UNIVERSITÉ DU BURUNDI
GRENIER DU SAVOIR DU BURUNDI
Dépôt institutionnel officiel
https://repository.ub.edu.bi

Conception et réalisation d'un entrepôt de données pour la mise en place d'un système de gestion de vente des produits théicoles : cas de l'Office du Thé du Burundi (OTB)

Havyarimana, Didier; Sous la direction de: Dr Ndayisaba Longin

2016
UB, FSI
https://repository.ub.edu.bi/handle/123456789/782

Document téléchargé depuis le **Grenier du Savoir du Burundi** — https://repository.ub.edu.bi

UNIVERSITE DU BURUNDI



FACULTE DE SCIENCES DE L'INGENIEUR

DEPARTEMENT DES TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

Mémoire de Mastère en Génie Informatique

CONCEPTION ET REALISATION D'UN ENTREPOT DE DONNEES POUR LA MISE EN PLACE D'UN SYTEME DE GESTION DE VENTE DES PRODUITS THEICOLES : cas de l'Office du Thé du Burundi(OTB)

Présenté et soutenu par :

HAVYARIMANA Didier

En vue d'obtention d'un diplôme de Mastère en Génie Informatique

Sous la direction de:

Dr NDAYISABA Longin Enseignant à l'Université du Burundi

Membres du Jury :

Dr MUKESHIMANA Michelle	: Président
Dr SAHINGUVU William	: Vice-président
Dr NDIKUMAGENGE Jérémie	: Secrétaire
Dr NDAYISABA Longin	: Directeur

Année Académique 2015 -2016

Dédicace

A Dieu tout puissant ;

A mes parents ;

A mes frères et sœurs ;

A mes cousins et cousines ;

A mes oncles et tantes ;

A tous ceux qui me sont chers ;

A tous mes amis et connaissances ;

HAVYARIMANA Didier

Remerciements

La réalisation de ce mémoire a été possible grâce aux efforts consentis par différentes personnes tant physiques que morales à qui je voudrais témoigner ma reconnaissance.

Je voudrais tout d'abord remercier notre Dieu tout puissant qui ne cesse de me protéger et de me guider dans ma vie quotidienne. Qu'il soit loué infiniment.

Je désire aussi adresser mes très sincères remerciements au directeur de ce mémoire, Mr Dr. NDAYISABA Longin, pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué à alimenter ma réflexion. Il a accepté de me guider dans l'accomplissement de mon travail malgré ses multiples responsabilités, sa franche collaboration et ses remarques m'ont permis de mener à terme de ce travail.

Je tiens à remercier mes très chers parents, qui ont toujours été là pour moi, et qui ont tout sacrifié n'épargnant ni santé ni effort. « Vous m'avez donné un magnifique modèle de labeur et de persévérance. Je suis redevable d'une éducation dont je suis fier ».

J'aimerais exprimer ma gratitude au corps professoral de l'UB en général, plus particulièrement celui de la FSI département des TIC dans l'option de GI pour la formation de qualité dont j'ai bénéficiée.

Mes remerciements sont adressés aux responsables de l'OTB en général pour leur appui lors de la documentation, plus particulièrement dans le service commercial, service de production et service d'informatique qui m'ont réservé un accueil chaleureux et qui m'ont prêté une oreille attentive.

Je remercie vivement les membres du jury d'avoir accepté d'évaluer ce travail.

Pour finir, toute personne qui, de près ou de loin, par sa parole, son écrit, ses conseils et ses critiques a contribué à la réalisation de ce travail nous disons « Merci à tous ».

Liste des abréviations

OTB	:	Office du thé du Burundi
DW	:	Data Warehousse
ETC	:	Extraction-Transformation-Chargement
ETL	:	Extract-Transform-Load
FSI	:	Faculté des Sciences de l'Ingenieur
GI	:	Génie Informatique
HOLAP	:	Hybrid On-Line Analytical Processing
MOLAP	:	Multidimensional On-Line Analytical Processing
OLTP	:	On-Line Transaction Processing
OOLAP	:	Object On-Line Analytical Processing
ROLAP	:	Relational On-Line Analytical Processing
SID	:	Système d'Information Décisionnel
TIC	:	Technologies de l'Information et de la Communication
UB	:	Université du Burundi
DES	:	Data Encryptions Standard
SQL		Structured Query Language

:

Liste des figures

Figure 1: Environnement Décisionnel au sein du système d'information.....16
Figure 2: Différentes composantes du système décisionnel.....17
Figure 3: Architecture des systèmes décisionnels.....19
Figure 4: Structure des données d'un entrepôt de données.....23
Figure 5**:** Achitecture de l'entrepot de données.....24
Figure 6: Exemple d'une table de fait.....26
Figure 7: Exemple d'une table de dimension.....27
Figure 8: Exemple du modèle en étoile.....28
Figure 9: Exemple de modèle en flocon.....29
Figure 10: Exemple d'un modèle en constellation.....30
Figure 11: Dimension Temps.....40
Figure 12: Dimension Type_Vente.....41
Figure 13: Dimension Type_Monnaie.....41
Figure 14: Dimension Quantite_The.....42
Figure 15: Dimension Usine.....43
Figure 16: Fait « Vente ».....44
Figure 17: Modèle en étoile de l'activité de vente « Vente ».....45
Figure 18: Dimension Transporteur.....46
Figure 19: Fait « Transport ».....47
Figure 20: Modèle en étoile de l'activite de transport « Transport ».....48
Figure 21: Dimension Type_Stock.....49
Figure 22: Dimension Details_Frais.....50
Figure 23: Fait « Depenses ».....51
Figure 24: Modèle en étoile de l'activité « Depense ».....52

Figure 25: Fait « Stock_Export »....54
Figure 26: Modèle en étoile de l'activité « Stock_Export »....55
Figure 27: Dimension Paquet....56
Figure 28: Fait « Stock_Local »....57
Figure 29: Modèle en étoile de l'activité « Stock_Local »....59
Figure 30: Structure d'extraction des données de notre entrepôt des données....60
Figure 31: Exemple de l'opération de sélection....64
Figure 32: Résultat de l'opération de projection sur la table Usine....65
Figure 33: Exemple de l'opération d'Union....67
Figure 34: Exemple de l'opération d'Intersection....69
Figure 35: Exemple de l'opération de différence....70
Figure 36: Exemple de l'opération du produit cartésien....72
Figure 37: Exemple de l'opération de jointure....73
Figure 38: Exemple de l'opération de jointure naturelle....75
Figure 39**:** Exemple de l'opération de Division....77
Figure 40: Opération de rotation sur le cube Vente « Quantité de the vendu par usine et par année »....82
Figure 41: Opération de permutation sur le cube Vente « quantité du thé vendu par usine et par année »....83
Figure 42: Opération de division sur le cube Vente « quantité du thé vendu par usine et par année »....85
Figure 43 : Principe du DES....94
Figure 44: Algorithme de chiffrement par DES....96

Liste des tableaux

Tableau 1**:** Différences entres les données opérationnelles et les données décisionnelles 18
Tableau 2: Avantages du Modèle en étoile et en Flocon [AR19].....30
Tableau 3: Avantages et inconvénients de ROLAP, MOLAP et HOLAP.....32
Tableau 4: Le tableau récapitulatif des besoins.....37
Tableau 5: Tableau descriptif de la dimension temps.....40
Tableau 6: Tableau descriptif de la dimension Type_Vente.....41
Tableau **7**: Tableau descriptif de la dimension Type_Monnaie.....42
Tableau 8: Tableau descriptif de la dimension Quantite_the.....42
Tableau 9: Tableau descriptif e la dimension usine.....43
Tableau 10: Tableau descriptif du fait Vente.....44
Tableau 11: Tableau descriptif de la dimension Transporteur.....46
Tableau 12: Les dimensions communes entre l'activité « Transport » et « Vente ».....47
Tableau 13: Tableau descriptif du fait « Transport ».....48
Tableau 14: Tableau descriptif de la dimension Type_Stock.....50
Tableau 15: Tableau descriptif de la dimension Details_Frais.....50
Tableau 16: Dimensions communes ou non entre l'activité « Vente » et « Depenses »51
Tableau 17: Tableau descriptif du fait « Depenses ».....52
Tableau 18: Dimensions communes ou non entre l'activité « Vente » , « Depenses » et « Stock_Export ».....53
Tableau 19: Tableau descriptif du fait « Stock_Export ».....54
Tableau 20: Tableau descriptif de la dimension « Paquet ».....56
Tableau 21: Dimensions communes ou non entre l'activité « Vente » et « Stock_Local» 57
Tableau 22: Tableau descriptif du fait « Stock_Local ».....58
Tableau 23: Modèles mathématiques actuellement utilisés pour modéliser le système informatique et leurs rôles.....62

Tableau 24: Liste des cubes.....79

Résumé du thème

Depuis sa création dans les années quatre-vingt-dix, le système d’information décisionnel a été caractérisé par une allure montante mais aujourd’hui, on observe une évolution croissante de ce dernier dans les systèmes d’information grâce au rôle qu’il joue dans le processus de prise des décisions. En effet, il permet de collecter, consolider, modéliser et restituer les données d'une entreprise en vue d'offrir un outil de prise des décisions.

L’objectif principal de mon projet au sein de l’office du thé du Burundi consiste à concevoir et réaliser un système d’information décisionnel permettant de faciliter et de supporter les activités de l’office surtout la vente de ses produits, en réalisant un outil de prise de décisions qui appuiera les décideurs, et aussi d’élaborer des rapports corrects en temps réel afin de gérer et corriger les erreurs dues au travail manuel.

Pour arriver à cette fin, nous avons commencé par la phase de collecte et d’analyse des données ce qui nous a permis de déterminer le modèle de l’entrepôt des données à utiliser. Ensuite, nous avons utilisé le crypto système de Data Encryptions Standard pour assurer la protection des données et l’algèbre relationnelle comme outil mathématique de modélisation du comportement du système. Enfin, l’implémentation a été réalisée grâce aux outils de Microsoft comme SQL server 2014 et Visual Studio 2013 professionnel.

Table des matières

Dédicace.....i
Remerciements.....ii
Liste des abréviations.....iii
Liste des figures.....iv
Liste des tableaux.....vi
Résumé du thème.....vii
CHAPITRE I : INTRODUCTION GENERALE ET PRESENTATION DE L'OFFICE DE THE DU BURUNDI.....1
I.1 Justification du projet.....1
I.2. Objectifs du travail.....2
I.2.1 Objectif global du travail.....2
I.2.2 Objectifs spécifiques.....2
I.3 Problématique.....2
I.3.1 Solutions proposées.....3
I.3.2 Les résultats attendus.....3
I.3.3 Apport scientifique et technologique.....3
I.3.4 Domaine d'application.....4
I.4 : PRESENTATION DE L'OFFICE DE THE DU BURUNDI.....4
I.4.1. Historique de l'Office du Thé du Burundi.....4
I.4.2. Organigramme de l'Office du Thé du Burundi.....6
I.4.3. Description des tâches par poste.....7
I.4.4 Département agronomique.....10
I.4.5 Département industriel.....10
I.5. Direction commerciale.....11

I.5.1 Conseiller commercial.....11
I.5.2. Service Marché international.....12
I.5.3 Responsable de la gestion des échantillons et appui à la dégustation.....12
I.5.4 Responsable de la section export.....12
I.5.5 Département financier.....13
I.6 Mission et Vision de l'Office du Thé du Burundi.....13
I.6.1 Mission.....13
I.6.2 Vision.....13
I.6.3 Chiffre d'affaire.....13
I.7 Métier de production du The.....14
I.8 Autres producteurs du thé.....14
I.9 Informatique au sein de l'O.T. B.....14
CHAPITRE II : GENERALITES SUR LES SYSTEMES DECISIONNELS.....15
II.1 Introduction sur les systèmes décisionnels.....15
II.2 Définition.....15
II.3. Place du décisionnel dans l'entreprise.....16
II.3.1. Différentes composantes du décisionnel.....17
II.3.2 Différences entre le système opérationnel et le système décisionnel.....17
II.3.3 Architecture des systèmes décisionnels.....19
II.3.4 Conclusion.....20
CHAPIPITRE III : ENTREPOT DE DONNEES.....21
III.1 Définition de l'Entrepôt de données.....21
III.1.1 Historique des entrepôts de données.....22
III.1.2 Structure des données d'un Entrepôt des données.....23
III.1.3 Eléments d'un entrepôt de données.....24
III.2 Modélisation dimensionnelle.....25
III.2.1 Introduction.....25
III.2.1.1 Niveau conceptuel.....26
III.2.1.2 Concept Dimension.....27
III.2.1.3 Modèles structuraux des entrepôts de données.....28
III.2.1.3.1 Modèle en étoile.....28

III.2.1.3.2 Modèle en flocon de neige.....29
III.2.1.3.3 Modèle en constellation.....30
III.3 Niveau logique.....31
III.3.1 ROLAP.....31
III.3.2 MOLAP.....31
III.3.3 HOLAP.....31
III.3.4 OOLAP.....33
III.4 Navigation hypercube.....33
III.4.1 Navigation dans les données.....34
III.4.1.1 Slice & Dice.....34
III.4.1.2 Drill-down & Roll-up.....34
III. 4.1.3 Rapports.....35
III.5 Phases de l'alimentation « E.T.L. ».....35
III.6 Conclusion.....35
CHAPITRE IV : CONSTRUCTION D'UN ENTREPOT DE DONNEES.....36
IV.1 Collecte et définition des besoins.....36
IV. 1.1 Introduction.....36
IV.1.2 Contact avec les agents du service commercial.....36
IV.1.3 Consultation des rapports et détection d'autres besoins.....37
IV.2 Modélisation multidimensionnelle.....38
IV.2.1Introduction.....38
IV.2.2 Présentation de l'activité de vente du thé sec « Vente ».....39
IV.2.2.1 Dimensions participantes du modèle.....39
VI.2.3 Présentation de l'activité "Transport".....45
IV.2.4 Presentation de l'activite Dépenses.....49
IV.2.5 Présentation de l'activité de stockage « Stock_Export ».....53
IV.3. Phases d'alimentation "ETL".....60
IV.3.1 Introduction.....60
IV.3.2 Découverte des données.....60
IV.3.3 Extraction des données.....60
IV.3.4 Transformation des données.....61

IV.3.5 Chargement des données.....61
CHAPITRE V. MODELE MATHEMATIQUE DU SYSTEME.....62
V.1. Introduction.....62
V.2 Généralités sur l'algèbre relationnelle.....63
V.3 Sélection.....63
V.4. Projection.....64
V.5. Operateur ensembliste « Union ».....65
V.6. Intersection.....67
V.7 Différence.....69
V.8 Produit cartésien.....70
V.9. Jointures.....72
V.9.1 Définition.....72
V.9.2 Équi-jointure.....74
V.9.3 Jointure naturelle.....74
V.10 Division.....76
V.11 Conclusion.....77
CHAPITRE VI : IMPLEMENTATION DE LA SOLUTION.....78
VI.1 Outils logiciels et matériels.....78
VI.2 Cubes dimensionnels.....78
VI.2.1 Introduction.....78
VI.2.2 Liste des cubes.....79
VI.3 Présentation des interfaces.....80
VI.3.1 Interface générale.....80
VI.3.2 Interface de navigation des données.....80
VI.4 Présentation multidimensionnelle des données de l'entrepôt.....80
VI.4.1 Introduction.....80
VI.4.2 Opérations OLAP.....81
VI.4.2.1 Opérations agissant sur la structure.....81
VI.4.2.1.1 Rotation.....81
VI.4.2.1.2 Permutation.....83
VI.4.2.1.3 Division.....84

VI.4.2.2 Opérations agissant sur le contenu ou la granularité.....85
VI.4.2.2.1 Roll-up.....85
VI.4.2.2.2 Slicing.....88
VI.4.2.2.3 Dicing.....90
CHAPITRE VII : SECURITE DES DONNEES.....92
VII.1. Définition.....92
VII.1.1 Critères fondamentaux.....92
VII.1.2 L'identification et l'authentification.....93
VII.2 Sécurité des données de l’entrepôt.....93
VII.2.1 Historique Crypto système DES.....93
VII.2.2 Principe de fonctionnement de DES.....94
VII.2.3 Algorithme du DES.....95
VII.3 Conclusion.....97
CONCLUSION GENERALE ET RECOMENDATIONS.....98
Références et bibliographies.....100
ANNEXE.....102

CHAPITRE I : INTRODUCTION GENERALE ET PRESENTATION DE L’OFFICE DE THE DU BURUNDI

I.1 Justification du projet

Dans le domaine des entreprises, toutes les activités que ça soient commerciales ou non doivent être décidées par les dirigeants de l’entreprise afin d’atteindre les meilleurs objectifs qu’ils se sont donnés et éviter la concurrence et la chute de leur chiffre d’affaire. Il s’agit de prendre de bonnes décisions pour réussir. Pour ce, ces décisions doivent être basées sur des informations statiques bien calculées, fiables et lisibles. L’obtention de ces dernières demeure un problème majeur dans le cas où la plupart des entreprises utilisent encore des bases de données transactionnelles antérieures ne fournissant pas toutes ces informations nécessaires.

C’est pour cette raison que les systèmes décisionnels sont très utiles dans le domaine industriel. Ils permettent aux décideurs d’avoir accès aux informations pertinentes sur lesquelles ils se basent pour faire leurs choix décisionnels. Pour ce faire, ces systèmes utilisent les entrepôts de données pour la mise en place d’un bon système décisionnel.

Pour mieux accomplir ses activités commerciales de ventes des produits théicoles, l'Office du the du Burundi traite une grande quantité de données possédant souvent des redondances et des agrégations. Ces données peuvent participer pour améliorer le processus de prise de décision par les responsables et les décideurs de l'Office du thé du Burundi.

Ce travail de fin du cycle de Mastère vise la mise en place d'un système décisionnel permettant de consolider les données issues des bases de données transactionnelles, et d'offrir des informations de qualité pour les décideurs. Pour cela, un tel système requiert la mise en place d'un entrepôt de données fiables contenant les informations nécessaires à l'accomplissement des processus décisionnels.

I.2. Objectifs du travail

I.2.1 Objectif global du travail

Ce projet a pour objectifs globaux suivants :

- Concevoir et réaliser un entrepôt de données pour la mise en place d'un système décisionnel de vente des produits industriels
- Fournir des informations nécessaires et fiables lors de la prise des décisions et la réalisation des rapports.
- Analyser les données selon différents critères afin d'améliorer la prise des décisions dans l'entreprise.

I.2.2 Objectifs spécifiques

Les objectifs spécifiques de ce projet sont :

- Concevoir une base de données afin de collecter les données

- Concevoir un entrepôt de données
- Concevoir une application décisionnelle fournissant des informations fiables aux décideurs.
- Génération de rapports corrects
- Réaliser une interface permettant aux décideurs d'analyser les informations selon les critères voulus
- Réaliser un système permettant de réduire le temps dédié à la rédaction des rapports.

I.3 Problématique

L'Office du Thé du Burundi est une institution publique qui produit beaucoup de thé et qui possède aussi des clients et des marchés variés.

Dans nos jours, une telle société a besoin d'un système décisionnel permettant de gérer sa clientèle et ses marchés, de réaliser des rapports corrects et de prendre des décisions basées sur des informations fiables et efficaces.

On peut résumer les difficultés de la manière suivante :

- Difficultés dans l'élaboration des rapports d'activité ;
- Lenteur de la procédure pendant l'élaboration des rapports ;
- Coûts des rapports élevés ;
- Manque des informations fiables pendant l'analyse des données pour la prise des décisions ;
- Manque d'interface permettant aux analystes et décideurs de faire des analyses.

I.3.1 Solutions proposées

- Génération automatique des rapports de ventes du thé sec.
- Recherche automatisée des meilleurs marchés
- Module d'aide de prise de décision pour les décideurs
- Module des statistiques des ventes par clients
- Former les utilisateurs pour le nouveau système décisionnel

I.3.2 Les résultats attendus

Les résultats attendus pour ce travail de thèse de Mater sont :

- Les décideurs de la société de l'Office du Thé du Burundi disposeront un système décisionnel ;
- Un outil de génération automatique de rapport sera déployé ;
- Les produits de l'OTB seront distribués selon les besoins des points de vente ;
- Un entrepôt de données sera mis en place pour l'OTB ;
- Les utilisateurs du système décisionnel seront formés ;

I.3.3 Apport scientifique et technologique

- Conception et implémentation des applications décisionnelles
- Intégration des outils d'analyse des services ;
- Développement des applications Windows ;
- Aide à la prise de décision ;
- Méthodes heuristiques de conception et de prise de décision ;

I.3.4 Domaine d'application

- Informatique décisionnelle ;
- Entrepôt des données ;
- Analyse des données ;
- Société de production de produits comestibles ;
- Système d'aide à la prise de décision ;

I.4 : PRESENTATION DE L'OFFICE DE THE DU BURUNDI

I.4.1. Historique de l'Office du Thé du Burundi

Ce n'est qu'en 1971 que l'Office du Thé du Burundi fut créé par décret présidentiel n°1/79 du 30 Juillet 1971. A la même année, le Projet Thé TEZA est transformé en une société théicole par ordonnance ministériel n°050/126 du 05 août 1971.

L'évolution de l'environnement institutionnel de l'OTB a été caractérisée par un recentrage continu de ses structures.

Les autres projets théicoles seront transformés en Sociétés théicoles ultérieurement, notamment RWEGURA en 1979 par ordonnance ministériel n°710/19 du 14/02/1979 et Tora en 1975, par ordonnance ministériel n°710/142 du 31/07/1975.

L'Office du Thé du Burundi était donc conçu comme une société Holding chargée de superviser, d'encadrer et de contrôler les Sociétés théicoles existantes, tout en poursuivant la conception et la réalisation de nouveaux projets théicoles qui deviendraient des sociétés à leur tour dès qu'ils seraient arrivés à maturité.

Après la mise en place de l'OTB et de certaines sociétés théicoles, d'autres réformes statutaires ont eu lieu notamment ? Le décret n°100/82 du 19/06/1994 portant modification des statuts de l'OTB pour les adapter au décret-loi n° 1/30 du 10/10/1978 portant cadre organique des Etablissements Publics.

L'encadrement de la production théicole est confié au Projet Cultures Villageoises en Haute Altitude (CVHA) et les sociétés théicoles restent des filiales de l'OTB dotées d'une certaine autonomie, tandis que tous les choix stratégiques sont du ressort des organes de l'OTB. Le décret n°100/066 du 19/04/1990 retire la personnalité juridique aux sociétés théicoles. Celles-ci deviennent des filiales de l'OTB appelés « Complexe théicole ». L'OTB subsiste comme seule Entreprise responsable de la production et de la commercialisation du thé.

Même l'encadrement de la théiculture est restitué à l'OTB en 1994 pour que celui-ci puisse maîtriser toute la filière théicole. Le décret n°100/157 du 05/09/1997 a été pris dans le souci d'harmoniser les statuts de l'Office du Thé du Burundi avec les dispositions du Codes des Sociétés de l'époque.

Ce texte renforce l'intégration des complexes au sein de l'OTB. Les gérants des complexes anciennement nommés par le Ministre de tutelle deviennent des cadres nommés par le Conseil d'Administration sur proposition du Directeur Général et placés sous l'autorité de ce dernier. Le Décret n°100/263 du 20 septembre 2007

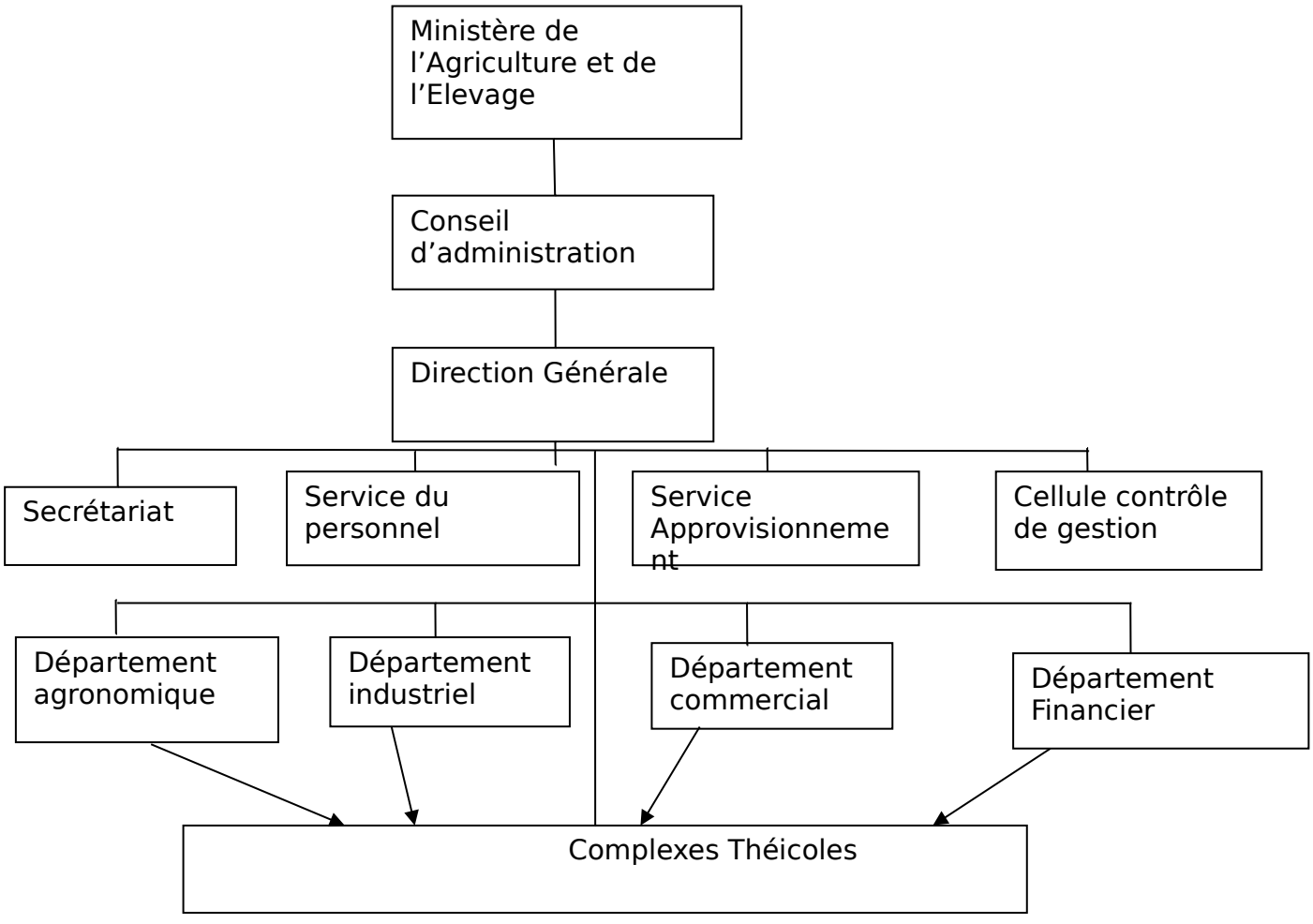
porte sur la libéralisation de la filière thé et le Décret n°100/01 du 08 janvier 2013 régissant actuellement l'OTB a été pris encore une fois dans le cadre de l'harmonisation des statuts de l'Office du Thé du Burundi, « OTB-SP » avec le code des sociétés privées et à participation publique.

Ce décret introduit deux nouveautés :

1)le capital social de l'OTB passe de 800 millions à 20 milliards par transformation de la créance de l'Etat en capital en plus de l'incorporation de quelques réserves dans le capital ;
2)la représentation des théiculteurs au Conseil d'Administration de l'OTB.

Noter que jusqu'aujourd'hui l'OTB comporte 5complexes théicoles à savoir : RWEGURA, TEZA, JENDA, TORA et BUHORO.

I.4.2. Organigramme de l'Office du Thé du Burundi



: Lien fonctionnel

: Lien hiérarchique

I.4.3. Description des tâches par poste

I.4.3.1. Conseil d'administration

Dans les limites de l'objet social, le conseil d'administration a le rôle de :

- définir les orientations de l'action de l'office et prendre toute décision nécessaire à son administration ;
- voter le budget prévisionnel, approuver les comptes de l'exercice écoulé sur rapport des commissaires aux comptes, décider de l'affectation du solde bénéficiaire et donner quitus au Directeur général après l'approbation des comptes de l'exercice écoulé ;
- déterminer les modalités et les limites dans lesquelles le Directeur Général peut recourir au crédit ;
- fixer la rémunération des administrateurs, du Directeur Général, des Commissaires aux comptes et du reviseur indépendant.

I.4.3.2. Direction Général de l'OTB siège

Dans le cadre des orientations fixées par le conseil d'administration, le Directeur Général :

- assure la gestion quotidienne de l'Office et exécute les décisions du conseil d'administration ;
- négocie les contrats d'assistance technique ou de financement dans le cadre de la coopération internationale et les soumet à l'approbation du conseil d'administration ;
- conclut et exécute les contrats tant avec les fournisseurs qu'avec les clients conformément aux lois et usage du commerce ;
- engage et révoque le personnel dans le respect des dispositions du code de travail ;
- Recourt au crédit bancaire dans les limites et selon les fixées par le conseil d'administration ;
- représente l'office auprès de l'administration, de la justice et des tiers ;

I.4.3.3. Secrétariat

Les principales tâches sont :

- le traitement du courrier (réception, dépouillement, enregistrement, acheminement à la direction général, pointage du courrier traité par le DG, distribution au sein des départements, complexes ou services) ;
- la dactylographie de tous les travaux de la direction général et du service du personnel, le classement et l'expédition du courrier (dans les différents départements complexes et services, dans les différents services partenaires en mairie de Bujumbura et à l'étranger ;
- l'envoi et la réception des messages fax ;
- la réception des communications téléphoniques du DG (les trier, les enregistrer et les lui annoncer) ;
- le suivi des abonnements à différents journaux ;
- l'accueil des visiteurs et la fixation des rendez-vous ;

I.4.3.4. Service du personnel

Ses principales fonctions sont :

- conseiller le directeur Général sur tous les aspects de la vie de l'OTB en général sur les aspects relatifs au personnel en particulier ;
- collaborer avec les départements techniques et les complexes, gérer les dossiers du personnel en général et du personnel de l'OTB siège en particulier ;
- mettre à jour les dossiers du personnel de l'OTB siège
- Vérifier quotidiennement les registres de présences ;
- veiller à l'établissement régulier des journaux de paie au calcul des diverses cotisations, et à leur versement ;
- proposer le programme de formation du personnel ;
- préparer des journaux de paie pour le personnel de l'OTB siège
- classer tous les dossiers du personnel de l'OTB et la préparation des décisions de congé ;
- saisir toutes les données relatives aux personnels du siège

I.4.3.5. Service Approvisionnement et gestion des stocks :

Au niveau des approvisionnements, on effectue l'ouverture et le suivi de tous les dossiers de commande pour l'import et le local c'est à dire préparation et lancement de l'avis d'appel d'offres ou des consultations restreintes auprès des fournisseurs potentiels.

- l'établissement des lettres de marché et de bons de commande
- le maintien des relations avec les fournisseurs, transitaires /transporteurs et avec les assureurs dans le cadre du suivi des commandes,
- organisation des réceptions de fournitures et transmission au département financier des demandes de paiement des factures y relatives ;
- le contrôle périodique de la petite caisse approvisionnée.
- la participation à toute procédure de passation de marchés, analyse des offres et attribution des marchés, réception des travaux et fournitures.

Au niveau de la gestion des stocks, le contrôle périodique de la tenue des fichiers de stocks et de la situation physique de ces derniers ;

- La transmission après vérification des rapports des mouvements des stocks (vérification des bons d'entrée et de bons d sortie ainsi que du rapport global des mouvements de stocks après chaque mois).
- La vérification de la valorisation pour les produits importés
- L'organisation de l'expédition des engrais chimiques aux différents complexes théicoles.

I.4.3.6 Cellule Contrôle de gestion

Cette cellule anime et coordonne toutes les activités de la cellule, établit les programmes de travail, leur suivi et gère le personnel de la cellule ,veille à ce que l'ensemble des budgets annuels de l'office soient élaborés dans les délais et qu'ils soient adoptés par le conseil d'administration avant le début de l'exercice concerné .Elle assure également le suivi mensuel, trimestriel et annuel de l'exécution de ces budgets et s'assure que les rapports périodiques sortent dans les délais et dégage périodiquement les couts de production du thé sec ainsi que les résultats d'exploitation de chaque unité de gestion de l'ensemble de l'Office :

- s'assure de la fiabilité des données du suivi budgétaire en faisant les rapprochements avec les données des comptes de gestion de la comptabilité ;
- fait analyse des résultats de gestion et soumet des propositions concrètes à la direction générale en vue d'améliorer la gestion aussi bien au niveau du siège que des complexes.

I.4.4 Département agronomique

Il Effectue le suivi de toutes les plantation théicoles, des boisements et des pistes et propose à la direction générale les stratégies à mettre en œuvre en vue de promouvoir le développement de la théiculture. Ce département participe à l'élaboration des conventions des collaborations avec l'ISABU et assure la mise en application de nouveaux paquets technologiques éprouvés

Il joue aussi le rôle de conseiller agronomique auprès des complexes théicoles, élabore des notes techniques sur les différents thèmes, évalue le travail du personnel agricole des complexes et assure la formation du personnel agricole :

- Veille à la disponibilisation et la bonne distribution du matériel et de l'équipements agricoles, pour couvrir les besoins en matière intrants ;
- Étudie les possibilités de mise en place des associations de théiculteurs ;
- Prépare le programme de travail et le budget annuels du département ;
- Élabore tout document et tout le rapport lui demandés par la direction générale.

I.4.5 Département industriel

Le département industriel comprend le secrétariat, le service usines et formation et service gestion du charroi. Il coordonne les activités des deux services, propose

annuellement un programme d'investissement à court et moyen termes (les extensions, les transformations et les acquisitions en nouveaux équipements des usines, la réhabilitation des équipements), vérifie les besoins en pièces détachées des usines (quantités et spécifications techniques) et les consolide avant la commande.

Ce département participe également à la réception des équipements et pièces détachées et assure le suivi de l'exécution des grands travaux d'investissement dans le domaine des usines, assure la formation des superviseurs techniques ou organise (au Burundi ou à l'étranger) ainsi que tout appui technique nécessaire aux usines

I.5. Direction commerciale

Les tâches assignées au Directeur commercial sont les suivantes :

- Suivi de la qualité du thé par des dégustations systématiques des échantillons lots fabriqués ;
- Commercialisation du thé sur les divers marchés (enchères, ventes privés, vente locale) négociations des contrats privés et recherche de nouvelles opportunités d'affaires ;
- Suivi permanent de l'évolution des frais commerciaux et des prestations des partenaires de l'office (transporteurs, entreposeurs, courtiers…)
- Élaboration des rapports périodiques sur la commercialisation du thé.

I.5.1 Conseiller commercial

Placé sous l'autorité du directeur Général, il appuie le département commercial. Il est chargé de :

- Déguster d'une façon systématique les échantillons des lots fabriqués.
- Relever les défauts de fabrication et les communiquer immédiatement aux usines en vue de procéder à leur correction ;
- Visiter régulièrement les unités de productions en vue d'inspecter le processus de fabrication dans le but de s'assurer du respect des normes d'usinage ;

- Veiller à ce que la qualité du thé fabriqué réponde aux besoins du marché par des communications régulières et échange d'échantillons avec les courtiers changés de vendre le thé burundais sur le marché international ;
- Former les responsables de productions (tea-makers) de toutes les usines en matière de techniques de dégustation et de contrôle de la qualité en vue de doter les unités de production des capacités autonomes de contrôle de la qualité en vue de doter les unités de production des capacités autonomes de contrôle de suivi de la qualité du thé en production ;
- Élaborer des rapports périodiques sur la qualité du thé burundais.

I.5.2. Service Marché international

Le chef de service marché internationale quant à lui doit contrôler la qualité du thé par des dégustations systématiques, veiller à l'expédition des échantillons aux courtiers et aux acheteurs privés, suivre la rotation hebdomadaire des stocks de thé entreposés à l'extérieur du Burundi, établir une situation des encaissements mensuels sur les ventes aux enchères et sur les ventes privées, établir une situation des encaissements mensuels sur les ventes aux enchères et sur les ventes privées.

C'est lui qui établit des programmes hebdomadaires de chargement en accord avec des transporteurs, produit des tableaux de bord périodiques (quantités de thé mensuellement réceptionnées à Bujumbura et exportées, situation mensuelle des stocks au Burundi et à l'étranger, des frais commerciaux et des encaissements), vérifie l'arrivée à destination des lots intacts et autorise le paiement des factures de transport. Il établit mensuellement les statistiques commerciales commentées, et suit l'évolution des frais commerciaux ;

I.5.3 Responsable de la gestion des échantillons et appui à la dégustation

Il est chargé de réceptionner les échantillons des lots (triages, classement) et les expédier en fonction des allocations (échantillons « types » pour les courtiers, échantillons pour les clients privés, les « type samples » etc…), d'enregistrer sur les fiches de contrôle des contrats privés la date d'expédition des échantillons ;

C’est lui qui prépare la dégustation, note et dactylographie les observations et les transmettre aux usines. Il range un exemple de chaque échantillon par ordre numérique, grade et usine et Classe tous les documents relatifs aux échantillons et entretient la salle et les équipements de dégustation. Il doit participer aux foires, expositions ….

I.5.4 Responsable de la section export

Il supervise le travail du responsable de l’entrepôt (chargements, déchargements) et du responsable de la sous-section « assurances » ainsi que les opérations de chargement à l’export (bons de chargement, état des camions et containers, fiches de chargement, programmes chargements). Il établit également les documents commerciaux (parking list, certificat d’origine, de poids phytosanitaire, facture commerciale) et tient un registre de suivi des encaissements en cours, dossiers en assurances …)

I.5.5 Département financier

Le rôle de ce dernier est d’animer et organiser le personnel du département, répartir et contrôler l’exécution des tâches du département. Il veille à la mise en application des procédures et supervise le travail du chef comptable OTB notamment au niveau des arrêtés de bilan,

Il est également responsable de l’optimisation de la trésorerie et des ressources financières de l’OTB, de la constitution et du suivi des dossiers de financement ; des relations avec les bailleurs de fonds, et de la gestion de la dette.

I.6 Mission et Vision de l’Office du Thé du Burundi

I.6.1 Mission

L'Office du Thé du Burundi (O.T.B) est une société publique agroalimentaire créée en 1971. Sa principale mission est la production de la culture du thé au Burundi à travers l'encadrement des petits producteurs, la gestion des plantations industrielles, la transformation du thé et la commercialisation du thé sec. L’O.T. B travaille en collaboration avec environ 60.000 Exploitants théicoles et centralise la production de 2239 Ha de plantations industrielles et 7087 Ha de plantations villageoises.

La production annuelle moyenne est de 10 .952 Tonnes de thé sec. La production du thé provient à 20% des plantations industrielles appartenant à l’OTB et à 80% des plantations villageoises appartenant aux paysans.

I.6.2 Vision

L’office de Thé du Burundi a comme vision de faire évoluer de la production de la feuille verte et thé sec (kg).

I.6.3 Chiffre d’affaire

Le capital social de l’OTB passe de 800 millions à 20 milliards par transformation de la créance de l’Etat en capital en plus de l’incorporation de quelques réserves dans le capital ;

I.7 Métier de production du The

L'OTB a pour rôle promouvoir la théiculture au Burundi et partant, de contribuer à l’apport en devises et de procurer des revenus aux théiculteurs. A travers ses départements agronomique, industriel, commercial et financier, le siège appuie les complexes théicoles pour ce qui concerne les questions techniques, commerciales, administratives et financières. Trois services sont rattachés directement à la direction générale : un service d’approvisionnement, un service de contrôle de gestion et un service de gestion du personnel. Huit cents personnes travaillent de manière permanente au siège de l’OTB. Celles-ci interviennent au niveau de tous les maillons de la filière. Outre ce personnel permanent, il y a également une main-d’œuvre temporaire équivalent à environ huit millions hommes-jours par an, ce qui équivaut approximativement à quatre mille travailleurs à temps plein. Cette main d’œuvre intervient surtout en période de pointe comme par exemple pendant la cueillette. Au niveau de la recherche, c’est l’Institut des Sciences Agronomiques du Burundi (ISABU) qui intervient dans la filière. L’OTB a aussi d’autres partenaires socioéconomiques avec lesquels il échange des biens et services.

La filière théicole du Burundi est organisée selon un schéma d’intégration verticale où l’OTB assure l’ensemble des fonctions, de la fourniture des plants et des engrais à l’exportation du produit fini en passant par la collecte et la transformation de la feuille verte.

I.8 Autres producteurs du thé

Cependant, actuellement, l’OTB n’est plus le seul acteur agro-industriel de la filière. En effet, suite au processus de libéralisation / privatisation engagé par le gouvernement depuis 2007, quatre sociétés ont été agréées pour intervenir dans le secteur et l’une d’entre elles (PROTHEM) a débuté son activité d’achat des feuilles vertes et de transformation en 2011.

I.9 Informatique au sein de l'O.T. B

La majorité des activités sont faites manuellement à l'office de thé du Burundi. Les données liées aux activités (achats, vente et production) sont complétées manuellement sur des fiches, puis saisies à la machine pour être conservées sous les fichiers de Microsoft Excel et Word.

CHAPITRE II : GENERALITES SUR LES SYSTEMES DECISIONNELS

II.1 Introduction sur les systèmes décisionnels

Dans ce chapitre nous allons définir le système décisionnel et donner quelques concepts relatifs au business intelligence avec des applications pour boucler par une conclusion puisque on a remarqué qu'il existe des problèmes au moment de la prise des décisions sur l'évolution de l'entreprise et problèmes y relatifs.

Notons qu'un nombre plus important d'acteurs des entreprises éprouvent des difficultés dans le processus de la prise des décisions notamment au niveau de stratégies d'organisation de l'entreprise. Il faut changer la stratégie de diriger l'entreprise en se référant sur des informations et données.

L'informatique décisionnelle désigne les moyens, les outils et les méthodes qui permettent de collecter, consolider, modéliser et restituer les données, matérielles ou immatérielles d'une entreprise en vue d'offrir une aide à la décision et de permettre aux responsables de prendre des stratégies pour l'entreprise et d'avoir une vue d'ensemble de l'activité traitée au sein de l'entreprise.

Les systèmes d'information décisionnels (SID) sont utilisés pour faciliter l'accès, l'interrogation et l'analyse de l'information d'une organisation pour ses décideurs. La dernière évolution notable des SID repose sur les concepts d'entrepôt de données ou data warehouse et d'OLAP ou On-Line Analytical Processing. [AR1]

II.2 Définition

Afin de mieux comprendre la finalité des systèmes décisionnels, nous nous devons de les placer dans leurs contextes et rappeler ce qu'est un système d'information.

« Le système d'information est l'ensemble des méthodes et moyens de recueil de contrôle et de distribution des informations nécessaires à l'exercice de l'activité en tout point de l'organisation. Il a pour fonction de produire et de mémoriser les informations, de l'activité du système opérant (système opérationnel), puis de les mettre à disposition du système de décision (système de pilotage) » [**Le Moigne, 1977**]

Les différences qui existent entre le système de pilotage et le système opérationnel, du point de vue fonctionnel ou des tâches à effectuer, conduisent à l'apparition des *« systèmes d'information décisionnels » (S.I.D.)*. Ces différences seront clairement illustrées un peu plus loin dans notre document.

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

II.3. Place du décisionnel dans l'entreprise

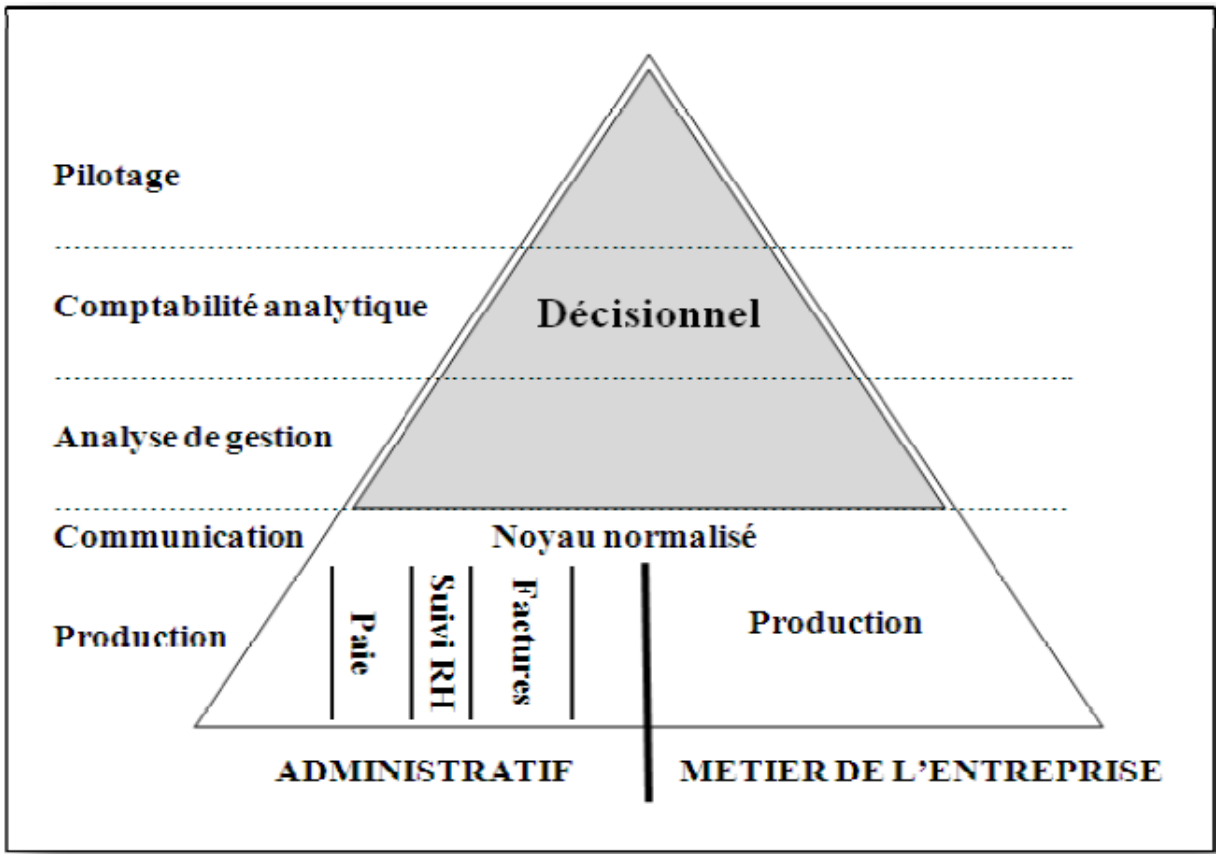


Figure 1: Environnement Décisionnel au sein du système d'information

La figure précédente illustre parfaitement la place qui revient au décisionnel au sein d'une entreprise. Cette place comprend plusieurs fonctions clés de l'entreprise. Les finalités décisionnelles étant différentes selon le poste et la fonction occupée, ont pour but d'engendrer plusieurs composantes. [AR2]

II.3.1. Différentes composantes du décisionnel

En relation étroite avec les nouvelles technologies de l'information et des télécommunications, le système décisionnel se manifeste à différents niveaux selon leurs utilités et leurs missions principales, comme illustré dans la figure suivante :

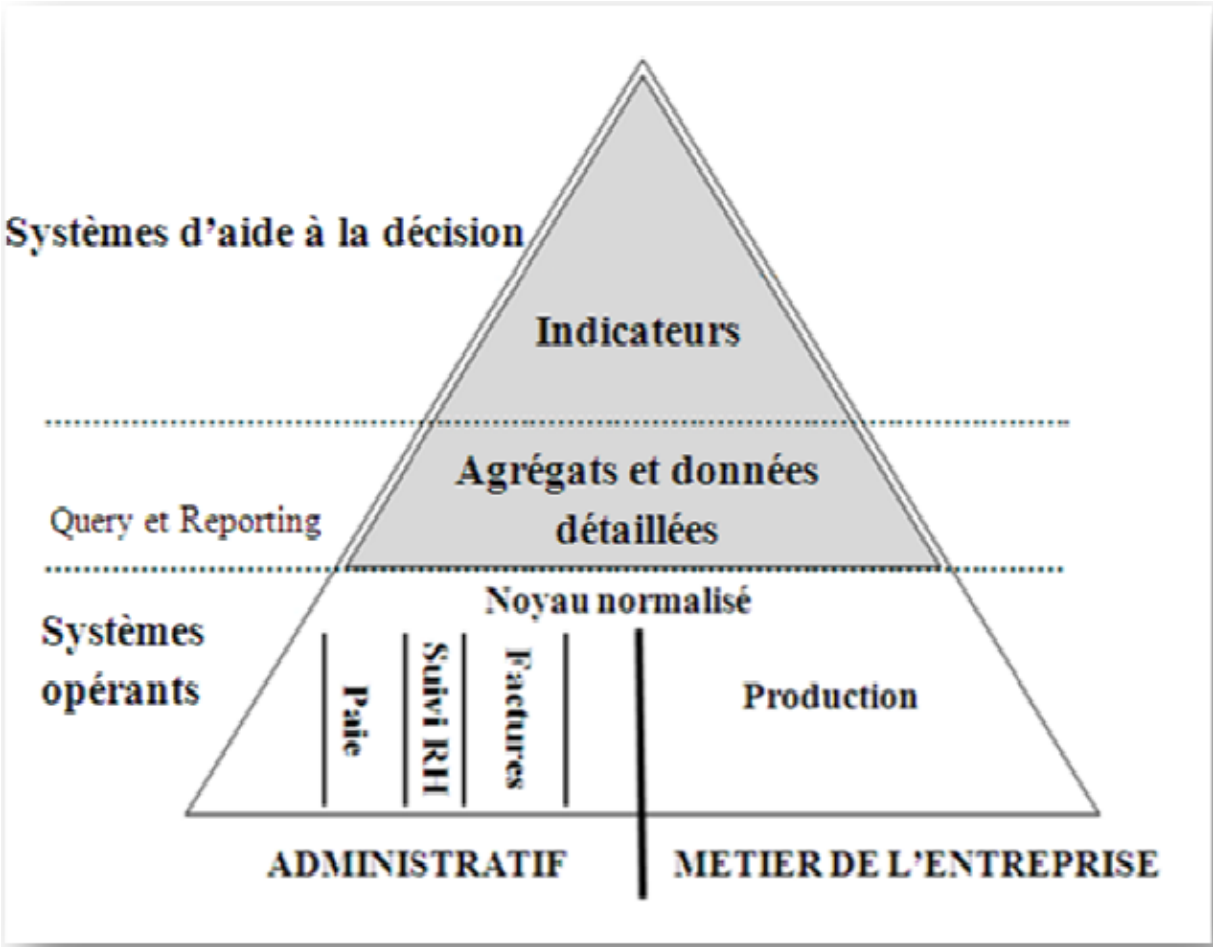


Figure 2: Différentes composantes du système décisionnel

La figure précédente illustre comment le système décisionnel se manifeste dans une entreprise sur différents niveaux selon leurs utilités et leurs missions principales. **[AR2]**

II.3.2 Différences entre le système opérationnel et le système décisionnel

Les systèmes « opérationnels » ou « de gestion », également appelés systèmes OLTP (on-line transaction processing), sont dédiés aux métiers de l'entreprise pour les assister dans leurs tâches quotidiennes et opérationnelles de gestion. La tendance est à l'utilisation de P.G.I. (progiciels de gestion intégrée) qui regroupent tous les logiciels

de gestion de l'entreprise – finances, ressources humaines, logistique, ventes, etc. – en un unique progiciel paramétrable aux règles de l'entreprise, organisé autour d'une base de données, réduisant ainsi les coûts de communications entre applications. **[AR3]**

Les systèmes « décisionnels », également appelés OLAP (*on-line analytical processing*), sont dédiés au management de l'entreprise pour l'aider au pilotage de l'activité, et donc indirectement opérationnels. Ils offrent aux décideurs une vision transversale de l'entreprise. La tendance pour réaliser un système décisionnel est à la mise en place d'un entrepôt de données.

Le tableau suivant résume de façon non exhaustive les différences qu'il peut y avoir entre les systèmes transactionnels et les systèmes décisionnels selon les données et l'usage fait des systèmes.

Tableau 1**:** Différences entres les données opérationnelles et les données décisionnelles

Données opérationnelles	**Données décisionnelles**
Orientées applications, détaillées, précises au moment de l'accès	Orientées activité (thèmes, sujet), condensées, représentent des données historiques
Mise à jour interactive possible de la part des utilisateurs	Pas de mise à jour interactive de la part des utilisateurs
Accédées de façon unitaire par une personne à la fois	Utilisées par l'ensemble des analystes, gérées par sous-ensembles
Haute disponibilité en continu	Exigence différente., haute disponibilité ponctuelle
Unique (pas de redondance en théorie)	Peuvent être redondantes
Petite quantité de données utilisées par un traitement	Grandes quantités de données utilisées par le traitement
Réalisation des opération au jour le jour	Cycle de vie diffèrent
Forte probabilité d'accès	Faible probabilité d'accès
Utilisées de façon répétitive	Utilisées de façon aléatoire

Ces différences font ressortir la nécessité de mettre en place un système répondant aux besoins décisionnels. Ce système n'est rien d'autre que le « ***Data Warehouse*** ». [AR3]

II.3.3 Architecture des systèmes décisionnels

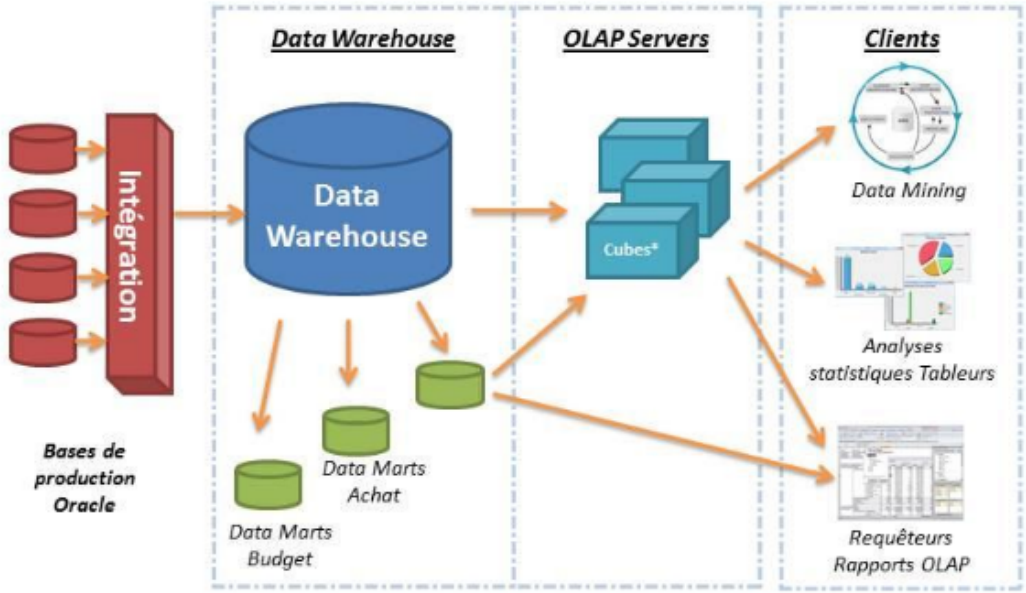


Figure 3: Architecture des systèmes décisionnels

Sur ce schéma de l'architecture du système décisionnel, même si on n'a pas montré d'autres systèmes de gestion des bases de données, on tient à vous informer qu'on peut implémenter non seulement en Oracle mais aussi en SQL Server, MySQL et d'autres. Tout dépend du choix du développeur.

L'entrepôt de données est le support nécessaire à la réalisation d'une architecture qualifiée de décisionnelle, qui va permettre, comme l'indique son nom, l'aide à la décision grâce au processus d'analyse qu'elle offre. Cette architecture décisionnelle, peut être représentée classiquement selon le schéma de la figure précédente. On peut distinguer la partie source, la partie gestion de données et enfin la partie analyse. On parle généralement d'architecture n-tiers en raison de différentes couches possibles pour gérer les données (entrepôt, magasin, etc.) et réaliser des analyses. **[AR4]**

II.3.4 Conclusion

Au cours ce chapitre, nous avons pu décrire le système décisionnel, nous avons donné une idée générale de l'entrepôt de données qui est l'objet principal de notre travail. Ce chapitre nous met dans un bain général des activités que nous allons réaliser dans la suite de ce projet. La bonne réalisation de l'entrepôt repose sur une bonne conception du schéma, car c'est ce dernier qui déterminera quelles sont les possibilités d'analyse des données du système décisionnel selon des axes différents pour faciliter le processus de prise de décisions.

CHAPIPITRE III : ENTREPOT DE DONNEES

III.1 Définition de l'Entrepôt de données

Selon Bill Inmon, dans son livre considéré comme étant la référence dans le domaine "Building the Data Warehouse" l'Entrepôt de données est défini comme suit :

« L'entrepôt de données est une collection de données orientées sujet, intégrées, non-volatiles et évolutives dans le temps, organisées pour le support d'un processus d'aide à la décision. » [AR7]

D'après la définition ci-haut donnée, voici les caractéristiques d'un entrepôt de données :

Il est Orienté sujet.

L'orientation sujet implique que l'Entrepôt de données est organisé autour des sujets majeurs de l'entreprise, contrairement à l'approche transactionnelle utilisée dans les systèmes opérationnels, qui sont conçus autour d'applications et de fonctions telles que : cartes bancaires, solvabilité client…, les Data Warehouse sont organisés autour de sujets majeurs de l'entreprise tels que : clientèle, ventes, produits…. Cette organisation affecte forcément la conception et l'implémentation des données contenues dans le Data Warehouse. Le contenu en données et en relations entre elles diffère aussi. Dans un système opérationnel, les données sont essentiellement destinées à satisfaire un processus fonctionnel et obéit à des règles de gestion, alors que celles d'un Data Warehouse sont destinées à un processus analytique.

Intégrée l'Entrepôt de données va intégrer des données en provenance de différentes sources. Cela nécessite la gestion de toute incohérence.

Evolutives dans le temps Dans un système décisionnel il est important de conserver les différentes valeurs d'une donnée, cela permet les comparaisons et le suivi de l'évolution des valeurs dans le temps, alors que dans un système opérationnel la valeur d'une donnée est simplement mise à jour. Dans un Entrepôt de données chaque valeur est associée à un moment.

Every key structure in the data warehouse contains - implicitly or explicitly -an element of time » **[AR7]**.

Non volatiles » : c'est ce qui est, en quelque sorte la conséquence de l'historisation décrite précédemment. Une donnée dans un environnement opérationnel peut être mise à jour ou supprimée, de telles opérations n'existent pas dans un environnement Entrepôt de données.

« Organisées pour le support d'un processus d'aide à la décision » : Les données de l'Entrepôt de données sont organisées de manière à permettre l'exécution des processus d'aide à la décision (Reporting, Data Mining…).

« Résumées » : Les informations issues des sources des données doivent être agrégées et réorganisées afin de faciliter le processus de prise des décisions.

« Disponibles pour l'interrogation et l'analyse » : Les utilisateurs doivent pouvoir consulter les données réorganisées de l'entrepôt en fonction de leurs droits d'accès au travers d'outils interactifs d'aide à la manipulation et l'analyse.

III.1.1 Historique des entrepôts de données

Les entrepôts de données ont été utilisés pour la première fois en 1990 par Bill Immon. Il était question de concevoir une base de données orientée sujet, intégrée et contenant des informations historisées, non volatiles et exclusivement destinés aux processus d'aide à la décision**. [AR7]**

En réalité, la raison de cette conception ne suffit pas pour pérenniser. Elle est un système ouvert sur son environnement au cœur des systèmes d'informations confrontées à des phénomènes économiques et sociaux lourds de conséquences. Pour faire face aux nouveaux enjeux, l'entreprise doit collecter, traiter, analyser les informations de son environnement pour suivre son évolution.

Il devient fondamental de rassembler et de centraliser les données afin de permettre l'analyse des indicateurs pertinents pour faciliter la prise de décision. Le but de l'entrepôt de données est de définir et d'intégrer une architecture qui sert de base aux applications décisionnelles.

III.1.2 Structure des données d'un Entrepôt des données

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

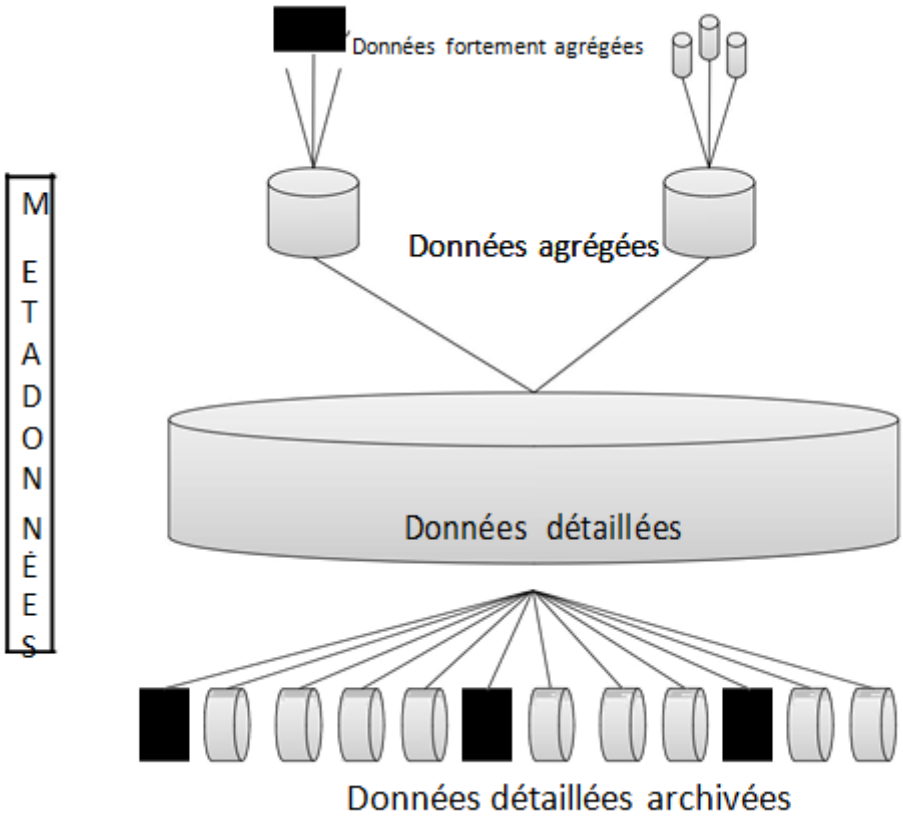


Figure 4: Structure des données d'un entrepôt de données.

Données détaillées

Ce sont les données qui reflètent les événements les plus récents, fréquemment consultées, généralement volumineuses car elles sont d'un niveau détaillé.

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

Données agrégées : données agrégées à partir des données détaillées.

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

Meta données : ce sont les informations relatives à la structure des données, les méthodes d'agrégation et le lien entre les données opérationnelles et celles d'un entrepôt de données. [AR2]

Le modèle de données,

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

Voici sur la figure ci-dessous l'architecture de d'un entrepôt de données : [AR8]

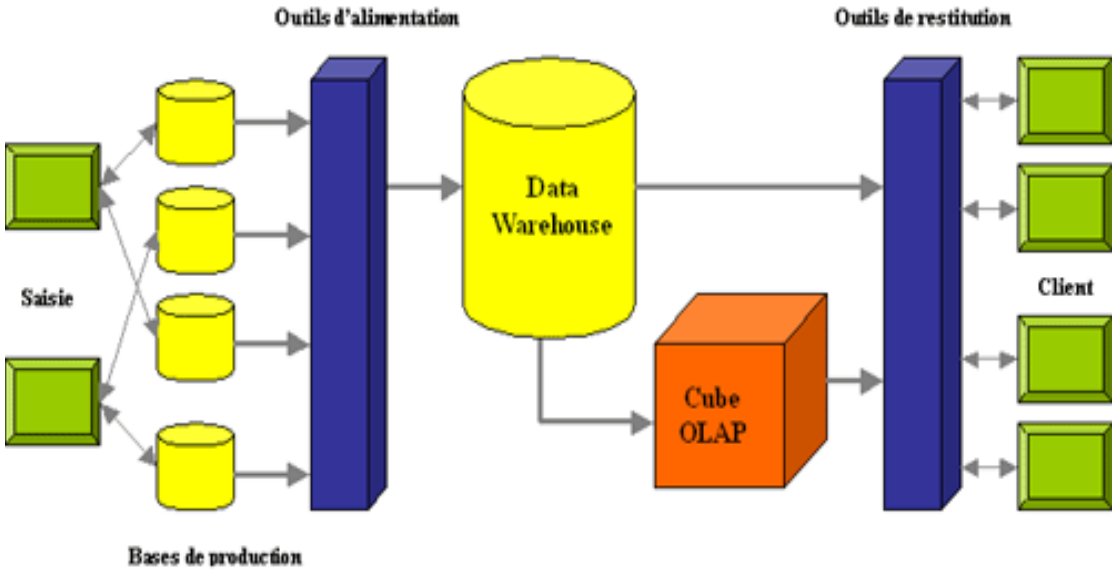


Figure 5**:** Achitecture de l'entrepot de données

III.1.3 Eléments d'un entrepôt de données

L'environnement d'un entrepôt de données est constitué essentiellement de quatre composantes : les applications opérationnelles, la zone de préparation des données, la présentation des données et les outils d'accès aux données.

Applications opérationnelles : ce sont les applications du système opérationnel de l'entreprise et dont la priorité est d'assurer le fonctionnement de ce dernier et sa performance. Ces applications sont extérieures à l'entrepôt de données

Préparation des données : la préparation englobe tout ce qu'il y a entre les applications opérationnelles et la présentation des données. Elle est constituée d'un ensemble de processus appelé ETL, « Extract, transform and Load », les données sont extraites et stockées pour subir les transformations nécessaires avant leur chargement. [AR2]

«Un point très important, dans l'aménagement d'un entrepôt de données, est d'interdire aux utilisateurs l'accès à la zone de préparation des données, qui ne fournit aucun service de requête ou de présentation » **[**AR10].

Présentation des données : c'est l'entrepôt où les données sont organisées et stockées. Si les données de la zone de préparation sont interdites aux utilisateurs, la zone de présentation est tout ce que l'utilisateur voit et touche par le biais des outils d'accès.

L'entrepôt de données est constitué d'un ensemble de Data Mart. Ce dernier est défini comme étant une miniaturisation d'un Data Warehouse, construit autour d'un sujet précis d'analyse ou consacré à un niveau départemental.

III.2 Modélisation dimensionnelle

III.2.1 Introduction

La modélisation dimensionnelle (modèle multidimensionnel) souvent appelée modélisation OLAP (Codd1993) se présente comme une alternative au modèle relationnel. Elle correspond mieux aux besoins du décideur tout en intégrant la modélisation par sujet c'est une méthode de conception logique qui vise à présenter les données sous une forme standardise et qui permet des accès hautement performant.

Cette Modélisation dimensionnelle consiste également à considérer un sujet d'analyse comme un cube à plusieurs dimensions, offrant des vues en tranches ou des analyses selon différents axes. Elle se présente sous deux niveaux :

- Conceptuel ;
- Logique ;

III.2.1.1 Niveau conceptuel

Conceptuellement, cette modélisation multidimensionnelle a donné naissance aux concepts de fait et de dimension. [AR10]

Le sujet analysé est représenté par le concept de fait. La table de fait est composée d'un ensemble de mesures qui représentent les différentes valeurs de l'activité analysée. Chaque table des faits contient au moins deux clés étrangères qui la connectent aux dimensions.

Chaque ligne dans la table des faits correspond à un événement mesurable, ce qui représente le grain (niveau de granularité). [AR9]

Le grain désigne la signification d'une ligne dans une table des faits. Les tables dimensions et faits doivent être cohérentes avec le grain défini ce qui crée une cohérence totale dans tout le schéma dimensionnel réalisé, chaque ligne dans la table des faits correspond à un événement mesurable, ce qui représente le grain (niveau de granularité).

La figure suivante illustre un exemple d'une table de fait :

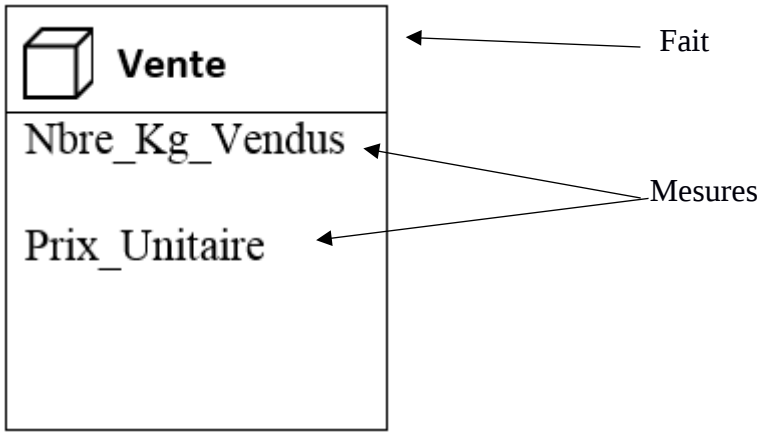


Figure 6: Exemple d'une table de fait

III.2.1.2 Concept Dimension

La dimension définit les différentes perspectives utilisées pour analyser le sujet c'est à dire le fait. Ces perspectives correspondent à une catégorie utilisée pour caractériser les mesures d'activité analysées.

Une dimension se compose de paramètres correspondant aux informations faisant varier les mesures de l'activité. On définit une hiérarchie de dimension comme le résultat de la décomposition d'une ou plusieurs dimensions en plusieurs niveaux [AR11]

Exemple d'une dimension :

La figure suivante montre la structure d'une table de dimension :

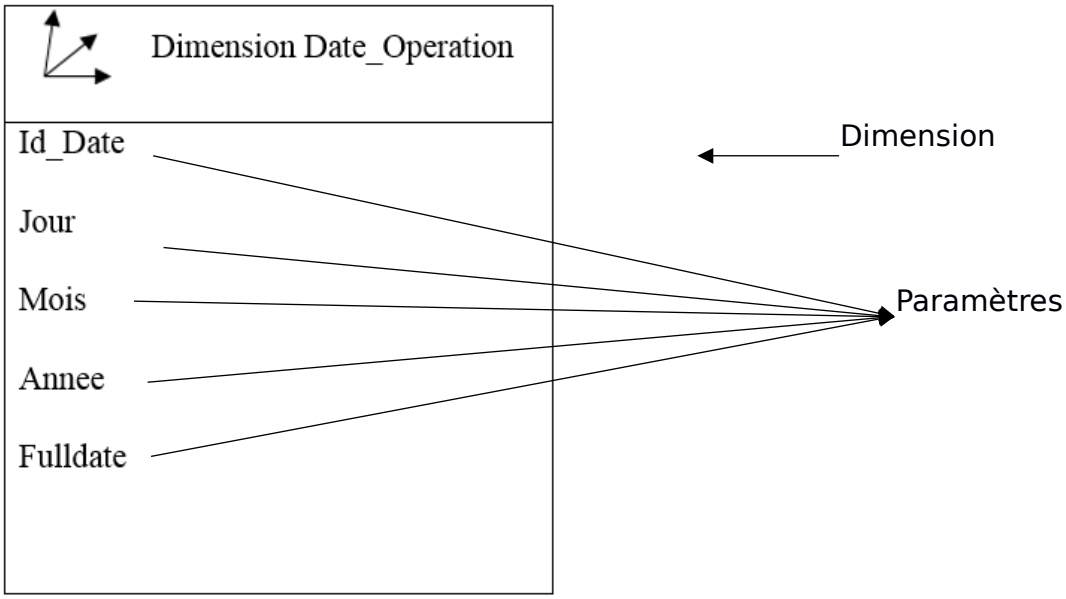


Figure 7: Exemple d'une table de dimension

III.2.1.3 Modèles structuraux des entrepôts de données

Il existe trois modèles structuraux sur lesquels on peut faire référence lors de la modélisation des entrepôts tels que le modèle en étoile, modèle en flacon et la constellation.

III.2.1.3.1 Modèle en étoile

Ce modèle se présente comme une étoile dont le centre n'est autre que la table des faits et les branches sont les tables de dimension. La force de ce type de modélisation est sa lisibilité et sa performance.

Il représente visuellement une étoile, on parle de modèle en étoile [AR10]

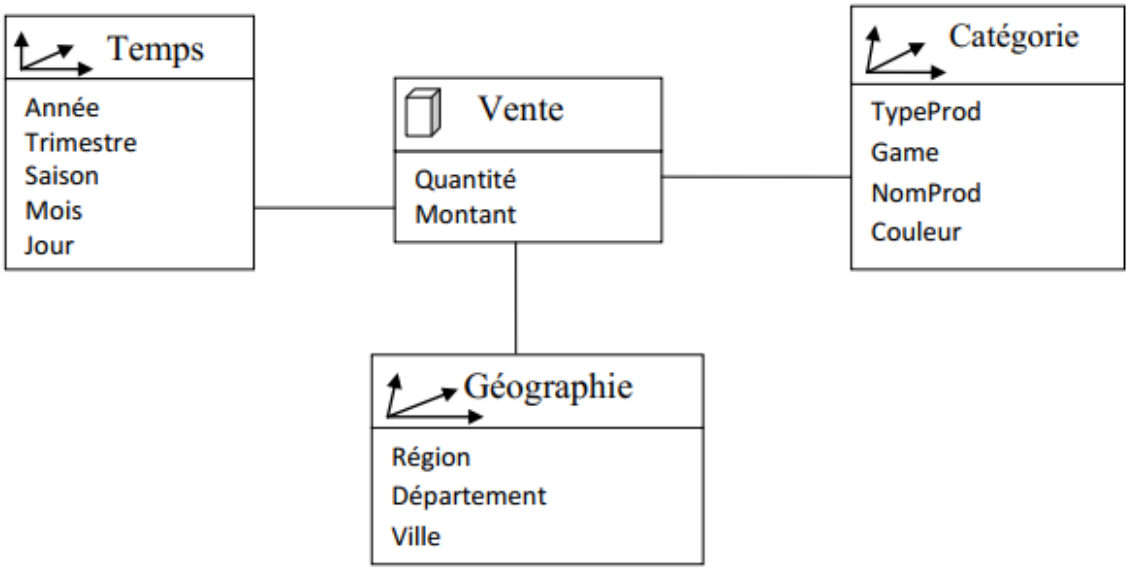


Figure 8: Exemple du modèle en étoile

III.2.1.3.2 Modèle en flocon de neige

Identique au modèle en étoile, sauf que ses branches sont éclatées en hiérarchies. Cette modélisation est généralement justifiée par l'économie d'espace de stockage, il consiste à décomposer les dimensions du modèle en étoile en sous hiérarchies. Ce modèle est donc une émanation du modèle en étoile puisque le fait est conservé et les dimensions sont éclatées conformément à sa hiérarchie des paramètres.

Cependant elle peut s'avérer moins compréhensible pour l'utilisateur final, et très couteuse en terme de performances.

La figure suivante illustre le modèle en flocon de neige :

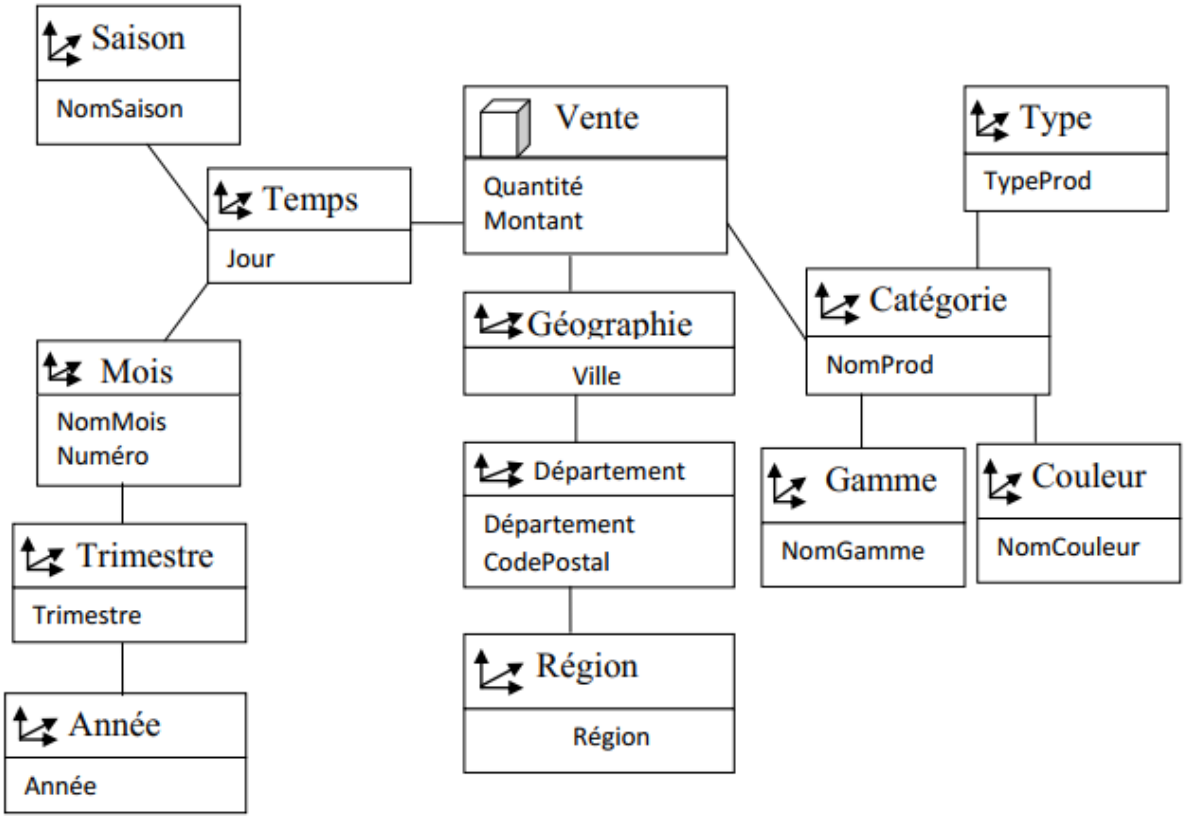


Figure 9: Exemple de modèle en flocon

Voici les différents avantages de deux modèles précédemment vus dans le tableau suivant :

Tableau 2: Avantages du Modèle en étoile et en Flocon [AR19]

	Etoile	**Flocon**
Avantages	- Lisibilité et performance des Requêtes	- Espace disque - Performances des mises à jouir

III.2.1.3.3 Modèle en constellation

Ce modèle est le résultat de la fusion entre plusieurs modèles en étoile qui utilisent des dimensions communes. Il comprend donc plusieurs faits et des dimensions communes ou non. [AR13]

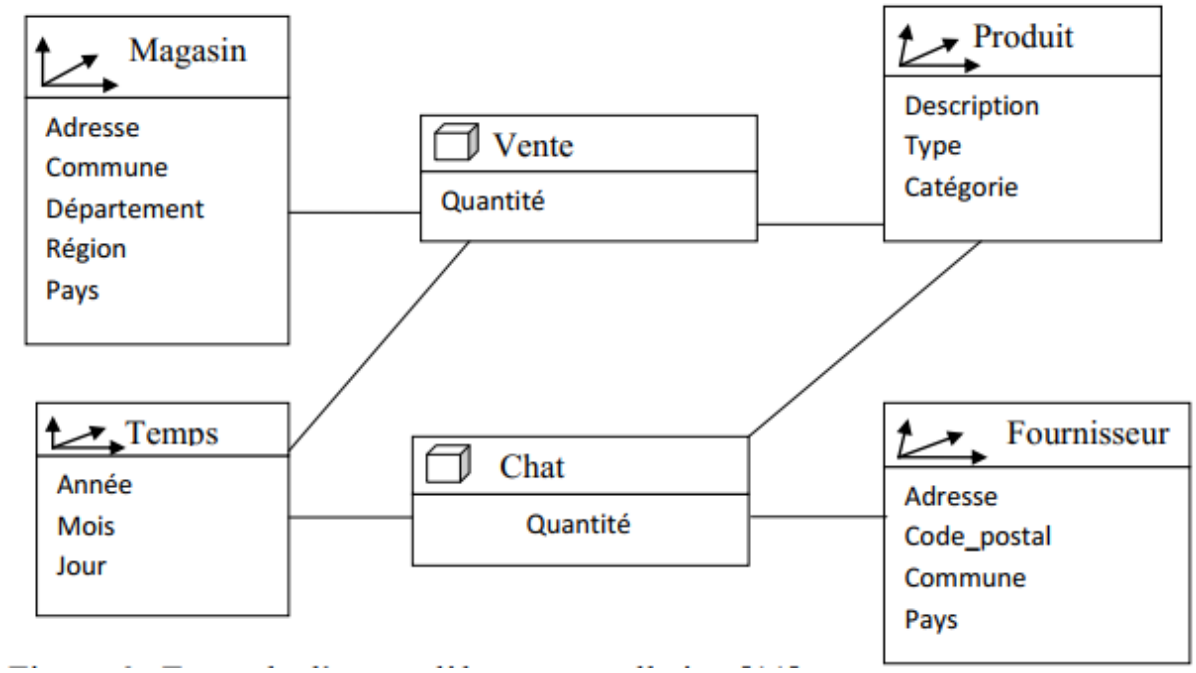


Figure 10: Exemple d'un modèle en constellation

III.3 Niveau logique

Il existe plusieurs possibilités de faire la modélisation multidimensionnelle. Les architectures les plus répandues et les plus adoptées par les fournisseurs de solutions OLAP sont ROLAP, MOLAP et HOLAP mais il y a d'autres systèmes qui sont basées sur des autres architectures comme architecture OOLAP.

III.3.1 ROLAP

Pour stocker le données, les systèmes ROLAP utilisent un SGBD relationnel. Chaque dimension est représentée par une table de dimension et chaque fait est représenté par une table de fait. Les mesures sont stockées dans les tables des faits [AR6].

III.3.2 MOLAP

Les systèmes MOLAP implémentent le cube (qui est une intersection de données dans un espace à plusieurs dimensions) sous forme d'un tableau multidimensionnel. Les données représentent chaque cellule sur le cube. C'est une solution de stockage qui garantit un temps de réponse rapide [AR14].

III.3.3 HOLAP

Les systèmes HOLAP sont des systèmes où les données fréquemment utilisées (données agrégées) sont maintenues par un SGBD multidimensionnel, et les données non fréquemment utilisées dans un SGBD relationnel [AR14].

Le tableau suivant décrit les avantages et les inconvénients de trois approches de l'OLAP vues précédemment.

Tableau 3: Avantages et inconvénients de ROLAP, MOLAP et HOLAP

	avantages	Inconvénients
ROLAP	-Standard, technologie prouvée et familière -Scalable (capacité d'expansion élevée) -Stockage de grande quantités de données	Temps de réponse élevé
MOLAP	-Modèle multidimensionnel (calcul automatique des agrégations des données) -Temps de réponse très court -Techniques d'indexation spécialisée	-Redondance des données -Non scalable -Couts élevés (les licences des bases multidimensionnelles et du développement) -Le besoin de définir des opérations pour manipuler les structures multidimensionnelles -Difficulté de la mise à jour et de la gestion du modèle
HOLAP	-Accès rapide pour différents niveaux d'agrégation -Stockage compact des agrégations	-Système complexe (serveurs HOLAP doit supporter des outils MOLAP et ROLAP) -Charge importante (entre

	-Temps de réponse assez court	les techniques d’optimisation ROLAP et les outils MOLAP)

III.3.4 OOLAP

Plus récemment, une autre approche s'appuie sur le paradigme objet ; on parle de l'approche OOLAP. Le modèle multidimensionnel se traduit ainsi :

« Chaque fait correspond à une classe, appelée classe de fait,

« Chaque dimension correspond à une classe, appelée classe de dimension.

III.4 Navigation hypercube

Une fois que le serveur OLAP a construit le cube multidimensionnel « ou simulé ce cube selon l’architecture du serveur », plusieurs opérations sont possibles sur ce dernier offrant ainsi la possibilité de naviguer dans les données qui le constituent. Ces opérations de navigation « Data Surfing » doivent être, d’une part, assez complexes pour adresser l’ensemble des données et, d’autre part, assez simples afin de permettre à l’utilisateur de circuler de manière libre et intuitive dans le modèle dimensionnel

III.4.1 Navigation dans les données

Une fois que le serveur OLAP a construit le cube multidimensionnel « ou simulé ce cube selon l'architecture du serveur », plusieurs opérations sont possibles sur ce dernier offrant ainsi la possibilité de naviguer dans les données qui le constituent. Ces opérations de navigation « Data Surfing » doivent être, d'une part, assez complexes pour adresser l'ensemble des données et, d'autre part, assez simples afin de permettre à l'utilisateur de circuler de manière libre et intuitive dans le modèle dimensionnel.

Afin de répondre à ces attentes, un ensemble de mécanismes est exploité, permettant une navigation par rapport à la dimension et par rapport à la granularité d'une dimension.

III.4.1.1 Slice & Dice

Le ***« Slicing »*** et le ***« Dicing »*** sont des techniques qui offrent la possibilité de faire des tranches « trancher » dans les données par rapport à des filtres de dimension bien précis, se classant de fait comme des opérations liées à la structure « se font sur les dimensions ». La différence entre eux se manifestent dans le fait que :

Le ***Slicing*** consiste à faire une sélection de tranches du cube selon des prédicats et selon une dimension *« filtrer une dimension selon une valeur »* .[AR14]

Le ***Dicing,*** quant à lui, peut être vu comme étant une extraction d'un sous cube.

III.4.1.2 Drill-down & Roll-up

Ces méthodes, appelées aussi « forage vers le bas/vers le haut », sont les méthodes les plus répandues pour une navigation dans un entrepôt de données. Elles consistent à représenter les données du cube à un niveau de granularité inférieur, dans le cas du « Drill - down », ou un niveau supérieur, c'est le « Roll-up ». En somme, ces deux opérations permettent de contrôler le niveau de détail des données du cube.

Ces opérations ne sont pas aussi faciles à implémenter car basées sur la notion d'une bonne hiérarchisation des attributs d'une dimension et la différenciation entre tous les niveaux de hiérarchie disponibles dans les différentes dimensions.

III. 4.1.3 Rapports

Le reporting est l'ensemble des comptes rendus permettant à une entreprise de suivre son activité et de s'évaluer grâce à la création périodique de rapports et de bilans analytiques de son activité. Ces rapports sont souvent destinés au manager ou au corps exécutif. [AR14]

III.5 Phases de l'alimentation « E.T.L. »

Les phases du processus E.T.L représentent la mécanique d'alimentation du DataWarehouse et comprennent les étapes suivantes :

- Extraction des données ;
- Transformation ;
- Chargement ;

Les outils ETL, en français ETC selon Kimball, sont des outils qui garantissent la faisabilité et facilitent le déroulement des trois phases citées précédemment d'où leur importance dans un projet Data Warehouse. [AR20]

III.6 Conclusion

Dans ce chapitre, nous avons pu définir l'entrepôt de données et décrire la structure des données qui le composent. Nous avons également montré les différents modèles utilisés pour modéliser l'entrepôt de données ainsi que les différents niveaux logiques et les opérations utilisées pour la navigation de données de l'entrepôt.

CHAPITRE IV : CONSTRUCTION D'UN ENTREPOT DE DONNEES

IV.1 Collecte et définition des besoins

IV. 1.1 Introduction

Pour concevoir l'entrepôt de données, on doit tenir compte des attentes des utilisateurs et pour cela, une collecte des données est nécessaire pour mieux commencer.

Pour notre travail, on a procédé par les étapes ci-dessous :

- Visite à l'OTB pour recueillir les informations nécessaires ;
- Faire des interviews avec les agents du département commercial ;
- Faire des interviews avec les agents du service informatique ;
- Consultation des rapports de ventes pour enrichir nos ressources de données en détectant d'autres besoin non identifies lors des interviews.
- Comme les agents du service informatique sont ceux qui maîtrisent plus l'état de chaque service au point de vu modélisation, nous devrions faire une interview sommaire pour savoir l'existant dans ce domaine.

IV.1.2 Contact avec les agents du service commercial

Lors de notre rencontre avec le service commercial, nous avons eu l'occasion d'identifier le processus de vente du thé depuis les usines de production jusqu'aux marchés locaux et étrangers. Alors pour qu'une vente soit réalisée complétement, il y a intervention de plusieurs services selon l'état d'avancement.

Après avoir été emballé dans les sacs ou les paquets selon la quantité voulue, il y a l'intervention du service de transport pour acheminer le the vers les différents stocks de l'OTB.

Jusqu'ici, on a constaté le besoin de bien analyser les entités Suivantes :

- Usines ;
- Transport ;
- Stock.

IV.1.3 Consultation des rapports et détection d'autres besoins

Ensuite on a eu le temps de consulter les rapports ce qui nous a permis de détecter d'autres besoins qui ne figurent pas dans les réponses aux interviews.

Cette étape nous permet de bien mener la construction de l'entrepôt des données entre autres la détermination des dimensions et les faits.

Le tableau suivant résume tous les besoins que nous avons recensés durant notre analyse.

Tableau 4: Le tableau récapitulatif des besoins

Catégorie	**Besoins constatés (critères d'analyse)**
Vente	-savoir la quantité du thé vendu selon les critères (Marché et Usine de production) - savoir l'évolution de ventes selon les critères (marché et temps) -Générer les rapports de vente selon le critères voulus (Temps et marché et usine de production)
Transport	-Savoir la place du the transport pour chaque société de transport selon le temps et la quantité transportée. - générer le rapport complet selon les critères voulus (temps, place et quantité)

Dépenses	-Savoir les couts des dépenses selon les critères (Types de dépense, temps et montant dépensé) - Générer le rapport complet selon les critères voulus (Frais, temps et stock)
Stock local	-Savoir la quantité du thé stocké selon les critères (temps, grade et Nombre de paque ou sacs) - Générer le rapport complet selon les critères choisis (Usine, temps, Nombre de paquet)
Stock Export	-Savoir la quantité du thé stocké selon les critères (temps, grade et Nombre de paque ou sacs) - Générer le rapport complet selon les critères choisis (Usine, temps, Nombre de paquet)

IV.2 Modélisation multidimensionnelle

IV.2.1Introduction

Pour bien modéliser notre entrepôt, nous préférons regrouper les besoins dans des activités telles que nous les avons identifiés dans le chapitre précédent. De ce qui suit, on précisera l'activité, la granularité, les dimensions, le fait ainsi que le modèle multidimensionnel.

IV.2.2 Présentation de l’activité de vente du thé sec « Vente »

« Une vente est la cession d’un bien ou d’un service en échange d’une somme d’argent convenue entre le vendeur, celui qui cède le bien ou le service, et l’acheteur, celui qui paie. » **[Larousse,2008]**

OTB, Par le biais de ses cinq usines, propose la vente des produits théicoles (the sec et the vert), vendus d’une part sur le territoire national et d’autre part livrés sur les marchés internationaux (Mombasa, Dar es lam, ...)

La vente du thé demeure comme l’activité principale de l’office du the du Burundi, réalisant la plus grande partir du chiffre d’affaire de l’office. Les chiffres lies aux ventes se présentent comme des indicateurs d’une grande signification par rapport à la performance de l’office. Ainsi la disponibilité de ces informations s’avère indispensable pour les décideurs de l’entreprise

L’OTB effectue alors cette opération de vente via son service commercial qui se charge de suivre tout le mouvement du the entre l’office et sa clientèle.

On pourra donc obtenir alors toutes les information lies à l’évolution des ventes dans les différents points de ventes et marches.

La granularité ou le niveau de détail le plus bas de cette activité correspond à une ligne de table de fait :« Le nombre de Kg de the vendu sur un marché à une date donnée ».

IV.2.2.1 Dimensions participantes du modèle

Les dimensions ont pour objectif de décrire le fait, donc on essaye de recenser toutes les informations qui décrivent une vente et qui peuvent intéresser les décideurs

- **Dimension Temps** :« La dimension temps est la seule dimension qui figure systématiquement dans tout l’entrepôt de données, car en pratique tout entrepôt de données est une série temporelle. Le temps est le plus souvent la première dimension dans le classement sous-jacent de la base de données » [AR13].

La dimension temps se présente comme suit :



Figure 11: Dimension Temps

Pour cette dimension, le niveau de détail le plus bas est la journée tandis que le niveau détail plus haut est l'année.

Tableau 5: Tableau descriptif de la dimension temps

Désignation	**Détails**
Id_Date	La clé primaire de la dimension
Jour	Jour du mois de l' année
Mois	Mois de l'année
Annee	L' année
Fulldate	La date du jour et heure de l'opération

- **Dimension Type_vente** : cette dimension permet de distinguer quel type de vente a été effectuée en se basant sur les critères « comment et où la vente a été réalisée ? »

La dimension Type_Vente se présente comme suit :

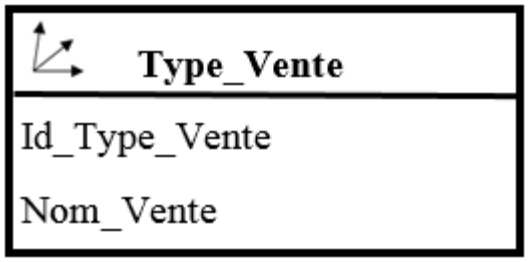


Figure 12: Dimension Type_Vente

Tableau 6: Tableau descriptif de la dimension Type_Vente

Désignation	**Détails**
Id_Type_Vente	Clé primaire de la dimension
Nom_Vente	Nom de type de vente effectuée

- **Dimension Type_Monnaie** : Lors de l'activité de « Vente », le client doit payer une certaine somme en échange contre le produit qu'il a choisi. Cette somme est exprimée soit en monnaie locale, soit en monnaie étrangère d'où l'intervention de cette dimension pour déterminer en type de monnaie équivaut la valeur du produit vendu.

La dimension Type_monnaie se présente comme suit :

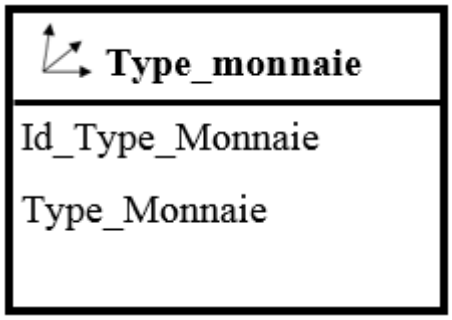


Figure 13: Dimension Type_Monnaie

Tableau 7: Tableau descriptif de la dimension Type_Monnaie

Désignation	**Détails**
Id_Type_Monnaie	La clé primaire de la dimension
Type_Monnaie	Type de la monnaie

- **Dimension Quantite_The** : Même si on trouve le mot quantité dans le nom de cette dimension, elle est là pour exprimer la qualité et la valeur du the selon les différents grades attribue à ce produit a la sortie de l'usine. Donc à l'aide de cette dimension, on pourra avoir des informations de plus sur l'activité « Vente ».

La dimension Quantite_The se présente comme suit :

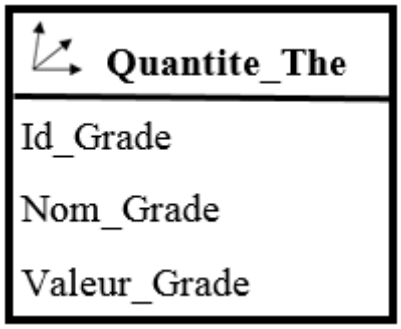


Figure 14: Dimension Quantite_The

Voici-ci dessous la description de la dimension Quantite_The :

Tableau 8: Tableau descriptif de la dimension Quantite_the

Désignation	**Détails**
Id_Quantite_the	Clé primaire de la dimension
Nom_Grade	Nom de la grade du thé
Valeur_Grade	Valeur de la grade du thé

- Dimension Usine : Dans les activités quotidiennes de l'OTB, l'entité usine joue un rôle primordial puisque toutes les opérations liées à la transformation ainsi qu'au paquetage du thé, se passent dans une usine. Notons aussi que cette assure le stockage the en attendant le transfert

vers l’entrepôt principal ou vers les marchés étrangers. D’où cette dimension a un rôle particulier dans l’activité « Vente ».

La dimension Usine se présente de la manière suivante :

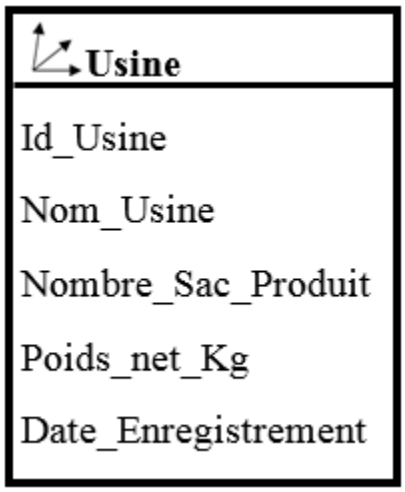


Figure 15: Dimension Usine

Voici-ci dessous la description de la dimension Usine :

Tableau 9: Tableau descriptif e la dimension usine

Désignation	**Détails**
Id_usine	Clé primaire de la dimension
Nom_Usine	Nom de l’Usine
Nombre_Sac_Produit	Nombre de sac de thé produit dans cette Usine
Poids_net_Kg	Poids net du thé produit dans cette usine
Date Enregistrement	Date de production du thé enregistrée

Le fait de l’activité « Vente » nous permet d’analyser les Ventes effectues .il comprend la mesure « Nombre de Kg vendu » qui permet de mesurer l’évolution de cette activité

Le fait Vente se présente comme suit :

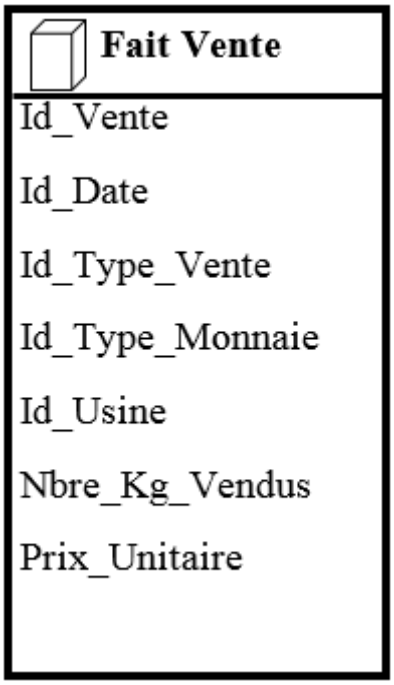


Figure 16: Fait « Vente »

Voici ci-dessous la description du fait Vente :

Tableau 10: Tableau descriptif du fait Vente

Désignation	**Détails**
Id_Vente	Clé primaire du fait
Id_Date	Attribut représentant la dimension Date_Operation
Id_Type_Vente	Attribut représentant la dimension Type_Vente
Id_Type_Monnaie	Attribut représentant la dimension Type_monnaie
Id Usine	Attribut représentant la dimension Usine
Nbre_Kg_Vendus	La mesure du fait qui le nombre de Kg de thé Vendu pour une usine et sur un marche
Prix_Unitaire	La mesure du fait qui est prix unitaire du thé vendu

Voici le modelé en étoile de l’activité Vente :

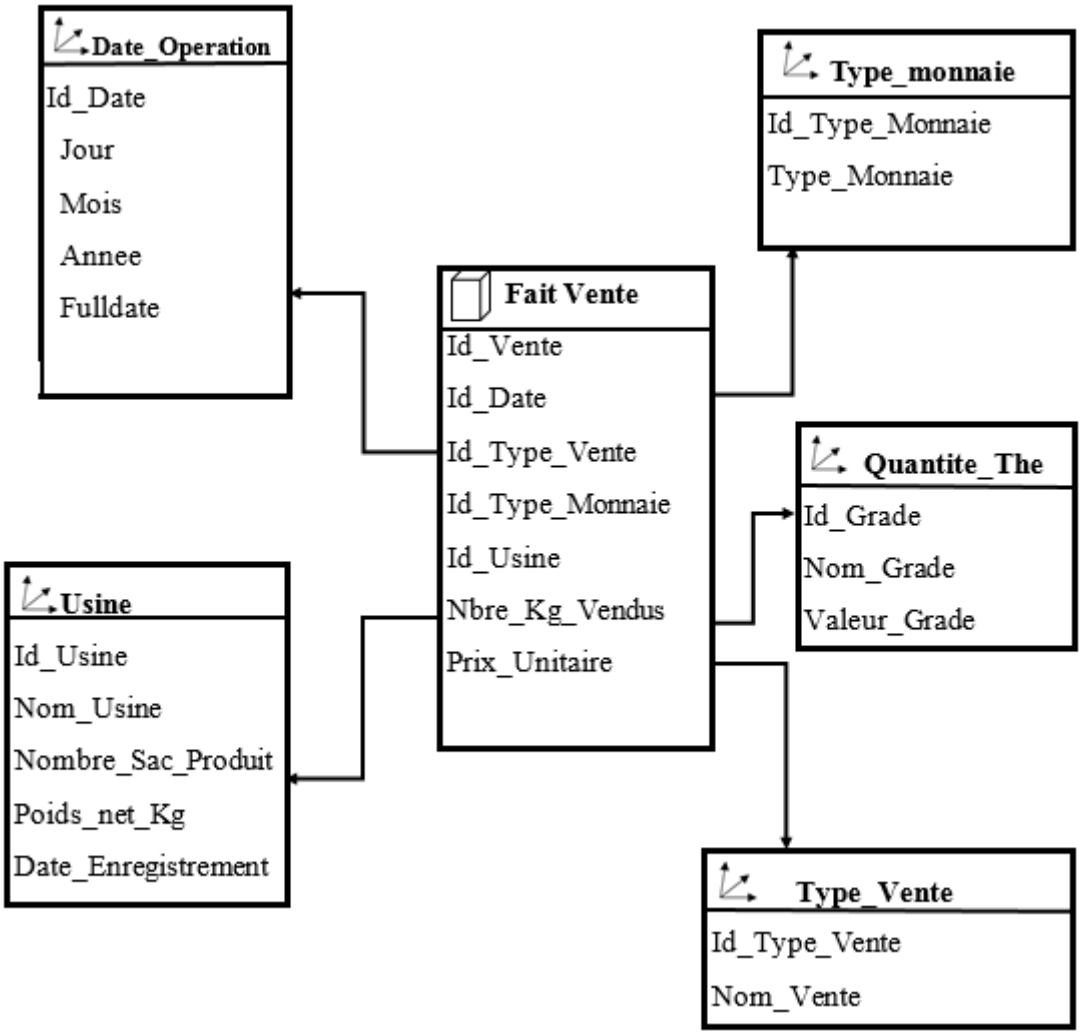


Figure 17: Modèle en étoile de l'activité de vente « Vente »

VI.2.3 Présentation de l'activité "Transport"

Après avoir transformé et emballé les produits finis, c'est le moment d'acheminer ces derniers dans le stock se trouvant au siège de l'OTB à Bujumbura ou bien on les envoie directement vers les marchés de l'étranger pour la vente aux enchères. Cette activité est d'une grande importance pour l'entreprise car elle garantit une livraison en bon ou en mauvais état, et cela a un impact non négligeable sur la clientèle.

L'OTB effectue alors cette activité de transport via son service interne (charroi) tout en collaborant avec d'autres sociétés privées de transport surtout pour acheminer ses produits à l'étranger.

En analysant cette activité, on pourra donc obtenir alors toutes les informations liées à l'acheminement du the depuis les différentes usines jusqu'aux points de ventes

La granularité ou le niveau de détail le plus bas de cette activité correspond à une ligne de table de fait :« Le nombre de Kg de the acheminés par un transporteur à une date donnée ».

Parmi les dimensions que nous allons utiliser dans ce modèle, il y a une qui est utilisée dans l'activité précédente. Ici nous allons ajouter une autre dimension Transporteur qui nous permet d'identifier les détails concernant chaque transporteur.

La dimension Transporteur se présente comme suit :

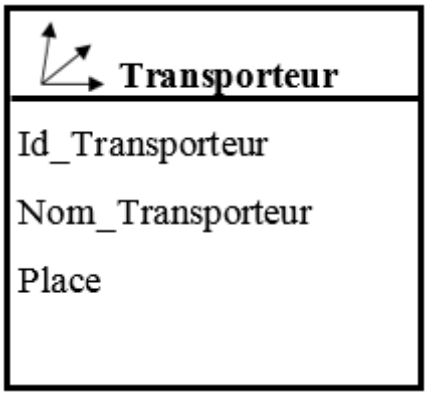


Figure 18: Dimension Transporteur

Tableau 11: Tableau descriptif de la dimension Transporteur

Désignation	Détails
Id_Transporteur	La clé primaire de la dimension
Nom_Transporteur	Le nom du transporteur
Place	La place occupée par le transporteur

Tableau 12: Les dimensions communes entre l'activité « Transport » et « Vente »

Dimension	Transporteur	Vente
Date_Operation	✓	✓
Transporteur	✓	✗
Type_vente	✗	✓
Quantite_The	✗	✓
Type_monnaie	✗	✓
Usine	✗	✓

Cette activité nous permet d'analyser la quantité de thé transportée. Il contient la mesure « Nbre_Kg_Transportes » qui permet d'identifier aussi le meilleur transporteur.

Ce fait est structuré de la manière suivante :

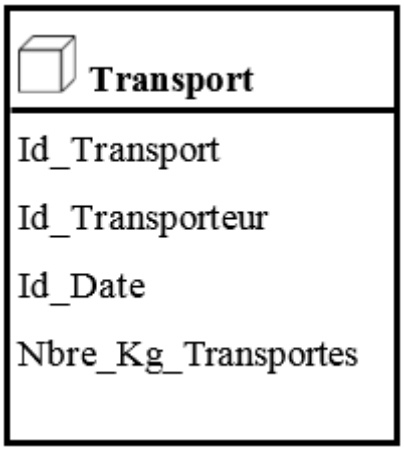


Figure 19: Fait « Transport »

Tableau 13: Tableau descriptif du fait « Transport »

Désignation	Détails

Id_Transport	La clé primaire du fait
Id_Transporteur	Attribut représentant la dimension Transporteur
Id_Date	Attribut représentant la dimension Date_operation
Nbre_Kg_Transportes	La mesure du fait qui le nombre de Kg de the transporté.

Voici le modèle en étoile de cette activité :

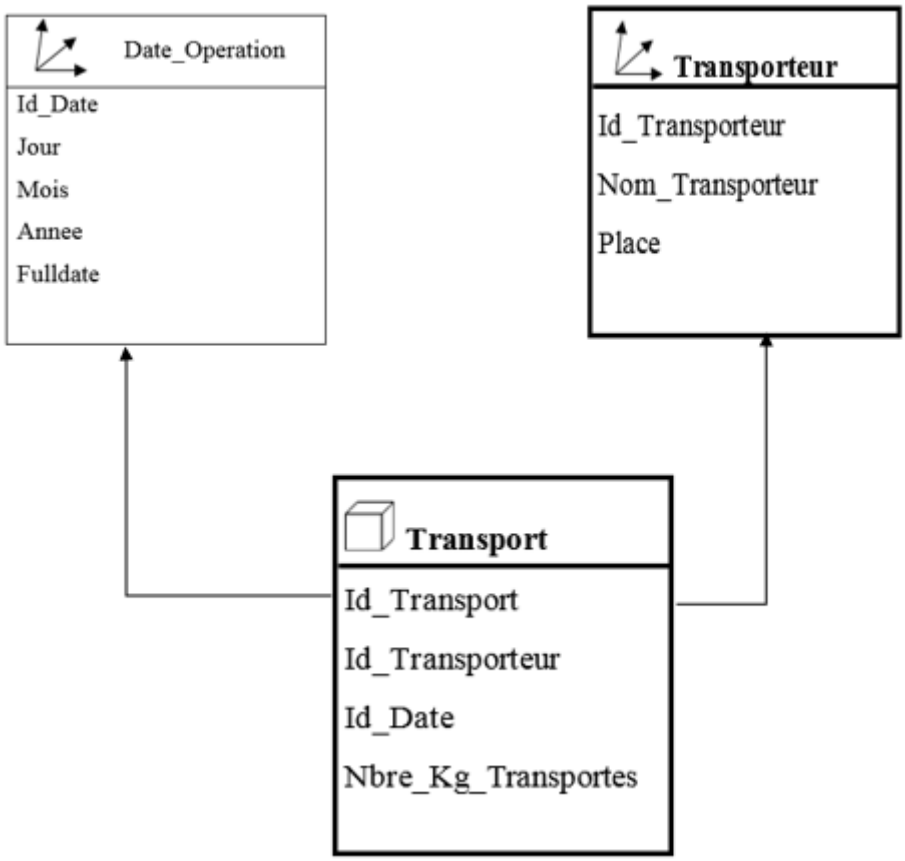


Figure 20: Modèle en étoile de l'activite de transport « Transport »

IV.2.4 Presentation de l'activite Dépenses

Dans notre cas, les dépenses peuvent etre définies comme une activité qui consiste à couvrir tous les frais nécessaires pour que les autres activités liées au transport et au vente des produits de l'OTB puissent être réalisées sans difficultés et sans retard.cette

activité nous permet alors de pouvoir analyser les frais octroyés durant tout le processus de transport et de vente du thé sur les marches locaux et internationaux.

La granularité ou le niveau de détail le plus bas de cette activité correspond à une ligne de table de fait :« calculer le Montant dépensé pour chaque stock à une date donnée ».

Parmi les dimensions que nous allons utiliser dans ce modèle, il y a une qui est utilisée dans les activité précédente. Ici nous allons ajouter deux autres dimensions Type_stock et Details_frais qui nous permettront d'identifier les détails concernant les stocks du thé et les types de frais.

La dimension Type_Stock se présente de la manière suivante :

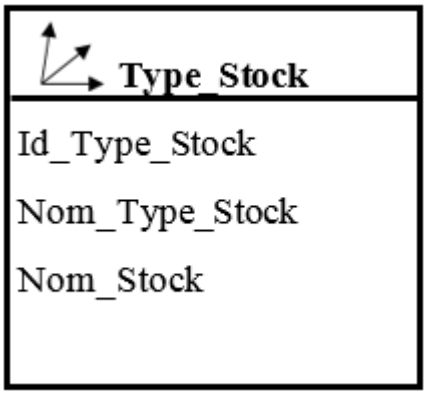


Figure 21: Dimension Type_Stock

Tableau 14: Tableau descriptif de la dimension Type_Stock

Désignation	Détails
Id_Type_Stock	La clé primaire de la dimension

Nom_Type_Stock	Le type de stock
Nom_Stock	Le nom du stock

La dimension Detail_Frais se présente comme suit :

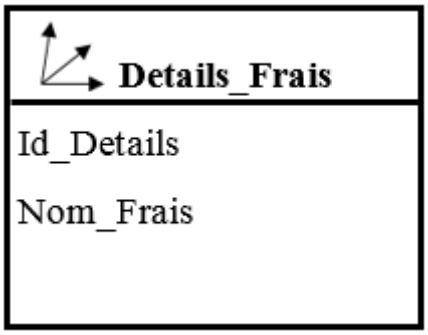


Figure 22: Dimension Details_Frais

Tableau 15: Tableau descriptif de la dimension Details_Frais

Deesignation	Details
Id_Details	La cle primaire de la dimension
Nom_Frais	Descritpion des frais

Tableau 16: Dimensions communes ou non entre l'activité « Vente » et « Depenses »

Dimension	vente	Dépenses
Date_Operation	✓	✓
Type_Vente	✓	✓

Type_monnaie	✓	✗
Usine	✓	✗
Quantite_The	✓	✗
Type_Stock	✗	✓
Detail_Frais	✗	✓

La structure de ce fait et la suivante :

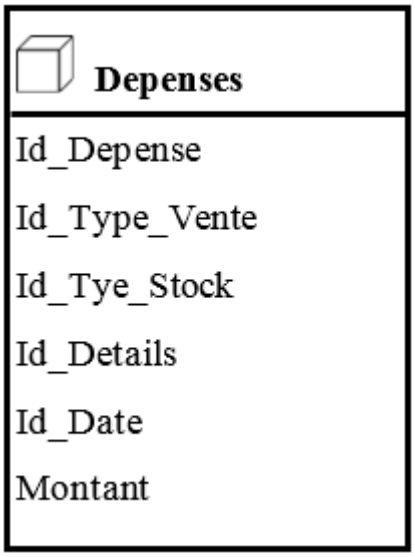


Figure 23: Fait « Depenses »

Tableau 17: Tableau descriptif du fait « Depenses »

Désignation	Détails
Id_Depenses	La clé primaire du fait
Id_Type_Vente	Attribut représentant la dimension Type_Vente
Id_Type_Stock	Attribut représentant la di,ension Type_Stock
Id_Details	Attribut representant la dimension Details_Frais

Id_Date	Attribut representant la dimension Date_Operation
Montant	La mesure du fait qui est le montant total des fais depensés

Le modele en Etoile de l'activite de « Depense »

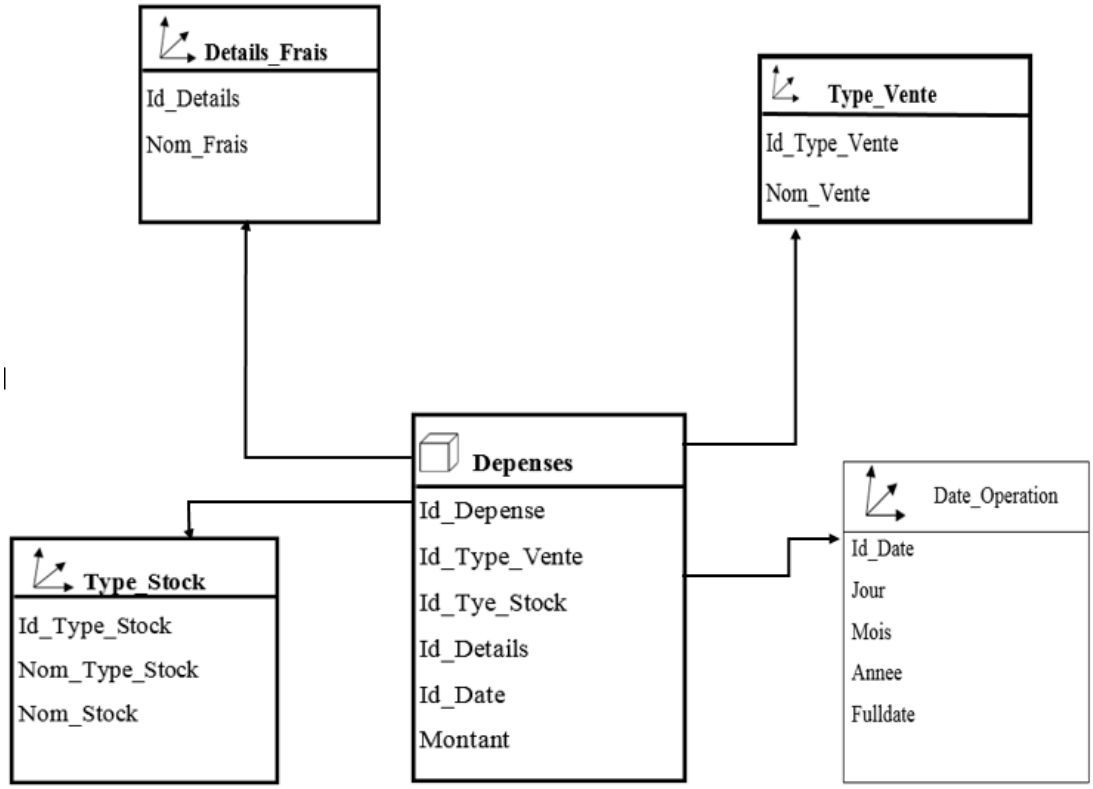


Figure 24: Modèle en étoile de l'activité « Depense »

IV.2.5 Présentation de l'activité de stockage « Stock_Export »

Toutes les entreprises de production doivent avoir un lieu dans lequel elles conservent leurs produits fabriqués en attente d'être vendus. C'est le cas également pour l'OTB. Après la phase de transformation de feuilles vertes en thé sec (Produits finis), l'étape suivante consiste au stockage ces derniers dans un lieu approprié. C'est le stock.

Dans notre projet, nous allons plus nous pencher sur le stock destiné à l'exportation vers les points de ventes étrangers. Cette activité de stockage sera illustrée dans la suite par le fait « Stock_Export ». Ce dernier nous permet d'analyser la quantité de thé par chaque stock à une date donnée.

La granularité ou le niveau de détail le plus bas de cette activité correspond à une ligne de table de fait :« calculer la quantité de thé par chaque stock pour l'export à une date donnée ».

Toutes les dimensions que nous allons utiliser dans ce modèle sont utilisées dans les activités précédentes.

Tableau 18: Dimensions communes ou non entre l'activité « Vente » , « Depenses » et « Stock_Export »

Dimension	vente	Dépenses	Stock_Export
Date_Operation	✓	✓	✓
Type_Vente	✓	✓	✗
Type_monnaie	✓	✗	✗
Usine	✓	✗	✗
Quantite_The	✓	✗	✓
Type_Stock	✗	✓	✓
Detail_Frais	✗	✓	✗

Le fait « Stock_Export » nous permet d'analyser la quantité de the dans chaque stock pour l'export.

La structure de ce fait est la suivante :

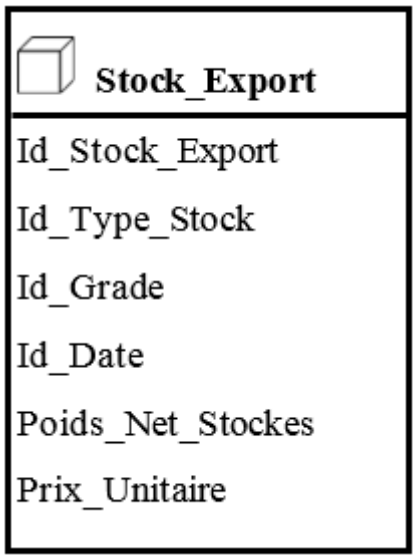


Figure 25: Fait « Stock_Export »

Tableau 19: Tableau descriptif du fait « Stock_Export »

Désignation	Détails
Id_Stock_Export	La clé primaire du fait
Id_Type_Stock	Attribut représentant la dimension Type_Stock
Id_Grade	Attribut représentant la dimension Quantite_The
Id_Date	Attribut représentant la dimension Date_Operation
Poids_Net_Stockes	La mesure de fait qui est la quantité de the stockée pour l’export
Prix_Unitaire	Attribut du fait.

Voici le modèle en étoile du fait « Stock_Export »

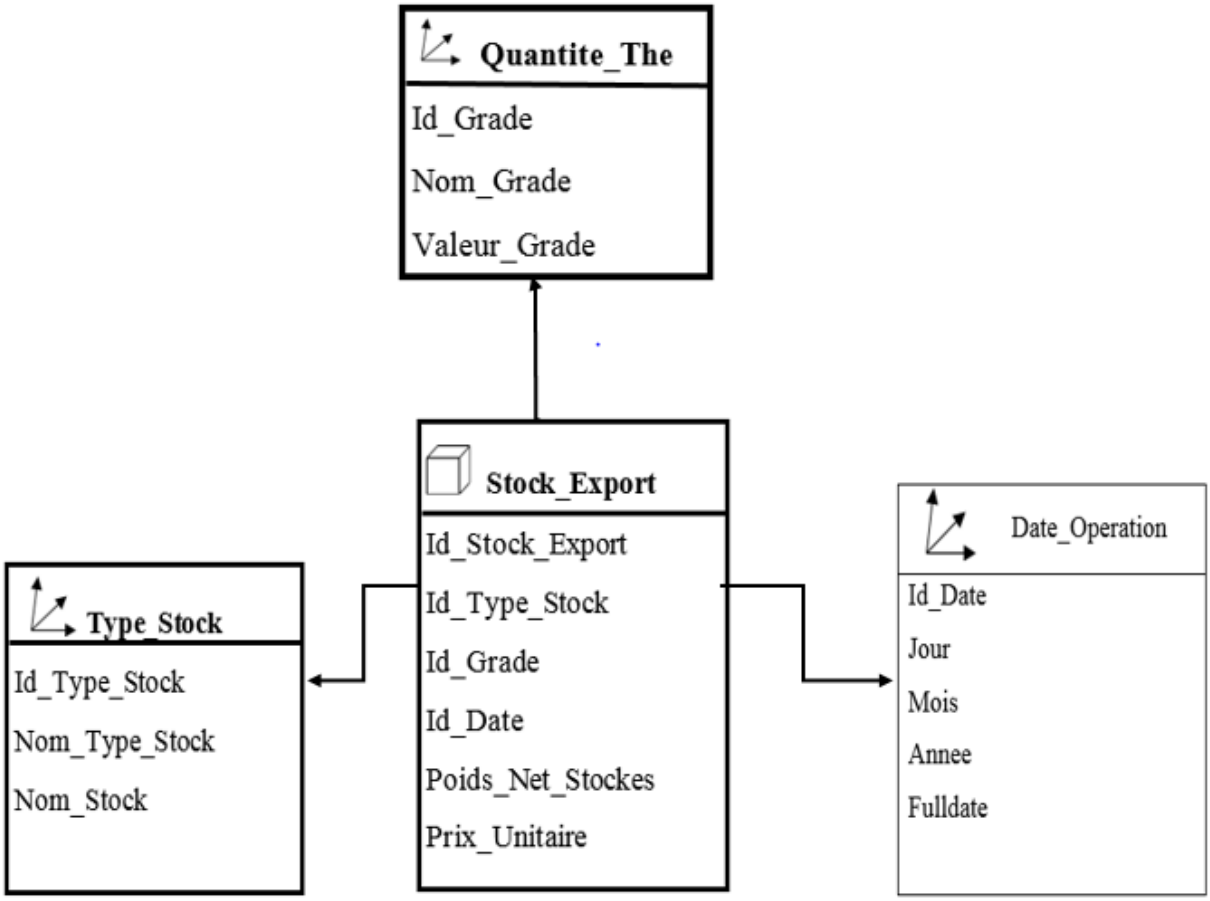


Figure 26: Modèle en étoile de l'activité « Stock_Export »

IV.2.6 Présentation de l'activité « Stock_Local »

Cette activité a la même signification que celle traitée au point précèdent sauf que celle-ci concerne le stockage des produits qui seront vendus sur les marches locaux. Elle va nous permettre alors d'analyser quantité de the stockée et destinée à être vendue localement.

La granularité ou le niveau de détail le plus bas de cette activité correspond à une ligne de table de fait :« calculer la quantité de thé stockée par chaque usine pour le marché local à une date donnée ».

Parmi les dimensions que nous allons utiliser dans ce modèle, il y a celles qui sont déjà utilisées dans les activités précédentes. Ici nous allons ajouter deux autres dimensions qui nous permettront d'identifier les détails concernant les stocks du thé.

La dimension « Paquet » se présente comme suit :

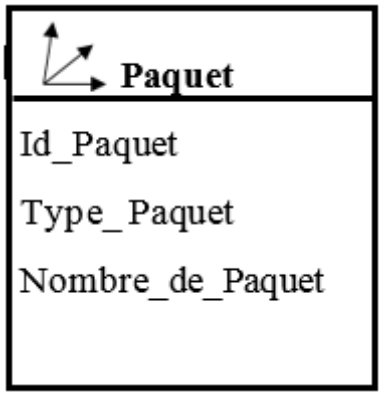


Figure 27: Dimension Paquet

Tableau 20: Tableau descriptif de la dimension « Paquet »

Désignation	Détails
Id_Paquet	La clé de la dimension
Type_Paquet	Description du paquet de thé
Nombre_de_Paquet	Le nombre de paquet de thé

Tableau 21: Dimensions communes ou non entre l'activité « Vente » et « Stock_Local»

Dimension	vente	Stock_Local

Date_Operation	✓	✓
Type_Vente	✓	✗
Type_monnaie	✓	✗
Usine	✓	✓
Quantite_The	✓	✓
Type_Stock	✗	✓
Paquet	✗	✓

Le fait « Stock_Local » nous permet d'analyser la quantité de thé dans chaque stock pour le marché local.

Le fait Stock_Export se présente de la manière suivante :

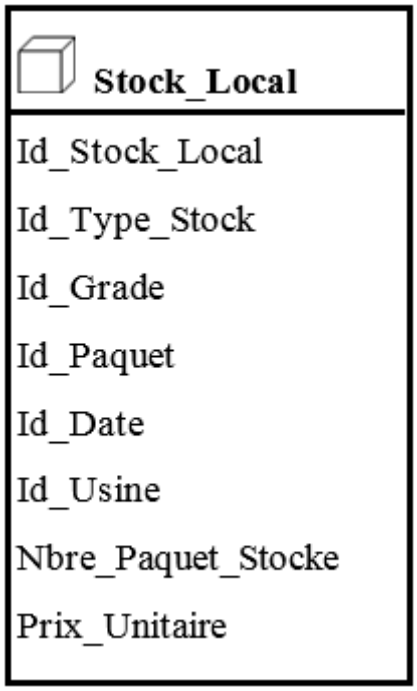


Figure 28: Fait « Stock_Local »

Tableau 22: Tableau descriptif du fait « Stock_Local »

Désignation	Détails
Id_Stock_Local	La clé primaire du fait
Id_Type_Stock	Attribut représentant la dimension Type_Stock

Id_paquet	Attribut représentant la dimension Paquet
Id_Date	Attribut représentant la dimension Date_Operation
Id_Usine	Attribut représentant la dimension Usine
Id_Grade	Attribut représentant la dimension Quantite_The
Nbre_Paquets_stocke	La mesure permettant de calculer la quantité stockée
Prix_Unitaire	La mesure désignant le prix unitaire par paquet

Voici le modèle en étoile du fait « Stock_Local » :

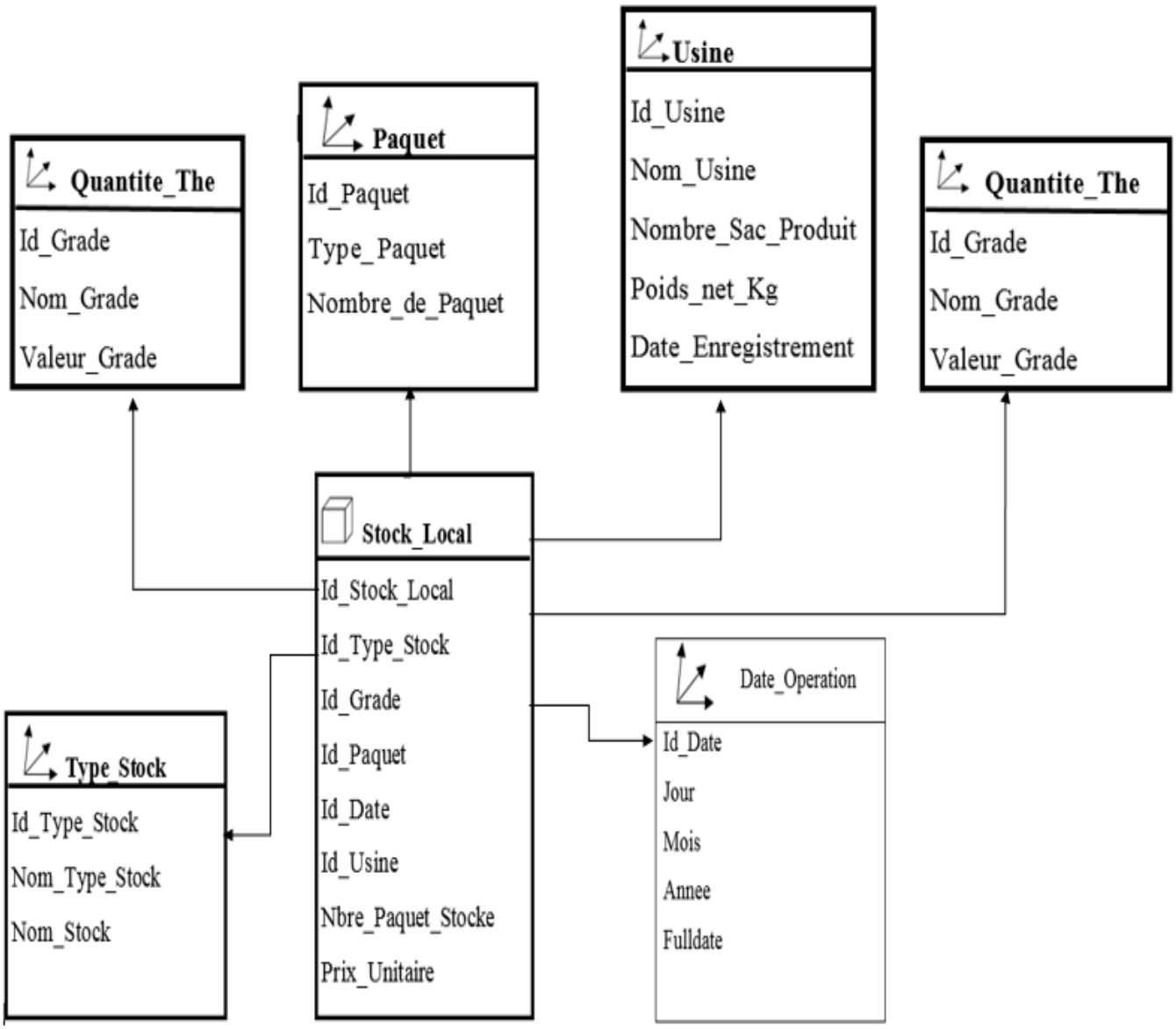


Figure 29: Modèle en étoile de l'activité « Stock_Local »

IV.3. Phases d’alimentation “ETL”

IV.3.1 Introduction

L'alimentation des données à partir des bases de production est une phase primordiale d'un datawarehouse. Des outils logiciels sont alors nécessaires pour intégrer les données dans le datawharehouse. On parle d'ELT (Extract, Transform, Load). Les phases de l'alimentation d'un datawarehouse sont les suivantes :

- Découverte des données
- Extraction des données
- Chargement des données
- Transformation des données

IV.3.2 Découverte des données

Il s'agit d'identifier dans les systèmes sources les données à importer dans le datawarehouse. Il faut prendre les données les plus judicieuses. Un mauvais choix peut considérablement compliquer les phases suivantes de l'alimentation.

IV.3.3 Extraction des données

Il s'agit de collecter les données utiles dans les systèmes de production. Il faut identifier les données ayant été modifiées afin d'importer le minimum de données dans le datawarehouse. Le but de cette étape est comme son nom l’indique la lecture et l'extraction des données du système provenant des différentes sources. Dans notre cas, les données proviennent d’une seule source, la base de donnée. La figure suivante montre la source de données pour notre entrepôt

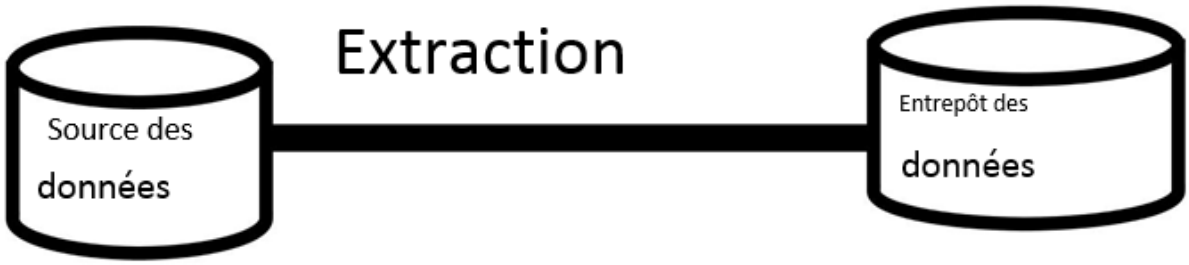


Figure 30: Structure d’extraction des données de notre entrepôt des données

IV.3.4 Transformation des données

Il faut rendre les données cohérentes avec la structure de l'entrepôt de données. On d'applique alors des filtres sur les données. Il peut être nécessaire de convertir le format des données (EBCDIC vers ASCII par exemple) ou d'harmoniser les formats de dates (jj/mm/aa). Il faut également associer les champs source avec les champs cibles. Un champ source « adresse » pourra ainsi par exemple être décomposé en « numéro », « rue », « code postal », « ville » ou l'inverse. Enfin des données des systèmes de production doivent être agrégées ou calculées avant leur chargement.

IV.3.5 Chargement des données

C'est la dernière phase de l'alimentation d'un entrepôt de données. Il s'agit d'insérer les données au sein de l'entrepôt. C'est une phase délicate car les quantités de données sont souvent très importantes.

CHAPITRE V. MODELE MATHEMATIQUE DU SYSTEME

V.1. Introduction

La notion de modèle en mathématiques permet de modéliser, c'est-à-dire de représenter, toutes sortes de situations, d'objets et de structures du monde réel. Il existe plusieurs outils mathématiques qui sont utilisés pour modéliser

le comportement du système informatique tels que l’algèbre relationnelle, théorie des graphes, réseau des neurones, file d’attente, réseaux de Pétri… [AR18].

Tableau 23: Μοδ⎛λεσ ματη⎡ματιθυεσ αχτυελλεμεντ υτιλισ⎡σ πουρ μοδ⎡λισερ λε σψστ⎛με ινφορματιθυε ετ λευρσ ρ⎢λεσ

Le nom modèle	**Rôle du modèle**
Réseau Pétri	Il sert à modéliser le comportement du système dynamique
Files d’attente	Elle consiste à modéliser le système dans lequel les serveurs sont soumis à un flux de requêtes qu’ils doivent traiter si les clients arrivent aléatoirement
Algèbre relationnelle	Elle décrit un ensemble d’opérations qui permettent de manipuler des relations considérées comme des ensembles de n-uplets
Théorie des graphes	Elle permet de décrire un ensemble d'objets et leurs relations, c'est à dire les liens entre les objets
Réseau des neurones	Un réseau neuronal s'inspire du fonctionnement des neurones biologiques et prend le corps dans un ordinateur sous forme d'un algorithme

Suite au tableau précédent, nous choisissons l’Algèbre relationnelle comme outil mathématique pour modéliser notre système car son rôle correspond correctement au comportement du modèle en étoile.

V.2 Généralités sur l’algèbre relationnelle

L’algèbre relationnelle est le support mathématique cohérent sur lequel repose le modèle relationnel. L’algèbre relationnelle propose un ensemble d’opérations élémentaires formelles sur les relations dans le but de créer de nouvelles relations. Ces permettent de représenter des requêtes sur la base de données dont le résultat s’exprime sous la forme d’une relation (i.e. table). C’est ce formalisme qui est au cœur du langage de requête de SQL. [*AR*21]

Nous pouvons distinguer trois familles d’opérateurs relationnels :

- **Les opérateurs unaires** (la sélection et la projection), qui sont les plus simples, permettent de produire une nouvelle table à partir d'une autre table.

- **Les opérateurs binaires ensemblistes** (l'union, l'intersection et la différence) permettent de produire une nouvelle relation à partir de deux relations de même degré et de même domaine.

- **Les opérateurs binaires ou n-aires** (le produit cartésien, la jointure et la division) permettent de produire une nouvelle table à partir de deux ou plusieurs autres tables.

Remarque : Les notations de l'algèbre relationnelle ne sont pas standardisées.

V.3 Sélection

Définition : La sélection génère une relation regroupant exclusivement toutes les occurrences de la relation R qui satisfont l'expression logique E.[AR21]

Notation : $\sigma_{(E)}\mathbf{R}$.

Il s'agit d'une opération unaire essentielle dont la signature est :

$relation \times expression\ logique \rightarrow relation$

La sélection permet ainsi de choisir (i.e. sélectionner) certaines lignes dans une table. Le résultat de la sélection est donc une nouvelle relation qui a les mêmes attributs que R. Si R est vide (c'est-à-dire sans aucune occurrence), la relation qui résulte de la sélection est vide.

Soit la relation Usine suivante :

Usine				
Id_Usine	Nom_Usine	Nombre_Sac_Produit	Poids_net_Kg	Date _Enregistrement
1	IJENDA	5000	50000	2016-04-22
4	BUHORO	6000	50000	2016-04-27
9	TORA	125	30000	2016-05-03
10	RWEGURA	250	12500	2016-05-08
11	MABAYI	260	14000	2016-05-09
12	TEZA	300	15000	2016-05-15

Ci-dessous un exemple de sélection sur la table Usine :

$\sigma_{(Id_Usine \geq 9)}$Usine : sélectionner toutes les occurrences de la relation Usine remplissant la condition **« Id_Usine≥9 »**

Usine				
Id_Usine	Nom_Usine	Nombre_Sac_Produit	Poids_net_Kg	Date _Enregistrement
9	TORA	125	30000	2016-05-03
10	RWEGURA	250	12500	2016-05-08
11	MABAYI	260	14000	2016-05-09
12	TEZA	300	15000	2016-05-15

Figure 31: Exemple de l'opération de sélection

La requête SQL correspondante est la suivante :

*SELECT * FROM Usine WHERE id_Usine >= 9;*

V.4. Projection

Définition : La *projection* consiste à supprimer les attributs autres que A1, A2, ..., An d'une relation et à éliminer les n-uplets en double apparaissant dans la nouvelle version. [AR21]

Notation : $\prod_{(A1, A2,..,An)}$**R**

Il s'agit d'une opération unaire essentielle dont la signature est :
relation × liste d'attributs →relation

En d'autres termes, la projection permet de choisir des colonnes dans une table. Si R est vide, la relation qui résulte de la projection est vide, mais pas forcément équivalente étant donné qu'elle contient généralement moins d'attributs.

Exemple de projection sur la table Usine définie au point VI.3 :

Usine
Nom_Usine
IJENDA
BUHORO
TORA
RWEGURA
MABAYI
TEZA

Figure 32: Résultat de l'opération de projection sur la table Usine

Voici la requête SQL correspondant à l'exemple précédent :

SELECT Nom_Usine FROM Usine;

La même requête en algèbre relationnelle :

$\Pi_{(Nom_Usine)}$ ***Usine***

V.5. Operateur ensembliste « Union »

Définition : l'*union* est une opération portant sur deux relations R1 et R2 ayant le même schéma et construisant une troisième relation constituée des n-uplets appartenant à l'une ou l'autre des deux relations R1 et R2 sans doublon. *[AR21]*

Notation : R1 ∪ R2.

Il s'agit d'une opération binaire ensembliste commutative essentielle dont la signature est :
relation × relation → relation

Comme nous l'avons déjà dit, R1 et R2 doivent avoir les mêmes attributs et si une même occurrence existe dans R1 et R2, elle n'apparaît qu'une seule fois dans le résultat de l'union. Le résultat de l'union est une nouvelle relation qui a les mêmes attributs que R1et R2. Si R1 et R2 sont vides, la relation qui résulte de l'union est vide. Si R1 (respectivement R2) est vide, la relation qui résulte de l'union est identique à R2 (respectivement R1).

Ci-dessous un exemple d'union :

Soient les deux relations Usine1 et Usine2

Usine1	
Nom_Usine	**Nombre_Sac_Produit**
IJENDA	2777
TORA	3035
RWEGURA	5680
MABAYI	2340

Usine2	
Nom_Usine	**Nombre_Sac_Produit**
IJENDA	2777
TORA	3035
BUHORO	2450
TEZA	3480

La requête en SQL est la suivante :

SELECT DISTINCT Nom_Usine, Nombre_Sac_Produit FROM Usine1 UNION SELECT Nom_Usine, Nombre_Sac_Produit FROM Usine2 ;

Voici son équivalent en algèbre relationnelle :

$\prod_{\text{Nom_Usine, Nombre_Sac_Produit}}(\textbf{Usine1}) \cup \prod_{\text{Nom_Usine, Nombre_Sac_Produit}} (\textbf{Usine2})$

Cette requête donne tous les noms des usines et le nombre de sacs produits appartenant à l'Usine1 ou à l'Usine2.

Voici ci-dessous le résultat de l'opération d'union appliquée entre les deux relations Usine1 et Usine2 :

Usine1 ∪ Usine2	
Nom_Usine	**Nombre_Sac_Produit**
IJENDA	2777

TORA	3035
RWEGURA	5680
MABAYI	2340
BUHORO	2450
TEZA	3480

Figure 33: Exemple de l'opération d'Union

V.6. Intersection

Définition : L'intersection est une opération portant sur deux relations R1 et R2 ayant le même schéma et construisant une troisième relation dont les n-uplets sont constitués de ceux appartenant aux deux relations. [AR21]

Notation : R1 ∩ R2.

Il s'agit d'une opération binaire ensembliste commutative dont la signature est :

$$\text{relation} \times \text{relation} \rightarrow \text{relation}$$

Comme nous l'avons déjà dit, R1 et R2 doivent avoir les mêmes attributs. Le résultat de l'intersection est une nouvelle relation qui a les mêmes attributs que R1 et R2. Si R1 ou R2 ou les deux sont vides, la relation qui résulte de l'intersection est vide.

Voici ci-dessous un exemple d'intersection :

Soient deux relations **Usine1** et **Usine2 :**

Usine1	
Nom_Usine	**Nombre_Sac_Produit**
IJENDA	2777
TORA	3035
RWEGURA	5680
MABAYI	2340

Usine2	
Nom_Usine	**Nombre_Sac_Produit**

IJENDA	2777
TORA	3035
BUHORO	2450
TEZA	3480

La requête en SQL est la suivante :

SELECT DISTINCT Nom_Usine, Nombre_Sac_Produit FROM Usine1 INTERSECT SELECT Nom_Usine1, Nombre_Sac_Produit FROM Usine2 ;

Son équivalent en algèbre relationnelle est la suivante :
$\prod_{\text{Nom_Usine, Nombre_Sac_Produit}}(\textbf{Usine1}) \cap \prod_{\text{Nom_Usine, Nombre_Sac_Produit}}(\textbf{Usine2})$

Voici la relation résultante de leur intersection :

Usine1 ∩ Usine2	
Nom_Usine	**Nombre_Sac_Produit**
IJENDA	2777
TORA	3035

Figure 34: Exemple de l'opération d'Intersection

V.7 Différence

Définition : La différence est une opération portant sur deux relations R1 et R2 ayant le même schéma et construisant une troisième relation dont les n-uplets sont constitués de ceux ne se trouvant que dans la relation R1. [AR21]

Notation : R1 - R2.

Il s'agit d'une opération binaire ensembliste non commutative essentielle dont la signature est :

$$\text{relation} \times \text{relation} \rightarrow \text{relation}$$

Comme nous l'avons déjà dit, R1 et R2 doivent avoir les mêmes attributs. Le résultat

de la différence est une nouvelle relation qui a les mêmes attributs que R1 et R2. Si R1 est vide, la relation qui résulte de la différence est vide aussi. Si R2 est vide, la relation qui résulte de la différence est identique à R1.

Voici ci-dessous l'exemple de différence entre deux relations **Transport1** et **Transport2** :

Transport1	
Transporteur	**Tonnage**
SOTRADE	337824
ETS NYANDWI	56895
ITRACOM	251229
KAYOMBE	22350
ATS BURUNDI	155240
GEBUTRA	171372

Transport2	
Transporteur	**Tonnage**
SOTRADE	337824
ETS NYANDWI	56895
GEBUTRA	171372

La requête en SQL est la suivante :

SELECT Transporteur, Tonnage FROM Transport EXCEPT SELECT Transporteur, Tonnage FROM Transport2 ;

Son équivalent en algèbre relationnelle est la suivante :

$\prod_{\text{Transporteur, Nombre_Sac_Produit}}(\textbf{Transport1}) - \prod_{\text{Transporteur, Transport}}(\textbf{Transport2})$

Transport1 – Transport2	
Transporteur	**Tonnage**

ITRACOM	251229
KAYOMBE	22350
ATS BURUNDI	155240

Figure 35: Exemple de l'opération de différence

V.8 Produit cartésien

Définition : Le *produit cartésien* est une opération portant sur deux relations R1 et R2 et qui construit une troisième relation regroupant exclusivement toutes les possibilités de combinaison des occurrences des relations R1 et R2. [AR21]
Notation : R1 × R2.

Il s'agit d'une opération binaire commutative essentielle dont la signature est :

$$\text{relation} \times \text{relation} \rightarrow \text{relation}$$

Le résultat du produit cartésien est une nouvelle relation qui a tous les attributs de R1 et tous ceux de R2. Si R1 ou R2 ou les deux sont vides, la relation qui résulte du produit cartésien est vide. Le nombre d'occurrences de la relation qui résulte du produit cartésien est le nombre d'occurrences de R1 multiplié par le nombre d'occurrences de R2.

Voici ci-dessous l'exemple du produit cartésien entre les relations **Details_Frais** et **Transporteur** :

Details_Frais	
ID_Details	**Nom_Frais**
6	Taxe
7	Manutention OTB
8	Frais carburant

Transporteur		
ID_Transporteur	**Nom_Transporteur**	**Place**
3	ITRACOM	2
5	ETS	3

7	SOTARDE	1

Voici la requete SQL Correspondant au produit cartésien :

*SELECT * FROM Details_Frais, Transporteur WHERE Deatails_Frais.ID_Details_frais = Transporteur.ID_Transporteur ;*

Son équivalent en algèbre relationnelle est la suivante :

Details_Frais × Transporteur

Details_Frais × Transporteur				
ID_Details	**Nom_Frais**	**ID_Transporteur**	**Nom_Transporteur**	**Place**
6	Taxe	3	ITRACOM	2
6	Taxe	5	ETS	3
6	Taxe	7	SOTARDE	1
7	Manutention OTB	3	ITRACOM	2
7	Manutention OTB	5	ETS	3
7	Manutention OTB	7	SOTARDE	1
8	Frais carburant	3	ITRACOM	2
8	Frais carburant	5	ETS	3
8	Frais carburant	7	SOTARDE	1

Figure 36: Exemple de l’opération du produit cartésien

V.9. Jointures

V.9.1 Définition

La jointure est une opération portant sur deux relations R1 et R2 qui construit une troisième relation regroupant exclusivement toutes les possibilités de combinaison des occurrences des relations R1 et R2 qui satisfont l'expression logique E. [AR21]

La jointure est notée : **R1** $\bowtie_E$ **R2**.

Il s'agit d'une opération binaire commutative dont la signature est :

relation × relation × expression logique → relation

Si R1 ou R2 ou les deux sont vides, alors la relation qui résulte de la jointure est vide. En fait, la jointure n'est rien d'autre qu'un produit cartésien suivi d'une sélection :

$$R1 \bowtie_E R2 = \boldsymbol{\sigma}_E (R1 \times R2)$$

Exemple de jointure entre la relation :

TRASPORTEUR		
Nom	**Prénom**	**Age**
Didier	Irakoze	6
Fubini	Claude	42
Donatien	Izere	16

CAMION		
âgeC	**Type**	**Tonnage**
50	Fuso	30
35	Ben	60
20	Actros	45
15	Dyna	15

TRANSPORTEUR ⋈ ((Age ≤ âgeT) ⋀ (Tonnage ≤ 50)) CAMION					
Nom	**Prénom**	**Age**	**âgeC**	**Type**	**Tonnage**
Didier	Irakoze	6	50	Fuso	30
Didier	Irakoze	6	20	Actros	45
Didier	Irakoze	6	15	Dyna	15
Donatien	Izere	16	50	Fuso	30
Donatien	Izere	16	20	Actros	45
Donatien	Izere	16	15	Dyna	15
Fubini	Claude	42	50	Fuso	30

Figure 37: Exemple de l'opération de jointure

La requête SQL correspondante est la suivante :

*SELECT * FROM Transporteur, Camion WHERE Transporteur.Age = Camion.âgeC AND Transporteur.Tonnage <= 50 ;*

Cette jointure permet de générer toutes les possibilités d'association entre un camion et un chauffeur en respectant l'âge maximum conseillé pour un conduire un camion et le poids de 50 tonnes à ne pas dépasser.

V.9.2 Équi-jointure

Définition : Une équi-jointure est une thêta-jointure dans laquelle l'expression logique E est un test d'égalité entre un attribut A1 de la relation R1 et un attribut A2 de la relation R2. L'équi-jointure est notée :

R1 $\bowtie_{A1=A2}$ **R2**.

Remarque : Il vaut mieux écrire R1 $\bowtie_{A1=A2}$R2 que R1 $\bowtie_{A1, A2}$R2, car cette dernière notation, bien que parfois dans la littérature, prête à confusion avec une jointure naturelle explicite.

V.9.3 Jointure naturelle

Définition : Une jointure naturelle est une jointure dans laquelle l'expression logique E est un test d'égalité entre les attributs qui portent le même nom dans les relations R1 et R2. Dans la relation construite, ces attributs ne sont pas dupliqués, mais fusionnés en une seule colonne par couple d'attributs. [AR21]

La jointure naturelle est notée R1 ⋈ R2.

Si la jointure ne doit porter que sur un sous-ensemble des attributs communs à R1 et R2 il faut préciser explicitement ces attributs de la manière suivante :
$R1 \bowtie_{A1, \ldots, An} R2$.

Généralement, R1 et R2 n'ont qu'un attribut en commun. Dans ce cas, une jointure naturelle est équivalente à une équi-jointure dans laquelle l'attribut de R1 et celui de R2 sont justement les deux attributs qui portent le même nom.

Pour effectuer une jointure naturelle entre R1 et R2 sur un attribut A1 commun à R1 et R2, il vaut mieux écrire $R1 \bowtie_{A1} R2$ que $R1 \bowtie R2$.

En effet, si R1 et R2 possèdent deux attributs portant un nom commun, A1 et A2, $R1 \bowtie_{A1} R2$ est bien une jointure naturelle sur l'attribut A1, mais $R1 \bowtie R2$ est une jointure naturelle sur le couple d'attributs A1,A2, ce qui produit un résultat très différent .

Exemple :

PERSONNE		
nom	**prénom**	**âge**
Fourt	**Lisa**	**6**
Juny	**Carole**	**40**
Fidus	**Laure**	**20**
Choupy	**Emma**	**6**

CADEAU		
âge	**article**	**prix**
40	livre	45
6	poupée	45
20	montre	87

Ci-dessous la jointure naturelle PERSONNE $\bowtie$ CADEAU qui peut également s'écrire :
PERSONNE $\bowtie_{âge}$ CADEAU.

PERSONNE ⋈ CADEAU				
nom	**prénom**	**âge**	**article**	**prix**
Fourt	Lisa	6	poupée	25

Juny	Carole	40	livre	45
Fidus	Laure	20	montre	87
Choupy	Emma	6	poupée	25

Figure 38: Exemple de l'opération de jointure naturelle

La requête SQL correspondante est la suivante :

*SELECT * FROM PERSONNE, CADEAU WHERE PERSONNE.âge = CADEAU.âge;*

V.10 Division

Définition : la division est une opération portant sur deux relations R1 et R2, telles que le schéma de R2 est strictement inclus dans celui de R1, qui génère une troisième relation regroupant toutes les parties d'occurrences de la relation R1 qui, associées à toutes les occurrences de la relation R2, se retrouvent dans R1.

Notation : R1÷ R2.

Il s'agit d'une opération binaire non commutative dont la signature est :

$$\text{relation} \times \text{relation} \rightarrow \text{relation}$$

Autrement dit, la division de R1 par R2 (R1÷ R2) génère une relation qui regroupe tous les n-uplets qui, concaténés à chacun des n-uplets de R2, donne toujours un n-uplet deR1.

La relation R2 ne peut pas être vide. Tous les attributs de R2 doivent être présents dans R1 et R2 doit posséder au moins un attribut de plus que R2

(Inclusion stricte). Le résultat de la division est une nouvelle relation qui a tous les attributs de R1 sans aucun de ceux de R2. Si R1 est vide, la relation qui résulte de la division est vide.

Ci-dessous, un exemple de division Transporteur ÷ Camion qui permet de dresser la table R de tous les Chauffeurs de la relation Transporteur qui conduisent tous les types de camions de la relation Camion.

Transporteur	
Nom	**Type_Camion**
Didier	Fuso

Fubini	Dyna
Donatien	Fuso
Didier	Dyna
Fubini	Fuso
Didier	Ben
Donatien	Ben

Camion
Type_Camion
Fuso
Dyna

Transporteur ÷ Camion
Nom
Didier
Fubini

Figure 39**:** Exemple de l'opération de Division

La requête SQL correspondante est la suivante :

SELECT * FROM Transporteur as t WHERE NOT EXIST ((SELECT Camion.Type_Camion) EXPECT (SELECT Transporteur.Type_Camion FROM Transporteur WHERE Camion.Tye_Camion =t.Type_Camion))

Son équivalent en algèbre relationnelle est la suivante :

$$\prod_{\text{Nom, Type_Camion}}(\textbf{Transport1}) \div \prod_{\text{Type_Camion}}(\textbf{Transport2})$$

V.11 Conclusion :

Au cours de ce chapitre nous avons pu donner la description de notre entrepôt de données tout en utilisant l'algèbre relationnelle comme modèle mathématique. Nous avons pu démontrer les relations qui existent les faits et les dimensions tout en

appliquant les opérations de l'algèbre relationnelles agissant sur les données de l'entrepôt.

CHAPITRE VI : IMPLEMENTATION DE LA SOLUTION

L'implémentation d'un entrepôt des données doit dépendre des plateformes utilisées car il y a des plateformes qui peuvent exiger certaines configurations :

Les plateformes utilisées sont :

- Le matériel (PC client, Serveur de stockage des données) ;
- Les logiciels (Système d'exploitation : Windows7 ; Microsoft Office Professionnel 2016 ; SQL server 2014 ; Business Intelligence intégré en Visual studio Professionnel 2013 en collaboration avec SQL server Data Tools for Visual 2013 qui est intégré dans SQL server 2014).

Un cube est une structure multidimensionnelle qui contient des dimensions et des mesures. Les dimensions définissent la structure sur le cube tandis que les mesures fournissent les valeurs numériques agrégées présentant un intérêt pour l'utilisateur final. Un cube permet à une application cliente d'extraire des valeurs comme si les cellules sur le cube définissaient toutes les valeurs synthétisées possibles. Les positions des cellules dans le cube sont définies par l'intersection des membres de dimension [AR15].

Un cube est une structure multidimensionnelle qui contient des dimensions et des mesures. Les dimensions définissent la structure sur le cube tandis que les mesures fournissent les valeurs numériques agrégées présentant un intérêt pour l'utilisateur final. Un cube permet à une application cliente d'extraire des valeurs comme si les cellules sur le cube définissaient toutes les valeurs synthétisées possibles. Les positions des cellules dans le cube sont définies par l'intersection des membres de dimension [AR15].

Nous adressons la liste des cubes tout en précisant ses dimensions et ses mesures.

Tableau 24: Liste des cubes

Nom sur le cube	Dimension	Mesures
Vente	Type vente Type_monnaie Quantité Usine Date-operation	Nombre de Kg vendus
Depenses	Type_vente Type_stok Date_Oprartion	Le montant des dépenses
Stock_local	Type_stock Quantite Paquet Date_operation	Quantité de the stock localement
Stock_Export	Type_stock Quantite Date_operation	Quantité de the pour le stock export
Transport	Date_Operation Transporteur	De kg transporte par un chauffeur

Pour bien manipuler les cubes de notre entrepôt, on a mis en place une interface qui va faciliter la communication entre l'entrepôt et l'utilisateur. Dans la suite elle sera utilisée sous le nom d'interface générale.

L’interface de navigation donne accès aux utilisateurs analystes pour pouvoir manipuler les données selon différents critères. Cette interface se trouve en annexe à ce document.

L’analyse multidimensionnelle exige une organisation spécifique des données car les données de l’entrepôt doivent être organisées sous forme d’un hypercube où chaque cellule de l’hypercube contient une donnée observée ou calculée.

Les cubes OLAP ont les caractéristiques suivantes : obtenir des informations déjà agrégées selon les besoins de l’utilisateur, simplicité et rapidité d’accès, capacité à manipuler les données agrégées selon différentes dimensions et un cube utilise les fonctions classiques d’agrégation : min, max, count, sum, avg. [AR16]

Dans la suite, nous allons déterminer les opérations attachées à la manipulation des données multidimensionnelles ainsi que les requêtes Ad hoc utilisées pour construire les tableaux croisés, nous prenons le cube « Vente » pour montrer les opérations liées au cube dimensionnel.

Les opérations agissant sur la structure multidimensionnelle visent à changer le point de vue des données observées.

La rotation (rotate) consiste à effectuer une rotation de l'hypercube afin de présenter une face différente [17].

- Le cube de quantité du thé vendu par Usine, par marché et par année

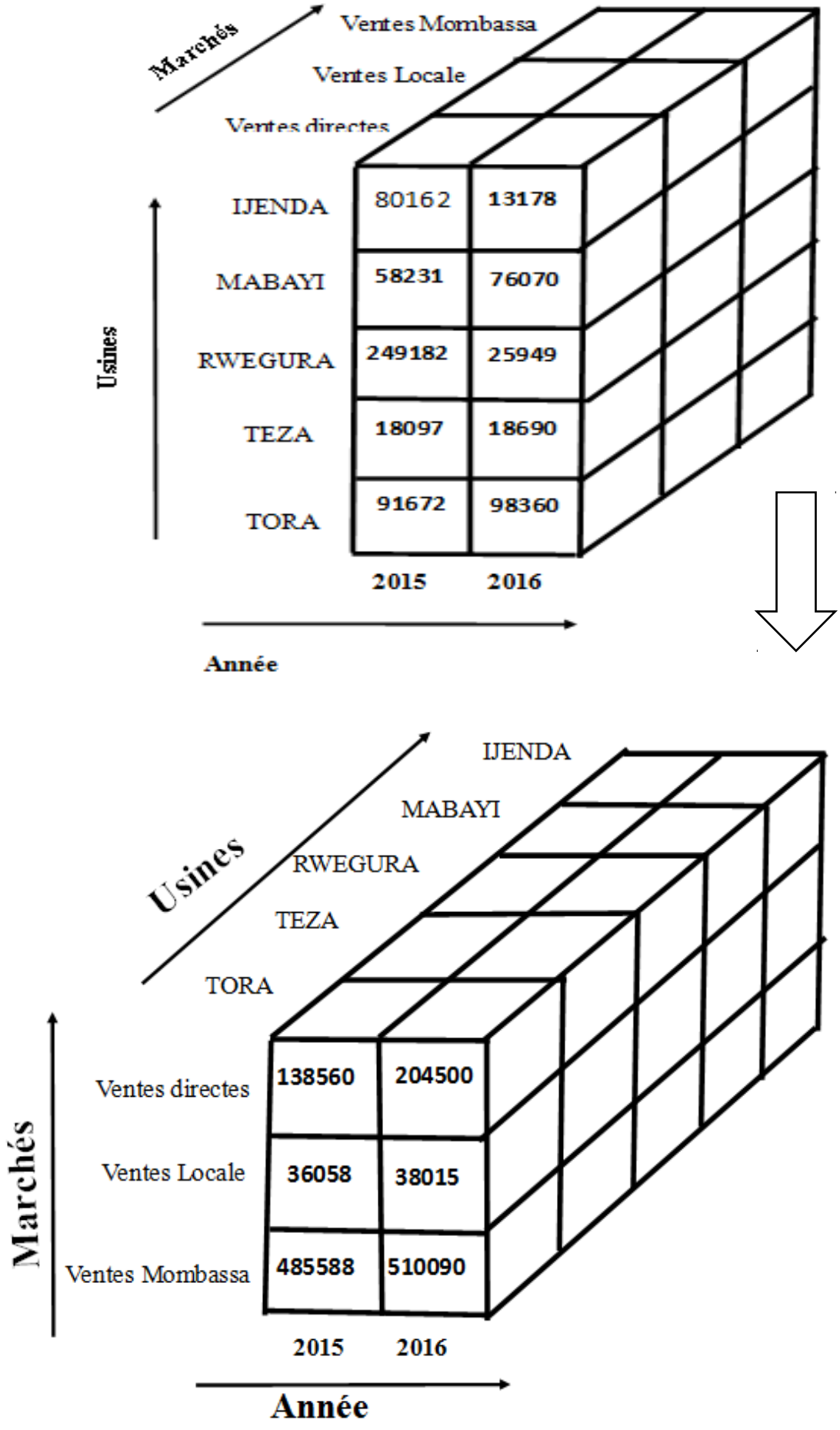


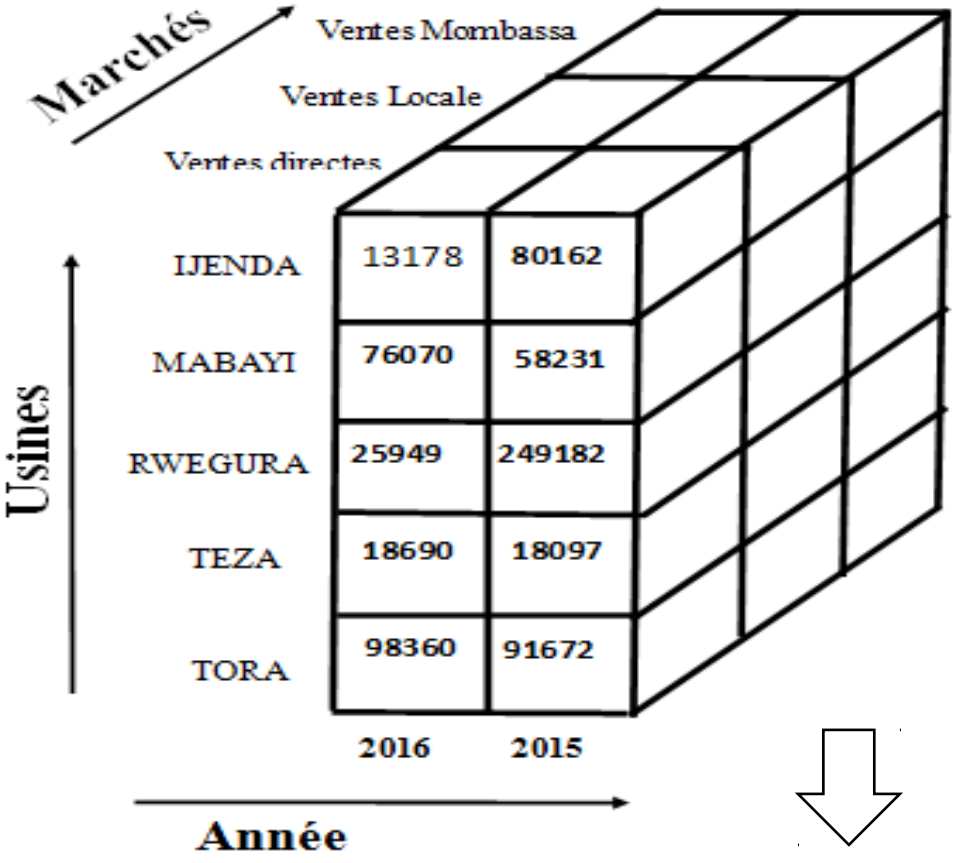
Figure 40: Opération de rotation sur le cube Vente « Quantité de the vendu par usine et par année »

Voici requête ad hoc correspondante a cette rotation

« Select Nom_vente, Annee, Nom_Usine, Sum(Nbre_Kg_Vendus) as QUANTITE from Vente, Type_Vente, Date_Operation, Usine where Vente.Id_Type_Vente= Type_Vente.Id_Type_Vente and Vente.Id_Date= Date_Operation.Id_Date and Vente.Id_Usine= Usine.Id_Usine and Annee = 2015 or Annee = 2016 group by Nom_vente, Annee, Nom_Usine ; »

Son équivalent en Algèbre relationnelle est la suivante :

π Nom_vente, Annee, Nom_Usine, $\gamma_{\text{Sum(Nbre_Kg_Vendus)}}$ (Type_**Vente**$\bowtie_{\text{Id_Type_Vente}}$ Vente $\bowtie_{\text{Id_Usine}}$Usine) $\wedge$ (Anne = 2015 V Annee = 2016)



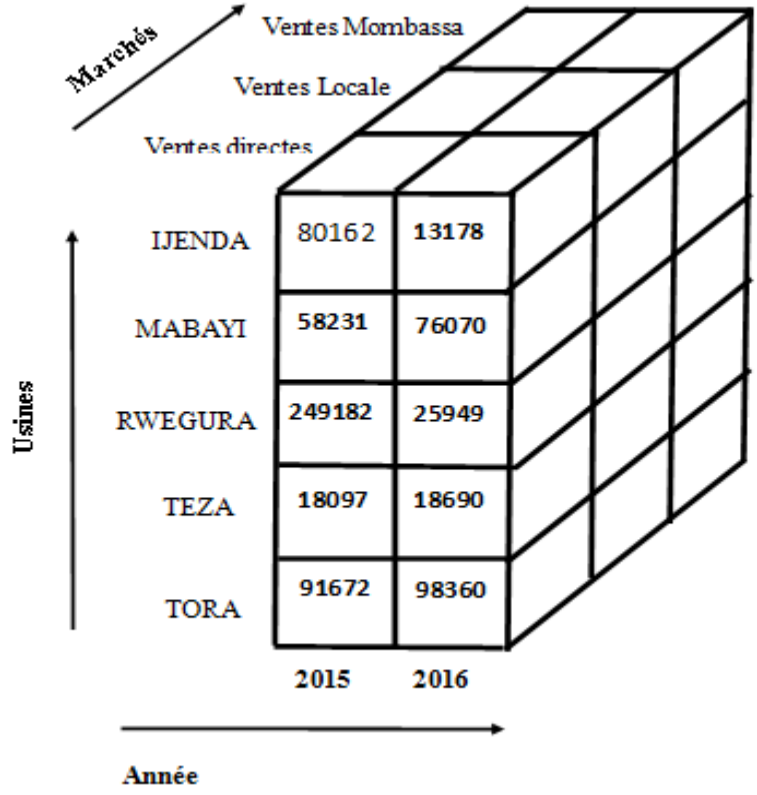


Figure 41: Opération de permutation sur le cube Vente « quantité du thé vendu par usine et par année »

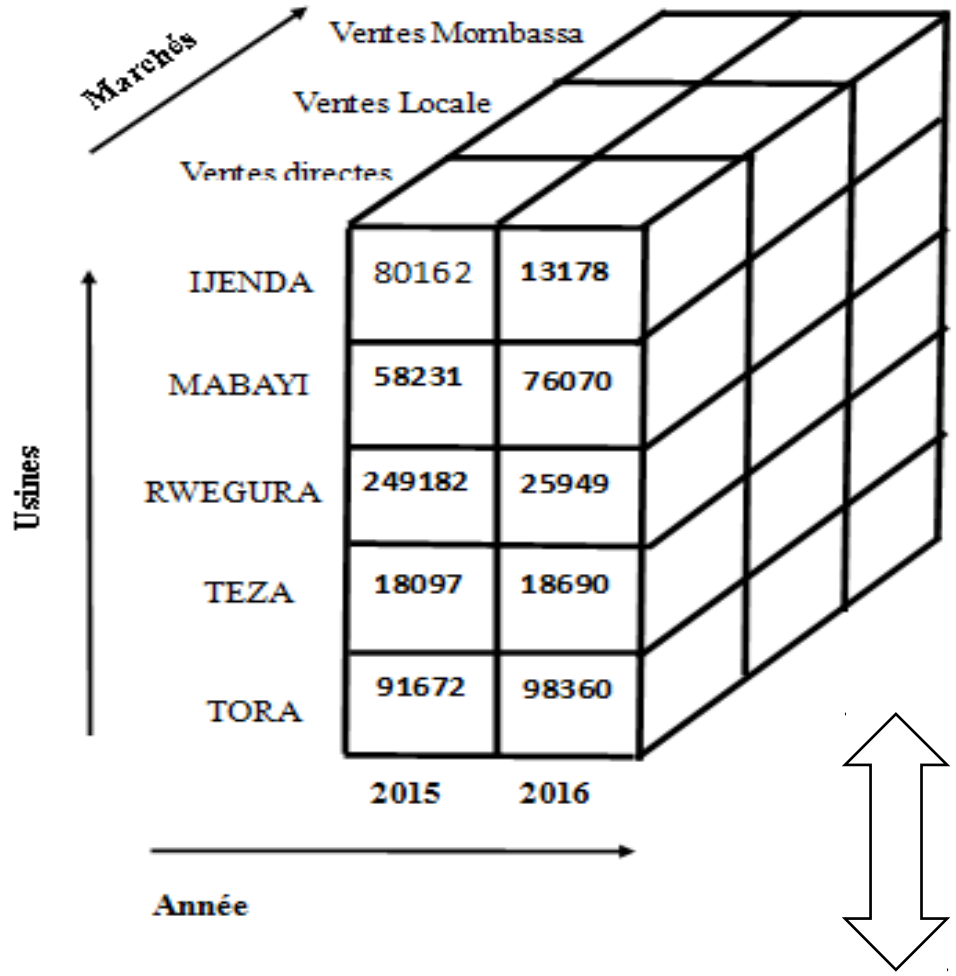
Voici requête ad hoc correspondante a cette permutation :

« Select Nom_vente, Annee, Nom_Usine, Sum(Nbre_Kg_Vendus) as QUANTITE from Vente, Type_Vente, Date_Operation, Usine where Vente.Id_Type_Vente= Type_Vente.Id_Type_Vente and Vente.Id_Date= Date_Operation.Id_Date and Vente.Id_Usine= Usine.Id_Usine and Annee = 2016 or Annee = 2015 group by Nom_vente, Annee, Nom_Usine ; »

Son équivalent en Algèbre relationnelle est la suivante :

π Nom_vente, Annee, Nom_Usine, $\gamma_{Sum(Nbre_Kg_Vendus)}$ (Type_**Vente**$\bowtie_{Id_Type_Vente}$ Vente $\bowtie_{Id_Usine}$Usine) $\wedge$ (Anne = 2016 V Annee = 2015)

La division (split) consiste à représenter chaque tranche de l'hypercube en
passant à une représentation tabulaire. Plus généralement, cette opération permet de réduire le nombre de dimensions



Usine	2015
JENDA	80162
MABAYI	58231
RWEGURA	249182
TEZA	18297
TORA	91672

Usine	2016
JENDA	13178
MABAYI	76070
RWEGURA	25949
TEZA	18690
TORA	98360

Figure 42: Opération de division sur le cube Vente « quantité du thé vendu par usine et par année »

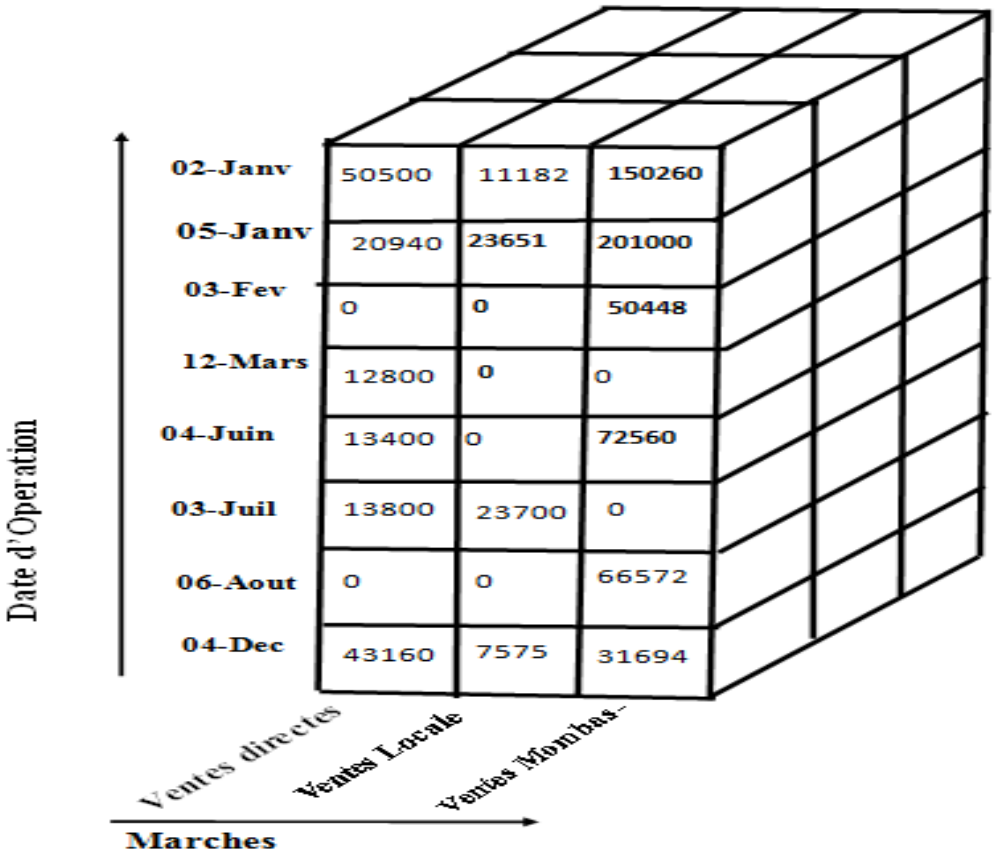
Les opérations agissant sur la granularité d'observation des données caractérisent la hiérarchie de navigation entre les différents niveaux.

Elles correspondent aux opérations suivantes :

- Roll- up
- Slicing
- Dicing

Roll- up (forage vers le haut) est une représentation des données à un niveau de granularité supérieure conformément à la hiérarchie définie sur la dimension.

Nous allons appliquer cette operation sur notre cube suivant :



Requete Adhoc correspondant :

```
Select distinct Date Enregistrement, Nom_vente,
Nom_Usine, sum(Nbre_Kg_Vendus) as quantite from
Type_Vente as tv, Vente as v, Usine as u where
tv.Id_Type_Vente=v.Id_Type_Vente and
u.Id_Usine=v.Id_Usine
group by Nom_vente, Date_Enregistrement, Nom_Usine ;
```

Requête en algèbre relationnelle correspondante :

$$\pi_{\text{Date_Enregistrement, Nom_vente, Nom_Usine,}}\ \gamma_{\text{sum(Nbre_Kg_Vendus)}}\ (\text{Type_Vente} \bowtie_{\text{Id_Type_Vente}} \text{Vente} \bowtie_{\text{Id_Usine}} \text{Usine})$$

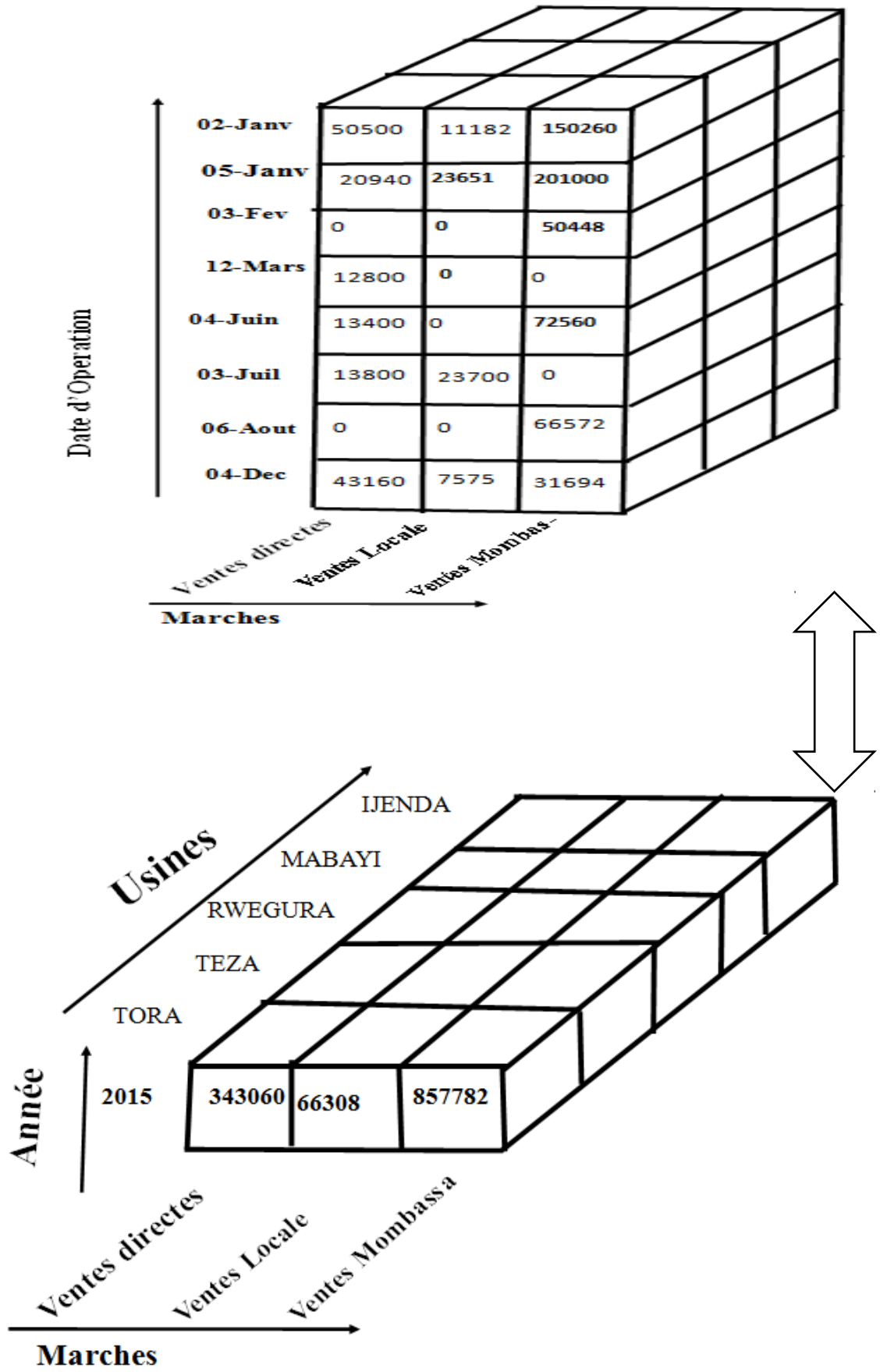
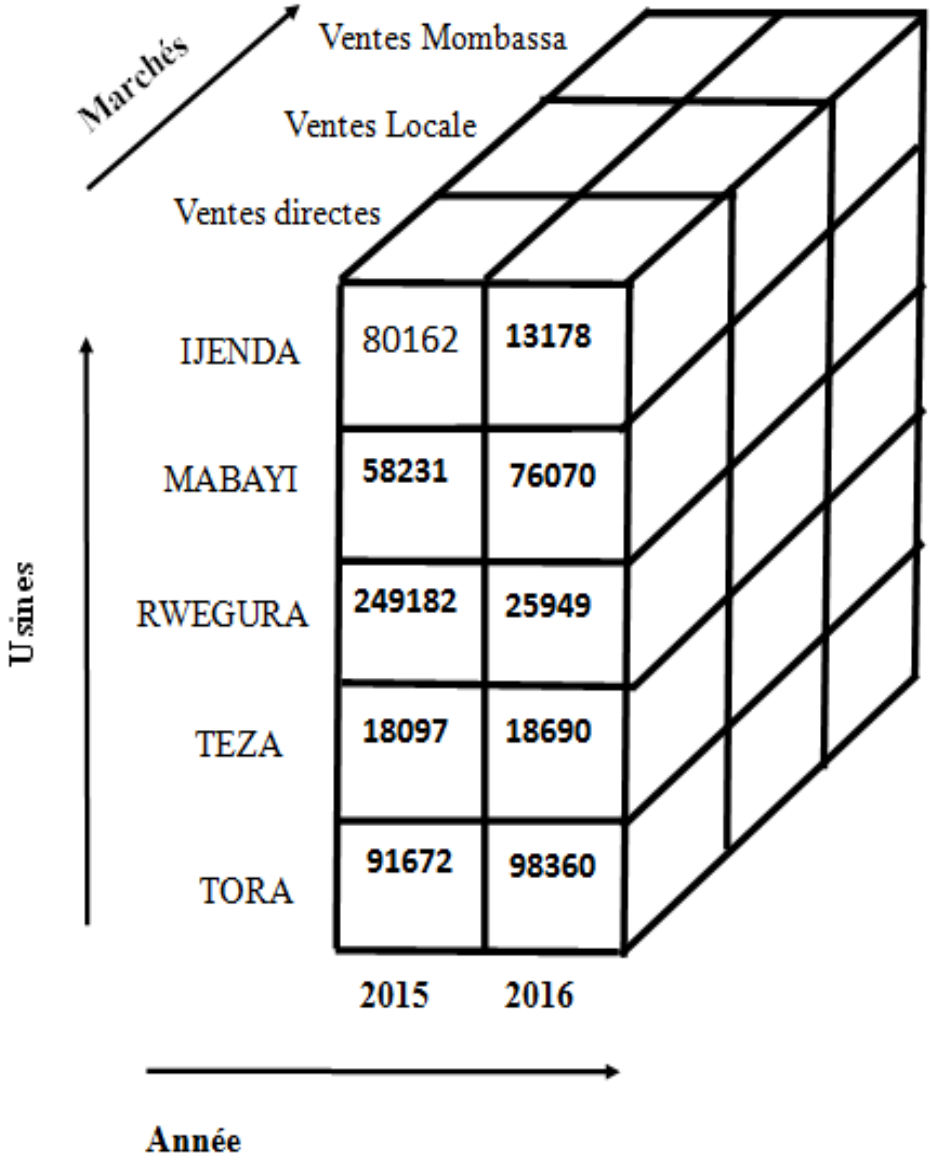
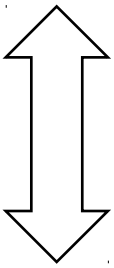


Figure22 : Opération de Roll up sur le cube Vente « quantité du thé vendu par usine et par année »

C’est une sélection de tranches de l’hypercube par des prédicats selon une dimension.





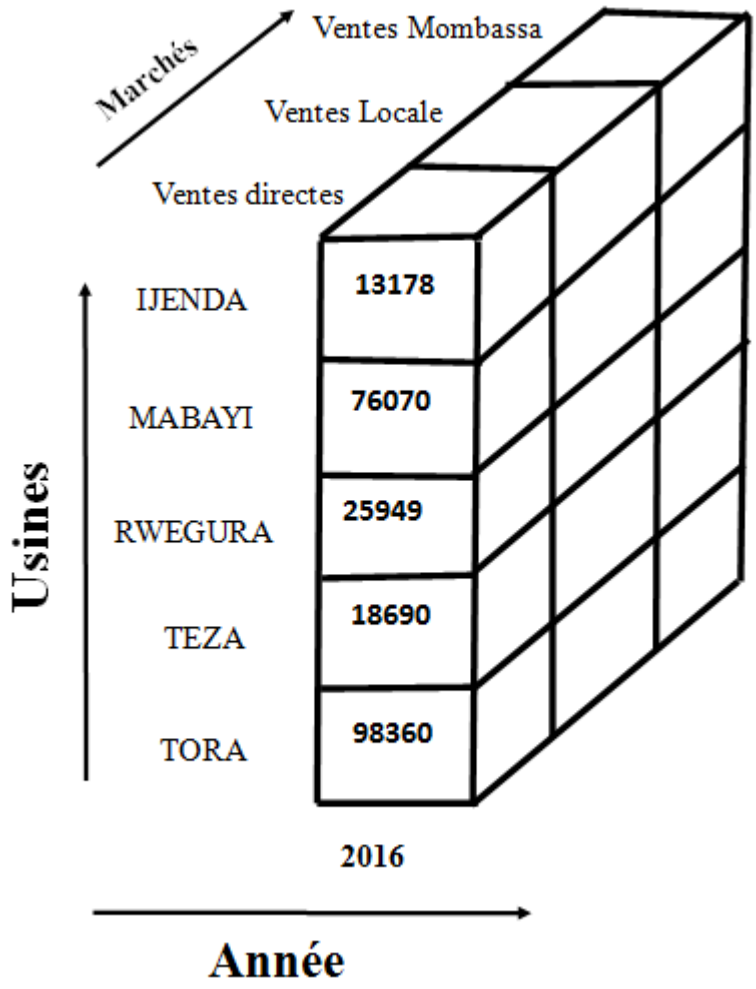
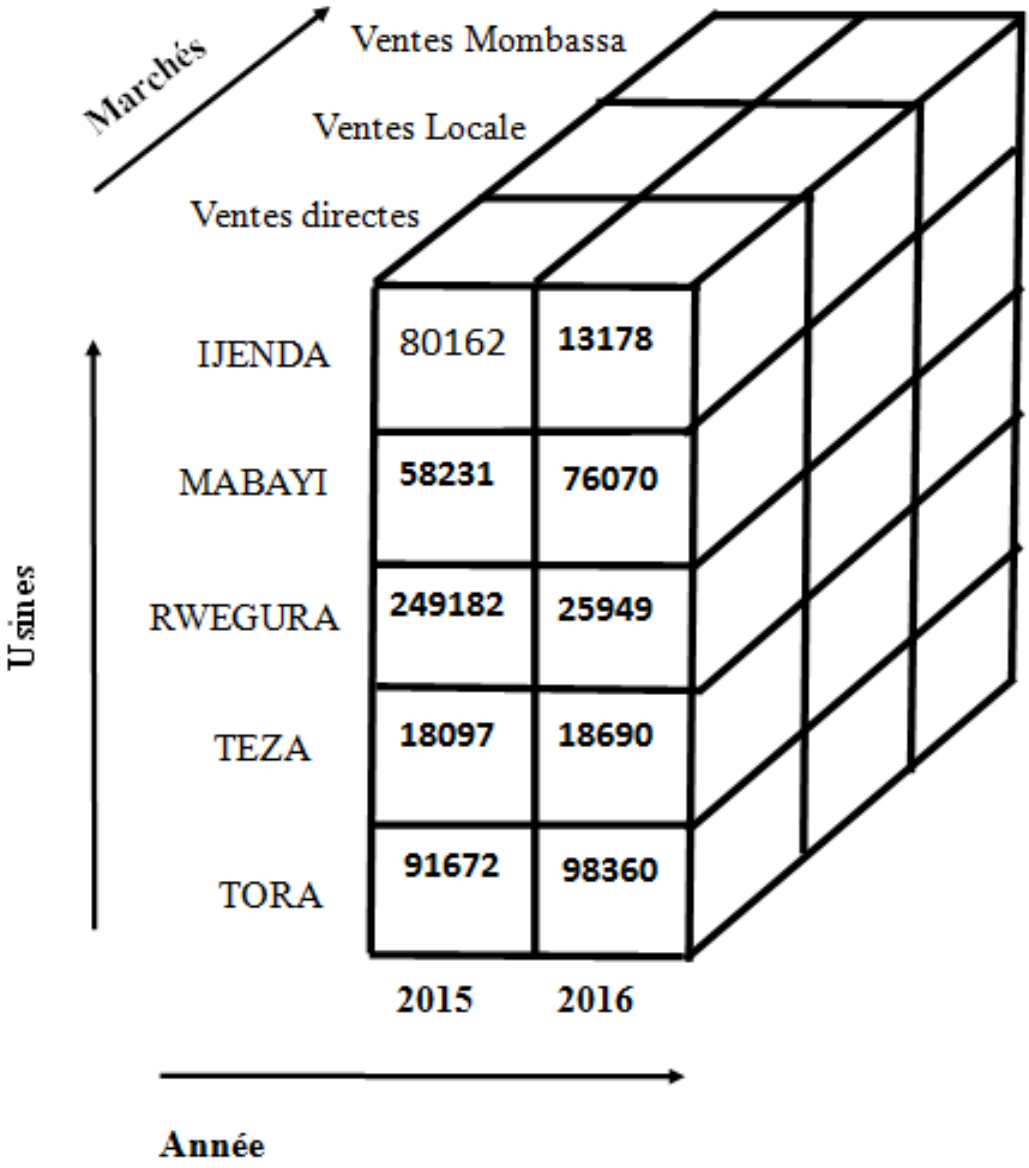


Figure23 : Opération de division sur le cube Vente « quantité du thé vendu par usine et par année »

VI.4.2.2.3 Dicing

Cette opération permet d'extraire d'un sous-hypercube c'est à dire une sorte de cumuls de sélection.



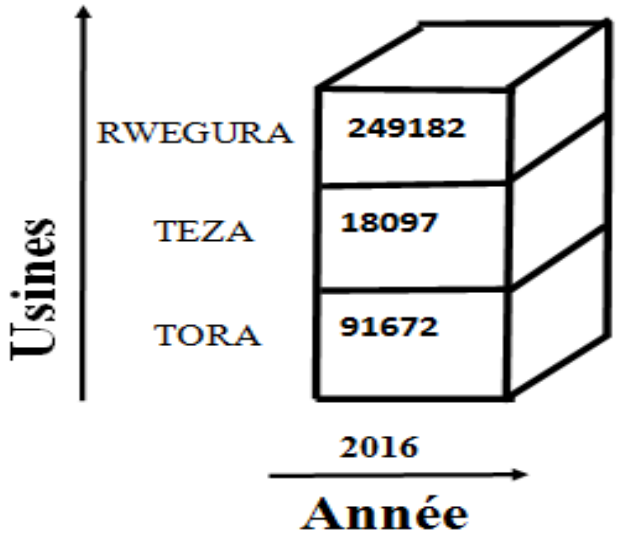


Figure 23 : Opération de Dicing sur le cube Vente « quantité du thé vendu par usine et par année »

VI.4.3 Conclusion

Dans ce chapitre nous avons pu montrer les opérations agissant sur le modèle multidimensionnel en se basant sur les données de notre entrepôt suivant les différents axes d'analyse.

CHAPITRE VII : SECURITE DES DONNEES

VII.1. Définition

La sécurité informatique est un domaine vaste qui peut appréhender dans plusieurs domaines :

- Les systèmes d'informations
- Les réseaux informatiques
- Les accès physiques à des salles machines

Pour notre projet, nous nous concentrerons sur les aspects relatifs aux systèmes informatiques.

VII.1.1 Critères fondamentaux

Les solutions de sécurité doivent contribuer à satisfaire au moins les critères suivant :

- La disponibilité : la probabilité de pouvoir mener correctement à terme une session de travail ;
- L'intégrité ;
- La confidentialité.

Dans la suite nous allons plus nous baser sur l'intégrité et la confidentialité.

❖ L'intégrité

- L'intégrité permet de certifier que les données, les traitements ou les services n'ont pas été modifiés, altérés ou détruits tant de façon intentionnelle qu'accidentelle.
- L'altération est principalement occasionnée par le média de transmission mais peut provenir du système d'informations

❖ La confidentialité

« La confidentialité est le maintien du secret des informations » **[Le Petit Robert]**

- Dans le cadre d'un système d'information, cela peut être vu comme une protection des données contre une divulgation non autorisée.

- 2 actions complémentaires permettant d'assurer la confidentialité des données :
  - ✓ Limiter leur accès par un mécanisme de contrôle d'accès
  - ✓ Transformer les données par des procédures de chiffrement

C’est par ce deuxième point qui consiste à transformer les donnes par chiffrement que nous allons assurer la sécurité de donnes de notre entrepôt de données :

VII.1.2 L'identification et l'authentification

L'identification de l'auteur d'un document peut être aisé , par contre être en mesure d'assurer l'authenticité du document est chose plus délicate.

- L'identification peut être vu comme un simple login de connexion sur un système
- L'authentification peut être un mot de passe connu seulement par l'utilisateur.

VII.2 Sécurité des données de l’entrepôt

Pour assurer la sécurité des données de notre entrepôt, nous avons choisi d’utiliser une politique d’identification et d’authentification avant de faire quoi que ce soit sur l’entrepôt. Afin renforce la sécurité, nous allons mettre en place un système de chiffrement par le crypto système DES pour les logins et les mots de passe des utilisateurs.

VII.2.1 Historique Crypto système DES

Le 15 mai 1973 le **NBS** (*National Bureau of Standards, aujourd'hui appelé NIST - National Institute of Standards and Technology*) a lancé un appel dans le Federal Register (*l'équivalent aux Etats-Unis du Journal Officiel en France*) pour la création d'un algorithme de chiffrement répondant aux critères suivants :

- Posséder un haut niveau de sécurité lié à une clé de petite taille servant au chiffrement et au déchiffrement
- Être compréhensible
- Ne pas dépendre de la confidentialité de l'algorithme
- Être adaptable et économique
- Être efficace et exportable

Fin 1974, IBM propose « Lucifer », qui, grâce à la NSA (National Security Agency), est modifié le 23 novembre 1976 pour donner le **DES** (*Data Encryption Standard*). Le DES a finalement été approuvé en 1978 par le NBS. Le DES fut normalisé par l'*ANSI* (*American National Standard Institute*) sous le nom de *ANSI X3.92*, plus connu sous la dénomination *DEA* (*Data Encryption Algorithm*).

VII.2.2 Principe de fonctionnement de DES

Le DES (Data Encryptions Standard) d'origine IBM (Karl Meyer 1977) est l'algorithme à clé symétrique le plus connu. Il consiste en une suite de substitutions (DES-S) et de transpositions, ou permutations (DES-P), par bloc de 64 bits. La figure suivante illustre de manière simple le principe d'un tel code. Utilisant une clé de 56 bits (64 bits dont 8 de parité), le DES est aujourd'hui facilement « cassable », il est remplacé par le triple DES (3DES, application de 3 DES successivement avec 3 clés indépendantes).

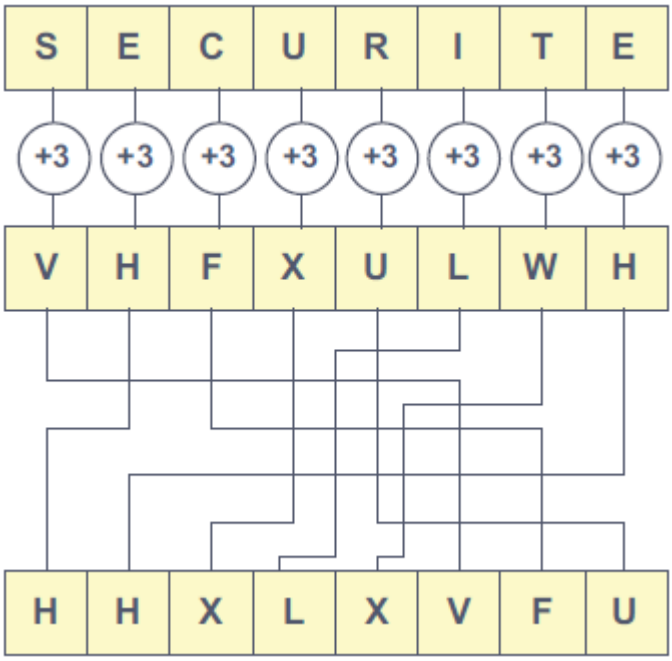


Figure 43 : Principe du DES.

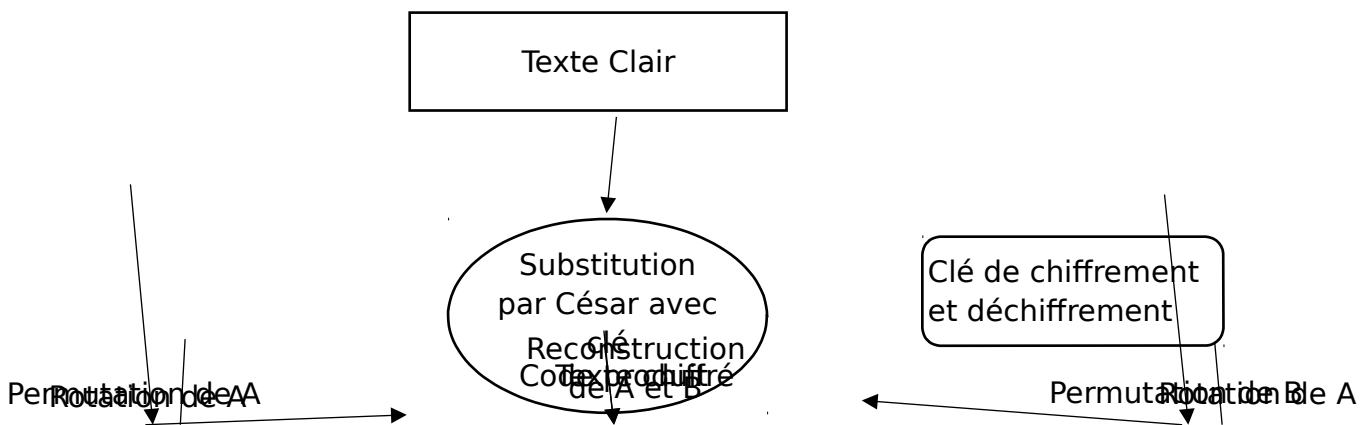
L'algorithme consiste à effectuer des combinaisons, des substitutions et des permutations entre le texte à chiffrer et la clé, en faisant en sorte que les opérations puissent se faire dans les deux sens (pour le déchiffrement). La combinaison entre substitutions et permutations est appelée **code produit**.

La clé est codée sur 64 bits et formée de 16 blocs de 4 bits, généralement notés k_1 à k_{16}. Etant donné que « seuls » 56 bits servent effectivement à chiffrer, il peut exister 2^{56} (soit $7.2*10^{16}$) clés différentes !

VII.2.3 Algorithme du DES

Les grandes lignes de l'algorithme DES sont les suivantes :

- Substitution du texte avec la clé par la méthode de César ;
- Découpage des blocs en deux parties : gauche et droite, nommées *A* et *B* ;
- Etapes de permutation et de rotation répétées (appelées **rondes**) ;
- Recollement des parties gauche et droite puis permutation initiale inverse.



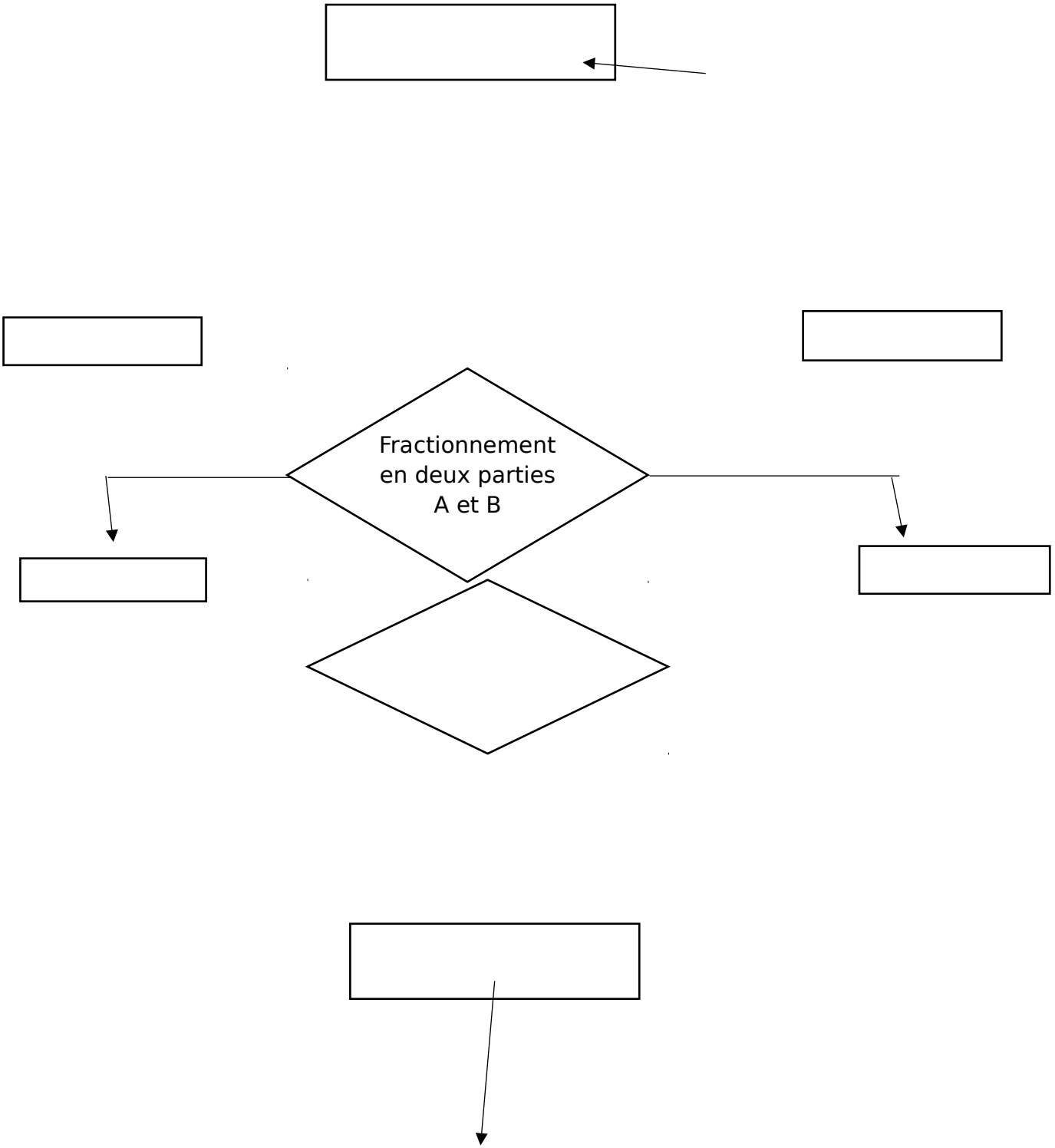


Figure 44: Algorithme de chiffrement par DES

Pour le déchiffrement on applique le même algorithme cette fois en sens inverse

VII.3 Conclusion

Tout au long de ce chapitre, nous avons pu montrer les grandes lignes et le principe de fonctionnement de l'algorithme de chiffrement et déchiffrement par DES. D'une manière générale notre objectif de sécuriser les données de notre entrepôt a été atteint.

CONCLUSION GENERALE ET RECOMENDATIONS

Au début de ce noble travail, on ne possédait que peu de connaissance en informatiques décisionnelle tout en connaissant beaucoup de choses sur les bases de

données transactionnelles. Mais puisque on avait une bonne intention sur les entrepôts de données, on s'est lancé dans ce domaine et on s'est intéressé au domaine du décisionnel dans lequel on souhaite renforcer l'expérience. On a fait des recherches sur internet, on a consulté beaucoup de livres sur les entrepôts de données et des mémoires défendus par d'autres étudiants comme vous allez le remarquer dans la partie « références » du travail réalisé. Le travail est constitué par 6 chapitres :

Dans le premier chapitre, on a essayé de faire la description du travail réalisé, les motivations et les argumentations. On a bien justifié le projet réalisé puis montré tous les objectifs du travail en mettant au clair les différentes problématiques et les solutions proposées.

Dans le deuxième chapitre, on a présenté l'entreprise « O.T.B », son organigramme et ses activités. On s'est focalisé sur sa mission principale et sa structure organisationnelle.

Dans le troisième chapitre, on a fait une introduction sur les systèmes décisionnels. On a introduit les systèmes décisionnels avec des définitions de l'architecture générale des systèmes décisionnels.

Dans le quatrième chapitre, qui est le point principal de notre projet, on a montré les différentes manipulations des données multidimensionnelles et on est entré en profondeur dans l'entrepôt de données qui constitue le cœur du système décisionnel jouant le rôle important au sein de l'entreprise puisqu'il permet de référencer les données souvent éparpillées dans les bases de données.

Dans le cinquième chapitre, c'est l'implémentation du modèle mathématique du travail réalisé. On a implémenté le modèle relationnel, pour voir qu'il y a une interaction entre les différents éléments du système. ARX

Dans le dernier chapitre, qui est un chapitre aussi important du fait qu'il montre la sécurité des données contre les accès non autorisés. On a implémenté le crypto système de chiffrement et déchiffrement DES qui est un système plus efficace dans le domaine de sécurité des données.

En se référant sur les efforts utilisés dans ce travail, nous recommandons à l'office ce qui suit :

- Informatiser le maximum possible des services de l'O.T. B ;
- Avoir des applications les évoluées permettant le traitement rapide de tous toutes les données dans l'entreprise surtout ces qui sont en rapport avec l'activité commerciale ;
- Mettre en place une politique de gestion de ses service d'une façon automatisée.

- Disposer des serveurs plus puissants capables de stocker et traiter de données importants qui serviront dans l'analyse des données de l'entreprise.

Aux futurs chercheurs, nous recommandons ce qui suit :

- D'approfondir la recherche afin de mieux enrichir ce domaine qui n'est pas encore très exploité dans notre pays.
- Mettre beaucoup d'efforts sur la sécurité puisque la sécurité de l'entrepôt de données reste toujours difficile à réaliser ;

En résumé, la contribution de ce travail était de mettre en place un entrepôt de données de gestion de vente des produits de l'OTB ; on peut dire que ce travail nous a inspiré à la recherche car il nous a permis d'acquérir une très bonne expérience professionnelle et de bien se concentrer dans ce domaine qui est intéressant.

Références et bibliographies

Ouvrage généraux

1. DONSEZ D. Systèmes d'information décisionnels (Data Warehouse /Data Mining). Université de Valenciennes, ISTV, 1996-2006 [AR17]

2. FILALI A; KEDJNANE S. Conception et réalisation d'un Data Warehouse pour la

mise en place d’un système décisionnel. Mémoire,Ecole Nationale Supérieur d’Information Ex.INI (Institut National de formation en Informatique) en Algérie 2009/2010 [AR13]

3. INMON. Building the Data Warehouse. Third Edition, Wiley Computer Publishing, 2002 [AR7]

4. KIMBALL R; CASERTA J. The Data warehouse ETL Toolkit. Wiley Publishing, INC 2004 [AR20]

5. KIMBALL R. Entrepôts de données : Guide pratique du concepteur de Data Warehouse. Wiley Computer Publishing 1996 [AR10]

6. KIMBALL R; MARGY R. The Data warehouse toolkit: The Definitive Guide to Dimensional Modeling. Third Edition, United States of America:John Wiley & Sons, 2013 [AR9]

7. POLETTO M. L’informatique décisionnelle. Thèse professionnelle en Informatique. France : Ecole Supérieure d’Informatique CESI EXIA, 2012 [AR14]

8. Rim Chaabane ; Base de données-Algebre relationnelle L3-Informatique Paris8 [AR21]

Mémoires et articles

9. Filali A. : Conception et réalisation d’un data warehouse pour la mise en place d’un système décisionnel, Ecole Normale Supérieur d’Informatique d’Algérie, 2009-2010 [AR2]

10. KHOURI S. Modélisation conceptuelle à base ontologique d’un entrepôt de données. Thèse de magistère en informatique. Alger : Institut National de Formation en Informatique (I.N.I), 2008 [AR6]

11. MARCEL P. Manipulation de Données Multidimensionnelles et

Langages de Règles. Thèse de Doctorat de l'Institut des Sciences

Appliquées de Lyon, 1998 [AR11]

Sites Web

12. , Mars

2018 [AR16]

13. e 2017 à 18h00; [AR3]

14. ; Fevrier 2018 [8]

15. . [AR1]

16. ;Janvier 2018 [AR15]

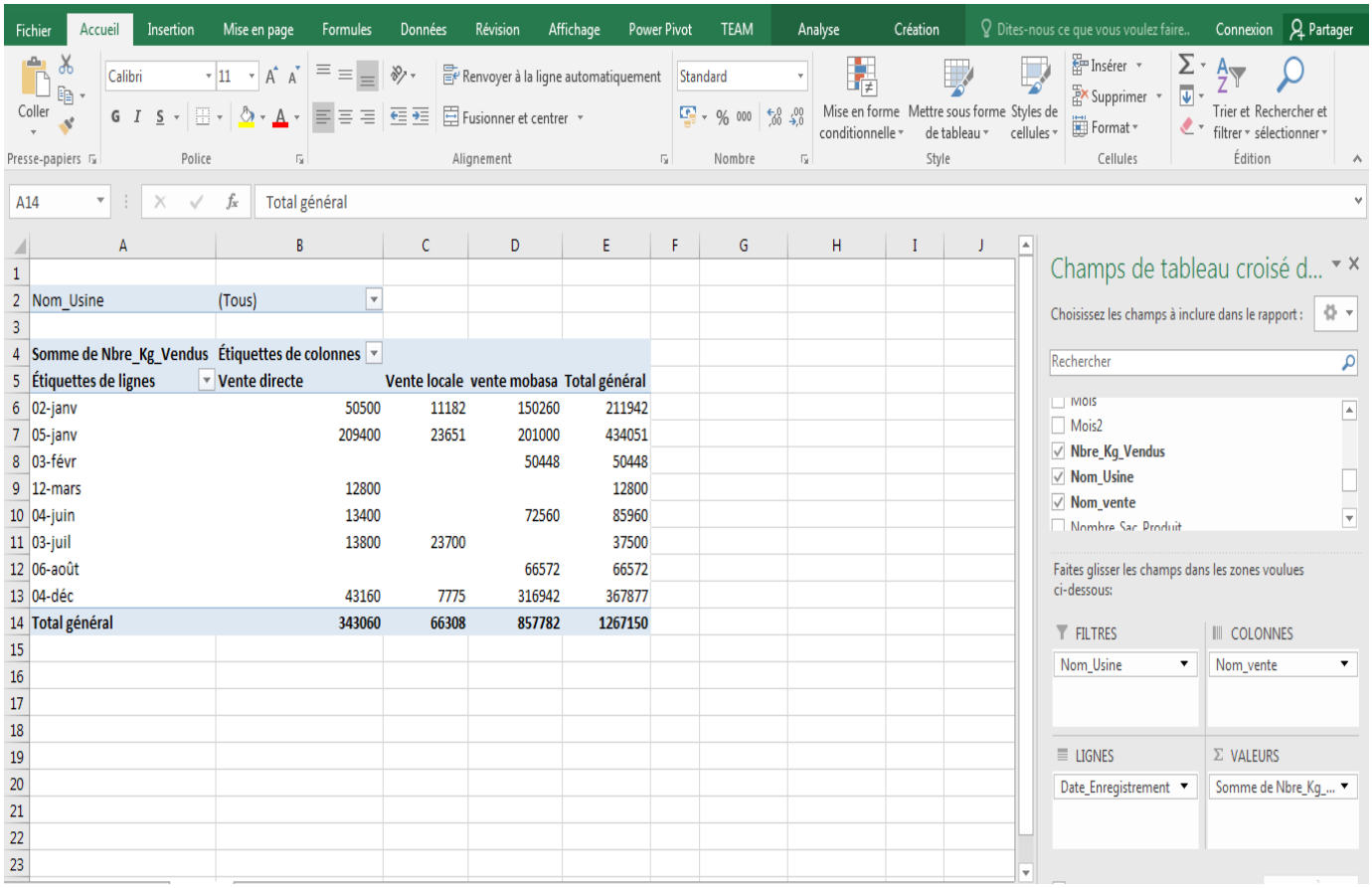
17. ;Février 2018 [AR18]

18. . Visité le 10 Avril 2017 à 23h00 ; [AR4]

19. https://www.youtube.com/watch?v=uzBbZP0NPCs/, Janvier 2017 [19] [AR19]

ANNEXE

Exemple d’une interface de navigation pour l’analyse des données



Nom_Usine	(Tous)				
Somme de Nbre_Kg_Vendus	**Étiquettes de colonnes**				
Étiquettes de lignes	**Vente directe**	**Vente locale**	**vente mobasa**	**Total général**	
02-janv	50500	11182	150260	211942	
05-janv	209400	23651	201000	434051	
03-févr			50448	50448	
12-mars	12800			12800	
04-juin	13400		72560	85960	
03-juil	13800	23700		37500	
06-août			66572	66572	
04-déc	43160	7775	316942	367877	
Total général	**343060**	**66308**	**857782**	**1267150**	