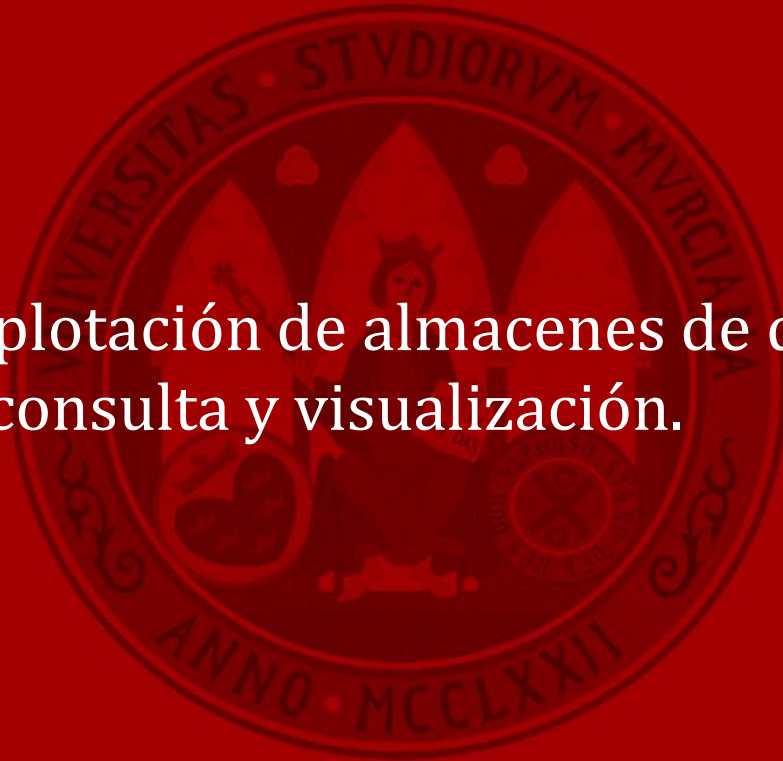


MDX

Unidad 3 – Explotación de almacenes de datos.
Lenguajes de consulta y visualización.
U3.1 - MDX



UNIVERSIDAD DE
MURCIA

Unidad 3. Explotación de almacenes de datos

Estructura

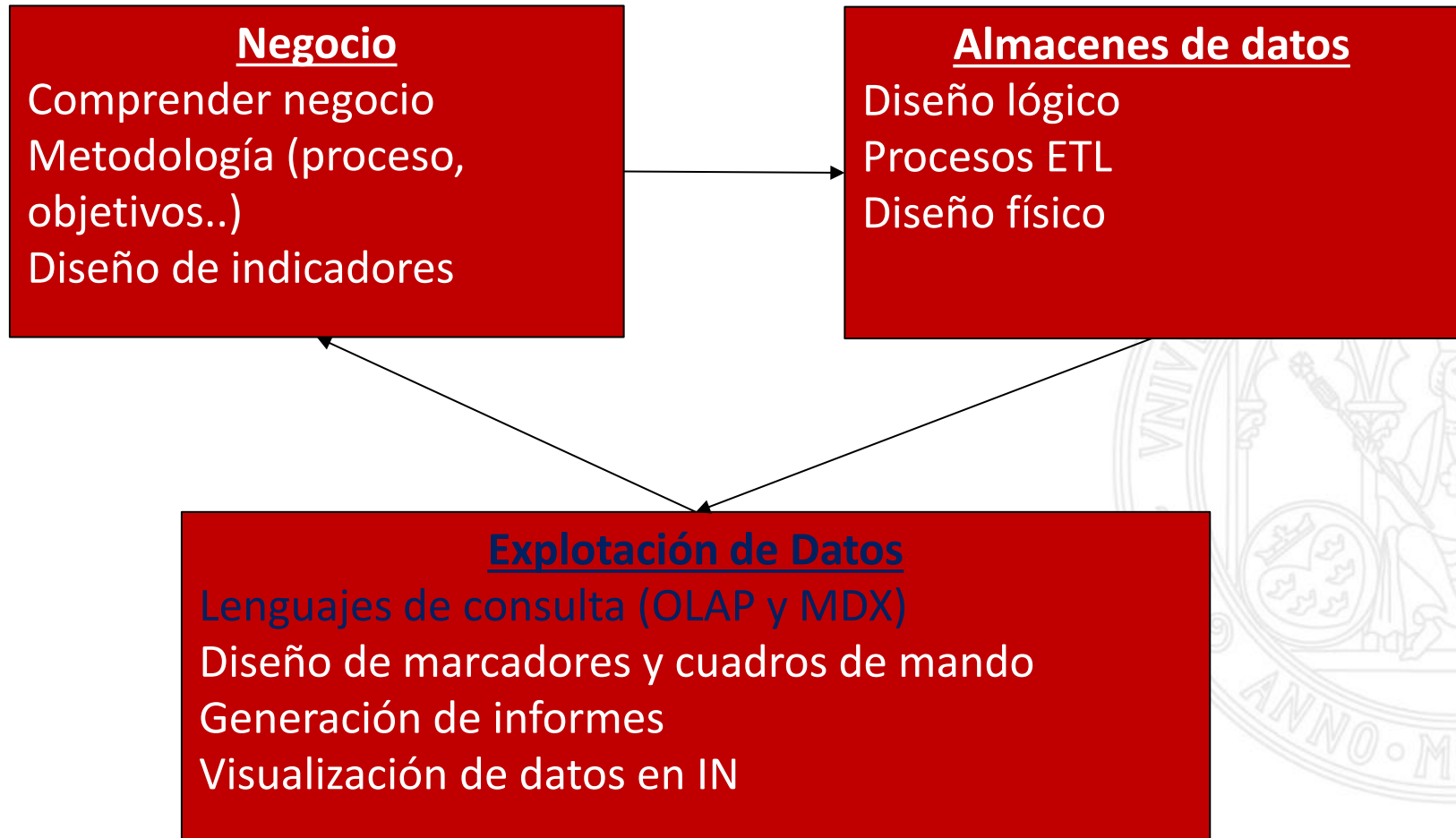
- Tuplas, conjuntos y celdas
- Términos MDX
- Sintaxis de consultas MDX



Inteligencia de negocio

Nuestros contenidos

UNIVERSIDAD DE
MURCIA



CUBOS

LENGUAJE

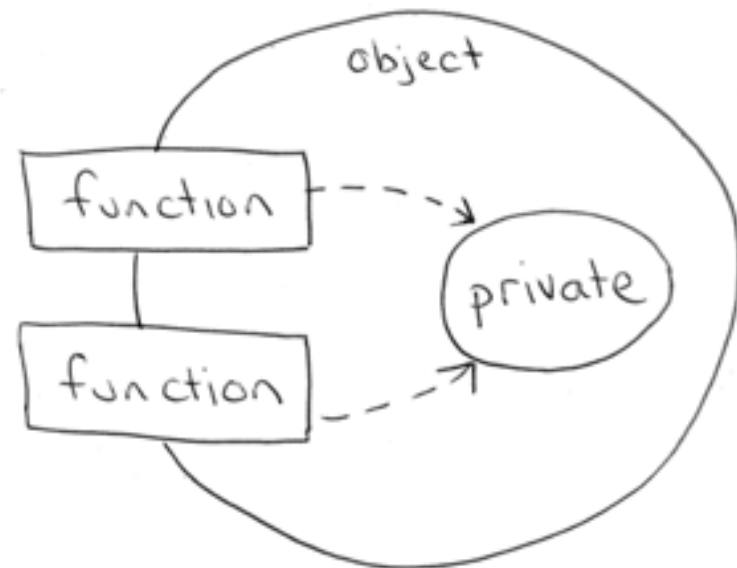
MultiDimensional **eX**pression.

Microsoft en 1997



Unidad 3. Explotación de almacenes de datos MDX

- ¿Recuerdas estudiar Programación Orientada a Objetos?
- Desde la filosofía de POO (agnóstica del lenguaje específico) hasta el código (dependiente del lenguaje de programación)
 - Nos enfocaremos en los conceptos, **NO** en la programación



Unidad 3. Explotación de almacenes de datos - MDX

TUPLAS, CONJUNTOS Y CELDAS



Unidad 3. Explotación de almacenes de datos

¿Qué es MDX?

- **MultiDimensional eXpression (MDX)** proporciona un **lenguaje** para manipular **cubos MDX. Microsoft (1997)**
- Los cubos están definidos por **Dimensiones y Medidas**, conteniendo una **Medida (o más)**, y donde cada Dimensión puede contener múltiples **Miembros**
- El valor de la **celda** se define como la intersección entre Miembros

UnitsSold

	Product			
Time	Sardines	Anchovies	Herrings	Pilchards
April	16	23	12	4
May	14	12	23	6
June	34	19	19	8
July	17	22	14	4

Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

Medidas | Dimensiones > Miembros

Ejemplo: cubo de dos dimensiones

1 medida: pacientes dados de alta.

Dimensión Tiempo con 4 miembros: Enero a Abril.

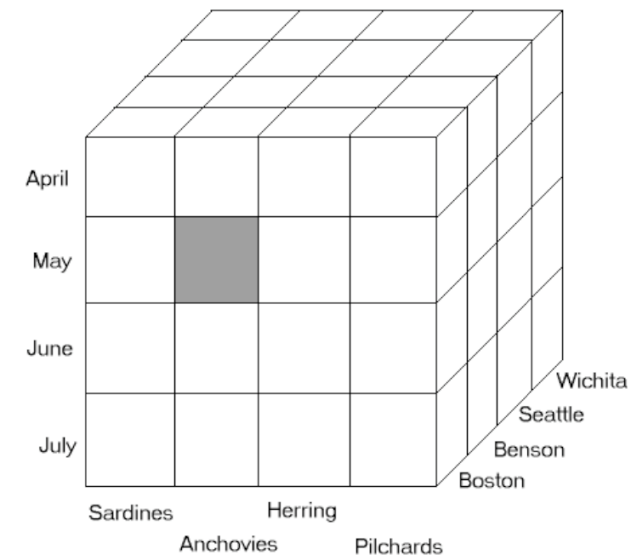
Dimensión Hospital con 4 miembros: H1,H2,H3,H4.

Discharged	H1	H2	H3	H4
January	20	44	81	44
February	15	32	78	32
March	23	65	88	65
April	19	67	67	67

Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

- Podemos generalizar de la siguiente forma:
 - Un cubo puede tener más de una medida por intersección
 - Un cubo puede tener más de dos dimensiones, generando un hipercubo (n-cubo)
 - Cubo en servicios de análisis: 1024 medidas, 128 dimensiones (con miles o millones de miembros)



UnitsSold/Profit

	Product						
Time	Sardines		Anchovies		Herrings		Pilchards
April	16	\$40.00	23	\$78.20	12	\$23.88	4 \$8.20
May	14	\$35.00	12	\$40.80	23	\$45.77	6 \$12.30
June	34	\$85.00	19	\$64.60	19	\$37.81	8 \$16.40
July	17	\$42.50	22	\$74.80	14	\$27.86	4 \$8.20

Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

Medidas | Dimensiones > Miembros

Ejemplo: cubo de 2 dimensiones

2 medidas: nº pacientes dados de alta, coste total (M€).

Dimensión Tiempo con 4 miembros: Enero a Abril.

Dimensión Hospital con 4 miembros: H1,H2,H3,H4.

Discharged	H1	H2	H3	H4
January	20 1.5M€	44 4.1M€	81 10.5M€	44 4.1M€
February	15 1.1M€	32 3.9M€	78 10.4M€	32 3.9M€
March	23 1.6M€	65 5.4M€	88 10.7M€	65 5.4M€
April	19 1.5M€	67 5.6M€	67 9.5M€	67 5.6M€

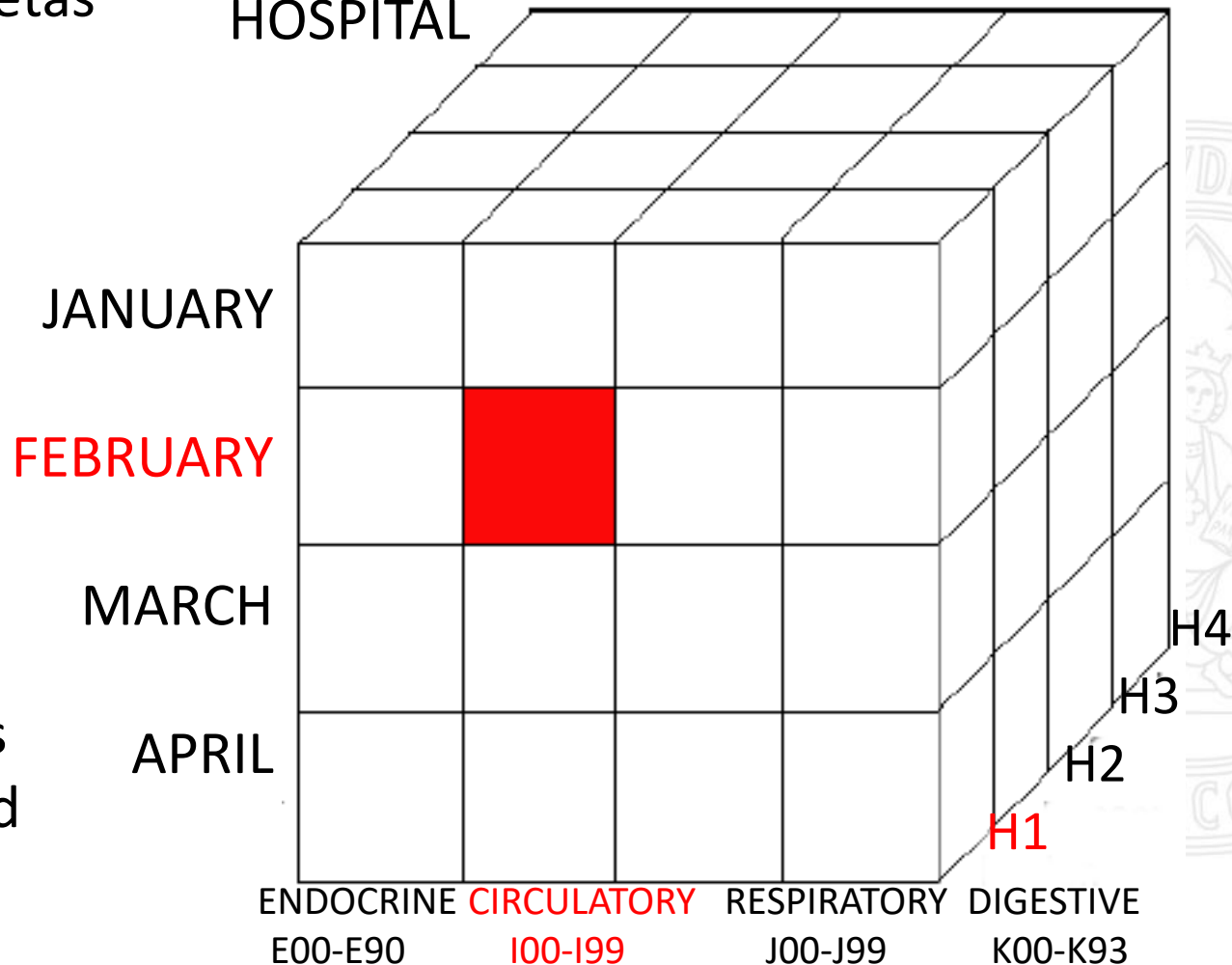
Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

UNIVERSIDAD DE
MURCIA

¿Qué son las etiquetas
en cada eje?

CUBO: 3 DIMENSIONES TIEMPO, DX,
HOSPITAL

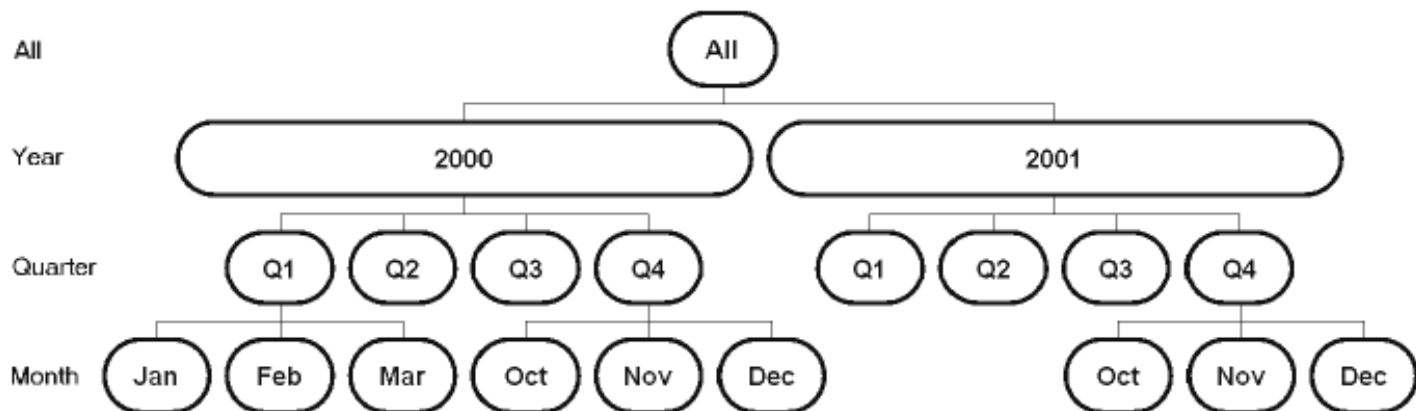


Pero... necesitamos
aún más flexibilidad
para un business
analyst

Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

- Una Dimensión puede tener **jerarquías**
 - No es obligatorio que todas las dimensiones las tengan
 - La mayoría de cubos tienen una dimensión Tiempo (Time) - siempre jerárquica.
- Una jerarquía tiene **niveles**: All, Year, Quarter, Month
 - Cada miembro puede dividirse en respectivos sub-miembros
 - **Estructura de Árbol** : desde el nodo **superior** ("All"), bajando hasta los nodos **hoja**.



Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

- Los valores de las celdas **agregan** los datos originales basándose en los niveles de jerarquía
- La agregación puede no estar presente en los datos originales
 - Podemos aplicar funciones como *percentage()* o *mean()*
 - El motor subyacente buscará optimizar el sistema
 - ¿Se parece a Excel?

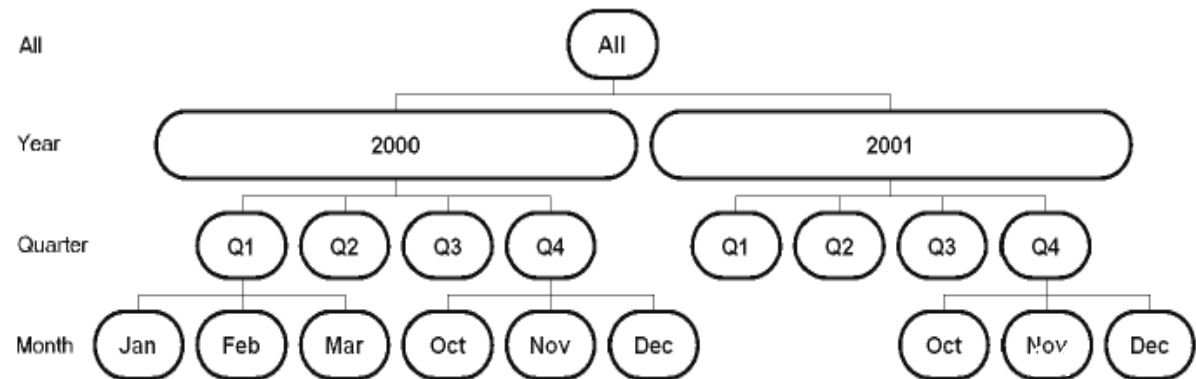
			Sardines	Anchovies
2000	Q3	July	17	22
		Aug	16	18
		Sept	12	19
	Q3 total		45	59
2000	Q4	Oct	27	19
		Nov	24	19
		Dec	21	12
	Q4 total		72	50
2000 total			242	199

Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

Convenciones de nomenclatura

- Enfoque detallado pero más confiable: [Time].[All].[2000].[Q4].[Oct]
- Enfoque más corto, pero dependiente de la nomenclatura: [Time].[October]
 - Puedes imponer convenciones de nombres: [Time].[October-2000]
- En grandes conjuntos de datos reales aún podrían surgir problemas, por ejemplo, con ciudades.
 - Member keys – Claves de miembros (unique)



Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

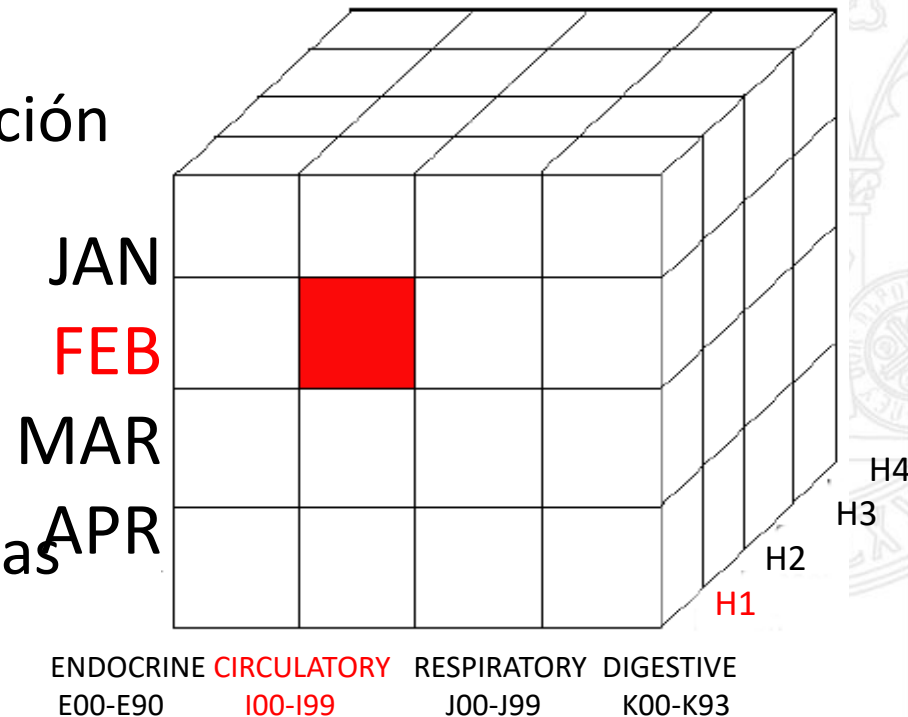
- Convenciones de nomenclatura: **Tupla**

Tupla en pseudo-MDX: $(x,y,z)=(y,z,x)$ (el orden no es importante)

$([Time].[Feb],[Dx].[Circ],[Hosp].[H1])$

- Def1: “Una **Tupla** es la intersección eligiendo **un miembro de cada dimensión**”

- Def2: “Una tupla es la Intersección de un (y solo un) miembro tomado de una o varias dimensiones en el cubo”



Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

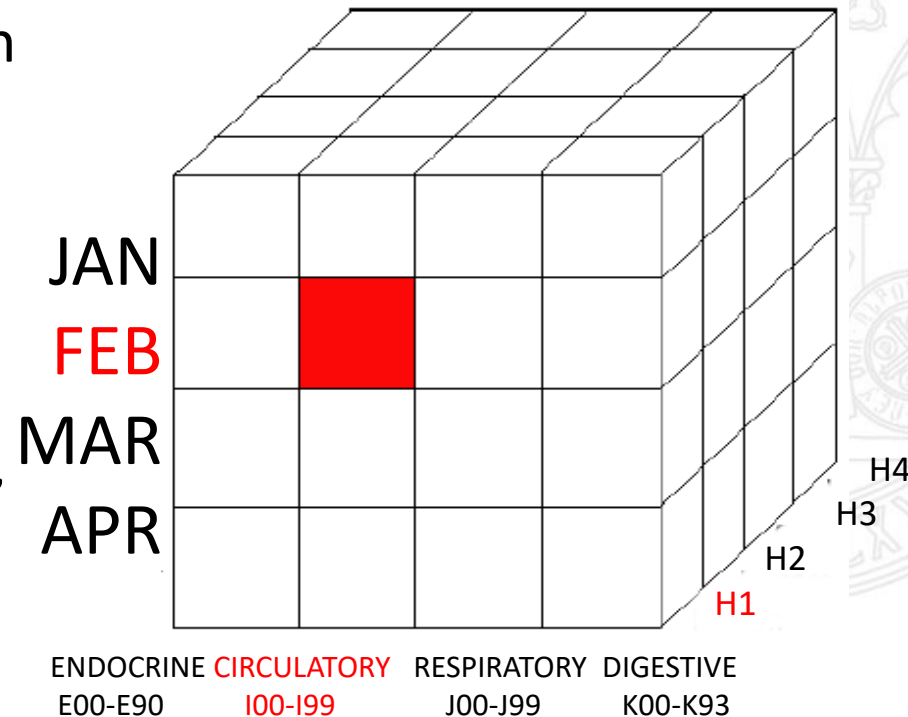
- Convenciones de nomenclatura: **Tupla**

Tupla en pseudo-MDX: $(x,y,z)=(y,z,x)$
([Time].[Feb],[Dx].[Circ],[Hosp].[H1])

Def1: “Una **Tupla** es la intersección eligiendo **un miembro de cada dimensión**”

Def2: “Una tupla es la intersección de un (y solo un) miembro tomado de **una o de varias dimensiones** en el cubo”

Tupla= ¿Es una sola celda en el cubo?



Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

- ¿Cómo definirías esta celda usando una Tupla?
- ¿Sería esta tupla 44?
- Importante distinguir entre la tupla vs el contenido de la celda
- Cada miembro debe provenir **de una dimensión diferente** para señalar a una única celda

	A	B	C
1	12	32	45
2	37	23	12
3	65	45	32
4	78	56	44
5	98	23	34
6	290	179	167

Unidad 3. Explotación de almacenes de datos

Tuplas, conjuntos y celdas

- Convenciones de nomenclatura: **Conjunto**

Conjunto en pseudo-MDX: $\{(x1,y1,z1),\dots,(xn,yn,zn)\}$

$\{([Time].[Feb],[Dx].[Circ],[Hosp].[H1]),$
 $([Time].[Mar],[Dx].[Circ],[Hosp].[H1])\}$

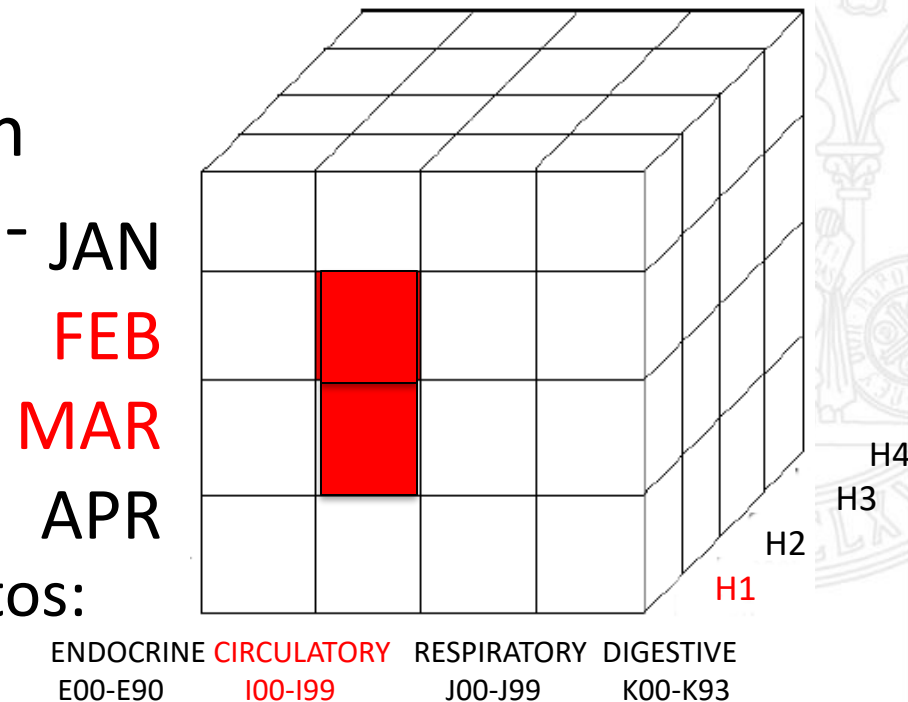
“Un **Conjunto** es una colección
De tuplas con la misma dimen-
sionalidad”

(conjunto de celdas en el cubo)

(podría tener 0(conjunto vacío) o 1 tupla)

Las funciones pueden usar conjuntos:

AVG(SET)→FLOAT



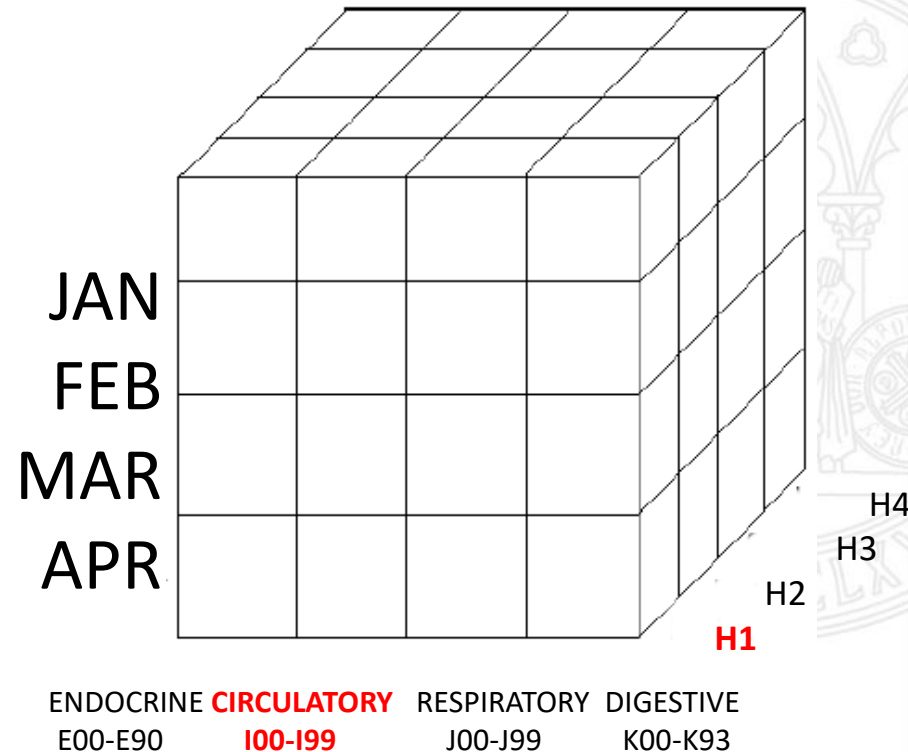
Unidad 3. Explotación de almacenes de datos

Pregunta 1: Tuplas, conjuntos y celdas

UNIVERSIDAD DE
MURCIA

Pregunta: ¿Tupla o Conjunto?

[DX].[Circulatory],[Hosp].[H1]



Unidad 3. Explotación de almacenes de datos

Pregunta 1: Tuplas, conjuntos y celdas

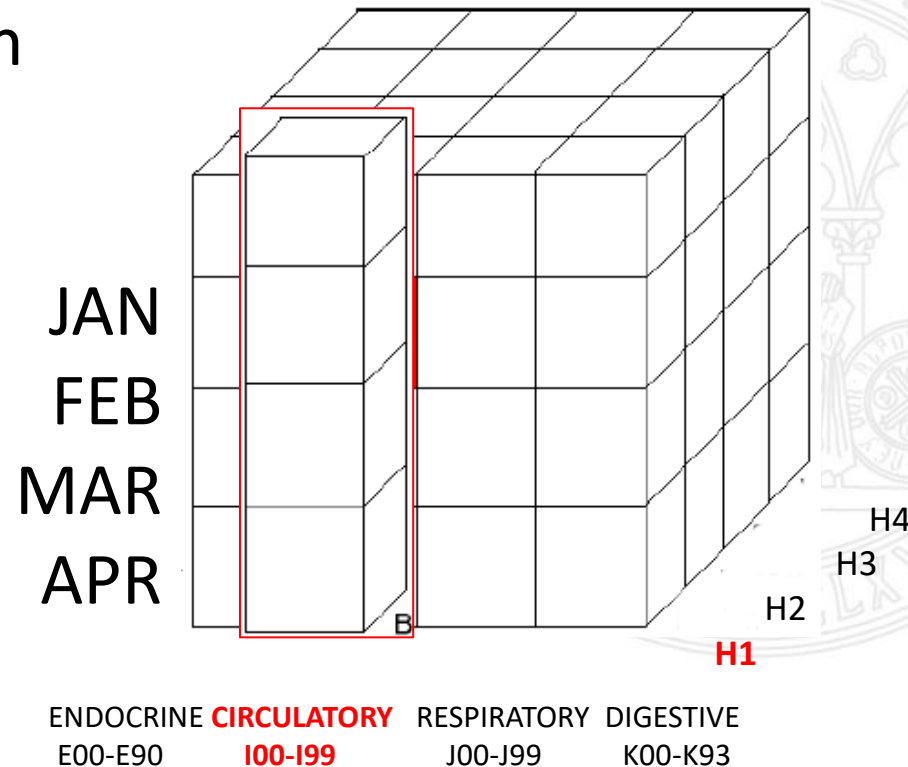
UNIVERSIDAD DE
MURCIA

Pregunta: ¿Tupla o Conjunto?

[DX].[Circulatory],[Hosp].[H1]

Sin restricciones en la dimensión faltante

¡Es una TUPLA!
(pero MUCHAS CELDAS)



Unidad 3. Explotación de almacenes de datos

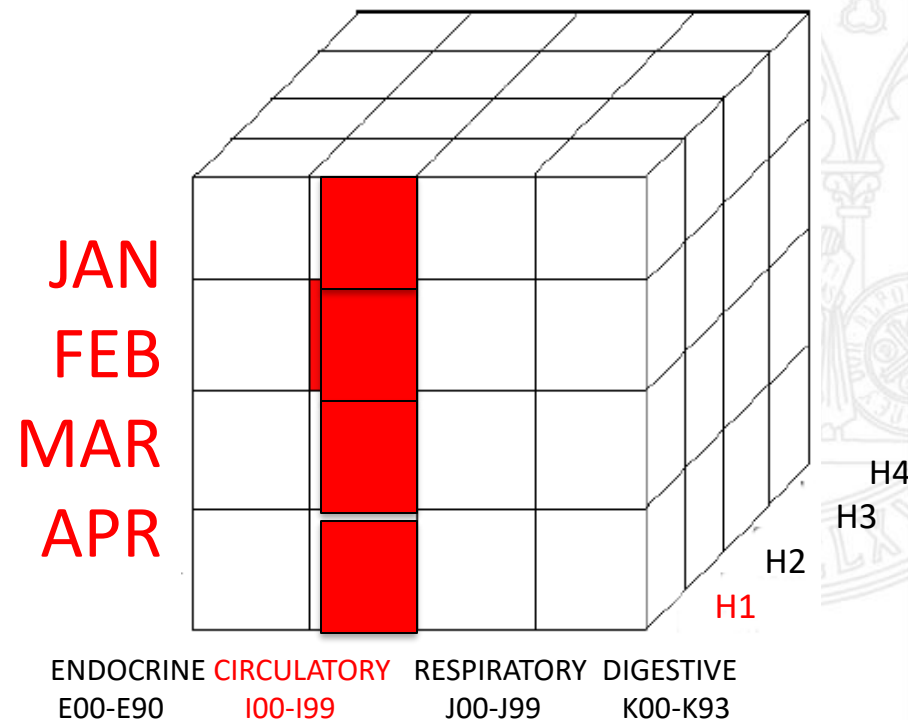
Pregunta 2: Tuplas, conjuntos y celdas

Pregunta. Diferencia entre:

a) $([Dx].[Circ],[Hosp].[H1])$

b)

$\{([Dx].[Circ],[Hosp].[H1],[Time].[Jan]),$
 $([Dx].[Circ],[Hosp].[H1],[Time].[Feb]),$
 $([Dx].[Circ],[Hosp].[H1],[Time].[Mar]),$
 $([Dx].[Circ],[Hosp].[H1],[Time].[Apr])$
 $\}$



Unidad 3. Explotación de almacenes de datos

Pregunta 2: Tuplas, conjuntos y celdas

Pregunta. Diferencia entre... ¿Mismo conjunto de celdas?

a) $([Dx].[Circ],[Hosp].[H1])$ **ES UNA TUPLA** (VER DEFINICIÓN 1)

b)

$\{([Dx].[Circ],[Hosp].[H1],[Time].[Jan]),$

$([Dx].[Circ],[Hosp].[H1],[Time].[Feb]),$

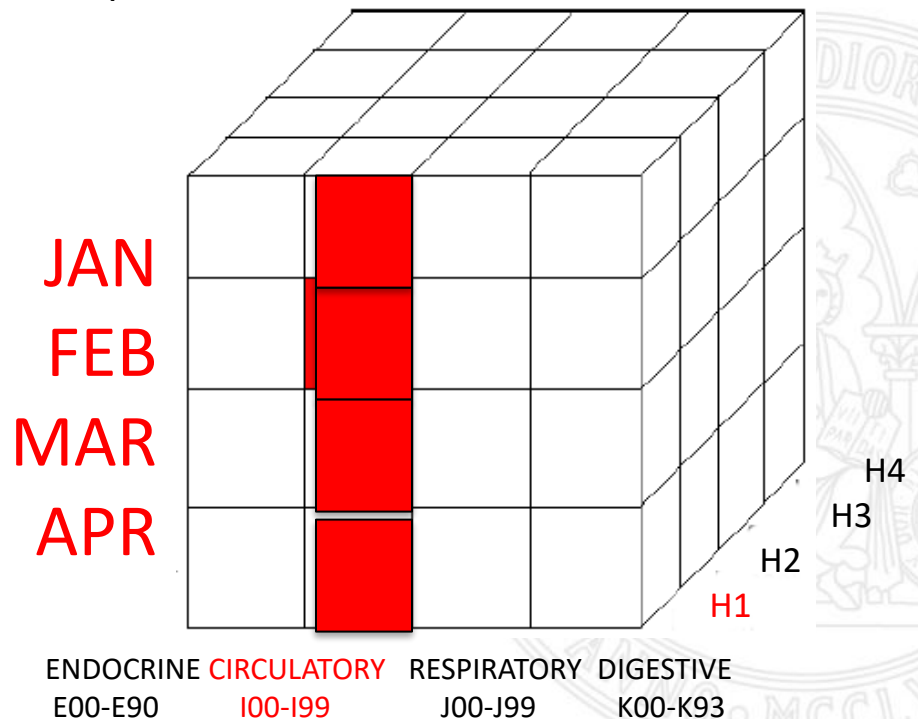
$([Dx].[Circ],[Hosp].[H1],[Time].[Mar]),$

$([Dx].[Circ],[Hosp].[H1],[Time].[Apr])$

} ES UN CONJUNTO

Motivo clave:

La tupla solo usa un miembro por dimensión



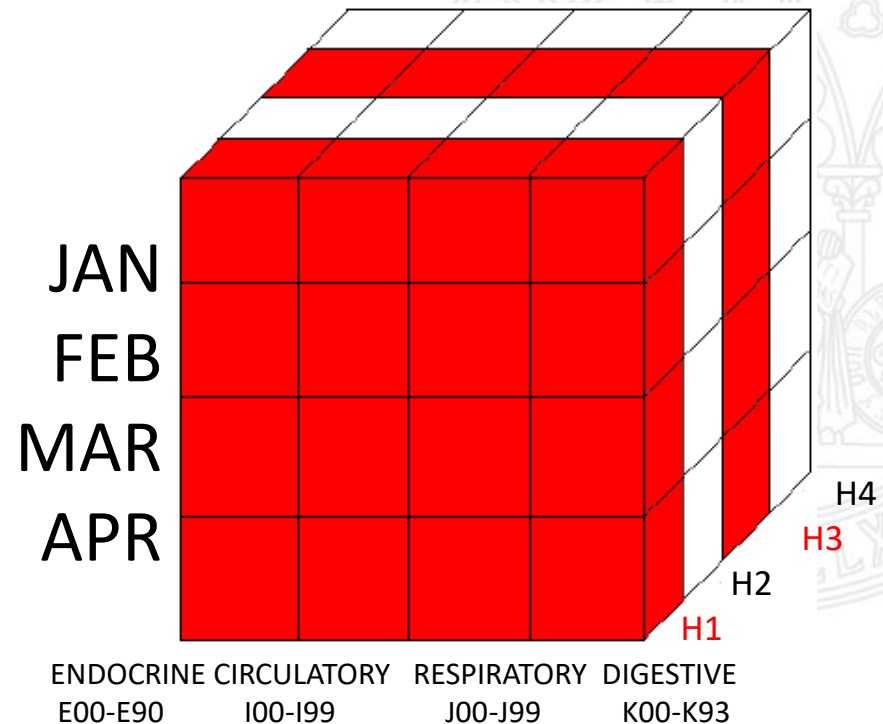
Unidad 3. Explotación de almacenes de datos

Pregunta 3: Tuplas, conjuntos y celdas

UNIVERSIDAD DE
MURCIA

Pregunta: ¿Tupla o Conjunto?

[Hosp].[H1], [Hosp].[H3]



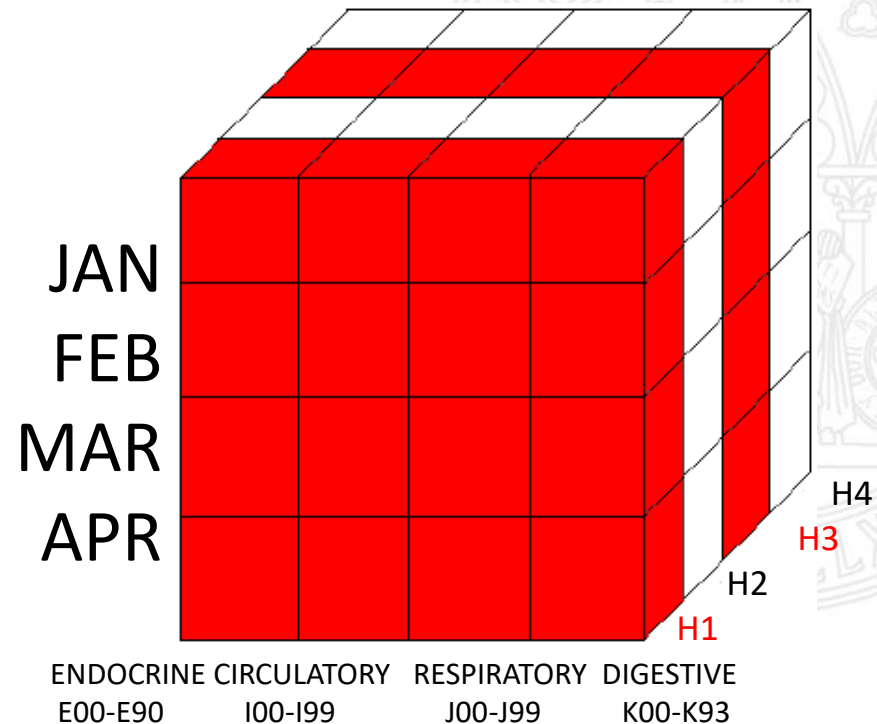
Unidad 3. Explotación de almacenes de datos

Pregunta 3: Tuplas, conjuntos y celdas

Pregunta: ¿Tupla o Conjunto?

[Hosp].[H1], [Hosp].[H3]

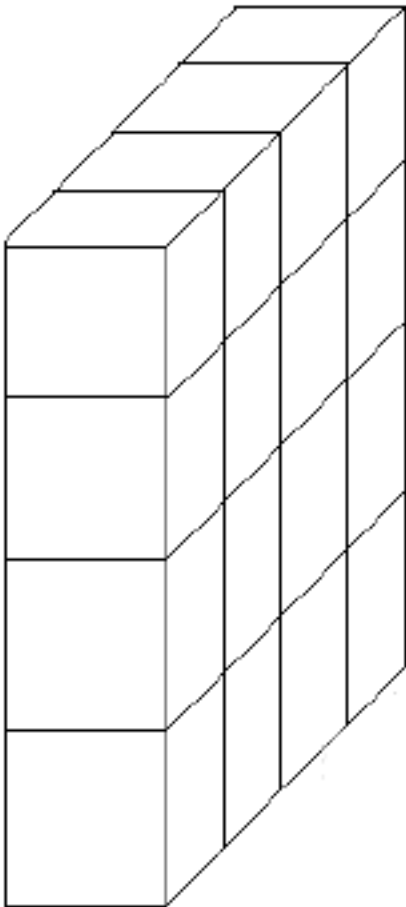
es un **Conjunto**
(ver Def2)



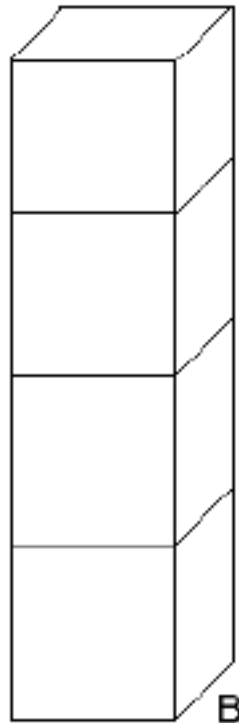
Unidad 3. Explotación de almacenes de datos

Resumen: Tuplas, conjuntos y celdas

[DX].[Circulatory]



[DX].[Circulatory],[Hosp].[H1]



TODO ESTO SON TUPLAS
YA QUE TIENEN LA
"CAPACIDAD PARA SEÑALAR
A UNA ÚNICA CELDA"

Unidad 3. Explotación de almacenes de datos

Resumen: Tuplas, conjuntos y celdas

- Def2: “Una tupla es la intersección de un (y solo un) miembro tomado de una o de varias dimensiones en el cubo. **Una tupla identifica (o tiene el potencial de identificar) una única celda en la matriz multidimensional**”
- Clave: un único miembro de cada dimensión
- **No es necesario definir cada dimensión** – Las dimensiones tienen un **miembro predeterminado** en MDX para asegurar que apunten a una única celda

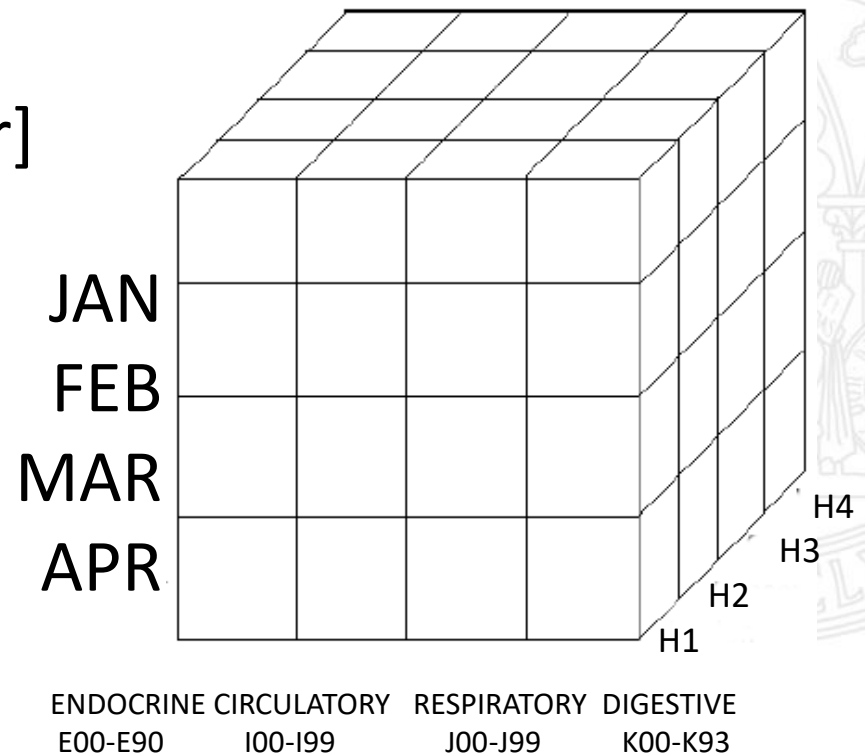


Unidad 3. Explotación de almacenes de datos

Pregunta 4: Tuplas, conjuntos y celdas

Pregunta: ¿Apuntan estas tuplas a una única celda?

[DX].[Circulatory],[Hosp].[H1], [Time].[Mar]
[DX].[Circulatory],[Hosp].[H1]
[DX].[Circulatory],[Time].[Mar]
[Hosp].[H1]



Unidad 3. Explotación de almacenes de datos

Pregunta 4: Tuplas, conjuntos y celdas

Pregunta: ¿Apuntan estas tuplas a una única celda?

[DX].[Circulatory],[Hosp].[H1], [Time].[Mar]

[DX].[Circulatory],[Hosp].[H1]

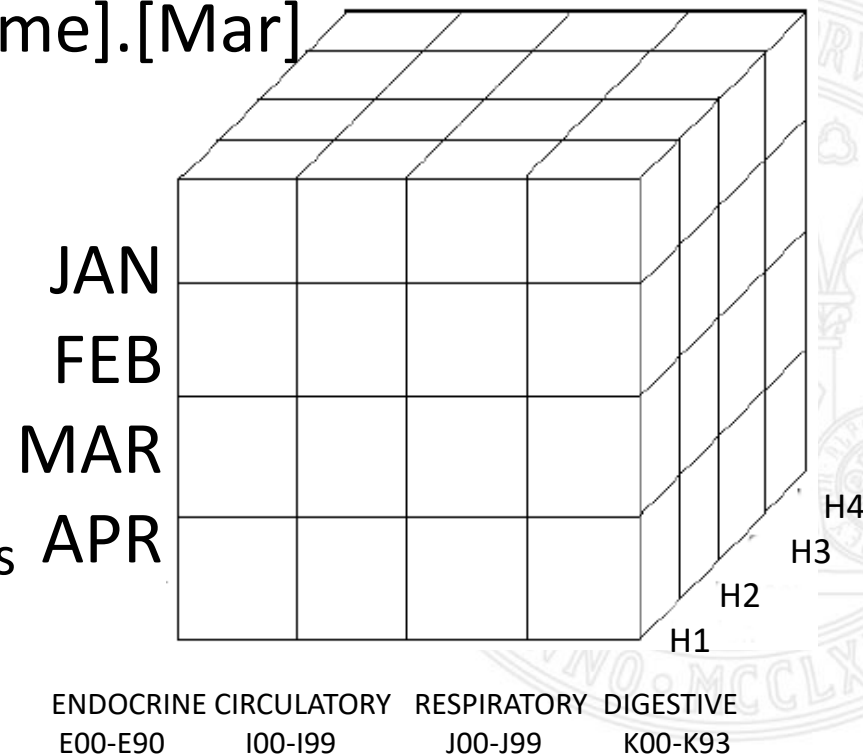
[DX].[Circulatory],[Time].[Mar]

[Hosp].[H1]

SI

Si consideramos que todas las dimensiones tienen un '**miembro predeterminado**'

En MDX si no especificas un miembro de una dimensión, se toma el **miembro por defecto**

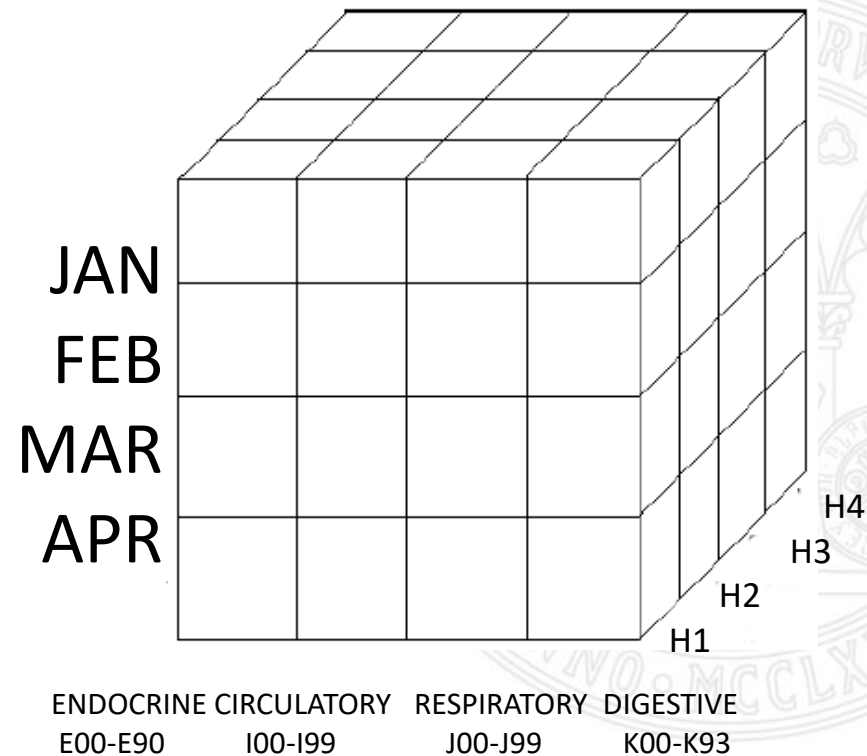


Unidad 3. Explotación de almacenes de datos

Tuplas y jerarquías

Pregunta: ¿Apunta esta tupla a una única celda?

[DX].[Circulatory],[Hosp].[H1], [Time].[2000]



Unidad 3. Explotación de almacenes de datos

Tuplas y jerarquías

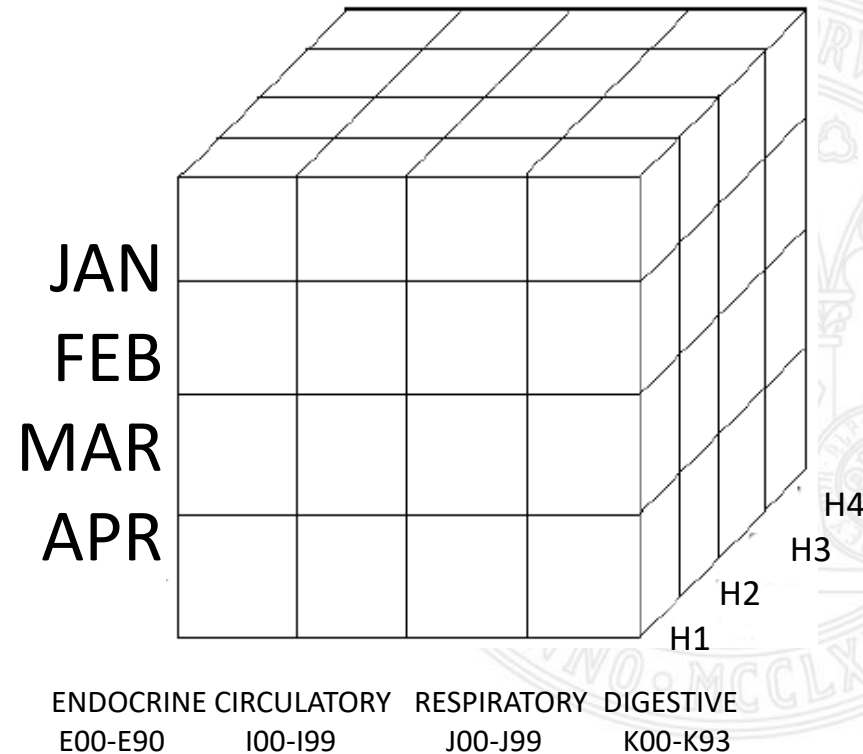
Pregunta: ¿Apunta esta tupla a una única celda?

[DX].[Circulatory],[Hosp].[H1], [Time].[2000]

SI

Tendremos una intersección de agregación para cada miembro de la jerarquía en algún lugar del CUBO

Toda la agregación puede no estar pre-calculada, pero aun así es una tupla (estén almacenadas o se calculen sobre la marcha)



Unidad 3. Explotación de almacenes de datos

Pregunta 5: Tuplas, conjuntos y celdas

UNIVERSIDAD DE
MURCIA

Medidas y Jerarquías

[Hos].[H1],[TIME].[2000].[Q1],[Dx].[Cir]

¿Conjunto?

¿Tupla?

¿Celda?

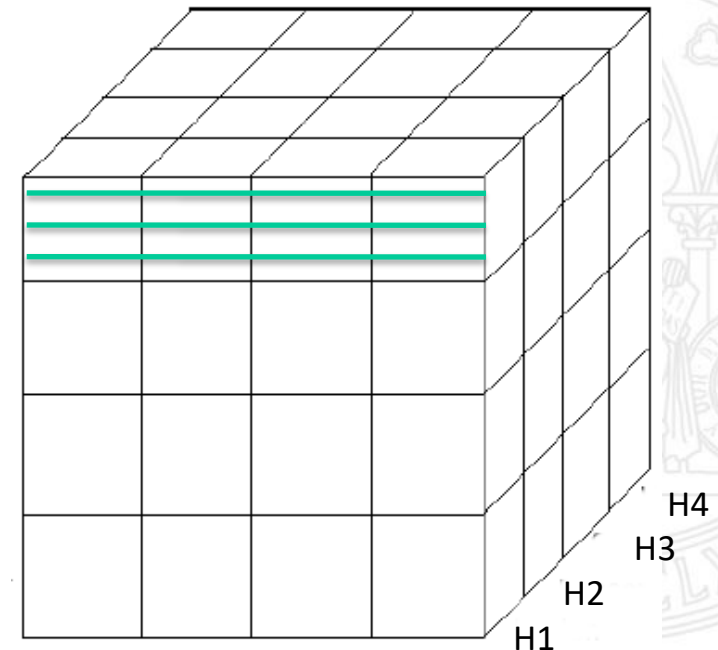
2000

Q1
JAN
FEB
MAR

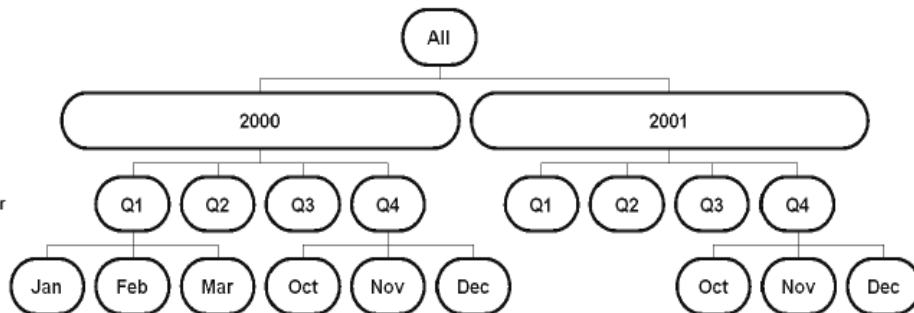
Q2
APR
MAY
JUN

Q3
JULY
...

Q4



ENDOCRINE CIRCULATORY RESPIRATORY DIGESTIVE
E00-E90 I00-I99 J00-J99 K00-K93



Unidad 3. Explotación de almacenes de datos

Pregunta 5: Tuplas, conjuntos y celdas

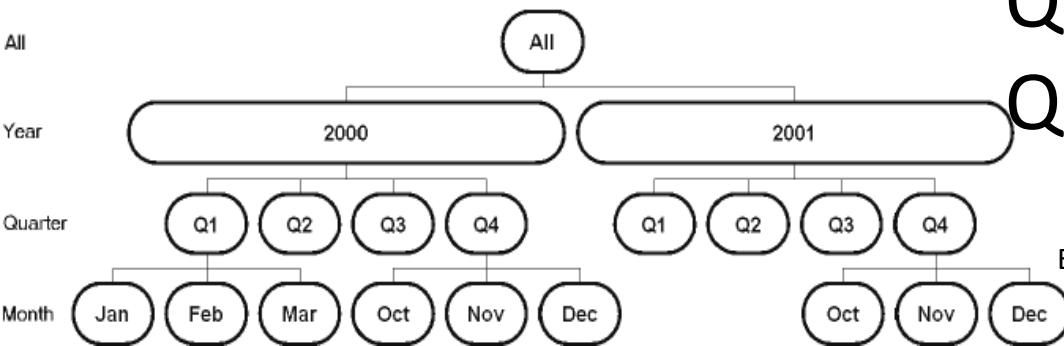
UNIVERSIDAD DE
MURCIA

Medidas y Jerarquías

[Hos].[H1],[**TIME**].[**Q1**],[Dx].[Cir]

¡TUPLA!

¡CELDA! (AGREGACIÓN)



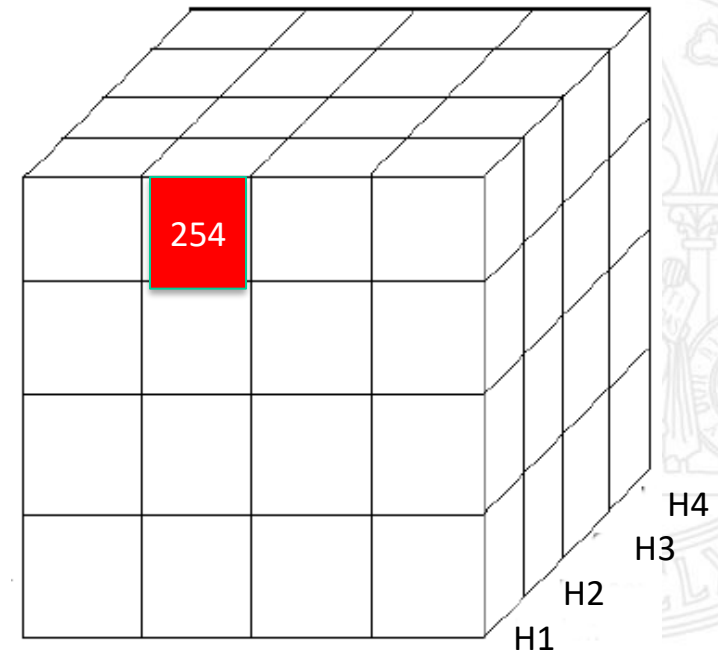
2000

Q1 JAN
FEB
MAR

Q2 APR
MAY
JUN

Q3 JULY
...

Q4



ENDOCRINE CIRCULATORY RESPIRATORY DIGESTIVE
E00-E90 I00-I99 J00-J99 K00-K93

Unidad 3. Explotación de almacenes de datos Y SI...

UNIVERSIDAD DE
MURCIA

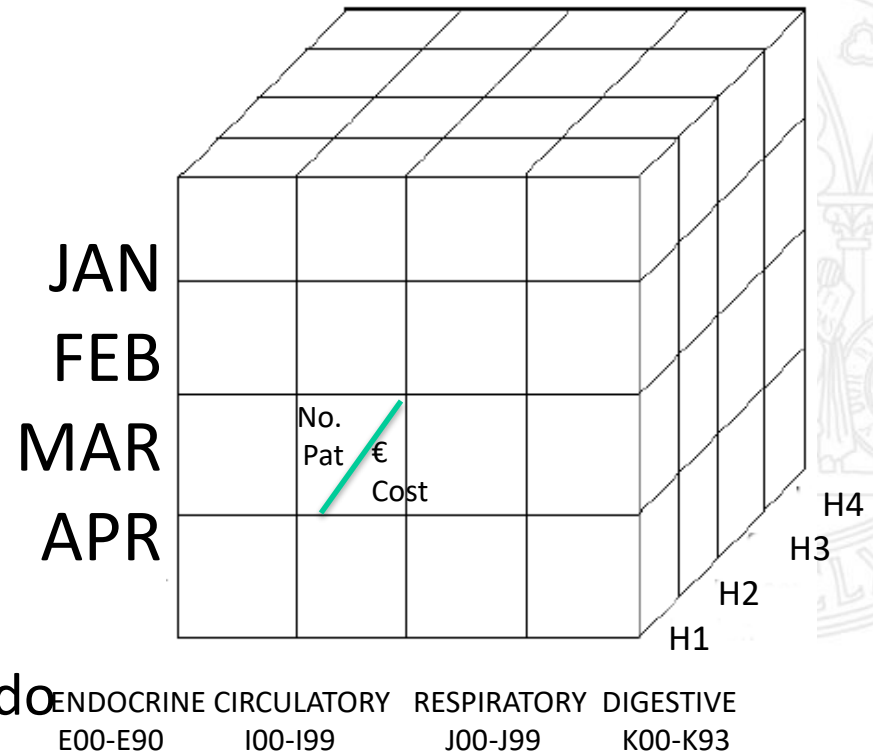
Medidas como dimensiones

Supón un cubo **con 2 medidas**:
Nº pacientes y coste.

[Hos].[H1],[TIME].[Mar],
[Dx].[Car],[**Medidas**].[**NumPac**]

Medida se comporta como miembro
de una dimensión

Si no se especifica, usa el predeterminado



Unidad 3. Explotación de almacenes de datos

Resumen final de definiciones

- Tupla: Intersección de un único miembro de cada dimensión
 - Miembros no añadidos de forma explícita: (actual) miembro implícito, o predeterminado o por defecto
 - Identifica o tiene el potencial para identificar una única celda (por agregación)
- Conjunto: Colección de tuplas con la misma dimensionalidad (puede ser 0 o 1 o varias)
- Medidas: Pueden tratarse como dimensiones. ¿Diferencias?
 - Normalmente numéricas y continuas
 - Propiedades especiales como Data Type o Format String
 - No son jerárquicas
 - Las dimensiones normalmente son discontinuas
- Los miembros pueden tener propiedades para añadir información
 - En nuestro ejemplo: el tamaño de un hospital concreto

Unidad 3. Explotación de almacenes de datos - MDX

TÉRMINOS MDX

SELECT ON_COLUMNS ON_CELLS FROM



Unidad 3. Explotación de almacenes de datos

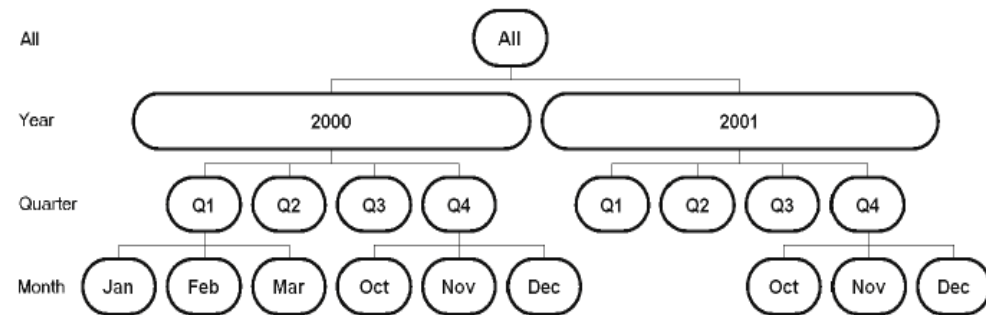
Términos MDX

- MDX puede parecer similar a SQL, pero tiene diferencias clave
 - Ambos son TLAs (Three Letter Acronyms)
 - Lenguajes diseñados específicamente para consultar estructuras de datos
 - Ambos usan cláusulas como SELECT, FROM, y WHERE
 - SQL no es solo un lenguaje de consulta, también para estructuras de datos
 - MDX puede hacer cosas que SQL no, por ejemplo, jerarquías relacionadas, estructura de cubo, ubicación de filas(e.g., día/mes anterior)
- ¡Es divertido aprender estas diferencias!
- Consultas MDX: declaraciones independientes para recuperar datos, muchas personas lo hacen sin conocimiento claro (e.g., tablas dinámicas en Excel)
- Expresiones MDX: Declaraciones parciales utilizadas para diversos fines, por ejemplo, crear un miembro calculado

Unidad 3. Explotación de almacenes de datos

Términos MDX

	ALL (TIME)
COSTE	45,300,000 €



SELECT

{[TIME].[ALL]} **ON COLUMNS**

(COL dimension)

{[Measure].[Cost]} **ON ROWS**

(ROW dimension)

FROM [MyCube]

(muestra el coste de[HOSP].[H1], miembro por defecto de HOSP)

(también para [Dx].[Circulatory])

Unidad 3. Explotación de almacenes de datos

Términos MDX

	COL 1	COL 2	COL 3
ROW A			
ROW B			
ROW C			

SELECT

{column headers} ON COLUMNS → CONJUNTO

{row headers} ON ROWS → CONJUNTO

FROM [cubo] → nombre

1. Un componente para especificar el column header (**conjunto**)
2. Un componente para especificar el row header (**conjunto**)
3. Un puntero al cubo que estamos usando (nombre)

Unidad 3. Explotación de almacenes de datos

Términos MDX

SELECT

{[Medida].[Paciente]} ON COLUMNS

{[Hospital].[Hosp1],

[Hospital]. [Hosp2],

[Hospital]. [Hosp3],

[Hospital]. [Hosp4]} ON ROWS

FROM [MiCubo]

(muestra el miembro por defecto de TIME)

	PACIENTE
HOSP 1	23
HOSP 2	65
HOSP 3	88
HOSP 4	65

Unidad 3. Explotación de almacenes de datos

Términos MDX

SELECT

{[Medida].[Paciente]} ON COLUMNS

{[Hospital].~~All Hosp~~} ON ROWS

FROM [MiCubo]

SELECT

{[Medida].[Paciente]} ON COLUMNS

{[Hospital].Children} ON ROWS

FROM [MiCubo]

	PACIENTE
HOSP 1	23
HOSP 2	65
HOSP 3	88
HOSP 4	65

Unidad 3. Explotación de almacenes de datos

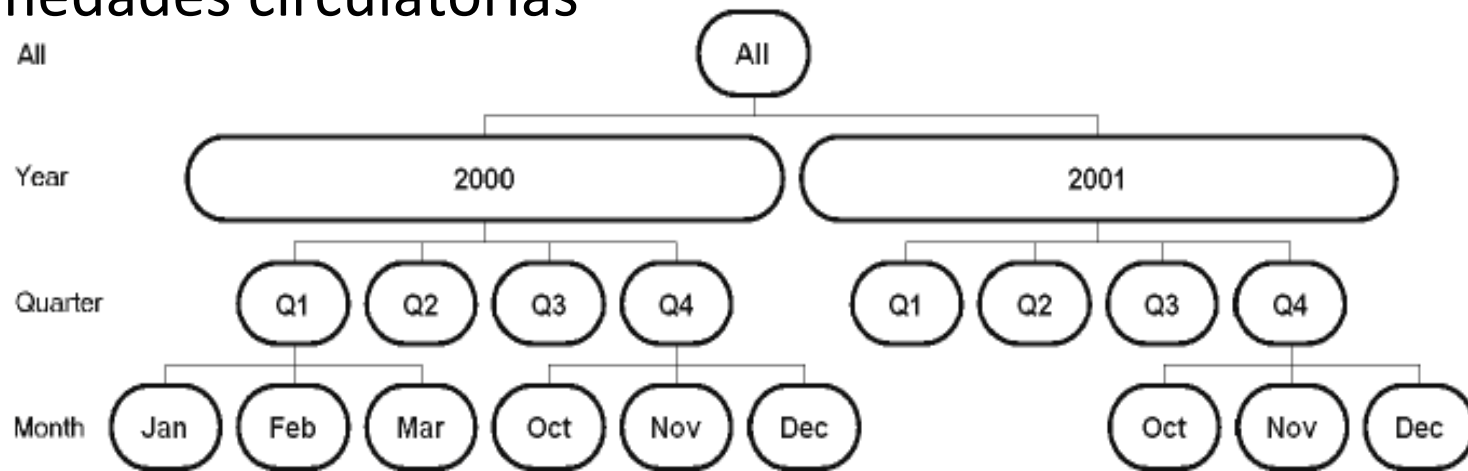
Términos MDX

PREGUNTA:

Coste en H1,H2

durante el año 2000 (por cada trimestre,
Q), de enfermedades circulatorias

Y2000	HOSP1	HOSP2
Q1	2M€	0.3M€
Q2	3.2M€	0.7M€
Q3	1.5M€	0.6M€
Q4	0.4M€	0.5M€



Pista:

coste/circulatorias son miembros por defecto

Unidad 3. Explotación de almacenes de datos

Términos MDX

PREGUNTA:

Coste en H1,H2

durante el año 2000 (por cada trimestre,
Q), de enfermedades circulatorias:

SELECT

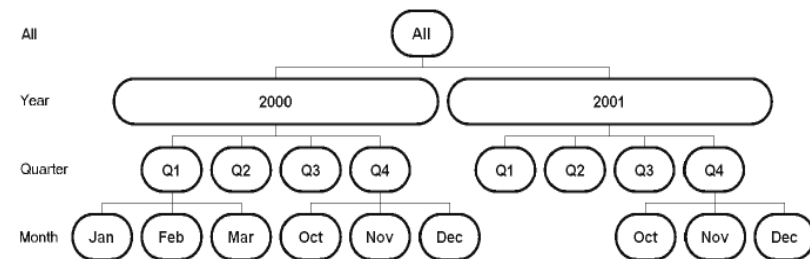
{[Hospital].[Hosp1],

[Hospital].[Hosp2]} ON COLUMNS

{[Time].[All].[2000].Children} ON ROW

FROM [MiCubo]

Y2000	HOSP1	HOSP2
Q1	2M€	0.3M€
Q2	3.2M€	0.7M€
Q3	1.5M€	0.6M€
Q4	0.4M€	0.5M€



Unidad 3. Explotación de almacenes de datos

Términos MDX

PREGUNTA:

Nº de pacientes en H1,H2

durante el año 2000 (por trimestre), para
enfermedades circulatorias

Y2000	HOSP1	HOSP2
Q1	121 pac	78 pac
Q2	165 pac	61 pac
Q3	115 pac	41 pac
Q4	120 pac	76 pac

Pista:

El número de pacientes NO es un miembro por defecto

Unidad 3. Explotación de almacenes de datos

Términos MDX

PREGUNTA:

Nº de pacientes en H1,H2

durante el año 2000 (por trimestre), para
enfermedades circulatorias.

SELECT

{[Hospital].[Hosp1],

[Hospital].[Hosp2]} ON COLUMNS

{[Time].[All].[2000].Children} ON ROWS

FROM [MiCubo]

WHERE ([Medidas].[NumPac])

Y2000	HOSP1	HOSP2
Q1	121 pac	78 pac
Q2	165 pac	61 pac
Q3	115 pac	41 pac
Q4	120 pac	76 pac

Unidad 3. Explotación de almacenes de datos

Términos MDX

Cláusula WHERE

No está restringida a medidas.

No está restringida a 1 dimensión.

Es un *SLICER/DICER* (*cortador/seccionador*)



Unidad 3. Explotación de almacenes de datos

Términos MDX

Cláusula WHERE

No está restringida a medidas.

SELECT

{[Hospital].[Hosp1],

[Hospital].[Hosp2]} ON COLUMNS

{[Time].[All].[2000].Children} ON ROWS

FROM [MiCubo]

WHERE ([Dx].[Respiratory])

Y2000	HOSP1	HOSP2
Q1	1M€	0.4M€
Q2	1.2M€	0.1M€
Q3	0.5M€	0.5M€
Q4	0.4M€	0.3M€

Unidad 3. Explotación de almacenes de datos

Términos MDX

Cláusula WHERE

No está restringida a 1 dimensión

SELECT

{[Hospital].[Hosp1],

[Hospital].[Hosp2]} ON COLUMNS

{[Time].[All].[2000].Children} ON ROWS

FROM [MiCubo]

WHERE ([Dx].[Respiratory],[Medidas].[NumPac])

Y2000	HOSP1	HOSP2
Q1	61 pac	28 pac
Q2	75 pac	41 pac
Q3	105 pac	11 pac
Q4	112 pac	56 pac

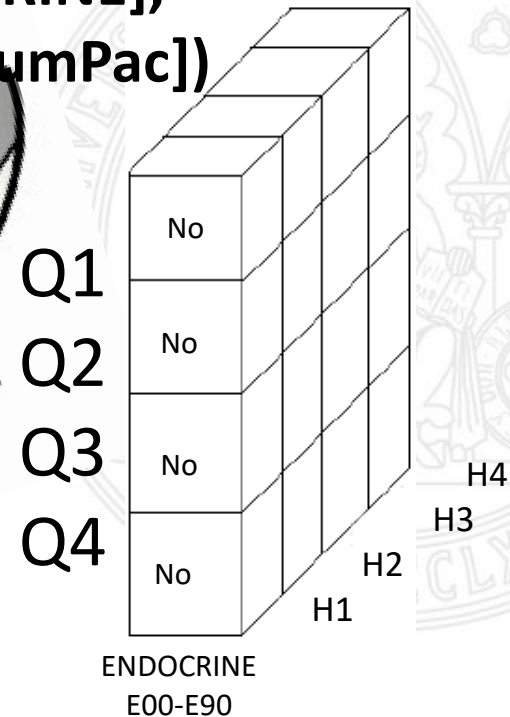
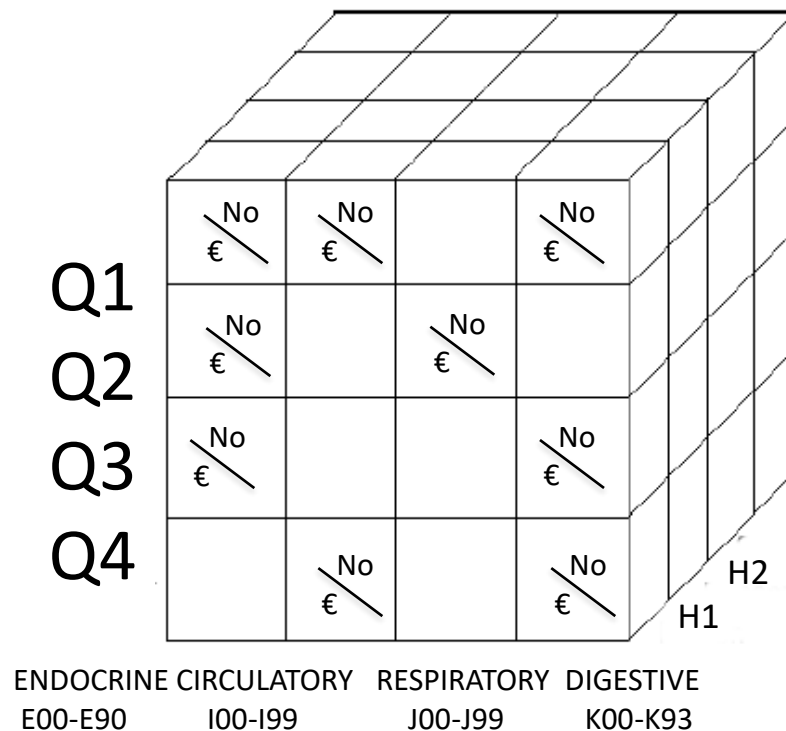
Unidad 3. Explotación de almacenes de datos

Términos MDX

Cláusula WHERE

Es un *SLICER/DICER* (cortador/seccionador)

WHERE
([Dx].[ENDOCRINE],
[Medidas].[NumPac])



Unidad 3. Explotación de almacenes de datos - MDX

SINTAXIS MDX

() [] {}



Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

Corchetes []

Dimensiones: [Time] y Miembros: [2000]

Solo es sintácticamente obligatorio de usar si hay números, espacios...(aunque es mejor usarlos siempre para evitar confusiones)

Puntos .

Separadores: [Time].[2000].[Q3]

Paréntesis ()

Tuplas: ([DX].[Circulatory],[Hosp].[H1])

Intersección de uno o más miembros, con una coma para separar a los distintos miembros

Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

Llaves {}

Conjuntos: {[Hosp].[H1], [Hosp].[H3]}

{[Dx].Children}

{ ([Dx].[Circ],[Hosp].[H1],[Time].[Jan]),

([Dx].[Circ],[Hosp].[H1],[Time].[Feb]),

([Dx].[Circ],[Hosp].[H1],[Time].[Mar]),

([Dx].[Circ],[Hosp].[H1],[Time].[Apr]) }

Una colección de tuplas con la misma dimensionalidad (pueden ser 0 o 1 o varias)

Hint: Si una consulta no funciona con una tupla, prueba añadiendo { }

Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

SELECT

{ CONJUNTO } ON COLUMNS

{ CONJUNTO } ON ROWS

FROM [cubo]

WHERE (TUPLA)



Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

Pregunta: ¿Es correcto? ¿Por qué?

```
SELECT  
([Medidas].[NumPacientes]) ON COLUMNS,  
{[Time].[2000].Children} ON ROWS  
FROM [MiCubo]
```



Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

Pregunta: ¿Es correcto? ¿Por qué?

```
SELECT  
{[Dx].Children} ON COLUMNS,  
{[Time].[2000].[Q1].[Mayo].Children} ON ROWS  
FROM [MiCubo]  
WHERE {[Medidas].[coste],[Hosp].[H2]}
```



Unidad 3. Explotación de almacenes de datos

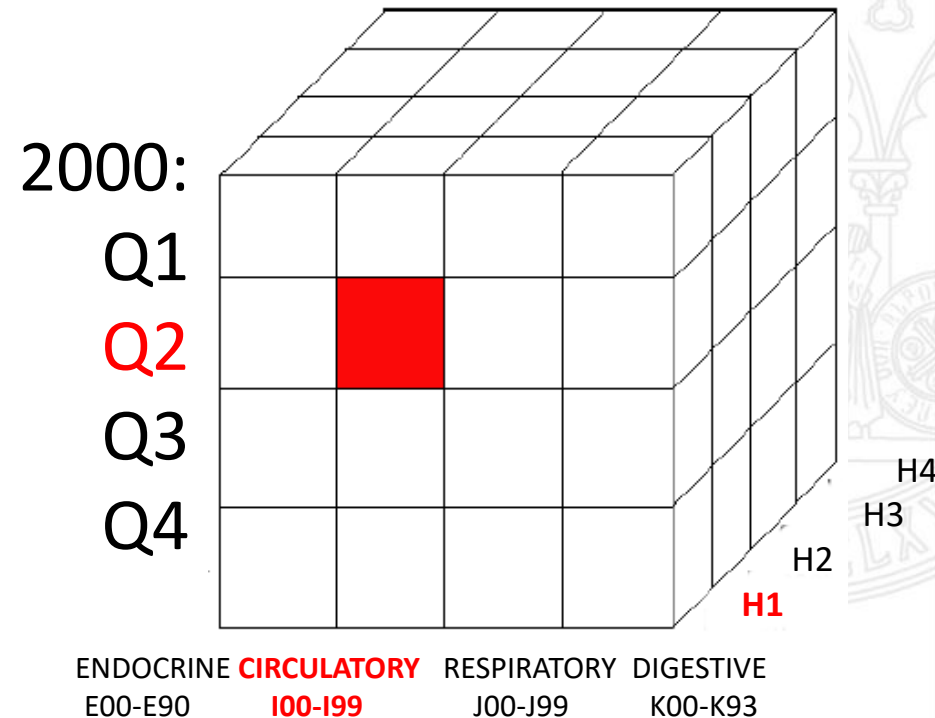
Sintaxis MDX

Nombre de una CELDA

En un cubo, cada celda tiene un nombre – una tupla.

El nombre de esta celda es:

([Time].[2000].[Q2],
[Dx].[Circulatory],
[Hospital].[H1])



Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

Nombre de una CELDA

El nombre de esta celda es:

([Time].[2000].[Q2].[Jun],

[Dx].[Circulatory],

[Hospital].[H1])

La celda actual y vecinas:

CurrentMember

2000:

Jan

Feb

Mar

Apr

May

Jun

July

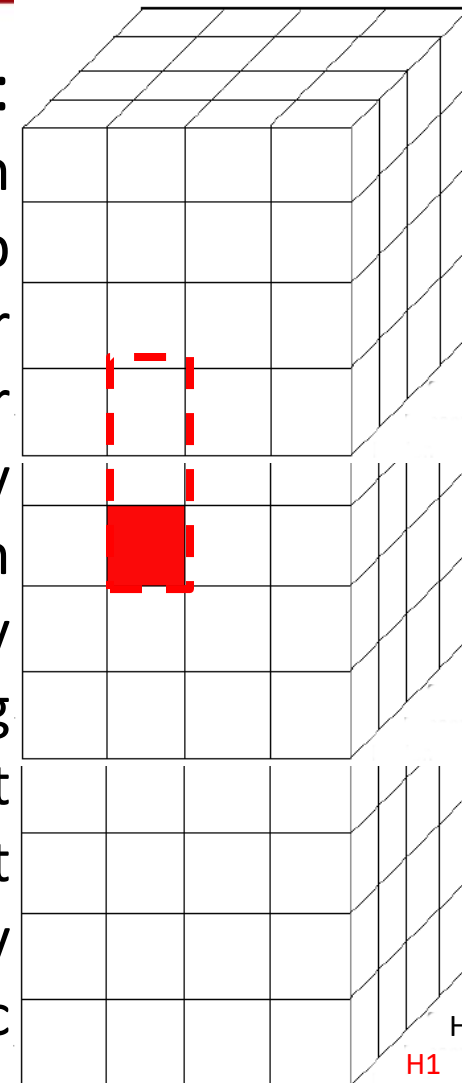
Aug

Sept

Oct

Nov

Dec



Unidad 3. Explotación de almacenes de datos

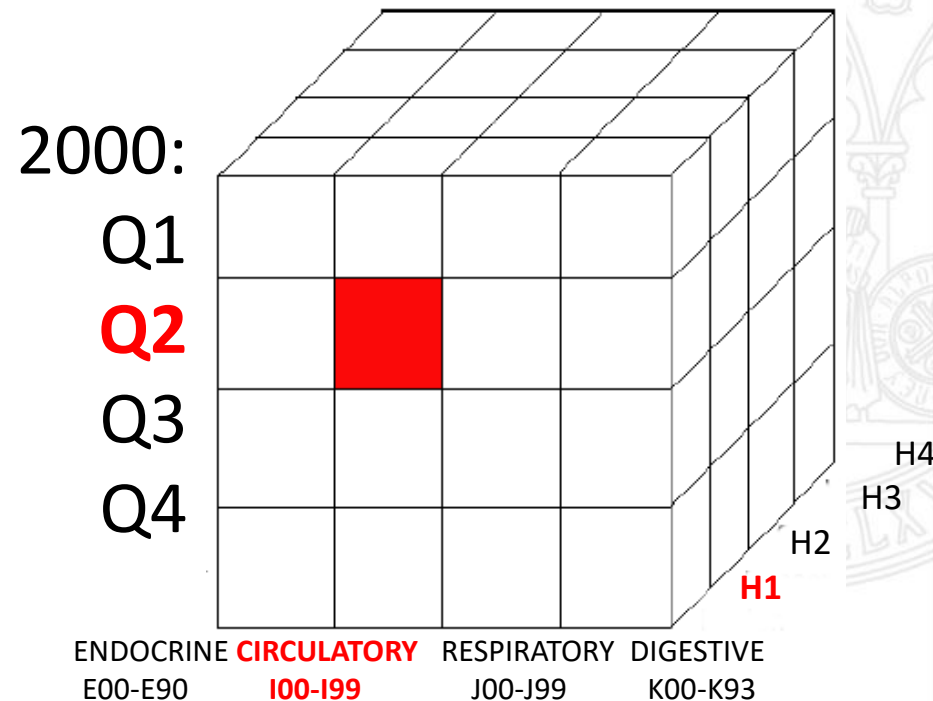
Sintaxis MDX

Referencia de celda relativa:

CurrentMember, PrevMember, NextMember.

El nombre de esta celda es:

([Time].[2000].[Q3].PrevMember,
[Dx].[Circulatory],
[Hospital].[H1])



Unidad 3. Explotación de almacenes de datos

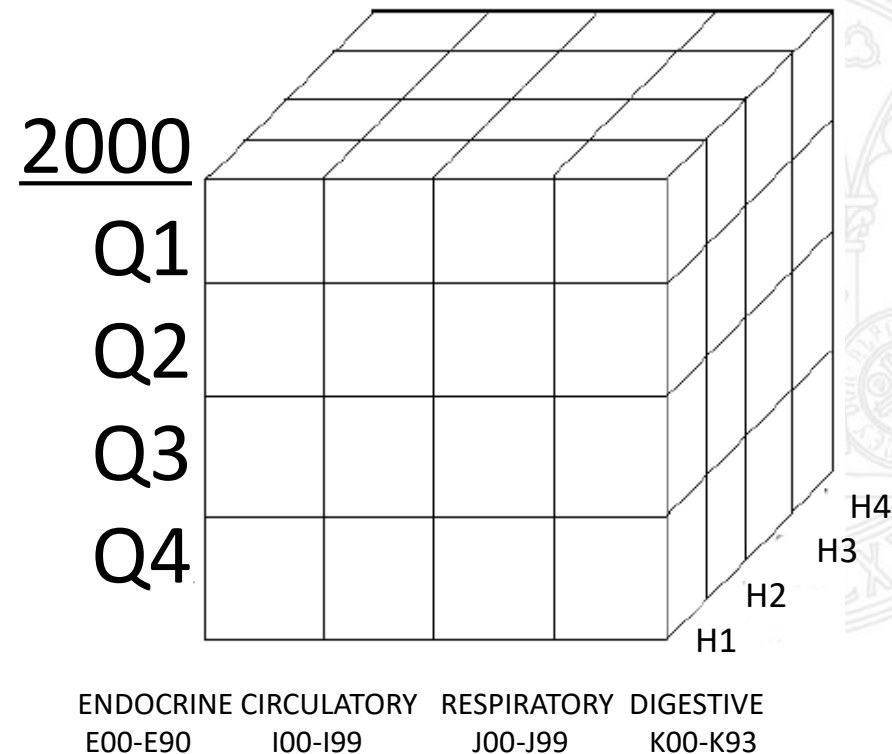
Sintaxis MDX

Referencia de celda relativa:

Las funciones `Lag()` y `Lead()` permiten generalizar el movimiento hacia atrás y hacia delante (¡sí, también se pueden usar negativos!)

El nombre de esta celda es:

`([Time].[2000].[Q3].Lag(2),
[Dx].[Circulatory].Lead(2),
[Hospital].[H1])`



Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

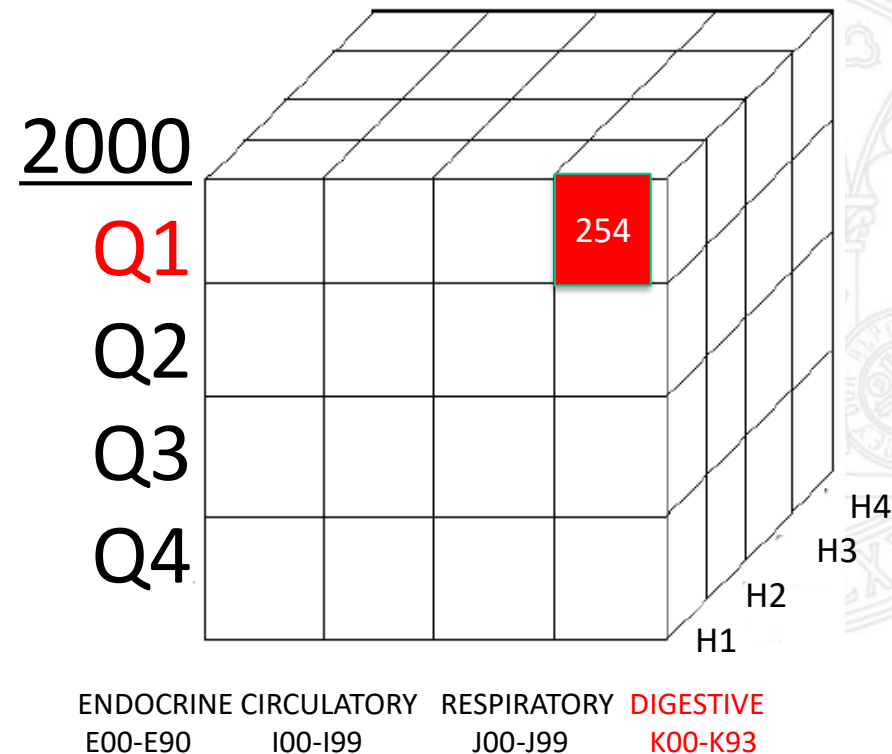
Referencia de celda relativa:

Las funciones `Lag()` y `Lead()` permiten generalizar el movimiento hacia delante y hacia atrás (¡sí, también se pueden usar negativos!)

El nombre de esta celda es:

`([Time].[2000].[Q3].Lag(2),
[Dx].[Circulatory].Lead(2),
[Hospital].[H1])`

¿Interesante desde la perspectiva
de negocio?



Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

Miembros calculados: +-*/ %

Calculados en base a una operación, no existían antes, ejemplo:

“Mejora de la atención a pacientes de enfermedades circulatorias en el primer trimestre de los años 1999 y 2000”.

Cálculo:

`([Hosp].[H1],[Dx].[Circ],[Time].[2000].[Q1],[Medidas].[NumPacientes])`

-

`([Hosp].[H1],[Dx].[Circ],[Time].[1999].[Q1],[Medidas].[NumPacientes])`

¿Se puede generalizar para que funcione con cualquier miembro de la dimensión Tiempo (Time)?

Unidad 3. Explotación de almacenes de datos

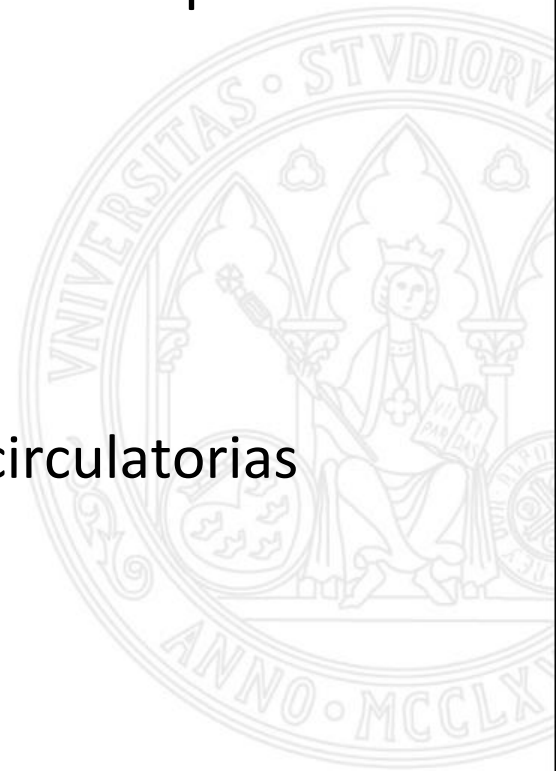
Sintaxis MDX

Miembros calculados: +-*/ %

“**Crecimiento del coste a lo largo del año 2000 en H1 para pacientes de enfermedades circulatorias**”.

Crecimiento del coste: $\text{coste}(t) - \text{coste}(t-1)$
(incremento/derivada)

Obviar: H1 para pacientes de enfermedades circulatorias



Unidad 3. Explotación de Sintaxis MDX

“Crecimiento del coste a lo largo del año 2000”.

Year	Quarter	Month	Coste M€	Nº Pacientes
2000			57	280
	Q1		15	90
		January	5	20
		February	5	30
		Mach	5	40
	Q2		10	60
		April	5	15
		Jun	3	15
		July	2	30
	Q3		12	50
		April	5	10
		Jun	5	10
		July	2	30
	Q4		20	80
		April	5	15
		Jun	5	15
		July	10	50

Unidad 3. Explotación de Sintaxis MDX

“Crecimiento del coste a lo largo del año 2000”.

Year	Quarter	Month	Coste M€	Nº Pacientes
2000			57	280
	Q1		15	90
		January	5	20
		February	5	30
		March	5	40
	Q2		10	60
		April	5	15
		June	3	15
		July	2	30
	Q3		12	30
		April	5	10
		June	5	10
		July	2	30
	Q4		20	80
		April	5	15
		June	5	15
		July	10	50

0

-1

-2

3

Unidad 3. Explotación de Sintaxis MDX

Year	Quarter	Month	Coste M€	Nº Pacientes
2000			57	280
	Q1		15	90
		January	5	20
		February	5	30
		March	5	40
	Q2		10	60
		April	5	15
		Jun	3	15
	Q4		20	80
		April	5	15
		Jun	5	15
		July	10	50

“Crecimiento del coste
a lo largo del año 2000”.

[Time].[CurrentMember],[Medidas].[Coste]

-

[Time].[PrevMember],[Medidas].[Coste]

Unidad 3. Explotación de almacenes de datos

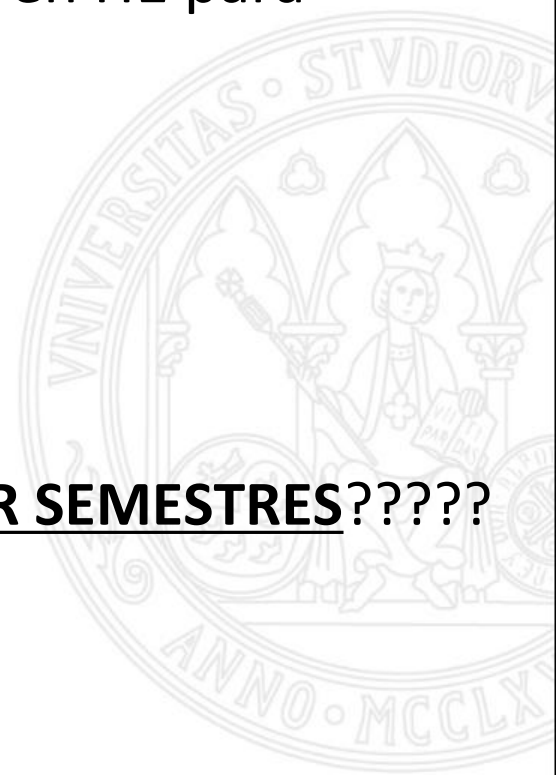
Sintaxis MDX

Miembros calculados: +-*/ %

“Crecimiento del coste a lo largo del año 2000 en H1 para pacientes de enfermedades circulatorias”.

([Time].[**CurrentMember**],[Medidas].[Coste]
- [Time].[**PrevMember**],[Medidas].[Coste])

¿Y SI NOS CENTRAMOS EN EL CRECIMIENTO POR SEMESTRES?????



Unidad 3. Explotación de Sintaxis MDX

Year	Quarter	Month	Coste M€	Nº Pacientes
2000			57	280
	Q1		15	90
		January	5	20
		February	5	30
		March	5	40
	Q2		10	60
		April	5	15
		June	3	15
		July	2	30
	Q3		12	50
		April	5	10
		June	5	10
		July	2	30
	Q4		20	80
		April	5	15
		June	5	15
		July	10	50

Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

Miembros calculados: +-*/ %

“Crecimiento del coste a lo largo del año 2000 en H1 para pacientes de enfermedades circulatorias”.

([Time].[**CurrentMember**],[Medidas].[Coste]
- [Time].[**PrevMember**],[Medidas].[Coste])

¿Y SI NOS CENTRAMOS EN EL CRECIMIENTO ANUAL?!

... MISMA EXPRESIÓN. ESO ES LO INTERESANTE. DEPENDE DE LA DIMENSIÓN TIEMPO, QUE DEFINE LA PROPIEDAD CURRENT MEMBER

Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

Otras funciones:

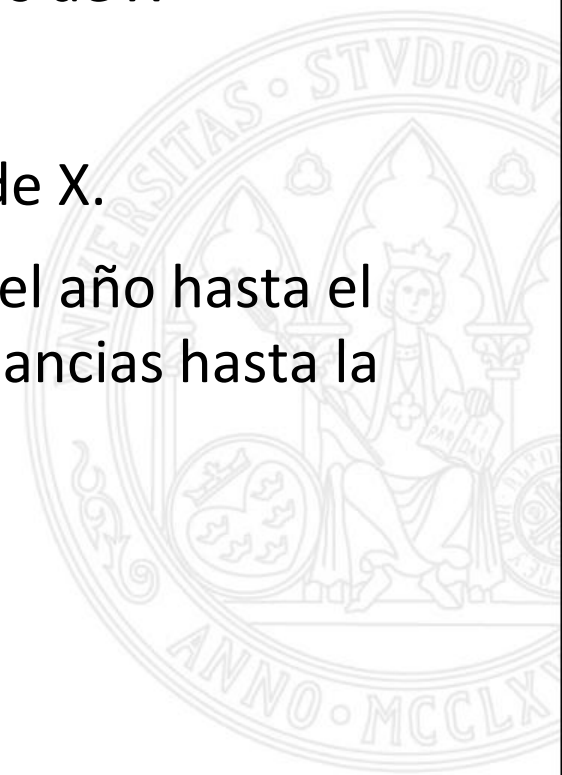
Sum (X) → Número : Suma todos los miembros de X

X.Lag(N) : N posiciones hacia atrás desde X

X.Lead(M) : M posiciones hacia adelante desde X.

YTD(X) → Conjunto: YearToDate: Miembros del año hasta el miembro X. (por ejemplo, para calcular ganancias hasta la fecha actual)

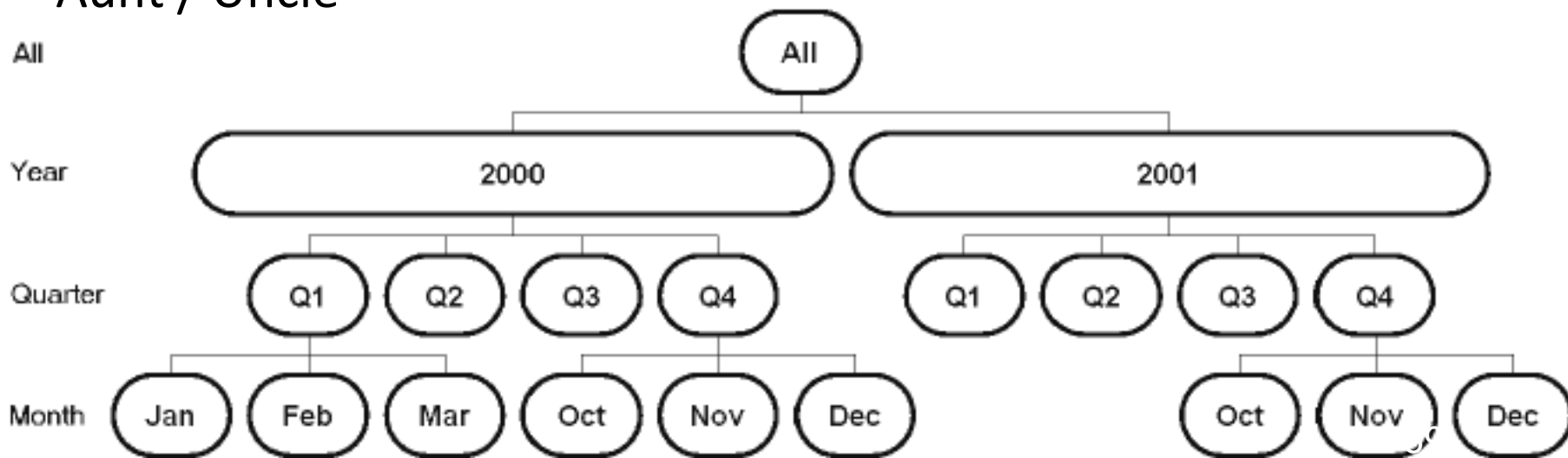
e.g. YTD(March) → {Jan, Feb, March}



Unidad 3. Explotación de almacenes de datos

Sintaxis MDX

- Current member permite análisis basados en la posición
- La navegación jerárquica permite el análisis basado en referencias jerárquicas relativas
- *Member.Children*, *Member.Parent*, *Member.FirstChild* / *LastChild*,
 - *Descendants(X,n)*, *Ancestors(X,n)* (no *Member.Grandchild...*)
 - *Siblings* / *Cousins*: Hermanos/Primos (se esperan niveles simétricos)
- ~~Aunt / Uncle~~



Unidad 3. Explotación de almacenes de datos MDX

Bibliografía y recursos:

- Mark Whitehorn et al. Fast Track to MDX (2nd Ed). Springer. 2004.
- Microsoft, “Key Concepts in MDX (Analysis Services)”, <https://docs.microsoft.com/en-us/analysis-services/multidimensional-models/mdx/key-concepts-in-mdx-analysis-services?view=asallproducts-allversions>
- InterSystems, “Introduction to MDX Queries”, https://docs.intersystems.com/irislatest/csp/docbook/DocBook.UI.Page.cls?KEY=D2GMDX_CH_MDX_INTRO