

Modelo Estrella Data Mart

Andres David Mesa Ospina

Nolberto Carmen Gomez

Institución Universitaria Digital De Antioquia

Base De Datos II

Mgtr. Victor Hugo Mercado Ramos

30 marzo de 2025

Introducción

En este documento expondremos el modelo estrella de una base de datos de un negocio de jardinería, compondremos el sistema de metodología del modelo estrella, el cual es de las mas utilizadas para los almacenes de datos (data Warehouse) el cual facilita la optimización de las consultas dinámicas en las empresas, por ello se expondrá un diseño basado en el modelo estrella de la base de datos del proyecto de jardinería. Agregamos la base de datos intermedia que hará el proceso de extracción de los datos y posteriormente la carga de datos para ser subidos a la ser subido al datamart final.

Objetivos

Los objetivos de proyecto es la implementación del modelo estrella en la base de datos existente por ello:

1. Diseñar modelo estrella para el análisis de la información de ventas
2. Definición de las dimensiones de la estructura.
3. Definición de la tabla de hechos, componentes como atributos y datos.
4. Presentación de la estructura de relaciones del sistema
5. Presentación del diagrama visual.
6. Presentación base de datos intermedio JDSTJardineria.
7. Presentación de la base Standing
8. Muestra del proceso de Transformación y carga.

Planteamiento Del Problema

La empresa de jardinería requiere analizar mejor los datos de ventas que se encuentran distribuidos en varias tablas con sus respectivas relaciones, al lograr implementar un modelo estrella se lograra estructurar mejor los datos y centralizarlos logrando consultas y reportes más eficientes.

Descripción Modelo Propuesto

El modelo que proponemos es utilizar la tabla detalle_pedido para utilizarlo como tabla de hechos la cual almacenara las transacciones de las ventas, y usaremos las tablas ya creadas que ya contienen información de los productos, las ventas, los clientes los empleados y las fechas como tablas de dimensiones agregando tambien la dimensión tiempo.

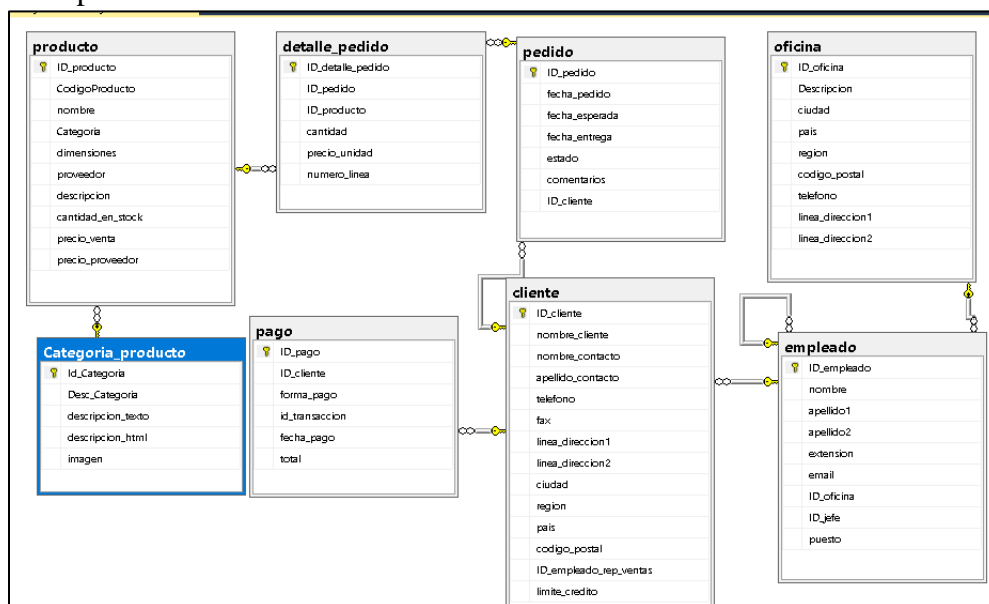


Diagrama Original

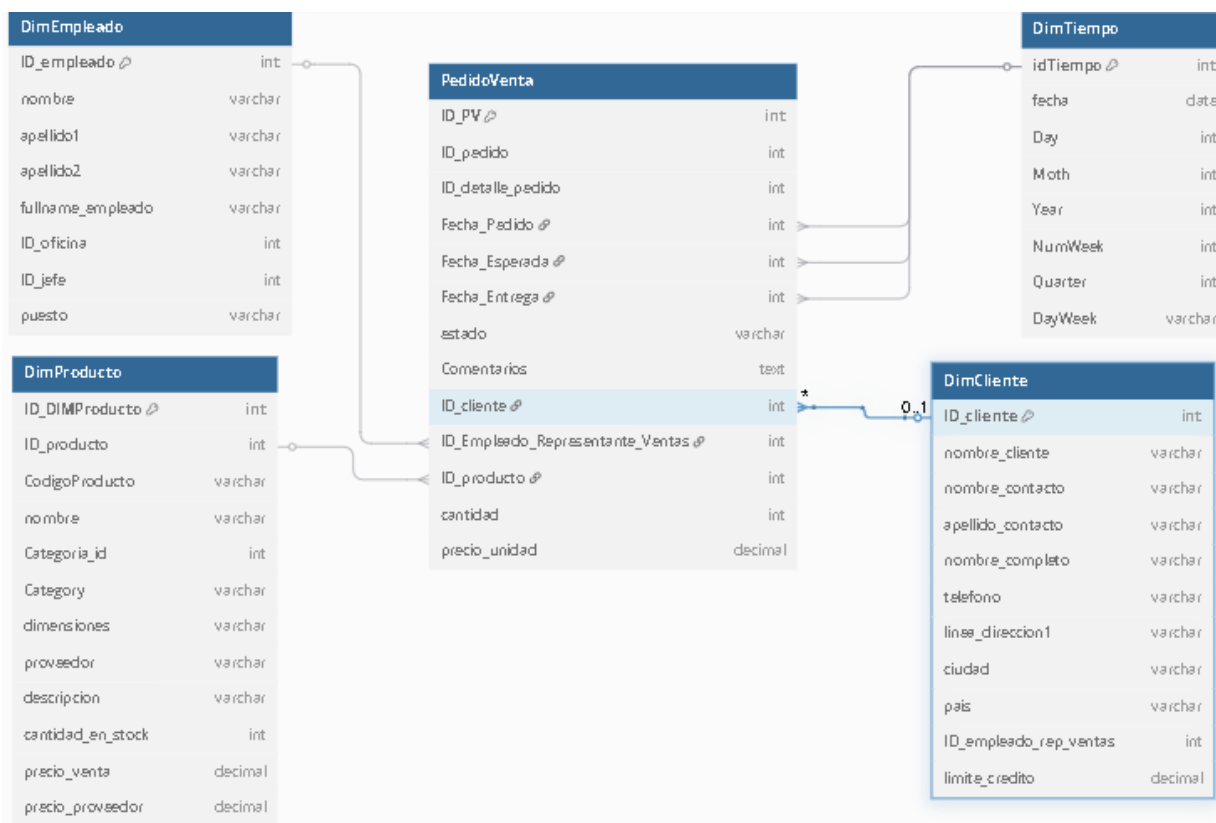


Diagrama Propuesto

Tabla de Hechos

- Pedido Venta:
- ID_PV (INT, PK)
- ID_pedido (INT)
- ID_detalle_pedido (INT)
- ID_producto (INT, FK)
- ID_cliente (INT, FK)


PedidoVenta	
ID_PV	int
ID_pedido	int
ID_detalle_pedido	int
Fecha_Pedido	int
Fecha_Esperada	int
Fecha_Entrega	int
estado	varchar
Comentarios	text
ID_cliente	int
ID_Empleado_Representante_Ventas	int
ID_producto	int
cantidad	int
precio_unidad	decimal

- ID_Empleado_Representante_Ventas (INT, FK)
- cantidad (INT)
- precio_unidad (DECIMAL)
- [Fecha_Pedido] (INT, FK → DimTiempo)
- [Fecha_Esperada] (INT, FK → DimTiempo)
- [Fecha_Entrega] (INT, FK → DimTiempo)
- estado (VARCHAR)
- comentarios (TEXT)

Lista de Dimensiones Propuesta

1. Dimensión Tiempo

- idTiempo (INT, PK)
- fecha (DATE)
- Day (INT)
- Moth (INT)
- Year (INT)

DimTiempo	
idTiempo 	int
fecha	date
Day	int
Moth	int
Year	int
NumWeek	int
Quarter	int
DayWeek	varchar

- NumWeek (INT)
- Quarter (INT)
- DayWeek (VARCHAR)

2. Dimensión Producto

ID_producto (INT, PK)

CodigoProducto (VARCHAR)

nombre (VARCHAR)

Categoria_id (INT)

Category (VARCHAR)

dimensiones (VARCHAR)

proveedor (VARCHAR)

descripcion (TEXT)

cantidad_en_stock (INT)

precio_venta (DECIMAL)

DimProducto	
ID_DIMProducto 	int
ID_producto	int
CodigoProducto	varchar
nombre	varchar
Categoria_id	int
Category	varchar
dimensiones	varchar
proveedor	varchar
descripcion	varchar
cantidad_en_stock	int
precio_venta	decimal
precio_proveedor	decimal

precio_proveedor (DECIMAL)


3. Dimensión Cliente

- ID_cliente (INT, PK)
- nombre_cliente (VARCHAR)
- nombre_contacto (VARCHAR)
- apellido_contacto (VARCHAR)
- nombre_completo (VARCHAR)
- telefono (VARCHAR)
- linea_direccion1 (VARCHAR)
- ciudad (VARCHAR)
- pais (VARCHAR)
- ID_empleado_rep_ventas (INT)
- limite_credito (DECIMAL)

DimCliente	
ID_cliente 	int
nombre_cliente	varchar
nombre_contacto	varchar
apellido_contacto	varchar
nombre_completo	varchar
telefono	varchar
linea_direccion1	varchar
ciudad	varchar
pais	varchar
ID_empleado_rep_ventas	int
limite_credito	decimal

4. Dimensión Empleado

- ID_empleado (INT, PK)
- nombre (VARCHAR)
- apellido1 (VARCHAR)
- apellido2 (VARCHAR)
- fullname_empleado (VARCHAR)
- ID_oficina (INT)
- ID_jefe (INT, FK)
- puesto (VARCHAR)

DimEmpleado	
ID_empleado 	int
nombre	varchar
apellido1	varchar
apellido2	varchar
fullname_empleado	varchar
ID_oficina	int
ID_jefe	int
puesto	varchar

Relaciones

Cada dimensión cuenta con su llave primaria y se relaciona con la tabla

PedidoVenta a través de llaves foráneas, lo que garantiza que cada venta esté asociada a un producto que exista, además de asegurar que no se registren pedidos con clientes o empleados inexistentes. Esto proporciona datos íntegros en las consultas respectivas y permite realizar análisis confiables sobre las ventas, los tiempos de entrega, la relación con los clientes y el desempeño de los empleados.

Base de datos Intermedia

Se realiza en visual Studio 2022 la creación de dos bases de datos previamente creadas en SQL Server Management Studio, en donde se agregaron al administrador de conexión dentro del entorno de integración de servicios SSIS, dentro del flujo de paquetes se procedió a configurar la extracción de los datos Jardinería los cuales serán posteriormente transformados, por el momento se crea un origen de datos que apunta a la base de datos original y el destino el cual corresponde a la base de datos “Stanging” en donde se almacenaran los datos temporalmente.

Durante este proceso, se definieran únicamente los campos relevantes para el modelo estrella propuesto, asegurando la carga de datos eficiente y controlada, el proceso forma parte del proceso de transformación.

Carga De Datos

Una vez completada la Transformación aplicando también normalización y condicionales se realiza el proceso de carga de datos desde la base de datos Stanging hacia las tablas del Data Mart diseñada bajo el esquema del modelo estrella, en esta fase los datos anteriormente transformados se dirigen a las tablas de dimensiones y de hechos, manteniendo la integridad mediante la correcta asignación de claves primarias y foráneas, la carga se realiza utilizando componentes del flujo logrando que los registros coincidan con los formatos y estructuras definidas para el datamart, y por medio de truncates evitando duplicados o datos inconsistentemente, logrando que se sincronicen adecuadamente las dimensiones; Este proceso permite consolidar la información para analizar futuros reportes, completando todo el proceso de ETL, desde el inicio al final en la base de datos original Jardinería a Stanging.

Conclusiones

El modelo estrella es un esquema de almacenamiento que organiza los datos mediante una tabla central (tabla de hechos) rodeada de tablas de dimensiones. Su estructura simple y eficiente permite segmentar la información de manera estratégica, optimizando el proceso de consulta y análisis.

En nuestro esquema, el modelo propuesto mejora el almacenamiento y facilita la generación de consultas rápidas y precisas, lo que permite responder con agilidad a informes clave en las empresas, como:

- ¿Cuáles son los mejores clientes?
- ¿Qué vendedor genera mayor facturación?
- ¿Qué productos tienen mayor volumen de ventas?

Gracias a esta optimización, se mejora el rendimiento en la entrega de datos, permitiendo una toma de decisiones más informada y efectiva.

Anexos

Código DBML:

```
Table DimEmpleado {  
    ID_empleado int [pk]  
    nombre varchar  
    apellido1 varchar  
    apellido2 varchar  
    fullname_empleado varchar  
    ID_oficina int  
    ID_jefe int  
    puesto varchar  
}  
  
Table DimCliente {  
    ID_cliente int [pk]  
    nombre_cliente varchar  
    nombre_contacto varchar  
    apellido_contacto varchar  
    nombre_completo varchar  
    telefono varchar  
    linea_direccion1 varchar  
    ciudad varchar  
    pais varchar  
    ID_empleado_rep_ventas int  
    limite_credito decimal
```

```
}
```

```
Table DimProducto {  
    ID_DIMProducto int [pk]  
    ID_producto int  
    CodigoProducto varchar  
    nombre varchar  
    Categoria_id int  
    Category varchar  
    dimensiones varchar  
    proveedor varchar  
    descripcion varchar  
    cantidad_en_stock int  
    precio_venta decimal  
    precio_proveedor decimal  
}
```

```
Table DimTiempo {  
    idTiempo int [pk]  
    fecha date  
    Day int  
    Moth int  
    Year int  
    NumWeek int  
    Quarter int  
    DayWeek varchar  
}
```

```

Table PedidoVenta {
    ID_PV int [pk]
    ID_pedido int
    ID_detalle_pedido int
    Fecha_Pedido int
    Fecha_Esperada int
    Fecha_Entrega int
    estado varchar
    Comentarios text
    ID_cliente int
    ID_Empleado_Representante_Ventas int
    ID_producto int
    cantidad int
    precio_unidad decimal
}

```

Ref: PedidoVenta.ID_cliente > DimCliente.ID_cliente

Ref: PedidoVenta.ID_Empleado_Representante_Ventas >
DimEmpleado.ID_empleado


Ref: PedidoVenta.ID_producto > DimProducto.ID_producto

Ref: PedidoVenta.Fecha_Pedido > DimTiempo.idTiempo

Ref: PedidoVenta.Fecha_Esperada > DimTiempo.idTiempo

Ref: PedidoVenta.Fecha_Entrega > DimTiempo.idTiempo

Bibliografía

1. dbdiagram.io. (s.f.). dbdiagram.io. Recuperado el 25 de febrero de 2025, de <https://dbdiagram.io/d>
2. Verne Academy. (s.f.). ¿Qué es Data Warehouse y Data Marts? Diferencias y diseño del modelo Recuperado el 25 de febrero de 2025, de <https://verneacademy.com/blog/articulos-data/que-es-data-warehouse-data-marts-diferencias/>
3. Yoti. (2022, 10 de octubre). Tabla de hechos y dimensiones | Modelo Estrella 
[DATA MART] [Video]. YouTube.
<https://www.youtube.com/watch?v=HvzO18fUjqY>