



Módulo 2

Introducción a Base de Datos

Sesión 3

Modelamiento de bases de datos analíticas OLAP

Introducción a OLAP

OLAP organiza los datos en cubos multidimensionales, donde cada dimensión representa un aspecto diferente de los datos (como tiempo, geografía, o producto).
Esto permite a los usuarios:



Explorar datos complejos de manera intuitiva, haciendo "drill down" para detalles o "drill up" para resúmenes.



Realizar análisis comparativos entre diferentes periodos, regiones, o segmentos de mercado.

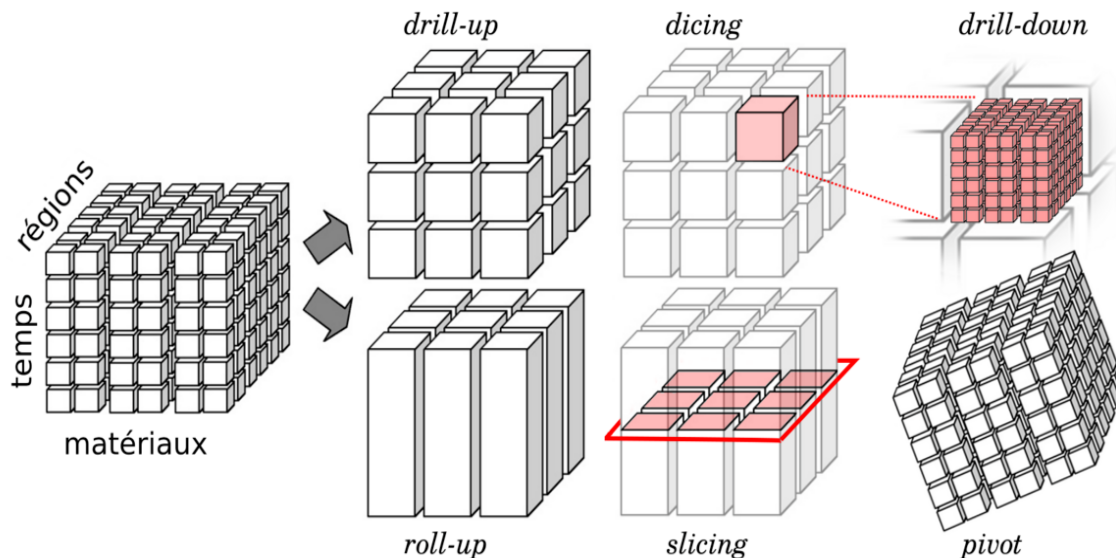


Identificar tendencias y patrones que no serían evidentes en sistemas de procesamiento transaccional (OLTP).

Drill Down vs Drill Up

El "drill down" es el proceso de desglosar los datos para ver información más detallada. Esta técnica permite a los usuarios moverse de un nivel de agregación alto a un nivel más bajo, explorando datos más específicos.

El "drill up" (también conocido como "roll up") es el proceso de consolidar los datos para ver información a un nivel más alto de agregación. Esta acción permite a los usuarios pasar de ver datos detallados a una vista más generalizada.



1

Análisis en tiempo real

Permite decisiones empresariales rápidas basadas en la última información disponible.

2

Flexible y escalable

Capaz de manejar desde pequeñas hasta enormes cantidades de datos.

3

Mejora la inteligencia empresarial

Al proporcionar un acceso fácil a los datos analíticos, mejora la capacidad de la organización para entender su rendimiento y planificar estratégicamente.

Tipos de bases de datos OLAP

MOLAP (Multidimensional OLAP)

Este modelo almacena los datos en un formato multidimensional, lo que permite consultas muy rápidas.

Es ideal para volúmenes de datos menores a medianos donde el rendimiento de la consulta es crítico.

ROLAP (Relacional OLAP)

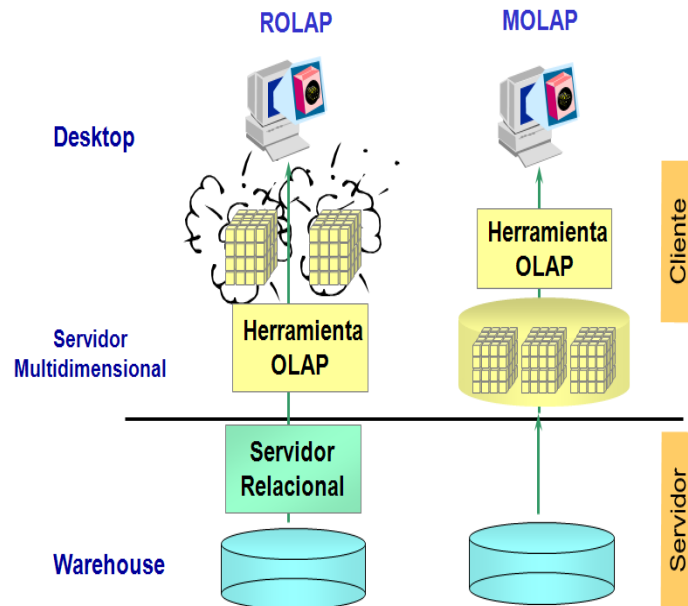
Utiliza bases de datos relacionales estándar y técnicas de modelado para almacenar datos y relaciones.

Es más escalable que MOLAP pero puede ser más lento debido a la falta de precálculo.

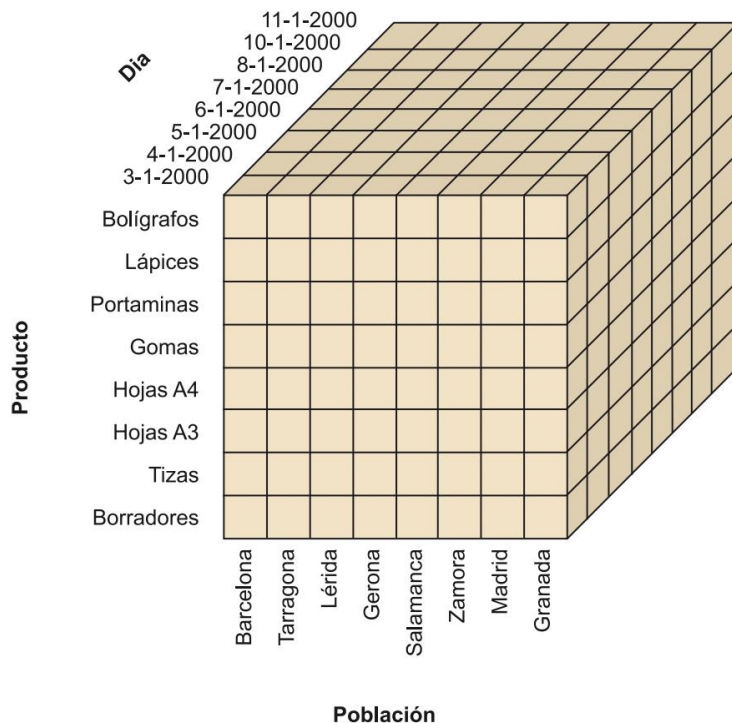
HOLAP (Híbrida OLAP)

Combina las tecnologías MOLAP y ROLAP para ofrecer un equilibrio entre almacenamiento eficiente y rendimiento de consulta.

Puede manejar grandes volúmenes de datos y proporciona un rendimiento razonablemente rápido.



Modelo multidimensional



Un modelo multidimensional
nos ayuda a analizar las
métricas desde distintas
perspectivas

Debido a que W. H. Inmon, es reconocido mundialmente como el padre del DW, la explicación de las características más sobresalientes de este concepto se basó en su definición.



Data Warehouse, características.

Cabe aclarar que los términos almacén de datos y depósito de datos, son análogos a DW, y se utilizarán de aquí en adelante para referirse al mismo.



Estructura de una base de datos multidimensional



datapath

Una base de datos multidimensional tiene 2 tipos de tablas.

1. Tabla de hechos
2. Tabla de dimensiones

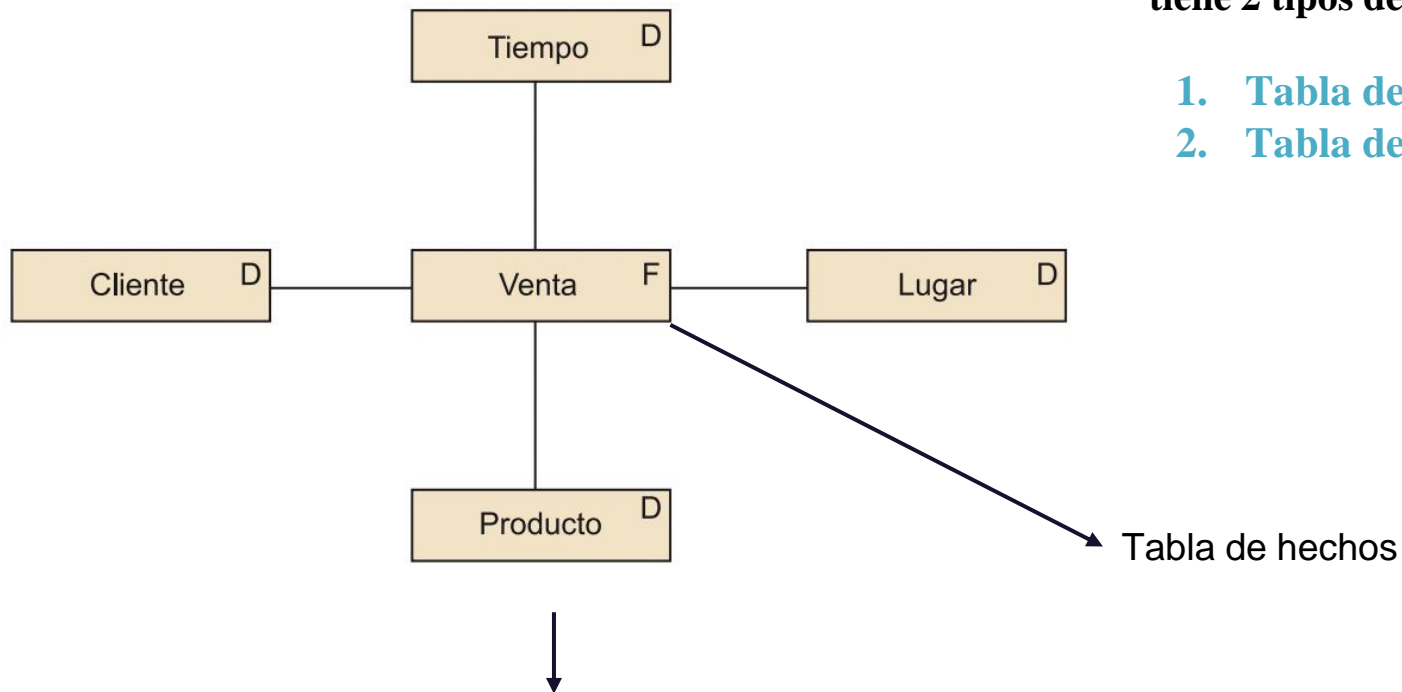
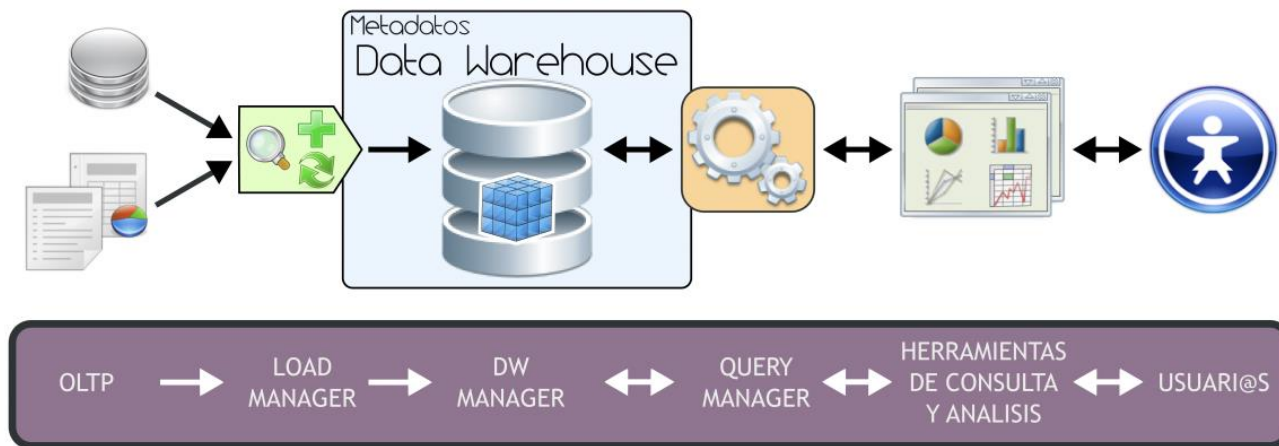


Tabla de dimensión

Arquitectura de un Data Warehouse

En este punto y teniendo en cuenta que ya se han detallado claramente las características generales del Data Warehousing, se definirán y describirán todos los componentes que intervienen en su arquitectura o ambiente.

A través del siguiente gráfico se explicitará la estructura del Data Warehousing:



Data Warehousing, arquitectura.

Modelado - Conceptos claves

Dimensiones

- Representan perspectivas diferentes desde las cuales se desea analizar los datos, como tiempo, geografía, productos, etc. Son los ejes alrededor de los cuales se organizan los informes.

Hechos

- Se refieren a las métricas y medidas cuantificables que se están analizando, como ventas, costos, y unidades vendidas. Los hechos se almacenan en tablas de hechos.

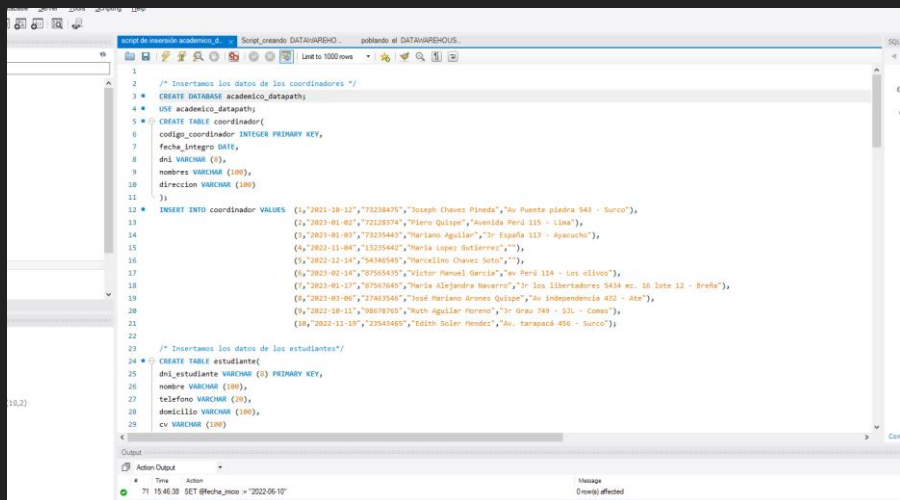
Métricas claves

- Son indicadores de rendimiento (KPIs) derivados de los datos, importantes para la toma de decisiones y análisis estratégico.

Interacción de componentes

- Cómo interactúan: Las dimensiones y los hechos se organizan en un esquema (como el esquema estrella o copo de nieve) que facilita la consulta y el análisis eficiente de los datos. Las métricas clave se calculan a partir de esta organización.

Implementación de un modelo multidimensional



```
1  -- Script de creación de la base de datos para el modelo multidimensional
2  /* Insertamos los datos de los coordinadores */
3  CREATE DATABASE academico_datapath;
4  USE academico_datapath;
5  CREATE TABLE coordinador(
6  codigo_coordinador INTEGER PRIMARY KEY,
7  fecha_inicio DATE,
8  dni VARCHAR (9),
9  nombres VARCHAR (100),
10 direccion VARCHAR (100)
11 );
12 INSERT INTO coordinador VALUES (1,"2021-10-10","73216475","Joseph Chavez Pineda","Av. Puente piedra 545 - Surco"),
13 (2,"2023-01-03","73212874","Diego Quiroga","Avenida Perú 115 - Lima"),
14 (3,"2023-01-03","73215443","Mariano Aguilar","Dr. Equiza 133 - Apuríhuco"),
15 (4,"2022-11-04","13215442","María López Gutiérrez",""),
16 (5,"2022-12-14","74636545","Marcelino Chavez Soto",""),
17 (6,"2023-02-14","87505435","Victor Manuel Garcia","Av. Perú 134 - Los Olivos"),
18 (7,"2023-01-17","87505445","María Alejandra Navarro","Dr. Los Libertadores 5434 ec. 10 Site 12 - Breña"),
19 (8,"2023-03-06","27463546","José Mariano Arones Quiroga","Av. Independencia 432 - Ate"),
20 (9,"2022-10-13","98078765","Ruth Aguilar Norero","Dr. Grau 748 - SCL - Comas"),
21 (10,"2022-11-28","23543485","Edith Soler Renteria","Av. Tarapacá 456 - Surco");
22
23 /* Insertamos los datos de los estudiantes */
24 CREATE TABLE estudiante(
25 dni_estudiante VARCHAR (9) PRIMARY KEY,
26 nombre VARCHAR (100),
27 telefono VARCHAR (20),
28 domicilio VARCHAR (100),
29 cv VARCHAR (100)
30 );
31
```

Output

Time	Action	Message
71:15:46:38	SET @fecha_inicio = '2022-06-10'	(rows) affected

Guía Construcción de un Datawarehouse:

<https://drive.google.com/file/d/12eRJb9w19l8RaYbxiNMpz0jomiO2r4lK/view?usp=sharing>



¡Gracias!

Aprende, aplica y crece