

# Universidad Autónoma del Estado de Hidalgo

Escuela Superior Huejutla





Área Académica: Sistemas Computacionales

Tema: Explotación de un Almacén de Datos:  
Herramientas OLAP

Profesor: Mtro Felipe de Jesus Nuñez Cardenas

Periodo: Agosto Noviembre 2011

Keywords

Herramientas OLAP, OLTP.





# Tema: Explotación de un Almacén de Datos: Herramientas OLAP

## Abstract

Datawarehouse are used by different OLAP tools, which help you get a better analysis of information from them reports we can generate.

Keywords:

Herramientas OLAP, OLTP





## Herramientas OLAP

- ✓ Las herramientas de OLAP presentan al usuario una visión multidimensional de los datos (esquema multidimensional) para cada actividad que es objeto de análisis.
- ✓ El usuario formula consultas a la herramienta OLAP seleccionando atributos de este esquema multidimensional sin conocer la estructura interna (esquema físico) del almacén de datos.
- ✓ La herramienta OLAP genera la correspondiente consulta y la envía al gestor de consultas del sistema (p.ej. mediante una sentencia SELECT).





# Herramientas OLAP

una consulta a un almacén de datos consiste generalmente en la obtención de **medidas** sobre los **hechos** parametrizadas por atributos de las **dimensiones** y restringidas por **condiciones** impuestas sobre las dimensiones



**Restricciones:** productos del departamento Bebidas, ventas durante este año

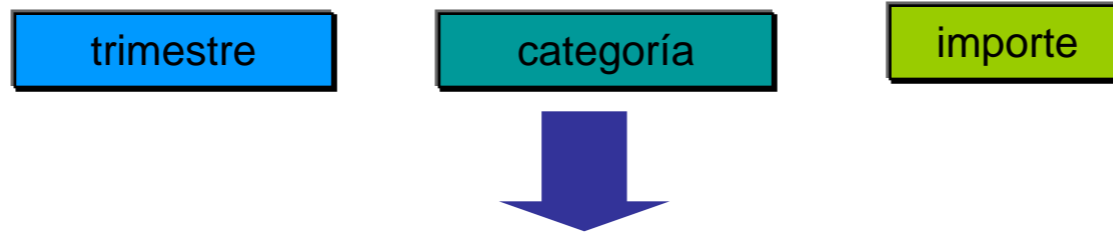
**Parámetros de la consulta:** por categoría de producto y por trimestre







# Herramientas OLAP



Oracle Discoverer - [Vidstr31.dis]

File Edit View Sheet Format Tools Graph Window Help

Min Max % + - x ÷ = ≠ < ≤ > ≥ ▾

Arial 10 B U

**Análisis de Alquileres y Ventas de video**  
15-NOV-01

Page Items: Departamento: Beverage ▾

	Beneficios			región en el	Porcentaje de cada año en la región		
	1995	1996	Total (1995-1996)		1995	1996	
Central	\$523	\$742	\$1,265	20%	31%	41%	59%
East	\$1,229	\$1,158	\$2,391	47%	48%	51%	49%
West	\$872	\$519	\$1,391	33%	21%	63%	37%
Total	\$2,624	\$2,391	\$5,047				

T-Beneficios (región-dpto-año) TC-Beneficios (Dpto-Region-fecha) mati Sheet 1 Sheet 3 Sheet 4 Sheet 5

**INFORME**





# Herramientas OLAP

Categoría	Trimestre	Ventas
Refrescos	T1	2000000
Refrescos	T2	1000000
Refrescos	T3	3000000
Refrescos	T4	2000000
Zumos	T1	1000000
Zumos	T2	1500000
Zumos	T3	8000000
Zumos	T4	2400000

Presentación tabular (relacional) de los datos seleccionados

Se asumen dos categorías en el departamento de *Bebidas*: Refrescos y Zumos.







# Herramientas OLAP

trimestre / categoría	T1	T2	T3	T4
Refrescos	2000000	1000000	3000000	2000000
Zumos	1000000	1500000	8000000	2400000

Presentación matricial (multidimensional) de los datos seleccionados

Los parámetros de la consulta (“por trimestre” y “por categoría”) determinan los criterios de agrupación de los datos seleccionados (ventas de productos del departamento *Bebidas* durante este año). La agrupación se realiza sobre dos dimensiones (**Producto, Tiempo**).





## Herramientas OLAP

- Lo interesante no es poder realizar consultas que, en cierto modo, se pueden hacer con selecciones, proyecciones, concatenaciones y agrupamientos tradicionales.
- Lo realmente interesante de las herramientas OLAP son sus **operadores de refinamiento o manipulación de consultas.**
  - DRILL
  - ROLL
  - SLICE & DICE
  - PIVOT





## Herramientas OLAP

El carácter agregado de las consultas en el Análisis de Datos, aconseja la definición de nuevos operadores que faciliten la agregación (consolidación) y la disgregación (división) de los datos:

- ✓ agregación (**roll**): permite eliminar un criterio de agrupación en el análisis, agregando los grupos actuales.
- ✓ disgregación (**drill**): permite introducir un nuevo criterio de agrupación en el análisis, disgregando los grupos actuales.





## Herramientas OLAP

Si se desea introducir la dimensión **Almacén** en el análisis anterior e incluir un nuevo criterio de agrupación sobre la ciudad del almacén:

¿ “Importe total de las ventas durante **este año** de los productos del **departamento Bebidas**, por **trimestre**, por **categorías** y por **ciudad del almacén**” ?.

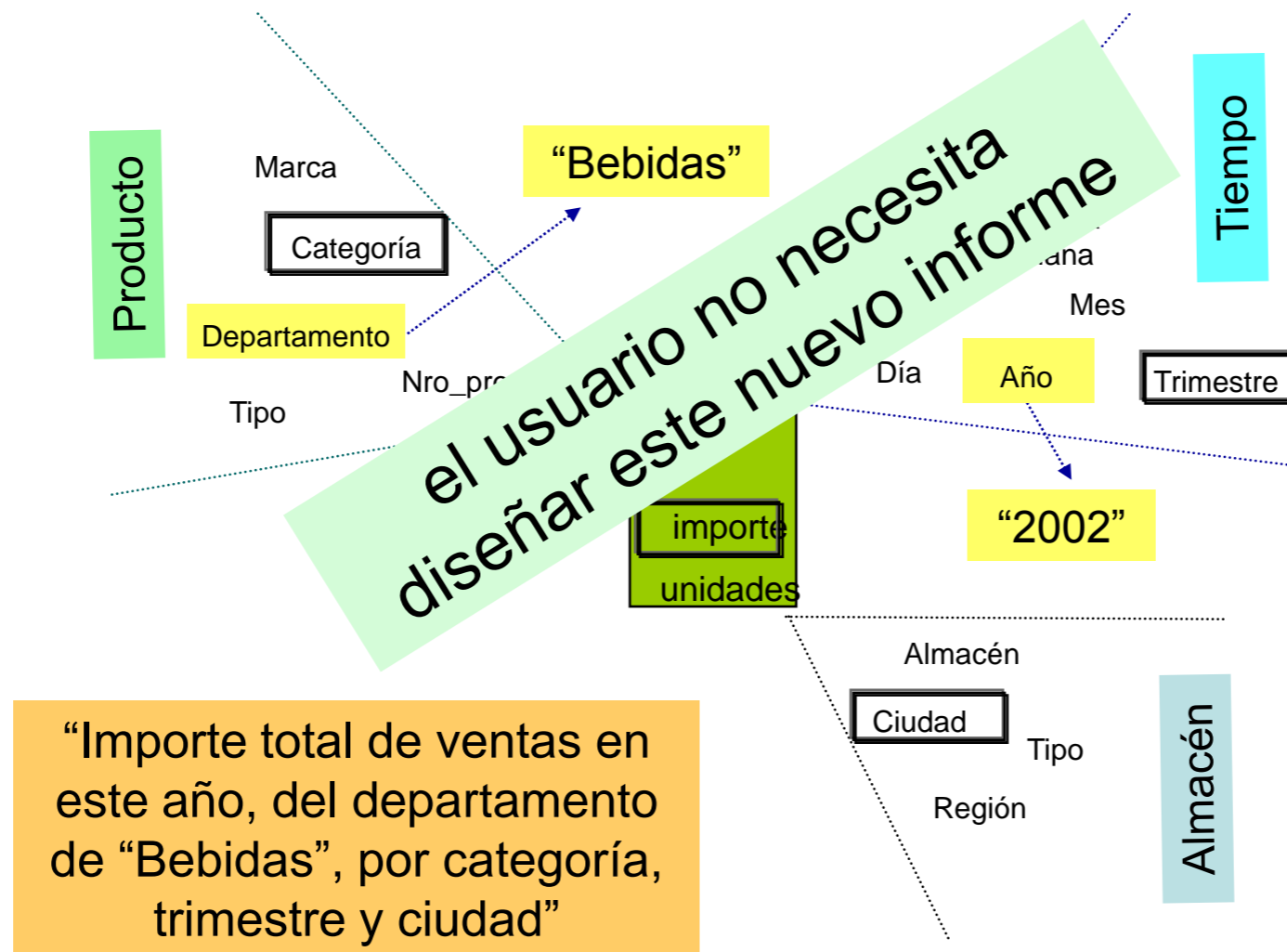
**Restricciones:** productos del departamento Bebidas, ventas durante este año

**Parámetros de la consulta:** por categoría de producto, por trimestre y por ciudad del almacén.





# Herramientas OLAP





# Herramientas OLAP

trimestre

categoría

importe

**DRILL ACROSS**  
Almacén (Ciudad)

	1995		1996	
Central	20%	31%	41%	59%
East	47%	48%	51%	49%
West	33%	21%	63%	37%
Total	\$2,391	\$5,047		

¡ la operación de DRILL se realiza sobre el informe original !

**Informe mas detallado**

	Beneficios		Porcentaje de cada región en el año		Porcentaje de cada año en la región	
	1995	1996	1995	1996	1995	1996
Central	\$523	\$1,265	20%	31%	41%	59%
East	\$1,203	\$2,391	47%	48%	51%	49%
West	\$1,391	\$1,391	33%	21%	63%	37%
Total	\$2,391	\$5,047				





# Herramientas OLAP

Categoría	Trimestre	Ventas
Refrescos	T1	2000000
Refrescos	T2	1000000
Refrescos	T3	3000000
Refrescos	T4	2000000
Zumos	T1	1000000
Zumos	T2	1500000
Zumos	T3	8000000
Zumos	T4	2400000

drill-across

Categoría	Trimestre	Ciudad	Ventas
Refrescos	T1	Valencia	1000000
Refrescos	T1	León	1000000
Refrescos	T2	Valencia	400000
Refrescos	T2	León	700000

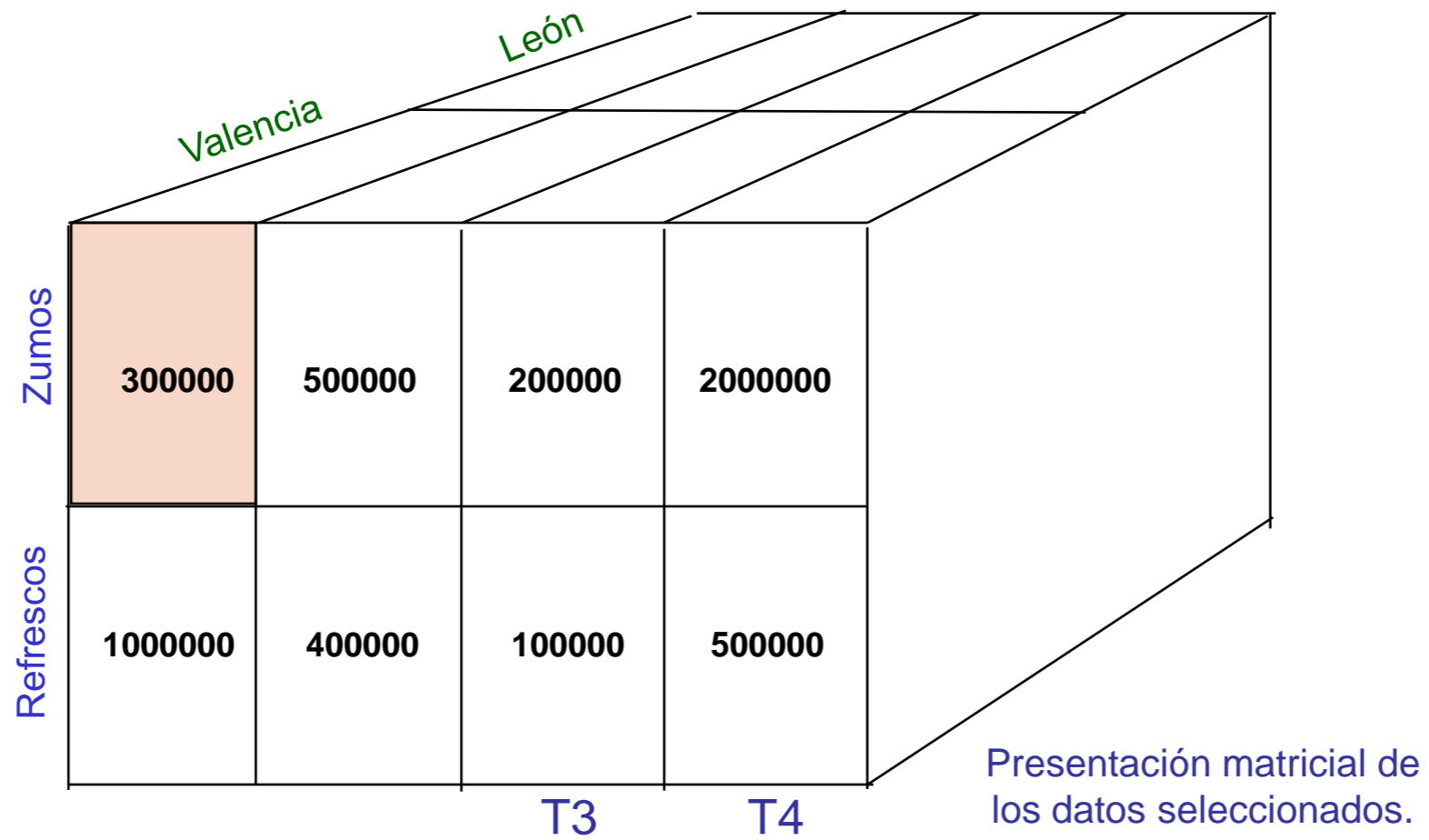
Cada grupo (categoría-trimestre) de la consulta original se disgrega en dos nuevos grupos (categoría-trimestre-ciudad) para las ciudades de León y Valencia.

\* Se asumen dos ciudades: Valencia y León.





# Herramientas OLAP







## Herramientas OLAP

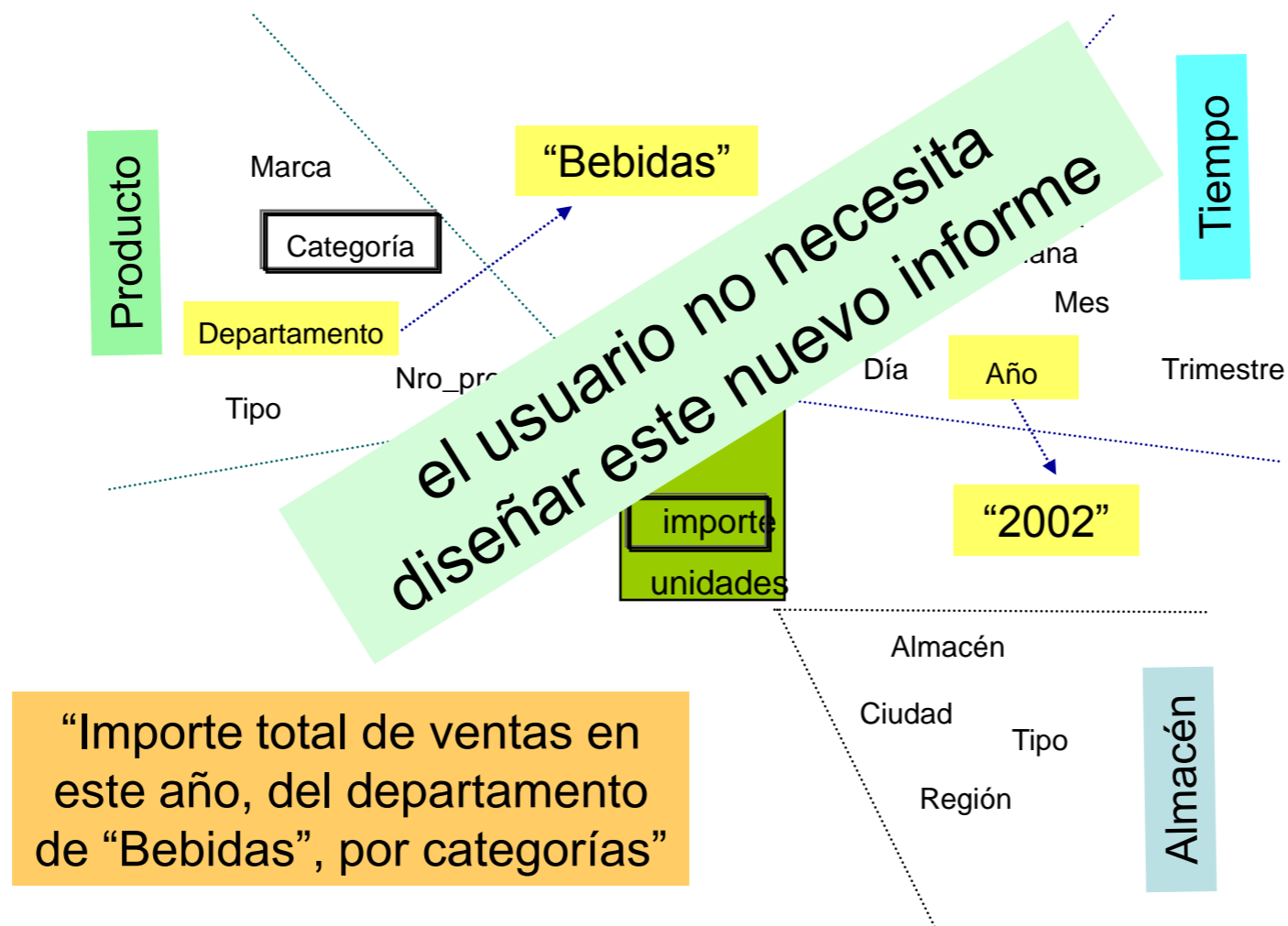
Si se desea eliminar el criterio de agrupación sobre la dimensión **Tiempo** en la consulta original:

¿ “Importe total de las ventas durante este año de los productos del departamento *Bebidas*, por **categorías**” ?





# Herramientas OLAP





# Herramientas OLAP

trimestre

categoría

importe



**ROLL ACROSS  
Tiempo (Trimestre)**

	1995	1996	1995	1996
Central	20%	31%	41%	50%
East	47%	48%	51%	49%
West	33%	21%	63%	37%
Total	\$5,047			

¡ la operación de ROLL se realiza sobre el informe original !



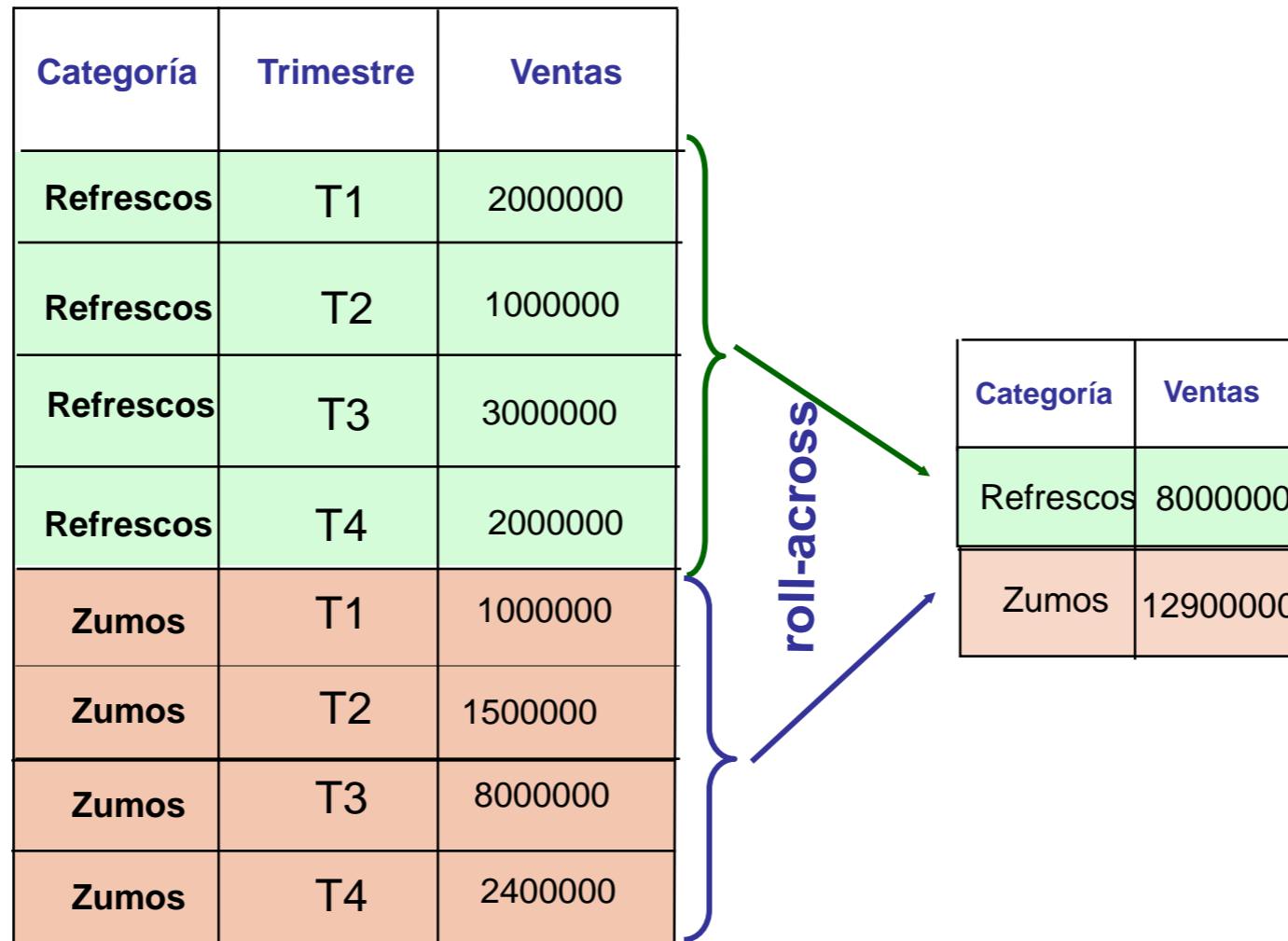
**Informe mas agregado**

	1995	1996	1995	1996
Central	\$208	\$742	31%	41%
East	\$1,929	\$1,929	48%	51%
West	\$607	\$210	21%	63%
Total	\$2,744	\$1,881		





# Herramientas OLAP





# Herramientas OLAP

Las operaciones de agregación (DRILL) y disgregación (ROLL) se pueden hacer sobre:

- ✓ atributos de una dimensión sobre los que se ha definido una jerarquía: **DRILL-DOWN, ROLL-UP**

departamento – categoría - producto (Producto)

año - trimestre – mes - día (Tiempo)

- ✓ sobre dimensiones independientes: **DRILL-ACROSS, ROLL-ACROSS**

Producto – Almacén -Tiempo





# Herramientas OLAP

trimestre

categoría

importe

**DRILL DOWN**  
Tiempo (mes)

	1995		1996	
Central	31%	41%	59%	
East	47%	48%	51%	49%
West	33%	21%	63%	37%
Total				

¡ la operación de DRILL se realiza sobre el informe original !

“Importe total de ventas en este año, del departamento de “Bebidas”, por categoría y mes”

	1995		1996		1996		1996	
Central	\$523	\$742			31%	41%	59%	
East	\$1,229	\$1,162			47%	48%	51%	49%
West	\$872	\$519			33%	21%	63%	37%
Total	\$2,624	\$2,423						





# Herramientas OLAP

Categoría	Trimestre	Ventas
Refrescos	T1	2000000
Refrescos	T2	1000000
Refrescos	T3	3000000
Refrescos	T4	2000000
Zumos	T1	1000000
Zumos	T2	1500000
Zumos	T3	8000000
Zumos	T4	2400000

drill-down

Categoría	Trimestre	Mes	Ventas
Refrescos	T1	Enero	1000000
Refrescos	T1	Febrero	500000
Refrescos	T1	Marzo	500000

Cada grupo (categoría-trimestre) de la consulta original se disgrega en dos nuevos grupos (categoría-trimestre-mes).





## Herramientas OLAP

### Otras operaciones de OLAP:

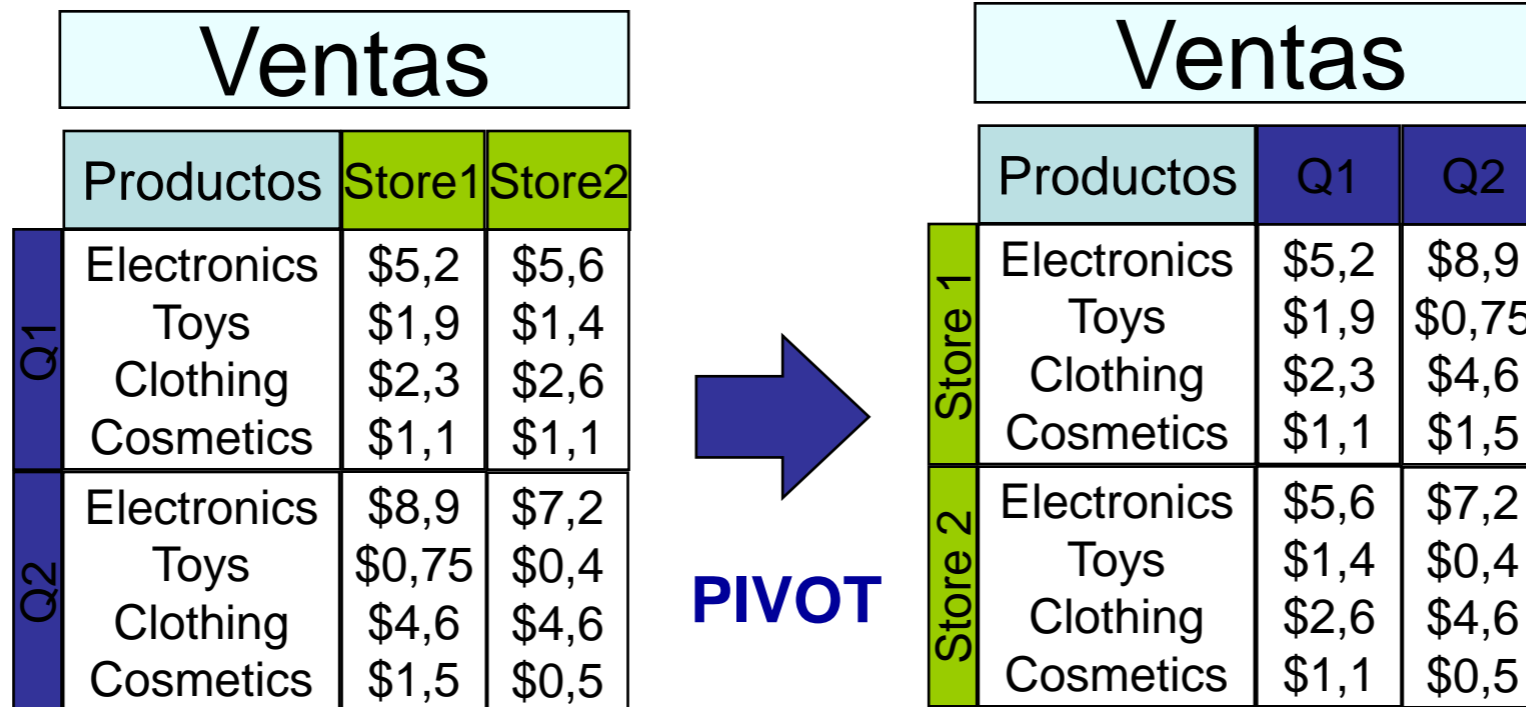
- ✓ SLICE & DICE: seleccionar y proyectar datos en el informe.
- ✓ PIVOT: reorientación de las dimensiones en el informe.





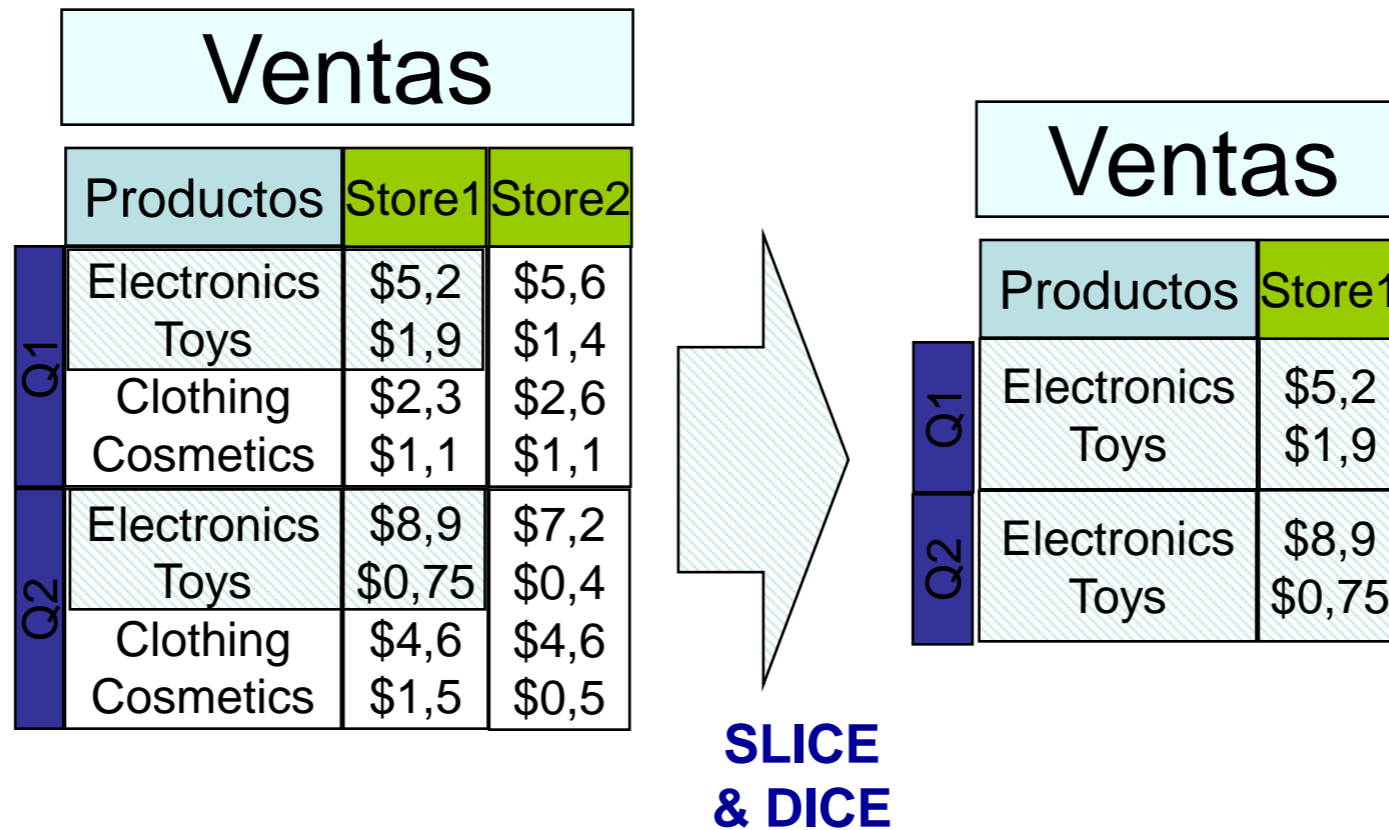


# Herramientas OLAP





# Herramientas OLAP





# Herramientas OLAP

## Las herramientas de OLAP se caracterizan\* por:

- ✓ ofrecer una visión multidimensional de los datos (matricial).
- ✓ no imponer restricciones sobre el número de dimensiones.
- ✓ ofrecer simetría para las dimensiones.
- ✓ permitir definir de forma flexible (sin limitaciones) sobre las dimensiones: restricciones, agregaciones y jerarquías entre ellas.
- ✓ ofrecer operadores intuitivos de manipulación: *drill-down*, *roll-up*, *slice-and-dice*, *pivot*.
- ✓ ser transparentes al tipo de tecnología que soporta el almacén de datos (ROLAP o MOLAP).

\*Subconjunto de las 12 reglas propuestas por E.F. Codd para A.D.





## Bibliografía

- Hand, D.J.; Mannila, H. and Smyth, P. “Principles of Data Mining”, The MIT Press, 2000.
- Hernández, J.; Ramírez, M.J.; Ferri, C. “Introducción a la Minería de Datos” Pearson Prentice Hall, 2004.
- Kosala, R.; Blockeel, H. “Web Mining Research: A Survey” ACM SIGKDD Explorations, Newsletter of the ACM SIG on Knowledge Discovery and Data Mining, June 2000, Vol. 2, nº1, pp. 1-15.
- Mena, Jesus “Data Mining Your Website”, Digital Press, July 1999.
- Mitchell, T.M. “Machine Learning” McGraw-Hill 1997.
- Pyle, D. “Data Preparation for Data Mining” Morgan Kaufmann, Harcourt Intl., 1999.
- Thuraisingham, B. “Data Mining. Technologies, Techniques, Tools, and Trends”, CRC Press, 1999.
- Witten, I.H.; Frank, E. “Tools for Data Mining”, Morgan Kaufmann, 1999.
- Wong, P. C. “Visual Data Mining”, Special Issue of *IEEE Computer Graphics and Applications*, Sep/ Oct 1999, pp. 20- 46.
- Material extraído del Análisis y Extracción de Conocimiento en Sistemas de Información: Datawarehouse y Datamining de **José Hernández Orallo**, **Universidad Politécnica de Valencia**

