

Práctica 2

1-. Copiar 6 consultas de la práctica anterior, y sombrear con **color naranja** los términos correspondientes a dimensiones y con **color verde** los términos correspondientes a medidas.

En nuestro caso, modificamos un poco las consultas para que sean más variadas y abarquen mas dimensiones y medidas.

- **Rol: Gerente de ventas**

- i. **Zonas** donde se registraron más **ventas** en un **periodo** dado.
- ii. Listado de **Fecha y hora** donde se registraron **más ventas**, ordenados en orden descendente.
- iii. Top 10 **zonas** donde se registraron los mayores **acumulados de descuentos por promoción**.

- **Rol: Gerente de proveedores**

- i. **Nombre y tipo de proveedores** que lograron más alcance en **distintos clientes**.
- ii. Top 3 **tipos de proveedores** que generaron mayor **cantidad de ventas** por **zona**.
- iii. **Proveedores** que realizaron ventas con **descuento por promoción mayor a \$5000** en un determinado **horario** por **zona**.

2-. Elija 4 consultas de la practica 1 y escriba las consultas SQL necesarias para responderlas.

Se van mencionando las consultas seleccionadas y su correspondiente resolución en SQL.

- Zonas donde se registraron más ventas en un periodo dado.

```
SELECT dz.nombre AS zona,
       COUNT(v.id) AS total_ventas
FROM venta v
JOIN cliente c ON v.cliente_id = c.id
JOIN domicilio d ON c.domicilio_id = d.id
JOIN domicilio_zona dz ON d.zona_id = dz.id
WHERE v.fecha_hora BETWEEN '2025-01-01' AND '2025-01-31'
GROUP BY dz.nombre
ORDER BY total_ventas DESC;
```

- Top 10 zonas donde se registraron los mayores acumulados de descuentos por promoción.

```

SELECT dz.nombre AS zona,
       SUM(v.descuento_promocion) AS total_descuento
FROM venta v
JOIN cliente c ON v.cliente_id = c.id
JOIN domicilio d ON c.domicilio_id = d.id
JOIN domicilio_zona dz ON d.zona_id = dz.id
GROUP BY dz.nombre
ORDER BY total_descuento DESC
LIMIT 10;

```

- Top 3 tipos de proveedores que generaron mayor cantidad de ventas por zona.

```

SELECT zona,
       tipo_proveedor,
       total_ventas
FROM (
    SELECT dz.nombre AS zona,
           tp.descripcion AS tipo_proveedor,
           COUNT(v.id) AS total_ventas,
           ROW_NUMBER() OVER (PARTITION BY dz.nombre ORDER BY COUNT(v.id) DESC) AS rn
    FROM venta v
    JOIN proveedor p ON v.proveedor_id = p.id
    JOIN tipo_proveedor tp ON p.tipo_id = tp.id
    JOIN cliente c ON v.cliente_id = c.id
    JOIN domicilio d ON c.domicilio_id = d.id
    JOIN domicilio_zona dz ON d.zona_id = dz.id
    GROUP BY dz.nombre, tp.descripcion
) sub
WHERE rn <= 3
ORDER BY zona, total_ventas DESC;

```

- Proveedores que realizaron ventas con descuento por promoción mayor a \$5000 en un determinado horario por zona.

```

SELECT p.nombre AS proveedor,
       dz.nombre AS zona,
       v.fecha_hora,
       v.descuento_promocion
FROM venta v
JOIN proveedor p ON v.proveedor_id = p.id
JOIN cliente c ON v.cliente_id = c.id
JOIN domicilio d ON c.domicilio_id = d.id
JOIN domicilio_zona dz ON d.zona_id = dz.id
WHERE v.descuento_promocion > 5000
       AND CAST(v.fecha_hora AS TIME) BETWEEN '18:00:00' AND '22:00:00'
ORDER BY v.fecha_hora DESC;
EN '18:00:00' AND '22:00:00'
ORDER BY v.fecha_hora DESC;

```

3-. Definir un modelo dimensional conceptual que contenga Hechos y Dimensiones, que sea suficiente para responder a las consultas del punto anterior. Tener presente que el modelo pueda eventualmente responder nuevas consultas no previstas inicialmente.

En nuestro caso decidimos emplear el **modelo copo de nieve**. Esto con la idea de normalizar parcialmente las dimensiones, así separando jerarquías o atributos repetidos en subtablas relacionadas.

Con esto creemos que se puede lograr mejorar la consistencia de datos, reducir redundancia y priorizar una estructuración y mantenibilidad.

Hecho principal: Venta

Medidas:

- monto_origen
- propina
- costo_servicio
- descuento_promocion
- monto_total
- cantidad_items (nuevo, útil para ticket promedio)

Grano: Una venta individual (por pedido realizado)

Dimensiones y subdimensiones:

Tiempo

- Tiempo: dim_tiempo (ID, Fecha, Día, Mes, Año, Trimestre, Semana del año, Día de semana, Festivo (S/N), Día laborable (S/N), Turno (mañana/tarde/noche))

Cliente (Con jerarquías geográficas y demográficas)

- dim_cliente: ID, Nombre completo, Sexo, Fecha nacimiento, ID_zona, ID_segmento
- dim_zona: ID_zona, Nombre zona, ID_ciudad
- dim_ciudad: ID_ciudad, Nombre ciudad, ID_region
- dim_region: ID_region, Nombre región, País
- dim_segmento: ID_segmento, Descripción (e.g., Frecuente, Ocasional, VIP), Criterio segmentación

Proveedor (Con tipo y categoría)

- dim_proveedor: ID, Nombre proveedor, ID_tipo, ID_categoria, ID_zona
- dim_tipo_proveedor: ID_tipo, Descripción (e.g., Restaurante, Tienda)
- dim_categoria_proveedor: ID_categoria, Descripción (e.g., Comida rápida, Gourmet)

Repartidor

- dim_repartidor: ID, Nombre, Sexo, Edad, ID_tipo_contrato, ID_zona
- dim_tipo_contrato: ID_tipo_contrato, Descripción (Fijo, Freelancer, Tercerizado)

Tipo de Pago

- dim_tipo_pago: ID, Descripción (Tarjeta, Efectivo, etc.), Plataforma (Visa, MasterCard, etc.), Tipo de dispositivo usado

Promoción

- dim_promocion: ID_promocion, Nombre, Tipo (Descuento %, cupón), Fecha inicio/fin, Canal, Campaña asociada

Algo a tener en cuenta es que se consideró eliminar Zona (está normalizada dentro de las dimensiones Cliente, Proveedor y Repartidor a través de jerarquías).

Ahora bien, ¿las consultas planteadas en los ejercicios anteriores se pueden responder con este modelo propuesto?. A continuación listamos las consultas y con que se pueden responder en base a lo planteado:

- Zonas donde se registraron más ventas en un periodo dado --> Podemos responderla con dim_zona y dim_tiempo

- Top 10 zonas donde se registraron los mayores acumulados de descuentos por promoción --> podemos responderla con dim_zona y medida descuento_promocion
- Top 3 tipos de proveedores que generaron mayor cantidad de ventas por zona --> se puede responder con dim_tipo_proveedor y dim_zona
- Proveedores que realizaron ventas con descuento por promoción mayor a \$5000 en un determinado horario por zona --> podemos responderla con descuento_promocion, dim_tiempo y dim_zona

4-. Especificar la granularidad del modelo.

La granularidad está dada por **HECHO_VENTA** contando así con ventas individuales realizada por un cliente a un proveedor, en un momento determinado, con un repartidor asignado, a través de un tipo de pago específico.