



Etude des facteurs associés à l'usage fréquent des transports en commun urbain

Nikita Gусаров, Tom Viellescazes

Niveau d'études :

Master 2

Parcours :

Chargé d'études économiques et statistiques

Tuteur :

Iragaël Joly

Université Grenoble Alpes
Faculté d'économie et gestion

2019 - 2020

Contents

Introduction	1
1 Revue de la littérature	1
2 Présentation des variables	3
3 Modélisation	6
4 Conclusion	10
Annexes	11
References	19

Introduction

Qu'est-ce qui pousse les gens à utiliser ou à préférer les transports en commun aux autres modes de transport? Qu'est-ce qui définit la fréquence d'utilisation de ces transports? Une compréhension des facteurs qui affectent le choix de fréquence et du mode est essentielle à la promotion de comportements plus durables et à la réalisation des objectifs de transport de l'État.

Cette étude vise à étudier les facteurs associés à l'usage fréquent (tous les jours) des transports en commun urbain afin de mieux comprendre les déterminants clefs du choix des usagers. En effet une meilleure connaissance de ces déterminants permet aux villes de savoir quel levier utilisé pour diminuer le nombre d'usagers privilégiant la voiture aux transports au commun ou aux autres alternatives. On peut remarquer que depuis un certain nombre d'années les grandes villes européennes essayent de limiter la circulation en ville et tout particulièrement en centre-ville afin de diminuer la pollution de l'air(particule fine). Grenoble étant particulièrement touchée par des problèmes de pollution, partiellement due à sa situation géographique, elle engage elle aussi des politiques de réduction de la pollution par le trafique à travers le système de vignette ne permettant pas aux voitures les plus polluantes de circuler lors de piques de pollution.

Nous allons commencer par donner un aperçu des études scientifiques qui forment la base pour notre étude. Puis nous allons décrire la base de données que nous avons utilisée en détaillant comment nous avons choisi la composition de cette base, la nature des données et les statistiques descriptives pour les variables clés. Dans la partie 3 nous construirons et explorerons les différents modèles économétriques. enfin nous conclurons.

1 Revue de la littérature

Nous avons commencé notre étude de la littérature par la recherche des articles méthodologiques et théoriques en lien avec le sujet traité.

Les ressources informatiques offertes par notre établissement (Université Grenoble Alpes) aussi bien que d'autres sources en libre accès nous ont permis de rechercher les articles d'intérêt. Les moteurs de recherche scientifique les plus utilisés ont été :

- Science Direct
- Google Scholar
- Elsevier
- SAGE journals

Ces outils informatiques nous ont permis de sélectionner les articles les plus pertinents selon le nombre des citations des articles.

1.1 La base théorique

Nous avons commencé notre étude de la littérature par une étude théorique des différents types du comportement existants dans le domaine de transport. Dans notre cas d'étude, adressant le problème de la modélisation de la fréquence d'utilisation du transport en commun, il nous paraissait que l'article Gärling and Axhausen (n.d.) fut le mieux adapté.

L'auteur remarque que les habitudes de déplacement quotidiennes ont tendance à se répéter (Muthyalagari and Pendyala 2001). Étant donné que le comportement habituel en matière de transport implique que les mêmes choix sont faits à maintes reprises, il n'est pas raisonnable de supposer que ces choix sont le résultat d'une délibération. Ce qui nous renvoie à notre cas d'étude, où nous risquerons selon cette logique d'estimer les effets non-répresentatives, si nous traitons des caractéristiques des individus au niveau très fin.

En effet, il serait complètement inapproprié d'interpréter les résultats obtenus lors d'estimation comme des variables ayant une influence directe. Le modèle ne sera qu'une approximation assez faible dans notre cas d'étude. Nous supposons davantage, que la marge d'erreur sera suffisamment large pour le modèle décrivant la fréquence d'utilisation du transport en commun, car une fréquence élevée reflète en soi un choix habituel. De plus, nous ignorons les choix habituels pour le reste des modes de transport.

Nous avons également essayé de positionner notre étude en relation avec les autres travaux évaluant la fréquence d'utilisation du transport. L'article de Truong and Somenahalli (2015) répondait à notre demande, car les auteurs effectuaient une étude empirique, explorant les facteurs influençant la fréquence d'utilisation des transports publics chez les personnes âgées à Adélaïde, en Australie. Plus précisément, ils construisent un modèle logistique multinomiale pour modéliser la fréquence d'utilisation des transports publics.

Nous avons également exploré les différents articles appliqués afin d'avoir un aperçu des variables utilisées dans les modèles décrivant le choix de transport. Parmi ces travaux on peut souligner l'ouvrage de Corpuz (2007). Les auteurs de cet article visent à identifier et étudier les variables déterminantes dans le choix du mode de déplacement. Dans ce dernier travail des approches univariées et qualitatives plus simples sont utilisées pour analyser tour à tour les effets des variables d'influence individuellement.

Pour terminer cette partie, portant sur la littérature utilisée dans ce travail, nous allons énumérer les différents résultats des articles que nous avons étudié.

Puisque nous étudions la fréquence d'utilisation de transport en commun, nous nous intéressons principalement aux résultats et découvertes reliés à ce domaine. Corpuz (2007) identifie les caractéristiques socio-démographiques des voyageurs, qui utilisent le transport en commun le plus :

- Les individus ayant de 11 à 20 ans ,
- Les étudiants,
- Ceux sans permis de conduire,
- Ceux sans véhicule disponible dans le ménage.

Selon Gärling and Axhausen (n.d.), dans la littérature la description des voyageurs se limite généralement à des variables sociodémographiques telles que :

- âge,
- sexe,
- éducation,
- le revenu.

Quand à Truong and Somenahalli (2015), les résultats de leurs étude montrent qu'une utilisation fréquente des transports en commun est étroitement associée à :

- une très grande importance perçue des transports en commun pour les zones résidentielles,
- un accès facile aux transports en commun dans les quartiers,
- l'utilisation hebdomadaire des transports publics est également légèrement liée à la possession de téléphones portables,
- aux groupes d'âge plus jeunes.

Bien qu'on a dans notre disposition la plupart de ces variables, figurant dans la base des données étudiée, nous ne pourrons pas contrôler pour certains facteurs (ex: accès aux transports en commun en termes de temps de marche).

1.2 Hypothèses

Pour conclure cette partie du papier nous allons lister les hypothèses clés que nous avons retenu des articles étudiés, qui peuvent s'appliquer dans le contexte des données étudiées.

D'abord nous étudions l'attitude envers un mode de transport unique, c'est-à-dire nous nous contenterons d'étudier les déterminants de la prise des transports en commun s'en prendre en compte des moyens de déplacement alternatif (vélo, trottinette électrique, et cetera). Puisque nous étudions la fréquence d'utilisation, nous n'étudierons pas non plus les différents moyens de locomotion utilisés à l'intérieur d'un déplacement (Truong and Somenahalli 2015) de plus les habitudes et la fréquence d'utilisation d'un mode de transport peut avoir des déterminants différents de ceux des transports en commun. Ce qui pousserait si on voulait étudier ses différents modes à ajouter un nombre conséquent de variable ce qui augmenterait les chances de colinéarité entre les variables explicatives.

Le transport en commun doit être disponible pour la zone unique et identique pour tous les individus afin que l'on ne compare pas des individus avec des caractéristiques trop différentes. Par exemple les individus habitant dans la campagne grenobloise et n'ayant pas ou peu accès aux transports en commun ne peuvent pas être comparé avec des individus en centre-ville qui peuvent aisément se déplacer sans voiture. On veut donc à l'intérieur de notre échantillon seulement les individus pour qui l'arbitrage entre transports en communs et autres moyen de déplacement est réel. Il faut donc que l'ensemble des individus puisse avoir accès aux transports en commun. Nous supposerons donc que l'accessibilité des transports en commun est la même pour l'ensemble des individus.

Nous supposons également que seules les caractéristiques spécifiques à l'individu jouent leurs rôles. En effet puisque nous étudions le comportement de déplacement au niveau de l'individu nous ne pouvons pas prendre en compte les caractéristiques de déplacement comme la durée ou l'heure de celui-ci. De plus nous ne prenons pas non plus en compte les caractéristiques du lieu d'arrivée ou de départ ce qui peut encore une fois jouer sur l'arbitrage entre voiture et transport en commun (Prix du stationnement, manque de parkings).

2 Présentation des variables

Passons maintenant à l'étude de la base des données. Les données qu'on explore décrivent les déplacements réalisés par les membres des ménages grenoblois en 2010. Ils ont renseigné des données relatives aux caractéristiques individuelles, aux caractéristiques du ménage (structure du ménage, motorisation, sexe, âge, statut professionnel, etc.) et aux pratiques de mobilité (nombres de déplacements, horaires des déplacements, localisations des déplacements, modes utilisés, motifs des déplacements, etc.).

Nous commencerons par une description de la façon dont nous créons notre échantillon à étudier. Ensuite, nous allons présenter les variables qui apparaîtront dans notre analyse.

2.1 Composition de l'échantillon

Puisque nous nous concentrons dans cette étude sur le comportement individuel en ce qui concerne les choix de déplacement quotidien dans l'agglomération grenobloise, et plus particulièrement celui des transports en commun (les transports faisant partie du réseau TAG), notre base des données contient des données au niveau individuel. C'est-à-dire, pour chaque individu dans la base de données initiale nous agrégeons les données suivant quelques règles décrites ci-après.

Nous nous intéressons seulement aux individus qui peuvent théoriquement utiliser le réseau TAG dans leurs déplacements. C'est pourquoi nous gardons dans notre échantillon final seulement les personnes qui ont effectué au moins un déplacement complet au sein de l'agglomération grenobloise. C'est-à-dire, ceux qui ont effectué un déplacement où les lieux de départ et d'arrivée se trouvent tous les deux dans l'agglomération grenobloise. De plus cela nous permettent d'exclure les individus trop loin des accès aux transports en commun afin de garder seulement des individus pouvant utiliser les transports en commun sans effort démesuré. Grâce à cela nous ne gardons seulement les individus réalisant un véritable arbitrage entre voitures et transports en commun.

De plus, afin d'exclure de notre échantillon les professionnels, nous excluons les individus qui effectuent plus de 11 déplacements dans leur journée. Cela nous permet d'éviter le biais apporté par les personnes qui gagnent leur vie en effectuant plusieurs déplacements dans la journée (ie: la livraison).

Finalement, nous excluons de notre échantillon les individus qui sont en cours d'obtention du permis de conduire. Ce groupe est peu nombreux et n'a pas de puissance statistique dans le cas de notre étude, sans compter que si ces individus sont gardés dans l'échantillon final, nous risquons biaiser nos estimateurs.

L'échantillon final comprend les données concernant 3050 individus.

La variable dépendante que nous visons à étudier décrit l'utilisation du réseau TAG par les habitants de Grenoble. Elle prend la valeur 1 si l'individu utilise le transport du réseau TAG chaque jour ouvert et 0 dans tous les autres cas. Dans l'annexe C1 le lecteur pourra observer la distribution spatiale des moyennes par zone géographique.

Les variables descriptives principales imposées par l'énoncé sont :

- le fait d'avoir un permis de conduire ;
- la taille du ménage.

Les deux variables rentrent dans l'idéologie générale des études des choix individuels dans l'utilisation du transport.

De plus, suite à nos explorations des études dans ce domaine, nous allons inclure les variables des caractéristiques individuelles suivantes :

- Age (Gärling and Axhausen, n.d.);
- Sexe (Gärling and Axhausen, n.d.);
- Nombre des voitures disponibles dans le ménage (Truong and Somenahalli 2015);
- Nombre des déplacements dans la journée (Truong and Somenahalli 2015);
- Activité de l'individu (Corpuz 2007).

Dans cette partie, nous allons étudier ces variables une par une en expliquant les fondements théoriques qui nous permettent de supposer un lien avec la variable d'intérêt. De plus, pour chaque variable nous effectuerons une analyse bivariée en les croisant avec la variable dépendante, ainsi que donné un aperçu des statistiques descriptives clés.

2.2 Presentation des variables explicatives

Les permis de conduire

Les permis de conduire Dans la base des données le fait d'avoir un permis de conduire est codé sous la forme d'une variable discrète avec trois niveaux différents :

- individu possède un permis de conduire ;
- individu n'a aucun permis de conduire ;
- individu est en période d'apprentissage et est en train de recevoir son permis.

Cette variable risque d'être déterminée par d'autres caractéristiques individuelles, telles que l'âge ou le nombre des véhicules dans le ménage. Ce fait risque dans le futur d'introduire un biais de multicolinéarité.

Dans le tableau suivant nous pouvons voir la répartition de notre échantillon par le type de permis:

Table 1: Permis de conduire par type du comportement

	Non (N=2560)	Oui (N=490)	Total (N=3050)	p value
Permis de conduire				< 0.001
Oui	1935 (75.6%)	267 (54.5%)	2202 (72.2%)	
Non	625 (24.4%)	223 (45.5%)	848 (27.8%)	

Nous pouvons voir que dans les individus n'utilisant pas les transports en commun de façon quotidienne 75,6% ont le permis tandis que dans le groupe des personnes utilisant les transports communs seulement 54.4% ont leur permis. Cet écart en utilisation des transports en commun devient encore plus flagrant lorsque l'on voit que seulement 12.1% des individus ayant le permis choisissent les transports en commun comme moyen de transport fréquent contre 26.3% pour les individus sans permis de conduire.

Nous pouvons attendre à ce que les individus sans permis de conduire sont plus susceptibles d'utiliser les transports en commun.

La taille de ménage

Dans la base des données initiale, la taille de ménage est codée par une variable discrète avec 8 niveaux différents, ce qui donne la possibilité, si besoin, de la traiter comme une variable continue. Pour nous il sera

plus facile de la traiter d'une telle façon, bien qu'il faudra justifier ce passage. Le comportement de cette variable risque d'être non linéaire ;

Identiquement à la variable précédente nous risquons d'avoir des problèmes de multicolinéarité lors de la combinaisons de cette variable avec des autres variables reliées dans notre modèle.

Dans le tableau suivant nous observons tout d'abord qu'il existe peu de ménage avec plus de 6 individus (23 observations) on ne peut donc pas réellement analyser leur comportement. Ensuite nous pouvons voir que les ménages utilisant le plus les transports en commun sont les ménages avec un seul individu avec près de 27.5% qui choisissent les transports en commun. Les ménages à six et trois individus sont ensuite les ménages qui utilisent le plus les transports publics avec respectivement 18.3 et 16.22% des individus utilisant les transports en commun. Enfin les ménages les moins disposés à utiliser les transports en commun sont les ménages de deux, cinq et quatre individus qui choisissent les transports en commun respectivement dans 13.9, 13.4 et 11.4%. On pourra donc dire que les effets de la taille du ménage semblent en effet ne pas être linéaires.

Table 2: La taille de ménage par type du comportement

	Non (N=2560)	Oui (N=490)	Total (N=3050)	p value
Taille de ménage				< 0.001
1	334 (13.0%)	127 (25.9%)	461 (15.1%)	
2	810 (31.6%)	131 (26.7%)	941 (30.9%)	
3	439 (17.1%)	85 (17.3%)	524 (17.2%)	
4	588 (23.0%)	79 (16.1%)	667 (21.9%)	
5	315 (12.3%)	49 (10.0%)	364 (11.9%)	
6	58 (2.3%)	13 (2.7%)	71 (2.3%)	
7	11 (0.4%)	6 (1.2%)	17 (0.6%)	
8+	5 (0.2%)	0 (0.0%)	5 (0.2%)	

Nous supposons, que les ménages de la taille moyenne ont plus de probabilité d'avoir une voiture à leur disposition, ce qui conduit à une faible utilisation de transports en commun par ses membres.

Age

La variable d'âge est la seule variable descriptive strictement continue dans notre étude. Les individus concernés par l'enquête de déplacement 2010 ont plus de 5 ans.

Nous pouvons supposer que la variable âge permettra de capter les effets de cycles de vie si introduite dans le modèle sous la forme non-linéaire. Par exemple, un polynôme de troisième degré nous permettra de capter les effets des périodes suivantes :

- l'enfance, quand le mode de transport choisi et déterminé par les choix des parents
- l'adolescence
- la maturité (l'étape de vie le plus active)
- la retraite et la diminution d'activité

L'âge moyen par zone est présenté dans l'annexe C2.

Sexe

Nous supposons que les femmes et les hommes ont des rôles sociaux différents. Cette idée, supporté par des résultats de plusieurs études antérieures nous amène à inclure cette caractéristique comme une déterminante de choix de fréquence d'utilisation de transport en commun.

Le tableau ci-dessous nous offre un aperçu des différences entre les différentes parties de notre échantillon.

Table 3: Genre par type du comportement

	Non (N=2560)	Oui (N=490)	Total (N=3050)	p value
Sexe				< 0.001
Homme	1314 (51.3%)	197 (40.2%)	1511 (49.5%)	
Femme	1246 (48.7%)	293 (59.8%)	1539 (50.5%)	

Nombre des voitures disponibles

Le nombre de voitures disponibles combiné avec le nombre des personnes dans le ménage nous permet de mieux jauger la situation avec la disponibilité individuelle des voitures dans le ménage. Par exemple, il y a une différence infinie entre un ménage avec deux voitures comptant cinq personnes et un ménage avec deux voitures, mais avec seulement deux individus.

En ce qui concerne la nature de cette variable, c'est toujours une variable binaire. Nous pouvons voir la répartition des individus selon le choix de mode de transport pour nombre des voitures disponibles.

Table 4: Disponibilité de voiture par type du comportement

	Non (N=2560)	Oui (N=490)	Total (N=3050)	p value
Nobre des voitures				< 0.001
0	197 (7.7%)	162 (33.1%)	359 (11.8%)	
1	1044 (40.8%)	204 (41.6%)	1248 (40.9%)	
2	1065 (41.6%)	102 (20.8%)	1167 (38.3%)	
3	219 (8.6%)	19 (3.9%)	238 (7.8%)	
4	35 (1.4%)	3 (0.6%)	38 (1.2%)	

Nombre de déplacements dans la journée

Les besoins à se déplacer définissent le choix et la fréquence d'utilisation de tel ou tel mode de transport. Les individus vont inévitablement ajuster leur comportement en fonction de besoins et de son rythme de vie.

Les statistiques descriptives pour le croisement de la variable dépendante avec le nombre de déplacements dans la journée sont présentés dans l'annexe A3. Nous allons considérer cette variable continue comprise entre 0 et 10, suite à la façon spécifique par laquelle nous avons déterminé notre sous-échantillon pour l'étude.

Occupation principale de la personne

Le type d'occupation détermine le comportement individuel. Par exemple, nous pouvons supposer que les travailleurs à temps plein doivent avoir un comportement spécifique, ayant une préférence forte pour l'utilisation de la voiture (une fréquence nulle d'utilisation de transport en commun). Ces différences ressortent si on observe les parts d'utilisateurs du réseaux TAG par occupation principale (voir annexe A2).

3 Modélisation

Ayant fini avec la description des variables et la création de la base des données, nous pouvons enfin passer à la modélisation. Dans cette section du travail nous allons passer des modèles les plus simples, vers leurs versions plus élaborées en manipulant le nombre des variables explicatives.

3.1 Modèle simple sous continuités

Nous commençons par l'élaboration d'un modèle simple comprenant seulement les trois variables données dans l'énoncé. Cela nous évite les problèmes de multicolinéarité ainsi que certaines complications liés à

l'endogénéité éventuelle des certaines variables.

A cette étape nous ne prenons en considération que les variables suivantes :

- le fait d'avoir un permis de conduire ;
- la taille de ménage.

Suite à un étude préliminaire des variables descriptives nous imposons des restrictions sur le comportement de la variable décrivant la taille de ménage. Nous posons cette variable d'être continue avec des effets non-linéaire sur la variable dépendante (un polynôme au deuxième degré). Nous proposons les résultats des estimations sous format d'un tableau :

Nous proposons les résultats des estimations sous format d'un tableau :

Table 5: Le modèle binaire sous continuité

<i>Dependent variable:</i>	
Utilisation fréquente de TCU	
N'a pas de permis	1.218*** (0.112)
Taille de ménage	-0.863*** (0.140)
Taille de ménage au carré	0.091*** (0.021)
Constant	-0.535*** (0.201)
Observations	3,050
Log Likelihood	-1,264.666
Akaike Inf. Crit.	2,537.333

Note:

*p<0.1; **p<0.05; ***p<0.01

Tous les coefficients sont significatifs. Nous voyons que nos hypothèses sur le comportement et les effets des variables d'intérêts sont confirmées.

Le fait de ne pas avoir un permis de conduire augmente significativement la probabilité que l'individu utilise le réseau de transport commun chaque jour.

En ce qui concerne la taille de ménage, cette variable a effectivement un effet non-linéaire sur la probabilité d'utiliser le réseau TAG. Cet effet est positif décroissant, car la taille des ménages facilite le covoiturage sous limitation raisonnables.

Le pseudo-R2 est de 0.059, ce qui nous servira en futur pour la comparaison des modèles différents.

Effets marginaux moyens

Passons maintenant à l'étude des effets marginaux pour ce modèle réduit. Ces effets, calculés au point moyen de l'échantillon, sont regroupé dans le tableau suivant :

Il faut préciser, que nous préférons estimer les effets moyens, car ils reflètent mieux la situation dans l'échantillon que les effets calculé pour le point moyen.

	AME	SE	z	p	lower	upper
N'a pas de permis	0.1820	0.0183	9.9472	0	0.1461	0.2178
Taille de ménage	-0.0459	0.0060	-7.6772	0	-0.0576	-0.0342

Nous pouvons noter que l'effet marginal moyen de ne pas avoir le permis est compris entre 0.1461 et 0.2178 c'est-à-dire que moyenne le fait de ne pas avoir de permis augmente les chances de prendre les transports en commun de 14.6 à 21.8% pour un intervalle de confiance au seuil de 5%. L'effet marginal de la taille du ménage est quand à elle comprise entre -0.0576 et -0.0342 ce qu'on peut interpréter par le fait qu'une personne de plus dans le ménage diminue les chance d'une utilisation de fréquente des transports en communs de 3.4 à 5.7% dans un intervalle de confiance au seuil de 5%.

Présentation graphique des effets

Il peut être également intéressant d'étudier les impacts sur les prédictions par une représentation graphique. Nous utilisons le package R *margins* afin de construire les deux figures suivantes. Sur ces graphiques il est facile d'observer les effets que nos variables descriptives ont sur la probabilité d'utiliser le transport en commun chaque jour.

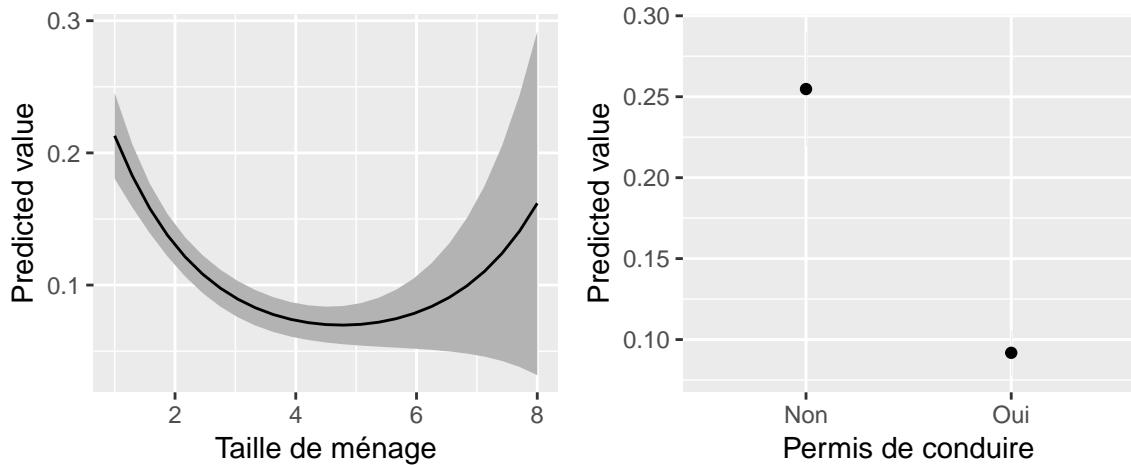


Figure 1: Les effets marginaux cas simple

Il faut noter, que bien que nous identifions correctement la nature des effets de la taille du ménage sur la variable dépendante, nous ne pouvons pas ignorer le taux d'erreur qui apparaîsse pour des valeurs limites de cette variable.

En ce qui concerne le permis de conduire il semble que le modèle permet mieux de prévoir le comportement des individus avec le permis que sans le permis. on distingue bien que l'intervalle de la valeur prédictée est plus grand pour les individus sans permis.

3.2 Modèle complexe

Maintenant nous pouvons passer à l'étude du modèle multivarié, qui prend en compte plusieurs facteurs possibles, influençant la fréquence d'utilisation des transports en communs.

Nous complétons le modèle obtenu précédemment par des variables explicatives supplémentaires, parmi lesquelles :

- Âge ;
- Sexe ;
- Nombre des voitures disponibles dans le ménage ;
- Nombre des déplacements dans la journée ;
- Activité de l'individu.

Nous respectons les hypothèses et les restrictions imposées sur les variables. C'est-à-dire, nous considérons que les effets d'âges ont une structure du polynôme de troisième degrés, que les effets du nombre de voitures

et le nombre des déplacements dans la journée sont linéaire, que seulement certains types d'activité ont un effet significatif sur la probabilité d'utiliser le réseau TAG chaque jours.

Ci-dessous nous offrons un aperçu des résultats obtenus lors des estimations (les résultats sous format complet, avec la variance des coefficients affichée, sont disponibles dans l'annexe B1).

Table 7: Les résultats du modèle avancé

<i>Dependent variable:</i>	
Utilisation fréquente de TCU	
N'a pas de permis	0.681***
Taille de ménage	-0.413**
Taille de ménage SQ	0.044*
Travaille temps plein	-1.036***
Travaille temps partiel	-1.297***
Age	0.292***
Age 2-ème	-0.007***
Age 3-ème	0.00005***
Sexe	0.354***
Nombre de déplacements	-0.101***
Nombre des voitures	-0.612***
Constant	-2.658***
Observations	3,050
Log Likelihood	-1,119.698
Akaike Inf. Crit.	2,263.396

Note:

*p<0.1; **p<0.05; ***p<0.01

Dans ce cas nous avons réussi à améliorer le pseudo-R2, qu'on utilise pour comparer les modèles. Particulièrement pour ce modèle sa valeur est de 0.167 ce qui démontre une amélioration significative par rapport au modèle simple. Pour une comparaison plus complète des modèles nous pourrions effectuer un analyse en utilisant des courbes ROC (les résultats de cette opération confirment une meilleure prédictivité du modèle se trouvent dans l'annexe B3).

Les effets des variables qu'on a étudiés précédemment restent inchangés aux niveau de leur signe. En ce qui concerne les autres variables, nous pouvons voir que la totalité de celle-ci sont significatives au seuil de 1% si ce n'est la taille des ménages au carré qui est significatif au seuil de 10%. Les variables qui vont alors jouer en faveur d'une prise des transports en commun de façon quotidienne sont l'âge, le sexe et la taille du ménage au carré avec des coefficients respectifs de 0.292, 0.354 et 0.044. L'âge au cube joue aussi un rôle positif extrêmement faible (ce qui est normal puisque valeur au cube) mais néanmoins significatif. Tandis que les variables qui diminuent les chances de l'utilisation fréquente des transports en commun sont le fait de travailler à temps plein, à temps partiel, le nombre de déplacements, le nombre de voitures ainsi que l'âge au carré.

En ce qui concerne les problèmes potentiels de ce modèle, nous pouvons nommer la multicollinéarité, laquelle est capté par un VIF test pour des variables non-linéaire (les résultats de ce test sont présentés dans l'annexe B3). En outre, le reste des variables ne rencontre pas ces difficultés.

Les effets

Par analogie avec le modèle précédent nous pouvons calculer les effets moyens pour l'ensemble des régresseurs.

Les effets par analogie avec le modèle précédent nous pouvons calculer les effets moyens pour l'ensemble des régresseurs. Dans ce cas nos intuitions sont également confirmées. Puisque les variables liées à l'activité ont bien un effet marginal moyen négatif. En effet le fait de travailler à temps plein diminue les chances de prendre les transports en commun de 11.6%, cet effet était attendu puisque les individus à temps plein risquent

d'avoir un besoin de flexibilité plus important, le nombre de déplacements a aussi un effet moyen négatif de 0.0113 pour les mêmes raisons. Les individus sans activité (Reste au foyer) ont quant à eux 14.56% de chance en moins de d'utiliser les transports en commun de façons quotidiennes ce qui peut être expliqué par un nombre de déplacements plus faibles que les autres individus et pour d'autres buts (course, aller chercher les enfants à l'école etc..). Le fait d'être une femme a comme prévu un effet positif sur la prise des transports en commun avec en moyenne 3.95% de chance de plus d'utiliser les transports en commun de façon quotidienne. Le nombre de voitures dans le ménage joue bien un rôle négatif avec un effet moyen de -0.0687. L'âge quant à lui est plus compliqué à interpréter comme nous l'avons vus dans le tableau précédent l'âge à plusieurs effets différents si on change sa forme ce qui est dû au fait que les utilisations des transports en commun sont plus liés à des classes d'âge qu'à l'âge de façon linéaire, ainsi les individus en âge de faire des études 18-25 ans ou à la retraite aura une tendance à utiliser les transports en commun de façon fréquente alors que le reste de la population utilise de préférence d'autre moyen de transport. Il reste que l'effet marginal moyen de l'âge est de 0.0028 ce qui peut s'interpréter : un individu à 0.28% de chance de plus d'utiliser les transports en commun de manière fréquente s'il a un an de plus.

Enfin on peut voir que les deux variables utilisées dans le modèle précédent gardent les mêmes signes mais que leur effet a diminué, l'effet moyen de la taille du ménage est passé de -0.0459 dans le premier modèle à 0.0182 dans le second. La possession du permis de conduire à un effet moyen de 0.0826 contre 0.1820 dans le premier modèle.

	AME	SE	z	p	lower	upper
Age	0.0028	0.0006	4.9161	0.0000	0.0017	0.0040
Nombre de déplacements	-0.0113	0.0033	-3.3993	0.0007	-0.0179	-0.0048
Reste à foyer	-0.1456	0.0403	-3.6102	0.0003	-0.2247	-0.0666
Femme	0.0395	0.0125	3.1653	0.0015	0.0151	0.0640
Temps pleins	-0.1163	0.0160	-7.2741	0.0000	-0.1476	-0.0849
Nombre de voitures	-0.0687	0.0089	-7.7409	0.0000	-0.0861	-0.0513
Pas de permis	0.0826	0.0202	4.0831	0.0000	0.0430	0.1223
Taille de ménage	-0.0182	0.0059	-3.1131	0.0019	-0.0297	-0.0068

De plus, de la même façon que pour le modèle simple nous nous intéressons aux effets des variables descriptives sur la probabilité d'utilisation du réseau de transport en commun par l'individu. Les résultats d'estimation des effets des différentes variables non-binaires sont regroupé dans l'annexe B2, ici nous n'offrons que quelque conclusions basées sur les résultats.

4 Conclusion

Ce travail présente les résultats d'une étude explorant les facteurs influençant la fréquence d'utilisation des transports publics chez les résidents du département de l'Isère. Les résultats ont montré qu'une utilisation plus fréquente des transports publics est étroitement associée aux caractéristiques individuelles telles que l'âge, le sexe, le fait d'avoir le permis ou non, le nombre de déplacements ou encore l'activité principale. Mais aussi à des caractéristiques du ménage telle que sa taille ou le nombre de voitures possédé. Il semble que les personnes qui sont le plus sensibles d'utiliser les transports en commun sont les jeunes autour des 25-26 (cf : Annexe B2 figures 2) puis les individus utilisent de moins en moins les transports en commun, après 75 ans les individus semblent redevenir plus enclins à leur utilisation. Enfin la taille du ménage ne semble elle aussi ne pas avoir un effet linéaire puisqu'on remarque que les ménages qui utilisent le moins fréquemment les transports en commun sont ceux de taille moyenne (4 à 5 individus).

Cette recherche a porté uniquement sur les facteurs affectant la fréquence d'utilisation des transports publics en négligeant la concurrence avec d'autres modes. Bien qu'il nous ayons inclus dans notre modèle des facteurs étant en relation avec l'accès à certains modes alternatifs, nous avons complètement ignoré des facteurs comme le choix de destination, l'objectif et la distance à parcourir. Il peut être intéressant de considérer ces variables dans les recherches futures pour permettre une exploration plus approfondie des facteurs d'influences.

Annexes

A Exploration des données

A1 La taille de ménage

Les résultats de la régression qui nous sert pour explorer les données sont présentées ci-dessous. la valeur de référence que nous utilisons est la taille de ménage égale à 1.

Table 9: Modélisation en fonction de l'occupation principale

<i>Dependent variable:</i>	
Utilisation fréquente de TCU	
Taille de ménage : 2	−0.855*** (0.140)
Taille de ménage : 3	−0.675*** (0.158)
Taille de ménage : 4	−1.040*** (0.159)
Taille de ménage : 5	−0.894*** (0.186)
Taille de ménage : 6	−0.529 (0.324)
Taille de ménage : 7	0.361 (0.518)
Taille de ménage : 8+	−12.599 (239.443)
Constante	−0.967*** (0.104)
Observations	3,050
Log Likelihood	−1,314.691
Akaike Inf. Crit.	2,645.381

Note: *p<0.1; **p<0.05; ***p<0.01

A2 L'occupation principale des individus

Table 10: Occupation principale par type du comportement

	Non (N=2560)	Oui (N=490)	Total (N=3050)	p value
Occupation principale				< 0.001
Temps plein	1090 (42.6%)	104 (21.2%)	1194 (39.1%)	
Temps partiel	206 (8.0%)	54 (11.0%)	260 (8.5%)	
Stage	20 (0.8%)	8 (1.6%)	28 (0.9%)	
Etudiant	112 (4.4%)	138 (28.2%)	250 (8.2%)	
Scolaire	432 (16.9%)	96 (19.6%)	528 (17.3%)	
Chomeur	87 (3.4%)	19 (3.9%)	106 (3.5%)	
Retraité	485 (18.9%)	57 (11.6%)	542 (17.8%)	
Reste à foyer	94 (3.7%)	10 (2.0%)	104 (3.4%)	
Autre	34 (1.3%)	4 (0.8%)	38 (1.2%)	

Les résultats de la régression qui nous sert pour explorer les données sont présentées ci-dessous. Nous utilisons la modalité “travail à temps plein” comme valeur de référence.

Table 11: Modélisation, taille de ménage

<i>Dependent variable:</i>	
Utilisation frequente de TCU	
Temps partiel	1.011*** (0.184)
Stage	1.433*** (0.431)
Etudiant	2.558*** (0.163)
Scolaire	0.845*** (0.153)
Chomeur	0.828*** (0.273)
Retraité	0.208 (0.174)
Reste à foyer	0.109 (0.348)
Autre	0.209 (0.538)
Constante	-2.350*** (0.103)
Observations	3,050
Log Likelihood	-1,202.844
Akaike Inf. Crit.	2,423.689

Note: *p<0.1; **p<0.05; ***p<0.01

A3 Nombre de déplacements dans la journée

Dans notre échantillon nous observons, que le nombre de déplacements dans la journée risque de ne pas être linéaire. Tout de même, nous ignorons cette évidence afin de simplifier le modèle.

Table 12: Nombre de déplacements par type du comportement

	Non (N=2560)	Oui (N=490)	Total (N=3050)	p value
Nombre de déplacements				< 0.001
1	8 (0.3%)	0 (0.0%)	8 (0.3%)	
2	518 (20.2%)	140 (28.6%)	658 (21.6%)	
3	293 (11.4%)	46 (9.4%)	339 (11.1%)	
4	673 (26.3%)	153 (31.2%)	826 (27.1%)	
5	328 (12.8%)	57 (11.6%)	385 (12.6%)	
6	315 (12.3%)	48 (9.8%)	363 (11.9%)	
7	189 (7.4%)	20 (4.1%)	209 (6.9%)	
8	121 (4.7%)	15 (3.1%)	136 (4.5%)	
9	70 (2.7%)	8 (1.6%)	78 (2.6%)	
10	45 (1.8%)	3 (0.6%)	48 (1.6%)	

B Modèle complexe

B1 *Summary* complet du modèle complexe

Cette annexe présente les résultats de la régression complexe sous son format complet.

Table 13: Les résultats du modèle avancé

<i>Dependent variable:</i>	
Utilisation frequente de TCU	
N'a pas de permis	0.681*** (0.155)
Taille de ménage	-0.413** (0.163)
Taille de ménage SQ	0.044* (0.024)
Travaille temps plein	-1.036*** (0.145)
Travaille temps partiel	-1.297*** (0.361)
Age	0.292*** (0.034)
Age 2-ème	-0.007*** (0.001)
Age 3-ème	0.00005*** (0.00001)
Sexe	0.354*** (0.113)
Nombre de deplacements	-0.101*** (0.030)
Nombre des voitures	-0.612*** (0.081)
Constant	-2.658*** (0.508)
Observations	3,050
Log Likelihood	-1,119.698
Akaike Inf. Crit.	2,263.396

Note:

*p<0.1; **p<0.05; ***p<0.01

B2 Présentation graphique des effets

La représentation graphique facilite la compréhension des effets marginaux sur la probabilité d'utiliser les transports en commun tous les jours :

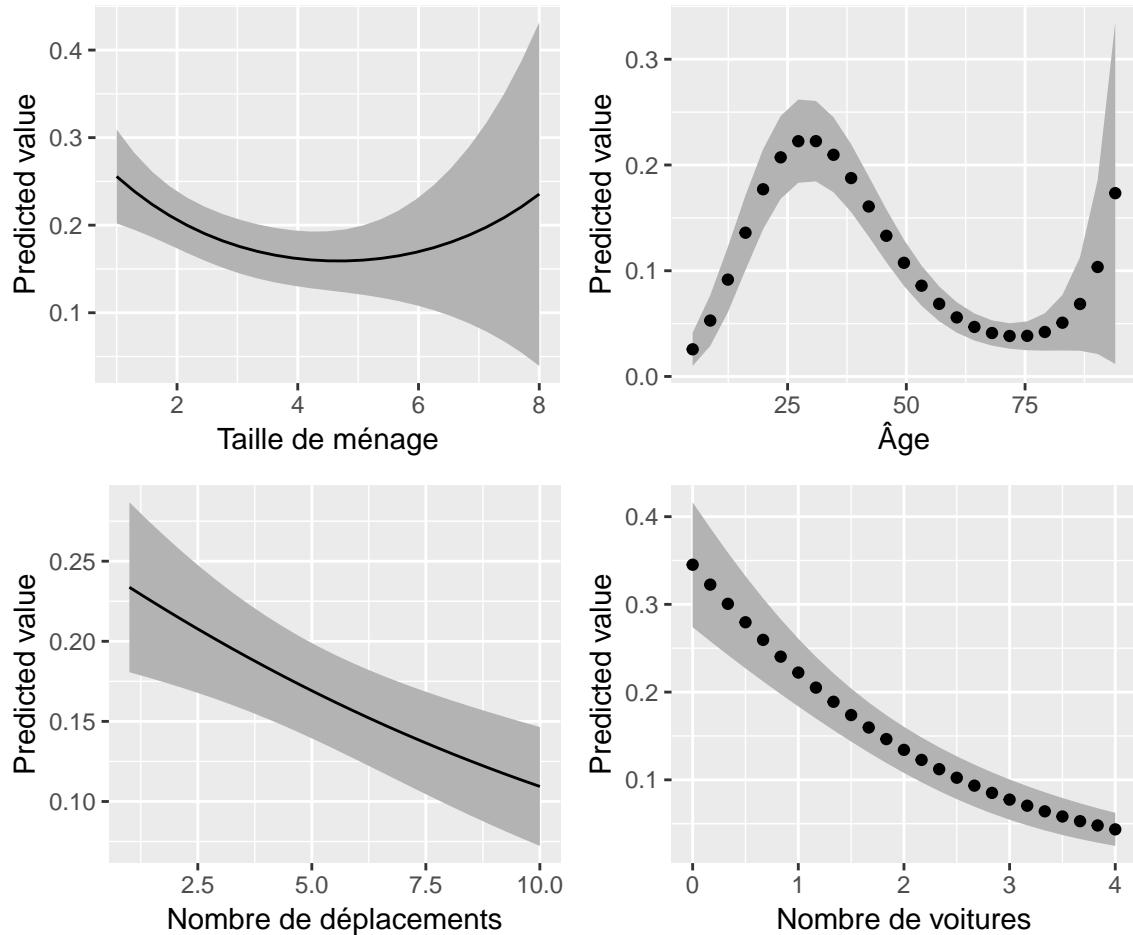


Figure 2: Les effets marginaux du modèle complexe

B3 Tests

VIF-test

Ce test nous permet d'avoir une intuition sur la présence de multicollinearité dans notre modèle. On n'observe les valeurs critiques que pour les variables non-linéaires.

Table 14: VIF test

VIF results	
N'a pas de permis	1.965
Taille de ménage	19.737
Taille de ménage SQ	18.198
Travaille temps plein	1.379
Travaille temps partiel	1.068
Age	163.146
Age 2-ème	696.675
Age 3-ème	226.571
Sexe	1.057
Nombre de déplacements	1.065
Nombre des voitures disponibles	1.418

ROC courbes

En faisant une simple analyse par les courbes ROC nous confirmons notre hypothèse d'une meilleure efficience du modèle complexe.

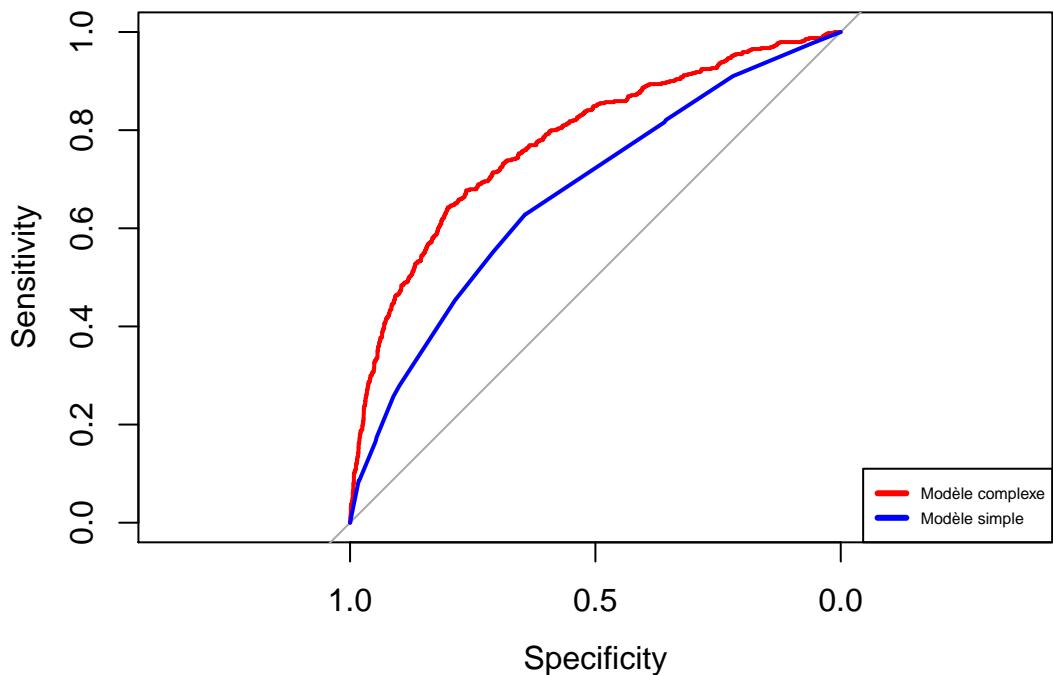


Figure 3: Courbes ROC pour les deux modèles

C Graphical analysis

C1 La variable dépendante

Il est facile de voir que la partie des individus venant en dehors de l'agglomération grenobloise n'est pas équilibré. Pour une des zones nous avons la part d'utilisateurs fréquents du transport en commun qui s'approche à 100%, ce qui n'est pas logique.

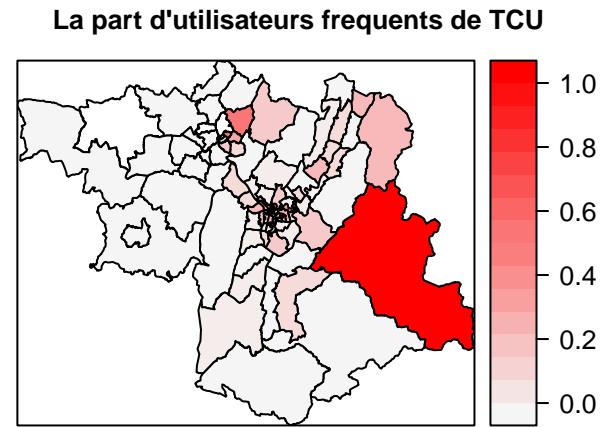


Figure 4: La part d'utilisateurs fréquents de TCU dans région

C'est pourquoi il est plus sensible d'étudier les données sous la représentation graphique par zone seulement pour le centre-ville. Sur le graphique résultant nous observons la présence de corrélation spatiale dans nos données.

Il est évident, que la disponibilité et accessibilité du transport en commun définit en partie les habitudes en terme de choix et de la fréquence d'utilisation de transport commun. C'est pourquoi nous observons une hausse de la part des utilisateurs fréquents du réseau TAG près du centre ville, où la densité et l'accessibilité des transports est plus élevée.

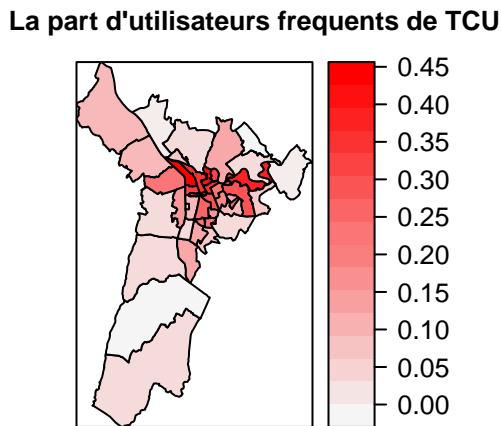


Figure 5: La part d'utilisateurs fréquents de TCU dans l'agglomération grenobloise

C2 Les variables explicatives

Dans cette partie nous présentons la répartition des autres caractéristiques dans les zones géographiques.

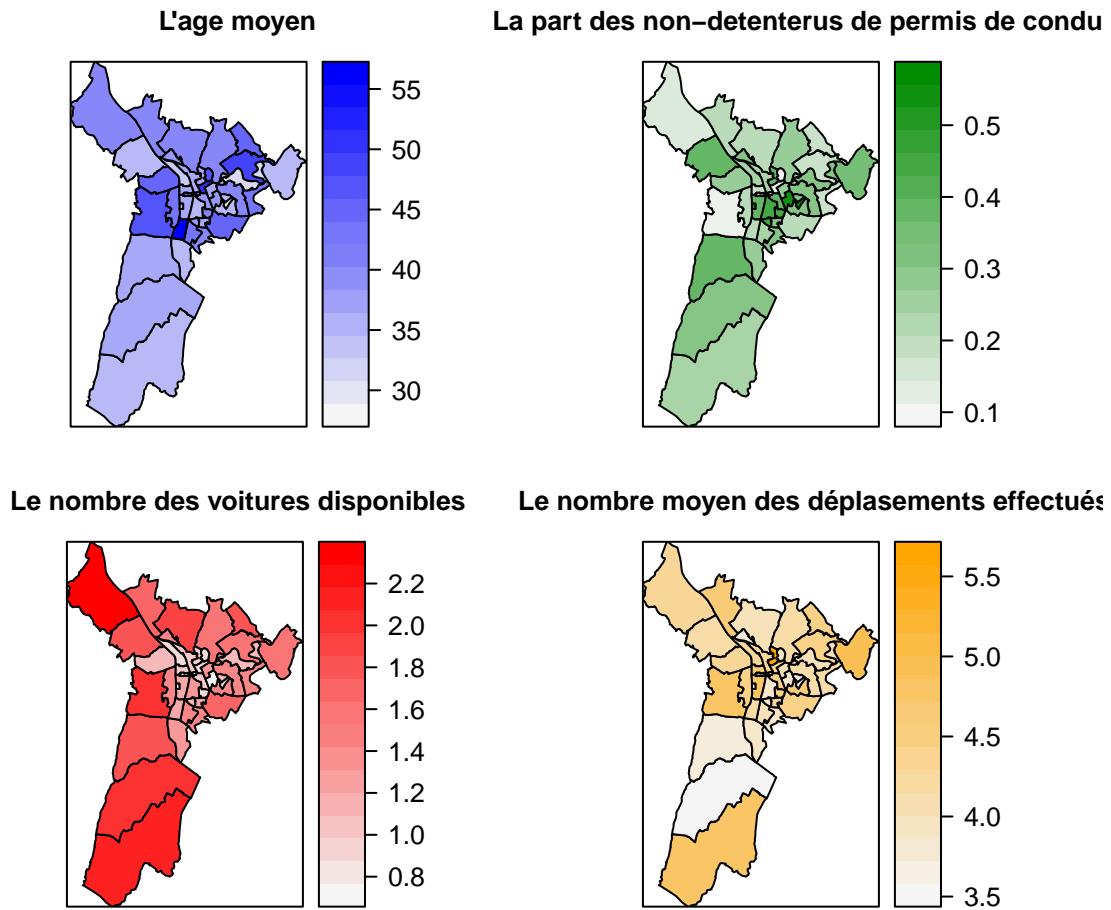


Figure 6: Les statistiques déscriptives par zone

Referéncies

- Corpuz, G. 2007. "Public Transport or Private Vehicle: Factors That Impact on Mode Choice." http://atrf.info/papers/2007/2007_Corpuz.pdf.
- Gärling, Tommy, and Kay W. Axhausen. n.d. "Introduction: Habitual Travel Choice," *Transportation*, 30 (1): 1–11. <https://doi.org/10.1023/A:1021230223001>.
- Muthyalagari, A. Parashar, G.R., and R.M. Pendyala. 2001. "Measuring Day-to-Day Variability in Travel Characteristics Using Gps Data." http://rampendyala.weebly.com/uploads/5/0/5/4/5054275/muthyalagari_parashar_pendyala_rev1_gpsmultidaydata_trb2001.pdf.
- Truong, Long Tien, and Sekhar V.C. Somenahalli. 2015. "Exploring Frequency of Public Transport Use Among Older Adults: A Study in Adelaide, Australia." *Travel Behaviour and Society* 2 (3): 148–55. <https://doi.org/https://doi.org/10.1016/j.tbs.2014.12.004>.