

DECISIONES DE DISEÑO

Data Warehouse de Análisis de Ventas

1. ¿Por qué Modelo Estrella?

Se seleccionó un **modelo estrella** por su simplicidad y rendimiento. Este diseño facilita consultas analíticas con menor cantidad de joins, es fácil de entender para usuarios de negocio, y es compatible con herramientas BI como Power BI. La alternativa de modelo copo de nieve fue descartada por agregar complejidad innecesaria para el volumen actual de datos.

2. ¿Cómo se Definieron las Dimensiones?

DimCustomer: Contiene información de clientes con SCD Tipo 2 para rastrear cambios históricos en datos demográficos (ciudad, país, contacto).

DimProduct: Almacena catálogo de productos con SCD Tipo 2 para mantener historial de precios, permitiendo análisis con precios del momento de venta.

DimTime: Dimensión precalculada (2020-2030) con atributos temporales (año, trimestre, mes, día) para facilitar agregaciones sin cálculos repetitivos.

DimStatus: Dimensión de referencia con 4 estados de órdenes (Pending, Shipped, Delivered, Cancelled).

3. ¿Cómo se Diseñó la Tabla de Hechos?

Grain: Una fila por cada línea de detalle de orden (un producto por orden). **Métricas:** Quantity, UnitPrice y TotalPrice (calculado). **Claves foráneas:** CustomerKey, ProductKey, TimeKey y StatusKey conectan a las cuatro dimensiones. Se incluyó OrderID como clave de negocio para trazabilidad al sistema transaccional origen.

4. ¿Por qué Incluir Capa de Staging?

Las **4 tablas staging** (StagingCustomers, StagingProducts, StagingOrders, StagingOrderDetails) desacoplan extracción de transformación en el proceso ETL. Esto permite validar y limpiar datos antes de cargarlos al Data Warehouse, mantener auditoría de datos crudos, y facilitar reprocesamiento en caso de errores.

5. ¿Por qué SCD Tipo 2?

Slowly Changing Dimensions Tipo 2 se implementó en DimCustomer y DimProduct usando campos StartDate, EndDate e IsCurrent. Esto permite mantener historial completo de cambios: cuando un producto cambia de precio, se crea un nuevo registro manteniendo el anterior. Las ventas conservan la FK al precio histórico correcto, garantizando precisión en análisis temporales.

6. ¿Cómo se Garantiza la Integridad?

Todas las relaciones tienen **Foreign Keys con ON DELETE RESTRICT**, previniendo eliminación accidental de dimensiones con ventas asociadas. ON UPDATE CASCADE permite actualizar claves si necesario. El orden de carga es: Dimensiones → DimTime (pre-poblada) → FactSales.

7. ¿Qué Optimizaciones se Aplicaron?

Índices: En todas las FKs de FactSales, claves de negocio en dimensiones, e índices compuestos para consultas frecuentes (TimeKey+CustomerKey, TimeKey+ProductKey). **Vistas:** 4 vistas analíticas pre-construidas (vw_CurrentSales, vw_SalesByProduct, vw_SalesByCustomer, vw_SalesByTime) simplifican consultas comunes para dashboards.

Resumen Técnico

Característica	Decisión
Modelo	Estrella (1 hecho + 4 dimensiones)

Motor BD	MySQL 8.0+ con InnoDB
SCD	Tipo 2 en DimCustomer y DimProduct
Staging	4 tablas para proceso ETL
Grain	Línea de detalle de orden
Métricas	Quantity, UnitPrice, TotalPrice