



Expert
EXPERT

Business Intelligence

avec **SQL Server 2019**
et **Power BI**

Maîtrisez les concepts
et réalisez un système décisionnel

En téléchargement



projets ETL, cube
et rapports



scripts et bases
de données



+ QUIZ

Version en ligne

OFFERTE !

pendant 1 an

Sébastien FANTINI



Les éléments à télécharger sont disponibles à l'adresse suivante :
<http://www.editions-eni.fr>
Saisissez la référence de l'ouvrage **EIBI19SQL** dans la zone de recherche
et validez. Cliquez sur le titre du livre puis sur le bouton de téléchargement.

Avant-propos

1. À qui s'adresse ce livre ? 9

2. Les prérequis 10

3. Les objectifs du livre 11

Chapitre 1
Introduction

1. Présentation du décisionnel 13

 1.1 La notion de décideur 14

 1.2 Les facteurs d'amélioration de la prise de décision 17

 1.3 L'informatique décisionnelle 20

2. La solution Microsoft BI 23

 2.1 L'offre Microsoft BI 23

 2.2 SQL Server 2019 25

 2.2.1 SQL Server Integration Services 26

 2.2.2 SQL Server Master Data Services 30

 2.2.3 SQL Server Data Quality Services 33

 2.2.4 SQL Server Analysis Services 35

 2.2.5 SQL Server Reporting Services 38

 2.3 Office 365 39

 2.4 Power BI 41

3. Microsoft BI et Azure 43

4. Notre étude de cas : Distrisys 44

5. Déroulement du projet 46

2 **Business Intelligence**

avec SQL Server 2019 et Power BI

Chapitre 2

Installation et découverte des outils SQL Server

1. Le choix des licences	51
1.1 SQL Server 2019	52
1.2 Office 365/Office 2019	53
1.3 SharePoint 2019	53
2. Architecture serveur	54
2.1 L'environnement de production	55
2.2 L'environnement de développement	61
2.3 Les environnements de recette et d'intégration	62
3. Découverte des outils SQL Server	62
3.1 SQL Server Management Studio	63
3.1.1 Connexion à des serveurs SQL Server	63
3.1.2 Modification des options de l'interface graphique	65
3.1.3 Restauration d'une base de données	66
3.2 SQL Server Data Tools	68

Chapitre 3

Réaliser son premier système décisionnel

1. Introduction	71
2. Création des tables de faits et de dimension	72
2.1 Création de l'entrepôt de données	72
2.2 Création d'une table de faits	73
2.3 Création des tables de type dimension	79
3. La dimension temps	89
4. Création des étoiles	104
5. Génération du jeu de test	112
6. Créer et utiliser simplement un cube brut	124

7.	Peaufiner le cube	145
7.1	Dimensions : hiérarchies et attributs	145
7.2	Mise en forme des mesures	169
7.3	Organisation des mesures	171
7.4	Mesures calculées	172
8.	Sécuriser l'accès au cube	177
8.1	Donner l'accès au cube	178
8.1.1	Donner l'accès aux administrateurs	178
8.1.2	Donner accès aux utilisateurs	180
8.2	Restreindre l'accès	186
8.2.1	Restreindre l'accès aux membres d'une dimension	187
8.2.2	Restreindre l'accès aux mesures	191
8.3	Pistes pour industrialiser la gestion des droits	193
9.	Le cube et la matrice dimensionnelle	195

Chapitre 4

La modélisation dimensionnelle

1.	La matrice dimensionnelle de Distrisys	199
2.	Facturation et commande client	204
2.1	Modélisation et schéma en étoile	204
2.2	Les factures	205
2.3	Le bilan de commande client	211
2.4	Le budget des ventes	219
2.5	Les perspectives	223
2.6	Les actions	228
2.7	Introduction au MDX	235
2.7.1	La requête MDX	235
2.7.2	Comparaison de valeurs à date	243
2.7.3	La somme cumulée	246

4 **Business Intelligence**

avec SQL Server 2019 et Power BI

3. Les stocks	250
3.1 Modélisation	250
3.2 Les mouvements de stock	251
3.3 La photo de stock	260
4. Autres exemples classiques de modélisation	269
4.1 Les achats	269
4.2 La finance	271
4.3 Les ressources humaines	274

Chapitre 5

Alimenter l'entrepôt de données avec SSIS

1. Découverte de SSIS	277
2. Réaliser son premier flux SSIS	283
2.1 Réaliser le chargement du budget d'un seul site	283
2.2 Charger les données de budget à partir de plusieurs fichiers Excel	297
3. Développer des flux ETL pour le décisionnel	304
3.1 Déroulement de l'exécution d'un processus ETL	304
3.2 Réaliser un flux pour charger le sas de données	306
3.3 Réaliser un flux pour charger une dimension	314
3.3.1 Cas d'une dimension standard	314
3.3.2 Cas d'une dimension en SCD	319
3.4 Réaliser un flux pour charger une table de faits	330
4. L'audit des flux ETL	336
4.1 Les objectifs de l'audit de flux ETL	336
4.2 Conception d'un système d'audit de flux	337
4.3 Exemple de flux avec audit	348
4.4 Activer la journalisation de SSIS	356

5. Gestion des paramètres de flux et mise en production	358
5.1 Paramétrage des flux	359
5.2 Création du catalogue Integration Services	362
5.3 Déploiement du projet SSIS sur le serveur de développement.	365
5.4 Les environnements	369
5.5 Mise en production du projet SSIS	375
5.6 Planifier un flux SSIS	378

Chapitre 6

Gérer les données de référence avec MDS

1. Découverte de MDS	383
1.1 Les données de référence	383
1.2 Données de référence et entrepôt de données	384
1.3 Les différents modes de gestion des données de référence.	385
1.4 Configurer Master Data Services.	389
2. Créer et organiser son premier référentiel.	396
2.1 Créer le modèle Produit	397
2.2 Créer l'entité Famille	400
2.3 Gérer les données d'une entité avec l'interface web	402
2.4 Créer un attribut basé sur un domaine.	406
2.5 Créer un attribut simple.	410
3. Travailler les données de référence avec Excel	413
3.1 Découvrir le client Excel.	413
3.2 Utiliser le client Excel.	418
3.2.1 Utiliser l'explorateur de données de référence	418
3.2.2 Alimenter l'entité Produit avec le client Excel	419
3.2.3 Gérer les requêtes	425
4. Assurer la qualité des données avec les règles d'entreprise.	427
4.1 Créer une règle simple d'unicité.	428
4.2 Créer une règle d'affectation de valeur par défaut	434
4.3 Créer une règle basée sur une condition	438

6 **Business Intelligence**

avec SQL Server 2019 et Power BI

5. Gérer la traçabilité des opérations réalisées sur le référentiel.	444
5.1 Les types de journalisation.	444
5.2 Journalisation par membre	445
5.3 Journalisation par attribut.	446
6. Mettre à disposition les données.	449
6.1 Les vues sur les entités	450
6.2 Les hiérarchies dérivées	454
7. Organiser les modèles et les entités.	458
7.1 Le cas de l'entité fantôme	461
7.2 Correspondance de référentiel	462
7.3 Référentiel Client et Géographie	464
8. Importer des données dans Master Data Services	467
8.1 Alimenter les tables intermédiaires	468
8.2 Intégrer les données des tables intermédiaires	474
8.3 Valider les données.	476
9. Gérer la sécurité	479
9.1 Donner accès à MDS	480
9.2 Définir les fonctions.	481
9.3 Définir les droits d'accès.	482

Chapitre 7

Restituer les données décisionnelles

1. Introduction	487
2. Excel	488
2.1 Utiliser Excel pour exploiter les données de son cube	489
2.2 Mettre en forme les données de son cube sous Excel.	493
2.3 Mettre en valeur des éléments précis	496
2.4 Créer des mesures calculées	498
2.5 Extraire des données de détail	502
2.6 Le mode Formule	504

3. Power BI	509
3.1 Le bon emploi de Power BI : le dashboard	509
3.2 Installer le serveur de rapports Power BI	510
3.3 Installer Power BI desktop	519
3.4 Créer un dashboard avec Power BI Desktop	521
4. Reporting Services	534
4.1 Le bon emploi de Reporting Services : le reporting automatisé de masse	534
4.2 Créer des rapports SSRS avec SSDT	536
4.2.1 Créer un projet de rapport	536
4.2.2 Créer un rapport avec l'assistant	538
4.2.3 L'environnement de travail	545
4.2.4 Travailler le rapport dans SSDT	548
4.2.5 Créer un rapport avec paramètres	552
4.2.6 Créer un lien entre deux rapports	556
4.2.7 Déployer les rapports sur le serveur de rapports	558
4.3 Créer un rapport SSRS avec le Générateur de rapports	561
4.3.1 Présentation et utilisation du Générateur de rapports	561
4.3.2 Créer le jeu de données	564
4.3.3 Créer la matrice de données	573
4.3.4 Créer un graphique de données	585
4.3.5 Les formats d'exportation du rapport	592
4.4 Le gestionnaire de rapports	593
4.5 Publication de masse	597
4.5.1 Définir la table des abonnés	598
4.5.2 Modifier la source de données	599
4.5.3 Définir l'abonnement piloté par les données	601
Conclusion et perspectives	609
Index	613

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 ces 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 Com-
mande 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 com-
mune de processus parfois très éloignés ou difficiles à rapprocher.

Au cours de ce chapitre nous allons mettre en œuvre, ensemble et progressi-
vement, 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.