



République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Mouloud Mammeri de Tizi-Ouzou
Faculté de Génie Electrique et d'Informatique
Département d'Informatique



Mémoire de fin d'étude

En vue de l'obtention du diplôme de master

Académique en informatique

Option : ingénierie des systèmes d'information

Thème

Mise en place d'une solution de business intelligence

Cas : société de vente de matériels informatique

Réalisé par:

HADJI Souad.
Ramdani Nassima.

Dirigé par :

Mme. Amirouche Fatiha.

Promotion 2019/2020

Remerciements

Nous tenons à témoigner notre reconnaissance à notre Dieu le tout puissant, de nous avoir donné le courage et la force de mener à terminer ce projet.

*Nous présentons nos remerciements à notre encadreur, Madame **Amirouche Fatiha**, d'avoir accepté de nous encadrer tout au long de ce travail.*

et ses remarques qui nous ont servi pour la réalisation du projet.

Nous tenons à exprimer notre gratitude et nos remerciements au membre de jury ;

Pour l'honneur qu'ils nous font en jugeant ce modeste travail.

Nous exprimant nos chaleureux remerciements à tous ceux qui ont contribué de près ou de

Loin à la réussite de ce travail.

Notre immense reconnaissance va à ceux qui ont plus particulièrement assuré le soutien

affectif de ce travail : nos famille



Dédicaces

On dédie ce modeste travail

A nos chers parents

*Pour leur soutien, leur patience, leur sacrifice,
vous qui avez fait de nos ce que nous somme maintenant.*

*On espère être l'image que vous avez fait de
nous, que Dieu vous garde et vous bénisse.*

Que ce travail soit pour vous le gage de notre profond amour

A tous nos frères et sœurs

A tous nos amis et mon binôme

A tous ceux qui nous ont aidés

Souad et Nassima



Résumé

Dans un cadre pédagogique nous avons créé une entreprise de vente de matérielles informatiques. Pour assurer plusieurs missions dans le domaine commerciale. Pour lui donner une image proche de la réalité. On lui a créé une application de vente, et une base de données. Comme toute entreprise commerciale, on a supposé que son application génère chaque jour une masse de données, qui n'est jamais exploitée, et ces masses de données sont toujours en croissance. Qui peuvent fournir une mine d'information sur l'évolution du chiffre d'affaires, l'évolution des opérations de vente, ainsi que le comportement de ces clients (dans le but de satisfaire leurs besoins Etc.

À ce titre, notre entreprise souhaite l'exploitation et l'analyse de ces données, dans l'objectif d'aide à la décision. Par contre l'élaboration des rapports d'analyse est manuelle et connaît des lenteurs qui n'arrangent pas les décideurs. Ces derniers ont besoin d'informations fiables dans des délais raisonnables.

Dans le monde technologique, on trouve beaucoup d'outils et de logiciels qui assurent ces objectifs, et qui facilitent l'exploitation et l'analyse de ces masses de données. Cela s'effectue à l'aide des outils de reporting, d'analyse OLAP (dite multidimensionnelle), le datamining ... Etc

Pour mettre en place notre solution, un projet est lancé qui s'intitule sur un système d'information d'aide à la décision « BI ». Ce système sera construit autour d'une base de données décisionnel dite « **Data Warehouse** » et répondant à tout les besoins d'analyse du groupe dans sa fonction de distribution.

Mots clés :

Système d'information d'aide à la décision, Indicateurs de performances, Reporting , Data Warehouse « Entrepôt de données », Business Intelligence « B.I », intégration de données, olap, datamining

Table des matières

Remerciements.....	ii
Dédicaces	iii
Résumer.....	iv
Table des matières.....	v
Liste des figures :	ix
Liste des tableaux :	xi
Liste des abréviations :	xii
Introduction générale	1
I. Introduction :.....	1
II. Problématique	1
III. Solution.....	2
IV. Objectifs et le périmètre du projet BI :	2
V. Structure du rapport :.....	3
chapitre I : Les systèmes d'information décisionnelle	4
I. Introduction :.....	5
II. Présentation de la Business Intelligence :	5
II.1. Historique de la BI :.....	5
II.2. Définition de la BI :	6
II.3. Les domaines d'utilisation de la BI :	6
II.4. L'architecture du processus de la BI et ces principales phases :	6
II.4.1. Les principes phases :	6
II.4.2. Le fonctionnement de ces phases.....	8
II.4.3. Les outils utilisés dans ces différentes phases :.....	7
II.4.3.1. Outils de collecte :	7
II.4.3.2. Outils de stockage :.....	8
1.Le Datamart :.....	8
2.LesDW	8
3.Les 4 caractéristiques d'un DW	8
4.Types du DW :.....	9
5. Modélisation du DW.....	9
6. Les modélisations dimensionnelles.....	10
7. Construction d'un Data Warehouse:.....	11
II.4.4. Restitution ou exploitation :	12

II.4.4.1. Les outils de type OLAP :	12
II.4.4.1.1. Représentation du cube de données :	12
II.4.4.1.2. L'architecture ROLAP	13
II.4.4.1.3. L'architecture MOLAP :	13
II.4.4.1.4. L'architecture HOLAP (Hybride OLAP)	14
II.4.4.2. Le Reporting	14
II.4.4.3. Les tableaux de bord	14
II.4.4.4. Le Datamining :	14
II.5. Les principaux avantages de la BI pour l'entreprise :	15
III. Informatique décisionnelle vs Informatique opérationnelle :	16
III.1. Informatique décisionnelle :	16
III.2. Informatique opérationnelle :	16
III.3. Décisionnel vs Opérationnel :	16
IV. Les limites de la BI :	18
IV.1. Les limites et les inconvénients de la BI :	18
IV.2. L'évolution de la BI :	18
IV.3. L'avenir de la BI :	19
Conclusion	20
Chapitre II: Spécification et analyse des besoins.....	22
I. Introduction :	22
II. Description de l'entreprise :	22
II.1. Définition:	22
II.2. L'environnement de l'entreprise	22
II.3. Fonctionnalités de l'entreprise	22
II.4. Identification du rôle des employés et des clients	23
II.5. Diagramme de classes :	24
III. Stratégie suivie :	25
III.1. Méthode de collecte d'informations	25
III.2. Analyse des besoins.....	25
III.3. Diagramme des cas d'utilisation	26
IV. Spécification des besoins	27
IV.1. Détermination des besoins fonctionnels	27
IV.2. Détermination des besoins techniques.....	28
IV.3. Détermination des besoins non fonctionnels	29
V. description de la solution :	30

V.1. Description générale de la solution basée sur un entrepôt de données :	29
V.2. Critique de la solution :	30
V.3. Structure détaillée de la solution	30
Conclusion.....	32
Chapitre III : Conception de la solution.....	34
A. Conception du DW :	34
I. Choix de l'approche de modélisation.....	34
II. Définition des activités à étudié	34
III. Choix d'Indicateurs :	35
IV. Schémas de modalisation	36
IV-1. La dimension date.....	37
IV-2. Construire les datamarts (les étoiles)	37
IV-2-1. Le suivi des produits :	37
IV-2-2. Le suivi commercial des ventes :	38
IV-2-3. le suivi des achats :.....	40
IV-2-4. le suivi du magasin :.....	40
IV-3 Regroupement en data warehouse.	43
B. Alimentation du DW :	43
I.Etapes de phase d'alimentation	43
I.1. Découverte des données :	43
I.2. Extraction des données :.....	43
I.3. Transformation des données :.....	43
I.4. Chargement des données :	44
C. Construction de cube OLAP :	44
I.Démarche d'utilisation d'OLAP.....	45
I.1. Operation de forage « roll-up » & « drill-down »).....	45
I.2. Opération de rotation.....	45
I.3. Autres Opération.....	45
Conclusion.....	46
Chapitre IV : Réalisation et mise en œuvre de la solution	48
A. Présentation des outils utilisés:	48
I. L'architecture et la construction de la solution :.....	48
II. Choix des outils (logiciels) :	48
II.1. Choix du SGBD :.....	48
II.2. les serveurs d'exécution	50

II.3. les langages de programmation	50
II.4. les outils de collecte	52
II.4.1. Talend :	51
II.4.2. Pentaho :	52
II.4.3. étude comparative des outils de collecte	52
III. Outils de restitution	53
III.1. Power BI	54
III.2. étude comparative des outils de restitutions.....	54
B. Mise en œuvre de la solution	55
I. Présentation de l'application de vente	53
II. Mise en œuvre de la solution décisionnelle	60
1. Business Models :	59
2. Etape de construction de la zone d'entreposage :	60
a. Chargement et mises à jour de toutes les tables de dimension reliées à la table de	60
b. Alimentation de la table de fait :	60
3. Construction et alimentation	60
3.1. Création des métadonnées	60
3.1.1. Outils_infos	60
3.1.2. La métadonnée staging_area	60
3.2. Création des jobs	60
3.2.1. Les composants de Talend.....	60
3.2.2. Création des jobs pour alimenter la Staging_area	60
3.2.2.1. Transformatin à l'aide des jobs	60
3.2.2.2. Alimentation du staging area	71
3.2.3. Création des jobs pour alimenter le DW	72
4. Construction des cubes multidimensionnelle et analyse OLAP.....	74
III. Construction de la zone de restitution :	74
conclusion :	77
Conclusion général.....	78

Liste des figures :

Figure 1 : historique d'évolution de la data.....	5
Figure 2 : Le processus ETL	7
Figure 3 : présentation du DW et data-mart	8
Figure 4: représentation de table de fait et table de dimension	9
Figure 5 : Modélisation en étoile	10
Figure 6 : Modélisation en flocons	10
Figure 7 : La modélisation en constellation	11
Figure 8 : Cube de données.....	12
Figure 9 : architecture du système OLAP.	13
Figure 10 : L'architecture du processus de la BI et ces phases	15
Figure 11: la place du décisionnel et d'opérationnel dans l'entreprise.	16
Figure 12 : diagramme de classe de notre BDD	24
Figure 13 : Diagramme (CU) des fonctionnalités générales de l'entreprise.	26
Figure 14 : Diagramme (CU) d'analyse commercial	27
Figure 15 : Diagramme (CU) d'analyse des produits	28
Figure 16 : architecture de la solution.	32
Figure 17 : architecture de l'approche de modélisation.....	34
Figure 18 : diagramme de priorité d'analyse	35
Figure 19 : schéma en étoile fact_produit.	38
Figure 20 : schéma en étoile fact_vente.....	39
Figure 21 : schéma en étoile fact_achat.....	40
Figure 22 : schéma en étoile fact_magasin.....	41
Figure 23 : schéma du DW en flocon	42
Figure 24 : phase de processus ETL	44
Figure 25: les logiciels choisis pour la solution décisionnelle.....	51
Figure 26: les logiciels choisis pour le reporting.....	54
Figure 27 : page d'accueil de note application de vente.....	56
Figure 28 : page pour choisir le produit.	56
Figure 29 : panier de l'utilisateur avec les produits choisis.....	57

Figure 30 : confirmer la demande.	57
Figure 31 : page d’inscription.....	58
Figure 32 : page de connexion.	58
Figure 33 : page de contact.....	59
Figure 34: présentation du business models.	59
Figure 35 : les métadonnées créés pour la réalisation de la solution.	61
Figure 36: fenêtre de la connexion à la bdd.	61
Figure 37 : fenêtre de récupération du schéma.	62
Figure 38 : fenêtre du schéma des tables	62
Figure 39 : fenêtre de la connexion au fichier Excel.	63
Figure 40 : présentation des jobs créés.	64
Figure 41 : les contextes	64
Figure 42 : job de transformation d’achat et stock.	67
Figure 43 : job pour faire une répartition des clients selon la facturation.	67
Figure 44 : job d'extraction du fait_produit et les tables relié à elle.	68
Figure 45 : job d'extraction et transformation table vente.....	68
Figure 46 : job de répartition des clients.	69
Figure 47 : Exemple de transformation des dimensions pour la table fait "vente"	69
Figure 48 : présentation du staging area après chargement.....	70
Figure 49 : fenêtre du schéma du staging area après récupération.....	70
Figure 50 : exemple d’alimentation des table dimension.	71
Figure 51 : exemple d’alimentation des tables relié avec fait_vente.....	71
Figure 52 : exemple d’alimentation fait_achat.....	72
Figure 53 : chargement du fait « suivi magasin ».....	72
Figure 54 : chargement du fait « suivi vente ».....	73
Figure 55 : chargement du fait « suivi produit ».....	73
Figure 56 : construction des cube OLAP	74
Figure 57 : répartition des client selon l’age.	75
Figure 58: tableau de bord répartition des clients.....	76
Figure 59: suivi de facturation	76
Figure 60 : tableau de bord répartition des clients selon le sexe	77

Liste des tableaux :

Tableau 1 : Comparaison entre le système OLAP et OLTP.	17
Tableau 2 : Identification du rôle des employés et des clients	23
Tableau 3 : présentation l'architecture de la solution	31
Tableau 4 : Choix d'indicateurs	36
Tableau 5 : matrice en bus pour identifier facilement les dimensions répartie	37
Tableau 6 : Table de dimensions pour le suivi des produits	37
Tableau 7 :description de la table de fait « produit ».....	38
Tableau 8 : Table de dimensions pour le suivi ventes	38
Tableau 9 : description de la table de fait « vente »	39
Tableau 10 : Table de dimensions pour le suivi des achats	40
Tableau 11 : description de la table de fait « achat »	40
Tableau 12 : Table de dimensions pour le suivi du magasin	41
Tableau 13 : description de la table de fait « magasin »	41
Tableau 14 : Comparaison entre Talend et Pentaho	53
Tableau 15 : comparatif entre les outils de restitution.	54
Tableau 16 : les composants Talend utilisés	66

Liste des abréviations :

BI :	Business Intelligence
BDD :	Base de données
BM :	Business Models
CU :	Cas d'Utilisation
DW:	Data Warehouse
ERP:	Enterprise Resource Planning
ETL:	Extract Transform Load
FTP:	File Transfer Protocol
HOLAP:	Hybride OLAP
Id :	Identifiant
IA :	Intelligence Artificielle
OLAP:	On Line Analytical Processing
OLTP:	On Line Transaction Processing
ROLAP:	Relational On Line Analytical Processing
MOLAP:	Multidimensional On Line Analytical Processing
SID:	système informatique décisionnelle
SQL:	Structured Query Language (for databases)
SGBD :	Système de Gestion de Bases de Données
SGBDR :	Système de Gestion de Bases de Données Relationnelles
TI :	Technologie Information
TTC :	tous taxes comprises
TVA :	taxe sur la valeur ajoutée

I. Introduction :

Le monde ne cesse d'évoluer dans une économie de consommation trop vaste, ce qui conduit à une masse de données. À ce titre leur exploitation, développement, et leur utilisation devient de plus en plus complexe, et plusieurs informations se perdent. Et cela pose des difficultés pour les entreprises, les organisations dans l'élaboration des rapports de leurs activités au quotidien. Ainsi que les opérations de reporting sont très longues, ce qui pénalise l'entreprise, et la prise de décision au moment voulu. Voir même le problème du coût de cette dernière.

Et cela a incité les différents acteurs économiques à orienter et s'intéresser dorénavant à leurs données et les informations qui circulent dans leur entreprise. Avant qu'ils soient attirés par la concurrence. Pour ce faire, les entreprises doivent tisser des liens étroits et établir une communication permanente avec chacun de ces informations afin de pouvoir par la suite les étudier et les analyser. Cette étude, permet à l'entreprise, de mieux cerner les besoins et ainsi être en mesure de les cibler. Ce qui suscite un intérêt de maximiser leur satisfaction d'une part et augmenter son chiffre d'affaire d'autre part, à travers des campagnes marketing de fidélisation.

Notre travail se porte sur la réalisation d'une solution de business intelligence. Cette dernière a pour rôle de collecter, de transformer et d'analyser des données extraites de différentes bases de production, via un outil nommé ETL. Avec l'existence d'autres outils, utilisés dans les différentes phases pour la réalisation de cette solution.

II. Problématique

L'entreprise dispose d'une grande quantité de données concernant ces activités, et elles sont générées automatiquement chaque jour. Si elle souhaite réaliser des rapports; il lui faut un traitement manuel (coûteux en termes de temps). Ces données ne sont généralement jamais exploitées ni utilisées. Ces données constituent des informations utiles ; et elles représentent un atout, pour lui permet de mieux effectuer l'opération d'analyse. De plus, elle fait face à une rude concurrence sur le marché dont elle risque de perdre la clientèle. Cependant que l'entreprise se retrouve donc face à deux principaux enjeux:

➤ **Problèmes lié aux bases de productions :**

- Problème d'historisation: les données sont gardées que sur une courte période. Ce qui produit une difficulté d'accès pour effectuer des comparaisons et des analyses.
- Des problèmes de qualité des données.

➤ Problèmes liés aux analyses et l'élaboration des rapports d'activité

- Un problème se pose lors du stockage et à l'analyse des données, la capacité de stockage des disques durs augmente mais le temps de lecture croît.
- Des problèmes liés à la difficulté de visualiser les données afin de prendre les bonnes décisions. Et ce problème est causé par : Le manque d'outils décisionnels et des technologies pour collecter et transformer les données.
- Manque d'autonomie, le décideur ne trouve aucun moyen d'élaborer les rapports en toute indépendance.
- Difficultés dans l'élaboration des rapports d'activité.
- Retard et Coût de la procédure de Reporting.

Pour répondre à ces besoins, l'entreprise s'engage dans un projet qui lui permet de comprendre et d'analyser ces activités à travers les données qu'elles génèrent sur ces derniers. Les résultats de cette analyse, vont lui offrir la possibilité de pouvoir répondre à leurs besoins dans le but ultime améliorer l'efficacité de ces campagnes marketing.

III. Solution

Notre travail consiste de mettre en place d'une solution BI pour répondre à une problématique décisionnelle ; dans le but d'améliorer l'efficacité de ces campagnes marketing. De plus, ce dernier doit traiter les tâches de collecte des données ainsi que transformation. En second lieu, nous allons exploiter ces données afin de concevoir et réaliser les différents rapports et l'affichage des résultats.

Afin de poursuivre, il est important de souligner quelques points déterminants pour la suite du projet. Ces points soulèvent des questions et requièrent des décisions à prendre. Nous tenterons donc de répondre aux questions suivantes :(au quels nous répondrons dans notre conception du système final).

1. Quelle sont les types de données à prendre en compte dans l'analyse?
2. Quel mécanisme adopter pour récupérer ces données ?
3. Quel serait la meilleure méthode à appliquer pour analyser ces données ?

IV. Objectifs et le périmètre du projet BI :

- ✓ Les analyseurs des ventes, des achats et du comportement des clients souhaitent stocker les données collectées dans une base de données informatique, appelée DW (DW).

- ✓ Disposer des outils performants adaptés pour analyser ; et rechercher toute information qui intéresse les domaines de vente (achat).
- ✓ Permettre d'avoir une vision claire sur les activités de vente et d'achat, et d'avoir une évaluation selon les exigences et les demandes des clients.
- ✓ Amélioration les rapports et la prise de décision.

I. Structure du rapport :

Afin de bien présenter notre travail, nous avons divisé le document en trois parties principales, résumées dans ce qui suit :

Chapitre I – : les systèmes d'information décisionnelle :

Dans ce chapitre on traite l'aspect théorique de la BI dans laquelle nous allons présenter le système d'aide à la décision BI. Ensuite, nous avons abordé ces phases ; ces concepts ; comme le data Warehouse, où nous allons définir les entrepôts de données ainsi que la modélisation multidimensionnelle. Pour en finir, par présenter l'ensemble des avantages ; et des limites de la BI ; et présenter son évolution dans le futur.

Chapitre II – : Spécification et analyse des besoins :

Dans ce chapitre, nous avons décrit l'accueil de notre entreprise dans un premier temps. Et dans un seconde on a traité l'analyse de l'existant et la spécification de l'entreprise. Au final, on a décrit la solution choisie et sa structure.

Chapitre III – : Conception de la solution :

Dans chapitre nous présentons l'aspect conceptuel de notre solution. En premier lieu la conception et l'alimentation du DW. Pour en finir avec la construction des cubes OLAP.

Chapitre IV – Réalisation et mise en œuvre de la solution :

Ce chapitre représente les outils choisis pour la réalisation en premier lieu. Ensuite, on présente l'application de vente réalisée. Et détailler les étapes de réalisation de la solution décisionnelle dans l'objectif du reporting.

Chapitre I :

Les systèmes d'information décisionnelle

I. Introduction :

Ce chapitre se décompose en trois sous-titres. Le premier se base sur la présentation de la BI. Le second titre s'intitule sur l'informatique décisionnelle vs Informatique opérationnelle. En dernier lieu nous allons définir les limites et l'évolution de la BI.

II. Présentation de la Business Intelligence :

II.1. Historique de la BI : [1]

La BI a vu le jour pour la toute première fois vers les années 60, précisément en 1958. Au cours de cette année, un informaticien nommé **Hans Peter Luhn**. Collaborant chez IBM, présenta une théorie pour le système robotisé conçu pour envoyer des informations à différents secteurs ; qu'elle soit industrielle, scientifique ou gouvernementale.

Ce système d'intelligence se base sur des machines de traitement de données ; dans le but d'extraire et transformer des documents. Et de créer des profils pour faciliter aux organisations l'utilisation des résultats retournés. Depuis cette année la data et la BI n'ont cessés de s'améliorer. Notamment avec l'apparition de nouvelles technologies qu'on citera ci-dessous :

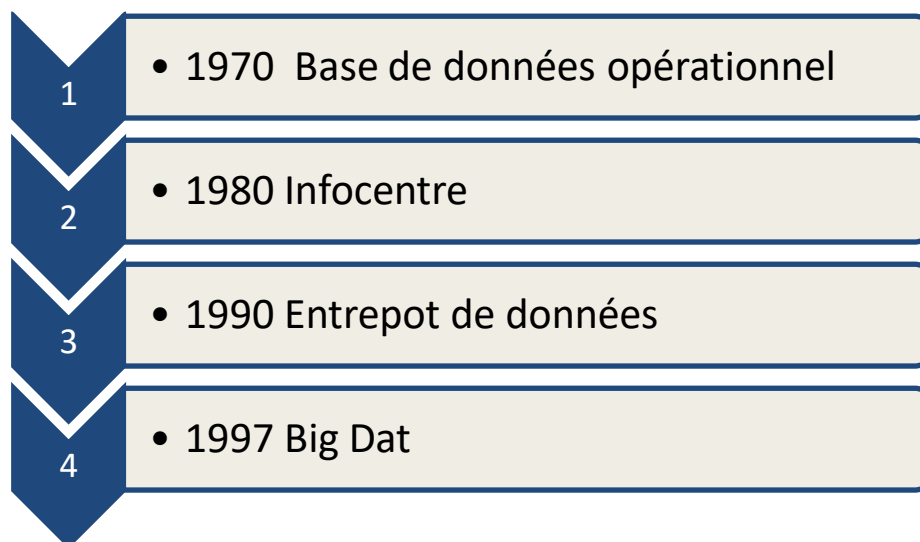


Figure 1 : Historique d'évolution de la da

En 1989, un autre informaticien du nom **Howard Dresner**, pose la problématique de pouvoir stocker les données dans des unités particulières, et améliorer les capacités analytiques. Et le besoin d'un environnement capable de gérer d'énormes quantités de données indépendamment de la production ; a donné naissance à un nouveau concept du nom « Data Warehouse » qui se base sur le principe ETL, ces deux derniers concepts seront détaillés dans les parties suivantes.

II.2. Définition de la BI : [2]

La Business Intelligence dite aussi Informatique décisionnelle est un ensemble d'outils et technologies qui se basent sur la collecte, l'analyse ; la restitution et la modélisation des données différentes et variées d'une entreprise. C'est un système dédié aux décideurs d'une organisation. Il permet aux utilisateurs d'obtenir des informations utiles à la prise de décision pour améliorer ces performances. et cela grâce à une base de données et une interface d'accès aux données.

Selon Inmon (Inmon B. , 1996) un système d'information décisionnel est défini comme étant «un regroupement de données orientées vers certains sujets, intégrées, dépendantes du temps, non volatiles, ayant pour but d'aider les gestionnaires dans leurs prises de décision ». Ainsi, Kimball le définit comme « une copie des données transactionnelles organisées spécialement pour l'interrogation et l'analyse ». [4][8]

II.3. Les domaines d'utilisation de la BI :

La BI est utilisée dans plusieurs domaines nous citerons :

- ✓ Finance, avec le reporting financiers et budgétaires.
- ✓ Vente et commercial, avec l'analyse de points de vente, l'analyse de la profitabilité.
- ✓ Marketing, avec la segmentation clients, les analyses comportementales.
- ✓ Logistique, avec l'optimisation de la gestion de stock, le suivi des livraisons.
- ✓ Ressources humaines, avec l'optimisation de l'allocation des ressources.

II.4. L'architecture du processus de la BI et ces principales phases :

Au niveau pratique et technique, l'intelligence d'affaires consiste en une suite d'outils, de logiciels informatiques, appelée chaîne décisionnelle. La chaîne décisionnelle est une chaîne de traitement de l'information, pour des fins décisionnelles. Et elle se traite selon des outils décomposés en quatre phases différentes qui sont cités ci-dessous :

II.4.1. Les principes phases :

- ✓ Phase de collecte et transformation.
- ✓ Phase d'intégration ou de stockage.
- ✓ Phase de restitution de données.
- ✓ Phase d'exploitation et d'analyse.

II.4.2. Le fonctionnement de ces phases : [2]

- ✚ **Phase de collecte (extract)** : Cette première phase consiste à collecter, nettoyer et consolider les données de l'entreprise venue de différentes sources. cela en utilisant les outils d'ETL pour Extract Transform Load.
- ✚ **Phase d'intégration ou de stockage** : Les données résultantes de la première phase seront ensuite stockées dans une base spécialisée. Nommé **DW** ou dans un **data mart** (qui est une version plus réduite du DW). Dans l'objectif de les préparer au rôle final dit analyse décisionnelle.
- ✚ **Phase de distribution de données** : Consiste à distribuer l'information auprès de l'ensemble des partenaires et la restituer sous une forme exploitable. C'est dans cette phase que les outils de restitution vont intervenir notamment les portails d'accès à des tableaux de bord, les outils de reporting... etc.
- ✚ **Phase de restitution ou d'exploitation** :
Une fois les données collectées, nettoyées, stockées, et accessibles elles peuvent être analysées pour faire ressortir des prévisions ; ou des estimations futures en utilisant les outils du datamining. Selon les besoins, différents types d'outils d'extraction et d'exploitation seront utilisés tels que :
 - ✓ OLAP pour les analyses multidimensionnelles, notamment analyser les données.
 - ✓ Le Datamining pour rechercher des corrélations.
 - ✓ Les tableaux de bord présentant les indicateurs clés de l'activité.
 - ✓ Le Reporting pour communiquer la performance.

II.4.3. Les outils utilisés dans ces différentes phases : [1]

II.4.3.1. Outils de collecte :

Les outils d'ETL : les outils ETL servent à récupérer les données de l'entreprise quelle que soient leurs sources. Ses données seront ensuite formatées, nettoyées et consolidées ; avant d'être chargées dans des bases spécialisées appelées DW ou Data-mart.



Figure 2 : Le processus ETL

II.4.3.2. Outils de stockage :[5]

1. **Le Data mart** : c'est une version plus réduite d'un DW, conçu pour répondre à un besoin métier spécifique.
2. **Les DW** : dits entrepôt de données sont des bases de données spécialisées utilisées pour stocker des données provenant de bases de données de production à l'aide des outils d'ETL.

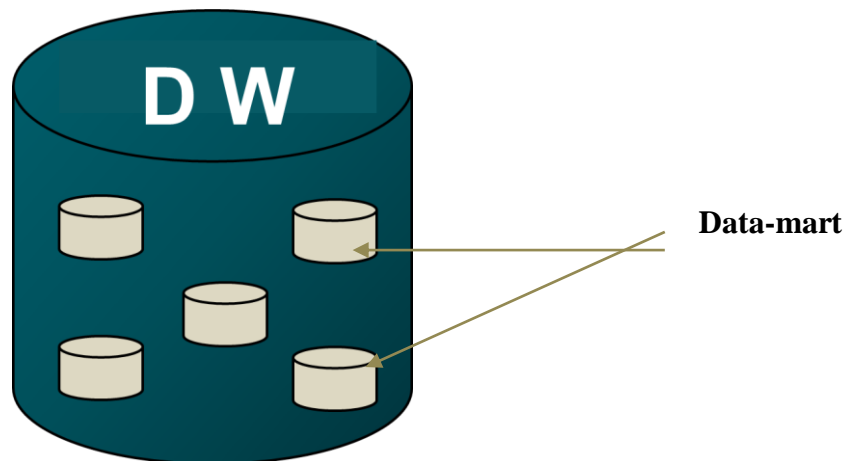


Figure 3 : présentation du DW et data-mart

3. Les 4 caractéristiques d'un DW : [6][10]

✚ Intégrées :

La DW est un système intégré, et non un simple regroupement de données provenant de diverses sources. Les données collectées grâce aux outils d'ETL sont alors transformées de manière à obtenir un ensemble homogène avant d'être stockées dans le DW, c'est une étape très complexe dans la mise en place d'un DW.

✚ Orientés sujets :

Les données d'un DW sont organisées par sujets et ceci pour faciliter leur traitement. Un DW peut ainsi être déployé spécialement pour analyser les données liées à un service spécifique de l'entreprise.

✚ Historisées :

Les données de DW ne sont jamais mises à jour, elles sont ajoutées au fur et à mesure, le rafraichissement d'un DW consiste alors à ajouter de nouvelles données sans avoir à modifier ou supprimer celles qui existent, et cela pour pouvoir visualiser au temps voulu l'évolution d'une valeur donnée. Un référentiel de temps doit ainsi être mis en place.

✚ Non volatiles :

Du fait que les données d'un DW sont historisées celles-ci ne sont jamais modifiées ou supprimées au fil du temps.

4. Types du DW : Nous distinguons quatre types de DW :

- ✓ Les systèmes de gestion de base de données (SGBD) relationnelles traditionnels qui sont les plus courantes.
- ✓ Les SGBD relationnelles spécialisés.
- ✓ Les DW Appliance.
- ✓ Et les DW Cloud.

5. Modélisation dimensionnelle, [7][14]

Nous distinguons deux types de tables. C'est le mode d'interconnexion de ses tables qui caractérise une modélisation. On trouve donc :

- ✓ **Les tables de dimensions** : elles sont utilisées pour présenter les données que l'on souhaite stocker dans le DW, chaque table de dimensions peut avoir plusieurs attributs.
- ✓ **Les tables de faits** : Elles contiennent les données que l'on souhaite voir dans des rapports d'analyse. Une table de faits se présente sous la forme d'un ensemble de colonnes stockant des valeurs dites mesures, et des clés étrangères (identifiants) qui sont généralement les clés primaires associées aux tables de dimensions.
- ✓ **Mesure** : les indicateurs d'analyse sont représentés sous forme de mesures, extraites par une ou plusieurs dimensions.

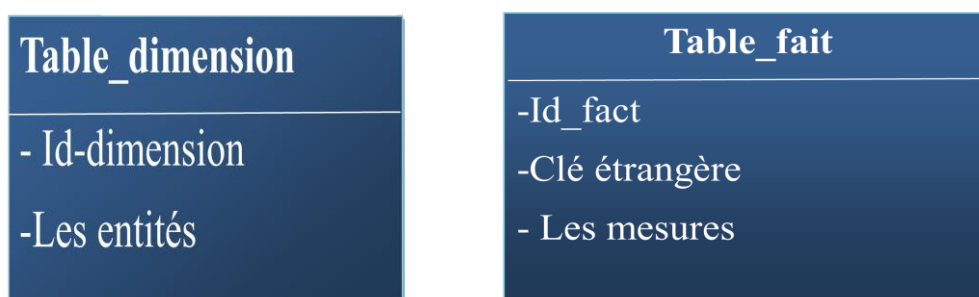


Figure 4: représentation de table de fait et table de dimension

6. Les modélisations dimensionnelles

Les modélisations possibles pour organiser les données stockées dans un DW; qui sont construites à partir des tables mentionnées ci-dessus sont :

6.a La modélisation en étoile :

Ce modèle tire son nom de sa configuration, en effet il se forme d'un objet central qui est la table des faits reliée à un ensemble de tables de dimensions. Présenté dans le schéma ci-dessous :

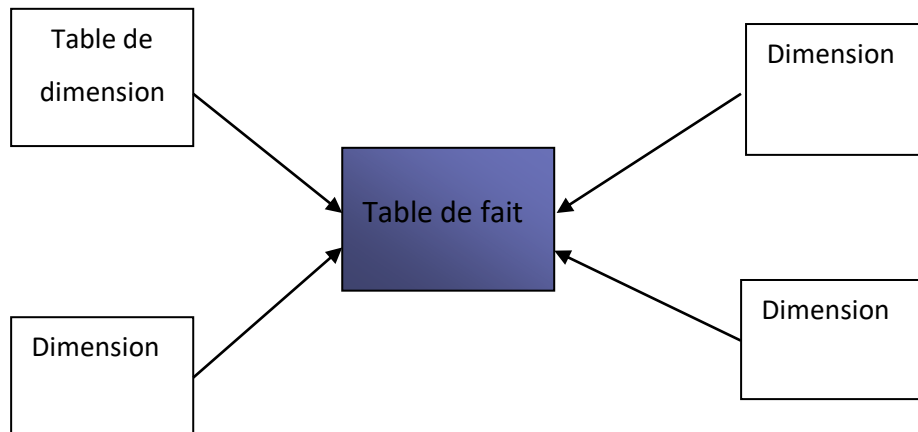


Figure 5 : Modélisation en étoile

6.b La modélisation en flocons :

Ce modèle est plus complexe de celui en étoile du fait qu'il contient beaucoup plus de tables. Le principe était qu'il peut exister des hiérarchies de dimensions ; qui sont reliées à la table de faits.

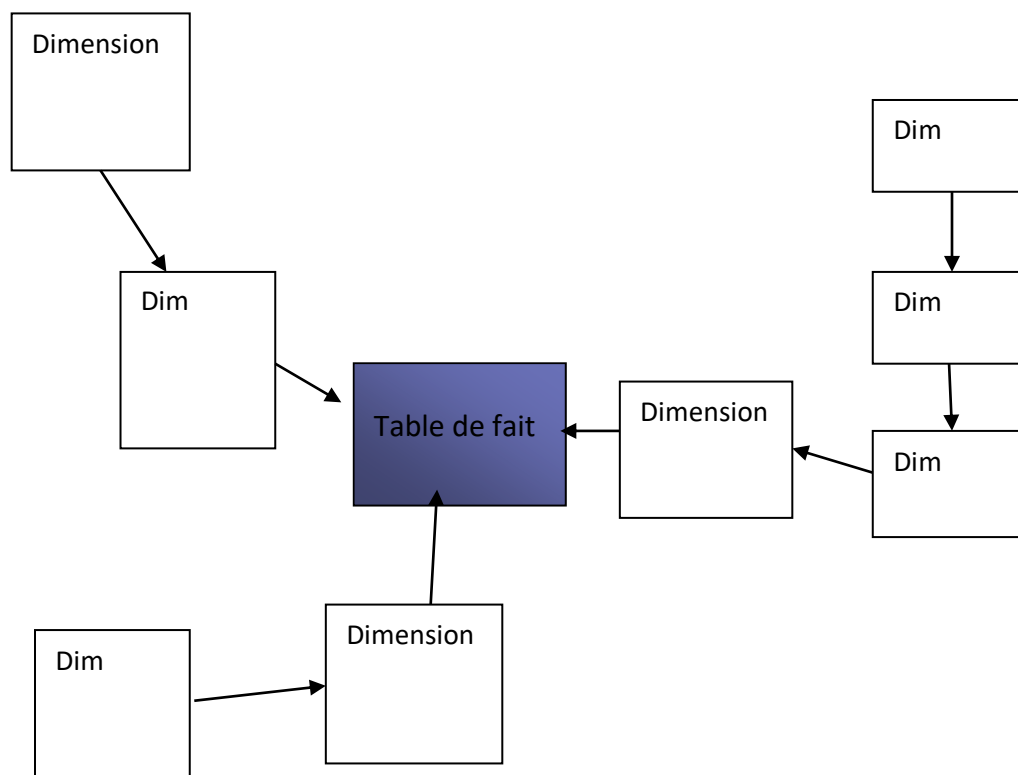


Figure 6 : Modélisation en flocons

6.c La modélisation en constellation :

Dérive de la modélisation en étoile, il représente plusieurs relations de faits partageants ou non des dimensions.

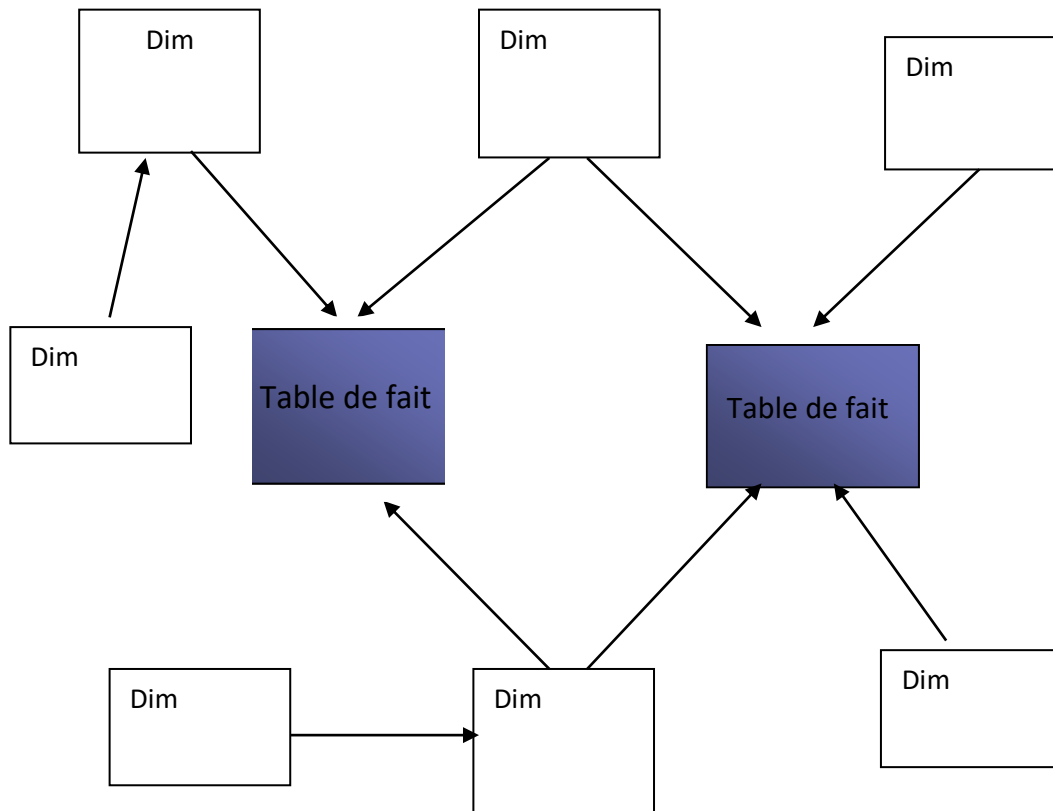


Figure 7 : La modélisation en constellation

7. Construction d'un Data Warehouse:

Pour construire un DW trois méthodes s'offrent à nous :

- ✓ **Top-Down** : c'est la méthode la plus lourde, la plus contraignante et la plus complète en même temps. Elle consiste une conception totale de l'entrepôt de données ; et il faut savoir à l'avance toutes les dimensions et tous les faits de l'entreprise avant sa réalisation. L'avantage de cette méthode est qu'elle nous donne une vision très claire et très conceptuelle des données de l'entreprise ainsi que du travail à faire.
- ✓ **Bottom-Up** : c'est l'approche inverse de la première méthode, Elle consiste à créer les modalisations dimensionnelles, puis les regrouper par des niveaux intermédiaires jusqu'à obtention d'un véritable entrepôt. L'avantage de cette méthode est qu'elle est simple à réaliser et l'inconvénient est le volume d'intégration pour obtenir un entrepôt de données.

- ✓ **Middle-Out** : c'est l'approche conseillée par les professionnels du BI. Elle consiste une conception totale de l'entrepôt de données. Puis créer des divisions plus petites et plus gérables et les mettre en œuvre. Cette méthode est la meilleure par rapport aux deux précédentes.

II.4.4. Restitution ou exploitation :

1. Les outils de type OLAP :

OLAP, dit traitement analytique en ligne, est un système de traitement informatique. Il permet aux utilisateurs de visualiser et d'extraire aisément les données à des fins de comparaison. Il s'agit d'un outil d'aide à la décision très utile. Les données sont stockées sur une base multidimensionnelle appelée Cubes OLAP. Pour simplifier ce type d'analyses un serveur OLAP est requis.

1.1 Représentation du cube de données :

La figure ci-dessous montre un cube de données sur trois axes (catégories de produits, magasin et des dates). L'intersection de ces axes, c'est-à-dire la cellule en bleu, représente une opération de vente. A noter également que la notion de cubes de données ne se limite pas en trois axes seulement, mais elle peut se généraliser en hypercube.

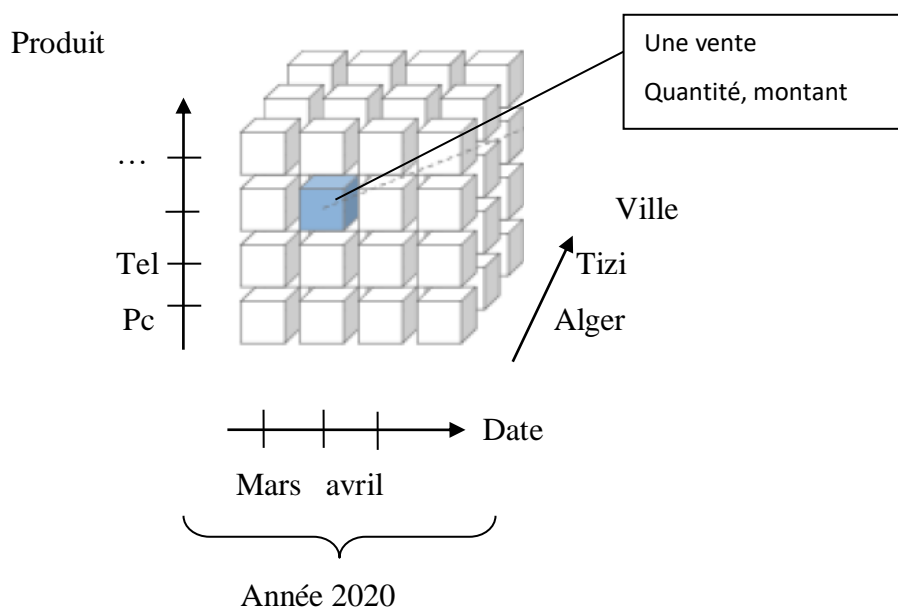


Figure 8 : Cube de données

Il existe plusieurs options pour stocker les données OLAP, à savoir ROLAP, MOLAP. HOLAP, ces abréviations désignent Relationnel. Multi dimensionnel et OLAP hybride respectivement.

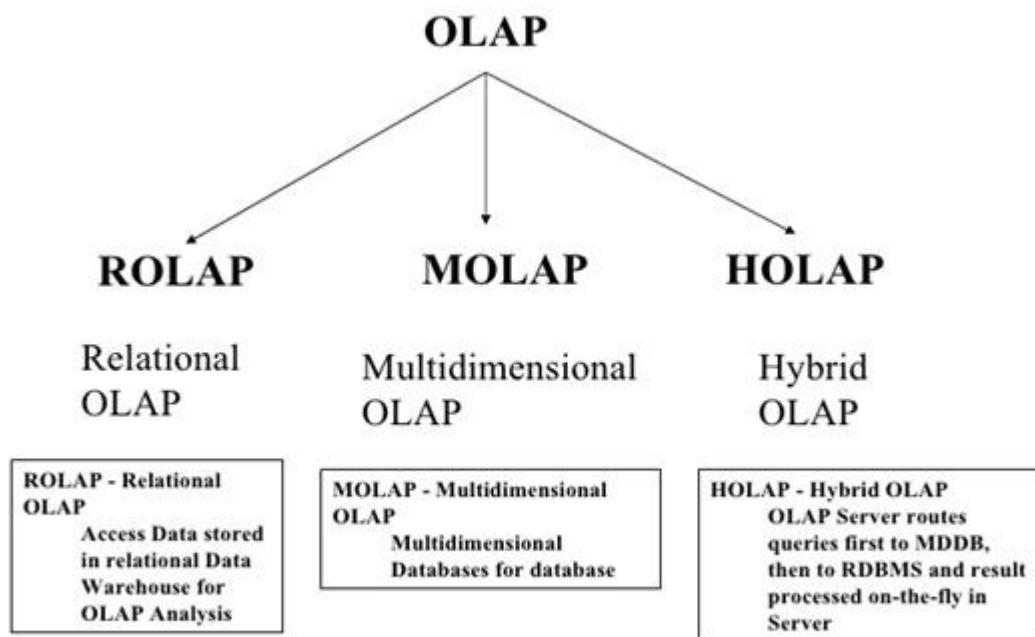


Figure 9 : architecture du système OLAP.

1.2. L'architecture ROLAP [12]

Elle a la même forme que (OLAP), elle effectue une analyse multidimensionnelle dynamique des données stockées dans des schémas BDD non-normalisée. L'utilisateur demande une analyse multidimensionnelle et le moteur ROLAP convertit la demande en SQL, pour la commander à la BDD. Ensuite, l'opération est exécutée en sens inverse : le moteur convertit les données résultantes de SQL en un format multidimensionnel, puis renvoyées au client pour être visualisées. L'avantage de cette approche est la possibilité que ces outils stockent et traitent de grandes quantités de données. Pour optimiser la requête de la BDD relationnelle, la sortie des requêtes peut être mise en cache dans la BDD ou dans une structure multidimensionnelle.

1.3. L'architecture MOLAP : [13]

Le traitement M-OLAP utilise une BDD multidimensionnelle pour le stockage des données. Elle dispose une structure optimisée pour l'analyse, souvent nommée cube ou hyper-cube. M-OLAP utilise un système de gestion de BDD, sous forme de matrices. Généralement plus performante que R-OLAP pour plusieurs raisons. Premier mot, l'accès aux données par l'indexation multidimensionnelle n'est pas coûteux. A savoir que les données sont organisées pour être directement exploitables par les outils d'analyse. Ensuite, les méthodes de compression réduisent au minimum l'espace de stockage. Ce qui permet d'occuper moins de place sur le disque.

1.4. L'architecture HOLAP (Hybride OLAP) : [14]

Elle représente un mélange entre les deux implémentations précédentes. Durant cette implémentation, un serveur H-OLAP accède à deux bases de données différentes, tel que : La première, multidimensionnelle, contient les données agrégées. La seconde, relationnelle, contient les données détaillées. L'implémentation H-OLAP utilise les caractéristiques des deux architectures R-OLAP ; et M-OLAP suivant le type de données manipulées par les requêtes ROLAP dans le cas de données détaillées et M-OLAP dans le cas de données agrégées. Elle est particulièrement attirante lorsque : La majeure partie des requêtes manipule des données agrégées (l'accès aux données dans M-OLAP est peu coûteux). Et que les données détaillées représentent un volume important (les SGBDR permettent de traiter un volume considérable de données).

2. Le Reporting :

C'est un terme rassemblant plusieurs outils destinés à assurer la réalisation, la publication et la diffusion de rapports d'activité selon un format prédéterminé et utilisé pour présenter les données obtenues sous forme simplifiée et plus compréhensible. Il permet de faciliter la communication des résultats obtenus.

3. Les tableaux de bord :

Ils donnent accès à différents indicateurs permettant de comprendre la situation, et d'aider à prendre une décision. Il permet de piloter l'entreprise.

4. Le Datamining :

C'est un terme rassemblant plusieurs outils facilitant l'analyse et l'exploitation des données contenues dans une base décisionnelle de type Data Warehouse ou Datamart. Il permet d'extraire des informations significatives et pertinentes depuis de grandes quantités de données et exploiter d'éventuelles corrélations peu visibles.

A la fin de ses quatre phases, les responsables de la prise de décision peuvent accéder à des informations organisées et présentables. Pour leurs permettrons de prendre de bonnes décisions.

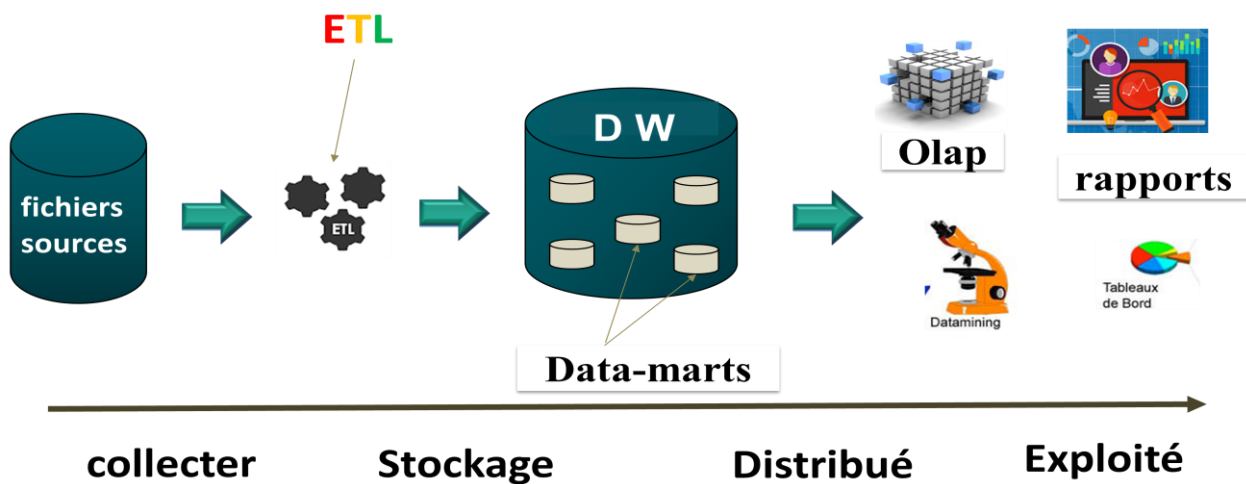


Figure 10 : L'architecture du processus de la BI et ces phases.

II.5. Les principaux avantages de la BI pour l'entreprise :[3]

- ✓ Réaliser des analyses prédictives.
- ✓ Présentation structurée et cohérente de l'information.
- ✓ Automatiser la prise de décision sur la base des indicateurs des entreprises.
- ✓ Une meilleure qualité des données.
- ✓ Maitrise de la relation client.
- ✓ Anticiper et réduire les défaillances de l'entreprise.
- ✓ Gestion plus simplifiée de l'entreprise.
- ✓ Prise de décision rapide et pertinente.
- ✓ Prendre de l'avantage sur la concurrence sur le marché.
- ✓ Augmenter la productivité de l'entreprise.
- ✓ Une visibilité plus claire sur les activités.
- ✓ Une interaction, et une compréhension facile avec les données.
- ✓ Aider les décideurs à appréhender la situation quels que soient les changements de contexte.
- ✓ **Éviter les dépenses** : il s'agit des coûts produits par les erreurs et du coût de problématique du personnel dus au déploiement de plusieurs supports d'analyse. Grouper les processus d'aide à la décision sur un seul support(DW) permet d'éviter ces dépenses.

III. Informatique décisionnelle vs Informatique opérationnelle :

III.1. Informatique décisionnelle :

Les systèmes « décisionnels », également appelés OLAP pour *on-line analytical processing* sont dédiés au management de l'entreprise pour l'aider au pilotage de l'activité, et donc indirectement opérationnels.

III.2. Informatique opérationnelle :

Les systèmes « opérationnels » ou « de gestion », également appelés systèmes OLTP pour *on-line transaction processing*, sont dédiés au métier de l'entreprise pour les assister dans leurs tâches de gestion quotidiennes et donc directement opérationnels.

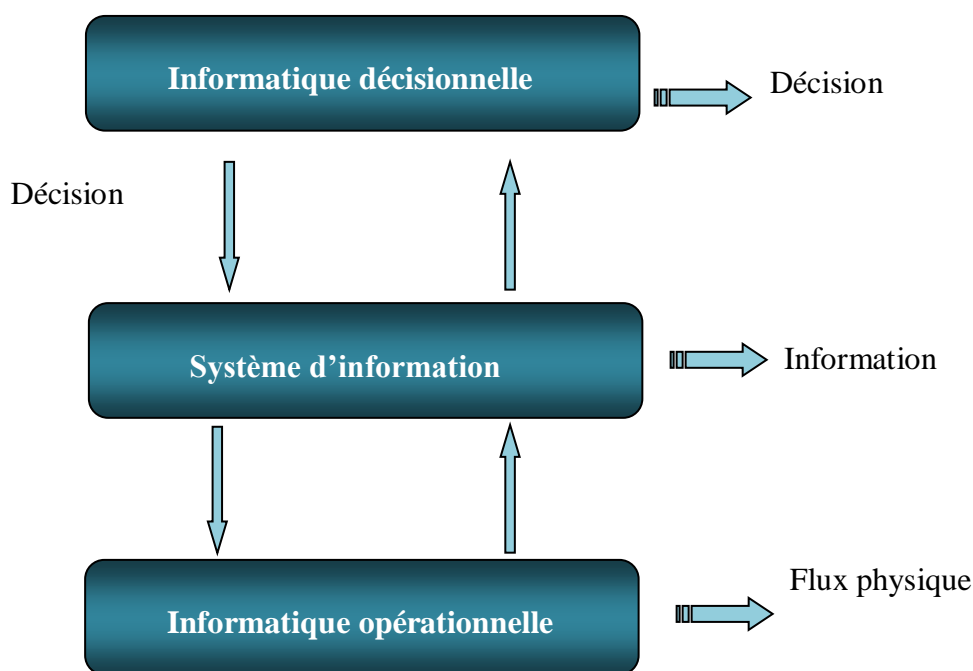


Figure 11: la place du décisionnel et d'opérationnel dans l'entreprise.

III.3. Décisionnel vs Opérationnel :

Un système opérationnel, appelé aussi système transactionnel est dédié aux emplois de l'entreprise .Pour participer dans leurs tâches d'administration quotidienne. Quant au système d'aide à la décision déjà défini, il est dédié au management de l'entreprise pour l'aider au pilotage de l'activité.

Le tableau ci-dessous présente une comparaison entre les deux systèmes :

	OLTP	OLAP
Système	Rapide	Très rapide
	Fermés : on ne laisse pas la place à l'imprévu dans les OLTP, les utilisateurs sont guidés dans le processus	Simplicités : les environnements d'un SID doivent permettre d'accéder le plus simplement possible aux données
	Transactionnels : OLTP fonctionne selon les principes de transaction	Non transactionnels : l'utilisateur doit pouvoir démarrer l'analyse, puis revenir en arrière
	Fragmentés : ou dispersé. Sauf le cas des ERP	Centralisés : toutes les données sont regroupées dans une seule source
Utilisateur	Grand public : destinés à toute personne	Petit public : sauf les décideurs
	Faible niveau de demande d'analyse	Niveau élevé de demande analytique
Données	Utilise des données récentes	Utilise des données historiques
	BDD normalisée (3eme forme) (SGBD)	BDD dénormalisée (DW)
	Lecture, écriture, et modification	Lecture seulement
	Hétérogènes : Les systèmes OLTP sont souvent différents dans la technologie utilisée	Homogènes
	Orientation linéaire des données	Orientation multidimensionnelle des données

Tableau 1 : Comparaison entre le système OLAP et OLTP.

IV. limites et évolution de LA BI :

Cette partie traite en général les limites de la BI que les utilisateurs de ce domaine ont pu remarquer au fur de ces dernières années, et la cause principale est due à et l'évolution de nouvelles technologies, à savoir leurs puissances et leurs avantages. Ce qui mène à une influence sur ce domaine. Une seconde partie, traitera l'évolution de BI. Au final nous allons présenter l'avenir de la BI, si la BI continuera d'exister ou serait-elle remplacée par d'autres (concepts) méthodes !

IV.1. Les limites et les inconvénients de la BI :

Malgré tous les avantages dont la BI dispose, la BI dispose de certain inconvénient et des limites :

- ✚ La non-possibilité d'utiliser toutes les données associées pour faire des rapports de ces activités. Cela signifie que certaines informations sont négligées d'affichages.
- ✚ Les indicateurs choisis pour faire leur système d'aide à la décision sont variant, ils changent au fur du temps, de nouvelles exigences peuvent apparaître (que ça soit de côté administrateur ou des clients).
- ✚ Les équipes découvrent de nouvelles fonctions
- ✚ Ainsi ils découvrent de nouveaux outils qui peuvent répondre en une grande précision à leurs besoins.
- ✚ La complexité croissante des activités.
- ✚ La solution BI avec entrepôt de données nécessite un serveur de BDD et un temps de mise en œuvre plus long.
- ✚ Le besoin de consulter les informations en temps réel est devenu très demandé, et les sources de données sont de plus en plus larges et non structurées ce qui présente un obstacle au niveau du traitement et leurs stockages.

IV.2. L'évolution de la BI :

La BI est dans une évolution croissante, et ne cesse de s'améliorer. Ci-dessous nous allons vous présenter les 6 évolutions les plus importantes que la BI a connues :

IV.2.1. Une ouverture vers les métiers :

À ses débuts la Business Intelligence était qu'aux secteurs de marketing et du commerce. Et aujourd'hui elle s'est répandue dans tous les métiers et en particulier au sein des directions financières et cela grâce aux nombreux avantages qu'apportent les solutions de BI.

III.2.2. Les représentations statiques ont laissé place à la Data Visualisation :

Avec l'amélioration du domaine BI et celui des interfaces, par conséquent on dispose aujourd'hui de restitutions graphiques interactives qui permettent l'exploration de données, et leur modélisation de façon intuitive.

IV.2.3. Les solutions agiles et modulaires prennent le pas:

Les solutions ciblées et modulaires en mode SaaS ont remplacé aujourd'hui les solutions coûteuses, lourdes, et exigeant des investissements substantiels en intégration proposée par les grands éditeurs historiques de la Business Intelligence.

IV.2.4. De l'analyse du passé à la prévision du futur :

Le monde change de plus en plus vite et le besoin d'anticipation est devenu un besoin vital pour toute organisation. De ce fait le principe de l'analyse prédictive est désormais intégré dans la plupart des solutions BI capable d'intégrer tendances historiques et surveillance des signaux faibles pour veiller et prévenir. Après que les solutions de BI étaient seulement focalisées sur les données du passé d'origine interne et structurée.

IV.3. L'avenir de la BI :

Avec l'apparition de nouveaux concepts qui ont bouleversé le marché des données et de l'analyse tel : **le cloud**, et **l'IA**, et **le Big Data**. La BI est incapable de passer à côté de ces concepts, surtout que ces nouvelles technologies occupent le monde d'information. Et l'extraction des données d'une base de production ; pour faire des rapports seront moins courants au fur du temps. Et la question se pose ; la BI a-t-elle un avenir ou sera-t-elle vite remplacé? Par contre la réalité suit un autre chemin, ces technologies vont réunir leurs forces, et cela présente un atout pour la BI.

IV.3.1. Le cloud computing :

N'est pas seulement une application, il peut non seulement stocker des données et des programmes systématiquement via un réseau de services distants ou internes, mais également fournir des services tels que SaaS, PaaS, LaaS.

IV.3.2. Big data :

Signifie grosses données ou méga-données c'est-à-dire ensemble de données à très grande échelle, le big data est très volumineux ce qui nécessite des outils de traitement spéciaux. Il ne permet à toute personne d'accéder à des bases de données en temps réel.

IV.3.3. Intelligence Artificielle :

Désigne l'ensemble de technologies mises en œuvre pour la réalisation des machines qui seront capables d'imiter l'être humain en simulant son intelligence et cela grâce à des programmes informatiques complexes.

IV.3.4. La BI et le cloud :

Pour les entreprises qui veulent migrer vers le cloud leurs systèmes sont automatisés à l'aide d'application cloud native qui permet d'extraire des connaissances, faire des recommandations et crée des tableaux plus clairs et visibles, Le but d'opter vers le cloud se résume:

- ✓ Améliorer l'agilité
- ✓ Les couts plus réduits
- ✓ Améliorer le développement des produits et l'innovation

IV.3.5. La BI et IA :

L'intelligence artificielle est reconnue pour le faite qu'elle figure dans la liste des révolutions qui touche tous les domaines. Les performances de l'IA ne cessent d'augmenter ce qui conduit vers une résolution des problèmes plus précise en encore plus facile. Ainsi, La BI et l'IA sont deux technologies complémentaires. Tant dit que la BI fournit des informations que l'utilisateur lui a demandées à travers les rapports et les tableaux de bords, l'IA répond à des questions qui ne sont pas encore posées. Donc l'idéal est d'utiliser dans un système d'information une solution dédiée à la Business Intelligence et le renforcer avec de l'Intelligence Artificielle.

IV.3.6. BI et Big Data

Les nouvelles technologies du Big Data permettent d'envisager de nouvelles solutions pour les systèmes décisionnels existants comme les méthodes de stockage avancées, le calcul distribué, les capacités temps réel, l'intégration de données. Le Big Data et ses outils permettent donc à la Business Intelligence d'évoluer vers l'agilité, la performance.

Conclusion :

Ce chapitre clôture l'aspect théorique de la BI, et ces déférentes phases et leur fonctionnement. On a présenté les avantages de la BI, ainsi que ces limites. Et dans le chapitre qui suit, on va définir l'accueil de notre entreprise, et faire une étude et analyse des besoins. Pour finir par exposer une solution.

Chapitre II :

Spécification et analyse des besoins

I. Introduction :

Dans ce chapitre, nous allons définir et présenter en premier lieu notre entreprise, et détailler son fonctionnement. Puis faire une analyse et spécifications de ces besoins. Comprendre ses objectifs et les exigences à respecter, améliorer les processus dans elle dispose. Afin de pouvoir construire une bonne stratégie (solution BI). Avant d'entamer l'analyse, nous allons commencer par définir la démarche suivie.

II. Description de l'entreprise :

II.1. Définition:

C'est une entreprise fictive dans le domaine de vente de matériels informatiques. Elle offre des avantages aux clients, telle la disponibilité à tout moment, multiple choix avec des prix raisonnables et d'autres privilèges supplémentaires.

II.2. L'environnement de l'entreprise

L'entreprise n'est pas indépendante, elle existe dans un environnement. Ce dernier lui impose des contraintes et des conditions lors du développement de ces activités. C'est l'un des cas qui poussent les structures internes de l'entreprise à s'adapter aux types et aux conditions de l'environnement, qui n'est ni statique, ni homogène. Cependant les directeurs doivent prendre en compte les risques qui viennent de l'extérieur tel : la concurrence, l'augmentation de prix des produits ou leurs réductions....

II.3. Fonctionnalités de l'entreprise

L'entreprise dispose d'un système opérationnel (OLTP) pour organiser ses fonctions. Elle participe dans les tâches de gestion quotidiennes (elle regroupe toutes les données de gestion comme les finances, les ventes, ...), pour les organiser dans une BDD.L'entreprise est chargée d'assurer la mise en œuvre de la vente à travers la prise en charge de ces activités. L'entreprise s'occupe :

- ✓ D'assurer la surveillance de la qualité des produits qu'elle achète.
- ✓ D'assurer la surveillance des prix et les quantités des produits en entrée et en sortie.
- ✓ Expliquer toute activité effectuée, par exemple le but des promotions.
- ✓ L'amélioration des outils qu'elle utilise dans le domaine de vente (site, application).
- ✓ La livraison qui doit être dans les temps et gratuite dans notre cas.
- ✓ Etablir une facturation pour les clients, après chaque confirmation de la commande.
- ✓ Le mode de paiement s'effectue lors de la livraison (main à main)

II.4. Identification du rôle des employés et des clients

Dans notre cas d'étude, nous avons créé une application de vente, pour laquelle on a créé une base de données (BDD). Nous nous sommes intéressés aux comportements de deux acteurs principaux qui sont : le client et l'administrateur, et par l'étude des échanges d'informations effectués entre eux comme exemple : l'échange établi lors d'achats d'un produit par le client.

Acteurs	Rôles
Clients	<ul style="list-style-type: none"> -S'inscrire, s'authentifier, gérer son profil (supprimer, modifier, rechercher) -Rechercher et/ou Réserver un produit -Confirmer la réservation
Administration	<ul style="list-style-type: none"> -Gérer les produits (consulter, ajouter, modifier, supprimer, rechercher) -Gérer les salaires -Gérer et piloter l'entreprise; le suivi commercial (achat et vente) -Organiser le déroulement des activités -Assurer la mise en œuvre et le suivi de toutes ces activités
Vendeur	<ul style="list-style-type: none"> -Confirmer les demandes des clients et établir des factures
Livreur	<ul style="list-style-type: none"> -Livrer la marchandise au destinataire dans les temps et dans un bon état (la livraison est gratuite)

Tableau 2 : Identification du rôle des employés et des clients.

II.5. Diagramme de classes :

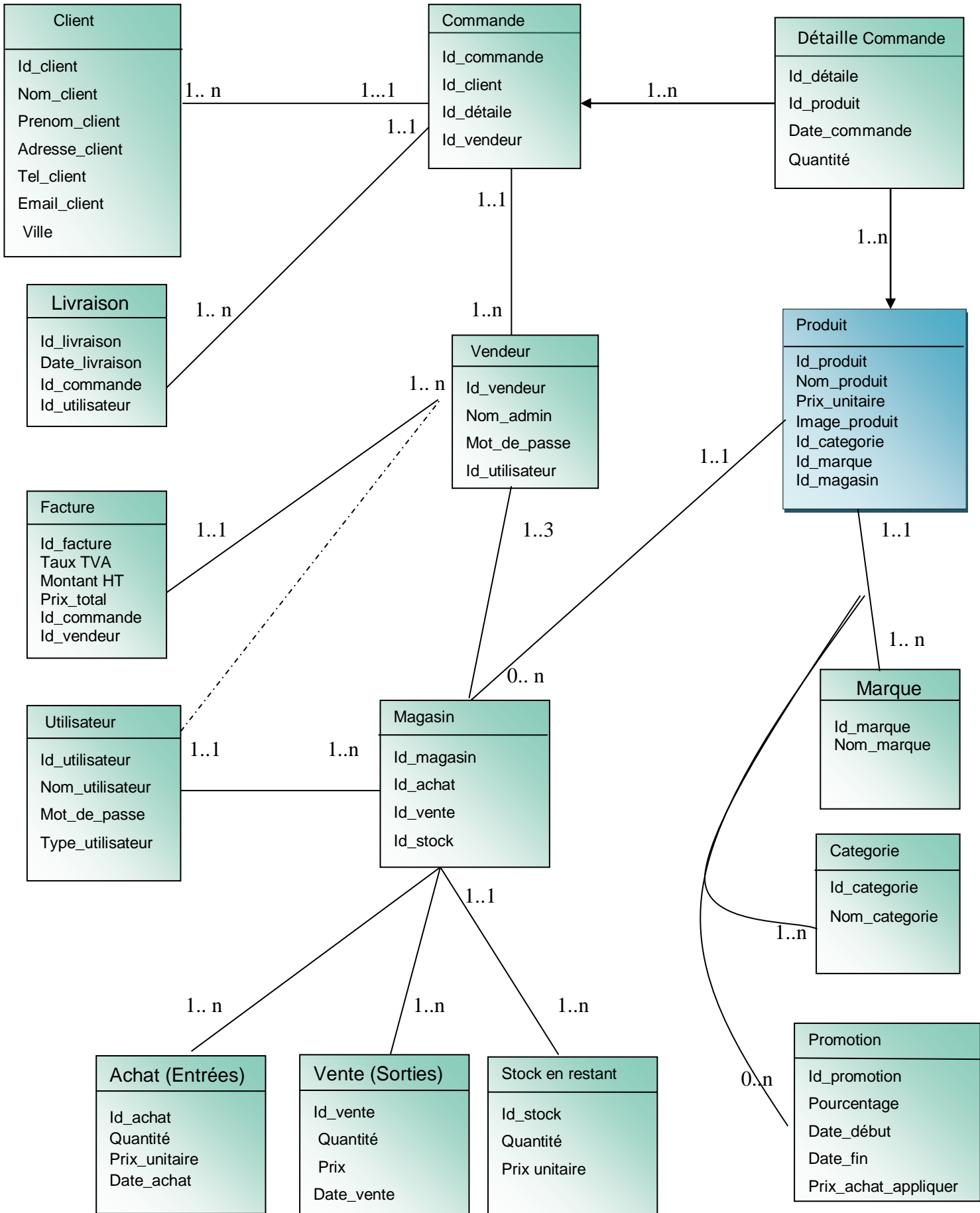


Figure 12 : diagramme de classe de notre BDD.

III. Stratégie suivie :

Avant de lancer n'importe quel projet BI il est porté à poser des questions l'une des plus importantes comment prioriser les besoins de mon projet BI ? Pour pouvoir répondre à cette question, il a fallu mettre en place une stratégie et se fixer une démarche.

III.1. Méthode de collecte d'informations dans l'entreprise

Il existe plusieurs techniques de collecte d'informations, les questionnaires, les entretiens, la documentation, et les réunions. Dans ce qui suit nous allons présenter les techniques auxquelles nous avons utilisé lors de notre étude :

a) Les réunions et les questionnaires :

En les a faits entre nous, et on a eu recours à des recherches et des comparaisons avec d'autres entreprises afin de bien clarifier l'objectif de notre travail.

b) La documentation :

Pour mieux comprendre les besoins et résoudre les problèmes rencontrés lors des réunions, nous avons utilisé plusieurs documents, qui donnent une vision claire sur les différents processus concernant notre projet, à savoir (la vente, les achats ...).

Notre démarche consiste à nous familiariser avec toute la littérature liée au fonctionnement de l'entreprise (vente, achat). Afin de pouvoir arriver à cibler quelques intérêts en fonction des besoins de l'administrateur, on a eu appel à l'analyse décrite ci-dessous :

III.2. Analyse des besoins :

« Analyse : c'est un processus d'examen de l'existant ». Après la collecte d'informations nous avons pu cibler quelques analyses que nous souhaitons étudiées présentées ci-dessus.

➤ Analyse des ventes (des achats) :

Le but d'analyse vente est d'avoir un affichage dynamique des statistiques et graphes pour les ventes (achats) effectuées dans les points de vente avec un filtrage par date, et prix TTC.

➤ Analyse des produits :

Répartition des produits par prix ainsi la quantité restant dans le stock avec un indicateur de marque qui représente l'état du stock pour chaque point de vente plus un filtrage dynamique des produits selon leurs dates d'expiration et promotion.

➤ **Analyse commerciale :**

Suivre les opérations de vente effectuées par le vendeur avec les détails et les opérations effectuées pour chaque caisse ainsi que l'évolution du chiffre d'affaires par année.

➤ **Analyse Clients :** Répartition des clients selon l'âge, le sexe, et le type d'abonnement.

III.3. Diagramme des cas d'utilisation

Si joignent des diagrammes de cas d'utilisation (CU) pour identifier tout type d'analyse effectuée par le participant principal dans l'entreprise.

III.3.1. Diagramme (CU) qui présente les fonctionnalités générales

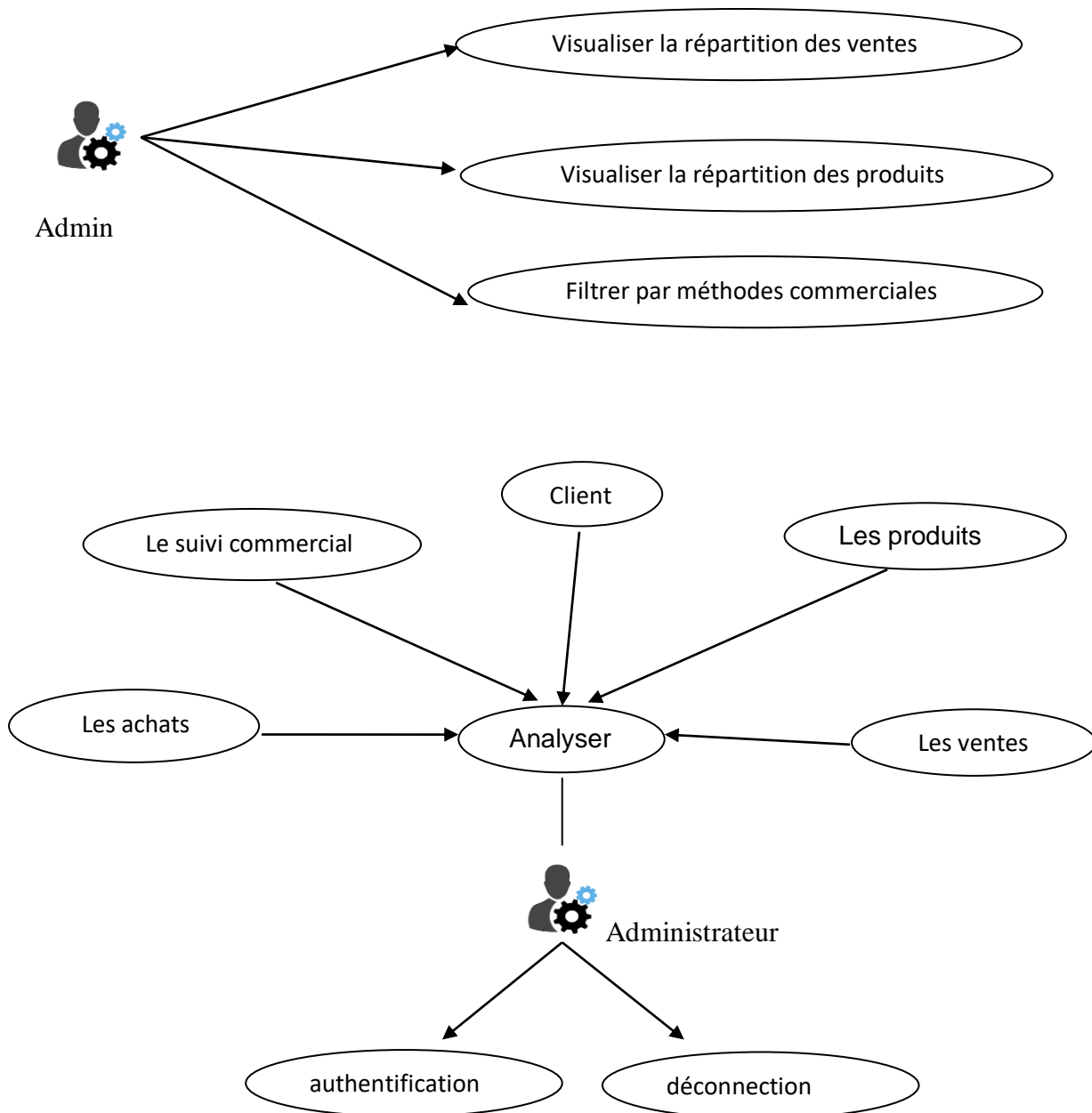


Figure 13 : Diagramme (CU) des fonctionnalités générales de l'entreprise.

III.2. Diagramme (CU) pour l'analyse commercial :

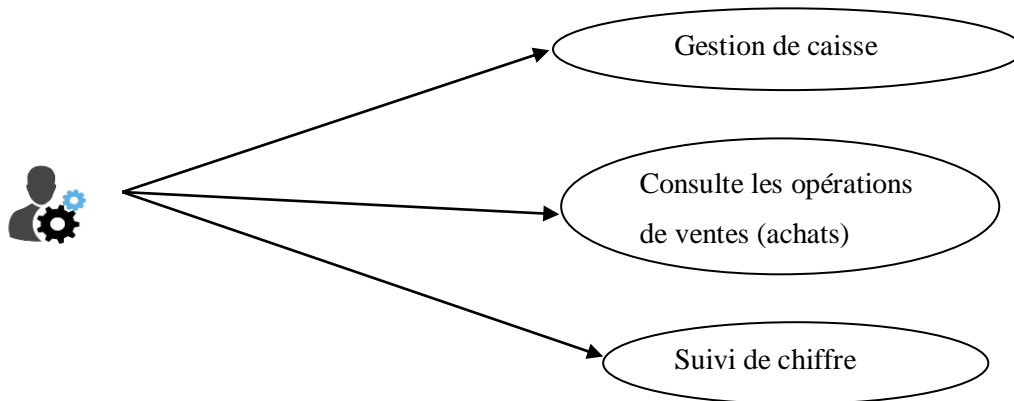


Figure 14 : Diagramme (CU) d'analyse commercial.

III.3. Analyse des produits :

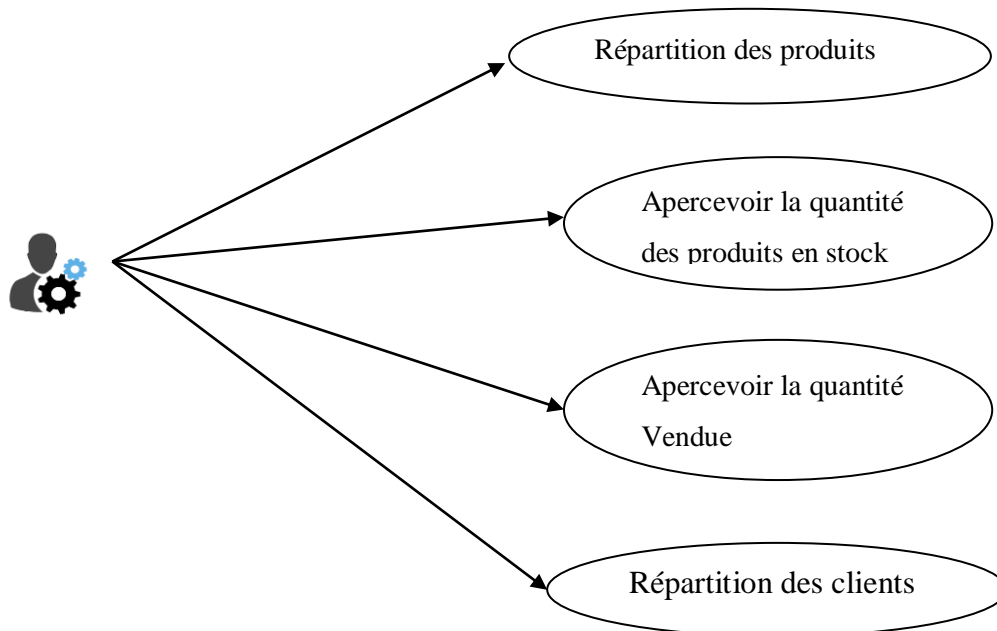


Figure 15 : Diagramme (CU) d'analyse des produits

IV. Spécification des besoins :

Dans ce qui suit nous nous sommes chargés de détailler les spécifications des besoins :

IV.1. Détermination des besoins fonctionnels :

✚ Les besoins de l'entrepôt de donnée

La conception de l'entrepôt doit permettre :

- ❖ Créer une seule source de données unique et non volatile.
- ❖ Meilleur contrôle des données : nous pouvons organiser les données selon nos besoins.

- ❖ Rendre la base de données plus évolutive et mieux adaptée à nos futurs besoins.
- ❖ Garder la traçabilité de chaque donnée : la source de ces informations.
- ❖ Effectuer une modélisation multidimensionnelle des données.
- ❖ Assurer que les données sont régulièrement mises à jour et actualisées régulièrement.
- ❖ Gérer un grand nombre de données différentes.
- ❖ Du datamining: l'outil doit permettre de faire du datamining, car même si ce n'est pas inclus dans notre projet, mais l'entité SI et qualité souhaite l'intégrer par la suite.

Les besoins de l'application

Les besoins de l'application se résument de manière à permettre :

- ❖ La disponibilité des données à n'importe quel instant.
- ❖ Interrogation les données de l'entrepôt à n'importe quel niveau hiérarchique.
- ❖ La consultation aisée des tableaux de bord fournis.
- ❖ L'utilisation facile des différentes opérations sur les données à travers des rapports et des tableaux de bords lisibles.
- ❖ La manipulation des interfaces par les décideurs afin d'explorer leurs données.
- ❖ Assurer la sécurité et la performance de l'entrepôt.
- ❖ Un ETL multi-sources.
- ❖ Des rapports statiques et dynamiques qui permettent de découvrir le total du projet.
- ❖ Des tableaux de bord.
- ❖ Des analyses multidimensionnelles.

IV.2. Détermination des besoins techniques

L'étude qui a conduit à l'élaboration de ces spécifications s'est basée sur les différentes exigences représentées ci-dessous :

- ❖ Les outils utilisés pour le développement de système doivent être des outils open source.
- ❖ Le système doit utiliser un système de gestion de la base de données.
- ❖ Le système doit être une solution web accessible via les navigateurs Mozilla Firefox, Google Chrome et Internet Explorer.
- ❖ Le système doit assurer un accès rapide à données simulées.
- ❖ Les accès au système doivent être gérés par une authentification.
Et une déconnexion automatique après 30 mn à 1h de non utilisation.
- ❖ Le système doit limiter l'accès aux données selon la hiérarchie de l'organisme.

IV.3. Détermination des besoins non fonctionnels :

Les besoins non fonctionnels se résument dans les points suivants :

- **Performance** : L'application répond à toutes les exigences des internautes d'une manière optimale.
- **Fiabilité** : Bon fonctionnement de l'application sans détection de défaillance.
- **Rapidité** : La navigation entre les pages doit être facile et rapide.
- **Sécurité** : Déconnexion après un tel temps d'inactivité.
- **Convivialité** :
 - ✓ Un design clair.
 - ✓ Une bonne interface qui donne aux internautes l'envie d'y utiliser.
 - ✓ Des notions faciles à comprendre.
 - ✓ Respect de l'ergonomie.
 - ✓ Positionnement du contenu sur les pages de la manière la plus accessible.
- **Portabilité** :
 - ✓ Elle fonctionne sur tout type de terminal Une application adaptative aux écrans.

V. Description de la solution

Dans ce chapitre nous allons présenter la solution qu'on a choisie pour réaliser l'objectif de notre travail. Après avoir étudié deux solutions et on a pu se familiariser pour l'une d'eux.

- La 1ere c'est de construire un système décisionnel en se basant sur l'entrepôt de données, on passe à travers le processus ETL.
- La 2eme consiste à prendre les données de la source et de les transformer directement en décisions (en tableaux de bords, en rapports ...).

La solution retenue est la 1ere voir quelle repend beaucoup plus à mos besoins. (Que ça soit de côté d'historisation, d'organisation des données par besoin, les données ne sont jamais supprimées ou modifiées (non volatile)). Ci-dessous on détail les étapes de cette solution :

V.1. Description générale de la solution basée sur un entrepôt de données :

La solution consiste à concevoir un DW qui sera alimenté à partir des différentes sources de données, dans notre cas ils sont sous forme de base de données et un fichier Excel pour le suivi des achats. Cette phase d'extraction et d'alimentation sera effectuée par un ETL, dans laquelle les données sont d'abord collectes, extraites des différents systèmes sources, chacune avec son propre format. Ensuite, transformées en un format qui peut être utilisable dans tout le système. Cette étape

consiste principalement à supprimer les données redondantes ou erronées, et la vérification de l'intégrité des données.

Les données converties, ne seront pas directement chargées dans l'entrepôt de données cible, mais plutôt dans une base de données de transfert appelé « Staging Area ». Ce qui évite leur perte, et facilite leur récupération en cas de problème. Les données sont chargées, et alimentées dans l'entrepôt de données. Enfin, un tableau de bord stratégique et un outil d'analyse OLAP sont déployés. Ces outils d'analyse permettent aux utilisateurs finaux d'exploiter ces données.

V.2. Critique de la solution :

Le seul inconvénient pour cette solution est la nécessité d'ajouter de nouvelles ressources matérielles en espace mémoire et de stockage. Car cette solution s'appuie sur l'historisation et la traçabilité des données ce qui génère un grand besoin d'espace mémoire. Cependant elle présente plusieurs avantages :

- ✓ Séparation du système décisionnel du système opérationnel,
- ✓ Les données ont une excellente qualité grâce aux traitements effectués dans le processus ETL, et cela rend les analyses plus convaincues.
- ✓ Exploiter de l'analyse OLAP, qui représente l'essentiel des besoins des utilisateurs finaux,
- ✓ Stockage des requêtes complexes dans le serveur OLAP pour le tableau de bord et les afficher sans revenir au système source.
- ✓ Adapter à toute nouvelle demande, qu'il s'agisse d'une source de données ou rapport d'analyse,
- ✓ Fusionner et agréger les données de toute l'entreprise, et réaliser des analyses multidimensionnelles à différents niveau d'agrégation.

V.3. Structure détaillée de la solution :

Les étapes de l'entrepôt de données doivent être conçues de façon que la bonne information soit facilement accessible au bon moment. L'architecture choisie est représentée par un modèle logique sous forme de 4 couches ou chaque couche jouera un rôle spécifique, produisant une sortie pour la couche suivante :

- Couche source de données,
- Couche de préparation de données (Data Staging area),
- Couche de stockage de données (DW),
- Couche de présentation et restitution de données.

Le tableau suivant présente l'architecture de la solution retenue :

Coches	Description
Source de données	<p>Cette couche fait représente différentes sources de données dans différents formats des systèmes opérationnels prêtes à être alimenter dans l'entrepôt de données.</p> <p>Dans notre cas (sous forme de BDD et fichier Excel)</p>
Data Staging area DSA	<p>Cette couche est celle pour la préparation de données (Data Staging Area) qui est principalement utilisée pour extraire rapidement les données des sources d'informations. Une fois les données chargées dans la DSA, l'étape de transformation entre en jeux (suppression de la redondance des données, filtrage des données incorrectes) et la validation des données avant de les transférer vers le DW.</p>
Stockage de données	<p>Elle représente la DW pour stocker des informations analytiques et historiques qui seront utilisées, Pour notre solution décisionnelle, l'approche de modélisation de la zone d'entreposage est détaillée dans le chapitre suivant.</p>
Restitution de données	<p>Cette couche fait référence aux informations qui parviennent aux utilisateurs finaux, un rapport par exemple, grâce à cette couche l'utilisateur sera capable de :</p> <ul style="list-style-type: none"> ✓ Analyser les données, notamment avec les outils de type OLAP pour les analyses multidimensionnelles, ✓ Aide à la décision des décideurs en situation avec les tableaux de bord présentant les indicateurs clés de l'activité, Transmettre l'efficacité avec le Reporting.

Tableau 3 : présentation l'architecture de la solution.

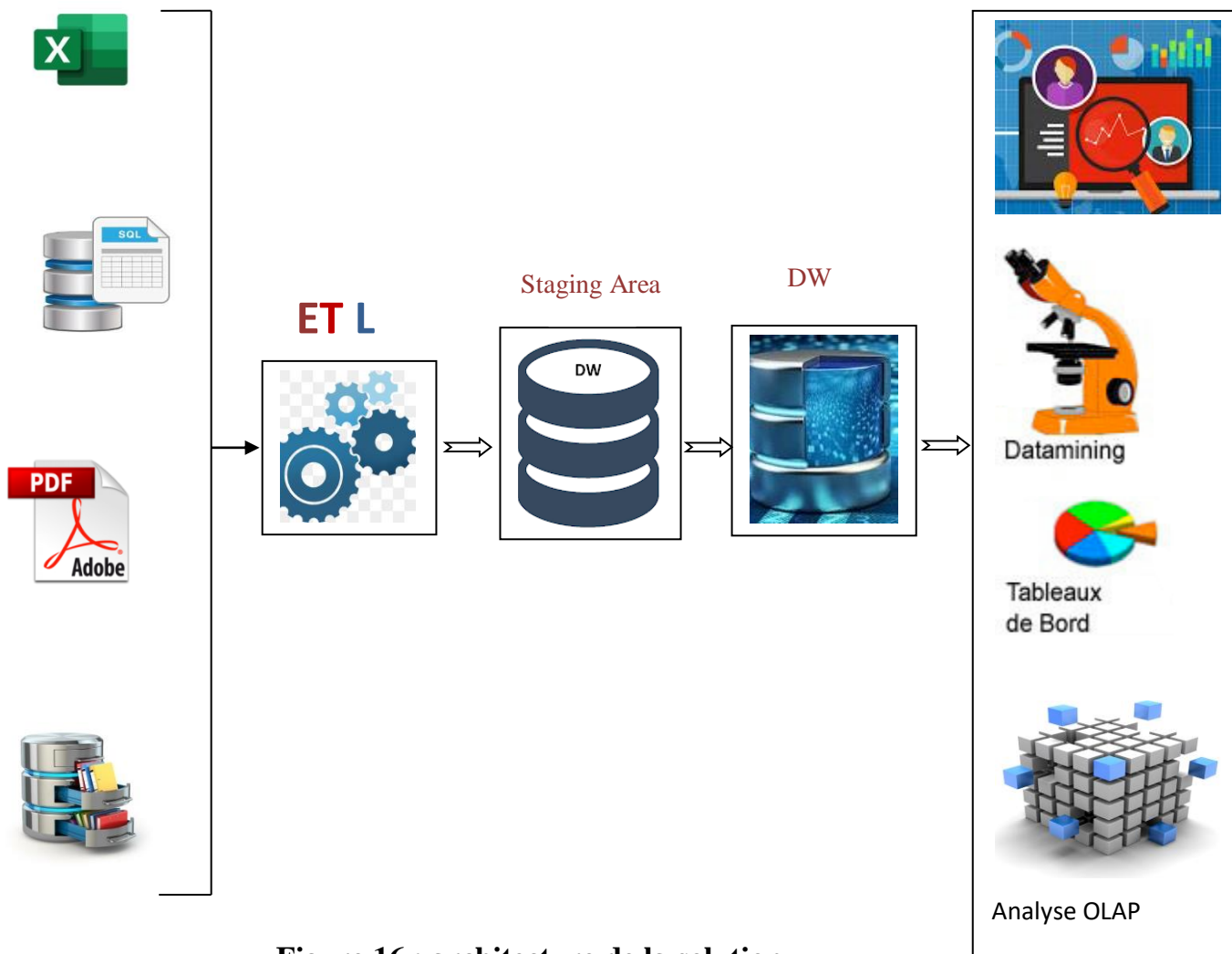


Figure 16 : architecture de la solution.

Conclusion

A travers ce chapitre nous avons discuté le choix de la solution pour répondre aux besoins attendus. Ensuite nous avons défini l’architecture de la solution retenue. Dans le prochain chapitre nous abordons la phase de conception.

Chapitre III :

Conception de la solution

A. Conception du DW :

Pour une meilleure conception nous avons fixé une démarche, et on a mis une stratégie à suivre. Suite à cette démarche nous avons choisis le schéma de modélisation dimensionnelle, à partir duquel nous avons sélectionné les activités à modéliser.

I. Choix de l'approche de modélisation

Avant de lancer la construction d'un DW il faut choisir une méthode de conception (top-down ou bottom-up ou middle-out). Selon l'analyse que nous avons fait, nous avons choisi de travailler avec l'approche **bottom-up** de Kimball déjà détaillée, dite ascendante. Dans le but des raisons suivantes :

- ❖ créer les data-marts, puis les rassembler par utilité ou par fonction afin de construire le DW.
- ❖ Rassembler tous les indicateurs clés; en se basant sur les besoins des utilisateurs.
- ❖ Moins couteux en termes de temps, par contre la réalisation d'un DW dure des années.
- ❖ Facilité de mise en œuvre.



Figure 17 : architecture de l'approche de modélisation.

II. Définition des activités à étudié

Selon l'étude de besoins, et le rassemblement des indicateurs clés, et nous sommes arrivés à définir les éléments d'analyses suivants :

- « Suivi et analyse des vente ».
- « Suivi des achats ».
- « Suivi des produits en stock ».
- « Suivi des produits ».

Le diagramme suivant présente la répartition des éléments d'analyses recensés selon les deux axes cités (importance, facilité) :

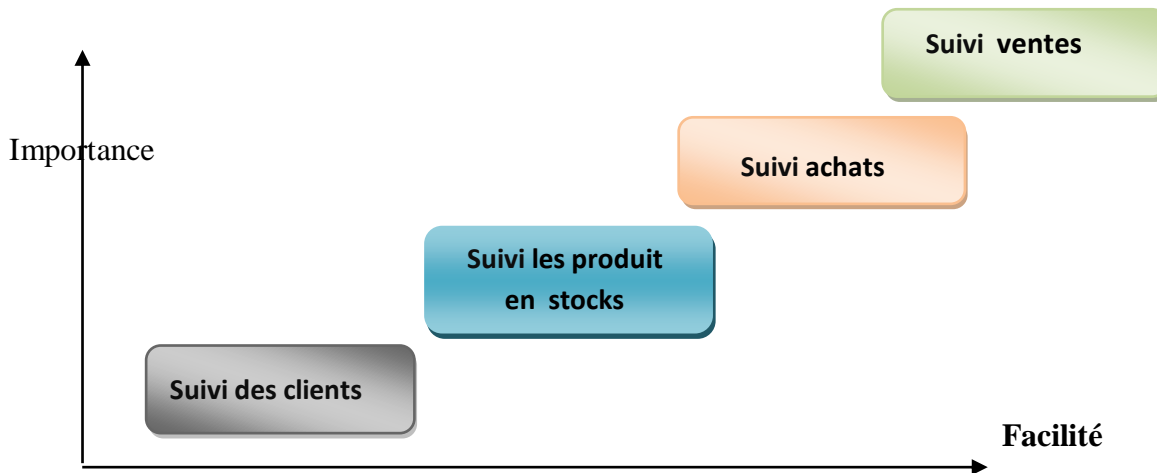


Figure 18 : diagramme de priorité d'analyse

III. Choix d'Indicateurs :

Premièrement, il est important d'établir des indicateurs corrects pour les activités que l'on veut étudier. Et sa nécessite une meilleur compréhension des informations. Chaque indicateur sera caractérisé, et sélectionné selon les critères suivant:

- ✓ Un indicateur doit être disponible en temps réel.
- ✓ L'indicateur doit mesurer un ou plusieurs objectifs.
- ✓ L'indicateur doit conduire à une action.
- ✓ L'indicateur doit être constructible.

ID	Indicateurs	Formule avec les attributs
1	Meilleur produits vendus par année et par trimestre	Max (quantite_vente) par produit
2	Division les produits vendu par catégorie	Σ (quantite_vente) par catégorie
3	Les clients non facturés	Aucune facture n'a été établie
4	Total de produit vendu par année	Σ (quantite_vente)
5	Le chiffre d'affaire par date	Σ montant ttc

6	Quantité en stocke par produit	Quantite_stock
7	Quantité acheté par date	Σ quantite_achat
8	Les clients facturés	Facture établie
9	Répartition des clients par région	Selon leur ville (dans notre cas Tizi-Ouzou, Alger et Bejaia)
10	TTC global	Σ TTC
11	Apport des clients	Le nombre total d'abonnés
12	Taux d'intégration d'abonnés	Le rapport entre le nombre des abonnés intégrés et le nombre des abonnés prévus à être intégrés.
13	Répartition des clients par âge	Dans notre cas on a divisé en 3 (≤ 30) ($>30, <50$) (≥ 50)
14	Répartition des clients par sexe	Homme ou femme

Tableau 4 : choix d'indicateurs.

IV. Schémas de modalisation :

Nous avons choisi le schéma en étoile défini comme étant une structure dimensionnelle, qui représente une seule table de faits entourée par un certain nombre de dimensions. Nous avons choisi l'utilisation d'un schéma en étoile pour les raisons suivante :

- ⊕ Une meilleure performance et la simplicité de compréhension,
- ⊕ une précision pour la mise en œuvre de la solution.
- ⊕ La création facile des cubes OLAP.
- ⊕ Le tableau ci-dessous est une étape préalable pour la construction du DW, il permet d'identifier facilement les dimensions réparties par différents processus.

Dimensions	Suivi des ventes	Suivi des achats	Suivi stock	Suivi des produits
Catégorie				X
Marque				X
Date	X	X		
Client	X			
Destination	X			
Détail_commande	X			
Magasin			X	
Stock			X	
Produit	X	X	X	X

Tableau 5 : matrice en bus pour identifier facilement les dimensions réparties.

Après avoir fini avec la construit de le matrice, nous présentons la conception suivante:

IV-1. La dimension date :

La dimension temps est une dimension faite pour assure l'historisation, elle contient toutes les dates qui synchronique avec un fait donné. Les instances de cette dimension ne sont pas extraites à partir des systèmes sources. Elles sont générées grâce à un script qui sera réalisé lors du chargement initial.

IV-2. Construire les datamarts (les étoiles)

À partir des indicateurs si dessous et des activités étudié on constate le faite de construire 4 étoiles (data marts) (Vente; Produit ; Achat ; magasin).

IV-2-1. Le suivi des produits :

a) Tables de dimensions :

Marque	Catégorie	Promotion
Id_marque	Id_categorie	Id_promotion
Nom_marque	Nom_categorie	Pourcentage
		Date-début
		Date_fin

Tableau 6 : Table de dimensions pour le suivi des produits.

b) **Table de fait :**

Table de fait	Clés
Fait_produit Clé (Id_produit)	Id_produit Id_marque Id_catégorie Prix_unitaire

Tableau 7 : description de la table de fait « produit »

c) **schéma en étoile (data mart) :**

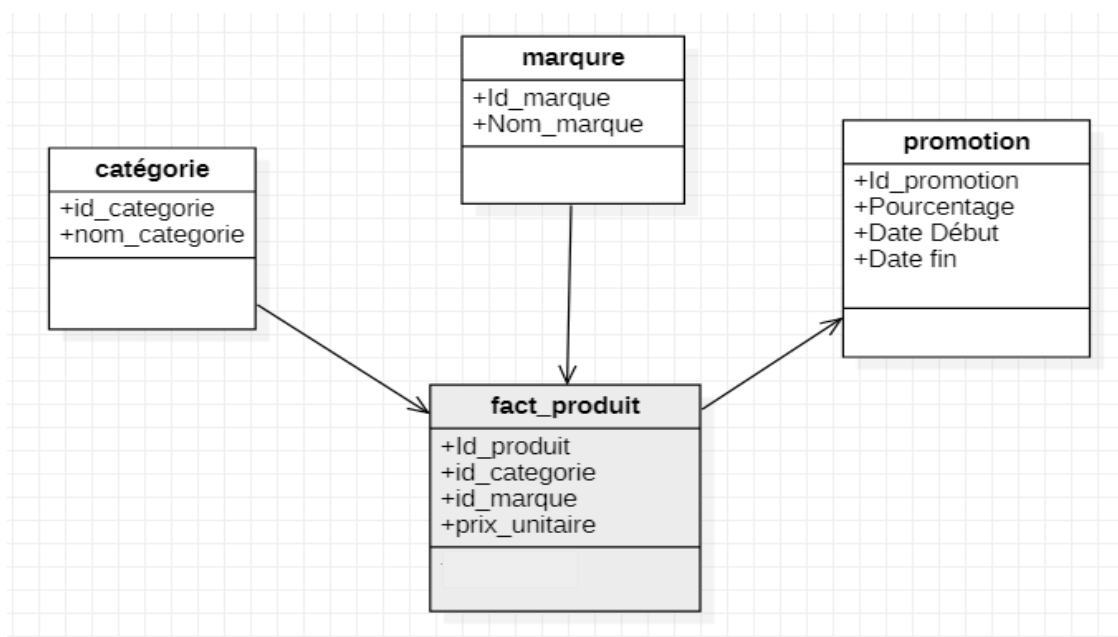


Figure 19 : schéma en étoile fact_produit.

IV-2-2. Le suivi commercial des ventes :

a) **Tables de dimensions :**

Produit	Date_vente	Client	Détail_command	Destination	Facture
Id_produit	Id_date	Id_client	Id_détail	Id_ville	Id_facture
Id_catégorie	Jour	Nom_client	Id_produit	Nom_ville	Taux_tva
prix_unitaire	Mois	Adresse	Quantité		Prix_total
Id_marque	Année	Age			Id_commande
Quantite()		Sexe			

Tableau 8 : Table de dimensions pour le suivi ventes.

b) **Table de fait :**

Table de fait	Clés	Mesures
Fait_vente Clé (Id_vente)	Id_vente ; Id_produit Id_client ; Id_detail Id_date_vente ; Id_distination Id_facture	Prix_total Quantité TTC Chiffre affaire

Tableau 9 : description de la table de fait « vente »

c) schéma en étoile (data mart)

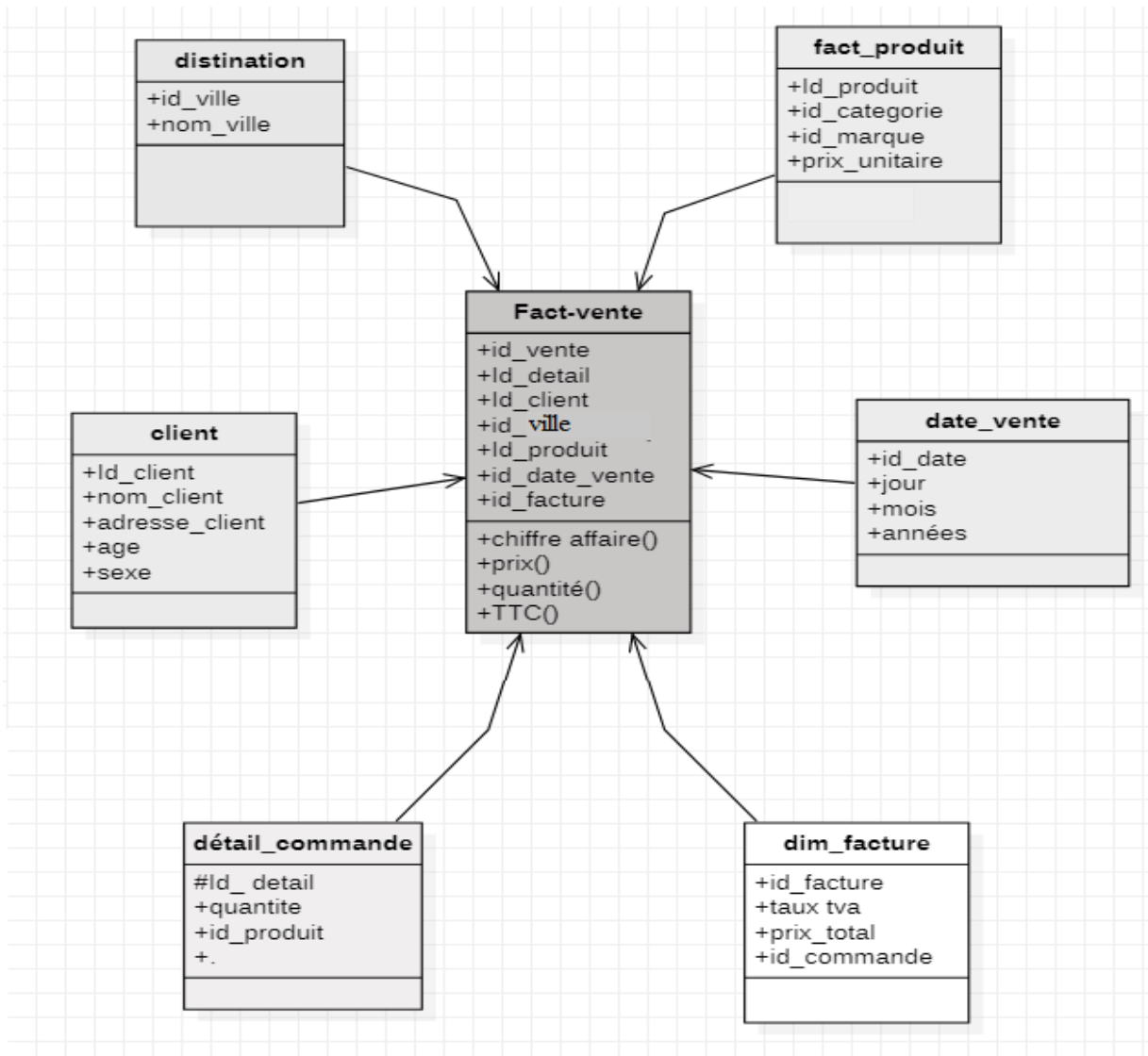


Figure 20 : schéma en étoile fact_vente.

V-2-3. le suivi des achats :

a)- Tables de dimensions :

Produit	Date_achat
Id_produit	Id_date
Id_catégorie	Jour
prix_unitaire	Mois
Id_marque	Année
Quantite()	

Tableau 10 :Table de dimensions pour le suivi des achats.

b)- Table de fait :

Table de fait	Clés étrangères	Mesures
Fait_achat	Id_achat	Prix_total
Clé (Id_achat)	Id_produit	Quantité
	Id_date_achat	
	Id_magasin	

Tableau 11 : description de la table de fait « achat »

c)- schéma en étoile (data mart)

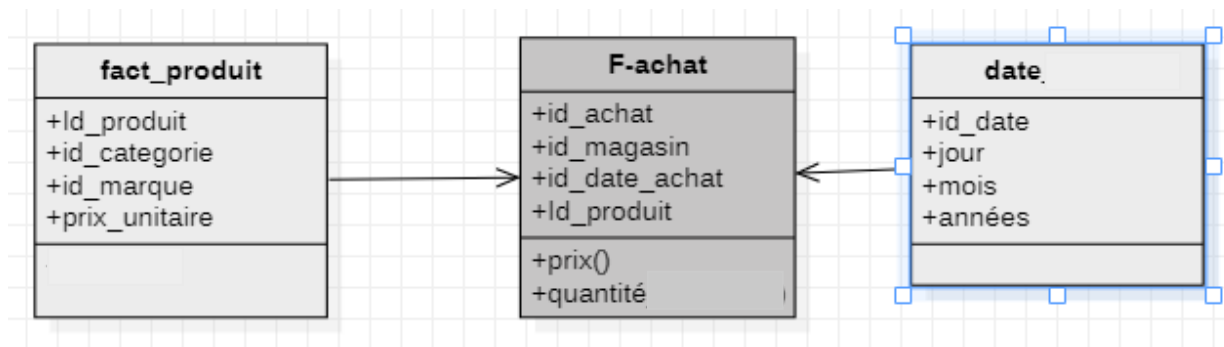


Figure 21 : schéma en étoile fact_achat.

V-2-4. Le suivi du magasin

a)- Tables de dimensions

Fact_produit	Stock
Id_produit	Id_stock
Id_categorie	Quantite_stock
prix_unitaire	Prix_unitaire
Id_marque	
Quantite()	

Tableau12 : Table de dimensions pour le suivi du magasin

b)- Table de fait :

Table de fait	Clés étrangères	Mesures
Fait_magasin Clé (Id_magasin)	Id_produit Id_stock Id_magasin	Prix_total Quantité_magasin

Tableau 13 : description de la table de fait « magasin »

c)- schéma en étoile (data mart)

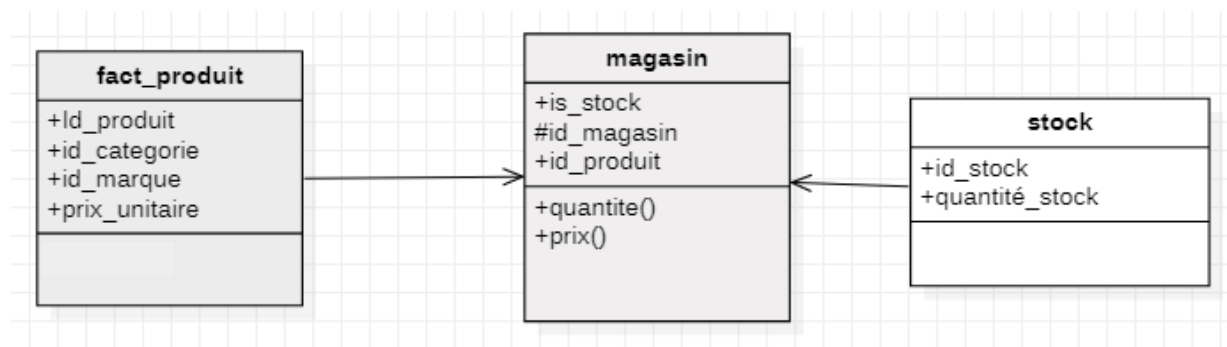


Figure 22 : schéma en étoile fact_magasin

V-3. Regroupé de DataWarehouse

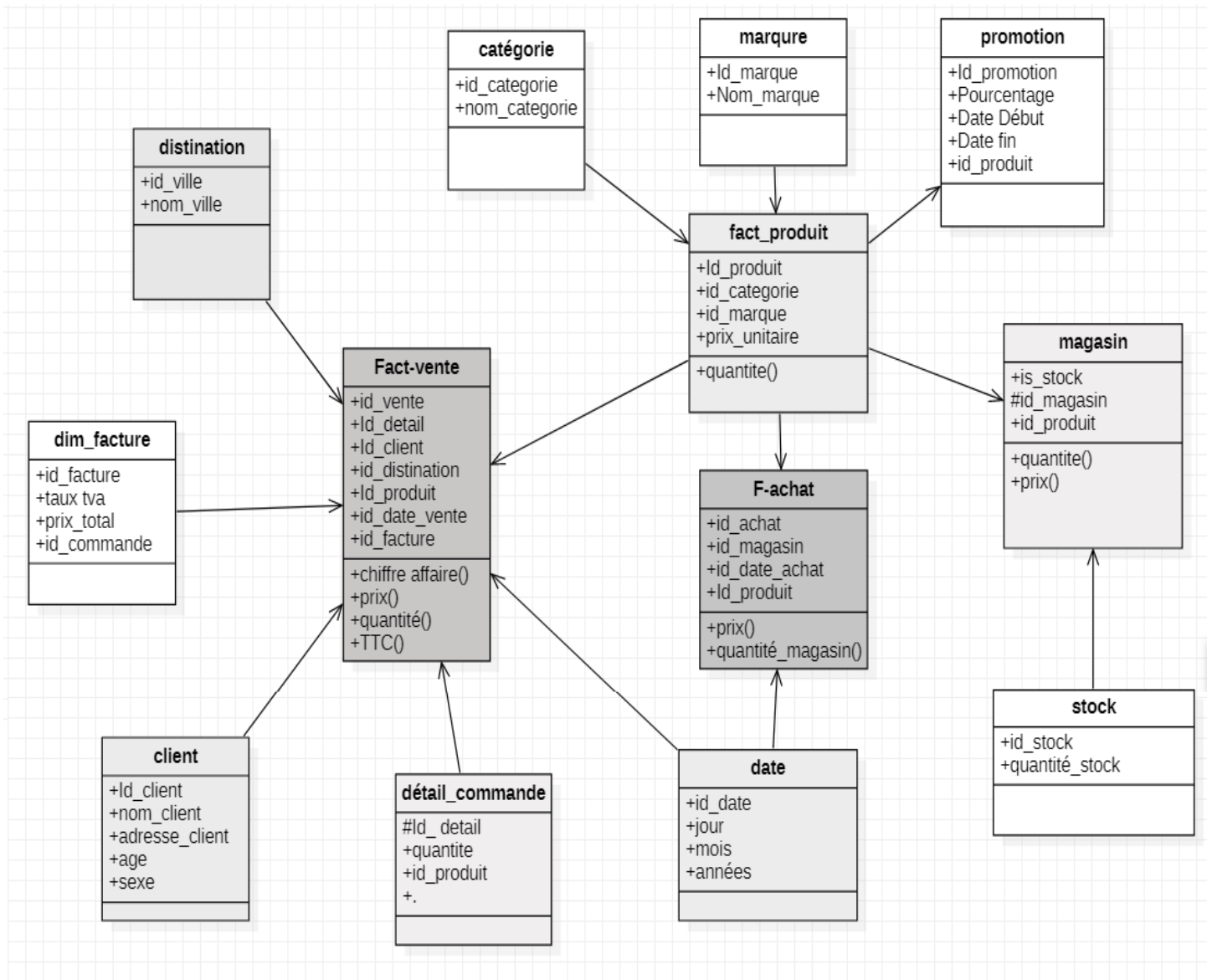


Figure 23 : schéma du DW en flocon

B. Alimentation du DW :

Ce passage est consacré pour détaillé le processus de la phase d'alimentations du DW. Le chargement de chacune des tables de faits et de dimensions, à partir de données venues d'une ou plusieurs sources, après avoir suber un certain nombre de transformation par le processus dit ETL.

I. Etapes de phase d'alimentation

La phase essentielle d'un entrepôt de données est l'alimentation des données provenant de déférentes bases ou ressources. La conception de cette phase est une tâche complexe plus de 60 % de la charge totale d'un projet DW. Pour que cette phase s'effectue avec l'utilisation d'un outil du nom d'ETL. L'alimentation d'un entrepôt est structurée selon les points suivants :

I.1. Découverte des données :

Avant de commencer la phase d'alimentation du DW il faut identifier des données à importer dans l'entrepôt de donnée qui se trouve, par exemple dans notre cas dans des BDD et fichier Excel. Cette étape demande un travail et une compréhension sur les bdds sources.

I.2. Extraction des données :

Collection des données importantes dans les systèmes de production. Périodique et Répétée Ne pas perturber les applications OLTP.

I.3. Transformation des données :

C'est une suite d'opérations qui a pour but de Rendre les données cohérentes

- ❖ Convertir / simplifier les noms des attributs.
- ❖ Fileter.
- ❖ Suppression des doublant.
- ❖ Effectuer des calculs pour les valeurs numériques.
- ❖ Ajouter des attributs si nécessaire.
- ❖ Nettoyage (valeurs manquantes, aberrantes...).
- ❖ Réduire les valeurs d'attributs.
- ❖ Homogènes avec la structure d'un entrepôt de données c'est-à-dire les données des systèmes de production doivent être agrégées ou calculées avant chargement.
- ❖ Etc....

II.4. Chargement des données :

- ❖ C'est la dernière étape du processus ETL, consiste à transférer et insérer des données nettoyées, préparées qui sont stocké au sein du **Staging Area** ver l'entrepôt de données.
- ❖ C'est une opération qui peut être longue.
- ❖ Mettre en place des stratégies pour assurer de bonnes conditions à sa réalisation.
- ❖ C'est une phase plutôt mécanique et la moins complexe.

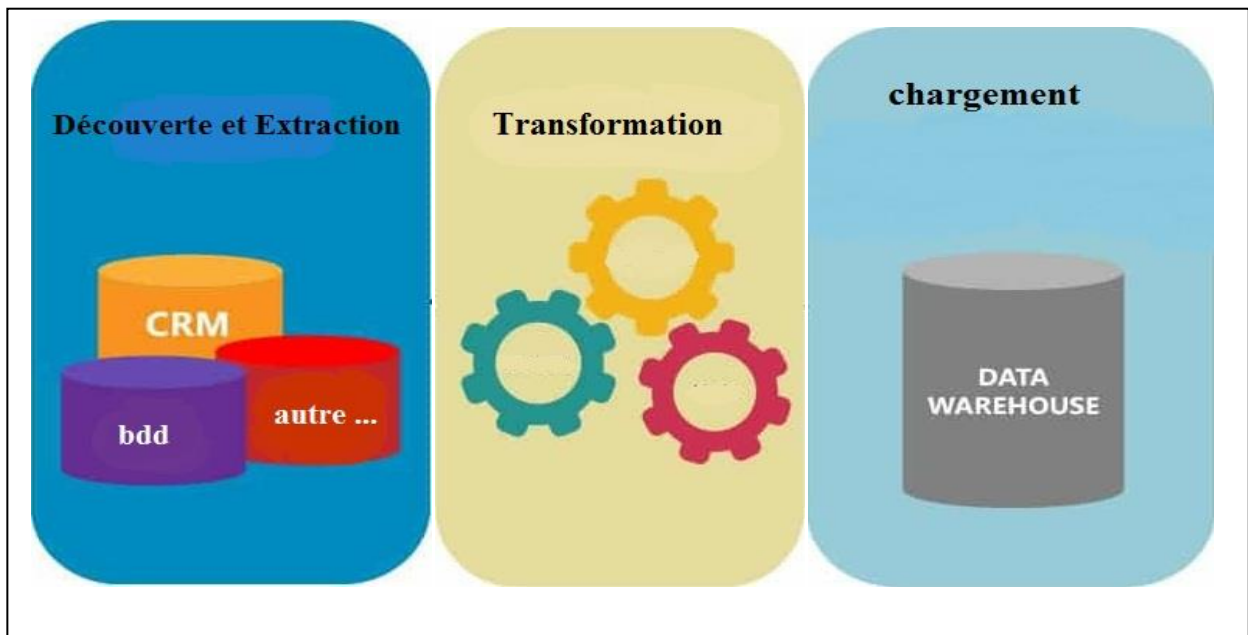


Figure 24 : phase de processus ETL.

C. Construction de cube OLAP :

L'idée est d'extraire les données constituant le cube une seule fois de l'entrepôt de données stocké en colonnes, et d'appliquer ensuite un ensemble de traitements pour calculer tous les agrégats possibles, à différents niveaux de granularité. Une fois le cube est calculé, il est matérialisé suivant l'architecture orientée colonnes pour permettre à l'analyse OLAP de bénéficier des avantages de cette architecture quant à la manipulation du cube (Slice, Dice, Drill-Down, Drill-Up, ..). Plus précisément, cette approche est exécutée en trois phases :

Phase 1 :

Extraction des données Cette phase consiste à extraire les données à partir de l'entrepôt de données. Le résultat de cette phase est une relation R composée de valeurs de dimensions et de la mesure qui satisfait les prédicats de la requête. Avec un entrepôt de données implémenté suivant l'approche orientée colonnes, cette phase est plus performante qu'avec un entrepôt de données implémenté

suivant l'approche orientée lignes. En effet, le mécanisme de stockage de l'architecture orientée colonnes permet un accès rapide aux données des différentes colonnes et diminue considérablement le coût de jointure comparé à celui de l'architecture orientée lignes.

Phase 2 :

Construction du treillis de cuboïdes Cette phase consiste à construire le treillis de cuboïdes et définir les combinaisons possibles du cube à calculer.

Phase3 :

Calcul et matérialisation du cube Cette phase consiste à parcourir le treillis de cuboïdes pour calculer le cube en fonction de la relation obtenue de la première phase, ainsi, le cube est matérialisé colonne par colonne.

I. Démarche d'utilisation d'OLAP :

De nombreuses recommandations impliquent la définition d'opérations à l'aide de cubes de données. Les opérations de forage et de rotation, reposent sur la métaphore d'un cube des données sont les plus difficiles.

I.1. Operation de forage « roll-up » & « drill-down »:

Les opérations de forage sont naviguées selon une hiérarchie d'axes d'analyses, afin de permettre une analyse plus ou moins précise des indicateurs La synthèse impliquent l'analyse des données à un niveau de granularité moins détaillé. Inversement, lorsque vous explorez vers le bas, vous pouvez effectuer une analyse plus fine des données.

I.2. Opération de rotation :

Les opérations de rotation réorientent une analyse. L'opération la plus courante consiste à changer l'axe d'analyse en cours d'utilisation (rotation de dimension). Un cas particulier de la rotation de dimension consiste à changer de perspective d'analyse (rotation de hiérarchie). Dans ce cas, l'axe en cours d'utilisation est maintenu, mais la manière de le graduer est changée.

I.3. Autres Opération :

Il existe autres opérations qui assurent la manipulation des données multidimensionnelles qui sont classées dans l'ordre suivant :

- ✓ **Les opérations de restriction « slice » et « dice »** : Permettent à un utilisateur de restreindre l'ensemble des données analysées en deux spécifications « slice » et « dice ». La première

assure une restriction sur l'un des axes d'analyse. Et la deuxième assure les données d'un indicateur d'analyse.

- ✓ **Les opérations de transformation « pull »** : Permettent l'ajout d'attributs de dimension en tant qu'indicateur d'analyse « push », ou de convertir un indicateur d'analyse en paramètre « pull ».
- ✓ **Les opérations d'ordonnement** : Permettent de changer l'ordre des valeurs (positions) des paramètres des dimensions (« switch ») ou de réordonner les paramètres d'une hiérarchie, cette dernière permet d'imbriquer un attribut dans une autre hiérarchie.

Conclusion :

Ce chapitre en globe la phase de conception de l'entrepôt de données. et nous avons présenté l'une des plus importantes étapes de conception d'un projet BI ; elle représente plus de 60 % de mise en œuvre. En effet, la conception de la zone d'alimentation garantira le transfert des données vers l'entrepôt de données. Et cela ce fait grâce à plusieurs procédures de conversion, de nettoyage, de correction et ce Mise en contexte des données.

Au début nous avons identifié les sources de données, leur endroit, ce qui nous a permis à définir l'architecture d'alimentation. Ensuite nous avons détaillé la stratégie d'extraction des données, puis nous avons présenté les techniques de transformation pour assurer le transfert des données vers l'entrepôt. Le contenu du prochain chapitre se porte sur la réalisation et mise en place de notre solution.

Chapitre IV :

Réalisation et mise en œuvre de la solution

A. Présentation des outils utilisés:

À travers ce chapitre nous allons détailler l'architecture et la construction de la solution. Puis étudier les différents outils qui existent pour la mise en œuvre de cette solution. Et choisir l'outil qui satisfait le plus nos besoins. Cette étape de comparaison on va l'appliquer que ça soit sur les outils ETLs ou de restitutions.

I. L'architecture et la construction de la solution :

La mise en œuvre de la solution s'effectue au niveau du poste de travail (réalisation). Pour déployer la solution on aura besoin de 2 serveurs :

- Le serveur MySQL du Wamp serveur pour héberger la DW et staging area, et la BDD de l'application.
- Un serveur de Reporting, accessible que par l'utilisateur final via les navigateurs web (chrome, mozilla, ...). Ou des applications de Reporting configuré et installer sur leur machine.
- Seul l'administrateur aura l'accès à ce serveur dans le but de garantir la sécurité des données.

Vu que dans notre cas, le travail se fait juste en local. Les données sont récupérées à partir des bases de données qui se trouvent dans le serveur MySQL (Phpmyadmin). On n'aura pas besoin d'installer des outils (logiciel) pour pouvoir récupérer ces données à distance.

Et dans le cas contraire, là où il y aura des données stockées à distance. La nécessité de cette installation s'impose ; pour extraire les données mises à jour pendant la non-activité du système, et les envoyer via un serveur FTP vers la BDD centre.

II. Choix des outils (logiciels) :

II.1. Choix du SGBD : nous avons opté pour **PHPmyadmin**.

PHPMyAdmin :

(PMA) est une application web de gestion pour les systèmes de gestion de base de données MySQL réalisée en PHP. PHPMyAdmin permet de :

- ✓ Créer/ supprimer de nouvelles bases.
- ✓ Créer/modifier/supprimer/copier des tables.
- ✓ Afficher/ajouter/modifier/supprimer des tuples dans des tables.
- ✓ Effectuer des sauvegardes de la structure et/ou effectuer n'importe quelle requête.

II.2. Les serveurs d'exécution:

1. Le serveur Wamp :

WampServer est un paquetage contenant à la fois deux serveurs (Apache et MySQL). Un interpréteur de script (PHP). Les deux bases SQL PHPMyAdmin et SQLite Manager pour gérer plus facilement les BDDs. Il permet d'installer automatiquement une plateforme permettant l'exploitation d'un site web en PHP qui éventuellement aurait besoin d'un accès à une base de données.

2. Le serveur MySQL:

MySQL est un serveur de base de données relationnel. Il fait partie d'un des logiciels de gestion de BDD les plus utilisés au monde ; par le grand public que par des professionnels. MySQL fait partie du quatuor LAMP : Linux, Apache, MySQL, PHP .Il appartient également à ses variantes WAMP(Windows). Le couple PHP/MySQL est très utilisé par les sites Web et proposé par la majorité des hébergeurs Web.

3. Le serveur Apache :

Est le serveur le plus répandu sur internet. Il a été créé pour une application conçue pour les systèmes d'exploitation de type Unix. Aujourd'hui, il est porté sur de nombreuses plates-formes, dont Windows. Logiciel open source, il est réputé pour sa fiabilité et sa stabilité, et il regroupe les applications suivantes :

- ✓ Le serveur web **Apache**.
- ✓ Le serveur de base de données **MySQL**.
- ✓ Apache Le serveur d'application **PHP**.
- ✓ L'outil **PHPMyAdmin**.

II.3. Les langages de programmation :

Nous avons utilisé les langages suivants pour la création de notre application de vente :

- ✓ Le langage HTML pour la partie statique.
- ✓ Le langage de requêtes MySQL pour l'interrogation de la base de données.
- ✓ Le langage javascript.
- ✓ CSS, pour la forme et le style.
- ✓ Les langages de script PHP.
- ✓ Bootstrap.

1. HTML :

L'**HTML** est un langage informatique utilisé pour créer des pages web. L'acronyme signifie ***HyperText MarkupLanguage***. Ce langage permet de réaliser de l'hypertexte à base d'une structure de balisage. Il permet également la structure sémantique ; logique du contenu des pages ; d'inclure des ressources multimédias dont des images, des formulaires de saisie. Il est souvent utilisé parallèlement avec les langages de programmation JavaScript et CSS.

2. Le PHP :

LE PHP est un le langage de script (côté serveur) utilisé pour la réalisation de sites Web dynamique. Langage très complet, il est principalement employé pour mettre des BDD en ligne dans les sites Web. Les principaux atouts du PHP sont :

- La simplicité d'écriture de scripts qui sont exécutés sur le serveur avant l'envoi aux internautes.
- La possibilité d'inclure le script PHP au sein d'une page HTML.

3. JavaScript :

JavaScript est un langage développé par Netscape. Les scripts sont ajoutés aux balises HTML s'exécutant dans le navigateur sans faire appel aux ressources du serveur permettant d'effectuer des contrôles de saisie pour valider les formulaires.

4. CSS :

Le CSS ou feuille de style en cascades (Cascading Style Sheets) est un langage qui sert à décrire la présentation des documents HTML et XML. CSS ont l'avantage de séparer les informations de la mise en forme du contenu et facilitent les mises à jour et les modifications des pages HTML.

5. Bootstrap :

Bootstrap est un Framework CSS ; qui embraque également des composants HTML et JavaScript. Il comporte un système de grille simple et efficace pour mettre en ordre l'aspect visuel d'une page web. Il permet de concevoir un site web rapidement.



Figure 25: les langages choisis pour l'application de vente.

II.4. Outil de collecte, ETL :

Il existe sur le marché beaucoup de solutions ETL open source :

- ✓ Talend open studio,
- ✓ Pentaho data integration,
- ✓ CloverETL ...

Qui s'occupe des phases : d'extraction des données des bases sources.

- ✓ puis transformation de ces données.
- ✓ alimenté la data warehouse.

Dans le but de choisir la solution la plus correcte, nous avons étudié deux outils, les plus fréquemment utiliser, talend et pentaho.

II.4.1.Talend :

Talend est un éditeur de logiciel qui permet l'intégration de données dans une seule plateforme unifiée, Permet : Il produit deux types de logiciel (commerciaux et open source (gratuit)). Et dans notre cas on va étudier talend for data integration TOS qui est de type open source

C'est un ETL contenant plus de 900 composants et connecteurs;

- ✓ L'extraction multi-sources: Oracle, SQL Server, MySQL, fichier Excel, fichier CSV...etc.
- ✓ Le chargement: Insertion ou mise à jour des tables, dimension à variation lente, alimentation fichier Excel,...etc.
- ✓ La transformation: Il permet de faire plusieurs types de transformations : fonctions de calculs, jointures, trie, recherche, exécution de script,...etc.
- ✓ Une interface intuitive.
- ✓ Concepteur de tâches.
- ✓ Modélisateur de métier.

L'IDE de développement est basé sur Eclipse et des connecteurs standards existent pour la majorité des sources de données, telles que les ERP, les BDD ou encore des solutions de commerce en ligne. Talend est capable de transformer les données vers presque tous les formats existants et si une transformation ou un connecteur vous manque il suffit de le développer très facilement en java.

II.4.2.Pentaho :

Pentaho est une suite de logicielle qui permet la distribution de fonctionnalités et documents décisionnels à des personnes par l'intermédiaire d'une interface Web ou par un outil de reporting. Il permet d'avoir les fonctionnalités suivantes:

- ❖ Un accès via une interface unique, aux différentes composantes décisionnelles de la suite.
- ❖ Un ETL contenant plus de 160 nœuds qui permettent :
 - ✓ L'extraction multi-sources: Oracle, SQL Server, MySQL, fichier Excel, fichier CSV...etc.
 - ✓ La transformation: avec plusieurs types : fonctions de calculs, jointures, trie, recherche, exécution de script,...etc.
 - ✓ Le chargement: Insertion ou mise à jour des tables, dimension, alimentation fichier Excel,...etc.
 - ✓ Une possibilité d'automatisation d'exécution du processus d'ETL à des horaires choisis.
 - ✓ Une conception et consultation des rapports statiques ou dynamiques, et une constitution des tableaux de bord.
 - ✓ Un envoi automatique par email d'un ensemble de rapports à leurs destinataires respectifs.

Il sera nécessaire d'étudier ces deux plateformes et de vérifier si elles répondent aux critères de nos besoins, avant de choisir une solution entre les deux. Même si leur fonctionnement est similaires mais il dispose d'un certain nombre de comparaison et le tableau ci-dessous illustre ces différences :

Caractéristique générales	Talend	Pentaho data integration
Années de création	2005 ou 2006	2004 ou 2006
Gestion de version	Oui	Non
Langage	Java et perl	Java
Version	Complète	Très complète
Création d'ETL	Oui	Oui
Jointures supportées	Interne et externe	Interne
Risque	Le niveau de risque est égal par rapport à Pentaho	Le niveau de risque est égal par rapport à Pentaho
Visualisation des traitements en temps réel	Oui	Non
Interface graphique	Oui	Oui
Gestion de contexte	Oui	Non
Analyse des données	Non	Non
Support technique	Libre-service	Non
Générateur de code	Oui	Non

Tableau 14 : Comparaison entre Talend et Pentaho.

D'après ce tableau de comparaison entre ces deux éditeurs, on a opté ver l'utilisation de Talend puisqu'il est l'outil le plus utilisé, et la plus part des entreprises lui fait plus confiance.

III. Outils de reporting :

Dans ce qui suit, on va définir trois outils open source les plus utilisés et qui peuvent répondre aux exigences des entreprises : Jasper Report, Pentaho, power BI. Pour bien choisir l'outil qui convient le plus, nous avons fait appel à l'étude comparative suivante.

III.1. Power BI : [15]

Power BI est une solution BI développée par Microsoft, il permet aux entreprises d'analyser et de visualiser des données résultant de différentes sources. Transforme les données en informations exploitables. Cela aide à prendre des décisions stratégiques. Microsoft Power BI est **un ensemble d'outils d'analyse décisionnelle**, il permet aussi de partager de façon sécurisée les informations sous forme de tableaux de bord. Cette solution permettre à tous les employés de comprendre les

données et de les exploiter ; pour fournir des rapports et des analyses à leurs entreprises d'augmenter leur productivité et leur créativité. L'outil Power BI Desktop permet notamment d'ordonner **des données en provenance de nombreuses sources** sur site ou sur le Cloud : bases de données, fichiers, services web. En utilisant l'outil Gateways, il est possible de connecter les bases de données SQL.

	Pentaho	Power BI	Jasper Report
ETL	√	X	√
Reporting	√	√	√
Tableaux de bord	√	√	X
Interface	√	√	√
Olap	√		√
Serveur BI	√	√	√
Gestion des utilisateurs	X		X

Tableau 15 : comparatif entre les outils de restitution

D'après le tableau comparatif ci-dessus, nous avons pu résumer que les trois outils fournissent une riche ensemble de fonctionnalités prêtes pour l'emploi de l'entreprise. C'est pour cela que nous avons choisi Power BI comme outil pour notre système.

Pour résumer, l'ensemble des outils que nous avons utilisés pour notre solution sont présentés dans les figures ci-dessous :



Figure 26: les logiciels choisis pour la solution décisionnelle.

B. mise en œuvre de la solution

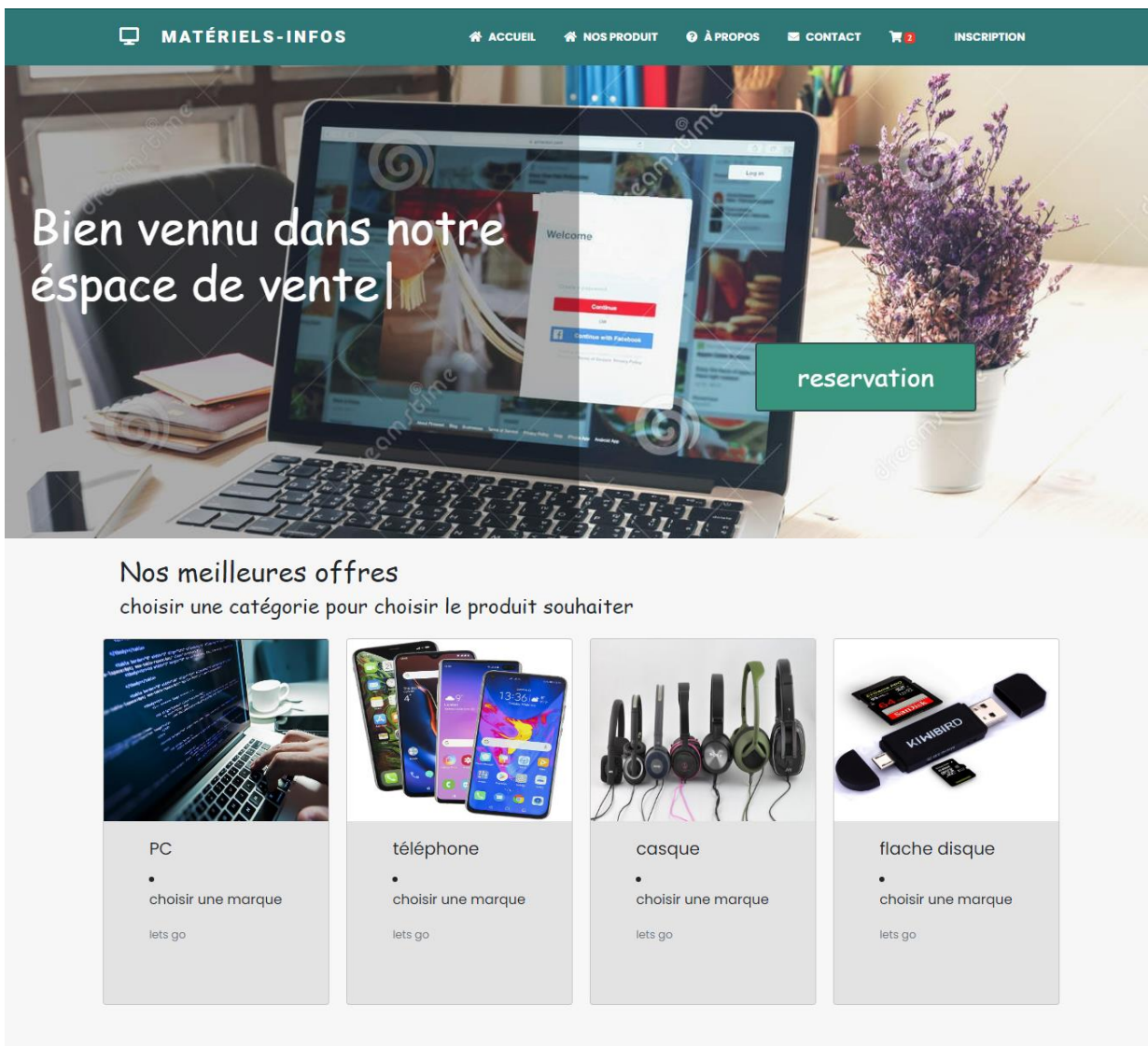
Ceci est le dernier chapitre, nous présentons le site créé pour notre entreprise dans un premier lieu. Puis nous abordons les étapes de l'implémentation de notre DW. Nous commençons par l'alimentation du DW à l'aide d'ETL ; et nous clôturerons par l'aspect de restitution (reporting_).

I. Présentation de l'application de vente :

Dans ce passage nous allons présenter notre application de vente, avec les fonctionnalités qu'elle dispose.

I.1. page d'accueil :

C'est la première page que l'utilisateur ouvre lors de sa connexion à notre site :



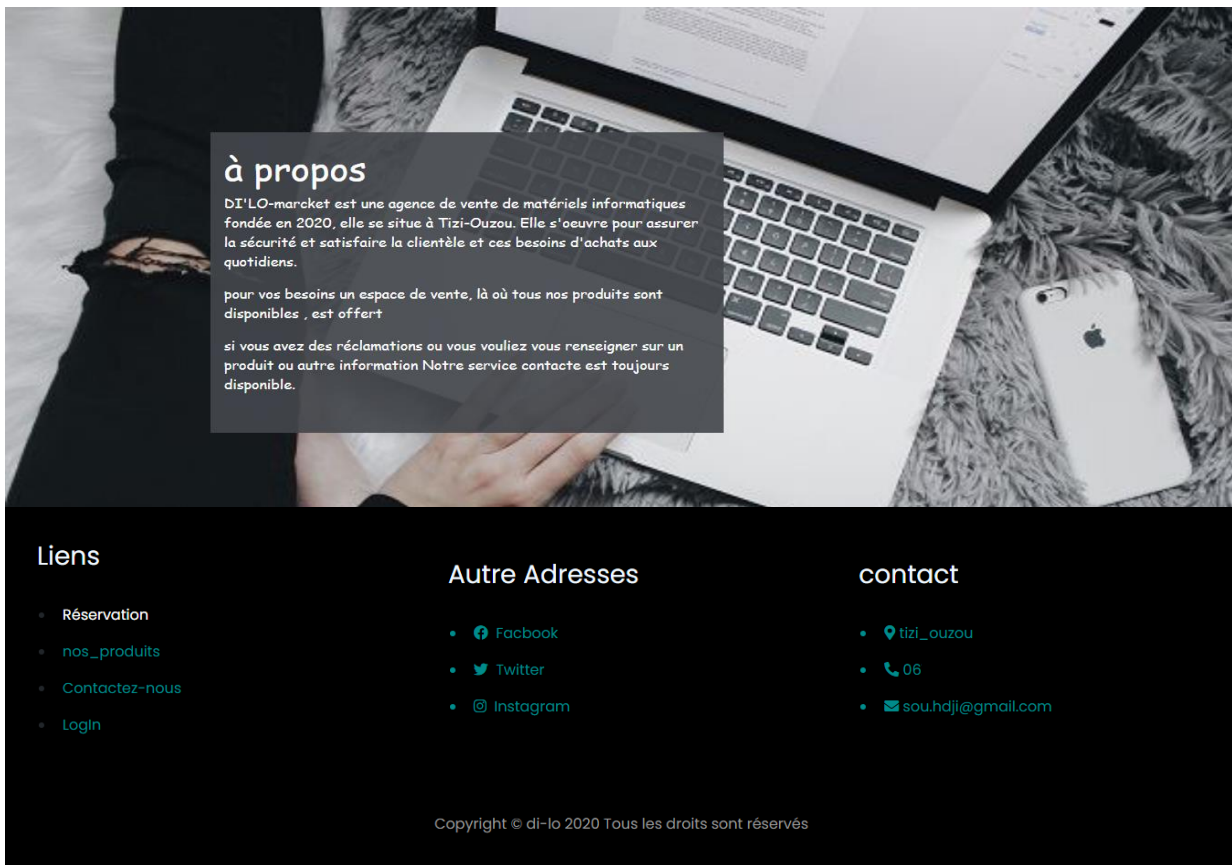


Figure 27 : page d'accueil de note application de vente.

I.2. effectuer une commande :

Si l'utilisateur choisi l'option **nos produit ...** le lien va ver la section **nos meilleurs offres pour choisir sa catégorie** (qui se trouve dans la page accueil présenter dans la figure ci-dessous). Une fois la catégorie et choisie la page si dessus s'affiche. Pour choisi le (ou) les produit(s) voulu, et l'ajouter au panier.

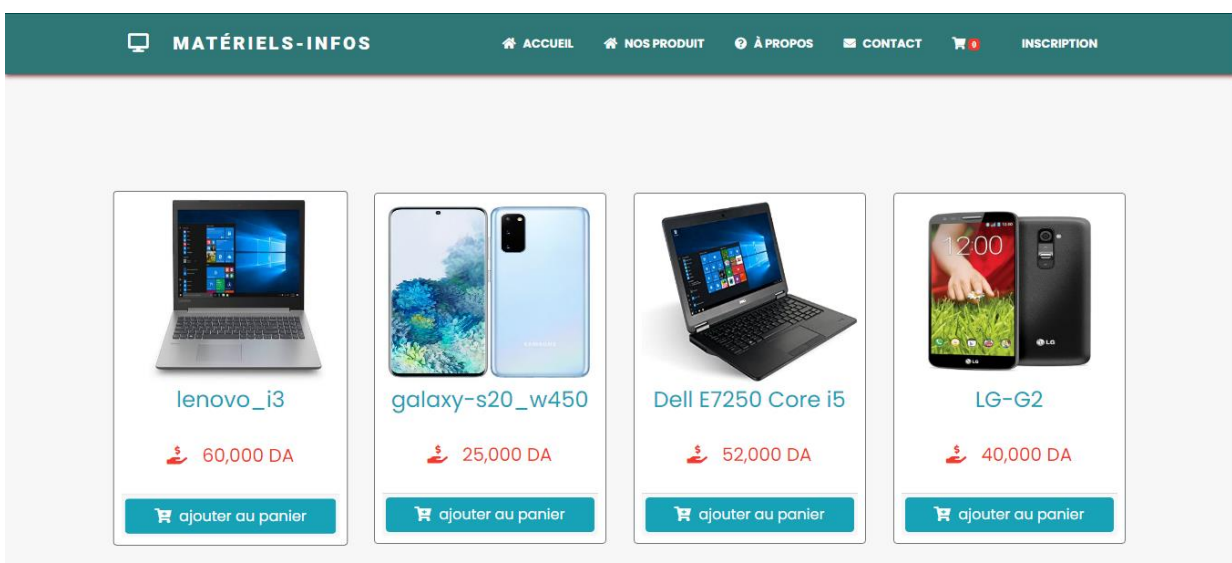


Figure 28 : page pour choisir le produit.

I.3. page pour le panier :

Une fois tous les produits voulus sont choisis. Il va vers sans panier pour le consulter et la page ci-dessus s'affiche. Avec les options de suppression d'un ou plusieurs produits ou ajouter des produits en plus. Et confirmer son ordre si il le souhaite :

ID	Image	Produit	Prix unitaire \$	Prix avec TVA \$	Quantités	montant HT	supp panier
3		Dell E7250 Core i5	52,000 DA	62,192 DA	4	248,768 DA	
4		LG-G2	40,000 DA	47,840 DA	2	95,680 DA	
6		Lenovo-Idpd-C340-15-Core-i3	65,000 DA	77,740 DA	1	77,740 DA	
			Grand Total 422,188 DA				

Figure 29 : panier de l'utilisateur avec les produits choisis.

Si on clique sur confirmer la page ci-dessus s'affiche pour confirmer son panier. Et remplir un certain nombre de champs pour lui valider sa demande.

Je dois ajouter la page après confirmation

completer votre ordre!

produit(s) : 4
frais de livraison: gratuit
Prix_total: 243,984DA

votre nom

l'email

le num de tel

votre adresse

confirmer

Figure 30 : confirmer la demande.

I.4. page d'inscription

Inscrivez-vous maintenant

Nom :

Prenom :

E-mail :

Mot de passe :

Confirmez mot de passe :

Sexe : Homme Femme

Age :



Figure 31 : page d'inscription.

I.4. page de connexion

matériels-infos accueil [Contacter-Nous](#)

Identifiez - Vous

E-mail :

Mot de passe :

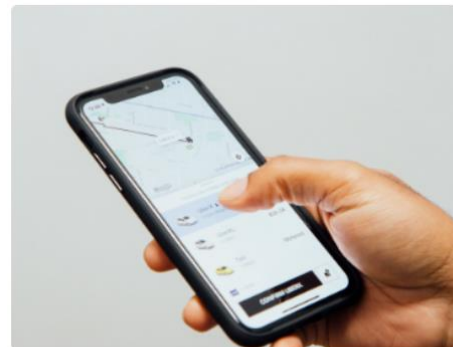


Figure 32 : page de connexion.

I.5. page contact :

Si le client ou l'utilisateur du site a une réclamation, ou désire se confirmer sur une information. Ou bien il veut faire une demande particulière. Le service contact est toujours à ça disposition.

Contactez - Nous

accueil

N'oubliez pas de nous écrire vos impressions sur notre organisme.

Nom :

Email :

Votre message :

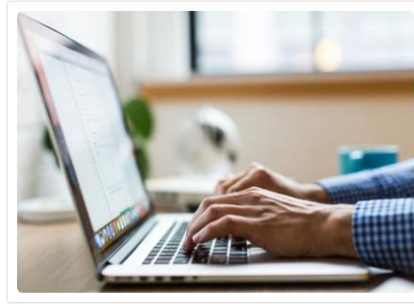


Figure 33 : page de contact.

II. Mise en œuvre de la solution décisionnelle :

1. Business Models :

Avant de lancer n'importe quel projet BI, il est préférable d'utiliser un business models (BM^o). BM permet aux entreprises de représenter graphiquement leur besoins sans avoir à se soucier de leur implémentation ; pour mieux les comprendre et les traduire en information ... BM intègre les ressources et les systèmes déjà en place. La conception du BM est une bonne pratique à adopter dès le début d'un projet. Et dans Talend on trouve la possibilité de schématiser ce BM à l'aide d'une boîte d'outils. Ci-dessous en présente le BM de notre projet BI :

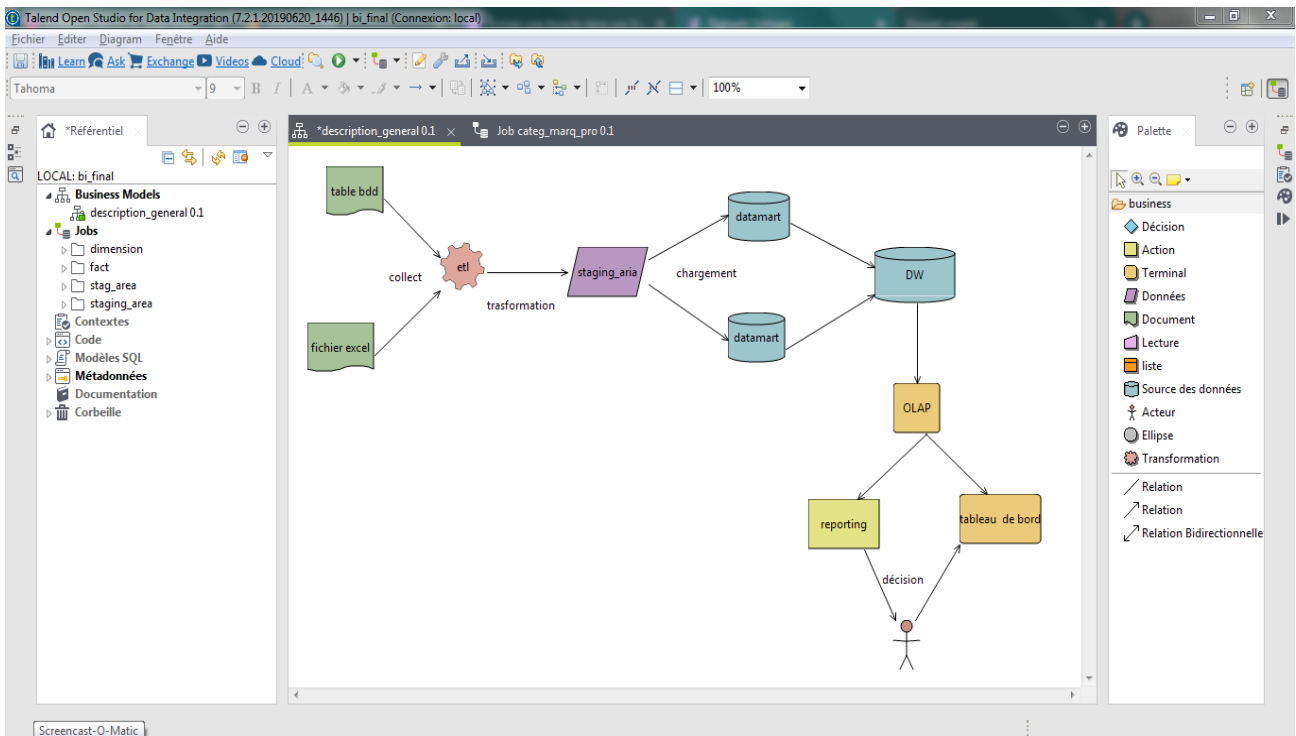


Figure 34 : présentation du business models.

2. Etape de construction de la zone d'entreposage :

Les ETL destinés au DW sont plus complexes car ils impliquent des modifications des schémas de données. Le processus suivant a été défini pour l'alimentation des tables de fait-dimension :

a. Chargement et mises à jour de toutes les tables de dimension reliées à la table de fait :

- ⊕ Extraction des données de la BDD, ou autre exemple (Excel, Access ...).
- ⊕ Conversion et nettoyage des données.
- ⊕ Recherche des non correspondances (supprimer les doublant).
- ⊕ Génération des métas donnés (datamart).
- ⊕ Ajout des nouveaux enregistrements.
- ⊕ Mise en œuvre de mécanisme de gestion des dimensions à variations lentes (Optionnel).
- ⊕ Gestion des erreurs et contrôles.

b. Alimentation de la table de fait :

Le chargement des tables de fait et la dernière étape à effectuer :

- ⊕ Extraction des données de la BDD.
- ⊕ Conversion et nettoyage des données.
- ⊕ Recherche des non correspondances.
- ⊕ Substitution des clefs.
- ⊕ Génération des métas donnés.
- ⊕ (Optionnel) Insertion des champs calculés.
- ⊕ (Optionnel) Agrégation des données.
- ⊕ Ajout des nouveaux enregistrements.
- ⊕ Gestion des erreurs et audits.

La construction de l'entrepôt de données consiste à créer une base de données relationnelle du nom « **data_warehouse0.1** ». Similaire aux schémas en étoile ; ceux que nous avons proposé dans la partie « conception de la zone d'entreposage ».

3. construction et d'alimentation : Pour la mise en place de la solution on suit les étapes suivantes:

3.1. Création des métadonnées :

3.1.1. Outils_infos :

Représente la BDD de notre entreprise, tout les informations sont enregistrer, plus un fichier Excel pour les achats. Après avoir créés la métadonnée. On fait une connexion ver la bdd qui se trouve dans le PHPMyadmin du server wamp. Pour récupérer (extraire) les différentes données de cette dernière. Avant de commence la construction de la zone d'alimentation de l'entrepôt de

données. Nous avons créé une bdd relationnelle sous SQL intermédiaire entre les systèmes sources et le DW nommé « **Staging_Area** ».

a. crée les métadonnées :

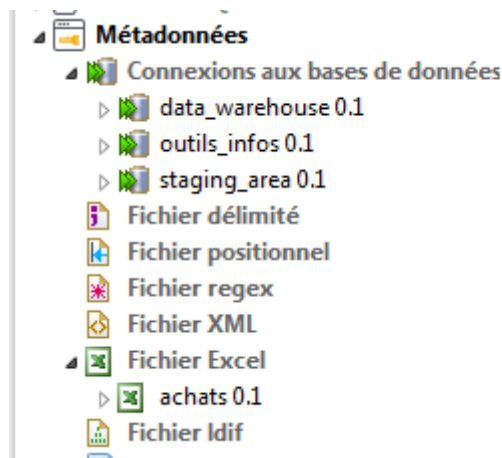


Figure 35 : les métadonnées crée pour la réalisation de la solution.

b. Editer une connexion à la base de données :

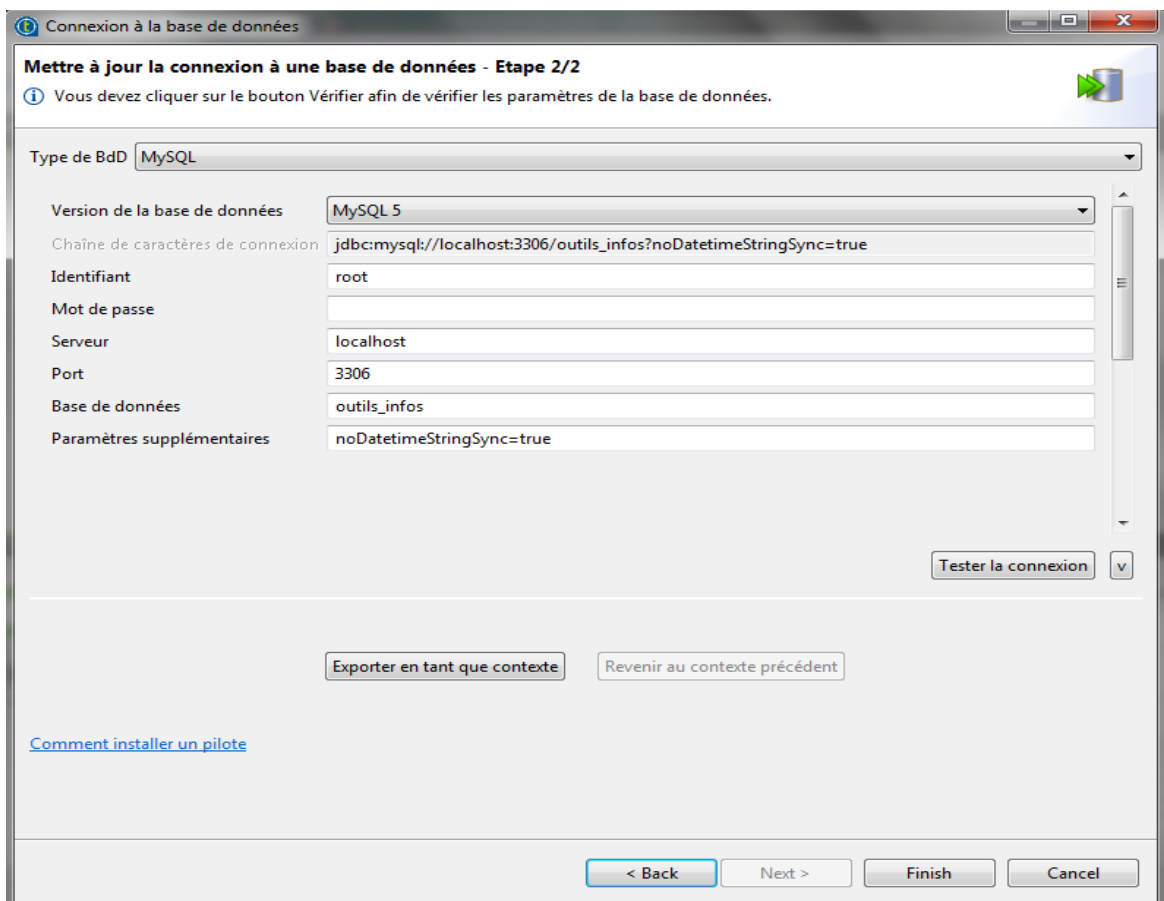


Figure 36: fenêtre de la connexion à la bdd.

c. Récupérer le schéma de la BDD

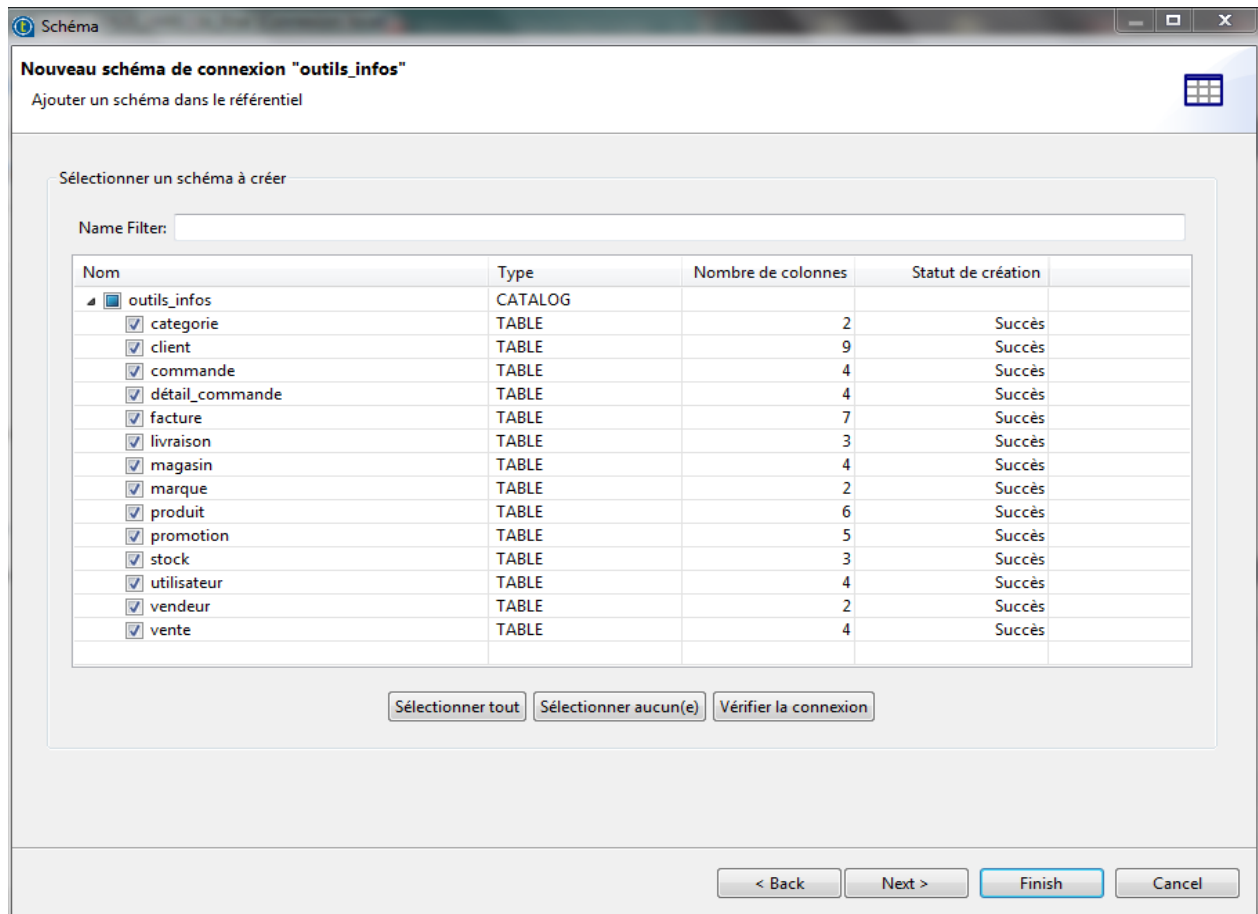


Figure 37 : fenêtre de récupération du schéma.

d. Schéma des tables de la BDD :

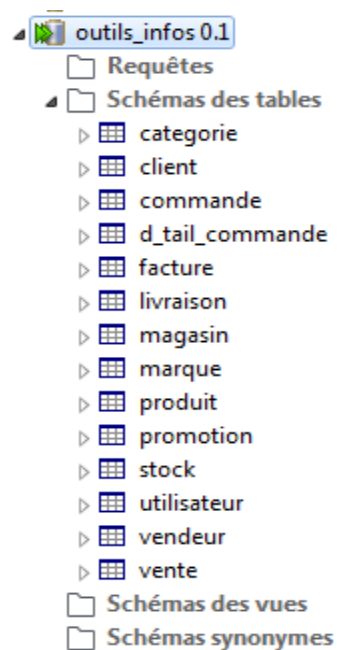


Figure 38 : fenêtre du schéma des tables.

e. Editer une connexion au fichier Excel :

Après avoir créé la métadonnée Excel, on édite une connexion vers le fichier source et récupérer son schéma :

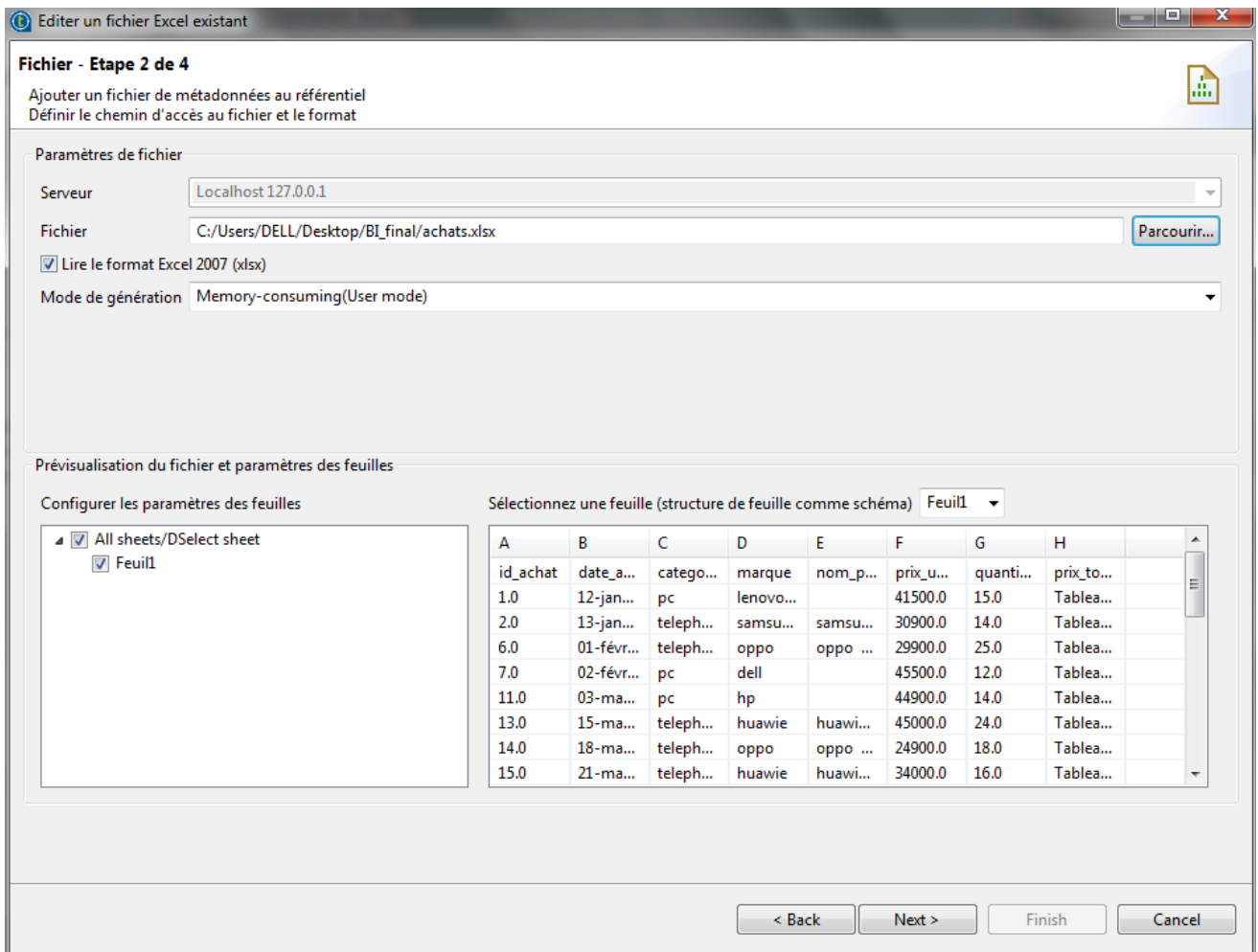


Figure 39 : fenêtre de la connexion au fichier Excel.

3.1.2. La métadonnée staging_area,

Elle représente une copie intégrale de la base de données opérationnelle. Avec quelques tables et modifications ajoutées pour la réalisation de la phase de transformation. Après avoir subi un certain nombre de transformation. Tel que cette transformation s'effectue à l'aide des jobs qu'on a créé sur le logiciel talend. Sur la figure suivante on présente des différentes jobs qu'on crée pour les différentes phases :

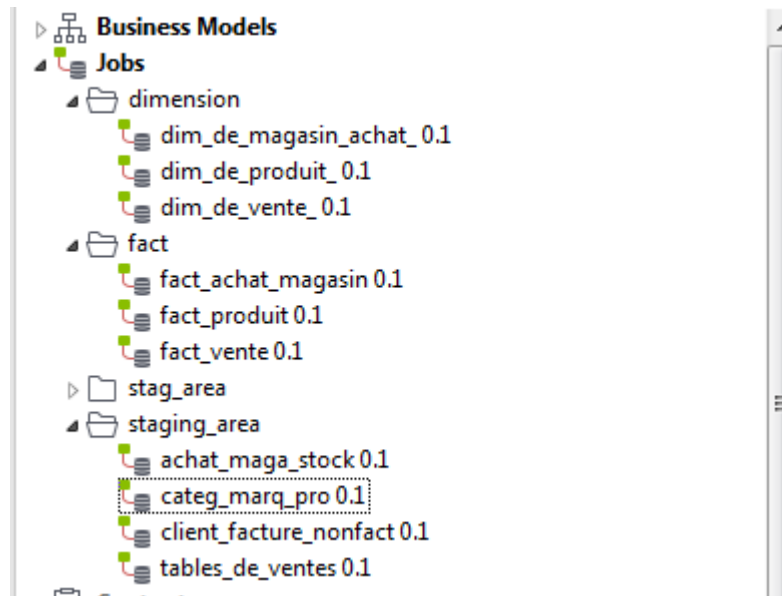


Figure 40 : présentation des jobs crée.

L'alimentation du **Staging_Area** consiste en premier lieu à créer les différentes connexions entre les sources des données (dans notre cas une BDD et un fichier Excel) ; et les métadonnées créées. (les étapes **a,b,c,d,e** du **3.1.1** ci-dessous). Pour définir les contextes qui assurent le basculement facile entre un environnement de développement et un environnement de production. En stockant les données de connexion dans des variables de contexte. Cela nous permet de déterminer sur quel environnement de travail, le projet doit être exécuté, sans avoir à recompiler ou à modifier le projet. La figure suivante présente les contextes que nous avons utilisés dans la construction d'ETL :

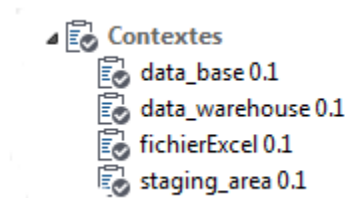




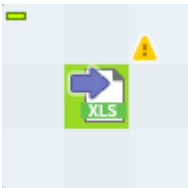

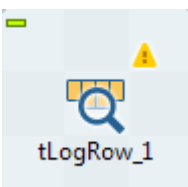
Figure 41 : les contextes.

3.2. création des jobs :

Cette étape consiste à reformater et à transformer les données les plus pertinentes selon le besoin des utilisateurs. Afin d'obtenir des données propres et prêtes à être charger dans l'entrepôt de données. Les figures suivantes présentes les différentes transformations effectuées sur différentes tables «qui se trouvent dans la BDD outils_infos » avant l'alimenter des différentes tables de faits et dimensions.

3.2.1. Les composants de Talend :

Ce tableau représente les différents composants que nous avons utilisés dans le processus ETL :

Composants	Nom	Fonction
	tMysqlInput	<p>Ce composant exécute une requête en base de données selon un ordre strict qui doit correspondre à celui défini dans le schéma.</p> <p>Le tMysqlInput lit une base de données et en extrait des champs à l'aide de requêtes. Les champs récupérés sont ensuite transmis au composant suivant via une connexion de flux (Main row).</p>
	tMysqlOutput	<p>Ce composant écrit, met à jour, modifie ou supprime les données d'une base de données.</p> <p>Le tMysqlOutput exécute l'action définie sur la table et/ou sur les données d'une table, en fonction du flux entrant provenant du composant précédent.</p>
	tFileInputExcel	<p>Ce composant lit un fichier Excel ligne par ligne pour le scinder en champs et envoie les champs comme défini dans le schéma au composant suivant.</p>
	tMap	<p>Ce composant transforme et route des données à partir d'une ou plusieurs source(s) de données vers une ou plusieurs destination(s).</p> <p>Le tMap est un composant avancé qui s'intègre au Studio Talend comme un plug-in.</p>
	tLogrow	<p>Ce composant affiche les données ou les résultats dans la console de la vue Run afin de monitorer les données traitées.</p> <p>Le composant tLogRow Standard appartient à la famille Logs et erreurs.</p>

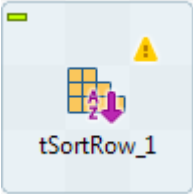
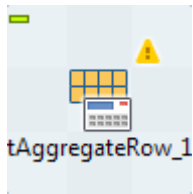
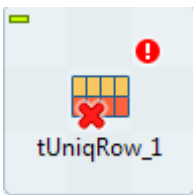


 <p>tSortRow_1</p>	tSortRow	<p>Ce composant établit des métriques et des tables de classification.</p> <p>tSortRow trie les données d'entrée basées sur une ou plusieurs colonnes, selon un type de tri et un ordre.</p>
 <p>tAggregateRow_1</p>	tAggregateRow	<p>Ce composant reçoit un flux de données et fait une agrégation basée sur une ou plusieurs colonnes.</p> <p>Pour chacune des lignes en sortie, une clé d'agrégation est fournie, ainsi que le résultat de l'opération d'agrégation correspondant (min, max, sum, etc.).</p> <p>Le tAggregateRow permet de fournir un jeu de métriques basés sur des valeurs ou des calculs.</p>
 <p>tUniqRow_1</p>	tUniqRow	<p>Ce composant assure une qualité de données des flux d'entrée et de sortie du Job.</p> <p>Le composant tUniqRow compare des entrées et trie les entrées en doublon du flux d'entrée.</p>
 <p>tFilterRow_1</p>	tFilterRow	<p>Ce composant filtre des lignes d'entrée en définissant une ou plusieurs condition(s) sur les colonnes sélectionnées.</p>
 <p>tReplicate_1</p>	tReplicate	<p>ce composant duplique le schéma entrant en deux flux identiques.</p> <p>Il effectue différentes opérations sur le même schéma.</p>

Tableau 16 : les composants talend utilisés

3.2.2. Création des jobs pour alimenter la staging area :

3.2.2.1 Transformation à l'aide des jobs :

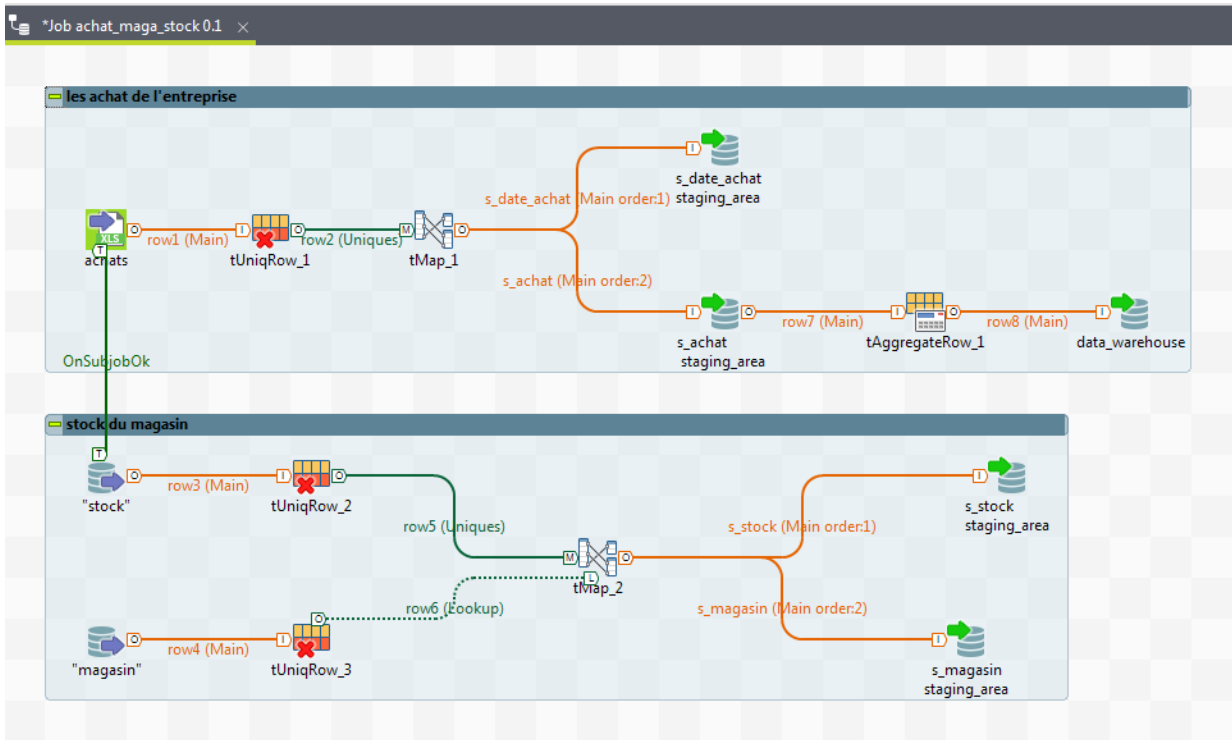


Figure 42 : job de transformation d'achat et stock.

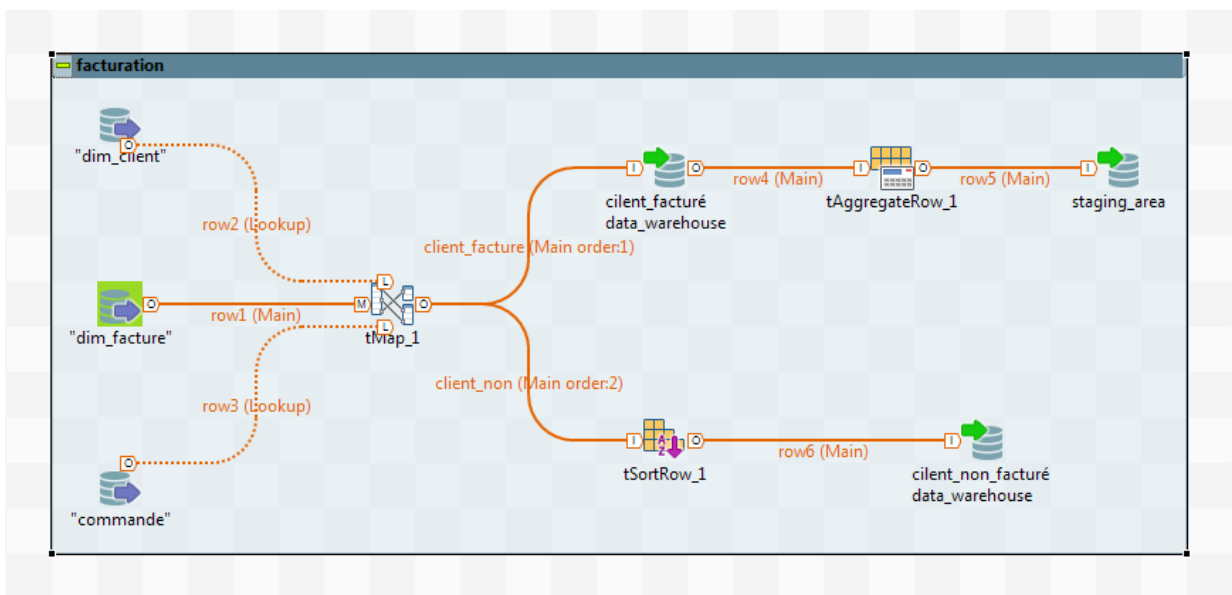


Figure 43 : job pour faire une répartition des clients selon la facturation.

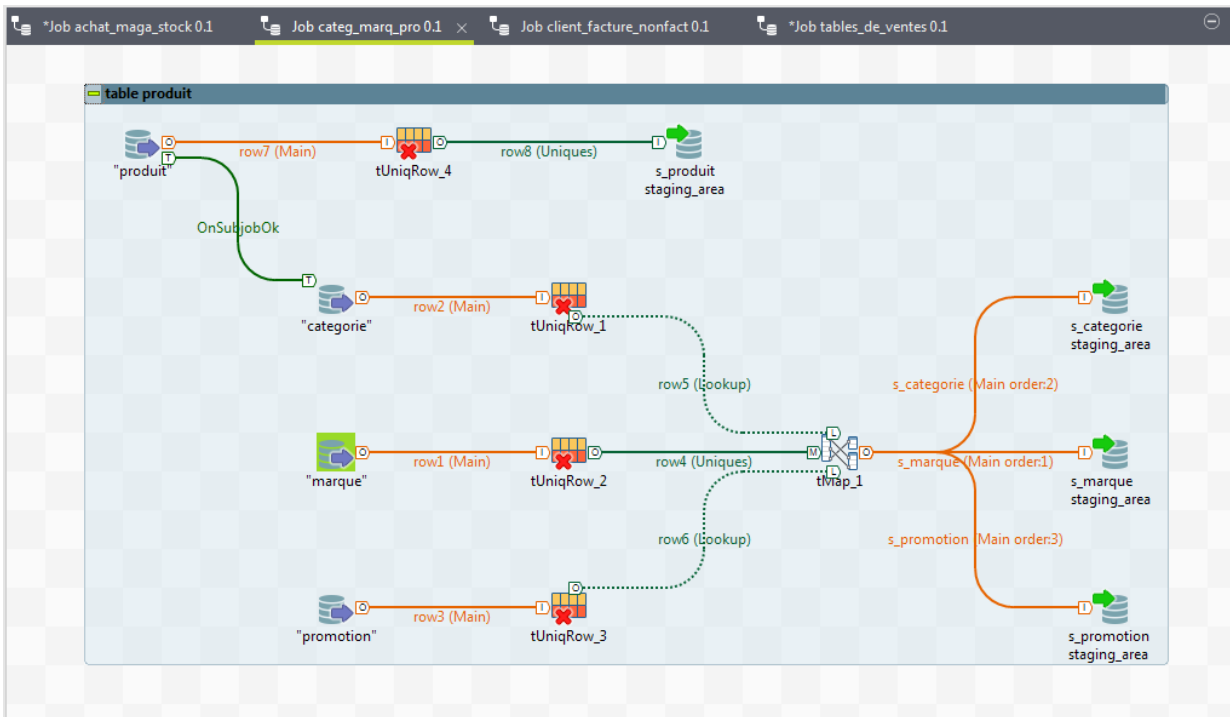


Figure 44 : job d'extraction du fait_produit et les tables relié à elle.

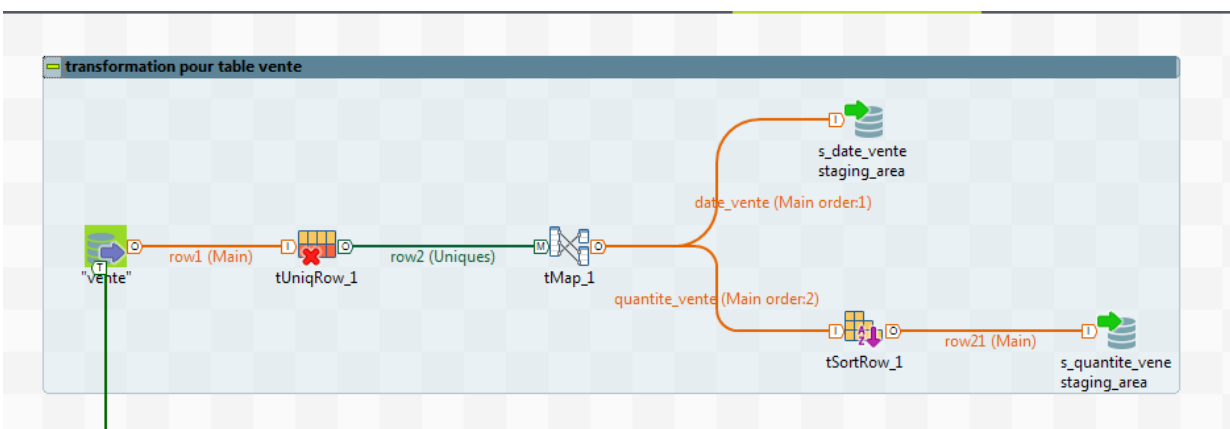


Figure 45 : job d'extraction et transformation table vente.

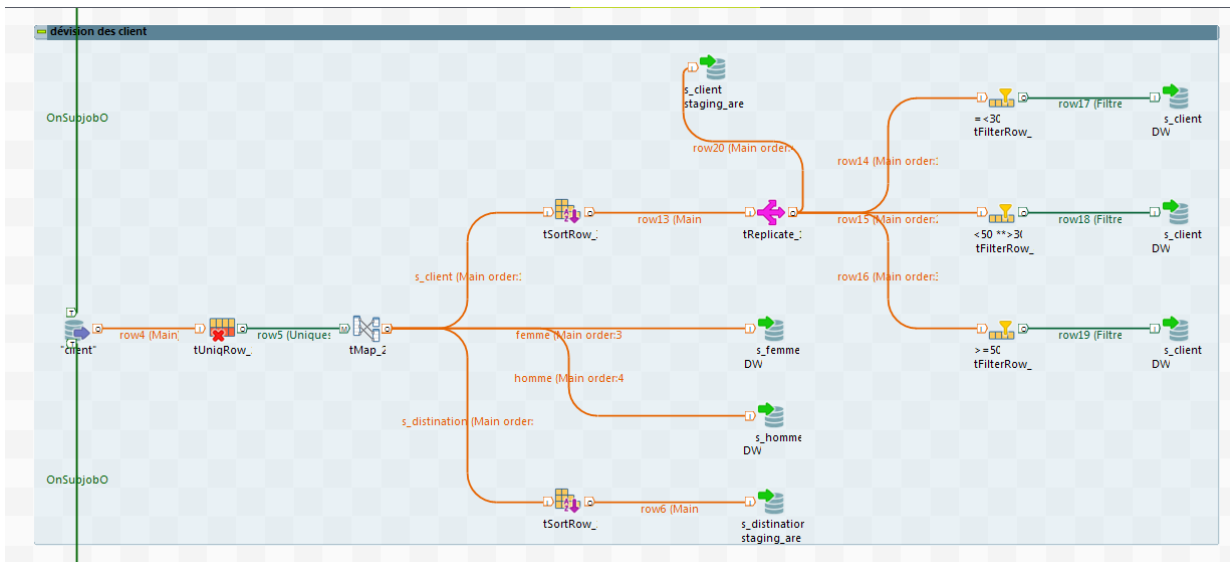


Figure 46 : job de répartition des clients.

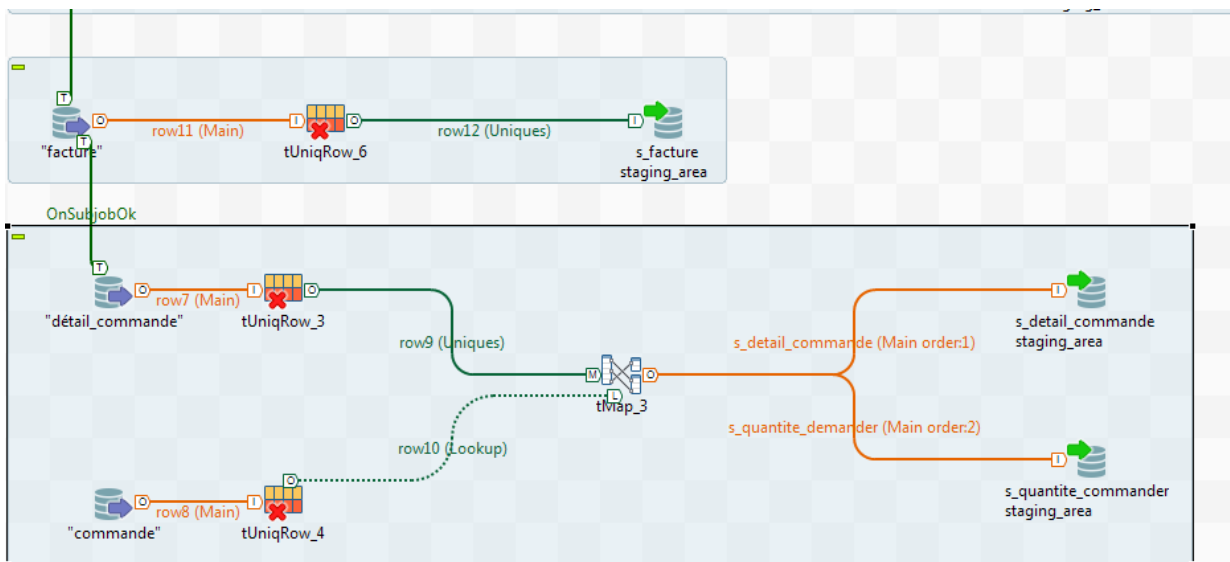


Figure 47: Exemple de transformation des dimensions pour la table fait "vente"

3.2.2.2. Lors de la transformation on stocke les données dans la BDD staging_area qui se trouve dans MySQL du serveur wamp. Puis on récupère les schémas des tables vers la métadonnée staging_area qui se trouve dans l'éditeur Talend :

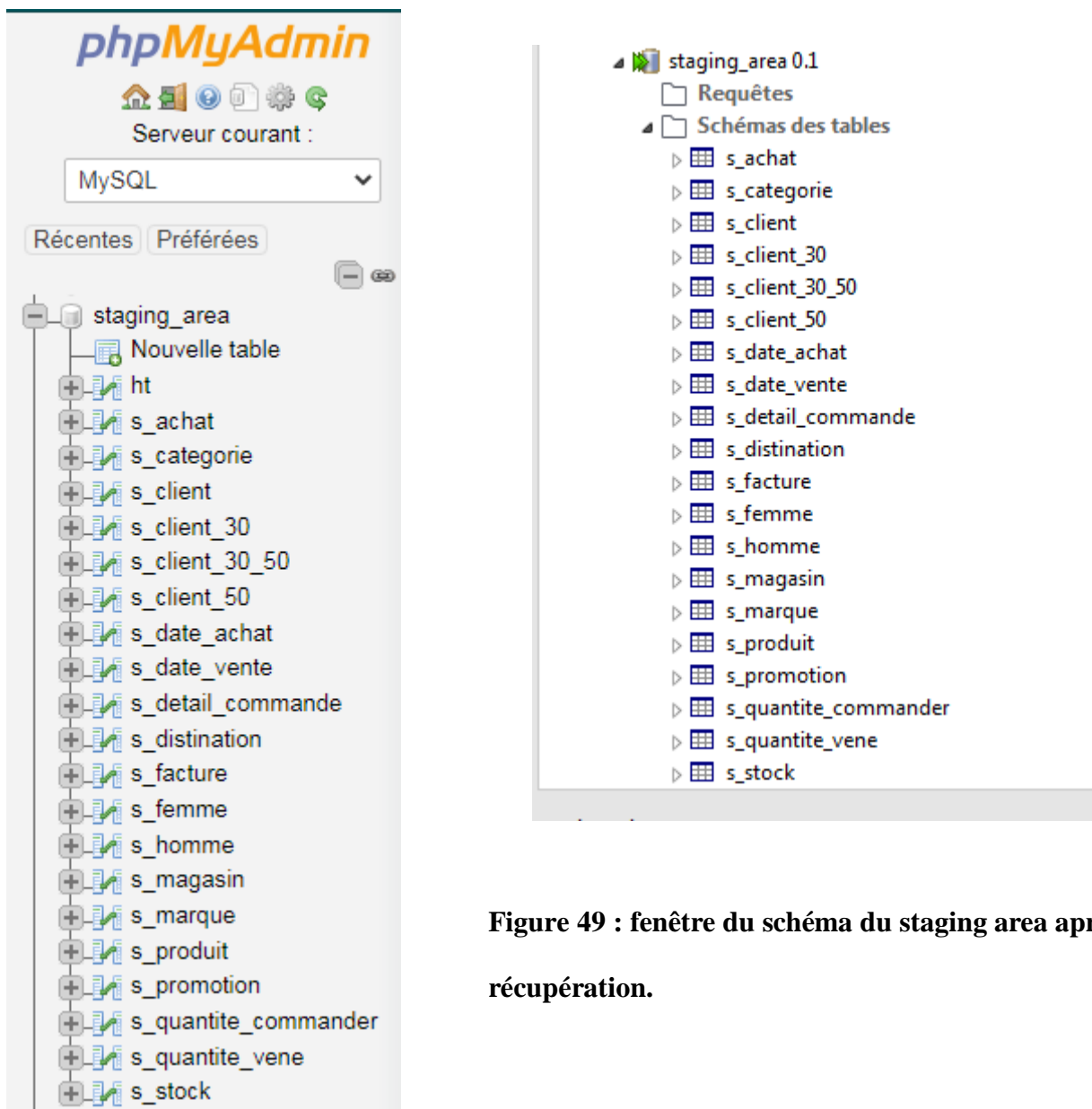


Figure 49 : fenêtre du schéma du staging area après récupération.

Figure 48: présentation du staging area chargement.

3.2.3. Création des jobs pour alimenter le DW (et les data-marts) :

Après avoir stocké les données transformer sur « staging_area » la dernière étape qui est l'alimentation du DW s'effectue ; Et cela en alimentation et chargement ces différentes tables dimensions et faits.

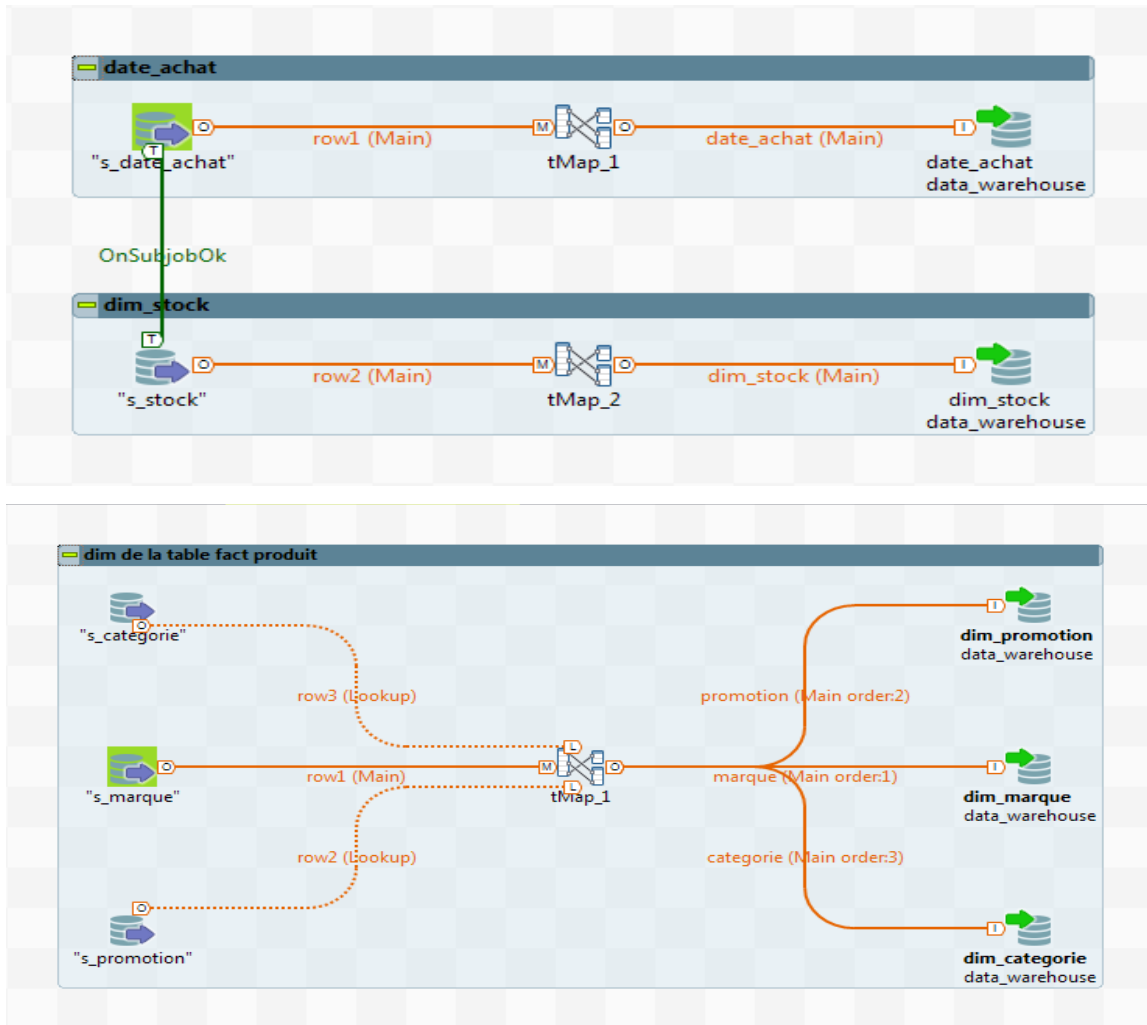


Figure 50 : exemple d'alimentation des tables de dimension.

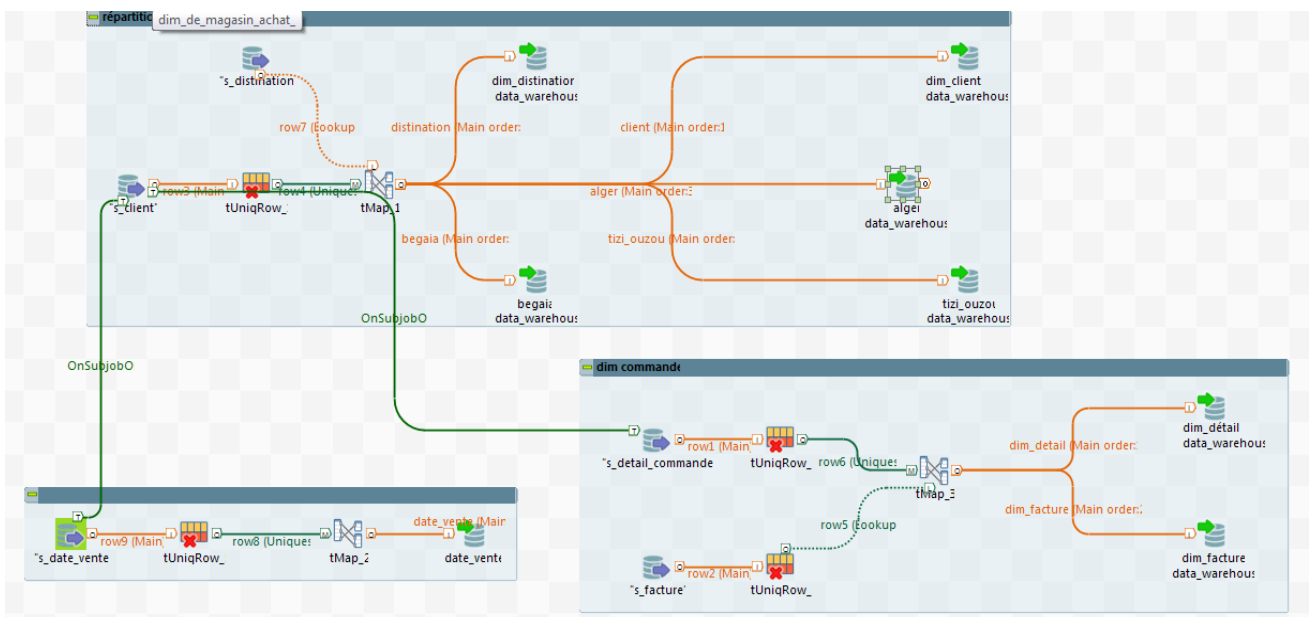


Figure 51 : exemple d'alimentation des tables relié avec fait_vente.

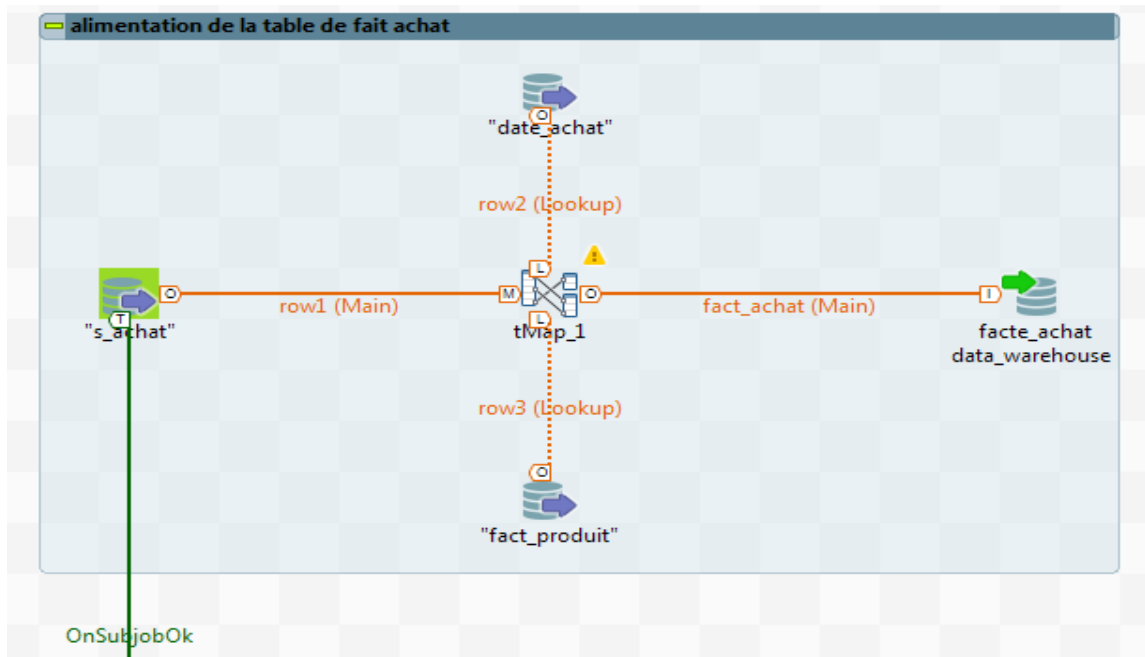


Figure 52 : exemple d'alimentation fait_achat.

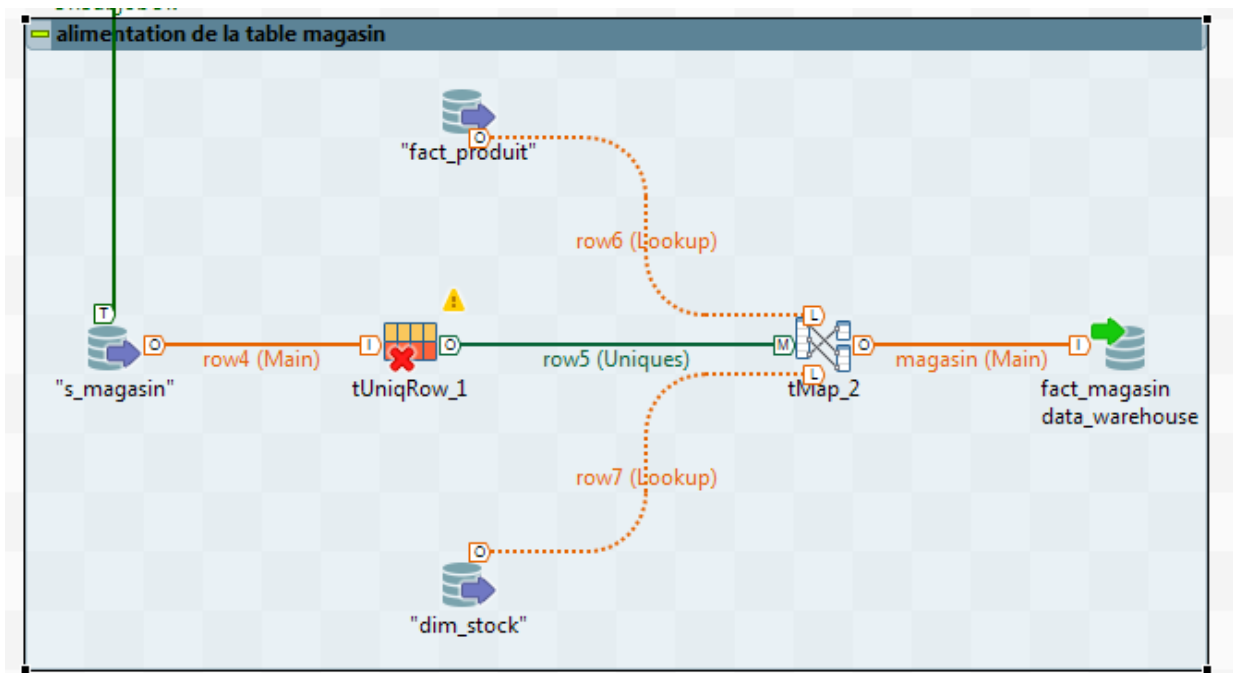


Figure 53 : chargement du fait « suivi magasin ».

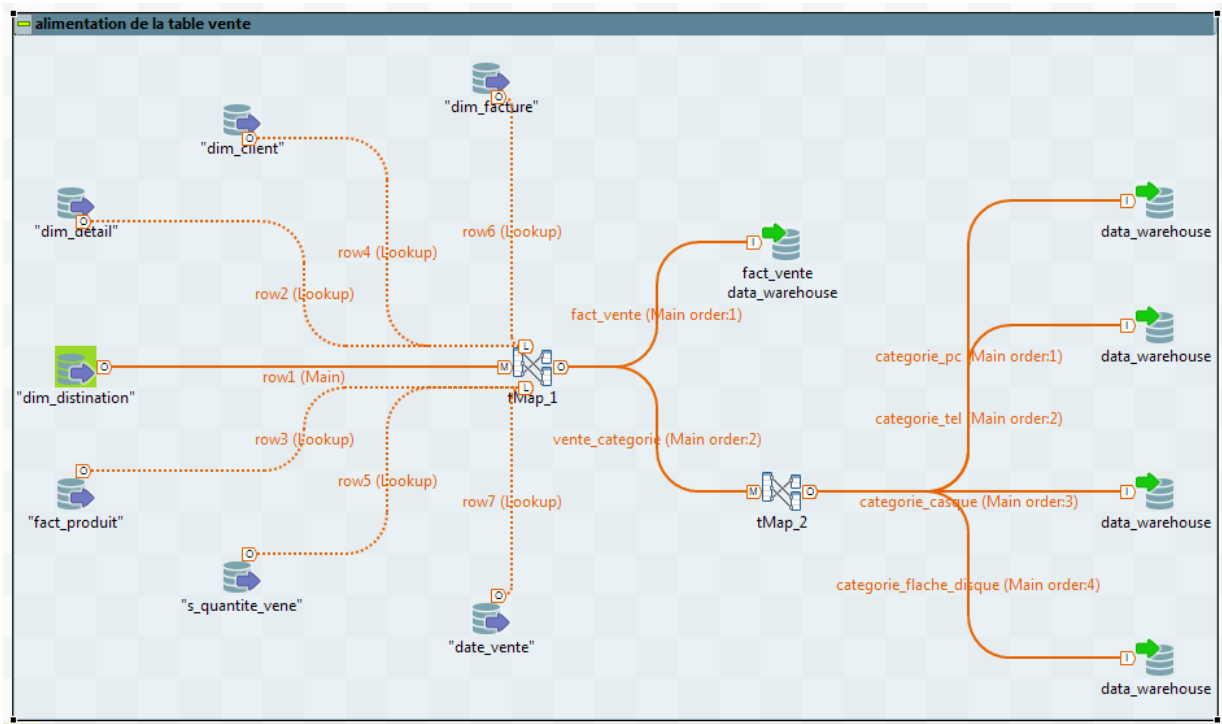


Figure 54 : chargement du fait « suivi vente ».

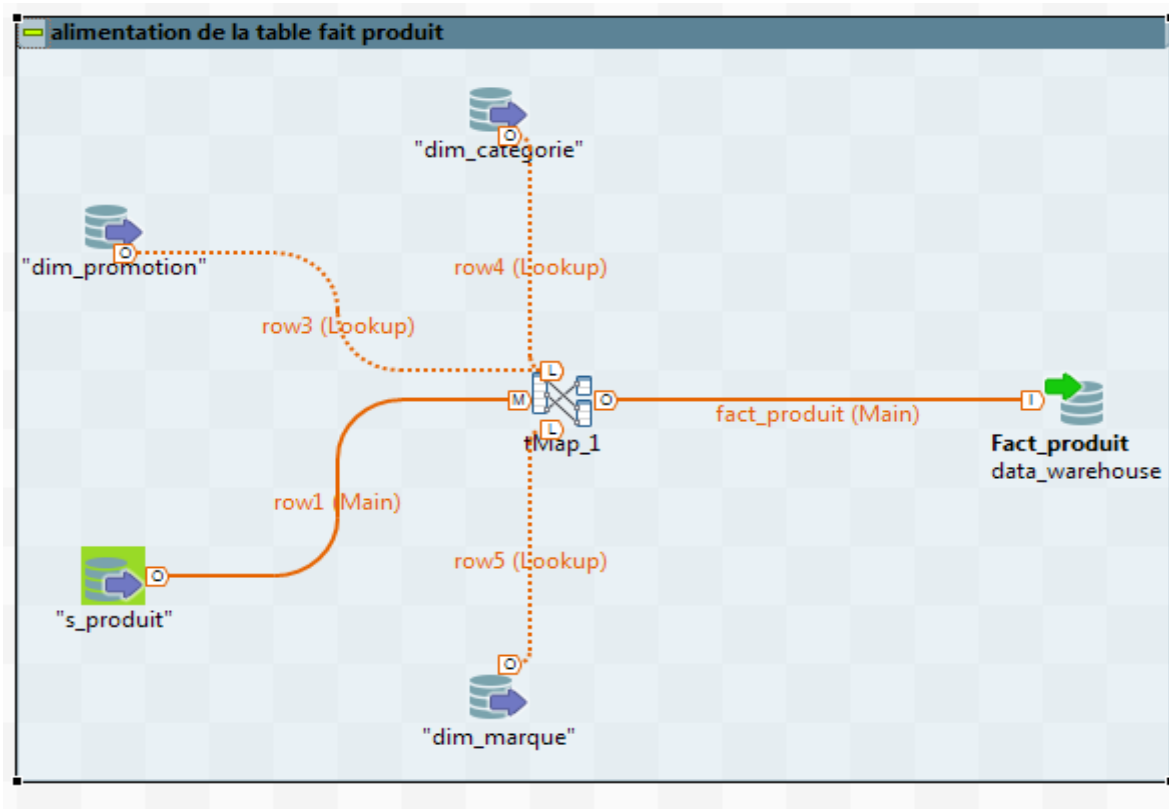


Figure 55 : chargement du fait « suivi produit ».

4. Construction des cubes multidimensionnelle et analyse OLAP :

Une fois les données sont traitées transformées et chargées dans l'entrepôt de données, nous avons procédé à la construction des cubes multidimensionnels. En utilisant **Mondrian Workbench** qui est un outil ROLAP. Il permet de mapper les structures OLAP, telles que les cubes, les dimensions, les faits, les attributs et les hiérarchies qui assurent la facilité de navigation dans le cube directement sur les tables et les colonnes de la base de données relationnelle via des fichiers XML. La figure suivante présente un exemple des cubes:

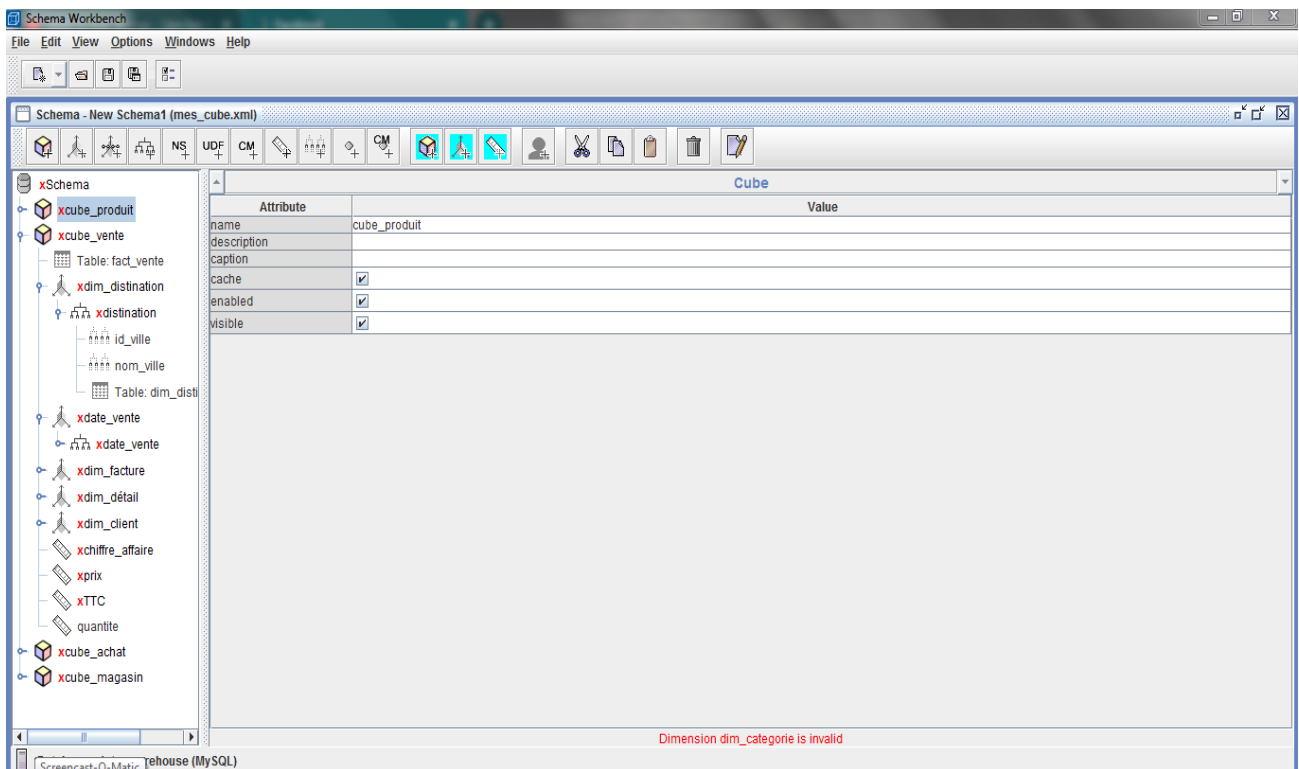


Figure 56 : construction des cube OLAP .

III. Construction de la zone de restitution :

Nous présentons dans les figures suivantes quelques captures des tableaux de bord réalisés :

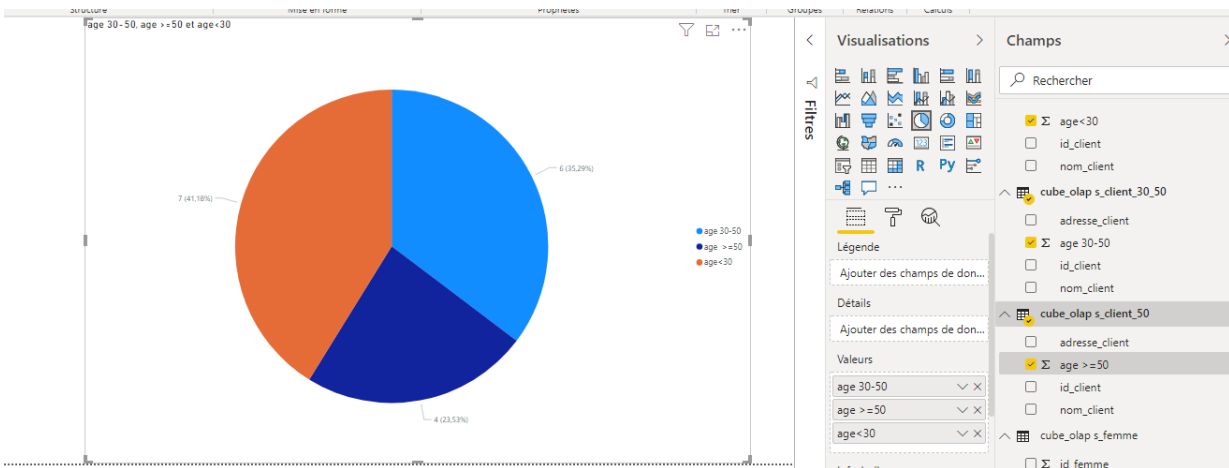
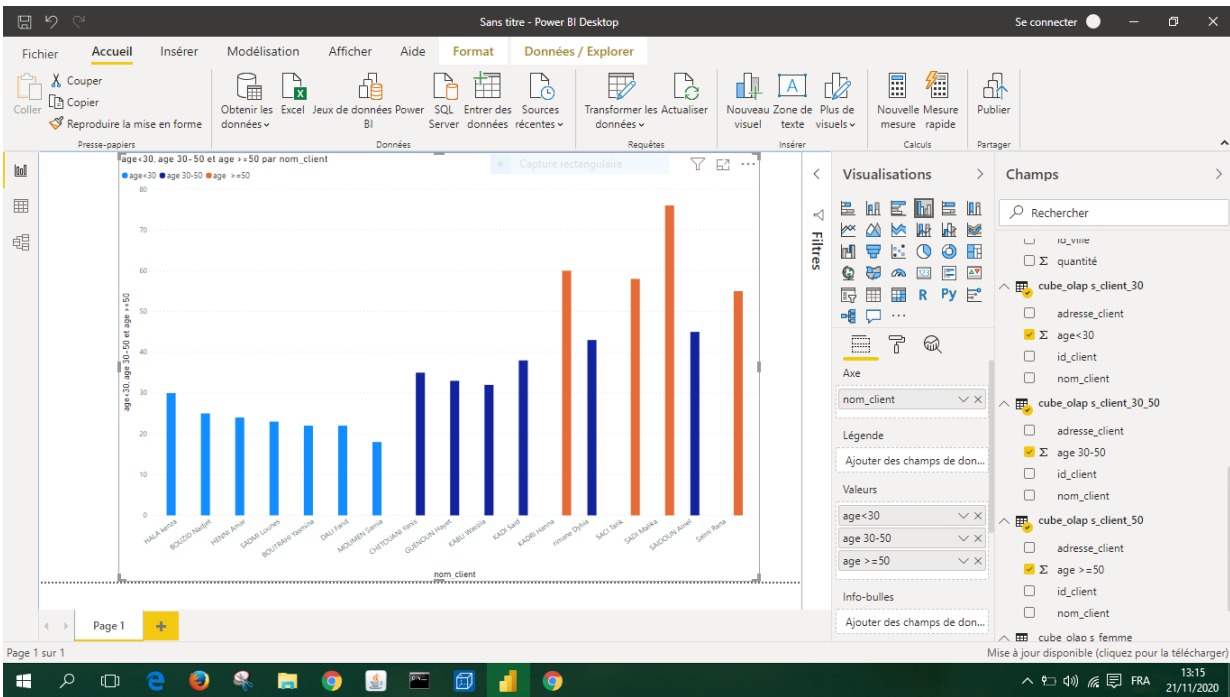


Figure 57 : répartition des client selon l'age.

Moyenne de id_ville_tiziouizou, Moyenne de id_ville_bejaia et Moyenne de id_ville_alger

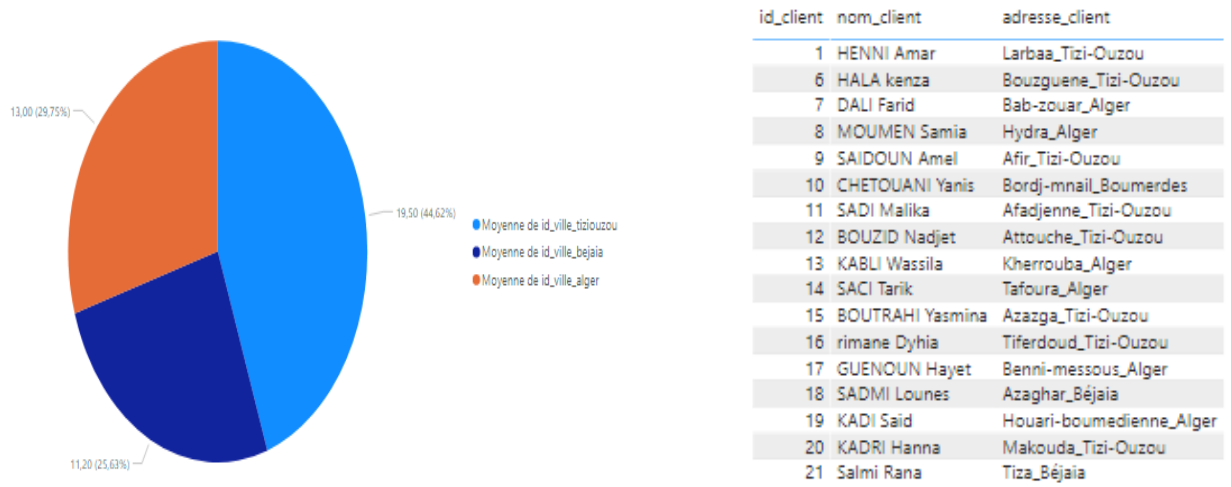


Figure 58 : tableau de bord répartition des clients

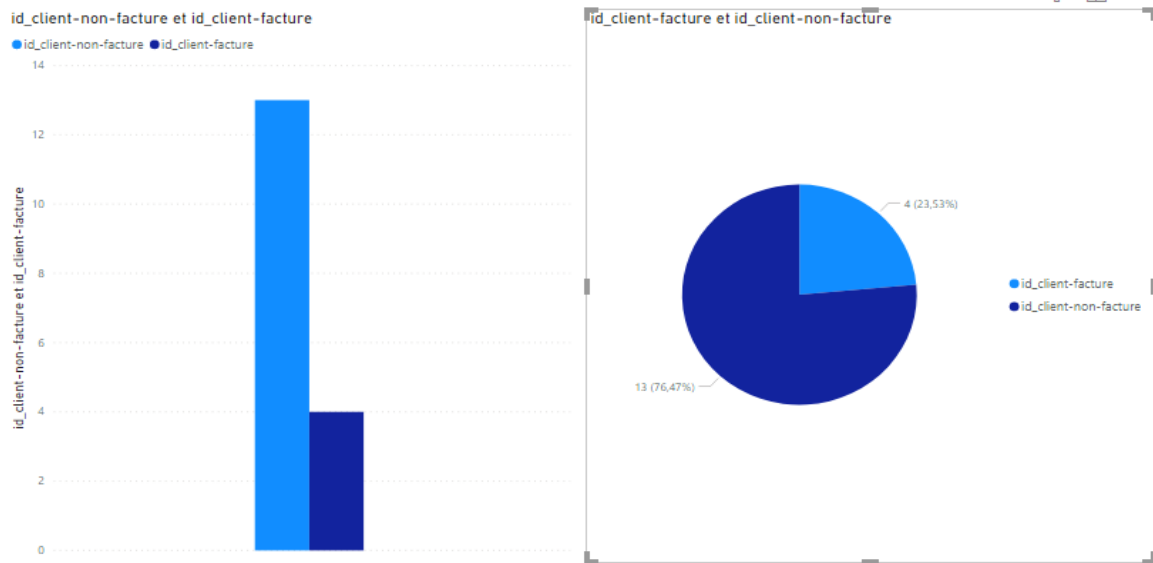


Figure 59 : suivi de facturation

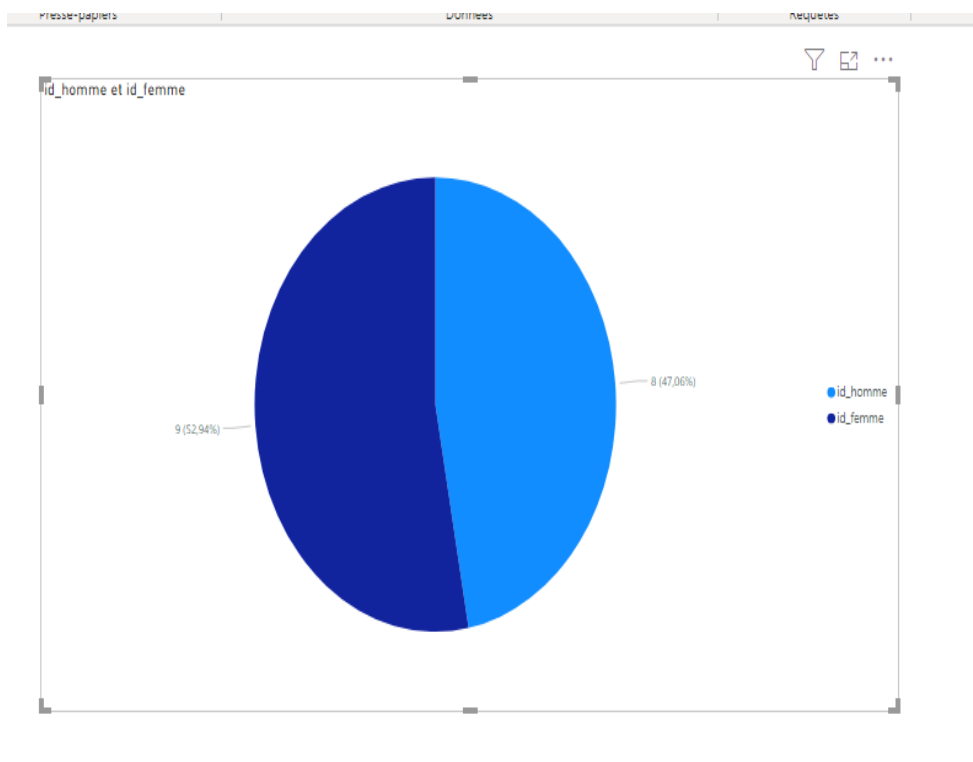


Figure 60 : tableau de bord répartition des clients selon le sexe.

Conclusion

Il a été en question dans ce chapitre d'aborder les différents aspects liés à l'implémentation de notre solution. En premier lieu nous avons présenté les étapes de construction des différentes zones de la solution à savoir la zone d'entreposage, la zone ETL, la réalisation des cubes multidimensionnels, et la zone de restitution.

Le travail présenté dans ce mémoire de fin d'étude s'est porté sur la conception et la réalisation d'un système d'information d'aide à la décision dédié à la gestion de commerciale. Ce projet rentre dans le cadre de la stratégie de renforcement de prise de décision d'une entreprise de vente. Le système mis en place permettra aux responsables et décideurs de l'entreprise, d'analyser soigneusement les données commerciales, pour avoir l'information voulue au moment voulu.

Dans le but de concrétiser notre solution, nous avons suivi une méthodologie, nous avons également dressé un programme qui évoque les principales étapes à suivre. Dans la première étape nous avons étudié les différents aspects théoriques liés aux systèmes décisionnels généralement et au data Warehouse et à la modélisation multidimensionnelle en particulier. Ensuite nous avons présenté l'organisme d'accueil, sa vision stratégique, ainsi qu'une description de ses processus métiers.

Par la suite, nous avons entamé la phase d'étude de besoins, qui est nécessaire à tout projet décisionnel. Dans laquelle nous avons récolté les besoins de l'entreprise ; en utilisant les différentes techniques de collecte de besoins. Finalement, nous avons analysé ces besoins à travers un diagnostic qui nous a aidé à identifier les faiblesses du processus décisionnel actuel et de trouver une solution qui satisfait les utilisateurs finaux.

A l'issue de cette partie, nous avons proposé notre solution, qui consistait à la mise en place d'un système décisionnel, basé sur un entrepôt de données, accompagné d'un ensemble d'outils d'analyse et de restitution adaptés aux besoins exprimés. En commençant par la zone d'entreposage dans laquelle nous avons défini les différents volets d'analyse qui couvrent tous les besoins des utilisateurs. Ensuite nous avons proposé une modélisation capable de répondre à ces besoins. Puis nous avons abordé la phase d'alimentation de l'entrepôt, où nous avons défini la politique de réalisation de ce processus. Et nous avons fini par la conception de la zone de restitution. Dans laquelle nous avons conçu les tableaux de bord, les rapports et les différents cubes OLAP qui permettent d'accélérer les analyses.

Bibliographie :

- [1] <https://biworks.fr/levolution-de-la-bi-et-les-tendances-pour-2018/>
- [2] <https://www.lebigdata.fr/business-intelligence-definition#:~:text=La%20Business%20Intelligence%2C%20ou%20informatique,dans%20leurs%20prises%20de%20d%C3%A9cisions.>
- [3] <https://www.lebigdata.fr/>
- [4] Inmon, B. (1996). Building the Data Warehouse.
- [5] Inmon, W. H. (2002). Building the Data Warehouse.
- [6] J.-F, G. (2001). La construction du datawarehouse : du datamart au dataweb. 2ème édition.
- [7] JOURNAL OFFICIEL DE LA REPUBLIQUE ALGERIENNE N° 08. (2008). DÈcret exÈcutif n° 08-54 du 2 Safar 1429 correspondant au 9 fÈvrier 2008 portant approbation du cahier des charges-type pour. JOURNAL OFFICIEL DE LA REPUBLIQUE ALGERIENNE N° 08.
- [8] Kimball. (2013). The data warehouse toolkit: The definitive guide to dimensional modeling.
- [9] Kimball, & Ross. (2008). The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling.
- [10] 5339c4bb665af.pdf
- [11] https://www.memoireonline.com/08/13/7308/m_Analyse-et-detection-de-lattrition-dans-une-entreprise-de-telecommunication5.html
- [12] Leonhardi. (2010). Augmenting olap exploration with dynamic advanced analytics. In Proceedings of the 13th International Conference.
- [13] (Bernard ESPINASSE . (2014, Septembre). Introduction aux Systèmes Interactifs d'Aide à la Décision. Marseille.
- [14] (METAIS, E. (2019, juin 8). Systèmes d'aide à la décision. Récupéré sur Encyclopædia Universalis: <http://www.universalis.fr/encyclopedie/systemes-informatiques-systemes-d->

aide-a-la-decision/

- [15] MEMOIRE_636953459723634666.pdf
- [16] sesnv160081.pdf
- [14] <https://www.digdash.com/fr/top-10-definitions-business-intelligence/>
- [17] <https://www.atinternet.com/glossaire/data-driven/#:~:text=Le%20Data%20Driven%2C%20%C3%A9galement%20appel%C3%A9,ses%20consommateurs%20et%20ses%20clients.>
- [18] <http://dihiainfos.blogspot.com/2016/10/le-systeme-operationnel-vs-le-systeme.html>
- [19] <https://grim.developpez.com/cours/businessintelligence/concepts/conception-datawarehouse/>
- [20] <https://www.lebigdata.fr/power-bi-microsoft>
- [21] Beaudoin. (1990). la gestion du changement, une approche stratégique pour l'entreprise en mutation.
- [22] Bellatrech, Giacomett, & Marce. (2005). A personalization framework for olap queries.
- [23] Chaudhuri, Dayal, & Narasayya. (2011). An overview of business intelligence technology.
- [24] Ghozzi, F. (2004). Contraintes pour modèle et langage multidimensionnels. Hermes – Lavoisier.
- [25] <https://help.talend.com/r/rqVGdHtCwjLQfXwWricOzQ/w5CH06hrLUwj0eR6GAZaEw>