# **ANALYTICS**

1. À qui s'adresse ce document ?

Ce document sur la Business Intelligence (BI) avec SQL Server 2017, s'adresse à tous **les membres de l’équipe décisionnelle.**

Le lecteur bénéficie des **retours d'expérience d’ainos** pour gagner en expertise sur les différentes étapes d'un projet décisionnel.

Ce document de spécification fonctionnel sera l'occasion pour le lecteur d'étudier les sujets suivants :  
  
- l'architecture des serveurs et le choix des licences  
- la modélisation de l'entrepôt de données  
- la conception du cube Analysis Services  
- la réalisation des différents types de flux d'alimentation ETL avec Integration Services  
- la mise en place de référentiels de données avec Master Data Services  
- l'utilisation d'Excel pour exploiter les données décisionnelles  
- la réalisation de rapports opérationnels et décisionnels avec Reporting ServicesWxw  
- et la création d'un dashboard interactif avec **Power BI**.

* + 2. Les prérequis

Le seul prérequis nécessaire est d’être familier avec la **base de données SQL Server**, sans toutefois en être un expert.

#### **Auteur(s)**

**Mohamed DIBASSI**, consultant décisionnel chez Ainos.

3. Les objectifs du spec fontionnel

Présenter le déroulement du projet.

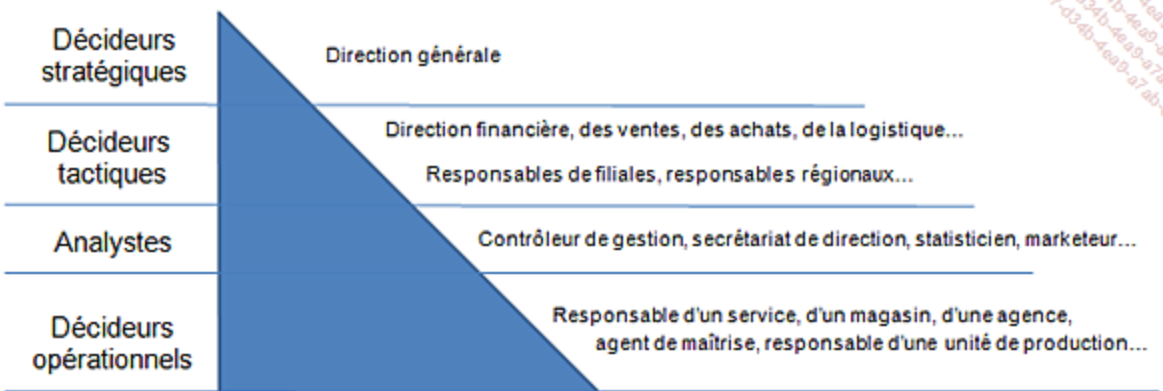
Pour cela, nous allons commencer par présenter le décisionnel, ses objectifs et les solutions délivrées par Microsoft pour y répondre.

* + 1. Présentation du décisionnel

L’informatique décisionnelle recouvre tous les moyens informatiques destinés à améliorer la prise de décision des décideurs d’une organisation.

1.1 La notion de décideur

qui sont les décideurs dans une entreprise ?



Classification des écideurs

Cette classification a de l’importance, car elle révèle de grosses différences dans le type d’outils dont chacun a besoin. Bien entendu, les choses ne sont pas aussi binaires : certaines personnes relèveront du stratégique/tactique, d’autres du tactique/opérationnel… Toutes les nuances sont permises. Ce qu’il faut bien comprendre, c’est que la notion de prise de décision pour chacun de ces décideurs n’a pas la même teinte et que chaque profil n’a pas les mêmes attentes vis-à-vis du décisionnel.

1.2 Les facteurs d'amélioration de la prise de décision

Généralement, on présente les trois facteurs de prise de décision comme étant :

* La connaissance et l’analyse du passé.
* La représentation du présent.
* L’anticipation du futur.

Les informations permettant d’appréhender ces facteurs peuvent être de deux natures différentes :

* Les informations quantitatives :
* Les informations qualitatives

Vous avez confiance en chacune des décisions que vous prenez, dès l’instant que vous avez :

* La certitude que vos données sont fiables, à jour et complètes.
* Les capacités de comprendre l’origine d’un succès ou d’une défaillance.
* Les moyens d’évaluer les plans d’action et les politiques mises en œuvre consécutivement à une décision.

Dès lors, la mise en place de l’outils décisionnels doit permettre de répondre progressivement à ces trois attentes :

* Améliorer l’accès et la qualité des données.
* Gagner en finesse d’analyse et de compréhension de données.
* Gérer les performances de l’organisation et de ses politiques.
  + - 1.3 L'informatique décisionnelle

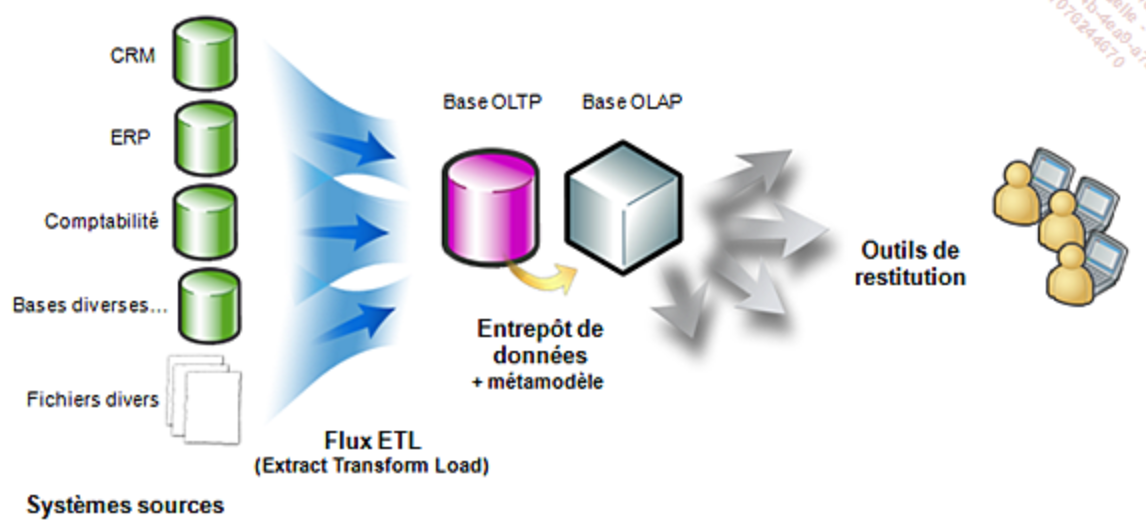
Partant du constat qu’il est difficile de croiser des données contenues dans des bases de données distinctes, le plus simple consiste à regrouper ces données éparses.

S’il est plus simple d’analyser ces données une fois qu’elles sont dans l’entrepôt de données, il n’en reste pas moins qu’il faut tout de même remplir l’entrepôt de données. L’extraction et le croisement des données des différents systèmes opérationnels puis le chargement dans l’entrepôt de données, ont fait émerger des outils dédiés à cette tâche, avec des concepts métiers qui leur sont propres : les outils d’ETL (*Extract Transform Load*).

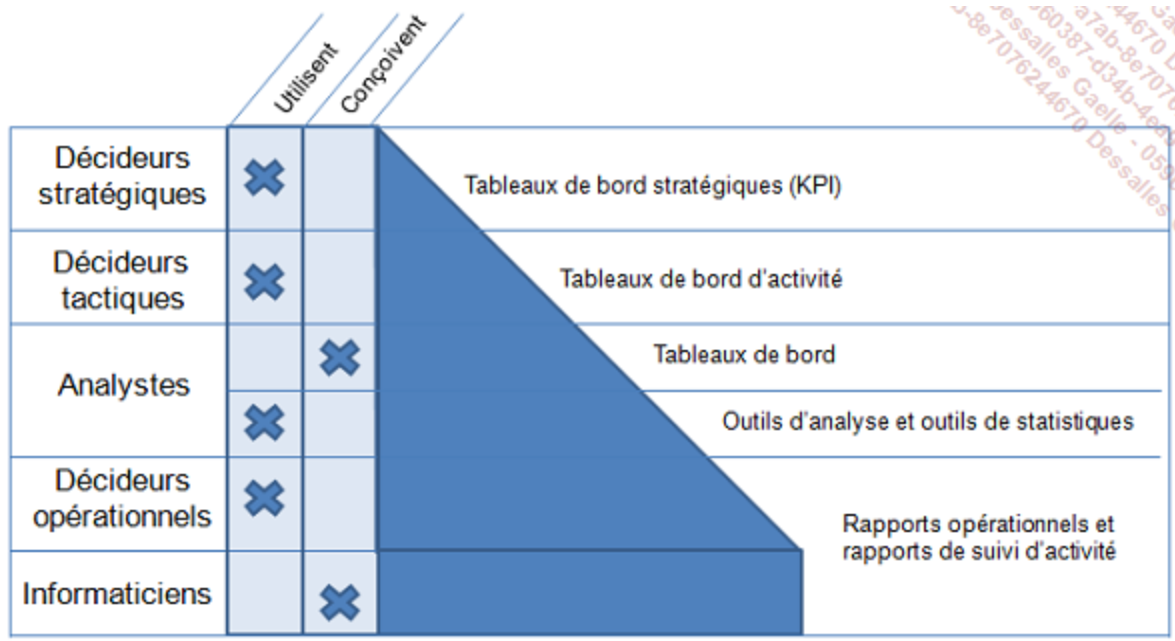
Si au début les requêtes d’analyses portaient sur une base relationnelle (dites OLTP pour *OnLine Transaction Processing*), le concept de base multidimensionnelle (dites OLAP pour *OnLine Analytical Processing*) s’est démocratisé. Ce concept de bases de données offre des performances très largement supérieures aux bases OLTP pour répondre à des requêtes d’analyse et permet également aux utilisateurs finaux de bénéficier d’une interface simplifiée d’accès aux données, beaucoup plus intuitive qu’une base de données OLTP.  On parle alors de métamodèle. Enfin, les éditeurs proposent des métamodèles montés directement en mémoire pour encore plus de performance.

NB : suivant les éditeurs, les métamodèles ne reposent pas forcément sur des bases multidimensionnelles. Les métamodèles peuvent utiliser les technologies OLTP (de moins en moins), OLAP ou "In Memory".

L’ensemble des moyens informatiques et techniques destiné à améliorer la prise de décision est appelé système décisionnel ou encore Système Informatique d’Aide à la Décision (SIAD).



*Architecture technique d’un système décisionnel*

* 
* *Les outils de restitution suivant les profils d’utilisateurs*
  + 2. La solution Microsoft BI
    - 2.1 L'offre Microsoft BI

Sur le fond, l’offre Microsoft BI est structurée autour des trois promesses du décisionnel :

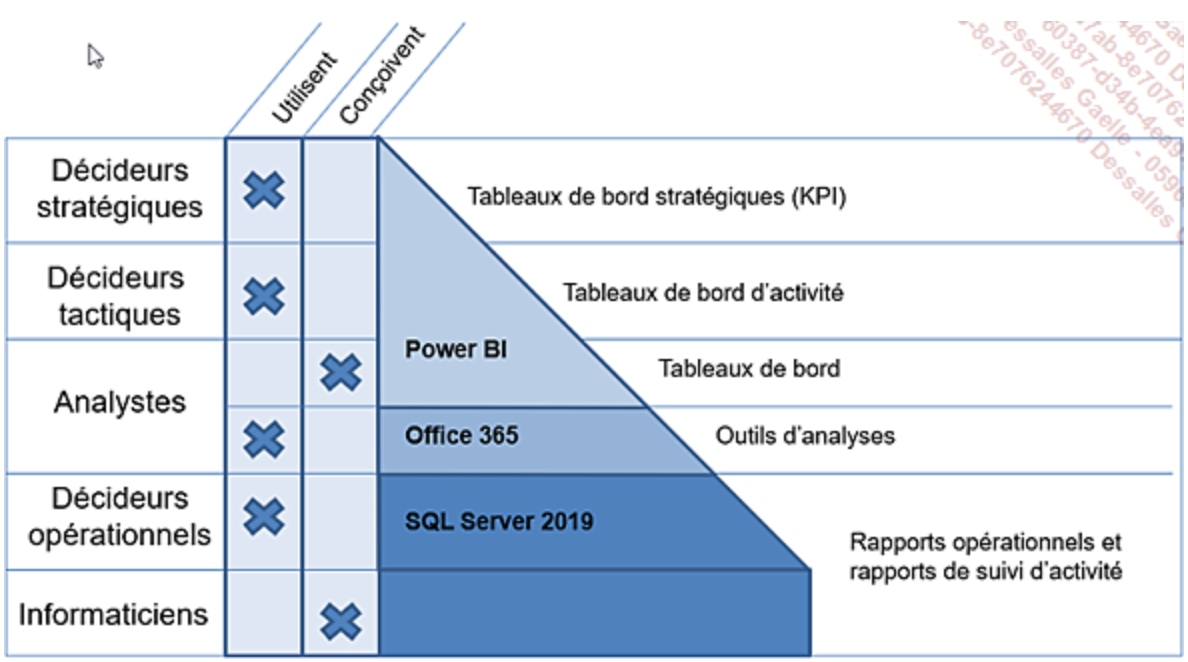
* Améliorer l’accès et la qualité des données :
* Gagner en finesse d’analyse et de compréhension de données
* Gérer les performances de l’organisation et de ses politiques :

Sur la forme, l’offre Microsoft BI est structurée au sein de trois licences. Chacune de ces licences contient de nombreux outils à usage décisionnel correspondant aux tâches énoncées plus haut :

* SQL Server 2017 : Integration Services, Master Data Services, Analysis Services, Reporting Services, Data Quality Services.
* Office 365 : Excel, SharePoint et Teams.
* Power Platform : Power BI, Power App, Power Automate.

Si les outils de la gamme SQL Server 2017 sont plutôt des produits techniques destinés, à l’usage et au service informatique, les outils Office 2019/Office 365 sont plus spécifiquement destinés aux utilisateurs finaux et les produits de la gamme Power Platform à des utilisateurs d’un niveau intermédiaire.

Dans la suite de cette partie, nous allons étudier le contenu, l’utilisation et le positionnement de chacune de ces solutions.



*L’utilisation des licences Microsoft suivant les profils utilisateurs*

2.2 SQL Server 2017

Nous avons vu précédemment que la construction d’un système décisionnel ne se limite pas à l’utilisation d’une base de données relationnelle. Avec la licence SQL Server, Microsoft nous livre toute la panoplie d’outils dont un service informatique a besoin pour bâtir un système décisionnel dans les règles de l’art.

Pour bâtir notre système d’aide à la décision, nous aurons besoin de :

* SQL Server Integration Services : l’ETL.
* SQL Server Master Data Services : le gestionnaire des données de référence.
* SQL Server Analysis Services : la base de données multidimensionnelle (OLAP) et le métamodèle.
* SQL Server Data Quality Services : la base de connaissances permettant de qualifier et nettoyer certaines données constitutives de vos référentiels de données ou des bases marketing.
* SQL Server Reporting Services : l’outil de reporting opérationnel et de reporting de masse.

Dans les sections qui suivent, nous allons présenter plus en détail chacun de ces outils.

* + - * 2.2.1 SQL Server Integration Services

La constitution des données de référence et l’alimentation d’un entrepôt de données représentent généralement près de 80 % de la charge du projet. De prime à bord, beaucoup de services informatiques ont tendance à **sous-estimer le temps nécessaire à récupérer l’information, mais aussi à la croiser**. Pour nous aider dans cette tâche, Microsoft met à notre disposition, dans la licence SQL Server 2017, un outil d’ETL (*Extract Transform Load*) : SQL Server Integration Services (SSIS).

* + - * 2.2.2 SQL Server Master Data Services

SQL Server Master Data Services (MDS) est une solution de MDM (*Master Data Management*) ou Management des données de référence.

2.2.3 SQL Server Data Quality Services

Cette solution permet de contrôler et d’améliorer la qualité des données qui circulent au sein de l’organisation à l’aide de bases de connaissances liées au contenu et à la structure de vos données.

2.2.4 SQL Server Analysis Services

SQL Server Analysis Services regroupe en fait plusieurs services destinés à faciliter l’exploration d’une quantité importante de données.

Les trois facettes de la solution sont :

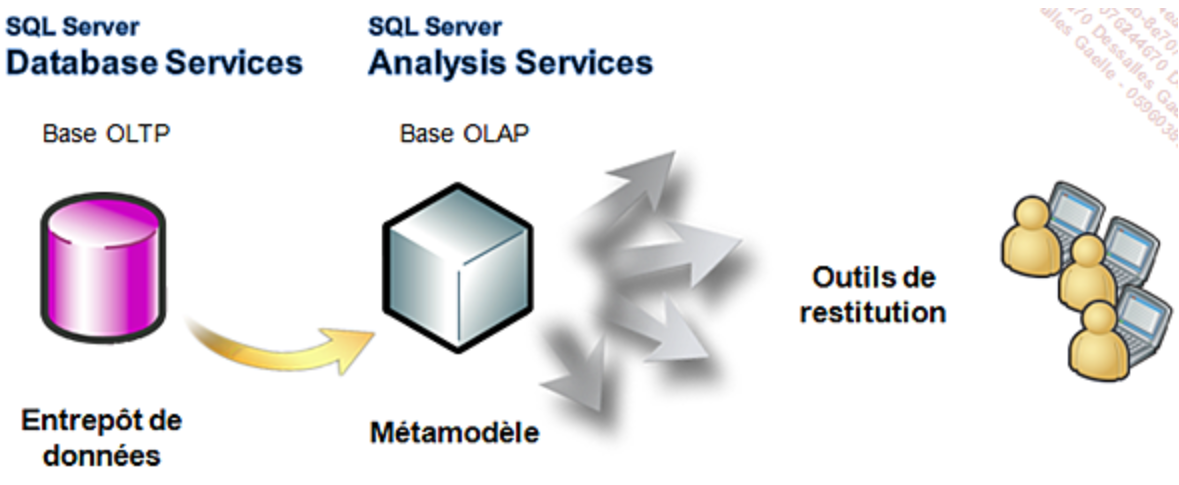
* Le mode multidimensionnel
* Le mode exploration de données ou data mining
* Le mode tabulaire

Le mode multidimensionnel

Attention toutefois à conserver une modélisation appropriée, car la performance de son moteur OLAP est directement liée à la modélisation de la base de données sous-jacente.

Nous ne saurions que trop insister tout au long sur l’importance de la modélisation de l’entrepôt de données.

Dans les faits, les utilisateurs n’accèdent pas directement à Analysis Services. Ce sont les outils de restitution qui proposent et consomment les données du cube. Il est donc indispensable, lorsque vous mettez en place SSAS, de proposer parallèlement aux utilisateurs, des outils de restitution adaptés (par exemple Excel).



*Analysis Services est l’interface entre l’entrepôt de données et les outils de restitution.*

Le mode tabulaire

Tout comme le mode multidimensionnel, le mode tabulaire offre un modèle sémantique qui se veut ouvert à tous les types de restitution : reporting, analyse, tableau de bord

Le mode tabulaire propose une autre façon de concevoir des modèles et de les stocker. Contrairement au mode multidimensionnel qui nécessite une modélisation dimensionnelle et un stockage OLAP, le mode tabulaire utilise une modélisation relationnelle et un moteur d’analyse en mémoire vive pour le stockage et le calcul.

De ce fait, le mode tabulaire va être plus proche des modèles relationnels classiques, tout en permettant une vitesse de restitution de données extrêmement rapide, de la donnée détaillée comme de la donnée agrégée.

**Pour utiliser le mode tabulaire, il vous faudra toutefois créer une instance dédiée et nous vous conseillons de dédier une machine physique et puissante à cette instance.**

C’est pourquoi nous vous suggérons de réaliser la nouvelle solution en mode multidimensionnel.

2.2.5 SQL Server Reporting Services

Enfin, la suite SQL Server dispose d’un serveur de rapports permettant d’afficher et de diffuser des informations.

Reporting Services est un excellent outil de reporting opérationnel. Elle est parfaite pour mettre en page une facture, un bon de livraison, un suivi de commande, un inventaire, un catalogue produit, la liste des clients à relancer… Tous les états dont une application de gestion a besoin sont parfaitement réalisables avec Reporting Services. Ce n’est pas un outil d’analyse destiné à des utilisateurs finaux. En revanche, il a totalement sa place dans la diffusion de rapports de masse, c’est-à-dire dans les rapports décisionnels destinés à de nombreux décideurs opérationnels.

*Par exemple, la diffusion par mail au format PDF du rapport mensuel de suivi des ventes à tous les commerciaux de l’entreprise.*

* + - 2.3 Office 2019/Office 365

Au sein de la solution BI Microsoft, Excel est l’outil d’analyse des utilisateurs finaux. Toutefois, Excel n’est utilisé que pour accéder, manipuler et naviguer dans les données d’Analysis Services. Les données ne sont plus contenues directement dans le fichier, mais sur un serveur. Et ces données ne sont plus traitées directement par les utilisateurs fonctionnels mais par le service informatique par le biais de l’ETL. Excel conserve toutefois toutes ses capacités de représentation graphique, de mise en page et de personnalisation à l’aide de formules.

2.4 Power BI

Power BI permet de fournir des interfaces interactives de visualisation des données. Il est tourné vers des utilisateurs non informaticiens, leur permettant à l’aide de l’outil Power BI desktop de concevoir des dashboards interactifs très riches. Contrairement à Excel, qui n’est pas conçu pour partager des visualisations à une équipe ou à son organisation, Power BI est vraiment pensé pour cela.

Nous restituerons les données décisionnelles - Power BI en déployant un serveur de rapports Power BI sans en passer par le cloud.

Cependant, Power BI n’est pas un outil d’analyse. Il est très efficace pour présenter l’information, mais très peu performant pour l’explorer. Excel reste en cela incontournable.

Il serait intéressant de prévoir une formation sur la création d’un dashboard Power BI, et à le partager au sein de votre organisation.

* + 3. Microsoft BI et Azure

La conception des projets décisionnels dans le cloud avec succès est aujourd’hui possible. Il y a toutefois des problématiques autour du transfert de données et de l’authentification qui s’ajoutent à ceux d’un projet On-Premise. Bien entendu, une fois ces problèmes surmontés, le système décisionnel bénéficie de toute la puissance du cloud en termes d’élasticité, de disponibilité et de sécurité.

C’est pourquoi, nous suggérons de nous consacrer à la mise en place d’un système décisionnel entièrement On-premise.

# **Déroulement du projet**

**Étant donné notre expérience, nous constatons généralement que la première intention d’un service informatique souhaitant réaliser un système décisionnel est de partir des données à leur disposition. Si ce réflexe est parfaitement compréhensible, il est l’opposé de ce qu’il faut faire si l’on souhaite mener à bien un projet et obtenir des résultats.**

**Un projet décisionnel nécessite, tout d’abord et forcément, d’être mené en trio : l’IT représenté par l’équipe BI, un membre du service fonctionnel concerné par le périmètre traité et enfin un membre du contrôle de gestion. Ce projet doit si possible s’inscrire dans un projet d’entreprise disposant du sponsoring ou du moins du soutien d’un membre influent de la direction.**

**Ce mode de fonctionnement est essentiel, car il est indispensable de mener chaque étape du projet du haut vers le bas, c’est-à-dire du besoin global, le plus stratégique, vers la donnée la plus détaillée**.

* + 1. À qui s'adresse ce livre ?
  + 2. Les prérequis
  + 3. Les objectifs du livre
* Introduction
  + 1. Présentation du décisionnel
    - 1.1 La notion de décideur
    - 1.2 Les facteurs d'amélioration de la prise de décision
    - 1.3 L'informatique décisionnelle
  + 2. La solution Microsoft BI
    - 2.1 L'offre Microsoft BI
    - 2.2 SQL Server 2017
      * 2.2.1 SQL Server Integration Services
      * 2.2.2 SQL Server Master Data Services
      * 2.2.3 SQL Server Data Quality Services
      * 2.2.4 SQL Server Analysis Services
      * 2.2.5 SQL Server Reporting Services
    - 2.3 Office 365
    - 2.4 Power BI
  + 3. Microsoft BI et Azure
  + 4. Notre étude de cas : Distrisys
  + 5. Déroulement du projet
* Installation et découverte des outils SQL Server
  + 1. Le choix des licences
    - 1.1 SQL Server 2017
    - 1.2 Office 365/Office 2017
    - 1.3 SharePoint 2017
  + 2. Architecture serveur
    - 2.1 L'environnement de production
    - 2.2 L'environnement de développement
    - 2.3 Les environnements de recette et d'intégration
  + 3. Découverte des outils SQL Server
    - 3.1 SQL Server Management Studio
      * 3.1.1 Connexion à des serveurs SQL Server
      * 3.1.2 Modification des options de l'interface graphique
      * 3.1.3 Restauration d'une base de données
    - 3.2 SQL Server Data Tools
* Réaliser son premier système décisionnel
  + 1. Introduction
  + 2. Création des tables de faits et de dimension
    - 2.1 Création de l'entrepôt de données
    - 2.2 Création d'une table de faits
    - 2.3 Création des tables de type dimension
  + 3. La dimension temps
  + 4. Création des étoiles
  + 5. Génération du jeu de test
  + 6. Créer et utiliser simplement un cube brut
  + 7. Peaufiner le cube
    - 7.1 Dimensions : hiérarchies et attributs
    - 7.2 Mise en forme des mesures
    - 7.3 Organisation des mesures
    - 7.4 Mesures calculées
  + 8. Sécuriser l'accès au cube
    - 8.1 Donner l’accès au cube
      * 8.1.1 Donner l’accès aux administrateurs
      * 8.1.2 Donner accès aux utilisateurs
    - 8.2 Restreindre l’accès
      * 8.2.1 Restreindre l'accès aux membres d’une dimension
      * 8.2.2 Restreindre l'accès aux mesures
    - 8.3 Pistes pour industrialiser la gestion des droits
  + 9. Le cube et la matrice dimensionnelle
* La modélisation dimensionnelle
  + 1. La matrice dimensionnelle de Distrisys
  + 2. Facturation et commande client
    - 2.1 Modélisation et schéma en étoile
    - 2.2 Les factures
    - 2.3 Le bilan de commande client
    - 2.4 Le budget des ventes
    - 2.5 Les perspectives
    - 2.6 Les actions
    - 2.7 Introduction au MDX
      * 2.7.1 La requête MDX
      * 2.7.2 Comparaison de valeurs à date
      * 2.7.3 La somme cumulée
  + 3. Les stocks
    - 3.1 Modélisation
    - 3.2 Les mouvements de stock
    - 3.3 La photo de stock
  + 4. Autres exemples classiques de modélisation
    - 4.1 Les achats
    - 4.2 La finance
    - 4.3 Les ressources humaines
* Alimenter l'entrepôt de données avec SSIS
  + 1. Découverte de SSIS
  + 2. Réaliser son premier flux SSIS
    - 2.1 Réaliser le chargement du budget d'un seul site
    - 2.2 Charger les données de budget à partir de plusieurs fichiers Excel
  + 3. Développer des flux ETL pour le décisionnel
    - 3.1 Déroulement de l'exécution d'un processus ETL
    - 3.2 Réaliser un flux pour charger le sas de données
    - 3.3 Réaliser un flux pour charger une dimension
      * 3.3.1 Cas d'une dimension standard
      * 3.3.2 Cas d'une dimension en SCD
    - 3.4 Réaliser un flux pour charger une table de faits
  + 4. L’audit des flux ETL
    - 4.1 Les objectifs de l'audit de flux ETL
    - 4.2 Conception d'un système d'audit de flux
    - 4.3 Exemple de flux avec audit
    - 4.4 Activer la journalisation de SSIS
  + 5. Gestion des paramètres de flux et mise en production
    - 5.1 Paramétrage des flux
    - 5.2 Création du catalogue Integration Services
    - 5.3 Déploiement du projet SSIS sur le serveur de développement
    - 5.4 Les environnements
    - 5.5 Mise en production du projet SSIS
    - 5.6 Planifier un flux SSIS
* Gérer les données de référence avec MDS
  + 1. Découverte de MDS
    - 1.1 Les données de référence
    - 1.2 Données de référence et entrepôt de données
    - 1.3 Les différents modes de gestion des données de référence
    - 1.4 Configurer Master Data Services
  + 2. Créer et organiser son premier référentiel
    - 2.1 Créer le modèle Produit
    - 2.2 Créer l'entité Famille
    - 2.3 Gérer les données d'une entité avec l'interface web
    - 2.4 Créer un attribut basé sur un domaine
    - 2.5 Créer un attribut simple
  + 3. Travailler les données de référence avec Excel
    - 3.1 Découvrir le client Excel
    - 3.2 Utiliser le client Excel
      * 3.2.1 Utiliser l'explorateur de données de référence
      * 3.2.2 Alimenter l'entité Produit avec le client Excel
      * 3.2.3 Gérer les requêtes
  + 4. Assurer la qualité des données avec les règles d'entreprise
    - 4.1 Créer une règle simple d'unicité
    - 4.2 Créer une règle d'affectation de valeur par défaut
    - 4.3 Créer une règle basée sur une condition
  + 5. Gérer la traçabilité des opérations réalisées sur le référentiel
    - 5.1 Les types de journalisation
    - 5.2 Journalisation par membre
    - 5.3 Journalisation par attribut
  + 6. Mettre à disposition les données
    - 6.1 Les vues sur les entités
    - 6.2 Les hiérarchies dérivées
  + 7. Organiser les modèles et les entités
    - 7.1 Le cas de l’entité fantôme
    - 7.2 Correspondance de référentiel
    - 7.3 Référentiel Client et Géographie
  + 8. Importer des données dans Master Data Services
    - 8.1 Alimenter les tables intermédiaires
    - 8.2 Intégrer les données des tables intermédiaires
    - 8.3 Valider les données
  + 9. Gérer la sécurité
    - 9.1 Donner accès à MDS
    - 9.2 Définir les fonctions
    - 9.3 Définir les droits d’accès
* Restituer les données décisionnelles
  + 1. Introduction
  + 2. Excel
    - 2.1 Utiliser Excel pour exploiter les données de son cube
    - 2.2 Mettre en forme les données de son cube sous Excel
    - 2.3 Mettre en valeur des éléments précis
    - 2.4 Créer des mesures calculées
    - 2.5 Extraire des données de détail
    - 2.6 Le mode Formule
  + 3. Power BI
    - 3.1 Le bon emploi de Power BI : le dashboard
    - 3.2 Installer le serveur de rapports Power BI
    - 3.3 Installer Power BI desktop
    - 3.4 Créer un dashboard avec Power BI Desktop
  + 4. Reporting Services
    - 4.1 Le bon emploi de Reporting Services : le reporting automatisé de masse
    - 4.2 Créer des rapports SSRS avec SSDT
      * 4.2.1 Créer un projet de rapport
      * 4.2.2 Créer un rapport avec l'assistant
      * 4.2.3 L'environnement de travail
      * 4.2.4 Travailler le rapport dans SSDT
      * 4.2.5 Créer un rapport avec paramètres
      * 4.2.6 Créer un lien entre deux rapports
      * 4.2.7 Déployer les rapports sur le serveur de rapports
    - 4.3 Créer un rapport SSRS avec le Générateur de rapports
      * 4.3.1 Présentation et utilisation du Générateur de rapports
      * 4.3.2 Créer le jeu de données
      * 4.3.3 Créer la matrice de données
      * 4.3.4 Créer un graphique de données
      * 4.3.5 Les formats d'exportation du rapport
    - 4.4 Le gestionnaire de rapports
    - 4.5 Publication de masse
      * 4.5.1 Définir la table des abonnés
      * 4.5.2 Modifier la source de données
      * 4.5.3 Définir l'abonnement piloté par les données
  + Conclusion et perspectives
  + Index