

# Créer et utiliser simplement un cube brut

Dans les sections précédentes, nous avons construit l'entrepôt de données sous SQL Server, puis nous avons chargé un jeu de test cohérent.

À présent, nous allons nous atteler à restituer l'information contenue dans l'entrepôt de données de manière simple, présentable et rapide.

Simple et présentable : cela signifie que l'utilisateur qui accède à l'information ne doit pas voir la complexité du traitement de l'information. Le fait que l'information provienne de multiples sources doit être totalement transparent. Le fait que la restitution de l'information suggère de nombreuses règles de gestion métier doit lui aussi être transparent. Un modèle réussi est un modèle que l'on peut donner à l'utilisateur sans avoir crainte qu'il ne se trompe, qu'il interprète mal les résultats ou qu'il se perde dans l'étendue de l'information.

Rapide signifie qu'un utilisateur, qui accède à des informations mises à disposition, doit avoir un temps d'attente de l'ordre de la seconde. Attendre 10 secondes peut déjà être considéré comme long.

➤ La base de données Analysis Services avec sa technologie multidimensionnelle peut nous aider à répondre à ces trois critères. Néanmoins la solution n'est pas magique. Ce n'est pas le fait d'employer la technologie qui va faire que l'information va être simple, rapide et présentable. L'outil va simplement nous donner les moyens de nos ambitions.

La modélisation que nous venons de mettre en œuvre dans les sections précédentes en est l'illustration. La bonne modélisation est la clé de l'utilisation optimale d'Analysis Services.

Nous allons voir ici qu'un entrepôt de données bien modélisé, c'est 80 % du travail accompli. Les 20 % restants relèvent de la finition.

➤ SQL Services Analysis Services (SSAS) est le service de bases de données multidimensionnelles. Attention, Analysis Services n'est pas un outil de restitution de données, mais une base de données orientée vers l'utilisateur.

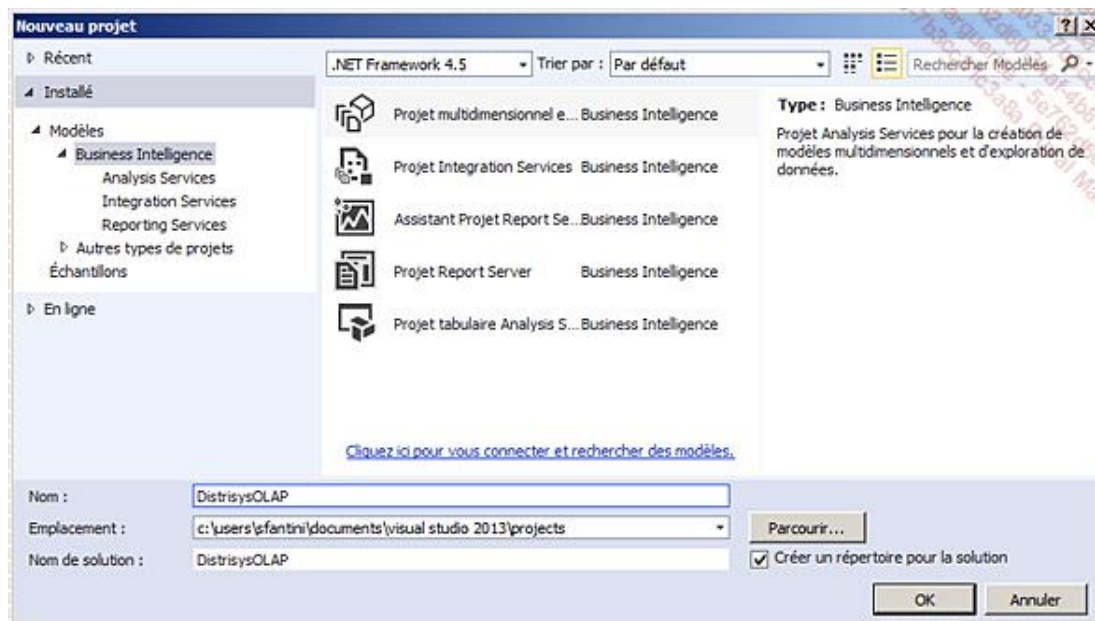
Dans les sections qui vont suivre, quand nous évoquerons SSAS, la base OLAP, la base multidimensionnelle, le cube ou le modèle, nous sous-entendrons toujours *la base de données Analysis Services*.

La construction du cube Analysis Services se réalise avec l'outil SQL Server Data Tools (ou SSDT) alors que la gestion et l'administration des modèles se réalisent avec l'outil SQL Server Management Studio (ou SSMS).

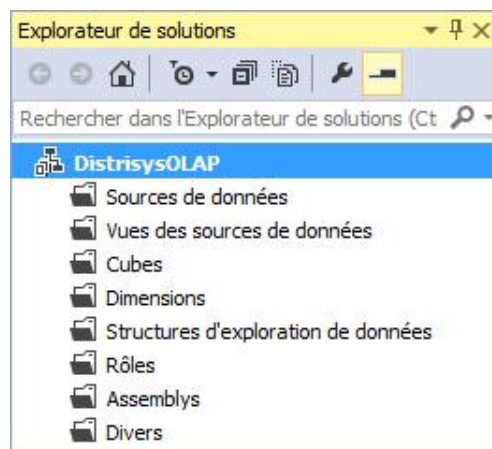
Pour commencer, nous allons créer un nouveau projet Analysis Services sous SQL Server Data Tools.

➔ Ouvrez SSDT.

➔ Créez un nouveau projet Analysis Services :



Un projet Analysis Services se présente ainsi :



Lorsque l'entrepôt de données est créé et bien modélisé, créer un cube d'entreprise est assez simple. Cela se fait en trois étapes :

1. Création de la source de données
2. Création de la vue de source de données (DSV)
3. Création du cube

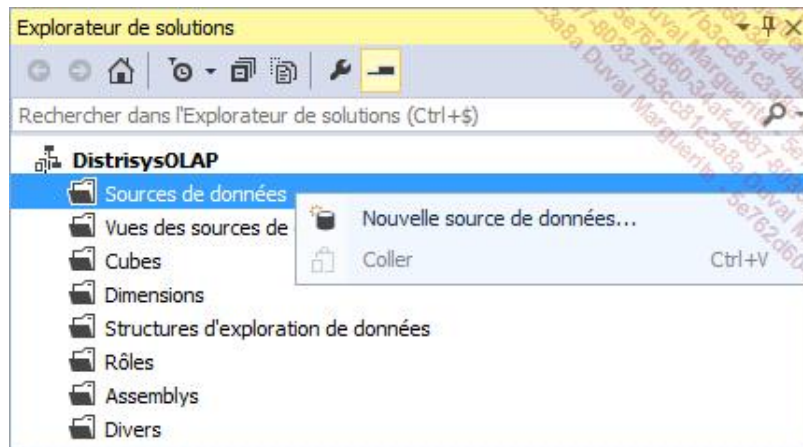
➤ Dans des conceptions avec SSDT, nous réaliserons surtout des cubes d'entreprise et, dans ce cas précis, nous retrouverons surtout un seul et grand cube. Les cubes d'entreprise sont des modèles reposant sur un entrepôt de données modélisé avec une approche définie par Ralph Kimball. Les cubes financiers ou de pilotage sont des cas particuliers nécessitant une approche et des outils de conception différents et complémentaires à l'approche évoquée par ce livre. Pour plus de détails sur la modification et la création des cubes de pilotage, veuillez vous référer à l'ouvrage "Business Intelligence avec SharePoint Server" paru aux Éditions ENI.

Dans un environnement sain, les choses sont simples :

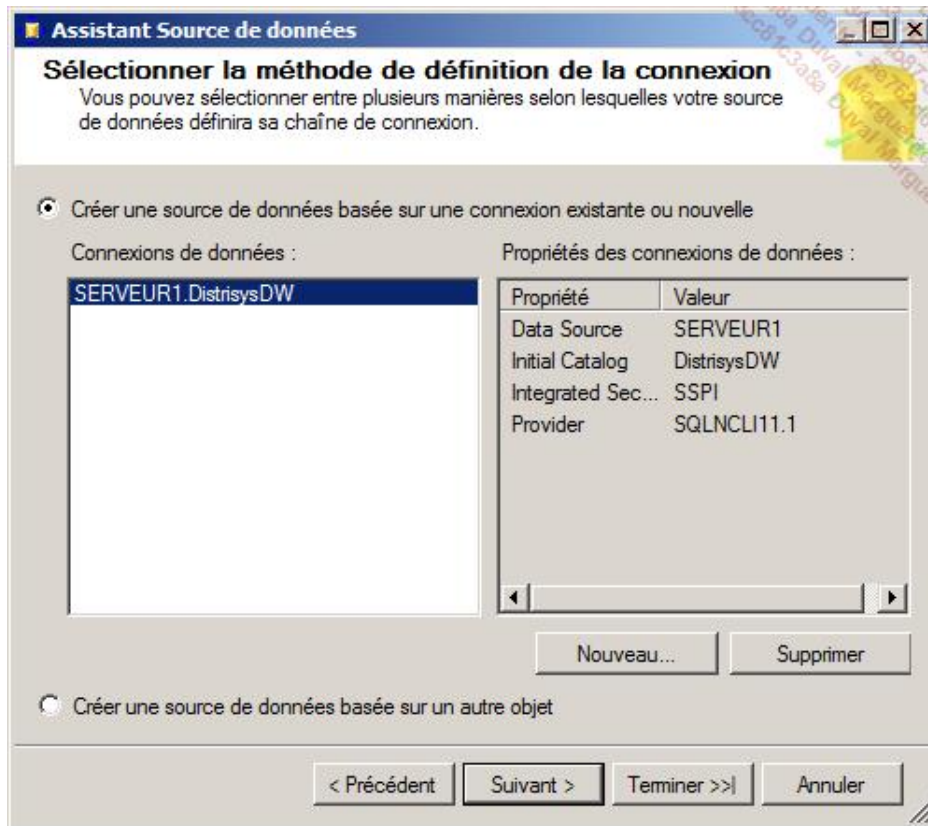
- Il n'y a qu'une seule source de données, il s'agit de l'entrepôt de données.
- Il n'y a qu'une seule vue de source de données et celle-ci est, autant que possible, le reflet de l'entrepôt de données.

Nous rappelons que le modèle de l'entrepôt de données doit être pensé, réfléchi et doit avoir une modélisation conforme. Toutes les transformations et manipulations de données devront être réalisées en amont lors de la phase d'ETL. La vue de sources de données ne doit pas être basée sur des vues ou basée sur des requêtes SQL, même si l'outil en donne la possibilité. Toute vue ou requête SQL présente dans un DSV est un symptôme d'une modélisation non conforme. Le DSV a une utilité dans des cas très précis qui relèvent de l'exception. Un de ces usages très particulier est illustré dans le chapitre La modélisation dimensionnelle - Facturation et commande client.

→ Créez la source de données :



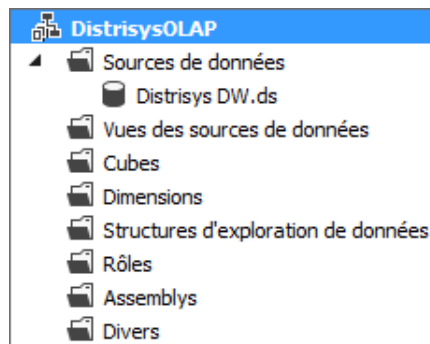
→ Sélectionnez l'entrepôt de données *DistrisysDW* créé précédemment, puis cliquez sur **Suivant** :



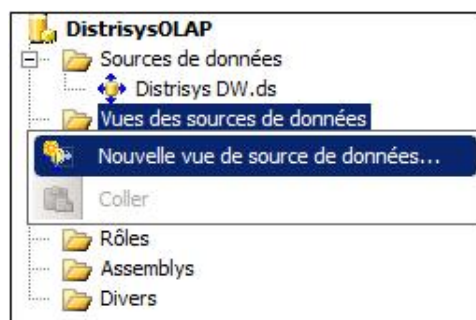
→ À la fenêtre de l'**Assistant Source de données**, sélectionnez **Utiliser le compte de service**, puis cliquez sur le bouton **Suivant**. Pour terminer l'assistant et créer la source de données, cliquez pour finir sur le bouton

**Terminer.**

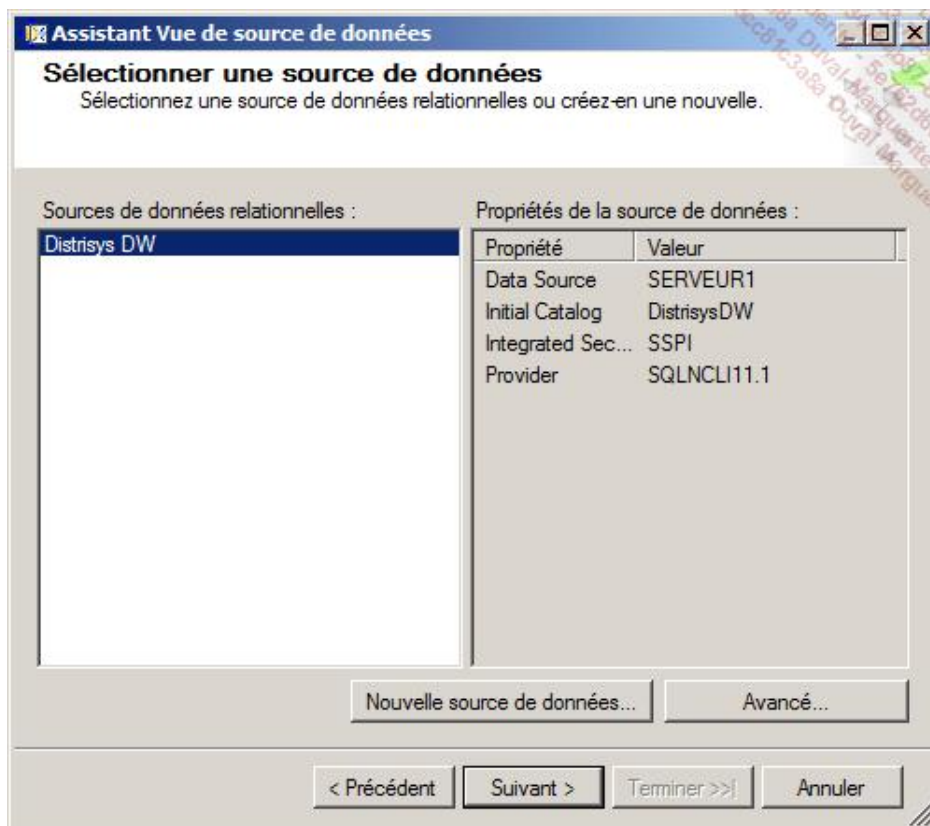
La source de données est créée :



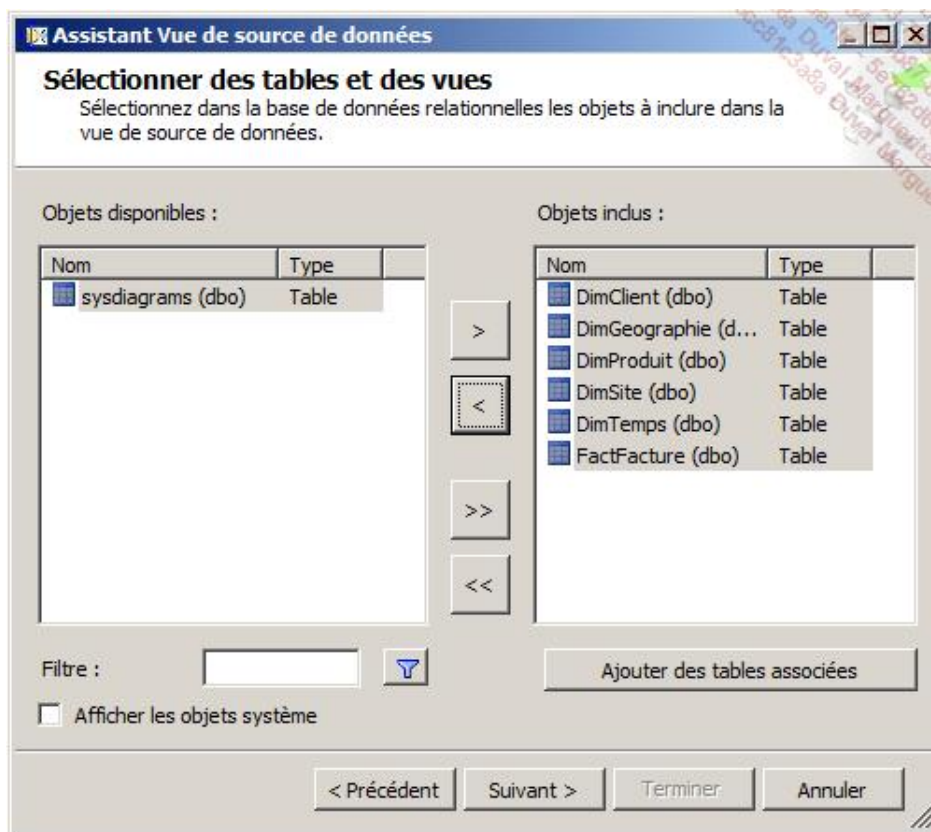
Maintenant créons la **Vue des sources de données** :



→ Sélectionnez la source de données précédemment créée :

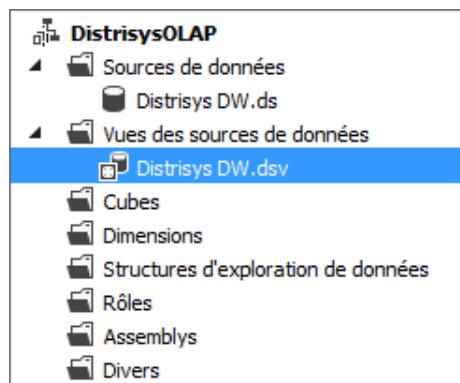


→ Sélectionnez toutes les tables préfixées **Dim** et **Fact** :



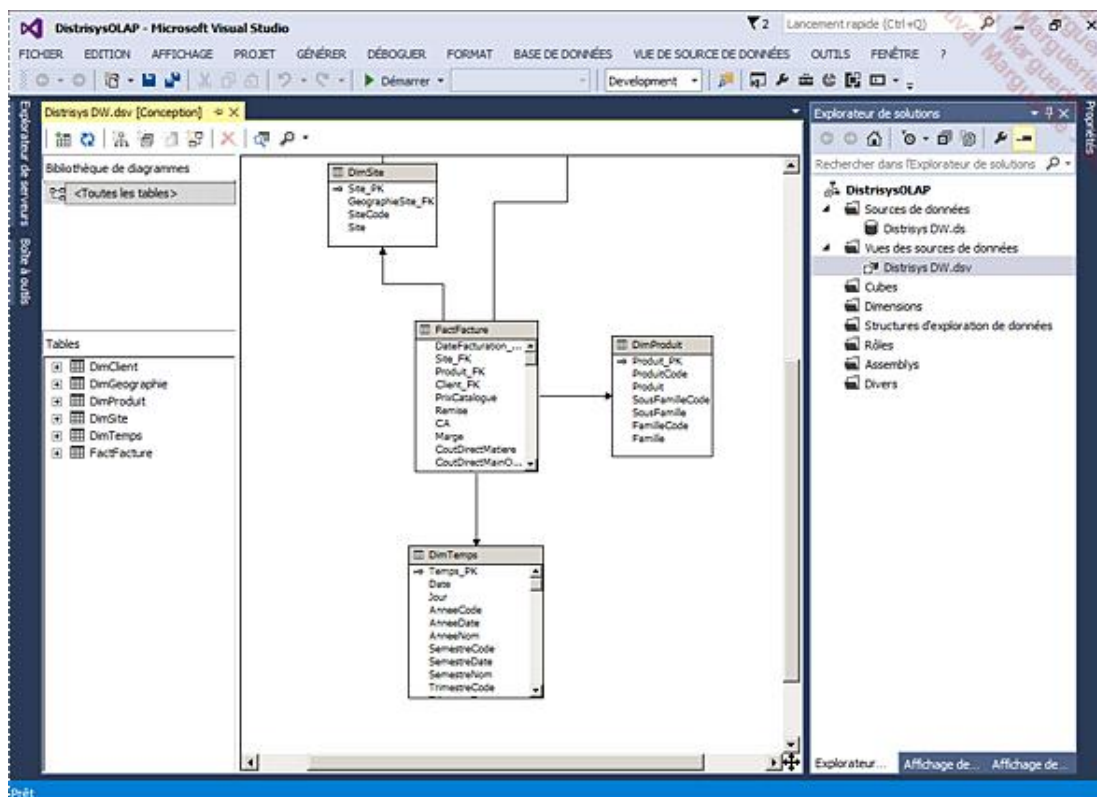
→ En fin d'assistant, cliquez sur **Terminer**.

La vue de sources de données est créée :



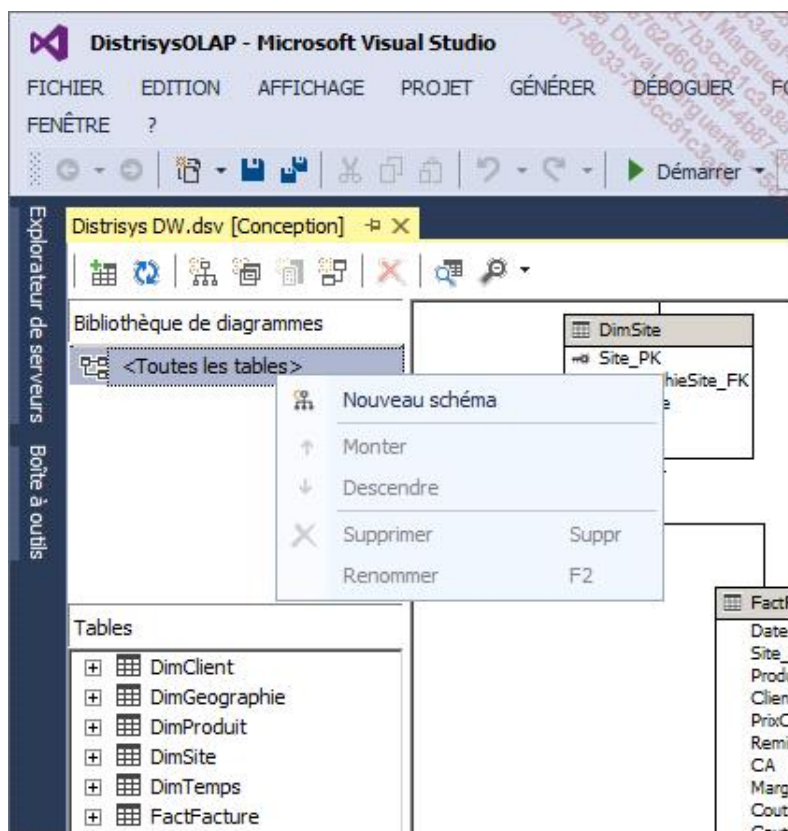
Les tables et leurs relations sont conservées :





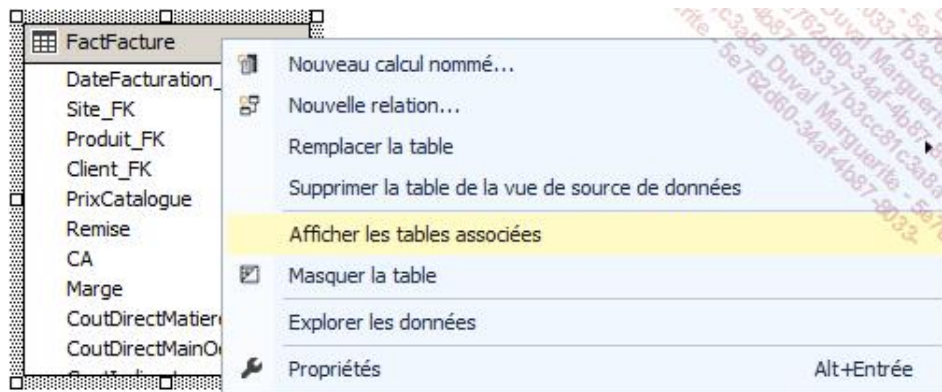
Le diagramme présenté est assez simple car nous n'avons qu'une seule table de faits et donc qu'une étoile. Pour prendre de bonnes habitudes, nous allons alors créer un nouveau diagramme *Facture Etoile*.

→ Effectuez un clic droit sur le diagramme par défaut pour créer un **Nouveau schéma** :

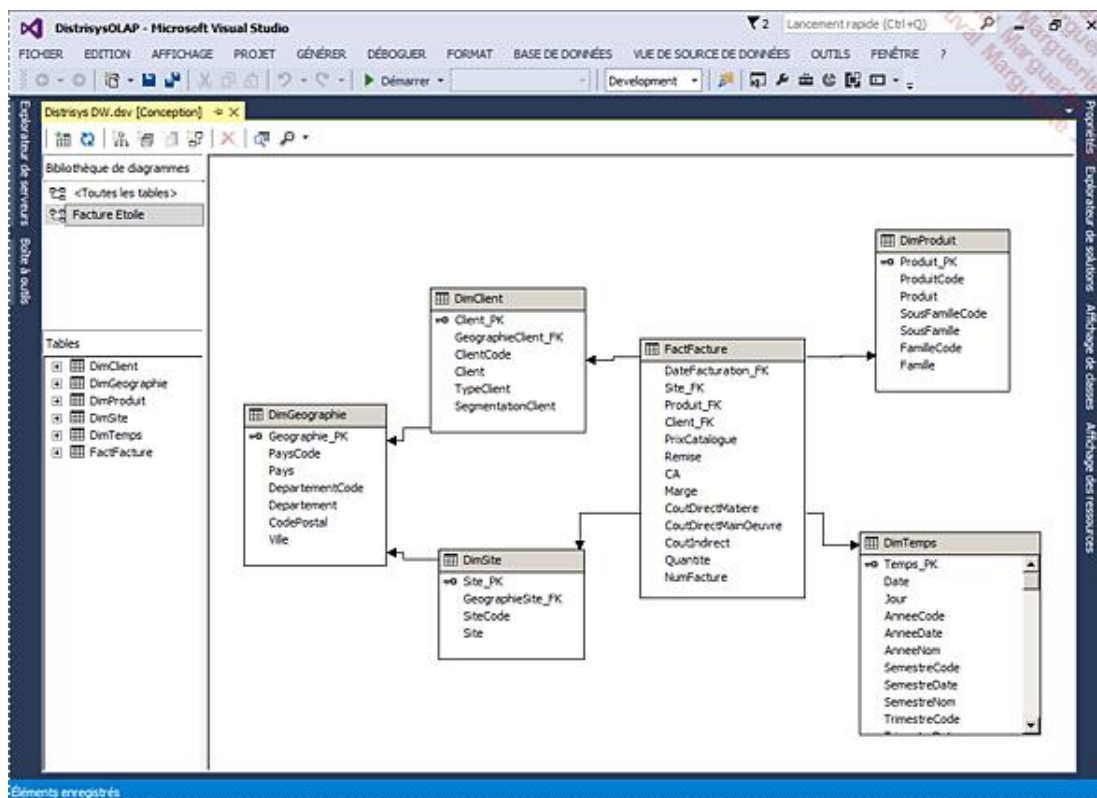


→ Renommez le nouveau diagramme **Facture Etoile**.

- Glissez la table de fait *FactFacture* dans l'espace de travail au centre, puis avec un clic droit sur la table, sélectionnez **Afficher les tables associées** :

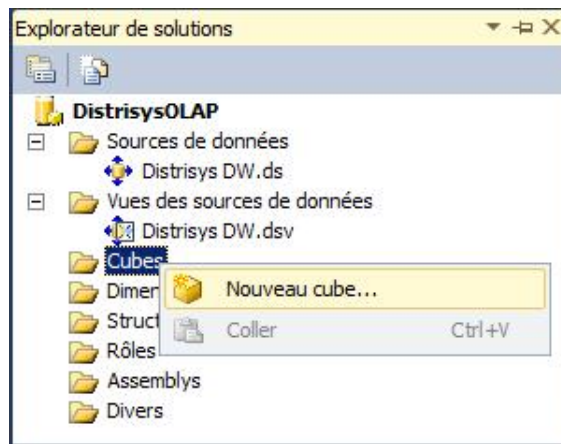


- Pour compléter l'étoile, glissez *DimGeographie* dans l'espace de travail et recréez l'étoile en vous assurant bien des liaisons entre les tables :

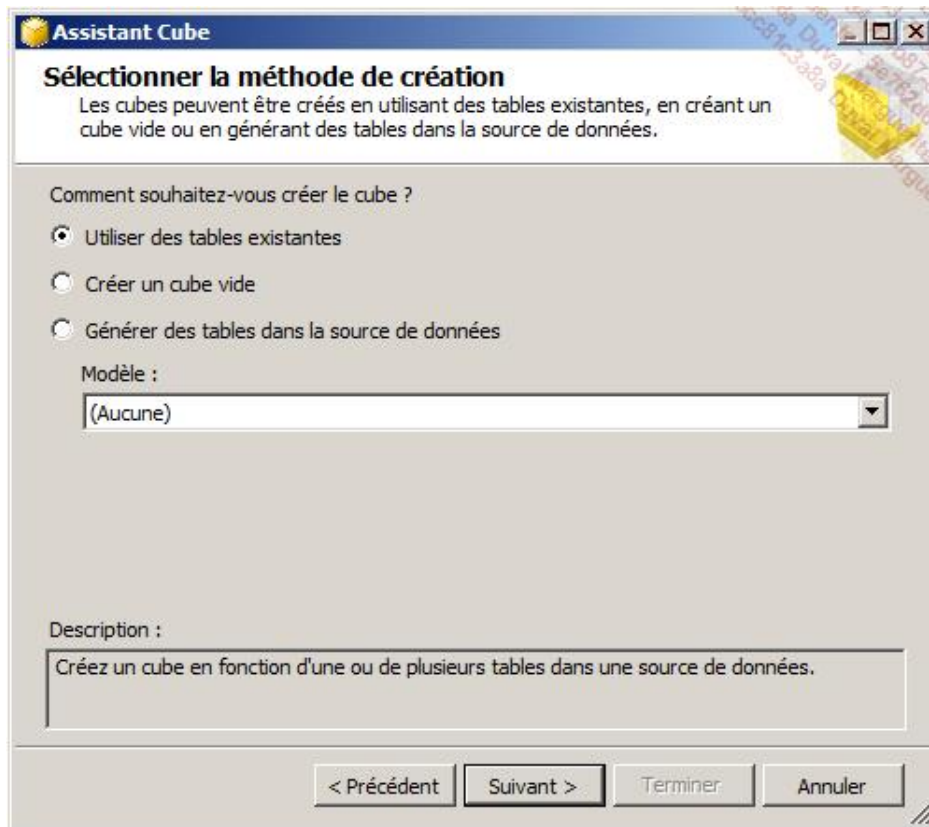


- Pour rappel, les diagrammes n'ont pas d'autre utilité que de vous permettre de bien vous assurer de la conformité de la modélisation.

- Enregistrez le projet.
- Créez le cube :



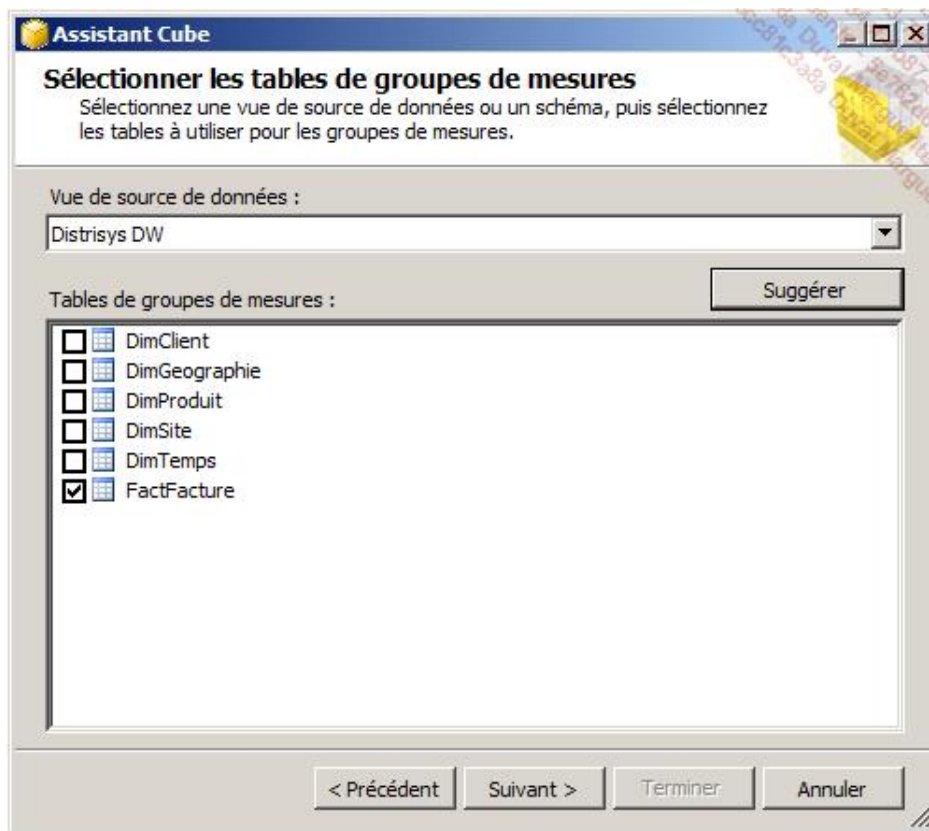
- Un assistant de création de cube s'ouvre. Dans la fenêtre **Sélectionner la méthode de création**, sélectionnez l'option **Utiliser des tables existantes**, puis cliquez sur le bouton **Suivant**.



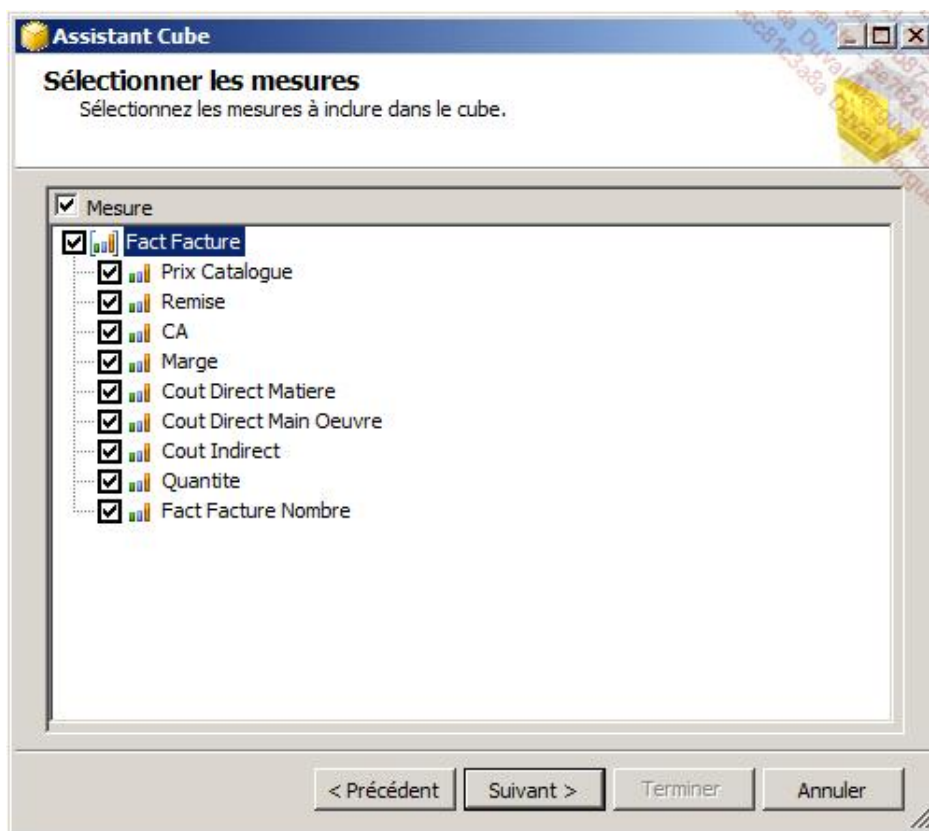
- Cliquez sur **Suggérer**.

L'assistant doit pouvoir trouver tout seul les tables de faits. Dans le cas contraire, nous vous conseillons de reprendre votre étoile et de vous assurer la cohérence du modèle :

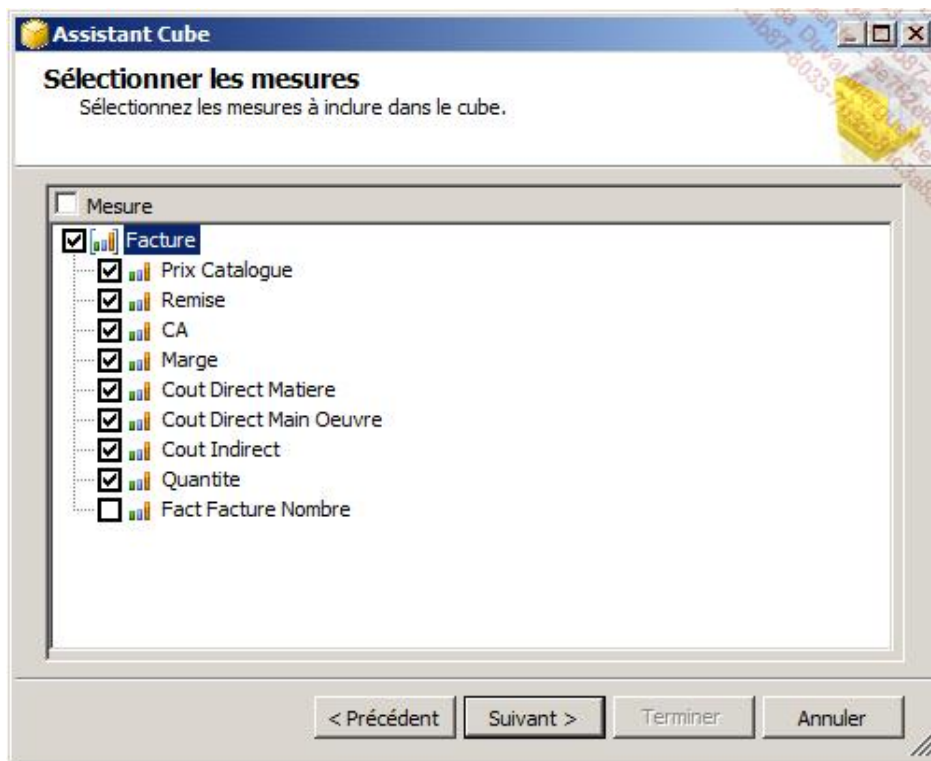




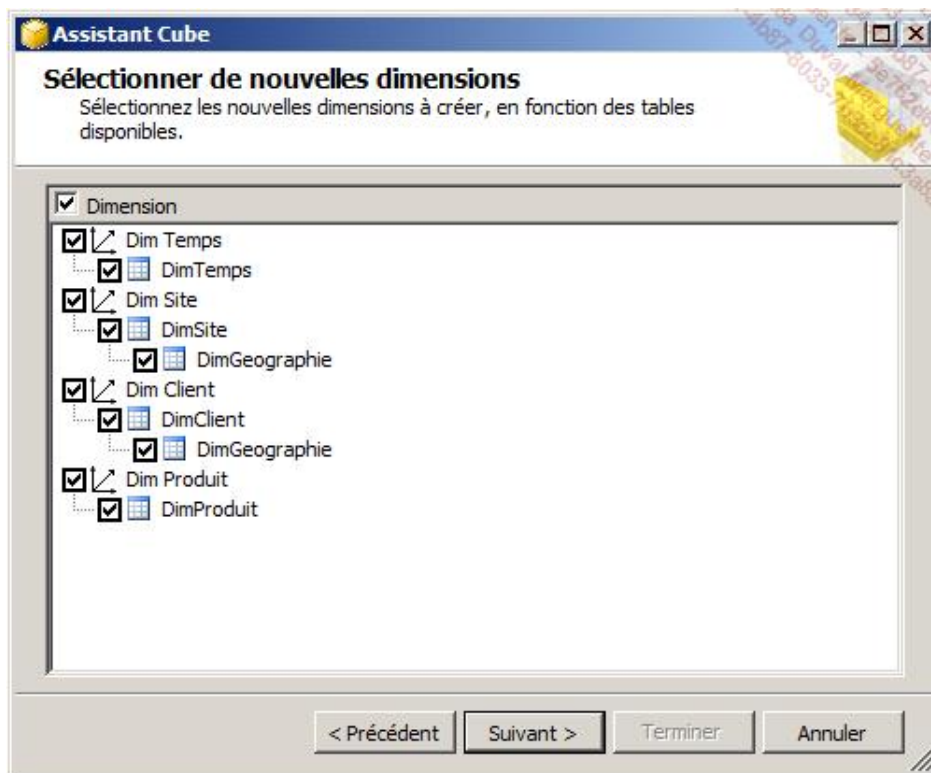
Dans l'écran suivant, l'assistant sélectionne de lui-même uniquement les mesures :



- Renommez le groupe de mesure *Fact Facture* en **Facture** et désélectionnez le compteur *Fact FactureNombre*.  
En effet, compter le nombre de lignes de la facture n'est pas une mesure pertinente dans notre cas.

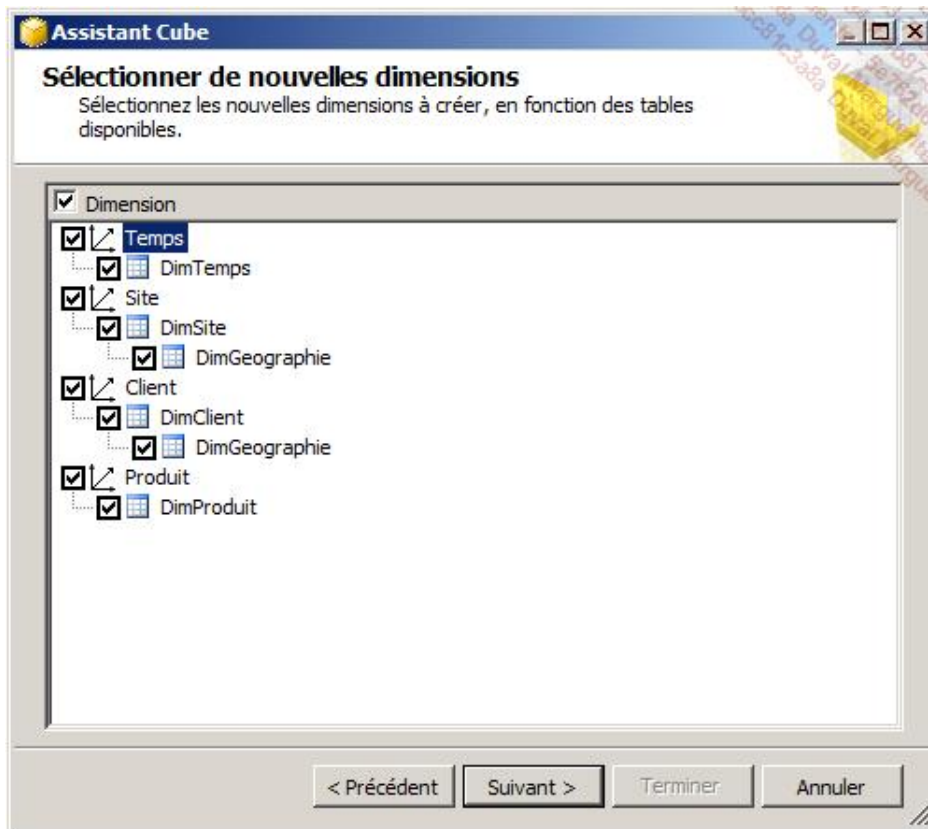


Les dimensions sont automatiquement détectées :

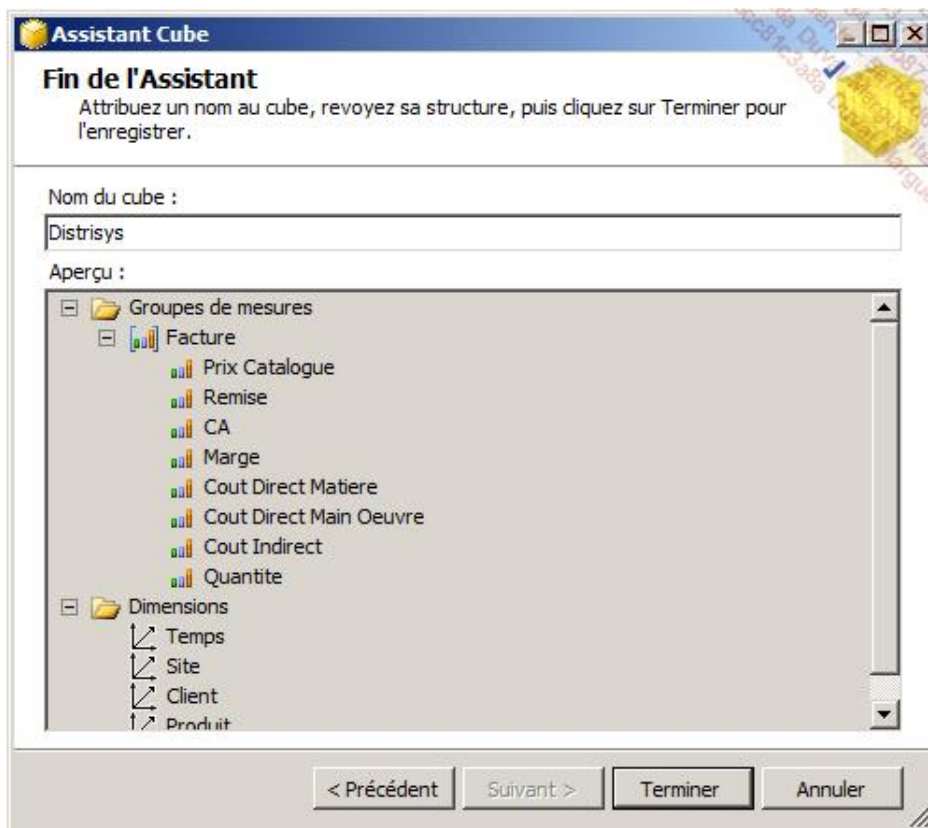


➤ Notez que la table de dimension **Geographie** n'est pas une dimension en tant que telle. Les informations de la tables seront disponibles au travers de la dimension **Client** et **Site**.

➔ Renommez les dimensions pour les rendre plus compréhensibles à un non-informaticien :



→ Renommez le cube si nécessaire :

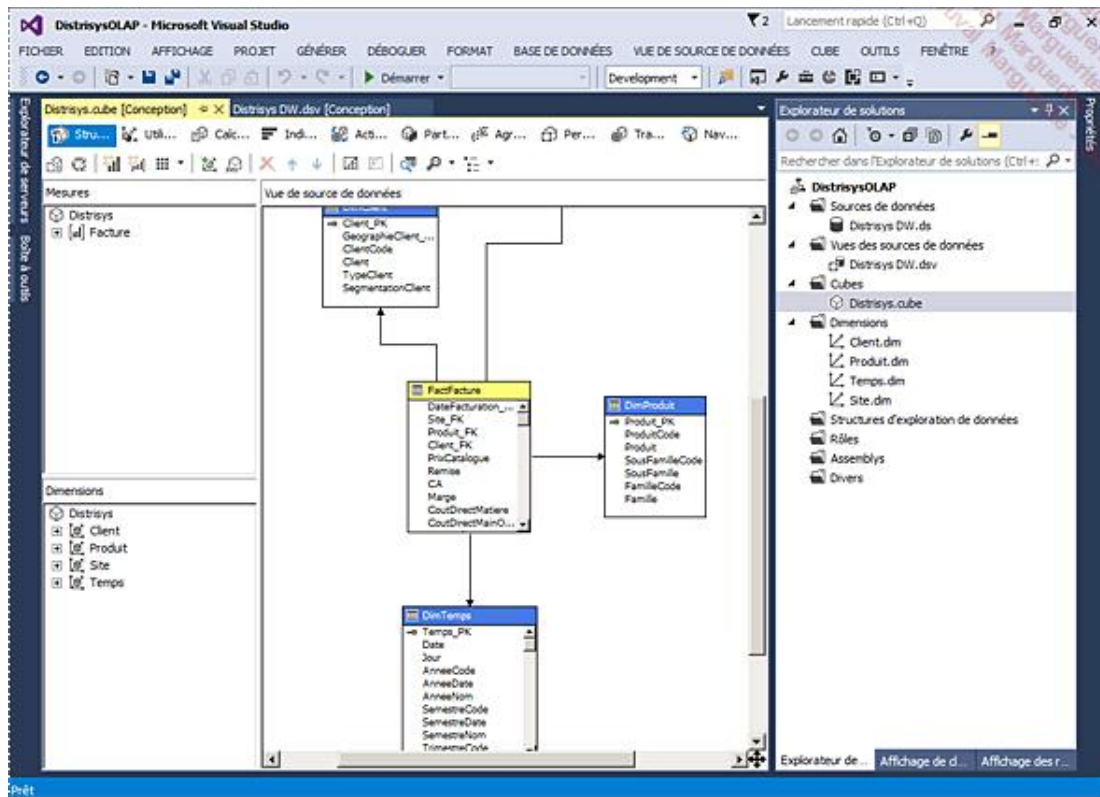


→ En cliquant sur **Terminer**, l'assistant va donc créer un cube *Distrisys*, comportant :

- Un groupe de mesures : *Facture* disposant des mesures CA, Prix Catalogue, remise, marge...

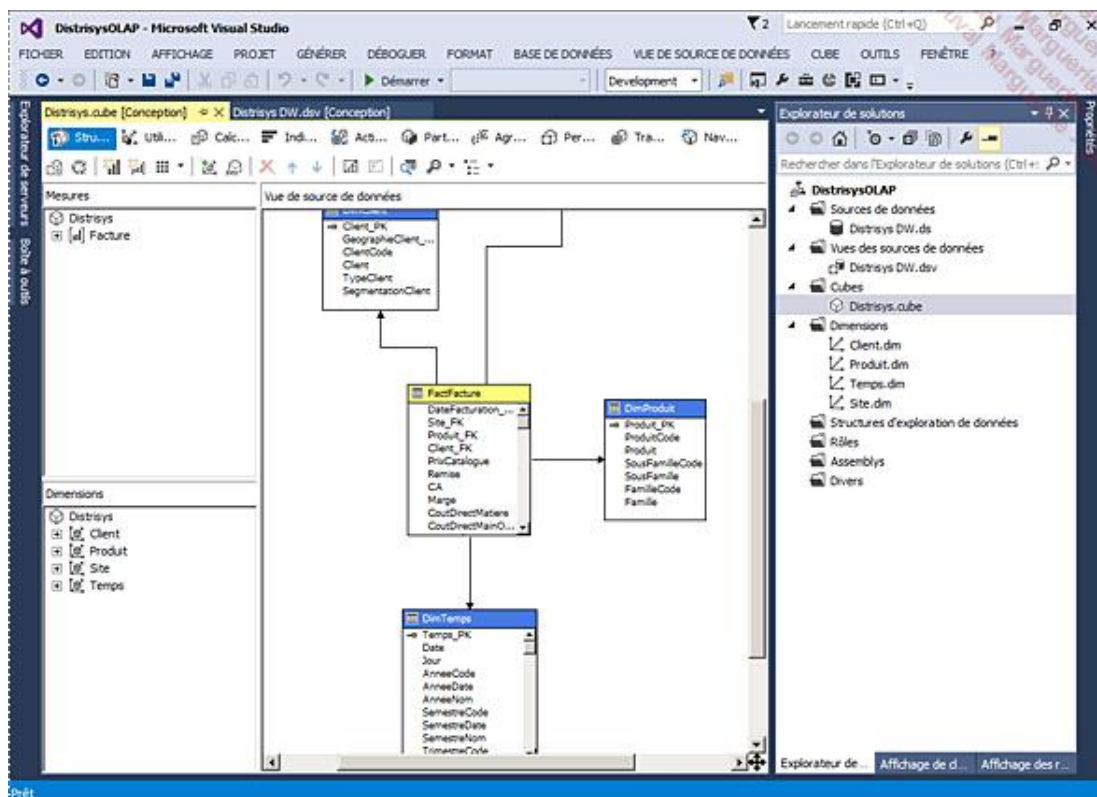
- Quatre dimensions : *Temps*, *Site*, *Client* et *Produit*

L'assistant est terminé, le cube est créé :

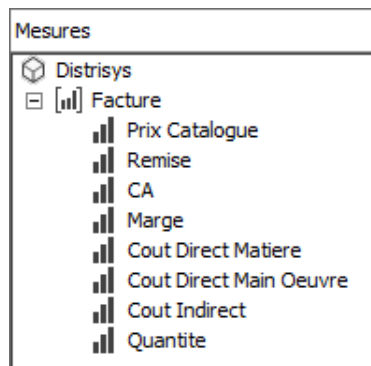


Dans cette interface concepteur, le contenu central, émanant de la vue de source de données, n'a que très peu d'intérêt.

Dans cette première interface, nous avons accès aux mesures regroupées par *groupe de mesures* (table de faits) :



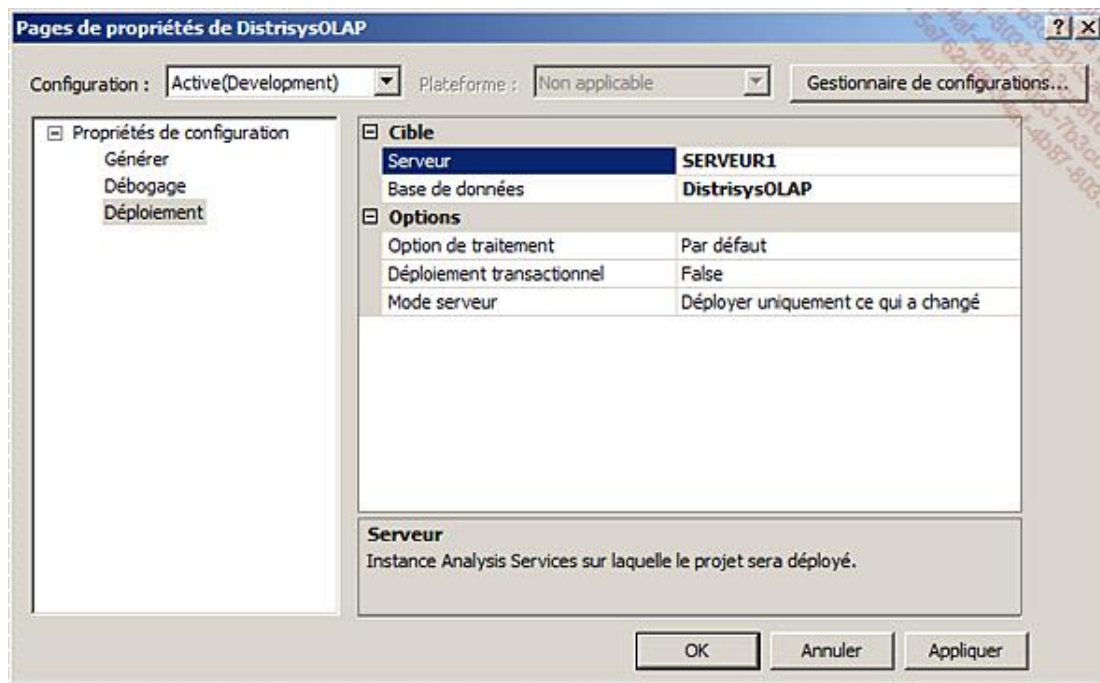
Ainsi qu'à la gestion des dimensions :




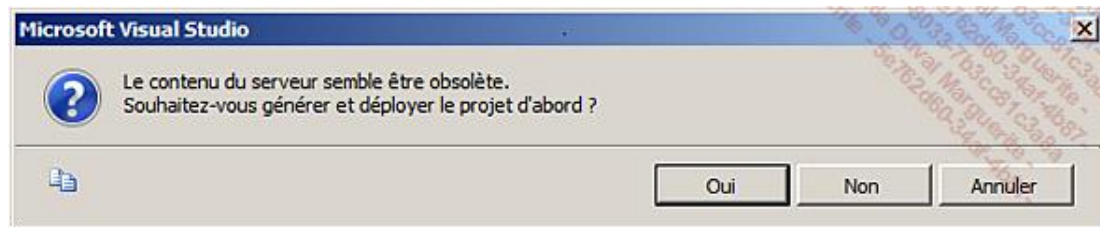
Avant de déployer le cube, nous allons nous assurer de le déployer sur le bon serveur Analysis Services.

- Dans la fenêtre **Explorateur de solutions**, faites un clic droit sur **DistrisysOLAP** puis cliquez sur **Propriétés**.
- Dans l'onglet **Déploiement**, au niveau du champ **Serveur**, entrez le nom et l'instance du serveur Analysis Services. Puis validez en cliquant sur **OK**.





- Visualisons sans plus tarder notre cube. Pour cela traitez le cube en cliquant sur l'icône .
- Au premier déploiement du cube ou à chaque modification de structure, le message ci-dessous apparaîtra, choisissez **Oui** :




Concrètement, le déploiement crée la structure de données.

- Après le déploiement du cube, la fenêtre de traitement apparaît, cliquez sur **Exécuter** :

**Traiter Cube - DISTRISYS**

Liste d'objets :

Nom de l'objet	Type	Options de traitement	Paramètres
 DISTRISYS	Cube	Traiter entièrement	

Résumé des paramètres du lot

Ordre de traitement :

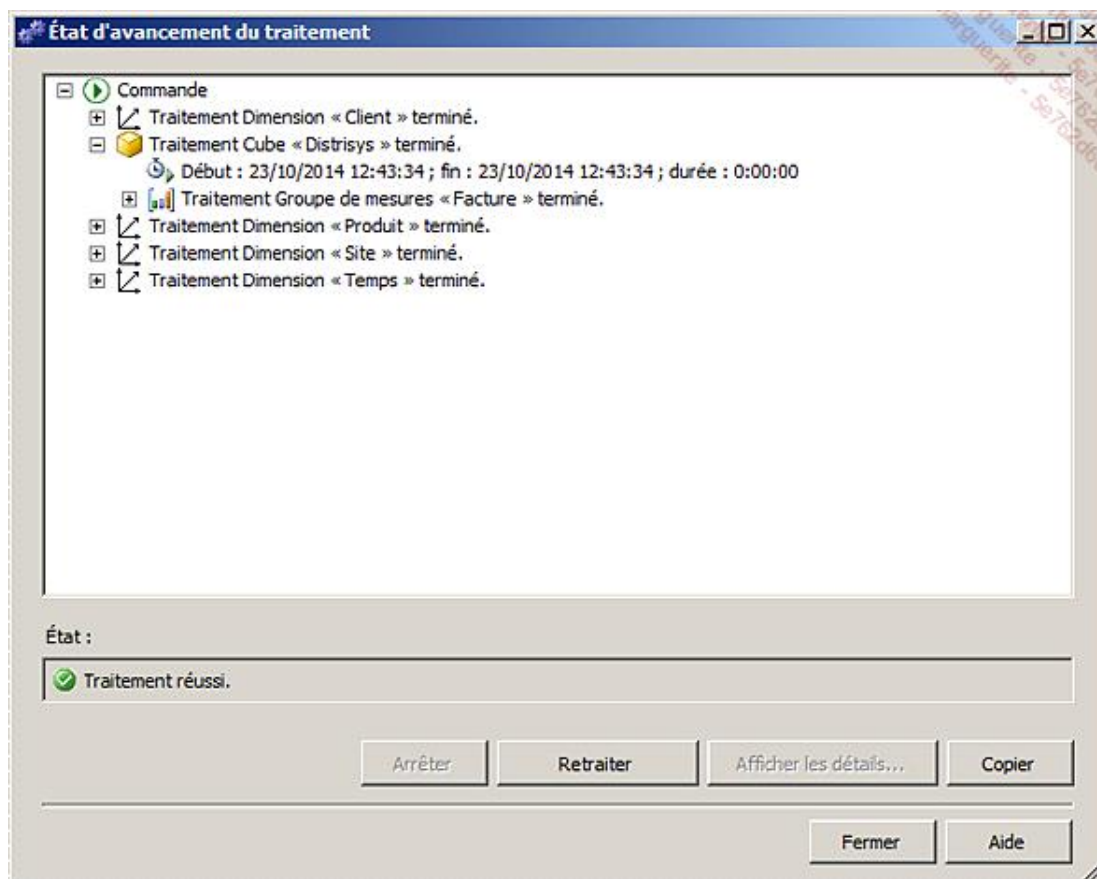
Mode de transaction :

Erreurs de la dimension :

Chemin du journal des erreurs de dé de dimension :

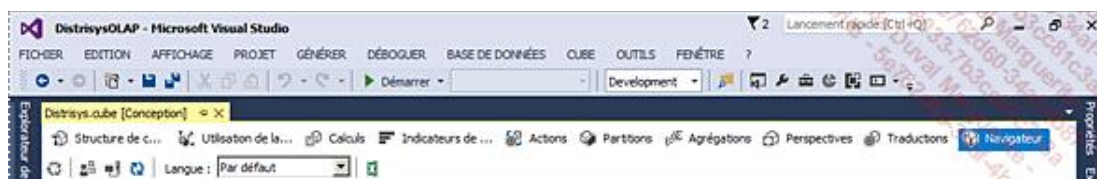
Traiter les objets affectés :

➤ Concrètement, traiter le cube revient à alimenter et à calculer les agrégats dans la structure déployée à l'étape précédente. Plus vous aurez de données, plus le temps de traitement va s'allonger. Ce n'est pas une tâche anodine, son optimisation relève du travail de l'administrateur Analysis Services. Étant donné que, dans notre cas, nous travaillons sur des données de test, le problème ne se pose pas et le traitement devrait se réaliser dans des délais assez courts.



Le traitement s'est terminé avec succès. Nous allons donc pouvoir manipuler le contenu du cube.

- Fermez toutes les fenêtres de traitement du cube.
- Cliquez sur l'onglet **Navigateur** du projet Analysis Services :



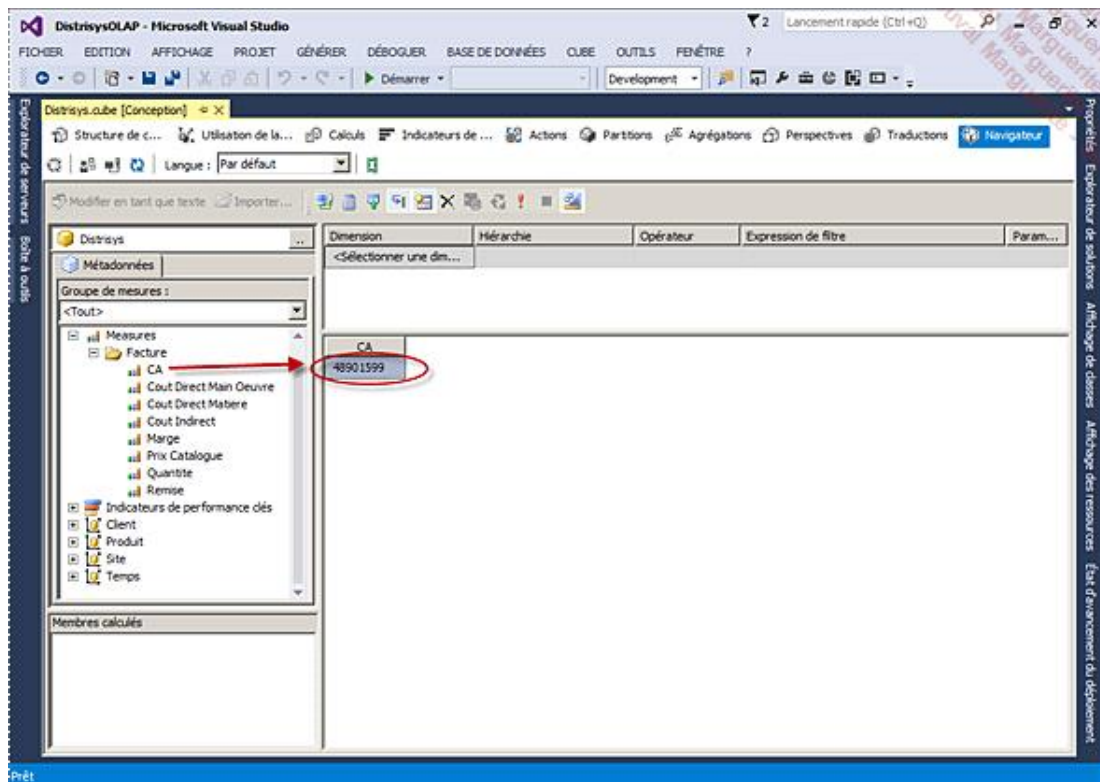
Le projet Analysis Services embarque un explorateur de cube afin de visualiser les modifications que vous effectuez. L'explorateur de cube le plus couramment employé est le *tableau croisé dynamique*.

Dans cette interface, l'explorateur de cube proposé est celui que vous retrouverez dans l'outil de création de rapport de Reporting Services. Cette interface ne permet pas de représentation en tableau croisé dynamique, mais préfère une représentation tabulaire.

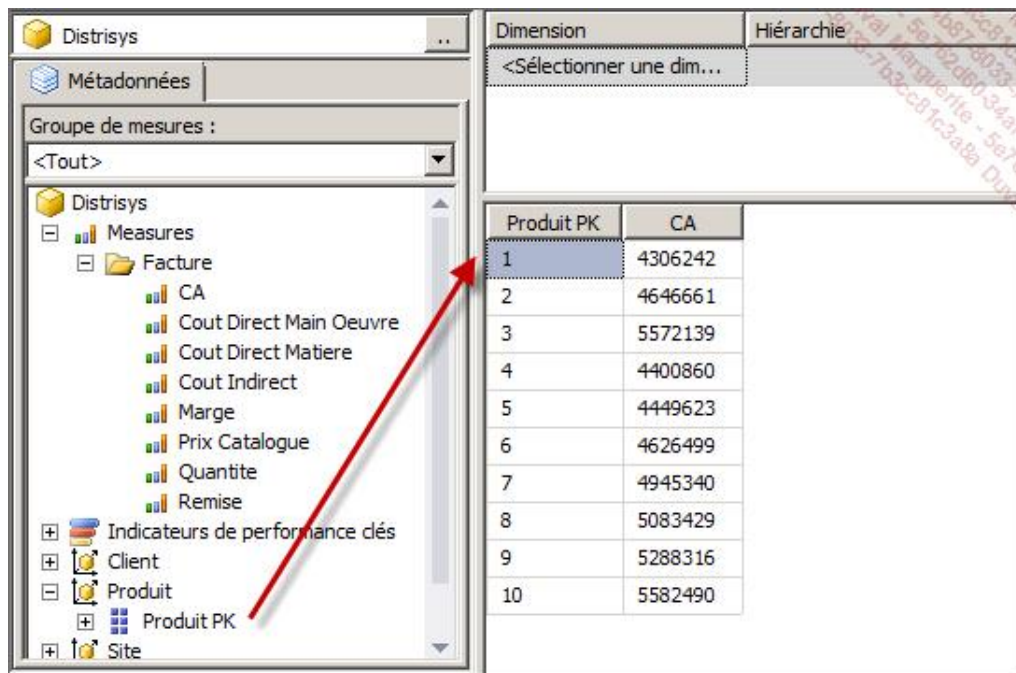
C'est une évolution d'interface de la version 2012. Un choix plutôt étonnant quand on sait que la représentation tabulaire est une représentation plutôt technique. Néanmoins, elle va dans un premier temps nous permettre de suivre les évolutions que nous allons effectuer au fur et à mesure sur le cube. Nous aborderons plus largement la représentation en tableau croisé dynamique dans le dernier chapitre avec notamment la restitution de données sous Microsoft Office Excel 2013.

Utilisons le navigateur embarqué dans SSDT afin d'afficher le contenu de notre cube.

- Glissez une mesure du cube dans l'espace central :



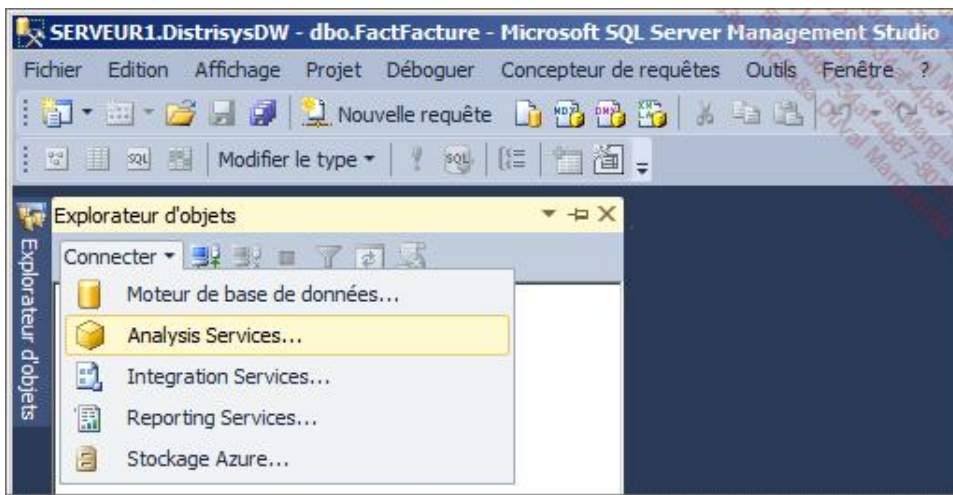
→ Puis glissez une dimension ou un de ses attributs :



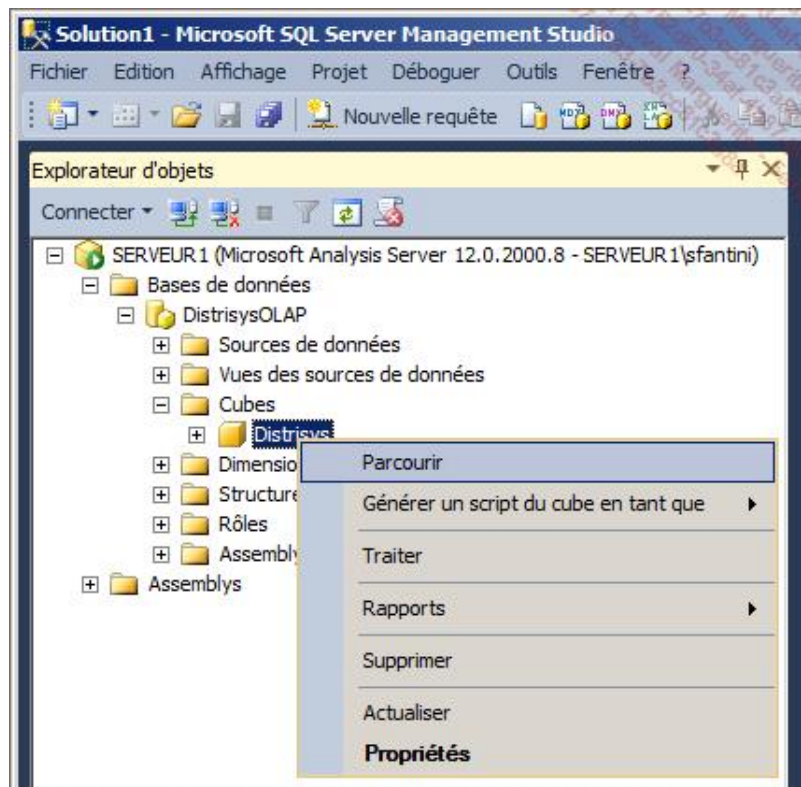
Votre cube est bien créé et vous avez la possibilité de l'explorer. Néanmoins, les dimensions ne sont accessibles que par leurs clés primaires.

Le même navigateur est disponible avec l'outil SQL Server Management Studio :

→ Ouvrez SSMS et connectez-vous à un serveur **Analysis Services** :

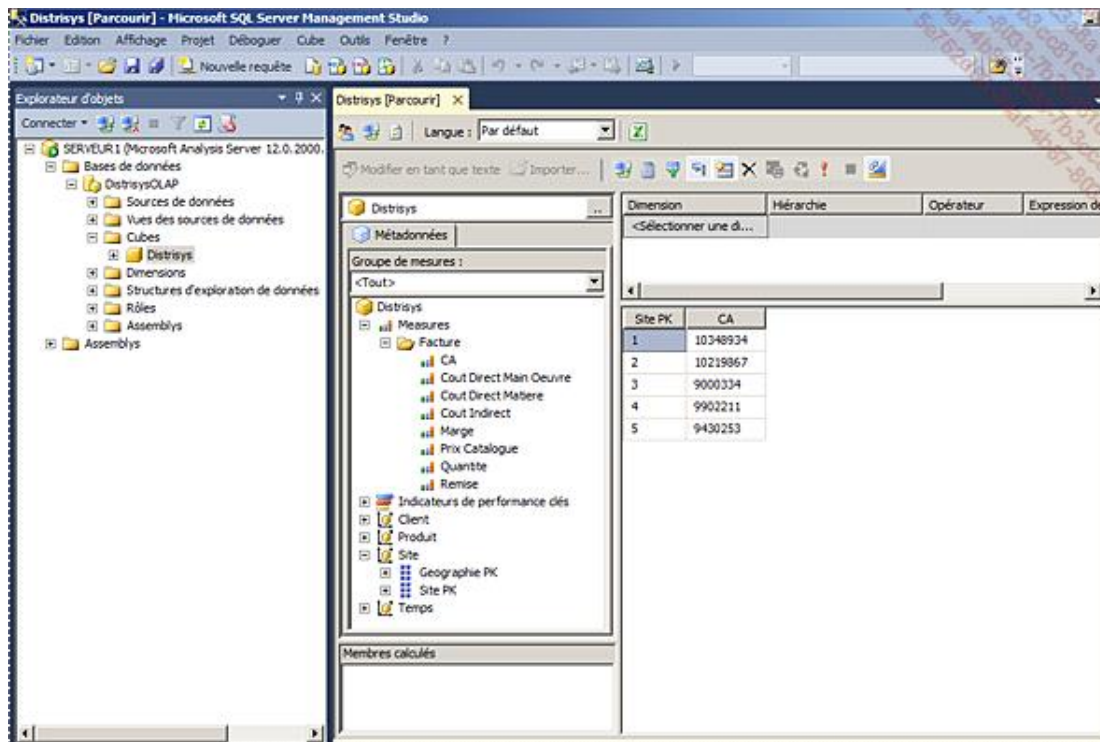


→ Ouvrez la base de données, puis le cube. À l'aide d'un clic droit sur **Parcourir**, vous avez accès à l'explorateur de cube :



Vous pouvez là encore accéder aux informations contenues dans votre cube :





Maintenant que vous savez créer un cube, au cours de la prochaine étape, nous allons apprendre à le peaufiner. Nous allons notamment travailler la présentation des dimensions et apprendre à créer des mesures complémentaires à celles présentes dans les tables de faits.