

N°Ordre...../F.S.S.A/UAMOB/2018

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
UNIVERSITE AKLI MOHAND OULHADJ-BOUIRA



Faculté des Sciences et des Sciences Appliquées
Département d'Informatique

Mémoire de fin d'étude

Présenté par :

Mlle. BOULIL Dalila

En vue de l'obtention du diplôme de **Master 02** en :

Filière : Informatique

Option : ISIL

Thème

**Analyse des données liées à la mobilité et à la transition
énergétique**

Devant le jury composé de :

MESSAOUDI Oussama
OUAHABI Abdeldjalil
ALOUIAT Wahiba
HAMBLI Chahrazad

UAMOB
UAMOB
UAMOB
UAMOB

Président
Encadreur
Examineur
Examineur

Année Universitaire 2018/2019

Remerciements

Je remercie DIEU tout puissant de m'avoir donné la force, la santé, le courage et la patience de pouvoir accomplir ce travail.

Un grand merci à toute ma famille surtout mes parents pour leur encouragement et leur suivi avec patience durant le déroulement de mon projet.

Mes remerciements s'étendent à mon promoteur le professeur OUAHABI Abdeldjalil pour sa patience, sa disponibilité et surtout pour ses judicieux conseils, qui ont contribué à alimenter mes réflexions.

Mes sincères remerciements s'adressent à mon cher frère BOULIL Kamal pour sa disponibilité et ces efforts et sa patience avec moi qui m'ont aidé à réaliser mon projet.

Je remercie aussi à tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.

Je tiens à remercier chacun des membres du jury monsieur BAL Kamel et mademoiselle AID Aicha pour m'avoir fait l'honneur d'examiner et d'évaluer mon travail et de l'enrichir par leurs propositions.

Dédicaces

A mes chers parents que le dieu les garde, qui m'ont donné un magnifique modèle de labeur et de persévérance.

A mes adorables frères Kamal et Mehdi et mes adorables sœurs Fadhila et Ouahiba qui m'ont souhaité un avenir prospère.

A toute ma famille « Boulil » sans exceptions.

A mes belles amies ALOUACHE Kahina et MEDJRI Kahina.

A tous mes amis et collègues.

A tous ceux qui m'ont aidé.

A tous les professeurs qui m'ont appris durant les cinq ans.

A tout le personnel du département informatique.

Je dédie cet humble travail.

Résumé

Avec le temps, la préservation de l'environnement devient une thématique plus importante qui doit être prise en considération dans tous les domaines d'activité. Parmi ces domaines le transport des personnes et des biens autrement dit la mobilité. Dans cette dernière étant proposé d'appliquer la transition énergétique qui désigne un ensemble de transformations pour aboutir à une mobilité durable et plus écologique.

Dans ce contexte, nous avons réalisé un système d'information décisionnel permettant d'évaluer et comparer plusieurs systèmes de mobilité selon un ensemble de critères et indicateurs. L'objectif de ce système étant de mettre en disposition des décideurs des rapports pour la prise des diverses décisions relatives à l'application de la transition énergétique sur les systèmes de mobilité.

Mots clés : transport, mobilité, transition énergétique, mobilité durable, système d'information décisionnel, rapports.

Abstract

Over time, the preservation of the environment becomes a more important issue that must be taken into consideration in all areas of activity. Among these areas is the transport of people and goods, in other words mobility. In the latter, it is proposed to apply the energy transition, which refers to a set of transformations to achieve sustainable and more ecological mobility.

In this context, we have developed a decision-making information system to evaluate and compare several mobility systems according to a set of criteria and indicators. The objective of this system is to provide decision-makers with reports for the various decisions relating to the application of the energy transition to mobility systems.

Key words : transport, mobility, energy transition, sustainable mobility, decision-making information system, reports.

Table de matières

Table de matières	ii
Liste des figures	v
Liste des tableaux	ix
Liste des abréviations.....	x
Introduction générale.....	1
1. Mobilité et transition énergétique.....	2
1.1. Introduction	2
1.2. Définitions.....	2
1.2.1. Mobilité	2
1.2.2. Transition énergétique.....	2
1.2.3. Mobilité durable.....	3
1.3. Variables d'analyse de mobilité.....	3
1.3.1. Indicateurs d'analyse.....	4
1.3.1.1. Indicateurs de mobilité propres à un système de mobilité	4
1.3.1.2. Indicateurs de transition liés au changement du système de mobilité	4
1.3.2. Critères d'analyse	4
1.4. Conclusion	5
2. Système d'information décisionnel	6
2.1. Introduction	6
2.2. Système d'information décisionnel	6
2.2.1. Définition	6
2.2.2. Architecture.....	7
2.3. Entrepôt de données	8
2.3.1. Définition	8
2.3.2. Modèle multidimensionnel	9
2.3.3. Modélisation multidimensionnelle.....	9
2.3.3.1. Définition des besoins.....	10
2.3.3.2. Modélisation multidimensionnelle conceptuelle.....	10
2.3.3.3. Modélisation multidimensionnelle logique.....	10

2.2.3.4. Modélisation multidimensionnelle physique.....	12
2.3.4. Approches de construction d'ED.....	12
2.3.4.1. Top-Down.....	12
2.3.4.2. Bottom-Up.....	12
2.3.4.3. Middle-Out.....	12
2.4. Méthodes d'analyse et restitution	13
2.4.1. OLAP	13
2.4.1.1. Définition.....	13
2.4.1.2. Les opérateurs OLAP	14
2.4.1.3. Comparaison entre OLAP et OLTP	17
2.4.2. Générateur des rapports ou Reporting.....	17
2.4.3. Analyse à la demande ou Ad hoc	18
2.4.4. Tableau de bord	18
2.4.5. Data Mining.....	19
2.5. Conclusion	19
3. Analyse des besoins de SID.....	20
3.1. Introduction	20
3.2. Collecte des besoins	20
3.2.1. Type d'approche	20
3.2.2. Sources de données	20
3.3. Spécification des besoins	20
3.3.1. Besoins d'affaires	20
3.3.2. Besoins fonctionnelles	20
3.3.3. Besoins techniques	21
3.4. Analyse des besoins.....	21
3.4.1. Fait.....	22
3.4.2. Indicateurs d'analyse.....	22
3.4.3. Dimensions d'analyse.....	22
3.5. Conclusion	23
4. Modélisation de SID	24
4.1. Introduction	24

4.2. Modélisation multidimensionnelle	24
4.2.1. Dimensions	24
4.2.2. Faits	29
4.2.3. Schéma en étoile	30
4.3. Conclusion	31
5. Implémentation de SID	32
5.1. Introduction	32
5.2. Architecture de SID	32
5.3. Création et alimentation de l'entrepôt de données	34
5.4. Création du cube OLAP	45
5.5. Analyse du cube OLAP	49
5.6. Exemples de Reporting	53
5.6.1. Analyse des déplacements selon les caractéristiques de l'individu	54
5.6.2. Analyse des déplacements selon le lieu de résidence	60
5.6.3. Analyse des déplacements selon le parc automobile	68
5.6.4. Analyse des flux de déplacements	70
5.7. Conclusion	72
Conclusion générale	73
Bibliographie	74

Liste des figures

<i>Figure 1-Architecture typique d'un SID (Malinowski, 2008).</i>	7
<i>Figure 2-Exemple d'un schéma en étoile.</i>	10
<i>Figure 3-Exemple d'un modèle en flocon.</i>	11
<i>Figure 4-Exemple d'un modèle en constellation.</i>	12
<i>Figure 5-Exemple du cube des ventes (Boulil, 2012).</i>	13
<i>Figure 6-Exemples des opérateurs roll-up et drill-down.</i>	14
<i>Figure 7-Exemple d'opérateur slice.</i>	15
<i>Figure 8-Exemple d'opérateur dice.</i>	16
<i>Figure 9-Exemple de l'opération pivot.</i>	16
<i>Figure 10-Exemple d'un outil de Reporting.</i>	18
<i>Figure 11-Table de dimension «Genre_individu ».</i>	24
<i>Figure 12-Table de dimension «Age_individu ».</i>	24
<i>Figure 13-Table de dimension «Nationalité_individu ».</i>	25
<i>Figure 14-Table de dimension « Etatmatlég_inividu ».</i>	25
<i>Figure 15-Table de dimension « CSP_individu ».</i>	25
<i>Figure 16-Table de dimension « Lieu_résidence ».</i>	26
<i>Figure 17-Table de dimension « Type_ménage ».</i>	26
<i>Figure 18-Table de dimension « Trancherevenu_ménage ».</i>	26
<i>Figure 19-Table de dimension « Date_départ».</i>	27
<i>Figure 20-Table de dimension « Date_arrivée».</i>	27
<i>Figure 21-Table de dimension «Lieu_départ».</i>	28
<i>Figure 22-Table de dimension «Lieu_destination».</i>	28
<i>Figure 23-Table de dimension «Motif_déplacement».</i>	28
<i>Figure 24-Table de dimension «Mode_transport».</i>	29
<i>Figure 25-Table de dimension «Type_carburant».</i>	29
<i>Figure 26-Table du fait « Déplacements ».</i>	30
<i>Figure 27-Schéma en étoile de notre SID.</i>	31
<i>Figure 28-Architecture technique de notre SID.</i>	32
<i>Figure 29-Création de base de données « Mobilité ».</i>	34
<i>Figure 30-Création du schéma « ods ».</i>	
<i>Figure 31-Création du schéma « dwh ».</i>	34
<i>Figure 32-Fichiers délimités et connexions des métadonnées.</i>	36

Figure 33-Job « TransformerGenre_individuEnSchemaods ».	36
Figure 34-Job « TransformerAge_individuEnSchemaods ».	37
Figure 35-Job « TransformerNationalite_individuEnSchemaods ».	37
Figure 36-Job « TransformerEtatmatrimoniallegalEnSchemaods ».	37
Figure 37-Job « TransformerCSP_individuEnSchemaods ».	37
Figure 38-Job « TransformerDépartementstRégions_ResidenceEnSchemaods ».	38
Figure 39-Job « TransformerType_ménageEnSchemaods ».	38
Figure 40-Job « TransformerTrancherevenu_ménageEnSchemaods ».	38
Figure 41-Job « TransformerDeptPaysDepart_DestinationEnSchemaods ».	38
Figure 42-Job « TransformerMotif_deplacementEnSchemaods ».	39
Figure 43-Job « TransformerMode_transportEnSchemaods ».	39
Figure 44-Job « TransformerType_carburantEnSchemaods ».	39
Figure 45-Job « TransformerDate_departarriveedeplocEnSchemaods ».	39
Figure 46-Job « TransformerDate_departarriveedepvoydetEnSchemaods ».	40
Figure 47-Job « TransformerDeplacements_localEnSchemaods ».	40
Figure 48-Job « TransformerDeplacements_voydetEnSchemaods ».	40
Figure 49-Tables du schéma «ods».	
	Figure 50-Tables du schéma
«dwh».	41
Figure 51-Job « TransformerGenre_individuEnSchemadwh ».	41
Figure 52-Job « TransformerAge_individuEnSchemadwh ».	41
Figure 53-Job « TransformerNationalité_individuEnSchemadwh ».	42
Figure 54-Job « TransformerEtatmatlég_individuEnSchemadwh ».	42
Figure 55-Job « TransformerCSP_individuEnSchemadwh ».	42
Figure 56-Job « TransformerLieu_destinationEnSchemadwh ».	42
Figure 57-Job « TransformerType_ménageEnSchemadwh ».	42
Figure 58-Job « TransformerTrancherevenu_ménageEnSchemadwh ».	43
Figure 59-Job « TransformerDate_departEnSchemadwh ».	43
Figure 60-Job « TransformerDate_arriveeEnSchemadwh ».	43
Figure 61-Job « TransformerLieu_departEnSchemadwh ».	43
Figure 62-Job « TransformerLieu_destinationEnSchemadwh ».	44
Figure 63-Job « TransformerMotif_deplacementEnSchemadwh ».	44
Figure 64-Job « TransformerMode_transportEnSchemadwh ».	44
Figure 65-Job « TransformerType_carburantEnSchemadwh ».	44
Figure 66-Job « TransformerDeplacementsEnSchemadwh ».	45

<i>Figure 67-Formulaire de configuration de base dans Schéma Workbench.</i>	<i>45</i>
<i>Figure 68- Fichier XML du cube « Cube de Déplacements ».</i>	<i>49</i>
<i>Figure 69-Fenêtre de la configuration de connexion de JRUBIK au serveur PostgreSQL. ...</i>	<i>50</i>
<i>Figure 70-Fenêtre de la configuration de source de données.</i>	<i>50</i>
<i>Figure 71-Cube « Cube de déplacements ».</i>	<i>51</i>
<i>Figure 72-Fenêtre « Discover » pour glisser les dimensions et les mesures.</i>	<i>52</i>
<i>Figure 73-Fenêtre « Naviguer » pour définir les niveaux des dimensions.</i>	<i>52</i>
<i>Figure 74-Exemple de résultat d'analyse sous forme d'un tableau.</i>	<i>53</i>
<i>Figure 75-Exemple de résultat d'analyse sous forme d'une requête MDX.</i>	<i>53</i>
<i>Figure 76- Nombre de déplacements par genre de l'individu et par mode de transport.</i>	<i>54</i>
<i>Figure 77-Nombre de déplacements par genre de l'individu et par motif de déplacement.</i>	<i>55</i>
<i>Figure 78- Nombre de déplacements par catégorie d'âge de l'individu et par mode de transport.</i>	<i>56</i>
<i>Figure 79-Nombre de déplacements par catégorie d'âge de l'individu et par motif de déplacement.</i>	<i>57</i>
<i>Figure 80- Nombre de déplacements par tranche du revenu mensuel net du ménage et par catégorie de mode de transport.</i>	<i>58</i>
<i>Figure 81-Nombre de déplacements par tranche du revenu mensuel net du ménage et par catégorie de motif de déplacement.</i>	<i>59</i>
<i>Figure 82-Nombre de déplacements par tranche du revenu mensuel net du ménage et par type de carburant.</i>	<i>60</i>
<i>Figure 83- Nombre des déplacements par région de résidence et par année.</i>	<i>61</i>
<i>Figure 84- Consommation totale de carburant par région de résidence et par année.</i>	<i>62</i>
<i>Figure 85- Emissions totales de CO2 par région de résidence et par année.</i>	<i>63</i>
<i>Figure 86-A1-Distance moyenne par la région de résidence pour le motif d'études.</i>	<i>64</i>
<i>Figure 87- A2-Distance moyenne par région de résidence pour les motifs d'achats.</i>	<i>65</i>
<i>Figure 88- A3-Distance moyenne par région de résidence domicile pour le motif de soins. ...</i>	<i>66</i>
<i>Figure 89- A4-Distance moyenne par région de résidence pour les motifs du travail.</i>	<i>67</i>
<i>Figure 90-Nombre des déplacements par région de résidence et par mode de transport (seulement les modes de transport en commun).</i>	<i>68</i>
<i>Figure 91- Nombre de déplacements par type de carburant et par année pour la catégorie de mode de transport (Automobile).</i>	<i>69</i>
<i>Figure 92- Emissions totales de CO2 par type de carburant et par année pour la catégorie de mode de transport (Automobile).</i>	<i>69</i>

<i>Figure 93- Nombre de déplacements par région de départ et par région de destination.....</i>	<i>70</i>
<i>Figure 94-Nombre de déplacements par région de départ et par région de destination et par mode de transport (seulement les modes de transport en commun régional ou urbain).</i>	<i>71</i>
<i>Figure 95- Nombre de déplacements par région de départ et par région de destination et par catégorie de motif de déplacement.</i>	<i>72</i>

Liste des tableaux

<i>Tableau 1 - Comparaison entre OLAP et OLTP (SGBD et Datawarehouse) (SGBD et Datawarehouse).</i>	<i>17</i>
--	-----------

Liste des abréviations

ARS Anti Récession des Soupapes

CSP Catégorie Socio-Professionnelle

DWH Data Warehouse

ED Entrepôt de Données

ENTD Enquête Nationale Transports et Déplacements

ETL Extract Transform Load

GES Gaz à Effet de Serre

GNV Gaz Naturel pour Véhicules

GPL Gaz de Pétrole Liquéfié

HOLAP Hybrid Online Analytical Processus

KPI Key Performance Indicator

MDX MultiDimensional eXpressions language

MOLAP Multidimensional Online Analytical Processus

ODS Operating Data Store

OLAP Online Analytical Processus

OLTP OnLine Transaction Processing

RER Réseau Express Régional

ROLAP Relational Online Analytical Processus

SGBDR Système de Gestion de Base de Données Relationnelle

SID Système d'Information Décisionnel

SNCF Société Nationale des Chemins de Fer

SQL Structured Query Language

TER Transport Express Régional

VAL Véhicule Automatique Léger

TVG Train à Grande Vitesse

VUL Voiture Utilitaire Légère

XML Extensible Markup Language.

Introduction générale

De nos jours, la transition énergétique est devenue un sujet très important. Elle vise à remplacer les anciens systèmes énergétiques par des systèmes durables et plus écologiques. La transition énergétique devrait être réalisée dans des différents domaines notamment le transport de biens et de personnes, autrement dit la mobilité.

Afin de mieux définir la politique de transition énergétique cible de mobilité, il est nécessaire d'évaluer et de comparer plusieurs systèmes énergétiques (ex. voiture classique vs voiture électrique) selon différents critères et indicateurs. Pour faciliter cette comparaison, une des solutions est de mettre en place un système d'information décisionnel. Ce dernier devrait permettre de collecter, intégrer et analyser des données de mobilité concernant plusieurs systèmes ; et calculer et restituer différents indicateurs de mobilité.

Dans notre travail, nous allons réaliser un système d'information décisionnel (SID) pour l'analyse des données liées à la mobilité et à la transition énergétique selon cinq chapitres :

- Dans le premier chapitre "Mobilité et transition énergétique", nous allons définir quelques concepts sur la mobilité et la transition énergétique.
- Dans le deuxième chapitre "Système d'information décisionnel", nous allons définir les concepts de base de système d'information décisionnel.
- Dans le troisième chapitre "Analyse des besoins de SID", nous allons décrire la phase d'analyse des besoins de notre SID.
- Dans le quatrième chapitre "Modélisation de SID", nous allons modéliser l'entrepôt de données ED de notre SID.
- Dans le cinquième chapitre "Implémentation de SID", nous allons implémenter l'architecture technique de notre SID en utilisant les différents technologies et logiciels.

L'objectif de notre SID est de mettre en disposition des décideurs (ex. le ministère de transport) des rapports permettant d'évaluer et de comparer les différents systèmes de mobilité selon un ensemble de critères et indicateurs pour la prise des diverses décisions relatives à l'application de la transition énergétique sur les systèmes de mobilité.

1. Mobilité et transition énergétique

1.1. Introduction

Chaque jour, des personnes se déplacent d'un endroit à autre par l'utilisation des différents modes de transport. Souvent les énergies consommées et émiées (fossiles) par ces modes ont des effets négatifs sur l'environnement, la santé de la population, etc. Pour réduire ces effets et améliorer les systèmes de mobilité, un nouveau concept apparu nommé « la transition énergétique».

Dans ce chapitre, nous allons expliquer plus détaillément sur la mobilité et la transition énergétique puis nous allons citer les principaux variables et indicateurs utilisés pour comparer et évaluer les systèmes de mobilité dans la littérature.

1.2. Définitions

1.2.1. Mobilité

Le terme de mobilité (Orfeuil, 1996), tout d'abord contient une idée de mise en mouvement. Il fait référence à une notion de déplacement. De manière très générale, le déplacement est le mouvement d'une personne, effectué pour un certain motif (activité), sur une voie publique, dans un temps voulu, entre une origine et une destination, à l'aide d'un ou de plusieurs modes de transport.

Chaque déplacement se caractérise par un motif et un seul et tout changement de motif entraîne un nouveau déplacement.

Le motif représente la raison pour lequel le déplacement est réalisé (ex. travailler, étudier et voyager, etc.).

Le mode de transport est le moyen de locomotion emprunté pour effectuer le déplacement (ex. le vélo, la voiture et le bus etc.).

1.2.2. Transition énergétique

Le concept de transition énergétique (Ortar, 2014) est apparu en 1980, en Allemagne et en Autriche. La transition énergétique désigne l'ensemble des transformations du système de production, de distribution et de consommation d'énergie effectuées sur un territoire dans le but de le rendre plus écologique. Elle vise à modifier un système énergétique pour minimiser son impact environnemental. Elle s'applique souvent dans différents secteurs d'activité (par exemple l'industrie, les transports, les bâtiments, etc.).

Un exemple de transition énergétique est de transformer un système de chauffage basé sur le gaz vers un système basé sur l'énergie éolien.

Les principaux objectifs de la transition énergétique sont (Comeen, 2017) :

- La lutte contre le réchauffement climatique provoqué par l'augmentation des émissions de gaz à effet de serre(GES)¹.
- La diminution de la consommation des énergies fossiles et l'exploitation des énergies renouvelables.
- La préservation de la santé de la population.
- La sécurisation des systèmes énergétiques.
- Etc.

1.2.3. Mobilité durable

La transition énergétique dans la mobilité permet de proposer des moyens et des solutions pour le développement d'une mobilité durable.

La mobilité durable ou écomobilité (Écomobilité, 2001) est l'un des enjeux majeurs de notre transition énergétique. Elle représente une politique d'aménagement et de gestion du territoire et de la ville qui favorise une mobilité pratique, peu polluante et respectueuse de l'environnement, ainsi que du cadre de vie (c'est à dire satisfaire la liberté des déplacements des personnes ou des biens tout en cherchant à diminuer l'impact des moyens de transport utilisés sur l'environnement).

Parmi les solutions proposées pour le développement de mobilité durable (CLER, 2018) :

- Réduire le nombre des déplacements (ex. par le développement de télétravail) et leurs distances (ex. le maintien de services et de commerces locaux).
- Encourager l'utilisation des modes de transports actifs (la marche, le vélo, etc.)
- Utiliser et développer les modes de transport collectifs (par exemple le train, le bus, etc.).
- Améliorer et organiser l'aménagement du territoire.
- Auto partage, covoiturage, etc.
- Utiliser et développer les véhicules électriques.
- Etc.

1.3. Variables d'analyse de mobilité

Dans la littérature, les données de mobilité sont analysées en utilisant différents types d'indicateurs (mesures) et de critères (Benjamin Méreau, 2019).

¹ GES (gaz à effet de serre) sont des gaz qui absorbent une partie des rayons solaires en les redistribuant sous la forme de radiations au sein de l'atmosphère terrestre, phénomène appelé effet de serre (Cogiterra, 2013).

1.3.1. Indicateurs d'analyse

Dans la littérature, on distingue généralement deux catégories d'indicateurs liés au thème de mobilité et transition énergétique : les indicateurs de mobilité propres à un système de mobilité et les indicateurs de transition liés au changement du système de mobilité.

1.3.1.1. Indicateurs de mobilité propres à un système de mobilité

Pour cette catégorie, la littérature distingue plusieurs sous catégories :

- Les indicateurs de durée, par exemple la durée moyenne en minutes des déplacements professionnels domicile-travail
- Les indicateurs de distance, par exemple la distance moyenne en km des déplacements scolaires
- Les indicateurs de coût, par exemple le coût annuel en des déplacements par ménage
- Les indicateurs de consommation de carburant en litre.
- Les indicateurs de comptage/fréquence, par exemple, le nombre de déplacements couteux en temps.
- Les indicateurs d'impact sur l'environnement, par exemple les émissions de gaz à effet de serre, l'empreinte carbone, etc.
- Les indicateurs complexes calculés à partir des indicateurs précédents, différents taux (ex. taux d'effort énergétique).
- Enfin, ces indicateurs sont étudiés et calculés à différents niveaux de détail : ménage, commune, pays, etc.

1.3.1.2. Indicateurs de transition liés au changement du système de mobilité

Cette catégorie regroupe les indicateurs/mesures induits par le passage d'un système de mobilité vers un autre. Par exemple les indicateurs de cout et durée d'investissement, etc.

Dans notre stage, nous nous sommes focalisés sur les indicateurs de la première catégorie, à savoir les indicateurs de mobilité qui servent à mesurer (évaluer) et comparer les différents systèmes de mobilité.

1.3.2. Critères d'analyse

Dans la littérature, les données de mobilité sont analysées selon différents critères. Les critères les plus utilisées sont :

- ✓ Les caractéristiques du ménage : taille du ménage, type du ménage, csp de la personne de référence du ménage, lieu de résidence, type de logement, etc.
- ✓ Les caractéristiques d'individu : sexe d'individu, catégorie d'âge d'individu, niveau d'études d'individu, etc.

- ✓ Le mode de transport
- ✓ Le motif de transport
- ✓ La date d'évaluation
- ✓ L'heure de départ
- ✓ L'heure d'arrivée
- ✓ Le lieu de départ
- ✓ Le lieu d'arrivée
- ✓ Etc.

1.4. Conclusion

Dans ce chapitre, nous avons défini les concepts liés à la mobilité et à la transition énergétique : la mobilité, la transition énergétique et la mobilité durable. Puis nous avons cité quelques indicateurs et variables suivant les quels les données de mobilité et transition énergétique sont analysées dans la littérature.

Dans le chapitre suivant, nous allons définir les concepts de base de la solution informatique que nous avons choisie pour l'analyse des données liées à la mobilité et à la transition énergétique nommée « le système d'information décisionnel ».

2. Système d'information décisionnel

2.1. Introduction

L'informatique décisionnelle désigne les moyens, les outils et les méthodes qui permettent de collecter, consolider, analyser et restituer les données d'une entreprise, pour permettre aux décideurs d'avoir une vue d'ensemble et une meilleure compréhension des activités de l'entreprise afin de les aider dans la prise de décision.

L'architecture typique d'un Système d'Information Décisionnel (SID) est constituée de quatre couches :

- La couche ETL (Extract, Transform and Load tools), outil qui permet l'extraction, transformation et le chargement des données dans l'entrepôt de données.
- La couche Entrepôt de Données(ED), base de données pour intégrer, stocker et historiser les données provenant de plusieurs sources hétérogènes.
- La couche d'analyse des données basée généralement sur un outil OLAP (Online Analytical Processus).
- La couche de restitution des données, interfaces utilisateur.

Pour la réalisation d'un système d'information décisionnel, on suit généralement un processus de trois étapes :

- L'analyse des besoins des utilisateurs.
- La conception (modélisation) du système d'information décisionnel.
- L'implémentation du système d'information décisionnel.

Dans ce chapitre, nous avons comme but de donner une idée générale sur le système d'information décisionnel et les différents concepts qui lui sont reliés.

2.2. Système d'information décisionnel

Dans cette section, nous allons définir la notion du système d'information décisionnel (SID) et son architecture.

2.2.1. Définition

Les systèmes d'information décisionnels (Smile, 2013) sont apparus dans le début des années 90. Ces systèmes sont des solutions informatiques permettant l'analyse et le requêtage des données de l'entreprise afin d'en dégager des informations nouvelles généralement synthétisées qui vont fonder les prises de décision. Ils permettent de restituer une vision large et unifiée de l'information de l'entreprise.

Ces systèmes offrent (Smile, 2013). :

- Des outils pour la collecte et l'intégration de gros volumes des données issues de plusieurs sources hétérogènes (outils ETL).
- Un lieu de stockage, d'intégration et d'historisation centralisé des données (entrepôt de données).
- Des outils d'analyse et de restitution des données (OLAP, Reporting, Data Mining...).

2.2.2. Architecture

L'architecture typique (Figure 2.1) de SID est constituée de quatre couches logicielles : ETL, stockage, analyse et restitution (Malinowski, 2008). En amont, on trouve les sources de données qui peuvent être internes ou externes à l'entreprise.

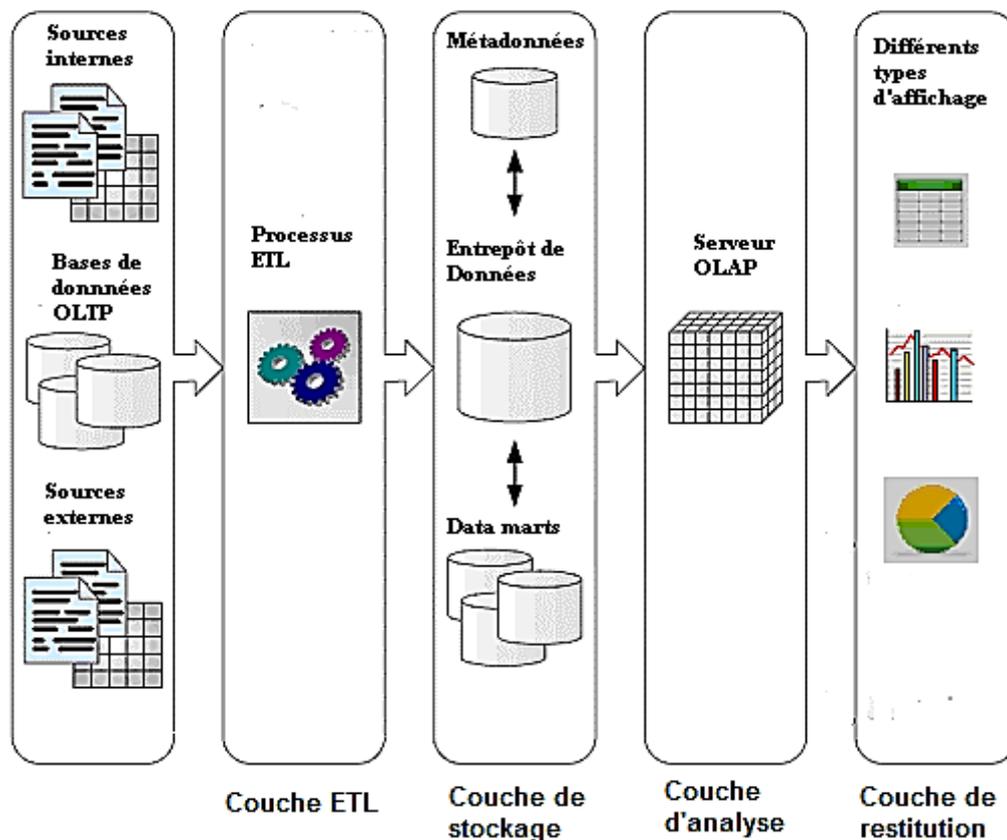


Figure 1-Architecture typique d'un SID (Malinowski, 2008).

La couche ETL

Cette couche permet l'extraction, la transformation et le chargement périodique de toutes les données source utiles à la prise de décision dans l'ED.

Selon (Jarke, 2003), cette couche comprend généralement quatre phases :

- Extraction : sélection et extraction régulière (par rapport à l'occurrence d'un événement temporel (ex. fin du mois) ou intemporel (ex. changement de données sources) des données sources utiles.
 - Nettoyage : le but de cette phase est d'améliorer la qualité des données qui peut être mauvaise dans les sources en supprimant/rectifiant certaines erreurs et inconsistances.
 - Transformation : permet la conversion des données nettoyées vers le schéma de l'ED.
- Chargement : permet le chargement des données intégrées dans l'ED.

La couche de stockage

Cette couche comprend (a) l'ED, (b) le référentiel de métadonnées et éventuellement (c) un ou plusieurs magasins de données (Data marts). L'ED est un lieu d'intégration et d'historisation pour toutes les données de l'entreprise et concernant tous les sujets d'analyse. Les data marts sont des petits EDs focalisés sur quelques sujets d'analyse. Le référentiel de métadonnées stocke toutes les métadonnées (données sur données) nécessaires à l'administration du système (par exemple, données sur les processus ETL, schémas de données, etc.).

La couche d'analyse

Elle permet le calcul de différents indicateurs d'analyse à partir des données de l'ED ou des data marts.

La couche de restitution

Cette couche permet de visualiser les indicateurs selon des axes d'analyses. Elle définit plusieurs interfaces utilisateur interactives et intuitives qui permettent d'explorer les données de l'ED en utilisant des différents types de visualisations : tableaux croisés dynamiques, histogrammes, diagrammes circulaires etc.

2.3. Entrepôt de données

Dans cette section, nous allons définir la notion d'entrepôt de données (ED), le modèle multidimensionnel, les étapes et les approches d'ED.

2.3.1. Définition

Selon (Inmon, 1996), l'entrepôt de données est une collection de données orientées sujet, intégrées, non volatiles et historisées, organisées pour le support d'un processus d'aide à la décision.

Orientées sujet : les données sont organisées par sujet d'analyse.

Intégrées : les données d'ED sont le résultat de l'intégration de diverses sources hétérogènes. Autrement dit toutes les données sources sont transformées vers le schéma de l'ED.

Historiques : les données sont datées (des références temporelles sont associées aux données pour suivre leur évolution dans le temps).

Non-volatiles : les données insérées dans l'ED ne sont jamais modifiées.

2.3.2. Modèle multidimensionnel

Le modèle multidimensionnel (Kimball, 2002) est l'une des techniques les plus utilisées pour la modélisation des EDs. Il consiste à représenter les données d'un sujet d'analyse dans un espace à plusieurs dimensions (axes). Les points de cet espace correspondent aux mesures d'analyse et les axes correspondent aux dimensions ou critères d'analyse.

Ce modèle repose sur les deux concepts : fait et dimension.

Les faits

Les **faits** représentent les sujets d'analyse (ex. les ventes, les achats). Les faits sont décrits par des **mesures**. Les mesures sont des attributs généralement numériques (ex. le fait ventes est décrit par les mesures quantité des produits vendus, montant des ventes, etc.).

Les dimensions

Les **dimensions** représentent les perspectives (critères) d'analyse par rapport auxquelles on observe le sujet analysé. Par exemple, le fait ventes peut être analysé suivant les dimensions temps, client, produit etc.

Les hiérarchies de dimension

Une **hiérarchie de dimension** organise les attributs d'une dimension hiérarchiquement selon plusieurs niveaux de détail qu'on appelle les **niveaux de dimension**. Par exemple, pour la dimension client, on peut définir la hiérarchie suivante : Client → Ville → Département → Région. Cette hiérarchie regroupe les clients par ville puis par les villes par département enfin les départements par région ; elle définit les niveaux de dimension client, ville, département et région.

La granularité des faits

La **granularité d'un fait** est le niveau de détail le plus fin par rapport auquel on analyse le fait ; elle est définie par la combinaison des niveaux les plus fins des dimensions. Par exemple, la granularité du fait vente est définie par la combinaison des niveaux mois, client et produit.

2.3.3. Modélisation multidimensionnelle

La modélisation multidimensionnelle (Boulil, 2012) définit les étapes pour la construction de l'entrepôt de données. Elle comprend quatre étapes : définition des besoins, modélisation multidimensionnelle conceptuelle, modélisation multidimensionnelle logique et modélisation multidimensionnelle physique.

2.3.3.1. Définition des besoins

L'objectif de cette étape est d'identifier, collecter, organiser et décrire les besoins d'analyse des utilisateurs dans un document de spécification des besoins.

2.3.3.2. Modélisation multidimensionnelle conceptuelle

L'objectif de cette étape est de créer le modèle multidimensionnelle conceptuel à partir de document de spécification des besoins et en utilisant un langage conceptuel pour représenter le modèle (exemple, un diagramme de classes).

2.3.3.3. Modélisation multidimensionnelle logique

L'objectif de cette étape est de traduire le modèle conceptuel en un modèle logique, en suivant un ensemble de règles de transformation.

A cette étape, le schéma en étoile (et ses variantes) (Kimball, 2002) est le plus utilisé. Dans la suite de cette section, nous allons définir ce schéma et ses variantes.

Le schéma en étoile

Ce schéma est composé d'une table du fait centrale qui référence plusieurs tables de dimension avec des clés étrangères. La clé primaire de table du fait est la combinaison des clés étrangères vers les tables de dimension. Les dimensions sont dénormalisées : tous les attributs d'une dimension sont représentés dans une seule table. Dans l'exemple suivant d'analyse des ventes (Figure 2), la table du fait Ventes référence les tables de dimension Temps, Produit et Magasin. Chaque dimension est représentée par une seule table.

Ce schéma est performant en termes de temps de réponse aux requêtes car le nombre de jointures est limité (que des jointures entre la table du fait et les tables de dimension). Par contre, il nécessite plus d'espace de stockage à cause de la redondance dans les tables de dimension (les tables de dimension sont dénormalisées).

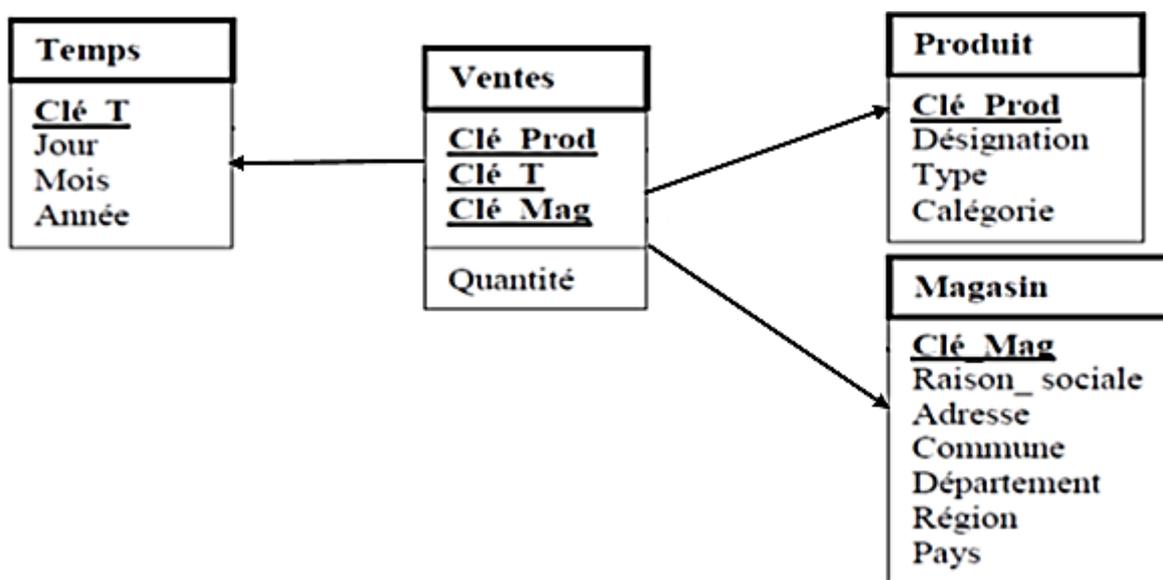


Figure 2-Exemple d'un schéma en étoile.

Le schéma en flocons de neige

Ce schéma est similaire au schéma en étoile sauf que les dimensions sont normalisées (une table de dimension par niveau de dimension). Dans l'exemple de la figure 3, la dimension Magasin est représentée par trois tables : Magasin, Département et Région.

Ce schéma est performant en termes d'espace de stockage car les dimensions sont normalisées. Par contre, il est moins performant en termes de temps de réponse aux requêtes à cause des jointures supplémentaires entre les niveaux de dimension.

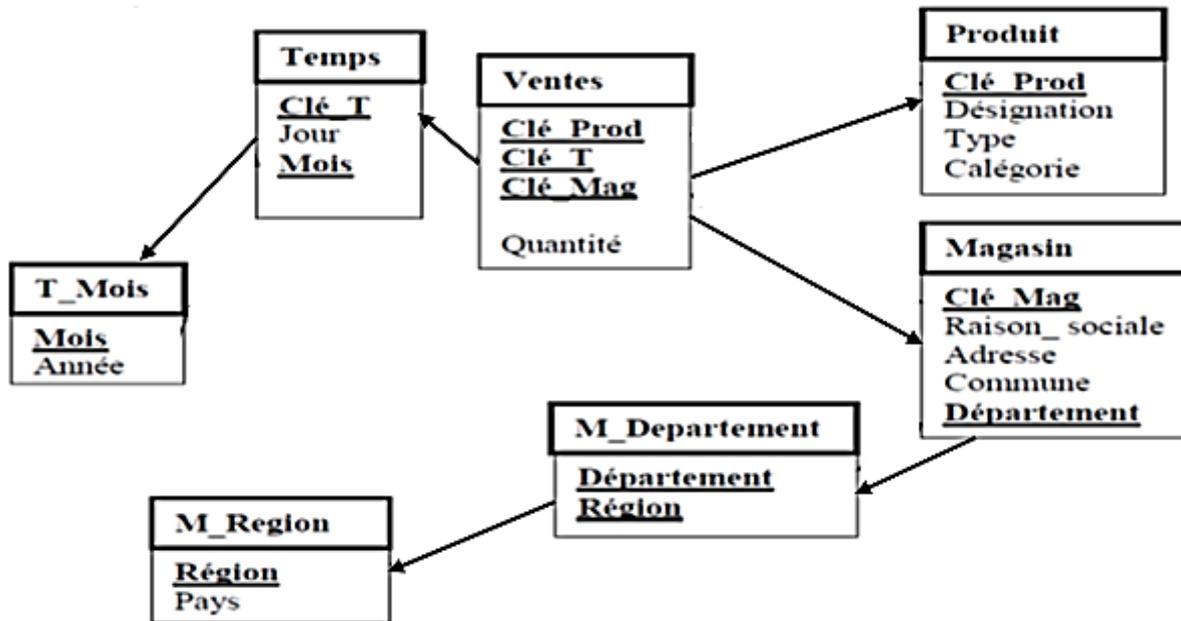


Figure 3-Exemple d'un modèle en flocon.

Le schéma en constellation

Ce schéma est composé de plusieurs schémas en étoile/flocons qui partagent des dimensions. Dans la figure suivante, les schémas Ventes et Achat partagent les dimensions Temps et Produit.

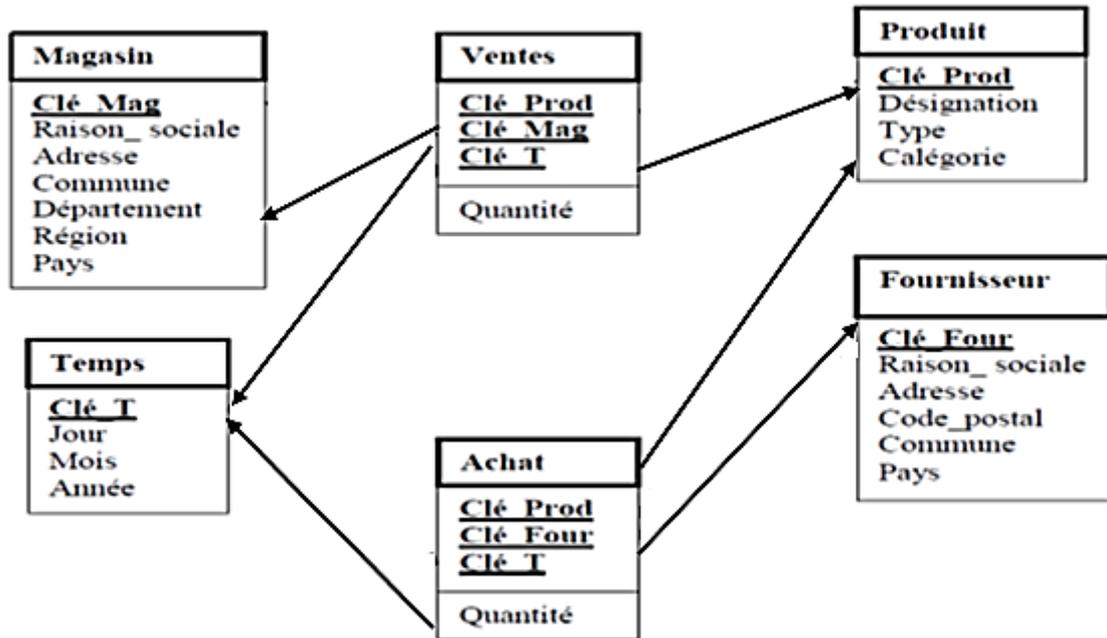


Figure 4-Exemple d'un modèle en constellation.

2.2.3.4. Modélisation multidimensionnelle physique

L'objectif de cette étape est de traduire le modèle multidimensionnelle logique en modèle multidimensionnelle physique dans le langage du SGBD cible (Oracle, PostgreSQL, etc.).

2.3.4. Approches de construction d'ED

Pour la construction d'un ED, il existe trois approches (Grim, 2019) :

2.3.4.1. Top-Down

Cette approche est proposée par William H. Inmon, elle consiste la conception de tout l'entrepôt. Elle est la plus complète car elle offre une vision très claire et très conceptuelle des données de l'entreprise. Par contre elle est la plus lourde, contraignante, nécessite plus de temps et une connaissance globale de l'entreprise.

2.3.4.2. Bottom-Up

C'est l'approche inverse de top down est proposée par Ralph Kimball, elle consiste à créer les étoiles (data marts) une par une, puis les regrouper par des niveaux intermédiaires jusqu'à obtention d'un véritable entrepôt avec une vision d'entreprise. Elle est simple à réaliser, offrir des résultats rapides. Par contre le volume de travail d'intégration est gros et redondance possible entre les étoiles.

2.3.4.3. Middle-Out

C'est approche hybride est conseillée par les professionnels du BI, elle consiste la conception totale de l'entrepôt de données, puis créer des divisions plus petites et plus gérables et les mettre en œuvre. Elle est la meilleure que les deux approches précédentes. Par contre elle implique,

parfois, des compromis de découpage (dupliquer des dimensions identiques pour des besoins pratiques).

2.4. Méthodes d'analyse et restitution

Dans la suite de ce chapitre, nous allons présenter brièvement les principales techniques d'analyse et restitution de données dans les systèmes décisionnels (Smile, 2013).

2.4.1. OLAP

2.4.1.1. Définition

Les **systèmes OLAP** (Online Analytical Processing) sont des systèmes d'analyse de données en ligne manipulant des cubes. Un **cube OLAP** est représentation multidimensionnelle des données ; les cellules du cube stockent les valeurs des mesures (agrégés ou non) ; les axes (coordonnées) du cube représentent les valeurs des dimensions. Par exemple, pour l'année "2004", on a vendu 7 (quantité) produits "Alca23" dans le magasin "Standa" pour un montant de 45.

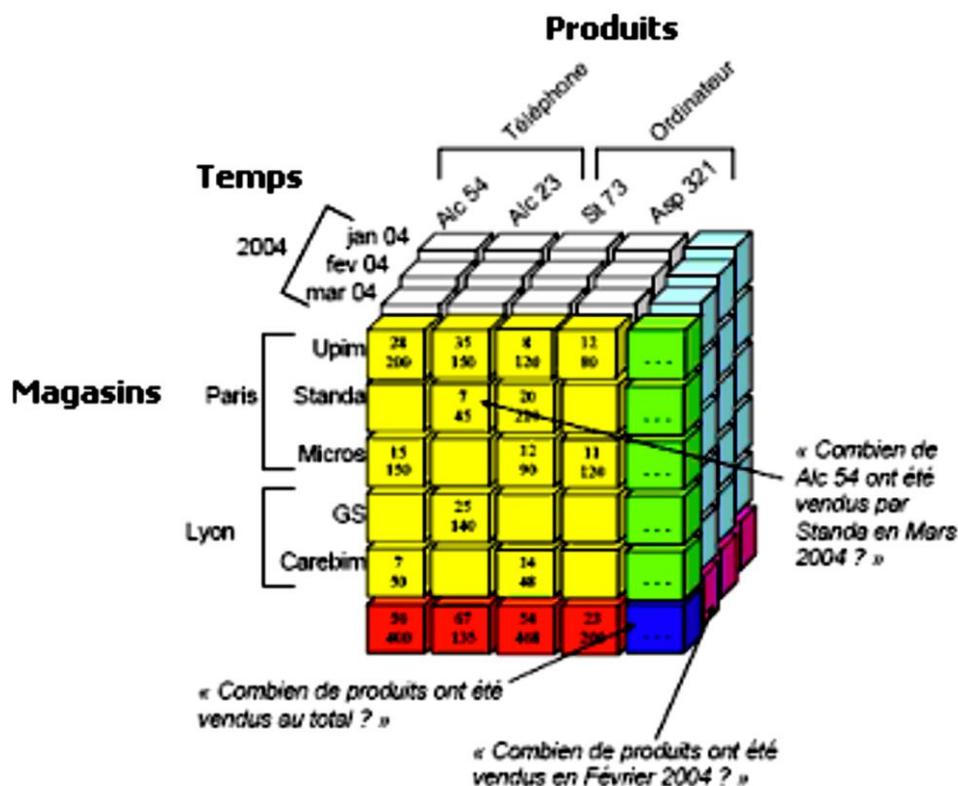


Figure 5-Exemple du cube des ventes (Boullil, 2012).

Il existe trois types d'implémentation OLAP : ROLAP (Relational OLAP) basé sur une base de données relationnelle, MOLAP (Multidimensional OLAP) basé sur une base de données multidimensionnelle et HOLAP (Hybrid OLAP) qui combine les deux.

Pour l'interrogation des cubes OLAP, on utilise souvent le langage MDX (MultiDimensional eXpressions language) (Rouse, 2019).

MDX (ESPINASSE, Entrepôts de données : Introduction au langage MDX (Multi-Dimensional eXtensions) pour l'OLAP, 2015) est un langage de requêtes OLAP pour les bases de données multidimensionnelles, inventé en 1997 par Mosha Pasumansky au sein de Microsoft. Il est fait pour naviguer dans les bases multidimensionnelles et pour exécuter des requêtes sur tous leurs objets (dimensions, hiérarchies, niveaux, membres et cellules) afin d'obtenir (simplement) une représentation sous forme de tableaux croisés.

2.4.1.2. Les opérateurs OLAP

Les systèmes OLAP implémentent un ensemble d'opérateurs de manipulation des cubes qui permettant d'exécuter rapidement les requêtes décisionnelles.

Les opérateurs OLAP sont classés en trois catégories (a) les opérateurs de forage liés à la granularité de dimension (roll-up et drill-down), (b) les opérateurs de sélection (slice et dice) et (c) les opérateurs de présentation (pivot) (Boulil, 2012):

➤ Roll-up

Il permet d'agréger les valeurs de mesure en montant dans une hiérarchie de dimension.

➤ Drill-down

C'est l'opérateur inverse à Roll-up, il permet de détailler les données en descendant le niveau hiérarchique d'une dimension.

Dans l'exemple suivant (Figure 6), l'opérateur roll-up agrège les données en montant dans la dimension Géographie ; il regroupe les départements par région. Par contre l'opérateur drill-down réalise l'inverse.

Quantité des ventes		Géographie.Département					
		Loiret	Eure et Loir	Indre	Indre et Loire	Loir et Cher	Gironde
Temps.Année	2005	100	200	100	150	1	2
	2006	250	210	120	200	2	2
	2007	260	230	250	200	4	5
	2008	280	220	220	200	5	4

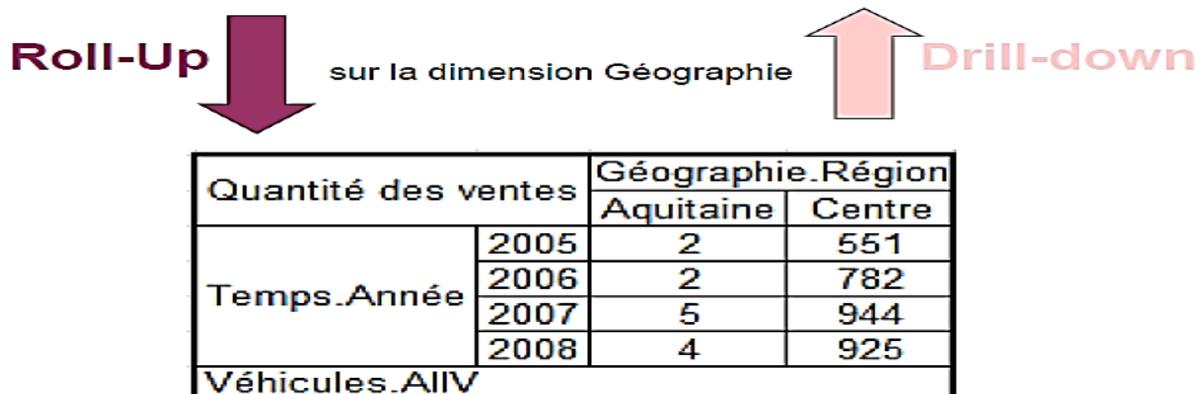


Figure 6-Exemples des opérateurs roll-up et drill-down.

➤ Slice

Il effectue une sélection des cellules du cube en utilisant une condition (prédicat) définie sur les valeurs d'une dimension, elle retourne un sous cube du cube initial.

Dans l'exemple suivant (Figure 7), l'opérateur slice permet de sélectionner un sous cube qui contient les valeurs où l'année = 2005 (une seule condition sur la dimension Temps).

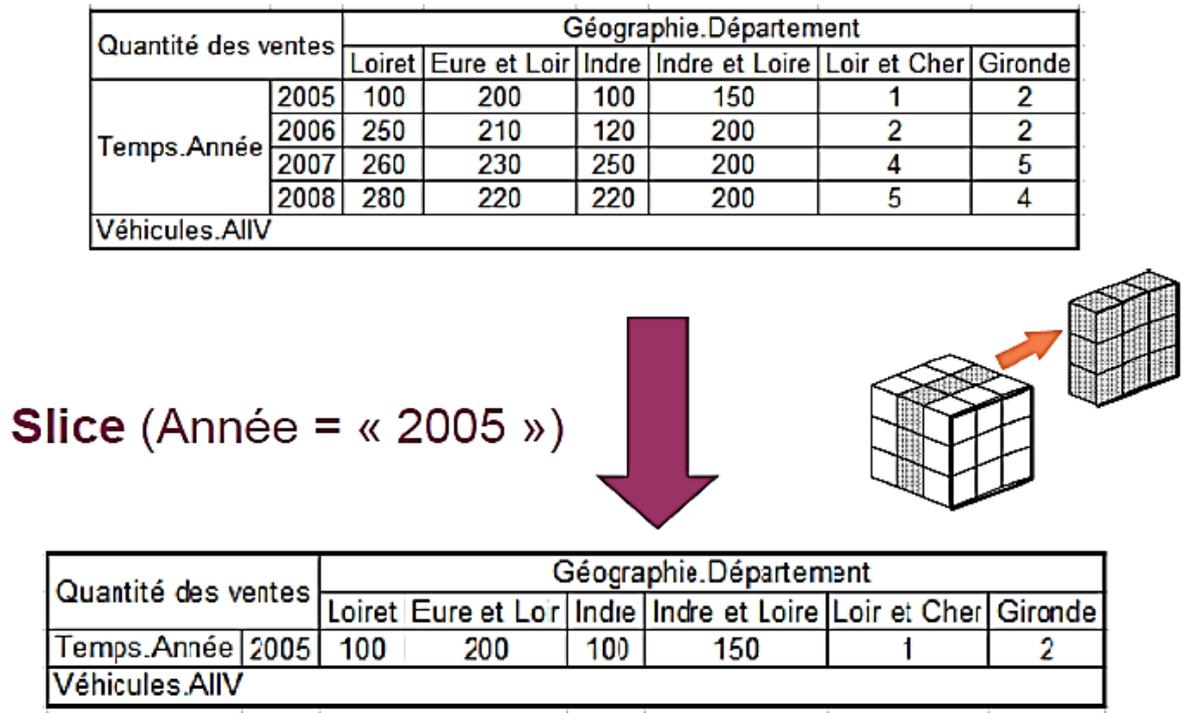


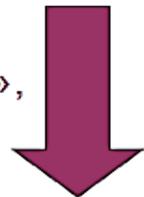
Figure 7-Exemple d'opérateur slice.

➤ **Dice**

Il applique des conditions de sélection sur plusieurs dimensions. Dans l'exemple suivant (Figure 9), l'opérateur dice permet de sélectionner un sous cube avec les conditions sur les dimensions Géographie et Temps : département = 'Loir et Cher' ou 'Gironde' Et année = 2005 ou 2008.

Quantité des ventes		Géographie.Département					
		Loiret	Eure et Loir	Indre	Indre et Loire	Loir et Cher	Gironde
Temps.Année	2005	100	200	100	150	1	2
	2006	250	210	120	200	2	2
	2007	260	230	250	200	4	5
	2008	280	220	220	200	5	4
Véhicules.AIIV							

Dice (Département = « Loir et Cher » ou « Gironde »,
Année = « 2007 » ou « 2008 »)



Quantité des ventes		Géographie.Département	
		Loir et Cher	Gironde
Temps.Année	2007	4	5
	2008	5	4
Véhicules.AIIV			

Figure 8-Exemple d'opérateur dice.

➤ Pivot

Il effectue un pivotage des axes du cube pour obtenir des d'autres alternatives de présentation des données. Dans l'exemple suivant (Figure 10), l'opérateur pivot permet de tourner le cube pour obtenir un autre cube.

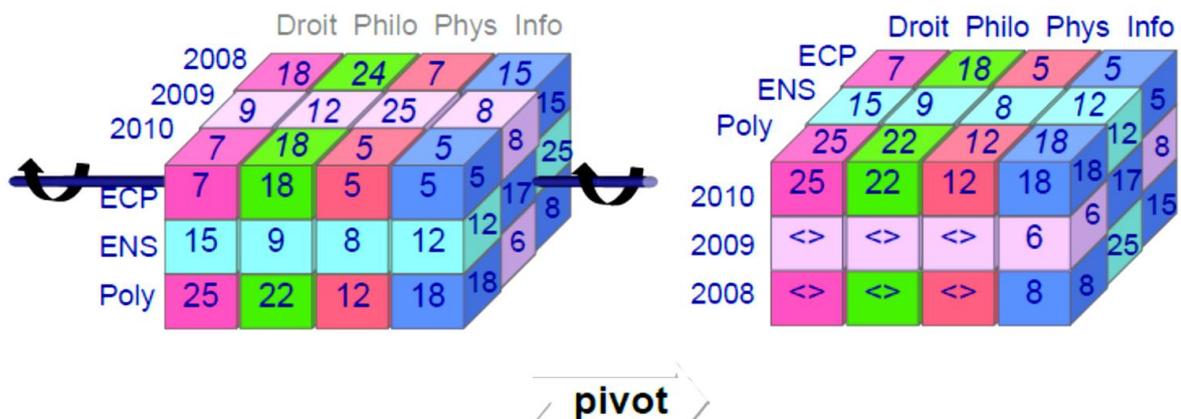


Figure 9-Exemple de l'opération pivot.

➤ Projection

Il permet de sélectionner un sous-ensemble de mesures du cube.

2.4.1.3. Comparaison entre OLAP et OLTP

Les principales différences entre les systèmes OLAP² et OLTP³ (On Line Transaction Processing) sont résumées dans le tableau suivant (Tableau 1) (Ferragu, 2013).

	OLTP	OLAP
Objectif	Support de réalisation des activités quotidiennes de l'entreprise.	Analyse des données de l'entreprise pour améliorer les performances des activités de l'entreprise.
Type de traitement	Transactionnel (requêtes simples, accès à quelques enregistrements)	Analytique (requêtes complexes, accès et agrégation de million d'enregistrements)
Données	Détaillées, courantes, dynamiques, orientées applications, non volumineuses et normalisées (3FN)	Agrégées, historisées, statiques, orientées sujet, volumineuses et dénormalisés
Opérations	Insertion, mise à jour, suppression et lecture	Insertion et lecture
Utilisateurs	Nombreux (milliers), agents opérationnels	Peu nombreux (centaines), décideurs et analystes

Tableau 2 - Comparaison entre OLAP et OLTP (SGBD et Datawarehouse) (SGBD et Datawarehouse).

2.4.2. Générateur des rapports ou Reporting

Un rapport est une présentation synthétique et lisible de données, généralement à des fins d'impression ou d'utilisation interactive. Il peut prendre plusieurs formes (tableaux, graphiques etc.).

Un générateur de rapports est un programme qui permet de définir un rapport selon ses besoins, avec un minimum de programmation, uniquement au moyen d'une interface interactive.

Un rapport peut être paramétrable. Les paramètres peuvent être définis par l'utilisateur ou bien définis automatiquement par des règles de gestion.

L'exemple suivant (Figure 11) représente un outil Excel qui permet de créer des différents rapports.

² OLAP (OnLine Analytical Processing) est un système de récupération et d'analyse de données en lignes (Différence entre OLTP et OLAP, 2018).

³ OLTP (OnLine Transaction Processing) est un système transactionnel en ligne qui sert à effectuer des modifications dans une base de données (Différence entre OLTP et OLAP, 2018).

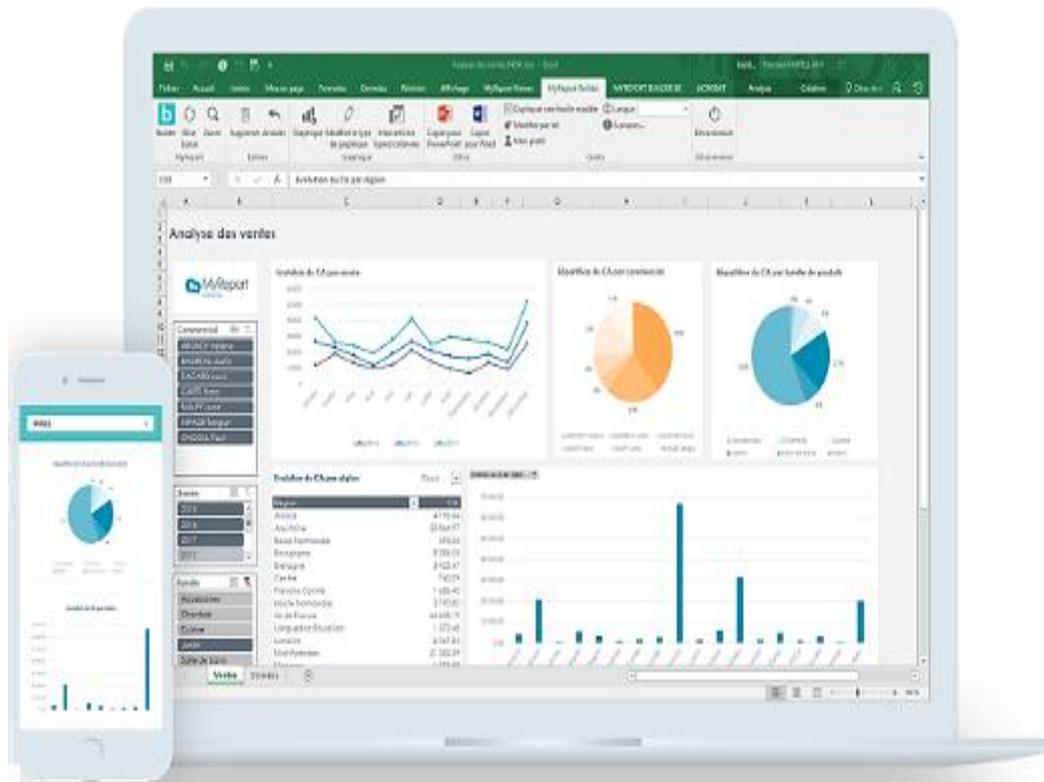


Figure 10-Exemple d'un outil de Reporting.

2.4.3. Analyse à la demande ou Ad hoc

Les outils de Reporting Ad hoc permettent aux utilisateurs finaux de réaliser rapidement et simplement des rapports simples sans connaissance technique (sans connaissance du langage SQL⁴ notamment). Ils requièrent d'excellents temps de réponse, afin de permettre à l'utilisateur d'affiner petit à petit sa requête à partir des résultats obtenus.

2.4.4. Tableau de bord

Un tableau de bord est une forme particulière de rapport très synthétique (tout doit tenir sur une feuille A4 ou sur un écran d'ordinateur). Il est le plus souvent composite, c'est à dire présentant plusieurs indicateurs, qui ensemble offrent une représentation complète de l'activité de l'entreprise. Il est souvent personnalisé, c'est à dire que chaque acteur consulte un tableau de bord qui concerne son activité.

Un tableau de bord présente souvent les indicateurs de performance clé (KPI⁵) d'une entreprise au niveau de la page d'accueil du portail décisionnel de l'entreprise.

⁴ SQL (Structured Query Language) est le langage commun à tous les SGBDR, ce qui permet de concevoir des BD relativement indépendamment des systèmes utilisés (CROZAT, 2016).

⁵ KPI (Key Performance Indicator) sont des mesures qui permettent aux entreprises d'évaluer la performance d'un domaine (Smile, 2013).

2.4.5. Data Mining

Le Data Mining est une technique qui permet de fouiller un ensemble de données afin de découvrir des informations statistiques utiles cachées.

Les outils de Data Mining recherchent de manière semi-automatisée, des corrélations, des règles statistiques, au travers de grands volumes de données, en s'aidant parfois d'une restitution graphique de ces règles, qui fera apparaître par exemple des nuages de points sur une représentation axiale des données.

2.5. Conclusion

Au cours de ce chapitre, nous avons défini les concepts de base de SID et nous allons parler brièvement sur le rôle et le fonctionnement des composants de chaque couche de l'architecture technique de SID.

Pour la réalisation de notre SID, nous allons suivre trois étapes principales : (a) l'analyse des besoins, (b) la conception (modélisation) de SID, (c) l'implémentation de SID.

Dans le chapitre suivant, nous allons décrire en détail l'étape d'analyse des besoins de notre SID dans le cadre de notre sujet « Analyse des données liées à la mobilité et à la transition énergétique ».

3. Analyse des besoins de SID

3.1. Introduction

Dans cette phase, nous allons décrire les différents besoins d'analyse de notre système d'information décisionnel (SID) puis nous allons citer les faits, les dimensions et les indicateurs sur lesquels est basée notre analyse.

3.2. Collecte des besoins

Dans cette section, nous allons décrire les éléments suivants :

3.2.1. Type d'approche

Nous avons suivi l'approche ascendante ou 'orientés données' qui se basé sur les référentiels des sources de données pour la collecte des besoins.

3.2.2. Sources de données

Nous avons extraire nos données (sous forme des fichiers.csv) d'après le site Données et études statiques depuis l'enquête nationale transports et déplacements(ENTD) 2008. Cette enquête décrit les déplacements des personnes, elle se déroule sur une année, d'avril 2007 à avril 2008 dans le champ géographique ' France métropolitaine' (qui désigne un ensemble des parties européennes de la France).

3.3. Spécification des besoins

Après avoir les sources de données liées à la mobilité et transition énergétique, nous allons spécifier les besoins suivants :

3.3.1. Besoins d'affaires

Représentent l'objectif de notre SID qui désigne l'application de la transition énergétique sur des systèmes de mobilité pour aboutir à des systèmes **durables** et **plus écologiques**.

3.3.2. Besoins fonctionnelles

Consistent à avoir des **informations décisionnelles** sur les systèmes de mobilité sous forme des **rappports**, afin d'appliquer sur lesquels la transition énergétique par exemple :

- Nombre de déplacements par caractéristiques de l'individu (genre, catégorie d'âge, etc.) et par mode de transport.
- Nombre de déplacements par caractéristiques de l'individu (genre, catégorie d'âge, etc.) et par motif de déplacement.

- Nombre de déplacements par caractéristiques du ménage (type, tranche du revenu, etc.) et par mode de transport.
- Nombre de déplacements par caractéristiques du ménage (type, tranche du revenu, etc.) et par motif de déplacement.
- Nombre de déplacements par lieu de résidence (région, département) et par année.
- Emissions totales par lieu de résidence (région, département) et par année.
- Nombre de déplacements par lieu de résidence (région, département) et par mode de transport.
- Distance moyenne par lieu de résidence (région, département) et par motif de déplacement (études, soins, travail, etc.).
- Consommation totale par type de carburant et par année pour les automobiles.
- Emissions totales par type de carburant et par année pour les automobiles.
- Nombre de déplacements par lieu de départ et par lieu de destination.
- Nombre de déplacements par lieu de départ et par lieu de destination et par mode de transport.
- Nombre de déplacements par lieu de départ et par lieu de destination et par motif de déplacement.
- Etc.

3.3.3. Besoins techniques

Décrivent les **outils** nécessaires pour l'**implémentation** de notre solution SID :

- Un outil pour l'extraction, transformation et le chargement des données.
- Un SGBDR⁶ pour la manipulation et le stockage des données.
- Un outil d'analyse multidimensionnelle des données.
- Un outil pour la génération/création des rapports.

3.4. Analyse des besoins

D'après les besoins précédents, nous allons déduire les éléments suivants.

⁶ SGBDR (système de gestion de base de données relationnelle) est un logiciel qui prend en charge la structuration, le stockage, la mise à jour et la maintenance d'une base de données. Il est l'unique interface entre les informaticiens et les données (définition des schémas, programmation des applications), ainsi qu'entre les utilisateurs et les données (consultation et mise à jour) (CROZAT, 2016).

3.4.1. Fait

Un fait représente un sujet d'analyse, il est décrit par un ensemble de mesures généralement numériques. Dans notre SID, le fait observable est « **Déplacements** » qui se décrivent par les mesures suivants :

- La distance en km.
- La durée en minutes.
- La consommation de carburant.
- Les émissions de CO2.
- Le nombre des personnes (égal au nombre des personnes accompagnés avec la personne qui déplace+1).

3.4.2. Indicateurs d'analyse

Un indicateur d'analyse est le résultat d'agrégations de mesures d'un fait. Dans notre cas, nous allons baser seulement sur les indicateurs de mobilité qui servent à évaluer et comparer les différents systèmes de mobilité.

- La distance moyenne en km : C'est le résultat de la fonction d'agrégation (avg) sur la mesure (distance).
- La consommation totale de carburant : C'est le résultat de la fonction d'agrégation (sum) sur la mesure (consommation de carburant).
- Les émissions totales de CO2 : C'est le résultat de la fonction d'agrégation (sum) sur la mesure (émissions de CO2).
- Le nombre des déplacements C'est le résultat de la fonction d'agrégation (distinct count) sur l'attribut (identifiant de déplacement) du fait (Déplacements).

3.4.3. Dimensions d'analyse

Les dimensions d'analyse représentent les critères/axes d'analyse d'un fait. Le fait « **Déplacements** » est analysé par rapport aux dimensions suivants :

- Le genre de l'individu.
- L'âge l'individu.
- La nationalité l'individu.
- L'état matrimonial légal l'individu.
- La catégorie socioprofessionnelle l'individu.
- Le lieu de résidence l'individu.
- Le type du ménage.
- La tranche du revenu mensuel net du ménage.
- La date de départ.

- La date d'arrivée.
- Lieu de départ.
- Lieu de destination.
- Le motif de transport.
- Le mode de transport.
- Le type de carburant.

3.5. Conclusion

Au cours de ce chapitre, nous avons décrit comment nous avons collecté, spécifié et analysé les besoins de notre SID, puis nous avons cité les faits, les dimensions et les indicateurs sur lesquels est basée notre analyse.

Dans le chapitre suivant, nous allons décrire en détail l'étape de modélisation (conception) de l'entrepôt de données (ED) de notre SID dans le cadre de notre sujet « Analyse des données liées à la mobilité et à la transition énergétique ».

4. Modélisation de SID

4.1. Introduction

Dans cette étape, nous allons concevoir le modèle multidimensionnel de l'entrepôt de données(ED) de notre système d'information décisionnel (SID).

4.2. Modélisation multidimensionnelle

Dans ce niveau, nous allons créer le modèle multidimensionnelle qui consiste à représenter les données d'un sujet d'analyse dans un espace à plusieurs dimensions (axes). Il se basée sur deux concepts : fait et dimension.

4.2.1. Dimensions

Les **dimensions** représentent les perspectives (critères) d'analyse par rapport auxquelles on observe le sujet analysé. Dans notre SID nous allons proposer 15 dimensions.

La table de dimension « Genre d'individu »

Elle contient la clé primaire `code_genre` et l'attribut `libellé_genre`.

La figure suivante montre la table de dimension « Genre_individu ».

Genre_individu			
	code_genre	varchar(2)	
	libellé_genre	varchar	
 Add field			

Figure 11-Table de dimension «Genre_individu ».

La dimension « Age individu »

Elle contient la clé primaire `age_individu` les attributs suivants `code_categorieage` et `libellé_categorieage`.

La figure suivante montre la table de dimension « Age_individu ».

Age_individu			
	age_individu	integer(3)	
	code_categorieage	varchar(2)	
	libellé_categorieage	varchar	

Figure 12-Table de dimension «Age_individu ».

La dimension « Nationalité individu »

Elle contient la clé primaire `code_nationalité` et l'attribut `libellé_nationalité`.

La figure suivante montre la table de dimension « Nationalité_individu ».

Nationalité_individu			
	code_nationalité	varchar(2)	
	libellé_nationalité	varchar	

Figure 13-Table de dimension «Nationalité_individu ».

La dimension « Etatmatlég_individu »

Elle contient la clé primaire `code_étatmatlég` et l'attribut `libellé_étatmatlég`.

La figure suivante montre la table de dimension « Etatmatlég_individu ».

Etatmatlég_individu			
	code_étatmatlég	varchar(2)	
	libellé_étatmatlég	varchar	

Figure 14-Table de dimension « Etatmatlég_individu ».

La dimension « CSP_individu »

Elle contient la clé primaire `code_cspdétaillée` et les attributs : `libellé_cspdétaillée`, `code_cspregroupée` et `libellé_cspregroupée`.

La figure suivante montre la table de dimension « CSP_individu ».

CSP_individu			
	code_cspdétaillée	varchar(2)	
	libellé_cspdétaillé	varchar	
	code_cspregroupée	varchar(2)	
	libellé_cspregroupée	varchar	

Figure 15-Table de dimension « CSP_individu ».

La dimension « Lieu_résidence »

Elle contient la clé primaire `code_départementrés` et les attributs : `libellé_départementrés`, `code_régionrés` et `libellé_régionrés`.

La figure suivante montre la table de dimension « Lieu_résidence ».

Lieu_résidence			
	code_départementrés	varchar(2)	
	nom_départementrés	varchar	
	code_régionrés	varchar(2)	
	nom_régionrés	varchar	

Figure 16-Table de dimension « Lieu_résidence ».

La dimension « Type_ménage »

Elle contient la clé primaire `code_typedétaillé` et les attributs : `libellé_typedétaillé`, `code_typeregroupé` et `libellé_typeregroupé`.

La figure suivante montre la table de dimension « Type_ménage ».

Type_ménage			
	code_typedétaillé	varchar(2)	
	libellé_typedétaillé	varchar	
	code_typeregroupé	varchar(2)	
	libellé_typeregroupé	varchar	

Figure 17-Table de dimension « Type_ménage ».

La dimension « Trancherevenu_ménage »

Elle contient la clé primaire `code_trancherevenu` et l'attribut `libellé_trancherevenu`.

La figure suivante montre la table de dimension « Trancherevenu_individu ».

Trancherevenu_ménage			
	code_trancherevenu	varchar(2)	
	libellé_trancherevenu	varchar	

Figure 18-Table de dimension « Trancherevenu_ménage ».

La dimension « Date_départ »

Elle contient la clé primaire `dateheure_départ` et les attributs : `code_jourdep`, `libellé_jourdep`, `code_moisdép`, `libellé_moisdép`, `code_annédép` et `heure_départ`.

La figure suivante montre la table de dimension « Date_départ ».

Date_départ			
	dateheure_départ	datetime(19)	
	code_jourdép	date(10)	
	libellé_jourdép	varchar	
	code_moisdép	varchar(6)	
	libellé_moisdép	varchar	
	code_annéedép	integer(4)	
	heure_départ	time(8)	

Figure 19-Table de dimension « Date_départ».

La dimension « Date_arrivée»

Elle contient la clé primaire `dateheure_arrivée` et les attributs : `code_jourarr`, `libellé_jourarr`, `code_moisarr`, `libellé_moisrr`, `code_annéearr` et `heure_arrivée`.

La figure suivante montre la table de dimension « Date_arrivée ».

Date_arrivée			
	dateheure_arrivée	datetime(19)	
	code_jourarr	date(10)	
	libellé_jourarr	varchar	
	code_moisarr	varchar(6)	
	libellé_moisarr	varchar	
	code_annéearr	integer(4)	
	heure_arrivée	time(8)	

Figure 20-Table de dimension « Date_arrivée».

La dimension « Lieu_départ »

Elle contient la clé primaire `code_lieudépart` et les attributs : `code_départementdép`, `libellé_départementdép`, `code_régiondép`, `libellé_régiondép`, `code_paysdép` et `libellé_paysdép`.

La figure suivante montre la table de dimension « Lieu_départ ».

Lieu_départ			
	code_lieudépart	varchar(6)	
	code_départementdép	varchar(2)	
	nom_départementdép	varchar	
	code_régiondép	varchar(2)	
	nom_régiondép	varchar	
	code_paysdép	varchar(3)	
	libellé_paysdép	varchar	
 Add field			

Figure 21-Table de dimension «Lieu_départ».

La dimension « Lieu_destination »

Elle contient la clé primaire `code_lieudestination` et les attributs : `code_départementdes`, `libellé_départementdes`, `code_régiondes`, `libellé_régiondes`, `code_paysdes` et `libellé_paysdes`.

La figure suivante montre la table de dimension « Lieu_destination ».

Lieu_destination			
	<code>code_lieudestination</code>	<code>varchar(6)</code>	
	<code>code_départementdes</code>	<code>varchar(3)</code>	
	<code>nom_départementdes</code>	<code>varchar</code>	
	<code>code_régiondes</code>	<code>varchar(2)</code>	
	<code>nom_régiondes</code>	<code>varchar</code>	
	<code>code_paysdes</code>	<code>varchar(3)</code>	
	<code>nom_paysdes</code>	<code>varchar</code>	
 Add field			

Figure 22-Table de dimension «Lieu_destination».

La dimension « Motif_déplacement »

Elle contient la clé primaire `code_motif` et les attributs : `libellé_motif`, `code_catégorie`, `libellé_catégorie`, `code_type` et `libellé_type`.

La figure suivante montre la table de dimension « Motif_déplacement ».

Motif_déplacement			
	<code>code_motif</code>	<code>varchar(4)</code>	
	<code>libellé_motif</code>	<code>varchar</code>	
	<code>code_catégorie</code>	<code>varchar(2)</code>	
	<code>libellé_catégorie</code>	<code>varchar</code>	
	<code>code_type</code>	<code>varchar(2)</code>	
	<code>libellé_type</code>	<code>varchar</code>	

Figure 23-Table de dimension «Motif_déplacement».

La dimension « Mode_transport»

Elle contient la clé primaire `code_mode` et les attributs : `libellé_mode`, `code_catégorie` et `libellé_catégorie`.

La figure suivante montre la table de dimension « Mode_transport».

Mode_transport			
	code_mode	varchar(4)	
	libellé_mode	varchar	
	code_catégorie	varchar(2)	
	catégorie_mode	varchar	

Figure 24-Table de dimension «Mode_transport».

La dimension « Type_carburant»

Elle contient la clé primaire `code_typecarb` et l'attribut : `libellé_typecarb`.

La figure suivante montre la table de dimension « Type_carburant».

Type_carburant			
	code_typecarb	varchar(2)	
	libellé_typecarb	varchar	

Figure 25-Table de dimension «Type_carburant».

4.2.2. Faits

Les faits représentent les sujets d'analyse. Dans notre SID nous allons proposer un seul fait.

Le fait «Déplacements »

Il contient les clés primaires `id_déplacement` et `id_voyage` et l'attribut `type_déplacement`.

Il contient les clés primaires étrangères vers les dimensions : `code_genre`, `age_individu`, `code_nationalité`, `code_étatmatlég`, `code_cspdétaillée`, `code_départementrés`, `code_typedétaillée`, `code_trancherevenu`, `dateheure_départ`, `dateheure_arrivée`, `code_lieudépart`, `code_lieudestination`, `code_motif`, `code_mode` et `code_typecarb`.

Il contient les mesures : `distance`, `durée`, `cosommation_carburant`, `émissions_CO2` et `nombre_personnes`.

Il contient les attributs des dimensions dégénérés : `id_inividu`, `id_ménage` et `id_véhicule`.

La figure suivante montre la table du fait «Déplacements».

Déplacements			
🔑	id_déplacement	varchar(15)	✎ ⚙
🔑	id_voyage	varchar(14)	✎ ⚙
	type_déplacement	varchar	✎ ⚙
🔑	code_genre	varchar(2)	✎ ⚙
🔑	age_individu	integer(3)	✎ ⚙
🔑	code_nationalité	varchar(2)	✎ ⚙
🔑	code_étatmatlég	varchar(2)	✎ ⚙
🔑	code_cspdétaillée	varchar(2)	✎ ⚙
🔑	code_départementrés	varchar(2)	✎ ⚙
🔑	code_typedétaillé	varchar(2)	✎ ⚙
🔑	code_trancherevenu	varchar(2)	✎ ⚙
🔑	dateheure_départ	datetime(19)	✎ ⚙
🔑	dateheure_arrivée	datetime(19)	✎ ⚙
🔑	code_lieudépart	varchar(6)	✎ ⚙
🔑	code_lieudestination	varchar(6)	✎ ⚙
🔑	code_motif	varchar(4)	✎ ⚙
🔑	code_mode	varchar(4)	✎ ⚙
🔑	code_typecarb	varchar(2)	✎ ⚙
	distance	float(12)	✎ ⚙
	durée	float(12)	✎ ⚙
	consommation_carburant	float(12)	✎ ⚙
	émissions_CO2	float(12)	✎ ⚙
	nombre_personnes	integer(2)	✎ ⚙
🔑	id_individu	varchar(13)	✎ ⚙
🔑	id_ménage	varchar(11)	✎ ⚙
🔑	id_véhicule	varchar(14)	✎ ⚙

Figure 26-Table du fait « Déplacements ».

4.2.3. Schéma en étoile

A partir de la table du fait et les tables des dimensions précédents, nous allons choisir de construire le schéma en étoile car il est le plus utilisé et performant en termes de temps de réponse aux requêtes (car le nombre de jointures est limité seulement entre la table du fait et les table des dimensions).

La figure suivante représente le schéma en étoile de l'ED de notre SID qui est composé d'une table fait centrale (Déplacements) qui référence plusieurs tables des dimensions avec des clés étrangères

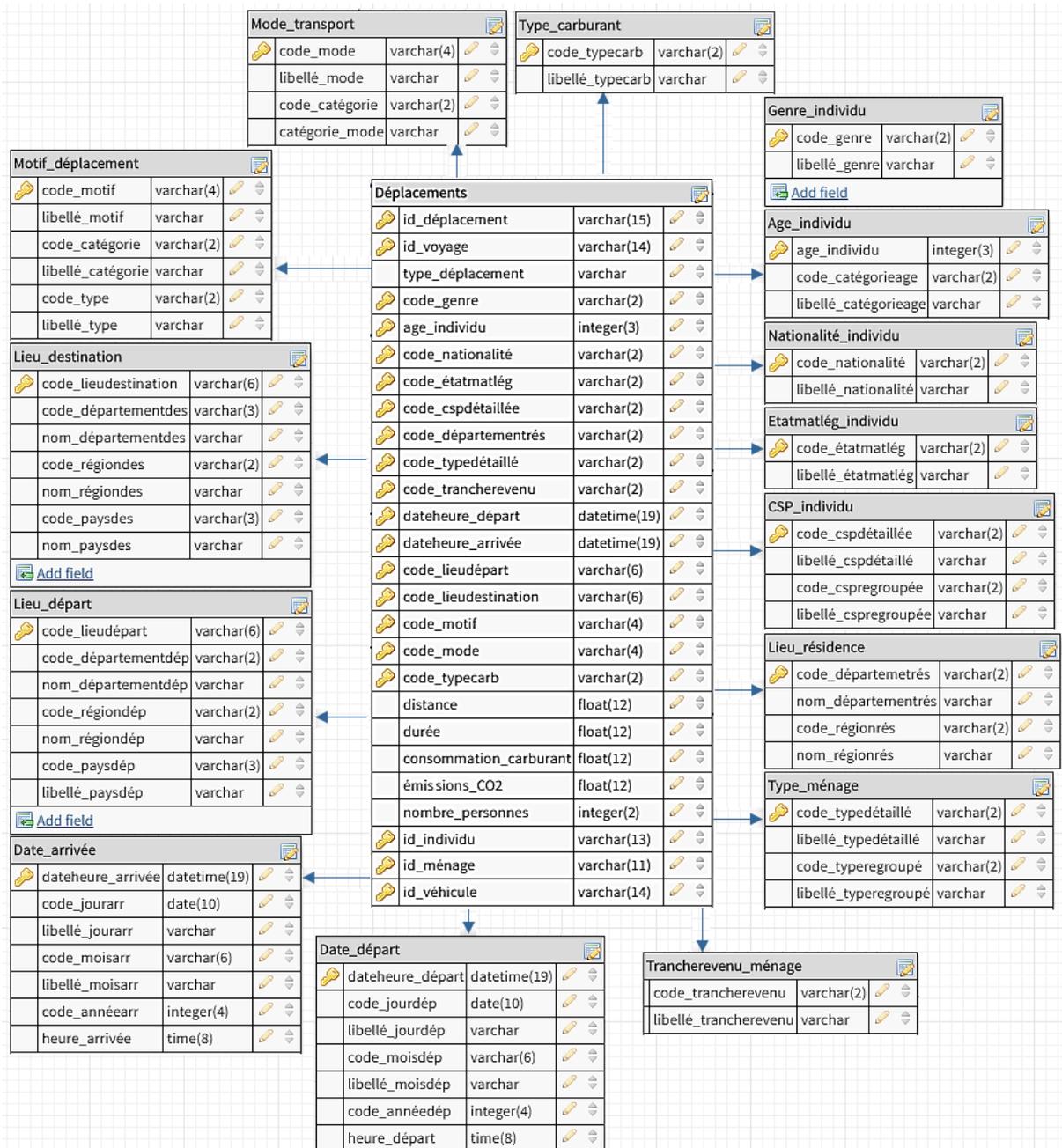


Figure 27-Schéma en étoile de notre SID.

4.3. Conclusion

Dans cette étape, nous avons modélisé l'ED de notre SID. Premièrement nous avons décrit les tables des dimensions et la table du fait, puis à partir de ces derniers nous avons créé le schéma en étoile qui sera ensuite transformé en un modèle physique implémenté dans un langage du SGBD cible.

Dans le chapitre suivant, nous allons décrire en détail l'implémentation de l'architecture technique de notre SID en utilisant les différents logiciels et technologies.

5. Implémentation de SID

5.1. Introduction

Dans ce dernier chapitre, nous allons montrer les principaux étapes pour l'implémentation de l'architecture technique de notre système d'information décisionnel(SID) en utilisant quelques logiciels et technologies afin d'établir quelques rapports permettant aux décideurs de prendre des divers décisions concernant l'application de la transition énergétique sur les systèmes de mobilité.

5.2. Architecture de SID

Dans la figure suivante, nous allons présenter l'architecture technique de notre SID qui est composée de quatre couches logicielles.

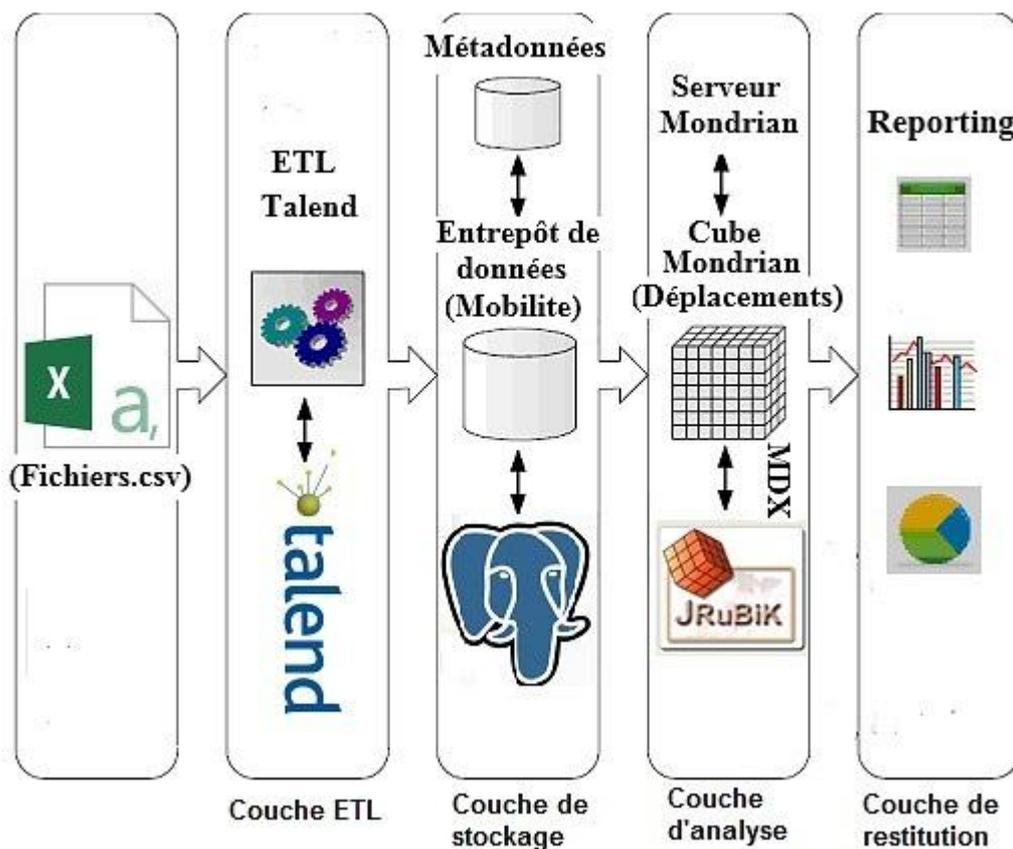


Figure 28-Architecture technique de notre SID.

- ❖ La couche ETL représente L'extraction, la transformation et le chargement des sources de données (fichies.csv) à l'aide de l'ETL Talend.
- ❖ La couche de stockage de données dans la base de l'entrepôt de données (ED) « Mobilité » implémenté dans le SGBDR « PostgreSQL ».
- ❖ La couche d'analyse du cube « Cube déplacements » qui déployé dans le serveur

« Mondrian » à l'aide de l'application « JRUBIK ».

- ❖ La couche de restitution représente la génération des exemples du Reporting en utilisant l'application « JRUBIK ».

Pour l'implémentation de cette architecture, tout d'abord nous allons choisir de travailler avec ces logiciels :

- ❖ Talend (Smile, 2013) Open Studio est un ETL open source, développé par la société Talend, basée en France. Talend est un ETL de type « générateur de code », c'est-à-dire qu'il permet de créer graphiquement des processus de manipulation et de transformation de données puis de générer l'exécutable correspondant sous forme de programme Java. Ce programme doit ensuite être déployé sur le serveur d'exécution.
- ❖ PostgreSQL (Lardière, 2007) est un gestionnaire de bases de données relationnelles (SGBDR) supportant le langage SQL. Il a été développé à partir du projet Postgres 4.2 initié par l'Université de Californie à Berkeley (UCB), département informatique, dès 1986. Postgres est lui-même dérivé d'Ingres. PostgreSQL est développé selon le mode « open source », sous licence BSD.
- ❖ Mondrian (ESPINASSE, Travail Pratique associé au cours « Entrepôts de données » Mise en œuvre de MONDRIAN et JRUBIK, 2017) est un serveur OLAP, développé en Open-Source depuis 2001 sous licence CPL. Ce serveur se présente sous la forme d'un ensemble de bibliothèques Java. Ces bibliothèques permettent tout d'abord la construction de structures multidimensionnelles (cubes notamment) à partir d'un entrepôt stocké dans une base de données relationnelle. Ces structures multidimensionnelles, ou cubes MONDRIAN, sont définies selon un schéma spécifié dans un fichier XML ⁷préalablement élaboré.
- ❖ JRUBIK (ESPINASSE, Travail Pratique associé au cours « Entrepôts de données » Mise en œuvre de MONDRIAN et JRUBIK, 2017) est une application Open-Source Java autonome permettant la connexion et l'analyse OLAP d'un cube MONDRIAN en langage MDX. L'utilisateur dispose avec JRUBIK d'une interface graphique à base d'onglets lui permettant d'effectuer de façon interactive des analyses, des requêtes OLAP en MDX sur un cube MONDRIAN et d'en visualiser les résultats. Pour cela JRUBIK doit tout d'abord charger, à partir de la base de

⁷ XML (Extensible Markup Language) est un méta-langage permettant de définir des langages à balises. Il est standardisé comme une recommandation par le W3C depuis 1998 (CROZAT, Introduction à XML : principes, syntaxe, schémas et manipulations, 2016).

données relationnelle de l'entrepôt de l'entrepôt, les données dans le cube MONDRIAN spécifié dans le schéma XML, et ensuite agréger ces données en mémoire cache.

Puis nous allons suivre les étapes suivantes :

5.3. Création et alimentation de l'entrepôt de données

Pour la création et l'alimentation de base de l'entrepôt de données(ED), nous allons suivre les traitements comme suit :

- La création de base d'ED nommée « Mobilite ».

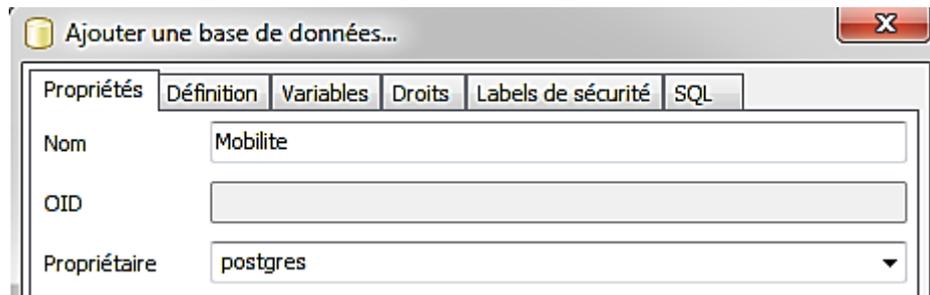


Figure 29-Création de base de données « Mobilite ».

- La création des schémas de base « Mobilite» : Dans notre base« Mobilite », nous avons créé deux schémas :
 - ✓ Le premier nommé « ods », il représente les tables des sources de données.
 - ✓ Le deuxième nommé « dwh», il représente les tables de notre modèle multidimensionnel (schéma en étoile).

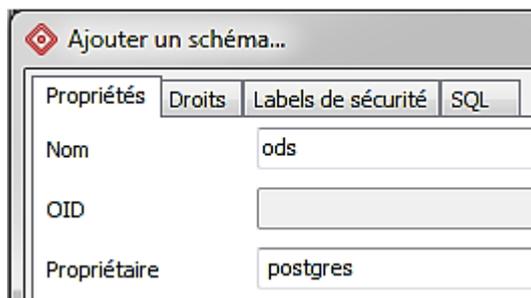


Figure 30-Création du schéma « ods ».

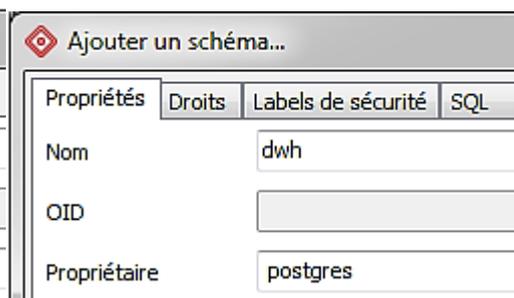


Figure 31-Création du schéma « dwh ».

- La création des tables du schéma « ods ».
Les tables du schéma « ods » sont étaient créés automatiquement à l'aide de l'ETL Talend lors les transformations des données.
- La création des tables du schéma « dwh».

Les tables du schéma « ods » sont étaiées créés à l'aide de le SGBDR PostgreSQL, premièrement nous avons créé les tables des dimensions puis la table du fait.

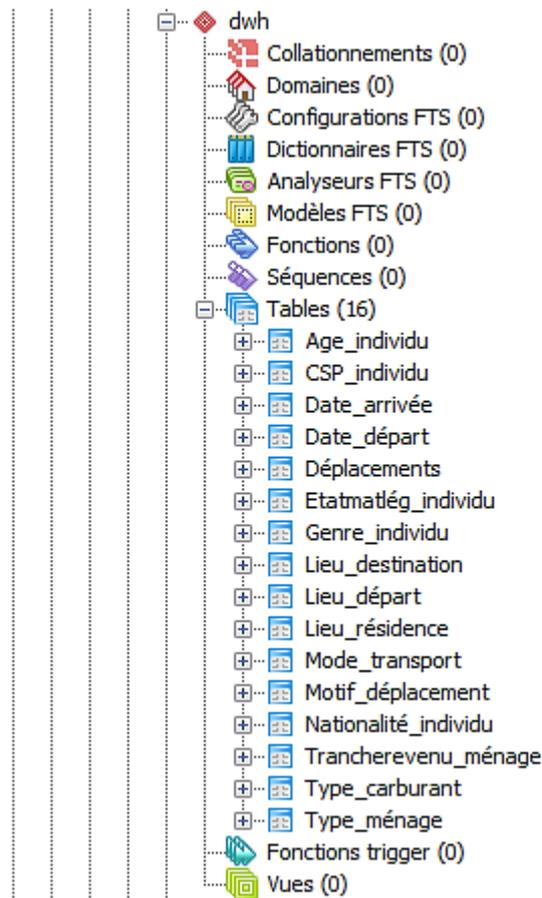


Figure 4-Liste de tables du schéma « dwh ».

- L'alimentation de base « Mobilité » : Pour alimenter les tables des schémas de base « Mobilité », nous allons utiliser l'ETL Talend Open Studio et nous allons suivre les traitements suivants :
 - ✓ On lance l'ETL Talend « TOS_DI-Win32-20170623_1246-V6.4.1 » et on crée un projet nommé « ProjetTalend ».
 - ✓ Puis on crée des fichiers délimités et des connexions aux bases de données de métadonnées (Figure 5).

Les fichiers délimités représentent les sources de données sous forme des fichiers (.csv). Pour les connexions aux bases de données, nous avons créé deux connexions l'une connectée au schéma « ods » et l'autre au schéma « dwh ».

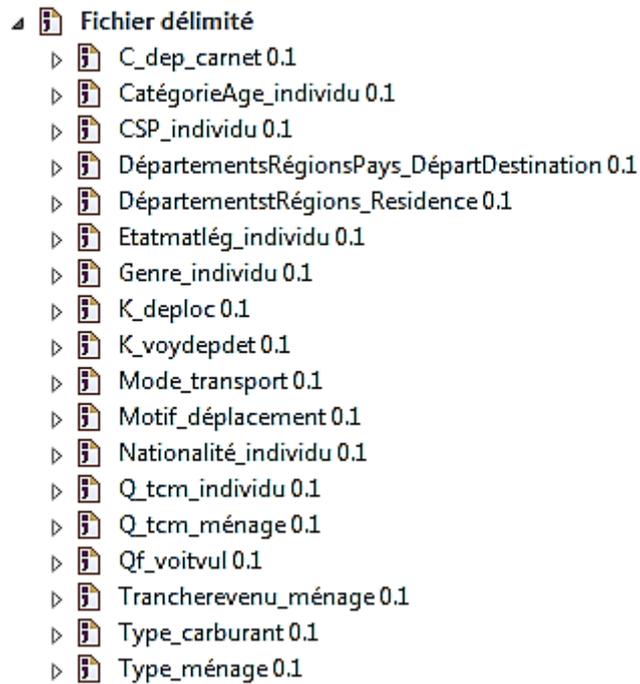


Figure 32-Fichiers délimités et connexions des métadonnées.

- ✓ La création des jobs : Dans cette phase nous avons classé les jobs en deux catégories.

La première catégorie représente les jobs (de Job01 jusqu'à Job16) permettant de la transformation et le chargement des tables des sources de données (les fichiers délimités) aux tables dans le schéma « ods » de base « Mobilité ».

La deuxième catégorie représente les jobs (de Job17 jusqu'à Job32) permettant la transformation et le chargement des données des tables du schéma « ods » dans les tables dans le schéma « dwh » de base « Mobilité ».

Job01 : la transformation et le chargement de fichier délimité « Genre_individu.csv » vers le schéma « ods » sous forme d'une table « Genre_individu ».

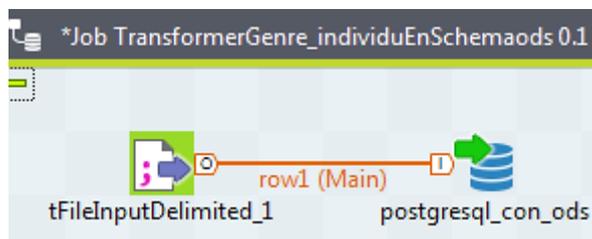


Figure 33-Job « TransformerGenre_individuEnSchemaods ».

Job02 : la transformation et le chargement de deux fichiers délimités « Q_tcm_individu.csv » et « Catégorieage_individu » vers le schéma « ods » sous forme d'une table « Age_individu ».

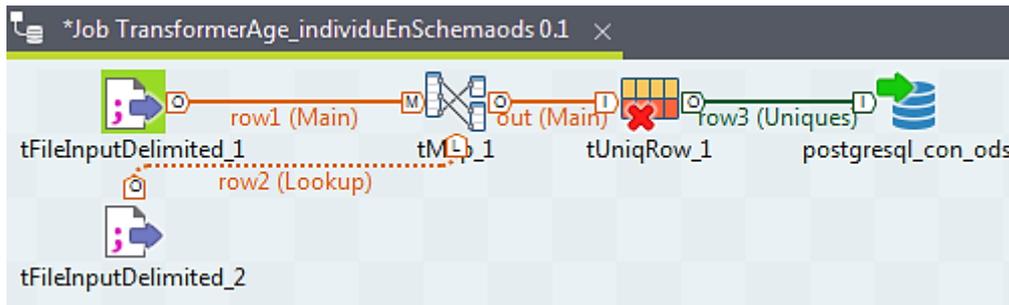


Figure 34-Job « TransformerAge_individuEnSchemaods ».

Job03: la transformation et le chargement de fichier délimité « Nationalité_individu.csv » vers le schéma « ods » sous forme d'une table « Nationalité_individu ».

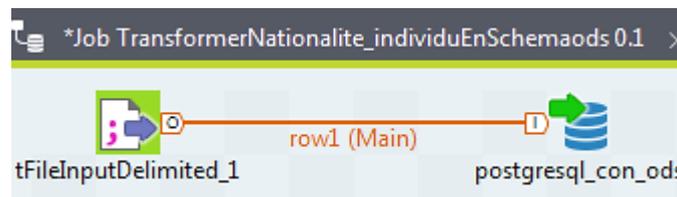


Figure 35-Job « TransformerNationalite_individuEnSchemaods ».

Job04: la transformation et le chargement de fichier délimité « Etatmatlég_individu.csv » vers le schéma « ods » sous forme d'une table « Etatmatlég_individu ».

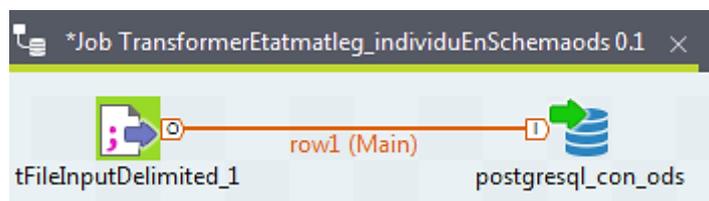


Figure 36-Job « TransformerEtatmatrimoniallegalEnSchemaods ».

Job05 : la transformation et le chargement de fichier délimité « CSP_individu.csv » vers le schéma « ods » sous forme d'une table « CSP_individu ».

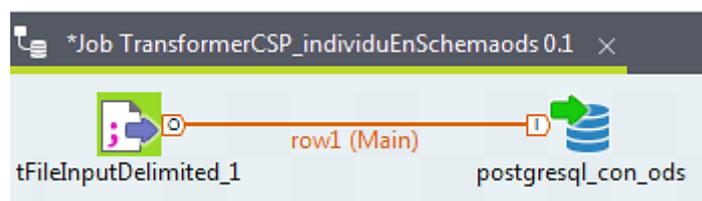


Figure 37-Job « TransformerCSP_individuEnSchemaods ».

Job06: la transformation et le chargement de fichier délimité « DépartementstRégionsRésidence.csv » vers le schéma « ods » sous forme d'une table « Lieu_destination ».

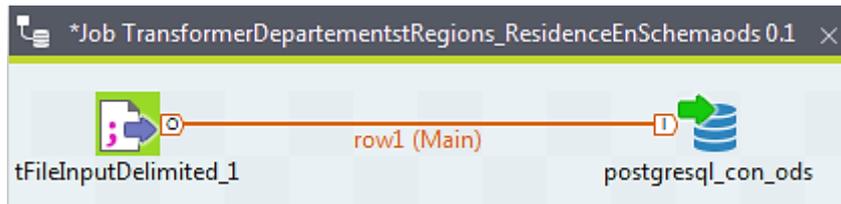


Figure 38-Job « TransformerDépartementstRégions_ResidenceEnSchemaods ».

Job07: la transformation et le chargement de fichier délimité « Type_ménage.csv »vers le schéma « ods » sous forme d'une table « Type_ménage ».

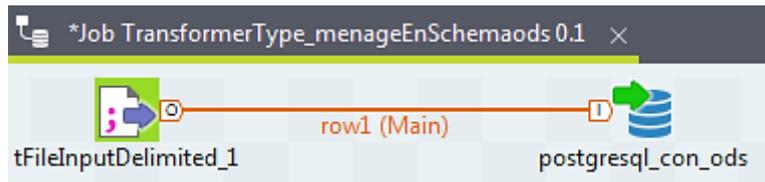


Figure 39-Job « TransformerType_ménageEnSchemaods ».

Job08: la transformation et le chargement de fichier délimité « Trancherevenu_ménage.csv»vers le schéma « ods » sous forme d'une table « Trancherevenu_ménage».

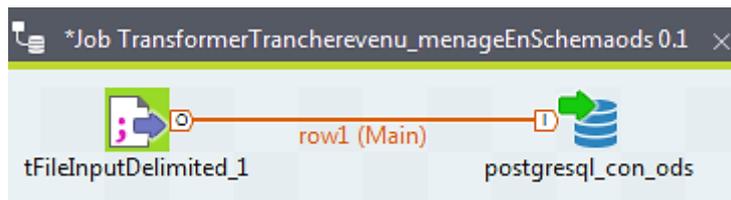


Figure 40-Job « TransformerTrancherevenu_ménageEnSchemaods ».

Job09: la transformation et le chargement de fichier délimité « DépartementsRégionsPays_DépartDestination.csv»vers le schéma « ods » sous forme de deux tables :« Lieu_départ » et :« Lieu_destination ».

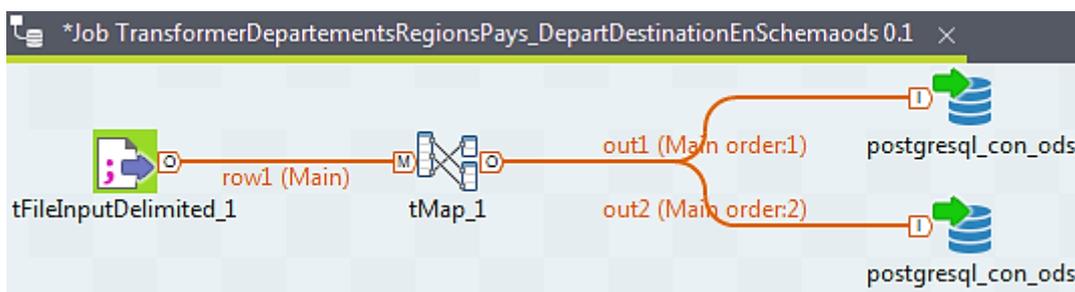


Figure 41-Job « TransformerDeptPaysDepart_DestinationEnSchemaods ».

Job10: la transformation et le chargement de fichier délimité « Motif_déplacement.csv»vers le schéma « ods » sous forme de table :« Motif_déplacement».

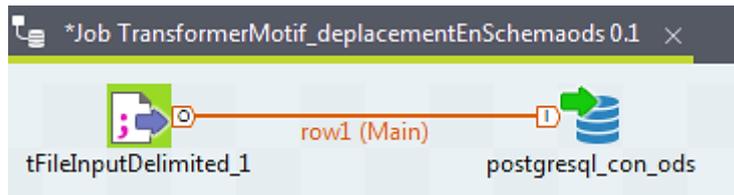


Figure 42-Job « TransformerMotif_deplacementEnSchemaods ».

Job11: la transformation et le chargement de fichier délimité « Mode_transport.csv » vers le schéma « ods » sous forme de table « Mode_transport ».

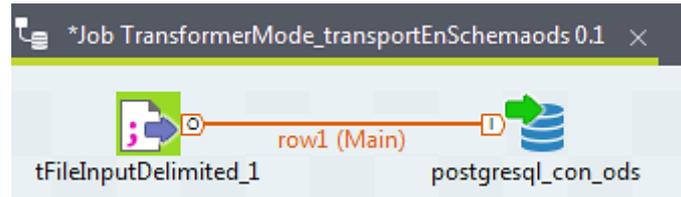


Figure 43-Job « TransformerMode_transportEnSchemaods ».

Job12: la transformation et le chargement de fichier délimité « Type_carburant.csv » vers le schéma « ods » sous forme de table « Type_carburant ».

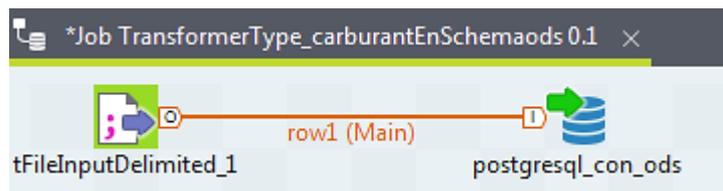


Figure 44-Job « TransformerType_carburantEnSchemaods ».

Job13: la transformation et le chargement de fichier délimité « K_deploc.csv » vers le schéma « ods » sous forme de deux tables : « Date_départdeploc » et « Date_arrivéedeploc ».

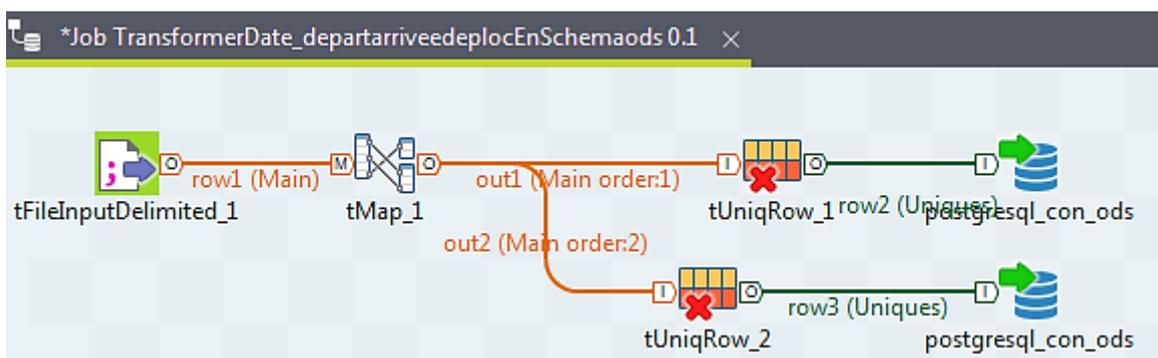


Figure 45-Job « TransformerDate_departarriveedeplocEnSchemaods ».

Job14: la transformation et le chargement de fichier délimité « K_voydet.csv » vers le schéma « ods » sous forme de deux tables : « Date_départvoydet » et « Date_arrivéevoydet ».

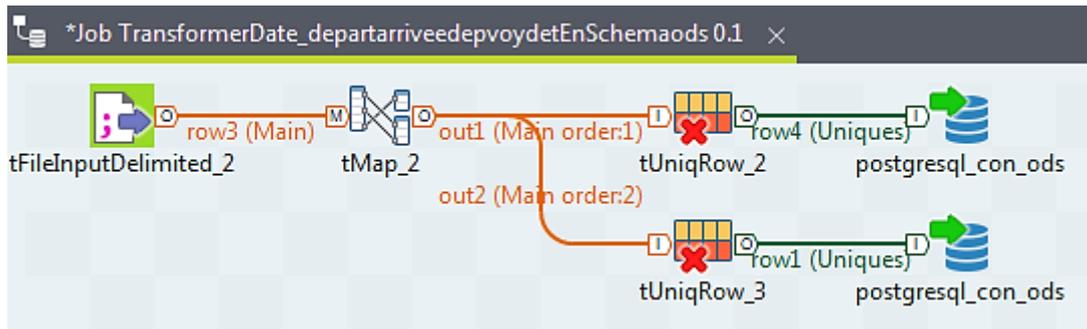


Figure 46-Job « TransformerDate_departarriveedepvoydetEnSchemaods ».

Job15: la transformation et le chargement des quatre fichiers délimités « K_deploc.csv » « Q_tcm_menage.csv », « Q_tcm_individu.csv » et « Q_tcm_menage.csv », « Q_tcm_individu.csv » vers le schéma « ods » sous forme de table « Déplacements_locals » .

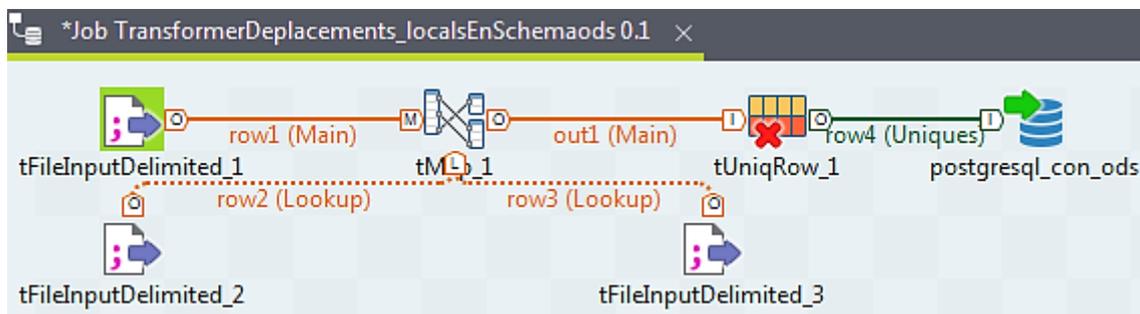


Figure 47-Job « TransformerDeplacements_localEnSchemaods ».

Job16: la transformation et le chargement des trois fichiers délimités «K_voydet.csv», « Q_tcm_menage.csv » et « Q_tcm_individu.csv » vers le schéma « ods » sous forme de table « Déplacements_voyages ».

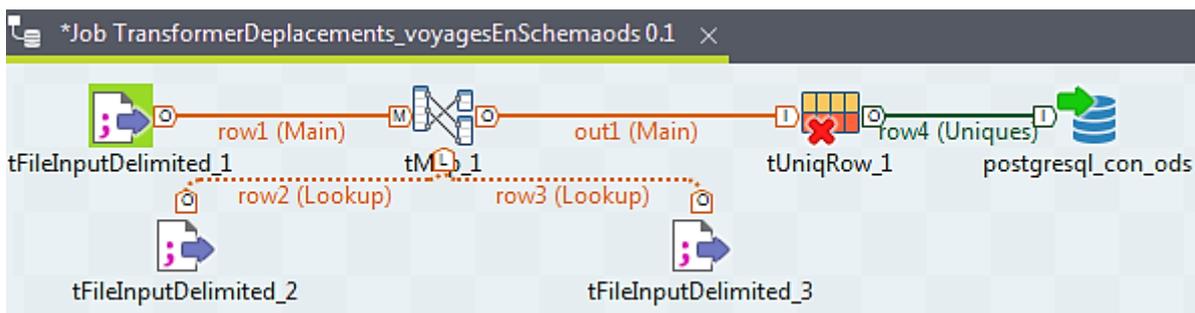


Figure 48-Job « TransformerDeplacements_voydetEnSchemaods ».

Après la transformation et le chargement des tables de sources de données dans le schéma «ods», nous allons récupérer les schémas des tables de deux schémas «ods» et «dwh», puis nous allons transformer et charger les données des tables du schéma «ods» dans les tables du schéma «dwh».

La figure suivante représente les schémas des tables de deux schémas «ods» et «dwh».

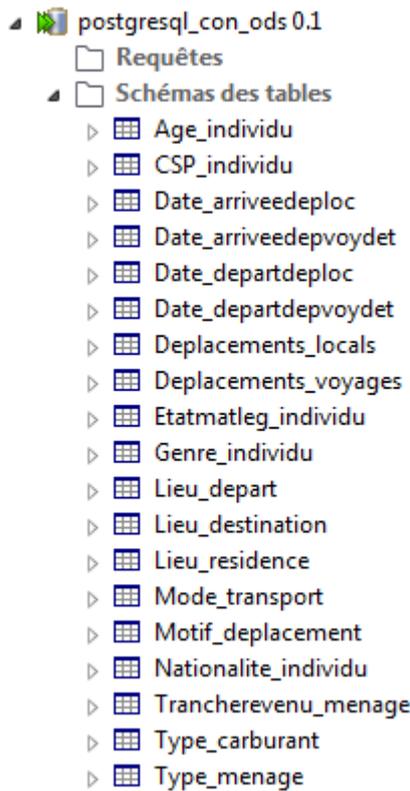


Figure 49-Tables du schéma «ods».

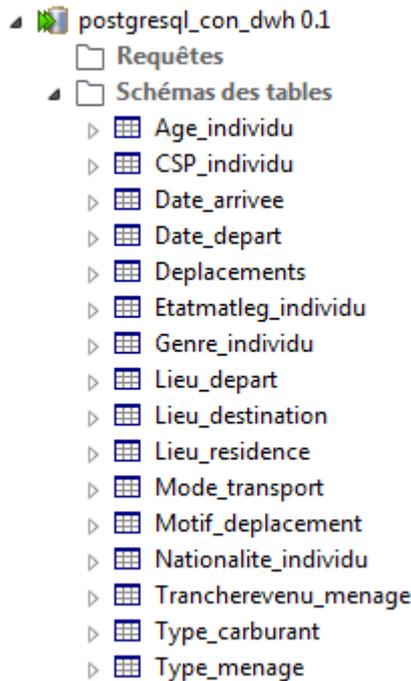


Figure 50-Tables du schéma «dwh».

Job17 : la transformation et le chargement des données de la table« Genre_individu »de schéma « ods » le dans la table « Genre_individu » de schéma « dwh ».

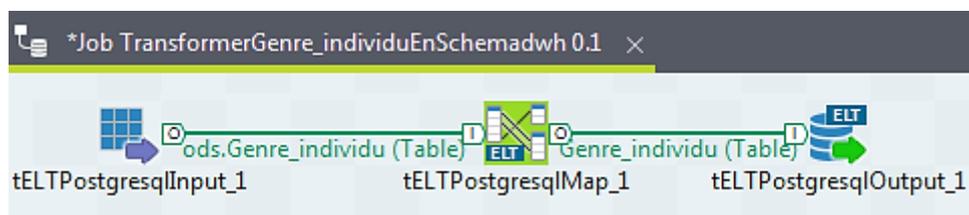


Figure 51-Job « TransformerGenre_individuEnSchemadwh ».

Job18: la transformation et le chargement des données de la table«Age_individu » de schéma « ods » dans la table « Age_individu » de schéma « dwh ».

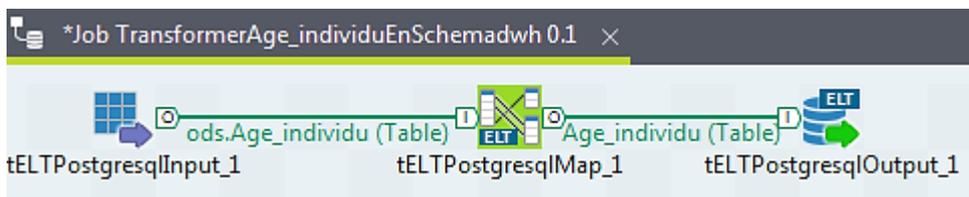


Figure 52-Job « TransformerAge_individuEnSchemadwh ».

Job19: la transformation et le chargement des données de la table« Nationalité_individu »de schéma « ods » dans la table « Nationalité__individuo » de schéma « dwh ».

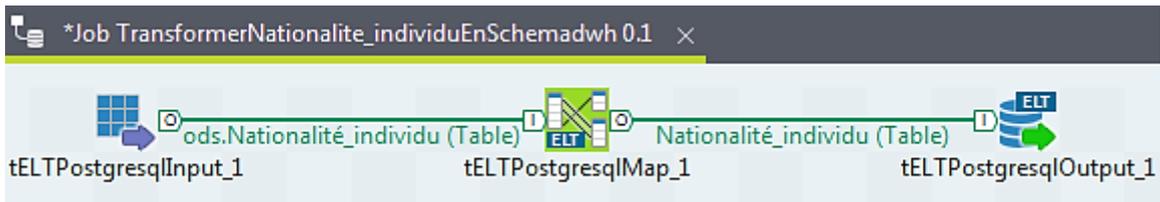


Figure 53-Job « TransformerNationalité_individuEnSchemadwh ».

Job20: la transformation et le chargement des données de la table« Etatmatlég_individu»de schéma « ods » dans la table « Etatmatlég_individu » de schéma « dwh ».

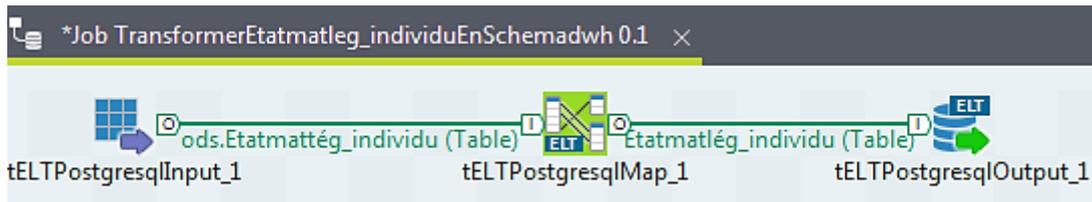


Figure 54-Job « TransformerEtatmatlég_individuEnSchemadwh ».

Job21 : la transformation et le chargement des données de la table« CSP_individu »de schéma « ods » dans la table «CSP_individu» de schéma « dwh».

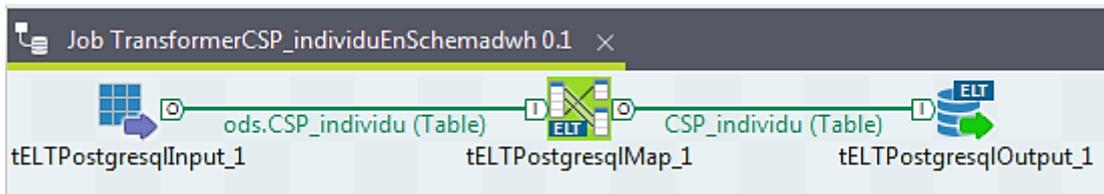


Figure 55-Job « TransformerCSP_individuEnSchemadwh ».

Job22 : la transformation et le chargement des données de la table« Lieu_résidence »de schéma « ods » dans la table «Lieu_résidence» de schéma « dwh ».

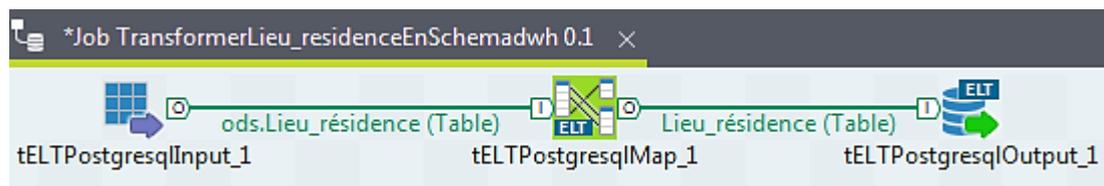


Figure 56-Job « TransformerLieu_destinationEnSchemadwh ».

Job23 : la transformation et le chargement des données de la table« Type_ménage »de schéma « ods » dans la table «Type_ménage » de schéma « dwh ».

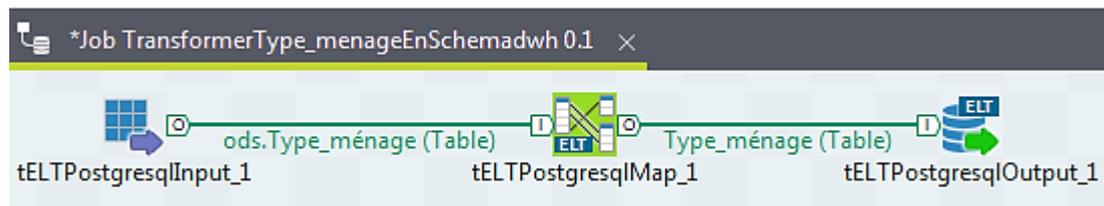


Figure 57-Job « TransformerType_menageEnSchemadwh ».

Job24 : la transformation et le chargement des données de la table « Trancherevenu_ménage » de schéma « ods » dans la table « Trancherevenu_ménage » de schéma « dwh ».

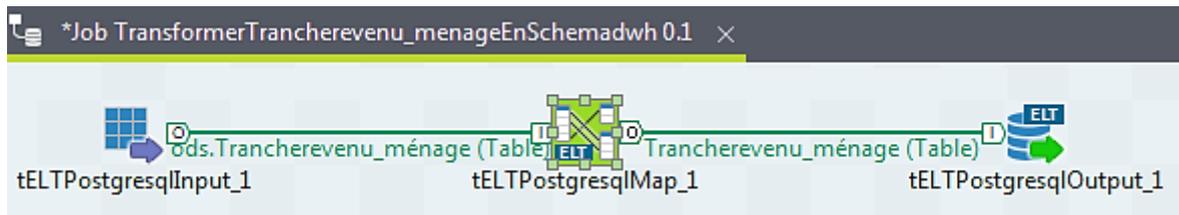


Figure 58-Job « TransformerTrancherevenu_ménageEnSchemadwh ».

Job25 : la transformation et le chargement des données des deux tables « Date_départdeploc » et « Date_départvoydet » de schéma « ods » dans la table « Date_départ » de schéma « dwh ».

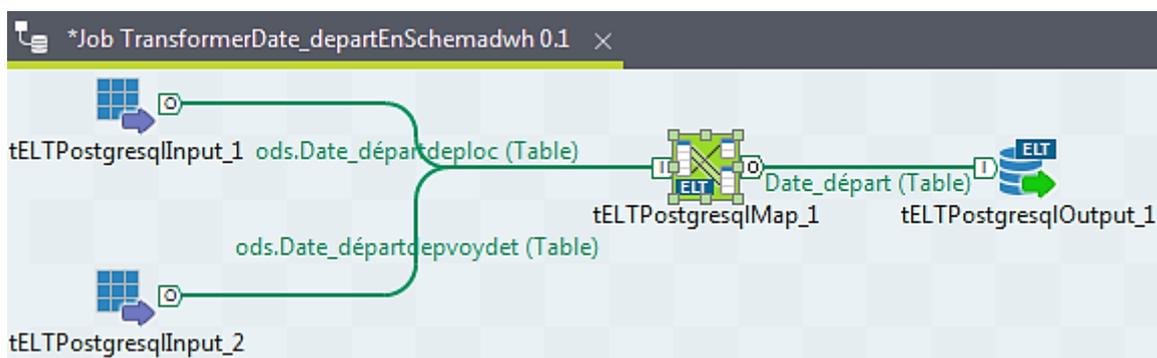


Figure 59-Job « TransformerDate_departEnSchemadwh ».

Job26 : la transformation et le chargement des données deux tables « Date_arrivéedeploc » et « Date_arrivéevoydet » de schéma « ods » dans la table « Date_arrivée » de schéma « dwh ».

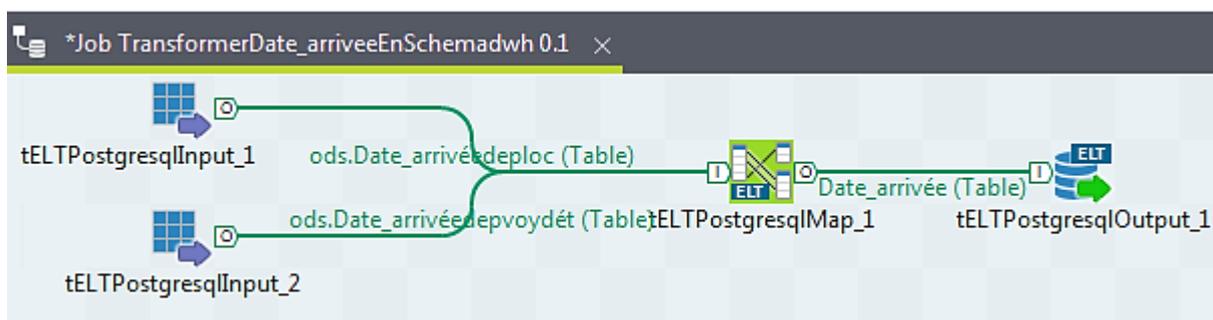


Figure 60-Job « TransformerDate_arriveeEnSchemadwh ».

Job27 : la transformation et le chargement des données de la table « Lieu_départ » de schéma « ods » dans la table « Lieu_départ » de schéma « dwh ».

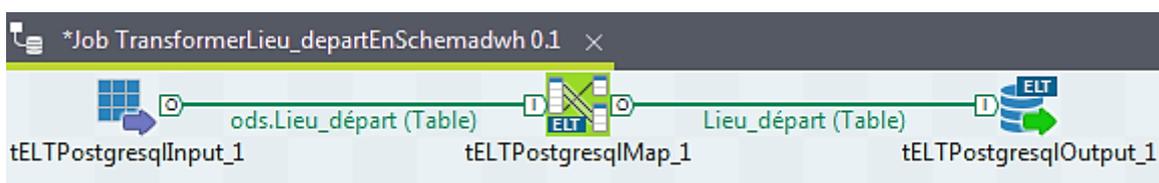


Figure 61-Job « TransformerLieu_departEnSchemadwh ».

Job28 : la transformation et le chargement des données de la table « Lieu_destination » de schéma « ods » dans la table « Lieu_destination » de schéma « dwh ».

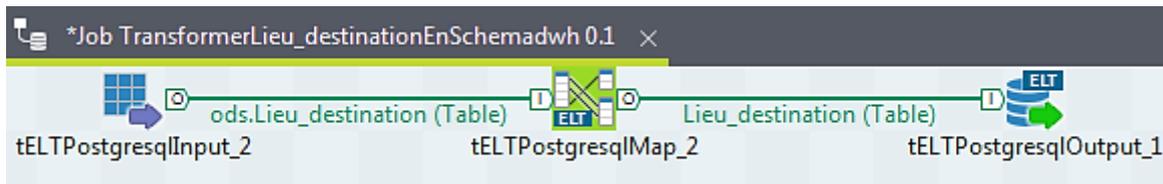


Figure 62-Job « TransformerLieu_destinationEnSchemadwh ».

Job29 : la transformation et le chargement des données de la table « Motif_déplacement » de schéma « ods » dans la table « Motif_déplacement » de schéma « dwh ».

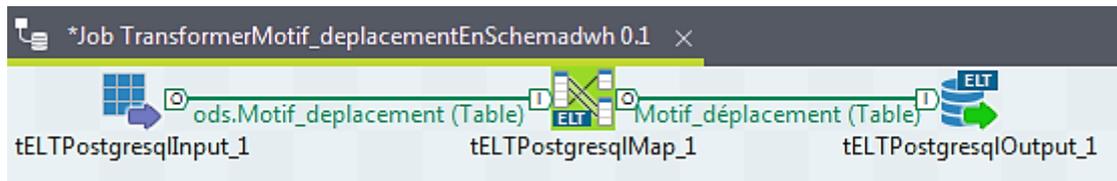


Figure 63-Job « TransformerMotif_deplacementEnSchemadwh ».

Job30 : la transformation et le chargement des données de la table « Mode_transport » de schéma « ods » dans la table « Mode_transport » de schéma « dwh ».

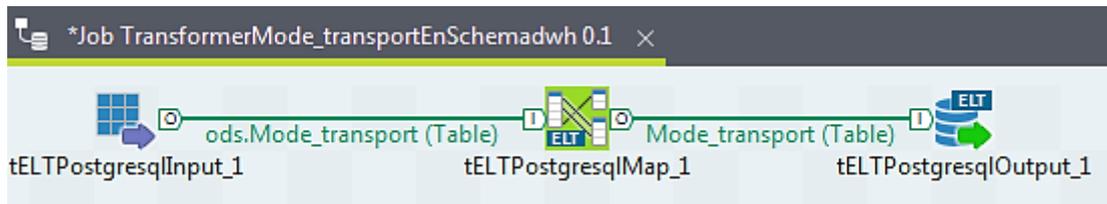


Figure 64-Job « TransformerMode_transportEnSchemadwh ».

Job31 : la transformation et le chargement des données de la table « Type_carburant » de schéma « ods » dans la table « Type_carburant » de schéma « dwh ».

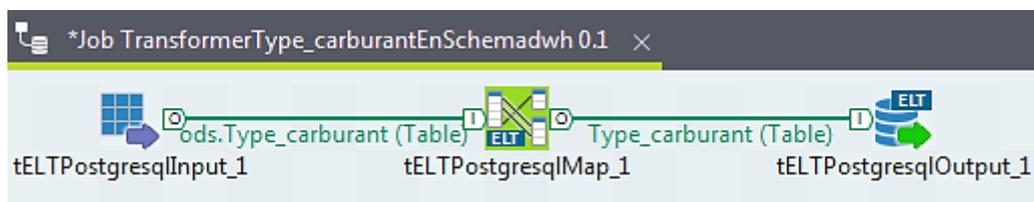


Figure 65-Job « TransformerType_carburantEnSchemadwh ».

Job32: la transformation et le chargement des données des deux « Déplacements_locaux » et « Déplacements_voyages » de schéma « ods » dans la table « Déplacements » de schéma « dwh ».

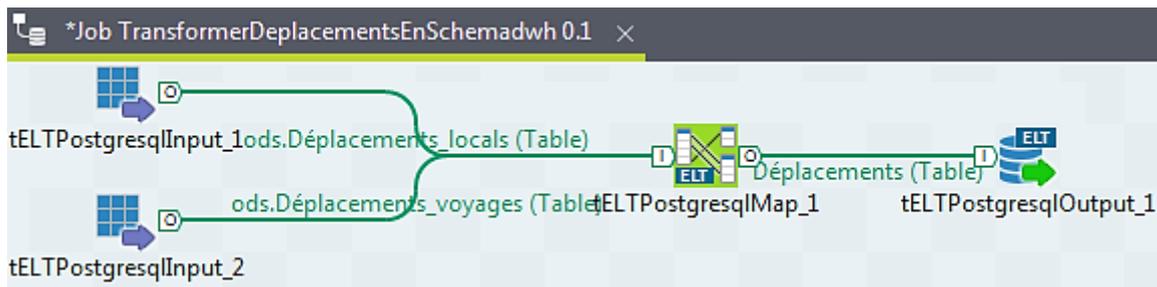


Figure 66-Job « TransformerDéplacementsEnSchemadwh ».

5.4. Création du cube OLAP

La création d'un cube au sens Mondrian est tout simplement la rédaction d'un fichier XML. Ce fichier permet de lier les informations du cube que nous souhaitons faire apparaître et la base de données de notre entrepôt.

Pour faciliter la création de ce fichier, nous allons utiliser l'outil Schéma Workbench qui offre une interface graphique pour effectuer cette tâche.

La figure suivante représente le formulaire de configuration de la base « Mobilite » sur laquelle nous souhaitons nous baser pour la création de cube.

Figure 67-Formulaire de configuration de base dans Schéma Workbench.

La figure suivante montre que pour la création de notre cube(ou bien la rédaction d'un fichier XML) nous avons besoins de :

- La table du fait « déplacements »
- Les tables des dimensions : « Genre_individu », « Age_individu », « Nationalité_individu », « Etatmatlég_individu », « CSP_individu », « Lieu_résidence », « Type_ménage », « Trancherevenu_ménage», « Date_départ», « Date_départ», « Lieu_départ», « Lieu_destination», « Motif_déplacement», « Mode_transport».
- Les hiérarchies (les niveaux de détail) pour chaque dimension.
- Les mesures (calculés et non calculés).

```

- <Schema name="CubeMultidimensionnel">
- <Dimension type="StandardDimension" foreignKey="" name="Motif de déplacement">
- <Hierarchy name="HierarchyMotif" hasAll="true" allMemberName="Tous" primaryKey="code_motif">
  <Table name="Motif_déplacement" schema="dwh" />
  <Level name="Type de motif" column="code_type" nameColumn="libellé_type" type="String"
    uniqueMembers="false" levelType="Regular" hideMemberIf="Never" />
  <Level name="Catégorie de motif" column="code_catégorie" nameColumn="libellé_catégorie"
    type="String" uniqueMembers="false" levelType="Regular" hideMemberIf="Never" />
  <Level name="Motif" column="code_motif" nameColumn="libellé_motif" type="String"
    uniqueMembers="false" levelType="Regular" hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Mode de transport">
- <Hierarchy name="HierarchyMode" hasAll="true" allMemberName="Tous" primaryKey="code_mode">
  <Table name="Mode_transport" schema="dwh" />
  <Level name="Catégorie de mode" column="code_catégoriemode"
    nameColumn="libellé_catégoriemode" type="String" uniqueMembers="false" levelType="Regular"
    hideMemberIf="Never" />
  <Level name="Mode" column="code_mode" nameColumn="libellé_mode" type="String"
    uniqueMembers="false" levelType="Regular" hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Type de carburant">
- <Hierarchy name="HierarchyTypeCarburant" hasAll="true" allMemberName="Tous"
  primaryKey="code_typecarb">
  <Table name="Type_carburant" schema="dwh" />
  <Level name="Type" column="code_typecarb" nameColumn="libellé_typecarb" type="String"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Genre d'individu">
- <Hierarchy name="HierarchyGenre" hasAll="true" allMemberName="Tous" primaryKey="code_genre">
  <Table name="Genre_individu" schema="dwh" />
  <Level name="Genre" column="code_genre" nameColumn="libellé_genre" type="String"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Age d'individu">
- <Hierarchy name="HierarchyAge" hasAll="true" allMemberName="Tous" primaryKey="age_individu">
  <Table name="Age_individu" schema="dwh" />
  <Level name="Categorie d'age" column="code_catégorieage"
    nameColumn="libellé_catégorieage" type="String" uniqueMembers="true" levelType="Regular"
    hideMemberIf="Never" />
  <Level name="Age" column="age_individu" type="Integer" uniqueMembers="true"
    levelType="Regular" hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Nationalité d'individu">
- <Hierarchy name="HierarchyNationalité" hasAll="true" allMemberName="Tous"
  primaryKey="code_nationalité">
  <Table name="Nationalité_individu" schema="dwh" />
  <Level name="Nationalite" column="code_nationalité" nameColumn="libellé_nationalité"
    type="String" uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Etat matrimonial légal d'individu">
- <Hierarchy name="HierarchyEtatMatrimonialLégal" hasAll="true" allMemberName="Tous"
  primaryKey="code_étatmatlég">
  <Table name="Etatmatlég_individu" schema="dwh" />
  <Level name="Etat matrimonial légal" column="code_étatmatlég"
    nameColumn="libellé_étatmatlég" type="String" uniqueMembers="true" levelType="Regular"
    hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="CSP d'individu">
- <Hierarchy name="HierarchyCSP" hasAll="true" allMemberName="Tous"
  primaryKey="code_cspdétaillée" primaryKeyTable="CSP_individu">
  <Table name="CSP_individu" schema="dwh" />
  <Level name="CSP regroupée" column="code_cspregroupée"
    nameColumn="libellé_cspregroupée" type="String" uniqueMembers="true" levelType="Regular"
    hideMemberIf="Never" />
  <Level name="CSP détaillée" column="code_cspdétaillée" nameColumn="libellé_cspdétaillée"
    type="String" uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  </Hierarchy>
</Dimension>

```

```

- <Dimension type="StandardDimension" name="Type de ménage">
- <Hierarchy name="HierarchyTypeMenage" hasAll="true" allMemberName="Tous"
  primaryKey="code_typedétaillé">
  <Table name="Type_ménage" schema="dwh" />
  <Level name="Type regroupé" column="code_typeregroupé" nameColumn="libellé_typeregroupé"
    type="String" uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  <Level name="Type détaillé;" column="code_typedétaillé" nameColumn="libellé_typedétaillé"
    type="String" uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Tranche de revenu mensuel net de ménage">
- <Hierarchy name="HierarchyTranchesRevenu" hasAll="true" allMemberName="Tous"
  primaryKey="code_trancherevenu">
  <Table name="Trancherevenu_ménage" schema="dwh" />
  <Level name="Tranche de revenu mensuel net" column="code_trancherevenu"
    nameColumn="libellé_trancherevenu" type="String" uniqueMembers="true" levelType="Regular"
    hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Lieu de résidence">
- <Hierarchy name="HierarchyLieuxRésidence" hasAll="true" allMemberName="Tous"
  primaryKey="code_départementrés">
  <Table name="Lieu_résidence" schema="dwh" />
  <Level name="Région" column="code_régionrés" nameColumn="nom_régionrés" type="String"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  <Level name="Département" column="code_départementrés"
    nameColumn="nom_départementrés" type="String" uniqueMembers="true" levelType="Regular"
    hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Lieu de départ">
- <Hierarchy name="HierarchyLieuxDépart" hasAll="true" allMemberName="Tous"
  primaryKey="code_lieudépart">
  <Table name="Lieu_départ1" schema="dwh" />
  <Level name="Pays" column="code_paysdép" nameColumn="nom_paysdép" type="String"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  <Level name="Région" column="code_régiondép" nameColumn="nom_régiondép" type="String"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  <Level name="Département" column="code_lieudépart" nameColumn="nom_départementdép"
    type="String" uniqueMembers="false" levelType="Regular" hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Lieu de destination">
- <Hierarchy name="HierarchyLieuxDestination" hasAll="true" allMemberName="Tous"
  primaryKey="code_lieudestination">
  <Table name="Lieu_destination1" schema="dwh" />
  <Level name="Pays" column="code_paysdes" nameColumn="nom_paysdes" type="String"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  <Level name="Région" column="code_régiondes" nameColumn="nom_régiondes" type="String"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  <Level name="Département" column="code_lieudestination"
    nameColumn="nom_départementdes" type="String" uniqueMembers="false" levelType="Regular"
    hideMemberIf="Never" />
  </Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Date de départ">
- <Hierarchy name="HierarchyDateDépartAMJH" hasAll="true" allMemberName="Tous"
  primaryKey="dateheure_départ">
  <Table name="Date_départ" schema="dwh" />
  <Level name="Année" column="code_annéedép" type="Integer" uniqueMembers="true"
    levelType="Regular" hideMemberIf="Never" />
  <Level name="Mois" column="code_moisdép" nameColumn="libellé_moisdép" type="String"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  <Level name="Jour" column="code_jourdép" nameColumn="libellé_jourdép" type="Date"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  <Level name="Heure" column="dateheure_départ" type="String" uniqueMembers="true">
  - <KeyExpression>
    <SQL dialect="postgres">to_char("Date_départ"."dateheure_départ", 'YYYY-MM-DD
      HH24:MI:SS')</SQL>
    </KeyExpression>
  - <NameExpression>
    <SQL dialect="postgres">to_char("Date_départ"."dateheure_départ", 'HH24:MI:SS')
    </SQL>
    </NameExpression>
  </Level>
</Hierarchy>

```

```

- <Hierarchy name="HierarchyDateDépartASJH" hasAll="true" allMemberName="Tous"
  primaryKey="dateheure_départ">
  <Table name="Date_départ" schema="dwh" />
  <Level name="Année" column="code_annéedép" type="Integer" uniqueMembers="true"
    levelType="Regular" hideMemberIf="Never" />
- <Level name="Semaine" type="String" uniqueMembers="true">
- <KeyExpression>
  <SQL dialect="postgres">Extract( 'week' from "Date_départ"."dateheure_départ")
  </SQL>
  </KeyExpression>
</Level>
<Level name="Jour" column="code_jourdep" nameColumn="libellé_jourdep" type="Date"
  uniqueMembers="true" />
- <Level name="Heure" column="dateheure_départ" type="String" uniqueMembers="true">
- <KeyExpression>
  <SQL dialect="postgres">to_char("Date_départ"."dateheure_départ", 'YYYY-MM-DD
  HH24:MI:SS')</SQL>
  </KeyExpression>
- <NameExpression>
  <SQL dialect="postgres">to_char("Date_départ"."dateheure_départ", 'HH24:MI:SS')
  </SQL>
  </NameExpression>
</Level>
</Hierarchy>
</Dimension>
- <Dimension type="StandardDimension" name="Date d'arrivée">
- <Hierarchy name="HierarchyDateArriveeAMJH" hasAll="true" allMemberName="Tous"
  primaryKey="dateheure_arrivée">
  <Table name="Date_arrivée" schema="dwh" />
  <Level name="Année" column="code_annéearr" type="Integer" levelType="Regular"
    hideMemberIf="Never" />
  <Level name="Mois" column="code_moisarr" nameColumn="libellé_moisarr" type="String"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
  <Level name="Jour" column="code_jourarr" nameColumn="libellé_jourarr" type="Date"
    uniqueMembers="true" levelType="Regular" hideMemberIf="Never" />
- <Level name="Heure" column="Date_arrivée" type="String" uniqueMembers="true">
- <KeyExpression>
  <SQL dialect="postgres">to_char("Date_arrivée"."dateheure_arrivée", 'YYYY-MM-DD
  HH24:MI:SS')</SQL>
  </KeyExpression>
- <NameExpression>
  <SQL dialect="postgres">to_char("Date_arrivée"."dateheure_arrivée", 'HH24:MI:SS')
  </SQL>
  </NameExpression>
</Level>
</Hierarchy>
- <Hierarchy name="HierarchyDateArrivéASJH" hasAll="true" allMemberName="Tous"
  primaryKey="dateheure_arrivée">
  <Table name="Date_arrivée" schema="dwh" />
  <Level name="Année" column="code_annéearr" type="Integer" uniqueMembers="true"
    levelType="Regular" hideMemberIf="Never" />
- <Level name="Semaine" type="String" uniqueMembers="false">
- <KeyExpression>
  <SQL dialect="postgres">Extract('week' from "Date_arrivée"."dateheure_arrivée")
  </SQL>
  </KeyExpression>
</Level>
<Level name="Jour" column="code_jourarr" nameColumn="libellé_jourarr" type="Date"
  uniqueMembers="true" />
- <Level name="Heure" column="Date_arrivée" type="String" uniqueMembers="true">
- <KeyExpression>
  <SQL dialect="postgres">to_char("Date_arrivée"."dateheure_arrivée", 'YYYY-MM-DD
  HH24:MI:SS')</SQL>
  </KeyExpression>
- <NameExpression>
  <SQL dialect="postgres">to_char("Date_arrivée"."dateheure_arrivée", 'HH24:MI:SS')
  </SQL>
  </NameExpression>
</Level>
</Hierarchy>
</Dimension>

```

```

- <Cube name="Cube de déplacements" cache="true" enabled="true">
  <Table name="Déplacements" schema="dwh" />
  <DimensionUsage source="Motif de déplacement" name="Motif de déplacement"
    foreignKey="code_motif" />
  <DimensionUsage source="Mode de transport" name="Mode de transport" foreignKey="code_mode" />
  <DimensionUsage source="Type de carburant" name="Type de carburant"
    foreignKey="code_typecarb" />
  <DimensionUsage source="Genre d'individu" name="Genre d'individu" foreignKey="code_genre" />
  <DimensionUsage source="Age d'individu" name="Age d'individu" foreignKey="age_individu" />
  <DimensionUsage source="Etat matrimonial légal d'individu" name="Etat matrimonial légal"
    foreignKey="code_étatmatlég" />
  <DimensionUsage source="CSP d'individu" name="CSP d'individu" foreignKey="code_cspdétaillée" />
  <DimensionUsage source="Type de ménage" name="Type de ménage"
    foreignKey="code_typedétaillé" />
  <DimensionUsage source="Tranche de revenu mensuel net de ménage" name="Tranche de revenu
    mensuel de ménage" foreignKey="code_trancherevenu" />
  <DimensionUsage source="Nationalité d'individu" name="Nationalité d'individu"
    foreignKey="code_nationalité" />
  <DimensionUsage source="Lieu de résidence" name="Lieu de résidence"
    foreignKey="code_départementrés" />
  <DimensionUsage source="Lieu de départ" name="Lieu de départ" foreignKey="code_lieudépart" />
  <DimensionUsage source="Lieu de destination" name="Lieu de destination"
    foreignKey="code_lieudestination" />
  <DimensionUsage source="Date de départ" name="Date de départ" foreignKey="dateheure_départ" />
  <DimensionUsage source="Date d'arrivée" name="Date d'arrivée" foreignKey="dateheure_arrivée" />
  <Measure name="distance moyenne" column="distance" datatype="Numeric"
    formatString="#,##0.000000" aggregator="avg" visible="true" />
  <Measure name="consommation totale de carburant" column="consommation_carburant"
    datatype="Numeric" formatString="#,###" aggregator="sum" visible="true" />
  <Measure name="émissions totales de CO2" column="émissions_CO2" datatype="Numeric"
    formatString="#,###" aggregator="sum" visible="true" />
  <Measure name="nombre des déplacements" column="id_déplacement" datatype="Integer"
    formatString="#,#" aggregator="distinct count" visible="true" />
</Cube>
</Schema>

```

Figure 68- Fichier XML du cube « Cube de Déplacements ».

5.5. Analyse du cube OLAP

Une fois le cube créé, nous allons déployer dans le serveur Mondrian, ce dernier connecté à la base relationnelle « Mobilite » de l'entrepôt de données (PostgreSQL) pour la récupération des données sous forme multidimensionnelle avec leurs niveaux de détail.

Pour l'exploitation du cube créé précédemment, nous allons suivre les étapes suivantes :

1. On lance l'application JRUBIK.
2. On clique sur Settings puis sur Wizard Drivers pour la configuration de la connexion au serveur PostgreSQL (c'est-à-dire indiquer à JRUBIK, à partir du poste de l'administrateur PostgreSQL (poste de l'utilisateur administrateur PostgreSQL), où se trouve la BD de l'entrepôt et comment récupérer les données.

Dans la figure suivante, nous allons monter la configuration de la connexion en spécifiant l'adresse, le port et le type de base de données.

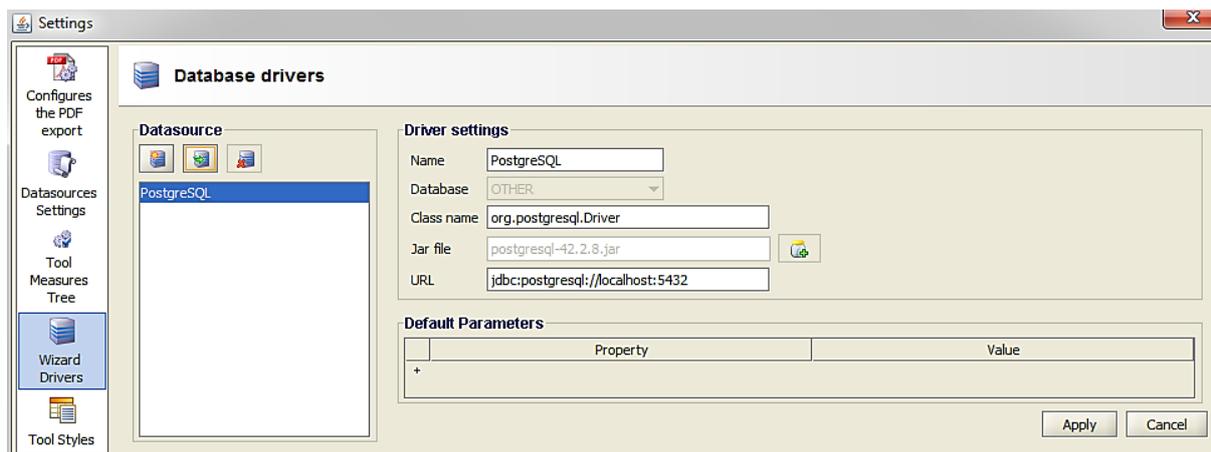


Figure 69-Fenêtre de la configuration de connexion de JRUBIK au serveur PostgreSQL.

3. On clique sur Settings puis Datasources Settings pour la configuration de source de données. C'est dans Data Source que JRUBIK saura où se trouve le fichier XML qui spécifie le cube, et où se trouve la base de données (en utilisant la connexion précédemment créée).

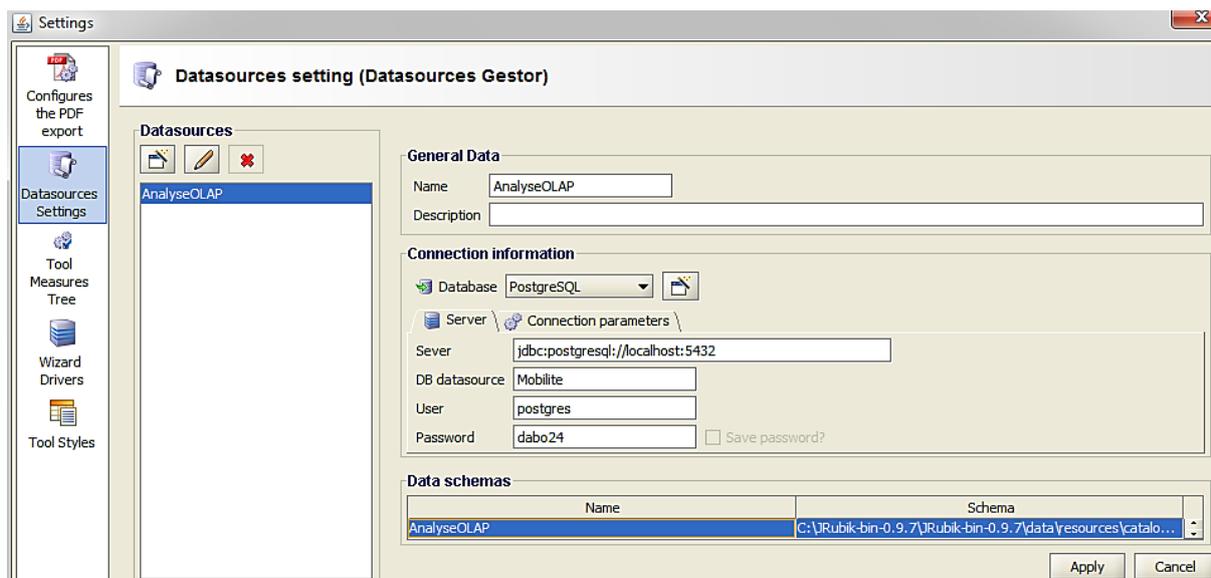


Figure 70-Fenêtre de la configuration de source de données.

4. Puis on clique sur apply pour appeler le cube créé dans le fichier XML.

La figure suivante représente le cube appelé « Cube de déplacements », ce cube se compose des dimensions avec leurs niveaux (hiérarchies) et des mesures.

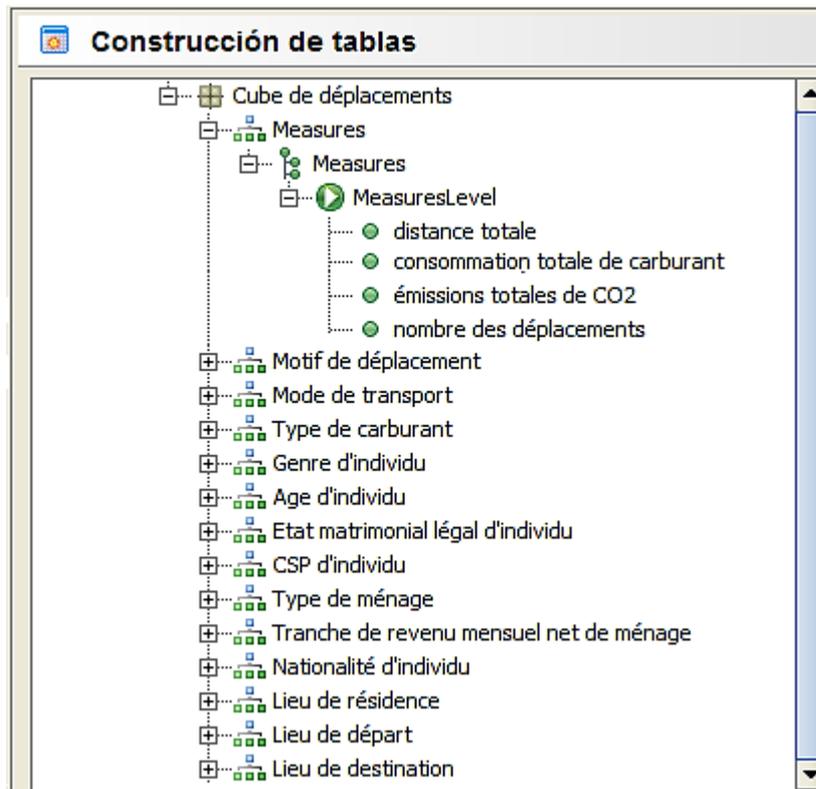


Figure 71-Cube « Cube de déplacements ».

Après avoir chargé à partir de la base de données relationnelle de l'entrepôt les données dans le cube MONDRIAN « Cube de déplacements » spécifié dans le schéma XML, et ensuite agréger ces données (les mesures) en mémoire cache.

Maintenant nous pouvons effectuer de façon interactive des analyses, des requêtes OLAP en MDX sur le cube « Cube de déplacements », et d'en visualiser les résultats soit sous forme des tables de données ou bien sous forme des graphiques.

Dans les figures suivantes, nous allons montrer un exemple d'analyse de « Cube de déplacements » dans JRUBIK qui représente le nombre de déplacements par genre de l'individu et par mode de transport.

Pour cela, nous allons suivre les étapes suivantes :

- ✓ On ouvre la fenêtre « Discover » puis on glisse les dimensions « Genre d'individu » et « Mode de transport » puis la mesure « nombre des déplacements ».

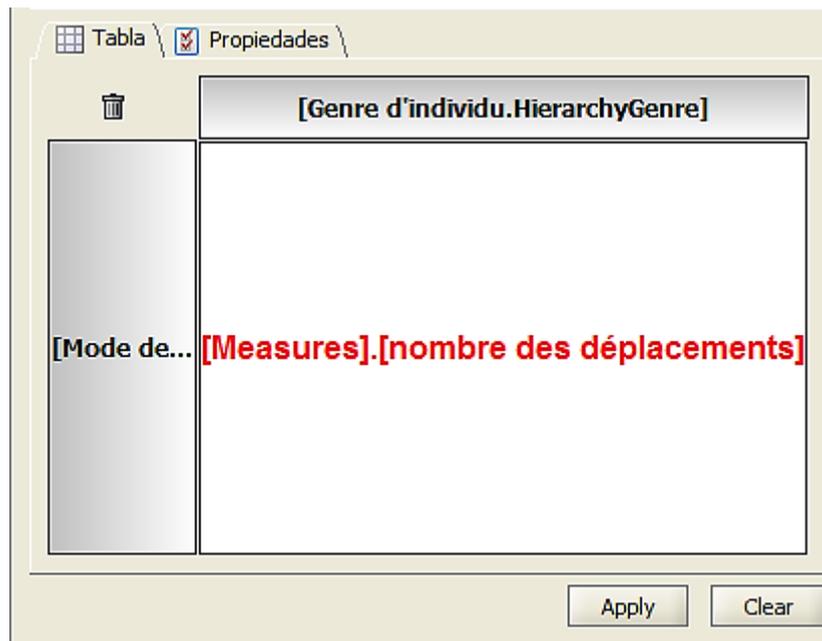


Figure 72-Fenêtre « Discover » pour glisser les dimensions et les mesures.

- ✓ On ouvre la fenêtre « Naviguer » puis on définit les niveaux de détail de chaque dimension.

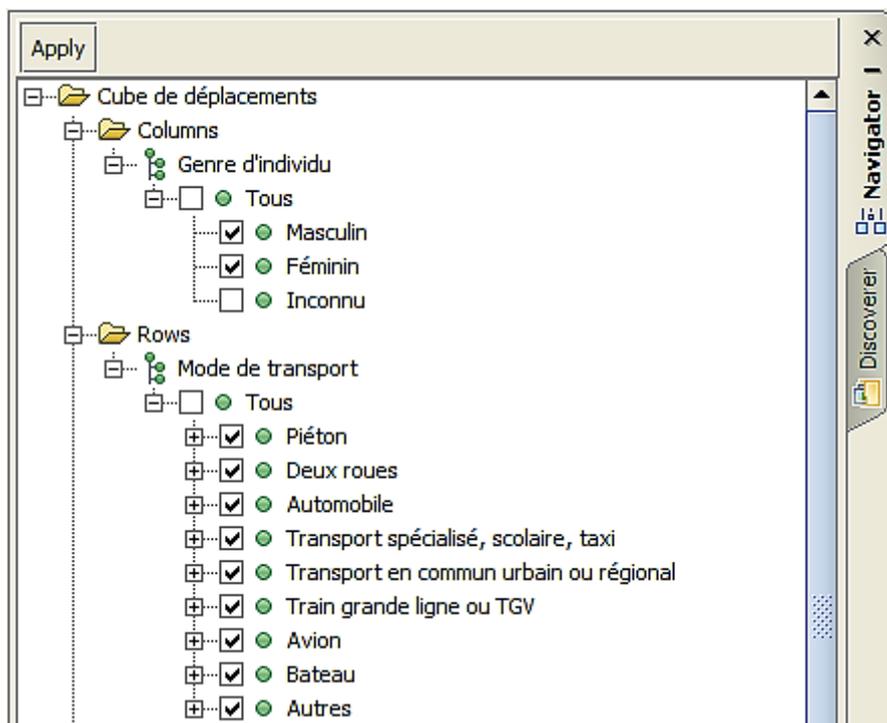


Figure 73-Fenêtre « Naviguer » pour définir les niveaux des dimensions.

- ✓ On clique sur apply, le résultat apparaitre sous forme d'un tableau(Figure74) et une requête MDX été généré(Figure75).

Mode de transport	Genre d'individu	
	Masculin	Féminin
+Piéton	56	56
+Deux roues	106	78
+Automobile	8 251	5 694
+Transport spécialisé, scolaire, taxi	100	80
+Transport en commun urbain ou régional	407	375
+Train grande ligne ou TGV	1 416	1 322
+Avion	697	513
+Bateau	23	25
+Autres	82	41

Figure 74-Exemple de résultat d'analyse sous forme d'un tableau.

```

MDX MDX query - x Editor
[Icons: File, Print, Refresh, Undo]

select {[Genre d'individu.HierarchyGenre].[Tous].[Masculin],
[Genre d'individu.HierarchyGenre].[Tous].[Féminin]}
ON COLUMNS,
{[Mode de transport.HierarchyMode].[Tous].[Piéton],
[Mode de transport.HierarchyMode].[Tous].[Deux roues],
[Mode de transport.HierarchyMode].[Tous].[Automobile],
[Mode de transport.HierarchyMode].[Tous].[Transport spécialisé, scolaire, taxi],
[Mode de transport.HierarchyMode].[Tous].[Transport en commun urbain ou régional],
[Mode de transport.HierarchyMode].[Tous].[Train grande ligne ou TGV],
[Mode de transport.HierarchyMode].[Tous].[Avion],
[Mode de transport.HierarchyMode].[Tous].[Bateau],
[Mode de transport.HierarchyMode].[Tous].[Autres]}
ON ROWS
from [Cube de déplacements]
where [Measures].[nombre des déplacements]

```

Figure 75-Exemple de résultat d'analyse sous forme d'une requête MDX.

On peut aussi visualiser le résultat sous forme d'un graphique en cliquant sur le bouton Chart



5.6. Exemples de Reporting

Dans cette section, nous allons présenter certains résultats d'analyse sous forme des rapports du « Cube de déplacements » concernant les sujets d'analyse suivants :

- Analyse des déplacements selon les caractéristiques de l'individu.
- Analyse des déplacements selon les caractéristiques du ménage.
- Analyse des déplacements selon le lieu de résidence.
- Analyse des déplacements selon le parc automobile.

➤ Analyse des flux de déplacements (cartographie des déplacements).

5.6.1. Analyse des déplacements selon les caractéristiques de l'individu

La figure 76 représente le nombre de déplacements par genre de l'individu et par mode de transport.

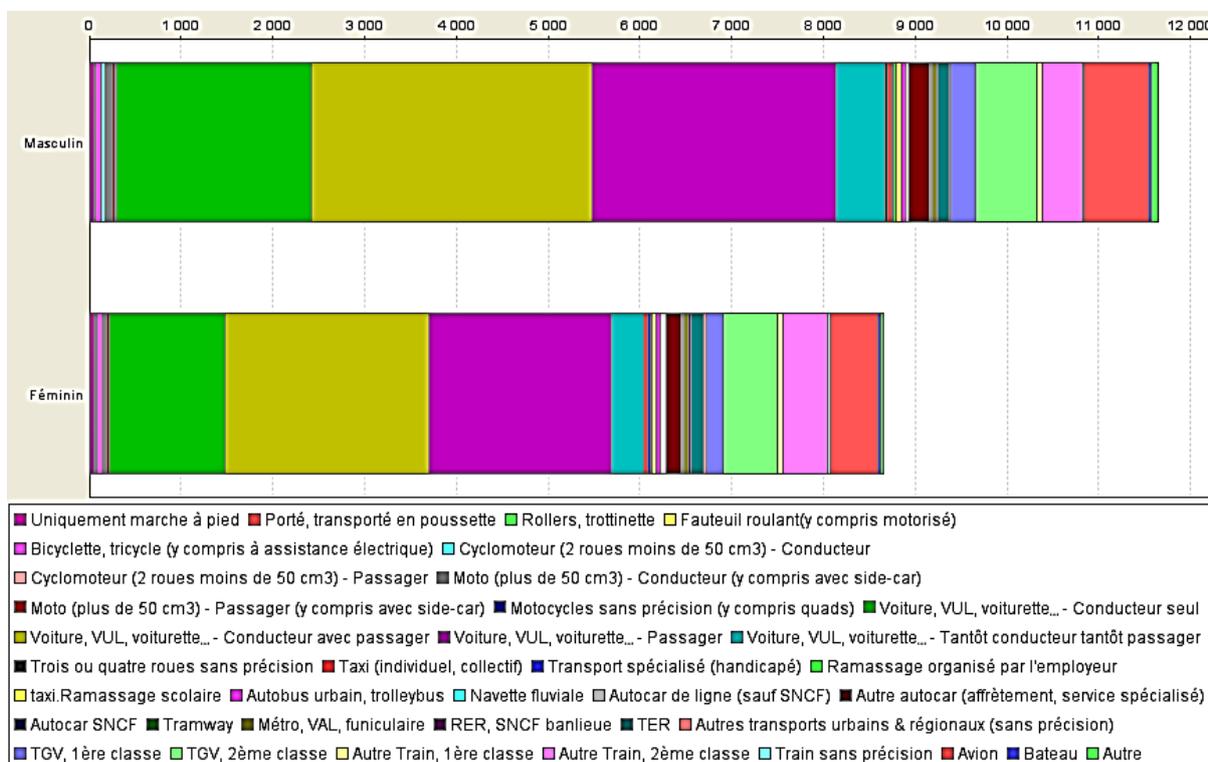


Figure 76- Nombre de déplacements par genre de l'individu et par mode de transport.

D'après la figure 76, nous remarquons que les hommes déplacent plus que les femmes et que les automobiles (les voitures, VULs et voiturettes... etc.) sont le mode le plus utilisé.

La figure 77 représente le nombre de déplacements par genre de l'individu et par motif de déplacement.

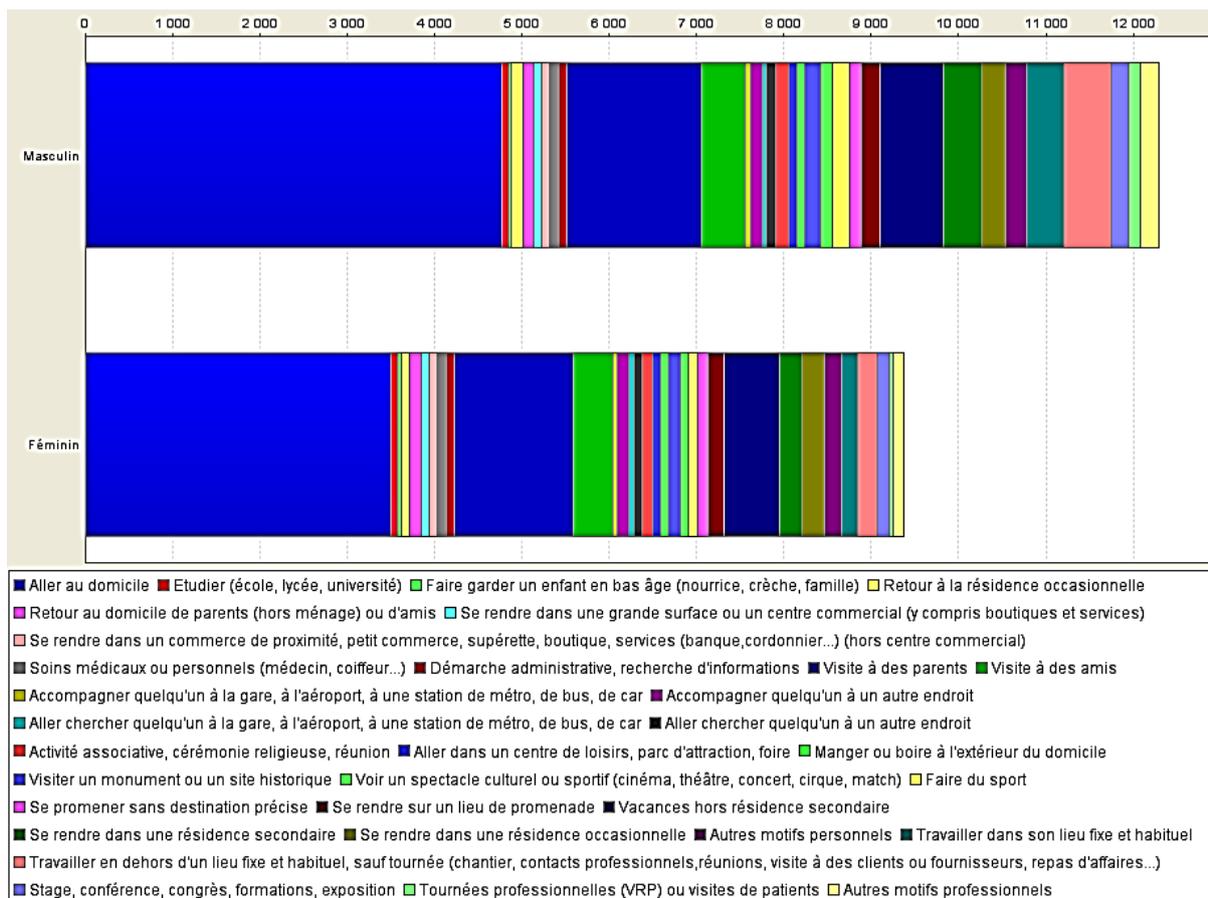


Figure 77-Nombre de déplacements par genre de l'individu et par motif de déplacement.

D'après la figure 77, nous remarquons que les motifs « Retour au domicile » et « Visite à des parents » représentent les raisons principales de la plupart des déplacements.

La figure 78 représente le nombre de déplacements par catégorie d'âge de l'individu et par mode de transport.

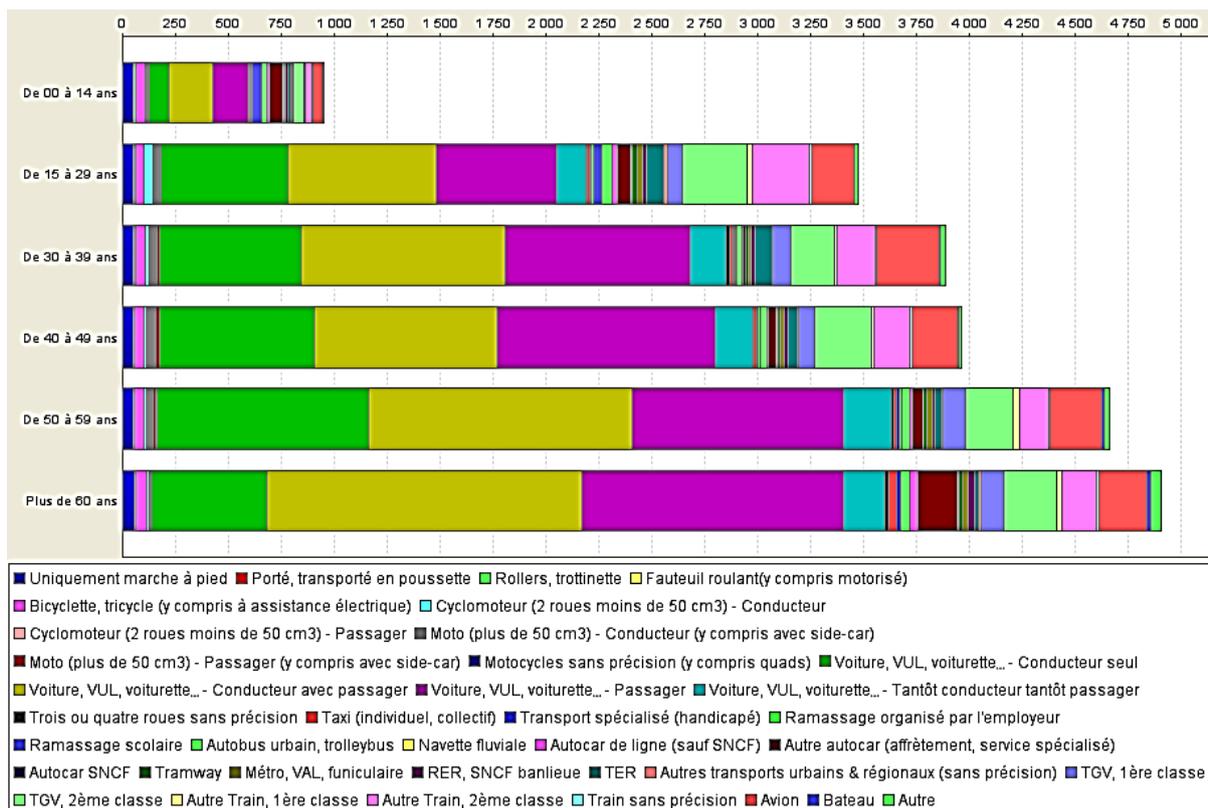


Figure 78- Nombre de déplacements par catégorie d'âge de l'individu et par mode de transport.

D'après la figure 78, nous remarquons que la catégorie d'âge « Plus de 60 ans » qui déplace plus par rapport aux autres catégories et que les automobiles (les voitures, VULs et voiturettes...etc.) sont le mode le plus utilisé.

La figure 79 représente le nombre de déplacements par catégorie d'âge de l'individu et par motif de déplacement.

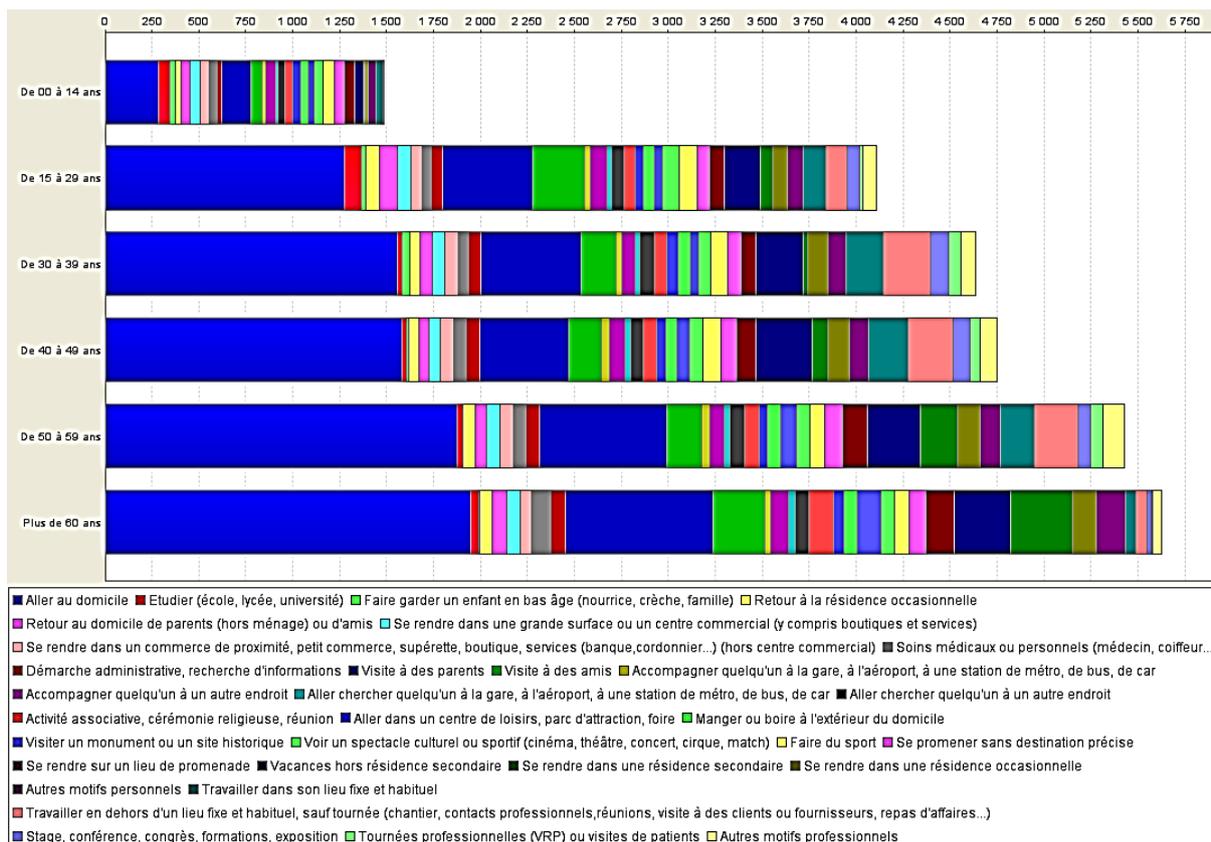


Figure 79- Nombre de déplacements par catégorie d'âge de l'individu et par motif de déplacement.

D'après la figure 79, nous remarquons que les motifs « Retour au domicile » et « Visite à des parents » représentent les raisons principales de la plupart des déplacements.

5.6.2. Analyse des déplacements selon les caractéristiques du ménage

La figure 80 représente le nombre de déplacements par tranche du revenu mensuel net du ménage et par catégorie de mode de transport.

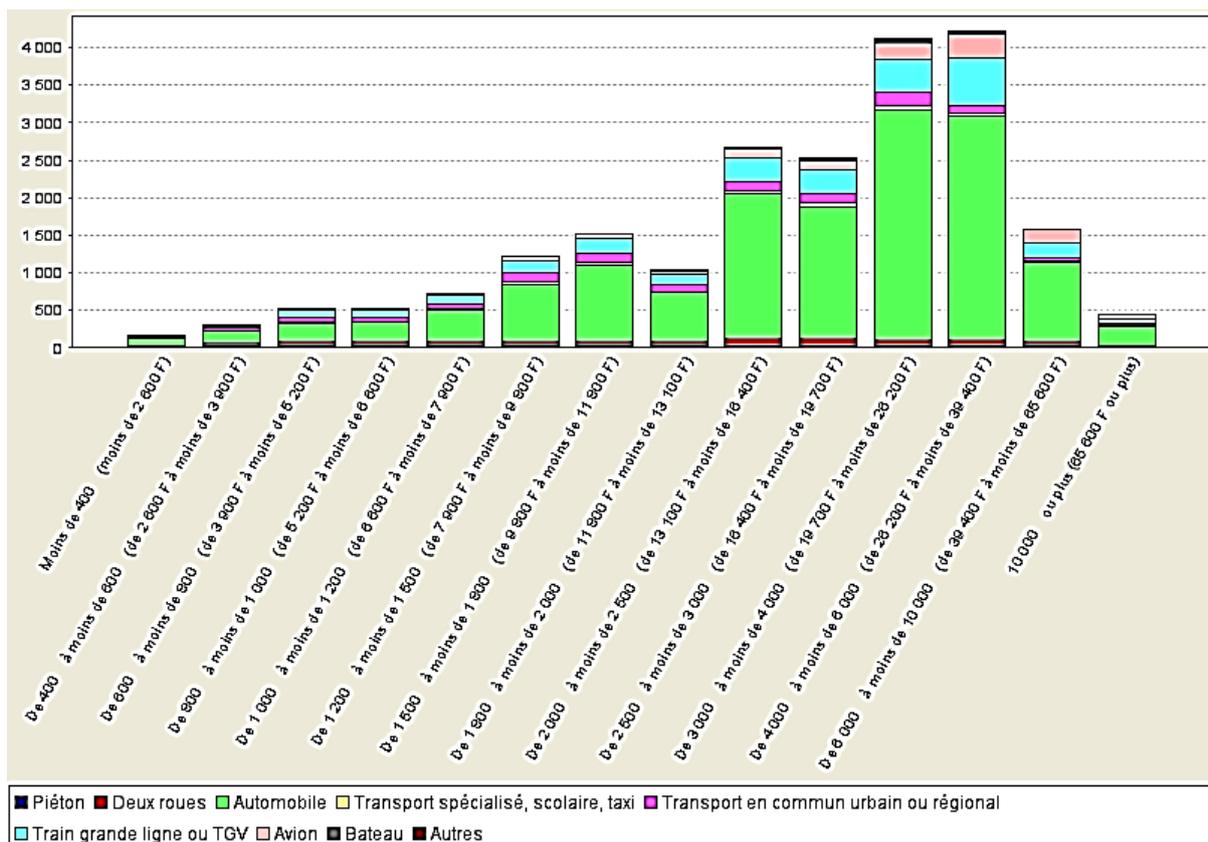


Figure 80- Nombre de déplacements par tranche du revenu mensuel net du ménage et par catégorie de mode de transport.

D’après la figure 80, nous remarquons que la tranche du revenu du ménage « De 4000 à moins de 6000 » déplace plus par rapport aux autres tranches et que la catégorie(Automobile) est le mode le plus utilisé.

La figure 81 représente le nombre de déplacements par tranche du revenu mensuel net du ménage et par catégorie de motif de déplacement.

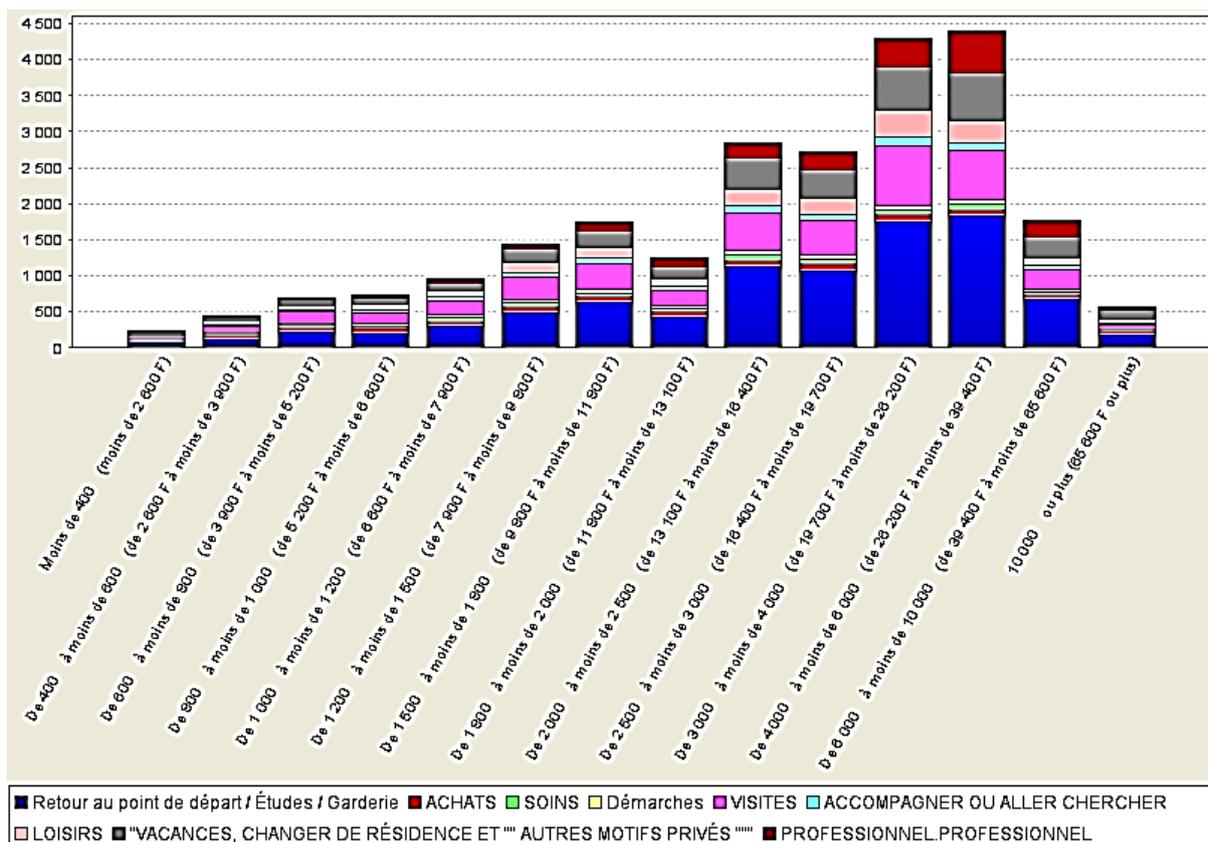


Figure 81- Nombre de déplacements par tranche du revenu mensuel net du ménage et par catégorie de motif de déplacement.

D'après la figure 81, nous remarquons que les catégories de motif « Retour au point de départ/Études/Garderie » et « VISITES » représentent les raisons principales de la plupart des déplacements.

La figure 82 représente le nombre de déplacements par tranche du revenu mensuel net du ménage et par type de carburant.

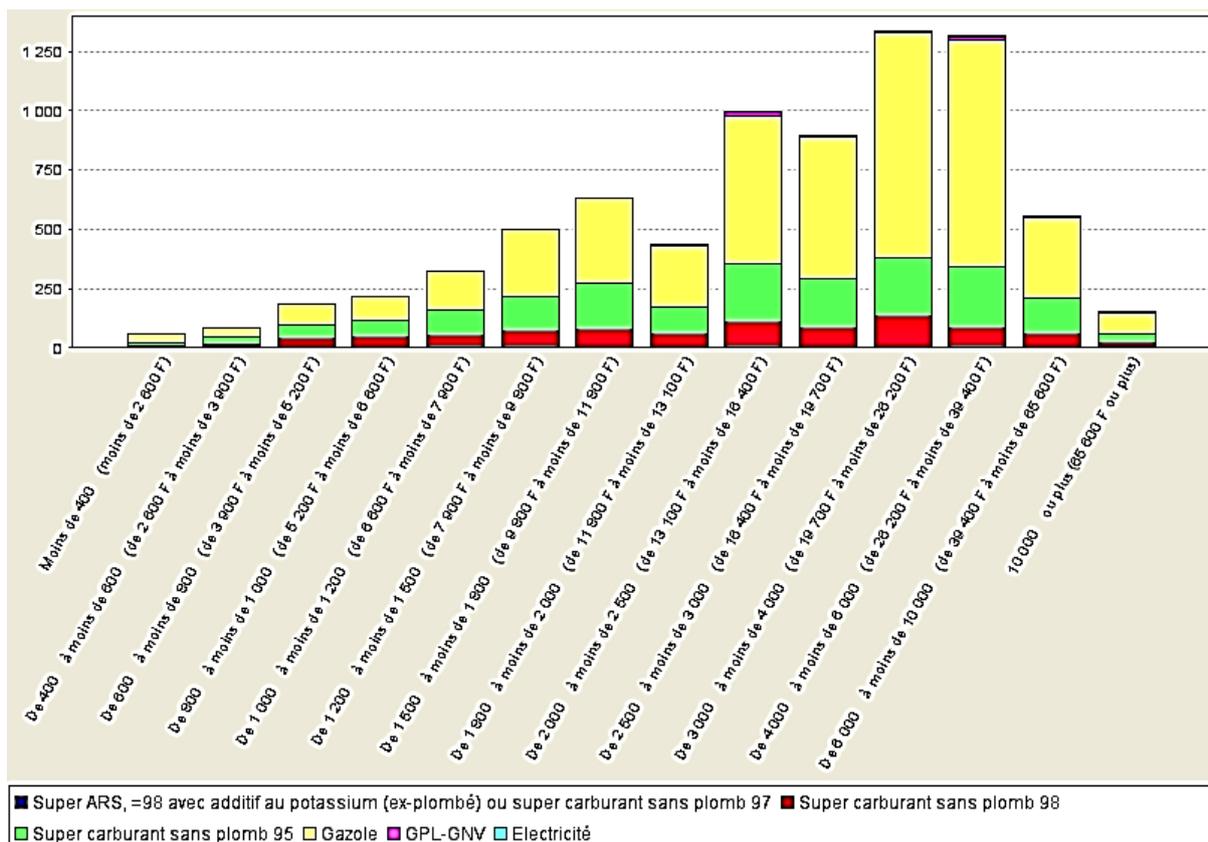


Figure 82-Nombre de déplacements par tranche du revenu mensuel net du ménage et par type de carburant.

D’après la figure 82, nous remarquons que le type de carburant « Gazole » est le carburant le plus utilisé par les différentes tranches du revenu.

5.6.2. Analyse des déplacements selon le lieu de résidence

La figure 83 représente le nombre des déplacements par région de résidence et par année.

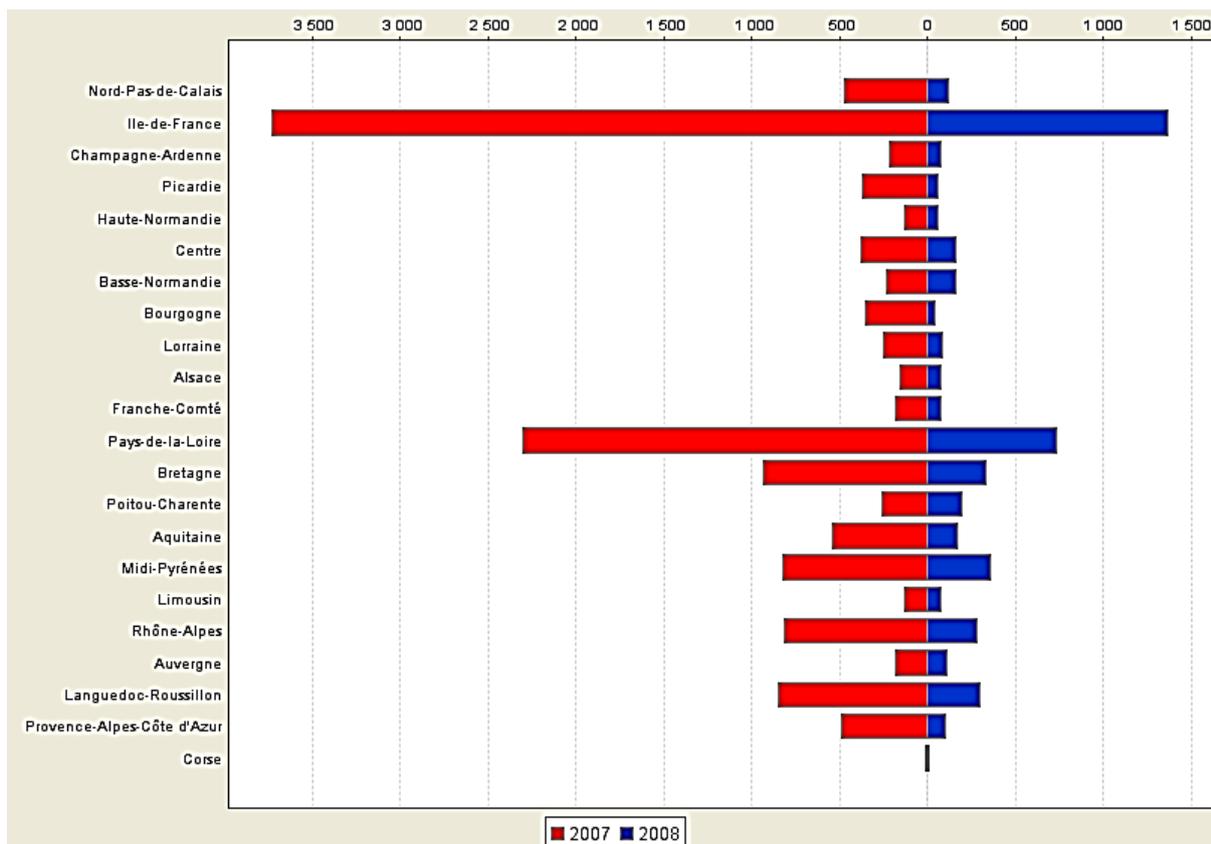


Figure 83- Nombre des déplacements par région de résidence et par année.

D'après la figure 83, nous remarquons que dans la région (Ile-de-France) les déplacements sont nombreux par rapport aux autres régions.

La figure 84 représente la consommation totale de carburant par région de résidence et par année.

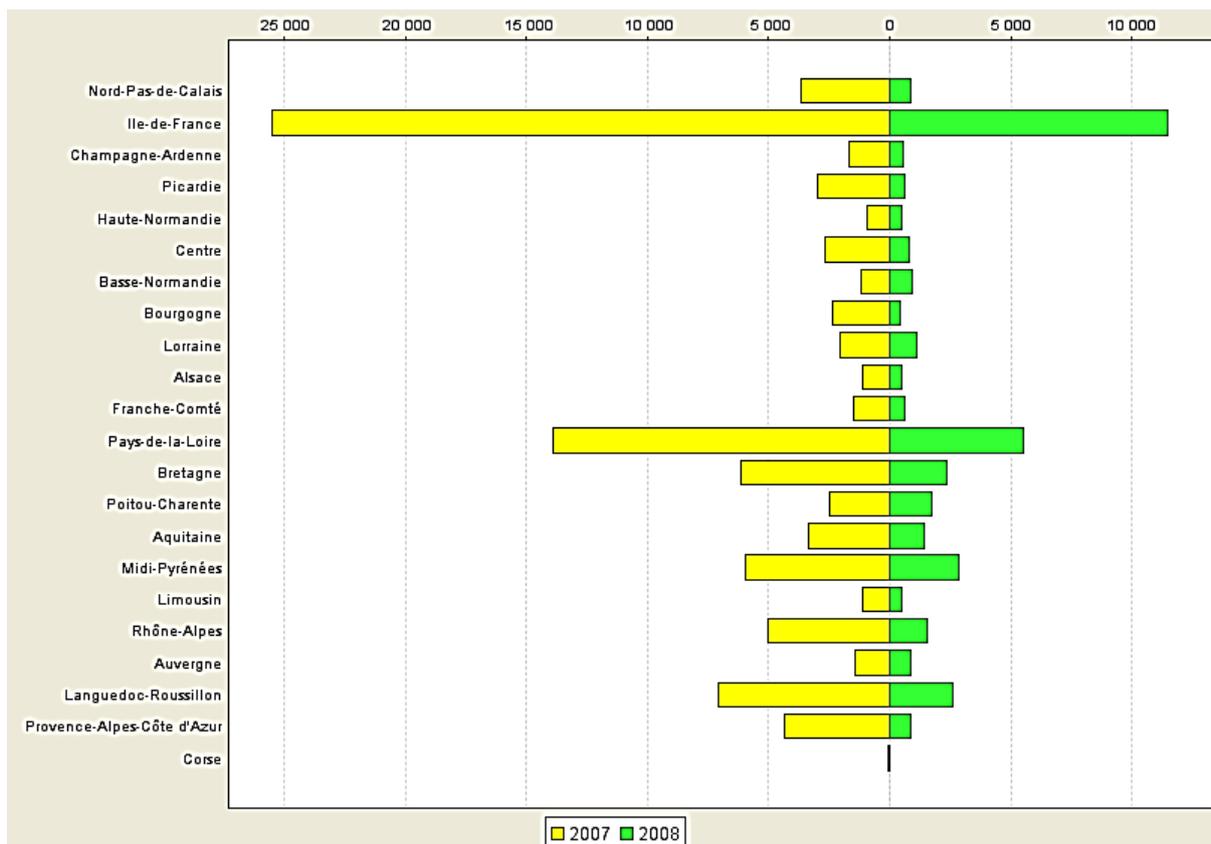


Figure 84- Consommation totale de carburant par région de résidence et par année.

D'après la figure 84, nous remarquons que dans la région (Ile-de-France) la consommation totale de carburant est élevée par rapport aux autres régions.

La figure 85 représente les émissions totales de CO2 par région de résidence et par année.

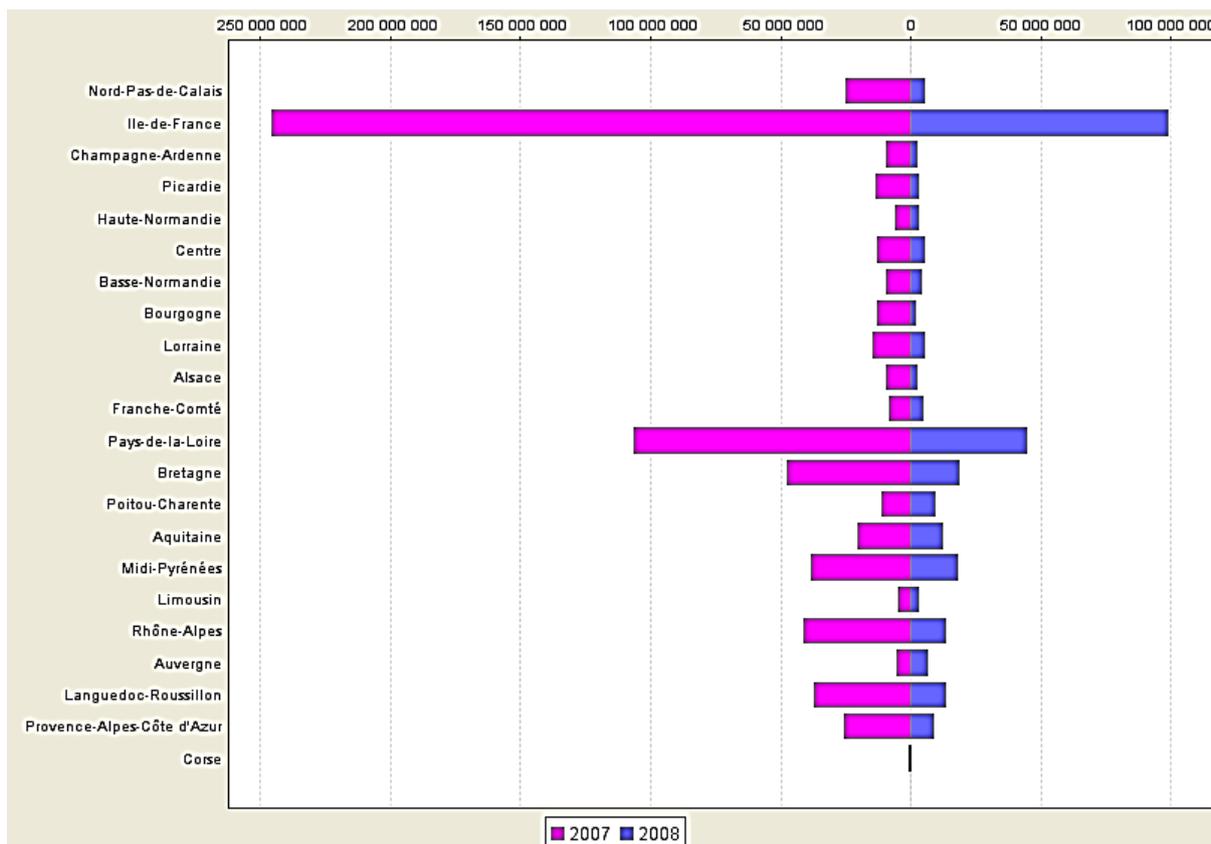


Figure 85- Emissions totales de CO2 par région de résidence et par année.

D'après la figure 85, nous remarquons que dans la région (Ile-de-France) les émissions totales de CO2 sont élevées par rapport aux autres régions.

La figure 86 représente la distance moyenne (en kilomètres) par région de résidence pour le motif d'études.

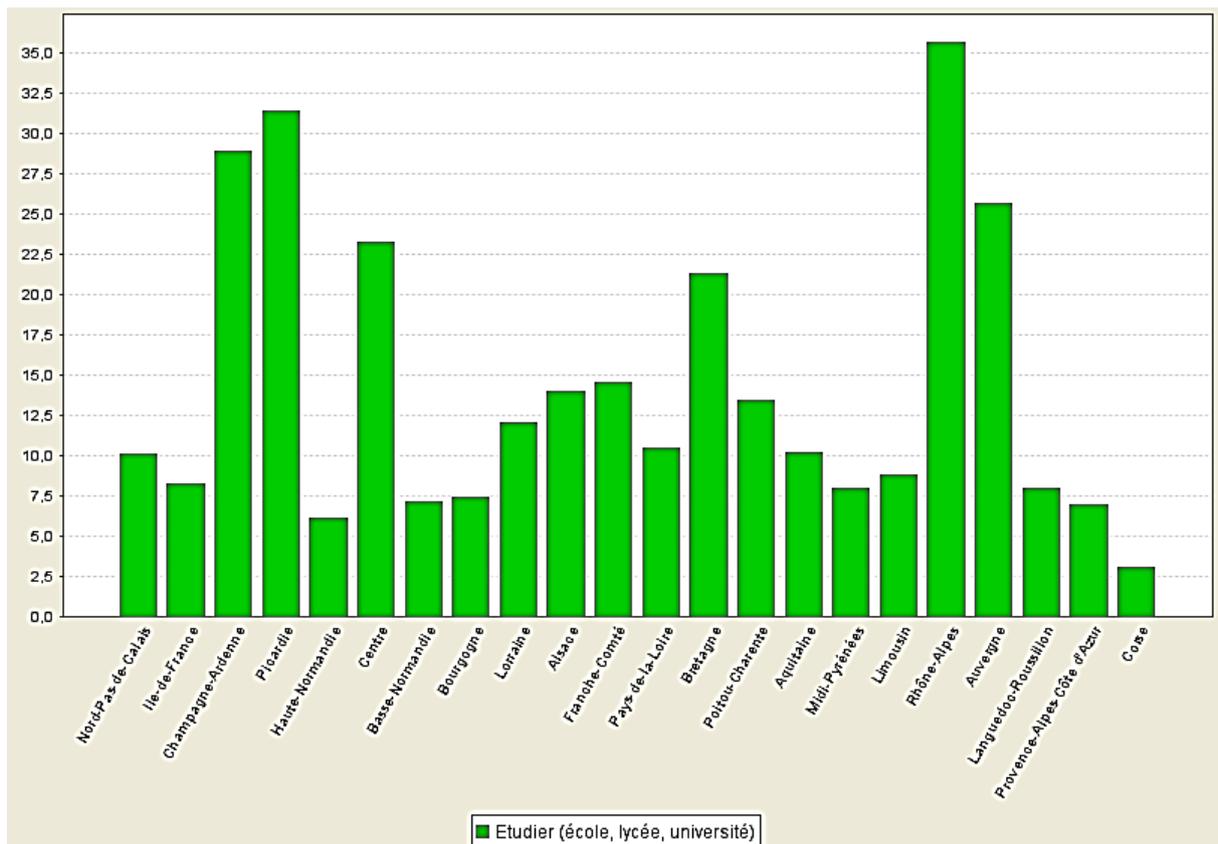


Figure 86-A1-Distance moyenne par la région de résidence pour le motif d'études.

D'après la figure 86, nous remarquons que la distance moyenne par la région (Rhône-Alpes) et pour le motif d'études est un peu grande (se dépasse 35 en kilomètres).

La figure 87 représente la distance moyenne (en kilomètres) par région de résidence pour les motifs d'achats.

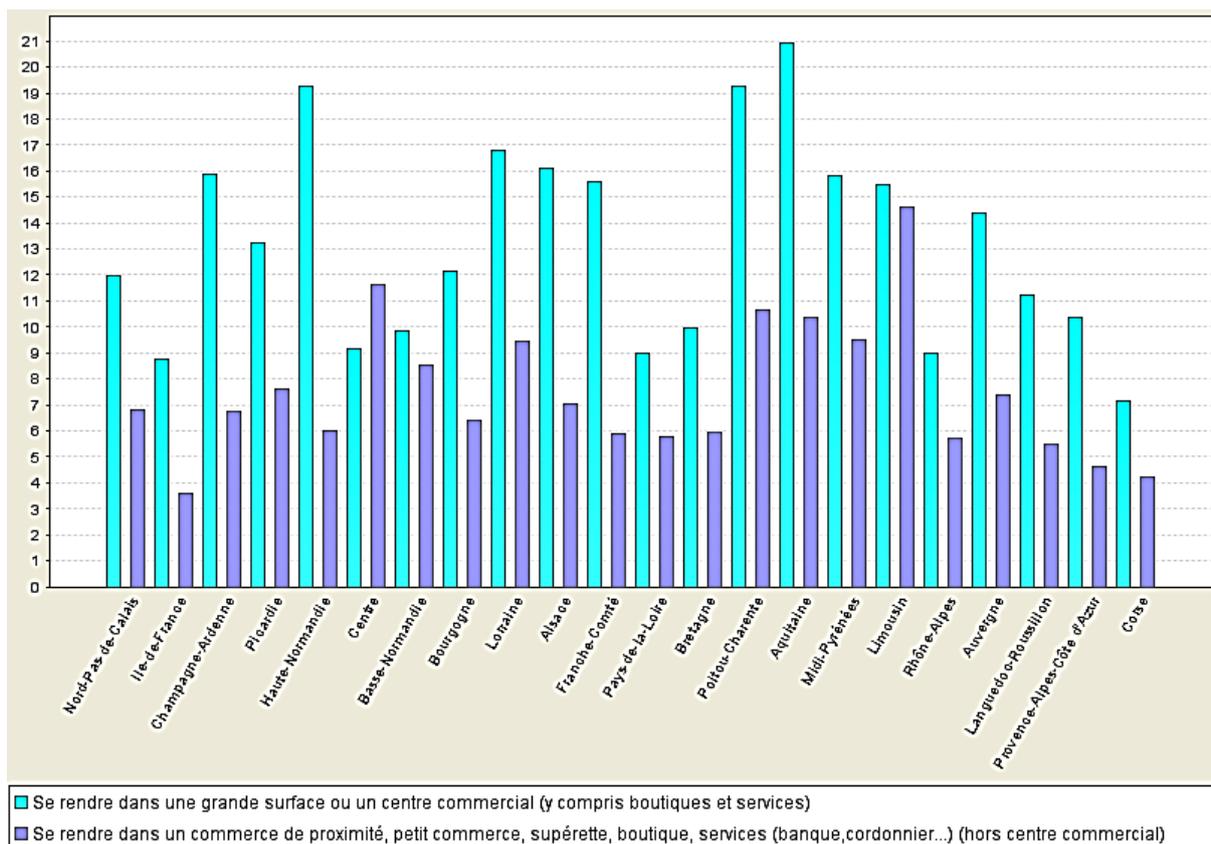


Figure 87- A2-Distance moyenne par région de résidence pour les motifs d'achats.

D'après la figure 87, nous remarquons que la distance moyenne par la région (Aquitaine) et pour le motif d'achat (Se rendre dans un commerce...) est un peu grande (se dépasse 20 en kilomètres).

La figure 88 représente distance moyenne (en kilomètres) selon la région de lieu de résidence pour le motif de soins.

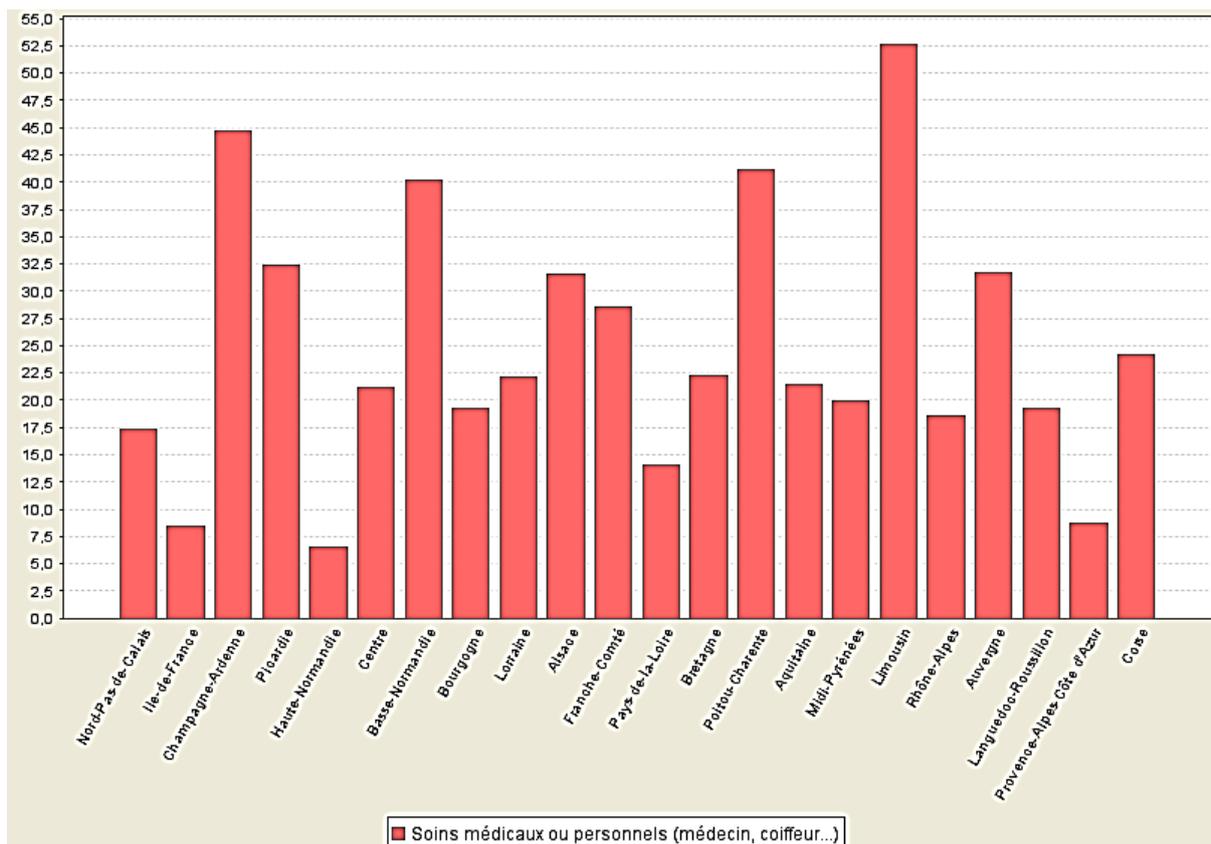


Figure 88- A3-Distance moyenne par région de résidence domicile pour le motif de soins.

D’après la figure 88, nous remarquons que la distance moyenne par la région (Limousin) et pour le motif de soins est assez grande (se dépasse 50 en kilomètres).

La figure 89 représente la distance moyenne (en kilomètres) par région de lieu de résidence pour les motifs du travail.

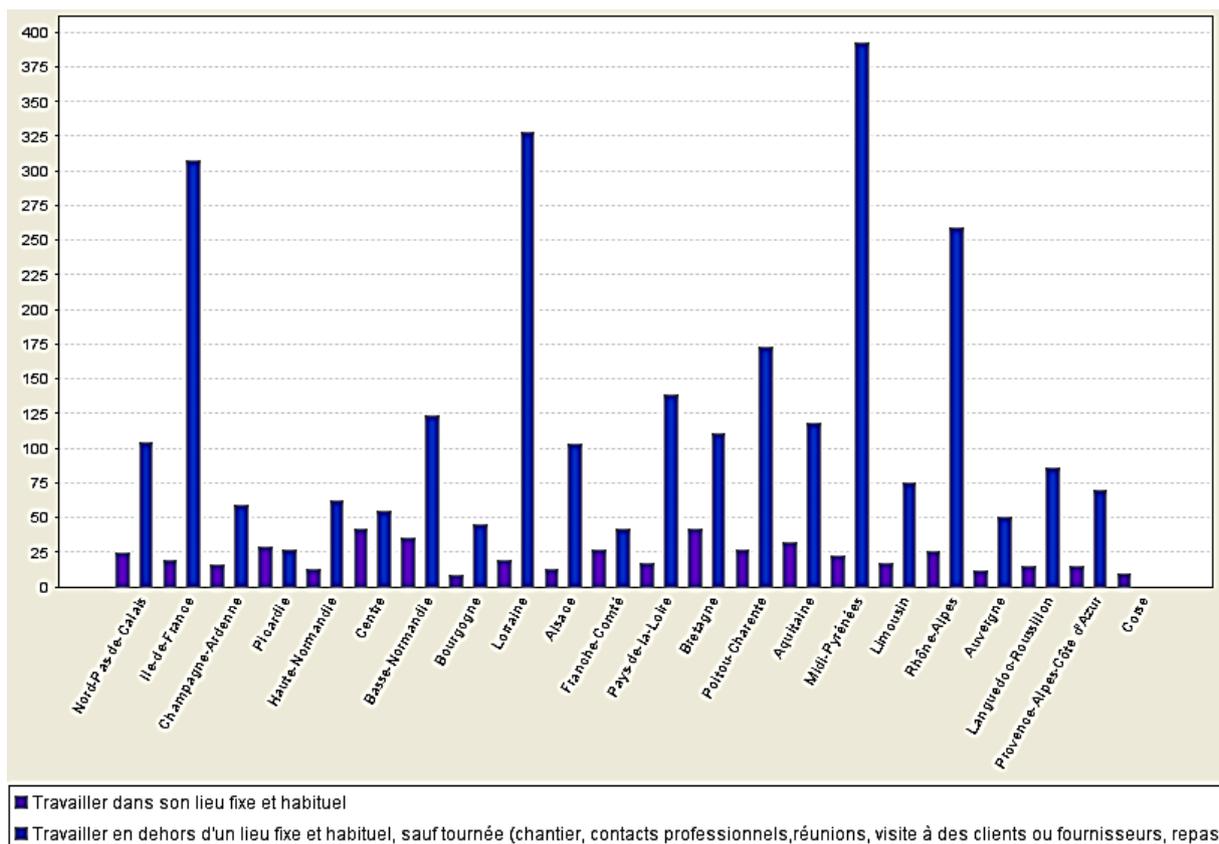


Figure 89- A4-Distance moyenne par région de résidence pour les motifs du travail.

D'après la figure 89, nous remarquons que la distance moyenne par certaines régions (ex. Midi-Pyrénées) et pour le motif du travail (Travailler en dehors d'un lieu fixe...) est assez beaucoup grande (se dépasse 100 en kilomètres).

La figure 90 représente le nombre de déplacements par région de résidence et par mode de transport (seulement les modes de transport en commun).

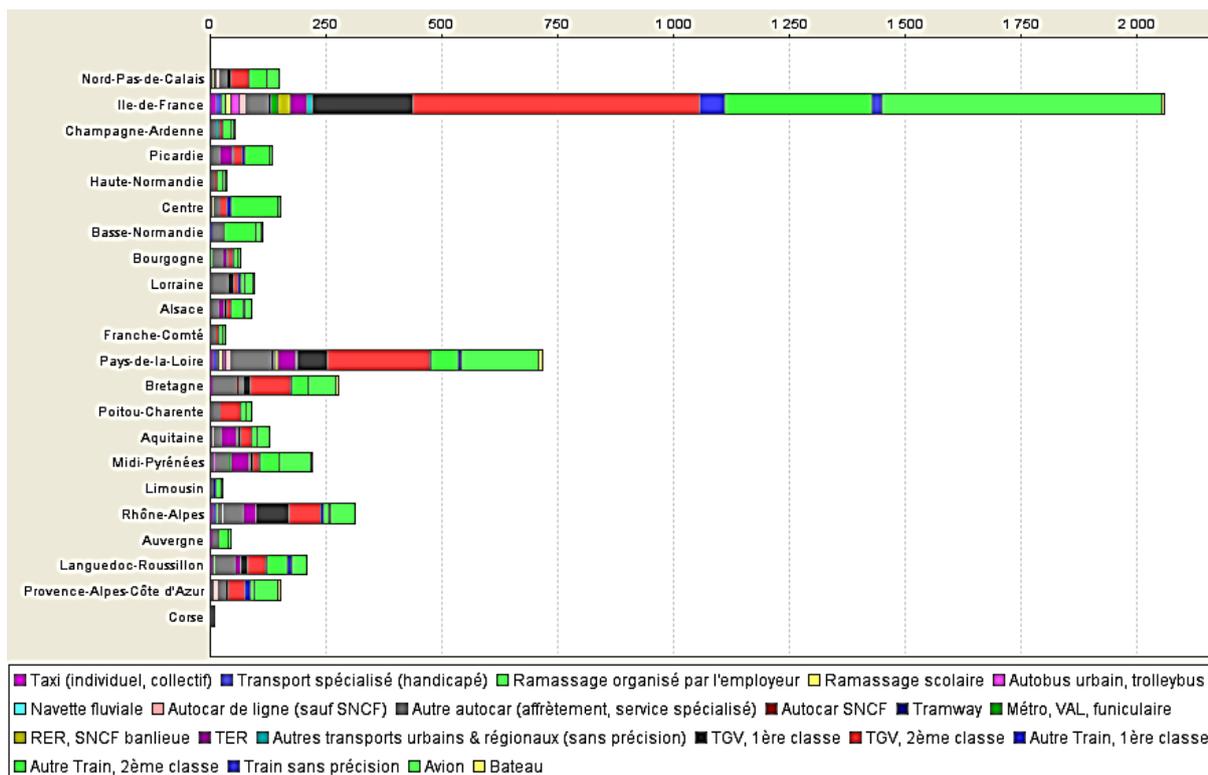


Figure 90- Nombre des déplacements par région de résidence et par mode de transport (seulement les modes de transport en commun).

D'après la figure 90, nous remarquons que dans certaines régions (ex. Corse) et (ex. Limousin) la plupart des modes en commun sont moins utilisés.

5.6.3. Analyse des déplacements selon le parc automobile

La figure 91 représente le nombre de déplacements par type de carburant et par année pour la catégorie de mode de transport (Automobile).

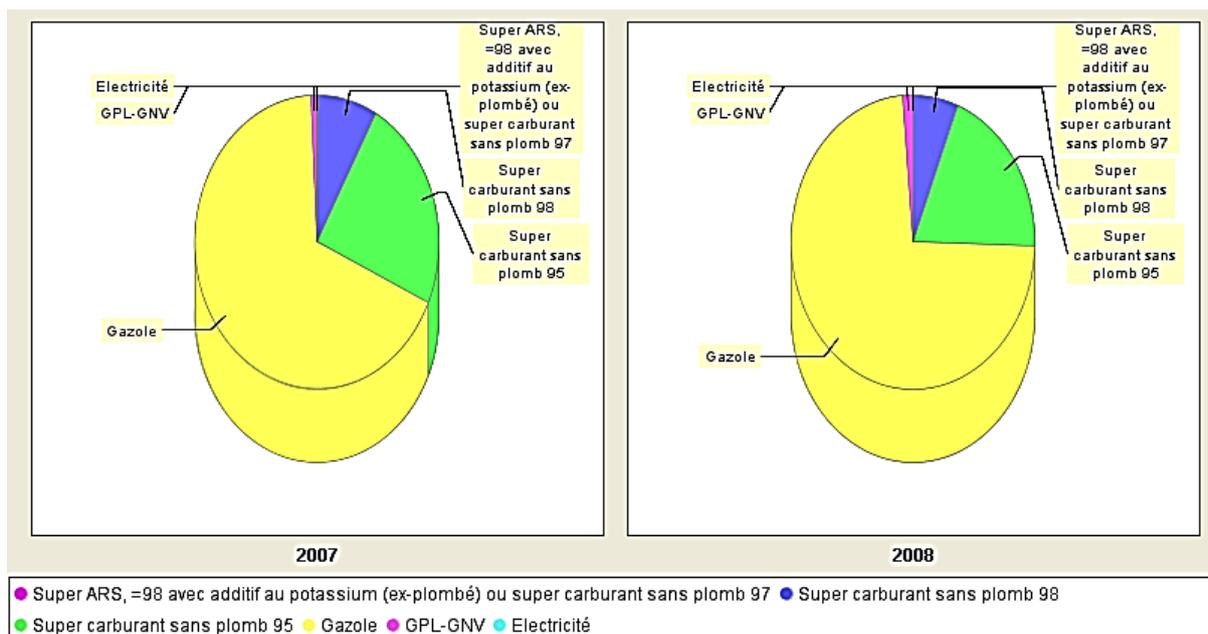


Figure 91- Nombre de déplacements par type de carburant et par année pour la catégorie de mode de transport (Automobile).

La figure 92 représente les émissions totales de CO₂ par type de carburant et par année pour la catégorie de mode de transport (Automobile).

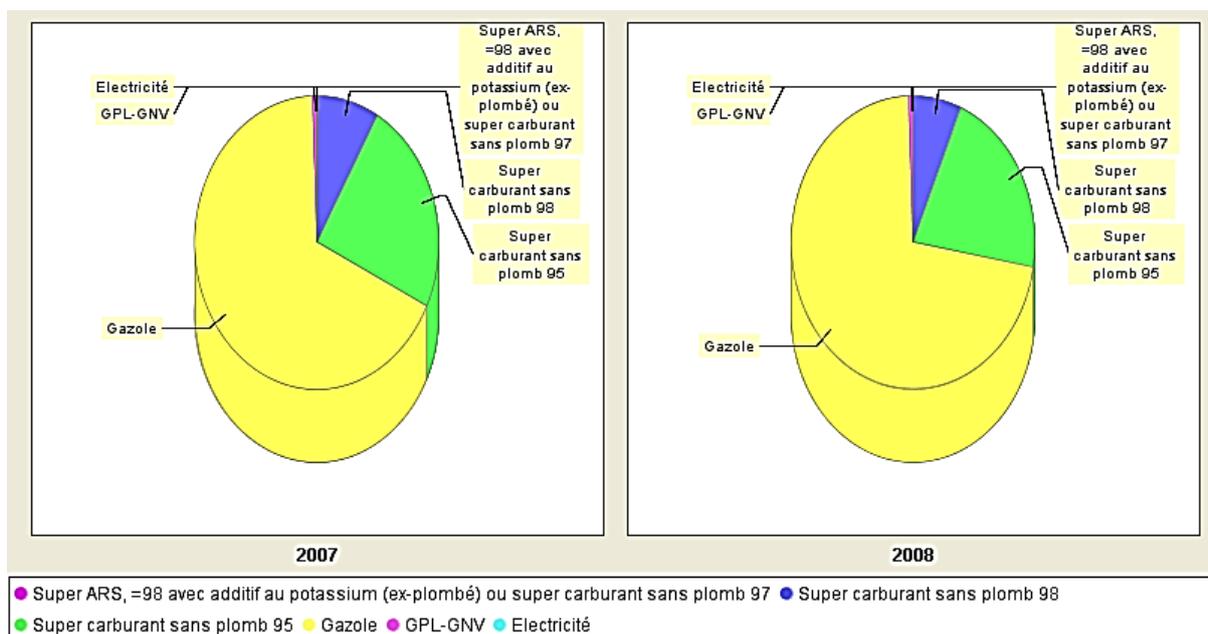


Figure 92- Emissions totales de CO₂ par type de carburant et par année pour la catégorie de mode de transport (Automobile).

D'après la figure 91, nous remarquons que le gazole est le plus utilisé dans les automobiles malgré ces émissions élevés (figure 92).

5.6.4. Analyse des flux de déplacements

La figure 93 représente le nombre de déplacements par région de départ (seulement les régions de France métropolitaine) et par région de destination (seulement les régions de France métropolitaine).

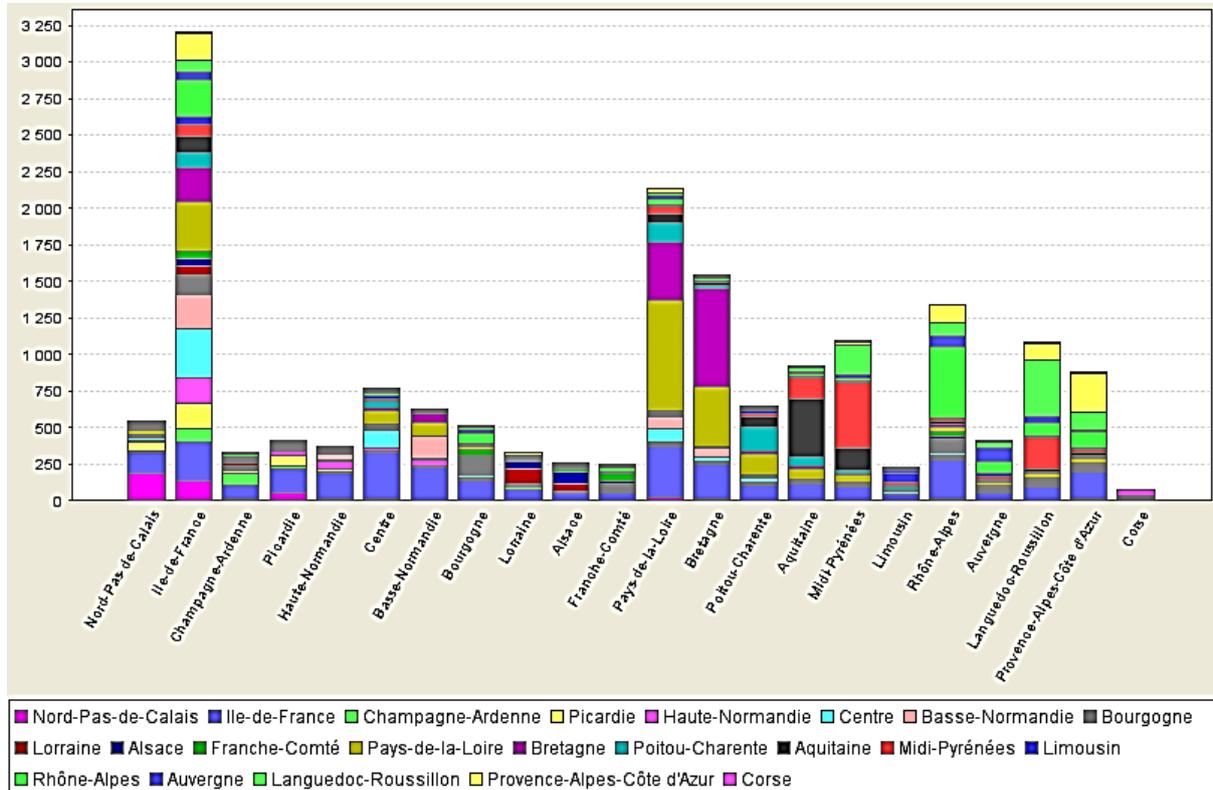


Figure 93- Nombre de déplacements par région de départ et par région de destination.

D'après la figure 93, nous remarquons que les déplacements de la région (Corse) à d'autres régions sont moins.

La figure 94 représente le nombre de déplacements par région de départ (seulement les régions de France métropolitaine) et par région de destination (seulement les régions de France métropolitaine) et par mode de transport (seulement les modes de transport en commun régional ou urbain).

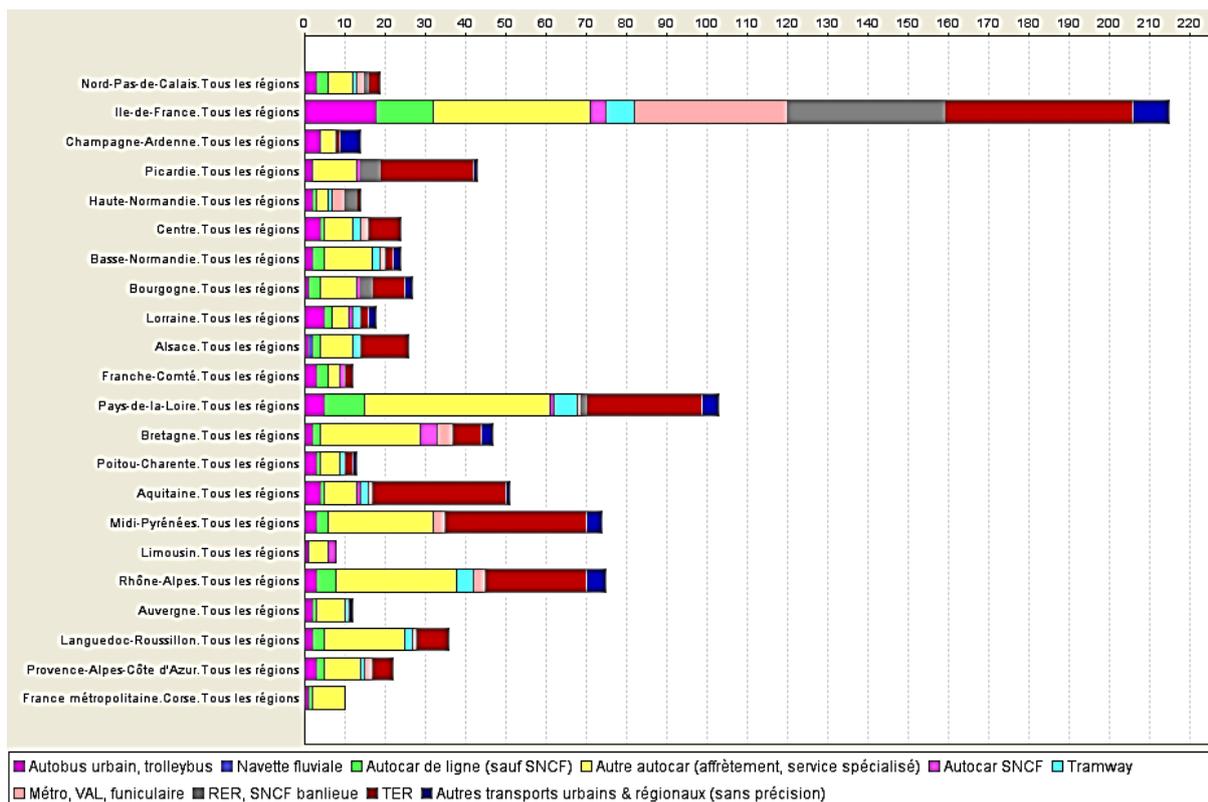


Figure 94- Nombre de déplacements par région de départ et par région de destination et par mode de transport (seulement les modes de transport en commun régional ou urbain).

D'après la figure 94, nous remarquons que dans les régions (Limousin) et (Corse) le mode (Tramway) n'est pas utilisé.

La figure 95 représente le nombre de déplacements par région de départ (seulement les régions de France métropolitaine) et par région de destination (seulement les régions de France métropolitaine) et par catégorie de motif de déplacement.

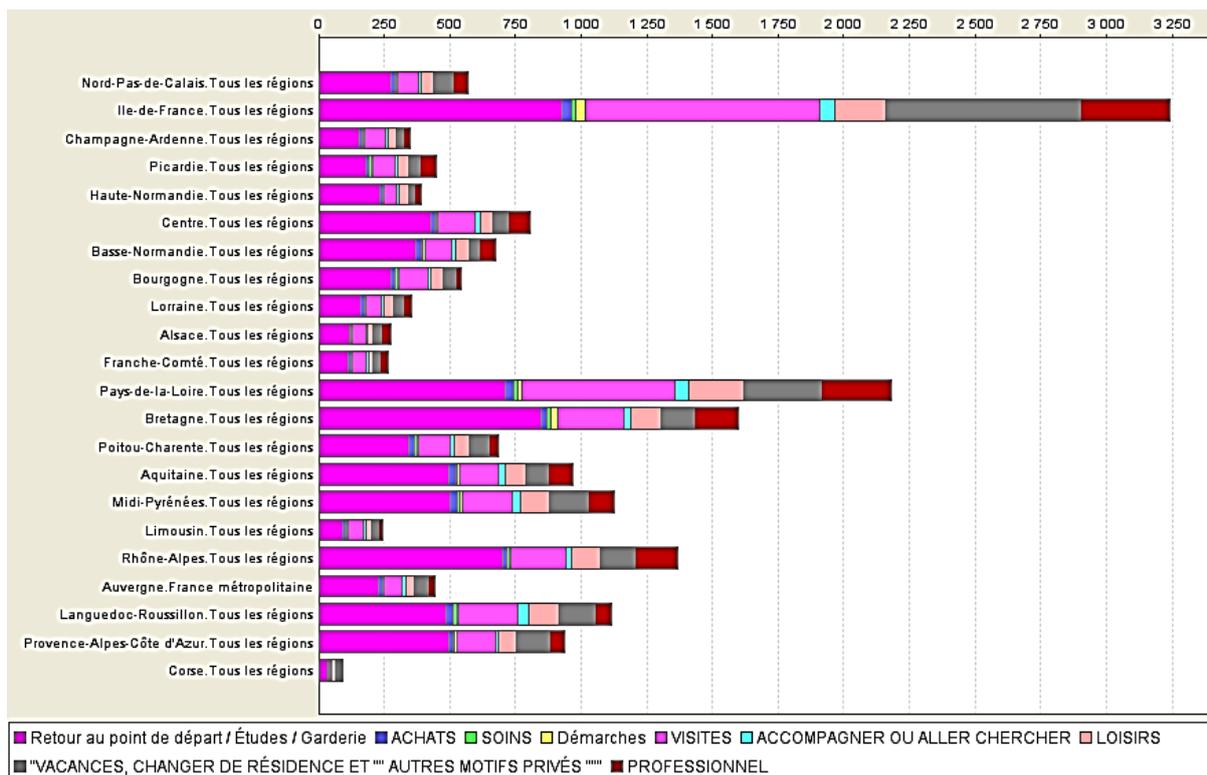


Figure 95- Nombre de déplacements par région de départ et par région de destination et par catégorie de motif de déplacement.

D'après la figure 95, nous avons remarqué que la plupart des déplacements sont d'après la région « Ile-de-France », et que les motifs « Retour au domicile » et « Visite à des parents » représentent les raisons principales entre les différents régions.

5.7. Conclusion

Au cours de dernier chapitre, nous avons implémenté l'architecture technique de notre SID à l'aide des différents logiciels (Talend Opens Studio, Mondrian et JRUBIK) avec le SGBDR PostgreSQL.

Tout d'abord nous avons commencé la création et alimentation de la base de l'entrepôt de données puis la création et l'analyse du cube multidimensionnel enfin la représentation des exemples de Reporting.

Conclusion générale

Au cours de notre travail, nous avons réalisé un système d'information décisionnel (SID) pour l'analyse des données liées à la mobilité et à la transition énergétique selon trois phases principales :

Dans la première phase "Analyse des besoins", nous avons collecté les besoins d'après une enquête nationale de transports et déplacements(ENTD) 2008 dans ' France métropolitaine'. Puis nous avons spécifié les différents besoins de notre SID. Enfin nous avons cité les sujets, les critères et les indicateurs sur lesquels est basée notre analyse.

Dans la deuxième phase "Modélisation de SID", nous avons créé le modèle multidimensionnel (schéma en étoile) de l'entrepôt de données qui se basé un seul table du fait et quinze tables des dimensions.

Dans la troisième phase "Implémentation de SID", nous avons implémenté l'architecture technique de notre SID par l'utilisation de différentes technologies et logiciels en suivant les étapes suivants :

- La création de base de l'entrepôt de données (ED) qui été implémenté par le SGBDR « PostgreSQL », puis son alimentation à l'aide de l'ETL « Talend ».
- La création du cube multidimensionnel sous forme d'un fichier XML qui déployé dans le serveur « Mondrian», puis son analyse grâce à l'application « JRUBIK » d'analyse OLAP.
- La représentation des exemples de Reporting en utilisent l'application « JRUBI ».

L'objectif de ce système étant d'aider les décideurs (ex. le ministère de transport) dans la prise des diverses décisions plus précisément concernant l'application de la transition énergétique sur les systèmes de mobilité.

Durant le déroulement de notre projet, nous avons acquis une bonne expérience dans le domaine des systèmes d'information décisionnels et nous avons bénéficié des connaissances et des compétences concernant le travaille avec certains outils décisionnels (Talend Opens Studio, PostgreSQL, Mondrian et JRUBIK).

En plus nous avons compris l'importance et la nécessité d'appliquer la transition énergétique dans certaines systèmes de mobilité pour passer à des systèmes durables et plus écologiques dont l'objectif principal est la préservation de l'environnement.

Bibliographie

- Benjamin Méreau, M. F. (2019, 07 28). *Les territoires à forte vulnérabilité énergétique ont les croissances de population les plus marquées*. Récupéré sur insee:
http://www.insee.fr/fr/themes/document.asp?reg_id=21&ref_id=19536
- Bouilil, K. (2012). *Une approche automatisée basée sur des contraintes d'intégrité définies en UML et OCL pour la vérification de la cohérence logique dans les systèmes SOLAP : Applications dans le domaine agri-environnemental*. Université Blaise Pascal-Clermont-Ferrand II: Thèse de doctorat.
- CLER. (2018). *MOBILITÉ , ÉNERGIE ET DÉVELOPPEMENT LOCAL*. publication, MONTREUIL. Consulté le 07 28, 2019, sur CLER: https://cler.org/wp-content/uploads/2018/02/mobilite2017_CLER_WEB.pdf
- Cogiterra (Éd.). (2013). *DICTIONNAIRE ENVIRONNEMENT*. Consulté le 11 22, 2019, sur ACTU ENVIRONNEMENT: https://www.actu-environnement.com/ae/dictionnaire_environnement/definition/gaz_a_effet_de_serre_ges.php4
- Comeen. (2017). *Transition énergétique : définition, enjeux et défis de la transition énergétique en France et dans le monde*. Consulté le 07 29, 2019, sur E-RSE: <https://e-rse.net/definitions/transition-energetique-definition-enjeux/#gs.ssifyc>
- CROZAT, S. (2016). *Introduction à XML : principes, syntaxe, schémas et manipulations*.
- CROZAT, S. (2016). *Introduction aux bases de données relationnelles*.
- Différence entre OLTP et OLAP*. (2018, 9 5). Consulté le 11 2019, 22, sur WayToLearnX: <https://waytolearnx.com/2018/09/difference-entre-oltp-et-olap.html>
- Écomobilité*. (2001). Consulté le 07 28, 2019, sur FUTURA PLANETE: <https://www.futura-sciences.com/planete/definitions/developpement-durable-ecomobilite-7529/>
- ESPINASSE, B. (2015, Décembre). *Entrepôts de données : Introduction au langage MDX (Multi-Dimensional eXtensions) pour l'OLAP*. Marseille: Ecole Polytechnique Universitaire de Marseille.

- ESPINASSE, B. (2017). Travail Pratique associé au cours « Entrepôts de données » Mise en œuvre de MONDRIAN et JRUBIK. Département Informatique, Marseille: Ecole Polytechnique Universitaire de Marseille.
- Ferragu, E. (2013). *MODÉLISATION DES SYSTÈMES D'INFORMATION DÉCISIONNELS*. Vuibert.
- Grim, Y. (2019). *Conception d'un entrepôt de données (Data Warehouse)*. Consulté le 07 20, 2019, sur Developpez.com: <https://grim.developpez.com/cours/businessintelligence/concepts/conception-datawarehouse>
- Inmon, W. H. (1996). The data warehouse and data mining. *Communications of the ACM*, 49-50.
- Jarke, M. M. (2003). *Fundamentals of data warehouses*. New York: Springer-Verlag.
- Kimball, R. a. (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling, 2nd Edition*. Wiley.
- Lardière, S. (2007). *PostgreSQL - Administration et exploitation d'une base de données (2ème édition)* (éd. 2ème édition). ENI Éditions.
- Malinowski, E. a. (2008). *Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications*. Springer.
- Orfeuil. (1996). La mobilité urbaine. Lyon, Université Lumière Lyon 2. Consulté le 07 24, 2019, sur [www.http://theses.univ-lyon2.fr](http://theses.univ-lyon2.fr): http://theses.univ-lyon2.fr/documents/getpart.php?id=lyon2.2000.petiot_r&part=29377
- Ortar, P. L. (2014). *La transition énergétique : vrais enjeux, faux départs ?* Consulté le 07 23, 2019, sur SHS Web of Conferences: https://www.shs-conferences.org/articles/shsconf/pdf/2014/06/shsconf_trans2013_01001.pdf
- Rouse, M. (2019). *OLAP (traitement analytique en ligne)*. Retrieved 07 20, 2019, from TechTarget: <https://searchdatamanagement.techtarget.com/definition/OLAP>

SGBD et Datawarehouse. (s.d.). Consulté le 07 22, 2019, sur Institute of electronics and computer science Gaspard-Monge (IGM): <http://www-igm.univ-mlv.fr/~dr/XPOSE2005/entrepot/sghd.html>

Smile. (2013). *DÉCISIONNEL - Les meilleures solutions open source.*