

Materia: Electiva II Bases de datos II

Docente: Victor Hugo Mercado

Módulo 1

Modelo estrella de un Data Mart



Grupo 5:

- **Carlos Andrés Bustamante Montoya**
- **Santiago Osorio Henao**

Especialización en Analítica y Big Data

Medellín, 16 de Noviembre de 2024

Diseño de un Modelo Estrella para el Análisis de Ventas en la Empresa "Jardinería"

Introducción

En un mercado altamente competitivo, las empresas deben ser capaces de analizar sus datos de manera eficiente para tomar decisiones informadas. En este sentido, un modelo estrella es una estructura ideal para el análisis de múltiples variables, ya que permite organizar los datos de manera que se facilite la consulta rápida de métricas clave. Este trabajo propone el diseño de un modelo estrella utilizando la base de datos "Jardinería", lo cual permitirá realizar análisis profundos sobre las ventas, identificando patrones y facilitando decisiones estratégicas basadas en datos. La estructura propuesta ofrece una visión detallada de los productos, clientes y fechas de transacciones, entre otros aspectos, organizados en torno a una tabla de hechos.

Objetivos

- Diseñar un modelo de datos en estrella para la base de datos de ventas de la empresa "Jardinería" que facilite el análisis de los datos de manera rápida y optima.
- Identificar y definir la tabla de hechos y las dimensiones clave para el modelo estrella.
- Proporcionar un diseño gráfico del modelo, incluyendo relaciones y tipos de datos.
- Detallar cada dimensión con sus campos específicos y tipos de datos.
- Asegurar que la estructura propuesta cumpla con la integridad referencial entre las tablas.

Planteamiento del problema

La base de datos "Jardinería" actualmente contiene información detallada sobre clientes, empleados, productos, y transacciones, pero no está optimizada para realizar análisis de ventas de manera eficiente limitando su capacidad para obtener información detallada sobre las transacciones, productos, clientes y empleados involucrados. La implementación de un modelo en estrella permitirá una organización más intuitiva y eficiente de los datos, adecuada para consultas de análisis.

Los responsables de la toma de decisiones requieren una herramienta que les permita conocer de forma rápida y precisa los productos más vendidos, las categorías que generan más ingresos y el comportamiento de las ventas por año. Para resolver este problema, se necesita estructurar los datos de manera que se facilite el acceso a métricas clave a través de un modelo de datos adecuado.

Análisis del problema

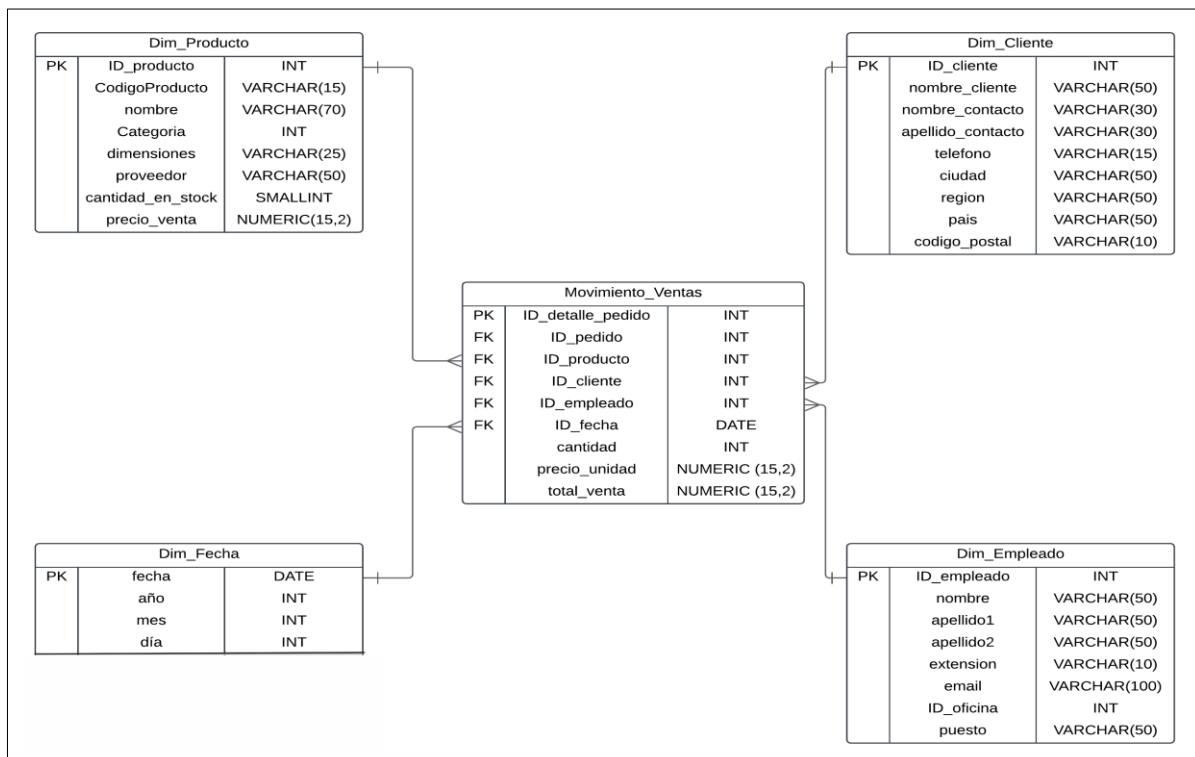
El principal desafío radica en la estructura actual de la base de datos, que almacena la información de forma relacional pero no está optimizada para consultas analíticas rápidas. Sin un modelo bien estructurado, los usuarios enfrentan dificultades para obtener respuestas rápidas sobre el desempeño de productos, categorías y periodos específicos. La implementación de un modelo estrella puede ayudar a organizar la información de manera eficiente y facilitar el análisis de los datos, lo cual es fundamental para una correcta toma de decisiones estratégicas.

Propuesta de solución

Descripción de modelo de estrella propuesto

El modelo estrella para "Jardinería" está centrado en la tabla de hechos llamada **"Movimiento_Ventas"**, que representa cada transacción de venta de productos. Esta tabla se relaciona con dimensiones que describen los detalles contextuales de cada venta, como el producto, el cliente, el empleado que gestionó la venta y la fecha en que se realizó la transacción. Cada dimensión almacena detalles específicos que permiten analizar las ventas desde diferentes perspectivas.

Diseño del modelo estrella



Lista de dimensiones propuestas

1. Dimensión de Producto (dim_producto)

Descripción: Proporciona detalles de cada producto vendido.

Columnas y tipos de datos:

- **ID_producto (INT)** - Llave primaria.
- **CodigoProducto (VARCHAR(15))** - Código único del producto.
- **nombre (VARCHAR(70))** - Nombre descriptivo del producto.
- **Categoria (INT)** - Identificador de la categoría del producto.
- **dimensiones (VARCHAR(25))** - Dimensiones del producto.
- **proveedor (VARCHAR(50))** - Nombre del proveedor del producto.
- **cantidad_en_stock (SMALLINT)** - Cantidad disponible en inventario.
- **precio_venta (NUMERIC(15,2))** - Precio de venta del producto.

2. Dimensión de Cliente (dim_cliente)

Descripción: Contiene información sobre los clientes que realizan las compras.

Columnas y tipos de datos:

- **ID_cliente (INT)** - Llave primaria.
- **nombre_cliente (VARCHAR(50))** - Nombre de la empresa o cliente.
- **nombre_contacto (VARCHAR(30))** - Nombre del contacto principal.
- **apellido_contacto (VARCHAR(30))** - Apellido del contacto.
- **telefono (VARCHAR(15))** - Teléfono de contacto.
- **ciudad (VARCHAR(50))** - Ciudad de residencia.
- **region (VARCHAR(50))** - Región de residencia.
- **pais (VARCHAR(50))** - País de residencia.
- **codigo_postal (VARCHAR(10))** - Código postal.

3. Dimensión de Empleado (dim_empleado)

Descripción: Registra información sobre los empleados que gestionan las ventas.

Columnas y tipos de datos:

- **ID_empleado (INT)** - Llave primaria.
- **nombre (VARCHAR(50))** - Nombre del empleado.
- **apellido1 (VARCHAR(50))** - Primer apellido.
- **apellido2 (VARCHAR(50))** - Segundo apellido.
- **extension (VARCHAR(10))** - Extensión telefónica.
- **email (VARCHAR(100))** - Correo electrónico.

- **ID_oficina (INT)** - Identificador de la oficina.
- **puesto (VARCHAR(50))** - Puesto del empleado.

4. Dimensión de Fecha (dim_fecha)

Descripción: Permite análisis mediante el desglose de la fecha de venta.

Columnas y tipos de datos:

- **fecha (DATE)** - Llave primaria.
- **año (INT)** - Año de la fecha.
- **mes (INT)** - Mes de la fecha.
- **día (INT)** - Día de la fecha.

Tabla Hechos

La tabla de hechos “**Movimiento_Ventas**” almacena las transacciones de venta

- **ID_detalle_pedido (INT (PK))**- Identificador único de la transacción de venta.
- **ID_pedido (INT)**- Relación con la tabla de pedidos.
- **ID_producto (INT)**- Identificador del producto vendido.
- **ID_cliente (INT)**- Identificador del cliente.
- **ID_empleado (INT)**- Identificador del empleado de ventas.
- **ID_fecha (DATE)**- Fecha de la venta.
- **Cantidad (INT)**- Cantidad de producto vendida.
- **precio_unidad (NUMERIC(15,2))**- Precio de venta unitario.
- **total_linea (NUMERIC(15,2))**- Total de la línea (cantidad * precio_unidad).

Conclusión

El modelo estrella propuesto permite a la empresa "Jardinería" mejorar el análisis de sus ventas, ofreciendo un marco robusto para realizar consultas más rápidas y optimas. Este diseño optimiza la capacidad de la empresa para identificar patrones en los datos de ventas, mejorando la toma de decisiones a través de una estructura eficiente, clara y orientada al análisis de grandes volúmenes de datos.

Además, responde de manera efectiva al objetivo de identificar el producto más vendido, la categoría con más productos vendidos y el año con el mayor volumen de ventas.

En resumen, este modelo estrella optimizado no solo mejora la estructura y eficiencia del análisis, sino que también facilita la obtención de datos clave, permitiendo a la empresa de jardinería tomar decisiones informadas y estratégicas basadas en datos precisos y bien organizados.