

La dimension temps

Dans un système décisionnel, la dimension *Temps* revêt une importance particulière et doit faire l'objet de la plus grande attention. La table *Temps* est unique et sa présence est obligatoire dans toute table de faits, quelle qu'elle soit. La présence de la dimension *Temps* dans une table de faits est une des caractéristiques d'un système décisionnel par rapport à un système transactionnel.

La table de dimension *Temps* a une granularité au *jour*.

- Dans certains cas, vous aurez besoin de faire des analyses à l'heure : il s'agira alors de créer une dimension *Heure*. Nous n'intégrerons jamais les heures au sein de la dimension *Temps*.
- Dans d'autres cas, vous aurez une table de faits à la granularité mois : il s'agira alors de considérer le premier jour ou le dernier jour du mois comme étant représentatif du mois. Nous aborderons ce cas-là ultérieurement, dans le chapitre La modélisation dimensionnelle - Facturation et commande client.

Le premier réflexe est de construire une table de dimension *Temps* assez simple, comme ci-dessous :

	Nom de la colonne	Type de données	Autoriser l...
?	Temps_PK	int	<input type="checkbox"/>
	Annee	varchar(20)	<input type="checkbox"/>
	Mois	varchar(20)	<input type="checkbox"/>
	Jour	varchar(20)	<input type="checkbox"/>

- Ne créez pas cette table car cette construction se révélera vite très insuffisante. Il faudrait la compléter par bon nombre d'attributs supplémentaires.

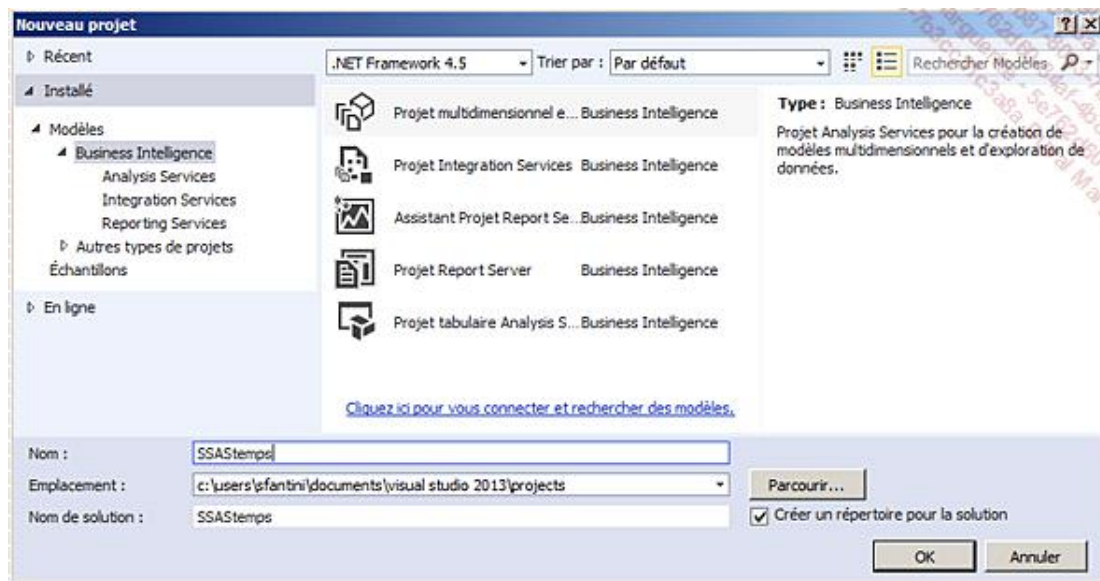
L'expérience vous apprendra qu'une table de dimension bien construite permettra d'anticiper bon nombre de complications ultérieures.

Au cours des pages suivantes, nous allons créer la table *Temps*, saisir son contenu, puis la peaufiner pour qu'elle corresponde à nos attentes. Néanmoins, la table *DimTemps* finale est disponible en téléchargement au format csv sur l'espace de téléchargement du livre.

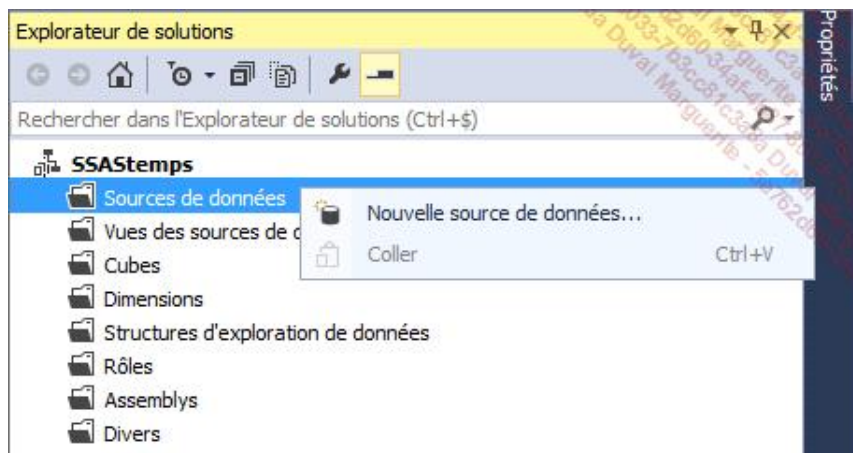
L'intérêt de la procédure qui suit est de vous permettre de vous constituer votre propre *DimTemps* qui reflétera les spécificités de votre organisation.

Pour commencer, nous allons nous servir d'un assistant de projet SSAS, afin de générer une première version de la table *DimTemps*. La création de la table de dimension *Temps* va nous permettre d'avoir un premier contact avec l'outil de création de cubes de Microsoft : SSDT (*SQL Server Data Tools*). Nous reviendrons, plus tard dans le chapitre et plus en détails, sur l'environnement de création de cube.

- ➔ Ouvrez l'outil SSDT et créez un nouveau projet.
- ➔ Sélectionnez un projet de type **Projet Analysis Services** :

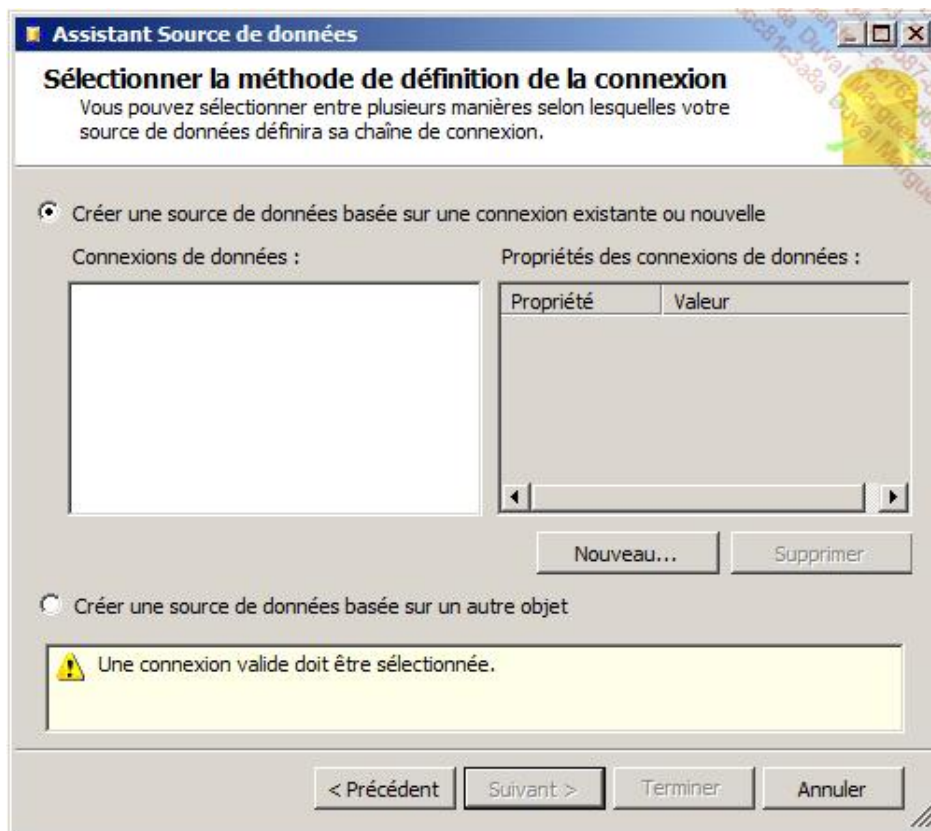


→ Créez une nouvelle source de données :



Nous allons suivre l'assistant de création d'une nouvelle source de données :

- L'assistant s'ouvre sur un écran d'accueil, cliquez sur le bouton **Suivant**.
- Aucune référence à une source de données n'étant encore définie, cliquez sur le bouton **Nouveau** pour en créer une nouvelle.



- Saisissez le nom de votre instance SQL Server, ainsi que le nom de l'entrepôt de données, puis cliquez sur **Tester la connexion** pour vérifier que les paramètres sont corrects.

Gestionnaire de connexions

Fournisseur : OLE DB natif\SQL Server Native Client 11.0

Nom du serveur : SERVEUR1 Actualiser

Connexion au serveur

☒ Utiliser l'authentification Windows

☐ Utiliser l'authentification SQL Server

Nom d'utilisateur :

Mot de passe :

☐ Enregistrer mon mot de passe

Connexion à la base de données

☒ Sélectionner ou entrer un nom de base de données :

DistrisysDW

☐ Attacher un fichier de base de données :

Parcourir...

Nom logique :

Tester la connexion OK Annuler Aide

→ Cliquez sur **Suivant** pour continuer la configuration de la source de données.

Assistant Source de données

Sélectionner la méthode de définition de la connexion

Vous pouvez sélectionner entre plusieurs manières selon lesquelles votre source de données définira sa chaîne de connexion.

☒ Créer une source de données basée sur une connexion existante ou nouvelle

Connexions de données :

Propriété	Valeur
Data Source	SERVEUR1
Initial Catalog	DistrisysDW
Integrated Sec...	SSPI
Provider	SQLNCLI11.1

Nouveau... Supprimer

☐ Créer une source de données basée sur un autre objet

< Précédent Suivant > Terminer >>| Annuler

→ Sélectionnez **Utiliser le compte de service** pour vous connecter à la base de données.

The screenshot shows the 'Assistant Source de données' window with the title bar 'Assistant Source de données'. The main heading is 'Informations d'emprunt d'identité'. Below it, a subtitle reads: 'Vous pouvez définir les informations d'identification Windows qu'Analysis Services doit utiliser pour se connecter à la source de données.' There are four radio button options: 'Utiliser un nom d'utilisateur et un mot de passe Windows spécifiques' (unselected), 'Utiliser le compte de service' (selected), 'Utiliser les infos d'identification de l'utilisateur actuel' (unselected), and 'Hériter' (unselected). Below the first option are two text boxes labeled 'Nom d'utilisateur :' and 'Mot de passe :'. At the bottom, there are four buttons: '< Précédent', 'Suivant >', 'Terminer', and 'Annuler'.

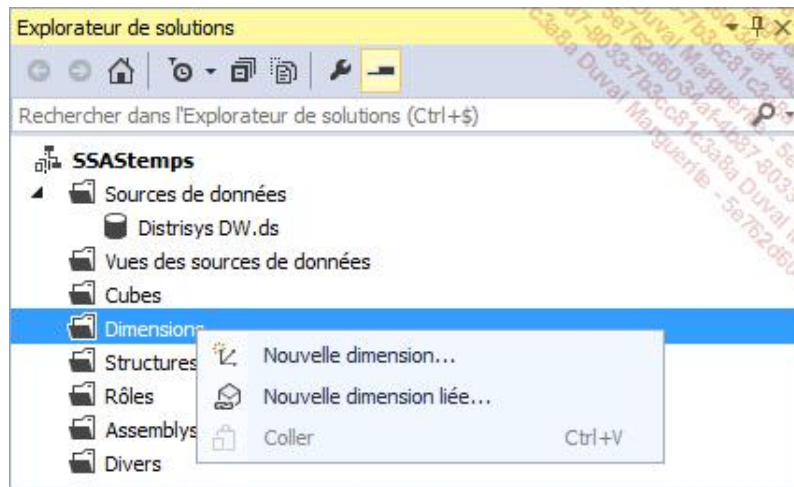
→ Cliquez sur **Terminer**.

The screenshot shows the 'Assistant Source de données' window with the title bar 'Assistant Source de données'. The main heading is 'Fin de l'Assistant'. Below it, a subtitle reads: 'Fournissez un nom, puis cliquez sur Terminer pour créer la source de données.' There is a text box labeled 'Nom de la source de données :' containing the text 'Distrisys DW'. Below this is a section labeled 'Aperçu :'. Inside this section, under the heading 'Chaîne de connexion :', the following text is displayed: 'Provider=SQLNCLI11.1;Data Source=SERVEUR1;Integrated Security=SSPI;Initial Catalog=DistrisysDW'. At the bottom, there are four buttons: '< Précédent', 'Suivant >', 'Terminer', and 'Annuler'.

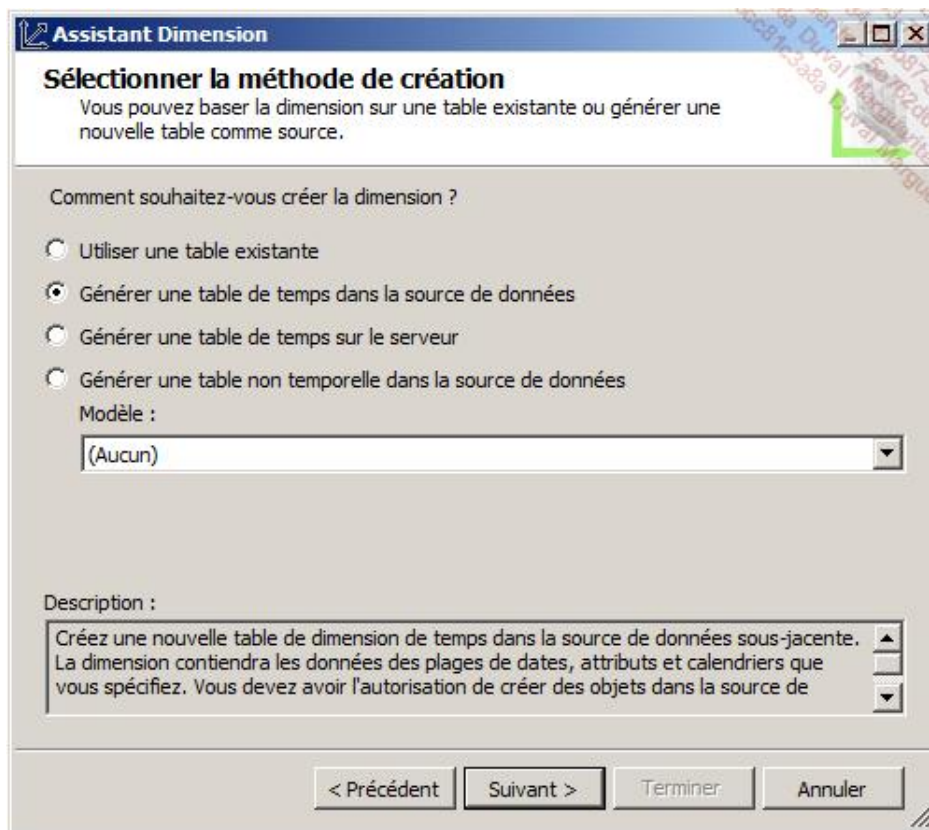
Maintenant que la source de données est créée, nous allons créer une dimension *Temps* dans le projet Analysis

Services.

- Créez une nouvelle dimension. Pour cela, dans l'**Explorateur de solutions**, faites un clic droit sur **Dimensions** puis, dans le menu contextuel, cliquez sur **Nouvelle dimension**.



- Un nouvel assistant s'ouvre. Dans la fenêtre **Sélectionner la méthode de création**, sélectionnez l'option **Générer une table de temps dans la source de données**.



- Sélectionnez la période qui vous intéresse : dans notre cas, nous souhaitons une plage allant du 1^{er} janvier 2009 au 31 décembre 2018.

Assistant Dimension

Définir des périodes

Sélectionnez les périodes à utiliser lors de la création des hiérarchies.

Premier jour civil : jeudi 1 janvier 2009

Demier jour civil : lundi 31 décembre 2018

Premier jour de la semaine : Dimanche

Périodes :

- ☒ Année
- ☒ Semestre
- ☒ Trimestre
- ☐ Quadrimestre
- ☒ Mois
- ☐ Décade
- ☒ Semaine
- ☒ Date

Langue pour les noms des membres de temps : Français (France)

< Précédent Suivant > Terminer >> Annuler

→ Sélectionnez le ou les types de calendrier qui sont supportés par votre axe temps.

➤ Dans notre cas, nous n'aurons besoin que du calendrier régulier. Dans la réalité, il est très probable que des services tels que le service financier, les ventes ou les services production et logistique aient chacun leur propre calendrier qui diffère du calendrier standard. Le calendrier standard commence le 1^{er} janvier et finit le 31 décembre. Le calendrier fiscal d'une entreprise peut par exemple commencer le 1^{er} septembre et terminer le 31 août de l'année suivante.

Assistant Dimension

Sélectionner des calendriers

Sélectionnez les calendriers pour lesquels vous souhaitez créer des hiérarchies.

☒ Calendrier régulier

☐ Calendrier fiscal

 Jour et mois de départ : 1 Janvier

 Convention de nom du calendrier fiscal : Nom de l'année civile + 1

☐ Calendrier rapports (ou marketing)

 Jour et mois de départ : 1 Janvier

 Modèle semaine du mois : Semaine 445

☐ Calendrier de fabrication

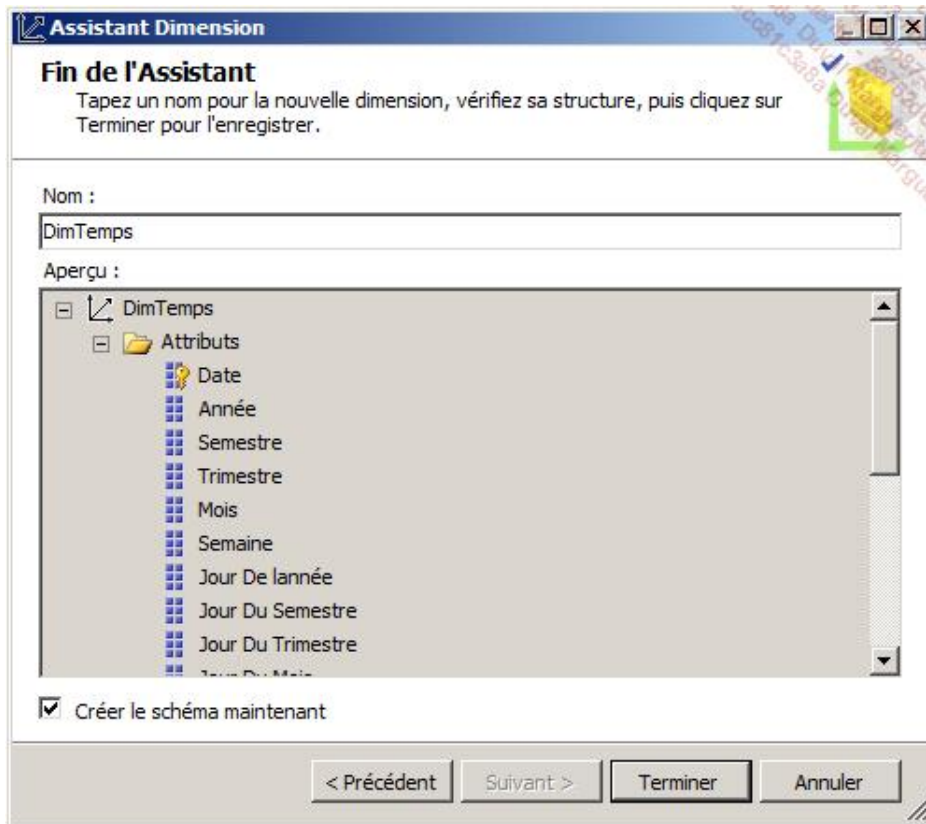
 Jour et mois de départ : 1 Janvier

 Trimestre avec des périodes suppl. : 4

☐ Calendrier ISO 8601

< Précédent Suivant > Terminer Annuler

- Cochez l'option **Créer le schéma maintenant** afin que l'assistant crée la structure et génère également les données. Changez le nom de la nouvelle dimension en **DimTemps** puis cliquez enfin sur le bouton **Terminer**.



- À l'affichage de l'assistant de génération de schéma, cliquez sur **Suivant**.
- Sélectionnez l'option **Créer une nouvelle vue de source de données** nommée arbitrairement **SSAS temps**.

Nous reviendrons plus en détail sur ces aspects du projet Analysis Services. Ce projet ci n'a pas d'autre vocation que de générer la dimension temps.

Assistant Génération de schéma

Spécifier la cible
Spécifiez la vue de source de données dans laquelle il faut produire le schéma.

☒ Créer une nouvelle vue de source de données

Nom de la vue de source de données :

Source de données :

Une nouvelle vue de source de données sera créée pour la source de données que vous spécifiez.

☐ Utiliser une vue de source de données existante

La vue de source de données sélectionnée et sa source de données seront utilisées.

< Précédent Suivant > Terminer Annuler

→ Sélectionnez l'option **Remplir** afin de générer les données de la table *DimTemps* :

Assistant Génération de schéma

Options du schéma de la BdD de la zone de sujet
Spécifiez les options de génération de schéma pour la base de données de la zone de sujet.

Schéma propriétaire :

☒ Créer des clés primaires sur les tables de la dimension

☒ Créer des index

☒ Appliquer l'intégrité référentielle

☒ Préserver les données lors de la régénération

Remplir la ou les tables de temps :

< Précédent Suivant > Terminer Annuler

→ L'assistant affiche un écran de paramétrage des conventions de nommage. À l'option **Séparateur**, spécifiez la valeur **Aucun** :

Assistant Génération de schéma

Spécifier les conventions de nom
Spécifiez les conventions de nom que vous souhaitez utiliser dans le nouveau schéma.

Option	Valeur
Séparateur	Aucun
Préfixe de la colonne de clé primaire	PK
Préfixe de la colonne de clé étrangère	FK
Suffixe du nom de l'attribut	Nom
Suffixe du cumul personnalisé	CustomRollup
Suffixe des propriétés de cumul personn...	CustomRollupProperties
Suffixe de l'opérateur unaire	UnaryOperator
Suffixe de niveaux ignorés	SkippedLevels
Suffixe de colonne de valeur	Valeur

< Précédent Suivant > Terminer Annuler

→ Achetez l'assistant en cliquant sur le bouton **Terminer**.

Assistant Génération de schéma

Fin de l'Assistant
Vérifiez les options de génération, puis cliquez sur Terminer.

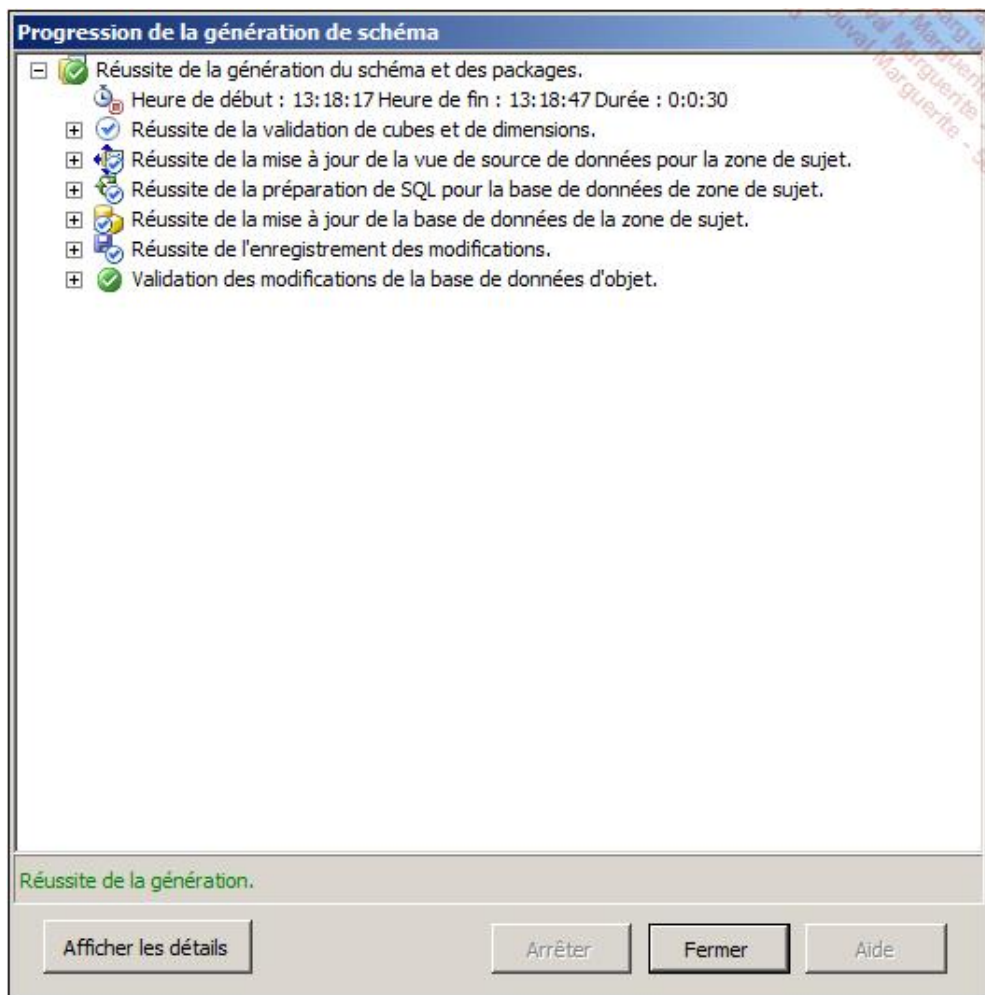
Résumé :

Zone de sujet

- Vue de source de données : SSAS temps
- Schéma : SSAS temps
- Les données seront conservées
- La table de durée sera remplie

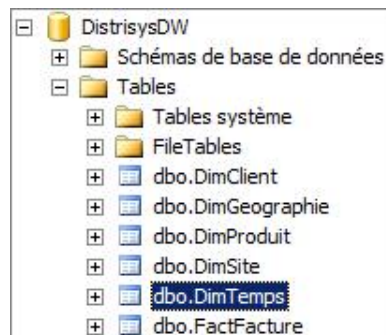
< Précédent Suivant > Terminer Annuler

→ À la fin du processus de génération du schéma, cliquez sur **Fermer**.



L'assistant est terminé, la dimension est créée. Vous pouvez quitter le projet.

Dans l'outil SQL Server Management Studio, vous pouvez constater que la table *DimTemps* vient d'être ajoutée :



La structure de la table *DimTemps* générée est la suivante :

SERVEUR 1.DistrisysDW - dbo.DimTemps X			
	Nom de la colonne	Type de données	Autoriser l...
?	PKDate	datetime	<input type="checkbox"/>
	DateNom	nvarchar(50)	<input checked="" type="checkbox"/>
	Année	datetime	<input checked="" type="checkbox"/>
	AnnéeNom	nvarchar(50)	<input checked="" type="checkbox"/>
	Semestre	datetime	<input checked="" type="checkbox"/>
	SemestreNom	nvarchar(50)	<input checked="" type="checkbox"/>
	Trimestre	datetime	<input checked="" type="checkbox"/>
	TrimestreNom	nvarchar(50)	<input checked="" type="checkbox"/>
	Mois	datetime	<input checked="" type="checkbox"/>
	MoisNom	nvarchar(50)	<input checked="" type="checkbox"/>
	Semaine	datetime	<input checked="" type="checkbox"/>
	SemaineNom	nvarchar(50)	<input checked="" type="checkbox"/>
	JourDelannée	int	<input checked="" type="checkbox"/>
	JourDelannéeNom	nvarchar(50)	<input checked="" type="checkbox"/>
	JourDuSemestre	int	<input checked="" type="checkbox"/>

Et vous pouvez vérifier que les lignes de la table *DimTemps* ont bien été saisies :

	PKDate	DateNom	Année	AnnéeNom	Semestre	Semestre
1	2009-01-01 00:00:00.000	jeudi, janvier 01 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
2	2009-01-02 00:00:00.000	vendredi, janvier 02 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
3	2009-01-03 00:00:00.000	samedi, janvier 03 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
4	2009-01-04 00:00:00.000	dimanche, janvier 04 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
5	2009-01-05 00:00:00.000	lundi, janvier 05 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
6	2009-01-06 00:00:00.000	mardi, janvier 06 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
7	2009-01-07 00:00:00.000	mercredi, janvier 07 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
8	2009-01-08 00:00:00.000	jeudi, janvier 08 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
9	2009-01-09 00:00:00.000	vendredi, janvier 09 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
10	2009-01-10 00:00:00.000	samedi, janvier 10 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
11	2009-01-11 00:00:00.000	dimanche, janvier 11 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
12	2009-01-12 00:00:00.000	lundi, janvier 12 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
13	2009-01-13 00:00:00.000	mardi, janvier 13 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
14	2009-01-14 00:00:00.000	mercredi, janvier 14 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr
15	2009-01-15 00:00:00.000	jeudi, janvier 15 2009	2009-01-01 00:00:00.000	Calendrier 2009	2009-01-01 00:00:00.000	Semestr

Nous disposons maintenant d'une table de dimension temps beaucoup plus conforme à nos attentes. Néanmoins, nous allons continuer d'apporter manuellement quelques améliorations qui nous seront profitables par la suite.

- Par expérience, nous vous recommandons d'utiliser une clé **Temps_PK** de type int au format **aaaammjj**. Cette clé vous évitera de nombreux écueils lors du chargement de vos données, les champs de type Date étant souvent sources de problème d'alimentation. D'autre part, ce type de clé vous permettra d'obtenir de meilleures performances.

Cette codification au format **aaaammjj** devra être généralisée à tous les niveaux de notre axe temps : année, semestre, trimestre, mois, semaine...

Nous vous suggérons donc que chaque niveau (année, semestre, trimestre, mois, semaine, jour) soit composé de trois attributs distincts :

- Code
- Date

- Nom

Par exemple, le mois devra être composé des attributs suivants :

Attribut	Type	Valeur exemple	Commentaire
MoisCode	Int	20151101	Format aaaammjj, par défaut doit toujours se référer à la 1 ^{ère} date de la période
MoisDate	SmallDateTime	2015-11-01 00:00:00	Format date
MoisNom	Varchar(50)	Novembre 2015	Valeur au format affichée


Les valeurs des attributs *Code* devront être déduites à partir des valeurs attributs *Date*, à l'aide d'un script SQL de mise à jour (requête UPDATE).

Afin de finaliser l'axe temps, nous vous conseillons d'exécuter le script *FinalisationDimTemps.sql*, téléchargeable sur la page Informations générales.

Ce script réalise les opérations suivantes :

- Il renomme un certain nombre de colonnes.
- Il modifie des types de données *DateTime* en *SmallDateTime*.
- Il modifie des types de données *nvarchar* en *varchar*.
- Il crée la clé *Temps_PK* ainsi que les colonnes *AnneeCode*, *SemestreCode*, *TrimestreCode*, *MoisCode* et *SemaineCode*.
- Il remplit les valeurs pour les champs créés précédemment.
- Il modifie les colonnes pour interdire la valeur null.
- Il repositionne la clé primaire sur *Temps_PK*.

Au final, après exécution du script, la table *DimTemps* devra avoir ce formalisme :

	Nom de la colonne	Type de données	Autoriser l...
	Temps_PK	int	<input type="checkbox"/>
	Date	smalldatetime	<input type="checkbox"/>
	Jour	varchar(50)	<input type="checkbox"/>
	AnneeCode	int	<input type="checkbox"/>
	AnneeDate	smalldatetime	<input type="checkbox"/>
	AnneeNom	varchar(50)	<input type="checkbox"/>
	SemestreCode	int	<input type="checkbox"/>
	SemestreDate	smalldatetime	<input type="checkbox"/>
	SemestreNom	varchar(50)	<input type="checkbox"/>
	TrimestreCode	int	<input type="checkbox"/>
	TrimestreDate	smalldatetime	<input type="checkbox"/>
	TrimestreNom	varchar(50)	<input type="checkbox"/>
	MoisCode	int	<input type="checkbox"/>
	MoisDate	smalldatetime	<input type="checkbox"/>

Le champ *Temps_PK* devra bien entendu être une clé primaire. En revanche, exceptionnellement, n'activez pas l'incrémentement automatique pour cette table.

→ Vérifiez que la table *DimTemps* est bien remplie :

	Temps_PK	Date	Jour	AnneeCode	AnneeDate	AnneeNom	SemestreCode
1	20130623	2013-06-23 00:00:00	dimanche, juin 23 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130101
2	20130624	2013-06-24 00:00:00	lundi, juin 24 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130101
3	20130625	2013-06-25 00:00:00	mardi, juin 25 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130101
4	20130626	2013-06-26 00:00:00	mercredi, juin 26 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130101
5	20130627	2013-06-27 00:00:00	jeudi, juin 27 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130101
6	20130628	2013-06-28 00:00:00	vendredi, juin 28 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130101
7	20130629	2013-06-29 00:00:00	samedi, juin 29 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130101
8	20130630	2013-06-30 00:00:00	dimanche, juin 30 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130101
9	20130701	2013-07-01 00:00:00	lundi, juillet 01 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130701
10	20130702	2013-07-02 00:00:00	mardi, juillet 02 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130701
11	20130703	2013-07-03 00:00:00	mercredi, juillet 03 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130701
12	20130704	2013-07-04 00:00:00	jeudi, juillet 04 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130701
13	20130705	2013-07-05 00:00:00	vendredi, juillet 05 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130701
14	20130706	2013-07-06 00:00:00	samedi, juillet 06 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130701
15	20130707	2013-07-07 00:00:00	dimanche, juillet 07 2013	20130101	2013-01-01 00:00:00	Calendrier 2013	20130701

➤ Conservez précieusement la table **DimTemps** ainsi que ses données. En effet, la table **DimTemps** que vous venez de créer est parfaitement standard et réutilisable, quels que soient vos futurs projets d'entrepôt de données.