

BI = Business Intelligence
Master Data-Science
Cours 5 - MDX

Ludovic DENOYER - ludovic.denoyer@lip6.fr

UPMC

23 février 2015

Plan

- Vision générale
- ETL
- Datawarehouse
- OLAP
- Reporting
- Data Mining

Définition

OLAP

En informatique, et plus particulièrement dans le domaine des bases de données, le traitement analytique en ligne (anglais online analytical processing, OLAP) est un type d'application informatique orienté vers l'analyse sur-le-champ d'informations selon plusieurs axes, dans le but d'obtenir des rapports de synthèse tels que ceux utilisés en analyse financière. Les applications de type OLAP sont couramment utilisées en informatique décisionnelle, dans le but d'aider la direction à avoir une vue transversale de l'activité d'une entreprise.

Source : wikipedia

OLAP s'oppose au traitement de transactions en ligne (*online transaction processing abr. OLTP*) qui s'inscrit dans un système opérationnel (en production).

Définition

Le MDX (de l'anglais Multidimensional Expressions, « expressions multidimensionnelles ») est un langage de requête pour les bases de données OLAP, analogue au rôle de SQL pour les bases de données relationnelles. C'est aussi un langage de calcul avec une syntaxe similaire à celle des tableurs.

Le langage des expressions multidimensionnelles possède une syntaxe appropriée à l'interrogation et manipulation des données multidimensionnelles mémorisées dans un cube OLAP¹. Bien qu'il soit possible de traduire certaines expressions dans le langage SQL traditionnel, cela nécessite une syntaxe SQL souvent maladroite même pour des expressions MDX très simples. MDX a été adopté par une large majorité de fournisseur de la technologie OLAP et est devenu un standard de facto pour les systèmes OLAP.

Source : wikipedia

MDX

```
SELECT Measures.[Unit Sales] ON COLUMNS,  
[Store].[All Stores] ON ROWS  
FROM [Sales]
```

Délimiteurs

Les délimiteurs [et] peuvent ne pas être mis si pas d'ambiguïté.

MDX

```
SELECT Measures.MEMBERS ON COLUMNS,  
Product.Style.CHILDREN ON ROWS  
FROM [Adventure Works]
```

équivalent à :

```
SELECT [Measures].MEMBERS ON COLUMNS,  
[Product].[Style].CHILDREN ON ROWS  
FROM [Adventure Works]
```

MDX

```
SELECT [Measures].MEMBERS ON COLUMNS,  
[Date].[Calendar Year].MEMBERS ON ROWS  
FROM  
(SELECT [Measures].[Internet Sales Amount] ON COLUMNS,  
[Date].[Calendar Year].&[2004] ON ROWS  
FROM [Adventure Works])
```

Expressions de sous-cube

Attention : ne marche pas dans tous les systèmes OLAP

MDX

```
SELECT Measures.MEMBERS ON COLUMNS,  
[Store].MEMBERS ON ROWS  
FROM [Sales]
```

MEMBERS

Retourne le jeu des membres d'une dimension, d'un niveau ou d'une hiérarchie.

MDX

```
SELECT Measures.MEMBERS ON COLUMNS,  
{[Store].[Store State].[CA], [Store].[Store State].[WA]}  
  ON ROWS  
FROM [Sales]
```

Exercice

Dessinez la sortie de cette requête ?

MDX

```
SELECT Measures.MEMBERS ON COLUMNS,  
{[Store].[Store State].[CA].CHILDREN,  
[Store].[Store State].[WA].CHILDREN}  
ON ROWS  
FROM [Sales]
```

CHILDREN

Retourne le jeu des enfants d'un membre spécifié.

```
SELECT Measures.MEMBERS ON COLUMNS,  
  {[Store].[Store State].[CA],  
  DESCENDANTS([Store].[Store State].[CA], [Store City])}  
  ON ROWS  
FROM [Sales]
```

DESCENDANTS

DESCENDANTS(member, level [, flags])

Retourne le jeu de descendants d'un membre à un niveau spécifié ou à une distance spécifiée, en incluant ou en excluant éventuellement des descendants dans d'autres niveaux.

DESCENDANTS

Le "flag" peut être BEFORE, AFTER, SELF ou bien BEFORE_AND_AFTER selon le niveau désiré

```
SELECT Measures.MEMBERS ON COLUMNS,  
  {[Store].[Store State].[CA],  
  DESCENDANTS([Store].[Store State].[CA], [Store City],AFTER)  
  ON ROWS  
FROM [Sales]
```

```
SELECT Measures.MEMBERS ON COLUMNS,  
  {[Store].[Store State].[CA],  
  DESCENDANTS([Store].[Store State], [Store City]  
  ,BEFORE_AND_AFTER)} ON ROWS FROM [Sales]
```

MDX

```
select AddCalculatedMembers([Measures].Members) ON COLUMNS  
      {[Store].[USA].[CA],  
Descendants([Store].[USA].[CA], [Store].[Store City])}  
ON ROWS  
from [Sales]
```

AddCalculatedMembers

Retourne un ensemble généré par l'ajout de membres calculés à un ensemble spécifié.

MDX

```
SELECT {[Store Type].[Store Type].MEMBERS} ON COLUMNS,  
        {[Store].[Store State].MEMBERS} ON ROWS  
FROM [Sales]  
WHERE (Measures.[Unit Sales])
```

Slicer

WHERE spécifie une "tranche dimensionnelle" dans le cube

MDX

```
SELECT {[Store Type].[Store Type].MEMBERS} ON COLUMNS,  
        {[Store].[Store State].MEMBERS} ON ROWS  
FROM [Sales]  
WHERE (Measures.[Unit Sales], [Time].[Year].[1997])
```

Slicer

WHERE spécifie une "tranche dimensionnelle" dans le cube

WITH SET

WITH SET permet la création d'ensembles

```
WITH SET [ChardonnayChablis] AS
    {[Product].[All Products].[Drink].[Good Chardonnay],
    [Product].[All Products].[Drink].[Pearl Chardonnay],
    [Product].[All Products].[Drink].[Portsmouth],
    [Product].[All Products].[Drink].[Walrus Chardonnay],
SELECT
    [ChardonnayChablis] ON COLUMNS,
    {Measures.[Unit Sales]} ON ROWS
FROM Sales
```


Calculated Members

WITH MEMBER

MDX permet de rajouter des calculs directement dans les requêtes MDX. Ceci s'effectue grâce à la syntaxe :

```
WITH MEMBER parent.name AS 'expression'
```

```
WITH MEMBER Measures.ProfitPercent AS  
'(Measures.[Store Sales] - Measures.[Store Cost]) /  
(Measures.[Store Cost])', FORMAT_STRING = '#.00%'
```

```
WITH
    MEMBER [Measures].[Special Discount] AS
        [Measures].[Discount Amount] * 1.5
SELECT
    [Measures].[Special Discount] on COLUMNS,
    NON EMPTY [Product].[Product].MEMBERS ON Rows
FROM [Adventure Works]
WHERE [Product].[Category].[Bikes]
```

Calculated Members

WITH MEMBER

MDX permet de rajouter des calculs directement dans les requêtes MDX. Ceci s'effectue grâce à la syntaxe :

```
WITH MEMBER parent.name AS 'expression'
```

```
WITH MEMBER [Time].[First Half 97] AS  
'[Time].[1997].[Q1] + [Time].[1997].[Q2]',  
MEMBER [Time].[Second Half 97] AS  
'[Time].[1997].[Q3] + [Time].[1997].[Q4]'
```

Calculated Members

```
WITH MEMBER [Time].[First Half 97] AS  
  '[Time].[1997].[Q1] + [Time].[1997].[Q2]'  
MEMBER [Time].[Second Half 97] AS  
  '[Time].[1997].[Q3] + [Time].[1997].[Q4]'  
SELECT {[Time].[First Half 97], [Time].[Second Half 97]},  
  [Time].[1997].CHILDREN} ON COLUMNS,  
  {[Store].[Store Name].MEMBERS} ON ROWS  
FROM [Sales]
```

GENERATE

Applique un jeu à chaque membre d'un autre jeu, puis effectue la jointure par union des jeux résultants.

```
Generate( Set_Expression1 , Set_Expression2 [ , ALL ] )
```

```
SELECT  
GENERATE( [Time].[Year].MEMBERS, {[Measures].[Unit Sales]}  
ON 0  
FROM [Sales]
```

GENERATE

Applique un jeu à chaque membre d'un autre jeu, puis effectue la jointure par union des jeux résultants.

```
Generate( Set_Expression1 , Set_Expression2 [ , ALL ] )
```

```
SELECT  
GENERATE( [Time].[Year].MEMBERS, {[Measures].MEMBERS}, ALL)  
ON 0  
FROM [Sales]
```

```
SELECT {GENERATE([Time].[Year].MEMBERS,  
{[Time].CURRENTMEMBER, [Time].CURRENTMEMBER.CHILDREN})}  
ON COLUMNS,  
[Promotions].[All Promotions].CHILDREN ON ROWS  
from Sales
```

CURRENTMEMBER

Retourne le membre actuel dans une hiérarchie spécifique au cours d'une itération.

```
SELECT {GENERATE({[Store].[USA].[CA],[Store].[USA].[WA]},  
DESCENDANTS([Store].CURRENTMEMBER, [Store Name]))}  
ON COLUMNS,  
[Promotions].[All Promotions].CHILDREN ON ROWS  
FROM [Sales]  
WHERE (Measures.[Unit Sales])
```



```
WITH MEMBER measures.X AS  
    [Product].children.count  
SELECT Measures.X ON 0  
FROM [Sales]
```

COUNT

Retourne le nombre de cellules d'un ensemble

```
SELECT Measures.MEMBERS ON COLUMNS,  
TOPPERCENT({[Store].[Store City].MEMBERS}, 50,  
Measures.[Sales Count]) ON ROWS  
FROM [Sales]
```

TOPPERCENT

Trie un jeu en ordre décroissant et retourne un jeu de tuples avec les valeurs les plus élevées dont le total cumulé est égal ou supérieur à un pourcentage spécifié.

```
WITH MEMBER Measures.[Maximum Sales] AS  
'MAX(DESCENDANTS([Time].CURRENTMEMBER, [Time].[Month]),  
Measures.[Unit Sales])'  
SELECT {[Time].[1997]} ON COLUMNS,  
[Product].[Product Category].MEMBERS ON ROWS  
FROM [Sales]  
WHERE (Measures.[Maximum Sales])
```

Autres mesures

Les fonctions disponibles sont AVG, MEDIAN, MAX, MIN, VAR, and STDDEV

References

- MDX Reference @ MSDN :
<http://msdn2.microsoft.com/en-us/library/ms145506.aspx>
- The Baker's Dozen : 13 Tips for Querying OLAP Databases with MDX :
<http://www.devx.com/codemag/Article/37460/1954?pf=true>
- Tutorial : Introduction to Multidimensional Expression (MDX). http://www.fing.edu.uy/inco/grupos/csi/esp/Cursos/cursos_act/2005/DAP_SistDW/Material/2-SDW-Laboratorio1-2005.pdf
- MDX resources : <http://www.mosha.com/msolap/mdx.htm>