

École Supérieure Privée de Management de Tunis
Esprit School of Business



RAPPORT DE STAGE

en vue de l'obtention du diplôme de Licence Fondamentale en
Informatique de Gestion (option Systèmes d'Information Décisionnels)

Réalisé par: **Oussama ZRIBI**

Mise en place d'un système d'aide à la décision au
niveau du système d'information MS Dynamics Nav

Avril – Juillet 2019

Maître de stage : Mohamed Amine Djobbi

Encadrant académique : Mohamed Aymen Haj Kacem

	<i>Nom & Prénom</i>	<i>Date et Signature</i>
Maître de stage	<i>Mohamed Amin Djobbi</i>	
Encadrant académique	Mohamed Aymen Hadj Kacem	

Avec tout respect et amour je dédie ce modeste travail

À mes chers parents, la source de ma réussite

À ma petite sœur Salima

Pour tout leur soutien moral, leur amour et affection

À tous mes amis

en Souvenir des plus beaux instants qu'on a passé ensemble

Aussi bien à tous ceux qui m'ont aidé

Merci

Remerciement

Je profite par le biais de ce rapport, pour exprimer mes vifs remerciements à toute personne contribuant de près ou de loin à l'élaboration de cet humble travail.

En premier lieu, je tiens à remercier mon maître de stage **M. Mohamed Amin Djobbi**, directeur de l'activité Microsoft. Un grand merci pour son accueil chaleureux, ainsi que pour sa patience et ses précieux conseils. Il m'a beaucoup appris sur l'entreprise et les défis qu'un directeur doit relever au quotidien.

Nous remercions également tous les employés de Discovery Informatique pour leur hospitalité et leur chaleureux accueil.

C'est avec une grande reconnaissance et gratitude que je tiens à remercier mon professeur M. Aymen Bel Hadj Kacem pour son aide et pour ses conseils enrichissants.

Je souhaite ensuite adresser mes remerciements au corps professoral et administratif de l'École Esprit School of Business, pour la qualité de l'enseignement offert et le soutien de l'équipe administrative.

A mon père **Mourad**,

Je veux te remercier mon père pour tout cet amour donné depuis ma naissance jusqu'à aujourd'hui, pour ses conseils et la confiance qu'il m'accorde quotidiennement, ainsi pour leur soutien inconditionnel, à la fois moral et économique

*A ma mère **Khadija**,*

Maman, il y a tant de choses que j'aimerais te dire mais les mots ont parfois du mal à franchir mes lèvres. Alors je te les écris. Même si tu sais ce que je ressens, j'éprouve aujourd'hui cette envie, ce besoin de laisser sur papier une trace de mon amour pour toi

Enfin, mes remerciements vont également aux membres du jury d'avoir accepté de juger mon travail.

Table des matières

REMERCIEMENT	5
INTRODUCTION GENERALE :	10
CHAPITRE 1 : CONTEXTE DE TRAVAIL	12
I. INTRODUCTION :	13
II. PRESENTATION DE L'ENTREPRISE :	13
III. DOMAINE D'ACTIVITE :	14
IV. CONTEXTE DE PROJET :	14
1. CADRE DE PROJET :	14
2. ETUDE DE L'EXISTANT :	15
3. PROBLEMATIQUE :	16
4. SOLUTION PROPOSEE :	16
CHAPITRE 2 : METHODOLOGIE DE TRAVAIL	17
INTRODUCTION	18
I. METHODOLOGIE DE TRAVAIL EXISTANTE.....	18
1. LA METHODE PREDICTIVE CLASSIQUE « CYCLE DE VIE EN V »	18
2. LA METHODE AGILE BI.....	19
3. COMPARAISON DES METHODES.....	20
4. PRESENTATION DE LA METHODE SCRUM	20
<i>i. Les intervenants dans SCRUM</i>	20
<i>ii. Les artéfacts du Scrum</i>	21
II. ANALYSE COMPARATIVE ENTRE L'APPROCHE TOP-DOWN ET BOTTOM-UP	22
1. L'APPROCHE DE BILL INMON	22
2. L'APPROCHE DE RALPH KIMBALL.....	23
3. ANALYSE COMPARATIVE ENTRE LES DEUX APPROCHES	24
III. ANALYSE COMPARATIVE ENTRE LES DIFFERENTS TYPES DE MODELE	25
1. MODELE EN ETOILE.....	25
2. MODELE EN FLOCON DE NEIGE.....	26
IV. ETUDE COMPARATIVE ENTRE LES DIFFERENTS Outils BI	27
1. QLIKVIEW	27
2. TABLEAU	27
3. POWER BI	28
V. RECAPITULATIF DE CHOIX	29
1. CHOIX DE METHODOLOGIE	29
2. CHOIX DE LA DEMARCHE	29
<i>i. Choix du type du type de modèle</i>	29
<i>ii. Choix de l'outil</i>	29
CONCLUSION	31
CHAPITRE 3 : DEFINITION DES BESOINS ET PLANIFICATION DU PROJET	32
INTRODUCTION	33

I. SPECIFICATION DE BESOIN	33
II. ORGANISATION DU PROJET A L'AIDE DE SCRUM	38
1. LE BACKLOG PRODUIT.....	38
2. PLANIFICATION DES SPRINTS.....	39
<i>i. Pilotage du projet avec SCRUM</i>	<i>39</i>
<i>ii. La partition de la solution en Sprints.....</i>	<i>40</i>
<i>iii. Diagramme de Gantt</i>	<i>40</i>
<i>iv. Les risques de projet.....</i>	<i>41</i>
<i>v. Les contraintes de projet</i>	<i>41</i>
<i>vi. Etude des données sources</i>	<i>41</i>
CONCLUSION	43
CHAPITRE 4	44
SPRINT 1 : MODELISATION ET MISE EN PLACE DU DATA WAREHOUSE (ETL)	44
INTRODUCTION	45
I. QU'EST-CE QU'UN ETL ?	45
II. EXTRACT, TRANSFORM, LOAD (ETL).....	45
III. LE SPRINT BACKLOG.....	46
1. EXTRACT	47
2. TRANSFORM	48
3. LOAD.....	48
IV. PACKAGE SSIS ET CONFIGURATION JOB ETL.....	49
1. STAGING AREA « STG »	50
2. OPERATIONAL DATA STORE « ODS »	53
3. DATA WAREHOUSE « DW ».....	57
4. DATAMARTS « DMT »	58
<i>i. L'avantage de datamart</i>	<i>58</i>
<i>ii. Mise en place des cubes.....</i>	<i>59</i>
V. PACKAGE SSAS ET MODELISATION DE DONNEES	62
1. PRESENTATION DE SQL SERVER ANALYSIS SERVICES (SSAS).....	62
2. PACKAGE SSAS.....	62
CHAPITRE 5	64
SPRINT 2 : ELABORATION DES TABLEAUX DE BORD	64
INTRODUCTION	65
1. PRESENTATION DE SQL SERVER REPORTING SERVICES (SSRS)	65
2. REALISATION DES TABLEAUX DE BORD.....	66
CONCLUSION GENERALE	67
REFERENCE	69
ANNEXES	70

Table des figures

Figure 1:logo Discovery Informatique	13
Figure 2 :Les solutions de gestion ERP	14
Figure 3: les prévisions de tableau de bord (dept. Direction générale)	15
Figure 4 : Cycle de vie en V	18
Figure 5: Méthodologie Agile appliquée à la BI	19
Figure 6: Déroulement projet Scrum	21
Figure 7 : Approche Bill Inmon	23
Figure 8 : Approche Ralph Kimball	24
Figure 9 :Schéma en étoile	25
Figure 10 : Schéma en flocon de neige	26
Figure 11 : Logo Qlikview	27
Figure 12 : logo Tableau	27
Figure 13 : classification des outils BI (Source : Gartner fév 2019)	28
Figure 14 : logo Visual Studio	29
Figure 15 : logo Microsoft Dynamics Navision	29
Figure 16 : Logo SSMS	30
Figure 17 : Logo Power BI	30
Figure 18 : récapitulatif du choix d'outils dans le processus décisionnel	30
Figure 19 : Les Sprints de projet	40
Figure 20 :Diagramme de Gantt	40
Figure 21 : Identifier la table par son numéro	41
Figure 22 : procédure de SSMS	42
Figure 23 : Extract, Transform, Load [15]	45
Figure 24 : Nouveau projet SSIS	49
Figure 25 : Processus de ETL	49
Figure 26 : Staging Area	51
Figure 27 : Extraction de donnée pour charge produit et comptabilité générale	51
Figure 28 : Connexion OLE DB pour Staging Area	52
Figure 29 : Connexion Operational Data Store	53
Figure 30 : Operational Data Store	54
Figure 31 : Requête SQL Truncate Table	54
Figure 32: Jointure des tables entête et ligne	55
Figure 33 : Table Purchase Reception ODS	55
Figure 34 : Nouvelle Table Chiffre d'affaires achat	56
Figure 35 : colonne dérivée avoir achat	56
Figure 36 : Data warehouse	57
Figure 37 : Datamarts	59
Figure 38 : Datamart date	59
Figure 39 : Fact Table	60
Figure 40 : Item datamart	60
Figure 41 : Création d'un nouveau Job	61
Figure 42 : Planification du lancement package	61
Figure 43 : Modélisation des données	62
Figure 44 :Sélection des dimensions et table de fait	65
Figure 45 : Tableau de bord Vente	Erreur ! Signet non défini.
Figure 46 : Etat de récapitulatif des Achat	Erreur ! Signet non défini.
Figure 47: Script Date	70

Table des tableaux

Tableau 1: Tableau comparatif entre la méthode traditionnelle et agile	20
Tableau 2 : Tableau comparatif des deux approches	24
Tableau 3 : Tableau besoin fonctionnel département contrôle de gestion	36
Tableau 4 : Tableau besoin fonctionnel Direction Général	37
Tableau 5 : Backlog produit	39
Tableau 6 : Backlog Sprint 1	46

0

Introduction générale :

Dans un monde ultra-concurrentiel où la recherche et l'innovation font la différence, toute entreprise quel que soit son secteur d'activité est de plus attentive à sa performance, pour être toujours en phase avec les attentes de ses clients, tout en maintenant une longueur d'avance sur ses concurrents actuels et potentiels.

Nous observons parfois comme un flux bruyant de nouvelles idées et produits, et souvent avec affirmations exagérées sur l'importance et l'effet qui peuvent métamorphoser totalement le mode de vie et les choix des clients.

C'est dans ce contexte que le domaine 'Business Intelligence' ou l'informatique décisionnel (en français) trouve son essence en offrant des passerelles vers les solutions IT apportant une aide à la décision aux professionnels avec, en bout de chaîne, des rapports et tableaux de bord de suivi des activités de l'entreprise à la fois analytiques et prospectifs leurs permettant ainsi de prendre des décisions fines et lancer les bonnes actions.

Les prémisses de la BI apparaissent fin 1970 (bien que le terme soit utilisé par **Hans Peter Luhn**¹ en 1958) avec l'apparition des premiers infocentres. Pour rappel un infocentre sert à prendre des décisions opérationnelles basées sur des valeurs courantes. L'informatique décisionnel est en général l'informatique à l'usage des décideurs et des dirigeants des entreprises. Les systèmes BI sont utilisés par les décideurs pour obtenir une connaissance approfondie de l'entreprise, définir et soutenir leurs stratégies d'affaires dans l'objectif d'acquérir un avantage concurrentiel, d'améliorer la performance de l'entreprise et de répondre plus rapidement aux changements.

Les mutations continuent d'affecter le paysage des entreprises, l'innovation des technologies et son évolution évoque des nouveaux outils pour gérer, analyser et décider dans une entreprise.

Pour faire simple les personnes décideuses doivent disposer d'un environnement simple d'utilisation, rapide pouvant traiter des volumes de données colossaux sans interaction avec la production, de cette problématique va naître la BI tel que nous la connaissons aujourd'hui.

¹ **Hans Peter Luhn** est un informaticien allemand.

Ce présent travail a comme objectif de réaliser en premier lieu, une solution BI pour l'entreprise :

Dans ce contexte, la première étape consiste à réaliser un état des lieux du système existant qui permettra de sensibiliser les principaux acteurs de l'entreprise aux enjeux du système de mesures qui va être créé.

Le diagnostic de départ permet d'identifier dès le début les failles ou limites du dispositif existant et de se mettre d'accord sur les caractéristiques du nouveau système.

Cette étude est faite pour le SI² de « Discovery Informatique » en exploitant les données de la base de données de l'ERP Microsoft Dynamics Navision qui sera par la suite adaptée aux besoins spécifiques créant ainsi un Data Warehouse permettant de contrôler les achats, les ventes et les projets tout en offrant une meilleure visibilité sur le chiffre d'affaire selon des indicateurs et des axes bien déterminés.

Ce rapport est organisé comme suit :

Le premier chapitre s'intitule 'Contexte de travail' qui contient la présentation de l'organisme d'accueil, l'étude de l'existant, problématique et la solution proposée.

Le deuxième chapitre 'Méthodologie de travail' contient la justification de choix et l'environnement de travail adopté et le type de modèle multidimensionnel.

Le troisième chapitre 'Définition des besoins et planification du projet' contiendra la définition des besoins fonctionnels ainsi que besoin de l'entreprise.

Les chapitres, qui suivent, décrivent les différentes Sprint en plus de détails.

Terminant, bien sûr, par une conclusion générale.

² SI : Système interne

Chapitre 1 : Contexte de travail

I. Introduction :

Dans ce chapitre, nous décrivons le contexte général du projet. Nous présentons d'abord l'entreprise au sein de laquelle se déroule le stage et nous enchainons par la suite par l'exposé du sujet en expliquant ses différents modules.

II. Présentation de l'entreprise :



Figure 1:logo Discovery Informatique

Depuis plus d'une vingtaine d'années, **Discovery Informatique** n'a cessé de développer le secteur de l'ingénierie logicielle et l'intégration des solutions de gestion intégrée ERP et des solutions de gestion d'entreprise. Elle est notamment la première à introduire le concept ERP en Tunisie en 1997.

Le Groupe Discovery est spécialisé dans la transformation digitale des entreprises et des institutions, et ce dans les régions du Maghreb, d'Afrique sub-saharienne et d'Europe.

Il opère dans l'intégration de solutions de management et d'ERP où il est Leader en Afrique, ainsi que dans le développement de plateformes et d'applications digitales pour le marché européen.

Fort de ses 150 ressources humaines hautement qualifiées et des certifications reconnues dans les nouvelles technologies digitales, le groupe offre à ses 500 clients des solutions digitales intelligentes pour mettre en œuvre une vraie transformation du système de management de leurs entreprises, combinant agilité, mobilité et performance. [1]

III. Domaine d'activité :

Les activités principales de Discovery sont l'intégration de solution de gestion d'entreprises ERP, le développement informatique et l'infogérance.

- Intégrations de solution de gestion d'entreprises ERP :



Figure 2 :Les solutions de gestion ERP

IV. Contexte de projet :

1. Cadre de projet :

Dans une démarche d'amélioration continue, l'entreprise nécessite des tableaux de bords pour chaque département selon leur besoin spécifique. En effet, c'est une implémentation d'une solution décisionnelle et la diffuser.

Ce projet se décompose en plusieurs parties dépendants ; la collection, la consolidation, la modélisation et la restitution des données, qui seront décrites en plus de détails, afin de permettre au décideur d'avoir une vue d'ensemble de l'activité traitée.

2. Etude de l'existant :

Au début, la vision n'était pas claire sur l'existant pour deux raisons ; les KPI³s sont de plus en plus vagues et c'était ma première expérience professionnelle dans ce domaine, ce qui m'a poussé à découvrir ce nouveau monde et à écarter cette obscurité.

Tout à commencer par des réunions avec les divers départements pour connaitre vraiment son besoin et savoir quels sont les KPIs plus pertinents pour son domaine d'activité qu'on termine par avoir un procès-verbal et nous les accordons une maquette jusqu'ils la valident. Chaque département a ces états qui sont élaborés manuellement sur des fichiers Excel issus des différentes structures métiers de l'entreprise. D'autre part, le fait que le système d'information exploité actuellement par l'entreprise n'est pas intégrée, les structures récupèrent donc les données de différentes solutions, puis elles les traitent sur Excel en rajoutant des formules. Ceci peut engendrer quelques limites, ce qui implique les interrogations suivantes de la part de l'entreprise :

- Quelle est l'activité la plus rentable par rapport à l'année X?
- Quelle est le projet le plus couteux de terme charge ?
- Quelle est le taux d'occupation des consultants pour chaque projet ?

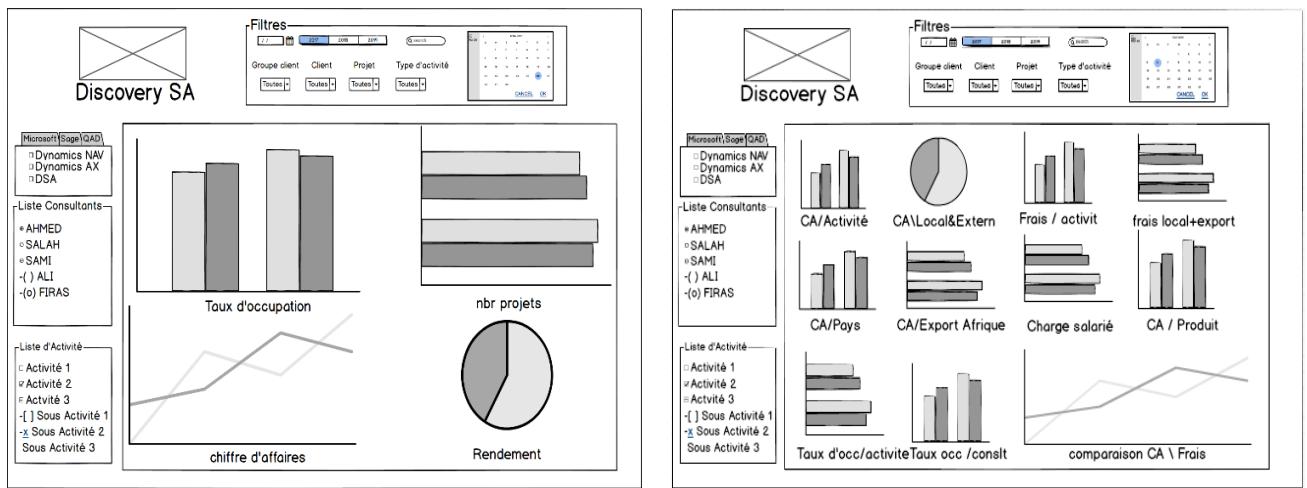


Figure 3: les prévisions de tableau de bord (dept. Direction générale)

³ KPI : indicateur de performance clé

3. Problématique :

A ce jour, de nombreuses entreprises, effectuent manuellement des requêtes assez complexes sur une ou plusieurs bases de données spécification en utilisant l'outil Excel, dont le but d'élaborer des tableaux de bords et d'extraire des informations parfois éparpillées dans plusieurs systèmes hybrides non centralisés.

L'administration générale demande des analyses sur plusieurs axes et des visualisations des données directement de le SI d'information actuel (microsoft dynamics NAV). En même temps, certes tâches et calculs seront à la fois automatisés et instantanés, ce qui rend à l'entreprise le gain de temps et la fiabilité des calculs.

4. Solution proposée :

Afin de faciliter le pilotage et la prise de décision, une solution doit être mise en œuvre, faciliter le suivi et la visualisation des données. Celle-ci va permettre d'avoir une vue globale sur toute la procédure. L'objectif étant de mettre en place une solution décisionnelle qui va contenir des tableaux de bord interactifs permettant aux décideurs de faire un suivi et une analyse de tout le processus, mais également d'accéder à toutes les informations qui seront bien organisées et toujours à portée de main.

5. Conclusion :

L'étude de l'existant et la problématique nous ont permis d'arracher une solution décisionnelle qui doit répondre à toute les besoins des décideurs.

Avant de présenter la solution, nous allons expliquer notre choix quant aux méthodes de gestion de travail ainsi que les outils utilisés.

Chapitre 2 : Méthodologie de travail

Introduction

Dans ce chapitre, nous allons faire une étude comparative entre plusieurs volets. Commençant par les méthodes de gestion des projets, les approches de la conception du Data Warehouse et enfin explication de choix.

I. Méthodologie de travail existante

Dans ce chapitre, nous posons la question suivante :

Quelle est la méthode la plus adapté pour ce projet ?

Parmi les méthodes fréquemment utilisées :

1. La méthode prédictive classique « cycle de vie en V »

Les méthodes traditionnelles se caractérisent par une approche prédictive. Cette méthode exige l'équipe technique à suivre à lettre le cahier de charge et travailler sur l'ensemble de projet. Il existe une seule condition qui permet d'adapter le projet aux méthodes prédictives, c'est lorsque le client possède des idées claires et bien précises. [2]

Le cycle en V se déroule en 3 phases :

- La phase de conception
- La phase de réalisation
- La phase de validation

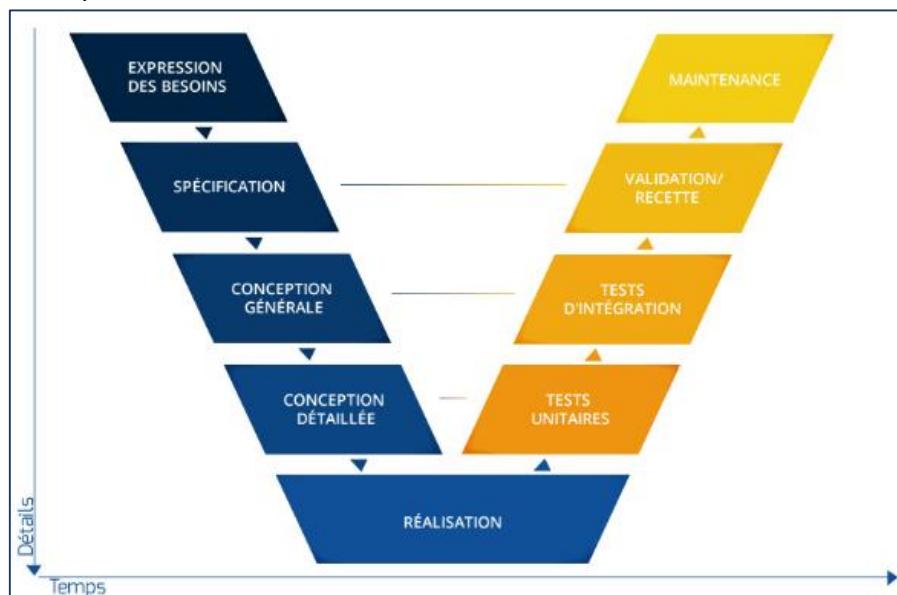


Figure 4 : Cycle de vie en V

Le seul inconvénient qui se pose en adoptant cette méthode est que les utilisateurs n'ont pas de véritables interventions sur leur propre produit puisque le projet sera livré à la fin de cycle.

2. La méthode Agile BI

L'agilité se définit comme une souplesse dans l'exécution des mouvements ou une vivacité intellectuelle. Les méthodes agiles utilisent le principe de développement itératif qui consiste à découper le projet en plusieurs étapes ou d'une série d'itération qu'on appelle « **Sprint** ».

Ces méthodes sont plus efficaces et moins rigides, elles sont particulièrement adaptées pour les projets nécessitant une grande flexibilité. En effet, la réalisation des tâches se fait avec des cycles courts généralement de deux à quatre semaines. [3]

La solution est livrée par itérations successives (Sprint) de courte durée avec des livraisons (validation successives) dont le contenu a été bien défini au préalable de commun accord entre les acteurs comme indique la figure ci-dessous

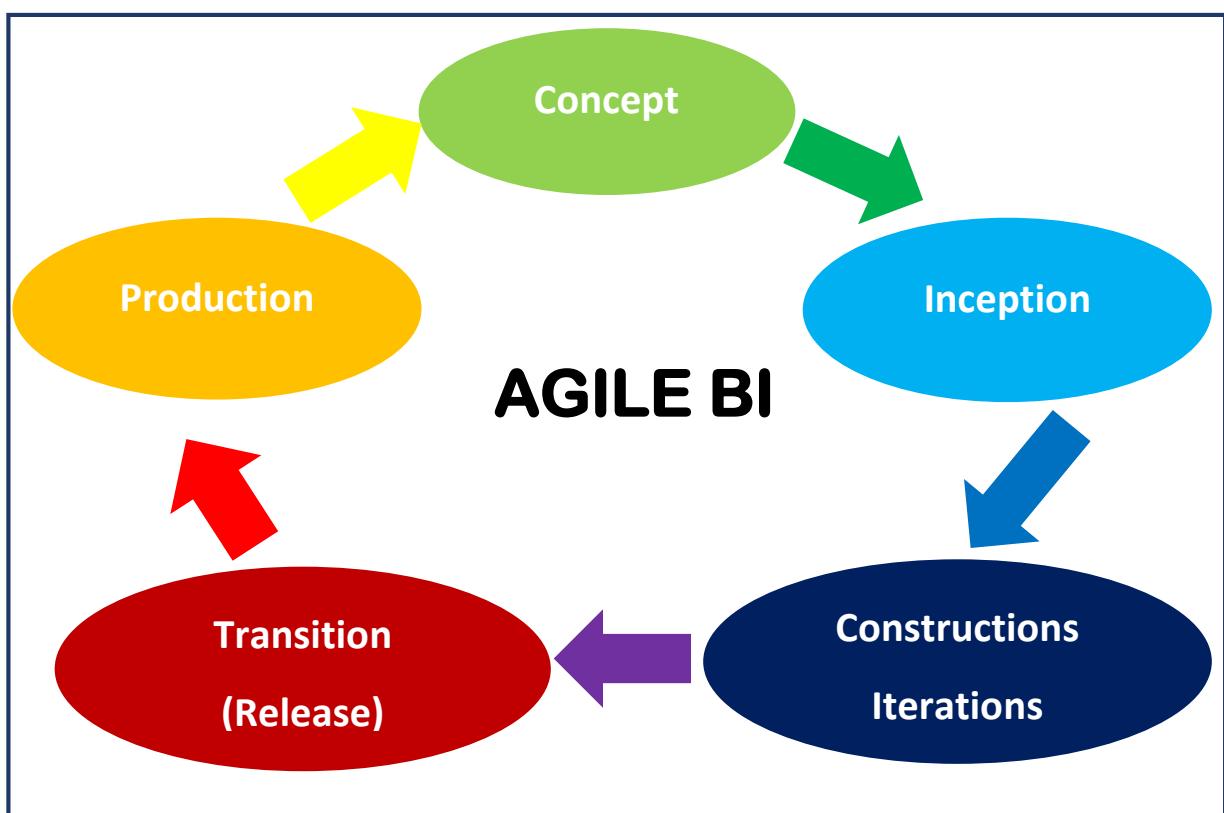


Figure 5: Méthodologie Agile appliquée à la BI

Les métiers s'approprient plus facilement le produit obtenu par cette méthodologie et nous risquons moins le phénomène de rejet et de défiance vis-à-vis de l'IT qui souvent est perçue comme une boîte noire qui coûte cher et qui délivre des solutions qui ne marchent jamais.

L'agilité en BI ne signifie pas non-respect des standards de développement ou d'application des best practices. Elle n'est donc pas en contradiction avec les principes de bonne gestion de projet et de gouvernance SI.

3. Comparaison des méthodes

Le tableau suivant montre quelques critères sur lesquels la comparaison est basée entre les différentes méthodes de gestion choisies

	Traditionnel	Agile
Processus	Exhaustif	Minimal
Dirigé par	Plan	Vision, valeur
Cycle de vie	En cascade, en V	Continu, Itératif
Rôle et collaboration	Directif, Fonctions	Participatif, polyvalent
Contrôle qualité	En fin de cycle	Continu
Planification	Prédictives	Adaptives

Tableau 1: Tableau comparative entre la méthode traditionnel et agile

4. Présentation de la méthode SCRUM

La méthode Scrum se réalise dans un milieu où le projet est considéré complexe. Elle est destinée à la gestion des projets informatiques de manière bien structurée et organisée.

La méthode agile est conçue pour l'interaction entre l'équipe et les gestionnaires en indiquant l'importance de l'intervention du client. [4]

i. Les intervenants dans SCRUM

La méthodologie Scrum définit trois rôles :

- **Product Owner** : Client dans notre cas c'est la direction générale
- **SCRUM Master** : L'encadrant professionnel et pédagogique, ce sont eux qui supervisent et qui garantissent l'avancement du projet.
- **Equipe de développement** : Nous représentons l'équipe de développement. Nous jouerons le rôle d'un architecte, intégrateur de données, réalisateur de tableau de bord.

ii. Les artefacts du Scrum

- **Le Backlog produit :**

Représente une liste des tâches qui vont être effectuées lors du projet. Chaque tâche est reliée aux critères exigés par le client. Chaque élément est classé selon la priorité mise par le client.

- **Le Backlog du Sprint :**

Il s'agit d'un engagement que chaque développeur doit suivre sur un temps établi pour chaque tâche. Chaque Sprint doit assurer l'atteinte des objectifs mis en place au début du sprint.

- **Sprint Burn-Down Chart :**

Il s'agit d'un graphique d'avancement, qui donne une vision globale sur l'évolution de la quantité de travail restante par rapport au travail engagé.

La figure suivante représente le processus de déroulement d'un projet Scrum

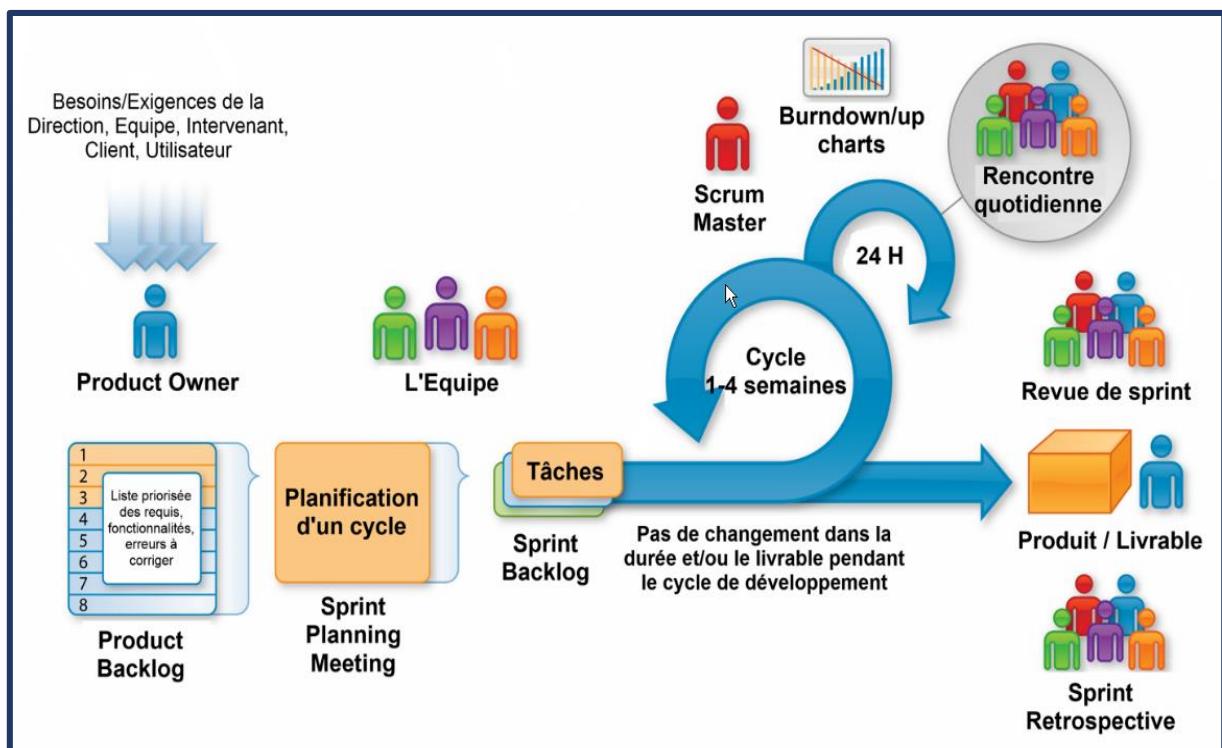


Figure 6: Déroulement projet Scrum

II. Analyse comparative entre l'approche Top-Down et Bottom-UP

Quand on parle de la conception du data Warehouse pour le projet, les deux méthodes les plus connues sont l'approche introduite par Bill Inmon (Top-down) et celle de Ralph Kimball (Bottom-up). Chacune d'entre elles a ses propres caractéristiques. Dans les paragraphes qui vont suivre, nous nous intéresserons de plus près à chaque approche pour faire une étude comparative entre les deux approches.

1. L'approche de Bill Inmon

D'après Inmon l'entrepôt de données fournit un cadre logique pour une prestation de business intelligence et gestion d'affaire.

Selon la théorie et l'approche de Bill Inmon l'entrepôt de données est :

- **Orienté vers le sujet** : les données de l'entrepôt de données sont organisées selon les sujets (table de fait) et les dimensions sont les axes d'analyse de ces derniers.
- **Non volatile** : les données de l'entrepôt de données ne sont jamais écrasées ou supprimées une fois engagées, les données sont statiques. Elles sont conservées pour les analyses futures.
- **Intégré** : l'entrepôt de données contient des données en provenance de la plupart ou la totalité des systèmes opérationnels et ces données sont rendues compatibles.
- **Varient sur le temps (Time-Variant)** : Pour un système d'exploitation, les données mémorisées contiennent la valeur actuelle.
- **Le data warehouse** : Il s'agit d'un entrepôt de données centralisé, qui a pour tâche d'enregistrer et d'organiser des informations dans un entrepôt de données à partir d'un organisme ou d'une société.

Dans ce cas l'entrepôt de données est une architecture de base relationnelle formée d'une approche dite « **Top-Down** » qui provient d'un modèle de données d'entreprise définit d'avance. Le data warehouse représente une vue centralisée et organisée sur toutes les informations, ce qui met en avant la massification d'un data warehouse. [5]

Cette figure illustre le processus du data warehouse selon l'approche de Bill Inmon

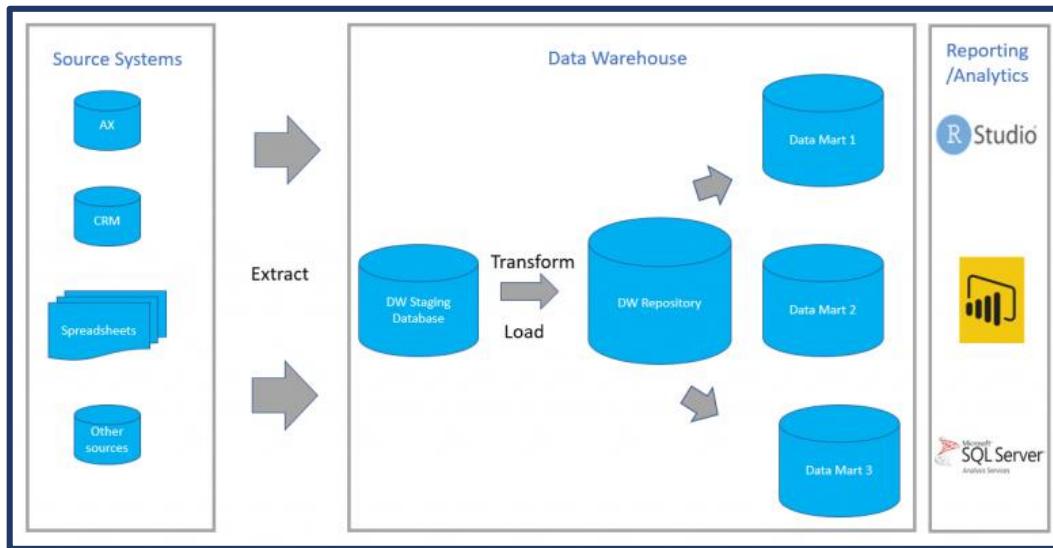


Figure 7 : Approche Bill Inmon

2. L'approche de Ralph Kimball

Une approche du data warehouse peut être prise en compte, c'est celle de Ralph Kimball selon qui l'entrepôt de données est l'union des datamarts⁴ cohérents entre eux grâce aux dimensions conformes. [6]

Cette méthode permet de mettre en place une conception appuyée sur :

- Des phases de fonctionnement clairement établies
- Définir la finalité d'un entrepôt de données
- Différentes phases :
 - Analyse et conception
 - Définition du modèle de données (modèle en étoile/flocon)
 - Acquisition des données
- Déterminer et recenser les données à entreposer :
 - Recherche des données dans les sources de l'entreprise
 - Nettoyage et organisation des données
 - Démarches d'alimentation
 - Définition les aspects techniques de la réalisation
 - Définition les modes de restitution, indicateur
 - Définition des stratégies d'administration, évolution, maintenance

⁴ Datamart : un sous-ensemble d'un data Warehouse destiné à fournir des données aux utilisateurs et souvent spécialisé vers un groupe ou un type d'affaire.

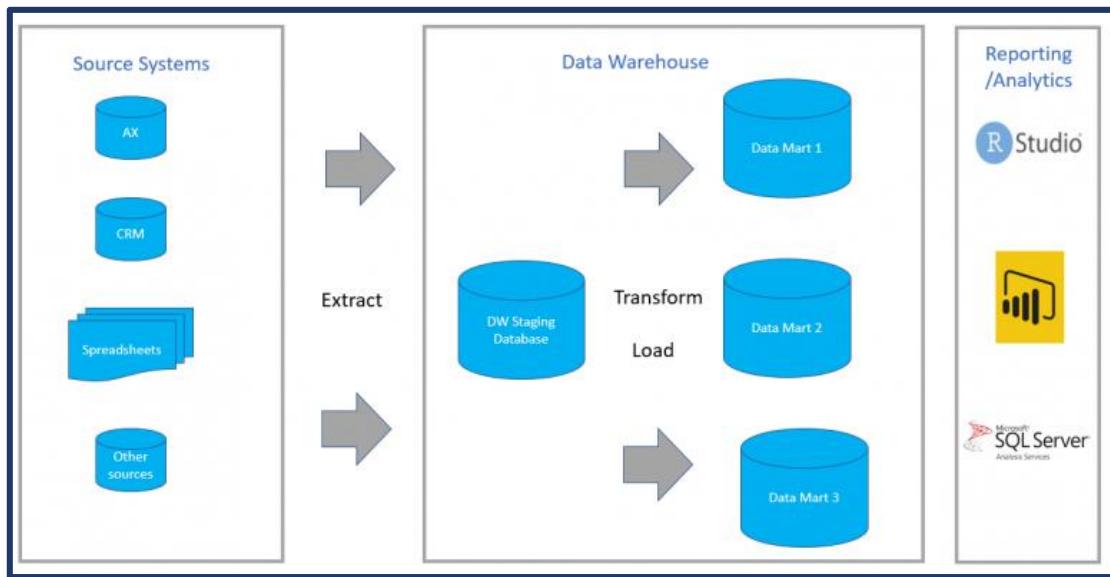


Figure 8 : Approche Ralph Kimball

3. Analyse comparative entre les deux approches

	TOP - DOWN	BOTTOM - UP
Partisan	Bill INMON	Ralph KIMBALL
Conception	Data warehouse	Data Mart
La structure de l'architecture	Un data warehouse regroupe toutes les données de l'entreprise	Un data mart est centré sur un sujet ou un métier particulier
La complexité de la méthode	Assez simple	Simple
Outils	Traditionnel (modèle entité relation)	Modélisation dimensionnelle qui part de la modélisation relationnelle
Autre	-Flexibilité -orienté données -Longue vie	-Restrictive -Orienté projet -Courte vie

Tableau 2 : Tableau comparatif des deux approches

IV. Analyse comparative entre les différents types de modèle

Avant de choisir le type de modèle que nous allons concevoir, il faut faire une étude comparative entre les types existants afin de connaître le modèle le plus approprié pour le projet.

1. Modèle en étoile

Il s'agit d'un modèle multidimensionnel qui est formé d'une table de faits liés par des dimensions comme le montre la figure si dessous. Pour choisir le modèle le plus adéquat, nous allons présenter les avantages et les inconvénients de chacun. [7]

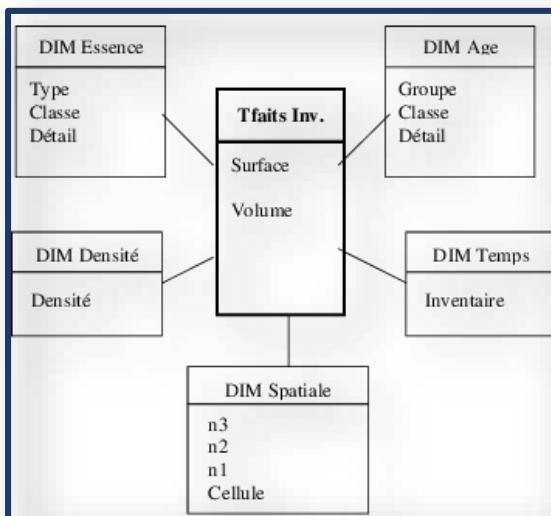


Figure 9 : Schéma en étoile

Avantages :

- Les requêtes sont plus simples car les jointures du schéma en étoile sont plus simples que celle du schéma normalisé
- La logique de reporting est simplifiée dans un schéma en étoile, telle que la génération de rapports sur une période ou un intervalle de temps
- La performance des applications est améliorée et amène à une lecture simple du rapport

Inconvénients :

- Il y a une redondance dans les dimensions
- La mise à jour peut entraîner des anomalies sur les données

2. Modèle en flocon de neige

La seule différence par rapport au modèle en étoile c'est le changement des dimensions. Dans le schéma en étoile, elles sont normalisées. [8]

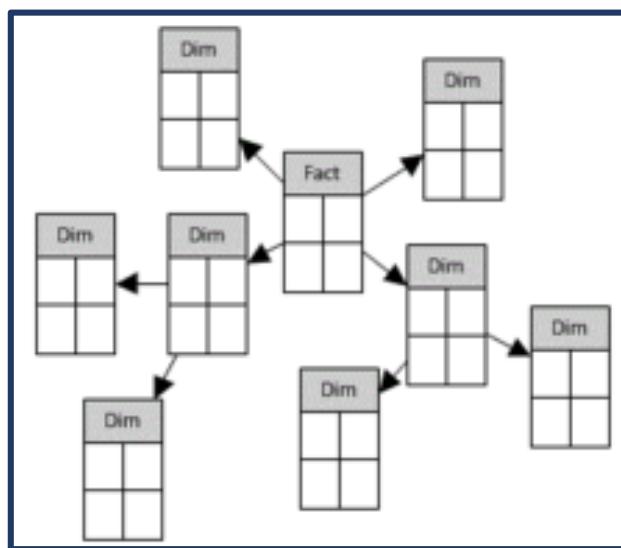


Figure 10 : Schéma en flocon de neige

Avantages :

- Il y a une économie de stockage grâce à la normalisation des attributs

Inconvénients :

- Difficulté de navigation à cause des nombreuses jointures
- Puisque l'objectif est de stocker plus, ceci implique une augmentation du coût au niveau temporel. La performance lors de la navigation sera également mauvaise

IV. Etude comparative entre les différents outils BI

1. QlikView

Qlikview est généralement populaire en tant qu'outil décisionnel utilisant une plate-forme de découverte d'entreprise essentielle. Il constitue l'un des outils de BI les plus attractifs pour les utilisateurs professionnels car il ne fait que satisfaire la gestion en libre-service des utilisateurs professionnels ordinaires de toute organisation. Comme tout autre outil de BI, QlikView peut analyser des données volumineuses et fournir des décisions variétales afin de générer un rapport analytique pour tout type d'organisation en analysant les données de leur dictionnaire de données.



Figure 11 : Logo Qlikview

2. Tableau

C'est l'outil de Business Intelligence le plus utilisé pour créer, afficher et partager des rapports interactifs avec les utilisateurs, à tout moment. Ils sont indépendants de la plate-forme et nécessitent un nombre inférieur d'experts techniques. Ils travaillent sur une architecture client-serveur pour des déploiements rapides. Ils conviennent parfaitement aux cadres supérieurs qui ont besoin d'une bonne BI. Tableau présente des avantages en termes d'accès rapide, de visualisation interactive, rentable et de bonne sécurité.



Figure 12 : logo Tableau

Caractéristiques de Tableau :

- Pour le partage de données, ils utilisent Tableau Public.
- Ils implémentent des visualisations de données interactives sur le Web.
- Leur performance est forte et fiable et fonctionne sur des données énormes.
- Ils sont compatibles avec les appareils mobiles et prennent en charge les versions complètes en ligne.

3. Power BI

Power BI est un outil de Business Intelligence qui permet de télécharger et de publier des données dans l'ensemble de nos sociétés. La Business Intelligence répond à toute requête et améliore la prise de décision. Ajout de puissance à l'entreprise pour une bonne visualisation des données. Une autre fonctionnalité de Power BI est Quick Insights dans laquelle nous pouvons rechercher dans un jeu de données des modèles intéressants et fournit une liste de graphiques pour une meilleure compréhension des données. Il utilise l'intelligence artificielle et l'exploration de données pour analyser les données.

Les fonctionnalités de Power BI sont :

- Visualisations graphiques riches à partir de données BI complexes.
- Rapport ad hoc
- Bon volet de navigation
- Comprend un ensemble de données avec des tableaux de bord personnalisables.

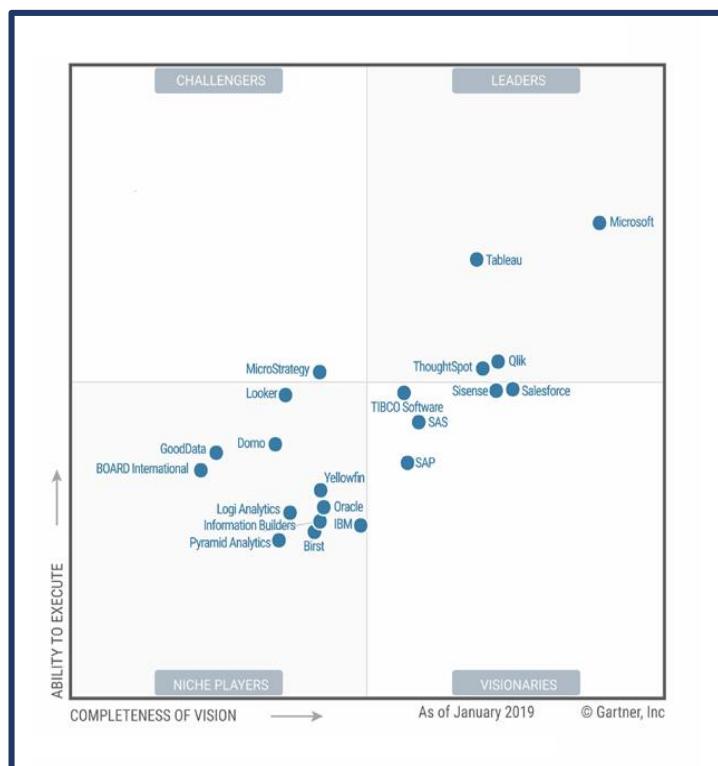


Figure 13 : classification des outils BI (Source : Gartner fév 2019)

V. Récapitulatif de choix

Toutes ces comparaisons mènent à un choix de méthodologie appropriée à notre projet. Dans les paragraphes qui suivent, nous allons expliquer nos choix.

1. Choix de méthodologie

En se basant sur la comparaison entre la méthode agile et la méthode classique, nous avons décidé de construire notre solution décisionnelle sous la supervision de la méthode agile ou plus précisément à l'aide de Scrum BI, qui va nous donner une vision sur tous les aspects du projet et ce grâce au Blacklog qui aide à organiser le temps et le mesurer par rapport à l'échéance.

2. Choix de la démarche

En se basant sur la comparaison entre l'approche top down de Bill Inmon et l'approche Bottom- up de Ralph Kimball, nous avons choisi celle de Ralph Kimball étant donné que nous disposons de plusieurs dimensions qui sont en relation l'une avec l'autre. Notre choix s'explique aussi par la facilité de l'analyse grâce à cette méthode.

i. Choix du type du type de modèle

Les deux modèles présentés sont intéressants, mais nous allons opter pour le modèle en flocon de neige qui va faciliter l'analyse grâce à la hiérarchisation des dimensions.

ii. Choix de l'outil

Concernant les outils, nous n'avons pas fait d'étude comparative, étant donné que la société nous a proposé de travailler avec les outils Microsoft commençant par Microsoft SQL Server Management Studio pour la conception du data warehouse, avec Visual Studio pour le processus ETL et enfin Power BI pour la conception du tableau de bord.

- **Microsoft Visual studio 2017**

Le SSDT est un outil moderne pour la création des lots SSIS (SQL Server Intégration services), la conception des SSAS (SQL Server Analyse Services). [4]



Figure 14 : logo Visual Studio

- **Microsoft Dynamics Navision 2017**

C'est le progiciel ERP utilisé par l'entreprise. Les tables sont tous défini dedans et qui possède un id. Son usage essentiellement pour savoir les noms des tables. [5]



Figure 15 : logo Microsoft Dynamics Navision

- **SQL Server Management Studio 2017**

Est un environnement intégré permettant de gérer toutes les infrastructures SQL. SSMS, fournir des outils pour configurer, surveiller et administrer des instances SQL Server et des bases de données. SQL Server Management est utilisé surtout pour mettre à niveau les composants de la couche de données utilisées par les applications. [4]



Figure 16 : Logo SSMS

- **Power BI**

Power bi est une application qui permet de créer des tableaux de bord personnalisés avec simplicité. Il peut créer tout un environnement décisionnel, mais nous l'avons utilisé que pour l'élaboration du Dashboard. [4]



Figure 17 : Logo Power BI

Le schéma ci-dessous représente le processus de la réalisation du projet BI avec les outils choisis :

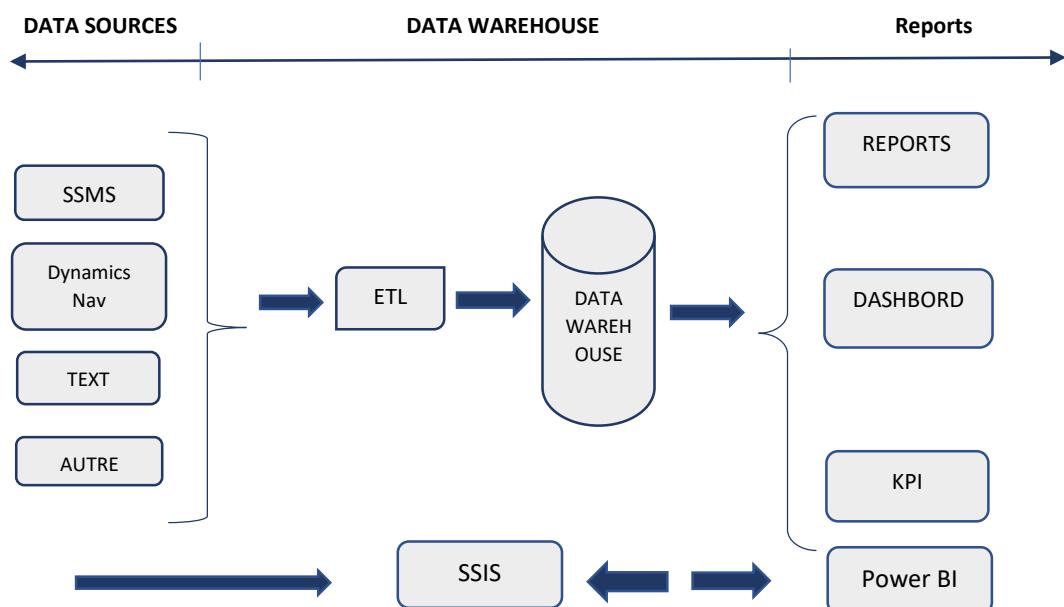


Figure 18 : récapitulatif du choix d'outils dans le processus décisionnel

Conclusion

Grâce à l'étude comparative réalisée et au choix de la méthode de gestion de notre projet ainsi que celui des outils utilisés. Nos choix techniques et organisationnels fixés, nous allons dans ce qui suit, planifier notre projet et démarrer par la suite la réalisation et la mise en place de notre solution.

Chapitre 3 :
Définition des besoins et planification
du projet

Introduction

Cette phase de préparation est aussi connue sous le nom de Sprint 0, elle est consacrée au lancement du projet.

Dans ce chapitre nous allons spécifier les besoins et étudier les besoins fonctionnels ainsi que les besoins non fonctionnels du système. Ensuite, nous allons élaborer le Backlog produit et mettre au point le découpage des sprints.

I. Spécification de besoin

L'objectif ultime de ce travail est de fournir des tableaux de bord à 4 départements ;

- Tableau de bord dédié au contrôle de gestion
- Tableau de bord dédié au directeur général administratif
- Tableau de bord dédié au président directeur général
- Tableau de bord dédié au direction activité Microsoft

Chaque tableau de bord possède ces indicateurs de performance clé spécifiques que je vais les citer en détail plus tard.

1. Les besoins d'affaires

L'objectif de l'entreprise est de concevoir un système décisionnel qui va subvenir à son besoin, qui consiste à :

- **Département contrôle de gestion**

Contrôle de gestion rencontre des difficultés au calculs de chiffre d'affaire par jour/activité.

Pour effectuer ses calculs, il possède un fichier Excel et il saisit les données manuellement ce qui prend beaucoup de temps en même temps ces calculs ne sont pas fiable. Plus la marge d'erreur est importante, moins on peut avoir confiance que les résultats du sondage sont proches des vrais résultats, et ainsi, de la réalité.

- **Direction opérationnel « Dev Microsoft »**

Bien que l'entreprise est un intégrateur des systèmes ERP, elle possède plein de projet et le directeur activité Microsoft veut avoir un suivi sur l'avancement des projets et le taux

d'occupation de ces consultants au temps réel pour savoir comment gérer leur affectation, et surtout pour respecter les délais des clients.

- **Département direction administratif**

Le directeur général administratif a besoin des mêmes calculs de département contrôle de gestion avec d'autres filtres par exemple le chiffre d'affaire de produit X pour la région Y ou d'un autre pays.

- **Département président directeur administratif**

Le PDG est le dirigeant de plus haut rang dans l'entreprise qui possède un niveau d'accès supérieur et un privilège ultime de coup il va avoir tous les tableaux de bords créés avec quelque KPIs supplémentaires.

2. Les acteurs de l'entreprise

Pendant tout le processus d'étude des besoins, nous nous sommes entretenus avec un grand nombre d'employés à **Discovery** afin d'avoir une solution satisfaisante.

C'est à partir de ces rencontres, que nous allons identifier les différents intervenants et utilisateurs du système décisionnel que nous allons élaborer.

Dans notre cas il y a deux types de cible ou utilisateur :

- **Utilisateur informaticien :**

Dès que la solution est mise en place, les techniciens informatiques dont l'objectif est de s'assurer de sa maintenance, disposeront d'une data Warehouse qui va leur permettre de générer des rapports et déployer les tableaux de bords.

- **Utilisateur administratif :**

Discovery Informatique n'emploie pas que des informaticiens, elle emploie aussi des comptables, des financiers des responsables ressources humaines, ce qui fait la solution élaborée doit être comprise par eux aussi ; il faut donc leurs procurer une solution facile, rapide et qui leur garantit un accès sécurisé aux rapports générés, offrant à chaque personne une vue spécifique selon son poste, son besoin et sa satisfaction.

3. Besoins fonctionnels

- Pour Chiffre d'affaire
 - ✓ Les visualisations demandées sont : CA⁵ par activité, par sous activité, par produit, par pays ou région, par marché (local, externe), par Client, export Afrique par région.
 - ✓ Calcule de Chiffre d'affaire :
[Facturé] + [PAF].[Année] – [PAF].[Année - 1]
- Pour les charges :
 - ✓ Charges directes : Achat de licence, salaire, autre achat, formation, publicité et communication, ainsi déduire la marge directe de chaque pôle, activité et produit.
 - ✓ Charge indirectes : clé de répartition (nombre par équipe, amortissement . . .) et déduire le résultat net.
- Recouvrement
 - ✓ Suivi des mesures et taux de réalisation
 - ✓ Suivi le recouvrement par consultant

Visualiser et filtrer le montant et le nombre de demande de financement selon :

- Les régions
- Les secteurs d'activité
- L'activité
- Le temps (année, mois)
- Type de projet
- Zones

⁵ CA : Chiffre d'affaires

Afin de satisfaire les attentes des décideurs, nous proposons de concevoir un système d'aide à la décision qui répond à leurs besoins, comme le montre le tableau suivant :

Département : Contrôle de gestion		
KPI	Objectif	Description
CA /Produit	Déduire le produit l plus rentable	Chiffre d'affaire pour chaque produit
CA / Activité	Déduire l'activité la plus rentable	Chiffre d'affaire pour chaque activité
CA/Région	Calculer les CA par région	Chiffre d'affaire pour chaque région
CA Local / Externe	Déduire les relations entre national et international	Chiffre d'affaire Local Chiffre d'affaire Externe
CA Export Afrique	L'apport du marché africain	Chiffre d'affaire l'export africain
Frais par activité	Déduire les activités les plus couteuses	Frais total par activité
Frais Local / Export	Frais totaux nationaux et internationaux	Frais total local et export
Comparaison Frais & CA	Visualisation de Compte de résultat	Frais et chiffre d'affaire totaux
Charge indirecte	Calculer Résultat net	Charge indirecte totale
Charge directe	Déduire la marge directe de chaque pôle, activité produit	Charge directe totale
Taux des consultations	Répartition en pourcentage la phase périodique des consultant	taux d'occupation des consultants
Recouvrement	L'affectation de chiffre d'affaire	Recouvrement par les projets

Tableau 3 : Tableau besoin fonctionnel département contrôle de gestion

Afin de satisfaire les attentes des décideurs, nous proposons de concevoir un système d'aide à la décision qui répond à leurs besoins, comme le montre le tableau suivant :

Département : Direction Générale		
KPI	Objectif	Description
CA/Activité	Déduire l'activité la plus rentable	Chiffre d'affaire pour chaque activité
CA/Sous Activité	Déduire la sous activité la plus rentable	Chiffre d'affaire pour chaque sous activité
CA/Pays	Calculer les CA par pays	Chiffre d'affaire pour chaque pays
CA Local / Externe	Déduire les relations entre national et international	Chiffre d'affaire Local Chiffre d'affaire Externe
CA Export Afrique	L'apport du marché africain	Chiffre d'affaire l'export africain
Frais par activité	Déduire les activités les plus couteuses	Frais filtré par activité
Frais Local / Export	Frais totaux nationaux et internationaux	Frais local et export totaux
Charge salariale Par activité	Charge salariale par activité	Charge salariale par rapport charge totale
Comparaison Frais & CA	Visualisation de Compte de résultat	Frais par rapport chiffre d'affaire
Taux d'occupation	Répartition en pourcentage la phase périodique des consultant	Taux d'occupation par consultant, par activité

Tableau 4 : Tableau besoin fonctionnel Direction Général

4. Besoins non fonctionnels

En plus des principaux points des besoins fonctionnels, notre solution doit répondre aux critères suivants :

- Clarté : des tableaux clairs lisibles et faciles à interpréter
- L'efficacité : l'analyse du tableau de bord doit donner un résultat à un pourcentage près des attentes des décideurs
- L'efficience : est un composant important de la mesure de la performance. C'est l'optimisation de la consommation des ressources utilisées la production d'un résultat.
- Scalabilité : la solution doit être extensible pour pouvoir suivre le rythme de l'entreprise en termes de données et histoire de donnée.

II. Organisation du projet à l'aide de SCRUM

Pour entamer le processus de la méthode agile, il faut tout d'abord élaborer le backlog produit. Ensuite, il faut identifier les points les plus importants du projet : les risques, les défaillances du projet.

1. Le Backlog Produit

ID	Thème	User Story	Priorité	Sprint
1	Vision globale sur le chiffre d'affaire	En tant que décideur, je veux avoir une vue globale sur les statistiques générales des Ventes et des Achats par projet, par secteur d'activité et par région.	Élevé	S2
2		En tant que décideur, je veux avoir le pouvoir de filtrer les Ventes et les achats selon le secteur d'activité, région, projet et activité	Élevé	S2
4		En tant que décideur je veux avoir le pouvoir de filtrer le chiffre d'affaire par période et fixer l'objectif pour chaque mois.	Élevé	S2
5		En tant que décideur je veux visualiser le chiffre d'affaire par pays(export), pour le marché tunisien, par date, par client, par projet, par activité	Élevé	S2
6		En tant que décideur je veux suivre le nombre de demande de financement.	Élevé	S2

ID	Thème	User Story	Priorité	Sprint
7	Suivi le taux d'occupation des consultant	En tant que décideur je veux savoir le taux d'occupation des consultant par projet, par secteur d'activité.	Elevé	S2
8		En tant que décideur je veux savoir le nombre de homme jours pour chaque projet, secteur d'activité et par consultant.	Elevé	S2
9	Suivi des charges et frais.	En tant que décideur je veux avoir le pouvoir de visualiser les charges et les frais par projet, par secteur d'activité, par activité, par région.	Elevé	S2
10	Suivi des durées des projets	En tant que décideur je veux visualiser les statistiques sur la durée de réalisation des projets par consultant, par client, par activité et sous activité.	Élevé	S2
11		En tant que décideur je veux filtrer les durées de projet sur le temps	Élevé	S2

Tableau 5 : Backlog produit

2. Planification des Sprints

i. Pilotage du projet avec SCRUM

Notre équipe et formé de 3 membres, représentée dans la figure ci-dessous :

Le Product Master représenté par Mr Aymen Bel Haj Kacem

Le Product Owner est représenté par Mr Med Amine Djobbi directeur d'activité Microsoft et Mr Amir Trabelsi ingénieur BI.

Le team membre présenté par Oussama Zribi.



Product Master

Med Aymen Haj Kacem



Product Owner

Med Amine Djobbi

Amir Trabelsi



Team members

Oussama Zribi

ii. La partition de la solution en Sprints

Afin d'entamer notre projet en bonne et due forme, nous avons fait la répartition des sprints sur la durée de notre stage, comme le montre la figure suivante :

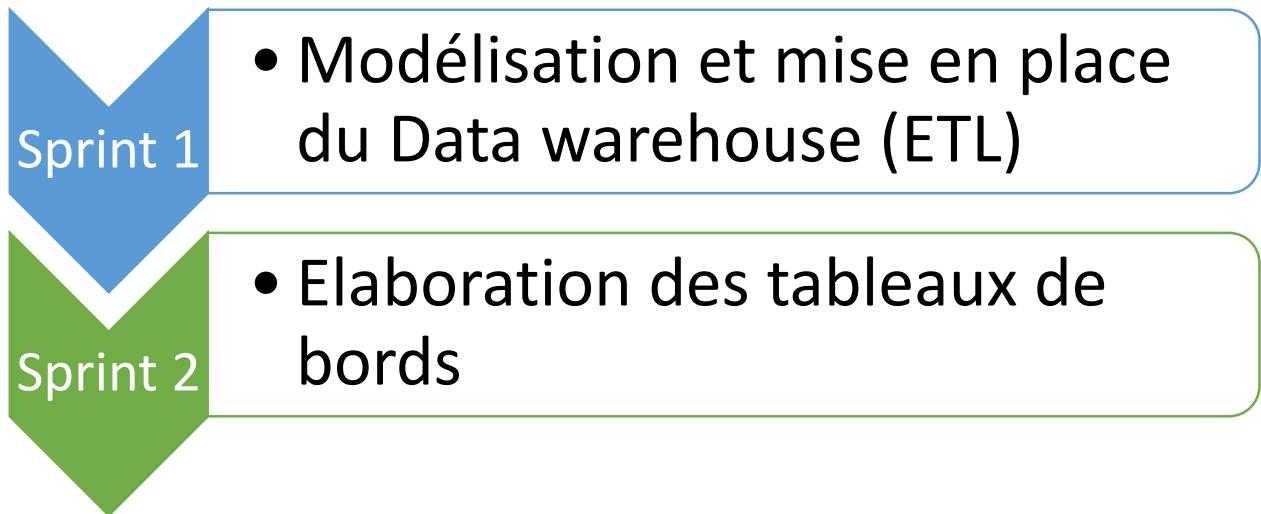


Figure 19 : Les Sprints de projet

iii. Diagramme de Gantt

Afin de mener à bien notre travail, nous avons élaboré un chronogramme, représenté dans la figure ci-dessous, qui englobe les différentes tâches à réaliser dans le projet :

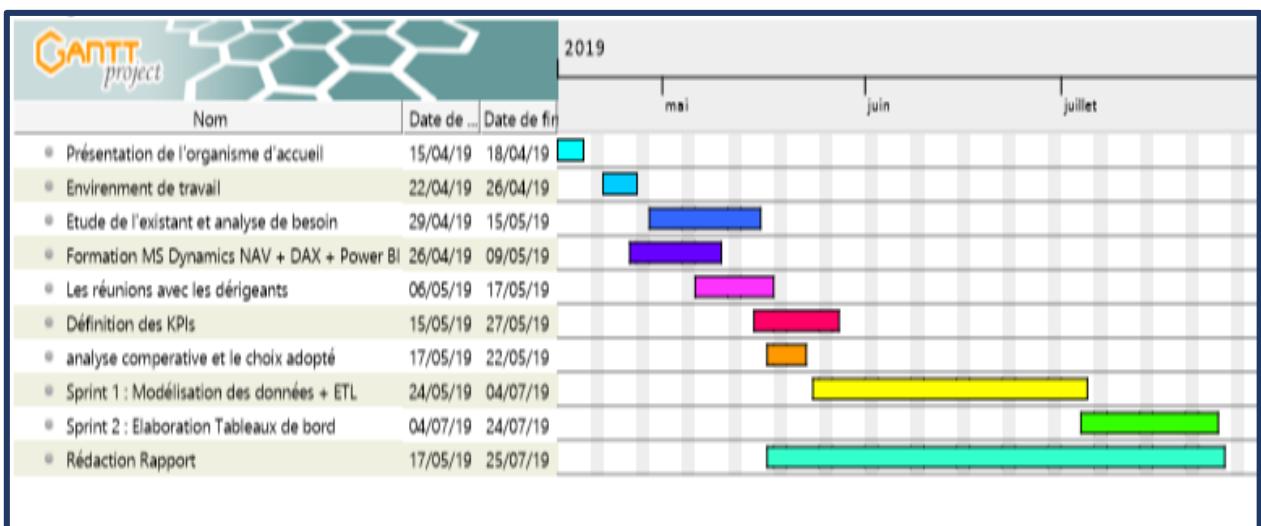


Figure 20 :Diagramme de Gantt

iv. Les risques de projet

Chaque projet peut présenter des risques. Scrum est la méthode la plus adéquate qui permet d'éviter tout problème. En effet, avec le Backlog produit, tous les obstacles sont regroupés et mis comme sujet de discussion afin de trouver une résolution dans la prochaine cérémonie Scrum avec le Scrum Master.

v. Les contraintes de projet

Dans chaque projet, nous pouvons rencontrer des défis à relever et pour pouvoir aboutir à un travail satisfaisant, nous devons spécifier dès le commencement, les contraintes possibles.

Dans notre cas, l'entreprise nous a fourni une base purgée au début, la structure des tables sans pour autant nous fournir les détails étant donné la confidentialité des données. La modélisation de donnée est faite par la base Démo de Navision (Cronus).

vi. Etude des données sources

Après la détermination des besoins fonctionnels et non fonctionnels du projet, il faut passer à la collecte des données. Dans cette partie, nous allons étudier les données sources collectées du système d'information de l'entreprise, avec lesquels nous allons alimenter notre entrepôt de données.

Dans ce cas nous avons eu une réunion avec Mr Mourad Smida, consultant ERP, pour aider à comprendre les données de la base et nous a fourni les numéros des tables nécessaires qui sont définis dans le MS Dynamics Navision. Nous avons eu recours à des données sous un fichier d'extension .bak et pour identifier les tables, il faut appliquer cette petite procédure, on prend l'exemple de la table client numéro 18, d'après son nom sur la base de donné, nous connectons à l'instance SSMS où nous avons parcouru notre fichier et cherchons la table numéro 18 intitulé '\$Customer'

Type	ID	Name	Modified	Version List	Date	Time	Compi
Table	6	Customer Price Group	NAWV10.00	25/10/16 12:00:00	✓	✓	✓
	7	Standard Text	NAWV10.00.00.15601	23/02/17 12:00:00	✓	✓	✓
	8	Language	NAWV19.00	15/09/15 12:00:00	✓	✓	✓
	9	Country/Region	NAWV19.00.NAVFR9.00	15/09/15 12:00:00	✓	✓	✓
	10	Shipment Method	NAWV11.00.NAVFR10.00	25/10/16 12:00:00	✓	✓	✓
	11	Salesperson/Purchaser	NAWV11.00.00	25/10/16 12:00:00	✓	✓	✓
	12	Location	NAWV11.00.00.17501	30/06/17 12:00:00	✓	✓	✓
	13	G/L Account	NAWV11.00.00.00.17972.NAVF...	04/08/17 12:00:00	✓	✓	✓
	14	G/L Entry	NAWV11.00.00.00.17972.NAVF...	04/08/17 12:00:00	✓	✓	✓
	15	Cust. Ledger Entry	NAWV11.00.00.00.18609.NAVF...	26/09/17 12:00:00	✓	✓	✓
	16		NAWV11.00.00.00.18609.NAVF...	07/09/12 12:00:00	✓	✓	✓
	17		NAWV11.00.00.00.18609.NAVF...	25/10/16 12:00:00	✓	✓	✓
	18		NAWV11.00.00.00.18609.NAVF...	27/10/17 12:00:00	✓	✓	✓
	19	Cust. Invoice Disc.	NAWV11.00.00.00.18976.NAVF...	07/09/12 12:00:00	✓	✓	✓
	20		NAWV11.00.00.00.18976.NAVF...	25/10/16 12:00:00	✓	✓	✓
	21	Cust. Ledger Entry	NAWV11.00.00.NAVFR10.00	27/10/17 12:00:00	✓	✓	✓
	22		NAWV11.00.00.NAVFR10.00	07/09/12 12:00:00	✓	✓	✓
	23	Vendor	NAWV11.00.00.00.18976.NAVF...	25/10/16 12:00:00	✓	✓	✓
	24	Vendor Invoice Disc.	NAWV11.00.00.00.18976.NAVF...	24/08/17 12:00:00	✓	✓	✓
	25	Vendor Ledger Entry	NAWV11.00.00.NAVFR10.00	24/08/17 12:00:00	✓	✓	✓
	26		NAWV11.00.00.NAVFR10.00	24/08/17 12:00:00	✓	✓	✓
	27	Item	NAWV11.00.00.00.18197	24/08/17 12:00:00	✓	✓	✓

Figure 21 : Identifier la table par son numéro

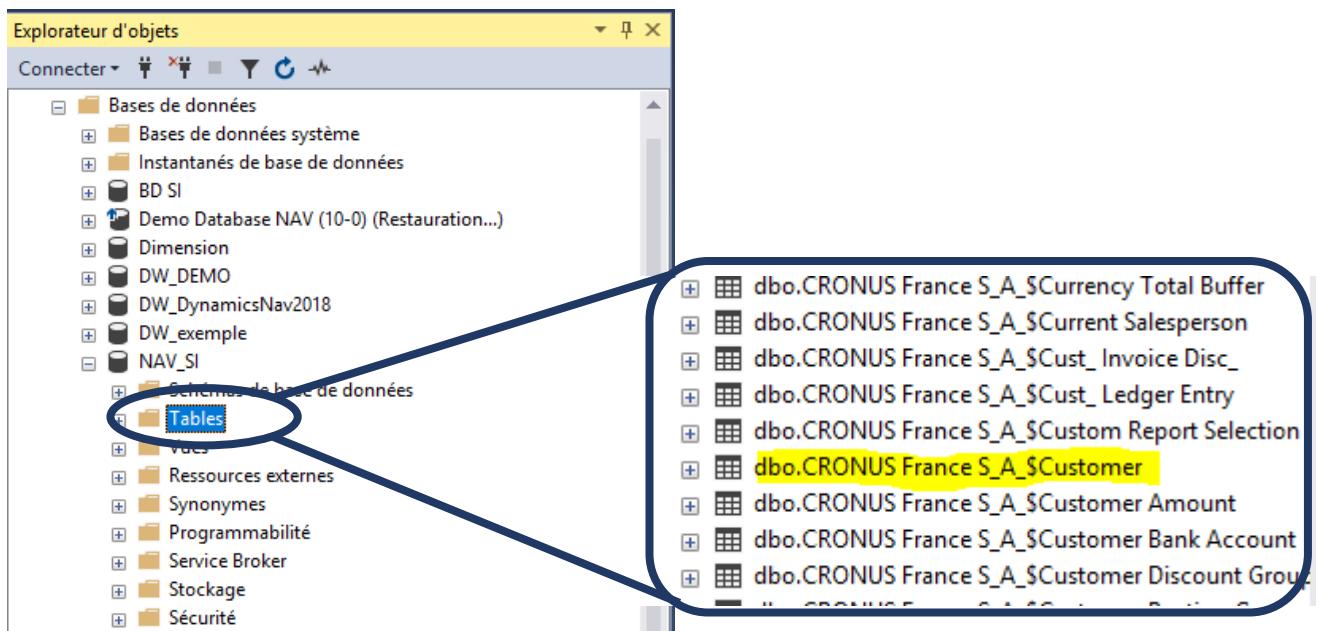


Figure 22 : procédure de SSMS

Les différentes tables nécessaires sont présentées dans le tableau suivant :

ID	Name	Description
10865	Payment Header	encaissement client
10866	Payment Line	Décaissement fournisseur
270	Bank Account	Banque et caisse (écriture)
271	Bank Account Ledger Entry	Banque et caisse
18	Customer	Client
23	Vendor	Fournisseur
17	G/L Entry	comptabilité générale
167	Job	Projet
5800	Item Charge	Frais annexe
27	Item	Produit
32	Item Ledger Entry	Charge produit
124	Purch. Cr. Memo Hdr.	Entête avoir achat
125	Purch. Cr. Memo Line	Ligne avoir achat
120	Purch. Rcpt. Header	Entête réception achat
121	Purch. Rcpt. Line	Ligne réception achat
25	Vendor Ledger Entry	Entête écriture fournisseur
380	Detailed Vendor Ledg. En	Ligne écriture fournisseur
32	Item Ledger Entry	Quantité de stock
5802	Value Entry	Valeur de stock

Tableau 5 : Les tables de projet

Conclusion

Dans ce chapitre, nous nous sommes intéressés aux acteurs du système, aux besoins fonctionnels ou non fonctionnels, aux priorités, aux contraintes et risques du projet.

Ensuite, nous nous sommes intéressés à l'élaboration du Backlog Produit qui constitue la partie la plus importante de ce chapitre ainsi que la planification du projet en découplant la solution en sprint.

Pour finir, nous avons étudié les données sources livrées par l'entreprise. Dans le prochain chapitre, nous allons commencer le premier sprint.

Chapitre 4

Sprint 1 : Modélisation et mise en place du
Data Warehouse (ETL)

Introduction

Ce chapitre sera consacré au traitement du premier sprint qui va présenter tout le processus d'un projet BI. Nous présentons dans un premier temps la définition d'un ETL. Puis nous présentons la conception multidimensionnelle de l'entrepôt de données et nous allons assister à la mise en place de l'entrepôts de données en commençant par la collection de données, jusqu'à arriver à la restitution des données.

I. Qu'est-ce qu'un ETL ?

ETL est un acronyme qui désigne les termes « Extract, Transform, Load ». Il s'agit d'un type de logiciel permettant de collecter des données en provenance de sources multiples pour ensuite les convertir dans un format adapté à une Data Warehouse et les y transférer. [15]

Un logiciel ETL (Extract Transform Load) permet d'extraire des données brutes depuis une base de données, pour ensuite les restructurer, et enfin les charger dans une Data Warehouse.

II. Extract, Transform, Load (ETL)

Comme expliqué précédemment, ETL signifie Extraction, Transformation, Chargement. Ce sont les trois étapes que doit impérativement implémenter un outil ETL.

Effectuer dans l'ordre, elle forment un traitement, une tâche ou un scénario (selon les diverses appellations des logiciels).

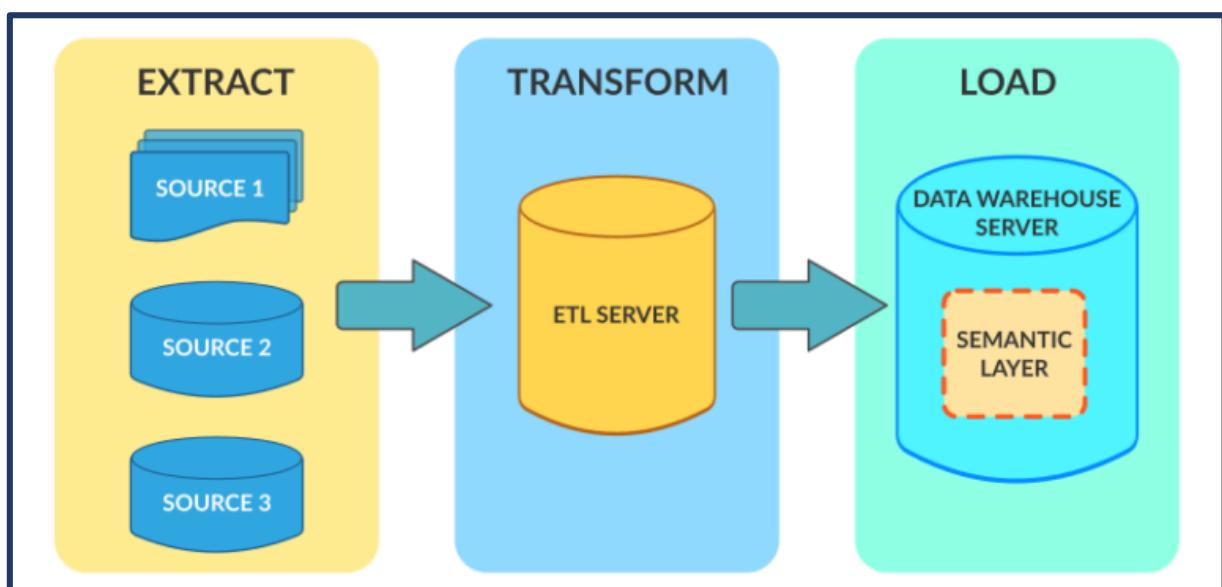


Figure 23 : Extract, Transform, Load [15]

III. Le Sprint Backlog

Le Sprint backlog est utilisé pour faciliter le travail et avoir une vue d'ensemble sur les tâches à faire. Dans notre cas, le sprint 1 contient les objectifs de la construction du data warehouse et le processus de ETL qui va permettre par la suite, l'élaboration du tableau de bord et des rapports.

Sprint 1 <u>Modélisation et mise en place du Data Warehouse (ETL)</u>			
User story	ID Tache	Tache	Durée
En tant que développeur je souhaite mettre en œuvre un data Warehouse.	1	Détermination des mesures d'après les données livrées	4
	2	Détermination des dimensions d'après les données livrées	2
	3	Détermination de l'architecture technique	3
	3	Modélisation du choix	1
	4	Chargement des données dans l'entrepôt de donnée selon le modèle choisi.	2
	5	Créer Staging Area	2
	6	Création de Operational Data Store Et manipulation les données dans l'ODS	10
	7	Création Data Warehouse	5

Tableau 6 : Backlog Sprint 1

L'ETL forme presque la totalité du travail et dispose trois phases :

- **La phase d'Extraction** consiste à collecter les données en provenance d'une ou plusieurs sources
- **La phase de transformation** consiste à reformater et à transformer les données
- **La phase de chargement (loading)** consiste à transférer les données transformées vers la Data Warehouse, le Data Store ou la base de données cible

1. Extract

La première étape concerne l'extraction des données qui sont la plupart du temps hétérogènes. Cela signifie qu'elles peuvent provenir de SGBD (MySQL, Oracle, SQL Server, etc.), de fichiers plats (Txt, Excel, XML, etc.), d'ERP (Entreprise Ressource Planning), de bases hiérarchiques (les ancêtres des SGBD) ou d'autres applications spécifiques. Nous pouvons déjà remarquer le premier obstacle aux ETL : la multitude formats sources possibles à gérer.

Cette étape doit permettre de se connecter aux bases, soit de façon native, soit via JDBC/ODBC, OLE DB, ou encore avec des connecteurs spéciaux.

Il est aussi important, lorsque nous extraireons des données, de pouvoir les analyser. Il faut donc connaître les clés primaires et étrangères permettant respectivement d'identifier une table de façon unique, et de garantir l'intégrité des données. Il faut aussi reconnaître les propriétés de celles-ci : savoir si par exemple cette donnée est de type entier ou chaîne de caractères et quelle est sa taille maximale. Cela peut paraître simple lorsque la source SGBD mais s'avère plus complexe lorsqu'elle provient d'un fichier plat.

Enfin, cette première étape doit être effectuée le plus rapidement possible en exploitant au minimum les ressources du système. Pour gagner du temps, l'objectif de cette étape est aussi de filtrer au maximum les données.

L'étape d'extraction est donc très importante. Elle doit être performante et complète pour pouvoir disposer d'un bon ETL.

2. Transform

Cette seconde étape a pour objectifs la transformation des données. Elle est bien évidemment indispensable si l'on veut obtenir des cibles différentes des sources.

C'est cette étape qui va permettre de joindre les différentes sources selon les clés précédemment spécifiées. Elle va aussi permettre de filtre les données. Le filtrage est bien différent de l'extraction puisque le nous filtre selon des critères à définir.

Une partie importante de l'étape de transformation est de pouvoir effectuer des calculs. Ils peuvent être simple comme une addition ou multiplication, mais peuvent être aussi plus complexes. Disposer d'un outils ETL (SSIS) proposant de nombreuses opérations par défaut est donc un plus. La transformation doit aussi s'occuper des différentes agrégations : effectuer les commandes SQL classiques tels que SUM (somme), COUNT (comptage), ou AVG (moyenne).

En bref, cette étape doit permettre d'effectuer toutes les transformations que le nous souhaite applique aux données sources. Il ne faut pas non plus oublier la sélection ou le découpage des colonnes, la traduction des valeurs (les différents formats de dates possibles ou encore le booléen 1 qui peut signifier M pour 'Masculin'), la fusion, les jointures, les recherches (lookups), la gestion des erreurs et encore de nombreuses autres fonctionnalités.

3. Load

La dernière étape, s'occupe de charger les données, préalablement extraites puis transformée, dans des cibles hétérogènes (le plus souvent des entrepôts de données qui pourront être structures selon un modèle bien précis).

Le chargement va permettre d'insérer ou de mettre à jour les données cibles, et, comme dans les deux étapes précédentes, il doit aussi gérer les erreurs (une chaîne de caractère ne doit pas être inséré dans un champ fait pour les entiers).

Le chargement n'est pas à négliger pour un bon outil ETL, il doit, là aussi, être complet et performant.

IV. Package SSIS et configuration job ETL

L'alimentation de la Datawarehouse se fait automatiquement à travers les Jobs ETL.

Afin de commencer le travail, nous avons créé notre projet en utilisant l'environnement Visual studio 2017. Ci-dessous une visualisation de l'outil utilisé dans laquelle nous avons sélectionné « Intégration Service ».

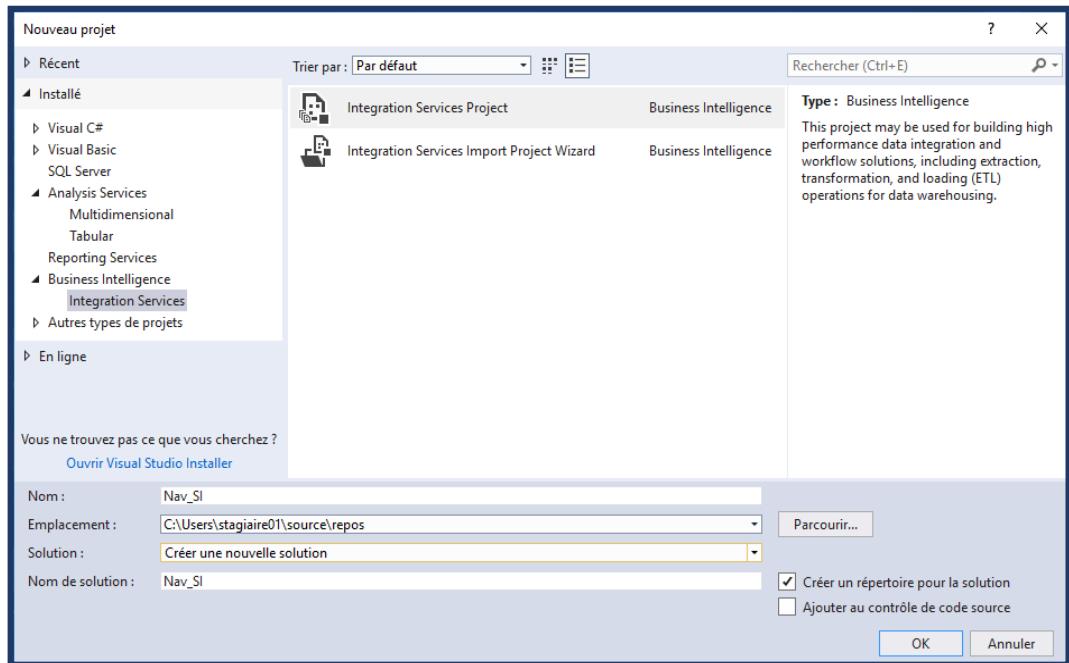


Figure 24: Nouveau projet SSIS

Notre projet se décompose en 4 grandes parties respectivement : Staging Area « STG »,

Operational Data Store « ODS », Data Warehouse « DW » et Datamarts « DMT ».

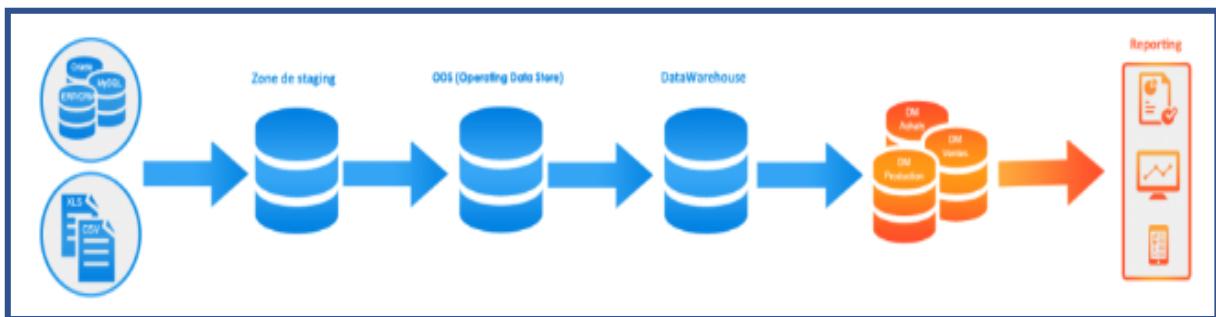


Figure 25 : Processus de ETL

1. Staging Area « STG »

Voici les quelques points que nous voudrons énumérer à propos de la zone de Staging Area.

La zone STG est un endroit où nous stockons des tables temporaires sur le serveur d'entrepôt de données. Les tables de transfert sont connectées à la zone de travail ou aux tables de faits. Nous avons essentiellement besoin d'une zone de stockage intermédiaire pour stocker les données et effectuer le nettoyage et la fusion des données avant de charger les données dans l'entrepôt. Elle peut être considérée comme une copie de la source.

En outre, il offre également une plate-forme pour effectuer le nettoyage des données.

Staging Area est une zone intermédiaire qui est un schéma temporaire utilisé pour :

- Effectuez un mappage à plat, c'est-à-dire renversez toutes les données sans appliquer de règles commerciales. L'introduction de données dans la mise en attente prendra moins de temps, car aucune règle ni transformation d'entreprise n'y est appliquée.
- Utilisé pour le nettoyage et la validation des données.

En fonction de la complexité de la règle de gestion, il se peut que nous ayons besoin d'une zone intermédiaire, le besoin fondamental d'une zone intermédiaire étant de nettoyer les données source et de les rassembler à un endroit. Fondamentalement, il s'agit d'une zone de base de données temporaire. Les données de la zone STG sont utilisées pour la suite du processus et peuvent ensuite être supprimées. [16]

Dans notre cas, Staging Area est un emplacement temporaire où les données des systèmes source sont copiées. Cette copie conforme des données de la base fournie par l'entreprise est traitée et nettoyée avant de les destiner à la zone de stockage intermédiaire de l'entrepôt de données.

En bref, toutes les données requises doivent être disponibles avant que les données puissent être intégrées à l'entrepôt de données. Elle contient une copie exacte des données source à partir desquelles nous pouvons extraire des données pour créer un ODS en appliquant des règles commerciales et des transformations et en intégrant des données.

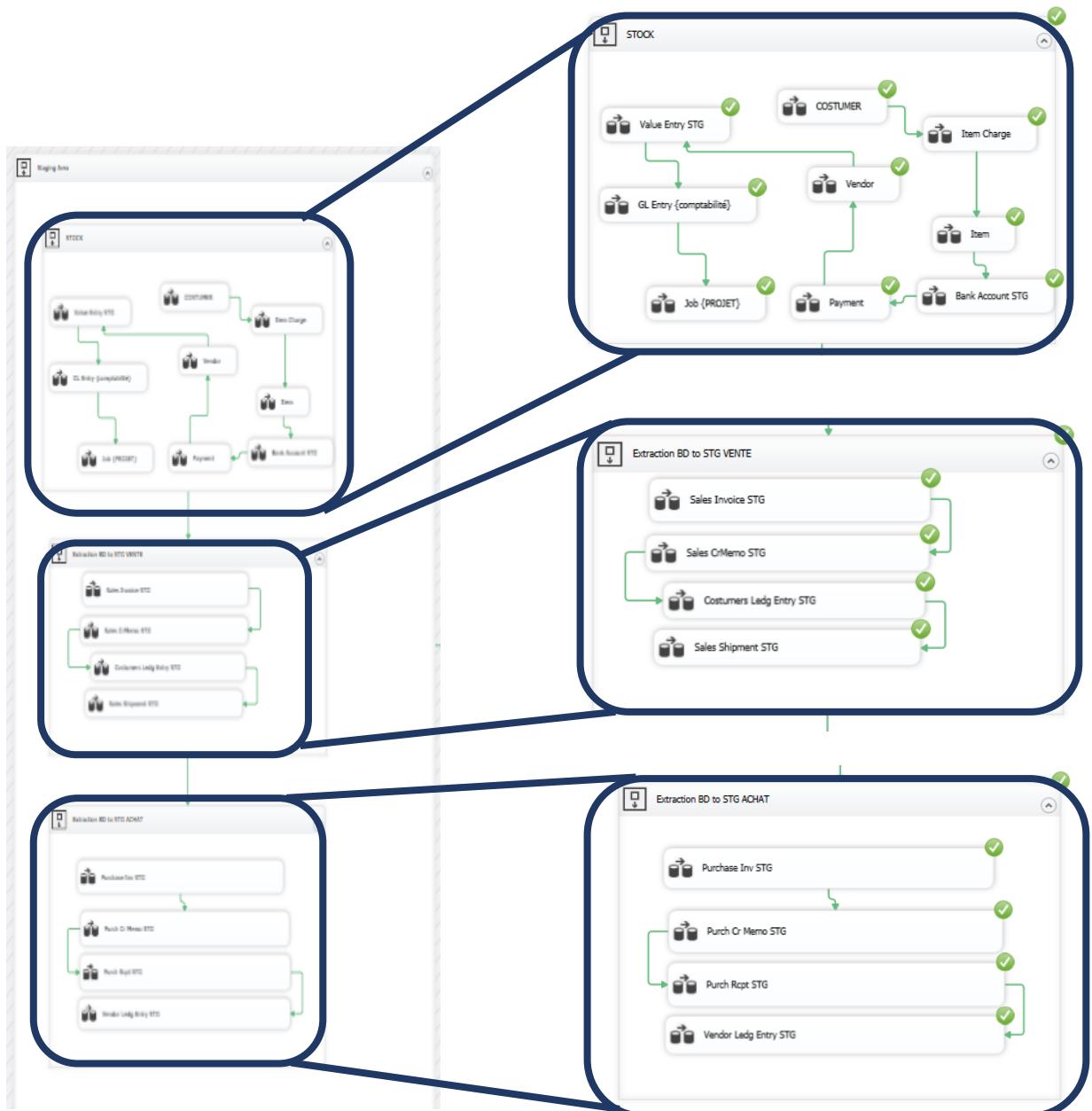


Figure 26 : Staging Area

Toutes les tables sont intégrées dans la zone temporaire ou le Staging Area et maintenant nous présentons quelque table pour expliquer le passage de donnée.



Figure 27 : Extraction de donnée pour charge produit et comptabilité générale

La source est la base fournie de l'entreprise et la destination correspond à la zone temporaire qui va avoir une copie conforme de la source.

Les connexions nécessaires pour cette étape sont des connexions OLE DB.

Source : NAV_SI

Destination : NAV_SI Staging Area

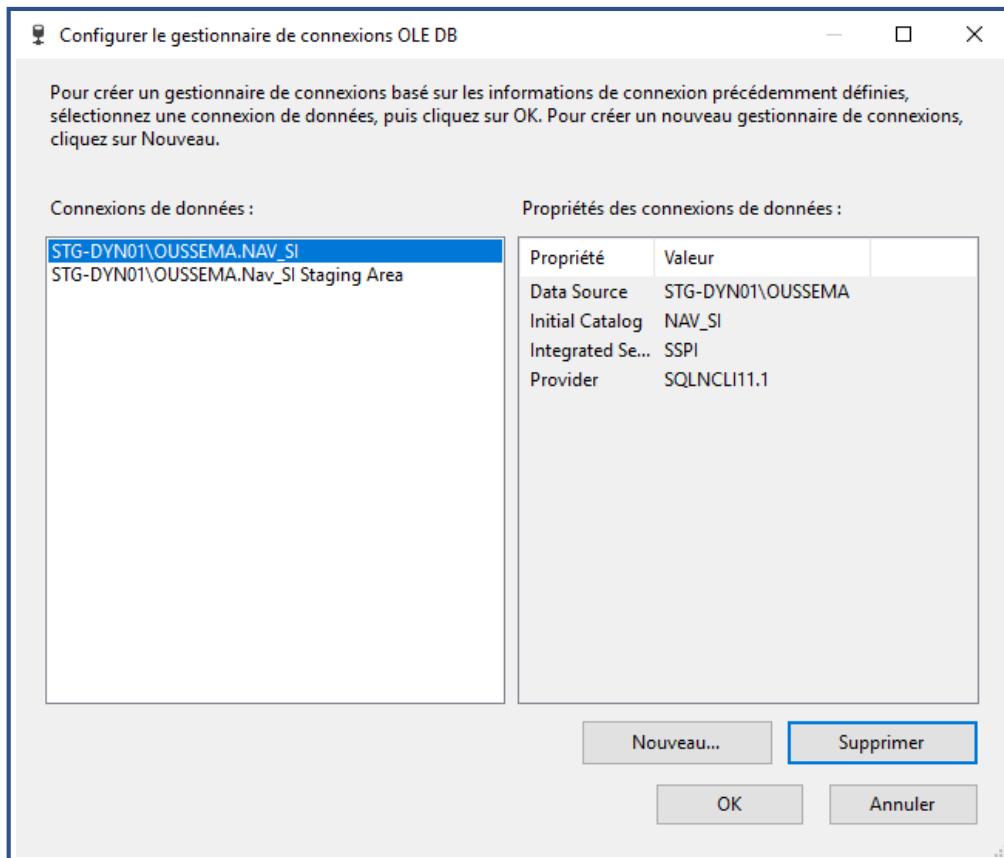


Figure 28 : Connexion OLE DB pour Staging Area

2. Operational Data Store « ODS »

L'Operational Data Store « ODS » est un type de base de données souvent utilisé comme zone logique intermédiaire pour un entrepôt de données.

Dans l'ODS, les données peuvent être effacées, résolues pour la redondance et vérifiées pour assurer leur conformité aux règles de gestion correspondantes. Un ODS peut être utilisé pour intégrer des données disparates provenant de sources multiples afin que les opérations commerciales, les analyses et les rapports puissent être effectués pendant les opérations commerciales. C'est l'endroit où sont stockées la plupart des données utilisées dans les opérations en cours avant leur transfert dans l'entrepôt de données pour un stockage ou un archivage à plus long terme.

L'ODS est conçu pour des requêtes relativement simples sur de petites quantités de données (telles que la recherche du statut d'une commande client), plutôt que sur des requêtes complexes sur de grandes quantités de données typiques de l'entrepôt de données. Un ODS est similaire à la mémoire à court terme en ce sens qu'il ne stocke que des informations très récentes ; en comparaison, l'entrepôt de données s'apparente davantage à une mémoire à long terme, en ce sens qu'il stocke des informations relativement permanentes. [17]

Les connexions nécessaires pour cette étape sont des connexions OLE DB.

Source : NAV_SI Staging Area

Destination : NAV_SI Operational Data Store

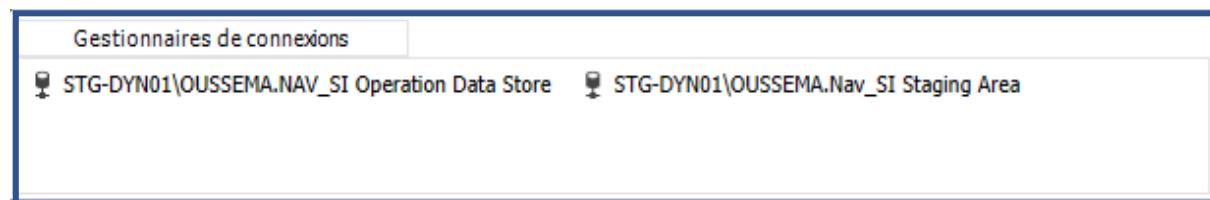


Figure 29 : Connexion Operational Data Store

Cette partie comporte deux parties de l'ETL qui sont la transformation des données et le chargement dans la zone ODS.

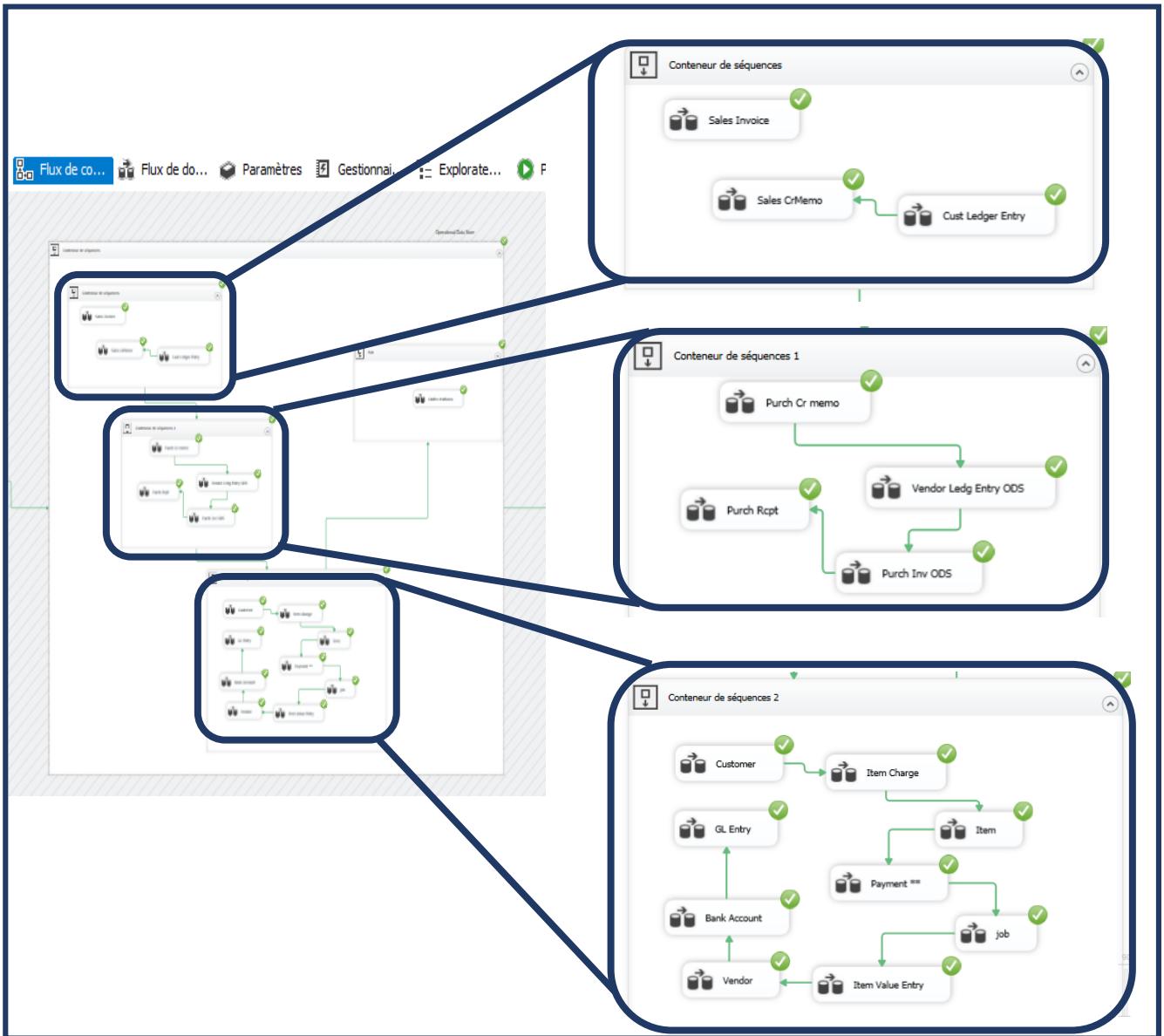


Figure 30 : Operational Data Store

Dans notre cas, nous avons eu recours à un seul package pour toutes les dimensions, nous avons utilisé « TRUNCATE TABLE » qui permet d'appliquer une requête SQL qui va supprimer le contenu de la table lors de chaque exécution de ce package pour éviter la redondance.

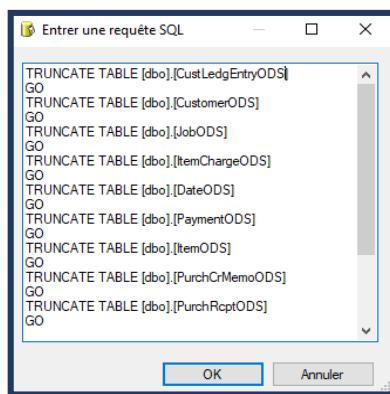


Figure 31 : Requête SQL Truncate Table

Afin de garantir le bon déroulement d'un ETL, il faut faire les changements et les transformations nécessaires sur les données recueillies.

Tout d'abords, nous possédons quelques données qui sont composées de deux tables (Entête et ligne). Pour faciliter l'intégration de ces données nous avons fait des jointures entre les deux tables pour obtenir une seule table avec la même clé primaire. Nous présentons l'exemple de la table avoir vente 'Sales Cr Memo' :

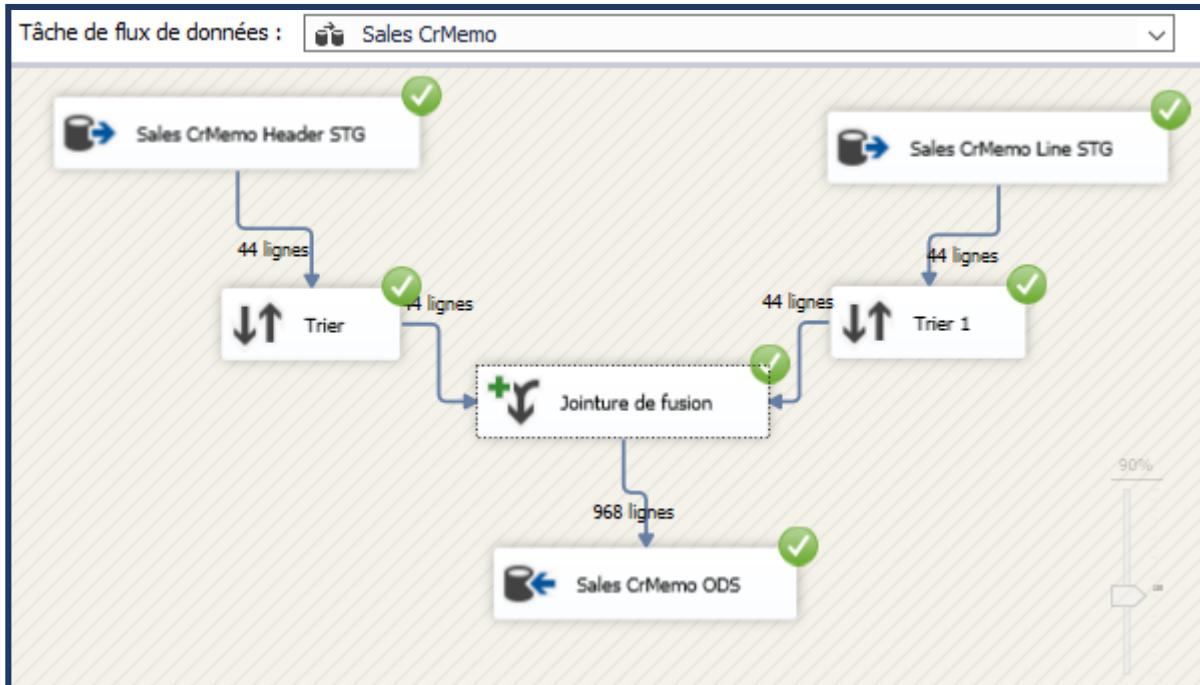


Figure 32: Jointure des tables entête et ligne

Nous avons aussi utilisé une autre forme de transformation qui est la fonction « dimension à variation lente » permettant d'effectuer une mise à jour pour les nouvelles données entrantes. Prenant l'exemple de la table expédition achat 'Purchase Reception' :

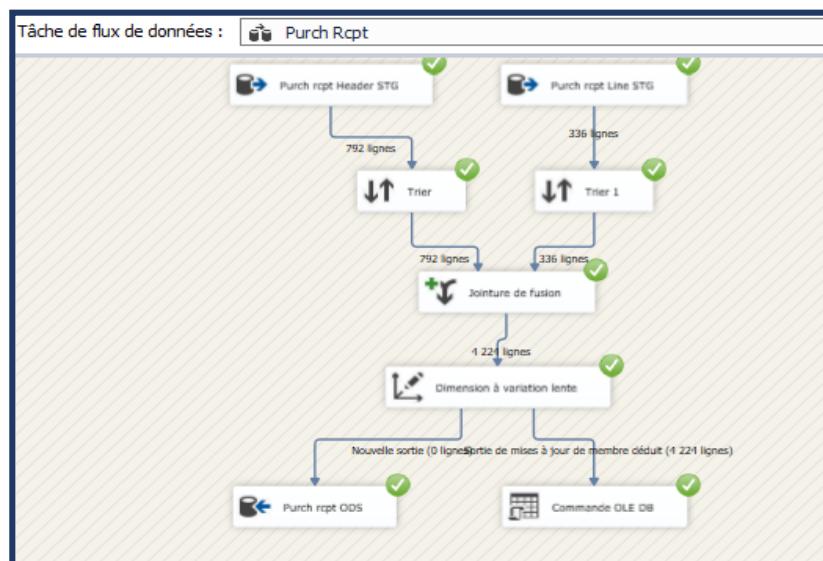


Figure 33 : Table Purchase Reception ODS

Pour que le tableau soit compréhensible et les rapports faciles à interpréter, il faut tout d'abord convertir les données, pour s'y faire nous devons changer de types de données en utilisant la fonction « colonne dérivé ». Nous avons aussi utilisé une autre forme de transformation qui est la fonction « agrégation » permettant d'effectuer des calculs en appliquant des fonctions telles que les opérations mathématiques et les combinaisons de variables.

Dans cette exemple, nous avons ajouté une nouvelle table intitulé ‘Chiffre d'affaires’ en appliquant les fonctions citées précédemment :

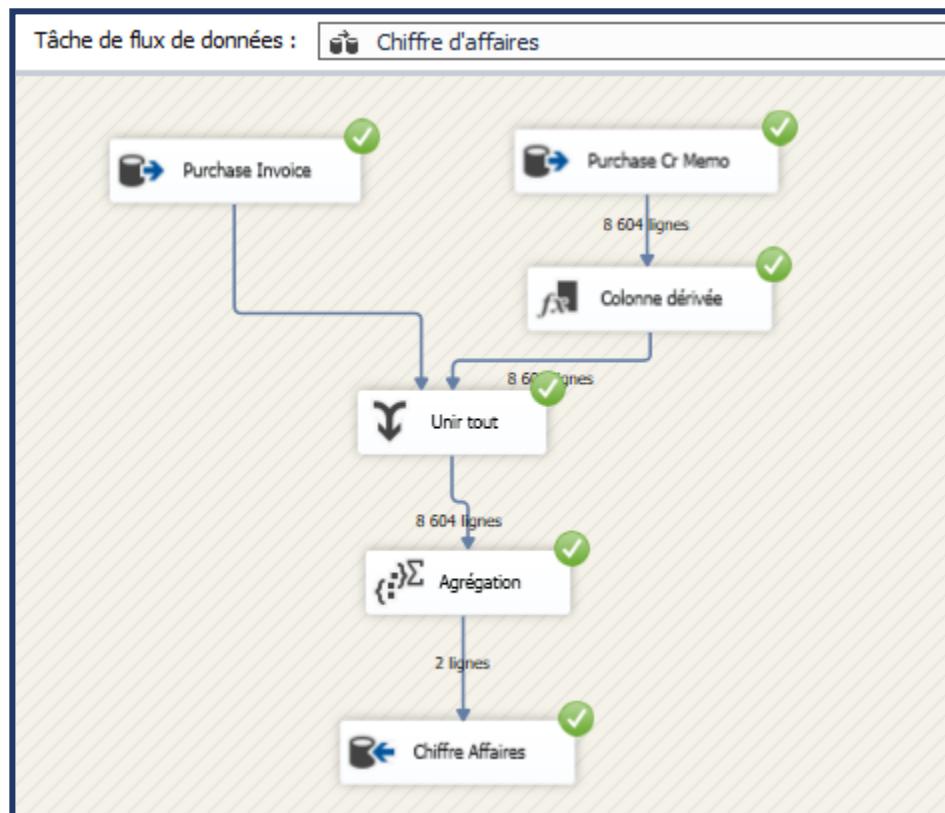


Figure 34 : Nouvelle Table Chiffre d'affaires achat

Les sources de donnée sont la table facture achat et la table avoir achat dont nous avons changé le signe des colonnes de la deuxième table comme présenté ci-dessous :

Nom de la colonne dérivée	Colonne dérivée	Expression
Quantity	Remplacer 'Quantity'	-Quantity
Amount	Remplacer 'Amount'	-Amount
Amount Including VAT	Remplacer 'Amount Includ...	-[Amount Including VAT]

Figure 35 : colonne dérivée avoir achat

3. Data Warehouse « DW »

Operational data store et les entrepôts de données « DW » stockent tous les deux des données opérationnelles, mais leurs similitudes s'arrêtent là et ils ont tous les deux un rôle à jouer dans les architectures d'analyse.

En fait, dans la plupart des implémentations de DW, nous utilisons toujours des classifications de processus métier qui nous aident à décomposer un DW dans des datamarts en fonction des questions commerciales que le BI posera à nos DW ou DMT (Data Marts). [18]

Après l'extraction « STG », la transformation des données et l'obtention des tables « ODS » qui répondent parfaitement à nos besoins, il faut charger ces données dans notre nouvelle structure qui n'est autre que notre entrepôt de données « DW » que nous avons au préalable réalisé dans SQL Server Management « Nav_SI DW » comme le montrent les deux figures 18 et 19.

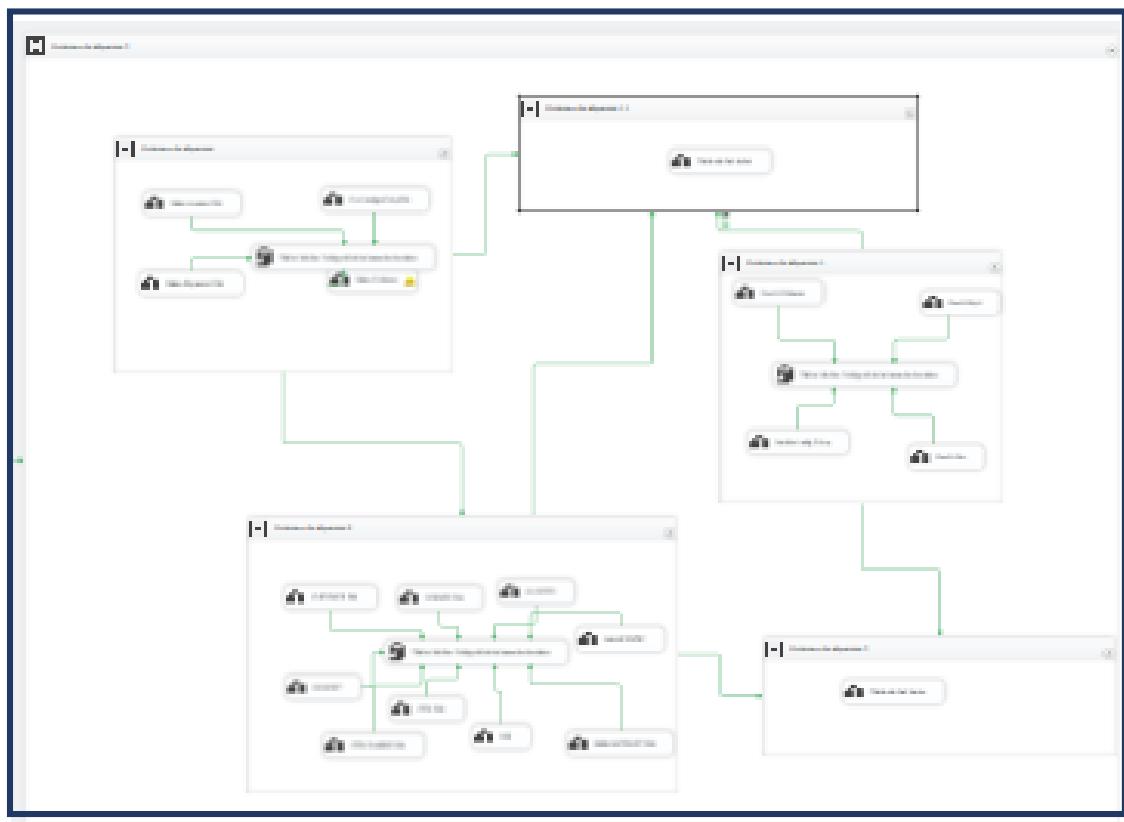


Figure 36 : Data warehouse

Nous avons utilisé une autre forme de transformation qui est une tâche pour vérifier l'intégrité de la base de données. Aussi, nous avons intégré deux tables faits : Table de fait vente et table de fait achat.

Les connexions nécessaires pour cette étape sont des connexions OLE DB.

Source : NAV_SI Operational Data Store

Destination : NAV_SI DW

Un entrepôt de données est un référentiel fédéré de toutes les données collectées par les différents systèmes opérationnels d'une entreprise. Le data warehousing met l'accent sur la saisie de données provenant de différentes sources pour un accès et une analyse.

4. Datamarts « DMT »

Datamart est un référentiel de données conçu pour servir une communauté particulière de travailleurs du savoir. Les datamarts permettent aux utilisateurs de récupérer des informations sur des départements ou des sujets particuliers, améliorant ainsi le temps de réponse des utilisateurs. Etant donné que les Datamarts cataloguent des données spécifiques, ils nécessitent souvent moins d'espace que les entrepôts de données d'entreprise, ce qui les rend plus faciles à rechercher et moins coûteux à exploiter. [19]

Le datamart n'utilise généralement que quelques sources de données, comparativement à un entrepôt de données. Ces datamarts sont de petite taille et sont plus flexibles par rapport à un entrepôt de données.

i. L'avantage de datamart

- Le Data Mart contribue à améliorer le temps de réponse de l'utilisateur grâce à la réduction du volume de données
- Il fournit un accès facile aux données fréquemment demandées.
- Les Data Marts sont plus simples à mettre en œuvre que les magasins de données d'entreprise. Dans le même temps, le coût de la mise en œuvre de Data Mart est certainement inférieur à celui de la mise en place d'un entrepôt de données complet.
- Comparé à Data Warehouse, un datamart est agile. En cas de changement de modèle, datamart peut être construit plus rapidement grâce à une taille plus petite.
- Un Data Mart est défini par un seul expert en la matière. Au contraire, l'entrepôt de données est défini par une PME interdisciplinaire appartenant à divers domaines. Par conséquent, le magasin de données est plus ouvert au changement que Datawarehouse.
- Les données sont partitionnées et permettent des priviléges de contrôle d'accès très granulaires.
- Les données peuvent être segmentées et stockées sur différentes plates-formes matérielles / logicielles.

ii. Mise en place des cubes

Les connexions nécessaires pour cette étape sont des connexions OLE DB.

Source : NAV_SI DW

Destination : NAV_SI Dimension

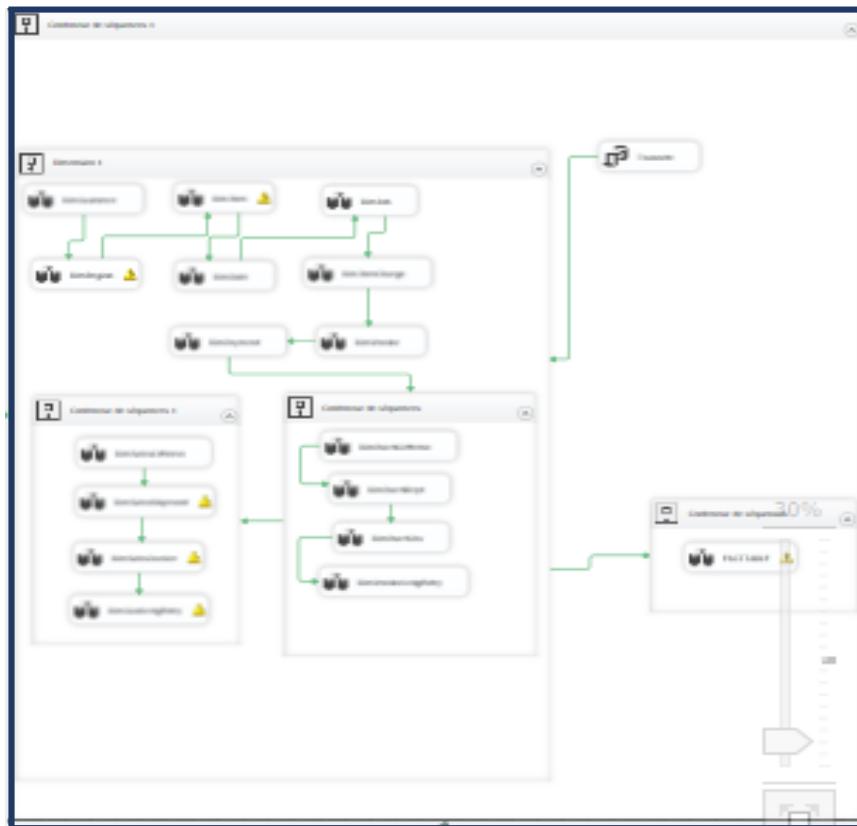


Figure 37 : Datamarts

Les cubes contiennent des informations plus pertinente que l'entrepôt de donnée, d'où nous devons sélectionner les détails généralement en se basant sur des axes particuliers.

A la création des datamarts, nous présentons l'exemple de dimension date, nous avons parcouru les dates de toutes les tables en ajoutant la DimTime qu'est une requête SQL sert à convertir de format de la date entré en format individuelle.

'JJ/MM/AAAA' devient jour 'JJ'

Mois 'MM'

Année 'AAAA'

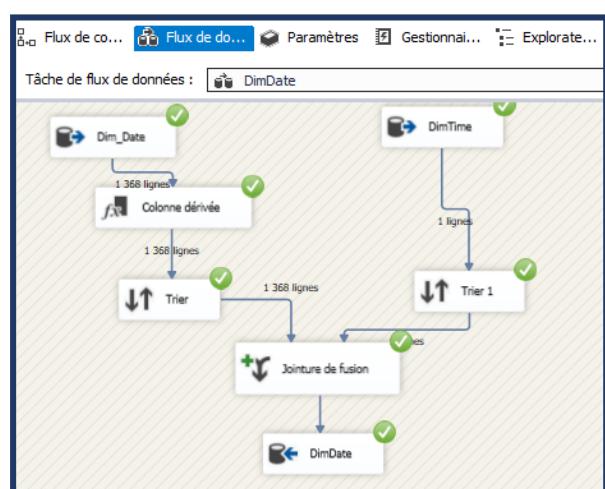


Figure 38 : Datamart date

Puisque nous avons plusieurs tables liées les unes aux autres, nous avons utilisé le composant « look up » en SSIS qui va permettre la recherche d'une valeur dans une autre source de données et de retourner les valeurs de colonnes associées ; nous l'avons aussi utilisé dans la table de faits, afin de trouver les colonnes associées à chaque clef étrangère

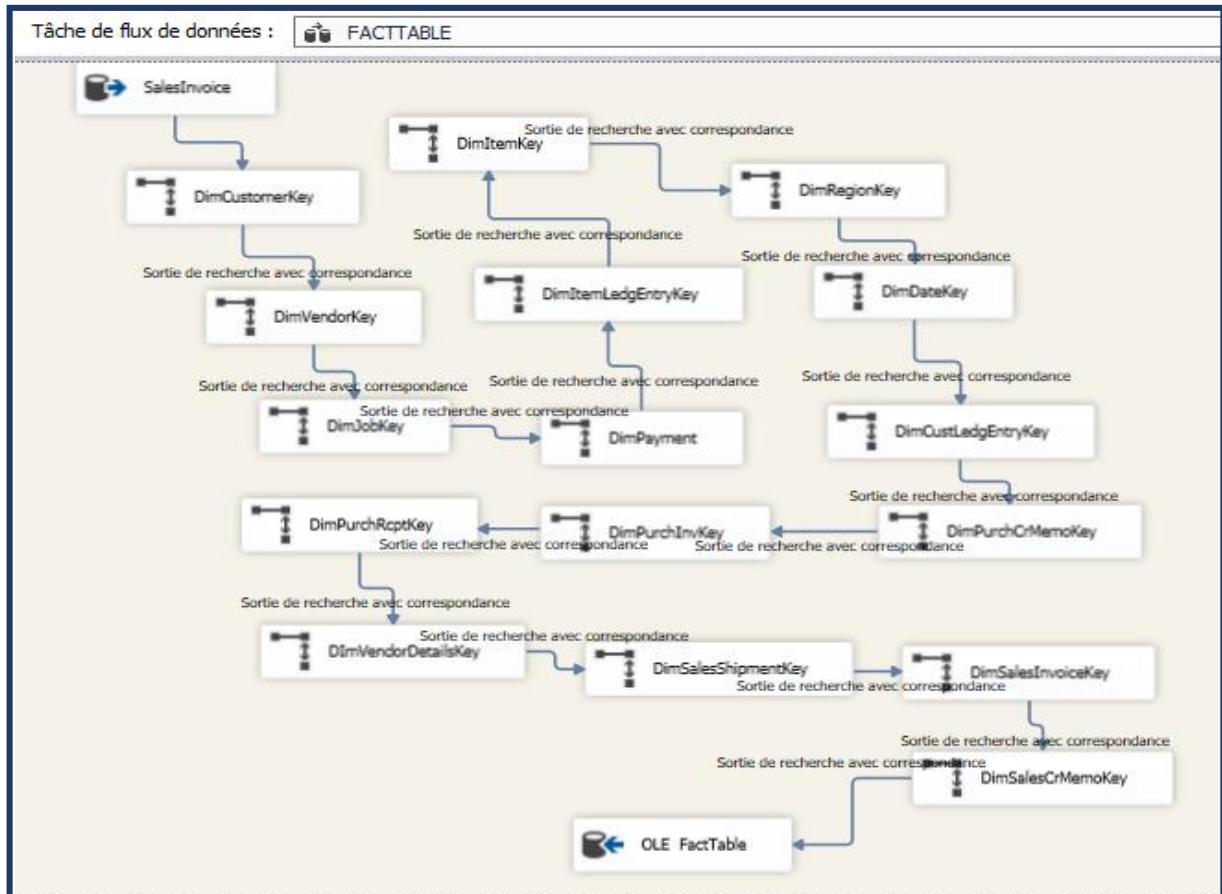


Figure 39 : Fact Table

Prenant l'exemple de la table produit 'Item'. Nous effectuons des lookups pour déterminer des clés étrangères que nous aurons besoin pour les produits. Aussi une nouvelle destination fichier plat pour les messages d'erreurs.

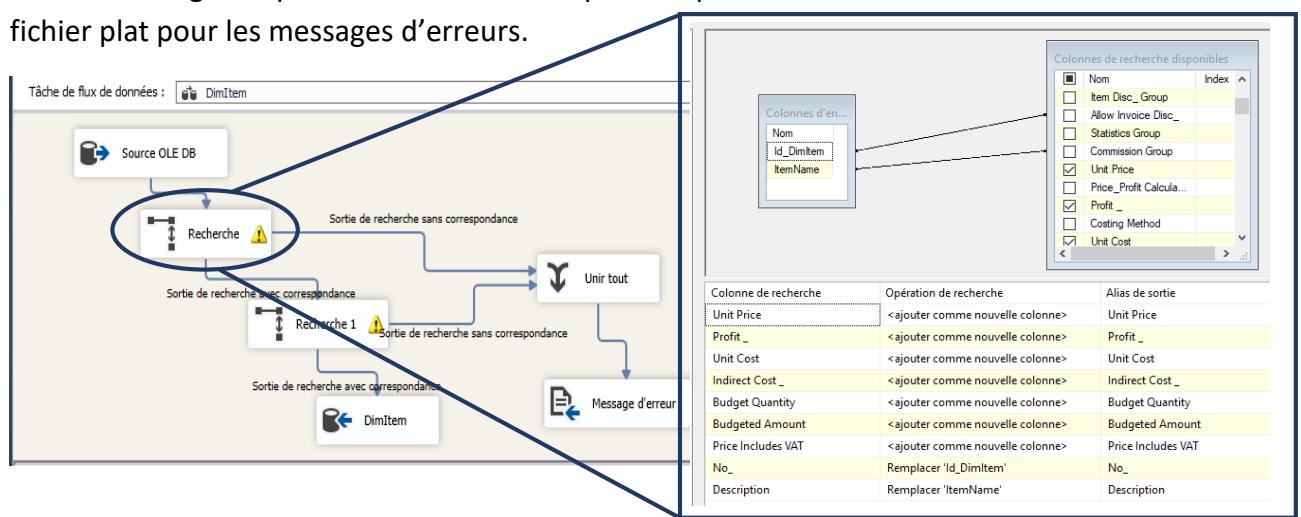


Figure 40 : Item datamart

Une fois le projet d'intégration de données est prêt, une planification automatique via un job SQL sera possible à l'aide de SQL Server Agent dans le SSMS.

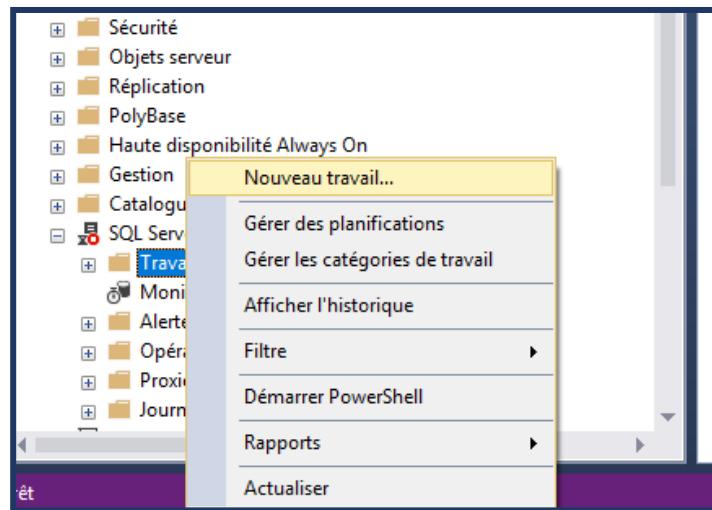


Figure 41 : Création d'un nouveau Job

Création et planification du Job dans SQL Agent :

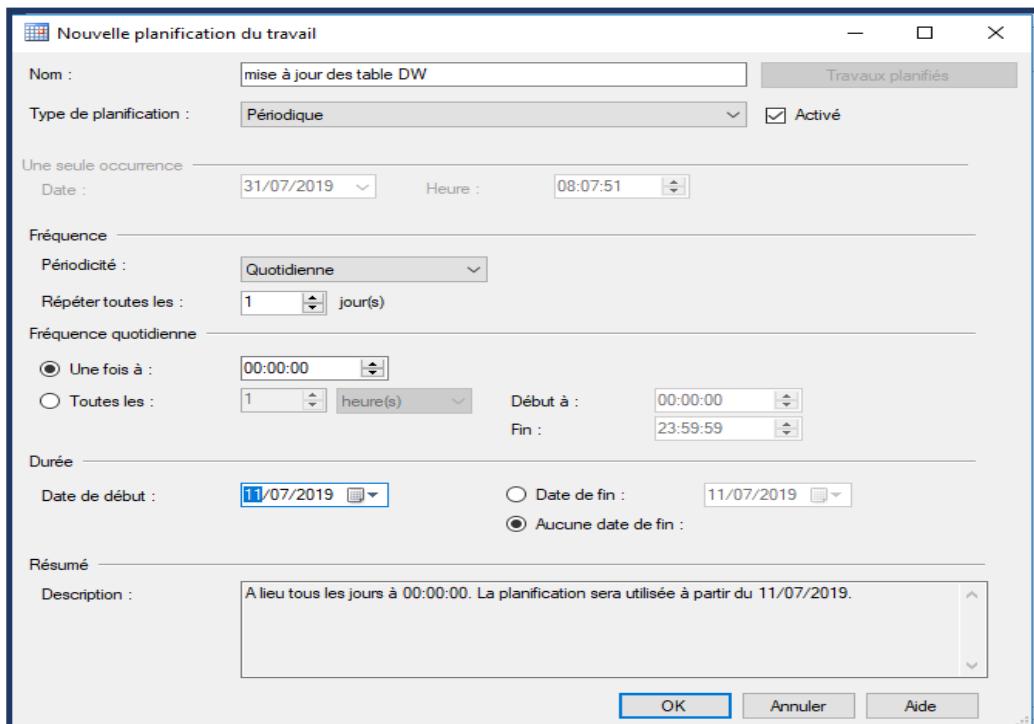


Figure 42 : Planification du lancement package

V. Package SSAS et modélisation de données

1. Présentation de SQL Server Analysis Services (SSAS)

Analysis Services fournit des fonctionnalités OLAP (Online Analytical Processing) et d'exploration de données pour les applications de Business Intelligence.

Analysis Services prend en charge OLAP et permet de concevoir, créer et gérer des structures multidimensionnelles qui contiennent des données agrégées à partir d'autres sources de données, telles que des bases de données relationnelles.

Pour les applications d'exploration de données, Analysis Services permet de concevoir, créer et visualiser des modèles d'exploration de données.

Ces modèles d'exploration de données peuvent être construits à partir d'autres sources de données au moyen d'un large éventail d'algorithmes standards d'exploration de données.

2. Package SSAS

Après la définition des mesures, des dimensions et de la table de faits, nous allons finalement procéder à la modélisation du schéma conceptuel de notre entrepôt de données qui se présente comme suit :

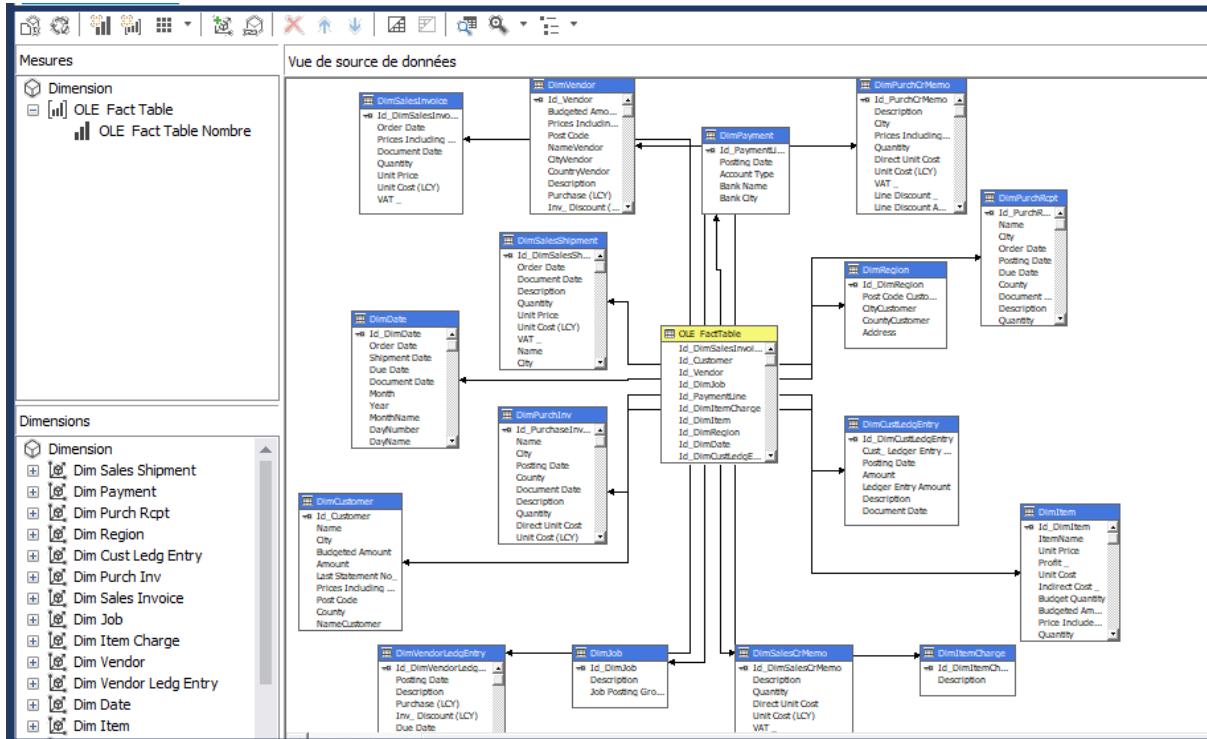


Figure 43 : Modélisation des données

Ce schéma évite les redondances d'informations mais nécessite des jointures lors des agrégats de ces dimensions.

La première composante : la table de Faits ‘OLE FactTable’ (Table des mesures)

C'est la table qui contient les données observables (les faits) que l'on possède sur un sujet et que l'on veut étudier, selon divers axes d'analyse (les dimensions).

Les « faits », dans un entrepôt de données, sont normalement numériques, puisque d'ordre quantitatif. Il peut s'agir du montant en argent des ventes, des achats, du nombre d'unités vendues d'un produit, etc.

La deuxième composante : les tables de Dimensions (Axe d'analyse)

Ce sont les tables qui contiennent les axes d'analyse (les dimensions) selon lesquels on veut étudier des données observables (les faits) qui, soumises à une analyse multidimensionnelle, donnent aux utilisateurs des renseignements nécessaires à la prise de décision.

Le schéma en flocons de neige est une variante du schéma en étoile.

Il est donc tout simplement question de mettre les attributs de chaque niveau hiérarchique dans une table de dimension à part. Cette seconde couche d'abstraction présente un autre avantage qui est la capacité à représenter la relation entre les attributs d'une dimension.

Toute dimension contient des attributs qui définissent une granularité plus ou moins importante. Il peut y avoir plusieurs hiérarchies dans une table de dimension.

Chapitre 5

Sprint 2 : Elaboration des tableaux de bord

Introduction

Comme mentionné dans le chapitre 2, Power BI représente notre choix pour la génération des rapports et des tableaux de bord. Afin de commencer la réalisation des tableaux de bord, nous avons connecté notre base de données multidimensionnelle, ce qui nous a permis d'avoir une visibilité sur les dimensions et ainsi les sélectionner, comme le montre la figure suivante :

The screenshot shows the 'Navigateur' (Navigator) window in Power BI. On the left, there is a tree view of available tables under 'STG-DYN01\OUSSEMA [15]'. Several dimensions are selected, indicated by checked checkboxes: DimCustLedgEntry, DimCustomer, DimDate, DimItem, DimJob, DimPayment, DimPurchCrMemo, DimPurchInv, DimPurchRcpt, and DimRegion. To the right, a preview of the 'DimCustomer' table is displayed, showing columns 'Name', 'Id_Customer', and 'City'. The preview is dated 'Aperçu téléchargé le jeudi 11 juillet 2019'. At the bottom, there are buttons for 'Sélectionner les tables associées' (Select associated tables), 'Charger' (Load), 'Modifier' (Edit), and 'Annuler' (Cancel).

Name	Id_Customer	City
Spotsmeyer's Furnishings	01121212	Miami
Spotsmeyer's Furnishings	01121212	Miami
Spotsmeyer's Furnishings	01121212	Miami
Progressive Home Furnishings	01445544	Chicago
Progressive Home Furnishings	01445544	Chicago
Progressive Home Furnishings	01445544	Chicago
New Concepts Furniture	01454545	Atlanta
New Concepts Furniture	01454545	Atlanta
New Concepts Furniture	01454545	Atlanta
Candoxy Canada Inc.	01905893	Thunder Bay
Candoxy Canada Inc.	01905893	Thunder Bay
Candoxy Canada Inc.	01905893	Thunder Bay
Elkhorn Airport	01905899	Elkhorn
Elkhorn Airport	01905899	Elkhorn
Elkhorn Airport	01905899	Elkhorn

Figure 44 :Sélection des dimensions et table de fait

1. Présentation de SQL Server Reporting Services (SSRS)

SQL Server Reporting Services fournit une gamme complète d'outils et de services prêts à l'emploi pour vous aider à créer, déployer et gérer des rapports pour votre organisation, ainsi que des fonctions de programmation pour vous permettre d'étendre et de personnaliser vos fonctionnalités de création de rapports. Reporting Services est une plateforme serveur qui fournit des fonctionnalités complètes de création de rapports pour différentes sources de données. Reporting Services inclut un jeu complet d'outils que vous pouvez utiliser pour créer, gérer et livrer des rapports, ainsi que les API qui permettent aux développeurs d'intégrer ou d'étendre des données et de traiter le rapport dans des applications personnalisées. Les outils

Reporting Services s'exécutent dans l'environnement Microsoft Visual Studio et sont totalement intégrés aux outils et composants SQL Server.

2. Réalisation des tableaux de bord

L'implémentation du projet a été faite en deux phases. Dans la première phase, nous avons préparé une maquette d'une version standard et la deuxième phase qui consiste à l'implémentation du spécifique pour l'entreprise.

Dans la version standard, nous avons utilisés différents types de graphs (Camembert, histogramme etc.) Nous analysons le chiffre d'affaires des ventes dans le tableau de bord à travers différents dimensions (pays, vendeur, catégorie, société) et nous visualisons l'évolution par rapport à l'année qui précède.

Dans le cas où l'évolution du chiffre d'affaires est décroissante c'est au décideurs d'intervenir. Ce Dashboard a été réalisé à l'aide de l'outil SSRS

Après la préparation et l'importation des données dans power BI, nous enchaînerons avec l'analyse des tableaux de bord générés. Cet état permet de visualiser les achats par Groupe et par engin. L'état dégage le coût de l'article directe (Facture Fournisseur) et coût de revient (Le cout direct avec les frais annexes qui lui sont affectées) Le pourcentage de frais permet de dégager les anomalies au niveau de la saisie des factures achats (Soit retard de saisie ou factures annexes non encore parvenues)

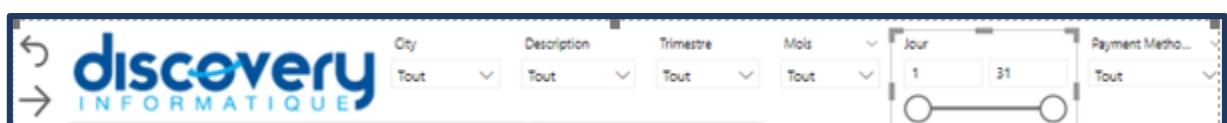


Figure 45: Filtres Dashboard

Nous avons élaboré les filtres des dashboard selon la demande de besoin et selon les classifications nécessaires pour visualisé les KPIs.

Conclusion générale

Ce stage représente pour nous une expérience enrichissante dans la mesure où il nous a permis d'enrichir nos compétences techniques et de découvrir de plus près notre domaine dans le milieu professionnel. Nous avons eu le privilège de travailler dans un milieu agréable et coopératif au sein du département des systèmes d'information.

Grâce à ce projet de fin d'études, nous avons appris à gérer notre stress et mener un projet du début à la fin. Le début était difficile, étant donné que nous étions dans l'obligation de se familiariser avec les métiers et le jargon de ce domaine. Nous avons commencé par la suite à analyser les technologies adoptées tout en mettant l'accent sur leurs différentes spécificités.

Ensuite, nous avons élaboré notre plan d'exécution avec Scrum BI. Mettant fin à la théorique, les phases de réalisation et implémentation des interfaces viennent enfin pour recueillir le fruit de ces études et analyses. Nous avons commencé par la compréhension du produit. Puis, nous avons fait une répartition, qui nous a permis de terminer le travail à temps. Non seulement ce travail nous a permis de connaître une nouvelle méthode de gestion de projet mais aussi de la modeler efficacement en fonction de notre projet. Pour finir, nous avons mis au point un data warehouse qui nous a mené à développer les tableaux de bord.

Comme dans chaque expérience, nous avons rencontré des obstacles que nous avons réussi à surmonter.

A présent, le Reporting peut répondre à tous les besoins en gestion des achats et des ventes de **Discovery Informatique** en offrant aux responsables informatiques locaux la possibilité de gérer les rapports selon leurs besoins.

Nous tenons à préciser que ce projet de fin d'étude nous a été très bénéfique, non seulement parce qu'il nous a offert l'opportunité d'améliorer nos connaissances techniques et de les approfondir dans le domaine de programmation en Business Intelligence, mais surtout parce que nous avons réussi à nous améliorer professionnellement (Soft-Skills : communication et sens de responsabilité) car nous étions confronté à plusieurs difficultés et il fallait gérer des réunions et collaborer avec différentes équipes ainsi que de prendre les bonnes décisions pour réussir le projet. Nous espérons que notre travail reflète bien la bonne formation SID⁶ que nous avons suivie au sein de l'école Esprit school of business en partenariat avec Claude Bernard Lyon 1.

Enfin nous tenons à exprimer que cette expérience nous a été très enrichissante et importante car elle marque la fin de cycle double licences et nous a permis d'être confronté aux responsabilités qui incombent à un consultant junior en informatique décisionnelle : faire face aux délais, au stress, aux contraintes du travail au sein de l'entreprise.

⁶ SID : Système d'information décisionnel

Référence :

- [1] : discoveryinformatique.com/fr/content/presentation
- [2] : geek-directeur-technique.com/2009/02/04/le-cycle-en-v
- [3] : piloter.org/business-intelligence/business-intelligence-agile.htm
- [4] : blog.keyrus.co.uk/scrum_training_for_bi_teams.html
- [5] : computerweekly.com/tip/Inmon-or-Kimball
- [6] : aerow.group/a16u1509/
- [7] : waytolearnx.com/2018/08/difference-entre-le-schema-en-etoile.html
- [8] : notrefamille.com/dictionnaire/definition/flocon/
- [9] : code.visualstudio.com/docs/editor/whyvscode
- [10] : cosmoconsult.com/products/erp/microsoft-dynamics-nav-navision/
- [11] : docs.microsoft.com/en-us/sql/ssms/sql-server-management-studio-2017
- [12] : searchcontentmanagement.techtarget.com/definition/Microsoft-Power-BI
- [13] : educba.com/power-bi-vs-tableau-vs-qlik/
- [14] : it.toolbox.com/question/need-of-ods-in-etl-031308
- [15] : lebigdata.fr/etl-definition
- [16] : dwhlaureate.blogspot.com/2012/06/what-is-staging-area-and-what-is.html
- [17] : searchoracle.techtarget.com/definition/operational-data-store
- [18] : panoply.io/data-warehouse-guide/
- [19] : searchdatamanagement.techtarget.com/definition/data-mart

Annexes

Script SQL pour la dimension Date

```
CREATE TABLE DimTime
(
    DateID      INT          NOT NULL  PRIMARY KEY
    ,MonthName   VARCHAR(9)   NOT NULL
    ,MonthNumber INT          NOT NULL
    ,DayNumber   INT          NOT NULL
    ,DayName     VARCHAR(9)   NOT NULL
    ,Quarter    INT          NOT NULL
    ,QuarterName CHAR(7)     NOT NULL
    ,Year        INT          NOT NULL
    ,FullDate DATETIME     NOT NULL
)
BEGIN
    SET DATEFIRST 1
    DECLARE @startDate  DATETIME
    DECLARE @endDate    DATETIME
    DECLARE @date       DATE
    DECLARE @DateId    INT
    DECLARE @cpt INT
    SET @startDate = '2009-01-01'
    SET @cpt = 0
    SET @date = DATEADD(dd, @cpt, @startdate)
BEGIN
    SET @DateId = CONVERT(VARCHAR(8), @date, 112)
    INSERT INTO DimTime(DateID,MonthName,MonthNumber,DayNumber,DayName,Quarter,QuarterName,Year,FullDate)
    VALUES ( @DateID
    ,DATENAME(mm, @date)
    ,DATEPART(mm, @date)
    ,DATEPART(dd, @date)
    ,DATENAME(DW, @date)
    ,DATEPART(qq, @date)
    , 'Q' + DATENAME(qq, @date)
    ,DATEPART(yy, @date)
    , @date )
    SET @cpt = @cpt + 1
    SET @date = DATEADD(dd, @cpt, @startdate)
END
END
```

Figure 45: Script Date

Exemple « langage MDX »

- Chiffre d'affaires année en cours mois-1

```
CREATE MEMBER CURENTCUBE.[Measures].[CA Mois N-1]
As ([Posting Date].[Mois].prevMember, [Measures].[CA]),
```

- Chiffre d'affaires année et mois en cours

```
CREATE MEMBER CURENTCUBE.[Measures].[CA Mois En Cours]
As ([Posting Date].[Année-Mois].currentMember, [Measures].[CA]),
```

- Quantité année en cours mois-1

```
CREATE MEMBER CURENTCUBE.[Measures].[Qte Mois N-1]
As ([Posting Date].[Année-Mois].prevMember, [Measures].[QTE]),
```

- Quantité année et mois en cours

```
CREATE MEMBER CURENTCUBE.[Measures].[Qte Mois En Cours]
As ([Posting Date].[Année-Mois].currentMember, [Measures].[QTE]),
```

- Chiffre d'affaires année-1

```
CREATE MEMBER CURENTCUBE.[Measures].[CA Mois N-1]
As ([Posting Date].[Mois].prevMember, [Measures].[CA]),
```

- Chiffre d'affaires par date

```
CREATE MEMBER CURENTCUBE.[Measures].[Cumul CA]
As Aggregate
( PeriodsToDate
[Posting Date].[AnneeMois].[ Annee] , [Date].[AnneeMois].currentMember ),
[Measures].[CA]),
```

- Requête MDX pour calculer l'évolution du CA

```
CREATE MEMBER CURRENTCUBE.[Measures].[Evol CA]
```

```
IIF( ISEMPTY( [Measures].[Cumul CA ME N-1] ) or [Measures].[Cumul CA ME N-1]=0 , null,
([Measures].[Cumul CA N]-[Measures].[Cumul CA ME N-1])/[Measures].[Cumul CA ME N-1] )
```

- Requête MDX pour calculer Production A DATE

```
Aggregate (lastperiods(10000, [Période].[Année-Mois].[Mois].HIERARCHY.CURRENTMEMBER) ,
[Measures].[Qte] )
```

Résumé :

Ce rapport se compose de cinq chapitres, au cours desquelles nous avons mentionné les phases de création de notre solution décisionnelle qui servira pour Discovery Informatique comme un outil d'aide à la décision. Commençant par la phase de modélisation, passant par la phase d'analyse et de restitution jusqu'à arriver à la dernière phase élaboration des Dashboard. Ce rapport comporte aussi les choix des outils, méthodologie de travail ainsi que la présentation de l'organisme d'accueil.

Mots-clés :

Informatique décisionnelle - Analyse de donnée – Base de donnée – Modélisation multidimensionnelle – cube – Data Warehouse – Scrum BI – Dashboard – Modèle en Etoile

Abstract:

This report is composed of five chapters, during which we mentioned the creation phases of our Business Intelligence solution that will be used by Discovery Informatique as a decision support tool. Starting from the modeling phase, through the analysis and rendering phase until reaching the final development phase of the Dashboard. This report also includes the choice of tools, work methodology and the presentation of the host organization.

Key words:

Business Intelligence - Data Analysis - Database - Multidimensional Modeling - Cube - Data Warehouse - Scrum BI - Dashboard - Star Model