

Diseño de Modelo Estrella para la base de datos "Jardinería"

Evidencia de aprendizaje 1

Daniela Ceferino Marín

Laura Victoria Navarro Arriola

Alexander Zapata Rico

Institución Universitaria Digital de Antioquia

Facultad de Ingeniería y Ciencias Agropecuarias

Especialización en Analítica y Big Data

Electiva II: Base de datos II

Docente: Victor Hugo Mercado

Semestre 2024-2 Bloque 2

TABLA DE CONTENIDO

Diseño de Modelo Estrella para la base de datos Jardinería.....	3
Introducción.....	3
Objetivos.....	3
Planteamiento del problema	4
Análisis de la problemática.	5
Propuesta de la solución	6
Descripción del modelo estrella propuesto.....	6
Diseño (Imagen) del modelo estrella donde se puedan observar las dimensiones, la tabla de hechos, sus campos, tipos de datos y relaciones.	6
Lista de dimensiones propuestas.....	7
Cada dimensión debe tener especificadas las columnas y qué tipos de datos van a almacenar.	9
Detalla la tabla de hechos, con sus campos y tipos de datos.	9
Conclusiones	9
Referencias	10

Diseño de Modelo Estrella para la base de datos Jardinería

Introducción

Toda empresa y/o organización necesita una estructura de base de datos optimizada para analizar sus transacciones de manera efectiva en sus diversas áreas. Es por ello que resulta necesario plantear el orden y estructura de la información para que al momento de desarrollar sus aplicaciones para los diferentes fines que se creen se pueda obtener la información en corto tiempo, estructurada, y fiel a los procesos que ocurren en tiempo real. De ahí el uso de modelos estructurados de data que parten de la selección de un dominio y unas entidades y relaciones que hagan una adecuada asociación entre ellas para realizar todo tipo de consultas.

En el proceso de aprendizaje académico, se ha tenido la oportunidad de interactuar y mediante la práctica comprender una serie de recursos fundamentales para el procesamiento, almacenamiento y modelado de datos, donde inicialmente se plantea el conocimiento de un modelo relacional que es útil para el almacenamiento pero dificulta la generación de reportes, rendimiento en recursos al hablar de una arquitectura de modelado y eficiencia en la posibilidad de incurrir en la redundancia de datos y repetición de los mismos. Por lo tanto, se hace interesante conocer otras metodologías y aquí es donde se introduce la necesidad de construir un modelo de datos dimensional a partir de una estructura relacional que permite análisis más eficientes, menor consumo de recursos en hardware y la implementación ligera de consultas.

Existen dos modelos dimensionales muy usados: el modelo estrella y el modelo copo de nieve, cada uno tiene diversas implicaciones, pero el más usado y el que será objeto de este trabajo es el modelo estrella, que permita el análisis de datos y transacciones facilitando la toma de decisiones mediante un esquema de datos centralizado y optimizado para consultas analíticas. (Sánchez, 2023)

Este proyecto aborda la reestructuración de un modelo datos, pasando de un esquema relacional a uno dimensional desnormalizado adecuado para el análisis de datos de la base de datos jardinería.

Objetivos

1. **Identificar las tablas y relaciones relevantes** en la base de datos "Jardinería" que permitan capturar toda la información pertinente a las ventas y transacciones de la empresa.
2. **Construir un modelo estrella** que centre los datos de ventas en una tabla de FACTS y proporcione tablas dimensionales adecuadas que enriquezcan la información de cada transacción.
3. **Diseñar la tabla de FACTS y las dimensiones** para representar el proceso de ventas, estableciendo relaciones que aseguren la integridad referencial y permitan consultas rápidas y eficaces.
4. **Generar un sistema de análisis** que permita identificar el producto más vendido.
Determinar la categoría con más productos e identificar el año con más ventas.

Planteamiento del problema

La empresa enfrenta problemas en la generación de reportes y análisis sobre ventas y transacciones debido a la estructura de su base de datos, actualmente diseñada bajo un modelo relacional ya que es una forma estándar de representar y consultar datos (¿Qué es una Base de Datos Relacional?, s. f.). Este modelo es eficiente para el almacenamiento y gestión transaccional, pero no está optimizado para el análisis de datos, lo que resulta en consultas un poco más complejas debido a la necesidad de implementar una gran cantidad de los joins que van ampliando el recurso de almacenamiento. Ante esta situación, es importante contar con una estructura de datos que permita realizar consultas de forma ágil y obtener insights valiosos de manera oportuna para apoyar la toma de decisiones.

Por lo tanto, se requiere transformar el modelo relacional de la empresa, migrando a un modelo dimensional ya que son estructuras desnormalizadas diseñadas para recuperar datos de un almacén de datos (Datamart) y utilizan tablas de hechos y dimensiones para mantener un registro de datos históricos en almacenes de datos (Datawarehouse) (Ahmed, 2024). Este nuevo enfoque permitirá centralizar y optimizar los datos de las ventas de la empresa de Jardinería en una estructura adecuada mejorando el acceso a la información y facilitando la extracción de indicadores claves por medio de la implementación de un modelo estrella.

Análisis de la problemática.

La base de datos “Jardinería” de la empresa, está organizada bajo un modelo relacional, lo cual implica que las tablas se encuentran normalizadas para asegurar una consistencia y reducir

redundancia. Sin embargo, esta organización resulta poco eficaz al momento de la generación de reportes analíticos en cuanto a las ventas:

- La información sobre las ventas se encuentra distribuida en varias tablas como: pedido, detalle de pedido, producto, categoría de producto, oficina, pago, entre otras y están relacionadas mediante claves lo que implica realizar múltiples uniones o joins en las consultas.
- Estas consultas ralentizan el proceso de generación de reportes y dificultan el acceso de métricas estratégicas para la empresa como son las ventas y sus analíticas.

Para optimizar el análisis de datos y la toma de decisiones, es necesario implementar el modelo dimensional ya que se organiza la tabla de hechos (Facts) que contiene los datos cuantitativos de ventas y varias tablas de dimensiones que contiene los atributos descriptivos relacionados: Dim Pedido, Dim Producto, Dim Cliente y Dim Empleado. Esto con el fin de centralizar la información de ventas, facilitar el análisis de los datos y la reducir la complejidad en las consultas, así pues, esta estructura le permitirá a la empresa acceder a información relevante de manera más rápida y fortalecer la toma de decisiones estratégicas.

Propuesta de la solución

Descripción del modelo estrella propuesto.

El modelo estrella que se propone se define a partir de la creación de una tabla FACT Ventas, la cual es la base de todas las consultas cuantitativas hacia las dimensiones (descriptivas-cualitativas) que en consenso grupal se decidieron tuvieran como columnas aquellas que tengan un aporte relevante y sean las variables insumo para comprender los indicadores necesarios para un cierre exitoso de ventas.

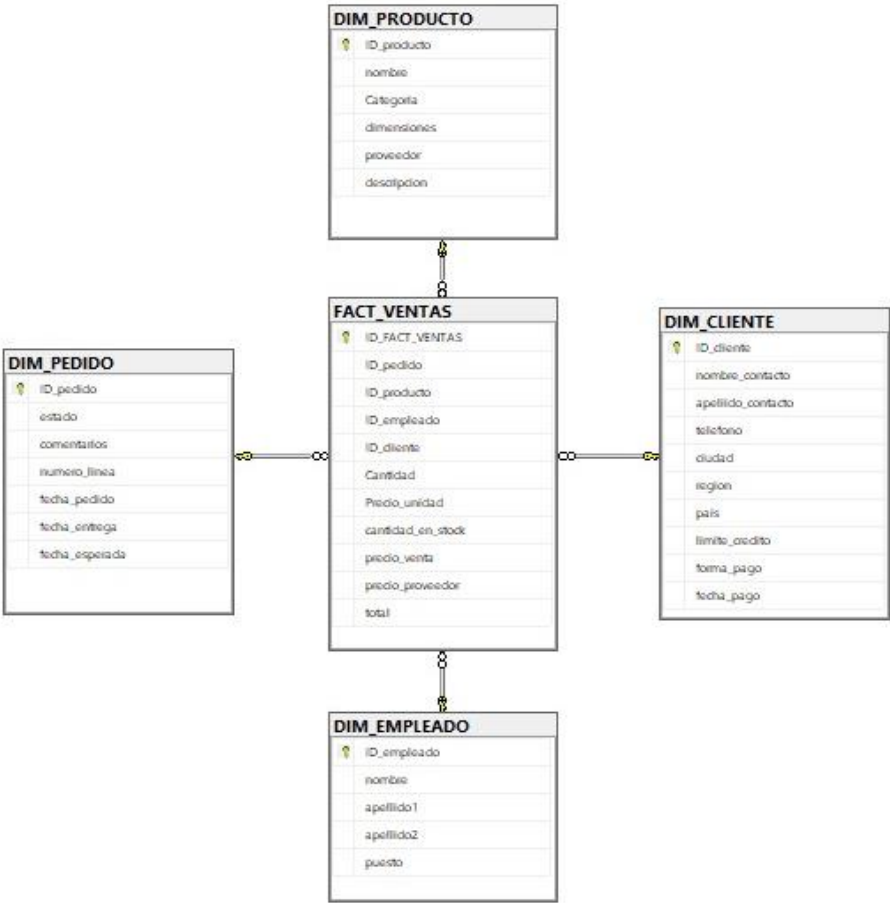
Así mismo la absorción o unificación de tablas conectadas por claves foráneas, que estén dentro de una estructura dimensional de jerarquía de datos en la tabla dimensión de mayor granularidad pueda a través de la conexión de dichas claves tener la posibilidad de generar unas respuestas a las consultas que se planteen mejorando su desempeño.

El modelo relacional de Jardinería constaba de 8 tablas relacionadas entre sí a partir de claves foráneas, unas tablas muy íntimamente relacionadas como la tabla clientes que desprendían de ella las tablas unidas entre sí: Pedido, detalle pedido, producto y categoría producto, de aquí pudimos detectar que era una de las tablas principales de dicho modelo relacional, pues de esta misma también salían otras tablas que no estaban relacionadas entre sí como: empleado, oficina y pago.

Al final el modelo dimensional planteado y basándonos en el concepto de jerarquía y granularidad la tabla DIM_PRODUCTO, quedó como aquella tabla que conecta a través de la clave foránea a la tabla detalle_pedido, así mismo la tabla DIM_PRODUCTO es quién tendrá relación con la tabla Categoría_producto que deja de aparecer en el modelo dimensional. En el caso de la tabla oficina, esta queda absorbida por la tabla DIM_EMPLEADO y la tabla pago en DIM_CLIENTE.

En nuestro análisis consideramos que el modelo estrella planteado es el que tiene mayor coherencia con la definición de las tablas y columnas que pueden describir de manera detallada queries del proceso de venta para Jardinería.

Diseño (Imagen) del modelo estrella donde se puedan observar las dimensiones, la tabla de hechos, sus campos, tipos de datos y relaciones.



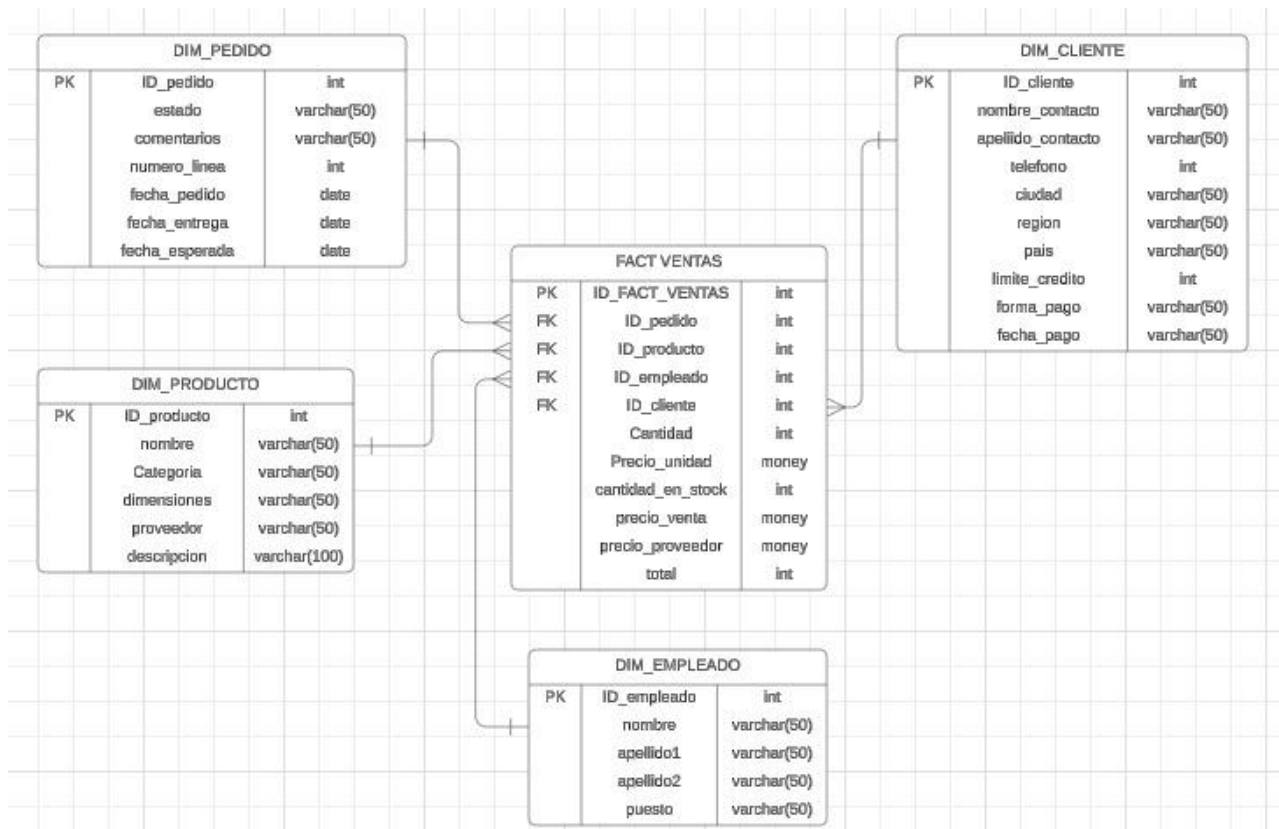
Lista de dimensiones propuestas.

Para la empresa se diseñó el modelo estrella con la base de datos de Jardinería y se tomó como base el modelo relacional para analizar e identificar qué dimensiones son las más pertinentes para análisis de las ventas, de esta manera las tablas dimensionales que se crearon fueron:

Dimensiones	Columnas	Tipo de datos
Producto	ID_producto, nombre, categoria, dimensiones, proveedor y descripción	Int y Varchar
Pedido	ID_pedido, estado, comentarios, numero_linea, fecha_pedido, fecha_entrega y fecha_esperada	Int, Varchar y date
Empleado	ID_empleado, nombre, apellido1, apellido2 y puesto	Int y Varchar
Cliente	ID_cliente, nombre_contacto, apellido_contacto, telefono, ciudad, region, pais, limite_credito, forma_pago y fecha_pago	Int y Varchar

- Cada dimensión debe tener especificadas las columnas y qué tipos de datos van a almacenar.
- Detalla la tabla de hechos, con sus campos y tipos de datos.

A continuación, se describe el modelo detallado de tablas DIMENSIONALES y de HECHOS con sus respectivos campos, claves y tipo de datos.



Conclusiones

- Los modelos dimensionales permiten una mejor visualización y comprensión del contexto con el que se definen las tablas que van a conformar la base de datos.
- Hablar de modelos dimensionales permite tener un mayor enfoque al momento de realizar las consultas disminuyendo joins y procedimientos más extensos que pueden llevar a errores.
- Para definir un modelo relacional es necesario plantear las relaciones entre tablas, y por ello es importante contar con un modelo Entidad Relación.

Referencias

Sánchez, M. G. (2023, julio 19). ¿Qué es el Modelo Estrella? Tecon. <https://www.tecon.es/que-es-el-modelo-estrella/>

¿Qué es una base de datos relacional? (s.f). Oracle.com. Obtenido de <https://www.oracle.com/co/database/what-is-a-relational-database/>

Ahmed, I. (2024, junio 07). What is dimensional data modeling? Examples, benefits & more. Astera. <https://www.astera.com/es/knowledge-center/dimensional-modeling-guide/>

Datha Learning. (2022, diciembre 12). Introducción al Modelado Dimensional. YouTube. <https://youtu.be/Ewh8WsCOFQ4?feature=shared>