POLYPUBLIE

Méthodologie d'analyse et de suivi d'un système de transport par taxi

Annick Lacombe

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

Date	2016
Type	Mémoire ou thèse / Dissertation or Thesis Mémoire de maîtrise/Master's thesis
URL de PolyPublie PolyPublie URL	https://publications.polymtl.ca/2261/
Conditions d'utilisation Terms of use	© Tous droits réservés / All rights reserved
Référence Citation	Lacombe, A. (2016). Méthodologie d'analyse et de suivi d'un système de transport par taxi [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/2261/

Détails de la thèse
Thesis details

Directeurs de recherche Advisors	Catherine Morency
Programme Program	Génie civil

Ce fichier a été téléchargé à partir de PolyPublie, le dépôt institutionnel de Polytechnique Montréal
This file has been downloaded from PolyPublie, the institutional repository of Polytechnique Montréal
https://publications.polymtl.ca

UNIVERSITÉ DE MONTRÉAL

MÉTHODOLOGIE D'ANALYSE ET DE SUIVI D'UN SYSTÈME DE TRANSPORT PAR TAXI

ANNICK LACOMBE

DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET DES MINES

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION

DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES

(GÉNIE CIVIL)

AOÛT 2016

© Annick Lacombe, 2016.

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

MÉTHODOLOGIE D'ANALYSE ET DE SUIVI D'UN SYSTÈME DE TRANSPORT PAR TAXI

présenté par : LACOMBE Annick

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. SAUNIER Nicolas, Ph. D., président

Mme MORENCY Catherine, Ph. D., membre et directrice de recherche

M. TRÉPANIER Martin, Ph. D., membre

REMERCIEMENTS

Je tiens à remercier Catherine Morency, ma directrice de recherche, qui a cru en moi depuis le début, m'a fourni de précieux conseils, a été une guide exceptionnelle et m'a offert plusieurs opportunités tout au long de ma maîtrise. Elle a fait croître mon intérêt pour le domaine des transports dès ses premiers enseignements au baccalauréat et m'a transmis sa passion pour cette discipline.

L'environnement de Polytechnique Montréal m'a permis de me développer tant sur le plan académique que professionnel et personnel. C'est pourquoi je désire remercier les gens hors pair que j'y ai côtoyés. Philippe Fortin a été d'un support moral formidable et a su rassurer les doutes que j'ai pu ressentir à propos de mon projet de recherche. Nous avons évolué ensemble ainsi que travaillé de longues journées au bureau, et pour tout cela je lui suis reconnaissante. Je remercie profondément Hubert Verreault pour sa présence à toutes les rencontres avec les partenaires et ses nombreux conseils méthodologiques. Enfin, tous les échanges inspirants et les activités passionnantes que j'ai vécus n'auraient pas été possibles sans mes collègues et amis Joanie Smith, Alexis Frappier, Gabriel Sicotte, Jean-Simon Bourdeau, Marc-André Tessier et Philippe Gaudette.

Ma famille (Françoise, Pierre, Claudie, Mireille, Geneviève et Étienne) a su être présente et m'apporter un support inestimable dans les meilleurs et les pires moments tout au long de cette aventure. La remarquable compréhension de Mathieu, mon amoureux, qui se traduisit par toutes les petites attentions (et les bons soupers) dont il m'a gratifiée, particulièrement dans les derniers miles, doit indéniablement être soulignée. Ma reconnaissance va aussi à mes amis proches qui ont démontré leur intérêt envers mon domaine d'étude, qui ont trouvé les moyens de me convaincre de lâcher-prise à certains moments pour participer à leurs activités et qui m'ont fait part de leurs encouragements à maintes reprises.

Finalement, parce que sans leur soutien je n'aurais pas pu autant développer et approfondir mon projet, je remercie le Bureau du Taxi de Montréal et Revenu Québec pour le financement de mon projet ainsi que Taxi Diamond pour l'accès aux données qu'il nous a fournies.

RÉSUMÉ

Le taxi est devenu, au cours des dernières années, un sujet actuel de réflexion, au cœur des enjeux de plusieurs villes. En effet, cette industrie a très peu évolué depuis de nombreuses décennies et se retrouve maintenant confrontée à l'arrivée de modes alternatifs à l'automobile. Elle doit maintenant se questionner sur ses méthodes de gestion de la flotte ainsi que sur sa position dans le système de transport. Ce projet de recherche vise à développer des indicateurs de performance et de suivi sur l'offre et la demande de déplacements par taxi grâce notamment à des données GPS. Cinq objectifs secondaires spécifient le but principal :

1. Évaluer le potentiel analytique des données provenant des systèmes embarqués (GPS) et d'autres sources de données administratives ;
2. Développer une méthode d'identification des objets du système (chauffeurs, taxis, courses et postes d'attente) basée sur les différentes bases de données intégrées ;
3. Proposer un ensemble d'indicateurs de performance et de suivi basés sur les objets identifiés ;
4. Développer un modèle de génération de la demande en taxi soulignant les principaux facteurs d'influence ;
5. Contribuer au développement d'un outil de consultation permettant le suivi et l'analyse des indicateurs sur les activités des taxis.

Premièrement, ce mémoire présente une revue de littérature axée sur le service de transport par taxi. Elle fait d'abord état de la définition de ce type de transport ainsi que d'un historique de son évolution à travers le temps. Puis, un survol du potentiel que pourraient apporter les données GPS à cette industrie est réalisé. Ensuite, le corps principal de la revue de littérature sur le taxi se constitue d'une description complète de son industrie, comprenant les caractéristiques de l'offre (règlementation et répartition de la flotte) et de la demande (types d'utilisateurs, facteurs d'influence et prédiction de la demande) ainsi que le suivi des activités. Enfin, le rôle du taxi dans la mobilité quotidienne des individus est exposé.

Deuxièmement, un portrait de l'environnement dans lequel se réalise l'étude de cas est tracé. Les principales infrastructures de transport, les centres d'activités et la mobilité de la ville de Montréal sont décrits. De plus, tout ce qui a trait à l'industrie du taxi de la ville est présenté, passant par sa

structure organisationnelle, ses différents services et ses aspects réglementaires. Ainsi, cette partie permet d'acquérir les connaissances nécessaires pour bien mettre en contexte l'étude de cas.

Troisièmement, la méthodologie liste et détaille les étapes techniques pour atteindre l'objectif principal de l'étude. Toutes les données utilisées dans cette recherche sont décrites, de même que le format et le type du système d'information sélectionné pour intégrer les objets dérivés de ces données. Le traitement de standardisation, de nettoyage et d'importation des données dans la base de données est aussi présenté. Les étapes de calcul des indicateurs sont également développées selon les quatre objets principaux du système (chauffeur, taxi, course et poste d'attente). La méthodologie permet donc de définir les activités d'une flotte de taxis, qui est un fondement dans la compréhension du rôle du taxi dans la mobilité quotidienne.

Quatrièmement, cette méthodologie est appliquée à la ville de Montréal pour l'année 2015. Les données GPS de Taxi Diamond, un intermédiaire en service de taxis opérant 25 % de la flotte dans l'agglomération 11 (centre de l'île), sont utilisées. Seuls les véhicules possédant un permis dans cette agglomération sont compris dans les calculs, et ce, pour la période de l'année 2015. Les indicateurs produits sont pondérés au niveau de la flotte et catégorisés selon leur liaison à l'offre et la demande. Les constats généraux sur les activités de la flotte montrent que le début et la fin des quarts de travail sont concentrés respectivement le matin et en début de soirée. La distance sans client que parcourt un taxi en service est en moyenne 60 % de sa distance totale et pour sa durée, cela atteint 72 %. Le Centre-Ville ainsi que le Plateau-Mont-Royal présentent en tout temps une forte demande, qui devient très concentrée à ces endroits le vendredi et samedi soir. Une course moyenne pour l'année est de 4,39 km et 11,4 minutes.

Finalement, afin de cibler les facteurs influençant le plus la demande en taxi à Montréal, un modèle de régression linéaire des origines et destinations des courses est estimé. Le nombre de courses pour chaque secteur de recensement et jour d'octobre 2014 est utilisé pour les taxis de l'agglomération 11. Une analyse de corrélation sert à sélectionner les variables et une analyse de la signification statistique accompagnée de celle du coefficient de détermination sert à choisir le modèle final. Les résultats montrent que les variables sur la population (âge et revenu) ainsi que celles sur l'utilisation du sol (zones de bureaux, de commerces et accessibilité au transport en commun) sont les principaux facteurs qui affectent la demande.

Pour conclure, ce projet développe une méthodologie pour le suivi des activités de transport par taxi. Les indicateurs, développés grâce aux données GPS, montrent les tendances générales de ce mode de transport. Grâce au système d'information élaboré, l'automatisation de tous les processus de la méthodologie peut être réalisée sans complexité afin de produire un outil autonome d'analyse des activités.

ABSTRACT

In the last years, Taxi has become an important topic of reflection, being at the heart of the transportation challenges in several cities. Indeed, this industry has remained unchanged for many decades, and now finds itself confronted with the arrival of innovative transportation alternatives. This industry must question its current management strategies and clearly assessed its role in the transportation system. This research project aims to develop performance and monitoring indicators on taxi supply and demand with GPS data. Five specific objectives complete the main one :

1. Assess the analytical potential of data from on-board systems (GPS) and other administrative data sources;
2. Develop a method to identify the objects of the system (driver, vehicle, ride and taxi stand) based on the different integrated databases;
3. Propose a set of performance and monitoring indicators based on these objects;
4. Develop a taxi trip generation model to highlight the main influential factors;
5. Contribute to the development of a consultation tool to monitor and analyse indicators on taxis activities.

First, this document presents a literature review focused on taxi services. It initially defines this type of transportation as well as a history of its evolution through time. Then, an overview of the potential that could provide GPS data for this industry is accomplished. The main body of the literature review on taxi is a complete description of its industry, including the characterization of supply (regulation and fleet dispatch) and demand (user types, influential factors and prediction of demand) as well as the monitoring of its activities. Finally, the role of taxis in the daily mobility is exposed.

Second, a portrait of the environment in which is conducted the case study is drawn. The main transportation infrastructure, the activity centers and the travel behaviours in Montreal are described. Moreover, everything that relates to the taxi industry in this city is presented, through its organizational structure, its various services and its regulatory aspects. Thus, this chapter provides the necessary knowledge to contextualize the case study.

Thirdly, the methodology and list details the technical steps to achieve the main objective of the study. All data used in this research are described, as well as the selected type and layout of the information system to integrate the objects derived from the data. The method used to standardize, clean and import the data in the database is also presented. Then, the steps to estimate the indicators are defined according to the four main objects of the system (driver, taxi, ride and taxi stand). The methodology allows to define the taxi fleet activities, which is a foundation in understanding the role of the taxi in daily mobility.

Fourth, this methodology is applied to the city of Montreal for the year 2015. The GPS data from Taxi Diamond, a taxi company operating 25% of the fleet in agglomeration 11 (center of the island), is used. Only vehicles with a licence in this agglomeration for 2015 are included in the calculations. The indicators are weighted in function of the entire fleet in the zone and categorized according to their connection to supply and demand. The general findings on the fleet activities show that the beginning and end of shifts are respectively concentrated in the morning and early in the evening. The unoccupied distance traveled by a taxi in operation is on average 60% of the total distance and for duration, it reaches 72%. Downtown and Plateau-Mont-Royal have, at all times, a strong demand, which becomes highly concentrated in these areas on Friday and Saturday night. An average ride for 2015 is about 4.39 km and 11.4 minutes.

Finally, to identify factors that most influence taxi demand in Montreal, a linear regression model for pickups and drop-offs is estimated. The number of rides for each census tract and day of October 2014 are used for the agglomeration 11. A correlation analysis is performed to select variables and an analysis of the statistical significance and the coefficient of determination is made to select the final model. The results show that variables on population (age and income) as well as land use (office areas, shops and accessibility to transit) are the main factors affecting demand.

In conclusion, this project develops a methodology for monitoring transportation activities from a taxi fleet. The indicators developed using GPS data show general figures of this mode of transportation. Thanks to the information system designed, the automatization of all processes can be performed without complexities in order to produce an autonomous tool for analysing taxi activities.

TABLE DES MATIÈRES

REMERCIEMENTS III

RÉSUMÉ IV

ABSTRACT VII

TABLE DES MATIÈRES IX

LISTE DES TABLEAUX XIII

LISTE DES FIGURES XV

LISTE DES SIGLES ET ABRÉVIATIONS XVII

LISTE DES ANNEXES XVIII

CHAPITRE 1 INTRODUCTION 1

1.1 Problématique 2

1.2 Objectifs 3

1.3 Structure du mémoire 3

CHAPITRE 2 REVUE DE LITTÉRATURE 5

2.1 Définition du taxi 5

2.2 Historique 8

2.3 Potentiel des données GPS 10

2.4 Industrie du taxi 12

2.4.1 Offre 12

2.4.1.1 Règlementation 12

2.4.1.2 Dérèglementation 13

2.4.1.3 Répartition de la flotte 14

2.4.2 Demande 15

2.4.2.1 Types d'utilisateurs 16

2.4.2.2 Facteurs d'influence de la demande 16
2.4.2.3 Prédiction de la demande 18
2.4.3 Suivi analytique 18
2.5 Rôle du taxi 21
CHAPITRE 3 MISE EN CONTEXTE 23
3.1 Description de la ville 23
3.1.1 Réseau de transport 24
3.1.2 Déplacements 25
3.2 Industrie du taxi 26
3.2.1 Structure organisationnelle 26
3.2.2 Services de taxi 27
3.2.3 Règlementation 28
CHAPITRE 4 MÉTHODOLOGIE 31
4.1 Données 33
4.1.1 Données de taxi 33
4.1.2 Données explicatives 35
4.1.3 Autres données 36
4.2 Système d'information 37
4.2.1 Type de base de données 37
4.2.2 Objets et entités 38
4.2.2.1 Objets du système de transport par taxi 39
4.2.2.2 Entités de la base de données 40
4.2.3 Schéma de la base de données 41
4.3 Traitement des données 43

4.3.1 Standardisation 43
4.3.2 Importation et nettoyage des données 45
4.3.3 Attribution du permis 48
4.4 Indicateurs 49
4.4.1 Pondération 50
4.4.2 Formes de visualisation 51
4.4.3 Objets d'étude 52
4.4.3.1 Chauffeur 52
4.4.3.2 Taxi 54
4.4.3.3 Course 56
4.4.3.4 Poste d'attente 58
CHAPITRE 5 DÉMONSTRATION DES INDICATEURS 61
5.1 Données utilisées 62
5.2 Représentativité 63
5.3 Indicateurs 65
5.3.1 Offre 65
5.3.2 Demande 77
5.4 Tableau de bord 93
CHAPITRE 6 MODÉLISATION DE LA DEMANDE 100
6.1.1 Données utilisées 100
6.1.2 Méthodologie 101
6.1.3 Résultats et analyses 104
CHAPITRE 7 CONCLUSION 108
7.1 Synthèse de la recherche 108

7.2 Contributions 110
7.3 Limitations 111
7.4 Perspectives 113
BIBLIOGRAPHIE 116
ANNEXES 122

LISTE DES TABLEAUX

Tableau 2-1 : Modes de transport parallèle public du véhicule privé au transport en commun 7

Tableau 2-2 : Revue des indicateurs 19

Tableau 4-1 : Information sur les courses 34

Tableau 5-1 : Traitement des données de courses pour l'année 2015 62

Tableau 5-2 : Traitement des données GPS pour l'année 2015 62

Tableau 5-3 : Attribution du permis aux véhicules de 2015 63

Tableau 5-4 : Facteurs d'expansion pour l'année 2015 65

Tableau 5-5 : Statistiques descriptives sur l'offre du service de taxi 66

Tableau 5-6 : Statistiques descriptives sur la demande du service de taxi 78

Tableau 6-1 : Description des variables 104

Tableau 6-2 : Modèles de régression pour l'origine et la destination des courses 105

Tableau A-1 : Table des points GPS 122

Tableau A-2 : Table des courses 123

Tableau B-1 : Table des intermédiaires 124

Tableau B-2 : Table des chauffeurs 124

Tableau B-3 : Table des véhicules 124

Tableau B-4 : Table des points GPS 125

Tableau B-5 : Table des courses 126

Tableau B-6 : Table des secteurs 127

Tableau B-7 : Table des arrondissements 127

Tableau B-8 : Table des agglomérations 127

Tableau B-9 : Table de l'utilisation du sol 128

Tableau B-10 : Table des postes d'attente de taxis 129

Tableau B-11 : Table du calendrier 130

Tableau B-12 : Table de la météo par jour 130

Tableau B-13 : Table de la météo aux 15 minutes..... 131

LISTE DES FIGURES

Figure 2-1 : Ligne du temps des évènements majeurs du taxi 8

Figure 3-1 : Population de la région métropolitaine de Montréal 23

Figure 3-2 : Principaux lieux et axes routiers de l'île de Montréal 24

Figure 3-3 : Parts modales des déplacements sur l'île de Montréal 25

Figure 3-4 : Agglomérations et nombre de permis de taxi sur l'île de Montréal (Bureau du Taxi de Montréal, 2015) 29

Figure 4-1 : Schéma méthodologique 32

Figure 4-2 : Types de courses (a) commandées et (b) hélées 40

Figure 4-3 : Diagramme relationnel de la base de données de taxis 42

Figure 4-4 : Étapes d'importation et de nettoyage des données de courses brutes 46

Figure 4-5 : Étapes d'importation et de nettoyage des données GPS brutes 47

Figure 4-6 : Formes de visualisation 51

Figure 5-1 : Construction des indicateurs et segmentation pour leur analyse 61

Figure 5-2 : Débuts des quarts de travail pour une semaine moyenne selon les saisons (incrément de 30 min) 67

Figure 5-3 : Fins des quarts de travail pour une semaine moyenne selon les saisons (incrément de 30 min) 69

Figure 5-4 : Taxis en service et taxis avec clients pour une semaine moyenne selon les saisons (incrément de 30 min) 70

Figure 5-5 : Disponibilité et utilisation des taxis pour une semaine moyenne d'automne (incrément 30 min) 71

Figure 5-6 : Densité horaire moyenne de disponibilité des taxis pour une journée moyenne de l'année 2015 73

Figure 5-7 : Distance parcourue par jour par taxi et moyenne de la distance sans client pour une semaine moyenne de l'année 2015 75

Figure 5-8 : Durée de travail totale et sans client par jour par taxi 76

Figure 5-9 : Courses effectuées par chauffeur par jour selon les saisons 80

Figure 5-10 : Courses effectuées par chauffeur par saison cumulée pour l'ensemble des chauffeurs 81

Figure 5-11 : Courses par jour pour l'année 2015 selon les précipitations et les jours fériés 82

Figure 5-12 : Origines et destinations des courses pour une semaine moyenne d'automne (incrément de 30 min) 84

Figure 5-13 : Origines et destinations des courses pour une journée moyenne de 2015 85

Figure 5-14 : Type de courses commandé pour une journée moyenne de 2015 88

Figure 5-15 : Distances de courses pour l'été 2015 89

Figure 5-16 : Distances de courses pour l'hiver 2015 90

Figure 5-17 : Durée de courses pour l'année 2015 91

Figure 5-18 : Courses débutant aux postes d'attente pour une journée moyenne de 2015 92

Figure 5-19 : Proposition d'un tableau de bord 94

LISTE DES SIGLES ET ABRÉVIATIONS

AMT	Agence Métropolitaine de Transport
BTM	Bureau du Taxi de Montréal
BTR	Bureau du Taxi et du Remorquage
CTQ	Commission des Transports du Québec
GPS	Global Positioning System
MTMDET	Ministère des Transports, de la Mobilité durable et de l'Électrification des transports
SAAQ	Société de l'Assurance Automobile du Québec
SPVM	Service de Police de la Ville de Montréal
STM	Société de Transport de Montréal
TD	Taxi Diamond
TATC	Temps d'accès au transport en commun

LISTE DES ANNEXES

Annexe A – Tables d'information de Taxi Diamond 122

Annexe B – Tables d'information de la base de données 124

CHAPITRE 1 INTRODUCTION

Le taxi représente la première forme de transport public ayant émergé dans les régions urbaines. Il fait donc partie de l'éventail des modes de transport depuis de nombreuses décennies. À partir des années 1980, l'industrie du taxi ralentit son évolution, ne modifiant aucun aspect de sa règlementation à l'exception du nombre de permis et de la tarification des courses. Ne s'adaptant plus à l'environnement urbain, elle se voit maintenant confrontée à diverses situations touchant la diversification des options de transport et le vieillissement de la population.

L'essor de nouveaux modes alternatifs à la possession automobile est de nos jours une tendance qui modifie la place du taxi dans le choix modal de la population. L'autopartage et le vélopartage sont des formes de services prenant de l'ampleur dans plusieurs villes du monde. Par exemple, à Montréal, ses modes sont opérés par des compagnies telles Car2Go, Communauto ou Bixi. Tout comme ces nouveaux modes, le taxi est une solution alternative à la possession automobile. Les besoins d'utiliser une automobile ne sont qu'occasionnels en milieu urbain et le taxi peut très bien répondre à ces besoins (Darbéra, 2010). Le taxi est aussi un mode complémentaire ou substitut selon les circonstances au transport collectif permettant ainsi aux individus de se déplacer lors de situations particulières (Austin Jr., 2011). Cela témoigne donc de son potentiel dans la mobilité durable.

Le vieillissement de la population va quant à lui changer les besoins de mobilité, bouleversant la demande en taxi. Combiné à l'étalement urbain, il engendrera une augmentation du marché potentiel sur une grande partie du territoire de la ville. En effet, « les hommes qui aujourd'hui ne peuvent plus conduire dans de bonnes conditions de sécurité parce qu'ils sont trop vieux sont la première génération qui a majoritairement possédé une voiture et organisé sa mobilité autour d'elle. » (Darbéra, 2010). Ainsi, la population habituée à la voiture ne pouvant plus conduire verra le taxi comme une option vers laquelle se tourner pour se déplacer. Cela montre donc l'importance du taxi pour ce type de clientèle.

Ce contexte exige un repositionnement de l'industrie. Elle doit maintenant se questionner sur la qualité de ses services, sa réglementation et d'éventuelles méthodes de gestion de la flotte. Bref, elle doit se questionner sur son avenir afin de trouver son rôle dans le système de transport.

1.1 Problématique

La problématique majeure de l'industrie du taxi est qu'elle ne connaît pas les activités de sa flotte de véhicules. Le nombre de taxis en service, la distance et la durée à vide ou les zones les plus achalandées sont des exemples d'informations mal connues de l'industrie. Cela rend difficile la compréhension de son rôle dans la mobilité quotidienne des individus, ce qui fragilise son maintien et son développement. Le manque de données et l'absence d'analyses systématiques en sont les principales causes.

Pour les villes où des compagnies privées assurent le service de taxi, la compétition au sein de l'industrie représente un frein au développement de connaissances sur les activités de la flotte. Chaque compagnie concurrence l'autre afin d'obtenir le plus d'appels de client. Cet esprit de compétition fait en sorte que ces intermédiaires en services de taxi sont quelques fois réticentes à fournir leurs données, ce qui en restreint l'accessibilité. Dans le cas positif où les compagnies fournissent leurs données, celles-ci sont malheureusement de faibles qualités. Elles ne contiennent pas toute l'information désirée (les courses hélées sur la rue n'en font généralement pas partie) et l'information est mal définie (la signification de certaines données est inconnue). Bref, de lourdes démarches sont à considérer afin d'obtenir et d'assimiler ces données.

Les analyses sur les taxis sont rares. Celles qui existent portent majoritairement sur les aspects contrôlés par l'instance gouvernementale responsable de la règlementation. Ainsi, des statistiques sur les chauffeurs, les permis, les inspections ainsi que le service à la clientèle peuvent être émises de manière périodique. Des articles sur la règlementation générale, c'est-à-dire les avantages et inconvénients que procure un tel système contrôlé, sont aussi publiés depuis quelques années. Toutefois, dû à l'absence de données, très peu d'analyses témoignent des activités des taxis.

Dans le but de remédier à ce manque, l'industrie du taxi a grand intérêt à tirer parti des progrès technologiques de notre époque. Elle doit se doter des instruments capables de recueillir et traiter l'information dont elle a besoin. Il n'est donc pas étonnant de voir aujourd'hui une hausse du nombre de compagnies implémenter des dispositifs à bord des véhicules récoltant des données pour assister ses opérations et fournir des informations au client.

L'information recueillie par ces nouveaux dispositifs technologiques ouvre des portes sur plusieurs champs d'études. Modéliser la demande et l'offre de taxi, répartir efficacement la flotte,

recommander aux chauffeurs les endroits les plus lucratifs ou assister les clients dans leur recherche de taxi sont des exemples de l'application des données GPS dans cette industrie. De nouvelles compagnies technologiques proposent aussi des applications mobiles pour commander un taxi. Les projets entrepris grâce à l'utilisation des données GPS aident également aux opérations ainsi qu'au processus de planification du service. Les différents usages de ces données permettent donc de rendre le service plus attractif pour la clientèle ainsi que de cheminer vers l'équilibre entre l'offre et la demande, qui est un élément critique pour améliorer le service.

1.2 Objectifs

L'objectif principal de cette recherche est de développer des indicateurs de performance et de suivi reflétant un système de transport par taxi. Ces indicateurs quantifient l'offre et la demande de taxis et sont développés à partir de données GPS provenant des systèmes embarqués à bord des véhicules. Grâce à leur analyse, les indicateurs contribuent à établir le portrait actuel de l'utilisation du taxi. Les représentants de l'industrie ainsi que ceux impliqués dans sa planification et sa réglementation auront une meilleure compréhension de son rôle et de son potentiel dans la mobilité.

Ce projet de recherche vise à atteindre les objectifs spécifiques suivants :

1. Évaluer le potentiel analytique des données provenant des systèmes embarqués (GPS) et d'autres sources de données administratives ;
2. Développer une méthode d'identification des objets du système (chauffeurs, taxis, courses et postes d'attente) basée sur les différentes bases de données intégrées ;
3. Proposer un ensemble d'indicateurs de performance et de suivi basés sur les objets identifiés ;
4. Développer un modèle de génération de la demande en taxi soulignant les principaux facteurs d'influence ;
5. Contribuer au développement d'un outil de consultation permettant le suivi et l'analyse des indicateurs sur les activités des taxis.

1.3 Structure du mémoire

Le présent document est composé de 7 chapitres.

Le premier chapitre est une introduction au domaine du taxi. Il met en contexte le taxi dans la société actuelle et surligne la problématique majeure à laquelle ce dernier fait face. La raison de l'existence du projet de recherche étant ainsi justifiée, les objectifs sont par la suite posés. Le chapitre 2 présente une revue de la littérature dans ce domaine. Il permet de mieux comprendre l'ensemble des aspects liés au taxi puisqu'il définit le taxi, relate les éléments majeurs de son histoire, montre le potentiel de ses données GPS, explique différents enjeux de son industrie et introduit son rôle dans la mobilité.

Le chapitre 3 présente la ville de Montréal qui sert d'étude de cas pour l'application des indicateurs développés. Une description des différents transports et de l'industrie du taxi de la ville permet donc de saisir l'environnement dans lequel le taxi évolue.

Le chapitre 4 établit la méthodologie pour traiter les données essentielles à ce projet. Il est nécessaire de décrire les données disponibles, le montage du système d'information utilisé ainsi que l'approche à suivre pour développer les indicateurs de performance et de suivi de l'offre de taxi. Les différents indicateurs choisis pour refléter le système de transport par taxi ainsi que les méthodes pour les calculer sont aussi présentés en détail. Le chapitre 5 sert à mettre en contexte l'ensemble des indicateurs pour la ville de Montréal. Les données utilisées sont spécifiées afin de mieux pouvoir interpréter les résultats obtenus.

Toujours pour la ville de Montréal, un modèle de génération de la demande en taxi est développé au chapitre 6. Les ensembles de données utilisés sont exposés, suivis de la méthodologie d'estimation et de l'analyse de deux régressions linéaires : une pour les origines des clients et une pour leurs destinations.

Le chapitre 7 conclut en résumant les points importants du projet de recherche, les différents niveaux de contributions, les limitations ainsi que les perspectives d'un tel projet pour l'industrie du taxi.

CHAPITRE 2 REVUE DE LITTÉRATURE

Ce chapitre propose une revue de la littérature sur le taxi afin de saisir les enjeux principaux sur ce sujet. Premièrement, il est nécessaire de bien établir la définition de ce service. Puis, un résumé de l'histoire du développement de ce mode de transport est exposé. Les aspects plus techniques sont ensuite abordés à propos du potentiel des données GPS à bord des véhicules et de l'industrie du taxi (règlementation, répartition de la flotte, utilisateurs, facteurs d'influence et indicateurs). Finalement, le taxi est pris en considération parmi les systèmes de transport urbain afin de comprendre son positionnement ainsi que son rôle.

2.1 Définition du taxi

Le terme taxi est assez large puisqu'il englobe de nombreuses formes de services, véhicules et structures organisationnelles. D'ailleurs, il n'est pas caractérisé de la même manière dans toutes les villes. Il est donc pertinent d'établir d'abord une définition générale du taxi.

Le taxi est un véhicule destiné au transport payant d'un ou plusieurs passagers et de leurs bagages. Le prix du déplacement varie selon la distance à parcourir, le temps passé dans le véhicule ainsi que d'autres paramètres variant selon les villes tels que le nombre de passagers, bagages, etc. Le taxi peut être pris à partir d'un poste d'attente de taxis, hélé dans la rue, commandé suite à un appel téléphonique à un centre de répartition ou commandé depuis une application mobile. Il déplace son passager d'un point d'origine à un point de destination, lui offrant ainsi un transport porte-à-porte direct. Le taxi-partage offre la possibilité de desservir plus d'une paire origine-destination et ce, à moindre coût pour le client. Pour cela, il doit emprunter des détours par rapport au chemin le plus direct en vue de prendre et déposer d'autres passagers. La limousine, souvent incluse dans cette industrie, offre un service de haute qualité pour ses passagers, mais a généralement un tarif plus élevé.

Une définition du taxi plus précise et appropriée pour cette étude est celle de la province de Québec, qui décrit le taxi comme une automobile utilisée pour offrir ou effectuer un transport privé et rémunéré de personnes (Gouvernement du Québec, 2011). La loi provinciale sur le transport par taxi présente la règlementation pour chaque intervenant direct, soit les propriétaires de permis, chauffeurs et intermédiaires. Elle informe quant au permis requis, aux obligations, à la tarification, aux inspections et aux entités règlementant ce milieu. Il est à noter qu'au Québec, l'appellation taxi

est utilisée pour tous les véhicules offrant un service de taxi, à l'exception des limousines. Il n'en est pas de même dans toutes les villes du pays et à l'international ; les villes anglophones, par exemple, nomment différemment les véhicules selon le type de services qu'ils fournissent et la manière dont ils sont commandés (Public Hire Vehicule, Taxi, Private Hire Vehicule, For Hire Vehicle, Taxibus).

Dans le système de transport global, le taxi se positionne dans le transport parallèle. La définition du transport parallèle, fournie par l'Office Québécois de la Langue Française (2010), est :

> « Système de transport instauré d'abord pour assurer la desserte de zones peu densément peuplées et ensuite pour régler une partie des problèmes de pollution et de congestion dus à l'usage croissant de l'automobile. Le transport parallèle utilise le réseau routier existant, mais est plus flexible que le transport en commun, puisque la tarification, les horaires et les parcours ne sont pas fixes. L'automobile, la fourgonnette et le minibus sont des types de véhicules utilisés fréquemment par le transport parallèle. »

Le tableau 2-1, inspiré d'Austin Jr. (2011), explique que le taxi se situe entre le transport privé (possession privée d'une automobile) et collectif (transport en commun). En effet, l'utilisation du taxi ne nécessite pas d'être propriétaire du véhicule. Il offre un service flexible desservant n'importe quelle paire origine-destination en région urbaine à n'importe quel moment de la journée. Cela fait en sorte que son tarif est plus élevé que les autres modes du transport parallèle et en commun. Il accueille aussi, à bord de ses véhicules, un nombre plus restreint de passagers que le transport en commun.

Le taxi est considéré comme un transport public puisqu'il est disponible pour n'importe quel passager contre un certain tarif et ne nécessite pas la possession d'une automobile (Austin Jr., 2011). Toutefois même si le taxi est un transport public et est lourdement règlementé par une institution publique, celui-ci reste opéré par des compagnies privées. Son intérêt est donc de réaliser un profit.

Tableau 2-1 : Modes de transport parallèle public du véhicule privé au transport en commun

		Type de services	Configuration du service (Origine à Destination)	Charge typique de passagers	Conducteur	Tarif	Règlementation juridique typique	Degré de restrictions règlementaires
VÉHICULE PRIVÉ		Sur demande (appartient à l'utilisateur)	Plusieurs à plusieurs	1-4	Propriétaire	Élevé	Province	Faible
TRANSPORT PARALLÈLE	**Autopartage**	Sur demande	Plusieurs à plusieurs	1-4	Abonné	Modéré	Ville/province	Faible à modéré
	Taxi	Sur demande, requête téléphonique ou hélée	Plusieurs à plusieurs	1-4	Professionnel	Élevé	Ville	Élevé
	Taxi spécialisé	Sur demande, requête téléphonique	Plusieurs à plusieurs	1-4	Professionnel	Élevé	Ville	Élevé
	Taxi-partage	Sur demande, requête téléphonique ou hélée	Plusieurs à plusieurs	2-4	Professionnel	Modéré	Ville	Élevé
	Minibus spécialisé	Sur réservation, requête téléphonique	Plusieurs à plusieurs	2-10	Professionnel	Faible	Ville	Élevé (appartient à la province)
	Minibus (circulatoire)	Route régulière, arrêts fixes	Route fixe/boucle	6-15	Professionnel	Faible	Ville	Modéré à élevé
	Minibus (répond à la demande)	Route régulière, requête hélée	Plusieurs à un seul	6-15	Professionnel	Faible	Ville	Modéré à élevé
	Minibus (étendu)	Route semi-régulière, requête hélée	Plusieurs à plusieurs	6-15	Professionnel	Faible	Ville	Modéré à élevé
	Van de banlieue	Sur réservation, planifié	Peu à un seul	6-15	Professionnel	Faible	Province	Faible
TRANSPORT EN COMMUN		Route régulière, arrêts fixes	Route fixe	10 ou plus	Professionnel	Faible	Région métropolitaine	Élevé (appartient à la province)

Source : Inspiré de Austin Jr. (2011) citant Cervero (1997)

2.2 Historique

Le taxi est une des plus vieilles formes de transport public encore en fonction de nos jours, remontant bien avant l'ère de la motorisation. La revue historique du taxi de Austin Jr. (2011) est une synthèse pertinente de la littérature. De plus, Cooper, Mundy et Nelson (2010) font une revue complète du développement du taxi au Royaume-Uni ainsi qu'aux États-Unis. Ils affirment qu'il n'y a aucun doute que le concept d'un véhicule opéré avec rémunération existe depuis l'origine des véhicules mêmes. La figure 2-1 présente la ligne du temps des évènements majeurs de l'industrie du taxi.

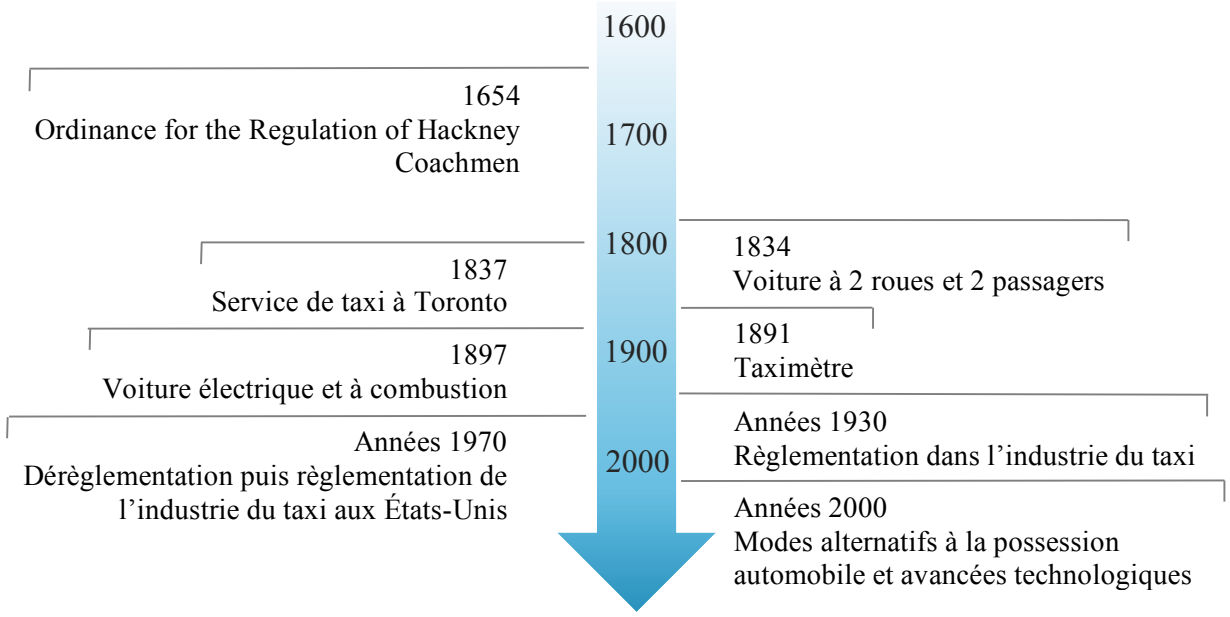


Figure 2-1 : Ligne du temps des évènements majeurs du taxi

Les premières traces de l'existence de taxi remontent au début du 17^{e} siècle, à Londres et Paris, où ces véhicules à disposition du public sont des voitures hippomobiles. À Londres, il prend le nom de hackney. Le texte règlementaire de 1654 « Ordinance for the Regulation of Hackney Coachmen » est la première preuve officielle de l'opération des hackneys (Cooper et al., 2010). À Paris, ce mode de transport se nomme voiture de louage ou plus communément fiacre, tiré de l'hôtel Le Saint-Fiacre en face duquel le premier poste d'attente des voitures de louage était situé. Le premier texte officiel faisant mention des voitures à louage est une ordonnance émise par le roi Louis XIV en 1657, dressant la règlementation de ce service de transport (Taxi de France, 2016).

Puis, les fiacres s'étendent en Europe et atteignent l'Amérique. Du milieu du 17e siècle jusqu'au début du 19e siècle, les fiacres utilisés sont des voitures à quatre roues tirées par des chevaux possédant de quatre à huit sièges passagers. Durant cette période, une séparation dans l'opération des fiacres se produit : ceux empruntant des routes fixes se distinguent alors de ceux offrant un service porte-à-porte. C'est donc à cette époque que la distinction entre les services de taxi et de transport en commun fait son apparition.

En 1834, la voiture à deux roues et deux places pour passagers est inventée. Celle-ci est beaucoup plus légère et permet des manœuvres plus précises. Elle est tirée par un seul cheval, ce qui réduit le coût des déplacements. Cela la rend populaire auprès des compagnies de taxi en Europe et en Amérique. Le premier service de taxi connu au Canada date de 1837, dans la ville de Toronto. En 1891, le taximètre est inventé, duquel découle l'appellation taxi.

L'ère de la motorisation débute en 1884 avec l'arrivée de la voiture électrique alimentée par une batterie. Des compagnies à Londres, Paris et New York se dotent de taxis propulsés à l'électricité en 1897. Malheureusement, pour des raisons de rentabilité, ces véhicules sont délaissés dans les années 1920. Les voitures à moteur à combustion font aussi leur apparition en 1885. Elles envahissent le marché du taxi lors du déclin des taxis électriques et remplacent rapidement les voitures tirées par chevaux.

Avec la vague d'arrivée des voitures à pétrole, certaines villes commencent à règlementer ce secteur. Londres, en 1907, émet des restrictions quant au type de véhicule utilisé comme taxi ; celui-ci doit pouvoir effectuer un demi-tour à l'intérieur d'une route de 25 pieds de large. Cette période se caractérise aussi par la croissance des petites compagnies de taxi vers des flottes de plus grande envergure opérant dans plusieurs villes (Austin Jr., 2011). La tendance des prix à diminuer et l'invention du téléphone et de la radio permettent à une plus grande partie de la population d'utiliser ce service.

La Grande Dépression a des effets importants sur l'industrie du taxi. Les emplois sont rares dans plusieurs secteurs ce qui amène une surabondance de chauffeurs de taxi puisqu'il est facile d'entrer dans cette industrie. Les conditions de compétition découlant de cette situation s'amplifient, menant même à la violence. Cela fait réaliser aux décideurs qu'il y a une nécessité de règlementer plus strictement. Ainsi, trois sphères de règlementation sont instaurées par la ville : le nombre

d'entrées (quantité), le tarif de déplacements (tarif) et l'aspect du véhicule (qualité). L'industrie est désormais opérée par le privé, mais règlementée par le public.

Dans les années 1970, le gouvernement fédéral des États-Unis impose une dérèglementation du transport de biens et personnes par train, avion et véhicule motorisé. Peu de temps après, les villes appliquent cette dérèglementation à leur industrie du taxi. Suite à plusieurs effets pervers sur les tarifs, les temps d'attente, la productivité, etc., certaines villes règlementent de nouveau leur industrie.

À partir de la fin du 20e siècle, des modes alternatifs à la possession automobile émergent. L'autopartage et le vélopartage sont des formes de nouveaux modes prenant de l'ampleur. Les avancées technologiques sont aussi nombreuses durant cette période, notamment internet, les technologies de localisation par satellite (GPS) et les téléphones intelligents qui permettent le développement d'applications mobiles pour commander un taxi. Cela fait donc se remettre en question l'industrie du taxi qui a, quant à elle, peu évolué au cours des dernières années.

2.3 Potentiel des données GPS

Avec les progrès technologiques des années 2000, les systèmes embarqués de localisation sont devenus de plus en plus abordables. De nombreux dispositifs GPS sont installés dans des véhicules de même que dans des appareils mobiles. Les données emmagasinées grâce à ces dispositifs ont un potentiel énorme puisqu'elles sont enregistrées en continu, ce qui procure un large ensemble de données à analyser. Le rôle des données GPS dans le domaine des transports est une aide au suivi opérationnel, à l'identification des comportements de déplacement ainsi qu'à la modélisation. Toutefois, la grande quantité d'information amassée par les différents médias nécessite de reconsidérer les méthodes d'analyse sur la qualité des données et la législation sur la confidentialité.

Le premier rôle des données GPS est d'assurer le suivi des opérations d'un système de transport. Que ce soit pour le transport de marchandises ou de personnes, par voies terrestres ou aériennes, les dispositifs GPS à bord des véhicules recueillent de l'information sur leurs déplacements. Les opérateurs et planificateurs ont donc la possibilité de superviser en tout temps la localisation de leur flotte grâce aux données en temps réel, en plus de constater l'évolution de leurs activités grâce à l'analyse des données historiques.

Le second rôle des données GPS dans le domaine du transport est d'identifier les comportements de mobilité. Les données GPS améliorent les enquêtes origine-destination puisqu'elles permettent, entre autres, de voir tous les mouvements de la personne tout en réduisant la durée de l'entrevue. Le traitement de l'information est cependant plus laborieux : des méthodes d'identification du déplacement (Schönfelder, Axhausen, Antille, & Bierlaire, 2002), du mode de transport (Zheng, Li, Chen, Xie, & Ma, 2008), du motif (Zheng et al., 2008) et autres caractéristiques du déplacement sont développées. Wolf, Bachman, Oliveira, Auld, Mohammadian, Vovsha, et al. (2014) proposent une revue des méthodes utilisées pour identifier les caractéristiques des déplacements. Les enquêtes origine-destination reflètent la mobilité quotidienne des personnes, mais l'identification des comportements pour des modes précis comme le vélo, l'automobile ou le camion est aussi envisageable.

Le troisième rôle des données GPS est le support à la modélisation. Ces données sont utilisées pour modéliser différents aspects du transport de manière à apporter une couverture spatiale et temporelle amplifiée et continue. En général, pour ce faire, l'identification des comportements de déplacement est préalablement accomplie. La demande en transport, la circulation (Liu, X., Gong, Gong, & Liu, 2015) et la congestion routière (Huang et al., 2016), les temps de déplacement (Li, Enam, Abou-Zeid, & Ben-Akiva, 2013) ainsi que le choix de route (Dai & Chen, 2014) sont des exemples de modélisation améliorés par l'emploi des données GPS.

Le dernier rôle consiste à soutenir les domaines connexes au transport. L'apport de ce type de données permet d'obtenir des informations non disponibles auparavant et plus précises. Ainsi, les études sur les émissions de gaz à effet de serre (Yu & Peng, 2013), le péage, l'activité physique et autres s'améliorent grâce à ces données.

Les médias de collecte de données GPS se diversifiant avec les avancées technologiques, il devient donc pertinent de procéder à des vérifications sur la validité de ces données (Wolf, Bachman, Oliveira, Auld, Mohammadian, & Vovsha, 2014). D'abord, selon l'appareil utilisé, un biais du répondant est à définir. Par exemple, les données GPS récoltées par les utilisateurs d'une application mobile sont biaisées, car seulement ceux possédant un appareil mobile et portant un intérêt pour cette application fournissent leurs informations. Ensuite, une vérification de la qualité des données GPS récoltées doit être effectuée. La qualité varie, entre autres, en fonction de la dispersion satellite, du nombre de satellites et du bruit. Les points aberrants induits par ces

variations causent des irrégularités dans le trajet GPS détectables grâce à l'utilisation de plusieurs indicateurs : le nombre de points, la distance, la somme des angles, la vitesse moyenne, etc. (Wang et al., 2015). Finalement, la confidentialité des données est un enjeu important pour l'individu fournissant ses informations. Les trois principales méthodes pour assurer la confidentialité sont : la signature d'une entente entre l'organisme récoltant les données et l'individu, la création de lois par le gouvernement et l'agrégation des données (Wolf, Bachman, Oliveira, Auld, Mohammadian, Vovsha, et al., 2014).

2.4 Industrie du taxi

L'industrie du taxi est définie par plusieurs aspects sociaux, règlementaires et opérationnels. Elle observe ses propres activités dans le but d'améliorer les différents services de taxi et d'augmenter les bénéfices sociaux des passagers, chauffeurs et propriétaires. Pour y arriver, les analyses se concentrent sur les deux éléments principaux de tout système de transport : l'offre et la demande. Afin de mieux comprendre ses activités et de suivre l'évolution de ces deux éléments, une analyse des indicateurs de performance est requise.

2.4.1 Offre

Afin de comprendre l'offre de taxis, il faut d'abord saisir la règlementation mise en place dans le but de contrôler les services de taxi ainsi que les enjeux de sa pertinence dans notre société. Par la suite, puisque les comportements de la flotte de véhicules dans le temps et l'espace reflètent l'accessibilité au service, il est utile de connaître les différentes méthodes de répartition des véhicules afin de proposer un service répondant à la demande.

2.4.1.1 Règlementation

La règlementation dans l'industrie du taxi sert à fixer l'offre de véhicules sur le territoire. Elle se fait à trois niveaux : la quantité, le tarif ainsi que la qualité.

La quantité de taxis est contrôlée à travers la délimitation du territoire en zones ainsi que la détermination du nombre de permis émis pour chaque zone. Les zones sont instaurées afin d'assurer une dispersion des taxis sur le territoire. Le nombre de permis pour chaque zone est fixé dans le but d'atteindre l'équilibre entre l'offre et la demande. Les modèles d'équilibre représentent une aide à la décision dans les industries règlementées (Yang, H., Wong, & Wong, 2002). Le choix

de la taille de la flotte est déterminant pour la disponibilité ainsi que la qualité des services de taxi (Schaller, 2005). Les modèles d'équilibre évaluent, pour la majorité, la demande ainsi que l'offre de taxis selon l'heure de la journée, les conditions météorologiques ainsi que l'emplacement des taxis (Chang, Tai, & Hsu, 2009). Ils sont aussi étudiés sous différents secteurs économiques du marché (monopole ou oligopole) avec l'ajout de facteurs comme le réseau routier, les patrons de paires origine-destination et, quelques fois, la congestion routière (Yang, H. et al., 2002; Zhu, Shuai, Huang, & Sun, 2013).

Le tarif des courses est contrôlé pour assurer au client une uniformité dans les déplacements par taxi. Le développement de modèles de détermination des tarifs des courses est un support pour les décideurs. Ces modèles montrent que les tarifs dépendent évidemment du nombre de taxis dans la flotte, mais prennent aussi en considération le coût d'opération, la localisation et l'heure de la journée (Cooper et al., 2010).

La qualité du service et des véhicules est contrôlée dans le but d'assurer un minimum d'uniformité dans la flotte. Elle est souvent étudiée à partir d'enquêtes de satisfaction chez les clients. De nombreuses études sont réalisées dans les villes par l'entremise de questionnaires fournis aux différents intervenants de l'industrie afin de recueillir leurs commentaires sur l'amélioration des services (Taxi Study Panel, 1999; Tennessee Transportation & Logistics Foundation, 2009, 2010a, 2010b, 2014). L'inspection des véhicules se fait couramment dans l'industrie pour vérifier la propreté et la sécurité des véhicules (New York City Taxi & Limousine Commission, 2014). Le contrôle de la qualité est jugé nécessaire afin d'assurer un service satisfaisant pour les passagers et chauffeurs à travers des normes de sécurité, d'entretien, de qualification des véhicules et chauffeurs (Organisation for Economic Co-operation and Development, 2007).

2.4.1.2 Dérèglementation

Au cours de la fin du XXe siècle, certaines villes ont expérimenté la dérèglementation de leur industrie afin de contrer les effets négatifs des lois établies. Suite à leur expérience, les analystes se sont trouvés divisés en deux communautés : ceux appuyant la règlementation et ceux la rejetant.

Un marché ouvert comporte comme avantages une diminution des délais d'attente et une baisse des tarifs ou du moins une stabilisation grâce à la plus grande concurrence entre les taxis en service. Bacache-Beauvallet et Janin (2012) montrent que la valeur des licences diminue avec la libéralisation du nombre d'entrées. Le British Office of Fair Trading affirme aussi que le contrôle

du nombre d'entrées a comme effet de réduire l'accessibilité, d'augmenter le temps d'attente et de nuire à la sécurité (Gwilliam, 2005).

Dérèglementer l'industrie du taxi apporte comme désavantages une congestion plus lourde sur les routes, rallongeant ainsi la durée du déplacement même si le temps d'attente est court (Gwilliam, 2005). Dans les zones à faible densité de population et selon l'heure de la journée, l'accessibilité aux taxis diminue et les tarifs augmentent (Gwilliam, 2005). Cairns et Liston-Heyes (1996) en viennent même à la conclusion que, selon leur modèle établi, l'offre et la demande n'atteindront jamais l'équilibre sous un marché ouvert.

Toutefois, l'expérience récente de nouvelles formes de transport par taxi laisse plusieurs penser que la décentralisation de ce mode de déplacement permettra d'approcher l'équilibre entre l'offre et la demande. Ainsi, ils jugent qu'une ouverture dans la règlementation devrait être instaurée, tenant compte de la variation de la demande dans les tarifs et le nombre de permis (Zhang, W. & Ukkusuri, 2016) et laissant un nombre approprié de compagnies s'implanter dans ce secteur (Harding, Kandlikar, & Gulati, 2015).

2.4.1.3 Répartition de la flotte

Dans un marché règlementé, situation dans laquelle la majeure partie des villes industrialisées se trouvent, l'atteinte de l'équilibre entre l'offre et la demande nécessite une répartition adéquate de la flotte. Les différents intermédiaires en services présents dans la ville reçoivent les commandes des clients et les répartissent à travers les véhicules disponibles. Toutefois, après avoir complété sa course, un chauffeur de taxi se voit confronté au problème du choix de sa prochaine localisation pour trouver son prochain client. De nombreux modèles tentent de décrire cette situation ainsi que de proposer des recommandations aux chauffeurs et aux clients quant au meilleur endroit où se diriger (Hu, Y., Yang, & Huang, 2015).

Un premier modèle fournit de l'information au client sur l'endroit où prendre un taxi. La probabilité d'obtenir un taxi à l'endroit le plus proche ou à un endroit et moment particulier est développée grâce à un modèle probabiliste conçu à partir de données GPS historiques (Yuan, J., Zheng, Zhang, Xie, & Sun, 2011; Yuan, N. J., Zheng, Zhang, & Xie, 2013). De plus, Xu, Zhou, Liu, Xu et Zha (2015) prédisent le temps d'attente estimé pour obtenir un taxi, comme le font aussi Deep Singh, Wu, Xiang et Krishnaswamy (2015).

Un deuxième modèle concerne directement les chauffeurs, leur recommandant les routes à prendre et endroits où aller pour trouver leur prochain client. Les chauffeurs se dirigent où ils estiment que la probabilité d'obtenir un passager est élevée (Wong, Szeto, & Wong, 2014). Pour qu'un système de taxi soit optimal au niveau de la distance parcourue par les véhicules, chaque chauffeur doit connaître où le prochain client sera. D'un côté, un modèle peut recommander au chauffeur les routes montrant la plus grande probabilité d'obtenir un passager (Hu, X., Gao, Chiu, & Lin, 2012). D'un autre côté, un modèle peut recommander l'endroit le plus proche ayant la plus grande probabilité, plus précisément les postes d'attente ou des lieux de concentration de demande historique (Dow, Chen, & Hwang, 2015; Hsueh, Hwang, & Chen, 2014; Shin, Park, Saha, Kwak, & Kim, 2009; Tang, Kerber, Huang, & Guibas, 2013). Les deux recommandations peuvent aussi être fournies au chauffeur : l'endroit le plus profitable (en termes de poste d'attente) ainsi que le chemin pour s'y rendre (Afian, Odoni, & Rus, 2015; Yang, T., Yang, Wong, & Sze, 2014; Yuan, J. et al., 2011; Yuan, N. J. et al., 2013).

Moreira-Matias, Gama, Ferreira, Mendes-Moreira et Damas (2013) améliorent ces modèles en utilisant des données GPS pour générer la demande aux postes d'attente et fournir un modèle en ligne se mettant à jour toutes les 24 heures. D'autres systèmes de répartition en temps réel sont développés (Ding, Xiong, & Liu, 2013) et peuvent même montrer qu'une diminution de 52 % des distances à vide est possible (Miao et al., 2015). Ces systèmes de répartition en temps réel peuvent aussi fournir de l'information aux clients tels que le temps d'attente et la géolocalisation du taxi associé à la course (Jianxin, Xiaomin, & Hongyu, 2009).

2.4.2 Demande

Acquérir une bonne compréhension la demande en taxi est important puisque la répartition de la flotte de véhicules s'effectue en fonction du comportement des utilisateurs. Ainsi, il devient pertinent de connaître les types d'utilisateurs ainsi que les facteurs influençant leur choix de prendre un taxi afin de cibler les améliorations à apporter pour satisfaire leurs besoins. Dans une optique opérationnelle, être en mesure de prédire où émergera la demande dans le temps et l'espace est aussi un enjeu pour mieux atteindre l'équilibre de l'offre et de la demande.

2.4.2.1 Types d'utilisateurs

Il existe deux grands groupes d'utilisateurs de taxis : les clients non captifs ainsi que les clients captifs. Tel que le décrivent Beimborn, Greenwald et Jin (2003), les clients non captifs sélectionnent le service de transport en commun ou d'automobile lorsqu'ils considèrent une option supérieure à l'autre, tandis que les clients captifs n'ont qu'une seule option pour se déplacer. D'un côté, les utilisateurs ont le loisir de décider leur mode de transport lorsque plusieurs modes sont accessibles pour effectuer leur déplacement. Ils basent leur décision sur l'accessibilité à l'origine et la destination, la durée, le coût, la commodité, le confort, la sécurité et la connectivité (Beimborn et al., 2003). De l'autre côté, la captivité dans un déplacement peut être permanente ou ponctuelle (Darbéra, 2010). Elle est due à des facteurs de la personne tels l'âge, l'infirmité, le revenu et la possession automobile, tout comme à des facteurs spatiotemporels tels que résider en banlieue ou se déplacer la nuit lorsque le service de transport en commun n'est pas fonctionnel.

Les caractéristiques des individus et des services de transport déterminent le type d'utilisateur (non-captif ou captif). Toutefois, des situations créées par les chauffeurs de taxi font aussi en sorte d'améliorer ou de limiter l'accès au service de taxi. En fait, chaque chauffeur cherche à maximiser son profit et donc à minimiser son temps de recherche de clients (Phithakkitnukoon, Veloso, Bento, Biderman, & Ratti, 2010; Yang, H. et al., 2002). Ceux-ci vont alors se diriger vers des zones connues de grande demande où les clients profiteront de meilleurs services (Austin Jr., 2011). D'autres zones se voient délaissées par les taxis, ce qui décourage les individus à utiliser ce mode de transport. Ainsi, le comportement des chauffeurs peut influencer les bénéfices d'un voyageur à utiliser le service de taxi.

New York est un exemple de ville qui réalise des enquêtes sur la clientèle de leur service de taxi. La New York City Taxi & Limousine Commission (2014) recueille des données tels l'âge, le sexe, la possession automobile, la possession de permis de conduire, le revenu, etc. Cela leur permet de déterminer que leurs principaux utilisateurs sont des clients non captifs puisqu'ils possèdent un revenu élevé et qu'ils effectuent des déplacements en taxi dans des zones où le transport en commun est en forte présence.

2.4.2.2 Facteurs d'influence de la demande

Les enquêtes auprès des clients sont des opérations coûteuses pour lesquelles il est difficile d'obtenir un échantillon représentatif de tous les types d'utilisateurs ainsi que des informations

exhaustives. Ainsi, les chercheurs se tournent souvent vers la modélisation de la demande à partir des données GPS afin d'obtenir une idée de base sur les facteurs d'influence les plus importants.

Une première évaluation pour cibler les facteurs d'influence de la demande en taxi se fait à partir des modèles de choix modal. Puisque le taxi et le transport en commun sont des modes de transport public, les facteurs liés au transport en commun sont également appropriés pour le taxi. L'âge, le sexe, la structure du ménage, le revenu, l'emploi, la possession automobile, la densité de population, l'utilisation du sol, la distance d'un arrêt de transport en commun, le niveau de service, le jour de la semaine et l'heure sont de possibles facteurs explicatifs de la demande (Mousavi, Bunker, & Lee, 2012; Racca & Ratledge, 2003). Chatman (2013) ajoute à cette liste des facteurs spécifiques d'utilisation du sol comme le nombre de places de stationnement disponibles, le type de logement et la distance du centre-ville. L'étude du choix du taxi comme mode de transport montre que le revenu, l'heure de la journée et le confort sont les variables les plus importantes (Gómez-Gélvez, Rodriguez-Valencia, & Restrepo, 2015).

La possession automobile, le revenu, l'emploi, les déplacements réguliers, les utilisateurs du transport en commun et les déplacements vers ou à partir de l'aéroport sont des facteurs liés directement à la demande en taxi (Kattan, de Barros, & Wirasinghe, 2010; Schaller, 2005; Zhang, D. & Song, 2009). Yang, C. et Gonzales (2014) étudient ce problème de compréhension de la demande en détail. Ils modélisent les origines et destinations des déplacements en taxi pour chaque heure et zone en utilisant une régression linéaire. Ils utilisent des données démographiques, d'emploi et de transport en commun. Leur conclusion est que la concentration de population, l'éducation, l'âge, le revenu, le temps d'accessibilité au transport en commun, le nombre d'emplois ainsi que le type d'emploi influencent le nombre d'origines et de destinations dans une zone. Qian et Ukkusuri (2015) utilisent, quant à eux, une régression pondérée géographiquement afin de cibler les facteurs d'influence. Ce type de régression, montrant un meilleur pouvoir explicatif que celle des moindres carrées, montre que le revenu médian, le niveau de scolarité (baccalauréat), les zones résidentielles, la densité de rues et l'accessibilité au métro sont les principaux facteurs explicatifs de la demande en taxi. Toutefois, il ne faut pas oublier que le marché du taxi diffère d'une ville à l'autre. Les facteurs d'influence varient donc aussi selon la ville.

2.4.2.3 Prédiction de la demande

Sachant que le temps et l'espace influencent la demande, celle-ci peut donc être prédite dans un contexte spatiotemporel. Des méthodes de regroupement sont généralement utilisées pour cerner les zones les plus achalandées durant la journée. Wan, Gao, Kang et Zhao (2013) prédisent les lieux de concentration des origines et des destinations des courses de taxis à partir de l'algorithme DBSCAN. Chang et al. (2009) appliquent plutôt trois méthodes de regroupement pour prédire les lieux de concentration et comparer leur performance, concluant que chaque méthode a ses avantages et inconvénients lorsqu'appliquée sur des données GPS. Liu, D., Cheng et Yang (2015) appliquent plutôt une grille sur la ville à l'étude pour laquelle la demande dans chaque cellule est modélisée par une loi de Poisson et un modèle de mélange de Gaussienne.

2.4.3 Suivi analytique

La compréhension du taxi dans son ensemble passe par l'état de ses activités et de son suivi dans le temps. Le développement d'indicateurs spatiotemporels assure cette compréhension et permet de suivre l'évolution des activités sur plusieurs années. Les différents objets du système de taxi ainsi que certaines conditions spatiales et temporelles sont ainsi étudiés. Une revue des indicateurs rencontrés dans la littérature est présentée au tableau 2-2.

Tableau 2-2 : Revue des indicateurs

OBJET	INDICATEUR	RÉFÉRENCES	
Course	Carte de chaleur des origines et destinations	• (Ferreira, Poco, Vo, Freire, & Silva, 2013) • (Pelé & Morency, 2014)	• (Zhang, L., Ahmadi, Pan, & Chang, 2012) • (Zhang, Y., 2014)
	Nombre d'origines et de destinations	• (Ferreira et al., 2013) • (King, Peters, & Daus, 2012) • (New York City Taxi & Limousine Commission, 2014)	• (Savage & Vo, 2013) • (Zhang, Y., 2014)
	Nombre de courses	• (Bischoff, Maciejewski, & Sohr, 2015) • (Ferreira et al., 2013)	• (New York City Taxi & Limousine Commission, 2014) • (Pelé & Morency, 2014) • (Savage & Vo, 2013)
	Distance et durée de courses	• (Bischoff et al., 2015) • (New York City Taxi & Limousine Commission, 2014)	• (Pelé & Morency, 2014) • (Savage & Vo, 2013) • (Zhang, D.-Z., Peng, & Sun, 2014)
	Vitesse de course	• (Pelé & Morency, 2014) • (Savage & Vo, 2013)	• (Zhang, D.-Z. et al., 2014)
	Nombre de passagers par course	• (Savage & Vo, 2013)	
Taxi	Distance et durée à vide	• (Bischoff et al., 2015) • (Jun, 2013)	• (Pelé & Morency, 2014) • (Zhang, D.-Z. et al., 2014)
	Temps en service	• (Pelé & Morency, 2014)	
	Nombre de courses	• (Pelé & Morency, 2014)	
	Taux d'occupation	• (Zhang, Y., 2014)	
	Nombre de taxis disponibles	• (Bischoff et al., 2015) (Ganti, Mohomed, Raghavendra, & Ranganathan, 2011)	• (New York City Taxi & Limousine Commission, 2014) • (Pelé & Morency, 2014) • (Zhang, D.-Z. et al., 2014)

Tableau 2-2 : Revue des indicateurs (suite)

Chauffeur	Temps travaillé	• (Zhang, D.-Z. et al., 2014)	
	Début et fin des quarts de travail	• (New York City Taxi & Limousine Commission, 2014)	
	Revenu ou profit	• (Ferreira et al., 2013) • (Jun, 2013) • (New York City Taxi & Limousine Commission, 2014)	• (Zhang, D.-Z. et al., 2014) • (Zhang, Y., 2014)
Zone	Temps de transition	• (Zhang, L. et al., 2012)	
	Nombre de transitions entre 2 régions	• (Zhang, L. et al., 2012)	
	Étude spécifique d'une zone (ex. : aéroport)	• (Bischoff et al., 2015)	• (New York City Taxi & Limousine Commission, 2014)
Client	Distance de marche entre l'origine et le centre du lieu de concentration	• (Zhang, D.-Z. et al., 2014)	
	Caractéristiques des clients (âge, sexe, etc.)	• (New York City Taxi & Limousine Commission, 2014)	

La course, c'est-à-dire le trajet en taxi de l'origine à la destination du client, est l'objet du système de taxi le plus étudié. En effet, les courses montrent la demande en taxi ainsi que sa variabilité dans le temps et l'espace, ce qui est essentiel à la compréhension et la planification du service. Le véhicule est souvent analysé pour déterminer sa performance ; c'est-à-dire pour connaître l'offre disponible sur le territoire selon l'heure de la journée ainsi que les distances et durées à vide. Le chauffeur, quant à lui, est plutôt étudié sous l'angle économique ; il est possible de retracer ses heures de travail ainsi que son revenu. La zone est l'objet le moins exploré dans le système. Elle est analysée pour connaître les transitions de véhicules, sinon elle est étudiée spécifiquement à travers les indicateurs de courses appliqués à sa zone. Le client est un objet difficile à cibler à partir uniquement des données provenant des taxis. Alors, seuls des indicateurs en rapport avec l'origine et la destination de sa course peuvent être calculés, dont la distance de marche jusqu'au centre du

lieu de concentration le plus proche. Dans le cas où des enquêtes auprès des clients sont conduites, il est alors possible d'obtenir des informations sociodémographiques à leur sujet.

Tous ces indicateurs sont évalués de manière spatiotemporelle et peuvent aussi être calculés lors de conditions particulières. Analyser les objets du système selon les conditions météorologiques et les jours fériés permet de comparer les effets de ces évènements sur la demande et l'offre. La neige n'occasionne aucun effet notable sur les activités de taxis, toutefois la pluie encourage un plus grand nombre de courses de plus courtes distances (Kamga, Yazici, & Singhal, 2013; Pelé & Morency, 2014). Les jours fériés montrent une tendance semblable à celle d'un dimanche ; autrement dit un faible nombre de courses sans pic d'achalandage le matin (Pelé & Morency, 2014), ainsi que des origines et destinations dispersées selon les évènements spéciaux dans la ville (Doraiswamy, Ferreira, Damoulas, Freire, & Silva, 2014).

2.5 Rôle du taxi

L'ensemble du système de transport par taxi n'englobe pas uniquement ses activités internes, mais aussi sa position dans la mobilité quotidienne d'une ville. Le positionnement du taxi parmi les nombreux modes de transport présents dans une région s'évalue en fonction des différents rôles qu'il endosse.

Le premier rôle du taxi est la complémentarité de celui-ci avec les autres modes de transport. En effet, il offre une mobilité spatiale et temporelle que n'offrent pas les autres modes, par exemple aux endroits et aux heures qui ne sont pas desservis par le transport en commun (Austin Jr., 2011), pour de longues distances non commodes en vélo ou pour des individus qui ne possèdent pas de permis de conduire. La complémentarité est donc circonstancielle et profite aux clients captifs du taxi puisqu'ils n'ont pas accès aux autres modes de déplacements.

Le second rôle du taxi est la substitution aux autres modes de transport. Les clients non captifs sont ceux qui utilisent le taxi dans son rôle de substitut. Ils trouvent plus attrayant d'utiliser le taxi que le transport en commun, la marche ou le vélo et cela même s'ils paient plus cher puisque, pour eux, le service est plus rapide, plus confortable et plus sécuritaire (Austin Jr., 2011). Les clients non captifs, en utilisant le taxi, permettent aussi une réduction du nombre de personnes engorgeant le transport en commun en heure de pointe.

À travers ces deux rôles, le taxi peut prendre une fonction de déplacement spécialisé. Cette fonction s'est développée autour de l'aspect social de la mobilité ainsi que du transport adapté (Cooper et al., 2010). Ce genre de taxi transporte des individus limités par leur condition physique, c'est-à-dire les personnes à mobilité réduite et celles nécessitant des besoins spécialisés. Le type de véhicule utilisé pour ces déplacements est donc adapté à cette clientèle.

Les clients captifs et non captifs, coexistant dans une ville, retirent chacun les avantages du rôle de complémentarité et de substitution. Cependant, la qualité du service n'est pas uniforme entre les groupes. Les zones les plus achalandées, souvent lorsque les clients sont non-captifs, reçoivent une meilleure couverture, tandis que les zones éloignées où la demande est moindre et le territoire plus étendu profitent d'une offre de taxi faible. Dans un souci d'équité sociale, il est important que les deux types de clients soient desservis à un niveau de service de base (Austin Jr., 2011).

Dans une vision de développement durable, un objectif à atteindre est la réduction de la possession d'automobile. Le taxi est un mode de transport alternatif à la voiture et est donc un service à encourager. José Viegas soulève que « quand un ménage se résout à acheter une voiture, c'est une clientèle définitivement perdue pour le transporteur » et que selon lui « pour beaucoup de ménages, la décision d'acheter une voiture ne tient qu'à des besoins occasionnels que le taxi pourrait très bien satisfaire » (Darbéra, 2010).

CHAPITRE 3 MISE EN CONTEXTE

À titre d'expérimentation, les indicateurs développés dans le cadre de cette recherche sont appliqués à la ville de Montréal. Afin de mieux saisir la situation actuelle de l'industrie du taxi dans cette ville, ce chapitre dresse le portrait du fonctionnement de l'industrie et des circonstances englobant son évolution. Premièrement, une description de Montréal, de ses réseaux de transport ainsi que des déplacements effectués sur l'île par les résidents de la région métropolitaine est exposée. Puis, les différents aspects de l'industrie du taxi de Montréal sont expliqués.

3.1 Description de la ville

Le territoire de l'île de Montréal comprend la ville de Montréal, segmentée en 19 arrondissements, ainsi que 14 autres villes indépendantes. L'île est entourée au sud et à l'est par le fleuve Saint-Laurent, à l'ouest par le lac des Deux Montagnes et au nord par la rivière des Prairies. La banlieue de Montréal est catégorisée selon trois appellations: Laval, la couronne nord et la couronne sud. La région métropolitaine de Montréal couvre le territoire indiqué à la figure 3-1 et sa population est de 3 824 221 habitants (Statistique Canada, 2016). Le Bureau du Taxi de Montréal (BTM) ne peut délivrer des permis que pour le territoire de l'île. Ainsi, les taxis ont le droit d'opérer uniquement sur l'île, à quelques exceptions près mentionnées à la section 3.2.3.

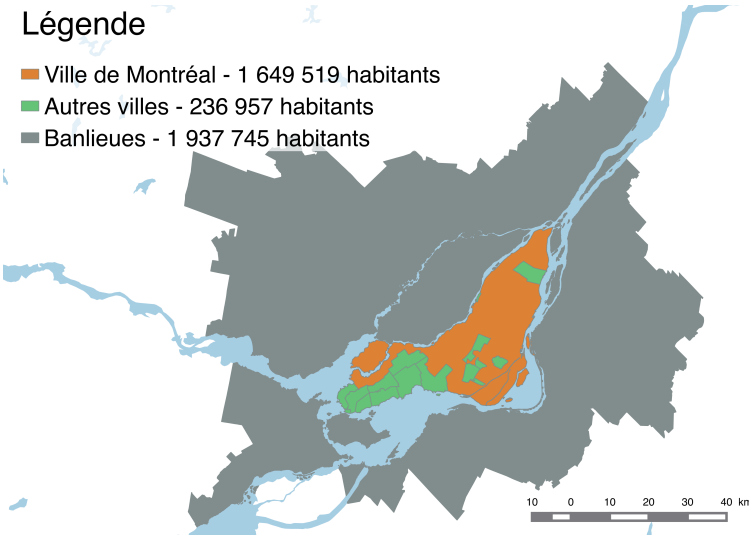


Figure 3-1 : Population de la région métropolitaine de Montréal

Afin de cerner la place et le rôle qu'occupe le taxi dans le système de transport de la ville, il est nécessaire de connaître l'environnement dans lequel il opère. Il est question ici d'informer sur les différents réseaux de transport présents sur l'île ainsi que sur les habitudes de déplacements des résidents de la région métropolitaine.

3.1.1 Réseau de transport

Le réseau routier est le plus étendu des réseaux. Il comprend environ 4 050 km de routes (Ville de Montréal). La plus grande entrave à ce réseau est le Mont-Royal, situé légèrement à l'est du centre de l'île ; une seule route le traverse. Les accès pour entrer ou sortir de la ville sont restreints par les 20 ponts à voies routières et ferrées traversant le fleuve Saint-Laurent et la rivière des Prairies. La figure 3-2 montre les autoroutes, les voies ferrées ainsi que les accès à l'île.

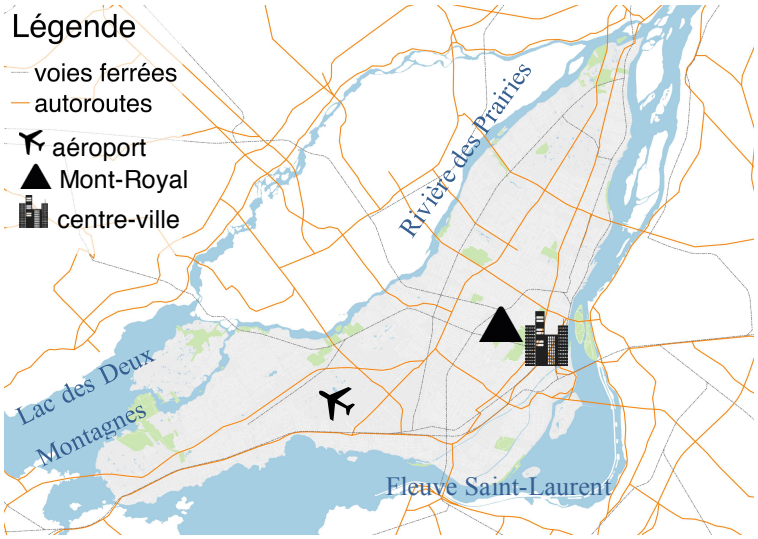


Figure 3-2 : Principaux lieux et axes routiers de l'île de Montréal

Un réseau de transport public est aussi implanté à Montréal. Celui-ci comprend 8 926 arrêts d'autobus, 68 stations de métro ainsi que 6 lignes de trains de banlieue (Société de transport de Montréal, 2015, 2016).

L'aéroport international Pierre-Eliott Trudeau est présent à même l'île de Montréal. Situées dans l'ouest de la ville, seules une ligne d'autobus ainsi que des autoroutes relient l'aéroport au centre-ville. Cette situation par rapport au système de transport encourage donc l'utilisation du taxi ou d'un véhicule privé.

3.1.2 Déplacements

Les habitudes de déplacement des résidents de la grande région de Montréal sont étudiées tous les cinq ans, à travers l'enquête Origine-Destination menée par un consortium d'institutions de transport de la région. La plus récente date de l'automne 2013. Cette enquête recueille les informations quant à la personne, au ménage ainsi qu'aux déplacements de chaque individu du ménage durant la journée ouvrable précédant l'entrevue. Provenant de cette enquête, quelques informations pertinentes pour l'étude de cas sur les taxis sont présentées ici (Agence Métropolitaine de Transport, 2015).

Le portrait général des déplacements ayant l'île de Montréal pour origine et destination montre qu'une grande concentration d'emplois et de services se trouve au centre-ville puisque 12,3 % des déplacements se destinent vers ce petit territoire. Cela est encore plus remarquable pour l'heure de pointe du matin (entre 7 h et 9 h), où 19,4 % des déplacements s'y dirigent. Les déplacements internes, c'est-à-dire ceux provenant et se destinant dans le même arrondissement, représentent 36,9 % des déplacements.

La figure 3-3 représente les parts de chaque mode dans les déplacements quotidiens. L'automobile est le mode de transport le plus utilisé sur l'île, avec une part modale de 53,5 %. Les déplacements ayant pour origine et destination la banlieue, où le service de transport public est moins présent, se font majoritairement en automobile, plus précisément à 80,9 %.

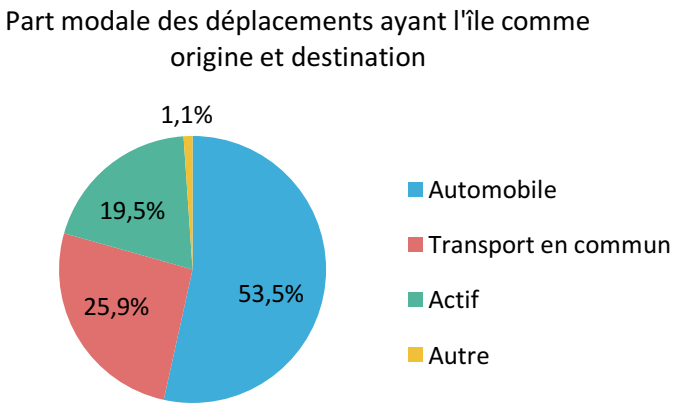


Figure 3-3 : Parts modales des déplacements sur l'île de Montréal

3.2 Industrie du taxi

Le BTM, dirigé par le Service de Police de la Ville de Montréal (SPVM), fut créé en 1986 et effectif dès 1987 (Bureau du Taxi de Montréal, 2015). Il a pour mandat d'œuvrer en lien direct avec l'industrie du taxi. En 2003, il étend sa juridiction aux activités de remorquage et devient le Bureau du Taxi et du Remorquage (BTR). Son mandat est ainsi modifié : il doit veiller à l'amélioration des services de taxi et de remorquage ainsi qu'au développement des industries concernées. Pour cela, ses champs d'action portent sur la sécurité des citoyens, la standardisation des véhicules et de l'‘équipement utilisés, la réception et le traitement des plaintes, l'encadrement des opérations de remorquage privé, l'émission et le renouvèlement des permis ainsi que le respect de la loi et des règlements.

En février 2013, suite aux recommandations de la Commission permanente du conseil municipal sur le transport, la gestion des infrastructures et l'environnement ainsi que la Commission du conseil d'agglomération sur l'environnement, le transport et les infrastructures, une société paramunicipale et désormais indépendante du SPVM est créée : le Bureau du taxi de Montréal (BTM). Il a pour mandat prioritaire le développement de l'industrie du taxi, tout en conservant son rôle de guichet unique, d'application règlementaire et de concertation. Son conseil d'administration est composé de représentants du grand public, des différentes clientèles et des différents intervenants présents dans l'industrie du taxi.

3.2.1 Structure organisationnelle

Les partenaires institutionnels du BTM au niveau de la règlementation sont au nombre de trois (Bureau du Taxi de Montréal, 2011). Le Ministère des Transports, de la Mobilité durable et de l'Électrification des transports (MTMDET) élabore les politiques de transport concernant entre autres les taxis. La Commission des Transports du Québec (CTQ) fixe les tarifs, définit les agglomérations d'opération et limite le nombre de permis en circulation. La Société de l'Assurance Automobile du Québec (SAAQ) délivre les permis de chauffeur de taxi et s'assure de la condition mécanique des véhicules. Toutefois, les activités concernant l'île de Montréal sont en majeure partie assurées par le BTM.

Les sources de revenus du BTM proviennent des droits annuels de permis, des contrats avec la SAAQ et Aéroports de Montréal, des revenus publicitaires et des vignettes. La compagnie FinTaxi

ainsi que les Caisses Desjardins sont les plus grands bailleurs de fonds pour l'acquisition des permis de taxi.

Les intervenants directs dans l'industrie du taxi sont l'intermédiaire en service, le propriétaire du véhicule et le chauffeur. L'intermédiaire en service offre des services de répartition des appels et de publicité aux propriétaires de taxi. Il doit être inscrit et payer des frais à la CTQ. Le propriétaire est celui qui possède le véhicule. Il peut conduire lui-même sa voiture, s'il possède un permis de conduire pour taxi, ou bien il peut autoriser quelqu'un qui possède un tel permis à conduire son véhicule. Le propriétaire doit acquérir, à la CTQ, un permis d'opération dans une agglomération précise de Montréal et l'afficher à bord du véhicule.

3.2.2 Services de taxi

Le service régulier de taxi est celui pour lequel les véhicules sont des berlines et le chauffeur a suivi les formations de base. Ce service est le plus utilisé parmi tous les services de transport de l'industrie du taxi.

Le service en transport adapté vise la clientèle à mobilité réduite. Plusieurs véhicules sont donc équipés de rampes facilitant l'accès aux fauteuils roulants. Les chauffeurs suivent une formation spécialisée pour l'utilisation de cet équipement et le service aux personnes à mobilité réduite. La STM et d'autres sociétés de transport en commun offrent aussi ce service à ses usagers pour le même tarif que leur service de transport. Pour cela, elle utilise des minibus adaptés ou bien fait affaire avec des intermédiaires en services de taxi adaptés.

Le service de taxi collectif est offert par la STM, en collaboration avec certains intermédiaires en service régulier. Il consiste à desservir les zones les moins denses de l'île de Montréal, où l'ajout d'une ligne d'autobus serait non rentable. Les taxis se synchronisent avec les arrivées et départs d'autobus ou de trains et suivent un trajet prédéterminé. Les titres de transport de la STM sont les droits de passages adoptés dans ce service.

Le service de taxi partage a été instauré à la fin octobre 2013 grâce à l'association entre le Jour de la Terre et Taxi Diamond. Ce service permettait aux clients de partager leur taxi avec d'autres personnes dont la course emprunte un trajet similaire. La demande d'un de ces véhicules se faisait par internet ou par une application pour mobile. Le prix de la course était affiché dès la commande du taxi. Ce service était une version bêta et il n'est plus en fonction depuis la fin 2014.

Le service de limousine est un service de taxi haut de gamme. Des compagnies autres que les intermédiaires en taxi régulier s'occupent de fournir ce service de transport. Ces véhicules peuvent être loués sur une longue période.

3.2.3 Règlementation

La CTQ règlemente l'industrie du taxi de Montréal sur trois principaux aspects : les agglomérations d'opération, les permis et les tarifs. Ce sont ces aspects qui impactent particulièrement les activités des taxis sur l'île.

L'île de Montréal est divisée en trois agglomérations distinctes telles que présentées à la figure 3-4. Chaque permis de taxi est lié à une de ces agglomérations. Ce dernier est contraint à prendre ou déposer un client dans l'agglomération de son permis. Trois formes de courses sont ainsi possibles :

- Prendre un client dans sa zone et le déposer dans sa zone ;
- Prendre un client dans sa zone et le déposer à l'extérieur de sa zone ;
- Prendre un client à l'extérieur de sa zone et le déposer dans sa zone.

Un taxi ne peut donc jamais accepter un client qui le hèle dans la rue lorsqu'il n'est pas dans son agglomération. Cette méthode d'opération vise autant les véhicules détenant un permis régulier (assurant principalement le service régulier de taxi) que ceux détenant un permis restreint (assurant principalement le service adapté de taxi). Toutefois, les taxis possédant des permis restreints ont plus de liberté que les taxis réguliers. Ils peuvent opérer dans toutes les agglomérations pour lesquelles leur intermédiaire en service est enregistré, ainsi que celles pour lesquelles aucun permis restreint n'est délivré.

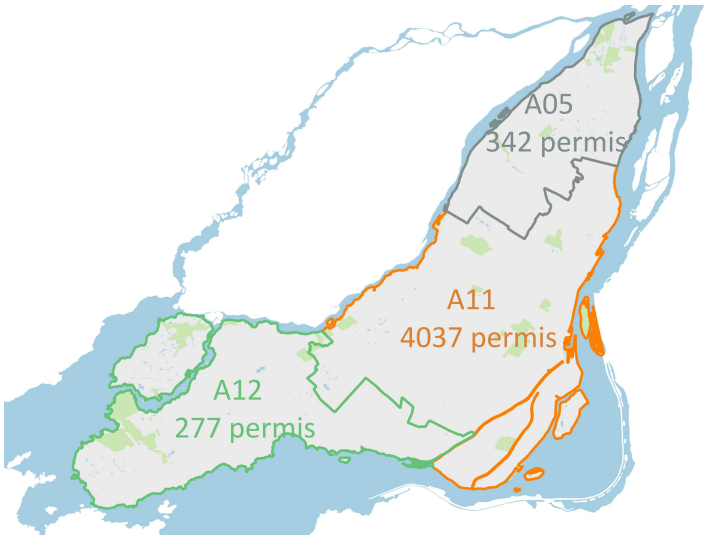


Figure 3-4 : Agglomérations et nombre de permis de taxi sur l'île de Montréal (Bureau du Taxi de Montréal, 2015)

Le nombre de permis émis dans chacune de ces agglomérations est indiqué à la figure 3-4 pour un total de 4 656 taxis œuvrant sur l'île, en date d'avril 2015. Parmi ce nombre, 219 sont des permis restreints. Le nombre de permis peut varier dans le temps en fonction des décisions de retrait ou d'émission de la CTQ. La plupart de ces taxis sont rattachés à un des 19 intermédiaires en services de taxi, les autres sont des chauffeurs de taxi indépendants, travaillant à leur compte. Des permis spéciaux pour prendre un client sans réservation à l'aéroport Pierre-Elliot-Trudeau sont tirés à chaque année parmi les propriétaires souhaitant participer au tirage. 310 permis spéciaux ont été attribués en octobre 2015 : 41 pour A05, 217 pour A11 et 52 pour A12.

En date du 31 décembre 2015 (Bureau du Taxi de Montréal, 2015), 10 353 chauffeurs de taxi sont enregistrés au BTM. Ils sont composés à 98,8 % d'homme et la moyenne d'âge est de 51,39 ans. De nombreux postes d'attente sont répartis sur le territoire montréalais permettant aux véhicules de s'y stationner en attendant leur prochain client. Il y a 29 postes communs à plusieurs intermédiaires, 120 postes privés et 306 postes publics pour un total de 455 postes.

La dernière révision des tarifs pour taxi s'est effectuée le 13 janvier 2012 (Bureau du Taxi de Montréal, 2015). Un tarif distinct de 40,00 $ est attribué pour les courses de l'aéroport Trudeau vers la zone du centre-ville et inversement. De plus, un prix minimum de 17,00 $ est exigé aux courses ayant comme point de départ l'aéroport Trudeau. Le tarif d'une course régulière de taxis est le suivant :

$$3{,}45\ \$\ au\ moment\ de\ la\ prise\ en\ charge\ du\ client + 1{,}70\ \$\ par\ kilomètre\ parcouru + 0{,}63\ \$\ par\ minute\ d'attente$$

Ainsi, un tarif de base est chargé lorsqu'un client entre dans le taxi. S'ajoute à cela un tarif par kilomètre parcouru qui devient un tarif par minute lorsque le taxi se déplace à très faible vitesse ou lorsqu'il ne bouge pas.

CHAPITRE 4 MÉTHODOLOGIE

Afin de développer des indicateurs de performance et de suivi d'une flotte de taxis, il est nécessaire de créer un système d'information qui regroupe plusieurs données pertinentes, dont les données opérationnelles, et qui permet le traitement rapide de l'information. D'abord, les données disponibles ayant un lien avec le système de taxi sont détaillées. Puis, ce qui a trait au système d'information, soit le type de base de données et ses objets, est décrit. Finalement, le traitement de standardisation, d'importation et de nettoyage des données est présenté. La figure 4-1 montre un aperçu de ce qui est couvert dans ce chapitre.

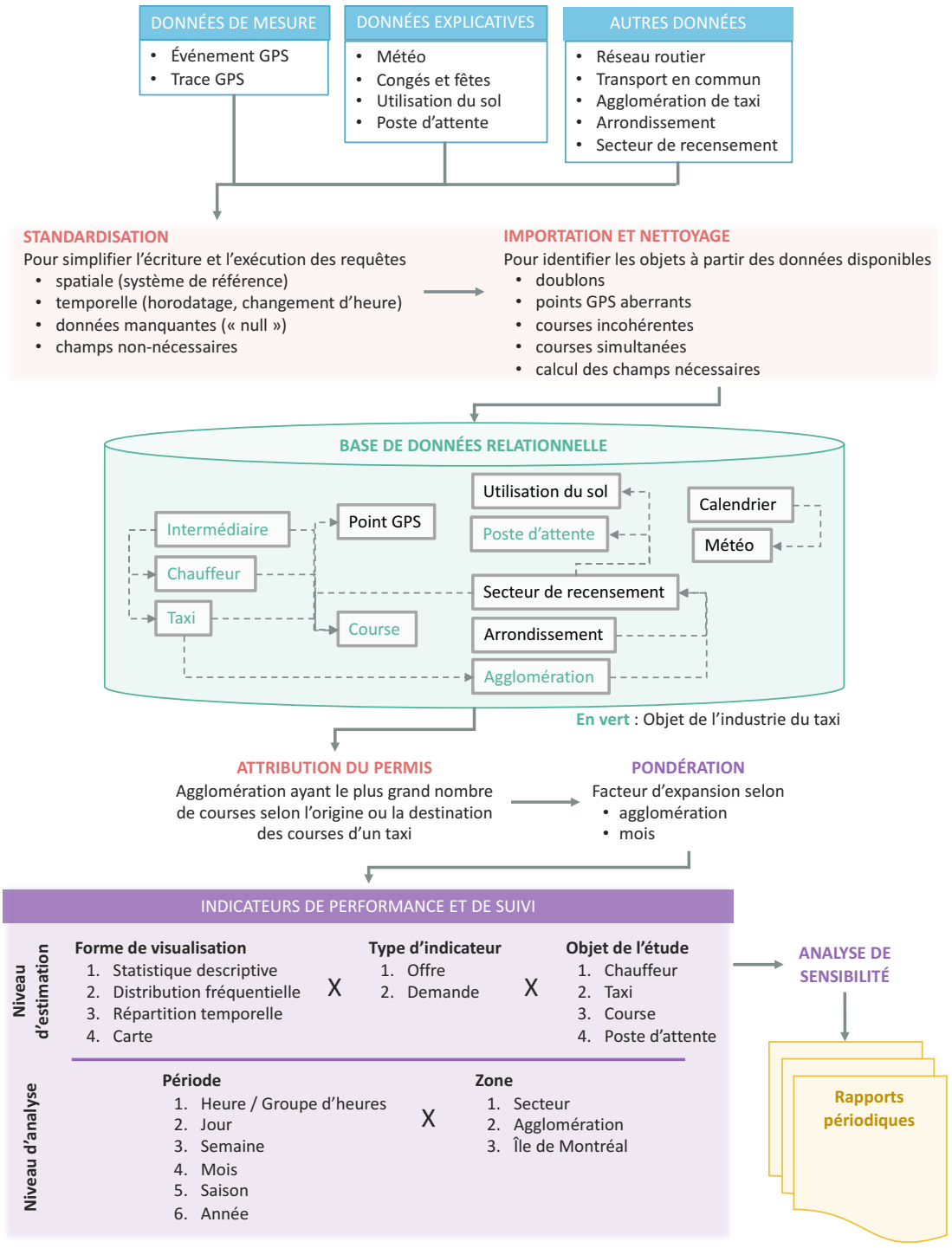


Figure 4-1 : Schéma méthodologique

4.1 Données

De nombreux ensembles de données peuvent aider à la compréhension d'un système de transport par taxi. Ces ensembles sont regroupés en trois types de données : les données de taxi, les données explicatives ainsi que les autres données. Une description complète incluant le format dans lequel elles sont disponibles est nécessaire pour mieux les intégrer au système d'information. Puisque l'étude de cas de cette recherche se produit à Montréal, les données disponibles par rapport à cet endroit seront décrites à titre d'exemple.

4.1.1 Données de taxi

Les données de taxi proviennent des intermédiaires en services récoltant de l'information sur leurs opérations. Afin d'évaluer les activités de la flotte de véhicule, seules les informations sur les courses et les véhicules sont recueillies. Les services de taxi régulier et une partie de taxi adapté sont couverts par cette information. Il faut cependant noter que les données varient d'un intermédiaire à l'autre, ne présentant donc pas toutes la même information.

Les données sur les courses sont liées aux appels des clients au centre de répartition et à leur point d'origine et de destination. Cette information recueillie par les répartiteurs (principalement pour le transport adapté) ou par des dispositifs GPS (principalement pour le transport régulier) est listée au tableau 4-1. Le répartiteur enregistre un endroit et un moment uniquement pour le point de départ et d'arrivée tandis que les dispositifs GPS le font pour tous les évènements composant la course. Un évènement est une action effectuée durant la course, par exemple lorsque le chauffeur arrive au lieu de départ du client ou lorsqu'il arrête le taximètre. Les caractéristiques spécifiques notées sont principalement utilisées par le service de transport adapté par taxi.

Tableau 4-1 : Information sur les courses

PRISE D'INFORMATION PAR LE **RÉPARTITEUR**	PRISE D'INFORMATION PAR LE DISPOSITIF **GPS (*coordonnées, horodatage*)**
Identifiant de la course	Identifiant de la course
Commande (*horodatage*)	Horodatage d'envoi de la commande aux chauffeurs
Départ (*adresse, horodatage, caractéristique spécifique*)	Taxi hélé sur la rue
Arrivée (*adresse, horodatage, caractéristique spécifique*)	Commande acceptée par le chauffeur
Identifiant du chauffeur	Commande refusée par le chauffeur
Caractéristiques spécifiques du client	Commande annulée
	Client ne s'est pas présenté
	Chauffeur arrive au lieu d'origine
	Taximètre démarré
	Taximètre arrêté
	Paiement complété

Les données sur les taxis sont en fait leurs traces GPS. Un point GPS est la localisation dans le temps et l'espace d'un véhicule. Il est enregistré à une fréquence variant pour chaque intermédiaire et permet ainsi de retracer le chemin d'un taxi dans le temps et l'espace. Les informations recueillies regroupent l'identifiant du point GPS, les coordonnées, l'horodatage, l'identifiant du véhicule, la vitesse instantanée, l'azimut, l'identifiant du chauffeur, le statut du chauffeur et quelques fois l'identifiant de la course. Le champ qui assure le lien entre les courses et les traces GPS varie selon l'intermédiaire ; ce peut être l'identifiant de la course tout comme l'identifiant du véhicule.

Plusieurs échantillons de quelques jours de données fournis par les différents intermédiaires sont examinés dans le cadre de cette recherche. Cependant, seules les données de l'intermédiaire en service Taxi Diamond (TD) sont disponibles en continu. Ce sont donc ces dernières qui sont utilisées pour la suite du projet puisqu'elles permettent une analyse plus approfondie. Les tableaux A-1 et A-2, présentés à l'annexe A, exposent l'information fournie par Taxi Diamond.

Taxi Diamond est le plus grand intermédiaire en service de Montréal, opérant près de 25 % des taxis principalement dans l'agglomération 11, ce qui assure une bonne représentativité des activités du centre de l'île. Ses données proviennent entièrement de dispositifs à bord des 1 200 véhicules de sa flotte. Ainsi, aucune intervention humaine n'est pratiquée dans la prise de données. L'information reçue à propos des courses est celle de la colonne de droite du tableau 4-1. Quant aux points GPS, ceux-ci sont enregistrés à tous les 600 mètres ou 5 minutes avant le printemps 2015, puis à chaque 500 mètres ou 1 minute depuis ce temps. L'identifiant de la course est le champ qui relie les deux ensembles de données. Il est à noter que la majeure partie du transport adapté n'est pas indiqué dans les données, donc un taxi qui n'est pas disponible peut en fait être sur une course adaptée.

4.1.2 Données explicatives

Les données explicatives représentent l'ensemble des données pouvant expliquer de façon directe les variances dans les activités de taxis. Elles proviennent de sources variées puisque de nombreux facteurs impactent la demande et l'offre de taxi. Les variables utilisées dans ce projet de recherche sont : les conditions météorologiques, les jours de congés et fêtes, l'utilisation du sol ainsi que la présence de postes d'attente.

Les conditions météorologiques sont enregistrées à plusieurs stations sur le territoire desservi par les taxis. Pour le cas de Montréal, Environnement Canada publie ces données. Grâce à ces stations météorologiques réparties dans la région, l'information sur la température, l'humidité, la pression atmosphérique, le vent, les précipitations de pluie et de neige ainsi que la visibilité est recueillie. Puisque la météo et les activités de taxis sont très variables dans le temps, il est nécessaire de collecter les données à une forte fréquence. Il est possible de les enregistrer à un intervalle de 15 minutes, toutefois ces dernières ne proviennent que de la station de l'aéroport. Une agrégation à la journée est aussi faite pour fin d'analyses plus générales.

Le calendrier permet de cibler les dates des jours fériés et jours de fête. Ces jours sont atypiques en termes de mobilité des personnes puisque ce ne sont pas des jours de travail ou de fin de semaine habituels. Le Gouvernement du Canada rend cette information disponible pour chaque année.

Les données sur l'utilisation du sol se trouvent sous forme de polygones couvrant un secteur précis. Ce secteur est caractérisé par une des 11 classes d'utilisation (résidentielle, commercial, etc.), le

nombre d'étages, le nombre de logements, l'année de construction ainsi que la valeur du terrain et du bâtiment. La CMM s'occupe de la gestion de ces données et les met à jour environ à chaque 2 ans. La dernière mise à jour date de 2014.

Les postes d'attente, où les taxis stationnent dans l'attente de leur prochain client, sont répartis partout sur l'île de Montréal. Les postes publics et privés sont déterminés respectivement par le BTM et les intermédiaires ayant des ententes avec certaines compagnies désirant avoir un espace pour les taxis sur leur propriété. L'information sur les postes d'attente de taxis comprend l'identifiant du poste, ses coordonnées, le nombre de places disponibles, son type (privé, publique ou commun) ainsi que son statut (actif, temporaire ou fermé). C'est le BTM qui fournit ces données et celles-ci datent d'avril 2015.

Une information difficile à obtenir, mais très intéressante pour l'analyse est celle de tous les évènements culturels, touristiques et d'affaire. Ce type d'évènement rassemble une grande quantité de gens au même endroit ce qui peut produire une demande importante en taxi.

4.1.3 Autres données

Les autres données sont complémentaires aux deux types de données précédentes. Elles servent principalement à établir une base d'information afin de fournir des connaissances primaires sur l'environnement dans lequel opèrent les véhicules de taxi. Ainsi, les données retrouvées dans ce groupe concernent les infrastructures de réseau de transport et les limites territoriales.

Les infrastructures de transport englobent tous les types de réseaux servant à la mobilité des personnes. Le réseau routier est l'ensemble des voies sur lesquelles circulent les automobiles. Chaque entité est un tronçon de route délimité par deux intersections. Il est caractérisé par un type de voie, un nom, une direction, un sens, une classe et la ville dans laquelle il se situe. Le réseau ferroviaire montre l'emplacement des voies ferrées sur le territoire. Le réseau de train de banlieue, de métro et d'autobus réunit les arrêts ainsi que les trajets empruntés par les véhicules de chaque réseau. Leur principale caractéristique est le numéro de la ligne de transport y passant. Ces données proviennent de l'AMT et de la STM à travers ses données GTFS. Le réseau cyclable, quant à lui, rassemble toutes les voies réservées aux vélos, qu'elles soient séparées physiquement ou pas de la circulation. Chaque ville s'occupe de son réseau cyclable et ce sont donc elles qui produisent les

données. Finalement, le réseau piétonnier représente les voies dédiées aux piétons, c'est-à-dire tous les trottoirs ou chemins dans les parcs.

Les limitations territoriales sont les différents découpages d'un territoire. Les agglomérations de taxi définissent la zone d'opération liée à un permis de taxi. Il existe trois agglomérations sur l'île de Montréal, chacune ayant un nombre limité de permis. Leurs frontières n'ont pas été modifiées depuis leur création. Les secteurs de recensement sont relativement stables au Canada. Ces secteurs sont composés de 2 500 à 8 000 habitants et appartiennent à une région métropolitaine dont le noyau est formé de plus de 50 000 résidents. Ils sont utilisés comme segmentation lors du recensement et de l'enquête nationale auprès des ménages à chaque cinq ans. Les plus récents ajustements ont été effectués en 2011, lors du dernier recensement. À cette date, 531 secteurs forment l'île de Montréal. Les arrondissements sont les divisions qu'utilise la ville de Montréal sur son territoire. Chacun a une mairie qui gère certains aspects publics non pris en charge par la Ville. Un arrondissement correspond à un regroupement de secteurs de recensement. Les limites d'arrondissement utilisées dans ce projet ont été mises à jour en 2016.

4.2 Système d'information

Le développement du système d'information sur le taxi est complexe dû à la quantité colossale de données ainsi qu'aux nombreux objets à considérer. Ainsi, il faut soigneusement choisir le type de base de données permettant le traitement rapide de l'information, définir avec précision les objets du système de taxi, normaliser les différents ensembles de données pour les incorporer dans une base de données unique et traiter les données afin d'en écarter les cas problématiques.

Le système d'information développé ici est un système d'information de gestion puisqu'il reçoit un volume élevé de données, les traite avec des analyses de base et des modèles simples pour, finalement, produire des rapports sommaires (Beynon-Davies, 2013).

4.2.1 Type de base de données

Une base de données relationnelle est utilisée pour monter le système d'information sur les taxis. Cette approche dérivée de l'algèbre relationnelle est axée sur la décomposition du système en entités simples et sur les associations entre ces entités (Hernandez, 2013). Une entité représente un objet matériel ou immatériel et est définie selon différents attributs qui prennent la forme de

colonnes dans la table de données. Afin de préserver l'unicité de chacun des enregistrements de la table, l'entité doit se doter d'une clé primaire. La clé primaire peut être une colonne unique ou bien la combinaison de plusieurs colonnes.

Les associations entre les entités sont représentées par les liens entre les tables de la base de données. La clé primaire d'une table assure la communication en se présentant comme clé étrangère dans une autre table. Les liens peuvent prendre trois cardinalités distinctes soit : relation de zéro à plusieurs, relation d'un à plusieurs ou relation de plusieurs à plusieurs. Dans le dernier cas, une table supplémentaire est créée dont l'enregistrement unique est défini par l'union des deux clés primaires.

Un tel type de base de données permet de structurer, conserver l'intégrité et manipuler les données. En effet, la structure de la base de données est fixée par les relations valides entre les entités. L'intégrité des tables et des colonnes est préservée grâce aux contraintes liées à la définition des données. La manipulation des données est gérée avec des opérateurs définis par la structure.

Une base de données relationnelle est avantageuse ici (Faucher, 2013). D'abord, elle contient peu de redondance des données dans une table pouvant alourdir les opérations sur les données. En normalisant les tables selon une des trois formes normales, elle peut même réduire la redondance au plus faible niveau. De plus, l'ajout d'informations ainsi que leur mise à jour est plus facile puisqu'une seule table est à modifier ; les liens assurent l'association de l'information entre les tables.

4.2.2 Objets et entités

Les objets du système de transport par taxi sont les éléments permettant de définir les opérations des taxis soit : le taxi, le chauffeur, l'intermédiaire, la course, le poste d'attente ainsi que l'agglomération. Ce sont donc sur ces objets que des indicateurs peuvent être développés. Les entités de la base de données sont les tables qui en composent la structure. Elles sont indispensables puisqu'elles recèlent l'information nécessaire au développement des indicateurs ainsi qu'aux analyses prévues. Ainsi, ces entités sont le point GPS, le calendrier, la météo par jour et par 15 minutes, le secteur de recensement, l'arrondissement et l'utilisation du sol.

4.2.2.1 Objets du système de transport par taxi

Le taxi est une automobile, conduite par un chauffeur, destinée au transport de personnes et de leurs bagages et pour laquelle est enregistré un permis de taxi. Il appartient à un seul propriétaire et peut être loué par différents chauffeurs qui le conduiront en alternance. L'information générale liée au véhicule est non nécessaire pour le développement des indicateurs, à l'exception de l'agglomération de son permis ainsi que de l'intermédiaire pour lequel il opère. Le taxi est un objet suivi dans le temps et l'espace grâce aux données GPS et de courses.

Le chauffeur est la personne conduisant le taxi. C'est un travailleur autonome ou salarié pouvant conduire différents véhicules de manière consécutive. La seule information à retenir concernant le chauffeur est l'intermédiaire pour lequel il travaille. Grâce à son identifiant unique dans les données GPS et les données de courses, il est possible de le suivre spatiotemporellement, comme c'est le cas avec le taxi.

L'intermédiaire en services de taxi est l'organisation s'occupant de répartir les commandes des clients aux véhicules disponibles sur le territoire. Il fait le lien entre le chauffeur et le client, indiquant au chauffeur où il doit se rendre pour ramasser le client. Plusieurs intermédiaires peuvent offrir leur service dans une même région. Les propriétaires de véhicules et les chauffeurs choisissent l'intermédiaire pour lequel ils désirent recevoir les services et ne sont pas limités à faire affaire avec un seul intermédiaire. Afin de les distinguer, l'unique information nécessaire dans la base de données est le nom de la compagnie.

La course est le trajet effectué par un taxi lorsqu'un client se trouve à bord, d'un point d'origine à une destination. Ainsi, une course est considérée dès le moment où le client entre dans le taxi jusqu'à ce que celui-ci paie le chauffeur et quitte le taxi. Il existe deux types de courses soit celles commandées d'avance via une plateforme de commande (centre de répartition téléphonique, application mobile, réservation par internet, etc.), soit celles hélées directement sur la rue. La figure 4-2 montre les différences entre ces types de courses. La course est aussi caractérisée par un chauffeur, véhicule et intermédiaire distinct. Toutefois, les courses effectuées dans le cadre du transport adapté peuvent nécessiter plus d'un chauffeur dans le but d'avoir un appui supplémentaire pour déplacer le client jusqu'au véhicule principal.

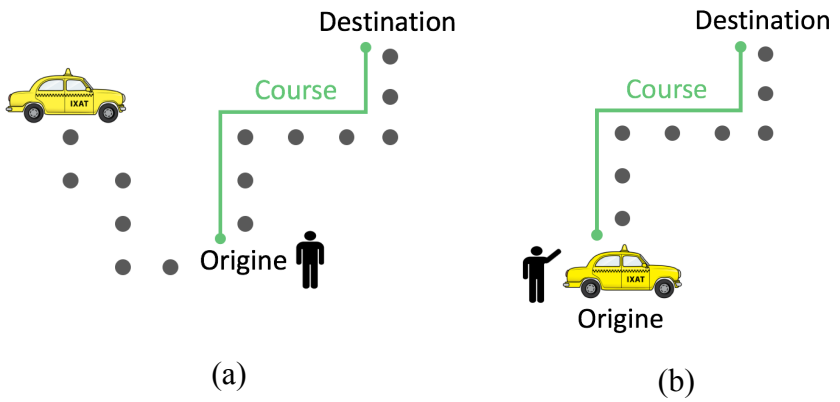


Figure 4-2 : Types de courses (a) commandées et (b) hélées

Le poste d'attente est l'endroit où les véhicules ont la possibilité de se stationner en attendant un prochain client. Il est composé d'un nombre limité de places et peut être public (n'importe quel taxi y a accès), privé (un intermédiaire paie pour qu'uniquement ses véhicules s'y stationnent) ou commun (seuls les véhicules de certains intermédiaires peuvent s'y garer). L'ordre d'arrivée au poste d'attente définit l'ordre d'attribution des courses. Ainsi, le véhicule étant le plus longtemps stationné au poste se trouve en avant de la file de voitures stationnées et se fait attribuer un client avant les autres.

L'agglomération est la zone dans laquelle le taxi est contraint à prendre ou déposer un client. La délimitation des agglomérations est instaurée, à Montréal, par la CTQ. Un nombre de permis précis est délivré pour chaque agglomération et la CTQ peut décider en tout temps d'en rajouter. Les chauffeurs et intermédiaires sont donc contraints à accepter les courses selon le permis associé aux taxis qu'ils opèrent.

4.2.2.2 Entités de la base de données

Tous les objets du système de transport par taxi se retrouvent comme entité dans la base de données. À cela s'ajoutent les entités qui complètent la structure des données et qui permettent de complémenter et segmenter les analyses. Le point GPS est l'entité centrale de ce système d'information. Grâce aux informations enregistrées à intervalle régulier pour chaque véhicule, les principaux indicateurs de performance sont développés. Les entités temporelles, c'est-à-dire le calendrier et les conditions météorologiques, ont un pouvoir explicatif sur la demande en taxi. Leur intégration avec les objets du taxi enrichit la compréhension des facteurs d'influence. Les entités spatiales, qui incluent le secteur de recensement, l'arrondissement et l'utilisation du sol, sont

principalement utilisées comme découpage simplifiant la visualisation des indicateurs. L'information sur la démographie et les infrastructures que contiennent ces entités appuie aussi le développement de la connaissance sur la demande en taxi.

4.2.3 Schéma de la base de données

Les interactions entre les entités forment la structure de la base de données de la figure 4-3. Les liens montrent la manière dont sont associées les entités, à savoir les relations de cardinalité qui caractérisent le nombre d'enregistrements impliqués. Dans le diagramme de la base de données de taxis, tous les liens sont de cardinalité un à plusieurs. Les attributs de chaque entité représentent l'information extraite des données brutes disponibles et leur format est indiqué. Les caractéristiques de chaque entité sont définies du tableau b- au tableau b-13 de l'annexe B. Afin de garder une traçabilité des données brutes, un attribut d'identifiant interne est créé pour chaque entité liée au système de transport par taxi.

Un grand avantage de ce type de structure est que l'ajout et la modification d'information y sont simples. Il suffit d'ajouter une colonne à une table existante pour intégrer une nouvelle information ou d'ajouter une table pour créer une nouvelle entité et de la connecter au reste de la structure en fonction des entités qui y sont associées. Dans l'optique d'évaluer la contribution du taxi au développement durable, de nouvelles tables telles le type de voitures ou les tournées du transport adapté et collectif peuvent facilement être intégrées à la base de données.

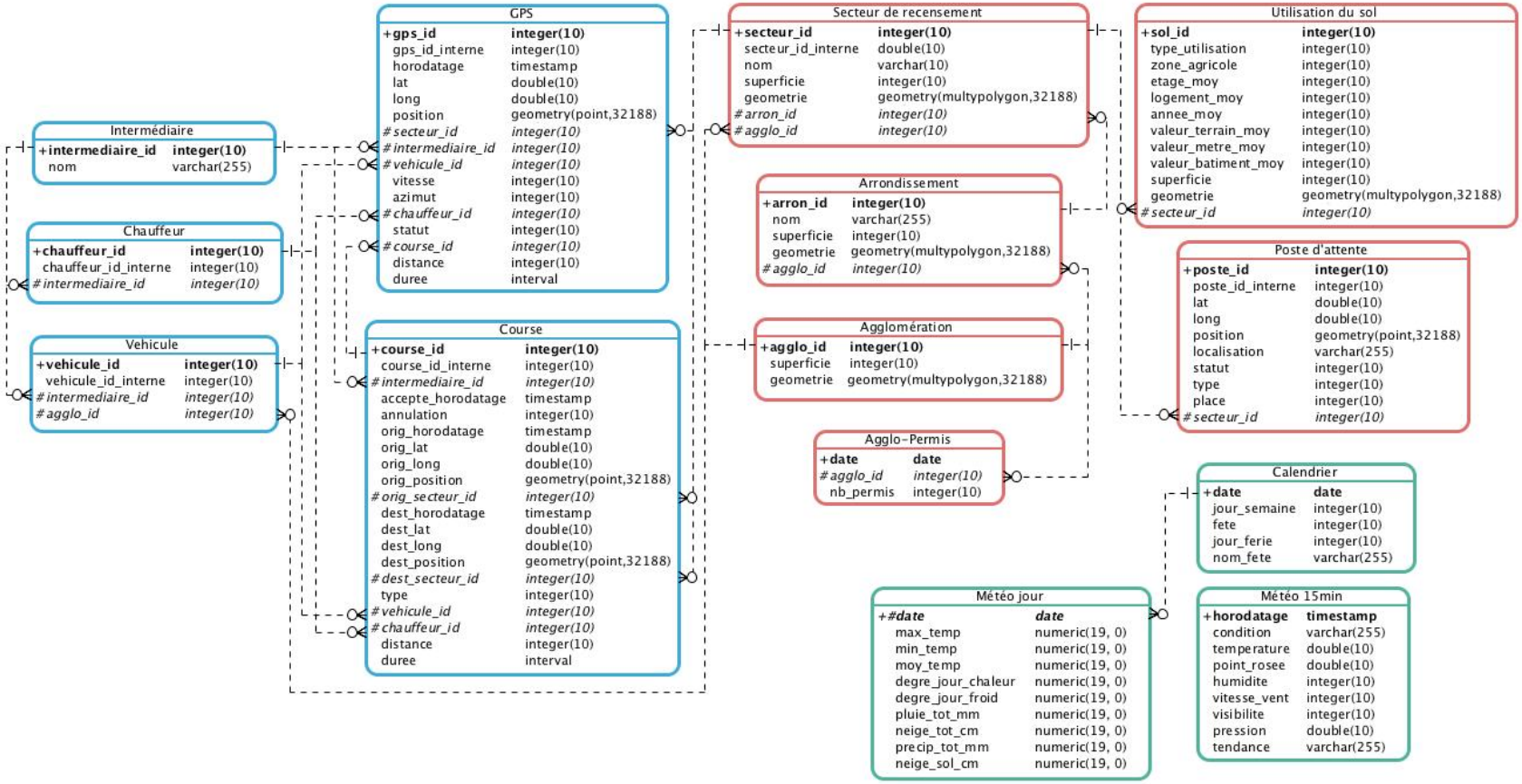


Figure 4-3 : Diagramme relationnel de la base de données de taxis

4.3 Traitement des données

4.3.1 Standardisation

Les données de taxis, provenant de plusieurs sources différentes, ne se présentent pas sous le même format. Il est donc important de leur faire suivre un processus de standardisation, afin de rassembler ces données dans une base de données unique permettant un traitement plus simple de l'information.

La méthode suivie pour standardiser les données consiste à créer une table unique pour chaque objet dans laquelle toutes les informations sont rassemblées et chaque instance est identifiée à l'intermédiaire dont elle provient. Même si la modification des données pour les adapter à la structure est complexe, cette méthode est utilisée puisqu'elle assure une bonne intégration des données (sans répétition) et accélère la durée d'exécution des requêtes sur l'information emmagasinée. Afin d'éviter un trop gros stockage d'information dans une même table, ralentissant par le fait même les performances du système, celles-ci sont partitionnées selon le principe d'hérédité (des tables enfants comportant les données et une table parent reliée à tous les enfants). Par exemple, une table est créée chaque mois pour recevoir les données GPS et est reliée à la table parent unique. Une requête sur la table parent parcourt donc les données GPS de tous les mois.

La standardisation implique la réalisation de plusieurs étapes permettant d'unifier les données d'une table à l'autre ainsi qu'à l'intérieur d'une même table. Les étapes à suivre pour ce projet sont les suivantes :

1. En ce qui concerne la localisation géographique, il est nécessaire de transformer toutes formes de localisation en coordonnées de latitude et longitude, puis de les convertir dans le système de référence spatial NAD83 MTM Zone 8 (EPSG 32188). Ce système de référence est celui attribué précisément à la zone de Montréal ce qui permet de diminuer la distorsion des surfaces sur le territoire de l'étude de cas. Les coordonnées sont en mètres ce qui simplifie le calcul des distances. Les requêtes spatiales sur les données peuvent ainsi être exécutées aisément.
2. En ce qui a trait au moment lié à la prise de données GPS, il faut s'assurer de combiner les champs d'année, de mois, de jour et d'heure afin de n'obtenir qu'un champ d'horodatage.

De cette manière, seule une colonne de la table de données est impliquée dans la demande d'information. De plus, le fuseau horaire doit être indiqué pour tous les champs d'horodatage de manière à ce que le système assimile le concept de changement d'heure ayant lieu le premier dimanche d'octobre ainsi que le second dimanche de mars.

3. Dans tous les cas, lorsqu'il y a des données manquantes, il faut qu'une valeur null la définisse. Ainsi, les requêtes sur un champ ne renverront pas de valeurs faussées par une donnée manquante fixée à zéro : ce qui est le cas, par exemple, pour l'identifiant de la course qui est fixé à zéro dans la table GPS des données brutes lorsqu'un point ne fait pas partie d'une course.
4. Taxi Diamond fournit ses données de manière continue par l'envoi d'un fichier par jour. À minuit, lorsque le fichier est extrait, certaines courses sont en train de se produire. Puisque celles-ci enjambent deux jours, elles apparaîtront donc dans les données de ces deux journées, ce qui crée des doublons dans la table des courses. Ces doublons sont repérables en vérifiant si tous les champs des différentes lignes de la table de données sont identiques. Une fois repéré, le doublon peut être supprimé.
5. Des imprécisions dans le dispositif de collecte de données à bord des véhicules engendrent des points GPS erronés. Si la couverture satellite est faible, par exemple le nombre de satellites disponibles n'est pas assez grand ou s'il y a trop d'obstacles empêchant les signaux d'atteindre le dispositif, ce dernier enregistre des valeurs erronées. En limitant les coordonnées à celles d'une zone autour de Montréal, cela permet de faire un tri des positions GPS. La zone carrée déterminée dans ce projet est assez large, incluant Toronto, New York ainsi que Sept-Îles, offrant ainsi la possibilité de capturer les courses interurbaines.

Une méthode plus précise de détection des points erronés serait à élaborer. En fait, une trace GPS peut contenir des points aberrants (un point ponctuel qui est décalé de la trace) ou décalés (à partir d'un point, toute la trace se poursuit et est décalée). Leur identification est complexe et fût tentée ici en utilisant un critère de distance et de vitesse maximale entre des points consécutifs. Toutefois, cette méthode ne fut pas concluante et c'est pourquoi un approfondissement est nécessaire.

Les données en général sont aussi normalisées de manière à ne retenir que les champs nécessaires au projet définis à la figure 4-3. Ainsi, certaines informations ne sont pas importées dans la base de données puisqu'elles n'apportent aucun bénéfice au développement des indicateurs.

4.3.2 Importation et nettoyage des données

L'importation des données dans les tables ainsi que le nettoyage de celles-ci correspondent au processus d'identification des objets du système. En effet, suite à la standardisation, les données ont atteint un format acceptable afin d'être insérées dans la base de données. Toutefois, des règles d'importation sont indispensables afin de s'assurer que tous les champs caractérisant l'objet ainsi que toutes les instances importées soient valides. Puisque chaque intermédiaire fournit ses propres données, le traitement de l'information est particulier de l'un à l'autre. Les données de taxis utilisées dans l'étude de cas proviennent uniquement de Taxi Diamond. Les règles d'insertion décrites dans cette section concernent donc cet intermédiaire.

La course est l'objet pour lequel de nombreuses analyses seront effectuées puisqu'elle représente la demande en taxi. Afin de détecter une vraie course parmi les données brutes fournies de l'intermédiaire, le processus décisionnel de la figure 4-4, résumant les règles d'importation, est suivi. La figure 4-4 montre aussi les valeurs que prennent les champs de la table selon les attributs bruts de la course (voir les tableaux A-1 et A-2 pour les attributs bruts et les tableaux B-1 à B-13 pour les attributs de la base de données).

Les principaux cas d'exclusion d'une course se dégageant du processus sont les suivantes :

- L'origine ou la destination de la course se trouve à l'extérieur de la zone de Montréal ;
- Une course annulée ne possède pas de moment où le chauffeur accepte la course ;
- Aucun véhicule et chauffeur n'est indiqué dans les points GPS de la course ;
- La distance ou la durée de la course est plus petite ou égale à 0 ;
- Une autre course se produit simultanément par le même véhicule.

Malgré ces contraintes, certaines courses semblent toujours illogiques. En prenant comme exemple la journée du 1er janvier 2015, près de 9 % des courses ont une durée de moins d'une minute et près de 7 % couvrent une distance de moins de 100 mètres entre l'origine et la destination. Un nettoyage plus poussé serait envisageable, avec une possibilité de considérer le secteur d'où provient la course. En effet, une course de petite distance s'effectuant au centre-ville semble plus valide qu'une même course s'effectuant à l'aéroport. Ainsi, la densité de population et de commerces jouerait un rôle dans la validation des courses.

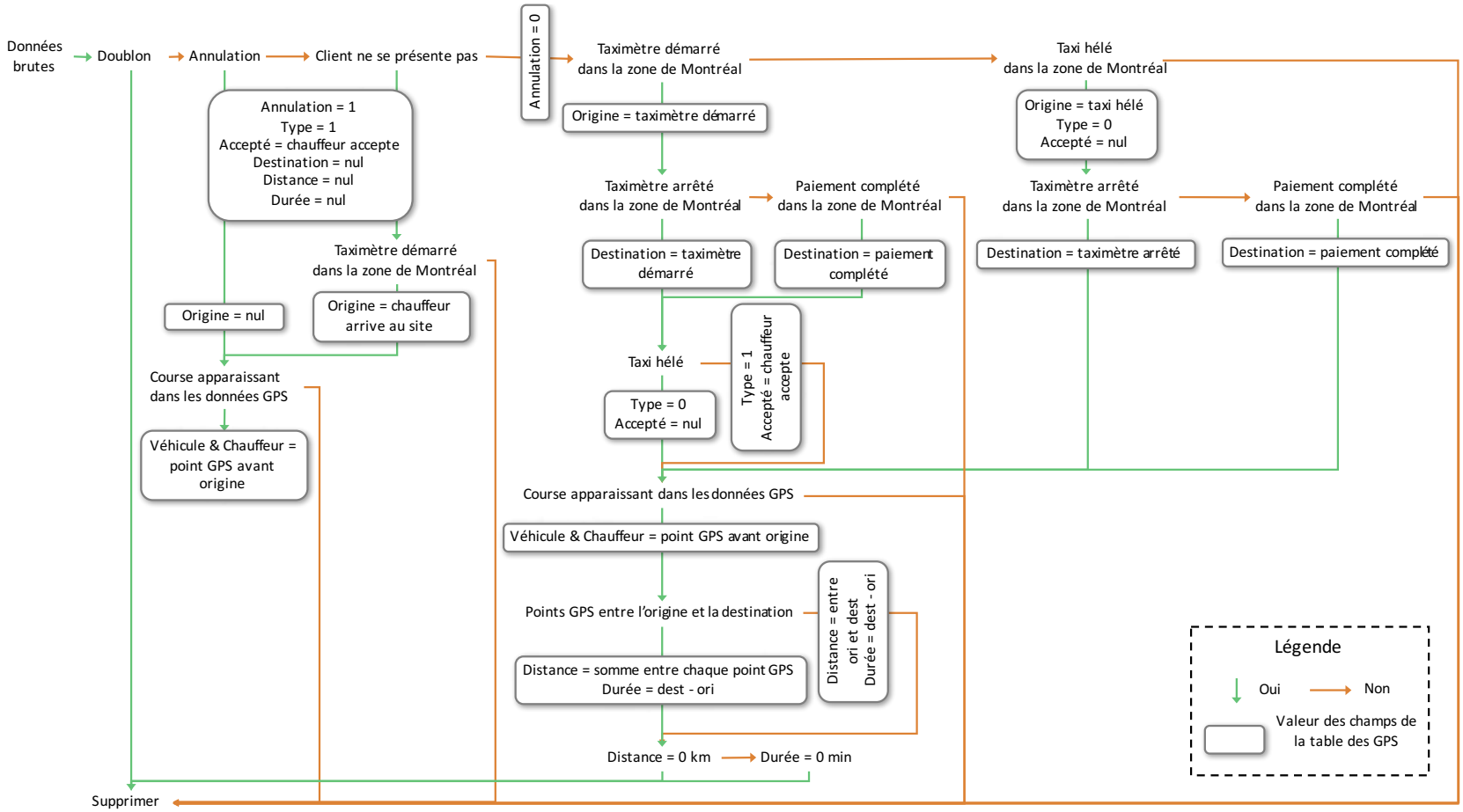


Figure 4-4 : Étapes d'importation et de nettoyage des données de courses brutes

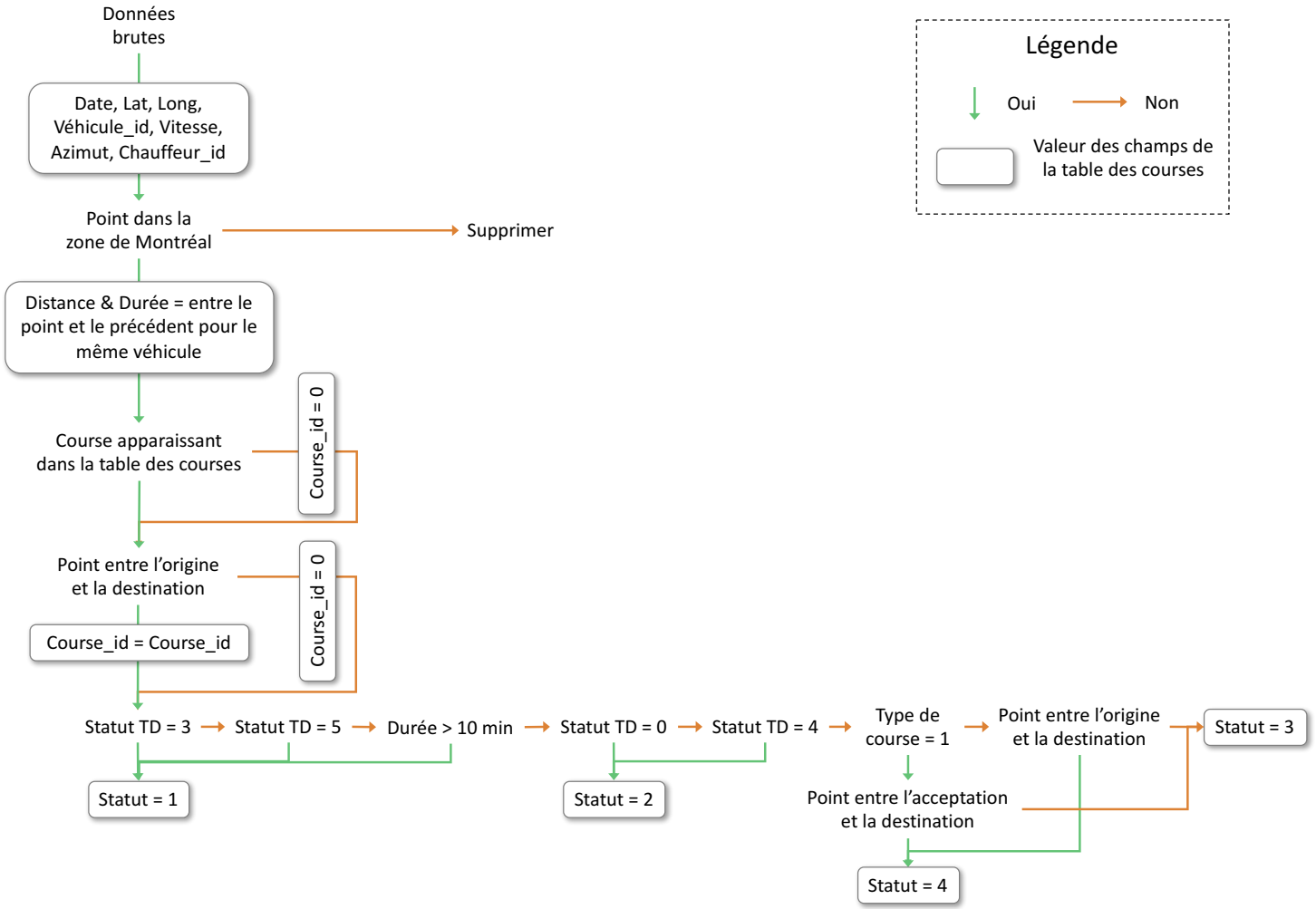


Figure 4-5 : Étapes d'importation et de nettoyage des données GPS brutes

Le point GPS est l'objet de base pour le calcul de plusieurs indicateurs. Afin d'aboutir à une table de données GPS valide, il faut avoir recours au processus décisionnel montré à la figure 4-5. La distance et la durée sont calculées pour chaque point, en fonction du point précédent pour le même véhicule. La détermination du statut du point GPS est principalement basée sur le statut déjà existant dans les données de l'intermédiaire. Seul le statut hors service est modifié de manière à considérer les instances ayant une durée de plus de 10 minutes avec la précédente. Ce critère est sélectionné grâce à l'analyse de la distribution des durées entre les points GPS. À partir de 10 minutes, la courbe de distribution s'abaisse drastiquement, ne laissant que 1 % des points au-delà de ce seuil, ce qui est logique selon la fréquence d'enregistrement (600 m ou 5 min) des dispositifs GPS.

Suite à l'importation des données GPS, un nettoyage doit avoir lieu. Les cas suivants montrent le type de modifications à apporter afin d'assurer une cohérence avec la table des courses.

- Imposer un identifiant de la course null au point GPS si cet identifiant ne se retrouve pas dans la table des courses ;
- Une course présentant un point GPS de statut hors service (statut=1) doit être supprimée de la table des courses ; suite à cela, le point GPS doit prendre un identifiant de course nul.

Les autres objets de la base de données sont bien moins complexes à identifier puisqu'ils ont des caractéristiques très similaires à leurs données brutes. Il ne suffit que de créer les champs utiles pour le calcul des indicateurs. La superficie ainsi que les champs géométriques sont ainsi rajoutés aux tables.

4.3.3 Attribution du permis

Les données fournies des intermédiaires ne contiennent pas l'agglomération dans laquelle peut opérer le taxi. Cette information est primordiale puisque le comportement d'un taxi dépend de son permis. Afin de le déterminer pour la période étudiée, la méthode suivante est élaborée. Il faut d'abord calculer le nombre de courses effectuées par le taxi pour chaque agglomération. Pour ce faire, il suffit de sommer les courses ayant une origine ou une destination dans celle-ci. Il est normal qu'une course débutant et finissant dans deux agglomérations différentes soit incluse dans la somme de chacune d'elles. Par la suite, il faut identifier l'agglomération ayant le plus grand nombre de courses pour qu'elle devienne celle du permis.

Deux limitations majeures engendrent toutefois un biais. D'abord, les véhicules possédant un permis restreint, pouvant prendre et déposer un client dans n'importe quelle agglomération de son intermédiaire et celles pour lesquelles aucun permis restreint n'est délivré, sont difficilement identifiables. Suivant la méthode d'attribution de permis, ils se voient associer l'agglomération dans laquelle ils opèrent le plus. Si cette agglomération fait partie de celles de l'intermédiaire en service, alors le taxi sera considéré au même ordre qu'un taxi régulier ; sinon, le taxi est retiré de l'analyse.

Par la suite, les véhicules détenant un permis spécial pour l'aéroport sont aussi difficiles à reconnaître. Ce permis donne au véhicule propriétaire l'accès au poste d'attente de l'aéroport (type de courses hélées). Les taxis ne possédant pas ce permis peuvent tout de même prendre des clients ayant commandé d'avance leur transport ainsi que déposer des clients. Ainsi, le critère d'origine des courses à l'aéroport ne peut être utilisable puisque ce n'est pas unique au permis spécial. De plus, si un taxi possédant un permis spécial ne fait que des courses depuis l'aéroport et non dans sa zone régulière, alors il se verra attribuer l'agglomération 12.

Bref, l'ajout par les intermédiaires de l'information sur le type de permis et de l'agglomération associée serait un atout majeur. Si cela ne s'avère pas possible, l'obtention de l'information sur un type de courses comprenant un type régulier, adapté et aéroport pourrait pallier cette problématique. Dans l'attente de ces nouveaux champs, une méthode approfondie pour la détection du permis spécial, prenant en compte le lieu précis du poste de taxis ainsi que la proportion de courses débutant à ce lieu, serait envisageable.

4.4 Indicateurs

La littérature n'est pas étoffée quant au développement des indicateurs du service de taxi. Diverses statistiques sont étudiées depuis maintes années, néanmoins elles concernent plutôt les aspects règlementés de l'industrie comme le nombre de permis émis, leur prix, les caractéristiques des véhicules ainsi que les rapports d'inspection. Produire un état des lieux des activités du taxi sur le territoire est un concept récemment envisageable grâce au potentiel qu'apportent les données GPS dans ce domaine. Ainsi, la conception d'indicateurs de base émerge tel qu'il a été montré au tableau 2-2.

Accéder à l'entièreté des données d'une flotte de véhicules est une démarche très laborieuse dans l'industrie du taxi. C'est pourquoi un échantillon de véhicules est utilisé et, grâce à des méthodes de pondération qui y sont appliquées, est en mesure de représenter la flotte. Dans le contexte montréalais, l'échantillon provient des véhicules opérés par Taxi Diamond et est pondéré pour représenter la flotte de l'île.

Suite à la mesure de la pondération, les indicateurs de base sont développés et amenés à l'échelle de l'ensemble de la flotte. Les formes de visualisation utilisées sont les statistiques descriptives, les distributions fréquentielles, les répartitions temporelles ainsi que les cartes. Afin de mieux structurer les indicateurs du système de taxi, une segmentation par objet est employée. Le chauffeur, le taxi, la course et le poste d'attente sont les principaux objets pour lesquels sont estimés les indicateurs. Ils sont analysés dans le temps (heure, jour, semaine et saison), dans l'espace (secteur et agglomération) et par rapport aux autres objets.

4.4.1 Pondération

La pondération est l'affectation d'un coefficient à une variable afin de modifier son incidence sur un résultat. Dans le but de produire, pour ce projet, un résultat représentant la flotte complète de taxis, le coefficient utilisé est un facteur d'expansion. Habituellement, ce facteur est mesuré selon les caractéristiques et les comportements mesurables de la flotte. Toutes les instances ayant ces mêmes aspects se voient attribuer le même facteur d'expansion.

Pour le cas de Montréal, seule l'information sur le nombre de permis par agglomération est disponible pour la population de taxis. Ainsi, le facteur d'expansion n'est mesuré qu'en fonction de cette information. Le type de permis (régulier, restreint ou spécial) n'est pas encore accessible, mais est une piste pour pondérer de manière plus précise.

La méthode de pondération est élaborée par agglomération et par mois. En effet, le nombre de permis émis est différent d'une agglomération à l'autre et varie dans le temps en fonction des décisions d'octroi ou de retrait de la CTQ. Ainsi, un taxi n'a pas la même incidence sur le résultat selon l'agglomération de son permis ainsi que le moment lorsqu'il est observé dans les données. L'équation 4-1 montre la manière de calculer le facteur d'expansion en fonction du nombre de permis. Ce facteur est appliqué aux véhicules faisant partie de la segmentation spatiale et temporelle établie.

$$facteur\ d'expansion = {}^{nombre\ total\ de\ permis}/_{nombre\ de\ permis\ observé} \quad \text{pour chaque agglomération et mois} \qquad \text{Équation 4-1}$$

La segmentation spatiale utilisée ne peut être que l'agglomération puisque c'est l'unique information spatiale disponible pour la flotte. Toutefois, la segmentation temporelle doit faire l'objet d'une réflexion. Le jour ne peut être utilisé pour la segmentation ; cela signifierait que tous les véhicules non observés dans l'échantillon de données pour cette journée précise travailleraient tout de même. Pourtant, il est logique de croire que le véhicule est en réparation ou simplement qu'il n'y a pas de chauffeur pour le conduire cette journée-là. La semaine semble plus appropriée puisqu'il est improbable de penser qu'un taxi ne sera pas présent sur les routes durant une semaine complète. Cependant, des cas rares comme de lourdes réparations ou l'achat d'un nouveau véhicule suite à un accident majeur pourraient entraîner le retrait du taxi pour une semaine. De plus, en raison des nombreux jours manquants dans les données de Taxi Diamond, une segmentation par mois est privilégiée.

Il est à noter qu'un taxi peut travailler pour plusieurs intermédiaires. Néanmoins, ces cas sont rares puisqu'un propriétaire ne désire pas nécessairement s'acquitter des frais de cotisation envers plus d'un intermédiaire. Puisque les données fournies sont anonymisées, il est difficile de capturer ces cas. Cela représente un défi d'analyse et complexifie le calcul des facteurs lorsque plus d'un intermédiaire est compris dans l'estimation des indicateurs. La méthode devra donc être raffinée.

4.4.2 Formes de visualisation

Afin de présenter les indicateurs, quatre formes de visualisation sont ciblées tel qu'en fait état la figure 4-6 : la statistique descriptive, la distribution fréquentielle, la répartition temporelle ainsi que la carte.

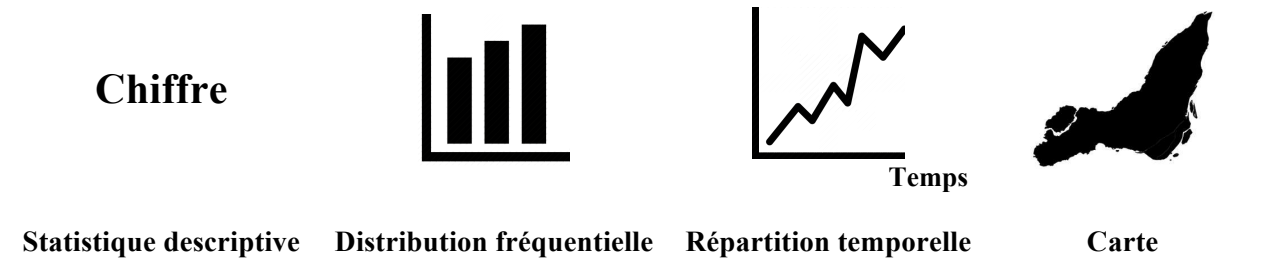


Figure 4-6 : Formes de visualisation

Premièrement, la statistique descriptive est en fait une représentation générale des données. Elle n'analyse pas en détail un objet, elle représente une caractéristique globale de l'objet à travers un chiffre. Les plus communément estimées sont le nombre, le pourcentage, la moyenne et l'écart-type (par exemple le nombre de taxis étudiés ou la moyenne de la distance parcourue par taxi accompagnée de son écart-type).

Deuxièmement, la distribution fréquentielle est le nombre de fois qu'apparaît un objet dans des catégories. La décomposition d'une caractéristique de l'objet selon les valeurs qu'elle peut prendre génère ces catégories. Elle est fréquemment représentée par des histogrammes ou des courbes. Dans le contexte du taxi, un exemple serait le nombre de courses distribuées selon des classes de distances.

Troisièmement, la répartition temporelle est utilisée pour une variable qui évolue dans le temps. Ainsi, la répartition est une évolution dans le temps d'un objet ou de ses caractéristiques. Elle prend communément la forme d'une courbe. Les courbes peuvent être segmentées selon une variable spatiale comme des zones ; celle-ci doit toutefois être limitée à un petit nombre de zones (par exemple les agglomérations), sinon les courbes seront confondues les unes par rapport aux autres.

Quatrièmement, la carte localise sur le territoire étudié les objets ainsi que leurs caractéristiques. Une division par zone assure la représentativité des données et est majoritairement utilisée dans ce projet. La carte est la forme la plus appréciée de visualisation puisqu'elle représente une manière très intuitive de comprendre un phénomène spatial et en le segmentant par période de temps, permet de montrer l'évolution du phénomène sur le territoire.

4.4.3 Objets d'étude

Dans une optique de planification, il est important de suivre les tendances globales d'évolution de la demande ainsi que les performances des acteurs du service de taxi. Les quatre objets du système de taxis pour lesquels sont développés les indicateurs de suivi et de performance sont le chauffeur, le taxi, la course ainsi que le poste d'attente.

4.4.3.1 Chauffeur

L'analyse du chauffeur porte principalement sur sa performance relativement aux courses effectuées. Cette dernière varie selon les habitudes de travail des chauffeurs. Ainsi, des indicateurs

sur leurs horaires de travail et sur le nombre de courses qu'ils accomplissent sont estimés. Des pistes de développement pour d'autres indicateurs plus détaillés sont aussi proposées.

Indicateur de quarts de travail

Dans un premier temps, il est possible de déduire quand les chauffeurs sont présents sur la route à travers leurs horaires de travail. Les moments où ils débutent et finissent déterminent les périodes où ils sont disponibles et pour lesquelles ils doivent travailler afin de gagner un revenu qui leur est acceptable.

Dans les données, le début et la fin d'un quart de travail sont identifiables grâce au changement de l'identifiant du chauffeur associé au véhicule. En effet, lorsque le chauffeur désire recevoir des appels, il doit se connecter au système par l'entremise d'une tablette. Au moment où celui-ci finit de travailler, il se déconnecte du système afin de ne plus recevoir de courses, permettant ainsi à un second chauffeur d'utiliser le véhicule, si c'est le cas. Pour chaque tranche de 30 minutes, la somme des débuts puis des fins de quarts de travail est calculée.

Un cas problématique, attestant de la mauvaise utilisation des tablettes par les chauffeurs, peut biaiser l'indicateur : le chauffeur ne se déconnecte pas du système à la fin de son quart de travail. Si aucun second chauffeur n'utilise le véhicule, son identifiant y reste lié jusqu'à ce qu'il se reconnecte. Son quart de travail sera donc d'une longue durée.

Indicateur de courses

Par la suite, un croisement des chauffeurs avec les courses effectuées permet d'évaluer la distribution des courses par chauffeur par jour. Ceux étant les plus et les moins actifs par rapport à la demande peuvent alors être identifiés. Dans la base de données, les courses effectuées par un chauffeur sont identifiables directement dans la table des courses où un identifiant du chauffeur est présent. Il faut s'assurer de n'utiliser que celles ayant été complétées. Le champ annulation permet cette vérification : si celui-ci présente une valeur de 0, alors la course s'est bien déroulée, sinon elle a été annulée.

Autres visualisations

D'autres indicateurs, autant de base que complexes, pourraient être évalués pour compléter l'évaluation de la performance des chauffeurs. Des exemples en fonction de leur forme de visualisation sont énumérés ci-dessous :

1. Le nombre total de chauffeurs ;
2. La répartition par jour des moyennes d'heures de travail par chauffeur sur une semaine ;
3. La distribution des courses par chauffeur en fonction de la durée de travail ;
4. La distribution des courses par chauffeur en fonction du début du quart de travail ;
5. La distribution des durées de travail en fonction de l'heure de début du quart de travail ;
6. La distribution de la distance parcourue par quart de travail ;
7. La distribution du nombre de postes visités sur un quart de travail.

Un champ d'intérêt pour les régulateurs de l'industrie concerne le revenu des chauffeurs. Cette information, non fournie dans les données, ne suit pas une relation proportionnelle avec le nombre de courses puisque le tarif varie en fonction de la distance et du temps d'attente. Ainsi, ce n'est pas parce qu'un chauffeur effectue de nombreuses courses que celui-ci est bien rémunéré. Le revenu pourrait être estimé, de manière rudimentaire, à partir des distances de courses.

4.4.3.2 Taxi

L'analyse du taxi est d'abord réalisée par un suivi de leur mouvement et de leur statut, puis par une évaluation de ses performances en termes de distances et durées parcourues sans clients. Ces indicateurs sont les plus essentiels pour un opérateur désirant optimiser son service. D'autres indicateurs peuvent aussi être estimés afin d'examiner d'autres aspects du véhicule de taxi.

Indicateur de disponibilité

La disponibilité se mesure dans le temps grâce à la répartition, sur une semaine, des taxis ayant un statut prêt à recevoir un client. La courbe de demande se rajoute facilement à cette répartition grâce au statut sur une course. Ce croisement avec la demande permet d'évaluer le taux d'utilisation des véhicules et de cibler les périodes où les taxis sont les plus utilisés. Cette représentation permet de

visualiser la fluctuation de la flotte et donc de l'offre de service à travers les heures des journées de la semaine.

C'est à partir du statut du véhicule dans la table GPS qu'il est possible d'identifier si le taxi est prêt à recevoir un client (statut=3) ou s'il est sur une course (statut=4). La résolution temporelle se fait par tranche de 30 minutes. Pour chacune d'elle, le nombre total de taxis en service donc ayant un statut 3 ou 4 est comptabilisé. Les taxis sur une course sont déterminés en additionnant tous les véhicules présentant au moins un point GPS de statut 4 dans la tranche de temps. Puis, en soustrayant le nombre de taxis sur une course du total, ceux disponibles sont obtenus. Le taux d'utilisation est simplement la division des taxis sur une course sur le total des véhicules.

La limitation de cette méthode de calcul concerne les courses de courtes durées enjambant les tranches de temps. En effet, un taxi effectuant ce type de course apparaît avec un client pour les deux tranches de 30 minutes tandis qu'en réalité, il est disponible la majeure partie du temps de ces deux tranches. Une manière de pallier ce problème est d'utiliser de plus petites tranches de temps, approchant la durée d'une course.

La disponibilité se mesure aussi dans l'espace par l'intermédiaire d'une carte de densité de temps. Le concept est de montrer le temps cumulé disponible de tous les véhicules étant passés dans un secteur. Les périodes utilisées pour afficher les résultats sont les groupes d'heures. Ceux choisis ici représentent les grands mouvements habituels de population soit : la nuit (0 h à 6 h), l'heure de pointe du matin (6 h – 9 h), le milieu de la journée (9 h – 15 h), l'heure de pointe de l'après-midi (15 h – 18 h) puis le soir (18 h – 24 h). Afin de calculer l'indicateur de disponibilité spatiale, il faut additionner, pour chaque groupe d'heures, la durée pour tous les points GPS ayant un statut 3 et un même secteur. Puis, la division de cette somme par le nombre d'heures du groupe et par la superficie du secteur permet la comparaison entre les groupes d'heures et les secteurs. Ainsi, les moments et les endroits où il est plus probable pour un client de croiser un taxi libre sont identifiés.

Indicateurs de distances et durées

Les distances et durées parcourues par taxi sont un indicateur de performance permettant de déduire les portions productives et non productives du service du véhicule. Elles sont représentées ici sous la forme d'une distribution journalière. La distance totale se mesure à partir de la table GPS, en sommant le champ de distance pour tous les points du véhicule possédant un statut 3 ou 4 pour chaque jour. L'opération est la même pour la distance productive, ou en course, à l'exception du

statut utilisé qui est uniquement celui en course. La soustraction entre la distance totale et la distance productive donne celle non productive. La méthode de calcul est similaire pour la durée.

Cependant, un cas limitatif existe. Les courses de transport adapté fournies par la STM ne font pas partie des informations reçues. Un véhicule effectuant une de ces courses n'est pas identifié par un statut 4, mais plutôt un statut pas prêt à recevoir des clients (statut=2). Ces points sont donc comptabilisés parmi les distances et durées à vide ce qui les surestime. Toutefois, Taxi Diamond n'est pas l'intermédiaire le plus important auquel la STM fournit du transport adapté ; c'est donc une minorité des véhicules qui est concernée.

Autres visualisations

Ces autres indicateurs offrent une description générale du taxi. Puisque son comportement est finement lié à celui du chauffeur, la performance du véhicule peut donc aussi être évaluée à travers les indicateurs sur les chauffeurs.

1. Le nombre total de taxis ;
2. Le nombre moyen de chauffeurs par taxi.

Les périodes de grande demande comme le jour de l'an, les grands festivals ou la fin de semaine de la F1 peuvent faire l'objet d'analyses évènementielles. Ces dernières peuvent être utilisées pour vérifier la couverture des taxis et inciter les chauffeurs à être en service durant ces périodes si historiquement ils n'étaient pas assez présents. Les périodes de moindre demande comme les jours fériés peuvent aussi être l'objet d'une telle analyse pour mieux ajuster l'offre à la demande.

4.4.3.3 Course

La course représente l'utilisation du taxi. Cet objet est donc de principale importance pour les planificateurs ayant pour objectif d'atteindre un équilibre entre l'offre et la demande. De plus, il offre une piste de réflexion sur la place du taxi dans les choix modaux faits par les individus en milieu urbain. Son analyse consiste en un suivi spatiotemporel détaillé. Ainsi, la répartition des débuts et fins de courses, la localisation des origines et destinations, la localisation du type de course ainsi que la distribution des distances et durées en plus d'autres indicateurs sont développées.

Indicateur de débuts et fins de courses

Les débuts et fins de courses répartis selon les heures de la semaine sont observables respectivement à partir des champs du moment d'origine et du moment de destination de la table des courses. Seules les courses non annulées (annulation=0) sont utilisées ici. Le nombre de débuts et fins est additionné pour chaque tranche de 30 minutes. En segmentant la demande par agglomération, les interactions entre ces zones peuvent être étudiées, en plus de savoir si une agglomération génère plus de courses qu'elle en produit.

Indicateur d'origines et de destinations

La carte localisant la densité d'origines et de destinations des courses permet de cibler dans l'espace les secteurs les plus achalandés. Grâce au secteur d'origine et de destination déjà identifié dans la table des courses, il est possible de sommer les courses des mêmes secteurs. Afin de visualiser la composante temporelle sur la carte, une division selon des groupes d'heures mentionnés précédemment est utilisée. Par la suite, cette quantité est divisée par le nombre d'heures du groupe ainsi que par la superficie du secteur à des fins de comparaison.

La limitation de cette méthode est en lien avec la superficie du secteur. Elle touche les vastes secteurs qui possèdent un point de grande concentration de la demande. Par exemple, l'aéroport est un grand pôle d'attraction des courses par taxi et puisque sa superficie est très étendue (car peu de gens habitent ce secteur), son secteur ne présente pas une forte densité d'origines ou de destinations. Une division du territoire par grille, préférablement de 1 km^2, serait une option à considérer. Cela permettrait de capter les grands pôles de demande en taxis.

Indicateur de types de courses

La localisation des types de courses offre une représentation spatiale du marché du taxi. En segmentant par groupes d'heures, l'évolution à travers une journée vient compléter l'analyse. À partir de la table des courses, le nombre de courses hélées (type=0) puis commandées (type=1) est comptabilisé pour chaque secteur d'origine et groupe d'heures. La proportion de courses commandées (commandé/(hélé+commandé)) est ensuite calculée proposant ainsi une manière simplifiée de représenter la tendance quant au type de courses.

Indicateur des distances et durées

La distribution des distances et durées de courses permet de cibler le type d'utilisation du taxi que font les clients. Pour chaque distance d'un kilomètre et groupe d'heures, le nombre de courses est compté. Ce nombre est par la suite divisé par le nombre total de courses pour obtenir un pourcentage. La même méthode est utilisée pour la distribution des durées, avec des tranches de durées de cinq minutes. La table des courses possède déjà les champs distance et durée, à partir desquels les calculs sont effectués.

Autres visualisations

Les mêmes caractéristiques des courses sont analysées sous plusieurs autres formes de visualisation, soulignant possiblement de nouvelles tendances de la demande. De plus, d'autres indicateurs sur les courses sont aussi pertinents à développer pour l'opération de la flotte de véhicules.

1. Nombre total de courses ;
2. Répartition du nombre de courses par jour ;
3. Répartition du type de courses selon les heures ;
4. Distribution du temps d'attente des clients ;
5. Distance moyenne de courses par zone ;
6. Nombre de courses par zone où le client ne s'est pas présenté ;
7. Répartition et localisation des distances pour aller chercher un client.

Croiser les caractéristiques du sol comme le type d'utilisation et la population avec la demande en taxi pourrait produire des visualisations intéressantes permettant de mieux comprendre d'où émergent les courses. Ce croisement est justement analysé au chapitre 6 dans le but de cibler les facteurs d'influence de la demande en taxis et, dans un même temps, prédire les secteurs où la demande apparaîtra.

4.4.3.4 Poste d'attente

Le poste d'attente est parfois situé assez près d'un pôle de demande pour que les individus n'aient qu'à se déplacer de quelques pas pour entrer dans le taxi. Puisqu'il peut ainsi être corrélé avec la

génération de la demande, les indicateurs développés pour cet objet portent sur sa performance. Les courses ayant pour point de départ un poste ainsi que le temps qu'y passent les chauffeurs sont les principaux indicateurs estimés.

Indicateur de courses

La localisation des courses émergeant des postes d'attente indique si les postes sont bien situés, c'est-à-dire près d'un endroit générateur de déplacements. Afin de capter les courses de n'importe quels taxis stationnés, une zone tampon d'un rayon de 25 mètres est définie autour du point GPS marquant les emplacements réservés. Par jointure spatiale entre les points d'origine du déplacement du client et les zones tampons, le nombre de courses pour chaque poste est calculé. Puisque ce nombre est assez faible, aucune segmentation temporelle n'est employée ; le calcul se fait sur la période complète. Les postes d'attente fermés (statut=2 dans la table concernée) sont évidemment exclus de l'analyse.

Le choix de la grandeur de la zone tampon peut se baser sur des critères plus précis. Une information fournie sur les postes est le nombre de places de stationnement disponibles. Celui-ci pourrait être combiné à la longueur moyenne d'un espace de stationnement, de manière à obtenir une zone propre à chaque poste.

Indicateur de chauffeurs

La somme du temps passé par les chauffeurs aux postes d'attente montre sur une carte le niveau d'utilisation de ceux-ci. Elle reflète, dans un même temps, la disponibilité des chauffeurs aux postes. Une jointure spatiale des zones tampons avec la table GPS est effectuée pour récolter tous les points des véhicules étant passés par le poste d'attente. Seuls les points présentant une vitesse instantanée de 5 km/h ou moins sont conservés afin d'exclure les véhicules ne faisant que circuler sur les rues captées par la zone tampon. Pour chaque poste d'attente, la somme des durées est calculée. L'indicateur reflète donc la popularité du poste, en termes de quantité de chauffeurs, ainsi que la proximité de générateurs de demande, en termes de durée avant d'obtenir une course. Une segmentation par groupes d'heures est envisageable dans le but de voir les périodes de la journée où ils sont utilisés.

Autres visualisations

Les autres indicateurs des postes d'attente portent principalement sur leurs performances, à savoir si leur situation est bénéfique à la demande. En effet, si les distances pour aller chercher un client sont longues, que peu de chauffeurs attendent à ce poste ou qu'un générateur n'est pas dans les alentours, il serait peut-être intéressant de considérer la relocalisation du poste. Des exemples de représentation de ces indicateurs sont :

1. Nombre total de postes ;
2. Distribution du nombre de chauffeurs visitant un poste ;
3. Distribution du temps passé par les chauffeurs à un poste ;
4. Distribution de la distance pour se rendre au client par poste ;
5. Distance moyenne par poste pour se rendre au client.

L'étude spécifique d'un poste d'attente pourrait être réalisée afin de déterminer s'il est profitable au service de taxi. Toutefois, sans les données de la flotte complète, il est difficile de faire l'étude de cas pour tous les postes d'attente. En effet, les postes privés détenus par d'autres intermédiaires que Taxi Diamond ainsi que les postes n'étant pas situés dans la zone d'opération de l'intermédiaire ne peuvent être analysés dans le contexte de ce projet. De plus, la localisation n'est pas l'unique facteur pour lequel un chauffeur choisit d'attendre à un poste en particulier. La présence de ses amis, la proximité d'un certain service ou même la convivialité de l'emplacement peuvent aussi en faire partie. Ainsi, ce n'est pas parce que l'utilisation du poste est faible que ce dernier n'est pas bien situé pour répondre à la demande.

CHAPITRE 5 DÉMONSTRATION DES INDICATEURS

Ce chapitre a pour objectif de mettre en contexte l'ensemble des indicateurs développés pour l'industrie du taxi de Montréal. Le cadre de développement des indicateurs présente deux niveaux tels que détaillés à la figure 5-1. D'abord, le niveau d'estimation cible la manière de développer l'indicateur selon la forme de visualisation, le type et l'objet de l'étude. Par exemple, la répartition temporelle de l'offre de taxis pourrait être estimée. Puis, le niveau d'analyse permet de déterminer la segmentation spatiale et temporelle à utiliser pour calculer l'indicateur. En reprenant l'exemple précédent, l'indicateur devient la répartition temporelle de l'offre de taxis pour une semaine dans l'agglomération 11.

NIVEAU D'ESTIMATION	**Forme de visualisation**	**X**	**Type d'indicateurs**	**X**	**Objet de l'étude**
	1. Statistique descriptive 2. Distribution fréquentielle 3. Répartition temporelle 4. Carte		1. Offre 2. Demande		1. Chauffeur 2. Taxi 3. Course 4. Poste d'attente
NIVEAU D'ANALYSE	**Période**	**X**	**Zone**		
	1. Heure/Groupe d'heures 2. Jour 3. Semaine 4. Mois 5. Saison 6. Année		1. Secteur 2. Agglomération 3. Île de Montréal		

Figure 5-1 : Construction des indicateurs et segmentation pour leur analyse

Tous les indicateurs de ce projet sont développés afin de montrer les tendances globales des activités des taxis sur une année. De manière plus spécifique, les indicateurs peuvent être employés pour évaluer des périodes et endroits précis. Des jours et des secteurs de forte demande, comme le jour de l'an et l'aéroport, sont pertinents à étudier pour les planificateurs.

Cette section présente d'abord les données utilisées pour développer les indicateurs. Puis, pour s'assurer que les données soient en mesure de dresser le portrait de la flotte, la représentativité de l'échantillon est mesurée. Par la suite, les indicateurs au niveau de l'offre et de la demande en taxis sont estimés et analysés.

5.1 Données utilisées

Afin de produire les indicateurs d'offre et de demande, une partie des données décrites à la section 4.1 sont utilisées. Plus précisément, l'information sur les taxis comprend les traces GPS et les courses ; l'information expliquant les activités de taxis inclut les postes d'attente, le calendrier et la météo ; puis l'information délimitant le territoire est composée des agglomérations et des secteurs de recensement.

Les données de taxis proviennent de Taxi Diamond et couvrent la période du 1er janvier 2015 au 31 décembre 2015. Au total, l'ensemble brut de 199 millions de points GPS et 4,4 millions de courses représentent 1 282 taxis et 2 093 chauffeurs uniques. Pour des raisons de représentativité expliquées à la section 5.2, uniquement les véhicules possédant un permis dans A11 sont utilisés, réduisant à 1 209 le nombre de taxis de l'échantillon. Le traitement des données de courses et GPS effectué pour aboutir à la base de données finale est respectivement détaillé au tableau 5-1 et au tableau 5-2. Chaque grande étape y est indiquée ainsi que la quantité d'information associée qui est retirée de la table.

Tableau 5-1 : Traitement des données de courses pour l'année 2015

COURSE			
Nombre de courses brutes	**4 428 820**		
Doublons	42 473	0,96 %	
Origine ou destination hors de la zone de Montréal	271 953	6,20 %	
Course hélée dans la rue pour laquelle le client ne se présente pas	38	0,00 %	**8,09 %**
Distance = 0 km ou Durée = 0 min	38 646	0,94 %	
Course simultanée	5 012	0,12 %	

Tableau 5-2 : Traitement des données GPS pour l'année 2015

GPS			
Nombre de points GPS bruts	**199 486 082**		
Hors de la zone de Montréal	265 246	0,13 %	
Numéro de course ne se retrouve pas dans la table course	6 389 719	3,21 %	**3,61 %**
Point se trouvant dans une course ayant un point hors service	539 023	0,28 %	

Parmi cet ensemble de données GPS, 89 journées ne possèdent aucun enregistrement. Cela engendre un problème dans le traitement des données, au niveau de l'attribution du numéro de véhicule et de chauffeur à la course. En effet, une course qui s'est fait rejeter par un chauffeur à 23h58 avant d'être acceptée par un autre à 0h01 (la seconde journée n'ayant pas de données) se voit attribuer le mauvais numéro de véhicule. Un autre problème engendré se situe au niveau de la distance des courses commençant ou finissant un jour vide. Celle-ci sera principalement mesurée à vol d'oiseau entre l'origine et la destination, selon la portion de la trace GPS se trouvant le jour sans données. Une piste pour pallier cette situation est d'estimer et d'appliquer un indice de tortuosité selon le quartier dans lequel la course se déroule.

Les conditions météorologiques enregistrées à une fréquence de 15 minutes ne sont pas utilisées ici. Ces dernières ne sont recueillies qu'à partir de juillet 2015 sur le site internet d'Environnement Canada ; elles ne couvrent donc pas la totalité de l'année 2015. La météo par jour de la station de l'aéroport Pierre-Elliott-Trudeau, tirée des données historiques d'Environnement Canada, est plutôt l'information utilisée pour cette étude.

5.2 Représentativité

Afin de produire des résultats valides pour l'ensemble de la flotte, il faut s'assurer de la représentativité de l'échantillon de données. Dans un premier temps, le nombre et la proportion de taxis opérés par Taxi Diamond dans chaque agglomération sont déterminés. La méthode d'attribution du permis, décrite à la section 4.3.3, aboutit au résultat du tableau 5-3.

Tableau 5-3 : Attribution du permis aux véhicules de 2015

AGGLOMÉRATION		NOMBRE DE VÉHICULES
A5	→	5
A11	→	1 209
A12	→	56
Hors Montréal	→	12

Taxi Diamond n'est pas enregistré pour opérer dans l'agglomération 5 ni à l'extérieur de l'île. De plus, les taxis dans l'agglomération 12 ne sont généralement pas dispersés partout sur le territoire ne pouvant pas offrir de service à toute la demande. C'est pourquoi même si Taxi Diamond atteint

une proportion de 20,2 %, cette zone est exclue de l'étude. Taxi Diamond opère 1 209 taxis dans l'agglomération 11, avec une proportion atteignant 29,9 %. Cet échantillon est plus que suffisant puisqu'il respecte la quantité minimum de 384 véhicules selon l'équation 5-1 (Morency & Saunier, communication personnelle, Automne 2013). L'écart-type le plus critique, soit de 0,5, est utilisé puisqu'aucune information n'est fournie à propos de la population à l'exception du nombre de permis par agglomération. L'échantillon possède donc une quantité de données suffisante, cependant sa qualité l'est moins puisque les permis restreints et spéciaux sont difficilement identifiables.

$$n = {t^2 \cdot p(1-p)}/{m^2} \qquad \text{Équation 5-1}$$

Où *n* est la taille de l'échantillon

t est le niveau de confiance (valeur habituelle = 1,96)

p est l'écart-type du paramètre à mesurer dans la population (valeur critique = 0,5)

m est la marge d'erreur acceptable (valeur habituelle = 0,05)

Dans un second temps, il faut mesurer les facteurs d'expansion de l'agglomération 11 pour chaque mois de l'année. Seuls les véhicules ayant fait au moins une course durant la période sont considérés. Le tableau 5-4 montre que les facteurs sont très semblables. En effet, le coefficient de variation de 1,2 % indique que la dispersion relativement à la moyenne est faible. Les valeurs sont donc très concentrées autour de la moyenne. Cela révèle qu'un facteur pour l'année aurait pu être appliqué.

Tableau 5-4 : Facteurs d'expansion pour l'année 2015

MOIS	NOMBRE DE VÉHICULES	FACTEUR D'EXPANSION
Janvier	1025	3,94
Février	1023	3,95
Mars	1008	4,00
Avril	1020	3,96
Mai	1024	3,94
Juin	1027	3,93
Juillet	1020	3,96
Août	1018	3,97
Septembre	1028	3,93
Octobre	1028	3,93
Novembre	1053	3,83
Décembre	1045	3,86

STATISTIQUES	
Moyenne	1027
Écart-type	11,94
Coefficient de variation	1,2 %

5.3 Indicateurs

Les indicateurs estimés dans l'étude de cas de Montréal se basent sur les véhicules détenant un permis pour l'agglomération 11 durant l'année 2015. Une période d'un an reflète à Montréal de nombreuses tendances dues aux variations de température. Ainsi, une analyse par saison est réalisée afin de les comparer entre elles ainsi que d'évaluer les tendances à l'intérieur d'une même saison. Les résultats sont pondérés par mois afin de représenter la flotte complète de l'agglomération 11.

Le but de cette section est de cibler et analyser les grandes tendances de déplacements par taxi. Ainsi, les jours fériés sont retirés de l'analyse puisque les comportements des utilisateurs ne sont pas habituels durant ces types de jours.

5.3.1 Offre

L'offre du service de taxi est identifiable à travers les indicateurs portant sur le chauffeur et le taxi. Leur analyse est très similaire puisque ces objets sont étroitement liés. En effet, un chauffeur est le conducteur du taxi ; ces deux objets évoluent donc ensemble dans le temps et l'espace. Le tableau 5-5 offre une description générale de ces objets pour l'année 2015.

Tableau 5-5 : Statistiques descriptives sur l'offre du service de taxi

INDICATEUR		MOYENNE	COEF. VAR.
Nombre de chauffeurs par taxi		2,69 chauffeurs	± 47 %
Distance parcourue par taxi			
Année	Total	30 191,3 km	± 49 %
Année	Sans client	18 127,8 km	± 51 %
Jour	Total	144,7 km	± 56 %
Jour	Sans client	86,9 km	± 58 %
Durée de travail par taxi			
Année	Total	2 128 h	± 48 %
Année	Sans client	1 539 h	± 47 %
Jour	Total	10,2 h	± 43 %
Jour	Sans client	7,4 h	± 45 %
Durée d'un quart de travail		7,1 h	± 66 %
Nombre de chauffeurs par jour			
	Lundi	4 396 chauffeurs	± 3,2 %
	Mardi	4 379 chauffeurs	± 3,7 %
	Mercredi	4 460 chauffeurs	± 3,4 %
	Jeudi	4 549 chauffeurs	± 3,5 %
	Vendredi	4 649 chauffeurs	± 4,5 %
	Samedi	4 197 chauffeurs	± 3,7 %
	Dimanche	3 942 chauffeurs	± 3,4 %

Le propriétaire d'un taxi choisit le nombre de chauffeurs qui utilisent son véhicule. En général, deux à trois chauffeurs se partageront le véhicule au cours de l'année. Un taxi moyen se retrouve la majeure partie de son temps (72 %) et de sa distance (60 %) sans client, que ce soit pour l'année ou pour une journée. Les durées et distances sont assez variables d'un taxi à l'autre avec un coefficient de variation de près de 50 %. Une plus grande variabilité est soulignée pour la distance parcourue par jour. Il serait logique de penser que le coefficient de variation des distances et durées soit partiellement lié au nombre de chauffeurs par taxi. Un véhicule à deux chauffeurs peut

travailler plus longtemps et faire plus de distance qu'un véhicule à chauffeur unique, d'où l'importante variabilité.

Un chauffeur peut travailler, dans une journée, durant une longue période sans se déconnecter du système de répartition ou durant plusieurs périodes consécutives entrecoupées de pauses au cours desquelles il se déconnecte du système. En moyenne, une période de travail est de 7,1 h. Son coefficient de variation de 66 % montre que les chauffeurs ont des habitudes de travail variées. Par exemple, certains préfèrent se déconnecter du système afin de prendre une pause dîner tandis que d'autres restent disponibles et refusent des courses pendant leur dîner, ce qui en résulte un nombre de quarts de travail différent. Il faut se rappeler que lorsqu'aucun point GPS n'est reçu en moins de 10 minutes, le véhicule et son conducteur sont considérés comme déconnectés. De plus, au cours d'une semaine moyenne, le nombre de chauffeurs disponibles par jour est très semblable, avec un écart de 707 entre les moyennes des jours extrêmes (dimanche et vendredi). Pour une même journée de semaine, le nombre de chauffeurs est constant durant l'année.

La répartition temporelle des débuts de quarts de travail, présentée à la figure 5-2, détaille les habitudes et comportements des chauffeurs. Il est attendu que la somme des quarts d'une journée ne soit pas égale au nombre de chauffeurs puisque ceux-ci peuvent se déconnecter et se connecter de nouveau à n'importe quel moment de la journée.

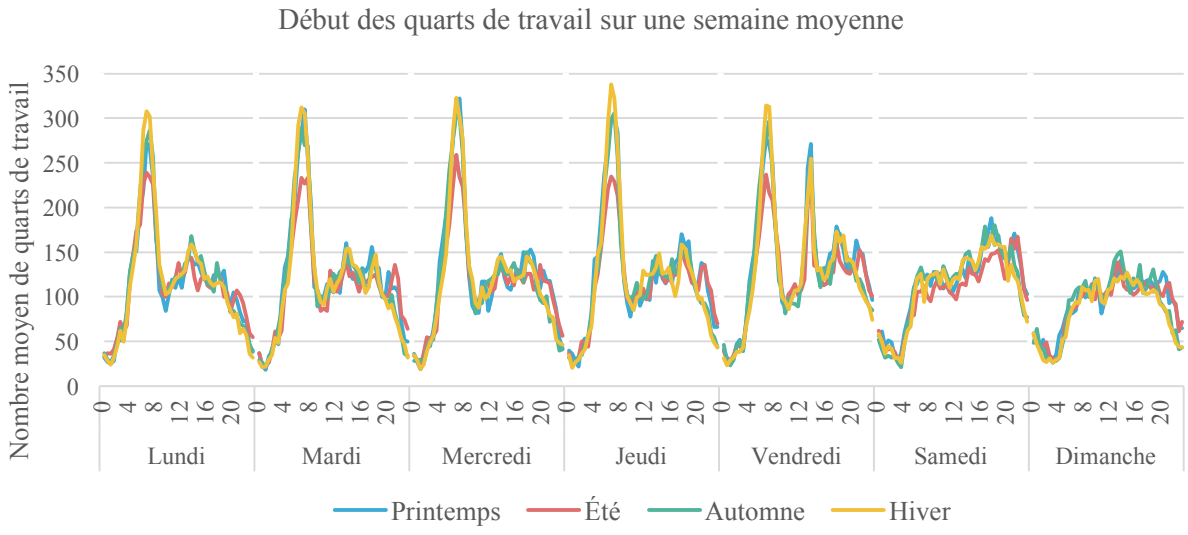


Figure 5-2 : Débuts des quarts de travail pour une semaine moyenne selon les saisons (incrément de 30 min)

La tendance de début de travail est assez similaire entre les saisons. En ce qui a trait aux jours de semaine (lundi au jeudi), une importante part des chauffeurs se met au travail entre 6 h et 8 h. Puis, les débuts de quarts oscillent entre 100 et 150 le reste du jour, pour présenter une légère augmentation vers 14 h. Cette augmentation de l'après-midi se décale au fur et à mesure qu'avance la semaine pour atteindre 18 h le jeudi. Le vendredi présente encore, le matin, de nombreux quarts débutant entre 6 h et 8 h. Un second pic important se produit entre 13 h 30 et 14 h 30. Le commencement de la soirée est marqué par un nombre variant autour de 150 quarts. Les jours de fin de semaine révèlent une quantité plus faible de quarts de travail, en lien avec le nombre plus faible de chauffeurs. Le samedi voit toutefois une hausse en soirée, à partir de 17 h. Peu importe la journée, très peu de chauffeurs commencent à travailler la nuit, passé 0 h 30.

Les différences majeures entre les saisons se situent le matin et en soirée. Durant la matinée, l'été présente moins de chauffeurs commençant leur travail tandis que les trois autres saisons en montrent généralement plus. En soirée, particulièrement de 21 h à 24 h, l'été montre le plus grand nombre de débuts de quarts de travail, suivi de près par le printemps, puis l'automne et finalement l'hiver qui montre le nombre le plus faible.

Les quarts de travail commençant à minuit ne peuvent pas être analysés, c'est pourquoi ils ont été enlevés de la figure 5-2. Ils sont biaisés par les jours de données manquantes. Par exemple, un chauffeur commence à travailler durant une journée pour laquelle les données ne sont pas fournies et finit la journée suivante où les données sont disponibles (travail de nuit). Le premier point GPS disponible, donc celui à minuit, se voit attribuer le statut hors service, car plus de 10 minutes se sont écoulées depuis sa dernière apparition dans les données. Puisque le début des quarts de travail se mesure grâce au changement du statut hors service vers un autre statut, tous les chauffeurs travaillant de nuit sont comptabilisés à minuit, ce qui, en réalité, n'est pas le cas.

L'analyse des périodes de travail des chauffeurs n'est pas entière sans l'élément complémentaire au début des quarts de travail : c'est-à-dire la fin des quarts de travail. La répartition temporelle de la figure 5-3 révèle les moments où les chauffeurs terminent leur quart et probablement certains changements de conducteurs dans les véhicules.

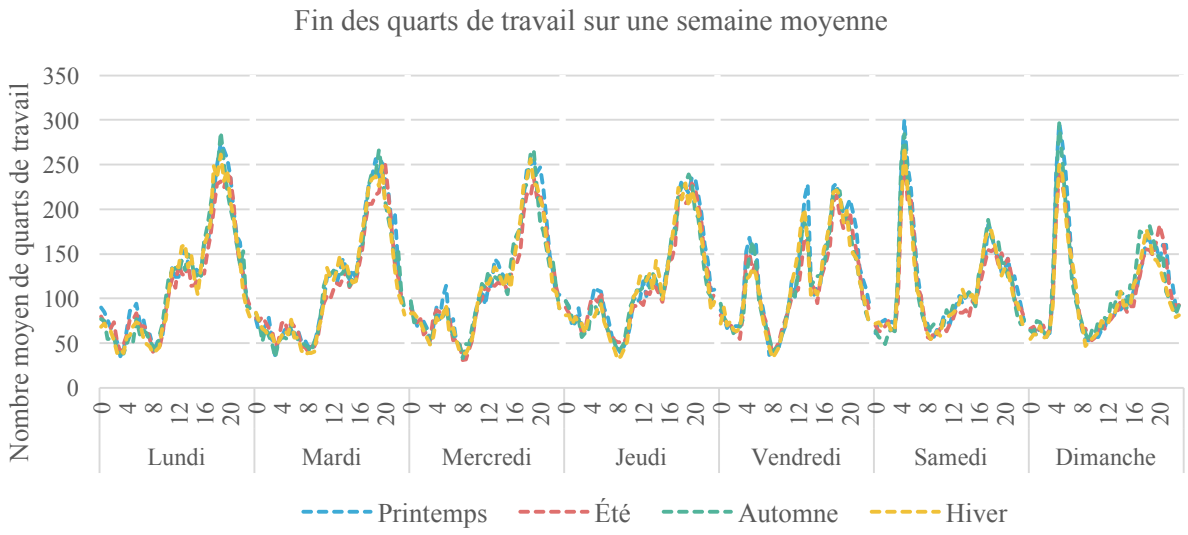


Figure 5-3 : Fins des quarts de travail pour une semaine moyenne selon les saisons (incrément de 30 min)

Tout comme pour la figure 5-2, les fins de quarts ont une tendance similaire au cours des différentes saisons. Du lundi au jeudi, une première vague relativement importante termine vers 13 h (ce qui porte à croire à une pause dîner ou un changement de chauffeur) puis une seconde vague majeure se produit entre 17 h 30 et 20 h. Ce pic du soir tend toutefois à diminuer plus la semaine avance. Il fait d'ailleurs penser à la masse importante de chauffeurs commençant avant 8 h, qui pourrait finir leur journée de travail à ce moment, soit 10 h plus tard. La fin de semaine montre une répartition temporelle complètement différente de celle des débuts de quarts. Des pics se produisent en fin d'après-midi de 17 h à 19 h.

Les chauffeurs de nuit finissent la majeure partie des jours de semaine vers 5 h, et plus la semaine avance, plus ils sont nombreux à finir à ce moment. Les activités nocturnes de la nuit du vendredi et samedi sont clairement identifiées par la plus grande quantité de chauffeurs finissant de travailler entre 4 h et 5 h du matin, après la fermeture des bars (3 h).

Pour la même raison que les débuts des quarts de travail, les fins se produisant, cette fois-ci, tout juste avant minuit (dans la plage de 23 h 30) ne peuvent pas être analysées. Elles ont donc été retirées de la figure 5-3.

La disponibilité des taxis est l'indicateur principal pour évaluer la couverture et la performance du service. Il est aussi un de ceux sur lequel un client se fie pour héler ou commander un taxi. D'abord,

la répartition temporelle de la figure 5-4 permet de cibler les périodes lorsqu'il est plus probable de croiser un taxi sur les routes. La différence entre la courbe des taxis en service et des taxis avec clients représente, en fait, les taxis disponibles. La répartition temporelle des taxis en service est assurément en relation avec les horaires de travail des chauffeurs (début et fin de quarts de travail).

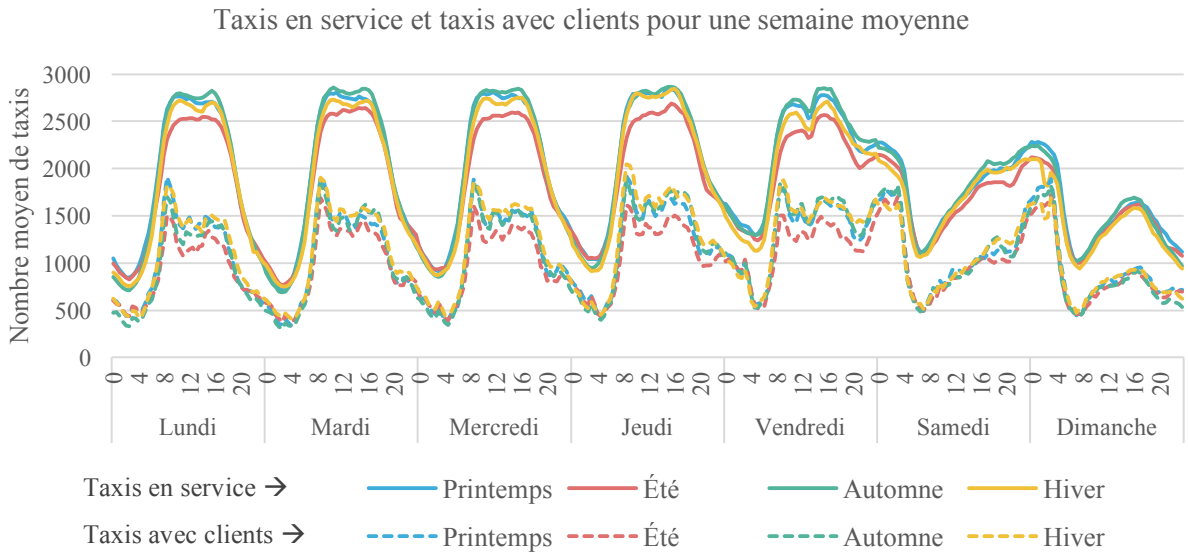


Figure 5-4 : Taxis en service et taxis avec clients pour une semaine moyenne selon les saisons (incrément de 30 min)

La tendance du nombre de taxis en service suit bien la combinaison des débuts et fins de quarts de travail des chauffeurs. En effet, la quantité de taxis sur les routes explose les matins de semaine, suite à la vague de chauffeurs qui commencent à travailler à partir de 6 h, pour atteindre plus de 2 500 taxis à 8 h 30. Puis, elle reste constante tout au long de la journée pour descendre graduellement à partir de 16 h, lorsque l'importante masse de chauffeurs finit son quart. La courbe passe sous la barre des 1 000 chauffeurs durant la nuit. La nuit du jeudi montre toutefois un nombre plus élevé de taxis.

La tendance du vendredi, en termes de nombre de taxis avec et sans client, débute de la même manière que les autres jours de la semaine. Toutefois, un second pic de même ampleur se produit en après-midi, et la soirée est marquée par plus de 2 000 taxis en service jusqu'à 3 h du matin. Le samedi et le dimanche révèlent un nombre moins important de taxis, atteignant respectivement au maximum 2 250 et 1 600. Ce maximum est atteint à minuit pour le samedi et en milieu d'après-midi pour le dimanche. L'ascension vers ce pic se fait progressivement à partir du matin.

La répartition dans le temps des taxis avec clients ressemble à celle des taxis en service. La nuit, moins de véhicules sont occupés que durant le jour. Un pic d'utilisation, au-dessus de 1 500 taxis, survient les matins de semaine, entre 8 h et 9 h. Un autre pic de moindre importance arrive aussi, cette fois-ci en après-midi, vers 15 h 30. Le vendredi montre une troisième période de forte utilisation en soirée, débutant à 23 h et se terminant à la fermeture des bars, à 3 h. Le samedi soir montre, comme le vendredi, plus de 1 500 taxis avec clients entre 23 h et 3 h. Durant la journée du samedi ainsi que le dimanche, le nombre de taxis occupés est faible, n'atteignant même pas 1 000 taxis le dimanche.

L'analyse en détails de la disponibilité des taxis pour une saison met aussi l'accent sur le taux d'utilisation des véhicules. La figure 5-5 fait état de la performance des taxis au niveau de leur utilisation par les clients tout au long d'une semaine d'automne. L'automne est choisi puisque cette saison représente les comportements moyens d'une année. De plus, les données utilisées dans la planification des systèmes de transport, celles provenant des grandes enquêtes régionales origine-destination, portent aussi sur la période d'automne.

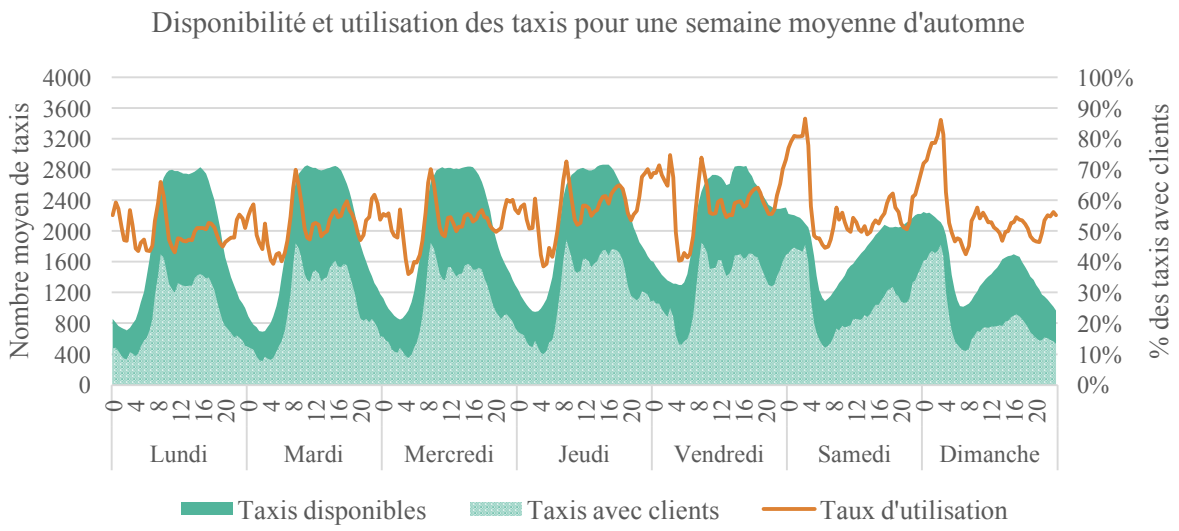


Figure 5-5 : Disponibilité et utilisation des taxis pour une semaine moyenne d'automne (incrément 30 min)

Le taux d'utilisation est le nombre de taxis avec clients divisé par le nombre de taxis en service. Ce pourcentage varie en fonction de l'heure et du jour de la semaine, tel que montré par la courbe à la figure 5-5. Les matins de semaine, vers 8 h, plus de 60 % des taxis se retrouvent avec un client.

Puis, le taux oscille autour de 50 % le reste de la journée avec un léger regain en soirée, particulièrement le jeudi où ce regain est en fait la période ayant le taux le plus élevé. Plus la semaine avance, plus le pourcentage des taxis avec clients est élevé pour les mêmes heures. Le vendredi présente la même tendance matinale que les jours de semaine, avec une utilisation atteignant 70 %. C'est la nuit qui est toutefois la période où la majorité des taxis en service sont utilisés, avec un taux de plus de 75 % entre minuit et 3 h 30. La même situation nocturne se produit le samedi. Pour le reste de la fin de semaine, les taux d'utilisation varient autour de 50 %.

La disponibilité des taxis est l'inverse du taux d'utilisation. Un planificateur cherche à ce que les taxis rencontrent la demande et soient ainsi le plus souvent occupés tout en assurant une disponibilité suffisante. Cela se produit le soir et la nuit du vendredi et samedi tandis que la situation contraire advient à partir de 4 h les matins de semaine.

Par la suite, afin de compléter la réflexion sur la disponibilité des taxis, la répartition spatiale est estimée, définissant ainsi les endroits où il est plus probable de trouver un véhicule libre. La figure 5-6 montre l'évolution des taxis disponibles sur le territoire de l'île de Montréal pour une journée moyenne de l'année 2015. Une segmentation par saison n'est pas utilisée ici puisque, suite au calcul de l'indicateur en fonction des saisons, les résultats étaient équivalents.

À chaque secteur de recensement est associée une densité d'heures moyenne (somme de la durée disponible de chaque taxi dans le secteur / superficie / nombre d'heures dans la période). Ainsi, un secteur ayant 1 h de disponibilité signifie que l'équivalent d'un taxi par km^2 est disponible durant toute la période.



(a) 0 h – 6 h

(b) 6 h – 9 h

(c) 9 h – 15 h

(d) 15 h – 18 h

(e) 18 h – 24 h

Figure 5-6 : Densité horaire moyenne de disponibilité des taxis pour une journée moyenne de l'année 2015

Les heures de pointe sont les périodes durant lesquelles la disponibilité des taxis est plus grande. Les endroits les plus achalandés par les taxis vides, où les densités d'heures sont de plus de 4 h/km^2, sont le Village, le Plateau-Mont-Royal, Villeray et Shaughnessy Village (près du Centre-Ville). En heures hors pointe, les mêmes endroits sont achalandés, toutefois de manière moins importante. Puisque le nombre de taxis en service est sensiblement le même durant le jour (voir la figure 5-4), cela signifie que les taxis se concentrent en des endroits précis en heure de pointe, tandis qu'en heure hors pointe, ils se dispersent sur tout le territoire de l'île. Les densités n'atteignent 4 h/km^2 que dans très peu de secteurs. Le soir et la nuit montrent des valeurs particulièrement faibles, ce qui est normal puisque les taxis en service sont moins nombreux en plus d'être beaucoup plus utilisés la nuit, à un taux dépassant 70 % le jeudi, vendredi et samedi.

Peu importe le moment de la journée, les taxis de A11 disponibles se tiennent plus à l'est ainsi qu'au nord-est du Mont-Royal que dans l'ouest. Cette situation peut s'expliquer par le fait que les centres d'activités sont nombreux dans ces zones ou que les destinations des courses s'y situent, obligeant le taxi à revenir vide de cet endroit. De plus, en supposant qu'un chauffeur se connecte au système dès qu'il quitte sa résidence et que cette profession ne fournit pas un salaire élevé, il se pourrait qu'une part importante des chauffeurs réside dans les quartiers de l'est qui sont moins dispendieux que ceux de l'ouest.

Un autre indicateur de performance du taxi a trait à la distance parcourue lorsqu'il est en service. Tel que mentionné au tableau 5-5, la distance par jour est en moyenne de 144,7 km avec un coefficient de variation de 56 % autour de la moyenne. Une analyse approfondie peut être faite par l'intermédiaire de la figure 5-7, où les distances totales ainsi que la moyenne de distance sans client selon le jour de la semaine sont estimées.

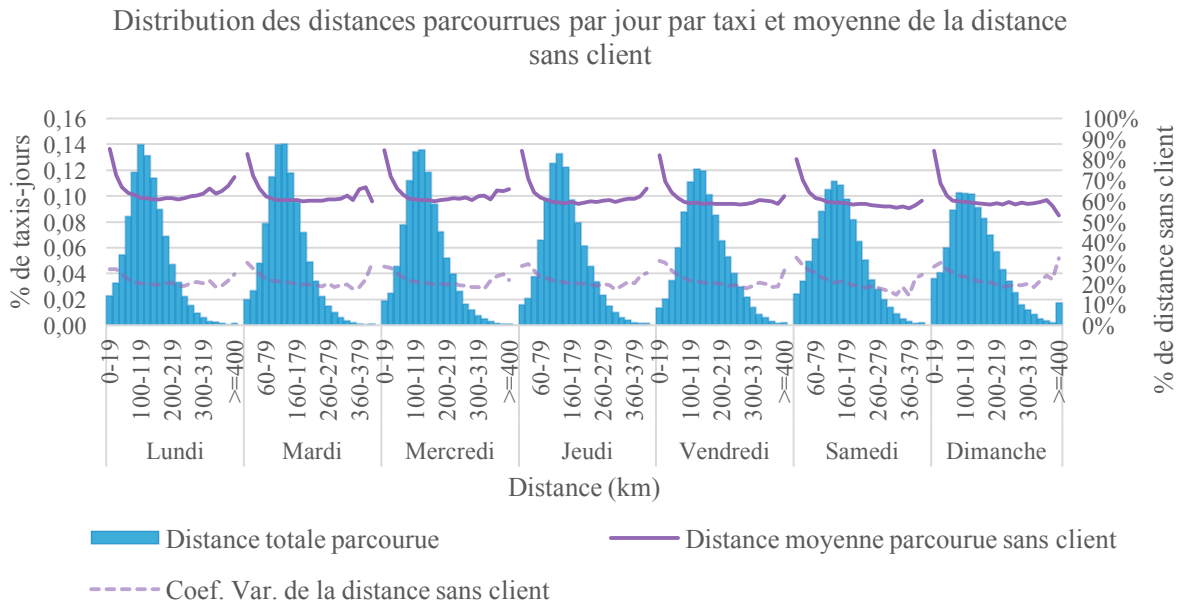


Figure 5-7 : Distance parcourue par jour par taxi et moyenne de la distance sans client pour une semaine moyenne de l'année 2015

La distance totale parcourue par un taxi en une journée semble augmenter plus la semaine avance. Le lundi, le groupe possédant le plus de taxis-jours est celui de 100 à 120 km, puis il devient celui de 120 à 140 km le reste de la semaine pour redescendre à 100 à 120 km le dimanche. Les distances moyennes par jour montrent aussi ce phénomène avec, du lundi au dimanche, des moyennes de 134 km, 136 km, 140 km, 148 km, 156 km, 150 km et 145 km. Dans le même sens, la dispersion des valeurs augmente au fur et à mesure qu'avance la semaine. Le lundi, les taxis-jours sont très concentrés autour de 100 km à 140 km tandis que le dimanche, l'étalement est plus grand, allant de 60 km à 160 km.

La moyenne de la distance sans client selon les groupes montre que celle-ci est assez constante, s'élevant à 60 %, à l'exception des petites et grandes distances. Cette relation révèle d'un côté que plus la distance parcourue par jour est petite, plus la portion à vide est élevée. Il est normal de croire que moins un véhicule travaille, moins il a la possibilité de faire des courses. D'un autre côté, pour les jours de semaine, plus la distance est grande, plus celle sans client est grande tandis que pour la fin de semaine, plus la distance est grande, plus la portion vide est faible. Pour la plupart des groupes de distance dans lequel se trouve un taxi-jour, sa distance sans client possède un coefficient de variation entre 20 % et 25 %.

Un taxi est vide plus de la moitié de sa distance puisque celui-ci doit se rendre à vide à l'origine de son client, puis, après l'avoir conduit à sa destination, il doit retourner à vide dans son agglomération ou bien à un endroit où il considère trouver facilement son prochain client. Il va sans dire que la réglementation en place, surtout en ce qui concerne les agglomérations, ne favorise pas la diminution des distances parcourues sans passager. Toutefois, cela peut aussi être attribué au comportement du chauffeur qui erre dans la rue à la recherche d'un client, augmentant ainsi sa distance parcourue sans client.

L'analyse des durées ressemble fortement à celle de la distance, c'est pourquoi la figure 5-8 montre une distribution des durées pour un jour moyen de 2015 et non pour chaque journée de semaine. La distribution de la durée sans client est aussi estimée.

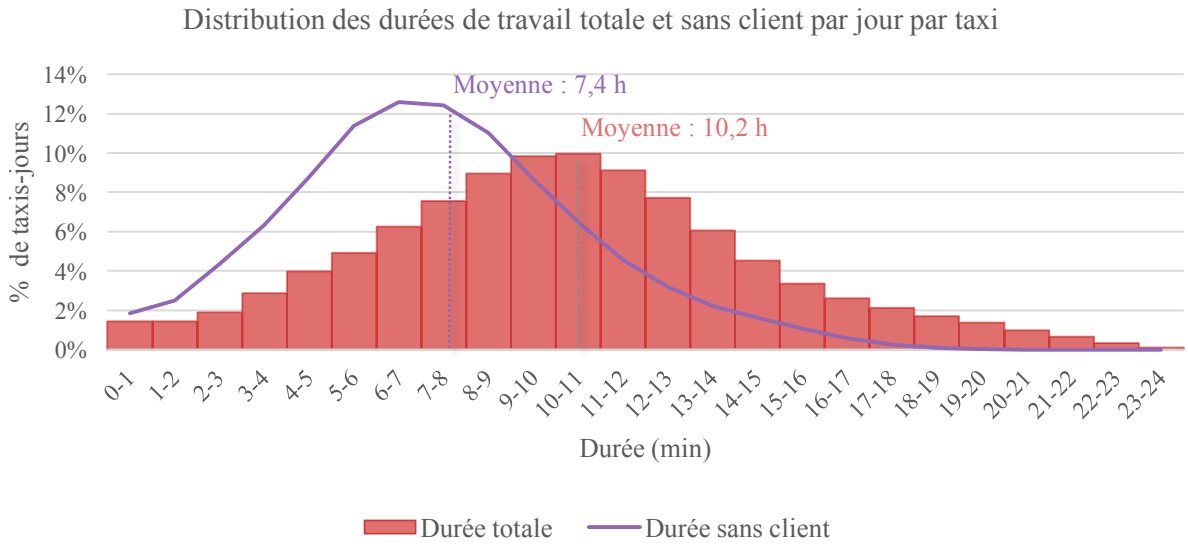


Figure 5-8 : Durée de travail totale et sans client par jour par taxi

Les durées de travail d'un taxi dans une journée ont une variation très large ; elles peuvent autant être minimes, de près de 0 h, qu'atteindre le maximum de 24 h. La moyenne est de 10,2 h et les taxis-jours se répartissent de manière symétrique autour de celle-ci. La durée sans client est naturellement plus courte que la durée totale de service. Sa moyenne est de 7,4 h et la plupart des valeurs se concentrent dans les classes plus courtes que la moyenne (la courbe ne semble donc pas suivre une distribution normale). La courbe de la durée sans client semble présenter une distribution plus concentrée autour de la moyenne que celle de la durée totale, avec des groupes atteignant respectivement pour chaque courbe 12 % et 10 % de taxis-jours près de la moyenne.

Il est attendu de voir des proportions plus élevées pour la durée sans client que pour de la distance sans client. En effet, un taxi ne se déplacera pas sans arrêt durant sa journée de travail. Ainsi, les périodes passées au poste d'attente contribuent à sa durée, mais non à sa distance à vide.

5.3.2 Demande

L'objet principal permettant l'étude de la demande en taxi est la course. Dès qu'une course se produit, cela signifie que la demande a été rencontrée et qu'un client se fait servir. Plusieurs caractéristiques des courses telles la distance, la durée, le type, etc. sont analysées ici. De plus, les chauffeurs, les taxis ainsi que les postes d'attente sont évalués en fonction des courses, offrant ainsi un regard sur leurs performances. Pour faire de meilleures comparaisons et analyses, une partie de cette évaluation a déjà été réalisée à la section 5.3.1. Elle concernait l'utilisation des taxis versus leur disponibilité ainsi que les distances et durées sans client. Le tableau 5-6 propose une description générale des caractéristiques des courses ainsi que des objets liés aux courses.

Tableau 5-6 : Statistiques descriptives sur la demande du service de taxi

INDICATEUR	MOYENNE	COEF. VAR.
Nombre de courses par chauffeur	10,4 courses	± 50 %
Nombre de courses par jour		
Lundi	38 365 courses	± 13 %
Mardi	40 777 courses	± 11 %
Mercredi	42 004 courses	± 10 %
Jeudi	47 406 courses	± 11 %
Vendredi	52 267 courses	± 14 %
Samedi	47 145 courses	± 13 %
Dimanche	39 412 courses	± 10 %
Distance d'une course	4,39 km	± 103 %
Durée d'une course	11,4 min	± 81 %
Proportion de courses s'amorçant aux postes d'attente par jour	11,7 %	

Indicateur de courses	Nuit	Pointe AM	Jour	Pointe PM	Soir
Distance (km)	5,05 ± 96 %	4,88 ± 99 %	4,01 ± 109 %	4,33 ± 109 %	4,08 ± 101 %
Durée (min)	10,9 ± 81 %	12,5 ± 81 %	11,2 ± 82 %	13 ± 82 %	10,8 ± 81 %
Proportion de courses s'amorçant aux postes d'attente	10,3 %	10,5 %	12,0 %	12,8 %	12,5 %
Proportion de courses commandées (commandé/hélé)	31 %	63 %	55 %	49 %	42 %

Un chauffeur fait en moyenne 10,4 courses dans sa journée de travail. Ce nombre est variable à 50 %, ce qui révèle que certains chauffeurs sont plus actifs que d'autres. Les courses débutant à un poste d'attente, et donc obligatoirement de type hélé, atteignent une proportion non négligeable par rapport au nombre de courses total d'une journée, soit en moyenne 11,7 %. L'évolution de cette proportion sur une journée moyenne montre que les individus vont moins prendre un taxi à partir d'un poste d'attente durant la nuit et les heures de pointe du matin que durant les heures de pointe de l'après-midi et la soirée.

L'évolution du nombre de courses par jour montre que, pour une semaine moyenne, la demande augmente pour atteindre un maximum le vendredi, puis redescend la fin de semaine. Ce sont principalement les activités nocturnes de la population le jeudi, vendredi et samedi soir qui font en sorte que ces jours sont de grande demande. La variabilité des courses ne change que très peu entre les journées, fluctuant de près de 12 %. Le changement de saisons en est une raison puisqu'il semble logique, lors de temps froid, que les individus préfèrent se déplacer au chaud à bord d'un véhicule tandis qu'en été, la marche est plus appréciée.

La distance moyenne d'une course est de 4,39 km, mais cette valeur varie énormément. La médiane se situe entre 2 et 3 km ce qui indique que la majorité des distances sont courtes. La durée moyenne est moins variable que la distance et a une moyenne de 11,4 min. Cela pourrait être expliqué par le fait que les petites courses du centre-ville subissent les effets de la congestion, ce qui fait en sorte que leur temps de parcours peut approcher celui des courses de moyenne distance. Les périodes de la journée influencent grandement la distance et la durée des courses. Un client utilise un taxi en général pour un plus long trajet la nuit que le jour, toutefois ce trajet dure le même temps. Il se déplace aussi pour une distance semblable la nuit et en heures de pointe, mais la durée est largement plus longue pour les heures de pointe. Ce phénomène est principalement dû à la lourde congestion routière durant ces périodes de la journée.

La manière de prendre un taxi varie aussi en fonction des heures de la journée. Le matin est particulièrement marqué par un marché de courses commandées, c'est-à-dire que les utilisateurs font la demande d'un taxi par téléphone ou via d'autres médias. Puis, le jour et l'heure de pointe de l'après-midi présentent une mixité de courses commandées et hélées. Finalement, le soir et la nuit sont plutôt de type hélé, puisque les taxis se positionnent aux pôles d'activités nocturnes et permettent aux individus de n'avoir qu'à sortir sur la rue pour trouver un taxi.

Afin d'évaluer la performance des chauffeurs, la figure 5-9 établit leur distribution selon le nombre de courses qu'ils effectuent par jour. Cette figure détaille les habitudes des chauffeurs selon les grands changements de mobilité dus aux saisons.

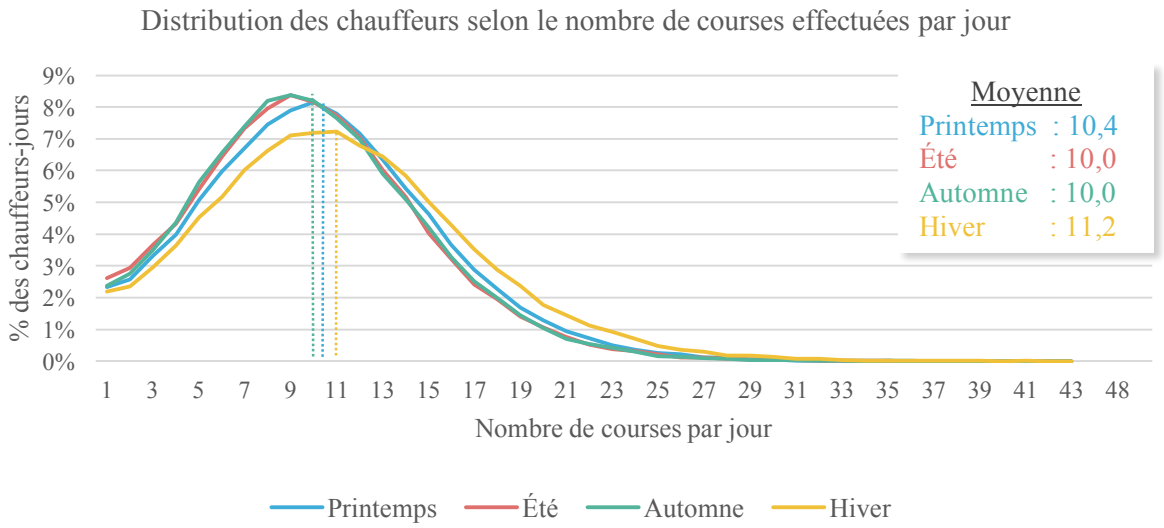


Figure 5-9 : Courses effectuées par chauffeur par jour selon les saisons

La segmentation par saison permet de constater que l'hiver a une distribution légèrement plus aplatie que les autres saisons, avec une valeur maximale de 1 point de pourcentage plus faible. Sa moyenne de 11,2 courses par jour est plus élevée que les autres saisons (10,0 et 10,4 courses par jour), ce qui est dû à la demande plus importante en hiver et au nombre de chauffeurs qui reste constant par rapport à l'automne et au printemps. En général, l'étalement des courbes est important, montrant qu'il est commun qu'un chauffeur effectue 5 ou 13 courses dans sa journée. Cette situation est confirmée par le coefficient de variation de 50 % de la valeur moyenne annuelle.

Le nombre de courses varie en fonction du nombre d'heures pour lesquelles le chauffeur est en service. Ayant déterminé que la durée moyenne d'un quart de travail est de 7,1 h et que son coefficient de variation est assez élevé (66 %), il est normal de voir un nombre de courses varier grandement d'un chauffeur-jour à l'autre. Cela va aussi de pair avec le comportement du chauffeur, si celui-ci est proactif et se positionne à des endroits stratégiques ou non. Bref, afin de comparer les performances entre les chauffeurs, prendre en considération le nombre d'heures travaillées serait nécessaire.

Laissant de côté les performances individuelles des chauffeurs, la figure 5-10 permet d'évaluer la performance globale de l'ensemble des conducteurs. Le nombre total de courses par chauffeur durant la saison est mesuré puis ordonné du plus actif au moins actif.

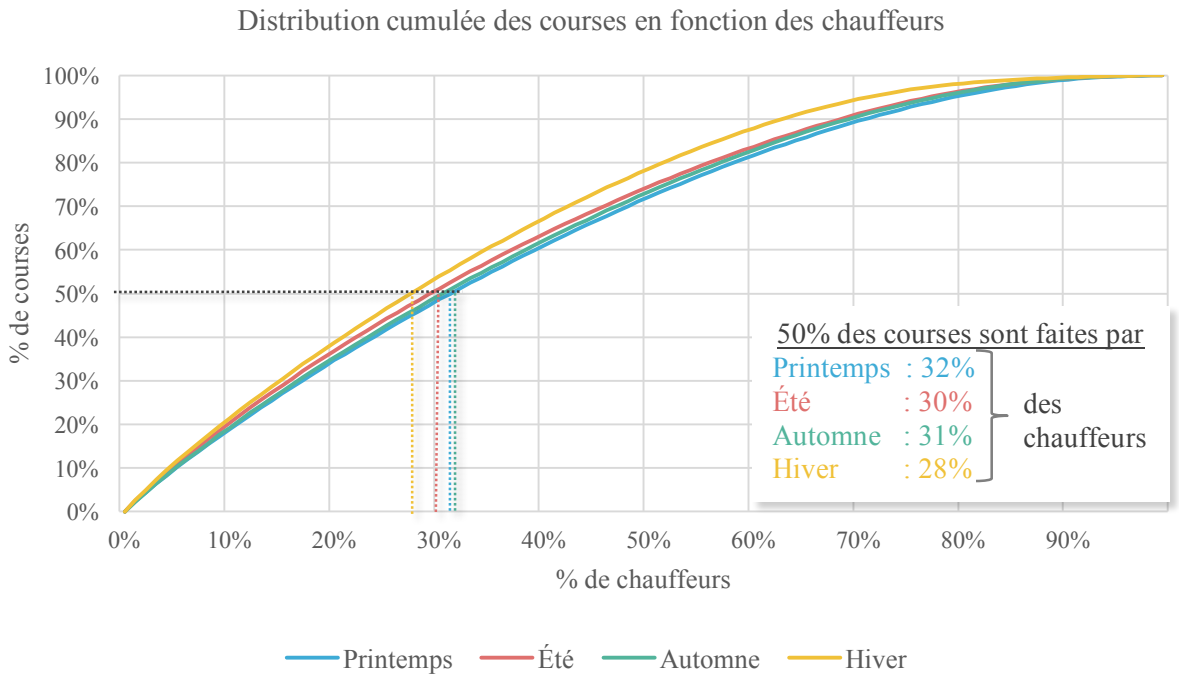


Figure 5-10 : Courses effectuées par chauffeur par saison cumulée pour l'ensemble des chauffeurs

Parallèlement à la figure 5-9 qui montre un étalement plus important vers les valeurs élevées de courses par jour pour l'hiver, cette même saison dévoile qu'un plus petit nombre de chauffeurs fait la majorité des courses. Tandis que l'été, l'automne et le printemps indiquent respectivement que 30 %, 31 % et 32 % des chauffeurs font 50 % des courses, l'hiver indique un pourcentage plus faible, soit 28 %. Peu importe la saison, environ un tiers des chauffeurs assurent la moitié de la demande.

Il est à noter que le nombre de courses n'est pas proportionnel au revenu du chauffeur. Par exemple, un chauffeur ayant fait 10 petites courses aura un salaire moins élevé qu'un chauffeur ayant fait le même nombre de courses de plus longue distance.

La répartition temporelle des courses par jour pour l'année 2015 dévoile des tendances majeures selon différents facteurs. La figure 5-11 met en évidence ces tendances en croisant les jours de fortes précipitations de pluie et de neige ainsi que les jours fériés avec la demande journalière.

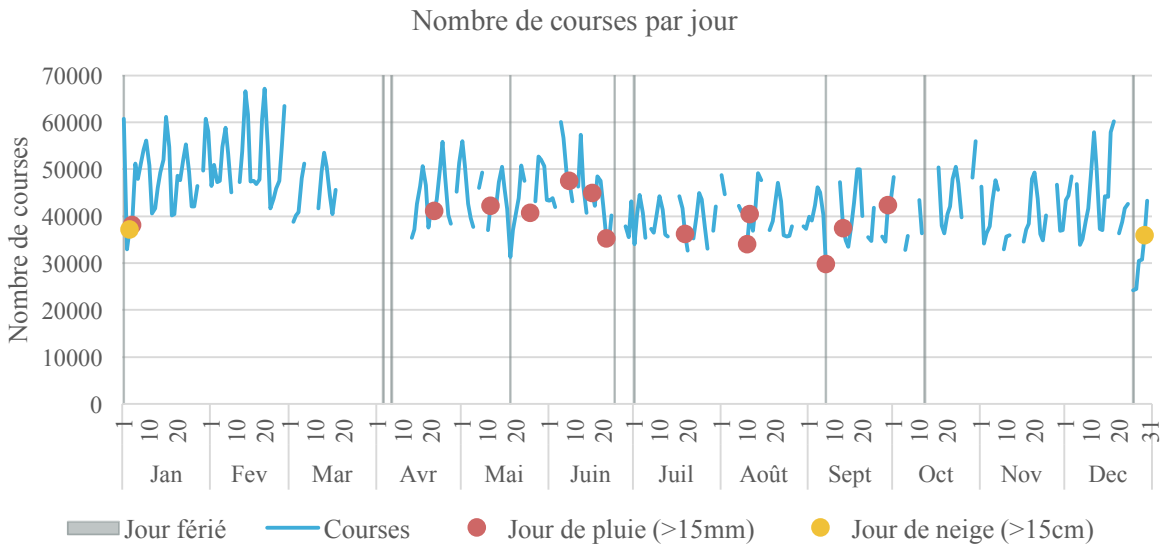


Figure 5-11 : Courses par jour pour l'année 2015 selon les précipitations et les jours fériés

D'abord, une tendance semble se développer selon les saisons. Les mois d'hiver (de fin décembre à fin mars) enregistrent un nombre de courses généralement élevé, pour atteindre à partir de la fin juin jusqu'à la fin septembre la quantité de courses la moins importante de l'année. Les mois de printemps et d'automne sont des mois de transition, présentant des valeurs moyennes. Ce phénomène est lié aux températures des saisons : le froid de l'hiver décourage les individus à effectuer des déplacements à l'extérieur tandis que la chaleur de l'été incite les individus à utiliser des modes qui leur font profiter du soleil, en plus d'être une période de vacances pour les travailleurs.

La journée présentant la plus grande demande est le vendredi 20 février, suivi de près par les autres vendredis de février ainsi que du jour de la Saint-Valentin. Le Jour de l'An est aussi marqué par une forte demande. La seule journée hors hiver pour laquelle le nombre de courses atteint le 60 000 est le 5 juin, le vendredi du Grand-Prix de la F1. En contraste à ces jours, la demande la plus faible se produit le jour de Noël, le 25 décembre, ainsi que les trois jours suivants Noël.

Les jours de fortes pluies semblent inciter les gens à utiliser le taxi. L'allure de la figure paraît refléter le contraire puisque les points roses sont souvent situés dans le bas de la courbe. Pourtant, en comparant avec les mêmes jours de semaine de la saison, la quantité de courses lors des jours de pluie est effectivement supérieure de quelques milliers. Les jours de tempêtes de neige figurent,

quant à eux, parmi les jours de plus faible demande de la saison. Les individus semblent éviter de se déplacer ou bien préfèrent utiliser d'autres modes de transport.

Quatre des cinq jours fériés, pour lesquels les données sont disponibles, montrent une faible quantité de courses. En effet, ils font partie des 15 journées de plus petite demande dans l'année. Il y a toutefois le Jour de l'An qui se positionne dans les jours de forte demande. Le fait que le service de raccompagnement Opération Nez rouge trouve difficilement des bénévoles pour cette nuit-là doit probablement influencer la demande en taxi. Les conclusions sont difficiles à poser à propos des jours fériés puisque près de la moitié font partie des journées où aucune donnée ne fut reçue.

La figure 5-11 révèle une tendance de la demande selon les jours de la semaine. Une analyse détaillée est donc produite à la figure 5-12 par segmentation de la composante temporelle au niveau de la demi-heure. Seul l'automne est présenté puisque cette saison exprime un comportement moyen de la mobilité et parce que les tendances de demande en taxis en fonction de l'heure de la journée sont similaires entre les saisons.

La demande en taxis est très rapprochée de l'utilisation de ces véhicules (taxis avec clients) de la figure 5-5. En effet, lorsqu'un véhicule est occupé, cela signifie qu'il a répondu à la demande. Toutefois, un client émerge à un moment et un endroit précis, tandis qu'un taxi sera occupé pendant une plus longue période de temps et sur un plus grand territoire. C'est pourquoi la répartition temporelle des origines et destinations sur une semaine est nécessaire à l'analyse de la demande.

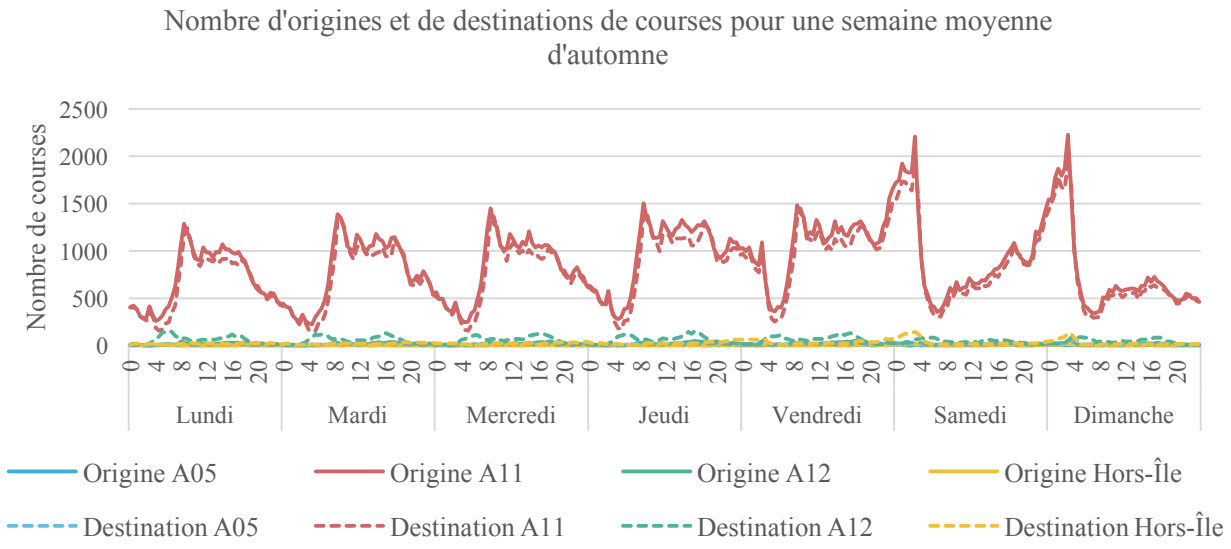


Figure 5-12 : Origines et destinations des courses pour une semaine moyenne d'automne (incrément de 30 min)

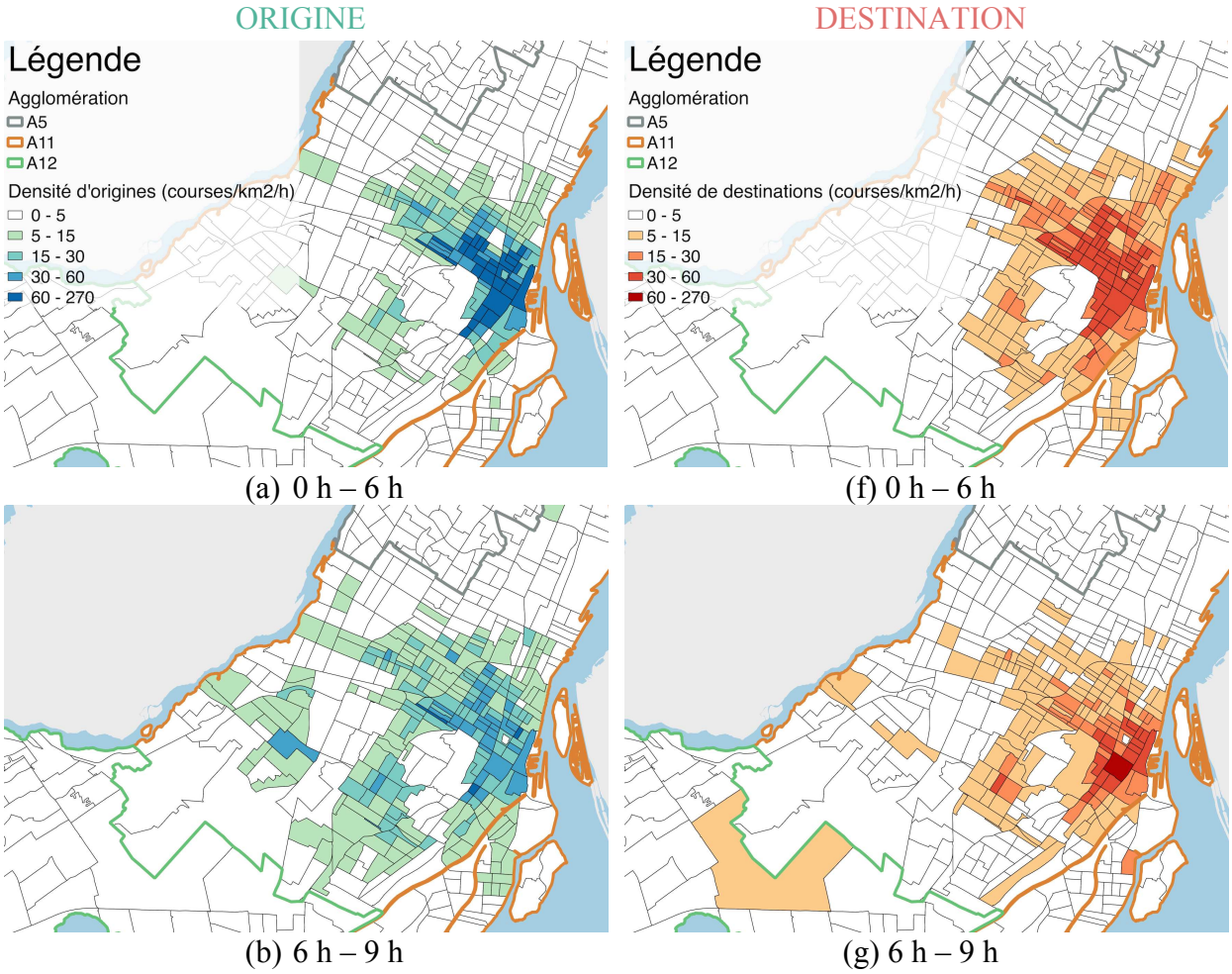
Le lundi, mardi et mercredi présentent une courbe de même allure. Le matin, un pic dépassant 1 250 courses se produit de 8 h 30 à 10 h. Il est légèrement décalé de l'heure de pointe traditionnelle, ce qui porte à croire que les utilisateurs se trouvent dans une situation pour laquelle ils désirent se déplacer plus rapidement qu'à l'habitude. Par la suite, jusqu'à 16 h, la demande oscille autour de 1 000 courses, pour ensuite diminuer progressivement vers un minimum d'environ 250 à 4 h du matin. Le jeudi présente sensiblement la même tendance avec un nombre de courses plus élevé. Il est toutefois marqué par une demande importante de 1 000 courses par demi-heure plus tard en soirée.

Le vendredi est le jour pour lequel le plus grand nombre de courses est enregistré. Tout au long de la journée, il varie entre 1 000 et 1 500. Le plus grand pic de la semaine, atteignant à son sommet plus de 2 000 courses par demi-heure, se situe le soir et la nuit du vendredi. La même situation se produit le samedi, en fin de soirée. Toutefois, durant le jour, le nombre de courses est assez faible, augmentant de 500 le matin jusqu'à 1 000 en après-midi. Le dimanche est la journée la plus tranquille, avec une demande variant autour de 600 courses par demi-heure.

La majorité des clients utilisent les taxis de A11 pour des déplacements à l'intérieur de cette agglomération. Quelques exceptions sont à mentionner. D'abord, le matin ainsi que l'après-midi montent que A12 est sollicitée par les clients. Ceux-ci se dirigent très probablement vers l'aéroport,

qui est le principal attrait de cette agglomération. L'extérieur de l'île est aussi la destination de certains clients qui profitent des activités nocturnes du vendredi et samedi. Une étude approfondie permettrait de déterminer si une rive est plus populaire qu'une autre.

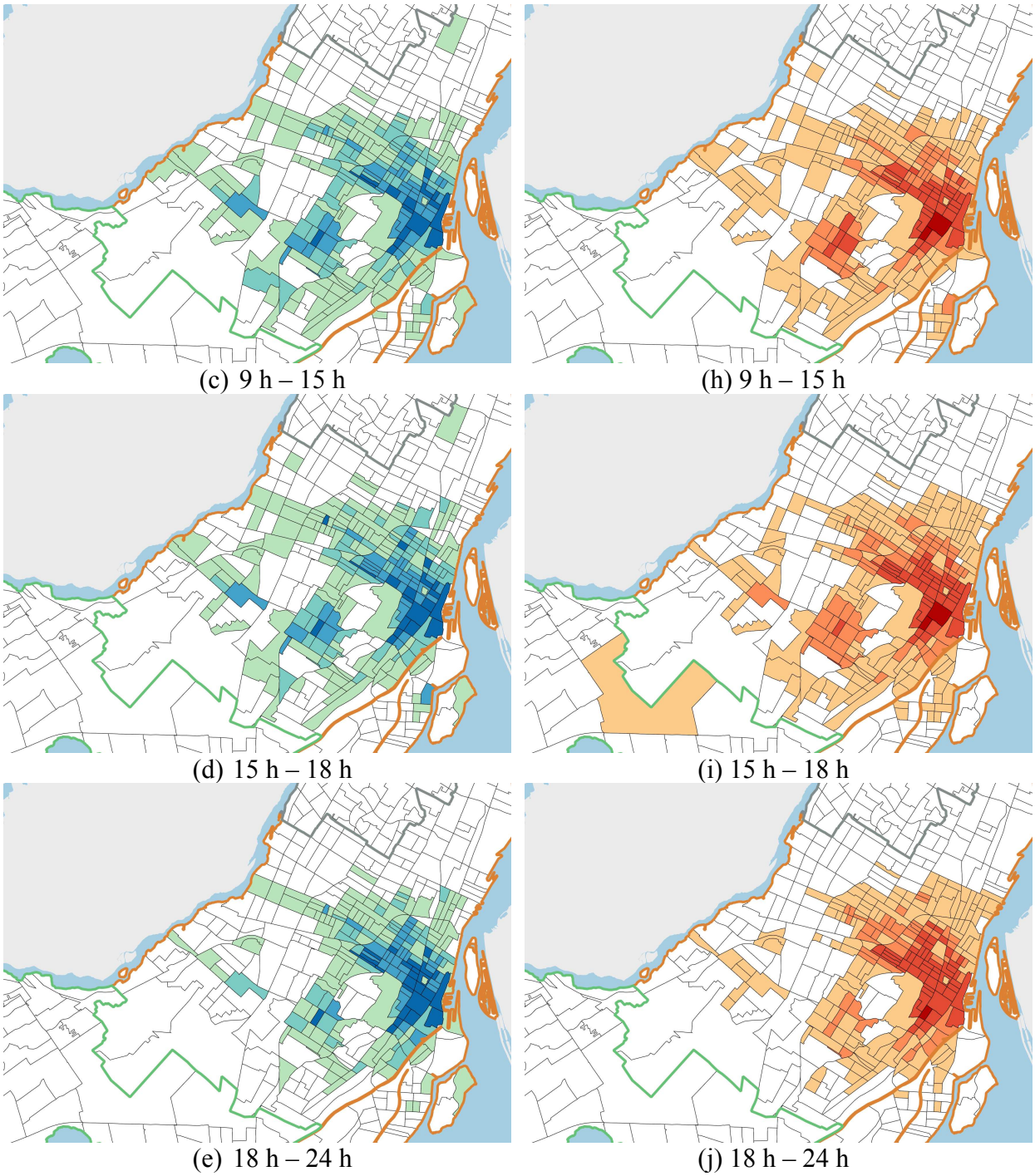
L'analyse de la demande en termes d'origines et de destinations dans le temps se complète par celle dans l'espace. La figure 5-13 présente la répartition spatiale des origines et destinations des courses. Puisque montrer l'évolution de la demande par jour pour tous les jours de la semaine implique un lourd visuel, une journée moyenne a été estimée. Afin de comparer les périodes de la journée entre elles, la densité de courses est rapportée par heure (nombre de courses/superficie/nombre d'heures dans la période).



(a) 0 h – 6 h (f) 0 h – 6 h

(b) 6 h – 9 h (g) 6 h – 9 h

Figure 5-13 : Origines et destinations des courses pour une journée moyenne de 2015



(c) 9 h – 15 h

(h) 9 h – 15 h

(d) 15 h – 18 h

(i) 15 h – 18 h

(e) 18 h – 24 h

(j) 18 h – 24 h

Figure 5-13 : Origines et destinations des courses pour une journée moyenne de 2015 (suite)

En général, les endroits de fortes origines et destinations se situent au Centre-Ville, au Vieux-Port ainsi que dans le Plateau-Mont-Royal. Côte-des-Neiges ainsi que les alentours du boulevard de la Côte-Vertu et Décarie se démarquent aussi en quelques occasions. Les destinations sont réparties sur le territoire de manière plus uniforme que les origines. Plus les secteurs s'éloignent du Centre-

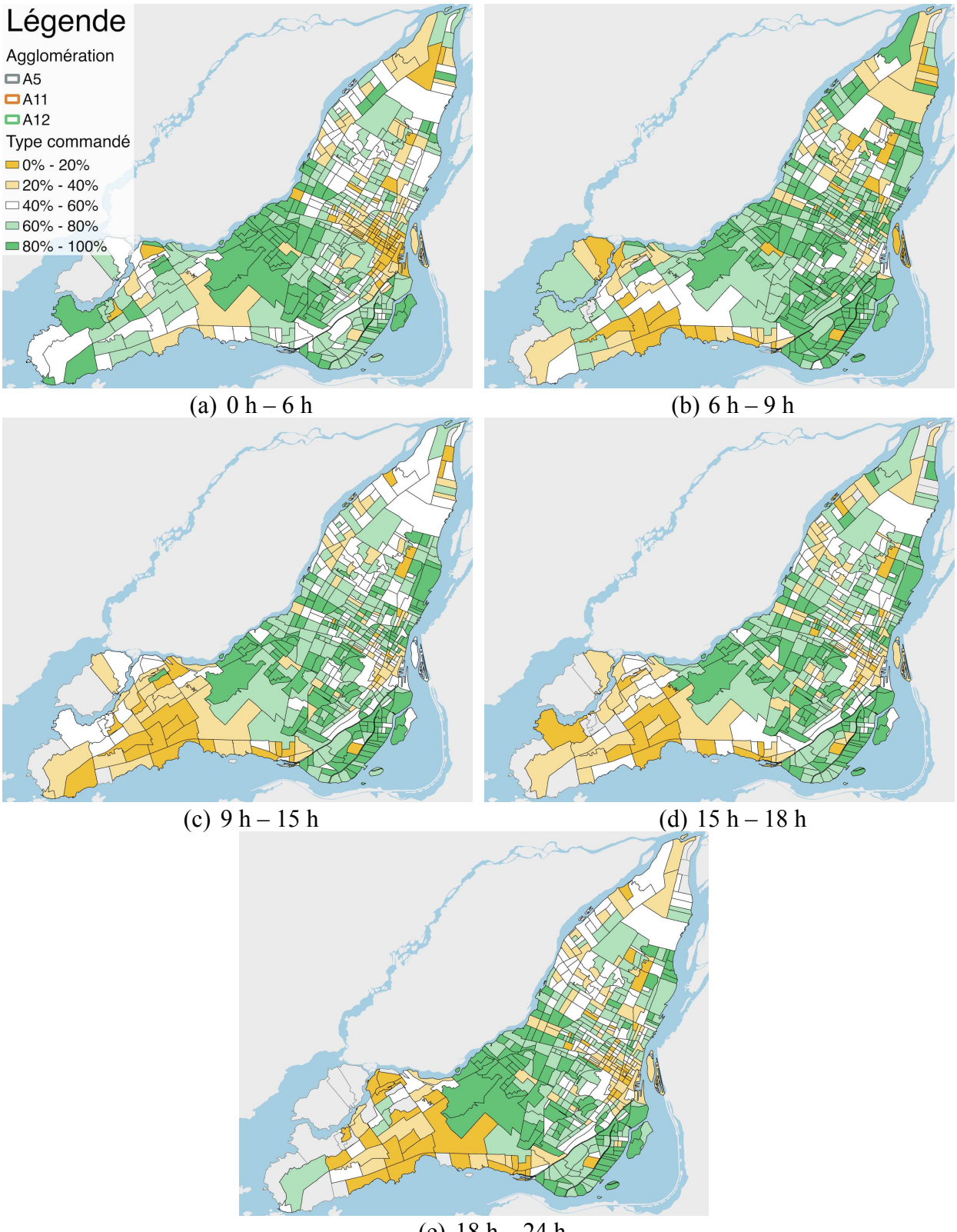
Ville ou de Côte-des-Neiges, moins il y a de destinations tandis que les origines indiquent plusieurs secteurs de forte demande assez éloignés les uns des autres. Une particularité des destinations est qu'elles se rendent jusqu'à l'aéroport, dans l'ouest de l'île.

L'évolution temporelle sur une journée révèle que la nuit, la plupart des individus quittent le Centre-Ville et le Plateau-Mont-Royal. L'inverse est observé durant l'heure de pointe matinale, lorsque les individus provenant de partout dans l'agglomération 11 se dirigent en majorité vers le Centre-Ville. Durant le reste de la journée, les déplacements s'équilibrent assez au niveau des secteurs d'origines et de destinations. Le soir est marqué par une importante concentration de courses au Centre-Ville et dans le Plateau-Mont-Royal.

La présence du métro semble inciter une certaine demande en taxis, particulièrement en origines, puisque les secteurs ont toujours une densité au-dessus de 5 à 15 courses/km^2. Les hôpitaux constituent aussi d'importants pôles de demande, influençant la quantité de courses dans son secteur. Les grands centres commerciaux sont aussi repérables sur la figure 5-13, par exemple la Place Versailles, situé à la frontière est de A11, montre plus de 5 courses/km^2 durant le jour.

Afin de déterminer si les taxis vides se positionnent dans la ville par rapport à la demande, l'analyse combinée de la figure 5-13 avec la figure 5-6 est réalisée. En général, la disponibilité des taxis reflète bien les zones où se destinent les clients, à quelques exceptions près. Toutefois, si les taxis cherchaient les clients aux endroits exacts où ils se trouvent, les secteurs de demande (Centre-Ville, Vieux-Port, Plateau-Mont-Royal et Côte-des-Neiges) verraient autant de véhicules disponibles que les autres secteurs, ce qui n'est pas le cas ici. En heure de pointe, une disponibilité beaucoup trop prononcée comparée à la demande apparaît dans le nord et à l'est de l'île.

Le type de marché du taxi évolue durant la journée. Il est déterminé au tableau 5-6 qu'une majorité de gens commandent par téléphone leur véhicule le matin, tandis que le soir et la nuit, cela se fait plutôt en hélant un taxi sur la rue. Afin d'examiner si une tendance existe aussi dans l'espace, la figure 5-14 est développée. Celle-ci présente la proportion de courses commandées par rapport au nombre de courses total pour différentes périodes d'une journée moyenne. Une segmentation par saison n'est pas nécessaire puisque les variations par saison sont très minimes.



(a) 0 h – 6 h

(b) 6 h – 9 h

(c) 9 h – 15 h

(d) 15 h – 18 h

(e) 18 h – 24 h

Figure 5-14 : Type de courses commandé pour une journée moyenne de 2015

Les changements les plus importants ont lieu au Centre-Ville ainsi que dans l'ouest de l'île. En effet, durant le soir et la nuit, le Centre-Ville fait état d'un marché de taxis hélé. Puis, le matin, il devient plutôt commandé, pour enfin se retrouver neutre le reste de la journée. L'ouest de l'île, quant à lui, est un marché plutôt commandé la nuit, puis devient dès l'après-midi un marché hélé dominant. L'est de l'île est assez variable d'un secteur à l'autre alors que le nord est principalement hélé, et ce, peu importe l'heure de la journée. Les secteurs éloignés du Centre-Ville ne contiennent que peu de courses, ce qui signifie que les résultats sont moins fiables.

La distance des courses est un facteur pouvant expliquer la concentration des origines et destinations dans les mêmes zones. Si celles-ci sont courtes, il est donc naturel d'observer cette concentration. La figure 5-15 et la figure 5-16 présentent une distribution des courses en fonction des périodes de la journée. L'été et l'hiver sont utilisés puisque ces saisons démontrent des comportements différents des moyennes du printemps et de l'automne.

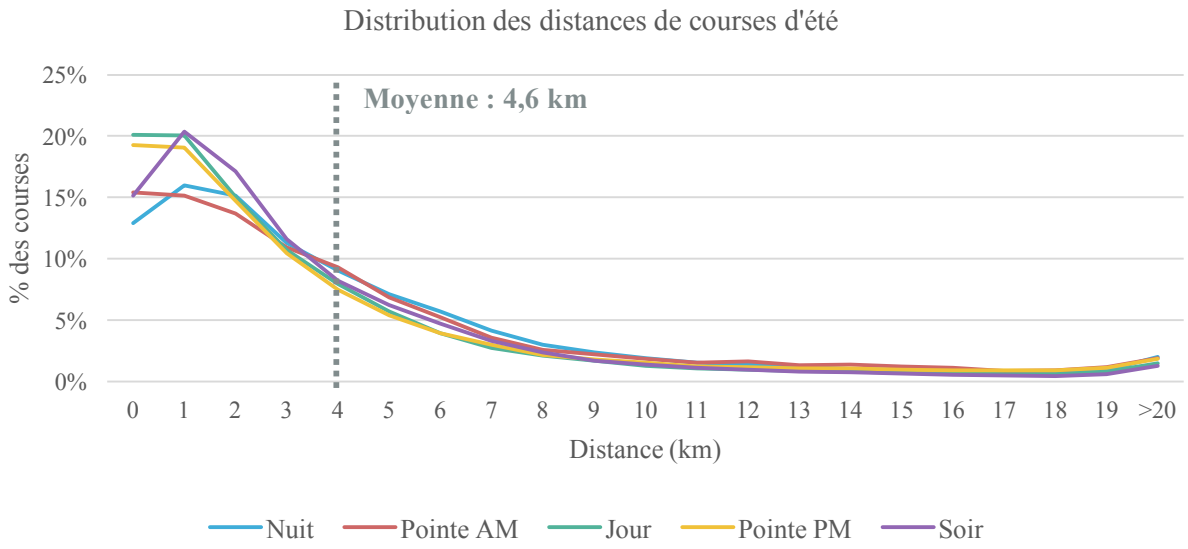


Figure 5-15 : Distances de courses pour l'été 2015

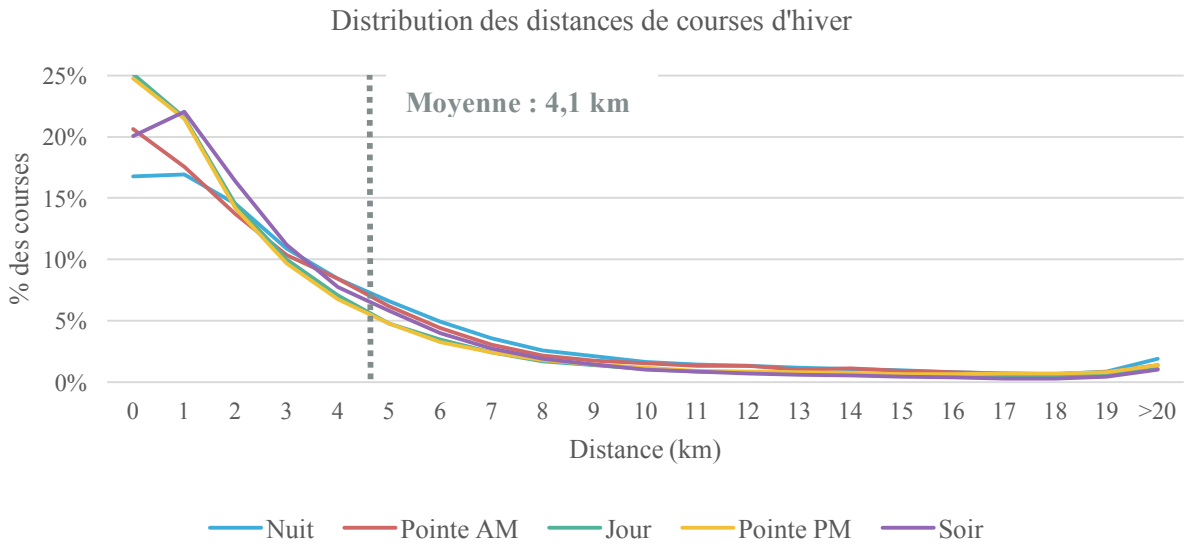


Figure 5-16 : Distances de courses pour l'hiver 2015

La distribution des courses est équivalente entre les saisons à partir d'une distance de 3 km. En effet, plus la distance augmente, plus le nombre de courses diminue pour toutes les périodes de la journée. Les différences entre l'été et l'hiver se situent dans les petites courses. La proportion en été est plus faible que celle en hiver, de l'ordre de 5 points de pourcentage pour chaque groupe d'heures. Cela transparaît dans la moyenne par saison de 4,1 km en été par rapport à 4,6 km en hiver.

Les principales différences entre les périodes se situent aussi dans les petites courses, celles de moins de 2 km. En prenant l'été comme référence, il est observé que ces courses sont en plus faible quantité la nuit (29 %) ainsi qu'en heure de pointe du matin (30 %). Puis, elles atteignent 35 % le soir, suivies respectivement de 38 % et 40 % en pointe de l'après-midi et durant le jour. C'est cette proportion de petites courses qui affecte la distance moyenne par groupe d'heures présentée au tableau 5-6.

Les durées de courses, aspect complémentaire à la distance, sont évaluées à la figure 5-17 pour les mêmes plages d'heures. Puisque le résultat pour chaque saison est semblable, les données pour toute l'année sont utilisées.

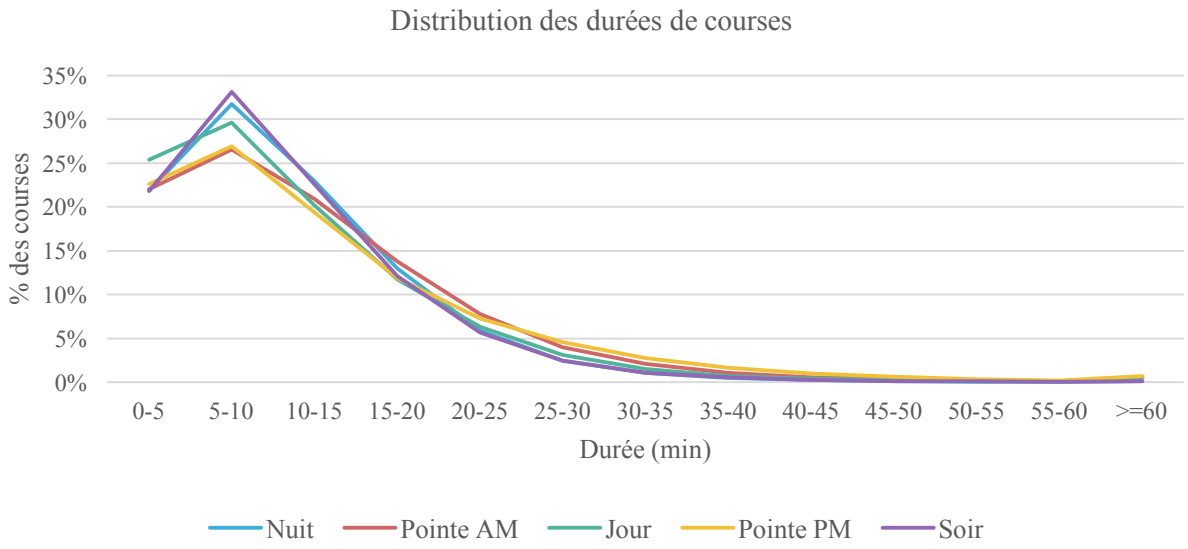


Figure 5-17 : Durée de courses pour l'année 2015

Comme décrit dans les statistiques du tableau 5-6, les heures de pointe ont des moyennes plus élevées, ce qui se reflète par un pourcentage plus petit des courses entre 0 min et 10 min (49 %). Des pourcentages légèrement plus élevés, à 54 % et 55 %, définissent respectivement la nuit ainsi que le jour et le soir. L'allure de ces distributions ne suit pas exactement celle des distances. En effet, la période montrant les plus longues distances est associée à une des plus courtes durées : il s'agit de la nuit. La pointe de l'après-midi montre la situation inverse, soit de courtes distances associées à de longues durées. Comme mentionné plus haut, cela est dû à la congestion routière plus intense durant la journée.

L'utilisation des postes d'attente est un indice pour évaluer si ceux-ci sont bien situés sur le territoire. Les résultats du calcul du nombre de courses débutant aux postes sont exposés à la figure 5-18. Une journée moyenne est utilisée puisque la segmentation par groupe d'heures n'aurait pas fourni assez de données par postes pour être représentative.

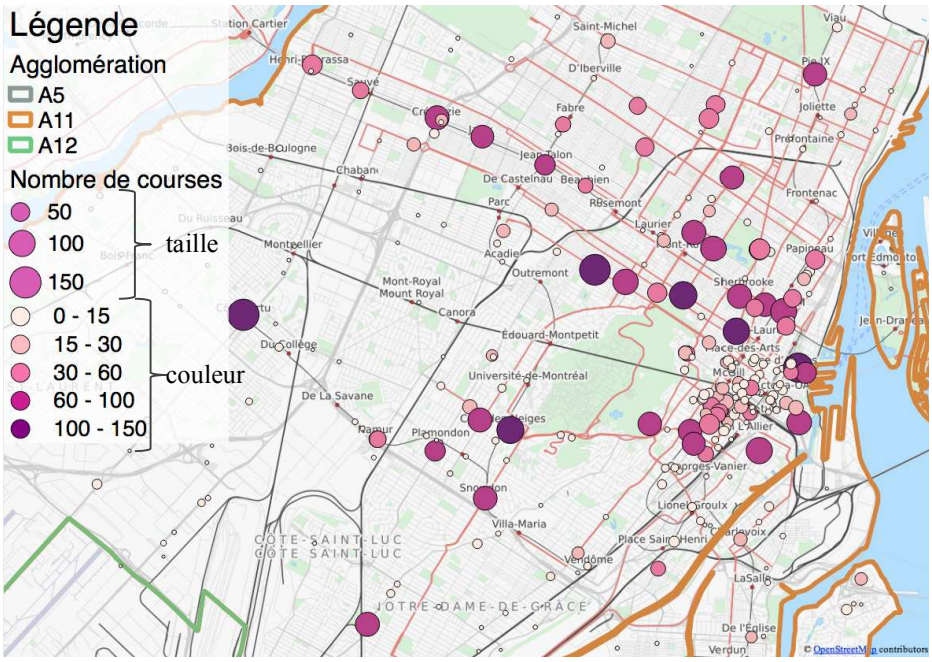


Figure 5-18 : Courses débutant aux postes d'attente pour une journée moyenne de 2015

Contrairement à la croyance générale, ce ne sont pas les postes d'attente du Centre-Ville qui sont les plus utilisés. Ce sont plutôt ceux placés à des endroits stratégiques. La majorité des stations de métro affiche un endroit à proximité où des taxis sont disponibles. Puisque plusieurs utilisateurs ne souhaitent pas prendre l'autobus après le métro, ceux-ci ont l'option de prendre un taxi. C'est ce qu'ils font, puisqu'une quinzaine de stations présente un nombre de courses plus grand que 30 par jour. Cela montre donc la complémentarité de ces modes de transport.

Les autres postes de forte demande semblent être situés un peu plus à l'écart des métros, dans Rosemont ainsi que près de l'avenue du Parc. Les alentours du Centre-Ville sont aussi populaires, plus particulièrement trois postes localisés au Vieux-Port. Des bâtiments spécifiques, comme les hôpitaux, sont aussi très utilisés par les clients qui désirent se déplacer de manière confortable suite à leur visite à l'hôpital.

Il est à noter que certains postes, pour lesquels peu de courses sont enregistrées, ne sont pas nécessairement mal positionnés. Ils peuvent faire partie des postes privés appartenant à d'autres intermédiaires que Taxi Diamond.

5.4 Tableau de bord

Un des objectifs spécifiques de ce projet de recherche consiste à contribuer à un outil de consultation des activités de taxi. La méthodologie développée précédemment établit la base technologique nécessaire pour produire les indicateurs de performance et de suivi. Il est essentiel, par la suite, de rassembler adéquatement ces indicateurs dans un tableau de bord permettant une consultation rapide et facile de l'information.

Un tableau de bord est proposé à la figure 5-19. Celui-ci regroupe les indicateurs selon l'objet du système auquel il appartient. Dus aux limites du média utilisé ici (le papier), certains indicateurs ne peuvent être exploités à leur maximum. Par exemple, les cartes doivent être présentées par journée ou selon des groupes d'heures. Toutefois, le médium envisagé pour la création de l'outil permettra une consultation dynamique de ces indicateurs. Ainsi, reprenant l'exemple précédent, un simple clic permettra de naviguer d'une heure à l'autre sur une carte. Un volet comparatif sera aussi implanté, afin de visualiser un même indicateur selon deux périodes différentes.

Le tableau de bord n'utilise, pour l'instant, que les indicateurs estimés à la section 5.3. De nombreux autres indicateurs ont été mentionnés à la section 4.4.3 et pourraient aisément être rajoutés dans une prochaine version du tableau de bord.

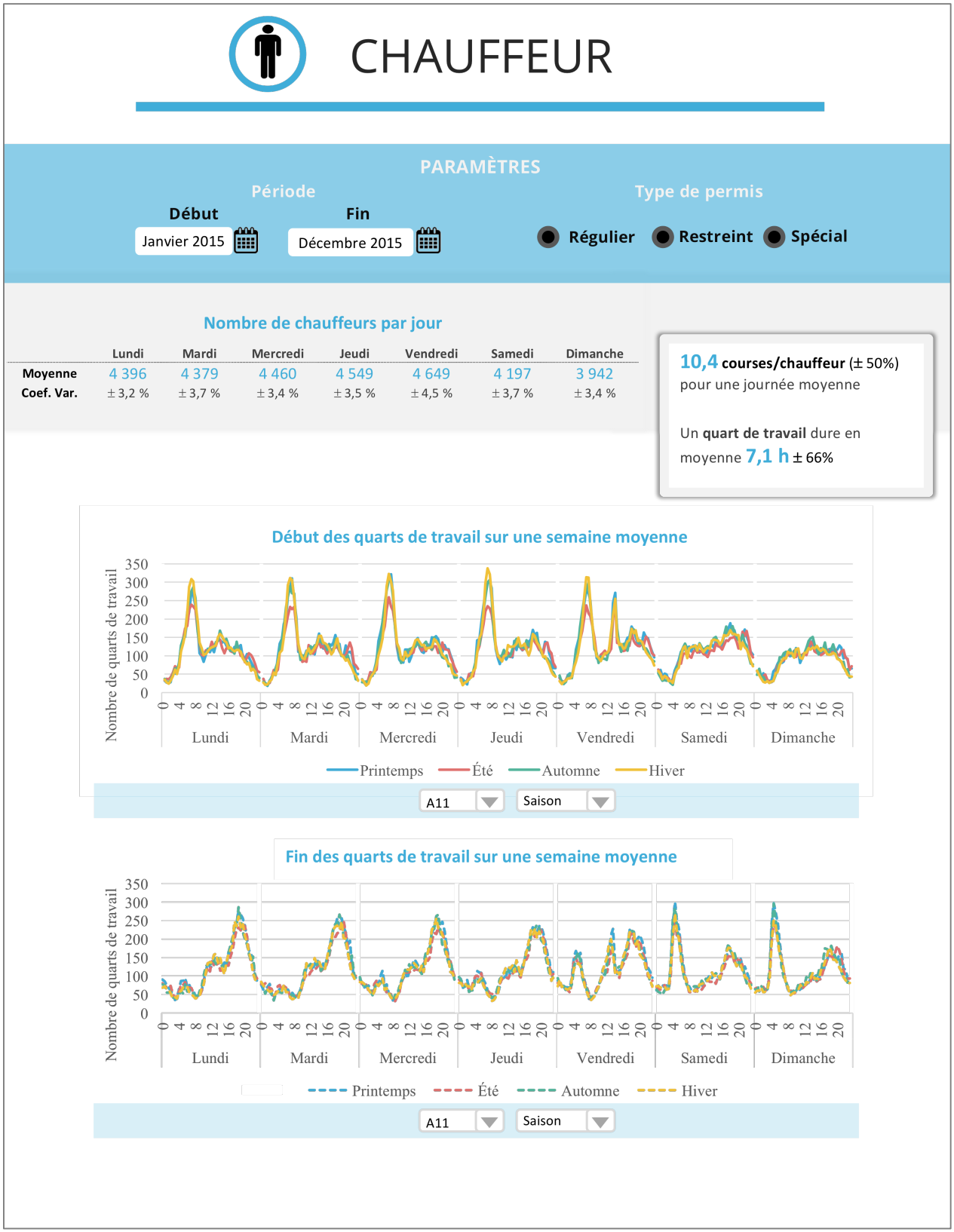


Figure 5-19 : Proposition d'un tableau de bord

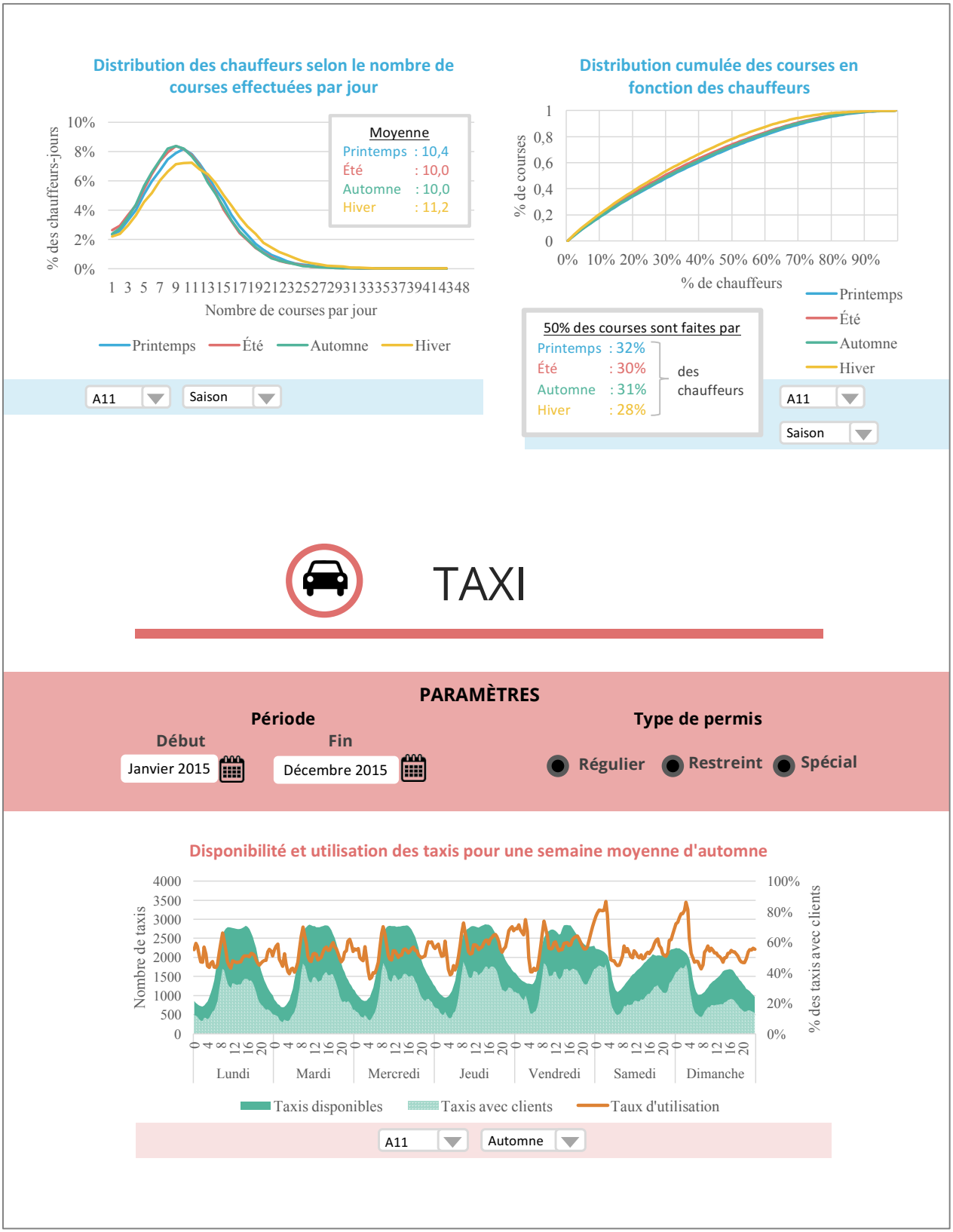


Figure 5-19 : Proposition d'un tableau de bord (suite)

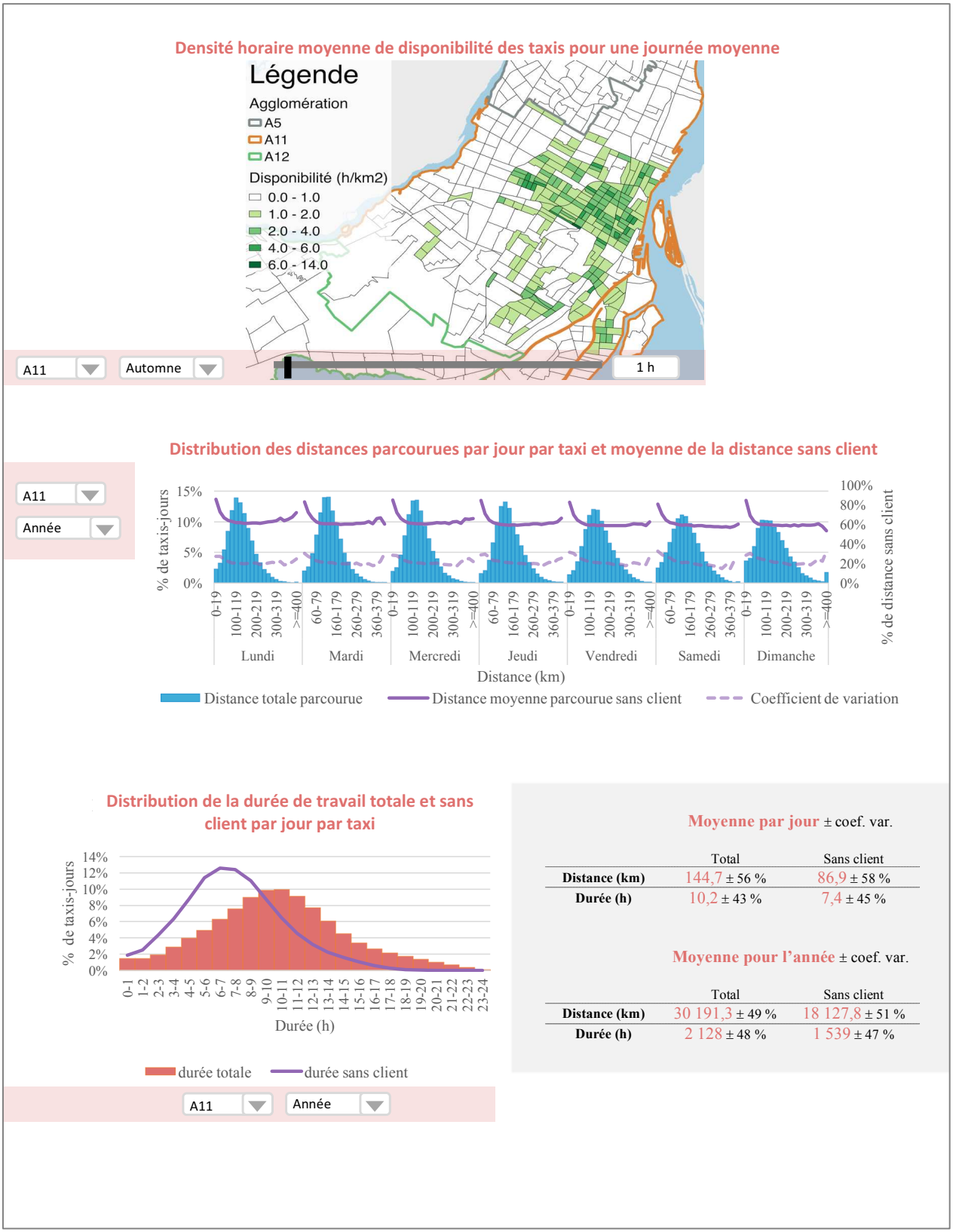


Figure 5-19 : Proposition d'un tableau de bord (suite)

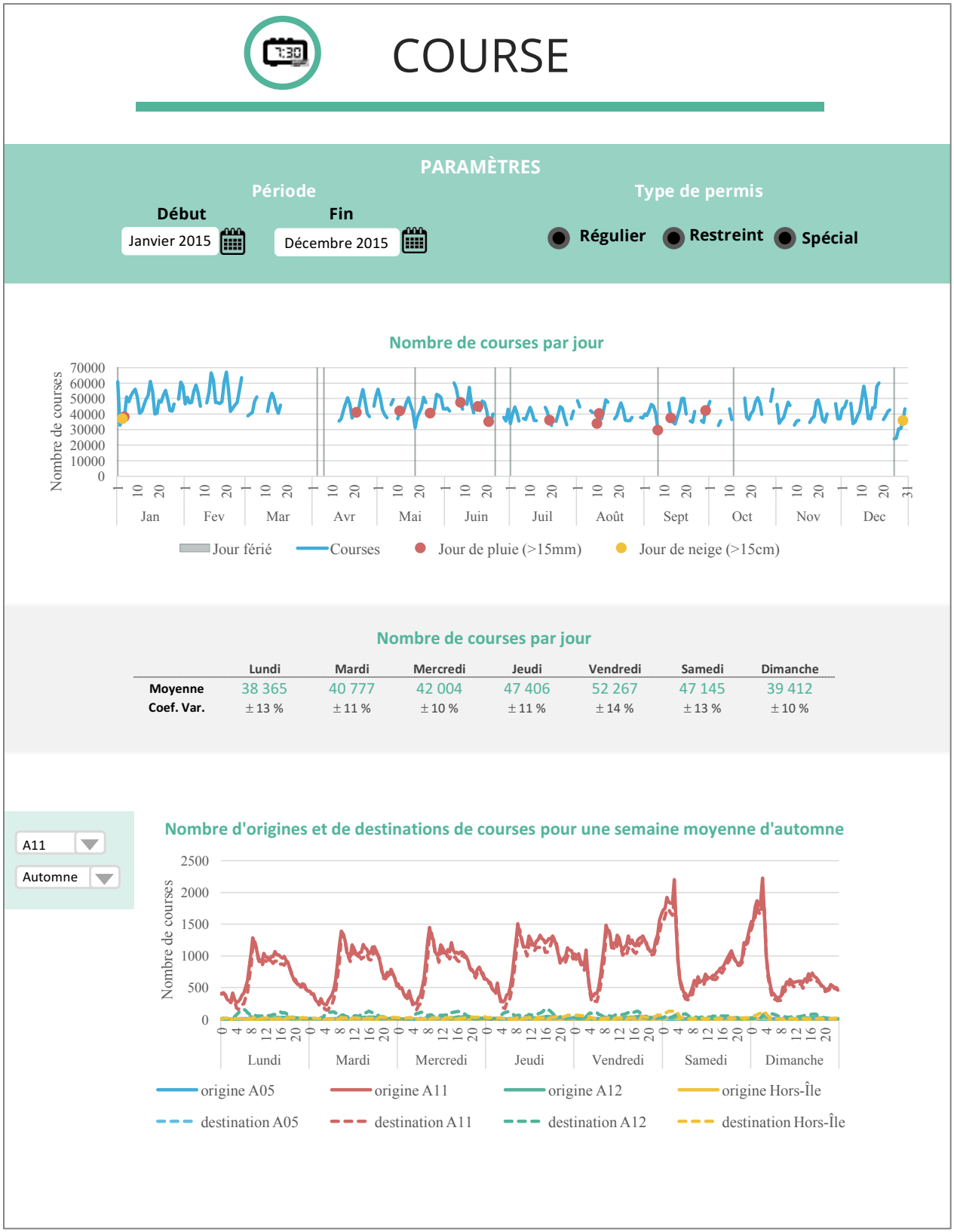


Figure 5-19 : Proposition d'un tableau de bord (suite)

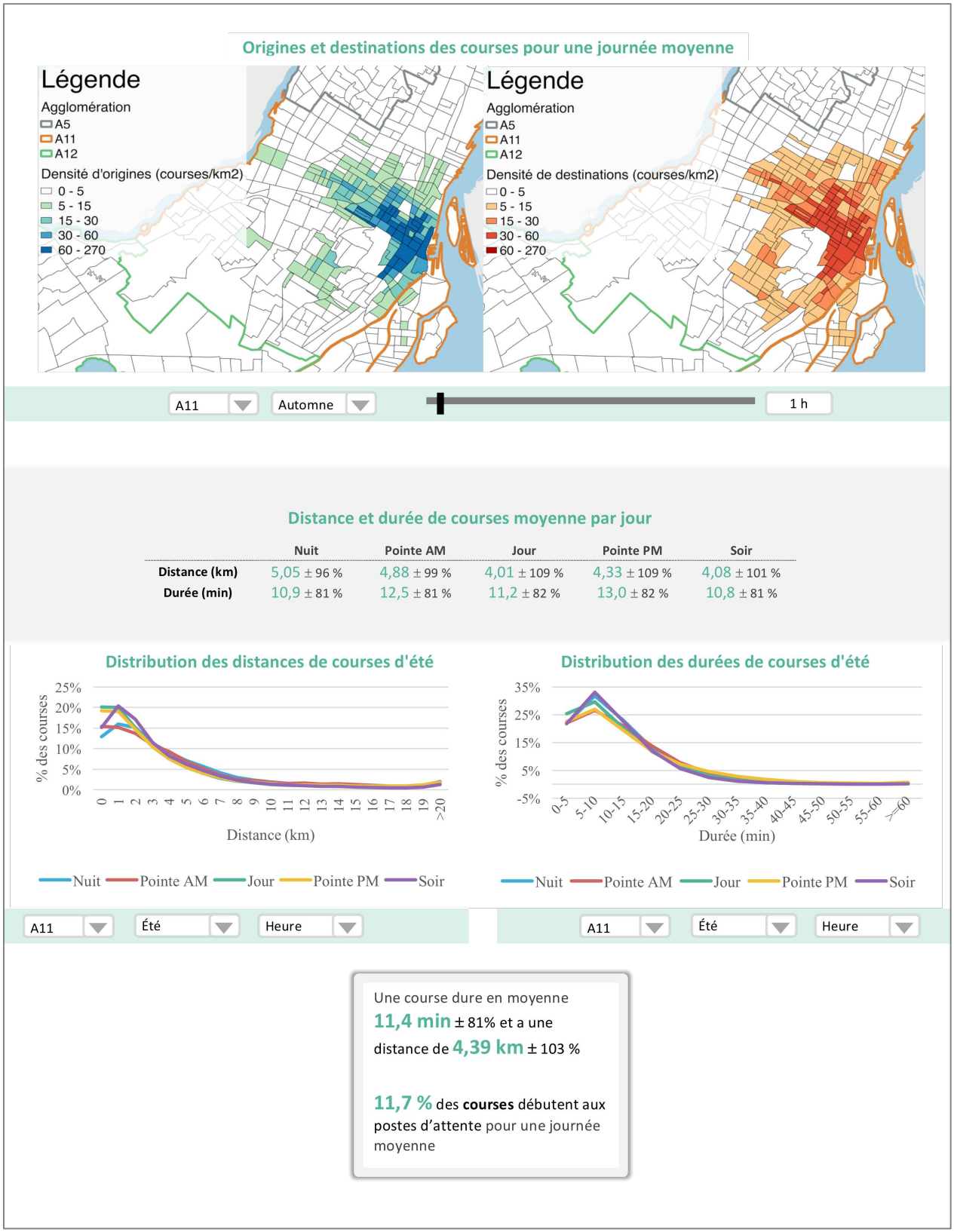
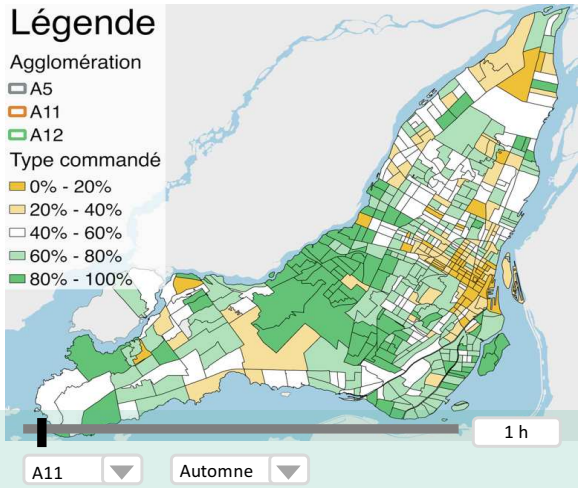


Figure 5-19 : Proposition d'un tableau de bord (suite)

Type de courses commandé pour une journée moyenne



Proportion moyenne de courses commandées par jour

Nuit	Pointe AM	Jour	Pointe PM	Soir
30,6 %	63,5 %	54,8 %	48,7 %	41,8 %

POSTE D'ATTENTE

PARAMÈTRES

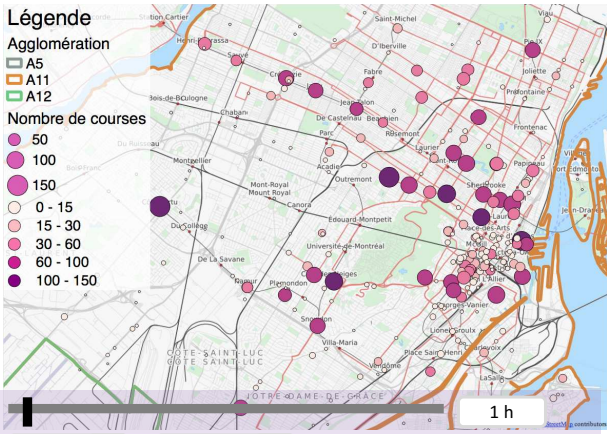
Période

Début: Janvier 2015

Fin: Décembre 2015

Type de permis: Régulier, Restreint, Spécial

Courses débutant aux postes d'attente pour une journée moyenne



Proportion moyenne de courses s'amorçant aux postes d'attente par jour

Nuit	Pointe AM	Jour	Pointe PM	Soir
10,3 %	10,5 %	12,0 %	12,8 %	12,5 %

Figure 5-19 : Proposition d'un tableau de bord (suite)

CHAPITRE 6 MODÉLISATION DE LA DEMANDE

Les données GPS provenant des dispositifs à bord des véhicules présentent un potentiel de support à la modélisation, dû à leur couverture spatiale et temporelle continue (section 2.3). Un des grands enjeux dans l'industrie du taxi est la compréhension de la demande, à savoir comment la clientèle utilise ce mode de transport. Modéliser un tel phénomène permet de prédire l'émergence de la demande ainsi que de cibler les facteurs principaux qui l'influencent.

La modélisation de la demande est un enjeu crucial pour les opérateurs et planificateurs de l'industrie du taxi. En effet, les chauffeurs ne savent pas toujours où se rendre pour trouver leur prochain client et errent dans les rues, participant ainsi à la congestion, consommant plus d'essence et ne fournissant donc pas de service aux voyageurs. S'ils étaient plus informés des lieux et moments de la journée où émerge la demande, leurs temps et kilométrage non productifs seraient réduits. En considérant tous les chauffeurs de la flotte, cette optimisation des services tendrait vers un équilibre de l'offre et la demande, élément majeur de la planification du taxi.

Cette section développe un modèle de génération des courses de taxis dans le but de mieux comprendre la demande. Deux régressions linéaires sont réalisées grâce aux données GPS ; une pour les origines et l'autre pour les destinations des courses. Elles sont agrégées au niveau du jour ainsi que du secteur de recensement et se basent sur des données démographiques, d'utilisation du sol, d'accessibilité au transport en commun ainsi que de conditions météorologiques pour expliquer ce phénomène. Bien que simples, ces modèles sont robustes et facilitent l'identification des variables significatives (Kennedy, 2008).

6.1.1 Données utilisées

Les données utilisées pour développer le modèle sont regroupées en cinq catégories : taxis, population, utilisation du sol, transport en commun et météo.

Les données sur le taxi sont celles de Taxi Diamond et couvrent la période du 1er au 31 octobre 2014. Seules les courses ayant une origine ou une destination dans A11 sont utilisées. Un total de 1 088 véhicules, 1 542 chauffeurs et 383 022 courses sont représentés par les 15 550 033 points GPS du mois.

Les données sur la population proviennent du recensement canadien conduit en 2011. Ce dernier, effectué tous les cinq ans, recueille de l'information sur les familles, les ménages et la population pour chaque secteur de recensement. Il y a 493 secteurs dans l'agglomération de taxi A11. L'enquête nationale auprès des ménages est conduite parallèlement au recensement. Le revenu, le niveau d'éducation ainsi que l'emploi des résidents sont les informations extraites de cette enquête.

Les données sur l'utilisation du sol de 2014, rassemblant l'information sur les zones résidentielles, commerciales et industrielles, sont utilisées.

Les données de transport en commun sont disponibles à partir des fichiers GTFS de la Société de Transport de Montréal. Elles sont utilisées pour estimer la fréquence du service dans le temps et l'espace.

Les données météorologiques historiques, regroupant la température et les précipitations, sont extraites par jour pour le mois d'octobre 2014.

6.1.2 Méthodologie

Trois étapes se succèdent dans le développement d'un modèle de base : d'abord le traitement des données et la définition des variables, ensuite la génération de la base de données, puis l'analyse de corrélation. Ces étapes permettent la sélection des variables pertinentes à inclure dans la régression.

La première étape est la manipulation des données dans le but de créer les variables désirées et d'atteindre le format de données requis. Le temps d'accès au transport en commun (TATC) est l'unique variable à calculer pour le modèle. Parmi les différentes méthodes existantes pour la détermination du TACT, celle utilisée ici est similaire à celle employée par Yang, C. et Gonzales (2014) et se base sur la distance de l'arrêt et le temps d'attente. En fait, le TATC est la somme du temps de marche à l'arrêt et du temps d'attente qui est la moitié de l'intervalle entre les arrivées d'autobus ou de métro. L'équation 6-1 montre la manière de calculer le TATC en minutes :

$$TATC = {}^{60D}/_{v_w} + {}^{30}/_{f} \qquad \text{Équation 6-1}$$

Où

D est la distance jusqu'à l'arrêt de transport en commun (m)

v_w est la vitesse de marche imposée à 5 km/h

f est la fréquence du service de transport en commun (passages/h)

Pour chaque secteur de recensement et jour, le TATC est calculé. Premièrement, à l'aide des données d'adresses (coordonnées pour chaque bâtiment de l'île), la distance à vol d'oiseau entre chaque bâtiment et arrêt de transport en commun dans le même secteur est estimée. La deuxième étape consiste à calculer la fréquence d'autobus ou métro par jour. Grâce aux données GTFS, le nombre de départs à un arrêt est estimé pour chaque type de service (jours de semaine, samedi, dimanche et jour férié). Troisièmement, le calcul du TATC se fait pour chaque paire de bâtiment et arrêt dans le secteur de recensement. Le TATC minimum pour chaque bâtiment est retenu. Finalement, la valeur moyenne de tous les TATC est estimée pour chaque secteur ; ceci correspond à l'information intégrée dans le modèle.

En utilisant les données d'adresse, cette méthode met l'accent sur les zones d'activités humaines. L'hypothèse envisagée est que les gens se déplacent d'un bâtiment à l'autre, ce qui est représentatif de l'attractivité générale des déplacements. Par conséquent, le TATC dépend de la distribution des distances entre les bâtiments et les points d'accès. Cette méthode comporte toutefois ses limitations. D'abord, il est supposé que les gens choisissent l'arrêt le plus proche en considérant la marche et le temps d'attente. Aucune différence n'est faite entre l'autobus et le métro même s'ils montrent des commodités différentes. Puis, la destination du déplacement n'est pas prise en compte. Ainsi, la connexion entre l'arrêt de transport en commun et une destination particulière n'est pas optimale. De plus, seuls les arrêts dans le même secteur de recensement que l'adresse à l'étude sont considérés. Ceci est utilisé pour simplifier la méthode en diminuant le temps de calcul et la mémoire nécessaires. Les limitations ne saisissent pas l'attractivité complète du transport en commun, ce qui laisse place à plusieurs champs d'amélioration dans le calcul du TATC.

Pour éviter le problème de non-linéarité dans la régression, chaque variable continue est analysée à travers un graphique en fonction du nombre de courses afin de détecter si une tendance autre que linéaire existe. Le cas échéant, elle est divisée en catégories présentant chacune la même tendance.

Autant de variables indicatrices sont créées que de catégories. Dans le cas contraire, la variable demeure continue, mais est transformée en pourcentage de sorte que l'interprétation et la comparaison des coefficients soient simplifiées. Cela implique que les variables déjà présentes dans les données, constituant les catégories d'une variable globale (par exemple les groupes d'âge), peuvent être combinées si elles révèlent une tendance similaire.

La seconde étape, la génération de la base de données pour la modélisation, comprend l'agrégation désirée au niveau du temps et de l'espace. Puisque le résultat de la régression est un nombre de courses par jour et par secteur de recensement, il est nécessaire que les variables temporelles et spatiales soient agrégées au niveau du jour et du secteur de recensement. Ainsi, seuls les données météorologiques ainsi que le TATC doivent être définis par jour en plus des données d'utilisation du sol qui doivent être jointes spatialement au secteur de recensement.

La troisième étape est l'analyse de corrélation entre les variables explicatives potentielles puisque cela pourrait biaiser les estimations (Kennedy, 2008). En fait, la corrélation entre les variables signifie qu'elles ont un effet combiné sur le résultat du modèle. Ainsi, l'effet du premier facteur est partiellement inclus dans le deuxième facteur et vice versa, ce qui amplifie ou diminue l'effet et est non désiré. Parmi les variables explicatives corrélées, celle qui présente le coefficient le plus faible avec la variable dépendante est généralement omise. Ici, un coefficient inférieur à | 0,4 | est considérée comme faible. Les facteurs omis sont la température moyenne par jour, les types d'emplois ainsi que certaines des variables indicatrices qui composent un facteur.

La corrélation avec la variable dépendante est également analysée pour guider la sélection des variables du modèle final. Celles ayant une faible corrélation avec la variable dépendante peuvent être omises sans provoquer un effet majeur sur la qualité du modèle. Un coefficient plus petit que | 0,1 | est considéré faible. Ainsi, les précipitations et la densité de population sont les facteurs retirés. Il est vivement conseillé d'ajouter les variables à fort coefficient afin d'améliorer le modèle. De plus, certaines variables, qui sont difficiles à expliquer, sont retirées. C'est le cas pour le type d'industrie de l'emploi des résidents du secteur.

Le tableau 6-1 présente la liste des variables explicatives potentielles. La sélection suite à l'analyse de corrélation est aussi spécifiée pour chaque régression.

Tableau 6-1 : Description des variables

Ensemble de données	Catégorie	Facteur	Nombre de variables	Type	Nombre de variables sélectionnées	
					Origine	Destination
		Jour de semaine	7	Binaire	6	6
Temporel	Taxi	Nombre d'origines	1	Continu	1	
		Nombre de destinations	1	Continu		1
	Météo	Température moyenne	3	Binaire		
		Précipitation (mm)	4	Binaire		
	Transport en commun	Temps d'accès	6	Binaire	1	1
Spatial	Population	Densité de résidents	6	Binaire		
		Âge (% de la population par tranche d'âge)	8	Continu	2	2
		Taille moyenne du ménage	4	Binaire	1	1
		Revenu (% de la population par tranche de revenu)	5	Continu	1	1
		Éducation (% de la population par diplôme obtenu)	6	Continu	1	1
		Type d'emploi (% de la population par type d'emploi)	10	Continu		
		Type d'industrie (% de la population par type d'industrie de leur emploi)	20	Continu		
	Utilisation du sol	Type de sol (% de la superficie par type de sol)	3	Continu	2	2

6.1.3 Résultats et analyses

L'objectif de cette modélisation est de mieux comprendre la demande en taxis. C'est pourquoi une régression linéaire multiple est choisie pour décrire la génération des origines et destinations des courses de taxis dans le temps et l'espace.

Un processus itératif est utilisé pour tester différents modèles : la signification statistique des variables ainsi que la qualité du modèle sont vérifiées à chaque itération. Les variables sont ainsi retirées jusqu'à l'atteinte d'un modèle satisfaisant. Le test de Student est utilisé pour examiner si la variable est significative à un niveau de confiance de 95 %. Celles présentant des signes de

colinéarité, c'est-à-dire celles ayant une forte corrélation avec les autres variables ou celles ayant un coefficient instable dans les multiples modèles, sont retirées de la régression. La qualité du modèle aide à choisir, entre plusieurs régressions linéaires, celle qui explique le mieux la variable dépendante. Le coefficient de détermination (R^2) est le critère le plus populaire pour évaluer un modèle. Il indique la variabilité du résultat par les facteurs explicatifs.

Les résultats du modèle final sont présentés dans le tableau 6-2 : pour chaque variable, le coefficient ainsi que l'effet moyen sur le nombre d'origines et de destinations sont indiqués. Le coefficient est utilisé pour analyser la façon dont les variables influencent chaque modèle. Puisque toutes les variables ne sont pas binaires, le coefficient est multiplié par la valeur moyenne de la variable afin d'obtenir l'impact moyen, permettant la comparaison en termes de nombre de courses.

Tableau 6-2 : Modèles de régression pour l'origine et la destination des courses

		Origines		**Destinations**	
Nombre d'observations		12 369		12 369	
R^2		0,28		0,27	
Variable	**Valeur moyenne**	**Coefficient**	**Impact moyen**	**Coefficient**	**Impact moyen**
Jeudi	0,16	5,79	0,93	5,58	0,90
Vendredi	0,16	8,28	1,34	7,86	1,27
Samedi	0,13	8,77	1,13	8,34	1,08
Dimanche	0,13	3,16	0,41	2,61	0,34
Temps d'accès au transport en commun de moins de 10 min	0,69	10,89	7,56	8,43	5,85
% de la population âgée de 0 à 15 ans	0,14	-29,33	-4,10	-30,56	-4,27
% de la population âgée de 20 à 29 ans	0,18	231,62	40,61	201,94	35,41
% de zone commerciale	0,06	116,41	6,66	110,51	6,33
% de zone de bureaux	0,13	66,31	8,43	67,82	8,63
% des ménages dont le revenu est de 80 000 $ et plus	0,11	117,13	13,20	106,05	11,95
% de la population possédant un diplôme de niveau collégial	0,34	-73,91	-25,34	-75,70	-25,96
Constante	-	-22,17	-	-14,72	-
Notes : - Tous les coefficients sont significatifs à un niveau de 95 %					

La qualité du modèle final est un R^2 de 0,28 pour les origines et 0,27 pour les destinations. Les résultats montrent que les variables ayant le plus grand impact sur la génération de la demande sont, en ordre d'importance : l'âge, le niveau d'éducation, le revenu, le type d'occupation du sol, l'accessibilité au transport en commun et le jour de la semaine. Les deux modèles sont très semblables. Cela montre que l'origine et la destination des courses de taxis sont assez symétriques en matière d'occupation du sol et de jour de la semaine. La différence majeure entre les modèles est que les variables indépendantes influencent la demande liée aux origines légèrement plus fortement que celle liée aux destinations. En effet, les impacts moyens (coefficient multiplié par la valeur moyenne de la variable) pour l'origine sont plus élevés que ceux pour la destination.

L'analyse détaillée des variables du modèle montre que :

- L'âge de la population influence grandement le nombre de courses. Plus les résidents d'un secteur sont âgés de 20 et 29 ans, plus le nombre de courses sera élevé, tandis que l'inverse est observé pour les jeunes âgés de 0 à 15 ans. Cet effet est relatif à la tranche des 30 à 39 ans qui est la catégorie de référence.
- La possession d'un diplôme de niveau collégial reflète fortement ici le salaire lié à ce niveau d'éducation. On remarque à travers le modèle que plus les résidents du secteur détiennent un diplôme collégial, donc plus leur salaire est faible, moins les courses émergeront de et convergeront vers ce secteur.
- Le revenu, troisième variable qui influence le plus la demande, montre que plus un secteur possède des ménages dont le salaire est de 80 000 $ et plus, plus le nombre de courses sera élevé. Cela va de pair avec le nombre moins élevé de courses engendrées par un salaire plus faible tel que montré avec le niveau d'éducation.
- Les zones de bureaux et de commerces font varier positivement le nombre de courses du secteur. Ainsi, plus la superficie de bureaux et commerces est grande, plus il y aura de courses originant et se destinant dans le secteur.
- L'accessibilité au transport en commun est un facteur qui engendre une plus grande demande en taxi. Un secteur ayant un temps d'accès de moins de 10 minutes voit son nombre de courses de taxi augmenter.
- La demande en taxi varie en fonction du jour de la semaine. Comparativement à un mardi, on remarque que le lundi et le mercredi sont non significatifs, c'est-à-dire qu'ils présentent

une distribution des courses trop semblable au mardi. Le jeudi et le dimanche sont aussi similaires en nombre de courses, quoique légèrement supérieurs. Le vendredi et le samedi ont un impact complètement différent, avec un nombre de courses plus élevé que le reste de la semaine.

Des démarches continues sont nécessaires pour améliorer le modèle puisque celui-ci n'explique qu'une part mineure de la variabilité de la demande en taxi (0,28 et 0,27). L'analyse des résidus permet d'établir certains constats dont il faut tenir compte pour cette amélioration. En effet, les graphiques des résidus en fonction des 11 variables explicatives du modèle attestent d'une bonne distribution. Il est donc juste de dire que ces variables ont un effet sur la demande. Par contre, l'analyse des résidus par rapport à la variable dépendante montre que les résidus ont une tendance linéaire positive par secteur, même si la moyenne est de zéro. Cela porte à croire que les variables temporelles utilisées ici n'arrivent pas à expliquer à elles seules la variation de la demande dans un secteur. Il est donc nécessaire d'approfondir les connaissances sur la composante temporelle de la demande afin d'améliorer le modèle. Les régressions élaborées dans ce projet procurent tout de même une vision globale des facteurs d'influence du taxi.

CHAPITRE 7 CONCLUSION

Ce chapitre conclut ce projet de recherche dont l'objectif est de développer des indicateurs de performance et de suivi reflétant un système de transport par taxi. Un retour sur la méthode suivie pour atteindre l'objectif ainsi que sur les résultats et analyses qui découlent de l'étude de cas est d'abord présenté. Par la suite, les contributions puis les limitations de la méthodologie et de l'analyse sont détaillées. Finalement, les perspectives des recherches sur l'analyse du service de taxis offrent un aperçu sur la portée d'un tel projet et les recherches futures.

7.1 Synthèse de la recherche

Premièrement, la revue de littérature a permis de comprendre ce qu'est le service de transport par taxi, depuis les aspects internes de l'industrie jusqu'à son rôle dans la mobilité quotidienne. La définition du service de taxi ainsi que les éléments marquants de son évolution ont mis la table afin de mieux saisir la définition de cette industrie. Ainsi, la littérature qui a trait aux données GPS, à la règlementation, à la répartition de la flotte, aux types d'utilisateurs, aux facteurs d'influence et aux indicateurs spatiotemporels est rassemblée et analysée dans le but de produire un portrait exhaustif des connaissances autour de l'industrie du taxi. Le positionnement du taxi parmi les autres modes de transport est aussi présenté de manière à mieux définir son rôle.

Deuxièmement, l'île de Montréal et son industrie du taxi ont été mises en contexte pour compléter le bagage d'informations nécessaires à l'analyse des activités des taxis de l'étude de cas. La répartition spatiale des infrastructures, principalement les routes, les stations de métro et les grands centres d'activités, en plus de la situation actuelle sur les déplacements dans l'île ont été établies à partir des données des sociétés et agences de transport. Selon l'enquête OD de 2013, le taxi représenterait moins de 1,1 % des déplacements faits quotidiennement (jour moyen de semaine d'automne) par les résidents du territoire. Puis, une description détaillée et spécifique à l'industrie du taxi de Montréal a été produite, incluant sa structure organisationnelle, les différents services offerts et ses aspects réglementaires (agglomérations, nombre de taxis et tarifs).

Troisièmement, la méthodologie d'estimation des indicateurs est développée. Elle a détaillé les différents types de données utilisés et expliqué le choix du système d'information comprenant les bases de données relationnelles. Les objets du système de taxi ont été définis afin de créer les tables associées dans la base de données. La méthodologie comprend aussi le traitement des données

brutes caractérisé par la standardisation, le nettoyage et l'importation des données ainsi que le calcul des champs nécessaires. Les objets ayant subi les plus lourds traitements sont les courses et les points GPS, pour lesquels les doublons, les points aberrants, les courses incohérentes et celles qui sont simultanées sont retirés. L'équation de la pondération des données ainsi que les étapes de calculs des indicateurs de performance et de suivi sont ensuite élaborées dans le développement des indicateurs. Ces indicateurs portent sur quatre objets du système : le chauffeur, le taxi, la course et le poste d'attente.

Quatrièmement, la méthodologie développée a été testée pour estimer les indicateurs sur un intermédiaire majeur en service de taxi pour l'année 2015. Uniquement les taxis opérant dans A11 sont utilisés. Dans un premier temps, les indicateurs de l'offre de service ont été mesurés. Ils ont révélé qu'une grande part des chauffeurs commence à travailler entre 6 h et 8 h le matin et finissent leur journée entre 17 h 30 et 20 h. Les taxis sont principalement disponibles durant le jour à l'est et au nord-est du Mont-Royal. Les taxis sont vides en moyenne 60 % de leur distance et 72 % de leur temps en service. Dans un second temps, les indicateurs de demande ont été calculés grâce aux données de courses. Les constats majeurs ont été que la demande la plus élevée se produit au Centre-Ville et au Plateau-Mont-Royal, le vendredi et samedi en fin de soirée et durant la nuit. Le marché en est un principalement de courses commandées le matin et tend progressivement vers des courses hélées durant la nuit. Les distances et durées moyennes d'une course sont respectivement de 4,39 km et 11,4 min.

Finalement, une modélisation de la demande par l'intermédiaire de régressions linéaires sur les origines et destinations des courses a été estimée. Ce modèle, ayant pour but de cibler les facteurs d'influence de la demande, a utilisé les jours ainsi que les secteurs de recensement comme segmentation temporelle et spatiale ainsi que les données d'octobre 2014 pour les taxis de A11. Les variables ont été sélectionnées suite à l'analyse de la corrélation, de la signification statistique ainsi que du coefficient de détermination. Le modèle final a dévoilé que l'âge, le revenu, les zones de bureaux et de commerces ainsi que l'accessibilité au transport en commun influencent grandement la demande en taxi.

7.2 Contributions

Ce projet de recherche, faisant état de l'industrie du taxi dans son ensemble, apporte des contributions variées à la littérature sur le sujet, à la méthodologie de développement d'indicateurs, à l'analyse des activités de taxis d'une ville ainsi qu'à une future opérationnalisation.

Tout d'abord, le taxi devient de plus en plus un sujet d'étude dans le domaine de la recherche. Les articles portent souvent sur un sujet précis de l'industrie, sans le prendre dans son ensemble et expliquer les impacts dans les différentes sphères du système de transport par taxi. La revue de littérature de cette recherche décrit en détail tous les aspects de l'industrie ainsi que les enjeux reliés. Cela est en soi, une contribution à la connaissance et à la compréhension du sujet. De plus, une liste d'indicateurs pertinents sur les activités des taxis est dressée, ce qui est rarement le cas dans la littérature sur le sujet.

Ensuite, les contributions méthodologiques se concentrent principalement au niveau de la reproductibilité de la méthodologie développée. Le choix et la création du système d'information, qui permet une manipulation facile des données, sont bénéfiques pour tout projet créé dans un but d'estimation d'indicateurs et de modélisation de phénomènes du système de taxis. Les données GPS ont habituellement des champs semblables peu importe la source, ce qui assure une reproductibilité du traitement développé ici. La définition de la course ainsi que les informations ciblées comme étant nécessaires représentent un apport à toutes méthodologies désirant faire l'étude de cet objet. Certains indicateurs n'ont jamais été produits auparavant : il est question ici de la disponibilité spatiale ainsi que la répartition spatiale du type de courses. Le modèle de demande peut aussi contribuer à la prédiction des courses de Montréal dans le temps et l'espace et peut être reproduit pour n'importe quelle ville possédant les données nécessaires.

Puis, l'analyse des indicateurs d'offre et de demande estimés permet de faire des constats sur les activités d'une flotte de taxis. À Montréal, très peu d'analyses avaient été faites à ce niveau. Ainsi, l'ensemble des résultats obtenus pour l'année 2015 est la première étude de grande ampleur pour Montréal. Elle dresse un meilleur portrait, plus représentatif de l'utilisation globale des taxis dans la ville que ceux précédemment établis. De plus, l'estimation de nouveaux indicateurs rajoute de la valeur à cette étude.

Enfin, le fait que cette méthodologie soit reproductible et créée dans le but de développer un outil de diagnostic périodique est une contribution opérationnelle. En effet, les indicateurs seront intégrés à cet outil afin d'assurer un suivi dans le temps et l'espace de l'offre et de la demande. Le tableau de bord proposé offre aussi un design sur lequel pourra se bâtir l'outil. Cet outil permettra de produire des résultats favorisant la planification et l'opération de ce service de transport.

7.3 Limitations

Les hypothèses posées tout au long de la recherche affectent la qualité des résultats et l'interprétation de l'analyse. Les principales limitations sont catégorisées ci-dessous :

1. Service de taxis

   Un taxi peut travailler pour plusieurs intermédiaires de manière alternée durant le mois ou simplement quitter son fournisseur pour en rejoindre un autre. Ces cas sont difficiles à capter puisque les données reçues sont anonymes. Lorsqu'une pondération doit être réalisée pour plusieurs intermédiaires, cela complexifie la méthode de calcul. Toutefois, les taxis travaillant de manière alternée sont rares, car les propriétaires ne désirent pas nécessairement payer les cotisations de plusieurs intermédiaires.

   La flotte de taxis n'est pas composée uniquement des véhicules liés à un intermédiaire en service. De nombreux propriétaires désirent travailler à leur compte, de manière indépendante. Leur comportement est probablement différent des autres taxis puisqu'ils ne reçoivent pas d'appels ou en reçoivent uniquement de clients privés. Cela devrait être pris en considération dans la pondération des données.

2. Source de données

   L'utilisation des données d'un seul intermédiaire est une grande limitation de la représentativité des résultats. Les différents types de permis ainsi que les agglomérations de l'île ne sont pas représentés adéquatement, ce qui empêche de conclure sur l'ensemble des activités des taxis. Ainsi, l'acquisition des données d'autres intermédiaires va permettre d'obtenir un échantillon représentatif du service de taxi à Montréal.

   Une source de données pour laquelle les informations ne sont pas encore obtenues est la STM, qui fournit des courses de transport adapté aux intermédiaires en service. Certaines

de ces courses sont payées à l'heure ce qui fait que les chauffeurs n'ont pas à démarrer leur taximètre. Les points GPS n'ont donc pas de numéro de course associé et sont considérés dans les portions de distances et durées sans client.

Le type de courses (régulier, adapté, etc.) n'est pas indiqué dans les données reçues. Cette information permettrait de mieux déterminer les permis restreints et fournirait une piste d'analyse sur les utilisateurs du service.

Certains évènements de la course semblent mal enregistrés dans les données brutes. Il s'avère que les chauffeurs utilisent mal leur outil de travail, indiquant, par exemple, qu'un paiement se produit avant un démarrage du taximètre. Ainsi, cette situation complexifie l'identification des courses et du statut du chauffeur.

Les conditions météorologiques représentent un facteur qui influence l'utilisation du taxi. Elles sont extraites à partir d'une seule station, cependant les précipitations sont variables sur le territoire de l'île de Montréal pour le même moment. Il serait donc pertinent de prendre en compte les données de plusieurs stations pour les joindre aux limites spatiales utilisées dans le projet.

3. Traitement des données

   Le processus général du traitement de données comporte des pistes d'améliorations possibles, particulièrement au niveau de la détection des points aberrants ainsi que des courses valides. Les points aberrants et décalés doivent être identifiés et retirés pour éviter de biaiser le calcul des distances. Les courses invalides, celles de très longues ou de très courtes distances et durées, doivent être étudiées pour déterminer lesquelles sont réelles et lesquelles proviennent d'erreurs de manipulation des chauffeurs. Leur exclusion du calcul des indicateurs de courses fournirait des résultats plus exacts.

4. Attribution du permis

   La méthode actuelle d'attribution des agglomérations de permis, tel qu'elle a été développée dans ce projet, ne garantit pas l'identification des permis restreints ni des permis spéciaux. Cela engendre un biais dans l'expansion des résultats au niveau de la flotte globale puisqu'il est ainsi supposé que tous les permis ont les mêmes comportements. L'obtention de cette information directement dans les données brutes serait idéale, toutefois

une solution temporaire pour détecter un permis spécial peut être envisagée en utilisant la localisation précise des origines de courses à l'aéroport.

5. Segmentation spatiale et temporelle

   Les nombreux jours manquants peuvent évidemment apporter une limitation dans l'interprétation des résultats puisqu'ils ne sont pas répartis uniformément dans l'année. Ainsi, il faut s'assurer que chaque groupe créé suite à la segmentation possède un nombre minimum d'observations. Par exemple, lors de l'analyse des jours de semaine d'une saison, il est nécessaire que chaque jour soit présent au moins une fois dans la saison.

   Il a été montré que la durée moyenne des courses est de 11,4 minutes. Cela remet donc en question l'incrément de 30 minutes utilisé dans l'indicateur de disponibilité des taxis puisque le véhicule ne sera pas avec un client la majorité de la plage de temps. Il serait donc pertinent de raccourcir l'incrément à 15 minutes. De cette manière, un taxi commençant une course à 13 h sera la majorité de la plage (13 h -13 h 15) avec un client. Aussi, un taxi commençant une course à 12 h 58 sera indiqué utilisé de 12 h 45 à 13 h 15 et non de 12 h 30 à 13 h 30 comme c'est le cas avec un incrément de 30 minutes, ce qui ajoute une précision à l'indicateur.

7.4 Perspectives

L'industrie du taxi est un domaine pour lequel l'intérêt des chercheurs émerge. Puisque peu d'études ont déjà été produites, les possibilités d'exploration dans ce milieu sont nombreuses. Les axes de perspectives spécifiques à ce projet concernent le suivi de l'industrie, l'ajout de données explicatives, le perfectionnement des indicateurs, l'amélioration du modèle de demande, la connaissance des utilisateurs ainsi que la contribution à la mobilité durable.

D'abord, les aspects règlementés de l'industrie du taxi ne restent pas stables dans le temps. Il est précisément question ici du nombre de permis émis par la CTQ. Celle-ci peut délivrer des permis et leur refuser le droit d'exploitation à n'importe quel moment. Le besoin de suivre ces données dans le temps devient nécessaire dans l'optique où les analyses peuvent couvrir une longue période.

Ensuite, les données explicatives de ce projet n'incluent pas toutes les données pouvant expliquer les variations dans les activités de taxis. Les évènements spéciaux, comme les spectacles, les compétitions sportives, les congrès ou les festivals, sont des moments spécifiques à un lieu où la

demande en taxi est forte. Cette information est difficile à capter, mais serait un grand atout pour le développement d'indicateurs et la modélisation de la demande.

Puis, les indicateurs de ce projet sont développés dans un objectif de planification, ils sont donc conçus pour montrer les tendances globales des activités de taxis. Toutefois, des indicateurs complexes seraient bénéfiques sur plusieurs plans. La combinaison de plusieurs objets dans un même indicateur montrerait l'interaction qui existe entre eux et assurerait une évaluation des performances plus complète que celle présentée. Des indicateurs de comparaison permettraient d'évaluer les performances de la flotte à des fins opérationnelles. D'autres indicateurs ciblant l'interaction entre les agglomérations et les différents centres d'activités seraient pertinents à produire pour émettre des recommandations sur la réglementation en place.

En ce qui a trait au modèle de demande en taxi, différentes perspectives sont envisageables pour son amélioration. De nouvelles variables explicatives devraient être incorporées au modèle pour tenter d'augmenter le R^2, par exemple la densité des emplois, le type d'emploi ou le nombre de places de stationnement dans le secteur. Le niveau d'agrégation des données devrait être plus précis (passer du jour à l'heure), puisque les heures de la journée montrent des tendances propres. La période d'un mois de données utilisée dans l'analyse devrait être allongée afin de couvrir la demande en taxi durant des périodes plus variables comme l'hiver. L'accès à des données GPS de plusieurs autres intermédiaires permettrait d'obtenir la demande complète en taxi sur l'île. Enfin, il serait nécessaire de modéliser l'interaction entre le taxi et les autres modes alternatifs afin de le prendre en compte dans le modèle.

Par la suite, la connaissance de la clientèle et de ses besoins est nécessaire pour faire progresser le service de taxis. Le modèle de demande développé a permis d'obtenir des facteurs d'influence qui représentent des indices sur les caractéristiques des utilisateurs du service grâce aux données sur la population. Une meilleure façon d'obtenir ce type de données serait d'élaborer une enquête auprès des clients à bord des véhicules. Ainsi, l'enquête permettrait d'obtenir des informations directes sur les clients et non dérivées à partir d'autres données.

Enfin, dans le contexte où plusieurs villes (dont la Ville de Montréal (2008)) misent sur le développement durable de leurs transports, les contributions du taxi à la mobilité durable devraient être mesurées. De nouvelles données seraient nécessaires, par exemple le type de véhicule, sa consommation d'essence, etc. Des indicateurs devraient être élaborés pour estimer cette

contribution, dont certains prendront en considération les interactions du taxi avec les autres modes de transport.

Élaborer une enquête auprès des clients et développer des indicateurs de développement durable font partie des principales perspectives du grand projet dans lequel s'inscrit cette recherche. Le projet de grande envergure vise à comprendre la mobilité par taxi, évaluer le marché actuel et le potentiel de ce mode dans la mobilité quotidienne ainsi que dresser les contributions de cette industrie au développement durable.

Finalement, les résultats de ce projet vont mener à des recommandations sur différents aspects de l'industrie pour en améliorer les services et l'impact environnemental. Il fournira aussi au BTM un outil de diagnostic périodique utile à la planification et à l'opérationnalisation du service de taxis. Cet outil automatisé permettra aux représentants de l'industrie d'améliorer leur compréhension des dynamiques de l'offre et de la demande en taxis.

BIBLIOGRAPHIE

Afian, A., Odoni, A., & Rus, D. (2015). *Inferring Unmet Demand from Taxi Probe Data.* Communication présentée à Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on (p. 861-868). doi: 10.1109/ITSC.2015.145

Agence Métropolitaine de Transport. (2015). *Enquête Origine-Destination 2013 : La mobilité des personnes dans la région de Montréal*. Tiré de https://www.amt.qc.ca/fr/a-propos/portrait-mobilite/enquete-od-2013

Austin Jr., A. B. (2011). *The Taxicab as Public Transportation in Boston.* (Massachusetts Institute of Technology).

Bacache-Beauvallet, M., & Janin, L. (2012). Taxicab licence value and market regulation. *Transport Policy, 19* 57-63. doi: 10.1016/j.tranpol.2011.08.001

Beimborn, E. A., Greenwald, M. J., & Jin, X. (2003). *Transit accessibility and connectivity impacts on transit choice and captivity*. Communication présentée à Transportation Research Board 82nd Annual Meeting, Washington, DC.

Beynon-Davies, P. (2013). *Business information systems* (2e éd.): Palgrave Macmillan.

Bischoff, J., Maciejewski, M., & Sohr, A. (2015). *Analysis of Berlin's taxi services by exploring GPS traces.* Communication présentée à Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2015 International Conference on (p. 209-215). doi: 10.1109/MTITS.2015.7223258

Bureau du Taxi de Montréal. (2011). *Planification triennale 2011-2013 : Une image de marque pour les taxis de Montréal*.

Bureau du Taxi de Montréal. (2015). Ville de Montréal - Bureau du taxi et du remorquage - Bureau du taxi de Montréal. Tiré de http://ville.montreal.qc.ca/portal/page?_pageid=8177,89603648&_dad=portal&_schema=PORTAL

Cairns, R. D., & Liston-Heyes, C. (1996). Competition and regulation in the taxi industry. *Journal of Public Economics, 59*(1), 1-15.

Cervero, R. (1997). *Paratransit in America: Redefining Mass Transportation*: Greenwood Publishing Group.

Chang, H.-w., Tai, Y.-c., & Hsu, J. Y.-j. (2009). Context-aware taxi demand hotspots prediction. *International Journal of Business Intelligence and Data Mining, 5*(1), 3-18. doi: 10.1504/IJBIDM.2010.030296

Chatman, D. G. (2013). Does TOD need the T? On the importance of factors other than rail access. *Journal of the American Planning Association, 79*(1), 17-31. doi: 10.1080/01944363.2013.791008

Cooper, J., Mundy, R., & Nelson, J. (2010). *Taxi! : Urban Economies and the Social and Transport Impacts of the Taxicab*: Ashgate Publishing.

Dai, J., & Chen, X. (2014). *Empirical Evidence on Taxis' Route Choice Set Using GPS Data.* Communication présentée à CICTP 2014@ sSafe, Smart, and Sustainable Multimodal Transportation Systems (p. 3117-3126). doi: 10.1061/9780784413623.298

Darbéra, R. (2010). Transports publics et taxis: concurrence ou complémentarité? *Ville, Rail et Transport* 34-39.

Deep Singh, A., Wu, W., Xiang, S., & Krishnaswamy, S. (2015). *Taxi trip time prediction using similar trips and road network data.* Communication présentée à Big Data (Big Data), 2015 IEEE International Conference on (p. 2892-2894). doi: 10.1109/BigData.2015.7364113

Ding, Y.-r., Xiong, J., & Liu, H.-j. (2013). The Self-Adapted Taxi Dispatch Platform Based on Geographic Information System. Dans *Emerging Technologies for Information Systems, Computing, and Management* (p. 771-779): Springer.

Doraiswamy, H., Ferreira, N., Damoulas, T., Freire, J., & Silva, C. T. (2014). Using topological analysis to support event-guided exploration in urban data. *Visualization and Computer Graphics, IEEE Transactions on, 20*(12), 2634-2643. doi: 10.1109/TVCG.2014.2346449

Dow, C.-R., Chen, H.-C., & Hwang, S.-F. (2015). *A hotspot aware taxi zone queuing system.* Communication présentée à Computing and Communication (IEMCON), 2015 International Conference and Workshop on (p. 1-4). doi: 10.1109/IEMCON.2015.7344451

Faucher, J. (2013). *Conception d'un système d'information pour soutenir l'analyse des impacts d'une nouvelle infrastructure de transport.* (École Polytechnique de Montréal, Montréal).

Ferreira, N., Poco, J., Vo, H. T., Freire, J., & Silva, C. T. (2013). Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips. *Visualization and Computer Graphics, IEEE Transactions on, 19*(12), 2149-2158. doi: 10.1109/TVCG.2013.226

Ganti, R., Mohomed, I., Raghavendra, R., & Ranganathan, A. (2011). Analysis of data from a taxi cab participatory sensor network. Dans *Mobile and ubiquitous systems: Computing, networking, and services* (p. 197-208): Springer.

Gómez-Gélvez, J. A., Rodriguez-Valencia, A., & Restrepo, N. (2015). *The Role of Taxis in Bogotá, Colombia: Analysis from a Disaggregate Mobility Survey.* Communication présentée à Transportation Research Board 94th Annual Meeting.

Loi concernant les services de transport par taxi, S-6.01 C.F.R. (2011).

Gwilliam, K. M. (2005). Regulation of taxi markets in developing countries: issues and options. *Transport Notes series, TRN*(3), 7.

Harding, S., Kandlikar, M., & Gulati, S. (2015). Taxi apps, regulation, and the market for taxi journeys: Vancouver Institute for Resources, Environment and Sustainability.

Hernandez, M. J. (2013). *Database Design for Mere Mortals: A Hands-On Guide to Relational Database Design* (3e éd.): Pearson Education.

Hsueh, Y.-L., Hwang, R.-H., & Chen, Y.-T. (2014). *An effective taxi recommender system based on a spatiotemporal factor analysis model.* Communication présentée à International Conference on Computing, Networking and Communications (ICNC), Honolulu, USA (p. 429-433). doi: 10.1109/ICCNC.2014.6785373

Hu, X., Gao, S., Chiu, Y.-C., & Lin, D.-Y. (2012). Modeling routing behavior for vacant taxicabs in urban traffic networks. *Transportation Research Record: Journal of the Transportation Research Board*(2284), 81-88. doi: 10.3141/2284-10

Hu, Y., Yang, Y., & Huang, B. (2015). *A comprehensive survey of recommendation system based on taxi GPS trajectory.* Communication présentée à 2015 International Conference on Service Science (ICSS) (p. 99-105). doi: 10.1109/ICSS.2015.31

Huang, X., Zhao, Y., Yang, J., Zhang, C., Ma, C., & Ye, X. (2016). TrajGraph: A Graph-Based Visual Analytics Approach to Studying Urban Network Centralities Using Taxi Trajectory Data. *Visualization and Computer Graphics, IEEE Transactions on, 22*(1), 160-169. doi: 10.1109/TVCG.2015.2467771

Jianxin, Y., Xiaomin, Z., & Hongyu, Z. (2009). *Design and implementation of taxi calling and dispatching system based on GPS mobile phone.* Communication présentée à 4th International Conference on Computer Science&Education, Nanning, China (p. 1163-1169). doi: 10.1109/ICCSE.2009.5228489

Jun, L. (2013). The Multidimensional Displaying and Analysising of Taxi meter Data Based on BI. *ICTIS 2013* pp.348-353. doi: 10.1061/9780784413036.048

Kamga, C., Yazici, M. A., & Singhal, A. (2013). *Hailing in the rain: Temporal and weather-related variations in taxi ridership and taxi demand-supply equilibrium.* Communication présentée à Transportation Research Board 92nd Annual Meeting.

Kattan, L., de Barros, A., & Wirasinghe, S. C. (2010). Analysis of work trips made by taxi in canadian cities. *Journal of Advanced Transportation, 44*(1), 11-18. doi: 10.1002/atr.102

Kennedy, P. (2008). *A guide to econometrics* (6e éd.). Malden, USA: Blackwell Publishing Ltd.

King, D. A., Peters, J. R., & Daus, M. W. (2012). *Taxicabs for Improved Urban Mobility: Are We Missing an Opportunity?* Communication présentée à Transportation Research Board 91st Annual Meeting.

Li, S., Enam, A., Abou-Zeid, M., & Ben-Akiva, M. E. (2013). *Travel Time Modeling with Taxi GPS and Household Survey Data.* Communication présentée à Transportation Research Board 92nd Annual Meeting, Washington D.C.

Liu, D., Cheng, S.-F., & Yang, Y. (2015). *Density Peaks Clustering Approach for Discovering Demand Hot Spots in City-scale Taxi Fleet Dataset.* Communication présentée à Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on (p. 1831-1836). doi: 10.1109/ITSC.2015.297

Liu, X., Gong, L., Gong, Y., & Liu, Y. (2015). Revealing travel patterns and city structure with taxi trip data. *Journal of Transport Geography, 43* 78-90. doi: 10.1016/j.jtrangeo.2015.01.016

Miao, F., Lin, S., Munir, S., Stankovic, J. A., Huang, H., Zhang, D., . . . Pappas, G. J. (2015). *Taxi dispatch with real-time sensing data in metropolitan areas: a receding horizon control approach.* Communication présentée à Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems (p. 100-109). doi: 10.1145/2735960.2735961

Moreira-Matias, L., Gama, J., Ferreira, M., Mendes-Moreira, J., & Damas, L. (2013). Predicting Taxi–Passenger Demand Using Streaming Data. *IEEE Transactions on Intelligent Transportation Systems, 14*(3), 1393-1402. doi: 10.1109/TITS.2013.2262376

Mousavi, A., Bunker, J. M., & Lee, B. (2012). *A new approach for trip generation estimation for use in traffic impact assessments.* Communication présentée à 25th ARRB Conference Proceedings, Perth, Australia. Tiré de http://eprints.qut.edu.au/53697/

New York City Taxi & Limousine Commission. (2014). *2014 Taxicab Fact Book*. Tiré de http://www.nyc.gov/html/tlc/downloads/pdf/2014_taxicab_fact_book.pdf

Office Québécois de la Langue Française. (Édit.). (2010). *Le Grand Dictionnaire Terminologique*.

Organisation for Economic Co-operation and Development. (2007). *Taxi Services: Competition and Regulation*. Organisation for Economic Co-operation and Development.

Pelé, N., & Morency, C. (2014). *When, Where and How Taxis Are Used in Montreal.* Communication présentée à Transportation Research Board 93rd Annual Meeting.

Phithakkitnukoon, S., Veloso, M., Bento, C., Biderman, A., & Ratti, C. (2010). Taxi-aware map: Identifying and predicting vacant taxis in the city. Dans *Ambient Intelligence* (p. 86-95): Springer.

Qian, X., & Ukkusuri, S. V. (2015). Spatial variation of the urban taxi ridership using GPS data. *Applied Geography, 59* 31-42. doi: 10.1016/j.apgeog.2015.02.011

Racca, D. P., & Ratledge, E. (2003). *Factors that affect and/or can alter mode choice*. Delaware Center for Transportation. Tiré de http://sites.udel.edu/dct/files/2013/10/Rpt.-159-Factors-that-Affect-and-or-Can-Alter-Mode-Choice-yzxqre.pdf

Savage, T. H., & Vo, H. T. (2013). *Yellow cabs as red corpuscles.* Communication présentée à Big Data, 2013 IEEE International Conference on (p. 22-28). doi: 10.1109/BigData.2013.6691773

Schaller, B. (2005). A regression model of the number of taxicabs in US cities. *Journal of Public Transportation, 8*(5), 63-78. doi: 10.5038/2375-0901.8.5.4

Schönfelder, S., Axhausen, K. W., Antille, N., & Bierlaire, M. (2002). *Exploring the Potentials of Automatically Collected GPS Data for Travel Behaviour Analysis: A Swedish Data Source*: Citeseer.

Shin, I.-H., Park, G.-L., Saha, A., Kwak, H.-y., & Kim, H. (2009). Analysis of Moving Patterns of Moving Objects with the Proposed Framework. Dans *Computational Science and Its Applications–ICCSA 2009* (p. 443-452): Springer.

Société de transport de Montréal. (2015). Développeurs | Société de transport de Montréal. Tiré de http://www.stm.info/fr/a-propos/developpeurs

Société de transport de Montréal. (2016). Métro | Société de transport de Montréal. Tiré de http://www.stm.info/fr/infos/reseaux/metro

Statistique Canada. (2016). Profil du recensement. Tiré de http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/index.cfm?Lang=F

Tang, H., Kerber, M., Huang, Q., & Guibas, L. (2013). *Locating lucrative passengers for taxicab drivers.* Communication présentée à 21st ACM SIGSPATIAL International Conference on

Advances in Geographic Information Systems, Orlando, USA (p. 504-507). doi: 10.1145/2525314.2525471

Taxi de France. (2016). Historique du Taxi. Tiré de http://www.taxis-de-france.com/historique/

Taxi Study Panel. (1999). *A Study of the Taxi Industry in British Columbia*. Tiré de http://www2.gov.bc.ca/assets/gov/driving-and-transportation/reports-and-reference/reports-and-studies/planning-strategy-economy/taxi_industry_bc_study.pdf

Tennessee Transportation & Logistics Foundation. (2009). *Winnipeg Taxi Study*. Tiré de http://www.gov.mb.ca/ia/taxicab/pdf/surveyvol1.pdf

Tennessee Transportation & Logistics Foundation. (2010a). *Regina Taxi Study*. Tiré de http://www.regina.ca/opencms/export/sites/regina.ca/residents/licences/.media/pdf/regina_taxi_industry_study_-_full_report.pdf

Tennessee Transportation & Logistics Foundation. (2010b). *Saskatoon Taxi Study*. Tiré de http://www.taxi-library.org/saskatoon_taxi_study_2010.pdf

Tennessee Transportation & Logistics Foundation. (2014). *Houston Taxi Study*. Tiré de https://www.houstontx.gov/ara/regaffairs/Houston_Taxi_Study_Results_20140120.pdf

Ville de Montréal.). Ville de Montréal - Info Travaux - Chaussées et trottoirs. Tiré de http://ville.montreal.qc.ca/portal/page?_dad=portal&_pageid=7097,71577618&_schema=PORTAL

Ville de Montréal. (2008). *Réinventer Montréal - Plan de transport 2008*. Tiré de http://servicesenligne.ville.montreal.qc.ca/sel/publications/PorteAccesTelechargement?lng=Fr&systemName=68235660&client=Serv_corp

Wan, X., Gao, M., Kang, J., & Zhao, J. (2013). *Taxi origin-destination areas of interest discovering based on functional region division.* Communication présentée à Third International Conference on Innovative Computing Technology (INTECH), London, England (p. 365-370). doi: 10.1109/INTECH.2013.6653677

Wang, Z., Yuan, X., Ye, T., Hao, Y., Chen, S., Liangk, J., . . . Wu, Y. (2015). Visual Data Quality Analysis for Taxi GPS Data. *2015 IEEE Conference on Visual Analytics Science and Technology (VAST)* 223 - 224. doi: 10.1109/VAST.2015.7347689

Wolf, J., Bachman, W., Oliveira, M. S., Auld, J., Mohammadian, A. K., & Vovsha, P. (2014). *Applying GPS Data to Understand Travel Behavior, Volume I: Background, Methods, and Tests* (vol. 1).

Wolf, J., Bachman, W., Oliveira, M. S., Auld, J., Mohammadian, A. K., Vovsha, P., & Zmud, J. (2014). *Applying GPS Data to Understand Travel Behavior, Volume II: Guidelines* (vol. 2).

Wong, R. C. P., Szeto, W. Y., & Wong, S. C. (2014). A cell-based logit-opportunity taxi customer-search model. *Transportation Research Part C: Emerging Technologies, 48* 84-96. doi: 10.1016/j.trc.2014.08.010

Xu, X., Zhou, J., Liu, Y., Xu, Z., & Zha, X. (2015). Taxi-RS: Taxi-Hunting Recommendation System Based on Taxi GPS Data. *IEEE Transactions on Intelligent Transportation Systems, PP*(99), 1-12. doi: 10.1109/TITS.2014.2371815

Yang, C., & Gonzales, E. J. (2014). Modeling Taxi Trip Demand by Time of Day in New York City. *Transportation Research Record: Journal of the Transportation Research Board, 2429*(1), 110-120. doi: 10.3141/2429-12

Yang, H., Wong, S. C., & Wong, K. I. (2002). Demand–supply equilibrium of taxi services in a network under competition and regulation. *Transportation Research Part B: Methodological, 36*(9), 799-819. doi: 10.1016/S0191-2615(01)00031-5

Yang, T., Yang, H., Wong, S. C., & Sze, N. N. (2014). Returns to scale in the production of taxi services: an empirical analysis. *Transportmetrica A: Transport Science, 10*(9), 775-790. doi: 10.1080/23249935.2013.794174

Yu, L.-J., & Peng, Z.-R. (2013). *A Better Understanding of Taxi Emissions in Shenzhen, China, Based on Floating-Car Data.* Communication présentée à Transportation Research Board 92nd Annual Meeting.

Yuan, J., Zheng, Y., Zhang, L., Xie, X., & Sun, G. (2011). *Where to find my next passenger.* Communication présentée à 13th international conference on Ubiquitous computing, Beijing, China (p. 109-118). doi: 10.1145/2030112.2030128

Yuan, N. J., Zheng, Y., Zhang, L., & Xie, X. (2013). T-Finder: A recommender system for finding passengers and vacant taxis. *IEEE Transactions on Knowledge and Data Engineering, 25*(10), 2390-2403. doi: 10.1109/TKDE.2012.153

Zhang, D., & Song, L. (2009). *Travel mode choice behavior of taxi passengers*. Communication présentée à Advanced Forum on Transportation of China (AFTC), Beijing, China.

Zhang, D.-Z., Peng, Z.-R., & Sun, D. J. (2014). *A Comprehensive Taxi Assessment Index Using Floating Car Data.* Communication présentée à Transportation Research Board 93rd Annual Meeting.

Zhang, L., Ahmadi, M., Pan, J., & Chang, L. (2012). *Metropolitan-scale taxicab mobility modeling.* Communication présentée à Global Communications Conference (GLOBECOM), 2012 IEEE (p. 5404-5409). doi: 10.1109/GLOCOM.2012.6503980

Zhang, W., & Ukkusuri, S. V. (2016). Optimal fleet size and fare setting in emerging taxi markets with stochastic demand. *Computer-Aided Civil and Infrastructure Engineering* 647-660. doi: 10.1111/mice.12203

Zhang, Y. (2014). *How Do Taxis Work in Beijing? An Exploratory Study of Spatio-Temporal Taxi Travel Pattern Using GPS Data.* (University of California).

Zheng, Y., Li, Q., Chen, Y., Xie, X., & Ma, W.-Y. (2008). *Understanding mobility based on GPS data.* Communication présentée à Proceedings of the 10th international conference on Ubiquitous computing (p. 312-321).

Zhu, J., Shuai, B., Huang, Z., & Sun, C. (2013). The Optimal Taxi Fleet Size Structure under Various Market Regimes When Charging Taxis with Link-Based Toll. *Journal of Applied Mathematics, 2013*. doi: 10.1155/2013/535878

ANNEXES

ANNEXE A – TABLES D'INFORMATION DE TAXI DIAMOND

Tableau A-1 : Table des points GPS

Champs	**Définition**	**Format**
GPS_id	Identifiant du point GPS	Entier
Horodatage	Date et heure (aaaa-mm-jj hh:mm:ss.sss)	Date-Heure
Lat	Latitude du point GPS	Nombre réel
Long	Longitude du point GPS	Nombre réel
Vehicule_id	Identifiant du véhicule	Entier
Vitesse	Vitesse instantanée du véhicule (km/h)	Entier
Azimut	Orientation par rapport au nord géographique (degré)	Entier
Chauffeur	Identifiant du chauffeur	Entier
Statut*	État du chauffeur : 0-Connecté, pas prêt à recevoir une course 1-Connecté, prêt à recevoir une course 2-En service, va chercher ou est avec des clients 3-En discipline (bloqué par l'administration) 4-À l'extérieur de son véhicule pour un maximum de 10 minutes 5-Pas connecté	Entier
Course_id	Identifiant de la course	Entier

*Appellation « Statut TD » employée à la figure 4-5

Tableau A-2 : Table des courses

Champs	**Définition**	**Format**
Course_id	Identifiant de la course	Entier
Hélage*	Taxi hélé sur la rue	
Acceptation*	Course acceptée par le chauffeur	
Rejet*	Course refusée par le chauffeur	
Annulation*	Course annulée par le client	
Vide*	Course transférée à un autre chauffeur (champ incertain)	
Absent*	Client ne s'est pas présenté au point d'origine	
Au site*	Chauffeur arrivé au point d'origine du client	
Taximètre démarré*	Taximètre démarré par le chauffeur	
Taximètre arrêté*	Taximètre arrêté par le chauffeur	
Paiement*	Paiement complété	

*Ces champs sont composés d'un Horodatage (Date-Heure), d'une Latitude (Nombre réel) et d'une Longitude (Nombre réel).

ANNEXE B – TABLES D'INFORMATION DE LA BASE DE DONNÉES

Tableau B-1 : Table des intermédiaires

Champs	Définition	Format
Intermediaire_id	Identifiant de l'intermédiaire	Entier
Nom	Nom de l'intermédiaire	Texte

Tableau B-2 : Table des chauffeurs

Champs	Définition	Format
Chauffeur_id	Identifiant du chauffeur	Entier
Chauffeur_id_interne	Identifiant du chauffeur attribué par l'intermédiaire	Entier
Intermediaire_id	Identifiant de l'intermédiaire	Entier

Tableau B-3 : Table des véhicules

Champs	Définition	Format
Vehicule_id	Identifiant du véhicule	Entier
Vehicule_id_interne	Identifiant du véhicule attribué par l'intermédiaire	Entier
Intermediaire_id	Identifiant de l'intermédiaire	Entier
Agglo_id	Identifiant de l'agglomération associé au permis du véhicule	Entier

Tableau B-4 : Table des points GPS

Champs	**Définition**	**Format**
Gps_id	Identifiant du point GPS	Entier
GPS_id_interne	Identifiant du point GPS attribué par l'intermédiaire	Entier
Horodatage	Date et heure (aaaa-mm-jj hh:mm:ss.sss)	Date-Heure
Lat	Latitude du point GPS	Nombre réel
Long	Longitude du point GPS	Nombre réel
Position	Géométrie du point GPS selon EPSG : 32188	Géométrie
Secteur_id	Identifiant du secteur de recensement dans lequel se trouve le point GPS	Entier
Intermediaire_id	Identifiant de l'intermédiaire	Entier
Vehicule_id	Identifiant du véhicule	Entier
Vitesse	Vitesse instantanée du véhicule (km/h)	Entier
Azimut	Orientation par rapport au nord géographique (degré)	Entier
Chauffeur_id	Identifiant du chauffeur	Entier
Statut	État du chauffeur : 1-Hors service 2-En service, pas prêt à recevoir des clients 3-En service, prêt à recevoir des clients 4-En service, va chercher ou est avec des clients	Entier
Course_id	Identifiant de la course	Entier
Distance	Distance entre le point et le point précédent dans le temps du même véhicule (mètres)	Entier
Duree	Durée entre le point et le point précédent dans le temps du même véhicule (hh:mm:ss.sss)	Intervalle de temps

Tableau B-5 : Table des courses

Champs	**Définition**	**Format**
Course_id	Identifiant de la course	Entier
Course_id_interne	Identifiant de la course attribué par l'intermédiaire	Entier
Intermediaire_id	Identifiant de l'intermédiaire	Entier
Accepte_horodatage	Moment où le chauffeur accepte la course (aaaa-mm-jj hh:mm:ss.sss)	Date-Heure
Annulation	Annulation de la course : 0 -Course complétée 1-Course annulée	Entier
Orig_horodatage	Moment du début de la course (aaaa-mm-jj hh:mm:ss.sss)	Date-Heure
Orig_lat	Latitude de l'origine	Nombre réel
Orig_long	Longitude de l'origine	Nombre réel
Orig_position	Géométrie du point d'origine selon EPSG : 32188	Géométrie
Orig_secteur_id	Identifiant du secteur de recensement dans lequel se trouve l'origine	Entier
Dest_horodatage	Moment de la fin de la course (aaaa-mm-jj hh:mm:ss.sss)	Date-Heure
Dest_lat	Latitude de la destination	Nombre réel
Dest_long	Longitude de la destination	Nombre réel
Dest_position	Géométrie du point de destination selon EPSG : 32188	Géométrie
Dest_secteur_id	Identifiant du secteur de recensement dans lequel se trouve la destination	Entier
Type	Type de commande : 0 -Course hélée 1-Course commandée	Entier
Vehicule_id	Identifiant du véhicule	Entier
Chauffeur_id	Identifiant du chauffeur	Entier
Distance	Distance de la course (mètres)	Entier
Duree	Durée de la course (hh:mm:ss)	Intervalle de temps

Tableau B-6 : Table des secteurs

Champs	Définition	Format
Secteur_id	Identifiant du secteur	Entier
Secteur_id_interne	Identifiant du secteur attribué par Statistique Canada	Entier
Nom	Identifiant de la région métropolitaine suivi du secteur attribué par Statistique Canada	Texte
Superficie	Superficie (m^2)	Entier
Geometrie	Géométrie du polygone du secteur selon EPSG : 32188	Géométrie
Arron_id	Identifiant de l'arrondissement dans lequel se trouve le secteur	Entier
Agglo_id	Identifiant de l'agglomération dans lequel se trouve le secteur	Entier

Tableau B-7 : Table des arrondissements

Champs	Définition	Format
Arron_id	Identifiant de l'arrondissement	Entier
Nom	Nom selon la Ville de Montréal	Texte
Superficie	Superficie (m^2)	Entier
Geometrie	Géométrie du polygone de l'arrondissement selon EPSG : 32188	Géométrie
Agglo_id	Identifiant de l'agglomération dans lequel se trouve l'arrondissement	Entier

Tableau B-8 : Table des agglomérations

Champs	Définition	Format
Agglo_id	Identifiant de l'agglomération	Entier
Superficie	Superficie (m^2)	Entier
Geometrie	Géométrie du polygone de l'agglomération selon EPSG : 32188	Géométrie

Tableau B-9 : Table de l'utilisation du sol

Champs	**Définition**	**Format**
Sol_id	Identifiant de la zone	Entier
Type_utilisation	Type d'utilisation du sol : 100 -Résidentiel 200-Commercial 300-Bureau 400-Industrie 500-Institutionnel 600-Parc ou espace vert 700-Utilité publique 800-Agricole 900-Terrain vacant 1000-Hydrographie 1100-Golf	Entier
Zone_agricole	La zone se situe : 0-À l'extérieur d'une zone agricole 1-À l'intérieur d'une zone agricole	Entier
Etage_moy	Nombre moyen d'étages	Entier
Logement_moy	Nombre moyen de logements	Entier
Annee_moy	Année moyenne de construction	Entier
Valeur_terrain_moy	Valeur moyenne du terrain	Entier
Valeur_metre_moy	Valeur moyenne du terrain au mètre2	Entier
Valeur_batiment_moy	Valeur moyenne des bâtiments	Entier
Superficie	Superficie (m^2)	Entier
Geometrie	Géométrie du polygone de la zone selon EPSG : 32188	Géométrie
Secteur_id	Identifiant du secteur de recensement dans lequel se trouve la zone	Entier

Tableau B-10 : Table des postes d'attente de taxis

Champs	**Définition**	**Format**
Poste_id	Identifiant du poste d'attente	Entier
Poste_id_interne	Identifiant du poste attribué par le BTM	Entier
Lat	Latitude du poste	Nombre réel
Long	Longitude du poste	Nombre réel
Position	Géométrie du point du poste selon EPSG : 32188	Géométrie
Localisation	Nom attribué à l'endroit où le poste est situé (numéro de rue, rues de l'intersection, place, etc.)	Texte
Statut	Statut d'activité du poste : 1 -Actif 2-Fermé 3-Temporaire	Entier
Type	Type d'utilisation du poste : 1 -Commun 2-Privé 3-Public	Entier
Place	Nombre de places de stationnement	Entier
Secteur_id	Identifiant du secteur de recensement dans lequel se trouve le poste	Entier

Tableau B-11 : Table du calendrier

Champs	Définition	Format
Date	Date dans le calendrier (aaaa-mm-jj)	Date
Jour_semaine	Numéro du jour de la semaine : 1-Lundi 5-Vendredi 2-Mardi 6-Samedi 3-Mercredi 7-Dimanche 4-Jeudi	Entier
Fete	Jour de fête : 0 -Jour habituel 1-Jour de fête	Entier
Jour_ferie	Jour férié : 0 -Jour habituel 1-Jour férié (congé)	Entier
Nom_fete	Nom de la fête du jour	Texte

Tableau B-123 : Table de la météo par jour

Champs	Définition	Format
Date	Année, mois et jour constituant la date	Date
Max_temp	Température maximale (°C)	Nombre réel
Min_temp	Température minimale (°C)	Nombre réel
Moy_temp	Température moyenne (°C)	Nombre réel
Degre_jour_chaleur	Nombre de degrés entre la température moyenne et 18°C (pour une température sous 18°C)	Nombre réel
Degre_jour_froid	Nombre de degrés entre la température moyenne et 18°C (pour une température au-dessus de 18°C)	Nombre réel
Pluie_tot_mm	Précipitations de pluie (mm)	Nombre réel
Neige_tot_cm	Précipitations de neige (cm)	Nombre réel
Precip_tot_mm	Précipitations combinées de pluie et de neige (mm)	Nombre réel
Neige_sol_cm	Accumulation de neige au sol (cm)	Nombre réel

Tableau B-13 : Table de la météo aux 15 minutes

Champs	**Définition**	**Format**
Horodatage	Date et heure (aaaa-mm-jj hh:mm:ss.sss)	Date-Heure
Condition	Description des conditions météorologiques	Texte
Temperature	Température (°C)	Nombre réel
Point_rose	Température du point de rosée (°C)	Nombre réel
Humidite	Humidité relative (%)	Entier
Vitesse_vent	Vitesse du vent (km/h)	Entier
Visibilite	Distance à laquelle des objets de taille convenables peuvent être identifiés (km)	Entier
Pression	Pression atmosphérique (kPa)	Nombre réel
Tendance	Description des conditions météorologiques dans les prochaines heures	Texte