

Guía completa OLAP: Drill-down, Roll-up, Slice & Dice, Pivot

Esta guía te acompaña paso a paso desde el armado de archivos CSV, la carga en Pentaho Data Integration 7.1, la estructura de la base, y la ejecución de todas las operaciones OLAP clásicas con datos en español. El caso de estudio simula ventas de una cadena de tiendas de ropa y tecnología a lo largo de varios meses y regiones.

1. Archivos CSV

Guarda estos archivos como .csv separados por comas, cada uno con su nombre indicado.

producto.csv

```
ID_Producto,Nombre,Categoria
1,Camisa Polo,Ropa
2,Pantalón Jeans,Ropa
3,Zapatos Deportivos,Calzado
4,Monitor X,Electrónica
5,Mochila Urbana,Accesorios
```

tienda.csv

```
ID_Tienda,Nombre,Ciudad,Región
1,Tienda Centro,Managua,Pacífico
2,Tienda Granada,Granada,Central
3,Tienda León,León,Occidente
```

fecha.csv

```
Fecha,Año,Mes,Dia,Nombre_mes
2024-10-01,2024,10,1,Octubre
2024-11-05,2024,11,5,Noviembre
2024-12-10,2024,12,10,Diciembre
2025-01-15,2025,1,15,Enero
2025-02-20,2025,2,20,Febrero
2025-03-03,2025,3,3,Marzo
```

venta.csv

ID_Venta	Fecha	ID_Producto	ID_Tienda	Cantidad	PrecioUnitario
1	2024-10-01	1	1	5	15
2	2024-10-01	2	2	3	25
3	2024-11-05	3	1	2	45
4	2024-11-05	4	2	1	180
5	2024-12-10	5	3	4	30
6	2025-01-15	1	2	6	15
7	2025-02-20	2	3	2	25
8	2025-03-03	3	1	1	45
9	2025-03-03	4	3	2	180
10	2025-03-03	5	1	3	30

2. Crear la base de datos en Pentaho

Estructura SQL para crear tablas

```
CREATE TABLE producto (  
  ID_Producto INT PRIMARY KEY,  
  Nombre VARCHAR(100),  
  Categoria VARCHAR(50)  
);
```

```
CREATE TABLE tienda (  
  ID_Tienda INT PRIMARY KEY,  
  Nombre VARCHAR(100),  
  Ciudad VARCHAR(50),  
  Región VARCHAR(50)  
);
```

```
CREATE TABLE fecha (  
  Fecha DATE PRIMARY KEY,  
  Año INT,  
  Mes INT,  
  Dia INT,  
  Nombre_mes VARCHAR(20)  
);
```

```
CREATE TABLE venta (  
  ID_Venta INT PRIMARY KEY,
```

Fecha DATE,
ID_Producto INT,
ID_Tienda INT,
Cantidad INT,
PrecioUnitario DECIMAL(10,2)
);

Ejecuta este script en un paso "Execute SQL script" sobre tu conexión HSQLDB.

3. Carga de datos en Pentaho

Para cada archivo:

- Usa Text File Input para leer el CSV.
- Mapéalo a la tabla correspondiente con Table Output.
- Ejecuta la transformación para cada tabla.

4. Operaciones OLAP

Drill-Down (profundizar en los detalles)

Pregunta: ¿Cuántas unidades se vendieron por producto en cada mes?

```
SELECT f.Nombre_mes, p.Nombre, SUM(v.Cantidad) AS UnidadesVendidas
FROM venta v
JOIN producto p ON v.ID_Producto = p.ID_Producto
JOIN fecha f ON v.Fecha = f.Fecha
GROUP BY f.Nombre_mes, p.Nombre
ORDER BY f.Nombre_mes, UnidadesVendidas DESC;
```

Roll-Up (resumir datos a niveles superiores)

Pregunta: ¿Cuántas unidades totales se vendieron por año?

```
SELECT f.Año, SUM(v.Cantidad) AS TotalUnidades
FROM venta v
JOIN fecha f ON v.Fecha = f.Fecha
GROUP BY f.Año
ORDER BY f.Año;
```

Slice & Dice (filtrar y pivotar)

Pregunta Slice: ¿Ventas (unidades) solo de Electrónica?

```
SELECT f.Nombre_mes, p.Nombre, SUM(v.Cantidad) AS UnidadesElectronica
```

```
FROM venta v
JOIN producto p ON v.ID_Producto = p.ID_Producto
JOIN fecha f ON v.Fecha = f.Fecha
WHERE p.Categoria = 'Electrónica'
GROUP BY f.Nombre_mes, p.Nombre;
```

Pregunta Dice: ¿Ventas por ciudad y categoría en Noviembre?

```
SELECT t.Ciudad, p.Categoria, SUM(v.Cantidad) AS UnidadesNov
FROM venta v
JOIN tienda t ON v.ID_Tienda = t.ID_Tienda
JOIN producto p ON v.ID_Producto = p.ID_Producto
JOIN fecha f ON v.Fecha = f.Fecha
WHERE f.Nombre_mes = 'Noviembre'
GROUP BY t.Ciudad, p.Categoria;
```

Pivoting (reorientar los ejes analíticos)

Pregunta: ¿Cuántas unidades vendidas por producto y por región?

```
SELECT p.Nombre,
       SUM(CASE WHEN t.Región = 'Pacífico' THEN v.Cantidad ELSE 0 END) AS
Pacífico,
       SUM(CASE WHEN t.Región = 'Central' THEN v.Cantidad ELSE 0 END) AS
Central,
       SUM(CASE WHEN t.Región = 'Occidente' THEN v.Cantidad ELSE 0 END) AS
Occidente
FROM venta v
JOIN producto p ON v.ID_Producto = p.ID_Producto
JOIN tienda t ON v.ID_Tienda = t.ID_Tienda
GROUP BY p.Nombre;
```

5. Exportación y visualización

- Conecta cada consulta en un paso Table Input → Text File Output para guardar en CSV.
- Graficar Drill-Down, Roll-up, Slice & Dice.

6. Explicación y análisis

- Drill-down: Permite ver por cada mes y producto cuántas unidades se vendieron, útil para profundizar en tendencias estacionales.
- Roll-up: Resume las ventas a nivel anual, útil para comparar periodos grandes.
- Slice & Dice: Obtienes cortes específicos sobre toda la base, por ejemplo sólo “Electrónica” o sólo “Noviembre” por ciudad y categoría.
- Pivoting: Permite presentar ventas cruzadas por producto y región, facilitando comparaciones visuales y análisis estratégico.