

Online Analytical Processing (OLAP)

Gestión de la Información
Grado en Ingeniería Informática
Universidad de Burgos



José Ignacio Santos, José Manuel Galán
jisantos@ubu.es, jmgalan@ubu.es

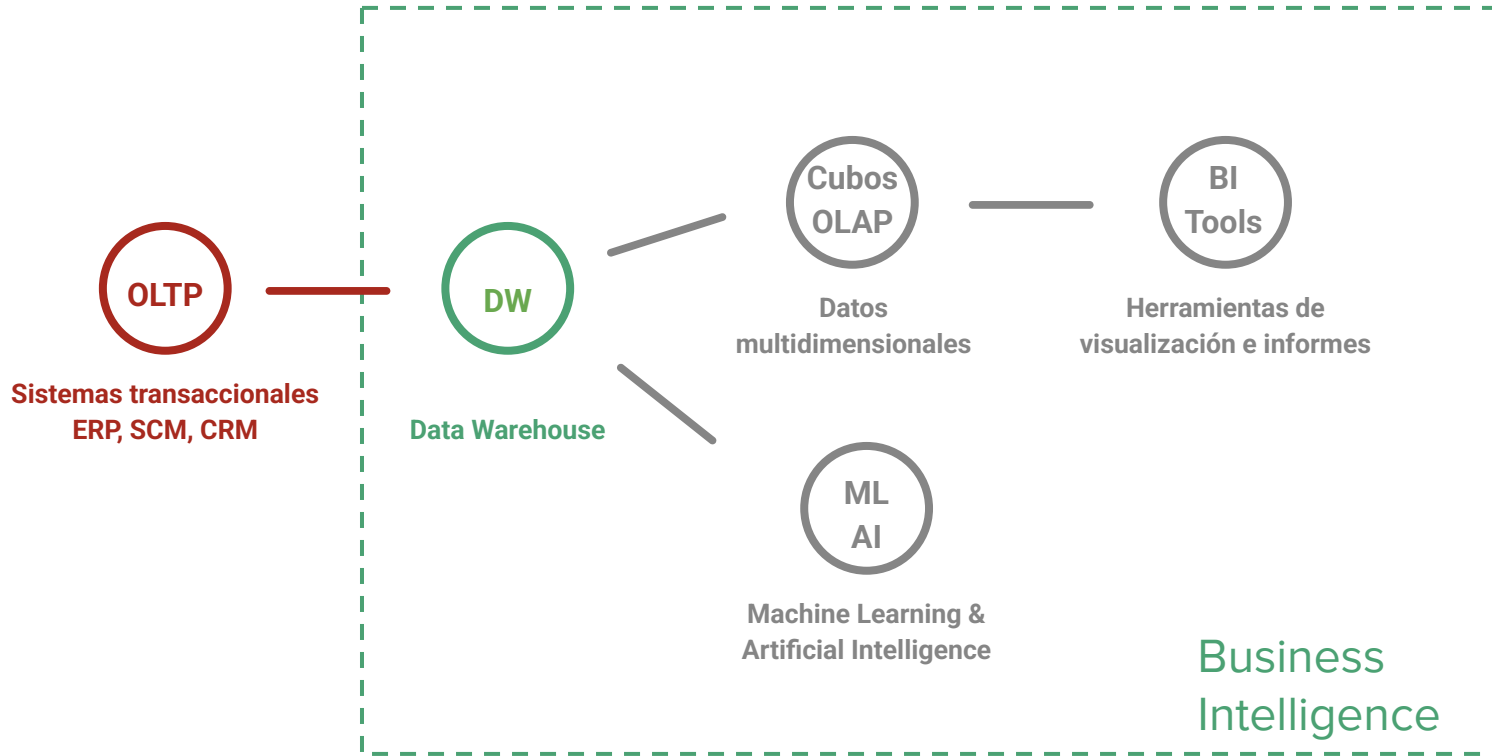
OLTP vs OLAP

	Online Transaction Processing (OLTP)	Online Analytical Processing (OLAP)
Finalidad	Procesamiento transaccional	Análisis y reporting
Tipo de datos	Detallados, en tiempo real	Agregados, históricos
Consultas	Rápidas y cortas	Complejas y multidimensionales
Base de datos	Relacional (SQL)	Multidimensional (cubos)
Ejemplo	Pedido de un cliente hoy	Ventas por región y año

ERP

BI

Ciclo de datos



Ciclo de datos (ejemplo)



OLTP gestiona la bases de datos **ERP** (e.j. tiendas donde se registran todas las operaciones de venta diarias)



Los datos transaccionales (y de otras fuentes) se **consolidan** en bases de datos independientes (**data warehouse** /data lake)



Un modelador define previamente la **estructura de los cubos**:

qué **dimensiones** (para segmentar los datos)

qué **métricas** (campos calculados mediante una función de agregación)

Periódicamente se **calculan/actualizan** los cubos

Los cubos alimentan objetos de visualización (tablas dinámicas, gráficos, scorecards, ...)

Cubo OLAP

Un **cubo** OLAP es una estructura de datos **multidimensional**:

Dimensiones: año, región, cliente

Métricas: ventas, costes, margen

Cada **celda** del cubo es una métrica indexada por los diferentes valores de las dimensiones del cubo

Los cubos OLAP son óptimos para consultas multidimensionales. Los informes y consultas deben ser **coherentes con la estructura** del cubo. Ejemplo de consulta OLAP:

```
SELECT SUM(ventas) FROM fact_ventas WHERE año=2024 AND region='Norte'
```

En el ejemplo el cubo OLAP fact_ventas tiene como dimensiones año y región y como métrica agregada ventas

Google Looker como metáfora de OLAP

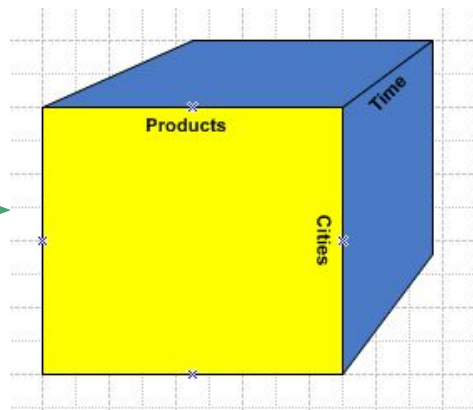
El diseño de Google Looker puede ayudarnos a comprender los cubos OLAP

En un **conector de Looker** definimos campos de tipo **dimensión** (que permiten segmentar los datos) y campos de tipo **métrica** (que se muestran calculados mediante alguna función de agregación), al igual que hacemos con la estructura de un cubo OLAP

La **tabla extendida de dimensiones y métricas** de Looker representaría un cubo multidimensional OLAP

	Año	Región	Nproducto	Importe venta
1.	2018	Europe	botellín 500cc	2,112.5
2.	2018	Australia and Ocea...	botella 5l	3,132
3.	2018	Central America an...	botellín 300cc	438.9
4.	2018	Australia and Ocea...	garrafa 8l	1,058.5
5.	2018	Middle East and N...	botella 0.5l	426
6.	2018	Asia	garrafa 8l	6,597.5
7.	2018	Central America an...	botella 0.5l	1,296

← dimensiones → ← métrica →



Google Looker como metáfora de OLAP

Pero Looker y OLAP son **diferentes**:

Looker no genera cubos, sólo traduce las selecciones del usuario en queries SQL contra las tablas subyacentes

Las agregaciones en Looker se calculan en tiempo de ejecución, no son precalculadas

En Looker tenemos más flexibilidad para construir las queries SQL, en cambio en OLAP las **consultas son limitadas** a la estructura multidimensional de los cubos:

- drill-down** (bajar al detalle)

- roll-up** (agrupar)

- slice** (filtrar una dimensión)

- dice** (cruzar varias dimensiones)

Machine Learning /AI

Las herramientas de ML/AI trabajan sobre **datos detallados del Data Warehouse** o Data Lake, no sobre cubos OLAP

OLAP sigue siendo óptimo para **análisis descriptivo**

ML/AI complementa el BI con analítica predictiva, no supervisada, etc.