**Entrepôt de données**

**Rapport Final**

***Sébastien Quiquerez***

***Johan Steiner***

Table des matières

[1. Introduction 4](#_Toc465153555)

[2. Analyse de l’existant 4](#_Toc465153556)

[3. Analyse des besoins 4](#_Toc465153557)

[4. Cahier des charges 4](#_Toc465153558)

[5. Modèle dimensionnel en étoile 5](#_Toc465153559)

[5.1. Explications des choix 6](#_Toc465153560)

[5.1.1. Tables de faits 6](#_Toc465153561)

[5.1.2. Dimension Commande 6](#_Toc465153562)

[5.1.3. Dimension Temps : Liaisons multiples 6](#_Toc465153563)

[5.1.4. Aucune dimension Devise 6](#_Toc465153564)

[5.1.5. Dimension Vendeur et table Vente : pas de liaison 6](#_Toc465153565)

[5.1.6. Dimension Géographie et table Vente : aucune liaison 6](#_Toc465153566)

[5.2. Hiérarchies 7](#_Toc465153567)

[5.2.1. Hiérarchie du temps 7](#_Toc465153568)

[5.2.2. Hiérarchie des vendeurs 7](#_Toc465153569)

[5.2.3. Hiérarchie des articles 7](#_Toc465153570)

[5.2.4. Hiérarchie de la géographie 8](#_Toc465153571)

[6. Données : importation/transormation 8](#_Toc465153572)

[6.1. Staging 8](#_Toc465153573)

[6.2. DataWarehouse 8](#_Toc465153574)

[6.2.1. Master data 8](#_Toc465153575)

[6.2.2. Qualité des données 8](#_Toc465153576)

[6.2.3. Audit 8](#_Toc465153577)

[6.3. Cube 8](#_Toc465153578)

[7. Indicateurs 8](#_Toc465153579)

[8. Conclusion 9](#_Toc465153580)

[9. Bibliographie 9](#_Toc465153581)

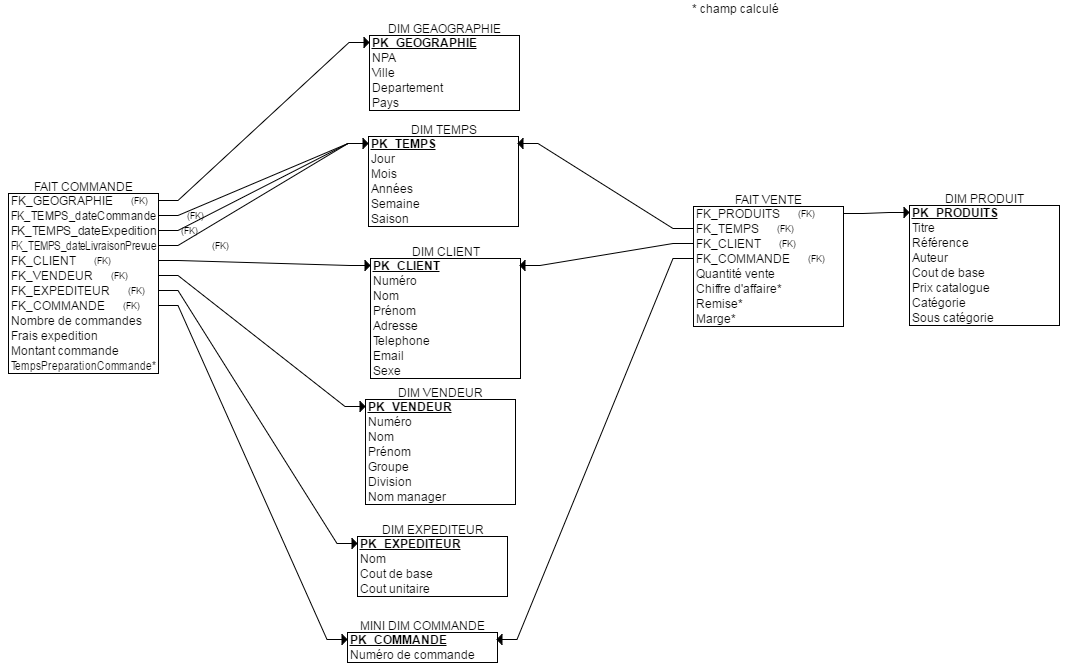
# Introduction

# Analyse de l’existant

# Analyse des besoins

# Cahier des charges

# Modèle dimensionnel en étoile



## Explications des choix

### Tables de faits

Notre schéma contient 2 tables de faits : la table Commande et la table Vente.

La table Vente contient les informations propres aux produits et plutôt à une ligne de la commande (quelle quantité pour un produit ? Quel pourcentage de remise sur chaque produit ? Quelle marge sur la vente ? Etc…) ainsi que le chiffre d’affaire.

La table Commande est plus orientée sur la commande en générale, par exemple pour connaître les dates de commande, livraison et expédition, l’expéditeur et ses frais ou encore des informations comme le nombre de commandes passées. On peut également

### Dimension Commande

Avant la création de cette dimension, il était impossible de lier une commande aux ventes. On ne pouvait par exemple pas connaître le prix total de la commande. La dimension Commande permet de créer cette liaison via le numéro de commande.

### Dimension Temps : Liaisons multiples

Nous avons créé 3 liaisons entre la dimension Temps et la table de faits Commande, car dans le cas, il est mentionné 3 types de dates distincts : la date de commande, la date de livraison prévue et la date d’expédition.

### Aucune dimension Devise

Nous avons omis la dimension Devise car nous prévoyons de convertir tous les prix et montant en une seule devise avant de les intégrer à notre entrepôt de données. Nous n’aurons donc qu’une seule devise présente dans notre Data Warehouse.

### Dimension Vendeur et table Vente : pas de liaison

Nous avons jugé inutile de relier un vendeur à la table Vente car nous sommes d’avis que le vendeur est plutôt lié à une commande et qu’il n’est pas très utile de déterminer les ventes d’un vendeur par produit par exemple.

### Dimension Géographie et table Vente : aucune liaison

Nous avons fait le choix de ne pas analyser les ventes de chaque produit par région. Nous n’allons pas chercher à savoir si un produit précis s’est mieux ou moins bien vendu dans une certaine région plutôt que dans une autre.

## Hiérarchies

### Hiérarchie du temps

**Année**

**Année**

**Semaine**

**Mois**

**Saison**

**Jour**

**Jour**

**Jour**

### Hiérarchie des vendeurs

**Groupe**

**Division**

**Manager**

**Vendeur**

### Hiérarchie des articles

**Catégorie**

**Sous-catégorie**

**Article**

### Hiérarchie de la géographie

**Pays**

**Département**

**Ville**

# Données : importation/transormation

L’importation des données dans notre Data Warehouse a été réalisée à l’aide de la méthode ETL (Extract, Transform, Load).

## Staging

La zone staging représente la zone contenant les données qui n’ont pas encore été transformées. Nous extrayons les données depuis nos différentes sources (fichiers Excel, fichiers Access, fichiers texte, etc…), les plaçons dans notre staging area, à la suite de quoi nous les transformons pour qu’elles correspondent à ce qu’attendent nos tables présentes dans notre Data Warehouse et les chargeons dans celui-ci.

### Types de données

L’importation des données dans la staging a été simplifiée au niveau des types de données. En effet, nous n’avons à ce moment-là pas encore procédé à une transformation des différents types, mais avons simplement importé toutes les données sous forme de chaînes de caractères (sauf les identifiants).

## Data Warehouse

### Master data

### Qualité des données

### Audit

## Cube

# Indicateurs

# Conclusion

# Bibliographie