



INSTITUTO POLITÉCNICO NACIONAL

ESCUELA SUPERIOR DE CÓMPUTO



Carrera: Licenciatura en Ciencia de Datos

Unidad de aprendizaje: Bases de datos avanzada

Laboratorio 8

Integrantes :

Angeles Librado Moncerrat

Grupo: 4AV1

Fecha : 19 de nov. de 24

Profesor:

Andres Garcia Floriano

Introducción

En este proyecto, se ha diseñado un esquema estrella en SQL para analizar las ventas de una librería en línea. El objetivo principal es facilitar consultas analíticas para obtener información detallada y multifacética sobre las ventas, con énfasis en las características de los libros, clientes, tiendas y el tiempo de la transacción. Este tipo de modelo, en el que se centraliza una tabla de hechos (ventas de libros) conectada con varias tablas de dimensiones, es especialmente útil en entornos de inteligencia de negocios, ya que permite el análisis rápido de grandes volúmenes de datos desde múltiples perspectivas.

Este reporte presenta el proceso de diseño e implementación del esquema estrella, la inserción de datos ficticios en las tablas, y una serie de consultas de análisis de ventas que pueden ayudar a mejorar la estrategia comercial de la librería.

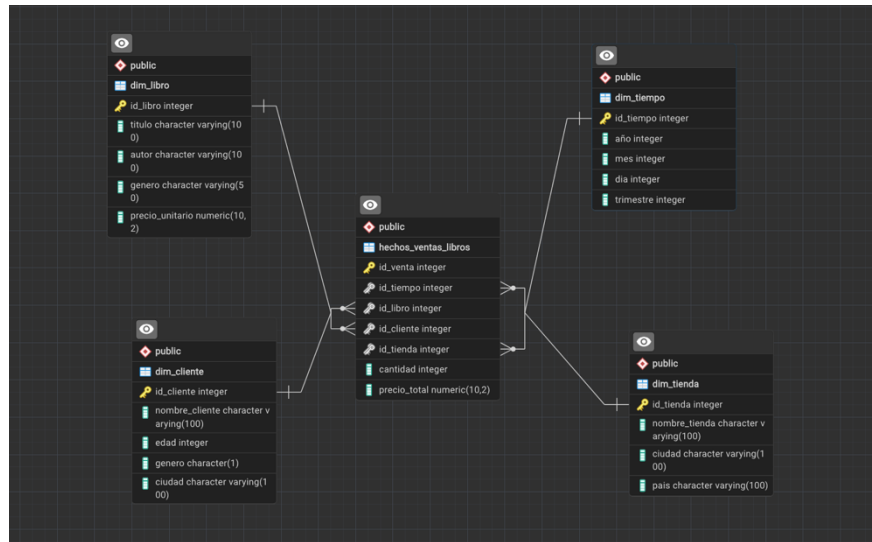
Desarrollo

Diseño del Esquema Estrella

Para el diseño, se ha creado una tabla de hechos denominada `hechos_ventas_libros` que almacena la información esencial de cada venta, incluyendo la cantidad de libros vendidos y el precio total de la venta. Además, se han implementado cuatro tablas de dimensiones que enriquecen los análisis:

1. **dim_tiempo:** Incluye detalles sobre el año, mes, día y trimestre, permitiendo el análisis temporal de las ventas.
2. **dim_libro:** Contiene información detallada sobre los libros, como título, autor, género y precio unitario, lo cual permite estudiar las ventas desde la perspectiva de los productos.
3. **dim_cliente:** Proporciona información sobre los clientes, incluyendo edad, género y ciudad, lo cual es fundamental para los análisis de segmentación de clientes.
4. **dim_tienda:** Incluye información de las tiendas, como el nombre, ciudad y país, para estudiar la distribución geográfica de las ventas.

Se implementaron claves foráneas en la tabla de hechos para relacionarla con cada tabla de dimensión, asegurando así la integridad referencial y el aprovechamiento de las relaciones en el análisis.



Inserción de Datos Ficticios

Para simular el funcionamiento del esquema, se insertaron registros ficticios en cada una de las tablas, logrando una base suficiente para realizar pruebas analíticas. Estos datos incluyen cinco libros, cinco clientes, tres tiendas y tres momentos de tiempo diferentes. Se ingresaron también diez registros de ventas en la tabla de hechos hechos_ventas_libros.

```

52 -- Insertar datos en dim_tiempo
53 INSERT INTO dim_tiempo (id_tiempo, año, mes, día, trimestre) VALUES
54 (1, 2023, 1, 15, 1),
55 (2, 2023, 2, 10, 1),
56 (3, 2023, 3, 5, 1);
57
58 -- Insertar datos en dim_libro
59 INSERT INTO dim_libro (id_libro, titulo, autor, genero, precio_unitario) VALUES
60 (1, 'Cien Años de Soledad', 'Gabriel García Márquez', 'Novela', 15.50),
61 (2, '1984', 'George Orwell', 'Distopía', 12.00),
62 (3, 'El Quijote', 'Miguel de Cervantes', 'Clásico', 18.75),
63 (4, 'Donde los Árboles Cantan', 'Laura Gallego', 'Fantasía', 9.99),
64 (5, 'Harry Potter y la Piedra Filosofal', 'J.K. Rowling', 'Fantasía', 14.50);
65
66 -- Insertar datos en dim_cliente
67 INSERT INTO dim_cliente (id_cliente, nombre_cliente, edad, genero, ciudad) VALUES
68 (1, 'Ana Martínez', 34, 'F', 'Madrid'),
69 (2, 'Carlos López', 28, 'M', 'Barcelona'),
70 (3, 'María Gómez', 45, 'F', 'Valencia'),
71 (4, 'José Ramírez', 51, 'M', 'Sevilla'),
72 (5, 'Lucía Fernández', 30, 'F', 'Zaragoza');
73
74 -- Insertar datos en dim_tienda
75 INSERT INTO dim_tienda (id_tienda, nombre_tienda, ciudad, pais) VALUES
76 (1, 'Librería Central', 'Madrid', 'España'),
77 (2, 'Libros del Sur', 'Barcelona', 'España'),
78 (3, 'Lecturas Vivas', 'Valencia', 'España');
79
80 -- Insertar datos en hechos_ventas_libros
81 INSERT INTO hechos_ventas_libros (id_venta, id_tiempo, id_libro, id_cliente, id_tienda, cantidad, precio_total) VALUES
82 (1, 1, 1, 1, 1, 2, 31.00),
83 (2, 1, 2, 2, 2, 1, 12.00),
84 (3, 1, 3, 3, 3, 1, 18.75),
85 (4, 2, 4, 4, 1, 3, 29.97),
86 (5, 2, 5, 5, 2, 1, 14.50),
87 (6, 2, 1, 1, 1, 2, 31.00),
88 (7, 3, 2, 2, 3, 2, 24.00);

```

Consultas de Análisis

Una vez creadas e insertadas las tablas, se realizaron consultas para extraer información valiosa:

1. **Total de ventas por género de libro y mes:** Esta consulta agrupa las ventas según el género de cada libro y el mes de la venta, permitiendo observar tendencias de ventas de cada género a lo largo del tiempo.

```
93 --Consulta 1: Total de ventas (precio_total) por género de libro y mes
94
95 SELECT
96     l.genero,
97     t.mes,
98     SUM(h.precio_total) AS total_ventas
99 FROM
100     hechos_ventas_libros h
101 JOIN
102     dim_libro l ON h.id_libro = l.id_libro
103 JOIN
104     dim_tiempo t ON h.id_tiempo = t.id_tiempo
105 GROUP BY
106     l.genero, t.mes
107 ORDER BY
108     t.mes, l.genero;
109
110
```

	genero character varying (50)	mes integer	total_ventas numeric
1	Clásico	1	18.75
2	Distopía	1	12.00
3	Novela	1	31.00
4	Fantasia	2	44.47
5	Novela	2	31.00
6	Clásico	3	18.75
7	Distopía	3	24.00
8	Fantasia	3	63.48

2. **Cantidad total de libros vendidos por tienda y autor:** Esta consulta identifica los autores y sus ventas en cada tienda, ayudando a determinar cuáles autores son más populares y en qué tiendas.

```
110 -- Consulta 2: Cantidad total de libros vendidos por tienda y autor
111
112 SELECT
113     ti.nombre_tienda,
114     l.autor,
115     SUM(h.cantidad) AS total_libros_vendidos
116 FROM
117     hechos_ventas_libros h
118 JOIN
119     dim_libro l ON h.id_libro = l.id_libro
120 JOIN
121     dim_tienda ti ON h.id_tienda = ti.id_tienda
122 GROUP BY
123     ti.nombre_tienda, l.autor
124 ORDER BY
125     ti.nombre_tienda, total_libros_vendidos DESC;
126
```

	nombre_tienda character varying (100)	autor character varying (100)	total_libros_vendidos bigint
1	Lecturas Vivas	J.K. Rowling	3
2	Lecturas Vivas	George Orwell	2
3	Lecturas Vivas	Miguel de Cervantes	1
4	Librería Central	Gabriel García Márquez	4
5	Librería Central	Laura Gallego	3
6	Librería Central	Miguel de Cervantes	1
7	Libros del Sur	Laura Gallego	2
8	Libros del Sur	George Orwell	1
9	Libros del Sur	J.K. Rowling	1

3. **Ingresos totales por ciudad de cliente y trimestre:** Con esta consulta se pueden observar las ciudades con mayores ingresos en diferentes trimestres, lo que permite identificar mercados rentables.

```
127 --Consulta 3: Ingresos totales por ciudad de cliente y trimestre
128
129 SELECT
130     c.ciudad,
131     t.trimestre,
132     SUM(h.precio_total) AS ingresos_totales
133 FROM
134     hechos_ventas_libros h
135 JOIN
136     dim_cliente c ON h.id_cliente = c.id_cliente
137 JOIN
138     dim_tiempo t ON h.id_tiempo = t.id_tiempo
139 GROUP BY
140     c.ciudad, t.trimestre
141 ORDER BY
142     c.ciudad, t.trimestre;
143
144
```

	ciudad character varying (100)	trimestre integer	ingresos_totales numeric
1	Barcelona	1	36.00
2	Madrid	1	62.00
3	Sevilla	1	49.95
4	Valencia	1	37.50
5	Zaragoza	1	58.00

4. **Total de ventas y número de libros comprados por cliente:** Esta consulta revela el gasto total y los libros adquiridos por cada cliente, útil para conocer a los clientes con mayor contribución a las ventas.

```
144 -- Consulta 4: Total de ventas de cada cliente (precio_total acumulado) y número de libros comprados
145
146 SELECT
147     c.nombre_cliente,
148     SUM(h.precio_total) AS total_gasto,
149     SUM(h.cantidad) AS total_libros_comprados
150 FROM
151     hechos_ventas_libros h
152 JOIN
153     dim_cliente c ON h.id_cliente = c.id_cliente
154 GROUP BY
155     c.nombre_cliente
156 ORDER BY
157     total_gasto DESC;

```

	nombre_cliente character varying (100)	total_gasto numeric	total_libros_comprados bigint
1	Ana Martínez	62.00	4
2	Lucía Fernández	58.00	4
3	José Ramírez	49.95	5
4	María Gómez	37.50	2
5	Carlos López	36.00	3

Conclusiones

El esquema estrella implementado cumple con el objetivo de organizar y estructurar los datos de ventas de la librería para permitir un análisis exhaustivo desde múltiples perspectivas. El modelo relacional empleado, junto con la implementación de las tablas de dimensiones, permite consultas que facilitan la identificación de patrones de compra y tendencias comerciales.

Con las consultas desarrolladas, la librería puede obtener información valiosa sobre sus ventas: los géneros de libros y autores más populares, las tiendas con mayores ventas, el desempeño por ciudad y trimestre, y los clientes más rentables. Esta información es esencial para tomar decisiones informadas, como personalizar promociones según los géneros más vendidos o concentrar esfuerzos de marketing en ciudades y clientes clave.

El esquema estrella es una herramienta eficaz para transformar datos de ventas en conocimiento accionable, y representa una base sólida para futuras expansiones, como la adición de nuevas dimensiones (por ejemplo, campañas de marketing) que permitan un análisis aún más detallado de la actividad de la librería.