

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Blida 1
Faculté des Sciences
Département d'Informatique



Mémoire de fin d'études

Pour l'obtention du diplôme de Master en Informatique

Option : Système Informatique et réseaux

Thème

**Développement d'une solution Business Intelligence
pour le Département Suivi et Inspection
DTE/RDC/SADEG**

Organisme d'accueil : **SONELGAZ**



Réalisé par :

Mlle GARZOU Fatima
Mlle HADJADJA Fatma

Promoteur : Mr M. BENYAHIA
Encadreur : Mr M. YAHOUI

Devant les jurys :

Président : Mme K. MESKALDJI
Examineur : Mme I. CHERFA

Promotion 2020/2021

Remerciements

En tout premier lieu, on remercie le bon dieu, tout puissant, de nous avoir donné la force pour survivre, ainsi que l'audace pour dépassé toutes les difficultés

Nos remerciements vont tout spécialement à nos familles, qui ont sus nous supporter et encourager tout au long de notre vie, ainsi que pour leur aide inestimable, leur patience et leur soutien indéfectible.

Nous tenons aussi, à remercier tout les enseignants qui ont contribué de près ou de loin à notre formation.

Nous remercions Mr M. Benyahia et Mr M. Yahoui pour avoir assuré l'encadrement de ce projet, qui n'a pas toujours été de tout repos.

On remercie monsieur Benyahia pour la gentillesse et la patience qu'il a manifestée à notre égard, son disponibilité et ces conseils pour l'accomplissement de ce travail.

Nous remercions monsieur Yahoui pour la confiance quel nous a accordé et de nous avoir donné l'opportunité de travailler sur un projet d'une tel envergure.

Nous nous devons de mentionner la précieuse et totale collaboration que nous avons reçu au sein de SONELGAZ, l'aide et le support apporté par l'ensemble des employés et des cadres.

On remercie vivement Mesdames et Messieurs les membres du jury d'avoir accepter d'évaluer ce travail.

Pour finir, et afin de n'oublier personne (amis, membre de la famille et tous ceux qui nous sont chers) nous utiliserons la formule : « Merci à ... ».

Fatima && Fatma

Dédicaces

Je dédie ce modeste travail à :

A mes chers parents Que Dieu ait pitié d'eux et leur accorde Jannat al-Firdaws.

A mes cher frères et sœurs,

En témoignage de mes sincères reconnaissances pour les efforts qu'ils ont consenti pour poursuivre mes études, je leur dédie ce modeste travail en témoignage de mon grand amour et ma gratitude infinie.

A tous mes amis,

Pour leur aide et leur soutien moral durant l'élaboration ce travail de fin d'étude

A tous ceux qui m'aiment...

Fatima

Dédicaces

*A celle qui a su me consolider durant les moments les plus difficiles de ma vie, mon mode
D'affection et de bonté*

..Ma mère, Ma mère, Ma mère

A celui qui m'a guidé par ses précieux conseils, qui a été toujours présent dans les pénibles

Moments.

..Mon père,

A mes frères, particulièrement Mohammed et sa femme, Abderrezak et sa femme Et Mounir

A mes très chères sœurs, Souad, Hassina, Karima et Souhila

*A Ma nièce katre ennada, tasnim, sirine, Anfalle, Sana, et mes neveux Oussama, ayoub et
Ibrahim, ahmed, wassim, iyad, abd rahim.*

A tous mes cousins, mes oncles, mes tentes et leurs familles,

A tous mes amis surtout : lamia, Ferial, asma, hamida, fathila, akila, fatima,

A mon binôme, fatima

*A tous ceux qui, de loin ou de près, n'ont cessé de m'apporter leur soutien tout au long de
mes études.*

Fatma

Table des matières

Résumer

ملخص

Abstract

Abréviation

Liste des figures

Liste des Tableaux

Introduction générale

1) Contexte général.....	16
2) Problématique	17
3) Objectifs du projet.....	18
Partie 1: Synthèse bibliographique	
I Introduction	21
II Les systèmes décisionnels	21
1) Définition Système d'information :.....	21
2) Définitions du décisionnel :.....	22
3) Composants des systèmes décisionnels.....	23
2.1) Les sources de données:.....	24
2.2) Outil d'extraction, transformation et de chargement des données sources :.....	24
2.3) L'entrepôt de données (Data Warehouse):.....	24
2.4) Les outils de présentation et d'analyse de données:.....	25
4) Système d'information décisionnel versus système d'information	25
III Le Data Warehouse support d'aide à la décision	26
1) Concepts et Définitions.....	26
2) Objectifs du Data Warehouse.....	27
3) Structure des données d'un Data Warehouse.....	30
3.1) Données détaillées :	30
3.2) Données détaillées archivées :.....	30
3.3) Données agrégées :	31
3.4) Données fortement agrégées :.....	31
3.5) Meta données :.....	31
4) Composants du Data Warehouse.....	31
4.1) Les systèmes sources.....	32
a) Données de production	32

b)	Données internes	32
c)	Données archivées	32
d)	Données externes.....	32
4.2)	La zone de préparation de données.....	33
4.3)	La zone de présentation de données.....	34
4.4)	Le portail de restitution de données.....	36
5)	Architecture d'un Data Warehouse.....	37
IV	Modélisation des données de Data warehouse.....	38
1)	Modélisation dimensionnelle et ses concepts.....	38
2)	Tables de faits et tables de dimensions :.....	39
2.1)	Table de fait.....	39
2.2)	Table de dimension :.....	39
3)	Différents modèles de la modélisation dimensionnelle :.....	40
3.1)	Modèle en étoile	40
3.2)	Modèle en flocon.....	41
3.3)	Modèle en constellation	41
3.4)	Modèle en grappe ou Modèle mixte.....	41
4)	Le cube de données :.....	41
5)	Le concept OLAP :.....	42
5.1)	Généralités	43
5.2)	Architecture des serveurs OLAP :.....	43
a)	Architecture MOLAP « Multidimensional OLAP » :.....	43
b)	Architecture ROLAP « Relational OLAP »:.....	44
c)	Architecture HOLAP « Hybrid OLAP »:.....	44
d)	Autres architectures OLAP :.....	45
e)	Comparaison MOLAP Vs. ROLAP Vs. HOLAP :.....	45
V	Modélisation et conception du Data Warehouse.....	46
1)	Les méthodologies conductrices à la mise en place d'un Data Warehouse :.....	46
1.1)	Top-Down :.....	46
1.2)	Bottom-Up :.....	46
1.3)	Middle-Out :.....	46
2)	Alimentation du Data Warehouse.....	47
2.1)	Définition d'ETL.....	47
2.2)	Phases constitutives d'un ETL.....	47
3)	Le déploiement et mise en oeuvre :.....	48
3.1)	Requête ad hoc :.....	48

3.2) Reporting:	48
3.3) Analyse dimensionnelle des données :.....	49
3.4) Tableau de bord :.....	49
3.5) Data mining :.....	49
4) Maintenance et croissance :.....	50
5) Conclusion	50

Partie 2: Cas pratique « Région de distribution centre RDC/SADEG filiale SONELGAZ »

Chapitre 1: Présentation de l'organisme d'accueil

1) Introduction	52
2) Présentation de SONELGAZ	53
2.1) Historique.....	53
2.2) Activités et missions de SONELGAZ :.....	57
3) Le métier de la distribution.....	58
3.1) Organisation de Région de Distribution Centre.....	59
3.2) Direction de Distribution.....	60
3.3) Département suivi et Inspection	61
4) Conclusion	62

Chapitre 2 : Etude de l'existant

1) Introduction	63
2) Processus de réalisation des programmes d'électrifications.....	63
2.1) Lancement des programmes d'électrification.....	63
2.2) Etablissement des listings des affaires au niveau des communes.....	63
2.3) Signature de convention entre ministère de l'énergie et la SADEG	64
2.4) Lancement des consultations	64
2.5) Suivi physique et financier des affaires.....	64
2.6) Clôture des affaires	65
3) Etat du décisionnel au sein de RDC(DSI)	66
4) Conclusion	68

Chapitre 3: Etude des besoins

1) Introduction	69
2) Démarche adopté pour l'identification des besoins.....	70
3) Identification des besoins	71
4) Les problèmes rencontrés.....	72
5) Conclusion	72

Chapitre 4: Conception de la solution

1) Introduction	73
2) Conception d'entrepôt de donnée	73
2.1) Modélisation dimensionnelle	73
2.2) Modélisation des activités	73
a) Activité « suivi des affaires ».....	73
b) Activité « suivi la clôture d'une affaire ».....	80
c) Activité « suivi des entreprise »	82
3) Alimentation de l'entrepôt de donnée.....	84
4) Le cube dimensionnel.....	84
5) Conclusion	88
Chapitre 5: Mise en œuvre de la solution	
1) Introduction	89
2) Outils de développements	89
3) Architecture technique de la solution	90
4) Implémentation de la solution	90
4.1) Le projet ETL	90
4.2) Projet création de Cube :.....	94
4.3) Mise en œuvre de la présentation des données	95
5) Conclusion	97
Conclusion générale perspective.....	99
Bibliographie.....	101

Résumer :

La SADEG (filiale de groupe SONELGAZ), par sa Région de Distribution Centre (RDC) lance chaque année des appels d'offre pour le choix des entreprises qui vont réaliser les programmes d'électrifications initié par l'état. La direction technique d'électricité à travers le département suivi et inspection assure la mise en œuvre de ces programmes du lancement jusqu'à l'achèvement sur les 14 concessions de distributions (CD) appartenant la région centre. Compte tenu de la grande quantité d'informations provenant de ces 14 concessions de distribution (sur les plan physique et financier), le risque d'erreurs est très important vu que la consolidation de ces informations se fait manuellement par le logiciel « Excel » ce qui peut induire en erreur les décideurs sur les actions à entreprendre et porter préjudice à la crédibilité du suivi et risque de retarder la clôture des programmes. Pour cela, les dirigeants du RDC nous ont confié la mise en place d'un système qui procure aux décideurs les informations nécessaires et fiables, les aidant ainsi à pendre dans les meilleurs délais les décisions les plus appropriées. Dans ce cadre, nous avons travaillé sur la mise en place d'une solution business intelligence en concevant et réalisant un système d'aide à la décision qui sera construit au tour d'une base de donnée dédié totalement au décisionnel un «Data Warehouse » et répondant aux besoins de RDC dans sa fonction d'avoir une vue d'ensemble sur la réalisation des programmes d'électrifications.

Mots clés : Business Intelligence « B.I », Décisionnel, Data Warehouse « Entrepôt de données », intégration de donnée, solutions « BI » Open Source.

الملخص
SADEG شركة تابعة لمجموعة (SONELGAZ)، من خلال منطقة التوزيع المركزية (RDC) تطلق سنويًا دعوات للمناقصات لاختيار الشركات التي ستنفذ برامج الكهرباء التي أطلقتها الدولة. يضمن التوجيه الفني للكهرباء من خلال إدارة المراقبة والتفتيش تنفيذ هذه البرامج من البداية وحتى الانتهاء على 14 امتياز توزيع (CD) تابعة للمنطقة الوسطى، مع الأخذ في الاعتبار الكمية الكبيرة من المعلومات الواردة من امتيازات التوزيع الـ 14 (المادية) والمصطلحات المالية)، فإن مخاطر الأخطاء عالية جدًا نظرًا لأن دمج هذه المعلومات يتم يدويًا بواسطة برنامج "Excel" يمكن أن يضل صناعات القرار بشأن ما يجب القيام به ويقوض مصداقية المراقبة ويخاطر بتأخير إغلاق البرنامج. للقيام بذلك كلفنا مسيري (RDC) بإنشاء نظام يزود صانعي القرار بالمعلومات الضرورية والموثوق بها، مما يساعدهم على اتخاذ القرارات الأكثر ملائمة في أسرع وقت ممكن. في هذا السياق، عملنا على تنفيذ حل ذكاء الأعمال من خلال تصميم وبناء نظام دعم القرار الذي سيتم بناؤه حول قاعدة بيانات مخصصة بالكامل لصنع القرار، "مستودع البيانات" والاستجابة لاحتياجات الشركة في وظيفتها المتمثلة في الحصول على نظرة عامة على تنفيذ برامج الكهرباء.

الكلمات المفتاحية: ذكاء الأعمال "BI"، مستودع البيانات، تكامل البيانات، حلول "BI" مفتوحة المصدر.

Abstract

SADEG (subsidiary of SONELGAZ group), through its Central Distribution Region (DRC) annually launches calls for tenders for the choice of companies that will carry out the electrification programs initiated by the state. The technical direction of electricity through the monitoring and inspection department ensures the implementation of these programs from launch to completion on the 14 concessions distribution (CD) belonging to the central region. Considering the large amount of information coming from these 14 distribution concessions (in physical and financial terms), the risk of errors is very high since the consolidation of this information is done manually by the "Excel" software. can mislead decision-makers about what to do and undermine the credibility of monitoring and risk delaying program closure. To do this, the leaders of the DRC have entrusted us with setting up a system that provides decision-makers with the necessary and reliable information, thus helping them to take the most appropriate decisions as soon as possible. In this context, we worked on the implementation of a business intelligence solution by designing and building a decision support system which will be built around a database totally dedicated to decision making, a "Data Warehouse" and responding to the needs of the DRC in its function of having an overview on the implementation of electrification programs.

Key words: Business Intelligence "BI", Data warehouse, data integration, Open Source "BI" solutions.

Abréviation

BI : Business Intelligence.

BT : Basse Tension.

CD : Concession de Distribution.

DD : Direction de Distribution.

DIM : Dimension.

DTE : Direction Technique Electricité

DSI : Département Suivi et Inspection

DET : Division Etude Travaux

DAM : Division Administration Marché

DDE : Direction d'Energie

DW : Data Warehouse (Entrepôt de données).

ER : Electrification Rurale

ETL : Extract, Transform and Load (ETC).

FK: Foreign Key.

HOLAP: Hybrid On Line Analytical Process.

MOLAP: Multidimensional On Line Analytical Process.

MT : Moyenne Tension.

ODS : Ordre de Service

OLAP : On Line Analytical Process.

OLTP: On Line Transactional Process.

PK : Primary Key.

PDG : Président Directeur General.

QLS : Quartiers et lotissements sociaux

RDC : Régional Direction Centre

ROLAP : Relational On Line Analytical Process.

SADEG : Société Algérienne de Distribution de l'Électricité et de Gaz

SI : Systèmes d'Information.

SID: Systèmes d'Information Décisionnels.

SGBD : Système de Gestion de Base de Données.

SONELGAZ : Société Nationale de l'Electricité et du GAZ.

SPA : Société Par Action.

SQL : Structured Query Language

Liste des figures

N°	Désignation	page
1	Le décisionnel au sein du Système d'information	22
2	Solution BI	23
3	Architecture standard pour la Business Intelligence	24
4	Structure de données d'un Data Warehouse	30
5	Place du Datamart dans le Data Warehouse	35
6	Architecture global d'un Data Warehouse	37
7	Considération d'un sujet d'analyse comme un cube à plusieurs dimensions	38
8	Exemple d'une table de fait	39
9	Exemple d'une table de dimension	39
10	Exemple d'un schéma étoile	40
11	Exemple d'un schéma flocon	41
12	Exemple d'un hyper-cube	42
13	Architecture MOLAP	43
14	Architecture ROLAP	44
15	Schéma organisationnel de groupe Sonelgaz	57
16	Schéma illustre différents région de distribution	58
17	Organisation des régions de distribution	59
18	Organisation des Directions de Distribution	60
19	Diagramme d'activité d'élaboration des rapports.	67
20	Place de l'étude des besoins dans le cycle de vie du projet DW/BI	69
21	Modèle en étoile de l'activité « Suivi des affaires »	79
22	Modèle en étoile de l'activité « suivi la clôture une affaire »	81

23	Modèle en étoile de l'activité « Suivi des entreprises	83
24	Cube dimensionnel « Suivi des affaires ».	85
25	Cube dimensionnel « Clôture des affaires ».	86
26	Cube dimensionnel « Suivi des entreprises ».	87
27	Architecture technique de la solution	90
28	Extraction des données à partir d'un fichier Excel	91
29	Vérification des colonnes	91
30	Transformation et conversion de données	92
31	Vérification le mappage des données	92
32	Chargement de donnée dans table affaire	93
33	Créer et charger la dimension Affaire à l'aide d'une dimension à variante lente	93
34	Création et chargement de la table des faits Factaffaire	94
35	Les éléments constitutifs de la base multidimensionnelle	95
36	Résultats d'importation de base de données	96
37	Tableau de bord réalisé par Power BI	96

Liste des Tableaux

N°	Désignation	page
1	Tableau comparatif entre les systèmes transactionnels et les systèmes décisionnels.	25
2	Comparaison entre les finalités des Data Warehouses et des Data Marts	36
3	Comparaison entre les différentes architectures d'OLAP	45
4	Synthèse des besoins détectés.	71
5	Description de la dimension « Temps	74
6	Description de la dimension « Entreprise ».	75
7	Description de la dimension « Facture ».	76
8	Description de la dimension « Situation ».	76
9	Description de la dimension « Direction distribution ».	77
10	Description de la dimension « Affaire ».	78
11	Description de la dimension « Garantie »	80
12	Description de la dimension « Catégorie ».	82

Introduction Générale

1) Contexte général

Les entreprises ont toujours cherché à mettre en application les meilleures pratiques de gestion afin d'être mieux outillées face à la concurrence. Ce climat de forte concurrence exige de ces entreprises une surveillance très étroite du marché afin de ne pas se laisser distancer par les concurrents et cela en répondant, le plus rapidement possible, aux attentes du marché, de leur clientèle et de leurs partenaires.

Pour se faire, les dirigeants de l'entreprise, quelque en soit d'ailleurs le domaine d'activité, doivent être en mesure de mener à bien les missions qui leur incombent en la matière. Ils devront prendre notamment les décisions les plus opportunes. Ces décisions, qui influenceront grandement sur l'avenir de l'entreprise, ne doivent pas être prises ni à la légère, ni de manière trop hâtive, vu leurs conséquences sur sa survie. Il s'agit de prendre des décisions fondées, basées sur des faits et informations claires, fiables et pertinentes.

L'identification des données importantes s'avère être une tâche difficile car les entreprises regorgent d'une part sous une masse considérable de données et que d'autre part les systèmes opérationnels s'avèrent limités, voire inaptes d'apporter des informations d'aide à la prise de décision.

L'apparition d'un ensemble de technique de modélisation ont apporté une solution à ces problèmes. Ces techniques sont souvent exprimées à travers les concepts de Business Intelligence (BI) de DataWarehouse et de Datamining. C'est dans ce contexte que les « systèmes décisionnels » ont vu le jour. Ils offrent aux décideurs des informations de qualité sur lesquelles ils peuvent se basé pour arrêter leurs choix décisionnels.

De part sa dimension économique et sa position sur le marché énergétique algérien, comme toute entreprise ayant une grande activité, SONELGAZ « Région de distribution centre (RDC) avec ces 14 (CD) » génère des données complexes et volumineuses qui représentent une source précieuse d'informations dont peuvent améliorer de façon significative le processus de prise de décision.

Le projet que nous abordant consiste à la mise au œuvre d'un système permettant une meilleur gestion des affaires de programmes d'électrifications en réalisation. Un tel système requiert la mise en place d'un entrepôt de données fiables contenant les informations nécessaires à l'accomplissement des processus décisionnels.

2) Problématique

Dans le cadre de la réalisation des programmes d'électrification initiés par les pouvoirs publics, à travers les conventions d'applications signées entre le Ministère de l'Energie et la Société SADEG (filiale du groupe Sonelgaz), la Région de Distribution Centre (RDC) lance chaque année des Appels d'Offres à travers ses quatorze(14) Concessions de Distribution (CD) pour le choix des entreprises qui vont réaliser ses programmes.

Ces programmes sont composés de deux opérations importantes :

- ✓ Le raccordement des centres ER (Electrification Rurale)
- ✓ Le raccordement des Quartiers et Lotissements Sociaux (QLS)

Pour le suivi de ces programmes, la Direction Technique électricité de la RDC, à travers le Département Suivi et Inspection, assure la mise en œuvre de ces programmes par :

- La définition des programmes à réaliser en coordination avec la SADEG par la transmission des listings des affaires à réaliser en coordination entre les Concessions de distribution et les Directions de l'Energie des Wilayas.
- La réalisation des études d'exécutions avec les Bureaux d'études au niveau des concessions de distribution.
- Lancement des Appels d'Offres pour les choix des entreprises de réalisation.
- Le Suivi de la réalisation des programmes sur les plans physique et financier à travers les reportings mensuels.
- L'établissement des différentes reportings et tableaux de bord pour le compte de la RDC et la SADEG.
- La clôture financière des programmes par l'établissement des Décomptes Généraux et définitifs et des avenants de clôture.

Vu le nombre important des Concessions de Distribution de la RDC (14 CD), et le nombre important d'informations parvenant de ces dernières (sur les plan physique et financier), le risque d'erreurs est très important vu que la consolidation de ces informations de fait manuellement par le logiciel « Excel ».

Ces anomalies peuvent induire en erreur les décideurs sur les actions à entreprendre et porter préjudice à la crédibilité du suivi et risque de retarder la clôture des programmes.

3) Objectifs du projet

Afin de palier aux problèmes précédemment cités, la RDC initié, à travers le DSI, le présent projet.

Ce projet a pour but d'introduire une informatique décisionnelle au sein du RDC, tout en proposant un système décisionnel qui assure le suivi d'ensemble des affaires des programmes d'électrifications à réalisé par les entreprise déjà retenue lors l'appel d'offre. Les principaux objectifs assignés au projet sont :

- Construire grâce aux techniques de la Business Intelligence un entrepôt de données pour RDC.
- Le suivie des réalisations physique et financières des programmes d'électrifications.
- Le suivie des opérations de clôture des programmes d'électrifications.
- Mettre en place à partir de l'entrepôt de données des tableaux de bord (tableaux, statistiques,graphiques, etc.) facilitant la prise de décision aux dirigeants du RDC/SADEG.
- Etablir mensuellement le Reporting à transmettre de RDC à la SADEG.
- Editer les différents états de suivi.

Les travaux effectués dans le cadre du présent mémoire est organisé en deux parties et se présente comme suit :

Après une introduction générale dans laquelle nous présentons le contexte général du projet, la problématique et les objectifs visés. La première partie est consacrée aux aspects théoriques du domaine des systèmes d'information décisionnels en évoquant des définitions et des concepts relatifs aux entrepôts de données et modélisation dimensionnel.

La deuxième partie traite l'aspect pratique de notre projet. Elle est divisée en cinq chapitres qui sont :

Chapitre un, présente l'organisme d'accueil, sa structure organisationnelle ainsi que ces fonctionnalités.

Chapitre deux, présente l'existant décisionnel au sein de l'entreprise d'accueil.

Chapitre trois, présente les techniques de collecte d'information utilisées, la démarche suivie pour identifier les besoins.

Chapitre quatre, consacré à la conception détaillé de notre solution. La conception de l'entrepôt de donnée, son alimentation et cube dimensionnel.

Chapitre cinq décrit les différents outils utilisés ainsi que la manière dont se passe la mise en œuvre de la solution.

Une conclusion générale est proposée afin de synthétiser le travail réalisé et abordé les perspectives.

Partie 1 :

Synthèse Bibliographique

I Introduction

Toutes les entreprises du monde disposent d'une masse de données plus ou moins Considérable proviennent soit de sources internes ou externes. L'excessivité de donnée, et l'incapacité des systèmes opérationnels à les exploiter à des fins d'analyse à conduit l'entreprise divergé vers une nouvelles informatique dite décisionnelle qui met l'accent sur la compréhension de l'environnement de l'entreprise et l'exploitation de ces données à bon escient.

II Les systèmes décisionnels

Afin de mieux comprendre qu'est ce qu'un système décisionnel, il est impérativement nécessaire de faire un rappel sur le système d'information.

1) Définition Système d'information :

«Le système d'information est l'ensemble des méthodes et moyens de recueil de contrôle de distribution des informations nécessaires à l'exercice de l'activité en tout point de l'organisation. Il a pour fonction de produire et de mémoriser les informations, de l'activité du système opérant (système opérationnel), puis de les mettre à disposition du système de décision (système de pilotage)». [4]

Le SI est le cerveau des entreprises où tous les acteurs véhiculent des informations. L'objectif principal d'un système d'information (SI) est de restituer l'information à la personne concernée sous une forme idoine et au moment opportun.

Les différences qui existent entre le système de pilotage et le système opérationnel, du point de vue fonctionnel ou des tâches à effectuer, conduit à l'apparition des « *systèmes d'information décisionnels* » (S.I.D.).

Les origines des SID remontent au début de l'informatique et des systèmes d'information qui ont, tous deux, connu une grande et complexe évolution liée notamment Cette évolution se poursuit à ce jour. [8]

La figure 1 illustre parfaitement l'importance du décisionnel au sein d'un système d'information.

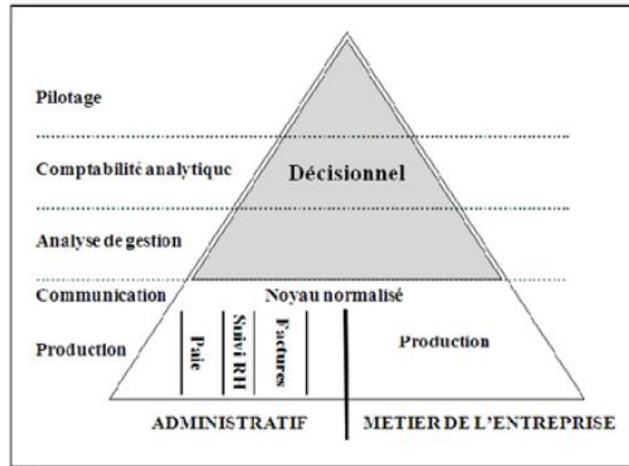


Figure 1 : Le décisionnel au sein du Système d'information. [6]

2) Définitions du décisionnel :

Parmi les différentes définitions du décisionnel « *business intelligence B.I.* » qui ont été données on trouve :

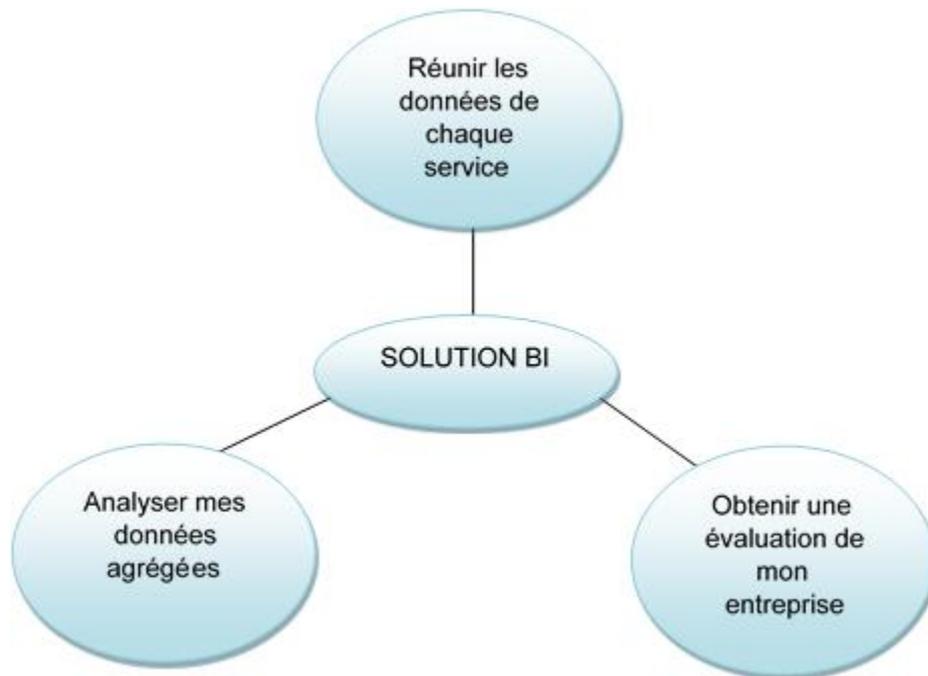
Def1 : « Le Décisionnel est le processus visant à transformer les données en informations et, par l'intermédiaire d'interrogations successives, transformer ces informations en connaissances ». [5]

Def2 : La business intelligence (informatique décisionnelle en français) est définie comme étant la capacité à disposer des bonnes informations, pour les bonnes personnes et au bon moment. Il s'agit de tout ce qui permet, à partir de données brutes des systèmes opérationnels, de créer l'intelligence qui sera la base de tous les processus de prise de décision au sein de l'entreprise. [2]

Les données et les informations sont pour l'entreprise et ses décideurs des éléments inertes. Simples enregistrements dans la base de données transactionnelle, les données sont filtrées, synthétisées, et agrégées ; elles sont contextualisées et ainsi désignées d' « informations ». [1]

L'intelligence élève l'information à un rang supérieur. Contrairement aux données et aux informations, l'intelligence est vivante. C'est la capacité collective de l'organisation à mettre en perspective les informations acquises, en les croisant avec les expériences passées et les stratégies futures. [1]

La figure2 illustre l'importance de la BI dans l'entreprise.



La figure2 : Solution BI. [7]

Les systèmes d'information décisionnels (SID) ont pour objectif d'aider à la prise de décision en permettant des analyses complexes (découverte des exigences et analyse des données). Les caractéristiques des fonctions d'analyse suivant plusieurs métiers ne peuvent pas toutes être représentées avec les modèles de SI classiques. Dans les milieux industriels et scientifiques, il est largement accepté que la modélisation de référence pour les SID est la modélisation multidimensionnelle. [9]

3) Composants des systèmes décisionnels

La définition de l'architecture du système décisionnel est une étape cruciale de tout projet de Business Intelligence. Une architecture solide permet un accès hautement performant et une analyse profonde des données.

Les systèmes décisionnel se compose généralement des données de l'organisation, d'un outil permettant leur extraction, leur transformation puis leur chargement dans la base de données d'aide à la décision, ou Data Warehouse, et d'outils de présentation et d'analyse de données. Ceci est illustré par la figure 3.

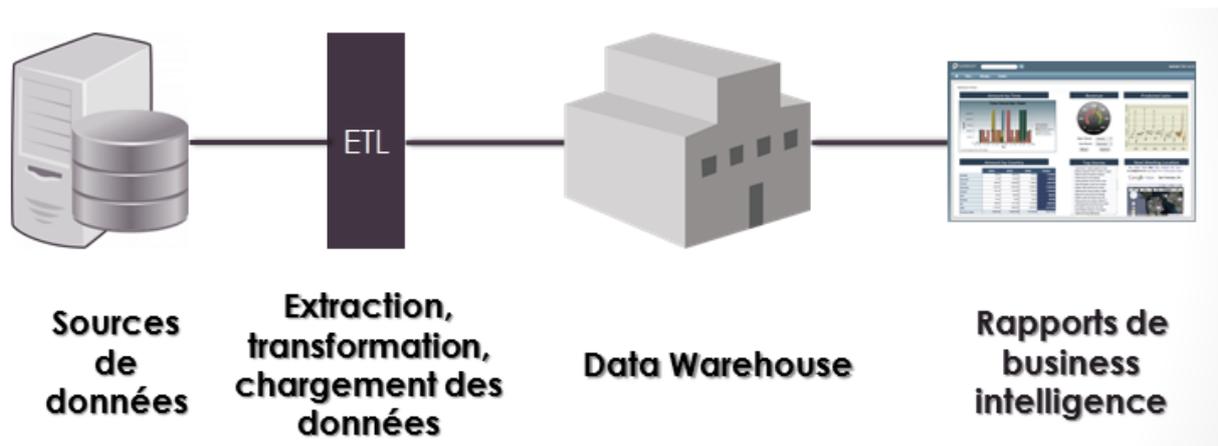


Figure 3 : Architecture standard pour la Business Intelligence. [10]

3.1) Les sources de données :

Ce sont majoritairement les données du système dit " *de production*" ; ces données sont internes à l'entreprise et sont stockées dans des bases de données de production ou externes à l'entreprise telles que les données fournies par les fournisseurs, les clients,...etc. De nos jours, les entreprises font face à certains défis concernant les informations provenant de ces sources de données : surabondance des données, incohérence, manque de fiabilité et inaccessibilité. [11]

3.2) Outil d'extraction, transformation et de chargement des données sources :

Avant d'être chargée dans le Data Warehouse, l'information doit être extraite, nettoyée et préparée. Cela se fait grâce à un outil spécifique, l'ETL.

L'ETL permet l'Extraction, la Transformation et le Chargement (Loading) des données. Cet outil permet de faire la transition entre les systèmes sources et le Data Warehouse en garantissant la qualité des données entreposées dans ce dernier. [12]

Cet outil sera détaillé en présentant les concepts de base du Data Warehouse.

3.3) L'entrepôt de données (Data Warehouse):

Considéré comme le cœur de la solution de Business Intelligence, l'entrepôt de données peut être défini comme étant une base de données d'aide à la décision. Il permet de stocker en un seul endroit toutes les informations provenant des différentes sources de données, utiles pour les dirigeants de l'organisation. [10]

3.4) Les outils de présentation et d'analyse de données:

Ces outils permettent aux utilisateurs d'accéder et d'analyser les données stockées dans le Data Warehouse afin d'obtenir les résultats pertinents pouvant supporter la prise de décision. [12]

4) Système d'information décisionnel versus système d'information

Le tableau suivant résume un ensemble de différences qu'il peut y avoir entre les systèmes transactionnels et les systèmes décisionnels selon les données et l'usage.

Différence	Système transactionnel	SID
Par les données	Orienté applications	Orienté thèmes et sujets
	Situation instantanée	Situation historique
	Donnée détaillées et codées non redondantes	Informations agrégées cohérentes souvent avec redondance
	Données changeantes constamment	Informations stables et synchronisées dans le temps
	Pas de référentiel commun	Un référentiel unique
L'usage	Assure l'activité au quotidien	Permet l'analyse et la prise de décision
	Pour les opérationnels	Pour les décideurs
	Mises à jour et requêtes simples	Lecture unique et requêtes complexes transparentes
	Temps de réponse immédiats	Temps de réponse moins critiques
	Faibles volumes à chaque transaction	Large volume manipulé
	Conçu pour la mise à jour	Conçue pour l'extraction
	Usage maîtrisé	Usage aléatoire

Tableau 1: Tableau comparatif entre les systèmes transactionnels et les systèmes décisionnels.

Ces différences font ressortir la nécessité de mettre en place un système répondant aux besoins décisionnels. Parmi les plus importantes composantes de ce système décisionnel est sa base de données d'aide à la décision, autrement dit le « *Data Warehouse* ».

Le Data Warehouse est véritablement le support de l'intelligence des entreprises puisqu'il est la source des données stratégiques qui font l'intelligence de l'organisation.

III Le Data Warehouse support d'aide à la décision

Le « Data Warehouse », « Entrepôt de données » en français, constitue une structure informatique et une fondation les plus incontournables pour la mise en place des systèmes décisionnels. Le concept de Data Warehouse, tel que connu aujourd'hui, est apparu pour la première fois en 1980. L'idée consistait alors à réaliser une base de données destinée exclusivement au processus décisionnel. Les nouveaux besoins de l'entreprise, les quantités importantes de données produites par les systèmes opérationnels et l'apparition des technologies aptes à sa mise en oeuvre ont contribué à l'apparition du concept «Data Warehouse» comme support aux systèmes décisionnels.

1) Concepts et Définitions

Ralph Kimball publie en 1996 son ouvrage « *Entrepôts de données: Guide pratique du concepteur de "Data Warehouse"* », dans lequel il définit l'entrepôt de données comme étant « *le conglomérat des données d'une organisation* ». [15]

Un entrepôt de données est une aire de rassemblement et de présentation où les données opérationnelles sont spécifiquement structurées pour garantir les performances des requêtes et faciliter l'analyse et l'utilisation.

« *Le Data Warehouse est un système qui extrait, nettoie, traite et rend conforme des données sources vers un espace de stockage multidimensionnel. Il permet ainsi la mise en oeuvre de l'interrogation et de l'analyse à des fins d'aide à la prise de décision* ». [14]

Ralph Kimball définit également les concepts de Data Warehouse d'entreprise et celui de Data Mart ; le Data Warehouse d'entreprise étant une agglomération des Data Mart.

« *Le Data Mart est un sous-ensemble, logique et physique de la présentation de l'entrepôt de données, organisé par processus métier ; c'est un ensemble flexible de données. Les Data Mart peuvent être attachés ensemble en utilisant des techniques de forage à travers leurs dimensions communes* ». [13]

Cette définition de l'entreposage de données n'est nullement unique, elle fut précédée par celle proposée en 1990 par William H Inmon dans son ouvrage « *Building the*

Data Warehouse » : « *Un entrepôt de données est une collection de données orientées sujet, intégrées, non volatiles, et historiées, servant de support à la prise de décision managériale. Les données de l'entrepôt de données sont aptes à être utilisées à de nombreuses fins, y compris des besoins futurs non encore connus*». [16]

Cette définition met en évidence quatre aspects du Data Warehouse :

- ❖ **Orienté sujet** : L'entrepôt de données est conçu autour d'une certaine activité, vente, assurances, services financiers...etc. Le Data Warehouse englobe tous les processus liés à l'activité pour laquelle il est conçu. Tout le domaine est couvert. Ceci est à contre-courant des systèmes transactionnels qui sont, quant à eux, orientés applications : commande, comptabilité...etc.
- ❖ **Données intégrées** : Le Data Warehouse va intégrer des données en provenance de différentes sources. Ces données doivent être cohérentes et consistantes.
- ❖ **Non volatilité** : Propriété d'une donnée qui, une fois enregistrée, n'est plus soumise à aucun changement. Une requête donnera toujours le même résultat peu importe le nombre de fois qu'elle est exécutée. Ceci constitue, encore une fois, un facteur de différenciation avec les bases de données relationnelles des systèmes opérationnels qui sont extrêmement volatiles [17]. Une requête ne produira jamais le même résultat si elle sollicite une table qui est fréquemment mise à jour.
- ❖ **Historisation** : Le Data Warehouse est vu comme une série temporelle où les données sont entreposées sous forme de couches successives. [14]
Cela signifie que les enregistrements dans l'entrepôt de données sont effectués de manière historiée, chaque élément de l'entrepôt est associé à la dimension temporelle. Cette propriété permet, entre autres, la prédiction par l'observation de la situation passée de l'organisation.

2) Objectifs du Data Warehouse

Pour définir et discuter des objectifs du Data Warehouse, il faut en premier lieu considérer le besoin des entreprises en informations stratégiques. Et cela, en considérant la crise d'information qu'elles connaissent et l'incapacité et l'inadéquation des technologies utilisées à procurer ses informations.

Avant tout, il serait judicieux de définir ce qu'est une information stratégique. « *L'information stratégique est le support sur lequel se basent les manager de l'entreprise,*

chargés de veiller au maintien de sa compétitivité, pour formuler les stratégies et actions à entreprendre, établir les objectifs de l'organisation et suivre de près les résultats. L'information stratégique doit être intégrée, correspondre aux contraintes de gestion établies au sein de l'organisation, accessible, crédible et historisée». [18]

Les objectifs du Data Warehouse peuvent être définis simplement en écoutant les discussions des dirigeants de l'organisation ; celles-ci tournent souvent autour de l'inaccessibilité des données, de leur incohérence, de la quasi-impossibilité de découper les données en tranches de manière simple et d'en tirer uniquement l'information essentielle. [15]

Le Data Warehouse, en intégrant, organisant et présentant les données de l'organisation permet aux décideurs d'avoir une vue globale de l'organisation et de leur métier. [14]

Le Data Warehouse offre les opportunités suivantes :

- **Accès aux informations de l'organisation**

La solution Data Warehouse doit être un outil facile à utiliser et permettre à n'importe quel utilisateur de se connecter aisément au Data Warehouse à partir de n'importe quel poste au sein de l'entreprise ou à partir d'un ordinateur personnel.

L'accès à l'information doit être direct et ne pas nécessiter l'intervention d'une tierce personne, en plus d'être immédiat, à la demande et de haute performance ; les requêtes les plus minuscules doivent s'exécuter en moins d'une seconde. [15]

- **Cohérence des données**

L'information ne doit exister au sein de l'organisation que sous une seule forme. Il sera, de ce fait, possible de disposer de Méta données¹. L'uniformité, ou intégration des données, assure qu'une requête lancée à un certain intervalle de temps concernant les mêmes paramètres donnera toujours le même résultat. Enfin, il sera également possible de savoir si toutes les données ont été chargées, et dans le cas contraire, quand cela sera-t-il fait. [15]

- **Analyse de données**

Effectuer des découpes sur les données, tranches et en dés (en anglais slice and dice) et ainsi pouvoir séparer et combiner les données de façon à obtenir les résultats voulus. [15]

¹ Définitions des éléments de données, permettant de savoir précisément ce qu'une requête permet de renvoyer comme information.

Le découpage en tranches et en dés est une description standard de la possibilité d'accéder à un entrepôt de données, avec la même facilité, par n'importe laquelle de ses dimensions. [19]

- **Publication de données**

Après avoir été rassemblées à partir de différentes sources situées dans différentes fonctions de l'organisation, les données sont nettoyées, leur qualité est vérifiée ; elles seront diffusées si elles sont jugées bonnes à être utilisées. [15]

La publication des données doit néanmoins se faire de manière sécurisée ; les informations d'une organisation constituent sa richesse. L'entrepôt de données doit permettre de contrôler l'accès aux informations confidentielles de l'entreprise. [14]

- **Garantir la qualité des informations produites**

Plus qu'un espace de stockage de données, le Data Warehouse fournit un nombre d'outils d'analyse et de préparation des données.

En outre, il est requis d'un Data Warehouse qu'il soit souple et adaptable. Les besoins des utilisateurs, les conditions commerciales, les données et la technologie étant en perpétuel changement, l'entrepôt de données doit être conçu pour gérer ce changement sans altérer ou invalider les données déjà présentes dans l'entrepôt et les applications implémentées. [14]

Malgré ces innombrables avantages fournis par l'entreposage de données, il existe néanmoins certains soucis et complexité liés à sa mise en place.

Premièrement, le fait que les données proviennent de systèmes opérationnels, nombreux et hétérogènes. Ces données doivent impérativement être intégrées pour être stockées dans le Data Warehouse ; ceci nécessite de rechercher, comprendre et analyser les données de toutes les applications opérationnelles existantes au sein de l'organisation pour mieux appréhender la signification de chaque donnée de l'organisation et donc sa véritable utilité. [20]

La réticence des utilisateurs face au changement d'environnement apporté par la mise en place du Data Warehouse peut également constituer un autre facteur problématique à l'aboutissement du projet.

3) Structure des données d'un Data Warehouse

Le Data Warehouse a une structure bien définie, selon différents niveaux d'agrégation et de détail des données. Cette structure est définie par Inmon [16] comme suit :

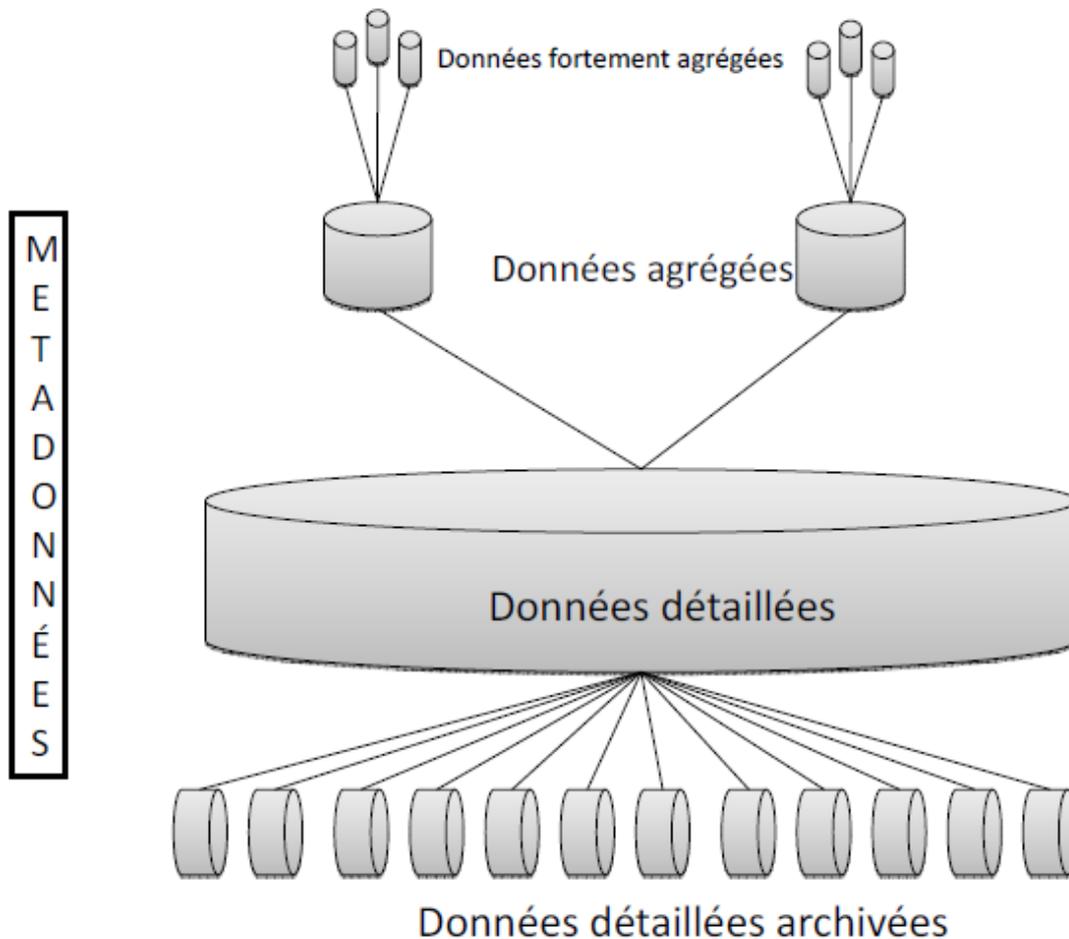


Figure 4: Structure de données d'un Data Warehouse. [16]

3.1) Données détaillées :

Les données détaillées reflètent les événements les plus récents. La gestion et le traitement de ces données nécessitent parfois des machines sophistiquées, car leur volume peut être très important.

3.2) Données détaillées archivées :

Anciennes données rarement sollicitées, généralement stockées dans un disque de stockage de masse, peu coûteux, à un même niveau de détail que les données détaillées.

3.3) **Données agrégées :**

Données agrégées à partir des données détaillées. Ces données sont très souvent utilisées, elles constituent déjà un résultat d'analyse et de synthèse de l'information contenue dans le système décisionnel.

3.4) **Données fortement agrégées :**

Données agrégées à partir des données détaillées, à un niveau d'agrégation plus élevé que les données agrégées.

3.5) **Meta données :**

Ce sont les informations relatives à la structure des données, les méthodes d'agrégation et le lien entre les données opérationnelles et celles du Data Warehouse. Les métadonnées doivent renseigner sur :

- ✓ Le modèle de données,
- ✓ La structure des données telle qu'elle est vue par les développeurs,
- ✓ La structure des données telle qu'elle est vue par les utilisateurs,
- ✓ Les sources des données,
- ✓ Les transformations nécessaires,
- ✓ Suivi des alimentations,

4) **Composants du Data Warehouse**

Les différentes approches de mise en place de Data Warehouse sont sujettes à de nombreux débats quant à l'efficacité et l'optimalité de l'une par rapport à l'autre. Le débat est néanmoins plus ou moins clos, et l'accent est davantage mis sur l'efficacité de l'approche de Ralph Kimball consistant à mettre en place des Data Mart départementaux dont la somme représente le Data Warehouse d'entreprise. La mise en place du Data Mart départemental étant plus rapide, simple et engendrant un risque relativement faible en comparaison avec l'approche visant à mettre en oeuvre le Data Warehouse d'entreprise comme étant une entité indivisible. [21]

La mise en place d'un Data Warehouse au sein d'une organisation nécessite la maîtrise totale des données de cette dernière et la bonne utilisation des techniques informatiques.

Un entrepôt de données est constitué de deux types de composants : *les outils de premier plan* et les *outils d'arrière-plan* (d'arrière guichet).

Les outils de premier plan sont les outils client qui récupèrent ou manipulent les données rangées dans le Data Warehouse. Il s'agit des requêtes d'analyse et de présentation

des données, des différents rapports générés et des tableaux de bord élaborés. Ces composants représentent **40%** des composants du Data Warehouse. [15]

Les composants d'arrière-guichet, représentent les données en elles-mêmes, les systèmes source desquels les données sont extraites et le matériel central, c'est-à-dire tous les logiciels résident à la fois sur le client et sur le serveur pour aider au processus d'extraction et de préparation des données. Les composants d'arrière-guichet représentent **60%** de la taille du Data Warehouse. [15]

On retrouve dans tout Data Warehouse les éléments suivants :

4.1) Les systèmes sources

Appelés communément applications de gestion ou systèmes opérationnels d'enregistrement. [13]

Leur fonction consiste à capturer les transactions quotidiennes liées à l'activité. Ils sont souvent peu homogénéisés les uns par rapport aux autres au sein d'une même organisation, n'attachent aucune importance à l'historisation des données et le reporting y constitue un fardeau.

Les sources de données peuvent être regroupées selon quatre catégories [18] :

a) Données de production

Ces données proviennent des systèmes opérationnels de l'entreprise. Ces systèmes sont nombreux, résident sur des plateformes différentes ; il s'agit du portefeuille des applications « métier » de l'entreprise.

b) Données internes

Représentent les divers documents que les employés d'une organisation gardent en leur possession. Feuilles de calcul, profils clients et bases de données départementales. Ces données peuvent être extraites des systèmes de production mais une grande partie d'entre-elles est contenue dans des fichiers privés au sein de l'organisation.

c) Données archivées

Les systèmes opérationnels étant destinés à la gestion actuelle de l'organisation ; ils sont régulièrement purgés des données les plus anciennement stockées vers des fichiers d'archivage.

d) Données externes

Parts de marché des concurrents, valeurs standards des indicateurs financiers et toutes autres statistiques relatives au secteur d'activité de l'entreprise, les données externes

constituent des informations stratégiques dans le processus de prise de décision. Ils permettent de comparer les performances et d'établir les tendances futures sur la base du présent et du passé du marché. Ces données doivent être reconçues car non conformes aux formats de l'entreprise.

4.2) La zone de préparation de données

« Ensemble des processus qui nettoient, transforment, combinent, archivent, suppriment les doublons, c'est-à-dire préparent les données sources en vue de leur intégration puis de leur exploitation au sein du Data Warehouse ». [13]

C'est à ce niveau que se fait le traitement des données brutes chargées à partir du système source. Elle est constituée d'un ensemble de processus appelé ETL, « Extract, transform and Load ». Le traitement de la qualité des données implique plusieurs étapes distinctes : vérification des valeurs valides, cohérences entre les valeurs, vérification des doublons... L'objectif final de la zone de préparation des données est l'obtention de données prêtes à être chargées sur un serveur de prestation (moteur OLAP ou SGBDR).

❖ Qualité des données

La qualité des données est un véritable souci pour toute organisation. La mauvaise qualité des informations stockées peut mettre en péril tout le projet Data Warehouse puisque son utilité est complètement remise en question si les données qui y sont stockées sont de faible qualité, induisant l'inexactitude des informations qu'il renvoie.

L'une des étapes les plus importantes du chargement des données dans le Data Warehouse est le contrôle de qualité. Ceci commence par la validation des données ; les données doivent être complètes, mise à jour et cohérentes. [22]

Après avoir été validées, les données doivent être intégrées au sein de l'entrepôt. La couche d'intégration assure la justesse des formats de données ; des données du même type doivent apparaître sous le même format. Par exemple des montants avec une certaine devise, des mesures avec une certaine unité. [22]

Une intégration de données réussie est aussi souple que possible, et adaptable au grès des changements des systèmes sources et des besoins de l'organisation en information. Pour cela, une grande importance devra être donnée aux Métadonnées.

La vérification de la qualité des données peut être mise en oeuvre sur la zone de préparation des données ou directement sur les systèmes sources eux-mêmes. [22]

4.3) **La zone de présentation de données**

« Machine cible sur laquelle l'entrepôt de données est stocké et organisé pour répondre en accès direct aux requêtes émises par des utilisateurs, les générateurs d'états et les autres applications. » [13]

C'est tout ce que l'utilisateur peut voir et toucher par le biais des outils d'accès. C'est l'entrepôt où les données sont organisées et stockées.

L'entrepôt de données est constitué d'un ensemble de Datamart. Le Datamart est une miniaturisation du Data Warehouse, construit autour d'un sujet d'analyse précis ou consacré à un niveau départemental.

❖ **Datamart :**

« Datamart est défini comme un sous-ensemble logique d'un entrepôt de données ». [14]

Le Datamart est une base de données plus réduite que le Data Warehouse, puisque destinée à quelques utilisateurs d'un département. Il séduit plus que le Data Warehouse les candidats au décisionnel.

C'est une petite structure très ciblée et pilotée par les besoins utilisateurs. Il a la même vocation que le Data Warehouse (fournir une architecture décisionnelle), mais ciblant un groupe de métier ou sujet spécifique.

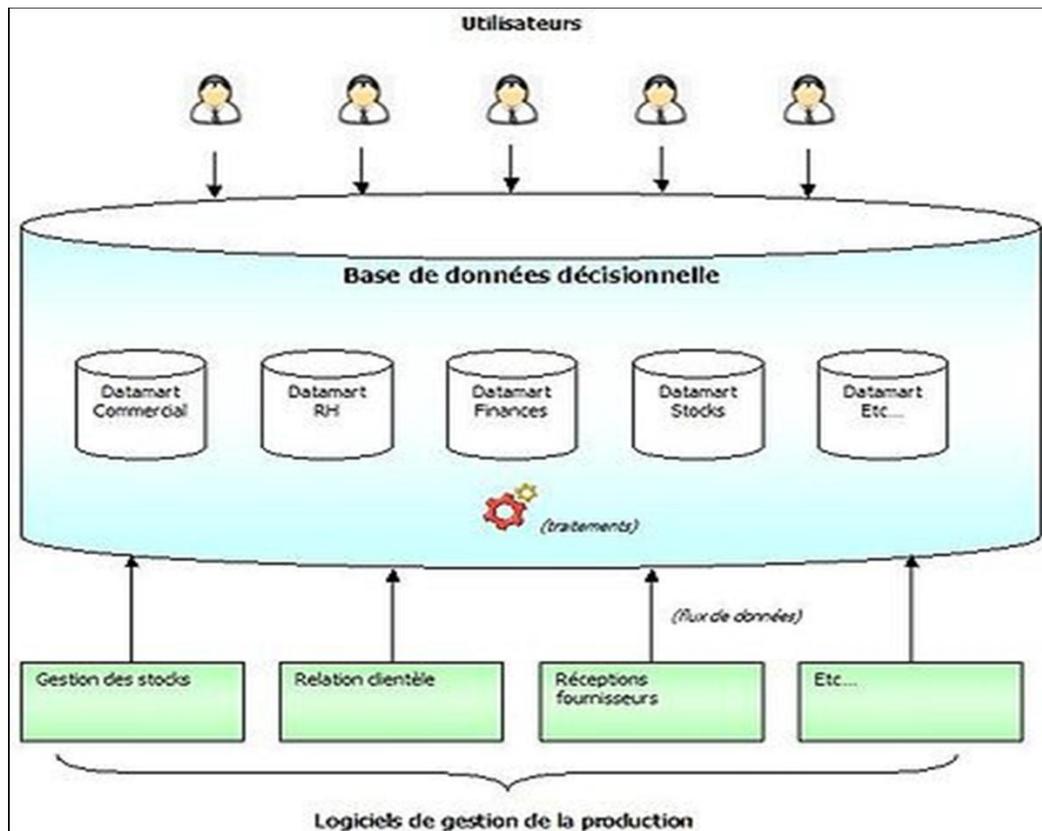


Figure 5: Place du Datamart dans le Data Warehouse².

² <https://fr.wikipedia.org/wiki/Datamart>

❖ Comparaison entre les finalités des Data Warehouse et des Data Mart :

	Data Warehouse	Data Mart
Cible utilisateur	Toute l'entreprise	Département
Implication du service informatique	Elevée	Faible ou moyenne
Modèles de données	A l'échelle de l'entreprise	Département
Champ applicatif	Multi sujets, neutre	Quelques sujets, spécifique
Sources de données	Multiples	Quelques-unes
Stockage	Base de données	Plusieurs bases distribuées
Taille	Centaine de GO et plus	Une à 2 dizaines de GO

Tableau 2 : Comparaison entre les finalités des Data Warehouses et des Data Marts. [23]

4.6) Le portail de restitution de données

Ce sont les outils utilisés par les utilisateurs finaux pour interroger le Data Warehouse et analyser les données des serveurs de présentation des données. [13] Les utilisateurs maintiennent des sessions ouvertes auprès du serveur de présentation en envoyant des flux de requêtes SQL, les données y sont présentées sur un écran sous forme d'états, de graphiques, de tableaux de bord...etc.

Le portail de restitution de données comprend des outils tels que les outils d'interrogation Ad hoc qui permettent aux utilisateurs de formuler leur requête d'interrogation, les outils d'analyse multidimensionnels, les outils de Data Mining et diverses applications de modélisation. [18]

5) Architecture d'un Data Warehouse

Après avoir exposé et défini chacun des éléments constituant l'environnement d'un Data Warehouse, il serait intéressant de connaître le positionnement de ces éléments dans une architecture globale d'un Data Warehouse .

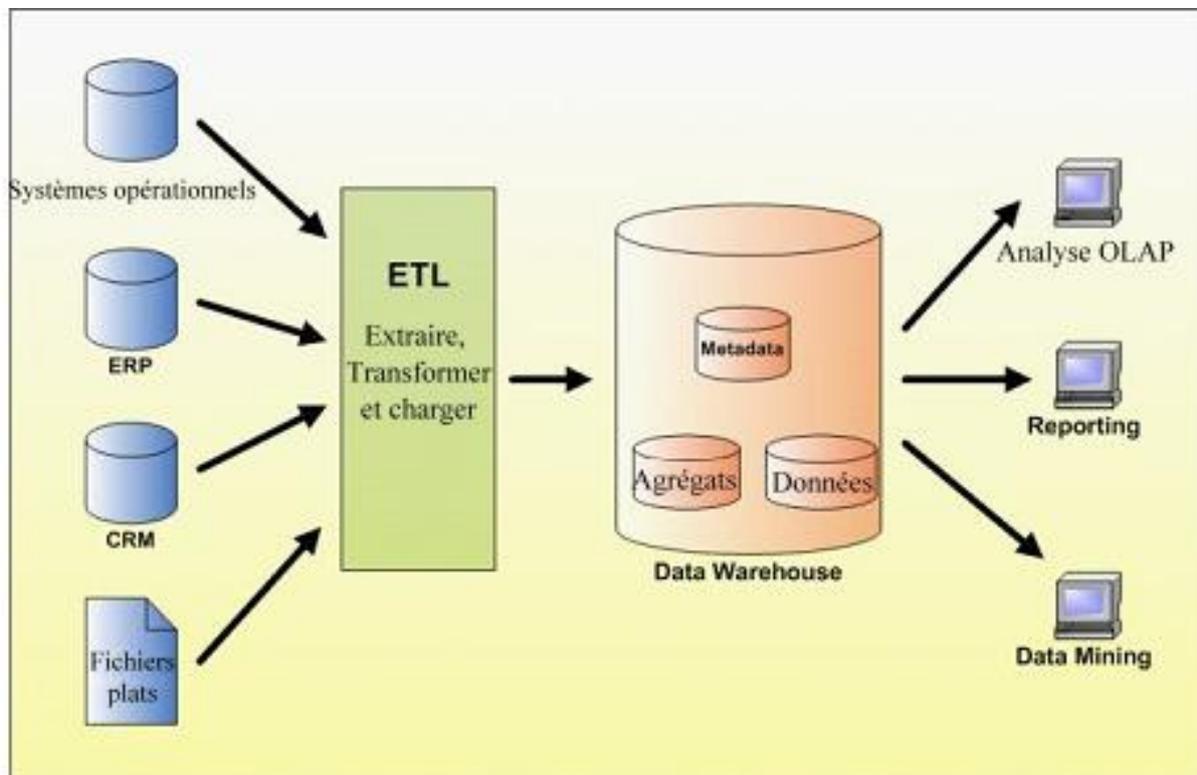


Figure 6 : Architecture globale d'un Data Warehouse³.

³ <https://www.semanticscholar.org/paper/A-STUDY-ON-DATA-WAREHOUSE-ARCHITECTURE-Arunachalam-Srigowthem>

IV Modélisation des données de Data warehouse

1) Modélisation dimensionnelle et ses concepts

Il ne faut pas perdre de vue que le Data Warehouse n'est qu'une étape dans la mise en place d'un système décisionnel. D'où le besoin d'une structure simple qui permet de représenter les données de manière visuelle pour les utilisateurs finaux. Cette structure est le modèle dimensionnel. Il consiste à représenter les données en tant que points dans un espace multidimensionnel (cube). Les données sont vues comme des sujets d'analyse (faits) étudiés selon plusieurs axes (dimensions).

La modélisation dimensionnelle permet non seulement de présenter les données sous une forme standardisée intuitive mais permet aussi des accès hautement performants. [24]

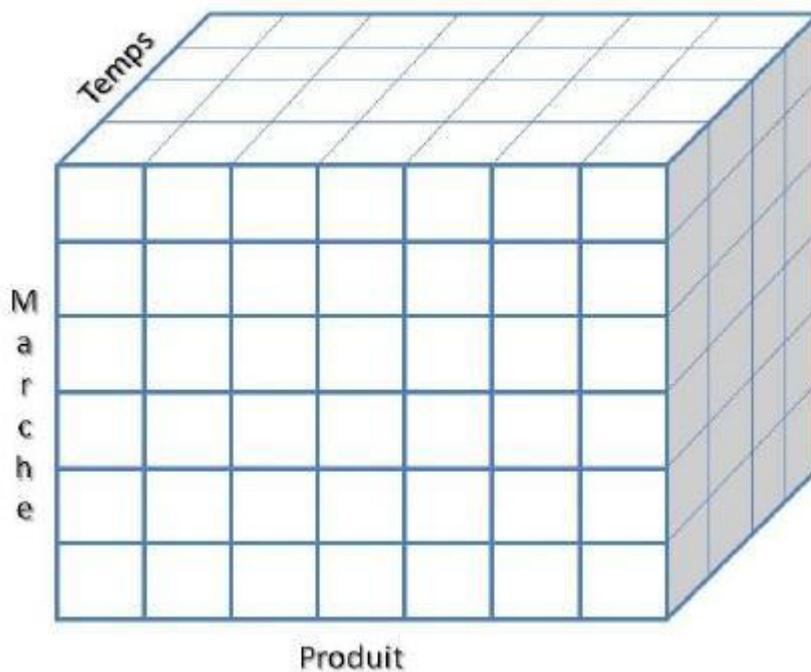


Figure 7: Considération d'un sujet d'analyse comme un cube à plusieurs dimensions. [17]

2) Tables de faits et tables de dimensions :

Le modèle multidimensionnel se compose de faits contenant les mesures à analyser et de dimensions contenant les paramètres de l'analyse.

2.1) Table de fait

La table de fait sert à stocker les mesures de l'activité. Chacune de ces mesures est prise à l'intersection de toutes les dimensions. Une ligne d'une table fait correspond à une mesure. Ces mesures sont généralement des valeurs numériques, additives.

Une table de faits assure les liens plusieurs à plusieurs entre les dimensions. Elles comportent des clés étrangères, qui ne sont autres que les clés primaires des tables de dimension.

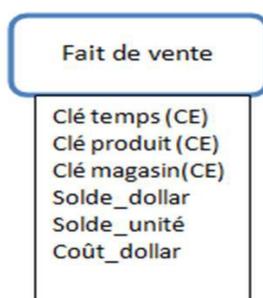


Figure 8 : Exemple d'une table de fait. [17]

2.2) Table de dimension :

Les tables de dimensions accompagnent une table fait, servent à enregistrer les descriptions textuelles des dimensions de l'activité. Elle est composée de plusieurs colonnes qui décrivent une ligne. C'est la table de faits qui rend le Data Warehouse compréhensible et utilisable.

« Une table de dimension établit l'interface homme / entrepôt, elle comporte une clé primaire » [14].



Figure 9 : Exemple d'une table de dimension. [17]

3) Différents modèles de la modélisation dimensionnelle :

3.1) Modèle en étoile

Appelé modèle en étoile car le diagramme qui le représente ressemble à une étoile. Il y a une grande table (table fait), dominant les autres, au centre du schéma. C'est la seule table du schéma comportant de multiples jointures la reliant aux autres tables (tables de dimensions).

Les autres tables n'ont qu'une seule jointure les reliant à la table de faits. C'est le modèle le plus utilisé pour sa lisibilité et performance.

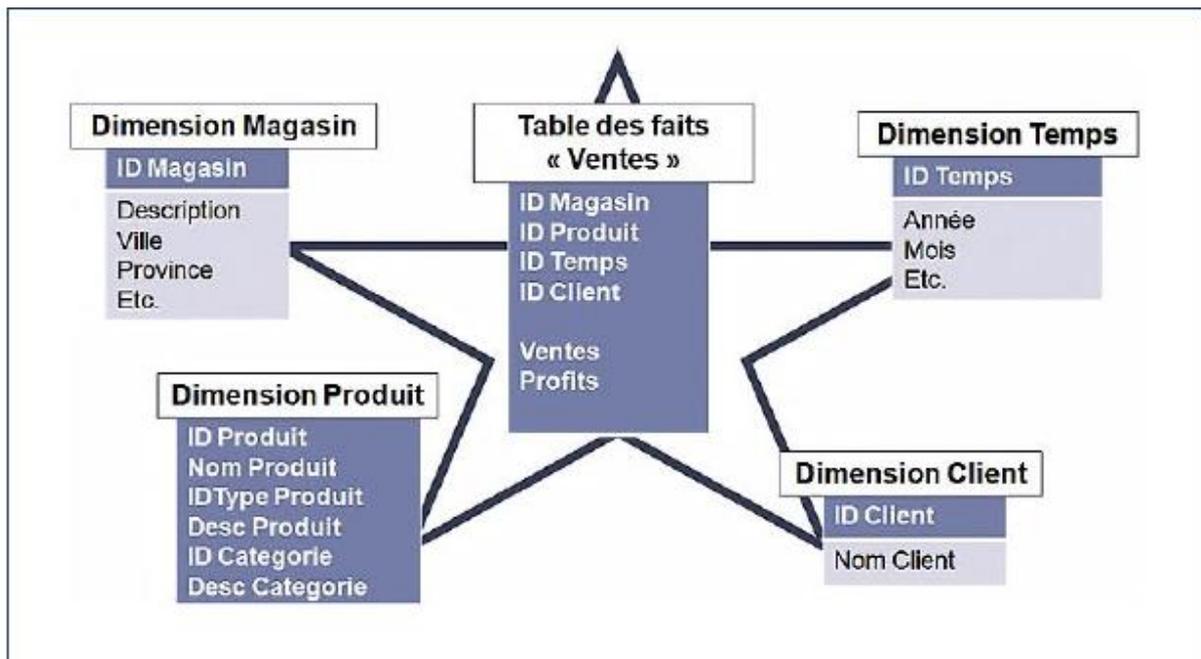


Figure 10: Exemple d'un schéma étoile⁴.

⁴ http://www-igm.univ-mlv.fr/~dr/XPOSE2009/informatique_decisionnelle_olap/md.html

3.2) Modèle en flocon

Identique au modèle en étoile la seule différence est que chaque attribut de niveau hiérarchique est mis dans une table dimension à part. Cette modélisation est généralement justifiée par l'économie d'espace de stockage, cependant elle peut s'avérer moins compréhensible pour l'utilisateur final, et très coûteuse en terme de performances. Prenons comme exemple la dimension temps :

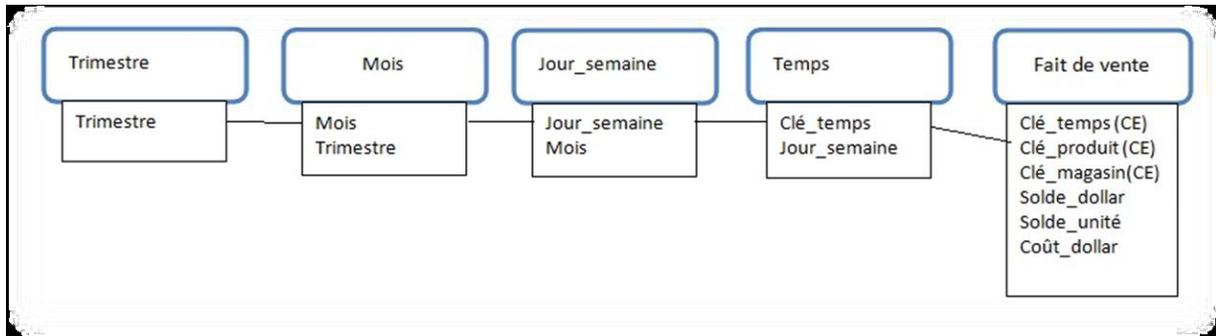


Figure 11: Exemple d'un schéma flocon. [17]

3.3) Modèle en constellation

Il est encore basé sur le modèle en étoile. Mais en rassemblant plusieurs tables de fait qui utilisent les mêmes dimensions.

3.4) Modèle en grappe ou Modèle mixte

Il fusionne des modèles en étoile et en flocon et consiste en une normalisation des grandes tables lorsqu'il y a trop de redondance.

4) Le cube de données :

La base de données multidimensionnelles s'appuie sur un hyper cube (cube à n dimensions). L'administrateur définit les dimensions, qui représentent une façon de trier l'information.

Cette modélisation permet à l'utilisateur sans l'aide de l'informaticien d'avoir l'information dont il a besoin. L'utilisateur choisit deux ou trois critères à visualiser sous forme de tableau ou de cube. Il peut également faire pivoter les axes d'analyse pour projeter les informations sous un angle différent. Ainsi, après avoir examiné les ventes par région, il peut permuter les axes pour une visualisation par mois. Ce critère est certainement le critère-clé du concept OLAP car il reflète la dimensionnalité de l'entreprise telle que la perçoivent ses membres qui ne sont autres que les utilisateurs du système.

Exemple : Sur le schéma ci-dessous on distingue les dimensions : produit, temps et région. Les faits sont la vente. Pour chaque combinaison des dimensions on peut accéder à la valeur numérique associée au fait vente.

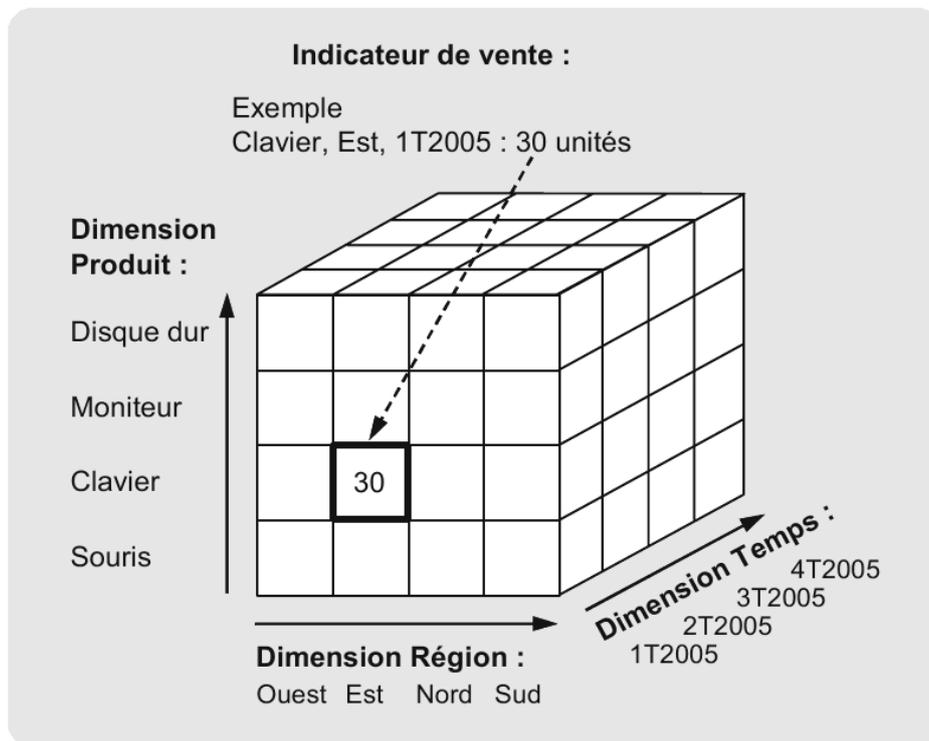


Figure 12: Exemple d'un hyper-cube⁵.

5) Le concept OLAP :

La différence d'objectifs entre les SGBD et les Data Warehouses a suscité l'utilisation de deux systèmes différents :

- OLTP (On Line Transaction Processing) pour les SGBD.
- OLAP (On Line Analytical Processing) pour les Data Warehouses

⁵ <https://business-intelligence.developpez.com/tutoriels/quest-ce-que-la-bi>

5.1) Généralités

En 1993, Edgar Frank Codd⁶ introduit le terme *On-Line Analytical processing* (OLAP) qui désigne : « *Activité globale de requêtage et de présentation de données textuelles et numériques contenues dans l'entrepôt de données ; style d'interrogation et de présentation spécifiquement dimensionnel* » [13]

Donc les Data Warehouses reposent sur le système OLAP, il est caractérisé par un volume très important de données. Il fonctionne seulement en lecture et a pour objectifs principaux de rassembler, gérer, traiter et de présenter des données multidimensionnelles et venant de sources diverses à des fins d'analyse et de décision, offrant ainsi à l'utilisateur une vue orientée métier.

5.2) Architecture des serveurs OLAP :

Le serveur OLAP est le noyau d'un système OLAP. C'est le lieu où les données sont stockées. L'architecture X-OLAP définit la façon dont les données sont stockées physiquement pour permettre des analyses multidimensionnelles. ces architectures peuvent être distinguées comme suit :

a) Architecture MOLAP « Multidimensional OLAP » :

MOLAP est la forme la plus classique d'OLAP. Elle s'appuie sur une base de données multidimensionnelle. Le stockage MOLAP est le plus approprié pour les cubes qui sont utilisés fréquemment et nécessitent un temps de réponse minimal. Son inconvénient est qu'il ne supporte pas de très grands volumes de données.

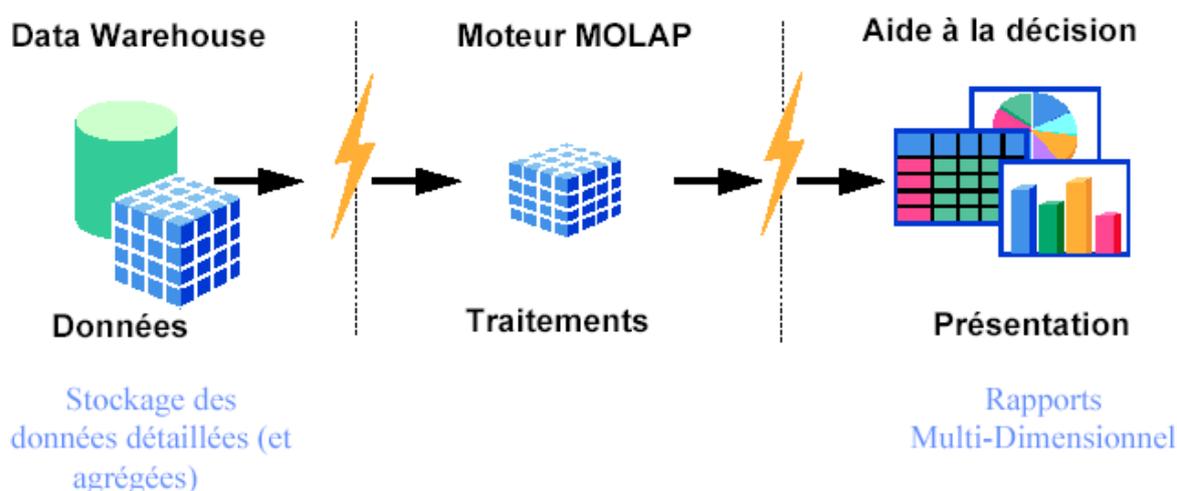


Figure 13 : Architecture MOLAP. [23]

⁶ Informaticien Britannique. Il est considéré comme Le père des bases de données relationnelles.

b) Architecture ROLAP « Relational OLAP »:

« Ensemble d'interfaces utilisateurs et d'applications qui donnent une vision dimensionnelle à des bases de données relationnelles ». [13]

Les données du cube et les agrégations sont stockées dans l'entrepôt existant. (Bases de données relationnelles). Les systèmes ROLAP sont en mesure de simuler le comportement d'un SGBD multidimensionnel en exploitant un SGBD relationnel. L'utilisateur aura ainsi l'impression d'interroger un cube multidimensionnel alors qu'en réalité il ne fait qu'adresser des requêtes sur une base de données relationnelles. En général il est utilisé pour les grands jeux de données qui ne sont pas sollicités fréquemment. Il permet aussi de stocker de gros volumes de données.

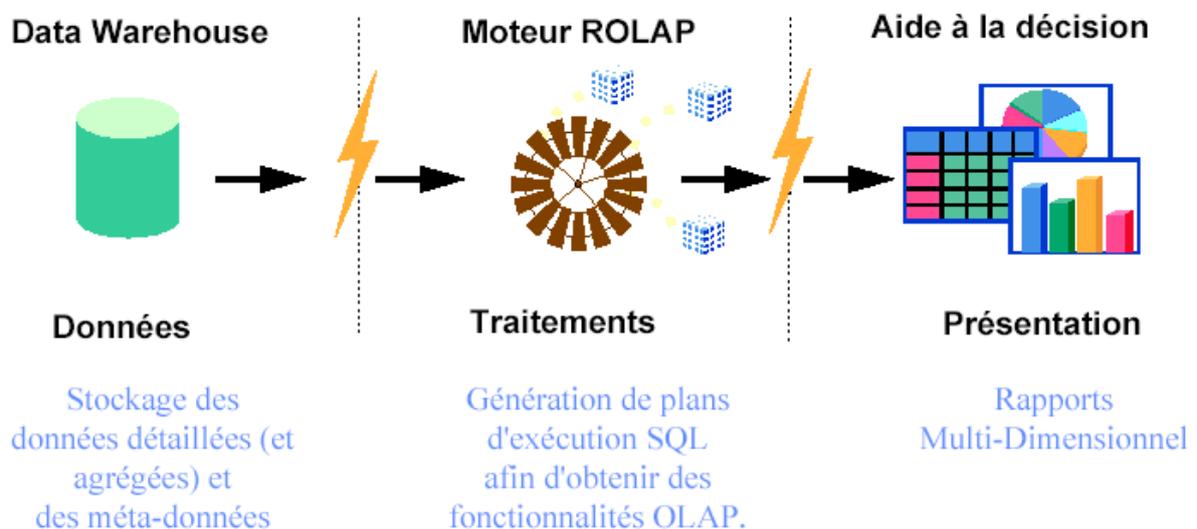


Figure 14 : Architecture ROLAP. [23]

c) Architecture HOLAP « Hybrid OLAP »:

Cette architecture est une sorte de compromis entre les deux architectures précédentes, en effet les données du cube sont stockées dans l'entrepôt relationnel existant, et les agrégations dans une structure multidimensionnelle. Cette combinaison donne à ce type de système les avantages du ROLAP et du MOLAP en utilisant tour à tour l'un ou l'autre selon le type de données.

d) Autres architectures OLAP :

Il existent d'autres architectures OLAP moins utilisées comme l'architecture OOLAP « Object OLAP », ou alors DOLAP « Desktop OLAP » qui décrivent une catégorie de produits qui ne sont pas nécessairement connectés à un serveur et qui s'appuient sur une source de données (un cube) construites, stockées et exploitées en local sur la machine de l'utilisateur.

e) Comparaison MOLAP Vs. ROLAP Vs. HOLAP :

	MOLAP	ROLAP	HOLAP
Stockage de données de base (détaillées)	BD multidimensionnelle	BD relationnelle	BD relationnelle
Stockage des agrégations	BD multidimensionnelle	BD relationnelle	BD multidimensionnelle
Performance requêtes	Elevée	Peu performante	Moyenne

Tableau 3: Comparaison entre les différentes architectures d'OLAP. [25]

V Modélisation et conception du Data Warehouse

Pour la construction du Data Warehouse on suivra les démarches suivantes :

- Méthodologies conductrices à la mise en place d'un Data Warehouse.
- Alimentation du Data Warehouse.
- Déploiement et mise en oeuvre.
- Maintenance et croissance.

1) Les méthodologies conductrices à la mise en place d'un Data Warehouse :

1.1) Top-Down :

C'est la méthode la plus lourde, la plus contraignante et la plus complète en même temps. Elle consiste à la conception de tout l'entrepôt (i.e. : toutes les étoiles), puis à la réalisation de ce dernier.

Imaginez le travail qu'une telle méthode implique : savoir à l'avance toutes les dimensions et tous les faits de l'entreprise, puis réaliser tout ça... Le seul avantage que cette méthode comporte est qu'elle offre une vision très claire et très conceptuelle des données de l'entreprise ainsi que du travail à faire.

1.2) Bottom-Up :

C'est l'approche inverse, elle consiste à créer les étoiles une par une, puis à les regrouper par des niveaux intermédiaires jusqu'à obtention d'un véritable entrepôt pyramidal avec une vision d'entreprise.

L'avantage de cette méthode est qu'elle est simple à réaliser (une étoile à la fois), l'inconvénient est le volume de travail d'intégration pour obtenir un entrepôt de données ainsi que la possibilité de redondances entre les étoiles (car elles sont faites indépendamment les unes des autres).

1.3) Middle-Out :

C'est l'approche hybride, et conseillée par les professionnels du BI. Elle consiste en la conception totale de l'entrepôt de données (i.e. : concevoir toutes les dimensions, tous les faits, toutes les relations), puis créer des divisions plus petites et plus gérables et les mettre en œuvre. Cela équivaut à découper notre conception par éléments en commun et réaliser les découpages un par un.

Cette méthode tire le meilleur des deux précédentes sans avoir les contraintes. Il faut juste noter que cette méthode implique, parfois, des compromis de découpage (dupliquer des dimensions identiques pour des besoins pratiques).⁷

2) Alimentation du Data Warehouse

L'outil d'alimentation permet de paramétrer des règles de gestion, propres à l'entreprise et à son secteur d'activité. Ces règles visent elles aussi à assurer la cohérence entre les données et à ne stocker dans l'entrepôt de données que des informations préalablement mises en relation les unes avec les autres.

Les ETL sont le plus souvent utilisés dans l'informatique décisionnelle afin de permettre l'alimentation des bases de données décisionnelles.

2.1) Définition d'ETL

« **Extract-Transform-Load** » est connu sous l'acronyme ETL (ou parfois : Datapumping). Il s'agit d'une technologie informatique middleware permettant d'effectuer des synchronisations massives d'information d'une banque de données vers une autre.

2.2) Phases constitutives d'un ETL

Les processus d'ETL regroupent trois étapes qui sont:

❖ L'extraction des données

« L'extraction est la première étape du processus d'apport de données à l'entrepôt de données. Extraire, cela veut dire lire et interpréter les données sources et les copier dans la zone de préparation en vue de manipulations ultérieures. ». [13]

Dans cette étape, il faut extraire les données à partir de différentes sources internes et externes, structurées et / ou non structurées à l'aide de plusieurs requêtes. Les flux de données vont donc venir rafraîchir la base décisionnelle. L'idéal est de ne recharger que les données modifiées ou ajoutées depuis la dernière extraction. Certains outils fournissent ce type de fonctionnalité, et s'appuie sur un mécanisme de marquage des données (date de la dernière modification associée à la donnée).

⁷ <https://grim.developpez.com/cours/businessintelligence/concepts/conception-datawarehouse/>

❖ **La transformation des données**

La transformation est la seconde phase du processus. Cette étape, qui du reste est très importante, assure en réalité plusieurs tâches qui garantissent la fiabilité des données et leurs qualités. Ces tâches sont :

- ✓ Consolidation des données.
- ✓ Correction des données et élimination de toute ambiguïté.
- ✓ Élimination des données redondantes.
- ✓ Compléter et renseigner les valeurs manquantes.

Cette opération se solde par la production d'informations signes d'intérêt pour l'entreprise qui sont prêtes à être entreposées.

❖ **Le chargement des données**

C'est la dernière étape et indispensable dans l'alimentation d'un entrepôt de données. Elle s'occupe de charger les données, préalablement extraites puis transformées, dans des entrepôts de données. Elle est toute fois très délicate et exige une certaine connaissance des structures du système de gestion de la base de données (tables et index) afin d'optimiser au mieux le processus. Le chargement va permettre d'insérer ou de mettre à jour les données cibles.

3) Le déploiement et mise en œuvre :

3.1) Requête ad hoc :

Aujourd'hui largement utilisé dans les entrepôts de données actifs, les requêtes ad hoc permettent d'accélérer la récupération des informations vitales pour répondre aux requêtes interactives dans une application à mission critique.

L'accès à ce genre de service peut se faire via différentes méthodes et outils. Cependant, les spécialistes préconisent de laisser la possibilité à l'utilisateur de choisir les outils qui lui paraissent les plus adéquats.

3.2) Reporting:

Le terme de « Reporting » désigne une famille d'outils de Business intelligence destinés à assurer la réalisation, la publication et la diffusion de rapports selon un format prédéterminé. Ils sont essentiellement destinés à faciliter la communication de résultats chiffrés ou d'un suivi d'avancement⁸.

⁸ <https://www.piloter.org/business-intelligence/reporting.htm>

3.3) Analyse dimensionnelle des données :

L'analyse dimensionnelle est sans doute celle qui exploite et fait ressortir au mieux les capacités de l'entrepôt de données. Le but par l'analyse dimensionnelle est d'offrir aux utilisateurs la possibilité d'analyser les données selon différents critères afin de confirmer une tendance ou suivre les performances de l'entreprise.

Cette analyse se fait selon le principe OLAP, offrant de ce fait aux utilisateurs les possibilités de recourir à différentes opérations facilitant la navigation dans les données. La mise en place de ces outils est une option très intéressante dans la mesure où les données seront accessibles en analyses instantanées. Plusieurs fournisseurs de solution OLAP existent sur le marché et offrent des solutions construites sur des méthodes et technologies différentes.

3.4) Tableau de bord :

Le tableau de bord est un ensemble d'indicateurs peu nombreux, conçus pour permettre aux gestionnaires de prendre connaissance de l'état et de l'évolution des systèmes qu'ils pilotent et d'identifier les tendances qui les influenceront sur un horizon cohérent avec la nature de leurs fonctions. [26]

C'est un outil de pilotage qui souligne l'état d'avancement dans lequel se trouve le processus afin de permettre au responsable de mettre en place des actions correctives.

3.5) Data mining :

Le Data Mining que l'on peut traduire par forage de données, Ensemble de techniques et de méthodes du domaine des statistiques, des mathématiques et de l'informatique consiste à chercher dans un grand volume de données afin d'en extraire des informations pertinentes pour le décideur.

Le point important du Data Mining est que l'utilisateur ne sait pas ce qu'il cherche. En effet, les outils de Data Mining recherchent, de manière automatique ou semi-automatique, des corrélations invisibles entre des données n'ayant à priori aucun lien entre elles.

L'utilisation industrielle ou opérationnelle de ce savoir dans le monde professionnel permet de résoudre des problématiques très diverses, allant de la gestion de relation client à la maintenance préventive, en passant par la détection de fraudes ou encore l'optimisation de sites web. [27]

4) Maintenance et croissance :

L'évolution du schéma du Data Warehouse nécessite un suivi constant compte tenu des besoins d'optimisation de performance ou d'expansion. Il est donc nécessaire d'investir dans les domaines suivants :

- ✓ **Support** : Il apporte une assistance de premier niveau aux utilisateurs dans l'utilisation du Data Warehouse pour la résolution et la correction des problèmes rencontrés.
- ✓ **Formation** : il est indispensable d'offrir un programme de formation permanent aux utilisateurs de Data Warehouse.
- ✓ **Support technique** : Le support technique surveille avec la plus grande vigilance les performances et les tendances concernant la charge du système.
- ✓ **Management de l'évolution** : il faut toujours s'assurer que l'implémentation répond aux besoins de l'entreprise. Les revues systématiques à certain point de contrôle sont un outil clé pour détecter et définir les possibilités d'amélioration. En plus du suivi et de la maintenance du Data Warehouse, des demandes d'expansion sont envisageables pour de nouveaux besoins, de nouvelles données ou pour des améliorations.

5) Conclusion

La mise en place d'une solution business intelligence « système décisionnel » est indispensable pour toute entreprise afin de consolider ses informations et d'avoir une aide à la décision.

La plus part de ces systèmes nécessitent la mise en place un Data Warehouse qui se représente comme un support de stockage de données exploitables à des fins décisionnelles.

Dans cette partie, nous avons défini les concepts de base en se qui concerne BI, les systèmes décisionnels, la modélisation dimensionnel ainsi que le Data warehouse .

Au cours de la seconde partie de cette étude, nous allons essayer d'utiliser les concepts présentés dans la synthèse bibliographique, et cela afin de mettre en œuvre notre Solution BI.

Partie 2 :
Cas pratique
« Région de Distribution
Centre RDC/SADEG filiale
SONELGAZ »

Chapitre 1 : Présentation de l'organisme D'accueil

1) Introduction

Le rôle essentiel d'une entreprise est de définir sa politique, puis de l'orienter dans le sens des objectifs prévus, pour assurer sa rentabilité et son développement dans les meilleures conditions Possibles compte tenu de son environnement socio-économique interne et externe.

Donc, gérer une entreprise, c'est organiser de façon productive un ensemble de ressources humaines, matérielles et financières. Cette organisation ne peut se faire sans que des informations ne soient disponibles sur :

- ✓ la situation de l'entreprise à un moment donné,
- ✓ les conséquences de l'activité passée,
- ✓ les effets possibles des décisions envisagées.

Toute organisation soucieuse de planifier son activité, doit :

- Se fixer un but à atteindre, compte tenu de ses désirs, de ses possibilités et des opportunités que lui offre son environnement,
- Se donner les moyens d'y parvenir, sur les plans : humain, matériels et financier,
- Décider des choix les plus appropriés et les mettre en œuvre,
- Contrôler à posteriori que le but était correctement défini, que les moyens adéquats ont été mis en œuvre correctement et mesurer le degré de réalisation.

Il est cependant nécessaire de contrôler la gestion de l'entreprise, c'est à dire maîtriser sa conduite en s'efforçant de prévoir les événements pour s'y préparer avec son équipe et s'adapter à une situation évolutive.

A cet effet, nous allons faire la présentation générale de la SONELGAZ et en particulier, la région de distribution centre(RDC) à travers son département suivie et inspection (DSI) qui ont constitué notre champ de prospection.

2) Présentation de SONELGAZ

2.1) Historique

SONELGAZ est l'opérateur historique dans le domaine de la fourniture des énergies électrique et gazière en Algérie. Ses missions principales sont la production, le transport et la distribution de l'électricité ainsi que le transport et la distribution du gaz par canalisations.

Son nouveau statut lui confère la possibilité d'intervenir dans d'autres segments d'activités présentant un intérêt pour l'entreprise et notamment dans le domaine de la commercialisation de l'électricité et du gaz à l'étranger. Durant son existence le groupe a connu des évolutions majeures qui peuvent être résumées comme suit :

- **1947**, Création d'EGA : C'est le décret du 5 juin 1947 qui a créé l'Etablissement Public National « Electricité et Gaz d'Algérie » (EGA par abréviation). Par décret du 16 août 1947, seize sociétés qui se partageaient les concessions électriques ont été transférées à EGA. Ces sociétés détenaient alors 90% des propriétés industrielles électriques et gazières du pays.
- **1962**, Le défi de la relève : Cette année représente la prise en main de l'Algérie indépendante de la SONELGAZ –alors Electricité et Gaz d'Algérie – et cela en faisant face au départ massif de cadres et techniciens français.
- **Période allant de 1962 à 1969** : Cette période a été caractérisée par la baisse de la consommation (une baisse de près de 33% en deux ans) dû au départ massive des étrangers qui représentaient plus de 87% de la clientèle. Par ailleurs, les tâches les plus urgentes ont été de reprendre le fichier des abonnés, reconstituer les plans des ouvrages et des réseaux, procéder au recrutement et à la formation dans tous les domaines et de ramener le niveau de consommation de l'énergie à celui de 1961.
- **1969, création de SONELGAZ**: C'est l'ordonnance n° 69-59 du 28 juillet 1969) portant dissolution de l'EGA et création de la nouvelle Société Nationale de l'Electricité et du GAZ - SONELGAZ-. Ce texte s'inscrit dans le cadre des mesures de nationalisation des secteurs clés de l'économie nationale. L'ordonnance précitée a attribué à l'entreprise le monopole de la production, du transport, de la distribution, de l'importation et de l'exportation de l'électricité et du gaz manufacturé (art. 4 et 7). L'ensemble des biens de l'ex-EGA lui a été légué.

- **1977**, le plan national d'électrification : Dès le milieu des années 70, l'Algérie s'est engagée dans un ambitieux plan national d'électrification qui avait objectif l'amélioration des conditions de vie des populations des campagnes tout en assurant un développement harmonieux de l'espace rural.
- **1983**, naissance des entreprises travaux : six entreprises autonomes voient le jour : KAHRIF pour l'électrification; KAHRAKIB-Infrastructures et installations électriques; KANAGAZ - Réalisation des réseaux gaz; INERGA - Génie civil; ETTERKIB - Montage industriel et l'entreprise ; AMC - Fabrication des compteurs et appareils de mesure et de contrôle.
- **1991**, SONELGAZ EPIC : SONELGAZ change de nature juridique et devient Etablissement Public à caractère Industriel et Commercial (EPIC) en vertu du décret exécutif n° 91-475 du 14 décembre 1991, portant transformation de la nature juridique de la Société Nationale de l'Electricité et du Gaz. Le décret exécutif n° 95-280 du 17 septembre 1995 confirme la nature de SONELGAZ en tant qu'Etablissement Public à caractère Industriel et Commercial.
- **1998**, création de filiales périphériques : Le 1er janvier 1998, neuf filiales périphériques ont vu le jour.
- **2002**, promulgation de la loi 02 / 01 du 5 février 2002 : Promulguée en février 2002, la nouvelle loi relative à l'électricité et à la distribution du gaz par canalisations est venue supprimer le monopole exercé jusque là par SONELGAZ, en ouvrant le secteur de l'électricité et du gaz à la concurrence, sauf pour les activités de Transport qui ont un caractère de monopole naturel.
- **Juin 2002, SONELGAZ SPA** : En vertu du décret présidentiel n° 02-195 du 1er juin 2002 portant statuts de la Société algérienne de l'électricité et du gaz dénommée "SONELGAZ. Spa", SONELGAZ est passé d'Etablissement Public à caractère Industriel et Commercial à une Société Par Actions dont le capital est détenu par l'Etat. Sur le plan de son fonctionnement, SONELGAZ Spa est dotée d'une Assemblée Générale et d'un Conseil d'Administration. Elle est dirigée par un Président directeur général.
- **2004** : SONELGAZ devient un holding ou groupe d'entreprises.

- **2004-2006** : Création des filiales :
 - Sonelgaz Production Electricité (SPE).
 - Gestion réseau transport électricité (GRTE).
 - Gestion réseau transport gaz (GRTG).
- **2006** : création nouvelles filiales du groupe industriel SONELGAZ.
- **OS** : Opérateur Système, chargée de la gestion du système de production et de transport d'électricité.
 - SDA : Sonelgaz distribution d'Alger.
 - SDC : Sonelgaz distribution Centre.
 - SDE : Sonelgaz distribution Est.
 - SDO : Sonelgaz distribution Ouest.

Les missions des filiales distribution :

- ✓ L'exploitation et l'entretien du réseau de distribution d'électricité et du gaz.
 - ✓ Le développement du réseau et l'assurance de la sécurité, l'efficacité et la qualité du service.
- **2009**: Parachève sa transformation en une Holding de sociétés déterminée à développer et renforcer ses infrastructures électriques et gazières.
 - **2011** : Sonelgaz. Spa est organisée en « SOCIETE HOLDING », sans création d'une personne morale nouvelle. La Holding Sonelgaz et ses sociétés filiales forment alors un ensemble dénommé « Groupe Sonelgaz ».
 - **2012** : Nouvel essor, Nouveau cap : les énergies renouvelables.
 - **2014 / 2015** : Le partenariat au cœur du développement
 - ✓ La création d'une société dénommée GEAT (General Electric Algeria Turbines), chargée de la réalisation et de l'exploitation d'un complexe industriel destinée à produire des TG et TV en partenariat avec Général Electric.
 - ✓ Partenariat avec Hyundai et Daewoo et création de la société dénommée HYENCO chargée de réaliser des prestations d'EPC (Engineering, Procurement and Construction) d'ouvrages énergétiques industriels.

- **22 mai 2017** : nouveau schéma organisationnel de l'activité de distribution décidé par les Pouvoirs Publics, a induit la création de la Société Algérienne de Distribution de l'Electricité et du Gaz, dénommée SADEG. Spa, résultat de la fusion - absorption des sociétés SDC, SDE, SDO et SDA.
- **Dimension internationale** : Création de sociétés mixtes avec des opérateurs internationaux d'envergure dont les plus récentes « Vijai Electricals Algérie », « Boilers Handassa Industrie Algérie » et « SEDIVER Algérie ».
- **2020-2021** : une nouvelle organisation a été mise en place sur deux niveaux
 - ❖ Niveau Groupe : La création, au sein du Groupe, de deux (02) holdings détenus à 100% par la holding mère Sonelgaz :
 - Une holding Engineering et Construction qui exercera un rôle hiérarchique sur les sociétés travaux et engineering.
 - Une holding Industries qui exercera un rôle hiérarchique sur les sociétés MEI, AMC, Rouiba Eclairage, VIJAI, BHI, SEDIVER et GEAT.
 - ❖ Niveau Holding :
 - Transformation des pôles d'activités en directions exécutives auxquelles seront rattachées, fonctionnellement, les sociétés du Groupe Sonelgaz regroupées par métier et activité, comme suit :
 - Direction Exécutive Production Conventiionnelle et Energies Renouve lables.
 - Direction Exécutive Transport et Distribution d'Energies.
 - Direction Exécutive des Activités de Services.
 Création d'un Secrétariat Général (SG).
 - Création d'une Direction Centrale chargée du pilotage de la performance et du contrôle de gestion (DCG).
 - Création d'une Direction des Affaires Juridiques (DAJ).
 - Création d'une Direction Centrale du Développement à l'International (DDI).
 - Transfert de la supervision et le contrôle de la société IFEG à la Direction Exécutive du Capital Humain et du Développement de l'Organisation (DCH).

**SCHÉMA ORGANISATIONNEL
DU GROUPE SONELGAZ**

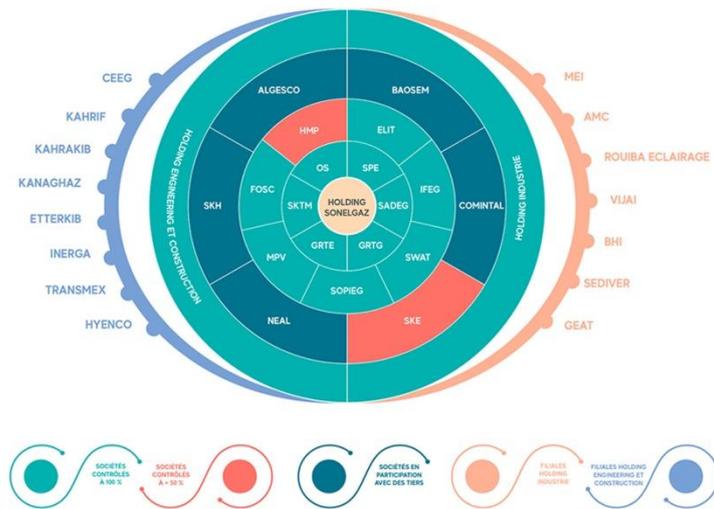


Figure 15 : Schéma organisationnel de groupe Sonelgaz⁹

2.3) Activités et missions de SONELGAZ :

SONELGAZ est une entreprise regroupant des moyens différents ; tendant vers un but unique qui consiste à mettre à la disposition des utilisateurs de l'électricité et du gaz sur tout le territoire national et selon une politique de tarification qui relève de l'autorité publique.

SADEG.Spa Ayant pour mission :

- La production, le transport, la distribution et la commercialisation de l'électricité tant en Algérie qu'à l'étranger.
- Le transport du gaz pour les besoins du marché national.
- La distribution et la commercialisation du gaz par canalisation tant en Algérie qu'à l'étranger.
- Le développement et la fourniture de toutes prestations en matière de service énergétiques.
- Développement par tout moyen de toute activité ayant un lien direct ou indirect avec les industries électriques et gazières et de toute activité pouvant engendrer un intérêt pour Sonelgaz.spa et généralement toute opération de quelque nature qu'elle soit pouvant se rattacher directement ou indirectement à

⁹ <https://www.sonelgaz.dz/fr/category/qui-s-ommes-nous>

son objet social, notamment la recherche, l'exploitation, la production et la distribution d'hydrocarbure.

- L'étude, la promotion et la valorisation de toutes formes et source d'énergie.
- Le développement de toute forme d'activités conjointes en Algérie et hors l'Algérie avec des sociétés algériennes ou étrangères.

3) Le métier de la distribution

L'un des métiers les plus importants du **SADEG.Spa**, et dans lequel s'inscrivent notre projet, est la fourniture et la distribution de l'énergie électrique et gazière. Ce métier, vu l'organisation du groupe, est assuré par quatre filiales qui sont les « Régions de Distribution ». Les régions sont :

- Région de Distribution d'Electricité et du Gaz de l'Ouest (RDO).
- Région de Distribution d'Electricité et du Gaz du Centre (RDC).
- Région de Distribution d'Electricité et du Gaz d'Alger (RDA).
- Région de Distribution d'Electricité et du Gaz de l'Est (RDE).

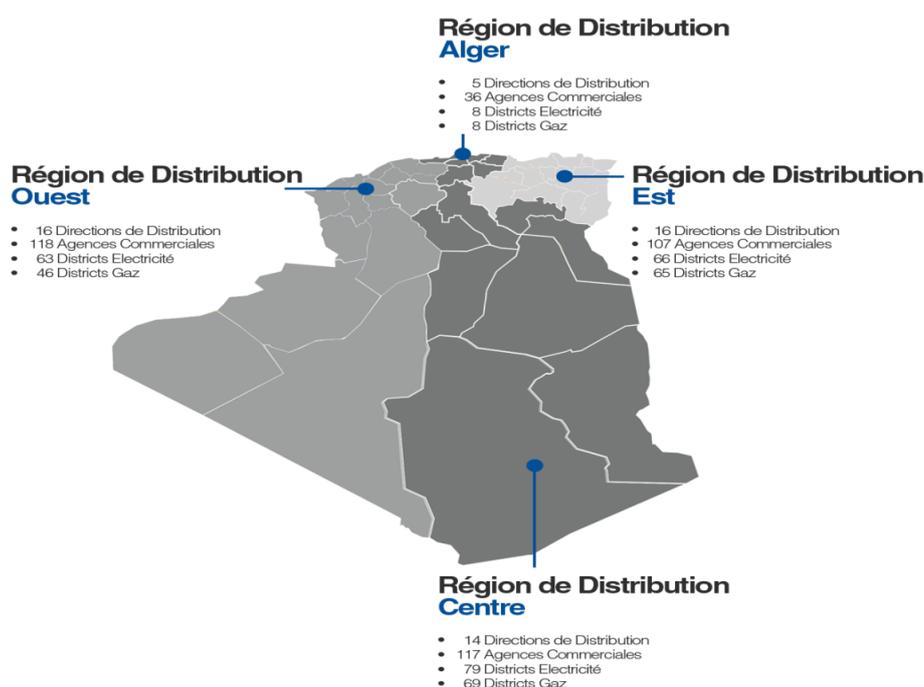


Figure 16 : Schéma illustre différents région de distribution¹⁰.

¹⁰ <https://www.sadeg.dz/?option=rubrique&i dr=43>

Les Régions de Distribution d'Electricité et du Gaz ont pour principales missions d'assurer :

- ✓ L'exploitation et la maintenance du réseau de distribution de l'électricité et du gaz
- ✓ Le développement des réseaux électricité et gaz permettant le raccordement des nouveaux clients.
- ✓ La commercialisation de l'électricité et du gaz.

3.1) Organisation de Région de Distribution Centre

SONELGAZ Région Distribution Centre (RDC) est chargée dans la limite de ses attributions, de la distribution de l'énergie électrique et gazière de la satisfaction des besoins de la clientèle aux conditions de coût, de qualité de service et de sécurité.

La RDC compte cinq directions centrales, situé au niveau de son siège, et gère 14 « **Direction de Distribution (DD)** ».

Ces 14 Concessions de Distribution (CD) sont : Blida, Biskra, Bouira, Boumerdès, Djelfa, ElOued, Ghardaïa, Illizi, Laghouat, Médéa, Ouargla (URBAIN, RURAL), Tamanrasset, Tipaza, Tizi-Ouzou.

L'organigramme suivant illustre :

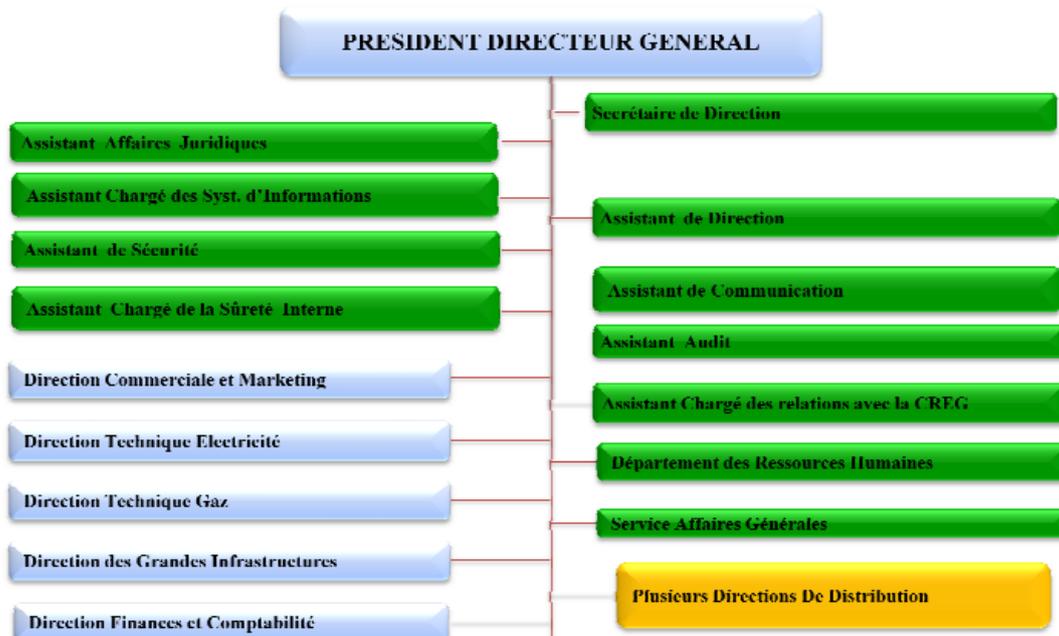


Figure 17 : Organisation des régions de distribution

➤ **Attributions**

1. Adapter le réseau d'aujourd'hui aux nouvelles technologies de demain afin de rester en adéquation permanente avec les besoins des clients.
2. Sécuriser les infrastructures vis-à-vis d'une part, de l'évolution des consommations, en particulier la consommation électrique (capacité du réseau et des postes sources, sécurisation des lignes aux aléas climatiques), et d'autre part, de la croissance de la production et les accidents de grande ampleur.
3. Accompagner et réaliser, dans le cadre du programme de l'Etat, les projets de développement, conformément aux missions de la Société.
4. Satisfaire aux meilleures conditions la demande de raccordement des clients électricité et gaz.
5. Contribuer à bâtir une politique sociale en cohérence avec nos valeurs et nos objectifs.

3.2) Direction de Distribution

Chacune des 14 **Directions de Distribution** gère plusieurs divisions et services. Leur organisation se présente comme suit :

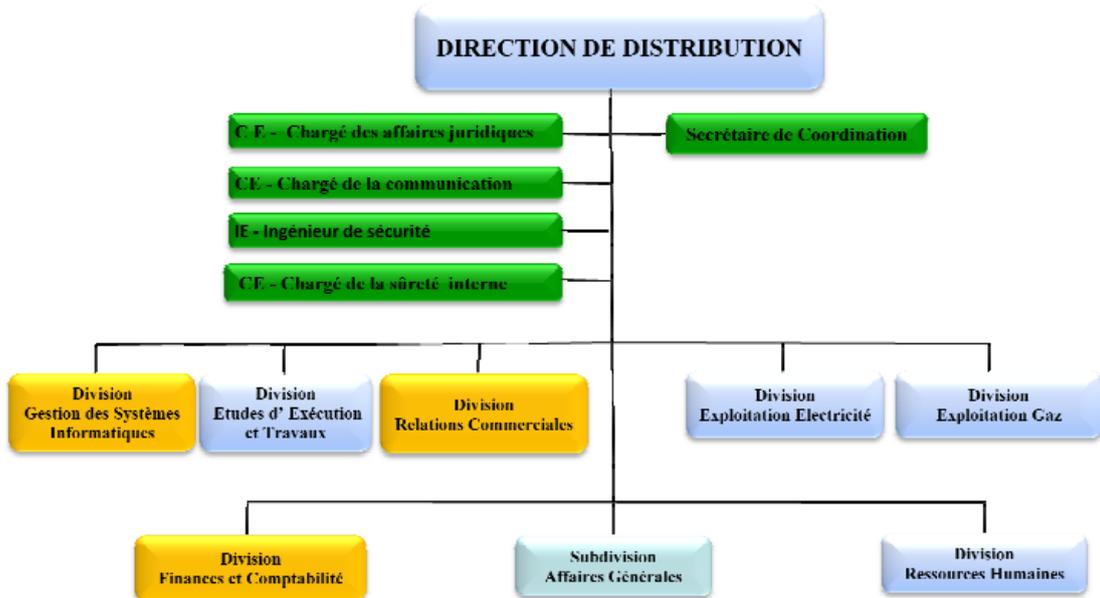


Figure 18 : Organisation des Directions de Distribution.

➤ **Attributions**

1. Participer à l'élaboration de la politique de la direction générale distribution (en matière de prestations rendus aux clients, développement des ventes, recouvrement des créances,.....).
2. Mettre en œuvre la politique commerciale de l'entreprise et en Contrôles l'application.
3. Satisfaire aux meilleures conditions de coûts et de délais la demande de raccordement des clients MT/BT, MP/BP et leur apporter conseil et assistance.
4. Assurer la gestion (conduite, exploitation, et maintenance) et le développement des réseaux MT/BT et MP/BP et des installations annexes.
5. Elaborer et mettre en œuvre le développement de la construction et la maintenance et l'exploitation des ouvrages.
6. Etablir les programmes travaux qui se rapportent à ses missions et en Assurer la maîtrise d'œuvre.
7. Assurer la gestion et le développement de la ressource humaine et des moyens matériels nécessaires au fonctionnement de la DR.
8. Assurer la sécurité des personnes et des biens en rapport avec les activités de la distribution.
9. Assurer la représentation de SONELGAZ au niveau local.

3.3) Département suivi et Inspection

La DTE à travers son département suivi et inspection (DSI) prend en charge l'étude, exécution et le suivi des différents projets de réalisation des programmes d'électrification. Ces projets d'électrifications sont répartir et validé sur les 14 directions de distributions de région centre.

➤ **Attribution**

1. Lancement d'un avis d'appel d'offre
2. Consultation des entreprises
3. Lancement des travaux
4. Suivi des travaux
5. Réception des travaux

4) Conclusion

La définition du contexte du projet est une étape indispensable.

La présentation de l'organisme d'accueil, son activité et ses structures permet de mieux appréhender le projet de mise en place d'un système décisionnel et d'aborder sa planification avec plus de facilité.

Au cours du chapitre suivant, on va entamer la conception de notre solution BI en faisant une étude approfondie du système existant à fin d'identifier l'ensemble des besoins que notre système décisionnel doit les satisfaire et procéder à la modélisation dimensionnel de ces activités.

Chapitre 2 : Etude de l'existant

1) Introduction

La RDC à travers le Département Suivi et Inspection relevant de la Direction Technique d'électricité veut, par le biais de ce projet, remédier à un manque important en matière de décisionnel. Ce manque se caractérise par le quasi inexistence de support d'aide à la décision, et l'indisponibilité de moyens de Reporting efficaces, en mesure de fournir des informations valides en temps réel.

De ce fait, nous allons présenter une analyse concernant l'existant décisionnel du groupe dans le cadre de sa fonction de distribution, ainsi que faire connaître les procédures et les méthodes de Reporting et de prise de décision relative à l'activité.

2) Processus de réalisation des programmes d'électrifications

2.1) Lancement des programmes d'électrification

Cela s'inscrit dans le cadre de programme d'état destinée pour les programmes quinquennal d'électrification au niveau national avec un mode de financement de programme partage de 75% de l'enveloppe global à la charge de l'état et 25% à la charge de SONELGAZ –SADEG dont le 1^{er} objectif est de recenser les affaires de programmes d'électrifications avec ces deux formules :

- Le raccordement des centres (Electrification Rurale « ER »).
- Le raccordement des Quartiers et lotissements sociaux (QLS).

2.2) Etablissement des listings des affaires au niveau des communes

La Direction Technique d'électricité à travers le Département Suivi et Inspection (DSI) informe les différents DD du lancement de programme d'électrification avec ces 2 formules pour se procédé avec les directions de l'énergie Wilaya (DMI) au recensement des besoins des commune de wilaya en ER/QLS en programmant des sortie sur terrain a fin de faire une étude préliminaire en définissant les besoins des familles et élaborer une fiche délimitation d'assiette en (MT ,BT, Branchement).

Cette fiche sera émise au bureau d'étude au niveau wilaya pour faire une étude d'exécution dont le but de validé la liste des différents programmes déjà recenser à travers l'ensemble de communes. Cette liste validé avec les consistances réelles en terme MT, BT, Branchement sera remettre au DTE précisément DSI pour faire l'estimation financière (cout/affaire) et effectuer des corrections suivant ce qui existe sur le marché.

2.3) Signature de convention entre ministère de l'énergie et la SADEG

Dès l'achèvement de l'étude technique, un rapport final validé sera envoyé par RDC au SADEG qui doit par la suite le transféré au ministère de l'énergie pour élaborer la convention sur la base des consistances physiques et financières. Après certification par le ministère et le PDG « SONELGAZ », la convention sera exécutable. Elle sera envoyée à la SADEG qui doit par la suite demandé de RDC à travers la DTE (DSI) à fin de lancer les appelq d' offre pour la réalisation des travaux.

2.4) Lancement des consultations

La DTE suivant le DSI demande aux DD de procédé au lancement de projet conformément à la convention. Les soumissionnaires doivent passer par 2 phases :

- 1- **phase technique** : Fournir un cahier de charge concernant les moyens de l'entreprise (personnel, ancienneté, qualification, matériels, engins...) qui doit dépasser 75 points requis/ paramètres du processus de notation pour arrivé à la deuxième phase.
- 2- **phase commercial** « déposer l'offre financière » : Lors l'ouverture des prix, une entreprise est choisi par le fait que son offre est le moins disant et qu'elle rentre dans l'estimation SONELGAZ. Ces résultats vont être validés par une commission de marché.

Ensuite, un ordre de service est établi (ODS) pour que les travaux puissent démarrer et un contrat est signé entre la division administration de marché (DAM) et l'entreprise concernée. Celle-ci est tenue de respecter les délais prescrits par la Direction de la distribution.

2.5) Suivi physique et financier des affaires

L'ODS établi un dossier appelé une chemise travaux qui sera remis au surveillant travaux. La mission de ce surveillant est de sortir avec l'entreprise pour lui montrer où se trouve-le lieu de l'affaire, l'oriente selon la minute d'étude, ce qui s'appelle piquetage. Ainsi suit l'entreprise pas à pas pour la bonne exécution des travaux depuis le début jusqu'à la fin.

Financièrement, l'entreprise doit facturer ces dépenses au fur et à mesure que les réalisation physique se déroule en élaborant des PV nommés Attachement qui doit être visé et validé par le surveillant travaux. Ce PV résume le progrès effectué des travaux de réalisation physique des affaires avec le cout de réalisation.

Il est à noter que l'entreprise réalisatrice doit impérativement respecter le planning de réalisation ainsi que les consistances physique et consistances financière fixés dans le cahier de charge sauf en cas d'opposition (changement des prix de marchandise, présence des obstructions,...), dans le cas contraire des pénalités seront appliquées.

Chaque mois, le surveillant prépare un rapport pour mise à jour le suivi des réalisations qui doit être donné à la Division Etude Travaux (DET) au niveau des DD pour être transféré au DTE (DSI).

Suivant ce rapport mensuel, la DSI élabore un Reporting pour le remettre à la SADEG, de cette dernière au ministère.

2.6) Clôture des affaires

Après constatation que tous les travaux étaient réalisés sans aucun problème, le surveillant travaux organise une journée de réception où il établira un avis de fin travaux (AFT) dont il envoie une copie à la DRC accompagné des documents y afférant :

- Plans
- PV de réception

Par conséquent : un ordre de mise en service sera établi pour le raccordement, effectuer des déplacements sur les lieux pour constater la conformité de la réalisation, s'il n'y a rien à signaler l'affaire sera mise en service (réception provisoire) qui consiste à :

- La vérification du travail exécuté par l'entreprise et le respect des normes.
- La pose du compteur électricité
- La vérification du disjoncteur (différentiel ou ampérage).

Et une Autorisation de Paiement (AP (budget global convention)) par affaire sera créée au niveau de comptabilité pour permettre sa dépense et effectuer le paiement de l'entreprise réalisatrice et une période de garantie sera appliquée avec consigne.

Dès l'achèvement de cette période sans signaler aucune réserve sur les travaux réalisés par l'entreprise, une réception définitive ainsi qu'une facture définitive seront accommodés.

3) Etat du décisionnel au sein de RDC(DSI)

Il est intéressant de signaler que la RDC à travers DSI, dans sa fonction de suivie et inspection des affaires d'électrification, ne dispose d'aucun système d'aide à la décision automatique ou semi-automatique. Aussi, tout processus d'analyse et de prise de décision à tous les niveaux se base essentiellement sur des rapports mensuels issus des 14 « DD » de la région centre et que la consolidation des données se fait d'une manière manuelle par le logiciel « Excel ».

Pour donner une vision claire de la manière dont sont consolidées les données et les rapports élaborés, la procédure se déroule comme suit :

Phase 1 : les utilisateurs de la RDC, principalement les analystes et décideurs de la DSI, formulent leurs demandes qui sont transmises au différentes DD de la région centre.

Phase 2 : La division DET au niveau de chaque DD, en recevant la demande, lance la procédure d'extraction des données des systèmes transactionnels à fin d'élaborer des rapports et les transmettre au manager de chaque DD. Ces rapports seront ensuite transférés au DSI.

Phase 3 : la consolidation est faite à partir des fichiers de données si existe, et en se basant aussi sur les rapports transmis par les DD, toujours d'une façon manuelle et monopolise les ingénieurs et techniciens de DSI. Cette consolidation permet d'élaborer les rapports voulus.

Phase 4 : Le rapport est validé par le chef de département DSI et envoyé aux utilisateurs finaux.

Remarque : la procédure d'élaboration des rapports se fait généralement chaque fin de mois. Elle peut aussi être lancée en cas de nécessité.

Le diagramme suivant résume les étapes d'élaboration du rapport.

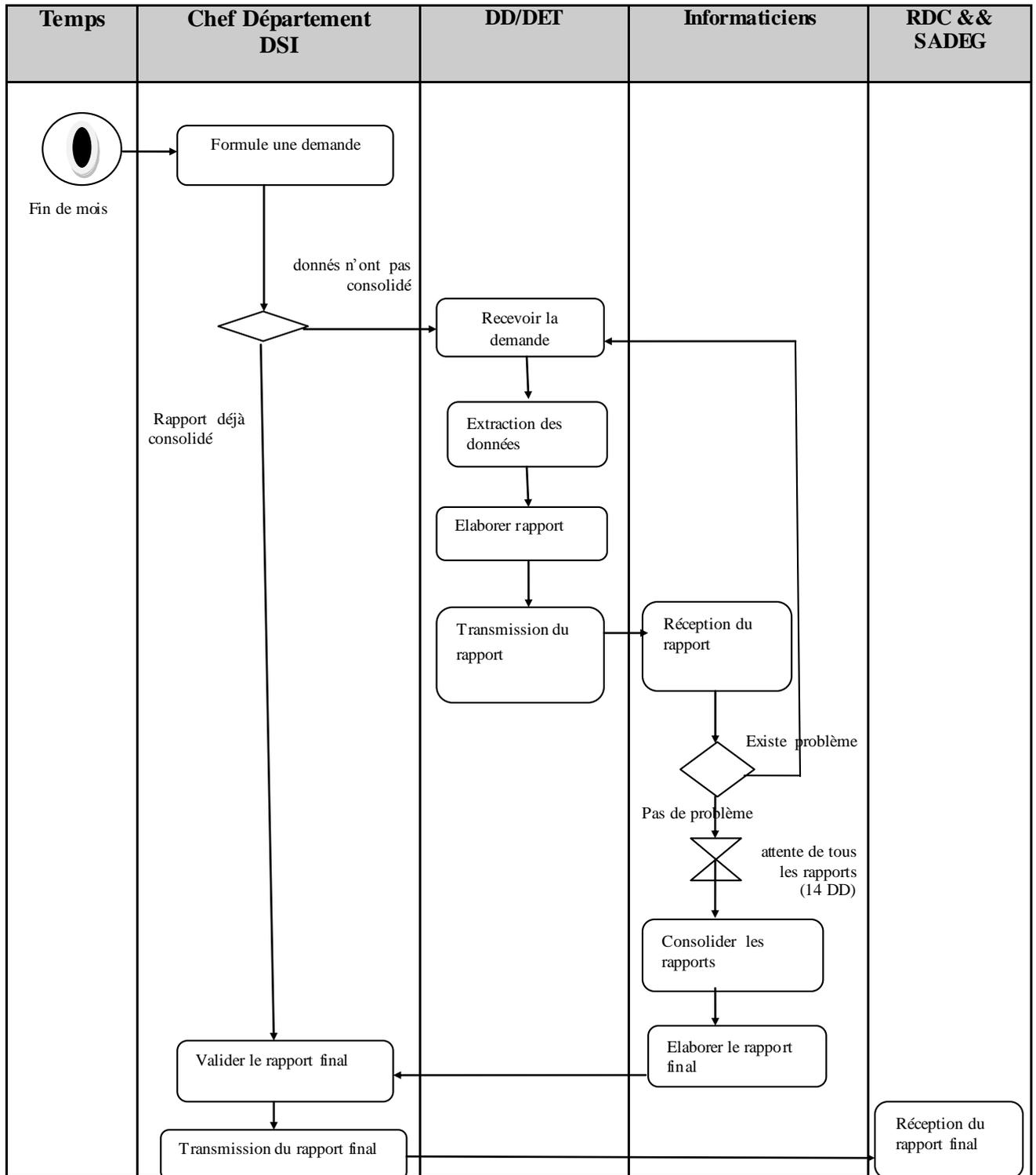


Figure 19 : Diagramme d'activité d'élaboration des rapports.

4) Conclusion

Cette étude nous a permis de savoir le processus de suivi des programmes d'électrifications et d'avoir une vision générale des procédures d'élaboration de rapports et de consolidation des données. Cela a conduit à définir comme premier point le périmètre général du projet et de l'étude des besoins en particulier ainsi que faire ressortir les insuffisances du système actuel en soulignant les points faibles.

Le chapitre suivant consacré à l'étude des besoins à fin de détecter ceux des utilisateurs de façon à pouvoir y répondre.

Chapitre 3 : Etude des besoins

1) Introduction

« Les chances de succès d'un Data Warehouse se trouvent considérablement accrues par la bonne compréhension des utilisateurs et de leurs besoins » [28].

Toute solution business intelligence « système décisionnel » doit être en mesure de répondre aux attentes des utilisateurs, cela ne peut se faire sans une étude approfondie de leurs besoins.

Comme tout projet informatique il est souvent très recommandé d'établir une bonne démarche de gestion de projet. Dans notre projet, nous nous sommes inspirées de la méthodologie pour les projets décisionnels préconisés par « Kimball » qui représente une succession des tâches nécessaires à la conception, au développement et au déploiement d'un Data Warehouse efficace. Nous avons opté pour cette méthodologie parce que :

- Elle comporte un cycle de vie bien défini qui offre la possibilité de décomposer certaines phases et éliminer certaines tâches.
- Elle se focalise sur les besoins de l'entreprise

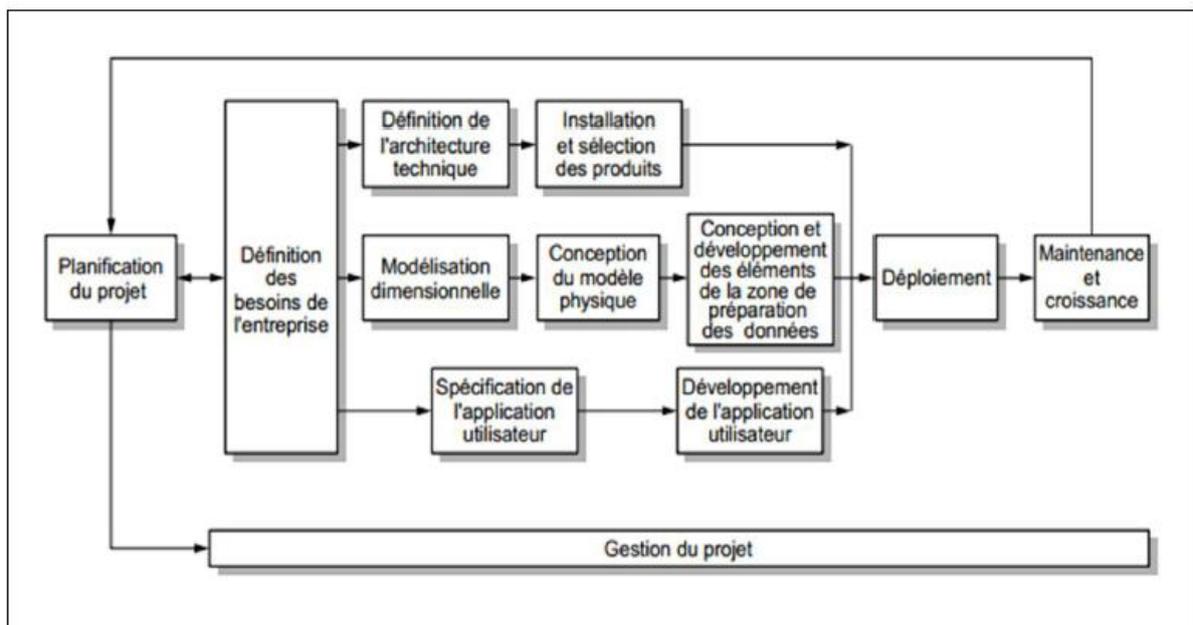


Figure 20 : Place de l'étude des besoins dans le cycle de vie du projet DW/BI. [28]

2) Démarche adopté pour l'identification des besoins

Afin de faire une étude complète et d'obtenir le maximum d'informations, nous avons concentré en premier lieu sur les besoins exprimés par les utilisateurs, puis, pour les enrichir, nous avons étudié les besoins relatifs aux données réelles.

Pour collecter l'ensemble d'informations auprès des utilisateurs, il existe plusieurs méthodes tel que les entretiens, les réunions, les questionnaires.... Chaque méthode présente des avantages et des inconvénients. Nous avons opté pour la technique d'entretien grâce à l'ensemble d'avantages qu'elle présente et qui peuvent se résumer en :

- Rapidité et simplicité d'utilisation
- Courts délais et faible coût: La préparation de l'entretien n'exige pas une durée très longue et ne demande pas trop de ressources.
- La participation des intervenants permet d'augmenter les chances de recueillir l'information voulue

D'autre part, cette méthode présente certains inconvénients dont notamment la difficulté d'interviewer les personnes qui ont une grande responsabilité à cause de la charge de leur emploi du temps et leur indisponibilité dans la plus part du temps au niveau de la direction, et le fait que les utilisateurs ne peuvent, même avec la meilleure volonté, exprimer tous leurs besoins. Pour cela nous avons eu recours à l'étude des rapports demandés auparavant et des données disponibles dans le but de fournir des informations exploitables.

L'étude des rapports, et données sert à détecté des besoins non déclarés par l'interviewé et qui peuvent apparaître ultérieurement, le but de cette démarche étant de construire un Data warehouse capable de répondre à des éventuels nouveaux besoins.

3) Identification des besoins

Les besoins recensés lors notre étude sont récapitulé comme suivant :

Volet détecté	Besoins enregistré
Suivi des entreprises	<p>Besoins : ce volet contient les informations relatives à l'entreprise réalisatrice des travaux. l'utilisateur devrait être en mesure de suivre ces entreprises selon différents critères d'analyse :</p> <p>(commune, temps, catégorie, direction distribution).</p>
Suivi des affaires	<p>Besoins : Avoir une vue d'ensemble de l'évolution des affaires et le suivi des progrès des travaux, cout et les délais de réalisations. L'utilisateur doit être en mesure de faire des analyses selon les critères : (commune, énergie, temps, entreprise)</p>
Suivi la clôture d'une affaire	<p>Besoins : Effectuer des analyses par rapport à ces axes : (situation, entreprise, temps, garantie)</p>

Tableau 4 : Synthèse des besoins détectés.

4) Les problèmes rencontrés

Lors de notre étude, nous avons rencontré plusieurs difficultés qui nous ont ralenties :

- Epidémie CORONA VIRUS que notre pays vive à ce moment nous a causé un véritable problème car on n'a pas pu planifier des interviews avec le personnel de plusieurs services qui peuvent intervenir avec efficacité dans notre domaine d'étude.
- L'emploi du temps chargé de l'interviewé,
- Annulation des rendez-vous à cause des imprévus (réunion,...)
- Interruption au milieu de l'interview à cause d'un appel important ou d'un travail qui ne peut être reporté.

5) Conclusion

La satisfaction des utilisateurs et la réussite des projets en général sont en effet largement dépendantes d'une bonne définition des besoins. Dans ce chapitre nous avons identifié, priorisé et affiné les besoins et attentes des utilisateurs finaux, et ce en utilisant l'entretien comme technique d'enquêtes et faire une étude de données et rapports existants.

Cette étude des besoins nous a permis de collecter beaucoup d'information et recensé les besoins qui nous seront utiles pour la prochaine étape, à savoir la construction et le développement de notre solution BI « système décisionnel » et son entrepôt de données.

Les besoins étant recensés, la construction du Data Warehouse peut alors commencer et cela on va le présenter dans le chapitre suivant.

Chapitre 4 : Conception de la solution

1) Introduction

Une fois les besoins des utilisateurs sont connus, nous pouvons débiter la conception des volets de notre Data warehouse. Pour cela nous avons eu recours à la modélisation dimensionnelle qui est souvent associés aux entrepôts de données compte tenue de ces avantages.

Dans cette partie, nous allons présenter la conception de notre entrepôt de données, son alimentation afin d'intégrer les données dans le DW finalement cubes dimensionnels dont le but de naviguer dans les données contenues dans le DW.

2) Conception d'entrepôt de donnée

2.1) Modélisation dimensionnelle

Selon Kimball, Le processus de modélisation dimensionnelle se décompose en général en quatre étapes :

- Sélection de l'activité d'entreprise à modéliser
- Déclaration du grain : Il répond à la question « comment décrivez-vous une ligne unique de la table de fait? ».
- Choix de dimension : « Comment les gestionnaires décrivent-ils des données qui résultent du processus concerné? ».
- Mesurable.

2.2) Modélisation des activités

a) Activité « suivi des affaires »

❖ Présentation de l'activité

La réalisation des programmes d'électrifications (ER, QLS) se fait à travers la mise en œuvre de plusieurs affaires d'électrification identifier dans la convention signée entre ministère et la SADEG.

La RDC suivant son DSI après avoir choisi les entreprises réalisatrices des programmes d'électrifications par le lancement de consultation, il prend en charge le suivi des progrès des travaux de réalisation des affaires sur les deux volets physique et financier.

❖ Grain de l'activité

Le grain de l'activité suivi des affaires représente, la possibilité de suivre une affaire selon des critères différents. Le grain peut alors être donné comme suit :

« Suivi dans le temps, la quantité d'énergie, le cout de réalisation d'une affaire concernant une entreprise activant dans une commune ».

❖ Les Dimensions participantes

Les dimensions ont pour objectif de décrire le fait, donc nous citons les différentes dimensions qui s'appliquent à chaque ligne de la table de fait et répondre au grain de l'activité.

1. Dimension temps

« C'est la seule dimension qui figure systématiquement dans tout entrepôt de données, car en pratique tout entrepôt de données est une série temporelle. Le temps est le plus souvent la première dimension dans le classement sous-jacent de la base de données » [14].

Libellé	Détails
Code_temps	Clé artificielle de la dimension
Date	La date au format complet
Jour	Position du Jour dans le mois.
Jour_semaine	Nom des jours de la semaine
Mois	Nom du mois.
Mois_annee	Numéro du mois dans l'année
Annee	Année de la date

Tableau 5 : Description de la dimension « Temps »

2. Dimension Entreprise

Cette dimension s'impose comme un élément important dans l'analyse. Elle contient l'ensemble des entreprises retenues pour la réalisation de programme d'électrification, en suivant leur comportement les décideurs peuvent faire de bonne analyse.

Libellé	Détails
Id_entreprise	Clé artificielle de la dimension
Code_entreprise	Code qui identifie l'entreprise
Num_registre_com	Numéro registre commerce
Libelle_entreprise	Intitulé de l'entreprise
Adresse	Adresse de l'entreprise
Code_postal	Code postal de l'entreprise
Tel	Téléphone de l'entreprise
Code_commune	Code commune siège social
Commune	Désignation commune siège social
Code_dd	Code de la direction de distribution
Nom_dd	Nom direction de distribution
Wilaya	Wilaya siège social
Secteur_activité	Spécialité

Tableau 6: Description de la dimension « Entreprise ».

3. Facture

Une facture est un document relatif au volet financier, elle contient un certain nombre d'informations intéressantes pour une analyse.

Libellé	Détails
Id_facture	Clé artificielle de la dimension
Num_facture	Le numéro de facture
Date_facture	La date de facturation
Type_facture	Le type de facture
Période_facturé	Période facturé
Montant_HTT	Le montant hors taxe
Montant_TTC	Le montant tous taxe

Tableau 7 : Description de la dimension « Facture ».

4. Situation

Cette dimension définit les différents états qui peuvent décrire une affaire sur laquelle on peut faire de bonne analyse.

Libellé	Détails
Id_situation	Clé artificielle de la dimension
Libelle_situation	Désignation de la situation

Tableau 8 : Description de la dimension « Situation ».

5. Direction distribution

Cette dimension décrit le lieu où se déroule ce fait. Elle contient plusieurs champs qui peuvent être l'objet d'une bonne analyse particulièrement la commune.

Libellé	Détails
Id_dd	Clé artificielle de la dimension
Code_dd	Code de la direction de distribution
Nom_dd	Nom de la direction de distribution
Code_commune	Code de la commune
Nom_commune	Nom de la commune
Code_wilaya	Code de la wilaya
Nom_wilaya	Nom de la wilaya
Adresse_dd	Adresse de la direction de distribution
Num_lot	Numéro lot dans la commune
Tel_dd	Téléphone de la direction de distribution

Tableau 9 : Description de la dimension « Direction distribution ».

6. Affaire

La dimension affaire est une dimension sur laquelle s'articule tout ce fait. Afin de bien suivre et faire une bonne analyse nous avons besoin de maximum d'information la concernant.

Libellé	Détails
Id_affaire	Clé artificielle de la dimension
Num_affaire	Numéro de l'affaire
Libelle_affaire	Intitulé de l'affaire
Code_commune	Code de la Commune
Intitulé_pgm	Type programme (ER, QLS)
MT	Consistance en moyen tension
BT	Consistance en basse tension
BRTS	Consistance branchement
Cout_etude	Cout d'étude
Cout_travaux	Cout des travaux
Cout_global	Cout global
Duree	Durée estimé de réalisation
Quotepart_etat	Participation d'Etat
Quoteparte_SDC	Participation SADEG
MT_realise	Moyen tension réalisé
BT_realise	Basse tension réalisé
BRTS_realise	Branchement réalisé

Tableau 10 : Description de la dimension « Affaire ».

❖ **Les mesurables**

Les mesurables qui correspondent à l'activité suivi des affaires et qui permettent de mesurer les performances de cette activité, est le « *cout de réalisation* ».

❖ Le modèle en étoile

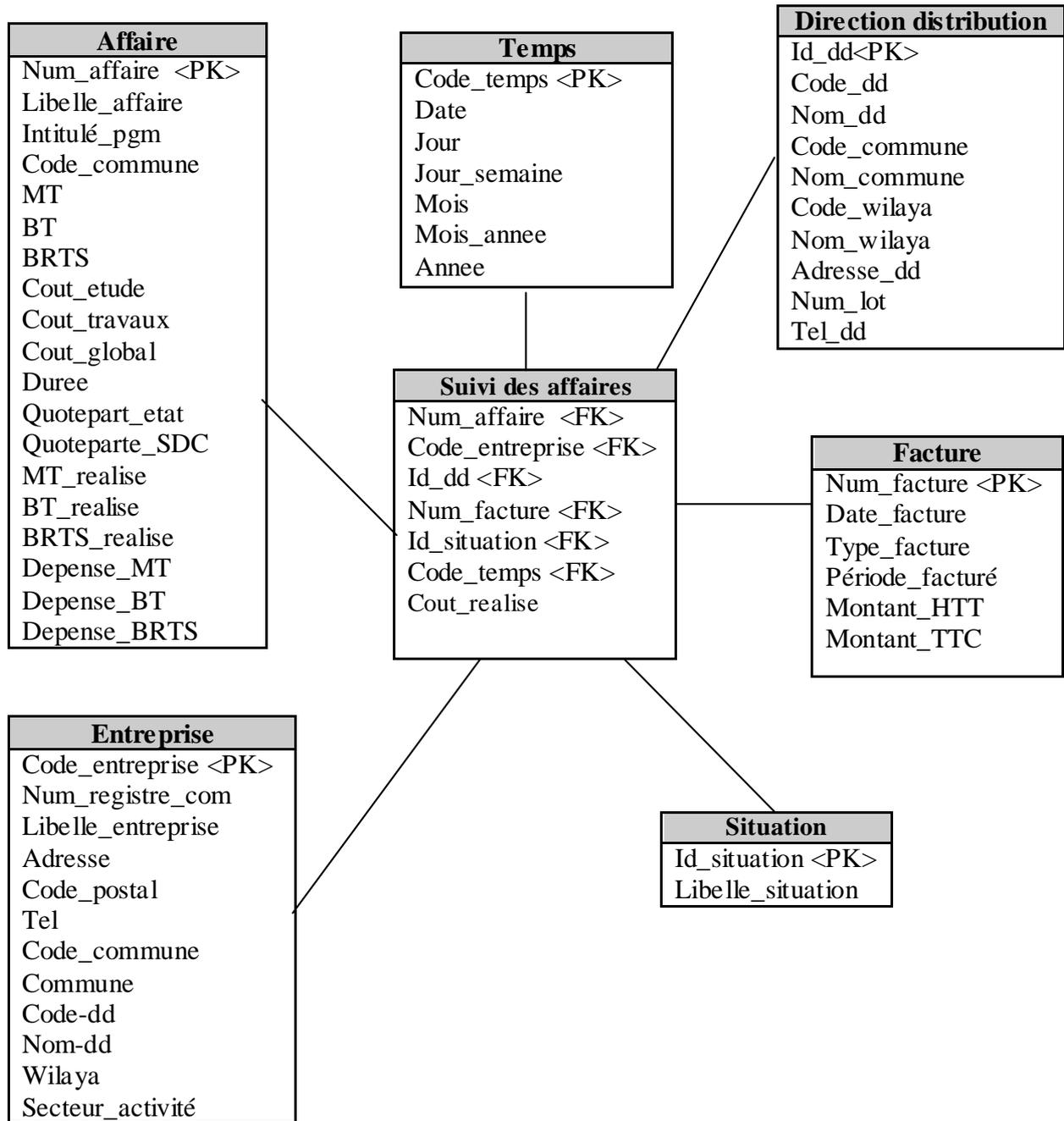


Figure 21 : Modèle en étoile de l'activité « Suivi des affaires ».

b) Activité « suivi la clôture d'une affaire »

❖ Présentation de l'activité

Dès l'achèvement les travaux de réalisation d'une affaire, une réception provisoire est effectuée ainsi que la mise en service. Une réception définitive sera programmée après une période de garantie.

❖ Grain de l'activité

Le grain le plus fin de l'activité correspond à :

« *Suivi une affaire par situation et garantie, a réalisé par une entreprise agissant dans une commune* ».

❖ Les dimensions participantes

En plus les dimensions « Temps », « Entreprise », « Facture », « Affaire », « Situation », « Direction distribution », déjà vu au cours de la modélisation de l'activité suivi des affaires, nous allons voir une autre dimension participante à la clôture d'une affaire qui est : « Garantie ».

1. Garantie

La dimension garantie renseigne la période de garantie qui sépare la réception provisoire et la réception définitive à fin d'assurer la clôture.

Libellé	Détails
Id_garantie	Clé artificielle de la dimension
Libelle_garantie	Désignation garantie

Tableau 11 : Description de la dimension « Garantie »

❖ Les mesurables

Le mesurable qui correspondent à l'activité suivi la clôture d'une affaires et qui permet de mesurer la performance de cette activité, est le « *Discount* ».

❖ Le modèle en étoile

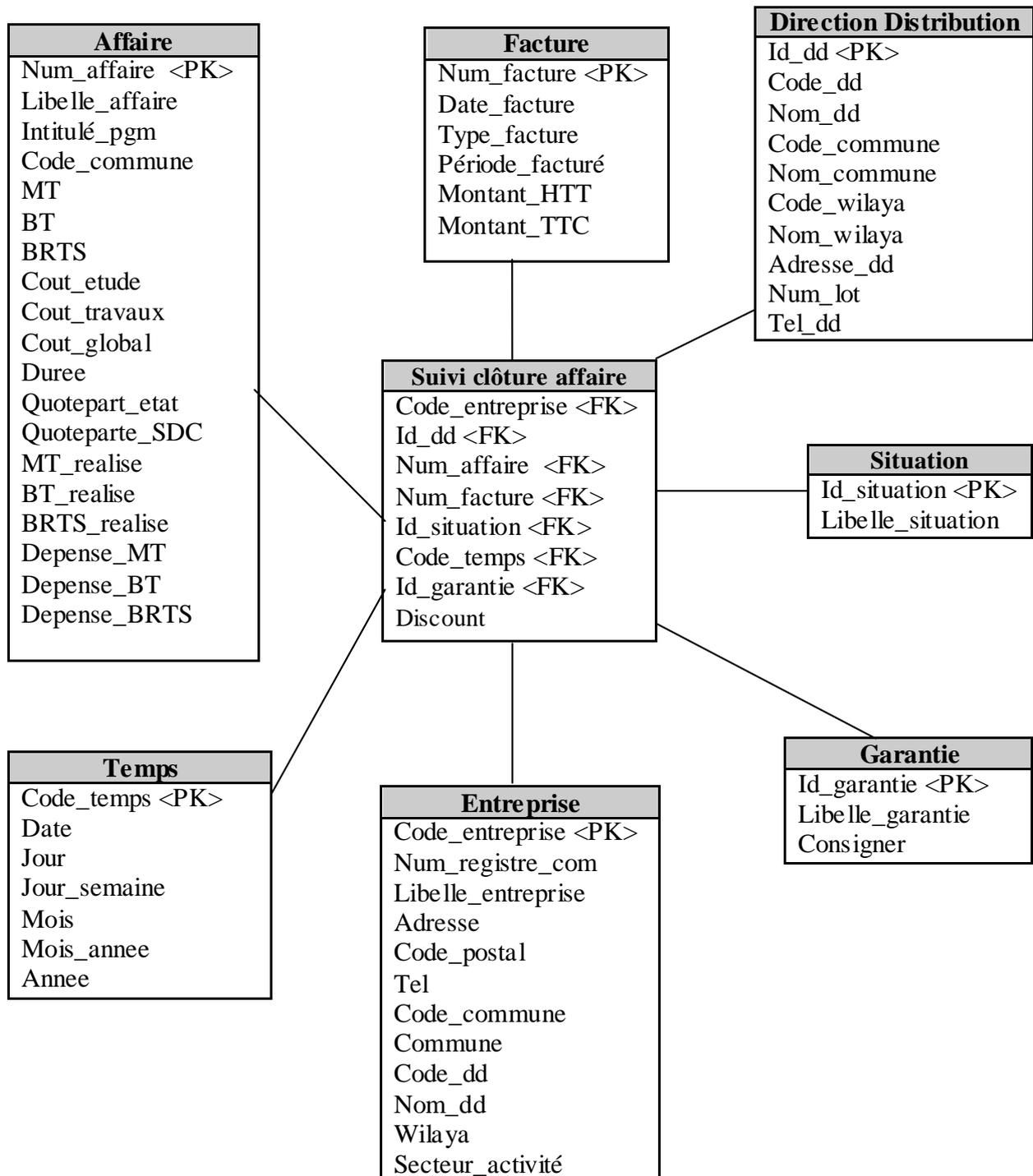


Figure 22 : Modèle en étoile de l'activité « suivi la clôture une affaire ».

c) Activité « suivi des entreprise »

❖ Présentation de l'activité

La RDC à travers son DSI a comme première vocation de lancer une consultation à fin de choisir des entreprises pour mettre en œuvre les programmes d'électrifications initiées par l'état. L'entreprise est la partie réalisatrice du projet d'électrification, en contrepartie il doit honorer ses engagements relatifs à la convention d'application, d'ou est née la nécessité de leur suivi.

❖ Grain de l'activité

Le grain le plus fin de l'activité correspond à :

« L'historique complet d'une entreprise par catégorie, affaire et direction distribution ».

❖ Les dimensions participantes

En plus les dimensions « Temps », « Entreprise », « Facture », « Affaire », « Direction Distribution » déjà vu au cours de la modélisation de l'activité suivi des affaires, nous allons voir une autre dimension participante au suivi de entreprise qui est : « Catégorie ».

Libellé	Détails
Code-catégorie	Clé artificielle de la dimension
Intitulé	Désignation état d'entreprise

Tableau 12 : Description de la dimension « Catégorie ».

❖ Les mesurables

Dans ce cas en se trouve avec une table de fait sans mesure, l'objectif est de savoir l'historique des entreprises.

❖ Le modèle en étoile

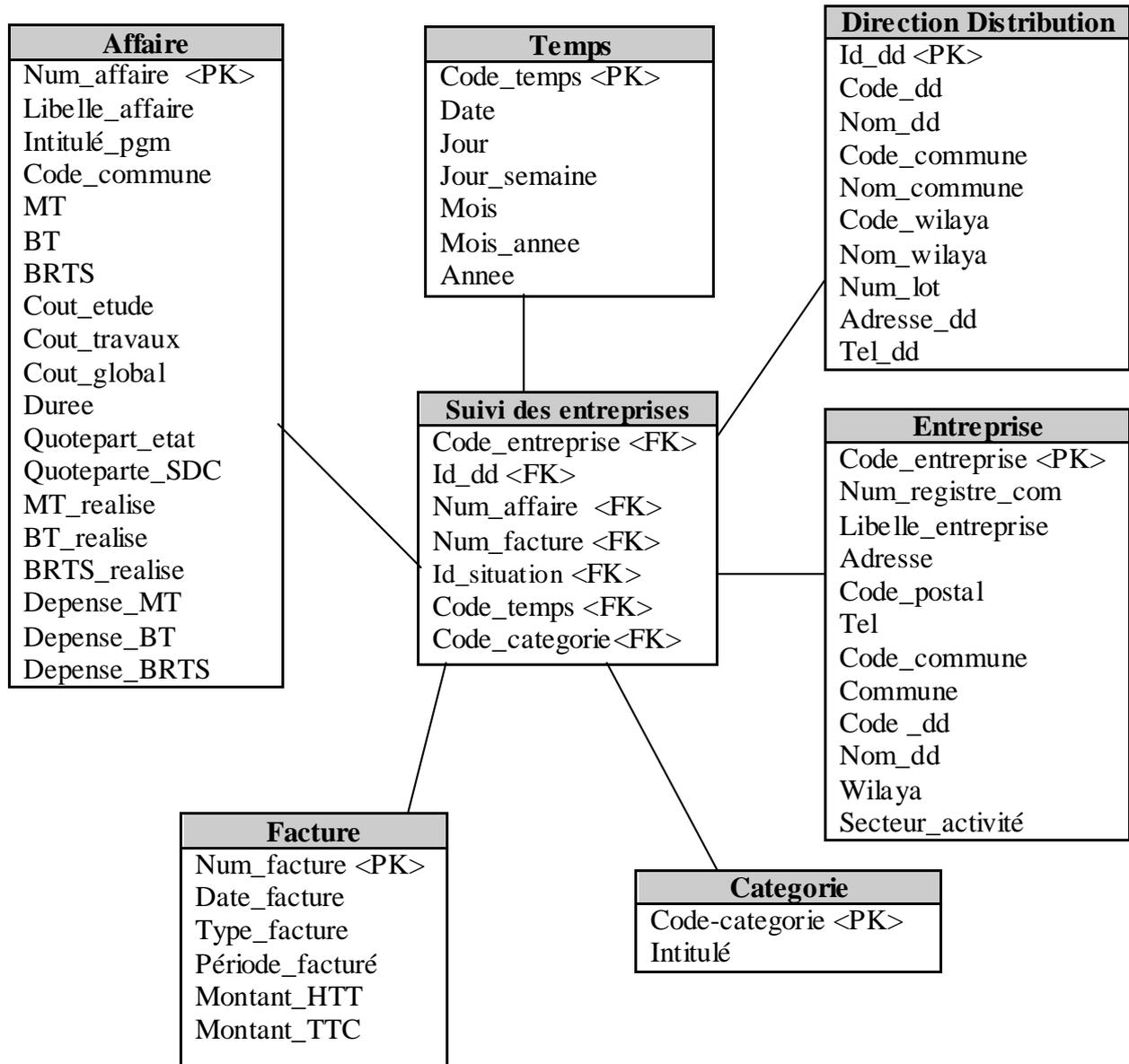


Figure 23 : Modèle en étoile de l'activité « Suivi des entreprises ».

3) Alimentation de l'entrepôt de donnée

L'alimentation d'un entrepôt de données est une phase la plus importante dans son déploiement, et constitue 70% environ d'un projet décisionnel .Cette phase se déroule essentiellement par la migration des données de la base de production vers l'entrepôt de données qui se fait à l'aide des outils ETL (Extracting, Transforming and Loading) où les données sont sélectionnées, nettoyés transformées, ainsi l'entreposage de ces dernières facilitera la prise de décision.

Pour remédier à ça, en premier, nous à été primordial de faire une étude de source de donnée. Cette étape nous a couté beaucoup de temps, car il nous a fallu étudier et analyser plusieurs rapports issus de différent DD, a fin de pouvoir déceler les informations utile à la prise de décision.

Ces informations seront extraites et injectées d'une manière complète dans zone de préparation de données « *Staging area* »¹¹. Ensuite, ils seront soumis à des opérations de transformation comme une préparation au chargement final dans la zone d'entreposage. Après chaque chargement, une mise à jour des Meta Data doit être faite.

Il est à noté que nous avons choisis de planifier un chargement mensuel J-1 pour effectuer des testes et détecter les changements grâce à la comparaison entre le dernier chargement et le chargement actuel.

4) Le cube dimensionnel

Un des principaux points de la Business Intelligence est l'analyse multidimensionnelle. Il devient primordial de pouvoir étudier une même information suivant différents axes. Par exemple, il est très intéressant pour DSI de connaitre l'état d'avancement des affaire par direction distribution, par entreprise par commune, par situation ...c'est pourquoi le moyen pour effectuer ces analyses est d'utiliser des cubes (ou hyper cube).

Les cubes dimensionnels sont présentés comme suivant :

¹¹ Ou seront opérés les différents traitements des données.

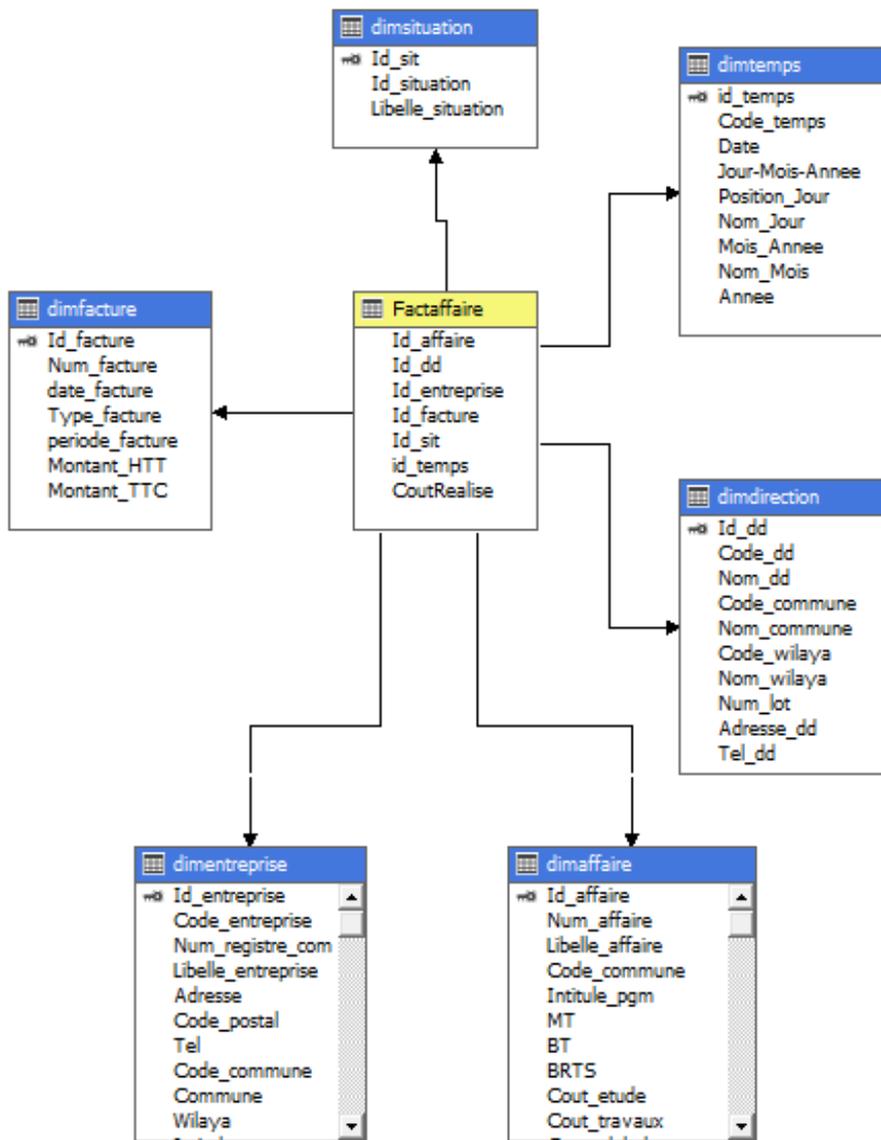


Figure 24 : Cube dimensionnel « Suivi des affaires ».

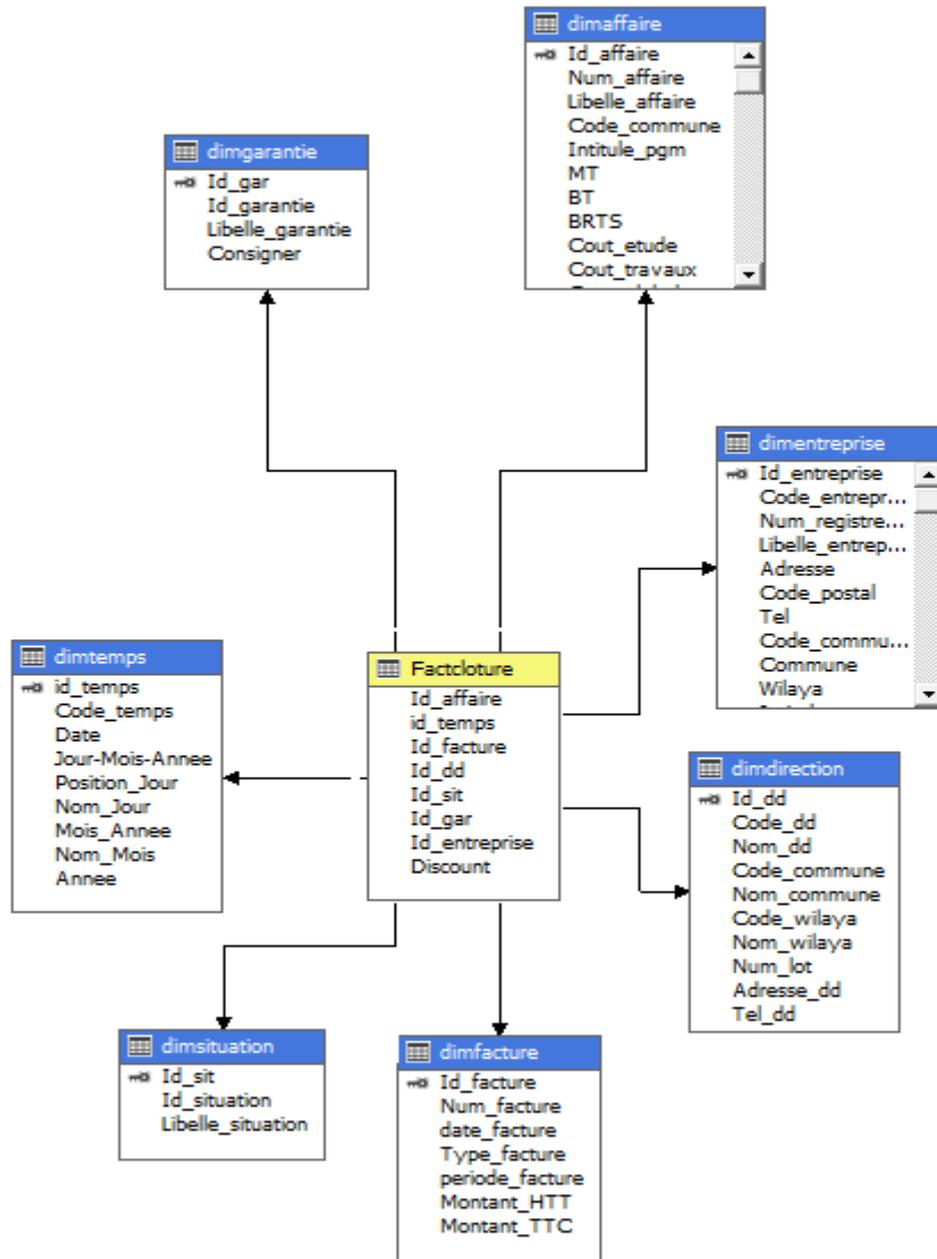


Figure 25 : Cube dimensionnel « Clôture des affaires ».

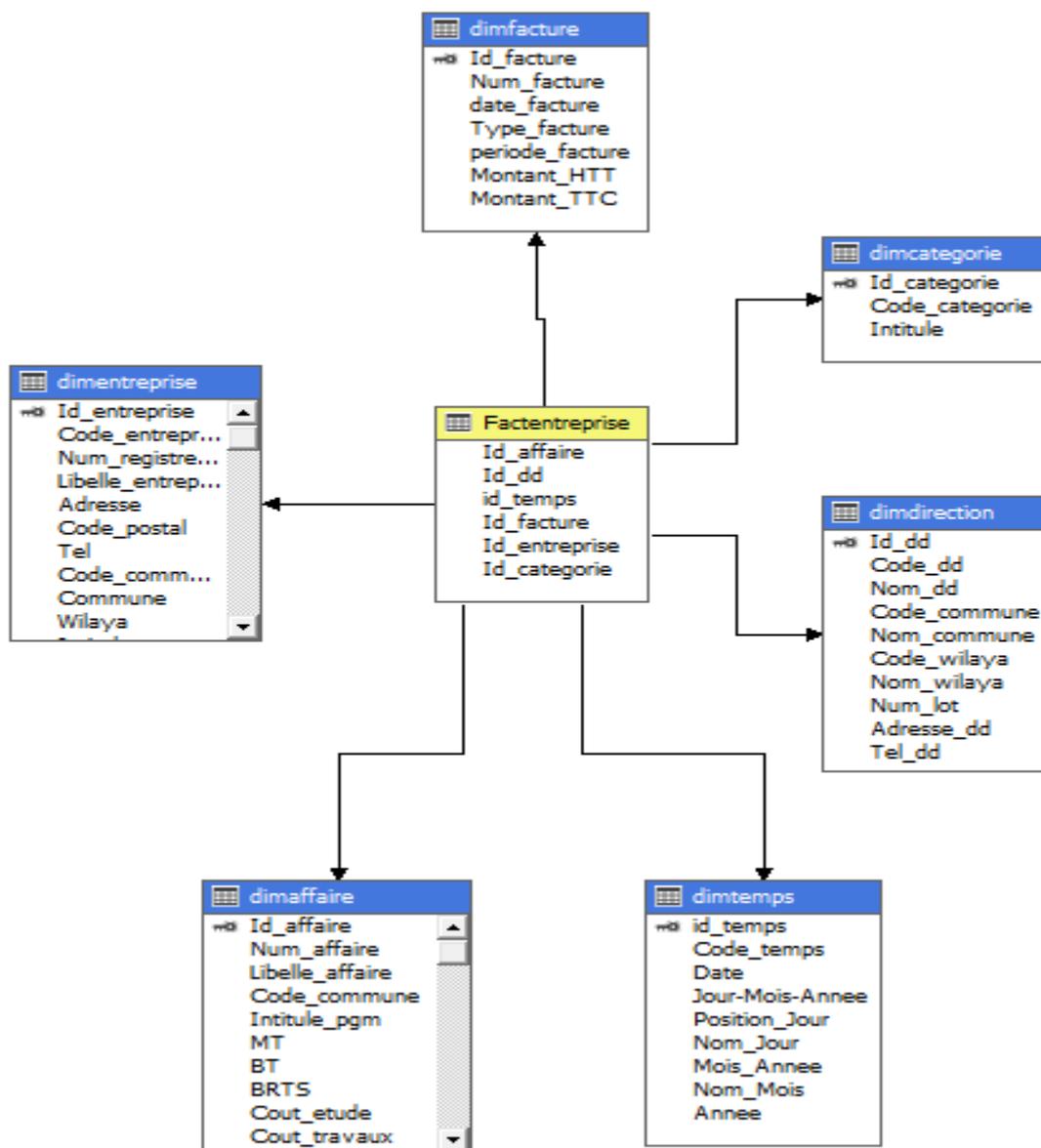


Figure 26 : Cube dimensionnel « Suivi des entreprises ».

5) Conclusion

Se chapitre étale au fond les étapes de conception de l'entrepôt qui se fait par la modélisation dimensionnelle. L'alimentation de Data warehouse se fait grâce à l'outil ETL, qui permet d'extraire, de transformer ensuite de chargé des données cohérente et homogène dans le DW.

Pour faciliter la navigation dans les données et exploiter de manière efficace les informations contenues dans le Data Warehouse, il est primordial de mettre en place les cubes dimensionnels.

La mise en œuvre de notre solution BI sera présenté dans le chapitre qui suit.

Chapitre 5 : Mise en œuvre de la solution

1) Introduction

Pour la réalisation et la mise en place de la solution, il a été nécessaire de recourir à un certain nombre d'outils et mettre en place des environnements d'exécution.

Ce chapitre a pour objectif, décrire l'environnement mis en place par les outils utilisés, l'architecture technique de la solution ensuite la mise en œuvre du système sous forme de prises d'écran avec des descriptions.

2) Outils de développements

Ici nous allons présenter les logiciels utilisés dans le cadre de la mise en place de l'entrepôt de données. Microsoft propose une offre de solutions de Business Intelligence riche et adaptable à de nombreux scénarios, quelles que soient la taille du projet et l'architecture choisie.

Cependant, dans le cadre de notre mémoire, nous avons fait le choix de développer la BI sur SQL Server 2017 qui, à lui présente les composants pour la mise en œuvre de la BI, depuis l'acquisition des données jusqu'à la restitution.

Notre choix de SQL Server s'explique pour les raisons suivantes :

- SQL Server est une plateforme de développement de Business Intelligence assurant la qualité des données avec de nouveaux outils comme Data Quality Services, qui ajouté à SQL Server Intégration Services de plus la gestion des données de référence avec Master Data Services, permettent de gérer la donnée pendant tout son cycle de vie afin de garantir son exactitude.
- SQL Server intègre par défaut des outils de gestion, d'administration et de développement de bases de données.
- Déploiement par un setup, mise en œuvre et administration par des interfaces graphiques intuitives.
- Gestion avancée de la sécurité, l'accès SQL sera avec authentification.
- Prise en compte des spécificités des projets décisionnels.

Comme outil d'analyse et reporting, notre choix est opté pour Microsoft Power BI qui est une plateforme de business intelligence permet d'agréger, d'analyser, de visualiser et de partager des données de façon gratuite via sa version Desktop.

3) Architecture technique de la solution

La figure suivant illustre l'architecture technique de notre solution.

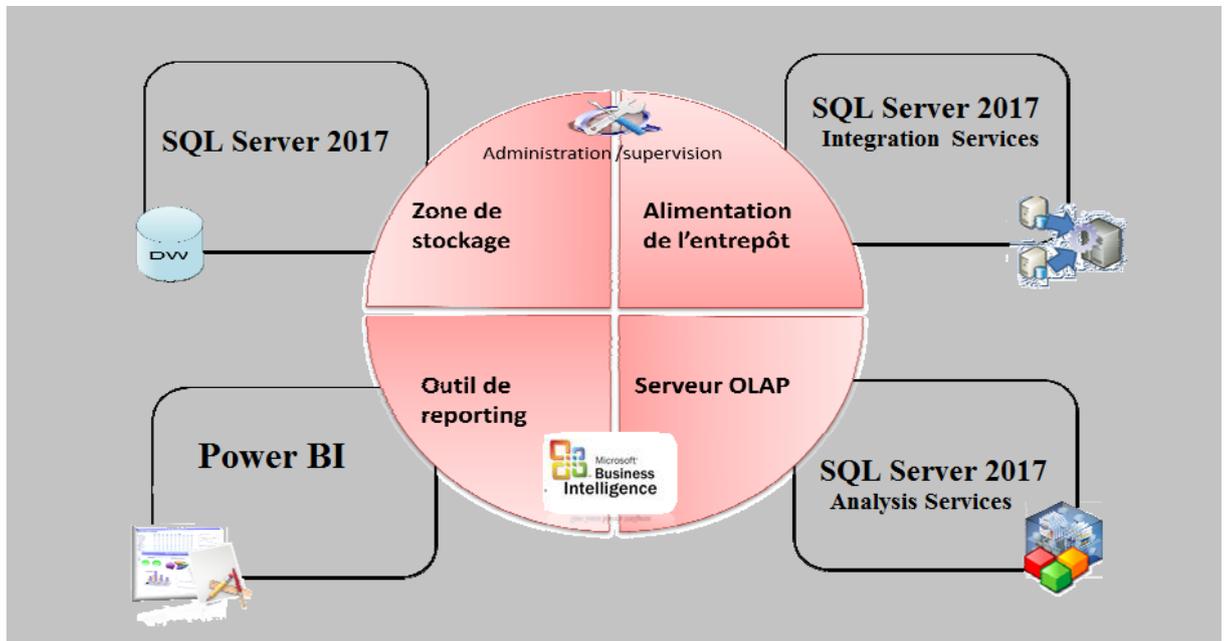


Figure 27 : Architecture technique de la solution

4) Implémentation de la solution

La réalisation de notre entrepôt de données ses déroulées essentiellement selon les étapes suivantes:

4.1) Le projet ETL

L'ETL est créé à partir d'un type de projet SQL Server Intégration Services dans SQL Server Data Tools. Ce dernier prend en charge toute la préparation des données, de leur extraction à leur chargement dans le Data warehouse. Il permet en outre :

- L'extraction des données dans notre cas à partir d'un fichier Excel,
- La mise en œuvre des opérations de transformation, de qualité des données et de mappage,
- Le chargement des données dans la zone d'entreposage,
- La conception et le déploiement et le chargement des cubes OLAP.

Les figures ci-dessous illustrent la configuration effectuée pour réaliser les étapes sus citées via un fichier Excel pour le chargement Dim affaire. Ces étapes se répètent pour tous les autres dimensions.

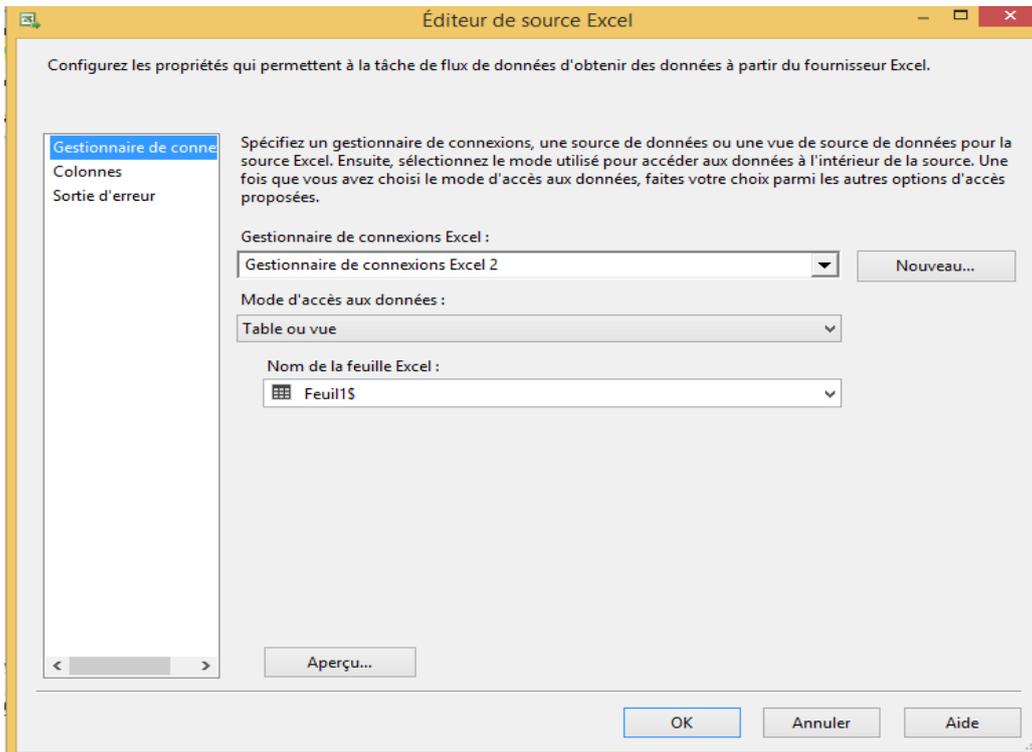


Figure 28 : Extraction des données à partir d'un fichier Excel.

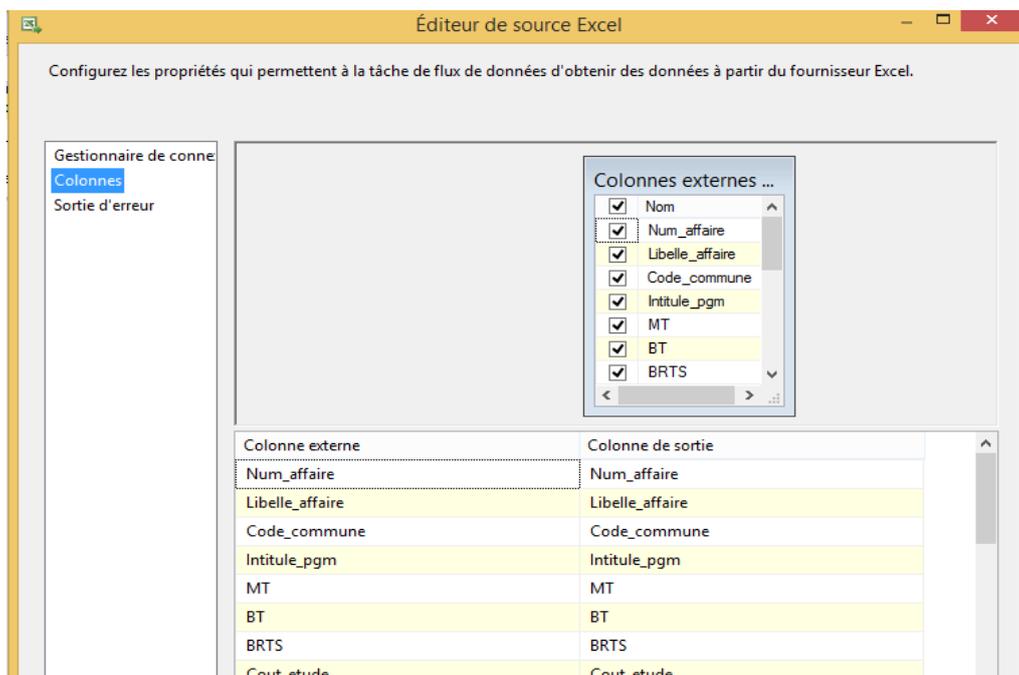


Figure 29 : Vérification des colonnes.

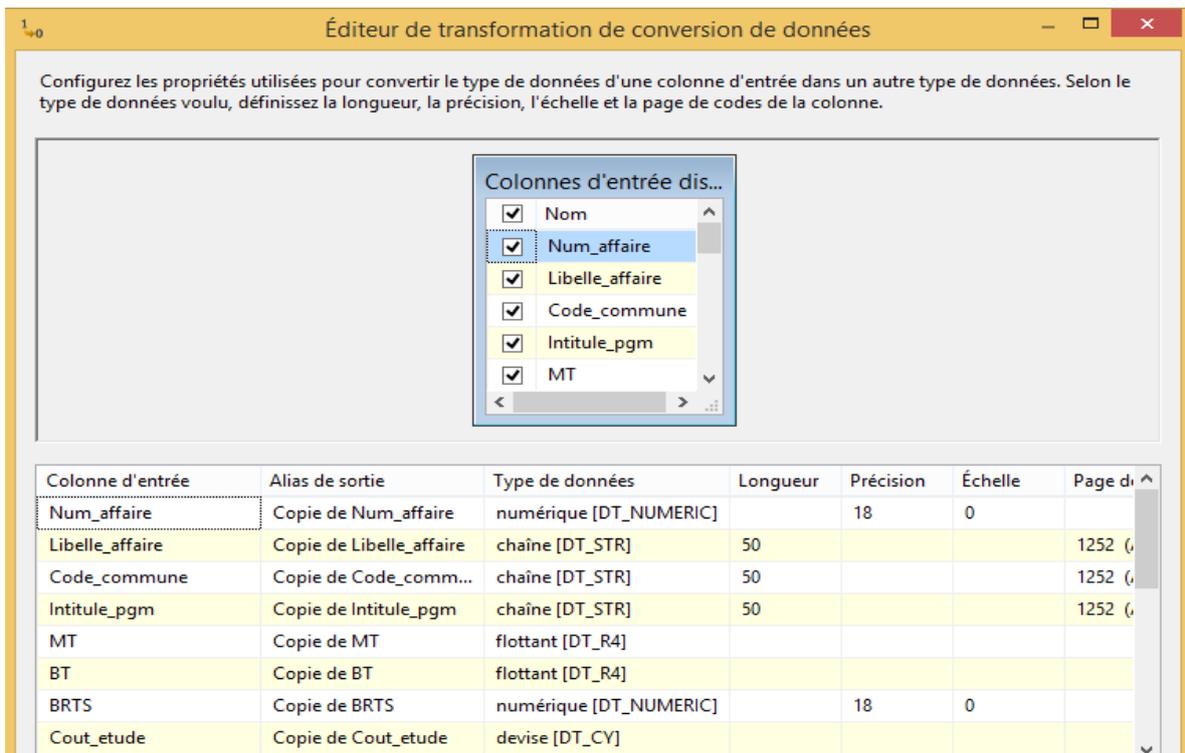


Figure 30 : Transformation et conversion de données.

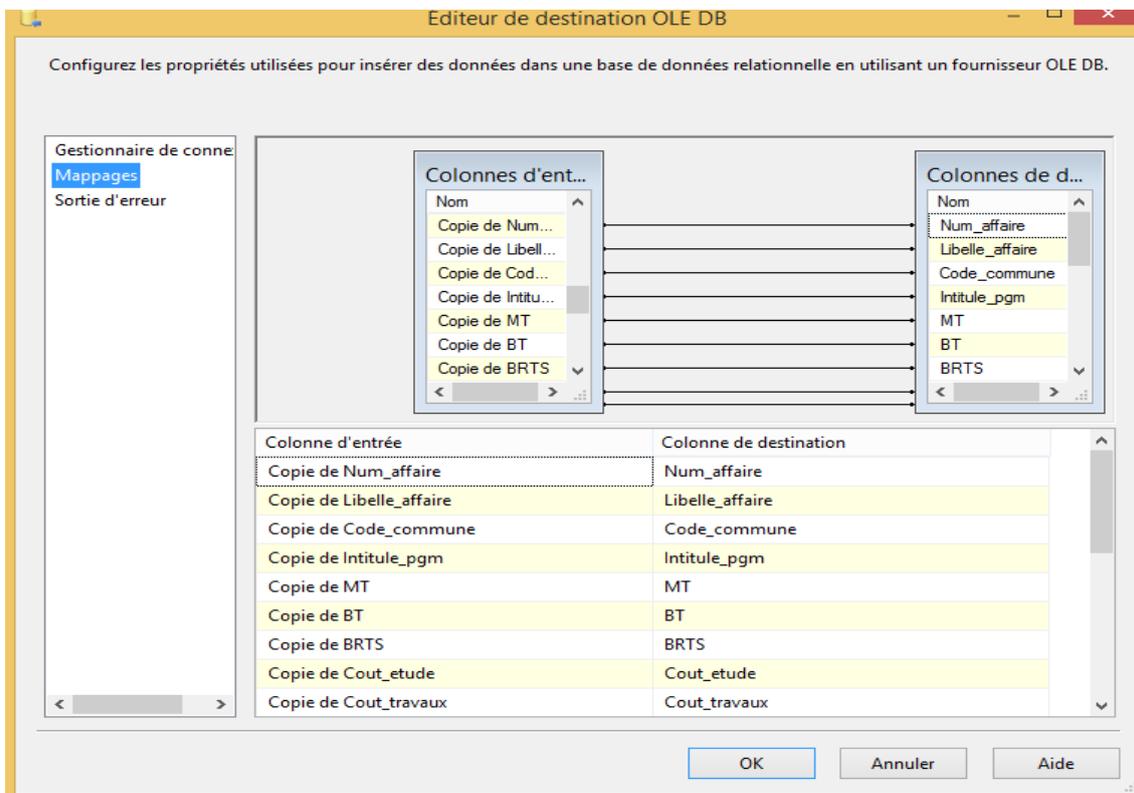


Figure 31 : Vérification le mappages des données.

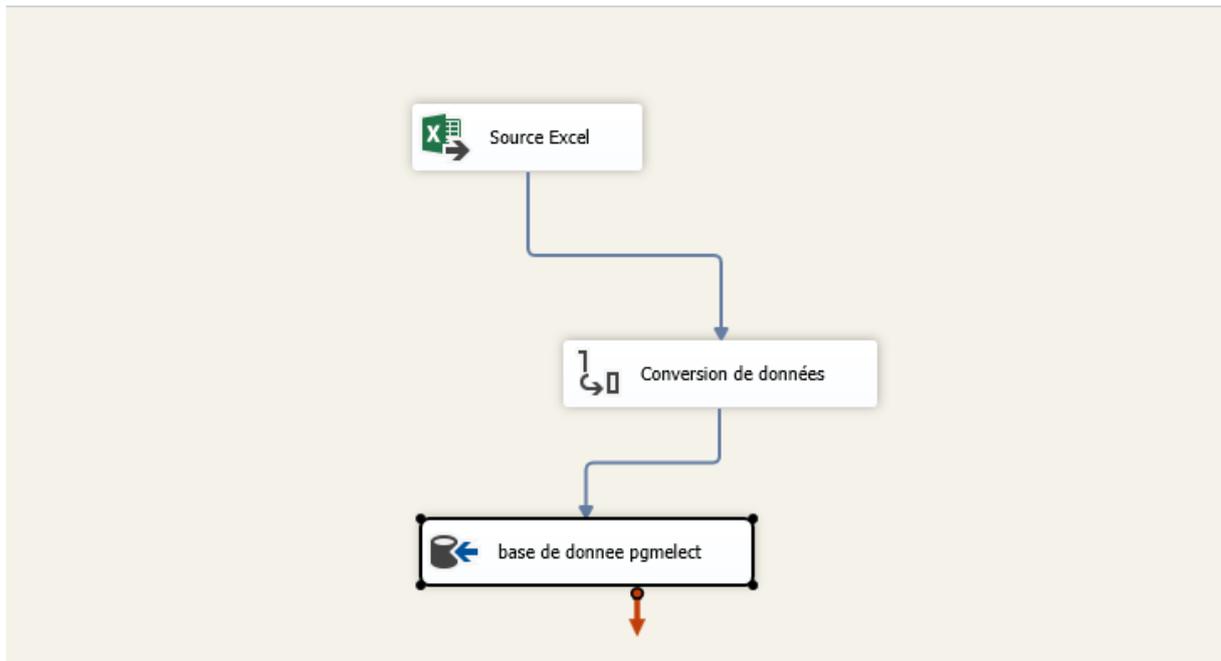


Figure 32 : Chargement de donnée dans table affaire.

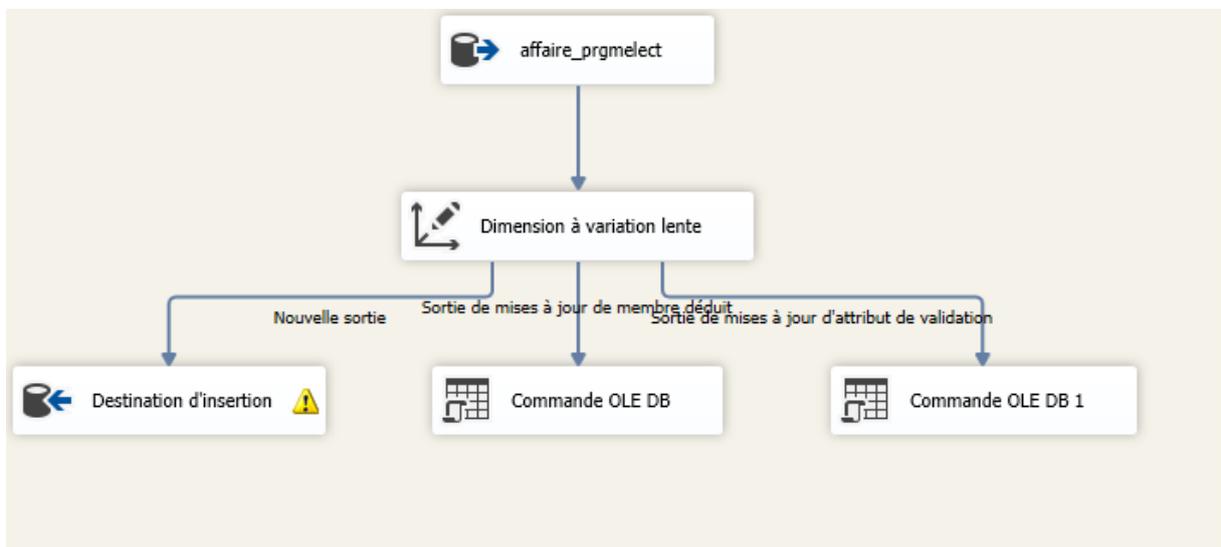


Figure 33 : Créer et charger la dimension Affaire à l'aide d'une dimension à variante lente.

Une fois que tous les tables de dimension sont créés, on peut procéder à la création de la table des faits. Pour effectuer son chargement, il faut identifier l'enregistrement concerné qui constitue la source données puis suivent des composants de recherche en vue de récupérer la clé artificiel de chaque table de dimension.

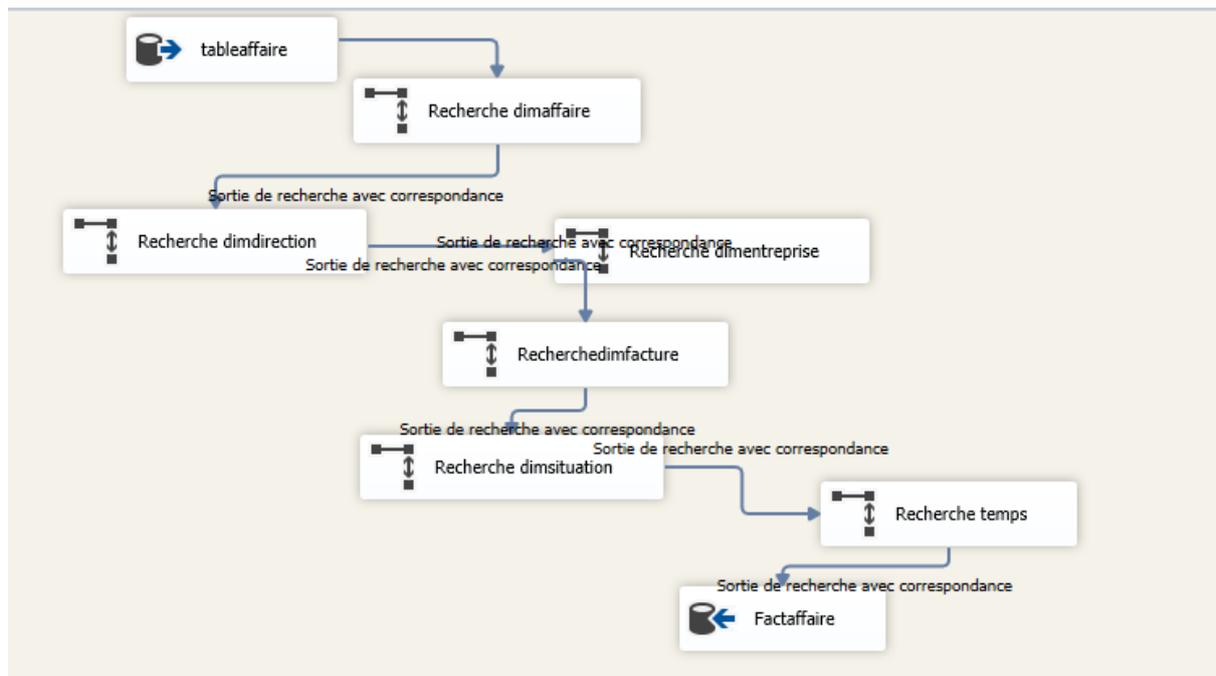


Figure 34 : Création et chargement de la table des faits Factaire.

4.2) **Projet création de Cube :**

L'objectif de ce projet est la création du cube et la base multidimensionnelle qui sera utilisée comme source d'analyse. C'est un projet Analysis Services qui est créé dans le SQL Server Data Tools. Pour cette réalisation nous avons :

- Créé la source de données DWH.ds qui fait référence au schéma relationnel de l'entrepôt créé au niveau du projet ETL,
- Créé la vue de source de données DWH.dsv qui est une couche logique relationnelle entre le moteur OLAP et les données sources,
- Créé les dimensions du cube en utilisant un Assistant Cube en définissant une vue de source de donnée et les tables intervenant dans sa création.

Une fois le cube créé, nous passons au déploiement du projet Analysis Service ou une base Analysis Services sera générée dans le moteur Analysis services de SQL Server.

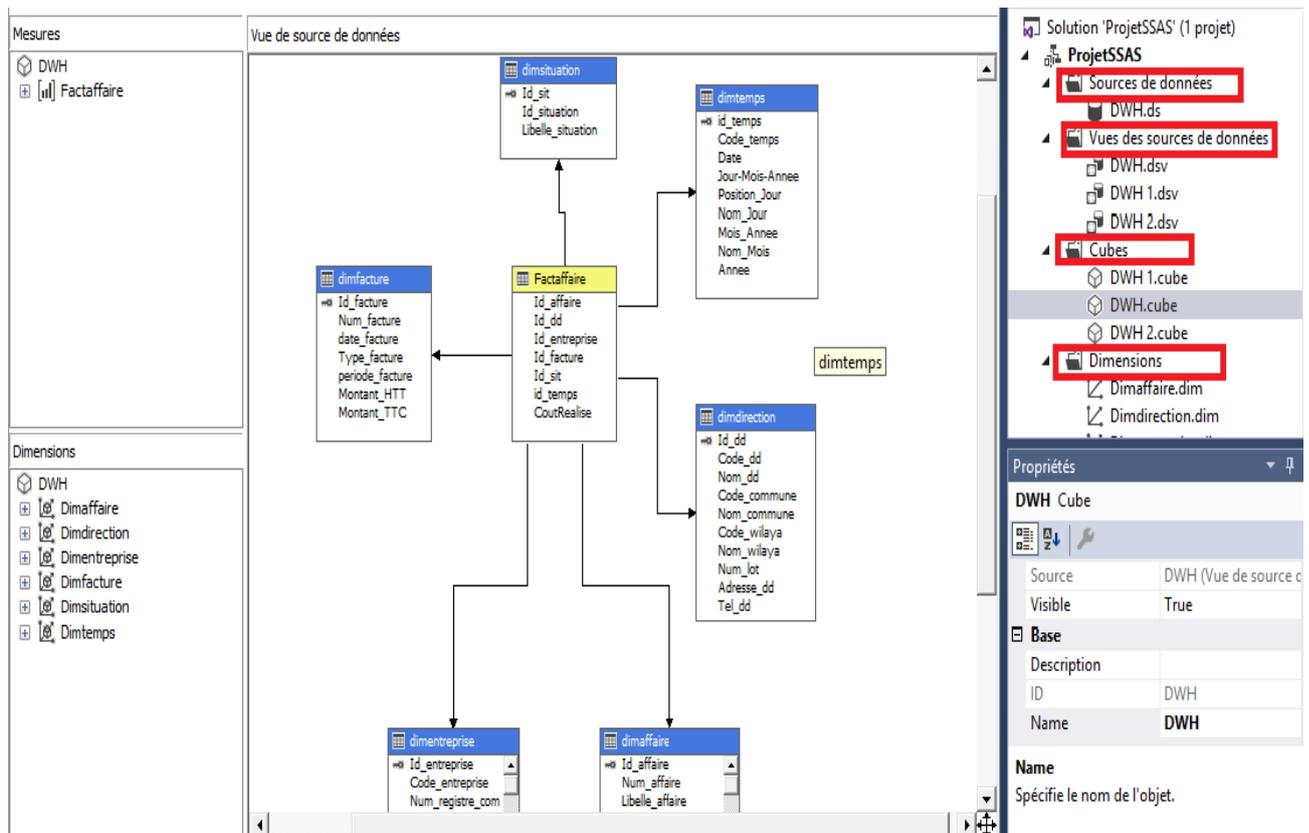


Figure 35 : Les éléments constitutifs de la base multidimensionnelle.

4.3) Mise en œuvre de la présentation des données

L'entrepôt de données permet aux décideurs d'avoir accès plus rapidement et plus facilement aux données stratégiques de l'entreprise. La présentation de ces données a été construite à l'aide de l'outil Microsoft Power BI.

Power BI est basé sur une architecture simple et intuitive. Elle offre une multitude de possibilités d'analyse et de présentation des données, permettant aux utilisateurs d'accéder à un ensemble de fonctionnalités natives à fin de créer leurs propres rapports et tableaux de bord.

Pour l'exploitation du cube, On a crée dans un premier temps une connexion entre Power BI desktop et moteur Analysis Services de SQL server pour récupérer la base de donnée de notre projet. Une fois la connexion établie, nous procédons à sélectionner les tables de la base de données ainsi que faire des modifications concernant les colonnes des tables si nécessaire.

Fichier Accueil Aide Outils de table

Nom DWH

Marquer comme table de dates
Gérer les relations
Nouvelle mesure rapide
Nouvelle mesure colonne
Nouvelle table

Structure Calendriers Relations Calculs

Champs

Rechercher

DWH

Adresse
Année
BRTS Réalise
BT Réalise
Code Commune
Commune
Cout Etude
Cout Global
Cout Réalise
Cout Travaux
Date
Date Facture
Depense BRTS
Depense BT

BRTS Réalise	BT Réalise	Code Commune	Cout Etude	Cout Global	Cout Travaux	Depense BRTS	Depense BT	Depense MT	Duree	Id Affaire	Intitule Pgm	Libelle Affair
10	2.365	0301	107207	3.0504332E+7	3.0397124E+7	100000	350000	15000	3 ans	3	QLS	VSA Benbadis
10	2.365	0301	10/20/	3.0504332E+7	3.0397124E+7	100000	350000	15000	3 ans	3	QLS	VSA Benbadis
23	20.899	0304	952800	3.2418493E+8	3.2323213E+8	320900	950800	1.050065E+7	3 ans	9	ER	HAY El menass
23	20.899	0304	952800	3.2418493E+8	3.2323213E+8	320900	950800	1.050065E+7	3 ans	9	ER	HAY El menass
36	7.2350001	0302	185749	5.503516E+7	5.4849408E+7	540500	720050	250000	3 ans	5	QLS	MALKA
36	7.2350001	0302	185/49	5.503516E+7	5.4849408E+7	540500	720050	250000	3 ans	5	QLS	MALKA
48	5.6350002	0304	187804	5.5363584E+7	5.517578E+7	500000	650500	520000	2 ans	7	ER	LOGTS Bureau
48	5.6350002	0304	187804	5.5363584E+7	5.517578E+7	500000	650500	520000	2 ans	7	ER	LOGTS Bureau
5	4.256	0305	186050	6.850276E+7	6.8316712E+7	45600	430000	3506000	3 ans	11	ER	Oued beddou
6	2.6259999	0301	109558	3.0640168E+7	3.053061E+7	50000	28000	20000	2 ans	1	ER	Amir AEK
6	2.6259999	0301	109558	3.0640168E+7	3.053061E+7	50000	28000	20000	2 ans	1	ER	Amir AEK

Figure 36 : Résultats d'importation de base de données.

Afin d'obtenir une information d'une manière simple à partir de notre cube, on a procédé à la création d'un tableau de bord qui peut faire l'analyse et le reporting en utilisant le référentiel Power BI.

La figure suivante représente une vue d'ensemble du tableau bord mis en place.

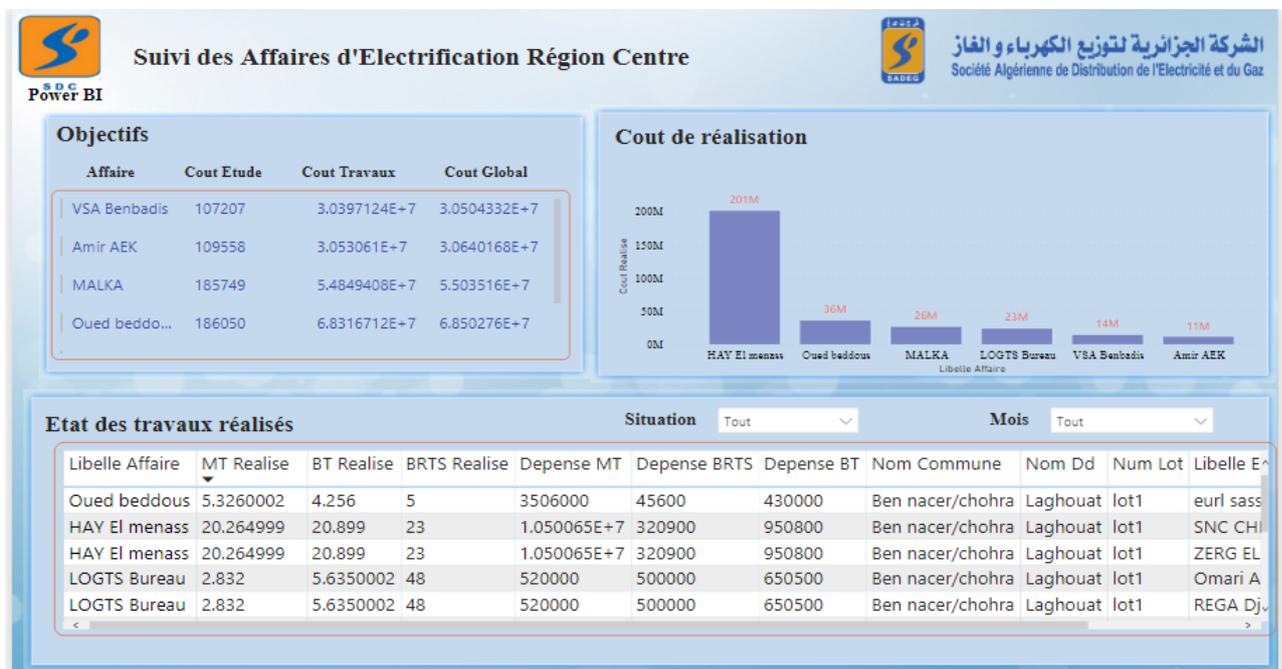


Figure 37 : Tableau de bord réalisé par Power BI.

5) Conclusion

Dans ce dernier chapitre, nous avons présenté les étapes liées à mise en œuvre notre projet BI en y joignant des captures d'écran de sa réalisation avec une description, tout en définissant l'architecture technique ainsi que les outils open source utilisés qui ont contribué à sa réalisation.

Conclusion générale et perspectives

Conclusion générale et Perspectives

La bonne information, à la bonne personne et au bon moment, ainsi est le credo du Business Intelligence.

Exploiter les données à disposition de l'entreprise afin de leur donner de la valeur ajoutée, est le défi des entreprises modernes.

Dans ce cadre, et afin de palier à des problèmes récurrents dans le processus de prise de décision, RDC à travers son département suivi et inspection a initié le projet de mettre en place un système décisionnel « BI » fiable et efficace basé sur Data warehouse.

Tout au long de notre travail de conception et de réalisation, nous avons essayé de suivre une démarche proposée par «R. Kimball » qui se focalise sur les besoins des utilisateurs. Cette démarche a été effectuée afin de collecter les besoins auprès des utilisateurs finaux et ce grâce aux techniques d'entretien et aux études des besoins déjà exprimés sous forme de rapports qui ont une grande utilité pour définir de manière claire le périmètre et les besoins réels des utilisateurs. Cette étude a fait ressortir les sujets d'analyse dignes d'intérêt suivants : suivi des affaires, suivi sa clôture et suivi des entreprises.

Dans un deuxième temps, la modélisation de la zone de stockage des données s'est faite grâce aux principes de la modélisation dimensionnelle. Cette modélisation offre une vision claire et une compréhension intuitive des modèles proposés. Nous avons de ce fait proposé des modèles en étoiles des trois activités recensés.

La partie d'alimentation de la zone de stockage « implémentation physique des modèles dimensionnels basé sur le modèle relationnel » a été sans nul doute la partie du projet la plus consommatrice en temps .Cette étape nous a permis de concevoir et de réaliser, grâce à des outils open source, les routines d'extraction, transformation et chargement des données.

L'utilisation d'outils et de solutions open source est un volet très important dans ce projet et c'est décidé de les utilisées depuis le début. Cette orientation, qui se fait ressentir durant les étapes sus citées, est plus présente dans la partie mise en œuvre de présentation des données grâce à l'intégration d'une plateforme « BI » offrant à l'utilisateur la possibilité d'exploiter les données de l'entrepôt pour des fins d'analyse et Reporting via le Tableau de bord.

Pour finir, et avant de citer les perspectives du projet, nous pouvons dire que ce stage au niveau de SONELGAZ nous a permis d'enrichir notre expérience dans le domaine professionnel et d'acquérir de nouvelles connaissances dans un domaine qui nous était totalement inconnu à savoir le domaine des systèmes décisionnels.

Comme un projet informatique n'est jamais complètement terminé, nous pouvons citer les perspectives et développements suivants :

- Intégrer d'autres Data Mart qui se grefferont à l'entrepôt de données afin d'assurer sa complétude.
- Suivre le déploiement actuel et recueillir les correctifs et remarques des utilisateurs.
- Continuer le développement du portail de restitution.

Bibliographie

- [1] Bernard LIAUTAUD, & Mark HAMMOND, « e-Business Intelligence, ». Maxima, 2001.
- [2] Gloria J. MILLER, & Dagmar BRÄUTIGAM, & Stefanie V. GERLACH, « Business Intelligence Competency Centers, A Team Approach to maximizing competitive advantage, » 1st edition. Hoboken, New Jersey, John Wiley & Sons, Inc., 2006.
- [3] Katarina URKO, & Mirjana PEJIC BACH, « Business Intelligence & Business Process Management in Banking Operation, », Thèse de Ph.D, University of Zagreb, Faculty of Economics, Department of Business Computing, Cavtat, Croatie, 2007.
- [4] Le Moigne J.L., « La théorie du système général, théorie de la modélisation », P.U.F., 1977.
- [5] H. Dresner ; « BI : Making the Data Make Sens » ; Gartner Group 2001.
- [6] J.F. Goglin; « La Construction du Datawarehouse : du Datamart au Dataweb »; Hermes 1998.
- [7] Valentin Pauline Introduction à la BI avec SQL Server 2008 par DotNet France Association
- [8] Abdenour Bouzghoub ; « Modélisation des Entrepôts de données XML : Application au domaine de la sécurité sociale » ; Thèse de Magistère Option : SISCSO ; Institut National de Formation en Informatique (I.N.I) 2008.
- [9] W. H. Inmon. Building the Data Warehouse», ed. Wiley 1ère édition : 1996,3ème édition :2002.
- [10] Y. Belhouchet,S.Hamid ; « Mise en place d'une Solution Data Warehouse pour le contrôle de gestion au sein de NATIXIS ALGERIE » ;Mémoire de Master option SI ;Ecole National d'informatique(ESI)2012.
- [11] M. TRINIDAD, & S. ENCINAS, « Entrepôts de données pour l'aide à la décision médicale : conception et expérimentation, », Thèse de Doctorat, Université Joseph Fourier, Grenoble, France, 2005.
- [12] Bernard LIAUTAUD, & Mark HAMMOND,« e-Business Intelligence, ». Maxima, 2001.
- [13] Ralph KIMBALL, Laura REEVES, Margy ROSS, & Warren THORNTHWAITE, « Le Data Warehouse, Guide de conduite de projet, », Paris, France, Eyrolles, 2005.

- [14] Ralph KIMBALL & Margy ROSS, « The Data Warehouse Toolkit. The Complete Guide to Dimensional Modeling, », 2nd ed, New York, USA: Robert Elliott, 2002.
- [15] Ralph KIMBALL, « Entrepôts de données: Guide pratique du concepteur de Data warehouse, » 1st ed. Paris, France, Hervé Soulard, 1997.
- [16] William H. INMON, « Building the Data Warehouse, », 4th ed, Indianapolis, Wiley Publishing Inc., 2005.
- [17] Jean-Michel FRANCO, & EDS-Institut Prométhéus, « Le Data Warehouse, Le Data Mining, ». Paris, France, Eyrolles, 1997
- [18] Paulraj PONNIAH, « Data Warehousing Fundamentals. A Comprehensive guide for IT professionals, », A Wiley-Interscience publication, John Wiley & SONS, Inc., New York, USA, John Wiley & Sons, Inc., 2001.
- [19] Lamri CHOUDER, « Entrepôt Distribué de Données », Thèse de Magistère Option SI, Institut National de Formation en Informatique (I.N.I), Alger, Algérie, 2007.
- [20] William H. INMON, Derek STRAUSS, & Genia NEUSHLOSS, « DW2.0, The Architecture for the Next Generation of Data Warehousing, », 1st ed. San Francisco, CA, USA, Morgan Kaufmann, 2008.
- [21] Chris Todman, « Designing a Data Warehouse: Supporting Customer Relationship Management, », 1st ed, Prentice Hall, 2000.
- [22] Vincent RAINARDI, « Building a Data Warehouse With Examples in SQL Server, », 1st ed, USA, Vincent RAINARDI, 2008.
- [23] Nakache, « Data Warehouse et Data mining », Mémoire présenté au CNAM (Conservatoire National des Arts et Métiers) Lille, 1998.
- [24] William H. Inmon « *DW 2.0 The architecture for the next generation of data warehousing* », Elsevier Press, 2008.
- [25] E. Grislin-Le Strugeon, « *Systèmes d'information décisionnel (Data Warehouse / Data Mining)* », Université de Valenciennes, ISTV, 2000.
- [26] Henri Bouquin, « *Le contrôle de gestion* », 3rd Edition, P.U.F, 2003.
- [27] A.Filali et S.kedjnane, « *Conception et réalisation d'un Data Warehouse pour la mise en place d'un système décisionnel* », 2010.
- [28] R. Kimball, M. Ross et W. Thornthwaite, « *Le Data Warehouse – Guide de conduite de projet* », 3rd Edition, Eyrolles, 2007.