# Project Report

# Pharmacy Management BI Solution


Realized by:

Ben Dahmen Ghazi
Chandoul Wejdane
Ghdiri Oussama
Khemiri Achref
Sabbagh Wiem

Supervised by:

Mrs. TRABELSI Dorra
Mrs. KALLEL Rania

# Table of contents

## Chapter 1 : Introduction


I. Project's context: ..... 2
1. Introduction ..... 2
2. Business Intelligence Solution ..... 2
II. Data source identification and description: ..... 3
III. Business objectives: ..... 6
IV. Data warehouse modeling ..... 7
1. Data warehouse ..... 7
2. Fact and dimension ..... 7
3. Data warehouse model ..... 7


## Chapter 2 : Data Integration


I. Used tools ..... 10
1. Business Intelligence Development Studio ..... 10
2. Python ..... 10
3. Jupyter Notebook ..... 11
4. MonogoDB ..... 11
II. Internal Data Integration ..... 11
1. Extract from source ..... 12
a. Staging area ..... 12
b. Staging Area Structure ..... 12
2. Transform the data ..... 14
3. Load the data ..... 14
a. Data warehouse ..... 14
b. Dimension tables ..... 14
c. Fact tables ..... 16
III. External Data Integration: ..... 17
1. Web Scraping ..... 17
2. Nosql Database ..... 18
3. Examples of external data ..... 18


## Chapter 3 : Data analysis and data visualization


I. OLAP Cube ..... 21

1. The cube's structure ..... 21
2. The dimensions' structure ..... 22
3. Hierarchies ..... 23
4. Calculations ..... 23
II. Data visualization ..... 24
1. Used tools ..... 24
2. Dashboard ..... 25
III. Data Mining ..... 25
1. Used Tools ..... 25
2. Association rule mining: ..... 27

**Chapter 4 : Conclusion and Perspectives**

Conclusion ..... 29

# List of illustarations

Figure 1: business intelligence architecture ..... 2
Figure 2 : Data warehouse model ..... 8
Figure 3: Sql server logo ..... 10
Figure 4: Sql server data tools logo ..... 10
Figure 5: Python logo ..... 11
Figure 6: Jupyter notebook and lab logos ..... 11
Figure 7: MongoDB logo ..... 11
Figure 8: Staging area structure ..... 12
Figure 9: Control flow of the table client in SA ..... 13
Figure 10: Data flow of the table client in SA ..... 13
Figure 11: Control flow of the table Article in SA ..... 13
Figure 12: Data flow of the table Article in SA ..... 14
Figure 13: Data flow of the table fournisseur in the data warehouse ..... 15
Figure 14: Data flow of the table article in the data warehouse ..... 15
Figure 15: Data flow of the fact stock ..... 16
Figure 16: Data flow of the fact vente ..... 16
Figure 17: Data warehouse structure ..... 17
Figure 18: MongoDb Explorer ..... 19
Figure 19: MongoDB explorer ..... 19
Figure 20: olap cube ..... 21
Figure 21: cube's solution explorateur ..... 21
Figure 22: Data sources' view ..... 22
Figure 23: dimensions' structure ..... 22
Figure 24: time dimension's hierarchy ..... 23
Figure 25: cube calculations ..... 23
Figure 26: cube calculations ..... 24
Figure 27: logo of SSRS ..... 24
Figure 28: logo of powerBi ..... 24
Figure 29: part of the dashboard ..... 25
Figure 30: the mobile application ..... 26
Figure 31: the mobile application ..... 27

# Chapter 1:

# Introduction

# I. Project's context:

## 1. Introduction

Pharmacies are today facing many challenges in terms of government regulation and competition, they are also facing many financial issues that can influence their business such as:

- Medicines shortage especially for chronic diseases
- Rise in price of medicines due to the weakness of the Tunisian currency
- The delay in payment of CNAM

## 2. Business Intelligence Solution

**Business Intelligence (BI)** refers to the tools, technologies, applications and practices used to collect, integrate, analyze, and present an organization's raw data in order to create insightful and actionable business information.
A business intelligence architecture is a framework for organizing the data, information management and technology components that are used to build business intelligence (BI) systems for reporting and data analytics.

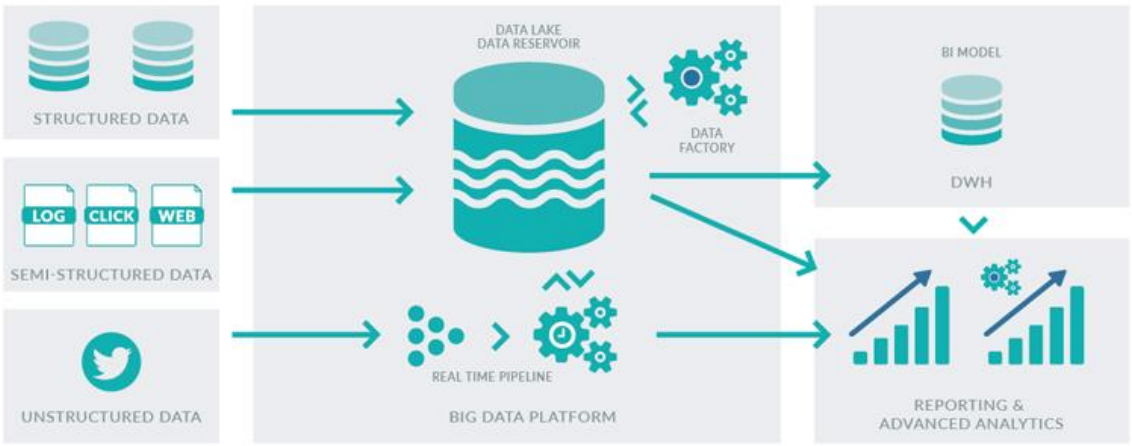


*Figure 1: business intelligence architecture*

Nowadays, many companies chose to use business intelligence solutions to analyze their data so that they can operate more efficiently and overcome their issues.

In this context, the Pharmacy of ***Mrs Amel Bey*** , a pharmacy that has 10 employees, located in Tunisia and more specifically in a suburb of Tunis: Kram,

is today making big investments in implementing BI solutions to improve its services.

The potential benefits of business intelligence solution are :

- Accelerating and making the best business decisions
- Spot inefficient business processes
- Optimizing internal business processes
- Increasing operational efficiency
- Driving new revenues by identifing new business opportunities
- Spotting business problems that need to be addressed

## II. Data source identification and description:

Our client gave us a file named 'pharma.bak' which is a backup file of the phamracy's database. It contains more than 158 tables described in the next table:

| | ***Tables*** | ***Description*** |
|---|---|---|
| Purchases | Achat | These tables provide information about the purchases made |
| | Achat détails | |
| | Achat Instance | |
| | Achat Instance détails, | |
| | Achat supprimés | |
| | Achat supprimés détails | |
| | Commande_Details | |
| | Commande_instance | |
| | Commande_instance_detail | |
| | Commande_instantannée | |
| | Fournisseur | |
| Sales | Client | These tables provide the sales and clients |
| | VENTE | |
| | VENTE_DETAILS | |
| | VENTE_HISTORIQUE | |
| | VENTE_HISTORIQUE_DETAIL | |
| | VENTE_DETAIL_IMAGE | |
| | VENTE_INSTANCE | |
| | VENTE_INSTANCE_DETAIL | |
| | VENTE_NUMERO | |
| | VENTE_SUPPRIME | |
| | VENTE_SUPPRIME_DETAILS | |
| | VENTE_ANNULER | |
| | ANNULER_DETAIL | |
| | Fiche patient | |

| | | |
|---|---|---|
| | Client_information | |
| | Client_famille | |
| | Client | |
| | Fiche patient | |
| | Client_information | |
| | SITUATION_CLIENT | |
| | Activation_ Client | |
| | Client_famille | |
| | SITUATION_CLIENT | |
| | Lien parenté | |
| Settlements | Facturation_ Client | These tables describe all settlements. |
| | Facturation _JOURNALIERE | |
| | Activation_ Client | |
| | Changement de prix | |
| | Règlement client | |
| | Règlement client numéro | |
| | Règlement client vente | |
| | Règlement mutuelle vente | |
| | Règlement mutuelle | |
| | Règlement fournisseur | |
| | Règlement fournisseur achat | |
| | TVA | |
| | Pret | |
| | TVA | |
| | Règlement CNAM | |
| | Règlement CNAM VENTE | |
| Products | Article | All tables provide information about products. |
| | Article_ignoré | |
| | Article_Mise à jour | |
| | Article_modification | |
| | Article_non_remboursable | |
| | Article_remboursable | |
| | Article_surveillé | |
| | Article1 | |
| | SITUATION_ARTICLE | |
| | Fractionnement | |
| | Catégorie | |
| | DCI | |
| | Niveau de gravité | |
| | Forme Article | |
| | FORMULE_PREPARATION | |
| | PRODUCTION_DETAILS_FORMULE PRODUCTION_DETAILS | |
| | Lot article | |

| | | |
|---|---|---|
| | ENTREE | |
| | ENTREE_DETAILS | |
| | NATURE_ENTREE | |
| | SORTIE_détail | |
| | Interaction | |
| Confidentiality | Autorisation | The table AUTORISATION provides information about the ability of each employee |
| | Profil_détails/ Points_contrôle | These tables provides every function done by employees like insert ,delete, updating…with a login |
| | ADRESSE_MAC | |
| | Historique update | |
| | Historique caisse | |
| | Historique accès | |
| | Log | These tables are used to provide the security of system. |
| | MODULES_SURVEILLANCE | |
| | Message Error | |
| Payment | Banque | This tables provides informations about Bank and checkout that are important for every sale or purchase |
| | Bon | |
| | Caisse | |
| | Emprunt | |
| | Emprunt_détails | |
| | Nature Règlement | |
| | Différence recap_caisse | |
| | Mouvement état | |
| Stock | Entrée | These tables describe stock and movement. |
| | Entrée_détails | |
| | Nature entrée | |
| | Nature sortie | |
| | Mouvement_article | |
| | Mouvement_lot_article | |
| | Inventaire | |
| | Inventaire_détails | |
| | SIMULATION_STOCK | |
| | STOCK_ANTERIEUR | |
| | SIMULATION_STOCK_DETAILS | |
| Reporting | MOIS_ANNEE_POUR_AFFICHAGE_ETAT | These tables provide some statics |
| | MOIS_ANNEE | |
| | STATISTIQUE_ARTICLE_ | |
| | STATISTIQUE_ARTICLE2011 | |

| | | |
|---|---|---|
| Employees | Parametre | These tables provide information about employees |
| | Preparation | |
| | Indemnité preparation | |
| | Ordonnancier | |
| | Ordonnancier ancien | |
| | Poste | |
| | Profil détails | |
| | Profil | |
| Collaborators | Medecin | These tables provide information about their collaborators |
| | SPECIALITE_MEDECIN | |
| | LABORATOIRE | |
| Insurance | Mutuelle | These tables contain information about the insurance of clients |
| | Mutuelle_Article | |
| | Mutuelle_Catégorie | |
| | Catégorie CNAM | |
| | Pharma_ODS_Mutuelle | |
| | APCI | |
| Orders | Commande | These tables describe orders realized. |
| | Commande Détail | |
| | Commande_instance | |
| | Commande_instance details | |
| | Commande_instanceanne | |
| | Mouvement commande | |

## III. Business objectives:

Survival and consistent profitability are major business goals for local businesses, and pharmacies are no exception so we tried to fix our Business objectives in order to guarantee a better performance:

- The bestseller / worst seller products by period in terms of Quantities and category.
- The most beneficial product.
- The most beneficial category of products.
- The bestselling category of products.
- The products that are usually sold together.
- The highest or the lowest profitability period per year.
- The number of sales per hour.
- The evolution of the number of clients with insurance.

- The favorite supplier.
- The disponibility of stock per product to overcome stock shortage.
- The shortest period of stock depletion per product.

# IV. Data warehouse modeling

## 1. Data warehouse

A data warehouse is a collection of data supporting management decisions. The data is subject oriented, integrated, nonvolatile, and time variant.
The data warehouse is the collection of snapshots from all of the operational environments and external sources. It is not up to the minute accurate, because it has to be extracted from the operational environment, which is done on a regular time base.

## 2. Fact and dimension

Facts and dimensions form the core of any business intelligence effort. These tables contain the basic data used to conduct detailed analyses and derive business value.

**What Are Facts and Facts Tables?**

Fact tables contain the data corresponding to a particular business process. Each row represents a single event associated with a process and contains the measurement data associated with that event.
The information contained within a fact table is typically numeric data, and it is often data that can be easily manipulated, particularly by summing together many thousands of rows.

**What Are Dimensions and Dimensions Tables?**

Dimensions describe the objects involved in a business intelligence effort. While facts correspond to events, dimensions correspond to people, items, or other objects.
Dimension tables contain details about each instance of an object. For example, the items dimension table would contain a record for each item sold in the store. It might include information such as the cost of the item, the supplier, color, sizes, and similar data.

## 3. Data warehouse model

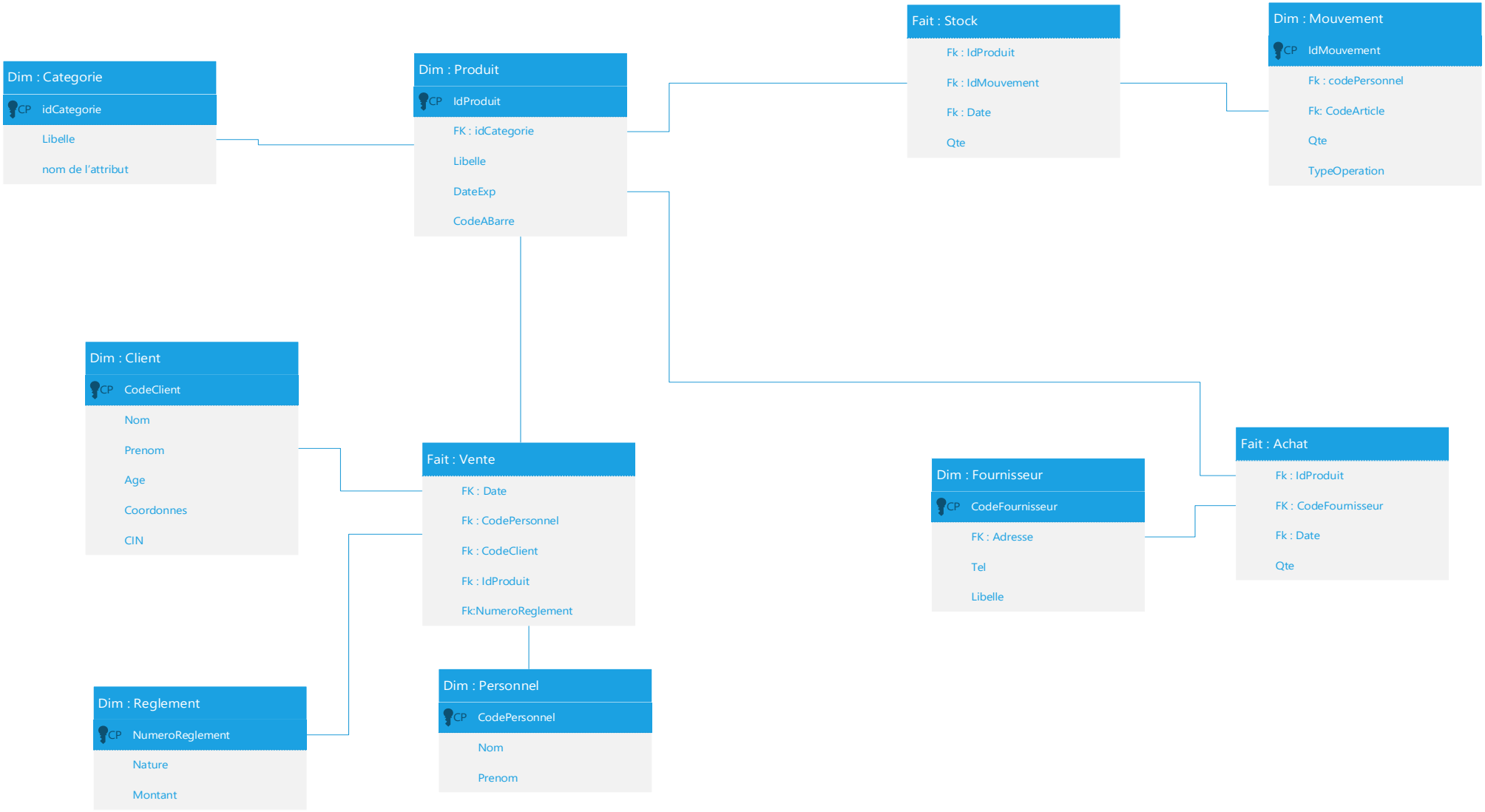


*Figure 2 : Data warehouse model*

# Chapter 2:

# Data Integration

# I. Used tools

## 1. Business Intelligence Development Studio

Business Intelligence Development Studio (BIDS) is the former IDE from Microsoft and was used to develop data analysis and Business Intelligence solutions utilizing the Microsoft SQL Server Analysis Services, Reporting Services and Integration Services. It is based on the Microsoft Visual Studio development environment but customized with the SQL Server services-specific extensions and project types, including tools, controls and projects for reports, ETL dataflows, OLAP cubes and data mining structure.

### a. SQL Server Management Studio

SQL Server Management Studio (SSMS) is an integrated environment for managing any SQL infrastructure. Use SSMS to access, configure, manage, administer, deploy and develop all components of SQL Server and SQL Data Warehouse.


*Figure 3: Sql server logo*

### b. SQL Server Data Tools

SQL Server Data Tools **(**SSDT) is a modern development tool to build SQL Server relational databases, Integration Services packages, Analysis Services data models, and Reporting Services reports. With SSDT, you can design and deploy any SQL Server content type easily.


*Figure 4: Sql server data tools logo*

## 2. Python

Python is an interpreted high-level programming language for general-purpose programming. Created by Guido van Rossum and first released in 1991, Python has a design philosophy that emphasizes code readability, notably using significant whitespace. It provides constructs that enable clear programming on both small and large scales.

Python features a dynamic type system and automatic memory management. It supports multiple programming paradigms, including object-oriented, imperative, functional and procedural, and has a large and comprehensive standard library.


*Figure 5: Python logo*

### 3. Jupyter Notebook

Jupyter notebooks are documents that combine live runnable code with narrative text (Markdown), equations (LaTeX), images, interactive visualizations and other rich output. Jupyter notebooks (.ipynb files) are fully supported in JupyterLab.


*Figure 6: Jupyter notebook and lab logos*

### 4. MonogoDB

MongoDB is a free and open-source cross-platform document-oriented database program. Classified as a NoSQL database program, MongoDB uses JSON-like documents with schemata. MongoDB is developed by MongoDB Inc., and is published under a combination of the Server Side Public License and the Apache License.


*Figure 7: MongoDB logo*

## II. Internal Data Integration

Data integration is the combination of technical and business processes used to combine data from disparate sources into meaningful and valuable information. A complete data integration solution delivers trusted data from various sources mainly the production database.

ETL is an abbreviation of the three words Extract, Transform and Load. It is an ETL process to extract data, mostly from different types of systems, transform it

into a structure that's more appropriate for reporting and analysis and finally load it into the database and or cube(s).

## 1. Extract from source

In this step we extract data from different internal and external sources, structured and/or unstructured. Plain queries are sent to the source systems, using native connections, message queuing, ODBC or OLE-DB middleware. The data will be put in a so-called Staging Area (SA), usually with the same structure as the source. In some cases we want only the data that is new or has been changed, the queries will only return the changes.

### a. Staging area

The Data Warehouse Staging Area is temporary location where data from source systems is copied. A staging area is mainly required in a Data Warehousing Architecture for timing reasons.

### b. Staging Area Structure

After the Data Comprehension phase we had to setup into our own Staging Area which has the following structure:

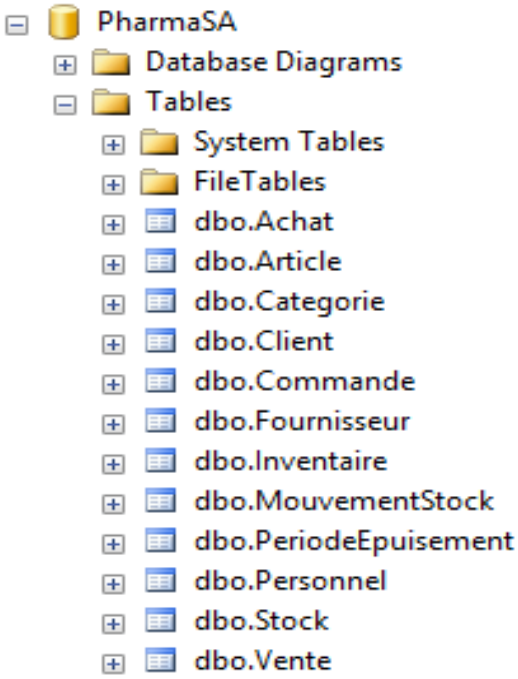


*Figure 8: Staging area structure*

The extraction of the data from the information system to the Staging Area was made thanks to this control flow:

- Example (Staging Area of table client):

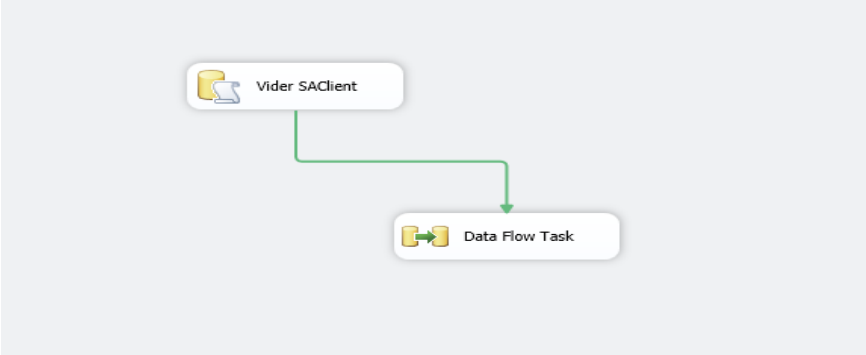


*Figure 9: Control flow of the table client in SA*

➔ We used an "Execute SQL Task" component to delete the content of the table client using an SQL query before executing the "Data Flow Task"

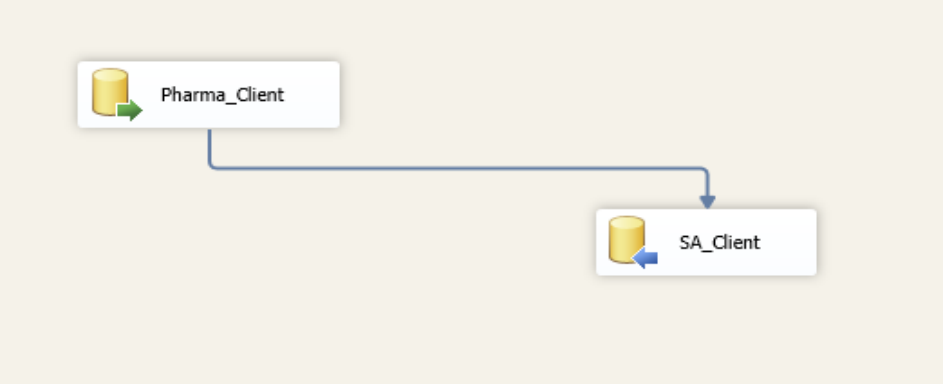


*Figure 10: Data flow of the table client in SA*

➔ We used an "OLDB source" to charge the data from the database "Pharma" and we put it into the database "PharmaSA" using an "OLDB destination"

- Example (Staging Area of table Article):

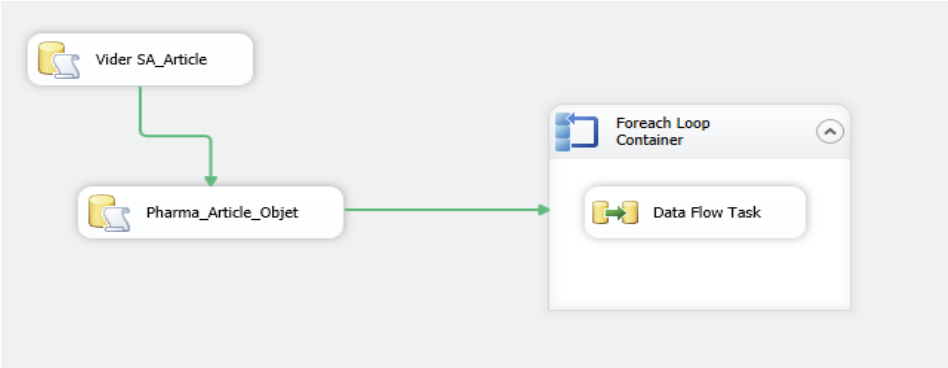


*Figure 11: Control flow of the table Article in SA*

➔ We used an “Execute SQL Task” component to delete the content of the table Article using an SQL query and then another “Execute SQL Task” component to extract all the article using the primary key into a result set. After that we will use the result set to look over all the article using a “Foreach Loop Container” to delete the duplications.

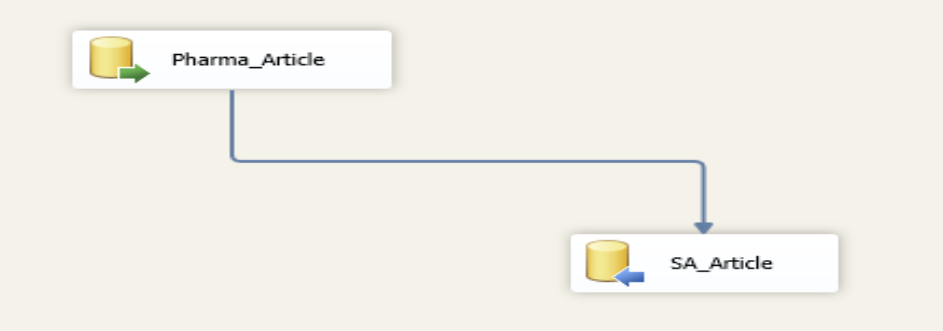


*Figure 12: Data flow of the table Article in SA*

➔ We used an “OLDB source” to charge the data from the database by executing a sql query “Pharma” and we put it into the database “PharmaSA” using an “OLDB destination”.

### 2. Transform the data

Once the data is available in the Staging Area, it is all on one platform and one database. So we can easily join and union tables, filter and sort the data using specific attributes, pivot to another structure and make business calculations. In this step of the ETL process, we can check on data quality and cleans the data if necessary.

### 3. Load the data

Once the data is available in the Staging Area, it is all on one platform and one database. So we can easily join and union tables, filter and sort the data using specific attributes, pivot to another structure and make business calculations. In this step of the ETL process, we can check on data quality and cleans the data if necessary.

#### a. Data warehouse

A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process.

#### b. Dimension tables

***Supplier dimension***

We extracted the data we need from the existing SAFournisseur source in the Staging Area. We checked with the Dim_Fournisseur already existing lines

using “Lookup” component, if there is a match between the two tables, we update the table otherwise we add the non-match to DimFournisseur table in the data warehouse.

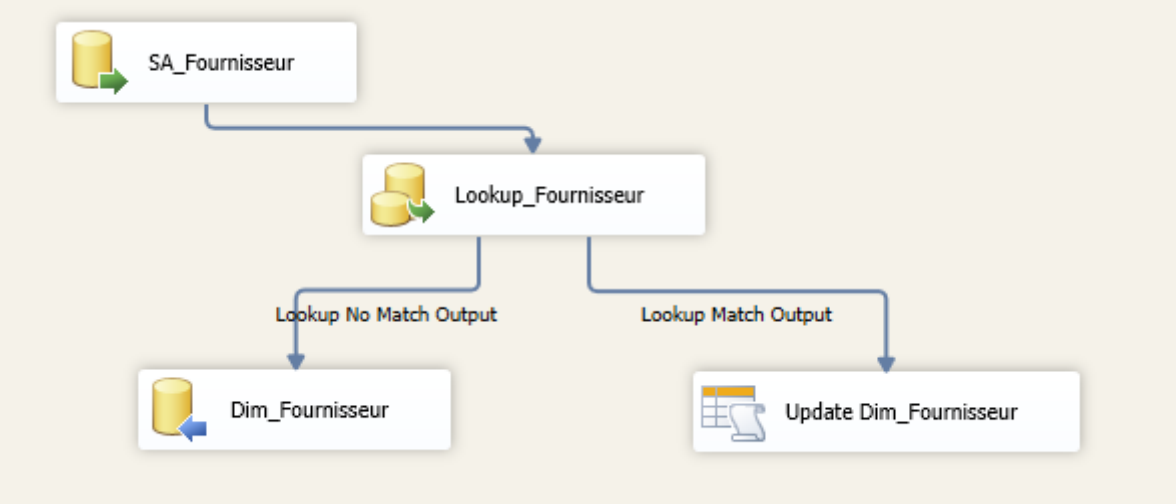


*Figure 13: Data flow of the table fournisseur in the data warehouse*

## *Product dimension*

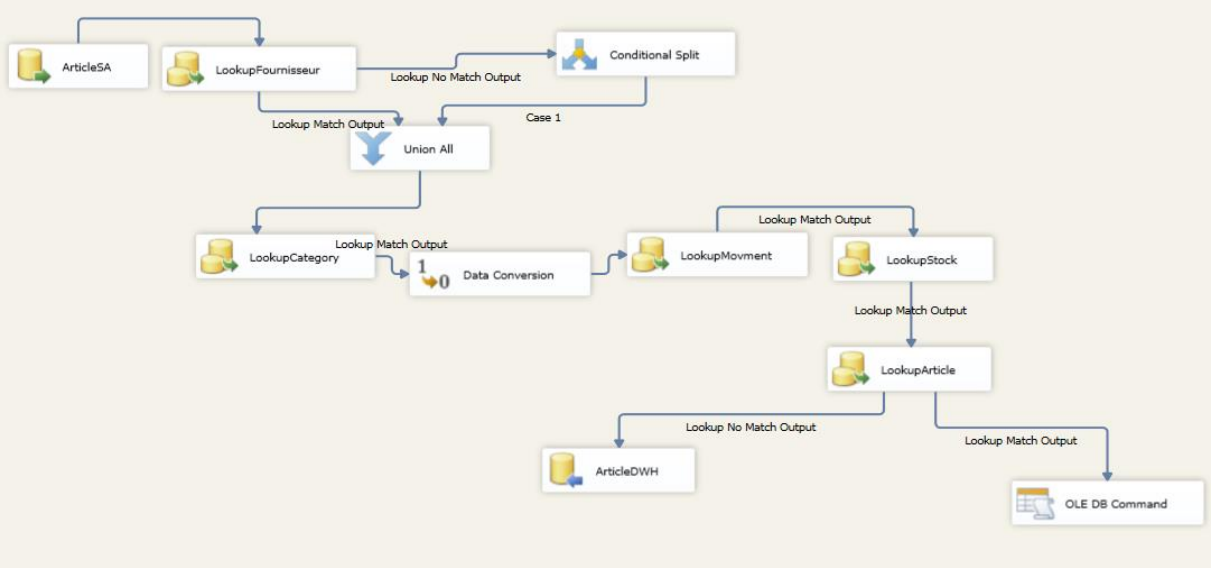


*Figure 14: Data flow of the table article in the data warehouse*

We extracted the data we need from the existing table ArticleSA source in the Staging Area. We checked if the supplier of the products exists or no using a “Lookup” with the table Fournisseur then the same with table category. And using the lookupMovment and lookupStock we updated the inventory and the last sold product. Finally, we check using “Lookup” component, if there is a match between the two tables, we update the table otherwise we add the non-match to DimArticle table in the data warehouse.

### c. Fact tables

Our data Warehouse is based on the fact that we have 3 different themes as we chose to work on 3 fact tables which are:
« fact_Stock » : For the Stock axis
« fact_Achat » : For the Sales axis
« fact_Vente » : For the Purchases axis

***Fact Stock:***

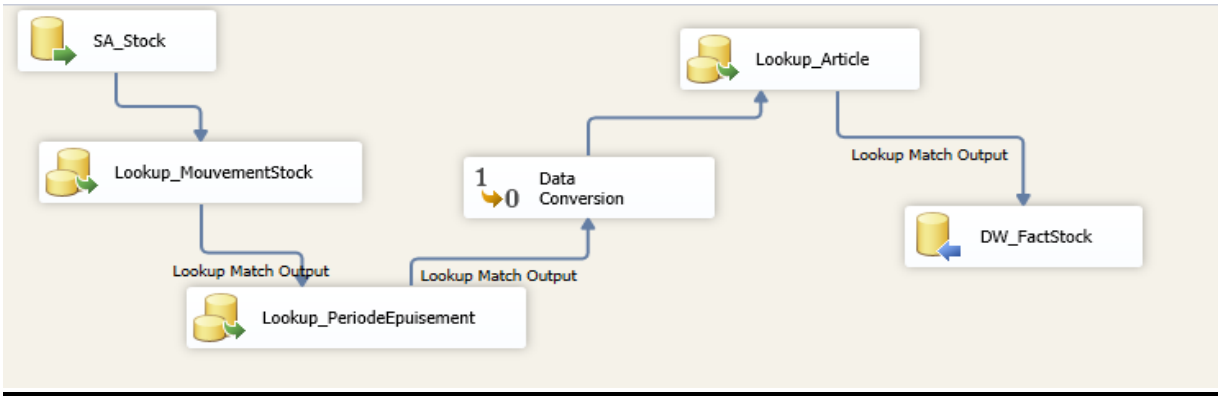


*Figure 15: Data flow of the fact stock*

-We extracted the data we need from the existing table SA_Stock source in the Staging Area
- At the data conversion component, the fields that are not compatible with the destination, so they have been converted.
- We checked with the Dim_MouvementStock if the movement already exists.
- We checked with the Dim_Article if the product already exists.

***Fact Vente:***

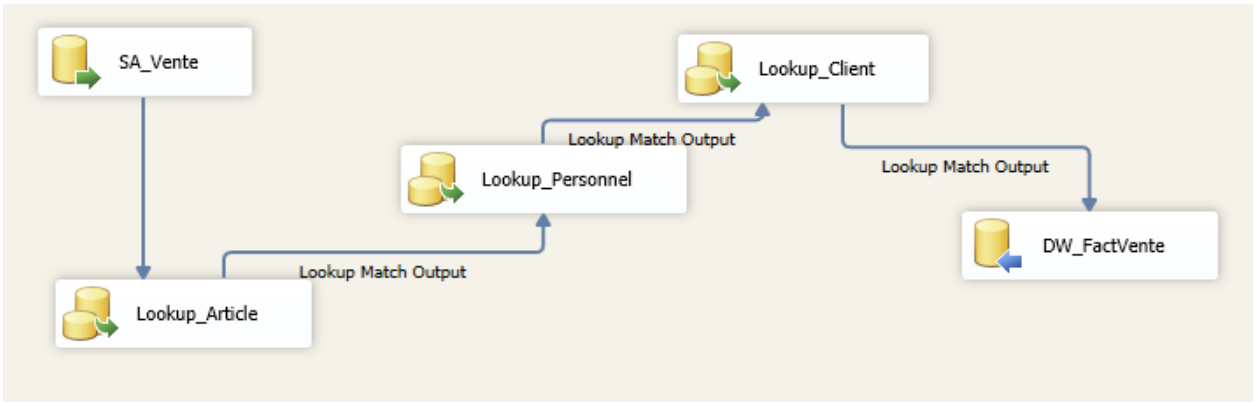


*Figure 16: Data flow of the fact vente*

-We extracted the data we need from the existing table SA_vente source in the Staging Area
- We checked with the Dim_Client already existing lines.
- We checked with the Dim_Article already existing lines.
- We checked with the Dim_Personnel already existing lines.

***Data ware house sturcture***

Finally, we got the Data Warehouse structure like the figure below:

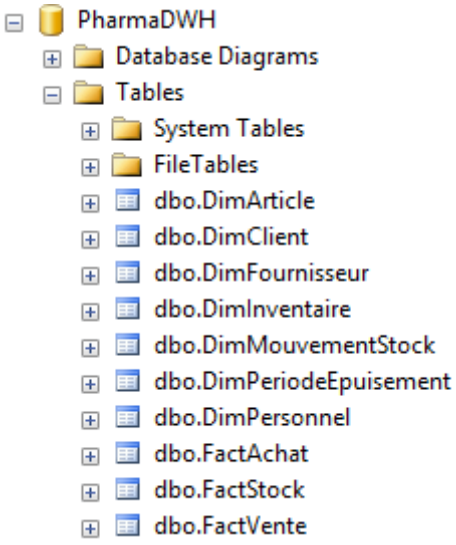


*Figure 17: Data warehouse structure*

After having gone through the entire Data preparing process : identifing the data, transforming it, building a data model, loading the data into a Data warehouse and verifing the results, comes the next step : Starting the analysing process.
Data analysis, also known as analysis of data or data analytics, is a process of inspecting data with the goal of discovering useful information.

# III. External Data Integration:

## 1. Web Scraping

Web scraping, web harvesting, or web data extraction is data scraping used for extracting data from websites. Web scraping software may access the World Wide Web directly using the Hypertext Transfer Protocol, or through a web browser. While web scraping can be done manually by a software user, the term typically refers to automated processes implemented using a bot or web crawler. It is a form of copying, in which specific data is gathered and copied from the web, typically into a central local database or spreadsheet, for later retrieval or analysis.

## 2. Nosql Database

NoSQL database provides a mechanism for storage and retrieval of data that is modeled in means other than the tabular relations used in relational databases. NoSQL databases are increasingly used in big data and real-time web applications. NoSQL systems are also sometimes called "Not only SQL" to emphasize that they may support SQL-like query languages, or sit alongside SQL database in a polyglot persistence architecture.

***Why MongoDB ?***

MongoDB is designed to meet the demands of modern apps with a technology foundation that enables you through:

1. The document data model: presenting you the best way to work with data.
2. A distributed systems design: allowing you to intelligently put data where you want it.
3. A unified experience that gives you the freedom to run anywhere: allowing you to future-proof your work and eliminate vendor lock-in.

## 3. Examples of external data

Using internal and external data is an essential element in good decision-making. Pharmacies that use more internal and external data sources possess a greater range of possibilities for data analysis. As we are done with internal data we have tried to collect the most useful external data.

- **Getting more suppliers:**

Source: www.santetunisie.rns.tn

Our internal data contains some suppliers but they don’t seem to be enough . We gathered more suppliers so the pharmacy could switch to them for more medecines.

We added them into a “fournisseurs” collection in our mongo database.

We use mongo explorer to visualize our mongo db database.

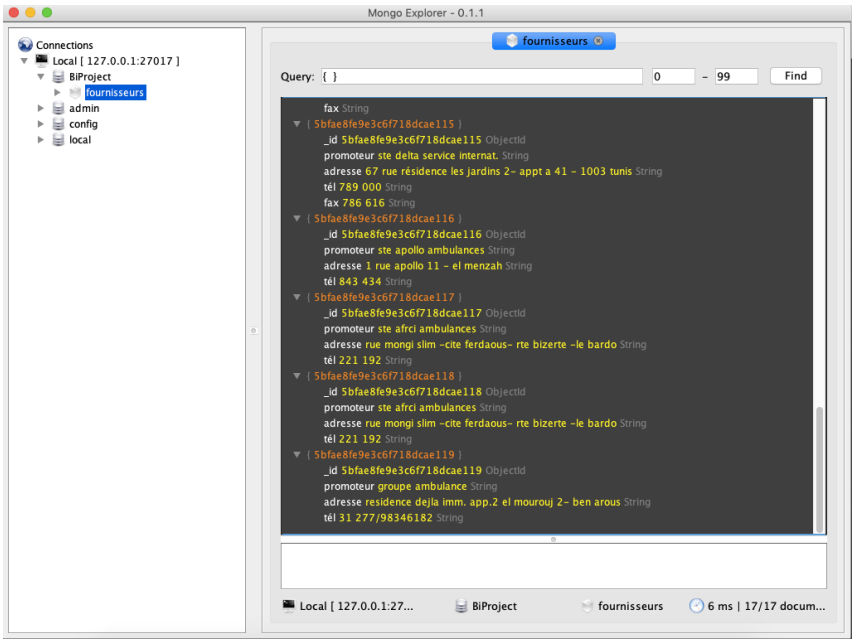


*Figure 18: MongoDb Explorer*

- **Getting medecines list**

Source : [www.pharmanity.com](http://www.pharmanity.com)

We added a list of 9542 medecines sold in france and converted their price from euros to dinars in case the pharmacy needs to import some of them and extend its range of products.
We added them into a “medicaments” collection in our mongo database.

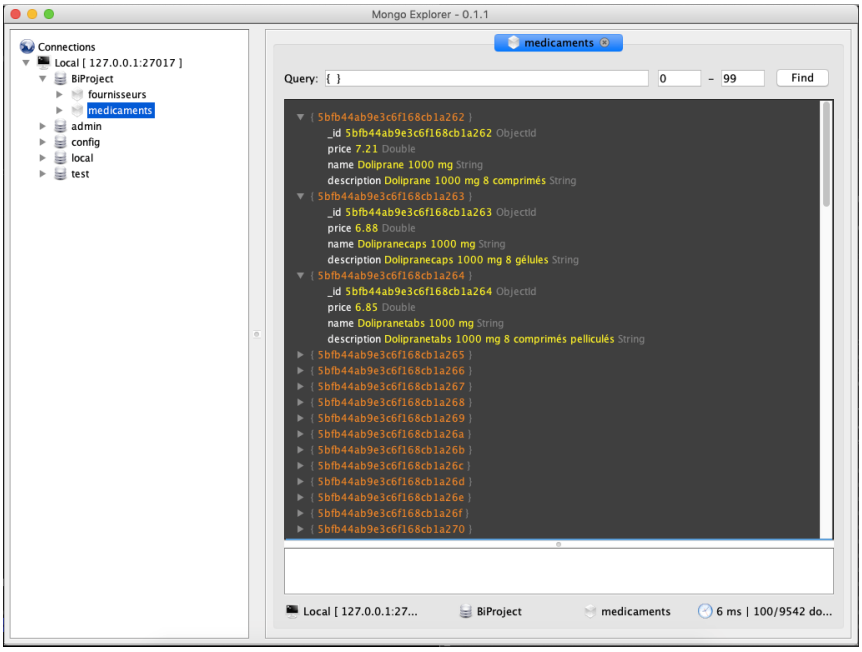


*Figure 19: MongoDB explorer*

# Chapter 3:

# Data Analysis and data visualization

# I. OLAP Cube

OLAP cube, an acronym for OnLine Analytical Processing, is a technology that stores data (data warehouse or other relational database) in an optimized way to provide a quick response to various types of complex queries by using dimensions and measures. Also, it provides the users to do their own analysis and reporting easily.

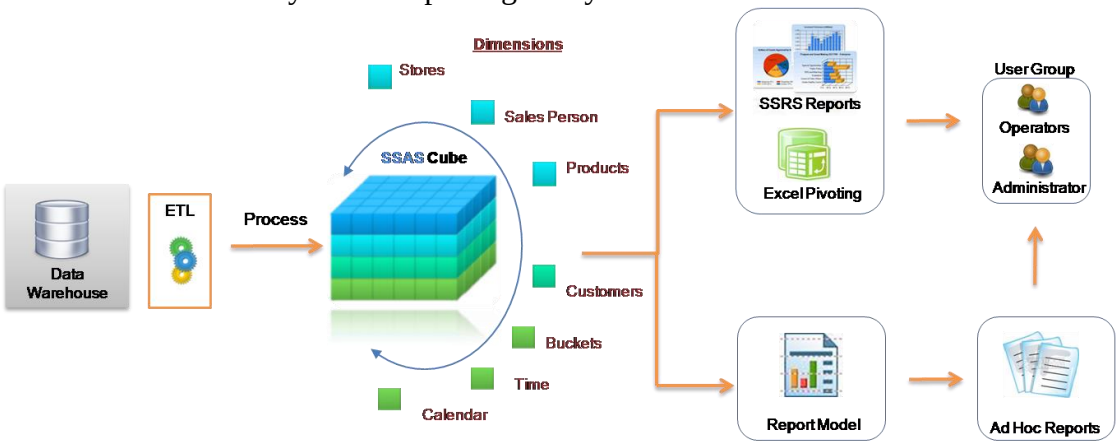


*Figure 20: olap cube*

OLAP offers some key benefits such as:

- Business-focused multidimensional data
- Business-focused calculations
- Trustworthy data and calculations
- Speed-of-thought analysis
- Flexible, self-service reporting

## 1. The cube's structure

We used SSAS to create the cube first by creating a new data source based on our datawarehouse then we created a new data source view based on the data source already loaded by selecting the dimensions and facts and finally, we created the olap cube.

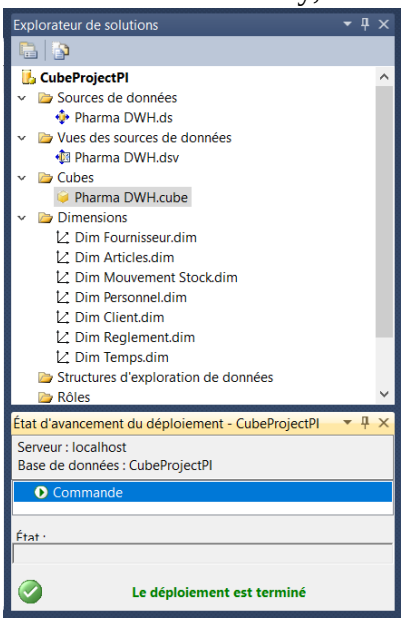


*Figure 21: cube's solution explorateur*

Here we can see that the view is the same as the data warehouse's model suggested at the beginning.

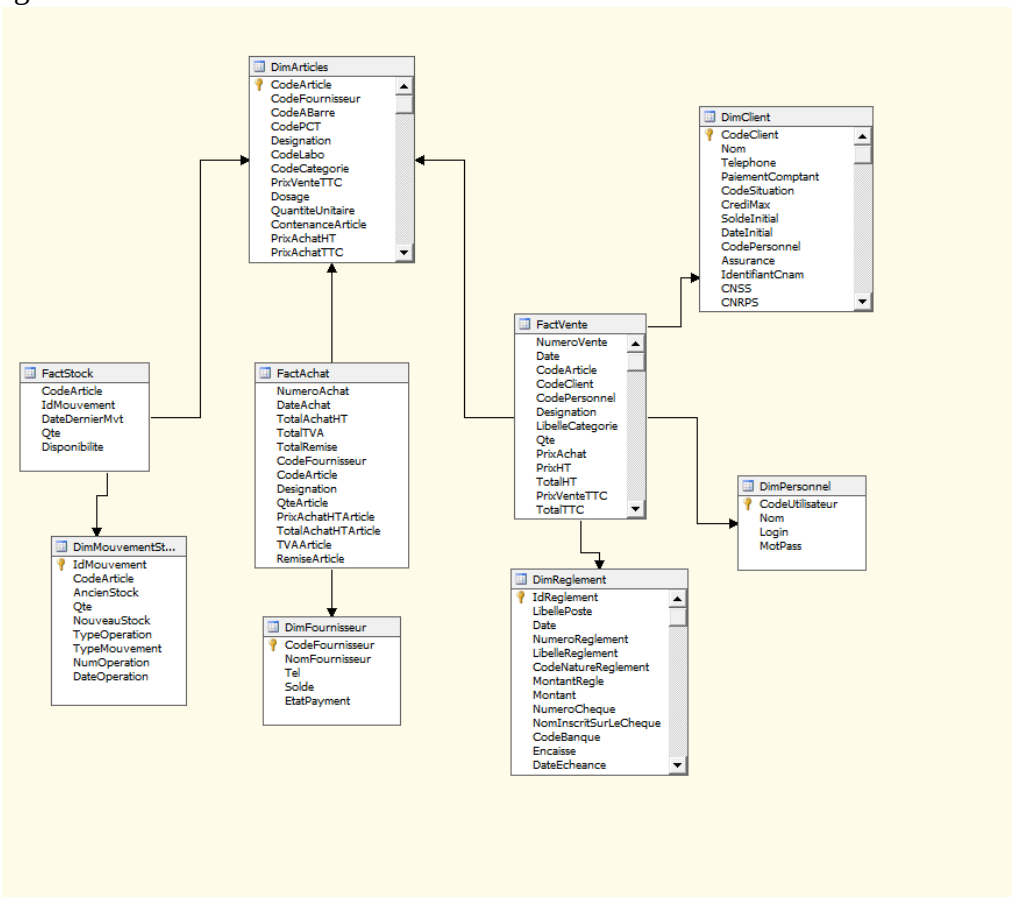


*Figure 22: Data sources' view*

## 2. The dimensions' structure

After the creation of the cube we notice that all the dimensions were added under the dimension folder. Afterwards we modified all the dimension depend on our need.

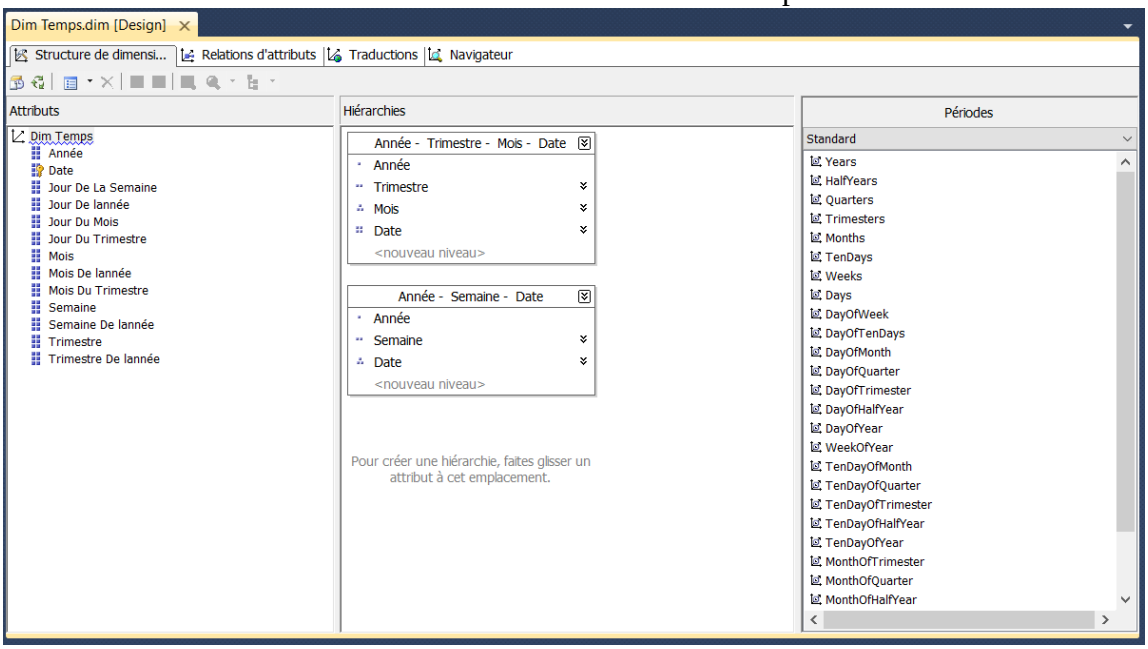


*Figure 23: dimensions' structure*

## 3. Hierarchies

A hierarchy lets you view aggregated fact data at multiple levels that's why we should define the hierarchies and the relation between the attributes for all the dimension.

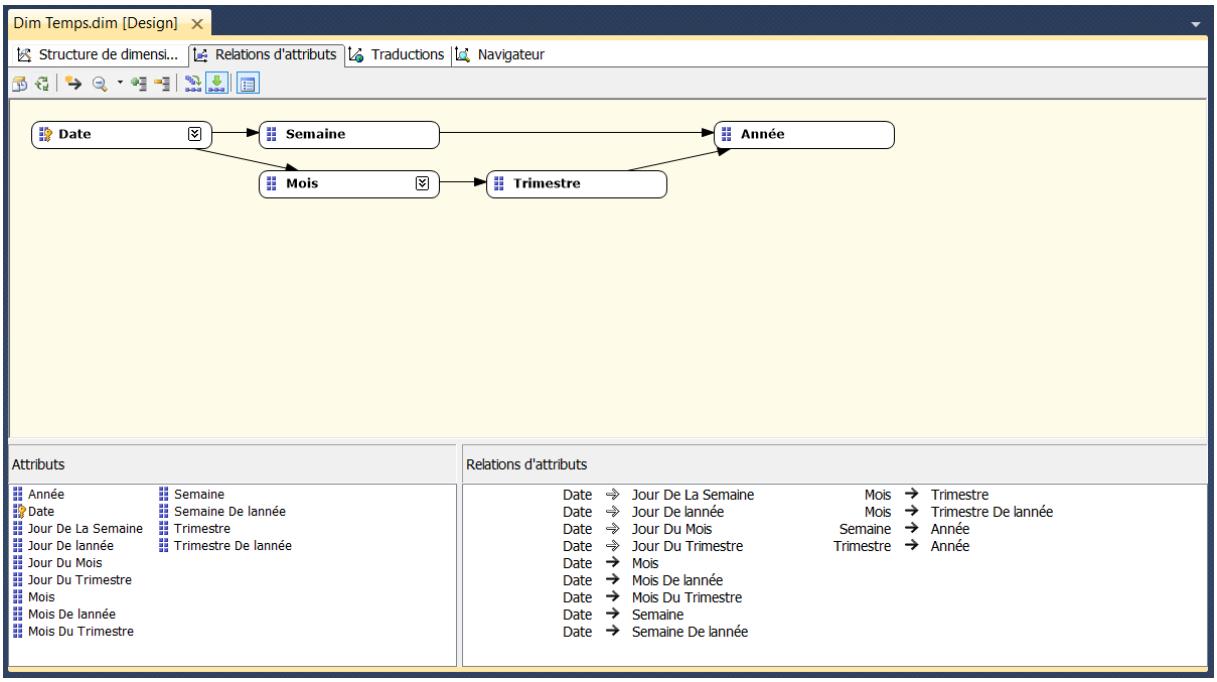


*Figure 24: time dimension's hierarchy*

## 4. Calculations

The most important task of the cube is calculation and measures. In this, we did all the needed calculation to reach our objectives for example we have entered the formula to calculate the Revenue.

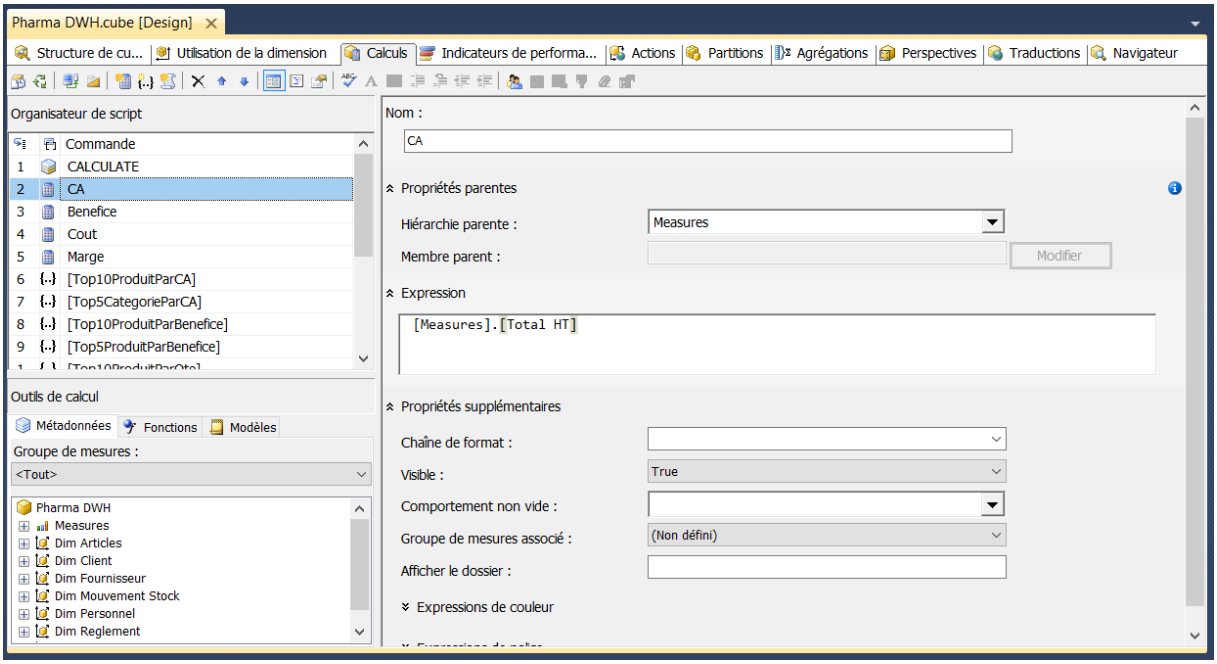


*Figure 25: cube calculations*

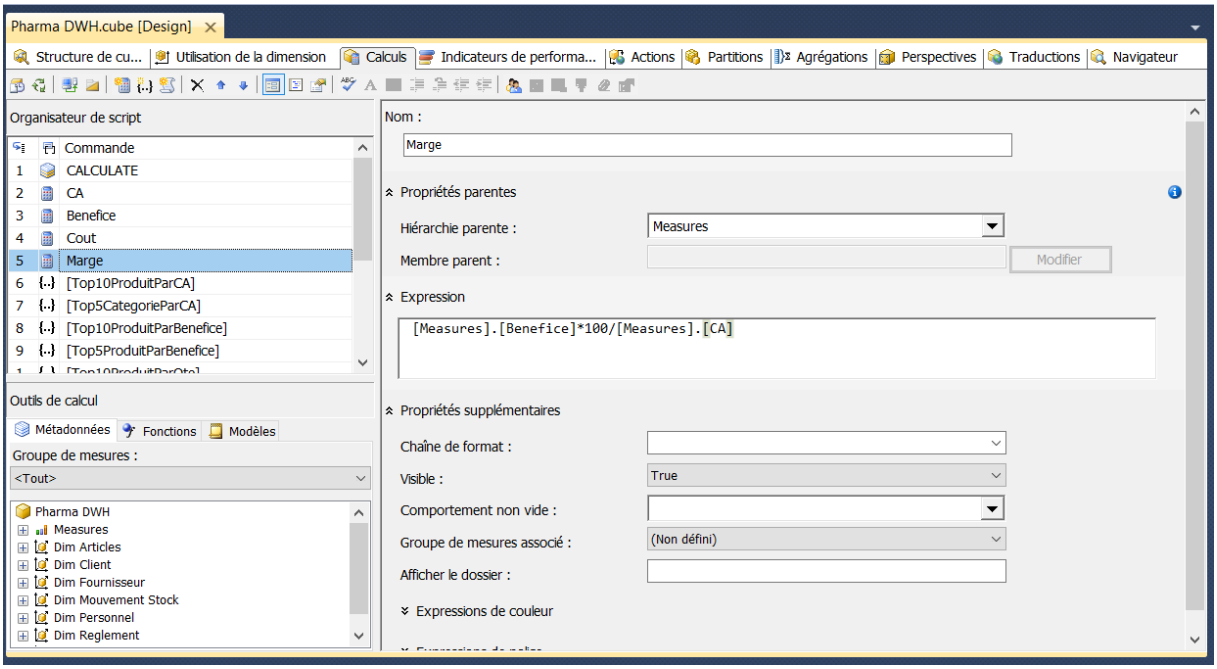


*Figure 26: cube calculations*

## II. Data visualization

Data visualization is a general term that describes any effort to help people understand the significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization softwares.

A report is a multi-perspective view into a dataset, with visualizations that represent different findings and insights from that dataset. A report can have a single visualization or pages full of visualizations. Depending on your job role, you may be someone who creates reports and/or you may be someone who consumes or uses reports.

### 1. Used tools

- **SQL Server Reporting Services**

SQL Server Reporting Services is a solution that customers deploy on their own premises for creating, publishing, and managing reports, then delivering them to the right users in different ways.


*Figure 27: logo of SSRS*

- **Power BI**

Power BI is a suite of marketing analysis tools offering insights across your organization. Produce wonderful reports, then publish them for your organization, get on the web and mobile devices.


*Figure 28: logo of powerBi*

## 2. Dashboard

A dashboard is a multi-perspective view into a dataset, with visualizations that represent different findings and insights from that dataset.

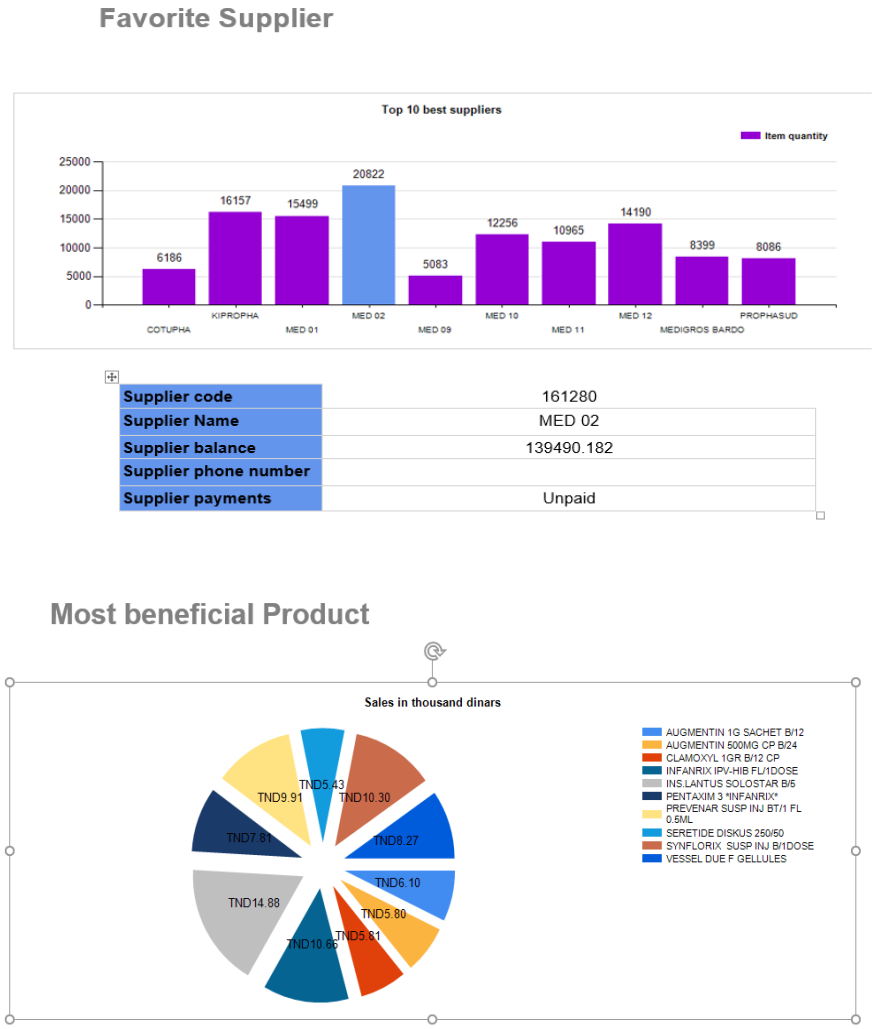


*Figure 29: part of the dashboard*

# III. Data Mining

Data mining is the process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems. Data mining is an interdisciplinary subfield of computer science and statistics with an overall goal to extract information (with intelligent methods) from a data set and transform the information into a comprehensible structure for further use.

## 1. Used Tools

- **Python and related libraries**

Python is a very used programming language in the data science community. There are several Python libraries which provide solid implementations of a range of machine learning algorithms. One of the best known is Scikit-Learn, a package that provides efficient versions

of a large number of common algorithms. **Scikit-Learn** is characterized by a clean, uniform, and streamlined API, as well as by very useful and complete online documentation.
We also used **Pandas** for data IO.
**FBProphet** was also a library we used which Implements a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with time series that have strong seasonal effects and several seasons of historical data.

- **React Native**

React Native is a JavaScript framework for writing real, natively rendering mobile applications for iOS and Android. It's based on React, Facebook's JavaScript library for building user interfaces. We used it to build our mobile application to display our data mining results .

- **Time Series**

Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values.
We used FBProphet to conduct such an analysis for our pharmacy and estimate how many sales will happen on a specific day.
This can be used to avoid giving a day off when the pharmacy is going to be busy etc.

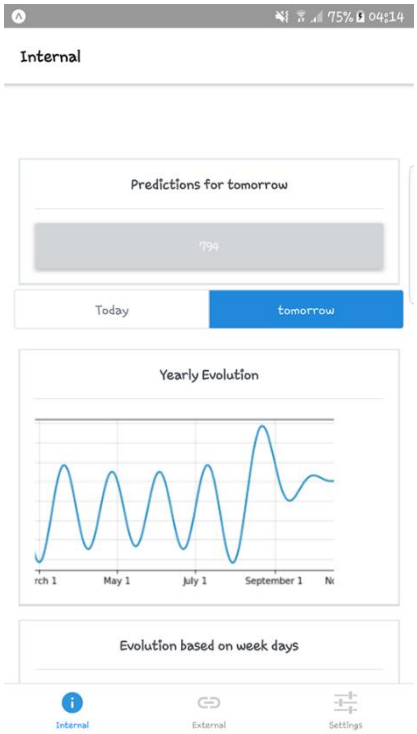


*Figure 30: the mobile application*

We show our forecasting using a mobile application as follows:

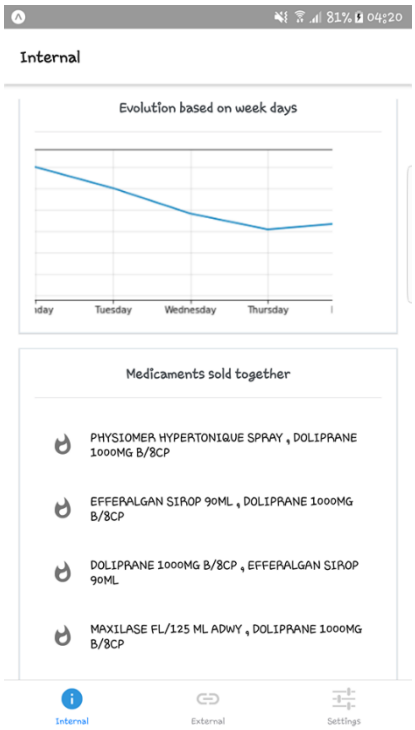


*Figure 31: the mobile application*

The pharmacist can use our app to determine the estimated number of transactions for today or tomorrow and has a graph that indicated the yearly and weekly trends of his pharmacy.

## 2. Association rule mining:

Association rule learning is a rule-based machine learning method for discovering interesting relations between variables in large databases. It is intended to identify strong rules discovered in databases using some measures of interestingness.This rule-based approach also generates new rules as it analyzes more data. The ultimate goal, assuming a large enough dataset, is to help a machine mimic the human brain's feature extraction and abstract association capabilities from new uncategorized data.

We used apriori algorithm to determine what medicaments are sold together so the pharmacy can be rearranged to have them next to each other.
***MAXILASE and PSYSIOMER HYPERTONIQUE SPRAY*** seem to really go well with ***Doliprane*** and ***Efferalgan*** who are the most sold drug (judging by the rules support). And should therefore be placed where they can be easily accessed.

# Chapter 4:

# Conclusion and perspectives

Creating Business Intelligence solutions opens up a rich and rapidly expanding overview of the current situation of the business. All that relaying on different techniques going from conventional analytics to Machine Learning, from reporting to advanced visualization.

Adaptive data virtualization solves what might best be described. More important, it gives CIOs a deep insight on how things are really going. Data virtualization becomes their ultimate risk-mitigation strategy while moving the enterprise into this new era of data processing.

Through this project, we tried to provide our client ”The pharmacy of Amel Bey” with the right information and numbers concerning mainly products, purchases, sales and suppliers. All that will help our client to take the right decisions to grow their business in an efficient way and know the real causes of the issues that they are facing today so that they can fix them.

Enterprises need to focus more on gathering data instead of technology and simply avoid getting distracted. Because it’s the ultimate way to know their business in depth and keep on growing and progress