

Diseño de Datamar

Análisis de ventas de un restaurante de pizza a domicilio que vende dos líneas de comida, la primera son las pizzas de diferentes ingredientes (jamón, champiñones, pepperoni, suprema, etc.), de 2 tipos de masas (delgada y gruesa) y de 4 tamaños diferentes. La segunda línea de productos incluye postres de diferentes tipos. Adicionalmente vende bebidas de tres tipos, jugos naturales, bebidas carbonatadas y smoothies. Posee un total de 35 restaurantes distribuidos en 3 países, y agrupados en cada país por regiones (centro, norte, sur, este, oeste).

El dueño quiere poder comparar cuánto venden por cada línea de productos así como cuál es el producto más vendido en cada restaurante y en cada región. Necesitan también analizar la cantidad de pedidos por hora en cada país, el monto de venta por producto en moneda local y en US\$ y el ticket promedio por restaurante. Necesita poder comparar el monto de venta mensual por restaurante contra el presupuesto mensual que se maneja también por restaurante en US\$. (El presupuesto de ventas está definido mensualmente por restaurante en US\$)

1. Identificar las dimensiones y métricas necesarias para definir un modelo, indique claramente las dimensiones necesarias, la o las tablas de hechos y las métricas correspondientes. determina si se requiere un modelo estrella, snowflake o starflake.
2. Defina de manera detallada las tablas y campos necesarios para crear el modelo anterior, especificando claramente las llaves primarias de cada tabla.

Dimensiones y Métricas

Decidí hacer cuatro tablas de hechos: *fact_ventas_bebidas*, *fact_ventas_pizzas*, *fact_ventas_postres*, *fact_planificacion*

Las ventas las decidí tener en una tabla de hechos por separado ya que las métricas ó propiedades no son las mismas, y esto podría causar valores nulos si todas llegaran a convivir en una sola tabla de hechos de cualquier tipo de modelo, sea estrella, copo de nieve o su combinación.

Con esta separación podre utilizar un modelo estrella y veremos en el diseño lógico como las tablas de hechos tendrán dimensiones en común, note que separé la localización geográfica, esto podría aumentar el tamaño del datawarehouse pero simplificaría los queries. Las dimensiones serán: `dim_mes_anio`, `dim_dia_hora`, `dim_precio`, `dim_restaurante`, `dim_pais`, `dim_region`, `dim_detalle_pizza`, `dim_detalle_jugo`, `dim_detalle_postre`

Definición Lógica del Modelo

El modelo gráficamente podría verse caótico pero la lógica detrás es muy simple, como los productos tienen propiedades diferentes se decidió crear tres tablas para evitar valores nulos, realmente las tres tablas que contienen la información de las ventas tienen llaves surrogadas que hacen referencia a las mismas tablas lo que ocasiona un poco de dificultad de entender visualmente utilizando la herramienta “drawdb”.

Cada tabla de ventas tiene solamente una llave surrogada específica para la descripción/detalle del producto, todas las demás llaves surrogadas refieren a una fila de la misma tabla, ósea que precio, restaurante, país, región, mes y año, día y hora, existen como llaves surrogadas en las tablas de hechos de ventas, esto nos permite hacer queries sin muchos joins y de esta manera optimizarlos.

También se puede observar que el hecho de plan, tiene al presupuesto como métrica, y hace referencia con sus llaves surrogadas respectivas a las mismas tablas de restaurante, mes y año, país, región.

Enlace al diagrama [aquí](#)

