

Initiation au décisionnel (Business Intelligence, DataWarehouse, OLAP)

par [Mohamed Taslimanka Sylla](#) ([Page personnelle](#))

Date de publication : 20/10/2007

Dernière mise à jour : 27/09/2011

Ce tutoriel a pour objectif de vous expliquer les objectifs du décisionnel ainsi que le jargon qui va avec. Dans une deuxième partie, on passera à un tutoriel qui traite de l'analyse OLAP, faisant partie intégrante du décisionnel.

I - Introduction.....	3
II - Aspect théorique.....	3
II-A - Pourquoi le décisionnel ?.....	3
II-B - Qui a besoin du décisionnel ?.....	3
II-C - Architecture des systèmes décisionnels.....	4
II-C-1 - Les sources de données.....	4
II-C-2 - L'entrepôt de données.....	5
II-C-3 - Le serveur OLAP ou serveur d'analyse.....	6
II-C-4 - Le générateur d'états.....	6
II-C-5 - Quelques termes usuels du décisionnel.....	6
III - Aspect pratique.....	7
III-A - Pré-requis.....	7
III-B - Créer une Source de données ODBC.....	8
III-C - Démarrer OLAP Manager.....	8
III-D - Créer votre base de données d'analyse.....	9
III-E - Créer une source de données OLAP.....	10
III-F - Ouvrir l'assistant création de Cube.....	10
III-G - Ajouter une mesure au cube.....	11
III-H - Construire la dimension Temps.....	12
III-I - Construire la dimension Produit.....	15
III-J - Construire la dimension Magasin.....	17
III-K - Construire la dimension Promotion.....	18
III-L - Terminer la création du cube.....	19
III-M - Editer le cube dans l'éditeur de cube.....	20
III-N - Ajouter un rôle au cube.....	22
III-O - Concevoir le type de stockage et traiter le cube.....	23
III-P - Visualiser les métadonnées du cube.....	25
III-Q - Naviguer sur les données du cube.....	25
IV - Conclusion.....	27
V - Remerciements.....	27

I - Introduction

Vous avez certainement dû entendre parler d'au moins un de ces trois termes qui sont intimement liés : Business intelligence, Datawarehouse et Analyse OLAP. En effet, depuis les années 2000-2001, le marché du décisionnel ne cesse d'exploser en France (et partout d'ailleurs) surtout avec l'investissement de plusieurs grandes sociétés qui souhaitent instaurer un système de Business Intelligence (B.I.) dans leur organisation. Ce système difficile à mettre en œuvre, demandant une expertise et nécessitant une maîtrise d'ouvrage de la part des informaticiens concepteurs du système rend souvent difficile le recrutement de ces profils. Mais c'est quoi le décisionnel ? Comment y débiter ? Eh bien je vais tenter dans cet article de donner un aperçu de tout ce jargon souvent méconnu et n'ayant souvent pas d'équivalents français et j'essaierai après un bref aspect théorique, de passer à un aspect pratique traitant l'analyse OLAP avec Analysis Services de Microsoft.

N.B. : la deuxième partie de cet article a été réalisée avec la version 7.0 de SQL Server. Étant encore nouveau dans la version 2005, j'attends encore un moment pour réaliser une nouvelle version de cette partie pratique avec SQL Server 2005.

II - Aspect théorique

Dans cet aspect théorique nous allons un peu expliquer l'utilité du décisionnel, des acteurs du décisionnel et des architectures usuelles.

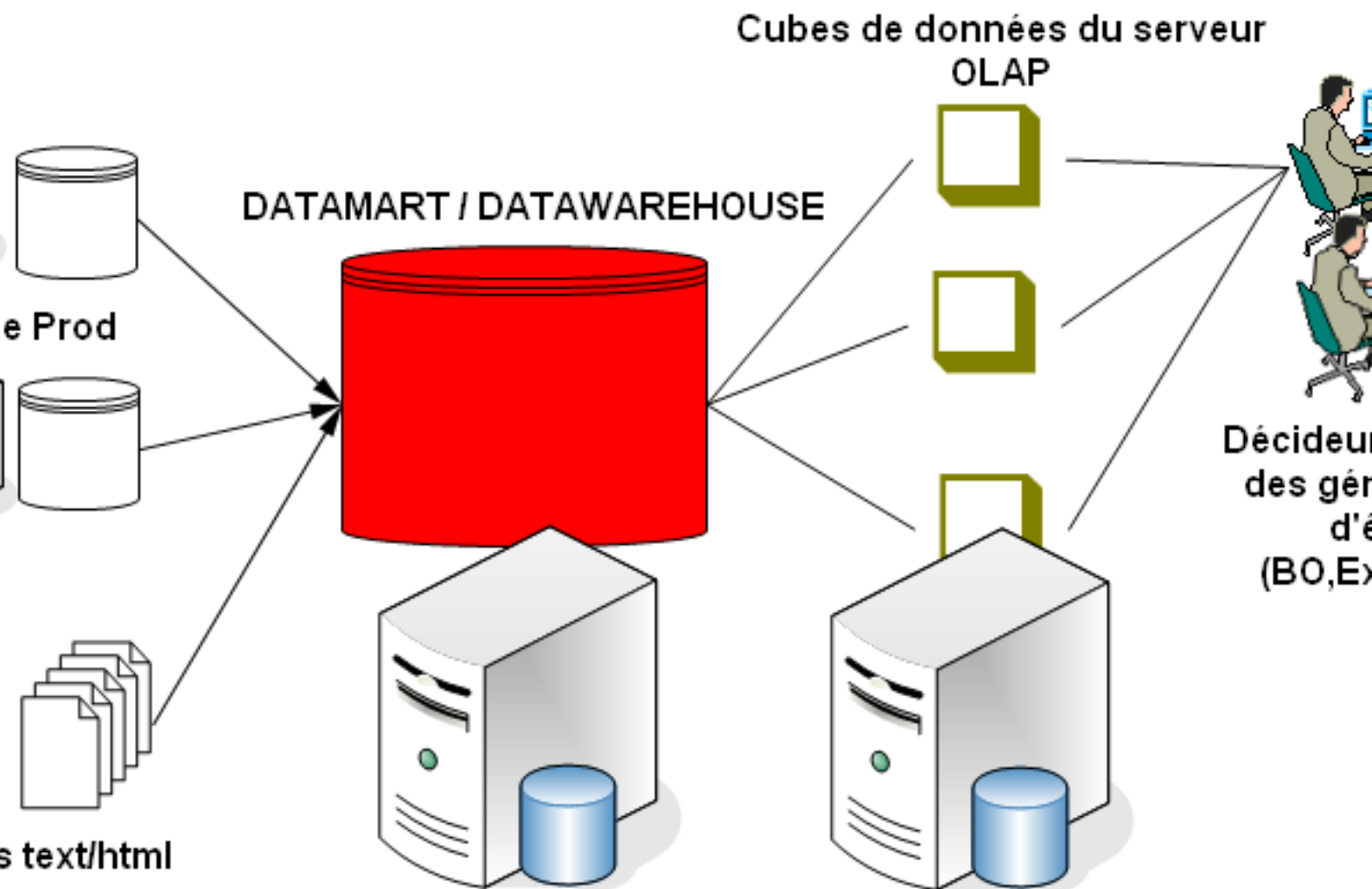
II-A - Pourquoi le décisionnel ?

Tout d'abord, rappelons-le, le décisionnel ne concerne souvent que les entreprises qui gèrent un historique de leurs événements passés (faits, transactions etc.). Les entreprises qui viennent de naître n'ont souvent pas besoin de faire du décisionnel car elles n'ont pas encore besoin de catégoriser ou de fidéliser leurs clients. Le souci majeur pour elles serait plutôt d'avoir le maximum de clients et c'est après en avoir récupéré un grand nombre qu'elles penseront certainement à les fidéliser et leur proposer d'autres produits susceptibles de les intéresser. C'est ce que l'on appelle Customer Relationship Management (CRM ou gestion des relations clients).

II-B - Qui a besoin du décisionnel ?

Comme cela peut se deviner, les décideurs sont les principaux utilisateurs des systèmes décisionnels. Les décideurs sont généralement des « marketeurs » ou analystes en général. Ces derniers établissent généralement des plans marketing qui leur permettent de mieux cibler leurs clients, de les fidéliser etc. Et pour cela, ils ont besoin d'indicateurs et des données résumées de leurs activités (ils n'ont souvent besoin de détail que pour des cas spécifiques). Par exemple, contrairement aux systèmes relationnels (ou base gestion) où les utilisateurs chercheront à connaître leurs transactions pour faire un bilan, les systèmes décisionnels quant à eux cherchent plutôt à donner un aperçu global pour connaître les tendances des clients (d'où l'opposition des deux modes [quantitatif contre qualitatif]).

II-C - Architecture des systèmes décisionnels



Exemple d'architecture décisionnelle.

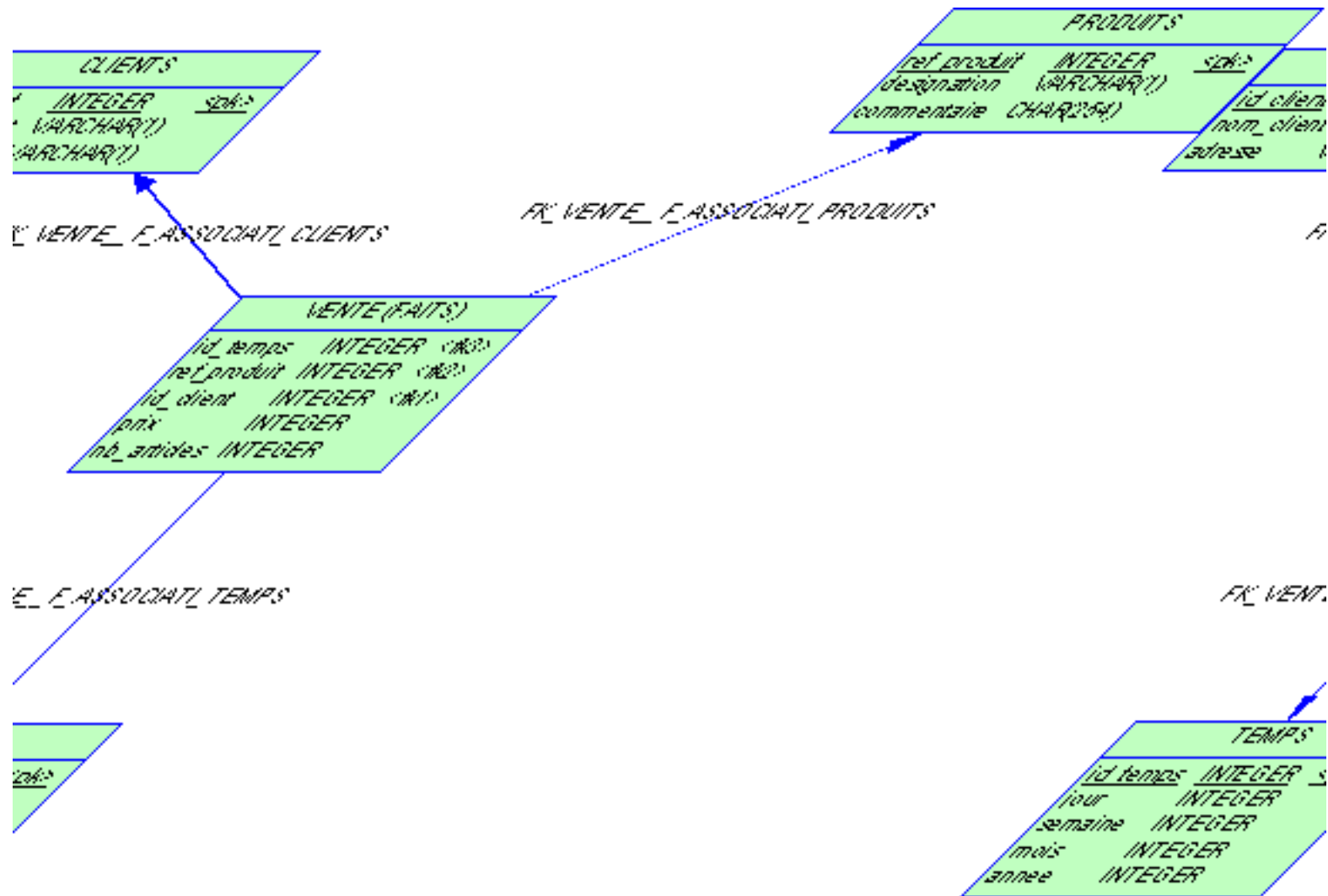
Voici une architecture de système décisionnel très utilisée. Dans cette architecture, on dispose d'un entrepôt de données ou DataWarehouse (généralement, il s'agit plutôt d'un datamart qui est plus petit que le DW et qui concerne un domaine bien particulier [finance, ressources humaines etc.]). L'entrepôt (ou encore info-centre !) centralise les données issues de plusieurs sources (bases de production de l'entreprise, fichiers textes, documents web [html, xml, sgml etc.] etc.). Ces données sont fusionnées dans l'entrepôt qui est généralement une grosse base de données (SQL Server, Oracle etc.). Ensuite, une fois l'entrepôt confectionné, des données sont extraites dans des serveurs d'analyse ou serveurs OLAP sous forme de cubes de données (Analysis Server, EssBase etc.) afin d'être analysées. Enfin, des générateurs d'états (Business Objects, Crystal Report etc.) sont utilisés afin de présenter l'étude aux utilisateurs finaux ou décideurs (ex: analystes marketing).

II-C-1 - Les sources de données

Les sources de données sont souvent diverses et variées et le but est de trouver des outils ETL (Extraction / Transformation / Loading) afin de les extraire, de les nettoyer, de les transformer et de les mettre dans l'entrepôt de données (DTS de SQL Server est un exemple d'outil ETL). Des outils comme Datastage ou Talend (monde open source) sont spécialisés en la matière.

II-C-2 - L'entrepôt de données

Il est le cœur du système décisionnel et demande une analyse profonde de la part de la maîtrise d'ouvrage. La conception d'un DataWarehouse diffère de la conception d'une base de données relationnelle. En effet, alors que les bases de données relationnelles tendent le plus souvent à être normalisées, les bases de données multidimensionnelles, elles, sont plutôt dénormalisées, respectant le modèle en étoile ou le modèle en flocon. Voici ci-dessous un exemple de schéma d'un entrepôt de données:



Modèle physique d'un schéma en étoile.

Le modèle physique ci-dessus contient une table centrale à laquelle toutes les autres tables sont liées (modèle en étoile). La table centrale (ici table VENTE) est appelée la table des faits et contient toutes les autres clés des autres tables. Cette table de faits contient aussi une ou plusieurs valeurs numériques particulières (ici prix et nb_articles) appelées mesures. Généralement un niveau de granularité est aussi défini pour la table des faits (regroupe-t-on par exemple un ensemble de ventes de même type pour en faire un enregistrement ? Les enregistrements sont-ils unitaires (un enregistrement par transaction ?). Les autres tables du modèle sont appelées tables de dimensions. Ici par exemple, on dispose des dimensions CLIENTS, PRODUITS et TEMPS. Une dimension Temps est presque toujours présente dans les bases multidimensionnelles tout simplement parce qu'on analyse les données dans le temps.

II-C-3 - Le serveur OLAP ou serveur d'analyse

OLAP (On-Line Analytical Processing) est opposé à OLTP (On-Line Transactional Processing) et a pour but d'organiser les données à analyser par domaine/thème et d'en ressortir des résultats pertinents pour le décideur. Les résultats sont donc des résumés et peuvent être obtenus par différents algorithmes de datamining (fouille de données) du serveur d'analyse. On peut par exemple établir le résultat suivant : « Les clients qui achètent généralement du beurre et du pain achètent aussi du lait ». Ces résultats pourraient amener l'organisation (ici en l'occurrence une grande distribution) à disposer ses rayons de telle façon qu'à côté de l'emplacement du beurre, elle mettra le pain et le lait.

II-C-4 - Le générateur d'états

Le générateur d'état permet seulement de mieux appréhender le résultat de l'analyse. L'utilisateur final n'étant pas forcément un informaticien, il aura plus de facilité dans des états Business Objects (ou même dans des feuilles de données Excel) avec des diagrammes et courbes statistiques que d'aller directement requêter dans le serveur d'analyse. Au passage, je rappelle que travaillant généralement avec SQL Server et Analysis Services, le langage de requêtage multidimensionnel a pour nom MDX (qui ressemble au SQL mais n'est pas du SQL). Les états permettent également de faire de l'exploration (navigation) de données (notamment du Rollup / Drill-Down).

II-C-5 - Quelques termes usuels du décisionnel

Datawarehouse : entrepôt de données

Datamart : petit entrepôt de données à l'échelle d'un département ou succursale d'une grande société. Généralement un datamart déverse ses données chez sa mère qui est le datawarehouse

OLTP : OnLine Transactional Processing. Il s'agit des traitements transactionnels. Par exemple, les logiciels des caisses enregistreuses des chaînes de magasins font du OLTP.

OLAP : OnLine Analytical Processing. Opposé à l'OLTP, faire de l'OLAP signifie faire de l'analyse de données. Analyser les ventes, détecter les fraudes, prospector des clients font partie du processus OLAP.

ETL : un outil ETL (Extraction / Transformation / Loading) permet à partir de diverses sources de données, d'extraire de l'information, de faire des transformations afin de nettoyer les données et de charger des données utiles dans l'entrepôt de données. Les sources de données peuvent être diverses (HTML, XML, base de données, fichiers texte, tableurs, ERP etc.).

Serveur d'analyse : un serveur d'analyse ou serveur OLAP est un serveur de base de données multidimensionnelle. Exemple : Analysis Server est un serveur de bases multidimensionnelles.

Base de données multidimensionnelle : une base de données multidimensionnelle par opposition à une base de données relationnelle est une base dénormalisée où il existe une table centrale (table de faits) liée à toutes les autres tables (tables de dimensions).

Table de faits : comme son nom l'indique, une table de faits est une table contenant tous les faits du SI et dont dépendent toutes les autres tables. Cette table ne contient que des clés étrangères venant des tables de dimensions et des valeurs numériques appelées mesures. Exemple de table de faits : table des ventes.

Tables de dimensions : les tables de dimensions sont des tables servant d'axes d'analyse. On peut par exemple analyser les ventes (table de faits) suivant l'axe des temps (table de dimensions) pour indiquer par exemple pendant quel trimestre de l'année les ventes ont explosé.

Mesure : une mesure est une quantité présente dans la table de faits qui permet de mesurer les faits. Par exemple, nombre de ventes ou prix unitaire sont des exemples de mesures.

Cube : un cube de données est une structure dimensionnelle comme une table est une structure relationnelle. Un cube est constitué d'une ou plusieurs tables de faits avec leurs tables de dimensions. On peut par exemple considérer un cube vente contenant sa table de faits « vente » et ses tables de dimensions « clients », « régions » et « temps ».

Niveau de hiérarchie : un niveau de hiérarchie se définit au niveau des tables de dimensions. Cela permet d'agréger les données. Par exemple, supposons qu'on ait la dimension région contenant la liste des villes, on pourrait faire un niveau de hiérarchie (niveau 1) classant les villes en régions, ensuite un niveau plus bas qui les classerait en départements (niveau 2).

Drill-down: faire un drill-down, c'est avoir un niveau de détails sur les données. Par exemple Supposons qu'on veuille voir le détail des ventes pour le premier trimestre de l'année 1997. On dit qu'on fait un drill-down sur l'axe (ou dimension) temps. C'est-à-dire qu'on ne veut pas voir seulement les données de l'année 1997 mais descendre à un niveau de détail plus bas.

Roll-up: rollup est le contraire de drill-down. C'est donc faire de l'agrégation (ou résumé) des données.

III - Aspect pratique

Pour l'aspect pratique, nous allons supposer que notre datawarehouse est déjà mis en place. Tout ce que nous allons faire, c'est faire de l'analyse OLAP sur notre datawarehouse.

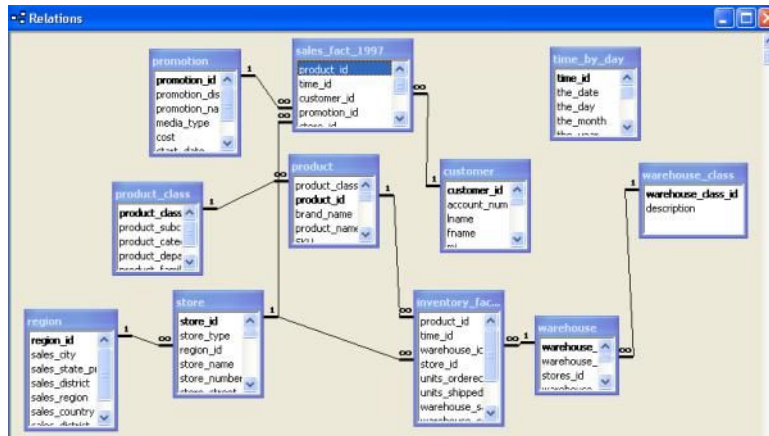
III-A - Pré-requis

Avant de commencer ce tutorial, vous devez posséder la base Access FoodMart.mdb qui est notre datawarehouse. Cette base est une base d'exemple fournie par Microsoft. Une remarque assez importante sur cette base est qu'ici on voit que la base Access est déjà sous format d'un datawarehouse avec des tables de faits (sales_fact_1997 ou sales_fact_1998) et tables de dimensions (product ou region ou store).



Ci-dessous on voit bien que la table sales_fact_1997 est une table de fait qui référence les autres tables. Cette table ne contient que des valeurs numériques (les identifiants des tables de dimension ainsi que les mesures).

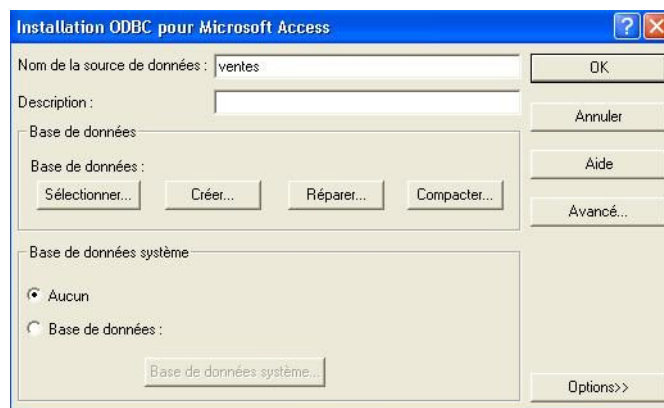
NB : Cette base de données exemple se situe sur C:\Program Files\OLAP Services\Samples\FoodMart.mdb (bien sûr pour la version 7 de SQL Server).



III-B - Créer une Source de données ODBC

Avant de commencer à travailler avec OLAP Manager, vous devez créer un lien ODBC sur vos données ici en l'occurrence sur la base Access FoodMart.mdb. Pour cela faire :

- A partir du menu démarrer, aller dans panneau de configuration, choisir Source de données ODBC.
- Dans l'onglet source de données system, cliquer sur ajouter.
- Choisir Microsoft Access Driver (*.mdb) et cliquer sur terminer.
- Entrer ventes pour le nom de la source de données.

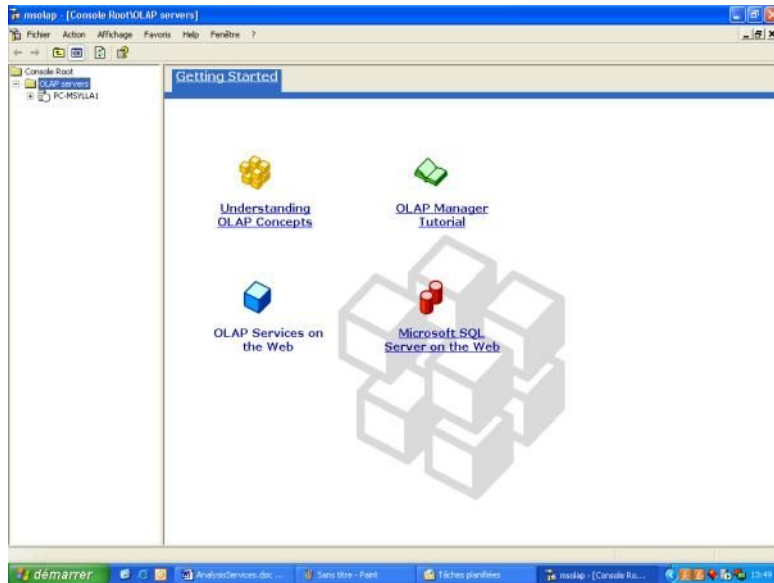


- Cliquer le bouton sélectionner et choisir dans la boîte de dialogue qui apparaît la base FoodMart.mdb qui se trouve dans C:\Program Files\OLAP.Cliquer ensuite sur OK.
- Cliquer sur OK pour fermer la boîte de dialogue de lien ODBC Access.
- Cliquer sur OK pour fermer la boîte de source de données ODBC.

III-C - Démarrer OLAP Manager

OLAP manager est un snap-in qui se situe sur la console MMC Microsoft Management Console. Pour démarrer OLAP manager faire :

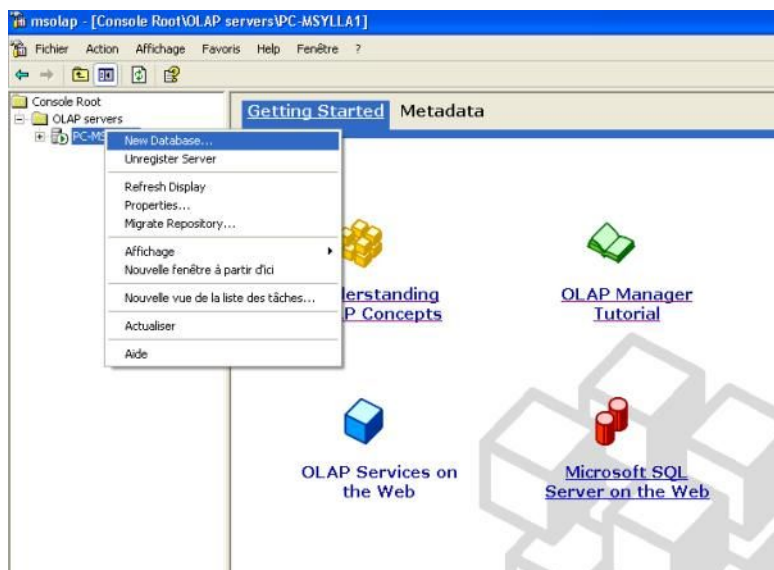
- A partir du menu démarrer, dans programmes, choisir Microsoft SQL Server 7.0, ensuite OLAP Services et choisir OLAP Manager.



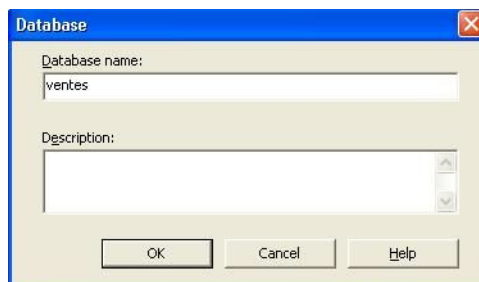
III-D - Créer votre base de données d'analyse

Maintenant vous pouvez travailler avec OLAP manager. Avant d'ouvrir votre cube en mode design, vous devez d'abords mettre en place une structure de données et vous connecter à la source de données créées plus haut. Pour mettre en place cette source de données faire :

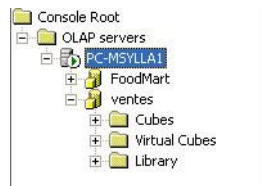
- Dans l'OLAP manager, défiler le noeud OLAP Servers
- Choisir le nom de votre serveur OLAP (ici PC-MSYLLA1).



- Faire un Click droit sur le serveur et choisir " Nouvelle base de donnée "



- Choisir un nom à donner à votre base OLAP et faire OK. Dans mon cas je l'appelle " ventes "
- Vous venez donc de créer votre base de données OLAP. Vous pouvez vous amuser à regarder ce que contient l'arborescence de votre base d'analyse. On peut donc y remarquer les snap-in : " Cubes ", " Virtual Cubes " et " Library ".

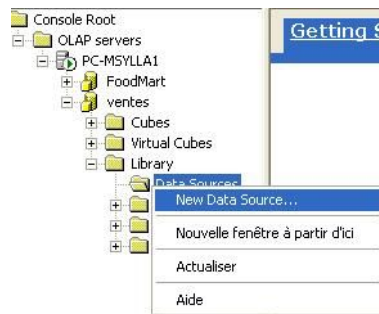


Maintenant que la base de données est créée, il va falloir se connecter à notre datawarehouse et pour cela il faudra créer une source de donnée sous OLAP manager et choisir notre source de données ODBC créées précédemment.

III-E - Créer une source de données OLAP

Pour cela faire :

- Dans l'arborescence du OLAP manager, défiler " Library " et faire un click droit sur " Source de données " et choisir Nouvelle source de données
- Choisir alors la source de données ODBC " ventes " que nous avons créées dans la section " créer une source ODBC ".



Maintenant que nous avons tout configuré, il est temps de construire notre cube de données. Pour cela nous allons considérer le scénario suivant :

Scénario : Vous êtes un DBA travaillant pour la société Food Mart. Food Mart est une large chaîne alimentaire avec des ventes enregistrées dans les 50 états des Etats-Unis. Le département de marketing voudrait alors analyser ses ventes réalisées pour la seule année 1997. Avec les données stockées dans le " datawarehouse ", vous êtes chargé de construire une structure multidimensionnelle (un cube) pour avoir des temps de réponse plus rapide lorsque les analystes marketing interrogent la base de données.

Rappel : Un cube de données contient des mesures (ou données qualitatifs comme les coûts ou le nombre de vente etc..) et des dimensions (ou données métiers descriptives comme les régions géographiques, le temps ou encore clients etc.).

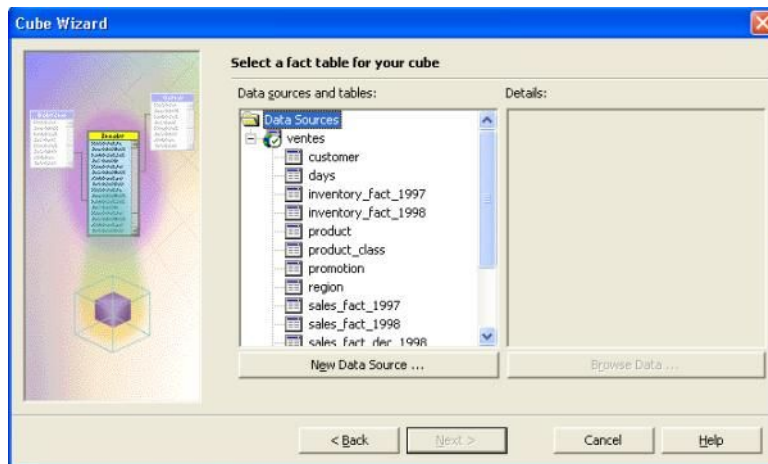
III-F - Ouvrir l'assistant création de Cube

Dans l'arborescence du OLAP manager, dans la base vente ", faire un click droit sur " Cube " et choisir " Nouveau cube " puis choisir le sous-menu " Assistant ".

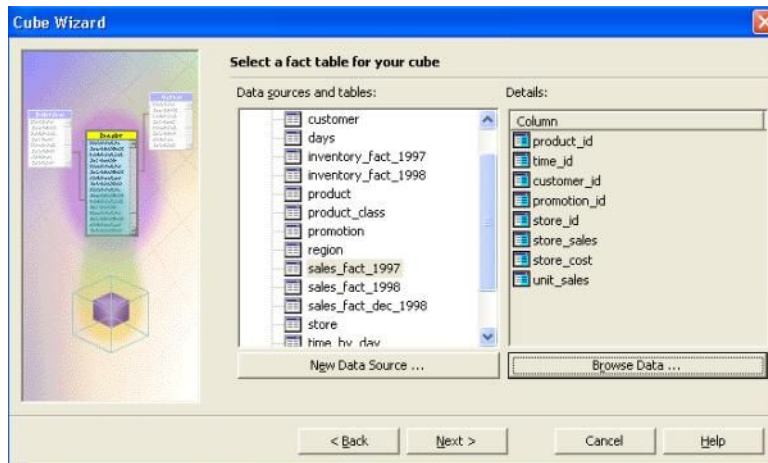


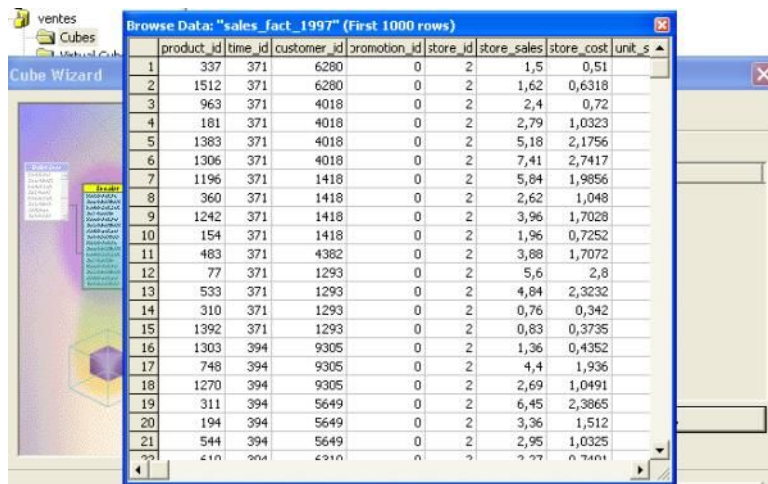
III-G - Ajouter une mesure au cube

- Dans l'écran de l'assistant de création du cube, faire suivant



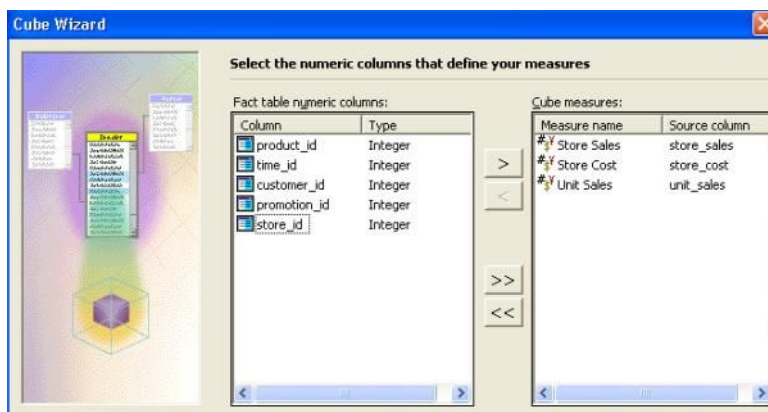
- Dans l'écran " Choisir une table de fait pour votre cube ",dérouler le data source " ventes "





	product_id	time_id	customer_id	promotion_id	store_id	store_sales	store_cost	unit_s
1	337	371	6280	0	2	1,5	0,51	
2	1512	371	6280	0	2	1,62	0,6318	
3	963	371	4018	0	2	2,4	0,72	
4	181	371	4018	0	2	2,79	1,0323	
5	1383	371	4018	0	2	5,18	2,1756	
6	1306	371	4018	0	2	7,41	2,7417	
7	1196	371	1418	0	2	5,84	1,9856	
8	360	371	1418	0	2	2,62	1,048	
9	1242	371	1418	0	2	3,96	1,7028	
10	154	371	1418	0	2	1,96	0,7252	
11	483	371	4382	0	2	3,88	1,7072	
12	77	371	1293	0	2	5,6	2,8	
13	533	371	1293	0	2	4,84	2,3232	
14	310	371	1293	0	2	0,76	0,342	
15	1392	371	1293	0	2	0,83	0,3735	
16	1303	394	9305	0	2	1,36	0,4352	
17	748	394	9305	0	2	4,4	1,936	
18	1270	394	9305	0	2	2,69	1,0491	
19	311	394	5649	0	2	6,45	2,3865	
20	194	394	5649	0	2	3,36	1,512	
21	544	394	5649	0	2	2,95	1,0325	

- Choisir " sales_fact_1997 ", représentant les ventes de 1997.Vous pouvez aussi visualiser les données contenues dans cette table.
- Cliquez sur " suivant "



Select the numeric columns that define your measures

Fact table numeric columns:

Column	Type
product_id	Integer
time_id	Integer
customer_id	Integer
promotion_id	Integer
store_id	Integer

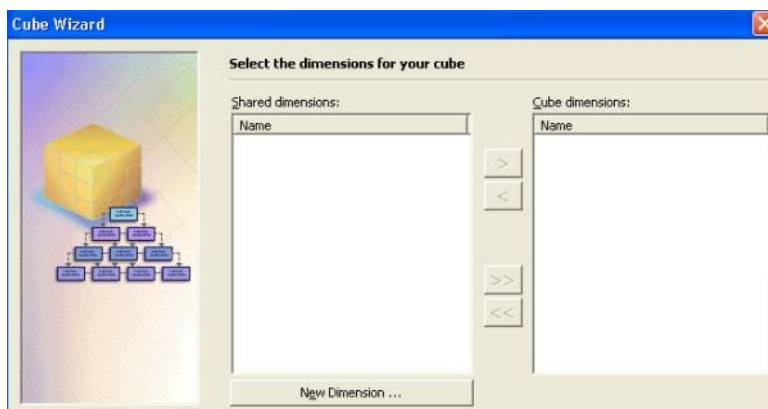
Cube measures:

Measure name	Source column
Store Sales	store_sales
Store Cost	store_cost
Unit Sales	unit_sales

- Pour définir les mesures pour votre cube, sous " Colonnes numériques de la table de fait ", double-cliquez sur store_sales. Faire de même pour les colonnes store_cost et unit_sales représentant respectivement le nombre de vente, le coût de la vente et le prix unitaire.Cliquez ensuite sur " suivant ".

III-H - Construire la dimension Temps

Dans les bases multidimensionnelles la dimension Temps est généralement utilisé. D'ailleurs dans des SGBD comme DB2, elle est même imposée. Pour construire cette dimension faire :



Select the dimensions for your cube

Shared dimensions:

Name

Cube dimensions:

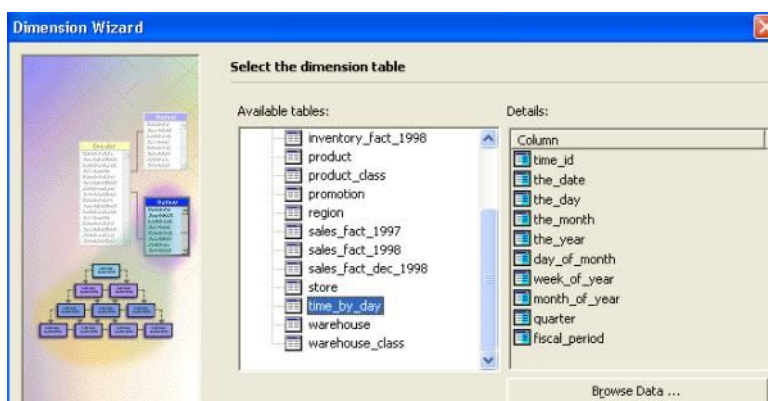
Name

New Dimension ...

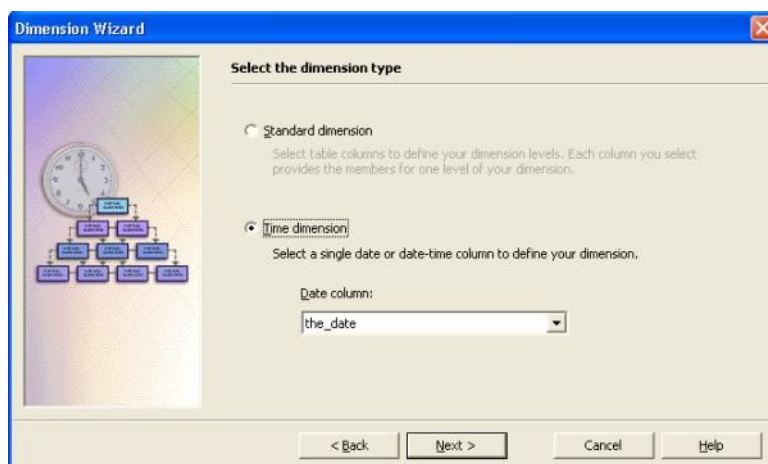
- Dans la boîte de dialogue " Sélectionnez les dimensions pour votre cube ", cliquez sur " Nouvelle Dimension ".



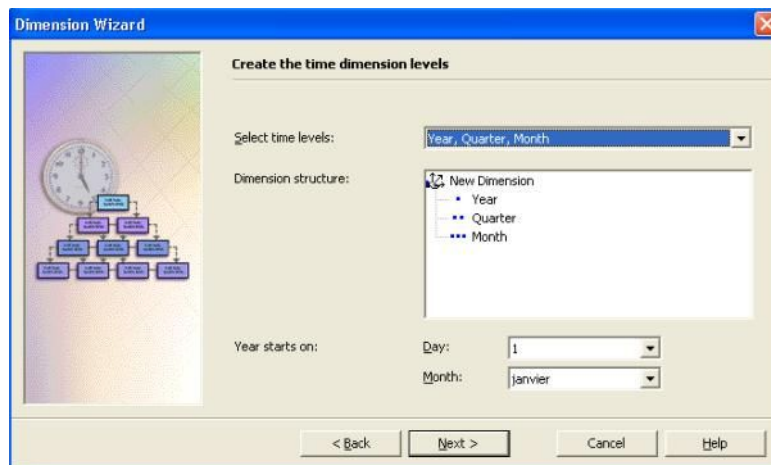
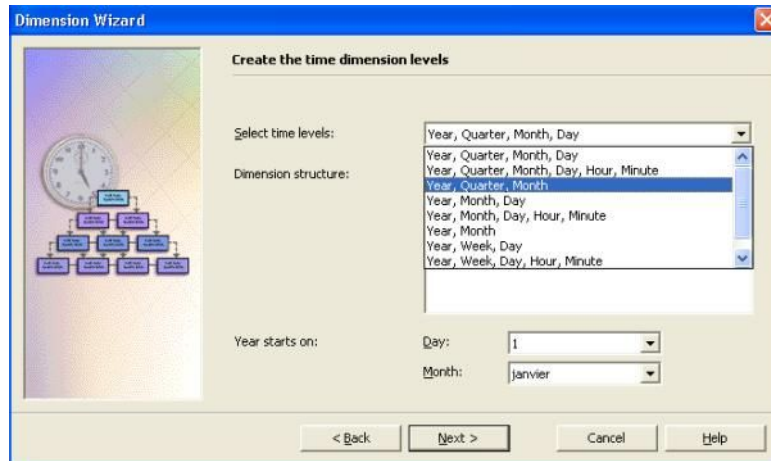
- Dans l'assistant de création de dimension, sélectionner " Une table de dimension simple" et cliquer sur " suivant ".



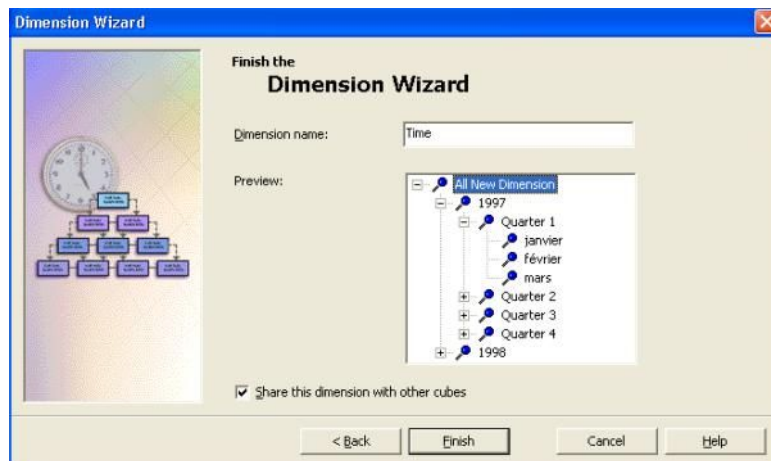
- Sur l'écran " sélectionner la table de dimension ",dérouler " ventes " et cliquer sur " time_by_day ".Vous pouvez voir le contenu de cette table en cliquant sur " parcourir les données ".On voit qu'une date fait partie d'une semaine, d'un mois d'un trimestre etc □
- Cliquer sur " suivant ".



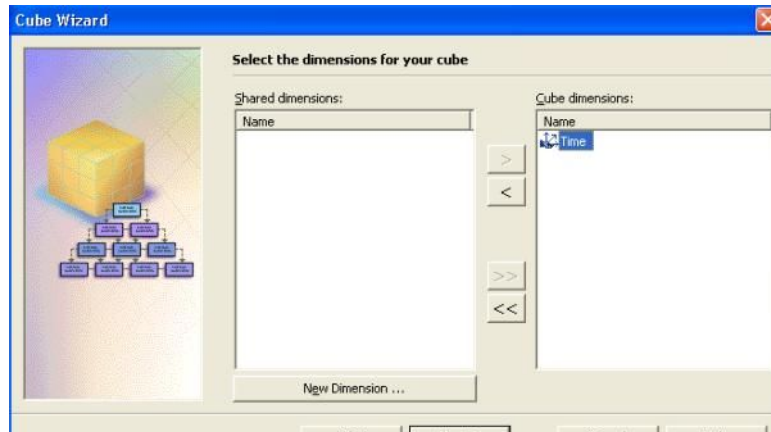
- Dans l'écran " sélectionner le type de dimension ", choisir " Dimension temps " puis cliquer sur " suivant ".



- Pour définir les niveaux de hiérarchie de votre dimension, cliquez sur " Choisir les hiérarchies du temps " et choisir Année, Trimestre, Mois comme niveau de hiérarchie et cliquer sur " suivant ".



- Donnez un nom à votre dimension. Ici je lui donne le nom " Time ". Ensuite cliquer sur " Terminer " pour retourner à l'assistant " création de cube ". Ici on voit bien que l'année 1997 a été décomposée en 4 trimestres et que le premier trimestre contient les mois de janvier, février et mars.

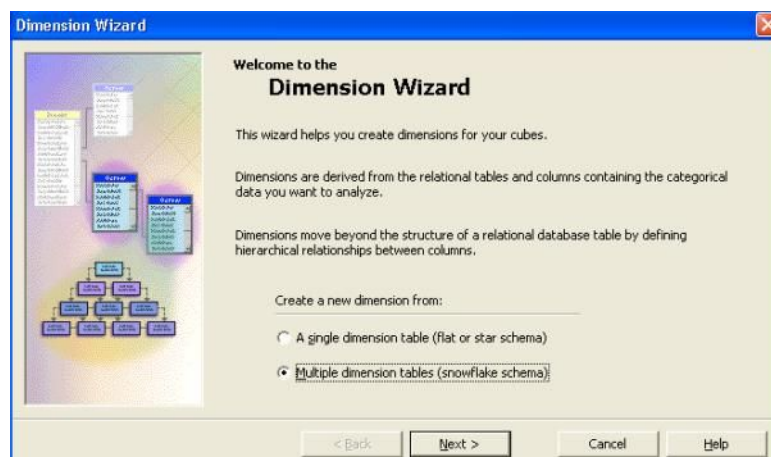


- Vous pouvez ainsi voir votre nouvelle dimension " Time " dans la liste des dimension de votre cube.

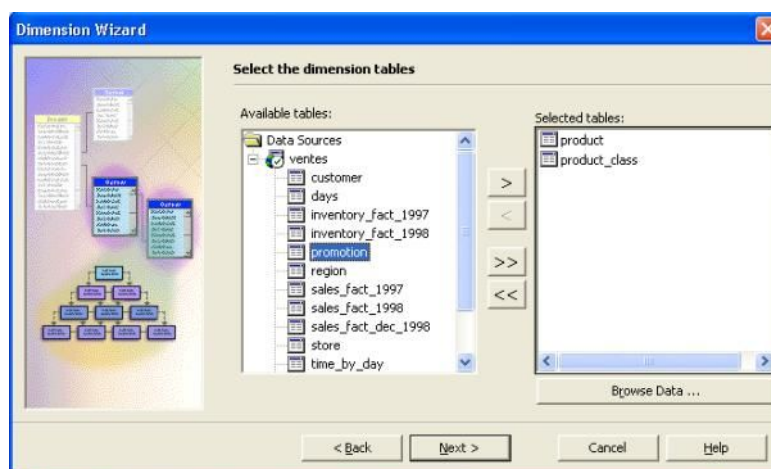
III-I - Construire la dimension Produit

Pour construire la dimension Produit :

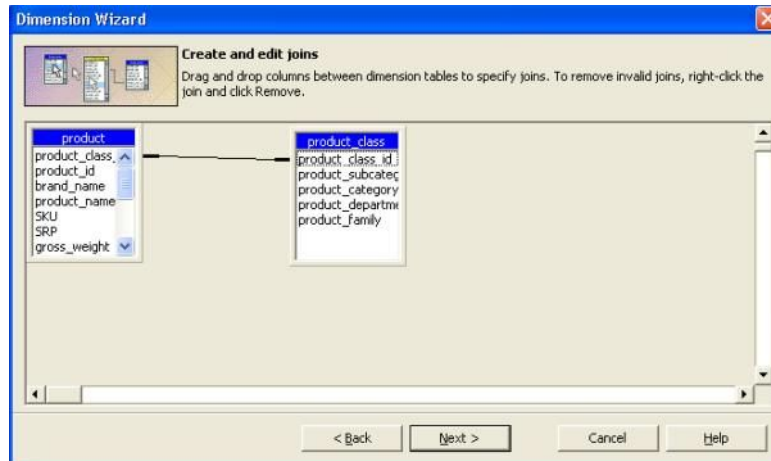
- Cliquer sur " Nouvelle dimension "



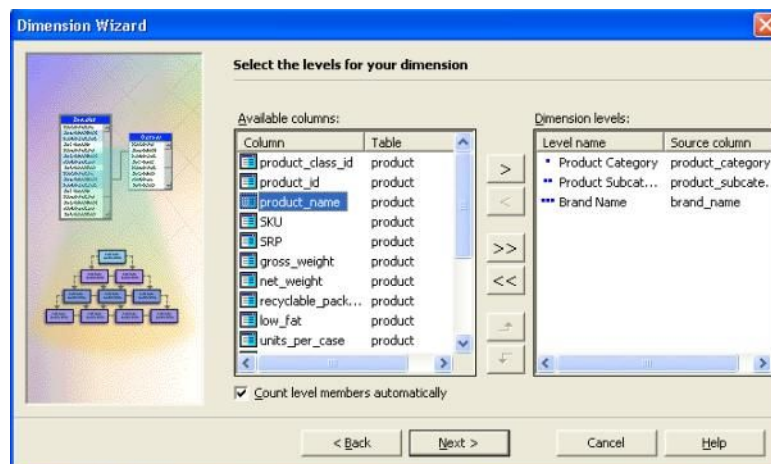
- Dans l'assistant de création de dimension, cliquer sur " tables de dimension multiples " et ensuite cliquer sur " suivant ".



- Sur l'écran " Sélectionner les tables de dimensions ", dérouler " ventes " et double-cliquer sur product et product_class pour les ajouter dans " tables sélectionnées ". Cliquez ensuite sur " suivant ".



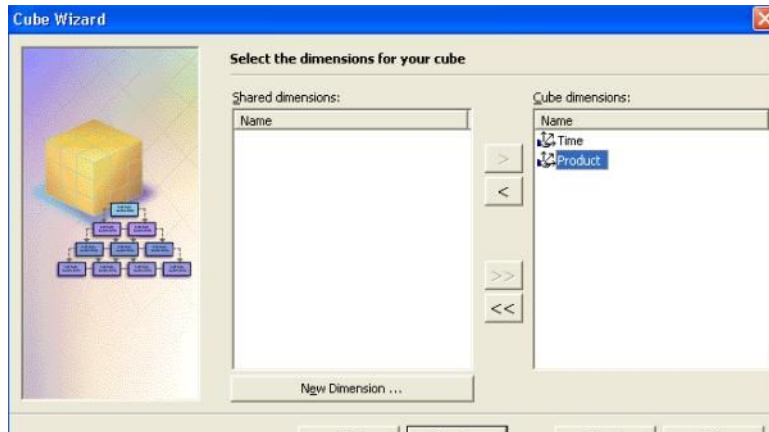
- Vous pouvez voir les deux tables sélectionnées et la jointure qui les lie. Ensuite cliquer sur " suivant ".



- Pour définir des niveaux de hiérarchie pour votre dimension, sous colonnes disponibles, double-cliquer dans cet ordre sur product_category, product_subcategory et brand_name (représentant respectivement la catégorie de produit, la sous-catégorie de produit et le nom de marque du produit). On voit ainsi le nom des niveaux de hiérarchie apparaître. Cliquez sur suivant.



- Donner le nom " Product " à la nouvelle dimension ainsi créée et laissez la case " Partagez cette dimension avec le autres cubes " cochée. Cliquez sur " Terminer ".

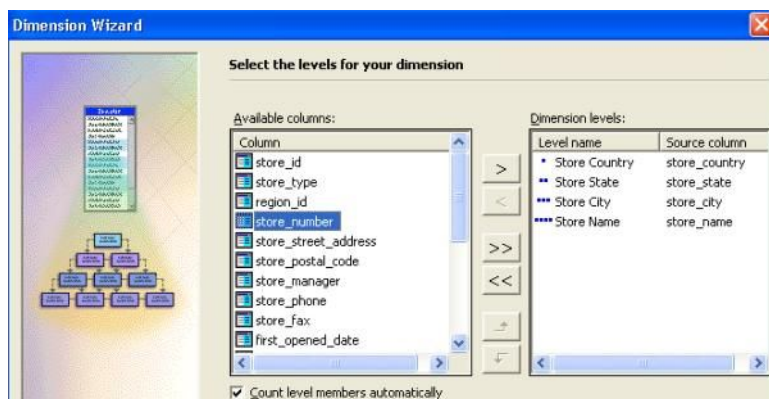


- vous pouvez alors voir la dimension " Product " dans la liste des dimensions du cube.

III-J - Construire la dimension Magasin

Pour construire la dimension Magasin :

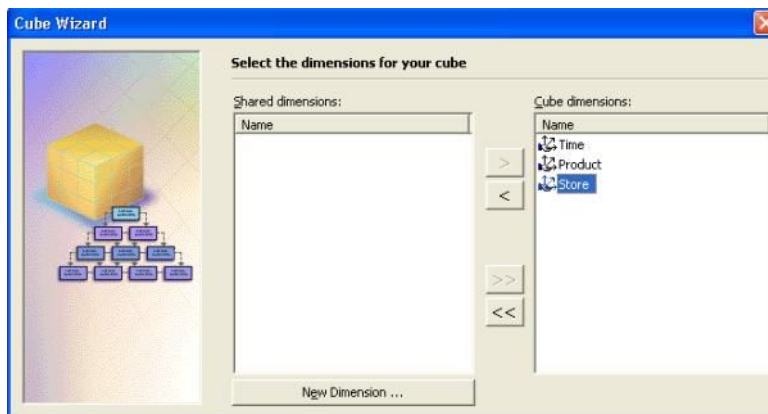
- Cliquer sur " Nouvelle dimension "
- Dans l'assistant de création de dimension, cliquer sur " table de dimension simple " et ensuite cliquer sur " suivant ".
- Sur l'écran " Sélectionner la table de dimension ", dérouler " ventes " et cliquer sur store. Cliquez ensuite sur " suivant ".
- Sur l'écran " sélectionnez le type de dimension ", cliquez sur " suivant ".



- Pour définir des niveaux de hiérarchie pour votre dimension, sous colonnes disponibles, double-cliquer dans cet ordre sur store_country, store_state, store_city et store_name (représentant respectivement le pays du magasin, l'état du magasin, la ville du magasin et le nom du magasin). On voit ainsi le nom des niveau de hiérarchie apparaître. Cliquez sur suivant.



- Donner le nom " Store " à la nouvelle dimension ainsi créée et laissez la case " Partagez cette dimension avec les autres cubes " cochée. Cliquez sur " Terminer ".

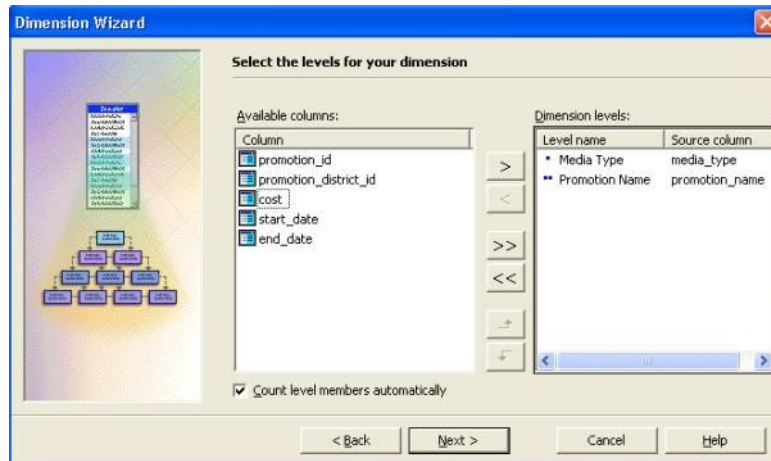


- Vous pouvez alors voir la dimension " Store " dans la liste des dimensions du cube.

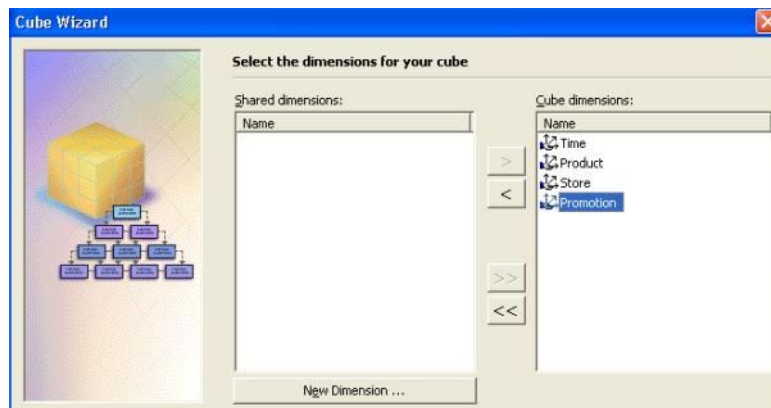
III-K - Construire la dimension Promotion

Pour construire la dimension promotion :

- Cliquer sur " Nouvelle dimension "
- Dans l'assistant de création de dimension, cliquer sur " table de dimension simple " et ensuite cliquer sur " suivant ".
- Sur l'écran " Sélectionner la table de dimension ", dérouler " ventes " et cliquer sur promotion. Cliquez ensuite sur " suivant ".
- Sur l'écran " sélectionnez le type de dimension ", cliquez sur " suivant ".



- Pour définir des niveaux de hiérarchie pour votre dimension, sous colonnes disponibles, double-cliquer dans cet ordre sur media_type, promotion_name (représentant respectivement le type de média et le nom de la promotion). On voit ainsi le nom des niveaux de hiérarchie apparaître. Cliquez sur suivant.
- Donner le nom "Promotion" à la nouvelle dimension ainsi créée et laissez la case "Partagez cette dimension avec les autres cubes" cochée. Cliquez sur "Terminer".



- Vous pouvez alors voir la dimension "Promotion" dans la liste des dimensions du cube.

III-L - Terminer la création du cube

Pour terminer la création du cube faire :

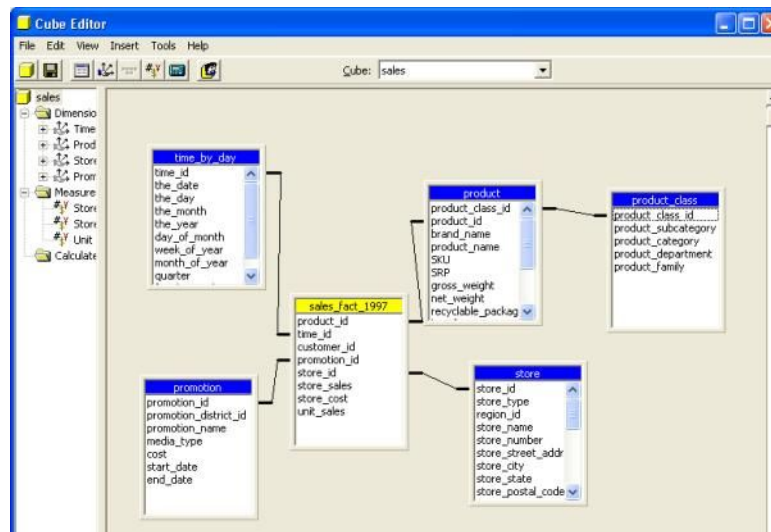
- Cliquer sur "suivant"



- Donner un nom à votre cube (ici " sales ") et cliquer sur " terminer ".

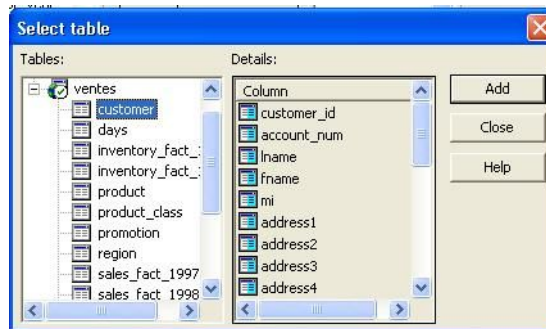
III-M - Editer le cube dans l'éditeur de cube

Dans le panneau schéma de l'éditeur de cube, vous pouvez voir la table de fait (avec sa barre de titre jaune) et les tables de dimension (avec leur barre de titre bleue). De plus, dans le panneau de gauche, vous pouvez voir la structure du cube. Vous pouvez éditer les propriétés du cube en cliquant sur le bouton " propriété ".

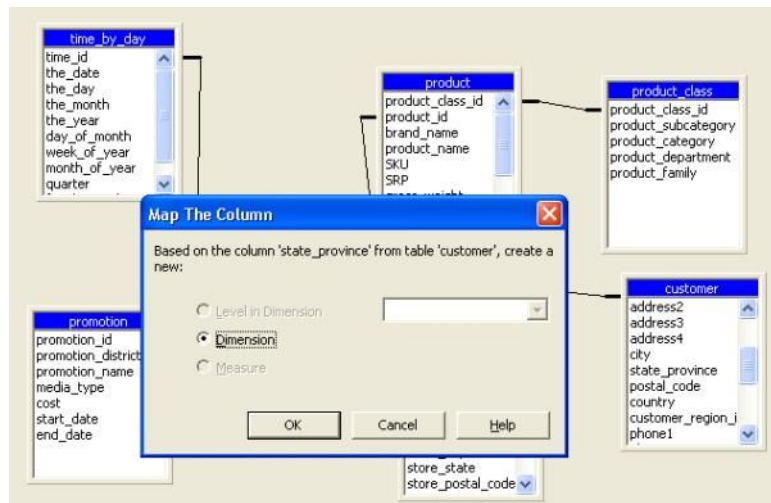


Supposons que maintenant vous ayez besoin d'une autre dimension qui vous donne des informations sur les clients. Vous pouvez facilement créer cette dimension. Cependant, les dimensions créées dans l'éditeur de cube sont privées c'est à dire qu'elles ne peuvent être utilisées qu'avec le cube avec lequel vous travaillez. Elles ne peuvent donc être partagées avec d'autres cubes.

- Dans le menu " insertion " de l'éditeur de cube, choisir " tables ".



- Dans la boîte de dialogue " Sélectionner une table ", dérouler la source de donnée " ventes ", double-cliquer sur la table " customer " représentant les clients ensuite cliquer sur " fermer ".
- Pour définir la nouvelle dimension, double-cliquer sur la colonne " state_province " de la table " customer ".



- Dans la boîte de dialogue Mapper la colonne, choisir " Dimension " puis cliquer sur " OK "
- Sélectionner la dimension " State Province " de l'arborescence.
- Choisir l'item " Rename " du menu " Edit "



- Donner le nom " Customer " puis appuyer sur " Entrée ".



- Faire un drag and drop de la colonne " city " de la table " customer " du panneau de schéma vers la nouvelle dimension renommée en " Customer " dans le menu gauche de l'éditeur de cube.
- Dérouler " Customer " pour voir les deux niveaux de hiérarchie créés pour la dimension.

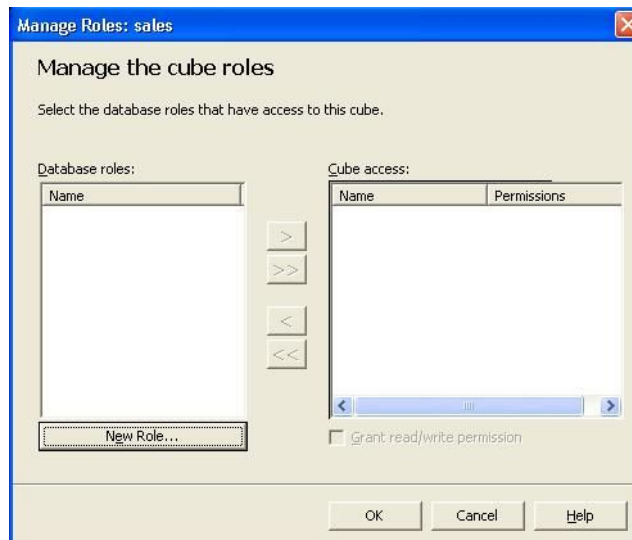
III-N - Ajouter un rôle au cube

Les rôles de cube définissent quels utilisateurs ou groupe d'utilisateurs ont accès au cube et peuvent y requêter. Maintenant que votre cube est totalement construit, vous allons ajouter un rôle au cube. Dans cet exemple, nous ajouterons le rôle " marketing ".

NB : Le DBA lui n'a pas besoin d'avoir des droits pour requêter sur le cube via OLAP manager. Seuls les utilisateurs qui utilisent un outil client (Excel, Business Objects etc.) sont concernés par les rôles définis.

Pour créer un nouveau rôle pour le cube faire :

- Dans l'éditeur de cube, choisir " Gérer les rôles " du menu " Outils "



- dans la boîte de dialogue " Rôle de cube ", cliquez sur " Nouveau rôle ".
- Dans la boîte de dialogue " Créer un rôle de base de données ", taper " marketing " dans " Nom du rôle " (c'est le nom que nous donnerons à notre rôle pour permettre aux analystes marketing de pouvoir interroger les données).



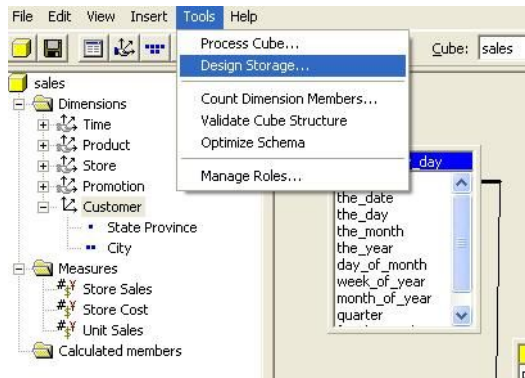
- Dans la partie " Utilisateurs et groupes ", mettre les utilisateurs ou groupes d'utilisateurs du réseau qui auront accès.
- Dans la boîte de dialogue " Rôles du cube ", le rôle " Marketing " apparaît dans la liste " Accès au cube ". Cliquez sur Ok.

- Dans l'éditeur de cube, choisir " enregistrer " à partir du menu " Fichier ".

III-O - Concevoir le type de stockage et traiter le cube

Les services OLAP permettent de choisir un type d'agrégations adéquat. Le choix du type d'agrégation est important car il influe beaucoup sur les temps de réponse des requêtes. Pour optimiser les performances de traitement des requêtes de votre cube, il faut utiliser " l'assistant design du stockage ".

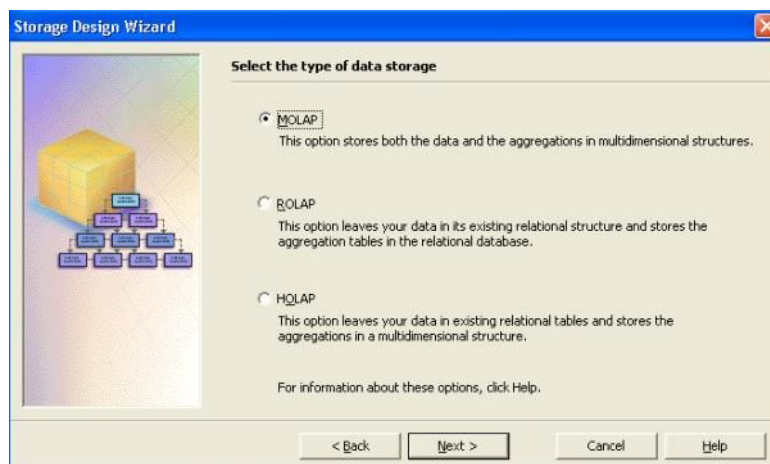
Pour démarrer " l'assistant design du stockage " faire :



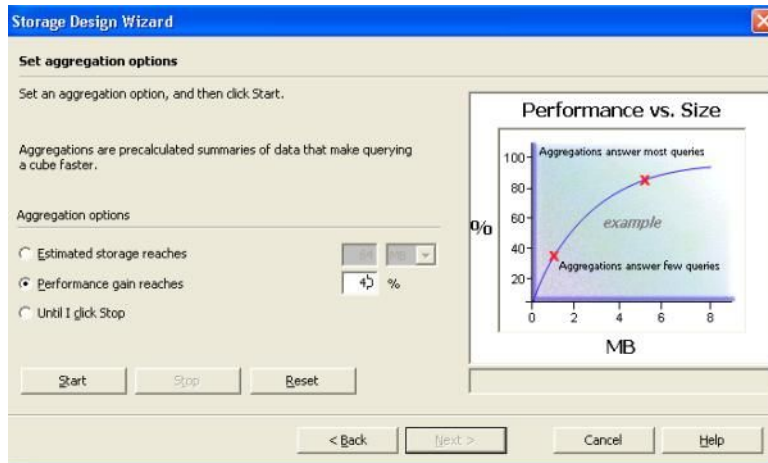
- A partir du menu " outils " de l'éditeur de cube, choisir " Design du stockage ".



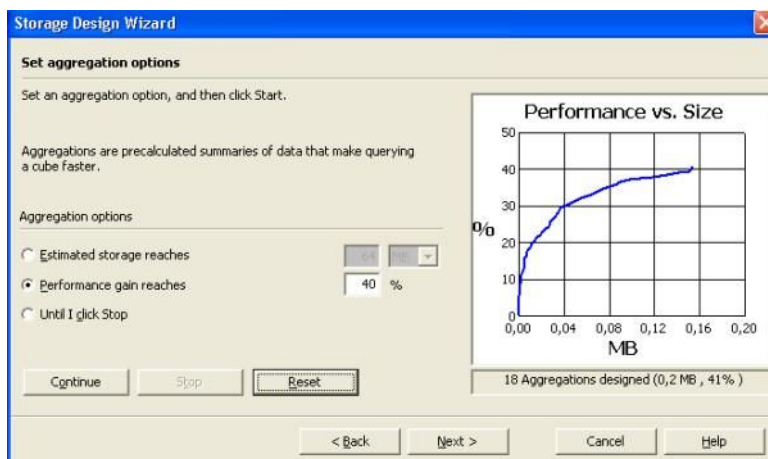
- Sur l'écran " Design du stockage ", cliquez sur suivant.



- Sélectionner le type de stockage MOLAP puis cliquez sur " suivant ". MOLAP est le type de stockage multidimensionnel qui permet de stocker les agrégations dans le moteur Analysis Service contrairement au ROLAP (Relational OLAP) où les données sont agrégées dans le moteur relationnel même. Pour un stockage de type MOLAP l'interrogation des données est rapide mais perd de son efficacité lorsque le volume de stockage devient très important.



- Sous " options d'agrégation ", sélectionner " gain de performance atteint ". Choisir alors un gain de performance de 40%.



- Ainsi donc on dit au moteur OLAP d'atteindre une performance de 40% même si on sait pas combien d'espace disque cela nécessiterait. Un DBA analyserait donc l'écran suivant pour voir l'impact que cela aurait sur l'espace disque requis. Cliquez donc sur " Démarrer ".
- Vous pouvez donc voir dans l'écran le rapport performance/espace. Ainsi on voit bien que plus on veut de la performance, plus on prend de la place. A la fin du traitement, faire " suivant ".



- Dans " que voulez-vous faire ? ", choisir " traiter maintenant ".
- Lorsque le traitement est terminé, un message apparaît signalant que " Le traitement s'est terminé avec succès ". Ensuite cliquer sur " fermer " pour retourner à l'éditeur de cube.
- A partir du menu fichier, choisir " quitter " pour fermer l'éditeur de cube et retourner à l'arborescence de l'OLAP manager.

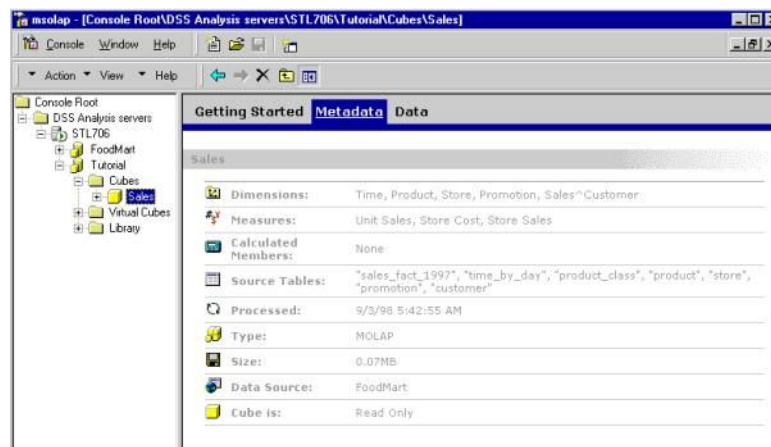
NB : Traiter les agrégations risque de prendre du temps.

III-P - Visualiser les métadonnées du cube

Les services OLAP permettent de voir les métadonnées du cube c'est à dire les informations détaillées de la configuration du cube de données. Ces informations apparaissent dans le panneau droit du OLAP manager.

Pour visualiser les métadonnées du cube " sales " faire :

- Dans l'arborescence de gauche du OLAP manager, dérouler " Cubes "
- Choisir le cube " Sales "
- Dans le panneau de droite du OLAP manager, cliquer sur " Metadata ".

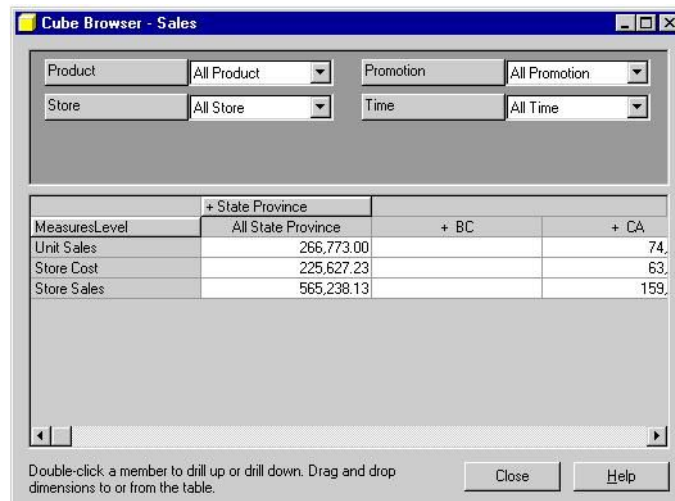


III-Q - Naviguer sur les données du cube

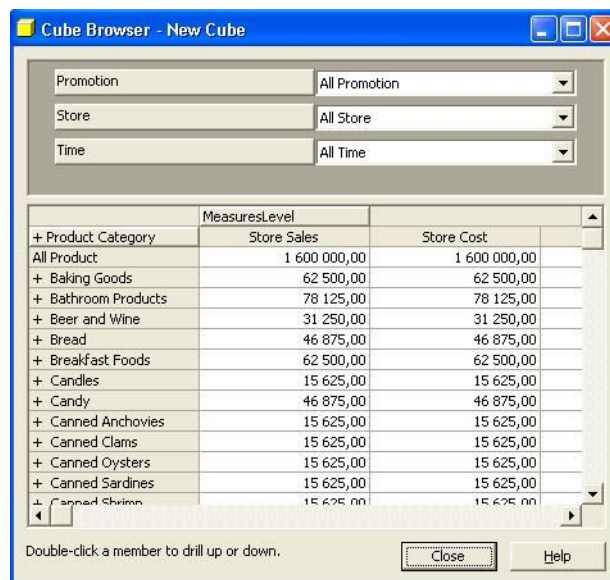
Maintenant, vous êtes prêt à naviguer sur les données de votre cube " Sales ". En utilisant le browser de cube, vous pouvez voir les données suivant différents axes, vous pouvez aussi filtrer la quantité de données visible des dimensions et vous pouvez aussi faire des " drill-down " pour voir les données à un niveau de détail plus fin (ou faire un " roll-up " pour voir les données agrégées). Pour naviguer faire :

- Dans l'arborescence du OLAP manager, faire un click-droit sur le cube " Sales " et choisir " Naviguer sur les données ".

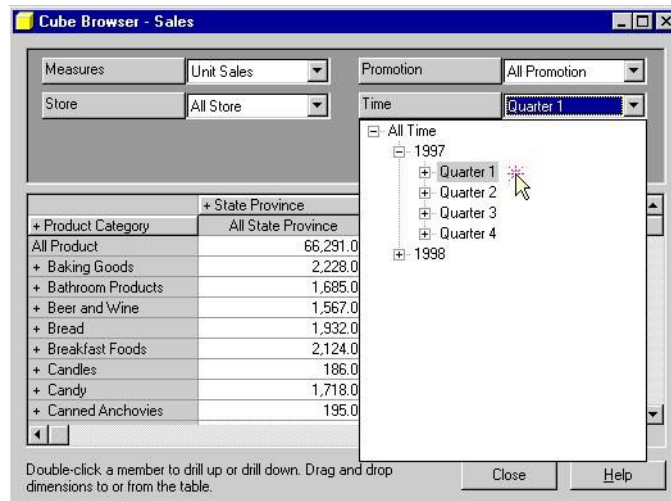
Ainsi le navigateur de cube apparaît faisant apparaître une partie supérieure contenant des dimensions et une partie inférieure contenant une grille avec les mesures et une dimension. Dans notre cas comme illustré dans la capture d'écran ci-dessous, on voit les quatre dimensions dans la partie de haut et une dimension et les mesures dans la partie basse. Pour remplacer une dimension par une autre, il suffit de faire un drag de la dimension à partir de la partie haute et faire le drop sur la partie basse.



Selectionnez donc la dimension Product faite un drag and drop sur la grille (partie basse) en faisant le drop sur là où se trouve les mesures (MeasuresLevel). Vous pouvez voir alors que product et MeasuresLevel sont intervertis comme sur l'écran suivant :



Maintenant, essayons de filtrer les données par date. Vous pouvez alors Cliquer sur la combo box de la dimension Time et vous avez la possibilité de voir toutes les données de l'année 1997 ou de voir par exemple les données du deuxième trimestre seulement de cette même année.



Vous pouvez aussi faire un "drill-down" des données, c'est à dire avoir un niveau de détail des données. Par exemple, pour la dimension Product, dérouler le produit "Baking Goods" (pâtisserie) de la grille, vous pouvez donc voir pour cette catégorie de produit, ses sous-catégories et voir les ventes pour voir les sous-catégories.



IV - Conclusion

Dans cet article, j'ai juste donné un aspect global d'un système décisionnel qui est à la base du Business intelligence (B.I). Car avant de se lancer dans la panoplie d'outils qui font du décisionnel, sans doute faudrait-il d'abord connaître son mode de fonctionnement et avoir un aperçu pratique avec Analysis Services. Cependant, il serait intéressant de traiter les outils ETL, les états Business Objects ou encore les APIs de programmation multidimensionnelles (ADOMD, javax.jolap etc).

V - Remerciements

Je remercie Fadace pour sa relecture de l'article.