**Análisis de caso – Modelamiento Multidimensional**

**1. Análisis del contexto**

La base de datos **OLTP (Procesamiento Analítico en Línea)** de Mercato está diseñada para operaciones diarias y transaccionales, como registrar una venta en tiempo real o actualizar el inventario. Su principal objetivo es la eficiencia en la inserción, actualización y eliminación de datos. Sin embargo, esto hace que las consultas complejas y que involucran grandes volúmenes de datos históricos (como "ventas totales del último año por región") sean lentas y consuman muchos recursos, esto principalmente porque el modelo OLTP no está diseñado para este tipo de consultas.

En cambio, un modelo **OLAP (Procesamiento Analítico en Línea)** está optimizado para el análisis de datos. Este tipo organiza la información de manera multidimensional, lo que permite a los analistas explorar y sumarizar datos desde diferentes perspectivas de forma rápida y sencilla. Esto es ideal para Mercato, ya que facilita la creación de reportes, dashboards y proyecciones sin afectar el rendimiento de sus operaciones diarias.

Clasificación de Hechos y Dimensiones

En este caso, las entidades del negocio se clasificarían de la siguiente manera:

* Hecho: La Venta. Es el evento principal que queremos analizar. Contiene métricas cuantificables como el monto, la cantidad y los descuentos.
* Dimensiones: Estas son las perspectivas desde las cuales se analiza el hecho de la venta. Se pueden explicar como los "quién, qué, dónde y cuándo" de la venta:
  + Cliente: Quién realizó la compra.
  + Producto: Qué se compró.
  + Tienda: Dónde se realizó la compra.
  + Fecha/Tiempo**:** Cuándo se realizó la compra.

**2. Diseño del modelo multidimensional**

**Esquema estrella para el proceso de ventas**

El esquema estrella es una excelente opción para este caso por su simplicidad y alto rendimiento en consultas. Se compone de una única tabla de hechos central y varias tablas de dimensiones que se conectan directamente a ella.

**Diagrama de Esquema Estrella:**

Diagrama, Esquemático

El contenido generado por IA puede ser incorrecto.

**3. Jerarquías y agregaciones**

**Jerarquías**

Las jerarquías permiten a los analistas navegar por los datos a diferentes niveles de granularidad, desde el detalle más fino hasta el más general.

* **Dim\_Tiempo:**
  + **Jerarquía:** hora g día g mes g año
  + **Niveles:** Un analista puede ver las ventas diarias y luego alejarse para ver el total mensual, trimestral o anual.
* **Dim\_Producto:**
  + **Jerarquía:** SKU g Subcategoría g Categoría g Marca
  + **Niveles:** Permite analizar las ventas de un producto específico (SKU), luego consolidarlas por categoría de producto y, finalmente, por marca.
* **Dim\_Tienda:**
  + **Jerarquía:** nombre\_tienda g Comuna g Región
  + **Niveles:** Permite analizar el rendimiento de una tienda individual, luego consolidar los datos para ver el rendimiento de una ciudad o región.

**Agregaciones:**

Las agregaciones son resúmenes precalculados de los datos que aceleran las consultas, se asume que serán consultas futuras a realizar de manera semanal o mensual por ejemplo como parte de la tarea de análisis de cada área.

1. **Ventas mensuales por región:** Se pueden precalcular las ventas totales (SUM(monto\_venta)) agrupadas por mes (de la dimensión Tiempo) y región (de la dimensión Tienda). Esto permite generar reportes de rendimiento regional sin tener que sumar millones de transacciones cada vez.
2. **Ventas por categoría de producto y mes:** De manera similar, se puede precalcular la suma del monto de venta (SUM(monto\_venta)) por categoría (de la dimensión Producto) y mes (de la dimensión Tiempo). Esto es útil para analizar qué categorías tienen mejor desempeño en diferentes épocas del año.

**4. Buenas prácticas de diseño**

Esquema estrella vs. copo de nieve

La razón para usar un esquema estrella en este caso es el mejor rendimiento y la simplicidad de su creación, por sobre los beneficios que podría entregar el esquema de copo de nieve.

* **Rendimiento:** Las consultas en un esquema estrella solo necesitan uniones entre la tabla de hechos y sus dimensiones. Esto la hace mucho más rápida que un esquema copo de nieve, que requiere múltiples uniones para llegar a los atributos de las dimensiones más profundas (por ejemplo, unir la tabla de hechos con Dim\_Producto, y luego esta con Dim\_Categoria).
* **Simplicidad:** La estructura es más intuitiva y fácil de entender para los analistas de negocio, lo que reduce la curva de aprendizaje y la complejidad en la creación de consultas.

**Normalización vs. desnormalización**

El esquema estrella se basa en la desnormalización de las dimensiones.

* Desnormalizar significa que los atributos se almacenan directamente en la tabla de la dimensión, incluso si se repiten. Por ejemplo, en Dim\_Producto, los campos categoria y subcategoria están en la misma tabla.
* Esto es una decisión intencional de diseño para OLAP. Al tener todos los atributos en una sola tabla (la de la dimensión), evitamos las uniones adicionales que ralentizarían las consultas. Por ejemplo, si la dimensión Producto estuviera normalizada (esquema copo de nieve), para obtener el nombre de la categoría necesitaríamos una unión adicional, lo que impactaría negativamente el rendimiento.

**5. Casos de uso**

* 1. **Total de ventas por categoría de producto y mes:**

Texto

El contenido generado por IA puede ser incorrecto.

* 1. **Ranking de clientes por monto de compra:**

Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

* 1. **Comparativo de ventas por tienda en distintos años:**

Texto

El contenido generado por IA puede ser incorrecto.