos Dimension_categoria

LAPTOP-5G1JGRV9...ension_Producto* SQLQuery3.sql - L...G1JGRV9\Lau (72))*

```

INSERT INTO Dimension_Producto (ID_producto, nombre, dimensiones, categoria)
SELECT ID_producto, nombre, dimensiones, Categoria FROM producto;

```

select * from Dimension_Producto;

82 %

Results Messages

	ID_producto	nombre	dimensiones	descripcion	categoria
1	1	Sierra de Poda 400MM	0,258	NULL	2
2	2	Pala	0,156	NULL	2
3	3	Rastrillo de Jardín	1,064	NULL	2
4	4	Azadón	0,168	NULL	2
5	5	Ajedrea	15-20	NULL	3
6	6	Lavándula Dentata	15-20	NULL	3
7	7	Mejorana	15-20	NULL	3

Insersión de datos Dimension_Producto

```
INSERT INTO Dimension_Tiempo (Id_tiempo, fecha_pedido)
SELECT DISTINCT ROW_NUMBER() OVER (ORDER BY fecha_pedido) AS Id_tiempo, fecha_pedido FROM pedido;
```

SQLQuery3.sql - L...G1JGRV9\Lau (52))

SQLQuery2.sql - L...G

```
select * from Dimension_Tiempo;
```

81 %

Results Messages

	Id_tiempo	fecha_pedido
1	1	2006-01-17
2	2	2006-05-25
3	3	2007-01-07

Insersión de datos Dimension_tiempo

```
INSERT INTO Hecho_Ventas (ID_producto, Id_categoria, Id_tiempo, cantidad, precio_unidad)
SELECT
    dp.ID_producto,
    p.Categoria AS Id_categoria,
    t.Id_tiempo,
    dp.cantidad,
    dp.precio_unidad
FROM detalle_pedido dp
JOIN producto p ON dp.ID_producto = p.ID_producto
JOIN pedido pe ON dp.ID_pedido = pe.ID_pedido
JOIN Dimension_Tiempo t ON pe.fecha_pedido = t.fecha_pedido;
```

```
select * from Hecho_Ventas;
```

81 %

Results Messages

	ID_producto	Id_categoria	Id_tiempo	cantidad	precio_unidad
1	87	4	1	10	70.00
2	151	5	1	40	4.00
3	165	5	1	25	4.00
4	265	5	1	15	19.00
5	276	5	1	23	14.00
6	57	4	8	3	29.00
7	58	4	8	7	8.00
8	164	5	8	50	4.00

Inserción de datos Tabla Hechos_ventas

Bibliografía

- GeeksforGeeks. (2025, 18 de julio). Difference between fact table and dimension table. Recuperado el 26 de agosto de 2025, de <https://www.geeksforgeeks.org/computer-networks/difference-between-fact-table-and-dimension-table/>
- IBM. (s. f.). Example of a fact table and its associated dimensions. En IBM Documentation. Recuperado el 22 de agosto de 2025, de <https://www.ibm.com/docs/en/tap/5.0.0?topic=tables-example-fact-table-associated-dimensions>