

**École Supérieure Privée de Management de Tunis**  
**Esprit School of Business**



# **RAPPORT DE STAGE**

En vue de l'obtention du diplôme de Licence en Business Computing

## **Parcours Business Intelligence**

---

Soutenu le : ...

Réalisé par : **Ouahbi HAJ KACEM**

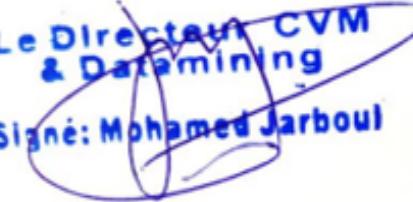
« Segmentation du parc Data et Proposition d'une solution reporting (BI) : cas de Tunisie Télécom »

17/02/2022 – 17/06/2022

**Maître de stage : Mr Mohamed JARBOUI**

**Encadrant académique : Mme. Ines BEN FREDJ**

Année universitaire 2021/2022

	Nom & Prénom	Date et Signature
Maître de stage	Mohamed JARBOUI	<p>Le Directeur CVM  &amp; DataMining</p> <p>Signé: Mohamed Jarboui</p> 
Encadrant académique	Ines BEN FREDJ	 <p>Le 29/06/2022</p>

## Dédicaces

Je dédie ce projet de fin d'études :

A mon père Sabeur HAJ KACEM et ma mère Madiha GATTOUSSI, pour tous leurs sacrifices, leur amour, leur soutien et leurs prières tout au long de mes études, à mon frère Yassine HAJ KACEM pour ses conseils, son soutien et ses encouragements, à toute ma famille pour leur support tout au long de mon parcours universitaire, que ce travail soit la réalisation de vos souhaits tant désirés et le fruit de votre soutien indéfectible, à tous mes amis et à tous ceux qui ont eu la gentillesse de m'aider et de me soutenir dans la réalisation de ce travail. Merci d'avoir toujours été là pour moi.

HAJ KACEM Ouahbi

## **Remerciements**

Je voudrais remercier tous ceux qui ont aidé, encouragé et contribué directement ou indirectement à cet humble projet.

Tout d'abord, nous nous adressons aux membres du jury et les remercions d'avoir accepté d'évaluer ce travail.

Nous adressons nos plus sincères remerciements à notre maître de stage à Esprit School of Business, Madame Ines BEN FREDJ pour sa disponibilité, pour son attention et pour tous les conseils judicieux qu'elle nous a prodigués pour accomplir correctement ce projet.

Nous tenons à remercier notre directeur de stage, Monsieur Mohamed JARBOUI et Madame Khawla BEN BAHRI, pour l'aide et les précieux conseils qu'ils nous ont apportés et pour leur accueil chaleureux tout au long de notre stage.

# Table des matières

<b>CHAPITRE 1</b>	<b>13</b>
<b>PRÉSENTATION DU CONTEXTE DU PROJET</b>	<b>13</b>
1. Introduction	13
2. Présentation de l'organisme d'accueil	13
2.1. Présentation générale de l'entreprise Tunisie Télécom	13
2.2. L'Organigramme de l'entreprise	13
3. Présentation du projet	14
3.1. Contexte du projet	14
3.2. Problématique	14
3.3. Solution proposée	15
4. Conclusion	15
<b>CHAPITRE 2</b>	<b>16</b>
<b>ETUDE DES MÉTHODOLOGIES</b>	<b>16</b>
1. Introduction	16
2. Présentation de la méthode de gestion de projet	16
2.1 Etude comparative entre les approches de gestion du projet	16
2.1.1. Les approches classiques	16
2.1.2 Les approches agiles	16
2.2. Etude comparative entre les méthodologies de gestion du projet	17
3. Architecture du système décisionnel :	18
4. Comparaison entre l'approche TOP-DOWN et BOTTOM-UP	19
4.1. Approche de Bill Inmon	19
4.2. Approche de Ralph Kimball	20
5. Composants d'un schéma dimensionnel	21
5.1. Tables de faits	21
5.2. Tables de dimensions	22
5.3. La différence entre tables des faits et tables des dimensions	22
5.4. Schémas de la modélisation dimensionnelle	23
5.4.1. Le modèle en étoile	23
5.4.2. Le modèle en flocon	24
5.4.3. Le modèle en constellation	24
5.5. Alimentation de l'entrepôt de données	25
5.6. Exploitation de l'entrepôt de données	25
5.6.1. Types de modélisation OLAP	26
5.6.2. Restitution	26
6. Récapitulatif des choix	27
6.1. Choix de méthodologie	27
6.2. Choix de la démarche	28
6.3. Choix du type de modèle	29

7. Planification prévisionnelle	29
8. Conclusion	29
<b>CHAPITRE 3</b>	<b>30</b>
<b>DÉFINITION DES BESOINS ET MODÉLISATION</b>	<b>30</b>
1. Introduction	30
2. Spécification des besoins	30
2.1. Acteurs	30
2.2. Besoins fonctionnels	30
2.3. Besoins non fonctionnels	31
2.4. Modélisation des besoins	31
3. Architecture	32
3.1. Architecture fonctionnelle du système	32
3.2. Architecture logique des données	32
3.3. Choix des indicateurs	33
3.3.1. Caractéristiques des indicateurs clé de performances	33
3.3.2. Le choix des indicateurs clé de performances	33
4. Modélisation de notre Data Warehouse	34
4.1. Description des dimensions	34
4.1.1. Dimension Date:	35
4.1.2. Dimension Client	35
4.1.3. Dimension Offre	36
4.1.4. Dimension Montant	36
4.1.5. Dimension Service	36
4.1.6. Dimension Consommation	37
4.1.7. Dimension Forfait	37
4.1.8. Dimension Transact	38
4.2. Description de Faits	38
4.2.1. Fait Trafic Offre	38
4.2.2. Fait Trafic Service	39
4.2.3. Fait Trafic Forfait	40
4.3. Vue d'ensemble de la conception de l'entrepôt	40
5. Processus de chargement ETL	44
5.1. Processus de chargement des dimensions	44
5.2. Processus de chargement des tables de faits	46
6. Conception des cubes dimensionnels OLAP	46
6.1. La conception des cubes OLAP	47
6.2. Liste des cubes OLAP	47
7. Conclusion	49
<b>CHAPITRE 4</b>	<b>50</b>
<b>RÉALISATION ET TESTS</b>	<b>50</b>
1. Introduction	50
2. Environnement	50

2.1. Environnement matériel	50
2.2. Environnement logiciel	50
2.3. Outils de développement	52
2.4. Langage de programmation	53
2.5. Langages descriptifs	53
2.6. Plateforme de développement	53
3. Implémentation ETL	53
3.1. Composants de SSIS	53
3.2. Sources de données	55
3.3. Nettoyage de données	55
3.4. Connexion à la base de données	56
4. Chargement des Datamarts	56
4.1. Datamart “Trafic_Forfait”	56
4.1.1. Chargement des dimensions	57
4.1.2. Chargement de la table de faits Trafic_Forfait :	59
4.2. Chargement des autres tables de faits	60
5. L'entrepôt de données	62
6. Mise en œuvre et analyse des données : Les cubes	64
6.1. Cube Trafic_Forfait	64
6.2. Cube Trafic_Offre	65
6.3. Cube Trafic_Service	65
7. Réalisation du reporting : Les tableaux de bord	66
7.1. Obtention des données à partir de la base de données SSAS	67
7.2. Tableau de bord “Trafic_Offre”	68
7.3. Tableau de bord “Trafic_Forfait”	69
7.4. Tableau de bord “Trafic_Service”	71
8. Initiative : Implémentation web	73
9. Conclusion	77

# Liste des figures

FIGURE 1: LOGO TUNISIE TÉLÉCOM	13
FIGURE 2: L'ORGANIGRAMME DE TT	14
FIGURE 4 : TOP-DOWN vs BOTTOM-UP	19
FIGURE 5: APPROCHE DE BILL INMON	20
FIGURE 6: APPROCHE DE RALPH KIMBALL	21
FIGURE 7: EXEMPLE D'UNE TABLE DE FAIT "VENTE"	22
FIGURE 8: EXEMPLE D'UNE TABLE DE DIMENSION "DATE"	22
FIGURE 9: MODÈLE EN ÉTOILE	24
FIGURE 10: MODÈLE EN FLOCON	24
FIGURE 11: MODÈLE EN CONSTELLATION	25
FIGURE 12: PROCESSUS ETL	25
FIGURE 13: MÉTHODE GIMSI	27
FIGURE 14: DIAGRAMME DE GANTT	29
FIGURE 15: DIAGRAMME DU CAS D'UTILISATION	31
FIGURE 16: ARCHITECTURE LOGIQUE DES DONNÉES	32
FIGURE 17 : DIM DATE	35
FIGURE 18: DIM CLIENT	35
FIGURE 19: DIM OFFRE	36
FIGURE 20: DIM MONTANT	36
FIGURE 21: DIM SERVICE	37
FIGURE 22: DIM CONSOMMATION	37
FIGURE 23: DIM FORFAIT	37
FIGURE 24: DIM TRANSACT	38
FIGURE 25: FACT TRAFIC OFFRE	39
FIGURE 26: FACT TRAFIC SERVICE	39
FIGURE 27: FACT TRAFIC FORFAIT	40
FIGURE 28: NOTRE DATA WAREHOUSE	41
FIGURE 29: DATAMART TRAFIC OFFRE	42
FIGURE 30: DATAMART TRAFIC FORFAIT	43
FIGURE 31: DATAMART TRAFIC SERVICE	43
FIGURE 32: CUBE OLAP "TRAFIC_SERVICE"	48
FIGURE 33: CUBE OLAP "TRAFIC_OFFRE"	48
FIGURE 34: CUBE OLAP "TRAFIC_FORFAIT"	49
FIGURE 35: STARUML	50
FIGURE 36: GANTT PROJECT	51
FIGURE 37: SQL SERVER 2019	51
FIGURE 38: SSMS 18	51
FIGURE 39: VISUAL STUDIO 2019	52

FIGURE 40: POWER BI DESKTOP	52
FIGURE 41: FICHIER SOURCE DE TYPE CSV	55
FIGURE 42: FICHIER SOURCE DE TYPE EXCEL	55
FIGURE 43: LA CONNEXION À L'ENTREPÔT DE DONNÉES DE DESTINATION « DWH_TT »	56
FIGURE 44: DIMENSION DATE	57
FIGURE 45: DIMENSION CLIENT	58
FIGURE 46: DIMENSION TRANSACT	58
FIGURE 47: DIMENSION FORFAIT	59
FIGURE 48: TABLE DE FAITS “TRAFFIC_FORFAIT”	60
FIGURE 49: TABLE DE FAITS “TRAFFIC_OFFRE”	61
FIGURE 50: TABLE DE FAITS “TRAFFIC_SERVICE”	61
FIGURE 51: EXÉCUTION FINALE DU PROCESSUS ETL	63
FIGURE 52: RÉALISATION ET DÉPLOIEMENT DU CUBE “TRAFFIC_FORFAIT”	64
FIGURE 53: RÉALISATION ET DÉPLOIEMENT DU CUBE “TRAFFIC_OFFRE”	65
FIGURE 54: RÉALISATION ET DÉPLOIEMENT DU CUBE “TRAFFIC_SERVICE”	66
FIGURE 55: CONNEXION AU SERVEUR “LOCALHOST”	67
FIGURE 56: EXEMPLE D'OBENTION DES DONNÉES DU CUBE “TRAFFIC_SERVICE”	67
FIGURE 57: PAGE 1 TDB “TRAFFIC_OFFRE”	68
FIGURE 58: PAGE 2 TDB “TRAFFIC_FORFAIT”	69
FIGURE 59: PAGE 1 TDB “TRAFFIC_FORFAIT”	70
FIGURE 60: PAGE 2 TDB “TRAFFIC_FORFAIT”	71
FIGURE 61: PAGE 1 TDB “TRAFFIC_SERVICE”	72
FIGURE 62: PAGE 2 TDB “TRAFFIC_SERVICE”	73
FIGURE 63: BOÎTE D'ALERTE	74
FIGURE 64: LE CONTRÔLE SUR LES CHAMPS DU FORMULAIRE DE CONNEXION	74
FIGURE 65 : PERMETTRE AU DIRECTEUR DU DÉPARTEMENT CVM DATAMINING DE SE CONNECTER	75
FIGURE 66 : CHOIX DU TABLEAU DE BORD À CONSULTER	75
FIGURE 67 : TABLEAU DE BORD TRAFFIC_FORFAIT	76
FIGURE 68 : TABLEAU DE BORD TRAFFIC_OFFRE	76
FIGURE 69 : TABLEAU DE BORD TRAFFIC_SERVICE	77

# Liste des tableaux

TABLEAU 3 : DIFFÉRENCES ENTRE TABLES DES FAITS ET TABLES DE DIMENSIONS	23
TABLEAU 4 : COMPARATIF ENTRE LES DEUX MODÈLES	26
TABLEAU 5 : LES ÉTAPES DE LA MÉTHODE DE TRAVAIL GIMSI	28
TABLEAU 6: INDICATEUR VOLUME TOTAL ATTRIBUÉ AU MOIS DE JANVIER 2022	33
TABLEAU 7: INDICATEUR MONTANT TOTAL DES TRANSACTIONS AU MOIS DE JANVIER 2022	33
TABLEAU 8: INDICATEUR VOLUME TOTAL CONSOMMÉ AU MOIS DE JANVIER 2022	34
TABLEAU 9: INDICATEUR REVENU TOTAL DES OFFRES AU MOIS DE JANVIER 2022	34
TABLEAU 10: DESCRIPTION DES DIMENSIONS	34
TABLEAU 11 : DESCRIPTION DIM DATE	35
TABLEAU 12: DESCRIPTION DIM CLIENT	35
TABLEAU 13 : DESCRIPTION DIM OFFRE	36
TABLEAU 14: DESCRIPTION DIM MONTANT	36
TABLEAU 15: DESCRIPTION DIM SERVICE	37
TABLEAU 16: DESCRIPTION DIM CONSOMMATION	37
TABLEAU 17: DESCRIPTION DIM FORFAIT	38
TABLEAU 18: DESCRIPTION DIM TRANSACT	38
TABLEAU 19 : DESCRIPTION FACT TRAFIC OFFRE	39
TABLEAU 20 : DESCRIPTION FACT TRAFIC SERVICE	40
TABLEAU 21 : DESCRIPTION FACT TRAFIC FORFAIT	40
TABLEAU 22 : DÉTECTOR LES DONNÉES SOURCE ET DE DESTINATIONS DE LA DIMENSION "DATE"	44
TABLEAU 23 : DÉTECTOR LES DONNÉES SOURCE ET DE DESTINATIONS DE LA DIMENSION "CLIENT"	44
TABLEAU 24 : DÉTECTOR LES DONNÉES SOURCE ET DE DESTINATIONS DE LA DIMENSION "FORFAIT"	44
TABLEAU 25 : DÉTECTOR LES DONNÉES SOURCE ET DE DESTINATIONS DE LA DIMENSION "SERVICE"	45
TABLEAU 26 : DÉTECTOR LES DONNÉES SOURCE ET DE DESTINATIONS DE LA DIMENSION "CONSOMMATION"	45
TABLEAU 27 : DÉTECTOR LES DONNÉES SOURCE ET DE DESTINATIONS DE LA DIMENSION "TRANSACT"	45
TABLEAU 28 : DÉTECTOR LES DONNÉES SOURCE ET DE DESTINATIONS DE LA DIMENSION "OFFRE"	45
TABLEAU 29 : DÉTECTOR LES DONNÉES SOURCE ET DE DESTINATIONS DE LA DIMENSION "MONTANT"	46
TABLEAU 30 : DÉTECTOR LES DONNÉES SOURCE ET DE DESTINATIONS DES TABLES DES FAITS	46
TABLEAU 31: LA LISTE DES CUBES OLAP	47
TABLEAU 32 : TABLEAU DE DESCRIPTION DES COMPOSANTS SSIS	54

# Introduction Générale

La plupart des entreprises évoluent dans un environnement hautement complexe et concurrentiel. En raison de cette forte concurrence, ces entreprises doivent suivre de près le marché afin de ne pas se laisser distancer par leurs concurrents, tout en ayant une réponse rapide aux attentes du marché, de leurs partenaires et de leurs clients.

Les dirigeants de l'entreprise doivent donc être en mesure de mener à bien les tâches qui leur sont confiées dans ce domaine, quel que soit le domaine d'activité. De plus, ils doivent prendre les décisions les plus appropriées. En effet, ces décisions ne doivent pas être prises à la légère et dans une trop grande précipitation en raison de leur impact important sur la stratégie et la survie de l'entreprise. Il s'agit de prendre des décisions basées sur des informations claires, pertinentes et fiables. Le principal problème est de pouvoir présenter ces informations à ceux qui y ont droit. En effet, les entreprises sont confrontées à une grande quantité de données et les systèmes opérationnels sont limités, voire incapables, de fournir ces informations.

C'est dans ce contexte que les systèmes d'aide à la décision sont apparus. Ils fournissent aux décideurs des informations de qualité sur lesquelles ils peuvent fonder leurs choix décisionnels. Ces systèmes utilisent donc un large éventail de méthodes et de technologies, parmi lesquelles les entrepôts de données constituent le principal moteur pour la mise en place d'un excellent système décisionnel.

L'activité quotidienne de Tunisie Telecom génère des informations importantes et très complexes en raison de sa dimension économique et de sa place dans le marché tunisien. Ces données constituent une source d'information importante qui peut améliorer considérablement le processus de prise de décision. Malgré cela, ces données n'ont pas été correctement exploitées, ce qui entrave le processus de prise de décision à tous les niveaux de l'entreprise.

C'est dans ce contexte que s'inscrit cette étude, qui vise à mettre en place un système qui consolide les données des systèmes transactionnels et fournit des rapports de qualité aux décideurs. Ainsi, les données sont mises à la disposition des décideurs pour faciliter la prise de décision rapide tout en connaissant la cause X. Ainsi, ce système nécessite la mise en place d'un entrepôt de données crédible qui contient les informations nécessaires à la réalisation des processus décisionnels.

Pour ce faire, notre présent rapport s'articule autour de quatre chapitres :

\*Le premier chapitre est intitulé "**Présentation du contexte du projet**". Dans ce chapitre, une brève présentation de l'organisme d'accueil, Tunisie Telecom, est faite, ainsi qu'une analyse détaillée du problème et de la solution proposée.

\*Le deuxième chapitre "**Etude des méthodologies**" : Dans ce chapitre, nous allons étudier les méthodologies et les fondements de la Business Intelligence et de la gestion de projet.

\*Le troisième chapitre "**Définition des besoins et modélisation**" : Dans ce chapitre, nous commencerons par une analyse complète des besoins fonctionnels et non fonctionnels, puis nous aborderons la partie modélisation de notre entrepôt de données.

\*Le quatrième chapitre "**Réalisation et tests**" nous permet de finaliser la présentation de notre projet en montrant l'environnement matériel et logiciel que nous avons adapté lors de la préparation de notre système BI.

Nous clôturons ce rapport en présentant les travaux réalisés.

# **CHAPITRE 1**

## **PRÉSENTATION DU CONTEXTE DU PROJET**

### **1. Introduction**

Au niveau de ce chapitre, nous allons mettre en avant le contexte général dans lequel s'inscrit notre projet, nous commencerons par la présentation de l'organisme d'accueil. Ensuite, nous aborderons la présentation du projet et nous terminerons ce chapitre par une étude de l'existant qui aboutit à une présentation de la solution proposée et nous conclurons par la méthodologie utilisée.

### **2. Présentation de l'organisme d'accueil**

#### **2.1. Présentation générale de l'entreprise Tunisie Télécom**



Figure 1: Logo Tunisie Télécom

Tunisie Télécom est parmi les pionniers des grands opérateurs Télécom de la région. C'est l'opérateur historique en Tunisie (plus de 150 ans). Tunisie Telecom offre des services de communication (téléphonie fixe, mobile et Internet) au grand public, entreprises, FSI et opérateurs tiers...

Fort d'un investissement substantiel dans une infrastructure moderne, des technologies de pointe très haut débit et une large connectivité à l'échelle internationale, Tunisie Telecom compte aujourd'hui plus de 6 millions de clients.

Depuis sa création, Tunisie Telecom a toujours assuré son rôle en tant qu'entreprise citoyenne pour contribuer au développement durable du pays. Tunisie Telecom n'a cessé d'être proche de ses clients en assurant un support soutenu à la culture, au sport, et aux activités liées à l'environnement, l'éducation, l'entrepreneuriat, etc. [Web1]

#### **2.2. L'Organigramme de l'entreprise**

Une vision globale de la structure administrative de l'entreprise est illustrée par l'organigramme suivant :

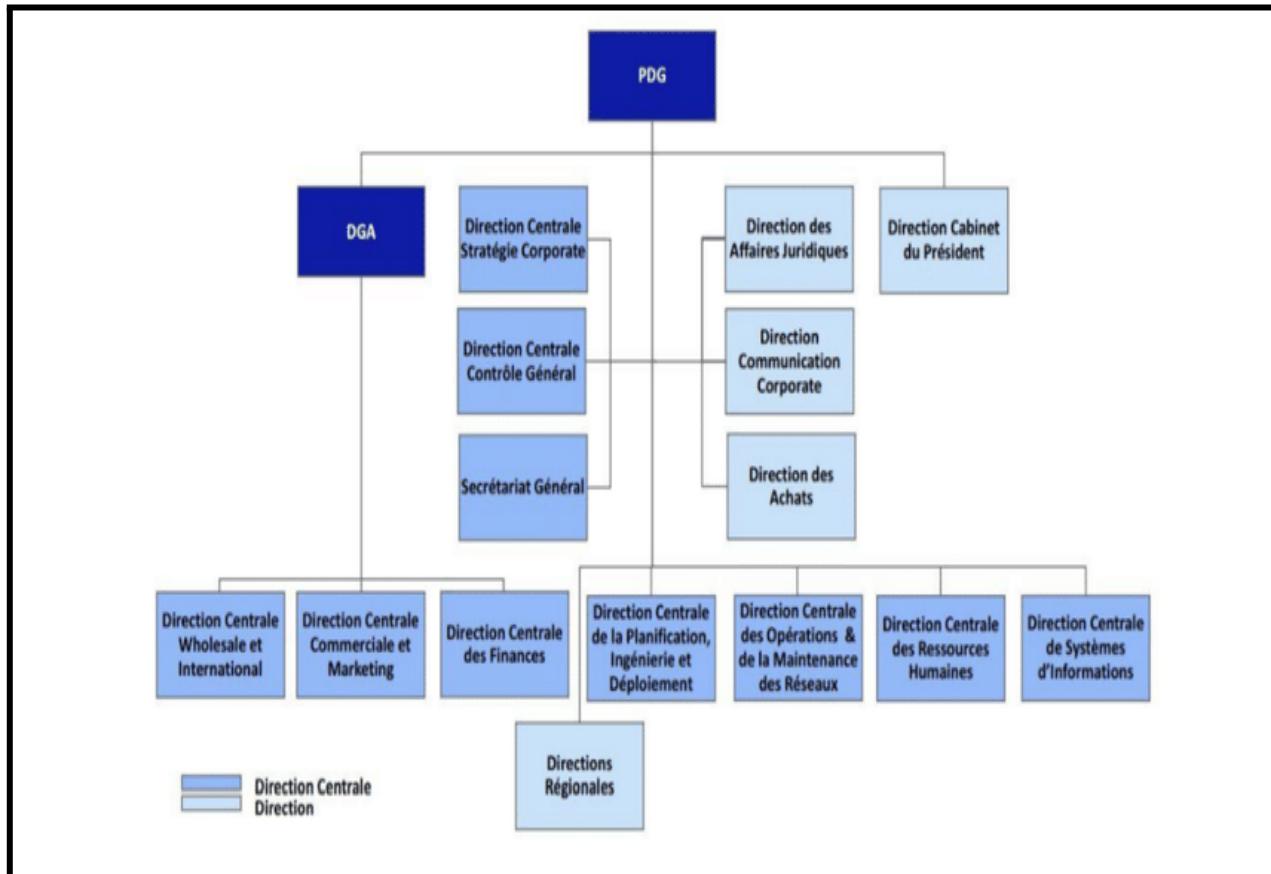


Figure 2: L'organigramme de TT

### 3. Présentation du projet

#### 3.1. Contexte du projet

Une énorme quantité de données circule chaque jour chez Tunisie Télécom. Chaque appel et chaque transaction génère plusieurs données telles que la durée, le numéro, la destination, le nombre d'appels, le volume de données consommées, le montant de la recharge, etc...

Ce projet vise à mettre en place un système de Dashboarding pour faciliter la visualisation des rapports et des indicateurs clés de performance.

#### 3.2. Problématique

Dans le contexte économique moderne caractérisé par la forte mutation qui affecte les différentes faces des marchés, notamment la mondialisation de l'économie, la diversification de la clientèle et la concurrence de plus en plus difficile, l'entreprise Tunisie Telecom veut faire face à tous ces changements instantanés. En effet, elle s'intéresse à une gestion stratégique qui permet de réaliser les tâches suivantes :

1. Extraction et collecte des informations.
2. Réalisation d'un ensemble de tableaux de bord comportant des indicateurs de performance.
3. Conception d'un modèle permettant une historisation des données.

Le traitement de ces besoins présente à l'heure actuelle plusieurs problèmes de limites et de faiblesses qui sont les suivants :

-La saisie des informations par Excel qui peut entraîner des erreurs de calcul et des pertes de temps.

- Une quantité importante de données copiées sur papier à partir des formulaires avec des possibilités d'erreur et un manque de sécurité, cette tâche ne devrait pas être générée manuellement en raison de son énorme besoin de temps et d'efforts physiques. Cette manière manuelle de gérer les données peut entraîner des pertes et des vols de données, car les informations saisies sont vastes, mal organisées et non sécurisées.

- Le système de reporting utilisé chez TT, qui est QlikView, n'est pas aussi clair et facile à apprendre et à utiliser que Power BI puisque Qlik nécessite un développeur pour travailler avec les rapports et les tableaux de bord. De plus, QlikView est un logiciel de BI payant alors que Power BI a une version Desktop qui est gratuite et accessible à tous.

Tous ces problèmes ont incité l'entreprise Tunisie Telecom à nous proposer de concevoir et de développer des tableaux de bord axés sur la détermination et le suivi de l'évolution des indicateurs de performance.

### **3.3. Solution proposée**

La solution que nous proposons sera de créer une application Dashboard d'un système décisionnel lié à la gestion des offres et services de TT qui vise à résoudre les problèmes soulevés et à donner à notre département CVM datamining une vision simple et claire de sa situation. Nous commencerons par extraire les données des différents fichiers sources pour ensuite les charger dans un entrepôt de données (DW) : cette étape est réalisée par un outil ETL. Nous finaliserons ce projet en développant des tableaux de bord qui présenteront les données de manière fiable, minimiseront les coûts d'élaboration des bilans en termes de temps et d'argent et enfin permettront aux dirigeants de prendre les bonnes décisions au bon moment.

## **4. Conclusion**

Le présent chapitre nous a permis de concevoir une solution décisionnelle qui permettra le suivi des indicateurs de Tunisie Telecom en temps réel.

Avant de présenter la conception de la solution, nous expliquerons notre choix des méthodes de gestion des travaux et les concepts de base de la Business Intelligence.

# **CHAPITRE 2**

## **ETUDE DES MÉTHODOLOGIES**

### **1. Introduction**

Dans ce chapitre, nous allons comparer plusieurs méthodologies. Nous commençons par les méthodes de gestion de projet. Nous abordons ensuite les approches de conception d'entrepôts de données, les concepts fondamentaux de la Business Intelligence et enfin nous justifions nos choix.

### **2. Présentation de la méthode de gestion de projet**

Un projet décisionnel n'est pas un projet informatique comme les autres. Il nécessite un savoir-faire approprié et une méthodologie adaptée à un projet dont la réussite dépend avant tout de l'implication des utilisateurs et des aspects fonctionnels. C'est pourquoi nous avons essayé d'adopter une méthodologie qui en tient compte.

#### **2.1 Etude comparative entre les approches de gestion du projet**

##### **2.1.1. Les approches classiques**

Les approches classiques sont divisées en plusieurs phases identifiées au début du projet (Tests et mise en production, définition des besoins, réalisation/développement, conception générale et détaillée).

La validation d'une phase entraîne le début de la suivante. Cette façon de travailler limite les allers-retours entre les phases. Les méthodologies de projet classiques sont axées sur la résistance au changement et la planification du projet de manière globale. Elles sont dites "prédictives".

##### **2.1.2 Les approches agiles**

L'approche agile est à l'opposé de l'approche traditionnelle. Elle est plus flexible et adaptative. Elle est à la fois itérative et incrémentale puisque les besoins du client sont placés au cœur des priorités du début à la fin du projet. Elle prend en compte le fait que le besoin ne peut être fixé et préconise plutôt de bien s'adapter à ses changements.

Un projet d'aide à la décision n'est jamais simple à mettre en œuvre car son objectif premier est de fournir à ses différentes parties prenantes les bonnes informations en temps réel grâce à des solutions qui les aident à prendre des décisions stratégiques pour l'organisation le plus rapidement possible dans un monde en constante évolution.

Le tableau 1 montre les principales différences entre les approches traditionnelles et agiles :

	<b>Approche agile</b>	<b>Approche classique</b>
<b>Objectif</b>	Satisfaire le client	Respecter les compromis et le besoin initial
<b>Changement</b>	Soutenir le changement	Non adapté au changement
<b>Equipe</b>	Livrer régulièrement	Livrer une fois seulement
<b>Communication</b>	Communiquer directement avec le staff opérationnel	Chacun fait sa part du travail
<b>Amélioration</b>	Intégrer la notion de progrès continu durant tout le projet	Possibilité d'amélioration, mais uniquement à la clôture du projet.

Tableau 1 : Etude Comparative entre les approches agiles et classiques

Nous pouvons constater que l'application d'une approche agile à notre projet ne peut être que favorable puisqu'elle contribue à la réduction des coûts de réalisation et à la scalabilité des besoins fonctionnels.

## 2.2. Etude comparative entre les méthodologies de gestion du projet

Nous avons choisi d'adopter une approche agile, il reste à choisir la méthode la plus appropriée pour notre projet.

Voici un tableau comparatif entre Scrum, qui est la méthode la plus utilisée, et GIMSI, une méthode pour effectuer et concevoir le système d'aide à la décision.

Critère	Scrum	GIMSI
<b>Planification</b>	Au début de chaque sprint.	Flux continu.
<b>Estimation de l'effort</b>	Au début de chaque sprint.	Facultatif, prévisibilité.
<b>Changement de périmètre</b>	Il faut attendre le prochain sprint	Selon les besoins.
<b>Rôles</b>	Scrum master, product owner et développeurs.	Chacun fait sa part du travail.
<b>Caractéristique principale</b>	Basé sur les sprints.	Pilotage visuel.
<b>Les bénéfices majeurs</b>	Productivité, évolutivité et engagement de l'équipe.	Mise en œuvre immédiate sans modifier les processus actuels et pilotage visuel.

Tableau 2: Etude comparative entre les 2 approches Scrum et GIMSI

En général, même si les deux méthodes sont différentes, elles ne sont pas vraiment opposées.

### Inconvénients de Scrum:

Comme tout cadre de travail, Scrum a aussi ses limites.

En effet, rien n'est parfait et la méthodologie Scrum ne fait pas exception. Dans certains cas, Scrum est associé à d'autres techniques de gestion de projet qui permettent de remédier à certains de ces inconvénients :

- Scrum mène fréquemment à la dérive du projet, en raison de l'absence d'une date de fin claire.

- Le risque d'échec du projet est fort si les personnes ne sont pas très engagées ou coopératives.
- Le cadre ne peut être réussi qu'avec des membres d'équipe expérimentés
- Les réunions journalières peuvent parfois frustrer les membres de l'équipe
- Le départ d'un membre de l'équipe au milieu d'un projet peut avoir un impact négatif important sur le projet.
- La qualité est difficile à mettre en place tant que l'équipe ne suit pas un processus de test dynamique.

### 3. Architecture du système décisionnel :

Tout système de BI est composé de trois couches, à savoir : l'ETL, l'entrepôt de données et la couche de restitution, qui sont présentées dans le schéma ci-dessous :

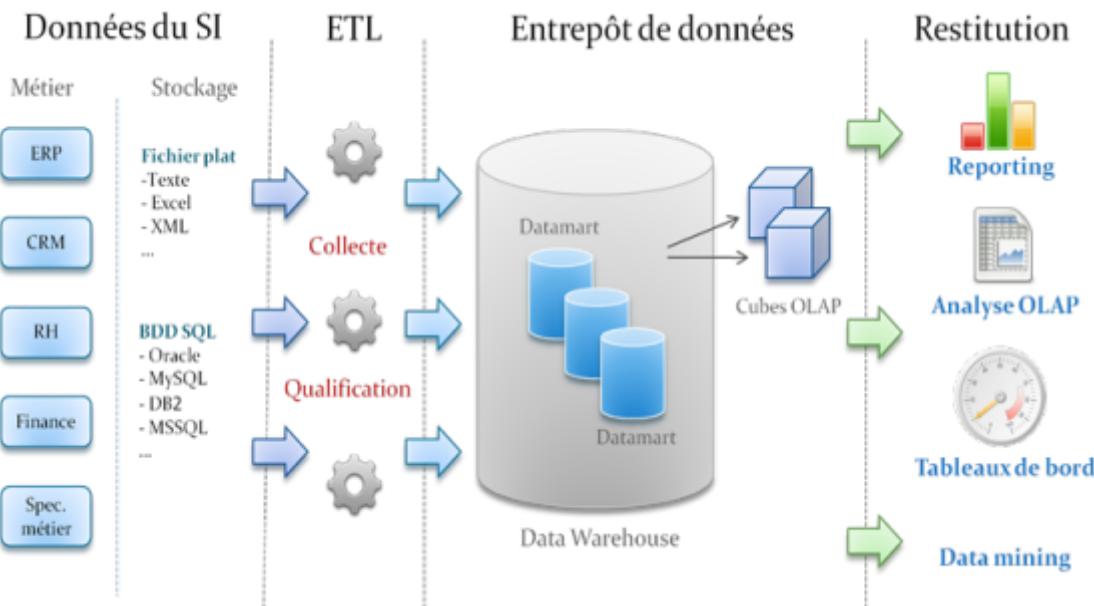


Figure 3: Architecture du système Décisionnel

- **L'ETL (Extract, Transform, Load)** est un processus complet de fusion de données qui envoie des informations brutes d'une source de données, les transforme et les nettoie en vue d'une utilisation future, puis les envoie vers un serveur cible ou un entrepôt de données. Dans ce phénomène, la transformation des données passe par un serveur intermédiaire avant d'être chargée dans la cible.

- **La zone de stockage des données** : Un entrepôt de données, ou autrement dit DataWarehouse, rassemble toutes les données de l'entreprise. C'est une structure dont l'objectif, contrairement aux bases de données, est de regrouper toutes les données de l'entreprise pour les analyser et faciliter la prise de décision stratégique. La décision stratégique est prise par les décideurs de l'entreprise et vise à améliorer les performances de l'entreprise, que ce soit en termes qualitatifs ou quantitatifs. Fondamentalement, il s'agit d'une énorme quantité d'informations bien organisée, historique et purgée provenant de multiples sources de données, qui facilite l'analyse et aide à la prise de décision. L'entrepôt de données est la base de l'aide à la décision.

En effet, l'entrepôt de données est le meilleur moyen que les élites ont trouvé pour modéliser l'information à des fins analytiques.

- **La zone de visualisation des données** qui comprend tous les outils permettant de générer des rapports ou des tableaux de bord. Il permet également de générer des données pour de futures tâches de machine learning ou de datamining.

#### 4. Comparaison entre l'approche TOP-DOWN et BOTTOM-UP

Lors de la conception d'un entrepôt de données, deux approches s'affrontent. Ces approches sont celles de Bill Inmon et de Ralph Kimball (figure 3).

Chacune de ces approches est adaptée à un environnement d'intégration et possède des avantages et des inconvénients.



Figure 4 : TOP-DOWN vs BOTTOM-UP

##### 4.1. Approche de Bill Inmon

Dans les années 1990, William Inmon propose une vision descendante de la structuration de la donnée de l'entreprise, dite "Top-down".

###### Avantages :

- Des données complètes
- Données cohérentes

###### Inconvénients :

- La mise en place prend beaucoup de temps
- Cette approche nécessite des experts pour bien gérer un entrepôt de données.

Selon l'approche Top-Down, l'entrepôt de données est :

- **Orienté sujet** : Les données sont triées par mesures et analysées par dimensions.

- **Non-volatile** : Toutes les données sont conservées, sans être écrasées pour l'analyse.
- **Intégré** : La collecte des données se fait dans tout le système opérationnel de l'entreprise.

### **Entrepôt de données “Data Warehouse”**

L'entrepôt de données est une technologie qui permet de regrouper des données structurées à partir d'une ou de plusieurs sources dans le but de les comparer et de les analyser.

Les entrepôts de données reposent sur une conception différente de celle des bases de données classiques. Ils sont optimisés pour maintenir une précision rigoureuse des données en les mettant à jour en temps réel. Les entrepôts de données sont destinés à fournir une vue à long terme des données dans le temps. Ils sacrifient le volume des transactions et se spécialisent plutôt dans lagrégation des données.

La figure suivante montre le processus BI selon l'approche de Bill Inmon

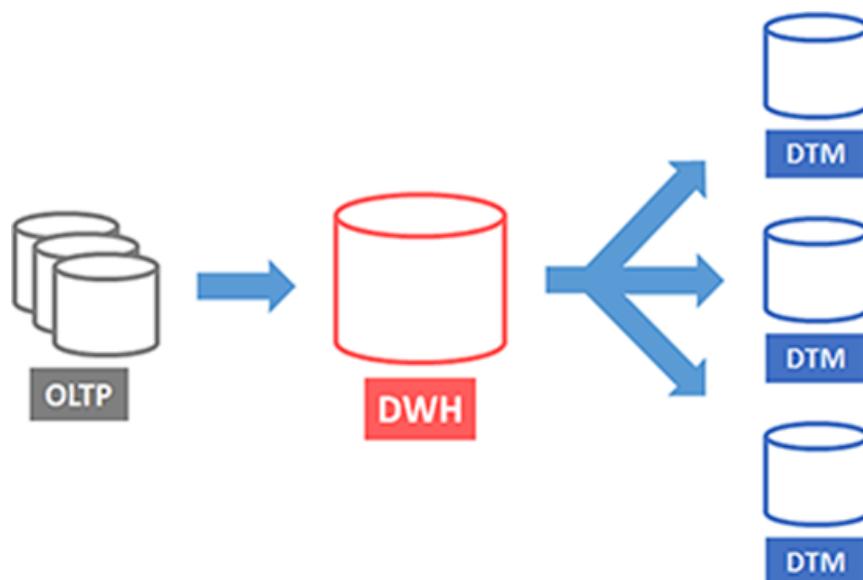


Figure 5: Approche de Bill Inmon

### **4.2. Approche de Ralph Kimball**

Ralph Kimball propose une vision radicalement différente et ascendante de la structuration des données, dans laquelle les besoins de l'entreprise sont représentés dans des magasins de données (Datamarts) et l'entrepôt de données est alors une simple union de ces magasins de données.

#### Avantages :

- Facile à réaliser
- Résultats rapides
- Efficace à court terme

#### Inconvénients :

- Pas très efficace à long terme

- La quantité de travail d'intégration pour réaliser un entrepôt de données
- Risque de redondances

### **Magasin de données “Datamart”**

Un magasin de données ou Datamart est une vue spécifique, orientée vers un sujet, d'une organisation. C'est un sous-ensemble d'un entrepôt de données conçu pour fournir des données aux utilisateurs et généralement spécialisé pour un groupe ou un type de travail. Il est constitué de tables détaillées ou groupées qui sont reliées entre elles.

Il peut y avoir plusieurs datamarts au sein d'une organisation.

La figure 5 illustre le processus BI selon l'approche de Ralph Kimball

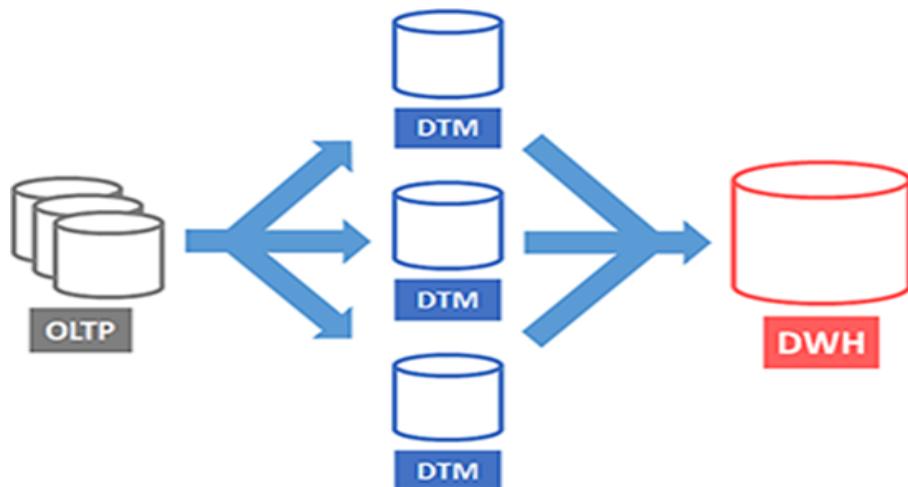


Figure 6: Approche de Ralph Kimball

## **5. Composants d'un schéma dimensionnel**

Chaque modèle dimensionnel est composé d'une table contenant la table de faits, et un ensemble de tables plus petites appelées tables dimensionnelles.

### **5.1. Tables de faits**

- Stocker les indicateurs analysés (ex : quantité vendue, prix de vente, chiffre d'affaires (CA)...).
- La table de faits est toujours au cœur du modèle. Ce sont des tables SQL avec des liens vers chacune des tables de dimension qui décrivent la ou les mesures de la même granularité.
- Un fait est tout ce que l'on veut analyser.
- Les mesures font partie des tables de faits : ce sont des valeurs quantitatives que les dimensions permettent d'analyser. Une mesure a sa propre fonction d'agrégation : SUM, MAX, COUNT ()...

Vente	
PK	ID vente
FK1	ID_produit
FK2	ID_magasin
FK3	ID_client
FK4	ID_promotion
FK5	ID_date quantite prix unitaire

Figure 7: Exemple d'une table de fait "Vente"

## 5.2. Tables de dimensions

Ce sont des tables SQL qui ont un identifiant à valeur unique (la clé de substitution), sans signification fonctionnelle, avec des colonnes d'attributs qui décrivent cet identifiant.

Exemple : dimension customer(customerId(INT, auto incremental), customerName, customerAddress, customerPurchases...).

Attribut de dimension : Un champ descriptif d'un élément de la dimension c'est le nom du produit, sa marque. Chaque produit a une valeur unique pour chacun de ces attributs. Il est nécessaire de respecter une organisation hiérarchique des attributs de dimension.

Exemple : un code postal se trouve dans un seul département, qui lui-même se trouve dans une seule région.

DimDate	
» DateKey	
FullDateAlternateKey	
DayNumberOfWeek	
EnglishDayNameOfWeek	
SpanishDayNameOfWeek	
FrenchDayNameOfWeek	

Figure 8: Exemple d'une table de dimension "Date"

## 5.3. La différence entre tables des faits et tables des dimensions

Le tableau 3 montre en détail la différence entre les tables de faits et de dimensions.

	<b>Table de faits</b>	<b>Table de dimensions</b>
<b>Clé primaire</b>	La table de faits contient une clé primaire qui est une concaténation de clés primaires de toutes les tables de dimensions	Chaque table de dimension contient sa clé primaire.
<b>Signification</b>	La table de faits contient les mesures avec des attributs d'une table de dimension.	La table de dimension contient les attributs avec lesquels la table de faits calcule la métrique.
<b>Taille de la table</b>	La table de faits se développe verticalement.	La table de dimensions se développe horizontalement.
<b>Attribut &amp; Record</b>	La table de faits contient moins d'attributs et plus d'enregistrements.	La table de dimension contient plus d'attributs et moins d'enregistrements
<b>Création</b>	La table de faits peut être créée uniquement lorsque les tables de dimensions sont complétées.	Les tables de dimension doivent être créées en premier.
<b>Schéma</b>	Un schéma contient moins de tables de faits.	Un schéma contient plus de tables de dimension.
<b>Les attributs</b>	La table de faits peut contenir des données au format numérique et au format textuel.	La table de dimension contient toujours des attributs au format textuel.

Tableau 3 : Différences entre tables des faits et tables de dimensions

## 5.4. Schémas de la modélisation dimensionnelle

Il existe différents types de modèles de données qui sont mis en œuvre dans les entrepôts de données et qui utilisent bien évidemment les notions de table de faits et de table de dimensions.

### 5.4.1. Le modèle en étoile

Un modèle en étoile est constitué d'une table centrale (Table de fait) et de nombreuses tables de dimensions autour.

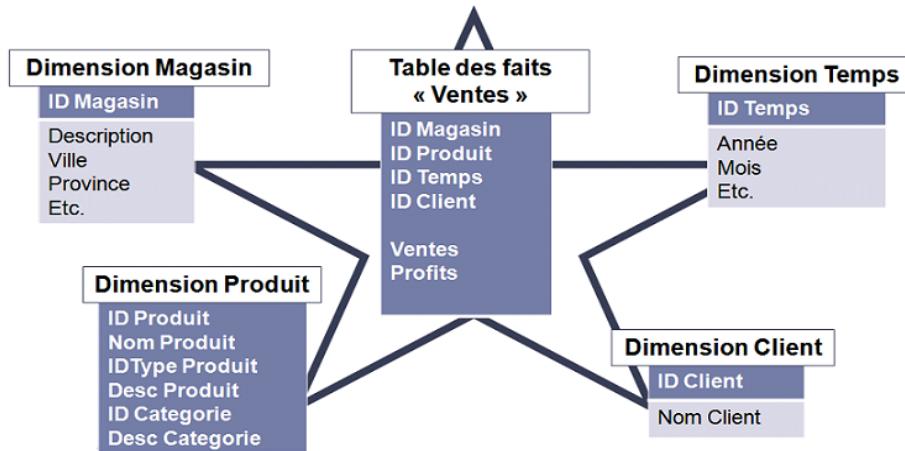


Figure 9: Modèle en étoile

#### 5.4.2. Le modèle en flocon

Un modèle en flocon se compose d'une table de faits qui se trouve toujours au centre du modèle et de tables de dimensions qui se situent autour de la table centrale. Le modèle en flocon évite la redondance des informations mais il nécessite des jointures.

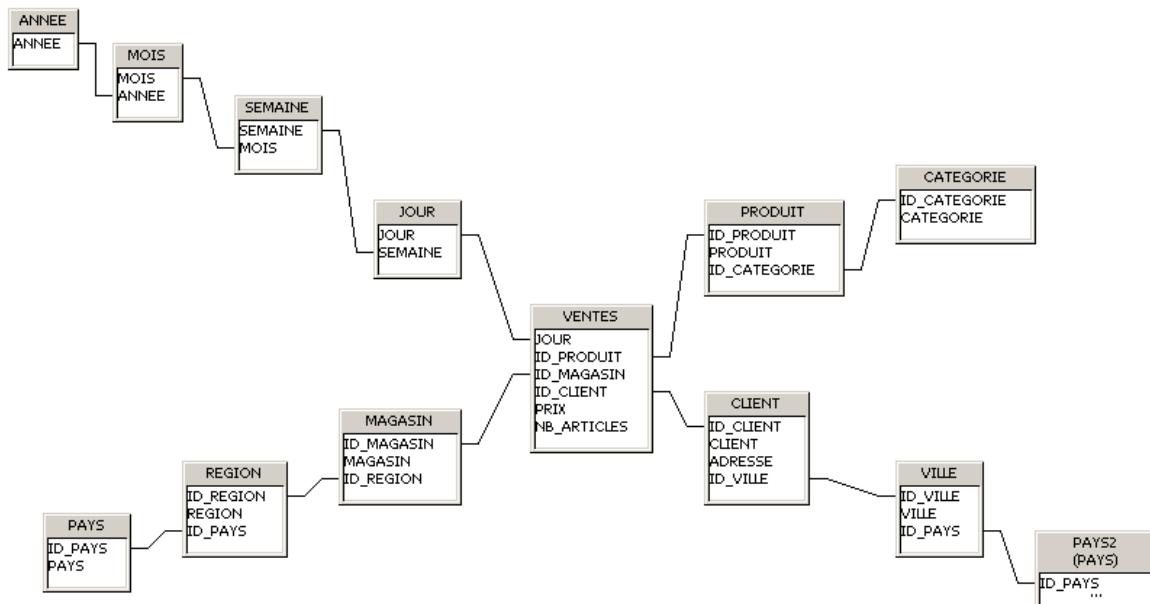


Figure 10: Modèle en flocon

#### 5.4.3. Le modèle en constellation

Le schéma de constellation se compose de plusieurs tables de faits qui partagent des tables de dimensions communes et chaque table de fait possède ses propres dimensions.

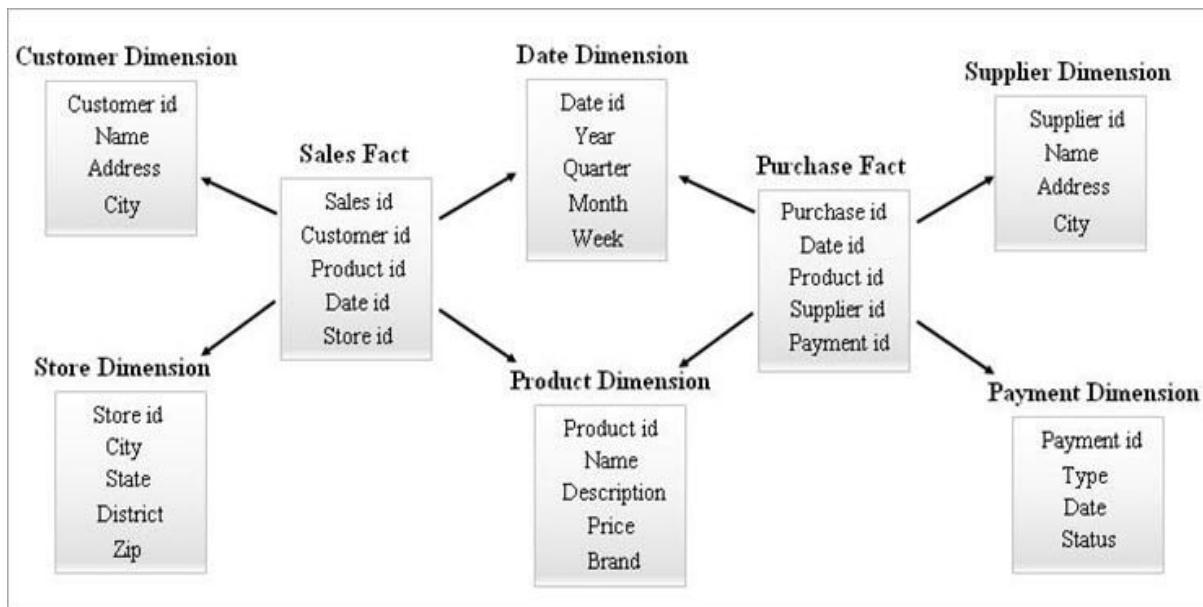


Figure 11: Modèle en constellation

## 5.5. Alimentation de l'entrepôt de données

L'alimentation des données à partir des bases de production est une phase primordiale d'un datawarehouse. Des outils logiciels sont alors nécessaires pour intégrer les données dans le datawarehouse. On parle d'ETL (Extract, Transform, Load).[Web3]

Il est utilisé pour alimenter l'entrepôt de données à partir des bases de données opérationnelles. Il est également utilisé pour importer, nettoyer et extraire des données de différentes sources de données et les charger régulièrement dans un entrepôt de données.

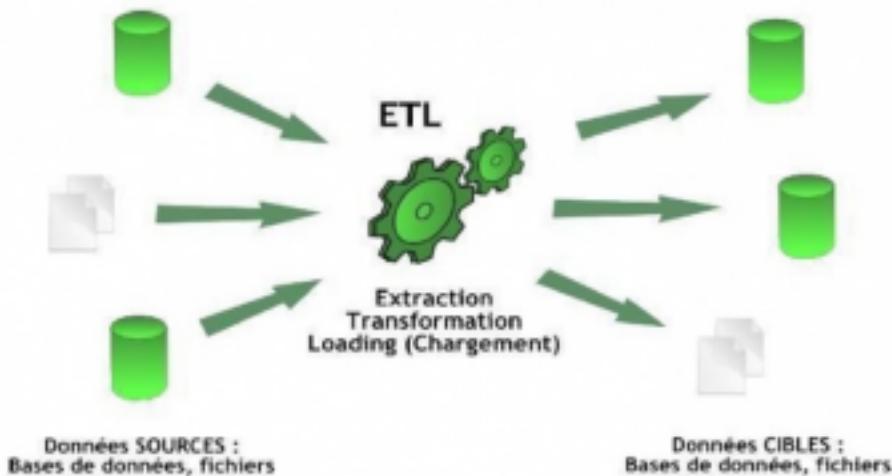


Figure 12: Processus ETL

## 5.6. Exploitation de l'entrepôt de données

C'est la phase d'analyse qui crée des cubes à partir de l'entrepôt de données. Cette phase contient des outils OLAP pour la restitution des données sous différentes formes (graphiques ou tableaux), le data mining et l'analyse des données.

### 5.6.1. Types de modélisation OLAP

SQL Server Analysis Services propose plusieurs méthodes pour concevoir un modèle de BI sémantique : tabulaire et multidimensionnel. Le fait de disposer de plusieurs approches offre une expérience de modélisation adaptée aux différents besoins des entreprises et des utilisateurs. Ces deux modèles sont très différents l'un de l'autre, il n'y a pas de choix prédefini ou de solution plus efficace qu'une autre. Chacun a ses propres spécificités qui vous permettront de choisir celui qui correspond à vos besoins et à votre environnement de travail.

- **Multidimensionnel** : avec le multidimensionnel, nous disposons d'une technologie éprouvée mais plus complexe permettant de gérer de grands volumes de données, avec la gestion de pré-agrégations et d'accès hiérarchiques. En revanche, elle nécessite plus de compétences pour un bon paramétrage et une utilisation de fonctions analytiques avancées.
- **Tabulaire** : pour le tabulaire, on se concentrera sur des scénarios plus relationnels avec des volumes plus raisonnables, et de meilleures performances au niveau final. Dans ce cas, il y a peu de place pour le paramétrage, avec des possibilités de calcul limitées et un moteur assez jeune, simple et facile à mettre en œuvre.

	Modèle tabulaire	Modèle multidimensionnel
Avantages	<ul style="list-style-type: none"><li>- Plus facile à développer que le modèle multidimensionnel.</li><li>- Technologie pour l'avenir</li></ul>	<ul style="list-style-type: none"><li>- Technologie mature.</li><li>- Technologie capable de gérer un très grand volume de données.</li><li>- Capable de faire face aux exigences de Modélisation et calcul avancées.</li></ul>
Inconvénients	<ul style="list-style-type: none"><li>- Nouveau produit donc encore des améliorations à venir.</li><li>- Limite à la RAM disponible.</li><li>- Il manque des calculs avancés disponibles avec MDX</li></ul>	<ul style="list-style-type: none"><li>- Aucune innovation majeure à attendre dans ce produit à l'avenir.</li><li>- Plus complexe que le tabulaire.</li></ul>

Tableau 4 : Comparatif entre les deux modèles

### 5.6.2. Restitution

La restitution est une partie très importante pour l'utilisateur. Les outils de reporting représentent l'interface par laquelle les utilisateurs interagissent avec le système BI afin d'obtenir des résultats.

En effet, l'objectif principal des systèmes BI est de fournir des résultats aux demandes des utilisateurs, avec un temps de réponse très court et sans connaissance avancée du domaine informatique.

Dans ce qui suit, nous proposons d'analyser quelques outils de navigation et de visualisation de données.

- **Tableau de bord** : Les tableaux de bord aident l'utilisateur à avoir une vue d'ensemble des informations les plus importantes en un coup d'œil. Ils peuvent donc fournir une solution efficace pour une énorme quantité de données et permettre une synthèse immédiate.

- **Reporting** : c'est un outil qui donne la possibilité aux managers d'avoir un regard approfondi sur une activité, un département ou l'ensemble de l'entreprise.

De manière générale, il s'agit de la production de chiffres et d'indicateurs bruts, en laissant la transposition et l'analyse aux décisionnaires.

## 6. Récapitulatif des choix

Après avoir terminé la comparaison entre les différentes méthodologies et modèles, nous procédons maintenant au choix de la méthodologie.

### 6.1. Choix de méthodologie

Après avoir sélectionné l'approche agile, nous nous sommes attachés à choisir la méthode la plus adaptée à notre projet.

Après recherche, nous avons choisi la méthodologie GIMSI car elle est la plus appropriée à la mise en œuvre de notre projet décisionnel. La méthode de conception GIMSI se focalise sur une question clé : "Comment faciliter la perception du risque dans l'incertitude pour prendre les meilleures décisions sur le terrain ?".

Face à la finesse des données, le problème des décisions doit nécessairement être revisité, en considérant l'importance primordiale des aspirations et des attentes personnelles de chacun des acteurs du terrain. C'est ainsi que les décideurs pourront prendre des décisions et évaluer les risques.

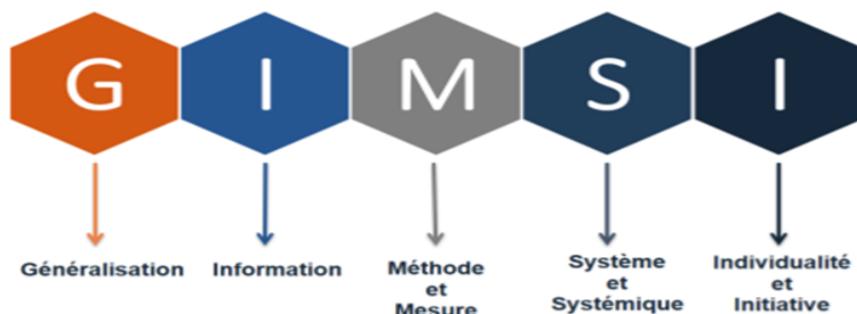


Figure 13: Méthode GIMSI

Un tableau de bord est constitué d'un ensemble d'indicateurs définis de manière consensuelle, de méthodes de collecte d'informations et de méthodes opérationnelles (utilisation des résultats) permettant d'atteindre les objectifs de gestion suivants :

- **Généralisation** : La méthode GIMSI est utilisée dans plusieurs domaines.
- **Information** : L'accès aux informations pertinentes est la base de l'aide à la décision.
- **Méthode et Mesure** : La GIMSI est une méthode, la mesure en est le principe.
- **Système et Systémique** : La méthode permet la construction du système de management et son intégration au cœur du système d'information. Elle s'appuie sur un concept systémique.
- **Individualité et Initiative** : La méthode GIMSI favorise l'autonomie de l'individu pour une prise d'initiative naturelle.

La méthode GIMSI est divisée en 10 étapes, représentées dans la figure suivante, dont chacune répond à une préoccupation particulière du projet de tableau de bord. Chacune des 10 étapes correspond à une phase identifiable dans l'évolution du système.

<b>Phase</b>	<b>N°</b>	<b>Etape</b>	<b>Objectifs</b>
<b>Identification</b> <b>Quel est le contexte ?</b>	<b>1</b>	Environnement de l'entreprise	Analyse de la stratégie et de l'environnement économique de l'entreprise afin de préciser la portée du projet.
	<b>2</b>	Identification de l'entreprise	Analyse des structures de l'entreprise afin de déterminer les processus, les activités et les acteurs impliqués.
<b>Conception</b> <b>Que faut-il faire ?</b>	<b>3</b>	Identification des objectifs	Sélection des objectifs tactiques à atteindre par chaque équipe
	<b>4</b>	Construction du tableau de bord	Définition d'un tableau de bord pour une aide à la décision efficace.
<b>Mise en œuvre</b> <b>Comment le faire ?</b>	<b>5</b>	Choix des indicateurs clés de performance	Choix des indicateurs en cohérence avec les objectifs sélectionnés
	<b>6</b>	Collecte des informations	Identification des informations requises pour le tableau de bord de pilotage.
	<b>7</b>	Le système de tableau de bord	Mise en place de tableaux de bord, vérification de la cohérence du système de pilotage pour une prise de décision efficace.
	<b>8</b>	Le choix des logiciels	Analyse des besoins technologiques dans le domaine de la Business Intelligence, développement d'une méthode de choix des outils d'aide à la décision.
<b>Amélioration continue</b>	<b>9</b>	Intégration et déploiement	Mise en œuvre de logiciels, déploiement dans l'entreprise
	<b>10</b>	Audit du système	Contrôle continu du système Le système répond-il toujours aux besoins exprimés ?

Tableau 5 : Les étapes de la méthode de travail GIMSI

Pour faciliter l'étude, les dix étapes peuvent être classées en 4 phases essentielles :

- **Identification** : répond à la question "Quel est le contexte ?", dans cette phase la stratégie de l'entreprise et l'environnement économique seront analysés.
- **Conception** : répond à la question " Que faut-il faire ? ", cette phase comprend la définition des besoins et des indicateurs, le choix des informations utiles à collecter et la création de tableaux de bord.
- **Mise en œuvre** : répond à la question " Comment faire ? " Cette phase inclut le choix du logiciel approprié ainsi que son installation, son déploiement et son implantation au sein de l'entreprise.
- **Amélioration continue** : répond à la question " Le système répond-il toujours aux besoins ? " Cette phase comprend un audit et un contrôle permanent du système.

## 6.2. Choix de la démarche

Après étude, nous avons constaté que l'approche de Ralph Kimball est la mieux adaptée à nos besoins. Elle nous fournit une copie parfaitement structurée des données de transaction, ce qui nous permet d'effectuer des requêtes d'analyse.

### 6.3. Choix du type de modèle

Pour le type de modèle, nous avons opté pour le modèle en constellation qui nous conduit à l'analyse par la hiérarchie des dimensions et puisque nos faits partagent plusieurs dimensions.

## 7. Planification prévisionnelle

La bonne planification du projet guide le bon déroulement de ses différentes tâches, c'est pourquoi la phase de planification est très importante. Elle nous permet de planifier et de spécifier les tâches du projet et d'estimer leurs charges respectives. Et afin de planifier notre projet nous avons utilisé le diagramme de GANTT qui nous aide à faciliter et à simplifier le suivi de notre projet et nous permet de visualiser l'ordre et la durée de chaque tâche tout au long de notre stage, la figure 14 représente les principales missions à accomplir durant notre projet

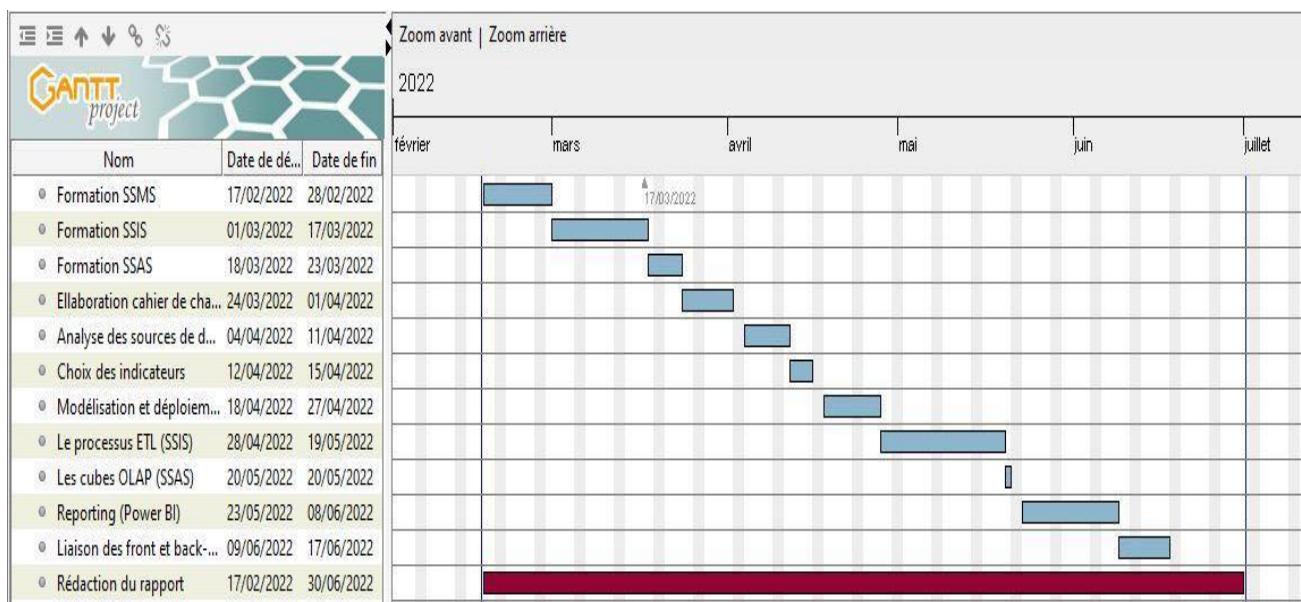


Figure 14: Diagramme de GANTT

## 8. Conclusion

Dans ce chapitre nous avons fait une étude comparative détaillée de toutes les notions liées aux méthodes de gestion de projet et à la Business Intelligence, ainsi que nous avons choisi et fixé les approches pour le bon déroulement du projet. Dans le prochain chapitre, nous définirons les besoins des utilisateurs et ensuite nous traiterons les étapes de chargement de notre ETL et des cubes OLAP.

# CHAPITRE 3

## DÉFINITION DES BESOINS ET MODÉLISATION

### 1. Introduction

Dans ce chapitre, nous allons spécifier les besoins fonctionnels, les besoins non-fonctionnels de l'entreprise ainsi que le diagramme des cas d'utilisation et par la suite nous nous occuperons des étapes de chargement de notre ETL et des cubes OLAP.

### 2. Spécification des besoins

La spécification des besoins est une phase initiale importante à réaliser avant de passer à la phase de modélisation. Elle permet de définir les besoins fonctionnels, les besoins non fonctionnels et le diagramme des cas d'utilisation afin de bien organiser notre projet.

#### 2.1. Acteurs

Dans notre application, les acteurs qui interagissent avec le système sont les suivants :

- **L'administrateur** : il est chargé de l'intégration des données et de la création des cubes OLAP et des tableaux de bord.
- **Le décideur** : il s'agit de la personne qui consultera les rapports et les informations résultant du processus ETL et apparaissant sur le tableau de bord de l'application afin de prendre les décisions nécessaires liées au département CVM datamining.

#### 2.2. Besoins fonctionnels

- **L'administrateur** :

- Il doit intégrer les données :

L'application du processus ETL est sa préoccupation fondamentale. En effet, cet acteur est chargé d'extraire d'abord les données de différentes sources puis de transformer ces données extraites. Le chargement de ces informations dans l'entrepôt de données a ensuite lieu.

Une fois l'intégration réalisée avec succès, l'administrateur devra passer à la sélection et à la spécification des axes d'analyse (cubes OLAP), ce qui lui permettra de rédiger les rapports puis de les charger au niveau des tableaux de bord, d'où l'administrateur s'occupe également de la gestion des rapports (ajout, suppression, modification, affichage).

- Il permet aux décideurs du département CVM d'accéder aux rapports réalisés via un formulaire à accès privé.

- **Le décideur** :

Cet acteur dispose de sa propre session lui permettant de visualiser les informations décisionnelles liées à la direction CVM datamining de TT, de modifier les filtres des tableaux de bords et de prendre les bonnes décisions au bon moment.

Ces informations sont classées en 3 catégories : Les offres, les forfaits et les services.

## 2.3. Besoins non fonctionnels

Notre solution requiert plusieurs autres besoins non fonctionnels, qui doivent être satisfaits, tels que :

- o La rapidité

La durée d'extraction et de modification des données est aussi proche que possible du temps réel. L'objectif principal de notre application est d'assurer un gain de temps.

- o La fiabilité

Les rapports produits seront précis, clairs et basés sur des données exactes.

- o La performance

Il s'agit de présenter toutes les fonctionnalités qui répondent à toutes les attentes des utilisateurs de manière optimale et d'assurer l'optimisation du temps de traitement des données.

## 2.4. Modélisation des besoins

Ce schéma représente le diagramme des cas d'utilisation pour structurer les besoins fonctionnels et non fonctionnels de Tunisie Télécom, comme le montre la figure 15

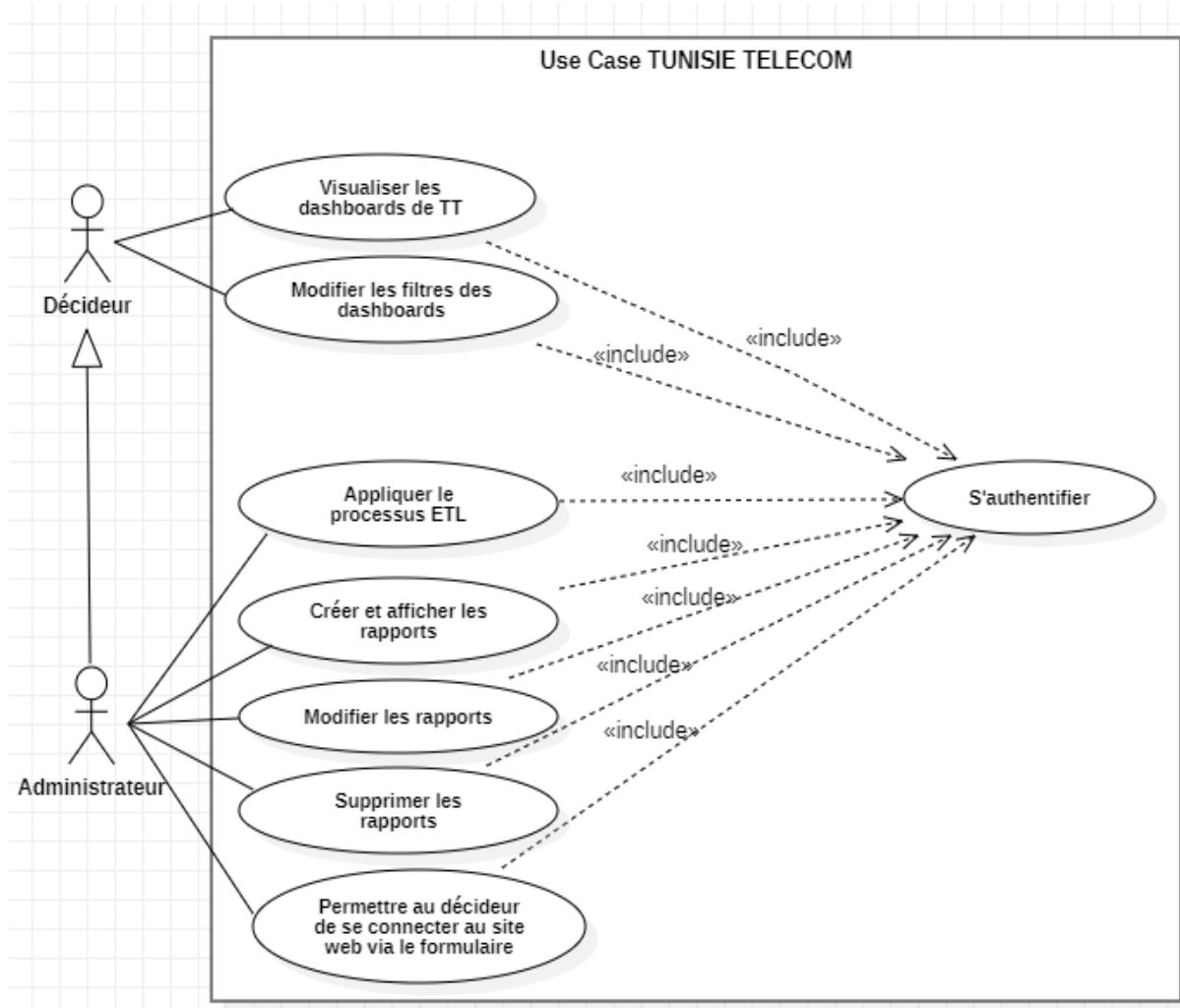


Figure 15: Diagramme du cas d'utilisation

### 3. Architecture

Notre solution couvre l'ensemble de la chaîne décisionnelle, de l'extraction des données au déploiement des états.

#### 3.1. Architecture fonctionnelle du système

L'architecture fonctionnelle permet ainsi de visualiser la corrélation entre les fonctions et services du système d'information et les processus métiers des utilisateurs. Elle est également utilisée pour initier la structuration à un niveau purement logique du système technique sous-jacent.

L'architecture fonctionnelle de notre système peut être représentée comme suit :

- Collecte des données de base à partir de différentes sources (excel, csv, texte).
- Extraction, transformation et chargement des données collectées dans un entrepôt de données dans le but de capturer des informations pour la prise de décision.
- L'entrepôt de données chargé est composé de trois Data Marts.
- Assurer la restitution des différentes données et déterminer comment les tableaux de bord seront présentés à l'utilisateur final en suivant ses besoins.

#### 3.2. Architecture logique des données

L'architecture logique détermine les composants logiciels qui doivent être utilisés pour mettre en œuvre une solution et décrit les relations entre ces composants. L'architecture logique et les besoins en matière de qualité de service établis lors de la phase de spécification technique forment un scénario de déploiement.

Quatre zones de données logiques sont définies et leur organisation est illustrée dans la figure 16.

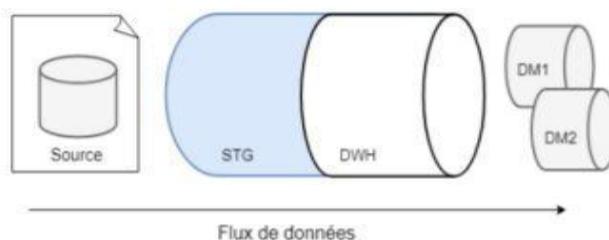


Figure 16: Architecture logique des données

- **Sources** : Les sources des données sont tous les systèmes externes qui leur fournissent des informations brutes.
- **STG (Staging)** : Il s'agit des métadonnées et des informations techniques nécessaires au traitement des données.
- **Entrepôt de données (DWH)** : Il s'agit de la principale zone de chargement. Toutes les données consolidées sont chargées ici.
- **Datamart** : Stocke les informations provenant de l'entrepôt de données et préparées à des fins spécifiques.

### **3.3. Choix des indicateurs**

#### **3.3.1. Caractéristiques des indicateurs clé de performances**

Un indicateur clé de performance (ICP, ou KPI en anglais) est une mesure quantitative qui vous permet de suivre la progression de votre équipe ou organisation au regard de vos objectifs commerciaux clés. [Web4]

Pour être efficace, votre ICP doit être à la fois spécifique et mesurable. Les ICP vous servent à établir clairement les objectifs que votre équipe veut atteindre et l'échéance visée, ainsi qu'à préciser la façon dont vous mesurerez leur accomplissement.

Pour être efficace, un ICP doit vous aider à atteindre vos objectifs stratégiques, guider votre plan de gestion des ressources, être mesurable et concerner un élément que vous pouvez contrôler et influencer... [Web4]

#### **3.3.2. Le choix des indicateurs clé de performances**

Les tableaux suivants présentent quelques indicateurs clés de performance à surveiller en permanence :

<b>Volume data attribué par TT par :</b>	<b>Forfait</b>
	<b>Client</b>
	<b>Age client</b>
	<b>Région</b>
	<b>Segment</b>
	<b>Handset</b>

Tableau 6: Indicateur Volume total attribué au mois de Janvier 2022

<b>Montant total des transactions par :</b>	<b>Forfait</b>
	<b>Client</b>
	<b>Age Client</b>
	<b>Région</b>
	<b>Segment</b>
	<b>Handset</b>

Tableau 7: Indicateur Montant total des transactions au mois de Janvier 2022

<b>Volume total consommé par :</b>	Service
	Client
	Age Client
	Région
	Segment
	Handset
	Client actif

Tableau 8: Indicateur volume total consommé au mois de Janvier 2022

<b>Revenu total des offres par :</b>	Offre
	Client
	Age Client
	Région
	Segment
	Handset
	Date

Tableau 9: Indicateur Revenu total des offres au mois de Janvier 2022

## 4. Modélisation de notre Data Warehouse

Nous avons choisi le modèle en constellation car il est le plus approprié pour la réalisation de notre projet.

### 4.1. Description des dimensions

A ce stade, différentes dimensions seront bien choisies afin qu'elles correspondent bien au même contexte des tables de faits.

Nous avons créé les dimensions et leurs descriptions dans le tableau ci-dessous

<b>Nom de la dimension</b>	<b>Description</b>
DIM_DATE	Représente la liste des dates
DIM_CLIENT	Représente les données des clients
DIM_OFFRE	Représente la liste des offres de TT
DIM_SERVICE	Représente la liste des services(Viber, Youtube,...)
DIM_FORFAIT	Représente la liste des forfaits TT
DIM_MONTANT	Représente les montants et les revenus des offres
DIM_TRANSACT	Représente les transactions effectuées par les clients
DIM_CONSUMMATION	Représente la consommation de services des utilisateurs

Tableau 10: Description des dimensions

#### 4.1.1. Dimension Date:

La dimension date est la seule dimension qui figure systématiquement dans tout entrepôt de données, car tout entrepôt de données est une série temporelle. Dans la plupart des cas, le temps est la première dimension dans la classification des bases de données BI.

Date	
⚠	ID_DATE
	DATE_ACTIVATION_DATA
	DERNIERE_DATE_DATA

Figure 17 : DIM Date

Nom du champ	Description
ID_DATE	Représente le code des dates
DATE_ACTIVATION_DATA	Représente la date d'activation data au mois de Janvier
DERNIERE_DATE_DATA	Représente la dernière date de connexion au mois de Janvier

Tableau 11 : Description DIM Date

#### 4.1.2. Dimension Client

La dimension client sera configurée comme une table de dimension qui contiendra les informations présentées dans la figure suivante

Client	
⚠	ID_CLIENT
	DATE_NAISSANCE
	AGE_CLIENT
	REGION
	SEGMENT
	Handset

Figure 18: DIM Client

Nom du champ	Description
ID_CLIENT	Représente l'identifiant du client
DATE_NAISSANCE	Représente la date de naissance du client
AGE_CLIENT	Représente l'age du client
REGION	Représente la région du client
SEGMENT	Représente les segments de clientèle
Handset	Représente le Handset du client (4G, 3G, etc...)

Tableau 12: Description DIM Client

#### 4.1.3. Dimension Offre

La dimension offre sera définie comme une table de dimension qui inclura les informations indiquées dans la figure ci-dessous

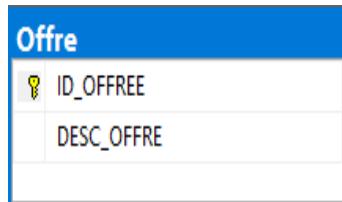


Figure 19: DIM Offre

Nom du champ	Description
ID_OFFRE	Représente l'identifiant de l'offre
DESC_OFFRE	Représente le nom complet de l'offre

Tableau 13 : Description DIM Offre

#### 4.1.4. Dimension Montant

La dimension montant sera établie comme un tableau de dimensions qui comprendra les informations indiquées dans la figure ci-dessous

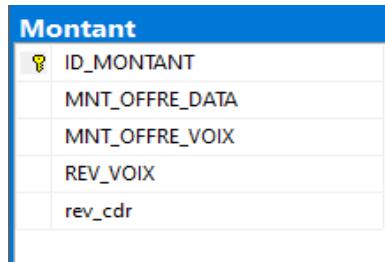


Figure 20: DIM Montant

Nom du champ	Description
ID_MONTANT	Représente l'identifiant du montant
MNT_OFFRE_DATA	Représente le montant des offres données mobiles
MNT_OFFRE_VOIX	Représente le montant des offres voix
REV_VOIX	Représente le revenu voix
rev_cdr	Représente le revenu data + sms

Tableau 14: Description DIM Montant

#### 4.1.5. Dimension Service

La dimension service sera établie comme un tableau de dimensions qui contiendra les informations indiquées dans la figure ci-dessous

Service	
KEY	ID_SERVICE_SVA
	DESC_SERVICE

Figure 21: DIM Service

Nom du champ	Description
ID_SERVICE	Représente l'identifiant du service
DESC_SERVICE	Représente le nom des services

Tableau 15: Description DIM Service

#### 4.1.6. Dimension Consommation

La dimension consommation sera définie comme un tableau de dimensions qui contiendra les informations indiquées dans la figure ci-dessous

Consommation	
KEY	ID_CONSUMMATION
	USER_SERVICE
	VOLUME_4G
	VOLUME_3G
	VOLUME_2G

Figure 22: DIM Consommation

Nom du champ	Description
ID_CONSUMMATION	Représente l'identifiant de consommation
USER_SERVICE	Représente l'utilisateur du service
VOLUME_4G	Représente le volume 4G consommé
VOLUME_3G	Représente le volume 3G consommé
VOLUME_2G	Représente le volume 2G consommé

Tableau 16: Description DIM Consommation

#### 4.1.7. Dimension Forfait

La table de dimension forfait sera configurée comme une table de dimension qui contiendra les informations présentées dans la figure suivante

Forfait	
KEY	ID_FORFAIT
	OPTION_FORFAIT_DESC

Figure 23: DIM Forfait

Nom du champ	Description
ID_FORFAIT	Représente l'identifiant du forfait
OPTION_FORFAIT_DESC	Représente la liste des forfaits TT

Tableau 17: Description DIM Forfait

#### 4.1.8. Dimension Transact

La table Transact sera définie comme une table de dimensions qui contiendra les informations montrées dans la figure ci-dessous.

Transact	
!	ID_TRANSACTION
	USER_TRANSACT
	VOLUME_DATA_MEGA_ATTRIBUE
	NB_JOUR_VALIDITE
	TRANSACT_AMOUNT

Figure 24: DIM Transact

Nom du champ	Description
ID_TRANSACTION	Représente l'identifiant de la transaction
USER_TRANSACT	Représente le client qui a effectué la transaction
VOLUME_DATA_MEGA_ATTRIBUE	Représente le volume data attribué en Mo
NB_JOUR_VALIDITE	Représente le nombre de jours de validité
TRANSACT_AMOUNT	Représente le montant de la transaction

Tableau 18: Description DIM Transact

### 4.2. Description de Faits

Une fois les tables de dimensions fixées, nous commençons la phase d'identification des tables de faits. En effet, pour construire correctement une table de faits, il faut tenir compte de certaines règles :

Une table de faits contient des valeurs numériques qui sont utilisées pour faire des SUM, AVG.

- Une table de faits porte les clés associées aux dimensions. Elles représenteront les clés étrangères dans la table de faits.

Dans notre cas, il y a trois tables de faits : la table de faits TRAFIC Offre, la table de faits TRAFIC Service et la table de faits TRAFIC Forfait.

#### 4.2.1. Fait Trafic Offre

La table de faits **TRAFIC Offre** (c'est la table pour le suivi des offres TT) sera définie comme une table de faits qui contiendra les informations indiquées dans la figure suivante:

Trafic_Offre
ID_CLIENT
ID_DATE
ID_OFFREE
ID_MONTANT
Score_churn
DUREE_APPEL
NB_APPELS
MNT_RECH
NB_RECH
NB_CDR
Montant_offre
Revenu_offre
Profit

Figure 25: Fact Trafic Offre

Nom du champ	Description
ID_CLIENT	Représente le code du client
ID_DATE	Représente la date de recharge
ID_MONTANT	Représente le code du montant
ID_OFFREE	Représente le code d'offre
Score_churn	Représente le score du perte de client au mois de Janvier
DUREE_APPEL	Représente la durée d'appel du client au mois de Janvier
NB_APPELS	Représente le nombre d'appels du client au mois de Janvier
MNT_RECH	Représente le montant de recharge du client au mois de Janvier
NB_RECH	Représente le nombre de recharge au mois de Janvier
NB_CDR	Représente le nombre d'utilisation des données au mois de Janvier

Tableau 19 : Description Fact TRAFIC Offre

#### 4.2.2. Fait Trafic Service

La table de faits **TRAFIC Service** (c'est la table pour le suivi des services TT) sera établie comme une table de faits qui présentera les informations illustrées dans la figure qui suit :

Trafic_Service
ID_CLIENT
ID_DATE
ID_SERVICE_SVA
ID_CONSUMMATION
STATUT_RGS30
STATUT_DATA_RGS30
VolumeConsomme4G
VolumeTotConsomme

Figure 26: Fact Trafic Service

Nom du champ	Description
ID_CLIENT	Représente le code du client
ID_DATE	Représente la date de recharge
ID_SERVICE_SVA	Représente le code du service
ID_CONSOMMATION	Représente le code de consommation
STATUT_RGS30	Représente le statut du client au mois de Janvier (Actif ou non)
STATUT_DATA_RGS30	Représente le statut data du client au mois de Janvier (Actif ou non)

Tableau 20 : Description Fact TRAFIC Service

#### 4.2.3. Fait Trafic Forfait

La table de faits du **TRAFIC Forfait** (il s'agit du tableau de suivi des forfaits TT) affichera les informations présentées dans la figure suivante :

Traffic_Forfait
ID_CLIENT
ID_FORFAIT
ID_TRANSACTION
ID_DATE
Flag_PROV_4G
NBR_JR_INACTIVITE_DATA
VolumeTotAttribue
MntTotTransactions
MoyTransaction
VolumeMoyAttribue

Figure 27: Fact Trafic Forfait

Nom du champ	Description
ID_CLIENT	Représente le code du client
ID_DATE	Représente le code des date
ID_FORFAIT	Représente le code du service
ID_TRANSACTION	Représente le code de consommation
FLAG_PROV_4G	Représente le flag de provisionnement du réseau 4G
NBR_JR_INACTIVITE_DATA	Représente le nombre de jours d'inactivité

Tableau 21 : Description Fact TRAFIC Forfait

#### 4.3. Vue d'ensemble de la conception de l'entrepôt

Chaque magasin a un segment d'analyse et tous les magasins partagent des dimensions communes entre eux.

Notre modèle de l'entrepôt de données complet est un modèle de constellation construit à partir des modèles en étoile et est illustré dans la figure 28.

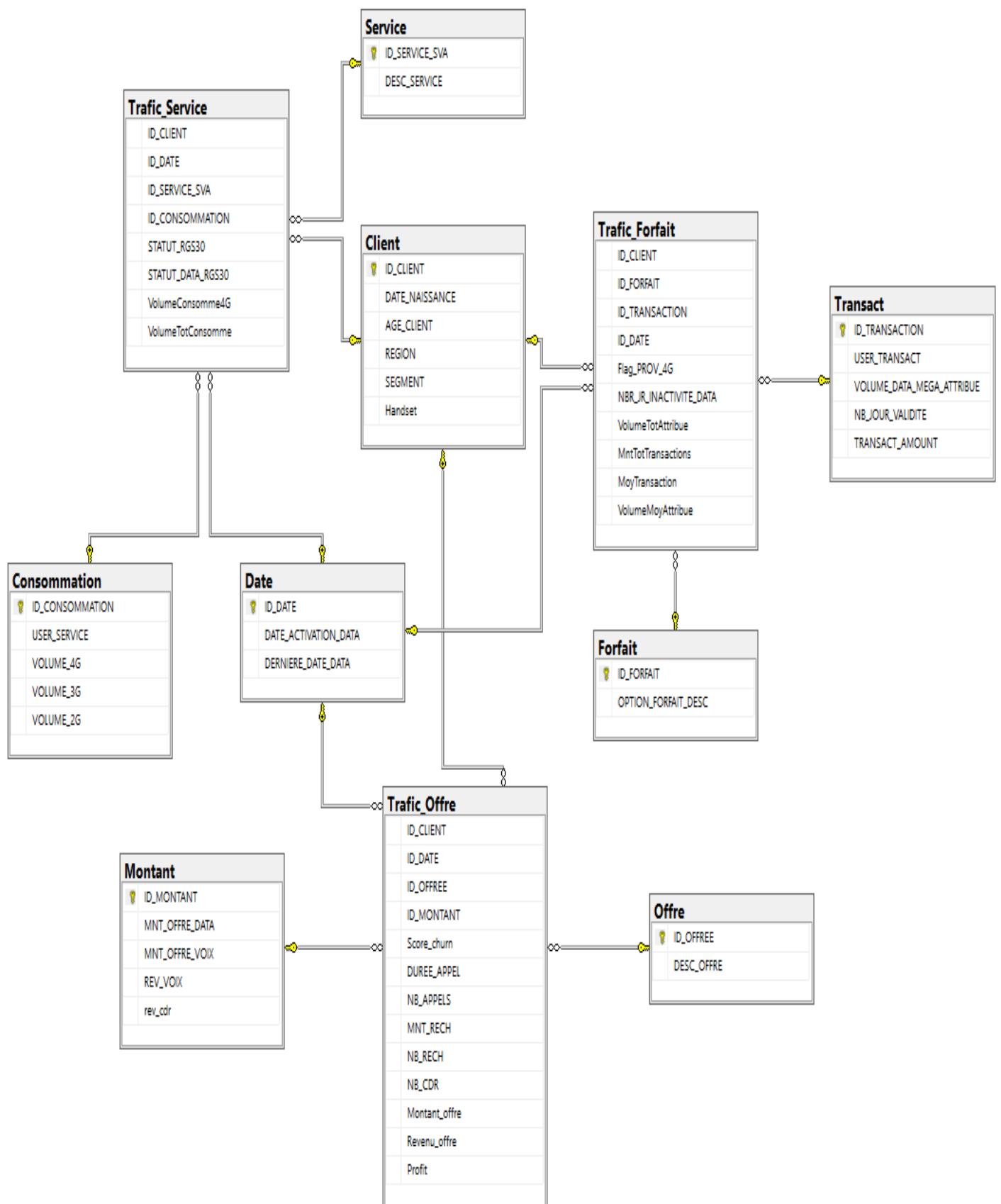


Figure 28: Notre Data Warehouse

Notre entrepôt de données est composé de trois datamarts intitulés "Trafic Offre", "Trafic Forfait", " Trafic Service".

#### - Datamart Trafic Offre

Ce schéma en étoile présente le Datamart Trafic Offre, dont les mesures sont le "montant des offres", le "revENU des offres" et le "profit du revenu".

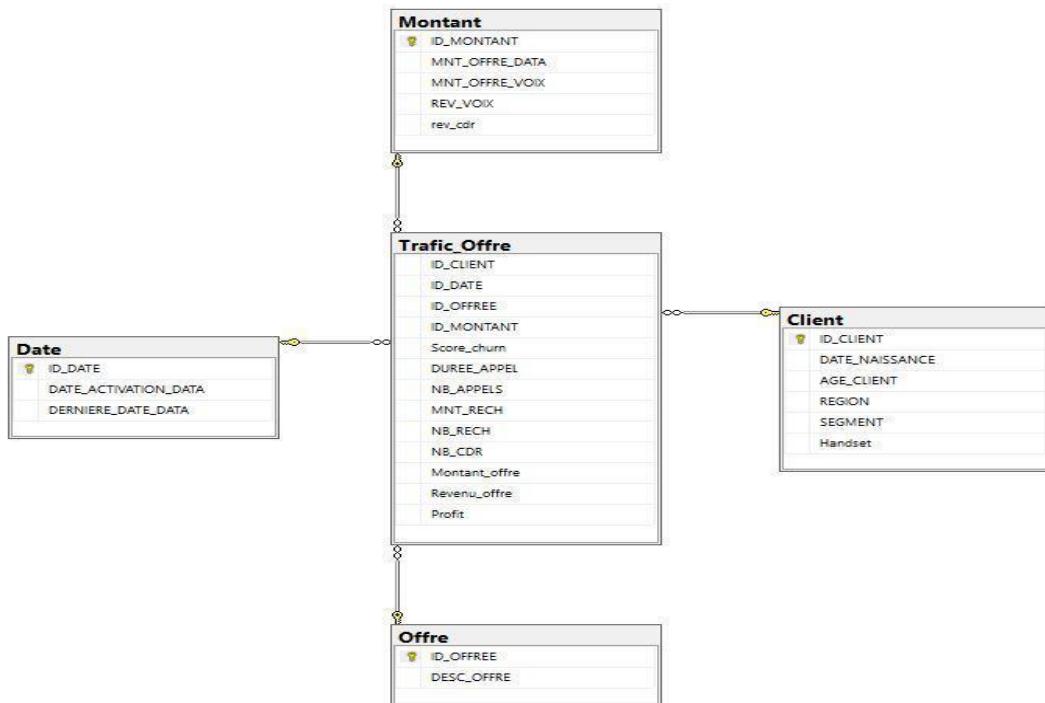


Figure 29: Datamart Trafic Offre

#### - Datamart Trafic Forfait

Ce schéma en étoile représente le Datamart de Forfait avec les mesures suivantes : volume total de données attribuées, total des transactions, transactions moyennes et volume moyen attribué.

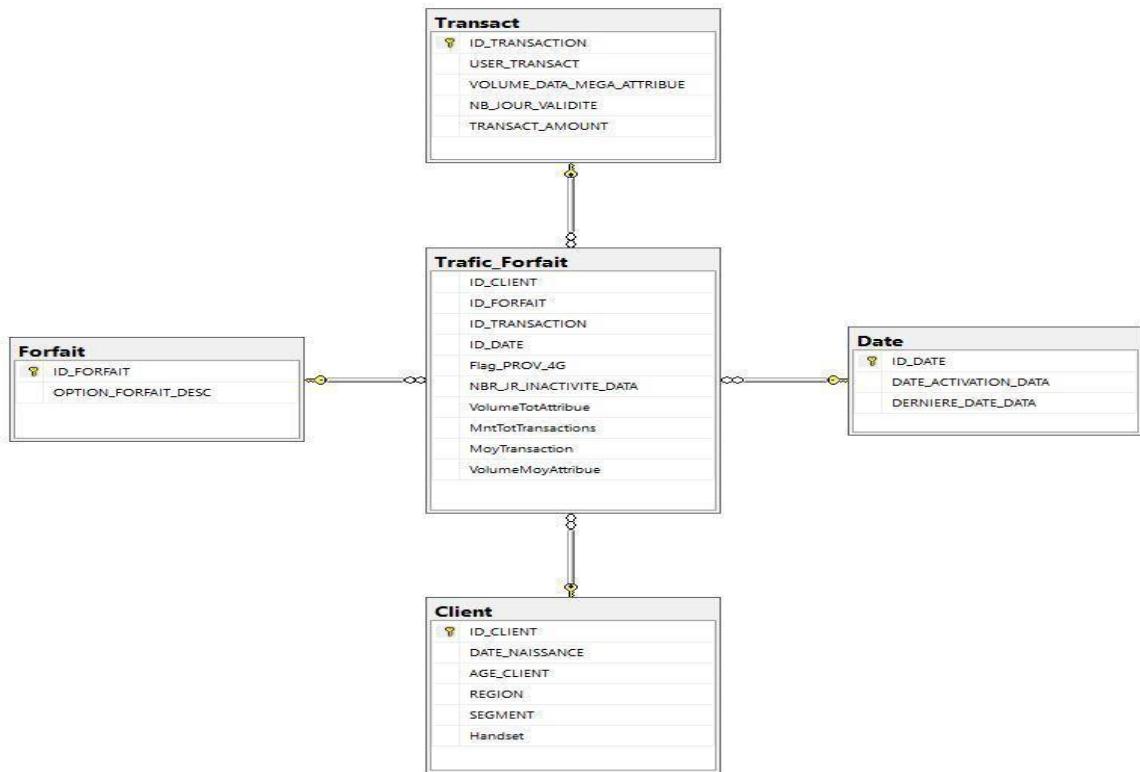


Figure 30: Datamart Trafic Forfait

#### - Datamart Trafic Service

Le datamart Trafic Service est présenté dans un schéma en étoile avec les mesures suivantes: Le volume total de 4G consommé et le volume total consommé.

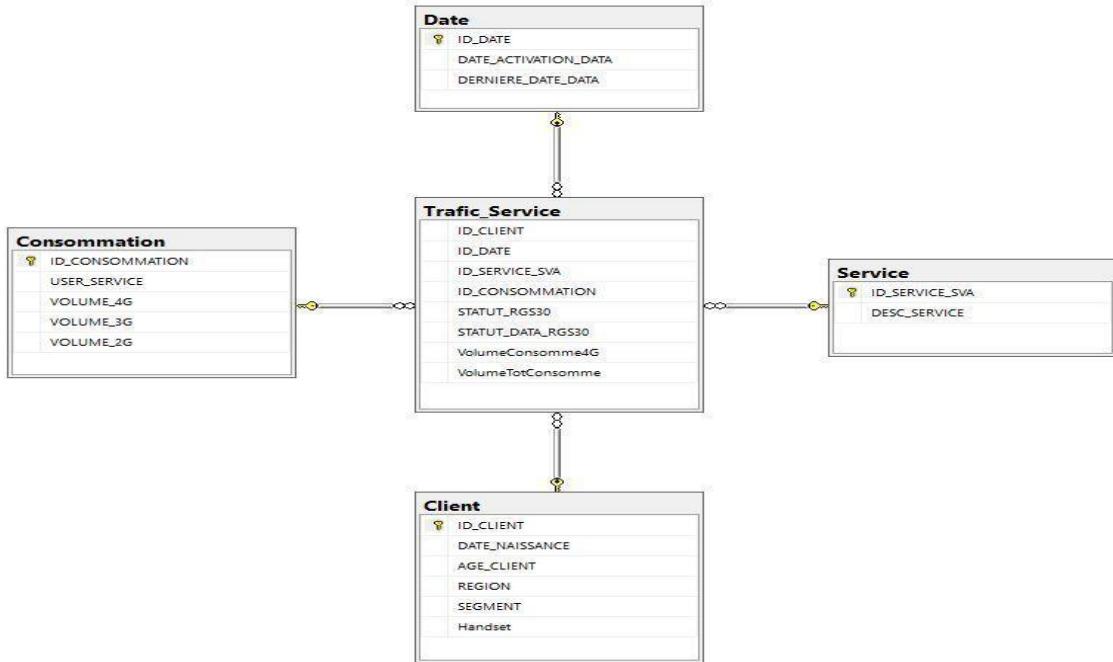


Figure 31: Datamart Trafic Service

## 5. Processus de chargement ETL

### 5.1. Processus de chargement des dimensions

La table de dimensions fournit le contexte de la table de faits. Elle représente le point d'entrée de l'entrepôt de données. Pour nos besoins, nous résumons la détection des données nécessaires au chargement des dimensions dans les tableaux ci-dessous.

#### - La dimension Date

Pour charger la dimension 'Date' présentée ci-après de notre entrepôt de données [DWH\_TT], nous avons utilisé la table 'DIM\_DATE' du fichier source .csv. Pour la colonne cible de l'entrepôt de données 'ID\_DATE' nous avons utilisé l'attribut 'ID\_DATE', pour la colonne 'DATE ACTIVATION DATA' nous avons utilisé l'attribut 'DATE ACTIVATION DATA' et pour la colonne 'DERNIÈRE DATE DATA' nous avons utilisé l'attribut 'DERNIÈRE DATE DATA'.

Colonne cible	Colonne source
[DWH_TT].[dbo].[Date].[ID_DATE]	[csv file].DIM_DATE.ID_DATE
[DWH_TT].[dbo].[Date].[DATE_ACTIVATION_DATA]	[csv file].DIM_DATE.DATE_ACTIVATION_DATA
[DWH_TT].[dbo].[Date].[DERNIERE_DATE_DATA]	[csv file].DIM_DATE.DERNIERE_DATE_DATA

Tableau 22 : Déterminer les données source et de destinations de la dimension "Date"

#### - La dimension Client

Pour charger la dimension 'Client' présentée ci-après de notre entrepôt de données [DWH\_TT], nous avons utilisé la table 'DIM\_DATE' du fichier Excel source. Pour la colonne cible de l'entrepôt de données 'AGE\_CLIENT', nous l'avons créée après dans le processus ETL en utilisant la différence entre la date actuelle et l'attribut 'DATE DE NAISSANCE'.

Colonne cible	Colonne source
[DWH_TT].[dbo].[Client].[ID_CLIENT]	[excel file].DIM_CLIENT.ID_CLIENT
[DWH_TT].[dbo].[Client].[DATE_NAISSANCE]	[excel file].DIM_CLIENT.DATE_NAISSANCE
[DWH_TT].[dbo].[Client].[AGE_CLIENT]	[excel file].DIM_CLIENT.AGE_CLIENT
[DWH_TT].[dbo].[Client].[REGION]	[excel file].DIM_CLIENT.REGION
[DWH_TT].[dbo].[Client].[SEGMENT]	[excel file].DIM_CLIENT.SEGMENT
[DWH_TT].[dbo].[Client].[Handset]	[excel file].DIM_CLIENT.Handset

Tableau 23 : Déterminer les données source et de destinations de la dimension "Client"

#### - La dimension Forfait

Pour charger la dimension 'Forfait' présentée ci-dessous de notre entrepôt de données [DWH\_TT], nous avons utilisé la table 'DIM\_FORFAIT' du fichier Excel source. Pour la colonne cible de l'entrepôt de données 'ID\_FORFAIT' nous avons utilisé l'attribut 'ID\_FORFAIT', et pour la colonne 'OPTION\_FORFAIT\_DESC' nous avons utilisé l'attribut 'OPTION\_FORFAIT\_DESC'.

Colonne cible	Colonne source
[DWH_TT].[dbo].[Forfait].[ID_FORFAIT]	[excel file].DIM_FORFAIT.ID_FORFAIT
[DWH_TT].[dbo].[Forfait].[OPTION_FORFAIT_DESC]	[excel file].DIM_FORFAIT.OPTION_FORFAIT_DESC

Tableau 24 : Déterminer les données source et de destinations de la dimension "Forfait"

### - La dimension Service

Pour charger la dimension 'Forfait' présentée ci-dessous de notre entrepôt de données [DWH\_TT], nous avons utilisé la table 'DIM\_SERVICE' du fichier Excel source. Pour la colonne cible de l'entrepôt de données 'ID\_SERVICE\_SVA' nous avons utilisé l'attribut 'ID\_SERVICE\_SVA', et pour la colonne 'DESC\_SERVICE' nous avons utilisé l'attribut 'DESC\_SERVICE'.

Colonne cible	Colonne source
[DWH_TT].[dbo].[Service].[ID_SERVICE_SVA]	[excel file].DIM_SERVICE.ID_SERVICE_SVA
[DWH_TT].[dbo].[Service].[DESC_SERVICE]	[excel file].DIM_SERVICE.DESC_SERVICE

Tableau 25 : Déterminer les données source et de destinations de la dimension "Service"

### - La dimension Consommation

Pour charger la dimension 'Consommation' de notre entrepôt de données [DWH\_TT], nous avons utilisé la table 'DIM\_CONSUM' dans le fichier source .csv.

Le tableau suivant montre les colonnes cibles et sources de la dimension 'Consommation'.

Colonne cible	Colonne source
[DWH_TT].[dbo].[Consommation].[ID_CONSUMMATION]	[csv file].DIM_CONSUM.ID_CONSUMMATION
[DWH_TT].[dbo].[Consommation].[USER_SERVICE]	[csv file].DIM_CONSUM.USER_SERVICE
[DWH_TT].[dbo].[Consommation].[VOLUME_4G]	[csv file].DIM_CONSUM.VOLUME_4G
[DWH_TT].[dbo].[Consommation].[VOLUME_3G]	[csv file].DIM_CONSUM.VOLUME_3G
[DWH_TT].[dbo].[Consommation].[VOLUME_2G]	[csv file].DIM_CONSUM.VOLUME_2G

Tableau 26 : Déterminer les données source et de destinations de la dimension "Consommation"

### - La dimension Transact

Pour charger la dimension 'Transact' de notre entrepôt de données [DWH\_TT], nous avons fait appel à la table 'DIM\_TRANSACT' dans le fichier source .csv.

Le tableau suivant présente les colonnes cibles et sources de la dimension 'Transact'.

Colonne cible	Colonne source
[DWH_TT].[dbo].[Transact].[ID_TRANSACTION]	[csv file].DIM_TRANSACT.ID_TRANSACTION
[DWH_TT].[dbo].[Transact].[USER_TRANSACT]	[csv file].DIM_TRANSACT.USER_TRANSACT
[DWH_TT].[dbo].[Transact].[VOLUME_DATA_MEGA_ATTRIBUE]	[csv file].DIM_TRANSACT.VOLUME_DATA_MEGA_ATTRIBUE
[DWH_TT].[dbo].[Transact].[NB_JOUR_VALIDITE]	[csv file].DIM_TRANSACT.NB_JOUR_VALIDITE
[DWH_TT].[dbo].[Transact].[TRANSACT_AMOUNT]	[csv file].DIM_TRANSACT.TRANSACT_AMOUNT

Tableau 27 : Déterminer les données source et de destinations de la dimension "Transact"

### - La dimension Offre

Pour charger la dimension 'Offre' illustrée ci-dessous de notre entrepôt de données [DWH\_TT], nous avons utilisé la table 'DIM\_OFFRE' du fichier Excel source. Pour la colonne cible 'ID\_OFFRE' de l'entrepôt de données, nous avons utilisé l'attribut 'ID\_OFFREE', et pour la colonne 'DESC\_OFFRE', nous avons utilisé l'attribut 'DESC\_OFFRE'.

Colonne cible	Colonne source
[DWH_TT].[dbo].[Offre].[ID_OFFREE]	[excel file].DIM_OFFRE.ID_OFFRE
[DWH_TT].[dbo].[Offre].[DESC_OFFRE]	[excel file].DIM_OFFRE.DESC_OFFRE

Tableau 28 : Déterminer les données source et de destinations de la dimension "Offre"

### - La dimension Montant

Pour charger la dimension 'Montant' de notre data warehouse [DWH\_TT], nous avons appelé la table 'DIM\_MONTANT' dans le fichier source .csv.

Le tableau qui suit montre les colonnes cibles et sources de la dimension 'Montant'.

Colonne cible	Colonne source
[DWH_TT].[dbo].[Montant].[ID_MONTANT]	[csv file].DIM_MONTANT.ID_MONTANT
[DWH_TT].[dbo].[Montant].[MNT_OFFRE_DATA]	[csv file].DIM_MONTANT.MNT_OFFRE_DATA
[DWH_TT].[dbo].[Montant].[MNT_OFFRE_VOIX]	[csv file].DIM_MONTANT.MNT_OFFRE_VOIX
[DWH_TT].[dbo].[Montant].[REV_VOIX]	[csv file].DIM_MONTANT.REV_VOIX
[DWH_TT].[dbo].[Montant].[rev_cdr]	[csv file].DIM_MONTANT.rev_cdr

Tableau 29 : Déterminer les données source et de destinations de la dimension "Montant"

### 5.2. Processus de chargement des tables de faits

Une fois que toutes les dimensions sont chargées, nous continuerons à charger les tables de faits. Pour chaque enregistrement de table de faits dans l'entrepôt de données, l'identifiant de la table de faits sera constitué des clés des dimensions pertinentes. La manière de déterminer les mesures est indiquée dans le tableau ci-après :

Nom de table	Mesure	Valeur dans la table source
Trafic_Forfait	[DWH_TT].[dbo].[Trafic_Forfait].[VolumeTotAttribue]	SUM (VOLUME_DATA_MEGA_ATTRIBUE)
	[DWH_TT].[dbo].[Trafic_Forfait].[MntTotTransactions]	SUM (TRANSACT_AMOUNT)
	[DWH_TT].[dbo].[Trafic_Forfait].[MoyTransaction]	SUM (TRANSACT_AMOUNT) / COUNT (ID_CLIENT)
	[DWH_TT].[dbo].[Trafic_Forfait].[VolumeMoyAttribue]	SUM (VOLUME_DATA_MEGA_ATTRIBUE) / COUNT (ID_CLIENT)
Trafic_Offre	[DWH_TT].[dbo].[Trafic_Offre].[Montant_offre]	SUM (MNT_OFFRE_DATA + MNT_OFFRE_VOIX)
	[DWH_TT].[dbo].[Trafic_Offre].[Revenu_offre]	SUM (REV_VOIX + rev_cdr)
	[DWH_TT].[dbo].[Trafic_Offre].[Profit]	SUM (REV_VOIX + rev_cdr) - MNT_RECH
Trafic_Service	[DWH_TT].[dbo].[Trafic_Service].[VolumeConsomme4G]	SUM (VOLUME 4G)
	[DWH_TT].[dbo].[Trafic_Service].[VolumeTotConsomme]	SUM (VOLUME 4G + VOLUME 3G + VOLUME 2G)

Tableau 30 : Déterminer les données source et de destinations des tables des faits

## 6. Conception des cubes dimensionnels OLAP

La modélisation de notre entrepôt de données nécessite une structure de stockage adoptée en OLAP, qui permet, à partir de l'entrepôt de données, d'effectuer une analyse des activités de l'entreprise avec des données statistiques exactes. On parle en effet d'une base de données multidimensionnelle.

Ces données sont analysables. Le cube présente la table de faits que l'on se propose de construire lors de la mise en place de l'entrepôt de données, et les données quantitatives sont utilisées pour faire des calculs.

## 6.1. La conception des cubes OLAP

Dans ce contenu, nous allons écrire la liste des cubes à placer.

On donnera les dimensions participantes et les valeurs mesurables qui apparaissent dans ces cubes.

Le tableau qui suit donne des détails sur la conception de ces cubes

Nom du cube	Mesures	Dimensions
DWH_TT_TRAFIC OFFRE CUBE	Montant_offre Revenu_offre Profit	Dim_Date Dim_Client Dim_Offre Dim_Montant
DWH_TT_TRAFIC FORFAIT CUBE	VolumeTotAttribue MntTotTransactions MoyTransaction VolumeMoyAttribue	Dim_Forfait Dim_Transact Dim_Date Dim_Client
DWH_TT_TRAFIC SERVICE CUBE	VolumeConsomme4G VolumeTotConsomme	Dim_Service Dim_Client Dim_Consumption Dim_Date

Tableau 31: La liste des cubes OLAP

## 6.2. Liste des cubes OLAP

Nous allons maintenant dresser la liste des cubes à placer. Pour chacun des cubes, nous allons indiquer les dimensions participantes et les valeurs mesurables qui y figurent.

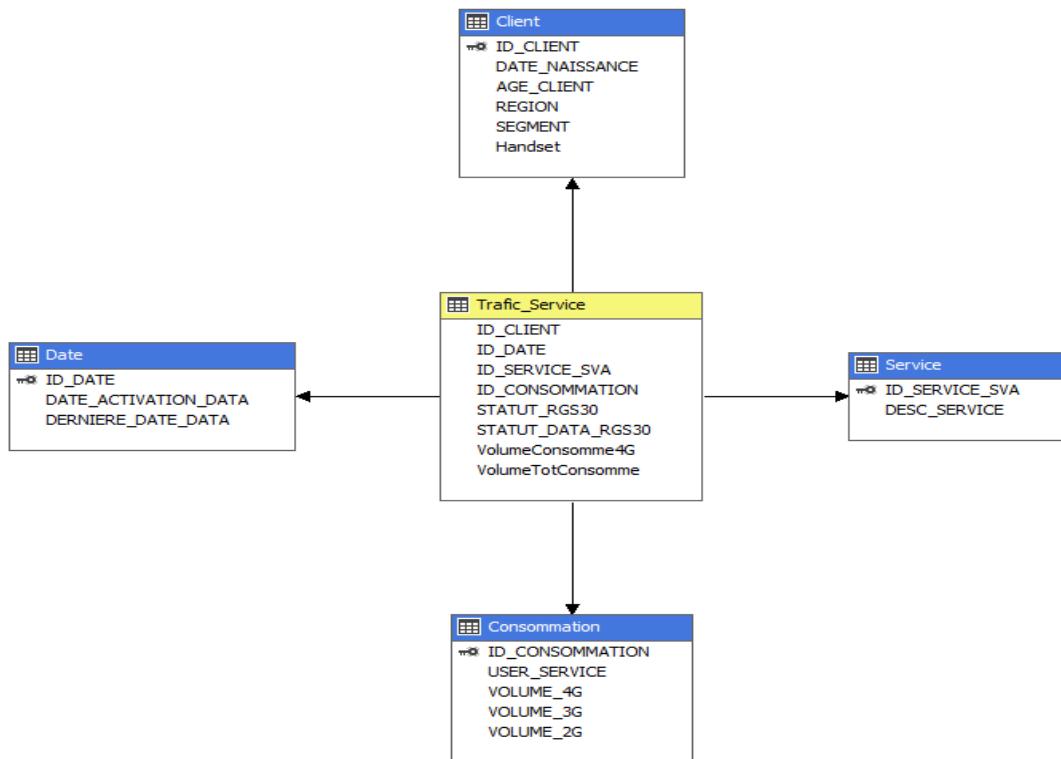


Figure 32: Cube OLAP “Trafic\_Service”

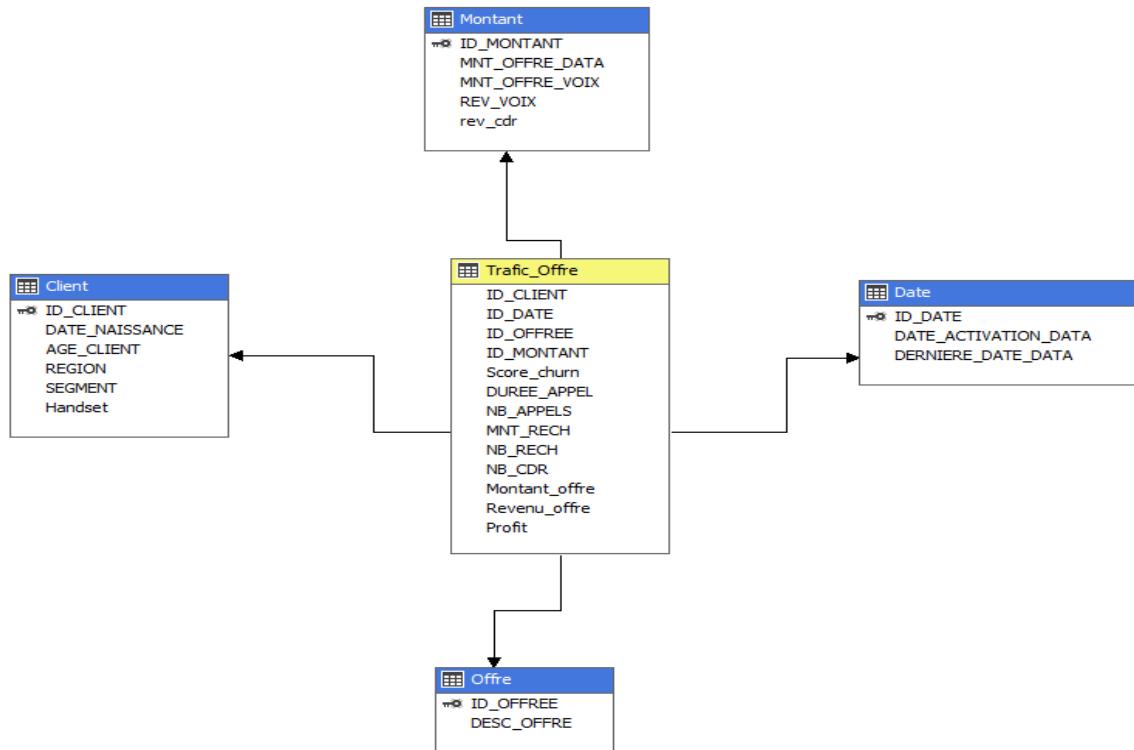


Figure 33: Cube OLAP “Trafic\_Offre”

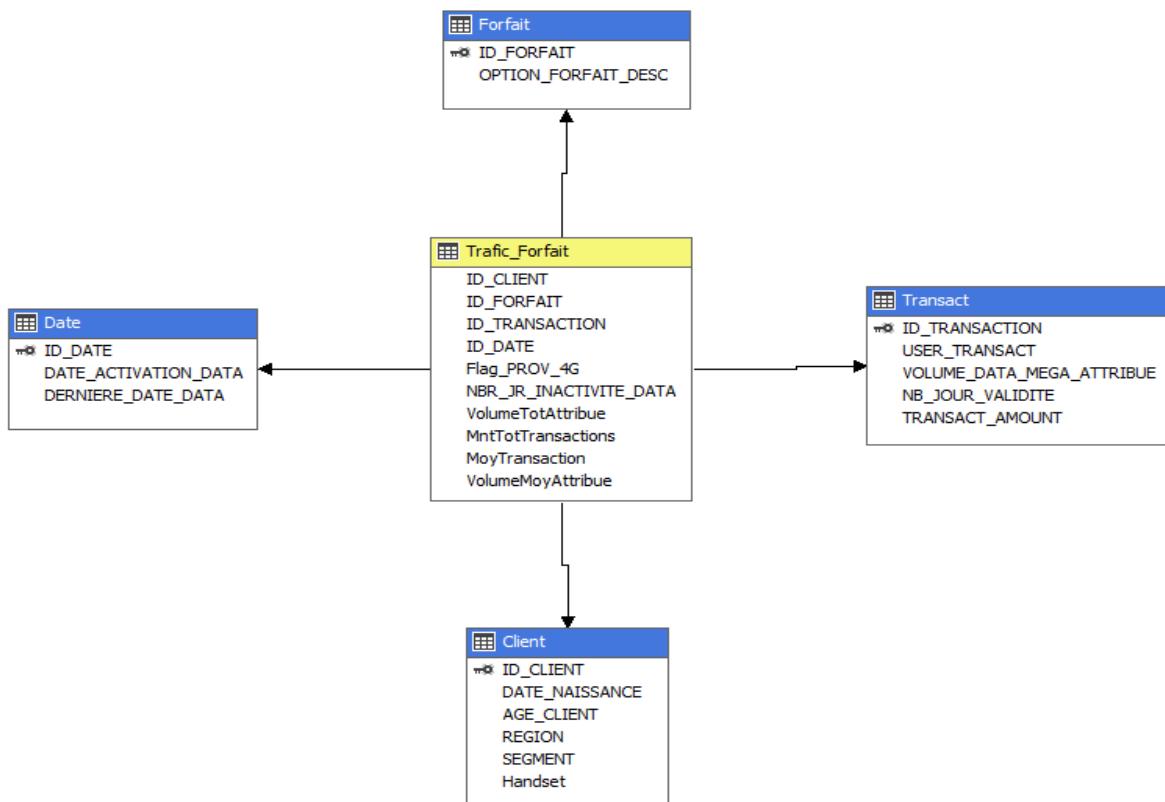


Figure 34: Cube OLAP “Trafic\_Forfait”

## 7. Conclusion

Tout au long de ce chapitre, nous avons spécifié et modélisé les besoins fonctionnels et non fonctionnels, puis présenté l'architecture suivie, les KPIs calculés, puis la modélisation de notre Data Warehouse et enfin la conception du chargement de l'ETL et des cubes OLAP.

Dans le chapitre qui suit, nous présenterons en détail l'environnement et les différents outils qui composent notre travail, et qui nous permettent ainsi de mettre en œuvre notre application.

# CHAPITRE 4

## RÉALISATION ET TESTS

### 1. Introduction

Après avoir terminé la phase de conception, nous arrivons à la dernière phase du processus GIMSI, la phase de réalisation et tests, qui détaille le travail nécessaire pour mettre en œuvre notre solution.

Nous présenterons d'abord l'environnement de travail, puis la sélection technique requise pour le développement et enfin les phases de construction.

### 2. Environnement

Tout projet informatique nécessite un environnement matériel et logiciel bien préparé.

#### 2.1. Environnement matériel

Nous avons réalisé notre projet en utilisant un ordinateur portable MSI dont la configuration est présentée ci-après :

- Disque Dur: 500 Go
- RAM: 16 Go
- Processeur: Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz 2.59 GHz
- Système d'exploitation : Windows 10 Professionnel 64 bits

#### 2.2. Environnement logiciel

Pour la réalisation de notre projet BI, nous avons utilisé plusieurs logiciels qui sont présentés dans ce qui suit :

- Outil de conception : StarUML



Figure 35: StarUML

StarUML est un outil de modélisation UML, publié en open source par son éditeur, à la fin de son implémentation, sous un certificat modifié de GNU à GPL.

StarUML gère la majorité des diagrammes spécifiés dans la norme UML 2.0.

### **- Gantt Project**

C'est un logiciel gratuit de gestion de projet qui permet de planifier un projet dans le temps grâce à la création d'un diagramme de Gantt.



Figure 36: Gantt Project

### **- Microsoft SQL SERVER 2019**

Microsoft SQL SERVER 2019 Developer est un système de gestion des données gratuit, puissant et fiable qui fournit une base de données riche et performante pour les sites Web et les applications de bureau open source.

---



Figure 37: SQL SERVER 2019

### **- SQL Server Management Studio 18 (SSMS 18)**

SQL Server Management Studio est l'un des logiciels de gestion de bases de données et de serveurs les plus couramment utilisés. Il vous permet de consulter, gérer, configurer et développer tous les composants de SQL Server et des bases de données SQL.

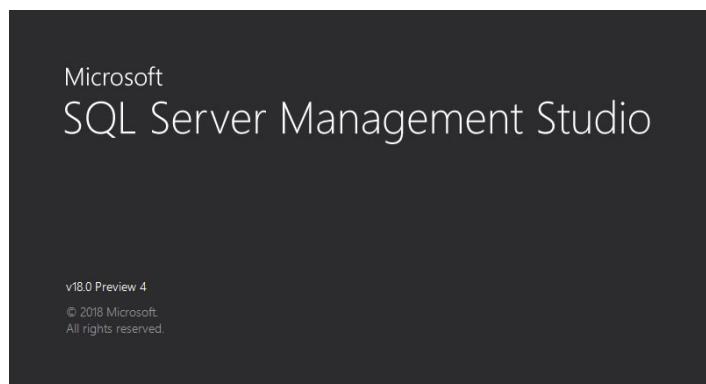


Figure 38: SSMS 18

### **- Visual Studio Community 2019**

Visual Studio Community est une version téléchargeable gratuitement pour les développeurs, équivalente en termes de fonctionnalités à Visual Studio Professional qui est une version payante. Elle permet de générer des applications web, des applications de stockage et de traitement des données, des applications de science et d'analyse des données, etc...

Nous allons télécharger les extensions SSIS et SSAS de Visual Studio 2019.



Figure 39: Visual Studio 2019

#### - SSIS

SQL Server Integration Services (SSIS) est un composant du logiciel de base de données Microsoft SQL Server. Il s'agit d'une plateforme de création de solutions de transformation et d'intégration de données au niveau de l'entreprise.

#### - SSAS

SQL Server Analysis Services (SSAS) est une plateforme de stockage et de restitution des données qui fait également partie de la suite Microsoft BI et qui est principalement dédiée à la mise en œuvre de cubes OLAP ou TABULAIRES.

#### - Power BI Desktop

Power BI Desktop est une application gratuite qui vous permet de vous connecter à plusieurs sources de données différentes, de les transformer et de les visualiser sous forme de rapports.



Power BI Desktop

Figure 40: Power BI Desktop

## 2.3. Outils de développement

#### - Visual Studio Code

Un éditeur de code gratuit, open source et multiplateforme qui prend en charge une douzaine de langages que nous avons utilisés pour développer avec le langage PHP.

## **2.4. Langage de programmation**

### **- PHP**

C'est un langage de programmation principalement utilisé pour produire des pages Web dynamiques via un serveur, mais pouvant également fonctionner comme n'importe quel langage interprété de façon locale. PHP est un langage impératif orienté objet. [Web6]

### **- Javascript**

JavaScript désigne un langage de développement informatique, et plus précisément un langage de script orienté objet. On le retrouve principalement dans les pages Internet. Il permet, entre autres, d'introduire sur une page web ou HTML des petites animations ou des effets. [Web10]

## **2.5. Langages descriptifs**

### **- HTML**

Le HTML est un langage descriptif qui sert à définir les différents éléments d'une page, à leur donner du sens. [Web8]

### **- CSS**

Le CSS va servir à mettre en forme les différents contenus définis par le HTML en leur appliquant des styles. [Web8]

## **2.6. Plateforme de développement**

### **- WampServer**

C'est une plateforme de développement Web de type WAMP, permettant de faire fonctionner localement (sans avoir à se connecter à un serveur externe) des scripts PHP. [Web7]

### **- PhpMyAdmin**

*PhpMyAdmin* (PMA) est une application Web de gestion pour les systèmes de gestion de base de données, réalisée principalement en PHP. Elle permet d'exécuter, très facilement et sans grandes connaissances en bases de données, des requêtes comme les créations de table de données, insertions, mises à jour, etc... [Web9]

## **3. Implémentation ETL**

### **3.1. Composants de SSIS**

Le tableau ci-après présente les composants que nous avons utilisés avec SQL Server Integration Services

Composants	Détails
Tâche de flux de données  Tâche de flux de données	Ce composant définit le moteur de flux de données qui transfère les données des sources aux destinations et fournit la fonctionnalité permettant de transformer, nettoyer et modifier les données lors de leur transfert entre les sources et les destinations.
Tâche d'exécution de requêtes SQL  Tâche d'exécution de requêtes SQL	La tâche d'exécution SQL exécute des instructions SQL ou des procédures stockées SQL à partir d'un package. Par exemple, le fait de tronquer une table ou une vue pour la préparer à l'insertion de données est une instruction SQL.
Source Excel  Source Excel	La source Excel permet d'extraire des données de feuilles de calcul ou de plages dans les classeurs Microsoft Excel.
Source du fichier plat  Source du fichier plat	La source Flat File lit les données d'un fichier texte. Ce dernier peut être au format délimité, à largeur fixe ou mixte.
Transformation Colonne dérivée  Colonne dérivée	Ce composant permet de générer de nouvelles valeurs de colonnes en appliquant des expressions aux colonnes d'entrée de la transformation.
Transformation Conversion de données  Conversion de données	Ce composant permet de convertir les données d'une colonne d'entrée vers un type de données différent, puis de copier ces données vers une nouvelle colonne de sortie.
Transformation Recherche  Recherche	La transformation Lookup effectue des recherches en joignant les données des colonnes d'entrée aux colonnes d'un jeu de données de référence.
Transformation jointure de fusion  Jointure de fusion	La transformation Merge Join fournit une sortie qui est générée en joignant deux jeux de données triés à l'aide d'une jointure FULL, LEFT ou INNER.
Transformation d'agrégation  Agrégation	La transformation Aggregate utilise des fonctions d'agrégation, telles que le calcul de la moyenne, sur les valeurs des colonnes et copie les résultats sur la sortie de la transformation. En outre, la transformation fournit la clause GROUP BY, que vous pouvez utiliser pour spécifier les différents groupes à agréger.
Transformation Trier  Trier	La transformation Trier classe les données d'entrée dans l'ordre croissant ou décroissant et copie les données classées dans la sortie de la transformation.
Destination OLE DB  Destination OLE DB	La destination OLE DB charge les données dans une variété de bases de données compatibles OLE DB en utilisant une table ou une vue de base de données ou une commande SQL.

Tableau 32 : Tableau de description des composants SSIS

### 3.2. Sources de données

Nous avons reçu des fichiers sources Excel et CSV dont nous allons présenter quelques-uns dans les 2 figures suivantes

DIM_CLIENT - DIM_CLIENT.csv - Bloc-notes				
Fichier	Édition	Format	Affichage	Aide
ID_CLIENT,DATE_NAISSANCE,REGION,SEGMENT,Handset				
ID104000712,1/4/1996,Kebili,DATA ADDICTED,4G				
ID104000866,12/4/1979,Kebili,VAS,4G				
ID104001440,27/8/2010,Kasserine,VIP,4G				
ID104001565,9/8/1982,Sidi Bouzid,YOUNG,4G				
ID104002025,21/3/1997,Sfax,VOICE,4G				
ID104002116,5/6/1955,Sfax,OLD,4G				
ID104002218,17/7/1991,Zaghouan,CONNECTED ADULTE,4G				
ID104002592,4/3/1954,Kebili,OLD,4G				
ID104002792,25/5/1983,Jendouba,DATA ADDICTED,4G				
ID104003138,14/3/1983,Siliana,YOUNG,4G				
ID104003141,20/8/2010,Kasserine,DATA ADDICTED,4G				
ID104003279,14/8/1981,Tunis,OLD,2G				
ID104003676,27/1/1982,Kebili,OLD,2G				
ID104003981,31/3/1984,Jendouba,OLD,4G				
ID104004638,1/9/1988,Ben Arous,YOUNG,2G				
ID104005257,19/3/1972,Tunis,VAS,4G				
ID104007397,5/8/1959,Kebili,CONNECTED ADULTE,4G				
ID104007969,21/11/1992,Kebili,VOICE,2G				
ID104008507,14/2/1985,Kasserine,VAS,3G				
ID104009259,9/10/1975,Sfax,OLD,2G				
ID104010408,27/2/1988,Tunis,YOUNG,4G				
ID104011115,27/8/1983,Jendouba,OLD,2G				
ID104013130,29/5/1966,Mannouba,YOUNG,4G				
ID104014316,21/1/1978,Tunis,VOICE,4G				
ID104014430,9/7/1955,Monastir,YOUNG,4G				
ID104015616,15/10/1971,Tunis,YOUNG,4G				
ID104015668,21/8/1993,Kasserine,VIP,4G				
ID104015803,29/11/1986,Sousse,VOICE,4G				
ID104016081,12/12/1990,Mahdia,YOUNG,3G				
ID104016354,24/11/1969,Medenine,CONNECTED ADULTE,4G				
ID104017870,27/3/1991,Gafsa,YOUNG,4G				
ID104021146,6/9/1993,Siliana,VOICE,2G				
ID104021390,25/9/1975,Beja,YOUNG,3G				
ID104022181,2/2/1990,Bizerte,VOICE,4G				
ID104022386,10/2/1953,Jendouba,OLD,2G				
T0104024185 3/3/1991 Monastir OLD 4G				

Figure 41: Fichier source de type csv

FACT_OFFRE.xls - Microsoft Excel													
FILE	HOME	INSERT	PAGE LAYOUT	FORMULAS	DATA	REVIEW	VIEW	Calibri	11	A	=	Wrap Text	General
Paste	Cut	Copy	Font	Font	Font	Font	Font	Font	Font	Font	Font	Font	Font
S26													
1	ID_CLIENT	ID_DATE	ID_MONTANT	ID_OFFRE	Score_churn	DUREE_APPEL	NB_APPELS	MNT_RECH	NB_RECH	NB_CDR			
2	ID104000712	1	0 offre#8	0.014575008	51.35	221	123	35	8738				
3	ID104000866	2	1 offre#24	0.013598623	62.2	133	10	2	1721				
4	ID104001440	3	2 offre#32	0.0130482507	679.03	516	96	17	19044				
5	ID104002025	4	3 offre#16	0.020370176	151.03	103	7	3	421				
6	ID104002025	5	4 offre#32	0.011476017	151.22	144	9	5	54				
7	ID104002116	6	5 offre#24	0.049717998	66.4	34	0	0	5968				
8	ID104002218	7	6 offre#24	0.0291290034	465.27	597	44	16	1247				
9	ID104002592	8	7 offre#24	0.0734236548	60.89	45	5	1	170				
10	ID104002792	9	8 offre#24	0.0160808224	34.77	57	30	6	11170				
11	ID104003138	10	9 offre#24	0.0159862281	29.74	66	10	2	1252				
12	ID104003141	11	10 offre#24	0.0107482507	303.45	380	52	16	9471				
13	ID104003279	12	11 offre#24	0.0238780959	33.13	60	8	4	355				
14	ID104003676	13	12 offre#32	0.1543540506	138.27	98	10	2	30				
15	ID104003981	14	13 offre#24	0.0523845503	29.6	57	5	1	129				
16	ID104004058	15	14 offre#32	0.020370176	250.05	263	10	2	634				
17	ID104004257	16	15 offre#32	0.070453174	21.08	68	6	6	160				
18	ID1040047397	17	16 offre#24	0.0217228282	77.23	127	39	11	3587				
19	ID104007969	18	17 offre#24	0.0112015373	383.64	506	38	18	0				
20	ID104008507	19	18 offre#48	0.0194888334	100.5	192	32	8	462				
21	ID104009259	20	19 offre#32	0.0123526327	104.54	104	9	5	4				
22	ID104010464	21	20 offre#24	0.0681567577	91.8	32	10	2	2674				
23	ID104011115	22	21 offre#24	0.0467354546	17.41	21	5	1	5				
24	ID104013130	23	22 offre#24	0.0626848549	4.98	22	8	4	1146				
25	ID104014316	24	23 offre#48	0.015377224	1298.78	525	5	1	903				
26	ID104014430	25	24 offre#24	0.044016131	53.8	41	5	1	763				
27	ID104015616	26	25 offre#32	0.0291290038	28.72	28	10	1	446				
28	ID104015848	27	26 offre#48	0.0479520074	243.82	276	56	12	5698				
29	ID104015903	28	27 offre#24	0.057613044	1.52	378	55	11	125				
30	ID104016081	29	28 offre#32	0.0103646336	122.27	226	17	17	469				
31	ID104016354	30	29 offre#24	0.0232703747	28.05	77	5	1	3969				
32	ID104017970	31	30 offre#32	0.0223351347	3.2	21	2	2	156				
33	ID104021146	32	31 offre#24	0.0592814778	49.21	216	6	2	0				
34	ID104021390	33	32 offre#48	0.0269596077	20.45	52	5	5	147				
35	ID104022181	34	33 offre#24	0.024029113	200.03	198	0	0	2152				
36	ID104022386	35	34 offre#24	0.0454050951	18.02	25	4	4	305				
37	ID104024185	36	35 offre#24	0.0583008374	36.55	25	0	0	2178				
38	ID104025451	37	36 offre#32	0.0107824725	411.43	456	61	13	18161				
39	ID104026081	38	37 offre#24	0.0549965012	30.49	72	10	6	32				

Figure 42: Fichier source de type Excel

### 3.3. Nettoyage de données

Il y a quelques problèmes principaux qui nous ont obligés à nettoyer les données comme :

- Nous avons trouvé beaucoup de colonnes de type numérique exact mais qui contiennent des cellules de type texte contenant la valeur "NULL", nous les avons donc remplacées par la valeur "0" pour les rendre de type numérique exact.
- Nous avons également trouvé des colonnes avec des valeurs de type "REAL" mais qui ont de grands chiffres après la virgule, donc après discussion avec le maître du stage, il a suggéré que nous arrondissions ces valeurs à des valeurs numériques exactes avec la fonction ROUND (Nombre à arrondir;0).
- Nous avons trouvé des colonnes Date de type personnalisées sous la forme "dd:mm:yyyy:00:00:00" donc nous les avons converties en type Date sous la forme "dd/mm/yyyy".

### 3.4. Connexion à la base de données

La figure suivante montre la configuration requise pour établir une connexion avec l'entrepôt de données de destination "DWH\_TT".

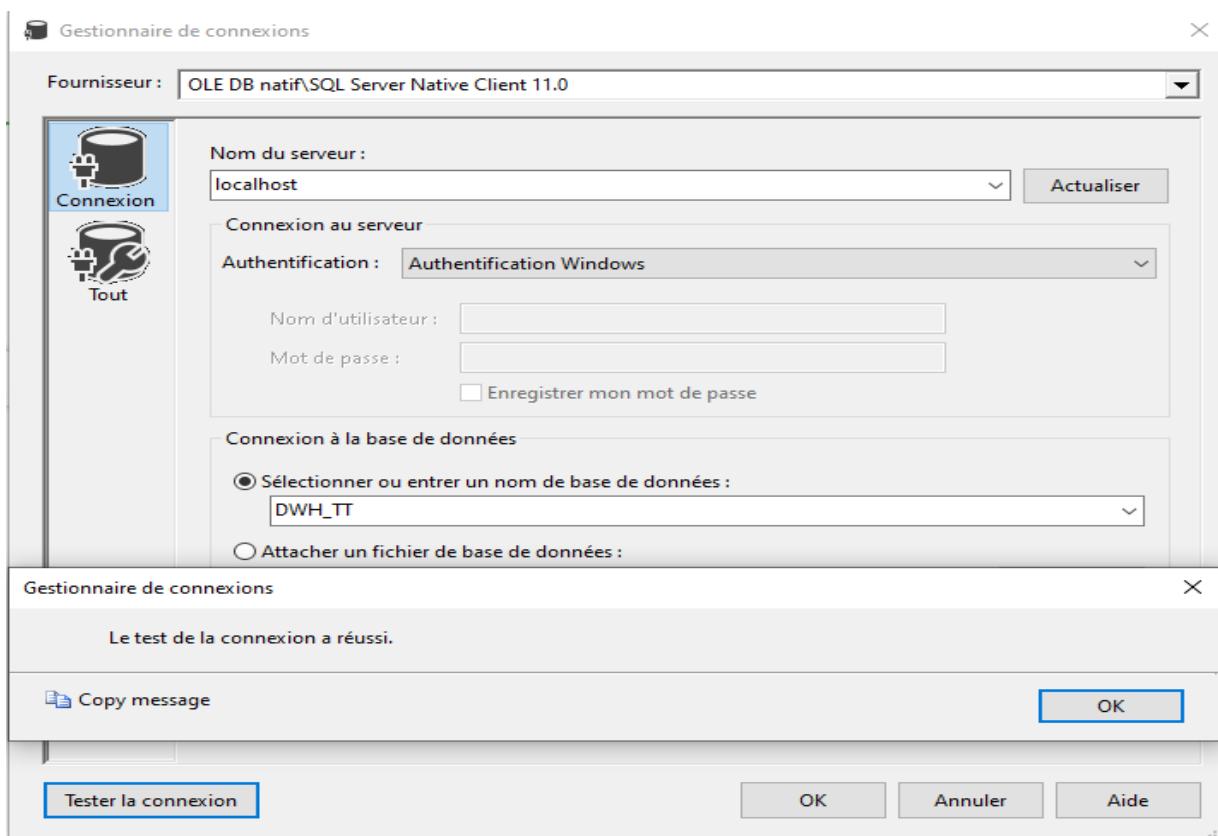


Figure 43: la connexion à l'entrepôt de données de destination « DWH\_TT »

## 4. Chargement des Datamarts

### 4.1. Datamart “Trafic\_Forfait”

Nous allons prendre l'exemple du datamart "Trafic\_Forfait" pour expliquer le processus de chargement d'un datamart.

#### 4.1.1. Chargement des dimensions

- **Dimension Date** : Cette dimension a pour rôle de stocker les informations relatives aux différentes dates.

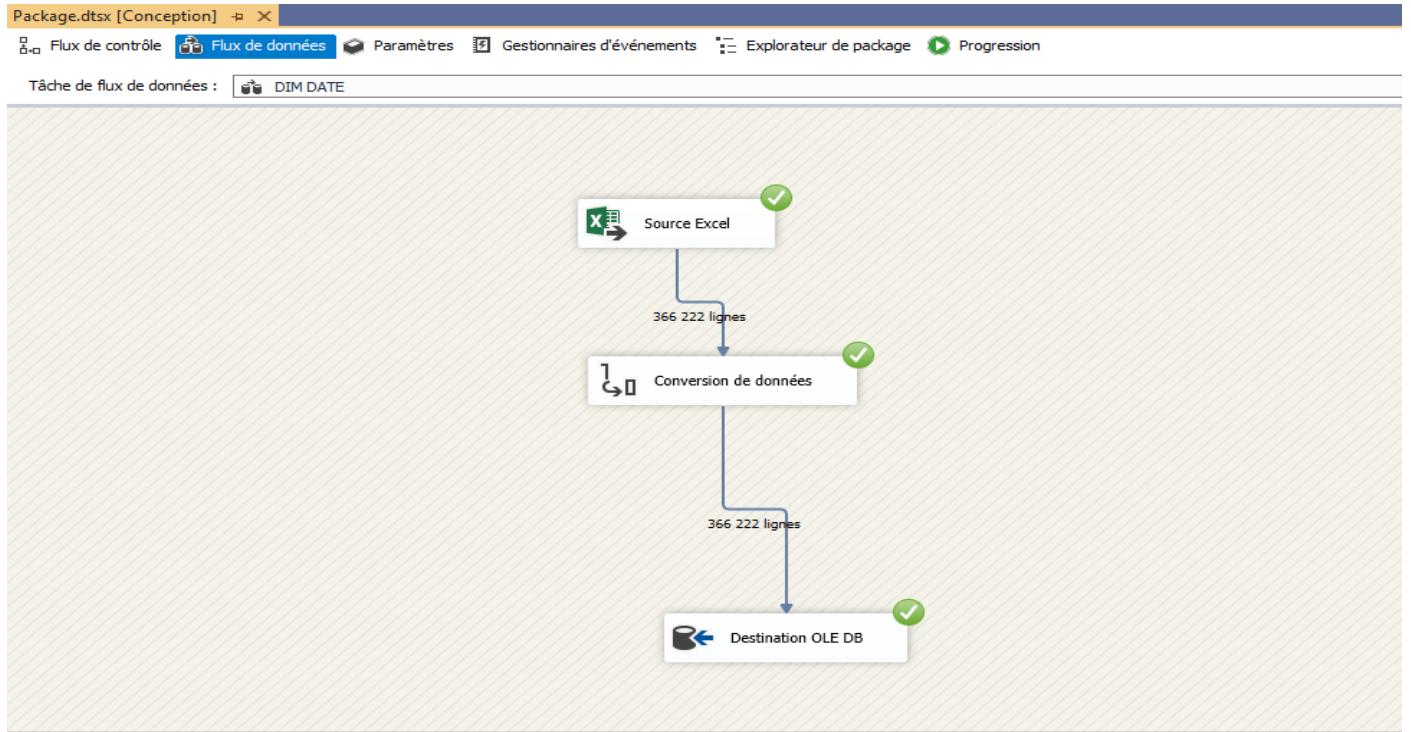


Figure 44: Dimension Date

- **Dimension Client** : Le rôle de cette dimension est de stocker des informations relatives aux clients, sachant que la colonne Age\_Client est calculée à l'aide de la transformation Colonne dérivée en utilisant l'expression “DATEDIFF (“yyyy”, DATE\_NAISSANCE, GETDATE())”.

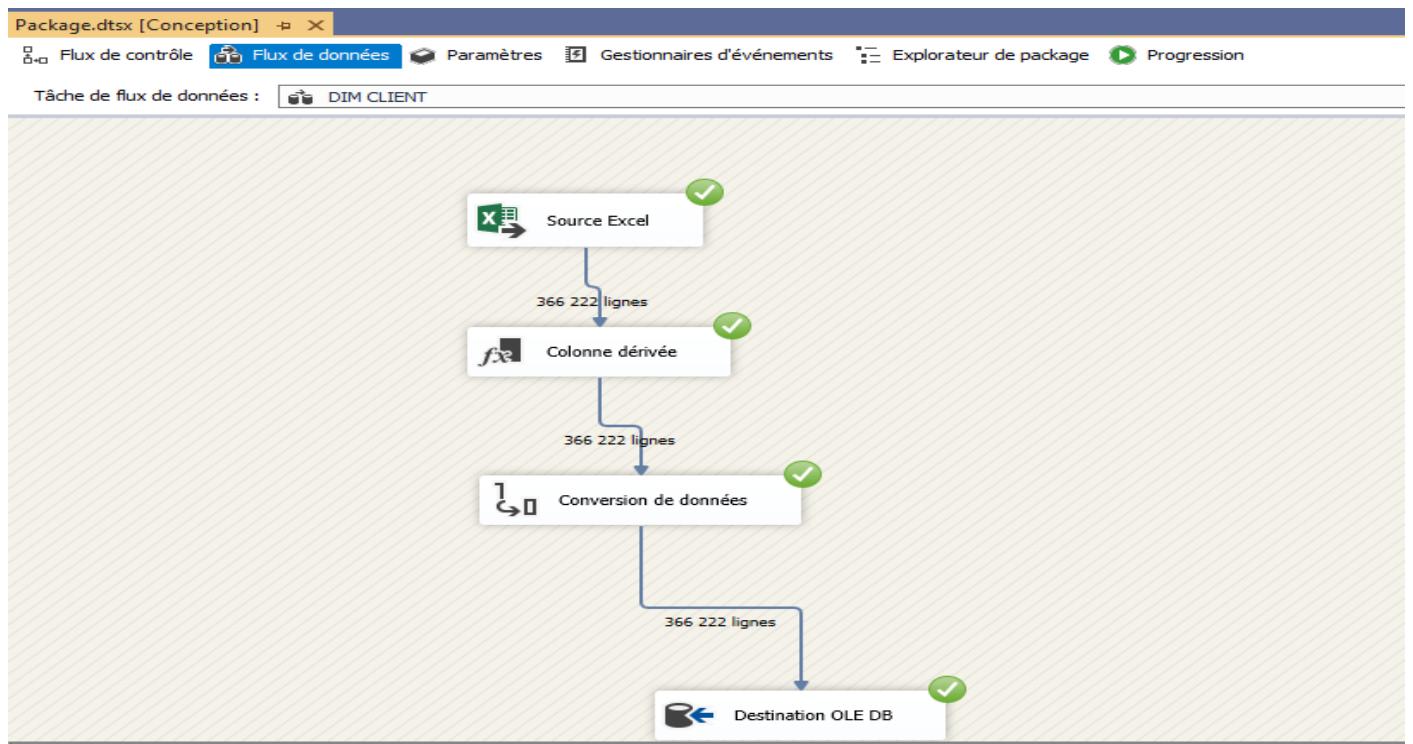


Figure 45: Dimension Client

**- Dimension Transact :** Le rôle de cette dimension est de stocker des informations sur les différentes transactions.



Figure 46: Dimension Transact

- **Dimension Forfait** : Le rôle de cette dimension est de stocker des informations relatives aux différents forfaits de TT.

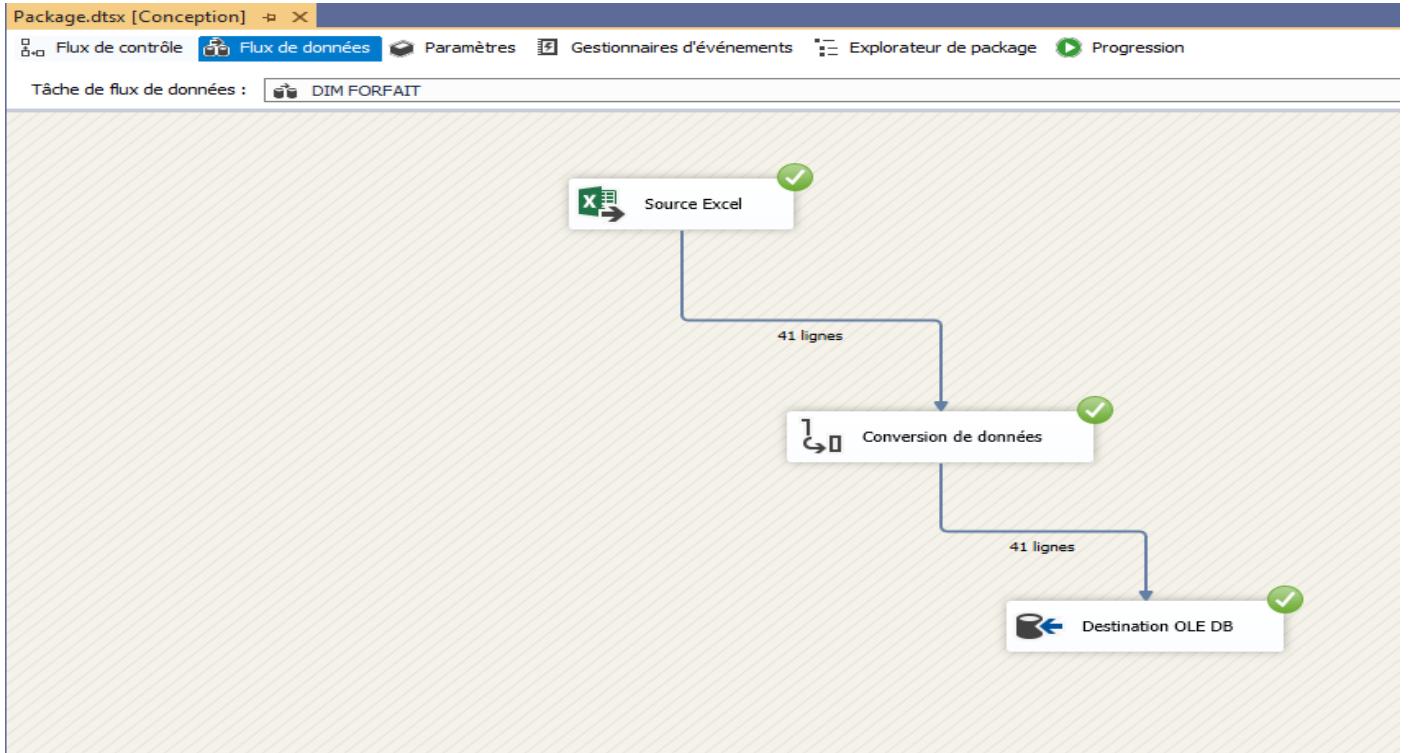


Figure 47: Dimension Forfait

#### 4.1.2. Chargement de la table de faits Trafic\_Forfait :

La table de faits "Trafic\_Forfait" assure le calcul des indicateurs liés aux forfaits de Tunisie Telecom, elle sera alimentée par la jointure de deux fichiers sources excel et aussi, par les dimensions qui lui sont liées.

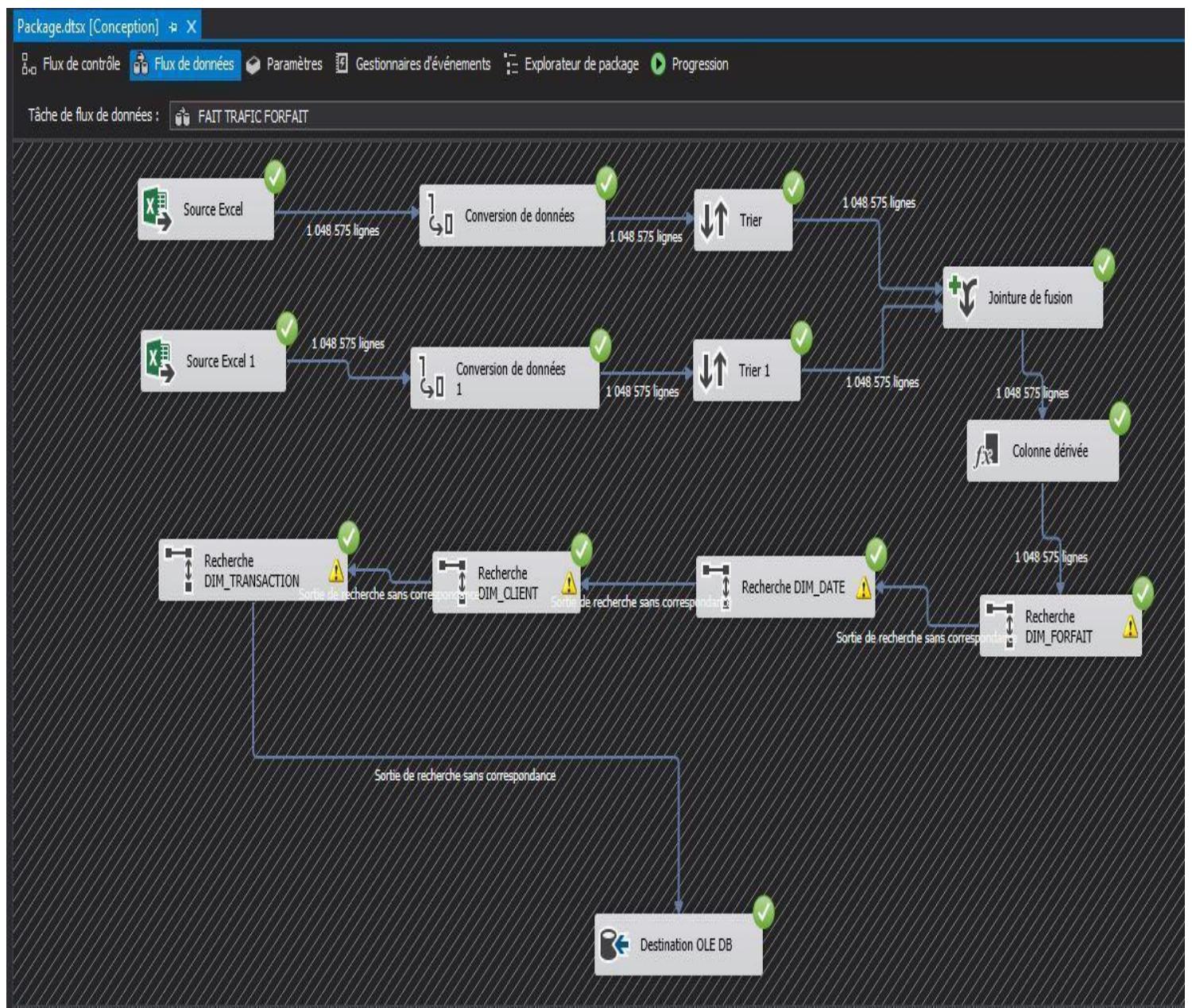


Figure 48: Table de faits “Trafic\_Forfait”

#### 4.2. Chargement des autres tables de faits

Les deux autres tables de faits " Trafic\_Offre " et " Trafic\_Service " seront présentées dans les figures suivantes :

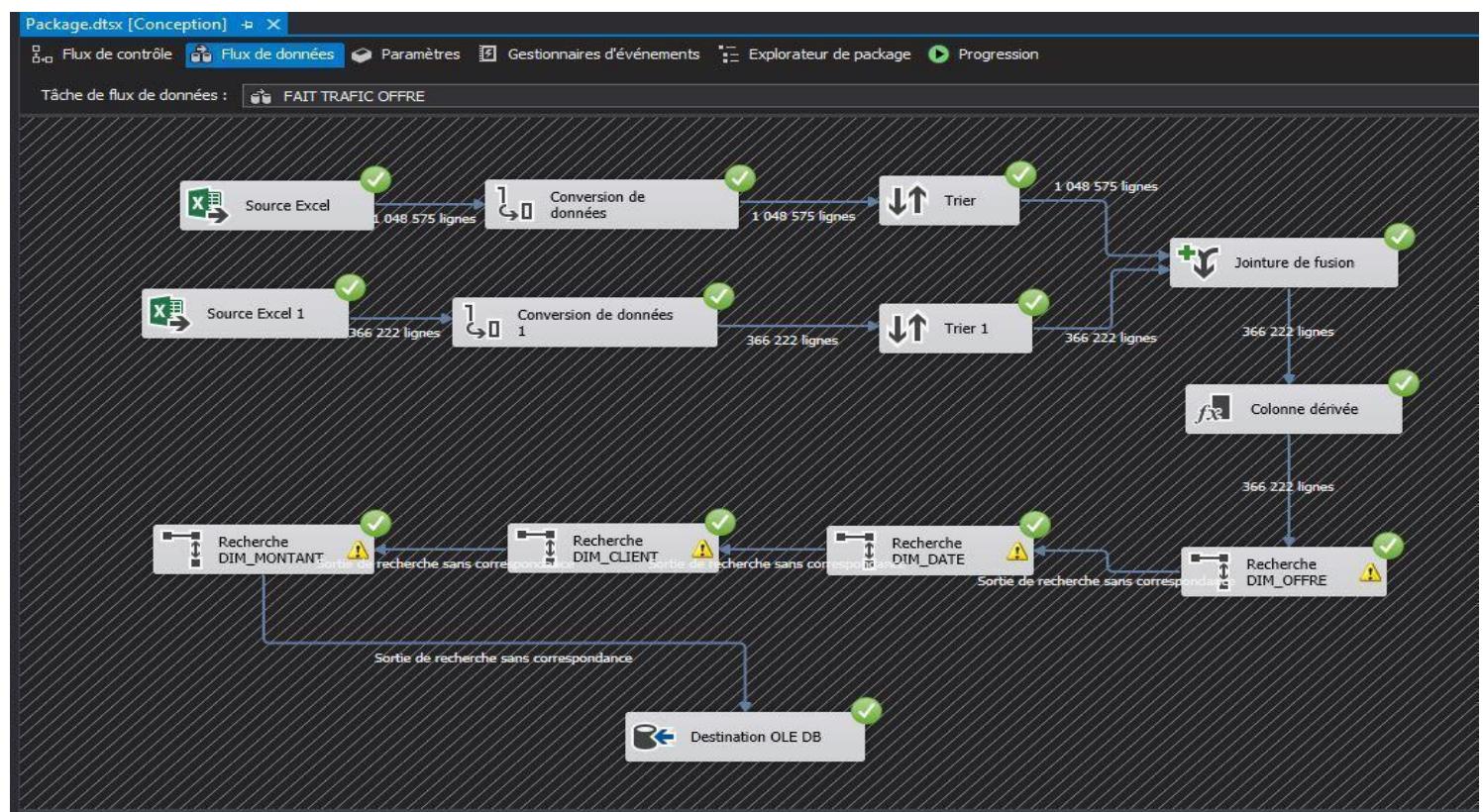


Figure 49: Table de faits “Trafic\_Offre”

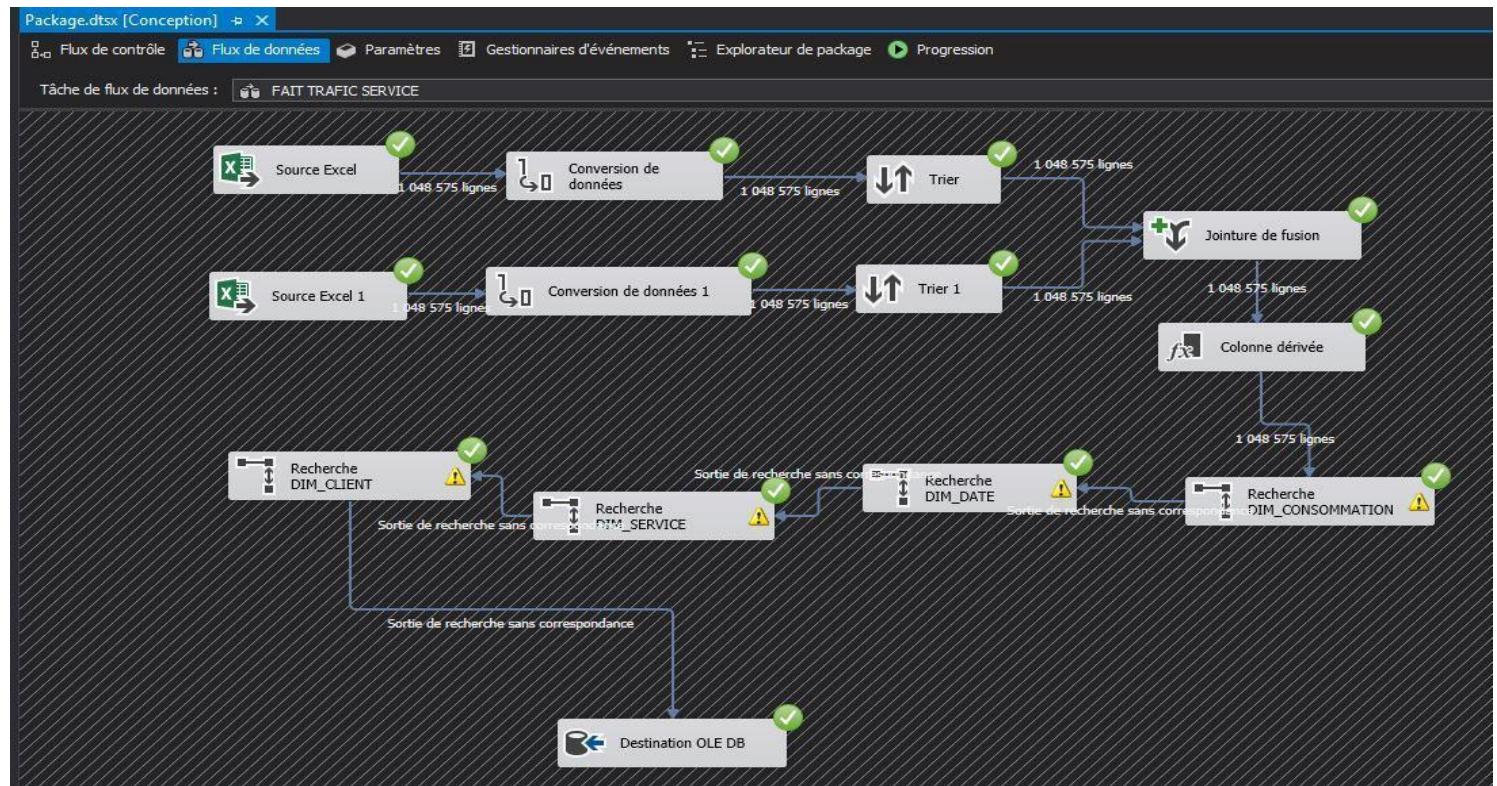


Figure 50: Table de faits “Trafic\_Service”

## **5. L'entrepôt de données**

Le processus d'alimentation commence par une tâche considérée comme la plus importante dans un projet BI, à savoir l'extraction des données de différentes sources. En effet, la difficulté de cette phase consiste essentiellement à sélectionner les données à extraire. Toutefois, dès leur extraction, ces données vont passer par un processus de traitement dont le but est uniquement de les convertir des données d'origine en informations utiles, et ces informations sont aussi significatives pour les décideurs de l'entreprise. Puis, nous poursuivons avec le chargement, nous considérons qu'il s'agit d'une étape relativement simple, car en suivant la stratégie de chargement, il suffit de transférer les données préalablement préparées vers la table cible. Ainsi, la première étape consiste à charger tout d'abord toutes les dimensions, puis toutes les tables de faits. Pour chaque magasin de données, il faut remplacer l'identifiant de la table de faits dans le système source par la clé de remplacement appropriée. Finalement, l'identifiant de la table de faits sera alors constitué des clés des dimensions qui font partie du même fait.

La figure 51 présente l'exécution du processus ETL de chargement de l'entrepôt global.

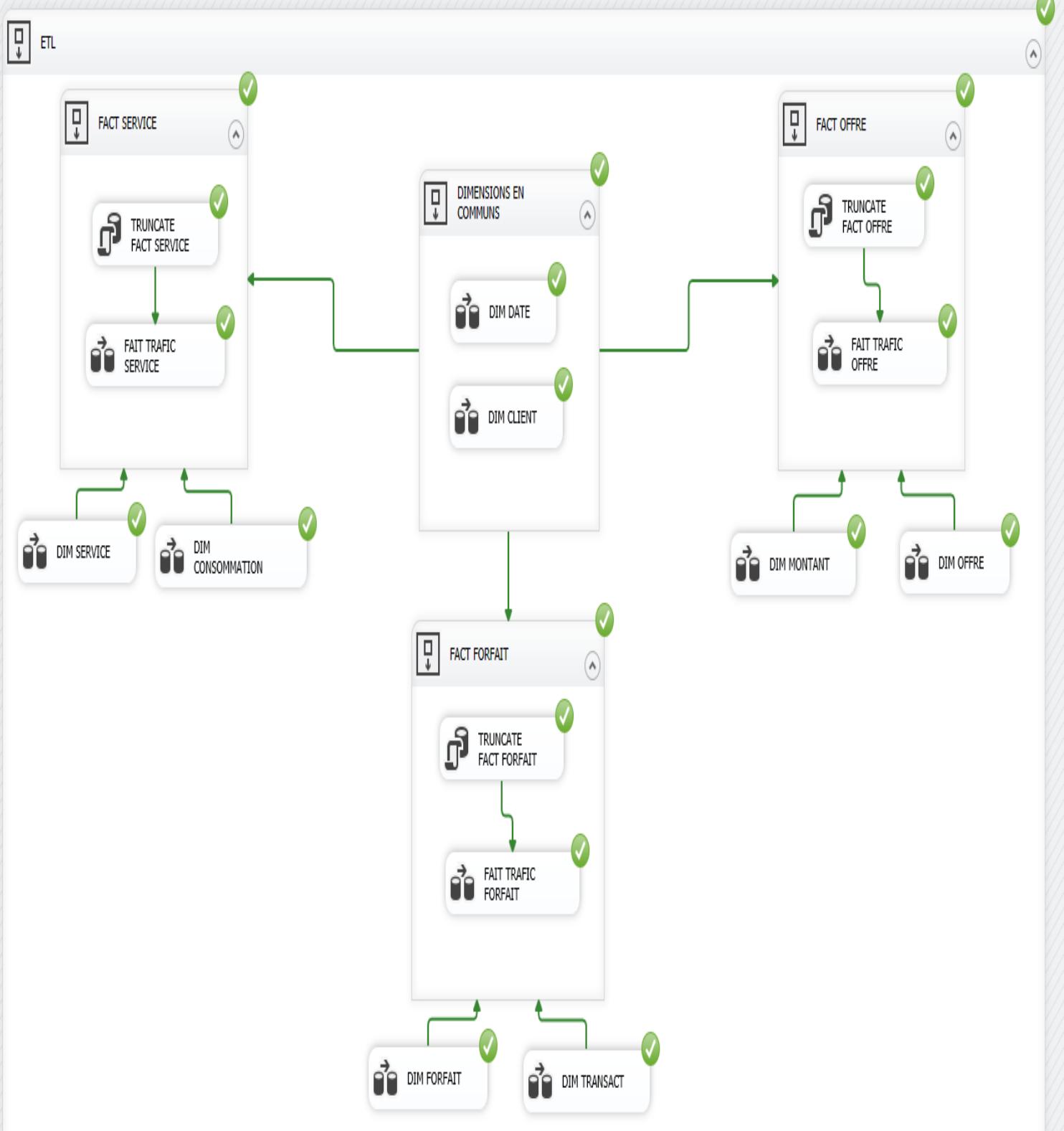


Figure 51: Exécution finale du processus ETL

## 6. Mise en œuvre et analyse des données : Les cubes

Après avoir expliqué le processus ETL, passons maintenant à la présentation de la réalisation des trois cubes OLAP.

### 6.1. Cube Trafic\_Forfait

La figure ci-dessous montre un cube simple avec quatre dimensions : Date, Client, Forfait et Transaction

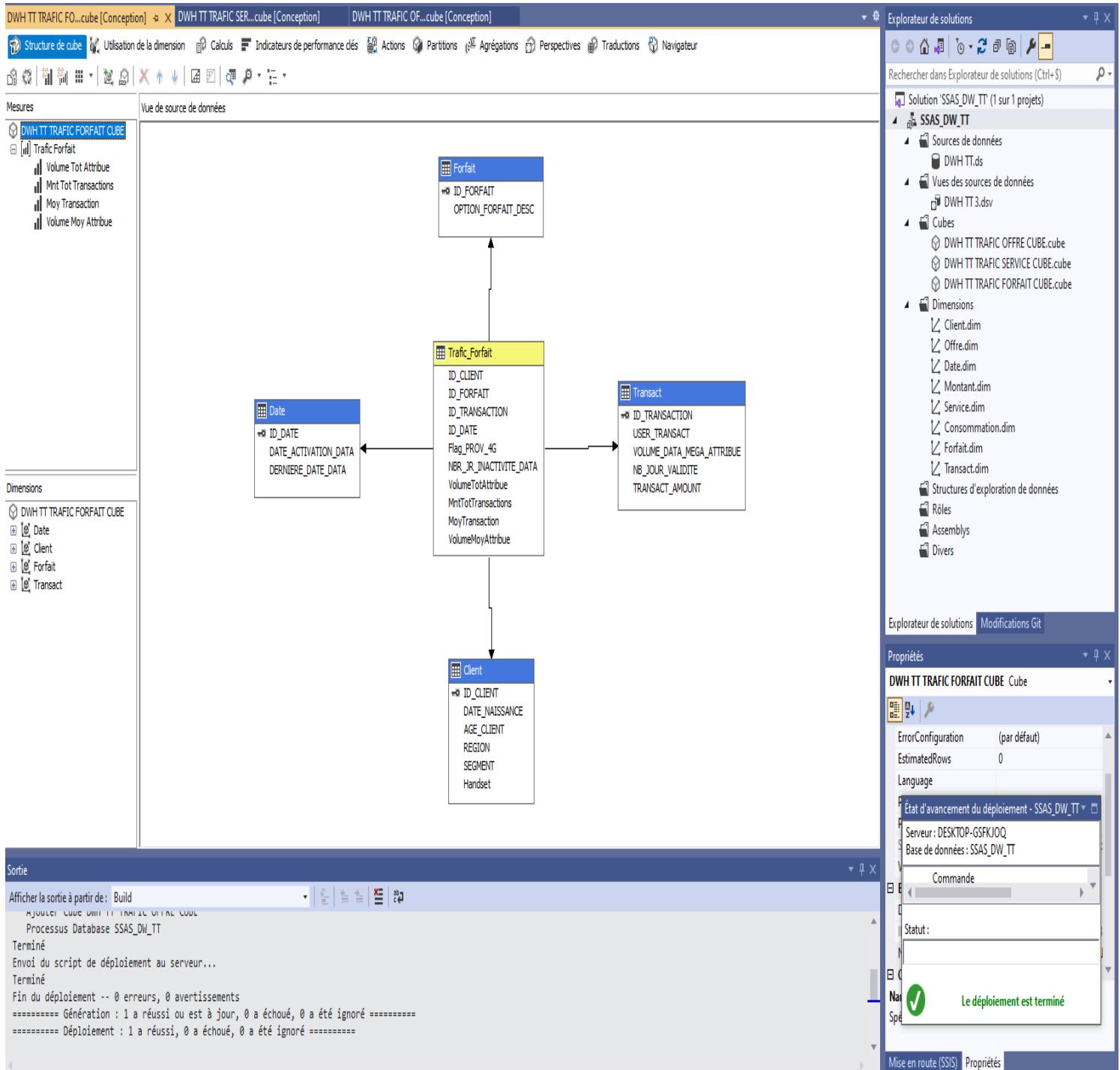


Figure 52: Réalisation et déploiement du cube "Trafic\_Forfait"

## 6.2. Cube Trafic\_Offre

La figure suivante illustre un cube simple comportant quatre dimensions : Date, Client, Offre et Montant

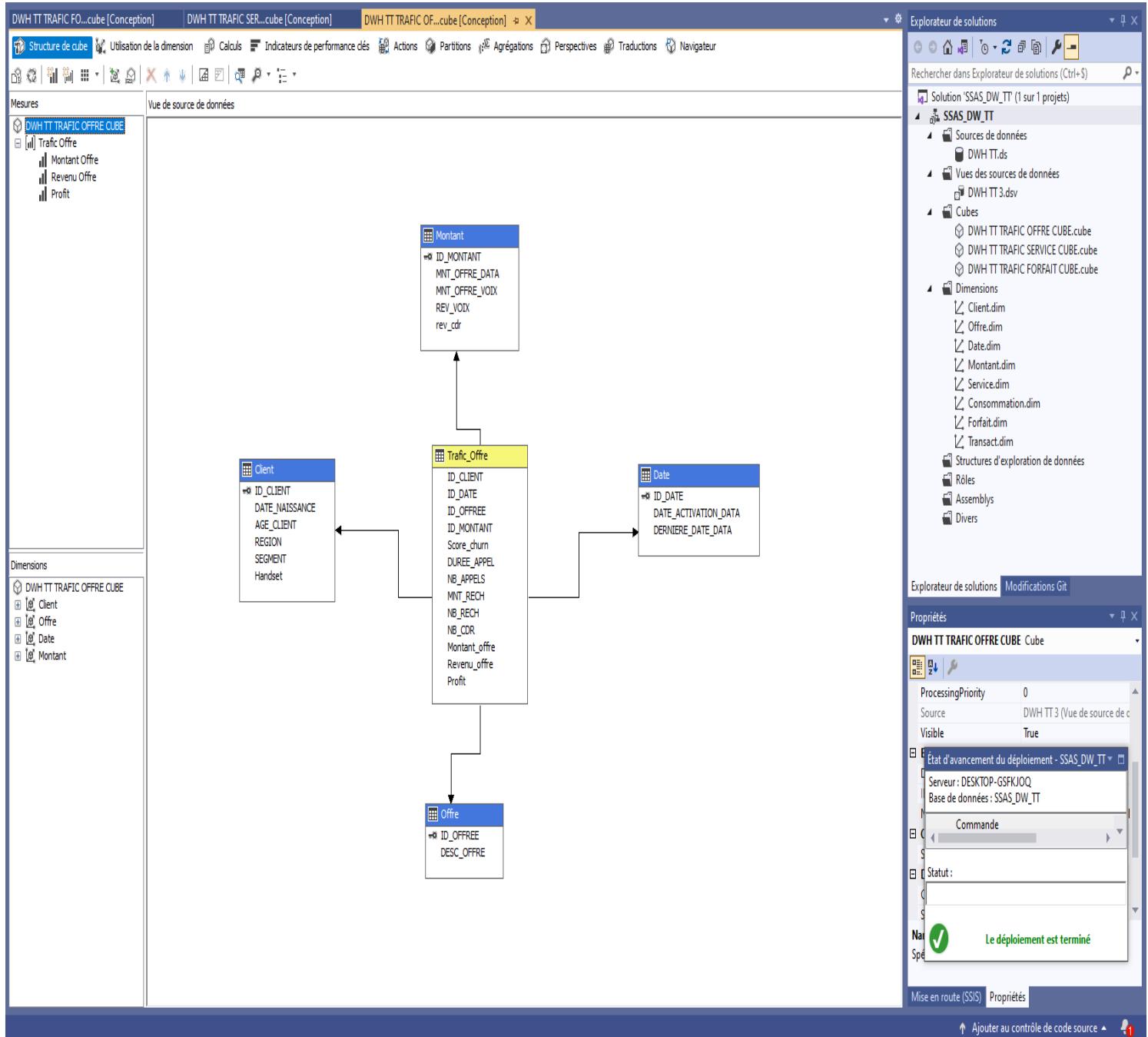


Figure 53: Réalisation et déploiement du cube “Trafic\_Offre”

## 6.3. Cube Trafic\_Service

La figure qui suit montre un cube simple comportant quatre dimensions : Date, Client, Service et Consommation.

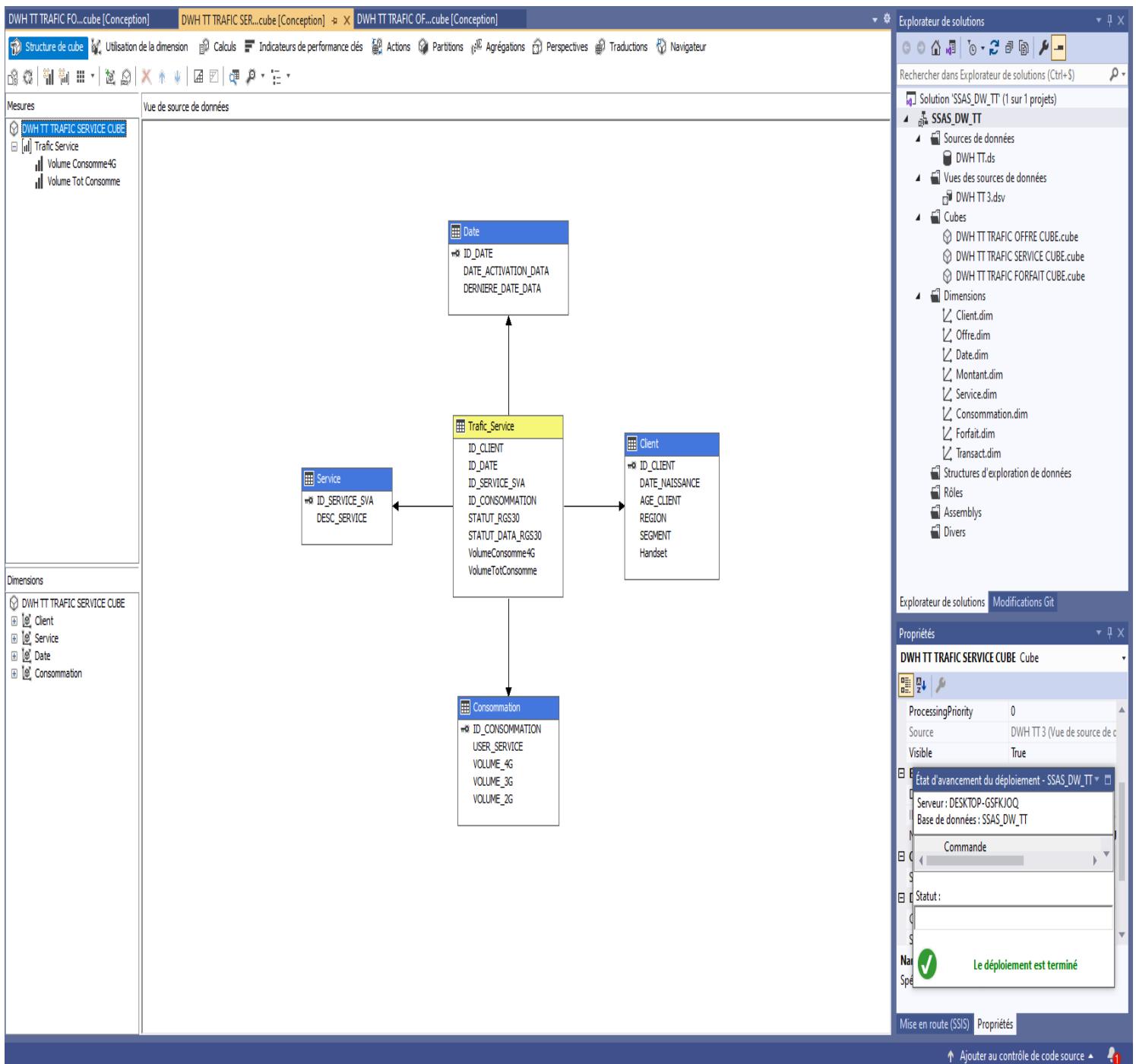


Figure 54: Réalisation et déploiement du cube “Trafic\_Service”

## 7. Réalisation du reporting : Les tableaux de bord

Notre expérience a prouvé que la visualisation de données est un outil efficace capable de prendre un grand ensemble de données, de l'analyser, de le quantifier et de le présenter d'une manière claire et facile à lire. Il n'est donc pas surprenant que la visualisation de données soit désormais indispensable.

Dans cette optique, nous allons présenter dans la section suivante des tableaux de bord générés avec Power BI.

## 7.1. Obtention des données à partir de la base de données SSAS

Les deux figures suivantes montrent les étapes à suivre pour obtenir les données de la base de données SQL Server Analysis Services.

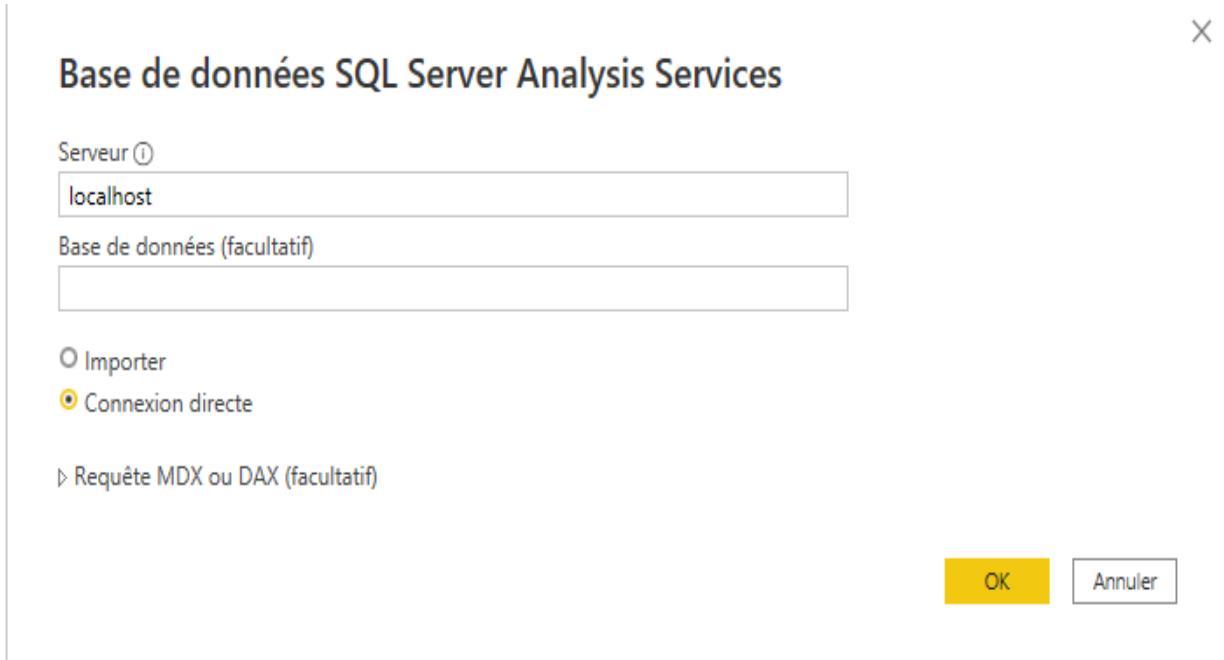


Figure 55: Connexion au serveur “localhost”

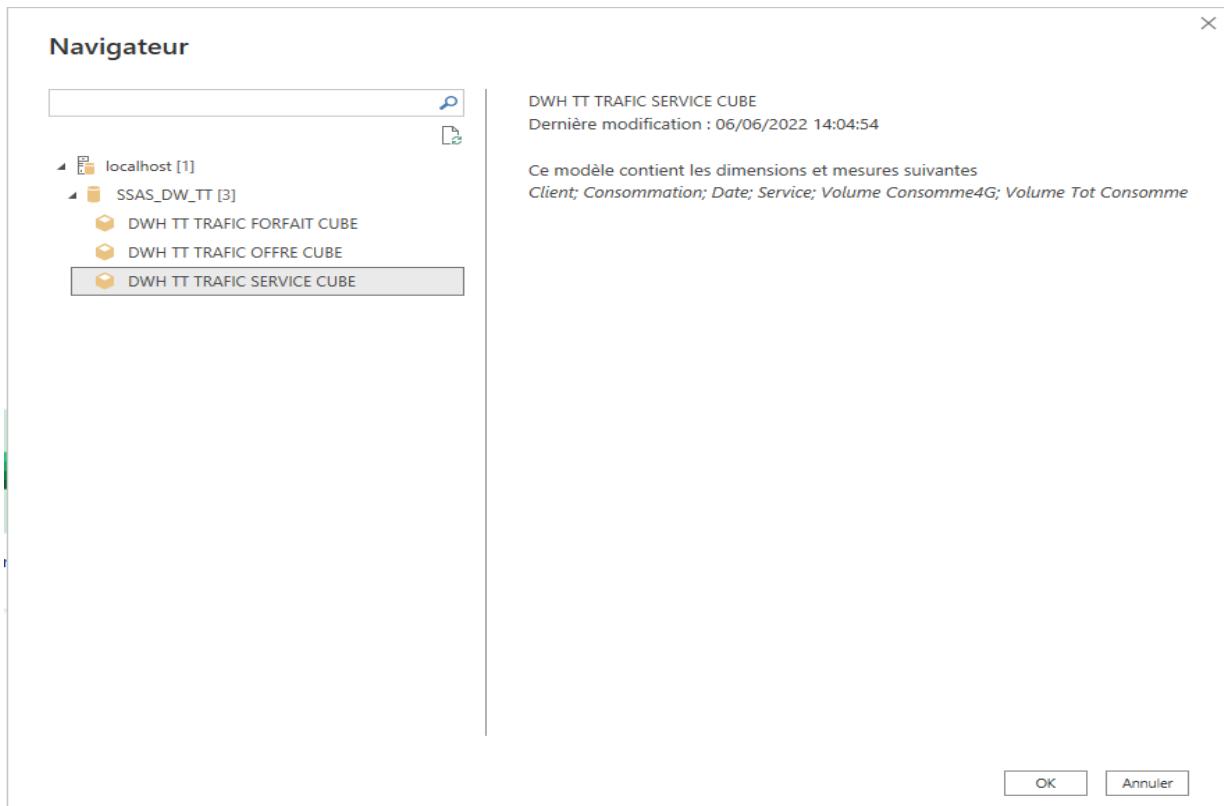


Figure 56: Exemple d’obtention des données du cube “Trafic\_Service”

## 7.2. Tableau de bord "Trafic\_Offre"

La figure 57 ci-dessous montre la première page du tableau de bord intitulé " TUNISIE TELECOM CVM DATAMINING DASHBOARDS ". Le tableau présente une vue globale de l'analyse des revenus des offres, illustrant des indicateurs filtrés par offre, client, âge du client, segment, région, date et handset.

Il est composé des visualisations suivantes :

- 5 cartes qui représentent les revenus des offres, le montant des offres, le nombre de clients, le montant des recharges et le profit des revenus.
- Un graphique en courbes et histogrammes groupés où l'on retrouve les revenus des offres (voix et cdr) par segment.
- un graphique en courbes et histogrammes empilés où l'on trouve les revenus des offres (voix et cdr) par âge du client.
- un graphique en aires qui représente les revenus des offres par région.
- un graphique en courbes et histogrammes groupés représentant les revenus des offres (voix et cdr) par handset.

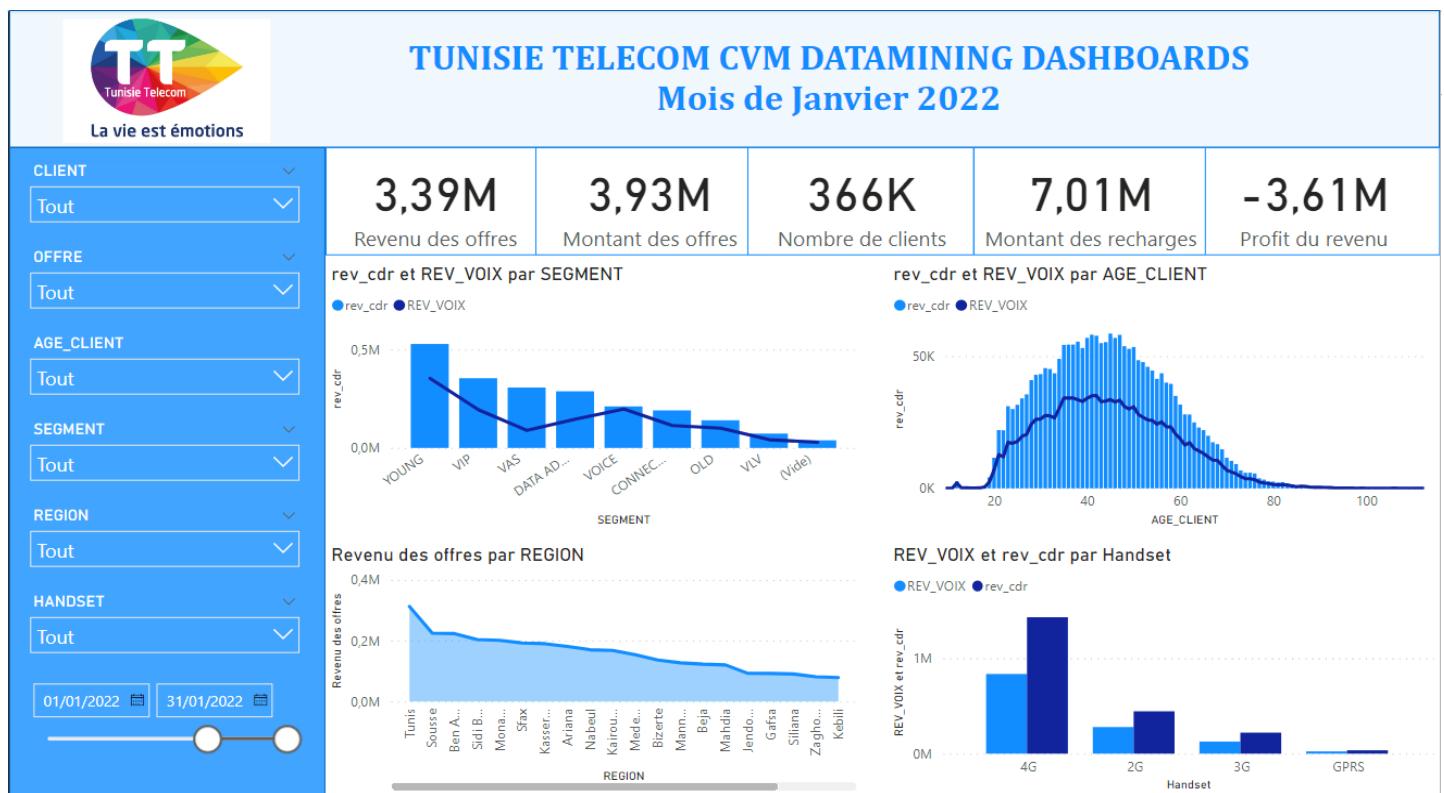


Figure 57: Page 1 TDB "Trafic\_Offre"

La figure 58 ci-après montre la deuxième page du tableau de bord intitulé "TUNISIE TELECOM CVM DATAMINING DASHBOARDS".

Elle est composée des visualisations suivantes :

- 3 graphiques représentant le nombre total de recharges, le nombre total de cdr et le nombre total d'appels.
- Un graphique à barres groupées représentant les revenus voix et cdr par handset.
- Un entonnoir représentant les 10 meilleures offres en janvier.
- Un graphique en courbe représentant le nombre d'appels, le nombre de cdr et le nombre de recharges par segment et par âge du client.

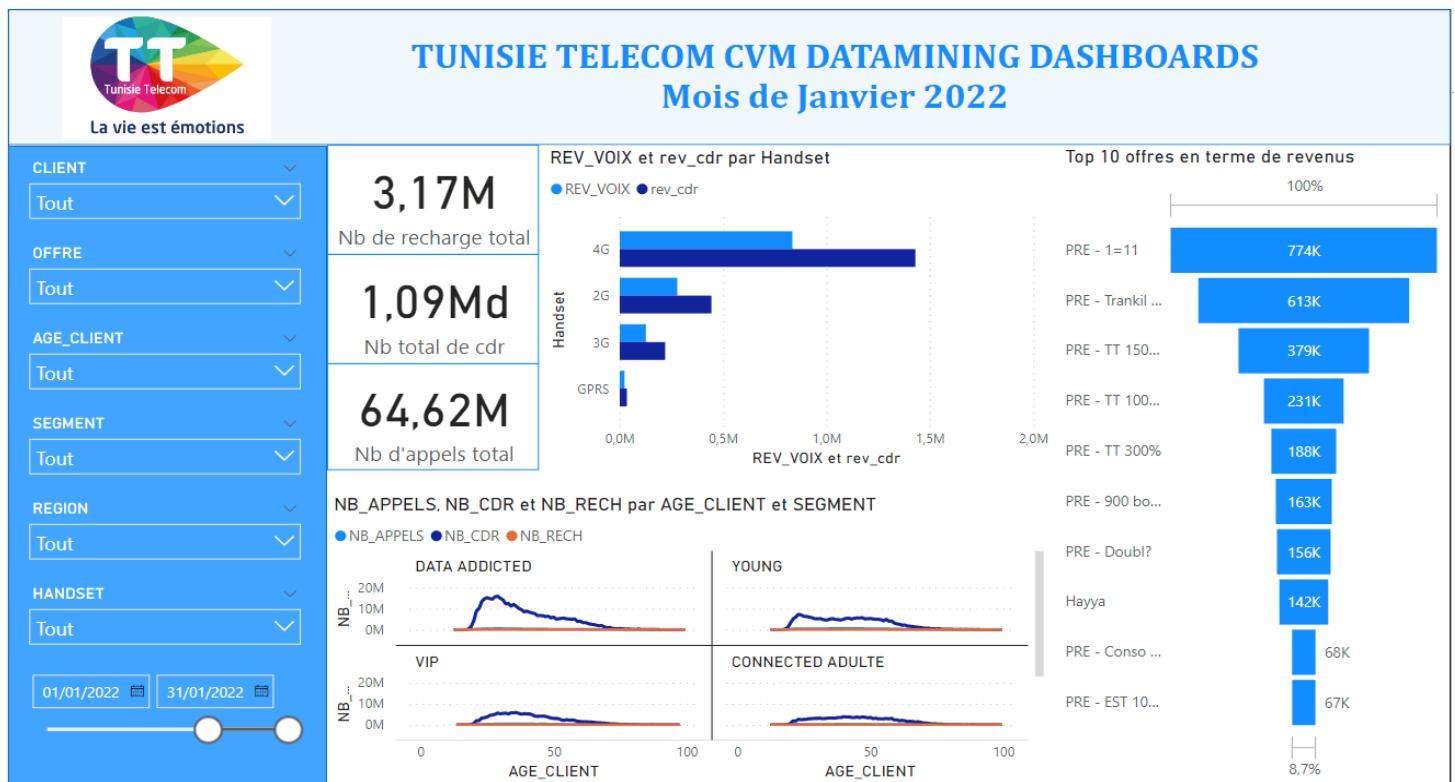


Figure 58: Page 2 TDB "Trafic\_Offre"

### 7.3. Tableau de bord "Trafic\_Forfait"

La figure 59 présente la première page du tableau de bord intitulé "TUNISIE TELECOM CVM DATAMINING DASHBOARDS". Le tableau de bord donne une vision globale des volumes attribués et des montants des transactions, illustrant des indicateurs filtrés par forfait, client, âge du client, segment, région, date et handset.

Il est composé des visualisations suivantes :

- 5 cartes qui représentent le revenu des offres, le montant total des transactions, le volume total attribué, le nombre de clients, la transaction moyenne et le volume moyen attribué par client.

- Un graphique en anneau qui représente le volume attribué par segment.
- Un graphique en aires qui représente le volume attribué par région.
- Un treemap qui représente les montants des transactions par forfait.
- Un graphique à barres empilées représentant le volume attribué par forfait.

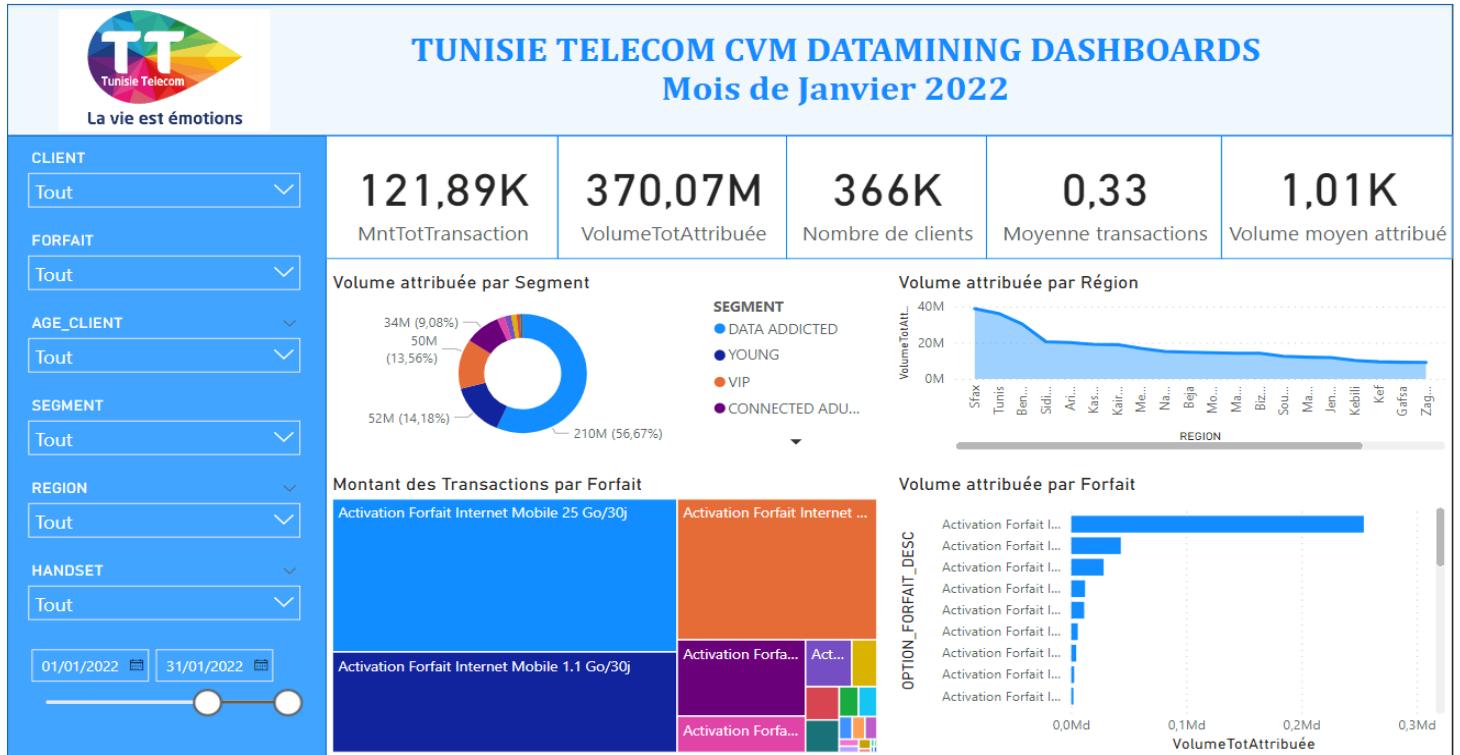


Figure 59: Page 1 TDB "Trafic\_Forfait"

La figure 60 ci-après illustre la deuxième page du tableau de bord intitulé "TUNISIE TELECOM CVM DATAMINING DASHBOARDS".

Elle est composée des visualisations qui suivent :

- Un graphique en courbe et un histogramme empilé représentant le volume attribué et le montant des transactions par âge du client.
- Un graphique en courbe et un histogramme groupé représentant le volume attribué et le montant des transactions par handset.
- Un graphique en cascade qui représente le montant des transactions par segment.
- Un camembert où l'on retrouve le montant des transactions par région.

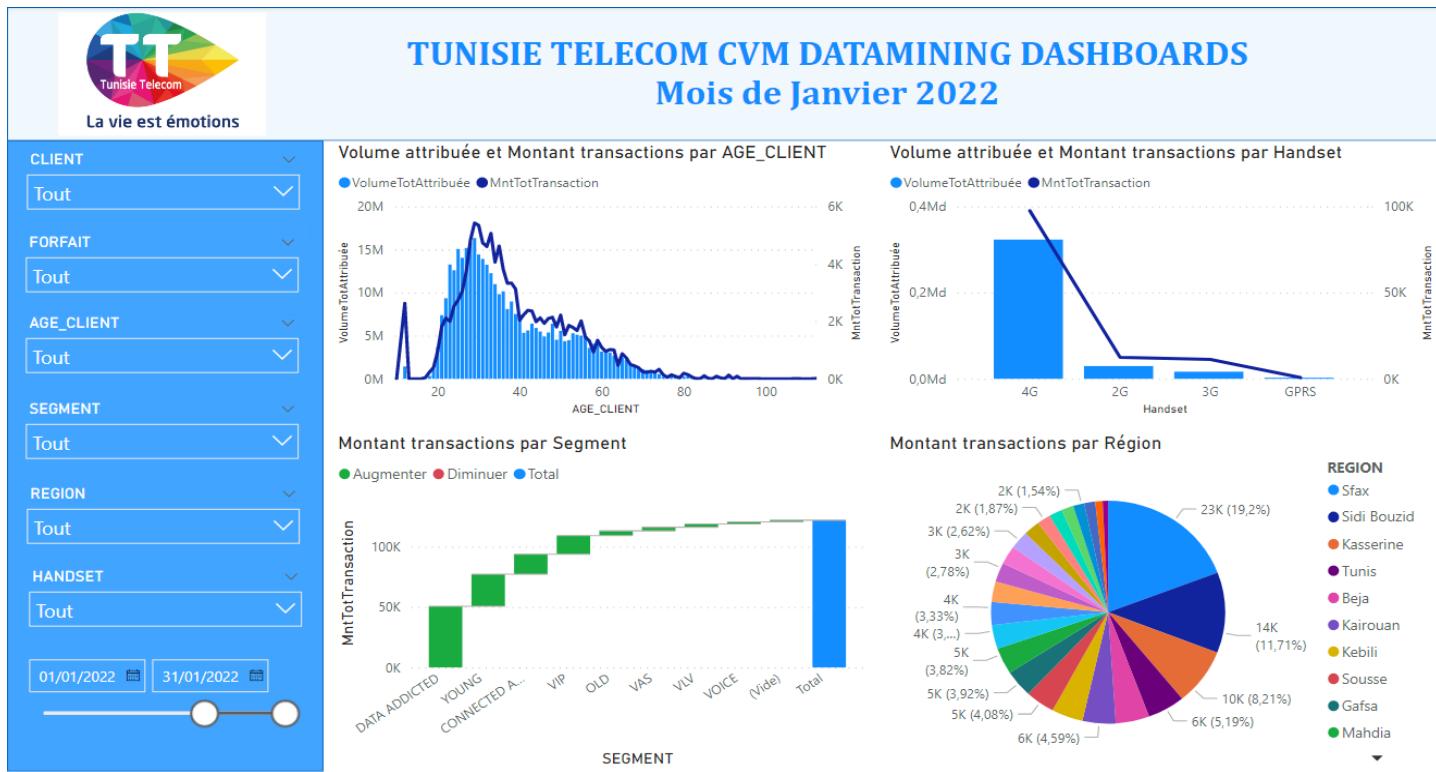


Figure 60: Page 2 TDB "Trafic\_Forfait"

#### 7.4. Tableau de bord "Trafic\_Service"

La figure 61 montre la première page du tableau de bord intitulé "TUNISIE TELECOM CVM DATAMINING DASHBOARDS". Ce tableau de bord donne une vision globale des volumes consommés par service, client, âge du client, segment, région, date et combiné.

Il est composé des visualisations suivantes :

- 5 graphiques qui représentent le volume consommé 4G, le volume consommé 3G, le volume consommé 2G, le volume total consommé et le nombre de clients.
- Un graphique en zone empilé représentant le volume consommé 4G, 3G et 2G par service.
- Un graphique de ruban représentant le volume consommé 4G, 3G et 2G par âge du client.
- Un histogramme empilé où l'on retrouve le volume consommé par région.

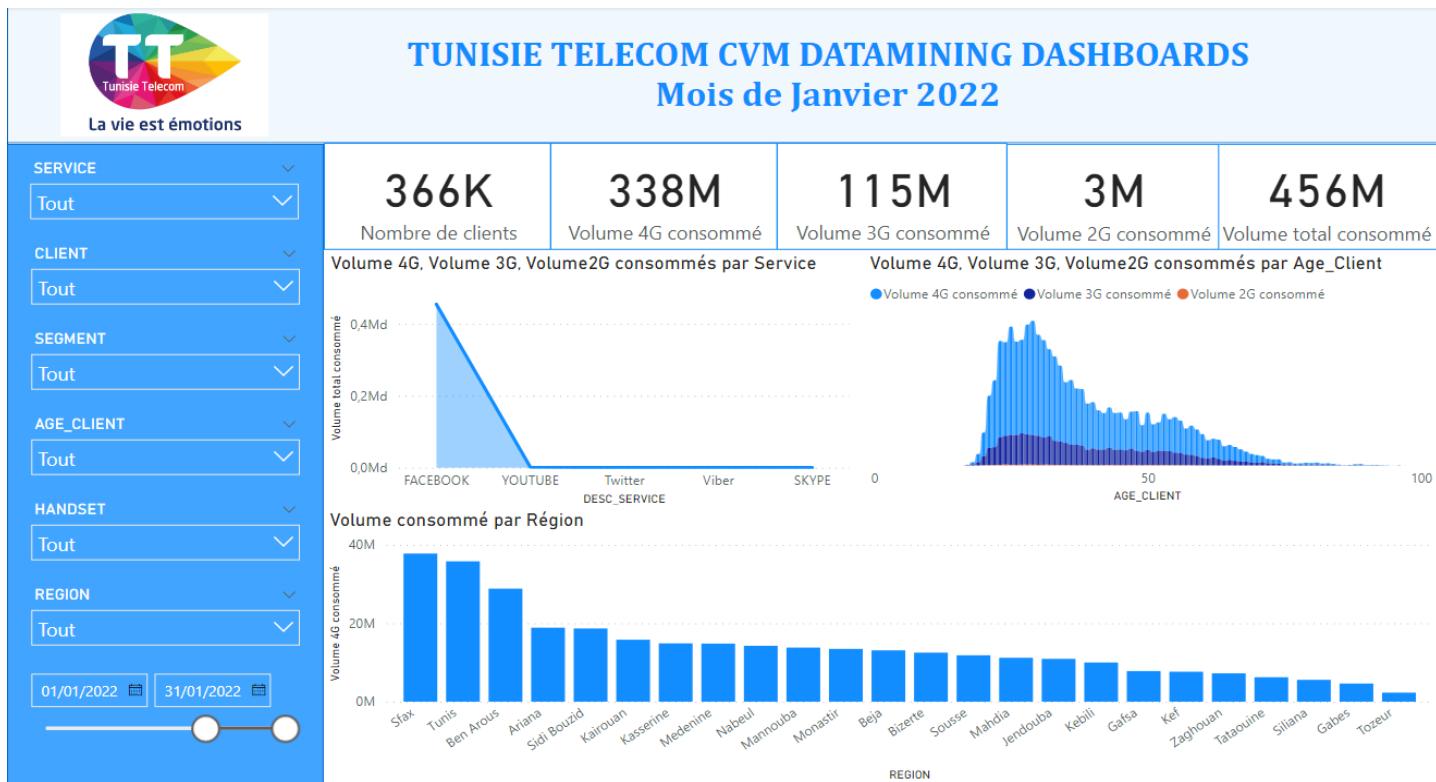


Figure 61: Page 1 TDB "Trafic\_Service"

La figure 62 ci-après montre la deuxième page du tableau de bord intitulé "TUNISIE TELECOM CVM DATAMINING DASHBOARDS".

Elle est composée des visualisations qui suivent :

- Un graphique en courbe et un histogramme groupé représentant le volume total consommé et le volume 4G consommé par handset.
- Un histogramme groupé qui représente le volume 4G, 3G, 2G consommés par segment.
- Deux jauge qui représentent le nombre de clients actifs et le nombre de clients actifs data.

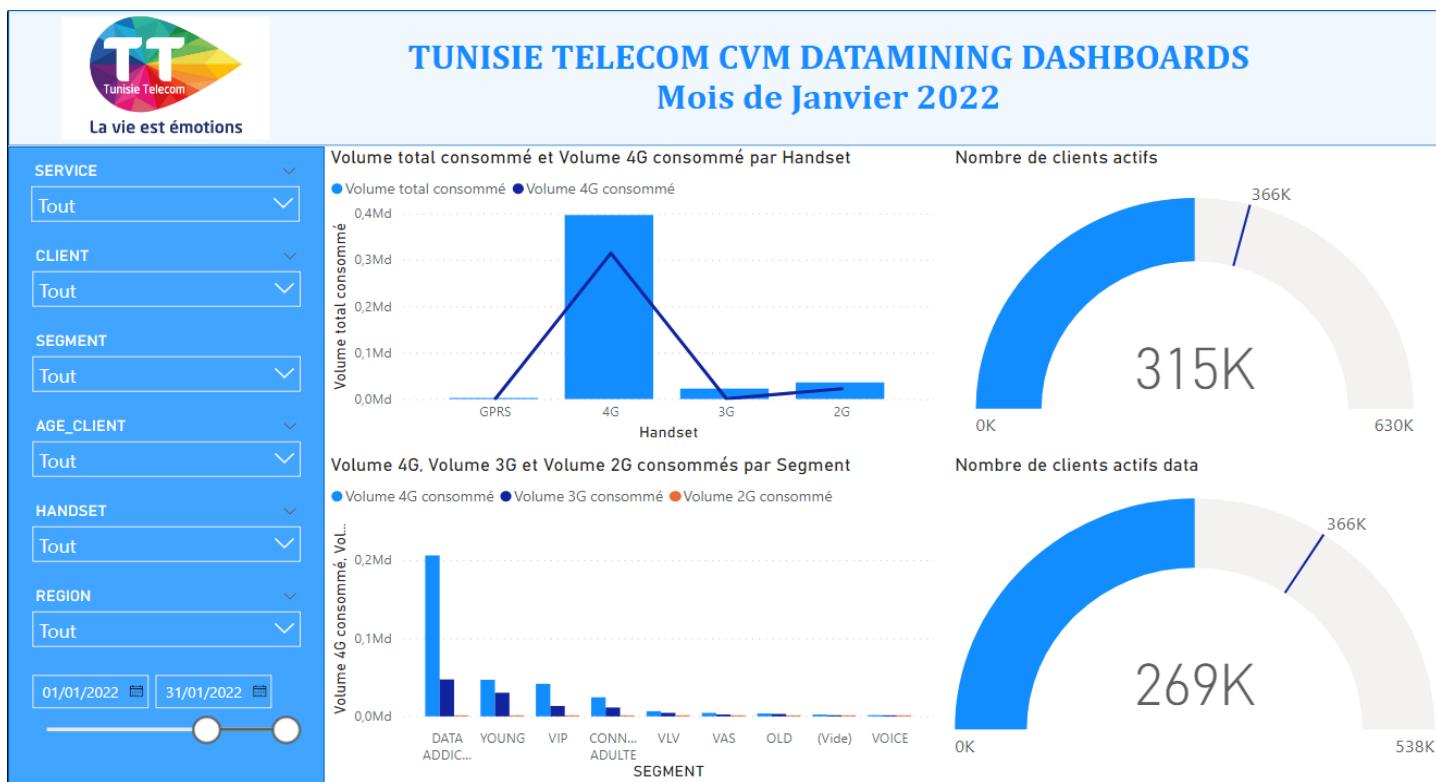


Figure 62: Page 2 TDB “Trafic\_Service”

## 8. Initiative : Implémentation web

Afin d'intégrer les rapports dans le processus de reporting, nous avons mis en place une implémentation web en développant une plateforme web avec un formulaire d'authentification avec un accès privé uniquement pour le directeur du département CVM datamining de Tunisie Télécom qui est M. Mohamed JARBOUI.

Nous avons créé une base de données dans phpmyadmin que nous avons la nommer "siteweb" dans laquelle nous avons créé une table nommée "login" contenant les champs du formulaire qui subissent un contrôle de sorte que le nom & prénom doivent être "Mohamed Jarboui", l'adresse email doit être "mohamed.jarboui@tunisietelecom.tn" et le mot de passe doit être "TT202122".

Si le décideur a inséré un champ incorrect, c'est-à-dire le nom et/ou l'adresse email et/ou le mot de passe, une boîte d'alerte contenant le message "You are not registered to access this content" et un bouton OK s'affichent, comme le montre la figure suivante.

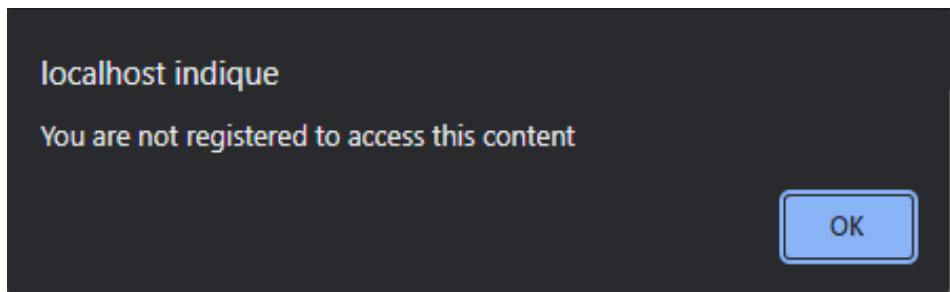


Figure 63: Boîte d'alerte

Dès qu'il se connecte, le directeur pourra choisir le tableau de bord à consulter et pourra modifier les filtres des tableaux de bord pour une bonne prise de décision.

Les figures suivantes montrent les étapes à suivre par le directeur pour consulter les tableaux de bord afin de prendre de bonnes décisions au bon moment.

	<b>id</b>	<b>username</b>	<b>email</b>	<b>password</b>
	2	Mohamed Jarboui	mohamed.jarboui@tunisitelecom.tn	TT202122

Figure 64: Le contrôle sur les champs du formulaire de connexion



La vie est émotions

**Accès privé**

Nom & Prénom :  
Mohamed Jarboui

E-mail :  
mohamed.jarboui@tunisitelecom.tn

Mot de passe :  
.....

**Se Connecter**

Figure 65 : Permettre au directeur du département CVM datamining de se connecter

## SUIVI DES TABLEAUX DE BORD TT

DASHBOARDS TT      [Trafic OFFRE](#)      [Trafic FORFAIT](#)      [Trafic SERVICE](#)

Figure 66 : Choix du tableau de bord à consulter

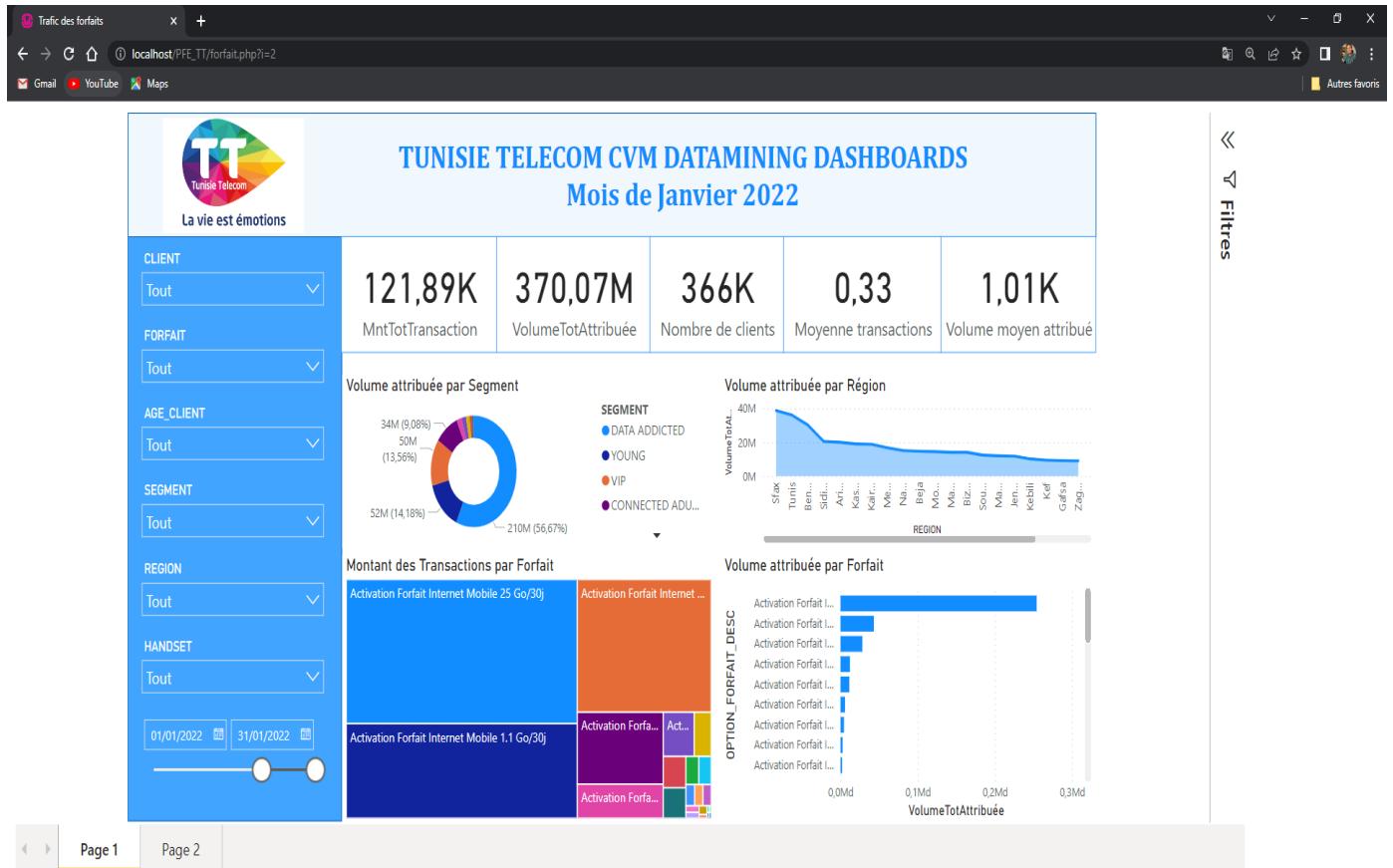


Figure 67 : Tableau de bord Trafic\_Forfait

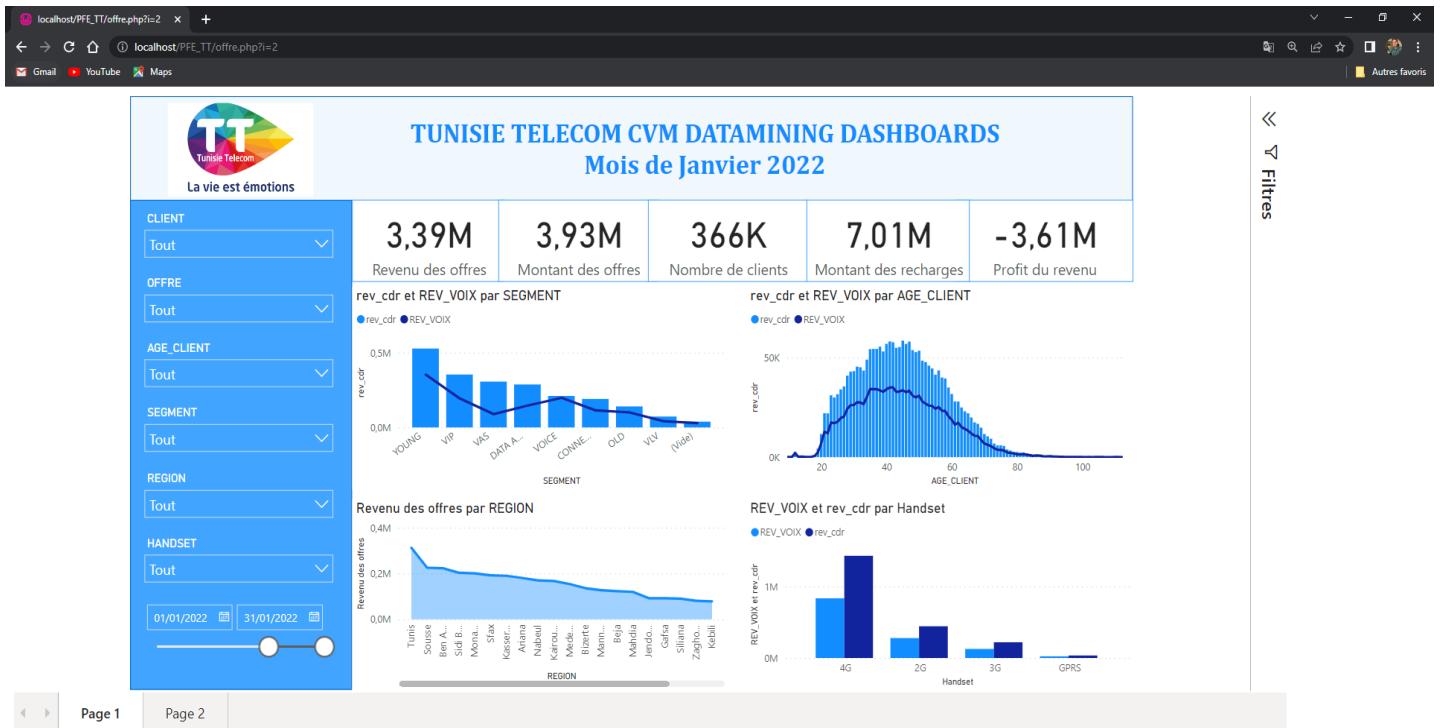


Figure 68 : Tableau de bord Trafic\_Offre

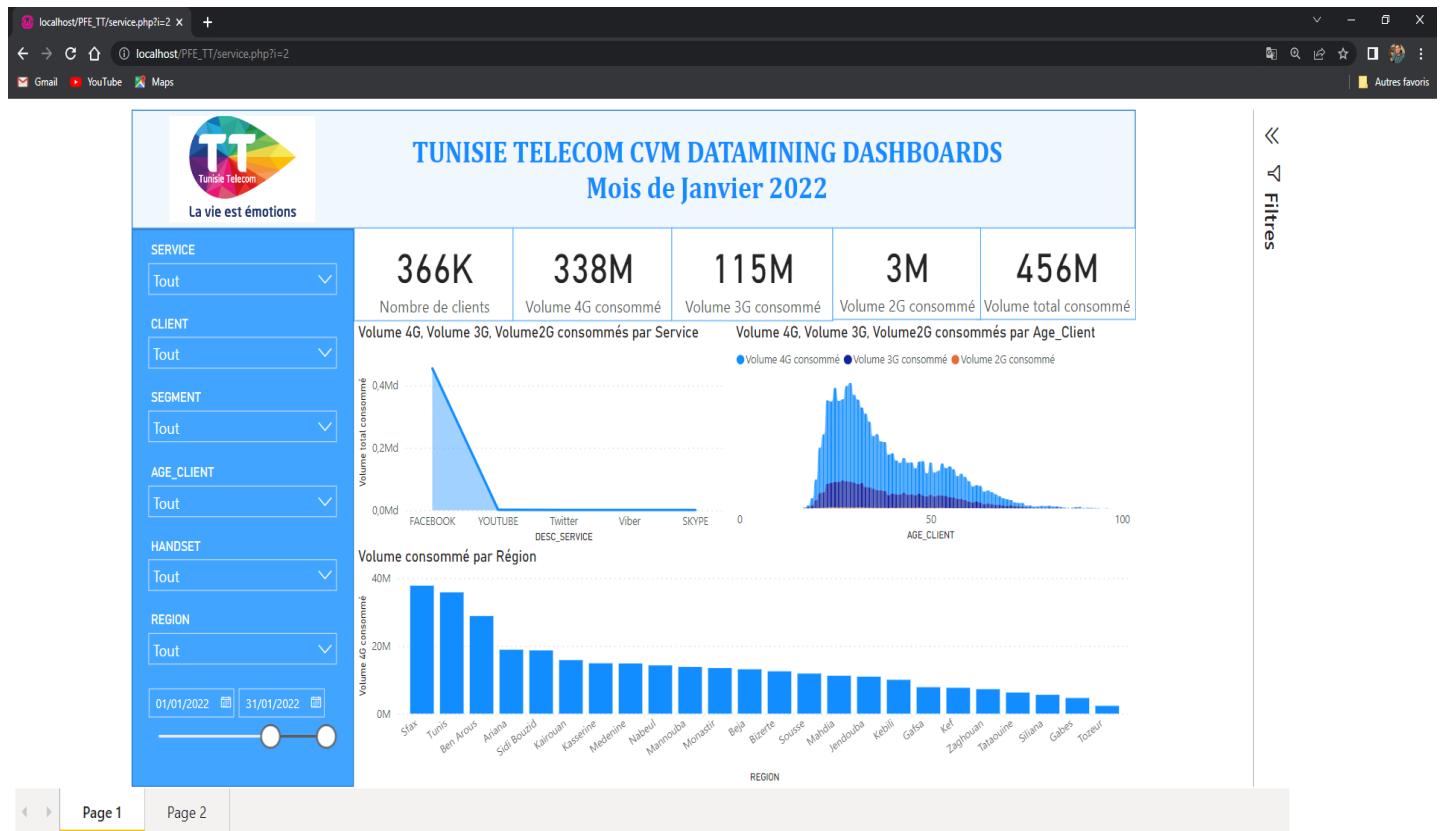


Figure 69 : Tableau de bord Trafic\_Service

## 9. Conclusion

Dans ce chapitre, nous concluons notre projet en démontrant notre environnement en termes de matériel et de logiciel ainsi que les outils et langages de développement mis en œuvre selon une architecture bien précise. Ensuite, nous avons présenté la mise en œuvre des outils ETL, des outils d'analyse et des outils de reporting.

Pour finir, nous avons développé une petite application qui est utilisée pour afficher des tableaux de bord au directeur du département CVM datamining.

# Conclusion générale

Tout au long de cette étude, nous nous sommes attachés à concevoir et à alimenter des tableaux de bord, des indicateurs de performance qui permettent de suivre le comportement des clients et les revenus à travers un système décisionnel. Ce dernier ne peut être conçu sans l'intermédiaire d'un entrepôt de données qui rassemble les différents types de transactions ainsi que les informations qui y sont liées, c'est pourquoi nous avons eu recours aux bases de données décisionnelles.

Aujourd'hui, les tableaux de bord sont un outil essentiel de la stratégie et des systèmes décisionnels des entreprises. En fait, ils sont la composante la plus importante d'une gestion de la performance bien gérée. La maîtrise de la conception des tableaux de bord est un facteur clé de la réussite de la mise en œuvre d'une stratégie efficace.

Les tableaux de bord aident les décideurs à tous les niveaux de l'organisation et leur fournissent la vue d'ensemble qu'ils souhaitent pour suivre l'état et les opportunités de l'entreprise. Ils peuvent également favoriser l'interaction avec les données, par exemple en creusant dans les détails des données pour en extraire des informations utiles.

En termes de mise en œuvre, nous avons largement atteint les objectifs fixés par l'opérateur malgré les difficultés rencontrées, qui concernent principalement le niveau technique. L'une de ces difficultés, qui a été la plus difficile, concerne la compréhension de la structure des sources de données. En effet, nous avons dû faire face à cette structure complexe composée de plusieurs tables, chacune provenant de systèmes différents et contenant plusieurs champs. Comme il existe peu ou pas de documentation sur la structure de ces sources de données, il nous a fallu beaucoup de temps pour comprendre la relation entre les tables et la signification de chaque attribut.

Pour développer cette solution, nous avons d'abord détaillé les besoins du tableau de bord à travers des réunions avec les responsables du département. Ensuite, nous avons conçu le modèle de données de notre entrepôt de données afin de répondre au mieux aux besoins des différents responsables. Enfin, nous avons mis en place des tableaux de bord contenant des indicateurs permettant de suivre les résultats des transactions et le nombre de clients et de minimiser l'effort requis par les décideurs dans le processus de prise de décision.

Ce stage nous a apporté beaucoup de connaissances, tant théoriques que pratiques. Du côté théorique, nous avons amélioré nos compétences et acquis des connaissances sur la BI, ce qui nous aidera dans nos études. Sur le plan pratique, ce cours nous a permis de compléter notre formation théorique et pour l'enrichir, nous avons pu maîtriser plusieurs logiciels.

Pour finir, plusieurs perspectives peuvent être ajoutées à notre solution complète, comme la configuration d'outils et de méthodes d'exploration de données pour optimiser l'utilisation des données, ainsi que la mise à disposition d'un entrepôt de données avec des données liées à d'autres départements, comme par exemple les données du département RH.

# Webographie

[Web1] : <https://www.linkedin.com/company/tunisie-t-l-com/about/> consulté le 28/02/2022

[Web3] : <http://www-igm.univ-mlv.fr/~dr/XPOSE2005/entrepot/archi.html> consulté le 04/04/2022

[Web4] : <https://asana.com/fr/resources/key-performance-indicator-kpi> consulté le 12/04/2022

[Web5] :

<https://www.editions-eni.fr/open/mediabook.aspx?idR=4567eb81ba2a8ce93a4406eae45742de> consulté le 25/03/2022

[Web6] : [PHP — Wikipédia \(wikipedia.org\)](#) consulté le 10/06/2022

[Web7] : <https://fr.wikipedia.org/wiki/WampServer> consulté le 10/06/2022

[Web8] :

<https://www.pierre-giraud.com/html-css-apprendre-coder-cours/definition-utilisation/#:~:text=Le%20CSS%20%C3%A0langage%20de%20styles&text=Le%20CSS%20vient%20r%C3%A9aliser%20un,en%20leur%20appliquant%20des%20styles> consulté le 26/06/2022

[Web9] : <https://fr.wikipedia.org/wiki/PhpMyAdmin> consulté le 26/06/2022

[Web10] :

<https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1203585-javascript/> consulté le 26/06/2022

## **Résumé :**

Ce travail s'inscrit dans le cadre du projet de fin d'études pour l'obtention du diplôme de la licence en Business Computing parcours Business Intelligence, vise à concevoir et mettre en œuvre une solution informatique d'un système décisionnel permettant d'intégrer les données de la direction CVM datamining, de créer un entrepôt de données et d'analyser les données pour obtenir comme résultats les rapports qui seront mis en œuvre dans un tableau de bord qui vise à automatiser le traitement des données de l'entreprise.

Notre solution vise à rendre la prise de décision plus rapide et plus éclairée au sein de l'entreprise Tunisie Télécom.

## **Mots-clés:**

Business Intelligence, DataWarehouse, SSIS, SSAS, Power BI, PHP

## **Abstract:**

This work is part of the graduation project of the license in Business Computing course Business Intelligence, aims to design and implement a computer solution of a decision-making system to integrate data from the direction CVM datamining, create a data warehouse and analyze the data to obtain as results the reports that will be implemented in a dashboard that aims to automate the processing of data from the company.

Our solution aims to make decision making faster and more informed within the company Tunisie Telecom.

## **Keywords:**

Business Intelligence, DataWarehouse, SSIS, SSAS, Power BI, PHP