



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

**UNIVERSITE IBN KHALDOUN - TIARET**

# MEMOIRE

Présenté à :

FACULTÉ MATHÉMATIQUES ET INFORMATIQUE  
DÉPARTEMENT D'INFORMATIQUE

Pour l'obtention du diplôme de :

**MASTER**

Spécialité : Réseaux et Télécommunications

Par :

**MEBKHOUT Khaira Sabrina**

**&**

**HEROUAL Imen**

Sur le thème

---

## **Développement d'une application pour l'interrogation d'un entrepôt de données**

---

Soutenu publiquement le 8/ 11 /2020 à Tiaret devant le jury composé de :

Mr BENGHENI Abdelmalek

M.C.B Université IBN-KHALDOUN Tiaret

Président

Mr BENAOUA Habib

M.C.A Université IBN-KHALDOUN Tiaret

Encadreur

Mr DJAFRI Laouni

M.C.B Université IBN-KHALDOUN Tiaret

Examineur

**Année Universitaire : 2019 - 2020**



## **Remerciements**


*Nous remercions **Dieu** de nous avoir accordé des connaissances de la science et de nous avoir aidés à réaliser ce travail. Au terme de ce modeste travail*

*Nous tenons à exprimer nos plus vifs remerciements à **Bengheni Abdelmalek** pour avoir accepté de présider le jury.*

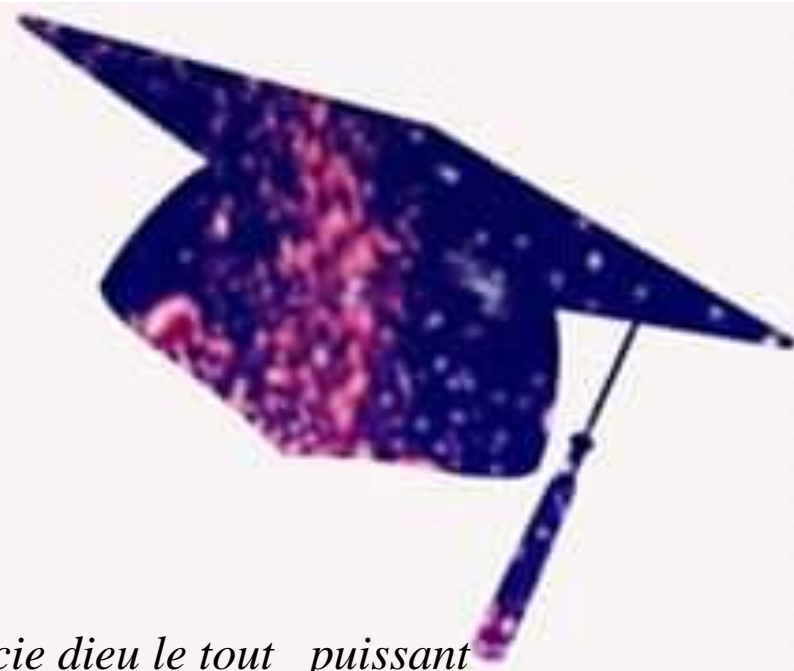
*Nos reconnaissances, et Nos sincères remerciements vont à Monsieur **Benaouda Habib** pour nous 'avoir dirigé tout au long de la réalisation de ce travail. Ses orientations, ses encouragements, sa compréhension, sa disponibilité constante m'ont été d'une précieuse aide.*

*Nous tenons à remercier également **Djafri Laouni** pour avoir accepté d'examiner ce travail et l'enrichir par ses propositions.*

*Et enfin, nos sincères remerciements et respects vont aux enseignants qui nous avons enseignés et qui par leurs compétences nous avons soutenu dans la poursuite de nos études.*



## *Dedicace*



*Avant tous, je remercie dieu le tout puissant de m'avoir donné le courage et la patience pour réaliser ce travail malgré toutes les difficultés rencontrées. Je dédie ce modeste travail*

*A mes chers parents et mes sœurs et mon frère, qui me toujours dirigée vers le chemin de la réussite, qui m'ont soutenue dans toutes mes aventures,*

*Qui me donne l'énergie d'avancer dans la bonne direction. Que Dieu leurs procure santé et longue vie.*

*Mes sœurs : Samira et Hiba pour leur aide*

*Mes amis : Djamila, Yassmine, Manel pour leurs soutiens et de précieux conseils.*

**Imen.H**



## *Dédicace*



*Grace a dieux le tout puissant, qui m'a donne la volonté et qui  
ma eclaire vers le bon chemin , Je tiens a dedier ce mémoire :*

*Ma mere . mon étoile, ma vie. qui a œuvré pour ma réussite, de par son  
amour, tous les sacrifices consentis et ses precieux conseils, pour  
toute son assistance et de présence dans ma vie , recois à travers ce  
travail aussi modeste soit-il ,l'expression de mes sentimments et de  
mon eternelle gratitude.*

*Mon père ,qui peut être fier de trouver ici le résultat de longues années  
de sacrefices pour m'aider à avancer dans la vie . Puisse difaire en  
sorte que ce travail porte son fruit , merci pour les valeurs nobles  
l'education et le soutient permanent venu de toi .*

*Mon frère saad et ma sœur Rahma qui n'ont cessé d'être pour moi  
des exemples de courage et de générosité , je vous souhaite un avenir  
plein de joie ,de bonheur et de reussite .*

*A tous mes chéres amies.*

*Merci du fond du mon cœur.*

**Sabrina.M**



## **Résumé**

Un entrepôt de données représente un dépôt central rempli d'informations pouvant être analysées dans le but de prendre des décisions plus éclairées. Les données stockées dans l'entrepôt de données proviennent de systèmes transactionnels de bases de données relationnelles et d'autres sources généralement à intervalle régulier. Les données et les analyses sont indispensables pour les entreprises qui souhaitent rester compétitives. Les analystes d'affaires, les spécialistes accèdent aux données via des outils de Business Intelligence comme les rapports, des tableaux de bord et des outils OLAP basés sur des requêtes Mdx pour exploiter leurs données, surveiller leurs performances et soutenir la prise de la bonne décision. Le présent travail a pour objectif d'implémenter une application qui permet d'interagir avec un entrepôt de données.

### **Mots clés**

Entrepôt de données, requête MDX, cube OLAP.

## **Abstract**

A data warehouse represents a central repository filled with information that can be analyzed in order to make more informed decisions. The data stored in the data warehouse comes from transactional systems relational databases and other sources usually at regular intervals. Data and analysis are essential for companies wishing to remain competitive Business analysts, specialists access data via Business Intelligence tools such as reports, dashboards and olap tools based on Mdx queries to use their data, monitor their performance and support the right decision. The purpose of this work is to implement an application that enables interaction with a data warehouse.

### **Keywords**

Data warehouse, MDX query, OLAP, Cube

## Liste des abréviations

- BI :** Business Intelligence
- OLAP:** On-Line Analytical Processing
- OLTP :** On-Line Transaction processing
- MDX :** Multi-Dimensionnelle eXpression
- ROLAP :** Relational Online Analytical Processing
- MOLAP :** Multidimensionnel Online Analytical Processing
- HOLAP :** Hybrid Online Analytical Processing
- ETL :** Extract Transform load

## Sommaire

<b>INTRODUCTION GENERALE.....</b>	<b>1</b>
<b>CHAPITRE 1 : L'ENTREPOT DE DONNEES CONCEPTS ET NOTION DE BASE.....</b>	<b>3</b>
Introduction.....	4
<b>1- La décision .....</b>	<b>4</b>
1.1 Le système décisionnel .....	4
1.2 Les Systèmes OLTP et OLAP .....	5
1.3 Les Application de L'Informatique Décisionnelle.....	6
1.4 Fonctions essentielles de l'informatique décisionnelle .....	6
<b>2- Entrepôt de donnée (Data Warehouse) .....</b>	<b>8</b>
2.1 Définition .....	8
2.2 Caractéristique d'un entrepôt de données .....	8
2.3 Structure des données de l'entrepôt.....	9
2.4 Objectifs d'un entrepôt de données.....	10
2.5 La différence entre une base de données et entrepôt de données .....	10
2.6 L'architectures d'entrepôts de données :.....	10
2.6.1 Approche théorique .....	10
2.6.2 Approche pratique.....	12
2.6.2.1 Les sources de données .....	12
2.6.2.2 L'entrepôt de données .....	12
2.6.2.3 Les magasins de données (datamarts) .....	13
2.6.2.4 Les outils d'analyses .....	13
2.7 Composants de base du data warehouse.....	13
2.7.1 Source de données .....	14
2.7.2 Zone de préparation des données.....	14
2.7.3 Zone de présentation.....	14
2.7.4 Zone de restitution .....	14
2.8 Alimentation d'un entrepôt de données .....	14
2.8.1 Processus ETL.....	15
2.9 Modélisation conceptuelle .....	15
2.9.1 Définitions : .....	15
2.9.2 Tables de faits : .....	16
2.9.3 Table de dimension .....	17
2.10 Modélisation d'un Data Warehouse.....	17
2.10.1 La modélisation en étoile .....	18
2.10.2 La modélisation en flocons .....	18
2.10.3 La modélisation en constellation.....	19
2.11 Modélisation logique .....	20
2.12 OLAP (Online Analytical Processing).....	20
2.12.1 Les différents types de systèmes [16].....	21
2.12.1.1 MOLAP : Multidimensionnel Online Analytical Processing .....	21
2.12.1.2 ROLAP : Relational Online Analytical Processing .....	21
2.12.1.3 HOLAP :Hybrid Online Analytical Processing .....	21

<b>Conclusion.....</b>	<b>22</b>
<b>CHAPITRE 2 : LES OUTILS D'INTERROGATION D'UN ENTREPOT DE DONNEES .....</b>	<b>23</b>
<b>Introduction .....</b>	<b>24</b>
<b>1. La phase d'interrogation d'un entrepôt de données .....</b>	<b>24</b>
<b>2. INTERROGATION ET ANALYSE : [18].....</b>	<b>24</b>
2.1 L'utilisation des techniques OLAP .....	25
2.1.1 Définition .....	25
2.1.2 S'utilise un cube OLAP ? .....	25
2.1.3 Les caractéristiques du cube OLAP .....	26
2.1.4 Catégories d'opérations.....	26
2.1.4.1 Opérations OLAP liées à la structure .....	26
2.1.4.2 Opérations OLAP liées à la granularité : .....	27
2.1.4.3 Les différents outils OLAP .....	30
2.1.5 Architecture Client / Serveur (Client Server Architecture) .....	30
2.1.6 Exploitation d'un cube Olap : .....	31
2.1.7 Les clients OLAP .....	31
2.1.7.1 Le web comme client OLAP .....	32
2.1.7.2 Les applications .....	33
2.1.8 Interrogation client/serveur .....	33
2.1.9 JIPIVOT client web .....	33
2.1.10 JRUBRIK application client .....	35
2.1.10.1 Création de requêtes OLAP sous JRUBRIK .....	35
2.1.11 FreeAnalysis.....	36
2.1.12 Palo : une excellente intégration à Excel.....	36
<b>2.2 Fouille de données (Data Mining) .....</b>	<b>37</b>
2.2.1 Définition .....	37
2.2.2 Les tâches du data mining .....	38
2.2.3 Les étapes du processus de data mining .....	38
<b>2.3 Outils reporting.....</b>	<b>38</b>
2.3.1 Définition .....	38
2.3.2 Principe.....	39
2.3.2.1 Le reporting ad hoc et de masse.....	39
2.3.2.1.1 Définition .....	39
<b>2.4 Outils de requêtes .....</b>	<b>39</b>
2.4.1 Les langages de données .....	40
2.4.2 Types de langages de données .....	40
2.4.3 Le langage de script de manipulation des données MDX.....	40
2.4.3.1 Définition .....	40
2.4.3.2 Caractéristiques de MDX : .....	40
2.4.3.3 Forme d'une requête OLAP sous MDX : .....	41
2.4.3.4 Syntaxe de requête MDX.....	41
2.4.3.5 Exemples:.....	41



2.4.3.6 Structure du résultat d'une requête MDX : .....	43
2.4.3.7 Comparaison entre SQL et MDX .....	44
2.4.3.8 Les avantages et inconvénients du langage MDX.....	44
<b>Conclusion.....</b>	<b>45</b>
<b>CHAPITRE 3 : <i>CONCEPTION, IMPLEMENTATION ET TEST</i>.....</b>	<b>46</b>
Introduction.....	47
<b>1. CONCEPTION .....</b>	<b>47</b>
1.1 Méthodologie de conception : .....	47
1.2 Les diagrammes des cas d'utilisation : .....	47
1.2.1 Composition du diagramme de cas d'utilisation .....	47
1.2.1.1 diagramme de cas d'utilisation de notre application ( <i>MdxQuery</i> ).....	48
1.3 Les diagramme de séquences.....	48
1.3.1 Composition d'un diagramme de séquence.....	48
1.3.2.1 diagramme de séquence d'authentification .....	49
1.3.2.2 diagramme de séquence de new query .....	49
1.3.2.3 diagramme de séquence de modifier une requête .....	50
1.3.2.4 Diagramme de séquence de supprimer une requête .....	51
1.3.2.5 diagramme de séquence de exécuter une requête .....	51
<b>2 L'ENVIRONNEMENT DE DEVELOPPEMENT :.....</b>	<b>52</b>
2.1 Mondrian .....	53
2.2 MySQL.....	53
2.3 JAVA .....	53
2.4 NetBeans .....	53
2.5 Tomcat.....	54
2.6 EasyPHP .....	54
2.7 Foodmart .....	54
3- Principales fenêtres de notre application .....	55
3.1 Fenêtre login membres .....	55
3.2 Fenêtre Welecome : .....	56
3.3 Fenêtre choose your data warehouse .....	56
3.4 Fenêtre select a cube.....	57
3.5 Fenêtre select dimension .....	58
3.6 Fenêtre select Levels .....	58
3.7 Fenêtre select membres .....	59
3.8 Fenêtre select Axes .....	60
3.9 Fenêtre generate query .....	60
Figure 3.34 Fenêtre generate query.....	60
3.10 Fenêtre d'exécution la requête .....	61
3.11 Fenêtre de my query .....	61
3.12 Fenêtre modifier une requête .....	62
3.13 Fenêtre renommer une requête .....	62
3.14 Fenêtre supprimer une requête .....	63
<b>Conclusion.....</b>	<b>63</b>

<b>CONCLUSION GENERALE .....</b>	<b>64</b>
<b>REFERENCES BIBLIOGRAPHIQUES .....</b>	<b>65</b>
<b>Webographie.....</b>	<b>66</b>

## Liste des figures

<b>Figure 1.1</b>	Processus global de traitement des données d'un système décisionnel [3] .....	7
<b>Figure 1. 2</b>	Approche de Bill Inmon [7] .....	11
<b>Figure 1.3</b>	Approche de Ralph Kimball [7] .....	11
<b>Figure 1.4</b>	Approche pratique du data warehouse [7] .....	12
<b>Figure 1-5</b>	Modélisation dimensionnelle (table de faits et dimension) exemple de vente [7]	13
<b>Figure 1.6</b>	Alimentation d'un DataWarehouse [10] .....	15
<b>Figure 1.7</b>	Modélisation dimensionnelle (table de faits et dimension) exemple de vente [3].	16
<b>Figure 1.8</b>	table de fait [3] .....	17
<b>Figure 1. 9</b>	Modèle en étoile exemple de vente [3] .....	18
<b>Figure 1.10</b>	modèle flocons [3] .....	19
<b>Figure 1.11</b>	Modèle en constellation [3] .....	19
<b>Figure 1.12</b>	Vue multidimensionnelle des données [7] .....	20
<b>Figure 2. 13</b>	Schéma représentent la phase restituée [32] .....	25
<b>Figure 2.14</b>	Architecture utilisée (client/serveur) [34] .....	33
<b>Figure 2.15</b>	Interface principale de JPivot [34] .....	34
<b>Figure 2.16</b>	Barre de commandes de JPIVOT [34] .....	34
<b>Figure 2.17</b>	Création d'une première requête OLAP en MDX sous JRUBIK.[35] .....	36
<b>Figure 2.18</b>	Data Mining [22] .....	37
<b>Figure 2.19</b>	Cube « Ventes »[36] .....	42
<b>Figure 2.20</b>	Résultat d'exécution requête d'exemple 2 [36] .....	43
<b>Figure 3.21</b>	Diagramme de cas d'utilisation de notre application .....	48
<b>Figure 3.22</b>	Diagramme de séquence d'authentification .....	49
<b>Figure 3.23</b>	Diagramme de séquence de création nouvelle requête .....	50
<b>Figure 3.24</b>	Diagramme de séquence de modifie une requête .....	50
<b>Figure 3.25</b>	Diagramme de séquence pour supprimer une requête .....	51
<b>Figure 3. 26</b>	Diagramme de séquence pour Exécuter une requête déjà existe .....	51
<b>Figure 3.27</b>	Fenêtre welecome .....	56
<b>Figure 3. 28</b>	Fenêtre choose your data warehouse .....	57
<b>Figure 3.29</b>	Fenêtre select a cube .....	57
<b>Figure 3.30</b>	Fenêtre select dimensions .....	58
<b>Figure 3.31</b>	Fenêtre select Levels .....	59
<b>Figure 3.32</b>	Fenêtre select members .....	59
<b>Figure 3. 33</b>	Fenêtre Select Axes .....	60
<b>Figure 3.34</b>	Fenêtre generate query .....	60
<b>Figure 3. 35</b>	Fenêtre d'exécution une nouvelle requête MDX .....	61
<b>Figure 3. 36</b>	Fenêtre my query .....	61
<b>Figure 3.37</b>	Fenêtre renommé le nom d'une requête .....	62
<b>Figure 3. 38</b>	Fenêtre supprimer une requête .....	63

## Liste des tableaux

<i>Tableau 1-1</i> base de donnees Vs entrepot de donnees [6] .....	10
<i>Tableau 1-2</i> comparatif de deux philosophies actuelles [7] .....	12
<i>Tableau 2-1</i> les Operations Olap liées à la structure [18].....	27
<i>Tableau2-2</i> Opérations OLAP liées à granularité [18] .....	28
<i>Tableau 2-3</i> résultat d'une requête MDX sur le cube 'vente' [36].....	43
<i>Tableau 2-4</i> différence entre requête SQL et une requête MDX [36] .....	44

## **Introduction générale**

Un des capitaux les plus importants d'une organisation sont les données qu'elle traite. Les applications qui manipulent ces données n'ont cessé de connaître des progrès permanents en termes de traitements, de stockage et de performances. Grâce aux bases de données exécutées des transactions journalières.

Ces applications utilisent des processus qui permettent de gérer des données variées et d'effectuer des traitements et transactions sur ces données.

Cependant, la plupart des entreprises disposent d'une masse d'information considérable, distribuées sur des systèmes opérationnels, ainsi, ces systèmes paraissent peu adaptés pour servir de support à la prise de décision. Nous nous posons donc la question suivante : comment procéder à une intégration décisionnelle réussite des données éparpillées et hétérogènes dans les bases de données opérationnelles ?

Face à cette inadéquation, l'informatique décisionnelle ou encore la Business intelligence (BI) est ainsi apparue durant les années 70, permettant d'exploiter de grandes quantités de données, manipulées dans un but d'analyse, afin d'en tirer les informations pertinentes et de définir des indications et des tableaux de bord qui facilitent la prise de décision.

Cependant, pour rendre cette analyse possible, les entreprises avaient besoin de nouvelles données agrégées, consolidées, historique et synthétisées selon plusieurs axes. Le besoin d'une nouvelle architecture, qui stocke et traite ce type de données s'est émergée du domaine de la BI, ce qui a donné naissance à l'architecture d'entrepôts de données (Data warehouse) et les datamarts (Magasin de données). Ces dernières sont des structures qui intègrent des données pour l'analyse et la représentation d'information pertinente.

Un entrepôt de données peut être vu comme une grosse base de données modélisée pour accueillir, après nettoyage et homogénéisation, les informations en provenance des différents systèmes de production de l'entreprise.

Il est vrai que la création de l'entrepôt n'est pas un objectif final. L'important, c'est son exploitation. La vocation de l'entrepôt de données est de supporter l'analyse des données. Ainsi, l'entrepôt de données constitue un support pour les outils d'analyse en ligne issus de la

technologie OLAP (On-Line Analytical Processing). Ce sont ces outils qui permettent d'accompagner les décideurs d'une entreprise en leur fournissant une possibilité de naviguer facilement dans les données

Ainsi l'entrepôt de données permet d'offrir des contextes d'analyses aux utilisateurs pour les aider dans leur prise de décision. Il s'avère alors que les utilisateurs peuvent avoir besoin de contextes d'analyses spécifiques, répondant à des besoins particuliers voir individuels d'où la personnalisation.

L'objectif de présent travail est de réaliser une application permettant à un utilisateur d'interrogé un entrepôt de données par le baie des requêtes MDX, et visualisée les résultats obtenus sous forme de tables croisées.

Notre mémoire est organisé selon le plan suivant :

Le premier chapitre est consacré, d'une part, aux notions et concepts de base pour les entrepôts de données, et de l'autre part, il décrit leurs méthodes de modélisation et d'implémentation.

Dans le deuxième chapitre, nous présentons les différents outils d'interrogation d'un entrepôt de données.

Dans le dernier chapitre, nous développons une application JAVA permettant à un utilisateur de formuler graphiquement des requêtes MDX, de l'exécuter sur un entrepôt de données de teste (FoodMart) et ensuite de visualiser graphiquement (table croisée) son résultat d'exécution.

**Chapitre 1 :**  
***L'entrepôt de données***  
***Concepts et notion***  
***De base***

## Introduction

La prise de décision est un processus important dans un organisme, et pour réussir cette procédure, il est important d'avoir les informations adéquates, c'est à ce moment que les systèmes d'informations décisionnels montrent leur efficacité. Avec le nombre d'informations que le système d'information doit stocker et analyser, les entrepôts de données ont fait leur apparition pour structurer les données, et aider les organismes à prendre les meilleures décisions.

Dans ce chapitre, nous allons définir les entrepôts de données, et nous allons présenter les notions et les concepts de base qu'ils caractérisent.

### 1- La décision

La décision peut être définie comme un acte de sélection d'une solution souhaitable pour un problème parmi une ou plusieurs alternatives disponibles pour guider une action choisie vers l'objectif désiré. Par ailleurs, quelque auteur (Carlisle, Stoner, Trewatha et Newport, Bedeian, Plunkett et Attner, Turban, Harrison) assimile la décision à un processus. Pour eux, la prise de décision est comme un processus de choix parmi divers plans d'action, dans le but de résoudre un problème ou atteindre une meilleure situation concernant les possibilités qui existent.

La décision ne peut être définie indépendamment de la notion de processus décisionnel. Prendre une décision, c'est mettre en œuvre un processus qui aboutira, à un choix. [1]

#### 1.1 Le système décisionnel

Un système décisionnel est un moyen qui a pour but de faciliter la définition et la mise en œuvre de stratégies gagnantes. Mais il ne s'agit pas de définir une stratégie une fois pour toute, mais d'être à même de continuellement s'adapter à son environnement, et de le faire plus vite que ses concurrents. Pour cela il convient de bien comprendre son environnement, d'ajuster ses interactions avec lui en faisant les meilleurs choix de cibles et d'actions.

Un système décisionnel va en particulier aider au pilotage des plans d'actions (prévision, planification, suivi), à l'apprentissage (acquisition de savoir faire, de connaissances, de compétences) et à la réalisation d'innovations incrémentales (adaptation du modèle d'affaires : produits/services, organisation, etc. ...). Les Systèmes décisionnels traditionnels permettent de faire l'analyse des activités déjà réalisées et d'en tirer des enseignements pour les activités futures, pour cela ils utilisent des données plus ou moins récentes (au Mieux mises à jour



quotidiennement). Les systèmes décisionnels plus avancés gèrent des données plus fraîches (certaines sont mises à jour en quasi temps réel), automatisent des décisions et supportent en temps réel des opérations (centre d'appels, web par exemple). [2]

Le système décisionnel a pour rôle de :

- Exploiter les données structurés et accumulées dans l'entreprise en les mettant à la disposition de tous,
- Sauvegarder et gérer les données dans le Data Warehouse.

## **1.2 Les Systèmes OLTP et OLAP**

Le terme OLTP (On Line Transaction processus) regroupe les concepts mis en place pour un système destiné à l'autorisation d'un processus de traitement de l'information. On utilise pour cela les bases de données transactionnelles, les fichiers, dont la mise à jour est faite en temps réel. Ces bases de données et fichiers sont utilisés dans les entreprises pour gérer les importants volumes d'informations contenus dans leurs systèmes opérationnels (ou transactionnel). Les systèmes transactionnels se caractérisent de la manière suivante :

- Ils sont nombreux au sein d'une entreprise,
- Ils concernent essentiellement la mise à jour des données,
- Ils traitent un nombre d'enregistrements réduit,
- Ils sont définis et exécutés par de nombreux utilisateurs.

De part leurs caractéristiques, les processus transactionnels en ligne regorgent une grande masse d'informations cachant une certaine connaissance dont eux-mêmes sont peu adaptés à pouvoir déceler. Pour remédier à ceci, il faudrait torturer cette masse d'informations jusqu'à ce qu'elles aillent avouer la connaissance qu'elles détiennent.

C'est dans cette perspective que plus de 95% d'entreprises européennes se sont donc mises à la recherche de systèmes supportant efficacement les applications d'aide à la décision à partir de la connaissance extraite dans les systèmes transactionnels. Ces applications décisionnelles utilisent des processus d'analyse en ligne de données (OLAP : "On-Line Analytical Processing". Ces processus OLAP se caractérisent de la manière suivante :

- Ils sont peu nombreux, mais leurs données et traitements sont complexes,
- Il s'agit uniquement de traitements semi-automatiques visant à interroger, visualiser et synthétiser les données,
- Ils concernent un nombre d'enregistrements importants aux structures hétérogènes,
- Ils sont définis et mis en œuvre par un nombre réduit d'utilisateurs qui sont les décideurs.

Une autre différence que nous pouvons établir entre les deux systèmes est que les OLTP sont orientés « Applications », ce qui fait que nous pouvons parler de la gestion de ressources humaines, des abonnés, de vente de cartes prépayées, alors que les systèmes OLAP sont orientés « Sujets ou thèmes ou encore Activités » et par exemple nous pouvons avoir comme sujet : la performance des employés, la vente, ... [2]

### **1.3 Les Application de L'Informatique Décisionnelle**

Ces applications :

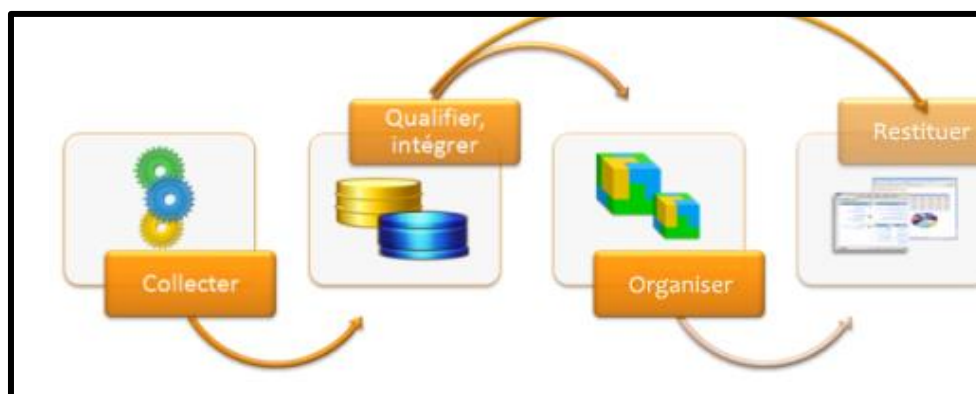
- Gestion de la relation client,
- Gestion de commandes,
- Gestion de stocks, les Prévisions de ventes,
- Définition de profil utilisateur,
- Analyse de transactions bancaires et la Détection de fraudes.

Les nouveaux outils de BI «business intelligence » facilitent la mise au point d'indicateurs fédérant toutes les couches de l'entreprise. La vision stratégique de la direction générale doit en effet être diffusée et déclinée dans tous les services de l'entreprise, il s'agit de fixer à chacun des objectifs qui lui « parlent ». Parce qu'un chef d'équipe sur une chaîne de montage se sentira infiniment plus concerné par un objectif de réduction des délais de mise en route de la chaîne ou par un ratio de défauts, que par un objectif d'accroissement de part de marché au plan mondial.[2]

### **1.4 Fonctions essentielles de l'informatique décisionnelle**

Un système d'information décisionnel assure quatre fonctions fondamentales, à savoir la collecte, l'intégration, la diffusion et la présentation des données à ces quatre fonctions s'ajoute une fonction de contrôle du système d'information décisionnelle [3]

- Collecte : La collecte est l'ensemble des tâches consistant à détecter, à sélectionner, à extraire et à filtrer les données brutes issues des environnements pertinentes compte tenu du périmètre du système d'information,
- Intégration : L'intégration consiste à concentrer les données collectées dans un espace unifié, dont le socle informatique essentiel est l'entrepôt. Élément centrale du dispositif, il permet aux applications décisionnelles de bénéficier d'une source d'information commune, homogène, normalisée et fiable, susceptible de masquer la diversité de l'origine des données.
- Diffusion ou la distribution : La diffusion met les données à la disposition des utilisateurs, selon des schémas correspondant au profil ou au métier de chacun, sachant que l'accès direct à l'entrepôt ne correspondrait généralement pas aux besoins d'un décideur ou d'une analyse.
- Présentation : Cette quatrième fonction, la plus visible pour l'utilisateur, régit les conditions d'accès de l'utilisateur aux informations. Elle assure le fonctionnement du poste de travail, le contrôle d'accès, la prise en charge des requêtes, la visualisation des résultats sous une forme ou une autre. Elle utilise les techniques de communication possibles comme les outils bureautiques, raquetteurs et générateurs d'états spécialisés, infrastructure web, télécommunications mobiles, etc.
- Administration : C'est la fonction transversale qui supervise la bonne exécution de toutes les autres, elle pilote le processus de mise à jour des données, la documentation sur les données



**Figure 1.1** Processus global de traitement des données d'un système décisionnel [3]

## 2- Entrepôt de donnée (Data Warehouse)

Selon **bill Inmon** « Un Data Warehouse ne s'achète pas, il se construit »

Un entrepôt de données est une base de données regroupant une partie ou l'ensemble des données fonctionnelles d'une entreprise. Il entre dans le cadre de l'informatique décisionnelle, son but est de fournir un ensemble de données servant de référence unique, utilisée pour la prise de décisions dans l'entreprise par le biais de statistiques et de rapports réalisés via des outils de reporting. D'un point de vue technique, il sert surtout à 'délester' les bases de données opérationnelles des requêtes pouvant nuire à leurs performances.

### 2.1 Définition

« Entrepôt de Données : système central pour la prise de décision, Un entrepôt est une collection de données orientées sujet, intégrées, non volatiles et historiées, organisées pour le support d'un processus d'aide à la décision » **Bill Inmon**,

Comme le Père du Data Warehouse le définit :

« L'entrepôt de données est la ressource de présentation interrogeable des données d'une entreprise et elle ne doit pas être organisée autour d'un modèle entité relation, qui lui ferait perdre sa clarté et ses performances, ces source de données interrogeable de l'entreprise. C'est l'union des DataMarts qui le composent » **Kimball**.

### 2.2 Caractéristique d'un entrepôt de données

Un data warehouse est une base de données organisée et structurée dans le cadre décisionnel. Elle n'est pas une copie des bases de production. Un data warehouse doit répondre aux quatre principes suivant selon, **Bill Inmon** :[3]

- **Orienté sujet** : Les bases de production sont, généralement, organisées par processus fonctionnels (facturation, comptabilité, RH etc.). Au cœur du Data warehouse, les données sont organisées par thème, l'intérêt dans cette organisation est de disposer d'un ensemble d'informations utiles sur un sujet transversal aux structures fonctionnelles et organisationnelles de l'entreprise,
- **Intégré** : Les données proviennent de plusieurs sources différentes. Avant d'être intégrées au sein du data warehouse elles doivent être mise en forme et unifiées afin

d'en assurer la cohérence. Cela nécessite une forte normalisation, de bénéficier d'un référentiel unique et cohérent ainsi que de bonnes règles de gestion,

- **Non volatile** : Un data warehouse veut conserver la traçabilité des informations et des décisions prises. Les données ne sont ni modifiées ni supprimées, et le rafraîchissement de l'entrepôt de données, consiste seulement à ajouter de nouvelles données sans modifier ou perdre celle qui existent. Ceci pour conserver la traçabilité des informations et des décisions prises,
- **Histories** : Contrairement au système de production, les données ne sont jamais mises à jour. Chaque nouvelle donnée est insérée. Un référentiel de temps doit être mis en place afin de pouvoir identifier chaque donnée dans le temps, un référentiel de temps est associé aux données, afin de permettre l'identification dans la durée de valeurs.

### 2.3 Structure des données de l'entrepôt

Les données de l'entrepôt sont structurées en quatre classes. Ces dernières sont organisées selon un axe historique et un axe de synthèse :[4]

- **Les données agrégées** : Ce sont les données qui correspondent à des éléments d'analyse représentant les besoins des utilisateurs,
- **Les données détaillées** : Reflètent les événements les plus récents. Les intégrations régulières des données issues des systèmes de production sont réalisées habituellement à ce niveau,
- **Les données historisées** : Chaque nouvelle insertion de données provenant du système de production ne détruit pas les anciennes valeurs, mais crée une nouvelle occurrence des données,
- **Les métadonnées** : Elles constituent l'ensemble des données qui décrivent des règles ou processus attachés à d'autres données, qui représente la finalité du système d'information.

## 2.4 Objectifs d'un entrepôt de données

- Les données de l'organisation doivent être accessible facilement (données parlantes, et signification évidente pas seulement pour le développeur mais surtout pour l'utilisateur),
- La présentation des informations de manière cohérente : les données doivent être assemblées à partir des différentes sources de l'entreprise et il faut contrôler la qualité et éviter de les mettre à disposition quand elles ne sont pas nettoyées. Attention si deux mesures sont identiques, elles doivent porter le même nom pour éviter toute confusion et inversement ne pas mettre le même nom pour deux mesures qui ne sont pas calculées,
- Il doit être adaptable,
- Il doit efficacement protégées les informations de l'entreprise,
- Il doit être utilisé lors de la prise de décision et par l'ensemble de la communauté,
- Il doit servir de stockage d'information. [5]

## 2.5 La différence entre une base de données et entrepôt de données

L'objectif premier d'un entrepôt de données est de stocker les données pertinentes aux besoins de prise de décision. Contrairement aux bases de données opérationnelles qui sont conçues pour supporter des opérations journalières, un entrepôt est conçu pour supporter des opérations d'analyse utile à la prise de décision. Le tableau suivant (Tableau 1.1) récapitule les principales différences entre les bases de données opérationnelles et les entrepôts de données. [6]

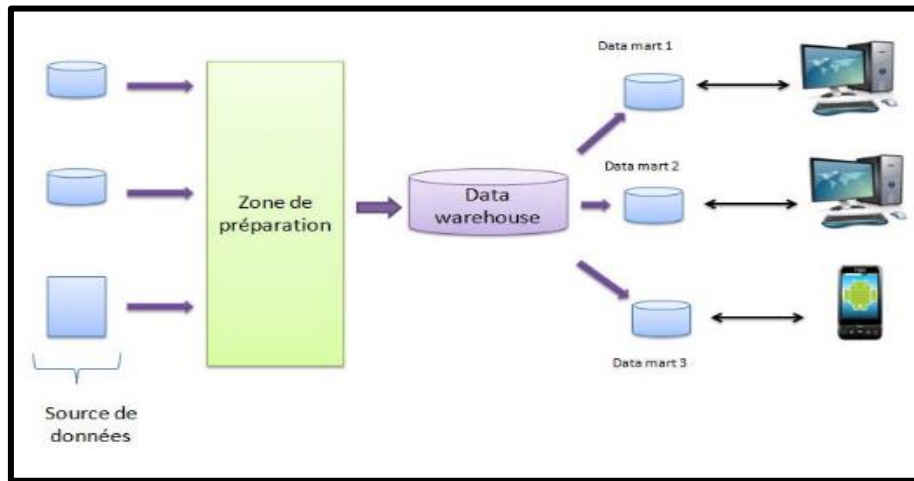
Caractéristiques	Base de données	Entrepôts de données
<b>But</b>	Exécution d'un processus métier	Exécution d'un processus métier
<b>Usage</b>	Support à la gestion courante	Support à la prise de décision
<b>Principe de Conception</b>	Troisième forme normale	Conception multidimensionnelle
<b>Données</b>	Actuelles, brutes	Historiques, agrégées
<b>Opérations</b>	Lecture et écriture	Lecture et rafraîchissement
<b>Utilisateurs</b>	Employé	Analyste et décideur
<b>Taille</b>	Des giga-octets	Plutôt des téra-octets

Tableau 1-1 base de données Vs entrepot de données [6]

## 2.6 L'architectures d'entrepôts de données :

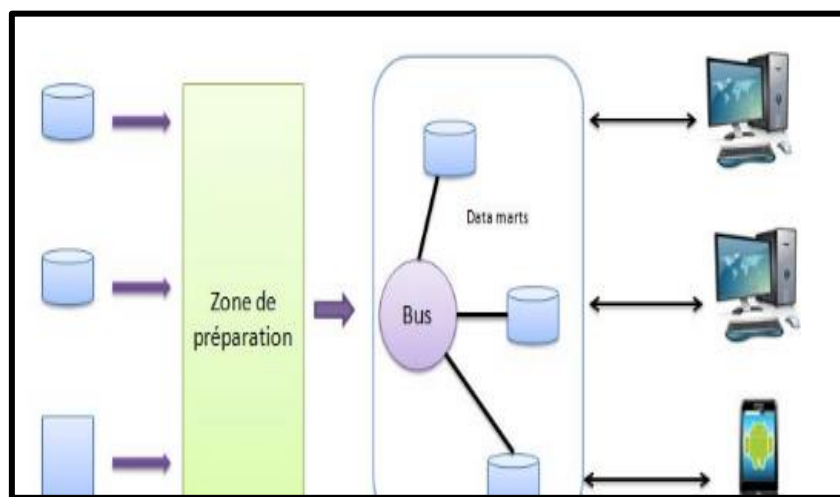
### 2.6.1 Approche théorique

L'architecture de haut en bas: selon **Bill Inmon**, l'entrepôt de données est une base de données au niveau détail, consistant en un référentiel global et centralisé de l'entreprise. En cela, il se distingue du DataMart, qui regroupe, agrège et cible fonctionnellement les données. [7]



**Figure1. 2** Approche de Bill Inmon [7]

- L'architecture de bas en haut: Pour **Ralph Kimball**, l'entrepôt de données est constitué peu à peu par les Data Mart de l'entreprise, regroupant ainsi différents niveaux d'agrégation et d'historisation de données au sein d'une même base.



**Figure 1.3** Approche de Ralph Kimball [7]

	Ralph Kimball	Bill Inmon
<b>Processus</b>	Bottom-up	Top-down
<b>Organisation</b>	Data mart	Data warehouse
<b>Schématisation</b>	Etoile	Flocon

Tableau 1-2 comparatif de deux philosophies actuelles [7]

### 2.6.2 Approche pratique

Le schéma suivant illustre l'architecture générique d'un système data warehouse [7]

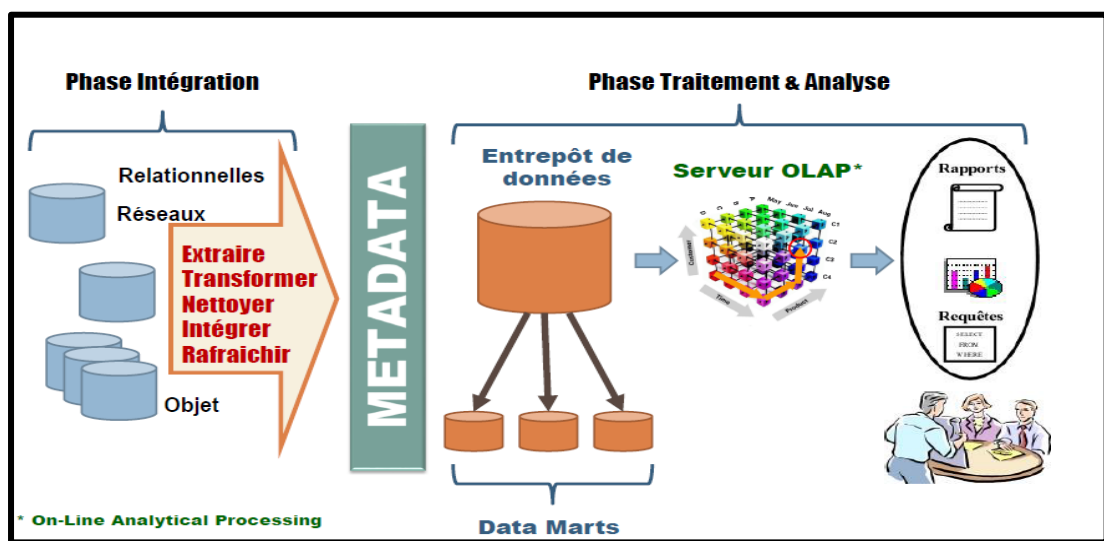


Figure 1.4 Approche pratique du data warehouse [7]

L'architecture des systèmes décisionnels met en jeu quatre éléments essentiels : les sources de données, l'entrepôt de données, les magasins de données et les outils d'analyse et d'interrogation

#### 2.6.2.1 Les sources de données

Sont nombreuses, variées, distribuées et autonomes. Elles peuvent être internes (bases de production) ou externes (internet) à l'entreprise.

#### 2.6.2.2 L'entrepôt de données

Est le lieu de stockage centralisé des informations utiles pour les décideurs.



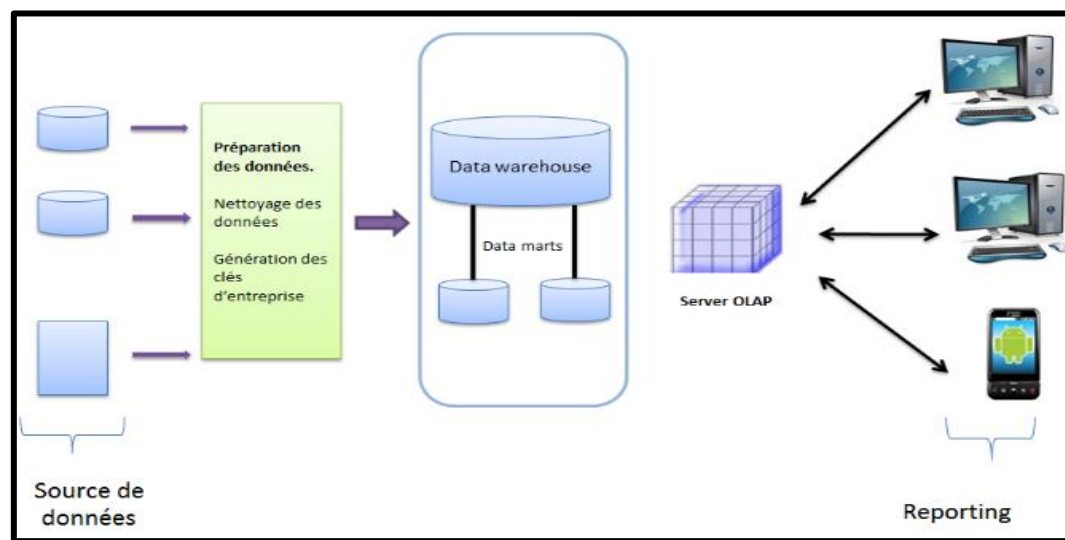
### 2.6.2.3 Les magasins de données (datamarts)

Est une vue partielle du data warehouse mais orientée sujet. C'est un sous-ensemble du data warehouse contenant des informations se rapportant à un secteur d'activité particulier de l'entreprise ou à un métier qui y est exercé. Il se situe en aval du data warehouse et est alimenté par celui-ci. On peut donc créer plusieurs datamart correspondant au différent besoin de l'utilisateur cela permet de réduire le nombre d'opération sur les bases de production. De plus cela permet d'offrir aux utilisateurs un outil spécifiquement adapté à leurs besoins. Cet outil sera plus petit et permettra donc un accès plus rapide à l'information. [8]

### 2.6.2.4 Les outils d'analyses

Permettent de manipuler les données suivant des axes d'analyses, l'information est visualisée à travers d'interface interactives et fonctionnelles dédiées à des décideurs souvent non informaticiens (directeurs).

## 2.7 Composants de base du data warehouse



**Figure 1-5** Modélisation dimensionnelle (table de faits et dimension) exemple de vente [7]

### 2.7.1 Source de données

L'entrepôt de données stocke des données qui proviennent de différentes sources d'informations hétérogènes et distribuées. Ces sources peuvent être des bases de données, des fichiers de données, des sources externes à l'entreprise etc. [5]

### 2.7.2 Zone de préparation des données

Cette partie est une zone de stockage, d'un ensemble de processus appelé ETC ou ETL (extract/transform/load). Cette zone permet de récupérer les données depuis les applications opérationnelles, puis de les nettoyer (résolution de conflit, correction, conversion en format standard), puis de les combiner pour les envoyer dans la partie de présentation des données. Cette zone ne doit pas être accessible aux requêtes.

### 2.7.3 Zone de présentation

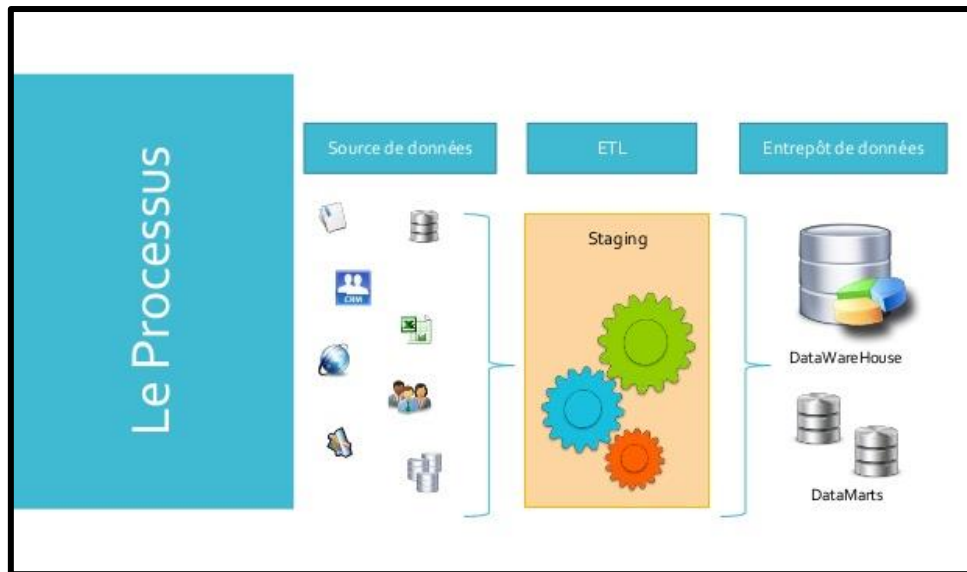
Cette zone est le lieu où toutes les données sont stockées et offertes aux requêtes des utilisateurs. Les données doivent être présentées, stockées et consultées sous forme de schémas dimensionnels. La modélisation normalisée apporte des gains de performances dans les traitements opérationnels mais les modèles normalisés ne sont pas adaptés aux requêtes d'entrepôts.

### 2.7.4 Zone de restitution

L'ensemble des outils qui permettent aux utilisateurs d'exploiter les données, il représente ce que voient les utilisateurs.

## 2.8 Alimentation d'un entrepôt de données

L'alimentation d'un entrepôt de données est une phase essentielle dans le processus d'entreposage. Elle se déroule en plusieurs étapes : extraction, transformation, chargement et rafraîchissement des données, qui sont prises en charge par le processus d'ETL (Extracting, Transforming and Loading). Ce processus constitue la phase de migration des données de production [10]



**Figure 1.6** Alimentation d'un DataWarehouse [10]

### 2.8.1 Processus ETL

Sont les termes (Extract, transform, load) se traduisant par (Extraction, Transformation, Chargement), c'est le processus qui permet de charger un DataWarehouse à partir de données externes généralement issues de bases transactionnelles. Son rôle est de récupérer ces données et de les traiter pour qu'elles correspondent aux besoins du modèle dimensionnel. En général les données sources doivent être "nettoyées" et aménagées pour être exploitables par les outils décisionnels. [11]

➤ Exemples de cas d'usage pour les outils ETL :

- Migrer des données d'une application à une autre,
- Stocker les données dans un Data Warehouse avant de les importer, les trier et les transformer dans un environnement de business intelligence,
- Synchronisation des systèmes critiques.

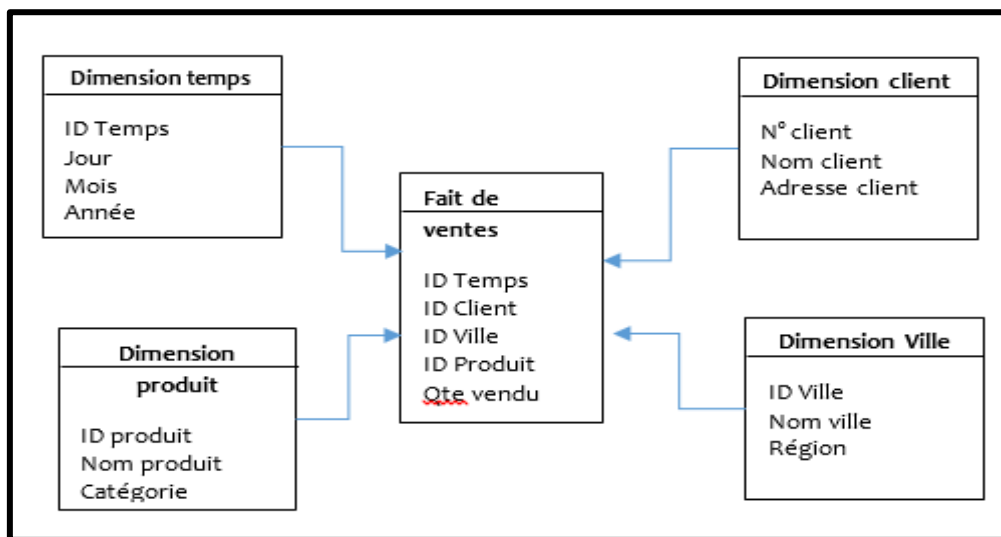
## 2.9 Modélisation conceptuelle

### 2.9.1 Définitions :

Technique de conception logique permettant de structurer les données de manière à les rendre intuitives aux utilisateurs et offrir une bonne performance aux requêtes.[13]

Modélisation dimensionnelle permet cela : elle consiste à considérer un sujet d'analyse comme un cube à plusieurs dimensions, offrant des vues en tranches ou des analyses selon différents axes.

La modélisation des bases de données relationnelles utilise les concepts d'entités et de relations afin de construire des tables. En business intelligence, la modélisation d'un data warehouse (entrepôt de données) utilise les notions de **table de faits** et **table de dimension**. [3]



**Figure 1.7** Modélisation dimensionnelle (table de faits et dimension) exemple de vente [3]

### 2.9.2 Tables de faits :

Ce que l'on souhaite mesurer Quantités vendues, montant des ventes [14] .

Une table de faits est la table centrale du modèle dimensionnel. Elle contient les informations observables (les mesures) sur ce qu'on veut analyser : table de faits des ventes par exemple.

Une ligne d'une table de faits correspond à une mesure. Ces mesures sont généralement des valeurs numériques, additives, cependant des mesures textuelles peuvent exister mais sont rares.

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

Fait de ventes
ID Temps
ID Client
ID Ville
ID Produit
<u>Qte vendu</u>
Montant Facture

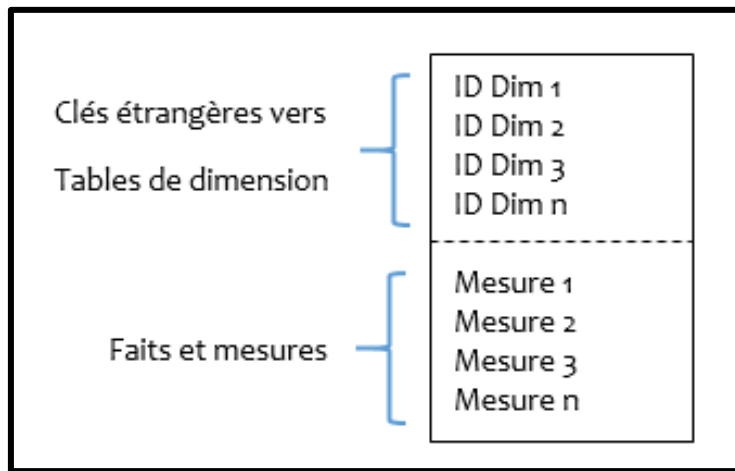


Figure 1.8 table de fait [3]

Il existe trois (3) types de faits : [14]

- **Additif** : additionnable suivant toutes les dimensions,
- **Semi additif** : additionnable suivant certaines dimensions,
- **Non additif** : fait non additionnable quelque soit la dimension.

### 2.9.3 Table de dimension

Les tables de dimension sont les tables qui accompagnent une table de faits, elles contiennent les descriptions textuelles de l'activité. Une table de dimension est constituée de nombreuses colonnes qui décrivent une ligne. C'est grâce à cette table que l'entrepôt de données est compréhensible et utilisable, Une dimension est généralement constituée : d'une clé naturelle et des attributs.

### 2.10 Modélisation d'un Data Warehouse

Dans un Data Warehouse les données et leurs relations sont organisées suivant un modèle de données spécifique. Le choix du modèle de données structure et définit le design du Data Warehouse. Nous avons vu qu'il existait trois modélisations possibles :

- La modélisation en étoile,
- La modélisation en flocons,
- La modélisation en constellation.

### 2.10.1 La modélisation en étoile

Le modèle de données en étoile doit son nom à sa forme. Ce modèle de conception privilégie l'approche utilisateur, l'orientation métier. La table de référence contient les faits. Les faits ou mesures sont les données chiffrées (du type résultats par secteur). Les tables satellites correspondent aux dimensions. Ce sont les axes d'analyse des utilisateurs. Ainsi, dès la conception, la base est orientée en tenant compte des types de recherche des utilisateurs.[15]

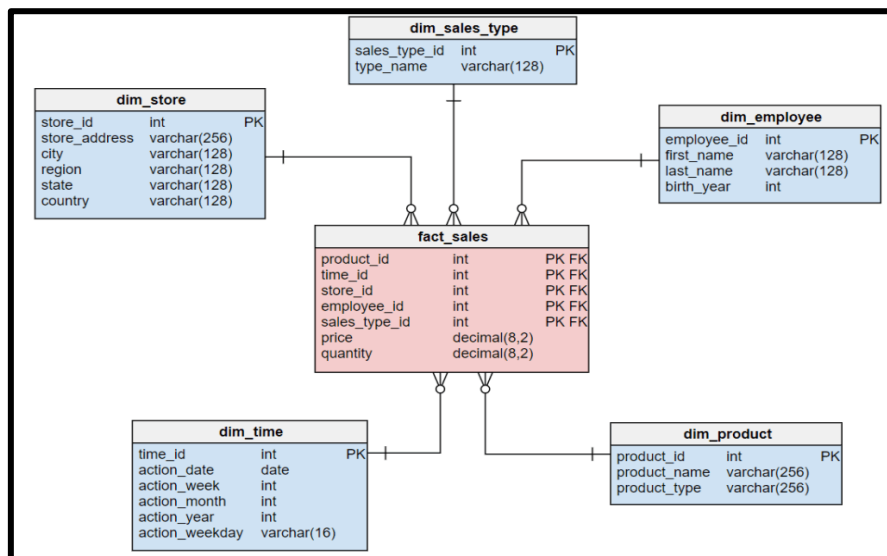


Figure1. 9 Modèle en étoile exemple de vente [3]

### 2.10.2 La modélisation en flocons

Un modèle en flocon est un modèle pour lequel chaque dimension est représentée avec plusieurs tables.[9]

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, cependant elle peut s'avérer moins compréhensible pour l'utilisateur final, et très couteux en termes de performances. [3]

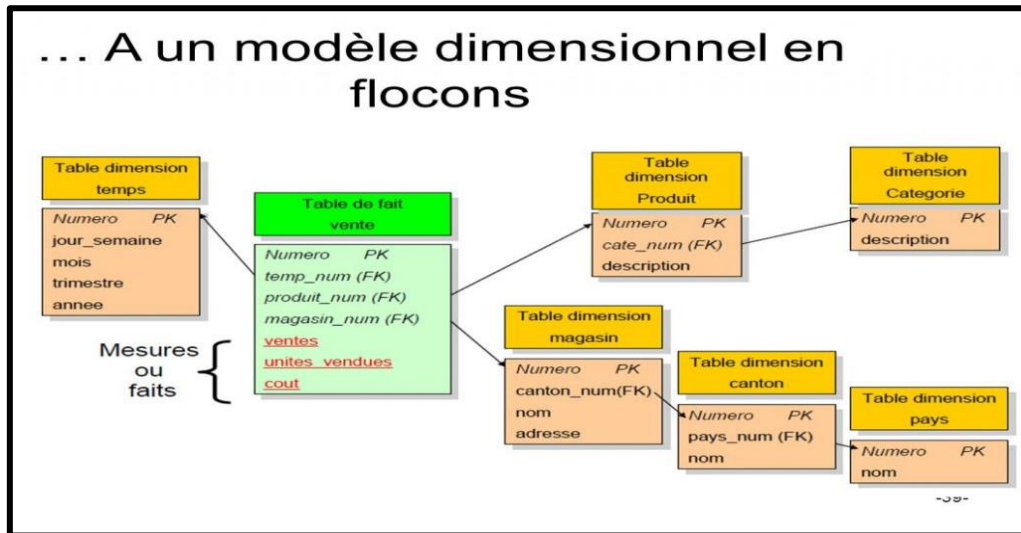


Figure 1.10 modèle flocons [3]

### 2.10.3 La modélisation en constellation

Ce n'est rien d'autre que plusieurs modèles en étoile liés entre eux par des dimensions communes.

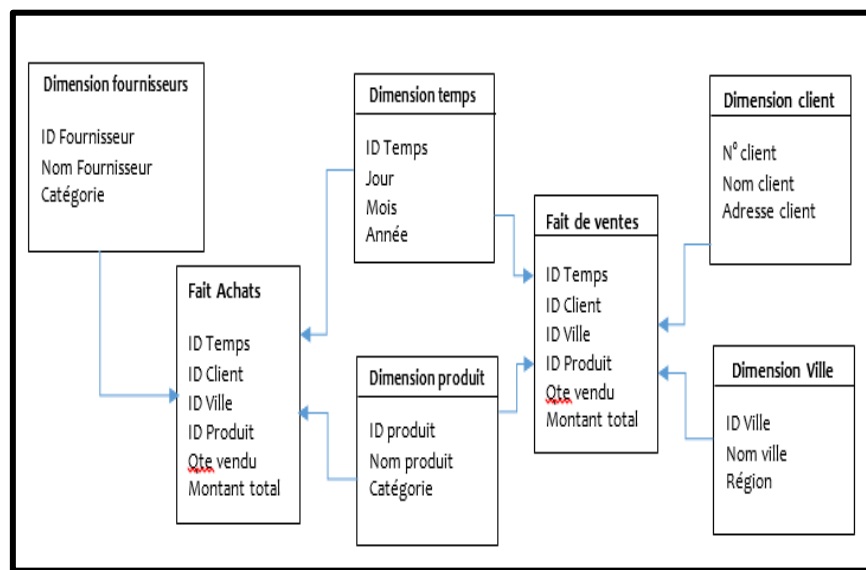


Figure 1.11 Modèle en constellation [3]

## 2.11 Modélisation logique

Au niveau logique plusieurs possibilités sont envisageables pour la modélisation multidimensionnelle. Il est possible d'utiliser :

- Un système de gestion de bases de données (SGBD) existant tel que les SGBD relationnels (ROLAP) ou bien les SGBD orientées objet (OOLAP),
- Un système de gestion de bases de données multidimensionnelles (MOLAP).[10]

## 2.12 OLAP (Online Analytical Processing)

Un cube OLAP est une base de données multidimensionnelle optimisée pour les Data Warehouses et les applications OLAP. Il s'agit d'une méthode permettant de stocker les données sous forme multidimensionnelle. En général, ces cubes sont pré-résumés pour accélérer le temps de requête par rapport aux bases de données relationnelles.

Pour effectuer des requêtes au sein des cubes OLAP, on utilise le langage **MDX (multidimensional expressions)**. Ce langage fut développé par Microsoft à la fin des années 1990 avant d'être adopté par les autres vendeurs de bases de données multidimensionnelles.

[1]

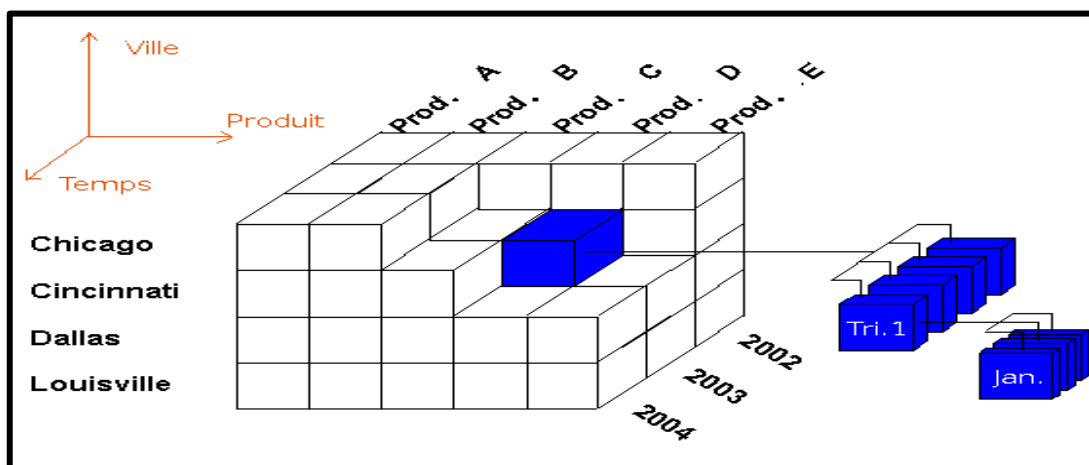


Figure 1.12 Vue multidimensionnelle des données [7]



## 2.12.1 Les différents types de systèmes [16]

### 2.12.1.1 MOLAP : Multidimensionnel Online Analytical Processing

C'est de l'OLAP Multidimensionnel. On stock les données dans un cube qui est en fait une base de données multidimensionnelles. De cette façon, le concept de relationnel disparaît. Tous les croisements envisageables et possibles sont pré calculés. C'est l'objectif principal de cette base de données multidimensionnelle car cela permet la restitution des données de façon instantanée. Les données étant stockées, le temps gagné pendant la restitution des données sera considérable.

➤ **Les inconvénients des cubes MOLAP :**

Le coût est important, en effet, elle nécessite souvent des licences pour les bases multidimensionnelles et des coûts pour le développement des cubes.

➤ **L'avantage des cubes MOLAP :**

Le temps de réponse est extrêmement rapide car la totalité des données sont stockées dans un cube, c'est à dire que les combinaisons sont pré-calculées durant le traitement du cube.

### 2.12.1.2 ROLAP : Relational Online Analytical Processing

Permettent d'accéder directement aux données stockées dans les bases de données relationnelles. Ainsi, les entreprises peuvent directement exploiter leurs SGBDR déjà en place. On effectue les requêtes avec un produit ROLAP en utilisant le langage SQL.[17]

➤ **Inconvénients de ROLAP :**

Les temps de réponse sont conséquents étant donné que les requêtes lisent les résultats directement depuis des tables relationnelles. Les bases sont donc utilisées à chaque exécution du rapport, sauf si le cache est activé. De plus les requêtes sont plus complexes, car l'agrégation et les formules se font à la volée.

➤ **Les avantages de ROLAP :**

Le coût est relativement faible, en effet, cette méthode utilise des ressources qui sont disponibles comme un serveur SQL. De plus les données sont toujours disponibles sans traitement pour pré-calculer et agréger les données.

### 2.12.1.3 HOLAP :Hybrid Online Analytical Processing

Combinent les meilleures fonctionnalités du MOLAP et du ROLAP dans une seule architecture. Ainsi, ces produits corrigent les inconvénients de ces deux types de produits. Ils peuvent être utilisés aussi bien sur une base de données multidimensionnelle que sur une base de données relationnelle.

- **Les inconvénients de HOLAP**

Il est difficile d'utiliser la méthode Hybride OLAP dans le cas de complexité élevée des rapports ou qu'ils fassent appel à trop de croisements de données.

- **Les avantages de HOLAP**

Un investissement financier moindre que la méthode MOLAP, en effet la partie développement sera beaucoup moins importante. De plus le temps de réponse est relativement court.

## **Conclusion**

Les entrepôts de données aident beaucoup à gérer une quantité importante de données, et c'est un outil central dans la prise de décision des organismes.

Dans ce chapitre nous avons évoqué les principaux concepts liés à la conception et l'exploitation des entrepôts de données où nous avons vu les caractéristiques des données de l'entrepôt et nous avons présenté l'architecture et les composants d'entrepôt de données.

Dans le chapitre suivant nous décrirons plus en détail les différents outils d'interrogation des entrepôts de données.

## **Chapitre 2 :**

***Les outils d'interrogation***

***D'un entrepôt***

***De données***

## Introduction

Une fois que les données sont traitées et chargées dans le data warehouse, il convient de les mettre à la disposition des utilisateurs finaux sous une forme synthétique et exploitable. On utilise pour cela des outils de reporting. Ces outils vont se connecter au data warehouse et proposer des interfaces et processus pour faciliter la création et la distribution d'analyses visuelles avec tableaux, graphes.

Dans ce cadre on va parler et détailler ces différents outils d'interrogation avec l'entrepôt de données.

### 1. La phase d'interrogation d'un entrepôt de données

La mise en œuvre du Data Warehouse est la dernière étape d'un projet d'entrepôt de données, soit son exploitation. L'interrogation des entrepôts de données est une tâche cruciale, qui permet de naviguer dans le cube de données afin d'exploiter les résultats obtenus. Il s'agit de manipuler des données. Il existe divers façons de naviguer au sein d'un cube OLAP afin de récupérer une information, pour se faire, un certain nombre d'outils sont disponibles, souvent spécifiques aux éditeurs du cube, qui offrent la possibilité de naviguer facilement au sein de celui-ci.

### 2. Interrogation et Analyse : [18]

L'exploitation de l'entrepôt, pour l'aide à la décision peut se faire de différentes façons, dont :

- L'interrogation à travers un langage de requêtes,
- La connexion à des composants de report,
- L'utilisation des techniques OLAP (OnLineAnalyticalProcess ),
- L'utilisation des techniques de fouille de données (Data Mining).

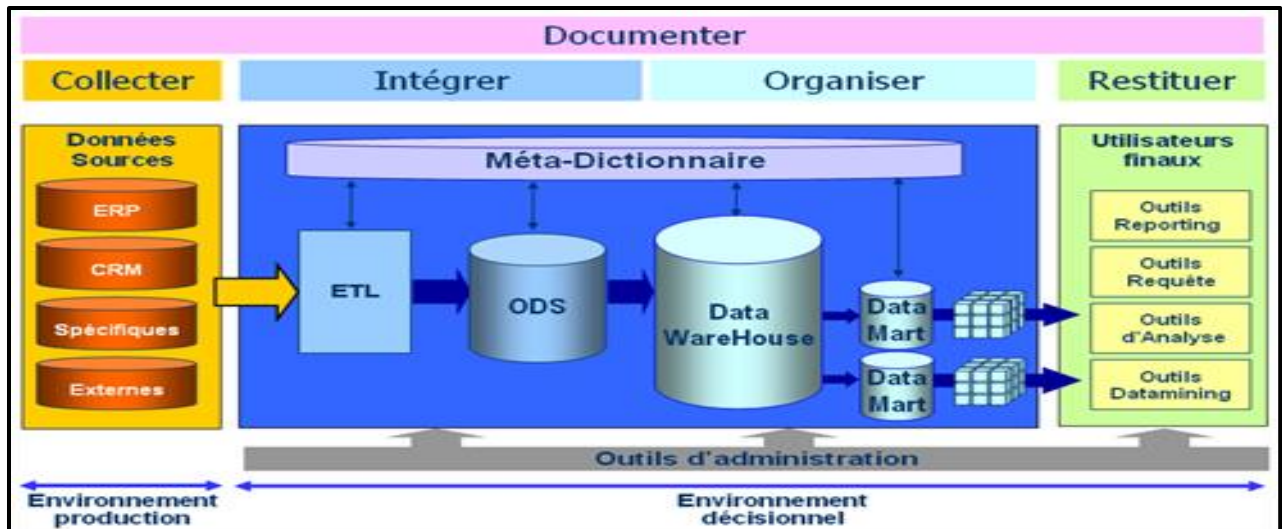


Figure2. 13 Schéma représentant la phase restituée [32]

## 2.1 L'utilisation des techniques OLAP :

### 2.1.1 Définition

Un cube OLAP est, avant toute chose, une base de données multidimensionnelle, dans laquelle les données sont pré-calculées. L'utilisateur peut ainsi en interroger un volume important sans impacter le temps de réponse qui est, la plupart du temps, instantané.

L'utilisateur a la possibilité de voir, donc de comprendre, les composantes d'un agrégat. L'agrégation fournie par l'OLAP ne détruit rien ce qui permet à l'utilisateur de remonter jusqu'à la source de l'information.[19] Un cube OLAP est une méthode de stockage de données sous forme multidimensionnelle, généralement à des fins de génération de rapports.

Dans les cubes OLAP, les données (ou mesures) sont classées par dimensions. Les cubes OLAP sont souvent pré-synthétisés entre les dimensions, ceci afin d'accélérer considérablement l'interrogation par rapport aux bases de données relationnelles.[20]

### 2.1.2 S'utilise un cube OLAP ?

Il peut être consulté directement dans Excel, comme un tableau croisé dynamique standard. Il peut également être consulté et manipulé directement dans des pages web, en HTML, grâce à certains composants.

**2.1.3 Les caractéristiques du cube OLAP:**

OLAP permet aux utilisateurs d'analyser des données présentes de plusieurs systèmes de bases de données en même temps. Alors que les bases de données relationnelles sont considérées comme bidimensionnelles, les données OLAP sont multidimensionnelles, ce qui signifie que l'information peut être comparée de nombreuses façons différentes (dimension). Par exemple, une entreprise peut comparer ses ventes d'ordinateurs en juin avec ses ventes en juillet, puis comparer ces résultats avec les ventes d'un autre endroit, qui pourraient être stockées dans une base de données différente.

Un serveur OLAP est nécessaire pour organiser et comparer les informations. Les clients peuvent analyser différents ensembles de données à l'aide des fonctions intégrées au serveur OLAP. En raison de ses puissantes capacités d'analyse de données, le traitement OLAP est souvent utilisé pour le data mining, qui vise à découvrir de nouvelles relations entre différents ensembles de données.

L'OLAP fournit aux analystes, aux gestionnaires et aux cadres supérieurs l'information dont ils ont besoin pour prendre des décisions efficaces sur les orientations stratégiques d'une organisation. L'OLAP peut fournir des informations précieuses sur la performance de leur entreprise, ainsi que sur les améliorations qu'ils peuvent apporter.

Les outils OLAP permettent aux utilisateurs d'analyser des données multidimensionnelles de manière interactive à partir de perspectives multiples. OLAP peut être utilisé pour trouver des tendances et obtenir une vue d'ensemble des données. Il peut également être utilisé pour des calculs complexes et pour créer des scénarios " et si " pour la planification prévisionnelle.

Les applications OLAP typiques incluent les rapports d'affaires pour les ventes, le marketing, les rapports de gestion, la gestion des processus d'affaires, la budgétisation et les prévisions, les rapports financiers, et plus encore.

**2.1.4 Catégories d'opérations****2.1.4.1 Opérations OLAP liées à la structure :**

Ce sont les opérations qui agissent sur la structure multidimensionnelle, qui visent à changer le point de vue des données.

Opérations	Description
Extraction d'un bloc de données (Dice)	Elle extrait un "sous-cube" du cube principal. Donc elle travaille que sous un sous-cube.
Pivot (Rotate)	Consiste à faire effectuer à un cube une rotation autour d'un de ces axes passant par le centre de deux faces opposées, de manière à présenter un ensemble différent.
Permutation (Switch)	Consiste à inter-changer la position des membres d'une dimension.
Extraction d'une Tranche du cube (Slice)	Consiste à sélectionner une dimension, il s'agit de couper une tranche du cube afin d'observer les données de la dimension.
Division (Split)	Consiste à présenter chaque tranche du cube et passer d'une présentation tridimensionnelle d'un cube à sa présentation sous forme d'un ensemble de tables. Sa généralisation permet de découper un hypercube de dimension 4 en cubes.
Emboitement (Nest)	Permet d'imbriquer des membres à partir du cube. L'intérêt de cette opération est qu'elle permet de grouper sur une même représentation bidimensionnelle toutes les informations (mesures et membres) d'un cube, quel que soit le nombre de ses dimensions.
Enfoncement (Push)	Permet de combiner les membres d'une dimension aux mesures du cube, i.e. de faire passer des membres comme contenu de cellules.

**Tableau 2-1** les Operations Olap liées à la structure [18]

#### 2.1.4.2 Opérations OLAP liées à la granularité :

Les opérations agissant sur la granularité d'observation des données, caractérisent la hiérarchie de navigation entre les différents niveaux. Elles correspondent aux opérations suivantes :

Opérations	Description
Forage vers le haut (Drill-up ou Roll-up ou Scale-up)	Consiste à représenter les données du cube à niveau de granularité supérieur conformément à la hiérarchie définie sur la dimension. Une fonction d'agrégation est utilisée (somme, moyenne, etc.) en paramètre à l'opération, indique comment sont calculés les valeurs du niveau supérieur à partir de celles du niveau inférieur.
Forage vers le bas (Drill-down ou Roll-down ou Scale-down)	Elle fait l'inverse de la précédente (Roll-up), elle consiste à représenter les données du cube à niveau de granularité de niveau inférieur, donc sous une forme plus détaillée (descendre dans la hiérarchie de dimension).

Tableau2-2 Opérations OLAP liées à granularité [18]

**Exemples :**

- **Roll\_up & drill-down :**

1) prenons l'exemple des données suivantes :

France	3
Espagne	1
Suisse	41
Canada	24
EU	14

Tel que la dimension c'est **pays**. Si on juge que cela fournit trop de détails inutiles, on peut demander un **roll-up** est le résultat sera :

Europe	45
Amérique du Nord	38

2) le même exemple (dimension pays), on souhaite décomposer les mesures selon une autre dimension comme **la couleur** :

France	Rouge	2
--------	-------	---



France	Blue	1
France	Orange	0
Espagne	Rouge	1
Espagne	Blue	0
Espagne	Orange	0
Suisse	Rouge	40
Suisse	Blue	1
Suisse	Orange	0
Canada	Rouge	20
Canada	Blue	4
Canada	Orange	0
EU	Rouge	10
EU	Blue	4
EU	Orange	0

Cette nouvelle table représente un **drill-down** par rapport à la table originale, parce qu'il y a plus de détails. L'opération inverse, celle qui consiste à faire disparaître l'attribut couleur, est un **roll-up**

- **Slice** : si on prendre l'exemple du pays, mais on intéresse qu'a la couleur rouge.

France	Rouge	2
Espagne	Rouge	1
Suisse	Rouge	40
Canada	Rouge	20
EU	Rouge	10

On se rend compte, par contre, que de conserver la dimension couleur ne sert pas a grand chose. Il vaut mieux l'omettre :

France	2
Espagne	1
Suisse	40

Canada	20
EU	10

### 2.1.4.3 Les différents outils OLAP

- **BIRT** : BIRT est un projet open source basé sur Eclipse qui s'intègre à des applications Java/J2EE pour produire des très bons rapports.
- **Pentaho** : Pentaho est une collection de projets open source, principalement axé sur la création, la production et la distribution d'un contenu riche et sophistiqué
- **Jasper Reports** : Jasper Reports est un outil de reporting Open source pour le langage Java. Il peut accéder aux données via JDBC, Table Models, JavaBeans, XML, Hibernate, CSV, Il génère des rapports au format PDF, RTF, XML, XLS, CSV, HTML, XHTML, texte, DOCX, et OpenOffice.
- **OpenRPT** : OpenRPT est un outil de reporting multiplateforme (pour Windows, Linux et Mac OS X) qui fait partie de l'ERP xTuple. Il dispose d'une interface écrite en Qt et utilise PostgreSQL comme système de gestion de base de données.
- **OpenReports**: OpenReports est un outil de web reporting puissant souple et facile à utiliser. Il supporte une variété de moteurs de reporting open source, comme par exemple JasperReports, JFreeReport, JXLS et Eclipse BIRT.
- **DataVision**: DataVision est un outil de reporting Open Source similaire à Crystal Reports. Il dispose d'une interface graphique qui permet de concevoir des rapports en utilisant le glisser-déposer. Il produit des résultats sous différents formats comme par exemple HTML, XML, PDF, Excel, LaTeX2e, DocBook.
- **IcCube**: est un moteur de traitement analytique multidimensionnel. Disponible depuis n'importe quel périphérique (ordinateurs, mobiles, tablettes) grâce à son implémentation en Java (normes J2EE).

### 2.1.5 Architecture Client / Serveur (Client Server Architecture)

La plupart des données OLAP sont stockées sur des systèmes puissants et sont accessibles via des PC. Il est donc nécessaire que les produits OLAP travaillent en environnement

client/serveur mais aussi que les composants serveur d'un produit OLAP intègrent facilement ses différents clients (avec un minimum d'effort et de programmation pour l'intégration).

### **2.1.6 Exploitation d'un cube Olap :**

Dans l'informatique décisionnelle, les cubes sont des composants importants. Ils permettent de visualiser les données suivant des dimensions prédéfinies et de réaliser des transformations. Il est possible de 'creuser' les données jusqu'au niveau de détail, de les faire pivoter autour d'axes temporels ou géographiques.

Dans le domaine du logiciel libre MONDRIAN est le projet phare pour l'agrégation des données en cube. Si MONDRIAN fournit cette structure en cube, il faut lui adjoindre un programme qui lance des requêtes sur ces cubes et qui en restitue les résultats. [33]

MONDRIAN est un serveur OLAP, développé en Open-Source depuis 2001 sous licence CPL. Ce serveur se présente sous la forme d'un ensemble de bibliothèques Java. Ces bibliothèques permettent tout d'abord la construction de structures multidimensionnelles (cubes notamment) à partir d'un entrepôt stocké dans une base de données relationnelle. Ces structures multidimensionnelles, ou cubes MONDRIAN, sont définies selon un schéma spécifié dans un fichier XML préalablement élaboré. Ce cube MONDRIAN peut être interrogé grâce à un client spécifique comme JPIVOT, JRUBRIK.

### **2.1.7 Les clients OLAP**

Le rôle du client est de générer de requêtes MDX, d'afficher le cube résultat et d'offrir des outils de manipulation des cubes.

Le client est le composant utilisé pour visualiser et manipuler les données d'une base OLAP. Un client peut être aussi simple qu'une feuille de calculs qui intègre des fonctionnalités OLAP telles le pivotement et le déroulement des données, il peut être un outil de reporting spécialisé ou encore une application OLAP spécifiquement développée pour des manipulations de données plus élaborées.

Le web est la forme la plus récente de client OLAP.

Ces applications clientes se caractérisent par une interface travaillée et conviviale, par la mise à disposition de fonctions et de structures prédéfinies, ainsi que par la proposition de solutions à des situations plus ou moins standards.

En particulier, les packages financiers sont très populaires. Une application financière préétablie permettra à beaucoup d'utilisateurs de se servir d'outils financiers les plus courants sans avoir à concevoir la structure de la base de données ni les rapports.

### **2.1.7.1 Le web comme client OLAP**

Le dernier arrivé dans la famille des clients OLAP est le web. Il y a un grand nombre d'avantages à déployer OLAP via le web. L'avantage le plus significatif c'est qu'aucun logiciel particulier n'est nécessaire à une personne pour accéder à l'information. Ce qui permet d'économiser beaucoup de temps et d'argent pour une société.

On trouve actuellement des produits Internet - OLAP très différents. Certains permettent de créer facilement des pages web, mais ne sont pas très flexibles. D'autres permettent de créer des vues de données, puis de les sauvegarder sous le format HTML. Ceux-là permettront de visualiser des données sur le web, mais ne permettront pas de les manipuler de manière interactive. D'autres produits par contre, sont entièrement interactifs et dynamiques ce qui les rendent complètement fonctionnels. L'utilisateur disposera de toutes les fonctionnalités OLAP.

Il est donc nécessaire de définir clairement les besoins avant de choisir un déploiement OLAP sur le web.

Il existe plusieurs raisons pour lesquelles le Web est jugé convenable à OLAP :

- 1- Le Web apporte une complète indépendance vis-à-vis de la plateforme. OLAP peut être déployé sur n'importe quel navigateur Web sur n'importe quelle plateforme,
- 2- Pour les administrateurs OLAP, la maintenance est plus facile. Sur le client/serveur OLAP, tout doit être installé sur le client ainsi que sur le serveur. Tandis que sur le Web, l'installation est effectuée sur le serveur uniquement,
- 3- Les techniques de navigation et de sélection d'informations sont les mêmes que pour toutes les autres applications basées Web,
- 4- L'interface du navigateur, avec des liens hypertextes et des boutons est immédiatement reconnue par les millions d'utilisateurs d'ordinateurs qui connaissent le Web.

### 2.1.7.2 Les applications

Les applications sont un type de client qui utilise des bases de données OLAP. Elles sont similaires à des outils de requêtes ou à des outils de reporting, mais elles comportent des fonctionnalités beaucoup plus approfondies. Les applications OLAP sont plus robustes qu'un simple outil de reporting.

### 2.1.8 Interrogation client/serveur

Une fois le cube MONDRIAN spécifié dans un fichier en XML, et rempli par les données de l'entrepôt, le client et le serveur MONDRIAN fonctionnent ainsi :

1. le client transmet une requête au format MDX écrite par l'utilisateur au serveur MONDRIAN,
2. Après validation du format de la requête MDX, MONDRIAN utilise le schéma XML pour transformer la requête MDX en requêtes SQL et passe la main à la couche de requêtage SQL.
3. la couche de requêtage SQL de MONDRIAN accède à la base de données relationnelle physique de l'entrepôt et récupère les données résultantes des requêtes SQL exécutées,
4. les résultats sont consolidés et transmis au client pour être présentés à l'utilisateur

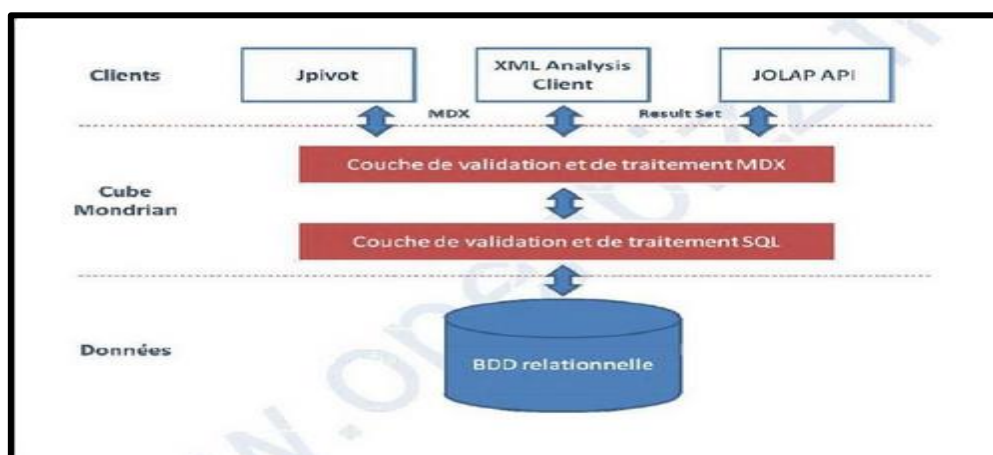


Figure 2.14 Architecture utilisée (client/serveur) [34]

### 2.1.9 JPIVOT client web

JPivot est une librairie Java permettant la connexion et l'interrogation d'un cube Mondrian à partir d'une définition d'accès à un cube existant. L'utilisateur dispose d'une interface web lui

permettant d'effectuer des analyses sur les « mesures » définies dans le « schéma XML »  
Fonctions de choix des axes à afficher Fonctions de slice et dice, drill down .

Les définitions des cubes, indicateurs, dimensions est définit dans un fichier de mapping xml.  
Nous avons prendre pour exemple la base de données foodmart, qui est livrée avec Mondrian  
[34]

**Test Query uses Mondrian OLAP**

Promotion Media	Product	Mesures		
		Unit Sales	Store Cost	Store Sales
-All Media	+All Products	266 773	225 627,23	565 238,13
Bulk Mail	+All Products	4 320	3 740,95	9 349,07
Cash Register Handout	+All Products	6 697	5 715,67	14 321,33
Daily Paper	+All Products	7 738	6 559,23	16 479,81
Daily Paper, Radio	+All Products	6 891	5 668,77	14 169,42
Daily Paper, Radio, TV	+All Products	9 513	8 055,22	20 173,97
In-Store Coupon	+All Products	3 798	3 263,11	8 162,46
No Media	+All Products	195 448	165 214,85	414 026,92
Product Attachment	+All Products	7 544	6 306,24	15 898,25
Radio	+All Products	2 454	2 087,51	5 213,61
Street Handout	+All Products	5 753	4 856,54	12 192,90
Sunday Paper	+All Products	4 339	3 673,86	9 092,89
Sunday Paper, Radio	+All Products	5 945	5 027,31	12 551,96
Sunday Paper, Radio, TV	+All Products	2 726	2 341,58	5 819,33
TV	+All Products	3 607	3 116,40	7 786,21

Slicer: [Year=1997]  
[back to index](#)

Figure 2.15 Interface principale de JPivot [34]

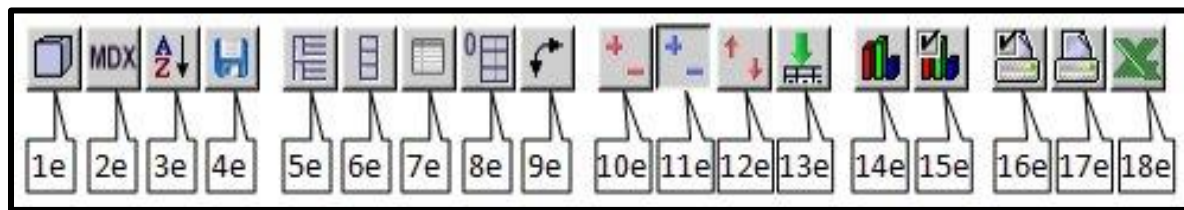


Figure 2.16 Barre de commandes de JPIVOT [34]

- 1e : Ajouter ou enlever des champs ,
- 2e : Editer la requête MDX ,
- 3e : Trier ,
- 4e : sauvegarde (dans spago BI uniquement) ,
- 5e et 6e :fusion de colonne ,
- 7e : a compléter ,
- 8e : Supprimer des lignes blanches ,

9e : Intervertir les lignes et les colonnes ,  
10e 11e 12e 13e: Drill (zoomer ,agreger,detailler) ,  
14e 15e : Générer des graphismes ,

Les 3 derniers icônes permettent de définir les options d'impression PDF et d'exporter en PDF ou en XLS (MS Excel ou openoffice).

JPIVOT est avant tout une librairie de TAG JSP et c'est à l'administrateur de développer son interface graphique. [33]

### 2.1.10 JRUBRIK application client

Pour réaliser des analyses multidimensionnelles OLAP en MDX, JRUBIK doit tout d'abord charger à partir de la base de données relationnelle de l'entrepôt, les données dans le cube MONDRIAN spécifié dans le schéma MONDRIAN XML, et ensuite agréger ces données dans une mémoire cache.

L'interface JRUBIK est à base d'onglets. Il est composé de :

- un menu principal,
- une barre d'outils,
- de plusieurs grandes fenêtres permettant chacune de :
  - choisir les objets à afficher et filtrer,
  - choisir les lignes et colonnes d'affichage,
  - d'afficher les résultats,
  - d'afficher le code MDX généré et permettre sa modification (édition).

#### 2.1.10.1 Création de requêtes OLAP sous JRUBRIK

Pour créer une requête OLAP dans JRUBIK, il est nécessaire d'utiliser l'onglet « discoverer ».

L'expression de la requête OLAP est faite à la souris, pour cela, faire glisser (Drag) les hiérarchies choisies (Seulement des hiérarchies) et les déposer (Drop) sur les boutons « ON Rows » ou «ON colonnes », appuyer enfin sur le bouton « Apply ».

On notera que :

- La fenêtre « Table\_X » de l'éditeur affiche les résultats de la requête,
- La fenêtre de l'onglet « MDX query\_X » de l'éditeur, affiche le code MDX.

Associé à la requête et généré par l'outil. Elle permet aussi de créer, modifier et lancer une requête OLAP en MDX.

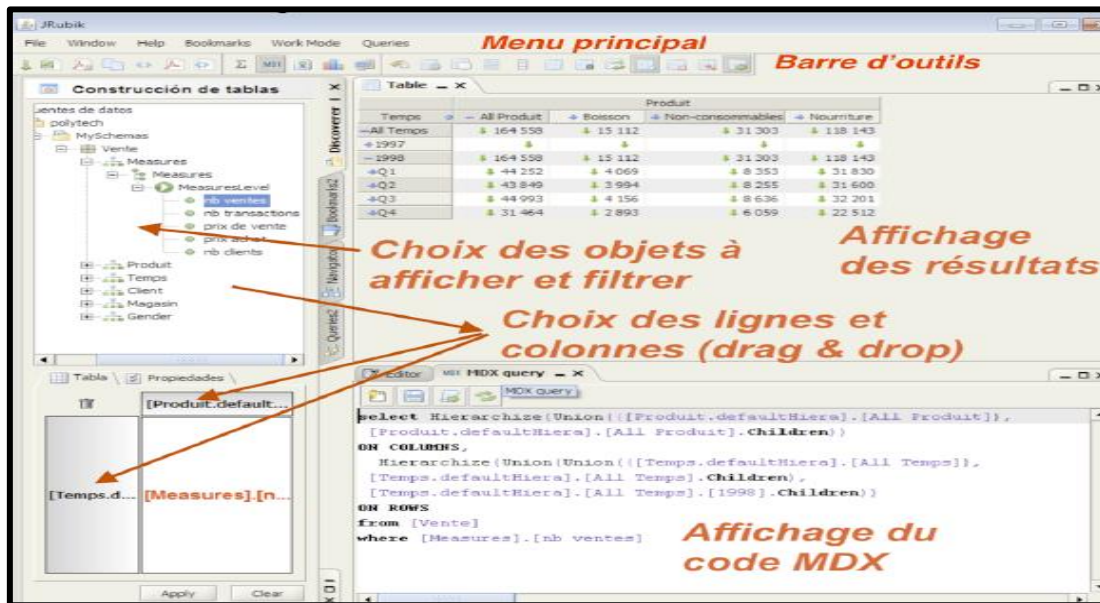


Figure 2.17 Création d'une première requête OLAP en MDX sous JRUBIK.[35]

### 2.1.11 FreeAnalysis

Anciennement nommé FreeOlap, **FreeAnalysis** est un outil de conception de cubes pour les développeurs. Il permet de manipuler les cubes Mondrian via une interface Web comme c'est également le cas pour JPivot.

FreeAnalysis propose un outil de conception de cubes baptisé Schéma Designer qui permet de créer et mettre à jour des cubes Mondrian sans passer par l'édition de fichiers XML.

**Avantages** : convivialité de l'interface avec des dimensions du cube présentées sous forme d'arbres, qualité de l'interface Ajax...

**Inconvénients** : manque d'indicateurs visuels de type **drill-replace**, icônes des tableaux croisés...

### 2.1.12 Palo : une excellente intégration à Excel

Distribué par la société allemande Jedox, Palo est un serveur M-OLAP développé en .Net et distribué sous licence GPL. Toutes les données gérées sont chargées en mémoire et calculées à la volée, ce qui permet d'obtenir des temps de réponse optimisés.



Palo permet de construire une application de saisie décentralisée permettant de consolider des données dont les valeurs évoluent dans le temps (budget initial, réalisé, révisé...).

**Avantages** : dans sa version gratuite, il s'intègre dans le tableur Excel ou Calc (OpenOffice) sous forme d'extension.

**Inconvénients** : nécessite l'utilisation d'applications clientes dédiées et ne supporte pas le langage MDX.

## 2.2 Fouille de données (Data Mining)

### 2.2.1 Définition

- Le Data Mining ou **fouille de données** est un ensemble de méthodes qui consistent à extraire un savoir ou une connaissance, à partir d'une base de donnée. Contrairement aux méthodes classiques, le Datamining est plus adapté à des données à grand volume et les informations récoltées doivent aider à prendre des décisions. [22]
- Data Mining est en fait un terme générique englobant toute une famille d'outils facilitant l'exploration et l'analyse des données contenues au sein d'une base décisionnelle de type DataWarehouse ou DataMart. Les techniques mises en action lors de l'utilisation de cet instrument d'analyse et de prospection sont particulièrement efficaces pour extraire des informations significatives depuis de grandes quantités de données. [23]

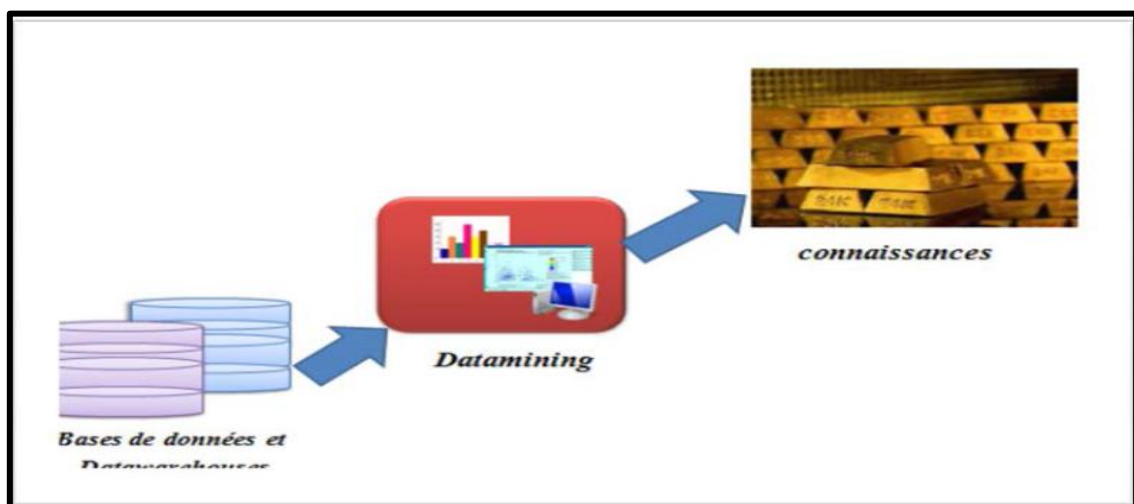


Figure 2.18 Data Mining [22]

**2.2.2 Les tâches du data mining**

Beaucoup de problèmes intellectuels, économiques ou même commerciaux peuvent être exprimés en termes des six (6) tâches suivantes :

- La classification,
- L'estimation,
- La prédiction,
- Le regroupement par similitude,
- L'analyse des clusters,
- La description.

**2.2.3 Les étapes du processus de data mining**

Plus précisément, le datamining peut se décomposer en huit (8) étapes : [27]

- a- Poser le problème,
- b- Rechercher des données,
- c- Sélectionner les données pertinentes,
- d- Nettoyer des données,
- e- Transformer les variables,
- f- Rechercher le modèle,
- g- Evaluer le résultat,
- h- Intégrer la connaissance,

**2.3 Outils reporting****2.3.1 Définition [23]**

Le terme "Reporting" désigne une famille d'outils de Business intelligence destinés à assurer la réalisation, la publication et la diffusion des rapports d'activité selon un format prédéterminé.

Ils sont essentiellement destinés à faciliter la communication de résultats chiffrés ou d'un suivi d'avancement.

**2.3.2 Principe**

L'outil de reporting assure l'interrogation des bases de données selon les requêtes SQL préparées lors de l'élaboration du modèle. Le rapport d'activité peut ensuite être publié sur l'internet, périodiquement en automatique ou ponctuellement à la demande.

L'outil offre bien entendu des fonctions spécifiques pour l'élaboration du modèle du rapport, des modules de calcul et de présentation (graphiques) afin de concevoir des comptes rendus particulièrement seyants et pertinents.

Il existe deux (2) grand domaines dans le reporting : le reporting ad hoc et le reporting de masse.

**2.3.2.1 Le reporting ad hoc et de masse****2.3.2.1.1 Définition [30]**

L'analyse ad hoc est un processus de Business Intelligence (BI) conçu pour répondre à une seule question métier spécifique. Cette analyse est effectuée par les utilisateurs professionnels selon les besoins pour répondre aux besoins d'analyse des données non satisfaits par les rapports statiques et réguliers de l'entreprise déjà effectués quotidiennement, mensuellement ou annuellement , peut également être appliquée aux méga données provenant de sources détenues en dehors de l'entreprise, par exemple, une entreprise peut exécuter une analyse ad hoc sur des sources de données internes pour voir si une baisse des ventes est liée à un ralentissement de la production ou à un récent changement de prix. Si aucun facteur interne n'est à blâmer, l'entreprise peut analyser les sources de méga données pour voir s'il y avait eu un problème hors du contrôle de l'entreprise, comme une plainte d'un client sur les réseaux sociaux qui a endommagé la marque ou un concurrent a baissé son prix, ce qui a eu un impact ventes.

Le reporting de masse permet de répondre rapidement à un besoin régulier de beaucoup d'utilisateurs comme tableau de bord. On peut définir un tableau de bord par un ensemble d'indicateurs peu nombreux conçus pour permettre aux gestionnaires de prendre connaissance de l'état et de l'évolution des systèmes qu'ils pilotent et d'identifier les tendances qui les influenceront sur un horizon cohérent avec la nature de leurs fonctions.

**2. 4 Outils de requêtes**

### 2.4.1 Les langages de données

La représentation multidimensionnelle des données a donné naissance aux nouveaux langages de données permettant de définir, et de manipuler les schémas des entrepôts de données d'une manière assimilable par la machine. Ces langages de données sont constitués d'un ensemble d'opérations qui peuvent être classées selon le type de langage de données.

### 2.4.2 Types de langages de données

Les langages de données dédiés aux entrepôts de données sont classés en trois types et qui se présentent comme suit :

#### a- Le langage de définition de données (LDD) :

Ce langage permet de définir les opérateurs nécessaires à la définition d'un schéma multidimensionnel comme la création des cubes (définition des faits), dimensions et hiérarchies ainsi que la mise à jour (la modification et / ou la suppression) de ses éléments.

#### b- Le langage de contrôle des données (LCD) :

Ce langage permet de gérer la sécurité de l'entrepôt de données, il définit les opérateurs nécessaires à l'affectation et à la suppression des droits des utilisateurs (les permissions) sur les éléments du schéma multidimensionnel.

#### c- Langage de manipulation des données (LMD) :

Ce langage adresse des requêtes à l'entrepôt de données et définit les opérateurs nécessaires à la consultation des données.

### 2.4.3 Le langage de script de manipulation des données MDX:

#### 2.4.3.1 Définition

MDX, acronyme de *Multidimensionnel eXpressions*, est un langage de requêtes qui sont exprimées sous forme d'un script à base d'instruction. Le MDX comprend une syntaxe qui prend en charge la définition le contrôle et la manipulation de données et d'objets multidimensionnels (bases de données multidimensionnelles) crée par Microsoft, de la même manière que SQL est utilisé pour les bases de données relationnelles(Le langage MDX est aux cubes ce que SQL est aux tables) .

#### 2.4.3.2 Caractéristiques de MDX :

Le MDX offre une syntaxe assez similaire à celle de SQL, mais il ne s'agit pas d'une extension du langage SQL, en fait, certaines des fonctionnalités prises en charge par MDX peuvent être fournies, avec certes moins d'efficacité ou de facilité, par SQL.

A l'instar d'une requête SQL, chaque requête MDX nécessite une demande de données

(La clause **SELECT**), un point de départ (la clause **FROM**) et un filtre (la clause **WHERE**).

Ces mots clés, ainsi que d'autres, fournissent les outils nécessaires à l'extraction de portions de données spécifiques d'un cube en vue d'une analyse. Le MDX fournit également un jeu robuste de fonctions pour la manipulation des données extraites, ainsi que la possibilité d'enrichir MDX avec des fonctions définies par l'utilisateur.

MDX, comme, SQL fournit un langage de définition de données, DDL (*Data Definition*

*Language*), pour la gestion des structures de données. Certaines commandes de MDX permettent de créer et de supprimer des cubes, des dimensions, des mesures, ainsi que les objets qui leur sont subordonnés.

#### 2.4.3.3 Forme d'une requête OLAP sous MDX :

Peut avoir plusieurs dimensions et la dimension Mesure est la seule que chaque cube doit contenir. Chaque dimension des membres provenant d'attributs et chaque membre représente un élément unique au sein d'une dimension. Les mesures sont les membres de la dimension mesures et ces mesures représentent les données organisées par les autres dimensions du cube.

Une requête MDX permet de rechercher des mesures à des niveaux différents de granularité, en formulant une expression unique et peut construire une visualisation pour l'ensemble des résultats.

#### 2.4.3.4 Syntaxe de requête MDX

**SELECT** les colonnes, les lignes et autres axes (sélectionner les axes d'analyse)

**FROM** le cube (le cube sur lequel porte l'analyse)

**WHERE** les dimensions tranchées (conditions de sélection d'axes d'analyse)

#### 2.4.3.5 Exemples:

**Exemple 1** si nous avons le cube 'Ventes' présenté dans cette figure :

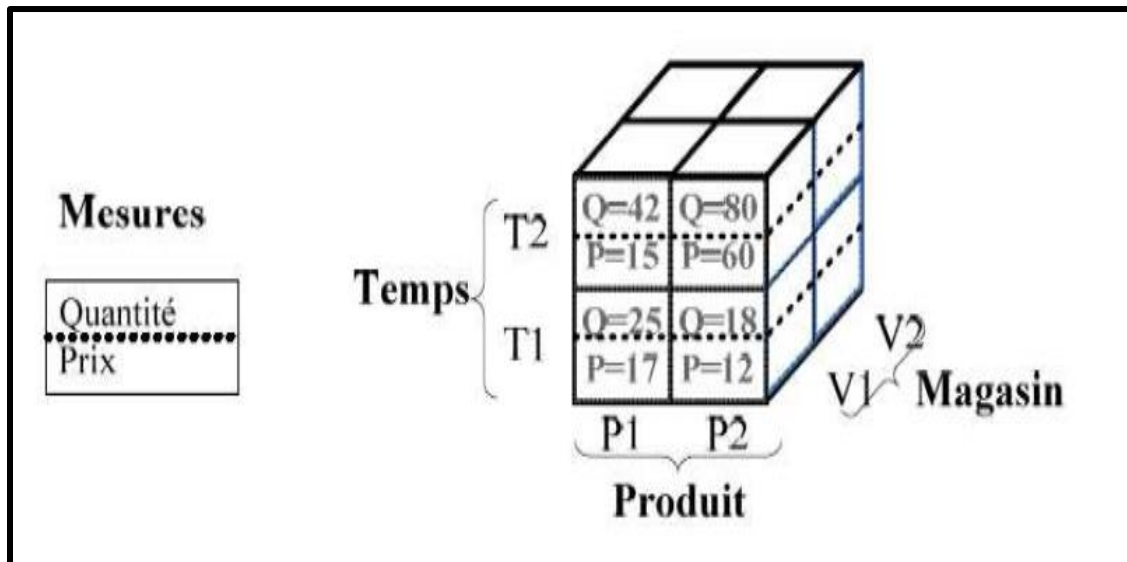


Figure 2.19 Cube « Ventes »[36]

• Ce cube est composé de trois dimensions qui sont Temps, Produit, Magasin et la Dimension mesures.

- La dimension Temps contient deux membres T1 et T2,
- La dimension Produit contient deux membres P1 et P2,
- La dimension Magasin contient deux membres V1 et V2,
- La dimension Mesures contient deux mesures (membres) quantité(q) et prix(p).

La requête MDX qui permet de voir les quantités vendues des produits P1 et P2 et qui correspondent aux temps T1 et T2 est la suivante :

```
SELECT {Produit.[P1],Produit.[P2]} ON COLUMNS,
      {Temps.[T1],Temps.[T2]} ON ROWS
FROM Ventes
```

### ➤ Exemple 2

```
SELECT {[Year].2003,[Year].2004,[Year].2005,[Year].2006} ON COLUMNS
      CROSSJOIN({[City].Tours,[City].Orleans}),
            {[Category].shoes,[Category].cloth,
            [Category].food,[Category].drink }) ON ROWS
FROM Sales
WHERE [Measures].quantity
```

		2003	2004	2005	2006
Tours	drink	77,00	54,00	55,00	33,00
	food	89,00	61,00	30,00	41,00
Orleans	drink	25,00	50,00	49,00	32,00
	food	33,00	44,00	59,00	27,00

Figure 2.20 Résultat d'exécution requête d'exemple 2 [36]

#### 2.4.3.6 Structure du résultat d'une requête MDX :

Le résultat d'une telle requête est un tableau multidimensionnel, dont les cellules ne contiennent qu'une valeur et dans lequel on peut naviguer avec les opérations décrites. Une requête MDX permet de rechercher des mesures à des niveaux différents de granularité en formulant une expression unique et de spécifier comment afficher la réponse sous forme de table croisée. De plus la syntaxe MDX distingue l'identification des axes d'affichage par 'COLUMNS' et 'ROWS'. Ainsi, l'emplacement des informations de la requête sur ces axes est précisé.

##### ➤ Exemple :

Si nous avons le cube 'Ventes' présenté ci après, Le résultat de la requête de l'exemple 1 précédent est une table croisée de deux lignes et deux colonnes et qui se présente ainsi :

	P1	P2
T1	25	18
T2	42	80

Tableau 2-3 résultat d'une requête MDX sur le cube 'vente' [36]

Tel que :

P1 et P2 : membres de la dimension Produit,

T1 et T2 : membres de la dimension Temps,

[25 , 18 , 42, 80 ] : les quantités vendues des produits P1 et P2 Correspondent aux temps T1 et T2 .

### 2.4.3.7 Comparaison entre SQL et MDX

MDX est fait pour naviguer dans les bases multidimensionnelles et pour exécuter des requêtes sur tous leurs objets (dimensions, hiérarchies, niveaux, membres et cellules) afin d'obtenir (simplement) une représentation sous forme de tableaux croisés [31]

SQL	MDX
Table	Cube
Plusieurs colonnes liées ou une table de dimension	Dimension
Colonne (chaîne de caractère ou valeur numérique)	Niveau (Level)
Valeur dans une colonne et une ligne particulière de la table	Membre de dimension (Dimension member)

**Tableau 2-4** différence entre requête SQL et une requête MDX [36]

### 2.4.3.8 Les avantages et inconvénients du langage MDX

Le langage MDX possède une syntaxe un peu ardue. Néanmoins le temps que nous passons à son assimilation est largement compensé par celui que nous gagnons lors de l'écriture des requêtes du langage SQL. De plus, les fondements et les formalisations théoriques de ce langage sont mal connus, et il présente l'inconvénient de ne pas être clos. Si l'utilisation interactive des cubes est intéressante, leur structure multidimensionnelle associée à la puissance de ce langage en fait des outils de prédilection pour le reporting opérationnel d'entreprise ou d'administration. Au-delà de cet aspect, le MDX permet d'explorer les données de cubes et crée la vraie valeur autour des données et c'est ici que se trouve la plus value. Il n'en reste pas moins que MDX reste un langage, qui n'est pas forcément simple à appréhender. Des éditeurs commencent à proposer des outils multidimensionnels.



**Conclusion**

L'interrogation des entrepôts de données est une tâche circulaire qui permet de naviguer dans l'entrepôt de données afin d'exploiter les résultats obtenus.

Dans ce chapitre on a parlé sur les différentes stratégies d'interrogation d'un Data warehouse tel que le cube OLAP avec ces différents clients (web et applications), qui permettent d'exploiter les données et de les afficher sous forme des cubes multidimensionnels et de naviguer dans les différentes dimensions , puis les outils reporting qui permettent la mise à disposition de rapports périodiques, pré-formatés et paramétrables par les opérationnels et les outils requêtes comme MDX qui est la partie la plus importante dans notre travail .

Dans le chapitre suivant, nous allons développer notre application Java qui permette d'interroger a un data warehouse (foodmart).

## **Chapitre 3 :**

# ***Conception, implémentation et test***

## Introduction

Ce dernier chapitre est consacré à la conception et développement de notre application Java (*MdxQuery*), à la présentation de différents outils mis en jeu dans sa réalisation et l'ensemble de fonctionnalités qu'offre cette dernière sous forme de prises d'écran accompagnées des descriptions pour bien comprendre le mode de d'utilisation de notre application développée (*MdxQuery*).

### 1. Conception

Dans le cycle de vie d'un projet, la conception représente une phase primordiale et déterminante pour produire une application de haute qualité.

#### 1.1 Méthodologie de conception :

Pour faciliter notre tâche nous avons recours langage de modélisation unifiée (UML :Unified Modeling Language ) c'est une notion qui permet de modéliser un problème de façon standard. Ce langage est né de la fusion de plusieurs méthodes existantes auparavant, et il est devenu une référence en terme de modélisation objet à un tel point que sa connaissance devienne indispensable pour un développeur.

#### 1.2 Les diagrammes des cas d'utilisation :

Le rôle des diagrammes de cas d'utilisation permettre l'expression des besoins de l'utilisateur de façon beaucoup plus simple. Ils permettent de recenser les principales fonctionnalités du système.

Le diagramme de cas d'utilisation apporte une vision utilisateur et absolument pas une vision informatique .Il nécessite aucune connaissance informatique et l'idéal serait qu'il soit réalisé par le client.

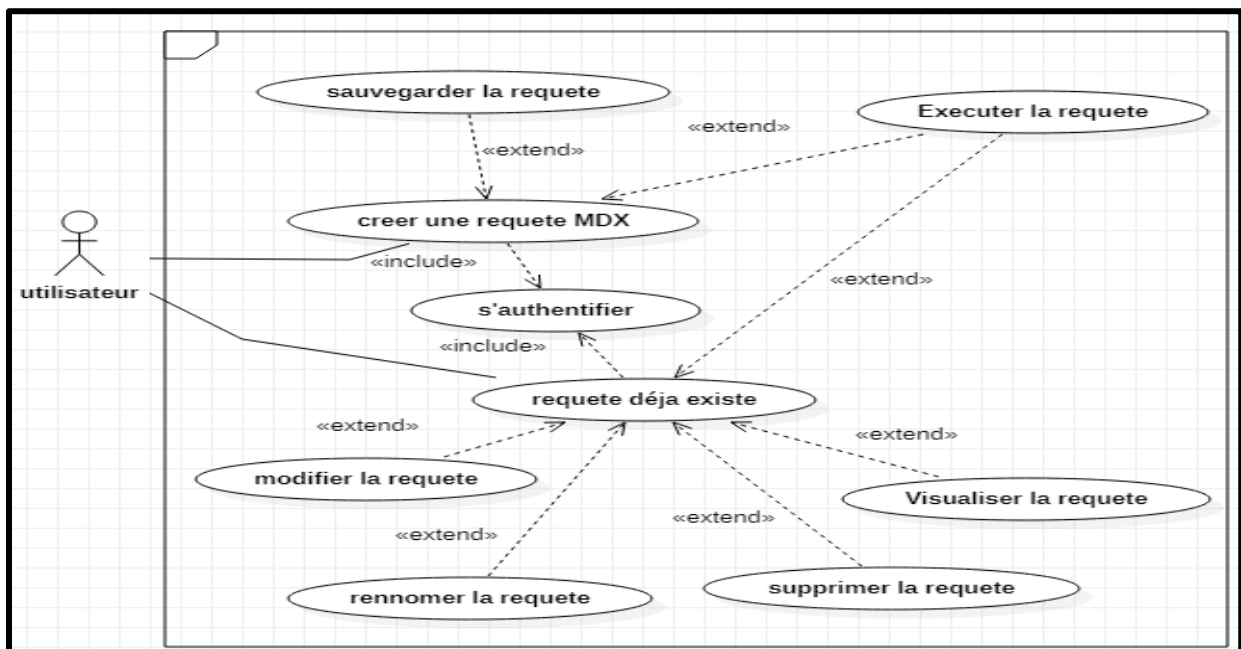
##### 1.2.1 Composition du diagramme de cas d'utilisation

Le diagramme de cas se compose de trois éléments principaux :[37]

- Un cas d'utilisation : est un service rendu à l'utilisateur, il implique des séries d'action plus élémentaires.

- Un acteur : est une entité extérieure au système modélise, et qui interagit directement avec lui.
- Relation : Entre cas d'utilisation et acteur : les acteurs implique dans un cas d'utilisation lui sont lies par association :
  - a) Inclusion : le cas A inclut le cas B (B est une partie obligatoire de A ),
  - b) Extension : le cas B étend le cas A (B est une partie optionnelle de A ) ,
  - c) Généralisation : le cas A est une généralisation du cas B (B est une sorte de A).

### 1.2.1.1 diagramme de cas d'utilisation de notre application (*MdxQuery*)



**Figure 3.21** Diagramme de cas d'utilisation de notre application

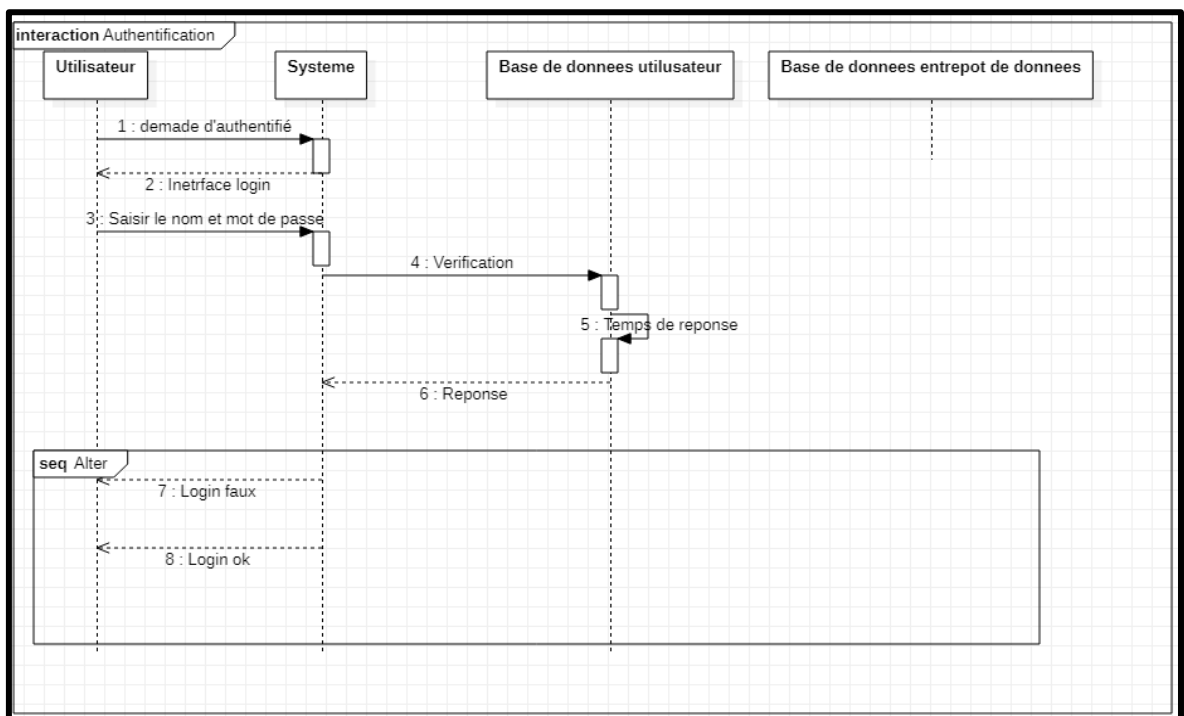
### 1.3 Les diagramme de séquences

- Représente les communications avec et au sein du logiciel,
- Représentation temporelle des interactions entre les objets,
- Chronologie des messages échangés entre les objets et avec les acteurs.

#### 1.3.1 Composition d'un diagramme de séquence

- Les lignes de vie : La ligne de vie indique les périodes d'activité de l'objet.
- Les messages : deux type de message, le premier est dit message synchrone bloque l'expéditeur jusqu'à la réponse du destinataire. Le deuxième est appelé message asynchrone, n'est pas bloquant pour l'expéditeur. Le message envoyé peut être pris en compte par le récepteur à tout moment ou ignore.
- Les occurrences d'exécution : représente la période d'exécution d'une opération.
- Les commentaires : un commentaire peut être joint a tout point sur une ligne de vie
- Les itérations : représente un message de réponse suite à une question de vérification.

**1.3.2.1 diagramme de séquence d'authentification**



**Figure 3.22** Diagramme de séquence d'authentification

**1.3.2.2 diagramme de séquence de new query**

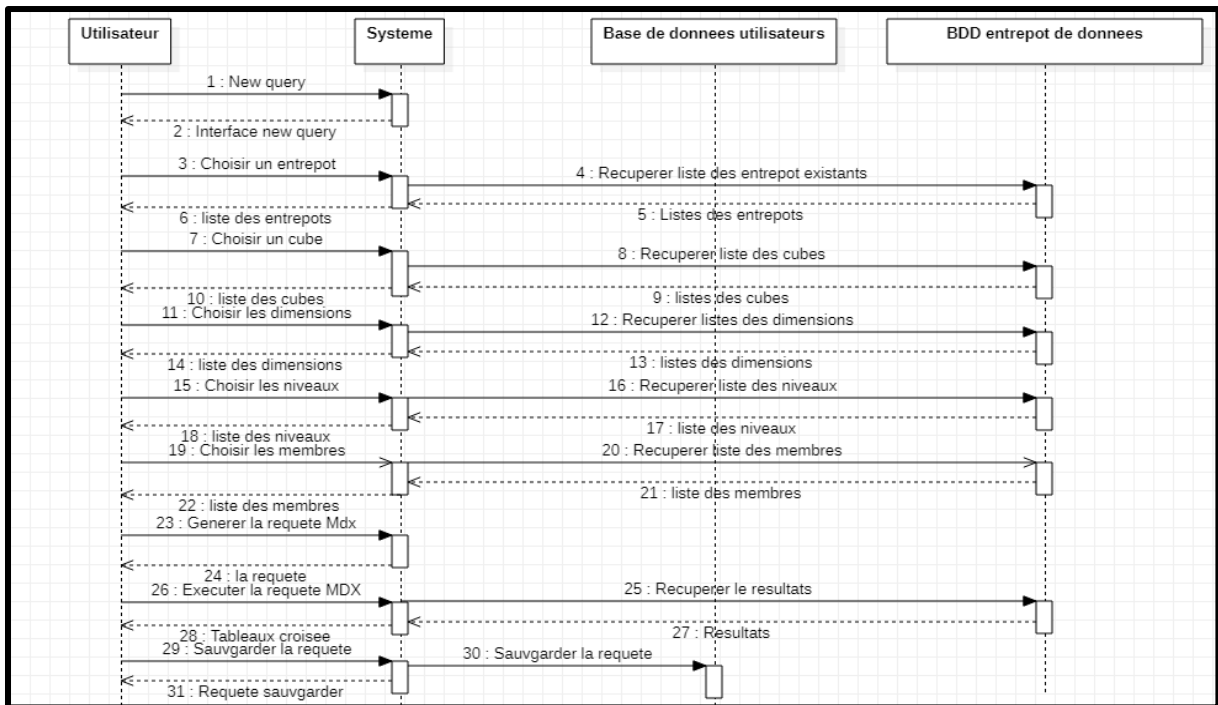


Figure 3.23 Diagramme de séquence de création nouvelle requête

1.3.2.3 diagramme de séquence de modifier une requête

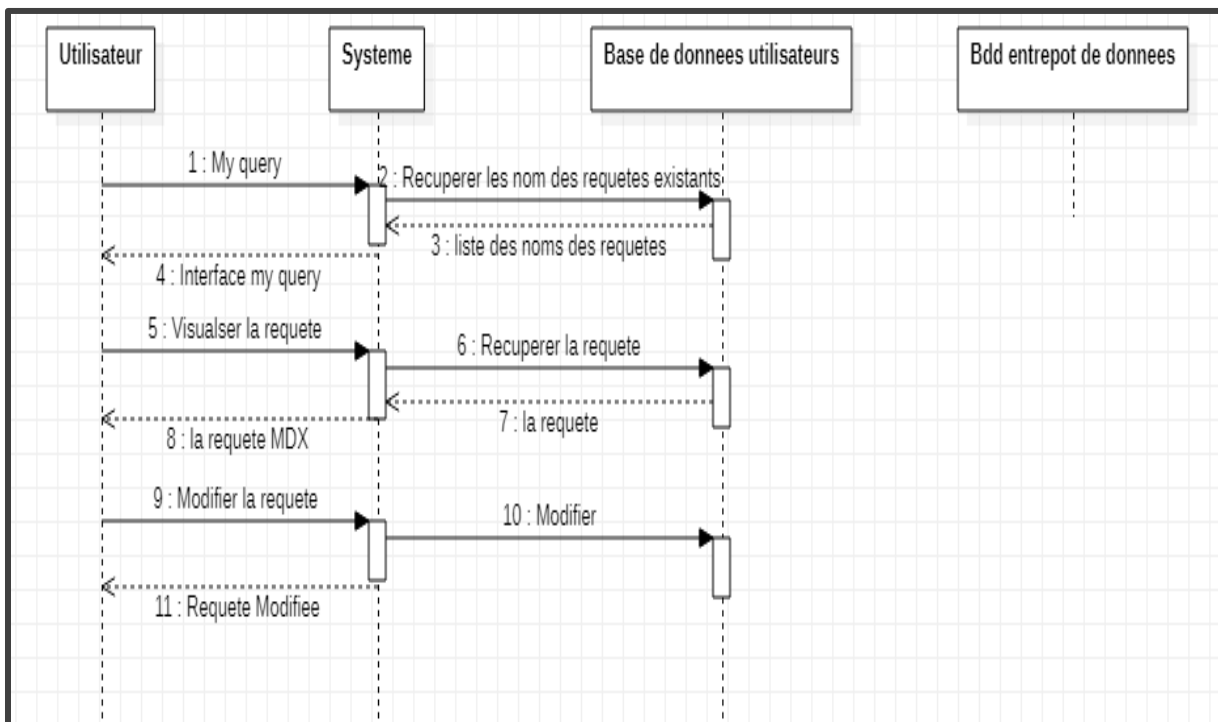


Figure 3.24 Diagramme de séquence de modifie une requête

1.3.2.4 Diagramme de séquence de supprimer une requête

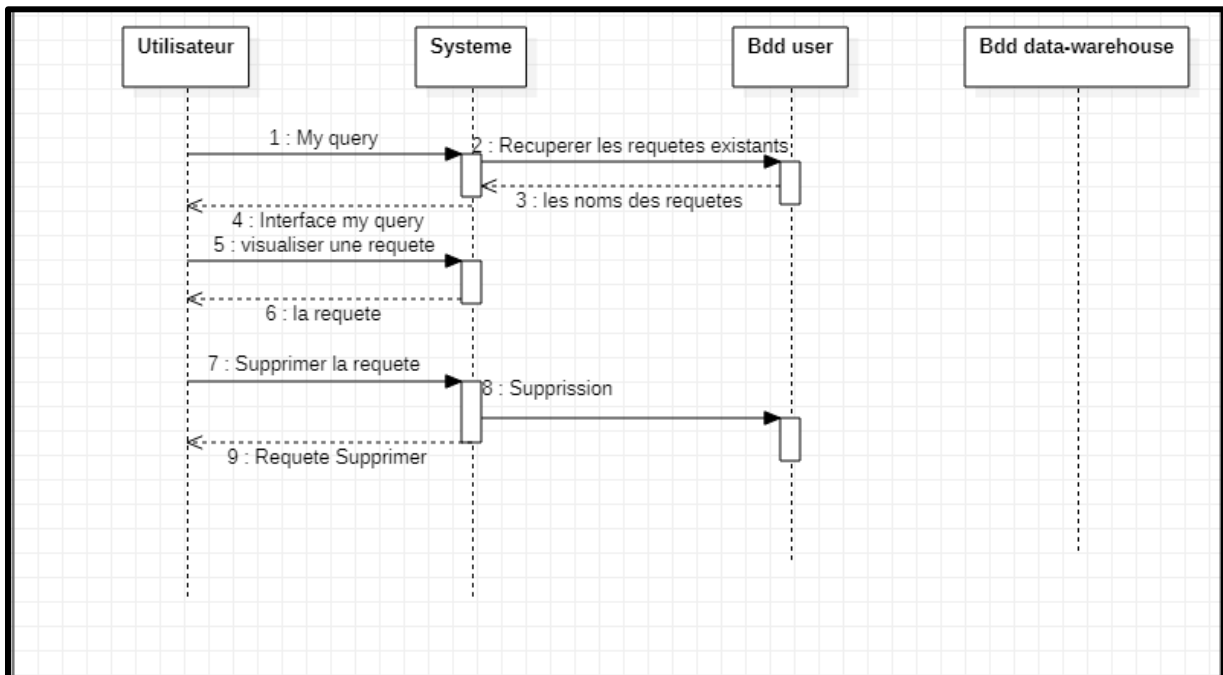


Figure 3.25 Diagramme de séquence pour supprimer une requête

1.3.2.5 diagramme de séquence de exécuter une requête

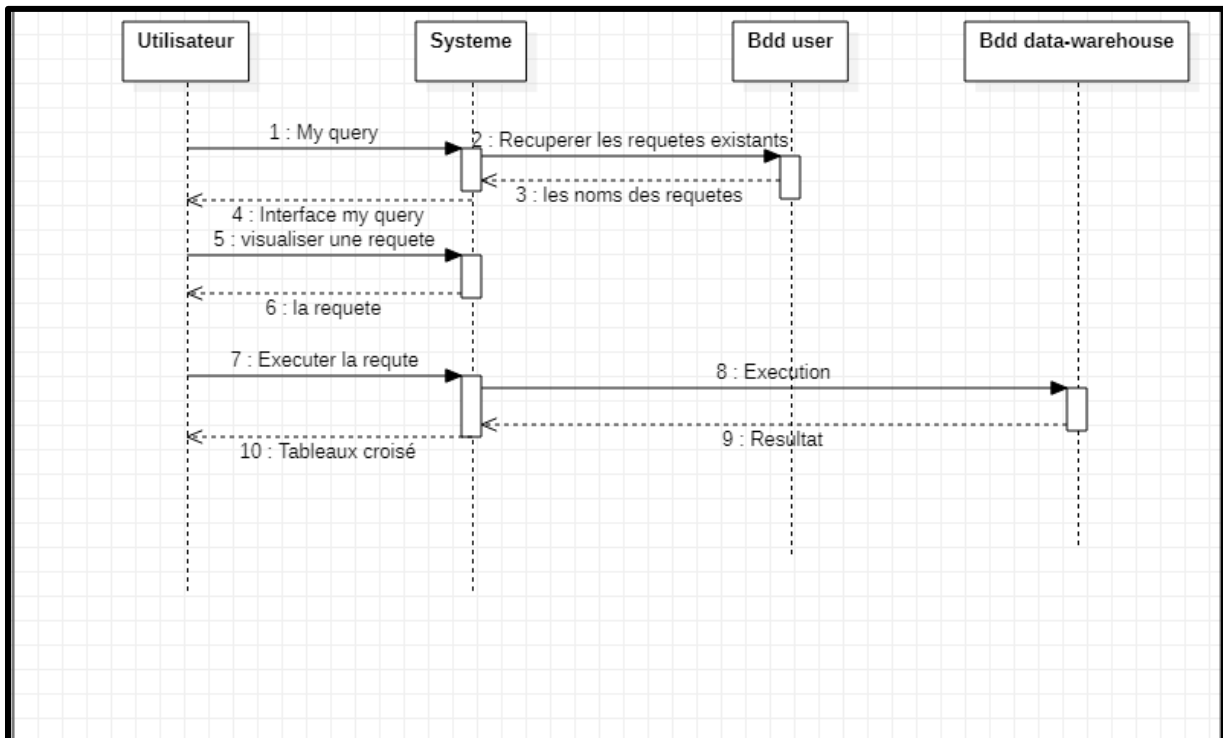
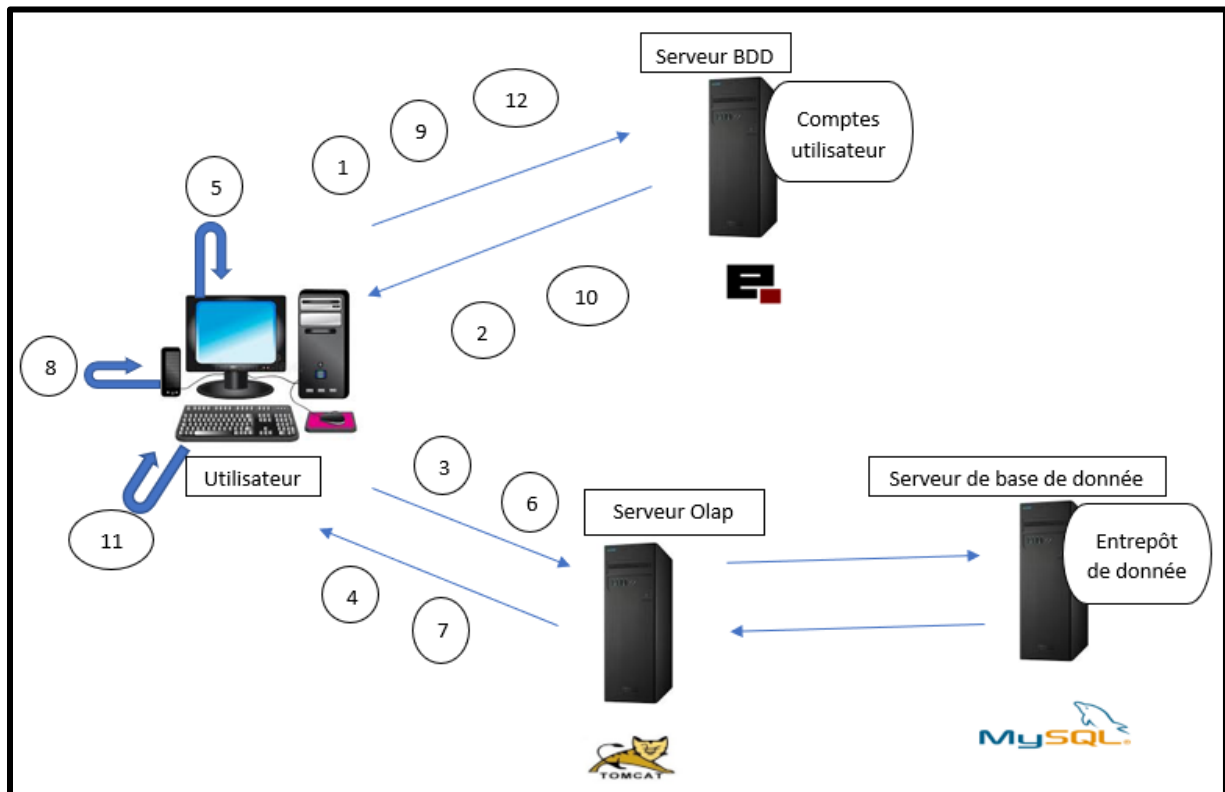


Figure 3. 26 Diagramme de séquence pour Exécuter une requête déjà existe

## 2 L'environnement de développement :

Au niveau de cette dernière partie, nous allons énumérer les outils que nous avons utilisé pour réaliser notre application

Le schéma suivant décrit le principe de base pour le fonctionnement de notre système :



1. Vérification (Nom d'utilisateur + mot de passe)
2. Affichage page d'accueil
3. Demande de Dimension, Levels (niveaux) et membres (valeurs)
4. Récupérer les (Dimension, Levels (niveaux) et membres (valeurs))
5. Construire une requête MDX
6. Exécuter une requête MDX
7. Envoie du résultat d'exécution d'une requête MDX
8. Affichage format de résultat sous forme d'une table croisé
9. Demande la liste des requêtes MDX stockés dans le compte de l'utilisateur
10. Envoie la liste des requêtes MDX stockés dans le compte de l'utilisateur
11. Modifier, renommer, supprimer une requête MDX
12. Sauvegarder, suppression, mise à jour une requête MDX



## 2.1 Mondrian

Mondrian est un moteur OLAP (Online Analytical Processing) écrit en Java par Julian Hyde qui permet la conception, la publication et le requêtage de cubes multidimensionnels.

Mondrian permet l'exécution de requêtes en langage MDX sur des entrepôts de données s'appuyant sur des SGBDR, d'où sa caractérisation de « ROLAP » (Relational OLAP). En matière de ROLAP, Mondrian est la référence open source.

## 2.2 MySQL

MySQL est un Système de Gestion de Base de Données (SGBD) parmi les plus populaires au monde. Il est distribué sous double licence, une licence publique générale GNU et une propriétaire selon l'utilisation qui en est faites. La première version de MySQL est apparue en 1995 et l'outil est régulièrement entretenu.

Ce système est particulièrement connu des développeurs pour faire partit des célèbres quatuors: WAMP (Windows, Apache, MySQL et PHP), LAMP (Linux) et MAMP (Mac). Ces packages sont si populaires et simples à mettre en œuvre que MySQL est largement connu et exploité comme système de gestion de base de données pour des applications utilisant PHP. C'est d'ailleurs pour cette raison que la plupart des hébergeurs web proposent PHP et MySQL.

## 2.3 JAVA

Le langage java, né en 1995, est un langage orienté objets, développé par Sun Microsystems Il permet de créer des logiciels compatibles avec de nombreux systèmes d'exploitations (Windows, Linux, Macintosh, Solaris). Java donne aussi la possibilité de développer des programmes pour téléphones portables et assistants personnels. Enfin, ce langage peut-être utilisé sur internet pour des petites applications intégrées à la page web (applet) ou encore comme langage serveur (jsp).

## 2.4 NetBeans

NetBeans est un environnement de développement intégré (IDE) pour Java, placé en open source par Sun en juin 2000 sous licence CDDL (Common Développent and Distribution License). En plus de Java, NetBeans permet également de supporter différents autres

langages, comme Python, C, C++, XML et HTML. Il comprend toutes les caractéristiques d'un IDE moderne (éditeur en couleur, projets multi-langage, re factoring, éditeur graphique d'interfaces et de pages web). NetBeans est disponible sous Windows, Linux, Solaris (sur x86 et SPARC), Mac OS X et Open VMS. NetBeans est lui-même développé en Java, ce qui peut le rendre assez lent et gourmand en ressources mémoires.

## 2.5 Tomcat

Apache-Tomcat est le serveur d'applications Java du projet Jakarta de la fondation Apache. Ce serveur libre, sous licence Apache permet d'exécuter des applications Web développées avec les technologies Java (Servlets, JSP...). Apache-Tomcat trouve ses origines au tout début de l'apparition des technologies Servlets et JSP Java lorsque Sun Microsystems décide de donner le code de son serveur Java Web Server à la fondation Apache (1999). Aujourd'hui, Tomcat est pour Sun Microsystems, le serveur de référence pour les technologies Java EE Servlet et JSP. Tomcat est un moteur de Servlets fiable, évolutif et adapté à l'utilisation professionnelle. Il est actuellement utilisé dans le monde entier et mis en application au sein de domaines très variés. [38]

## 2.6 EasyPHP

Il s'agit d'une plateforme de développement web, permettant de faire fonctionner localement. C'est un environnement comprenant deux serveur (un serveur web apache et un serveur de bases de données MySQL). Il dispose d'une interface d'administration servant à gérer les alias (dossiers virtuels disponibles sous Apache), et le démarrage/arrêt des serveurs. Il permet par conséquent d'installer en une seule fois tout l'indispensable au développement local du PHP. Par défaut, le serveur Apache crée un nom de domaine virtuel (en local) `http://127.0.0.1` ou `http://localhost`. Ainsi, lorsque on choisit "Web local" dans le menu d'EasyPHP, le navigateur s'ouvre sur cette URL et affiche la page index. PHP de ce site qui correspond en fait au contenu du dossier d'EasyPHP.[39]

## 2.7 Foodmart

Les données de cet entrepôt (base de données de teste FoodMart) sont stockés sur le serveur de base de données utilisant le SGBD MySql, avec l'ensemble des tables de dimensions suivant :

- **Customers** : stocke les informations concernant les clients,
- **Education Level** : stocke le niveau d'éducation pour chaque client,
- **Gender** : le sexe de client, male ou female,
- **Marital Status** : stocke la situation maritale pour chaque client «S»: célibataire ou «M» : marié,
- **Product** : stocke les produits a vendre dans FoodMart,
- **Promotion media** : stocke les medias utilisées pour une promotion, tel que : les journaux, la tv ou le radio,
- **Promotions** : identifier la promotion qui a déclenchée le vente,
- **Store** : stocke la localisation de différentes stockes (dépôts) dans la chaîne (pays, état, cité),
- **Time** : stocke pour chaque vente, la période quand le vente était déclenché.

Et la table de fait :

- **Measures** : cette table contient plusieurs mesures tel que :
  - ✓ **Unit Sales** : nombre d'unités vendues,
  - ✓ **Store Cost** : coût de marchandises vendues,
  - ✓ **Store Sales** : valeur des transactions de ventes,
  - ✓ **Sales Count** : nombre des transactions de ventes.

### 3- Principales fenêtres de notre application

#### 3.1 Fenêtre login membres

Chaque utilisateur devrait avoir un compte, nom d'utilisateur et mot de passe, pour avoir accès à l'application et bénéficier de leurs diverses fonctionnalités.

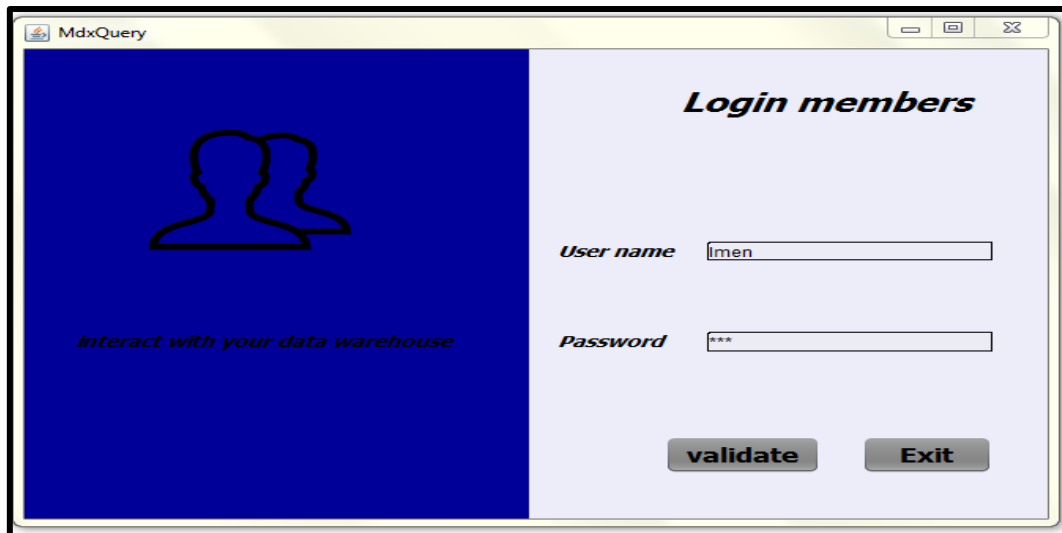


Figure 3.30 Fenêtre login members

### 3.2 Fenêtre Welcome :

De cet endroit l'utilisateur peut commencer la navigation sur notre système. Il a deux parties. La première (*My query*) lui permet de créer une nouvelle (à zéro) requête MDX. Tandis que, dans la seconde (*New query*), il aura la possibilité de consulter et mettre à jour (renommer, modifier, supprimer) et exécuter ses requêtes MDX stockées déjà dans son compte.

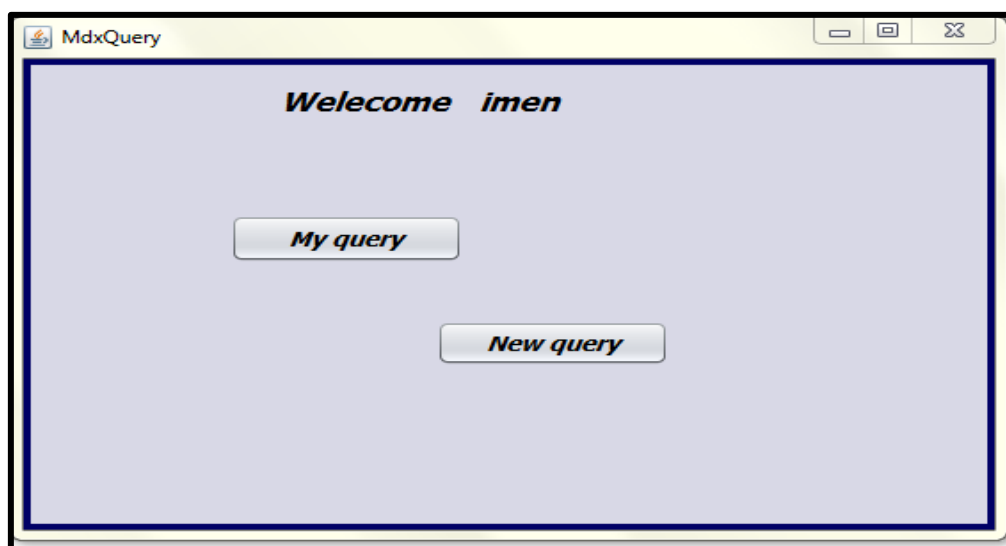


Figure 3.27 Fenêtre welcome

### 3.3 Fenêtre choose your data warehouse

La première étape dans la procédure de création d'une nouvelle requête MDX, est le choix de l'entrepôt de données (c'est *Foodmart* dans notre cas) sur quoi cette requête MDX va s'exécuter.



Figure 3. 28 Fenêtre choose your data warehouse

### 3.4 Fenêtre select a cube

Cette étape (pour la clause *FROM*) permet d'indiquer le cube (notre base de données Foodmart comporte plusieurs Cubes tels que Sales, Warehouse, .. etc.) qui représente la source de données.

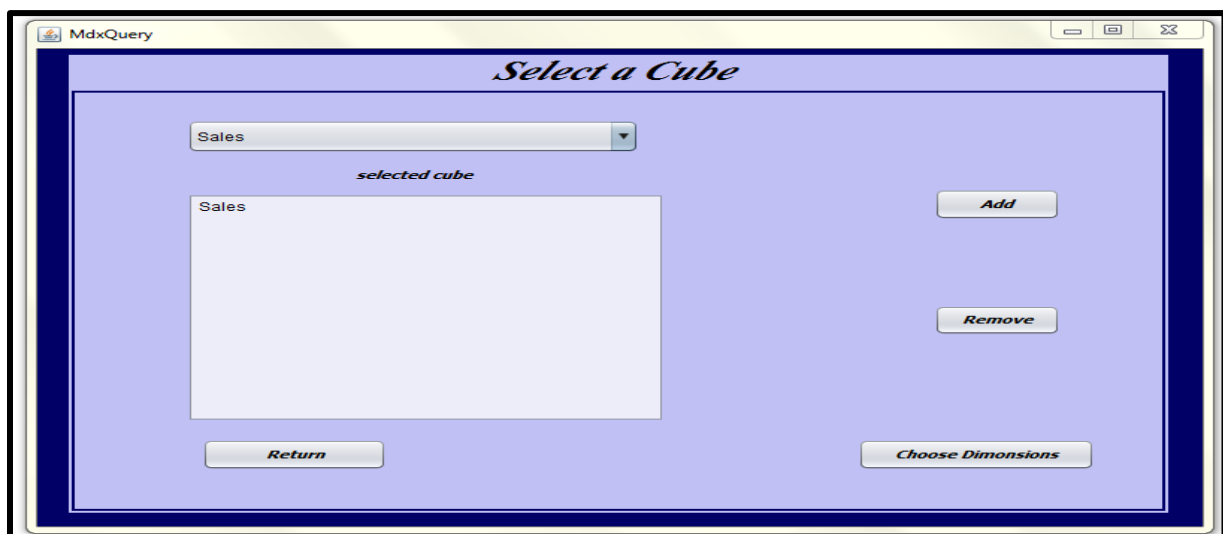


Figure 3.29 Fenêtre select a cube

### 3.5 Fenêtre select dimension

Dans cette phase de création d'une nouvelle requête MDX, l'utilisateur doit sélectionner, parmi toutes les tables (tables de dimensions et tables de fait) de l'entrepôt de données, l'ensemble de dimensions mise en jeu dans la clause SELECT et la clause WHERE. S'il veut, par exemple, créer une requête MDX pour connaître la quantité des produits les plus vendus par année et par qui sont achetés, alors il a besoin de sélectionner les trois dimensions suivant : *Product*, *Time* et *Gender*.

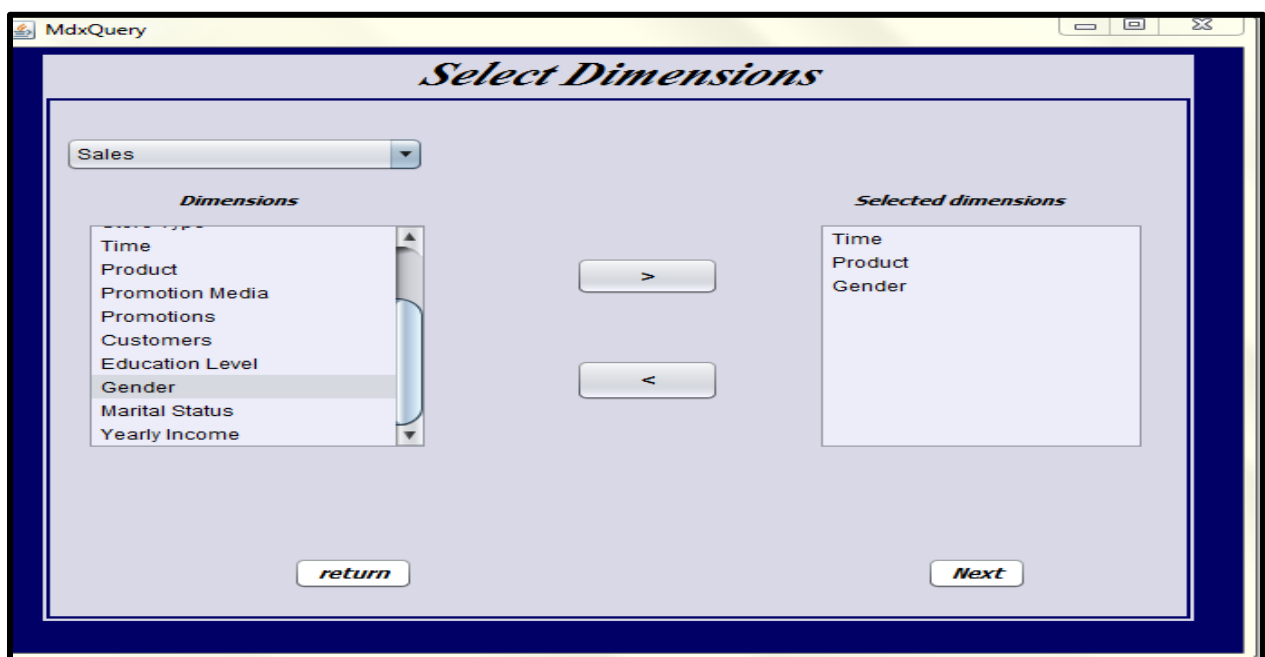


Figure 3.30 Fenêtre select dimensions

### 3.6 Fenêtre select Levels

Après avoir sélectionné les dimensions mises en jeu dans la création d'une nouvelle requête MDX dans l'étape précédente et de fait que chaque dimension est organisée sous forme d'une hiérarchie de niveaux (*Levels*), l'utilisateur devrait préciser, pour chaque dimension sélectionnée, leurs niveaux concernés par le sujet de création. Par exemple, pour la dimension *Product*, l'utilisateur choisit le niveau (*Level*) Familles des produits (*Product Family*) et pour la dimension *Time* il sera le niveau (*Level*) *Year*.

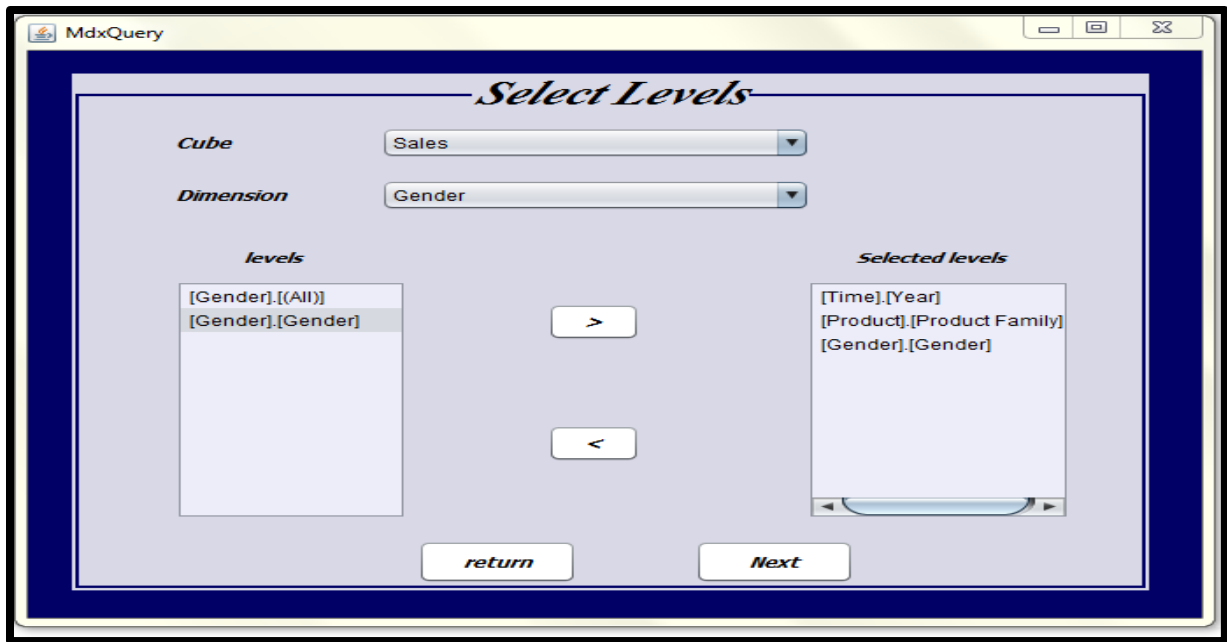


Figure 3.31 Fenêtre select Levels

### 3.7 Fenêtre select membres

Après la sélection des levels , cet dernier est également organisée sous forme d'une hiérarchie des membres , donc l'utilisateur doit choisir pour chaque levels sélectionnés les membres désirées , dans notre exemple pour le niveau (level) *Time* , l'utilisateur choisit le membre (1997) , et pour *product family* , il a choisi les deux membres ( Food et Drink ) ,et (M : male, f : femelle) pour le level *gender* .

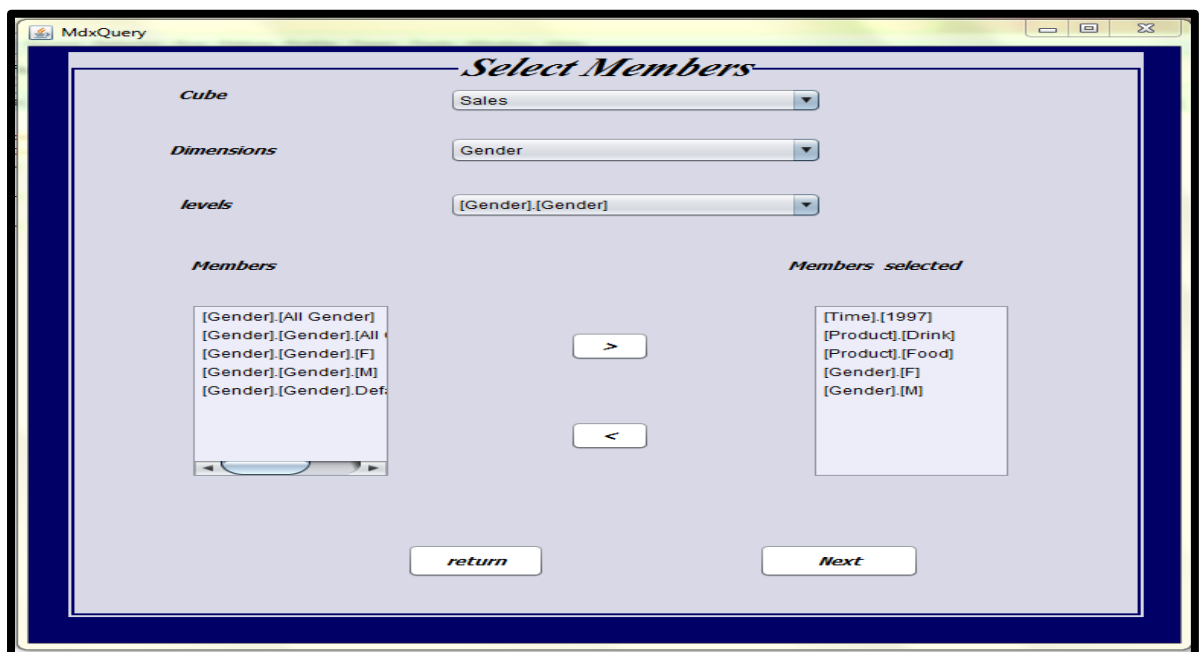


Figure 3.32 Fenêtre select members

### 3.8 Fenêtre select Axes

Après avoir terminé l'étape de sélection, l'utilisateur organise ses choix en axes, il y a deux axes *Columns* et *Rows* , Par exemple en plaçant les membres (*product food ,product family* ) dans l'axe *Columns* , et ( *gender male , femelle*) dans l'axe *Rows*.

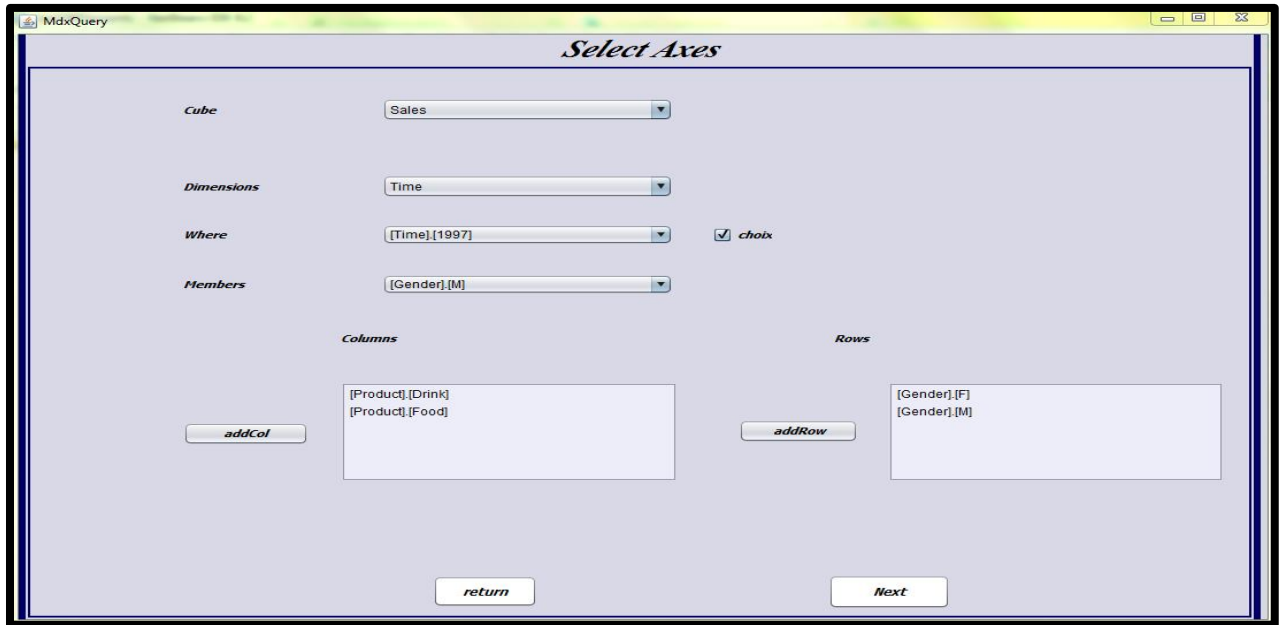


Figure 3. 33 Fenêtre Select Axes

### 3.9 Fenêtre generate query

Dans cette fenêtre l'utilisateur obtient la structure finale de leur requête, qui correspondante a tous ses Choix, en commençant par le cube jusqu'aux nombres. L'utilisateur peut exécuter la Requête et la sauvegardé

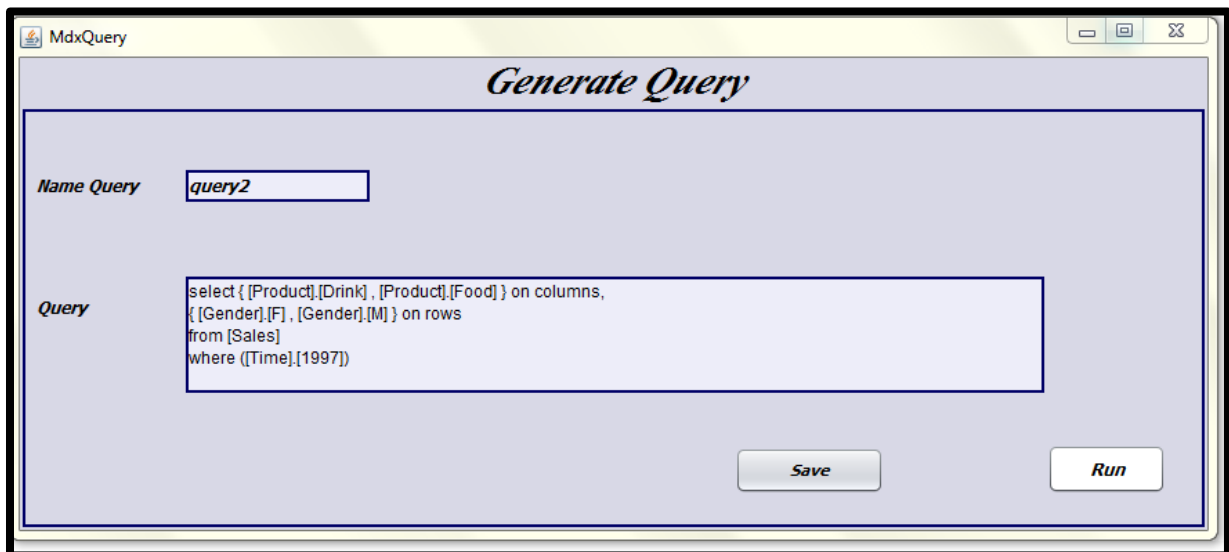


Figure 3.34 Fenêtre generate query



### 3.10 Fenêtre d'exécution la requête

C'est la fenêtre qui montre à l'utilisateur le résultat (sous forme d'un tableau croisé) d'exécution de son requête MDX (créée nouvellement où stockée dans son compte).

[Time].[1997]	Drink	Food
F	12202.0	94814.0
M	12395.0	97126.0

Figure 3. 35 Fenêtre d'exécution une nouvelle requête MDX

### 3.11 Fenêtre de my query

Dans cette fenêtre, l'utilisateur peut voir la liste de l'ensemble des requêtes MDX stockées dans son compte. On sélectionnant un nom de cette liste, l'utilisateur peut voir la définition de la requête MDX correspondante.

Ainsi, l'utilisateur peut modifier, renommer, supprimer ou exécuter la requête MDX sélectionnée.

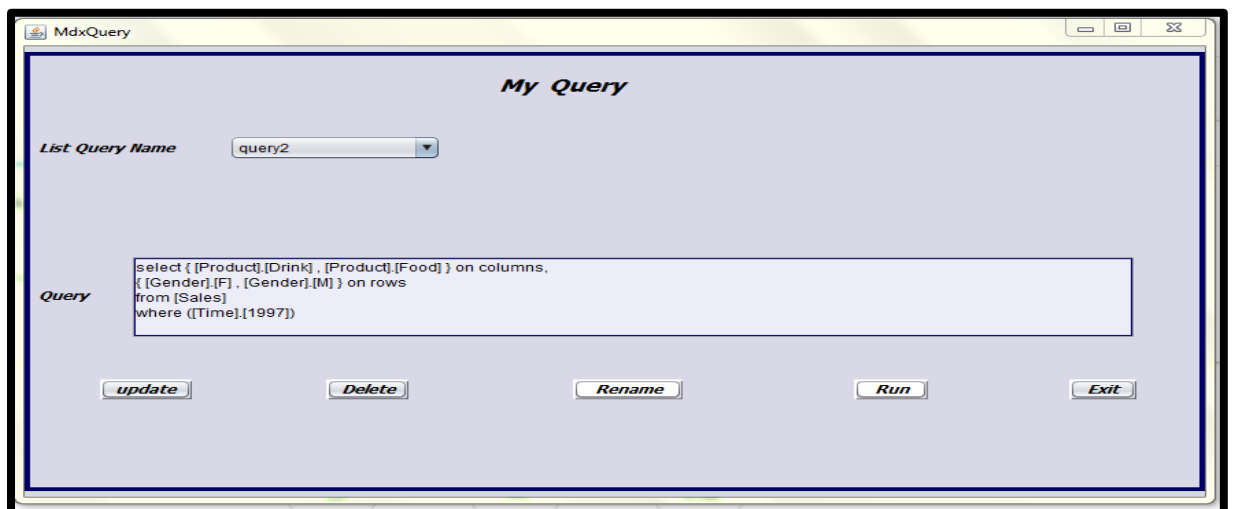


Figure3. 36 Fenêtre my query

### 3.12 Fenêtre modifier une requête

Pour modifier une requête, d'abord l'utilisateur choisit le nom de la requête qu'il souhaite modifier, la requête apparaît dans les champs (*query*), puis l'utilisateur modifie la requête comme il la souhaite et avant de confirmer sa modification, une boîte de confirmation apparaît pour lui, (veut-il vraiment modifier la requête), s'il choisit (oui) les modifications sont sauvegardées.

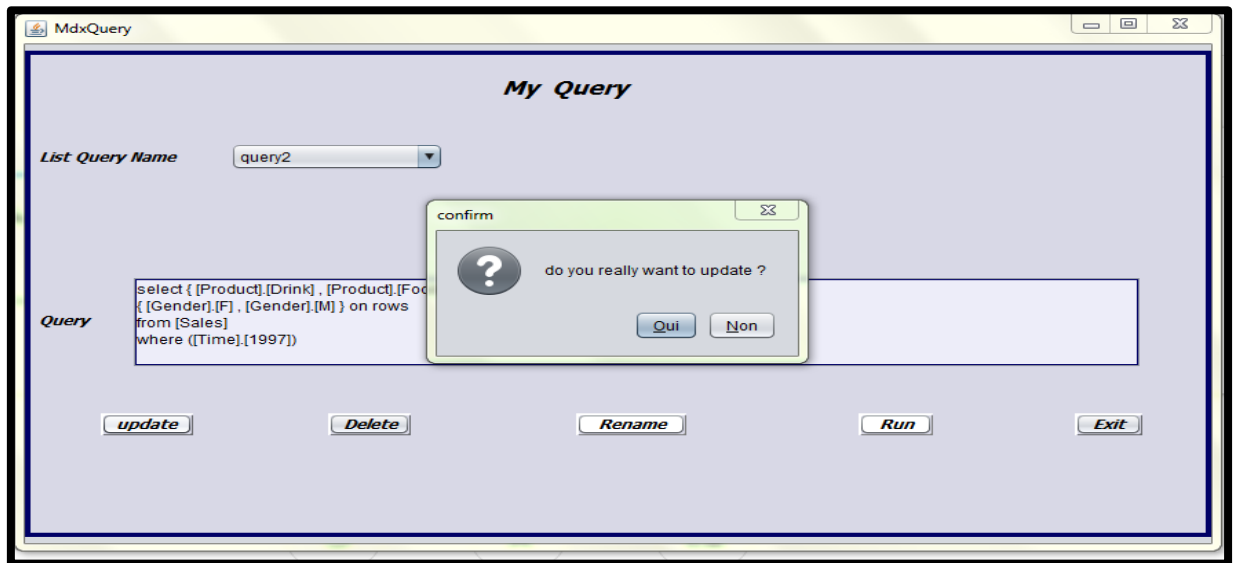


Figure 3.39 Fenêtre modifier une requête

### 3.13 Fenêtre renommer une requête

Cette fenêtre apparaît si l'utilisateur appuie sur le bouton (*Rename*) de l'interface (*my query*), donc il sélectionne le nom de la requête qu'il veut la renommer, et il écrit le nouveau nom dans le champ (*query name*), s'il a confirmé l'action, le nom sera modifié et enregistré dans leur compte.

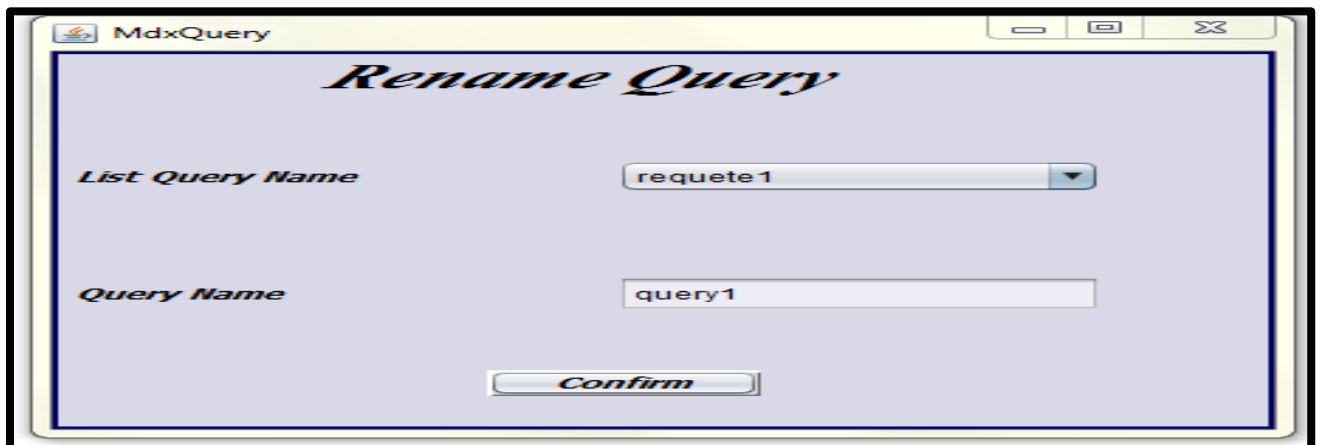
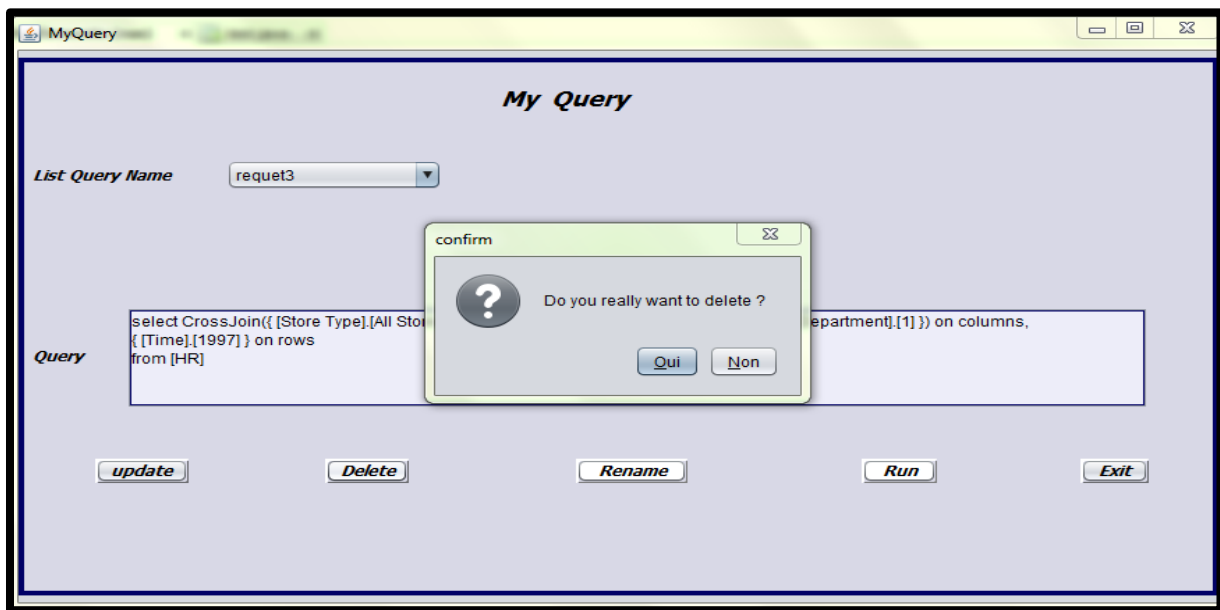


Figure 3.37 Fenêtre renommé le nom d'une requête

### 3.14 Fenêtre supprimer une requête

Afin de supprimer une requête, l'utilisateur choisit le nom de la requête qu'il veut la supprimer, lorsque elle apparaît dans le champ (query), il appuie sur le bouton supprimer, une boîte de confirmation apparaît, s'il a confirmé action par (oui), la requête sera supprimée de son compte.



**Figure 3. 38** Fenêtre supprimer une requête

### Conclusion

Dans ce dernier chapitre nous avons détaillé toutes les étapes liées au développement de notre application et nous avons expliquée exhaustivement les toutes les fonctionnalités (à travers les prises d'écran des fenêtres correspondantes) de notre système.

## Conclusion générale

L'informatique décisionnelle est un domaine en pleine évolution pour l'aide à la prise de décisions aux seins des entreprises. Elle est devenue un besoin capital, car la simple logique de produire pour répondre à une demande, ne suffit plus pour pérenniser l'activité de celle-ci, de ce fait sont apparus les entrepôts de donnée.

Dans un data Warehouse la restitution des données est une phase très importante, car le décideur a besoin de visualiser les résultats d'une manière simple, claire, lui permet de prendre la bonne décision.

Dans le présent travail, nous avons développé une application Java qui permet de construire des requêtes MDX (OLAP), et de l'exécuter sur un entrepôt de données (dans notre cas c'est la base de données de teste Food Mart) pour avoir des résultats sous forme de tables croisées. Pour réaliser ce travail nous avons, tout d'abord, commencé par de l'étude la notion de l'entrepôt de données, son architecture, et comment ses informations sont stockées. Après, nous avons donné un aperçu sur les différents outils d'interrogation d'un entrepôt de données (web et application). Notre application est basée sur la connexion à un entrepôt de données (base de données de teste FoodMart, stockée dans un SGBD relationnel) via un serveur (Mondrian) pour exécuter la requête créer par l'utilisateur et visualiser son résultat sous une forme un peu complexe (table croisée).

Et enfin ce thème a été pour nous une grande opportunité pour développer nos connaissances théoriques sur les entrepôts de données par la recherche documentaire, et nos connaissances pratiques par la programmation.

Dans les perspectives à venir, nous projetons de :

- ✓ Sophistiquer notre interface graphique,
- ✓ Visualiser l'exécution de notre requête sous forme d'un cube réel (visualisation 3D).
- ✓ Rendre notre application plus dynamique, avec la possibilité de supporter la navigation sur plus d'un entrepôt de données.

## Références Bibliographiques



[1] YAHIAOUI Brahim, Contribution du système d'information à la prise de décision : Cas des entreprises de Bejaia

[4] JF Desnos, Entrepôt de données-introduction, Spécialité Double Compétence : informatique et Sciences Sociales, Université de Grenoble

[6] OUHAB Abdallah & Ramtani Tarik , Etude et Configuration d'un entrepôt de données cas hôpital Khelil Amrane

[7] Mise en place d'un data warehouse et d'une application de webmapping pour la gestion du réseau routier. Cas de la république démocratique du Congo

[12] Entrepôts de données Thierry Hamon Bureau H202 Institut Galil\_ee - Universit\_e Paris 13 & LIMSI-CNRS

[18] Olivier TESTE , Modélisation et manipulation d'entrepôts de données complexes et historiées

[24] J. HAN, M. KAMBER, Data Mining: Concepts and Techniques, Simon Fraser University

[26] O. R. ZAIANE, Principles of Knowledge Discovery in Databases, CMPUT690, University of Alberta,

[28] G. CALAS, Études des principaux algorithmes de data mining, Spécialisation Sciences Cognitives et Informatique Avancée, France.

[31] Bernard ESPINASSE, Entrepôts de données : Introduction au langage MDX (Multi-Dimensionale Extensions) pour l'OLAP, Ecole Polytechnique, support Décembre 201

[36] Plioui wahiba , Conception et exploitation d'un entrepôt de donnée au sein de la banque BADR

[37] Pierre gennord , 2008 , UML2 modelisation oriente objet de système logiciel.

## Webographie

- [2] Decision, « [Https://www.decideo.fr/bruley/docs/B374.pdf](https://www.decideo.fr/bruley/docs/B374.pdf) » (consulté le 25/02/2020)
- [3] datawarehouse-comprendre-identifier, « [Https://www.saadrachid.net/bi-big-data/datawarehouse-comprendre-identifier](https://www.saadrachid.net/bi-big-data/datawarehouse-comprendre-identifier) » (consulté le 25/02/2020)
- [5] Objectifsdatawarehouse, « [Https://binieuses.wordpress.com/article/datawarehouse/objectifs-datawarehouse/](https://binieuses.wordpress.com/article/datawarehouse/objectifs-datawarehouse/) » (consulté le 28/02/2020)
- [8] datamart.html, « [Http://www-igm.univ-mlv.fr/~dr/XPOSE2005/entrepot/datamart.html](http://www-igm.univ-mlv.fr/~dr/XPOSE2005/entrepot/datamart.html) » (consulté le 04/03/2020)
- [9] data-mining-definition-exemples, « [Https://www.lebigdata.fr/data-mining-definition](https://www.lebigdata.fr/data-mining-definition) » (consulté le 5/03/2020)
- [10] « [Https://hal.archives-ouvertes.fr/hal-01273589/document](https://hal.archives-ouvertes.fr/hal-01273589/document) », (consulté le 7/03/2020)
- [11] « [Https://stph.scenari-community.org/dwh/mod2.prs/co/mod2UC023.html?mode=html](https://stph.scenari-community.org/dwh/mod2.prs/co/mod2UC023.html?mode=html) », (consulté le 7/03/2020)
- [13] « [Cours.etsmtl.ca](https://cours.etsmtl.ca) », (consulté le 8/03/2020)
- [14] 4395—datawarehouse, « [www.supinfo.com/articles/single/4395—datawarehouse](http://www.supinfo.com/articles/single/4395—datawarehouse) », (consulté le 9/03/2020)
- [15] data-warehouse-modelisation-etoile, « <https://www.piloter.org/business-intelligence/modelisation-datawarehouse.htm> », (consulté le 20/03/2020)
- [16] cubes-rolap-molap-holap-multidimensionnel, « [Https://expert-only.com/concepts/cubes-rolap-molap-holap-multidimensionnel/](https://expert-only.com/concepts/cubes-rolap-molap-holap-multidimensionnel/) », (consulté le 25/03/2020)
- [17] data-warehouse-entrepot-donnees-definition, « [Https://www.lebigdata.fr/data-warehouse-entrepot-donnees-definition](https://www.lebigdata.fr/data-warehouse-entrepot-donnees-definition) », (consulté le 7/04/2020)
- [19] « [Https://www.decideo.fr/Les-cubes-OLAP-Une-revolution-pour-le-reporting\\_a9149.html](https://www.decideo.fr/Les-cubes-OLAP-Une-revolution-pour-le-reporting_a9149.html) », (consulté le 20/04/2020)
- [20] Cube-OLAP, « [Https://whatis.techtarget.com/fr/definition/Cube-OLAP](https://whatis.techtarget.com/fr/definition/Cube-OLAP) », (consulté le 24/04/2020)
- [21] informatique\_decisionnelle\_olap, « [Http://igm.univmlv.fr/~dr/XPOSE2009/informatique\\_decisionnelle\\_olap/colap.html](http://igm.univmlv.fr/~dr/XPOSE2009/informatique_decisionnelle_olap/colap.html) », (consulté le 24/04/2020)
- [22] « [Https://fr.wikiversity.org/wiki/Datamining](https://fr.wikiversity.org/wiki/Datamining) », (consulté le 2/05/2020)
- [23] Datamining.htm, « [Https://www.piloter.org/business-intelligence/datamining.htm](https://www.piloter.org/business-intelligence/datamining.htm) », (consulté le 15/05/2020)

- [25] « [Http://dictionnaire.sensagent.leparisien.fr](http://dictionnaire.sensagent.leparisien.fr) », (consulté le 18/05/2020)
- [27] « [Https://www.ceremade.dauphine.fr/SODAS/EXEMPLES/](https://www.ceremade.dauphine.fr/SODAS/EXEMPLES/) », (consulté le 25/05/2020)
- [29] « [Https://mrmint.fr/](https://mrmint.fr/) », (consulté le 15/06/2020)
- [30] « [Https://searchbusinessanalytics.techtarget.com/definition/ad-hoc-analysis

\[32\] « \[Https://perso.univ-lyon1.fr/haytham.elghazel/BI/presentation.html\]\(https://perso.univ-lyon1.fr/haytham.elghazel/BI/presentation.html\) », \(consulté le 30/06/2020\)

\[33\] « \[Http://germanlinux.blogspot.com/2009/06/cubes-mondrian-et-ses-clients.html\]\(http://germanlinux.blogspot.com/2009/06/cubes-mondrian-et-ses-clients.html\) », \(consulté le 7/07/2020\)

\[34\] « \[Https://blog.developpez.com/talend/p2947/technologie/navigation\\\_multidimensionnelle\\\_avec\\\_jpi\]\(https://blog.developpez.com/talend/p2947/technologie/navigation\_multidimensionnelle\_avec\_jpi\) », \(consulté le 20/07/2020\)

\[35\] « \[Http://rubik.sourceforge.net/jrubik/intro.html\]\(http://rubik.sourceforge.net/jrubik/intro.html\) », \(consulté le 27/07/2020\)

\[38\] « \[www.editionseni.fr/open/mediabook.aspx?idR=98db25384380c51b79e46e170641bf45\]\(http://www.editionseni.fr/open/mediabook.aspx?idR=98db25384380c51b79e46e170641bf45\) », \(consulté le 20/08/2020\)

\[39\] « \[Http://www.standard-du-web.com/easyphp.php\]\(http://www.standard-du-web.com/easyphp.php\) », \(consulté le 27/08/2020\)](https://searchbusinessanalytics.techtarget.com/definition/ad-hoc-analysis)