

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Caractérisation des typologies d'utilisateurs d'un service de taxis électriques

VINCENT CHABIN

Département des génies civil, géologique et des mines

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie Civil

Juillet 2019

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé:

Caractérisation des typologies d'utilisateurs d'un service de taxis électriques

Présenté par **Vincent CHABIN**

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

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

Martin TRÉPANIÉ, président

Nicolas SAUNIER, membre et directeur de recherche

Catherine MORENCY, membre et codirectrice de recherche

Simon BOIVIN, membre externe

REMERCIEMENTS

Je souhaite remercier Nicolas Saunier et Catherine Morency qui ont su me conseiller et me guider tout au long de ce projet. Le temps précieux qu'ils ont réussi à m'accorder durant cette recherche m'a été d'une grande aide. Leurs expertises respectives et leur passion pour le domaine des transports ont su profiter à ce projet.

Ma maîtrise n'aurait pas été la même sans toute l'équipe de recherche de la Chaire Mobilité. Je tiens particulièrement à remercier Jean-Simon Bourdeau, Pierre-Léo Bourbonnais et Hubert Verreault qui ont toujours répondu à mes questions avec beaucoup d'intérêt et qui m'ont permis de développer de nombreuses compétences.

Je souhaite également remercier chaleureusement tous les étudiants des différents bureaux de transport. L'atmosphère d'entraide qui y règne permet à chacun de profiter des connaissances de tous et fait évoluer nos projets respectifs. Il ne faut pas oublier de mentionner l'ambiance chaleureuse du bureau ainsi que les agréables moments de détente entre étudiants.

Il est également important de mentionner tous les partenaires sans qui cette recherche n'aurait pas été possible. Je souhaite donc remercier InnovÉÉ, le CRSNG, Fraxion et Taxelco d'avoir rendu ce projet réalisable. Merci à Taxelco / Téo Taxi pour leurs contributions et l'accès aux données pour fins de recherche.

Finalement, je tiens à remercier l'École Spéciale des Travaux Publics de m'avoir donné l'opportunité d'étudier à Polytechnique Montréal.

RÉSUMÉ

Les introductions répétées de nouveaux services de mobilité à la demande à partir de la fin des années 2000 ont forcé l'industrie du taxi à se réinventer. Active de 2015 à 2019, l'entreprise Téo Taxi a opéré à Montréal une flotte de taxis électriques. Une des particularités de Téo Taxi était d'offrir à ses clients la possibilité de commander un taxi par l'intermédiaire d'une application mobile. Plus que de modifier l'expérience client, cette technologie permet de caractériser le comportement des utilisateurs de ce service de taxis. Si ce type d'analyse sur les usagers est depuis longtemps fait pour d'autres modes de transport comme le transport collectif avec l'utilisation des données de carte à puce, il était jusqu'à présent impossible de le faire pour le taxi.

Le but principal de ce projet est de dresser un portrait d'utilisation des clients de Téo Taxi. Pour cela, plusieurs sous-objectifs ont été définis :

- Définir des indicateurs pertinents d'utilisation du taxi à l'échelle de l'utilisateur;
- Identifier des comportements d'utilisation sur une échelle de temps à la fois relativement courte pour observer de possibles changements d'utilisation dans le temps, mais également plus longue prenant en compte les possibles effets de saisonnalité dans les analyses;
- Explorer les facteurs favorisant l'utilisation du taxi du point de vue de l'utilisateur.

Une courte revue de littérature permet de comprendre quels sont les enjeux actuels auxquels le taxi doit faire face. Cette revue présente tout d'abord une définition d'un système de transport par taxi et de son évolution au fil du temps. L'intégration du taxi dans les systèmes de transport urbain est également présentée dans cette revue. Enfin cette revue aborde également les questions relatives à l'étude des utilisateurs du taxi notamment par le biais d'enquêtes réalisées auprès des utilisateurs de ce mode.

Comme dans tout projet de fouille de données, les hypothèses relatives au nettoyage et au filtrage des données peuvent avoir un impact important. Ces hypothèses ainsi que des éléments permettant de justifier les choix méthodologiques sont présentés. Les données nettoyées et filtrées sont ensuite étudiées. Avant de caractériser les utilisateurs de Téo Taxi, une analyse descriptive de l'utilisation de Téo Taxi et du jeu de données à notre disposition est faite.

Ensuite, la méthodologie développée pour caractériser l'utilisation de Téo Taxi par un client est introduite. Cette méthodologie permet de caractériser le comportement d'utilisation des clients de

Téo Taxi à la fois sur une courte période de temps (quatre semaines) et sur une période plus longue (durée de vie de l'utilisateur ou période d'activité de Téo Taxi). Cette méthodologie nous a permis notamment de mettre en lumière cinq groupes d'utilisations sur une période de quatre semaines. Ces cinq groupes vont de l'usage exceptionnel (une course sur la période) à des usages beaucoup plus intensifs (dix courses et plus sur une période). La répartition des courses dans le temps est également différente pour chaque groupe d'utilisation. L'étude des comportements des utilisateurs sur une période plus longue a permis de catégoriser certains usagers spécifiques de Téo Taxi. Cette étude à long terme a également permis de montrer une stabilité temporelle relative dans l'utilisation de Téo Taxi.

Enfin, pour approfondir nos connaissances sur les utilisateurs de Téo Taxi, un sondage a été réalisé auprès de ces derniers durant le mois de décembre 2018. Près de 1 200 utilisateurs ont complété entièrement l'enquête. Ce sondage a permis d'approfondir nos connaissances sur l'utilisation d'une part et sur la caractérisation des usagers (habitudes de mobilité, profil socio-économique) d'autre part. Le questionnaire ainsi que les résultats du sondage sont détaillés. Des analyses croisées entre les réponses à l'enquête et les caractéristiques d'utilisation de Téo Taxi ont également été réalisées. Cela nous a permis de mieux comprendre l'influence du comportement long terme et court terme des utilisateurs ainsi que celle de l'heure de la course sur les réponses à cette enquête. L'influence est par exemple particulièrement marquée sur le motif du déplacement réalisé en taxi.

ABSTRACT

Repeated introductions of new mobility services such as Transportation network companies since the late 2000s forced the taxi industry to reinvent itself. Active from 2015 to 2019, Téo Taxi operated a fleet of electric taxis in Montreal. One particularity of Téo Taxi is that a trip can be ordered by the customers via a mobile app. More than improving the customer experience, this technology is an opportunity to explore and characterize Téo Taxi users' behaviour. While this type of user analysis has long been done for other modes of transportation such as public transportation with the use of smart card data, it has not been possible for the taxi due to the lack of data.

The main goal of this project is to characterize Téo Taxi's customers' use of the service. For this purpose, several sub-objectives have been defined:

- Define relevant indicators of taxi use at the user level;
- Identify usage behaviours both on a relatively short time scale to observe possible changes in use over time and for a longer time scale considering possible seasonality effects in the analyzes;
- Explore the factors that favour the use of the taxi from the point of view of the user.

First, a short literature review helps us better understand the current issues with the taxi industry. This review starts with a small definition of taxi transportation system and summarizes the historical development of this industry. The role of such systems within an urban transportation system is also explored in this review. Finally, this review focuses on questions concerning the taxi user especially surveys conducted among taxi customers.

As in any data mining project, choices about data cleaning and filtering can have a significant impact on the results. Those choices as well as elements justifying the methodological choices are presented. Before characterising Téo Taxi users, a descriptive analysis of the use of Téo Taxi in its entirety is proposed.

Then the methodology developed to characterize Téo Taxi usage for each client is introduced. This methodology allows to characterize the usage behaviour over both a short period (four weeks) and longer periods (user's lifetime or Téo Taxi's period of activity). The application of this methodology on the dataset highlights five types of usage among customers over a four week-

period. These five typologies range from exceptional use (one trip on average over the period) to much more intensive use (ten trips and more on average over a period). The temporal distributions of the trips also differ for each type of usage. The study of users' behaviors over a longer period also made it possible to categorize certain specific long-time users of Téo Taxi. This long-term study also showed relative temporal stability in the general use of Téo Taxi.

Finally, to deepen our knowledge of Téo Taxi users, a survey was conducted during the December 2018 month. Nearly 1,200 users completed the survey. This survey allowed to learn more on their usage of Téo Taxi but also to gather general information about them (general mobility habits or socio-economic profile). The online questionnaire and the results are detailed. Cross-analyzes between survey responses and Téo Taxi usage characteristics were also conducted. This allowed us to better understand the influence of users long-term and short-term behavior as well as the time of the trip on the responses to this survey. The influence is for example particularly marked on the purpose of the trip made by taxi.

TABLE DES MATIÈRES

| | |
|---|----------|
| REMERCIEMENTS | III |
| RÉSUMÉ..... | IV |
| ABSTRACT | VI |
| TABLE DES MATIÈRES | VIII |
| LISTE DES TABLEAUX..... | XII |
| LISTE DES FIGURES..... | XIII |
| LISTE DES SIGLES ET ABRÉVIATIONS | XIX |
| LISTE DES ANNEXES..... | XX |
| CHAPITRE 1 INTRODUCTION..... | 1 |
| 1.1 Problématique..... | 1 |
| 1.2 Objectifs | 2 |
| 1.3 Structure du mémoire | 3 |
| CHAPITRE 2 REVUE DE LITTÉRATURE | 5 |
| 2.1 Brève histoire du taxi et de son cadre réglementaire | 5 |
| 2.1.1 Définition et développement historique du taxi | 5 |
| 2.1.2 Le cadre réglementaire | 7 |
| 2.2 Le rôle du taxi dans un système de transport urbain | 8 |
| 2.2.1 Le taxi : un mode de substitution | 9 |
| 2.2.2 Le taxi : un système complémentaire au transport public | 9 |
| 2.2.3 L'industrie du taxi comme facteur d'intégration sociale | 11 |
| 2.3 Un secteur en pleine mutation..... | 11 |
| 2.3.1 Impact de ces nouveaux acteurs sur l'équilibre du marché..... | 12 |
| 2.3.2 Les réactions de l'industrie | 13 |

| | | |
|---|---|-----------|
| 2.3.3 | Le futur de l'industrie..... | 14 |
| 2.4 | Les utilisateurs..... | 15 |
| 2.4.1 | Catégories d'utilisateurs..... | 15 |
| 2.4.2 | Les enquêtes..... | 16 |
| 2.5 | Classification comportementale..... | 20 |
| CHAPITRE 3 MONTRÉAL ET LE FONCTIONNEMENT PARTICULIER DE TÉO TAXI | | 22 |
| 3.1 | La mobilité montréalaise..... | 22 |
| 3.1.1 | Le territoire..... | 22 |
| 3.1.2 | Le réseau de transport..... | 23 |
| 3.1.3 | La mobilité sur le territoire..... | 25 |
| 3.2 | Le taxi à Montréal..... | 27 |
| 3.3 | L'expérimentation de Téo Taxi..... | 29 |
| 3.3.1 | Présentation de l'entreprise..... | 29 |
| 3.3.2 | Bilan du projet Téo Taxi..... | 31 |
| CHAPITRE 4 CADRE MÉTHODOLOGIQUE DU PROJET ET DESCRIPTION DES DONNÉES | | 33 |
| 4.1 | Les grandes étapes méthodologiques..... | 33 |
| 4.1.1 | Construction de typologies d'utilisation..... | 35 |
| 4.1.2 | Compréhension des facteurs influençant l'utilisation du taxi..... | 36 |
| 4.2 | Description des données utilisées..... | 36 |
| 4.2.1 | Les données provenant de Téo Taxi..... | 36 |
| 4.2.2 | Autres sources de données..... | 37 |
| 4.3 | Filtrage préliminaire des données..... | 37 |
| 4.3.1 | Définition des critères de filtrage..... | 38 |
| 4.3.2 | Durée de la course..... | 39 |

| | | |
|---|--|-----------|
| 4.3.3 | Distance de la course..... | 40 |
| 4.3.4 | Vitesse de la course..... | 41 |
| 4.3.5 | Temps intercourse..... | 42 |
| 4.3.6 | Synthèse du processus de filtrage..... | 43 |
| 4.4 | Analyses descriptives de la base de données..... | 44 |
| 4.4.1 | Répartition temporelle des courses..... | 45 |
| 4.4.2 | Répartition spatiale des courses..... | 47 |
| 4.4.3 | Nombre de courses par utilisateur..... | 49 |
| 4.4.4 | Distribution des grandeurs caractéristiques des courses..... | 50 |
| CHAPITRE 5 CARACTÉRISATION DE L'UTILISATION DE TÉO TAXI..... | | 54 |
| 5.1 | Enjeux méthodologiques..... | 54 |
| 5.2 | Méthodologie d'analyse comportementale sur une période..... | 55 |
| 5.2.1 | Choix de la durée de la période..... | 55 |
| 5.2.2 | Création des usagers-périodes..... | 58 |
| 5.2.3 | Classification des usagers-périodes..... | 62 |
| 5.3 | Analyse des résultats de la classification des usagers-périodes..... | 69 |
| 5.3.1 | Description des centroïdes des groupes..... | 71 |
| 5.3.2 | Analyse de l'usage des groupes sur une période..... | 75 |
| 5.4 | Analyses de l'utilisation de Téo Taxi sur une période de longue durée..... | 81 |
| 5.4.1 | Évolution de l'utilisation globale de Téo Taxi..... | 81 |
| 5.4.2 | Analyses des super-utilisateurs..... | 83 |
| 5.5 | Synthèse de l'utilisation de Téo Taxi..... | 89 |
| CHAPITRE 6 ENQUÊTE AUPRÈS DES UTILISATEURS..... | | 91 |
| 6.1 | Modalités méthodologiques..... | 91 |

| | | |
|---|--|------------|
| 6.1.1 | Prise de contact avec le répondant | 91 |
| 6.1.2 | Présentation de la plateforme utilisée..... | 92 |
| 6.1.3 | Politique de confidentialité..... | 93 |
| 6.2 | Construction du questionnaire..... | 93 |
| 6.3 | Résultats | 95 |
| 6.3.1 | Nombre de réponses | 95 |
| 6.3.2 | Représentativité de la population | 96 |
| 6.3.3 | Analyses croisées des résultats..... | 109 |
| 6.4 | Synthèse de l'enquête mené auprès des utilisateurs | 128 |
| CHAPITRE 7 CONCLUSION ET RECOMMANDATIONS | | 129 |
| 7.1 | Synthèse de la recherche | 129 |
| 7.2 | Contributions..... | 130 |
| 7.3 | Limitations | 131 |
| 7.4 | Perspectives..... | 132 |
| BIBLIOGRAPHIE | | 133 |
| ANNEXES | | 139 |

LISTE DES TABLEAUX

| | |
|--|----|
| Tableau 2-1 Tableau comparatif des enquêtes menées auprès des utilisateurs du taxi et de service de mobilité à la demande..... | 17 |
| Tableau 4-1 Tableau comparatif des filtres appliqués aux données brutes..... | 43 |
| Tableau 4-2 Caractéristiques principales du jeu de données..... | 44 |
| Tableau 4-3 Indicateurs statistiques de la distribution du nombre de courses par usager..... | 50 |
| Tableau 4-4 Indicateurs statistiques des distributions des grandeurs caractéristiques des courses après l'application des filtres..... | 53 |
| Tableau 5-1 Exemple de construction d'usager-périodes..... | 60 |
| Tableau 5-2 Différences des écarts-types..... | 68 |
| Tableau 5-3 Distance entre les groupes..... | 69 |
| Tableau 5-4 Caractéristiques principales des super-utilisateurs..... | 84 |
| Tableau 5-5 nombre de groupes d'usagers-périodes différents pendant la durée de vie des super-utilisateurs..... | 88 |
| Tableau 5-6 Statistiques descriptive du nombre de changements de groupe par période active ... | 89 |

LISTE DES FIGURES

| | |
|---|----|
| Figure 3-1 Carte de la communauté métropolitaine de Montréal tirée de (Mamh, 2019) | 23 |
| Figure 3-2 Carte du réseau routier principal de l'île de Montréal tirée de (Lacombe, 2016) | 24 |
| Figure 3-3 Plan du réseau de métro opéré par la STM tiré de (Société de Transport de Montréal, 2018a)..... | 25 |
| Figure 3-4 Part des déplacements produits selon les motifs sur l'ensemble du territoire de l'enquête (chiffres de l'enquête O-D 2013 version 13.2b) | 26 |
| Figure 3-5 Part des déplacements selon les modes sur l'ensemble des trajets effectués par les habitants du territoire de l'enquête (chiffres l'enquête O-D 2013)..... | 27 |
| Figure 3-6 Carte des trois agglomérations de taxi sur l'île de Montréal tirée de (Ville de Montréal, 2019a)..... | 28 |
| Figure 3-7 Tarification des courses tirée de (Ville de Montréal, 2019c) | 29 |
| Figure 3-8 Capture d'écran de l'application mobile de Téo Taxi tirée de (Taxelco, 2018b)..... | 30 |
| Figure 3-9 Zone de couverture du service Téo Taxi à la fin de son activité tirée de (Taxelco, 2018a) | 31 |
| Figure 4-1 Schéma méthodologique du projet de recherche | 34 |
| Figure 4-2 Illustration des caractéristiques du point d'inflexion | 39 |
| Figure 4-3 Distribution cumulée de la durée des courses pour les enregistrements d'une durée inférieure à 200 s | 40 |
| Figure 4-4 Distribution cumulée de la distance des courses en fonction pour les enregistrements d'une distance inférieure à 800 m | 41 |
| Figure 4-5 Distribution cumulée du temps à la prochaine course pour les enregistrements d'un temps inférieur à 200 s | 42 |
| Figure 4-6 Taux de données filtrées par mois sur l'ensemble de la période d'activité de Téo Taxi | 44 |

| | |
|--|----|
| Figure 4-7 Nombre de courses réalisées par mois..... | 45 |
| Figure 4-8 Répartition des courses en fonction de l'heure et du jour de la semaine | 46 |
| Figure 4-9 Carte de chaleur des points de départ des courses..... | 47 |
| Figure 4-10 Carte de chaleur des points de destination des courses | 48 |
| Figure 4-11 Distribution fréquentielle du nombre de courses par utilisateur | 49 |
| Figure 4-12 Distribution cumulée de la distance des courses | 51 |
| Figure 4-13 Distribution cumulée de la durée des courses | 52 |
| Figure 4-14 Distribution cumulée de la vitesse des courses | 53 |
| Figure 5-1 Schéma méthodologique de la construction des groupes d'usagers-périodes | 55 |
| Figure 5-2 Distribution cumulée du temps intercourse..... | 57 |
| Figure 5-3 Comparaison des différents algorithmes de classification en fonction de la structure du jeu de données tirée de (Scikit-learn, 2019a) | 64 |
| Figure 5-4 Graphique de la méthode du coude | 66 |
| Figure 5-5 Dendrogramme de la classification des 1000 centroïdes | 67 |
| Figure 5-6 Représentation des centroïdes des cinq groupes obtenus | 70 |
| Figure 5-7 Distribution du nombre de courses associées à chaque groupe..... | 71 |
| Figure 5-8 Exemple de diagramme boîte à moustache | 71 |
| Figure 5-9 Diagramme boîte à moustache du groupe aller-retour | 72 |
| Figure 5-10 Diagramme boîte à moustache du groupe usage faible | 73 |
| Figure 5-11 Diagramme boîte à moustache du groupe usage moyen | 73 |
| Figure 5-12 Diagramme boîte à moustache du groupe usage exceptionnel..... | 74 |
| Figure 5-13 Diagramme boîte à moustache du groupe usage intensif | 75 |
| Figure 5-14 Distribution des courses du groupe aller-retour en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)..... | 76 |

| | |
|--|----|
| Figure 5-15 Distribution des courses du groupe usage faible en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)..... | 77 |
| Figure 5-16 Distribution des courses du groupe usage moyen en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)..... | 78 |
| Figure 5-17 Distribution des courses du groupe usage exceptionnel en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)..... | 79 |
| Figure 5-18 Distribution des courses du groupe usage intensif en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)..... | 80 |
| Figure 5-19 Distribution des courses pour l'ensemble des groupes en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)..... | 81 |
| Figure 5-20 Population relative des groupes d'usagers-périodes en fonction de la période | 82 |
| Figure 5-21 Nombre moyen de trajets par utilisateur pour chaque période | 83 |
| Figure 5-22 Distribution des périodes de naissance et de fin d'utilisation des super-utilisateurs | 86 |
| Figure 5-23 Population des super-utilisateurs (en usagers-périodes) pour chaque groupe (à gauche) et proportion d'usagers-périodes de super-utilisateurs pour chaque groupe (à droite) | 87 |
| Figure 5-24 Distribution du nombre de changements de groupe par période active | 89 |
| Figure 6-1 Courriel de prise de contact avec le répondant (en anglais à gauche et en français à droite) | 92 |
| Figure 6-2 Aperçu de la plateforme de sondage | 93 |
| Figure 6-3 Distribution des réponses à la question « Quel était le motif de votre déplacement ? » | 97 |
| Figure 6-4 Distribution des réponses aux questions « Avez-vous partagé cette course de taxi avec d'autres passagers ? » (à gauche) et « Est-ce que tous les passagers ont terminé leur course au même endroit ? » (à droite) | 98 |
| Figure 6-5 Distribution des réponses à la question « Qu'auriez-vous fait s'il vous avait été impossible d'effectuer votre course en Téo Taxi ? » | 99 |

| | |
|---|-----|
| Figure 6-6 Distribution des réponses à la question « Quel(s) mode(s) de transport auriez-vous été susceptible de choisir? »..... | 99 |
| Figure 6-7 Distribution des réponses à la question « Avez-vous utilisé uniquement le taxi comme mode de transport pour vous rendre à votre destination finale ? » | 100 |
| Figure 6-8 Distribution des réponses à la question « Quelles sont les raisons qui vous ont poussé à effectuer ce déplacement avec Téo Taxi ? » | 100 |
| Figure 6-9 Distribution des réponses à la question « Qui a payé la course ? »..... | 101 |
| Figure 6-10 Distribution des réponses à la question « Sélectionnez les services dont vous êtes membre / abonné ou que vous utilisez : » | 102 |
| Figure 6-11 Distribution des réponses aux questions « À quelle fréquence avez-vous pris le taxi (autre que Uber ou Téo Taxi) dans les six derniers mois ? » (à gauche) et « À quelle fréquence avez-vous utilisé Uber dans les six derniers mois ? » (à droite) | 103 |
| Figure 6-12 Distribution des réponses à la question « Connaissez-vous l'application mobile Transit ? »..... | 103 |
| Figure 6-13 Distribution des réponses aux questions « Quel âge avez-vous ? » (à gauche) et « Genre : » (à droite) | 104 |
| Figure 6-14 Répartition de la population montréalaise âgée de 20 ans à 84 ans en fonction de l'âge en 2016 (données issues de (Ville de Montréal, 2018))..... | 105 |
| Figure 6-15 Distribution des réponses au question « Quel est votre occupation principale ? » .. | 105 |
| Figure 6-16 Distribution des réponses aux questions « Possédez-vous un permis de conduire ? » (à gauche) et « Combien de véhicules sont à la disposition d'un ou des membres de votre ménage ? » (à droite)..... | 106 |
| Figure 6-17 Carte de lieux de domicile des répondants | 107 |
| Figure 6-18 Distribution des réponses à la question « Combien de personnes habitent votre domicile de façon permanente, y compris vous-même ? | 107 |
| Figure 6-19 Distribution des réponses à la question « Quelle était la tranche de revenu de votre ménage avant impôt en 2017 ? »..... | 108 |

| | |
|---|-----|
| Figure 6-20 Répartition de la population montréalaise en fonction de revenu annuel par ménage en 2016 (données issues de (Ville de Montréal, 2018))..... | 108 |
| Figure 6-21 Distributions de la taille des ménages en fonction du type d'usager (à gauche) et du nombre de voitures par ménages en fonction du type d'usager (à droite) | 112 |
| Figure 6-22 Distribution du revenu par ménage en fonction du type d'usager | 112 |
| Figure 6-23 Distribution des fréquences d'utilisation du taxi en fonction du type d'usager | 113 |
| Figure 6-24 Distribution du motif du déplacement en fonction du type d'usager..... | 114 |
| Figure 6-25 Distribution du nombre de passagers en plus du répondant dans le taxi en fonction du type d'usager..... | 114 |
| Figure 6-26 Distribution des fréquences des abonnements aux différents services de mobilité en fonction du type d'usager | 115 |
| Figure 6-27 Distribution des fréquences des abonnements aux différents services de mobilité en fonction du type d'usager | 116 |
| Figure 6-28 Distribution des réponses à la question « Qui a payé la course ? » en fonction de l'heure de la course..... | 117 |
| Figure 6-29 Distribution des réponses à la question « Est-ce que tous les passagers ont terminé leur course au même endroit ? » en fonction de l'heure de la course | 118 |
| Figure 6-30 Distribution de l'âge des répondants en fonction de l'heure de la course | 118 |
| Figure 6-31 Distribution du nombre de véhicules par ménage en fonction de l'heure de la course | 119 |
| Figure 6-32 Distribution des raisons du choix de Téo Taxi en fonction de l'heure de la course . | 120 |
| Figure 6-33 Distribution du motif du déplacement en fonction de l'heure de la course | 120 |
| Figure 6-34 Distribution de l'âge des répondants en fonction de l'utilisation de Téo Taxi sur la période..... | 122 |
| Figure 6-35 Distribution de l'occupation principale en fonction de l'utilisation de Téo Taxi sur la période..... | 122 |

| | |
|---|-----|
| Figure 6-36 Distribution du revenu annuel par ménage en fonction de l'utilisation de Téo Taxi sur la période..... | 123 |
| Figure 6-37 Distribution de la taille du ménage en fonction de l'utilisation de Téo Taxi sur la période..... | 124 |
| Figure 6-38 Distribution des fréquences des abonnements aux différents services de mobilité en fonction de l'utilisation de Téo Taxi sur la période | 124 |
| Figure 6-39 Distribution des modes alternatifs envisagés en fonction de l'utilisation de Téo Taxi sur la période | 125 |
| Figure 6-40 Distribution de l'utilisation d'autres services de taxi en fonction de l'utilisation de Téo Taxi sur la période..... | 126 |
| Figure 6-41 Distribution des raisons du choix de Téo Taxi en fonction de l'utilisation de Téo Taxi sur la période | 126 |
| Figure 6-42 Distribution du nombre de passagers en plus du répondant dans le taxi en fonction de l'utilisation de Téo Taxi sur la période | 127 |
| Figure 6-43 Distribution du motif du déplacement en fonction de l'utilisation de Téo Taxi sur la période..... | 127 |

LISTE DES SIGLES ET ABRÉVIATIONS

| | |
|--------|--|
| ARTM | Autorité Régionale de Transport Métropolitain |
| BTM | Bureau de Taxi de Montréal |
| CMM | Communauté Métropolitaine de Montréal |
| IPART | Independent Pricing and Regulatory Tribunal of New South Wales |
| NCST | National Center for Sustainable Transportation |
| NYCTLC | New York City Taxi & Limousine Commission |
| RACQ | Royal Automobile Club of Queensland |
| STM | Société de Transport de Montréal |
| URL | Uniform Resource Locator |

LISTE DES ANNEXES

| | |
|--|-----|
| Annexe A – Extrait de la loi concernant les services de transport par taxi, RLRQ c S-6.01 (définition du terme services de transport par taxi) | 139 |
| Annexe B – Valeur des coordonnées de centroïdes et écart type des composantes de chaque groupe | 141 |
| Annexe C – Page d'accueil de la plateforme de sondage | 142 |
| Annexe D – Formulaire d'information et de consentement du sondage..... | 143 |
| Annexe E – Table de valeurs du test du χ^2 | 147 |
| Annexe F – Tableau de résultat du test du χ^2 entre la distribution des réponses des super-utilisateurs et des utilisateurs normaux | 148 |
| Annexe G – Tableau de résultat du test du χ^2 entre la distribution des réponses pour les courses réalisées en journée et en soirée | 149 |
| Annexe H – Tableau de résultat du test du χ^2 entre la distribution des réponses pour les courses afiliées à aux groupes de forte intensité d'utilisation et aux groupes de faible intensité d'utilisation..... | 150 |

CHAPITRE 1 INTRODUCTION

Depuis les premières apparitions du taxi dans les capitales européennes à la fin du XVIIIe siècle, le taxi a subi de nombreuses transformations notamment grâce aux progrès de l'automobile (Cooper, Mundy et Nelson, 2010). Le taxi est passé de la voiture à cheval à la voiture électrique à la fin du XIXe siècle, puis de la voiture électrique à la voiture thermique au début du siècle suivant. Aujourd'hui, la voiture électrique redevient une solution compétitive. Les progrès importants dans les télécommunications ont également fait évoluer cette industrie; on peut penser à l'arrivée des systèmes de communication par radio embarqués dans les véhicules de taxi, ou encore plus récemment aux nouvelles possibilités offertes grâce l'introduction massive des téléphones intelligents dans la population. Cette évolution ne semble pas s'arrêter, nous pouvons dès à présent être les observateurs des premières expérimentations de service de transport de personnes par véhicules sans conducteur (« Waymo », 2019).

L'introduction des technologies de l'information et la hausse du taux de possession de téléphone intelligent ont permis au paysage de l'industrie du taxi, et plus largement de l'industrie du transport à la demande, de s'élargir et de proposer de nouveaux services (Gössling, 2018). Plus que l'optimisation opérationnelle des systèmes de transport par taxi cette évolution est également centrée sur le client et sur la bonification de l'expérience utilisateur.

1.1 Problématique

L'écosystème de la mobilité urbaine est en pleine mutation. Depuis quelques années, l'introduction de nouvelles technologies et de nouveaux acteurs dans le secteur de la mobilité ont modifié la manière dont nous pensons nos déplacements (Gössling, 2018). Dans l'industrie du taxi montréalaise, l'implantation progressive de GPS embarqués dans les véhicules a permis d'élargir les possibilités d'analyse des comportements de transport. Ainsi plusieurs recherches ont été rendues possibles grâce à ces données (Bouchard, 2018; Geneste, 2017; Lacombe, 2016; Laviolette, 2017). Ces recherches ont porté sur des sujets très diversifiés allant du développement d'indicateurs de suivi pour l'industrie au perfectionnement de modèles de consommation énergétique pour des véhicules électriques. Aucun de ces travaux n'a cependant porté sur la caractérisation des utilisateurs du taxi.

Dans la région de Montréal, les usagers d'autres modes ont fait l'objet d'études centrées sur leurs usagers (Deschaintres, 2018; Wielinski, 2018). Ces recherches sont rendues possibles le plus souvent grâce à la présence d'informations relatives à l'utilisateur et à un identifiant permettant de regrouper les différents trajets réalisés par un même usager. Explorer l'usage que les utilisateurs font d'un mode permet d'approfondir notre compréhension de ce mode.

Depuis sa création en 2015, l'entreprise de taxi montréalaise Téo Taxi recueille différentes informations sur les déplacements des usagers de son service. Le fonctionnement particulier de cette compagnie est détaillé dans le Chapitre 3. Grâce à ces données, il est désormais possible d'étudier une partie de l'écosystème du taxi montréalais (les courses de taxi faites par l'intermédiaire de Téo Taxi) sous l'angle de « l'utilisateur ». Cela constitue donc une opportunité pour améliorer la compréhension des utilisateurs et de l'utilisation du taxi à Montréal.

Plus que l'utilisation du taxi, les facteurs liés à son utilisation restent également peu étudiés dans la région de Montréal. Bien que la région possède de nombreuses informations sur la mobilité de ses résidents grâce notamment aux Enquêtes Origine-Destination réalisées depuis 1970 (Exo, 2015), il est difficile d'étudier avec certitude le taxi en raison du très faible nombre d'observations recueillies concernant le taxi dans ces enquêtes.

On peut donc apercevoir ici un des freins à la compréhension d'un système de transport par taxi et de son intégration dans un écosystème de mobilité urbaine. Il nous apparaît essentiel de pouvoir mieux comprendre les usagers du taxi et répondre à de nombreuses questions déjà étudiées pour d'autres modes de transport. Qui sont les utilisateurs du taxi? Comment l'utilisent-ils? Pourquoi ont-ils choisi ce mode de transport? Prendre un taxi est-il un geste habituel pour eux ou alors un comportement exceptionnel?

1.2 Objectifs

Ce mémoire est donc centré sur l'étude des utilisateurs d'un service de taxis électriques dans la région de Montréal. Son objectif principal est de définir et caractériser l'utilisation de ce même service de taxi. De plus, la méthodologie d'analyse développée dans ce mémoire doit être facilement reproductible et applicable à d'autres services de transport de personnes dès lors que les informations recueillies par ce système sont comparables à celles utilisées pour ce mémoire. L'objectif principal peut se décomposer en une liste de plusieurs sous-objectifs :

- Définir des indicateurs pertinents d'utilisation du taxi à l'échelle de l'utilisateur;
- Identifier des comportements d'utilisation à la fois sur une échelle de temps relativement courte pour observer de possibles changements d'utilisation dans le temps, mais également sur une échelle de temps plus longue prenant en compte les possibles effets de saisonnalité dans les analyses;
- Explorer les facteurs favorisant l'utilisation du taxi du point de vue de l'utilisateur.

1.3 Structure du mémoire

Ce mémoire est composé de sept chapitres différents. Le Chapitre 1 (ce chapitre) constitue l'introduction de ce projet. Il permet de définir le cadre de l'étude et poser la problématique principale de la recherche. Le Chapitre 2 présente une revue de littérature sur les systèmes de transport par taxi. Cette revue permet de comprendre tout d'abord ce qu'est un tel système de transport, puis quels sont les enjeux et les problématiques inhérentes à ces systèmes. La forte évolution observée dans l'industrie du taxi depuis quelques années est également abordée dans cette revue.

Le Chapitre 3 ancre ce projet dans le contexte géographique de la région de Montréal. Pour comprendre l'utilisation d'un système de transport par taxi, il est important de connaître les spécificités locales du taxi. L'opérateur de taxi Téo Taxi et son modèle de fonctionnement spécifique sont également présentés dans ce chapitre. Les données utilisées pour la suite du projet proviennent de cet opérateur.

La méthodologie du projet de recherche est introduite dans le Chapitre 4. La méthodologie développée comporte deux aspects qui seront tous deux détaillés dans des chapitres distincts. Les hypothèses initiales de traitements et de filtrages des données sont explicitées dans ce chapitre. Enfin, les données issues du processus de filtrage et leurs principales caractéristiques sont également présentées dans ce chapitre.

Les deux chapitres suivants présentent les deux dimensions de la recherche. Dans le Chapitre 5, l'usage de Téo Taxi est exploré en fonction des courses réalisées par les utilisateurs. Cet usage est étudié via le prisme de l'utilisateur. Une méthodologie détaillée de cette étude est présentée dans

cette partie pour compléter l'introduction méthodologique faite au chapitre précédent. Les résultats obtenus suite à l'application de cette méthode sont également présentés dans ce chapitre.

Le Chapitre 6 se concentre sur la description d'une enquête réalisée auprès des utilisateurs. Les éléments méthodologiques relatifs à la construction et la réalisation de ce sondage sont présentés dans le début de ce chapitre. La présentation des résultats et des analyses réalisées à partir des réponses à l'enquête est faite dans la suite du chapitre.

Enfin le Chapitre 7 conclut ce mémoire. Dans ce chapitre, les principaux résultats sont regroupés. Les contributions du projet sont également détaillées. Enfin les limitations et les perspectives de recherche, qui sont deux éléments importants dans ce projet de recherche, sont également décrites dans ce chapitre.

CHAPITRE 2 REVUE DE LITTÉRATURE

Cette revue de littérature vise à mettre en contexte et introduire les enjeux relatifs à l'industrie du taxi. Le taxi dans la région de Montréal a déjà été le sujet de plusieurs recherches (Geneste, 2017; Lacombe, 2016; Laviolette, 2017). Ces mémoires présentent de manière détaillée l'histoire du taxi depuis ses débuts au XVIIIe siècle ainsi que l'évolution détaillée du cadre législatif qui régit le marché nord-américain dans leur revue de littérature respective. De plus, d'autres ouvrages et rapports externes au contexte montréalais détaillent également l'historique et le cadre réglementaire de cette industrie (Austin et Zegras, 2012; Cooper et al., 2010; Harding, Kandlikar et Gulati, 2016; Schaller, 2007). Bien qu'il soit nécessaire d'aborder ces points de la revue de littérature, ils ne sont pas approfondis ici. L'introduction qui est faite dans le début de la revue de littérature permet une meilleure compréhension des problématiques inhérentes à cette industrie. Dans la suite, nous explorons le rôle d'un système de taxi dans l'organisation des transports d'un territoire. Les nouveaux acteurs du transport de personnes à la demande ainsi que leurs relations avec l'industrie du taxi seront abordés dans une autre partie. Enfin, les questions d'identification des utilisateurs et de leur comportement seront abordées.

2.1 Brève histoire du taxi et de son cadre réglementaire

2.1.1 Définition et développement historique du taxi

Le taxi a vu le jour vers le milieu du XVIIIe siècle dans les capitales européennes (Cooper et al., 2010). D'après l'Office québécois de la langue française, le taxi peut être défini comme :

« Voiture automobile de place, munie d'un compteur qui indique le prix de la course. »
(Office québécois de la langue française, 1986)

Le terme « voiture automobile de place » dans cette définition renvoie à un véhicule conçu pour transporter des personnes et non de la marchandise. Cette définition à l'avantage d'être relativement concise, elle ne permet cependant pas de rendre compte de la complexité réglementaire qui se cache derrière la définition d'un système de transport par taxi. Il existe donc de nombreuses définitions extrêmement précises du taxi dans les textes de loi qui régissent cette industrie. La définition d'un « service de transport par taxi » en vigueur au Québec (extraite de la loi concernant les services de transport par taxi, RLRQ c S-6.01) est disponible à titre d'exemple

dans l'Annexe A. La longueur et la précision de cette définition illustrent bien la complexité de caractériser avec exactitude le terme « taxi ».

Le mot « taxi » lui-même est apparu plusieurs années après la genèse de ce mode de transport il faudra attendre 1891 avec l'invention du « taximètre » par Wilhelm Bruhn. C'est l'usage généralisé de cette invention dans les véhicules qui a donné son nom au taxi (Cooper et al., 2010). Un taximètre est un compteur horokilométrique qui permet de calculer le prix exact d'une course en fonction de sa distance et de sa durée. Le principe de tarification basée sur la distance et la durée de la course ainsi que ce type d'appareil sont encore utilisés aujourd'hui. Très souvent, la réglementation en vigueur impose aux conducteurs de taxis de disposer d'un taximètre à bord de leur véhicule et de l'utiliser pour déterminer la tarification des courses, ce qui est le cas au Québec (*Règlement sur les services de transport par taxi*, s. d.).

Le taxi est un système de transport dit porte-à-porte. Il n'y a pas de routes ou d'horaires prédéfinis à l'avance. Pour effectuer une course en taxi, un client dispose de plusieurs alternatives. Il peut entre autres héler un taxi dans la rue, réserver une course auprès d'un opérateur ou un chauffeur de taxi et enfin se rendre à un poste d'attente de taxi (Aarhaug et Skollerud, 2014). La répartition des différentes méthodes d'utilisation du taxi dépend du contexte. Dans les zones où la densité de l'offre de taxi est importante, un client n'a pas à attendre très longtemps dans la rue avant de trouver un taxi. Le nombre de courses hélées est important dans ce contexte (Schaller, 2007). A certains endroits très spécifiques ou lorsque la densité de l'offre diminue, le nombre de courses débutant à un poste d'attente du taxi est important. Ces points sont notamment les aéroports, les gares routières et ferroviaires, les environs de certains grands hôtels et centres de congrès (Aarhaug et Skollerud, 2014). Enfin, la proportion de courses préalablement commandées est la plus importante dans des milieux où le nombre d'alternatives de transport est faible ou lorsque l'accessibilité aux éventuelles postes d'attente est complexe. Dans cette situation, la seule solution pour un client d'être certain de pouvoir effectuer son déplacement en taxi est d'en commander un (Schaller, 2007).

Un système de transport par taxi n'est pas un système de transport de masse, sa capacité est relativement faible. Le nombre de passagers par voiture est limité et le nombre de taxis sur la route est également limité par des mesures réglementaires (Austin et Zegras, 2012). Un système de transport de taxi présente cependant l'avantage de ne pas nécessiter d'infrastructure spécifique pour fonctionner. Cette propriété couplée au fait que sa capacité est relativement faible font qu'il est une

des seules alternatives à l'automobile dans des milieux peu denses où l'implantation d'un réseau de bus ne serait pas viable (Austin et Zegras, 2012) citant (Gilbert et Samuels, 1982).

2.1.2 Le cadre réglementaire

La première trace historique de réglementation sur le taxi date de 1654 à Londres (Cooper et al., 2010). La réglementation est utilisée comme un outil permettant de garantir au client un haut niveau de service à un prix abordable. Le cadre législatif peut être divisé en trois axes majeurs : le contrôle de la qualité, le contrôle de la quantité et le contrôle économique (Cooper et al., 2010). La réglementation dans ce secteur est également importante. Le fonctionnement du taxi est basé sur la confiance que le client a envers le système. Lorsqu'un passager rentre dans un taxi, il peut connaître le niveau de service offert par le chauffeur et son intégrité professionnel uniquement en se basant sur la réglementation à laquelle ce dernier doit obéir (Harding et al., 2016). Au fur et à mesure des progrès technologiques et de l'évolution de la société, le cadre législatif n'a cessé d'évoluer pour faire également évoluer le taxi avec son époque et ne pas faire disparaître cette industrie (Cooper et al., 2010).

Deux décennies ont marqué des tournants majeurs dans les cadres législatifs régulant ce secteur. Dans les années 1930, le cadre législatif est devenu plus stricte afin de protéger le marché à la suite des conséquences de la grande récession. Dans cette période de crise économique, deux phénomènes ont été observés sur le marché. Il y a tout d'abord une baisse importante de la demande, car beaucoup de personnes n'ont soudainement soit plus eu le besoin de prendre le taxi (course pour raisons professionnelles notamment) soit plus eu les moyens (baisse du niveau de vie). Cette baisse de la demande s'est accompagnée d'une forte hausse de l'offre, car beaucoup de personnes ayant perdu leur emploi ont choisi d'offrir des services de transport de personnes afin de subvenir à leurs besoins. Ces deux phénomènes ont déséquilibré le marché et fortement accru la compétition. La qualité du service s'est vue fortement impactée face à cette situation (Harding et al., 2016). Pour garantir une qualité de service suffisante, le cadre réglementaire est devenu plus contraignant en augmentant nettement les exigences de qualité de service et également en mettant des barrières d'entrée avec l'instauration des systèmes de « licence » (Laviolette, 2017) citant (Frankena, 1984).

Dans les années 1970 et 1980, le phénomène inverse s'est produit. Les courants politiques et économiques de l'époque prônant le marché libre et la liberté de concurrence ont mené à une

dérégulation du marché. D'après les défenseurs de la libre économie, le cadre législatif trop restreint dans l'industrie du taxi est un frein à l'atteinte de l'équilibre réel du marché (Schaller, 2007). Selon eux, la libre concurrence permettra d'améliorer le service offert aux utilisateurs tout en réduisant dans le même temps le prix des prestations (Harding et al., 2016). Des expériences de dérégulation du marché ont été menées dans de nombreuses villes nord-américaines. Ces expériences ont généralement été infructueuses et les bénéfices présumés de la dérégulation n'ont pas été observés (Laviolette, 2017) citant (Price Waterhouse, 1993). L'échec de ces expériences est notamment dû à deux spécificités du marché : tout d'abord le taxi est un marché basé sur la confiance puisque le client ne peut connaître à l'avance les performances du vendeur (le conducteur de taxi), ensuite ce marché est dit « étroit » dans le sens où il y a un nombre relativement faible d'acheteurs et de vendeurs. Lorsque de tels marchés sont faiblement régulés, la volatilité des prix est grande et la qualité du service n'est pas homogène (Harding et al., 2016).

2.2 Le rôle du taxi dans un système de transport urbain

En raison du faible nombre de courses effectuées en taxi comparé au nombre de déplacements tous modes confondus sur le même territoire, le marché du taxi peut être vu comme un marché de niche (Harding et al., 2016). Cependant, la faible part modale du taxi ne minimise pas l'importance du taxi dans un système de transport urbain (Rayle, Dai, Chan, Cervero et Shaheen, 2016). Dans de nombreuses villes, le taxi peut être utilisé en complément du réseau de transport collectif. Il va permettre (lorsqu'il est utilisé à bon escient) de combler les besoins auxquels le réseau de transport collectif ne peut pas répondre comme dans les zones de faible densité. Dans ces zones, la viabilité d'un réseau de transport collectif est difficilement atteignable. Même dans les zones à forte densité, le taxi peut être complémentaire aux autres modes notamment à des moments de la journée où la fréquence du réseau de transport collectif est faible, voire nulle (en particulier la nuit) (Austin et Zegras, 2012). En plus de ce rôle de complémentarité, le système de transport par taxi peut être utilisé comme mode de substitution aux autres modes de transport. Cela est particulièrement fréquent dans le cas de déplacements où le confort et la rapidité priment sur le coût du déplacement (Aarhaug et Skollerud, 2014).

2.2.1 Le taxi : un mode de substitution

L'utilisation du taxi comme un mode de substitution est particulièrement observable pour deux types de courses : les courses desservant un aéroport et les courses effectuées dans les quartiers centraux et quartiers d'affaires. Les courses qui se destinent ou partent des aéroports sont une part très importante du marché pour le taxi (Cooper et al., 2010). Les aéroports sont généralement situés en zone périphérique des grandes villes et s'y rendre en transport en commun peut être long. De plus même si l'accès à l'aéroport est facile et rapide en transport en commun, le taxi reste une option de déplacement compétitive grâce à sa commodité (Rayle et al., 2016). Transporter des objets encombrants comme des valises peut s'avérer très complexe dans le réseau de transport collectif particulièrement en période de fort achalandage. Il est très fréquent que l'accès aux zones aéroportuaires fasse l'objet de réglementation spécifique (tarification, zone de dépôt des passagers, zone d'embarquement, etc.) et d'une gestion spécifique de la part de l'aéroport (gestion des flux de passagers, création de postes d'attente pour les chauffeurs, etc.) (Cooper et al., 2010).

De nombreuses courses de taxis sont également effectuées en centre-ville et quartiers d'affaires, dans des zones où la desserte en transport en commun est généralement bonne et peuvent donc être considérées comme une utilisation de substitution par le taxi (Austin et Zegras, 2012). Ce type d'utilisation est observable dans les zones où la densité d'emplois est importante et des enquêtes ont révélé que la proportion de déplacements effectués pour des motifs professionnels est plus importante dans ces zones que dans les quartiers périphériques (IPART, 2017). Les critères déterminant le choix du mode de transport peuvent varier en fonction des caractéristiques des personnes et de motifs du déplacement (Jara-Díaz et Guevara, 2001). Lorsque les notions de rapidité, de confort et d'intimité prennent une grande importance et que la capacité à payer est grande, le taxi devient un mode compétitif.

2.2.2 Le taxi : un système complémentaire au transport public

L'utilisation complémentaire du taxi est très différente de celle décrite dans la section précédente. La complémentarité entre un système de transport par taxi et le réseau de transport public peut prendre la forme d'une complémentarité spatiale entre ces deux systèmes (Austin et Zegras, 2012). L'absence de construction d'infrastructures spécifiques ainsi que l'absence de trajets prédéfinis sont entre autres des facteurs qui permettent à un système de taxi d'opérer dans des zones à faible

densité. L'implantation, dans ces situations, d'une ligne de bus viable avec une fréquence suffisante pour être attractive est très complexe (Austin et Zegras, 2012). Certains opérateurs de transport public s'appuient sur cette complémentarité en subventionnant le transport par taxi pour effectuer du rabattement vers leur réseau. Ce type de service est par exemple opéré à Montréal par la *Société de Transport de Montréal* (STM) (« Taxi collectif », 2019). Cela permet aux opérateurs d'étendre leur zone de desserte et d'augmenter la mobilité des résidents des quartiers à faible densité. Cependant, dans de nombreux quartiers mal reliés aux réseaux de transport en commun, la desserte proposée par le taxi ne permet pas de combler le manque d'options de mobilité (Atkinson-Palombo, Varone et Garrick, 2019). Une des propriétés intrinsèques des systèmes de transport par taxi est que chaque chauffeur va chercher à optimiser ses gains particuliers (Lacombe, 2016) citant (Phithakkitnukoon, Veloso, Bento, Biderman et Ratti, 2010; Yang, Wong et Wong, 2002). Même si le cadre législatif permet d'assurer la qualité et l'offre sur l'ensemble du territoire, la multiplication des acteurs privés dans le secteur implique que l'optimisation globale des gains est complexe, voire impossible. Ainsi, un chauffeur de taxi va privilégier les zones où la densité de la demande est la plus forte (Austin et Zegras, 2012). Cela a pour conséquence d'augmenter le temps d'attente du taxi dans les quartiers les moins bien desservis et donc de diminuer son attractivité.

Il ne faut pas oublier que la complémentarité avec le transport en commun n'est pas uniquement géographique. En effet, même dans des zones urbaines denses et bien desservies par les réseaux, certaines personnes à mobilité réduite ne peuvent pas facilement utiliser les réseaux de transport public (Austin et Zegras, 2012). Même si les politiques de développement des réseaux mettent tout en œuvre pour que l'accessibilité soit la meilleure possible, pouvoir offrir un service utilisable par l'intégralité de la population sans exception reste un défi pour les villes (Société de Transport de Montréal, 2016). Pour certaines personnes à mobilité réduite ou en situation de handicap, le taxi reste aujourd'hui une des seules alternatives disponibles pour effectuer leurs déplacements. De plus, le vieillissement de la population, que de nombreux pays occidentaux sont actuellement en train d'observer, aura également pour conséquence d'augmenter le nombre de personnes dans l'incapacité de conduire une voiture ou d'utiliser le réseau de transport public (Lacombe, 2016). Le nombre de personnes dépendantes du taxi et des autres modes de transport parallèle n'est donc pas amené à décroître.

2.2.3 L'industrie du taxi comme facteur d'intégration sociale

Enfin, le taxi n'a pas uniquement un rôle de système de transport, il a également un rôle social. Le taxi permet d'offrir une alternative de transport (même si elle reste relativement onéreuse) à des ménages à faible revenu ne possédant pas de véhicule pour des raisons financières (Rayle et al., 2016) citant (Webster, Weiner et Wells, 1974). Pour ces ménages, avoir accès au taxi signifie avoir accès à un mode de transport pour répondre à des besoins occasionnels, ce qui est moins coûteux que de posséder une automobile. Il a été historiquement observé que le taxi est à la fois utilisé par des ménages à haut revenu et des ménages à très faible revenu (Austin et Zegras, 2012) citant (Frankena, 1984). L'introduction de nouveaux acteurs technologiques (présentée dans le paragraphe suivant) a renforcé ce rôle joué par le taxi et les transports parallèles (Atkinson-Palombo et al., 2019).

L'industrie du taxi peut également être vue comme un facteur d'intégration sociale. Cette industrie offre des opportunités d'emploi à de nombreuses personnes issues de l'immigration (NYC Taxi & Limousine Commission, 2014, 2016, 2018). La commission municipale responsable de l'organisation des systèmes de transport parallèle dans la région métropolitaine de New York (*New York City Taxi & Limousin Commission* (NYTLC)) publie régulièrement des communiqués de presse précisant l'origine ethnique des conducteurs. Seule une part très faible des conducteurs est originaire des États-Unis (8,5 % en 2018) (NYC Taxi & Limousine Commission, 2018). Il faut rester cependant mesuré face à ces chiffres. La presse américaine publie très fréquemment des enquêtes décrivant les méthodes mafieuses de certains acteurs du secteur profitant de la situation de détresse de ces personnes pour les exploiter financièrement en les poussant notamment à contracter des prêts frauduleux (Rosenthal, 2019). De plus, certains conducteurs issus de l'immigration possèdent un très haut niveau de qualification. Si les opportunités offertes dans un premier temps par l'industrie du taxi peuvent leur être bénéfiques, il faut s'assurer qu'elles ne sont pas un frein à la pleine exploitation de leur haut niveau de compétence sur le long terme (L. Xu, 2012).

2.3 Un secteur en pleine mutation

Nous assistons actuellement à une modification profonde de l'équilibre du marché du transport à la demande et du transport parallèle. La hausse du taux de possession de téléphones intelligents a

rendu possible l'introduction de nouveaux acteurs technologiques comme Uber ou encore Lyft pour n'en citer que deux (Harding et al., 2016). Ces services sont très récents (création de Uber en 2009 et de Lyft en 2012) et connaissent une très forte croissance. En décembre 2018 Uber opérait des véhicules dans plus de 700 villes réparties dans 63 pays (Uber, 2018a). Aujourd'hui, Lyft estime que ses services sont accessibles à 95 % de la population américaine (Lyft, 2018). Il n'existe pas de terminologie officielle pour ce segment de l'industrie du transport de personnes par automobile. Un terme fréquemment rencontré dans la littérature anglophone qui ne présente pas d'équivalent français est *Transportation Network Company* (TNC). Il existe également d'autres termes anglophones pour désigner ces entreprises comme *ride-sourcing services* par exemple. Dans la littérature francophone, ces entreprises peuvent être désignées comme des acteurs de la mobilité à la demande. Le terme de Véhicule de Tourisme avec Chauffeur (VTC) est issu du cadre réglementaire français et peut également être rencontré.

2.3.1 Impact de ces nouveaux acteurs sur l'équilibre du marché

Les services offerts par ces nouveaux opérateurs de mobilité à la demande sont, dans beaucoup de points, comparables aux services proposés par l'industrie du taxi. Pour effectuer une course, le client doit la réserver depuis une application mobile. Il demeure cependant interdit aux chauffeurs de ces services d'accepter des courses hélées directement dans la rue dans la plupart des villes. Seuls les chauffeurs détenteurs d'un permis de taxi sont autorisés à réaliser ce type de course (Rayle et al., 2016). Une autre différence importante avec le taxi est la tarification des courses. Puisque ces services ne sont pas concernés par le cadre réglementaire strict qui encadre les systèmes de transport par taxi, ils ne sont pas tenus de respecter la politique de tarification en vigueur dans la zone d'activité. De nombreux acteurs ont opté pour une tarification dynamique (plus onéreuse en période de forte demande et plus abordable en période de faible demande) qui pour à effet de mieux équilibrer l'offre et la demande (Ngo, 2015).

Des comparaisons ont montré que les services proposés par Uber peuvent être plus rapides et moins dispendieux que le taxi pour des courses similaires et ce, particulièrement dans les secteurs les plus défavorisés où comme nous l'avons vu la demande peut être faible (Atkinson-Palombo et al., 2019). L'analyse des données des taximètres dans la ville de Los Angeles a révélé qu'entre 2009 et 2014, le nombre de courses effectuées en taxi a diminué de 18 % et les revenus générés par l'industrie du taxi ont baissé de 9 %. Cette diminution de l'activité a débuté lorsque les entreprises

Uber et Lyft ont commencé à offrir leurs services dans la région en 2012 (Waheed et al., 2015). La diminution du nombre de courses effectuées en taxi après l'introduction d'un acteur de mobilité à la demande dans le territoire a également été observée dans de nombreuses villes nord-américaines (Ngo, 2015). Le marché de ces nouveaux services est en concurrence directe avec le taxi. Des enquêtes sur la mobilité menées en Californie ont révélé que pour la majorité des utilisateurs de services de mobilité à la demande, l'utilisation du taxi est l'alternative de transport qu'ils auraient choisie dans le cas d'une impossibilité de faire leur déplacement via les applications Uber ou Lyft (Circella et al., 2017; Rayle et al., 2016). Il est également intéressant de connaître dans quelle mesure la relation inverse est également vérifiée. Quelle proportion des utilisateurs de services de taxis réguliers se reporterait sur des services de mobilité à la demande en cas de non-disponibilité du service? Une enquête menée à Sydney nous permet d'avoir un début de réponse à cette question. Cette enquête montre que, dans un contexte urbain, plus de la moitié des répondants qui n'ont pas réussi à faire un déplacement en taxi ont choisi de prendre le réseau de transport en commun. Seul un pourcent des répondants ont utilisé un service de mobilité à la demande (IPART, 2017). Le fort ancrage des problématiques de transport dans leur contexte géographique implique que ces résultats, observés à Sydney ne peuvent être extrapolés dans d'autres contextes. Il faut également garder à l'esprit qu'en transport, il est très délicat d'attribuer avec certitude des liens de causalité, car le système étudié n'est pas isolé. Rien ne nous permet d'affirmer que la décroissance du secteur est uniquement due à l'introduction de services de mobilité à la demande. D'autres événements ou phénomènes socio-économiques qui ont eu lieu à la même période peuvent également avoir leur part de responsabilité.

2.3.2 Les réactions de l'industrie

Face à cette nouvelle concurrence, de nombreux mouvements de contestation ont émergé parmi les chauffeurs de taxi. Des exemples de protestation parfois violente sont régulièrement rapportés dans les médias populaires (Le Figaro, 2019). Au-delà des mouvements de protestation, des acteurs de l'industrie du taxi réclament des modifications du cadre réglementaire qui leur permettraient une concurrence plus équitable entre ces deux industries (Yıldızgöz et Murat Çelik, 2019). Le cadre réglementaire très contraignant peut être vu comme un obstacle rendant très difficile l'adaptation du marché du taxi à ce nouveau marché. Les réactions des autorités régulatrices ou des acteurs principaux du marché sont diverses. Alors que certains acteurs continuent à ignorer ces nouveaux

concurrents, d'autres en revanche essayent de concurrencer les services de mobilité à la demande et ainsi conserver une part du marché du transport de personne par automobile la plus grande possible. Une des solutions qui est également explorées par certaines villes comme Singapour est de créer une coopération vertueuse entre ces deux services dans le but de ne pas pénaliser les passagers (Yıldızgöz et Murat Çelik, 2019).

Les acteurs du secteur sont également conscients de l'attrait que ces nouveaux services ont vis-à-vis du public et de l'impact que l'implantation de nouvelles technologies peut avoir sur la qualité du service. Afin de mieux répondre aux attentes des consommateurs, de nombreuses applications permettant aux utilisateurs de commander un taxi grâce à leur téléphone intelligent ont été créées. Ces services sont donc très comparables pour l'utilisateur à ceux offerts par les entreprises de mobilité à la demande (Le Taxi, 2019; Yıldızgöz et Murat Çelik, 2019). Certaines applications incluent également des fonctionnalités de paiement. Pour utiliser ces fonctionnalités, le client doit préalablement fournir ses coordonnées bancaires et il sera ensuite automatiquement prélevé du montant à payer à la fin de chaque course. La facilité de paiement ainsi que la facilité de réservation des courses sont deux des raisons d'utilisation les plus importantes pour les utilisateurs de services de mobilité à la demande (Frazzani, Grea et Zamboni, 2016). Ces deux fonctionnalités semblent clefs pour la réussite de ces projets.

2.3.3 Le futur de l'industrie

Avec les modifications de l'équilibre fragile du marché du taxi, certains avaient prédit dès 2012 la « fin du taxi » ou au moins du taxi tel que nous le connaissions en 2012 (Mathis, 2012). Le taxi va-t-il disparaître? Ce mode de transport est à un moment charnière de son histoire. Pour faire face à ce défi, de nombreux gouvernements tentent de réformer ce secteur. Le gouvernement du Québec a récemment présenté le projet de loi 17 (Assemblée nationale du Québec, 2019). Ce projet de loi qui vise à *moderniser l'industrie du taxi* propose notamment une refonte du cadre tarifaire en permettant aux chauffeurs de taxi d'appliquer une tarification dynamique lorsqu'une course remplit un certain nombre de conditions. Ce projet s'inscrit dans une logique de déréglementation du marché du taxi. Il pourrait avoir pour effet d'atténuer les différences entre les services offerts par les acteurs historiques de l'industrie et les acteurs de la mobilité à la demande comme cela a été fait à Singapour (Yıldızgöz et Murat Çelik, 2019). Il est également intéressant de voir que les entreprises Uber et Didi (première entreprise de mobilité à la demande sur le marché chinois)

proposent toutes deux des services de taxi régulier intégrés dans leurs applications, même si cette activité reste marginale (DIDI, 2019; Uber, 2018b). Ces constats nous laissent à penser que la séparation entre le taxi et les services de mobilité à la demande est amenée à s'estomper.

2.4 Les utilisateurs

Si nous comparons à d'autres modes de transport urbain (transport en commun, vélopartage, etc.), les informations sur les utilisateurs des systèmes de transport par taxi sont très minces. En effet lorsqu'un client fait une course de taxi, dans la plupart des cas, nous ne disposons pas d'information sur ce dernier. De nombreuses informations sur la course peuvent être enregistrées grâce à la multiplication des systèmes de récolte de données GPS embarqués dans les véhicules (Lacombe, 2016), mais aucun lien n'est généralement possible entre l'utilisateur et la course. Cet utilisateur fait-il régulièrement cette course ? Est-ce un utilisateur qui utilise fréquemment le taxi, pour effectuer des courses différentes ou est-ce au contraire un utilisateur occasionnel ? Répondre à ces questions permettrait d'améliorer notre compréhension de l'utilisation du taxi. Il est cependant difficile d'y répondre étant donnée la complexité de l'accès aux données.

2.4.1 Catégories d'utilisateurs

Une classification simple des utilisateurs du taxi, qui peut être applicable à l'ensemble des modes de transport, peut-être la suivante : usagers captifs et usagers non-captifs (Austin et Zegras, 2012) citant (Beimborn, Greenwald et Jin, 2003). Nous entendons ici par usagers captifs les utilisateurs qui n'ont aucune autre alternative que le taxi pour effectuer leur déplacement. À l'inverse, un usager non captif possède plusieurs alternatives de transport et choisit le taxi, car c'est le mode qui lui apparaît comme le plus adapté pour le déplacement qu'il effectue (Austin et Zegras, 2012) citant (Beimborn et al., 2003).

Paradoxalement, pour de nombreux usagers qui sont captifs du taxi, l'accès à ce service de transport est complexe. Comme nous l'avons soulevé précédemment, la qualité du service de taxi peut être moindre (en particulier le temps d'attente) dans les zones de moins forte demande et/ou excentrées du centre-ville. Or ce sont notamment dans ces zones que les alternatives à l'automobile sont les moins nombreuses. Pour les populations les plus vulnérables n'ayant pas d'accès à l'automobile résidant dans ces secteurs, la situation est problématique (Atkinson-Palombo et al., 2019).

2.4.2 Les enquêtes

Afin de mieux connaître la clientèle du taxi, il est possible d'effectuer des enquêtes auprès des utilisateurs. Nous avons recensé, dans cette revue, huit enquêtes menées auprès des utilisateurs de services de taxi ou alors auprès d'un échantillon représentatif de la population et abordant des problématiques liées au taxi. Le Tableau 2-1 permet de comparer les principales caractéristiques des enquêtes recensées. Ces enquêtes sont, dans la plupart des cas, effectuées par les opérateurs ou des autorités responsables de la régulation du marché (Frazzani et al., 2016; Hara Associates Inc. et Corey, Canapary & Galanis, 2013; IPART, 2017; NYC Taxi & Limousine Commission, 2011, 2014; RACQ Public Policy Department, 2016). D'autres enquêtes ont été effectuées par le milieu de la recherche universitaire soit traitant uniquement du taxi et des autres services de mobilité parallèle (Rayle et al., 2016), soit s'intégrant à des enquêtes plus larges sur la mobilité des personnes dont une des parties s'intéresse à l'utilisation du taxi et des services de mobilité à la demande (Circella et al., 2017). Les méthodologies pour ces enquêtes sont diverses. Certaines enquêtes ont été conduites par internet et/ou entretiens téléphoniques. Dans ce cas, le recrutement est effectué par l'intermédiaire de panels représentatifs de la population (Circella et al., 2017; Hara Associates Inc. et Corey, Canapary & Galanis, 2013; IPART, 2017; RACQ Public Policy Department, 2016). Il est également possible, pour ces enquêtes, de recruter uniquement des utilisateurs du taxi (Frazzani et al., 2016). Pour certaines enquêtes, les données ont été récoltées par entrevues faites dans la rue en interceptant des clients ou des aéroports. Les répondants ont été recrutés dans des lieux où l'activité du taxi est forte (Hara Associates Inc. et Corey, Canapary & Galanis, 2013). Enfin, une dernière méthodologie utilisée à New York par la NYTLC, est de poser des questions aux utilisateurs directement à l'intérieur du véhicule. Cette méthodologie de sondage est possible à New York, car chaque taxi est équipé d'un système de divertissement interactif par l'intermédiaire duquel il est possible de poser des questions (NYC Taxi & Limousine Commission, 2014).

Si les opérateurs de taxis ont pour habitude de rendre publics les résultats de leurs enquêtes, il en est tout autre pour les opérateurs de systèmes de transport à la demande qui ne rendent que très rarement leurs données publiques (Ngo, 2015). Le sondage mené par l'Université de la Californie à Berkeley est donc très intéressant, car c'est une des rares sources de données s'intéressant exclusivement aux utilisateurs de services de mobilité à la demande (Rayle et al., 2016).

Tableau 2-1 Tableau comparatif des enquêtes menées auprès des utilisateurs du taxi et de service de mobilité à la demande

| Commanditaire | Année de l'enquête | Localisation | Sujet principal de l'enquête | Méthode d'enquête | Nombre de répondants |
|---|---------------------------|----------------------------|--|--|-----------------------------|
| NYCTLC | 2011 | New York (États-Unis) | Utilisation du taxi | - | 22 613 |
| NYCTLC | 2012-2013 | New York (États-Unis) | Profil socio-économique des utilisateurs | Enquête à bord des véhicules | - |
| SFMTA | 2013 | San Francisco (États-Unis) | Utilisation du taxi | Entretiens téléphoniques | 649 |
| | | | | Entrevues à l'aéroport | 1383 |
| | | | | Entrevues dans les lieux touristiques | 365 |
| University of California, Berkeley | 2014 | San Francisco (États-Unis) | Utilisation des services de mobilité à la demande | Entrevues dans les lieux de forte activité | 380 |
| RACQ | 2015 | Queensland (Australie) | Utilisation du taxi et des services de mobilité à la demande | Enquête web | 1009 |
| NCST | 2015 | Californie (États-Unis) | Mobilité des milléniaux | Enquête web | 2400 |
| IPART | 2016 | Sydney (Australie) | Utilisation du taxi | Enquête web | 2776 |
| Commission européenne | 2016 | Capitales européennes | Utilisation du taxi et des services de mobilité à la demande | Enquête web | 510 |

Ces enquêtes s'articulent autour de questions pouvant être regroupées en plusieurs catégories. Quatre catégories de questions reviennent régulièrement dans les enquêtes recensées. Un des objectifs premiers de ces initiatives est de connaître les utilisateurs. Ainsi, la plupart des enquêtes mentionnées intègrent des questions permettant de construire le profil socio-économique des répondants (Circella et al., 2017; Frazzani et al., 2016; Hara Associates Inc. et Corey, Canapary & Galanis, 2013; NYC Taxi & Limousine Commission, 2011, 2014; RACQ Public Policy Department, 2016; Rayle et al., 2016). Ensuite, une autre partie des questions permet d'explorer l'utilisation que les répondants font du taxi ainsi que les facteurs qui influencent cette utilisation (Circella et al., 2017; Frazzani et al., 2016; Hara Associates Inc. et Corey, Canapary & Galanis, 2013; IPART, 2017; NYC Taxi & Limousine Commission, 2011; RACQ Public Policy Department, 2016). Pour des opérateurs de l'industrie du taxi, il est également très important de connaître la satisfaction des clients vis-à-vis des services offerts. Certaines de ces enquêtes explorent donc cette problématique (Frazzani et al., 2016; Hara Associates Inc. et Corey, Canapary & Galanis, 2013; IPART, 2017; NYC Taxi & Limousine Commission, 2011; RACQ Public Policy Department, 2016). Enfin, les questions de concurrence avec les services de mobilité à la demande sont également un sujet d'importance dans ces enquêtes (Circella et al., 2017; Frazzani et al., 2016; IPART, 2017; RACQ Public Policy Department, 2016; Rayle et al., 2016).

2.4.2.1 Le profil socio-économique des répondants

Même si les questions posées dans les différents sondages sont semblables (âge, genre, revenu, accès à l'automobile, etc.), il est très délicat de comparer les réponses car l'influence du contexte géographique est très forte. Il est cependant possible de dresser plusieurs constats concernant le profil des utilisateurs du taxi. Le premier fait marquant est que pour les études dont la sélection des répondants n'est pas faite à partir d'un échantillon représentatif de la population dans son ensemble (Circella et al., 2017; Frazzani et al., 2016; NYC Taxi & Limousine Commission, 2011, 2014), la part des hommes dans les répondants est systématiquement plus importante. Dans le rapport publié par la Commission européenne par exemple, 61 % des répondants se sont identifiés comme des hommes et 39 % comme des femmes. Il est plus difficile de comparer directement les revenus des répondants, car cela dépend du contexte dans lequel le sondage a été mené. Un revenu annuel de 80 000 \$ (US) par foyer peut ne pas être très élevé à San Francisco, mais être relativement conséquent à Bratislava. Les sondages menés à Sydney (IPART) et à New York en 2011 semblent

cependant montrer que les populations à hauts revenus sont plus susceptibles d'utiliser fréquemment le taxi (IPART, 2017; NYC Taxi & Limousine Commission, 2011).

2.4.2.2 L'utilisation du taxi

À l'exception de New York, où en 2011 plus de la moitié des répondants utilisaient le taxi une fois par semaine et plus, la fréquence d'utilisation du taxi reste faible. Dans le sondage de l'Union Européenne, plus de la moitié des répondants utilisent le taxi à une fréquence mensuelle ou plus faible. Des constats similaires sont également observés dans les autres sondages. Les répondants sont également interrogés sur la manière dont ils commandent généralement le taxi (héler directement, appel à un centre de répartition, station de taxis, utilisation d'application mobile, etc.). Les moyens utilisés pour commander un taxi varient en fonction du contexte. Cela n'est pas surprenant, car ces différentes utilisations du taxi représentent des segments de marché différents (IPART, 2017; NYC Taxi & Limousine Commission, 2011).

2.4.2.3 La satisfaction envers le service

Connaître le niveau de satisfaction des clients est important pour un opérateur et les sondages sont un outil adapté à cette tâche. En plus que de simplement chercher à connaître la satisfaction de clients, des pistes d'amélioration du service sont également étudiées à travers ces sondages. Le sondage mené à New York en 2011 interroge par exemple les répondants sur les fonctionnalités qu'ils désireraient voir dans les futurs véhicules des flottes de taxi de la ville. Cependant, de même que pour les profils des utilisateurs, ces questions dépendent fortement du contexte géographique de l'étude et des politiques mises en place sur le territoire concerné. Dans certaines villes, les utilisateurs n'expriment pas d'insatisfaction marquée concernant le service qui leur est offert (Hara Associates Inc. et Corey, Canary & Galanis, 2013; NYC Taxi & Limousine Commission, 2011). Dans d'autres contextes en revanche, la qualité des systèmes en transport n'est pas suffisante pour les utilisateurs (IPART, 2017). La question du prix d'une course de taxi en est une des rares questions qui fait l'unanimité dans tous les sondages. Dans l'ensemble, les répondants estiment que la tarification des courses de taxi est trop élevée.

2.4.2.4 Comparaison avec l'utilisation des services de mobilité à la demande

Comme il a été introduit précédemment, l'introduction récente de ces nouveaux acteurs dans le secteur du transport de personnes par automobile a fortement impacté l'industrie du taxi. Certains

commanditaires de ces enquêtes ont donc profité de cette occasion pour mieux comprendre les interactions entre ces deux systèmes de transport. Là encore, il est très difficile de comparer directement les résultats de l'ensemble des sondages, car la dépendance géographique est également très forte. Si à Amsterdam seulement 19 % des répondants au sondage de la Commission européenne utilisent à la fois des services de taxis et des services de mobilité à la demande, à Paris ou à Stockholm, ce chiffre est de 47 % d'après ce même sondage. La perception de la tarification des courses effectuées par le biais de ces nouveaux intermédiaires est meilleure que celle pour l'industrie du taxi.

2.5 Classification comportementale

Dans les autres modes de transport, il est également fréquent d'étudier les usagers en fonction de leur utilisation de ces modes. Ainsi des méthodes de classification spécifiques ont été développées pour créer des portraits types d'utilisateur et d'utilisation (Ghaemi, Agard, Trépanier et Nia, 2017; Vogel et al., 2014). Une des conditions nécessaires pour effectuer ce type d'analyse est de disposer de données permettant de relier un utilisateur aux trajets qu'il a effectués par ce mode. C'est notamment le cas dans les réseaux de transport en commun équipés d'un système de validation par carte à puce. Chaque validation est alors reliée au numéro de carte de l'usager (Agard, Nia et Trépanier, 2013). Plus généralement, ces données sont disponibles dans tous les systèmes de transport nécessitant l'identification du client pour l'utiliser. Les enregistrements transactionnels de ce système sont alors reliés au numéro de la carte à puce utilisée par le client (Morency, Trepanier, Agard, Martin et Quashie, 2007). Des analyses de ce type pour le taxi ne sont pas présentes dans la littérature en raison de l'absence ou de la difficulté d'accès à ce type de données.

Cette méthodologie d'analyse de données permet de créer des groupes d'usage type. Tous les utilisateurs au sein d'un même groupe possèdent des caractéristiques d'utilisation similaires. Ces analyses peuvent également être utilisées pour améliorer la compréhension de l'utilisation d'un réseau de transport (El Mahrsi, Come, Oukhellou et Verleysen, 2017). Les attributs utilisés pour construire ces groupes varient. Ils peuvent être la fréquence d'utilisation, la répartition temporelle des trajets effectués, la répartition spatiale des points d'origine et/ou de destination, etc. La période d'analyse peut elle aussi varier en fonction des objectifs du projet de recherche et des données disponibles. Il est également possible d'effectuer cette méthodologie sur plusieurs périodes

consécutives et d'évaluer l'évolution des comportements en comparant les résultats obtenus sur chacune des périodes (Deschaintres, 2018).

CHAPITRE 3 MONTRÉAL ET LE FONCTIONNEMENT PARTICULIER DE TÉO TAXI

Dans ce chapitre, nous allons tout d’abord présenter très brièvement le portrait général de la mobilité montréalaise. Par la suite, nous présenterons plus en détail le fonctionnement du système de taxi dans la région de Montréal. Une analyse descriptive des données provenant de l’opérateur de taxi montréalais Téo Taxi est également faite dans ce chapitre. Pour toute la suite du mémoire, les analyses seront faites sur les données nettoyées. Ce chapitre permet d’ancrer ce mémoire dans le contexte de Montréal et de mieux comprendre les spécificités que le cadre géographique peut apporter dans la suite des analyses. Comme nous l’avons vu dans le Chapitre 2, l’étude d’un système de taxi est indissociable de son ancrage territorial.

3.1 La mobilité montréalaise

3.1.1 Le territoire

Le territoire de la communauté métropolitaine de Montréal (CMM) permet de décrire la région urbaine de Montréal. En 2019, il comprend 82 municipalités réparties sur une superficie de 3837,15 km². Sa population est de 3 979 402 habitants (Mamh, 2019). La Figure 3-1 présente les limites géographiques de ce territoire.

L’île de Montréal compte 1 999 883 habitants, la moitié de la population de la CMM et a une superficie de 497,05 km² (Mamh, 2019). L’île de Montréal est séparée du reste du territoire par le fleuve Saint-Laurent au Sud et à l’Est, le lac des Deux Montagnes à l’Ouest et la rivière des Prairies au Nord.

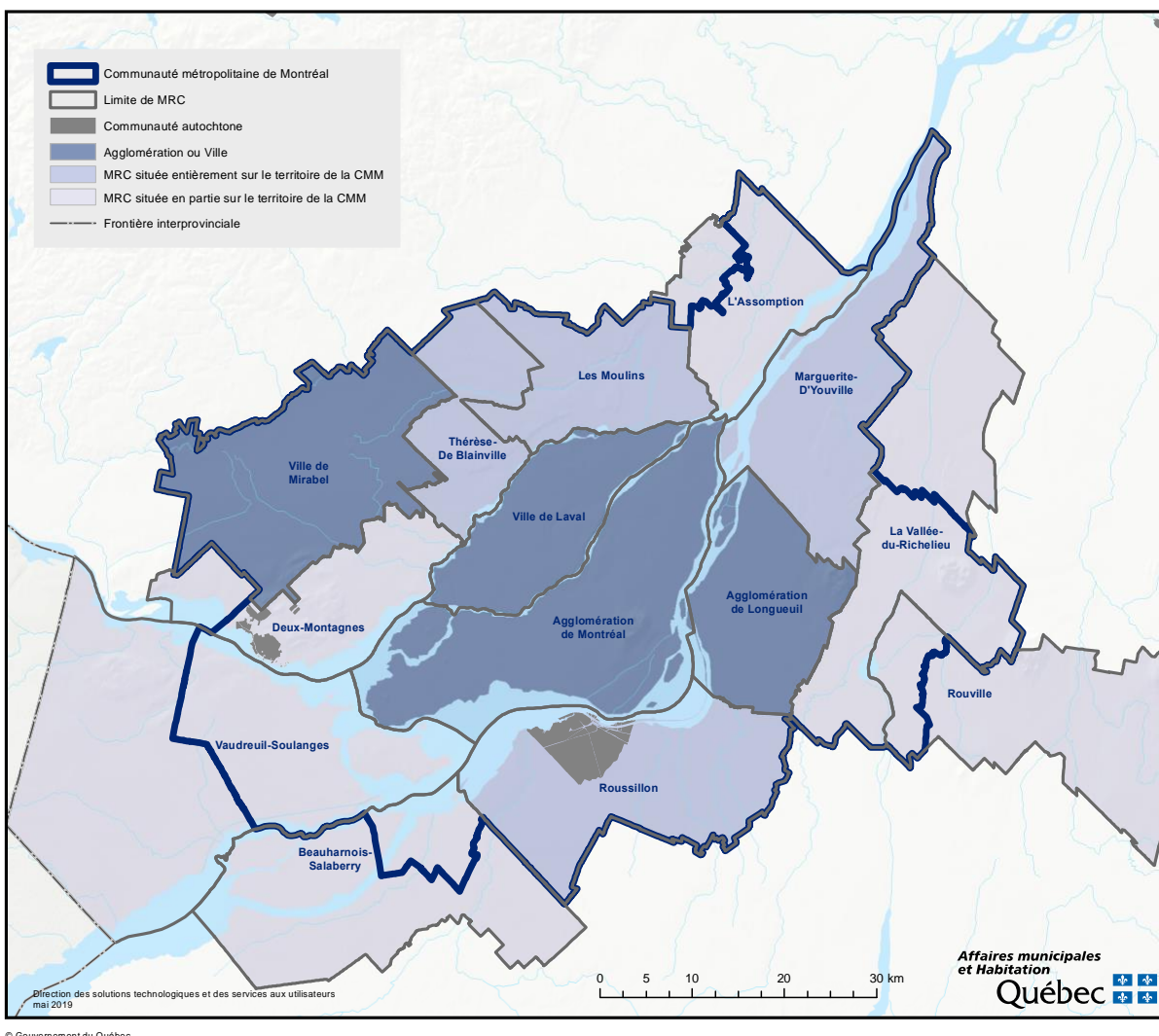


Figure 3-1 Carte de la communauté métropolitaine de Montréal tirée de (Mamh, 2019)

3.1.2 Le réseau de transport

L'île de Montréal est reliée au reste du territoire urbain par de nombreux liens routiers et autoroutiers. Ces liens sont visibles sur la Figure 3-2. L'aéroport et le centre-ville, qui sont deux zones d'activité fortes pour le taxi, sont également visibles sur cette carte. Il est également possible de se rendre sur l'île de Montréal en transport en commun depuis les autres agglomérations de la CMM. Nous pouvons voir sur la Figure 3-3 les deux lignes de métro qui permettent de se rendre à Montréal (ligne orange et ligne jaune) ainsi que les six lignes de trains de banlieue qui desservent l'île. Des lignes de bus permettent également de traverser le fleuve Saint-Laurent et la rivière des Prairies.



Figure 3-2 Carte du réseau routier principal de l'île de Montréal tirée de (Lacombe, 2016)

La Société de transport de Montréal (STM) est l'opérateur de transport en commun sur l'île de Montréal. Le territoire de desserte de la STM est de 500 km² (Société de Transport de Montréal, 2019). La STM opère 68 stations de métro réparties sur quatre lignes (Figure 3-3) et plus de 200 lignes de bus (Société de Transport de Montréal, 2018b). La coordination et l'organisation des différents acteurs du transport collectif dans la CMM sont gérées par l'Autorité régionale des transports métropolitains (ARTM).



Figure 3-3 Plan du réseau de métro opéré par la STM tiré de (Société de Transport de Montréal, 2018a)

3.1.3 La mobilité sur le territoire

Afin de mieux connaître la mobilité des Montréalais, l'ARTM organise régulièrement des Enquêtes Origine-Destination auprès des ménages de la CMM. Ces enquêtes, réalisées aux cinq ans, permettent de caractériser les déplacements effectués lors d'un jour moyen de semaine. Dans la suite de ce chapitre, les résultats de l'enquête O-D 2013 sont utilisés (Exo, 2015).

En 2013, les résidents habitant dans la zone de couverture de l'enquête effectuent 9 392 000 déplacements par jours, dont 27 % pendant la période de pointe du matin. Cela représente une moyenne de 2,2 déplacements par personne. La Figure 3-4 montre les motifs de ces déplacements sur l'ensemble de la journée. Après le motif « Retour au domicile » qui représente 42,7 % des déplacements produits, le motif le plus fréquent est « Travail » qui représente 18,5 % des déplacements produits.

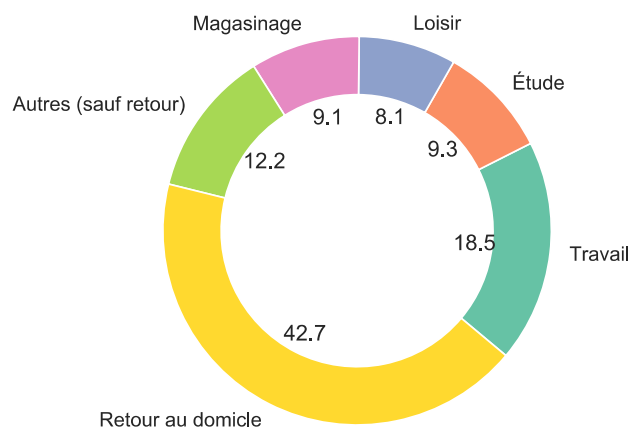


Figure 3-4 Part des déplacements produits selon les motifs sur l'ensemble du territoire de l'enquête (chiffres de l'enquête O-D 2013 version 13.2b)

Le premier mode utilisé est l'automobile (conducteur et passager). Nous pouvons voir sur la Figure 3-5 que 67,8 % des déplacements réalisés par l'ensemble des habitants du territoire de l'enquête ont été effectués avec ce mode. Le faible nombre d'observations concernant le taxi dans cette enquête (1 346 observations) limite les analyses possibles sur ce mode à partir de l'enquête. Le taxi est généralement associé à la catégorie « Autre » en raison de ce faible nombre d'observations. La part modale du taxi est donc inférieure à 1,1 %. Cependant comme il a été discuté dans le Chapitre 2, le faible nombre de déplacements réalisés en taxi ne signifie pas qu'il ne joue pas un rôle important dans le réseau de transport de la région.

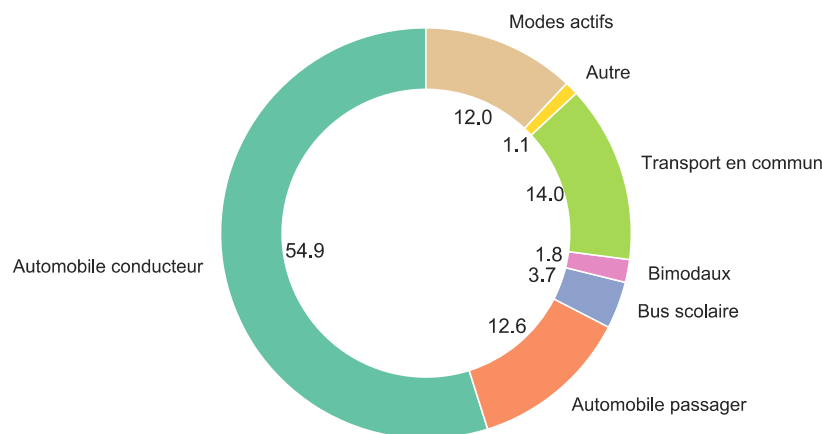


Figure 3-5 Part des déplacements selon les modes sur l'ensemble des trajets effectués par les habitants du territoire de l'enquête (chiffres l'enquête O-D 2013)

3.2 Le taxi à Montréal

Depuis sa création en 1987, le Bureau de taxi de Montréal (BTM) est l'autorité responsable de l'industrie de transport par taxi et de l'application du chapitre de loi S-6.01 *Loi concernant le transport par taxi* (Ville de Montréal, 2019b) sur le territoire de l'île de Montréal. Le BTM s'occupe entre autres de vérifier la qualité du service offert par les chauffeurs, de délivrer des permis dans les trois agglomérations de l'île et de définir la politique de tarification en vigueur.

La carte de la Figure 3-6 représente les limites géographiques des trois agglomérations de taxi du territoire du BTM (A.12 à l'Ouest, A.11 au centre, et A.5 à l'Est) ainsi que la répartition du nombre de permis sur ces trois zones. Le détenteur d'un permis peut effectuer toutes les courses dont l'origine se situe dans les limites géographiques de la zone et ce peu importe la localisation du lieu de destination ou le mode de réservation de la course. Un chauffeur de taxi est également autorisé à effectuer des courses dont l'origine se situe à l'extérieur de la zone de son permis à condition que cette course ne soit pas hélée dans la rue et que la destination de la course soit à l'intérieur de la zone de desserte de son permis.

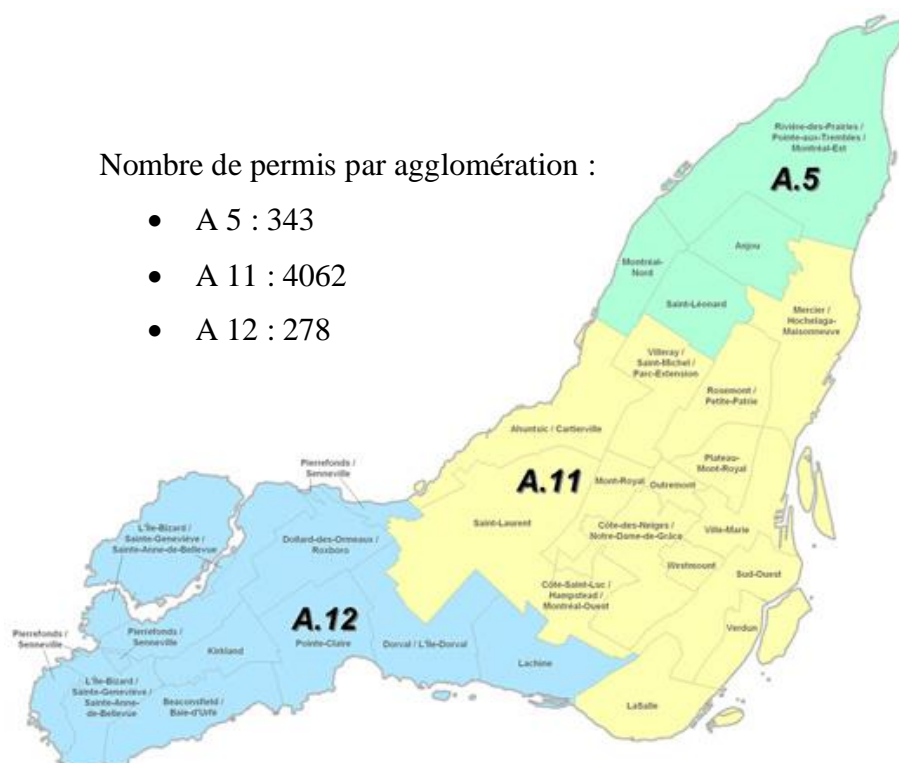


Figure 3-6 Carte des trois agglomérations de taxi sur l'île de Montréal tirée de (Ville de Montréal, 2019a)

La tarification en vigueur à Montréal est une fonction de la durée et de la distance de la course. La tarification précise est décrite sur la Figure 3-7. Les chauffeurs de taxi ont l'obligation d'accepter les cartes de débit et crédit comme moyen de paiement. La tarification depuis et vers l'aéroport est différente puisqu'elle fait d'une part l'objet d'un tarif fixe depuis le centre-ville de 41\$ et d'autre part d'un prix minimum de 17,45 \$ pour les autres zones de desserte.

La zone aéroportuaire fait également l'objet d'autres spécificités. En effet, pour pouvoir effectuer des courses ayant pour origine l'aéroport de Montréal, le chauffeur doit disposer d'un permis de taxi spécifique pour cette zone. Depuis 2018, ces permis sont au nombre de 350 répartis sur les trois territoires (A.11, A.12, A.5). La flotte est composée de 250 véhicules dits verts (empâtement minimal de 271 cm et capacité du coffre minimal de 410 litres) et 100 minivans (capacité minimum de six personnes) (Réseau québécois des aéroports, 2017). Pour obtenir un permis, un chauffeur doit effectuer une demande auprès de l'aéroport de Montréal. En 2018, 2500 chauffeurs ont déposé une demande. Les 350 permis aéroports ont été distribués par tirage au sort.

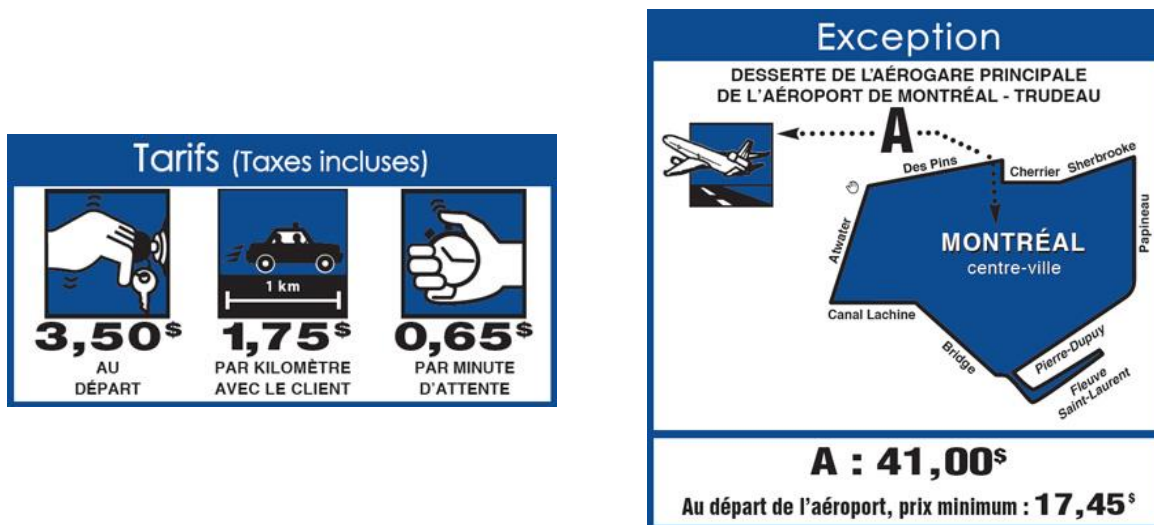


Figure 3-7 Tarification des courses tirée de (Ville de Montréal, 2019c)

3.3 L'expérimentation de Téo Taxi

Afin de favoriser l'innovation entre autres dans le secteur de transport par taxi, le ministre des Transports est autorisé à mettre en œuvre des projets pilotes. En 2019, le nombre de projets pilotes de ce secteur est de cinq (Commission des transports Québec, 2019). La création de Téo Taxi a été rendue possible grâce à un de ces projets pilotes. L'entreprise Téo Taxi a été créée en 2015 et a fermé ses portes en janvier 2019. Suite à cela, Téo Taxi a été racheté. Au moment de la rédaction de ce mémoire, les activités de Téo Taxi n'avaient pas encore été relancées.

3.3.1 Présentation de l'entreprise

L'entreprise Téo Taxi a débuté ses activités en novembre 2015. Dès sa création, il a été possible pour Téo Taxi, grâce à la flexibilité offerte par le projet de loi, de proposer à ses utilisateurs un service différent de celui des acteurs traditionnels du secteur. Trois différences opérationnelles majeures distinguent Téo Taxi du reste de l'industrie. Premièrement la flotte de véhicules de Téo était composée intégralement de véhicules électriques. Début 2019, la flotte de Téo comprenait 192 véhicules (Taxelco, 2019). Cette situation implique des contraintes opérationnelles fortes puisque le temps de recharge d'un véhicule électrique est relativement long et ne peut être effectué que dans certains points spécifiques (Geneste, 2017). Le moment, le lieu et la durée de la recharge doivent donc être anticipés. Ensuite, les chauffeurs de Téo Taxi n'étaient pas des travailleurs

indépendants, mais des employés de Téo Taxi. Téo Taxi garantissait un salaire brut minimum de 15 \$(CA) de l'heure dollar à ses chauffeurs, et la rémunération de beaucoup d'entre eux était supérieure à ce seuil grâce aux pourboires distribués par les clients. À la fin de la période d'activité du service, le salaire horaire moyen des chauffeurs était de 16,15 \$ (Taxelco, 2019). Enfin, pour emprunter un Téo Taxi, les utilisateurs avaient trois solutions : hélér directement un taxi de l'entreprise dans la rue, commander une course par le biais d'une application, commander une course par le biais d'un site internet. Une capture d'écran de l'interface utilisateur de l'application mobile est visible sur la Figure 3-8. Pour commander une course par l'application, l'utilisateur doit préalablement créer un compte utilisateur et fournir ses informations bancaires pour utiliser les options de paiement intégrées. Lors de la création de son compte, le client se voyait également crédité d'un rabais de 10\$ sur sa première course. Il était aussi possible pour le client de spécifier sur cette application le type de véhicule qu'il souhaitait utiliser (véhicule de la marque Tesla ou non) ainsi que de commander des véhicules appartenant à des flottes de taxi affiliées à Téo Taxi.

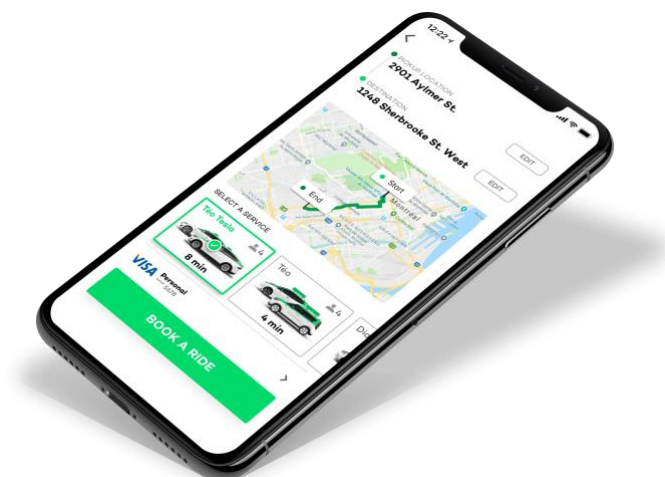


Figure 3-8 Capture d'écran de l'application mobile de Téo Taxi tirée de (Taxelco, 2018b)

L'utilisation de l'application est comparable à celle d'acteurs de la mobilité à la demande comme Uber ou Lyft. Tout d'abord, l'utilisateur doit définir le point d'origine et de destination de la course. L'opérateur (Téo Taxi) lui indique ensuite une estimation du temps d'attente avant l'arrivée du taxi. Lorsque la course est terminée, le paiement est automatiquement effectué via l'application. Le client reçoit également la facture de la transaction par courriel. Puisque Téo Taxi est un service de transport par taxi, la tarification utilisée est celle définie par le BTM (Figure 3-7). Pour utiliser Téo Taxi en le hélant directement, le client n'a pas besoin de posséder l'application ou de compte

client. La transaction est directement effectuée dans le taxi via un terminal de paiement présent à bord du véhicule. Cette utilisation de Téo Taxi n'était pas disponible au début du lancement du Projet Téo Taxi.

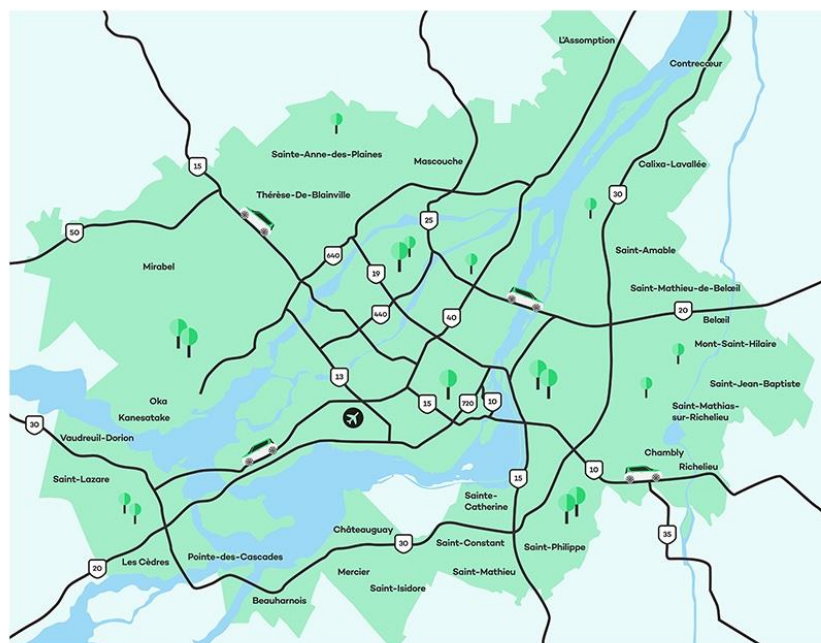


Figure 3-9 Zone de couverture du service Téo Taxi à la fin de son activité tirée de (Taxelco, 2018a)

À la fin de ses activités, la zone de desserte couvrait une grande partie du territoire de la CMM et s'étendait au-delà du territoire du BTM. Cependant, la réglementation en vigueur ne permet pas à Téo Taxi d'opérer l'intégralité des courses sur cette zone de desserte, certaines conditions doivent être respectées. Téo Taxi possédait des permis rattachés à la zone A.11. Pour respecter la réglementation en vigueur, il faut que l'origine et/ou la destination de la course se situe dans l'agglomération de taxi A.11 (Figure 3-6). Dans le cas contraire, l'entreprise ne peut pas réaliser la course. Pour les mêmes raisons réglementaires, aucune course hélée dont le point d'origine est externe à l'agglomération A.11 ne peut être opérée par Téo Taxi. Toutes les courses dont le point de départ se situe en dehors de la zone A.11 doivent être réservées par l'application.

3.3.2 Bilan du projet Téo Taxi

Si le modèle proposé par Téo Taxi était novateur, les nombreuses contraintes opérationnelles qu'il impliquait ont rendu l'atteinte d'un équilibre économique très complexe. Après trois ans d'opération, la société Téo Taxi a mis fin à ses activités à la fin janvier 2019. À la suite de la

fermeture du service, Téo Taxi a rendu public un mémoire présentant un retour d'expérience sur les trois ans d'activité du service (Taxelco, 2019). Certains éléments du modèle spécifique de Téo sont un succès. La marque et sa qualité de service étaient reconnues par ses utilisateurs. Le modèle de salariat des chauffeurs également permis aux employés de Téo de bénéficier d'une plus grande protection sociale que leurs confrères qui ont un statut de travailleur indépendant. Cependant, Téo Taxi identifie également d'autres problématiques qui ont rendu l'opération de Téo beaucoup plus complexe. Tout d'abord, le choix d'utiliser une flotte de taxis électriques est relativement onéreux à l'achat d'une part et à l'exploitation d'autre part. De plus, la performance de ces véhicules peut être fortement impactée lors des épisodes de grand froid que peut connaître Montréal en hiver. Le système de rémunération des chauffeurs implique également que ces derniers doivent avoir une productivité (minimisation du temps sans client à bord) plus importante que celle des acteurs traditionnels.

Dans ce rapport, une analyse des éléments de l'écosystème du transport par taxi qui ont limité le développement de Téo Taxi est également présentée. Le cadre législatif actuel au Québec ne permet pas de valoriser l'implantation d'un service de taxi novateur proposant un service différent. La structure des permis de taxi mis en place par le BTM est relativement rigide et constitue un handicap d'une part pour opérer des véhicules électriques et hybrides, mais également pour coordonner une équipe de chauffeurs salariés d'autre part.

CHAPITRE 4 CADRE MÉTHODOLOGIQUE DU PROJET ET DESCRIPTION DES DONNÉES

4.1 Les grandes étapes méthodologiques

Dans cette partie, nous présentons la méthodologie globale de ce projet. Sur le schéma méthodologique de ce projet (Figure 4-1), nous pouvons distinguer deux parties distinctes. D'un côté, nous avons développé une méthodologie permettant de construire des typologies d'utilisation de Téo Taxi. Cette partie du projet est détaillée dans le Chapitre 5. D'un autre côté, nous avons souhaité approfondir la compréhension des facteurs influençant l'utilisation du taxi et la description du profil sociodémographique des utilisateurs ainsi que leurs habitudes de mobilité générales grâce aux résultats d'un sondage qui a été conduit auprès des clients de Téo Taxi. Le Chapitre 6 détaille cette partie de la recherche. Ces deux parties ne sont pas indépendantes, au contraire, elles sont complémentaires. Des analyses croisées présentées à la fin du mémoire de ces deux parties nous ont permis de mettre en évidence un certain nombre de constats.

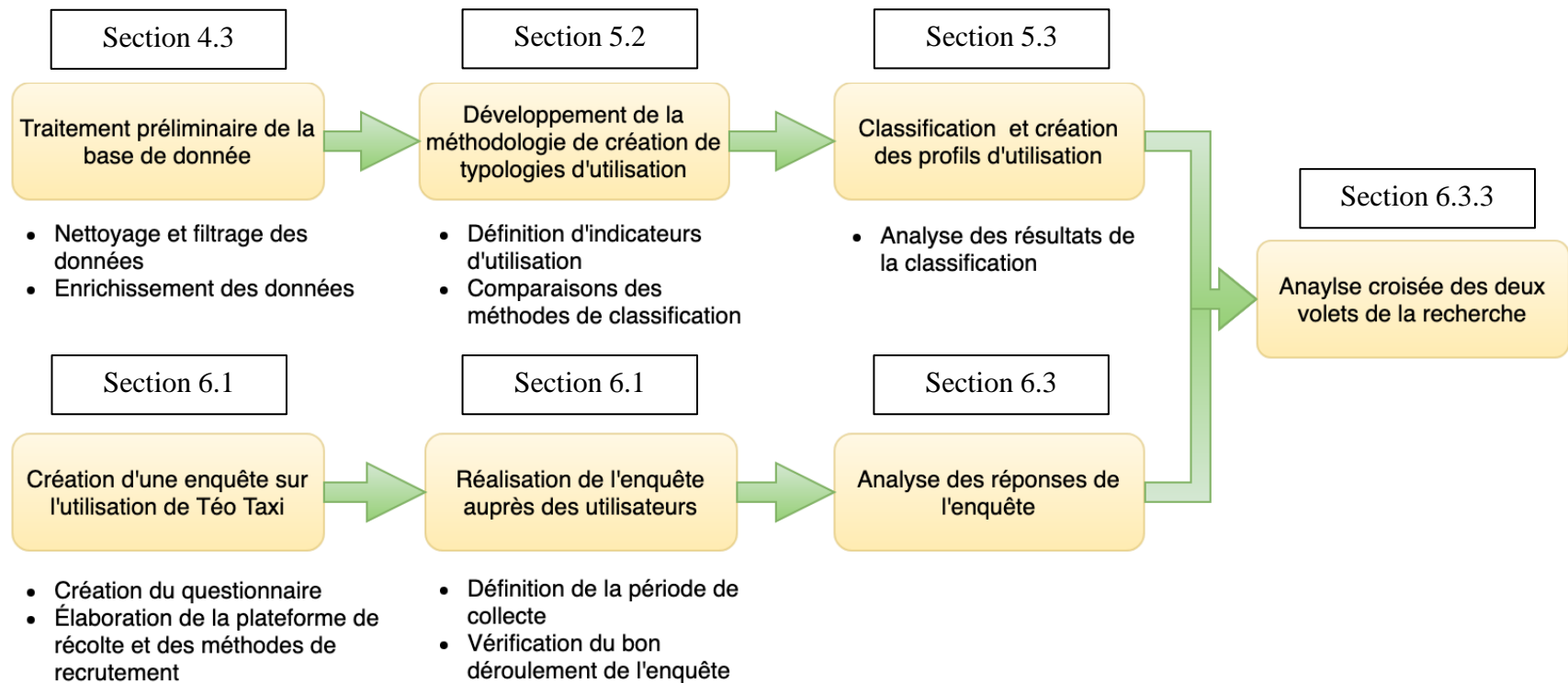


Figure 4-1 Schéma méthodologique du projet de recherche

4.1.1 Construction de typologies d'utilisation

4.1.1.1 Contraintes méthodologiques

Pour construire des typologies d'utilisation de Téo Taxi et étudier le comportement d'usage des clients, nous avons dû faire face à plusieurs défis méthodologiques.

Tout d'abord, la fréquence d'utilisation du taxi est très faible comparée aux autres modes pour lesquels des méthodologies similaires ont été développées. Sur l'ensemble des données (période allant du 5 avril 2016 au 28 janvier 2019), le nombre de courses moyen par utilisateur est de 8. À titre comparatif, le nombre de validations moyen par détenteur d'une carte OPUS durant l'année 2016 dans le métro uniquement est de plus de 200 (Deschaintres, 2018). Bien que les informations dont nous disposons sur une course soient plus précises (temps de déplacement, lieux d'arrivée, lieux de départ, etc.) que pour la plupart des autres modes de transport urbain, l'utilisation générale du taxi diffère trop de celle de ces modes pour appliquer directement les techniques d'analyse de données précédemment développées pour caractériser les utilisateurs.

De plus, la variabilité d'utilisation du taxi est très forte au sein des usagers. Alors que certains usagers ont effectué plus de 1300 courses sur leur durée de vie (période entre leur première et dernière utilisation de Téo Taxi), une grande partie des utilisateurs utilise Téo Taxi de manière très occasionnelle. Plus de 37,8 % des utilisateurs de la base de données n'ont effectué qu'une seule course avec le système. La méthodologie retenue doit donc également pouvoir rendre compte de cette diversité d'utilisation.

Il faut également noter que le service Téo Taxi a connu une très forte croissance tout au long de son activité. Si dans les débuts de son activité, le nombre de courses effectuées par mois est entre 10 000 et 20 000, sur la fin de sa période d'activité Téo opérait plus de 50 000 courses par mois. Cette contrainte implique donc que la méthodologie puisse prendre en compte les modifications d'utilisation liées à cette croissance.

Enfin lorsque cette méthodologie a été construite, le service Téo Taxi était encore en activité. Nous avons fait en sorte que la méthodologie développée puisse régulièrement être répliquée en prenant en considération les nouveaux enregistrements provenant des dernières courses réalisées. La méthode doit donc pouvoir traiter toutes périodes d'analyse.

4.1.1.2 Choix méthodologiques

La solution méthodologique retenue pour faire face à ces nombreuses contraintes a été de ne pas étudier les usagers sur l'ensemble de leur durée de vie, mais plutôt sur une succession de périodes consécutives et comparables. Le comportement d'un utilisateur sur sa durée de vie peut être étudié en analysant l'union de son comportement sur chacune des périodes. La justification des choix concernant les caractéristiques de ces périodes est détaillée dans le Chapitre 5. Dans la suite de ce mémoire, le terme d'usager-période sera utilisé pour définir le comportement d'un usager sur une période. Une définition précise d'un usager-période est proposée dans la section 5.2.2.

4.1.2 Compréhension des facteurs influençant l'utilisation du taxi

Pour cette partie du mémoire, nous avons conduit un sondage auprès des utilisateurs de Téo Taxi. Ce sondage a été réalisé sur une semaine de décembre 2018. Les modalités méthodologiques de la mise en place de ce sondage sont détaillées dans le Chapitre 6. En concevant le sondage, nous avons fait en sorte que chaque réponse puisse être associée à l'enregistrement d'un déplacement spécifique et à l'usager qui a réalisé ce déplacement. Cette disposition nous permet de relier chaque réponse avec une course précise et également l'usager qui a effectué la course et répondu à l'enquête.

4.2 Description des données utilisées

4.2.1 Les données provenant de Téo Taxi

Ce projet a été rendu possible grâce au fonctionnement spécifique du service Téo Taxi détaillé dans le Chapitre 3 et la structure particulière de leur base de données. Plus spécifiquement, nous avons utilisé les données de l'entreprise relatives aux courses de taxi réalisées par leurs véhicules et leurs chauffeurs sur l'ensemble de la durée de vie du service (de novembre 2015 à janvier 2019). Nous disposons d'un enregistrement pour chaque course qui a été commandée par l'application ou hélée dans la rue, ainsi que des caractéristiques de ces courses. Pour chaque enregistrement, nous disposons normalement des attributs suivants :

- L'identifiant unique de la course
- L'identifiant unique du client (lorsque la course a été commandée par l'application)

- L'heure de départ de la course
- L'adresse de départ de la course
- L'heure de fin de la course
- L'adresse de destination de la course
- La durée de la course
- La distance parcourue
- La vitesse moyenne de la course
- Un champ informant si la course a été annulée ou réalisée

Il est important de préciser que nous ne pouvons pas identifier formellement les utilisateurs à partir des données en notre possession. En effet, la seule information dont nous disposons sur l'utilisateur est son numéro d'utilisateur (numéro compris entre 9 et 196 619) et les informations relatives aux courses qu'il a effectuées ou annulées. Nous ne disposons pas non plus d'information sur les modalités de facturation de la course au client ni même sur le prix de la course.

4.2.2 Autres sources de données

Deux jeux de données provenant de la géobase d'Adresses Québec ont également été utilisés dans ce projet : les données AQPro (<https://adressesquebec.gouv.qc.ca/aqpro.asp>) et AQGéobati (<https://adressesquebec.gouv.qc.ca/aqgeobati.asp>) qui est le complément de AQPro. Ces deux jeux de données ont été utilisés pour géolocaliser les points d'origine et de destination des courses.

Les données qui ont été récoltées pendant le sondage réalisé dans le cadre de ce projet ont également été utilisées. Ces données sont décrites dans le Chapitre 6.

4.3 Filtrage préliminaire des données

Dans tous les projets reposant sur le traitement et l'analyse de données, le filtrage des données est une étape cruciale. La qualité des résultats d'un tel projet repose en partie sur la qualité des données (Rahm et Do, s. d.). Utiliser une méthodologie de filtrage adéquat permet donc de maximiser la qualité des futures analyses.

Lors de l'exploration préliminaire des données, nous avons détecté plusieurs anomalies telles que des courses d'une durée très faible (moins d'une minute), des courses d'une distance trop petite ou trop grande (plus de mille kilomètres), ect. Quatre critères ont finalement été retenus pour filtrer les données :

- Durée comprise entre cent secondes et une heure
- Distance comprise entre trois cents mètres et quatre-vingt-cinq kilomètres
- Vitesse positive et inférieure à 34 m.s^{-1}
- Temps à la prochaine course supérieure à quatre-vingt-dix secondes

Comme de nombreux enregistrements ne possèdent pas de renseignement sur la vitesse de la course, il a été choisi de définir ici la vitesse comme la vitesse moyenne de la course c'est-à-dire le rapport entre la distance totale et la durée totale de la course. Le temps à la prochaine course est défini comme la différence de temps entre la fin de la course et le début de la prochaine course effectuée par le même client. Pour qu'un temps à la prochaine course soit défini, il faut donc que le client de la course effectuée une autre course. Dans le cas contraire, cette information est laissée comme non renseignée et n'est pas utilisée comme critère discriminant.

Avant d'être filtrée, la base de données comporte 1 819 286 enregistrements dont 490 599 sont des courses qui ont été annulées. À la suite du processus de filtrage, 1 215 949 courses ont été retenues dont 128 193 ont été hélées dans la rue et 1 087 756 ont été réservées par l'application et son reliés à un numéro de client. Le taux de données écartées est de 8,48 %. Toutes les analyses présentées dans la suite du mémoire ont été effectuées sur les données filtrées.

4.3.1 Définition des critères de filtrage

Comme il a été mentionné précédemment, quatre critères ont été définis pour filtrer le jeu de données brutes. Pour trouver les valeurs seuils de ces filtres, nous nous sommes appuyés sur l'analyse des distributions de ces grandeurs ainsi que sur leur signification physique. Lorsque cela est possible, trouver des discontinuités dans la distribution des grandeurs permet de définir une valeur seuil pour les filtres numériques.

Deux méthodes ont été utilisées pour identifier des points de discontinuité. La première méthode consiste à identifier un point d'inflexion de nombre de courses cumulées en fonction de la grandeur

étudiée. La recherche du point d'inflexion est faite dans la plage de grandeurs potentielles du filtre. D'un point de vue mathématique, un point d'infection permet de caractériser un point sur une courbe où un changement de concavité est observé. D'un point de vue graphique, cela se traduit par le fait que la tangente à la courbe traverse cette dernière en ce point. Une caractéristique du point d'inflexion d'une courbe est que ce point est également un extremum local de la dérivée de cette courbe. La Figure 4-2 illustre ces propriétés en prenant comme exemple la tangente de la fonction $f(x)=\tan(x)$ entre -1 et 1. D'un point de vue de la distribution des données, ce point a également des caractéristiques spécifiques. La dérivée de la distribution cumulée d'une grandeur représente la densité de cette grandeur. Le minimum local de la densité d'une grandeur représente un trou dans la distribution, soit une valeur pour laquelle le nombre ou la proportion d'observations est plus faible que pour les valeurs adjacentes. La deuxième méthode utilisée consiste à repérer les changements de caractéristique de la distribution cumulée des grandeurs comme un changement de tendance.

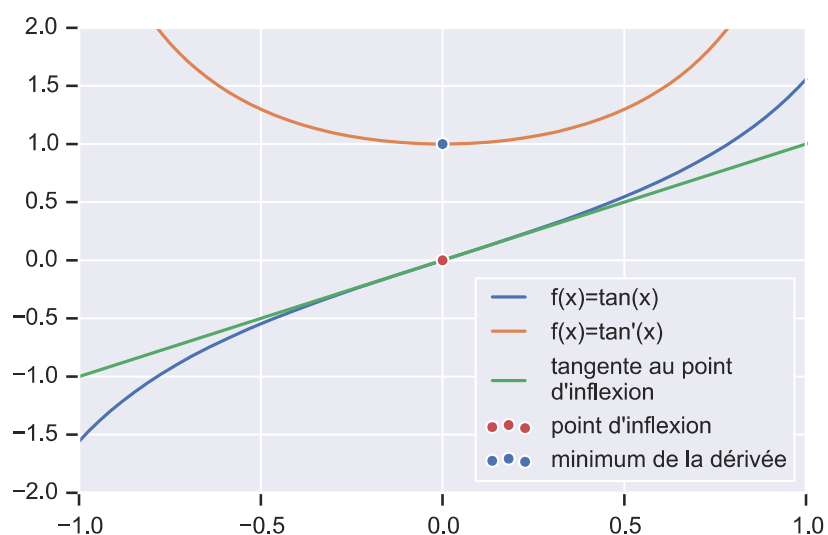


Figure 4-2 Illustration des caractéristiques du point d'inflexion

4.3.2 Durée de la course

La Figure 4-3 montre le début de la distribution cumulée de la durée des courses (entre 0 s et 200 s). On peut apercevoir sur cette courbe un point d'inflexion à 100 s. Cette borne inférieure fait également sens d'un point de vue logique. En effet, il est difficile d'imaginer une course de taxi réelle avec une durée inférieure à 100 s. Pour ce filtre, nous avons également défini une durée

maximale à 3600 s (ou une heure). De même il est très rare qu'une course de taxi soit d'une durée supérieure à une heure. Même si certaines courses peuvent atteindre cette limite, ce phénomène reste exceptionnel dans le cas de forte congestion par exemple. Puisque la grande majorité des courses d'une durée supérieure à une heure sont probablement le fruit d'une erreur, nous avons préféré supprimer toutes les courses de plus d'une heure afin de favoriser la qualité de notre jeu de données. Ce filtre nous permet de supprimer 19 892 courses soit 1,5 % des enregistrements.

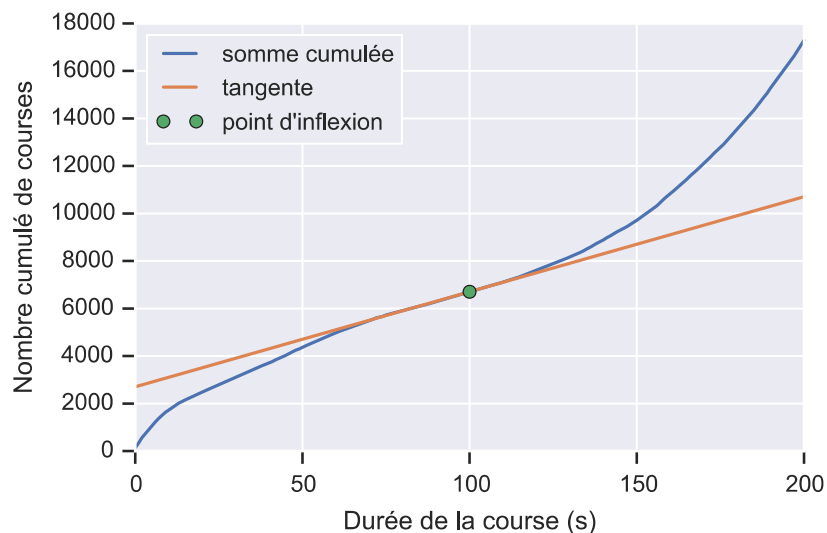


Figure 4-3 Distribution cumulée de la durée des courses pour les enregistrements d'une durée inférieure à 200 s

4.3.3 Distance de la course

La même logique a été suivie pour définir les bornes du filtre sur la distance totale de la course (distance réellement effectuée par le taxi). La Figure 4-4 montre la distribution cumulée des distances totales des courses (entre 0 m et 800 m). Un point d'inflexion est observable à 300 m. Cette borne inférieure a elle aussi un sens concret. Les distances inférieures à 300 mètres sont des distances qui peut être parcourues à pied en moins de cinq minutes pour une grande partie de la population. De même que pour la durée de la course, certaines courses d'une distance inférieure à 300 mètres peuvent ne pas être le fruit d'une erreur, mais cela reste minoritaire. Dans le même principe de maximisation de la qualité des données, nous avons choisi de laisser ce filtre à 300 mètres. La borne supérieure a, quant à elle été fixée à 85 km ce qui correspond à la distance réseau maximale qu'il est possible d'effectuer entre la zone de permis A.11 (Figure 3-6) et l'extrémité de

la zone de desserte de Téo Taxi (Figure 3-9). Ce filtre nous permet d'écarter 107 301 courses soit 8,1 % des enregistrements. Ce nombre est important mais nous avons fait le choix de privilégier la qualité des données étudiées sur la quantité.

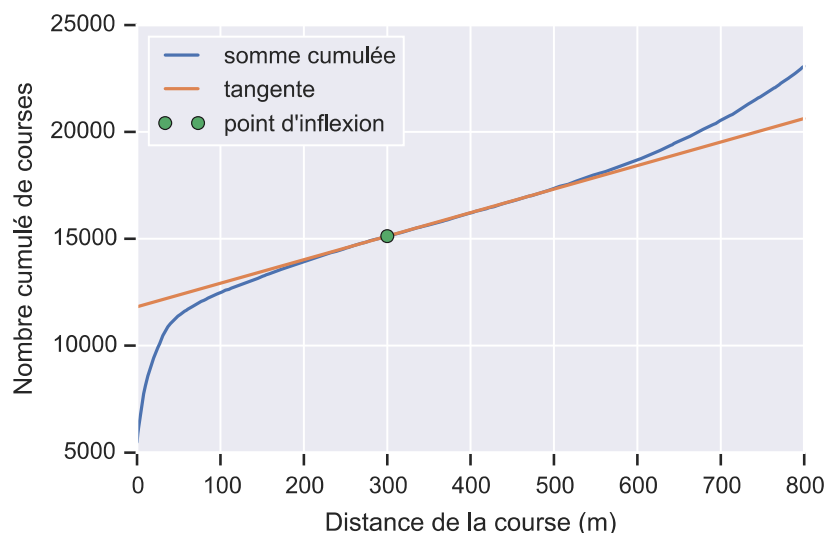


Figure 4-4 Distribution cumulée de la distance des courses en fonction pour les enregistrements d'une distance inférieure à 800 m

4.3.4 Vitesse de la course

Le choix du filtre maximal de vitesse ne provient pas directement de l'analyse de notre jeu de données. Au Québec, la limitation de vitesse est de 100 km.h^{-1} sur les autoroutes même si généralement la vitesse observée dans des conditions de circulation non congestionnée est plus proche de 120 km.h^{-1} . Nous avons donc choisi 34 m.s^{-1} (soit 122 km.h^{-1}) comme valeur pour la vitesse maximale du filtre. La vitesse que nous utilisons étant une vitesse moyenne, de plus en milieu urbain, de la course de taxi, il est impossible que cette dernière dépasse 34 m.s^{-1} . Nous n'avons pas choisi d'appliquer de filtre minimal pour la vitesse, car la vitesse d'une course effectuée dans le centre-ville lorsque la congestion est forte peut être très basse. Nous avons également vérifié que toutes vitesses soient positives. Dans le contexte de cette base de données une vitesse négative signifierait que la date d'arrivée de la course est antérieur à la date de départ ce qui est impossible. Ce filtre nous permet de supprimer 92 558 courses soit 7,0 % des enregistrements.

Le nombre de courses écartées par les filtres de vitesse et de distance est relativement élevé. Cependant, sur l'ensemble de 107 301 données écartées par le filtre de distance, 92 048 ont été mises de côté parce que le champ de distance était vide ou en raison d'une valeur aberrante. Cela représente 85 % des données concernées par le filtre de distance. Ces mêmes 92 048 enregistrements ont également été écartés par le filtre de vitesse pour les mêmes raisons.

4.3.5 Temps intercourse

Comme mentionné précédemment, le temps intercourse représente la durée qui sépare la fin de la course au début de la prochaine course réalisée par le même utilisateur. Ce temps ne peut donc pas être négatif. Cependant, dans les enregistrements, certains clients ont débuté une course avant d'avoir terminé la course précédente. La Figure 4-5 montre la distribution cumulée de ces temps à la prochaine course. On peut observer un changement de tendance à partir de 100 secondes. Entre 100 s et 200 s, la distribution cumulée est très semblable à une distribution linéaire. La borne inférieure de ce filtre a donc été définie à 100 s. Ce filtre ne comporte pas de borne supérieure, car il n'y a pas de limitation logique à cette grandeur. Lorsque qu'un client n'effectue qu'une seule course avec Téo Taxi, il n'est pas possible de définir un temps intercourse. Le filtre ne peut donc pas être appliqué pour ces courses. Ce filtre nous permet de supprimer 3975 courses soit 0,3 % des enregistrements.

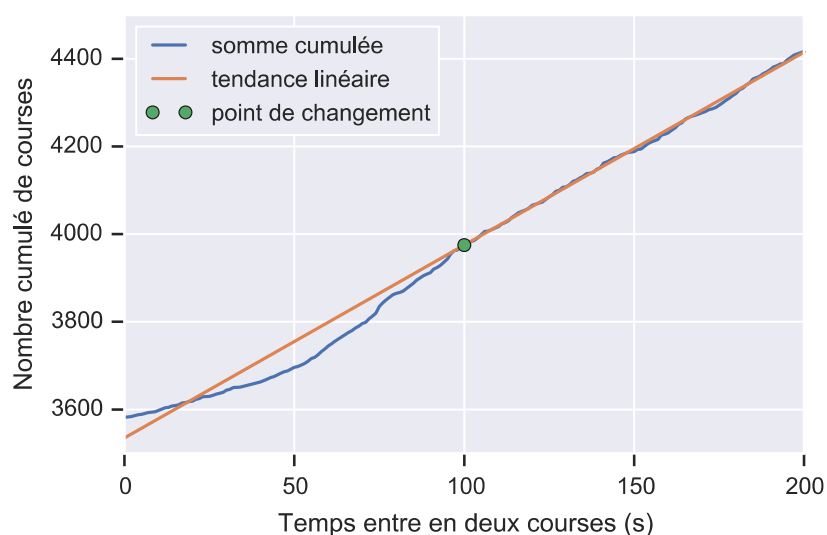


Figure 4-5 Distribution cumulée du temps à la prochaine course pour les enregistrements d'un temps inférieur à 200 s

4.3.6 Synthèse du processus de filtrage

La mise en place de l'ensemble des filtres sur les données brutes a permis d'écartier 112 738 courses sur les 1 328 687 recensées soit 8,5 % des courses. Ce chiffre ne correspond pas à la somme des chiffres présentés pour chacun des filtres car beaucoup de courses ont été filtrées par plusieurs critères (comme la vitesse et la distance). Le nombre de courses retenu est de 1 215 949 courses dont 128 193 ont été héléées et 1 087 756 réservées par l'application. Le tableau résume le nombre de données filtrées pour chacun des critères. Le Tableau 4-1 permet de comparer le nombre d'enregistrements concernés pour chacun des filtres. La somme des nombres d'enregistrements sélectionnés par chaque filtre n'est pas égale au nombre total de données écartées. Dans un premier temps, chacun des filtres a été appliqué de manière indépendante aux données. La condition pour écartier un enregistrement est alors qu'il ne respecte pas au moins l'un des critères de ces quatre filtres.

Tableau 4-1 Tableau comparatif des filtres appliqués aux données brutes

| <i>Filtre</i> | <i>Nombre de courses filtrées</i> | <i>Pourcentage de courses filtrées</i> |
|-----------------------------|-----------------------------------|--|
| <i>Durée</i> | 19 892 | 1,50 % |
| <i>Distance</i> | 107 301 | 8,08 % |
| <i>Vitesse</i> | 92 558 | 6,97 % |
| <i>Temps intercourse</i> | 3975 | 0,30 % |
| <i>Ensemble des filtres</i> | 112 738 | 8,48 % |

La Figure 4-6 permet de voir l'évolution du taux de données filtrées tout au long de la période d'activité de Téo Taxi. Le taux de filtrage des données est de 100 % pour toute la période antérieure à avril 2016, car nous ne disposons pas d'information sur la distance de la course pour ces premiers enregistrements. Le taux de filtrage est par la suite inférieur à 15 %. Une part anormalement élevée de données est filtrée au mois de novembre 2016. Cela est en partie dû à la défaillance du système informatique de Téo sur cette période. La part de donnée écartée pendant ce mois est de 48,3 %. En faisant abstraction de cet évènement, la tendance de la part de données écartées est à la baisse.

La part de données filtrées est inférieure à 10 % à partir de juillet 2017. Pour les deux derniers mois d'activité de Téo Taxi, les taux de données écartées étaient respectivement de 4,4 % et 4,0 %.

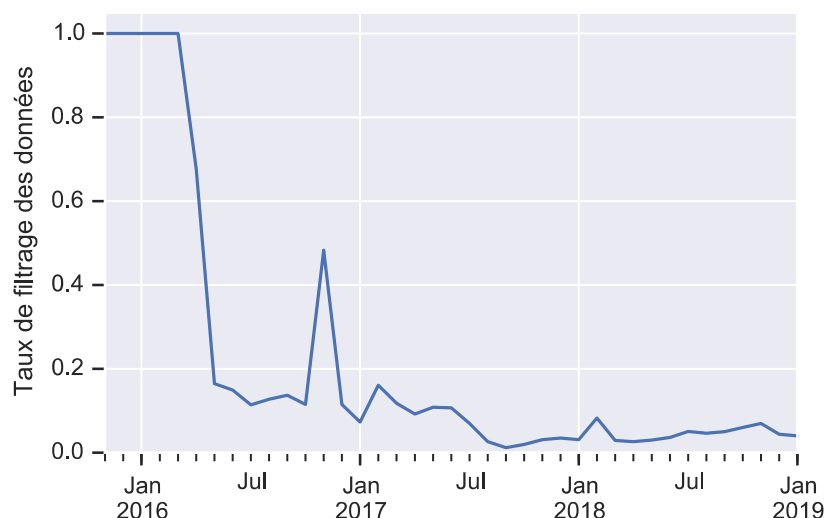


Figure 4-6 Taux de données filtrées par mois sur l'ensemble de la période d'activité de Téo Taxi

4.4 Analyses descriptives de la base de données

Tableau 4-2 Caractéristiques principales du jeu de données

| | |
|--|------------|
| <i>Nombre de courses effectuées après une réservation via l'application mobile</i> | 1 087 756 |
| <i>Nombre de courses hélées (sans informations liées de clients)</i> | 128 193 |
| <i>Nombre total d'enregistrements</i> | 1 215 949 |
| <i>Date du premier enregistrement</i> | 2016-04-11 |
| <i>Date du dernier enregistrement</i> | 2019-01-28 |
| <i>Nombre d'utilisateurs distincts</i> | 134 047 |
| <i>Nombre de courses moyen par utilisateur sur l'ensemble de la période</i> | 8,01 |

Cette partie présente une analyse descriptive des différentes caractéristiques du jeu de données. Le Tableau 4-2 résume la dimension du jeu de données à la suite de l'étape de filtrage. Les analyses

présentées dans cette partie sont réalisées à la fois sur les courses réservées par l'application et celles héliées dans la rue. La date du premier enregistrement est avril 2016 car toutes les courses antérieures à cette date ont été filtrées.

4.4.1 Répartition temporelle des courses

La Figure 4-7 montre le nombre de courses effectuées par mois entre le 11 avril 2016 et le 28 janvier 2019. L'anomalie du mois de novembre 2016, détectée dans la partie précédente est également visible. Au-delà de cette anomalie, le phénomène de croissance de l'entreprise ressort très nettement sur ce graphique. En mai 2016 (premier mois de récolte complet des données), 15 940 courses ont été réalisées. Le mois de novembre 2018 est le mois avec l'utilisation la plus intense avec 58 016 courses réalisées pendant cette période. Le jour d'utilisation maximale sur l'ensemble de la base de données est le vendredi 16 novembre 2018; 2475 courses ont été effectuées ce jour-là. L'utilisation moyenne de Téo Taxi est de 1182 courses par jour.

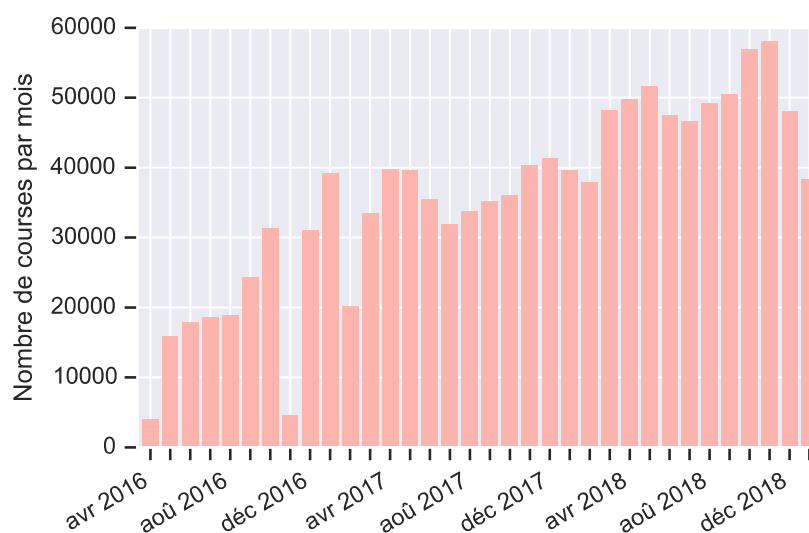


Figure 4-7 Nombre de courses réalisées par mois

La Figure 4-8 représente la répartition moyenne des courses en fonction de l'heure dans une semaine. Pour pouvoir comparer les semaines entre elles, le nombre de courses par heure a été normalisé par le nombre total de courses effectuées au total dans la semaine. Ainsi pour chacune des semaines (représentées en bleu) ou pour la semaine moyenne (représentée en rouge), l'échelle est la même. Cette normalisation des données est nécessaire pour pouvoir comparer les semaines sur l'ensemble de la période d'activité de Téo Taxi. Comme l'illustre la Figure 4-7, la croissance

importante du service fait que les nombres de courses réalisées par semaines diffèrent énormément. En excluant les semaines incomplètes (première semaine et dernière semaine des données) ainsi que le mois de novembre 2016, la semaine avec le moins de courses est la semaine allant du 20 février 2017 au 26 février 2017, 1 340 ont été faites sur ces sept jours. La semaine avec le plus de courses va du 5 novembre 2018 au 11 novembre 2018, avec un total de 13 640 courses. Le rapport d'intensité d'utilisation entre ces deux périodes est supérieur à dix.

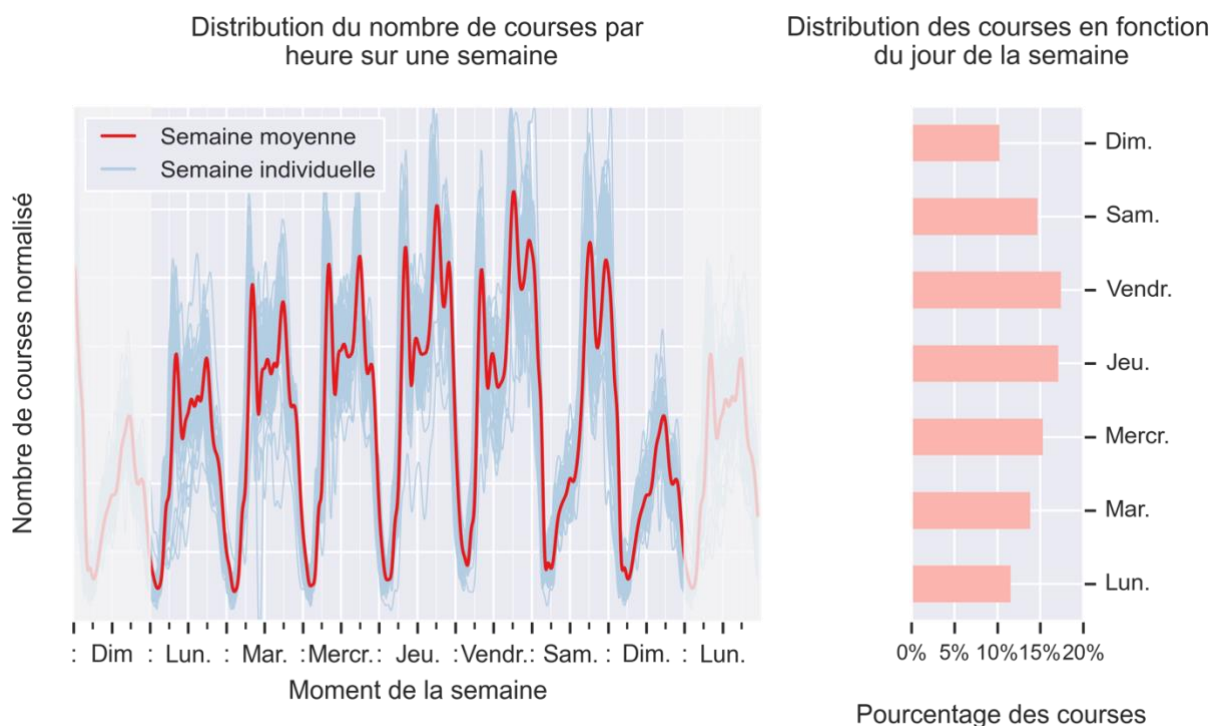


Figure 4-8 Répartition des courses en fonction de l'heure et du jour de la semaine

Sur la Figure 4-8, nous pouvons remarquer que chaque jour moyen de la semaine possède un profil unique. En semaine, deux points sont observables : la pointe du matin à partir de huit heures ainsi qu'une autre pointe en fin d'après-midi à partir de dix-huit heures. Alors que pendant le début de la semaine, le nombre de trajets effectués en pointe du matin est plus important relativement à ceux effectués en pointe de l'après-midi, ce rapport s'inverse au fil de la semaine. Le mercredi, le nombre de trajets en pointe de l'après-midi est légèrement plus important que le matin. Le jeudi et le vendredi en revanche, le nombre de trajets en pointe du soir est prédominant comparé à ceux réalisés le matin de la même journée. Le samedi et le dimanche, il n'y a plus de pic d'utilisation en matinée. À partir du jeudi et jusqu'au samedi soir, le nombre de trajets effectués le soir (au-delà de

22h) est également important. Ce nombre reste d'un ordre de grandeur similaire sur les trois soirs, et ce, malgré les différences d'intensité d'utilisation sur les trois jours.

L'intensité d'utilisation selon le jour de la semaine varie également. Le dimanche est le jour de plus faible intensité. Seulement 10,2 % des courses sont réalisées ce jour-là. Si la distribution du nombre de courses était uniforme selon le jour de la semaine, la part de déplacements réalisés chaque jour serait de 14,3 %. L'intensité d'utilisation croît du lundi au vendredi. L'utilisation est maximale le vendredi avec 17,4 % des courses. Le samedi, l'utilisation décroît, mais reste supérieure à la moyenne (14,7 %). L'utilisation est supérieure à la moyenne du mercredi au samedi.

4.4.2 Répartition spatiale des courses

Les Figure 4-9 et Figure 4-10 représentent respectivement les cartes de chaleurs des points d'origine et de destination des courses. Ces deux cartes ont été construites avec la même méthodologie et les échelles de couleurs sont donc comparables entre ces figures. Cela signifie que la même nuance de bleu représente la même densité d'observation sur les deux cartes. Plus la nuance de bleu est foncée, plus la densité de points de départ ou d'origine de courses est importante. Nous disposons d'information sur le lieu de départ pour 87,7 % des courses et sur le lieu d'arrivée pour 86,0 % des courses. Les lieux de départ et d'arrivée d'une même course sont connus pour 78,9 % des enregistrements.

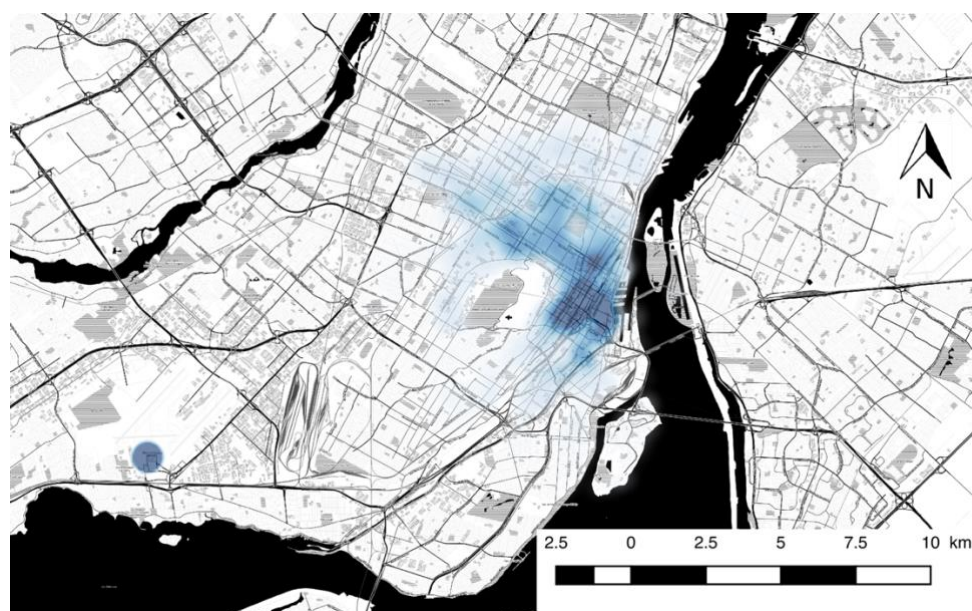


Figure 4-9 Carte de chaleur des points de départ des courses

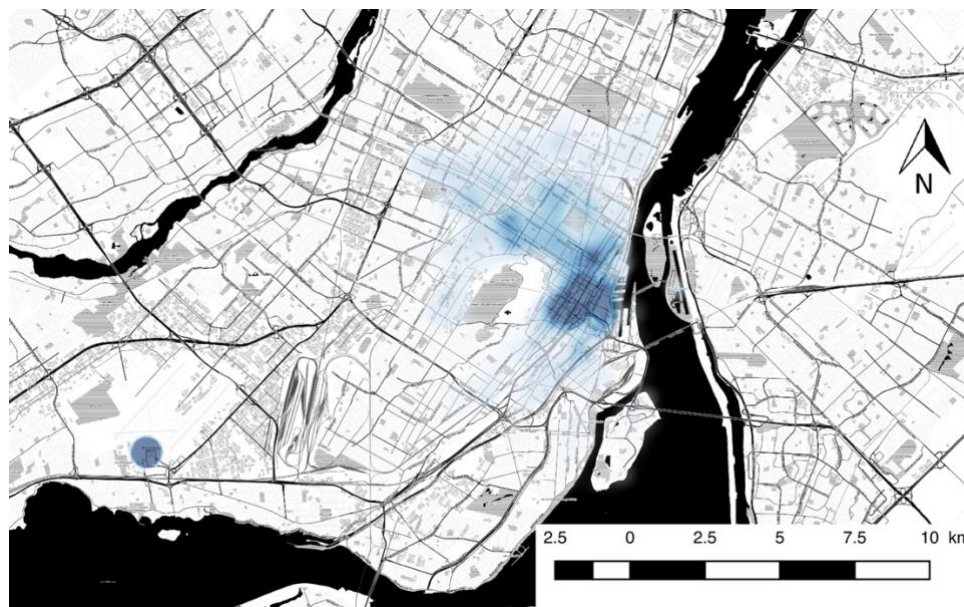


Figure 4-10 Carte de chaleur des points de destination des courses

En comparant ces deux cartes, il apparaît que la répartition des destinations des courses est légèrement plus diffuse que celle des origines. Pour les destinations, un nombre plus important de destinations se situent dans des quartiers au profil plus résidentiel avec une densité d'emplois et de lieux de divertissement plus faible. La part de courses ayant pour origine l'aéroport international Pierre-Elliott Trudeau de Montréal est de 9,7 %, ce qui est légèrement inférieur à la part des courses ayant pour destination l'aéroport (10,5 %).

Tant pour les points d'origine que pour les points de destination, deux zones se démarquent par leur forte densité. La première est l'aéroport et la deuxième est le centre-ville. Ces deux zones sont marquées sur la Figure 3-2. Bien que la zone de desserte de Téo Taxi soit vaste (zone présentée dans la Figure 3-9), la densité d'origine et de destination est beaucoup plus faible en dehors des quartiers centraux. L'activité du taxi est également importante autour de la branche Est de la ligne de métro orange entre les stations Berri-UQAM et Jean-Talon (la carte du plan de métro est disponible à la Figure 3-3).

4.4.3 Nombre de courses par utilisateur

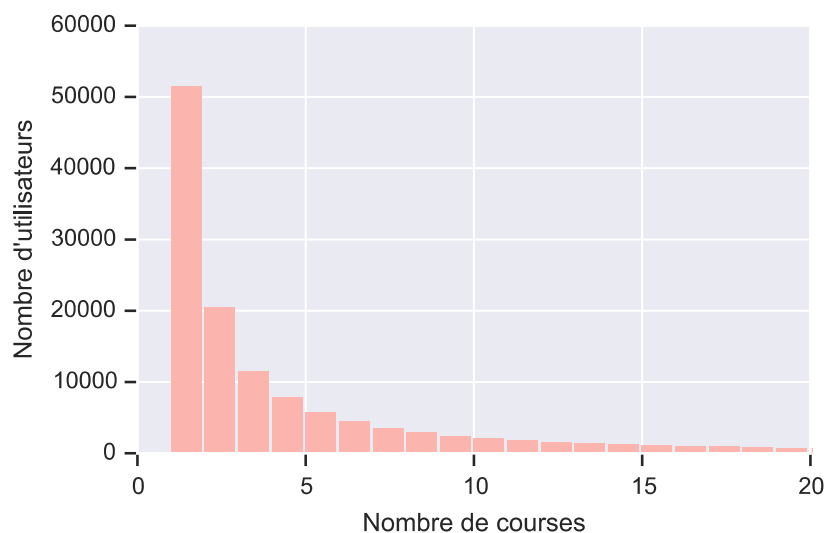


Figure 4-11 Distribution fréquentielle du nombre de courses par utilisateur

Les 1 086 938 courses réservées par le biais de l'application ont été réalisées par 135 508 utilisateurs distincts. Cela représente une moyenne de 8,0 courses par utilisateur. Ce chiffre peut sembler important, mais la fréquence d'utilisation annuelle moyenne sur l'ensemble des 1027 jours du jeu de données est de 2,8 courses par an. La médiane du nombre de trajets par utilisateur est beaucoup plus faible : elle est de 2 courses par utilisateur. Cette médiane est plus faible que la moyenne, car la queue de la distribution du nombre de courses par utilisateur est très importante. La Figure 4-11 présente le début de cette distribution. D'un côté, un grand nombre d'utilisateurs effectue un petit nombre de courses (37,9 % des utilisateurs ont réalisé seulement une course) et d'un autre côté un très petit nombre d'utilisateurs fait un très grand nombre de courses. Le nombre de courses maximales par utilisateur est de 1360 (ce qui représente une fréquence d'utilisation moyenne de 1,4 course par jour sur la durée de vie de cet utilisateur). L'opération promotionnel de Téo offrant un rabais de 10\$ sur la première course d'un client pourrait avoir eu comme effet d'augmenter le nombre de client ayant réalisé une seule course.

Le Tableau 4-3 présente les principaux indicateurs statistiques de cette distribution. La valeur importante de l'écart type traduit la grande diversité d'utilisation entre les usagers les plus fréquents et les usagers plus occasionnels. La différence entre le 95^e centile et la valeur maximale montre

également le poids important que jouent les valeurs extrêmes sur les caractéristiques de la distribution.

Tableau 4-3 Indicateurs statistiques de la distribution du nombre de courses par usager

| | |
|--------------------------------|---------|
| <i>Nombre d'utilisateurs</i> | 135 508 |
| <i>Moyenne</i> | 8,02 |
| <i>Écart type</i> | 22,0 |
| <i>Minimum</i> | 1 |
| <i>1^{er} quartile</i> | 1 |
| <i>Médiane</i> | 2 |
| <i>3e quartile</i> | 7 |
| <i>95e percentile</i> | 32 |
| <i>Maximum</i> | 1360 |

4.4.4 Distribution des grandeurs caractéristiques des courses

4.4.4.1 Distribution des courses en fonction de leur distance

Le filtre que nous avons appliqué sur la distance des courses a pour seuil minimal 300 m et pour seuil maximal 85 km. Il convient également de mentionner que la distance présentée ici est bien la distance réelle du trajet fait en taxi et non la distance à vol d'oiseau entre l'origine et la destination. Les principaux indicateurs statistiques de cette distribution sont dans le Tableau 4-4. La distance minimale est de 300 m en raison du filtre. La Figure 4-12 montre la distribution cumulée de la distance des courses. On peut voir que la proportion de courses de faible longueur (inférieure à 8 km) est très importante (70,6 %). Un deuxième pic de densité se détache également dans cette distribution. La densité de courses dont la distance est comprise entre 15 km et 25 km est importante. La présence de ce deuxième pic s'explique par le nombre important de trajets qui relie le centre-ville à l'aéroport. La distance réseau entre la station de métro Bonaventure (située dans le centre-ville) et l'aéroport est de 17,5 km.

De la même manière que pour le nombre de courses par utilisateur, la différence entre la médiane et la moyenne est importante, car la distance totale de la course est faible pour un très grand nombre d'entre elles (premier quartile à 2,9 km) et très grande pour un petit nombre de courses.

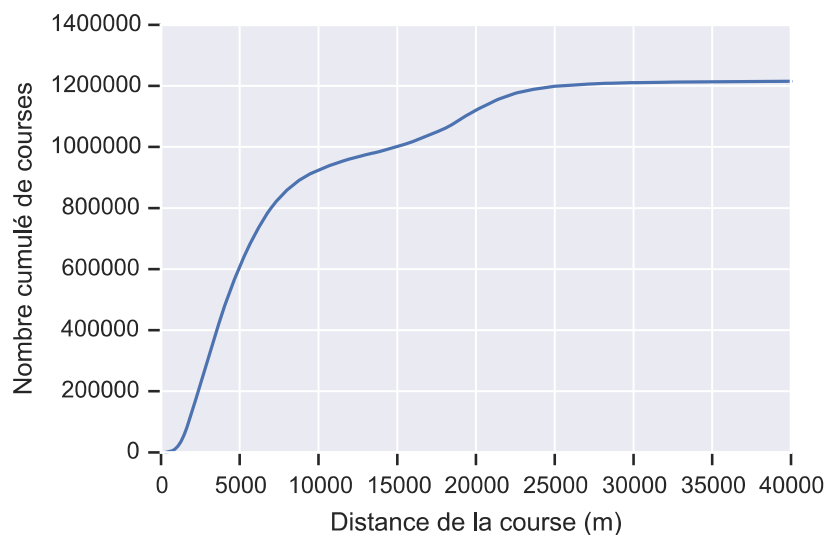


Figure 4-12 Distribution cumulée de la distance des courses

4.4.4.2 Distribution des courses en fonction de leur durée

En raison du filtre appliqué sur les durées des courses, la durée minimale d'une course est de 100 s et la durée maximale de 3600 s. La distribution cumulée de cette grandeur est sur la Figure 4-13. Contrairement à la distribution des distances des courses, la distribution des durées des courses ne présente qu'un seul pic de distribution. La distribution des durées est également plus homogène que celle des distances. Le coefficient de variation de cette distribution est de 0,558 alors que pour la distribution de la distance cette valeur est de 0,867. Rappelons que le coefficient de variation est l'écart type de la distribution normalisé sur sa moyenne. C'est une mesure de dispersion sans unité. Plus la valeur de ce coefficient est grande, plus la dispersion relativement à la moyenne est importante. Il faut également être prudent dans l'utilisation de ce coefficient lorsque la moyenne est proche de zéro, car ce coefficient va tendre vers l'infini. L'homogénéité de cette distribution est un effet de la différence de vitesse entre les courses. Si la vitesse moyenne de toutes les courses était identique, les caractéristiques des distributions des durées et des distances seraient identiques à un facteur près. Bien que la moyenne des vitesses soit légèrement plus élevée que la médiane,

l'écart relatif entre ces deux valeurs est beaucoup plus faible que pour les distributions étudiées précédemment.

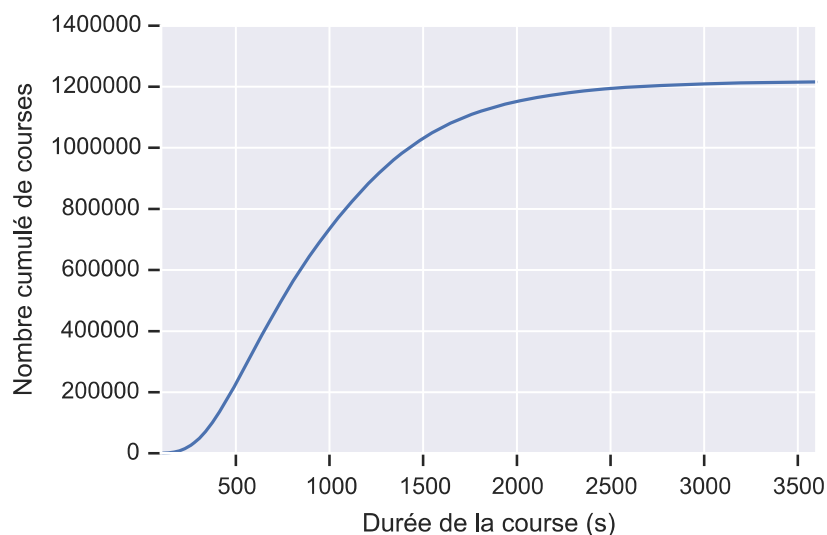


Figure 4-13 Distribution cumulée de la durée des courses

4.4.4.3 Distribution des courses en fonction de leur vitesse moyenne

La Figure 4-14 montre la distribution cumulée de la vitesse des courses. Des trois distributions cumulées, c'est celle dont le coefficient de variation est le plus faible (0,481). L'ensemble des indicateurs statistiques calculés pour cette distribution est regroupé dans le Tableau 4-4. La vitesse médiane est de $6,17 \text{ m.s}^{-1}$ (22 km.h^{-1}) ce qui est un ordre de grandeur raisonnable pour un déplacement en automobile dans un contexte urbain. La vitesse moyenne d'une course a été obtenue à partir de la durée totale de la course et de la distance réelle du trajet. Un certain nombre de courses possède une vitesse qui semble relativement élevée. La vitesse moyenne maximale observée est de $31,8 \text{ m.s}^{-1}$ (114 km.h^{-1}). En regardant spécifiquement ces courses, on s'aperçoit que beaucoup d'entre elles ont leur point d'origine et de destination proche de points d'accès au réseau routier supérieur de la région. De plus, ces courses à vitesses élevées ne sont généralement pas réalisées en période de pointe (du matin ou de l'après-midi). La distribution de la vitesse est également très intéressante car contrairement à la distance et la durée, cette grandeur est beaucoup moins sensible aux caractéristiques de la course. Des éléments extérieurs comme l'état de congestion du réseau vont avoir beaucoup d'influence sur cette grandeur.

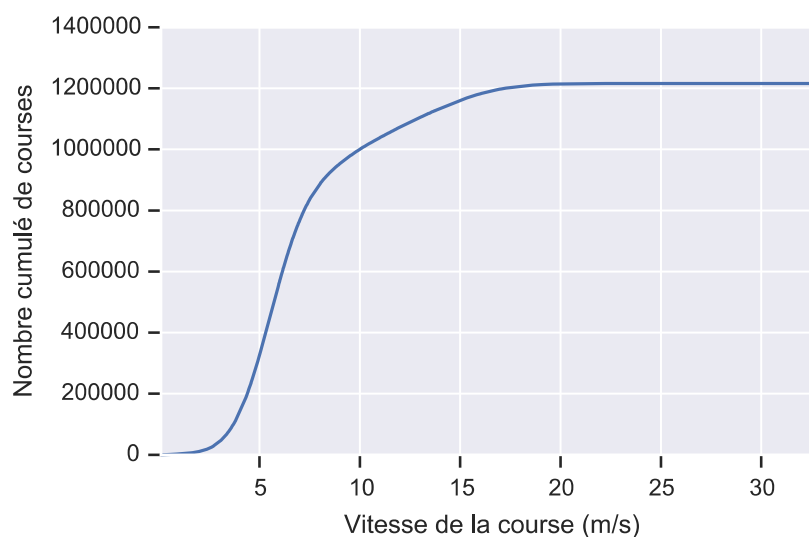


Figure 4-14 Distribution cumulée de la vitesse des courses

Tableau 4-4 Indicateurs statistiques des distributions des grandeurs caractéristiques des courses après l'application des filtres

| | <i>Distance (m)</i> | <i>Durée (s)</i> | <i>Vitesse (m.s⁻¹)</i> |
|---------------------------------|---------------------|------------------|-----------------------------------|
| <i>Nombre d'utilisateurs</i> | 1 214 772 | 1 214 772 | 1 214 772 |
| <i>Moyenne</i> | 7611 | 970 | 7,15 |
| <i>Écart type</i> | 6675 | 541 | 3,44 |
| <i>Minimum</i> | 300 | 100 | 0 |
| <i>1^{er} quartile</i> | 2960 | 567 | 4,90 |
| <i>Médiane</i> | 4992 | 852 | 6,17 |
| <i>3e quartile</i> | 9466 | 1253 | 8,18 |
| <i>Maximum</i> | 79 155 | 3600 | 31,8 |
| <i>Coefficient de variation</i> | 0,877 | 0,558 | 0,481 |

CHAPITRE 5 CARACTÉRISATION DE L'UTILISATION DE TÉO TAXI

5.1 Enjeux méthodologiques

Des enjeux méthodologiques ont été introduits dans la section 4.1.1.1. La présence de ces enjeux est expliquée d'une part par la spécificité des systèmes de transport par taxi et par les particularités opérationnelles propres à Téo Taxi d'autre part. Quatre points ont été soulevés :

- La fréquence d'utilisation du taxi est très faible comparée à d'autres modes
- La variabilité dans l'utilisation du taxi entre les usagers est grande
- La croissance du service Téo Taxi a été très importante
- La période d'analyse doit être facilement modifiable

Une des solutions utilisées en réponse à ces problématiques a été d'étudier et caractériser le comportement des usagers selon deux échelles de temps différentes. Une échelle à court terme (période de plusieurs semaines) et une échelle à long terme (intégralité de la vie du client sur l'ensemble du jeu de données). Ce type de méthodologie peut être utilisé dans l'analyse des données de validation de carte à puce dans un réseau de transport en commun (Deschaintres, 2018).

L'étude de l'utilisation de Téo Taxi sur une période est détaillée dans la partie 5.3. Le choix d'étudier les comportements sur une courte période permet de prendre en compte les problématiques liées à la croissance et l'évolution du service. En effet, la durée d'une période étant relativement courte, l'impact de l'évolution du service au sein de chaque période est très faible. De plus, ces analyses peuvent être facilement complétées avec de nouvelles données dès lors que nous disposons d'enregistrements sur l'intégralité d'une nouvelle période.

Étudier le comportement d'un utilisateur sur sa durée de vie revient alors à étudier l'union de ces comportements sur chaque période. Nous pouvons également caractériser l'évolution du système de taxi en regardant les différences comportementales de l'ensemble des utilisateurs en fonction de la période. Ces analyses à long terme sont regroupées dans la section 5.4.

5.2 Méthodologie d'analyse comportementale sur une période

La classification des usagers-périodes peut se décomposer en trois grandes étapes méthodologiques. La Figure 5-1 présente ces trois étapes ainsi que les points clés pour chacune d'entre elles. Cette méthodologie mène à la création de groupes types d'usagers-périodes. Ces groupes seront étudiés dans la section 5.3. Chacune de ces étapes est détaillée dans les trois sous-parties de cette section.

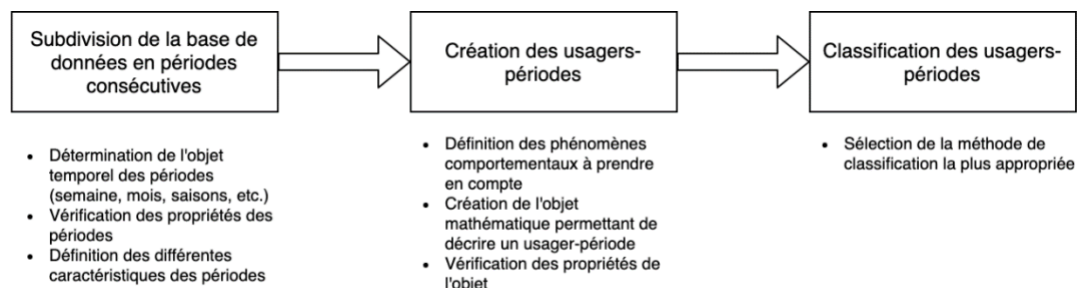


Figure 5-1 Schéma méthodologique de la construction des groupes d'usagers-périodes

5.2.1 Choix de la durée de la période

5.2.1.1 Définition d'une période

Pour segmenter temporellement un jeu de données, de nombreuses solutions existent. Une solution courante est de choisir une unité de temps existante (jour, semaine, mois, etc.) et de l'utiliser comme critère de partition. Pour les raisons présentées dans la partie 5.2.1.2, l'unité la plus adaptée dans notre étude serait le mois. Dans certaines situations (comme dans la Figure 4-7), cette unité est parfaitement adaptée. Elle présente l'avantage d'être simple et naturelle. Un inconvénient majeur de cette unité est qu'il est très difficile de comparer des mois entre eux pour faire des analyses fines. Tous les mois ne présentent pas les mêmes caractéristiques. Le nombre de jours varie selon les mois et la répartition des jours de la semaine au sein des mois n'est pas non plus la même.

Une des caractéristiques les plus importantes des périodes est qu'elles doivent être semblables et comparables. La suite de la méthodologie repose sur le fait que nous pouvons comparer les comportements des utilisateurs entre périodes. Pour cela, il apparaît essentiel de s'assurer que chaque période comprend le même nombre de jours ainsi que la même répartition des jours de la

semaine. Il est important d'avoir le même nombre de jours de la semaine au sein de chaque période. Comme le montre la Figure 4-8, la variabilité de l'utilisation du taxi est importante selon le jour de la semaine.

Pour s'assurer que ces critères soient remplis, la durée d'une période doit être définie comme un nombre de jours consécutifs où ce nombre est un multiple de sept. Le premier jour des périodes construites est un lundi. Le choix de ce jour n'a pas d'influence sur les résultats puisque les périodes sont constituées de semaines entières. Le nombre de chaque jour de la semaine dans la période est donc égal.

5.2.1.2 Détermination de la durée d'une période

Le choix de la durée des différentes périodes est crucial, car il conditionne la qualité des résultats obtenus. D'un côté, la durée de cette période doit être assez courte pour obtenir une précision temporelle permettant de détecter les évolutions comportementales potentielles. Mais cette durée doit également être suffisamment grande pour permettre d'observer une variabilité des comportements des différents utilisateurs. Dans le cas où la durée de la période choisie serait trop grande, nous pourrions observer une grande diversité de comportement au sein de chaque période. En revanche, il serait plus difficile de détecter les évolutions temporelles (différences entre les périodes). Si au contraire la durée de la période est trop petite, nous pouvons alors observer de grandes différences entre les périodes, mais les comportements au sein de chaque période sont semblables en raison de la faible fréquence d'utilisation par un grand nombre d'utilisateurs. Dans cette situation, seule une minorité de clients aurait effectué plus qu'une ou deux courses de taxi durant chaque période. Il est donc important de trouver un équilibre entre ces deux enjeux.

Afin de déterminer la durée optimale de la période, nous avons utilisé comme indicateur la durée entre deux courses effectuées par le même client. Cet indicateur nous permet d'estimer la durée minimale d'une période pour qu'une diversité intéressante d'usages soit observable au sein de cette période. La Figure 5-2 présente la distribution cumulée du temps intercourse. Le choix d'une période de quatre semaines (ou 28 jours) apparaît ici comme un bon compromis. Le fait que plus de 77 % du temps entre deux courses soit inférieur à quatre semaines indique qu'une diversité de comportements pourra être observée au sein de ces périodes. Il convient cependant de rester prudent dans l'interprétation de cet indicateur. Cela ne signifie pas 77 % des usagers réalisant

plusieurs courses le font à moins de quatre semaines d'intervalle. Un grand nombre de temps intercourses faible proviennent des utilisateurs dont la fréquence d'utilisation est la plus élevée.

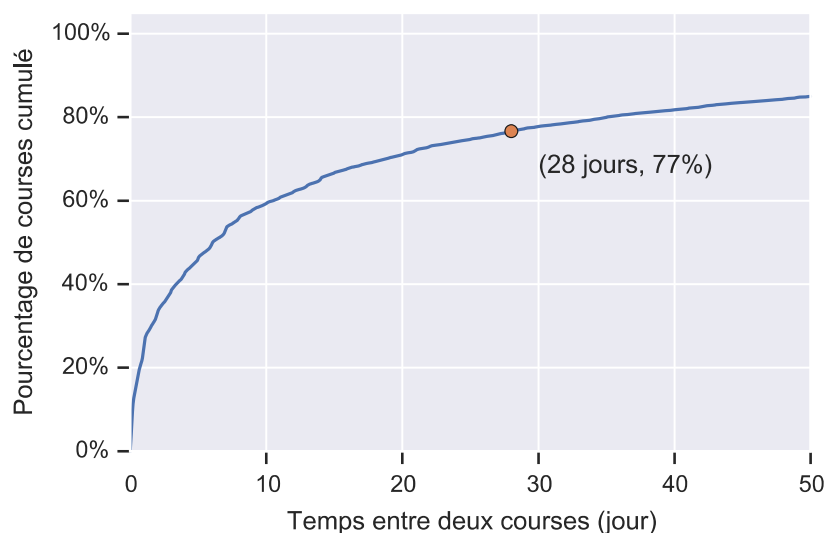


Figure 5-2 Distribution cumulée du temps intercourse

5.2.1.3 Construction des périodes

Les critères retenus pour construire les périodes sont les suivants :

- Jour de début de la période le lundi
- Durée de la période de quatre semaines (ou 28 jours)

La date de début de la première période est le premier lundi pour lequel nous disposons d'enregistrements dans la base de données. Il s'agit du lundi 11 avril 2016. Les périodes suivantes ont ensuite été construites successivement jusqu'à ce que nous ne puissions plus construire une période complète avec des enregistrements pour chaque jour de la période. La dernière période se termine le dimanche 13 janvier 2019. Au total, 36 périodes de quatre semaines ont été construites. Sur cette durée de 1 008 jours, 1 069 075 courses ont été réalisées par 134 236 clients différents. Le choix a été fait de conserver les jours fériés présents dans les périodes. En effet les services de Téo Taxi n'étant pas uniquement destinés à des professionnels, un certain nombre de courses sont également réalisées pendant les jours fériés.

5.2.2 Création des usagers-périodes

Le terme d'usager-période renvoie au comportement d'un usager pendant une période. Pour chaque usager, un usager-période est construit sur chacune des périodes durant lesquelles il a été actif. Un usager est défini comme actif sur une période s'il a réalisé un minimum d'une course sur cette période. Dans le cas contraire, aucun usager-période n'est construit, car son utilisation de Téo Taxi est nulle. Ainsi, l'ensemble des usagers-périodes affiliés à un utilisateur permet de définir son comportement sur l'ensemble de ses périodes d'activité.

Cet usager-période prend la forme d'un vecteur. Ce choix est une conséquence des méthodes de classification utilisées dans la suite de la méthodologie. Comme il est détaillé dans la partie 5.2.3.1, la méthode de classification implique la possibilité de définir une distance ou du moins une similarité entre les objets utilisés. Il est courant pour cela d'utiliser des vecteurs multidimensionnels.

La définition de l'objet « usager-période » provient de l'adaptation d'une méthodologie développée à Lyon (France) pour les utilisateurs du système de vélo en libre-service basé station Vélo'v (Vogel et al., 2014). Dans leurs travaux publiés en 2014, Vogel et al., étudient le comportement d'utilisation temporel des utilisateurs de Vélo'v sur une période d'une année. Pour cela, les auteurs ont construit un vecteur de dimension 21 permettant de quantifier la répartition des courses en fonction des jours de la semaine et des mois de l'année.

Pour les raisons expliquées dans la partie 5.2.1.2, dans le cadre de notre projet, les usagers-périodes ne décrivent pas le comportement des usagers sur un an, mais sur une période de quatre semaines. Pour cela, nous avons défini un vecteur de dimension huit permettant de retranscrire l'intensité d'utilisation sur cette période ainsi que la répartition de l'utilisation dans la semaine. La définition de chacune des composantes du vecteur (nommée de x_1 à x_8) est la suivante :

- x_1, \dots, x_7 représente le nombre de courses faites pour chaque jour de la semaine et triées dans l'ordre décroissant. Ainsi la première composante du vecteur ne représente pas les courses effectuées un lundi, mais le jour de la semaine durant lequel l'utilisateur a effectué le plus de courses. Ces sept premières composantes sont normalisées par le nombre total de courses effectuées par l'utilisateur pendant la période. La somme de ces sept premières composantes est donc toujours de 1.

- x_8 est le terme d'intensité d'usage pendant la période. Il représente le nombre de courses effectuées par l'utilisateur pendant la période.

Pour illustrer de manière plus concrète cette méthodologie, un exemple de construction d'usagers-périodes est présenté dans la sous-partie 5.2.2.1.

Le nombre maximum d'usagers-périodes qu'il est possible de former pour les 134 236 usagers sur les 36 périodes est de 4 832 496. Cependant, la plupart des usagers ne sont actifs que pendant quelques-unes des 36 périodes. Ainsi, cette méthodologie nous a permis de construire 507 018 vecteurs d'usager-période. Le grand nombre de vecteurs couplé à leur complexité (les vecteurs sont de dimension huit) rend toute analyse descriptive de l'ensemble des usagers-périodes très complexe. L'utilisation d'algorithmes de regroupement est ici nécessaire pour extraire des comportements d'utilisation et effectuer une analyse pertinente des résultats.

5.2.2.1 Exemple de construction d'usagers-périodes

Afin de mieux comprendre la construction et la signification du vecteur d'usager-période, nous présentons ici le détail de construction de trois usagers-périodes fictifs. Le Tableau 5-1 regroupe les usagers-périodes de ces trois utilisateurs. Voici les courses effectuées par ces derniers sur la période :

- Utilisateur 1 :
 - Une course le mercredi 20 avril 2016
- Utilisateur 2 :
 - Une course le mardi 12 avril 2016
 - Une course le samedi 16 avril 2016
 - Une course le mardi 19 avril 2016
 - Une course le jeudi 21 avril 2016
 - Une course le mardi 26 avril 2016
- Utilisateur 3 :
 - Deux courses le vendredi 29 avril 2016

Le premier utilisateur a effectué une course un mercredi. Le nombre de courses réalisées sur les autres jours de la semaine est nul. Ainsi, pour les composantes de x_1 à x_7 , seule la première composante (x_1) est non nulle puisque les courses sont réparties sur un unique jour de la semaine. Le nombre normalisé de courses sur ce jour de la semaine est 1. L'utilisateur 1 a effectué une seule course sur la semaine. Le terme d'intensité d'utilisation (x_8) est donc 1.

Le deuxième utilisateur a réalisé plus de course sur la période (cinq au total). Sur ces cinq courses, trois ont été faites un mardi, une un jeudi et une un samedi. Les trois premières composantes du vecteur d'usager-période représentent donc respectivement l'utilisation du taxi les mardis, les jeudis et les samedis. Chaque composante est égale au nombre de courses réalisées par ce jour de semaine normalisée par le nombre de courses total sur la période, donc cinq pour cet usager-période. La dernière composante est ici égale à cinq (nombre de courses sur la période).

Le troisième et dernier utilisateur a réalisé deux courses dans la même journée. Toutes ces courses sont donc réparties sur un même jour de la semaine. De même que pour le premier utilisateur sur les sept premières composantes de son usager-période, seule la première ne sera pas nulle et sera égale à un. Puisque cet utilisateur a complété deux courses sur la période, la dernière composante est égale à deux.

Tableau 5-1 Exemple de construction d'usager-périodes

| | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | x_8 |
|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| <i>Utilisateur 1</i> | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| <i>Utilisateur 2</i> | 0,6 | 0,2 | 0,2 | 0 | 0 | 0 | 0 | 5 |
| <i>Utilisateur 3</i> | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |

5.2.2.2 Vérification des propriétés des usagers-périodes

Le fonctionnement des algorithmes de regroupement est basé sur des mesures de distance ou de dissimilarité entre les objets à classer (Maimon et Rokach, 2010). Dans la méthodologie de classification utilisée par la suite (détaillée dans la section 5.2.3.2) la distance choisie est la distance euclidienne. Avant d'appliquer ces algorithmes, il est important que cette distance permette de caractériser les différences d'utilisation entre les usagers-périodes. Il est fréquent, lorsqu'un

algorithme de regroupement est appliqué sur des données complexes de devoir utiliser des mesures de similarité spécifiques (Deschaintres, 2018). S'assurer de la pertinence de la distance euclidienne est donc nécessaire.

Trois situations doivent être caractérisées :

- Deux usagers ont effectué le même nombre de courses dans une période mais réparties différemment en proportion différentes sur les jours de la semaine;
- Deux usagers ont effectué un nombre de courses différent, mais réparties en proportion identique sur les jours de la semaine;
- Deux usagers ont effectué un nombre de courses différent avec une réparties en proportion identique sur les jours de la semaine.

La mesure de dissimilarité doit pouvoir différencier les deux usagers-périodes dans ces trois situations. De plus cette mesure doit être nulle lorsque les deux usagers ont effectué le même nombre de courses avec la même répartition sur les jours de la semaine. Nous allons vérifier qu'une mesure de distance euclidienne respecte toutes ces contraintes.

La distance euclidienne entre deux vecteurs de dimension est définie par la formule suivante :

$$\forall x \in \mathbb{R}^8 \text{ et } y \in \mathbb{R}^8, d(x, y) = \left| \sqrt{\sum_{i=1}^8 (x_i - y_i)^2} \right|$$

5.2.2.2.1 Cas d'une utilisation identique

Dans ce cas, les huit composantes des deux vecteurs d'usagers-périodes sont identiques. La distance euclidienne entre ces deux vecteurs est alors nulle.

5.2.2.2.2 Cas du même nombre de courses dans la période réparties sur des jours de semaine différents

Lorsque la répartition entre les jours de la semaine est différente au moins pour deux des sept premières composantes des usagers-périodes. La huitième composante est identique pour ces deux vecteurs. La différence de valeur sur les premiers termes implique une distance entre ces deux vecteurs non-nulle.

5.2.2.2.3 *Cas d'un nombre de courses différent avec la même répartition sur les jours de la semaine*

Dans cette situation, les sept premières composantes sont identiques. La huitième composante sera différente entre ces deux vecteurs. La mesure de distance est alors égale à la valeur absolue de la différence entre ces deux composantes.

5.2.2.2.4 *Cas d'un nombre de courses différent avec une répartition différente sur les jours de la semaine*

Dans ce cas, d'une part au moins deux composantes sur les sept premières diffèrent entre les deux vecteurs, mais d'autre part la dernière composante est également différente. La valeur de la distance entre ces deux vecteurs sera alors strictement supérieure à la valeur absolue de la différence de leur dernière composante.

En conclusion, la distance euclidienne permet bien de différencier les usagers-périodes et de quantifier les différences d'utilisation entre deux usagers-périodes. Utiliser cette distance comme mesure de similarité dans un algorithme de classification est cohérent et permet de créer des typologies d'utilisation différenciant les différents comportements d'usage de Téo Taxi sur une période.

5.2.3 Classification des usagers-périodes

5.2.3.1 Principe des algorithmes de classification

Le terme de « classification » est ambigu en français : il peut être utilisé pour désigner deux termes anglophones référant à des procédés d'apprentissage automatique : le terme « *classification* » qui renvoie à des méthodes d'apprentissage supervisé (algorithmes adaptés à des jeux de données préalablement étiqueté, dont on connaît donc la ou les variables à prédire) et le terme « *clustering* » qui renvoie à des méthodes d'apprentissage non supervisé (A. K. Jain, Murty et Flynn, 1999). Dans ce mémoire l'expression de « classification » désigne les méthodes de « *clustering* ». Les termes de « regroupement » et de « partitionnement des données » peuvent également être rencontrés dans la littérature francophone.

La classification est un procédé de science des données qui permet de regrouper les enregistrements d'un jeu de données qui possède des caractéristiques similaires. Le développement des premiers

algorithmes de classification est relativement ancien. Des références à cette méthodologie ont été recensées dans des articles datant des années 1950 (Anil K. Jain, 2010).

La notion de mesure de similarité ou de dissimilarité est indissociable des méthodes de classification. Ces méthodologies se basent sur des mesures quantitatives de la ressemblance ou de la différence entre les entités. Bien qu'il soit courant d'utiliser une distance pour cela, cette condition n'est pas obligatoire. Il est également possible d'utiliser des mesures de similarité, la dissimilarité étant l'inverse (ou l'opposé) d'une mesure de similarité. Si comme pour les distances, ces mesures sont symétriques, elles ne sont pas tenues de respecter l'inégalité triangulaire (Maimon et Rokach, 2010). L'inégalité triangulaire implique que pour trois points dans un même espace (A, B et C), la distance de A vers B est inférieure ou égale à la somme des distances de A vers C et de C vers B.

Il n'existe pas de définition rigoureuse d'un algorithme de classification. Cependant, un tel algorithme doit répondre aux critères suivants (D. Xu et Tian, 2015) :

- Minimiser la variabilité interne à chaque groupe (tous les points appartenant au même groupe doivent être les plus semblables possible);
- Maximiser la variabilité entre les groupes (chaque groupe doit être le plus différents possibles des autres);
- Posséder une mesure de similarité qui a un sens par rapport aux données sur lesquelles elle est appliquée.

Enfin, il est également important de s'assurer que les résultats obtenus soient cohérents avec le contexte des données et peuvent avoir une explication pratique.

Il existe de nombreux types d'algorithmes de partitionnement (D. Xu et Tian, 2015). En fonction de la structure du jeu de données, ces différents algorithmes n'ont pas les mêmes performances. La Figure 5-3 permet de comparer différents algorithmes en fonction des caractéristiques de répartition d'un jeu de données à deux dimensions. Enfin, si certaines méthodes de classification déterminent automatiquement un nombre de groupe optimal (comme l'algorithme DBSCAN), cette information doit être choisie avant d'appliquer l'algorithme pour la plupart d'entre elles. En fonction de la méthode de classification utilisée, plusieurs indicateurs permettent d'aider à la détermination de cette caractéristique.

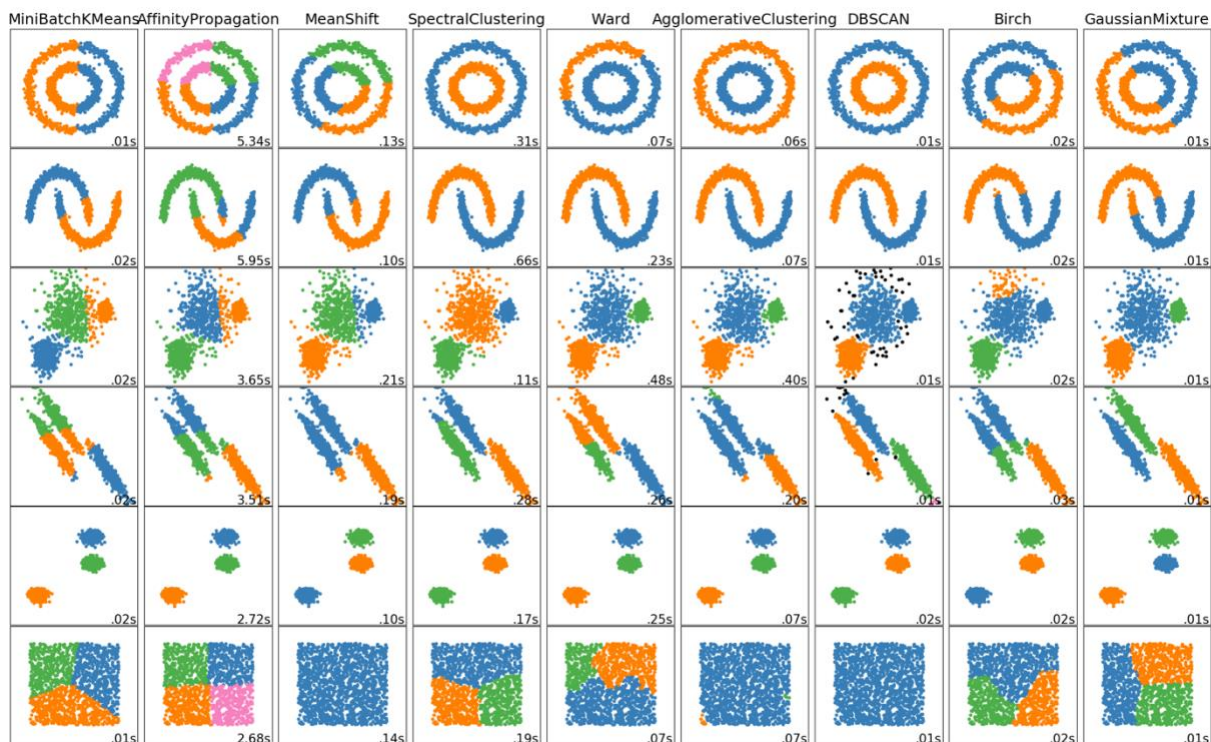


Figure 5-3 Comparaison des différents algorithmes de classification en fonction de la structure du jeu de données tirée de (Scikit-learn, 2019a)

La Figure 5-3 permet de comparer les résultats de neuf algorithmes de classifications (un par colonne) sur six jeux de données différents (un par ligne). Chaque jeu de données est de dimension deux et présente des caractéristiques spécifiques. Les groupes obtenus par chaque méthodologie sont représentés par des couleurs différentes. Les temps d'exécution sont également disponibles dans le coin inférieur droit de chaque vignette. La Figure 5-3 permet de se rendre compte de l'importance de choisir une méthodologie de classification adaptée à la structure des données étudiés. Les performances de l'algorithme et la qualité des résultats dépendent fortement de la méthode choisie.

5.2.3.2 Méthodologie de classification mise en place

Comme le montre la Figure 5-3, différentes méthodologies peuvent mener à différents résultats. En raison de la dimension importante des données, il est complexe de représenter graphiquement les résultats afin de valider les performances de l'algorithme comme dans la Figure 5-3. Pour valider la qualité de nos résultats, nous avons choisi de comparer deux méthodologies (méthode

des k-moyennes et méthode de classification hiérarchique de Ward). Les indicateurs utilisés pour identifier la meilleure méthodologie sont détaillés dans le paragraphe 5.2.3.2.4.

Les implémentations des algorithmes de classification utilisés sont celles du module *cluster* de la librairie Python d'apprentissage automatique scikit-learn (<https://scikit-learn.org/stable/modules/clustering.html#clustering>).

5.2.3.2.1 *Étape de traitement préliminaire*

Avant d'appliquer un algorithme de classification, deux étapes de traitement préliminaires ont été effectuées sur l'ensemble des usagers-périodes. Ces deux étapes ont pour but d'améliorer la performance de l'algorithme tant au niveau du temps d'exécution que de la qualité des résultats produits.

La première étape a été de créer manuellement deux groupes d'utilisation triviaux de l'ensemble des données fournies à l'algorithme. Le premier groupe est constitué des usagers-périodes qui ne sont composés que d'une seule course. Le nombre d'usagers-périodes représentant une seule course est de 291 278. Cela représente 57,4 % des 507 018 vecteurs construits. Un deuxième groupe d'usagers-périodes a également été classifié séparément; il s'agit des usagers-périodes représentant une utilisation intensive de Téo Taxi. Ce groupe est composé de tous les usagers-période pour lesquels l'utilisation est supérieure à dix courses sur la période. Cette classe d'usagers-périodes regroupe 7336 vecteurs, soit 1,4% de l'ensemble des usagers-périodes.

La deuxième étape a été de normaliser la composante d'intensité d'utilisation des vecteurs. Normaliser les données est une étape cruciale en science des données (Maimon et Rokach, 2010). Cette étape n'est pas nécessaire pour les sept premières composantes (composante de variation temporelle de l'utilisation), car elles sont déjà normalisées par construction. Ne pas normaliser la dernière composante aurait pour effet de minimiser l'importance de la répartition des courses dans le calcul de la distance entre deux vecteurs. Cette composante a été normalisée par le 95ème percentile de sa distribution. La valeur de la composante d'intensité est donc obtenue avec l'équation suivante :

$$x_8 = \frac{\text{Nombre de courses de l'utilisateur dans la période}}{31}$$

5.2.3.2.2 Méthode des k-moyennes

La méthode des k-moyennes est un algorithme itératif. À chaque itération, l'ensemble des points est attribué à un groupe. L'algorithme va alors chercher à minimiser la somme des carrés des distances entre le centroïde du groupe et chaque point de ce groupe. Le centroïde (ou centre du groupe) est défini comme la moyenne arithmétique des points du groupe. Cette méthodologie présente l'avantage de ne nécessiter que très peu de paramètres. Seuls la définition de la distance et le nombre de groupes sont nécessaires (Anil K. Jain, 2010). Un autre avantage calculatoire de cette méthodologie est de ne pas avoir besoin de calculer une matrice de distance entre l'ensemble des points. Le calcul de distance entre les points et les centroïdes uniquement est suffisant.

Pour déterminer le nombre de groupe optimal, il est fréquent, pour ces algorithmes, d'utiliser la méthode dite du coude. Cette méthode consiste à représenter graphiquement l'inertie en fonction du nombre de groupes. L'inertie est la somme des carrés des distances entre les points appartenant au groupe et leur centroïde respectif (Maimon et Rokach, 2010). Ce graphique est visible sur la Figure 5-4. Pour déterminer le nombre de groupes optimal, il faut localiser la valeur pour laquelle la variation de la pente de la courbe est importante. Ici le nombre de groupes optimal est de trois.

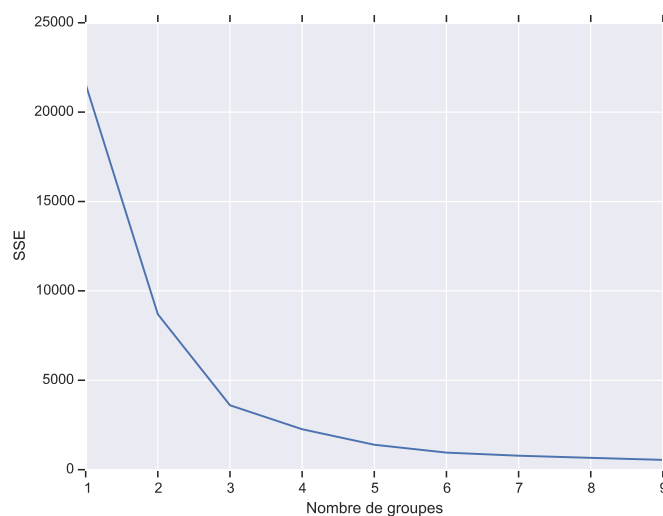


Figure 5-4 Graphique de la méthode du coude

5.2.3.2.3 Méthode de classification hiérarchique de Ward

Contrairement à l'algorithme des k-moyennes, la méthode de classification hiérarchique de Ward construit les groupes au fur et à mesure. Au départ, chaque point représente un groupe à part entière.

Les groupes sont ensuite fusionnés un par un en fonction de leur distance. À chaque étape, les deux groupes les plus proches sont fusionnés et ce, jusqu'à ce que le nombre de groupes désiré soit atteint. À chaque fois que deux groupes sont fusionnés, le centroïde du nouveau groupe obtenu est recalculé. La distance prise en compte dans la méthode de Ward pour fusionner deux groupes est la distance entre les centroïdes de ces groupes (Maimon et Rokach, 2010).

Cette méthodologie implique de calculer une matrice de distance de l'ensemble des points. Cela est très pénalisant d'un point de vue calculatoire. Lorsque le nombre de points de départ est trop important relativement aux ressources informatiques à notre disposition, il est possible de diminuer le nombre de points en effectuant une préclassification à l'aide de l'algorithme des k-moyenne (Agard, Morency et Trépanier, 2006). Cette solution a dû être utilisée dans le cadre de ce projet de recherche. Nous avons d'abord calculé 1000 groupes avec les k-moyennes. L'algorithme de classification hiérarchique de Ward a ensuite été appliqué sur les centroïdes de ces 1000 groupes.

Pour déterminer le nombre de groupes, il est fréquent d'utiliser un dendrogramme. Un dendrogramme permet de représenter la distance entre un groupe résultant de l'union de deux sous-groupes et chacun de ces deux sous-groupes. Le dendrogramme de la classification des 1000 centroïdes est présent sur la Figure 5-5. Pour déterminer le nombre de groupes, il faut repérer dans la figure une hausse importante de la distance entre le centroïde d'un nouveau groupe et celui des deux groupes fusionnés. Le nombre de groupes préconisé pour cette méthodologie est également de trois.

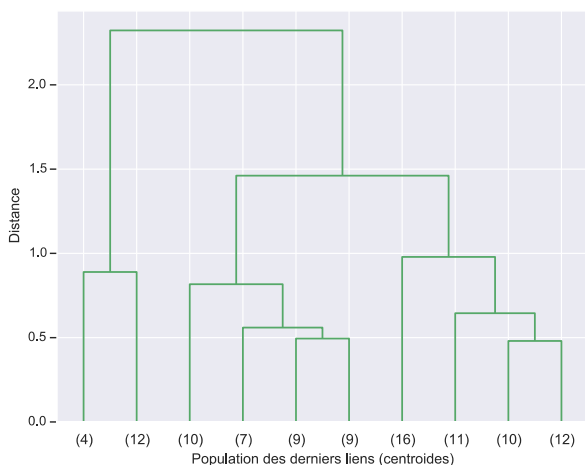


Figure 5-5 Dendrogramme de la classification des 1000 centroïdes

5.2.3.2.4 Comparaisons des deux méthodes

Les trois groupes obtenus par ces deux méthodes sont très proches. Les centroïdes obtenus pour chacun des groupes sont semblables et font ressortir les mêmes comportements dans l'ensemble. Cela nous permet de pouvoir comparer les groupes entre eux et déterminer laquelle de ces deux méthodologies est la plus performante. Ainsi la désignation des groupes de 1 à 3 dans les Tableau 5-2 et Tableau 5-3 fait référence aux groupes les plus semblables des deux méthodologies.

Pour comparer ces deux méthodes, nous avons utilisé les deux premiers critères énoncés dans le paragraphe 5.2.3.1 à savoir minimiser la variabilité interne à chaque groupe et maximiser la variabilité entre les groupes. Le Tableau 5-2 permet de comparer le premier critère. Ce tableau représente la différence entre l'écart type de chaque composante au sein des groupes entre les groupes obtenus avec les k-moyennes et les groupes obtenus par la méthode de Ward. Lorsqu'une valeur est verte, cela signifie que l'écart-type de cette composante est plus faible avec les k-moyennes. Pour minimiser la variabilité interne, nous cherchons donc à minimiser l'écart-type de chacune des composantes. Dans l'ensemble, le critère de minimisation est mieux respecté avec l'algorithme des k-moyennes.

Tableau 5-2 Différences des écarts-types

| | x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 |
|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| groupe 1 | -3,56E-02 | -8,19E-02 | -1,22E-01 | -5,09E-02 | -2,08E-02 | -6,26E-03 | -6,15E-04 | -2,79E-02 |
| groupe 2 | -8,16E-02 | -8,01E-02 | -1,45E-02 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | -4,44E-03 |
| groupe 3 | 4,54E-02 | 2,83E-02 | 1,99E-02 | 3,10E-02 | -2,28E-02 | -1,93E-02 | -7,02E-03 | -1,06E-02 |

Le Tableau 5-3 permet quant à lui de comparer le critère de maximisation de la variabilité entre les groupes. Ce tableau représente la différence des matrices de distance des centroïdes des groupes entre les k-moyenne et la méthode de Ward. Là encore, une valeur verte signifie une meilleure performance de l'algorithme des k-moyennes. Pour maximiser ce critère, il faut que la distance entre les groupes soit la plus importante possible. La distance utilisée pour obtenir ce tableau est la distance euclidienne puisque c'est également celle qui a été utilisée pour appliquer ces deux méthodologies.

Tableau 5-3 Distance entre les groupes

| | groupe 1 | groupe 2 | groupe 3 |
|----------|----------|----------|----------|
| groupe 1 | 0,00E+00 | 1,39E-01 | 4,05E-02 |
| groupe 2 | 1,39E-01 | 0,00E+00 | 7,50E-03 |
| groupe 3 | 4,05E-02 | 7,50E-03 | 0,00E+00 |

Au regard de ces deux critères, la méthode de classification par l’algorithme des k-moyennes a donc été privilégiée par rapport à la classification avec la méthode de Ward. Dans la suite du mémoire, tous les groupes analysés ont été obtenus par l’algorithme de classification des k-moyennes, plus spécifiquement l’implémentation dite « k-mean++ » (Scikit-learn, 2019b).

5.3 Analyse des résultats de la classification des usagers-périodes

L’application de la méthodologie décrite dans la section 5.2 sur l’ensemble des 507 018 usagers-périodes nous a permis d’extraire cinq typologies d’utilisations sur une période. Deux typologies proviennent du traitement préliminaire des données et trois sont issues de l’application de l’algorithme des k-moyennes. Pour améliorer la compréhension des résultats, le nom de chaque groupe a été choisi de sorte qu’il reflète l’utilisation au sein de ce dernier. Voici les cinq noms retenus :

- Aller-retour;
- Usage faible;
- Usage moyen;
- Usage exceptionnel;
- Usage intensif.

Les centroïdes de ces cinq groupes sont sur la Figure 5-6. Le nombre d’usagers-périodes attribués à chacun des groupes est également visible sur cette figure. Les valeurs des coordonnées des centroïdes ainsi des écarts-types de chaque composante sont présentées dans le tableau disponible dans l’annexe B.

Représentation des centroïdes des groupes d'usagers-périodes

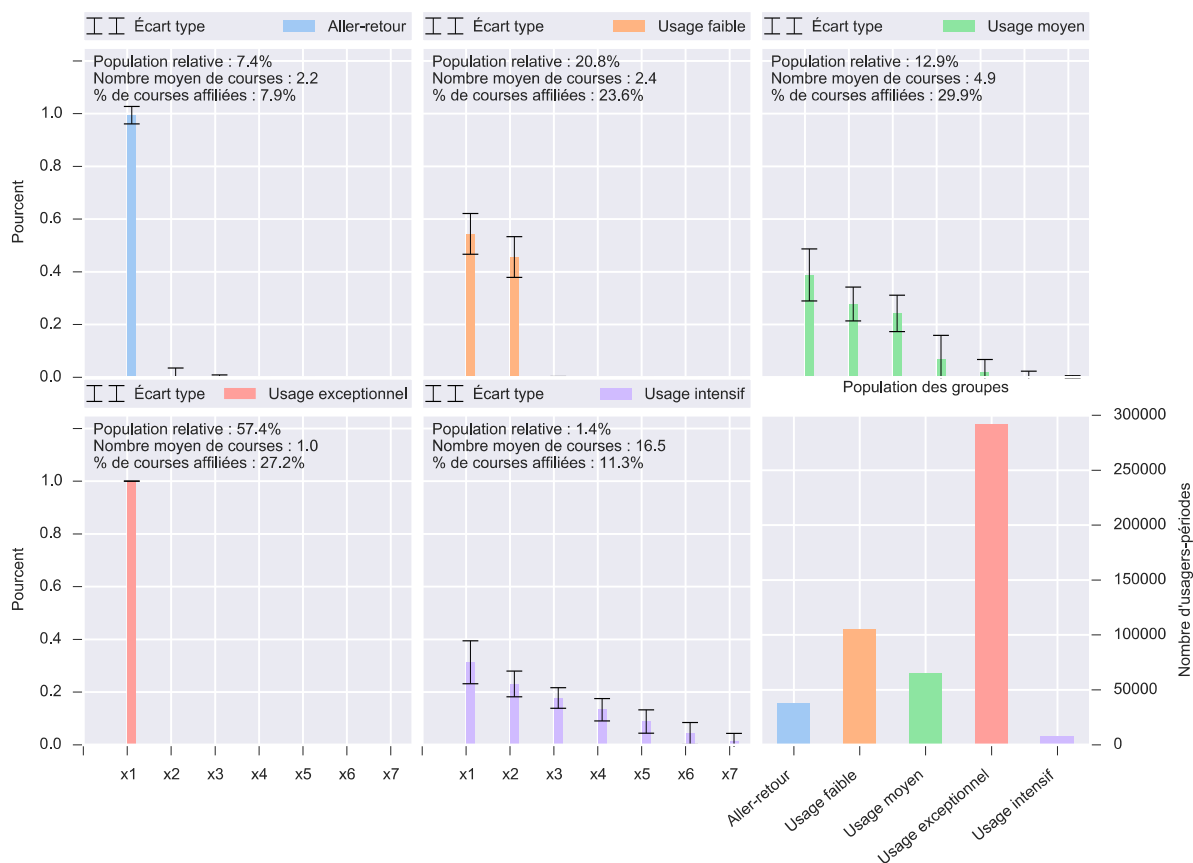


Figure 5-6 Représentation des centroïdes des cinq groupes obtenus

L'analyse de la population de chaque groupe semble montrer un fort déséquilibre entre les typologies. Le rapport entre la population du groupe le plus représenté (usage exceptionnel) et le groupe le moins représenté (usage intensif) est de 39,7. Cependant, en regardant l'importance relative de chaque groupe en fonction de son utilisation (nombre de course moyen dans la période), le poids relatif de chaque groupe est beaucoup plus équilibré. Pour cela, la Figure 5-7 montre pour chaque groupe le nombre de courses qui lui est associé. Le rapport du nombre de courses entre le groupe le plus important (usage moyen) et le groupe le plus faible (aller-retour) n'est plus que de 3,8. Ainsi, même si les différences dans l'utilisation des groupes sont importantes, l'ordre de grandeur du nombre de courses associées est comparable pour chacun des groupes. Cela nous permet d'étudier chaque groupe avec la même importance. L'homogénéité du nombre de courses pour chaque groupe s'explique par le fait que d'une part, les groupes à la population la plus faible représentent les usagers-périodes avec la plus forte intensité d'utilisation et, d'autre part, que les groupes à la plus forte population représentent des usagers-périodes de faible utilisation.

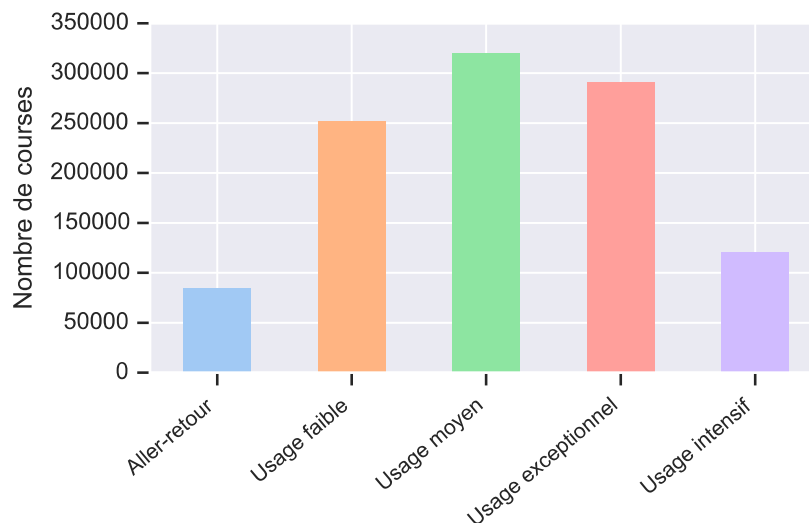


Figure 5-7 Distribution du nombre de courses associées à chaque groupe

5.3.1 Description des centroïdes des groupes

Pour compléter les graphiques de la Figure 5-6, nous présentons dans cette partie les diagrammes boîtes à moustache de chaque composante de l'ensemble des groupes. Pour rendre les graphiques plus faciles à comprendre, la dernière composante n'est pas le nombre de courses normalisées sur la période, mais le nombre des courses brutes sur la période. Les caractéristiques des diagrammes boîtes à moustache sont présentés dans la Figure 5-8.

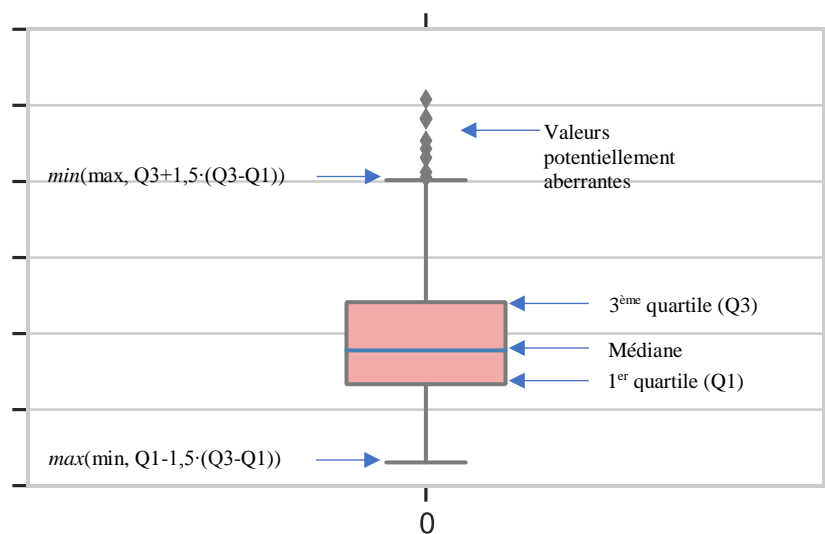


Figure 5-8 Exemple de diagramme boîte à moustache

5.3.1.1 Groupe des allers-retours

Le premier groupe est celui des allers-retours. Ce groupe est composé à plus de 75 % d'usagers-périodes représentant deux courses sur la période. Quelques usagers-périodes de ce groupe représentent plus que deux de déplacements (jusqu'à dix). Le nombre moyen de courses par période pour ce groupe est de 2,2. Les courses de ces groupes sont presque exclusivement réparties sur un seul jour de la semaine. La population de ce groupe est de 37 617 usagers-périodes. Le nombre total de trajets qui a été effectué par le groupe est 84 382. Il s'agit du groupe avec le moins de courses associées. Les usagers-périodes qui composent ce groupe ne représentent pas uniquement des allers-retours, mais c'est un des usages possibles dans ce groupe.

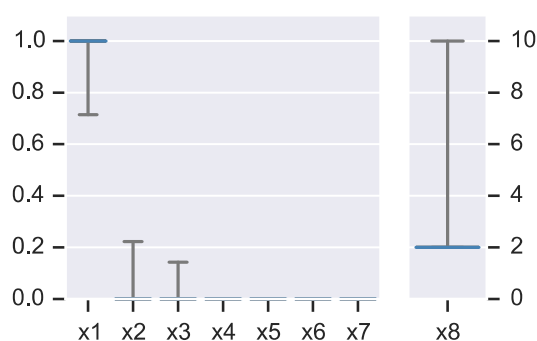


Figure 5-9 Diagramme boîte à moustache du groupe aller-retour

5.3.1.2 Groupe d'usage faible

Le groupe d'usage faible est composé d'usagers-périodes représentant majoritairement deux ou trois courses (la médiane ainsi que la valeur minimale du nombre de courses étant de deux et le troisième quartile de trois). Contrairement au groupe des allers-retours, ces courses ne sont plus réparties sur une seule journée, mais sur deux jours de semaine différents. Les médianes des deux premières composantes sont 0,5. Ce groupe comptabilise 105 458 usagers-périodes ce qui représente 252 439 courses.

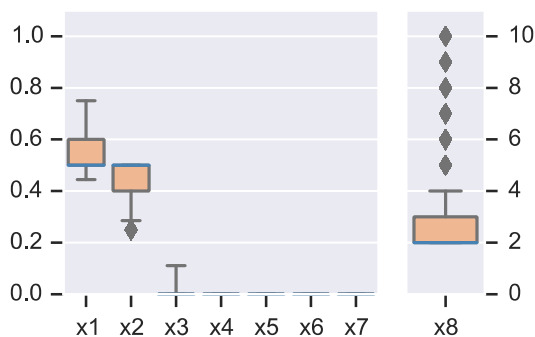


Figure 5-10 Diagramme boîte à moustache du groupe usage faible

5.3.1.3 Groupe d'usage moyen

Le groupe d'utilisateur faible est le troisième et dernier groupe obtenu par la classification réalisée par la méthode des k-moyennes. L'utilisation moyenne de ce groupe est la plus importante de ces trois premiers groupes. L'utilisation moyenne est de 4,9 courses par période et l'utilisation minimale de 3 courses par période. La moitié des usagers-périodes de ce groupe représente trois ou quatre courses. Ces courses sont principalement réparties sur trois jours de la semaine. Plus d'un quart des usagers ont cependant réparti leurs courses sur quatre jours et plus. La population du groupe est de 65 329 usagers-périodes. C'est également le groupe qui représente la plus grande utilisation du taxi avec 320 065 courses.

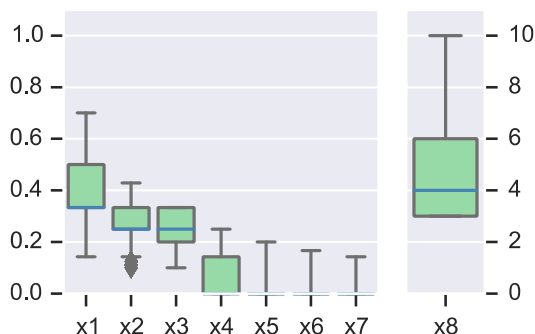


Figure 5-11 Diagramme boîte à moustache du groupe usage moyen

5.3.1.4 Groupe d'usage exceptionnel

Ce groupe représente les utilisations uniques sur une période. Sans surprise, c'est le groupe qui comprend la plus grande population : 291 287 usagers-périodes qui représentent exactement le même nombre de courses.

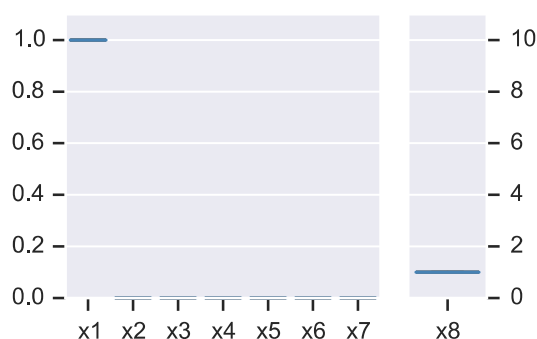


Figure 5-12 Diagramme boîte à moustache du groupe usage exceptionnel

5.3.1.5 Groupe d'usage intensif

Ce dernier groupe est également issu du traitement effectué en amont de la méthode des k-moyennes. C'est le groupe d'usage moyen le plus important. Il regroupe tous les usagers-périodes comprenant dix courses et plus. Le nombre moyen de courses est de 16,5 par période. L'usage maximal observé dépasse la centaine de courses sur une période. Il n'est pas étonnant non plus que ce groupe soit celui dont la répartition sur les jours de la semaine est la plus importante. Pour plus de la moitié des usagers-périodes, les courses sont réparties sur six jours de la semaine. La forte intensité d'utilisation de ce groupe pourrait en partie être expliquée par la présence de comptes corporatifs. Il est cependant impossible de vérifier cette hypothèse car les données utilisées pour ce projet ne contiennent pas cette information. La population de ce groupe est faible comparée à celle des autres groupes (seulement 7 336 usagers-périodes), mais en raison de leur forte intensité d'utilisation le nombre de courses représentées par ces usagers-périodes est important : 120 911 courses.

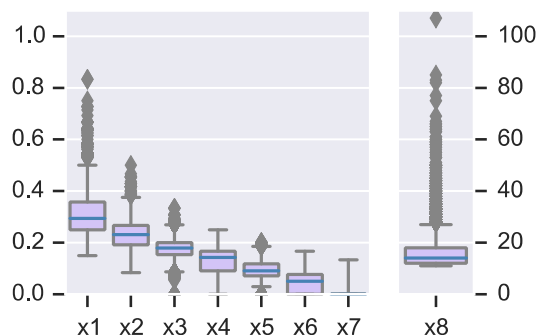


Figure 5-13 Diagramme boîte à moustache du groupe usage intensif

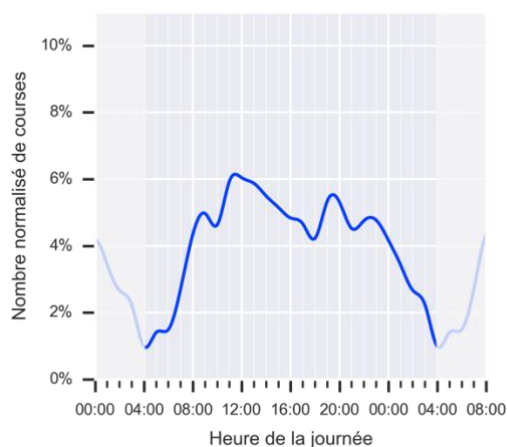
5.3.2 Analyse de l'usage des groupes sur une période

Dans la section précédente, les caractéristiques intrinsèques de groupes d'utilisateurs-périodes ont été analysées, mais nous nous sommes peu intéressés aux particularités de l'utilisation de Téo Taxi faite par chacun de ces groupes. Cette section se concentre sur la répartition temporelle des courses réalisées au sein de chacun des groupes. Pour cela, nous avons pour chaque groupe étudié la distribution des courses en fonction de l'heure de départ de la course et également la distribution des courses et selon le jour de la semaine.

5.3.2.1 Groupe des allers-retours

Les moments des courses des utilisateurs-périodes de ce groupe sont stables dans la journée. L'utilisation est soutenue entre 8h du matin et minuit. De l'ensemble des groupes, c'est celui dont la variation d'utilisation pendant la journée est la plus faible. Entre 11h du matin et 13h, l'utilisation est plus forte, mais sans pour autant que l'on observe un pic d'utilisation très marqué. Ce groupe est plus actif le vendredi et le samedi et très peu actif le dimanche et le lundi. Les deux jours d'utilisation maximale sont le vendredi et le samedi alors que le dimanche et le lundi sont les jours de plus faible activité. La période d'utilisation maximale de ce groupe sur l'ensemble de la semaine est le jeudi à 20h.

Distribution des courses en fonction de l'heure de la journée du groupe



Distribution des courses en fonction du jour de la semaine du groupe

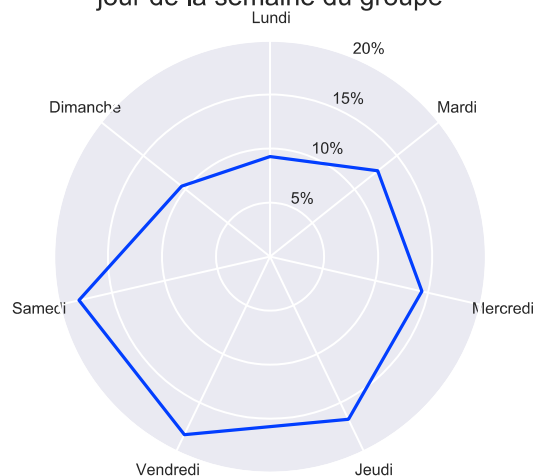
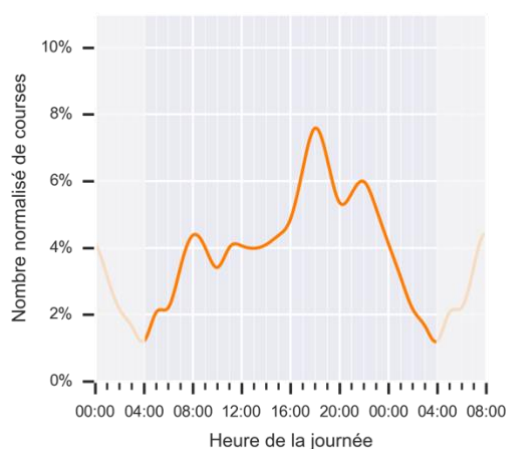


Figure 5-14 Distribution des courses du groupe aller-retour en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)

5.3.2.2 Groupe d'usage faible

Contrairement au groupe aller-retour, la répartition des courses des usagers-périodes de ce groupe n'est pas constante dans la journée. À partir de 7h du matin, l'utilisation est relativement faible et ce, jusqu'à 14h. Au-delà, l'utilisation croît fortement pour atteindre un maximum à 18h. Il faut ensuite attendre minuit pour que l'utilisation redevienne inférieure à celle de la matinée. L'évolution de l'utilisation en fonction du jour de la semaine est comparable à celle du groupe aller-retour à la différence près que l'écart entre le jour d'utilisation la plus forte (le vendredi) et le jour d'utilisation la plus faible (le dimanche) est moins important. L'écart type de la proportion de courses par jour de la semaine est de 0,026 contre 0,034 pour le groupe aller-retour. C'est le groupe avec la répartition des courses en fonction du jour de la semaine la plus équilibrée. La période d'utilisation maximale pour ce groupe est le vendredi, mais à 18h.

Distribution des courses en fonction de l'heure de la journée du groupe



Distribution des courses en fonction du jour de la semaine du groupe

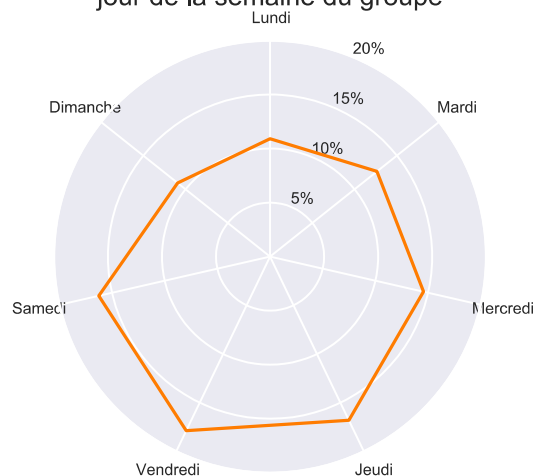
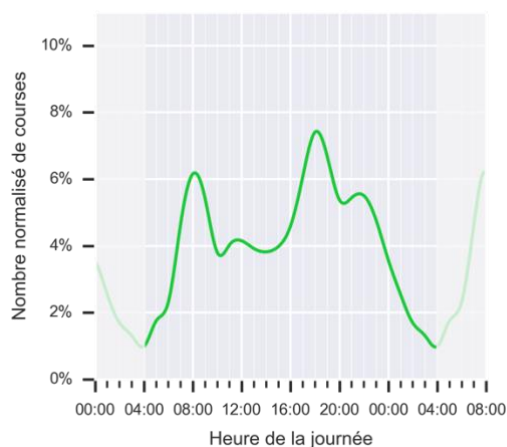


Figure 5-15 Distribution des courses du groupe usage faible en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)

5.3.2.3 Groupe d'usage moyen

La particularité de ce groupe par rapport aux deux précédents est d'avoir deux points d'utilisation dans la journée: une le matin à 8h et une l'après-midi à 18h. La pointe d'utilisation du matin est plus faible et moins longue que celle de l'après-midi. La répartition des courses en fonction du jour de la semaine diffère également des deux groupes précédents. L'utilisation est ici plus forte en semaine, particulièrement le lundi et le mardi. Le jour d'utilisation minimale est également le dimanche, mais l'utilisation maximale est observée le jeudi et le vendredi. La période d'utilisation maximale sur la semaine est le jeudi à 18h.

Distribution des courses en fonction de l'heure de la journée du groupe



Distribution des courses en fonction du jour de la semaine du groupe

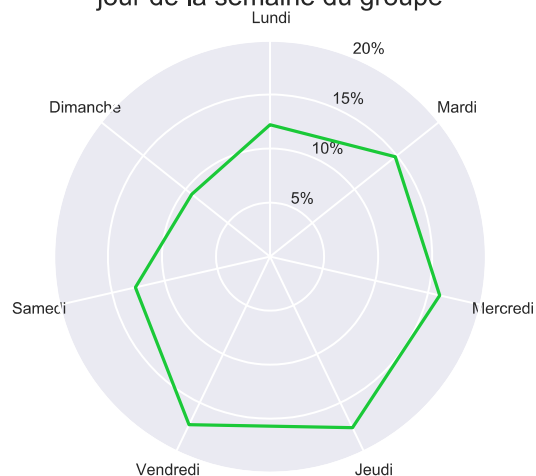
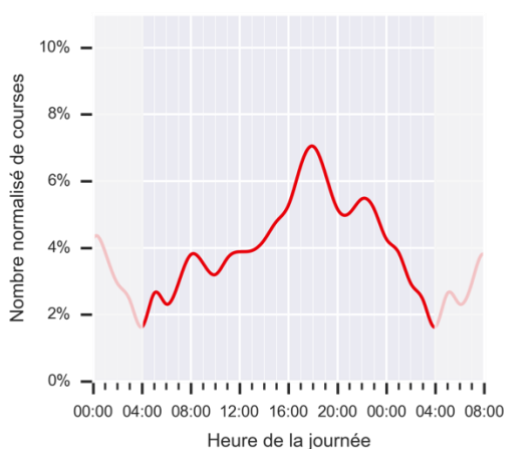


Figure 5-16 Distribution des courses du groupe usage moyen en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)

5.3.2.4 Groupe d'usage exceptionnel

La distribution de l'utilisation de ce groupe dans une journée est relativement proche de celle du groupe d'usage faible. L'utilisation est faible en matinée puis augmente en début d'après-midi pour arriver à un maximum à 18h. La largeur du pic d'utilisation cependant est plus importante pour ce groupe. Le profil d'utilisation selon le jour de la semaine est proche de celui du groupe aller-retour. Le nombre de courses est plus important en fin de semaine (vendredi, samedi) et est plus faible en début de semaine (lundi, mardi). La période d'usage maximale du taxi reflète le rôle de complémentarité du transport collectif introduit dans le paragraphe 2.2.2. L'utilisation maximale est observée le samedi à 1h du matin lorsque le service offert par les réseaux de transport en commun est relativement faible.

Distribution des courses en fonction de l'heure de la journée du groupe



Distribution des courses en fonction du jour de la semaine du groupe

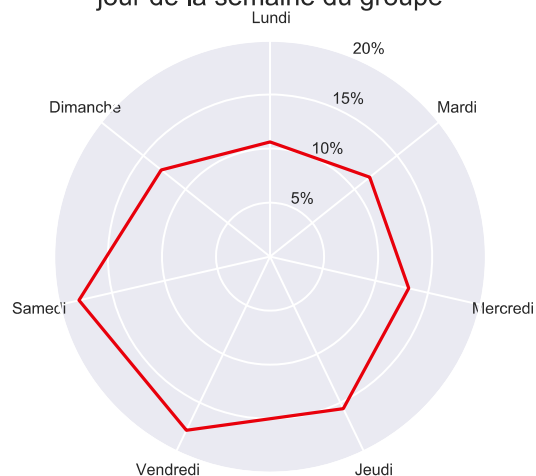
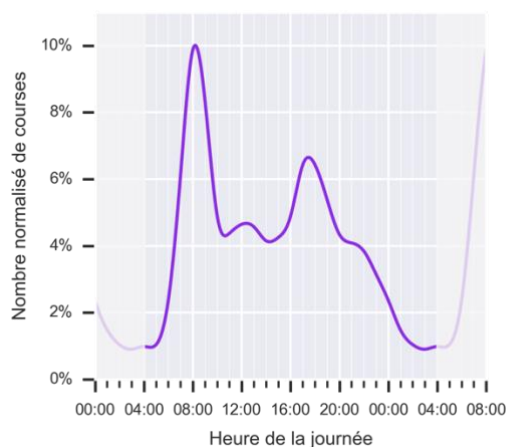


Figure 5-17 Distribution des courses du groupe usage exceptionnel en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)

5.3.2.5 Groupe d'usage intensif

Le portrait d'utilisation de ce groupe se détache de ceux des quatre groupes précédents. Comme pour le groupe d'usage moyen, deux périodes de forte utilisation durant la journée sont observables : une le matin à 8h et l'autre le soir à 18h. Cependant, contrairement au groupe d'usage moyen, l'utilisation est beaucoup plus forte lors de la pointe du matin. La répartition des courses sur les jours de la semaine est également très différente. Ce groupe est le seul à avoir une forte utilisation du taxi en semaine (du lundi au vendredi) et une utilisation relativement faible les samedis et dimanches. Ce groupe est également le seul à avoir une période de plus forte utilisation sur la semaine en matinée : le mercredi à 8h. L'utilisation de ce groupe ressemble beaucoup à une utilisation du taxi pour des raisons professionnelles.

Distribution des courses en fonction de l'heure de la journée du groupe



Distribution des courses en fonction du jour de la semaine du groupe

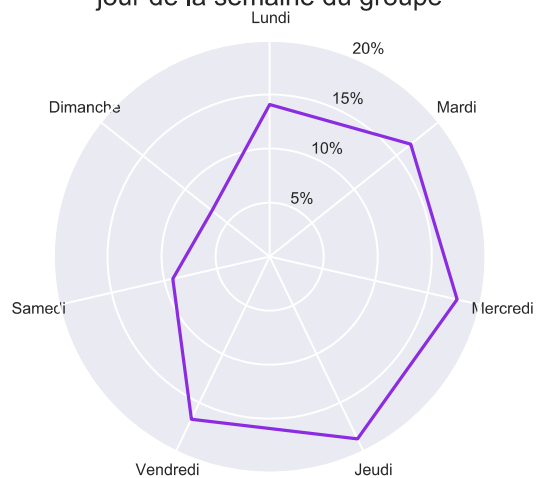


Figure 5-18 Distribution des courses du groupe usage intensif en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)

5.3.2.6 Comparaison des groupes

La Figure 5-19 montre une superposition de ces graphiques pour tous les groupes. Cette superposition permet de comparer facilement les comportements de chaque groupe. Ainsi, les spécificités du groupe d'usage intensif sont très visibles. La pointe d'utilisation du matin ressort très clairement. Le fort usage de ce groupe en début de semaine est également très marqué. On remarque également la forte ressemblance entre les groupes d'usage faible et les groupes d'usage exceptionnel particulièrement du point de vue de la distribution des courses en fonction de l'heure de la journée. Enfin, la répartition des courses du groupe d'usage moyen apparaît dans cette figure comme intermédiaire entre le groupe d'usage intensif et le groupe d'usage exceptionnel. Cela est observable à la fois pour les répartitions des courses en fonction de l'heure de la journée et pour la répartition selon le jour de la semaine.

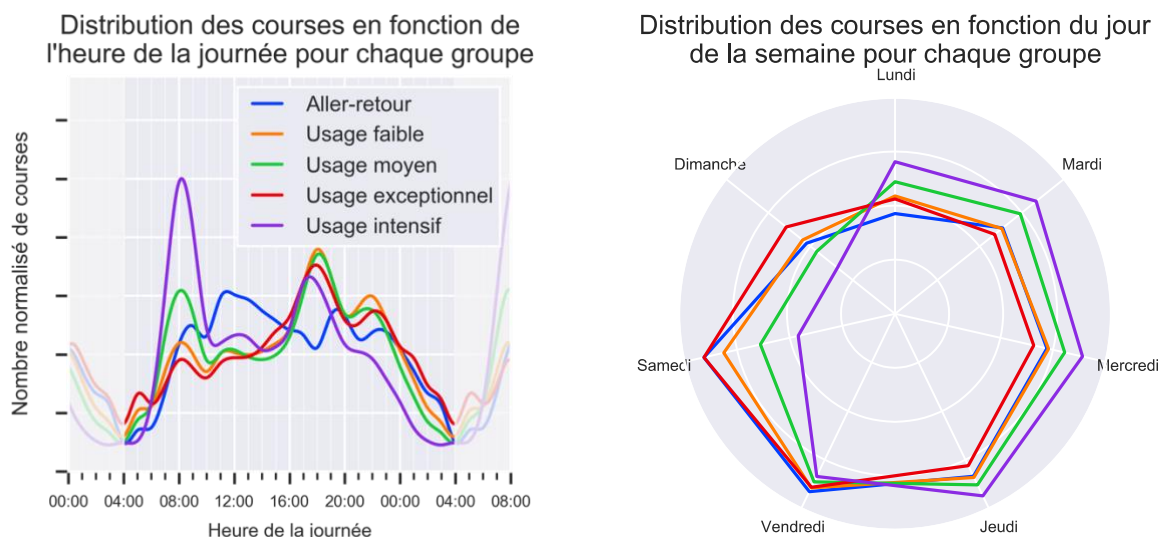


Figure 5-19 Distribution des courses pour l'ensemble des groupes en fonction de l'heure de la journée (gauche) et du jour de la semaine (droite)

5.4 Analyses de l'utilisation de Téo Taxi sur une période de longue durée

La section 5.3 a permis de dresser un portrait d'utilisation de Téo Taxi sur une période de quatre semaines. Dans cette section, en revanche, nous étudions l'utilisation du service sur une durée plus longue. Définir une période d'étude plus grande permet de détecter les éventuelles évolutions dans l'utilisation au cours du temps. Pour cela, nous nous intéresserons d'abord à l'utilisation du système de taxi dans son ensemble puis au comportement des usagers.

5.4.1 Évolution de l'utilisation globale de Téo Taxi

La Figure 5-20 permet de se rendre compte de l'évolution des profils d'usagers-périodes en fonction de la période sur l'intégralité de la durée de vie de Téo Taxi. Tout d'abord, trois périodes de quatre semaines se détachent. Ces trois périodes de quatre semaines incluent les mois de novembre et décembre 2016 et février 2017. Une défaillance du système informatique de Téo avait été observée sur ces périodes et a malheureusement partiellement endommagé les données récoltées. Ces deux mois sont spécifiques et peuvent également être considérés comme anormaux pour d'autres distributions. Dans la Figure 4-7, représentant le nombre de courses réalisées par

mois par exemple, ces deux mois apparaissent comme des mois avec une activité relativement faible comparativement à la tendance générale.

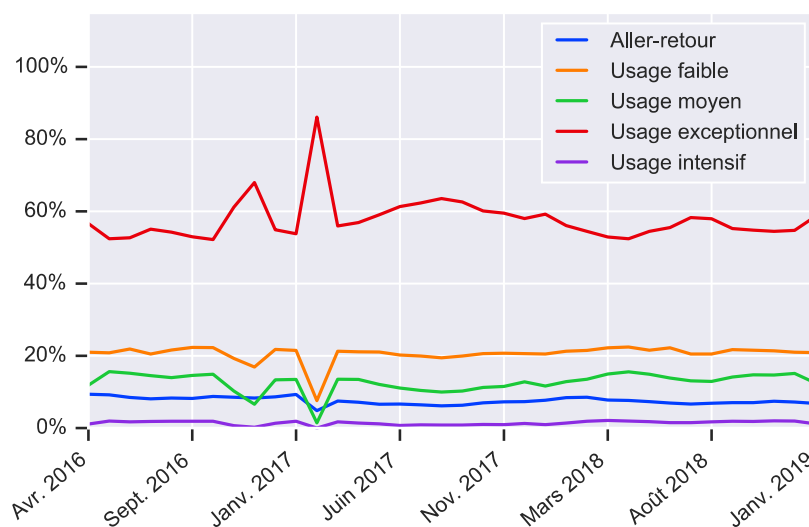


Figure 5-20 Population relative des groupes d'utilisateurs-périodes en fonction de la période

Au-delà de ces trois anomalies, la répartition de la proportion de chaque groupe est particulièrement stable. En excluant ces trois périodes, l'écart entre la valeur maximale et la valeur minimale est respectivement pour chaque groupe de : 3,2; 3,0; 5,7; 11,4 et 1,4 points de % (dans l'ordre de la légende de la Figure 5-20). Au regard de la forte croissance qu'a connue Téo Taxi, ce constat est particulièrement intéressant. La croissance a donc été homogène pour tous les usages de Téo et non seulement pour une ou deux typologies d'utilisation spécifiques.

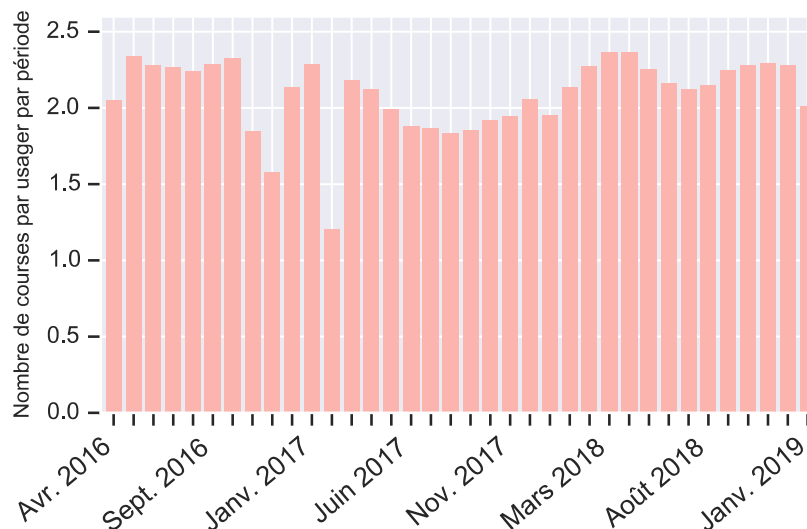


Figure 5-21 Nombre moyen de trajets par utilisateur pour chaque période

Des constats similaires peuvent être faits sur le nombre de trajets moyen par utilisateur actif et par période qui est représenté sur la Figure 5-21. Les trois mêmes périodes que précédemment (pendant les mois d’avril 2016 et février 2017) reflètent une utilisation plus faible que le reste. Le nombre moyen de courses par utilisateur et par période est de 2,1. En dehors de ces trois périodes, le nombre de courses par utilisateur est relativement stable sur l’ensemble des périodes. En excluant ces trois périodes d’utilisation plus faible, le coefficient de variation du nombre de trajets moyen par usager par mois est alors de 7,7 %.

5.4.2 Analyses des super-utilisateurs

Nous sommes également intéressés à comprendre l’utilisation par d’un utilisateur sur l’intégralité de sa durée de vie. Pour que cette analyse soit pertinente, nous avons dans un premier temps défini des « super-utilisateurs ». Cette étape est décrite dans le paragraphe 5.4.2.1. Ces super-utilisateurs regroupent les usagers qui ont une utilisation prolongée et régulière de Téo Taxi. Comme nous l’avons vu dans le Chapitre 4, un grand nombre d’usagers ont une utilisation très faible de Téo Taxi. Réaliser ce type d’analyse sur les 50 832 utilisateurs qui n’ont fait qu’une course entre avril 2016 et janvier 2019 ne serait pas pertinent.

5.4.2.1 Définition des super-utilisateurs

La définition des super-utilisateurs permet d'extraire les usagers qui ont une durée de vie relativement longue pour ne pas considérer des utilisateurs plus ponctuels de Téo. La vie d'un utilisateur est définie comme la période entre sa première et sa dernière utilisation de Téo Taxi. De plus l'utilisation par ces super-utilisateur de Téo doit être répartie sur l'intégralité de sa vie. Cette dernière condition évite de définir comme super-utilisateur un client qui aurait effectué des courses pendant la première partie et la dernière partie de sa vie uniquement.

Pour s'assurer que ces deux critères soient bien respectés, deux conditions ont été définies pour qu'un usager puisse être qualifié de super-utilisateur. Pour ce faire, un usager doit :

- Avoir une durée de vie d'au moins quatre mois (la durée de vie étant la durée entre la première et la dernière utilisation). Ce critère permet de s'assurer que la première contrainte sur la durée de vie est respectée.
- Être actif au moins la moitié des périodes dont les dates de début ou de fin sont comprises entre la date de sa première et dernière utilisation. Un utilisateur est dit actif sur une période s'il effectue au moins une course dans la période. Ce critère permet de respecter la deuxième contrainte qui est d'avoir une utilisation équilibrée dans le temps.

Ainsi un super-utilisateur devra au minimum avoir une durée de vie de quatre mois et être actif deux périodes sur ces quatre mois. Aucune condition n'a été définie sur l'intensité de l'utilisation pendant les périodes ou le nombre de courses minimum est réalisé par le client, car cela aurait pu conduire à écarter certains comportements d'utilisation. La condition sur la part de période active impose cependant que le nombre de courses faites par l'utilisateur soit supérieur ou égal à deux. La conséquence de ce critère sur le nombre de courses n'est pas limitative. Nous ne pouvons pas en effet caractériser une évolution de comportement si une seule course a été réalisée.

Tableau 5-4 Caractéristiques principales des super-utilisateurs

| | |
|--|-------------------------|
| <i>Nombre de super-utilisateurs</i> | 21 830 |
| <i>Nombre de courses moyen par usager</i> | 29,4 courses par usager |
| <i>Nombre de courses moyen par usager par période active</i> | 2,8 courses |

| | |
|---|---------------|
| <i>Durée de vie moyenne</i> | 410,7 jours |
| <i>Nombre moyen de périodes actives</i> | 10,4 périodes |
| <i>Nombre moyen de périodes inactives</i> | 4,9 périodes |

Au total, 21 830 usagers sont définis comme des super-utilisateurs, soit 16,3 % des clients. Le Tableau 5-4 regroupe les caractéristiques principales des super-utilisateurs. Ces super-utilisateurs ont réalisé 642 296 courses. Cela représente 60,1 % du total des courses réalisées. Le nombre moyen de courses par super-utilisateur est donc sans surprise supérieur au nombre de courses moyen pour l'ensemble des utilisateurs (29,4 courses contre 8,0 courses). Le nombre de courses par période active des super-utilisateurs est en revanche beaucoup plus proche du nombre de courses par période pour l'ensemble des usagers (2,8 courses pour les super-utilisateurs contre 2,1 courses pour l'ensemble des usagers-périodes). La durée de vie moyenne de ces utilisateurs est relativement longue (plus de 410 jours ou 60 semaines). Les caractéristiques des super-utilisateurs laissent à penser que cette catégorie regroupe les utilisateurs les plus intensifs des usagers de Téo bien que l'intensité d'utilisation ne soit pas un critère.

La Figure 5-22 montre la répartition des périodes de naissance et de décès des super-utilisateurs sur l'ensemble de la durée d'activité de Téo. Les deux répartitions présentent des caractéristiques très différentes. Le nombre de naissances de super-utilisateurs est très important dans les premières périodes. Il faut rappeler que pour les raisons que nous avons évoquées dans la section 4.3, les premiers mois d'enregistrement possèdent un taux de filtrage de 100 %. Le nombre plus important de naissances sur ces premières périodes peut être interprété comme une conséquence de ce taux de filtrage. Un certain nombre de super-utilisateurs utilisaient déjà Téo Taxi pendant ces premiers mois non retenus, mais sont quand même considérés comme de nouveaux utilisateurs lors de leur première utilisation de Téo sur ces périodes. La variation du nombre de naissances avant mars 2017 est importante. La croissance de Téo Taxi pendant cette époque était élevée. Cette période correspond à une période de bonification du service de Téo Taxi avec une hausse du nombre de véhicules et une hausse du nombre de permis de taxi ainsi que l'élargissement du territoire de desserte (Geneste, 2017). Cela explique donc le nombre important de naissances dans le début des activités.

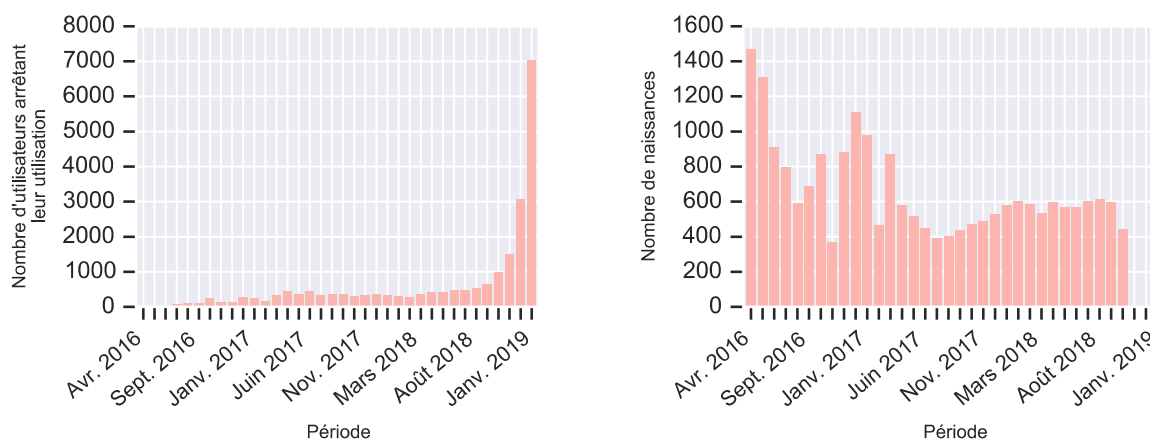


Figure 5-22 Distribution des périodes de naissance et de fin d'utilisation des super-utilisateurs

La répartition des périodes de fin d'utilisation est différente. Alors que le nombre d'utilisateur arrêtant leur utilisation de Téo est relativement faible et stable pour le début des périodes, il est beaucoup plus important pour les trois dernières périodes. Plus de 53 % des utilisateurs cessent leur utilisation de Téo durant ces trois dernières périodes.

5.4.2.2 Analyse de l'utilisation des super-utilisateurs

5.4.2.2.1 Répartition des super-utilisateurs au sein des groupes d'utilisateurs-périodes

Bien que les super-utilisateurs ne représentent que 16,3 % des utilisateurs, 44,9% des usagers-périodes construits appartiennent à des super-utilisateurs. Le nombre total d'usagers-périodes appartenant à des super-utilisateurs est de 227 543. Comparer les usagers-périodes des super-utilisateurs avec ceux de l'ensemble de la population permet de mieux comprendre les spécificités d'utilisation de ces super-utilisateurs. La Figure 5-23 présente la répartition de leurs usagers-périodes au sein des cinq groupes. La part de super-usagers-périodes au sein des groupes est présentée dans ce graphique. Puisque 44,89 % des usagers-périodes appartiennent à des super-utilisateurs, un chiffre supérieur à 44,9 % signifie une surreprésentation des super-utilisateurs dans ce groupe et un chiffre inférieur une sous-représentation. Bien que le groupe d'usage exceptionnel soit celui avec la part interne de super-utilisateurs la plus faible (32,7 %), il reste celui avec la population la plus importante. Sa population est de 95 362 usagers-périodes. Le groupe avec la surreprésentation de super-utilisateurs la plus forte est celui d'usage intensif. Au total, les 6 643 usagers-périodes des super-utilisateurs de ce groupe représentent 90,6 % de sa population. La

population des super-utilisateurs regroupe une grande partie du « petit nombre d'utilisateurs faisant un grand nombre de courses » présenté dans le Chapitre 4. La part du groupe d'usage faible est la plus proche de la part réelle de super-utilisateurs. Cela n'est pas surprenant, car l'utilisation par ce groupe est relativement proche de celle du groupe d'usage exceptionnel. De la même manière, la part du groupe d'usage moyen est relativement importante. L'usage du groupe d'usage moyen est très proche de celui de l'usage intensif.

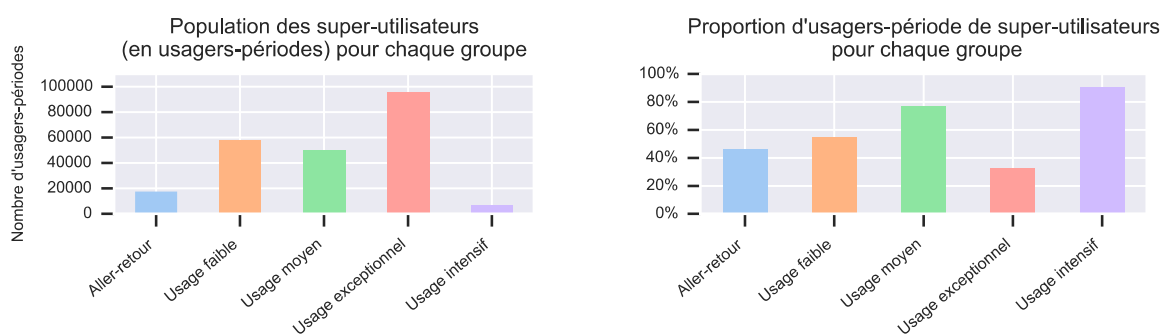


Figure 5-23 Population des super-utilisateurs (en usagers-périodes) pour chaque groupe (à gauche) et proportion d'usagers-périodes de super-utilisateurs pour chaque groupe (à droite)

5.4.2.2.2 Stabilité d'utilisation par les super-utilisateurs

Pour étudier la stabilité dans l'utilisation par les super-utilisateurs, nous avons utilisé deux indicateurs : le premier est le nombre de groupes d'usagers-périodes différents auxquels les super-utilisateurs ont été associés pendant leur vie, le deuxième est le nombre de transitions d'un groupe à l'autre divisé par le nombre de périodes actives pour chaque super utilisateur. Le Tableau 5-5 présente cet indicateur. Le nombre de groupe maximal est de cinq, car nous avons construit, dans la section 5.2, cinq groupes d'utilisations différents. Le comportement d'utilisation des super-utilisateurs n'est pas forcément régulier. Bien que 38,6 % des super-utilisateurs n'adoptent qu'un ou deux comportements d'usage différents, l'utilisation par 61,4 % d'entre eux reflète trois ou quatre comportements différents. Cet indicateur ne permet pas cependant de prendre en compte la durée de vie de l'utilisateur. Plus la durée de vie d'un usager est longue, plus la probabilité qu'il utilise le taxi de manière différente est importante.

Tableau 5-5 nombre de groupes d'utilisateurs-périodes différents pendant la durée de vie des super-utilisateurs

| <i>Nombre de groupes différents</i> | <i>Pourcentage de super-utilisateurs</i> |
|-------------------------------------|--|
| <i>1 groupe</i> | 13,6 % |
| <i>2 groupes</i> | 25,0 % |
| <i>3 groupes</i> | 28,6 % |
| <i>4 groupes</i> | 28,5 % |
| <i>5 groupes</i> | 4,3 % |

Le deuxième indicateur permet de prendre en compte la durée de vie de l'utilisateur. Une transition d'un groupe à l'autre est observée lorsque le groupe du n ème usager-période d'un utilisateur est différent du groupe de l'usager-période précédant du même utilisateur. Si un usager passe d'une période active à une période non active, aucune transition n'est observée. De la même manière nous n'avons pas compté de transition lorsqu'un utilisateur passe d'une période inactive à une période active. Le choix a été fait de s'intéresser aux périodes actives uniquement. Ce nombre permet de quantifier la stabilité du comportement d'un usager lorsque celui-ci est actif. Le nombre de transitions est ensuite divisé par le nombre de périodes actives de ce même usager. Le résultat obtenu permet ainsi de caractériser la stabilité du comportement d'utilisation par usager. Un nombre de changements par période égal à zéro signifie que l'utilisateur n'a pas changé de comportement d'utilisation tout au long de sa vie. Une valeur d'un signifie en revanche une forte instabilité dans le comportement d'utilisation. Pour ces usagers, les comportements d'utilisation d'une période à l'autre sont tous différents. Le Tableau 5-6 présente les caractéristiques de la distribution de cet indicateur. Nous pouvons voir sur la Figure 5-24 qu'un nombre important de super-utilisateurs a un comportement d'utilisation très stable (nombre de changements par période inférieur à 0,1). Plus de la moitié des utilisateurs possède également un nombre de changement d'utilisation supérieur à 0,5, ce qui signifie que le comportement d'utilisation change en moyenne toutes les deux périodes ou moins.

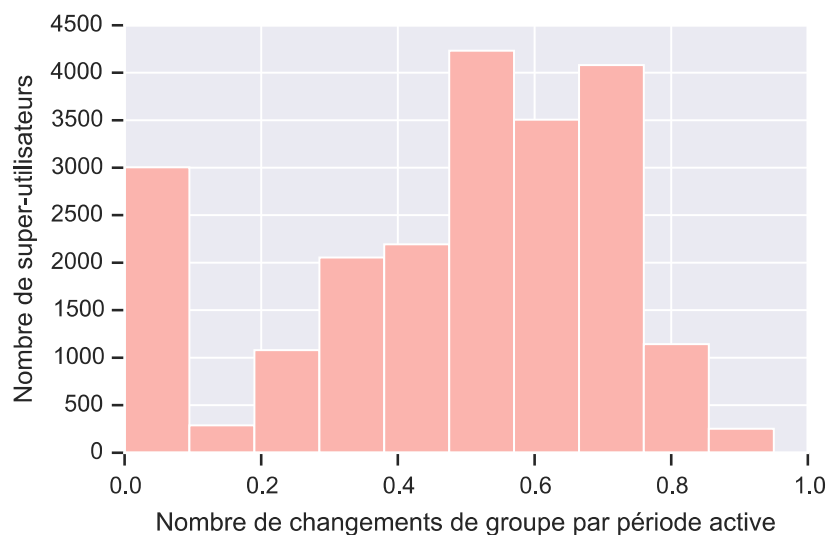


Figure 5-24 Distribution du nombre de changements de groupe par période active

Tableau 5-6 Statistiques descriptive du nombre de changements de groupe par période active

| | |
|--------------------------------|-------|
| <i>Moyenne</i> | 0,468 |
| <i>Écart type</i> | 0,340 |
| <i>Minimum</i> | 0,000 |
| <i>1^{er} quartile</i> | 0,333 |
| <i>Médiane</i> | 0,500 |
| <i>3e quartile</i> | 0,667 |
| <i>Maximum</i> | 1.00 |

5.5 Synthèse de l'utilisation de Téo Taxi

Nous avons pu voir dans ce chapitre la variabilité importante dans les comportements des utilisateurs à la fois sur une période de quatre semaines et sur l'intégralité de leur durée de vie. Concernant l'utilisation de Téo Taxi sur une période de quatre semaines, cinq typologies d'utilisations ont été construites :

- Groupe des aller-retours;

- Groupe d'usage faible;
- Groupe d'usage moyen;
- Groupe d'usage exceptionnel;
- Groupe d'usage intensif.

Des différences importantes existent dans l'utilisation de Téo Taxi pour chacun de ces groupes. Cependant la répartition de ces cinq comportements d'utilisation de Téo Taxi est relativement stable dans le temps. De même le nombre de courses réalisées par utilisateur durant une période de quatre semaine est également stable. La forte croissance de Téo Taxi ne semble avoir eu qu'un impact limité sur le comportement individuel de chaque utilisateur. Une petite partie des utilisateurs utilisaient également Téo Taxi de manière fréquente. L'utilisation de Téo Taxi faite par ces clients était plus intense que celle de l'ensemble des utilisateurs.

CHAPITRE 6 ENQUÊTE AUPRÈS DES UTILISATEURS

Comme il a été introduit dans le Chapitre 2, les enquêtes auprès des utilisateurs sont des outils utilisés pour connaître les utilisateurs d'un service de taxi. Dans ce chapitre, les résultats d'une enquête menée auprès des utilisateurs de Téo Taxi sont détaillés. De plus, les particularités de Téo Taxi nous permettent de pouvoir mettre en regard les réponses à cette enquête avec l'utilisation que les répondant ont du service de taxi.

6.1 Modalités méthodologiques

L'enquête a duré deux semaines. Elle a été menée du lundi 3 décembre 2018 à 14h au lundi 17 décembre 2018 à 14h, soit durant une période de deux semaines. Pour répondre au questionnaire, le répondant devait remplir deux critères : avoir effectué au moins une course en la réservant par l'application mobile de Téo Taxi pendant la période d'enquête et être âgé de 18 ans ou plus.

6.1.1 Prise de contact avec le répondant

La prise de contact avec le répondant s'est faite par l'intermédiaire d'un courriel. Ce courriel a été envoyé soit en français, soit en anglais en fonction de la langue paramétrée par le client dans les réglages de l'application mobile. Les versions anglaise et française de ce courriel sont présentées dans la Figure 6-1. Le courriel a été envoyé cinq minutes après l'envoi du courriel de facturation de la course au client par le système d'envoi de courriel interne de Téo Taxi. Nous n'avons pas, dans cette enquête, eu accès aux adresses courriel des répondants ou à toute information permettant d'identifier formellement les utilisateurs. Le client pouvait alors cliquer directement sur le lien présent dans le message pour accéder à la plateforme d'enquête. Pour que nous puissions ensuite relier la réponse à la course réalisée et à l'utilisateur concerné, les informations sur le numéro de la course, le lieu de départ ainsi que la date de départ nous étaient transmises via l'adresse URL du lien de réponse. Si un client effectuait plusieurs courses pendant la période de réponse, une seule invitation à répondre lui était envoyée (celle pour la première course qu'il a réalisée pendant la période de réponse).

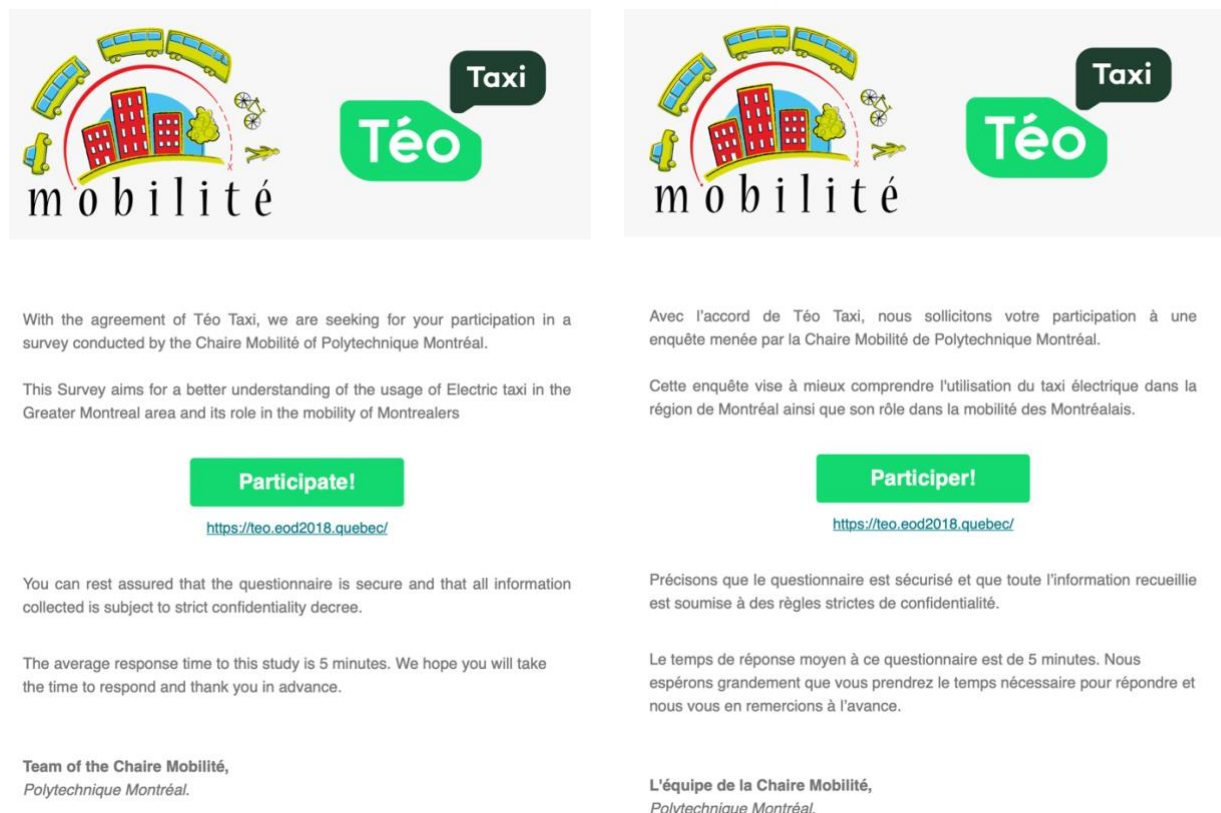


Figure 6-1 Courriel de prise de contact avec le répondant (en anglais à gauche et en français à droite)

6.1.2 Présentation de la plateforme utilisée

Le questionnaire a été réalisé sur internet. Nous avons utilisé pour cela le module de sondage Évolution de la plateforme Transition actuellement en développement (<http://transition.city/>). L'enquête était disponible en français et en anglais. Le répondant avait la possibilité de changer la langue à tout moment de l'enquête. En cliquant sur le lien présent dans le courriel, le répondant était redirigé vers la page d'accueil présentée dans l'annexe C. Aucune identification de la part du répondant n'était requise, l'identification était faite de manière automatique à partir des informations envoyées via l'adresse URL. Chaque répondant ne pouvait répondre qu'une seule fois à l'enquête. La Figure 6-2 est un aperçu de la plateforme de réponse une fois que le répondant avait débuté l'enquête.

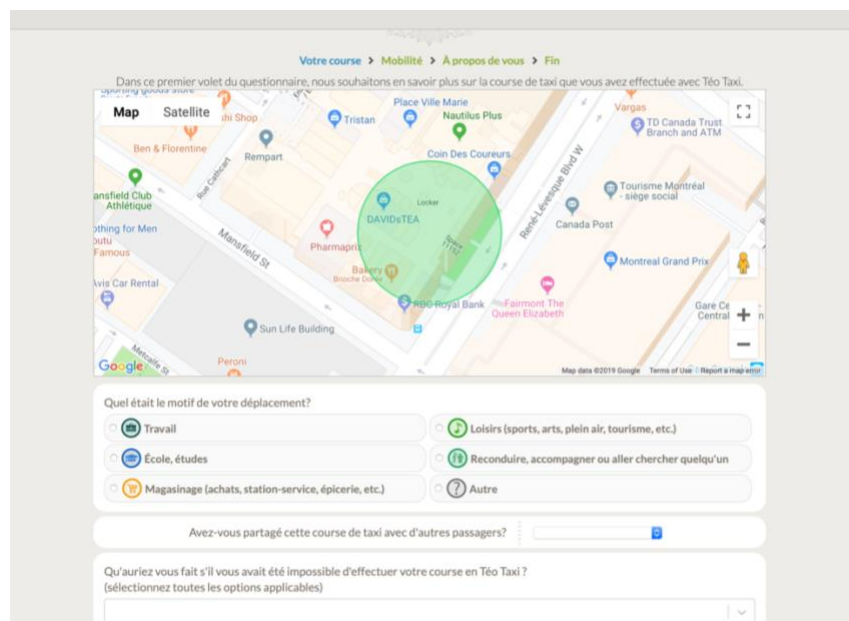


Figure 6-2 Aperçu de la plateforme de sondage

6.1.3 Politique de confidentialité

Comme nous pouvons le voir sur la présentation de la page d'accueil du sondage disponible dans l'annexe C, les répondants devaient accepter les conditions générales d'information et de consentement relatives à ce projet. Le formulaire était accessible depuis cette page via des liens hypertextes et était disponible en français et en anglais. La version française des conditions générales d'information et de consentement est disponible dans l'annexe D. Ce formulaire permettait aux participants de connaître les enjeux du projet de recherche dans lequel s'inscrit ce sondage ainsi que d'obtenir des informations sur les données récoltées. Comme nous l'avons mentionné précédemment, nous avons été attentifs à ce qu'aucune donnée permettant d'identifier formellement l'utilisateur ne soit récoltée. Bien que les réponses soient reliées au compte utilisateur du client, ces dernières restent anonymes.

6.2 Construction du questionnaire

Durant l'élaboration du questionnaire, nous avons été particulièrement attentifs à ce que le questionnaire soit relativement court et facile à remplir à la fois depuis un ordinateur ou depuis des plateformes nomades comme des téléphones intelligents ou des tablettes. Pour cela, nous avons limité au maximum l'usage de questions nécessitant un remplissage manuel par l'utilisateur en

privilégiant des formats de réponses avec des options à cocher. Les questions du questionnaire peuvent être regroupées en trois catégories : les questions relatives à la course effectuée, les questions relatives à la mobilité du répondant et les questions permettant de dresser le portrait socio-économique du répondant. Au total, le sondage comprend 18 questions :

1. Quel était le motif de votre déplacement ?
2. Avez-vous partagé cette course de taxi avec d'autres passagers ?
 - a. (Si réponse positive à 2) Est-ce que tous les passagers ont terminé leur course au même endroit ?
3. Qu'auriez-vous fait s'il vous avait été impossible d'effectuer votre course en Téo Taxi ?
 - a. (Si autre mode de transport choisi à 3) Quel(s) mode(s) de transport auriez-vous été susceptible de choisir ?
4. Quelles sont les raisons qui vous ont poussé à effectuer ce déplacement avec Téo Taxi ?
5. Avez-vous utilisé uniquement le taxi comme mode de transport pour vous rendre à votre destination finale ?
6. Qui a payé la course ?
7. Sélectionnez les services dont vous êtes membre / abonné ou que vous utilisez :
8. À quelle fréquence avez-vous pris le taxi (autre que Uber ou Téo Taxi) dans les six derniers mois ?
9. À quelle fréquence avez-vous utilisé Uber dans les six derniers mois ?
10. Connaissez-vous l'application mobile Transit ?
11. Quel âge avez-vous ?
12. Quel est votre occupation principale ?
13. Comment vous identifiez-vous ?
14. Possédez-vous un permis de conduire ?
15. Résidez-vous au Canada ?
 - a. (Si 15 positif) Code postal de votre domicile :

16. Combien de personnes habitent votre domicile de façon permanente, y compris vous-même ?

17. Combien de véhicules sont à la disposition d'un ou des membres de votre ménage ?

18. Quelle était la tranche de revenu de votre ménage avant impôt en 2017 ?

Les choix de réponses possibles sont détaillés dans la section 6.3.2.1. dans cette partie, la présentations des réponses à chaque question est faite.

La réponse à chacune des questions est optionnelle. Cela permet de réduire le taux d'abandon des répondants si ces derniers ne désirent pas nous transmettre la réponse concernant une des questions. Certaines questions contiennent également l'option de réponse « Je ne souhaite pas répondre ». Cette catégorie de réponses a été traité différencié d'une non-réponse, dans la suite cette option est désignée par le label « refus ». La question n°10 « Connaissez-vous l'application Transit ? » est utilisé comme un marqueur de la connaissance générale du répondant sur les différentes alternatives à sa disposition pour effectuer un déplacement. Cette application donne des informations sur les modes de transport disponibles autour de soi et leurs horaires et permet le calcul de chemin multimodal. Elle est développée par une compagnie dont le siège social est basé à Montréal. L'utilisation de cette application ainsi que l'intégration des différentes alternatives de transport sont importantes dans la région de Montréal (Transit, 2019).

6.3 Résultats

6.3.1 Nombre de réponses

Pendant la période de récolte de données, 25 780 courses ont été effectuées par 12 755 clients différents.

Lors de la phase de récolte, 1 178 entrevues ont été complétées entièrement sur les 1 361 qui ont été commencées. Cela correspond à un taux de complétion de 87%. À titre comparatif, le taux de complétion d'enquêtes sur la mobilité plus longues comme la version web de l'enquête Origine-Destination régionale par exemple est autour de 60%). Le taux de réponse de l'enquête est de 9 % (un courriel a été envoyé à chacun des 12 755 utilisateurs). Nous avons considéré une entrevue comme complétée lorsque le répondant se rendait à la dernière section du sondage.

Le temps moyen de réponse des répondants ayant complété l'enquête dans son intégralité est de 251 s (un peu plus que quatre minutes).

6.3.2 Représentativité de la population

Avant de commencer l'analyse des résultats de l'enquête, il est important de préciser que toutes les données présentées ne sont pas pondérées et que la population des répondants ne peut être considérée comme représentative. Nous avons vu dans la section 2.4, qu'il est généralement très difficile de connaître la population des utilisateurs du taxi. Cela est également le cas dans la région de Montréal. Il est en effet impossible de s'assurer de la représentativité des réponses ou de définir des critères de pondération si la population de référence est inconnue.

Il faut donc rester très prudent dans l'interprétation des résultats de l'enquête. Il est impossible de dire avec certitude que ces résultats peuvent décrire de manière globale l'utilisation de Téo Taxi.

6.3.2.1 Présentation des réponses

Pour présenter les résultats du sondage, nous avons regroupé les questions posées au répondant selon les trois thématiques abordées (caractéristiques de la course, habitudes de mobilité du répondant et profil socio-économique). Les résultats présentés dans cette section sont ceux de l'ensemble des répondants qu'ils aient terminé l'entrevue ou non. Pour chaque question, le nombre de réponses recueillies est inscrit. Ce nombre peut être supérieur ou inférieur au nombre d'entrevues complétées, car d'une part les réponses des répondants n'ayant pas complété intégralement l'enquête sont comptabilisées et d'autre part les répondants qui ont terminé l'entrevue ont pu choisir de ne pas répondre à certaines questions.

Afin d'améliorer la compréhension des résultats, les graphiques suivants présentent quelques particularités. Tout d'abord, la couleur du graphique permet de différencier le type de question. Les histogrammes dont la couleur est bleue (Figure 6-3) représentent des questions où le répondant ne pouvait sélectionner qu'une seule option de réponse. Les histogrammes de couleur verte (Figure 6-5) représentent des questions à choix multiples. Dans la plateforme d'enquête, il était également clairement précisé que les répondants devaient choisir toutes les réponses applicables pour ces questions. Lorsque cela était cohérent, les réponses ont également été triées par ordre de décroissance d'apparition. Ainsi, les options de réponse présentes à gauche du graphique sont les options les plus sélectionnées et celles à droite les moins sélectionnées.

6.3.2.2 Caractéristiques de la course

La première section comptabilise au total six questions (questions de 1 à 6 de la liste présentée dans le paragraphe 6.2). Cette partie permet de mieux comprendre pourquoi et comment les utilisateurs utilisent Téo Taxi.

La Figure 6-3 présente les motifs des déplacements effectués par les répondants. La raison de déplacement la plus récurrente est le motif professionnel (42,2 % des réponses). Le motif retour au domicile est absent de la liste des réponses possibles, car dans la cas d'une course dont la destination est le domicile du répondant, nous souhaitons connaître le motif pour lequel le répondant c'est rendu au lieu de départ de la course. Le motif professionnel est particulièrement important comparé aux déplacements de la région de Montréal. En 2013, le nombre de déplacements réalisés pour motif professionnel (sans le motif retour à domicile) est de 32 % (Exo, 2015).

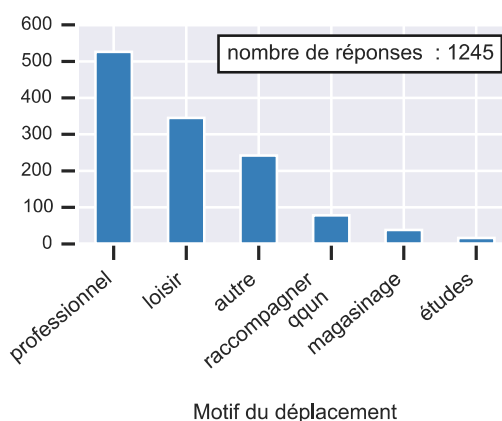


Figure 6-3 Distribution des réponses à la question « Quel était le motif de votre déplacement ? »

La Figure 6-4 permet de connaître le nombre d'utilisateurs également présents dans le taxi. Plus de 30 % des utilisateurs ont partagé leurs courses avec une ou plusieurs personnes. Le nombre de passagers maximal dans les véhicules de Téo Taxi étant de quatre (un utilisateur ne peut pas partager sa course avec plus de trois personnes). Dans le cas d'une course de taxi partagée, la question « Est-ce que tous les passagers ont terminé leur course au même endroit ? » était également posé au répondant. La réponse à cette question est également disponible sur la Figure 6-4. Une grande partie des utilisateurs terminent leur course au même endroit (83,2 % des réponses).

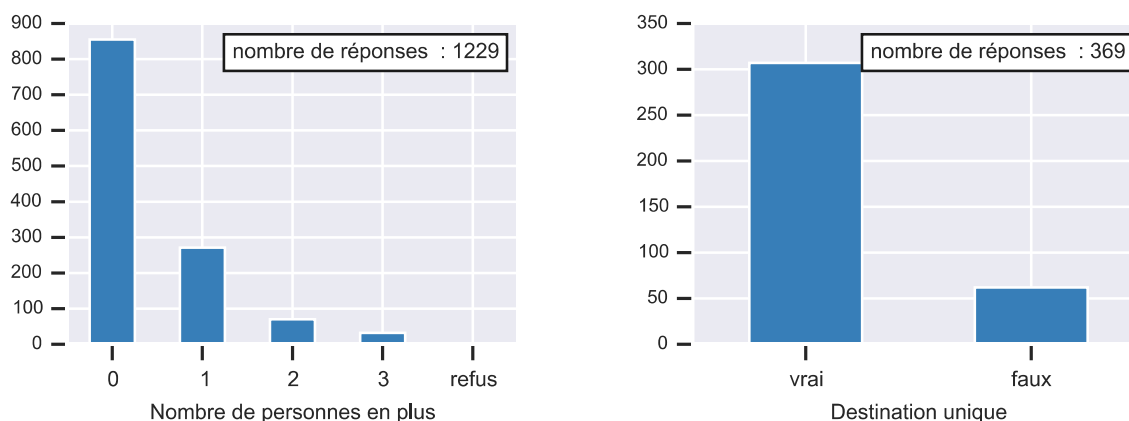


Figure 6-4 Distribution des réponses aux questions « Avez-vous partagé cette course de taxi avec d'autres passagers ? » (à gauche) et « Est-ce que tous les passagers ont terminé leur course au même endroit ? » (à droite)

Comme il a été mentionné dans le Chapitre 2, la question du rôle du taxi renvoie à la concurrence et à la complémentarité entre le taxi et les autres modes de transport. Cette question est double avec Téo Taxi; le fonctionnement particulier de l'entreprise fait que Téo Taxi peut également être perçu comme un service différent du taxi régulier. La Figure 6-5 montre cependant que l'alternative principale à Téo Taxi est l'utilisation d'un service de taxi traditionnel. Bien que les répondants eussent la possibilité de choisir plusieurs réponses, 55,98 % d'entre eux ont sélectionné comme unique réponse qu'ils auraient pris un autre service de taxi. Au total, 67,4 % des utilisateurs ont coché cette option. Pour 21,9 % des répondants, le taxi est utilisé comme une alternative à d'autres modes de transports. Enfin, seulement 34 répondants auraient choisi de changer les caractéristiques de leur déplacement (changement d'heure, changement de destination ou même annulation du déplacement). La Figure 6-6 permet de connaître les autres modes envisagés. Dans 76;9 % des cas, le transport en commun est un des modes alternatifs envisagés.

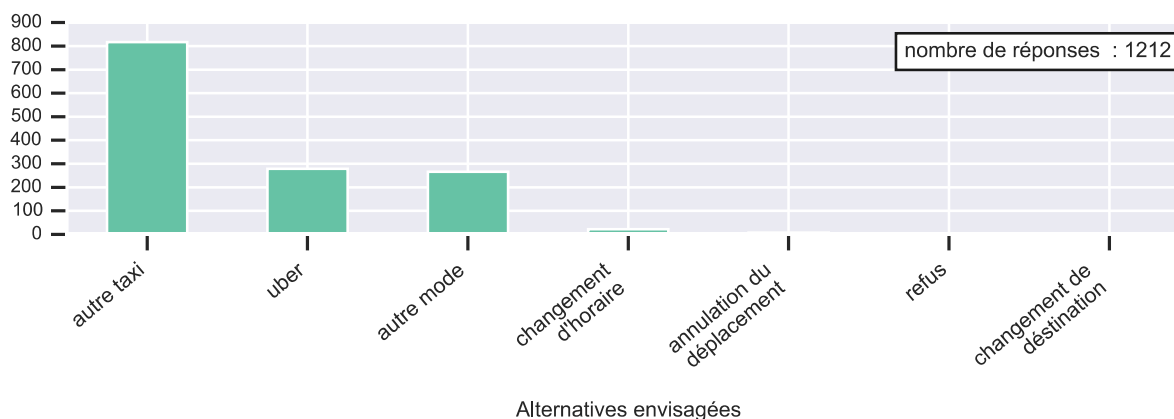


Figure 6-5 Distribution des réponses à la question « Qu'auriez-vous fait s'il vous avait été impossible d'effectuer votre course en Téo Taxi ? »

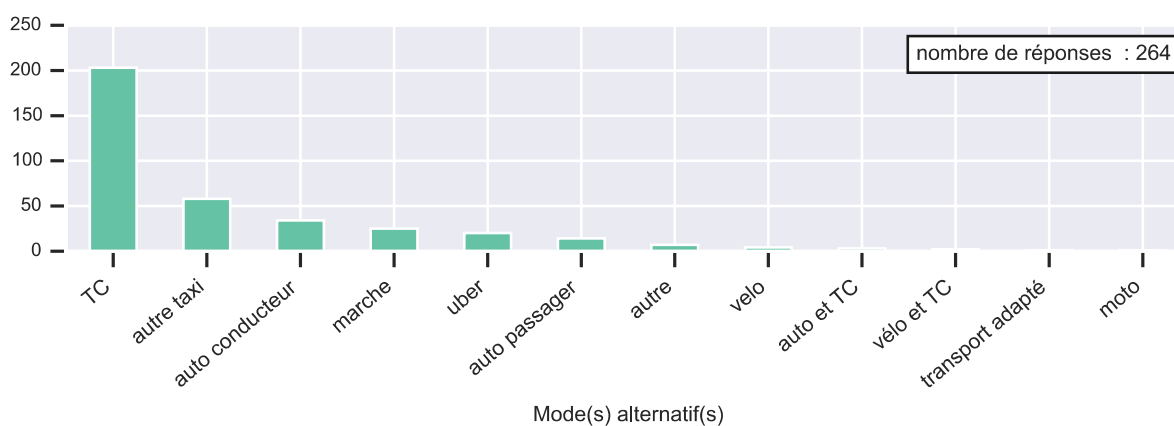


Figure 6-6 Distribution des réponses à la question « Quel(s) mode(s) de transport auriez-vous été susceptible de choisir? »

La Figure 6-7 permet également d'explorer la complémentarité entre les modes. Il a été dit dans le Chapitre 2 qu'un exemple de complémentarité entre le taxi et les autres modes de transport est d'utiliser le taxi comme un moyen de rabattement vers les autres modes. Cependant, dans notre enquête, le trajet en taxi s'inscrit dans un déplacement multimodal plus complexe pour seulement 12 % des répondants.

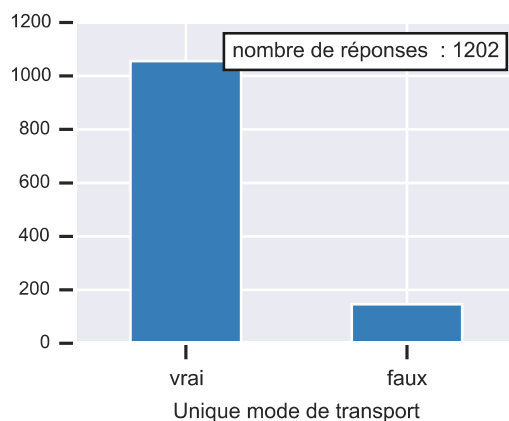


Figure 6-7 Distribution des réponses à la question « Avez-vous utilisé uniquement le taxi comme mode de transport pour vous rendre à votre destination finale ? »

Un autre moyen de connaître la différenciation que les utilisateurs font entre Téo Taxi est les services de taxi traditionnel (et donc la concurrence ou non entre ces deux services) est de connaître les raisons d'utilisation qui ont poussé les utilisateurs à choisir Téo Taxi plutôt que le taxi traditionnel. La Figure 6-8 montre que les trois premières raisons de choix de Téo Taxi sont spécifiques à ce service. Ces trois raisons ont été sélectionnées par respectivement 53,9 %, et 48,5 % et 43,5 % des répondants.

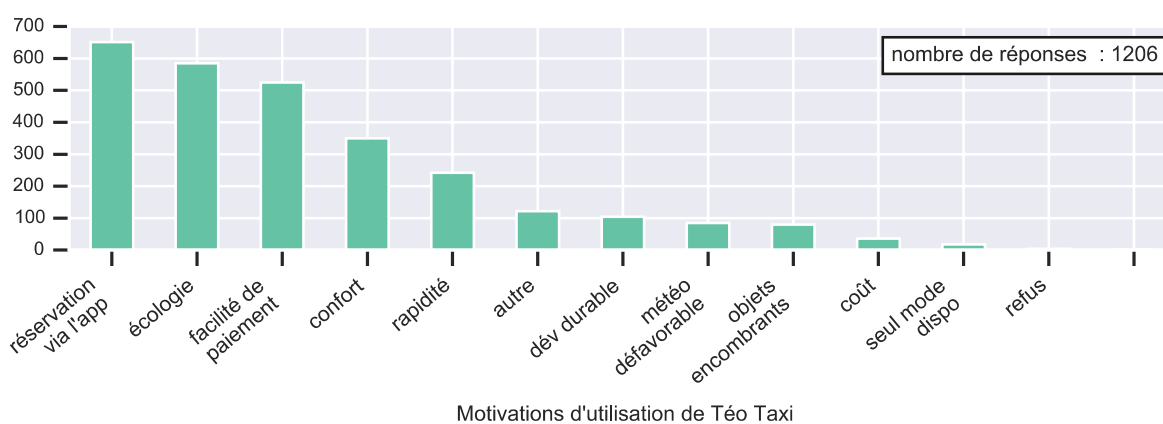


Figure 6-8 Distribution des réponses à la question « Quelles sont les raisons qui vous ont poussé à effectuer ce déplacement avec Téo Taxi ? »

Enfin la Figure 6-9 permet de savoir qui a payé la course de taxi. Si cette question est triviale dans le cas d'un déplacement de type loisir où le passager est seul dans le taxi, elle l'est moins dans le

cas d'une course pour raisons professionnelles (où l'employeur peut payer) et dans le cas d'une course partagée (où l'ensemble des frais peut être réparti). Dans le cas d'une course réalisée pour motif professionnel, l'employeur a payé pour la course dans 53 % des cas. Dans le cas d'une course partagée, il est rare que les frais soient partagés. Seulement 2 % des répondants ayant partagé leur course ont également réparti les frais.

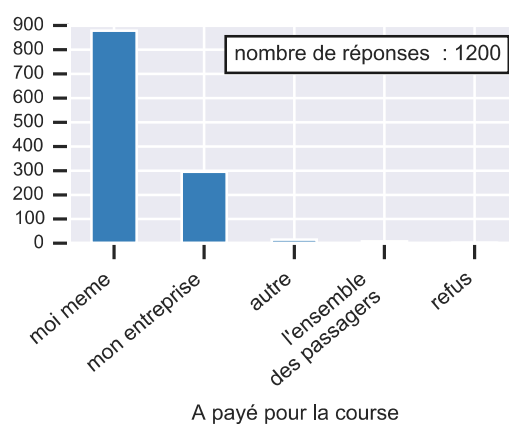


Figure 6-9 Distribution des réponses à la question « Qui a payé la course ? »

6.3.2.3 Mobilité du répondant

La section concernant les habitudes de mobilité du répondant regroupe quatre questions (de 7 à 10 de la liste présentée dans le paragraphe 6.2).

La Figure 6-10 permet de connaître les services de mobilité auxquelles les répondants sont abonnés ou qu'ils utilisent. Une part importante des répondants (43 %) possèdent un abonnement mensuel ou annuel de transport en commun. Une part également importante des utilisateurs sont inscrits au service de mobilité à la demande Uber (30 %). Les services car2go et Communauto sont des services de véhicules en libre-service implantés sur l'île de Montréal. Bixi est le service de vélo en libre-service basé station de la région de Montréal. Une part non négligeable des répondants n'utilisent aucun de ces services (23,7 %).

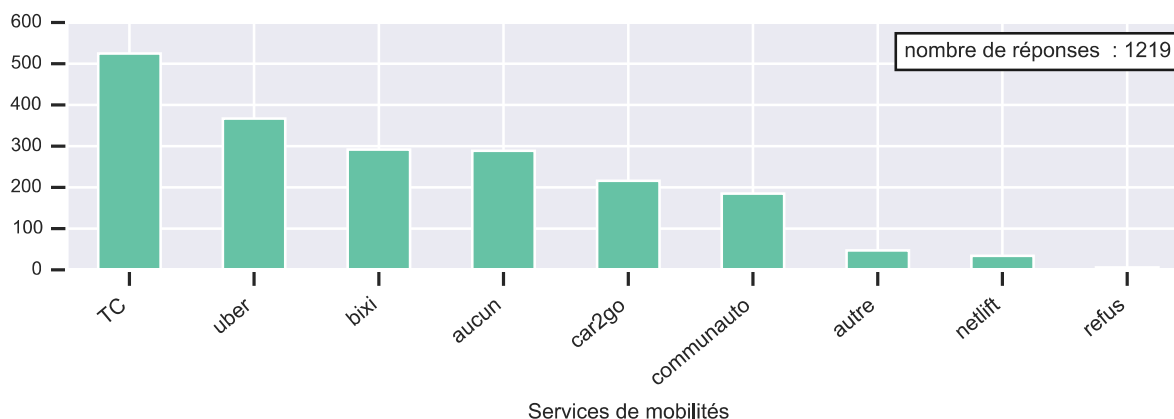


Figure 6-10 Distribution des réponses à la question « Sélectionnez les services dont vous êtes membre / abonné ou que vous utilisez : »

La Figure 6-11 permet de comparer la fréquence d'utilisation des services de taxi régulier et de Uber au cours des six derniers mois. Cette différence est très importante, alors que seulement 17 % des répondants n'ont pas utilisé un service de taxi régulier ces six derniers mois, ce chiffre est de 64 % pour l'utilisation d'Uber. La faible utilisation d'Uber n'est pas surprenante si elle est mise en parallèle avec la Figure 6-10. Seuls 30 % des répondants ont déclaré avoir un compte sur l'application mobile d'Uber. Cette faible utilisation est cependant plus étonnante si on la compare avec les raisons pour lesquelles les utilisateurs choisissent Téo Taxi. Deux des trois premières raisons de l'utilisation (à savoir la possibilité de réserver par une application et la facilité de paiement grâce à l'application également) sont également des fonctionnalités présentes dans les services offerts par Uber.

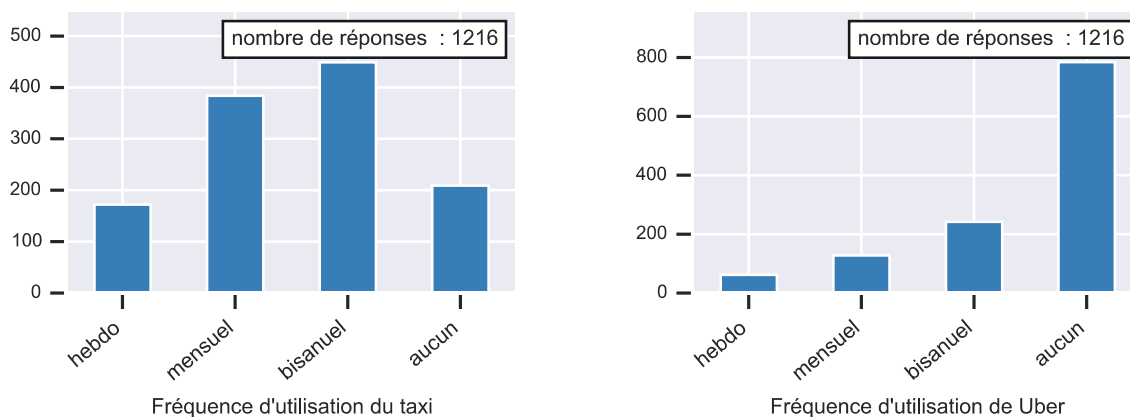


Figure 6-11 Distribution des réponses aux questions « À quelle fréquence avez-vous pris le taxi (autre que Uber ou Téo Taxi) dans les six derniers mois ? » (à gauche) et « À quelle fréquence avez-vous utilisé Uber dans les six derniers mois ? » (à droite)

Enfin, la Figure 6-12 nous permet d'estimer le niveau d'information que les personnes sur leurs alternatives de mobilité. Comme nous l'avons mentionné, la connaissance de l'application mobile Transit est utilisée comme un indicateur du niveau de connaissance des répondants en matière de mobilité.

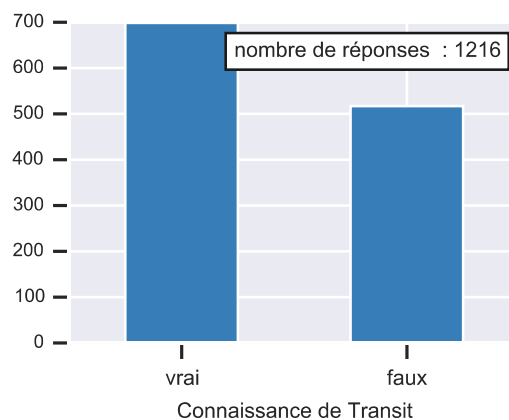


Figure 6-12 Distribution des réponses à la question « Connaissez-vous l'application mobile Transit ? »

6.3.2.4 Profil socio-économique des répondants

Cette dernière section du questionnaire permet de définir un portrait plus précis des répondants et de son ménage. Cette section regroupe les questions 11 à 18 de la liste présentée dans le paragraphe 6.2.

Nous pouvons voir sur la Figure 6-13 que la population des répondants est relativement jeune (58,9 % des répondants ont 44 ans ou moins). L'âge moyen des répondants est 42,8 ans, et l'âge médian est 41 ans. L'âge médian des habitants dans la région de Montréal âgés de 20 ans à 84 ans est entre 35 ans et 39 ans (Ville de Montréal, 2018). La Figure 6-14 présente la répartition de la population montréalaise en fonction de son âge en 2016. On peut observer chez les utilisateurs de Téo Taxi, une plus forte proportion d'utilisateurs entre 30 ans et 45 ans comparativement à la population montréalaise. On remarque également un nombre plus important d'utilisateurs masculins (51,8 % contre 46,6 %). Cette observation n'est pas propre à Téo Taxi, mais est fréquente dans beaucoup de systèmes de transport par taxi pour lesquels des études sur la population des utilisateurs ont été faites comme le montre le paragraphe 2.4.2.1. À l'échelle de Montréal, la proportion de femmes dans la population était en 2016 de 51,3% et de 48,7 % pour les hommes (Ville de Montréal, 2018).

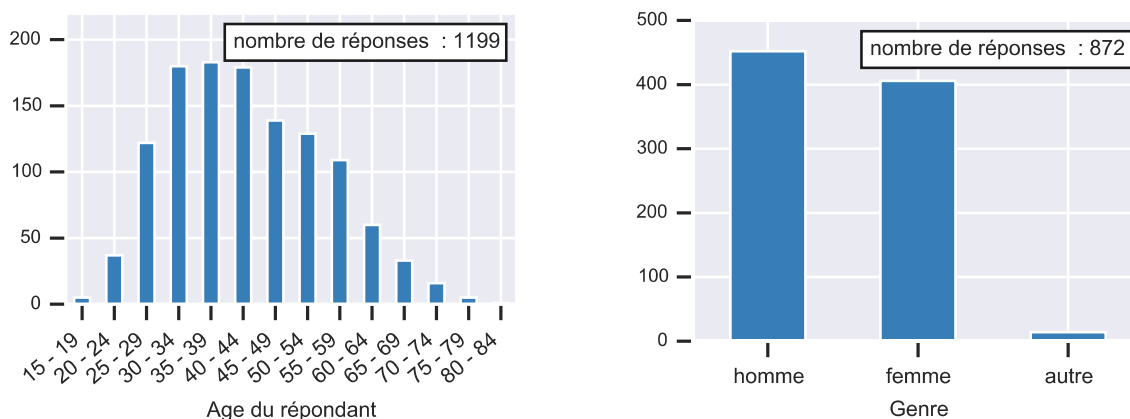


Figure 6-13 Distribution des réponses aux questions « Quel âge avez-vous ? » (à gauche) et « Genre : » (à droite)



Figure 6-14 Répartition de la population montréalaise âgée de 20 ans à 84 ans en fonction de l'âge en 2016 (données issues de (Ville de Montréal, 2018))

La Figure 6-15 montre la distribution de l'occupation principale des répondants. On peut observer que les répondants sont très majoritairement des travailleurs à temps plein (85,1 %). Ce chiffre est élevé puisque le taux d'emploi des personnes âgées de 15 ans et plus de la ville de Montréal était de 58,3% en 2016 (Ville de Montréal, 2018). La deuxième catégorie la plus fréquente est celle des retraités avec 3,5 % des réponses. Les étudiants (travail temps plein, travail temps partiel et travail et étude) représentent 3,3 % des répondants.

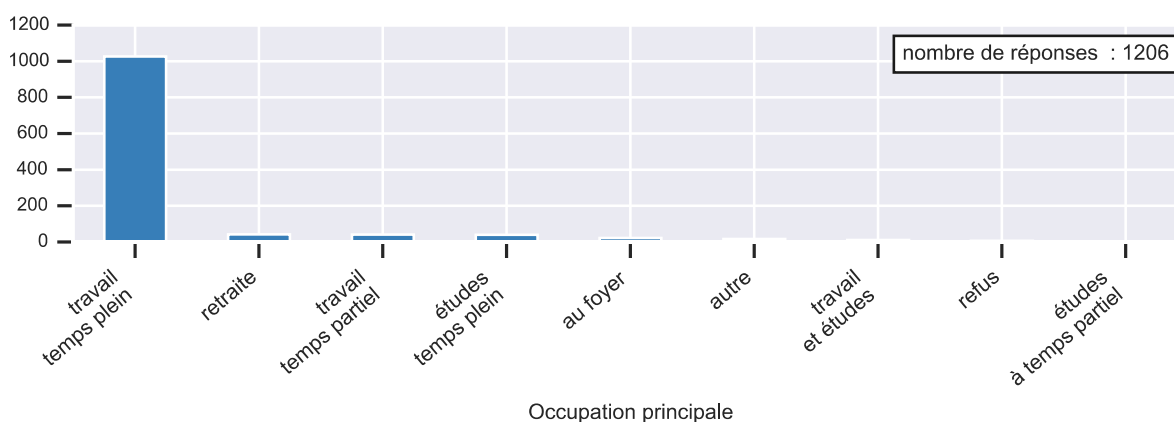


Figure 6-15 Distribution des réponses au question « Quel est votre occupation principale ? »

Le Figure 6-16 montre que le taux de possession du permis de conduire est important (90 %). Le taux de possession automobile est également relativement proche de celui observé sur l'île de Montréal en 2013. Alors que 30,9 % des ménages des répondants ne possèdent pas de voitures,

46,4 % ont un véhicule à leur disposition et 18,7 % en ont deux. Ces deux chiffres étaient sur l'île de Montréal en 2013 respectivement de 29,5 %, 47,4 % et 18,6 % (Exo, 2015).

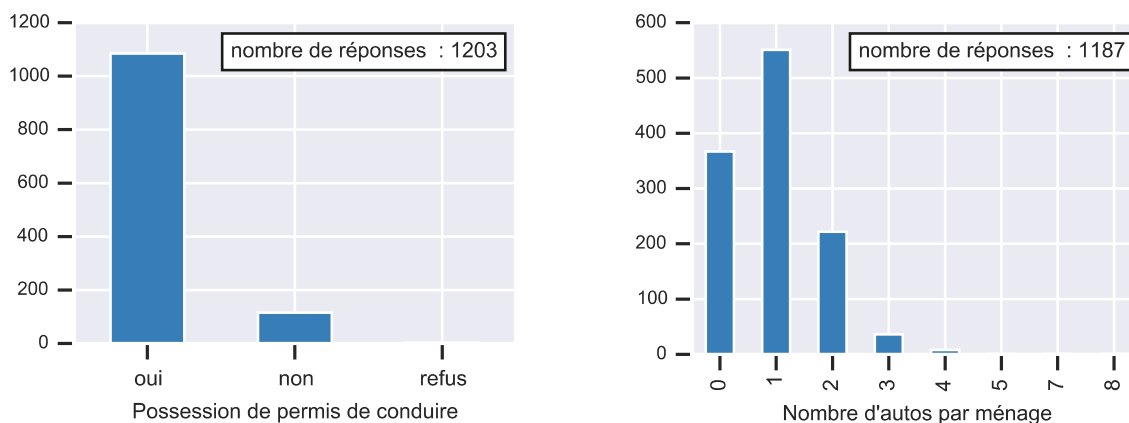


Figure 6-16 Distribution des réponses aux questions « Possédez-vous un permis de conduire ? » (à gauche) et « Combien de véhicules sont à la disposition d'un ou des membres de votre ménage ? » (à droite)

La Figure 6-17 est la carte des lieux de domicile des répondants. On remarque que la très grande majorité des répondants habite dans les quartiers centraux de la ville. Sur les 1171 répondants qui ont renseigné leur code postal, 1048 (soit 89,5 %) habitent sur l'île de Montréal. Seulement 13 répondants ont déclaré ne pas habiter au Canada.



Figure 6-17 Carte de lieux de domicile des répondants

Le nombre de personnes par ménage, présenté dans la Figure 6-18, est quant à lui légèrement plus élevé que pour l'île de Montréal. Ce nombre est de 2,43 personnes par ménage pour les répondants contre 2,22 pour l'île de Montréal) (Exo, 2015). Le portrait de la classe de revenu par ménage (Figure 6-19) est très différent de la répartition observée dans la région de Montréal. La classe de revenu la plus représentée au sein des répondants est la classe la plus haute (210 000 \$ et plus). Cette classe regroupe 18,2 % des répondants.

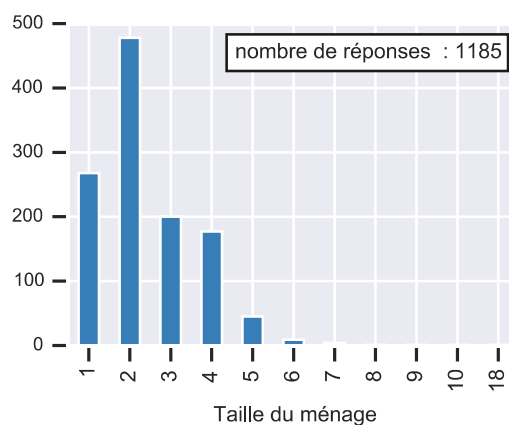


Figure 6-18 Distribution des réponses à la question « Combien de personnes habitent votre domicile de façon permanente, y compris vous-même ? »

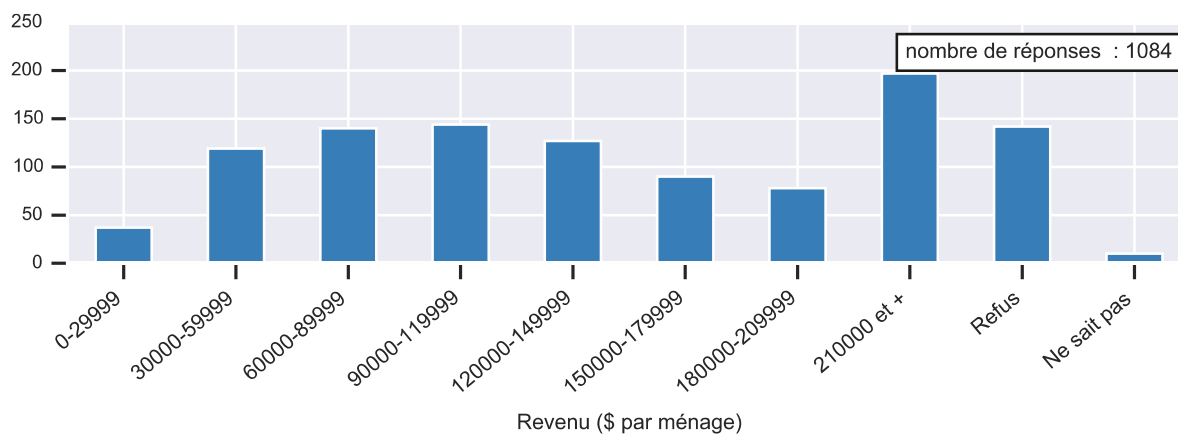


Figure 6-19 Distribution des réponses à la question « Quelle était la tranche de revenu de votre ménage avant impôt en 2017 ? »

Sur la Figure 6-20, nous pouvons également voir la répartition des revenus par ménage à Montréal en 2016. Les différences entre ces deux distributions sont importantes. D'une part les ménages aux revenus les plus faibles (59 999\$ et moins par année) sont relativement moins nombreux chez les utilisateurs de Téo et d'autre part les ménages aux revenus les plus élevés (210 000\$ et plus) sont beaucoup plus fréquents chez les répondants. Il faut cependant être prudent en comparant ces deux distributions car le niveau de résolution l'enquête est la personne alors que c'est le ménage pour la Figure 6-20. Cela signifie que deux personnes habitant dans le même ménage pourraient avoir répondu à l'enquête.

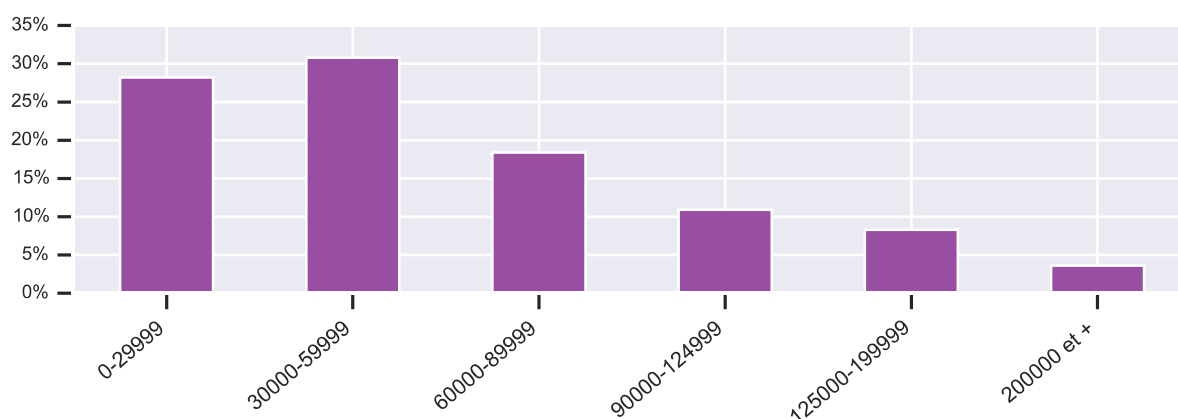


Figure 6-20 Répartition de la population montréalaise en fonction de revenue annuel par ménage en 2016 (données issues de (Ville de Montréal, 2018))

Les différences remarquées notamment sur le revenu entre le profil socio-économique des répondants et la population générale de Montréal marquent bien la particularité des utilisateurs de Téo Taxi.

6.3.3 Analyses croisées des résultats

6.3.3.1 Test du χ^2

Dans cette section, nous allons faire une analyse plus fine des résultats du sondage en regardant quelles variables ont une influence sur les réponses des utilisateurs. Il a été possible de relier les réponses du sondage au numéro de la course pour laquelle le courriel d'invitation a été envoyé et donc également à l'identifiant du client qui a réalisé cette course. Ce lien entre les réponses et les clients permet également de relier les caractéristiques du client (super-utilisateur, intensité d'utilisation, etc.) et de la course (heure de début et de fin) à chacune des réponses de l'enquête.

Pour vérifier l'indépendance des distributions en fonction d'une variable, nous avons utilisé le test du χ^2 . Ce test permet de comparer deux distributions et d'estimer la significativité de la différence entre ces deux distributions. Ce test permet pour un seuil de risque acceptable de rejeter ou non l'hypothèse nulle (le risque de première espèce est la probabilité de se tromper en rejetant l'hypothèse nulle). L'hypothèse nulle pour ce test est que les échantillons ont été générés par la même distribution. La condition pour rejeter l'hypothèse nulle est que la valeur ou statistique du test du χ^2 obtenue soit supérieure à une valeur seuil. Si la valeur du test du χ^2 est supérieure à la valeur seuil, on peut rejeter l'hypothèse nulle. Cette valeur dépend du nombre de degrés de liberté de la distribution du χ^2 suivie par la statistique du test sous l'hypothèse nulle ainsi que du risque de première espèce que nous considérons comme acceptable. Le nombre de degrés de liberté est $d = n - 1 - p$ ou n est le nombre de catégories de la distribution et p est le nombre de paramètres de la distribution de référence. Dans le cas d'une distribution de référence empirique, le nombre de paramètres de la distribution est nul. La table de valeurs seuil pour le test est disponible dans l'annexe E.

Avant d'appliquer le test du χ^2 , il faut vérifier que le nombre d'observations pour chaque valeur de la variable aléatoire soit supérieur à cinq. Si cela n'est pas le cas, il convient de regrouper des

classes jusqu'à ce que le nombre d'observations de la nouvelle classe obtenue soit supérieur ou égal à cinq.

Plus la valeur du test χ^2 est importante, plus la différence entre les deux distributions comparées est grande. Une valeur du test inférieure au seuil nous indique que nous ne pouvons pas rejeter l'hypothèse d'identité des distributions. Cependant, il est possible que, même lorsque l'hypothèse nulle est rejetée, ces deux distributions présentent des caractéristiques proches, surtout lorsque le nombre d'observations est grand.

6.3.3.2 Influence de l'utilisation sur les réponses

Pour cette première analyse, nous nous sommes intéressés à l'influence de l'utilisation. Pour quantifier les différences d'utilisation entre les utilisateurs, nous avons distingué les super-utilisateurs et les autres utilisateurs (dans la suite ces utilisateurs sont désigné par l'expression « utilisateurs normaux »). Les super-utilisateurs constituent les usagers de Téo Taxi qui ont une utilisation relativement régulière et importante de Téo Taxi. La définition des super-utilisateurs est faite dans la section 5.4.2. La distribution de référence utilisée pour les tests χ^2 du présentés dans cette partie est celle de la population des utilisateurs normaux.

Il a été possible d'identifier 554 super-utilisateurs dans les répondants qui ont complété l'enquête. Le nombre d'utilisateurs normaux qui ont complété l'enquête est de 488. La somme des utilisateurs normaux et des super-utilisateurs est de 1 042 ce qui est inférieur au nombre de réponses complètes total qui est de 1 178. Cette différence s'explique par le fait que nous avons été dans l'incapacité d'associer une entrevue à un utilisateur pour 136 répondants ayant terminé l'enquête en raison de données erronées ou manquantes. Ces 136 réponses n'ont donc pas été utilisées dans cette partie.

Le test du χ^2 a été appliqué sur chaque question à l'exception de la question « Résidez-vous au Canada ? », car il y a moins de cinq super-utilisateurs qui ne résident pas au Canada. Le test permet de comparer les réponses de la population des usagers normaux et des super-utilisateurs. Le tableau présenté dans l'annexe F présente, pour chaque question, la statistique du test, la valeur p représente le risque de première espèce, et le seuil choisi correspond à un risque de première espèce de 5 % (soit une valeur p de 0,05). Le numéro de question dans le tableau présenté dans l'annexe F renvoie aux numéros de la liste des questions dans le paragraphe 6.2. Les résultats sont classés par ordre de décroissance de leur valeur p.

Sur les dix-neuf tests effectués, huit sont significatifs avec un risque de première espèce inférieure à 0,05 dont six avec une valeur p inférieure à 0,01. Les huit questions présentant un test significatif sont les suivantes :

- Taille du ménage (valeur p de 0,035)
- Revenu du ménage (valeur p de 0,013)
- Fréquence d'utilisation du taxi (valeur p de 0,002)
- Motif du déplacement (valeur p de 0,002)
- Nombre de voitures dans le ménage (valeur p de 0,001)
- Occupation principale (valeur p < 0,001)
- Abonnement aux différents services de mobilité (valeur p < 0,001)
- Nombre de passagers dans le taxi (valeur p < 0,001)

La Figure 6-21 présente la distribution des tailles des ménages des super-utilisateurs et des utilisateurs normaux. Les proportions de ménages à une, deux et trois personnes sont très proches pour les deux populations. En revanche plus de super-utilisateurs habitent dans des ménages à quatre personnes que les usagers normaux (16,4 % pour les super-utilisateurs contre 13,3 % pour les utilisateurs normaux). La part d'utilisateur normaux habitant dans des ménages à cinq ou six personnes est plus importante (3,2 % pour les super-utilisateurs contre 5,9 % pour les utilisateurs normaux). Le nombre de voitures par ménage est également représenté sur la Figure 6-21. On observe que les super-utilisateur possède un nombre moins important de voitures par ménage. Plus de super-utilisateurs ne possèdent qu'une ou zéro voiture (80,4% pour les super-utilisateurs contre 74,1 % pour les utilisateurs normaux). Ainsi, la part d'utilisateurs normaux possédant deux voitures ou plus par ménage est plus importante.

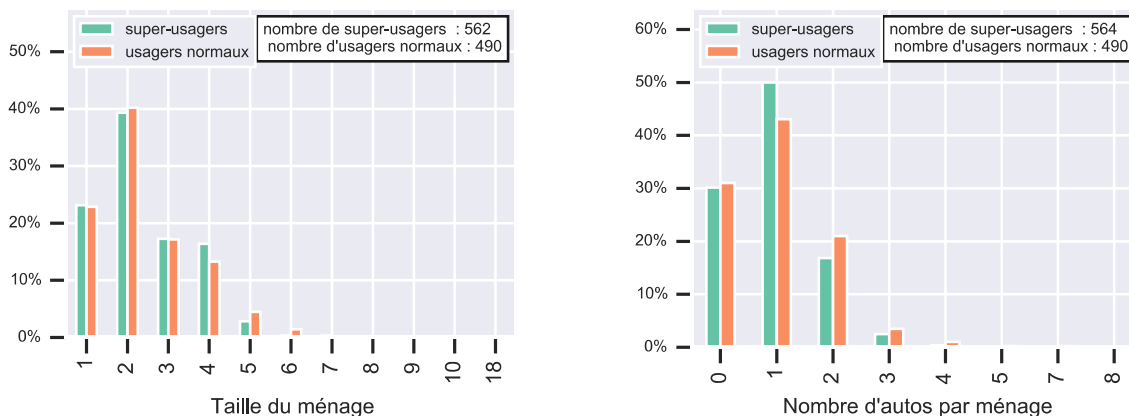


Figure 6-21 Distributions de la taille des ménages en fonction du type d'utilisateur (à gauche) et du nombre de voitures par ménage en fonction du type d'utilisateur (à droite)

La distribution du revenu annuel par ménage est présentée dans la Figure 6-22. Nous pouvons remarquer sur cette figure que la part de ménage avec les revenus annuels les plus faibles (moins de 60 000 \$) est plus faible chez les super-utilisateurs (13,0 % pour les super-utilisateurs contre 16,3 % pour les utilisateurs normaux). La part d'utilisateurs normaux dont les revenus sont compris entre 90 000 \$ et 120 000 \$ est plus faible (27,2 % pour les super-utilisateurs contre 23,8 % pour les utilisateurs normaux). Pour les revenus compris entre 120 000 \$ et 210 000 \$, la part des utilisateurs normaux est plus importante (43,4 % pour les super-utilisateurs contre 47,3 % pour les utilisateurs normaux). Une part plus importante de super-utilisateurs a préféré ne pas transmettre cette information.



Figure 6-22 Distribution du revenu par ménage en fonction du type d'utilisateur

La Figure 6-23 permet de se rendre compte de la différence d'utilisation des autres services de taxi. La fréquence d'utilisation du taxi est plus importante pour les super-utilisateurs. La part de super-utilisateur dont la fréquence d'utilisation est hebdomadaire est plus importante. De plus la part d'utilisateurs normaux ayant utilisé le taxi une fois au cours des six derniers mois est plus importante. Enfin la proportion d'utilisateurs n'ayant pas utilisé le taxi au cours des six derniers mois est plus faible pour les super-utilisateur.

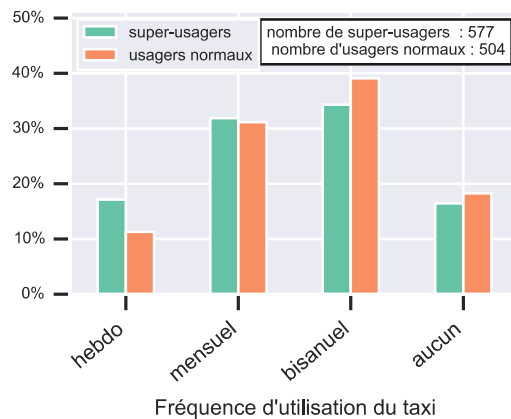


Figure 6-23 Distribution des fréquences d'utilisation du taxi en fonction du type d'utilisateur

Dans la Figure 6-24 le motif du déplacement est observable. La part de déplacements réalisés pour raisons professionnelles est plus importante chez les super-utilisateurs (45,4 % pour les super-utilisateurs contre 38,7 % pour les utilisateurs normaux). Conjointement les super-utilisateurs réalisent moins de déplacements pour le motif loisir (24,2 % pour les super-utilisateurs contre 31,7 % pour les utilisateurs normaux).

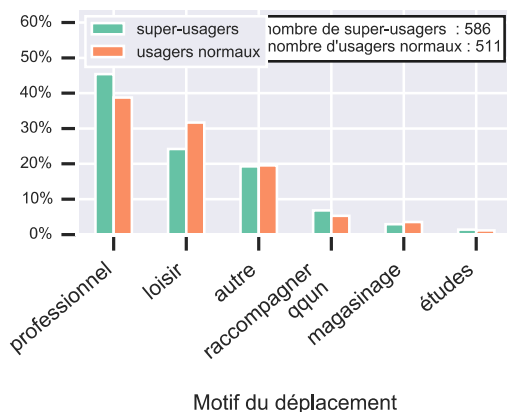


Figure 6-24 Distribution du motif du déplacement en fonction du type d'utilisateur

Dans la Figure 6-25, nous pouvons voir que les super-utilisateurs ont tendance à moins partager leur course avec d'autres passagers. Une part plus importante de super-utilisateurs était seul à bord du véhicule lorsqu'ils ont réalisé leur course avec Téo Taxi (72,1 % pour les super-utilisateurs contre 65,1 % pour les utilisateurs normaux).

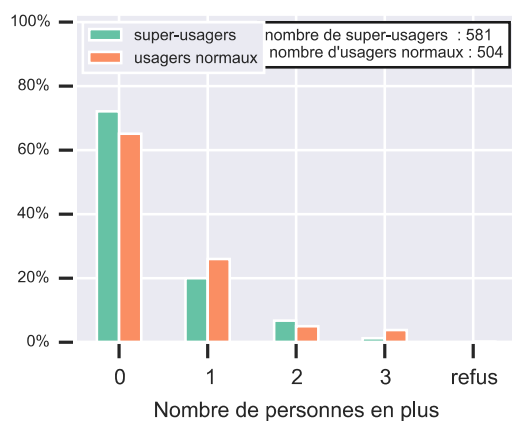


Figure 6-25 Distribution du nombre de passagers en plus du répondant dans le taxi en fonction du type d'utilisateur

La différence entre l'occupation des répondants selon les groupes est présentée dans la Figure 6-26. La proportion de travailleurs temps plein est légèrement plus importante chez les super-utilisateurs que chez les usagers normaux (86,2 % pour les super-utilisateurs contre 82,0 % pour les utilisateurs normaux). Le nombre de retraités chez les utilisateurs normaux est plus important (1,9 % pour les

super-utilisateurs contre 4,8 % pour les utilisateurs normaux). Malgré ces quelques différences, le nombre de travailleurs à temps plein reste largement majoritaire pour les deux populations.

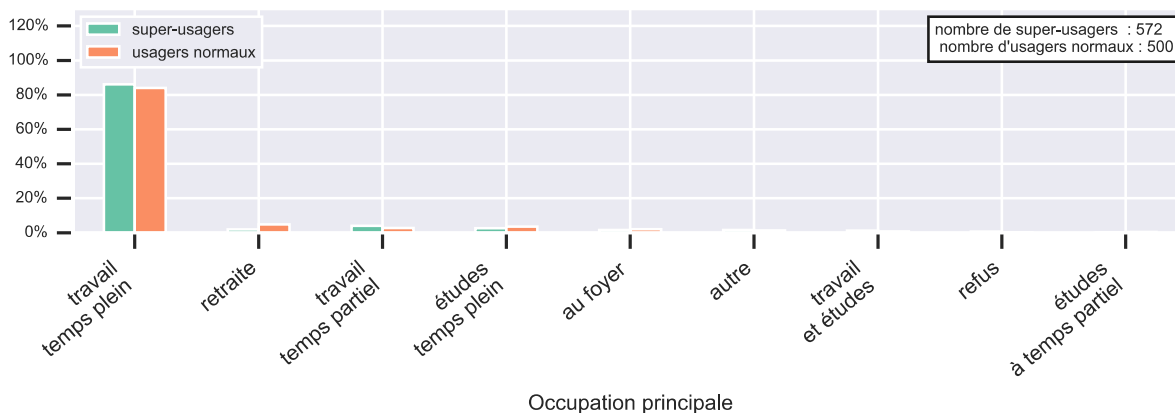


Figure 6-26 Distribution des fréquences des abonnements aux différents services de mobilité en fonction du type d'utilisateur

La Figure 6-27 présente les abonnements que possèdent les répondants aux différents services de mobilité. Pour les deux populations, le transport en commun est la réponse la plus fréquente. Cependant, une part plus faible de super-utilisateurs possède un abonnement annuel ou mensuel de transport en commun (26,5 % pour les super-utilisateurs contre 27,9 % pour les utilisateurs normaux). Les super-utilisateurs sont plus nombreux à posséder un abonnement au service de vélopartage Bixi (17,5 % pour les super-utilisateurs contre 11,8 % pour les utilisateurs normaux). Enfin une part moins importante des super-utilisateurs a répondu qu'il ne possédait aucun abonnement aux différents services de mobilité (13,4 % pour les super-utilisateurs contre 15,6 % pour les utilisateurs normaux).

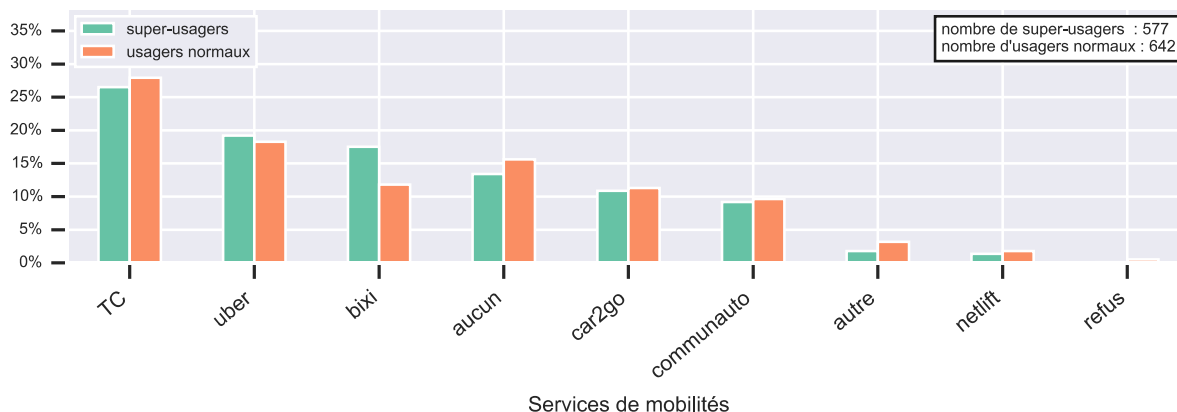


Figure 6-27 Distribution des fréquences des abonnements aux différents services de mobilité en fonction du type d'utilisateur

Regarder uniquement les différences entre les deux populations serait une comparaison incomplète. En effet, les points communs entre les distributions sont tout aussi importants. Au total, le test d'indépendance des distributions n'est pas significatif pour onze questions. Les différences d'âge et de genre entre les répondants de chaque groupe ne sont pas significatives par exemple.

Même s'il existe des différences significatives dans la répartition des réponses aux questions entre les usagers normaux et les super-utilisateurs, ces dernières restent mineures. Dans les résultats présentés dans ce paragraphe, les mêmes tendances générales sont observables pour les deux populations.

6.3.3.3 Influence de l'heure de la course sur les résultats

De la même manière, nous nous sommes intéressés à l'influence de l'heure de la course réalisée sur les réponses. Nous avons séparé la journée de 24h en deux catégories. Les courses réalisées en journée (entre 8h et 19h) et les courses réalisées la nuit (entre 19h et 8h). La même méthodologie a été appliquée aux deux populations. Au total, pour 556 réponses complétées, la course a été réalisée au cours de la journée. La course a été réalisée la nuit pour 486 répondants. Les résultats du test d'identité des distributions sont disponibles dans l'annexe G. Pendant la période de l'enquête 60,7 % ont été réalisés la journée. Dans notre échantillon des 1042 réponses complétées ce chiffre est de 53,3 %.

Le test a été significatif pour six questions avec un risque d'erreur de première espèce de 5 %. Ces questions sont les suivantes :

- Personne ayant payé pour la course (valeur p de 0,037)
- Nombre de voitures dans le ménage (valeur p de 0,001)
- Âge du répondant (valeur p < 0,001)
- Destinations différentes pour les passagers dans le véhicule (valeur p < 0,001)
- Raisons du choix de Téo Taxi (valeur p < 0,001)
- Motif du déplacement (valeur p < 0,001)

La Figure 6-28 permet de voir la personne qui a payé pour la course de taxi selon qu'elle a été réalisée la journée ou la nuit. Il est plus fréquent la journée que la course de taxi soit payée par l'employeur du répondant (26,4 % durant la journée contre 22,2 % la nuit). La nuit en revanche c'est plus généralement le répondant lui-même qui a payé la course (71,6 % durant la journée contre 76,4 % la nuit). La part de répondant ayant partagé les frais de la course entre avec l'ensemble de passagers dans le cas d'une course à plusieurs passagers est légèrement plus important la journée mais reste du même ordre de grandeur. Cette part est de 2,3 % la journée et 1,8 % la nuit.

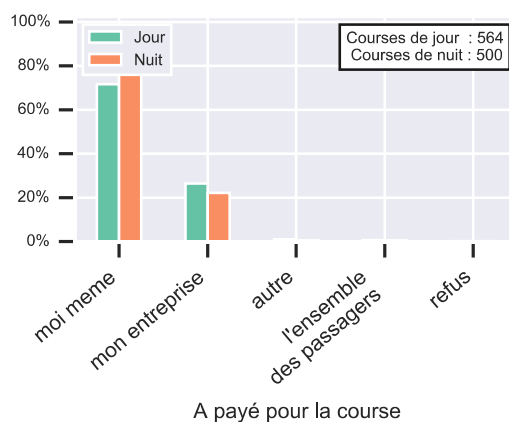


Figure 6-28 Distribution des réponses à la question « Qui a payé la course ? » en fonction de l'heure de la course

Dans le cas d'une course partagé, la part de courses dont tous les passagers n'avaient pas la même destination est plus importante la nuit. Pour les courses de nuit, 22,4 % des répondants dont la course était partagée ont déclaré que les passagers n'avaient pas la même destination. Ce chiffre est de 10,1 % pour les courses effectuées la journée.

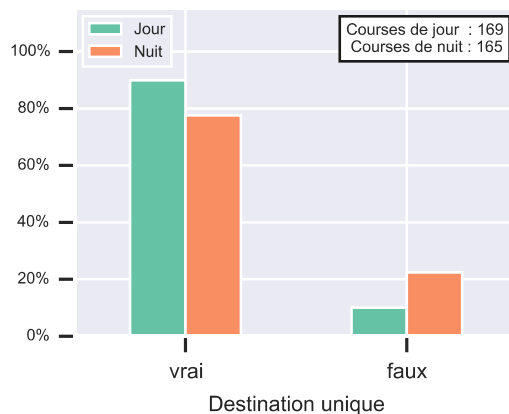


Figure 6-29 Distribution des réponses à la question « Est-ce que tous les passagers ont terminé leur course au même endroit ? » en fonction de l'heure de la course

Sur la Figure 6-30, nous pouvons voir l'âge des répondants selon le moment auquel la course a été réalisée. Les répondants réalisant des courses de nuits sont plus jeunes que ceux qui ont effectué une course la journée. 30,9 % des répondants nocturnes ont moins de 35 ans contre 25,3 % pour les utilisateurs diurnes. Les utilisateurs nocturnes âgés de 60 ans et plus sont 7,2 % contre 12,1 % pour la journée.

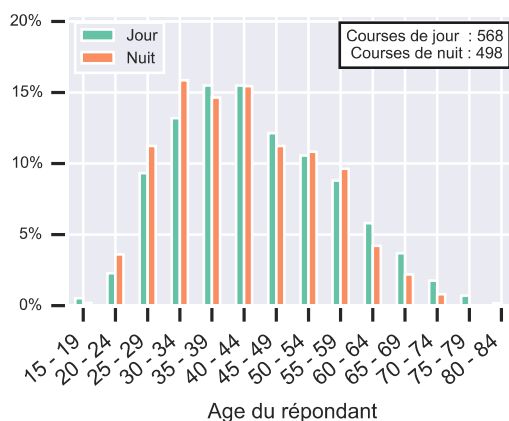


Figure 6-30 Distribution de l'âge des répondants en fonction de l'heure de la course

Le nombre de voitures par ménage dépendamment de l'heure de la course est également représenté sur la Figure 6-31. Les utilisateurs de nuit possèdent un nombre moins important de voitures par ménage. La part d'utilisateur nocturne possédant une ou zéro voiture est de 80,8 % pour la nuit contre 74,8 % pour les utilisateurs en journée. De même la part d'utilisateurs diurnes possédant

deux ou plus véhicules est plus importante (25,6 % la journée contre 19,2 % pour les utilisateurs nocturnes).

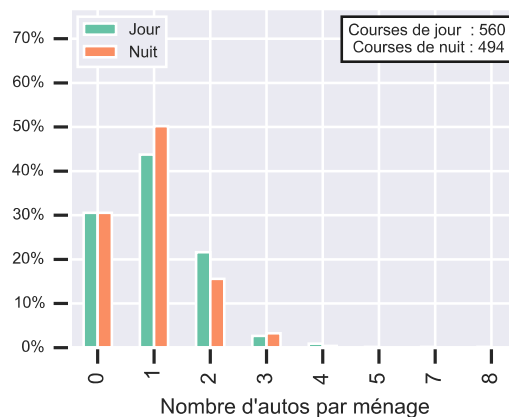


Figure 6-31 Distribution du nombre de véhicules par ménage en fonction de l'heure de la course

Les motivations d'utilisations sont présentées dans la Figure 6-32. La possibilité de réserver par l'application est important pour les deux groupes, cependant la part de répondants nocturnes ayant sélectionné cette option est légèrement plus importante (22,2 % durant la journée contre 24,2 % la nuit). Ce constat est également valable pour l'aspect écologique de Téo Taxi (20,2 % durant la journée contre 21,5 % la nuit). Deux éléments semblent plus importants en journée, le premier est la rapidité (9,9 % durant la journée contre 7,9 % la nuit) et le second est la possibilité de transportant facilement des objets encombrants en taxi (4,0 % durant la journée contre 1,6 % la nuit). Enfin le confort semble être un critère plus important pour les utilisateurs nocturnes (11,6 % durant la journée contre 13,7 % la nuit).

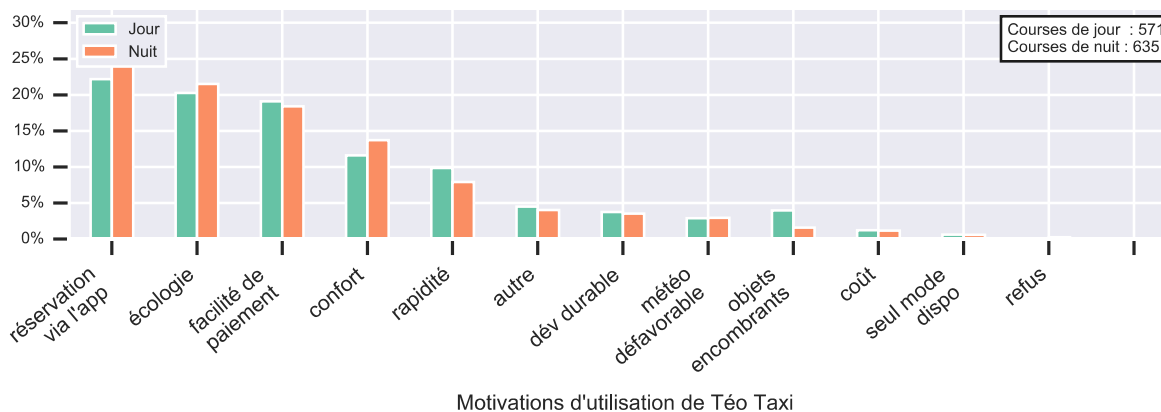


Figure 6-32 Distribution des raisons du choix de Téo Taxi en fonction de l'heure de la course

Enfin les différences de motif sont marquées en ces deux populations. La journée la part de courses réalisées pour des raisons professionnelles est plus importante (48,5 % durant la journée contre 35,2 % la nuit). La nuit, une part beaucoup plus importante de courses est réalisé pour le motif loisir (20,1 % durant la journée contre 36,4 % la nuit).

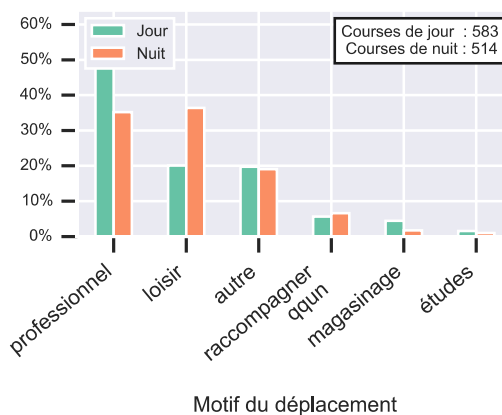


Figure 6-33 Distribution du motif du déplacement en fonction de l'heure de la course

6.3.3.4 Influence de l'utilisation de Téo Taxi sur la période

La dernière analyse croisée que nous avons réalisée concerne l'influence de l'utilisation de Téo Taxi sur la période pendant laquelle la course a été effectuée. Pour cela nous nous sommes basés sur les typologies d'utilisation développées dans le Chapitre 5. Les cinq typologies ont été séparées en deux catégories: la catégorie d'utilisation faible (composée des groupe aller-retour, usage faible et usage exceptionnel) et la catégorie d'usage fort (composée des groupes d'usage moyen et

d'usage intensif). Le nombre de répondants ayant complété l'enquête dont la période appartient à chacun des groupes est respectivement de 741 et 301. Les résultats des tests statistiques sont présentés dans l'annexe H.

Dans cette partition des répondants, onze tests se sont révélés significatifs dont sept avec une valeur p inférieure à 0,01. Des trois segmentations réalisées celle-ci est celle pour laquelle le plus grand nombre de tests du χ^2 sont significatif avec un risque d'erreur de première espèce inférieur à 0,05.

Voici les onze questions :

- Le mode alternatif envisagé (valeur p de 0,034)
- L'âge du répondant (valeur p de 0,034)
- La multimodalité du trajet (valeur p de 0,030)
- Le revenu annuel du ménage (valeur p de 0,010)
- Les raisons d'utilisation de Téo Taxi (valeur p de 0,005)
- L'occupation principale (valeur p de 0,002)
- Le nombre de passagers dans le véhicule (valeur p < 0,001)
- La taille du ménage (valeur p < 0,001)
- Les abonnements aux différents services de mobilité (valeur p < 0,001)
- La fréquence d'utilisation du taxi (valeur p < 0,001)
- Le motif du déplacement (valeur p < 0,001)

La Figure 6-34 montre l'âge du répondant en fonction de son utilisation sur la période. La part des utilisateurs les plus intensifs dans la trentaine (entre 30 et 39 ans) est plus importante (32,7 % pour les usagers intensifs contre 28,2 % pour les usagers modérés). La part d'utilisateur à l'usage le plus modéré âgée de plus de 55 ans est plus importante (18,1 % pour les usagers intensifs contre 19,4 % pour les usagers modérés).

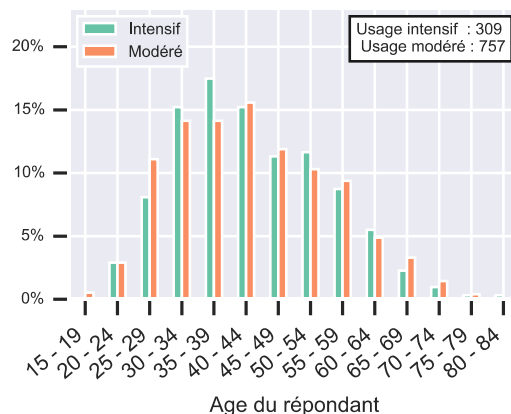


Figure 6-34 Distribution de l'âge des répondants en fonction de l'utilisation de Téo Taxi sur la période

Nous pouvons observer sur la Figure 6-35 un nombre légèrement plus important de travailleurs à temps plein chez les utilisateurs intensifs (87,1 % pour les usagers intensifs contre 84,4 % pour les usagers modérés) et une part plus importante de retraités chez les utilisateurs modérés (1,9 % pour les usagers intensifs contre 3,8 % pour les usagers modérés). Ce constat est cohérent avec les différences dans la répartition de l'âge deux répondants des deux populations présentées dans le paragraphe précédent.

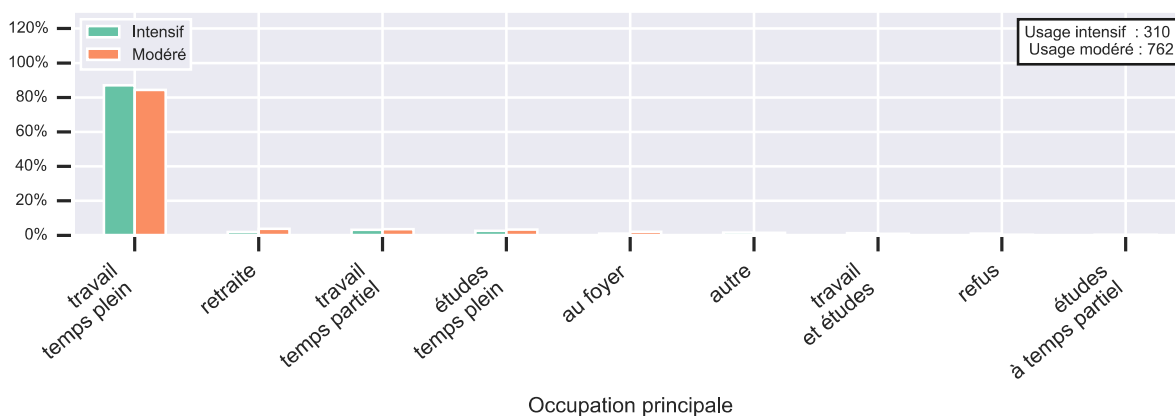


Figure 6-35 Distribution de l'occupation principale en fonction de l'utilisation de Téo Taxi sur la période

Dans la Figure 6-36 nous pouvons observer les distributions des revenus annuels des ménages. La part des classes de revenu les plus faibles (revenu inférieur à 60 000 \$) est plus importante pour les

utilisateurs modéré (12,0 % pour les usagers intensifs contre 15,4 % pour les usagers modérés). La part d'utilisateur modéré est également plus importante pour les classes de revenus comprises entre 150 000 \$ et 210 000 \$ (12,0 % pour les usagers intensifs contre 15,9 % pour les usagers modérés). Enfin les ménages à haut revenu (plus de 210 000 \$) est plus importante chez les utilisateurs intensifs (21,5 % pour les usagers intensifs contre 17,7 % pour les usagers modérés).

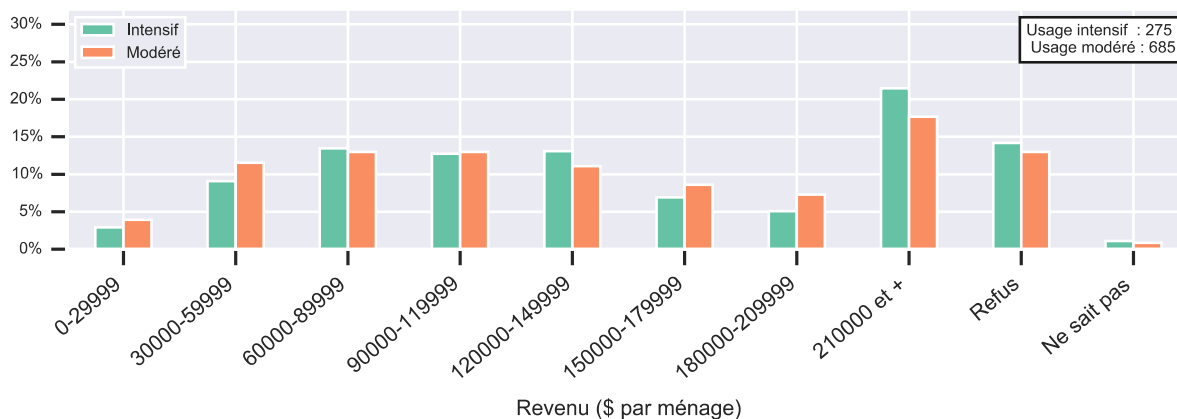


Figure 6-36 Distribution du revenu annuel par ménage en fonction de l'utilisation de Téo Taxi sur la période

La Figure 6-37 montre la répartition des tailles des ménages pour les deux populations. La part de ménages à deux ou trois personnes est plus importante chez les utilisateurs modérés (52,5 % pour les usagers intensifs contre 58,7 % pour les usagers modérés). Une part légèrement plus importante d'utilisateurs intensif habitent seuls (25,1 % pour les usagers intensifs contre 22,2 % pour les usagers modérés). Enfin une part plus importante d'utilisateurs intensifs habitent dans des ménages à quatre personnes (18,5 % pour les usagers intensifs contre 13,5 % pour les usagers modérés).

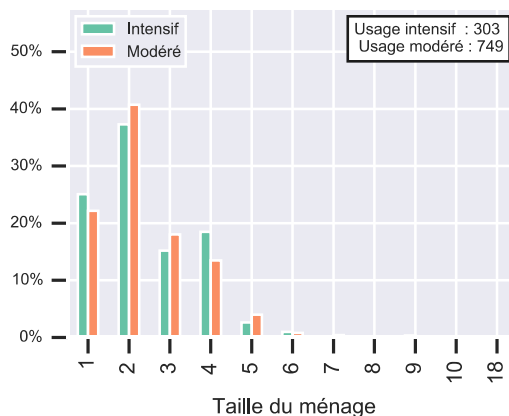


Figure 6-37 Distribution de la taille du ménage en fonction de l'utilisation de Téo Taxi sur la période

Les abonnements aux différents services de mobilité des deux populations sont visibles dans la Figure 6-38. La part d'utilisateur intensif possédant un abonnement de transport en commun est plus faible (25,1 % pour les usagers intensifs contre 28,4 % pour les usagers modérés). Cependant, ces utilisateurs sont proportionnellement moins nombreux à ne posséder aucun abonnement (12,8 % pour les usagers intensifs contre 15,1 % pour les usagers modérés). Les parts d'utilisateurs intensifs possédant un compte Uber (20,6 % pour les usagers intensifs contre 18,0 % pour les usagers modérés), un abonnement de vélo partage (Bixi) (16,4 % pour les usagers intensifs contre 14,3 % pour les usagers modérés) ou de covoiturage (car2go) (12,6 % pour les usagers intensifs contre 10,4 % pour les usagers modérés) sont plus importantes.

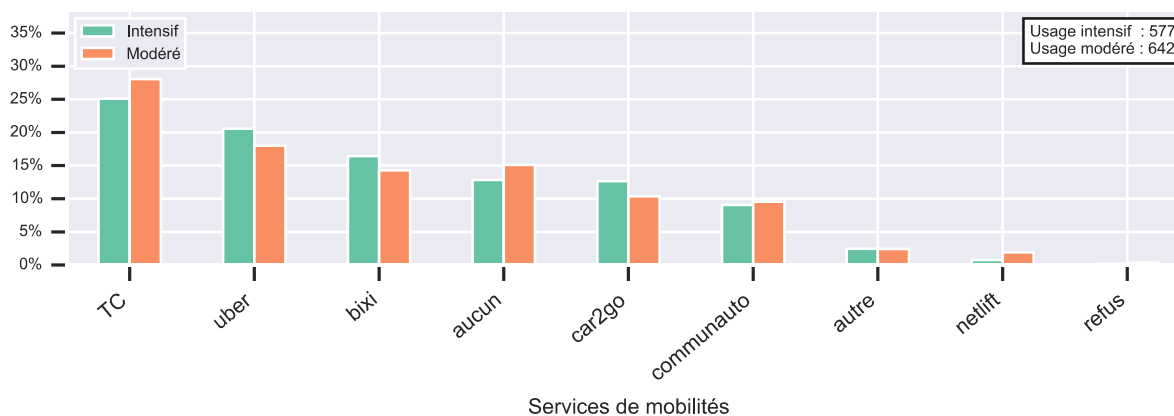


Figure 6-38 Distribution des fréquences des abonnements aux différents services de mobilité en fonction de l'utilisation de Téo Taxi sur la période

Dans la Figure 6-39, nous pouvons voir les modes alternatifs envisagés dans le cas où il aurait été impossible de réaliser la course avec Téo Taxi. Les utilisateurs intensifs auraient moins pris le transport en commun (50,0 % pour les usagers intensifs contre 58,4 % pour les usagers modérés) et plus un autre taxi (différent de Téo Taxi) (19,1 % pour les usagers intensifs contre 13,6 % pour les usagers modérés).

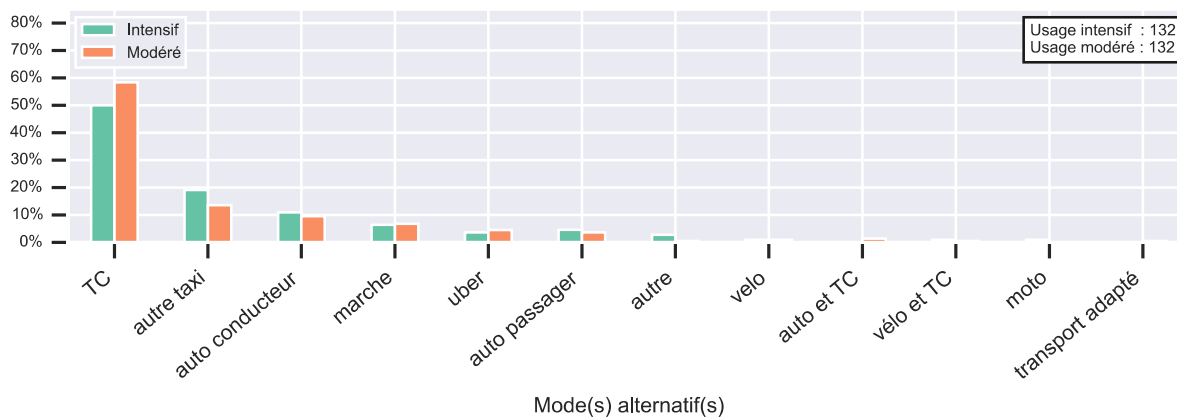


Figure 6-39 Distribution des modes alternatifs envisagés en fonction de l'utilisation de Téo Taxi sur la période

L'utilisation du taxi autre que Téo Taxi au cours des six derniers mois est présentée dans la Figure 6-40. Bien que le nombre d'utilisateurs n'ayant pas utilisé le taxi est semblable pour les deux populations (17,3 % pour les usagers intensifs et 17,3 % pour les usagers modérés), la fréquence d'utilisation de services de taxi différents de Téo Taxi des utilisateurs intensifs est plus importante. La part d'utilisateurs intensifs avec une fréquence hebdomadaire ou mensuelle est de 32,2 % pour les utilisateurs intensifs contre 24,7 % pour les utilisateurs modérés.

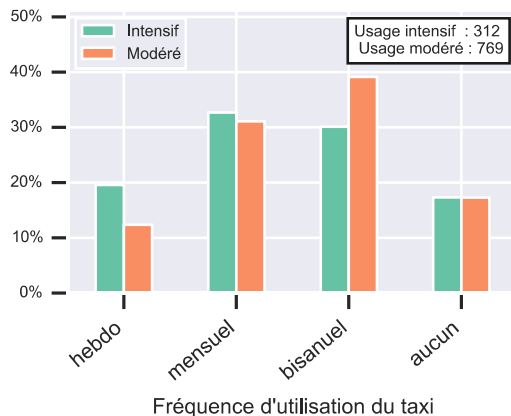


Figure 6-40 Distribution de l'utilisation d'autres services de taxi en fonction de l'utilisation de Téo Taxi sur la période

La Figure 6-41 présente les raisons du choix de Téo Taxi en fonction du comportement des utilisateurs sur la période. La facilité de paiement a été sélectionnée par plus d'utilisateurs intensifs (20,1 % pour les usagers intensifs contre 18,8 % pour les usagers modérés). Pour les utilisateurs modérés, le fait de pouvoir commander son taxi via l'application (22,4 % pour les usagers intensifs contre 23,4 % pour les usagers modérés) et la dimension écologique de Téo Taxi (20,5 % pour les usagers intensifs contre 21,0 % pour les usagers modérés) ont été plus sélectionnés. Les importances relatives de chaque option sont cependant très proches pour les deux populations.

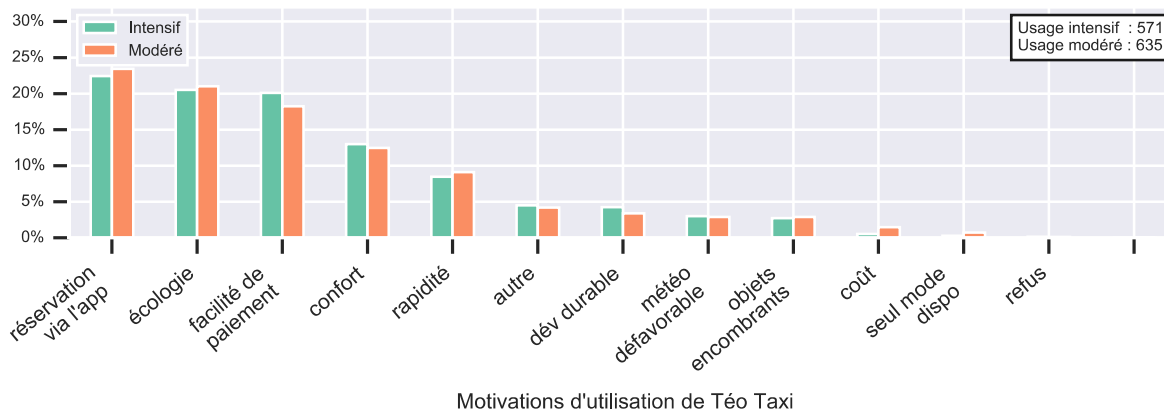


Figure 6-41 Distribution des raisons du choix de Téo Taxi en fonction de l'utilisation de Téo Taxi sur la période

Dans la Figure 6-42, nous pouvons voir que les utilisateurs intensifs ont tendance à moins partager leur course avec d'autres passagers. Une part plus importante d'utilisateurs intensifs étaient seuls à bord du véhicule lorsqu'ils ont réalisé leur course avec Téo Taxi (70,9 % pour les usagers intensifs contre 68,0 % pour les usagers modérés).

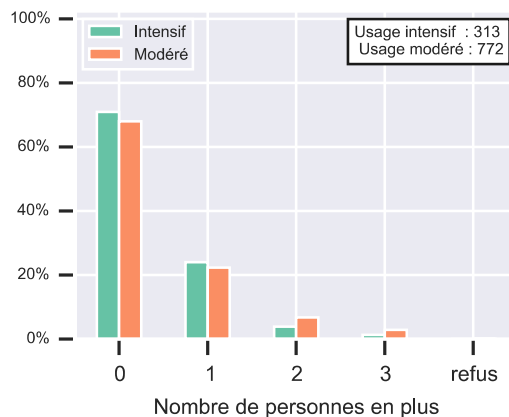


Figure 6-42 Distribution du nombre de passagers en plus du répondant dans le taxi en fonction de l'utilisation de Téo Taxi sur la période

La différence de motif du déplacement entre les deux groupes est marquée. La Figure 6-43 montre la distribution des motifs pour les deux populations. Les utilisateurs intensifs ont fait la course enquêtée plus pour des raisons professionnelles (60,0 % pour les usagers intensifs contre 38,8 % pour les usagers modérés). Les utilisateurs modérés en revanche ont fait la course enquêtée pour des loisirs (18,1 % pour les usagers intensifs contre 31,2 % pour les usagers modérés).

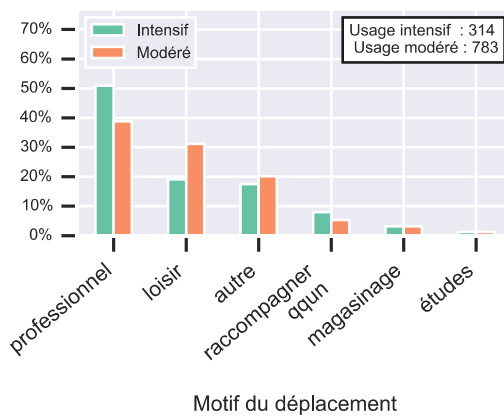


Figure 6-43 Distribution du motif du déplacement en fonction de l'utilisation de Téo Taxi sur la période

6.4 Synthèse de l'enquête menée auprès des utilisateurs

Cette enquête nous a permis de mieux comprendre l'utilisation de Téo Taxi pour 9 % des utilisateurs pendant la période allant du 8 décembre 2018 au 13 décembre 2018. Des informations sur les caractéristiques des déplacements, les habitudes de mobilité et le portrait sociodémographique des répondants ont pu être récoltées. Nous avons pu ainsi mieux comprendre les facteurs favorisant l'utilisation de Téo Taxi comme l'utilisation d'une application, l'aspect écologique de l'utilisation de véhicules électriques et la facilité de paiement. Malgré certains points communs entre l'utilisation de Téo Taxi et Uber, les répondants tendent plus utiliser les autres services de taxi disponibles que Uber. Enfin les répondants sont très majoritairement des travailleurs à temps plein qui utilisent Téo Taxi principalement pour des déplacements professionnels.

L'influence sur les réponses du comportement d'utilisation des répondants à long terme, de la période de l'enquête ainsi que l'heure à laquelle la course a été réalisée n'est pas négligeable. Parmi ces trois critères, l'utilisation de Téo Taxi sur la période de l'enquête est celui qui a un impact significatif sur les réponses au plus grand nombre de questions. Cette influence est particulièrement forte sur le motif du trajet réalisé.

Les résultats de cette enquête auraient également pu permettre à Téo Taxi de mieux connaître leurs clients et l'utilisation qu'ils faisaient du taxi. Ces informations auraient été précieuses pour continuer à améliorer leur service et mieux répondre aux besoins de leur clientèle.

CHAPITRE 7 CONCLUSION ET RECOMMANDATIONS

Ce chapitre conclut ce mémoire. Ce projet de recherche est consacré à l'étude de l'utilisation du service de taxi électrique : Téo Taxi. Téo Taxi était un opérateur de taxi qui a été actif dans la région de Montréal entre 2016 et début 2019. La particularité de cet opérateur est de permettre à ses clients d'utiliser une application mobile pour utiliser ses services et d'avoir une flotte de véhicules exclusivement électriques.

7.1 Synthèse de la recherche

Tout d'abord, la revue de littérature permet de mieux comprendre l'environnement complexe de l'industrie du taxi. Depuis l'introduction des nouveaux acteurs de mobilité à la demande, le marché de l'industrie s'est vu modifiée Les principes de fonctionnement et de régulation de l'industrie qui découlent de l'évolution de ce mode de transport sont remis aujourd'hui en question par les nouvelles formes que prend ce marché. Même si la part modale du taxi est généralement faible, le rôle de ce mode dans le système de transport est loin d'être négligeable. Il est très difficile cependant d'étudier les utilisateurs du taxi, car il est très difficile d'avoir accès à des données permettant de le faire. Il est fréquent de recourir à des enquêtes lorsque des recherches sur les utilisateurs du taxi sont menées.

Ensuite, il est important de comprendre le contexte général dans lequel un système de transport par taxi évolue. Le contexte géographique joue en effet un rôle prépondérant dans cette industrie. Il est donc nécessaire de connaître la particularité de la région de Montréal d'un point de vue de la mobilité des personnes. Le fonctionnement de l'entreprise Téo Taxi qui a opéré une flotte de taxis électriques pendant trois ans est très particulier. Trois spécificités notables de cette entreprise étaient particulièrement notables : la flotte de taxis était entièrement électrique, l'utilisation du service est basée sur l'utilisation d'une application et les chauffeurs étaient salariés de l'entreprise. Cette particularité permet d'étudier directement les utilisateurs de cet opérateur.

L'étude de ces utilisateurs nécessite le développement d'une méthodologie spécifique. La faible utilisation du service rend peu adaptées les méthodologies d'études des utilisateurs construites pour d'autres modes de transport comme le transport en commun. Une partie primordiale de la

méthodologie développée est le filtrage et le nettoyage des données brutes. L'utilisation de Téo Taxi est répartie principalement dans les secteurs centraux et autour de l'aéroport de Montréal. La fréquence d'utilisation de Téo Taxi est relativement faible (2,8 courses par an et par utilisateur en moyenne).

Les comportements d'utilisation des clients sur une courte période (quatre semaines) sont diversifiés. Bien qu'une grande partie des utilisateurs utilisent le taxi exceptionnellement sur ces périodes, un éventail d'utilisations peut être observé. L'intensité d'utilisation, la répartition des courses réalisées dans la journée ou dans la semaine sont entre autres les caractéristiques qui varient entre les différents comportements d'utilisation. Bien que les comportements d'utilisation par l'ensemble des utilisateurs soient relativement stables dans le temps, des comportements différents sur des durées d'étude plus longues peuvent également être observés au niveau de l'utilisateur.

Enfin, un sondage mené auprès des utilisateurs a permis de mieux comprendre certains facteurs associés à l'utilisation de Téo Taxi qui ne peuvent pas être déterminés avec l'étude des courses effectuées par l'utilisateur. Un portrait socio-économique des utilisateurs obtenu avec les réponses à cette enquête permet également de mieux connaître la population des utilisateurs de Téo Taxi. Même s'il existe des différences entre les répondants en fonction de leur utilisation de Téo Taxi, ces différences restent dans l'ensemble relativement faibles.

7.2 Contributions

Les contributions apportées par ce mémoire sont doubles. Ce projet de recherche a permis d'améliorer la connaissance des systèmes de transport par taxi. L'écosystème du taxi montréalais a déjà fait l'objet de plusieurs études, mais aucune d'entre elles n'a défini l'utilisateur comme objet d'étude. Ce projet de recherche a donc permis de combler ce manque. Cela a été rendu possible grâce aux données particulières d'un des opérateurs de taxi de la région.

Les contributions méthodologiques sont également importantes. La méthodologie développée dans ce mémoire a été conçue pour être facilement reproductible et adaptable. À condition de disposer de données permettant de relier les caractéristiques essentielles d'un déplacement à l'utilisateur qui l'a réalisé, il est possible de reproduire cette méthodologie. En construisant cette méthodologie d'analyse, nous avons été soucieux de la rendre la plus générale possible. Aucun ancrage

géographique n'est nécessaire pour appliquer cette méthodologie. La durée de la période d'étude peut également être facilement modifiée.

7.3 Limitations

Ce projet présente cependant un certain nombre de limitations et d'hypothèses qui ont dû être posées pour réaliser cette recherche.

Tout d'abord, les données à notre disposition ne proviennent que de l'opérateur Téo Taxi. De nombreux autres opérateurs de taxi sont également présents sur l'île de Montréal, mais ce type d'analyses centrées sur les utilisateurs n'est pas applicable, pour le taxi, que sur les activités de Téo Taxi. En effet, les données récoltées par les autres intermédiaires de service de taxi ne permettent pas encore d'identifier les clients et d'étudier leur utilisation du taxi. Les études présentées dans ce projet de recherche ne peuvent donc pas être étendues à l'ensemble des utilisateurs du taxi à Montréal, du moins pas pour le moment.

De plus, le caractère spécifique du modèle de Téo Taxi pose un problème de représentativité de l'utilisation du taxi. Les fonctionnalités et le fonctionnement particulier de Téo Taxi pourrait avoir un effet sur l'utilisation du taxi. Nous ne pouvons donc pas être certain que le portrait général d'utilisation de Téo Taxi est identique à celui du reste de l'industrie montréalaise. Les résultats présentés dans ce projet ne peuvent donc pas être interprétés pour l'ensemble de l'industrie du taxi dans la région de Montréal, mais uniquement pour l'opérateur Téo Taxi.

Le pourcentage de données incomplètes dans la base de données brute utilisée est important. Le choix de privilégier la qualité des données à la quantité a donc conduit à un nombre important de données filtrées. Dans ces conditions, les hypothèses et les choix méthodologiques de filtrage peuvent ne pas s'appliquer dans le cas d'un jeu de données présentant des caractéristiques différentes.

Enfin, puisqu'il s'agit de la première étude sur les utilisateurs de taxi dans la région de Montréal, très peu d'informations sur ces derniers est disponible. Il n'est donc pas possible de connaître les caractéristiques de la population de référence des usagers de Téo Taxi et du taxi en général. Les résultats d'une enquête menée auprès de cette population ne peuvent donc pas être considérés comme représentatifs dans cette situation. Nous ne disposons en effet d'aucunes données pour valider la représentativité de l'échantillon des répondants.

7.4 Perspectives

La richesse des données récoltées par Téo Taxi permet de nombreuses analyses. Plusieurs volets de recherche possibles n'ont pas été abordés. Les données sur les courses comprennent les lieux d'origine et de destination de chaque course. Il est ainsi possible de caractériser l'utilisation de Téo taxi en fonction de la répartition spatiale des courses réalisées par les clients. Le rôle du taxi pouvant être très différent selon les caractéristiques des quartiers, il est particulièrement intéressant de mieux comprendre les différences dans le comportement spatial d'utilisation du taxi par différents types de clients ou pour différentes périodes.

Les effets liés aux phénomènes de saisonnalité peuvent être importants dans les comportements de mobilité des personnes. Il serait donc très pertinent de réaliser le même sondage dans plusieurs contextes temporels (printemps, été, automne) pour estimer les effets de saisonnalité sur les réponses. Ces sondages étaient initialement prévus dans le cadre de ce projet, cependant, la fin précipitée des opérations de Téo Taxi ne nous a pas permis de les réaliser. Un sondage équivalent pourrait également être mené auprès des utilisateurs du taxi régulier. Ces deux sondages nous permettraient de comparer les utilisateurs de ces deux types de services.

BIBLIOGRAPHIE

- Aarhaug, J. et Skollerud, K. (2014). Taxi: Different Solutions in Different Segments. *Transportation Research Procedia*, 1(1), 276-283. doi:10.1016/j.trpro.2014.07.027
- Agard, B., Morency, C. et Trépanier, M. (2006). Mining Public Transport User Behaviour From Smart Card Data. *IFAC Proceedings Volumes*, 39(3), 399-404. doi:10.3182/20060517-3-FR-2903.00211
- Agard, B., Nia, V. P. et Trépanier, M. (2013). Assessing Public Transport Travel Behaviour From Smart Card Data With Advanced Data Mining Techniques, 13.
- Assemblée nationale du Québec. (2019). Projet de loi n°17 : Loi concernant le transport rémunéré de personnes par automobile - Assemblée nationale du Québec. Repéré à <http://www.assnat.qc.ca/fr/travaux-parlementaires/projets-loi/projet-loi-17-42-1.html>
- Atkinson-Palombo, C., Varone, L. et Garrick, N. (2019). Understanding the Surprising and Oversized Use of Ridesourcing Services in Poor Neighborhoods in New York City. *Transportation Research Record: Journal of the Transportation Research Board*, 036119811983580. doi:10.1177/0361198119835809
- Austin, D. et Zegras, P. C. (2012). Taxicabs as Public Transportation in Boston, Massachusetts. *Transportation Research Record*, 2277(1), 65-74. doi:10.3141/2277-08
- Beimborn, E. A., Greenwald, M. J. et Jin, X. (2003). Accessibility, Connectivity, and Captivity: Impacts on Transit Choice. *Transportation Research Record*, 1835(1), 1-9. doi:10.3141/1835-01
- Bouchard, É. (2018, août). *Analyse de la consommation énergétique de véhicules électriques* (masters, École Polytechnique de Montréal). Repéré à <https://publications.polymtl.ca/3297/>
- Circella, D. G., Alemi, F., Tiedeman, K., Berliner, R. M., Lee, Y., Fulton, D. L., ... Handy, S. (2017). What Affects Millennials' Mobility? PART II: The Impact of Residential Location, Individual Preferences and Lifestyles on Young Adults' Travel Behavior in California, 93.
- Commission des transports Québec. (2019). Projets pilotes dans le secteur du transport par taxi - CTQ - Commission des transports du Québec. Repéré à https://www.ctq.gouv.qc.ca/taxi/projets_pilotes_dans_le_secteur_du_transport_par_taxi.html
- Cooper, J., Mundy, R. et Nelson, J. (2010). Historical Development of the Taxi. Dans *Taxi!: Urban Economies and the Social and Transport Impacts of the Taxicab* (p. 1-14). Ashgate Publishing, Ltd.
- Deschaintres, E. (2018, août). *Analyse de la variabilité individuelle d'utilisation du transport en commun à l'aide de données de cartes à puce* (masters, École Polytechnique de Montréal). Repéré à <https://publications.polymtl.ca/3284/>
- DIDI. (2019). DiDi Taxi- Express Mobility - DiDi official website. Repéré à <https://www.didiglobal.com/travel-service/taxi>

- El Mahrsi, M. K., Come, E., Oukhellou, L. et Verleysen, M. (2017). Clustering Smart Card Data for Urban Mobility Analysis. *IEEE Transactions on Intelligent Transportation Systems*, 18(3), 712-728. doi:10.1109/TITS.2016.2600515
- Exo. (2015). *Enquête Origine-Destination 2013 de la région de Montréal, version 13.2b*.
- Frankena, M. W. (1984). An Economic Analysis of Taxicab Regulation, 183.
- Frazzani, S., Grea, G. et Zamboni, A. (2016). *Study on passenger transport by taxi, hire car with driver and ridesharing in the EU*. European Commission. Repéré à <https://ec.europa.eu/transport/sites/transport/files/2016-09-26-pax-transport-taxi-hirecar-w-driver-ridesharing-final-report.pdf>
- Geneste, A. (2017, juin). *Caractérisation de l'offre et de la demande de transport dans un système de taxis électriques* (masters, École Polytechnique de Montréal). Repéré à <https://publications.polymtl.ca/2600/>
- Ghaemi, M. S., Agard, B., Trépanier, M. et Nia, V. P. (2017). A visual segmentation method for temporal smart card data. *Transportmetrica A: Transport Science*, 13(5), 381-404. doi:10.1080/23249935.2016.1273273
- Gilbert, G. et Samuels, R. E. (1982). The Taxicab: an Urban Transportation Survivor. Repéré à <https://trid.trb.org/view/177870>
- Gössling, S. (2018). ICT and transport behavior: A conceptual review. *International Journal of Sustainable Transportation*, 12(3), 153-164. doi:10.1080/15568318.2017.1338318
- Hara Associates Inc. et Corey, Canapary & Galanis. (2013). *Best Practices Studies of Taxi Regulation Taxi User Surveys (Draft)*. San Francisco Municipal Transportation Agency. Repéré à <https://www.sfmta.com/sites/default/files/Draft%20SF%20UserSurvey%2055%20WEB%20version04042013.pdf>
- Harding, S., Kandlikar, M. et Gulati, S. (2016). Taxi apps, regulation, and the market for taxi journeys. *Transportation Research Part A: Policy and Practice*, 88, 15-25. doi:10.1016/j.tra.2016.03.009
- IPART. (2017). *Survey of Point to Point Transport Use*. Sydney, NSW, Australia : Independent Pricing and Regulatory Tribunal New South Wales. Repéré à <https://www.ipart.nsw.gov.au/files/sharedassets/website/shared-files/investigation-administrative-taxi-passenger-survey/consultant-report-survey-of-point-to-point-transport-use-february-2017.pdf>
- Jain, A. K., Murty, M. N. et Flynn, P. J. (1999). Data clustering: a review. *ACM Computing Surveys*, 31(3), 264-323. doi:10.1145/331499.331504
- Jain, Anil K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651-666. doi:10.1016/j.patrec.2009.09.011
- Jara-Díaz, S. R. et Guevara, C. A. (2001). *Behind the Subjective Value of Travel Time Savings*.
- Lacombe, A. (2016, août). *Méthodologie d'analyse et de suivi d'un système de transport par taxi* (masters, École Polytechnique de Montréal). Repéré à <https://publications.polymtl.ca/2261/>

- Laviolette, J. (2017, août). *Planification stratégique d'un système de transport par taxi* (masters, École Polytechnique de Montréal). Repéré à <https://publications.polymtl.ca/2738/>
- Le Figaro. (2019, 23 janvier). Espagne: taxis en grève, Uber menace de se retirer de Barcelone. Repéré à <http://www.lefigaro.fr/flash-eco/espagne-taxis-en-greve-uber-menace-de-se-retirer-de-barcelone-20190123>
- Le Taxi. (2019). Le Taxi : Un clic, un taxi. Repéré à <https://le.taxi/>
- Lyft. (2018). 2018 in Review: Putting Our Vision Into Action. *Lyft Blog*. Repéré à <https://blog.lyft.com/posts/2018/12/19/2018-year-in-review>
- Maimon, O. et Rokach, L. (dir.). (2010). A survey of Clustering Algorithms. Dans *Data Mining and Knowledge Discovery Handbook* (2^e éd., p. 269-298). Springer US. Repéré à <http://www.springer.com/la/book/9780387098227>
- Mamh, D. (2019). *Communauté métropolitaine de Montréal* (p. 2). Repéré à <https://www.mamh.gouv.qc.ca/fileadmin/cartes/cm/663.pdf>
- Mathis, S. (2012). The End of Taxis. *CityLab*. Repéré à <http://www.theatlanticcities.com/commute/2012/07/end-taxis/2353/>
- Morency, C., Trepanier, M., Agard, B., Martin, B. et Quashie, J. (2007). *Car sharing system: what transaction datasets reveal on users' behaviors*. Communication présentée au 2007 IEEE Intelligent Transportation Systems Conference, Bellevue, WA, USA (p. 284-289). doi:10.1109/ITSC.2007.4357656
- Ngo, V. D. (2015, octobre). *Transportation network companies and the ridesourcing industry : a review of impacts and emerging regulatory frameworks for Uber* (University of British Columbia). Repéré à <https://open.library.ubc.ca/cIRcle/collections/graduateresearch/42591/items/1.0220795>
- NYC Taxi & Limousine Commission. (2011). *Taxi of Tomorrow Survey Results*. New York, NY, USA : NYC Taxi & Limousine Commission. Repéré à https://www1.nyc.gov/assets/tlc/downloads/pdf/tot_survey_results_02_10_11.pdf
- NYC Taxi & Limousine Commission. (2014). *Taxicab Fact Book*. New York, NY, USA. Repéré à https://www1.nyc.gov/assets/tlc/downloads/pdf/2014_tlc_factbook.pdf
- NYC Taxi & Limousine Commission. (2016). *Taxicab Fact Book*. New York, NY, USA. Repéré à https://www1.nyc.gov/assets/tlc/downloads/pdf/2016_tlc_factbook.pdf
- NYC Taxi & Limousine Commission. (2018). *Taxicab Fact Book*. New York, NY, USA. Repéré à https://www1.nyc.gov/assets/tlc/downloads/pdf/2018_tlc_factbook.pdf
- Office québécois de la langue française. (1986). taxi. Repéré à http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8414318
- Phithakkitnukoon, S., Veloso, M., Bento, C., Biderman, A. et Ratti, C. (2010). *Taxi-Aware Map: Identifying and Predicting Vacant Taxis in the City*. B. de Ruyter, R. Wichert, D. V. Keyson, P. Markopoulos, N. Streitz, M. Divitini, ... A. Mana Gomez (dir.), Communication présentée au Ambient Intelligence (p. 86-95).
- Price Waterhouse. (1993). *Analysis of Taxicab Deregulation and Re-Regulation*. Washington, DC, USA. Repéré à <https://www.colorado.gov/pacific/sites/default/files/0708TransAttachI.pdf>

- RACQ Public Policy Department. (2016). *RACQ Mobility Survey - Taxis and Rideshare*. QLD, Australia : RACQ Public Policy Department. Repéré à <https://www.racq.com.au/-/media/pdf/racq-pdfs/cars-and-driving/representing-qld-drivers/racq-submission-to-infrastructure-planning-and-ersource-committee-re-offences.pdf?la=en&hash=6FF050D61201B30C5D1EC5818ADB2B6F43CF0D8A>
- Rahm, E. et Do, H. H. (s. d.). *Data Cleaning: Problems and Current Approaches*, 11.
- Rayle, L., Dai, D., Chan, N., Cervero, R. et Shaheen, S. (2016). Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. *Transport Policy*, 45, 168-178. doi:<http://dx.doi.org/10.1016/j.tranpol.2015.10.004>
- Règlement sur les services de transport par taxi. , RLRQ c S-6.01, r. 3. Repéré à <http://www.legisquebec.gouv.qc.ca/fr/showdoc/cr/S-6.01,%20r.%203>
- Réseau québécois des aéroports. (2017). Aéroports de Montréal (ADM) – permis de taxi. *CAQ*. Repéré à <https://reseauquebecoisdesaeroports.ca/new/aeroports-de-montreal-adm-permis-de-taxi/>
- Rosenthal, B. M. (2019). As Thousands of Taxi Drivers Were Trapped in Loans, Top Officials Counted the Money - The New York Times. Repéré à <https://www.nytimes.com/2019/05/19/nyregion/taxi-medallions.html?action=click&module=Top%20Stories&pgtype=Homepage>
- Schaller, B. (2007). Entry controls in taxi regulation: Implications of US and Canadian experience for taxi regulation and deregulation. *Transport Policy*, 14(6), 490-506. doi:10.1016/j.tranpol.2007.04.010
- Scikit-learn. (2019a). Comparing different clustering algorithms on toy datasets — scikit-learn 0.21.2 documentation. Repéré à https://scikit-learn.org/stable/auto_examples/cluster/plot_cluster_comparison.html
- Scikit-learn. (2019b). sklearn.cluster.KMeans — scikit-learn 0.21.2 documentation. Repéré à <https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>
- Scipy. (2019). scipy.stats.chisquare — SciPy v1.3.0 Reference Guide. Repéré à <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.chisquare.html>
- Société de Transport de Montréal. (2016). *Plan de développement d'accessibilité universelle 2016-2020*. Montréal, QC, Canada. Repéré à https://www.stm.info/sites/default/files/pdf/fr/planaccessibilite_2016-2020.pdf
- Société de Transport de Montréal. (2018a). Plan du métro. Repéré à http://www.stm.info/sites/default/files/pictures/plan_metro_blanc.pdf
- Société de Transport de Montréal. (2018b). *Rapport annuel 2017*. Repéré à https://www.stm.info/sites/default/files/affairespubliques/Communiques/Rapport_annuel_2017/ap_rapport_annuel_2017_final.pdf
- Société de Transport de Montréal. (2019). La STM en bref. *Société de transport de Montréal*. Repéré à <http://www.stm.info/fr/presse/ressources-medias-0/propos-de-la-stm>
- Taxelco. (2018a). Foire aux questions sur notre zone de couverture. *Téo Taxi*. Repéré à <https://teo.taxi/foire-questions/zone-couverture/>

- Taxelco. (2018b). Téléchargez l'application | Téo taxi | Commandez 24/7. *Téo Taxi*. Repéré à <https://teo.taxi/telechargement/>
- Taxelco. (2019). *Téo Taxi au service des passagers*. Repéré à <https://s3.amazonaws.com/production-teomtl/memoire/memoire-public-taxelco-janvier-2019.pdf>
- Taxi collectif. (2019). *Société de transport de Montréal*. Repéré à <http://www.stm.info/fr/taxi>
- Transit. (2019). Transit - Montréal. Repéré à <https://transitapp.com/region/montreal>
- Uber. (2018a). Company Information | Uber Newsroom Canada. *Uber Newsroom*. Repéré à <https://www.uber.com/en-CA/newsroom/company-info/>
- Uber. (2018b). Get an Uber Ride - Download the Passenger App. *Uber*. Repéré à <https://www.uber.com/ca/en/ride/>
- Ville de Montréal. (2018). *Profil Sciodémographique, recensement 2016*. Montréal, QC, Canada. Repéré à http://ville.montreal.qc.ca/pls/portal/docs/PAGE/MTL_STATS_FR/MEDIA/DOCUMENTS/PROFIL_SOCIOD%20MO_VILLE%20DE%20MONTR%20AL%202016.PDF
- Ville de Montréal. (2019a). Nombre de voitures taxi par agglomération. Repéré à https://ville.montreal.qc.ca/pls/portal/docs/page/bur_taxi_fr/media/images/public/ph_carte_taxi.jpg
- Ville de Montréal. (2019b). Ville de Montréal - Bureau du taxi de Montréal - Mission [Web page]. Repéré à https://ville.montreal.qc.ca/portal/page?_pageid=8177,92141681&_dad=portal&_schema=PORTAL
- Ville de Montréal. (2019c). Ville de Montréal - Bureau du taxi de Montréal - Tarification [Web page]. Repéré à http://ville.montreal.qc.ca/portal/page?_pageid=8177,92215612&_dad=portal&_schema=PORTAL
- Vogel, M., Hamon, R., Lozenguez, G., Merchez, L., Abry, P., Barnier, J., ... Robardet, C. (2014). From bicycle sharing system movements to users: a typology of Vélo'v cyclists in Lyon based on large-scale behavioural dataset. *Journal of Transport Geography*, 41, 280-291. doi:10.1016/j.jtrangeo.2014.07.005
- Waheed, S., Herrera, L., Ritoper, S., Mehta, J., Romero, H. et Narro, V. (2015). Ridesharing or Ridestealing? Changes in Taxi Ridership and Revenue in Los Angeles 2009-2014. Repéré à <https://escholarship.org/uc/item/55f107vf>
- Waymo. (2019). *Waymo*. Repéré à <https://waymo.com/>
- Webster, A. L., Weiner, E. et Wells, J. D. (1974). Role Of Taxicabs In Urban Transportation. Repéré à <https://trid.trb.org/view/38341>
- Wielinski, G. (2018, décembre). *Modélisation des membres et de leur comportement dans un écosystème de services d'autopartage à Montréal* (phd, École Polytechnique de Montréal). Repéré à <https://publications.polymtl.ca/3781/>

- Xu, D. et Tian, Y. (2015). A Comprehensive Survey of Clustering Algorithms. *Annals of Data Science*, 2(2), 165-193. doi:10.1007/s40745-015-0040-1
- Xu, L. (2012). *Who Drives a Taxi in Canada?* (p. 15). Citoyenneté et Immigration Canada.
- Yang, H., Wong, S. C. et 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
- Yıldızgöz, K. et Murat Çelik, H. (2019). Critical Moment for Taxi Sector: What Should Be Done by Traditional Taxi Sector After the TNC Disruption?: Proceedings of 4th Conference on Sustainable Urban Mobility (CSUM2018), 24 - 25 May, Skiathos Island, Greece (p. 453-460). doi:10.1007/978-3-030-02305-8_55

ANNEXE A – EXTRAIT DE LA LOI CONCERNANT LES SERVICES DE TRANSPORT PAR TAXI, RLRQ C S-6.01 (DÉFINITION DU TERME SERVICES DE TRANSPORT PAR TAXI)

La définition suivante est issue de la loi RLRP C S-6.01 accessible au line suivant :

<http://www.legisquebec.gouv.qc.ca/fr/showdoc/cs/S-6.01>

« services de transport par taxi », tout service de transport rémunéré de personnes par automobile, à l'exception des suivants:

- a) le covoiturage effectué sur une partie ou l'ensemble d'un même parcours, à la condition que:
 - i. l'automobile utilisée soit un véhicule de promenade au sens de l'article 4 du Code de la sécurité routière ([chapitre C-24.2](#));
 - ii. le conducteur décide de la destination finale et que la prise de passagers à bord soit accessoire à la raison pour laquelle il se déplace;
 - iii. le transport soit offert moyennant une contribution financière qui se limite, quel que soit le nombre de personnes à bord de l'automobile, aux frais d'utilisation de celle-ci et dont le montant total n'excède pas celui de l'indemnité accordée à un employé d'un ministère ou d'un organisme dont le personnel est nommé suivant la Loi sur la fonction publique ([chapitre F-3.1.1](#)) pour l'utilisation de son véhicule personnel;
- b) le transport scolaire prévu dans la Loi sur l'instruction publique ([chapitre I-13.3](#)), dans la Loi sur l'instruction publique pour les autochtones cris, inuit et naskapis ([chapitre I-14](#)), dans la Loi sur l'enseignement privé ([chapitre E-9.1](#)) et dans la Loi sur les collèges d'enseignement général et professionnel ([chapitre C-29](#)) et le transport des élèves d'une institution dont le régime d'enseignement est l'objet d'une entente internationale au sens de la Loi sur le ministère des Relations internationales ([chapitre M-25.1.1](#));
- c) le transport effectué par un conducteur bénévole œuvrant sous le contrôle d'un organisme humanitaire reconnu par le ministère de la Santé et des Services sociaux dans le cadre de l'un de ses programmes de soutien ou d'accompagnement, à la condition que:
 - i. le transport soit offert moyennant une contribution financière qui se limite, quel que soit le nombre de personnes à bord de l'automobile, aux frais d'utilisation de celle-ci qui ont été fixés par le conseil d'administration de l'organisme et dont le montant total n'excède pas celui déterminé par règlement du gouvernement;
 - ii. soit maintenu par l'organisme un registre permanent des transports qui identifie, pour chaque transport effectué, le conducteur, le client et, s'il y a lieu, l'accompagnateur et qui indique la date, le point d'origine, la distance parcourue et la destination de la course;
- d) le transport effectué par une entreprise d'économie sociale financée par un programme gouvernemental pour offrir des services d'accompagnement, notamment aux personnes âgées, handicapées, malades ou en perte d'autonomie, à la condition que soit maintenu par l'entreprise un registre permanent des transports qui identifie, pour chaque transport effectué, le conducteur, le client et, s'il y a lieu, l'accompagnateur et qui indique la date, le point d'origine, la distance parcourue et la destination de la course;
- e) le transport de personnes ayant les facultés affaiblies effectué par un conducteur bénévole œuvrant sous le contrôle d'un organisme ou d'une personne morale sans but lucratif ou par un conducteur rémunéré par une entreprise, à la condition que:
 - i. le transport de l'automobile de la personne transportée soit aussi effectué;
 - ii. le transport effectué par un conducteur bénévole soit sans intention de faire un gain pécuniaire;



- iii. soit maintenu par l'organisme ou la personne morale sans but lucratif ou l'entreprise concernée un registre permanent des transports qui identifie, pour chaque transport effectué, le conducteur, le client et l'accompagnateur et qui indique la date, le point d'origine, la distance parcourue et la destination de la course;
- f) le transport de courtoisie effectué par un conducteur rémunéré par une entreprise, mais offert gracieusement aux clients de celle-ci;
- g) le transport effectué dans un but d'entraide communautaire pour venir en aide ou accompagner une personne à la condition que ce transport soit offert moyennant une contribution financière qui se limite, quel que soit le nombre de personnes à bord de l'automobile, aux frais d'utilisation de celle-ci et dont le montant total n'excède pas celui déterminé par règlement du gouvernement;
- h) le transport de personnes à l'occasion de baptêmes, de mariages ou de funérailles ou le transport de personnes par automobile antique de plus de 30 ans;
- i) le transport par ambulance ou par corbillard.

**ANNEXE B – VALEUR DES COORDONNÉES DE CENTROÏDES ET ÉCART TYPE DES
COMPOSANTES DE CHAQUE GROUPE**

| | | <i>x1</i> | <i>x2</i> | <i>x3</i> | <i>x4</i> | <i>x5</i> | <i>x6</i> | <i>x7</i> | <i>x8</i> |
|---------------------------|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| <i>Aller-retour</i> | moyenne | 9,95E-01 | 4,92E-03 | 4,89E-04 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 7,24E-02 |
| | écart type | 3,31E-02 | 3,00E-02 | 8,18E-03 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 2,36E-02 |
| <i>Usage faible</i> | moyenne | 5,44E-01 | 4,56E-01 | 7,33E-05 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 7,72E-02 |
| | écart type | 7,74E-02 | 7,74E-02 | 2,77E-03 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 2,45E-02 |
| <i>Usage moyen</i> | moyenne | 3,88E-01 | 2,78E-01 | 2,43E-01 | 6,89E-02 | 1,86E-02 | 3,39E-03 | 2,96E-04 | 1,58E-01 |
| | écart type | 9,85E-02 | 6,41E-02 | 6,92E-02 | 9,03E-02 | 4,86E-02 | 2,02E-02 | 5,70E-03 | 6,07E-02 |
| <i>Usage exceptionnel</i> | moyenne | 1,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 3,23E-02 |
| | écart type | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 | 0,00E+00 |
| <i>Usager intensif</i> | moyenne | 3,13E-01 | 2,31E-01 | 1,77E-01 | 1,33E-01 | 8,82E-02 | 4,28E-02 | 1,45E-02 | 5,32E-01 |
| | écart type | 8,13E-02 | 4,88E-02 | 3,89E-02 | 4,23E-02 | 4,43E-02 | 4,19E-02 | 2,86E-02 | 2,30E-01 |

ANNEXE C – PAGE D'ACCUEIL DE LA PLATEFORME DE SONDAGE

Étude Mobilité English



La Chaire Mobilité de Polytechnique Montréal sollicite votre participation à une étude sur l'utilisation du taxi dans la région de Montréal dans le cadre d'un projet de recherche. Ce projet vise à mieux comprendre l'usage du taxi électrique à Montréal ainsi que son rôle dans la mobilité des Montréalais.

J'accepte les conditions du formulaire de consentement


Pour répondre à cette enquête, vous devez avoir au moins 18 ans et accepter les conditions présentées dans le [formulaire de consentement](#).

Attention : ce lien est actif une seule fois. Vous ne pourrez plus vous reconnecter une fois que vous aurez commencé à remplir ce questionnaire.

Confidentialité

Les informations recueillies par le biais de cette enquête serviront uniquement pour fins de recherche dans le cadre du projet de recherche : "Méthodologie de design et d'évaluation d'un service de taxi par véhicules électriques". Aucune donnée ne pourra être utilisée à des fins commerciales. Aucune publication ou communication scientifique ne renfermera quelque information que ce soit pouvant permettre de vous identifier. Si vous désirez en savoir plus, merci de consulter le [formulaire d'information et de consentement](#).

[English](#)



ANNEXE D – FORMULAIRE D’INFORMATION ET DE CONSENTEMENT DU SONDAGE



Titre du projet de recherche :

Méthodologie de design et d'évaluation d'un service de taxi par véhicules électriques
(projet de recherche financé par le CRNSG, InnovÉE, Taxelco et Fraxion)

Équipe de recherche :

Vincent Chabin
Candidat à la maîtrise en planification des transports
École Polytechnique de Montréal
C.P. 6079, succ. Centre-ville
Montréal (Québec)
H3C 3A7
Tél. (514) 340-4711 poste 4157
Adresse courriel : vincent.chabin@polymtl.ca

Nicolas Saunier
École Polytechnique de Montréal
C.P. 6079, succ. Centre-ville
Montréal (Québec)
H3C 3A7
Tél. (514) 340-4711 poste 4962
Adresse courriel : nicolas.saunier@polymtl.ca

Catherine Morency
École Polytechnique de Montréal
C.P. 6079, succ. Centre-ville
Montréal (Québec)
H3C 3A7
Tél. (514) 340-4711 poste 4502
Adresse courriel : catherine.morency@polymtl.ca

Préambule :

Nous vous invitons à participer à un projet de recherche qui vise à mieux connaître l'utilisation du service de taxi électrique Téo Taxi dans la région de Montréal car vous avez effectué un trajet de taxi en le commandant par l'application Téo Taxi pendant une des 3 phases de collecte de données d'une durée de 3 semaines (automne 2018, hiver 2019, été 2019).

Présentation du projet de recherche et de ses objectifs :

Ce questionnaire s'inscrit dans le cadre du projet de recherche :

« Méthodologie de design et d'évaluation d'un service de taxi par véhicules électriques »

Cette recherche a pour but d'améliorer la compréhension de l'utilisation d'un service de taxi électrique en se focalisant sur les typologies des utilisateurs. Les objectifs de l'étude que nous réalisons sont les suivants : i) Définir des indicateurs de variabilité spatio-temporelle des usages qui soient adaptés à l'utilisation du taxi; ii) Développer des typologies d'utilisateurs basées sur la variabilité spatio-temporelle des usages et l'influence d'autres facteurs exogènes ; et iii) Effectuer une enquête permettant de mieux comprendre les motifs déterminant l'usage du taxi électrique.

Afin d'atteindre nos objectifs, nous effectuons une enquête auprès des clients du service Téo Taxi afin de mieux comprendre les facteurs liés au choix de prendre un taxi électrique pour un déplacement en particulier. Cette enquête est développée sur plate-forme Web.

Nature et durée de votre participation au projet de recherche :

Dans le cadre de ce projet, votre participation consiste à répondre à une série de questions portant sur votre expérience d'utilisation du service Téo Taxi pour effectuer des déplacements en particulier. Toutes les réponses que vous fournirez seront confidentielles. Nous estimons que vous aurez besoin d'environ dix minutes pour compléter le questionnaire disponible à l'adresse suivante : [sondage](#). Si vous acceptez de participer au présent projet, veuillez simplement compléter le questionnaire sur le site internet ([sondage](#)) et valider vos réponses.

Dans le cadre de ce projet, seules les personnes âgées de 18 ans et plus et ayant effectué un trajet de Téo Taxi pendant les 3 phases de collecte de données, sont invitées à participer à la présente étude.

Avantages pouvant découler de votre participation au projet de recherche :

Vous ne retirerez aucun bénéfice personnel de votre participation au présent projet de recherche. Toutefois, les connaissances acquises grâce à votre participation permettront d'améliorer nos connaissances sur l'usage du taxi électrique dans la région de Montréal.

Inconvénients pouvant découler de votre participation au projet de recherche :

Votre participation au projet de recherche nécessitera environ dix minutes de votre temps.

Risques pouvant découler de votre participation au projet de recherche :

Vous pourriez éventuellement ne pas savoir quoi répondre à une question. Dans une telle éventualité, sentez-vous bien à l'aise de ne pas répondre à la question.

Il existe un risque de fuite des données confidentielles du répondant. Cependant ce risque reste limité puisque les données seront hébergées sur un serveur sécurisé par un accès SSH. Seuls les administrateurs du serveur (l'équipe de recherche) auront accès à ce serveur.

De plus pour prévenir tout risque de perte, vol, ou fuite de données, la plateforme utilisée est sécurisée par le protocole de communication HTTPS. Si vous avez plus de question concernant la sécurisation des échanges de données, vous pouvez vous référer au certificat de sécurité de la page web.

Compensation financière :

Vous ne recevrez aucune compensation financière pour votre participation au projet de recherche.

Indemnisation en cas de préjudice et droits du participant :

Si vous deviez subir quelque préjudice que ce soit par suite de votre participation à ce projet de recherche, vous ne renoncez à aucun de vos droits ni ne libérez les chercheurs, l'organisme subventionnaire ou l'établissement de leurs responsabilités légales et professionnelles.

Participation volontaire et possibilité de retrait :

Votre participation à ce projet de recherche est volontaire. Vous êtes donc libre de refuser d'y participer en ne répondant pas au questionnaire.

Les chercheurs ou le comité de la recherche de l'École Polytechnique pourront retirer les participants sans leur consentement, s'ils ne respectent pas les consignes du projet de recherche ou s'il existe des raisons administratives d'abandonner le projet, notamment pour des raisons de sécurité et de faisabilité.

Confidentialité :

Sur réception du questionnaire dûment rempli, le chercheur responsable du projet recueillera et consignera les réponses aux questions auxquelles vous aurez répondu.

Seuls les renseignements nécessaires à la bonne conduite du projet de recherche seront recueillis dans le cadre de la présente étude. Ces renseignements peuvent comprendre les informations concernant votre position, votre expérience de mobilité, ou encore des informations sur votre profil sociodémographique. Tous les renseignements recueillis au cours du projet de recherche demeureront strictement confidentiels dans les limites prévues par la loi. Afin de préserver votre identité et la confidentialité de ces renseignements, vous ne serez identifié que par un code. Les données collectées par le chercheur principal seront conservées dans un ordinateur protégé par un code et à accès limité dans son bureau à l'École Polytechnique de Montréal. Le chercheur responsable utilisera les données du projet de recherche pour les simples fins du projet de recherche.

Les résultats du projet de recherche pourront être publiés dans des revues scientifiques ou partagés avec d'autres personnes lors de discussions scientifiques. Toutefois, aucune publication ou communication scientifique ne renfermera quelque information que ce soit pouvant permettre de vous identifier. À des fins de surveillance et de contrôle, vos réponses au questionnaire pourra être consulté par une personne mandatée par le Comité d'éthique de la recherche de l'École Polytechnique de Montréal ou encore une personne mandatée par les organismes subventionnaires de recherche. Toutes ces personnes et ces organismes adhèrent à une politique de confidentialité.

Vous avez le droit de consulter vos réponses au questionnaire pour vérifier l'exactitude des renseignements recueillis aussi longtemps que le chercheur responsable du projet de recherche, ou l'établissement détiennent ces informations. Cependant, afin de préserver l'intégrité scientifique du projet de recherche, vous n'aurez accès à certaines de ces informations qu'une fois l'étude terminée. Les données seront conservées pour une durée de 10 ans et détruites après cette période.

Personnes ressource :

Si vous avez des questions concernant le projet de recherche, vous pouvez communiquer avec Vincent Chabin au (514) 340-4711, poste 4157 ou encore par courriel à vincent.chabin@polymtl.ca.

Si vous avez des questions concernant votre participation au projet de recherche, vous pouvez communiquer avec le président du Comité d'éthique de la recherche de l'École Polytechnique, M Yuvin Chinniah, au (514) 340-4711, poste 2268 ou encore par courriel à yuvn.chinniah@polymtl.ca

Consentement :

En soumettant ce sondage dûment complété, vous acceptez de participer à ce projet de recherche sujet aux conditions précisées ci-dessus.

ANNEXE E – TABLE DE VALEURS DU TEST DU χ^2

| | 0.995 | 0.99 | 0.975 | 0.95 | 0.9 | 0.75 | 0.5 | 0.25 | 0.1 | 0.05 | 0.025 | 0.01 | 0.005 |
|----|-----------|-----------|-----------|----------|---------|--------|--------|-------|-------|-------|-------|-------|-------|
| 1 | 3.927e-05 | 0.0001571 | 0.0009821 | 0.003932 | 0.01579 | 0.1015 | 0.4549 | 1.323 | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 |
| 2 | 0.01003 | 0.0201 | 0.05064 | 0.1026 | 0.2107 | 0.5754 | 1.386 | 2.773 | 4.605 | 5.991 | 7.378 | 9.21 | 10.6 |
| 3 | 0.07172 | 0.1148 | 0.2158 | 0.3518 | 0.5844 | 1.213 | 2.366 | 4.108 | 6.251 | 7.815 | 9.348 | 11.34 | 12.84 |
| 4 | 0.207 | 0.2971 | 0.4844 | 0.7107 | 1.064 | 1.923 | 3.357 | 5.385 | 7.779 | 9.488 | 11.14 | 13.28 | 14.86 |
| 5 | 0.4117 | 0.5543 | 0.8312 | 1.145 | 1.61 | 2.675 | 4.351 | 6.626 | 9.236 | 11.07 | 12.83 | 15.09 | 16.75 |
| 6 | 0.6757 | 0.8721 | 1.237 | 1.635 | 2.204 | 3.455 | 5.348 | 7.841 | 10.64 | 12.59 | 14.45 | 16.81 | 18.55 |
| 7 | 0.9893 | 1.239 | 1.69 | 2.167 | 2.833 | 4.255 | 6.346 | 9.037 | 12.02 | 14.07 | 16.01 | 18.48 | 20.28 |
| 8 | 1.344 | 1.646 | 2.18 | 2.733 | 3.49 | 5.071 | 7.344 | 10.22 | 13.36 | 15.51 | 17.53 | 20.09 | 21.95 |
| 9 | 1.735 | 2.088 | 2.7 | 3.325 | 4.168 | 5.899 | 8.343 | 11.39 | 14.68 | 16.92 | 19.02 | 21.67 | 23.59 |
| 10 | 2.156 | 2.558 | 3.247 | 3.94 | 4.865 | 6.737 | 9.342 | 12.55 | 15.99 | 18.31 | 20.48 | 23.21 | 25.19 |
| 11 | 2.603 | 3.053 | 3.816 | 4.575 | 5.578 | 7.584 | 10.34 | 13.7 | 17.28 | 19.68 | 21.92 | 24.72 | 26.76 |
| 12 | 3.074 | 3.571 | 4.404 | 5.226 | 6.304 | 8.438 | 11.34 | 14.85 | 18.55 | 21.03 | 23.34 | 26.22 | 28.3 |
| 13 | 3.565 | 4.107 | 5.009 | 5.892 | 7.042 | 9.299 | 12.34 | 15.98 | 19.81 | 22.36 | 24.74 | 27.69 | 29.82 |
| 14 | 4.075 | 4.66 | 5.629 | 6.571 | 7.79 | 10.17 | 13.34 | 17.12 | 21.06 | 23.68 | 26.12 | 29.14 | 31.32 |
| 15 | 4.601 | 5.229 | 6.262 | 7.261 | 8.547 | 11.04 | 14.34 | 18.25 | 22.31 | 25 | 27.49 | 30.58 | 32.8 |
| 16 | 5.142 | 5.812 | 6.908 | 7.962 | 9.312 | 11.91 | 15.34 | 19.37 | 23.54 | 26.3 | 28.85 | 32 | 34.27 |
| 17 | 5.697 | 6.408 | 7.564 | 8.672 | 10.09 | 12.79 | 16.34 | 20.49 | 24.77 | 27.59 | 30.19 | 33.41 | 35.72 |
| 18 | 6.265 | 7.015 | 8.231 | 9.39 | 10.86 | 13.68 | 17.34 | 21.6 | 25.99 | 28.87 | 31.53 | 34.81 | 37.16 |
| 19 | 6.844 | 7.633 | 8.907 | 10.12 | 11.65 | 14.56 | 18.34 | 22.72 | 27.2 | 30.14 | 32.85 | 36.19 | 38.58 |
| 20 | 7.434 | 8.26 | 9.591 | 10.85 | 12.44 | 15.45 | 19.34 | 23.83 | 28.41 | 31.41 | 34.17 | 37.57 | 40 |
| 21 | 8.034 | 8.897 | 10.28 | 11.59 | 13.24 | 16.34 | 20.34 | 24.93 | 29.62 | 32.67 | 35.48 | 38.93 | 41.4 |
| 22 | 8.643 | 9.542 | 10.98 | 12.34 | 14.04 | 17.24 | 21.34 | 26.04 | 30.81 | 33.92 | 36.78 | 40.29 | 42.8 |
| 23 | 9.26 | 10.2 | 11.69 | 13.09 | 14.85 | 18.14 | 22.34 | 27.14 | 32.01 | 35.17 | 38.08 | 41.64 | 44.18 |
| 24 | 9.886 | 10.86 | 12.4 | 13.85 | 15.66 | 19.04 | 23.34 | 28.24 | 33.2 | 36.42 | 39.36 | 42.98 | 45.56 |
| 25 | 10.52 | 11.52 | 13.12 | 14.61 | 16.47 | 19.94 | 24.34 | 29.34 | 34.38 | 37.65 | 40.65 | 44.31 | 46.93 |
| 26 | 11.16 | 12.2 | 13.84 | 15.38 | 17.29 | 20.84 | 25.34 | 30.43 | 35.56 | 38.89 | 41.92 | 45.64 | 48.29 |
| 27 | 11.81 | 12.88 | 14.57 | 16.15 | 18.11 | 21.75 | 26.34 | 31.53 | 36.74 | 40.11 | 43.19 | 46.96 | 49.64 |
| 28 | 12.46 | 13.56 | 15.31 | 16.93 | 18.94 | 22.66 | 27.34 | 32.62 | 37.92 | 41.34 | 44.46 | 48.28 | 50.99 |
| 29 | 13.12 | 14.26 | 16.05 | 17.71 | 19.77 | 23.57 | 28.34 | 33.71 | 39.09 | 42.56 | 45.72 | 49.59 | 52.34 |
| 30 | 13.79 | 14.95 | 16.79 | 18.49 | 20.6 | 24.48 | 29.34 | 34.8 | 40.26 | 43.77 | 46.98 | 50.89 | 53.67 |

TABLE 2 – Table de la loi du χ^2 : les lignes correspondent au nombre de degrés de liberté (indiqué dans la première colonne), les colonnes au risque de faire une erreur de type I (indiqué sur la première ligne) et le risque de première espèce est la probabilité qu'une variable aléatoire X suivant la loi du χ^2 à d degré de liberté soit supérieure à la valeur correspondante dans le tableau.

Valeurs obtenues à partir de (Scipy, 2019)

ANNEXE F – TABLEAU DE RÉSULTAT DU TEST DU X² ENTRE LA DISTRIBUTION DES RÉPONSES DES SUPER-UTILISATEURS ET DES UTILISATEURS NORMAUX

| N° de question | Question | Statistique du test | Valeur p | Valeur seuil à p=0,05 |
|----------------|---|---------------------|----------|-----------------------|
| 2,a | <i>Tous les passagers ont la même destination</i> | 0,01 | 0,935 | 3,81 |
| 5 | <i>Trajet multimodal</i> | 0,04 | 0,837 | 3,81 |
| 3,a | <i>Mode alternatives envisagées</i> | 3,16 | 0,676 | 11,07 |
| 10 | <i>Connaissance de Transit</i> | 0,37 | 0,542 | 3,81 |
| 3 | <i>Alternatives envisagées</i> | 3,14 | 0,535 | 9,49 |
| 6 | <i>Qui a payé pour la course ?</i> | 1,94 | 0,379 | 5,99 |
| 14 | <i>Possession du permis de conduire</i> | 1,52 | 0,218 | 3,81 |
| 4 | <i>Raisons du choix de Téo Taxi</i> | 14,41 | 0,155 | 18,31 |
| 13 | <i>Genre</i> | 3,77 | 0,152 | 5,99 |
| 9 | <i>Fréquence d'utilisation d'Uber</i> | 6,98 | 7,25E-02 | 7,82 |
| 11 | <i>Âge</i> | 17,68 | 6,06E-02 | 18,31 |
| 16 | <i>Taille du ménage</i> | 11,99 | 3,49E-02 | 11,07 |
| 18 | <i>Revenu du ménage</i> | 19,32 | 1,32E-02 | 15,51 |
| 8 | <i>Fréquence d'utilisation du taxi</i> | 14,46 | 2,34E-03 | 7,82 |
| 1 | <i>Motif du déplacement</i> | 19,36 | 1,65E-03 | 11,07 |
| 17 | <i>Nombre de voitures dans le ménage</i> | 15,69 | 1,31E-03 | 7,82 |
| 12 | <i>Occupation principale</i> | 27,31 | 1,27E-04 | 18,31 |
| 7 | <i>Service de mobilité</i> | 30,64 | 7,24E-05 | 14,07 |
| 2 | <i>Courses avec plusieurs passager</i> | 46,88 | 3,69E-10 | 7,82 |

Valeur en rouge : valeur non significative avec un risque de première espèce de 0,05

ANNEXE G – TABLEAU DE RÉSULTAT DU TEST DU X² ENTRE LA DISTRIBUTION DES RÉPONSES POUR LES COURSES RÉALISÉES EN JOURNÉE ET EN SOIRÉE

| N° de question | Question | Statistique du test | Valeur p | Valeur seuil à p=0,05 |
|----------------|---|---------------------|----------|-----------------------|
| 10 | <i>Connaissance de Transit</i> | 0,02 | 0,889 | 3,81 |
| 5 | <i>Trajet multimodal</i> | 0,11 | 0,735 | 3,81 |
| 13 | <i>Genre</i> | 1,61 | 0,448 | 5,99 |
| 14 | <i>Possession du permis de conduire</i> | 0,66 | 0,418 | 3,81 |
| 7 | <i>Service de mobilité</i> | 7,93 | 0,339 | 14,07 |
| 3,a | <i>Mode alternatives envisagées</i> | 5,58 | 0,232 | 9,49 |
| 8 | <i>Fréquence d'utilisation du taxi</i> | 4,35 | 0,226 | 7,82 |
| 16 | <i>Taille du ménage</i> | 7,07 | 0,215 | 11,07 |
| 18 | <i>Revenu du ménage</i> | 11,12 | 0,195 | 15,51 |
| 2 | <i>Courses avec plusieurs passager</i> | 5,05 | 0,168 | 7,82 |
| 12 | <i>Occupation principale</i> | 9,49 | 0,148 | 12,59 |
| 3 | <i>Alternatives envisagées</i> | 9,14 | 0,104 | 11,07 |
| 9 | <i>Fréquence d'utilisation d'Uber</i> | 6,75 | 8,03E-02 | 7,82 |
| 6 | <i>Qui a payé pour la course ?</i> | 6,58 | 3,73E-02 | 5,99 |
| 17 | <i>Nombre de voitures dans le ménage</i> | 15,87 | 1,20E-03 | 7,82 |
| 11 | <i>Âge</i> | 30,94 | 3,03E-04 | 16,92 |
| 2,a | <i>Tous les passagers ont la même destination</i> | 14,50 | 1,40E-04 | 3,81 |
| 4 | <i>Raisons du choix de Téo Taxi</i> | 56,97 | 1,35E-08 | 18,31 |
| 1 | <i>Motif du déplacement</i> | 84,65 | 1,80E-17 | 9,49 |

Valeur en rouge : valeur non significative avec un risque de première espèce de 0,05

**ANNEXE H – TABLEAU DE RÉSULTAT DU TEST DU X² ENTRE LA
DISTRIBUTION DES RÉPONSES POUR LES COURSES AFILIÉES À
AUX GROUPES DE FORTE INTENSITÉ D’UTILISATION ET AUX
GROUPES DE FAIBLE INTENSITÉ D’UTILISATION**

| <i>N° de question</i> | <i>Question</i> | <i>Statistique du test</i> | <i>Valeur p</i> | <i>Valeur seuil à p=0,05</i> |
|-----------------------|---|----------------------------|-----------------|------------------------------|
| 2,a | <i>Tous les passagers ont la même destination</i> | 0,06 | 0,801 | 3,81 |
| 9 | <i>Fréquence d'utilisation d'Uber</i> | 2,06 | 0,559 | 7,82 |
| 13 | <i>Genre</i> | 1,78 | 0,410 | 5,99 |
| 10 | <i>Connaissance de Transit</i> | 0,75 | 0,386 | 3,81 |
| 14 | <i>Possession du permis de conduire</i> | 1,12 | 0,289 | 3,81 |
| 6 | <i>Qui a payé pour la course ?</i> | 3,15 | 0,207 | 5,99 |
| 17 | <i>Nombre de voitures dans le ménage</i> | 5,18 | 0,159 | 7,82 |
| 3 | <i>Alternatives envisagées</i> | 11,09 | 8,57E-02 | 12,59 |
| 3,a | <i>Mode alternatives envisagées</i> | 10,41 | 3,40E-02 | 9,49 |
| 11 | <i>Âge</i> | 19,58 | 3,35E-02 | 18,31 |
| 5 | <i>Trajet multimodal</i> | 4,70 | 3,01E-02 | 3,81 |
| 18 | <i>Revenu du ménage</i> | 21,57 | 1,04E-02 | 16,92 |
| 4 | <i>Raisons du choix de Téo Taxi</i> | 24,99 | 5,36E-03 | 18,31 |
| 12 | <i>Occupation principale</i> | 20,65 | 2,12E-03 | 12,59 |
| 2 | <i>Courses avec plusieurs passager</i> | 19,10 | 2,61E-04 | 7,82 |
| 16 | <i>Taille du ménage</i> | 26,13 | 8,44E-05 | 11,07 |
| 7 | <i>Service de mobilité</i> | 31,46 | 5,11E-05 | 14,07 |
| 8 | <i>Fréquence d'utilisation du taxi</i> | 49,04 | 1,28E-10 | 7,82 |
| 1 | <i>Motif du déplacement</i> | 78,79 | 1,51E-15 | 11,07 |

Valeur en rouge : valeur non significative avec un risque de première espèce de 0,05