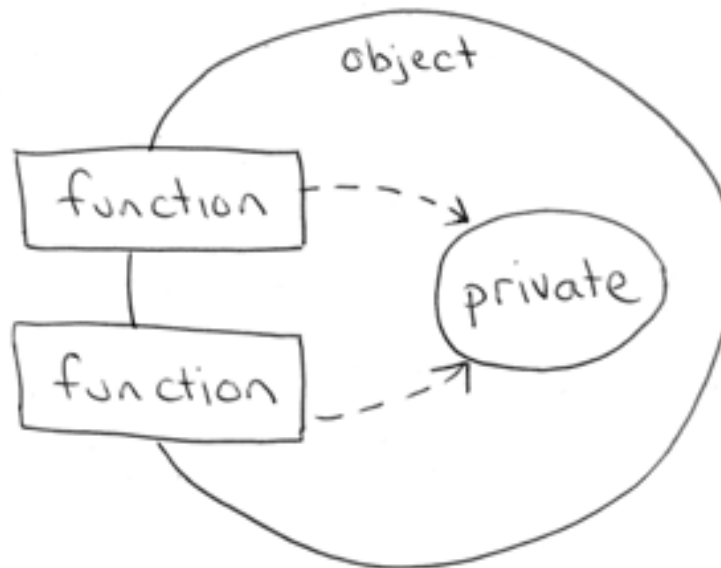# REPORTING

Unit 3 – Data exploitation. Query languages and visualization
S3 –2 – MDX

- **CUBES**

- **LANGUAGE**

- **M**ulti**D**imensional e**X**pression.

- Microsoft in 1997

- Do you remember studying Object-Oriented Programming?

1. Tuples, Sets & Cells
2. MDX Spells
3. MDX Query Syntax

# 1. Tuples, Sets & Cells

- Measures | Dimensions>Members
- E.g.: 2 dimension cube
  - **1 measure**: discharged patients.
  - Time Dimension with 4 members: Jan to April.
  - Hospital Dimension with 4 members: 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

- Measures | Dimensions>Members
- E.g.: 2 dimension cube
  - **2 measure**: no. discharged patients, total cost (M€).
  - Time Dimension with 4 members: Jan to April.
  - Hospital Dimension with 4 members: H1,H2,H3, H4.

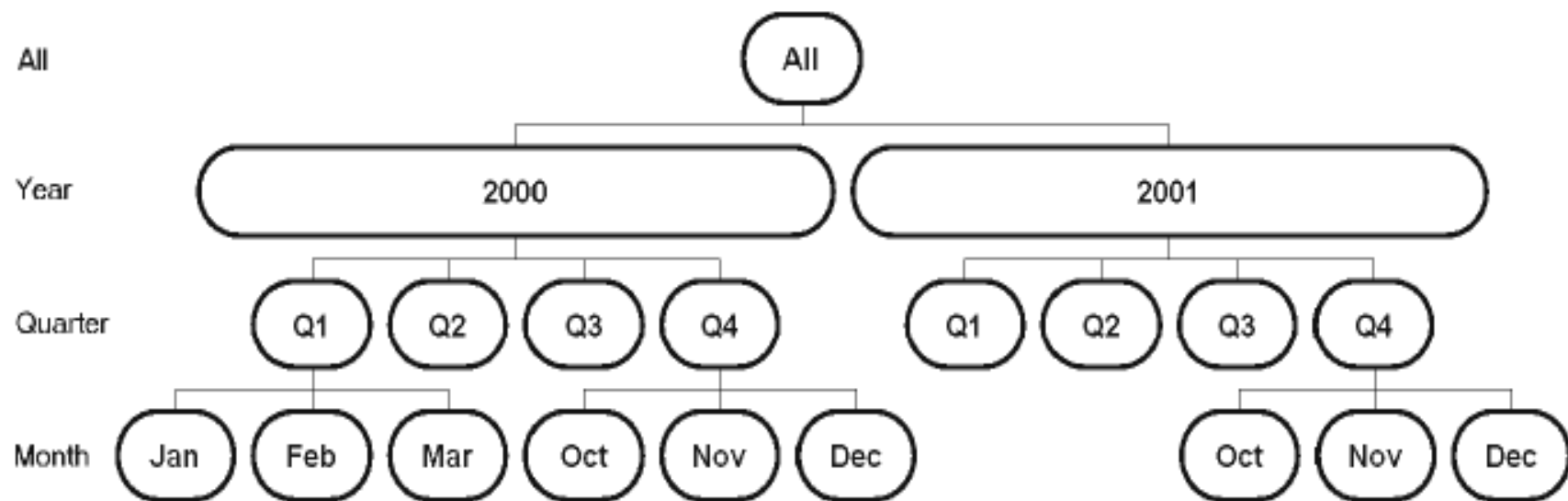| Discharged	H1	H2	H3	H4
January	20 1.5M€	44 4.1M€	81 10.5M€	44 4.1M€
February	15 1.1M€	32 3.9M€	78 10.4M€	32 3.9M€
March	23 1.6M€	65 5.4M€	88 10.7M€	65 5.4M€
April	19 1.5M€	67 5.6M€	67 9.5M€	67 5.6M€
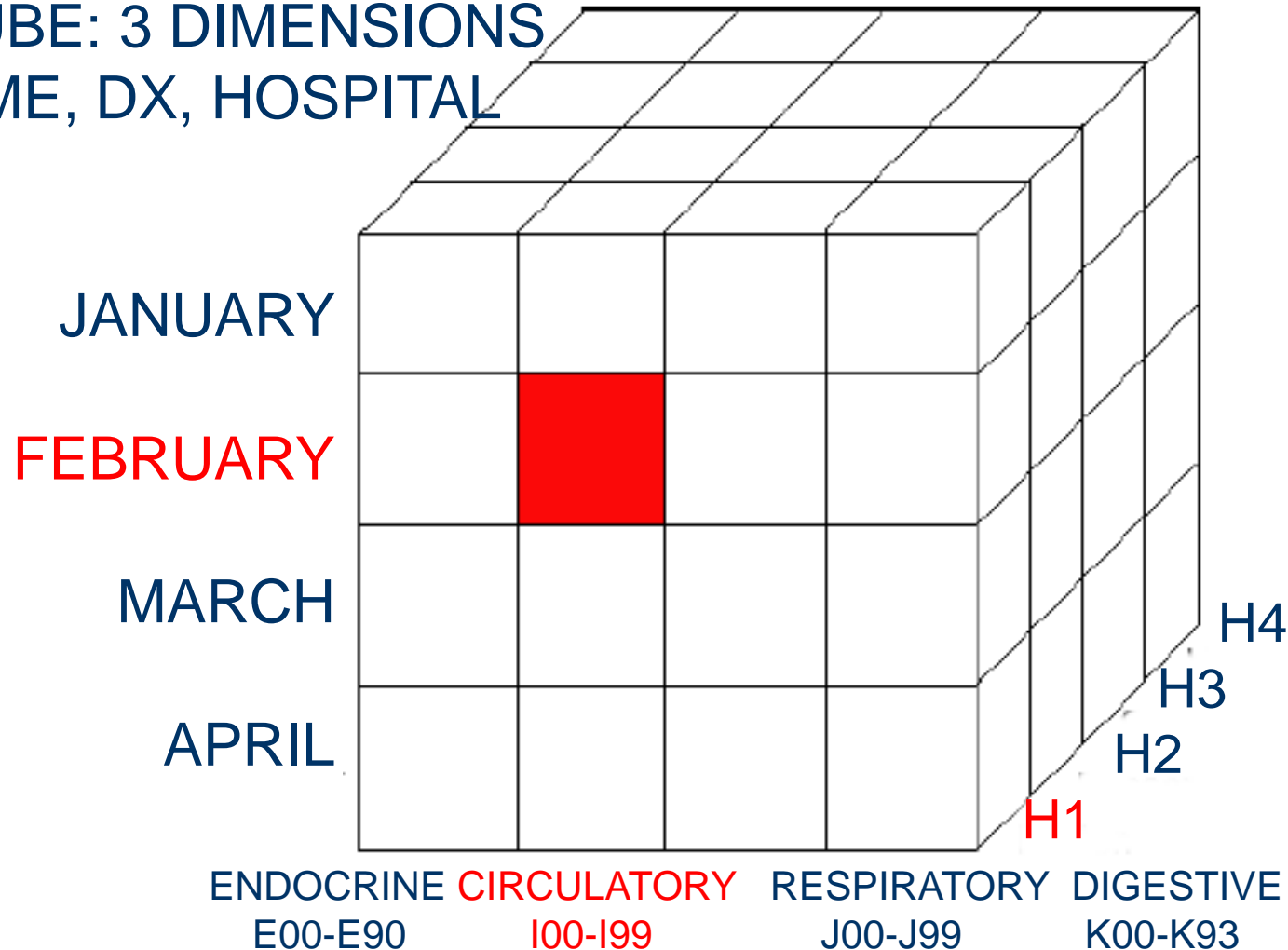
- Dimension has **hierarchies**
- Hierarchy has **levels**: All, Year, Quarter, Month
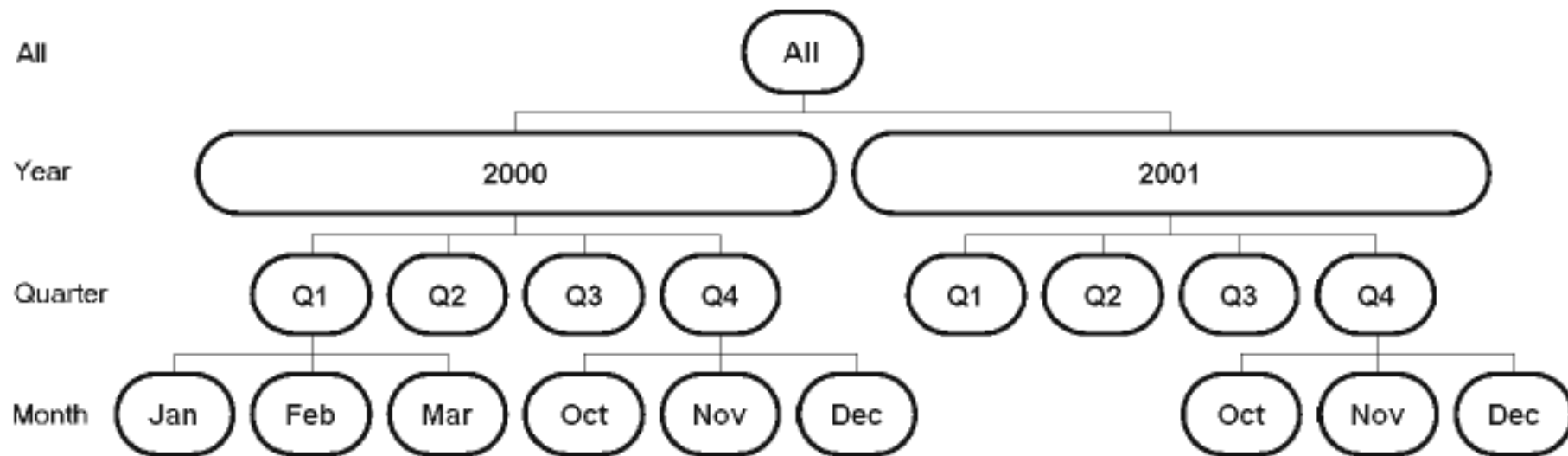
- Naming Conventions
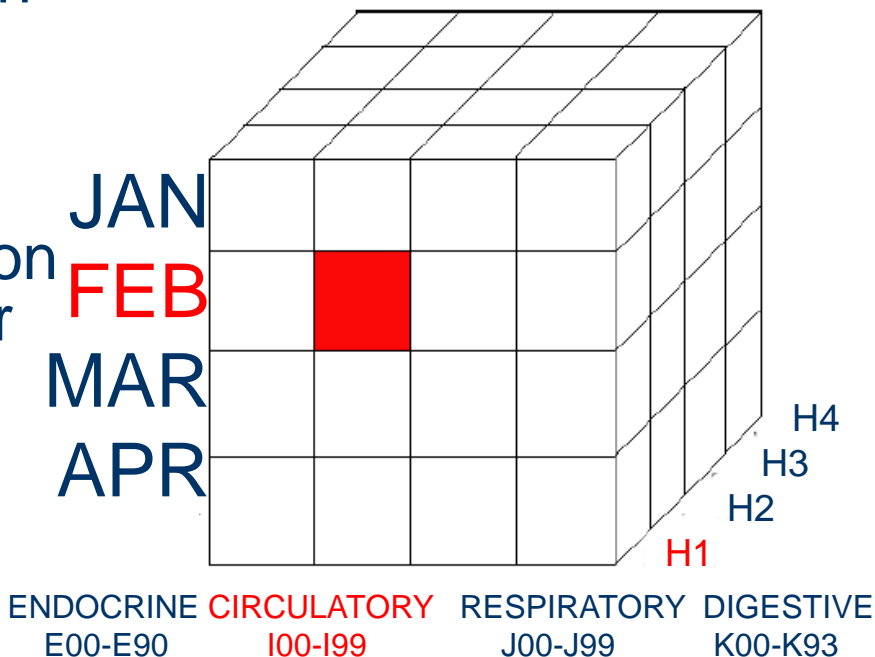
[Time].[All].[2000].[Q4].[Oct] = [Time].[Oct-2000]

- Naming Conventions: **Tuple**

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

Def1: "***Tuple*** is the intersection choosing **one member of each dimension**"

Def2: "A tuple is the intersection of one (<u>and only one</u>) member taken from one or several of the dimensions in the cube."

(tuple=single cell in the cube ??)

JAN
FEB
MAR
APR

H4
H3
H2
H1

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

- Naming Conventions: **Tuple**

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

Def1: "***Tuple*** is the intersection
choosing **one member
of each dimension**"

Def2: "A tuple is the intersection
of one (<u>and only one</u>) member
taken from one or several of
the dimensions in the cube."
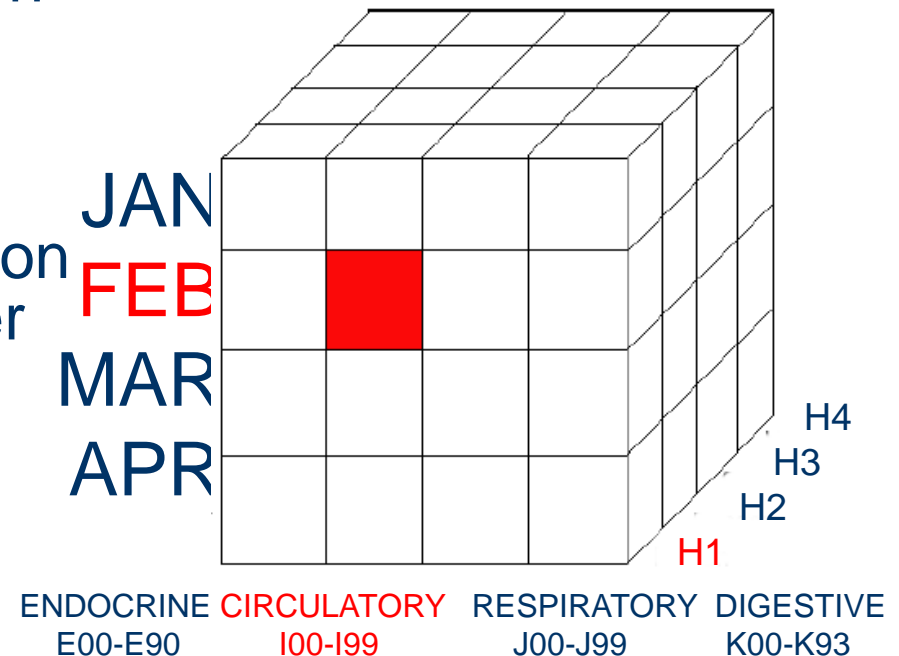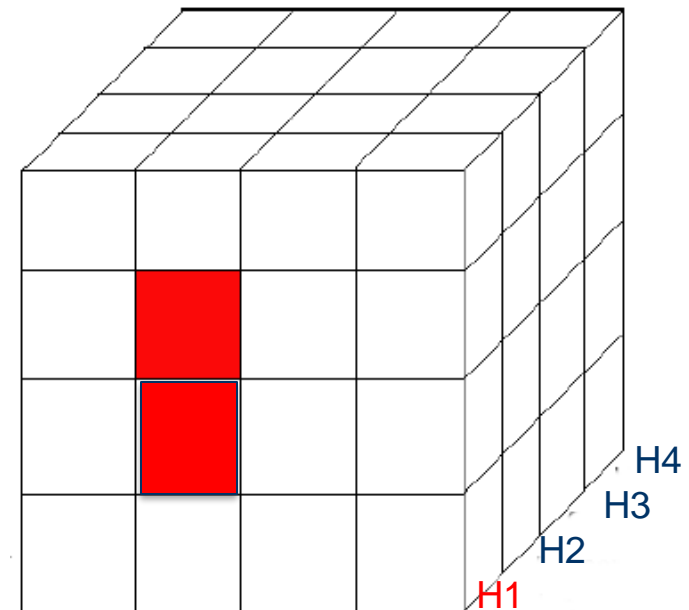
(tuple=single cell in the cube ??)

JAN
FEB
MAR
APR

H4
H3
H2
H1

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

- Naming Conventions: **Set**

Set in pseudo-MDX: {(x1,y1,z1),..,(xn,yn,zn)}
{([Time].[Feb],[Dx].[Circ],[Hosp].[H1]),
([Time].[Mar],[Dx].[Circ],[Hosp].[H1])}

"***Set*** is a set of tuples with the same dimensionality"
(set of cells in the cube)
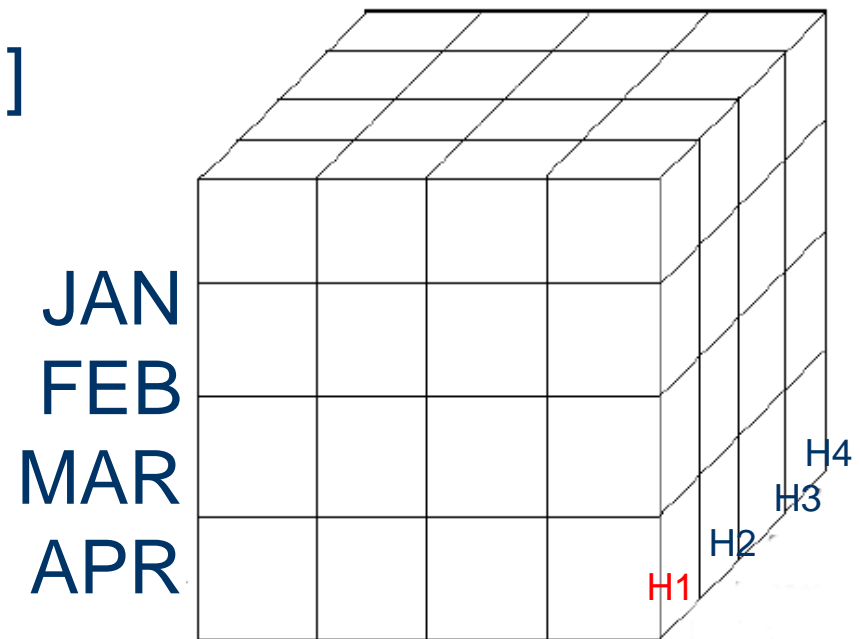
AVG(SET)→FLOAT

JAN
FEB
MAR
APR

H4
H3
H2
H1

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

13

- Question: Tuple or Set?

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

JAN
FEB
MAR
APR

H4
H3
H2
H1

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

14

- Question: Tuple or Set?

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

Is a TUPLE!
(but MANY CELLS!)

JAN			
FEB			
MAR			H4
APR			H3
			H2
			H1
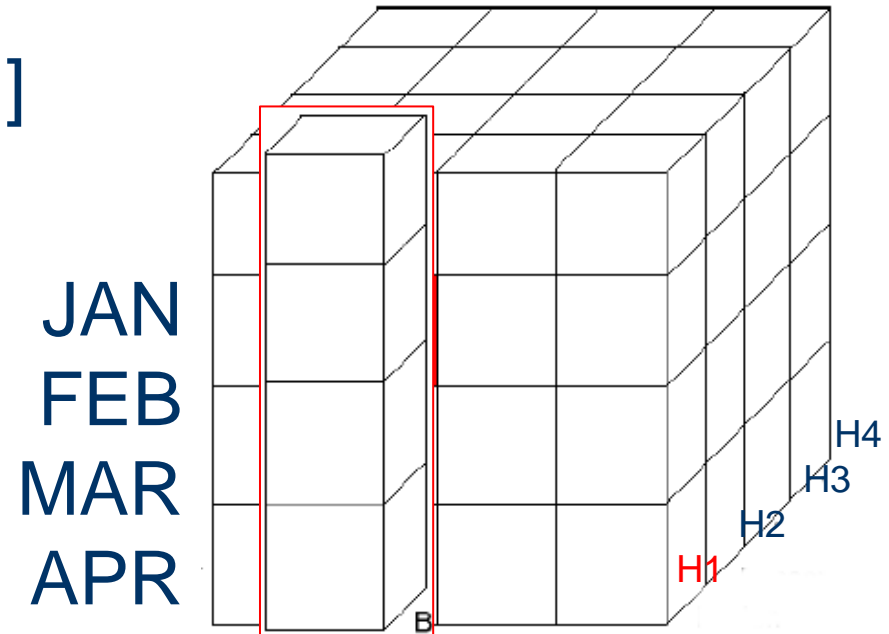
ENDOCRINE	CIRCULATORY	RESPIRATORY	DIGESTIVE
E00-E90	I00-I99	J00-J99	K00-K93

- Question: Difference between?
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])
}

JAN
FEB
MAR
APR

H4
H3
H2
H1

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

16
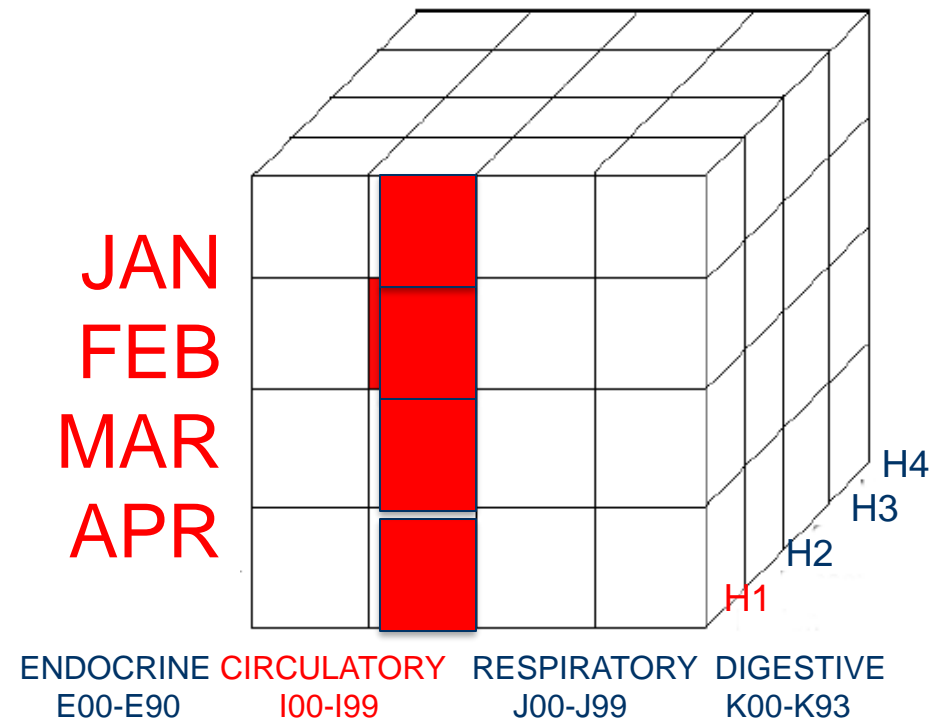
- Question: Difference between?

a)   ([Dx].[Circ],[Hosp].[H1]) **IS A TUPLE** (SEE DEFINITION 1)

b)
{([Dx].[Circ],[Hosp].[H1],[Time].[Jan]),
([Dx].[Circ],[Hosp].[H1],[Time].[Feb]),
([Dx].[Circ],[Hosp].[H1],[Time].[Mar]),
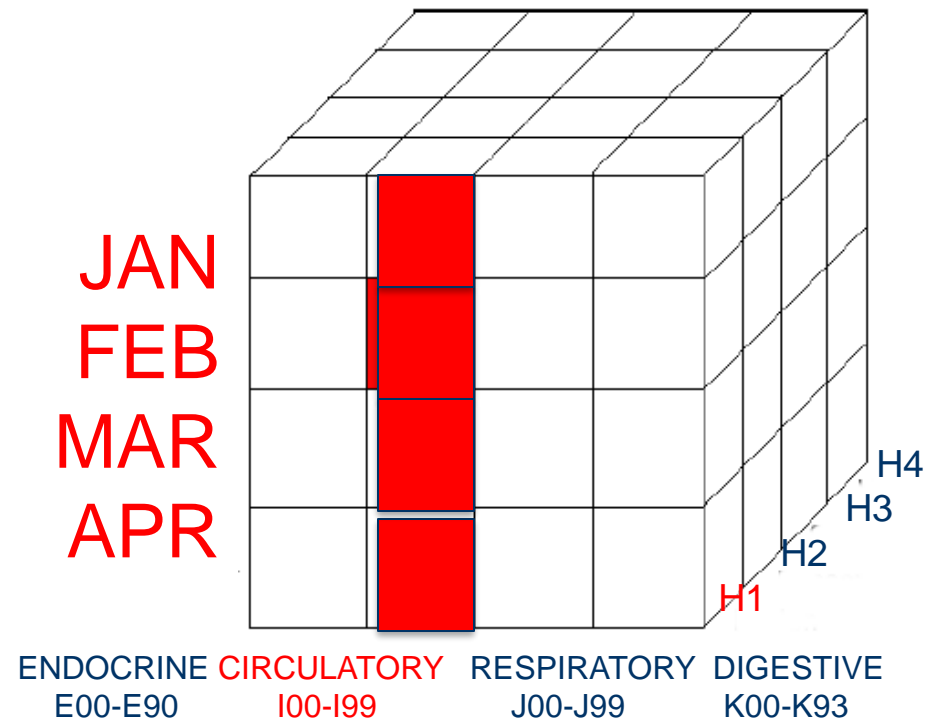([Dx].[Circ],[Hosp].[H1],[Time].[Apr])
}   **IS A SET**

JAN
FEB
MAR
APR

H4
H3
H2
H1

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

17

- Question: Tuple or Set?
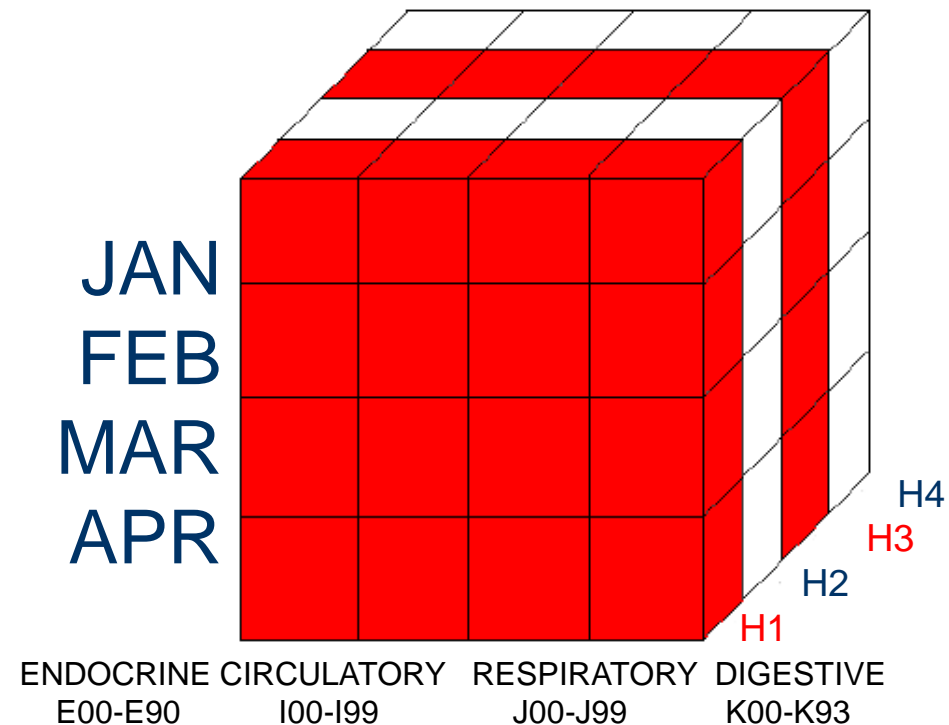
[Hosp].[H1], [Hosp].[H3]

- Question: Tuple or Set?

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

is a Set
(see Def2!)

JAN
FEB
MAR
APR

H4
H3
H2
H1

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

[DX].[Circulatory]   [DX].[Circulatory],[Hosp].[H1]
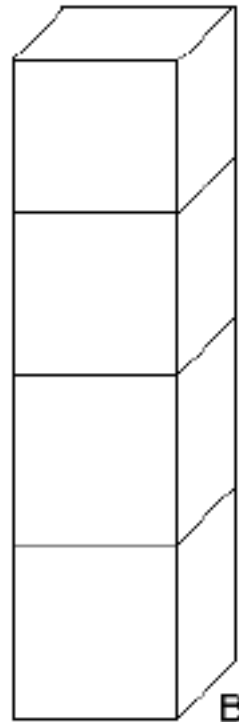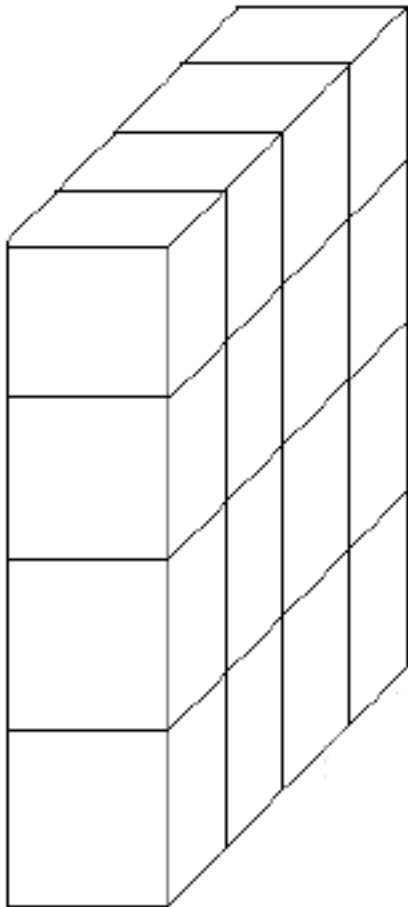
ALL THESE ARE TUPLES SINCE THEY HAVE THE "CAPACITY TO POINT TO A SINGLE CELL" (actually they don't)

- Question: Do these tuples point to a single cell?
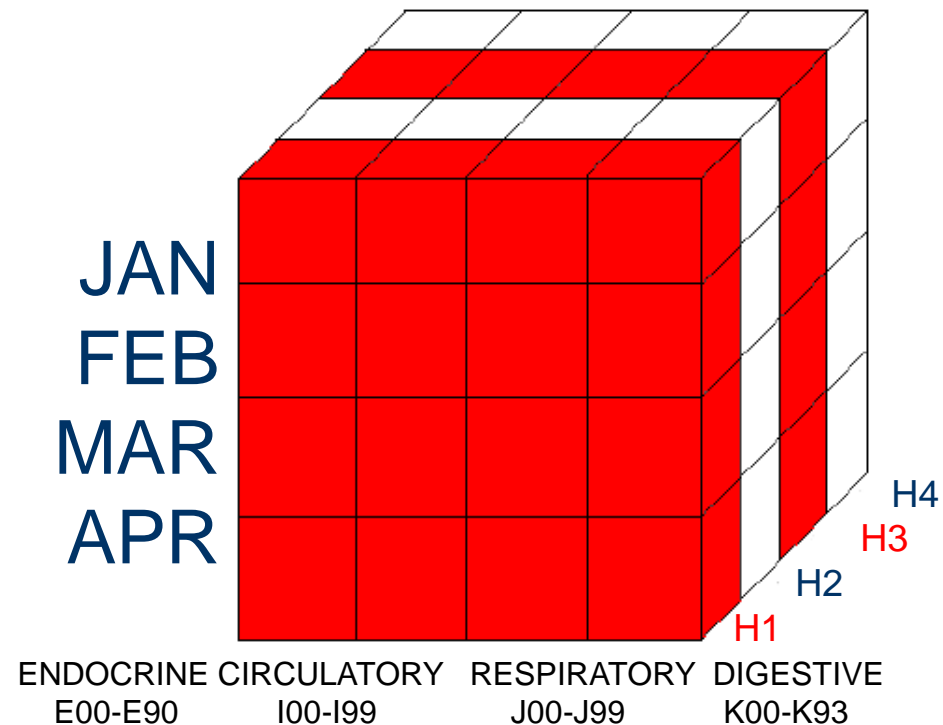
[DX].[Circulatory],[Hosp].[H1], [Time].[Mar]
[DX].[Circulatory],[Hosp].[H1]
[DX].[Circulatory],[Time].[Mar]
[Hosp].[H1]

JAN
FEB
MAR
APR

H4
H3
H2
H1

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

21

- Question: Do these tuples point to a single cell?

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

YES

If we consider that all dimensions have a '*DEFAULT MEMBER*'
In MDX if you don't specify a member of a dimension the default member is implied

JAN
FEB
MAR
APR

H4
H3
H2
H1
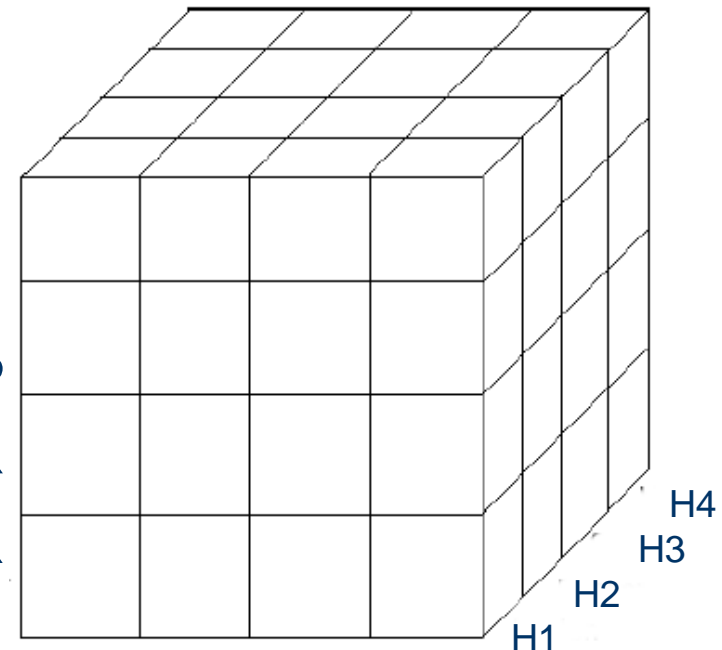
ENDOCRINE CIRCULATORY RESPIRATORY DIGESTIVE
E00-E90   I00-I99   J00-J99   K00-K93

- Measures like dimensions

Suppose a cube **with 2 measures**:
No.Patients and €Cost.

[Hos].[H1],[TIME].[Mar],
[Dx].[Car],**[Measures].[NoPat]**

Measure behaves like member of a dimension

JAN
FEB
MAR
APR

No. Pat € Cost

H4
H3
H2
H1

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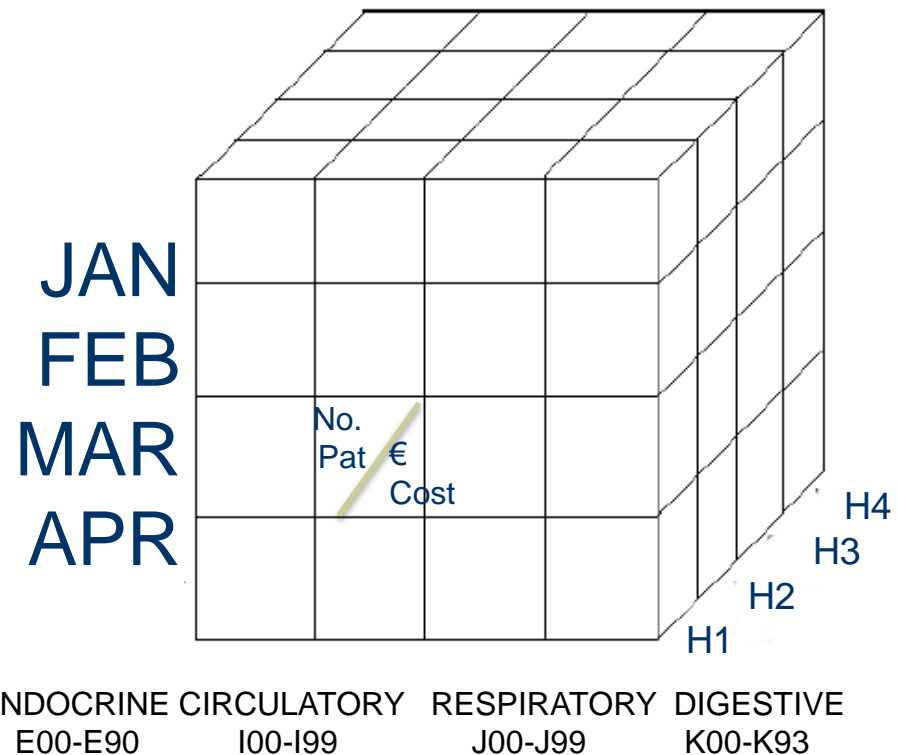
- Measures & Hierarchies

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

SET?
TUPLE?
CELL?

2000

Q1 JAN
   FEB
   MAR

Q2 APR
   MAY
   JUN

Q3 JULY
   ...

Q4

H4
H3
H2
H1

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

All

2000   2001

Q1 Q2 Q3 Q4   Q1 Q2 Q3 Q4

Jan Feb Mar Oct Nov Dec   Oct Nov Dec

- Measures & Hierarchies

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

TUPLE!
CELL! (AGGREGATION)

2000

Q1 JAN
FEB
MAR
Q2 APR
MAY
JUN
Q3 JULY
...
Q4
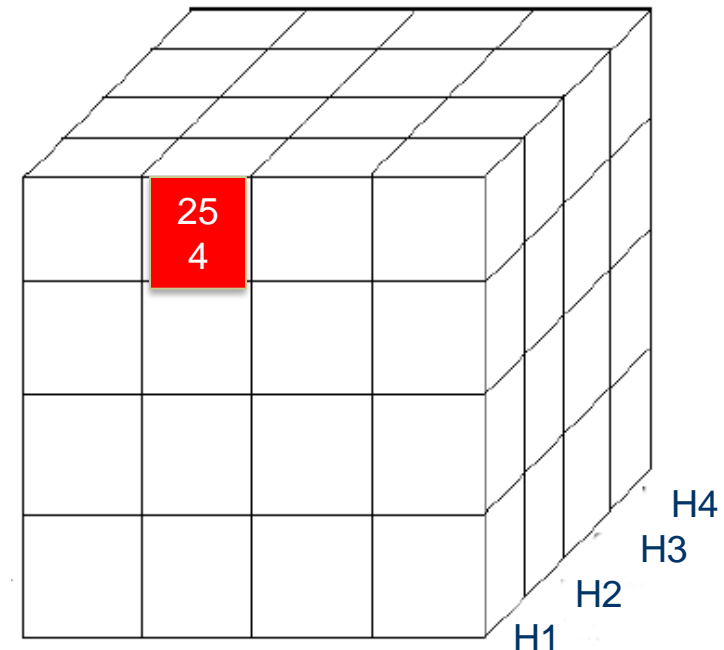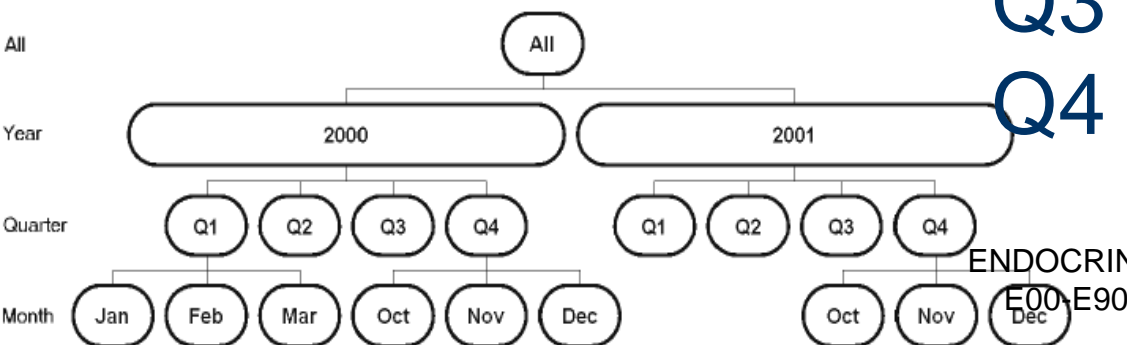
25
4

H4
H3
H2
H1

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

All

Year   2000   2001

Quarter   Q1 Q2 Q3 Q4   Q1 Q2 Q3 Q4

Month   Jan Feb Mar   Oct Nov Dec   Oct Nov Dec

25

2. MDX Spells

MDX ≠ SQL ?

	ALL (TIME)
COST	45,300,000 €
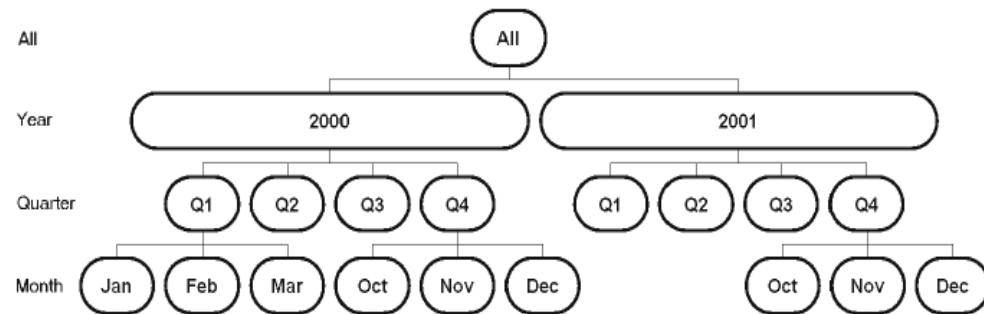
- **SELECT**
  {[TIME].[*ALL*]} **ON *COLUMNS***   *(COL dimension)*
  {[Measure].[Cost]} **ON *ROWS***   *(ROW dimension)*
  FROM [MyCube]
  (shows costs of [HOSP].[H1], default member of HOSP)
  (also for [Dx].[Circulatory])

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

SELECT
   {column headers} ON COLUMNS → SET
   {row headers} ON ROWS   → SET
   FROM [cube]   →name

SELECT
   {[Measure].[Patient]} ON COLUMNS
   {[Hospital].[Hosp1],
    [Hospital]. [Hosp2],
    [Hospital]. [Hosp3],
    [Hospital]. [Hosp4]}   ON ROWS
   FROM [MyCube]
(shows a default member of TIME)

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

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

SELECT

  {[Measure].[Patient]} ON COLUMNS

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

  FROM [MyCube]

SELECT

  {[Measure].[Patient]} ON COLUMNS

  {[Hospital].Children}   ON ROWS

  FROM [MyCube]

•

- QUESTION:
Cost in H1,H2
during 2000 (by Q), for
circulatory diseases.

Y2000	HOSP1	HOSP2
Q1	2M€	0.3M€
Q2	3.2M€	0.7M€
Q3	1.5M€	0.6M€
Q4	0.4M€	0.5M€
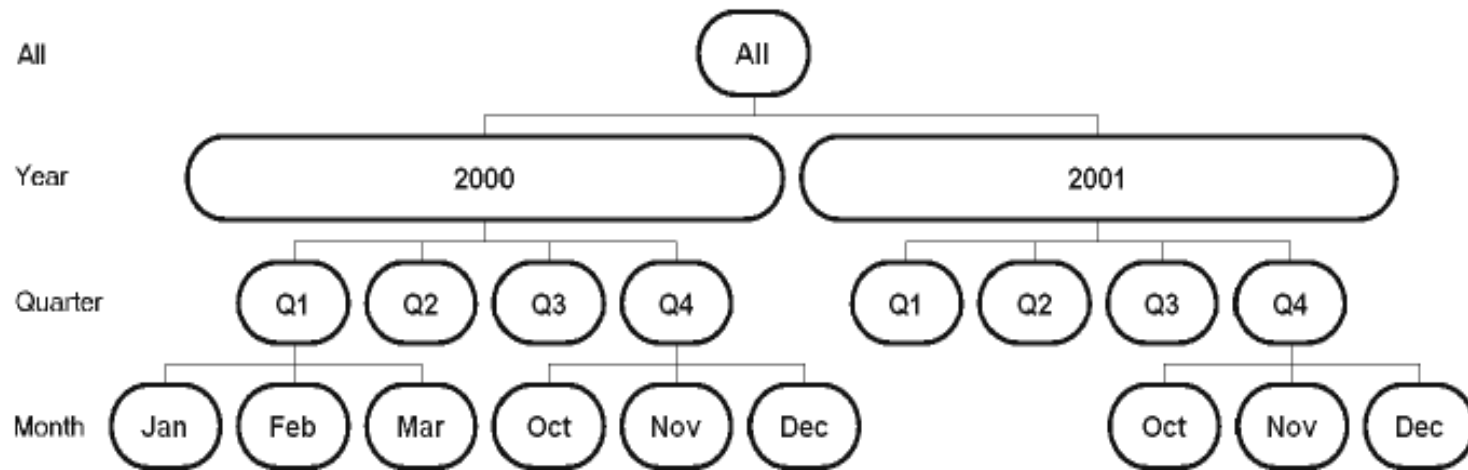
Hint:
cost/circulatory are default members

QUESTION:

Cost in H1,H2 in 2000 (by Q), for

circulatory diseases.

SELECT

   {[Hospital].[Hosp1],

    [Hospital].[Hosp2]} ON COLUMNS

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

   FROM [MyCube]

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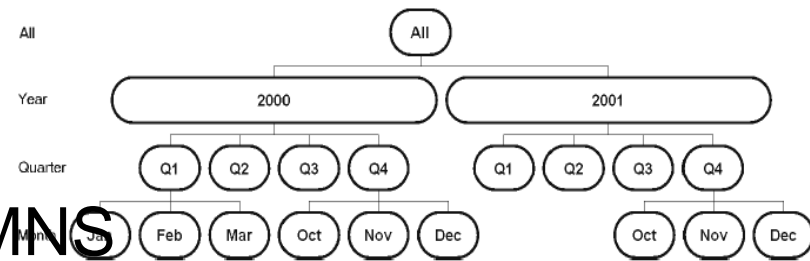
- QUESTION:

Patients no. in H1,H2 during 2000 (by Q), for circulatory diseases.

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

Hint:

Number of patients is NOT a default member

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

QUESTION:

Patients in H1,H2

during 2000 (by Q), for

circulatory diseases.

SELECT

    {[Hospital].[Hosp1],

     [Hospital].[Hosp2]} ON COLUMNS

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

FROM [MyCube]

WHERE ([Measures].[NoPat])

Hint:

Number of patients is NOT a default member

WHERE clause

    Not restricted to measures.

    Not restricted to 1 dimension.

    It is a *SLICER/DICER*.

WHERE clause

   Not restricted to measures.

SELECT

   {[Hospital].[Hosp1],

   [Hospital].[Hosp2]} ON COLUMNS

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

   FROM [MyCube]

   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€

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

WHERE clause

Not restricted to 1 dimension.

SELECT

{[Hospital].[Hosp1],

[Hospital].[Hosp2]} ON COLUMNS

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

FROM [MyCube]

WHERE ([Dx].[Respiratory],[Measures].[NoPat])

- **WHERE clause**
  - It is a *SLICER/DICER*.

WHERE ([Dx].[ENDOCRIN], [Measures].[NoPat])
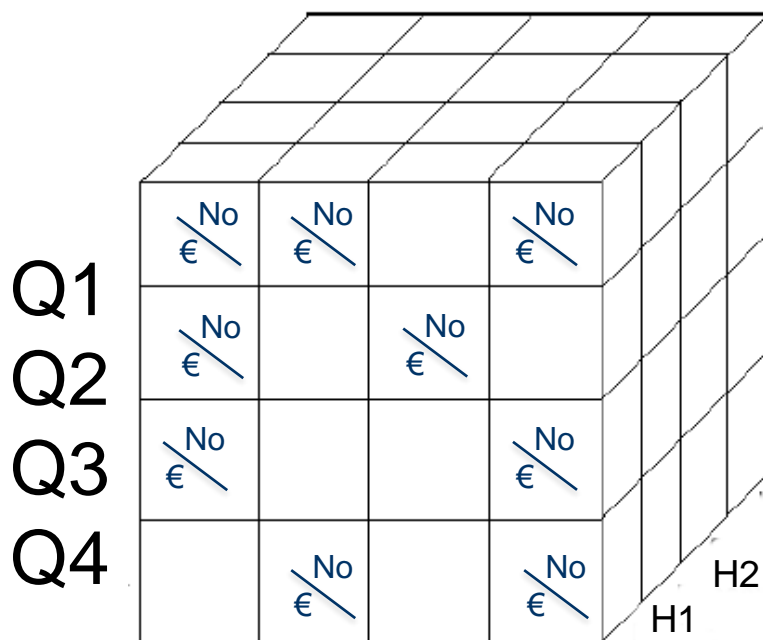
Q1
Q2
Q3
Q4

No € | No € | No €
No € | | No €
No € | | No €
| No € | No €

H4
H3
H2
H1

ENDOCRINE
E00-E90

CIRCULATORY
I00-I99

RESPIRATORY
J00-J99

DIGESTIVE
K00-K93

Q1 — No
Q2 — No
Q3 — No
Q4 — No

H4
H3
H2
H1
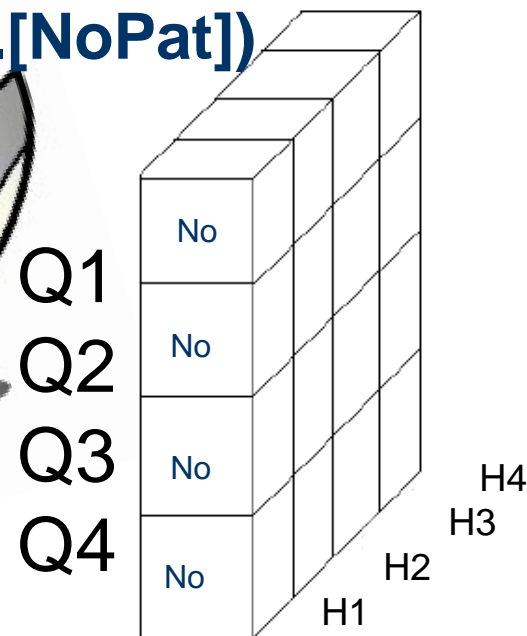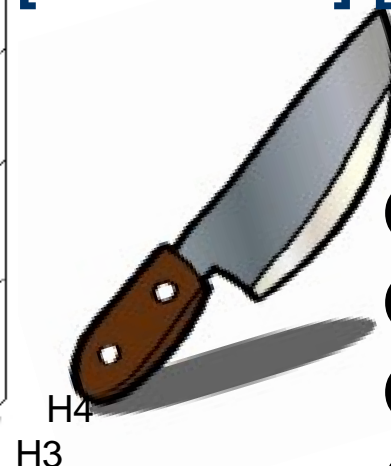
ENDOCRINE
E00-E90

3. MDX syntax

() [] {} .

- Brackets []
  - Dimensions: [Time]
  - Members: [2000]
- Dots .
  - Separators: [Time].[2000].[Q3]
- Braces ()
  - Tuples: ([DX].[Circulatory],[Hosp].[H1])

- Curly Braces {}
  - Sets: {[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]) }

SELECT
    { SET } **ON COLUMNS**
    { SET } **ON ROWS**
    FROM [cube]
    WHERE (TUPLE)

- QUESTION: Correct? Why?

SELECT
([Measures].[NoPatients]) ON COLUMNS,
{[Time].[2000].Children} ON ROWS
FROM [MyCube]

- QUESTION: Correct? Why?

SELECT
{[Dx].Children} ON COLUMNS,
{[Time]. [2000].[Q1].[May].Children} ON ROWS
FROM [MyCube]
WHERE {[Measure].[cost],[Hosp].[H2]}

- Name of a CELL.
  - In a cube, each cell has a name.

The name of this cell is:
([Time].[2000].[Q2],
[Dx].[Circulatory],
[Hospital].[H1])

2000:
Q1
Q2
Q3
Q4

H4
H3
H2
H1

ENDOCRINE
E00-E90

CIRCULATORY
I00-I99

RESPIRATORY
J00-J99

DIGESTIVE
K00-K93
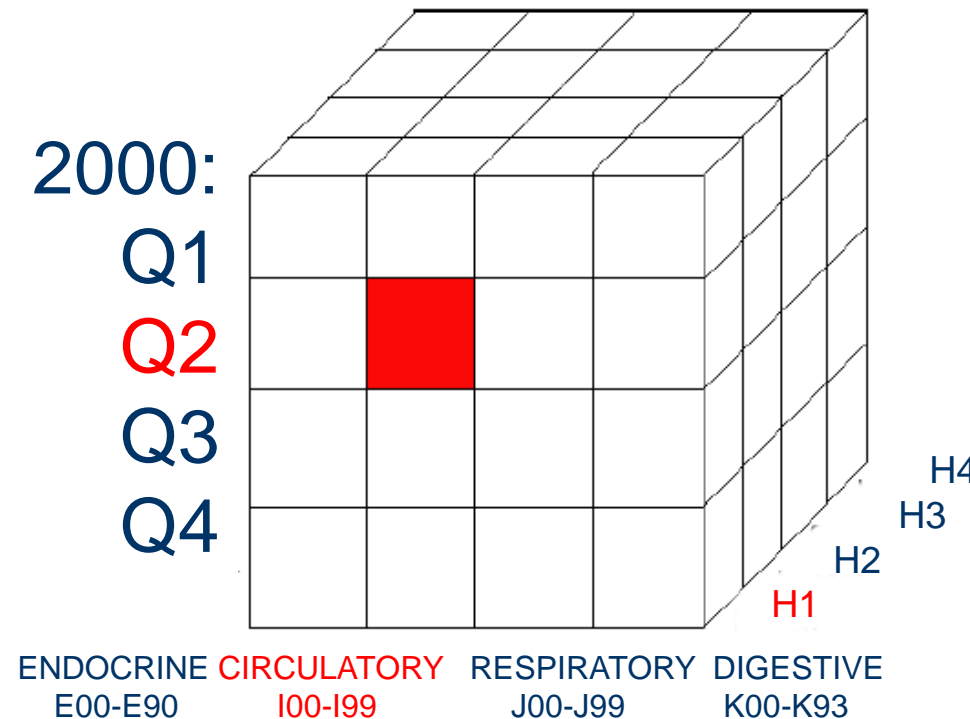
- Name of a CELL.

The name of this cell is:
([Time].[2000].[Q2].[Jun],
[Dx].[Circulatory],
[Hospital].[H1])

2000:
Jan
Feb
Mar
Apr
May
Jun
July
Aug
Sept
Oct
Nov
Dec

H4
H3
H2
H1

ENDOCRINE   CIRCULATORY   RESPIRATORY   DIGESTIVE
E00-E90   I00-I99   J00-J99   K00-K93
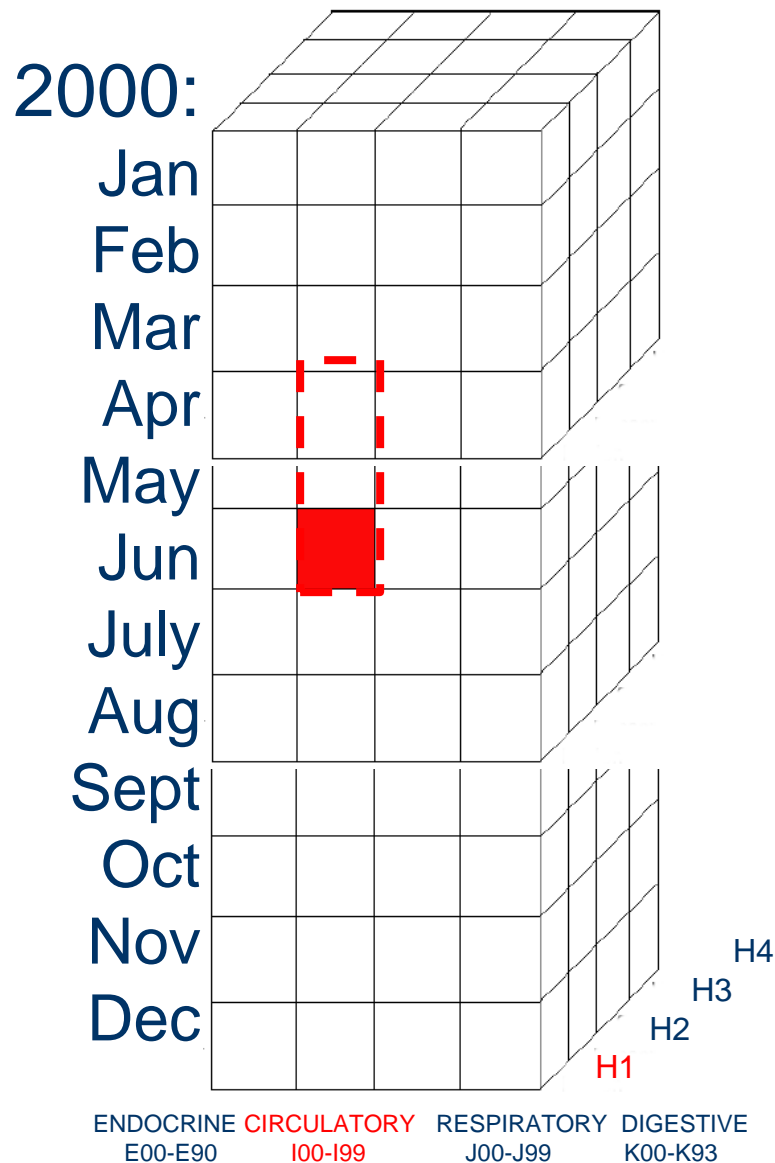
- Relative Cell Referencing:
  - CurrentMember, PrevMember, NextMember.

The name of this cell is:

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

2000:
Q1
Q2
Q3
Q4

H4
H3
H2
H1

ENDOCRINE   CIRCULATORY   RESPIRATORY   DIGESTIVE
E00-E90   I00-I99   J00-J99   K00-K93
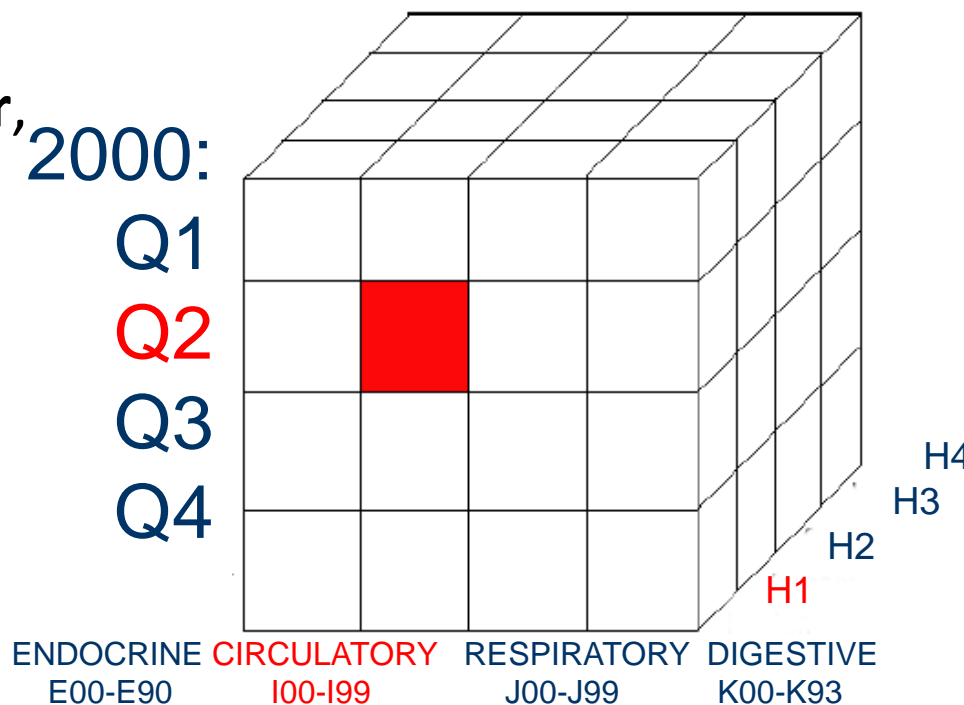
- Calculated Members: +-*/ %

"Attention improvement on circulatory patients of the 1st quarter of years 1999 and 2000".

Calculus:

([Hosp].[H1],[Dx].[Circ],[Time].[2000].[Q1],[Measure].[NoPatient])

-

([Hosp].[H1],[Dx].[Circ],[Time].[1999].[Q1],[Measure].[NoPatient])

- Calculated Members: +-*/ %

"Growth cost throughout year 2000 in H1 for Circulatory patients".

- Calculated Members: +-*/ %

"**<u>Growth</u>** cost **<u>throughout</u>** year 2000 in H1 for Circulatory patients".

  - Growth cost: cost(t)- cost(t-1)   (increment/derivate)
  - Obviate: H1 for Circulatory patients

MDX Syntax

"Growth cost throughout year 2000".

Year	Quarter	Month	Cost M€	No Patient
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

"Growth cost throughout year 2000".

Year	Quarter	Month	Cost M€	No Patient
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**	**30**
		April	5	10
		Jun	5	10
		July	2	30
	Q4		**20**	**80**
		April	5	15
		Jun	5	15
		July	10	50

0 0 0 -2 -1 3

52

Year	Quarter	Month	Cost M€	No Patient
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
	Q4		**20**	**80**
		April	5	15
		Jun	5	15
		July	10	50

"Growth cost throughout year 2000".

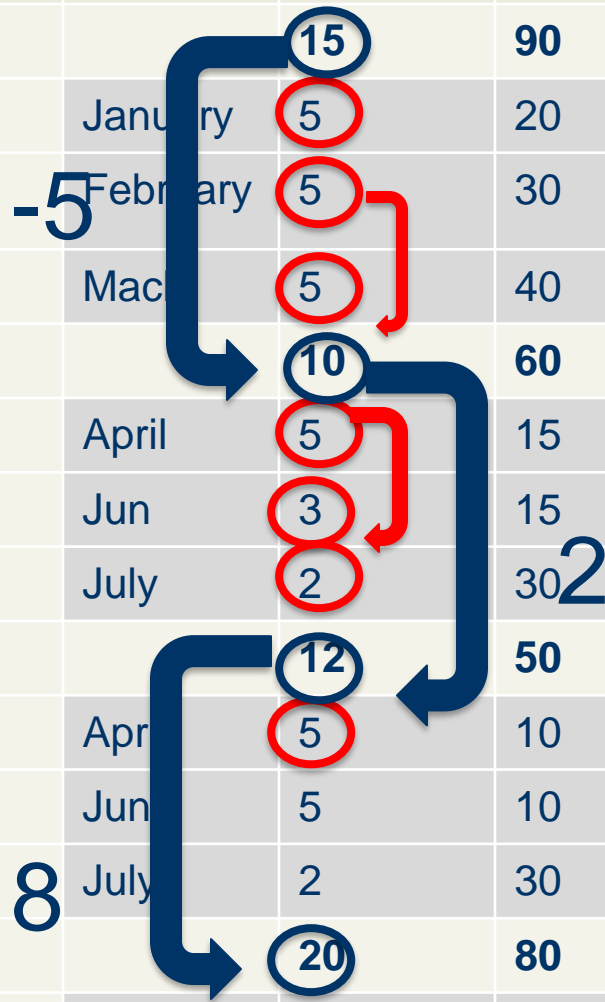[Time].[**CurrentMember**],[Measures].[Cost]

-

[Time].[**PrevMember**],[Measures].[Cost]

- Calculated Members: +-*/ %

"**Growth** cost **throughout** year 2000 in H1 for Circulatory patients".

- ([Time].[**CurrentMember**],[Measures].[Cost]
  - [Time].[**PrevMember**],[Measures].[Cost])

YES… WHAT IF WE FOCUS ON **THE QUARTER GROWTH** ?????

MDX Syntax

Year	Quarter	Month	Cost M€	No Patient
2000			57	280
	Q1		15	90
		January	5	20
		February	5	30
		March	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

-5

2

8

- Calculated Members: +-*/ % AVG SUM …

"**Growth** cost **throughout** year 2000 in H1 for Circulatory patients".

- ([Time].[**CurrentMember**],[Measures].[Cost]
 - [Time].[**PrevMember**],[Measures].[Cost])

YES… WHAT IF WE FOCUS ON **THE SEMESTER GROWTH** !!!

… **SAME EXPRESSION**. THAT'S THE COOL THING. DEPENDS ON THE TIME DIMENSION, WHICH DEFINES THE *CURRENT MEMBER* PROPERTY

- OTHER FUNCTIONS:
  - Sum (X)→Number : Sums all members of X
  - X.Lag(N) : N positions back from X.
  - X.Lead(M) : M position forward from X.
  - YTD(X)→Set : YearToDate: Members of the Year until member X.
    e.g. YTD(March) → {Jan, Feb, March}

  ClosingPeriod, OpeningPeriod, ParallelPeriod,

- Hiearchy navitagion:
  - *Member*.Children
  - *Member*.Parent
  - *Member*.FirstChild / LastChild
  - Descendants(X,n)
  - Ancestors(X,n)
  - Siblings / Cousins
  - ~~Aunt / Uncle~~
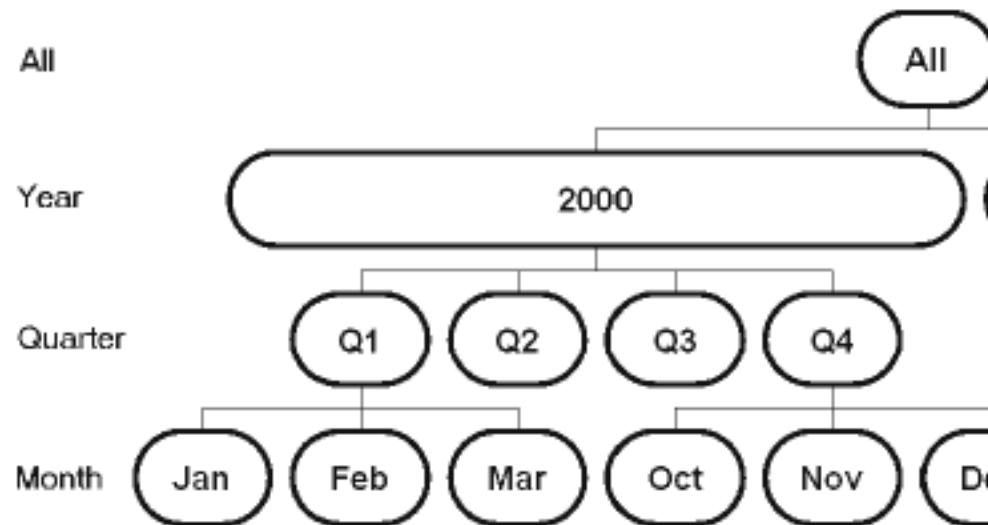
- Bibliography and Resources:

- 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