Desafío estudiante en práctica

# Desafío 1 (20% de la evaluación total)

**Objetivo:** Diseñar un Data Warehouse (DW) que permita analizar las transacciones de una empresa de retail desde múltiples perspectivas: productos, tiendas, proveedores, promociones, medios de pago, clientes, canales de venta y tiempo. El modelo sigue la estructura de un esquema en estrella (Star Schema), donde la tabla de hechos central almacena los datos transaccionales y las dimensiones entregan contexto.

**Estructura General:**

* Base de datos:



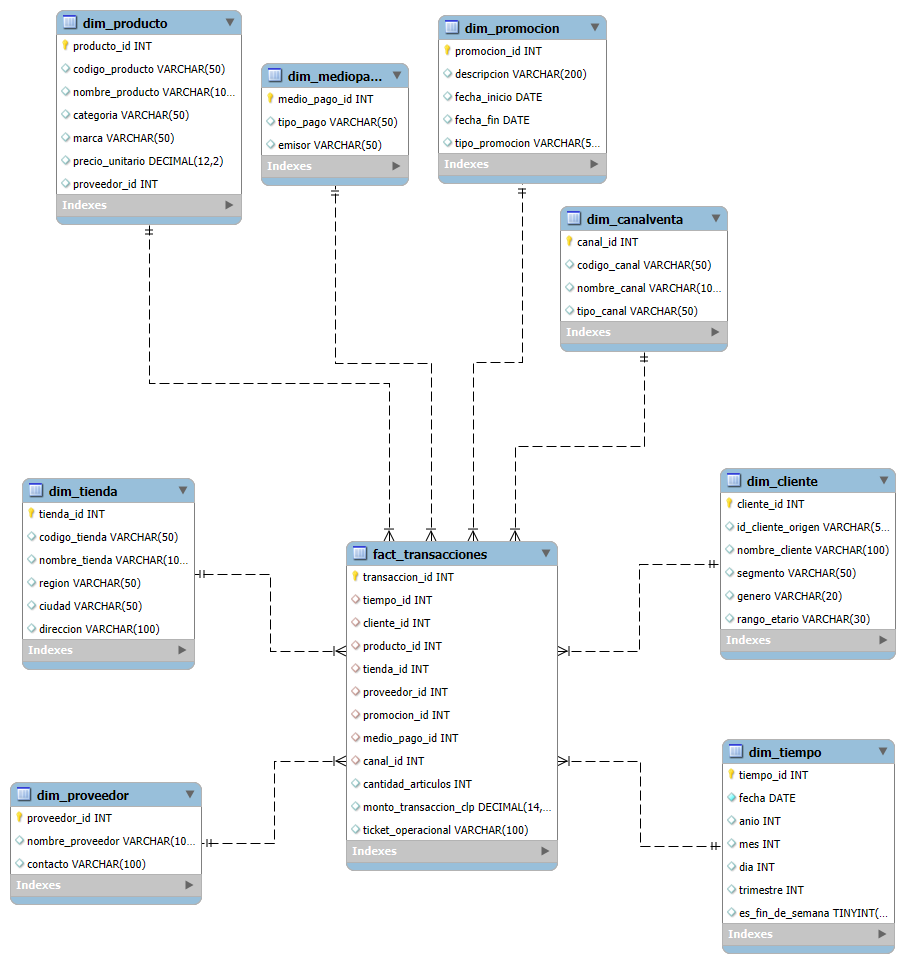
* Tipo de modelo: Dimensional (esquema estrella)
* Tabla de hechos: Fact\_Transacciones.

Almacena las métricas de ventas y claves foráneas hacia las dimensiones.

* Tablas dimensión:
* Dim\_Tiempo
* Dim\_Cliente
* Dim\_Producto
* Dim\_Tienda
* Dim\_Proveedor
* Dim\_Promocion
* Dim\_MedioPago
* Dim\_CanalVenta

**Diagrama Entidad–Relación (EER)**

Cada tabla de dimensión se conecta con la tabla de hechos Fact\_Transacciones a través de su llave primaria (PK), que se convierte en llave foránea (FK) dentro de la tabla de hechos.

*Generado mediante MySQL Workbench – Revese Engineer Database*

# Desafío 2 (20% de la evaluación total)

Documentación de dos entidades: Dim\_Producto y Fact\_Transacciones.

El propósito de esta documentación técnica es describir de forma clara y resumida dos entidades del modelo de Data Warehouse retail\_dw, diseñado para una empresa de retail. Este modelo sigue una arquitectura en estrella (Star Schema) y tiene como objetivo facilitar el análisis de ventas desde múltiples perspectivas: producto, cliente, tiempo, tienda, proveedor, promoción, canal y medio de pago.

**Dim\_Producto**

|  |  |
| --- | --- |
| Elemento | Descripción |
| Tipo de entidad | Dimensión |
| Propósito | Almacenar información descriptiva de los productos comercializados por la empresa. Permite analizar las ventas por categorías, marcas o rangos de precios. |
| Campos Principales | producto\_id (PK), codigo\_producto, nombre\_producto, categoria, marca, precio\_unitario, proveedor\_id (FK) |
| Relaciones | Se conecta con la tabla Dim\_Proveedor mediante proveedor\_id, y con la tabla de hechos Fact\_Transacciones como llave foránea. |
| Reglas de negocio | Cada producto pertenece a un proveedor principal. Si un producto cambia de proveedor o categoría, se registra una nueva versión (histórico tipo SCD2). |
| Fuente de datos | Sistema ERP o catálogo maestro de productos. |
| Frecuencia de Carga | Semanal o cuando existan actualizaciones relevantes. |
| Owner | Catalina Ormeño Rivera – Creadora del modelo y desarrolladora de la base de datos (transacciones\_dw.sql). |
| Versión | v1.0 – 22/10/2025 |

**Fact\_Transacciones**

|  |  |
| --- | --- |
| Elemento | Descripción |
| Tipo de entidad | Tabla de hechos |
| Propósito | Registrar todas las ventas realizadas por la empresa, vinculando las dimensiones del modelo para realizar análisis multidimensionales. |
| Campos Principales | transaccion\_id (PK), tiempo\_id (FK), cliente\_id (FK), producto\_id (FK), tienda\_id (FK), proveedor\_id (FK), promocion\_id (FK), medio\_pago\_id (FK), canal\_id (FK), cantidad\_articulos, monto\_transaccion\_clp, ticket\_operacional |
| Relaciones | tiempo\_id → Dim\_Tiempo (1:N)  cliente\_id → Dim\_Cliente (1:N)  producto\_id → Dim\_Producto (1:N)  tienda\_id → Dim\_Tienda (1:N)  proveedor\_id → Dim\_Proveedor (1:N)  promocion\_id → Dim\_Promocion (1:N)  medio\_pago\_id → Dim\_MedioPago (1:N)  canal\_id → Dim\_CanalVenta (1:N) |
| Reglas de negocio | Cada registro corresponde a una transacción única. Las métricas se pueden agregar por cualquier dimensión. No se permiten transacciones sin fecha o cliente. |
| Fuente de datos | Sistemas POS, plataformas de e-commerce y canales de venta telefónicos |
| Frecuencia de Carga | Diaria o en procesos batch. |
| Owner | Catalina Ormeño Rivera – Creadora del modelo y desarrolladora de la base de datos (transacciones\_dw.sql). |
| Versión | v1.0 – 22/10/2025 |

# Desafío 3 (60% de la evaluación total)

Objetivo: Transformar y analizar transacciones reales usando herramientas de procesamiento de datos (Pandas) y servicios en la nube (Google Cloud BigQuery).

Se adjunta desarrollo completo en archivo **desafio3.ipynb**

Una captura de pantalla de una computadora

El contenido generado por IA puede ser incorrecto.

**Creación del dataset retail\_dw\_dataset en BigQuery Sandbox**

Captura de pantalla de computadora

El contenido generado por IA puede ser incorrecto.

Captura de pantalla de computadora

El contenido generado por IA puede ser incorrecto.

Se realizó la carga del archivo transacciones\_clean.csv directamente en BigQuery Sandbox para generar posterior consulta SQL

Captura de pantalla de computadora

El contenido generado por IA puede ser incorrecto.

Captura de pantalla de computadora

El contenido generado por IA puede ser incorrecto.