



REPUBLIQUE ALGERIENNE  
DEMOCRATIQUE ET POPULAIRE



Ministère de l'Enseignement Supérieur et de la  
Recherche Scientifique

Ecole Nationale Polytechnique d'Alger  
Département de Génie Industriel

Mémoire de Projet de Fin d'Etudes  
Pour l'obtention du diplôme d'Ingénieur d'Etat en Génie Industriel

**Thème**

**Contribution à la mise en place d'un système de Business  
Intelligence basé sur la démarche de Kimball  
Application : Unilever Algérie**

Présenté par : Mlle Amira Bouchra HADDOUCHE (Management Industriel)  
Mlle Sonia MAKHLOUF (Management de l'Innovation)

Sous la direction de : Mme Bahia BOUCHAFAA (ENP)  
Mr Islam ZERROUT (Unilever Algérie)

Présenté et soutenu publiquement le : 06/09/2020

### Composition du Jury

Président : Mr. R.GOURINE Maître de conférences B (ENP)  
Promoteur : Mme B.BOUCHAFAA Maître de conférences A (ENP)  
Examineur : Mr. M.AIB Maître de conférences B (ENP)

**ENP 2020**





REPUBLIQUE ALGERIENNE  
DEMOCRATIQUE ET POPULAIRE



Ministère de l'Enseignement Supérieur et de la  
Recherche Scientifique

Ecole Nationale Polytechnique d'Alger  
Département de Génie Industriel

Mémoire de Projet de Fin d'Etudes  
Pour l'obtention du diplôme d'Ingénieur d'Etat en Génie Industriel

**Thème**

**Contribution à la mise en place d'un système de Business  
Intelligence basé sur la démarche de Kimball  
Application : Unilever Algérie**

Présenté par : Mlle Amira Bouchra HADDOUCHE (Management Industriel)  
Mlle Sonia MAKHLOUF (Management de l'Innovation)

Sous la direction de : Mme Bahia BOUCHAFAA (ENP)  
Mr Islam ZERROUT (Unilever Algérie)

Présenté et soutenu publiquement le : 06/09/2020

#### **Composition du Jury**

Président : Mr. R.GOURINE Maître de conférences B (ENP)  
Promoteur : Mme B.BOUCHAFAA Maître de conférences A (ENP)  
Examineur : Mr. M.AIB Maître de conférences B (ENP)

**ENP 2020**

## *Remerciements*

Nous remercions en premier lieu nos parents, sans qui nous n'aurions jamais pu atteindre ce point dans nos vies.

Nous adressons nos sincères remerciements à notre encadrante, Mme Bouchafaa, pour son suivi ainsi que les conseils qu'elle nous a prodigués.

Nous tenons également à remercier notre promoteur au sein d'Unilever Algeria, Mr. Islam Zerrouk, ainsi que les membres de son équipe, Mrs : Amine Hanafia, Idir Ait Menguellet et Chaouche Ramdane Kheir-eddine, pour nous avoir accueillies, soutenues, aidées et suivies tout au long de l'exécution de ce travail.

Sans oublier de remercier Mme Affaf Chergui ainsi que Mr Iskandar Zouaghi sans qui nous n'aurions pas pu avoir accès à cette opportunité.

Nous exprimons également notre gratitude à Yasmine Haddouche, pour ses recommandations techniques et les conseils qu'elle nous a offerts.

Nos remerciements s'adressent aussi aux enseignants du département Génie Industriel, pour tous les efforts qu'ils ont consentis pour notre formation.

Pour finir, nos remerciements s'adressent aussi aux membres du jury qui nous font l'honneur d'évaluer notre travail.

*Bouchra, Sonia*

## *Dédicaces*

*Je dédie ce travail à mes parents, qui se sont toujours sacrifiés pour que je puisse étudier dans les meilleures conditions possibles.*

*A ma soeur, qui n'hésite jamais à s'oublier pour m'aider.*

*A mes amies : Belkisse, Ablem, Hasna, Selma, Melissa, et d'autres encore, dont l'amitié m'a toujours été précieuse.*

*Au Génie industriel et à mes amies prépa, qui font de mes années de polytech une part inoubliable de mon existence.*

*A Sonia, ma binôme, sans qui ce travail n'aurait jamais pu voir le jour, elle qui, malgré toutes les contraintes auxquelles elle dut faire face, n'a jamais abandonné et s'est pleinement investie.*

*Bouchra*

## *Dédicaces*

*A ma mère qui m'a donné la vie, qui m'a appris les premiers pas et les premiers mots. A toi mon étoile qui illumine mon chemin et qui me donne sans retour. A toi mon idole qui m'a accompagné dans mon bonheur et mon malheur et qui a fait de moi la femme d'aujourd'hui.*

*A mon père qui a tant sacrifié pour mon bonheur, qui m'a forgé et m'a donné la force pour affronter la vie. A toi mon ange gardien qui a toujours été derrière moi, mon guide, mon soutien et ma raison dans la vie.*

*A ma sœur qui est une deuxième mère pour moi. A celle qui me donne la confiance et en qui j'ai le plus confiance dans ma vie et sur laquelle je peux compter dans les moments difficiles.*

*A toi Zineb, ma sœur du cœur, ma jumelle et ma moitié. A toi ma meilleure amie pour le pire et pour le meilleure.*

*A tous mes cousins et cousine, et à toutes ma famille qui m'a toujours encouragé pour aller de l'avant.*

*A tous mes amis qui m'ont soutenu, spécialement Rokaya, Sid-Ahmed, Wissem, Afaf, Sarah, Amira, Fatima, Lysa, Asma, Fella et Sarha.*

*A toute la famille indus et innov avec qui j'ai passé les meilleurs moments durant ces trois dernières années.*

*A tous mes profs qui ont contribué à ma réussite depuis le début de mon parcours académique.*

*Et spécialement à ma binôme Bouchra qui m'a accompagné durant tout le PFE, avec qui j'ai partagé tous les moments de ce projet et avec qui j'ai appris pas mal de choses.*

*Merci à vous*

*Sonia*

## ملخص:

الهدف من هذا العمل هو وضع تصميم وتفعيل نظام معلوماتي يسمح باتخاذ القرارات، الامر الذي من شأنه ان يساعد موظفي قسم خدمة العملاء والزبائن في شركة اونيليفر الجزائر في اعداد التقارير كما يسمح لهم بالقيام بتحليل البيانات بشكل أكثر ثراء.

وسعياً لذلك، اعتمدنا على مقارنة تتمثل في تفعيل مختلف مراحل مشروع التقنيات المعلوماتية، مما نتج عنه تنميط عدة مراكز صغرى لتخزين البيانات وتفعيل جزء منها، وهذا من شأنه تسهيل عمليات اعداد التقارير والتحليل.

**الكلمات المفتاحية :** ذكاء الاعمال، التقارير، مراكز صغرى لتخزين البيانات، تصميم نظام معلوماتي، OLAP

ETL

## Abstract :

This work aims to design and implement a Business Intelligence System that helps the Unilever Algeria Customer Service staff members to perform their reporting tasks and enables them to make richer data analyses.

To that end, we followed the various steps of a BI project implementation, resulting in the modeling of several datamarts and a partial implementation, which eased the reporting and analysis activities.

**Keywords :** Business intelligence, reporting, datamart, dimensional modeling, OLAP, ETL

## Résumé :

L'objectif de ce travail est de modéliser et d'implémenter un système décisionnel permettant de faciliter l'activité de Reporting et d'étendre les possibilités d'analyse des données du personnel du service Customer Service de l'entreprise Unilever Algeria.

Pour ce faire, nous avons adopté la démarche de réalisation d'un projet BI à travers ses différentes phases afin d'aboutir à la modélisation de différents datamarts ainsi que l'implémentation d'une partie d'entre eux, facilitant ainsi au service la réalisation de rapports et d'analyses.

**Mots clés :** Business intelligence, reporting, datamart, modélisation dimensionnelle, OLAP, ETL

# Table des matières

Liste des Tableaux

Liste des figures

Liste des abréviations

Introduction générale.....	15
Chapitre I : Etat de l'art.....	16
1. Introduction .....	17
2. Business Intelligence : (Ballard, 2006) .....	17
2.1. Définition : (Ballard, 2006) (Sharma et al, 2015).....	17
2.2. Phases d'un projet BI : (Captio,2015) (N'goran, 2016).....	17
3. Conclusion.....	39
Chapitre II : Etat des lieux.....	40
1. Introduction .....	41
2. Présentation du secteur des biens de grande consommation (FMCG).....	41
2.1. Dans le monde.....	41
2.2. En Algérie .....	42
3. Présentation d'Unilever.....	43
3.1. Unilever Monde .....	43
3.2. Présentation d'Unilever Algerie : (Bendiab et Sattouf, 2018) (Zennir,2019).....	44
4. Etude de l'existant .....	49
4.1. Processus.....	49
4.2. Rapports .....	54
4.3. Audit de la qualité des données .....	56
4.4. KPIs.....	58
4.5. Enoncé de la problématique .....	59
4.6. Méthodologie .....	59
5. Conclusion.....	60
Chapitre III : Conception de la solution BI.....	61
1. Introduction .....	62
2. Choix de la méthodologie adoptée : (YESSAD, et al. , 2016).....	62
3. Identification des besoins .....	63
3.1. Classification des besoins en thèmes .....	64
4. Modélisation dimensionnelle .....	66

4.1. Etape 1 : Choix des sous-processus .....	66
4.2. Etape 2 : Choix du grain .....	70
4.3. Etape 3 : Choix des dimensions .....	70
4.4. Etape 4 : Choix des mesures .....	71
5. Choix de l'architecture adoptée : (KimballGroup, SD) .....	73
5.1. La matrice de bus dimensionnel .....	74
6. Choix du schéma pour la modélisation .....	75
6.1. Modèles schémas multidimensionnels des datamarts .....	75
6.2. Modèle schéma multidimensionnel du processus Order To Cash .....	79
7. Phase de stockage .....	81
8. Conclusion .....	81
Chapitre IV : Réalisation de la solution BI .....	82
1. Introduction .....	83
2. Choix des datamarts .....	83
3. Implémentation de la solution .....	86
3.1. Phase d'alimentation .....	86
3.2. Phase d'exploitation .....	96
4. Exemples de rapports types .....	101
5. Pistes d'amélioration .....	105
6. Conclusion .....	105
Conclusion générale .....	106
Bibliographie .....	108
Annexes .....	114

# Liste des Tableaux

Tableau 1 : Comparaison Full Load vs Delta Load .....	19
Tableau 2 : Comparaison entre les différents types de tables de faits.....	22
Tableau 3 : Comparaison Inmon vs Kimball .....	25
Tableau 4 : Comparaison schéma en flocon vs schéma en étoile .....	29
Tableau 5 : Produits d'Unilever commercialisés en Algérie.....	45
Tableau 6 : Rapports concernant les ventes .....	54
Tableau 7 : Rapports résumant l'activité .....	54
Tableau 8 : Rapport concernant les réclamations .....	55
Tableau 9 : Rapport concernant la situation financière des clients .....	55
Tableau 10 : Rapports concernant les stocks .....	56
Tableau 11 : Rapports concernant les stocks .....	57
Tableau 12 : Liste et caractéristiques des KPIs liés à l'activité du service .....	58
Tableau 13 : comparaison entre les trois méthodologies du projet BI.....	62
Tableau 14: Description des fichiers récoltés en utilisant la méthode orientée Source .....	63
Tableau 15: Tableau illustrant la classification des besoins en thèmes .....	64
Tableau 16 : Faisabilité et Importance des sous-processus .....	68
Tableau 17: Affectation des thèmes de suivi aux sous-processus.....	69
Tableau 18: Détermination des grains des sous-processus .....	70
Tableau 19: Choix des dimensions des sous-processus .....	71
Tableau 20: Mesures de base des sous-processus .....	72
Tableau 21: Mesures dérivées des sous-processus .....	73
Tableau 22: Matrice de bus multidimensionnel .....	75
Tableau 23 : Récapitulatif Datamart Gestion des Commandes .....	84
Tableau 24 : Récapitulatif Datamart Réclamations en caisses.....	85
Tableau 25 : Récapitulatif Datamart Réclamations en pièces.....	85
Tableau 26 : Récapitulatif Datamart Gestion des Remboursements TTS .....	85
Tableau 27 : comparaison entre les leaders du marché des logiciels de l'ETL .....	86
Tableau 28 : Formules de création des colonnes dérivées .....	97
Tableau 29 : Formules de calcul des mesures dérivées .....	98
Tableau 30 : Suivi des réclamations.....	99
Tableau 31 : Suivi de la GSV .....	100
Tableau 32 : Suivi des ventes primaires.....	100
Tableau 33 : Suivi des remboursements TTS .....	101
Tableau 34 : Questionnaire sur les activités et les besoins du service .....	115
Tableau 35 : Dimension Produit.....	117
Tableau 36 : Dimension Client (KD) .....	117
Tableau 37 : Dimension TTS .....	117
Tableau 38 : Dimension Date.....	118
Tableau 39 : Dimension Entrepôt .....	118
Tableau 40 : Dimension CCFOT Reasons .....	118
Tableau 41 : Dimension Statut_Finance .....	118

Tableau 42 : Dimension Statut_Customer_Service .....	118
Tableau 43 : Dimension Statut_Reclamation .....	119
Tableau 44 : Dimension Type_Reclamation .....	119
Tableau 45 : Dimension Client_Secondaire .....	119

# Liste des figures

Figure 1 : Schéma expliquant le déroulement du processus ETL .....	20
Figure 2 : Modèle dimensionnel de type Datavault .....	25
Figure 3 : Magasins de données indépendants .....	26
Figure 4 : Bus de magasins de données.....	26
Figure 5 : Architecture Hub-and-spoke.....	27
Figure 6 : Entrepôt de données centralisé .....	27
Figure 7 : Architecture fédérée .....	28
Figure 8 : Exemple de schéma en étoile.....	28
Figure 9 : Exemple de schéma en flocon .....	29
Figure 10 : Exemple de schéma galaxy.....	30
Figure 11 : Hiérarchisation des processus .....	32
Figure 12 : Exemple de matrice de bus dimensionnel .....	34
Figure 13 : Opérations réalisées par OLAP.....	38
Figure 14 : Segmentation du marché des bien de grande consommation 2017 .....	41
Figure 15 : Top-6 des entreprises marché FMCG selon leur CA de 2019 .....	42
Figure 16: Top-5 des entreprises du marché algérien Beauty and Personal Care.....	42
Figure 17 : Organigramme d’Unilever Algeria .....	45
Figure 18 : Organigramme de la division Supply Chain.....	46
Figure 19 : Organisation de la distribution d’Unilever Algérie .....	48
Figure 20 Modélisation du processus Order-to-cash.....	51
Figure 21 Modélisation du processus Order-to-cash.....	52
Figure 22 : Modélisation du processus Order-to-cash .....	53
Figure 23: Hiérarchisation des sous-processus .....	69
Figure 24: Architecture en bus de magasin de données .....	74
Figure 25 : Datamart Gestion des stocks.....	76
Figure 26 : Datamart Gestion des réclamations .....	76
Figure 27: Datamart gestion des Provisions_TTS .....	77
Figure 28: Datamart gestion des Remboursements_TTS.....	78
Figure 29: Datamart Gestion des commandes.....	78
Figure 30: Datamart Forecast .....	79
Figure 31: Datamart Ventés secondaires.....	79
Figure 32: Schéma regroupant les datamarts .....	80
Figure 33 : Schéma ETL de la dimension Produit .....	88
Figure 34 : Schéma de l’ETL de la dimension KD .....	89
Figure 35 : ETL de la dimension TTS.....	90
Figure 36 : Schéma de l’ETL des remboursements TTS .....	91
Figure 37 : Schéma de l’ETL de la table de fait Commandes - 1 .....	93
Figure 38 : Schéma de l’ETL de la table de fait Commandes - 2 .....	93
Figure 39 : Schéma de l’ETL de la table de fait Réclamations - 1 .....	95
Figure 40 : Schéma de l’ETL de la table de fait Réclamations - 2 .....	95
Figure 41 : Schéma de l’ETL de la table de fait Réclamations - 3 .....	95

Figure 42 : Rapport Réclamations.....	102
Figure 43 : Rapport Commande - GSV.....	103
Figure 44 : Rapport Commandes - Les Ventes .....	104
Figure 45 : Norme BPMN.....	114
Figure 46 : Today's orders.....	120
Figure 47 : Claim Tracker .....	120
Figure 48 : KD : KD file .....	121
Figure 49 : Price .....	121
Figure 50 : Material.....	121
Figure 51 : CN JUL 19.....	121
Figure 52 : Création table KD .....	122
Figure 53 : Création Table Produit.....	122
Figure 54 : Création Table Réclamations_caisses .....	122
Figure 55 : Création Table Réclamations_pieces.....	123
Figure 56 : Création Table Commande .....	123
Figure 57 : Création Table Statut_Financier – 1 .....	123
Figure 58 : Création Table Statut_Financier - 2 .....	124
Figure 59 : Création Table Statut_Customer_Service - 1 .....	124
Figure 60 : Création Table Statut_Customer_Service - 2 .....	124
Figure 61 : Création Table Entrepôt - 1 .....	124
Figure 62 : Création Table Entrepôt - 2 .....	124
Figure 63 : Creation Table Date – 1 .....	125
Figure 64 : Creation Table Date – 2 .....	125
Figure 65 : Création table TTS.....	126
Figure 66 : Creation Table Remboursements TTS.....	126
Figure 67 : Creation Table Type Réclamation - 1 .....	126
Figure 68 : Contenu de la table Type Réclamation - 2.....	126
Figure 69 : Creation Table Statut Réclamation – 1 .....	126
Figure 70 : Contenu de la table Statut Réclamation - 2.....	127
Figure 71 : Creation Table CCFOT Reasons - 1 .....	127
Figure 72 : Contenu de la table CCFOT Reasons - 2 .....	127
Figure 73 : Architecture du logiciel SSIS .....	128
Figure 74 : Vue sur la fenêtre Dataflow .....	129
Figure 75 : Choix des colonnes à extraire de la source .....	130
Figure 76 : Colonne dérivée : Suppression des espaces de la chaîne.....	130
Figure 77 : Colonne dérivée : Remplacement des '/' par des '-'.....	131
Figure 78 : Conversion des données : du float à l'entier.....	131
Figure 79 : Fractionnement conditionnel en quatre flux selon les factures .....	132
Figure 80 : Recherche : Correspondance des codes SKUs de la source et de la table Produit .....	132
Figure 81 : Fractionnement conditionnel en 5 flux selon les régions .....	133
Figure 82 : Jointure de fusion.....	133
Figure 83 : Choix des colonnes de la table à extraire.....	134
Figure 84 : Fractionnement conditionnel : Nouvelles et anciennes réclamations.....	134

Figure 85 : Choix des colonnes à insérer dans la table .....	135
Figure 86 : Requête de la mise à jour la table Réclamations_caisse .....	135
Figure 87 : Tri des données extraites de la table Réclamations_caisses .....	136
Figure 88 : Modèle du datamart Réclamations sur Power Pivot .....	137
Figure 89 : Modèle du datamart Commandes sur Power Pivot .....	138
Figure 90 : Modèle du datamart Remboursement TTS sur Power Pivot .....	138

# Liste des abréviations

**BDD** : Base De Données

**BI** : Business Intelligence

**BPMN** : Business Process Modeling Norme

**CA** : Chiffre d'Affaires

**CCFOT** : Customer Cases Fill on Time

**CD** : Costumer Development

**DD** : Dimension Dégénérée

**DR** : Dispatch Rate

**DW** : Data Warehouse

**DZD** : Dinar Algérien

**ETL** : Extract, Transform, Load

**FK** : Foreign Key

**FMCG** : Fast Moving Consumer Goods

**GSV** : Gross Sales Value

**GTM** : Go To Market

**HT** : Hors Taxe

**IT** : Technologie de l'information

**KD** : Key Distributor

**KPI** : Key Performance Indicators

**MENA** : Middle East and North Africa

**OLAP** : On Line Analytical Processing

**OMR** : Order Management Report

**PO** : Purchase Order

**R&D** : Recherche et Développement

**RH** : Ressources Humaines

**RSP** : Retail Sale Price

**SAP** : Systems, Applications and Products for data processing

**SAS** : Statistical Analysis System

**SHEOH** : Security, Hygiene, Environement, Occupational, Health

**SK** : Surrogate Key

**SKU** : Stock Keeping Unit

**SLA** : Service Level Agreement

**SLOBS** : SLow moving / OBsoletes

**SO** : Sales Order

**SQL** : Structured Query Lannguage

**TDB** : Tableaux De Bord

**TRUB** : Turquie- Russie - Ukraine - Biélorussie

**TTC** : Toute Taxe Comprise

**TTS** : Trade Terms Specialist

**USD** : United States Dollar

## Introduction générale

Ayant duré 5 mois, notre stage de fin d'études s'est déroulé dans l'entreprise Unilever Algérie, dans le service Customer Service, il eut pour finalité la mise en place d'un système de business intelligence, en s'appuyant sur l'application de la méthodologie de Kimball, afin de faciliter l'activité de reporting et l'exploitation des nombreuses données générées par le service.

Les entreprises de production de biens de grande consommation (FMCG) génèrent d'énormes volumes de données tous les jours, néanmoins, elles ne sont pas toutes en mesure de toutes les utiliser pour en extraire des informations utiles. L'inclusion de la Business Intelligence dans la stratégie de l'entreprise permet justement de trier et sortir la bonne information de la bonne donnée. Cela conduit ces entreprises à avoir de l'avantage concurrentiel, en leur offrant une aide précieuse pour l'analyse de leur activité, la compréhension de leur marché et de leurs concurrents ainsi que l'identification d'opportunités.

C'est dans cette optique que nous avons proposé au service Customer Service, très actif en matière de reporting et disposant de grands volumes de données, l'implémentation d'un système de business intelligence afin de faciliter l'activité de reporting et de permettre de réaliser des analyses plus détaillées.

Ce qui nous a conduit à formuler notre problématique de la manière suivante :

**« Comment réduire le temps de création de rapports, tout en assurant un meilleur visuel et offrant plus de possibilités d'analyse aux utilisateurs finaux ? »**

Pour répondre à la problématique, nous proposons une solution que nous développons en de deux parties : la première partie a pour résultat la conception de datamarts et consiste en l'identification et l'analyse des besoins, les étapes de la modélisation dimensionnelle, ainsi que la création d'un espace de stockage. La seconde explique en détail l'implémentation réelle des datamarts à travers l'explication des étapes de l'ETL à suivre pour leur chargement, ainsi que les analyses que cette implémentation rend possibles.

A cette fin, nous avons divisé notre travail en quatre chapitres :

Le premier chapitre présente les fondements théoriques sur lesquels nous nous sommes basées, structurés selon les quatre phases du projet BI : Phase d'alimentation, Phase de modélisation et de stockage, Phase de restitution et Phase d'analyse.

Le second présente une étude de l'existant de l'entreprise et du service, avec premièrement une présentation de l'entreprise dans le monde et en Algérie, et secondement une étude de l'existant du service à travers l'étude de son processus, de son activité de reporting, de sa qualité des données et de ses KPIs.

Le troisième présente en détail les différentes étapes suivies pour la conception de la solution, depuis l'identification et l'analyse des besoins, le déroulement des quatre étapes de la modélisation dimensionnelle de Kimball qui nous ont permis d'obtenir les schémas des datamarts, jusqu'à leur stockage dans des tables SQL.

Le quatrième présente les deux phases d'alimentation et d'analyse, à travers le déroulement des étapes du processus ETL et la mise en relief des différentes analyses possibles après chargement des datamarts.

Pour finir, nous proposons quelques pistes d'améliorations pour notre travail.

# **Chapitre I : Etat de l'art**

## 1. Introduction

Ce chapitre sera dédié à la présentation des aspects théoriques sur lesquels repose notre travail. Il détaillera les quatre phases d'un projet BI.

En premier lieu nous détaillerons la phase d'alimentation, en expliquant le processus ETL.

Puis nous aborderons la phase de modélisation, pour laquelle nous introduirons les concepts de base du data warehousing et les différentes approches de modélisation.

Ensuite, nous présenterons la phase de restitution des données à travers l'introduction des TDBs et du reporting.

Pour finir, nous exposerons dans la phase d'analyse deux outils pour l'analyse des données : OLAP et le data mining.

## 2. Business Intelligence : (Ballard, 2006)

Dans un environnement où la concurrence est de plus en plus accrue, la survie des entreprises dépend de leur capacité à agir rapidement face aux dynamiques de changement et défis du marché. Elles doivent anticiper les tendances, identifier les opportunités, changer leurs stratégies et réorienter leurs ressources afin de garder une longueur d'avance sur la concurrence. La clé pour réussir dans un tel environnement est l'information.

Les entreprises collectent des volumes importants de données internes et externes. Elles ont donc besoin de transformer ces données brutes en informations exploitables, ce processus passe par plusieurs étapes : extraction, consolidation, organisation des données, puis stockage, distribution, analyse et assurance d'un accès facile aux informations obtenues. L'obtention de ces informations représente un avantage mais aussi un défi pour l'entreprise. Le but de la BI est de créer de la connaissance à partir des informations afin d'améliorer le processus de la prise de décision et de transformer ces décisions en actions.

### 2.1.Définition : (Ballard, 2006) (Sharma et al, 2015)

La BI est une solution informatique permettant d'analyser les données historisées internes ou externes à l'entreprise, telles que les données démographiques ou celles relatives aux conditions climatiques, afin d'étudier une fonction particulière de l'entreprise.

Les données sont utilisées pour comprendre les tendances, les forces, les faiblesses de l'entreprise ainsi que pour l'analyse de la situation du marché et des concurrents.

Avant l'existence de la BI, les entreprises utilisaient des méthodes conventionnelles standard pour transformer les données en informations exploitables, chose qui nécessite beaucoup de ressources et de temps, et peut occasionner des erreurs dues au facteur humain.

### 2.2.Phases d'un projet BI : (Captio,2015)\_(N'goran, 2016)

Le projet BI a la caractéristique d'être unique, il répond aux besoins spécifiques d'un processus donné. Néanmoins, il est possible de définir les caractéristiques et les phases communes entre les différents projets BI.

Le but principal d'un tel projet est d'assurer une information appropriée et précise pour la prise de décision. La meilleure façon de faire est de structurer le projet en un ensemble de phases liées entre elles et ayant leurs propres objectifs.

L'on peut donc structurer un projet BI en 04 phases :

- Phase d'alimentation.
- Phase de modélisation et de stockage.
- Phase de restitution ou de distribution.
- Phase d'analyse.

### **2.2.1. Phase d'alimentation : (IBM,2020) (Nida, 2020)**

Le but de cette phase est de collecter les données nécessaires à partir des différents systèmes sources. Il s'agit principalement du processus ETL qui se charge de l'extraction, de la transformation et du chargement des données.

#### **2.2.1.1.Extract : (GRIM, 2014)**

Regrouper les sources desquelles l'on souhaite extraire les données, puis les organiser en termes de date, de taille et de source pour faciliter la prochaine étape.

Après cela, il est nécessaire de :

- Déterminer les types de données présentes.
- Identifier les clés primaires et étrangères.
- Détecter les erreurs de sources.
- Filtrer les données en fonction du besoin.

Puis les données sont transférées dans la staging area pour y subir des transformations sans impacter significativement les systèmes sources.

Ce transfert peut s'effectuer de trois différentes manières :

- **Push** : Le système de production pousse les données vers la Staging Area lorsqu'il n'est pas occupé, cela peut être problématique si le système de production est tout le temps occupé.
- **Pull** : Les données sont tirées du système de production vers la Staging Area, cette méthode risque de surcharger le système de production dans le cas où il serait déjà utilisé
- **Push-Pull** : Mix des deux méthodes, la source prépare les données et en informe la Staging Area qui ne les récupère que lorsque la source est libre.

### 2.2.1.2.Transform : (IBM,2020) (Nida, 2020)

Durant cette étape les données se trouvent dans la Staging Area, elles subissent plusieurs transformations pour qu'elles puissent s'adapter aux schémas dimensionnels et aux besoins des utilisateurs. L'étendue des transformations dépend de ces besoins et de la nature des données source.

Les transformations peuvent se présenter sous forme de :

- Nettoyage et validation des données, ainsi que la suppression des doublons.
- Modifications sur les colonnes et les lignes : Type de données, appellations, concaténation, etc.
- Génération de clés de substitution et dénormalisation des données.
- Filtrage et tri des données.
- Jonction des systèmes sources avec les clés.

### 2.2.1.3.Load : (IBM,2020) (Nida, 2020)

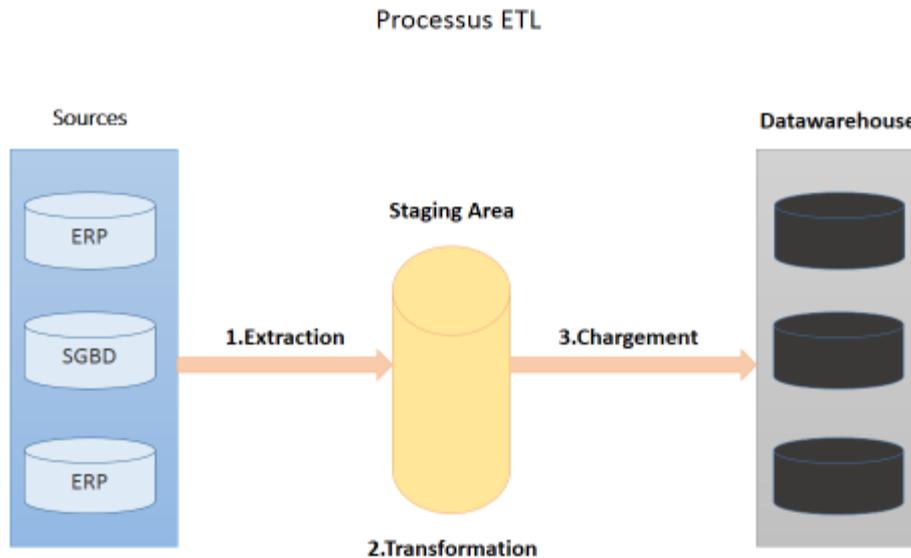
Dans cette étape les données précédemment transformées sont chargées dans le data warehouse ou les datamarts. En général, on commence par un premier chargement de toutes les données, puis on programme des rechargements périodiques pour ajouter les nouvelles données.

Il existe deux méthodes pour charger les données dans l'ETL, le full Load et le delta Load, le tableau ci-dessous montre les particularités de chaque méthode.

**Tableau 1 : Comparaison Full Load vs Delta Load (Zouaghi, 2019)**

	<b>Full Load</b>	<b>Delta Load</b>
<b>Démarche</b>	Chargement de toutes les données et effacement des anciennes données	Chargement des nouvelles données
<b>Durée</b>	Longue	Courte
<b>Risque de perte des données</b>	Présent	Absent
<b>Intégrité des données</b>	Garantie	Incertaine

La figure ci-dessous résume les étapes du processus ETL:



**Figure 1 : Schéma expliquant le déroulement du processus ETL (Adeoinformatique, 2019) (Zouaghi, 2019)**

### 2.2.2. Phase de modélisation et de stockage

Dans cette phase, les données sont stockées dans les data warehouses ou les datamarts qui se chargent de la centralisation des données.

#### 2.2.2.1. Définition d'un datawarehouse : (Godin, et al, 2011) (Bellaachia, SD)

Le data warehouse est une base de données conçue à partir des multiples sources de l'entreprise (internes et externes), principalement les systèmes transactionnels, et permettant la réalisation d'analyses optimisées.

Il représente le système décisionnel de l'entreprise destiné à l'analyse des données consolidées et historisées.

Dans son livre, Inmon décrit le data warehouse comme étant une collection de données **thématiques, intégrées, non volatiles** et **historiques**, organisées pour le support des processus d'aide à la décision.

Les données d'un data warehouse sont caractérisées par :

- **Thématiques** : Les données sont organisées par sujet (ex : clients, produits, ventes, etc.).
- **Intégrées** : Les données, qui proviennent de diverses sources hétérogènes, sont consolidées et intégrées dans l'entrepôt.
- **Non-volatiles** : une fois insérées dans l'entrepôt, les données ne sont jamais modifiées ou effacées, elles sont conservées pour des analyses futures. (Inmon, 2005)

- **Historiques** : les données sont représentées dans une certaine période, c'est-à-dire lié à un référentiel temporel. (Zouaghi, 2019)

Le data warehouse permet de : (Godin, et al, 2011)

- Réaliser des analyses élaborées sur différents sujets d'affaires.
- Donner une vue unique et consolidée des données de l'entreprise.
- Assurer la rapidité et la qualité de l'information.
- Accéder plus facilement aux données.

### 2.2.2.2. Définition d'un datamart

Un datamart représente une partie du data warehouse conçue pour satisfaire les besoins de processus, départements ou utilisateurs déterminés. Il est construit à partir des données historisées et capturées d'une façon agrégée ou détaillée, relatives à ceux dont on satisfait le besoin, ou alors à partir des données d'un data warehouse pré-établi. (Ballard, 2006)

Autrement dit, un datamart est destiné à pré-agréger des données disponibles de façon plus détaillée dans le data warehouse, afin de traiter plus facilement certaines question spécifiques. (Crozat, 2017)

Les caractéristiques d'un datamart : (Godin, et al, 2011)

- Il comporte une partie du contenu du datawarehouse.
- L'on y traite un unique sujet d'analyse.
- Les analyses qui y sont effectuées sont simples et spécifiques.
- Il est alimenté à partir d'un nombre limité de sources.
- Il est modélisé suivant le schéma en étoile.

Après avoir défini le data warehouse et le datamart, il est essentiel d'introduire les notions qui leur sont sous-jacentes.

- **Faits :**

Une table de fait contient des faits, c'est-à-dire les mesures numériques produites par un événement opérationnel de mesure dans le monde réel. En plus de cela, elle contient toujours les clés étrangères de chacune de ses dimensions associées, ainsi que celles des dimensions dégénérées.

Les tables de faits sont utilisées pour les calculs et les agrégations réalisées à travers des requêtes.

## ❖ Types de tables de faits :

Il existe trois types de tables de faits :

- Transactionnelle : c'est la plus courante, le grain dans ce type de table correspond à une ligne par transaction. Les données sont dans le plus haut niveau de détail, ce qui justifie le nombre important de dimensions associées. (ZenTut, SD)
- Périodique : Dans ce type de table de faits une ligne résume un ensemble de mesures opérationnelles ayant eu lieu dans une période donnée, tels qu'un jour, une semaine, ou un mois. Le grain est représenté par la période et non par la transaction. (Kimball, SD )
- Accumulative : elle décrit l'activité d'un business process qui a un début et une fin bien déterminés. Dans ce type de table on peut trouver différentes colonnes « Date » qui représentent les dates auxquelles ont lieu les événements du processus. (ZenTut, SD)

## ❖ Comparaison entre les différents types de tables de faits :

Le tableau ci-dessous présente une comparaison entre les trois types de tables de faits :

**Tableau 2 : Comparaison entre les différents types de tables de faits (IBM, 2017)**

	<b>Transactionnelle</b>	<b>Périodique</b>	<b>Accumulative</b>
<b>Grain</b>	Une ligne par transaction	Une ligne par période de temps	Une ligne par événement
<b>Dimension</b>	La dimension « date » au niveau de granularité le plus bas	La dimension « date » au niveau de granularité de la fin de période	Plusieurs dimensions « date »
<b>Mesures</b>	Liées aux activités de la transaction	Liées aux activités de la période	Liées aux activités qui ont un début et une fin bien définis
<b>Performance</b>	Bonne performance, peut être améliorée en choisissant le grain le plus détaillé	Meilleure performance	Bonne performance
<b>Evolution de la taille table de faits</b>	Très rapide	Moyennement rapide	Lente
<b>Besoin d'agrégation des tables</b>	Élevé, car les données sont stockées à un niveau de détail très élevé	Faible besoin d'agrégation, les données étant déjà agrégées	Moins élevé que la table transactionnelle

### • Mesure :

C'est une valeur numérique qui représente une métrique ayant un sens pour l'activité de l'entreprise. (STERRET, 2009)

Les mesures peuvent être de deux natures : de base ou dérivées. (Wibas, 2015)

- Celles de base sont obtenues par une mesure opérationnelle directe.
- Quant à celles dérivées, les données dont elles dérivent proviennent d'autres mesures, en général en combinant plusieurs mesures de base.

❖ **Types de mesures :** (Zouaghi, 2019)

Ces mêmes mesures peuvent être de trois types :

- Additives : ce sont les mesures qui peuvent être agrégées par rapport à toutes les dimensions.
- Semi-additive : ce sont les mesures qui peuvent être agrégées par rapport à un certain nombre de dimensions.
- Non additive : mesures qui ne peuvent pas être agrégées.

- **Dimensions :**

Une dimension consiste en un regroupement d'attributs ayant une clé primaire unique. Cette clé primaire est présente dans la table de faits en tant que clé étrangère afin de décrire le contexte descriptif des faits. Les attributs de la table sont soumis à des contraintes à travers des requêtes pour répondre à des questions relatives à l'activité de l'entreprise. (Kimball, et al., 2013).

❖ **Types de dimensions :**

Il existe plusieurs types de dimensions :

- Dimension conforme : Il s'agit d'une dimension utilisée par plusieurs Data Marts à la fois. (DWgeek, 2018)
- Dimension débarras : C'est une combinaison d'attributs qui ne sont pas liés à une dimension particulière. (DWgeek, 2018)
- Dimension dégénérée : Clé de dimension sans attributs stockée dans la table de faits. (Zouaghi, 2019)
- Role-playing dimension : Dimension utilisée pour plusieurs usages comme une dimension date représentant plusieurs dates selon la relation. (DWgeek, 2018)
- Dimension statique : c'est une dimension qui n'est pas extraite de la source de données mais qui est créée dans le contexte de l'entrepôt de données. (Zouaghi, 2019)
- Dimension à évolution lente : c'est une dimension dont les attributs évoluent lentement dans le temps. (DWgeek, 2018)
- Dimension à évolution rapide : c'est une dimension dont les attributs évoluent rapidement dans le temps. (DWgeek, 2018)

### 2.2.2.3.Approches de modélisation

Il existe trois principales approches pour modéliser un data warehouse :

- **Inmon** : (Tehreem, 2020) (Zouaghi, 2019)

L'approche d'Inmon pour la conception d'un Datawarehouse est basée sur la méthodologie Top-down, elle permet d'optimiser l'espace de stockage et l'intégrité des données mais n'assure pas une très grande performance de requête.

Elle passe par :

- 1) La conception du modèle de données corporate qui identifie tous les sujets d'analyse et entités de l'entreprise comme client, produit, vendeur.
- 2) La création d'un modèle logique pour chaque entité primaire en s'assurant qu'il soit normalisé pour éviter la redondance à travers une modélisation en flocon.
- 3) La création du modèle physique, le "datawarehouse", unique source de vérité pour l'entreprise.
- 4) Création de datamarts spécifiques à certains services, dont les données dérivent du datawarehouse de l'entreprise, pour faciliter leurs requêtes.

- **Kimball** : (Tehreem, 2020) (Zouaghi, 2019)

L'approche de Kimball suit une méthodologie bottom-up, elle permet une optimisation de la performance des requêtes mais nécessite beaucoup d'espace de stockage et cause quelques risques de redondance des données.

Elle passe par les étapes suivantes :

- 1) Conception de datamarts dénormalisés (schémas en étoile) suivant une architecture en bus, c.a.d. liés par des dimensions conformes, à partir des besoins des utilisateurs finaux.
- 2) Identification des sources de données, pour ensuite extraire, transformer et charger les données dans les modèles dimensionnels.

Le tableau ci-dessous présente une comparaison des deux méthodes à travers leurs différentes caractéristiques :

**Tableau 3 : Comparaison Inmon vs Kimball (Breslin, 2004)**

	<b>Inmon</b>	<b>Kimball</b>
<b>Approche</b>	Top-down	Bottom-up
<b>Architecture</b>	Hub-and-spoke	Bus de magasin de données
<b>Schéma dimensionnel</b>	En flocon	En étoile
<b>Complexité de la méthodologie</b>	Complexe	Simple
<b>Détermination des Données</b>	Par sujet	Par process
<b>Facilité d'usage</b>	Difficile	Facile
<b>Public visé</b>	Service IT	Utilisateurs du processus
<b>Coût</b>	Élevé au départ, faible pour le développement	Bas

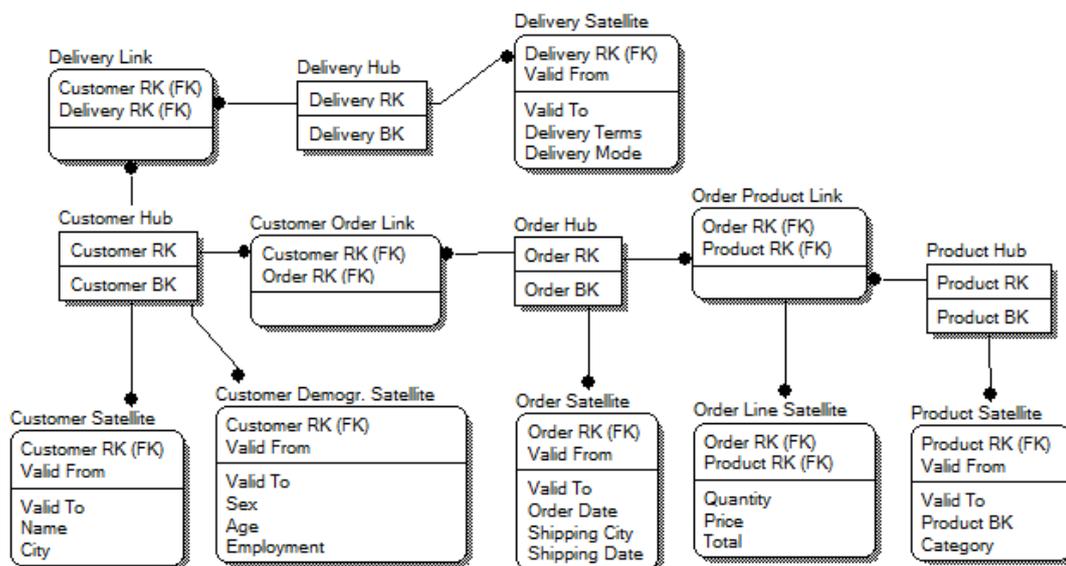
- **Datavault :** (Insidebigdata, 2017)

Data Vault est une approche de modélisation hybride représentant un mélange entre les deux approches précédemment citées.

Elle est composée de trois principaux éléments : les Hubs, les Links et les Satellites.

- Les Hubs sont des clés uniques.
- Les Links représentent les relations entre les hubs.
- les Satellites représentent des attributs des Hubs et des Links.

Son principal avantage est que son design est très flexible et peut facilement s'adapter aux changements de **besoins de l'entreprise**.



**Figure 2 : Modèle dimensionnel de type Datavault (Bukhantsov, 2012)**

#### 2.2.2.4. Architectures de données : (Chafki, et al, 2011)

On distingue cinq types d'architectures permettant d'organiser les données selon l'approche de modélisation adoptée.

- **Magasins de données indépendants :**

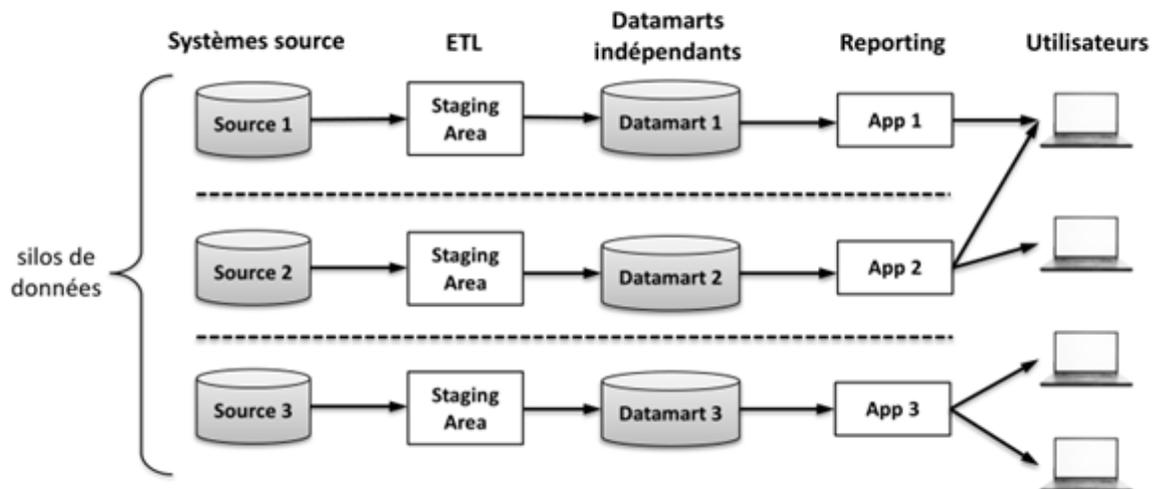


Figure 3 : Magasins de données indépendants (Chafki, et al, 2011)

C'est l'architecture la plus simple à développer. Elle est constituée de datamarts développés et opérant de manière indépendante, utilisant des données disposées en 'silos fonctionnels'.

Le manque de dimensions conformes entraîne des redondances et l'indépendance peut rendre difficile la réalisation d'analyses interfonctionnelles.

- **Bus de magasins de données :**

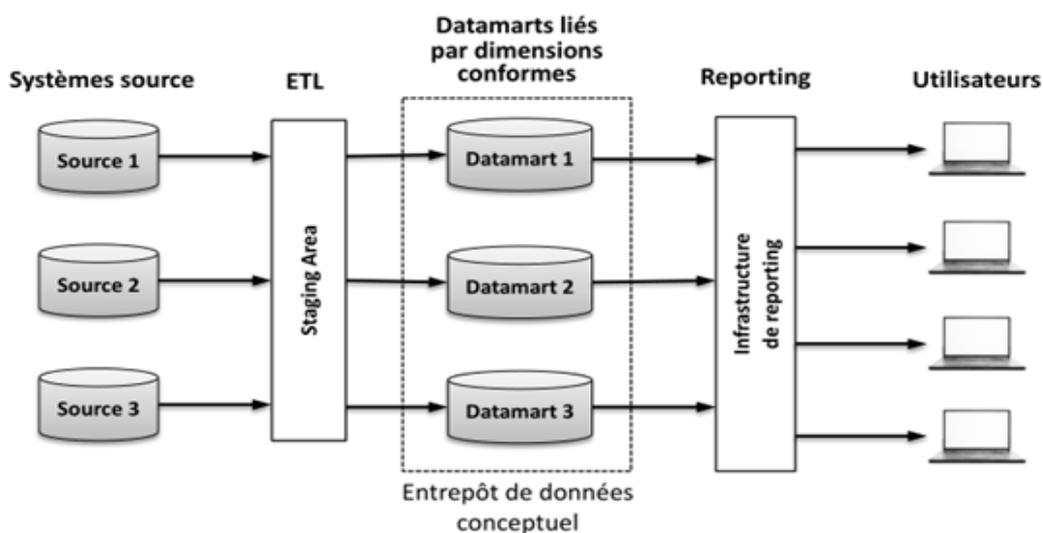
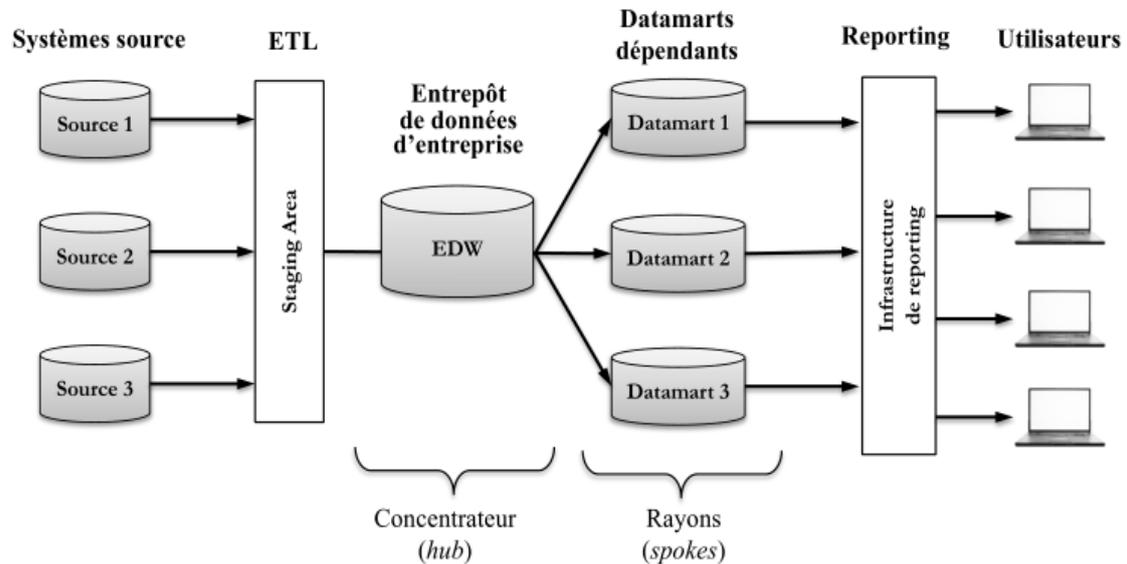


Figure 4 : Bus de magasins de données (Chafki, et al, 2011)

C'est architecture proposée par R. Kimball et basée sur le modèle en étoile dénormalisé. Elle adopte une approche incrémentale (Bottom-up) dans le développement des datamarts, qui

sont développés par processus d'affaires et sont inter-reliés par des dimensions conformes, formant ainsi un entrepôt de données conceptuel.

- **Architecture Hub-and-spoke :**

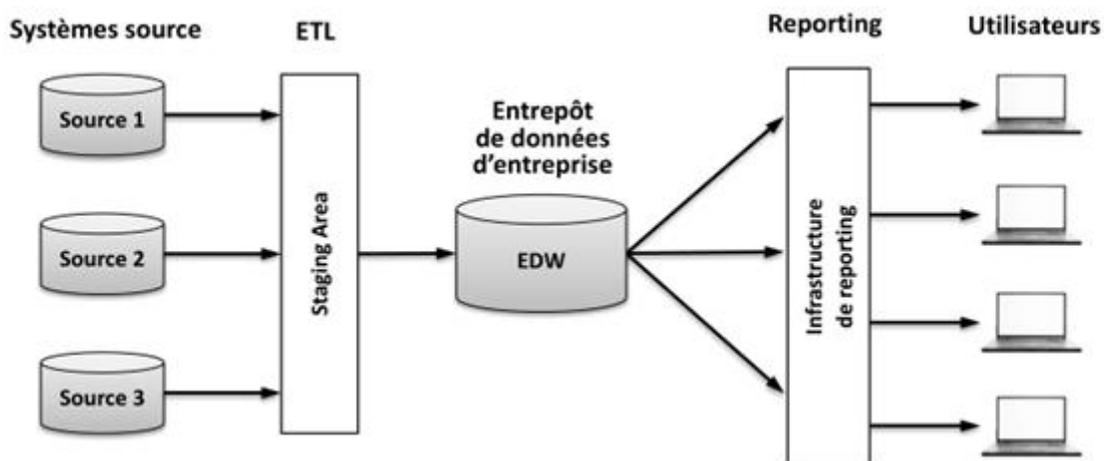


**Figure 5 : Architecture Hub-and-spoke (Chafki, et al, 2011)**

Proposée par B. Inmon et basée sur le modèle en flocon normalisé. Cette architecture assure une intégration complète des données grâce à l'adoption de l'approche Top-down.

Elle permet de construire des entrepôts avec des données au niveau de détail le plus fin, et qui sont agrégées dans les datamarts.

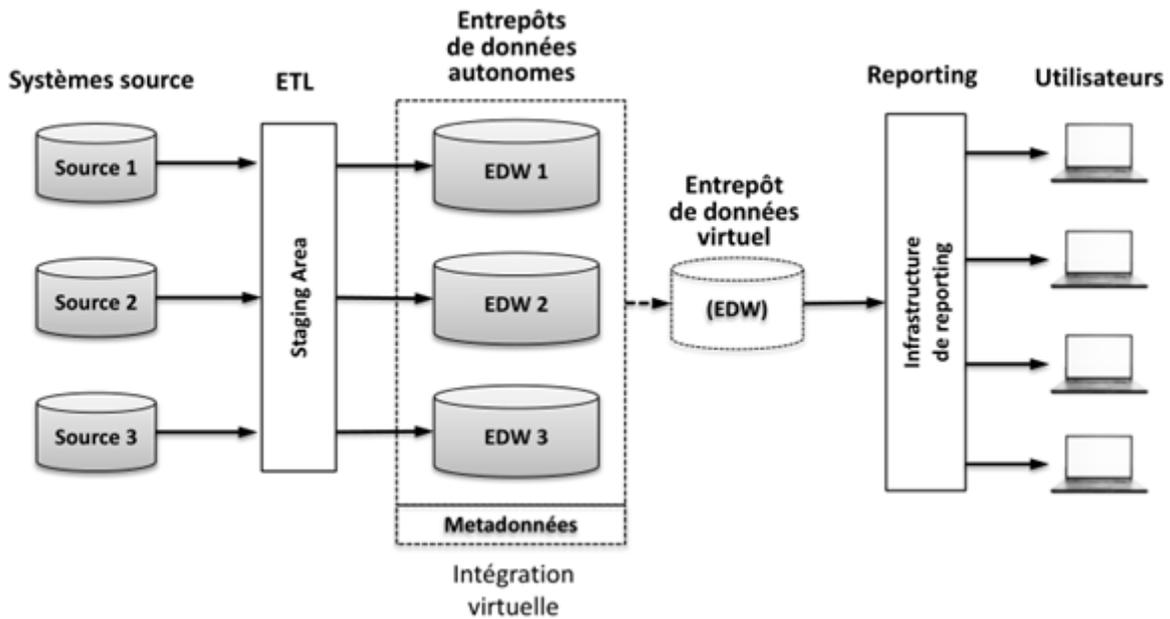
- **Entrepôt de données centralisé :**



**Figure 6 : Entrepôt de données centralisé (Chafki, et al, 2011)**

Cette architecture permet de faciliter l'intégration et la maintenance des données grâce à un unique entrepôt de données fonctionnant sans les datamarts, cela assure une performance optimale et un accès à toutes les données de l'entreprise mais peut être coûteux pour l'entreprise.

- **Architecture fédérée :**



**Figure 7 : Architecture fédérée (Chafki, et al., 2011)**

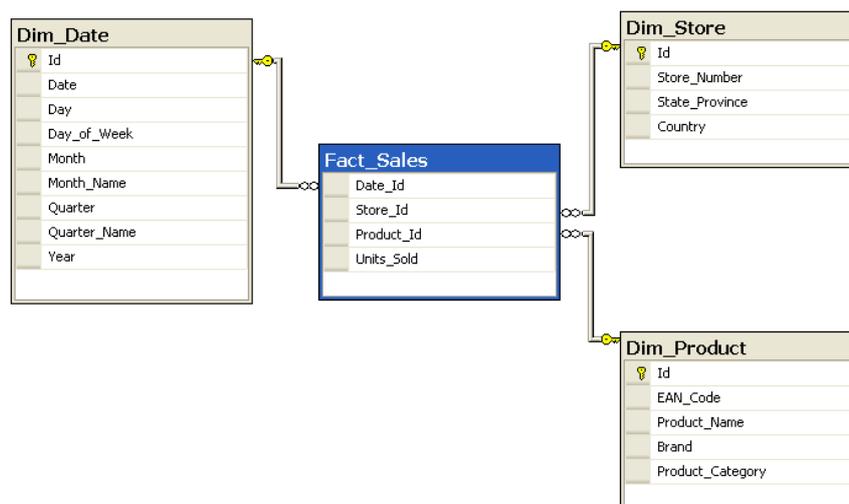
Architecture utile lorsqu'il y a déjà un entrepôt en place. Ce dernier est distribué sur plusieurs systèmes hétérogènes. Il assure une intégration logique ou physique des données à l'aide des métadonnées mais engendre une complexité supplémentaire et est peu performant.

#### 2.2.2.5. Schémas de modélisation d'un datawarehouse : (Torben, 2009)

La modélisation multidimensionnelle est le processus de modélisation des données à travers un modèle dimensionnel, qui catégorise les données en faits auxquels on associe des mesures ou en dimensions caractérisées par des attributs donnant un contexte aux faits.

Il existe trois types de schémas multidimensionnels :

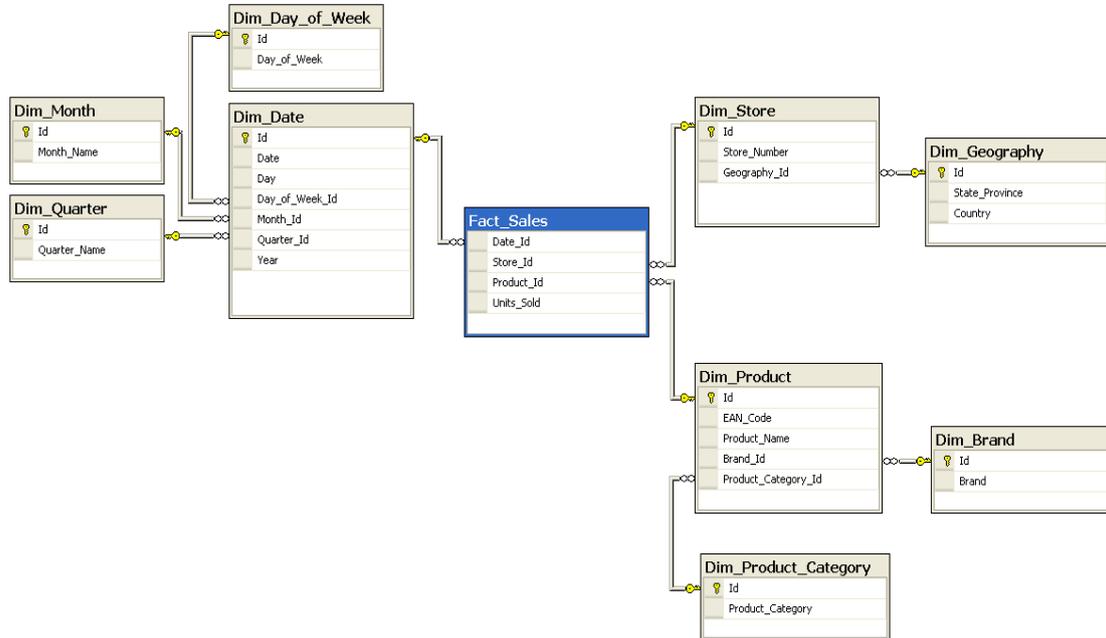
- **Schéma en étoile :** (Guru99, SD)



**Figure 8 : Exemple de schéma en étoile (Diffen, 2014)**

Comme son nom l'indique, le schéma en étoile est constitué d'une table de faits située au centre de l'étoile, et d'un ensemble de dimensions associées à cette table. C'est le modèle le plus simple de data warehouses permettant une performance de requêtes optimisée.

- **Schéma en flocon :** (Guru99, SD)



**Figure 9 : Exemple de schéma en flocon (diffen, 2014)**

C'est une extension du schéma en étoile, avec des dimensions normalisées et divisées en sous-dimensions. Il est plus complexe que le schéma en étoile, il permet de gagner en espace mais entraîne une perte de performance.

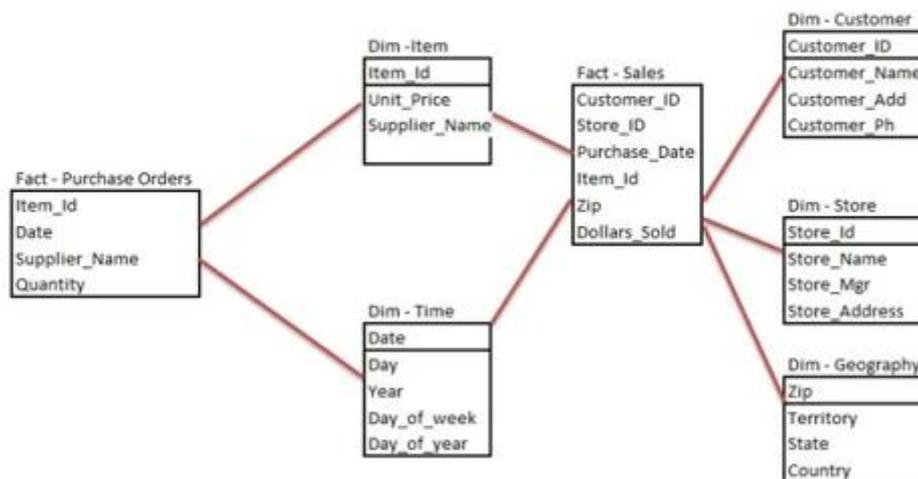
- ❖ **Schéma en étoile vs Schéma en flocon :** (diffen, 2014)

Le tableau ci-dessous présente une comparaison du schéma en flocon et du schéma en étoile :

**Tableau 4 : Comparaison schéma en flocon vs schéma en étoile (diffen, 2014)**

Caractéristiques	Schéma en flocon	Schéma en étoile
<b>Facilité de modification</b>	Il est facile d'y introduire des modifications car sans redondance	Moins facile à maintenir à cause de la redondance
<b>Facilité d'usage</b>	Un peu difficile à utiliser à cause de la complexité des requêtes	Facile à utiliser car les requêtes sont simples
<b>Performance de la requête</b>	Faible performance à cause du nombre élevé de jointures	Performance élevée
<b>Espace de stockage</b>	Optimise l'espace de stockage	Nécessite beaucoup d'espace de stockage
<b>Nombre de table par dimension</b>	Une ou plus	Une seule et unique table
<b>Normalisation</b>	Les dimensions sont normalisées	Les dimensions sont dénormalisées
<b>Approche</b>	Top-down (Inmon)	Bottom-up (Kimball)

- **Schéma Galaxy** : (Geeksforgeeks, SD)



**Figure 10 : Exemple de schéma galaxy (Dwbiloung, 2015)**

C'est une extension du schéma en étoile qui consiste en un ensemble de tables de faits ayant en commun quelques dimensions. Cela peut être vu comme une collection de plusieurs schémas en étoile, d'où l'appellation " Galaxy ".

### 2.2.2.6.Approche de Kimball

L'approche de modélisation de Kimball permet de réaliser des datamarts, et ce en suivant les étapes suivantes :

#### 2.2.2.6.1. Identification et analyse des besoins

L'identification et l'analyse des besoins métier sont une étape antérieure à la modélisation dimensionnelle, leur bon déroulement est crucial à la validité de cette dernière.

Pour réussir à identifier les besoins métier plusieurs questions peuvent être posées :

- Quelles fonctions doivent être analysées ?
- Comment mesure-t-on leur performance ?
- Quels sont les critères de succès ou d'échec ?
- Quelles données sont nécessaires et à quel moment ?
- Quelles données manquent à la prise de décision ?
- A quel niveau de détail cherche-t-on à analyser les données ?

Il existe deux méthodes d'identification des données : Orientée-source et orientée utilisateur :

- **Orientée source :**

Cette méthode se base sur l'utilisation des données sources disponibles dans les systèmes opérationnels, et ce à travers l'analyse du modèle entité-relation ou le modèle physique, en sélectionnant les éléments qui présentent un intérêt pour l'analyse.

L'avantage de cette méthode est qu'elle permet de savoir dès le départ quelles données sont disponibles, faisant ainsi gagner du temps. Son inconvénient est le manque d'interaction avec les utilisateurs.

- **Orientée utilisateur :**

Cette méthode se base sur la conduite d'entretiens avec les responsables des processus, où il leur est demandé de décrire les analyses qu'ils font ou souhaiteraient faire.

L'avantage de cette méthode est qu'elle permet de faire ressortir ce qui est réellement nécessaire. Son inconvénient est que les attentes des utilisateurs peuvent ne pas être réalisables à cause de l'indisponibilité de certaines données.

Une fois les besoins identifiés, ils sont analysés afin d'en tirer des entités et mesures préliminaires.

#### **2.2.2.6.2. Etapes modélisation dimensionnelle**

La modélisation est la technique préconisée par Kimball pour concevoir des modèles dimensionnels, elle consiste en 4 étapes :

##### **Etape 1 : Identifier les processus :**

- **Identification :** (IBM, 2014)

Dans cette étape on sélectionne les business process pour lesquels le modèle dimensionnel sera conçu. Basés sur cette sélection, les besoins de chaque processus sont déterminés. Un business process peut être traduit en plusieurs modèles dimensionnels. (Ballard, 2006)

Il faut préciser que les business process ne sont pas toujours équivalents aux départements, se baser uniquement sur les départements peut causer des problèmes relatifs aux données qui se retrouveront dupliquées.

Après avoir identifié les business process et les besoins, on procède à l'affectation des besoins aux business processus. (Benameur et Ferguene, 2016)

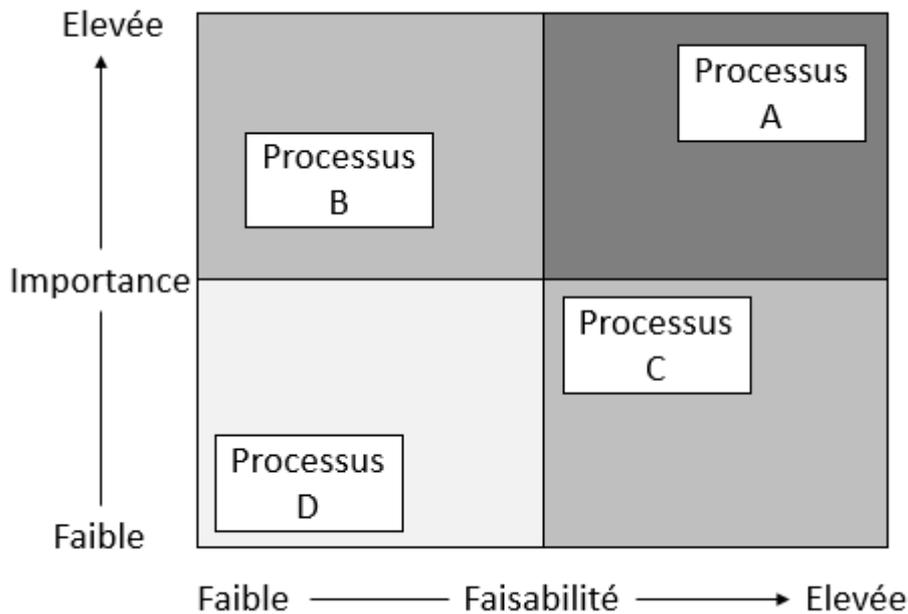
Les modèles multidimensionnels ne peuvent pas tous être réalisés en même temps dès la première itération, c'est pour cela qu'il est nécessaire de prioriser pour savoir quels processus seront réalisés en premier.

- **Priorisation :** (Kimball, et al. , 2013)

Ces processus sont jugés selon deux axes par les utilisateurs des processus : la faisabilité et l'importance par rapport au business, après cela ils sont présentés dans un schéma dans lequel la faisabilité est l'abscisse et l'impact l'ordonnée.

Les projets à forte faisabilité et importance sont ceux dont la réalisation peut être immédiatement entamée, ceux dans la partie droite en bas peuvent être intéressants sur le long

terme, ceux dans la partie gauche en haut nécessitent une concertation avec l'équipe IT pour résoudre le problème de faisabilité, tandis que ceux en bas à gauche sont à éviter.



**Figure 11 : Hiérarchisation des processus (Kimball et al. , 2007)**

**Etape 2 : Identifier le grain : (IBM, 2014)**

L'étape où l'on identifie le grain, qui selon Kimball est : « *La manière avec laquelle l'entreprise définit une mesure d'événement qui crée un enregistrement dans la table. Le grain dépend principalement de la structure des sources de données.* » (Kimball, 2007)

Pour l'identifier, il faut déterminer le niveau de détail et le contenu que l'on souhaite représenter dans la table de fait. Le résultat est que toute ligne de la table de fait contient le même type de données.

Plus on inclut de détail et plus le niveau de granularité est bas.

Cette étape est cruciale car si on s'y trompe il faudra reprendre la modélisation depuis cette étape.

Un business process peut comporter plusieurs grains, dans ce cas l'idéal serait de créer plusieurs tables de faits, une pour chaque grain.

**Etape 3 : Identifier les dimensions : (IBM, 2014)**

C'est l'étape où l'on identifie les dimensions qui permettent de caractériser les tables de fait.

Dans cette étape on passe par trois phases :

- 1) Identification de dimensions s'adaptant au grain des différentes tables de fait.

- 2) Identification des colonnes des dimensions, c-à-d des attributs ainsi que des hiérarchies liant les différents attributs.

#### **Etape 4 : Identifier les mesures : (IBM, 2014)**

C'est l'étape où l'on identifie les mesures qui s'adaptent au grain choisi pour la table de fait et répond au besoin des utilisateurs. Il faudra également déterminer leurs natures, leurs types ainsi que leurs méthodes de calcul s'il s'agit de mesures dérivées.

#### **La matrice de bus dimensionnel** : (Kimball, et al, 2013)

Cette matrice est recommandée pour communiquer la structure de l'architecture en Bus de Kimball.

Les lignes de cette matrice résument l'ensemble des processus de l'entreprise précédemment identifiés. Elles traduisent les activités principales de l'entreprise et ses principaux événements.

Les colonnes représentent les différentes dimensions utilisées dans l'entreprise. Il est souvent utile de lister toutes les dimensions avant de remplir la matrice, afin de pouvoir estimer si une dimension donnée peut être associée à un business process.

Lors du développement d'un DW, il est généralement recommandé de commencer par une ligne de la matrice, cela permet de se concentrer sur les résultats d'un seul processus et les systèmes source qui le définissent, ce qui réduit considérablement les risques de développement de l'ETL.

BUSINESS PROCESSES	COMMON DIMENSIONS						
	Date	Product	Warehouse	Store	Promotion	Customer	Employee
Issue Purchase Orders	X	X	X				
Receive Warehouse Deliveries	X	X	X				X
Warehouse Inventory	X	X	X				
Receive Store Deliveries	X	X	X	X			X
Store Inventory	X	X		X			
Retail Sales	X	X		X	X	X	X
Retail Sales Forecast	X	X		X			
Retail Promotion Tracking	X	X		X	X		
Customer Returns	X	X		X	X	X	X
Returns to Vendor	X	X		X			X
Frequent Shopper Sign-Ups	X			X		X	X

**Figure 12 : Exemple de matrice de bus dimensionnel (Kimball, et al, 2013)**

### 2.2.3. Phase de restitution ou de distribution

Cette phase permet de restituer les données adaptées à travers les outils de reporting, les tableaux de bord ... ce qui permet de faciliter l'accessibilité à l'information.

#### 2.2.3.1. Rapports

Le reporting est un ensemble d'outils de la BI. Plus précisément des outils de récupération d'informations, de réalisation de rapports et de leur partage avec les utilisateurs finaux, afin de leur fournir une base solide pour la prise de décision. (Techopedia, SD) (Durcevic, 2020) (Fernandez, SD)

Les rapports constituent un outil essentiel pour toute entreprise. Ils permettent de suivre et d'analyser la performance de l'entreprise tout en identifiant des champs d'amélioration et des opportunités de croissance.

L'objectif d'un rapport est de permettre une analyse critique de l'activité de l'entreprise, il permet également de guider la prise de décision et aide les employés ainsi que les managers à identifier et à résoudre leurs problèmes.

#### ❖ Comment sont-ils réalisés ?

A travers l'alimentation et la révision d'informations au sujet de domaines fonctionnels spécifiques à l'entreprise. Une fois ces informations regroupées et révisées, des conclusions peuvent être déduites du rapport et des décisions prises. (Unleashed Software, 2019)

## ❖ **Types de reporting :**

Il en existe deux types :

- **Reporting de masse :** ( Marshall, 2020)

Il s'agit de rapports caractérisés par leur réutilisabilité, ils sont basés sur des données répondant à un ensemble récurrent de spécifications ou de centres d'intérêts de l'entreprise.

Ils ciblent généralement une large audience, ont un visuel soigné et couvrent une période temporelle déterminée comme une semaine ou un mois.

Ces rapports ont souvent un format prédéterminé et sont donc aisés à recréer avec les outils BI.

- **Reporting ad-hoc :** (Izenda, 2019)

Il s'agit de rapports dynamiques créés en temps réels pour répondre à un besoin. Ils sont en général créés pour répondre à des questions spécifiques soulevées par un événement à travers la réalisation de requêtes personnalisées.

## ❖ **Importance des rapports :** (Bizfluent, 2018) (Unleashed Software, 2019)

- Offrent un aperçu utile au management sur des informations concernant les dépenses, les profits et la croissance, ce qui pourra être utilisé par la suite pour développer de futures prévisions, des plans marketing, une planification du budget ainsi que pour faciliter de la prise de décision.
- Permettent un suivi du progrès, identification de tendances ou d'irrégularités sur une période déterminée, tout en facilitant l'audit des activités de l'entreprise.
- Simplifient des idées compliquées et les exposent d'une manière claire et compréhensible.
- Peuvent constituer un outil de persuasion, par exemple pour convaincre des investisseurs d'investir leur capital en une idée.
- Servent de supports à d'importantes décisions et à la résolution de problèmes.

### **2.2.3.2. Tableaux de bord**

Le tableau de bord est un outil complémentaire aux activités de reporting. Il est conçu de manière consensuelle et périodique à partir des résultats réalisés par l'entreprise, qui sont représentés sous forme d'indicateurs de performance de l'activité.

Le but du tableau de bord est de faciliter le pilotage de l'entreprise à travers une représentation simplifiée de ses résultats, afin de permettre aux managers de prendre de meilleures décisions à tous les niveaux, pour ensuite décliner ces décisions en plans d'action visant l'atteinte des objectifs de performance. (Usinenouvelle, 2015)

Les indicateurs constituant un tableau de bord doivent être :

- **Pertinents** : c-à-d pouvant répondre, au moment voulu, aux différents besoins pour lesquels le TDB a été conçu.
- **Rapidement obtenus** : Cela leur permet une rapidité de mise en place des actions correctives.
- **Synthétiques** : ils doivent montrer une image complète de l'entreprise ou de l'activité du responsable.
- **Contingents** : permettent de répondre aux attentes du moment.

❖ **Rôle du tableau de bord** : (Compta-facile, 2016)

- Aide au contrôle de gestion en donnant une vue sur le niveau de performance atteint et en mettant en relief les dysfonctionnements.
- Permet aux managers de mieux communiquer entre eux.
- Rend plus simple la prise de décision et la mise en œuvre d'actions correctives.

❖ **Types de tableaux de bord**

Il existe trois types de tableaux de bord :

**1) Tableau de bord stratégique**

Il s'agit d'une traduction des objectifs stratégiques de l'entreprise en objectifs opérationnels afin de permettre la mise en place de plans d'action. Il constitue un outil de pilotage sur le long terme.

**2) Tableau de bord budgétaire**

C'est un tableau de bord utilisé pour mettre en exergue l'écart entre les prévisions et les résultats réels du point de vue financier. Il doit permettre l'analyse de ces écarts pour aider à la mise en place d'actions correctives. Il constitue un outil de pilotage au moyen terme.

**3) Tableau de bord opérationnel**

Il s'agit d'un tableau de bord utilisé pour mesurer l'avancement et la performance des plans d'actions mis en place. Il constitue un outil de pilotage au court terme.

**2.2.3.3.KPIs (Key Performance Indicators): (KPI,SD)**

Les indicateurs clés de performance sont des indicateurs critiques qui montrent la progression vers un résultat attendu.

Ils permettent de se concentrer sur l'amélioration de l'activité tant sur le niveau stratégique que sur le niveau opérationnel, de créer une base analytique pour la prise de décision et de se baser sur les faits les plus importants.

Manager en utilisant les KPIs comprend la fixation d'objectifs et le suivi de la progression vers l'atteinte de ces objectifs.

#### ❖ **Caractéristiques des bons KPIs**

- Donnent des preuves objectives de la progression pour la réalisation du résultat désiré.
- Mesurent ce qui doit être mesuré pour améliorer la prise de décision.
- Permettent de mesurer le degré de changement de la performance durant le temps.
- Peuvent suivre l'efficacité, l'efficacit , la qualit , la ponctualit , la gouvernance, la conformit , les comportements, la performance...

#### **2.2.4. Phase d'analyse**

C'est la phase o  les utilisateurs finaux peuvent analyser les informations qui leurs sont fournies   travers les outils de statistiques.

##### **2.2.4.1.OLAP**

Un cube OLAP est une structure dimensionnelle impl ment e dans une BDD multidimensionnelle, son contenu peut  tre  quivalent au sch ma en  toile ou en d river. Le cube dimensionnel (hypercube) contient les faits   analyser repr sent s comme des points dans l'espace multidimensionnel, et les dimensions qui repr sentent les axes d'analyse. (Kimball, et al. ,2013) (Darmont, et al, 2017)

C'est l' tape finale du d ploiement d'un syst me d cisionnel, il permet d'analyser les donn es   travers l'ensemble d'op rations suivantes : (Godin, et al, 2011) (Tutorialspoint, SD)

- **Op rations de s lection :**

- Slice : c'est une s lection appliqu e sur une dimension, qui a pour output une tranche du cube selon la valeur qu'on souhaite analyser.
- Dice : il s'agit d'un slice appliqu  sur deux dimensions ou plus, permettant ainsi de former un sous-cube du cube dimensionnel.

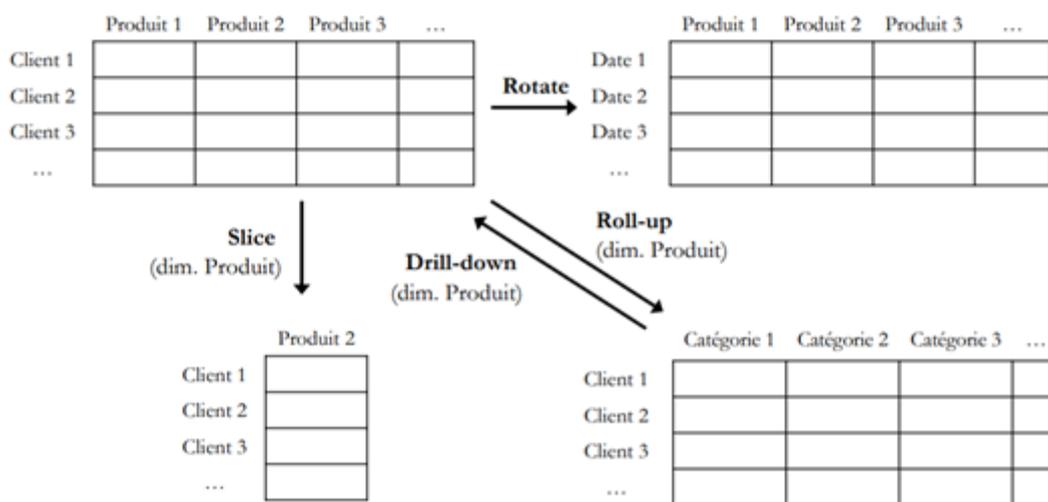
- **Op rations de forage :**

- Roll-up : permet d'appliquer une agr gation sur le cube en montant dans le niveau de d tail d'une dimension.
- Drill Down : c'est l'op ration inverse du Roll-up, permettant de descendre dans le niveau de d tail d'une dimension.

- **Opérations de rotation :**

- Rotation (pivot) : permet d'avoir une autre présentation des données en faisant tourner le cube.
- Drill-across : cette opération permet d'analyser différemment les tables de faits et ce à travers les dimensions conformes. Elle est divisée en deux étapes : il s'agit d'appliquer des requêtes séparées sur des tables de faits partageant une même dimension conforme pour obtenir les données, puis combiner les résultats obtenus.

Le schéma de la figure suivante résume les opérations assurées par OLAP :



**Figure 13 : Opérations réalisées par OLAP (Godin et al, 2011)**

#### 2.2.4.2. Data mining : (Berry, 2004)

Le data mining est l'exploration et l'analyse de larges quantités de données afin de découvrir des patterns et lois significatifs, il peut aider les entreprises en leur permettant d'améliorer leur marketing, leurs ventes ou leur compréhension de leurs clients.

Le data mining permet de construire des modèles, c'est-à-dire un algorithme ou ensemble de lois connectant un ensemble d'éléments en entrée à un élément de sortie particulier.

Ses principales tâches sont :

- **Classification :** Le but est d'assigner des éléments d'entrée à valeurs discrètes à des classes prédéterminées, par exemple pour assigner des demandeurs de crédit à des classes selon leur risque.
- **Estimation :** Pour estimer la valeur d'une fonction continue comme le revenu ou la valeur d'un compte bancaire, souvent utilisée avant une classification pour pouvoir y utiliser les valeurs obtenues.

- **Prédiction** : Cette tâche constitue une sorte de classification mais qui se base sur des valeurs futures des variables, elle est par exemple utilisée pour savoir quels clients quitteront l'entreprise dans un intervalle de temps donné.
- **Règle d'association** : Permet de déterminer quels éléments vont ensemble, par exemple pour savoir quels produits sont achetés en même temps ainsi que choisir quels éléments faire figurer ensemble sur l'étagère d'un supermarché.
- **Clustering** : Permet de segmenter une population hétérogène en un certain nombre de "clusters" plus homogènes. Contrairement à la classification, les classes ne sont pas prédéterminées, il appartient à l'analyste de donner un sens à chaque classe. Elle peut par exemple être utilisée pour regrouper les clients qui ont les mêmes habitudes de consommation.
- **Profiling** : Permet de mieux décrire la variable ciblée en dégagant un profil et ainsi en avoir une meilleure compréhension.

#### ❖ Quel lien avec le datawarehousing ?

Le datawarehouse permet une application plus simple et plus aisée du data mining, et ce pour plusieurs raisons :

- 1) Le data mining nécessite de larges quantités de données sur une large période de temps, généralement accessibles dans le datawarehouse.
- 2) Le data mining utilise des données nettoyées et fiables, ce qui représente un critère pour les données chargées lors de l'ETL.
- 3) Le datawarehouse permet de tester les hypothèses formulées lors du data mining.
- 4) Contrairement aux systèmes opérationnels, le datawarehouse offre une disponibilité continue, faisant ainsi gagner un temps précieux pour les calculs volumineux du data mining.

### 3. Conclusion

Dans ce chapitre nous avons expliqué les différentes notions relatives à la Business Intelligence, et ce, en détaillant les différentes phases d'un projet BI, depuis l'alimentation jusqu'à l'analyse, en passant par la modélisation et la restitution,

Ces notions nous permettront par la suite de bien formuler notre problématique, ainsi que de concevoir et mettre en place une solution adéquate et pratique.

## **Chapitre II : Etat des lieux**

## 1. Introduction

Ce chapitre sera dédié à l'état des lieux de l'entreprise Unilever. Ceci se fera en deux temps.

Dans un premier lieu, nous présenterons le secteur de la grande distribution dans le monde et en Algérie, ensuite nous ferons une présentation de l'entreprise Unilever et du service, le Customer Service, où nous avons effectué notre stage.

Ensuite, nous allons réaliser une étude de l'existant où nous allons retracer les processus du service, faire un inventaire de ses KPIs et de ses activités de reporting et auditer la qualité des données. Cette étude de l'existant nous permettra de mieux identifier les dysfonctionnements auxquels nous tenterons de remédier à travers la conception et l'implémentation d'une solution dans les prochains chapitres.

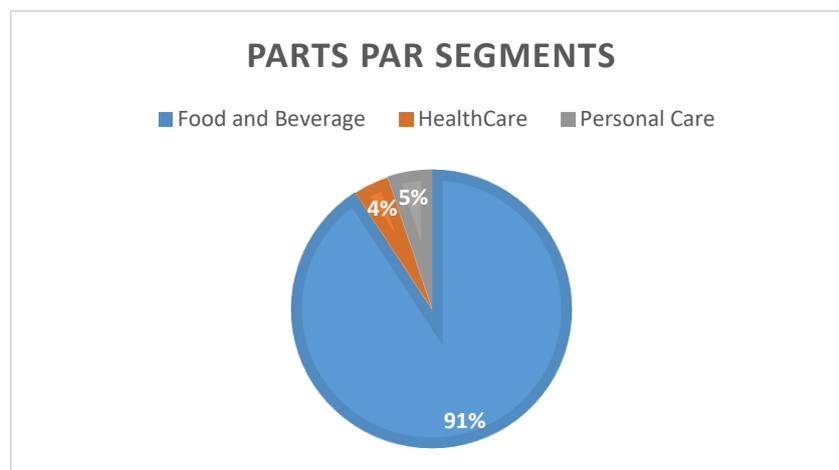
## 2. Présentation du secteur des biens de grande consommation (FMCG)

### 2.1. Dans le monde

Le marché des biens de grande consommation (FMCG) est un marché qui a connu une grande croissance lors des précédentes décennies, et ce après avoir bénéficié de l'essor de la R&D, du shopping en ligne et de l'amélioration de la situation financière des consommateurs. Néanmoins il rencontre quelques défis à cause de la concurrence entre les principaux acteurs et la conscience grandissante des populations à la qualité des biens qu'elles consomment.

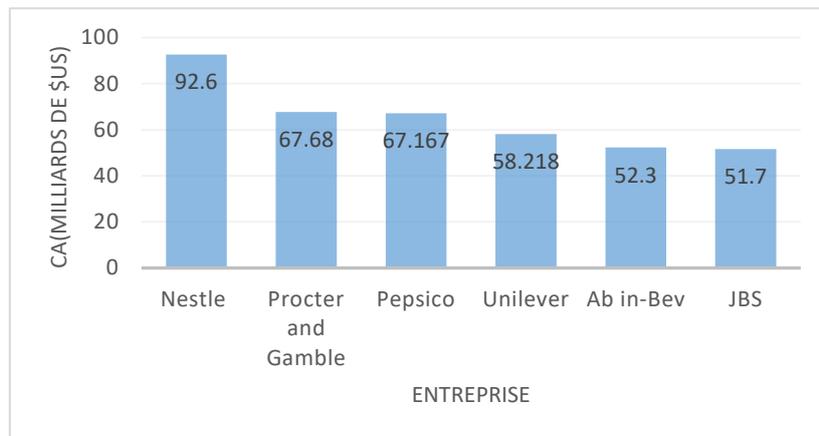
En 2017, les parts de ce marché étaient divisées entre les trois principaux segments comme suit :

- 91 % pour le segment Food and Beverages.
- 4 % pour le segment Healthcare.
- 5 % pour le segment Personal Care. (Alliedmarketresearch,2018)



**Figure 14 : Segmentation du marché des bien de grande consommation 2017 - Réalisé par nous-mêmes à partir des données de (Alliedmarketresearch,2018)**

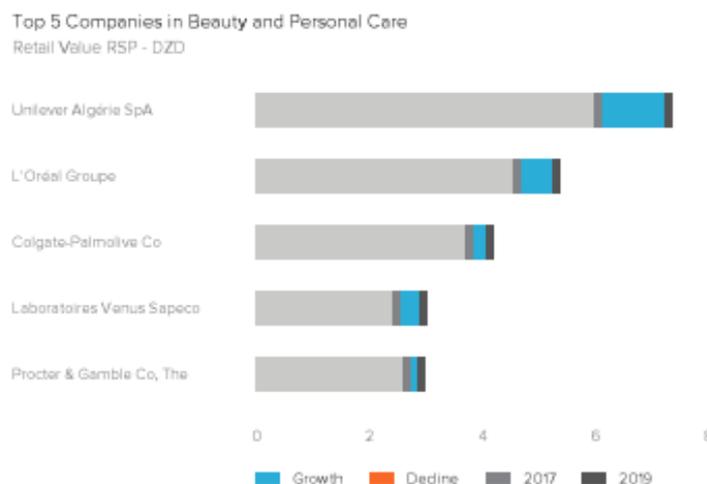
Les principaux leaders du marché sont représentés dans la figure ci-dessous avec leur CA de l'année 2019. (Graber, 2020) (Reiff, 2020) (Macrotrends,2020) (Macrotrends,2020) (forbes, 2019) (Statista, 2019)



**Figure 15 : Top-6 des entreprises marché FMCG selon leur CA de 2019 - Réalisé par nous-mêmes à partir des données de : (Graber,2020) (Reiff ,2020) (Macrotrends,2020) (Macrotrends,2020) (forbes, 2019) (Statista, 2019)**

Pour l'année 2019, Unilever a réalisé un chiffre d'affaires de 58.22 B\$, ce qui lui a permis de se placer au quatrième rang dans le classement des entreprises du marché FMCG derrière Nestle, Procter & Gamble et Pepsico.

## 2.2.En Algérie



**Figure 16: Top-5 des entreprises du marché algérien Beauty and Personal Care (Euromonitor, 2020)**

Unilever ayant trois portefeuilles de produit, elle est donc en concurrence avec des entreprises sur les marchés de chacun de ces Portefeuilles.

Sur le marché Beauty and Personal Care, Unilever Algeria est leader, elle est l'entreprise qui a enregistré ou affiché la plus grande croissance du chiffre d'affaires. Plusieurs autres entreprises partagent avec elle ce podium : L'oréal Groupe, Colgate-Palmolive Co, Laboratoires Venus Sapeco, Procter & Gamble Co.

### **3. Présentation d'Unilever**

#### **3.1. Unilever Monde**

Créée en 1929 suite à la fusion des deux entreprises : Margarine Unie et Levers Brothers, fabricant de savon, Unilever est un géant mondial spécialisé dans les biens de consommation et dont les sièges sociaux sont situés à Londres et à Rotterdam. Son chiffre d'affaires pour l'année 2019 s'est élevé à 52 milliards de dollars. Elle emploie plus de 160000 personnes à travers le monde en étant présente dans 190 pays avec plus de 400 marques.

Ces marques représentent des marques internationales comme Lipton, Knorr, Dove, Omo, ou des marques locales comme Brillhante au Brésil et Brooke Bond en Inde. (Unilever, SD)

##### **3.1.1. Objectifs et Activités : (Unilever, SD)**

- **Objectifs :** Unilever vise à promouvoir le développement durable, et ce en visant la réalisation de trois objectifs :
  - Aider plus d'un milliard de personnes à améliorer leur santé et leur bien-être en leur fournissant des produits d'hygiène.
  - Réduire de moitié l'impact environnemental associé à la fabrication et à l'utilisation de ses produits.
  - Améliorer les conditions de vie de millions de personnes en leur offrant davantage de possibilités d'emploi.
- **Activités :** Représentée par plusieurs centaines de marques, les produits d'Unilever sont répartis en trois portefeuilles de produits :
  - Le portefeuille Beauty and Personal Care orientés vers les soins de l'hygiène corporelle (42% du chiffre d'affaires en 2019).
  - Le portefeuille Food and Refreshments orientés vers la nutrition (37% du chiffre d'affaires en 2019).
  - Le portefeuille Home Care orientés vers les produits d'entretien ménager (21% du chiffre d'affaires en 2019).

##### **3.1.2. Historique : (Sutori, 2018) (CompaniesHistory, SD)**

L'histoire d'Unilever a été marquée par plusieurs événements, qu'il est possible de résumer par :

- **Création d'Unilever en 1929** suite à la fusion de deux entreprises : Margarine Unie et Levers Brothers, fabricant de savon.
- **1940-1949** : Concentration sur les marchés locaux.
- **1950-1959** : Expansion dans de nouveaux marchés de masse : Afrique et Asie.
- **1960-1969** : Nouveaux produits, nouveaux marchés dans un cadre économique favorable.
- **1970-1979** : Baisse des ventes due à un cadre économique défavorable.
- **1990-1999** : Restructuration et consolidation : Diminution des catégories de produits de 50 à 13.
- **2000-2009** : Transformations stratégiques et organisationnelles.
- **2010-Actuel** : Développement d'une stratégie axée sur le développement durable et l'impact positif.

### **3.2.Présentation d'Unilever Algerie :** (Bendiab et Sattouf, 2018) (Zennir,2019)

Unilever Monde est présente dans 7 grandes régions :

- MENA ( Middle East and North Africa);
- Arabia (Pays du golfe) ;
- Asia ;
- TRUB (Turquie- Russie - Ukraine - Biélorussie) ;
- America ;
- Europe ;

Unilever Algérie fait partie d'Unilever Maghreb qui est une filiale d'Unilever MENA (Middle East and North Africa).

Les activités d'Unilever Algérie se sont lancées en 2002 au niveau d'Oran, où se trouve son unité de production, dans la zone industrielle de Hassi Bounif. Il a été aussi implanté un bureau de liaison à Alger-afin de gérer les activités de Marketing, Supply & Demand planning aussi bien pour les produits locaux que pour les produits importés.

#### **3.2.1. Produits**

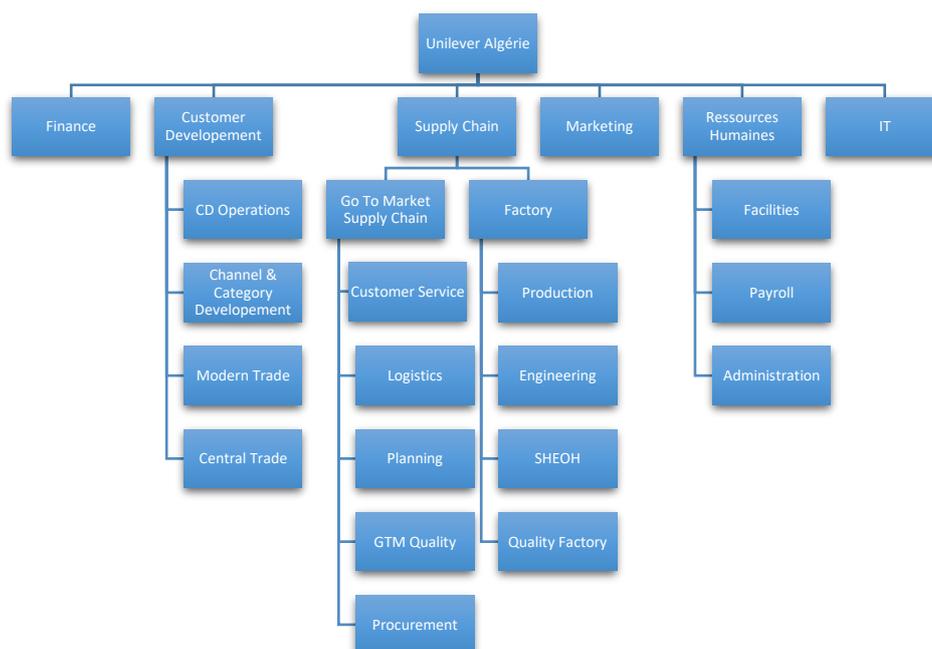
Unilever fabrique et commercialise trois catégories de produits en Algérie, ces derniers sont représentés le tableau ci-dessous :

**Tableau 5 : Produits d’Unilever commercialisés en Algérie (Zennir,2019)**

Catégories	Sous-catégories	Type	Marque
Home care	Détergent	Détergent Linge	Omo
		Multi-usage	JIF Surf
Personal Care	Hair Care	Shampoing	Dove
		Après-shampoing	Clear Sunsilk
	Skin Cleansing	Savon	LUX, Dove
		Gel-douche	
Déodorant		Rexona Axe Dove	
	Stick	Rexona Dove	
	Oral	Dentifrice Brosse-à-dent	Signal
Food	Rafraîchissements		Lipton

### 3.2.2. Organigramme

Unilever Algérie est organisée en six directions, chacune d’entre elles est divisée en sous directions comme montre l’organigramme ci-dessous :



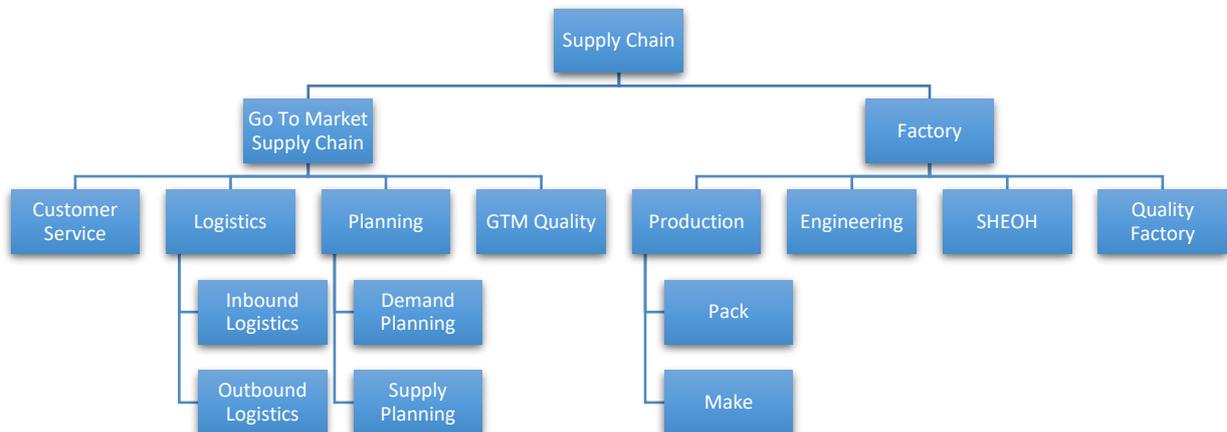
**Figure 17 : Organigramme d’Unilever Algeria (Zennir,2019)**

Unilever Algérie est constituée de 6 départements : (Zennir, 2019), (Entretien avec le Process Control and logistics claims Specialist le 23 Juin 2020)

- **IT** : Gère le hardware et le software et intervient en cas de problème informatique.
- **RH** : Gère les recrutements, les formations, l'administration des payes ainsi que les conditions de travail et les évènements.
- **Marketing** : Elabore les stratégies de vente des produits et supporte les actions de Customer Development.
- **Customer Development** : Il s'agit du commercial, il gère et suit les ventes secondaires et le feedback du marché. En plus de développer le portefeuille client dans le secondaire et la stratégie commerciale.
- **Finance** : Gère la comptabilité, la trésorerie, le budget et les taxes.
- **Supply chain** : Gère la planification et la réalisation de la production destinée à la vente ainsi que le suivi des commandes depuis la réception jusqu'à la livraison.

### 3.2.2.1. Organigramme de la supply chain : (Zennir,2019), (Entretien réalisé avec le Process Control and Claims Specialist le 23 Juin 2020)

L'organigramme ci-dessous représente la division Supply Chain d'Unilever Algérie :



**Figure 18 : Organigramme de la division Supply Chain (Zennir,2019)**

La Division supply chain effectue deux fonctions : Factory et Go-To-Market Supply chain.

- Factory est composée de 4 sous-divisions :

- **Production** : Assure la transformation de la matière première en produits finis.
  - **Engineering** : Gère la maintenance de l'usine.
  - **SHEOH** : Assure la sécurité, l'hygiène et le respect de l'environnement dans l'usine.
  - **Quality Factory** : Assure le respect des standards de qualité et l'audit qualité.
- Go-to-market Supply chain est également constituée de 4 sous-divisions :
    - **Planning** :
      - 1) Demand Planning : Assure la prévision des ventes et la planification de la production.
      - 2) Supply Planning : Assure la disponibilité des produits finis et des matières premières nécessaires à la satisfaction des commandes.
    - **Logistics** :
      - 1) Inbound : Assure la réception, le stockage et le transport des matières premières et des produits importés aux usines, ainsi que les formalités douanières.
      - 2) Outbound : Assure le transport des ventes aux distributeurs.
    - **GTM Quality** : Assure la qualité du point de vue logistique.
    - **Customer Service** : Où nous avons eu l'occasion de réaliser notre travail que nous présenterons ci-dessous.

### 3.2.2.1.1. Présentation du service Customer Service :

Ce service est l'interface de l'entreprise avec le client. Il prend en charge les commandes des produits depuis leur réception jusqu'à leur livraison, ainsi que la facturation et la gestion des réclamations et des retours.

Le service se charge des ventes primaires : ses clients sont des distributeurs répartis dans 5 grandes régions de l'Algérie, qui à leur tour se chargent d'assurer les ventes secondaires.

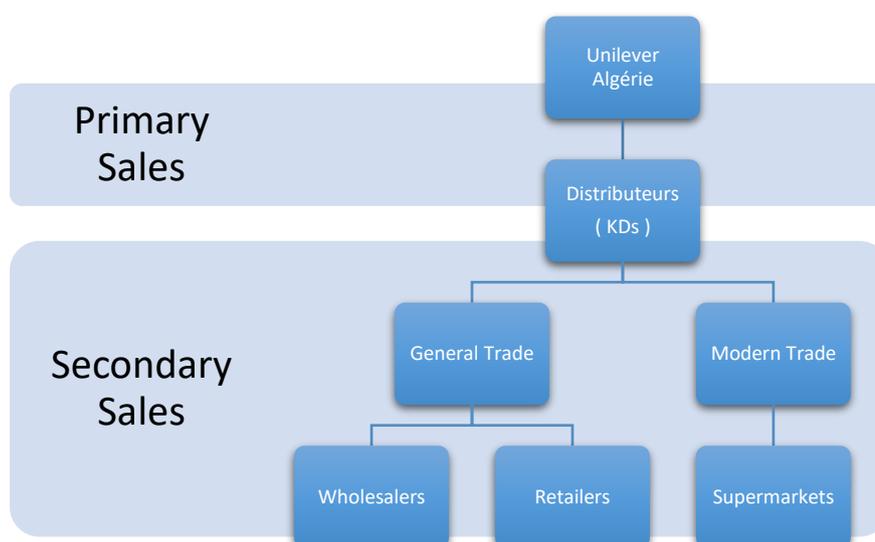
L'équipe est constituée de 8 membres :

- **Manager** : Chargé du suivi des tâches opérationnelles et du management de l'équipe
- **02 Customer Facing** : Chargés de recevoir les commandes et réclamations des clients

- **Stock Management Specialist** : Chargé de l'allocation du stock des commandes.
- **Order Entry Validation** : Chargé de l'entrée des commandes sur système.
- **Billing Specialist** : Chargé de la facturation et de l'archivage des factures.
- **Process Control and logistics claims Specialist**: Chargé du contrôle du processus, de la gestion des réclamations.
- **TTS Specialist** : Chargé de la gestion des remises.

### 3.2.3. Organisation de la distribution : (Bendiab et Sattouf, 2018) (Zennir,2019)

La figure suivante est une représentation de l'organisation de la distribution d'Unilever Algérie



**Figure 19 : Organisation de la distribution d'Unilever Algérie (Zennir,2019)**

Le réseau de distribution d'Unilever couvre toute l'Algérie et se divise en deux : Ventes primaires (Primary sales) et Ventes secondaires. ( Secondary sales)

Les ventes primaires sont gérées par le customer service, il s'agit de 27 distributeurs, répartis sur les 5 régions (Centre Est, Centre Ouest, Est, Ouest, Extrême-Est) que couvre Unilever.

Quant aux ventes secondaires elles sont gérées par le Customer Development operations et sont divisées en deux :

- **General Trade** : Concerne les grossistes, semi-grossistes et détaillants.
- **Modern Trade** : Concerne les supermarchés comme Ardis, Uno, Family Shop.

#### 4. Etude de l'existant

Dans la section suivante nous allons présenter les résultats de l'étude de l'existant que nous avons effectuée dans le service Customer Service qui avait pour but de cerner le processus et ses activités, les rapports établis par le service, la qualité des données exploitées ainsi que les indicateurs de performance. (Cf Annexe 2)

Nous avons été sollicitées par le service dans le but d'essayer de faciliter l'activité du reporting, par l'implémentation d'un système décisionnel ; chose qui est impossible à réaliser sans une bonne connaissance des processus ainsi que des activités de reporting et de la qualité des données disponibles. C'est pourquoi nous avons jugé impératif d'effectuer une étude de l'existant au sein de l'entreprise.

L'objectif de ce projet pour les utilisateurs est de pouvoir faire des analyses plus poussées et de pouvoir réaliser des rapports en moins de temps qu'ils n'en mettent dans la situation actuelle.

Pour la réalisation de l'étude de l'existant nous avons utilisé la méthode d'entretiens individuels en ligne avec différents membres du service, et ce pour plusieurs raisons : (Critzr, 2016)

- L'impossibilité d'aller au siège de l'entreprise à cause de la pandémie de coronavirus.
- Les horaires de disponibilité des employés étaient différents.
- les employés se sentent plus libres de s'exprimer.

##### 4.1. Processus

Dans ce qui suit, nous allons faire une description du processus **Order To Cash** géré par le Customer service :

- 1) Chaque début du mois un fichier Excel est partagé avec tous les distributeurs : le **KD Order**, avec le minimum Order quantity.
- 2) Le Customer Facing reçoit des **Purchase Orders** de la part des distributeurs, sous forme de fichiers excel, dont il vérifie d'abord les quantités. Si elles ne sont pas valides il recontacte le client pour qu'il reprenne commande, sinon il continue et divise les commandes selon plusieurs critères :
  - **Contraintes logistiques** : L'unité de commande est un camion recevant 33 palettes.
  - **Contraintes légales** : Les produits locaux et ceux importés sont séparés, les produits cosmétiques et ceux détergents également.

Ces fichiers sont par la suite partagés avec l'Order Entry.

- 3) L'Order Entry vérifie le fichier partagé par le Customer facing puis crée les **Sales Order** correspondants sur SAP.

L'Order Entry le partage par la suite avec le Stock Management Specialist pour vérifier les disponibilités. Avant de procéder à cela, une vérification de la situation financière du client est nécessaire, pour savoir si sa commande est bloquée ou non chez les finances.

4) Ici l'étape change selon le type de client :

- **Client récent** : La relation n'étant pas une relation de confiance, le Bill Specialist fait une facture pro-forma et la partage avec le client. Le client paye, une fois le paiement reçu il est demandé aux finances de débloquer le client.
- **Ancien client** : On travaille à crédit s'il n'a pas dépassé sa limite et il a 20 jours pour payer le SO, sinon, on lui demande de réguler sa situation.

5) Le Stock Management Specialist confirme les disponibilités des produits, puis alloue les stocks des produits à chaque commande.

Si les produits ne sont pas disponibles, il le partage comme déficit de stock et il faudra faire un retour vers le client, auquel il est demandé par le Customer Facing de changer sa commande. Cela est réitéré jusqu'à ce que les produits que souhaite commander le client soient tous disponibles.

Sinon, le stock est disponible et la préparation de la livraison commence.

- 6) Des Delivery Orders exprimant le besoin en termes de nombre de camions, de destinations, de lieux de stockage, de date de livraison sont créés. Ils sont par la suite partagés avec l'équipe logistique, qui est une flotte externalisées, puis est réalisée l'allocation finale des stocks, c'est-à-dire par rapport à la disponibilité physique réelle.
- 7) Une facture est générée pour chaque bon de commande.
- 8) La marchandise accompagnée de la facture est livrée au distributeur, qui réceptionne la marchandise, opère un premier cheking puis délivre un **Proof-of-delivery**.
- 9) Le client a 48h pour faire une réclamation. S'il en fait une, elle est transmise à l'équipe Logistique pour l'approuver ou la refuser. Si elle est acceptée il est procédé à des retours de marchandise.

En parallèle, un autre **processus** a lieu : celui concernant les remboursements suite aux remises et promotions effectuées par Unilever.

Il s'agit d'un cycle mensuel, chaque début de mois, les remises et promotions décidées en fin de mois précédent sont appliquées sur système. Elles consistent en un pourcentage prélevé sur le chiffre d'affaire brut (GSV) après atteinte de ses objectifs ou selon ce qu'il a réalisé en accord avec les SLA (Service Level Agreement) qui le lient à Unilever.

Au fil du mois, ces pourcentages finissent par constituer des provisions qui seront utilisées pour rembourser le client le mois prochain.

Donc lors de chaque mois  $m$ , le TTS Specialist réalise des provisions, c'est-dire des prélèvements sur la GSV, pour rembourser lors du mois  $m+1$ .

La figure suivante représente la modélisation du processus Order To Cash, créée avec le logiciel de modélisation « Camunda Modeler » à l'aide de la norme de modélisation BPMN. Pour plus de détails sur la norme BPMN, (Cf Annexe 1).

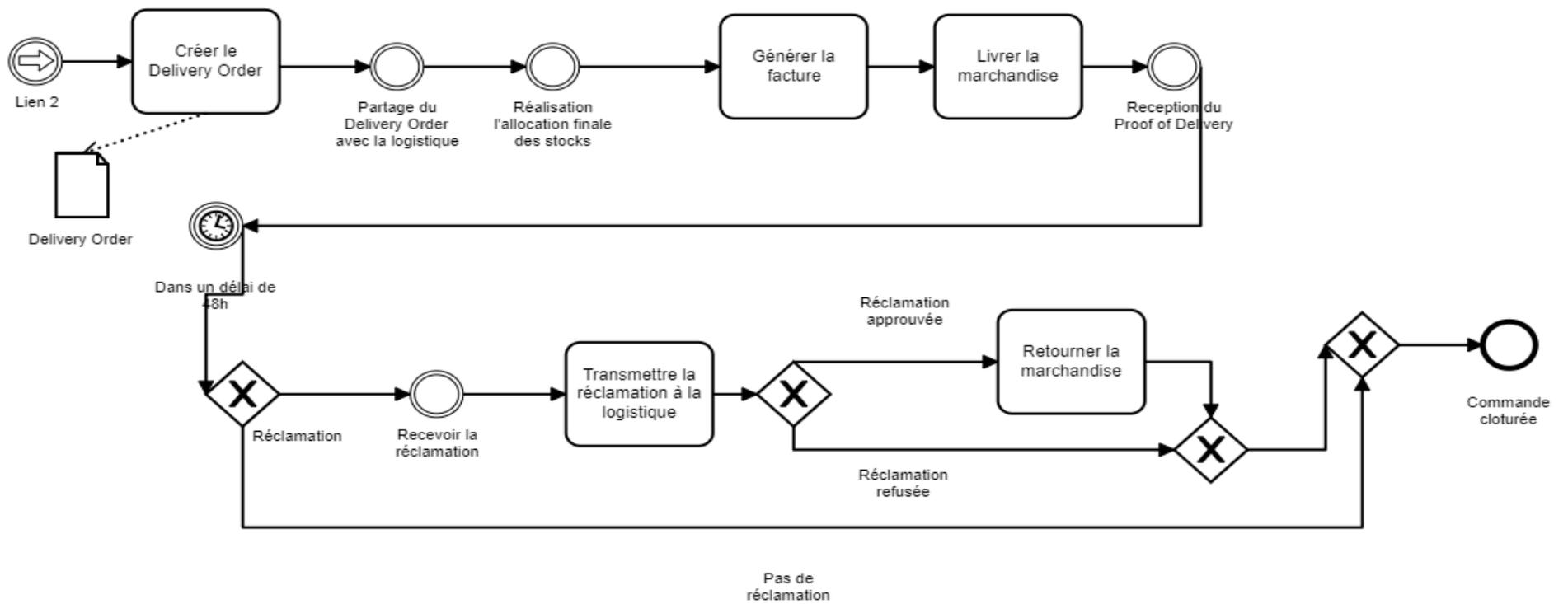


Figure 20 Modélisation du processus Order-to-cash réalisée par nous-mêmes sur Camunda Modeler - 1 (Nous-mêmes, 2020)

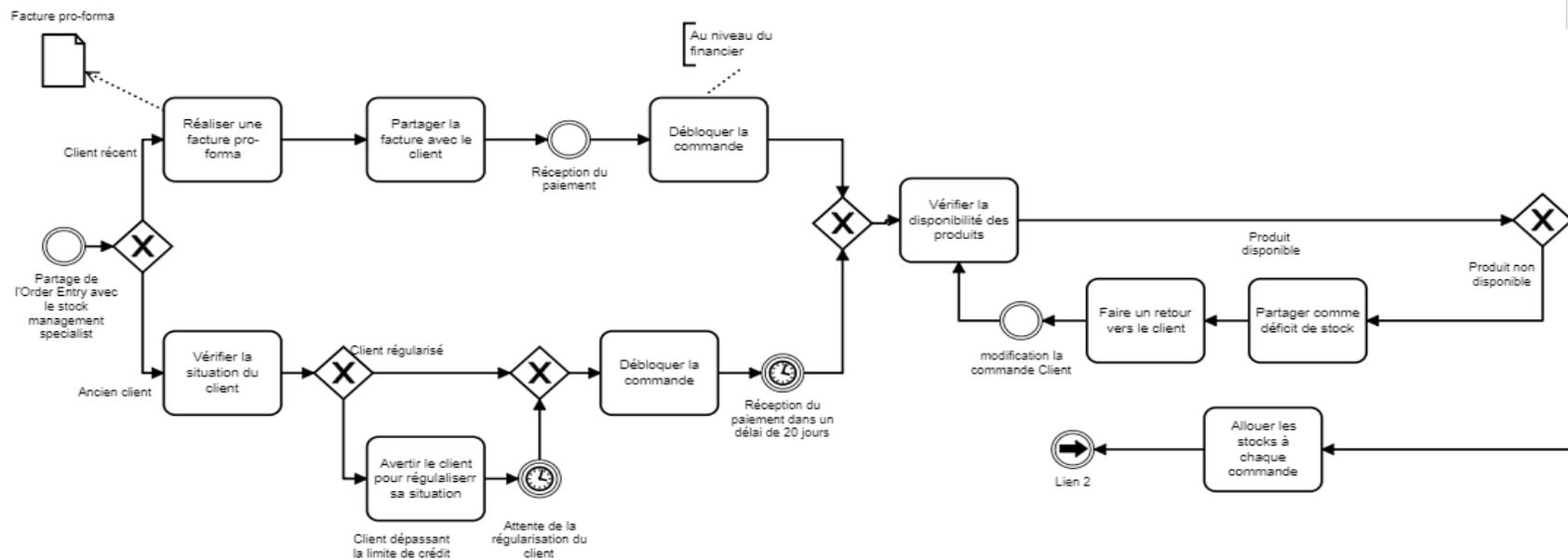


Figure 21 Modélisation du processus Order-to-cash réalisée par nous-mêmes sur Camunda Modeler - 2 (Nous-mêmes, 2020)

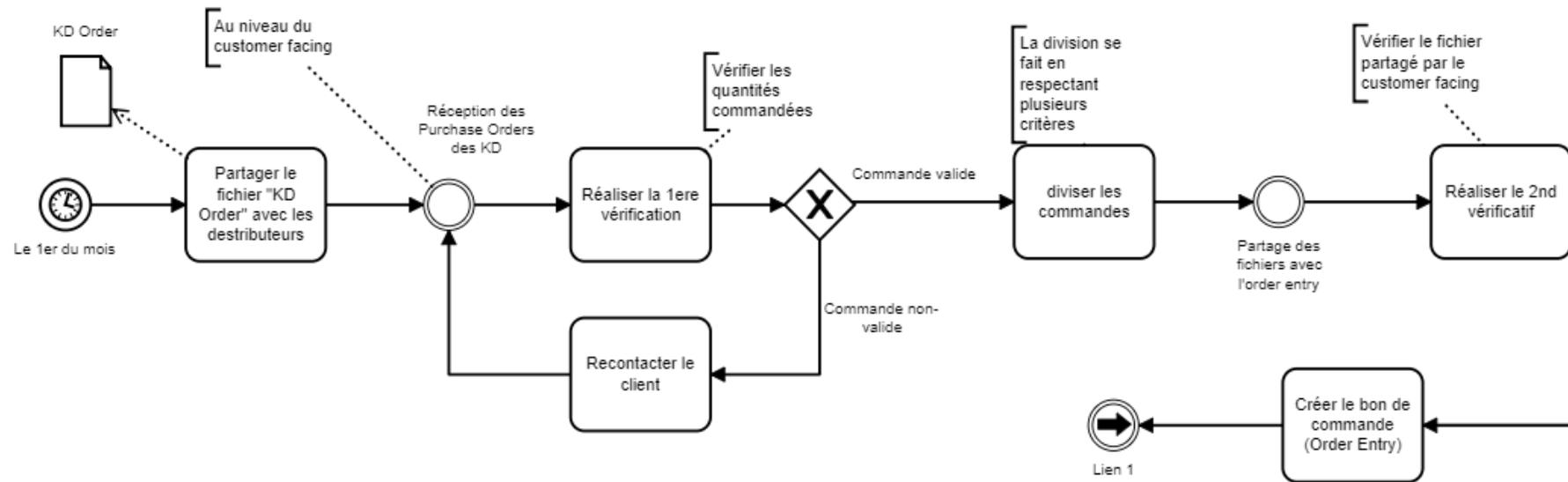


Figure 22 : Modélisation du processus Order-to-cash réalisée par nous-mêmes sur Camunda Modeler - 3 (Nous-mêmes, 2020)

## 4.2.Rapports

Dans le but de mieux suivre ses activités, le service produit une multitude de rapports. Lors de nos entretiens, nous en avons noté les principaux en nous concentrant sur leur contenu, leur méthode d'élaboration, leur fréquence et les personnes ciblées par ces rapports.

De ce fait nous avons classé ces derniers selon les thèmes qu'ils concernaient.

### 4.2.1. Rapports concernant les ventes

Le tableau ci-dessous présente les noms, les contenus, les techniques et fréquences d'élaboration ainsi que la cible des rapports concernant les ventes :

**Tableau 6 : Rapports concernant les ventes (Nous-même, 2020)**

Nom	Contenu	Elaboration	Fréquence	Cible
<b>Landing file</b>	Suivi des ventes et du turnover du point de vue business par rapport à l'objectif journalier et à l'objectif mensuel	Invoiced : les commandes facturées	Journalier	Unilever Algeria
		Les commandes ouvertes (c-à-d non encore facturées)		
<b>Daily sales tracker</b>	Suivi des ventes du point de vue logistique par rapport à l'objectif journalier et à l'objectif mensuel	Invoiced : les commandes facturées	Journalier	L'équipe logistique
		Les commandes ouvertes (c-à-d non encore facturées)		

### 4.2.2. Rapports résumant l'activité

Le tableau ci-dessous présente les noms, les contenus, les techniques et fréquences d'élaboration ainsi que la cible des rapports concernant l'activité du service

**Tableau 7 : Rapports résumant l'activité (Nous-même, 2020)**

Nom	Contenu	Elaboration	Fréquence	Cible
Weekly report	Résumé des opérations de la semaine : Création de commandes, Order Status, Delivery, Commandes ouvertes	Extraction de l'OMR ; une vue sur l'avancement de toutes les commandes, sur SAP	Hebdomadaire	Unilever Algeria
Monthly Report	Résumé des opérations du mois : Delivery, Finances, DR, TTS, Cash Collection	Flash report : partagé par les finances, après la clôture d'un mois, tout ce qui a été réalisé	Mensuel	Unilever Maghreb
		Extraction de l'OMR, une vue sur l'avancement de toutes les commandes, sur SAP (les commandes facturées)		
		follow-up fichier des Customer Facing		
		Extraction du DR catégorie et par brand		
		Cash Collection (fichier de la trésorerie)		

		Credit Note Report envoyé par TTS specialist		
--	--	--	--	--

#### 4.2.3. Rapport concernant les réclamations

Le tableau ci-dessous présente les noms, les contenus, les techniques et fréquences d'élaboration ainsi que la cible des rapports concernant les réclamations :

**Tableau 8 : Rapport concernant les réclamations (Nous-même, 2020)**

Nom	Contenu	Elaboration	Fréquence	Cible
Claims Tracker	Situation des réclamations	Réclamation envoyée via Outlook par le client uploadée sur <i>Sharepoint</i> <sup>1</sup> par le Customer Facing.	Avant chaque fin du mois	- Costumer service - Logistique - Quality - Client

#### 4.2.4. Rapport concernant la situation financière des clients

Le tableau ci-dessous présente les noms, les contenus, les techniques et fréquences d'élaboration ainsi que la cible des rapports concernant la situation financière des clients :

**Tableau 9 : Rapport concernant la situation financière des clients (Nous-même, 2020)**

Nom	Contenu	Elaboration	Fréquence	Cible
Over Due Report (facultatif)	Situation financière des clients	Extracation de Balance client sur SAP Extracation de Credit limit sur SAP	Hebdomadaire	- Equipe Finances - Equipe Commercial.
Credit Note Report	Rapport des notes de crédit ayant été remboursées	Extraction de la balance client sur SAP	Mensuel	- Customer Facing (régions+clients) - Finances - Commercial (CD)
Rebate ajustements	- Etat des provisions Remboursements à réaliser - Ajustements entre les deux montants	SAP : - Référence de chaque rebate - Ref de chaque provision Montant pour chacune SAS : - Ventes secondaires - Montant à rembourser pour chaque achievement	Mensuel	L'équipe finances
Accruals vs Actuals	Remboursements effectifs - Comparaison entre remboursements prévus et effectifs	Actuals à partir de : Credit Notes Accruals : Rebate ajustements Récapitulatif des deux rapports	Mensuel	Département finances Maghreb

<sup>1</sup> Il s'agit d'un logiciel permettant de créer un site web propre à l'entreprise afin de stocker, partager, organiser et accéder à l'information entre membres d'une même équipe.

#### 4.2.5. Rapport concernant les stocks

Le tableau ci-dessous présente les noms, les contenus, les techniques et fréquences d'élaboration ainsi que la cible des rapports concernant les stocks. Les données nécessaires à l'élaboration de ces rapports sont extraites sur SAP :

**Tableau 10 : Rapports concernant les stocks (Nous-même, 2020)**

Nom	Contenu	Elaboration	Fréquence	Cible
Opening Stock	Disponibilités du stock	Stock prêt à être commandé.	Quotidien (tous les matins)	Unilever Maghreb
		Stock total : Stock prêt à être commandé+ Stock physique		
		Produits cosmétiques disponibles en attente de contrôle qualité.		
		Produits détergents liquides disponibles en attente de contrôle qualité		
Pending Orders	Commandes débloquées par les finances en attente de chargement ou de modification	Commandes en attente de modification.	Deux fois par semaine	- Customer Channel Development  - Customer Service
		Commandes en attente de chargement.		
Out-of-stock	Justification pour les produits en rupture de stock	Produits en rupture de stock	Journalier (fin de journée)	Supply chain
Expiry-dates	Liste des produits avec leurs dates d'expiration	Références et noms produits avec date d'expiration.	Mensuel (début du mois)	Unilever Algeria
SLOBS	Liste des produits à faible rotation avec les dates d'expiration	Référence et noms des produits à faible rotation	Mensuel (début du mois)	Unilever Algeria
		Références et noms des produits avec date d'expiration.		

#### 4.3. Audit de la qualité des données

Dans cette partie nous allons auditer la qualité des données de l'entreprise Unilever, et ce, en nous basant sur plusieurs critères : (Zouaghi, 2019)

- 1) La correspondance avec la réalité : Les données sont-elles adaptées à la réalité et ont-elles trait aux éléments recherchés dans le cadre du travail ?
- 2) L'exhaustivité et la complétude : Les données nécessaires sont-elles toutes présentes ?
- 3) La validité et l'exactitude : Les données proviennent-elles de sources sûres ? Sont-elles fidèles à la réalité ?
- 4) L'actualité : Les données sont-elles disponibles au moment voulu ? Sont-elles à jour ?

- 5) L'accessibilité : Les données sont-elles d'un accès facile ? Sont-elles compréhensibles ?

Cet audit est le résultat des entretiens que nous avons conduit avec les collaborateurs du Customer Service :

**Tableau 11 : Rapports concernant les stocks (Nous-même, 2020)**

Critère de l'analyse	Etat des données
Correspondance avec la réalité	Les données correspondent globalement à la réalité, surtout en ce qui concerne les ventes, il y a parfois des décalages entre la production et la logistique mais ils sont mineurs.
Exhaustivité et complétude	Les données utilisées sont bien celles nécessaires et elles sont toutes présentes.
Validité et exactitude	Les données proviennent de sources sûres (SAP et SAS) et sont exactes.
Actualité	Les données sont globalement disponibles au moment voulu sauf en cas de problème de connexion, elles sont également la plupart du temps à jour, avec quelques légers décalages dus aux aléas de l'opérationnel comme des erreurs ou des oublis de la logistique.
Accessibilité	Les données sont accessibles et rapides à obtenir, elles ne sont pas toujours compréhensibles sous leur forme brute (fichiers présents sur SAP) mais le sont une fois traduites sous forme de rapports Excel.

S'assurer de la qualité des données est une étape importante voire primordiale, car c'est ce qui constitue la base de l'élaboration des rapports et le calcul des indicateurs de performance.

Avoir une donnée réelle, complète, valide, exacte, à jour et accessible, permet d'avoir des résultats et indicateurs plus corrects ainsi que d'élaborer des rapports fournissant davantage d'informations.

Les données utilisées pour l'élaboration des rapports du Customer Service sont généralement extraites directement à partir de SAP ou SAS, ou bien reçues à travers des fichiers élaborés par les autres services.

Pour les données extraites à partir de SAP cela se fait à travers l'utilisation de T-Codes (Transaction Codes), par exemple pour extraire les quotas ou les paiements par client pour l'Over Due Report.

Pour les fichiers partagés par les collaborateurs dans d'autres services, il est possible de citer le Flash Report partagé par les financiers et la Cash Collection qui provient de la trésorerie, qui sont utilisés pour établir le Monthly Report cité plus haut dans les rapports résumant l'activité.

#### 4.4.KPIs

Nous faisons ici l'inventaire des différents KPIs utilisés par les membres du service, qui ont pour principale mission de les orienter dans leur prise de décision afin qu'elle soit adéquate et faites au moment opportun.

Le tableau suivant résume tous les indicateurs repérés :

**Tableau 12 : Liste et caractéristiques des KPIs liés à l'activité du service (Nous-même, 2020)**

Nom	Sens	Type	Méthode de Calcul	Valeurs tolérées
DR: Dispatch Rate	Rapport entre quantités commandées et celles livrées	Global	$DR = \frac{\text{Quantités livrées}}{\text{Quantités commandées}}$	- 95% pour les produits importés - 98% pour les produits locaux.
CCFOT	Combien on a livré à temps au client par rapport à la quantité commandée	Global	$CCFOT = DR + \text{OnTime}$	98%, le même que ce soit pour les produits locaux ou importés
Delay Delivery: SLA	Combien de commandes ont pu être facturées avant 48h	Interne	$\frac{\text{Commandes facturées avant 48H}}{\text{Total des commandes}}$	Tourne autour de 60%.
Orders processin g: SLA	Combien de commandes ont dû être créées/modifiées de manière manuelle	interne	$\frac{\text{Commandes créées ou modifiées manuelle}}{\text{Total des commandes}}$	L'objectif est 0%.

Les données nécessaires au calcul de ces KPIs sont présentes dans le fichier Order Management Report (sur SAP), puis le calcul est effectué par un logiciel utilisé par le service, qui permet également d'obtenir des résultats détaillés par produit pour chaque KPI.

Lorsque les valeurs de ces KPIs se retrouvent en dehors de la fourchette tolérée, cela veut dire qu'il y a problème. Il devient alors nécessaire de justifier les pertes dans leurs valeurs, dans ce cas, il existe une série de justifications à prendre en considération afin de régler le problème ayant causé les pertes.

Ces KPIs permettent de renseigner sur l'état de l'activité et son avancement en vue d'un pilotage efficace. Malgré leur efficacité, ils ne suffisent pas à couvrir tout l'activité du service, les employés ressentant le besoin de mieux suivre de leurs activités, notamment pour le suivi des réapprovisionnements des stocks des clients du circuit secondaire, à travers l'élaboration éventuelle de tableaux de bords.

#### **4.5.Énoncé de la problématique**

Après avoir identifié le processus géré par le service, les différents rapports, données et KPIs utilisés, nous avons pu observer deux principaux dysfonctionnements. Le premier concerne un manque de tableau de bords pour le suivi de Replenishment des clients du circuit secondaire. Le second concerne la difficulté lors de la réalisation des rapports.

En effet, nous avons remarqué une absence de KPIs pour le suivi des réapprovisionnements des clients pour les ventes secondaires, ces KPIs seraient très utiles car une connaissance de la baisse des stocks chez le secondaire aiderait à mieux identifier le besoin chez les distributeurs et permettrait donc de le satisfaire.

De plus, les rapports sont créés sur Excel, après exécution de plusieurs modifications sur des données extraites à partir de SAP ; ce qui rend la tâche relativement longue et laborieuse, empêchant ainsi une bonne utilisation du temps, en plus de plusieurs autres défauts qui peuvent être reprochés à Excel et qui sont :

- L'aspect esthétique des rapports obtenus via Excel n'est pas satisfaisant, ce qui rend leur compréhension plus difficile.
- Le filtrage des données offert est limité.
- Les analyses offertes sont limitées, les requêtes qu'il est possible de faire sont pauvres et n'atteignent pas le niveau de performance souhaité.

De nombreux autres logiciels permettent de simplifier la tâche du reporting tout en proposant une multitude d'outils d'analyse, tels que les outils de la Business Intelligence, qu'il est possible d'utiliser une fois un modèle dimensionnel réalisé et implémenté.

Notons que, par manque d'accès aux données, de par les limites de la durée de réalisation de notre travail, et la priorité qu'accorde le service au second dysfonctionnement, nous nous sommes concentrées sur ce dernier pour l'élaboration de notre solution.

Nous nous sommes donc posé la question suivante :

**« Comment réduire le temps de création de rapports, tout en assurant un meilleur visuel et en offrant plus de possibilités d'analyse aux utilisateurs finaux ? »**

#### **4.6.Méthodologie**

Pour aborder notre problème, nous tenterons de réaliser des datamarts spécifiques aux processus concernant le service que nous implémenterons à travers un outil BI.

Nous avons suivi les étapes suivantes :

- Identification et analyse de besoins des utilisateurs à travers la conduite d'entretiens.
- Modélisation dimensionnelle avec ses quatre étapes, résultant dans la création de schémas multidimensionnels.
- Création de tables SQL pour le stockage des tables de faits et des dimensions.
- Développement des différentes étapes suivies lors de l'ETL pour le chargement des datamarts.
- Création des mesures et colonnes dérivées, ainsi que le développement des analyses possibles.
- Elaboration d'un exemple de rapport.

## **5. Conclusion**

Ce chapitre nous a permis de mieux connaître l'entreprise et le service où nous avons effectué notre stage, ainsi que le secteur d'activité de l'entreprise.

Il nous a également permis de cerner des dysfonctionnements. Ces dysfonctionnements ont été formulés sous forme d'une problématique dont la solution sera la conception et l'implémentation d'une solution BI, chose qui sera explicitée dans les prochains chapitres.

## **Chapitre III : Conception de la solution BI**

## 1. Introduction

Après avoir mené notre étude de l'existant, nous allons procéder à la conception de la solution. En premier lieu, nous allons justifier le choix de la méthodologie adoptée, puis nous allons dérouler ses différentes phases :

- Le recueil des besoins des utilisateurs.
- La modélisation dimensionnelle avec ses (4) étapes :
  - 1) L'identification des sous-processus à modéliser.
  - 2) Le choix du grain adapté à chaque sous-processus.
  - 3) La détermination des dimensions.
  - 4) La détermination des mesures.

Une fois la modélisation dimensionnelle réalisée, nous allons faire le choix de l'architecture de modélisation à adopter, puis réaliser la matrice en bus afin de déterminer les datamarts à mettre en œuvre et les dimensions conformes. Par la suite, nous allons choisir le type des schémas multidimensionnels et en réaliser un pour chaque datamart.

Une fois les schémas réalisés, nous allons procéder à la création des tables SQL afin de les stocker.

## 2. Choix de la méthodologie adoptée : (YESSAD, et al. , 2016)

Nous avons fait une comparaison entre trois méthodologies afin de choisir la plus adaptée à notre projet. Le tableau suivant montre la comparaison entre les trois méthodologies :

**Tableau 13 : comparaison entre les trois méthodologies du projet BI**

Caractéristiques/ Démarches	Inmon	Kimball	Datavault
Acteurs	Informaticiens	Utilisateurs finaux	Utilisateurs finaux
Complexité	Assez complexe	Simple	Simple
Méthodologie	En spirale	Quatre étapes	Agile
Coûts	Coûté élevé au départ	Coût bas	Coût le plus bas
Temps de développement	Long	Court	Court
Performance de requête	Les requêtes sont peu performantes	Les requêtes sont très performantes	Les requêtes sont peu performantes

Suite à cette comparaison, nous avons choisi d'exclure Inmon du choix car la méthode est coûteuse et demande une longue période et des compétences en informatique pour la réalisation.

En comparant Kimball et le Datavault, on peut voir que quoiqu'elles soient toutes les deux simples et à bas coût, Kimball permet d'avoir une meilleure performance. De plus, le datavault requiert une approche agile, chose assez difficile à réaliser pour un premier projet en télétravail.

Nous avons donc choisi Kimball pour la conception de notre solution.

### 3. Identification des besoins

Pour identifier les besoins nous nous sommes basées sur un mix des deux méthodes précédemment citées ; à savoir, la méthode orientée utilisateur et la méthode orientée source.

- **Méthode orientée utilisateur :**

Pour la méthode orientée utilisateur, nous avons organisé des entretiens avec le personnel en connaissance du problème objet de notre travail, dans le service, à savoir :

- Manager ;
- Process control & Claims Specialist ;
- TTS Specialist ;
- Stock management Specialist ;

Lors de ces entretiens, nous leur avons posé des questions sur les points suivants :

- Les besoins en termes d'analyse et de suivi ;
- Les axes par rapport auxquels les analyses sont effectuées ;
- Le niveau de détail des analyses effectuées.

Le questionnaire type utilisé lors des entretiens est présenté en Annexe 2.

- **Méthode orientée Source :**

Pour la méthode orientée source, nous avons récolté plusieurs fichiers utilisés au cours du processus ainsi que des rapports produits par les utilisateurs, documents que nous résumons dans le tableau ci-après :

**Tableau 14: Description des fichiers récoltés en utilisant la méthode orientée Source (Nous-mêmes, 2020)**

<b>Fichier</b>	<b>Description</b>
<b>Order file</b>	Fichier donnant le stock disponible pour chaque produit, sert de PO au client.
<b>Facture</b>	Echantillon d'une facture, elle accompagne les produits d'un SO.
<b>Claim sheet</b>	Echantillon d'une réclamation.
<b>Claims tracker</b>	Fichier montrant le statut et les différents détails des réclamations reçues.
<b>Delivery order</b>	Bon de livraison du SO.
<b>Daily sales tracker</b>	Fichier montrant l'avancement des ventes du point de vue logistique.
<b>Orders and Shipment</b>	Fichier montrant le suivi des commandes.
<b>Landing file</b>	Fichier détaillant le suivi des ventes en termes de tonnage et de GSV, quel que soit leur état : bloquées ou débloquées, confirmées, chargées, facturées.
<b>Weekly report</b>	Rapport résumant les opérations de la semaine.
<b>Monthly report</b>	Rapport résumant les opérations du mois.

<b>Rebate Adjustments</b>	Fichier montrant l'ajustement des provisions sur système aux montants dus aux clients.
<b>Credit Note</b>	Fichier montrant les remboursements effectifs relatifs aux différentes remises.
<b>Provisions followup</b>	Fichier comparant entre les provisions et ce qui a été effectivement été remboursé aux clients.
<b>Opening Stock</b>	Fichier donnant les stocks des différents produits selon leur statut : Disponible, en attente de contrôle qualité ou en attente de réception.
<b>Pending Orders</b>	Fichier montrant les commandes en attente de confirmation.
<b>Losses</b>	Fichier montrant les pertes pour out-of-stock ou dépassement de quota.
<b>DPM Forecast</b>	Fichier détaillant les prévisions de ventes des 3 prochaines années.

Une fois ces fichiers regroupés, nous les avons étudiés et analysés, pour en tirer les informations nécessaires au déroulement de la méthodologie.

### 3.1. Classification des besoins en thèmes

Après identification des besoins grâce aux deux méthodes, nous les avons regroupés en thèmes, et extrait les différents axes autour desquels s'articule chaque thème pour chaque utilisateur, et ce, en se basant sur les besoins de suivi les plus récurrents du service. Pour finalement obtenir le tableau suivant :

**Tableau 15: Tableau illustrant la classification des besoins en thèmes (Nous-mêmes, 2020)**

<b>Thème</b>	<b>Besoin de suivi</b>	<b>Personnes concernées</b>
<b>Suivi des stocks</b>	Besoin de suivre la quantité en stock en caisses, en palettes et en tonnes :  1) Le produit : Code, catégorie, sous-catégorie, marque, SKU. 2) Le statut	- Stock management Specialist
<b>Suivi des réclamations</b>	Besoin de suivre les montants des réclamations HT, les quantités réclamées ainsi que le nombre de réconciliations, selon :  1) Le type de réclamation. 2) Le statut de la réclamation. 3) Les clients concernés : Code, Nom, région et Wilaya. 4) Le produit : Code, catégorie, sous-catégorie, marque, SKU. 5) L'unité de réclamation. 6) L'unité de prix de la réclamation. 7) La date : Année, Mois, Semaine, Jour. 8) La facture concernée.	- Manager ; - Process control & Claims Specialist ;

<b>Suivi des provisions TTS</b>	Besoin de suivre les montants des provisions avant et après ajustement en DZD et en Keuros selon : <ol style="list-style-type: none"> <li>1) Les clients concernés : Code, Nom, région et Wilaya.</li> <li>2) Le produit : Code, catégorie, sous-catégorie, marque, SKU.</li> <li>3) Le TTS.</li> <li>4) La remise : la famille de provisions à laquelle appartient la provision.</li> <li>5) InternalOrder : Le compte bancaire d'où le remboursement est effectué.</li> <li>6) La date : Année, Mois, Semaine, Jour.</li> </ol>	<ul style="list-style-type: none"> <li>- TTS Specialist</li> </ul>
<b>Suivi des remboursements TTS</b>	Besoin de suivre les remboursements effectués selon : <ol style="list-style-type: none"> <li>1) La note de crédit.</li> <li>2) Les clients concernés : Code, Nom, région et Wilaya.</li> <li>3) Le TTS.</li> <li>4) La date : Année, Mois, Semaine, Jour.</li> </ol>	<ul style="list-style-type: none"> <li>- TTS Specialist</li> </ul>
<b>Suivi de la GSV</b>	Besoin de connaître la GSV TTC en DZD et en euros selon : <ol style="list-style-type: none"> <li>1) Le produit : Code, catégorie, sous-catégorie, marque, SKU.</li> <li>2) Le client : Code, Nom, Région, Wilaya.</li> <li>3) Le statut dans le Customer Service.</li> <li>4) Le statut financier.</li> <li>5) La date : Année, Mois, Semaine, Jour.</li> </ol>	<ul style="list-style-type: none"> <li>- Manager ;</li> <li>- Process control &amp; Claims Specialist ;</li> </ul>

<b>Suivi des ventes primaires</b>	<p>Suivre les quantités vendues au primaire en caisses et en tonnes, selon :</p> <ol style="list-style-type: none"> <li>1) Le produit : Code, catégorie, sous-catégorie, marque, SKU.</li> <li>2) Le client : Code, Nom, Région, Wilaya.</li> <li>3) Le statut dans le Customer Service.</li> <li>4) Le statut financier.</li> <li>5) La date : Année, Mois, Semaine, Jour</li> </ol>	<ul style="list-style-type: none"> <li>- Manager ;</li> <li>- Process control &amp; Claims Specialist ;</li> </ul>
<b>Suivi des prévisions (Forecast)</b>	<p>Besoins de connaître les quantités à vendre prévues selon :</p> <ol style="list-style-type: none"> <li>1) Le produit : Code, catégorie, sous-catégorie, marque, SKU.</li> <li>2) La date : Année, Mois, Semaine, Jour.</li> </ol>	<ul style="list-style-type: none"> <li>- Manager ;</li> <li>- Process control &amp; Claims Specialist.</li> <li>- Stock Management Specialist.</li> </ul>
<b>Suivi des ventes secondaires</b>	<p>Besoin de connaître les quantités vendues au circuit secondaire selon :</p> <ol style="list-style-type: none"> <li>1) Le produit : Code, catégorie, sous-catégorie, marque, SKU.</li> <li>2) Le client : Code, Nom, Région, Wilaya.</li> <li>3) La date : Année, Mois, Semaine, Jour.</li> </ol>	<ol style="list-style-type: none"> <li>3) Manager ;</li> <li>4) Process control &amp; Claims Specialist .</li> </ol>

#### 4. Modélisation dimensionnelle

##### 4.1. Etape 1 : Choix des sous-processus

Dans cette étape nous avons fait le choix des sous-processus à modéliser.

##### 4.1.1. Découpage du processus Order-To-Cash

Nous avons précédemment retracé le processus Order-to-cash et le processus relatif au TTS dans l'étude de l'existant, mais il s'est avéré nécessaire de les découper davantage pour les modéliser, et ce car la modélisation de tout le processus aurait eu pour résultat un modèle difficile à mettre en œuvre, incompréhensible et inutilisable.

Avant d'opérer ce découpage, nous avons réalisé quelques entretiens complémentaires avec les membres précédemment cités dans l'identification des besoins, et ce afin d'avoir une connaissance plus détaillée du processus.

Le processus Order-to-cash a été divisé en les sous-processus suivants : Gestion des stocks, Commande et Gestion des réclamations.

Le processus relatif au TTS a été divisé en deux sous-processus : Un pour les provisions et un pour les remboursements.

- **Gestion des stocks :**

Les stocks sont suivis quotidiennement par le Stock Management Specialist afin d'allouer les stocks aux commandes et d'avoir une connaissance des quantités disponibles quel que soit leur statut : Unrestricted (Stock physique total), disponible, sous-inspection qualité, en attente de réception, Confirmed Sales order, Planned stock replenishment.

- **Gestion des réclamations :**

Lorsque le client est livré il a 48 heures pour faire une réclamation. Dans ce cas, l'équipe logistique l'approuve ou la refuse. En cas d'acceptation le Customer Service la prend en charge et procède à des retours de marchandise.

- **Gestion des provisions des TTS :**

Au cours du mois des remises et promotions sont appliquées aux différents achats des clients. Suite à l'application de ces remises, des provisions sont tirées de la GSV afin de rembourser plus tard les clients.

A la fin du mois ces provisions sont comparées avec ce qui est réellement dû aux clients et les taux des remises sont ajustés pour se rapprocher des montants à rembourser.

- **Gestion des remboursements des TTS :**

Une fois le mois écoulé les clients sont remboursés avec ce qui a été provisionné au cours du mois. Cela se fait à travers la réalisation de notes de crédit.

- **Gestion de la commande :**

Lorsqu'un client passe une commande, un Purchase Order est créé puis divisé en plusieurs Sales Order. Chaque Sales order concerne un certain nombre de produits en une certaine quantité pour chacun. Chaque ligne du sales order est donc traitée individuellement.

Avant de passer à l'allocation des quantités commandées sur système, le statut du crédit de la ligne de la commande doit être débloqué, c.à.d. la limite de crédit du client non dépassée ou bien le montant prépayé, si ce n'est pas le cas, la commande est considérée comme étant **Under Finance**.

Une fois le statut du crédit de la commande débloqué, une certaine quantité de produit est allouée selon la disponibilité du produit. Cette quantité est appelée quantité confirmée et peut-être différente de la quantité commandée. Lors de cette opération la ligne de la commande est considérée sous le statut **Under Treatment**.

Son statut change après l'élaboration du Delivery Order, spécifiant les quantités des différentes lignes du Sales Order à picker pour remplir les camions. A ce stade-là, le statut des différentes lignes de la commande devient **Waiting for Shipment**.

Après cela, la facture est générée par rapport au Sales Order et son contenu livré, les quantités des différents produits sont appelées quantités facturées, le statut des différentes lignes du SO passe à **Invoiced**.

- **Forecast**

Les prévisions des trois prochaines années sont élaborées durant l'année en cours et sont corrigées au fur et à mesure en fonction des ventes réalisées, elles concernent chaque SKU

- **Ventes secondaires**

Une fois les commandes livrées aux KDs, ces derniers se chargent d'écouler leur contenu au circuit secondaire.

#### 4.1.2. Hiérarchisation des sous-processus

Une fois les sous-processus définis, ils ont été jugés par rapport à leur faisabilité et à leur importance par rapport aux utilisateurs.

Puis ces mêmes sous-processus ont été classés dans un schéma selon ces deux critères, les abscisses représentant la faisabilité et les ordonnées l'importance.

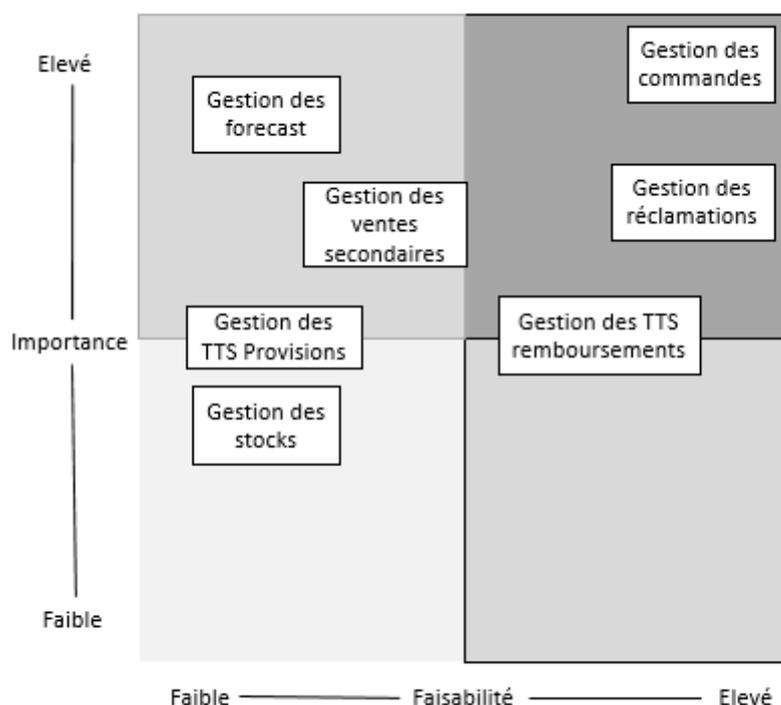
- La faisabilité a été jugée selon la disponibilité et la facilité d'accès aux données.
- L'importance a été jugée par rapport à la place qu'occupe le sous-processus dans le service en matière de reporting.

Le tableau et la figure suivants évaluent la faisabilité et l'importance de chaque sous processus :

**Tableau 16 : Faisabilité et Importance des sous-processus (Nous-mêmes, 2020)**

Processus	Faisabilité	Importance
Gestion des stocks	Il a une faible faisabilité car il nécessite des données fréquemment actualisées sur tous les produits disponibles, ce qui rend l'ETL difficile.	Il a une importance moyenne, il permet d'avoir une visibilité sur les disponibilités en stocks.
Gestion des réclamations	Il a une bonne faisabilité car les données sont disponibles, mais elles sont transmises via mail, elles ne sont donc pas automatiquement relevées sur système.	Il a une très grande importance, il permet de suivre l'avancement du statut des réclamations et le suivi se fait par la plupart du personnel du service.
Gestion des provisions TTS	Il a une faible faisabilité car elle nécessite l'accès au système d'information des ventes secondaires, SAS.	Il a une grande importance car les provisions font partie du calcul du chiffre d'affaires.
Gestion des remboursements TTS	Il a une bonne faisabilité car les données sont regroupées mensuellement dans le Credit Note Report.	Il a une grande importance car les provisions sont déduites du chiffre d'affaires.
Commande	Il a une grande faisabilité car les données sont disponibles et actualisées en temps réel. (SAP).	Il a une très grande importance car c'est une très grande contribution au chiffre d'affaires.
Forecast	Il a une faible faisabilité car les données sont confidentielles.	Il a une très grande importance car il permet de comparer les résultats réalisés à ceux prévus et donne une vue sur les objectifs de l'entreprise.

Ventes secondaires	Il a une faisabilité modérée car il faut passer par le SAS pour obtenir les données.	Il a une grande importance car il permet de savoir de quelle manière s'écoule le stock des KDs et de connaître le besoin réel des consommateurs finaux.
--------------------	--	---



**Figure 23: Hiérarchisation des sous-processus (Nous-mêmes, 2020)**

Vu le temps limité dont nous disposons, nous avons accordé la priorité aux processus « Commande », « Réclamations » et « Remboursements TTS » lors de l'implémentation de leurs modèles dimensionnels car affichant un ordre d'importance et de faisabilité le plus élevé.

#### 4.1.3. Affectation des besoins aux sous-processus

Enfin, nous avons affecté les besoins identifiés plus haut aux sous-processus comme montré dans le tableau ci-dessous, afin de nous en inspirer par la suite pour la détermination du grain, des dimensions et des mesures.

**Tableau 17: Affectation des thèmes de suivi aux sous-processus (Nous-mêmes, 2020)**

Sous-processus	Thèmes affectés
Gestion des stocks	Suivi des stocks
Gestion des réclamations	Suivi des réclamations
Gestion des provisions TTS	Suivi des provisions TTS
Gestion des remboursements TTS	Suivi des remboursements TTS

<b>Commande</b>	Suivi de la GSV
	Suivi des ventes
<b>Forecast</b>	Suivi des prévisions
<b>Ventes secondaires</b>	Suivi des ventes secondaires

#### 4.2. Etape 2 : Choix du grain

Dans cette étape nous avons défini le grain par rapport aux sous-processus précédemment déterminés comme le montre le tableau suivant, ce choix est décisif car le niveau de détail atteint dans les analyses en dépendra.

Il est à rappeler que si l'on se trompe dans ce choix il sera nécessaire de revenir à cette étape pour changer le grain.

**Tableau 18: Détermination des grains des sous-processus (Nous-mêmes, 2020)**

<b>Processus</b>	<b>Grain</b>
<b>Gestion des stocks</b>	Chaque ligne de la table de fait représente les quantités en caisses en stock d'un produit, selon les 6 différents statuts et en une date donnée.
<b>Les réclamations</b>	Chaque ligne de la table de fait représente le montant à rembourser d'une réclamation d'un certain type, au sujet d'un seul produit, faite par un seul client, en une date donnée accompagnée ou non d'une facture.
<b>Processus des provisions TTS</b>	Chaque ligne de la table de fait représente une provision réalisée par rapport à un produit sous une remise donnée, pour un client donné, à partir d'un certain compte bancaire en une date donnée.
<b>Processus remboursements TTS</b>	Chaque ligne de la table de fait représente une note de crédit au sujet d'un certain TTS, concernant un seul client en une date donnée.
<b>Processus Commande</b>	Chaque ligne de la table de fait représente une ligne d'un Sales Order, c.a.d une certaine quantité commandée en caisse, une certaine quantité allouée sur système, une quantité facturée pour un produit à un client particulier en une date particulière.
<b>Forecast</b>	Une ligne de la table de faits représente la quantité prévue d'un SKU en une semaine ou en un mois particulier.
<b>Ventes secondaires</b>	Chaque ligne de la table de faits représente une vente réalisée d'un certain produit d'une certaine quantité en caisses à un client secondaire donné et en une date donnée.

#### 4.3. Etape 3 : Choix des dimensions

Dans cette étape nous avons identifié les dimensions relatives au grain précédemment défini que nous avons résumées dans le tableau ci-dessous, ces dimensions permettent de donner un contexte descriptif aux faits que l'on souhaite suivre.

Les attributs de dimension, leurs significations et leurs types sont présentés en Annexe 3.

**Tableau 19: Choix des dimensions des sous-processus (Nous-mêmes, 2020)**

<b>Processus</b>	<b>Dimensions</b>
Gestion des stocks	Produit
	Statut_Stock
	Date
Gestion des réclamations	Client (KD)
	Date
	Produit
	N° de la réclamation (DD)
	Type de la réclamation
	Statut de la réclamation
	N° de la facture (DD)
Processus des provisions TTS	N° de la remise (DD)
	Compte bancaire(DD)
	Client (KD)
	Produit
	TTS
	Date
Processus remboursements TTS	N° de la note de crédit(DD)
	Client (KD)
	Date
	TTS
Commande	Produit
	Client (KD)
	Date
	Statut_Customer_Service
	CCFOT_Reason
	Statut_Financier
	Entrepôt
	N° Sales Order (DD)
Forecast	Produit
	Date
Ventes secondaires	Produit
	Date
	Client_Secondaire

#### **4.4.Etape 4 : Choix des mesures**

Dans cette étape nous avons déterminé les mesures nécessaires à la réalisation des analyses, ces mesures sont de deux natures :

- **De base** : Venant directement de mesures opérationnelles.
- **Dérivées** : Calculées grâce à d'autres mesures, celles de base par exemple.

Ces mêmes mesures peuvent être additives, semi-additives ou non-additives selon qu'elles puissent être agrégées ou non par rapport aux dimensions.

Les tableaux suivants représentent les mesures de base et les mesures additives de chaque sous processus :

**Tableau 20: Mesures de base des sous-processus (Nous-mêmes, 2020)**

<b>processus</b>	<b>Mesures de base</b>	<b>Signification</b>	<b>Type de la mesure</b>
<b>Gestion des stocks</b>	Unrestricted_Stock_Qty_Caisse	Quantité en caisse totale en stock quel que soit le statut	Semi-additive
	Stock_Disponible_Qty_Caisse	Quantité en stock disponible en caisse	Semi-additive
	Stock_Inspection_Qualité_Qty_Caisse	Quantité en stock sous inspection qualité en caisse	Semi-additive
	Stock_Attente_Reception_Qty_Caisse	Quantité en stock en attente de réception en caisse	Semi-additive
	Stock_Confirmed_Sales_Order_Qty_Caisse	Quantité en stock déjà vendue en caisse	Semi-additive
	Stock_Planned_Replenishment_Qty_Caisse	Quantité approvisionnée en caisse	Semi-additive
<b>Les réclamations</b>	Quantité_réclamée	Quantité réclamée	Semi-Additive
	Unité_Réclamation	Unité dans laquelle la réclamation a été faite	Non-additive
	Unité_Prix_réclamation	Unité dans laquelle le prix de la réclamation a été enregistré	Non-additive
<b>Processus des provisions TTS</b>	Montant_d'achat	Montant sur lequel la remise est appliquée	Additive
	% du TTS	% de la remise appliquée	Non additive
<b>Processus remboursements TTS</b>	Montant de la note de crédit HT	Montant du remboursement sans taxe	Additive
	Montant de la note de crédit TTC	Montant du remboursement toute taxe comprise	Additive
<b>Processus Commande</b>	Sales_Type	Type de commande	Non-additive
	N°Facture	N° de la facture accompagnant la commande	Non-additive
	N° Purchase Order	N° du PO	Non-additive
	N° Proforma	N° de la proforma accompagnant la commande	Non-additive
	N° Delivery_Order	N° du delivery_order accompagnant la commande	Non-additive
	Order_Qty	Quantité commandée	Additive
	Confirmed_Qty	Quantité confirmée	Additive
	Billed_Qty	Quantité facturée	Additive
	Credit_Status	Statut du crédit de ligne du Sales Order	Non-additive
	Delivery_Block	Livraison de la commande bloquée ou non	Non-additive
Forecast	Quantité_prévue	Quantité prévue	Additive
Ventes secondaires	Quantité_vendue	Quantité vendue	Additive

**Tableau 21: Mesures dérivées des sous-processus (Nous-mêmes, 2020)**

Processus	Mesures dérivées	Signification	Type de la mesure
<b>Gestion des stocks</b>	Unrestricted_Stock_Qty_Tonne	Quantité en tonnes totale en stock quel que soit le statut	Semi-additive
	Stock_Disponible_Qty_Tonne	Quantité en stock disponible en tonnes	Semi-additive
	Stock_Inspection_Qualité_Qty_Tonne	Quantité en stock sous inspection qualité en tonnes	Semi-additive
	Stock_Attente_Reception_Qty_Tonne	Quantité en stock en attente de réception en tonnes	Semi-additive
	Stock_Confirmed_Sales_Order_Qty_Tonne	Quantité en stock déjà vendue en tonnes	Semi-additive
	Stock_Planned_Replenishment_Qty_Tonne	Quantité approvisionnée en tonne	Semi-additive
<b>Les réclamations</b>	Montant_Réclamé_DZD	Montant réclamé en dinars algériens	Additive
	Somme_Montant_Réclamation	Somme totale des montants d'un ensemble de réclamations	Additive
	Nombre Réclamation	Nombre de réclamations reçues	Additive
<b>Processus des Provisions TTS</b>	Montant_provisionné	Montant provisionné	Additive
<b>Processus Commande</b>	Nombre_Commande	Nombre de commandes reçues	Additive
	Montant_Facturé_HT	Montant facturé HT	Additive
	Montant_Confirmé_HT	Montant confirmé HT	Additive
	Montant_Commandé_HT	Montant commandé HT	Additive
	Tonnage facturé (Billed Ton)	Tonnage facturé	Additive
	Tonnage_Confirmé	Tonnage confirmé	Additive
	Tonnage_commandé	Tonnage commandé	Additive
	Working_days	nombre de jours ouvrables en un mois	Semi-Additive
	Time_lapsed	Jours passés du mois	Non-Additive
<b>Remboursements TTS</b>	MTD	Quantité facturée depuis le début du mois jusqu'à ce jour	Additive
	Nombre_Note_Credit	Nombre de notes de crédit	Additive
	Somme_Montant_Note_Credit_HT	Somme des montants d'un ensemble de notes de crédit en HT	Additive
	Somme_Montant_Note_Credit_TTC	Somme des montants d'un ensemble de notes de crédit TTC	Additive

## 5. Choix de l'architecture adoptée : (KimballGroup, SD)

Nous avons choisi l'architecture en bus de magasin de données représentée dans la figure ci-dessous car sa structure permet de décomposer la création du data warehouse en plusieurs parties, plus faciles à gérer, prenant en considération à la fois les business process et les dimensions conformes associées.

Les dimensions conformes sont des dimensions communes, standardisées et créées en une fois lors du processus ETL, et après cela réutilisées par les différentes tables de faits.

Ces dimensions conformes permettent l'intégration des données à travers les différents business process en regroupant des attributs communs. Elles permettent également de réduire le temps de réalisation du modèle dimensionnel en réduisant les redondances, chaque dimension n'étant générée qu'une fois.

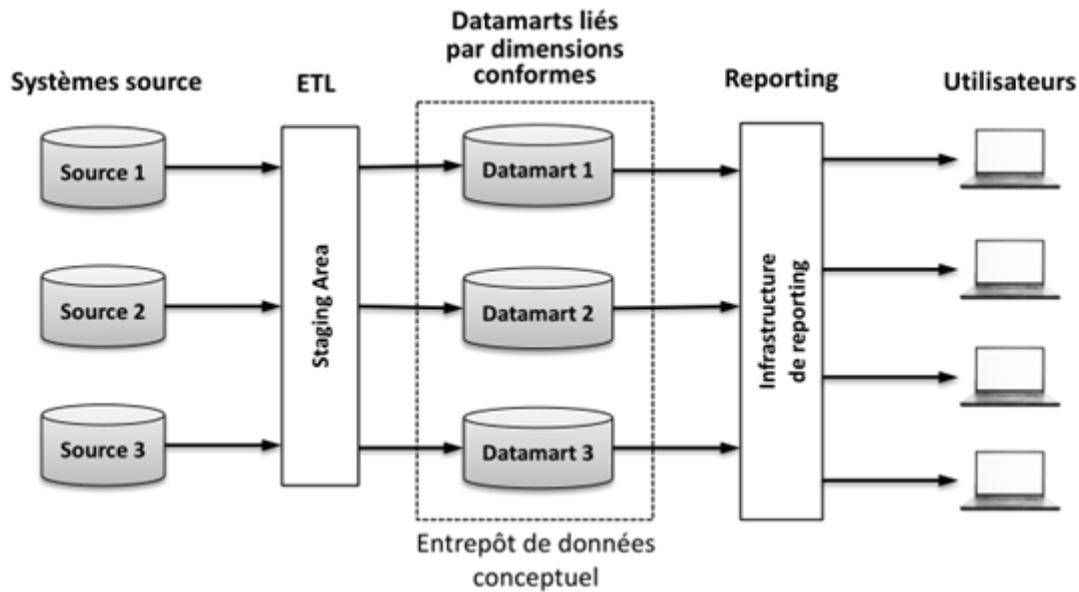


Figure 24: Architecture en bus de magasin de données (Chafki, et al, 2011)

### 5.1. La matrice de bus dimensionnel

Afin d'identifier les dimensions conformes et les datamarts réalisés, nous avons réalisé la matrice en bus dimensionnel, où :

- Les lignes représentent les sous-processus que nous avons définis.
- Les colonnes représentent les différentes dimensions.

Lorsqu'une certaine dimension est associée à un certain sous-processus, leur intersection est cochée.

Le tableau suivant est une illustration de la matrice de bus multidimensionnel du processus Order To Cash.

**Tableau 22: Matrice de bus multidimensionnel (Nous-mêmes, 2020)**

	Client	Date	Produit	Client_ Secondaire	Statut_ Stock	Type de la réclamation	Statut de la réclamation	Statut_Customer _Service	CCFOT_Reason	Statut_Financier	Entrepôt	TTS
Gestion des stocks		×	×		×							
Gestion des réclamations	×	×	×			×	×					
Gestion des provisions TTS	×	×	×									×
Gestion des remboursements des TTS	×	×										×
Gestion de commande	×	×	×					×	×	×	×	
Forecast		×	×									
Ventes secondaires		×	×	×								

Nous avons donc abouti à **7 datamarts** :

- le datamart gestion des stocks.
- le datamart gestion des réclamations.
- le datamart gestion des provisions TTS.
- le datamart gestion des remboursements des TTS.
- le datamart gestion des commandes
- le datamart gestion des forecasts.
- le datamart gestion des ventes secondaires.

## 6. Choix du schéma pour la modélisation

Nous avons opté pour le schéma en étoile pour la modélisation des différents datamarts, et ce pour trois principales raisons :

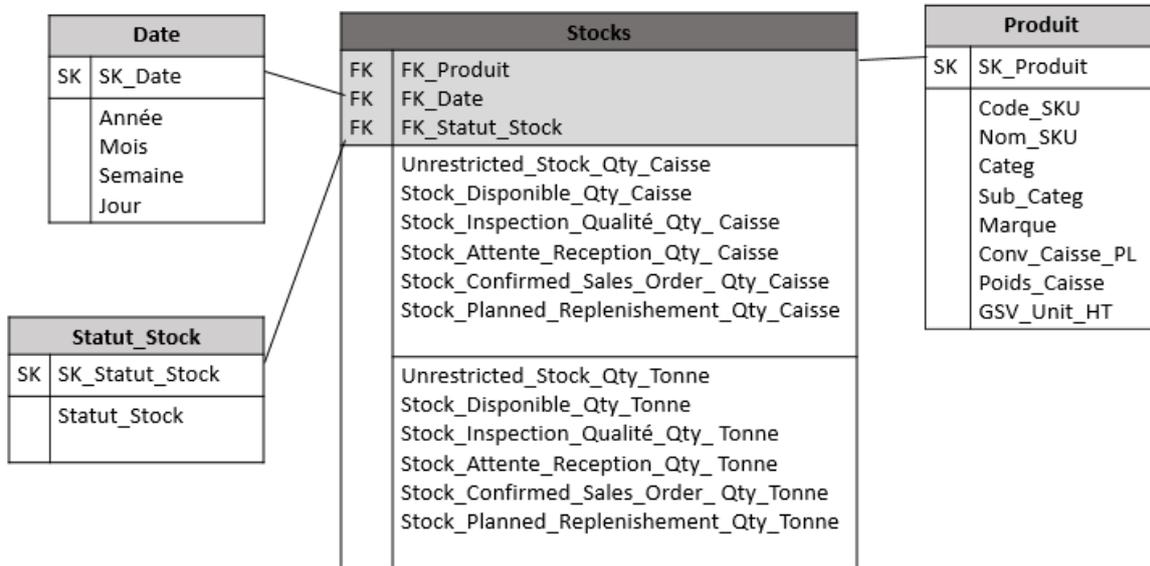
- La performance des requêtes est optimisée, car on y utilise peu de jointures. (Zouaghi, 2019)
- Ce schéma est facile à expliquer et à utiliser pour les utilisateurs finaux. (Tehreem, 2020)
- C'est la méthode préconisée dans la méthodologie de Kimball.

### 6.1. Modèles schémas multidimensionnels des datamarts

Les figures ci-dessous représentent le résultat de la modélisation dimensionnelle des différents sous processus :

- Le datamart gestion des stocks :

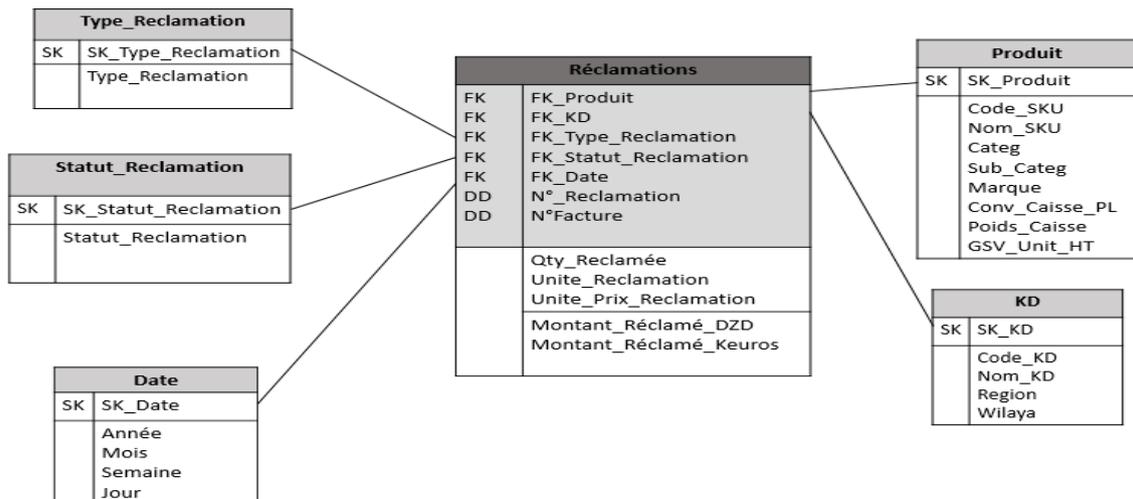
Le datamart gestion des stocks contient une table de fait : Stocks, et trois dimensions : Produit, Date et Statut\_Stock. Ces dimensions sont liées à la table de fait selon une schématisation en étoile.



**Figure 25 : Datamart Gestion des stocks (Nous-mêmes, 2020)**

- Le datamart gestion des réclamations :

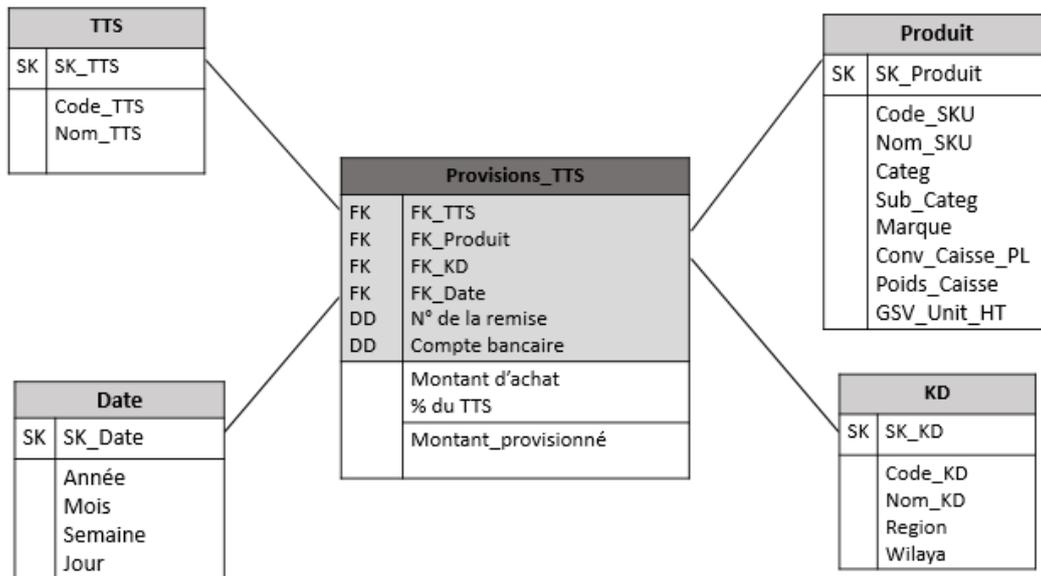
Le datamart gestion des réclamations contient une table de fait : Réclamations, et cinq dimensions : Produit, Date, Statut\_Reclamation, Type\_Reclamation, KD. Ces dimensions sont liées à la table de fait selon une schématisation en étoile.



**Figure 26 : Datamart Gestion des réclamations (Nous-mêmes, 2020)**

- Le datamart gestion des provisions TTS :

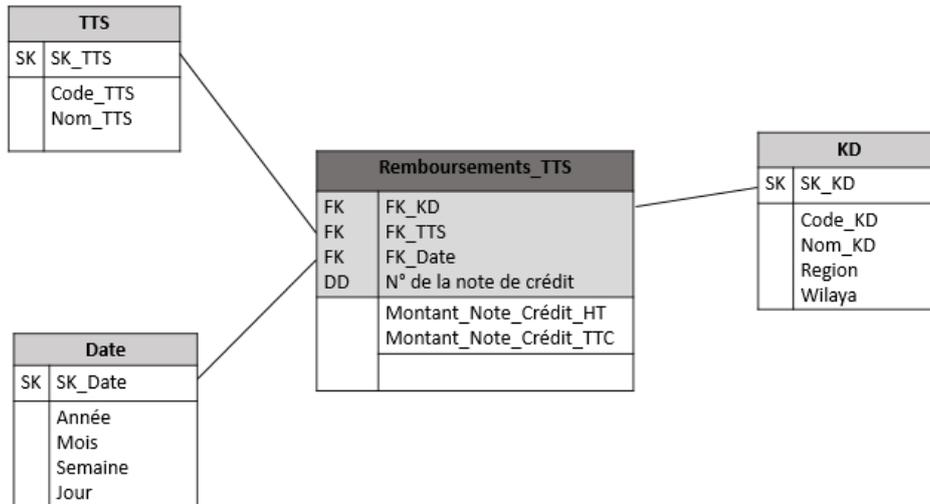
Le datamart gestion des provisions TTS contient une table de fait : Provisions\_TTS, et quatre dimensions : Produit, Date, KD et TTS. Ces dimensions sont liées à la table de fait selon une schématisation en étoile.



**Figure 27: Datamart gestion des Provisions\_TTS (Nous-mêmes, 2020)**

- Le datamart gestion des remboursements des TTS :

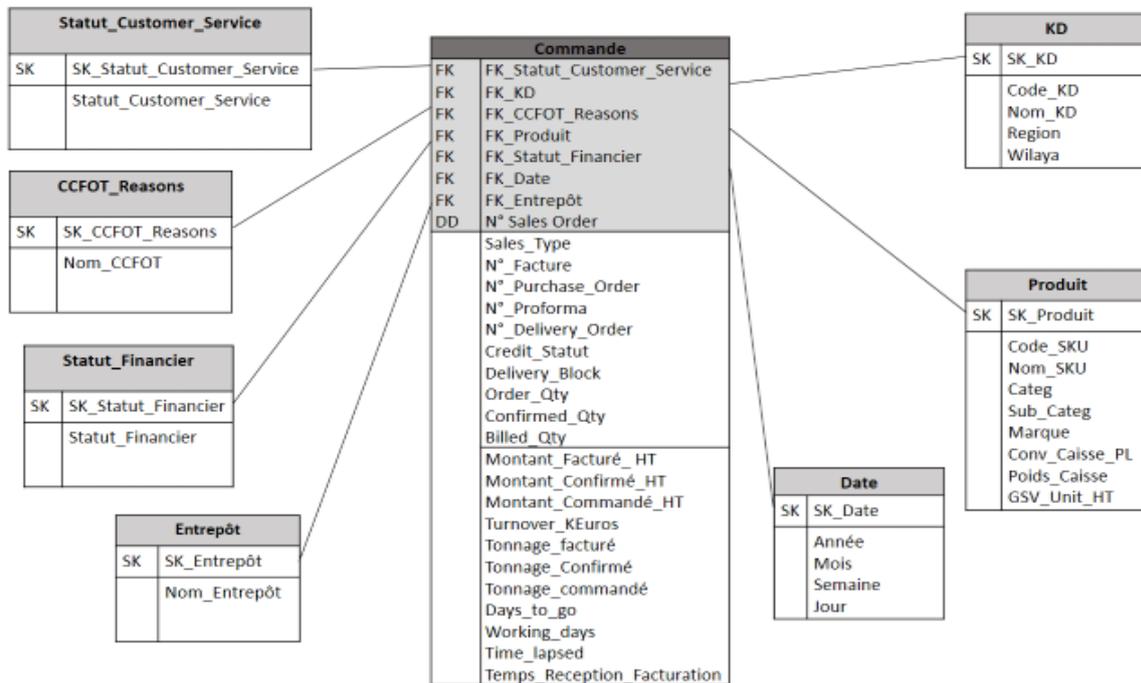
Le datamart gestion des remboursements des TTS contient une table de fait : Remboursements\_TTS, et trois dimensions : Date, KD et TTS. Ces dimensions sont liées à la table de fait selon une schématisation en étoile.



**Figure 28: Datamart gestion des Remboursements\_TTS (Nous-mêmes, 2020)**

- Le datamart gestion des commandes :

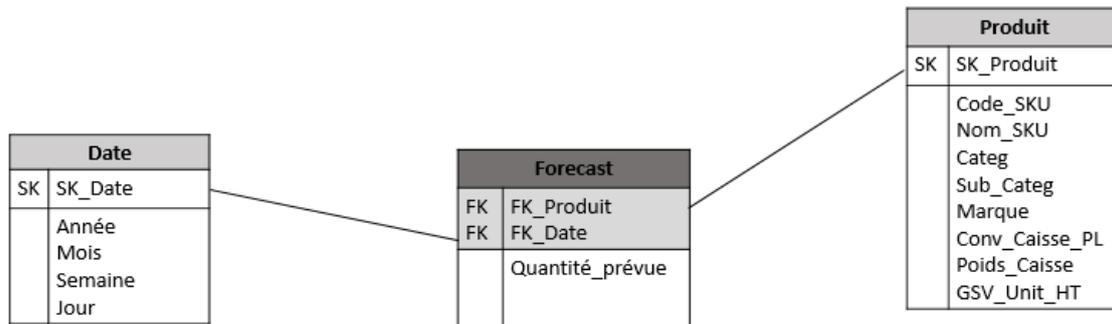
Le datamart gestion des commandes contient une table de fait : Commandes, et sept dimensions : Date, KD, Produit, Entrepôt, Statut\_Financier, CCFOT\_Reasons et Statut\_Customer\_Service. Ces dimensions sont liées à la table de fait selon une schématisation en étoile.



**Figure 29: Datamart Gestion des commandes (Nous-mêmes, 2020)**

- Le datamart gestion des forecasts :

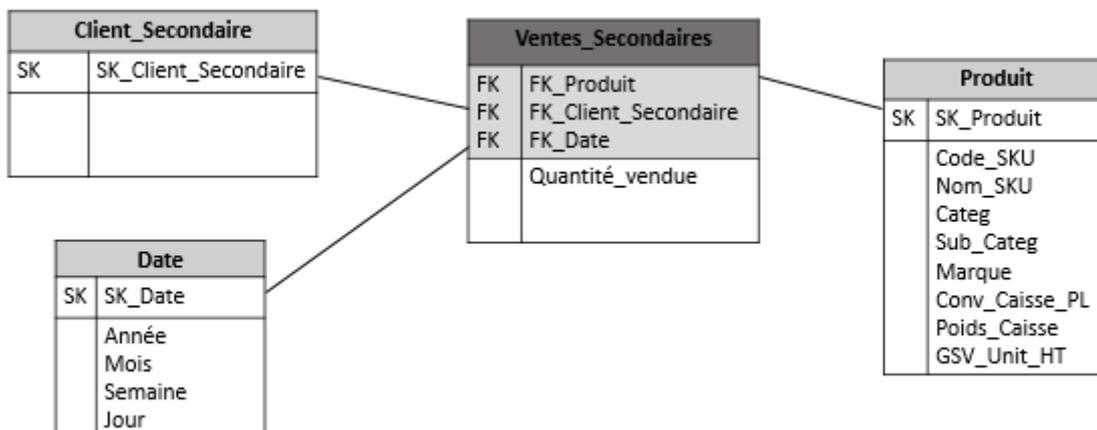
Le datamart gestion des remboursements des TTS contient une table de fait : Forecast, et deux dimensions : Date et Produit. Ces dimensions sont liées à la table de fait selon une schématisation en étoile.



**Figure 30: Datamart Forecast (Nous-mêmes, 2020)**

- Le datamart gestion des ventes secondaires :

Le datamart gestion des ventes secondaires contient une table de fait : Ventes\_secondaires et trois dimensions : Date, Client\_Secondaire et Produit. Ces dimensions sont liées à la table de fait selon une schématisation en étoile.



**Figure 31: Datamart Ventes secondaires (Nous-mêmes, 2020)**

## 6.2. Modèle schéma multidimensionnel du processus Order To Cash

La liaison des datamarts des différents sous process donne le datamart global du processus Order To Cash représenté avec le schéma galaxy :



## **7. Phase de stockage**

Dans cette phase nous avons créé les tables SQL des datamarts que nous avons décidé d'implémenter, à savoir : Gestion des commandes, Gestion des réclamations et Gestion des remboursements TTS.

Nous avons créé les tables SQL des tables de fait et des dimensions à l'aide de requêtes SQL (les requêtes de créations des tables dans SQL sont présentes dans l'Annexe 5). Nous avons par la suite défini les surrogate keys de chaque dimension comme étant des clés primaires et y avons lié les tables de fait à travers la définition de clés étrangères. Pour finir, nous avons rempli les champs de certaines dimensions peu peuplées.

## **8. Conclusion**

Dans ce chapitre nous avons modélisé le système décisionnel du service Customer Service. Pour ce faire, nous avons choisi comme méthode de conception la démarche de Kimball, qui se divise en deux principales phases : l'identification des besoins et la modélisation dimensionnelle.

A travers la conduite d'entretiens et l'analyse des différents rapports, nous avons pu identifier les besoins propres aux membres de l'équipe et les classer en thèmes.

Nous avons par la suite déroulé les quatre étapes de la modélisation dimensionnelle. En premier lieu, nous avons découpé le processus Order To Cash en sept sous-processus afin de faciliter la modélisation, et, après une hiérarchisation selon l'importance et la faisabilité, nous avons affecté les besoins à chaque sous-processus. Ensuite, nous avons déterminé le grain, choisi les dimensions et les mesures ainsi qu'une architecture de données adéquate à notre démarche.

Cela nous a permis d'obtenir sept datamarts reliés par des dimensions conformes, dont trois ont été stockés dans des tables SQL. Ces datamarts seront implémentés dans le prochain chapitre.

# **Chapitre IV : Réalisation de la solution BI**

## 1. Introduction

Ce chapitre sera dédié à la réalisation de notre solution BI que nous avons conçue dans le chapitre précédent.

Nous allons entamer ce chapitre par le choix des datamarts à réaliser, puis nous aborderons la phase d'alimentation qui consiste principalement en le déroulement du processus ETL.

Ensuite, nous passerons à la phase d'exploitation où nous réaliserons quelques mesures et expliciterons les opérations qu'il est possible de réaliser sur les besoins de suivi.

Pour finir, nous présenterons des exemples de rapport ainsi que des pistes d'amélioration possibles pour notre travail.

## 2. Choix des datamarts

Nous avons précédemment réalisé sept schémas multidimensionnels grâce à la modélisation dimensionnelle, ne pouvant tous les implémenter par contraintes de temps et de disponibilité de données, nous avons été contraintes d'en choisir trois pour notre implémentation.

Nous tenons à rappeler que nous avons choisi d'implémenter les datamarts suivants : Gestion des Commandes, Gestion des Réclamations et Gestion des Remboursements TTS, car ils sont à la fois faisables grâce à la disponibilité et importants pour le service.

Nous avons néanmoins dû opérer quelques changements à nos datamarts :

- **Datamart Gestion des Réclamations**

- 1) La granularité n'est pas respectée dans la table de fait Réclamations car deux réclamations y figurant peuvent avoir une granularité différente, une en pièce et l'autre en caisse. Nous avons donc décidé de scinder la table de fait Réclamations en deux : Réclamations\_Caisses pour les réclamations dont l'unité est la caisse et Réclamations\_Pièces pour les réclamations dont l'unité est la pièce. Cela ajoute une complexité supplémentaire pour l'analyse mais cela est inévitable car la granularité doit être respectée et la donnée permettant de convertir de la pièce à la caisse n'est pas disponible.
- 2) Nous avons également décidé d'enlever l'unité de prix de la réclamation car ce champ comptait beaucoup de champs vides dans le fichier source, le prix et son unité sont présents dans la dimension Produit.
- 3) Nous avons remarqué qu'une réclamation pouvait être accompagnée de plusieurs factures, pouvant aller jusqu'à quatre, nous avons donc décidé de diviser la colonne N°facture en quatre colonnes, une pour le code de chaque facture. Nous l'avons également déplacée de la catégorie des dimensions à celles des mesures de base.

- **Dimension Produit**

1) L'ajout d'une colonne 'Price Unit' donnant l'unité dans laquelle est exprimée le prix du produit. (Prix par caisse ou par pièce.)

### 2.1.Récapitulatif des datamarts à réaliser

Les tableaux ci-dessous résument la structure finale des tables de fait des datamarts que nous avons décidé de réaliser :

- **Datamart Gestion des Commandes**

**Tableau 23 : Récapitulatif Datamart Gestion des Commandes (Nous-même, 2020)**

Table de fait	Grain	Dimensions	Mesures	
<b>Commandes</b>	Chaque ligne de la table de fait représente une ligne d'un Sales Order, c.à.d une certaine quantité commandée en caisse, une certaine quantité allouée sur système, une quantité facturée pour un produit à un client particulier en une date particulière.	Produit	<b>De base</b>	Sales_Type
		Client		N°Facture
		Date		N° Purchase Order
		Statut_Customer_Service		N° Proforma
		CCFOT_Reason		N° Delivery_Order
		Statut_Financier		Order_Qty
		Entrepôt		Confirmed_Qty
		SO_Number (DD)		Billed_Qty (Quantité facturée)
			Credit_Status (Statut du crédit de ligne du Sales Order)	
			Delivery_Block	
			<b>Dérivés</b>	Nombre_commande
				MTD
				Montant_Facturé_HT
				Montant_Confirmé_HT
				Montant_Commandé_HT
				Tonnage facturé
				Tonnage_Confirmé
		Tonnage_commandé		
		Working_days		
		Time_lapsed		

- **Datamart Gestion des Réclamations en caisses**

**Tableau 24 : Récapitulatif Datamart Réclamations en caisses (Nous-même, 2020)**

Table de fait	Grain	Dimensions	Mesures	
<b>Réclamations_caisses</b>	Chaque ligne de la table de fait représente le montant à rembourser d'une réclamation en caisses d'un certain type, au sujet d'un seul produit, faite par un seul client, en une date donnée accompagnée ou non d'une à quatre factures.	Client	<b>De base</b>	Quantité_ réclamée
		Date		
		Produit	<b>Dérivées</b>	Montant_ Réclamé_ DZD
		Type de la réclamation		

- **Datamart Gestion des Réclamations en pièces**

**Tableau 25 : Récapitulatif Datamart Réclamations en pièces (Nous-même, 2020)**

Table de fait	Grain	Dimensions	Mesures		
<b>Réclamations_pièces</b>	Chaque ligne de la table de fait représente le montant à rembourser d'une réclamation en pièces d'un certain type, au sujet d'un seul produit, faite par un seul client, en une date donnée accompagnée ou non d'une à quatre factures.	Client	<b>De base</b>	Quantité_ réclamée	
		Date			
		Produit	<b>Dérivées</b>	Montant_ Réclamé_ DZD	
		N° de la réclamation (DD)			Nombre_ Réclamati on
		Type de la réclamation			
		Statut de la réclamation			

- **Datamart Gestion des Remboursements TTS**

**Tableau 26 : Récapitulatif Datamart Gestion des Remboursements TTS (Nous-même, 2020)**

Table de fait	Grain	Dimensions	Mesures	
<b>Remboursements_TTS</b>	Chaque ligne de la table de fait représente une note de crédit au sujet d'un certain TTS, concernant un seul client en une date donnée.	N° de la note de crédit(DD)	<b>De base</b>	Montant de la note de crédit HT
				Montant de la note de crédit TTC
		Date	<b>Dérivée</b>	Nombre_Credit_Note
		TTS		Somme_Montant_Not e_Credit_HT
				Somme_Montant_Not e_Credit_TTC

### 3. Implémentation de la solution

Après avoir conçu les datamarts, nous sommes passées à la phase d'implémentation de notre solution :

#### 3.1. Phase d'alimentation

Cette phase est la phase d'alimentation, dans laquelle est réalisé l'ETL :

- **Extract** : L'extraction de données, réalisée à partir de fichiers sources, sous un certain format.
- **Transform** : L'application de transformations, telles que des unions, des conditions ou des jointures, aux données récoltées à partir des sources.
- **Load** : Le chargement des données traitées dans les datamarts.

##### 3.1.1. Choix du logiciel pour réaliser l'ETL

Pour effectuer notre ETL de la manière la plus efficace nous avons dû faire le choix d'un logiciel adapté.

Pour choisir l'outil ETL, nous avons évalué 3 logiciels figurant parmi les leaders du marché en fonction de la longueur, de leur version d'essai et de la disponibilité de tutoriels pour l'apprentissage de leur utilisation. Le tableau suivant résume la comparaison entre les trois logiciels :

**Tableau 27 : comparaison entre les leaders du marché des logiciels de l'ETL**

Logiciel	Informatica	SSIS	Talend
Longueur de la version d'essai	30 jours	Version community illimitée pour les étudiants et les environnements d'apprentissage en entreprise.	15 jours
Disponibilité de tutoriels	Oui	Oui	Oui

Nous avons donc choisi SSIS pour la disponibilité des tutoriels et l'accès à la version Community.

Après notre départ, les utilisateurs de l'entreprise peuvent se tourner vers des options moins coûteuses ou gratuites, selon leurs besoins, ils auront néanmoins la liste des opérations à réaliser pour l'ETL et pourront utiliser la version Community de SSIS pour apprendre.

SSIS est un logiciel créé par Microsoft, qui offre des outils de nettoyage et de transformation des données. Il permet de réaliser des workflows qui ont pour but l'extraction, la transformation et le changement des données (ETL) de diverses sources vers diverses destinations. L'architecture du logiciel est présentée en Annexe 6.

- **Avantages** : (WORT, sd)

Selon le scénario d'utilisation, SSIS présente des solutions permettant de faciliter la gestion des données dans l'ETL. Les principaux avantages de SSIS sont :

- **La possibilité de dénormaliser les données**, c.a.d de combiner des données provenant de différentes sources de données. SSIS se charge d'homogénéiser et de fusionner ces données en un seul groupe de données.
- **La maintenance des data warehouses et des datamarts** : SSIS permet de charger de grandes volumes de données obtenues à partir de fichiers plats dans des tables de bases de données, tout en intégrant la possibilité d'anticiper sur les erreurs pouvant survenir durant l'ETL, réduisant ainsi son temps d'exécution.
- **La gestion de l'historique** : à l'aide de l'Assistant Dimension à évolution lente. Cet assistant crée et configure un ensemble de tâches de transformation de données utilisées pour gérer l'insertion et la mise à jour des enregistrements ainsi que l'ajout de nouvelles colonnes aux tables pour assurer le suivi de l'historique.
- **Transformations dynamiques des données**, grâce à un riche ensemble de transformations utiles pour le fractionnement ou la fusion de données selon des conditions précises, ou encore l'application de calculs sur des parties d'un ensemble de données.
- **Inconvénients** : SSIS présente néanmoins quelques inconvénients, dont :
  - Certaines fonctionnalités ne sont pas compatibles avec les processeurs de type 64 bits.
  - La conversion des données est parfois difficile à mettre en œuvre.
  - Le zoom de l'écran n'est pas facile à régler.
  - Le logiciel ne fonctionne que sur Microsoft.
  - Le logiciel a un usage conséquent de la mémoire. (DESAI, 2016) (CHRIS, 2016)

### 3.1.2. Mode de chargement des données

Tous les chargements ont eu en lieu en mode PUSH, c'est-à-dire que le système transactionnel poussait les données vers la Staging Area.

### 3.1.3. Déroulement de l'ETL

Dans ce qui suit, nous détaillons les étapes suivies pour la réalisation de l'ETL des différentes dimensions et tables de faits de nos datamarts. Les sources de données utilisées sont présentées en Annexe 4.

#### 3.1.3.1. Dimension Produit

Nous avons précisé les sources à partir desquelles nous avons extraits les données, les transformations opérées et les données à charger.

##### 3.1.3.1.1. Extraction

Pour réaliser la dimension Produit nous avons extrait les données de deux feuilles Excel :

- **Material** : Feuille contenant le code du SKU, son nom, sa catégorie et sa marque.
- **List Price** : Feuille contenant le code du SKU, son prix et l'unité dans laquelle son prix est exprimé.
- Les champs Sub\_Catégorie et Poids\_en\_Caisse ne sont pas remplis lors de cet ETL car nous n'avons pas accès aux données, néanmoins, nous les avons gardés car ils pourront être regroupés par la suite par les utilisateurs.

### 3.1.3.1.2. Transformation

- 1) SSIS affecte des types de données aux colonnes des fichiers sources, il a donc identifié la colonne du code SAP du produit comme étant un float, par conséquent, il a fallu convertir le code SKU en entier, car son type dans la table Produit est entier et l'intégrité des données doit être respectée.
- 2) Puis nous avons trié les deux groupes de données selon le code\_SKU pour pouvoir les joindre avec une jointure de type Left Outer Join, où l'entrée gauche est liée aux valeurs correspondantes présentes dans l'entrée. Cette jointure affecte le prix et unité de prix des produits, représentant l'entrée gauche, aux produits présents dans la feuille Material, et ignore ceux non-présents.
- 3) Nous avons par la suite réalisé une autre jointure de type Left Outer Join, cette fois-ci pour joindre les données précédemment obtenues avec celles déjà présentes dans la table Produit.
- 4) Nous avons utilisé un fractionnement conditionnel afin de séparer les nouvelles lignes et les lignes modifiées de celles déjà présentes dans la dimension Produit ou n'ayant pas subi de modification.

### 3.1.3.1.3. Chargement

Les nouvelles lignes sont insérées dans la dimension Produit, tandis que les champs des lignes ayant subi une modification sont modifiés.

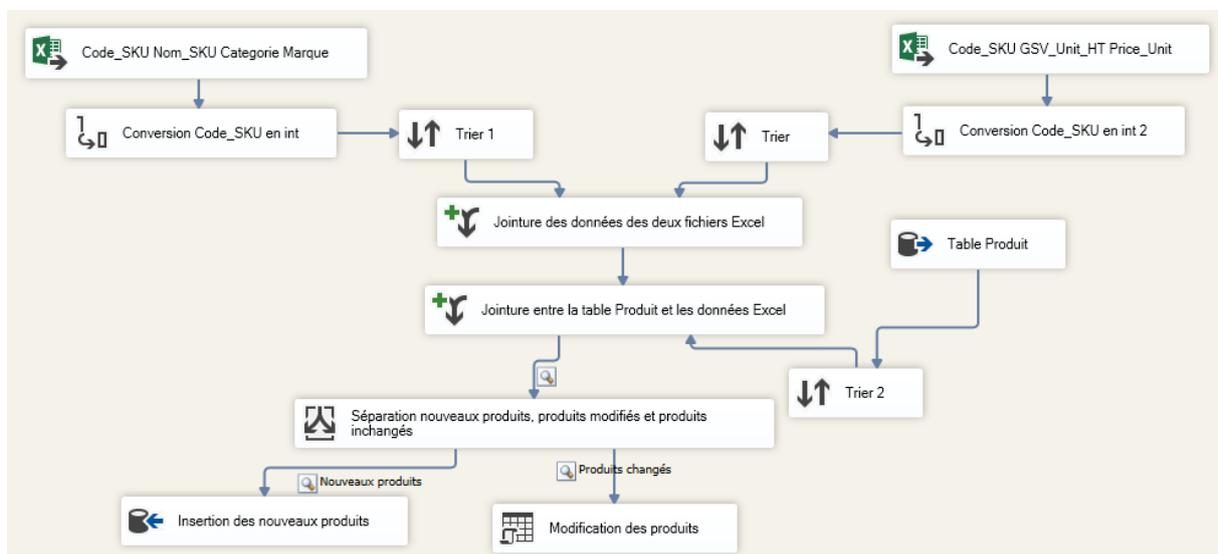


Figure 33 : Schéma ETL de la dimension Produit (Nous-même, 2020)

### 3.1.3.2. Dimension KD

Nous avons précisé les sources à partir desquelles nous avons extraits les données, les transformations opérées et les données à charger.

#### 3.1.3.2.1. Extraction

Pour réaliser l'ETL de la dimension KD nous avons utilisé les données de deux feuilles Excel :

- **KD List** : Contenant la liste des KDs, il s'agit d'une feuille avec deux colonnes : le code SAP et le nom des KDs.
- **Region** : Comportant la liste des KDs, avec uniquement le code SAP et la région comme colonnes.
- Le champ Wilaya n'est pas rempli lors de cet ETL car nous n'avons pas accès aux données, néanmoins, nous l'avons gardé car il pourra être rempli par la suite par les utilisateurs.

### 3.1.3.2.2. Transformation

- 1) Nous avons commencé par récupérer les données des deux feuilles et par convertir le code\_KD, détecté comme étant un float, en entier. Nous également remplacé un vide dans la liste de **Region** par le code KD correspondant.
- 2) Nous avons effectué deux tri préliminaires à la jointure des deux sources de données, il s'agit d'un Left Outer Join, l'entrée gauche étant la feuille **KD list**, pour identifier les KDs dont la région est connue et ceux dont la région n'est pas connue.
- 3) Nous avons par la suite utilisé un fractionnement conditionnel pour séparer les KDs sans région de ceux avec région.
- 4) Nous avons de nouveau effectué un tri, cette fois-ci sur les données avec région obtenues après le fractionnement conditionnel, afin de les joindre avec les données déjà présentes dans la table Client.
- 5) Pour finir, nous avons utilisé un fractionnement conditionnel afin de séparer les nouvelles lignes de celles déjà présentes dans la dimension Client.

### 3.1.3.2.3. Chargement

Les nouvelles lignes sont insérées dans la dimension KD.

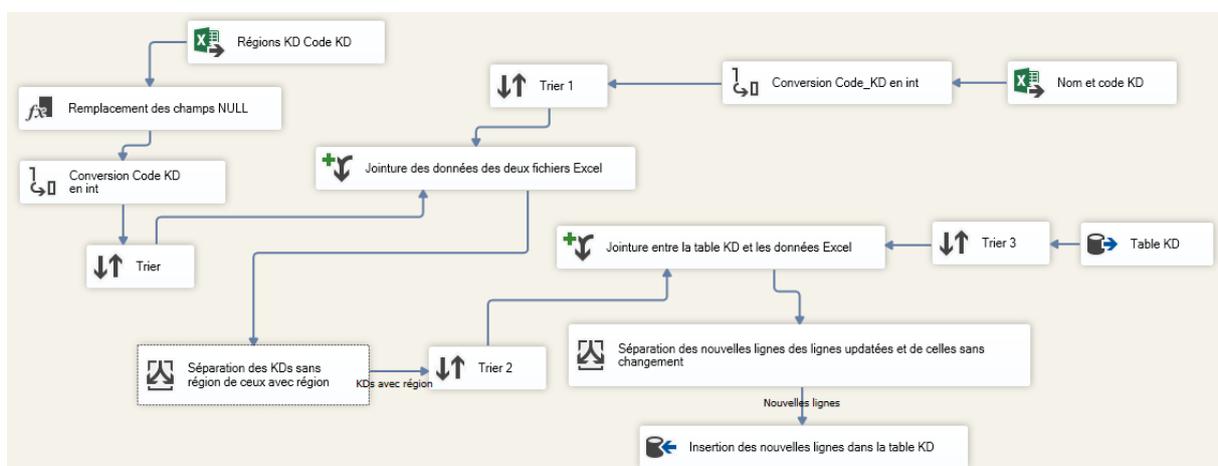


Figure 34 : Schéma de l'ETL de la dimension KD (Nous-mêmes, 2020)

### 3.1.3.3.Dimension TTS

Nous avons précisé les sources à partir desquelles nous avons extraits les données, les transformations opérées et les données à charger.

#### 3.1.3.3.1. Extraction

Pour réaliser l'ETL de la dimension TTS nous avons utilisé les données d'un unique fichier Excel **CN JUL 2019**, dont nous avons tiré le nom des TTS à partir de la colonne DESIGNATION.

#### 3.1.3.3.2. Transformation

- 1) Après avoir récupéré les données de la colonne DESIGNATION du fichier CN JUL 2019, nous avons procédé à un tri qui nous a permis d'éliminer les doublons figurant dans la colonne.
- 2) Par la suite, nous avons trié des données provenant de la table TTS, afin de les joindre avec celles précédemment récoltées du fichier Excel, grâce à un Left Outer Join où les données Excel ont représenté l'entrée gauche.
- 3) Nous avons utilisé un fractionnement conditionnel pour séparer les nouvelles lignes de celles déjà présentes dans la table TTS.

#### 3.1.3.3.3. Chargement

Les nouvelles lignes sont insérées dans la dimension TTS.

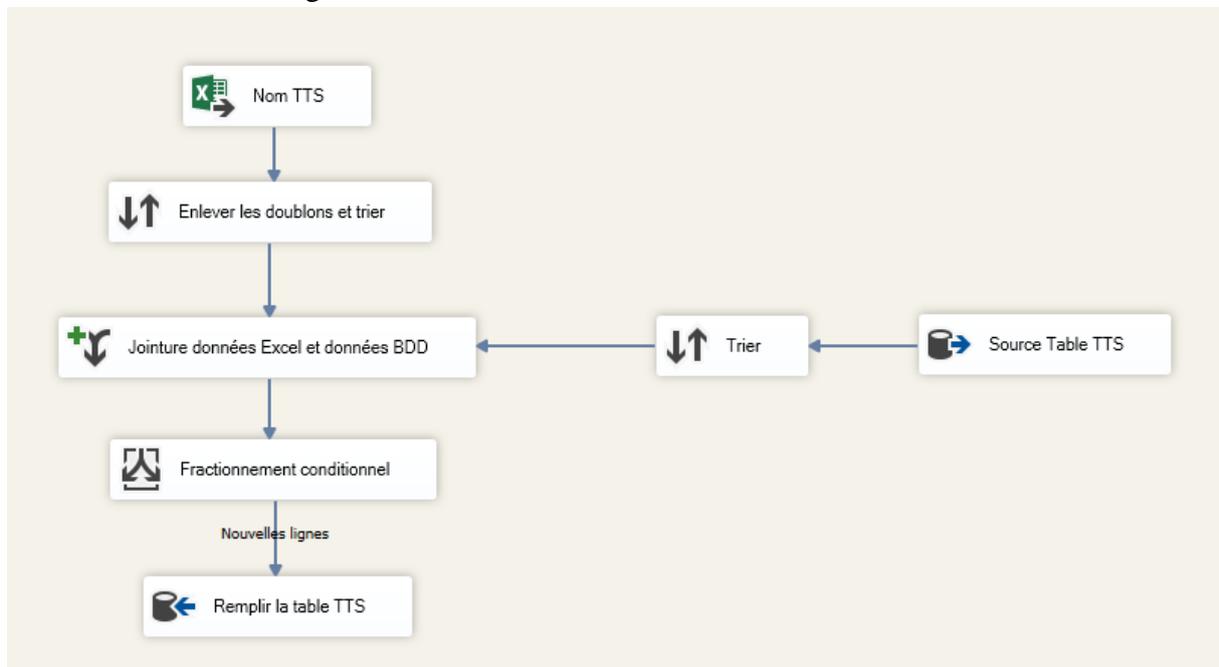


Figure 35 : ETL de la dimension TTS (Nous-même, 2020)

#### 3.1.3.4.Datamart Remboursements TTS

Nous avons précisé les sources à partir desquelles nous avons extraits les données, les transformations opérées et les données à charger.

### 3.1.3.4.1. Extraction

Pour réaliser la table de fait Remboursements\_TTS nous avons extrait les données de la feuille Excel **CN 2019**, dont nous avons extrait le code SAP du KD, la date où le remboursement a eu lieu ainsi que le nom du TTS appliqué. La source est mise à jour d'une manière mensuelle avec les remboursements du mois.

### 3.1.3.4.2. Transformation

- 1) Nous avons d'abord converti le code SAP du KD en entier et la date en format de base de données, pour s'ajuster aux types de données de la table TTS\_Remboursements.
- 2) Nous avons par la suite effectué trois transformations de type Recherche, transformations qui chargent une table SQL, puis comparent des valeurs des données extraites à celles présentes dans la table afin de détecter les correspondances, et, en cas de correspondance, d'affecter une certaine valeur présente dans une autre colonne de la table, tel que le code de la surrogate key de la table.

Ces trois transformations nous ont permis de créer trois colonnes : FK\_KD, FK\_Date et FK\_TTS, dont les valeurs représentent les correspondances des surrogate key des trois dimensions avec les données extraites. Ces colonnes s'ajoutent aux données extraites.

- 3) Ensuite, nous avons trié les données obtenues afin de les joindre avec les données déjà présentes dans la table Remboursements\_TTS, grâce à un Left Outer Join où les données de la table représentent l'entrée droite.
- 4) Nous avons séparé les nouvelles lignes de celles déjà présentes grâce à un fractionnement conditionnel.

### 3.1.3.4.3. Chargement

Nous avons chargé les nouvelles lignes dans la table Remboursements TTS.

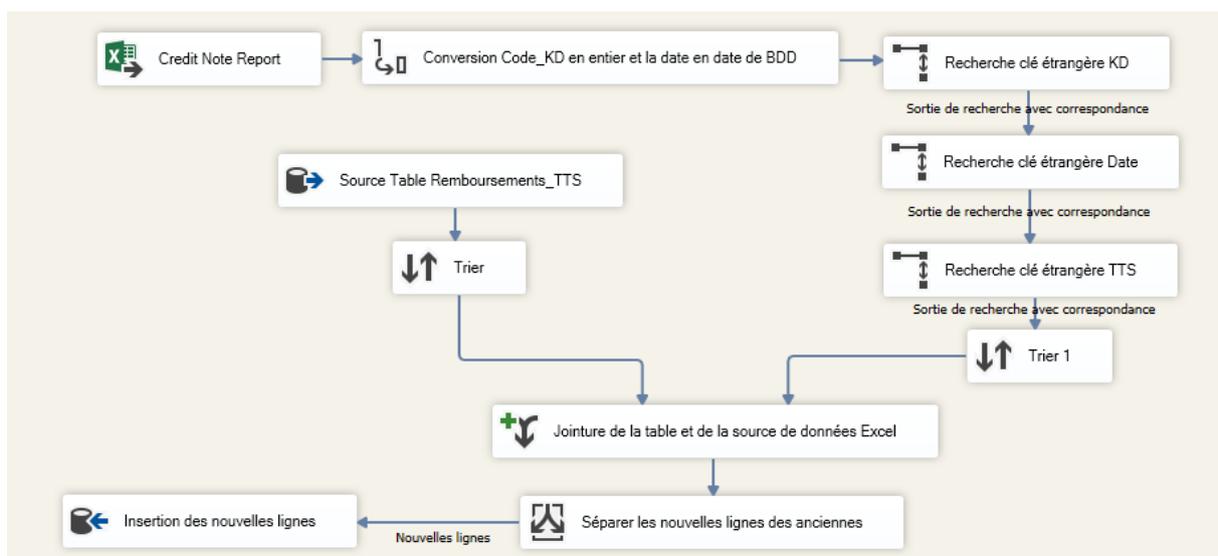


Figure 36 : Schéma de l'ETL des remboursements TTS (Nous-même, 2020)

### **3.1.3.5.Datamart commandes**

Nous avons précisé les sources à partir desquelles nous avons extraits les données, les transformations opérées et les données à charger. La source est mise à jour à chaque changement opéré sur la commande.

#### **3.1.3.5.1. Extraction**

Pour réaliser la table de fait Commandes, nous avons extrait les données de la feuille Excel Today's Orders, dont nous avons extrait plusieurs colonnes.

#### **3.1.3.5.2. Transformation**

- 1) Nous avons d'abord utilisé deux fractionnements conditionnels, le premier afin de séparer les lignes selon leur statut financier, Under Finance ou Not Under finance. Le second pour séparer les lignes selon leur statut dans le Customer Service : In treatment, Waiting for shipment ou encore Invoiced, tout en prenant soin d'unir de nouveau les lignes après chaque fractionnement conditionnel.
- 2) Ensuite, nous avons utilisé deux conversions afin de convertir le code KD et le code de l'SKU en entier puis deux Recherches pour affecter les clés étrangères correspondantes. Après quoi nous avons de nouveau utilisé quatre autres Recherches afin d'affecter les clés étrangères correspondantes au Statut\_financier, au Statut\_customer service, à l'entrepôt ainsi qu'à la CCFOT\_reasons.
- 3) Nous avons remplacé les points par des tirets dans les deux dates, et nous les avons converties au format date de bases de données, puis nous avons utilisé deux dernières Recherches pour affecter les clés étrangères propres à Bill\_Date et SOCreation\_Date, en prenant en considération le cas où il n'y a pas encore de Bill\_Date car la commande n'a pas encore été facturée.
- 4) Nous avons opéré un tri sur les données obtenues afin de les joindre avec les données déjà présentes dans la table Commande, grâce à un Left Outer Join où les données de la table représentent l'entrée droite.
- 5) Pour finir, nous avons utilisé un fractionnement conditionnel pour séparer les données en trois flux : les données déjà présentes dans la table, celles ayant subi un changement et celles qui n'y ont pas encore été insérées.



### **3.1.3.6.1. Extraction**

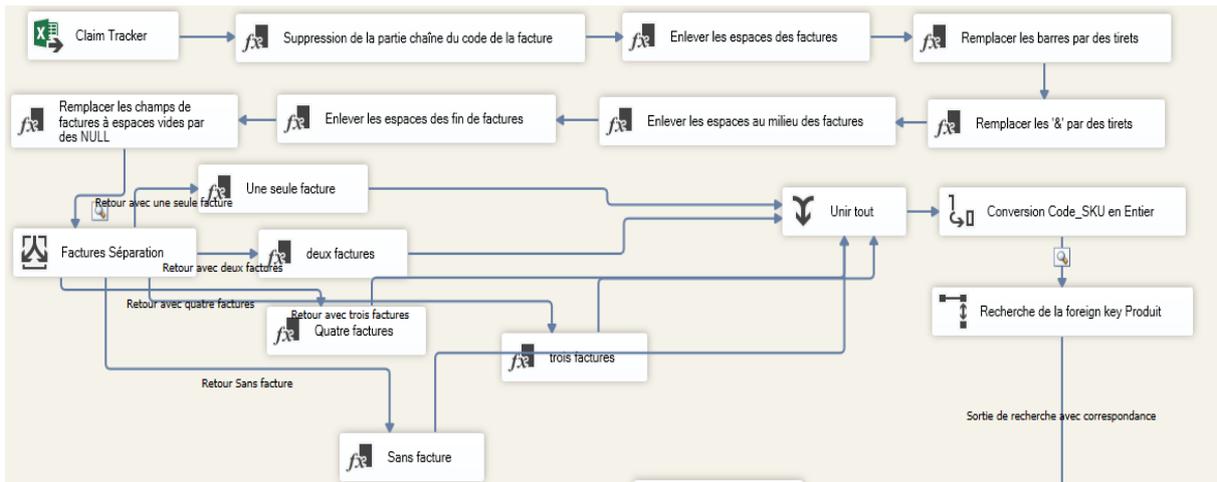
Pour réaliser les tables de faits Réclamations\_caisses et Réclamations\_pièces, nous avons utilisé le fichier Excel Claim Tracker, qui regroupe toutes les réclamations reçues et qui est actualisé à chaque fois que le statut d'une réclamation change.

### **3.1.3.6.2. Transformation**

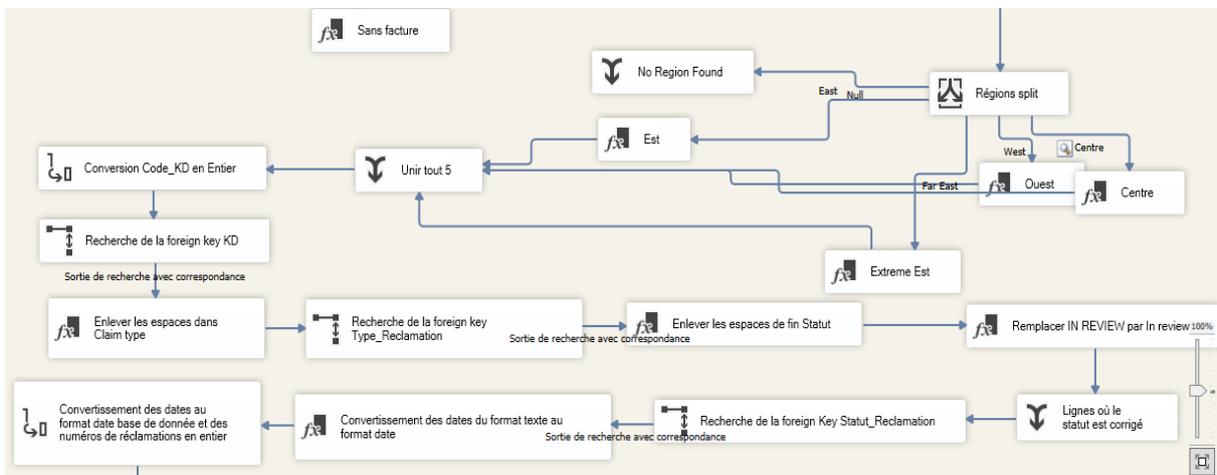
- 1) Nous avons commencé par une série de traitements appliqués aux colonnes des sources :
  - Suppression de la partie chaîne de caractères du code de la facture, que nous avons ajoutée pour que le logiciel reconnaisse les données de la colonne comme étant de type chaîne plutôt que float, car dans ce dernier cas les lignes contenant des chaînes plutôt que des chiffres étaient remplacées par des NULL.
  - Suppression des espaces en début/ fin des codes de factures
  - Remplacement des '/' et des '-' présents dans les codes de factures par des '-'.
  - Suppression des espaces présents au milieu des codes des factures.
  - Remplacement des champs de factures à espaces vides par des NULL.
- 2) A l'aide d'un fractionnement conditionnel, nous avons séparé les lignes selon le nombre de factures accompagnant la réclamation, et avons créé une colonne récupérant le code de chacune, puis regroupé toutes ces lignes avec une union.
- 3) Afin d'obtenir la colonne FK\_Produit, nous avons converti le code\_sku en entier et utilisé la transformation Recherche.
- 4) Nous avons séparé les lignes selon la valeur de la région pour la faire correspondre avec les régions présentes dans la dimension KD.
- 5) Nous avons procédé à la conversion du Code\_KD en entier, et utilisé une fonction recherche pour créer FK\_KD, en utilisant les colonnes Code\_KD\_int et Region.
- 6) Nous avons supprimé les espaces de fin de chaîne dans Claim Type et Statut, et affecté les clés étrangères correspondantes.
- 7) Nous avons converti les dates du format texte au format date, puis du format date au format BDD, puis affecté la clé étrangère correspondante.
- 8) Par la suite, nous avons corrigé les champs de la colonne Claims Unit pour qu'elles soient unifiées en 'CS' et 'PC'.
- 9) Puis nous avons utilisé un fractionnement conditionnel pour séparer les réclamations en caisses de celles en pièces.
- 10) Au final, nous avons utilisé le même traitement pour ces deux types de réclamations : Jointure avec les données déjà présentes dans les tables Réclamations\_caisses et Réclamations\_pièces, suivie d'un fractionnement conditionnel pour séparer les nouvelles lignes, les lignes dont le statut a changé et les lignes déjà présentes dans les tables et n'ayant subi aucun changement.

### **3.1.3.6.3. Chargement**

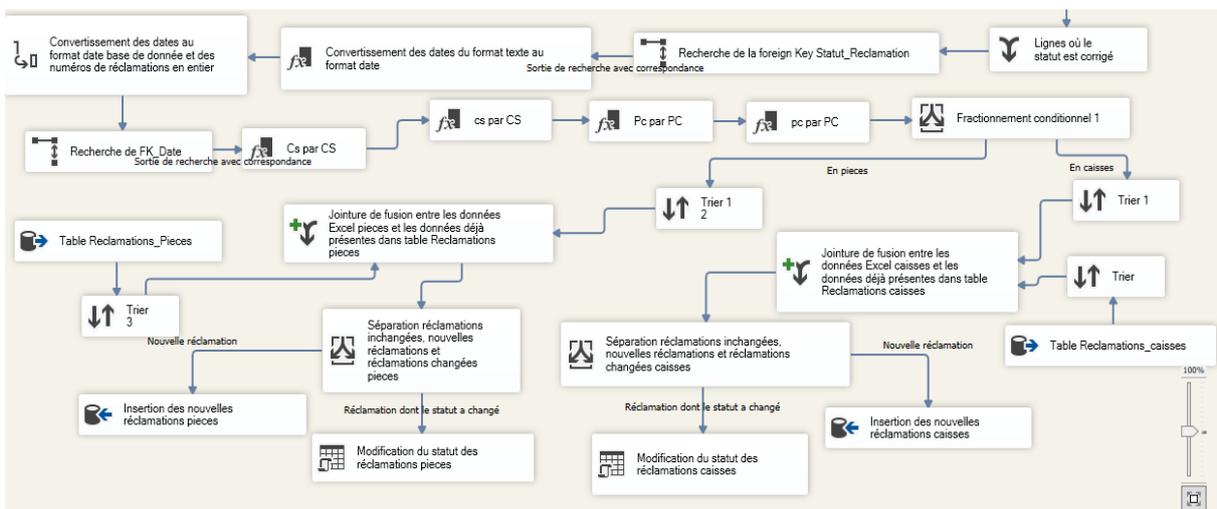
Nous avons chargé les nouvelles lignes dans les tables Réclamations\_caisses et Réclamations\_pièces et nous avons modifié le statut des lignes dont le statut avait changé.



**Figure 39 : Schéma de l'ETL de la table de fait Réclamations - 1 (Nous-même, 2020)**



**Figure 40 : Schéma de l'ETL de la table de fait Réclamations - 2 (Nous-même, 2020)**



**Figure 41 : Schéma de l'ETL de la table de fait Réclamations - 3 (Nous-même, 2020)**

Selon le détail de certaines transformations et opérations effectuées, des exemples sont présentés en Annexe 7.

### 3.2. Phase d'exploitation

Il s'agit du calcul des mesures et de l'élaboration des axes de suivi pour la préparation des rapports.

#### 3.2.1. Choix du logiciel pour réaliser les analyses (Desai, 2020) (data-flair, 2018)

Power BI est une plateforme de Business Intelligence réalisée par Microsoft. Il permet d'analyser en profondeur les données et de partager des informations pertinentes. Nous l'avons choisi car le service a déjà une licence pour son utilisation.

- **Avantages :** Power BI présente de nombreux avantages, tels que :
  - **Facilité d'utilisation :** Power BI a une interface très simple et facile à utiliser qui ne nécessite aucune expérience en programmation. Il possède une intelligence intégrée qui aide à sélectionner les meilleures formes à utiliser pour la création de rapports.
  - **Facilité d'apprentissage :** Power BI suit une approche similaire à celle d'Excel pour la conception des rapports. Excel étant un logiciel largement utilisé, cela facilite l'apprentissage de Power BI. En plus de cela, la modélisation des données est dérivée des principes fondamentaux de Microsoft SQL Server et de la base de données Microsoft Access. Ainsi, les utilisateurs et les programmeurs peuvent très facilement utiliser la modélisation des données de Power BI.
  - **Facilité de partage :** Power BI permet à l'utilisateur de collaborer avec ses collègues pour créer des rapports et des tableaux de bord interactifs.
  - **Rentabilité :** Power BI desktop est gratuit et permet néanmoins de créer des rapports très complexes. Les licences professionnelles de Power BI offrent des fonctionnalités supplémentaires mais demeurent dans une fourchette abordable (10 USD par mois).
  - **Large couverture des sources de données :** Power BI est livré avec une large gamme de connecteurs à des sources de données telles que Microsoft Excel, SQL Server, base de données Oracle, base de données IBM DB2....
  - **Outil puissant :**
    - ❖ **Visualisation :** Power BI dispose d'une énorme bibliothèque pour la création de modèles de visualisations personnalisés.
    - ❖ **Mise en forme des données :** Power BI propose un éditeur de requêtes (Query Editor) très flexible et puissant avec diverses fonctionnalités. Sa puissance est renforcée par l'usage du langage DAX.
    - ❖ **La modélisation des données :** Power BI offre des options de modélisation de données très efficaces inspirées de la modélisation SQL et des cubes multidimensionnels.
- **Inconvénients :** Power BI présente néanmoins quelques inconvénients, on peut citer :
  - Le langage DAX n'est pas facile à apprendre pour les utilisateurs.
  - Les tables ne peuvent pas avoir plus d'une relation active les liant.

### 3.2.2. Création de colonnes et mesures dérivées

Nous allons procéder à la création de mesures dérivées. Pour en réaliser certaines, il nous faudra préalablement créer des colonnes dérivées, qu'il est possible de créer sur Power BI en ajoutant une colonne.

Pour créer ces colonnes et mesures on utilise un langage de programmation propre à Power BI, il s'agit de DAX.

DAX permet de manipuler aussi bien des variables numériques que non numériques, et offre toute une série de fonctions pour la création de colonnes et mesures dérivées :

- Fonctions d'agrégation : telles que Min, Max, Average, Sum...
- Fonctions de comptage : telles que Count, DistinctCount, CountRows, CountBlank...
- Fonctions logiques : telles que AND, OR, NOT, IFERROR...
- Fonctions de dates : telles que DATE, WEEKDAY, NOW...
- Fonctions de texte : REPLACE, UPPER, CONCATENATE...

#### 3.2.2.1. Colonnes dérivées

Nous avons d'abord chargé les tables de faits ainsi que leurs dimensions associées depuis SQL Server vers Power BI. Les modèles des datamarts sont présentés en Annexe 8.

Dans le tableau suivant, nous avons résumé les colonnes dérivées qui nous ont permis de calculer les mesures :

**Tableau 28 : Formules de création des colonnes dérivées (Nous-même, 2020)**

Colonne dérivée	Formule
Montant_Comm andé	'Commandes'[Order_qty] * 'Commandes'[GSV_Unit]
Montant_Confir mé	Commandes'[Confirmed_qty] * 'Commandes'[GSV_Unit]
Montant_Factor é	'Commandes'[Billed_qty] * 'Commandes'[GSV_Unit]
Price_Unit	RELATED(Produit[Price_Unit])
GSV_Unit	IF('Commandes'[Price_Unit]= "CS";RELATED(Produit[GSV_Unit_ HT]) ;0)
Working day	IF(NOT('Date'[DayofTheWeek]==6  ('Date'[DayofTheWeek]==7)), "Travail", "Conge")
DayOfTheWeek	DayofTheWeek = weekday('Date'[TheDate],1)

#### 3.2.2.2. Mesures dérivées :

Le tableau ci-dessous présente les différentes mesures dérivées :

**Tableau 29 : Formules de calcul des mesures dérivées (Nous-même, 2020)**

Mesure dérivée	Formule
Nombre_Reclamation_CS	DISTINCTCOUNT('Réclamations_Caisses'[Numero_Reclamation])
Nombre_Reclamation_PC	DISTINCTCOUNT('Réclamations_Pieces'[Numero_Reclamation])
Nombre_Total_Réclamations	Nombre_Reclamation_CS+Nombre_Reclamation_PC
Somme_Montant_Réclamations_CS	SUM('Réclamations_Caisses'[Montant_Reclamation_CS])
Somme_Montant_Réclamations_PC	SUM('Réclamations_Caisses'[Montant_Reclamation_PC])
Nombre_Commande	DISTINCTCOUNT(Commandes[SO_Number])
Montant_Facturé_Total = GSV	SUM('Commandes'[Montant_Facturé])
MTD	MTD = TOTALMTD(SUM(Commandes[Confirmed_qty]); 'Date'[TheDate])
NombreJoursOuvrables	COUNT ('Date'[WorkingDay])
Time_lapsed	DATEDIFF ( STARTOFMONTH('Date'[TheDate]) ; Today(); DAY)
Somme_Montant_Note_Credit_TTC	SUM(Remboursements_TTS[Montant_Note_Credit_TTC])
Somme_Montant_Note_Credit_HT	SUM(Remboursements_TTS[Montant_Note_Credit_HT])
Nombre_Note_Credit	DISTINCTCOUNT(Remboursements_TTS[Numero_Note_De_Credit])

### 3.2.3 Elaboration des analyses des besoins de suivi

Power BI permet, en utilisant les visualisations, de réaliser des analyses sur les données à travers les opérations de :

- **Roll-up** : c'est le fait de monter à travers les différents niveaux d'une hiérarchie.
- **Drill-down** : c'est le fait de descendre à travers les différents niveaux d'une hiérarchie.
- **Slice** : il s'agit d'une sélection en fixant la valeur d'un attribut d'une dimension.
- **Dice** : c'est un slice sur deux ou plusieurs dimensions.

Les tableaux suivants détaillent les analyses possibles sur les différents besoins de suivi précédemment identifiés :

- Suivi des réclamations

**Tableau 30 : Suivi des réclamations (Nous-même, 2020)**

Besoin	Axes d'analyse	Valeur analysée	Opérations d'analyse
Connaître le nombre de réclamations de chaque type de réclamations	Type de réclamation	Nombre_ Réclamations	<p><b>Slice</b> en se fixant sur un type de réclamation particulier.</p> <p><b>Dice</b> en fixant un type de réclamation particulier et un attribut de KD et/ou de Produit et/ou de date particulier.</p> <p><b>Drill-down ou Roll-Up :</b></p> <ul style="list-style-type: none"> <li>➤ Suivant les hiérarchies suivantes :</li> </ul> <p><b>Dimension Produit :</b> Catégorie-&gt; Subcatégorie-&gt;Marque-&gt; SKU</p> <p><b>Dimension KD :</b> Région-&gt; Wilaya-&gt; Nom KD.</p> <p><b>Dimension Date :</b> Année-&gt; Trimestre -&gt; Mois-&gt; Date.</p>
	Produit		
	KD		
	Date		
Connaître le nombre de réclamations par statut	Statut de réclamation	Nombre_ Réclamations	<p><b>Slice</b> en se fixant sur un statut de réclamation particulier.</p> <p><b>Dice</b> en fixant un statut de réclamation particulier et un attribut de KD et/ou de Produit et/ou de date particulier.</p> <p><b>Drill-down ou Roll-Up :</b></p> <ul style="list-style-type: none"> <li>➤ Suivant les hiérarchies suivantes :</li> </ul> <p><b>Dimension Produit :</b> Catégorie-&gt; Subcatégorie-&gt;Marque-&gt; SKU</p> <p><b>Dimension KD :</b> Région-&gt; Wilaya-&gt; Nom KD.</p> <p><b>Dimension Date :</b> Année-&gt; Trimestre -&gt; Mois-&gt; Date.</p>
	Produit		
	KD		
	Date		
Connaître le montant réclamé	KD	Montant_ Réclamation	<p><b>Slice</b> en se fixant sur un attribut de la hiérarchie Produit.</p> <p><b>Dice</b> en fixant aussi un attribut de KD et/ou de date particulier.</p> <p><b>Drill-down ou Roll-Up :</b></p> <ul style="list-style-type: none"> <li>➤ Suivant les hiérarchies suivantes :</li> </ul> <p><b>Dimension Produit :</b> Catégorie-&gt; Subcatégorie-&gt;Marque-&gt; SKU</p> <p><b>Dimension KD :</b> Région-&gt; Wilaya-&gt; Nom KD.</p> <p><b>Dimension Date :</b> Année-&gt; Trimestre -&gt; Mois-&gt; Date.</p>
	Produit		
	Date		
Connaître la quantité réclamée	KD	Quantité réclamée	<p><b>Slice</b> en se fixant sur un attribut de la hiérarchie Produit.</p> <p><b>Dice</b> en fixant aussi un attribut de KD et/ou de date particulier.</p> <p><b>Drill-down ou Roll-Up :</b></p> <ul style="list-style-type: none"> <li>➤ Suivant les hiérarchies suivantes :</li> </ul> <p><b>Dimension Produit :</b> Catégorie-&gt; Subcatégorie-&gt;Marque-&gt; SKU</p> <p><b>Dimension KD :</b> Région-&gt; Wilaya-&gt; Nom KD.</p> <p><b>Dimension Date :</b> Année-&gt; Trimestre -&gt; Mois-&gt;Date.</p>
	Produit		
	Date		

- Suivi de la GSV

**Tableau 31 : Suivi de la GSV (Nous-même, 2020)**

Besoin	Axes d'analyse	Valeur analysée	Opérations d'analyse
Connaître la contribution d'une catégorie, sous-catégorie, marque ou SKU à la GSV.	Produit	GSV	<p><b>Slice</b> en se fixant sur un attribut de la hiérarchie Produit.</p> <p><b>Drill-down ou Roll-Up :</b></p> <ul style="list-style-type: none"> <li>➤ Suivant les hiérarchies suivantes :</li> </ul> <p><b>Dimension Produit :</b> Catégorie-&gt; Subcatégorie-&gt;Marque-&gt;SKU</p> <p><b>Dimension Date :</b> Année-&gt; Trimestre -&gt; Mois-&gt; Date.</p>
Connaître la contribution de chaque région, wilaya ou KD à la GSV.	KD	GSV	<p><b>Slice</b> en se fixant sur un attribut de la hiérarchie KD.</p> <p><b>Drill-down ou Roll-Up :</b></p> <ul style="list-style-type: none"> <li>➤ Suivant les hiérarchies suivantes :</li> </ul> <p><b>Dimension KD :</b> Région-&gt; Wilaya-&gt; Nom KD.</p> <p><b>Dimension Date :</b> Année-&gt; Trimestre -&gt; Mois-&gt; Date.</p>

- Suivi des ventes primaires

**Tableau 32 : Suivi des ventes primaires (Nous-même, 2020)**

Besoin	Axes d'analyse	Valeur analysée	Opérations d'analyse
Connaître le nombre de commandes réalisées, celles confirmées ou bien livrées.	Produit	Nombre_ commande	<p><b>Slice</b> en se fixant sur un statut du customer service particulier.</p> <p><b>Dice</b> en fixant aussi un attribut de Produit et/ou KD et/ou de date particulier.</p> <p><b>Drill-down ou Roll-Up :</b></p> <ul style="list-style-type: none"> <li>➤ Suivant les hiérarchies suivantes :</li> </ul> <p><b>Dimension Produit :</b> Catégorie-&gt; Subcatégorie-&gt;Marque-&gt; SKU</p> <p><b>Dimension KD :</b> Région-&gt; Wilaya-&gt; Nom KD.</p> <p><b>Dimension Date :</b> Année-&gt;Trimestre -&gt;Mois-&gt; Date.</p>
	KD		
	Date		
	Statut_Customer_Service		
Connaître le nombre de commandes bloquées chez les finances.	Produit	Nombre_ commande	<p><b>Slice</b> en se fixant sur un statut financier.</p> <p><b>Dice</b> en fixant aussi un attribut de Produit et/ou KD et/ou de date particulier.</p> <p><b>Drill-down ou Roll-Up :</b></p> <ul style="list-style-type: none"> <li>➤ Suivant les hiérarchies suivantes :</li> </ul> <p><b>Dimension Produit :</b> Catégorie-&gt; Subcatégorie-&gt;Marque-&gt; SKU</p> <p><b>Dimension KD :</b> Région-&gt; Wilaya-&gt; Nom KD.</p> <p><b>Dimension Date :</b> Année-&gt;Trimestre -&gt;Mois-&gt; Date.</p>
	KD		
	Date		
	Statut_Financier		
Connaître le nombre de commandes ayant rencontré un retard dans leur traitement grâce à CCFOTReasons.	Produit	Nombre_ commande	<p><b>Slice</b> en se fixant sur une CCFOTReason particulière.</p> <p><b>Dice</b> en fixant aussi un attribut de Produit et/ou KD et/ou de date particulier.</p> <p><b>Drill-down ou Roll-Up :</b></p> <ul style="list-style-type: none"> <li>Suivant les hiérarchies suivantes :</li> </ul> <p><b>Dimension Produit :</b> Catégorie-&gt; Subcatégorie-&gt;Marque-&gt; SKU</p> <p><b>Dimension KD :</b> Région-&gt; Wilaya-&gt; Nom KD.</p> <p><b>Dimension Date :</b> Année-&gt;Trimestre -&gt;Mois-&gt; Date.</p>
	KD		
	Date		
	CCFOTReasons		

- **Suivi des remboursements TTS**

**Tableau 33 : Suivi des remboursements TTS (Nous-même, 2020)**

Besoin	Axes d'analyse	Valeur analysée	Opérations d'analyse
Connaître le nombre de notes de crédit adressées à chaque KD	<b>KD</b>	Nombre _Note _Credit	<b>Slice</b> en se fixant sur un Nom_KD particulier. <b>Dice</b> en fixant aussi un TTS particulier et/ou un attribut de la date. <b>Roll-Up</b> : Nom KD-> Wilaya-> Region <b>Drill-Down</b> : Année-> Trimestre -> Mois-> Date.
	<b>Date</b>		
Connaître la contribution de chaque région au montant des notes de crédit.	<b>KD</b>	Montant _Note _Credit	<b>Slice</b> en se fixant sur une région particulière. <b>Dice</b> en fixant aussi un TTS particulier et/ou un attribut de la date. <b>Drill down</b> : <b>Dimension KD</b> : Région-> Wilaya-> Nom KD. <b>Dimension Date</b> : Année-> Trimestre ->Mois->Date.
	<b>Date</b>		
Connaître la contribution de chaque TTS au montant total des remboursements.	<b>TTS</b>	Montant _Note_ Credit	<b>Slice</b> en se fixant sur un TTS particulier. <b>Dice</b> en fixant aussi un attribut de KD particulier et/ou un attribut de la date. <b>Drill down</b> : <b>Dimension KD</b> : Région-> Wilaya-> Nom KD. <b>Dimension Date</b> : Année-> Trimestre ->Mois->Date.
	<b>KD</b>		
	<b>Date</b>		

#### 4. Exemples de rapports types

Dans cette partie, nous présentons des exemples de rapports pour les datamarts Réclamations et Commandes.

- **Rapport Réclamations**

Ci-dessous un exemple de rapport pour les réclamations. Il traduit certains des besoins de suivi des réclamations.

On peut y voir deux graphiques représentatifs des montants des réclamations, selon les SKUs et selon les KDs.

Ainsi que trois graphiques représentatifs du nombre de réclamations reçues, selon leur statut, leur type, et la catégorie de produits concernés.

- **Rapport Commandes**

Ceci est un exemple d'un rapport en deux pages, la première traduisant les besoins de suivi de la GSV et la seconde traduisant les besoins de suivi des ventes.

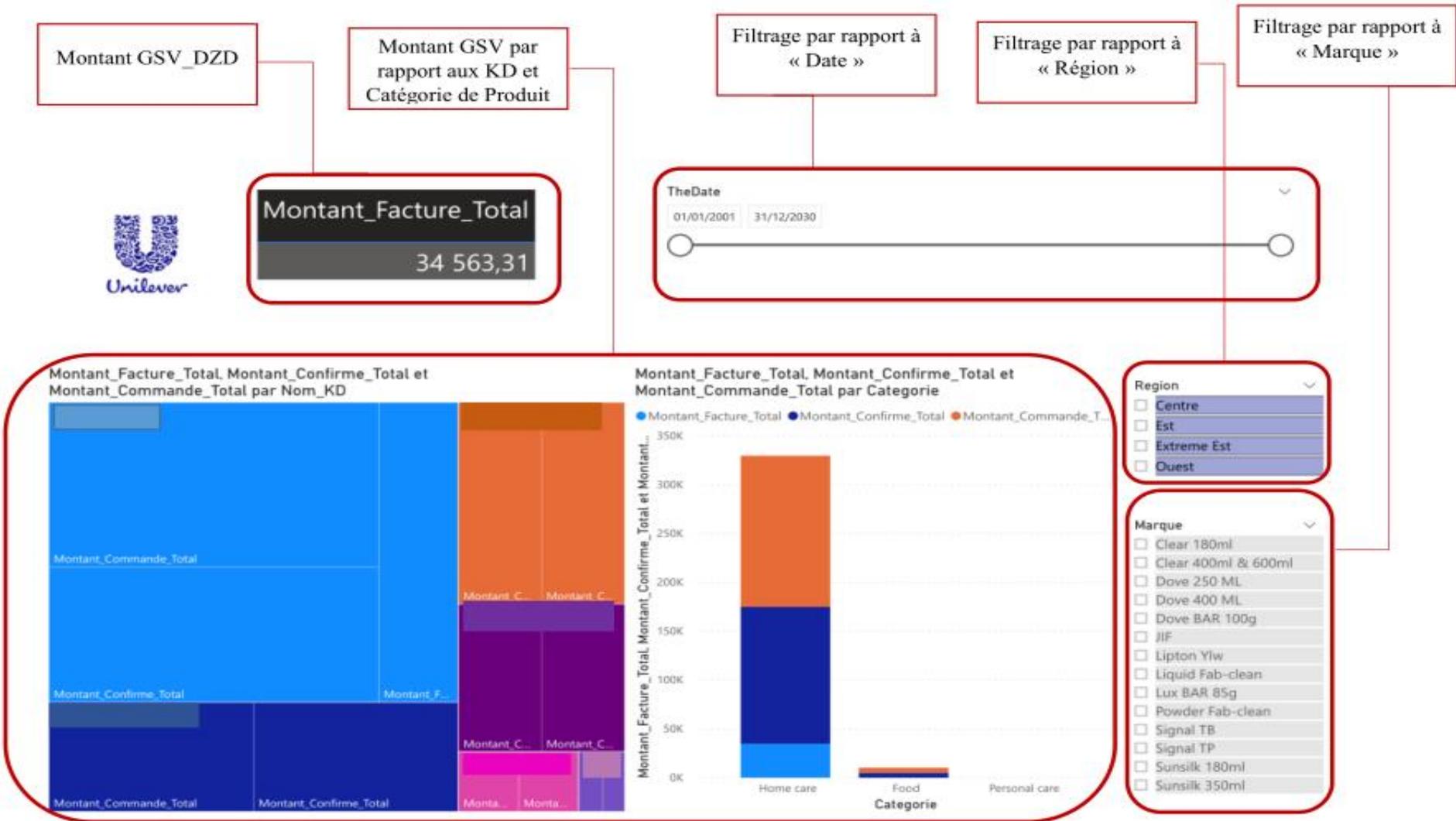
On peut y voir des figures représentatives des montants facturés selon le KD et la catégorie du produit, ainsi que des figures représentatives des quantités commandées, confirmées et facturée selon les KDs, les catégories, le statut Customer Service et le statut financier.

Pour chacun de ces rapports, nous avons inséré trois slicers : un pour la date, un pour la région et un pour la marque du produit, rendant ainsi possible le choix d'une ou de plusieurs dates/ régions/ marques particulières par rapport auxquelles les graphiques changeront.

Par conséquent, une fois le squelette d'un rapport réalisé, et s'il s'agit d'une forme de rapport récurrente, il suffit de changer les paramètres des slicers pour l'adapter au besoin du moment.



Figure 42 : Rapport Réclamations



**Figure 43 : Rapport Commande - GSV**

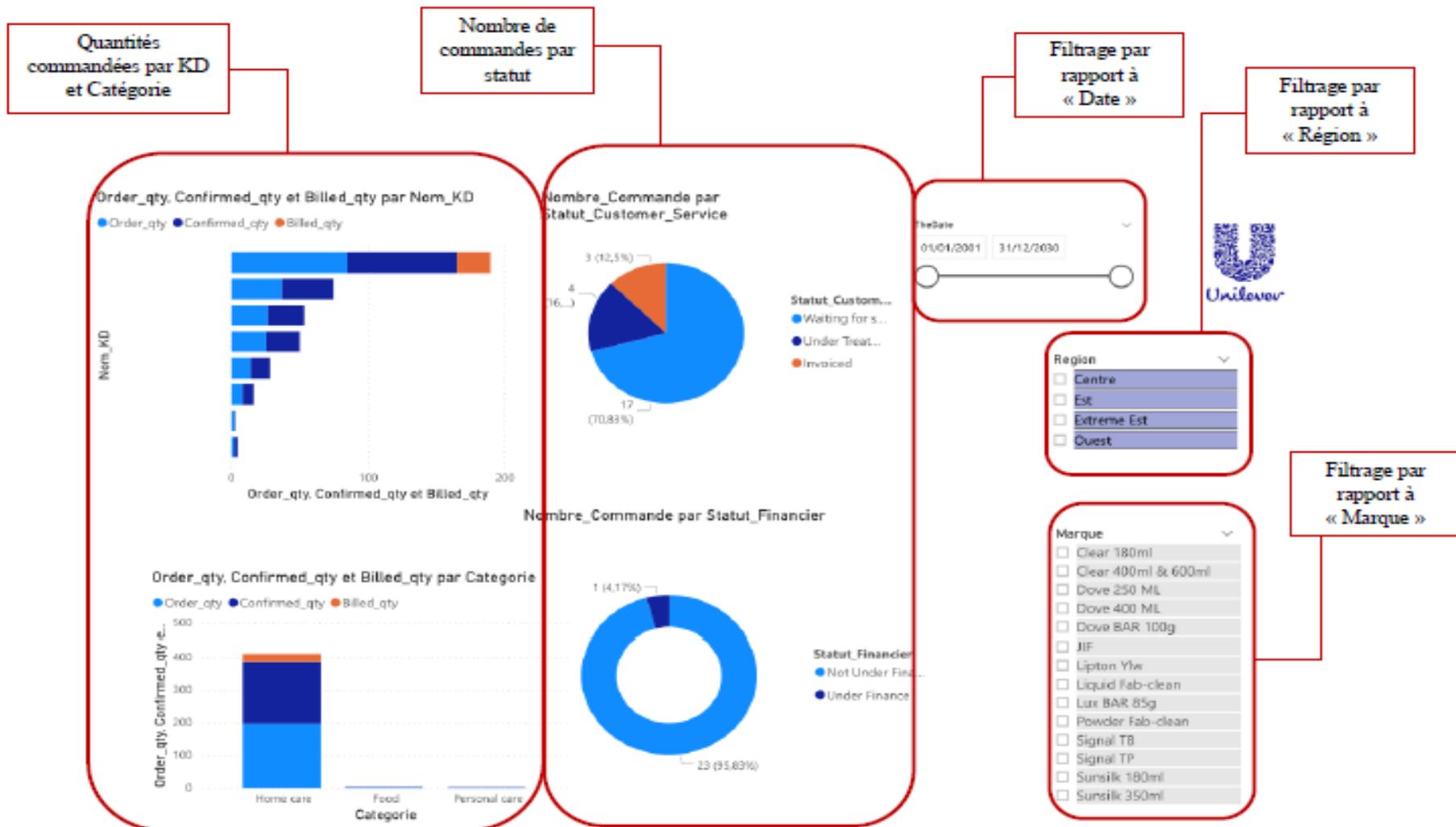


Figure 44 : Rapport Commandes - Les Ventes

## 5. Pistes d'amélioration

Comme complément à notre travail, nous proposons quelques pistes d'amélioration :

- Finalisation du datamart global Order-to-cash à travers la réalisation des datamarts restants afin de pouvoir réaliser des analyses plus complexes.
- Automatiser certaines requêtes SAP pour la collecte de données ; chaque requête ne permet d'obtenir des informations que sur un seul SKU, il faut donc un temps considérable pour les collecter.
- Data mining : afin de construire des modèles permettant d'acquérir de nouvelles informations sur les données récoltées.

## 6. Conclusion

Dans ce chapitre, nous avons expliqué les étapes que nous avons suivies pour la réalisation de notre solution.

Nous avons commencé par rappeler quels datamarts nous avons choisi de réaliser et indiqué les changements que nous y avons apportés. Ensuite, nous sommes passées à la phase d'alimentation où nous avons détaillé les processus ETL.

Par la suite, nous sommes passées à la phase d'exploitation, et ce, en créant les mesures et colonnes dérivées liées aux tables de faits ainsi qu'en présentant les analyses OLAP.

Afin de clore notre projet BI, nous avons présenté des exemples de rapports pour les datamart « Réclamations » et « commandes ». Enfin, nous avons proposé quelques pistes d'amélioration susceptibles de renforcer l'apport de notre travail.

Nous n'avons pas encore l'avis des utilisateurs sur les résultats de notre projet, et ce de par les difficultés communicationnelles en ces temps de pandémie et à cause de la charge de travail conséquence qui ne leur a pas permis d'évaluer notre travail.

Nous pouvons néanmoins affirmer que désormais ils ont accès à davantage d'axes pour analyser leurs données et qu'ils peuvent les analyses rapidement par rapport à plusieurs axes à la fois, et ce grâce à la technologie OLAP. Ils peuvent également créer des modèles de rapports d'une manière simple.

# Conclusion générale

L'objectif de notre travail est la réalisation et l'implémentation d'un système décisionnel permettant au service d'améliorer la qualité et de réduire le temps de réalisation de ses rapports, ainsi que de réaliser des analyses plus poussées.

Nous avons entamé notre projet en réalisant une étude de l'existant du service, afin de cerner les dysfonctionnements à résoudre, pour ce faire nous avons identifié le processus suivi par le personnel du service, leurs activités de reporting, leurs KPIs et nous avons également audité la qualité de leurs données.

A l'issue de cette étude de l'existant, nous avons pu formuler un principal dysfonctionnement, lié aux difficultés en termes de temps et d'analyses que rencontre le service dans son activité de reporting, ce qui nous a conduit à formuler notre problématique de la manière suivante :

**« Comment réduire le temps de création de rapports, tout en assurant un meilleur visuel et en offrant plus de possibilités d'analyse aux utilisateurs finaux ? »**

Pour remédier à ce dysfonctionnement, nous avons proposé la mise en place d'une solution BI, consistant en la conception et l'implémentation de datamarts. Lors de la conception de cette solution, nous avons utilisé la méthodologie de Kimball, où nous avons d'abord identifié et analysé les besoins de suivi du personnel du service, et ce en utilisant deux méthodes :

- Une méthode orientée source basée sur l'étude des documents du service.
- Une méthode orientée utilisateur basée sur la conduite d'entretiens.

Une fois ces besoins identifiés, nous avons suivi les 4 principales étapes de la modélisation dimensionnelle :

- 1) Découpage du processus en sous-processus, évaluation de la faisabilité et de l'importance de ces sous-processus ainsi que l'affectation des besoins précédemment identifiés aux sous-processus.
- 2) Identification du grain pour chaque sous-processus, afin de déterminer les tables de faits à réaliser.
- 3) Identification des dimensions et de leurs attributs.
- 4) Identification des mesures de base et dérivées.

Nous avons achevé cette modélisation dimensionnelle par le choix de l'architecture en bus de Kimball, qui nous a permis de lier les datamarts par des dimensions conformes. Une fois les datamarts modélisés, nous nous sommes occupées du stockage en créant des tables SQL pour les tables de faits et dimensions.

Une fois la conception de la solution terminée, nous sommes passées à son implémentation, qui s'est principalement basée sur l'ETL :

- Extract: Extraction des données à partir des fichiers sources.
- Transform: Application de transformation sur ces données afin de les rendre adaptées au format des datamarts.
- Load: Chargement des données transformées dans les datamarts.

Nous avons par la suite procédé à la création de mesures et détaillé les différentes analyses offertes par l'utilisation de cubes multidimensionnels sur les datamarts obtenus. Pour finir, nous avons présenté un exemple de rapport pour la gestion des réclamations et des commandes.

Cette solution que nous avons proposée a eu un impact sur l'entreprise, qui peut être résumé par les points suivants :

- Modélisation de plusieurs datamarts, faisant ainsi gagner à l'IT le temps qui y aurait été consacré lors de leur future implémentation ;
- Implémentation de trois datamarts contenant trois dimensions conformes pouvant être utilisées par la suite lors de l'élargissement du datawarehouse ;
- Élargissement des possibilités de l'activité de reporting, pouvant maintenant être effectuée sur des bases solides à l'aide d'outils de reporting ;
- Signalement de plusieurs problèmes liés à l'intégrité des données, découverts lors de l'exécution de notre ETL.

Ce projet nous a également permis d'acquérir des connaissances et des compétences personnelles, telles que :

- ❖ L'acquisition de connaissances dans la manipulation de logiciels tels que SSIS, Power BI ;
- ❖ L'acquisition de soft-skills telles que la gestion du temps et la conduite d'entretiens ;
- ❖ L'application des connaissances théoriques accumulées lors de nos études ainsi que la découverte de la réalité du monde du travail.

# Bibliographie

## Ouvrages :

1. **BALLARD, Chuck.** *Dimensional Modeling : In a Business intelligence Environment.* s.l. : USA: Vervante, 2006, 668p. ISBN 978-0738496443.
2. **BERRY, Michael J.A.** *Data Mining Techniques.* 2nde Edition. USA : Wiley, 2004, 643p. ISBN 471-47064-3.
3. **GODIN.R, DESROSIERS.C.** *Bases de données de haute performance Les entrepôts de données et l'analyse de données.* France: ETS, Dpt de génie logiciel et des TI, 2011, 61p : s.n.
4. **INMON, William H.** *Building the Data Warehouse.* 3ème édition. USA: Wiley, 2005, 576p. 978-0764599446 : s.n.
5. **KIMBALL, Ralph et ROSS, Mergy.** *The Data Warehouse Lifecycle Toolkit, 3eme edition.* USA : Wiley, 2013, 608 p. 978-1-118-53080-1.
6. **KIMBALL, Ralph et ROSS, Mergy.** *The Data Warehouse Lifecycle Toolkit 2eme édition.* USA: Wiley, 2007, 673p. 978-0470149775 : s.n.
7. **SHARMA, Subham et GOMES, Kévin et ACIER, Maximilien.** *Business Intelligence | State of the art.* s.l. : France: EPITA, 2015, 55p.
8. **ZOUAGHI, Iskander.** *Cours Business Intelligence ENP 2019.* Alger : ENP: Département Génie Industriel, 2019-2020, 248p.

## Mémoires et thèses :

9. **Diab, BENDIAB Rima et SATTOUF.** Contribution à l'élaboration d'un plan de production d'une nouvelle ligne industrielle par la modélisation mathématique - Application: Unilever Algérie-. Alger : s.n., 2018. [consulté le 30/04/2020]. Disponible à l'adresse: [https://drive.google.com/drive/folders/19\\_Wz0fgVWPBlvsxBWnXmFT9DITV8i8PZ?fbclid=IwAR3OnfyoS5ehNRzp1e0Jln4i2NKyySDMqT2TazalvSdPdnZSoZCHwto9178](https://drive.google.com/drive/folders/19_Wz0fgVWPBlvsxBWnXmFT9DITV8i8PZ?fbclid=IwAR3OnfyoS5ehNRzp1e0Jln4i2NKyySDMqT2TazalvSdPdnZSoZCHwto9178).
10. **ZENNIR, Ines.** The impact of green supply chain management on. Oran : s.n., 2019. [consulté le 30/04/2020]. Disponible à l'adresse : <https://drive.google.com/file/d/1GJjrhra6SEoMkCte90BCSaYIjVbrr7QX/view?usp=sharing>.

## Sites web et documents en ligne :

11. **A propos d'Unilever | Unilever France. Unilever.** [En ligne] [Citation : 28 Avril 2020.] <https://www.unilever.fr/notre-entreprise/>.
12. **Adeoinformatique.** etl-process. [En ligne] 2019. [Citation : 26 Juin 2020.] <https://www.adeoinformatique.com/nouveaute-extract-transform-load-etl/etl-process/>.
13. **Alliedmarketresearch.** Global FMCG Market Expected to Reach \$15,361.8 Billion by 2025 | AMR. [En ligne] [Citation : 04 Juin 2020.] <https://www.alliedmarketresearch.com/press-release/fmcg-market.html>.

- 14. BELLACHIA, Abdelghani.** *Data Warehousing and OLAP Technology*. Washington DC : George Washington University, 48p. [Citation : 02 Juillet 2020]. Disponible sur : [https://www2.seas.gwu.edu/~bell/csci243/lectures/data\\_warehousing.pdf](https://www2.seas.gwu.edu/~bell/csci243/lectures/data_warehousing.pdf).
- 15. Bizfluent.** Definition of Business Reports. [En ligne] 2018. [Citation : 5 Juillet 2020.] <https://bizfluent.com/info-7758187-definition-business-reports.html>.
- 16. BRESLIN, Mary.** Mary Breslin Data Warehousing Battle of the Giants: Comparing the Basics of the Kimball and Inmon Models. *BUSINESS INTELLIGENCE JOURNAL [En ligne]*. 2004 , 20p, Disponible sur : [https://pdfs.semanticscholar.org/c80f/8aaea5bf58846b0125b460401fed8230c2d2.pdf?\\_ga=2.249729796.309338956.1595011795-918448827.1592318981](https://pdfs.semanticscholar.org/c80f/8aaea5bf58846b0125b460401fed8230c2d2.pdf?_ga=2.249729796.309338956.1595011795-918448827.1592318981).
- 17. Bukhantsov.** What is Data Vault? [En ligne] [Citation : 26 Juin 2020.] <http://bukhantsov.org/2012/04/what-is-data-vault/>.
- 18. Captio.** What is a Business Intelligence (BI) project? *Captio*. [En ligne] 2015. <https://www.captio.com/blog/what-is-a-businss-intelligence-project>.
- 19. CHAFKI.S, DESROSIERS.C.** MTI820 – Entrepôts de données et intelligence d'affaires Architecture des entrepôts de données. [En ligne] 2011. [Citation : 25 Juin 2020.], Disponible sur : [https://cours.etsmtl.ca/mti820/public\\_docs/acetates/MTI820-Acetates-ArchitectureDW\\_1pp.pdf](https://cours.etsmtl.ca/mti820/public_docs/acetates/MTI820-Acetates-ArchitectureDW_1pp.pdf).
- 20. CHRIS, Morgan.** SSIS is terrific for all your ETL needs. *Trustradius*. [Online] 2016. [Cited: Septembre 9, 2020.] <https://www.trustradius.com/reviews/sql-server-integration-services-2016-05-25-19-17-48>.
- 21. CompaniesHistory.** Unilever. [En ligne] [Citation : 2020 Avril 28.] <https://www.companieshistory.com/unilever/>.
- 22. Compta-Facile.** Quels sont les différents tableaux de bord ? [En ligne] 2016. [Citation : 10 Juillet 2020.] <https://www.compta-facile.com/differents-tableaux-de-bord/>.
- 23. Critizr | Leader de la satisfaction et l'expérience client pour le retail. Critizr.** [En ligne] 2016. [Citation : 1 Juin 2020.] <https://business.critizr.com/blog/etude-qualitative-tout-savoir-sur-entretien-individuel>.
- 24. CROZAT, Stéphane.** *Introduction aux data warehouses : la modélisation en étoile*. Compiègne : UTC , 2017, 31 pages . [Consulté le 03/07/2020 ]. Disponible sur : <https://stph.scenari-community.org/bdd/dwh1.pdf> : s.n.
- 25. DARMONT Jérôme et MARCEL, Patrick.** Entrepôts de données et OLAP, analyse et décision dans l'entreprise. *Researchgate*. [En ligne] 2017. [Citation : 22 Juin 2020.] , Disponible sur : [https://www.researchgate.net/publication/316182904\\_Entrepots\\_de\\_donnees\\_et\\_OLAP\\_analyse\\_et\\_decision\\_dans\\_l'entreprise](https://www.researchgate.net/publication/316182904_Entrepots_de_donnees_et_OLAP_analyse_et_decision_dans_l'entreprise).
- 26. data-flair.** Pros and Cons of Power BI - The Bright & the Dull side of visualization suite. *data-flair*. [Online] 2018. [Cited: Septembre 9, 2020.] <https://data-flair.training/blogs/power-bi-advantages-and-disadvantages/>.
- 27. Desai, Shaishav.** 6 Reasons To Use Power BI AS Business Intelligence Solution. *Clairion Technologie*. [En ligne] [Citation : 21 08 2020.] , Disponible sur : <https://www.clariontech.com/blog/6-reasons-to-use-power-bi-as-your-business-intelligence-solution>.

- 28. DESAI, Shilpa.** What is SSIS? Its advantages and disadvantages. *Sarjen*. [Online] 2016. [Cited: Septembre 9, 2020.] <https://www.trustradius.com/reviews/sql-server-integration-services-2016-05-25-19-17-48>.
- 29. Dikken.** Snowflake Schema vs Star Schema Difference and Comparison | Dikken. [En ligne] 2014. [Citation : 15 Juin 2020.] [https://www.dikken.com/difference/Snowflake\\_Schema\\_vs\\_Star\\_Schema](https://www.dikken.com/difference/Snowflake_Schema_vs_Star_Schema).
- 30. DURCEVIC, Sandra.** Business Intelligence Reporting – Capitalize on Your BI Reports. *Datapine*. [En ligne] 2020. [Citation : 23 Juillet 2020.] <https://www.datapine.com/blog/business-intelligence-reporting/>.
- 31. Dwilounge.** Schemas implemented in Datawarehousing. [En ligne] 2015. [Citation : 26 Juin 2020.] <https://dwilounge.wordpress.com/2015/02/01/schemas-implemented-in-datawarehousing/>.
- 32. DWgeek.** Types of Dimension Tables in a Data Warehouse. [En ligne] 2018. [Citation : 22 Juin 2020.] <https://dwgeek.com/types-of-dimension-tables-data-warehouse.html/>.
- 33. Euromonitor.** Beauty and Personal Care in Algeria | Market Research Report | Euromonitor. Londres: Euromonitor, 2020. [En ligne] [Citation : 14 Juin 2020.] <https://www.euromonitor.com/beauty-and-personal-care-in-algeria/report>.
- 34. FERNANDEZ, Alain.** Le Reporting expliqué, rapports d'activité. *piloter*. [En ligne] [Citation : 23 Juillet 2020.] <https://www.piloter.org/business-intelligence/reporting.htm>.
- 35. forbes.** Can Procter & Gamble's Revenue Cross \$72 billion by 2021? [En ligne] 2019. [Citation : 4 Juin 2020.] <https://www.forbes.com/sites/greatspeculations/2019/10/31/can-procter--gambles-revenue-cross-72-billion-by-2021/#461849d1aee9>.
- 36. geeksforgeeks.** Fact Constellation in datawarehouse modelling. [En ligne] [Citation : 24 Juin 2020.] <https://www.geeksforgeeks.org/fact-constellation-in-data-warehouse-modelling/>.
- 37. Graber, Roy.** JBS reports record net revenue in 2019 | WATTAgNet. [En ligne] 2020. [Citation : 4 Juin 2020.] [https://www.wattagnet.com/articles/39910-jbs-reports-record-net-revenue-in-2019?v=preview#:~:text=Company's%20net%20revenue%20for%20past%20year%20improved%2012.6%25%20when%20compared%20to%202018&text=Net%20revenue%20for%20JBS%20reached,billion\)%20for%20th](https://www.wattagnet.com/articles/39910-jbs-reports-record-net-revenue-in-2019?v=preview#:~:text=Company's%20net%20revenue%20for%20past%20year%20improved%2012.6%25%20when%20compared%20to%202018&text=Net%20revenue%20for%20JBS%20reached,billion)%20for%20th).
- 38. GRIM, Yazid et BLAIN, Fleur-Anne.** ETL, les questions à se poser. [En ligne] 2014. [Citation : 24 Juin 2020.] <https://grim.developpez.com/articles/concepts/etl/>.
- 39. Guru99.** Star and Snowflake Schema in Datawarehouse. [En ligne] [Citation : 15 Juin 2020.] <https://www.guru99.com/star-snowflake-data-warehousing.html>.
- 40. IBM.** Fact tables and entities. [En ligne] 2017. [Citation : 22 Juin 2020.] [https://www.ibm.com/support/knowledgecenter/SS9UM9\\_9.1.0/com.ibm.datatools dimensional.ui.doc/topics/c\\_dm\\_fact\\_tables.html](https://www.ibm.com/support/knowledgecenter/SS9UM9_9.1.0/com.ibm.datatools dimensional.ui.doc/topics/c_dm_fact_tables.html).
- 41. IBM.** Step 1: Identify business process requirements. *IBM*. [En ligne] 2014. [Citation : 25 Juillet 2020.] [www.ibm.com/support/knowledgecenter/es/ss9um9\\_9.1.0/com.ibm.datatools dimensional.ui.doc/topics/c\\_dm\\_design\\_cycle\\_1\\_idbusprocreqs.html](http://www.ibm.com/support/knowledgecenter/es/ss9um9_9.1.0/com.ibm.datatools dimensional.ui.doc/topics/c_dm_design_cycle_1_idbusprocreqs.html).

- 42. IBM.** Step 2: Identify the grain. [En ligne] 2014. [Citation : 24 Juin 2020.] [www.ibm.com/support/knowledgecenter/es/ss9um9\\_9.1.0/com.ibm.datatools.dimensionai.ui.doc/topics/c\\_dm\\_design\\_cycle\\_2\\_idgrain.html](http://www.ibm.com/support/knowledgecenter/es/ss9um9_9.1.0/com.ibm.datatools.dimensionai.ui.doc/topics/c_dm_design_cycle_2_idgrain.html).
- 43. IBM.** Step 3: Identify the dimensions. [En ligne] 2014. [Citation : 24 Juin 2020.] [www.ibm.com/support/knowledgecenter/es/ss9um9\\_9.1.0/com.ibm.datatools.dimensionai.ui.doc/topics/c\\_dm\\_design\\_cycle\\_3\\_iddim.html](http://www.ibm.com/support/knowledgecenter/es/ss9um9_9.1.0/com.ibm.datatools.dimensionai.ui.doc/topics/c_dm_design_cycle_3_iddim.html).
- 44. IBM.** Step 4: Identify the measures. *IBM*. [En ligne] 2014. [Citation : 24 Juin 2020.] [www.ibm.com/support/knowledgecenter/es/ss9um9\\_9.1.0/com.ibm.datatools.dimensionai.ui.doc/topics/c\\_dm\\_design\\_cycle\\_4.html](http://www.ibm.com/support/knowledgecenter/es/ss9um9_9.1.0/com.ibm.datatools.dimensionai.ui.doc/topics/c_dm_design_cycle_4.html).
- 45. IBM.** What is ETL (Extract, Transform, Load)? | IBM. *IBM*. [En ligne] 2020. [Citation : 24 Juin 2020.] <https://www.ibm.com/cloud/learn/etl>.
- 46. Insidebigdata.** Is Data vault Modeling a Good Choice for Your Organisation? . [En ligne] 2017. [Citation : 22 Juin 2020.] <https://insidebigdata.com/2017/07/28/data-vault-modeling-good-choice-organization/#:~:text=Data%20Vault%20architecture%20is%20an,data%20coming%20into%20the%20database..>
- 47. Izenda.** What is Ad Hoc Reporting? Definitions, Benefits, Use Cases. [En ligne] 2019. [Citation : 5 Juillet 2020.] <https://www.izenda.com/ad-hoc-reporting/>.
- 48. KIMBALL, Ralph.** Fact Table Structure | Kimball Dimensional Modeling Techniques. *Kimballgroup*. [En ligne] 2007. [Citation : 22 Juin 2020.] <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/fact-table-structure/>.
- 49. KIMBALL, Ralph.** Keep to the Grain in Dimensional Modeling. *Kimballgroup*. [En ligne] 2007. [Citation : 24 Juin 2020.] <https://www.kimballgroup.com/2007/07/keep-to-the-grain-in-dimensional-modeling/>.
- 50. KIMBALL, Ralph.** Period Snapshot Fact Tables | Kimball Group. [En ligne] [Citation : 22 Juin 2020.] <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/periodic-snapshot-fact-table/>.
- 51. KPI.** What is a Key Performance Indicator (KPI)? *KPI*. [En ligne] [Citation : 22 Juillet 2020.] <https://kpi.org/KPI-Basics>.
- 52. Macrotrends.** Unilever Revenue 2006-2019 | UL | Macrotrends. [En ligne] 2020. [Citation : 6 Juin 2020.] <https://www.macrotrends.net/stocks/charts/UL/unilever/revenue>.
- 53. Mactrotrends.** PepsiCo Revenue 2006-2020 | PEP | MacroTrends. [En ligne] 2020. [Citation : 4 Juin 2020.] <https://www.macrotrends.net/stocks/charts/PEP/pepsico/revenue>.
- 54. MARSHALL, Brian.** Canned Reports vs. Ad-hoc Analysis: Which to Use and When [Infographic]. [En ligne] 2020. [Citation : 5 Juillet 2020.] <https://www.us-analytics.com/hyperionblog/canned-reports-vs-ad-hoc-analysis-which-to-use-and-when-infographic>.
- 55. N'GORAN, Angoi Ludovic N'guessan.** COMPRENDRE LES ETAPES DU PROCESSUS BI | SUPINFO, École Supérieure d'Informatique. *Supinfo*. [En ligne] 2016. [Citation : 22 Juillet 2020.] <https://www.supinfo.com/articles/single/3548-comprendre-etapes-processus-bi>.

- 56. NIDA, Fatima.** ETL Process and the Steps for its Implementation - Astera Software. [En ligne] 2020. [Citation : 24 Juin 2020.] <https://www.astera.com/type/blog/etl-process-and-steps/>.
- 57. Reiff, Nathan.** Nestlé's Top Companies and Brands. [En ligne] 2020. [Citation : 4 juin 2020.] <https://www.investopedia.com/articles/markets/122215/top-4-companies-owned-nestle.asp>.
- 58. Software, Unleashed.** Why Business Reporting Is Important For Business Success - Unleashed Software. [En ligne] 2019. [Citation : 5 Juillet 2020.] <https://www.unleashedsoftware.com/blog/why-business-reporting-is-important-for-business-success>.
- 59. Statista.** Revenue of Anheuser-Busch InBev worldwide, 2019 | Statistica. [En ligne] 2020. [Citation : 4 Juin 2020.] <https://www.statista.com/statistics/269112/revenue-of-anheuser-busch-inbev-worldwide/#:~:text=The%20beer%20manufacturer%20Anheuser%20Busch,but%20it%20has%20since%20decreased..>
- 60. STERRET, John.** Data Warehouse: Facts and Measures. [En ligne] 2009. [Citation : 24 Juin 2020.] <https://johnsterrett.com/2009/11/05/data-warehouse-facts-and-measures/>.
- 61. Sutori.** The history of Unilever | Sutori. [En ligne] 2018. [Citation : 28 Avril 2020.] <https://www.sutori.com/story/the-history-of-unilever--k2rovXCayw1Hv5M12sdUdVzb>.
- 62. Techopedia.** What is Business Intelligence Reporting (BI Reporting)? - Definition from Techopedia. *Techopedia*. [En ligne] [Citation : 23 07 2020.] <https://www.techopedia.com/definition/30217/business-intelligence-reporting-bi-reporting>.
- 63. TEHREEM, Naeem.** Datawarehouse concepts : Kimball vs Inmon Approach | Astera. *Astera*. [En ligne] 2020. [Citation : 23 Juin 2020.] <https://www.astera.com/type/blog/data-warehouse-concepts/>.
- 64. TORBEN, Pedersen Bach.** Multidimensionnal Modeling | SpringerLink. [En ligne] 2009. [Citation : 24 Juin 2020.] [https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-39940-9\\_229](https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-39940-9_229).
- 65. Tutorialspoint.** Data Warehousing - OLAP - Tutorialspoint. [En ligne] [Citation : 22 Juin 2020.] [https://www.tutorialspoint.com/dwh/dwh\\_olap.htm](https://www.tutorialspoint.com/dwh/dwh_olap.htm).
- 66. Unilever.** About Unilever | About | Unilever. *Unilever*. [En ligne] [Citation : 09 06 2020.] <https://www.unilever.com/about/who-we-are/about-Unilever/>.
- 67. Usinenouvelle.** Tableau de bord et reporting : instruments de pilotage des managers. [En ligne] 2015. [Citation : 10 Juillet 2020.] <https://www.usinenouvelle.com/expo/guides-d-achat/tableau-de-bord-et-reporting-223>.
- 68. Wibas.** MA.SP 1.2 Specify Measures. [En ligne] 2015. [Citation : 22 Juin 2020.] <https://www.wibas.com/cmimi/masp-12-specify-measures#:~:text=Measures%20can%20be%20either%20base,two%20or%20more%20base%20measures>.
- 69. WORT, Steven.** Uses for SQL Server Integration Services (SSIS). *Logicalread*. [En ligne] [Citation : 21 08 2020.], Disponible sur : <https://logicalread.com/sql-server-ssis-best-uses-w02/#.XyqyTigzbtQ><https://www.c-sharpcorner.com/article/introduction-to-ssis>.
- 70. YESSAD, Lamia et LABIOD, Aissa.** Comparative study of data warehouses modeling approaches: Inmon, Kimball and Data Vault. *Researchgate*. [Online] 2016. [Cited: Septembre 9, 2020.]

[https://www.researchgate.net/publication/312486486\\_Comparative\\_study\\_of\\_data\\_warehouses\\_modeling\\_approaches\\_Inmon\\_Kimball\\_and\\_Data\\_Vault](https://www.researchgate.net/publication/312486486_Comparative_study_of_data_warehouses_modeling_approaches_Inmon_Kimball_and_Data_Vault).

**71. ZenTut.** Fact Table - Definitions, Examples and Four Steps Design by Kimball. [En ligne] [Citation : 22 Juin 2020.] <https://www.zentut.com/data-warehouse/fact-table/>.

# Annexes

## Annexe 1 : Norme BPMN

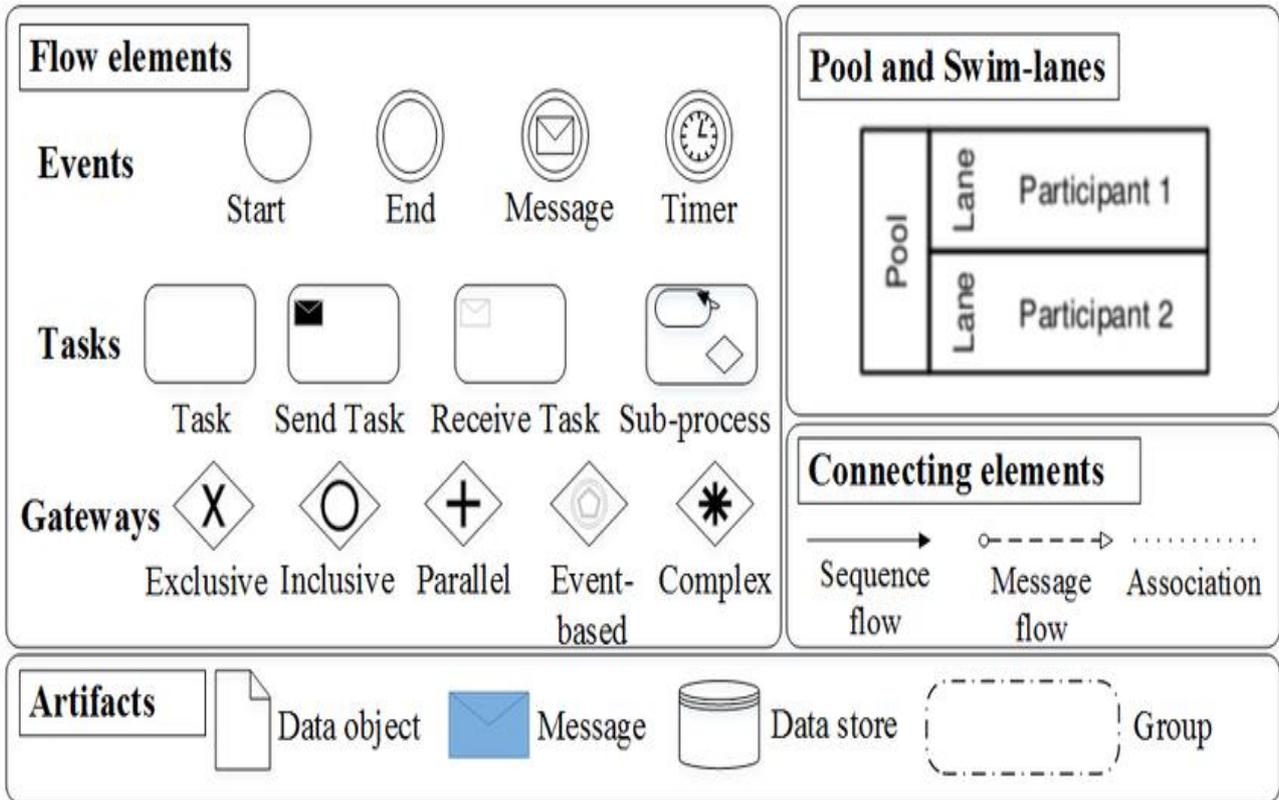


Figure 45 : Norme BPMN

## Annexe 2 : Questionnaires types

Tableau 34 : Questionnaire sur les activités et les besoins du service (Nous-même, 2020)

### **Thème 1 : Questionnaire de l'analyse des activités du service**

#### **Partie 1** : Questions relatives aux processus

**Q1** : Quels sont vos principaux processus ?

**Q2** : Quel est votre rôle dans le processus ? Quelles sont vos tâches habituelles ?

**Q3** : Comment savez-vous que les tâches ont été correctement faites/finies ?

**Q4** : Comment suivez-vous vos activités ?

**Q5** : Est-ce que les processus se déroulent de la manière dont ils devraient ?

#### **Partie 2** : Questions relatives aux activités de reporting

**Q6** : Quelles sont les activités de Reporting que vous faites ?

**Q7** : Quelles sont les activités de Reporting que vous avez besoin de faire ?

**Q8** : Comment réalisez-vous ces rapports avec la configuration actuelle ?

**Q9** : Quelles analyses faites-vous ?

**Q10** : Quelles analyses avez-vous besoin de faire ?

**Q11** : est-ce que la réalisation des rapports prend un temps considérable ?

**Q12** : Quelle est la fréquence de réalisation de chaque rapport ?

**Q13** : Rencontrez-vous des problèmes lors de la réalisation de vos rapports ? Quels sont ces problèmes ?

#### **Partie 3** : Questions relatives aux données

**Q14** : Quelles données utilisez-vous ? A partir de quelle base de données les extrayez-vous ?

**Q15** : Est-ce que les données que vous utilisez correspondent à la réalité ?

**Q16** : Est-ce que les données dont vous avez besoin sont disponibles ? Sinon comment faites-vous pour palier à ce problème ?

**Q17** : Les sources de données sont-elles sûres ?

**Q18** : Les données sont-elles à jour ?

**Partie 4 : Questions relatives aux KPIs**

**Q20 :** Quels-sont vos KPIs ?

**Q21 :** Comment sont-ils calculés ? A partir de quelles données ?

**Q22 :** les données nécessaires au calcul du KPI sont extraites à partir de quels fichiers ?

**Q23 :** pour chaque KPI, quelle est la fourchette de valeurs acceptable ?

**Q24 :** Quelles décisions permettent de prendre ces KPI ?

**Q25 :** Sont-t-il suffisants pour cela ?

**Thème 2 : Questionnaire de l'analyse des besoins**

**Q26 :** Avez-vous besoin de suivre / d'analyser l'un ou plusieurs des thèmes suivants ? Y en a-t-il d'autres que vous avez besoin de suivre / d'analyser ?

- Ventes primaires
- GSV
- Réclamations
- Paiements
- Facturation
- Stocks
- Suivi des remboursements TTS
- Suivi des provisions

**Q27 :** Selon quels axes les analysez-vous ? (Par ex : par rapport aux KD ou aux produits concernés)

**Q28 :** Jusqu'à quel niveau de détail avez-vous besoin de réaliser ces analyses ?

### Annexe 3 : Attribut de dimension

Tableau 35 : Dimension Produit (Nous-même, 2020)

<b>Dimension</b>	<b>Attribut</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Produit</b>	Code_SKU	Code SAP	Code SAP du produit	Entier
	Nom_SKU	Designation / Description	Nom et volume du produit	Chaîne
	Categorie	Category	Catégorie à laquelle appartient le SKU	Chaîne
	Sub_Categorie	Subcategory	Sous-catégorie à laquelle appartient le SKU	Chaîne
	Marque	Brand	Nom de la marque du produit	Chaîne
	Poids_Caisse_en_Tonne	Weight caisse	Poids en tonne d'une caisse	Réel
	GSV_Unit_HT (caisse)	GSV Price (Net Price)	Prix d'une caisse du produit sans taxe	Réel

Tableau 36 : Dimension Client (KD) (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Client (KD)</b>	Code_KD	Customer SAP Code/ KD Code	Code SAP du KD	Entier
	Nom_KD	Company Name / Ship-to-party/ Bill-to party/ KD Name	Nom du KD	Chaîne
	Région	Region	Région où se trouve le KD	Chaîne
	Wilaya	Wilaya	Wilaya où se trouve le KD	Chaîne

Tableau 37 : Dimension TTS (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Types_TTS</b>	TTS	TTS	Nom de la remise	Chaîne

Tableau 38 : Dimension Date (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Date</b>	TheDate	/	Date complète	Date (dd-mm-yyyy)
	TheWeek	/	Numéro de la semaine dans l'année	Entier
	TheMonth	/	Numéro du mois dans l'année	Entier
	TheMonthName	/	Nom du mois	Entier
	TheYear	/	Année	Entier

Tableau 39 : Dimension Entrepôt (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Entrepôt</b>	Nom_Entrepôt	Sloc	Nom de l'entrepôt où est stocké le produit	Chaîne

Tableau 40 : Dimension CCFOT Reasons (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>CCFOT Reasons</b>	CCFOT_Reasons	CCFOT_Reasons	Raison du retard qu'a eu la ligne de commande	Chaîne

Tableau 41 : Dimension Statut\_Finance (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Statut_Finance</b>	Statut_Financier	/	Statut financier de la commande : Bloquée ou non.(Under Finance ou non)	Chaîne

Tableau 42 : Dimension Statut\_Customer\_Service (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Statut_Customer_Service</b>	Statut_Customer_Service	/	Statut de la commande dans le pipeline : In Treatment, Waiting for shipment, Invoiced.	Chaîne

Tableau 43 : Dimension Statut\_Reclamation (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Statut_Reclamation</b>	Statut_Reclamation	Statut	Statut de la réclamation, pouvant être : - In review(en plein traitement) - Resolved - Open (pas encore traitée) - Rejected	Chaîne

Tableau 44 : Dimension Type\_Reclamation (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Type_Reclamation</b>	Type_Reclamation	Claims type	Type de la réclamation, pouvant être - Empty - Excess - Partially Empty - Shortage - Quality Issue	Chaîne

Tableau 45 : Dimension Client\_Secondaire (Nous-même, 2020)

<b>Dimension</b>	<b>Attributs</b>	<b>Appellation d'origine</b>	<b>Signification</b>	<b>Domaine</b>
<b>Client_secondaire</b>	Code_client_secondaire	Client_Code	Code du client secondaire	Entier
	Nom_client_secondaire	Client_Name	Nom du client secondaire	Chaîne
	Secteur_client_secondaire	Sales_Sector	Type de magasin du client secondaire	Chaîne

## Annexe 4 : Les sources de données

A	B	C	D	E	F	G	H	I	J	K	L	M
SaTy	Bill-to party Code	Bill-to party Name	Sales Doc.	PO Number	Order Crea	Pro-Forma C	Delivery Code	DIBI	Bill.Doc.	Doc. Date	Bill. Date	Credit Sta
####	####	####	####	####	####	####	####	####	####	####	####	####
####	####	####	####	####	####	####	####	####	####	####	####	####
####	####	####	####	####	####	####	####	####	####	####	####	####
####	####	####	####	####	####	####	####	####	####	####	####	####

E	F	G	H	I	J	K	L	M	N	O	P	Q	R
PO Number	Order Crea	Pro-Forma C	Delivery Code	DIBI	Bill.Doc.	Doc. Date	Bill. Date	Credit Sta	CCFOT reas	Material	Material Des	Order qty	Cum.cf.qty
####	####	####	####	####	####	####	####	####	####	####	####	####	####
####	####	####	####	####	####	####	####	####	####	####	####	####	####
####	####	####	####	####	####	####	####	####	####	####	####	####	####
####	####	####	####	####	####	####	####	####	####	####	####	####	####

Figure 46 : Todays orders

Statut	N°	KD code	DATE	InvoiceChaine	InvoiceEcrit	INVOICE	SKU Code
Resolved	984	###		01/03/2019	####	####	###
In review	991	###		01/03/2019	####	####	###
In review	991	###		01/03/2019	####	####	###
In review	991	###		01/03/2019	####	####	###
In review	991	###		01/03/2019	####	####	###

InvoiceChaine	InvoiceEcrit	INVOICE	SKU Code	SKU Description	Qty	Claim type	Claim UOM	Region
####	####	####	####	####	0,19	Empty	Cs	West
####	####	####	####	####	1	Empty	PC	Centre_West
####	####	####	####	####	1	Empty	PC	Centre_West
####	####	####	####	####	8	Empty	PC	Centre_West
####	####	####	####	####	6	Empty	Cs	Centre_West

Figure 47 : Claim Tracker

KD NAME	KD code
####	####
####	####
####	####
####	####
####	####

Figure 48 : KD : KD file

Material	Material Number	Net Price	UoM
####	####	####	CS
####	####	####	CS
####	####	####	CS
####	####	####	CS
####	####	####	CS
####	####	####	CS

Figure 49 : Price

Material	Description	category	Brand
####	####	Home care	####
####	####	Home care	####
####	####	Personal care	####
####	####	Personal care	####
####	####	Personal care	####

Figure 50 : Material

DocumentNo	KD CODE	KD NAME	Amount TTC	HT	Date	DESCRIPTION
###	###	###	###	###	###	DC
###	###	###	###	###	###	DC
###	###	###	###	###	###	DC
###	###	###	###	###	###	DC
###	###	###	###	###	###	DC

Figure 51 : CN JUL 19

## Annexe 5 : Requêtes de création des tables SQL

```
USE [BDDPfe]
GO

CREATE TABLE [KD](
    [SK_KD] [int] IDENTITY(1,1) NOT NULL,
    [Nom_KD] [nvarchar](255) NULL,
    [Code_KD] [int] NOT NULL,
    [Wilaya] [nvarchar](255) NULL,
    [Region] [nvarchar](255) NULL
);
```

Figure 52 : Création table KD (Nous-même, 2020)

```
USE [BDDPFE]
GO

CREATE TABLE [dbo].[Produit](
    [SK_Produit] [int] IDENTITY(1,1) NOT NULL,
    [Code_SKU] [int] NOT NULL,
    [Nom_SKU] [nvarchar](255) NULL,
    [Categorie] [nvarchar](255) NULL,
    [Sub_Categorie] [nvarchar](255) NULL,
    [Marque] [nvarchar](255) NULL,
    [Poid_Caisse_en_Tonne] [float] NULL,
    [GSV_Unit_HT] [float] NULL,
    [Price_Unit] [nvarchar](255) NULL
);
```

Figure 53 : Création Table Produit (Nous-même, 2020)

```
USE [BDDPfe]
CREATE TABLE [dbo].[Réclamations_Caisses](
    [FK_KD] [int] NOT NULL,
    [FK_Produit] [int] NOT NULL,
    [FK_Type_Reclamation] [int] NOT NULL,
    [FK_Statut_Reclamation] [int] NOT NULL,
    [Fk_Date_Reclamation] [int] NULL,
    [Quantité_reclamee] [float] NOT NULL,
    [Code_Première_Facture] [nvarchar](255) NULL,
    [Code_Seconde_Facture] [nvarchar](255) NULL,
    [Code_Troisième_Facture] [nvarchar](255) NULL,
    [Code_Quatrième_Facture] [nvarchar](255) NULL,
    [Numero_Reclamation] [int] NOT NULL,
);
```

Figure 54 : Création Table Réclamations\_caisses (Nous-même, 2020)

```

USE [BDDPfe]
CREATE TABLE [dbo].[Réclamations_Pieces](
    [FK_KD] [int] NOT NULL,
    [FK_Produit] [int] NOT NULL,
    [FK_Type_Reclamation] [int] NOT NULL,
    [FK_Statut_Reclamation] [int] NOT NULL,
    [Fk_Date_Reclamation] [int] NULL,
    [Quantité_reclamee] [float] NOT NULL,
    [Code_Première_Facture] [nvarchar](255) NULL,
    [Code_Seconde_Facture] [nvarchar](255) NULL,
    [Code_Troisième_Facture] [nvarchar](255) NULL,
    [Code_Quatrième_Facture] [nvarchar](255) NULL,
    [Numero_Reclamation] [int] NOT NULL,
)

```

Figure 55 : Création Table Réclamations\_pieces (Nous-même, 2020)

```

USE [BDDPfe]

CREATE TABLE [dbo].[Commandes](
    [FK_KD] [int] NOT NULL,
    [FK_Produit] [int] NOT NULL,
    [FK_Statut_Financier] [int] NOT NULL,
    [FK_Statut_Customer_Service] [int] NOT NULL,
    [FK_Entrepôt] [int] NOT NULL,
    [FK_CCFOT_Reasons] [int] NOT NULL,
    [FK_DateCreationSO] [int] NULL,
    [FK_DateBillCreation] [int] NULL,
    [Sales_Type] [nvarchar](255) NOT NULL,
    [PO_Number] [nvarchar](255) NOT NULL,
    [SO_Number] [float] NOT NULL,
    [Proforma_Code] [float] NULL,
    [Delivery_Order_Code] [float] NULL,
    [Delivery_Block] [nvarchar](255) NULL,
    [Bill_Document_Code] [float] NULL,
    [Credit_Status] [nvarchar](255) NULL,
    [Order_qty] [float] NOT NULL,
    [Confirmed_qty] [float] NULL,
    [Billed_qty] [float] NULL)

```

Figure 56 : Création Table Commande (Nous-même, 2020)

```

Use BDDPfe
CREATE TABLE dbo.Statut_Financier
(
    [SK_Statut_Financier] INT IDENTITY(1,1) NOT NULL
    , [Statut_Financier] nvarchar(255) NULL
);

```

Figure 57 : Création Table Statut\_Financier – 1 (Nous-même, 2020)

SK_Statut_Financier	Statut_Financier
1	Under Finance
2	Not Under Finance

Figure 58 : Création Table Statut\_Financier - 2 (Nous-même, 2020)

```
Use BDDPfe
CREATE TABLE dbo.Statut_Customer_Service
(
  [SK_Statut_Customer_Service] INT IDENTITY(1,1) NOT NULL
  , [Statut_Customer_Service] nvarchar(255) NULL
);
```

Figure 59 : Création Table Statut\_Customer\_Service - 1 (Nous-même, 2020)

SK_Statut_Customer_Service	Statut_Customer_Service
1	Under Treatment
2	Waiting for shipment
3	Invoiced

Figure 60 : Création Table Statut\_Customer\_Service - 2 (Nous-même, 2020)

```
Use BDDPfe
CREATE TABLE dbo.Entrepot
(
  [SK_Entrepot] INT IDENTITY(1,1) NOT NULL
  , [Statut_Entrepot] nvarchar(255) NULL
);
```

Figure 61 : Création Table Entrepôt - 1 (Nous-même, 2020)

SK_Entrepot	Statut_Entrepot
1	C001
2	C002
3	C003

Figure 62 : Création Table Entrepôt - 2 (Nous-même, 2020)

```

SET DATEFIRST 7
DATEFORMAT dmy,
LANGUAGE FRENCH;
DECLARE @StartDate date = CAST('01-01-2001' as Date)
DECLARE @CutoffDate date = DATEADD(DAY, -1, DATEADD(YEAR, 30, @StartDate));
Use BDDPfe
select @startdate;
;WITH seq(n) AS
(
SELECT 0 UNION ALL SELECT n + 1 FROM seq
WHERE n < DATEDIFF(DAY, @StartDate, @CutoffDate)
),
d(d) AS
(
SELECT DATEADD(DAY, n, @StartDate) FROM seq
),
src AS
(SELECT
TheDate          = CONVERT(date, d),
TheWeek          = DATEPART(WEEK,    d),
TheMonth         = DATEPART(MONTH,   d),
TheMonthName     = DATENAME(MONTH,   d),
TheQuarter       = DATEPART(Quarter, d),
TheYear          = DATEPART(YEAR,    d)
FROM d )
SELECT * INTO dbo.Date FROM src

ORDER BY TheDate
OPTION (MAXRECURSION 0)
CREATE UNIQUE CLUSTERED INDEX PK_Date ON dbo.Date(TheDate);

```

Figure 63 : Creation Table Date – 1 (Nous-même, 2020)

	TheDate	TheWeek	TheMonth	TheMonthName	TheQuarter	TheYear	SK_Date
1	NULL	NULL	NULL	NULL	NULL	NULL	10958
2	2001-01-01	1	1	janvier	1	2001	1
3	2001-01-02	1	1	janvier	1	2001	2
4	2001-01-03	1	1	janvier	1	2001	3
5	2001-01-04	1	1	janvier	1	2001	4
6	2001-01-05	1	1	janvier	1	2001	5
7	2001-01-06	1	1	janvier	1	2001	6
8	2001-01-07	2	1	janvier	1	2001	7
9	2001-01-08	2	1	janvier	1	2001	8
10	2001-01-09	2	1	janvier	1	2001	9
11	2001-01-10	2	1	janvier	1	2001	10
12	2001-01-11	2	1	janvier	1	2001	11

✓ Exécution de requête réussie. | USER-PC (15.0 RTM) | user-PC\user (60) | BDDPfe | 00:00:00 | 1 000 lignes

Figure 64 : Creation Table Date – 2 (Nous-même, 2020)

Tiré et modifié de : <https://www.mssqltips.com/sqlservertip/4054/creating-a-date-dimension-or-calendar-table-in-sql-server/>

```

USE [BDDPfe]

CREATE TABLE [dbo].[TTS](
  [SK_TTS] [int] IDENTITY(1,1) NOT NULL,
  [TTS] [nvarchar](255) NULL,

```

Figure 65 : Création table TTS

```

USE [BDDPfe]

CREATE TABLE [dbo].[Remboursements_TTS](
  [FK_KD] [int] NOT NULL,
  [FK_TTS] [int] NOT NULL,
  [FK_Date] [int] NOT NULL,
  [Numero_Note_De_Credit] [float] NOT NULL,
  [Montant_Note_Credit_HT] [float] NULL,
  [Montant_Note_Credit_TTC] [float] NULL)

```

Figure 66 : Creation Table Remboursements TTS (Nous-même, 2020)

```

USE [BDDPfe]

CREATE TABLE [dbo].[Type_Reclamation](
  [SK_Type_Reclamation] [int] IDENTITY(1,1) NOT NULL,
  [Type_Reclamation] [nvarchar](255) NULL)

```

Figure 67 : Creation Table Type Réclamation - 1 (Nous-même, 2020)

SK_Type_Reclamation	Type_Reclamation
1	Empty
2	Partially Empty
3	Avarie
4	Shortage
5	Quality Issue

Figure 68 : Contenu de la table Type Réclamation - 2 (Nous-même, 2020)

```

USE [BDDPfe]

CREATE TABLE [dbo].[Statut_Reclamation](
  [SK_statut_Reclamation] [int] IDENTITY(1,1) NOT NULL,
  [Statut_Reclamation] [nvarchar](255) NULL)

```

Figure 69 : Creation Table Statut Réclamation – 1 (Nous-même, 2020)

SK_Statut_Reclamation	Statut_Reclamation
6	In review
7	Resolved
8	Open
9	Rejected

Figure 70 : Contenu de la table Statut Réclamation - 2 (Nous-même, 2020)

```
USE [BDDPfe]

CREATE TABLE [dbo].[CCFOT_Reasons](
  [SK_CCFOT_Reasons] [int] IDENTITY(1,1) NOT NULL,
  [CCFOT_Reasons] [nvarchar](255) NULL)
```

Figure 71 : Creation Table CCFOT Reasons - 1 (Nous-même, 2020)

SK_CCFOT_Reasons	CCFOT_Reasons
1	SSXX
2	SPXX
3	NULL
4	Valeur non inscrite dans la BDD

Figure 72 : Contenu de la table CCFOT Reasons - 2 (Nous-même, 2020)

## Annexe 6 : Architecture du logiciel SSIS

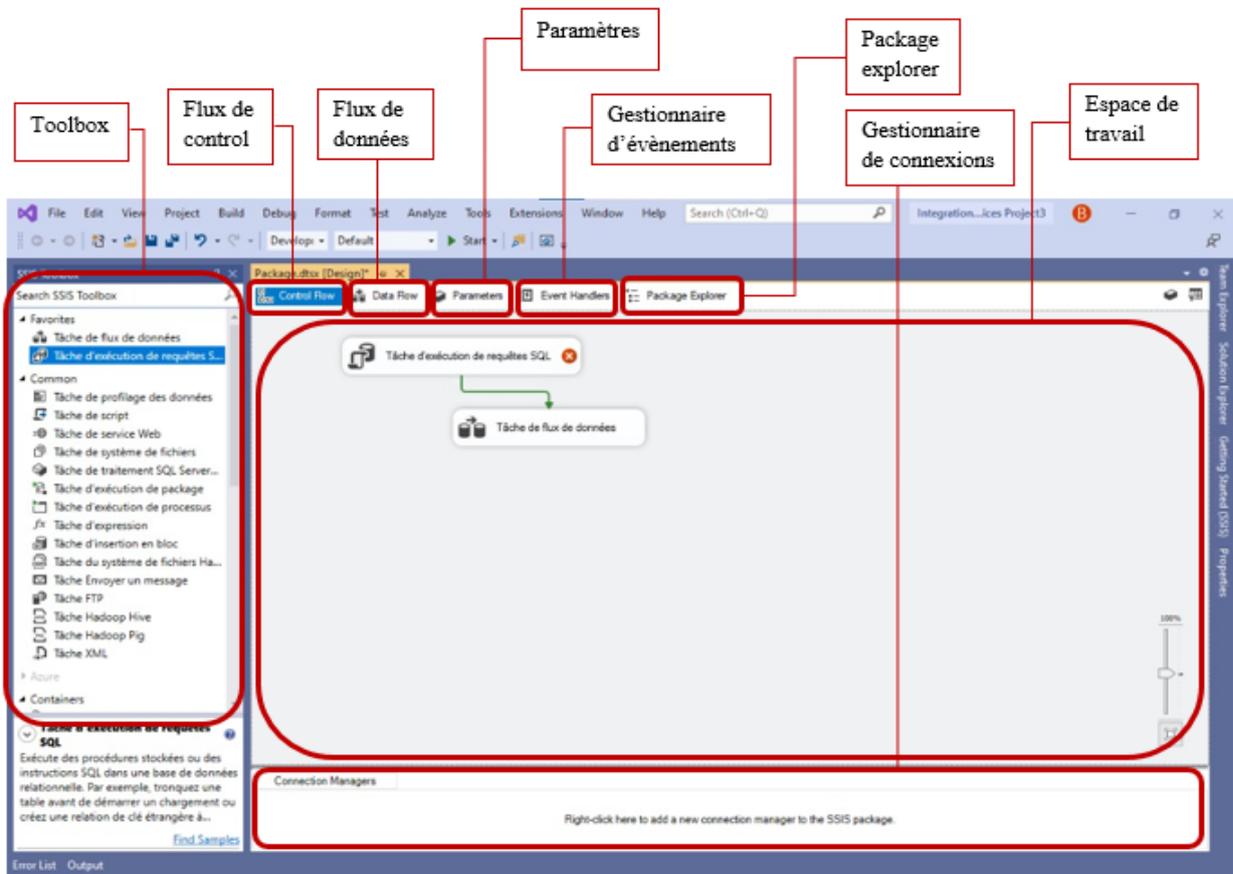


Figure 73 : Architecture du logiciel SSIS

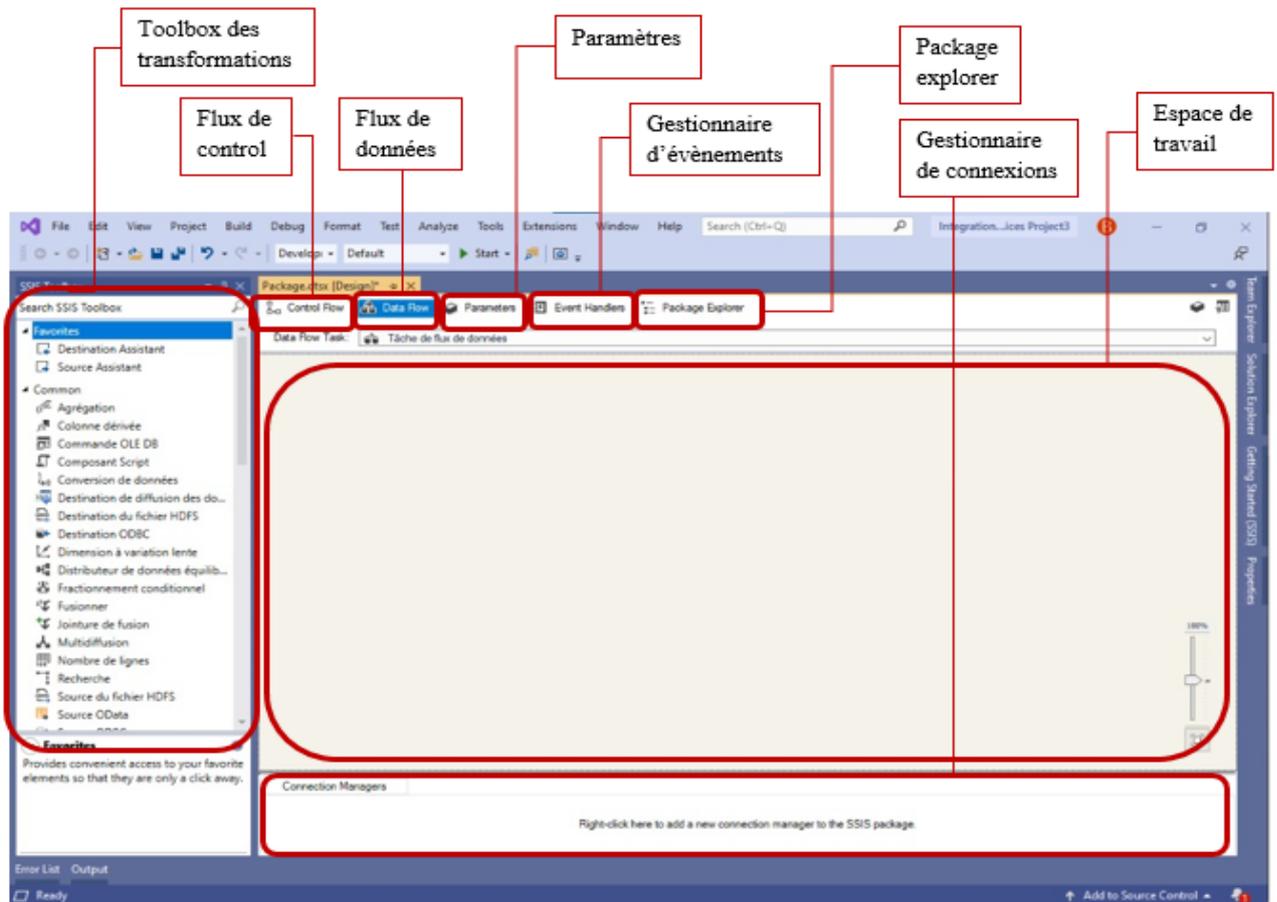


Figure 74 : Vue sur la fenêtre Dataflow

## Annexe 7 : Exemples de transformations dans SSIS

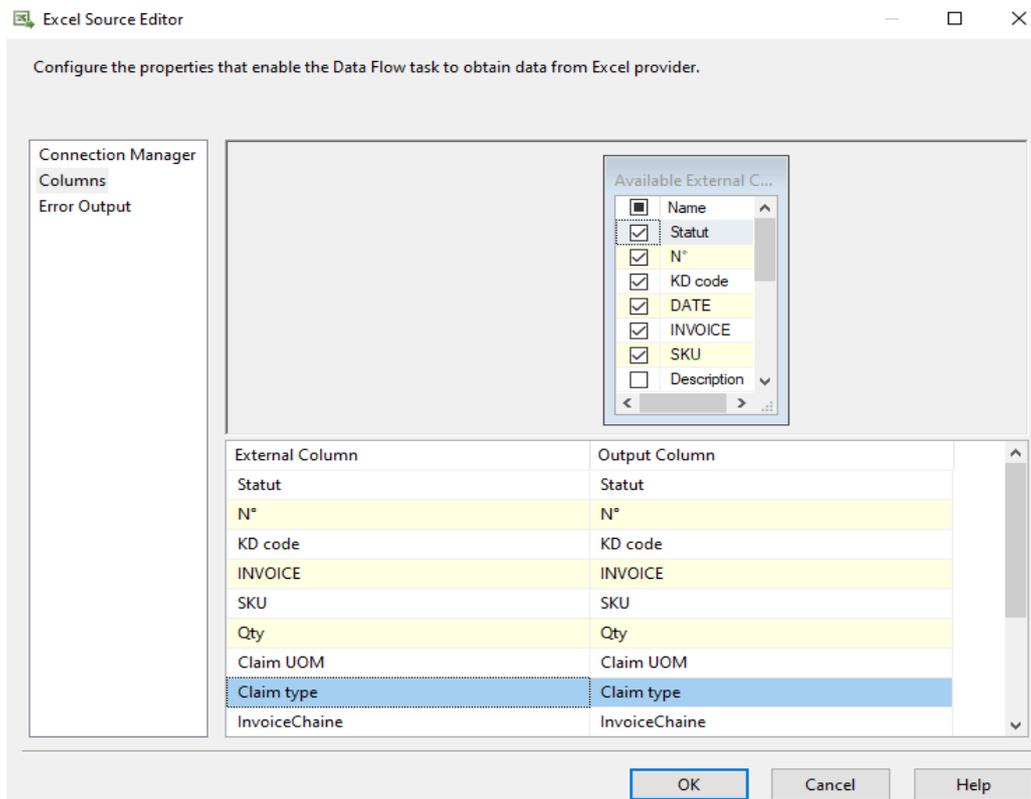


Figure 75 : Choix des colonnes à extraire de la source

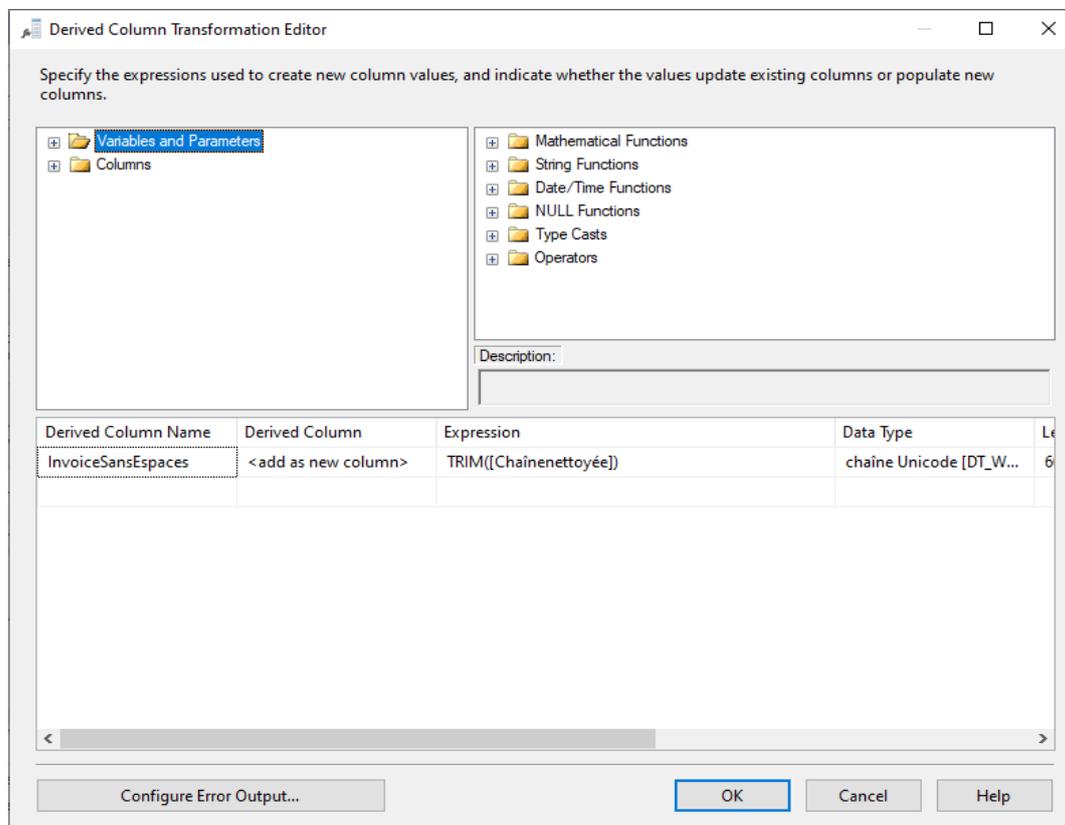


Figure 76 : Colonne dérivée : Suppression des espaces de la chaîne

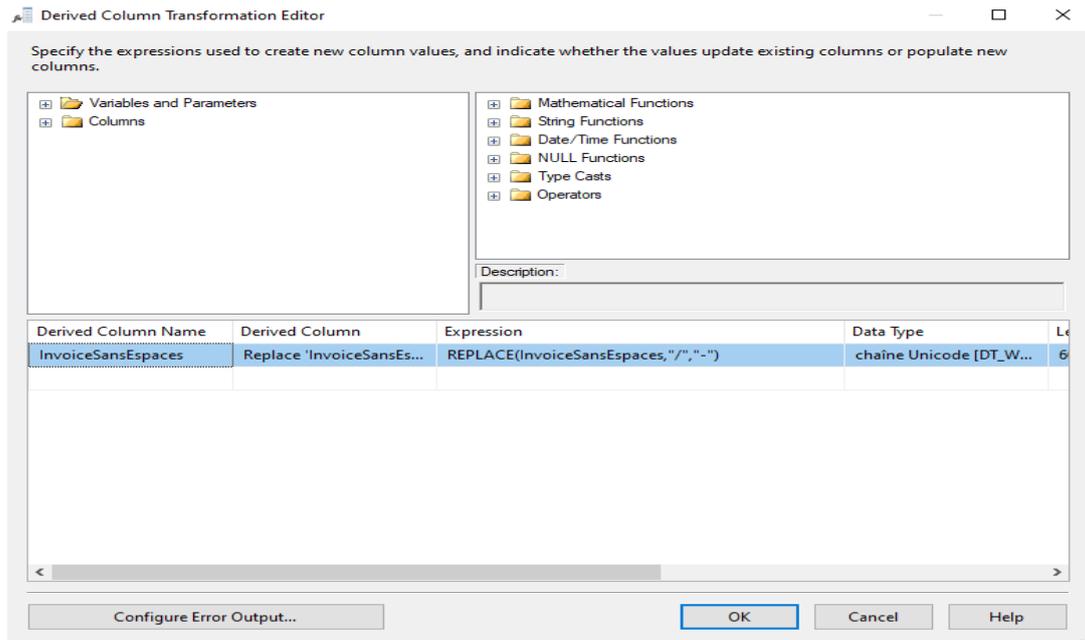


Figure 77 : Colonne dérivée : Remplacement des '/' par des '-'

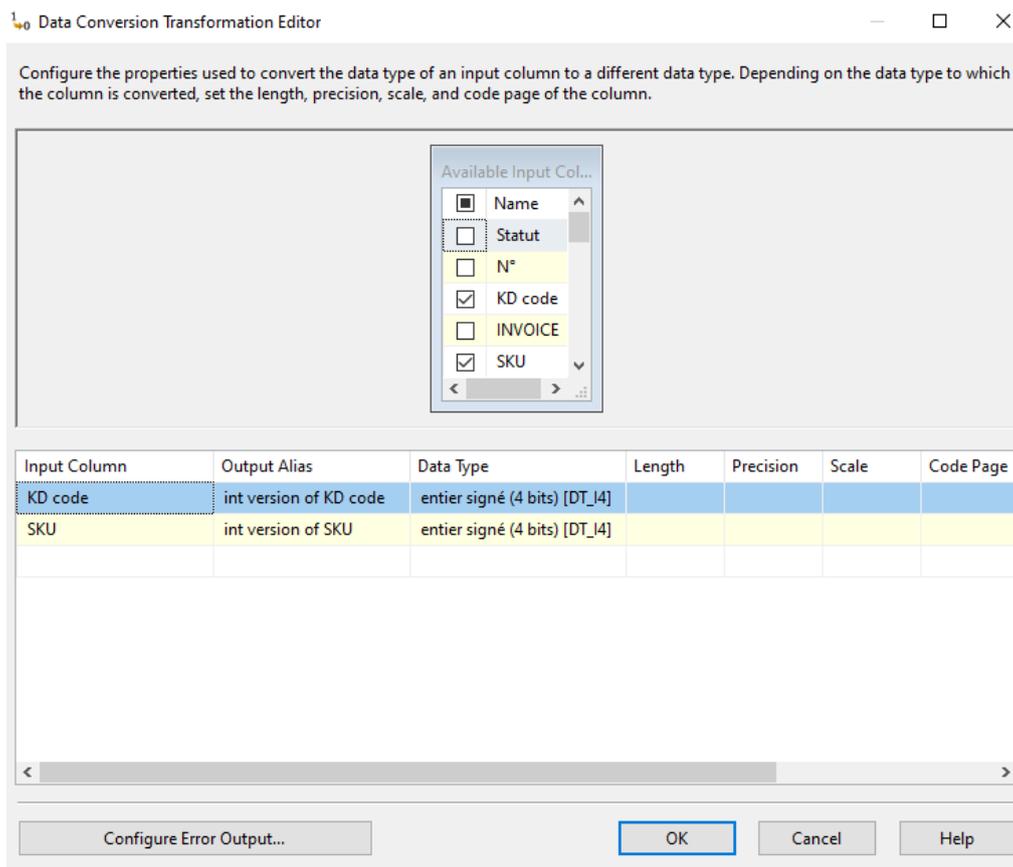


Figure 78 : Conversion des données : du float à l'entier

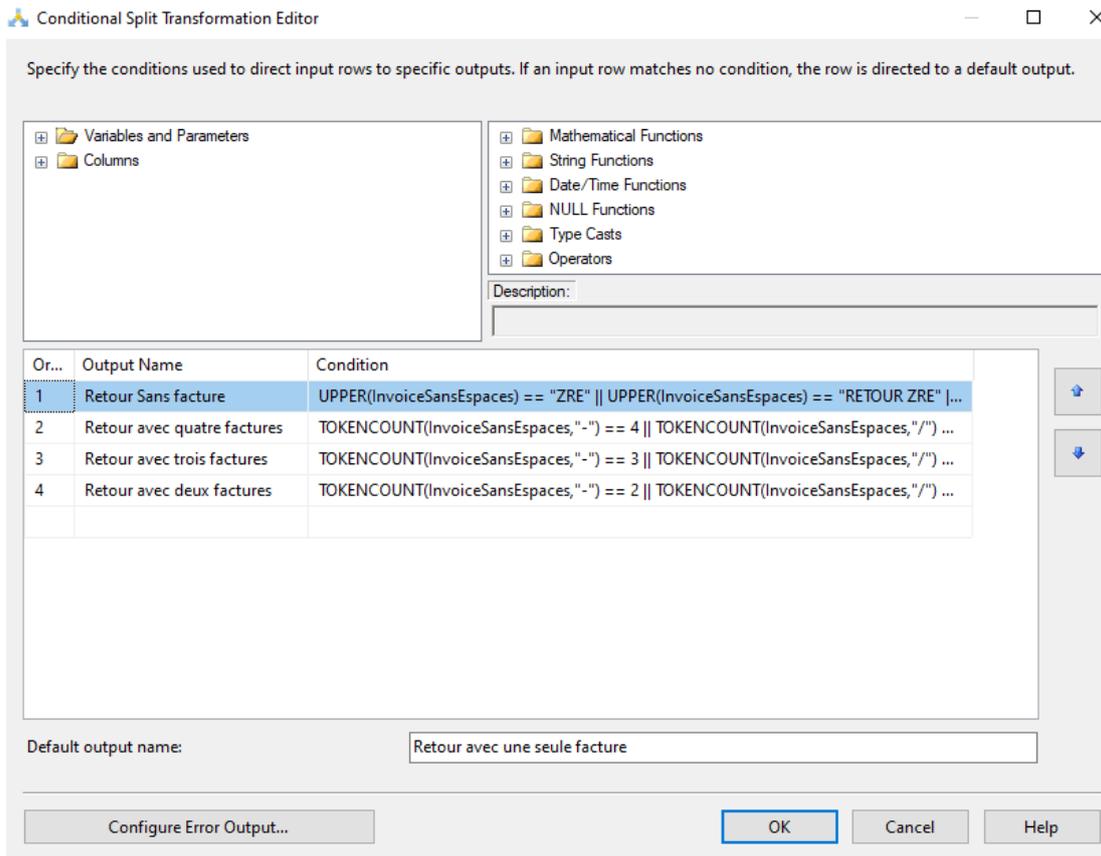


Figure 79 : Fractionnement conditionnel en quatre flux selon les factures

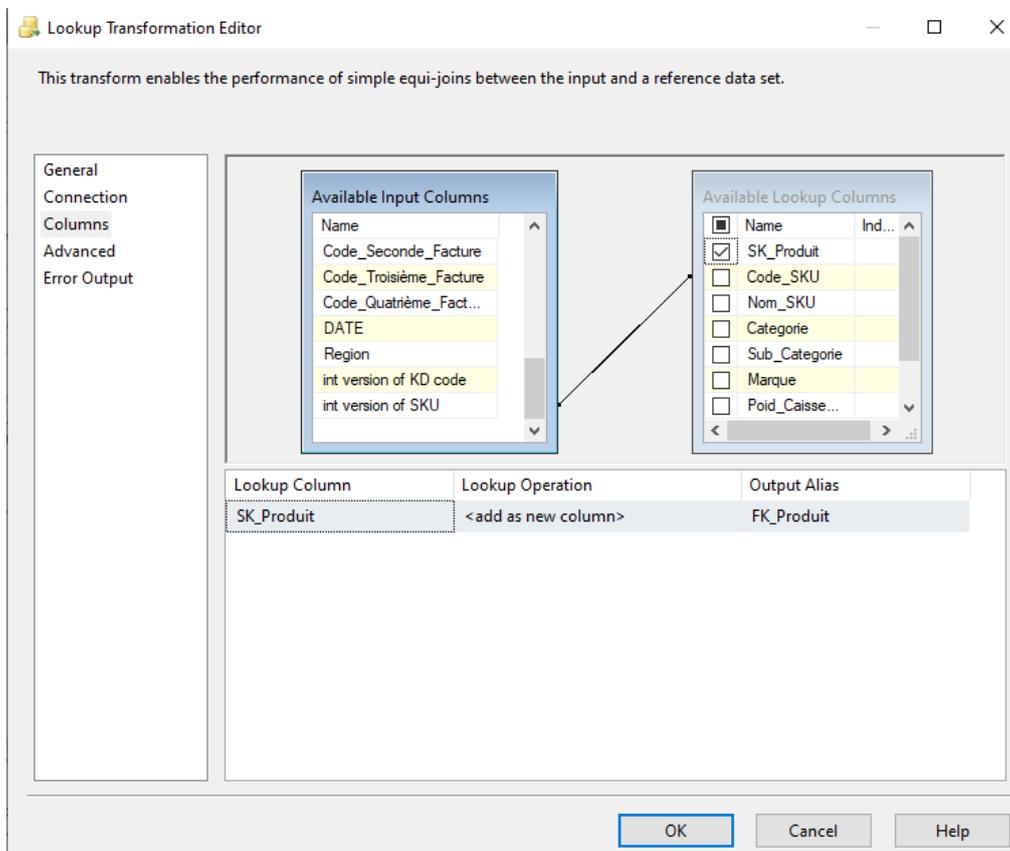


Figure 80 : Recherche : Correspondance des codes SKUs de la source et de la table Produit

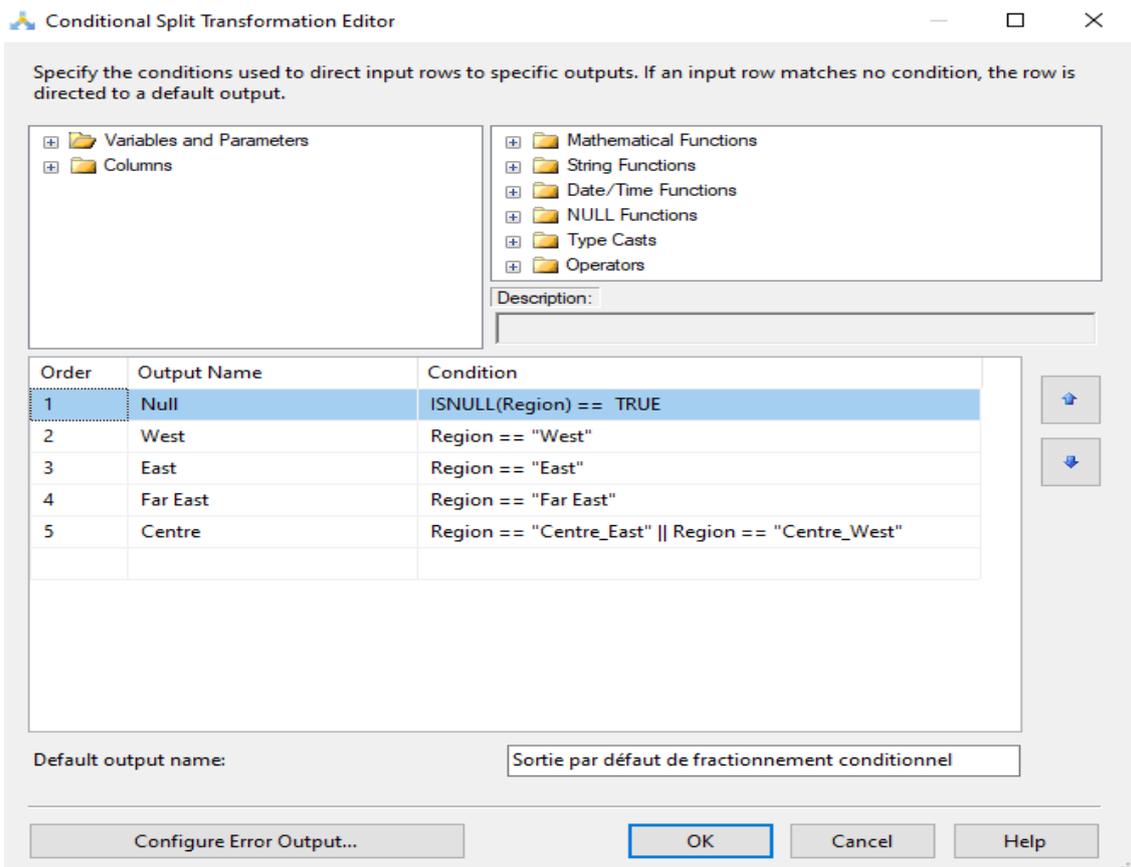


Figure 81 : Fractionnement conditionnel en 5 flux selon les régions

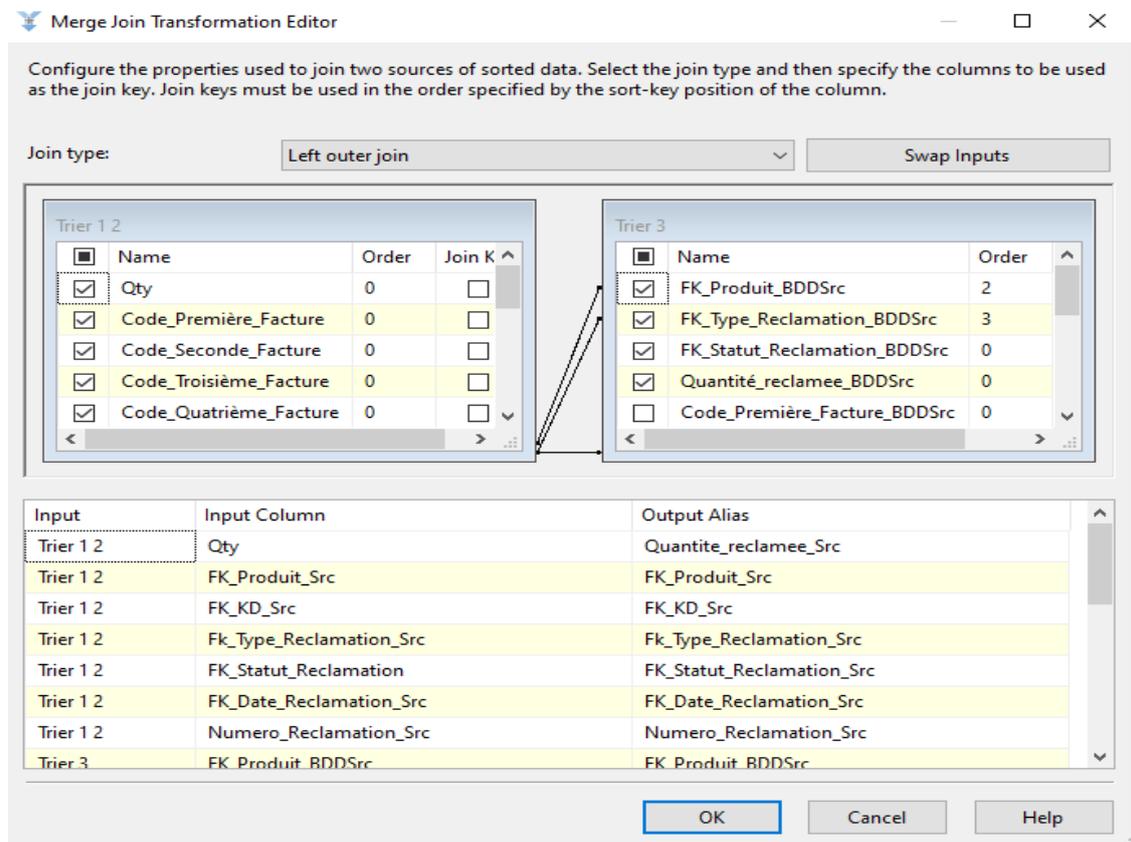


Figure 82 : Jointure de fusion

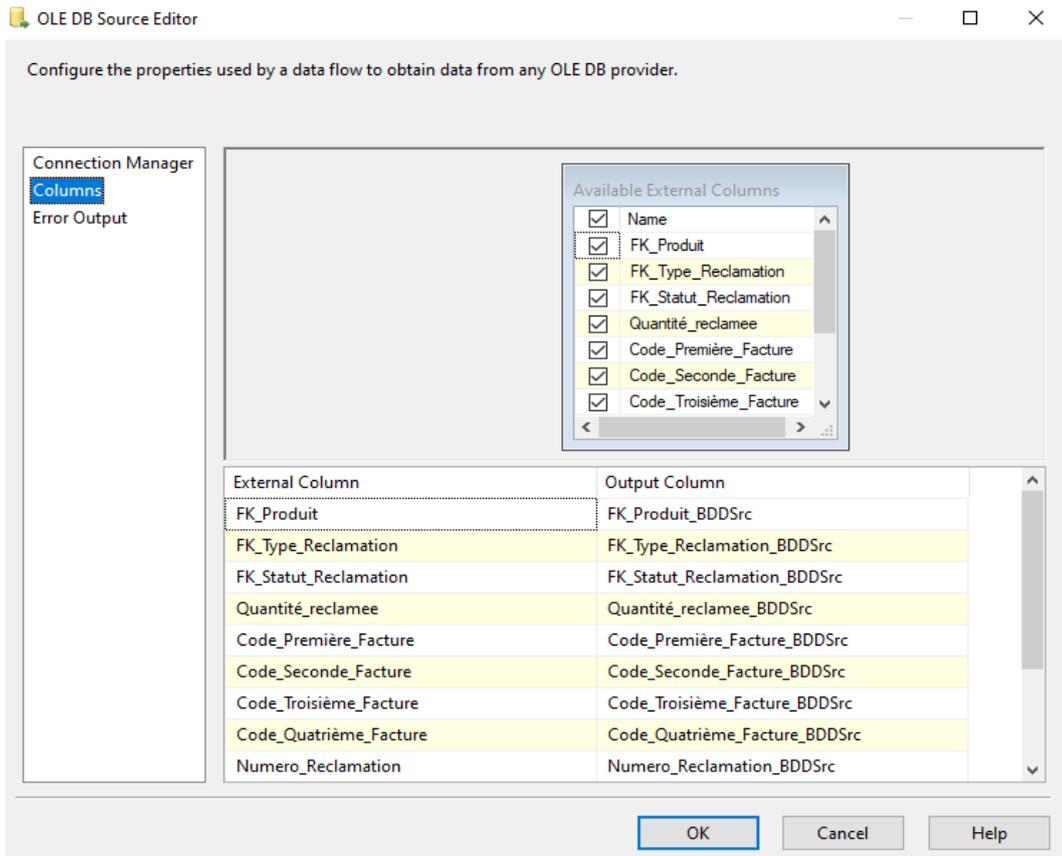


Figure 83 : Choix des colonnes de la table à extraire

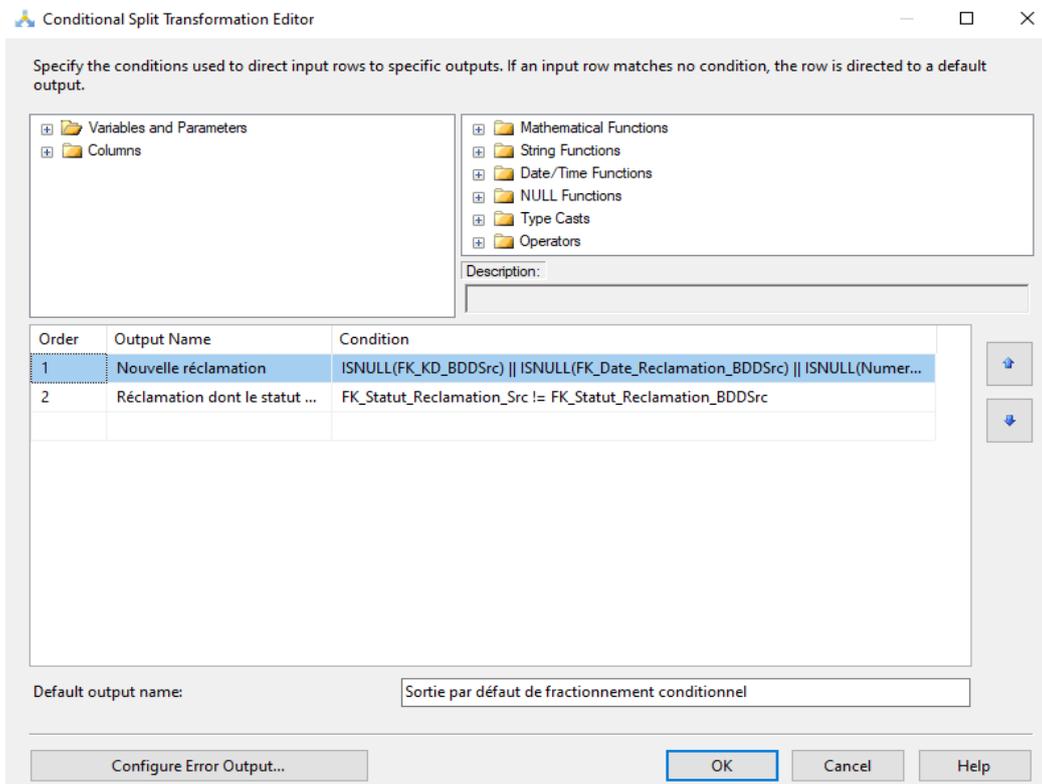


Figure 84 : Fractionnement conditionnel : Nouvelles et anciennes réclamations

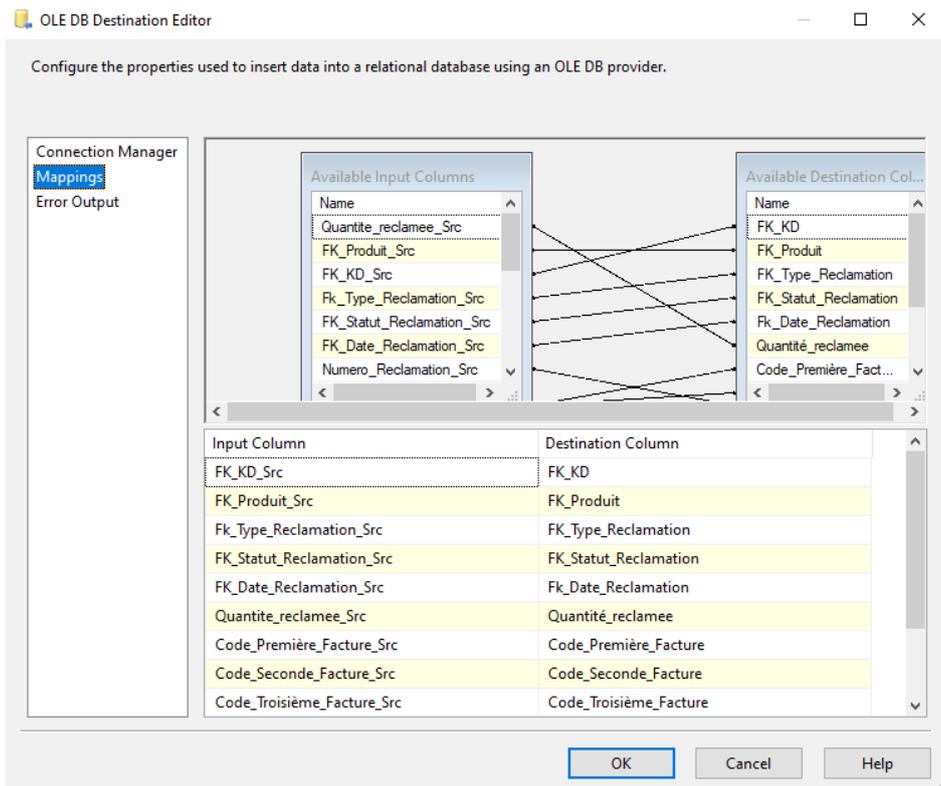


Figure 85 : Choix des colonnes à insérer dans la table

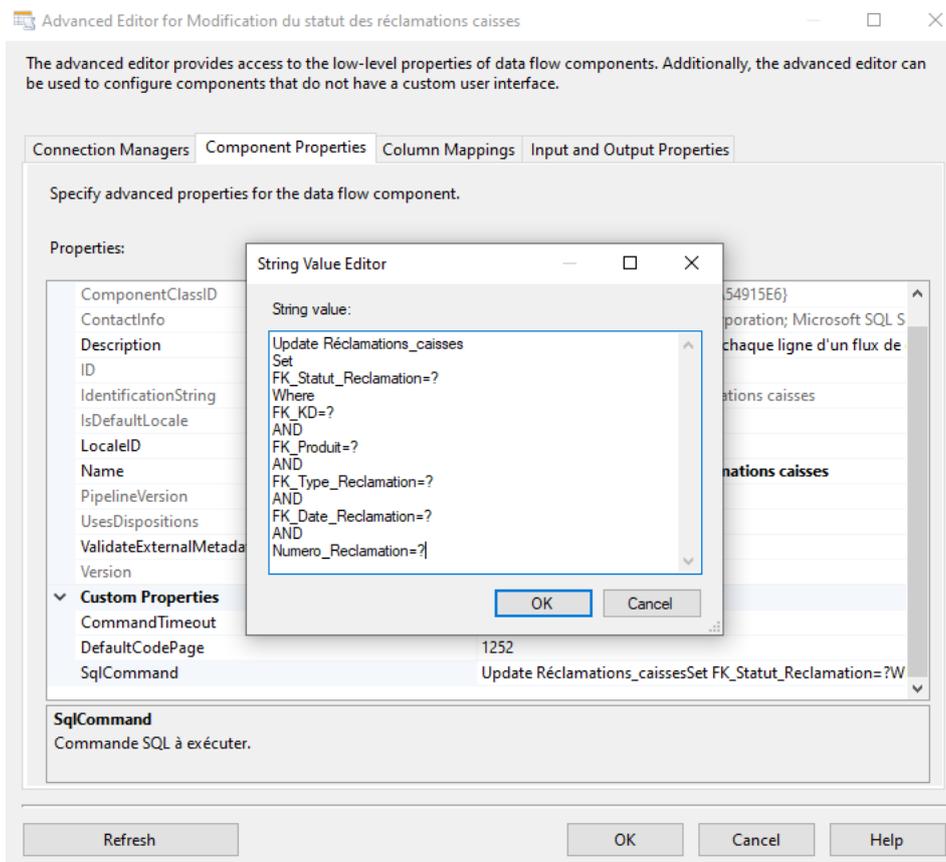


Figure 86 : Requête de la mise à jour la table Réclamations\_caisse

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

<input type="checkbox"/>	Name	Pass Throu...
<input checked="" type="checkbox"/>	FK_Produit_BDDSrc	<input type="checkbox"/>
<input checked="" type="checkbox"/>	FK_Type_Reclamation_BDDSrc	<input type="checkbox"/>
<input type="checkbox"/>	FK_Statut_Reclamation_BDDSrc	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Quantité_reclamee_BDDSrc	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Code_Première_Facture_BDDSrc	<input checked="" type="checkbox"/>

Input Column	Output Alias	Sort Type	Sort Order	Con
FK_Produit_BDDSrc	FK_Produit_BDDSrc	ascending	2	
FK_Type_Reclamation_BDDSrc	FK_Type_Reclamation_BDDSrc	ascending	3	
Numero_Reclamation_BDDSrc	Numero_Reclamation_BDDSrc	ascending	4	
FK_KD_BDDSrc	FK_KD_BDDSrc	ascending	1	
Fk_Date_Reclamation_BDDSrc	Fk_Date_Reclamation_BDDSrc	ascending	5	

Remove rows with duplicate sort values

Figure 87 : Tri des données extraites de la table Réclamations\_caisses

## Annexe 8 : Modèles des datamarts sur Power Pivot

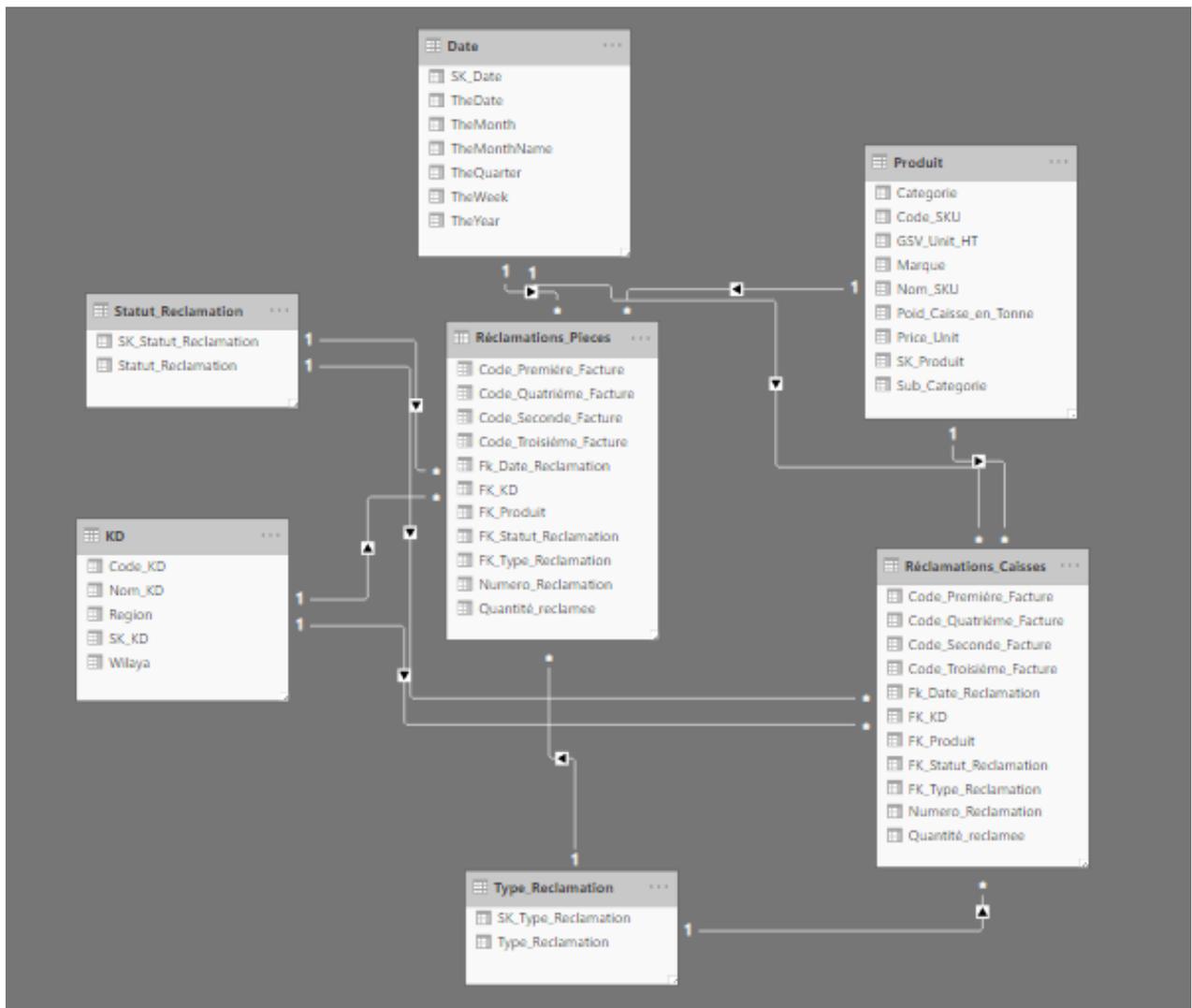


Figure 88 : Modèle du datamart Réclamations sur Power Pivot (Nous-même, 2020)

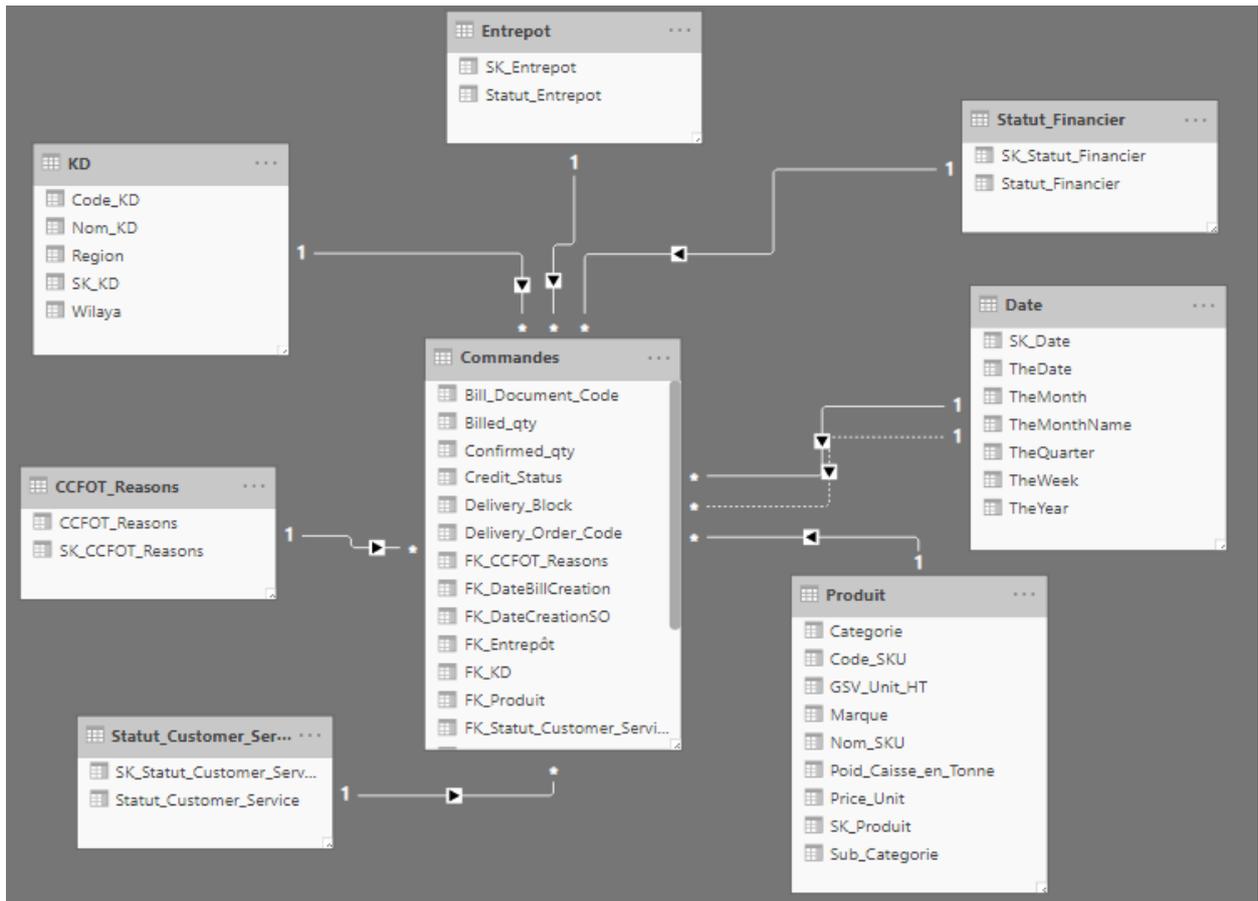


Figure 89 : Modèle du datamart Commandes sur Power Pivot (Nous-même, 2020)

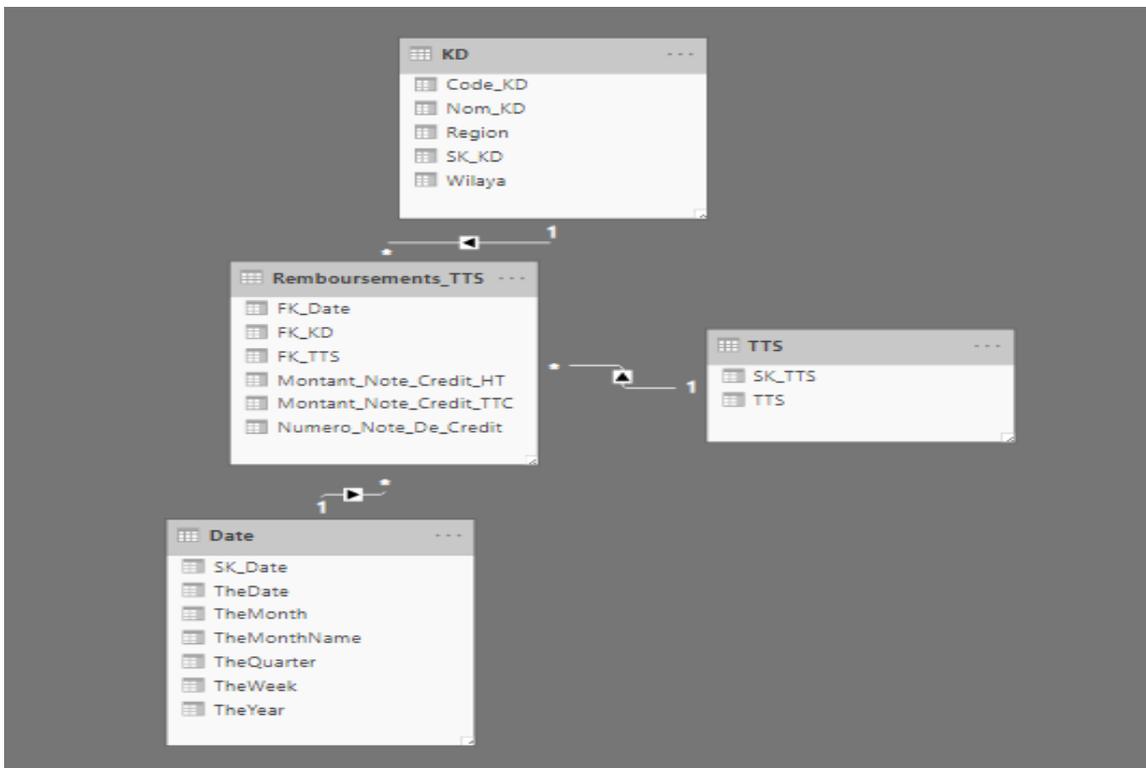


Figure 90 : Modèle du datamart Remboursement TTS sur Power Pivot (Nous-même, 2020)