

Chapitre 4

La modélisation dimensionnelle

1. La matrice dimensionnelle de Distrisys

Au chapitre Réaliser son premier système décisionnel, vous avez découvert les rudiments pour commencer à créer l'entrepôt de données de Distrisys. Ce dernier chapitre est un préalable nécessaire au chapitre que nous allons à présent aborder, notamment si vous n'êtes pas un utilisateur aguerri d'Analysis Services ou un concepteur décisionnel averti.

Lors de ce chapitre, nous verrons la mise en place de l'entrepôt de données Distrisys par rapport aux domaines fonctionnels **Vente** et **Stock**.

Nous porterons notre attention en particulier sur ces deux domaines, car il s'agit de domaines fonctionnels classiques et assez standard pour la plupart des entreprises. Nous profiterons de ces domaines fonctionnels pour aborder des concepts décisionnels et des fonctions avancées d'Analysis Services.

Pour finir, nous balayerons les modélisations de périmètres assez classiques et récurrent tel que les achats, la comptabilité ou les ressources humaines.

■ Remarque

Même si le périmètre fonctionnel proposé dans cette étude de cas est assez éloigné d'un organisme public, d'une banque ou d'une association, les concepts qui seront mis en œuvre vous permettront tout de même de relever le défi de création de votre propre entrepôt de données.

Pour commencer à modéliser un entrepôt de données, il faut d'abord s'attacher à établir la matrice dimensionnelle de l'organisation.

La matrice dimensionnelle décrit les processus stratégiques ou du moins les plus importants de l'organisation. Ces descriptions de processus apparaissent en colonnes dans la matrice dimensionnelle sous forme de groupes de mesures (ou tables de faits).

Nous verrons au cours du chapitre qu'il existe trois types de tables de faits :

- Les tables de faits de type **transaction** : il s'agit de décrire en détail l'étape d'un processus (l'évènement).
- Les tables de faits de type **bilan** : il s'agit de faire le récapitulatif de certaines étapes du déroulement d'un processus.
- Les tables de faits de type **photo** : il s'agit de faire l'état des lieux d'un processus en un instant T (l'inventaire).

Prenons l'exemple d'une situation de la vie quotidienne, comme l'acte d'achat d'un article sur Internet. Le client navigue sur un site Internet, il détecte le produit qui lui plairait, il commande cet article. Trois jours plus tard, n'ayant toujours pas reçu le colis, il contacte le support qui le rassure. Le lendemain, en effet, le client reçoit son colis, il signe un bon de réception. Son colis est accompagné de la facture et d'un bon de livraison.

Voyons maintenant du côté de l'entrepôt de données de l'entreprise comment serait traduite cette situation :

Tout d'abord, chacune des pages vues du site Internet pourrait faire l'objet d'une ligne (de faits) dans la table de faits de Navigation du Site Internet (transaction).

En fin de session Internet, une ligne de faits, dans la table de faits Session Site Internet, pourrait faire le bilan du temps passé par l'internaute. Elle pourrait également comptabiliser le nombre de pages totales vues, le nombre d'articles consultés et pourrait noter si le client potentiel a acheté... Il s'agit alors d'une table de faits de type Bilan.

La commande de l'article par le client ajouterait une ligne de faits à la table de faits Commande (transaction).

Au sein de l'entreprise, la commande serait préparée, la facture éditée (transaction), l'article serait sorti du stock (transaction) et, au moment du départ vers le livreur, le bon de livraison serait édité (transaction).

Pendant ce temps, le client appelle le support de l'entreprise : ajout d'une ligne de faits dans la table des appels entrants du support (transaction).

Puis, le livreur fait signer au client un bon de réception électronique, acte qui permet d'ajouter une ligne de faits dans la table de faits des réceptions (transaction).

À la réception du colis, une nouvelle ligne de faits est ajoutée à la table de faits de bilan de commande, permettant de savoir comment s'est déroulé le processus de vente : délai écoulé, retard, nombre d'incidents, temps passé avec le support, coûts additionnels... (Bilan).

En fin de mois, un inventaire des stocks est réalisé tant en quantité qu'en valeur (Photo). Une photo des clients est également faite afin de comptabiliser leur nombre sous différents aspects : segmentation, comportement d'achat, localisation géographique... Il s'agit là aussi d'une table de faits de type photo.

Tout élément d'un processus peut être sujet à mesure, et dans une entreprise, les processus peuvent être très nombreux.

C'est pour cela que chez Distrisys, nous avons entrepris d'interviewer les membres du comité de direction. En effet, nous souhaitons une vision globale du projet mais également obtenir les grandes orientations de l'entreprise afin de nous concentrer sur les processus les plus stratégiques. Des documents complémentaires tels que le plan stratégique ou la cartographie des processus de l'entreprise sont des documents majeurs pour nous aider lors de cette étape.

Cette vision globale est très importante car elle permet :

- De se concentrer sur les processus les plus importants et donc d'aider à prioriser la réalisation du projet. En effet, si la modélisation doit être globale, la réalisation de l'entrepôt de données doit se faire étape par étape.
- D'établir la matrice dimensionnelle et ainsi d'avoir une vision exhaustive des dimensions qui doivent croiser un processus.

■ Remarque

Un processus n'est pas la propriété d'une activité. L'évaluation des stocks intéresse autant le service de gestion des stocks, que le service financier, de vente, d'achat, de production... même s'il est presque certain que ces différents services n'analyseront pas les stocks avec le même angle de vue. Néanmoins, pour la bonne marche de l'entreprise, aucun de ces services ne peut avoir une vision prépondérante sur ses voisins.

Par exemple, il n'est pas rare de voir des responsables marketing, production, achat ou de service financier parler d'un axe produit qui n'a de commun, au premier abord, que le nom... Une des grandes tâches du projet sera alors de travailler de concert, afin que la remontée d'un même processus puisse permettre à tous les services de faire les analyses spécifiques de leur activité, tout en retrouvant et comprenant les analyses de l'activité voisine.

- D'établir la matrice dimensionnelle et ainsi de mettre en exergue les difficultés à venir au niveau des dimensions.

■ Remarque

Contrairement à ce que l'on peut croire, la difficulté de réalisation d'un bon entrepôt de données vient de la volonté de vouloir croiser un axe d'analyse avec deux processus parfaitement distincts et pris en charge par deux services à la vision et à la culture opposée. Il est là, le vrai challenge de l'entrepôt de données.

Le fait de s'assurer un appui de la direction générale vous aidera assurément à relever ces challenges.

■ Remarque

Certains lecteurs s'étonneront sûrement sur le fait que nous n'avons pas abordé l'étude des systèmes sources. Cela est normal, car le système source, qu'il soit grand ERP, ERP modeste, système propriétaire ou réseau de fichier Excel, n'impacte pas la modélisation de votre entrepôt de données. Le seul intérêt de l'étude des systèmes sources, à ce stade, serait d'appréhender la disponibilité de l'information.

Après ce travail de synthèse, dressons la matrice dimensionnelle de Distrisys, disponible ci-dessous :

	Facture Entete	Facture	Budget Vente	Bilan Commande Client	Stock Photo	Stock Mouvement	Facture fournisseur	Bilan Commande Achat	Commande Achat en transit	Retour Fournisseur
Temps	Jour	Jour	Mois	Jour	Jour	Jour	Jour	Jour	Jour	Jour
Date demande				Jour				Jour	Jour	
Date prévue de livraison				Jour				Jour	Jour	
Date réelle de livraison				Jour				Jour		
Heure						X				
Produit		Produit	Produit		Produit	Produit	Produit		Produit	Produit
Site	Site	Site	Site	Site	Site	Site	Site	Site	Site	Site
Client	Client	Client		Client						
Fournisseur							Fournisseur	Fournisseur	Fournisseur	Fournisseur
Stock Mouvement						X				

Matrice dimensionnelle de Distrisys

La matrice dimensionnelle ci-dessus couvre les domaines fonctionnels suivants :

- Activité commerciale : Facture Entête, Facture, Budget Vente, Bilan Commande Client.
- Les stocks : Stock Photo et Stock Mouvement.
- Le service achat : Facture fournisseur, Bilan Commande Achat, Commande Achat en transit et Retour Fournisseur.

L'intérêt d'une telle démarche est d'ores et déjà d'annoncer que les principaux axes *Temps*, *Produit*, *Site*, *Client* et *Fournisseur*, permettront une analyse commune de processus parfois très éloignés ou difficiles à rapprocher.

Au cours de ce chapitre nous allons mettre en œuvre, ensemble et progressivement, cette matrice dimensionnelle afin qu'elle devienne un entrepôt de données, puis un cube.

2. Facturation et commande client

2.1 Modélisation et schéma en étoile

Pour commencer, nous allons nous intéresser aux tables de faits et aux dimensions relevant de l'activité commerciale. Ce domaine fonctionnel peut s'illustrer par la matrice dimensionnelle ci-dessous :

	Facture Entete	Facture	Budget Vente	Bilan Commande Client
Temps	Jour	Jour	Mois	Jour
Date demande				Jour
Date prévue de livraison				Jour
Date réelle de livraison				Jour
Produit		Produit	Produit	
Site	Site	Site	Site	Site
Client	Client	Client		Client

Matrice dimensionnelle du périmètre des ventes

Quelques explications s'imposent quant aux nouvelles tables de faits :

- **Facture**, que l'on aurait pu nommer FactureLigne pour la distinguer de Facture Entête, est la table de faits que nous avons mise en œuvre au chapitre Réaliser son premier système décisionnel. Cette table, qui enregistre une ligne de facture par fait, nous a permis de mesurer le CA, la marge, la remise et de détailler les différents coûts au niveau produit.
- **Facture Entete** est la table de faits identifiant une facture par fait. Étant donné qu'une facture peut comporter plusieurs produits, nous sommes dans l'incapacité, avec la seule table de faits FactFacture, de pouvoir compter le nombre de factures émises. Utilisée pour elle-même, la mesure Nombre de factures n'est pas une mesure très intéressante. En revanche, elle le devient si elle permet la création de mesures calculées, telles que Prix moyen facturé ou Nombre d'articles moyen par facture.

Chapitre 7

Rapports avec graphique

1. Introduction

Les graphiques sont des éléments particulièrement utiles pour présenter les données de façon concise et efficace. Camemberts (graphiques de type secteurs), barres verticales ou horizontales, lignes, etc. sont utilisés par les outils de reporting en complément de tableaux de données ou seuls.

2. Graphique de type barres verticales

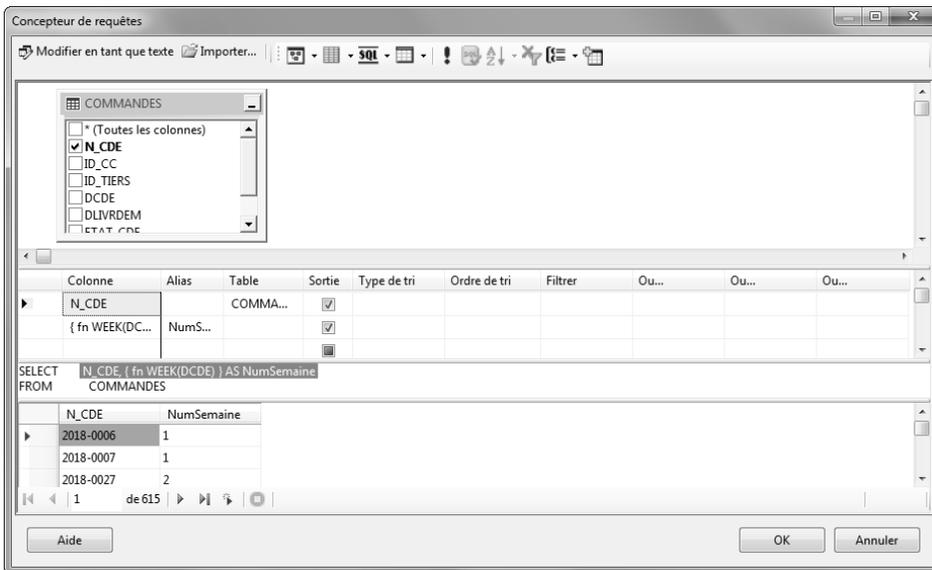
On souhaite établir un rapport qui présente l'évolution du nombre de commandes en fonction du mois ou de la semaine, avec un graphique de type barres verticales.

Créer un nouveau rapport, avec une source de données et un dataset qui collecte les informations qui serviront à établir le graphique.

328 _____ SQL Server Reporting Services

Conception, déploiement et distribution de rapports

Exemple :



Pour récupérer le numéro de semaine, on utilise la fonction **Week**. On peut saisir la fonction dans le volet de la requête SQL : **WEEK(DCDE)**.

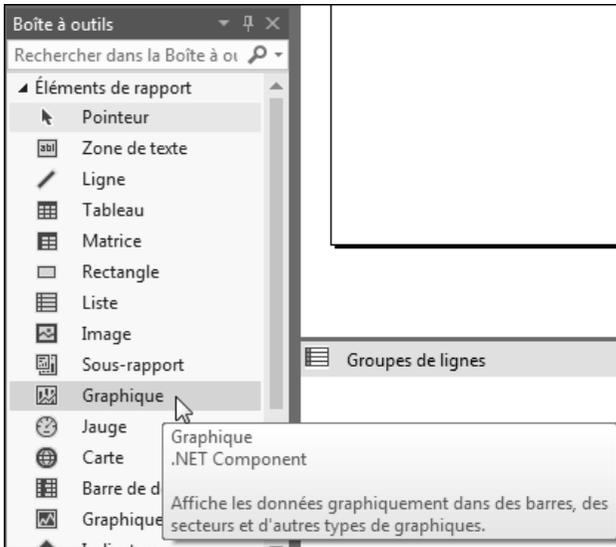
Le concepteur ajoute un encadrement : **{ fn WEEK(DCDE) }**

La requête est donc :

```
SELECT N_CDE, { fn WEEK(DCDE) } AS NumSemaine FROM COMMANDES
```

On ne collecte pas les données en utilisant les fonctions d'agrégat. C'est le rapport qui s'en chargera. Inutile donc d'utiliser la fonction **Count** et l'instruction **Group by**.

Cliquer sur **OK** et, en mode conception, faire glisser un composant **Graphique** à partir de la boîte à outils vers le corps du rapport.

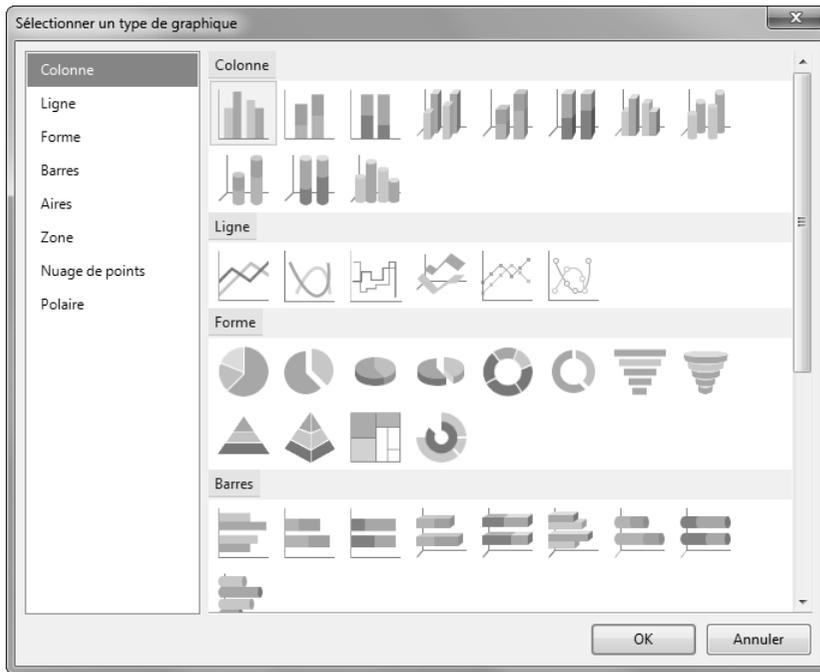


La fenêtre **Sélectionner un type de graphique** est affichée.

Les possibilités sont nombreuses. Il y a huit catégories, et pour chacune plusieurs types de graphiques. On peut balayer les icônes pour visualiser les textes d'info-bulles.

330 _____ SQL Server Reporting Services

Conception, déploiement et distribution de rapports



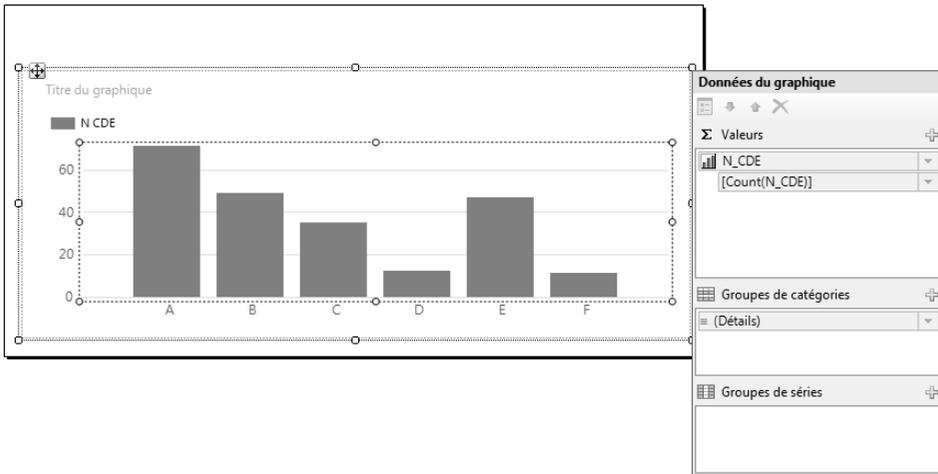
Choisir la première icône de la rubrique **Colonne** et cliquer sur **OK**.

Étirer le rectangle contenant le graphique pour obtenir une taille adaptée.

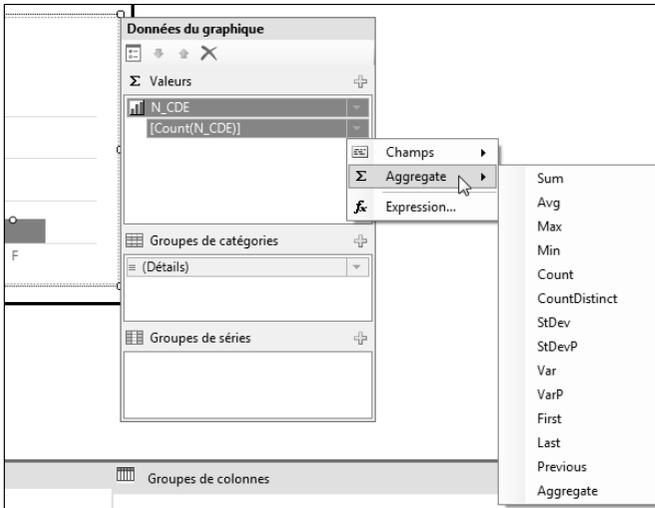
Cliquer une ou deux fois sur le graphique (une fois seulement s'il est déjà sélectionné) pour faire apparaître une fenêtre **Données du graphique**. Cette fenêtre contient trois sections : Σ **Valeurs**, **Groupes de catégories** et **Groupes de séries**. Chacune de ces sections peut contenir une ou plusieurs valeurs.

Pour notre besoin, nous souhaitons voir les numéros de semaines sur l'axe horizontal, et le nombre de commandes correspondant indiqué pour chaque semaine.

À partir de la fenêtre **Données du rapport**, glisser le champ à compter (ici **N_CDE**) dans la section Σ **Valeurs**. La fonction **Count** est automatiquement sélectionnée.



On peut bien sûr choisir la fonction d'agrégat à utiliser. Sélectionner celle choisie et cliquer sur la liste déroulante, comme pour en changer :



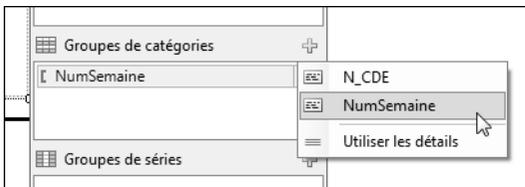
La fonction **Count** convient. Nous allons la conserver.

Glisser le champ **NumSemaine** dans la section **Groupes de catégories**.

332 _____ SQL Server Reporting Services

Conception, déploiement et distribution de rapports

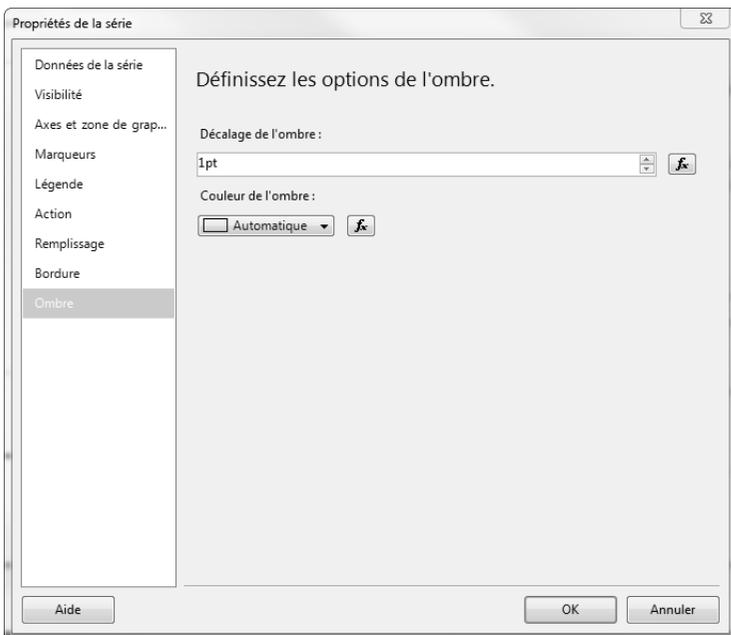
On peut aussi utiliser le signe **+** placé en haut à droite de la section pour faire apparaître les champs à sélectionner.



2.1 Propriétés de la série

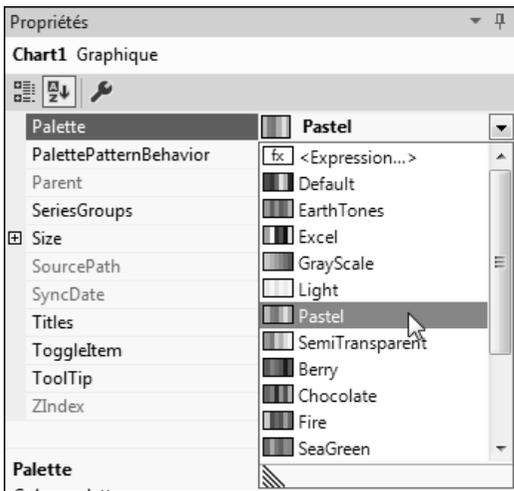
Sélectionner la donnée placée dans la section Σ **Valeurs**. Utiliser le menu contextuel pour choisir **Propriétés de la série**.

Sélectionner la section **Remplissage**, conserver le remplissage plein mais changer la couleur. Puis sélectionner la section **Ombre**, indiquer un décalage de **1pt** et choisir une couleur.



Cliquer sur **OK**.

Nota : pour les couleurs, on peut laisser la valeur à **Automatique** et laisser faire le concepteur. Dans ce cas, la couleur sera choisie dans la palette associée au graphique. Pour modifier cette palette, choisir le graphique et dans sa fenêtre **Propriétés**, modifier l'attribut **Palette** :



Il est également possible de définir sa propre palette et de limiter les couleurs ainsi utilisées.

2.2 Titres des axes

Remplacer le titre du graphique par **Évolution hebdomadaire du nombre de commandes**.

Sélectionner l'axe horizontal et remplacer son titre par **Semaines**. Remplacer le titre de l'axe vertical par **Nb Cdes**.