

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Modélisation des systèmes d'activités pour fins de prévision de la demande de
transport**

JULIEN DOUVILLE

Département des génies civil, géologique et des mines

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

Génie civil

Août 2020

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

Modélisation des systèmes d'activités pour fins de prévision de la demande de transport

Présenté par **Julien DOUVILLE**

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

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

Francesco CIARI, président

Catherine MORENCY, membre et directrice de recherche

Louiselle SIOUI, membre

DÉDICACE

À mes parents, qui m'ont tout donné

REMERCIEMENTS

J'aimerais tout d'abord remercier ma directrice de recherche, Catherine Morency, pour son engagement et dévouement inébranlable envers ce projet et ma réussite en général. Son énergie et son enthousiasme sont sans commune mesure et auront été d'une aide inestimable lors de la réalisation d'un projet aussi exigeant.

Bien entendu, je tiens à remercier l'organisme sans qui ce projet ne serait pas possible, le ministère des Transports du Québec.

J'aimerais également remercier l'ensemble des personnes que j'ai côtoyées lors de mon passage à Polytechnique. Merci aux associés de recherche Hubert Verreault, Pierre-Léo Bourbonnais et Jean-Simon Bourdeau pour leur soutien, non seulement lors de la réalisation de ce mémoire, mais également lors des cours qui ont pavé le chemin vers cette expérience unique. Merci aussi au soutien essentiel des étudiants du B-330, qui ont su rendre l'écriture de ce mémoire inoubliable. Merci à Gabriel Lefebvre-Ropars, Élodie Deschaintres, Marc-Antoine Gauthier et Anjeli Narrainen.

Finalement, je souhaite remercier ma famille, dont la fierté envers mes accomplissements est une source de motivation exceptionnelle. Marie-Josée, Gilles et Véronique, merci pour votre soutien inconditionnel. Enfin, je ne remercierai jamais assez ma copine Kim, dont la présence lors de cette maîtrise a été indispensable, grâce à son soutien infaillible.

RÉSUMÉ

Ce mémoire vise à enrichir le processus de prévision de la demande en transport du ministère des Transports du Québec. Tout comme les premiers modèles de prévision apparus dans les années 1950, celui du ministère a comme principal objet de modéliser les déplacements. En revanche, ce qui incite une personne à effectuer un déplacement sont les activités qu'elle veut réaliser à l'extérieur de son domicile. C'est pourquoi la volonté de concevoir des modèles basés sur le désir de réaliser des activités a émergé au courant des années 1970. Toutefois, l'implantation de tels modèles dans la pratique tarde à se matérialiser. L'objectif poursuivi par ce mémoire consiste donc à contribuer à la modernisation des pratiques de planification des transports au Québec. Cette contribution vise à prendre en compte le système d'activités des individus dans le processus de prévision de la demande de transport au Québec en développant une typologie de patrons typiques d'activités. Plus particulièrement, le but ultime de ce projet de recherche est de développer une méthode permettant de générer une séquence d'activités pour chaque individu.

Une revue de littérature permet tout d'abord de cerner les principales limites du modèle que vise à améliorer ce mémoire, soit celui du ministère, mais également celles de la procédure séquentielle classique, dont l'utilisation est largement répandue dans le monde depuis sa création dans les années 1960. La pertinence d'utiliser les modèles basés sur les activités dans l'amélioration de la procédure du MTQ est donc mise en évidence grâce à la mise en lumière de ces limites.

Les deux composants de base des modèles d'activités, qui consistent à générer et programmer les activités, sont par la suite présentés. Les diverses tentatives ayant pour but d'améliorer le module de génération sont également énumérées, en décrivant les multiples essais visant à détecter, au sein de la population, des patrons typiques d'activités. La revue de littérature se conclut par la présentation des principaux liens entre les attributs sociodémographiques des individus et la structure de leurs systèmes d'activité. Les tendances concernant ces liens sont également exposées.

La principale contribution de ce mémoire concerne la création de patrons typiques d'activités. La détection de ces patrons repose sur un partitionnement de la population sur la base de la similarité des systèmes d'activités entre individus. Trois vecteurs sont créés dans le but de rencontrer cet objectif. Selon la nature de chaque vecteur, des algorithmes de segmentation différents sont utilisés afin de regrouper les personnes présentes dans les échantillons de quatre enquêtes Origine-Destination différentes, soit celles de 1998, 2003, 2008 et 2013. Les patrons obtenus à la suite du

partitionnement sont ensuite disséqués, afin de comprendre les particularités de chacun d'entre eux. Cette exploration des patrons permet également d'identifier les forces et les faiblesses de chacun des vecteurs en vue de l'intégration des résultats découlant de ce mémoire au processus de prévision de la demande en transport.

La deuxième grande étape de ce mémoire concerne le développement d'un modèle visant à prédire le patron typique d'activités qu'un individu risque d'adopter. L'interprétation de ce modèle permet en premier lieu de comprendre les liens entre les patrons typiques et les attributs des personnes et des ménages. Les attributs utilisés sont l'âge et le genre au niveau individuel, de même que le type de ménage et la région du domicile pour la seconde entité. De plus, le modèle contribue également à alimenter le processus de génération des activités développé dans le cadre de ce mémoire. En effet, son utilisation permet d'assigner un patron à chaque individu.

La mise en commun de quatre enquêtes permet aussi d'analyser et modéliser l'évolution de la structure des systèmes d'activités des citoyens du Grand Montréal. Cette analyse permet d'évaluer la pertinence des patrons identifiés face à leur utilisation dans un contexte prévisionnel. À la suite de cette exploration, un exercice de modélisation est complété afin de quantifier ces tendances et les projeter dans le futur. Cet exercice consiste essentiellement à altérer le modèle créé à l'étape précédente.

Finalement, les constats ressortant de l'analyse temporelle confirment l'existence de tendances en lien avec l'évolution des systèmes d'activités et de la population en général. Tout d'abord, le vieillissement de la population est bien présent alors que le nombre d'étudiants est en baisse, à l'inverse des retraités. Ces derniers sont également de moins en moins mobiles qu'auparavant, ce qui est confirmé par l'analyse des patrons. De plus, les jeunes adultes présents dans les familles sont de plus en plus nombreux à réaliser des activités de type « autres » avant et après le travail. Finalement, le nombre de personnes quittant leur domicile le matin de façon hâtive est en hausse, comme le confirme l'exploration de l'évolution temporelle des patrons.

La principale limite de ce mémoire concerne l'allocation des patrons d'activités aux individus. En effet, le modèle permettant d'y parvenir n'a pas un taux de réussite élevé. Afin de pallier cette limite, la principale piste de solution consisterait à construire et assigner des semaines typiques d'activité aux individus, composé des patrons typiques journalier.

ABSTRACT

The purpose of this research project is to contribute to the enhancement of the ministère des Transports du Québec's transportation demand forecasting process. Like the first forecasting models that appeared in the 1950s, this model is based on trips. However, what drives people to make these trips are the activities they want to carry out outside their homes. For this reason, the desire to develop models based on the desire to carry out these activities emerged in the 1970s. However, the implementation of such models in practice has been slow to materialize. The goal of this project is therefore to contribute to the modernization of transportation planning practices in Quebec. This contribution aims to consider the activity system of individuals in the MTQ's process by developing a typology of typical activity patterns. More specifically, the goal of this research project is to develop a method for generating a sequence of activities for each individual.

A review of the literature identifies the main limitations of the model that aims this project aims to improve, namely the one of the MTQ, but also the four-stage models, which have been widely used around the world since their creation in the 1960s. The relevance of using activity-based models to improve the MTQ's procedure is thus highlighted by the presentation of these limitations.

The two basic components of activity-based models, which consist of generating and programming activities, are then presented. The various attempts to improve the generation component are also listed, describing the multiple attempts to detect typical activity patterns in the population. The literature review concludes with a presentation of the main links between the socio-demographic attributes of individuals and the structure of people's activity systems. Trends in these relationships are also presented.

The main contribution of this paper is the creation of typical activity patterns. The detection of these patterns is based on the segmentation of the population that rely on the similarity of the activity systems of individuals. Three vectors are created to meet this objective. Depending on the nature of each vector, different segmentation algorithms are used to group the individuals present in the samples of four different Origin-Destination surveys, i.e. the 1998, 2003, 2008 and 2013 surveys. The patterns obtained from the partitioning are then dissected to understand the particularities of each one of them. This exploration of the patterns makes it possible to identify the strengths and weaknesses of each of the vectors with the goal to integrate the results of this project into the MTQ's demand forecasting process.

The second major step in this project is the development of a model to predict the typical pattern that an individual is likely to adopt. The interpretation of this model first allows us to understand the links between typical patterns and the attributes of individuals and households. The attributes used are the age and gender at the individual level, as well as the household type and region of residence for the second entity. In addition, the model also helps to feed the activity generation process developed in this project. Indeed, its use allows the assignment of a pattern to each individual.

The pooling of four surveys makes it possible to analyze and model the evolution of the structure of the activity systems of the citizens of the Greater Montreal Area. This analysis makes it possible to evaluate the relevance of the patterns identified with respect to their use in a forecasting context. Following this exploration, a modelling exercise is completed to quantify these trends and project them into the future. This exercise essentially consists of altering the model created in the previous step.

Finally, the findings of the temporal analysis confirm the existence of trends related to the evolution of activity systems and the population in general. First, the aging of the population is very present, while the number of students is declining, unlike the number of retirees. The retirees are also less mobile than before, which is confirmed by the analysis of patterns. In addition, more and more young adults in families are engaged in "other" activities before and after work. Finally, the number of people leaving home early is on the rise, as confirmed by the exploration of the temporal evolution of patterns.

The main limitation of this project concerns the allocation of activity patterns to individuals. Indeed, the model allowing the allocation does not have a high success rate. In order to overcome this limitation, the main solution would be to construct and assign typical weeks of activity to individuals, composed of typical daily patterns.

TABLE DES MATIÈRES

DÉDICACE.....	III
REMERCIEMENTS	IV
RÉSUMÉ.....	V
ABSTRACT	VII
TABLE DES MATIÈRES	IX
LISTE DES TABLEAUX.....	XIII
LISTE DES FIGURES	XVIII
LISTE DES SIGLES ET ABRÉVIATIONS	XXI
LISTE DES ANNEXES	XXII
CHAPITRE 1 INTRODUCTION.....	1
1.1 Mise en contexte et problématique.....	2
1.1.1 Les limites de la PSC et l'arrivé de nouveaux modèles	2
1.1.2 L'émergence des modèles d'activités	3
1.1.3 L'intégration à l'environnement de modélisation québécois	4
1.2 Objectifs	5
1.3 Structure du mémoire	6
CHAPITRE 2 REVUE DE LITTÉRATURE	7
2.1 Procédure séquentielle classique.....	9
2.1.1 Génération	9
2.1.2 Distribution.....	10
2.1.3 Choix modal	10
2.1.4 Affectation.....	11
2.1.5 Limites de la PSC.....	11

2.2	Approche utilisée au MTQ.....	12
2.2.1	L'enquête Origine-Destination.....	14
2.2.2	Projection de la population.....	14
2.2.3	Projection des déplacements	15
2.2.4	Transfert modal	17
2.2.5	Affectation des déplacements.....	17
2.2.6	Limite de la procédure du MTQ.....	18
2.3	Les modèles d'activités	18
2.3.1	La structure des modèles d'activités	21
2.3.2	La génération des activités	24
2.3.3	Les individus, les ménages et les systèmes d'activités	27
2.3.4	L'évolution des systèmes d'activités.....	29
CHAPITRE 3 MÉTHODOLOGIE.....		31
3.1	Méthodologie générale	31
3.2	Description des données.....	34
3.2.1	L'enquête Origine-Destination.....	34
3.2.2	Traitement et transformation des données	40
3.2.3	Typologie des ménages	42
3.3	Segmentation de la population	44
3.3.1	Création de trois types de vecteurs.....	44
3.3.2	Segmentation avec l'algorithme de K-moyennes.....	46
3.3.3	Segmentation hiérarchique	48
3.4	Distributions des attributs des activités.....	50
3.4.1	Comparaison des distributions	51

3.4.2	Utilisation des distributions afin de générer les activités	53
3.5	Modèle d'allocation des personnes à un patron d'activité	57
CHAPITRE 4 CRÉATION DE TROIS TYPOLOGIES DE PATRONS D'ACTIVITÉS INDIVIDUELLES		59
4.1	Analyse des systèmes d'activités	59
4.2	Première segmentation	61
4.2.1	Choix du nombre de groupes	61
4.2.2	Analyse des groupes.....	63
4.2.3	Analyse des distributions	71
4.3	Deuxième segmentation	73
4.3.1	Analyse des groupes.....	75
4.3.2	Analyse des distributions	81
4.4	Troisième segmentation	83
4.4.1	Analyse des groupes.....	83
4.4.2	Analyse des distributions	88
CHAPITRE 5 MODÈLE D'ALLOCATION DES PATRONS D'ACTIVITÉS AUX INDIVIDUS		91
5.1	Description sociodémographique des groupes.....	91
5.1.1	Première segmentation (V1).....	91
5.1.2	Deuxième segmentation (V2).....	94
5.1.3	Deuxième segmentation (V3).....	96
5.2	Description et transformation des variables	98
5.2.1	Analyse de corrélation.....	103
5.2.2	Analyse des variables choisies	106
5.3	Première segmentation	107

5.3.1	Analyses des modèles.....	107
5.3.2	Validation des modèles	125
5.4	Deuxième segmentation	128
5.4.1	Analyse des modèles	128
5.4.2	Validation des modèles	132
5.5	Troisième segmentation	133
5.5.1	Analyse des modèles	134
5.5.2	Validation des modèles	136
5.6	Comparaison des méthodes.....	137
5.6.1	Sélection des personnes pour fin de comparaison.....	137
5.6.2	Analyse des activités générées	139
CHAPITRE 6	ANALYSE DES TRANSFORMATIONS DES SYSTÈMES D'ACTIVITÉS	141
6.1	Évolution temporelle des patrons.....	141
6.1.1	Première segmentation	141
6.1.2	Deuxième segmentation	147
6.1.3	Troisième segmentation	148
6.2	Modélisation de l'évolution de l'appartenance aux groupes.....	150
CHAPITRE 7	CONCLUSION	153
7.1	Synthèse	153
7.2	Limites.....	156
7.3	Perspectives.....	158
RÉFÉRENCES	160
ANNEXES	165

LISTE DES TABLEAUX

Tableau 3-1 Caractéristiques des enquêtes de 1998 à 2013	36
Tableau 3-2 Données utilisées dans le cadre de ce mémoire	38
Tableau 3-3 Liste des motifs et des types d'activités.....	41
Tableau 3-4 Caractéristique des huit types de ménages.....	43
Tableau 3-5 Proportion de ménages pour chaque catégorie	43
Tableau 3-6 Exemple de chacun des vecteurs avant normalisation	44
Tableau 4-1 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité	72
Tableau 4-2 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité	73
Tableau 4-3 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité – E1, T1 et T2.....	81
Tableau 4-4 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité – T2, T3, T4, T5, L1, L2 et A1.....	82
Tableau 4-5 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité – E1 et T1	88
Tableau 4-6 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité T1, T2, T3, T4, L1, L2, M1, A1 et MA1.....	89
Tableau 5-1 Caractéristiques sociodémographiques des individus de chacun des groupes.....	92
Tableau 5-2 Caractéristiques des ménages auxquels chaque individu appartient.....	93
Tableau 5-3 Caractéristiques sociodémographiques des individus de chacun des groupes.....	94
Tableau 5-4 Caractéristiques des ménages auxquels chaque individu appartient.....	95
Tableau 5-5 Caractéristiques sociodémographiques des individus de chacun des groupes.....	96
Tableau 5-6 Caractéristiques des ménages auxquels chaque individu appartient.....	97
Tableau 5-7 Variables utilisées pour l'estimation du modèle	99

Tableau 5-8 Description des variables explicatives des modèles 1 et 2.....	100
Tableau 5-9 Description des variables explicatives des modèles 3 et 4.....	101
Tableau 5-10 Description des variables explicatives du modèle 5	102
Tableau 5-11 Analyse de corrélation du modèle 1	103
Tableau 5-12 Analyse de corrélation du modèle 2.....	104
Tableau 5-13 Analyse de corrélation du modèle 4.....	105
Tableau 5-14 Analyse des variables explicatives du modèle 2	106
Tableau 5-15 Rapport de probabilités du modèle 1 – Activités <i>études et travail</i>	108
Tableau 5-16 Rapport de probabilités du modèle 1 – Activités <i>loisir, magasinage et autres</i>	109
Tableau 5-17 Rapport de probabilités du modèle 2 – Activités <i>études et travail</i>	110
Tableau 5-18 Rapport de probabilités du modèle 2 – Activités <i>loisir, magasinage et autres</i>	111
Tableau 5-19 Rapport de probabilités du modèle 3 – <i>Études et travail</i> – Variable âge et constante	113
Tableau 5-20 Rapport de probabilités du modèle 3 – <i>Études et travail</i> – Variable ménage.....	114
Tableau 5-21 Rapport de probabilités du modèle 3 – <i>Loisir, magasinage et autres</i> – Var. ménage	114
Tableau 5-22 Rapport de probabilités du modèle 3 – <i>Loisir, magasinage et autres</i> – Variable âge	115
Tableau 5-23 Rapport de probabilités du modèle 4 – <i>Études et travail</i>	116
Tableau 5-24 Rapport de probabilités du modèle 4 – <i>Loisir, magasinage et autres</i>	117
Tableau 5-25 Rapport de probabilités du modèle 5 – <i>Études et travail</i>	119
Tableau 5-26 Rapport de probabilités du modèle 5 – <i>Loisir, magasinage et autres</i>	120
Tableau 5-27 Extrait de la base de données des personnes permettant d’identifier la vraisemblance de chaque observation (modèle 3).....	121

Tableau 5-28 Extrait de la base de données des personnes permettant d'identifier la probabilité maximale de chaque observation (modèle 3).....	122
Tableau 5-29 Proportion d'individus selon le groupe correspondant à la probabilité maximale (modèle 3)	123
Tableau 5-30 AIC des 5 modèles	124
Tableau 5-31 Taux de réussite de chacun des 5 modèles.....	125
Tableau 5-32 Proportion d'individus observée et prédite par le modèle 3 dans chaque groupe..	126
Tableau 5-33 Matrice de confusion du modèle 3 (V1)	127
Tableau 5-34 Rapport de probabilités du modèle 2 – <i>Études et travail</i>	129
Tableau 5-35 Rapport de probabilités du modèle 2 – <i>Loisir, magasinage et autres</i>	131
Tableau 5-36 AIC des 5 modèles	132
Tableau 5-37 Taux de réussite de chacun des 5 modèles.....	132
Tableau 5-38 Matrice de confusion du modèle 2 (V2)	133
Tableau 5-39 Rapport de probabilités du modèle 2 – <i>Études et travail</i>	134
Tableau 5-40 Rapport de probabilités du modèle 2 – <i>Loisir, magasinage et autres</i>	135
Tableau 5-41 AIC des 5 modèles	136
Tableau 5-42 Taux de réussite de chacun des 5 modèles.....	136
Tableau 5-43 Matrice de confusion du modèle 2 (V3)	137
Tableau 5-44 Proportion d'activités générées parmi celles observées.....	139
Tableau 5-45 Proportion d'activités générées par rapport à celles observées par segment de population et par type d'activité – Segmentation V1 – Modèle 3.....	140
Tableau 6-1 Résultats du modèle 2 – variable « année »	151
Tableau A-1 V1 – Heure de début – E1, E1, T1, T2 et T3	165
Tableau A-2 V1 – Heure de début T4, L2, L1 et M1	166
Tableau A-3 V1 – Durée E1 et E2	166

Tableau A-4 V1 – Durée T1, T2, T3, T4, L1, L2 et M1	167
Tableau A-5 V2 – Heure de début E1, T1, T2, T3 et T4	168
Tableau A-6 V2 – Heure de début T5, L1, L2 et A1	169
Tableau A-7 V2 – Durée E1 et T1	169
Tableau A-8 V2 – Durée T2, T3, T4, T5, L1, L2 et A1	170
Tableau A-9 V3 – Heure de début E1, T1, T2, T3 et T4	171
Tableau A-10 V3 – Heure de début T5, L1, A1 et MA1	172
Tableau A-11 V3 – Durée – E1 et T1	172
Tableau A-12 V3 – Durée T2, T3, T4, T5, L1, A1 et MA1	173
Tableau B-13 Matrice de corrélation du modèle 3 – Variable ménage	174
Tableau B-14 Matrice de corrélation du modèle 3 – Variable d’interaction – 54 ans et moins ..	175
Tableau B-15 Matrice de corrélation du modèle 3 – Variable d’interaction – 55 ans et plus.....	176
Tableau B-16 Matrice de corrélation du modèle 5 – Âge 5-14 et 45-64	177
Tableau B-17 Matrice de corrélation du modèle 5 – Âge 15-24, 25-44 et 65+	178
Tableau C-18 V2 – Modèle 1 – <i>Études et travail</i>	179
Tableau C-19 V2 – Modèle 1 – <i>Loisir, magasinage et autres</i>	179
Tableau C-20 V2 – Modèle 3 – <i>Études et travail</i> – Variable âge	180
Tableau C-21 V2 – Modèle 3 – <i>Études et travail</i> – Variable type de ménage.....	181
Tableau C-22 V2 – Modèle 3 – <i>Loisir, magasinage et autres</i> – Variable type de ménage	181
Tableau C-23 V2 – Modèle 3 – <i>Loisir, magasinage et autres</i> – Variable âge.....	182
Tableau C-24 V2 – Modèle 4 – <i>Études et travail</i>	183
Tableau C-25 V2 – Modèle 4 – <i>Loisir, magasinage et autres</i>	184
Tableau C-26 V2 – Modèle 5 – <i>Études et travail</i>	185
Tableau C-27 V2 – Modèle 5 – <i>Loisir, magasinage et autres</i>	186

Tableau C-28 V3 – Modèle 1 – <i>Études et travail</i>	187
Tableau C-29 V3 – Modèle 1 – <i>Loisir, magasinage et autres</i>	187
Tableau C-30 V3 – Modèle 3 – <i>Études et travail</i> – Variable âge	188
Tableau C-31 V3 – Modèle 3 – <i>Études et travail</i> – Variable type de ménage.....	189
Tableau C-32 V3 – Modèle 3 – <i>Loisir, magasinage et autres</i> – Variable type de ménage	189
Tableau C-33 V3 – Modèle 3 – <i>Loisir, magasinage et autres</i> – Variable âge.....	190
Tableau C-34 V3 – Modèle 4 – <i>Études et travail</i>	191
Tableau C-35 V3 – Modèle 4 – <i>Loisir, magasinage et autres</i>	192
Tableau C-36 V3 – Modèle 5 – <i>Études et travail</i>	193
Tableau C-37 V3 – Modèle 5 – <i>Loisir, magasinage et autres</i>	194
Tableau D-38 Nombre d'activités observées et générées et ratio entre ces nombres avec V1 – Modèle 3.....	195
Tableau E-39 Rapport de probabilités – <i>Études et travail</i>	196
Tableau E-40 Rapport de probabilités – <i>Loisir, magasinage et autres</i>	197

LISTE DES FIGURES

Figure 2-1 Évolution des modèles de prévision de la demande en transport	8
Figure 2-2 Les grandes étapes de la procédure séquentielle classique	9
Figure 2-3 Schéma général du MOTREM (MTQ, 2017b).....	13
Figure 2-4 Trajectoire d'un individu dans l'espace-temps.....	19
Figure 3-1 Schéma méthodologique	32
Figure 3-2 Le territoire de l'enquête de 2013 et les huit régions qui le composent.....	35
Figure 3-3 Structure des données de l'enquête OD.....	37
Figure 3-4 Transformation des données des déplacements.....	40
Figure 3-5 Typologie des ménages	42
Figure 3-6 Opérations possibles pour le calcul de la distance de Damerau-Levenshtein	49
Figure 3-7 Exemple de distribution à des fins de comparaison	51
Figure 3-8 Identification de la distance maximale entre deux distributions	52
Figure 3-9 Exemple de l'utilisation de la méthode de la transformée inverse	54
Figure 3-10 Distribution du pourcentage de personnes selon les attributs des activités permettant la génération des activités	55
Figure 4-1 Distribution du pourcentage d'activités selon le type d'activité et selon l'âge de la personne qui a réalisé l'activité.....	60
Figure 4-2 Distribution du pourcentage d'activités selon le nombre d'activités	60
Figure 4-3 Durée et fréquence moyenne de chaque type d'activité selon l'âge.....	61
Figure 4-4 Inertie expliquée en fonction du nombre de groupes	62
Figure 4-5 Distribution de l'échantillon selon le groupe	63
Figure 4-6 Distribution du pourcentage de personnes dans chacun des groupes.....	65
Figure 4-7 Centres des groupes de la segmentation avec V1.....	66

Figure 4-8 Proportion d'individus, à l'intérieur de chaque groupe, réalisant un des cinq types d'activités selon l'heure de la journée.....	68
Figure 4-9 Proportion d'individus selon l'heure de début de l'activité dominante	70
Figure 4-10 Proportion de personnes dans chacun des groupes selon le jour de la semaine	71
Figure 4-11 Distribution du pourcentage de personnes à la suite de la création de trois groupes .	74
Figure 4-12 Distribution du pourcentage de personnes avec le choix final du nombre de groupes	76
Figure 4-13 Centres des groupes de la segmentation avec V2.....	77
Figure 4-14 Proportion d'individus, à l'intérieur de chaque groupe, réalisant un des cinq types d'activités selon l'heure de la journée.....	79
Figure 4-15 Proportion d'individus selon l'heure de départ de l'activité dominante	80
Figure 4-16 Distribution de l'échantillon selon le groupe	83
Figure 4-17 Centre des groupes de la segmentation avec V3	85
Figure 4-18 Proportion d'individus, à l'intérieur de chaque groupe, réalisant un des cinq types d'activités selon l'heure de la journée.....	86
Figure 4-19 Heure de départ de l'activité dominante.....	87
Figure 5-1 Vraisemblance de chaque observation selon le groupe observé (modèle 3).....	122
Figure 5-2 Distribution de la proportion d'individus selon la probabilité maximale (modèle 3)	123
Figure 6-1 Évolution des proportions dans chacun des groupes avec la segmentation 1	142
Figure 6-2 Évolution de la proportion de personnes dans V1-E1 et V1-E2 pour chaque segment de la population.....	144
Figure 6-3 Évolution de la proportion de personnes dans V1-T1, V1-T2, V1-T3 et V1-T4	145
Figure 6-4 Évolution de la proportion de personnes dans V1-L1, V1-L2, V1-M1 et V1-A1.....	146
Figure 6-5 Évolution temporelle des proportions d'individus dans chacun des groupes.....	147
Figure 6-6 Évolution temporelle des proportions d'individus dans chacun des groupes.....	149

Figure 6-7 Proportion de personnes dans chacun des groupes d'âge parmi les membres du groupe V3-T3	150
Figure 6-8 Évolution de la probabilité d'appartenir aux groupes dont le <i>travail</i> est l'activité dominante pour les hommes âgés entre 30 et 34 ans.....	152

LISTE DES SIGLES ET ABRÉVIATIONS

AC	Automobile conducteur
AIC	<i>Akaike Information Criterion</i>
AP	Automobile passager
ARTM	Autorité régionale de transport métropolitain
DMST	Direction de la modélisation des systèmes de transport
ES-3	Entrant-sortant à trois niveaux
Freq.	Fréquence
GRM	Grande région de Montréal
GTFS	General transit feeds specification
KS	Kolmogorov-Smirnov
MADITUC	Modèle d'Analyse Désagrégée des Itinéraires de Transport Urbain Collectif
MADEDE	Méthode de projection désagrégée de la demande
MOTREM	Modèle de transport de la région de Montréal
MTQ	Ministère des Transports du Québec
OD	Origine-Destination
PSC	Procédure séquentielle classique
SAM	Sequence Alignment Method
STM	Société de transport de Montréal
TC	Transport en commun
TASHA	Travel Activity Scheduler for Household Agents
Var.	Variable

LISTE DES ANNEXES

Annexe A	Résultat des tests de Kolmogorov-Smirnov.....	165
Annexe B	Analyse de corrélation des modèles.....	174
Annexe C	Rapports de probabilité des modèles 1 à 5 pour V2 et V3.....	179
Annexe D	Comparaison du nombre d'activités observées et générées avec V1	195
Annexe E	Rapports de probabilité du modèle 6 – V1.....	196

CHAPITRE 1 INTRODUCTION

L'histoire récente des villes est marquée par une augmentation considérable de la population citadine et de la mobilité des biens et personnes. En effet, depuis l'industrialisation de notre société, le nombre de personnes vivant en milieu urbain a explosé, entraînant une croissance significative du nombre de déplacements et une expansion sans commune mesure de l'aire urbanisée. Différents modes de transport ont dominé de façon successive le paysage urbain en structurant le développement des villes. Alors que l'espace urbain s'est principalement développé autour du tramway électrique à la fin du XIXe siècle et au début du XXe siècle, c'est l'automobile qui a joué ce rôle par la suite (Newman & Kenworthy, 1996). Elle apparaît au tout début du XXe siècle, mais c'est à partir des années 1930 qu'elle se démocratise pour ensuite devenir le mode dominant autour duquel l'espace urbain va s'articuler.

Ce nouveau mode de déplacement transforme de façon radicale la relation que les individus entretiennent avec le territoire, en étant omniprésent dans l'espace urbain. L'automobile permet de parcourir des distances de plus en plus longues et ainsi couvrir un territoire toujours plus vaste nécessaire à la réalisation des activités hors du domicile. Cette augmentation des distances est indissociable de la mise en place d'infrastructures lourdes de transport. En effet, durant les décennies qui suivent la démocratisation du véhicule privé, les révolutions technologiques propres aux moyens de transports et de communications ainsi que la transformation de l'économie en une économie d'après-guerre mène à une requalification de la planification urbaine. La volonté d'implanter des infrastructures permettant d'accommoder un flux grandissant d'automobiles est au cœur de cette requalification, qui est également marquée par une approche différente de la planification. La modernisation de la société provoque en effet l'apparition de nouveaux outils et de nouvelles techniques permettant de planifier et prévoir la construction d'infrastructures. Alors que l'accent est mis sur les projets routiers, et dans une moindre mesure sur le transport collectif, la naissance d'outils permettant de mesurer, quantifier, modéliser et prévoir les flux de circulation automobile devient inévitable.

C'est dans les années 1950 que les tout premiers modèles de prévision de la demande en transports prennent forme (Bonnel, 2004). Alors que ces modèles se concentrent principalement sur la modélisation des flux automobiles, une volonté naît de comprendre davantage les effets de l'utilisation du sol et de la distribution de la population sur l'utilisation du réseau routier qui s'étend

sur le territoire grandissant des villes. Combiné à un intérêt marqué pour le développement régional des transports et des investissements massifs dans des projets d'infrastructures lourdes, c'est dans ce contexte que la procédure séquentielle classique (PSC) apparait entre 1960 et 1965 (Masson, 1998). La création de ce modèle est également accompagnée de l'apparition des premières enquêtes Origine-Destination afin de soutenir cette démarche, qui a pour principal but de comprendre davantage les liens entre les caractéristiques de la population et l'utilisation des réseaux de transport pour ainsi améliorer la fluidité de la circulation automobile.

1.1 Mise en contexte et problématique

Les contrecoups des décisions prises durant les années 1950 et 1960 se font sentir à partir de la décennie suivante. Les politiques visant à favoriser le développement de l'automobile provoquent l'émergence de tendances lourdes en ce qui a trait aux caractéristiques des systèmes de transport urbain. Les ménages connaissent une motorisation sans précédent, les distances parcourues en automobile augmentent de façon significative, l'utilisation des transports collectifs baisse et le rapport au territoire change complètement, alors que l'étalement urbain devient problématique et que la distribution des lieux d'emploi est caractérisée par une forte dispersion (Newman & Kenworthy, 1996). Au moment où ces tendances surgissent, plusieurs voix s'élèvent pour critiquer la PSC, qui n'a pas été conçue pour évaluer les solutions qui permettraient de faire face à ces enjeux.

1.1.1 Les limites de la PSC et l'arrivée de nouveaux modèles

Les limites de la procédure séquentielle classique sont nombreuses. De manière générale, elle n'a pas été conçue dans une logique de modéliser le comportement des personnes et de simuler les décisions prises au niveau individuel, mais plutôt dans le but d'accompagner les investissements routiers (Masson, 1998). Alors que la volonté de réduire la congestion et d'intégrer le développement durable dans la planification des villes commence à émerger, de nouveaux modèles apparaissent dans le but d'évaluer les politiques publiques visant à atteindre ces objectifs (McNally Michael & Rindt Craig, 2007). Ces modèles tentent de pallier les principales limites de la PSC, à savoir la nature agrégative de la méthode ainsi que le fait que le déplacement soit l'objet principal de modélisation.

L'effort est donc tout d'abord porté sur la désagrégation de la méthode, mais également sur la nature statique de la procédure qui n'est pas en mesure d'appréhender l'évolution des comportements de mobilité à long terme. En effet, la méthode n'a pas la capacité d'intégrer les multiples phénomènes qui engendrent des changements importants dans les comportements de mobilité des personnes sur différents horizons temporels. La société connaît notamment un vieillissement de la population considérable et une évolution drastique du statut de la femme, qui l'amène à intégrer davantage le marché du travail et à occuper un nouveau rôle au sein des ménages. Les villes subissent également des phénomènes importants, tels que la dépendance à l'automobile ou la reconfiguration spatiale des lieux d'activités. Les nouveaux modèles proposés tentent de capturer ces phénomènes en modélisant les comportements de mobilité au niveau des individus. Mais avec la complexification de la structure des déplacements découlant de ces phénomènes, le cœur du problème demeure la nature de l'objet de modélisation qui est le déplacement. En effet, les modèles dont le déplacement est au centre du processus prévisionnel ne permettent pas de considérer les multiples liens entre les différents déplacements réalisés au cours d'une journée ou d'une semaine ainsi que les contraintes qui régissent le choix de ces déplacements au niveau individuel (Roorda, 2005).

1.1.2 L'émergence des modèles d'activités

C'est dans ce contexte que le concept des modèles basés sur les activités émerge au courant des années 1970. Le fait que les individus ne se déplacent pas uniquement pour le plaisir, mais plutôt dans le but de réaliser des activités en dehors du domicile est l'essentiel des motivations derrière la naissance de ce concept. Bien que les modèles d'activités soient présents dans le monde de la recherche depuis de nombreuses décennies, ce n'est que récemment que la volonté de les intégrer dans la pratique est apparue. En effet, la nécessité de développer des modèles plus sensibles à l'évolution des comportements de mobilité est devenue indispensable. Le fait de modéliser les activités permet de mieux intégrer les changements sociodémographiques et les effets en découlant pour ainsi mieux évaluer les politiques publiques visant à résoudre les problèmes qui sévissent dans les régions urbaines.

De manière générale, les modèles d'activité sont divisés en deux composants, soit la génération des activités et la programmation de celles-ci. Alors que la recherche sur le deuxième composant est très active depuis les dernières années, celle sur la génération des activités ne l'est pas autant.

Dans plusieurs modèles, la génération nécessite une segmentation de la population dans le but de mieux prédire les activités auxquelles les individus risquent de s'adonner. Dans certains cas, cette segmentation n'est pas le résultat d'une méthode dont la prise en compte des caractéristiques des systèmes d'activités est au cœur de la démarche. Autrement dit, le regroupement des individus n'a pas comme assise la similarité des systèmes d'activités.

Par conséquent, ce mémoire vise à améliorer le processus de génération d'activités individuelles en proposant une méthode qui tient compte de la structure des systèmes d'activités des individus. L'élaboration d'une telle méthode nécessite également la prise en compte des caractéristiques sociodémographiques des personnes et des ménages dans le processus de génération. Alors qu'il est évident que les attributs définissant les personnes et les ménages ont une influence sur les activités réalisées au cours d'une journée ou d'une semaine, il devient nécessaire de comprendre les liens entre ces attributs et la structure des systèmes d'activités. Il s'agit donc d'analyser et de caractériser ces systèmes dans le but de comprendre les liens qui les unissent avec les personnes qui les performant.

De plus, le processus de génération des activités de certains modèles n'a pas de composante évolutive. Autrement dit, les méthodes permettant de générer des activités à divers horizons temporels négligent le fait que la structure du système d'activité des individus puisse évoluer dans le temps pour différents segments de la population.

1.1.3 L'intégration à l'environnement de modélisation québécois

Au Québec, la responsabilité de prévoir de la demande en transport revient à la Direction de la modélisation des systèmes de transport (DMST) du ministère des Transports du Québec (MTQ). La méthode utilisée au MTQ est ancrée dans les données disponibles dans les enquêtes Origine-Destination, dont la première a été réalisée en 1970 dans la Grande région de Montréal (GRM) et qui a lieu à chaque cinq ans environ. Depuis quelques décennies, l'enquête est également réalisée dans d'autres villes de la province. La réalisation de ces enquêtes permet aux différentes régions urbaines de la province d'avoir accès à des données détaillées et représentatives de l'ensemble de la population. L'envergure de ces enquêtes, particulièrement dans le cas de Montréal, est pratiquement unique au monde. Toutefois, malgré la richesse des données provenant de l'enquête, la méthode prévisionnelle du MTQ n'intègre pas les pratiques les plus récentes en matière de

prévision de la demande en transport. En effet, les déplacements, et non les activités, sont au cœur de la démarche de prévision.

C'est pourquoi la DMST souhaite enrichir son modèle afin de pallier certaines lacunes. Le principal objectif est d'améliorer le processus afin qu'il soit en mesure de mieux évaluer l'impact de nouveaux scénarios d'offre de transport de même que mieux appréhender les interactions entre la demande en transport et l'aménagement du territoire. Afin d'y parvenir, le projet de recherche visant à remplir cet objectif se divise en trois volets, soit la synthèse de la population, ainsi que la prise en compte des caractéristiques des ménages et des systèmes d'activités. La synthèse de la population consiste à générer une population fictive pour l'ensemble de la région métropolitaine dont le poids de chaque individu est de 1. Le deuxième volet a pour but de créer une typologie des ménages pertinente dans le but d'intégrer l'impact que peut avoir la composition d'un ménage sur les comportements de mobilité. Finalement, ce mémoire concerne le dernier volet de ce projet de recherche, soit l'intégration des modèles d'activités dans l'environnement prévisionnel du MTQ.

Un projet de recherche a déjà été mené à Polytechnique Montréal et avait pour but d'enrichir le processus de modélisation de la demande de transport à l'aide des modèles d'activités (Morency & Trépanier, 2013). Ce projet a notamment permis de tester le générateur d'activités TASHA, développé à Toronto, dans la GRM, et a également mis en lumière les principales lacunes du modèle du MTQ.

1.2 Objectifs

L'objectif principal de ce mémoire est de développer une méthodologie permettant de caractériser les systèmes d'activités individuels et leur évolution. Les trois objectifs suivants sont ceux qui guident l'ensemble de la démarche de ce projet de recherche.

- Regrouper les individus qui composent l'échantillon de l'enquête OD selon la similarité de leurs systèmes d'activités à l'aide d'un algorithme de segmentation basé sur les différentes définitions d'un système d'activités et des stratégies de normalisation.
- Développer un modèle qui évalue la probabilité qu'un individu adopte un certain type de système d'activités en fonction de ses caractéristiques et de celles de son ménage d'appartenance.

- Modéliser les transformations dans les systèmes d'activités.

À travers ces trois objectifs, le but ultime de ce mémoire est de développer une méthode de génération d'activités individuelles. Le but est donc d'attribuer une série d'activités à chaque individu composant la population synthétique issu de l'étape précédente du projet de recherche global. Ces activités ont un type, une heure de début ainsi qu'une durée.

1.3 Structure du mémoire

Ce mémoire se divise en sept chapitres, incluant l'introduction qui compose le présent chapitre et qui permet de cerner le contexte dans lequel les premiers modèles de transport sont apparus ainsi que les enjeux qui remettent en cause leurs structures.

Le chapitre suivant permet d'approfondir le fonctionnement de ces modèles, soit la procédure séquentielle classique, le modèle utilisé au MTQ et les modèles d'activités. Les avancés et les tentatives d'amélioration des modèles d'activités sont également présentés dans ce chapitre.

Le troisième chapitre présente tout d'abord la principale source de données utilisée dans le cadre de ce mémoire, soit l'enquête Origine-Destination. La méthodologie permettant d'exploiter ces données est par la suite exposée.

Le quatrième chapitre contient les résultats de la première étape de ce projet de recherche, soit la segmentation des individus composant l'échantillon de l'enquête OD sur la base de la similarité de leur système d'activité. Trois différentes méthodes sont testées afin de comparer le système d'activités des individus et identifier des patrons d'activités typiques.

Le cinquième chapitre propose un modèle qui a pour but de prédire le groupe auquel un individu appartient sur la base de ses caractéristiques sociodémographiques. Il s'agit donc de trouver une méthode d'allocation des patrons d'activités typiques en fonction de l'âge, du genre, du type de ménage et de la localisation du domicile d'un individu.

Le sixième chapitre a pour objectif de quantifier et de mesurer l'évolution des systèmes d'activités.

Le septième et dernier chapitre conclut ce mémoire en résumant les principales contributions issues de la réalisation de ce projet. Les limites de la méthode ainsi que les perspectives de recherche y sont également présentées.

CHAPITRE 2 REVUE DE LITTÉRATURE

Les modèles de prévision de la demande en transport sont des outils indispensables au développement d'une région métropolitaine. Elles amènent un éclairage crucial lors de l'analyse de différents scénarios d'offre de transport ou de mise en place de politiques publiques. En effet, la modélisation des comportements de mobilité est un outil essentiel afin que les planificateurs urbains puissent être en mesure de prendre des décisions face aux enjeux de mobilité. Elles n'ont pas comme objectif de prédire le futur, mais bien d'évaluer les impacts potentiels de divers scénarios sur l'utilisation des réseaux de transport pour ensuite mesurer les effets de ces scénarios grâce à plusieurs indicateurs, tant au niveau social, environnemental et économique (MTQ, 2007).

Par définition, un modèle est une représentation simplifiée d'un phénomène complexe (Verreault, 2018). Les modèles en transport servent donc à simuler les comportements de mobilité, dont la complexité peut être élevée, en utilisant par exemple des équations mathématiques pour tenter de répliquer le processus décisionnel d'un individu. Partant de cette définition, les différents modèles de prévision développés jusqu'à présent ont tous des structures, des hypothèses et des théories différentes sur lesquelles s'appuyer, et ont également tous des objectifs et des raisons d'être tout aussi diversifiés.

Le chapitre qui suit a pour but de décrire les principaux modèles qui se sont succédés depuis leur toute première apparition ainsi que les diverses tentatives visant à les améliorer dans le monde de la recherche. Un élément essentiel à l'amélioration de ces modèles est la compréhension des liens qui existent entre les caractéristiques des personnes et des ménages et les comportements de mobilité qu'ils risquent d'adopter. C'est pourquoi les liens qui unissent ces deux éléments sont également exposés dans le présent chapitre.

Depuis la création des premiers modèles dans les années 1950, ceux-ci ont beaucoup évolué jusqu'à aujourd'hui. La Figure 2-1 présente leur évolution, en indiquant l'objet et le niveau de modélisation pour chaque génération de modèles.

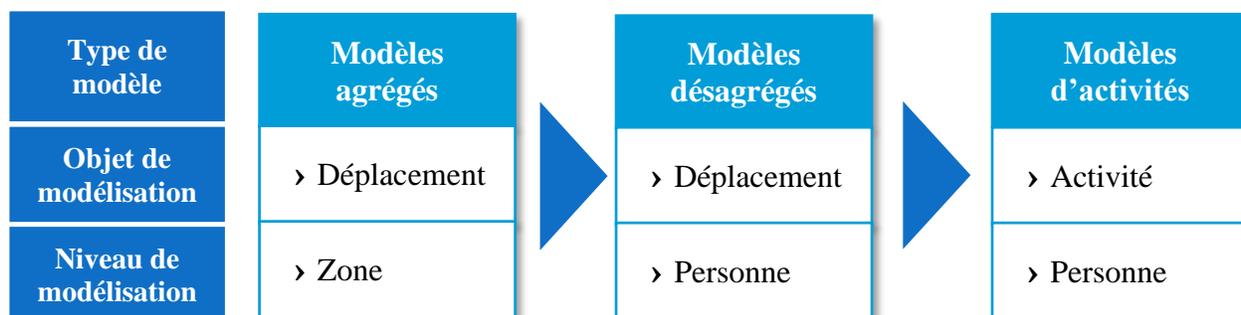


Figure 2-1 Évolution des modèles de prévision de la demande en transport

Comme mentionné dans l'introduction, la naissance des premiers modèles remonte au milieu du XX^{ème} siècle et précède la création de la PSC au début des années 1960 (Masson, 1998). Cette première génération de modèle est agrégée, ce qui signifie que la région urbaine où ce genre de modèle est appliqué nécessite une division spatiale de l'espace dans le but de créer des secteurs. C'est donc à ce niveau que la modélisation est réalisée.

Afin de pallier les limitations de la PSC, les modèles désagrégés ont émergé dans les années 1970 pour corriger le principal défaut de la PSC, soit sa nature agrégative. Plusieurs modèles se sont développés dans lesquels l'individu était au centre de l'exercice de modélisation. Toutefois, un autre enjeu majeur demeurait, soit le fait que ce sont les déplacements qui sont modélisés et non ce qui motivent les personnes à se déplacer, les activités. Ainsi, puisque l'objet principal de ces méthodes est le déplacement, les liens qui peuvent exister entre les différents déplacements ne sont généralement pas considérés dans l'exercice de modélisation. Les premières tentatives d'intégrer cette interdépendance entre les déplacements d'une personne se manifestent à travers l'apparition des modèles basés sur les chaînes de déplacements. Ces modèles, qui précèdent tout juste les modèles d'activités, permettent d'intégrer certains liens qui peuvent exister entre les déplacements d'une même chaîne, comme l'utilisation d'un mode dans l'ensemble de la chaîne. Toutefois, ces modèles ne considèrent pas les liens entre les différentes chaînes qui peuvent être réalisés au cours d'une journée.

Les modèles dont les activités étaient au cœur du processus prévisionnel sont apparus dans les années 1980. Le fait que la demande en transport est une demande dérivée est bien reconnu dans le monde de la recherche et ce, depuis des décennies. Malgré cette reconnaissance, les modèles d'activités tardent à s'imposer dans la pratique.

2.1 Procédure séquentielle classique

La PSC est une approche de modélisation agrégée, où l'information de l'ensemble des individus est rassemblée pour chaque zone composant la région à l'étude (Bonnell, 2004). Les données sur l'économie ou l'utilisation du sol, qui sont nécessaires à la réalisation de la procédure, sont également agrégées. À la suite du découpage spatial, la PSC peut être exécutée grâce à quatre grandes étapes, soit la génération, la distribution, le choix modal et l'affectation. Celles-ci sont illustrées à la Figure 2-2.

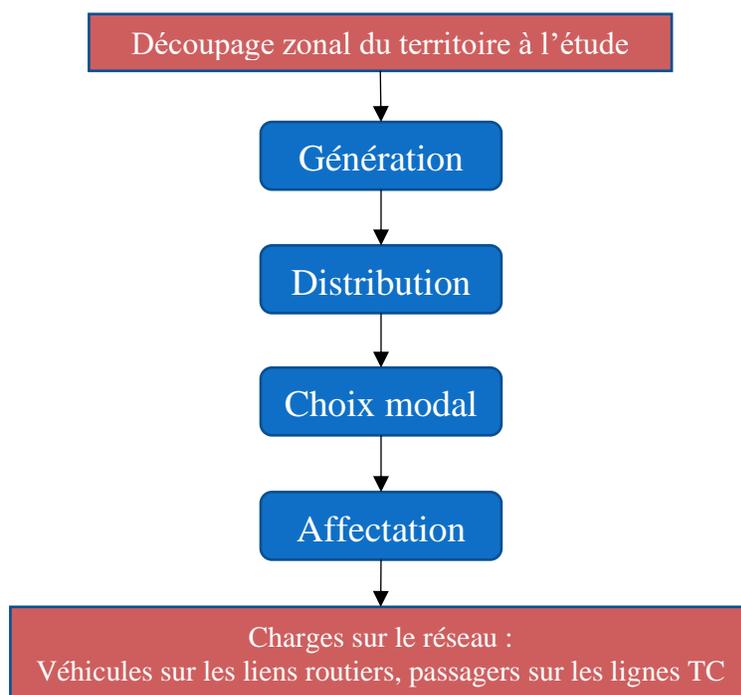


Figure 2-2 Les grandes étapes de la procédure séquentielle classique

L'exécution des quatre étapes, qui sont détaillées ci-dessous, permet d'obtenir les charges sur le réseau de transport du territoire à l'étude. Il est ainsi possible d'évaluer le nombre de véhicules sur les tronçons du réseau routier ainsi que le nombre d'utilisateurs de transport en commun sur les différents composants du réseau de transport collectif de la région en question.

2.1.1 Génération

L'étape de génération permet d'établir le nombre de déplacements produits et attirés par chacune des zones du territoire étudié. À partir des données sociodémographiques agrégées par zone, le

modèle tente d'estimer le nombre de déplacements que l'ensemble des individus de chaque zone vont réaliser et ce, pour chaque motif de déplacement. Cette estimation peut se faire avec une régression linéaire dont la variable à modéliser est le nombre de déplacements.

Le nombre de déplacements qu'une zone va attirer peut être déterminé grâce aux données relatives à l'occupation du sol ou à l'économie, comme le nombre d'emplois ou le nombre d'inscriptions scolaires. Préalablement aux prochaines étapes de la procédure, une matrice origine-destination est créée permettant de calculer le nombre de déplacements entre chacune des zones. L'étape de génération permet de remplir les marges de cette matrice où le nombre de déplacements produits et attirés est inscrit.

2.1.2 Distribution

Une fois que le nombre de déplacements produits et attirés par chacune des zones est connu, ceux-ci peuvent être distribués afin de connaître le nombre de déplacements entre chacune des zones. Cette étape correspond donc à remplir la matrice mentionnée précédemment. Il existe plusieurs modèles permettant de réaliser cette étape, comme le modèle proportionnel, le modèle à facteur de croissance ou le modèle d'opportunité ou d'entropie.

Le modèle le plus utilisé est cependant le modèle gravitaire. Outre le nombre de déplacements, la principale information sur laquelle repose ce modèle est le coût qui sépare chacune des zones, qui peut correspondre à la distance euclidienne ou au temps de parcours qui sépare le centroïde des différentes zones. Dans ce modèle-ci, la distribution des déplacements entre deux zones est influencée par l'éloignement entre elles ainsi que par la quantité de déplacements produits et attirés. Autrement dit, plus le coût entre deux zones est faible et plus la quantité de déplacements produits ou attirés par l'une ou l'autre des zones est grande, plus il y aura de déplacements entre les deux zones. L'inverse est également vrai. Un facteur d'ajustement peut également être estimé pour prendre en compte les autres effets qui peuvent influencer le nombre de déplacements entre deux zones.

2.1.3 Choix modal

Alors que l'étape précédente permet de déterminer la zone d'origine et de destination des déplacements, cette étape-ci permet de déterminer la proportion de déplacements réalisés pour

chacun des modes. Plusieurs méthodes existent afin d'y parvenir, dont une largement utilisée qui consiste à évaluer les probabilités d'utiliser chacun des modes grâce à une régression logistique multinomiale. La variable expliquée devient donc le mode de déplacements, alors que les variables explicatives peuvent être reliées aux caractéristiques des individus et des ménages, comme l'âge, le genre, la possession automobile et également aux attributs des déplacements, comme le temps de parcours pour chacun des modes. Ce dernier peut être déterminé grâce à l'étape suivante, dans le cas où l'application de l'ensemble de la PSC se fait de façon itérative.

2.1.4 Affectation

Une fois que l'origine, la destination et le mode de transport des déplacements entre chacune des zones sont connus, il est possible d'affecter ceux-ci sur le réseau de transport. Le chemin le plus court pour chaque déplacement est ainsi calculé sur le réseau, grâce à un coût généralisé qui peut comprendre entre autres le temps de déplacement et le coût monétaire. Toutefois, le chemin choisi peut dépendre du niveau de congestion sur le réseau, qui ne peut être déterminé qu'à la toute fin du processus d'affectation. C'est pourquoi il est nécessaire d'affecter les déplacements sur le réseau de façon itérative, afin d'amener le niveau d'utilisation des tronçons des différents réseaux à un équilibre.

2.1.5 Limites de la PSC

Plusieurs tentatives ont été menées pour améliorer la procédure, en créant par exemple des boucles d'interactions entre les différentes étapes. Ainsi, les résultats finaux de la procédure peuvent avoir une influence considérable sur les résultats des étapes intermédiaires. C'est pourquoi les améliorations apportées avaient pour but de tenir compte du niveau de congestion lors de l'application des trois dernières étapes.

Cependant, la principale limite de la PSC réside dans le fait que celle-ci est agrégée. En effet, le fait d'agréger de l'information rend la procédure incapable de considérer les dynamiques qui influencent la prise de décision individuelle. Ainsi, plusieurs facteurs ayant une incidence considérable sur les comportements de mobilité ne peuvent être pris en compte.

Finalement, l'objet principal de modélisation de la PSC est le déplacement, ce qui constitue également une limite de la procédure puisque, comme mentionné précédemment, la demande

en transport est une demande dérivée. C'est donc cette limite que vise à pallier ce mémoire, en démontrant, grâce aux données collectées dans la GRM, que l'application d'une approche basée activités permet de mieux cerner les comportements de mobilité.

2.2 Approche utilisée au MTQ

Le rôle des grandes étapes de la procédure utilisée au MTQ est similaire à celles de la PSC. Cette dernière a été modifiée afin de tirer profit au maximum des données de l'enquête Origine-Destination (MTQ, 2007). Ainsi, l'approche utilisée au ministère, du moins pour certaines étapes majeures, est en lien direct avec l'enquête OD et la façon dont celle-ci est structurée. L'approche mise de l'avant par le ministère tente d'intégrer les diverses tendances sociodémographiques pour évaluer leurs impacts sur l'utilisation des réseaux de transport et tenir compte de leur effet lors de l'analyse de différents scénarios à l'étude. La Figure 2-3 illustre le processus de l'ensemble du modèle.

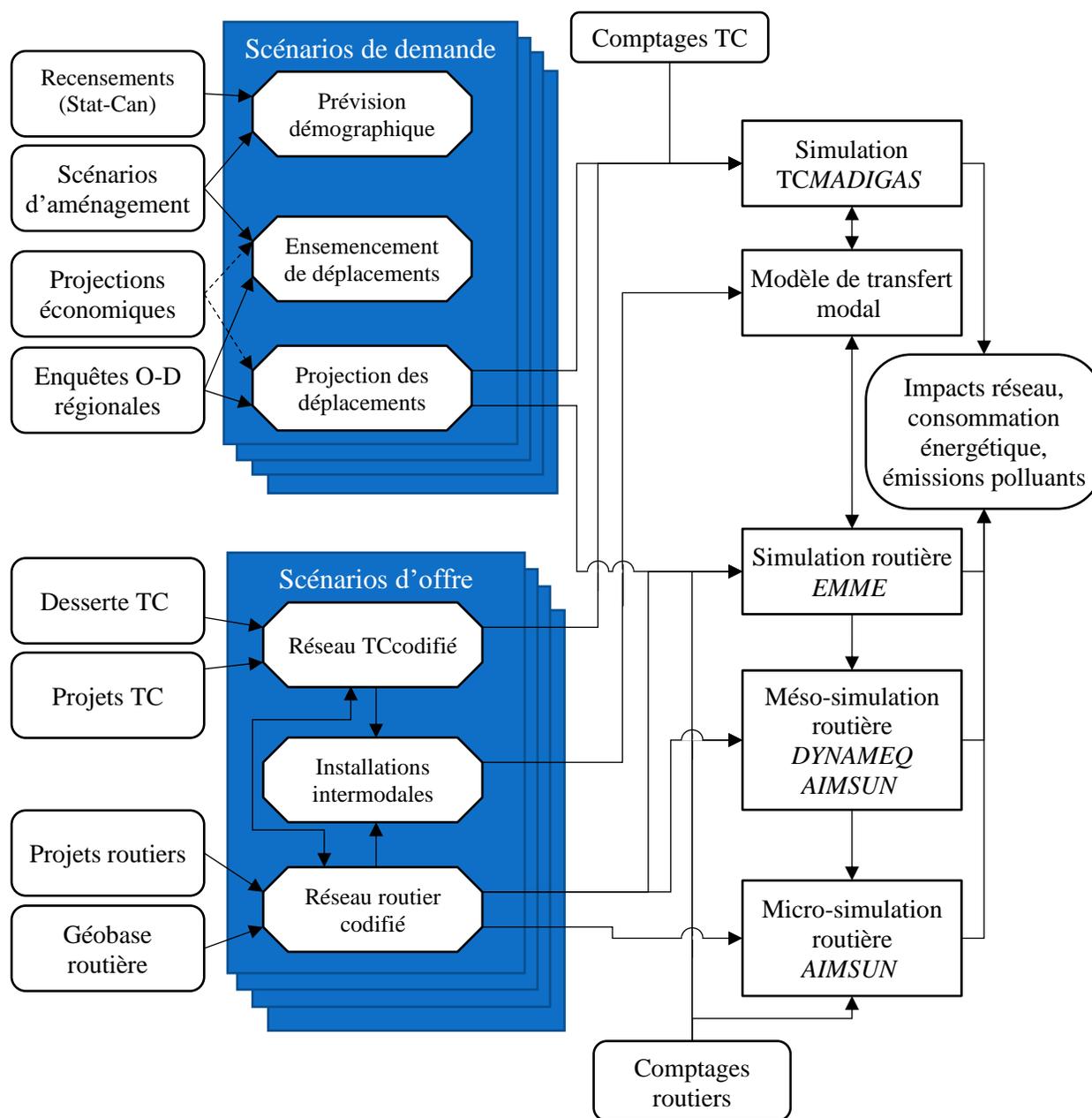


Figure 2-3 Schéma général du MOTREM (MTQ, 2017b)

Les éléments à gauche représentent l'ensemble des jeux de données nécessaires au fonctionnement du modèle. Outre l'enquête OD, le recensement canadien, les scénarios d'aménagement et les projections économiques permettent d'évaluer la demande de déplacements à travers plusieurs étapes. Tout d'abord, la projection de la population est réalisée pour l'ensemble du Québec de façon agrégée. Par la suite, grâce aux données provenant de l'enquête OD et des projections

démographiques, la projection des déplacements est réalisée de manière désagrégée pour l'ensemble de la GRM.

Pour ce qui est de l'offre, le réseau de transport en commun et le réseau routier sont tous deux codifiés grâce aux données TC et à la géobase routière. À cela s'ajoutent les différents projets de développement pour les deux réseaux. La combinaison de ces projets constitue les différents scénarios d'offre qui sont testés lors de l'exécution du modèle.

Grâce à la codification de l'offre et l'estimation de la demande, il est ainsi possible d'affecter la demande sur les réseaux TC et routier. L'affectation de la demande routière se fait à trois différents niveaux, soit au niveau microscopique, mésoscopique et macroscopique. De plus, un modèle de transfert modal est utilisé pour déterminer le nouveau mode des déplacements projetés à la suite de l'affectation de ceux-ci sur les réseaux. Grâce à l'ensemble de ces résultats, il est possible de déterminer les charges sur le réseau de transport collectif de même que le réseau routier. La consommation énergétique due aux déplacements générés par le modèle peut également être estimée, tout comme les émissions polluantes qui en découlent. Finalement, l'estimation de ces deux éléments permet de quantifier et d'analyser l'impact des différents scénarios.

2.2.1 L'enquête Origine-Destination

Une présentation plus détaillée de cette enquête est présentée à la section 3.2.1. En somme, cette enquête téléphonique se déroule tous les cinq ans depuis 1970 dans la Grande région de Montréal (GRM) et vise à recueillir des informations sur la mobilité des résidents de la région métropolitaine. Un élément essentiel de cette enquête est le facteur de pondération, qui permet d'indiquer le poids de chaque ménage et individu dans l'échantillon pour ainsi être en mesure de représenter l'ensemble de la population. Ces facteurs sont au cœur de la démarche visant à projeter et modéliser la demande de déplacements.

2.2.2 Projection de la population

Afin de projeter la demande de déplacements à divers horizons temporels, il est également nécessaire de projeter la population. Les projections démographiques sont produites grâce à l'outil ES-3 (Entrants-sortants à trois niveaux) qui a été développé par le service de l'économie et du plan directeur en transport du MTQ en 1995-96. Ce modèle permet de projeter la population de

l'ensemble du Québec à travers 1700 zones de projection selon l'âge et le genre (Desgagnés & Bédard, 2005). En se basant sur les données des deux derniers recensements, la population de chacun de ces secteurs peut être projetée grâce aux taux de mortalité et de fécondité, de même qu'avec les données sur les migrations intermunicipales. Les résultats de la procédure sont fournis pour chaque segment de population, défini par l'âge et le genre. L'âge est divisé en catégories de cinq ans d'intervalle.

2.2.3 Projection des déplacements

La projection de la demande en transport s'effectue avec le logiciel MADEDE créé par le groupe de recherche MATIDUC à l'École Polytechnique de Montréal au milieu des années 90. Les facteurs de pondération de l'enquête OD sont à la base du processus qui vise à modifier leur poids à divers horizons temporels selon l'évolution de facteurs déterminants (MTQ, 2017b). Les facteurs qui ont une influence sur la demande en transport sont donc sélectionnés et des tendances sont dégagées pour ensuite les appliquer aux facteurs de pondération de la façon suivante :

$$F_{tsg}^{2036} = F_{tsg}^{2013} * \frac{P_{tsg}^{2036}}{P_{tsg}^{2013}} * \frac{S_{tsg}^{2036}}{S_{tsg}^{2013}} * \frac{M_{tsg}^{2036}}{M_{tsg}^{2013}} * \frac{PM_{tsg}^{2036}}{PM_{tsg}^{2013}}$$

Le facteur de pondération de chaque déplacement de l'enquête OD la plus récente, dans ce cas-ci celle de 2013, est donc ajusté selon l'évolution hypothétique des effectifs dans les quatre catégories suivantes : la population (P), le statut d'activité (S), la motorisation (M) et le profil modal (PM). Ainsi, pour chacune des catégories, le ratio entre les effectifs de 2013 et d'une année ultérieure, par exemple 2036, peut être obtenu et utilisé dans la formule ci-haut. Cet ajustement se fait également par segment de population, selon le territoire de domicile (t), le sexe (s) et le groupe d'âge (g) de chaque individu réalisant le déplacement dont la pondération est altérée. Deux autres transformations sont également appliquées à certains segments de la demande afin de tenir compte de l'étalement de la période de pointe du matin ainsi que de la reconfiguration des lieux d'emplois dans la région métropolitaine.

2.2.3.1 La démographie

Les tendances en matière démographique sont obtenues grâce aux résultats du modèle ES-3, décrit à la section 2.2.2. Il est ainsi possible d'obtenir le nombre de personnes à une certaine année selon le territoire du domicile, le genre et l'âge.

2.2.3.2 Le statut d'activité

Le statut d'activité est divisé en quatre catégories, soit les travailleurs, les étudiants, les autres mobiles et les non-mobiles. Le nombre de personnes appartenant à chacune de ces catégories pour une année ultérieure est déterminé grâce à l'évaluation des tendances observées dans les enquêtes OD précédant la plus récente. Des hypothèses sont donc posées pour obtenir les effectifs futurs dans chaque segment de la population et ainsi calculer le ratio nécessaire à la repondération des déplacements.

2.2.3.3 La motorisation

Pour la motorisation, la catégorisation des individus se fait selon la situation qui prévaut dans le ménage. Trois situations peuvent être observées : le ménage ne possède aucun véhicule et chaque individu est considéré comme non-motorisé; il y a moins de véhicules que de personnes pouvant conduire, ceux-ci sont en auto-négociation; il y a autant de personnes pouvant conduire que de véhicules, les individus aptes à conduire entrent dans la catégorie des auto-suffisants. Le nombre de personnes appartenant à chacune de ces catégories est donc comptabilisé et ce pour chaque tranche de population, permettant ainsi de calculer le ratio. En plus du territoire de domicile, du genre et de l'âge, la population est également subdivisée selon la présence d'un enfant de moins de 15 ans dans le ménage.

2.2.3.4 Le profil modal

Six catégories ont été créées pour segmenter les personnes selon les modes qu'ils utilisent au cours d'une journée : automobile conducteur seulement (AC), transport collectif seulement (TC), automobile passager seulement (AP), transport collectif multimodal (TC-AC-Marche), vélo seulement, autre (toute autre combinaison de modes). Par suite de l'attribution d'un profil, le nombre d'individus correspondant à chacun d'entre eux est dénombré et ce pour chaque tranche de

la population, qui est définie par le territoire du domicile, le genre, l'âge, la motorisation et la présence d'enfants de moins de 15 ans dans le ménage.

2.2.3.5 L'étalement de la période de pointe

Afin de considérer l'étalement de la période de pointe, un segment de la demande en transport est ajusté. Cette transformation s'applique uniquement aux déplacements faits en pointe du matin pour le motif études et travail ainsi qu'aux déplacements réalisés par des personnes conduisant leur véhicule. Elle consiste à calculer le poids des déplacements faits entre 5h et 6h par rapport à l'ensemble des déplacements faits en période de pointe, soit de 5h à 9h59. Ce poids est calculé sur plusieurs enquêtes pour ensuite être projeté dans le temps.

2.2.3.6 La redistribution des emplois

Un ajustement de type Furness est également apporté à la demande pour tenir compte de l'augmentation du nombre de déplacements à motif travail vers des pôles d'emplois autres que le centre-ville de Montréal. Cette reconfiguration spatiale des lieux d'emplois diminue l'attractivité du centre-ville au profit des pôles présents en banlieue. L'ajustement permet d'altérer le poids des déplacements afin d'intégrer cette tendance dans la prévision.

2.2.4 Transfert modal

Alors que la PSC inclut un modèle de choix modal afin de déterminer le mode emprunté par les déplacements générés et distribués, le modèle du MTQ mise plutôt sur un modèle de transfert modal (MTQ, 2002). Le principe de ce modèle consiste à transférer le mode observé de chaque déplacement vers un autre mode selon un certain critère. L'automobile et le transport en commun sont les deux seuls modes considérés dans la démarche du MTQ. Le transfert s'effectue sur la base du temps de parcours, incluant le temps de transfert et d'attente dans le cas du TC, obtenu grâce à l'affectation des déplacements sur les réseaux de transport.

2.2.5 Affectation des déplacements

Une fois que les déplacements sont projetés dans le temps à divers horizons, ceux-ci sont affectés sur le réseau routier et le réseau de transport en commun. L'affectation routière s'effectue avec trois logiciels, soit EMME, DYNAMIQ et AIMSUN, selon le niveau de précision de l'affectation

(MTQ, 2002). Ce niveau peut être microscopique, soit à la plus petite échelle, mais également à l'échelle d'une sous-région, qui correspond à l'échelle mésoscopique, ou finalement pour l'ensemble de la GRM. L'affectation routière se fait de manière agrégée, contrairement à celle du transport collectif, qui est désagrégé. La simulation du transport en commun se fait grâce au logiciel MADIGAS de la suite MADITUC. Ce logiciel a été développé dans le cadre d'un projet de recherche mené à Polytechnique pour le ministère. La force de ce logiciel réside dans son arrimage avec les données de l'enquête OD, qui lui permet de tenir compte des préférences des individus lors du choix d'un itinéraire de transport collectif.

2.2.6 Limite de la procédure du MTQ

La première limite concerne l'influence de l'offre de transport. En effet, cette méthode prévisionnelle se base sur le fait que ce sont les changements sociodémographiques qui sont à la source des transformations de la demande de déplacements. Les différents scénarios d'offre de transport ne sont donc pas considérés lors de la prévision de la demande, mais uniquement dans les étapes subséquentes, lors de l'affectation et du transfert modal.

Une des limites de la procédure du ministère est semblable à celle de la PSC, soit que ce sont les déplacements qui sont au cœur de la démarche et non les activités. De plus, la méthode de prévision de la demande consiste essentiellement à altérer le poids de chaque déplacement de l'enquête OD. Ce poids est donc ajusté selon diverses tendances sociodémographiques, mais la structure de chacun de ces déplacements reste inchangée pour chaque segment de la population (Morency & Trépanier, 2013). Cette limite a pour effet de négliger certaines tendances au niveau des caractéristiques des systèmes d'activités des individus, comme la durée ou l'heure de début des activités. C'est pourquoi il devient pertinent d'intégrer la modélisation des activités dans le processus du MTQ.

2.3 Les modèles d'activités

Les fondements des modèles d'activités prennent racine dans les travaux d'Hägerstrand (1970) ainsi que ceux de Chapin (1974). Les travaux d'Hägerstrand s'articulent autour du prisme espace-temps, qui permet de conceptualiser le système d'activités d'un individu dans le temps et dans l'espace (Hägerstrand, 1970). Il existe trois contraintes au prisme espace-temps. La première se

définit comme étant la capacité physique d'un individu, soit la limite de se mouvoir selon les contraintes physiologiques d'un être humain qui est propre à chaque personne. La deuxième contrainte concerne les ressources humaines et matérielles qui sont nécessaires à la réalisation d'une activité, comme la présence d'un membre de la famille, un ami, un client ou l'utilisation d'une automobile. Finalement, la troisième contrainte se définit par l'ensemble des normes, lois, règlements qui régissent notre société, comme l'heure d'ouverture d'un magasin, les restrictions concernant le bruit ou l'utilisation par une autre personne d'une cabine téléphonique. Bref, chaque individu est contraint par le prisme défini par ces trois facettes lors de la réalisation d'une activité hors de son domicile. Le principe du prisme espace-temps est présenté sur la Figure 2-4 qui illustre la trajectoire d'un individu au cours d'une journée. Le prisme permet de visualiser les contraintes qui dictent les heures de départ et d'arrivée possibles du déplacement vers le lieu de travail.

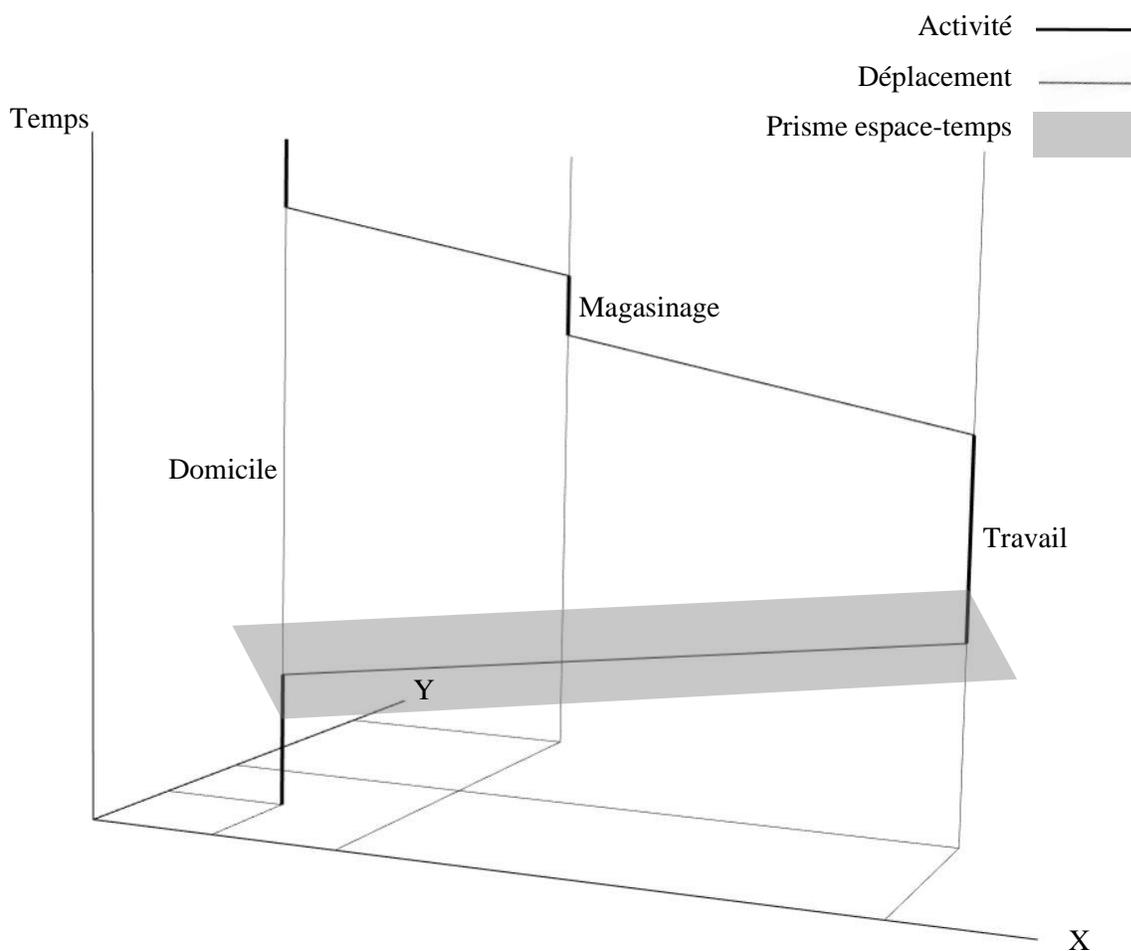


Figure 2-4 Trajectoire d'un individu dans l'espace-temps

Ce schéma permet donc de visualiser comment un individu s'approprie le temps et l'espace nécessaire à la réalisation de son programme d'activité. Il permet également de comprendre que c'est dans le temps et l'espace que les trois contraintes mentionnées précédemment prennent forme et s'imposent aux individus. Le domicile d'une personne se définit donc comme l'ancrage à partir duquel un potentiel spatial s'offre à une personne. Ce potentiel permet à chaque individu de s'approprier l'espace nécessaire à l'assouvissement de ses besoins et ses désirs en réalisant des activités hors domicile. De plus, cette appropriation doit se faire à l'intérieur d'une quantité de temps déterminée, par exemple vingt-quatre heures lors d'une journée ou sept jours lors d'une semaine. Il s'agit donc, pour chaque individu, de s'emparer de cet espace-temps et d'en maximiser l'utilité afin d'accomplir les activités souhaitées.

Cette maximisation se traduit par une série de décisions auxquelles chaque personne fait face au cours d'une journée ou d'une semaine. Ces décisions concernent d'abord et avant tout le choix de faire ou non une activité et, par la suite, le lieu et le moment de la réalisation de celle-ci. Ce processus décisionnel s'appuie sur un principe fondamental de la demande en transport, soit que le fait de se déplacer est souvent perçu comme un mal nécessaire à la réalisation d'activités hors domicile. Un déplacement est effectué lorsque la désutilité engendrée par le temps consacré à se mouvoir est moindre que l'utilité associée à la réalisation de l'activité.

De plus, cette désutilité dépend également des autres déplacements effectués au cours d'une journée. En effet, selon la loi de Zahavi, le temps consacré aux déplacements dans l'ensemble d'une région métropolitaine est en moyenne constant à travers le temps (Joly, Crozet, Bonnel, & Raux, 2002). Cela implique que chaque personne a un budget temps qui limite la durée totale des déplacements effectués lors d'une certaine période. Autrement dit, chaque individu est prêt à consacrer un certain temps à se déplacer et les décisions prises au niveau individuel amène chaque personne à atteindre de cette durée maximale, sans la dépasser, mais sans pour autant la gaspiller. Ainsi, la décision de réaliser une activité ou non dépend donc des autres activités réalisées au cours d'une journée ou d'une semaine, puisque le temps consacré à accéder au lieu des autres activités augmente la désutilité des déplacements nécessaires à la réalisation d'une activité supplémentaire.

Bref, le processus décisionnel d'une personne face aux choix qui façonnent son système d'activités est donc non seulement contraint par un budget temps limité, mais également par la volonté de combler ses besoins et ses désirs. Cette volonté se traduit tout d'abord par une planification de la

séquence des activités réalisées au cours d'une période. Cette planification se bute parfois aux diverses contraintes du prisme espace-temps et amène les individus à modifier ce qu'ils avaient prévu initialement. L'atteinte du budget temps, la présence d'un autre membre du ménage lors d'une activité, la nécessité d'utiliser une automobile pour accéder à un lieu, l'heure de fermeture d'un magasin, sont toutes des contraintes qui régissent le comportement des personnes et les choix qu'elles font à tout moment. L'effet de ces contraintes sur les décisions individuelles permet de comprendre un des fondements essentiels dans les modèles d'activités, soit que les différentes activités, et par conséquent les déplacements, qu'un individu choisit de réaliser au cours d'une journée ou au cours d'une semaine sont tous interreliées. C'est pourquoi il importe de ne pas modéliser uniquement les déplacements de façon isolée, mais bien l'ensemble du système d'activités de chaque individu.

2.3.1 La structure des modèles d'activités

Un modèle d'activité a la capacité d'intégrer l'ensemble des dynamiques qui régissent les comportements et les décisions des individus. Sa structure permet de tenir compte des liens qui existent entre les différentes activités réalisées par un individu lors d'une certaine période (Roorda, 2005). Ces modèles sont généralement divisés en deux modules, soit la génération des activités ainsi que la programmation de celle-ci (Chandra R. Bhat & Frank S. Koppelman, 1993).

Comme son nom l'indique, le premier module a comme objectif de générer une série d'activités pour chaque individu. Selon les modèles, chacune de ces activités a généralement un type, une heure de début, une durée et une localisation. Le processus par lequel les activités sont générées se base essentiellement sur la volonté d'assouvir ses besoins et ses désirs en réalisant des activités. Cependant, le module de génération de certains modèles d'activités, tels que CARLA (Jones, Dix, Clarke, & Heggie, 1983), AMOS (Kitamura, Lula, & Pas) et SMASH (Ettema, Borgers, & Timmermans, 1993), consiste uniquement à tirer aléatoirement parmi les séquences d'activités observées dans la population. Il n'y a donc aucun processus de modélisation concernant la génération des activités qu'un individu risque de faire. D'autres modèles génèrent des activités en créant des distributions des attributs des activités, comme la fréquence ou la durée, pour ensuite piger de façon aléatoire dans ces distributions. C'est notamment le cas de FAMOS (Pendyala, Kitamura, Kikuchi, Yamamoto, & Fujii, 2005) et TASHA (Roorda, 2005).

Le rôle du second module est de produire, à partir de ces activités générées, un horaire réalisable pour chaque personne en fonction de diverses contraintes temporelles, dont celles créées par les résultats du module précédent. En effet, les activités générées peuvent parfois se chevaucher, ce qui constitue une des principales contraintes nécessitant leur programmation. Ainsi, les deux modules sont interreliés, mais seuls certains modèles intègrent cette interrelation (C. R. Bhat & F. S. Koppelman, 1993).

De façon générale, parmi les modèles basés sur les activités développées jusqu'à présent, il existe deux catégories de modèles, soit les modèles économétriques ainsi que les modèles basés sur les règles. Les modèles dont les processus de génération et de programmation sont interreliés développent généralement les deux modules au sein de la même catégorie. Autrement dit, les deux composants se fondent tous les deux sur des modèles économétriques par exemple afin de faciliter leur intégration. Il existe également des modèles combinant les deux approches, économétriques et basées sur les règles et ce, à travers les deux modules.

2.3.1.1 Modèles économétriques

Les modèles économétriques se basent sur différents modèles statistiques afin de simuler les choix que les individus doivent prendre face aux décisions auxquelles ils sont confrontés. Ces modèles reposent sur la maximisation de l'utilité associée à la réalisation d'une activité, non seulement pour générer les activités, mais également pour les programmer. Cette programmation se fait en fonction des contraintes qui dictent le comportement des individus lorsque ceux-ci planifient la séquence des activités dont la génération a été faite préalablement.

Par exemple, des régressions logistiques imbriquées (nested logit) sont utilisées pour simuler les choix auxquels sont confrontés les individus. C'est le cas des modèles de Bowman and Ben-Akiva (2001) qui tentent de répliquer les comportements des personnes quant à la réalisation d'un certain patron d'activités parmi une série de patrons. Ces modèles se basent sur l'utilité associée à l'adoption que peut procurer chacun d'entre eux. Selon le patron sélectionné, le moment et la localisation spatiale de chacune des activités, de même que le mode pour y accéder, sont déterminés par la suite grâce aux modèles imbriqués.

Le fondement théorique des modèles économétriques repose sur la volonté, pour un individu, de maximiser l'utilité associée à l'ensemble des activités réalisées au cours d'une certaine période de

temps, soit une journée ou une semaine. La conception de certains modèles repose sur l'hypothèse qu'une personne planifie la réalisation de l'ensemble de ses activités au tout début de la période, sans modifier cette planification par la suite. Toutefois, en réalité, le processus décisionnel d'un individu ne correspond pas tout à fait à cette définition, mais correspond plutôt à une série de décisions venant altérer la planification préalablement établie. C'est pourquoi des modèles dynamiques sont apparus, visant à refléter cette réalité (Theo A. Arentze, Ettema, & Timmermans, 2011) et à considérer le fait qu'au cours d'une journée ou d'une semaine, les individus peuvent prendre des décisions spontanées et modifier les activités qui avaient été prévues à l'origine. L'aspect dynamique des modèles permet de modifier les décisions prises à l'origine par les individus lors de la simulation (Theo A. Arentze et al., 2011).

Un aspect essentiel aux modèles dynamiques est la catégorisation des activités, soit celles qui sont obligatoires et routinières, comme aller travailler, ou plus spontanées et moins essentielles, comme aller au cinéma. Cette division des types d'activités se fait en accordant une utilité plus grande aux activités obligatoires (Roorda, 2005).

2.3.1.2 Modèles basés sur les règles

Les modèles basés sur les règles utilisent principalement des arbres de décision afin de simuler le processus décisionnel d'un individu. L'objectif est de répliquer ce processus grâce à une série de règles qui imitent les choix des personnes auxquelles elles sont confrontées. Ainsi, lors de l'application de ce genre de modèle, un individu prend une certaine décision et adopte un certain comportement de mobilité selon diverses conditions. C'est le cas de TASHA, dont le processus de génération des activités se base sur les distributions des attributs des activités, mais dont le module de programmation utilise des arbres de décisions. Dans le cas de ce modèle, la programmation des activités avec les règles de décision se fait sous certaines contraintes, comme la disponibilité des véhicules au sein d'un ménage.

Deux types de modèles basés sur les règles ont été définis par le passé (Theo A Arentze & Timmermans, 2004). Selon cette catégorisation, les modèles peuvent soit être « faibles » ou « forts ». Les modèles de la première catégorie empruntent des théories propres aux modèles économétriques comme la maximisation de l'utilité associée à la réalisation d'une activité. Ces modèles sont essentiellement basés sur les règles, mais certaines de ces règles sont appliquées à

l'aide de modèles économétriques. Les modèles de type « forts » sont uniquement basés sur les règles et aucun modèle économétrique n'est introduit dans leurs processus.

2.3.2 La génération des activités

Bien que la recherche sur le module de génération ne soit pas aussi intense que celle sur la programmation des activités, diverses études ont tenté d'améliorer et de développer des méthodes de génération. Plusieurs d'entre elles avaient comme fondement la maximisation de l'utilité associée à la réalisation d'une activité. Ces méthodes sont donc destinées à être intégrées dans des modèles économétriques grâce à l'assignation de patrons d'activités. Outre ces types de modèles, ceux dont la génération des activités repose sur le tirage aléatoire des attributs des activités dans des distributions ont également fait l'objet de tentatives d'amélioration. Ces tentatives nécessitaient également l'utilisation de patrons typique d'activités afin de construire des distributions qui sont davantage pertinentes. Ainsi, la détection de patrons typiques ainsi que l'évaluation du niveau de similarité entre les systèmes d'activités des individus sont des tâches fondamentales dans l'amélioration des modèles d'activités. Plusieurs tentatives ont été réalisées par le passé afin de détecter des patrons similaires au sein de la population.

2.3.2.1 La méthode d'alignement des séquences

Plusieurs méthodes de détection utilisent la méthode d'alignement des séquences (Sequence Alignment Method – SAM), qui est largement utilisée dans le monde de la biologie depuis 1970. La première fois que cette méthode a été utilisée dans le but de mesurer la similarité entre les systèmes d'activités des individus est en 1998 (Wilson, 1998). Depuis l'introduction de cette méthode dans le monde de la recherche sur les modèles d'activités, elle est largement utilisée dans plusieurs études.

Cette méthode permet d'aligner deux séquences au moyen de trois opérations, soit la substitution, le retrait et l'insertion. Typiquement, une séquence représente la journée d'un individu en indiquant, à chaque moment de la journée, si une activité est réalisée ou non et si oui, de quel type est l'activité. La division de la journée varie d'une étude à l'autre, pouvant aller de cinq minutes à une heure. Ainsi, le système d'activités de chaque individu est représenté par un vecteur contenant par exemple 24 blocs dans le cas où la division de la journée est par heure. Chaque bloc contient un caractère indiquant le type d'activité réalisé par l'individu, incluant le fait d'être à son domicile.

La SAM permet, avec l'aide de ce vecteur, de comparer deux individus afin de mesurer le niveau de similarité de leurs systèmes d'activité. La méthode consiste à aligner un des vecteurs avec l'autre, en substituant, retirant ou insérant un caractère dans le premier vecteur afin qu'il soit identique au deuxième. Chaque modification implique un coût, dont la somme totale représente la distance qui sépare les deux vecteurs.

Le résultat de l'alignement peut ensuite être utilisé pour regrouper les personnes en fonction de la distance qui les sépare. Dans le cas de Wilson (1998), l'algorithme utilisé pour partitionner les individus est le même que celui utilisé lors de la première utilisation de la SAM (Needleman & Wunsch, 1970). Plusieurs études ont également utilisé d'autres algorithmes, telles que celui des K-moyennes.

La SAM a par la suite été améliorée en développant une méthode multidimensionnelle (Joh, Arentze, Hofman, & Timmermans, 2002). Alors que l'alignement de deux séquences avec la méthode conventionnelle permet uniquement de comparer deux individus selon les différents types d'activités réalisés à différents moments au cours d'une journée, cette nouvelle méthode permet d'introduire de nouvelles variables permettant de décrire le système d'activités d'un individu. Ces variables peuvent par exemple être le lieu de l'activité, le mode de déplacement ou la présence d'une autre personne. La méthode multidimensionnelle débute tout d'abord par aligner les différents vecteurs correspondant à la même dimension, puis à comparer l'alignement de chacune de ces dimensions. La SAM multidimensionnelle a également été comparée à différentes méthodes de détection des patrons (Joh, Arentze, & Timmermans, 2001)

Alors que la majorité des études utilisant la SAM se base uniquement sur une seule journée pour comparer le système d'activités de deux individus, certaines ont tenté d'utiliser cette méthode sur une semaine (Saneinejad & Roorda, 2009). Tel que mentionné précédemment, le processus par lequel un individu choisit les activités qu'il va faire peut s'étaler sur une journée, mais également sur une semaine complète. C'est pourquoi la SAM a été appliquée à cette fenêtre temporelle. La première étude du genre a été réalisée en 1986 (Hirsh, Prashke, & Ben-Akiva, 1986).

La SAM a également été intégrée à une méthode permettant non seulement de comparer les individus, mais également de leur attribuer un patron typique (Wongchavalidkul & Piantanakulchai, 2015). Ainsi, la méthode d'alignement est utilisée de façon conventionnelle, pour ainsi mesurer le niveau de similarité des systèmes d'activités de chaque individu. Par la suite, le

résultat de la procédure est utilisé pour bâtir un arbre de décision qui permet d'attribuer un patron typique à chaque individu. Différents attributs sociodémographiques sont utilisés pour les critères à chacun des nœuds et ces derniers sont établis selon le niveau de similarité des individus, obtenu grâce à la SAM.

2.3.2.2 Autres méthodes de mesures de similarité des systèmes d'activités

Outre la SAM, plusieurs tentatives se sont succédé dans le but d'identifier des patrons typiques d'activités. Les différentes études ayant cet objectif reposent tout d'abord sur une transformation des données dans le but de comparer chaque personne, suivie par l'application d'un algorithme visant à rassembler les personnes ayant des systèmes d'activités similaires. Dans certaines études, la transformation des données consiste à créer un vecteur qui indique, à chaque moment de la journée, si une activité d'un certain type est réalisée grâce à une variable binaire d'une valeur de 0 ou 1. C'est le cas d'une étude dont la journée de chaque individu était divisée en 288 blocs d'une durée de cinq minutes et dont l'algorithme utilisé afin de regrouper les personnes est la segmentation des C-moyennes flou (« fuzzy C-means ») (Hafezi, Liu, & Millward, 2017).

Plusieurs auteurs ont également tenté d'intégrer une dimension spatiale à la mesure du niveau de similarité des systèmes d'activités. Cette intégration consiste à attribuer un haut niveau de similarité à des systèmes d'activités dont les lieux d'activités des différentes personnes sont à proximité les uns des autres (Zhang & Thill, 2017). Plusieurs autres études ont également introduit une composante spatiale à l'identification de patrons d'activités (Allahviranloo & Priol, 2018; Chen et al., 2011; Shen & Cheng, 2016)

L'utilisation d'indicateurs pour définir le niveau de similarité des systèmes d'activité est également une méthode qui a été testée (McBride, Davis, & Goulas, 2019). Cette étude se base sur le calcul de deux indicateurs permettant de caractériser le système d'activités de chaque personne. Le premier indicateur a pour but de mesurer la variabilité présente dans le système d'activités, en mesurant l'entropie de ce dernier en fonction du nombre de types d'activité observé. Le second indicateur a comme objectif d'établir le niveau de complexité de la structure du système, en considérant la durée des activités et l'ordre dans lequel elles sont réalisées.

D'autres auteurs ont également utilisé des modèles de Markov caché afin de caractériser des séquences d'activités (Liu, Janssens, Cui, Wets, & Cools, 2015). En outre, plusieurs études ont été

menées visant à identifier des patrons typiques (Hafezi, Liu, & Millward, 2019; Jiang, Ferreira, & González, 2012; Lin, Lo, & Chen, 2009). Des méthodes ont également été développées dans le but de détecter des patrons d'activités propres à un ménage entier. Ainsi, les dynamiques présentes dans un ménage concernant la structure du système d'activités des membres le composant peuvent être considérées dans la caractérisation des patrons (Hannes et al., 2012).

Bref, plusieurs tentatives se sont succédé dans les dernières années afin d'analyser la similarité des systèmes d'activités des individus. Les différents paramètres propres à chacune de ces tentatives varient grandement d'une étude à l'autre. La variabilité de ces paramètres concerne tout d'abord les caractéristiques des données utilisées dans le cadre de ces études, comme le nombre d'observations, la diversité et le niveau de détails des types d'activités ou l'étendue temporelle de la période de temps observée. En effet, le nombre d'individus sondés varie de quelques centaines à quelques milliers d'individus selon les différentes études. Les types d'activités observés peuvent se résumer aux principaux motifs de déplacement, comme le travail ou les études, mais peuvent également inclure des types très précis, comme dormir, manger, faire le ménage ou s'affairer à des activités de détente au domicile. De plus, les personnes sondées peuvent également déclarer les caractéristiques des activités réalisées lors d'une journée, mais également lors d'une semaine complète, incluant ou non la fin de semaine. Finalement, les patrons d'activités peuvent également concerner l'ensemble du ménage plutôt qu'un seul individu.

2.3.3 Les individus, les ménages et les systèmes d'activités

Un autre élément essentiel dans l'élaboration d'une méthode de génération des activités est la compréhension de l'influence des attributs individuels sur la structure des systèmes d'activités. En effet, une hypothèse fondamentale dans les modèles d'activités est que les patrons typiques d'activités que les individus adoptent sont influencés par les caractéristiques des personnes et des ménages. L'influence des caractéristiques sociodémographiques sur les activités réalisées, de même que sur les comportements de mobilité en général, fait l'objet de beaucoup de recherche dans la littérature.

2.3.3.1 Tendances sociodémographiques et comportementales

Préalablement à l'analyse des liens entre les activités réalisées et les attributs d'une personne, il est également nécessaire d'identifier les tendances qui émergent au niveau de la composition

démographique de la population ainsi que sur les changements de comportement qu'il est possible d'observer dans la société. L'élément le plus marquant en ce qui concerne l'évolution démographique est le vieillissement de la population (Nation Unies, 2017). En effet, la proportion de personnes âgées dans la majorité des pays industrialisés est grandissante. Un taux de natalité très élevée durant les décennies suivant la Seconde Guerre mondiale, suivi par la suite d'un taux de natalité diminuant sans cesse, explique ce phénomène.

Un autre phénomène important concerne la composition des ménages dont la taille moyenne a diminué de façon importante depuis la fin de la Seconde Guerre mondiale. La baisse a été particulièrement importante au Québec, passant de 4,6 en 1951 à 2,3 en 2016 (Québec, 2018). La diminution du taux de natalité explique également ce phénomène. Alors que le nombre moyen d'enfants par ménage est de plus en plus faible, le nombre de personnes vivant seules est en hausse, contribuant à la baisse de la taille moyenne des ménages.

L'augmentation du nombre de femmes sur le marché du travail constitue également un changement important dans la société (Fortin, Godbout, & St-Cerny, 2012). De manière générale, les changements de comportements observés chez les femmes ont des répercussions sur la dynamique au sein des ménages. Les différences entre les genres tendent à se dissiper, alors que les femmes et les hommes participent de la même façon à la vie familiale. Ce changement de dynamique a également un impact sur le mode de vie des enfants, qui sont de moins en moins nombreux à quitter l'école pour aller dîner à la maison.

2.3.3.2 L'impact des attributs individuels et des ménages sur les systèmes d'activités

Dans le cas des systèmes d'activités, l'influence du genre, de l'âge et de la présence d'enfants dans le ménage sur les patrons d'activités est bien reconnue dans la littérature (Pas, 1984). Par exemple, certains auteurs ont constaté que la présence d'un enfant au sein d'un ménage dont les deux parents travaillent a pour effet de diminuer le temps consacré aux activités faites au domicile de même que celles réalisées à l'extérieur du domicile autres que le travail (Bernardo, Paleti, Hoklas, & Bhat, 2015). Cette étude démontre également que la durée des activités autres que celles réalisées au domicile ou celles reliées au travail est plus faible lorsqu'il y a la présence d'un enfant dans le ménage.

Une autre étude a tenté de comprendre la différence qui peut exister entre les hommes et les femmes concernant la réalisation des activités (Sweet & Kanaroglou, 2016). Cette analyse se base entre autres sur l'effet de la réalisation d'une activité sur la qualité de vie d'une personne et révèle qu'il existe des différences entre les genres concernant cet indicateur. En effet, les femmes semblent plus enclines à réaliser des activités qui procurent une plus grande satisfaction lors de la participation à ces activités.

D'autres recherches ont également démontré que la différence entre les femmes et les hommes concernant la structure du système d'activités des individus se manifeste uniquement lorsqu'il y a la présence d'un enfant au sein du ménage. En effet, l'utilisation du temps par les femmes et les hommes est différente dans ce genre de ménage. De façon générale, les hommes réalisent des activités d'une durée plus élevée que les femmes et ce, pour les activités de type travail, loisir et magasinage (Allard & Janes, 2008).

Un autre élément important de la mobilité des ménages incluant des enfants concerne les habitudes de déplacement de ces derniers. Depuis quelques décennies, les enfants effectuent la majorité de leurs déplacements en automobile généralement grâce aux parents qui les reconduisent à l'école (Brownson & Boehmer, 2004). Conséquemment, les enfants jouissent d'une autonomie de moins en moins grande concernant leurs déplacements.

Finalement, l'âge d'une personne a également une influence importante sur les comportements de mobilité. En effet, le nombre de déplacements effectués par les personnes âgées est en baisse, particulièrement pour les personnes de 80 ans et plus (Alsnih & Hensher, 2003).

2.3.4 L'évolution des systèmes d'activités

La structure des systèmes d'activités des individus a connu d'importantes évolutions au cours des dernières décennies. Le projet de recherche précédent celui dans lequel s'inscrit ce mémoire a permis d'identifier des tendances dans la GRM grâce aux enquêtes Origine-Destination. Tout d'abord, la proportion de déplacements à motif travail dont le départ est avant 7h a augmenté de 1998 à 2008 (Morency & Trépanier, 2013). Cette hausse pourrait être attribuable à la congestion présente sur les réseaux de transport qui incite les navetteurs à quitter leur domicile plus tôt afin d'éviter la congestion. De plus, le temps moyen consacré aux activités de type travail ou

magasinage, incluant le temps d'accès au lieu d'activité, a augmenté de 1998 à 2008 pour la majorité des segments de la population.

Ces tendances ont également été confirmées par un autre volet de ce projet de recherche. Ce volet a permis de comparer les distributions des attributs des activités de 1998 à 2008 avec les mêmes données, soit celles des enquêtes Origine-Destination (Yasmin, 2016). L'évolution des fréquences, des heures de début et des durées d'activités a pu être analysée pour chaque type d'activités. Des distributions ont été construites pour chacun de ces trois attributs. Elles ont par la suite été comparées grâce au test de Kolmogorov-Smirnov, dont l'application a révélé que la majorité des distributions sont significativement différentes d'une année à l'autre. Ces différences confirment donc que la structure des systèmes d'activités évolue au fil du temps.

Bref, la structure des systèmes d'activités connaît une évolution importante depuis les dernières décennies et a une influence considérable sur les changements des comportements de mobilité. C'est pourquoi il est indispensable de considérer la transformation de ces structures dans les exercices de prévision de la demande.

CHAPITRE 3 MÉTHODOLOGIE

Le but ultime de la méthode proposée dans ce mémoire est d'attribuer une séquence d'activités à chaque individu composant une population entière, projetée dans le futur. Les différentes étapes du développement de cette méthode doivent permettre d'élaborer un processus de génération des activités qui est pertinent pour la modélisation des comportements de mobilité. Ce processus doit donc tenir compte de la structure du système d'activité des individus et des liens qui existent entre ces systèmes et les caractéristiques des personnes.

La création de ce processus doit également permettre l'intégration d'une composante temporelle dans le but de l'appliquer à une population projetée dans le futur. Cet ajout consiste tout d'abord à comprendre l'évolution des systèmes d'activités et par la suite à développer une méthode permettant de quantifier et modéliser cette évolution. Il sera ensuite possible de l'appliquer à une population projetée à différents horizons.

3.1 Méthodologie générale

L'ensemble des étapes menant à la création du processus de génération des activités est présenté à la Figure 3-1. Ce schéma décrit l'ensemble de la démarche poursuivie lors de la réalisation de ce mémoire.

La principale source de données provient des enquêtes Origine-Destination qui sont décrites à la section suivante. La typologie des ménages développée dans le cadre de ce projet de recherche est également intégrée aux données provenant des enquêtes. Les données de quatre enquêtes différentes sont rassemblées au sein d'un même jeu de données afin d'éventuellement analyser l'évolution temporelle des systèmes d'activités. Ce jeu de données doit subir plusieurs traitements afin de modéliser ces systèmes. Le principal traitement consiste à transformer les déplacements observés durant l'enquête en activités. Il devient ensuite possible de décrire la journée de chaque individu en fonction des activités réalisées. Cette description prend la forme d'un vecteur résumant la journée de chaque personne.

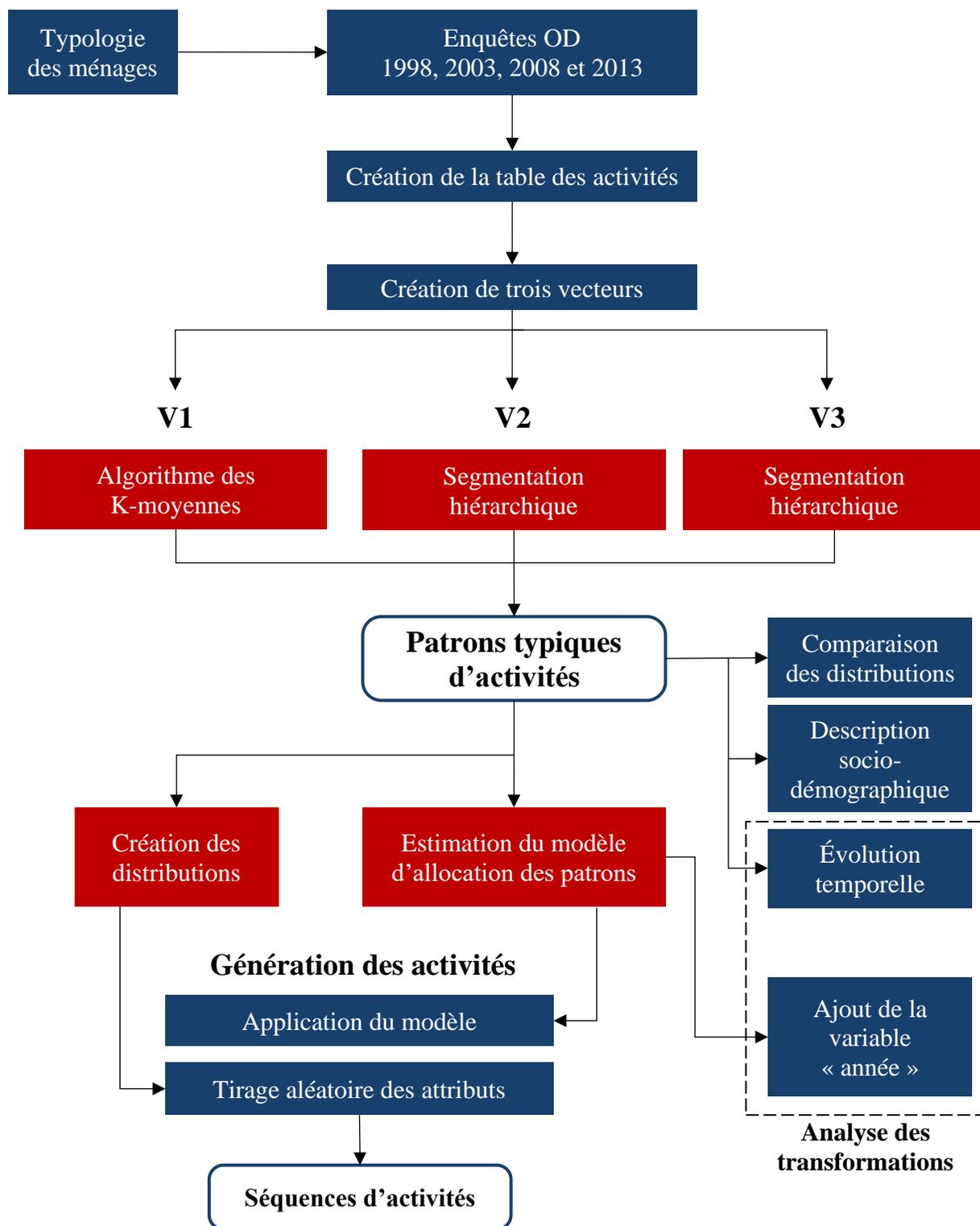


Figure 3-1 Schéma méthodologique

Trois différents vecteurs, nommés V1, V2 et V3, sont testés dans le cadre de ce mémoire. Ces vecteurs, présentés à la section 3.3.1, permettent de capter de diverses manières les multiples patrons typiques qui peuvent exister au sein de la population. Chacun d'entre eux est structuré d'une manière différente et possède des forces et des faiblesses quant à sa capacité à servir d'outil pour la génération des activités. Ces forces et faiblesses sont donc explorées tout au long de ce mémoire à travers le développement du processus de génération.

Grâce à la création de ces vecteurs, la première grande étape de ce mémoire, présentée au Chapitre 4, peut être réalisée. Cette étape concerne la segmentation de la population menant à la création des patrons typiques. Le choix de l'algorithme de segmentation repose sur la nature de chacun des vecteurs. Dans le cas de V1, le choix est porté sur l'algorithme des K-moyennes, alors qu'un algorithme de segmentation hiérarchique agglomératif est utilisé pour les vecteurs V2 et V3.

À la suite de la segmentation, diverses analyses peuvent être complétées. Tout d'abord, les distributions des attributs des activités, soit la fréquence, l'heure de départ et la durée, peuvent être construites pour chacun des groupes créés. Il est par la suite possible de les comparer pour s'assurer que les comportements sont significativement différents d'un groupe à l'autre. De plus, les individus présents dans chaque groupe peuvent être décrits selon leurs caractéristiques sociodémographiques. Aussi, puisque la segmentation est réalisée avec les quatre enquêtes, l'évolution de la proportion de personnes dans chacun des groupes peut également être réalisée.

La deuxième grande étape de ce projet, présenté au cinquième chapitre, consiste à estimer le modèle d'allocation des patrons aux individus. La variable expliquée de ce modèle devient donc la typologie précédemment développée, alors que les variables explicatives sont celles des caractéristiques sociodémographiques des individus.

Afin de générer des activités, des distributions similaires à celles permettant de comparer les comportements des individus composant les différents groupes peuvent être développées par suite de la segmentation. Grâce au modèle et à ces distributions, il devient possible de générer des activités à partir d'une certaine population. Il suffit d'appliquer le modèle aux membres de celle-ci, ce qui permet d'attribuer un patron typique à chaque individu la composant. Selon le type prédit, chaque individu se voit attribuer une série d'activités en pigeant aléatoirement dans les distributions créées à partir des individus appartenant au type en question.

Finalement, la modélisation de l'évolution des patrons est essentiellement réalisée avec l'ajout d'une variable au modèle d'allocation, soit le nombre d'années qui s'écoulent depuis la première enquête étudiée (1998). Il devient ainsi possible d'estimer un modèle avec les enquêtes de 1998, 2003, 2008 et 2013 et d'éventuellement l'appliquer à une population générée pour les années subséquentes (projections).

3.2 Description des données

Les enquêtes Origine-Destination réalisées dans la Grande Région de Montréal de même que les données récoltées lors de celles-ci sont décrites dans la section suivante. De plus, la typologie des ménages développée précédemment dans le cadre d'un projet de recherche antérieur est également présentée.

3.2.1 L'enquête Origine-Destination

À Montréal, les enquêtes Origine-Destination sont effectuées de façon périodique depuis 1970 et ce à un intervalle d'environ cinq ans (MTQ, 2017a). Ces enquêtes visent à récolter les données relatives aux déplacements réalisés par des milliers de ménages, sélectionnés afin de représenter l'ensemble des ménages présents dans la région métropolitaine de Montréal. Les informations sont généralement récoltées au moyen d'une entrevue téléphonique, dont le répondant est un membre du ménage enquêté. Ce répondant doit donc transmettre les informations relatives à ses déplacements, mais également celles relatives aux déplacements des autres membres du ménage. Les déplacements déclarés sont ceux réalisés la journée ouvrable précédant le jour de l'entrevue. De plus, les caractéristiques du ménage et des membres de celui-ci sont recueillies.

L'enquête se déroule habituellement du mois de septembre au mois de décembre. Ainsi, les déplacements observés dans le cadre de ces enquêtes permettent de représenter le portrait d'une journée typique de la semaine dans le Grand Montréal, à un moment où les réseaux de transport sont typiquement les plus utilisés, soit durant l'automne.

La Figure 3-2 présente le territoire de l'enquête OD de 2013 qui est divisé en huit régions. Ces régions sont également subdivisées en secteurs municipaux, dont la subdivision, qui est propre au besoin de l'enquête, est inspiré des secteurs de recensement canadien et des divisions municipales dans le Grand Montréal, soit les limites des villes ou municipalités, des municipalités régionales

de comtés ou agglomérations, et des arrondissements. Les secteurs municipaux sont également compatibles avec les aires de diffusion du recensement.

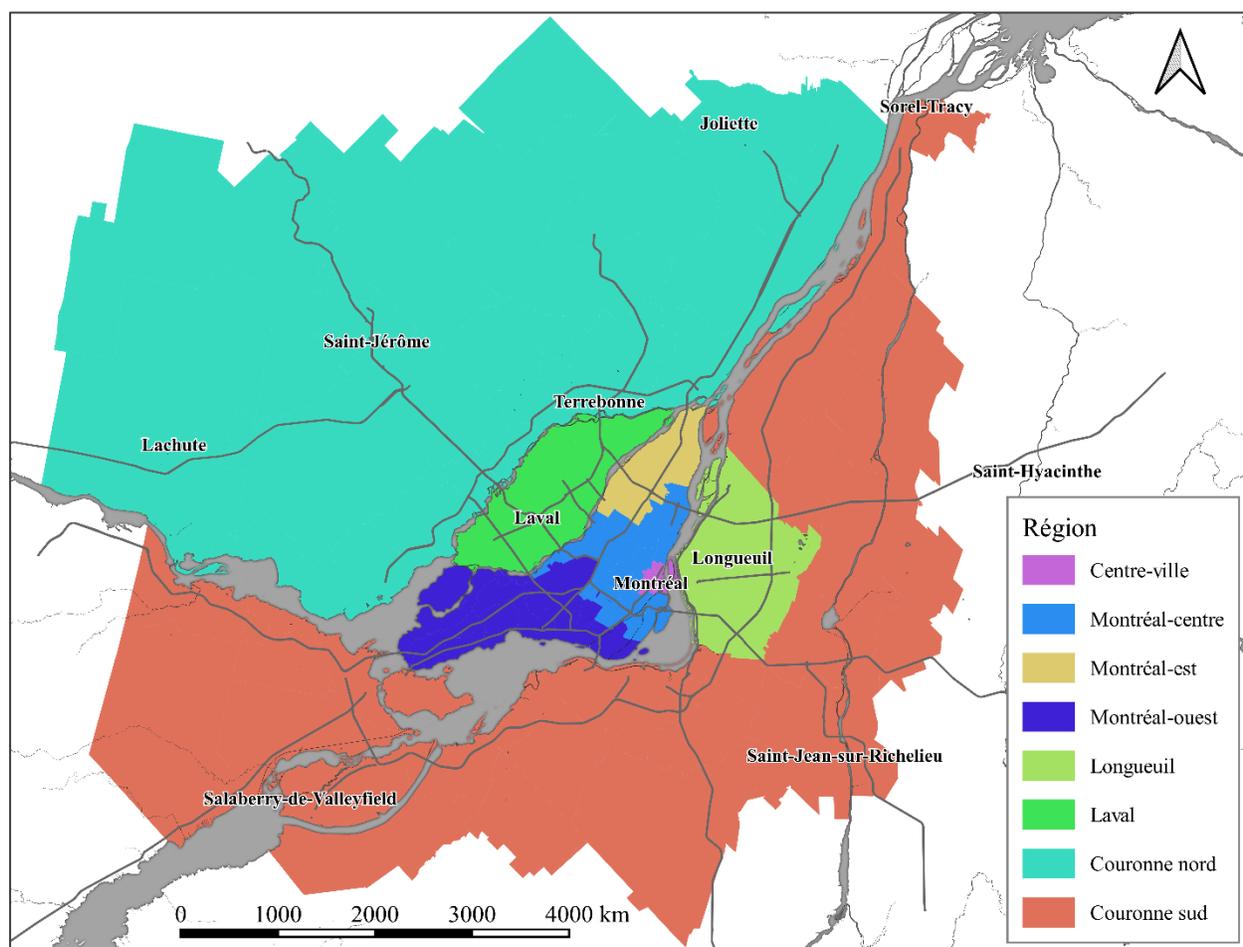


Figure 3-2 Le territoire de l'enquête de 2013 et les huit régions qui le composent

Le territoire couvert par l'enquête OD 2013 a significativement augmenté depuis 1998. Le Tableau 3-1 montre les informations critiques décrivant chaque enquête.

Tableau 3-1 Caractéristiques des enquêtes de 1998 à 2013

Année de l'enquête	1998	2003	2008	2013	Total
Nombre de ménages (échantillon)	65 227	58 000	66 124	78 731	268 082
Nombre de personnes (échantillon)	164 075	139 527	156 720	188 746	649 068
Taux d'échantillonnage (personne - %)	4,7	3,8	4	4,4	-
Nombre de déplacements (échantillon)	384 945	306 558	319 949	410 741	1 422 193
Nombre d'activités (échantillon)	215 846	169 781	176 860	233 609	796 096

Les données issues de l'enquête réalisée en 2003 ne sont que partielles. En effet, la réalisation de l'enquête s'est poursuivie lors du mois de janvier puisque le taux d'échantillon n'était pas assez élevé au moment où l'enquête se termine habituellement, soit au mois de décembre. Dans le cadre de ce mémoire, l'objectif est de comparer les comportements observés dans les enquêtes afin d'en dégager des tendances, c'est pourquoi seules les données récoltées lors de la période de collecte habituelle sont utilisées.

Un nombre élevé d'informations est récolté lors de la réalisation des entrevues téléphoniques. Les autoroutes empruntées ou les lignes de métro utilisées sont déclarées, la possession d'une passe mensuelle de transport en commun est observée, la langue dans laquelle l'entrevue se déroule est enregistrée. Ces informations sont quelques exemples de la panoplie de données auxquelles il est possible d'accéder. Dans le cadre de ce mémoire, seule une partie de ces informations est utilisée. La structure de ces informations est présentée à la Figure 3-3, qui énonce également d'autres informations de base qui permettent de caractériser les données récoltées durant l'enquête.

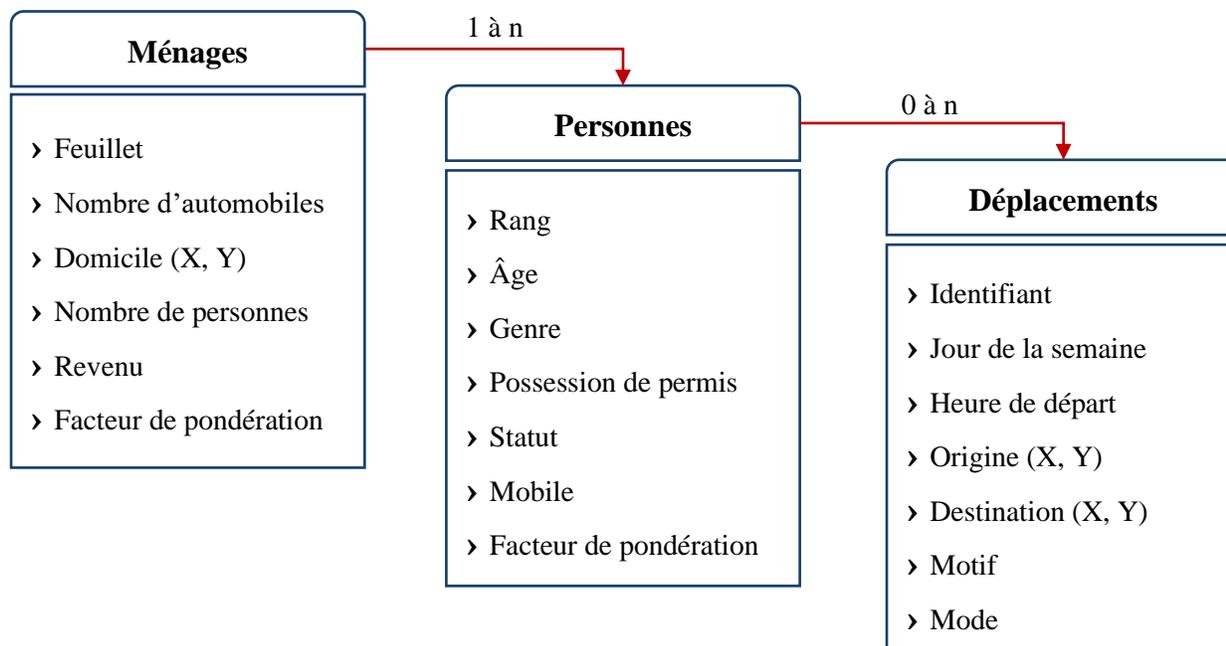


Figure 3-3 Structure des données de l'enquête OD

L'unité de base de l'enquête OD est le ménage, qui est composé d'au minimum une personne. Le feuillet permet d'identifier chaque ménage à travers l'ensemble des enquêtes¹ alors que le rang est propre à chaque individu à l'intérieur d'un ménage. Les individus composant les ménages peuvent être soit immobiles, en ne faisant aucun déplacement, ou faire un déplacement ou plus. Les informations concernant chacun de ces déplacements comprennent entre autres le motif ainsi que l'heure de départ. L'heure d'arrivée n'est pas collectée, à l'exception des déplacements du répondant qui transmet les informations relatives à l'ensemble du ménage auquel ils appartiennent. La durée de chaque déplacement n'est donc pas disponible.

Les caractéristiques des données utilisées dans le cadre de ce mémoire sont présentées dans le Tableau 3-2, en indiquant à quel objet les variables sont reliées, soit le ménage, la personne ou le déplacement.

¹ Les numéros de feuillet sont uniques pour les enquêtes de 2003, 2008 et 2013, mais certains numéros de feuillet dans l'enquête de 1998 sont présents dans celle de 2003. De nouveaux numéros ont été générés pour ces cas, numéros qui n'existent pas dans aucune autre enquête.

Tableau 3-2 Données utilisées dans le cadre de ce mémoire

Variables relatives au ménage		
Feuillet	Identifiant unique	Identifiant unique de chaque ménage à travers les enquêtes
Type de ménage	Catégorielle	Typologie de ménages – 8 catégories : définies à la section suivante
Domicile	Numérique	Coordonnées (X, Y) du domicile où réside le ménage
Variables relatives aux personnes		
Rang	Identifiant	Identifiant d'une personne au sein d'un ménage
Âge	Entier	Âge de chaque personne – 0 à 99 ans
Genre	Catégorielle	Genre de chaque personne – Homme ou femme
Statut	Catégorielle	Statut d'occupation de la personne – 8 catégories : <ul style="list-style-type: none"> › Travailleur à temps plein › Travailleurs à temps partiel › Étudiant / élève › Retraité › Autre › N/A : enfant de 4 ans et moins › À la maison › Refus
Facteur de pondération	Numérique	Poids de chaque personne à l'intérieur de chaque enquête – 1,7 à 261,19 (moyenne de 23,65)
Mobile	Catégorielle	Variable indiquant si la personne s'est déplacée – Mobile ou immobile
Variables relatives aux déplacements		
ID	Identifiant	Identifiant d'un déplacement pour une personne
Motif	Catégorielle	Motif détaillé du déplacement – 14 catégories : définies à la section suivante
Heure de départ	Entier	Heure à laquelle le déplacement a débuté – 0 à 28 heures (déclaré à la minute près)

Le type de ménage provient d'une autre source de données additionnelle à celle de l'enquête qui est présentée à la section 3.2.3. De plus, les coordonnées du domicile de chacun des ménages sont également disponibles. Grâce à cette information, il est possible de déterminer dans quelle région réside chaque ménage parmi les huit divisant le territoire de la GRM, illustrées à la Figure 3-2.

Concernant les personnes, outre l'identifiant, cinq variables sont utilisées dans cette étude. Trois d'entre elles permettent de décrire un individu, soit l'âge, le genre et le statut d'occupation. Huit catégories sont utilisées pour définir le statut d'une personne.

La quatrième variable, le facteur de pondération, est essentielle lors de l'utilisation des données provenant de l'enquête. Ces facteurs, qui sont au cœur de la démarche du MTQ permettant de projeter les déplacements, indique le poids de chaque personne, et par conséquent le poids des déplacements réalisés par celle-ci, à l'intérieur de l'échantillon. Ainsi, afin d'être en mesure de représenter l'ensemble de la population du Grand Montréal, chaque individu à l'intérieur des ménages échantillonnés se voit attribuer un poids en fonction de ses caractéristiques sociodémographiques. Ces poids permettent donc d'indiquer le nombre de personnes dans la population réelle qu'une personne échantillonnée représente. La prise en compte des caractéristiques permet donc de s'assurer que certains segments de la population ne sont pas surreprésentés ou sous-représentés. Pour chaque enquête, la population réelle est obtenue grâce aux données les plus récentes concernant les populations estimées par l'ISQ. Ces données sont estimées à partir du recensement canadien le plus récent, qui a généralement lieu deux ans avant la réalisation de l'enquête.

En prenant en compte ce facteur, il est par conséquent possible d'effectuer des calculs à partir des individus échantillonnés dont les constats s'appliquent à l'ensemble de la population réelle. Selon la même logique, des facteurs sont également déterminés pour les ménages, mais ils ne sont pas utilisés dans le cadre de ce mémoire.

La dernière variable relative à la personne est celle qui indique si une personne s'est déplacée ou si elle est demeurée à son domicile. Elle permet donc d'éliminer les personnes n'ayant réalisé aucun déplacement et donc aucune activité hors-domicile.

Pour ce qui est des variables relatives aux déplacements, deux variables sont utiles à ce projet de recherche. Tout d'abord, l'heure à laquelle chaque déplacement débute est déclarée et ce à la minute près. De plus, l'heure de départ de l'ensemble des déplacements observés durant les quatre enquêtes est entre minuit et 28 heures, soit l'équivalent de quatre heures du matin la journée qui suit celle où le ménage est enquêté. La journée typique qui est observée pour chaque personne est donc d'une durée de 28 heures, soit de minuit à quatre heures du matin.

Les motifs détaillés de déplacement sont composés de 14 catégories et sont présentés à la section suivante, tout comme la transformation appliquée à ces catégories pour les besoins de ce mémoire et la transformation appliquée à l'heure de départ des déplacements.

3.2.2 Traitement et transformation des données

Afin d'utiliser les données provenant de l'enquête OD dans le but de modéliser les systèmes d'activités, certains traitements sont nécessaires. Tout d'abord, tous les enfants ayant quatre ans et moins sont exclus de toutes les analyses faites dans le présent mémoire puisque leur mobilité est considérée comme non observée. Alors qu'un total de 649 068 personnes sont enquêtées en 1998, 2003, 2008 et 2013, 613 997 personnes sont sélectionnées à la suite de l'application de ce filtre. De plus, parmi ces individus, certains d'entre eux n'ont réalisé aucun déplacement visant à réaliser une activité lors de la journée où ils ont été sondés. En éliminant ces personnes, la base de données à partir de laquelle les transformations suivantes sont effectuées compte 504 776 individus.

Les informations contenues dans la base de données issue de l'enquête OD peuvent être transformées afin de passer d'une table des déplacements à une table des activités. La Figure 3-4 permet d'illustrer cette transformation.

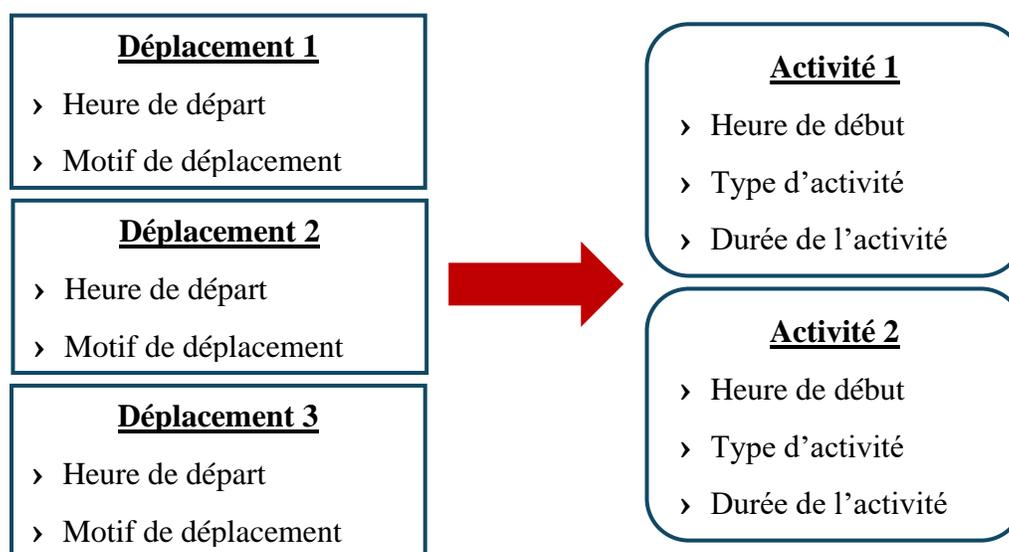


Figure 3-4 Transformation des données des déplacements

À partir de l'heure de départ et du motif d'un déplacement, il est possible d'en dériver les informations permettant de décrire le système d'activités d'un individu. Le motif de déplacement permet de déduire le type d'activité et l'heure de départ permet de déterminer l'heure à laquelle l'activité débute. De plus, la durée de chacune des activités peut être déterminée grâce à l'heure de départ du déplacement suivant. Comme mentionné à la section précédente, la durée des déplacements n'est pas connue. L'heure de début d'une activité correspond donc à l'heure de départ

du déplacement pour accéder au lieu d'activité et la durée de celle-ci inclut la durée de ce déplacement. Une fois l'ensemble de ces informations colligées pour chaque individu des ménages enquêtés lors des enquêtes de 1998, 2003, 2008 et 2013, il est possible d'énumérer l'ensemble des activités réalisées par chaque personne et ainsi décrire le système d'activités de chacune d'entre elles.

De plus, les motifs déclarés lors des enquêtes de 1998 à 2013 sont composés de 14 catégories, mais une partie seulement est utilisée dans le cadre de ce mémoire. La transformation de ces catégories est présentée dans le Tableau 3-3.

Tableau 3-3 Liste des motifs et des types d'activités

Motif détaillé	Motif agrégé	Type d'activité
Travail Rendez-vous d'affaires Sur la route	Travail	<i>Travail</i>
Étude/École	Étude	<i>Étude</i>
Loisir	Loisir	<i>Loisir</i>
Magasinage	Magasinage	<i>Magasinage</i>
Visites d'ami(e)s et/ou de la parenté Santé Reconduire quelqu'un Chercher quelqu'un Autre	Autre	<i>Autre</i>
Indéterminé/refus/NSP	Indéterminé	<i>Autre</i>
Retour au domicile	Retour au domicile	-

Les motifs détaillés correspondent à ceux déclarés par les ménages interrogés lors de la réalisation de l'enquête OD. Lors du traitement préliminaire des données par l'équipe chargée de mener l'enquête, les motifs détaillés sont agrégés en sept catégories. Dans le cadre de ce mémoire, le traitement des données s'est fait à partir de ces motifs agrégés. C'est donc sur la base de ces sept catégories que les types d'activité sont dérivés des motifs de déplacements, soit en éliminant le motif « retour au domicile » et en agrégeant les motifs « autre » et « indéterminé » pour ainsi former le type d'activité *autre*.

À la suite de l'ensemble de ces transformations, les données récoltées lors de l'enquête OD peuvent ensuite être utilisées pour modéliser les systèmes d'activités et comprendre les liens entre ces

systèmes et les caractéristiques des personnes et des ménages. Afin de comprendre ces liens, une autre source de données est nécessaire.

3.2.3 Typologie des ménages

La source de données présentée dans cette section correspond à la typologie des ménages développée dans le cadre des travaux de Fabre (2018). Comme mentionné au Chapitre 2, la composition d'un ménage a une grande influence sur les activités réalisées par un individu et la structure temporelle de ses activités. Catégoriser les ménages permet donc de tenir compte de cette influence en identifiant les types de ménages qui permettent d'observer des comportements de mobilité significativement différents d'un individu à l'autre. Le développement de cette typologie a permis d'obtenir huit types de ménages pertinents pour la prévision de la demande en transport. La Figure 3-5 permet d'observer les principales différences qui subsistent entre ces types, soit les jeunes couples, colocataires et célibataires, les très jeunes familles, les jeunes familles, les familles traditionnelles, les adultes et personnes âgées, les cinquantenaires, les jeunes retraités, les vieux retraités et les familles multigénérationnelles.

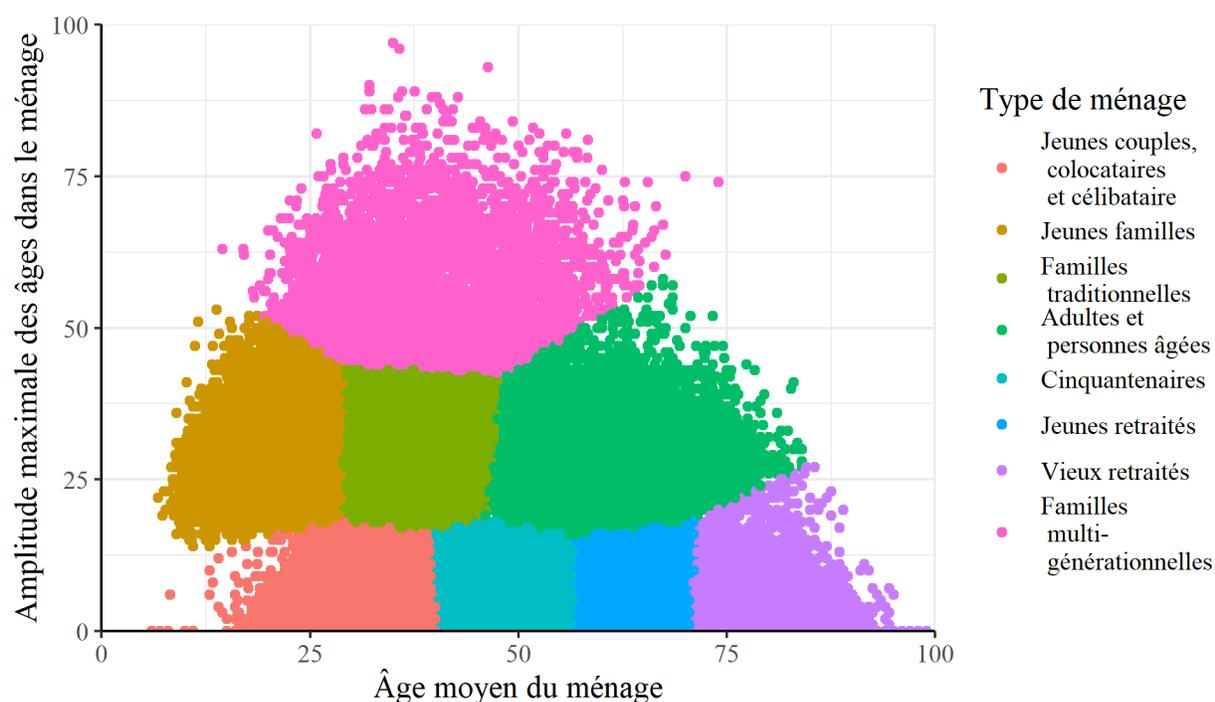


Figure 3-5 Typologie des ménages

Les principales caractéristiques qui distinguent les huit types sont l'âge moyen du ménage ainsi que l'amplitude maximale des âges. Le nombre de personnes, le nombre d'enfants de moins de 12

ans et le nombre de personnes âgées de 76 à 84 ans ont également été des variables utilisées lors de l'application de l'algorithme des K-moyennes qui a mené à la création de la typologie. Le Tableau 3-4 présente les centres des huit groupes, soit la valeur moyenne de ces cinq variables pour chaque type de ménage, alors que le Tableau 3-5 présente la proportion de ménages selon le type.

Tableau 3-4 Caractéristique des huit types de ménages

Type de ménage	Nombre moyen de personnes	Âge moyen	Amplitude des âges maximums	Nombre moyen d'enfants de moins de 12 ans	Nombre moyen de personnes entre 76 et 84 ans
Jeunes couples, colocataires et célibataire	1.6	30.7	2.1	0.0	0.0
Jeunes familles	3.9	22.8	32.7	1.3	0.0
Familles traditionnelles	3.2	37.4	32.7	0.0	0.0
Adultes et personnes âgées	2.6	59.3	30.1	0.0	0.2
Cinquantennaires	1.6	51.2	2.2	0.0	0.0
Jeunes retraités	1.6	65.1	2.2	0.0	0.0
Vieux retraités	1.4	79.3	1.5	0.0	0.6
Familles multi-générationnelles	4.1	38.9	54.2	0.7	0.1

Tableau 3-5 Proportion de ménages pour chaque catégorie

Type de ménage	Proportion de ménages (%)
Jeunes couples, colocataires et célibataire	14%
Jeunes familles	23%
Familles traditionnelles	13%
Adultes et personnes âgées	3%
Cinquantennaires	16%
Jeunes retraités	18%
Vieux retraités	9%
Familles multigénérationnelles	3%

Avec les données provenant des enquêtes OD de 1998 à 2013, une typologie différente a été créée pour chacune des enquêtes. L'analyse de ces typologies a permis de constater que certains types de ménages sont présents pour certaines enquêtes, mais disparaissent au profit d'un autre dans les enquêtes subséquentes. Dans un contexte prévisionnel, la typologie retenue est celle créée avec les données de l'ensemble des enquêtes regroupées au sein d'un seul jeu de données. Il est ainsi possible dans le cadre de ce mémoire d'utiliser une typologie unique à travers l'ensemble des enquêtes. Plus de détails sur la création de cette typologie sont disponibles dans le mémoire de Fabre (2018).

Le premier vecteur (**V1**) est composé de la durée totale, qui est indiquée en minutes, et de la fréquence de réalisation de chaque type d'activité. Ainsi, pour chaque individu enquêté, deux colonnes sont créées pour chacun des cinq types, soit la somme des durées de l'ensemble des activités du type en question et le nombre de fois que ce type d'activité a été réalisé lors de la journée où les informations ont été récoltées. Le vecteur est donc composé de dix variables, qui peuvent par la suite être transformées et normalisées préalablement à la segmentation. Tout d'abord, les dix variables X_i , i variant de 1 à 10, sont transformées selon la formule suivante où la fonction logarithmique est appliquée :

$$X_{i,trans} = \log(X_i + 1)$$

Cette transformation permet de modifier la distribution de certaines variables qui peuvent être asymétriques. Puisque les durées et les fréquences peuvent être nulles, la valeur de 1 est ajoutée à ces variables afin d'être en mesure d'appliquer la fonction logarithmique. Par la suite, la normalisation des données peut se faire avec la formule suivante :

$$X_{i,trans,norm} = \frac{X_{i,trans} - \min_i X_{i,trans}}{\max_i X_{i,trans} - \min_i X_{i,trans}}$$

Ainsi, les dix variables sont ramenées à des valeurs entre 0 et 1 selon le minimum et le maximum de chacune d'entre elles. Elles peuvent donc être comparées entre elles selon des échelles similaires et l'algorithme de segmentation peut être appliqué (Mohamad & Usman, 2013).

Le deuxième vecteur (**V2**) est composé de variables binaires indiquant si une activité est réalisée ou non. Pour chacun des cinq types d'activité, 56 variables sont créées contenant une valeur de 1 ou 0. Ainsi, la journée de chaque personne est divisée en 56 blocs représentant un intervalle de 30 minutes chacune, pour un total de 28 heures. Le vecteur est ainsi composé de 280 variables binaires qui nécessitent aucune transformation en vue de la segmentation.

Le troisième vecteur (**V3**) est composé de variables catégorielles. Il nécessite également une division de la journée de chaque individu par intervalle de 30 minutes. Le vecteur se résume donc à 56 variables qui, selon l'activité qui est réalisée durant chaque bloc de 30 minutes, peuvent prendre les valeurs de T, E, L, M, A et D. Ces lettres signifient *travail*, *études*, *loisir*, *magasinage*, *autres* et « domicile ». Ainsi, chaque bloc peut prendre la valeur d'un des types d'activité, mais peut également prendre la valeur D, qui indique que la personne est à son domicile. Toutefois,

certains individus présents dans l'échantillon des enquêtes ont fait plus d'une activité lors d'un des blocs de 30 minutes. Le type d'activité sélectionné dans ce bloc correspond au type de l'activité dont la durée est la plus longue à l'intérieur du bloc en question.

3.3.2 Segmentation avec l'algorithme de K-moyennes

L'algorithme de segmentation utilisé avec V1 est l'algorithme des K-moyennes, qui permet de partitionner des vecteurs selon la distance qui les sépare en créant plusieurs groupes contenant des vecteurs similaires. Le principe de base de l'algorithme est de maximiser l'homogénéité à l'intérieur de chacun des groupes et de maximiser l'hétérogénéité entre ceux-ci. L'homogénéité à l'intérieur d'un groupe peut être exprimée par la variabilité intragroupe et elle peut être évaluée par la somme des distances au carré entre chaque observation composant le groupe et la moyenne du groupe en question (Steinley, 2006). Pour un vecteur de dimension P et un groupe contenant n observations, cette somme peut être calculée grâce à la formule suivante :

$$\sum_{i=1}^n \sum_{j=1}^P (x_{ij} - \bar{x}_j)^2$$

Où x_{ij} est la valeur de la dimension j et de l'observation i , tandis que \bar{x}_j est la moyenne de la dimension j dans le groupe dont la variabilité est évaluée. La moyenne peut être calculée par la formule suivante :

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$$

Ainsi, la somme des distances au carré doit être minimisée afin de regrouper des observations qui ont une grande similarité. Les sommes calculées pour chacun des groupes peuvent être additionnées pour évaluer la variabilité intragroupe totale, soit pour l'ensemble des groupes. L'objectif du partitionnement réside dans la minimisation de cette somme pour ainsi créer des groupes contenant des vecteurs dont la valeur des variables varie le moins possible. Un algorithme est utilisé pour minimiser cette somme et il est appliqué avec le langage R, grâce à la fonction *kmeans*.

Plusieurs types d'algorithmes existent et le choix est porté sur celui de McQueen, qui peut se résumer comme suit :

1. Initialisation des k centres de façon aléatoire
2. Itérer tant que la solution n'est pas stable :
 - › Création de groupes en assignant tous les vecteurs au centre le plus proche
 - › Calcul du centre de chaque groupe
 - › Si aucun centre n'est modifié, la solution est considérée comme stable sinon, itérer de nouveau

La solution n'est jamais complètement stable, c'est pourquoi un nombre d'itérations maximum peut être indiqué lors de l'appel de la fonction *kmeans*. De plus, la nature aléatoire de l'initialisation de l'algorithme implique que les groupes issus du partitionnement des données ne sont pas uniques (Saunier, 2017). Autrement dit, les groupes créés peuvent être différents selon les centres initiaux choisis, ce qui peut mener à la possibilité d'obtenir un optimum local et non un optimum global (Steinley, 2006). En effet, les groupes obtenus à la suite de la segmentation pourraient ne pas être optimaux, c'est pourquoi l'application de l'algorithme peut être réalisée plusieurs fois afin de s'assurer d'une certaine stabilité et ainsi obtenir un optimum global.

Ces problèmes peuvent également être atténués par un choix non aléatoire du centre des groupes. En effet, le choix initial peut être fait de façon judicieuse en choisissant de meilleurs points de départ.

La métrique permettant de mesurer la distance entre chaque vecteur et le centre d'un groupe lors de l'exécution de la méthode est la distance euclidienne. Dans le cas de V1, la distance D entre deux observations i et j peut être calculée grâce à la formule suivante :

$$D = \sqrt{(Durée_{Travail,i} - Durée_{Travail,j})^2 + (Fréq_{Travail,i} - Fréq_{Travail,j})^2 + \dots}$$

Grâce à la transformation et à la normalisation des variables décrites précédemment, l'importance des dix variables contenues dans V1 est du même ordre de grandeur. La distance euclidienne est donc insensible aux différences d'amplitude qui peuvent exister entre les variables contenues dans le vecteur. De plus, les groupes créés suite à l'application de l'algorithme ne sont pas caractérisés par les variables ayant une importance accrue avant la normalisation, mais bien par l'ensemble d'entre elles.

3.3.3 Segmentation hiérarchique

La méthode de partitionnement utilisée avec les deux derniers vecteurs est un algorithme de segmentation hiérarchique. Ce type d'algorithme nécessite la création d'une matrice de distances préalablement à son application. Cette matrice contient la distance entre chaque vecteur destiné à être regroupé. Ainsi, pour n vecteur, la matrice a une dimension de $n \times n$. Dans le cas de V2 et V3, 504 776 individus, ou vecteurs, sont sélectionnés pour l'exécution de cet algorithme, ce qui porte à $504\,776 \times 504\,776$ la dimension d'une matrice. Les fonctions utilisées pour créer la matrice et appliquer l'algorithme dans le langage R n'ont pas la capacité de traiter autant d'information. C'est pourquoi environ 3 % des personnes, soit 14 965 vecteurs, sont sélectionnées pour créer les groupes. Cette sélection se fait en respectant les proportions observées dans le jeu de données original. Ces proportions sont déterminées selon le genre, l'âge ainsi que l'année de l'enquête à laquelle chaque individu appartient. Le sous-échantillon sélectionné, de même que les groupes et constats qui en découlent est donc, dans une certaine mesure, représentatif de la population selon les attributs sociodémographiques, mais également selon l'année de l'enquête, ce qui est indispensable dans un contexte où l'objectif de la typologie est de capturer l'évolution à travers les années des comportements en matière des systèmes d'activités. Afin de s'assurer de la stabilité des résultats, l'algorithme est appliqué plusieurs fois, avec différents échantillons choisis aléatoirement selon les proportions définies préalablement.

Afin de créer la matrice des distances entre individus, deux types de distances sont utilisées selon la nature des variables contenues dans V2 et V3.

3.3.3.1 Distance de Hamming

La distance de Hamming est utilisée lorsque les vecteurs sont composés d'une série de variables binaires. Elle permet donc de calculer la distance entre des vecteurs composés de variables binaires tels que V2. La formule ci-dessous indique comment la distance D est calculée entre les individus i et j .

$$D = \sum_{k=1}^n \delta_k, \text{ avec } \begin{cases} \delta_k = 1, \text{ si } heure1_{travail,i} = heure1_{travail,j} \\ \delta_k = 0 \text{ sinon} \end{cases}$$

Où n correspond à la variable binaire indiquant si une activité est réalisée ou pas. Le calcul de la distance peut donc se résumer à l'addition du nombre de variables ayant la même valeur, donc au nombre de blocs de trente minutes où deux individus pratiquent la même activité.

3.3.3.2 Distance de Damerau-Levenshtein

La distance de Damerau-Levenshtein s'applique à des vecteurs composés de variables catégorielles tels que V3. Similaire à la SAM, mentionnée à la section 2.3.2.1, cette distance correspond à la somme du coût des différentes opérations nécessaires à l'alignement de deux séquences, soit une série de variables catégorielles. Les opérations, au nombre de quatre, consistent au retrait ou à l'insertion d'une variable, ou à la substitution ou la transposition de deux variables. Ces opérations sont illustrées à la Figure 3-6.

Insertion					Substitution				
D	D	D	T	T	D	D	T	T	D
D	D	T	T	T	D	D	T	T	T
Retrait					Transposition				
D	D	T	T	D	D	D	T	M	T
D	T	T	T	D	D	D	T	T	M

Figure 3-6 Opérations possibles pour le calcul de la distance de Damerau-Levenshtein

Un poids peut également être attribué à chacune de ces opérations pour ainsi accorder une plus grande importance ou une moins grande importance à certaines d'entre elles.

3.3.3.3 Application de l'algorithme

À la suite de la création de la matrice des distances, les individus, représentés par des vecteurs, peuvent être regroupés ensemble. Plusieurs types de segmentation hiérarchique existent. Celui choisi pour les vecteurs V2 et V3 est de type agglomératif. Ce type d'algorithme implique une agglomération graduelle des observations tout au long de son application (Jain, Murty, & Flynn, 1999). Cela signifie qu'à l'initialisation de la procédure, chaque observation forme un groupe et, de façon successive, celles-ci sont agglomérées dans le but de permettre la formation des groupes. Ainsi, à chaque itération, deux groupes, qui peuvent contenir une seule observation, sont réunis pour n'en former qu'un et, à la fin de la procédure, l'ensemble des observations ne forme plus

qu'un seul groupe. C'est de cette façon qu'un dendrogramme est créé, soit un arbre illustrant le niveau de dissimilarité entre chaque groupe. L'arbre, dont la base est constituée de l'ensemble des observations, permet de visualiser la manière dont les groupes formés à chaque itération sont imbriqués l'un dans l'autre, pour éventuellement ne former qu'un seul groupe au sommet.

Le dendrogramme peut par la suite être coupé pour ainsi récupérer les groupes qui se sont formés à une certaine itération. La coupure se situe au niveau de dissimilarité désiré, qui correspond à la distance entre les groupes.

Le regroupement de deux observations se fait sur la base de la distance qui les sépare, disponible dans la matrice préalablement calculée. Cependant, l'agglomération d'un groupe et d'une observation, ou de deux groupes composés de plusieurs observations, peut se faire de plusieurs façons. Tout d'abord, il est possible de calculer la distance entre deux groupes grâce à trois métriques différentes, soit la distance maximale, la distance minimale et la distance moyenne. Outre ces trois distances, d'autres méthodes existent afin d'agglomérer des groupes à chaque itération. La méthode utilisée dans ce mémoire a été développée par Ward qui, plutôt que de sélectionner les deux groupes à rassembler sur la base de la distance qui les sépare, sélectionne les deux groupes dont le regroupement fait le moins augmenter la somme des variances intragroupe, soit la variance à l'intérieur d'un groupe (Ward Jr, 1963). Le choix de la méthode s'est arrêté sur celle de Ward puisque les groupes issus de la segmentation utilisant les autres méthodes n'étaient pas utilisables dans le cadre de ce mémoire. En effet, l'utilisation des autres méthodes menait à l'obtention de groupes composés de moins de 0,5 % d'individus. La fonction *agnes* disponible dans le langage R est utilisée pour l'application de l'algorithme.

3.4 Distributions des attributs des activités

À la suite de la segmentation de la population, chaque personne se voit attribuer un groupe dont les membres ont des systèmes d'activités similaires. À l'aide de cette variable additionnelle et des données disponibles dans l'enquête OD, il est possible de générer des distributions pour les attributs des activités, soit la fréquence, l'heure de début et la durée. Ainsi, pour chacun des groupes issus de la segmentation, la distribution de ces trois attributs peut être réalisée. Elles peuvent ensuite être utilisées pour comparer les groupes entre eux de même que générer des activités.

3.4.1 Comparaison des distributions

En générant des distributions pour chacun des groupes, il est possible de comparer les comportements des individus qui les composent afin de s'assurer que ces comportements sont significativement différents d'un groupe à l'autre. Il s'agit donc d'appliquer des tests statistiques sur les distributions afin de s'en assurer.

3.4.1.1 Construction des distributions

Les distributions créées pour fin de comparaison sont construites tel qu'illustré sur la Figure 3-7, grâce aux groupes créés lors de l'étape de la segmentation.

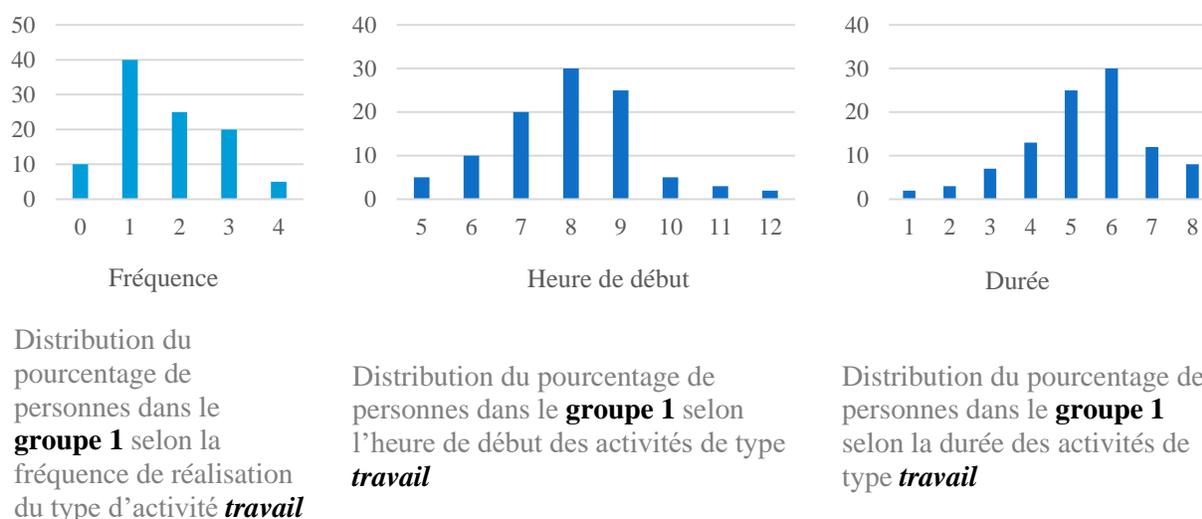


Figure 3-7 Exemple de distribution à des fins de comparaison

Ainsi, pour chacun des groupes et pour chaque type d'activité, la distribution du nombre de personnes selon la fréquence de réalisation des activités lors de la journée où chaque personne est sondée est créée. Pour chaque groupe et chaque activité également, la distribution du nombre de personnes selon l'heure de début d'une activité est réalisée et ce, peu importe le nombre d'activités de chaque type fait par ces personnes. Les heures de début sont regroupées en tranche de 60 minutes, tout comme les durées d'activités, dont la distribution est également créée pour chaque groupe et chaque type d'activité, mais sans tenir compte du nombre d'activités et de l'heure de début des activités réalisées par les personnes.

3.4.1.2 Tests de Kolmogorov-Smirnov

Le test choisi afin de comparer les distributions est celui de Kolmogorov-Smirnov. Ce test permet d'évaluer si une distribution empirique d'une variable aléatoire X a la même fonction de répartition qu'une loi théorique (Desgraupes, 2014). Ce test se base sur la distance maximale entre la fonction de répartition de la distribution empirique et celle de la loi théorique. La formule suivante permet de déterminer cette distance :

$$D_{KS}(F_i, F_j) = \max (F_i(X) - F_j(X))$$

Où F_i est la fonction de répartition de la distribution empirique et F_j est la distribution de la loi théorique. La Figure 3-8 permet de visualiser cette distance.

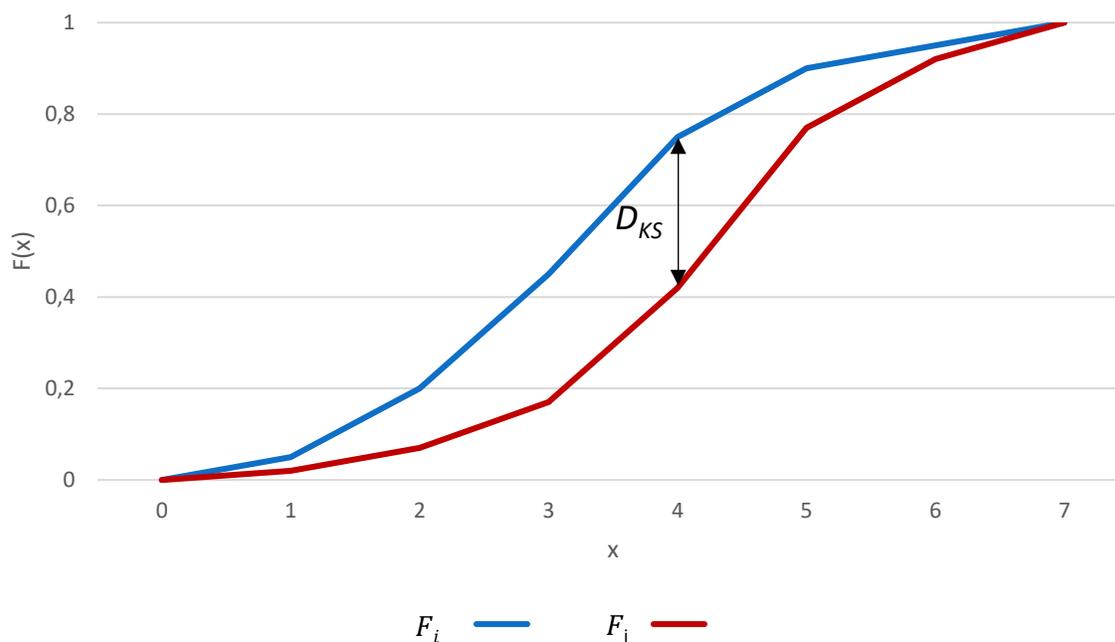


Figure 3-8 Identification de la distance maximale entre deux distributions

Grâce à cette distance, il est possible d'effectuer le test, dont l'hypothèse nulle est que la fonction de répartition empirique est la même que celle théorique. Une faible distance permet ultimement de rejeter l'hypothèse alternative afin d'accepter l'hypothèse nulle et conclure que les deux distributions sont sensiblement identiques.

Ce test peut toutefois être utilisé avec deux distributions empiriques. Ainsi, la fonction de répartition de la loi théorique peut être remplacée par celle d'une autre distribution empirique. Le test permet donc de vérifier si les deux distributions ont les mêmes fonctions de répartition.

3.4.2 Utilisation des distributions afin de générer les activités

Les distributions construites pour la génération des activités, qui sont légèrement différentes de celles utilisées pour des fins de comparaison, peuvent être utilisées pour attribuer un programme d'activités à un individu selon le groupe auquel il appartient. Les activités sont générées en sélectionnant de façon aléatoire les attributs des activités dans les distributions grâce à la méthode de la transformée inverse, qui est présentée dans la section suivante. La façon d'utiliser cette méthode dans la génération des attributs est par la suite présentée.

3.4.2.1 Méthode de la transformée inverse

Cette méthode permet de générer des nombres sélectionnés de façon aléatoire à partir de la fonction de répartition d'une distribution de probabilité (Le Gauffre, 2016). L'objectif de la méthode est de générer des nombres qui ont la même la fonction de répartition que la distribution de probabilité d'une certaine variable aléatoire discrète X . Cette distribution de probabilité $f(X)$ peut par exemple prendre la forme suivante :

$$\left\{ \begin{array}{l} P(X = 1) = 10\% \\ P(X = 2) = 40\% \\ P(X = 3) = 40\% \\ P(X = 4) = 10\% \end{array} \right.$$

Ainsi, chaque valeur de la variable aléatoire X a une probabilité qui lui est associée. À partir de la distribution de ces probabilités, il est possible d'obtenir la fonction de répartition $F(X)$ de la variable aléatoire X . Cette fonction est obtenue en cumulant la valeur des probabilités de chaque valeur de la variable aléatoire X . La fonction, qui varie donc de 0 à 1, permet de déterminer la valeur de n nombres générés aléatoirement selon la valeur d'un nombre aléatoire N_i , avec i variant de 1 à n .

$$\left\{ \begin{array}{l} 0 \leq N_i < 0,1 \rightarrow x_i = 1 \\ 0,1 \leq N_i < 0,5 \rightarrow x_i = 2 \\ 0,5 \leq N_i < 0,9 \rightarrow x_i = 3 \\ 0,9 \leq N_i < 1 \rightarrow x_i = 4 \end{array} \right.$$

Les nombres générés, représentés par x_i , suivent donc la même fonction de répartition que la variable aléatoire X si la valeur de n , soit la quantité de nombres générés, est élevée. Afin de les générer, la première étape de la méthode consiste à générer des nombres aléatoires N_i à partir d'une distribution de probabilité $f(N)$ uniforme. Dans l'exemple suivant, présenté à la Figure 3-9, quatre nombres aléatoires N_i sont générés à partir de cette distribution dans le but de générer quatre nombres x_i .

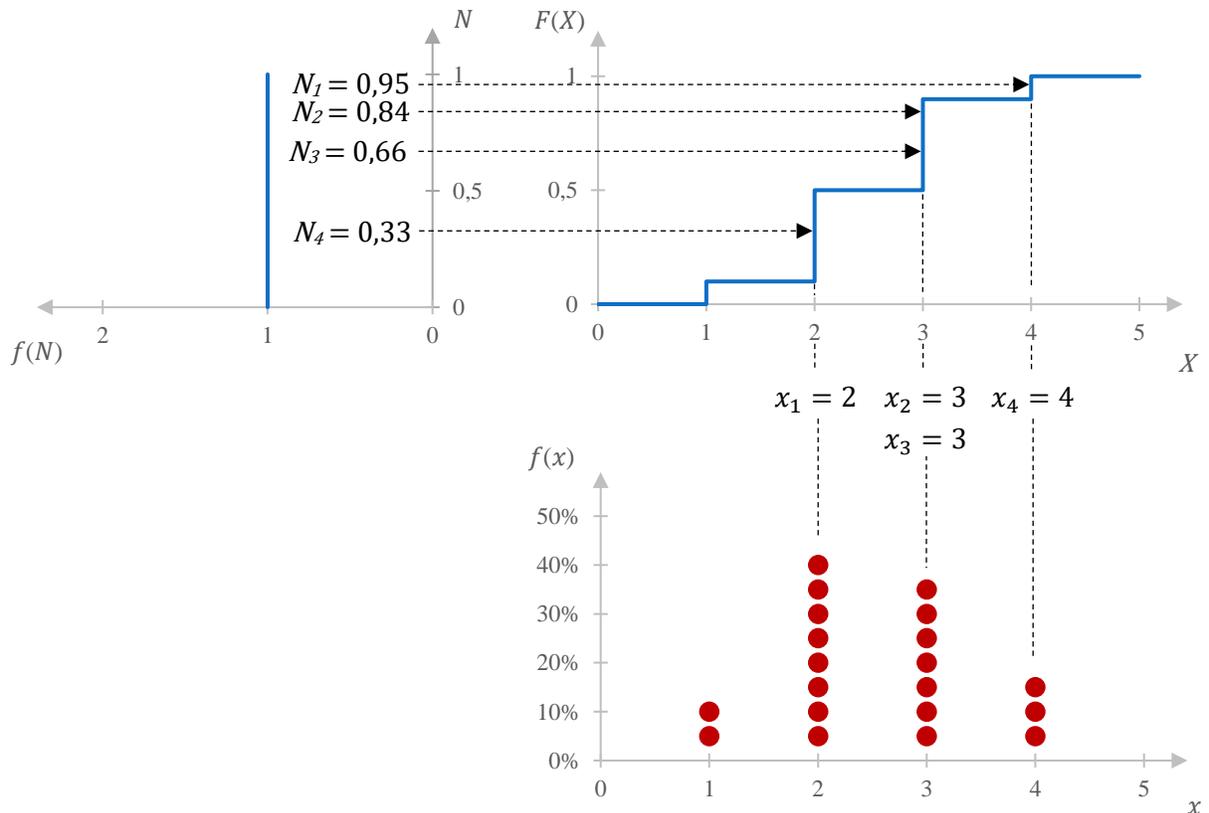


Figure 3-9 Exemple de l'utilisation de la méthode de la transformée inverse

La valeur des quatre nombres aléatoires N_i est comparée avec la fonction de répartition $F(X)$ de la valeur aléatoire X . Les nombres générés x_i prennent les valeurs de la variable aléatoire X à la suite de cette comparaison, ce qui permet de créer une distribution ayant une fonction de répartition $f(x)$ semblable à celle de X .

La méthode de la transformée inverse peut également s'appliquer à une variable aléatoire continue, mais seule des variables aléatoires discrètes sont utilisées pour l'application de cette méthode dans ce mémoire.

3.4.2.2 Génération des activités

Comme mentionné précédemment, les distributions utilisées pour la génération des activités sont différentes de celles utilisées pour la comparaison des groupes. Pour la fréquence des activités, elles sont identiques. Toutefois, les distributions des heures de début et des durées sont légèrement différentes, tel qu'illustré sur la Figure 3-10 et expliqué par après. Cette figure présente également le processus permettant de sélectionner aléatoirement les trois attributs des activités pour ainsi générer un programme d'activités pour un certain individu selon le groupe, c'est-à-dire le patron typique, qui lui est assigné. Cet individu peut par exemple faire partie d'une population fictive, alors que les individus ayant servi à la construction des distributions sont ceux composant l'échantillon de l'enquête OD. Dans l'exemple qui suit, la personne fictive en question appartiendrait au groupe 1.

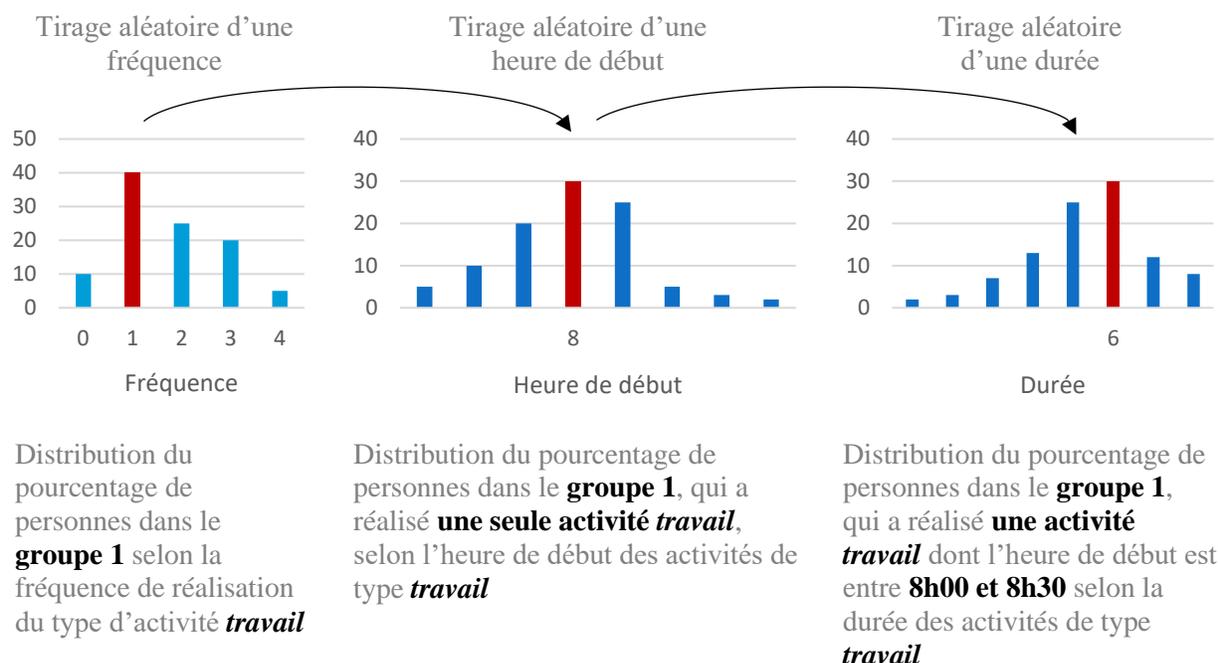


Figure 3-10 Distribution du pourcentage de personnes selon les attributs des activités permettant la génération des activités

La distribution des fréquences de réalisation d'activités est donc construite pour chaque type d'activité. Ainsi, pour chacun des groupes issus de la segmentation présentée au quatrième chapitre, la distribution du nombre de personnes selon le nombre d'un certain type d'activité réalisé est construite. La méthode de la transformée inverse est ensuite utilisée pour sélectionner aléatoirement une fréquence pour chaque individu, à qui un nombre aléatoire est généré pour les

besoins de la méthode. Dans l'exemple ci-haut, la fréquence de 1 est sélectionnée pour le type d'activité *travail*.

Ensuite, l'heure de début est sélectionnée à partir de la distribution de cet attribut. La distribution utilisée pour cette sélection dépend du résultat de l'étape précédente, soit la sélection d'une fréquence. Dans l'exemple ci-haut, alors qu'une fréquence de 1 est sélectionnée pour le type d'activité *travail*, l'heure de départ de cette activité générée est tirée aléatoirement depuis la distribution du nombre de personnes dans le groupe 1 ayant réalisé une seule activité *travail* selon l'heure de départ de ce type d'activité. Ainsi, une distribution selon l'heure de départ est réalisée pour chaque type d'activité, mais également pour chaque fréquence d'activité. De plus, cette distribution est réalisée en regroupant les heures de début des activités par intervalle. La distribution correspond donc au nombre de personnes débutant un certain type d'activité entre 8h00 et 8h30, 8h30 et 9h00 et ainsi de suite. Dans l'exemple présenté dans cette section, l'activité de type *travail* de l'individu fictif débute entre 8h00 et 8h30 à la suite du tirage réalisé par l'entremise de la méthode de la transformée inverse.

Le même processus s'applique à la sélection des durées. Dans l'exemple ci-haut, la durée de cette activité de type *travail* débutant entre 8h00 et 8h30 est sélectionnée dans la distribution propre à la valeur de ces attributs. La distribution utilisée correspond à celle du nombre de personnes ayant réalisé une seule activité travail et ayant débuté cette activité entre 8h00 et 8h30.

Dans le cas où la fréquence sélectionnée lors du premier tirage est de plus de 1, c'est-à-dire que le nombre d'activités d'un certain type attribué à un individu est de deux par exemple, l'heure de début et la durée de ces deux activités sont sélectionnées selon la valeur de ce premier tirage. Les distributions dans lesquelles ces deux attributs sont tirés sont celles qui sont construites à partir des individus ayant réalisé deux activités du type en question. Ainsi, deux tirages différents permettent de sélectionner l'heure de début de chacune des activités, tout comme la durée de celles-ci.

Contrairement aux distributions permettant de comparer les comportements des membres des groupes, les distributions utilisées pour générer des activités sont calculées avec le facteur de pondération. Ainsi, le nombre de personnes ne correspond pas au nombre d'individus dans l'échantillon des enquêtes, mais bien au nombre de personnes dans l'ensemble de la population.

3.5 Modèle d'allocation des personnes à un patron d'activité

Parallèlement à la création des distributions, le modèle d'allocation des patrons peut être estimé. Afin d'attribuer un patron d'activités typique à chaque individu, un modèle de choix discret est utilisé. Appartenant à la famille des régressions logistiques, le modèle choisi est une régression logistique multinomiale qui permet de modéliser une variable catégorielle (Hurlin, 2003). Ainsi, grâce à ce modèle, la variable expliquée, soit l'un des différents patrons d'activités typiques, peut être modélisée en fonction de variables explicatives décrivant une personne. L'équation suivante permet d'évaluer la probabilité qu'un individu i adopte le patron d'activité j , parmi les m patrons, selon les différentes variables explicatives, dénotées x_i .

$$P(y_i = j) = \frac{e^{\beta_j x_i}}{\sum_{k=0}^m e^{\beta_k x_i}} \quad (\text{Éq. 1})$$

La variable β_j représente les différents coefficients propres à chaque variable explicative, variant selon la modalité j . En normalisant les coefficients de la modalité $j = 0$, c'est-à-dire en posant $\beta_0 = 0$, celle-ci devient la modalité de référence à partir duquel le modèle peut être estimé. Dans ce cas-ci, cette modalité peut correspondre au fait d'être immobile et de ne réaliser aucune activité. À la suite de cette transformation, l'équation 1 devient la suivante :

$$P(y_i = j) = \frac{e^{\beta_j x_i}}{1 + \sum_{k=1}^m e^{\beta_k x_i}} \quad (\text{Éq. 2})$$

L'objectif est donc d'estimer la valeur de ces coefficients afin de comprendre l'influence des variables explicatives sur l'adoption d'un patron typique d'activités. Afin de les estimer, la méthode de la maximisation du log-vraisemblance est utilisée. La vraisemblance du modèle correspond au produit de la vraisemblance de chaque observation, soit la probabilité que la valeur prédite soit celle qui a été observée. Afin d'identifier le maximum de la vraisemblance, le logarithme est appliqué à cette fonction pour ainsi transformer la multiplication en une somme, qui est plus facilement dérivable, ce qui est nécessaire pour la recherche du maximum de la fonction de vraisemblance. Pour la fonction de vraisemblance $L(y, \beta_1, \beta_2, \dots, \beta_m)$, la fonction log-vraisemblance correspondante est la suivante :

$$\log L(y, \beta_1, \beta_2, \dots, \beta_m) = \sum_{i=1}^N \sum_{j=0}^m y_{i,j} \log[P(y_i = j)]$$

Où $y_{i,j} = 0$ lorsque que la probabilité correspond à celle de la modalité de référence, sinon $y_{i,j} = 0$ lorsque $y_i = j$, soit lorsque la probabilité calculée est celle d'une autre modalité.

Ainsi, en dérivant cette fonction logarithmique et en identifiant son maximum, les coefficients β_j qui permettent de maximiser la vraisemblance peuvent être identifiés.

De plus, une condition essentielle lors de l'application de ce modèle est l'indépendance des modalités, qui représentent ici les patrons d'activités typiques (W Hosmer & Lemeshow, 1989).

La fonction *multinom* disponible dans le langage R est utilisée pour estimer le modèle de régression logistique multinomiale.

CHAPITRE 4 CRÉATION DE TROIS TYPOLOGIES DE PATRONS D'ACTIVITÉS INDIVIDUELLES

Le but de ce chapitre est de développer une typologie qui consiste à regrouper les individus dont les systèmes d'activités sont similaires. Le processus doit donc permettre de créer des groupes au sein de l'échantillon des quatre enquêtes OD dont les personnes ont des comportements semblables afin de représenter le mieux possible les différents patrons typiques qui existent au sein de la population. En effet, il existe une variété de comportements en ce qui a trait aux activités réalisées par les individus et l'identification de patrons typiques parmi ces comportements est essentielle afin de mieux prédire quelles activités les individus risquent de réaliser.

Des dizaines de patrons typiques pourraient être détectés à travers la population en regard de la pluralité des comportements qui existent au sein de la population en termes de systèmes d'activités. Le défi consiste donc à capturer les patrons les plus pertinents afin de représenter au mieux leur diversité. Cette représentation doit donc avoir la capacité d'intégrer des patrons dont l'évolution à travers le temps reflète les changements dans le comportement des individus. Ces changements peuvent avoir un effet considérable sur l'utilisation des réseaux de transport, c'est pourquoi les patrons résultant de la segmentation doivent être en mesure de capter adéquatement les particularités de la structure des systèmes d'activités des individus et ainsi, intégrer les décisions que prennent les personnes et qui découlent de cette structure.

4.1 Analyse des systèmes d'activités

Préalablement à la création de la typologie, le système d'activités des individus est analysé dans la section qui suit. Tout d'abord, la Figure 4-1 présente la proportion de chaque type d'activités selon l'âge des personnes qui réalisent chacune d'entre elles. La majorité des activités sont de type *travail* et elles sont réalisées principalement par des personnes entre 20 et 60 ans. Celles de types études sont majoritairement réalisées par des personnes de moins de 20 ans. Finalement, les quatre autres types d'activités sont réalisés par des personnes appartenant à tous les groupes d'âge.

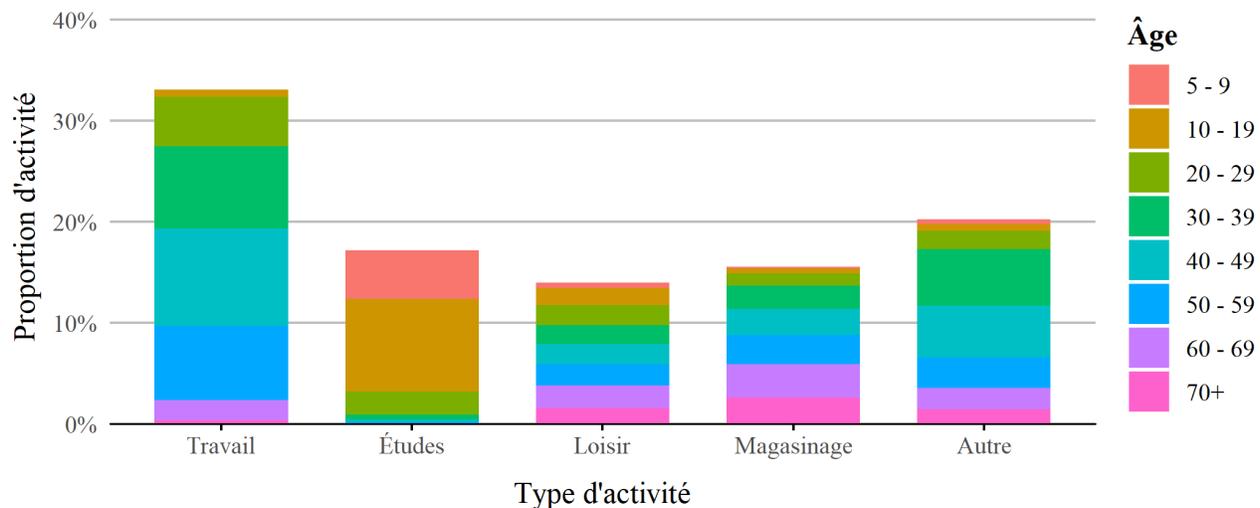


Figure 4-1 Distribution du pourcentage d'activités selon le type d'activité et selon l'âge de la personne qui a réalisé l'activité

La Figure 4-2 présente quant à elle la distribution de la proportion de personnes selon le nombre d'activités réalisées lors de la journée où chaque personne a été sondée.

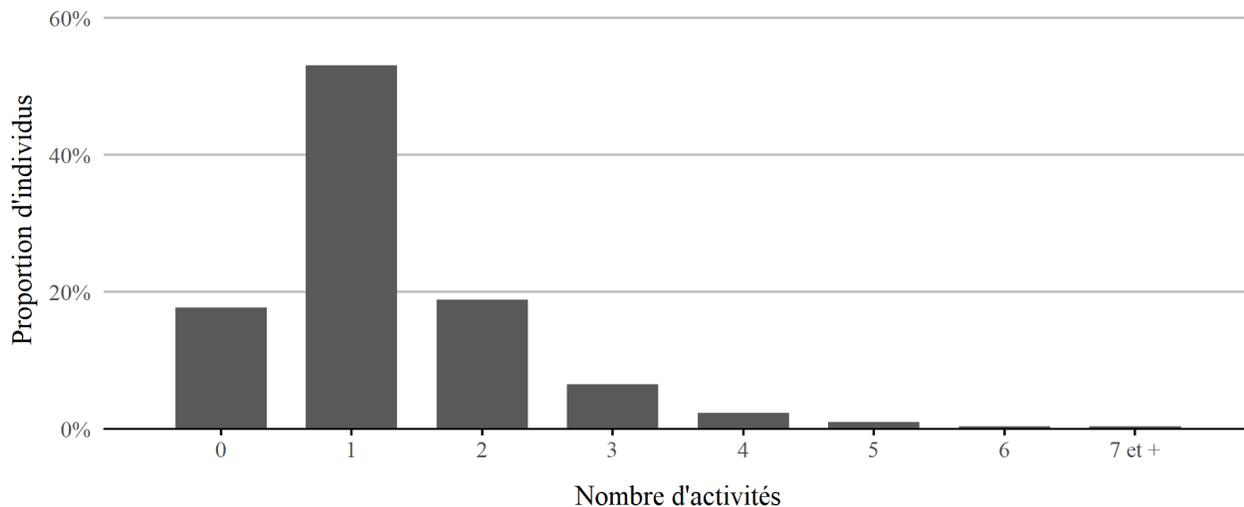


Figure 4-2 Distribution du pourcentage d'activités selon le nombre d'activités

Plus de la moitié des individus, soit 53,1 %, font une seule activité par jour, alors qu'ils sont 18,8 % à en faire deux. De plus, le nombre moyen d'activités est de 1,29 par jour. Ainsi, la majorité des individus ont un système d'activités composé d'aucune activité ou une seule activité.

La Figure 4-3 présente, selon l'âge, la durée moyenne d'activité de même que la fréquence moyenne de réalisation et ce, pour chaque type d'activité

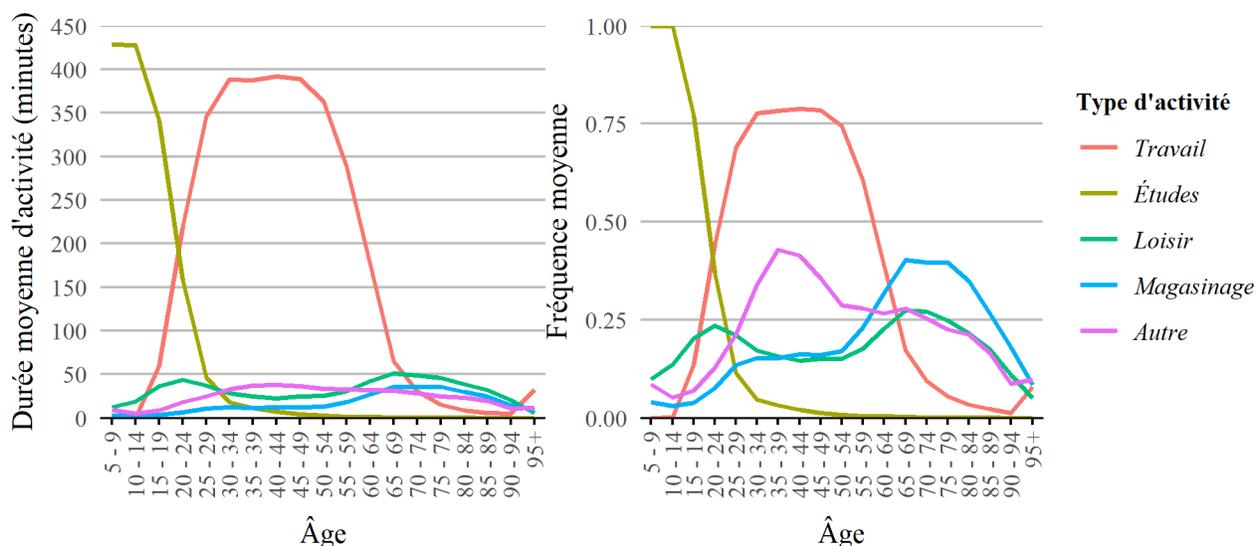


Figure 4-3 Durée et fréquence moyenne de chaque type d'activité selon l'âge

De manière générale, les durées et fréquences moyennes fluctuent énormément selon l'âge. Les durées et fréquences concernant les activités de type études chutent rapidement à partir de l'âge de 15 ans, contrairement aux activités de type travail, dont la valeur des attributs augmente à partir de cet âge. Les attributs des trois autres types d'activités connaissent également une variation importante. Bref, cette analyse permet de constater qu'il existe une variabilité importante dans la population concernant les caractéristiques des systèmes d'activités. C'est cette variabilité qui vise à être captée par les typologies développées dans le présent chapitre.

4.2 Première segmentation

La section suivante présente les résultats de la segmentation réalisée avec le premier vecteur (V1). Tel que mentionné dans le chapitre méthodologique, l'algorithme appliqué à ce vecteur est celui des K-moyennes. Préalablement à cette application, le nombre de groupes doit être déterminé.

4.2.1 Choix du nombre de groupes

Le choix du nombre de groupes repose sur la définition même de l'objectif de l'algorithme des K-moyennes, qui est de maximiser l'homogénéité à l'intérieur de chacun des groupes et de maximiser

l'hétérogénéité entre ceux-ci. La variabilité entre les observations d'un même groupe doit donc être minimale alors que celle entre les centres de chacun d'entre eux doit être maximisée. Le choix du nombre de groupes se base sur ce principe. En effet, afin de faire ce choix, l'algorithme des K-moyennes est appliqué de façon itérative en testant plusieurs nombres de groupe. L'algorithme est ainsi appliqué avec deux, trois, quatre groupes et ainsi de suite. Afin d'évaluer la pertinence de chacune de ces valeurs, le pourcentage de l'inertie expliquée est évalué. Ce pourcentage se traduit par le ratio entre la somme des distances au carré entre le centre des observations, soit la variabilité inter-groupe, qui permet de mesurer la variabilité entre les groupes, et la somme des distances au carré totale. Cette dernière correspond à l'addition de la somme des distances au carré inter-groupe et celle intragroupe, dont la valeur permet de mesurer la variabilité à l'intérieur de chacun des groupes. Il est ainsi possible de calculer la proportion de la variabilité inter-groupe sur la variabilité totale, c'est-à-dire l'inertie expliquée.

Grâce à cet indicateur, différentes valeurs quant au nombre de groupes peuvent être testées sur la base de la variabilité inter et intra groupe. L'approche itérative permet d'évaluer si l'ajout d'un groupe apporte une quantité suffisante d'information. L'inertie expliquée permet d'évaluer l'apport de cette quantité grâce au ratio entre la variabilité inter-groupe et la variabilité totale. Si ce ratio n'augmente pas significativement avec l'ajout d'un groupe, cet ajout n'est donc pas nécessaire et il est inutile de créer un groupe supplémentaire (Madhulatha, 2012). La Figure 4-4 présente ce ratio pour différentes valeurs du nombre de groupes.

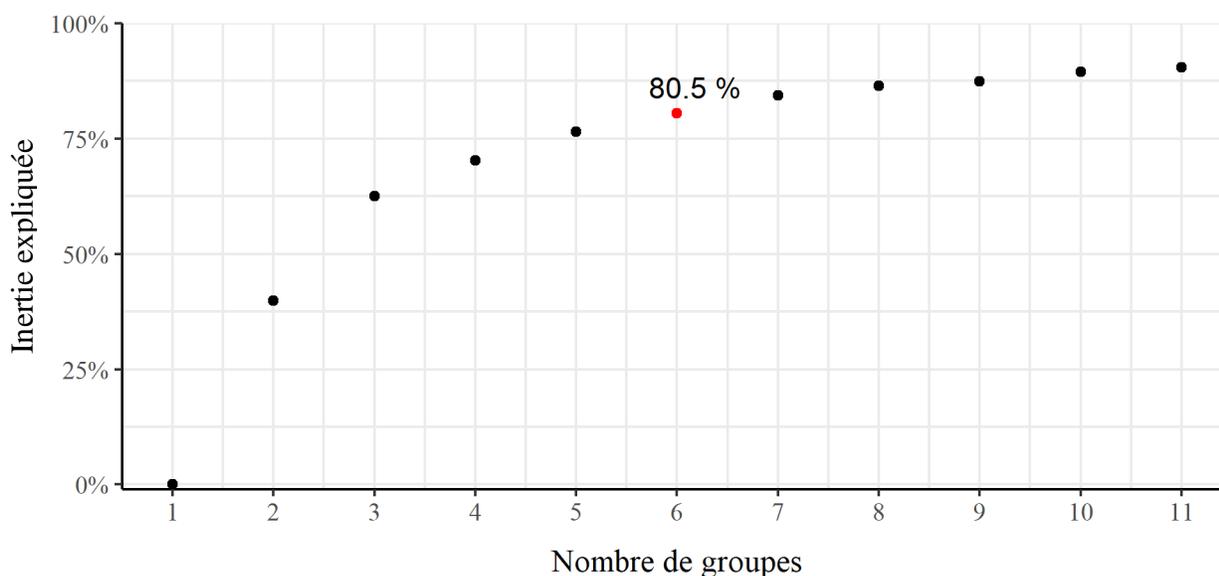


Figure 4-4 Inertie expliquée en fonction du nombre de groupes

Ainsi, avec la création de six groupes, la variabilité inter-groupe correspond à 80,5 % de la variabilité totale. Grâce au « critère du coude », qui permet d'identifier le point d'inflexion dans le graphique ci-haut, il est possible de constater que l'ajout d'autres groupes n'augmenterait pas de façon significative la valeur de ce ratio et par conséquent le gain d'information. En effet, la valeur de ce ratio plafonne à environ 88 %. Il est donc nécessaire de créer au moins six groupes pour garantir une variabilité maximale entre les groupes et une variabilité minimale à l'intérieur de chacun d'entre eux.

4.2.2 Analyse des groupes

La Figure 4-5 présente la distribution du nombre de personnes dans chacun des groupes à la suite de l'application de la première méthode avec le vecteur V1. Conséquemment à l'analyse faite à la section précédente, six groupes sont créés. L'activité dominante correspond à l'activité la plus longue qu'un individu a réalisée au cours de la journée au cours duquel son ménage a été sondé. Le nom des groupes correspond au nom du vecteur (V1) suivi du type de l'activité dominante, soit E pour *études*, T pour *travail*, L pour *loisir*, M pour *magasinage* et A pour *autres*.

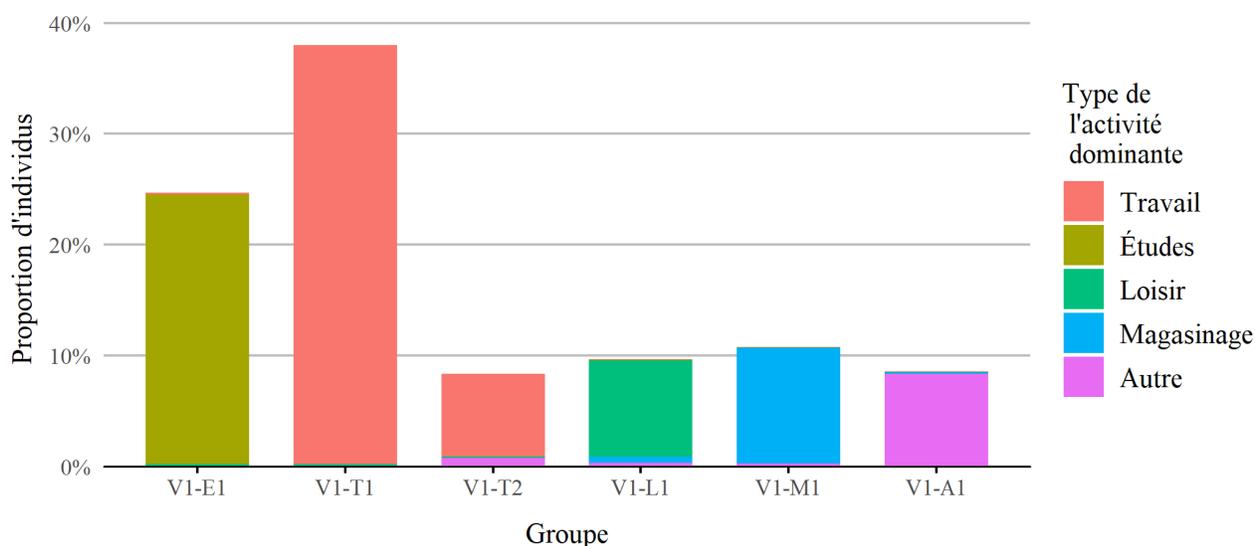


Figure 4-5 Distribution de l'échantillon selon le groupe

Pratiquement tous les individus dans chacun des groupes ont le même type d'activité dominante et chaque type est représenté par un groupe, à l'exception du *travail*, qui en compte deux. Ainsi, en divisant la population en six, ce partitionnement permet essentiellement de regrouper les individus en fonction du type de leur activité principale. L'augmentation du nombre de groupes ne conduirait

pas à une plus grande variabilité inter-groupe, mais permettrait de détecter d'autres comportements à l'intérieur de ces groupes. Ces comportements peuvent échapper à cette segmentation et peuvent être pertinents pour la modélisation des systèmes d'activités. C'est pourquoi il a été décidé de créer un plus grand nombre de groupes afin d'obtenir une meilleure granularité des données.

Le choix final du nombre de groupes se base essentiellement sur la stabilité des groupes à la suite de plusieurs essais, réalisés selon la méthodologie suivante. Comme mentionné dans la méthodologie à la section 3.3.2, le résultat de l'algorithme des K-moyennes n'est pas unique et peut parfois mener à un optimum qui n'est pas global. C'est pourquoi il est nécessaire d'appliquer l'algorithme plusieurs fois afin d'identifier les groupes les plus pertinents.

Ainsi, l'algorithme est appliqué à dix reprises pour différentes valeurs du nombre de groupes, variant de 9 à 12. Un total de 40 simulations est donc exécuté et lors de l'initialisation de chacune d'entre elles, dix centres aléatoires sont utilisés. Le résultat de chacun de ces 40 essais est choisi parmi les dix partitionnements obtenus grâce aux dix centres aléatoires, testés par la fonction *kmeans* du langage R.

L'analyse des 40 différentes segmentations démontre que certains groupes sont présents dans la majorité des partitionnements. Par exemple, un grand nombre de simulations mène à la création de deux groupes où l'activité principale est de type *études*. La différence entre les deux groupes est la réalisation ou non d'une autre activité de type *loisir* après l'école. Des situations similaires étaient observées dans le cas des groupes où le *travail* est l'activité dominante. Bref, plusieurs groupes semblent plus stables que d'autres en étant présents dans la majorité des simulations réalisées, alors que certains groupes sont présents de façon isolée, puisqu'ils sont observés dans très peu de simulations, ce qui laisse croire que leur présence est due à l'atteinte d'un optimum local et non global.

À la suite de cette analyse, le choix final de la typologie est fait en sélectionnant, parmi les 40 simulations réalisées, celles qui contiennent les groupes dont l'observation est la plus fréquente. La Figure 4-6 permet d'observer la distribution du pourcentage de personnes dans chaque groupe pour la typologie obtenue à la suite de la sélection de cette simulation.

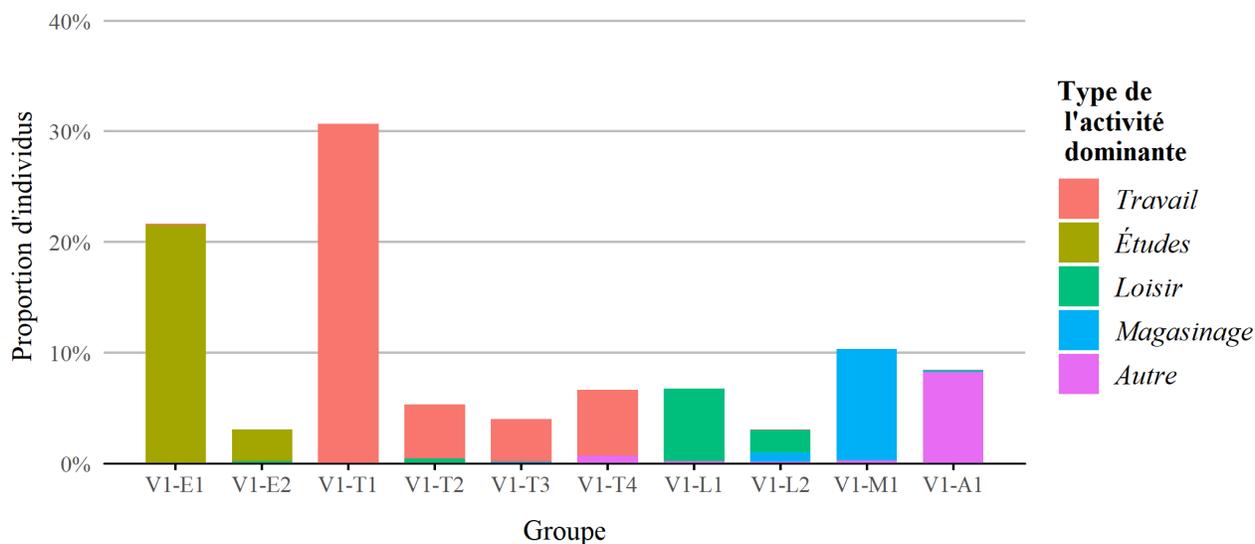


Figure 4-6 Distribution du pourcentage de personnes dans chacun des groupes

Tout comme la segmentation composée de six groupes, la totalité des groupes est composée d'une majorité d'individus ayant le même type d'activité dominante. Certains types sont également présents une seule fois, soit le *magasinage* et le type *autre*. Cependant, les autres types sont maintenant représentés par des groupes supplémentaires, soit un de plus pour les *études* et le *loisir*, pour un total de deux groupes chaque, et deux de plus pour le *travail*, qui est maintenant représenté par quatre groupes. L'ensemble des individus sont bien répartis dans l'ensemble des groupes, alors que la proportion de personnes varie de 2,9 à 10,2 %, à l'exception de V1-T1 et V1-E1, qui accaparent respectivement 31% et 22,3 % de l'échantillon des quatre enquêtes OD. Ces deux groupes se caractérisent par la réalisation d'une seule activité, comme le démontre la Figure 4-7, qui illustre le centre des groupes.

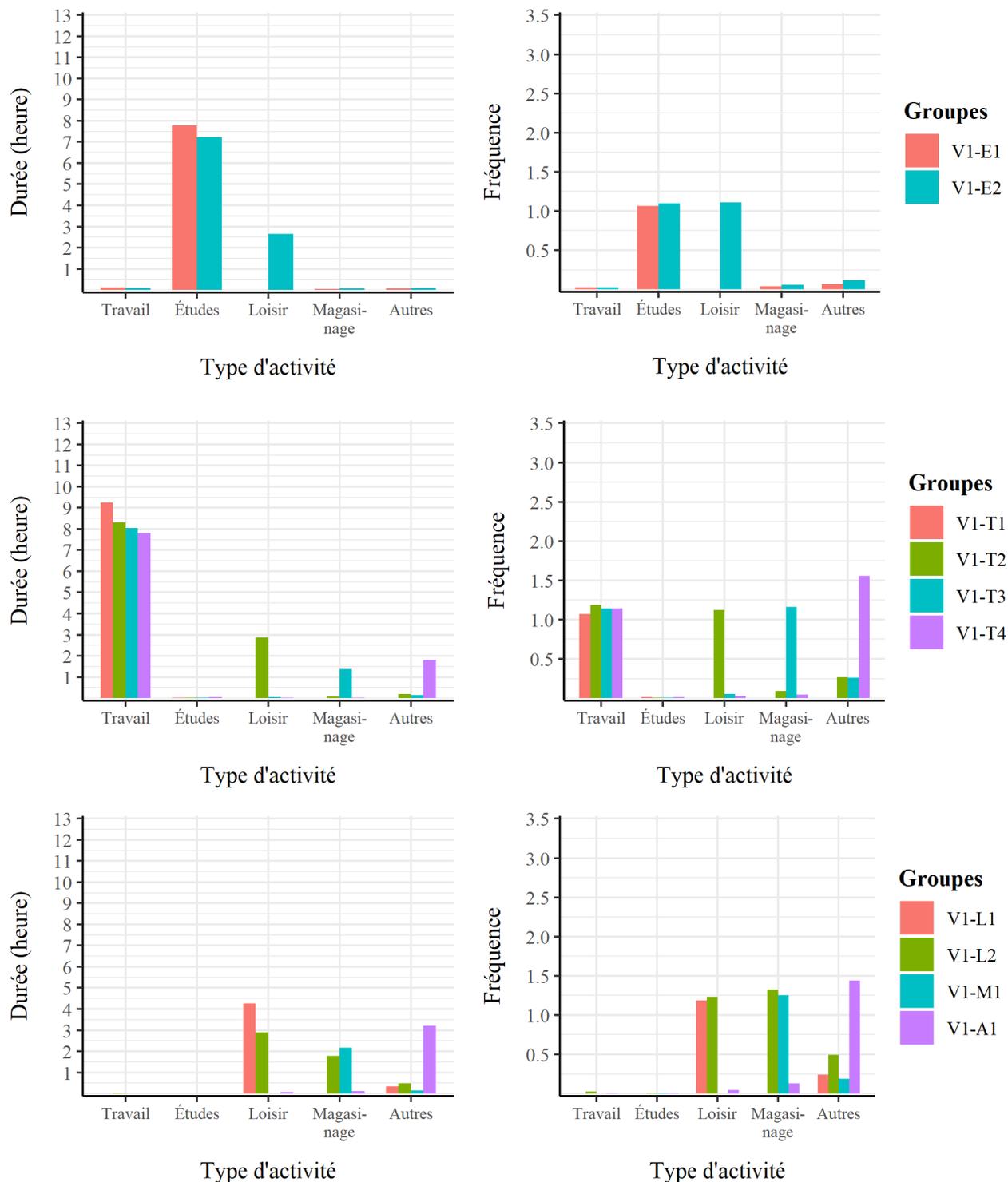


Figure 4-7 Centres des groupes de la segmentation avec V1

La principale différence entre V1-E1 et V1-E2 est la réalisation d'une activité de type *loisir*. Les membres du premier groupe n'en réalisent aucune alors que ceux du deuxième en réalisent une

d'une durée moyenne de 2,9 heures, avec une fréquence moyenne de 1,1 activité. Les membres de V1-E2 passent également moins de temps en moyenne sur les bancs d'école.

Une situation similaire est observée pour les groupes dont l'activité dominante est le *travail*. Le premier groupe, V1-T1, est composé d'individus qui ne réalisent aucune autre activité que l'activité principale, dont la durée moyenne est plus élevée que les autres groupes. Ces derniers se différencient par une activité secondaire réalisée en plus de l'activité principale, soit le *loisir* pour V1-T2, le *magasinage* pour V1-T3 et le type *autre* pour V1-T4. Les durées moyennes de ces activités supplémentaires varient de 1,4 à 2,9 heures. La fréquence moyenne de l'ensemble de ces activités secondaires est un peu plus de 1, à l'exception de V1-T4, dont la fréquence moyenne est de 1,6, ce qui indique qu'une grande proportion des membres de ce groupe réalise au moins deux activités de type *autre* lors de la journée où leur ménage a été sondé.

Les deux groupes dont l'activité dominante est le *loisir* se différencient par la durée moyenne de l'activité dominante, soit 4,2 heures pour V1-L1 et 2,9 heures pour V1-L2, mais également par la réalisation d'un autre type d'activité. En effet, les membres de V1-L2 font également d'autres activités, soit le *magasinage* et, dans une moindre mesure, une activité de type *autre*. La durée moyenne de ces deux activités supplémentaires est de 100 et 30 minutes respectivement et la fréquence moyenne est 1,1 et 0,5, ce qui signifie qu'une minorité de personnes présentes dans V1-L2 fait une activité de type *autre*.

Les membres des deux derniers groupes font des activités de courte durée et une seule fois par jour pour la majorité d'entre eux. En effet, l'activité dominante des personnes composant le groupe V1-M1 est caractérisée par une durée moyenne de 3 heures et une fréquence moyenne de 1,1. De leur côté, les membres de V1-A1 font en moyenne 1,4 activité de type *autres* pendant 190 minutes en moyenne.

La Figure 4-8 présente, pour chacun des groupes, la proportion d'individus qui font un des cinq types d'activité et ce, selon l'heure de la journée. Alors que les centres des groupes présentés précédemment illustrent la durée et la fréquence moyenne de chaque type d'activité pour chacun des groupes, cette figure permet de comprendre à quel moment de la journée les différentes activités se situent dans le temps.

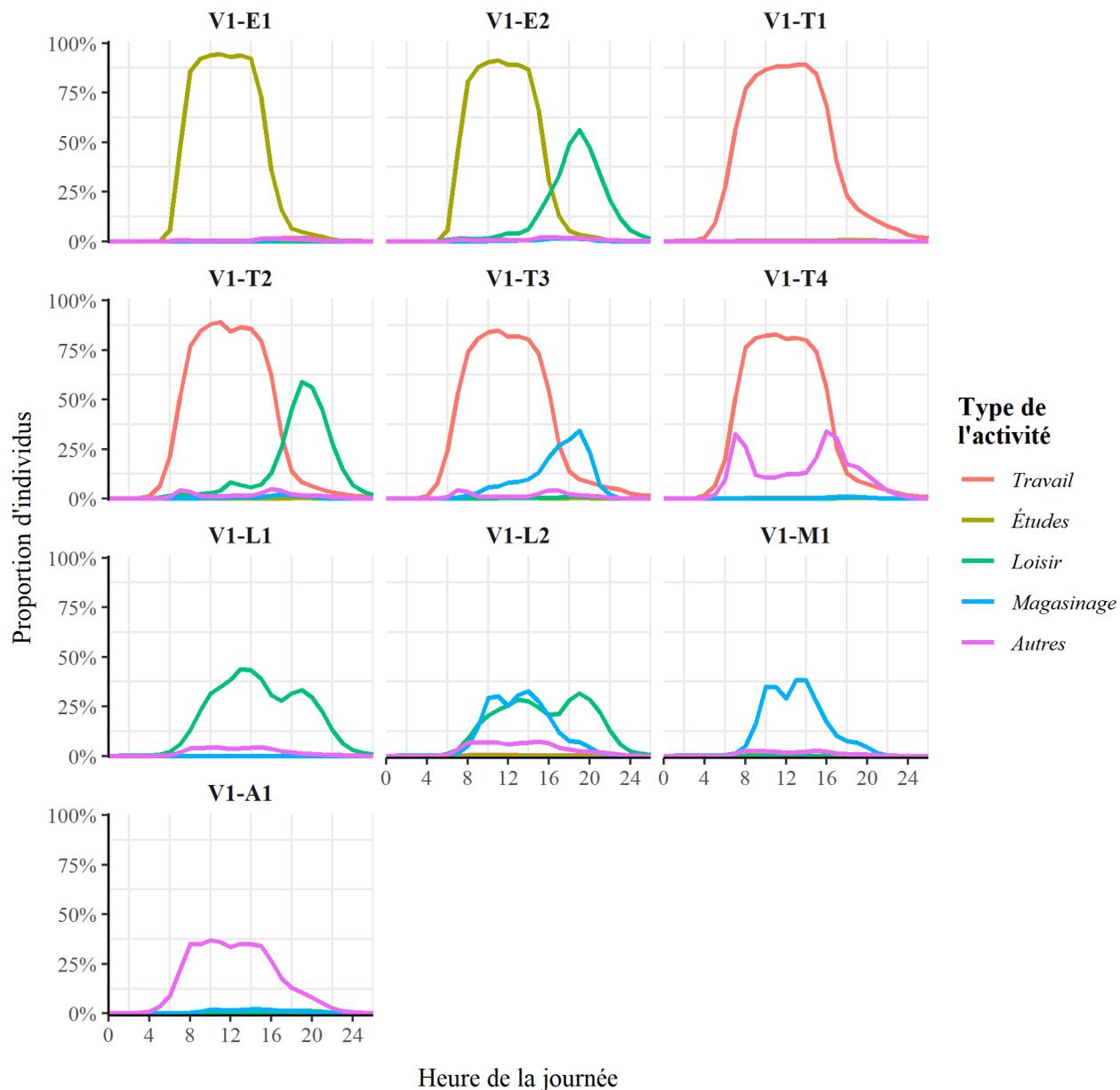


Figure 4-8 Proportion d'individus, à l'intérieur de chaque groupe, réalisant un des cinq types d'activités selon l'heure de la journée

La figure ci-haut permet tout d'abord de constater que plus de 80 % des membres dont l'activité dominante est le *travail* ou les *études* font leur activité principale entre 10h et 14h. De plus, alors que le centre des groupes confirmait la réalisation d'une activité de type *loisir* par les membres de V1-E2, ce qui n'est pas le cas des membres de V1-E1, l'évolution temporelle ci-haut permet de constater que cette activité se déroule principalement le soir après l'école, alors que 55 % des individus composant ce groupe font une activité de type *loisir* entre 19h et 19h30.

La même situation est observable pour les groupes où le *travail* est l'activité dominante. En effet, les membres de V1-T2 et V1-T3 réalisent leur activité secondaire, soit le *magasinage* et les activités de type *loisir* respectivement, le soir, après le travail. La pointe concernant la réalisation de ces types d'activité est également entre 19h et 19h30. Cependant, la structure temporelle du parton typique des membres de V1-T4 est passablement différente. Alors que le centre des groupes confirmait qu'une grande proportion des membres de ce groupe-ci faisait une activité de type *autre* plus d'une fois par jour, la Figure 4-8 apporte un éclairage supplémentaire à ce constat, en révélant que ces deux activités se déroulent le matin et le soir, soit avant et après le *travail*.

La Figure 4-8 permet également de constater l'apparition de pointes au cours de la journée concernant les activités réalisées par les individus composant les quatre derniers groupes. En effet, les personnes regroupées au sein de V1-L1 sont nombreuses à faire leur activité principale entre 12h et 13h de même qu'entre 18h et 20h. Les périodes de pointe pour les membres de V1-L2 sont similaires, mais celle entre 18h et 20h est plus élevée que celle de l'après-midi, où d'autres types d'activités sont réalisés par les membres de ce groupe, soit le *magasinage* et, dans une moindre mesure, les activités de type *autre*. La proportion de personnes de V1-L2 réalisant du *magasinage* est également assujettie à des pointes, qui sont situées avant et après le dîner. Les mêmes pointes sont observées pour V1-M1, dont la majorité des membres ne font qu'une seule activité *magasinage*, contrairement aux personnes appartenant à V1-L2 dont la majorité réalise une activité de *loisir* en plus. Le dernier groupe, V1-A1, regroupe les individus ayant réalisé des activités de type *autres* principalement entre 8h et 16h. Toutefois, la majorité d'entre eux débute la réalisation de cette activité avant midi, tel qu'illustrée sur la Figure 4-9, qui illustre la distribution de la proportion d'individus selon l'heure de départ de l'activité dominante.

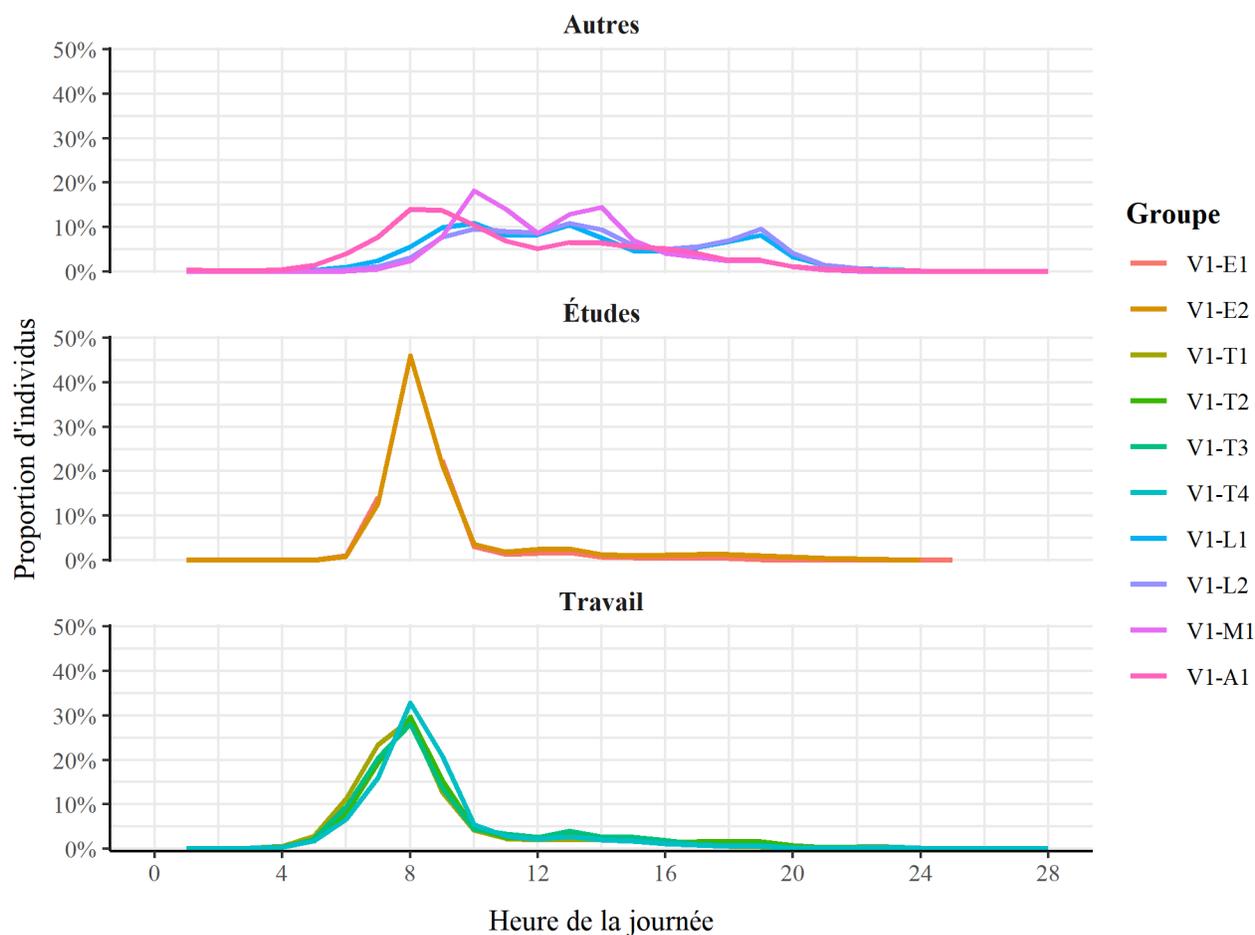


Figure 4-9 Proportion d'individus selon l'heure de début de l'activité dominante

Pour la majorité des membres dont l'activité principale est le *travail* ou les *études*, l'heure de début de cette activité est durant l'heure de pointe habituelle, soit le matin entre 6h et 9h. Une pointe est également visible entre 8h et 10h pour les membres de V1-A1, mais certains d'entre eux débutent leur activité de type *autres* tout au long de la journée et ce jusqu'à 20h. La distribution selon les heures de début est également étalée tout au long de la journée pour les deux groupes où l'activité dominante est de type *loisir*, avec une pointe le soir entre 19h et 20h. La majorité des membres de V1-M1 débute leur *magasinage* avant et après le dîner, conformément à ce qui est observé sur la Figure 4-8.

Le dernier élément de l'analyse de la typologie est présenté à la Figure 4-10, qui indique la proportion d'individus dans chacun des groupes, selon le jour de la semaine.

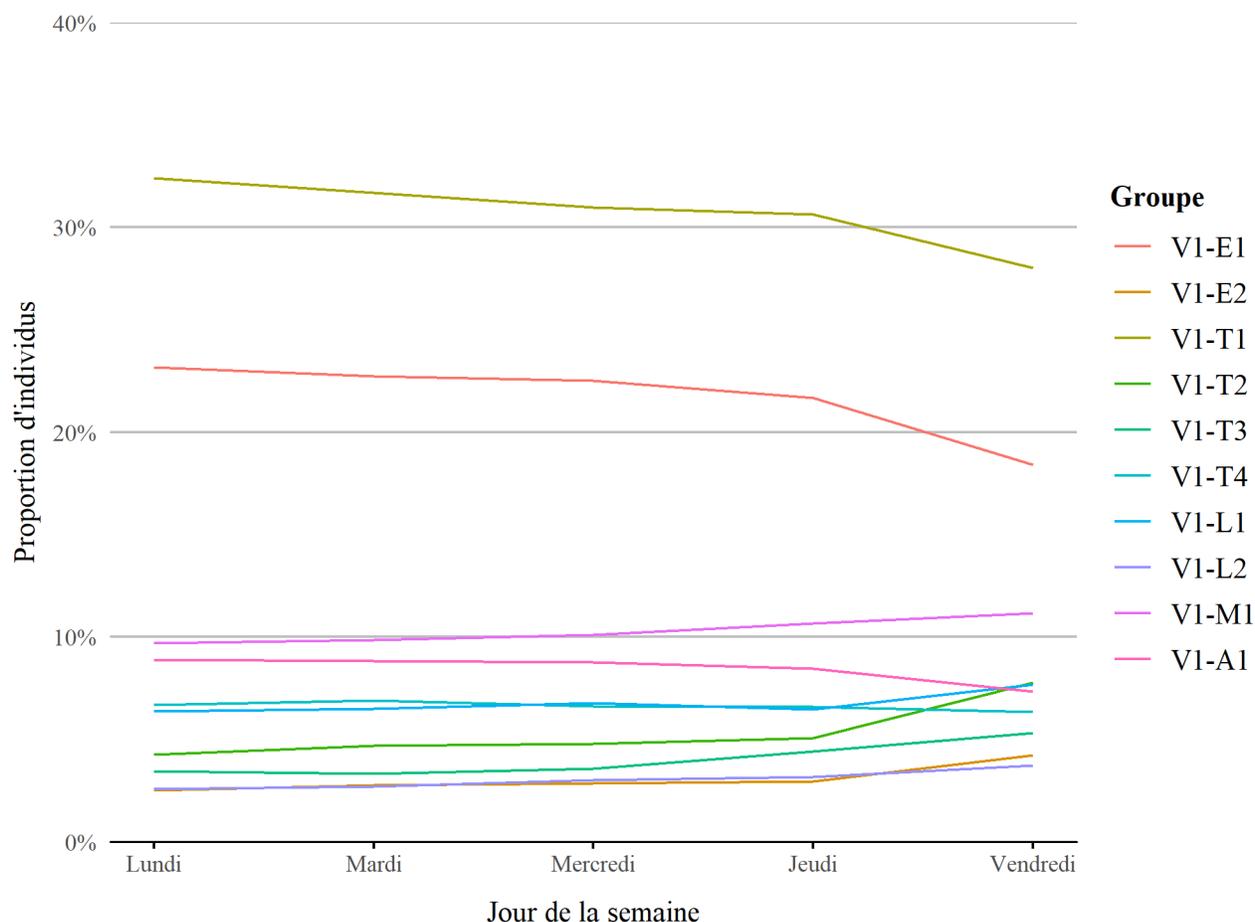


Figure 4-10 Proportion de personnes dans chacun des groupes selon le jour de la semaine

Il est possible de constater que les proportions varient tout au long de la semaine, alors que certains groupes voient leur proportion diminuer chaque jour et d'autres voient plutôt leur proportion augmenter. Ainsi, au fil de l'avancement de la semaine, de moins en moins de personnes réalisent les patrons V1-E1, V1-T1, V1-T4 et V1-A1, contrairement aux autres groupes, qui sont dominés par des personnes qui réalisent des activités de type *loisir* ou *magasinage*.

4.2.3 Analyse des distributions

Afin de s'assurer que les comportements sont significativement différents d'un groupe à l'autre, les tests de Kolmogorov-Smirnov sont appliqués aux distributions construites selon la méthode présentée à la section 3.4.1.1. Le Tableau 4-1 présente le résultat des tests pour les fréquences d'activités. La conclusion des résultats est représentée par trois lettres, soit D, I et N, qui signifient différent, identique et non concluant. Ces trois catégories sont attribuées selon le résultat du test,

qui permet de vérifier si l'hypothèse nulle, à savoir que les distributions sont identiques, peut être rejetée ou acceptée.

Tableau 4-1 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité

Groupe comparés	Travail	Études	Loisir	Magasinage	Autres
V1-E1 & V1-E2	I	D	D	D	D
V1-E1 & V1-T1	D	D	I	D	D
V1-E1 & V1-T2	D	D	D	D	D
V1-E1 & V1-T3	D	D	D	D	D
V1-E1 & V1-T4	D	D	D	D	D
V1-E1 & V1-L1	D	D	D	D	D
V1-E1 & V1-L2	I	D	D	D	D
V1-E1 & V1-M1	D	D	I	D	D
V1-E1 & V1-A1	D	D	D	D	D
V1-E2 & V1-T1	D	D	D	D	D
V1-E2 & V1-T2	D	D	N	D	D
V1-E2 & V1-T3	D	D	D	D	D
V1-E2 & V1-T4	D	D	D	N	D
V1-E2 & V1-L1	D	D	D	D	D
V1-E2 & V1-L2	I	D	D	D	D
V1-E2 & V1-M1	D	D	D	D	D
V1-E2 & V1-A1	D	D	D	D	D
V1-T1 & V1-T2	D	I	D	D	D
V1-T1 & V1-T3	D	I	D	D	D
V1-T1 & V1-T4	D	I	D	D	D
V1-T1 & V1-L1	D	D	D	I	D
V1-T1 & V1-L2	D	I	D	D	D
V1-T1 & V1-M1	D	D	I	D	D
V1-T1 & V1-A1	D	N	D	D	D
V1-T2 & V1-T3	D	I	D	D	I
V1-T2 & V1-T4	D	I	D	D	D
V1-T2 & V1-L1	D	N	D	D	D
V1-T2 & V1-L2	D	I	D	D	D
V1-T2 & V1-M1	D	N	D	D	D
V1-T2 & V1-A1	D	I	D	D	D
V1-T3 & V1-T4	I	I	D	D	D
V1-T3 & V1-L1	D	N	D	D	D
V1-T3 & V1-L2	D	I	D	D	D
V1-T3 & V1-M1	D	N	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes

I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques

N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau 4-2 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité

Groupe comparés	<i>Travail</i>	<i>Études</i>	<i>Loisir</i>	<i>Magasinage</i>	<i>Autres</i>
V1-T3 & V1-A1	D	I	N	D	D
V1-T4 & V1-L1	D	D	D	D	D
V1-T4 & V1-L2	D	I	D	D	D
V1-T4 & V1-M1	D	N	D	D	D
V1-T4 & V1-A1	D	N	D	D	D
V1-L1 & V1-L2	D	N	D	D	D
V1-L1 & V1-M1	I	I	D	D	D
V1-L1 & V1-A1	N	N	D	D	D
V1-L2 & V1-M1	D	I	D	D	D
V1-L2 & V1-A1	D	I	D	D	D
V1-M & V1-A1	D	I	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

La conclusion de la majorité des tests confirme que les comportements en lien avec la fréquence de réalisation des activités ne sont pas similaires pour les individus appartenant à des groupes différents. Les tests dont l'hypothèse nulle ne peut être rejetée permettent également de confirmer les particularités de chacun des groupes. Tout d'abord, la majorité des groupes dont les *études* ne sont pas l'activité dominante ont des distributions identiques selon la fréquence de réalisation des activités de type *études*. De plus, les groupes V1-E1, V1-T1 et V1-M1 ont tous des distributions identiques en ce qui concerne la fréquence de réalisation des activités de type *loisir*, c'est-à-dire qu'aucun individu ne réalise ce type d'activité. Les individus réalisant ce type d'activité, en plus du *travail*, des *études* et du *magasinage*, sont plutôt compris dans les groupes V1-E2, V1-T2 et V1-L2. C'est pourquoi leurs distributions selon la fréquence des activités *loisir* sont identiques.

La conclusion des tests en lien avec les distributions selon les heures de début et les durées sont présentées dans le Tableau A-1 dans l'Annexe A. Encore une fois, les résultats des tests permettent de conclure que les comportements sont significativement différents d'un groupe à l'autre.

4.3 Deuxième segmentation

La section suivante présente le résultat de l'application de l'algorithme hiérarchique agglomératif sur le deuxième vecteur (V2). Tel que mentionné dans le chapitre méthodologique, le principal résultat de cette application est un dendrogramme, qui illustre comment les vecteurs sont

agglomérés entre eux. La création des groupes nécessite une coupure de ce dendrogramme à un certain niveau, soit au niveau de dissimilarité désiré. Ce choix demeure une tâche délicate puisqu'en théorie, il est possible de créer de 14 965 groupes à un seul groupe, puisqu'il y a, à la base du dendrogramme, autant de groupes que d'observations, et l'ensemble de ces derniers sont réunis au sein d'un seul groupe au sommet de l'arbre. Ainsi, dû à la nature du dendrogramme, dont les groupes sont tous imbriqués l'un dans l'autre, les groupes choisis dans le cadre de ce mémoire sont en fin de compte l'agglomération d'autres groupes. Autrement dit, un certain patron typique, représenté par un groupe, renferme également d'autres patrons typiques, soit d'autres comportements qui sont assez similaires, mais qui se différencient par de légères variations concernant un certain aspect du comportement en question. Par exemple, en scindant le dendrogramme à un niveau de dissimilarité très élevé, trois groupes sont créés. La distribution du nombre de personnes dans chacun de ces groupes est présentée à la Figure 4-11.

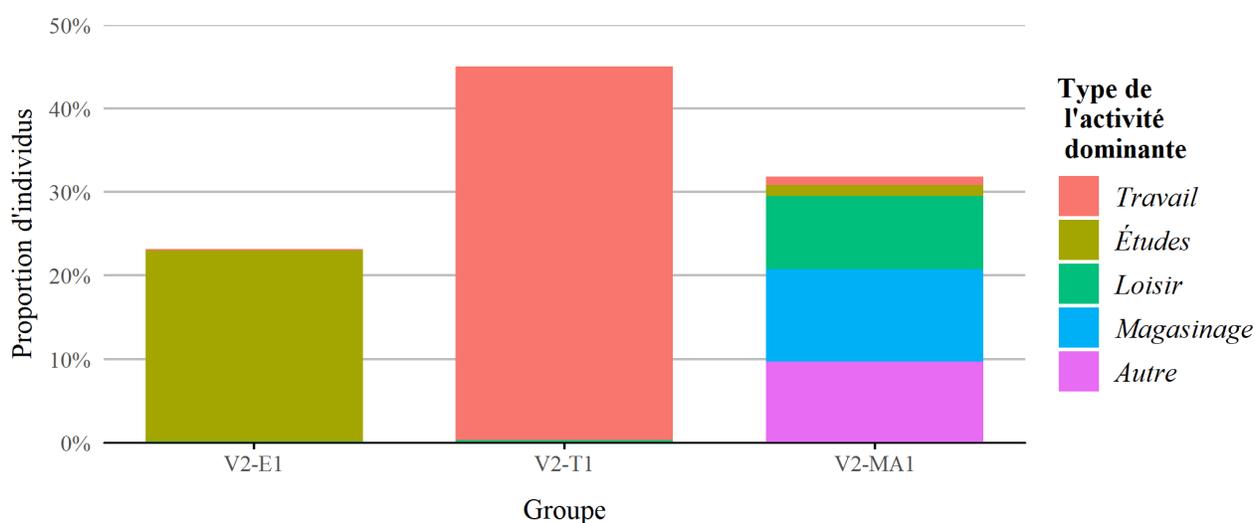


Figure 4-11 Distribution du pourcentage de personnes à la suite de la création de trois groupes

La création de trois groupes permet principalement de regrouper des individus ayant le même type d'activité dominante dans le cas des *études* et du *travail* et de réunir tous les autres individus dont l'activité principale est d'un autre type. Ainsi, l'algorithme, tous juste avant de réunir l'ensemble des observations au sein d'un seul groupe, réussit très bien à réunir les personnes dont l'activité principale est le *travail* ou les *études*. En revanche, chacun de ces trois groupes renferme une quantité innombrable de groupes, qui peuvent être révélés un à un en coupant le dendrogramme à un niveau de dissimilarité toujours plus bas. Le groupe supplémentaire issu de ce coupage itératif

représente donc une variante d'un des groupes déjà présents, puisqu'en abaissant le niveau, un des groupes déjà existants se divise en deux pour ainsi former un groupe supplémentaire.

Afin d'identifier le nombre optimal de groupe, le dendrogramme a été coupé à plusieurs niveaux, chaque coupure menant à la création d'un nombre différent de groupes. Les nombres de groupes testés sont de huit à douze. L'analyse des groupes créés à la suite de ces différentes coupures porte le choix sur dix.

4.3.1 Analyse des groupes

À la suite de la coupure du dendrogramme menant à la création de dix groupes, les individus composant V2-T1 font maintenant partie de cinq différents groupes, alors que les membres de V2-MA1, qui signifie multiactivité, sont divisés en quatre groupes. Ces trois groupes supplémentaires sont composés de moins de 5 % des individus et sont tous les trois caractérisés par des personnes ayant la même activité dominante, soit le type *loisir* et *autre*, tel que présenté sur la Figure 4-12. Ainsi, le groupe V2-MA1 issu de la coupure menant à la création de trois groupes reste pratiquement intact. Les trois groupes supplémentaires correspondent donc à une petite portion des personnes qui était présentes dans V2-MA1. Cela signifie que malgré la création de dix groupes, la majorité des membres de V2-MA1 ont des patrons d'activités assez similaires pour rester dans la même partition.

La raison pour laquelle dix groupes sont créés est que la création de deux groupes supplémentaires permettrait uniquement de diviser davantage le groupe V2-T1. Afin de diviser V2-MA1, treize groupes doivent être créés. La prochaine étape de ce mémoire est d'estimer un modèle d'allocation à ces groupes, ce qui peut s'avérer ardu avec un trop grand nombre de groupes.

De plus, le groupe V2-E1 est toujours présent. Pour diviser ce groupe, il faut couper le dendrogramme à un niveau où douze groupes sont créés. Cela signifie que les individus à l'intérieur de ce groupe sont davantage similaires entre eux que ceux composant les groupes où le *travail* est le type dominant. La Figure 4-12 permet d'observer la proportion de personnes dans V2-E1, de même que dans les autres groupes issus du partitionnement final composé de dix groupes.

Afin de s'assurer de la stabilité des résultats, une analyse de sensibilité a été réalisée. L'algorithme a ainsi été réalisé plusieurs fois avec plusieurs échantillons de 14 965 personnes afin de s'assurer de la constance de la typologie.

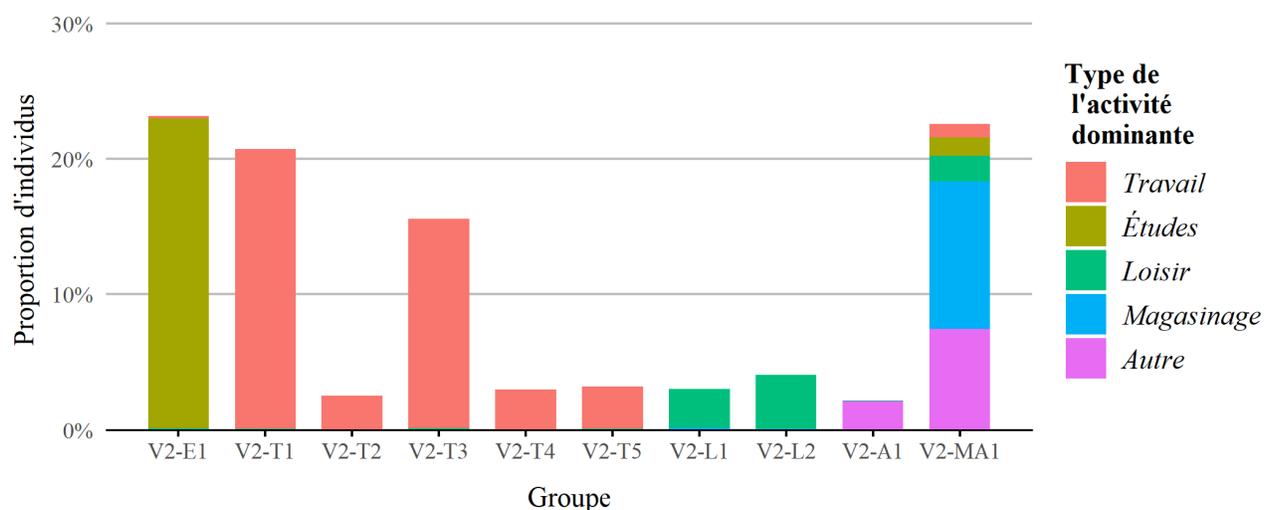


Figure 4-12 Distribution du pourcentage de personnes avec le choix final du nombre de groupes

La majorité des personnes dont l'activité principale est le *travail* sont dans les groupes V2-T1 et V2-T3, alors que moins de 5 % des individus sont dans les groupes V2-T2, V2-T4 et V2-T5. Les trois groupes supplémentaires issus de V2-MA1 incluent des personnes dont l'activité dominante est le *loisir* dans le cas de V2-L1 et V2-L2 et de type *autre* pour V2-A1. Ces groupes sont également composés de moins de 5 % des individus chacun. La Figure 4-13 permet de comprendre ce qui différencie l'ensemble des groupes grâce à la durée totale et le fréquence moyenne de chaque type d'activité.

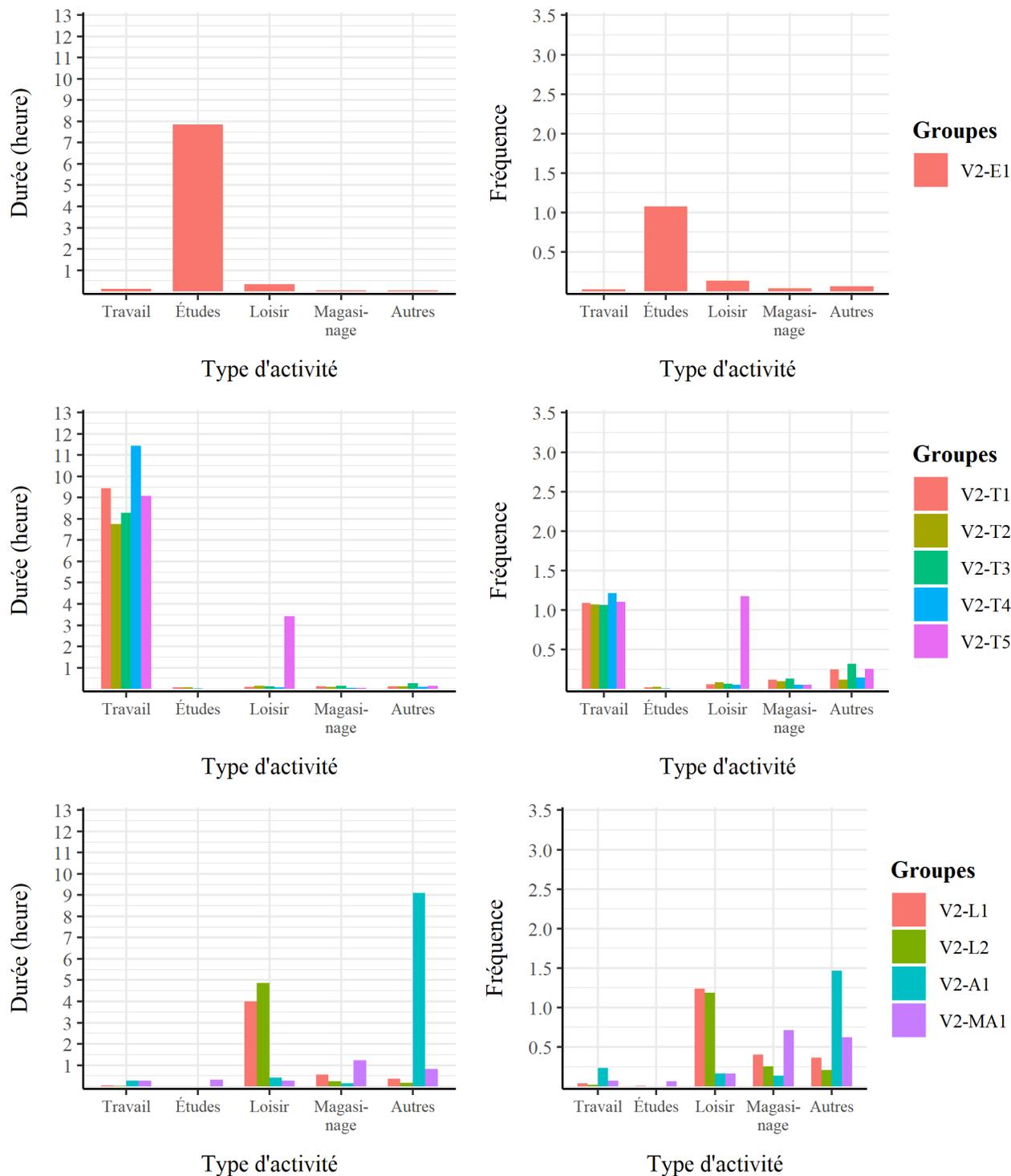


Figure 4-13 Centres des groupes de la segmentation avec V2

Le V2-E1 rassemble pratiquement toutes les personnes dont la principale activité est les études. La figure ci-haut permet simplement de confirmer que les étudiants passent en moyenne moins de huit heures par jour à l'école et ce une fois par jour.

À l'exception de V2-T5, la différence entre les groupes dont la principale activité est le *travail* est la durée. En effet, les membres de V2-T2 et V2-T3 semblent passer moins de temps en moyenne à travailler que ceux de V2-T1, alors que les membres de V2-T4 sont en moyenne beaucoup plus longtemps au *travail* que les autres. Seul V2-T5 se différencie par la réalisation d'une l'activité supplémentaire, soit une activité de type *loisir*. Ainsi, alors que la principale différence entre les groupes issue de la première segmentation est la réalisation d'une activité secondaire, ce qui différencie les groupes du deuxième partitionnement est davantage la durée totale moyenne, à l'exception de V2-T5.

Le système d'activité des membres de V2-MA1 semble être caractérisé par la présence de plusieurs patrons typiques. Le niveau de dissimilarité choisie lors de la coupure du dendrogramme n'était pas suffisant afin de les détecter et de scinder V2-MA1 en plusieurs groupes. Cependant, la structure du système des membres de cette partition semble, dans une certaine mesure, avoir un point commun, soit la réalisation d'activités de courte durée. La durée moyenne des activités de type *magasinage* et *autres* est en revanche plus élevée que celle des autres types d'activités.

Les trois autres groupes, qui étaient à l'origine à l'intérieur de V2-MA1 avant l'abaissement du niveau de dissimilarité menant à la création de dix segments, se différencient par le type de l'activité principale. Dans le cas de V2-A1, c'est le type *autre* qui domine. Les deux autres groupes, où l'activité dominante est le *loisir*, semblent très similaires selon la figure ci-haut, à l'exception de la différence entre les durées moyennes, qui est de moins d'une heure. Toutefois, en calculant la proportion d'individus réalisant ce type d'activité chaque heure de la journée, dont le résultat est présenté à la Figure 4-14, il est possible de constater que la différence entre les deux groupes concerne plutôt le moment de la journée où est réalisée cette activité. En effet, 88 % des membres de V2-L1 performant une activité de type *loisir* entre 19h et 20h, alors que 80 % des individus composant V2-L2 font ce type d'activité entre 13h et 15h.

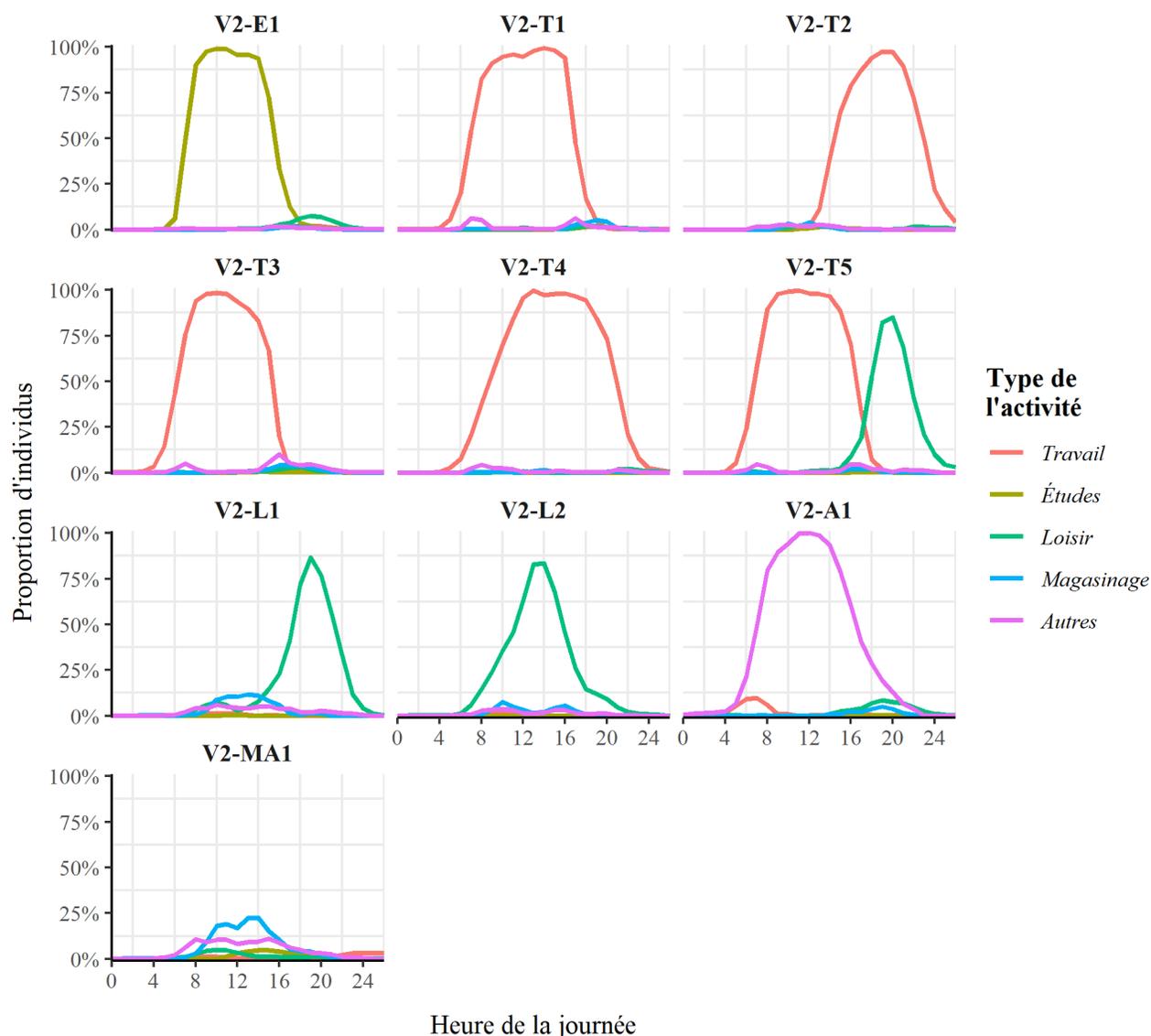


Figure 4-14 Proportion d'individus, à l'intérieur de chaque groupe, réalisant un des cinq types d'activités selon l'heure de la journée

Outre les groupes où le *travail* et les *loisirs* sont l'activité dominante, la figure ci-haut n'apporte pas d'informations supplémentaires sur les autres groupes. Presque tous les membres de V2-E1 et V2-A1 font leur activité dominante au milieu de la journée. Quant à V2-MA1, la figure ci-haut confirme que ses membres réalisent des activités de *magasinage* et de type *autres* de courtes durées et ce, tout au long de la journée.

Les individus composant V2-T5 ont un système d'activités très similaire à celui du groupe V1-T2, issue de la première segmentation, puisque l'activité secondaire de type *loisirs* est réalisée en soirée.

La Figure 4-14 permet également de constater la particularité du système d'activités des membres de V2-T2. En effet, la majorité de ces personnes font leur activité dominante, soit le *travail*, en soirée, entre 18h et 21h. Les trois autres groupes dont la principale activité est le *travail* semblent similaires selon la Figure 4-14. La Figure 4-15, qui illustre la distribution selon l'heure de début de l'activité dominante, permet de comprendre ce qui les différencie.

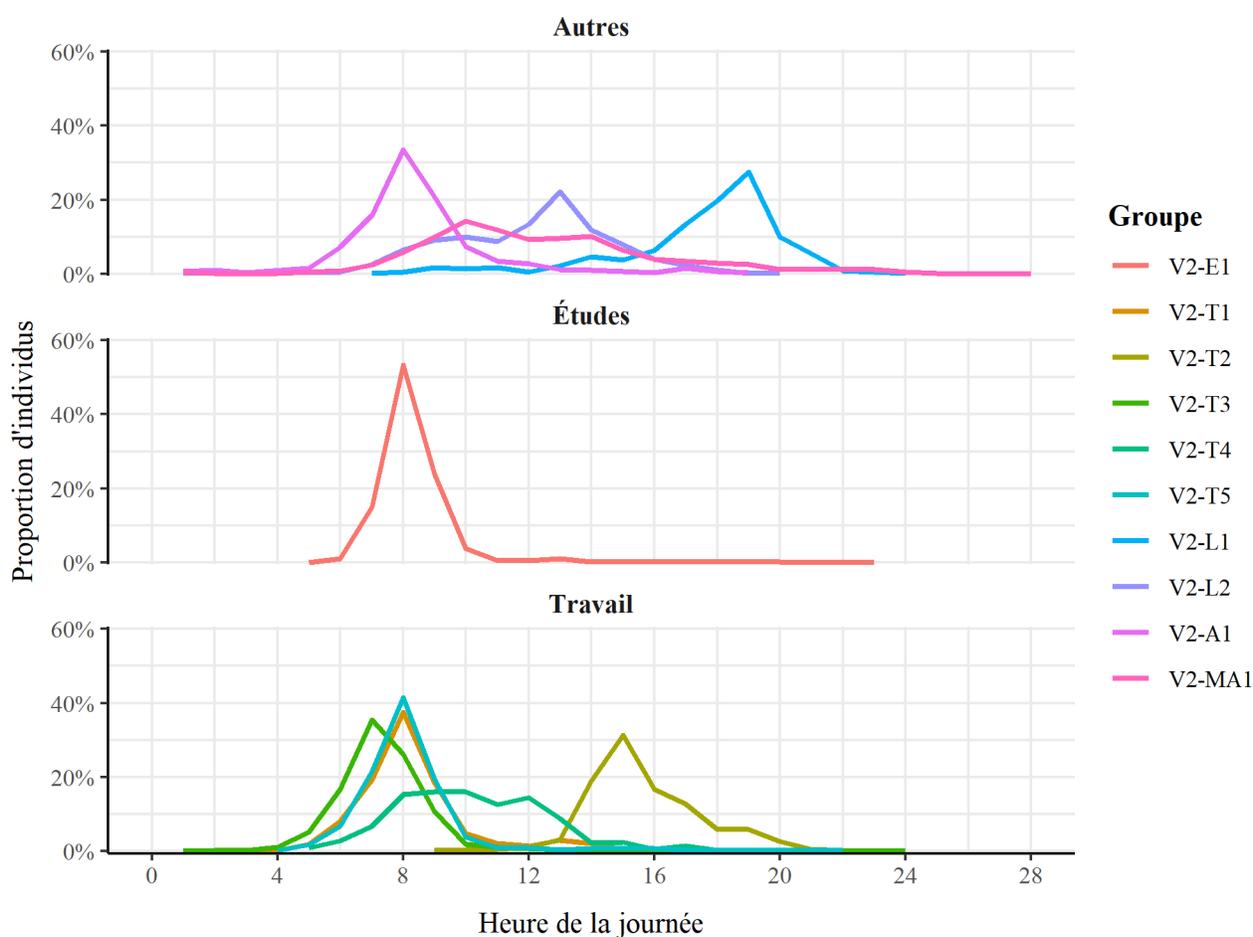


Figure 4-15 Proportion d'individus selon l'heure de départ de l'activité dominante

Ainsi, la différence entre V2-T1, V2-T3 et V2-T4, outre la durée totale, est l'heure de début de l'activité dominante. En effet, la pointe concernant la distribution selon l'heure de début est située entre 5h et 6h pour V2-T3, alors celle de V2-T1, tout comme V2-T5, est plus tard dans la matinée, soit entre 7h et 8h. La distribution de l'heure de début pour V2-T4, dont les membres consacrent plus de temps au *travail*, a une pointe plus étalée, soit entre 7h et 12h. De plus, la Figure 4-15 confirme également que les membres de V2-T2 se rendent au *travail* le soir.

Quant aux autres groupes, la distribution selon l'heure de début permet de confirmer les observations découlant de l'analyse de la Figure 4-14. En effet, la majorité des membres de V2-E1 et V2-A1 quittent leur domicile entre 7h et 8h tandis que les groupes dont les *loisirs* est l'activité dominante ont des pointes similaires à celles observées à la Figure 4-14, mais décalées un peu plus tôt dans la journée, puisque ce sont des heures de début. Finalement, les heures auxquelles débute la principale activité des membres de V2-MA1 sont étalées tout au long de la journée, ce qui confirme les observations obtenues précédemment.

4.3.2 Analyse des distributions

Le Tableau 4-3 présente les résultats des tests de Kolmogorov-Smirnov menés sur les paires de distributions des fréquences de réalisation des activités. Ainsi, les groupes sont comparés deux à deux sur la base de la distribution de la fréquence des activités et ce pour chaque type d'activité.

Tableau 4-3 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité – E1, T1 et T2

Groupe comparés	<i>Travail</i>	<i>Études</i>	<i>Loisir</i>	<i>Magasinage</i>	<i>Autres</i>
V2-E1 & V2-T1	D	D	D	D	D
V2-E1 & V2-T2	D	D	N	N	N
V2-E1 & V2-T3	D	D	D	D	D
V2-E1 & V2-T4	D	D	D	I	D
V2-E1 & V2-T5	D	D	D	I	D
V2-E1 & V2-L1	I	D	D	D	D
V2-E1 & V2-L2	I	D	D	D	D
V2-E1 & V2-A1	D	D	I	D	D
V2-E1 & V2-MA1	D	D	N	D	D
V2-T1 & V2-T2	I	I	I	I	N
V2-T1 & V2-T3	N	I	I	I	D
V2-T1 & V2-T4	D	I	I	N	N
V2-T1 & V2-T5	I	I	D	D	I
V2-T1 & V2-L1	D	I	D	D	D
V2-T1 & V2-L2	D	I	D	D	I
V2-T1 & V2-A1	D	I	D	I	D
V2-T1 & V2-MA1	D	D	D	D	D
V2-T2 & V2-T3	I	I	I	N	D
V2-T2 & V2-T4	D	I	I	I	I
V2-T2 & V2-T5	I	I	D	I	N

D	: $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I	: $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N	: $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau 4-4 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité – T2, T3, T4, T5, L1, L2 et A1

Groupe comparés	Travail	Études	Loisir	Magasinage	Autres
V2-T2 & V2-L1	D	I	D	D	D
V2-T2 & V2-L2	D	I	D	D	N
V2-T2 & V2-A1	D	I	N	I	D
V2-T2 & V2-MA1	D	N	N	D	D
V2-T3 & V2-T4	D	I	I	D	D
V2-T3 & V2-T5	N	I	D	D	N
V2-T3 & V2-L1	D	I	D	D	N
V2-T3 & V2-L2	D	I	D	D	D
V2-T3 & V2-A1	D	I	D	I	D
V2-T3 & V2-MA1	D	D	D	D	D
V2-T4 & V2-T5	D	I	D	I	N
V2-T4 & V2-L1	D	I	D	D	D
V2-T4 & V2-L2	D	I	D	D	I
V2-T4 & V2-A1	D	I	D	N	D
V2-T4 & V2-MA1	D	N	D	D	D
V2-T5 & V2-L1	D	I	N	D	N
V2-T5 & V2-L2	D	I	I	D	I
V2-T5 & V2-A1	D	I	D	N	D
V2-T5 & V2-MA1	D	D	D	D	D
V2-L1 & V2-L2	I	I	I	D	D
V2-L1 & V2-A1	D	I	D	D	D
V2-L1 & V2-MA1	N	N	D	D	D
V2-L2 & V2-A1	D	I	D	N	D
V2-L2 & V2-MA1	N	D	D	D	D
V2-A1 & V2-MA1	D	N	I	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

La majorité des conclusions du résultat des tests confirme que les distributions sont différentes. Les résultats indiquant que l'hypothèse nulle ne peut être acceptée confirment également les caractéristiques qui distinguent les groupes entre eux. Par exemple, la majorité des groupes dont l'activité dominante n'est pas les *études* ont des distributions identiques concernant celles indiquant la fréquence d'activité de type *études*.

Les tests pour les deux autres attributs, soit l'heure de début et la durée des activités, permettent également de conclure que les comportements observés d'un groupe à l'autre sont

significativement différents. La conclusion de ces tests est présentée dans les Tableau A-5 et Tableau A-7 de l'Annexe A.

4.4 Troisième segmentation

L'application de l'algorithme hiérarchique sur le vecteur V3 permet d'obtenir un dendrogramme, tout comme l'application de cet algorithme sur le vecteur V2. Ainsi, les mêmes constats sont valables pour le dendrogramme obtenu avec V3, tel qu'il sera possible de le constater dans la section suivante. De plus, la démarche décrite à la section précédente s'applique également aux résultats obtenus avec V3 afin d'identifier le nombre optimal de groupes.

4.4.1 Analyse des groupes

La coupure du dendrogramme menant à la création de trois groupes mène aux mêmes résultats que la segmentation avec V2. Cette coupure permet uniquement de partitionner le sous-échantillon provenant des quatre enquêtes OD selon l'activité dominante des individus dont la principale activité est le *travail* ou les *études*. L'ensemble des autres individus sont par conséquent regroupés dans un autre groupe. Après l'analyse des groupes issue de coupures à différents niveaux de dissimilarité, le même nombre de groupes que la segmentation précédente est choisi, soit dix. La distribution du pourcentage de personnes dans chacun des segments est présentée dans la Figure 4-16.

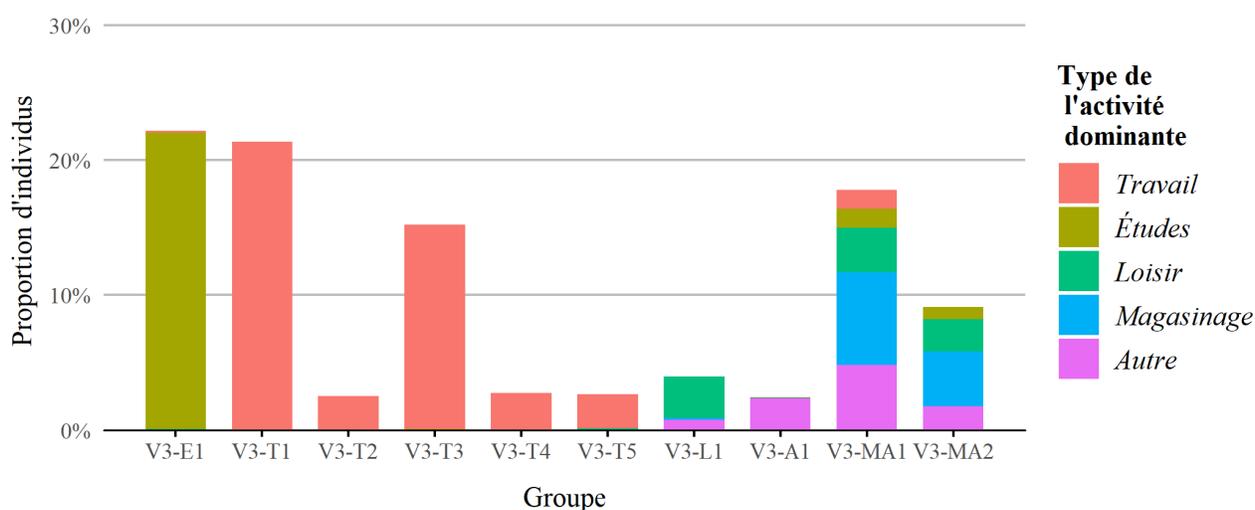


Figure 4-16 Distribution de l'échantillon selon le groupe

Les résultats sont similaires à ceux obtenus avec la segmentation du vecteur V2. Un seul groupe contient l'ensemble des élèves et étudiants. Cinq groupes rassemblent les personnes dont la principale activité est le *travail* et les proportions observées dans chacun d'entre eux sont similaires à celles observées dans la segmentation précédente. La principale différence entre les deux partitionnements concerne les quatre derniers groupes. Alors que celle avec V2 ne parvient pas à scinder de façon significative le groupe contenant la majorité des personnes ayant une activité dominante autre que le *travail* et les *études*, soit V2-MA1, la segmentation avec V3 réussit à le faire. En effet, deux groupes ont maintenant des membres avec une multitude d'activités dominantes. Deux autres groupes sont également créés avec cette segmentation, dont le type de l'activité dominante est le *loisir* pour V3-L1 et *autre* pour V3-A1. De plus, aucune des deux segmentations hiérarchiques ne parvient à créer un groupe dont la majorité des membres ont comme activité principale le *magasinage*. La durée totale et la fréquence moyenne de chacun des groupes sont présentées à la Figure 4-17.

Les attributs des cinq groupes où le *travail* est la principale activité sont similaires à ceux issus du partitionnement avec V2. Seuls les membres de V3-T5 réalisent une activité secondaire alors que c'est la durée totale qui distingue les autres groupes. Cette durée est plus élevée pour V3-T4, et plus faible pour V3-T2 et V3-T3.

La Figure 4-17 ne permet pas de décrire davantage V3-L1 et V3-A1. Cependant, la fréquence moyenne de 1,5 pour les membres de V3-A1 indique qu'une certaine portion d'entre eux font une activité par jour, mais une portion significative de personnes réalise également deux activités et plus par jour.

La principale différence entre V3-MA1 et V3-MA2 semble être la durée totale, qui est plus faible pour V3-MA1 que pour V3-MA2. La Figure 4-18, qui illustre la proportion de personnes réalisant une activité selon l'heure de la journée, permet de comprendre davantage les différences qui existent entre V3-MA1 et V3-MA2.

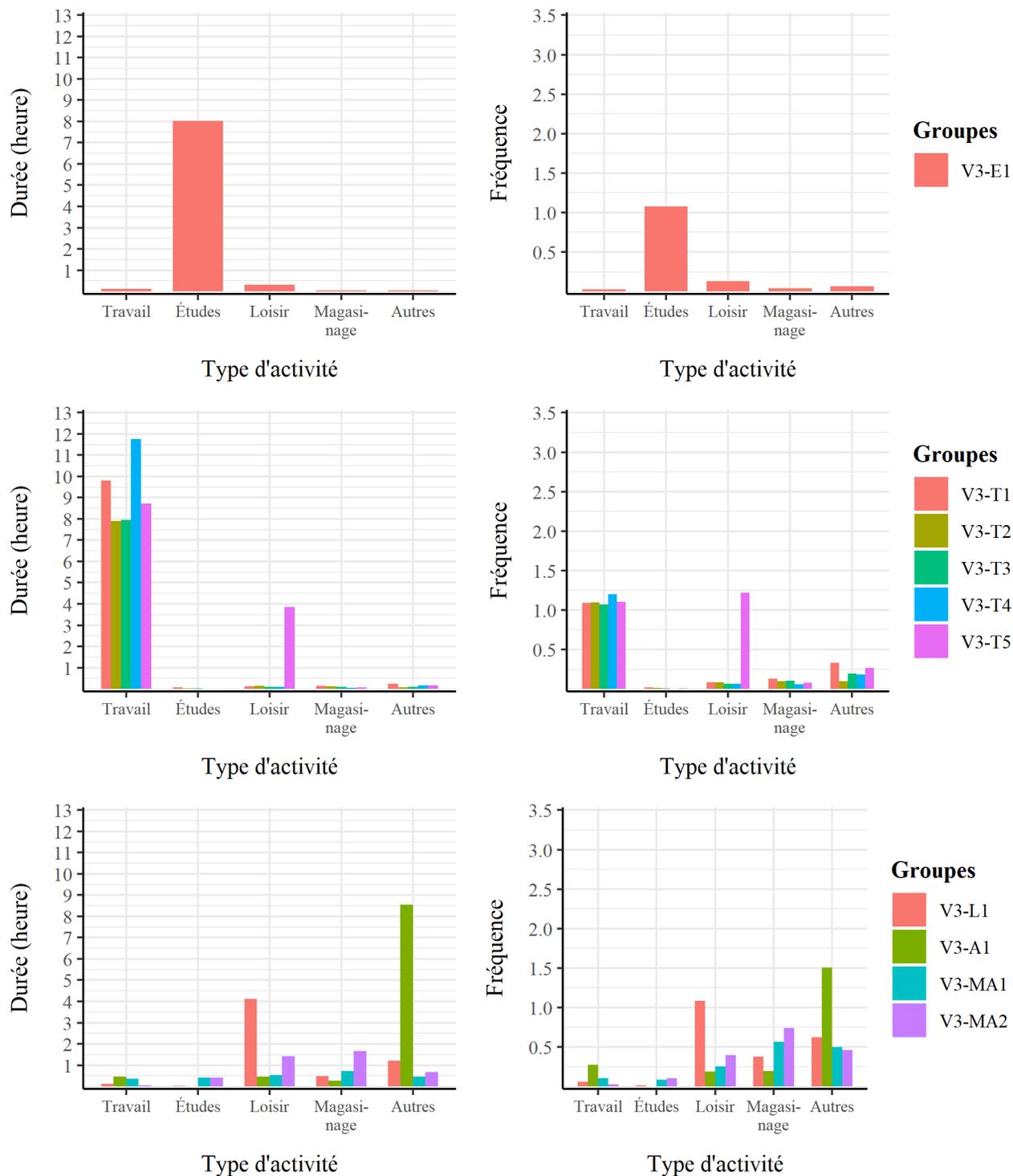


Figure 4-17 Centre des groupes de la segmentation avec V3

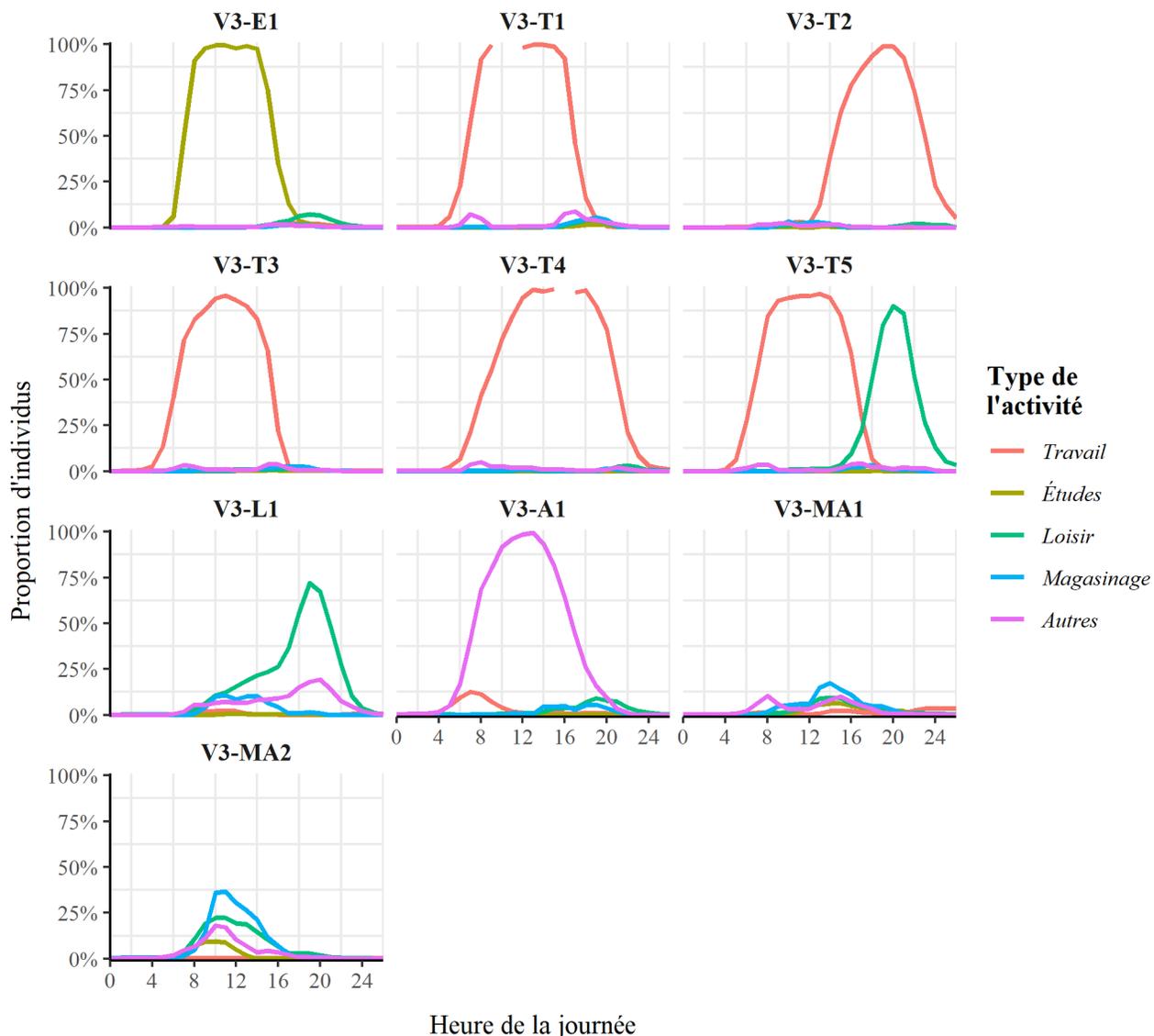


Figure 4-18 Proportion d'individus, à l'intérieur de chaque groupe, réalisant un des cinq types d'activités selon l'heure de la journée

La proportion de personnes réalisant les différents types d'activités dans le cas de V3-MA1 est plus faible que V3-MA2, ce qui signifie que les membres du premier groupe réalisent des activités tout au long de la journée, contrairement au second, où ils sont très nombreux à réaliser des activités au même moment, entre 8h et 13h, avec une pointe à 10h.

De plus, la majorité des membres de V3-L1 réalisent une activité de type *loisir* principalement le soir, entre 18h et 20h. La Figure 4-18 permet de constater que les autres groupes sont similaires à ceux issus du partitionnement avec le vecteur V2. En effet, les membres des cinq groupes dont le

travail est l'activité dominante semblent avoir le même patron d'activités que les cinq groupes issus de la deuxième segmentation, tout comme V3-E1 et V3-A1, qui sont similaires à V2-E1 et V2-A1. La Figure 4-19 confirme ces similarités. Elle illustre la distribution du pourcentage de personnes selon l'heure à laquelle débute l'activité dominante.

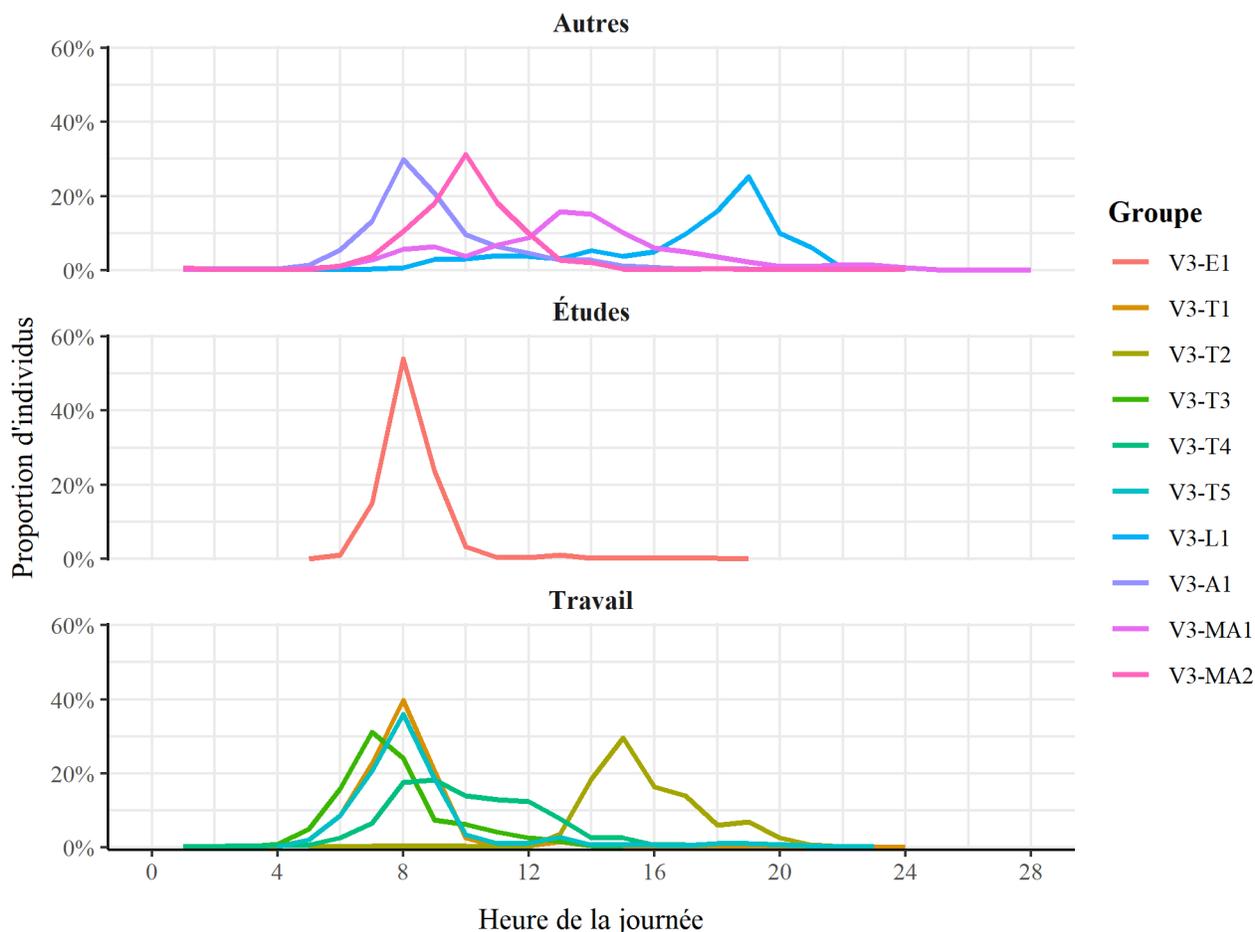


Figure 4-19 Heure de départ de l'activité dominante

La distribution selon les heures de départ confirme la similarité entre les deux segmentations hiérarchiques concernant les groupes où le *travail*, les *études* et le type *autre* sont l'activité dominante. En effet, l'élément qui différencie les groupes où le *travail* domine est non seulement la durée totale mais également la pointe de la distribution de l'heure de début, qui est étalée pour V3-T4, plus tôt pour V3-T3 et en soirée pour V3-T2.

Toutefois, la Figure 4-19 apporte un éclairage différent sur la structure du système d'activités des membres de V3-MA2. En effet, ces derniers sont nombreux, à la hauteur de 30 %, à débiter leur

activité entre 9h et 10h, alors que les membres de V3-MA1 débutent leurs activités à toute heure de la journée, avec une petite pointe entre 12h et 14h. Bref, les individus composant V3-MA2 semblent réaliser des activités de longue durée qui débutent en matinée, alors que les membres de V3-MA1 réalisent plutôt des activités de courtes durées et ce peu importe le moment de la journée. Ainsi, les différences entre les deux groupes concernent davantage la structure temporelle des systèmes d'activités de leurs membres que la fréquence de réalisation des différents types d'activités. Ce constat met en évidence la force de la segmentation avec le vecteur V3 et de façon plus générale des deux partitionnements hiérarchiques, par rapport à la segmentation avec le premier vecteur (V1). En effet, les vecteurs V2 et V3, ainsi que les distances qui permettent de mesurer leurs similarités, permettent de capter des patrons davantage caractérisés par la durée et l'heure de début. Les patrons détectés par la première segmentation (V1) sont plutôt caractérisés par la fréquence de réalisation de différents types d'activité et dans une moindre mesure, la durée des activités.

4.4.2 Analyse des distributions

Tout comme les deux premières segmentations, les distributions des attributs des activités sont comparées entre chacun des groupes afin de s'assurer de la pertinence des groupes créés. La conclusion du résultat des tests sur les distributions des fréquences est présentée dans le Tableau 4-5.

Tableau 4-5 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité – E1 et T1

Groupe comparés	<i>Travail</i>	<i>Études</i>	<i>Loisir</i>	<i>Magasinage</i>	<i>Autres</i>
V3-E1 & V3-T1	D	D	D	D	D
V3-E1 & V3-T2	D	D	N	N	N
V3-E1 & V3-T3	D	D	D	D	D
V3-E1 & V3-T4	D	D	N	I	D
V3-E1 & V3-T5	D	D	D	N	D
V3-E1 & V3-L1	N	D	D	D	D
V3-E1 & V3-A1	D	D	N	D	D
V3-E1 & V3-MA1	D	D	D	D	D
V3-E1 & V3-MA2	I	D	D	D	D

D	: $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I	: $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N	: $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau 4-6 Résultats des tests de Kolmogorov-Smirnov sur les distributions des fréquences de réalisation des différents types d'activité T1, T2, T3, T4, L1, L2, M1, A1 et MA1

Groupe comparés	Travail	Études	Loisir	Magasinage	Autres
V3-T1 & V3-T2	I	I	I	N	D
V3-T1 & V3-T3	N	I	N	N	D
V3-T1 & V3-T4	D	I	I	N	D
V3-T1 & V3-T5	I	I	D	N	N
V3-T1 & V3-L1	D	I	D	D	D
V3-T1 & V3-A1	D	I	D	N	D
V3-T1 & V3-MA1	D	D	D	D	D
V3-T1 & V3-MA2	D	D	D	D	D
V3-T2 & V3-T3	I	I	I	I	N
V3-T2 & V3-T4	D	I	I	I	N
V3-T2 & V3-T5	I	I	D	I	D
V3-T2 & V3-L1	D	I	D	D	D
V3-T2 & V3-A1	D	I	N	N	D
V3-T2 & V3-MA1	D	D	D	D	D
V3-T2 & V3-MA2	D	D	D	D	D
V3-T3 & V3-T4	D	I	I	N	I
V3-T3 & V3-T5	I	I	D	I	N
V3-T3 & V3-L1	D	I	D	D	D
V3-T3 & V3-A1	D	I	D	D	D
V3-T3 & V3-MA1	D	D	D	D	D
V3-T3 & V3-MA2	D	D	D	D	D
V3-T4 & V3-T5	N	I	D	I	N
V3-T4 & V3-L1	D	I	D	D	D
V3-T4 & V3-A1	D	I	D	D	D
V3-T4 & V3-MA1	D	D	D	D	D
V3-T4 & V3-MA2	D	D	D	D	D
V3-T5 & V3-L1	D	I	D	D	D
V3-T5 & V3-A1	D	I	D	D	D
V3-T5 & V3-MA1	D	D	D	D	D
V3-T5 & V3-MA2	D	D	D	D	D
V3-L1 & V3-A1	D	I	D	D	D
V3-L1 & V3-MA1	N	D	D	D	N
V3-L1 & V3-MA2	N	D	D	D	D
V3-A1 & V3-MA1	D	D	N	D	D
V3-A1 & V3-MA2	D	D	D	D	D
V3-MA1 & V3-MA2	D	I	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes

I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques

N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tout comme les résultats avec les deux premières segmentations, la majorité des tests permettent de conclure que la distribution de la fréquence est significativement différente entre les différents groupes. Les tests indiquant une conclusion différente concernent entre autres les comparaisons entre les distributions du type d'activité *études* du groupe V1-E1 et les autres groupes.

La conclusion des tests appliqués aux distributions concernant l'heure de début et la durée des activités indique également que les comportements observés dans chacun des groupes sont significativement différents. La conclusion de ces tests est présentée dans les Tableau A-9 et Tableau A-11 de l'Annexe A.

CHAPITRE 5 MODÈLE D'ALLOCATION DES PATRONS D'ACTIVITÉS AUX INDIVIDUS

Le but de ce chapitre est de développer un modèle permettant de modéliser les liens entre les caractéristiques sociodémographiques des individus et les patrons typiques qu'ils risquent d'adopter. Préalablement à cet exercice de modélisation, il est nécessaire de caractériser les membres de chacun des groupes issus des partitionnements réalisés au chapitre précédent.

5.1 Description sociodémographique des groupes

L'analyse sociodémographique des groupes se fait en deux temps. Tout d'abord, pour chacune des trois segmentations, la description des membres de chacun des groupes est faite selon les caractéristiques individuelles, soit le genre, l'âge et le statut d'occupation. Ensuite, les membres sont décrits selon les attributs des ménages auxquels ceux-ci appartiennent, soit le type de ménage et la région du domicile.

5.1.1 Première segmentation (V1)

Les descriptions sociodémographiques des groupes issus du premier partitionnement (V1) sont présentées dans le Tableau 5-1. Il indique la proportion de personnes dans chacune des catégories des trois caractéristiques sociodémographiques, soit le genre, le statut d'occupation et l'âge.

Les personnes composant les deux groupes où l'*étude* est le type de l'activité dominante, soit V1-E1 et V1-E2, sont presque exclusivement des étudiants âgés de 30 ans et moins, majoritairement entre 10 et 20 ans. Les individus présents dans V1-E2 semblent être plus âgés et également plus nombreux à être des hommes.

Les quatre groupes dont l'activité dominante est le travail semblent également être composés d'individus avec les mêmes caractéristiques. Entre 86 et 90 % d'entre eux sont des travailleurs à temps plein et certains d'entre eux le sont également, mais à temps partiel. Toutefois, l'âge des membres des quatre groupes semble différent. En effet, les individus composant V1-T2 semblent légèrement plus jeunes, alors que 77 % d'entre eux ont entre 20 et 49 ans. La différence la plus marquée concerne le groupe V1-T4, dont 70 % des membres sont âgées de 30 et 49 ans. De plus, les hommes sont majoritaires dans l'ensemble des groupes où la principale activité est le *travail*, à l'exception de celui où l'activité secondaire est le *magasinage*, soit V1-T3.

Tableau 5-1 Caractéristiques sociodémographiques des individus de chacun des groupes

	V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-L1	V1-L2	V1-M1	V1-A1
Genre										
Homme	50%	53%	55%	55%	49%	54%	47%	42%	40%	53%
Statut										
Travailleur à temps plein	1%	1%	90%	87%	86%	90%	20%	22%	19%	41%
Travailleur à temps partiel	1%	1%	6%	7%	9%	6%	6%	7%	6%	7%
Étudiant	98%	97%	2%	3%	2%	1%	10%	6%	4%	6%
Retraité	0%	1%	1%	2%	3%	1%	50%	51%	55%	31%
Autre	0%	1%	1%	1%	1%	1%	14%	14%	16%	15%
Âge										
5-9	34%	23%	0%	0%	0%	0%	2%	1%	1%	2%
10-19	50%	57%	3%	4%	1%	1%	7%	3%	2%	3%
20-29	12%	16%	20%	26%	16%	14%	12%	10%	8%	11%
30-39	2%	2%	25%	25%	26%	39%	10%	13%	11%	20%
40-49	1%	1%	29%	26%	31%	31%	10%	12%	12%	19%
50-59	0%	1%	19%	15%	19%	12%	16%	18%	18%	16%
60-69	0%	0%	4%	4%	5%	3%	23%	26%	26%	15%
70-79	0%	0%	1%	1%	1%	0%	15%	15%	18%	10%
80-89	0%	0%	0%	0%	0%	0%	4%	3%	5%	3%
90+	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Les groupes où l'activité dominante est le *magasinage* ou le *loisir* sont constitués principalement de personnes âgées entre 50 et 79 ans, et plus de la moitié d'entre eux sont retraités. En outre, ces groupes sont majoritairement composés de femmes, alors que la proportion d'hommes varie de 47 à 40 %. Finalement, le dernier groupe (V1-A1) semble contenir une variété de personnes, alors que la distribution selon l'âge est étalée et que les proportions de travailleurs et de retraités sont toutes deux considérables, avec 41 et 31 % d'individus de chaque statut respectivement.

Le Tableau 5-2 présente les caractéristiques des ménages auxquels les membres des différents groupes appartiennent.

Tableau 5-2 Caractéristiques des ménages auxquels chaque individu appartient

	V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-L1	V1-L2	V1-M1	V1-A1
Type du ménage										
Jeunes couples, colocataires et célibataire	6%	10%	19%	27%	21%	10%	10%	10%	7%	8%
Jeunes familles	66%	57%	27%	25%	30%	57%	16%	18%	17%	32%
Familles traditionnelles	21%	27%	25%	22%	19%	16%	14%	10%	11%	15%
Adultes et personnes âgées	0%	0%	3%	2%	2%	1%	5%	4%	6%	5%
Cinquantennaires	0%	1%	16%	16%	19%	9%	12%	13%	12%	11%
Jeunes retraités	0%	0%	5%	5%	6%	3%	27%	30%	29%	16%
Vieux retraités	0%	0%	0%	0%	1%	0%	13%	12%	15%	8%
Familles multi-générationnelles	6%	5%	4%	3%	3%	4%	3%	2%	4%	5%
Région du domicile										
Centre-ville	1%	2%	1%	2%	2%	1%	2%	2%	1%	1%
Montréal-centre	19%	20%	20%	22%	20%	18%	21%	20%	19%	20%
Montréal-est	6%	5%	6%	4%	5%	6%	6%	5%	6%	6%
Montréal-ouest	10%	9%	9%	8%	8%	10%	9%	9%	9%	9%
Laval	7%	7%	8%	7%	7%	8%	7%	8%	8%	7%
Longueuil	8%	6%	8%	6%	7%	8%	6%	6%	7%	7%
Couronne nord	15%	13%	15%	11%	11%	15%	13%	12%	14%	15%
Couronne sud	13%	11%	13%	11%	10%	13%	12%	11%	11%	12%

La région où les personnes résident semble avoir peu d'influence sur l'appartenance à un des groupes. Les proportions sont relativement stables d'un groupe à l'autre. En revanche, les proportions varient énormément en ce qui concerne les types de ménage. Sans surprise, les personnes appartenant aux « Jeunes retraités » représentent un peu moins du tiers des individus composant les groupes où l'activité dominante est le *magasinage* ou le type *loisir*.

Les membres de V1-E1 et V1-E2 appartiennent principalement aux « Jeunes familles » et aux « Familles traditionnelles » et, dans une moindre mesure, aux « Familles multigénérationnelles » et « Jeunes couples, colocataires et célibataire ».

Les trois premiers groupes où le *travail* est l'activité dominante sont composées essentiellement d'individus appartenant aux trois premiers types de ménage ainsi qu'aux « Cinquantenaires ». Toutefois, plus de 57 % des membres de V1-T4 vivent dans des ménages de type « Jeunes familles ».

5.1.2 Deuxième segmentation (V2)

Les caractéristiques individuelles des membres des groupes issus de la segmentation de l'échantillon des enquêtes OD avec V2 sont présentées dans le Tableau 5-3.

Tableau 5-3 Caractéristiques sociodémographiques des individus de chacun des groupes

	V2-E1	V2-T1	V2-T2	V2-T3	V2-T4	V2-T5	V2-L1	V2-L2	V2-A1	V2-MA1
Genre										
Homme	50%	55%	59%	50%	59%	58%	43%	43%	62%	47%
Statut										
Travailleur à temps plein	1%	92%	80%	90%	86%	92%	25%	15%	58%	29%
Travailleur à temps partiel	0%	5%	12%	6%	8%	5%	8%	5%	5%	6%
Étudiant	98%	2%	6%	2%	4%	3%	13%	9%	9%	10%
Retraité	0%	1%	1%	1%	1%	0%	40%	59%	21%	40%
Autre	0%	1%	1%	1%	1%	0%	15%	13%	8%	14%
Âge										
5-9	35%	0%	0%	0%	0%	0%	2%	2%	3%	1%
10-19	52%	2%	9%	2%	4%	3%	8%	5%	4%	5%
20-29	10%	18%	26%	17%	29%	27%	13%	10%	13%	13%
30-39	2%	28%	23%	26%	27%	26%	13%	8%	22%	15%
40-49	1%	30%	25%	31%	24%	28%	12%	8%	23%	14%
50-59	0%	17%	13%	19%	12%	13%	17%	16%	16%	16%
60-69	0%	3%	3%	4%	4%	3%	21%	28%	12%	20%
70-79	0%	0%	1%	1%	1%	0%	12%	18%	5%	13%
80-89	0%	0%	0%	0%	0%	0%	1%	5%	2%	3%
90+	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Puisqu'un seul groupe où l'*étude* est l'activité dominante est issu du deuxième partitionnement, peu d'analyses peuvent être effectuées sur ses membres, surtout que la majorité d'entre eux ont entre 10 et 19 ans et sont pratiquement tous des étudiants.

Pour ce qui est des groupes dont le *travail* est la principale activité des membres, une différence notable est la proportion de travailleurs à temps partiel dans V2-T2. Ce groupe est caractérisé par

une distribution des heures de départ dont la pointe est en après-midi, ce qui indique que ces travailleurs sont susceptibles d'avoir des horaires atypiques. Les membres de ce groupe, tout comme ceux de V2-T4 et V2-T5, semblent également plus jeunes que les personnes composant V2-T1 et V2-T3. De plus, à l'exception de V2-T3, les hommes sont tous plus nombreux dans les groupes où le *travail* est l'activité dominante.

C'est cependant l'inverse dans le cas des groupes où les retraités sont les plus nombreux, soit V1-L1, V1-L2 et V2-MA1. Ces derniers sont composés majoritairement de femmes et leurs membres sont plus âgées, bien que la distribution selon l'âge soit étalée. Le Tableau 5-4 présente les attributs des ménages auxquels font partie les membres de chacun des groupes.

Tableau 5-4 Caractéristiques des ménages auxquels chaque individu appartient

	V2-E1	V2-T1	V2-T2	V2-T3	V2-T4	V2-T5	V2-L1	V2-L2	V2-A1	V2-MA1
Type du ménage										
Jeunes couples, colocataires et célibataire	5%	19%	20%	14%	21%	27%	14%	7%	11%	9%
Jeunes familles	68%	32%	29%	36%	30%	25%	20%	14%	32%	23%
Familles traditionnelles	21%	23%	25%	23%	26%	24%	13%	11%	17%	15%
Adultes et personnes âgées	0%	3%	4%	3%	2%	3%	4%	6%	5%	5%
Cinquantennaires	0%	16%	11%	16%	11%	15%	12%	12%	13%	11%
Jeunes retraités	0%	4%	5%	5%	5%	4%	25%	34%	13%	23%
Vieux retraités	0%	0%	0%	0%	1%	0%	9%	15%	5%	10%
Familles multigénérationnelles	5%	3%	5%	3%	4%	3%	3%	3%	5%	4%
Région du domicile										
Centre-ville	1%	2%	0%	1%	1%	3%	2%	1%	1%	2%
Montréal-centre	19%	20%	25%	17%	26%	22%	19%	20%	16%	20%
Montréal-est	6%	5%	7%	6%	3%	5%	8%	6%	9%	6%
Montréal-ouest	10%	9%	6%	8%	7%	7%	10%	8%	11%	9%
Laval	7%	8%	6%	8%	6%	8%	8%	7%	6%	8%
Longueuil	7%	8%	8%	8%	9%	6%	5%	8%	7%	8%
Couronne nord	15%	13%	12%	17%	14%	12%	13%	14%	17%	14%
Couronne sud	12%	12%	10%	15%	10%	9%	10%	11%	10%	12%

Les proportions d'individus vivant dans les différentes régions de la GRM semblent varier davantage que celles des groupes de la première segmentation (V1). Les personnes composant V2-T3 sont plus nombreuses à vivre en couronne que les autres groupes. De plus, la distribution du

pourcentage de personnes selon la région pour les membres de V1-A1 semble quelque peu différente des autres groupes.

Pour les types de ménages, les personnes dont la principale activité est l'études vivent principalement dans les « Jeunes familles », alors les groupes V2-L1 et V2-L2 comptent davantage de retraités. Dans le cas des autres groupes, les distributions semblent similaires.

5.1.3 Deuxième segmentation (V3)

La composition sociodémographique des groupes issus de la troisième segmentation est similaire à celle des groupes du partitionnement précédent, tel qu'il est possible de le constater dans le Tableau 5-5.

Tableau 5-5 Caractéristiques sociodémographiques des individus de chacun des groupes

	V3-E1	V3-T1	V3-T2	V3-T3	V3-T4	V3-T5	V3-L1	V3-A1	V3-MA1	V3-MA2
Genre										
Homme	50%	54%	60%	51%	60%	59%	42%	62%	50%	42%
Statut										
Travailleur à temps plein	1%	95%	81%	88%	86%	90%	27%	57%	30%	18%
Travailleur à temps partiel	0%	3%	12%	7%	8%	6%	7%	5%	6%	5%
Étudiant	99%	1%	4%	2%	4%	4%	11%	8%	14%	14%
Retraité	0%	0%	1%	2%	1%	0%	38%	22%	37%	50%
Autre	0%	0%	1%	1%	1%	0%	18%	8%	13%	13%
Âge										
5-9	36%	0%	0%	0%	0%	0%	1%	3%	1%	2%
10-19	52%	2%	8%	3%	4%	4%	7%	3%	7%	7%
20-29	9%	19%	26%	16%	31%	26%	15%	13%	14%	10%
30-39	2%	30%	23%	25%	26%	26%	14%	22%	14%	11%
40-49	1%	29%	25%	31%	24%	28%	12%	21%	14%	11%
50-59	0%	17%	14%	20%	11%	13%	15%	19%	16%	16%
60-69	0%	3%	2%	4%	4%	4%	22%	12%	18%	25%
70-79	0%	0%	1%	1%	1%	0%	12%	5%	13%	14%
80-89	0%	0%	0%	0%	0%	0%	1%	2%	3%	5%
90+	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

La différence majeure entre les deux segmentations hiérarchiques (V2 et V3) est l'apparition d'un deuxième groupe composé de personnes dont la principale activité est multiple (V3-MA1 et V3-

MA2). La différence entre ces groupes, au niveau de la composition sociodémographique, concerne la présence plus marquée de retraités dans V3-MA2, alors que V3-MA1 semble contenir une variété de personnes, avec un nombre considérable de travailleurs tout comme de retraités. Cette différence est également observable dans la composition des groupes au niveau des types de ménages dont les membres font partie. Cette composition est présentée dans le Tableau 5-6.

Tableau 5-6 Caractéristiques des ménages auxquels chaque individu appartient

	V3-E1	V3-T1	V3-T2	V3-T3	V3-T4	V3-T5	V3-L1	V3-A1	V3-MA1	V3-MA2
Type du ménage										
Jeunes couples, colocataires et célibataire	5%	19%	20%	14%	23%	26%	15%	10%	10%	7%
Jeunes familles	69%	33%	30%	34%	29%	23%	21%	31%	23%	20%
Familles traditionnelles	20%	23%	25%	23%	25%	25%	14%	17%	16%	14%
Adultes et personnes âgées	0%	3%	5%	3%	3%	2%	4%	4%	5%	5%
Cinquantennaires	0%	15%	10%	17%	12%	15%	10%	15%	12%	11%
Jeunes retraités	0%	4%	5%	5%	5%	4%	26%	14%	21%	26%
Vieux retraités	0%	0%	1%	0%	1%	0%	8%	5%	9%	14%
Familles multigénérationnelles	5%	4%	5%	3%	4%	4%	3%	4%	4%	3%
Région du domicile										
Centre-ville	1%	2%	0%	1%	1%	3%	2%	1%	2%	2%
Montréal-centre	18%	20%	26%	18%	24%	23%	19%	16%	21%	20%
Montréal-est	5%	6%	7%	6%	3%	5%	7%	8%	5%	7%
Montréal-ouest	10%	9%	5%	8%	8%	6%	9%	11%	9%	9%
Laval	7%	8%	6%	9%	5%	7%	9%	6%	8%	7%
Longueuil	7%	8%	8%	8%	9%	6%	5%	7%	8%	8%
Couronne nord	16%	13%	12%	16%	14%	14%	12%	17%	13%	15%
Couronne sud	12%	12%	11%	15%	10%	10%	9%	11%	12%	11%

Les membres de V3-MA2 sont en effet plus nombreux à être membres de ménages qualifiés de retraités, contrairement au membre de V3-MA1 où les membres des familles sont davantage présents.

Encore une fois, il n'existe aucune différence majeure par rapport au lieu de résidence des membres de chacun des dix groupes.

5.2 Description et transformation des variables

Les trois variables explicatives choisies afin d'expliquer le patron d'activités observé chez chaque individu sont l'âge, le genre et la région dans laquelle réside la personne. Seules ces trois variables sont utilisées puisque le modèle doit s'intégrer dans la démarche du MTQ. En effet, le modèle vise à être appliqué sur une population qui est projetée à divers horizons temporels. La première étape du processus en place au ministère a la capacité de prévoir uniquement certaines caractéristiques de cette population, soit l'âge, le genre et la région du domicile.

Le projet dans lequel s'inscrit ce mémoire inclut également un volet qui consiste à estimer un modèle permettant d'allouer un type de ménage à chaque individu. Les types sont ceux décrits à la section 3.2.3 du chapitre méthodologique. La population projetée peut ainsi avoir une variable additionnelle, soit le type de ménage auquel chaque individu appartient. Cependant, les variables explicatives utilisées pour prédire ce type sont également celles qui sont disponibles pour la population projetée, soit l'âge, le genre et la région du domicile. Les variables explicatives du modèle estimé dans le cadre de ce mémoire incluent donc une variable, le type de ménage, qui est prédite à partir des autres variables explicatives.

Les variables utilisées sont donc l'âge, le genre et le type de ménage auquel chaque individu appartient, de même que la région de domicile du ménage. La structure de ces données, incluant le type de chaque variable, est présentée dans le Tableau 5-7.

À la suite de l'estimation des modèles décrits ci-bas, il a été observé que la variable indiquant la région dans laquelle un individu réside n'a pas une influence significative sur le type de patron d'activité qu'il risque d'adopter (quand l'âge, le genre et le type de ménage sont considérés). C'est pourquoi cette variable a été exclue de l'exercice de modélisation dans le cadre de ce mémoire.

Tableau 5-7 Variables utilisées pour l'estimation du modèle

Variables explicatives		
Âge	Continue	0 à 99
Genre	Catégorielle	Homme Femme
Appartenance à un type de ménage	Catégorielle	Jeunes couples, colocataires et célibataire Jeunes familles Familles traditionnelles Adultes et personnes âgées Cinquantennaires Jeunes retraités Vieux retraités Familles multigénérationnelles
Région du domicile	Catégorielle	Centre-ville Montréal-centre Montréal-est Montréal-ouest Laval Longueuil Couronne nord Couronne sud
Variable expliquée		
Typologie des systèmes d'activités	Catégorielle	Résultats du Chapitre 4

Les variables choisies doivent être transformées afin de les intégrer dans un modèle de choix discret, mais peuvent également être transformées afin d'analyser les multiples effets que peuvent avoir ces variables. En effet, les liens qui unissent les caractéristiques sociodémographiques des individus et les patrons typiques d'activités peuvent prendre différentes formes. L'influence de l'âge, du genre et de la situation qui prévaut dans un ménage sur les activités réalisées lors d'une journée typique de semaine peut être directe, mais peut également se combiner entre les variables. C'est pourquoi cinq différents modèles sont testés afin de comprendre l'influence de ces caractéristiques sur la structure des systèmes d'activités.

Les deux premiers modèles se basent sur l'hypothèse que les trois variables ont un effet direct sur le type de patron d'activités d'un individu. Le Tableau 5-8 présente la forme des variables explicatives des modèles 1 et 2.

Tableau 5-8 Description des variables explicatives des modèles 1 et 2

Modèle 1			Modèle 2		
Âge	Genre	Appartenance à un type de ménage	Âge	Genre	Appartenance à un type de ménage
0 à 99	Homme Femme	Jeunes couples, colocataires et célibataire	5-9	Homme Femme	Jeunes couples, colocataires et célibataire
		Jeunes familles	10-14		Jeunes familles
	Familles traditionnelles	15-19	Familles traditionnelles		
	Adultes et personnes âgées	20-24	Adultes et personnes âgées		
	Cinquantennaires	25-29	Cinquantennaires		
	Jeunes retraités	30-34	Jeunes retraités		
	Vieux retraités	35-39	Vieux retraités		
	Familles multigénérationnelles	40-44	Familles multigénérationnelles		
		45-49			
		50-54			
	55-59				
	60-64				
	65-69				
	70-74				
	75-79				
	80-84				
	85-89				
	90+				

Le premier modèle n'implique aucune transformation des trois variables explicatives. L'âge est donc considéré comme une variable continue et le genre et le type de ménage sont des variables catégorielles. En considérant l'âge comme une variable continue, cela implique que l'effet de l'âge sur le risque d'adopter un des patrons typiques est linéaire. Autrement dit, à travers l'ensemble de la population, l'augmentation d'une année de l'âge a le même effet sur la variation de la probabilité d'adopter un des patrons.

Le deuxième modèle se différencie par la transformation de la variable de l'âge, qui devient également une variable discrète. En discrétisant l'âge, l'hypothèse que l'effet de l'âge est non linéaire est donc émise. Dix-huit catégories sont créées, qui correspondent à une division par bond de cinq ans de l'âge. Le choix de cette division repose sur l'analyse faite à la section 4.1, qui permet de constater qu'il existe une fluctuation importante de la durée et de la fréquence moyenne des activités selon l'âge des individus, catégorisé par bond de cinq ans.

Les trois autres modèles tentent d'explorer l'interaction qui existe entre ces trois variables explicatives sur les attributs des activités. L'interprétation de ces trois derniers modèles pourrait par exemple révéler que la variation de l'âge n'a pas le même effet sur la réalisation d'un certain patron pour les hommes et pour les femmes. De multiples interactions du genre peuvent exister et

être révélées par les trois modèles suivants. Chacun d'entre eux inclut deux variables, dont une qui est la combinaison de deux des trois variables originales et qui tente d'estimer l'effet de l'interaction entre ces deux variables. La variable d'interaction du modèle 3 est la combinaison de l'âge et du genre, alors que la variable d'interaction du quatrième modèle combine le genre et le type de ménage. Le dernier modèle combine les deux variables restantes, soit l'âge et le type de ménage.

Tableau 5-9 Description des variables explicatives des modèles 3 et 4

Modèle 3		Modèle 4	
Âge et genre	Appartenance à un type de ménage	Âge	Genre et appartenance à un type de ménage
Homme de	5-14 ans	5-9	Homme dans jeunes couples, colocataires et célibataire
	15-19	10-14	
	20-24	15-19	
	25-29	20-24	
	30-34	25-29	
	35-39	30-34	
	40-44	35-39	
	45-49	40-44	
	50-54	45-49	
	55-59	50-54	
	60-64	55-59	
	65-69	60-64	
	70-74	65-69	
Femme de	75-79	70-74	Femme dans jeunes couples, colocataires et célibataire
	80-84	75-79	
	85-89	80-84	
	90+	85-89	
	15-19	90+	
	20-24		
	25-29		
	30-34		
	35-39		
	40-44		
	45-49		
	50-54		
	55-59		
	60-64		
	65-69		
	70-74		
	75-79		
80-84			
85-89			
90+			

Tableau 5-10 Description des variables explicatives du modèle 5

Modèle 5				
Genre	Âge et appartenance à un type de ménage			
Homme Femme	5-14 dans	jeunes couples, colocataires et célibataire jeunes familles ...	15-24 dans	jeunes couples, colocataires et célibataire jeunes familles ...
	25-44 dans	jeunes couples, colocataires et célibataire jeunes familles ...	45-64 dans	jeunes couples, colocataires et célibataire jeunes familles ...
	65-99	jeunes couples, colocataires et célibataire jeunes familles ...		

Dans le cas du modèle 3, les personnes âgées de 5 à 14 ans sont toutes regroupées au sein d'une seule catégorie, peu importe le genre. De plus, les catégories d'âges concernant le dernier modèle sont différentes des modèles 2, 3 et 4 pour limiter le nombre total de catégories.

À l'exception de la variable de l'âge dans le premier modèle, l'ensemble des variables explicatives des cinq modèles sont catégorielles. Il est donc nécessaire de les transformer en variables binaires. En effet, les variables explicatives doivent absolument être numériques afin de pouvoir être insérées dans l'équation d'utilité d'une régression. En transformant ces variables discrètes en variables pouvant prendre la valeur de 0 ou 1, il est possible de les intégrer au modèle tout en prenant soin d'identifier une catégorie de référence. Puisqu'un individu ne peut correspondre à plus d'une catégorie d'une certaine variable discrète, l'une d'entre elles doit servir de référence pour ainsi être en mesure d'estimer et d'interpréter les résultats du modèle. La catégorie de référence de chacune des variables est indiquée dans la prochaine section.

Cette transformation des variables implique l'existence d'un individu de référence, dont les caractéristiques correspondent aux catégories de référence de chacune des variables. En effet, lorsque l'ensemble des variables prennent la valeur de 0, l'équation d'utilité correspond ainsi à

celle de l'individu de référence. Seule la constante est encore présente dans l'équation puisque les autres coefficients sont annulés, ce qui signifie que la constante est propre à l'individu de référence.

5.2.1 Analyse de corrélation

À la suite de la formulation des modèles, il est indispensable de s'assurer que les variables indépendantes ne soient pas fortement corrélées entre elles afin d'éviter de biaiser l'interprétation des résultats. Le coefficient de corrélation linéaire de Pearson est utilisé afin de mesurer le niveau de corrélation entre les variables. Le seuil sous lequel le niveau de corrélation est jugé acceptable est de 0,4. Il doit donc n'y avoir aucune paire de variables ayant un coefficient plus élevé que ce seuil.

Le choix de la catégorie de référence pour chacune des variables discrètes correspond à celle qui est la plus fréquente dans les données. Ainsi, la catégorie qui est la plus susceptible d'être fortement corrélée avec les autres variables est éliminée. Dans le cas du premier modèle, les catégories « Jeunes familles » et « Femmes » sont celles étant les plus fréquentes dans l'échantillon de l'enquête pour les variables du type de ménage et du genre. Le Tableau 5-11 présente l'analyse de corrélation du premier modèle.

Tableau 5-11 Analyse de corrélation du modèle 1

Jeunes couples, colocataires et célibataire	-0.17								
Familles traditionnelles	-0.09	-0.16							
Adultes et personnes âgées	0.17	-0.07	-0.1						
Cinquantennaires	0.18	-0.12	-0.17	-0.07					
Jeunes retraités	0.44	-0.13	-0.18	-0.08	-0.13				
Vieux retraités	0.44	-0.08	-0.12	-0.05	-0.09	-0.09			
Familles multigénérationnelles	-0.01	-0.07	-0.1	-0.04	-0.08	-0.08	-0.05		
Homme	-0.04	0.03	0.02	0	0.01	-0.03	-0.05	-0.01	
	Âge	Jeunes couples, colocataires et célibataire	Familles traditionnelles	Adultes et personnes âgées	Cinquantennaires	Jeunes retraités	Vieux retraités	Familles multigénérationnelles	

Puisque les variables indiquant qu'un individu appartient à un ménage de type « Jeunes retraités » ou « Vieux retraités » sont trop corrélées avec la variable de l'âge, la catégorie de référence choisie préalablement pour les variables concernant le type de ménage est remplacée par le type « retraité », qui inclut les deux catégories corrélées. Après vérification, la catégorie de référence

pour les femmes. Afin d'éviter la présence d'un niveau de corrélation aussi élevé, cette catégorie devient la catégorie de référence dans le cas de la variable d'interaction entre le genre et le type de ménage.

Les coefficients de corrélation des variables du dernier modèle sont présentés à l'Annexe B, dans le Tableau B-16. Aucun coefficient n'indique la présence d'une trop forte corrélation entre les variables explicatives. Les catégories de référence sont donc les femmes pour le genre, ainsi que les personnes de 25 à 44 ans présentes dans les « Jeunes familles » pour la variable d'interaction.

5.2.2 Analyse des variables choisies

La dernière étape précédant l'estimation des modèles concerne l'analyse des variables explicatives. Cette analyse inclue diverses statistiques qui sont présentées dans le Tableau 5-14. Les variables qui y sont analysées sont celles du modèle 2.

Tableau 5-14 Analyse des variables explicatives du modèle 2

	Maximum	Minimum	Moyenne	Écart-type	Coefficient de variation
Jeunes couples, colocataires et célibataire	1	0	0.1	0.3	300 %
Familles traditionnelles	1	0	0.19	0.39	205 %
Adultes et personnes âgées	1	0	0.04	0.19	475 %
Cinquantennaires	1	0	0.11	0.31	282 %
Jeunes retraités	1	0	0.13	0.33	254 %
Vieux retraités	1	0	0.05	0.23	460 %
Familles multigénérationnelles	1	0	0.04	0.21	525 %
Âge	5-9	1	0.06	0.24	400 %
	10-14	1	0.06	0.25	417 %
	15-19	1	0.07	0.25	357 %
	20-24	1	0.06	0.24	400 %
	25-29	1	0.06	0.23	383 %
	30-34	1	0.07	0.26	371 %
	35-39	1	0.08	0.27	338 %
	40-44	1	0.09	0.28	311 %
	45-49	1	0.08	0.28	350 %
	50-54	1	0.09	0.28	311 %
	55-59	1	0.07	0.25	357 %
Homme	1	0	0.49	0.5	102 %

L'ensemble des variables explicatives du modèle 2 peuvent uniquement prendre la valeur de 0 ou 1. De plus, les moyennes des variables indiquant le type de ménage auquel chaque personne appartient reflètent la proportion de personnes dans chacun des ménages.

5.3 Première segmentation

Tel que mentionné dans le Chapitre 3, un échantillon de 504 776 personnes est utilisé pour la segmentation avec V1 à la suite de la sélection des individus âgés de cinq ans et plus ayant réalisé au moins une activité. Ce partitionnement a mené à la création de dix groupes. Dans le but de modéliser l'appartenance d'un individu à chacun de ces groupes, un onzième est créé, composé des personnes âgées de cinq ans et plus n'ayant fait aucune activité durant la journée où ils ont été sondés. En ajoutant ce groupe de personnes à l'échantillon utilisé pour la segmentation, le jeu de données nécessaire à la modélisation de l'appartenance aux groupes contient 613 997 personnes.

Cet exercice de modélisation nécessite également une division du jeu de données en deux segments, soit un jeu destiné à l'estimation des paramètres du modèle, de même qu'un autre pour tester la validité de celui-ci. Les personnes sélectionnées pour l'estimation du modèle sont au nombre de 491 198, soit 80 % du jeu de données initial, alors que 20 % des individus sont utilisés pour valider le modèle à la suite de son estimation.

5.3.1 Analyses des modèles

L'estimation des modèles consiste à calculer les coefficients des différentes variables explicatives. Grâce à ces coefficients, il est possible d'en dériver les rapports de probabilités de chacune des variables. La valeur du rapport équivaut à l'exponentiel du coefficient. L'interprétation de ces rapports, qui constitue la première portion de l'analyse des modèles, permet de comprendre l'influence de chacune des variables indépendantes sur la variable dépendante. Les coefficients peuvent également être utilisés pour calculer les probabilités d'adopter chacun des patrons typiques. Grâce à l'analyse de ces probabilités, il est possible d'évaluer l'ensemble d'un modèle, qui fait l'objet de la seconde portion de cette analyse.

5.3.1.1 Rapports de probabilités

Les rapports de probabilités expriment la chance d'adopter un certain patron typique par rapport au patron de référence, soit le fait d'être immobile, de demeurer à son domicile et de ne réaliser

aucune activité. Cette catégorie de la variable expliquée correspond à celle dont le coefficient est nul. De plus, la chance d'adopter un patron correspond à l'effet de l'augmentation de 1 d'une variable explicative. Dans le cas des variables catégorielles, puisque celles-ci peuvent uniquement prendre les valeurs de 0 ou 1, les chances découlant d'un rapport de probabilités correspondent au fait d'appartenir à une certaine catégorie par rapport à la catégorie de référence. Les rapports de probabilités du premier modèle sont présentés dans le Tableau 5-15 et le Tableau 5-16.

Tableau 5-15 Rapport de probabilités du modèle 1 – Activités *études* et *travail*

	V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-I
Constante	16850.78***	2284.74***	0.29***	0.08***	0.02***	0.01***	-
Âge	0.79***	0.82***	0.99***	0.99***	1.01***	1.01***	-
Jeunes couples, colocataires et célibataire	0.02***	0.02***	10.65***	12.01***	17.96***	23.41***	-
Jeunes familles	0.03***	0.02***	6.95***	5.62***	11.39***	54.25***	-
Familles traditionnelles	0.03***	0.02***	8.37***	6.41***	8.44***	17.53***	-
Adultes et personnes âgées	0.01***	0.01***	2.34***	1.61***	1.54***	2.01***	-
Cinquantennaires	0.12***	0.09***	8.24***	7.42***	10.65***	11.58***	-
Retraités	-	-	-	-	-	-	-
Familles multigénérationnelles	0.02***	0.01***	3.46***	2.48***	3.65***	10.51***	-
Homme	1.26***	1.46***	1.7***	1.72***	1.37***	1.7***	-
Femme	-	-	-	-	-	-	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

La majorité des rapports sont significatifs à $p < 0.01$. Concernant la variable de l'âge, les rapports de probabilités révèlent qu'une augmentation d'une unité de l'âge diminue de 21 % et de 18 % les chances d'adopter les patrons V1-E1 et V1-E2 respectivement et ce par rapport au patron de référence, V1-I. Autrement dit, plus une personne vieillit, moins elle a de chance de réaliser un patron typique où l'*étude* est le type d'activité dominante par rapport au fait de rester à son domicile et ne faire aucune activité.

L'effet de l'appartenance aux différents types de ménage sur l'adoption des patrons où l'*étude* est la principale activité peut sembler contradictoire. Toutefois, la valeur très élevée du rapport de probabilités de la constante permet d'expliquer cette contradiction. À titre de rappel, la constante représente l'équation d'utilité de l'individu de référence, dont les attributs correspondent aux

catégories de références, soit lorsque les coefficients de l'ensemble des variables explicatives sont nuls.

Tableau 5-16 Rapport de probabilités du modèle 1 – Activités *loisir, magasinage* et *autres*

	V1-L1	V1-L2	V1-M1	V1-A1	V1-I
Constante	0.72***	0.17***	0.26***	0.16***	-
Âge	0.99***	1	1.01***	1.01***	-
Jeunes couples, colocataires et célibataire	0.67***	1.01	1.03	1.7***	-
Jeunes familles	0.53***	0.82***	1.13***	2.84***	-
Familles traditionnelles	0.58***	0.61***	0.83***	1.65***	-
Adultes et personnes âgées	0.56***	0.52***	0.78***	0.98	-
Cinquantennaires	0.79***	1.05	1.14***	1.64***	-
Retraités	-	-	-	-	-
Familles multigénérationnelles	0.44***	0.4***	0.7***	1.38***	-
Homme	1.24***	1	0.93***	1.52***	-
Femme	-	-	-	-	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

En outre, les personnes composant les ménages de type « Retraités » ont moins de chance que les membres appartenant aux autres types de réaliser un patron dont la principale activité est de type *travail* ou *autres*. Ces chances s'avèrent plus élevées dans le cas des membres des « Jeunes familles » concernant les chances de réaliser le patron V1-T4. Dans le cas des trois patrons dont les activités dominantes sont le *loisir* et le *magasinage*, les « Retraités » ont généralement plus de chances de les réaliser par rapport aux membres des autres types de ménage.

L'effet de l'âge sur la réalisation de la majorité des patrons semble minimal selon la valeur des rapports de probabilités. En effet, à l'exception de V1-E1 et V1-E2, l'augmentation de l'âge induit une variation de 1 % des chances de réaliser un de ces patrons par rapport à celui de référence. Cette faible variation met en évidence la problématique créée par l'insertion de la variable de l'âge dans le modèle comme une variable numérique. En ne transformant pas cette variable en variable catégorielle, l'effet sur l'adoption d'un patron est donc considéré comme linéaire. Cette hypothèse peut être infirmée ou confirmée par les résultats du deuxième modèle, où l'âge est transformé en variable catégorielle. Les rapports de probabilités de ce modèle sont présentés dans le Tableau 5-17. La majorité des rapports sont significatifs.

Tableau 5-17 Rapport de probabilités du modèle 2 – Activités *études et travail*

	V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-I
Constante	0***	0.01***	0.48***	0.08***	0.06***	0.49***	-
5-9	2612.94***	91.75***	0.02***	0.04***	0.11***	0***	-
10-14	3994.45***	212.84***	0.03***	0.03***	0.03***	0***	-
15-19	1530.76***	99.08***	0.93**	1.45***	0.55***	0.08***	-
20-24	443.01***	30.25***	2.34***	5.9***	2.06***	0.56***	-
25-29	144.09***	5.29***	3.49***	5.9***	4.67***	0.99	-
30-34	83.41***	0.72	4.31***	5.34***	6.58***	1.76***	-
35-39	94.16***	0.16***	4.79***	7.13***	8.86***	2.04***	-
40-44	50.85***	0.01***	5.91***	7.15***	9.68***	2.13***	-
45-49	27.07***	10.38***	5.7***	5.89***	8.03***	2.79***	-
50-54	13.72***	0.04***	4.84***	6.44***	6.82***	2.45***	-
55-59	7.21***	8.37***	3.64***	3.78***	4.7***	1.31***	-
60+	-	-	-	-	-	-	-
Jeunes couples, colocataires et célibataire	1.06**	1.28***	1.42***	1.82***	1.14***	0.59***	-
Jeunes familles	-	-	-	-	-	-	-
Familles traditionnelles	1.2***	1.1***	1.12***	0.97	0.98	0.36***	-
Adultes et personnes âgées	0.35***	0.29***	0.43***	0.37***	0.25***	0.06***	-
Cinquantenaires	0.39***	0.46***	0.9***	0.95*	0.88***	0.17***	-
Jeunes retraités	0***	0.48***	0.56***	0.43***	0.54***	0.04***	-
Vieux retraités	3.53***	3.66***	0.06***	0***	0***	0***	-
Familles multigénérationnelles	0.71***	0.71***	0.53***	0.4***	0.45***	0.27***	-
Homme	1.18***	1.04*	1.92***	2***	1.4***	2.37***	-
Femme	-	-	-	-	-	-	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Tableau 5-18 Rapport de probabilités du modèle 2 – Activités *loisir, magasinage et autres*

	V1-L1	V1-L2	V1-M1	V1-A1	V1-I
Constante	0.38***	0.3***	0.76***	0.52***	-
5-9	0.51***	0.15***	0.13***	0.28***	-
10-14	1.07	0.01***	0.34***	0.12***	-
15-19	0.81***	0.29***	0.46***	0.25***	-
20-24	0.8***	0.55***	0.41***	0.67***	-
25-29	0.74***	0.86**	0.54***	0.75***	-
30-34	0.61***	0.64***	0.78***	0.97	-
35-39	0.47***	0.6***	0.72***	1.35***	-
40-44	0.67***	0.73***	0.99	1.69***	-
45-49	0.6***	0.74***	0.91***	1.74***	-
50-54	0.91**	0.71***	1.02	1.32***	-
55-59	0.79***	0.66***	1.27***	1.41***	-
60+	-	-	-	-	-
Jeunes couples, colocataires et célibataire	1.09**	0.92*	1.26***	0.76***	-
Jeunes familles	-	-	-	-	-
Familles traditionnelles	1.02	0.58***	0.61***	0.52***	-
Adultes et personnes âgées	0.66***	0.23***	0.65***	0.3***	-
Cinquantennaires	0.9***	0.83***	0.78***	0.43***	-
Jeunes retraités	1.19***	0.76***	0.86***	0.48***	-
Vieux retraités	0.46***	0.27***	0.51***	0.2***	-
Familles multigénérationnelles	0.44***	0.23***	0.58***	0.34***	-
Homme	1.45***	0.95**	1.03**	2.49***	-
Femme	-	-	-	-	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Les chances que les personnes de moins de 60 ans, par rapport à celles plus âgées, adoptent un des patrons où l'*étude* est l'activité dominante diminuent avec l'âge. Ces chances sont très élevées pour les personnes de moins de 19 ans, ce qui met évidence le fait que la grande majorité des jeunes vont à l'école, alors qu'une portion considérable de personnes âgées ne pratiquent aucune activité.

Dans le cas des patrons V1-T1 à V1-T4, l'évolution des rapports de probabilités de la variable de l'âge confirme la pertinence de discrétiser cette variable et infirme l'hypothèse nécessaire à l'estimation du premier modèle, soit que la variation de l'âge a un effet linéaire sur les chances d'adopter un patron typique. En effet, alors que pour les jeunes, les chances sont moindres par

rapport aux personnes âgées de réaliser un patron dont la principale activité est le *travail*, ces chances augmentent de plus en plus avec l'âge jusqu'à un certain point, pour ensuite diminuer.

L'influence de l'âge sur l'adoption des quatre derniers patrons est caractérisée par des chances moindres pour les personnes de moins de 60 ans par rapport à celles plus âgées, à l'exception des personnes entre 35 et 59 ans, qui ont plus de chances de réaliser V1-A1 plutôt que de rester à la maison, et des individus âgés entre 55 et 59, dont les chances d'adopter V1-M1 sont plus élevées que les personnes âgées.

De manière générale, une personne appartenant à un ménage dont le type est différent des « Jeunes familles » voit ses chances diminuer d'adopter la majorité des patrons plutôt que de ne réaliser aucune activité. Cette situation démontre que les personnes composant les « Jeunes familles » sont peu nombreuses à rester à la maison durant la journée où elles ont été sondées. Toutefois, les personnes composant les « Jeunes couples, colocataires et célibataires » ont plus de chances que celles à l'intérieur des « Jeunes familles » de réaliser la majorité des patrons, à l'exception de V1-T4, V1-L2 et V1-A1. Ainsi, les chances qu'un individu appartenant au type de ménage de référence réalise une activité de type *autre* avant et/ou après le *travail*, soit le patron V1-T4, sont plus élevées que celles des membres des ménages d'autres types. Ce constat est également valable pour V1-L2 et V1-A1. De plus, les membres des « Jeunes retraités » ont davantage de chance que ceux des « Jeunes familles » de réaliser une activité de type *loisir* durant la journée plutôt que de rester à la maison, contrairement aux membres des « Vieux retraités ».

Finalement, les femmes semblent avoir plus de chances que les hommes de rester à la maison plutôt que de réaliser un des patrons, à l'exception de V1-L2. Cependant, ces chances peuvent varier en fonction de l'âge des personnes et ces différences peuvent être captées par le troisième modèle, dont les rapports de probabilités sont présentés dans les Tableau 5-19 et Tableau 5-21. La catégorie de référence dans le cas de la variable d'interaction entre l'âge est le genre est une femme âgée de 75 et 79 ans. De plus, la majorité des rapports sont significatifs.

Tableau 5-19 Rapport de probabilités du modèle 3 – *Études et travail* – Variable âge et constante

		V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-I
Constante		0.01***	0***	0.07***	0.01***	0.02***	0.03***	-
5-14		2116.25***	413.94***	0***	0	0.17***	0	-
15-19	F	1319***	346.77***	8.47***	21.94***	4.99***	3.95***	-
	H	1097.84***	347.32***	11.12***	21.01***	0.15***	2.16**	-
20-24	F	386.03***	92.58***	22.48***	52.23***	10.32***	10.57***	-
	H	280.51***	102.51***	26.86***	64.92***	9.59***	12.26***	-
25-29	F	71.95***	17.43***	28.21***	56.5***	16.41***	27.25***	-
	H	108.32***	31.05***	52.07***	95.04***	22.02***	29.88***	-
30-34	F	42***	7.85**	28.61***	44.04***	15.79***	38.14***	-
	H	53.88***	9.57**	58.6***	87.97***	36.69***	55.53***	-
35-39	F	28.76***	3.51	33.78***	55.25***	19.91***	42.08***	-
	H	38.46***	3.85	67.86***	108.09***	30.61***	74.12***	-
40-44	F	26.57***	4.53	40.54***	73.47***	29.04***	41.48***	-
	H	28.62***	2.74	65.62***	105.87***	44.51***	70.17***	-
45-49	F	16.45***	3.41	38.28***	67.07***	33.05***	35.11***	-
	H	10.19***	2.92	63.78***	104.82***	45.84***	76.37***	-
50-54	F	6.27***	0.57	36.7***	47.32***	29.17***	24.84***	-
	H	0	2.56	51.25***	85.66***	33.8***	59.13***	-
55-59	F	5.27***	3.58	23.54***	29.99***	18.05***	13.2***	-
	H	6.19***	0.17	39.62***	44.66***	23.56***	36.98***	-
60-64	F	0.06***	14.01***	11.89***	11.48***	9.24***	6.41***	-
	H	3.59**	5.23*	22.93***	24.06***	15.13***	20.64***	-
65-69	F	0	22.68***	2.64***	3.47***	3.39***	1.61	-
	H	0	22.13***	7.47***	9.48***	6.42***	7.05***	-
70-74	F	0***	14.64***	1.41***	1.69*	2.02***	0.46**	-
	H	0	12.31***	3.27***	4.88***	2.94***	2.09**	-
75-79	F	-	-	-	-	-	-	-
	H	0.67	13.95***	2.66***	2.1**	2.24***	2.56***	-
80-84	F	0.02	6.02*	0.05***	0.05**	0.41**	0.93	-
	H	0.01	10.92**	1.82***	1.74	3.94***	4.93***	-
85-89	F	0.01	42.06***	0.02***	0.08*	0.45	3.38***	-
	H	0.01	0.59	0.54*	0.88	5.65***	3.5***	-
90+	F	0***	0.2	0.2***	2.19*	12.22***	0.76	-
	H	0.04	0.02	2.01**	0.2	0.03	0.31	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Tableau 5-20 Rapport de probabilités du modèle 3 – *Études et travail* – Variable ménage

	V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-I
Jeunes couples, colocataires et célibataire	0.83***	1.08	1.19***	1.65***	1.31***	0.34***	-
Jeunes familles	-	-	-	-	-	-	-
Familles traditionnelles	1.07***	1.18***	1.13***	1.1***	0.69***	0.43***	-
Adultes et personnes âgées	0.45***	0.13***	0.58***	0.49***	0.35***	0.12***	-
Cinquantennaires	0.86**	0.98	0.9***	1	0.78***	0.22***	-
Jeunes retraités	1.81***	0.07***	0.72***	0.96	0.46***	0.25***	-
Vieux retraités	0*	0.04***	0.33***	0.73*	0.27***	0.11***	-
Familles multigénérationnelles	0.88***	0.57***	0.71***	0.57***	0.61***	0.42***	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Les rapports de probabilités de la variable d'interaction intégrant l'âge et le genre permettent de constater que les chances d'adopter un des patrons typiques où le *travail* est l'activité dominante n'évoluent pas au même rythme selon le genre. En effet, pour les femmes âgées de 15 à 74 ans, les chances de réaliser le patron V1-T1 plutôt que de rester à la maison sont de 8,47 à 1,41 fois plus élevées que celle pour une femme âgée entre 75 et 79 ans. Ces chances augmentent avec l'âge, atteignant 40,54 pour les femmes de 40 et 44 ans, pour ensuite diminuer. De manière générale, les hommes ont également plus de chances que les femmes de 75 à 79 ans d'adopter ce patron plutôt qu'être immobile. Toutefois, le rapport de probabilités n'évolue pas au même rythme. Les chances de réaliser V1-T1 pour les hommes augmentent plus rapidement avec l'âge que les femmes, variant de 11,12 pour les hommes de 15 à 19 ans à 67,86 pour ceux âgés de 35 à 39 ans.

Bref, autant les femmes que les hommes entre 15 et 74 ans ont plus de chances que les femmes de 75 à 79 de réaliser un patron dont la principale activité est le *travail*. En revanche, ces chances n'évoluent pas au même rythme selon le genre et elles sont moins élevées pour les femmes, puisqu'elles sont plus nombreuses que les hommes à rester à la maison et ne faire aucune activité.

Tableau 5-21 Rapport de probabilités du modèle 3 – *Loisir, magasinage et autres* – Var. ménage

	V1-L1	V1-L2	V1-M1	V1-A1	V1-I
Constante	0.16***	0.09***	0.62***	0.33***	-
Jeunes couples, colocataires et célibataire	1.13***	0.96	0.72***	0.46***	-
Jeunes familles	-	-	-	-	-
Familles traditionnelles	0.96	0.58***	0.62***	0.62***	-
Adultes et personnes âgées	0.95	0.72***	0.66***	0.43***	-
Cinquantennaires	1.34***	1.01	0.8***	0.56***	-
Jeunes retraités	1.55***	1.09	0.69***	0.5***	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Tableau 5-22 Rapport de probabilités du modèle 3 – *Loisir, magasinage et autres* – Variable âge

		V1-L1	V1-L2	V1-M1	V1-A1	V1-I
Familles multigénérationnelles		0.79***	0.52***	0.51***	0.54***	-
5-14		1.47***	0.5***	0.17***	0.51***	-
15-19	F	2.25***	1.66***	0.31***	0.53***	-
	H	2.32***	0.23***	0.06***	0.63***	-
20-24	F	1.73***	1.44***	0.58***	0.92	-
	H	2.32***	1.07	0.28***	1.18**	-
25-29	F	1.82***	1.73***	0.9*	1.85***	-
	H	2.25***	1.65***	0.55***	2.82***	-
30-34	F	1.83***	1.76***	1.09	2.28***	-
	H	1.82***	1.74***	0.75***	3.03***	-
35-39	F	1.55***	2***	1.09	2.35***	-
	H	2.01***	1.78***	0.75***	3.43***	-
40-44	F	1.93***	1.96***	1.22***	2.3***	-
	H	1.92***	1.88***	0.69***	3.42***	-
45-49	F	1.51***	1.66***	1.28***	1.81***	-
	H	1.89***	1.84***	0.76***	3.55***	-
50-54	F	1.66***	1.95***	1.25***	1.77***	-
	H	1.74***	2.02***	1.06	3.33***	-
55-59	F	1.72***	2***	1.41***	1.64***	-
	H	2.01***	1.86***	1.25***	2.94***	-
60-64	F	1.76***	1.87***	1.39***	1.65***	-
	H	2.33***	1.96***	1.58***	2.71***	-
65-69	F	1.67***	1.77***	1.31***	1.55***	-
	H	2.08***	1.92***	1.57***	2.1***	-
70-74	F	1.36***	1.46***	1.06	1.37***	-
	H	1.85***	1.85***	1.36***	1.75***	-
75-79	F	-	-	-	-	-
	H	1.7***	1.56***	1.26***	1.92***	-
80-84	F	0.74***	0.63***	0.62***	0.86**	-
	H	1.36***	1.34***	0.96	1.87***	-
85-89	F	0.66***	0.81*	0.39***	0.59***	-
	H	1.1	1.54***	0.82**	1.22**	-
90+	F	0.49***	0***	0.25***	0.4***	-
	H	1.14	2.16***	0.77**	0.54***	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

La variation, en fonction de l'âge, des chances de réaliser le patron V1-M1 plutôt qu'être immobile semble être différente pour les hommes et les femmes. En effet, les hommes de 40 à 49 ans ont 24 et 31 % moins de chance que les femmes de 75 à 79 ans de magasiner durant la journée plutôt que de rester à la maison, alors que l'inverse est observé pour les femmes de 40 à 49 ans, qui ont 22 et

28 % plus de chance d'adopter le patron V1-M1 par rapport au groupe de référence. De plus, cette différence semble se dissiper pour les personnes âgées de 50 à 75 ans, dont les rapports de probabilités pour les deux genres sont similaires. Cette variation des rapports démontre la pertinence de créer la variable d'interaction intégrant l'âge et le genre. L'influence du genre sur les systèmes d'activités des individus peut également varier selon le type de ménage auquel une personne appartient. C'est l'exploration de cette variation qui incite au développement du quatrième modèle. Les rapports de probabilités de ce dernier sont présentés dans les Tableau 5-23 et Tableau 5-24. Encore une fois, les coefficients et les rapports de probabilités qui en découlent sont en majorité significatifs à $p < 0,01$.

Tableau 5-23 Rapport de probabilités du modèle 4 – *Études et travail*

		V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-I
Constante		0***	0.01***	0.18***	0.09***	0.25***	0***	-
5-14		15726.42***	113.23***	0.1***	0.12***	0.06***	0.02***	-
15-24		4481.93***	53.5***	0.55***	0.84***	0.28***	0.21***	-
25-44		379.73***	3.36***	1.14***	1.32***	0.98	1.14***	-
45-64		-	-	-	-	-	-	-
65+		8.35***	0.45***	0.14***	0.11***	0.11***	0.13***	-
Jeunes couples, colocataires et célibataire	F	37.98***	1.63*	16.42***	6.89***	2.24***	1939.6***	-
	H	43.22***	2.08***	22.31***	9.56***	2.47***	3098.19***	-
Jeunes familles	F	65.61***	1.94**	10.65***	3.17***	1.69***	6446.51***	-
	H	81.27***	2.8***	24.59***	7.48***	2.73***	10893.35***	-
Familles traditionnelles	F	60.42***	1.92**	12.94***	3.35***	1.27**	2033.27***	-
	H	62.75***	2.48***	19.37***	5.05***	1.28**	3560.9***	-
Adultes et personnes âgées	F	13.1**	0.67	4.84***	0.96	0.36***	608.73***	-
	H	15.95***	0.7	7.85***	1.84***	0.51***	602.97***	-
Cinquantennaires	F	35.27***	0.36***	9.48***	3.15***	1.19*	1016.34***	-
	H	43.59***	1.2	13.61***	4.54***	1.61***	2174.86***	-
Jeunes retraités	F	11.63**	0.57**	3.84***	0.98	0.42***	363.15**	-
	H	0.5	0.63	8.13***	2.26***	0.77***	1125.81***	-
Vieux retraités	F	-	-	-	-	-	-	-
	H	0***	0.8	2.68***	0.66*	0.32***	396.63**	-
Familles multigénérationnelles	F	43.47***	1.21	6.42***	1.49***	0.73***	2037.82***	-
	H	55.95***	1.66*	12.46***	3.05***	0.8*	3651.01***	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Encore une fois, les chances qu'un homme adopte un patron dont la principale activité est le *travail* par rapport au groupe de référence sont plus élevées que celles des femmes et ce, pour la majorité des types de ménage. En revanche, les chances d'adopter le patron V1-T3 sont similaires pour les deux genres dans le cas des ménages de type « Jeunes couples, colocataires et célibataires », « Familles traditionnelles », et « Cinquantenaires ». Une situation similaire est observée dans le cas des « Adultes et personnes âgées » quant aux chances de réaliser le patron V1-T4 plutôt que de rester à la maison.

Tableau 5-24 Rapport de probabilités du modèle 4 – *Loisir, magasinage et autres*

		V1-L1	V1-L2	V1-M1	V1-A1	V1-I
Constante		0.28***	0.08***	0.56***	0.36***	-
5-14		0.96	0.54***	0.22***	0.3***	-
15-24		1.34***	0.57***	0.34***	0.38***	-
25-44		0.89***	0.73***	0.8***	1.07***	-
45-64		-	-	-	-	-
65+		0.76***	0.85***	0.78***	0.6***	-
Jeunes couples, colocataires et célibataire	F	1.27***	2.86***	1.26***	0.76***	-
	H	1.74***	3.17***	1.06	1.54***	-
Jeunes familles	F	1.19***	3.39***	1.98***	2.29***	-
	H	1.2***	2.16***	1.06	2.96***	-
Familles traditionnelles	F	1	1.78***	1.06	1.13***	-
	H	1.19***	1.56***	0.69***	2.05***	-
Adultes et personnes âgées	F	0.93	0.92	0.88***	0.7***	-
	H	1.29***	1.32***	1.09**	1.27***	-
Cinquantenaires	F	1.19***	1.98***	1.1***	0.91**	-
	H	1.65***	2.73***	1.03	1.64***	-
Jeunes retraités	F	1.8***	2.98***	1.36***	1.09**	-
	H	2.53***	3.19***	1.65***	1.64***	-
Vieux retraités	F	-	-	-	-	-
	H	1.64***	1.76***	1.35***	1.31***	-
Familles multigénérationnelles	F	0.77***	0.76***	0.86***	1.06	-
	H	1.12**	1.1	0.74***	1.95***	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Les chances d'adopter un des deux patrons où le *loisir* est la principale activité varient peu selon le genre à l'intérieur des différents types de ménage. L'inverse est observé dans le cas de V1-A1,

où des différences entre les hommes et les femmes sont observées pour certains types de ménage quant à leurs chances d'adopter ce patron.

Ainsi, au sein d'un même type de ménage, les hommes et les femmes peuvent adopter le même comportement quant aux activités réalisées au cours d'une journée typique, mais dans certains cas, il subsiste des différences entre les deux genres. Des différences entre les personnes d'un même type de ménage peuvent également être observées selon l'âge des personnes, comme le démontre les résultats du dernier modèle, dont les rapports de probabilités sont présentés dans les Tableau 5-25 et Tableau 5-26.

Les chances d'adopter les patrons V1-T1 à V1-T4 sont similaires pour les personnes de moins de 14 ans, peu importe le type de ménage auxquelles ils appartiennent. Le même constat s'applique aux personnes de plus de 65 ans et à celles âgées de 15 à 24 ans, à l'exception d'un type de ménage. En effet, dans le cas de ce dernier groupe d'âge, le fait d'appartenir à un ménage de type « Jeunes couples, colocataires et célibataires » semble faire une différence quant aux chances d'adopter V1-T2 par rapport au fait de rester à la maison.

Des différences entre types de ménages sont également observées pour les personnes de 25 à 44 ans. Le fait d'appartenir à des types de ménage différents, malgré un âge similaire, peut avoir un effet sur les activités réalisées au cours d'une journée. Cet effet semble être le même pour les « Jeunes couples, colocataires et célibataires » et les « Cinquantenaires », qui ont de 27 à 86 % plus de chances que les personnes du même âge membres d'une « Jeune famille » d'adopter V1-T1, V1-T2 et V1-T3. De plus, le lien fort entre les « Jeunes familles » et V1-T4 se reflète également dans la valeur de ces rapports de probabilités, dont la valeur est toujours plus faible que 1 et ce pour l'ensemble des types de ménages. En outre, les personnes dans les « Jeunes familles » ne semblent pas avoir le même comportement selon leur âge, alors que celles ayant entre 45 et 64 ans ont davantage de chances que les personnes entre 25 et 44 ans d'adopter les patrons où le *travail* est l'activité dominante.

Tableau 5-25 Rapport de probabilités du modèle 5 – *Études et travail*

		V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-I
	Constante	0.17***	0***	2.08***	0.36***	0.35***	0.94***	-
	Homme	1.18***	1.4***	1.82***	1.99***	1.46***	2.03***	-
	Femme	-	-	-	-	-	-	-
5-14	Jeunes couples, colocataires et célibataire	1659812.52***	88695.8***	0***	0.06***	0.07***	0***	-
	Jeunes familles	81.02***	289.17***	0.01***	0.15***	0***	0	-
	Familles traditionnelles	67.26***	257.41***	0.03***	0.01***	0.02***	0.01***	-
	Familles multigénérationnelles	47.8***	199.46***	0.06***	0.01***	0.05***	0.02***	-
15-24	Jeunes couples, colocataires et célibataire	9.73***	101.47***	0.85***	1.14**	0.71***	0.18***	-
	Jeunes familles	37.91***	168.28***	0.35***	0.46***	0.13***	0.07***	-
	Familles traditionnelles	25.34***	166.36***	0.67***	0.71***	0.17***	0.07***	-
	Adultes et personnes âgées	3.32**	1.02	0.65	0.32	0.04	0.02*	-
	Familles multigénérationnelles	25.19***	139.28***	0.48***	0.28***	0.24***	0.07***	-
25-44	Jeunes couples, colocataires et célibataire	2.16***	9.87***	1.49***	1.86***	1.43***	0.28***	-
	Jeunes familles	-	-	-	-	-	-	-
	Familles traditionnelles	1.9***	8.53***	1.11***	0.9**	0.72***	0.29***	-
	Adultes et personnes âgées	0.98	5.18***	0.82***	0.67***	0.41***	0.1***	-
	Cinquantennaires	0.55***	4.38***	1.41***	1.47***	1.27***	0.35***	-
	Familles multigénérationnelles	1.34***	5.32***	0.92*	0.78***	0.65***	0.59***	-
45-64	Jeunes couples, colocataires et célibataire	0.49	0.19	1.37	1.27	1.5	0.24***	-
	Jeunes familles	0.64***	3.23***	1.35***	1.38***	1.44***	1.12***	-
	Familles traditionnelles	0.19***	0.65*	1.16***	0.8***	0.95	0.43***	-
	Adultes et personnes âgées	0***	1.61*	0.39***	0.27***	0.33***	0.08***	-
	Cinquantennaires	0.17***	1.06	0.86***	0.94**	0.69***	0.17***	-
	Jeunes retraités	0.09***	0.96	0.25***	0.19***	0.19***	0.05***	-
	Familles multigénérationnelles	0.31***	2.41***	0.65***	0.41***	0.63***	0.34***	-
65+	Familles traditionnelles	0	0.01	0.06***	0***	0.01***	0**	-
	Adultes et personnes âgées	0***	0.53*	0.03***	0.03***	0.04***	0.01***	-
	Cinquantennaires	0.17	0.91	0.96	0.16*	0.61	0	-
	Jeunes retraités	0.02***	0.64**	0.06***	0.05***	0.08***	0.02***	-
	Vieux retraités	0***	0.28***	0.01***	0.02***	0.02***	0***	-
	Familles multigénérationnelles	0***	3.5***	0.04***	0.02***	0.06***	0.02***	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Tableau 5-26 Rapport de probabilités du modèle 5 – *Loisir, magasinage et autres*

		V1-L1	V1-L2	V1-M1	V1-A1	V1-I
Constante		0.26***	0.13***	0.51***	0.58***	-
Homme		1.34***	1.02	0.96***	1.56***	-
Femme		-	-	-	-	-
5-14	Jeunes couples, colocataires et célibataire	0.2***	0.41***	1.22***	0.04***	-
	Jeunes familles	0.63***	0.33***	0.14***	0.17***	-
	Familles traditionnelles	0.66***	0.07***	0***	0.01***	-
	Familles multigénérationnelles	0.76**	0.4***		0.06***	-
15-24	Jeunes couples, colocataires et célibataire	1.41***	1.14	0.5***	0.28***	-
	Jeunes familles	1.18***	0.77***	0.38***	0.27***	-
	Familles traditionnelles	1.07*	0.49***	0.2***	0.25***	-
	Adultes et personnes âgées	2.46*	0.31	1.56	0.02	-
	Familles multigénérationnelles	1.51***	0.13***	0.25***	0.21***	-
25-44	Jeunes couples, colocataires et célibataire	1.72***	1.19***	0.74***	0.55***	-
	Jeunes familles	-	-	-	-	-
	Familles traditionnelles	0.89**	0.73***	0.63***	0.42***	-
	Adultes et personnes âgées	0.95	0.53***	0.49***	0.43***	-
	Cinquantenaires	1.19**	1.59***	1.07	0.54***	-
	Familles multigénérationnelles	0.83**	0.92	0.99	0.76***	-
45-64	Jeunes couples, colocataires et célibataire	3.32***	0.37	1.55	0.55*	-
	Jeunes familles	0.9*	1.34***	1	1.13***	-
	Familles traditionnelles	0.95	0.98	0.92***	0.76***	-
	Adultes et personnes âgées	0.97	0.97	1.06	0.4***	-
	Cinquantenaires	1.16***	1.41***	1.11***	0.65***	-
	Jeunes retraités	1.68***	1.55***	1.17***	0.39***	-
	Familles multigénérationnelles	0.75***	0.91	0.84***	0.72***	-
65+	Familles traditionnelles	0.62***	0.7	0.48***	0.17***	-
	Adultes et personnes âgées	0.64***	0.59***	0.78***	0.21***	-
	Cinquantenaires	0.59	1.1	0.28	0.11**	-
	Jeunes retraités	1.32***	1.45***	1.25***	0.38***	-
	Vieux retraités	0.97	0.91**	0.92***	0.26***	-
	Familles multigénérationnelles	0.52***	0.32***	0.46***	0.16***	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Les membres des « Jeunes couples, colocataires et célibataire » et des « Cinquantenaires » âgés de 25 à 64 ans semblent avoir des chances similaires d'adopter les patrons dont la principale activité est le *loisir*. De plus, parmi l'ensemble des personnes âgées de 25 à 44 ans, les chances d'adopter les patrons V1-L1, V1-L2 et V1-M1 diffèrent selon le type de ménage. Une situation similaire est observée pour les personnes de 45 à 64 ans et celles de 65 ans et plus, ce qui confirme la pertinence

d'intégrer la variable d'interaction entre l'âge et le type de ménage. De plus, ces observations confirment également que malgré un âge similaire, le fait d'appartenir à un type de ménage différent a une influence sur les activités réalisées, ce qui met de l'avant l'importance de prendre en compte la structure des ménages dans la prévision de la demande en transport.

5.3.1.2 Analyse des probabilités

Outre les rapports de probabilités, les coefficients estimés grâce à la technique de la maximisation du log-vraisemblance permettent également d'évaluer, pour chaque individu, les probabilités d'appartenir à chacun des groupes selon la valeur des variables explicatives propres à chaque personne. Ces probabilités peuvent ensuite être analysées dans le but d'évaluer la pertinence d'un modèle. Le premier élément de cette analyse correspond à la vraisemblance. Comme mentionné dans la méthodologie, à la section 3.5, la vraisemblance d'une observation correspond à la probabilité d'appartenir au type observé. Le Tableau 5-27 permet d'identifier cette probabilité, qui est en rouge, pour un extrait de la base de données à la suite de l'application du modèle.

Tableau 5-27 Extrait de la base de données des personnes permettant d'identifier la vraisemblance de chaque observation (modèle 3)

IDPERS	Groupe observé	Probabilité d'appartenir à chaque groupe										
		V1-I	V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-L1	V1-L2	V1-M1	V1-A1
1	V1-E1	0,1	0,65	0,16	0	0	0	0,06	0	0	0,01	0,02
2	V1-L2	0,33	0	0	0,09	0	0	0,01	0,11	0,17	0,2	0,1
3	V1-E1	0,05	0,86	0,07	0	0	0	0	0,02	0	0	0
4	V1-T4	0,15	0	0,01	0,34	0,11	0,18	0,2	0	0	0,01	0
5	V1-T2	0,19	0	0	0,55	0,15	0,06	0,05	0	0	0	0
6	V1-E1	0,14	0,61	0,13	0	0	0	0	0	0	0	0,12

La Figure 5-1 présente la distribution de la vraisemblance de chaque observation dans le jeu d'entraînement. Le modèle utilisé afin de calculer les probabilités illustrées dans ce graphique est le modèle 3.

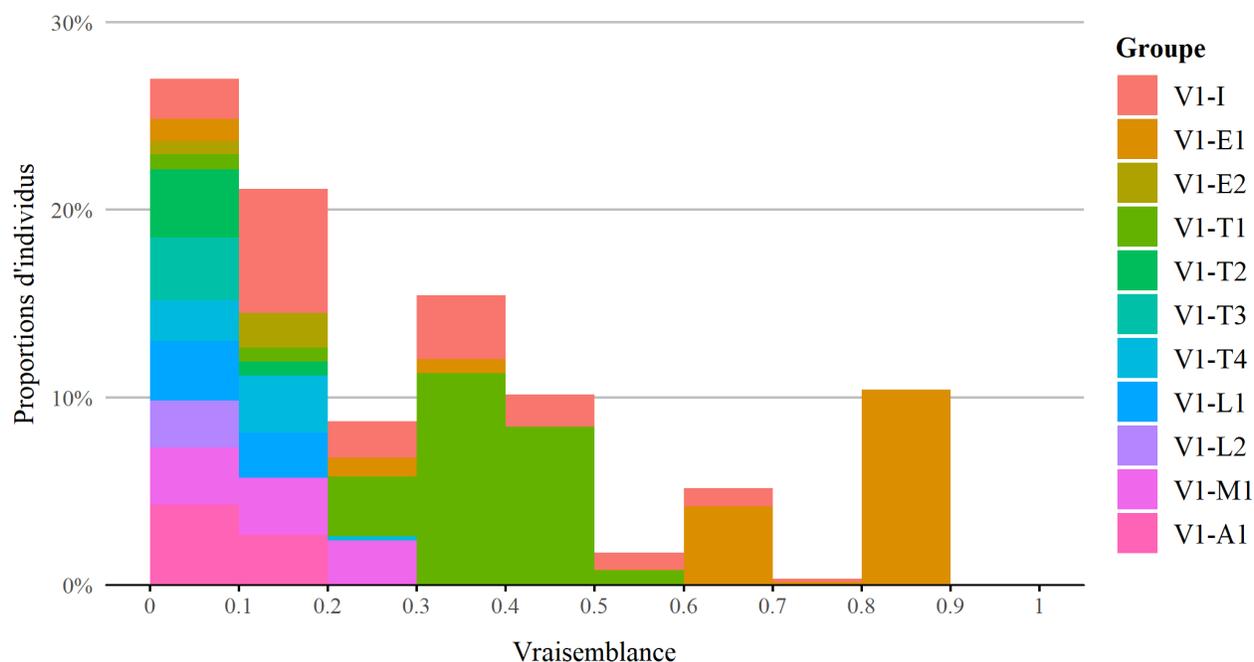


Figure 5-1 Vraisemblance de chaque observation selon le groupe observé (modèle 3)

Plus de la moitié des individus ont une vraisemblance de moins de 0,3, alors que seuls certains membres des groupes V1-E1, V1-T1 et V1-I ont une vraisemblance plus élevée. Ces trois groupes correspondent à ceux ayant les plus grandes proportions d'individus. En outre, la probabilité d'appartenir au groupe observé ne correspond pas nécessairement à la probabilité la plus élevée. Le Tableau 5-28 permet d'identifier cette dernière grâce à un extrait de la base de données.

Tableau 5-28 Extrait de la base de données des personnes permettant d'identifier la probabilité maximale de chaque observation (modèle 3)

IDPERS	Groupe observé	Probabilité d'appartenir à chaque groupe										
		V1-I	V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-L1	V1-L2	V1-M1	V1-A1
1	V1-E1	0,1	0,65	0,16	0	0	0	0,06	0	0	0,01	0,02
2	V1-L2	0,33	0	0	0,09	0	0	0,01	0,11	0,17	0,2	0,1
3	V1-E1	0,05	0,86	0,07	0	0	0	0	0,02	0	0	0
4	V1-L1	0,15	0	0,01	0,34	0,11	0,18	0,2	0	0	0,01	0
5	V1-T2	0,19	0	0	0,55	0,15	0,06	0,05	0	0	0	0
6	V1-E1	0,14	0,61	0,13	0	0	0	0	0	0	0	0,12

Parmi l'ensemble des individus présent dans le jeu de données destiné à l'entraînement du modèle, le groupe de la probabilité la plus élevée correspond au groupe observé pour 45,9 % des

observations. La Figure 5-2 présente la distribution de la proportion d'individus selon la valeur de la probabilité maximale.

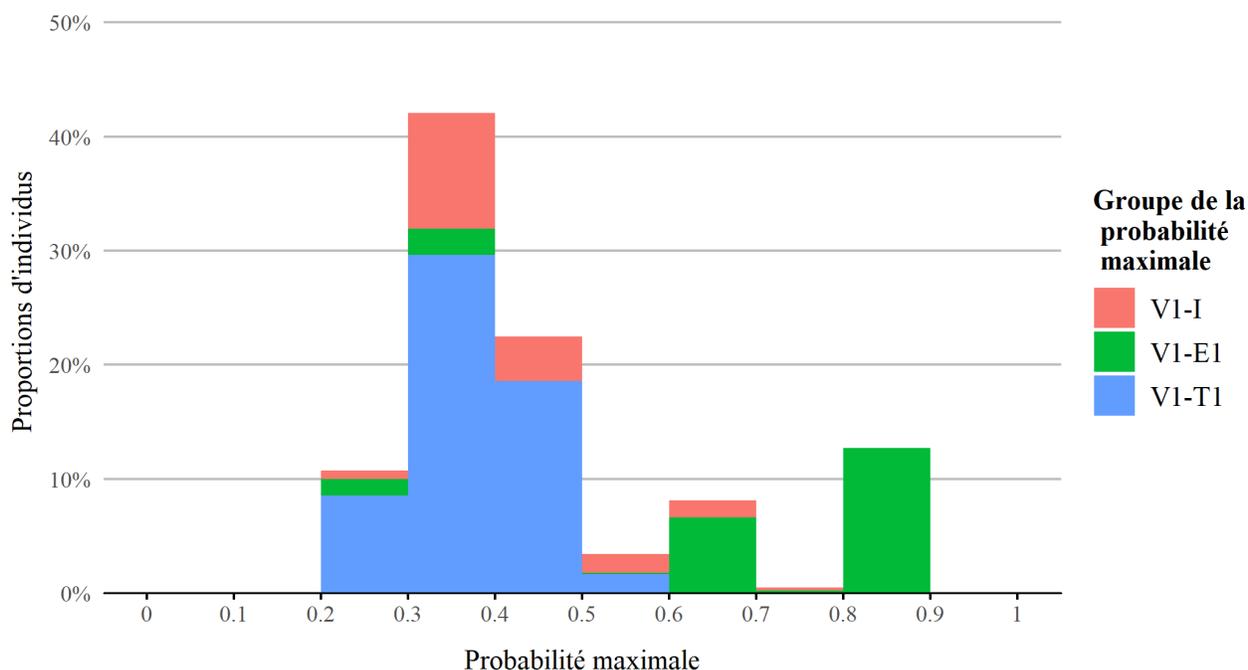


Figure 5-2 Distribution de la proportion d'individus selon la probabilité maximale (modèle 3)

Pour l'ensemble des individus, le groupe de la probabilité maximale est uniquement compris parmi les trois groupes les plus peuplés, soit V1-E1, V1-T1 et V1-I. La probabilité d'appartenir au groupe où l'*étude* est l'activité dominante est particulièrement élevée pour près de 20 % des personnes du jeu d'entraînement, alors que pour la majorité des observations, la probabilité maximale se situe entre 0,2 et 0,5. En effet, 58,4 % des individus composant cet ensemble de données a une probabilité maximale correspondant à V1-T1, tel qu'observé dans le Tableau 5-29, qui indique la proportion d'individus selon le groupe de la probabilité la plus élevée.

Tableau 5-29 Proportion d'individus selon le groupe correspondant à la probabilité maximale (modèle 3)

Groupe	Proportion d'individus
V1-I	18.3%
V1-E1	23.3%
V1-T1	58.4%

Bref, à partir du type de ménage et de l'âge et du genre combiné, le modèle 3 parvient principalement à comprendre les liens entre ces attributs et l'adoption des trois principaux patrons

d'activités, soit étudier, travailler ou demeurer à son domicile. Les probabilités maximales de chaque personne correspondent uniquement à ces trois patrons, soit V1-E1, V1-T1 et V1-I et seule la vraisemblance des individus appartenant réellement à ces trois groupes est supérieure à 0,3.

La vraisemblance de chaque observation peut également être utilisée pour calculer celle de l'ensemble du modèle. Ainsi, les cinq modèles utilisés pour modéliser l'appartenance aux différents groupes issus de la segmentation avec V1 peuvent être comparés sur la base de leur vraisemblance. Cette comparaison peut se faire avec l'aide du critère d'information d'Akaike (AIC). Ce critère évalue l'ajustement statistique relatif d'un modèle et permet de comparer différents modèles dont la variable expliquée est commune à chacun. Ce critère peut être calculé à l'aide de la formule suivante :

$$AIC = -2L + 2P$$

Les deux variables utilisées pour l'évaluation de ce critère sont L et P . La première utilise la fonction de vraisemblance, qui permet d'évaluer l'ajustement statistique du modèle. La valeur de la variable L correspond au maximum du logarithme de cette fonction. La nature relative de ce critère réside entre autres dans l'ajout du nombre de variables explicatives P , qui permet de comparer les différents modèles sur un pied d'égalité. Une condition essentielle à l'application de ce critère est l'utilisation d'un jeu de données unique pour l'estimation des différents modèles. Alors que 80 % du jeu de données initial est sélectionné pour entraîner un modèle, cet échantillon est constant pour l'estimation de chacun des cinq modèles et il est ainsi possible de les comparer sur la base de la valeur de l'AIC. Le modèle dont la valeur du critère est la plus faible est celui qui est le mieux ajusté aux données (Burnham & Anderson, 2004). Le Tableau 5-30 présente la valeur de ce critère pour chacun des cinq modèles.

Tableau 5-30 AIC des 5 modèles

Modèle	1	2	3	4	5
AIC	1 661 642	1 643 172	1 620 363	1 645 041	1 640 628

Le troisième modèle est donc celui qui est le mieux ajusté aux données, à l'inverse du premier qui est celui le moins ajusté. Ainsi, l'hypothèse sur laquelle repose le premier modèle, soit que l'effet de l'âge est linéaire sur l'adoption d'un patron, ne semble pas correspondre à la réalité. En effet, le modèle 2 est davantage ajusté aux données d'entraînement que le premier, grâce à la transformation de la variable de l'âge en variable catégorielle. De plus, le modèle incluant la variable combinant

l'âge et le genre est encore mieux ajusté aux données, confirmant que l'influence du genre peut différer selon l'âge d'une personne et inversement. Finalement, les deux derniers modèles sont également mieux ajustés que le premier, sans l'être autant que le troisième.

5.3.2 Validation des modèles

À la suite de l'estimation des modèles, il est possible de les valider en les appliquant sur le jeu de données destiné à les tester, qui correspond à 20 % des 613 997 personnes qui ont servi à estimer les modèles. L'application du modèle consiste à évaluer les probabilités d'appartenir à chaque groupe, de la même manière que sur le jeu d'entraînement à la section précédente. Ces probabilités peuvent ensuite être cumulées, jusqu'à la valeur de 1. Cette addition correspond à la somme des probabilités pour un individu, soit la fonction de répartition des probabilités. Chaque personne présente dans le jeu de validation a ainsi une probabilité cumulée propre à chacune des catégories de la variable dépendante, soit les groupes issus du Chapitre 4.

Afin de simuler l'appartenance à un des groupes, la méthode de la transformée inverse est utilisée. Ainsi, l'assignation d'un groupe à chaque personne nécessite la génération d'un nombre aléatoire entre 0 et 1. La valeur de ce nombre est ensuite comparée aux probabilités cumulées et le groupe dont la probabilité correspond à ce nombre est sélectionné. Chaque personne dans le jeu de validation se voit donc assigner un patron d'activités typique. Dû à la nature aléatoire de la procédure, ce processus est appliqué dix fois afin de s'assurer de la stabilité des résultats de simulation.

À la suite de l'application du modèle, chaque individu présent dans le jeu de validation se voit assigner un groupe, qui correspond au groupe prédit, s'ajoutant au groupe observé. Il devient donc possible d'analyser la performance du modèle avec la réalité. Le premier élément de cette analyse est présenté dans le Tableau 5-31, qui indique la proportion d'individus dont le groupe observé est le même que celui prédit par le modèle, soit le taux de réussite.

Tableau 5-31 Taux de réussite de chacun des 5 modèles

Modèle	1	2	3	4	5
Taux de réussite	28.6%	28.9%	29.5%	28.6%	29.7%

L'ensemble des modèles ont donc un taux de réussite similaire, variant de 28,6 à 29,7 %. Ainsi, malgré les forces et les faiblesses de chacun d'entre eux, les taux ne sont pas significativement

différents d'un modèle à l'autre. Bref, aucun modèle n'arrive à prédire correctement le groupe pour plus de 70 % des individus.

Les analyses subséquentes sont celles issues de la simulation réalisée avec le modèle 3. Toutefois, les constats issus de ces analyses sont similaires pour les autres modèles. Le Tableau 5-32 indique le pourcentage d'individu dans chacun des groupes. Les proportions présentées correspondent au groupe observé et prédit par le modèle.

Tableau 5-32 Proportion d'individus observée et prédite par le modèle 3 dans chaque groupe

Groupe	Observé	Prédit
V1-I	17.8%	17.7%
V1-E1	17.9%	18.1%
V1-E2	2.5%	2.7%
V1-T1	25.3%	24.9%
V1-T2	4.3%	4.4%
V1-T3	3.3%	3.5%
V1-T4	5.5%	5.6%
V1-L1	5.5%	5.8%
V1-L2	2.5%	2.4%
V1-M1	8.5%	8.2%
V1-A1	7.0%	6.7%
Total	100%	100%

Malgré un faible taux de réussite, le modèle parvient tout de même à prédire le bon nombre d'individus dans chacun des groupes. Toutefois, en disséquant ces proportions à l'aide d'une matrice de confusion, il est possible de comprendre pourquoi le modèle ne parvient pas à prédire le bon groupe pour plus de 70 % des personnes composant le jeu de validation. La matrice est présentée dans le Tableau 5-33. Elle indique le nombre d'individus selon le groupe observé et le groupe prédit. Le rappel et la précision sont également indiqués pour chacun des groupes. Le rappel correspond au pourcentage de personnes dont le groupe prédit est le même que celui observé, parmi les individus appartenant au même groupe observé. Alors que la précision, comme son nom l'indique, permet d'évaluer à quel point le modèle est précis. Il correspond à la proportion de personnes, parmi l'ensemble des individus dont le groupe prédit est le même, dont le groupe observé est identique à celui assigné par le modèle.

Tableau 5-33 Matrice de confusion du modèle 3 (V1)

Prédit	Observé											Précision
	I	E1	E2	T1	T2	T3	T4	L1	L2	M1	A1	
I	5678	1734	275	4647	764	626	874	1658	794	2872	1747	26%
E1	1818	14710	1978	1623	339	129	245	502	97	334	440	66%
E2	322	2116	279	259	63	31	42	94	29	49	70	8%
T1	4612	1442	250	11446	1929	1475	2463	1474	708	2309	2432	37%
T2	804	338	64	2007	356	275	437	242	103	383	383	7%
T3	664	149	32	1608	260	206	381	200	127	323	355	5%
T4	931	220	36	2602	444	326	777	293	141	466	628	11%
L1	1831	508	84	1531	253	216	294	595	283	946	589	8%
L2	782	122	24	662	123	102	120	241	123	408	273	4%
M1	2603	292	41	2322	357	333	489	899	433	1401	915	14%
A1	1763	371	45	2301	409	294	624	580	256	923	699	8%
Rappel	26%	67%	9%	37%	7%	5%	12%	9%	4%	13%	8%	

Parmi l'ensemble des personnes appartenant réellement au groupe V1-E1, le groupe prédit par le modèle est le même que celui observé pour 14 710 d'entre eux, ce qui correspond à 67 % des personnes dont le groupe observé est V1-E1, soit le rappel. De plus, parmi l'ensemble des personnes dont le groupe prédit par le modèle est V1-E1, 14 710 d'entre eux font réellement partie de ce groupe, ce qui équivaut à une précision de 66 %. Ainsi, le modèle offre une bonne précision concernant l'assignation du groupe V1-E1, contrairement au second groupe où l'*étude* est l'activité dominante, V1-E2, dont le modèle est précis à uniquement 8 %. La majorité des membres composant réellement ce groupe, soit 1 978 personnes, se sont vu attribuer V1-E1. Bref, le modèle réussit bien à prédire si une personne risque de réaliser un patron d'activité dont la principale activité est l'*étude*, comme en fait foi la vraisemblance élevée des observations composant le groupe V1-E1. Toutefois, le modèle ne parvient pas à bien classer les individus entre V1-E1 ou V1-E2. Autrement dit, à partir de l'âge, du genre et du type de ménage, le modèle est incapable de prédire si une personne risque de réaliser une activité de type *loisir* ou non après l'école.

Le même raisonnement s'applique aux groupes où le *travail* est la principale activité. Le modèle est relativement précis dans le cas du groupe V1-T1, avec un taux de 37 %, mais la précision chute à 7 % pour V1-T2, 5 % pour V1-T3 et 11 % pour V1-T4. Ainsi, le modèle assigne à la majorité des personnes dans ces trois groupes le patron V1-T1. La précision plus élevée dans le cas de V1-T4 peut être expliquée par la forte proportion de personnes dans ce groupe dont le type de ménage

correspond aux « Jeunes familles ». Cette caractéristique permettrait donc au modèle de mieux prédire l'adoption du patron V1-T4.

La classification des individus dans les quatre derniers groupes n'offre également pas une grande précision. Les patrons prédits pour les membres de ces groupes sont principalement V1-T1 et V1-I. Bref, bien que le modèle réussisse à répliquer les distributions observées du nombre d'individus dans chaque groupe, il ne réussit pas à bien prédire le patron réellement réalisé pour la majorité des personnes.

5.4 Deuxième segmentation

La segmentation réalisée avec le vecteur V2 a été appliquée à un échantillon de 14 965 personnes, soit environ 3 % des 504 776 personnes qui ont réalisé au moins une activité selon les enquêtes OD de 1998 à 2013. Afin d'estimer le modèle permettant d'évaluer les probabilités d'appartenir aux groupes issus de ce partitionnement, il est nécessaire d'inclure des individus n'ayant réalisé aucune activité. Un groupe supplémentaire est ainsi ajouté, tout comme pour le modèle estimé avec la première méthode de segmentation. Ces individus sont sélectionnés aléatoirement parmi les 109 221 personnes qui ont été immobiles lors de la journée où leur ménage a été sondé. Le même taux est utilisé pour sélectionner ces individus, soit 3 %, ce qui correspond à 3 379 individus. À la suite de l'ajout de ce groupe, il est possible de scinder ce jeu de données en deux, pour l'entraînement et la validation du modèle. Les mêmes proportions que celles utilisées lors de l'estimation des modèles avec la première segmentation sont appliquées, soit 80 et 20 %. Finalement, l'analyse de corrélation des variables avec ce jeu d'entraînement révèle les mêmes constats que les données avec V1.

5.4.1 Analyse des modèles

Les cinq modèles sont estimés avec le jeu de données d'entraînement. Chacun des modèles révèle des liens intéressants entre les patrons typiques et les attributs individuels. Cependant, les rapports de probabilités d'un seul d'entre eux sont analysés, soit ceux du deuxième modèle, puisque la majorité des coefficients et les rapports des autres modèles ne sont pas significatifs. Les rapports de probabilité du modèle 2 sont présentés dans les Tableau 5-34 et Tableau 5-35.

Tableau 5-34 Rapport de probabilités du modèle 2 – *Études et travail*

	V2-E1	V2-T1	V2-T2	V2-T3	V2-T4	V2-T5	V2-I
Constante	0.01***	0.23***	0.02***	0.24***	0.03***	0.02***	-
5-9	2110.01***	0.02**	0.04	0.02**	0.03	0.09	-
10-14	2324.6***	0.02**	0.04	0.02**	0.02	0.07	-
15-19	815.19***	0.77	4.01***	0.73	1.64	1.17	-
20-24	244.76***	2.99***	10.46***	2.98***	4.51***	7.41***	-
25-29	80.28***	5.87***	9.08***	5.01***	8.31***	15.55***	-
30-34	39.99***	9.01***	9.25***	7.61***	7.09***	14.76***	-
35-39	20.1***	7.43***	5.43***	5.97***	7.6***	10.19***	-
40-44	21.64***	10.13***	10.3***	9.1***	6.64***	15.03***	-
45-49	7.36**	8.46***	7.32***	7.47***	4.68***	11.59***	-
50-54	3.56	7.42***	7.59***	6.58***	3.42***	8.47***	-
55-59	2.23	4.47***	3.04***	4.4***	2.92***	4.37***	-
60+	-	-	-	-	-	-	-
Jeunes couples, colocataires et célibataire	0.72**	1.11	1.37	0.8*	1.27	1.79***	-
Jeunes familles	-	-	-	-	-	-	-
Familles traditionnelles	0.97	0.95	1.13	0.83*	1.33	1.36	-
Adultes et personnes âgées	0.56	0.36***	0.87	0.35***	0.33***	0.5*	-
Cinquantennaires	0.7	0.75**	0.93	0.69***	0.76	1.18	-
Jeunes retraités	0.38	0.64**	0.9	0.57***	0.76	0.88	-
Vieux retraités	0.07	0.02***	0.2**	0.05***	0.16***	0.01	-
Familles multigénérationnelles	0.61***	0.63***	1.01	0.46***	0.87	0.71	-
Homme	0.96	1.9***	2.1***	1.54***	2.02***	2.05***	-
Femme	-	-	-	-	-	-	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

La majorité des rapports pour la variable de l'âge sont significatifs à $p < 0.01$, contrairement à ceux de la variable indiquant le type de ménage, dont une minorité est significative. Tout comme le modèle 2 de la première segmentation, les chances d'adopter le patron où l'activité dominante est l'étude, soit V2-E1, plutôt que d'adopter V2-I, diminuent avec l'augmentation de l'âge.

Les chances d'adopter les patrons où le travail est l'activité dominante plutôt que de rester à la maison selon l'âge sont a priori similaires pour l'ensemble des patrons. Toutefois, quelques différences sont observables entre les différents patrons. Tout d'abord, les individus entre 20 et 24 ans ont 10,46 fois plus de chances que ceux ayant plus de 60 ans d'adopter V2-T2 plutôt que V2-I. Ces chances sont considérablement plus faibles pour les autres patrons, indiquant que les personnes de cet âge ont davantage tendance à travailler de soir, qui est la principale caractéristique

de V2-T2. Une situation similaire est observée pour les personnes de 45 à 59 ans, dont les rapports correspondant à V2-T4 sont plus faibles que les autres patrons, ce qui révèle que ces individus ont moins de chances de réaliser des activités de type *travail* d'une longue durée.

Les personnes vivant au sein de ménages d'un type autre que celui de référence ont généralement moins de chances que celles appartenant aux « Jeunes familles » de réaliser un patron où le *travail* ou l'*étude* est l'activité dominante. Ce n'est cependant pas le cas des « Jeunes couples, colocataires et célibataires » qui sont plus enclins à réaliser une activité de type *loisir* après le *travail* (V2-T5) que les membres des « Jeunes familles », avec 79 % plus de chance de réaliser ce patron plutôt que d'être immobile.

Finalement, les hommes ont de 54 à 110 % plus de chances que les femmes d'adopter les patrons V2-T1 à V2-T5 plutôt que V2-I, confirmant que les femmes ont plus de chances de ne faire aucune activité par rapport aux hommes. Le rapport le moins élevé correspond à celui de V2-T3, qui se différencie par une heure de début de l'activité dominante plus hâtive que celle des autres patrons.

Dans le cas des autres patrons, encore une fois, les membres des « Jeunes familles » semblent avoir généralement davantage de chances de réaliser les patrons où l'activité principale est de type *loisir*, *magasinage* ou *autres* que de demeurer au domicile, à l'exception de V2-L2. En effet, les « Jeunes retraités » ont 2,35 fois plus de chances que les membres du type de ménage de référence d'adopter ce patron.

Les individus ayant entre 25 et 59 ans ont de 2,11 à 3,66 fois plus de chances que ceux ayant plus de 60 ans d'adopter V2-A1 plutôt que V2-I. Une situation similaire est observée pour V2-MA1, mais qui ne peut être complètement confirmée puisque certains rapports ne sont pas significatifs. De plus, seuls trois rapports sont significatifs dans le cas de l'âge et des patrons où le *loisir* est la principale activité. Ces rapports indiquent que des personnes d'un certain âge, 50 à 59 ans pour V2-L1 et 25 à 29 ans pour V2-L2, ont davantage de chances que celles ayant plus de 60 ans de réaliser ces patrons. Cette situation pourrait être expliquée par le grand nombre de personnes de cet âge qui restent à la maison.

Tableau 5-35 Rapport de probabilités du modèle 2 – *Loisir, magasinage et autres*

	V2-L1	V2-L2	V2-A1	V2-MA1	V2-I
Constante	0.15***	0.13***	0.05***	1.12	-
5-9	0.94	1.59	1.89	0.33***	-
10-14	0.95	1.09	0.12	0.17***	-
15-19	1.01	0.93	0.65	0.66**	-
20-24	1.58	0.77	1.76	1.29	-
25-29	1.09	1.73*	2.11*	1.17	-
30-34	1.13	1.17	2.13*	1.67***	-
35-39	1.14	0.73	2.58**	1.29	-
40-44	1.57	0.89	3.09***	1.42**	-
45-49	1.05	0.89	3.66***	1.16	-
50-54	1.99**	0.75	2.76***	1.36**	-
55-59	1.61*	1.26	2.28**	1.2	-
60+	-	-	-	-	-
Jeunes couples, colocataires et célibataire	1.34	1.08	0.69	0.73**	-
Jeunes familles	-	-	-	-	-
Familles traditionnelles	0.63**	0.99	0.59**	0.76**	-
Adultes et personnes âgées	0.47**	1.04	0.42**	0.52***	-
Cinquantennaires	0.65	1.56	0.42***	0.76**	-
Jeunes retraités	1.18	2.35***	0.69	0.98	-
Vieux retraités	0.58	1.24	0.46*	0.6***	-
Familles multigénérationnelles	0.36***	0.86	0.63	0.58***	-
Homme	0.96	1.17	2.57***	1.18***	-
Femme	-	-	-	-	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Les rapports de probabilités des autres modèles sont présentés dans l'Annexe C, dans les Tableau C-18 à Tableau C-25. Ces modèles, malgré une majorité de coefficients et de rapports non significatifs, permettent quand même de tirer des conclusions quant à l'influence des attributs individuels sur les systèmes d'activités. Par exemple, le troisième modèle, tout comme pour les patrons de la première segmentation, révèle que l'influence de l'âge varie en fonction du genre dans le cas de l'adoption des patrons où le *travail* est la principale activité. Cette variation, qui n'est toutefois pas observée pour les autres patrons, permet de comprendre que l'hypothèse sur laquelle repose le modèle 2 n'est pas fondée, du moins dans le cas des patrons V2-T1 à V2-T5. Il est également possible de tirer la même conclusion avec l'AIC des modèles, qui sont présentés dans le Tableau 5-36.

Tableau 5-36 AIC des 5 modèles

Modèle	1	2	3	4	5
AIC	47 941	46 756	46 750	47 135	47 088

Ainsi, le modèle le mieux ajusté est le troisième, tout juste devant le modèle 2. Ce dernier est également mieux ajusté que le premier modèle, dont les rapports de la variable de l'âge, à l'exception du patron V2-E1, varient entre 0,97 et 1, ce qui représente une faible variation.

Les deux derniers modèles, dont l'ajustement statistique est de meilleure qualité que le premier modèle sans l'être autant que les modèles 2 et 3, permettent d'observer des différences concernant l'influence du type de ménage selon le genre et l'âge. Les rapports du modèle 4 indiquent que les hommes ont davantage de chances que les femmes d'adopter les patrons dominés par le *travail*, à l'exception des membres des « Jeunes couples, colocataires et célibataires », qui peu importe le genre, ont autant de chances d'adopter V2-T2 plutôt que V2-I. Finalement, le dernier modèle révèle que les personnes entre 25 et 44 ans n'adoptent pas les mêmes comportements selon le type de ménage dans lequel ils vivent.

5.4.2 Validation des modèles

Grâce au jeu de validation, il est possible de valider les cinq modèles de la deuxième segmentation, de la même manière que les modèles de la première segmentation ont été validés. Le premier élément de cette validation concerne le taux de réussite de chacun des modèles, présenté dans le Tableau 5-37.

Tableau 5-37 Taux de réussite de chacun des 5 modèles

Modèle	1	2	3	4	5
Taux de réussite	30.7%	31.1%	30.3%	30.1%	30.1%

Le taux de chacun des modèles est similaire, variant de 30,1 à 31,1 %. Ainsi, l'ensemble des modèles sont a priori tout aussi performants, comme le prouvent les proportions de personnes dans chacun des groupes, observés et prédits par les modèles. En effet, chacun d'entre eux réussit à répliquer la distribution du pourcentage de personnes dans chacun des groupes. En outre, la matrice de confusion du modèle 2, qui est présentée dans le Tableau 5-38, est également similaire à celle des autres modèles.

Tableau 5-38 Matrice de confusion du modèle 2 (V2)

Prédit	Observé											Précision
	I	E1	T1	T2	T3	T4	T5	L1	L2	A1	MA1	
I	167	56	78	19	79	10	15	15	35	11	136	27%
E1	62	541	22	9	15	1	6	10	9	4	45	75%
T1	92	16	160	18	136	22	24	15	15	11	128	25%
T2	13	2	18	1	8	7	4	2	0	2	17	1%
T3	57	18	115	6	71	12	17	14	16	13	87	17%
T4	13	2	23	3	15	6	2	2	1	2	11	8%
T5	11	4	33	3	14	6	5	2	1	1	21	5%
L1	25	6	23	0	9	4	1	4	6	1	23	4%
L2	35	9	18	5	14	3	4	4	8	6	24	6%
A1	9	6	8	1	6	1	0	2	0	2	11	4%
MA1	163	33	122	10	99	17	17	21	30	11	172	25%
Rappel	26%	78%	26%	1%	15%	7%	5%	4%	7%	3%	25%	

Les groupes où le modèle 2 est le plus précis sont les plus nombreux, soit V2-I, V2-E1, V2-T1, V2-T3 et V2-MA1. La précision concernant V2-E1 est la plus élevée, alors que celle des autres groupes est similaire. Parmi les membres de V2-L2, le groupe prédit pour la majorité d'entre eux est soit celui observé, soit V2-I. La forte proportion de retraités dans ce groupe permet de comprendre pourquoi le modèle est confus entre ces deux groupes uniquement, alors qu'une plus faible proportion d'entre eux se voit attribuer un patron dont la principale activité est le *travail*. C'est cependant l'inverse qui est observé dans le cas de V2-L1, V2-A1, V2-MA1, où le modèle n'arrive pas bien à prédire si leurs membres vont travailler, rester à la maison ou réaliser une activité de type *magasinage, loisir* ou *autres*.

Tout comme les modèles de la segmentation avec V1, le modèle 2 du partitionnement avec V2 ne parvient pas à différencier les membres des groupes où le *travail* est la principale activité. Toutefois, le modèle assigne à une proportion considérable de membres de V2-T2 le groupe V2-I, ce qui signifie que le modèle confond les personnes restant à la maison avec celles travaillant le soir, qui est la principale caractéristique du système d'activités des membres de V2-T2.

5.5 Troisième segmentation

L'analyse et la validation des modèles avec la segmentation de V3 sont présentés dans cette section.

5.5.1 Analyse des modèles

Tout comme la deuxième segmentation, seuls les résultats du modèle 2 sont présentés dans cette section. Les rapports de probabilité de ce modèle sont présentés dans le Tableau 5-39.

Tableau 5-39 Rapport de probabilités du modèle 2 – *Études et travail*

	V3-E1	V3-T1	V3-T2	V3-T3	V3-T4	V3-T5	V3-I
Constante	0.01***	0.23***	0.02***	0.21***	0.03***	0.02***	-
5-9	3152.65***	0.03**	0.05	0.03**	0.03	0.04	-
10-14	3377.77***	0.03**	0.06	0.03**	0.03	0.04	-
15-19	1186.03***	0.68	4.65***	0.95	1.32	1.03	-
20-24	317.88***	3.3***	11.98***	2.88***	5.05***	5.69***	-
25-29	106.59***	6.51***	10.64***	5.18***	9.36***	11.22***	-
30-34	47.78***	9.82***	10.98***	7.65***	6.99***	12.11***	-
35-39	24.57***	8.36***	6.02***	6.16***	7.18***	7.83***	-
40-44	22.01***	11.12***	12.46***	9.51***	6.55***	10.95***	-
45-49	5.95	8.89***	9.22***	7.97***	4.51***	8.35***	-
50-54	5.72	7.55***	9.83***	7.46***	2.77**	6.35***	-
55-59	3.65	4.7***	4.07***	4.66***	2.76***	3.17***	-
60+	-	-	-	-	-	-	-
Jeunes couples, colocataires et célibataire	0.71**	1.04	1.38	0.84	1.27	1.86***	-
Jeunes familles	-	-	-	-	-	-	-
Familles traditionnelles	0.94	0.94	1.11	0.85	1.3	1.69***	-
Adultes et personnes âgées	0.52	0.34***	1	0.41***	0.38**	0.46*	-
Cinquantennaires	0.8	0.71***	0.81	0.75**	0.96	1.41	-
Jeunes retraités	0.38	0.57***	0.95	0.66**	0.71	1.06	-
Vieux retraités	0.2	0.01***	0.33	0.07***	0.1***	0.01	-
Familles multigénérationnelles	0.59***	0.65***	0.97	0.46***	0.69	0.88	-
Homme	0.95	1.83***	2.19***	1.59***	2.14***	2.1***	-
Femme	-	-	-	-	-	-	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Les patrons où les types travail et études sont l'activité dominante sont très similaires à ceux de la deuxième segmentation. Les rapports de probabilité correspondant aux patrons V3-E1 et V3-T1 à V3-T5 sont également similaires à ceux de V2-E1 et V2-T1 à V2-T5. Les rapports concernant les patrons caractérisés par le *travail* révèlent donc les mêmes constats, soit que les personnes entre 20 et 59 ans ont davantage de chances que les personnes plus âgées d'adopter ces patrons. Les rapports concernant le genre et le type de ménage sont également similaires.

Tableau 5-40 Rapport de probabilités du modèle 2 – *Loisir, magasinage et autres*

	V3-L1	V3-A1	V3-MA1	V3-MA2	V3-I
Constante	0.22***	0.07***	0.68**	0.51***	-
5-9	0.64	1.52	0.36***	0.71	-
10-14	0.65	0.05	0.27***	0.68	-
15-19	0.91	0.45	1.08	0.83	-
20-24	1.38	1.69	1.8***	0.98	-
25-29	1.14	2.32**	1.59**	0.75	-
30-34	1	2.19*	1.85***	1.3	-
35-39	1.25	2.42**	1.39*	0.96	-
40-44	1.27	2.37**	1.77***	1.02	-
45-49	1	3.33***	1.4**	0.89	-
50-54	1.72*	3.08***	1.46**	0.94	-
55-59	1.18	2.59***	1.27*	1.25	-
60+	-	-	-	-	-
Jeunes couples, colocataires et célibataire	1.33	0.62*	0.85	0.64**	-
Jeunes familles	-	-	-	-	-
Familles traditionnelles	0.66*	0.55***	0.8*	0.8	-
Adultes et personnes âgées	0.52**	0.29***	0.59***	0.55***	-
Cinquantennaires	0.54**	0.41***	0.87	0.84	-
Jeunes retraités	1.21	0.67	1.14	1.07	-
Vieux retraités	0.48**	0.38**	0.63**	0.73	-
Familles multigénérationnelles	0.3***	0.56*	0.72*	0.47***	-
Homme	0.93	2.56***	1.32***	1	-
Femme	-	-	-	-	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

La principale caractéristique qui distingue la segmentation V3 par rapport à celle de V2 est la création de deux groupes où les membres ont de multiples activités dominantes, dont les types sont principalement *autres*, *magasinage* ou *loisir*. Les rapports de probabilités de la variable de l'âge ne sont pas significatifs dans le cas de V3-MA2, mais le sont dans le cas de V3-MA1. Ceux-ci révèlent que, tout comme les patrons dominés par le *travail*, que les personnes de 20 à 59 ont davantage de chances que les personnes plus âgées de réaliser V3-MA1 plutôt que d'être immobile. Malgré le fait qu'un grand nombre de retraités sont présents dans le groupe V3-MA1 à la suite de la segmentation, ceux-ci sont également nombreux à ne réaliser aucune activité. Une situation similaire est observée dans le cas de V3-A1.

Finalement, les rapports des autres variables indiquent que les membres des « Jeunes familles » ont moins de chances que les membres des autres types de ménage de ne faire aucune activité. De plus,

les hommes ont également 156 et 32 % plus de chances que les femmes d'adopter V3-A1 et V3-MA1 respectivement.

Les rapports de probabilité des autres modèles, présentés à l'Annexe C dans les Tableau C-28 et Tableau C-37, sont similaires aux modèles réalisés avec la segmentation de V2. De plus, les constats concernant l'ajustement statistique des cinq modèles obtenus avec V3 sont également similaires à ceux estimés avec V2. L'AIC de ces modèles sont présentés dans le Tableau 5-41.

Tableau 5-41 AIC des 5 modèles

Modèle	1	2	3	4	5
AIC	49 609	48 428	48 448	48 822	48 779

Les différences entre l'AIC des modèles estimés avec V3 sont les mêmes qu'avec ceux de V2. Ainsi, les modèles 2 et 3 sont les mieux ajustés pour les données d'entraînement, alors que le premier est celui qu'il l'est le moins.

5.5.2 Validation des modèles

Les cinq modèles estimés avec V3 ont également été validés suivant la même méthodologie ayant permis de valider les modèles de la segmentation précédente. Les taux de réussite de ces modèles sont présentés dans le Tableau 5-42.

Tableau 5-42 Taux de réussite de chacun des 5 modèles

Modèle	1	2	3	4	5
Taux de réussite	27.7%	30.3%	29.6%	27.8%	29.2%

Le taux de réussite sont similaires pour l'ensemble des cinq modèles. Ces taux sont également similaires à ceux des dix autres modèles estimés dans les sections précédentes, soit environ 30 %. La matrice de confusion du modèle 2 est présentée dans le Tableau 5-43.

Tableau 5-43 Matrice de confusion du modèle 2 (V3)

Prédit	Observé											Précision
	I	E1	T1	T2	T3	T4	T5	L1	A1	MA1	MA2	
I	162	58	110	10	72	15	10	27	17	138	63	24%
E1	58	497	28	5	17	3	7	8	2	32	23	73%
T1	89	18	179	17	133	26	21	23	14	75	41	28%
T2	13	4	14	1	4	3	3	3	3	16	3	1%
T3	54	4	114	10	73	14	11	17	9	65	33	18%
T4	16	8	19	3	16	4	4	0	1	15	4	4%
T5	10	5	13	2	17	5	2	4	3	17	6	2%
L1	32	11	22	1	17	0	2	7	1	22	8	6%
A1	12	7	15	1	9	1	3	3	6	9	7	8%
MA1	133	29	78	18	83	7	9	14	9	96	54	18%
MA2	68	17	42	8	29	5	4	13	5	47	29	11%
Rappel	25%	76%	28%	1%	16%	5%	3%	6%	9%	18%	11%	

Plusieurs constats découlant de l'analyse de la matrice de confusion du deuxième modèle avec V2 sont également applicables à cette matrice. Les patrons les plus observés dans le sous-échantillon sont ceux ayant la meilleure précision. De plus, la création de deux groupes où l'activité dominante est multiple (V3-MA1 et V3-MA2), contrairement à la deuxième segmentation où il n'y en avait qu'un, induit une baisse de précision pour ces deux groupes. Cette précision est plus faible que celle de V2-MA1.

5.6 Comparaison des méthodes

Grâce à l'estimation des modèles, les patrons typiques peuvent être utilisés pour générer des activités à l'aide des distributions développées selon la méthode présentée à la section 3.4.2. Ainsi, il devient possible de comparer les différents modèles, estimés à partir des segmentations obtenues avec les trois vecteurs, sur la base de leur capacité à générer une série d'activités pour chaque individu.

5.6.1 Sélection des personnes pour fin de comparaison

Le principal critère de comparaison concerne la capacité des patrons typiques et des modèles qui en découlent, à reproduire les activités observées dans l'enquête OD. La pertinence de cette comparaison repose sur une division appropriée des données. En effet, il serait inutile de comparer des activités générées avec celles observées si ces dernières ont servi au développement de la

méthode visant à les générer. C'est pourquoi il est nécessaire de partitionner les données provenant des enquêtes en deux segments distincts. Le premier segment vise à estimer les modèles et à construire les distributions des attributs des activités. Le modèle peut alors être appliqué sur le second segment et les distributions peuvent être utilisées pour générer les activités des personnes composant ce second segment.

Comme mentionné dans la méthodologie, 613 977 personnes composent les échantillons des quatre enquêtes OD et sont âgées de cinq ans et plus. Lors de l'estimation des modèles de la segmentation avec V1, ce jeu de données fut divisé en deux groupes afin d'estimer et valider les cinq modèles. Le jeu de validation peut également être utilisé pour tester la procédure de génération des activités. De plus, le jeu d'entraînement, à partir duquel les cinq modèles sont estimés, peut également servir à la construction des distributions des attributs des activités. À titre de rappel, ce jeu comprend 80 % des individus présent parmi les 613 977 personnes provenant des quatre échantillons. Par la suite, les modèles peuvent être appliqués sur le jeu de validation et les distributions peuvent être utilisées afin de générer des activités. Chaque individu composant ce jeu de données se voit attribuer un patron typique ainsi qu'une série d'activités, dont les attributs sont pigés aléatoirement dans les distributions. Par la suite, il est possible de comparer ces activités générées avec celles que les individus présents dans le jeu de validation ont réellement réalisées lors de la journée où leur ménage a été sondé.

Dans le cas de V2 et V3, le partitionnement de ces vecteurs est effectué avec uniquement 14 965 personnes due à la nature de l'algorithme de segmentation. Ce groupe d'individus, auquel un ensemble de personnes n'ayant réalisé aucune activité a été ajouté, est utilisé pour estimer et valider les cinq modèles de ces deux vecteurs. Un ensemble de données composé de 18 344 personnes est ainsi utilisé. Toutefois, celles-ci ne font pas partie du jeu de données visant à valider les modèles du premier vecteur. Cette exclusion permet d'utiliser ce jeu de validation pour la comparaison des quinze modèles, soit les cinq modèles des trois vecteurs. Dans le cas des deux derniers vecteurs, les distributions sont construites à partir des 14 965 individus sur qui la segmentation hiérarchique est appliquée. Ces distributions peuvent par la suite être utilisées pour générer les activités des individus composant le jeu de validation des modèles de V1, sur qui les modèles de V2 et V3 sont également appliqués.

5.6.2 Analyse des activités générées

À la suite de l'application de l'ensemble de la procédure, chaque personne présente dans le jeu de validation des modèles du premier vecteur se voit attribuer une série d'activités. Celles-ci peuvent être comparées avec les activités observées lors des enquêtes. Au total, les personnes composant ce jeu de validation, qui sont au nombre de 122 800, ont réalisé un total de 183 310 activités d'une durée totale de 508 774 heures. Le Tableau 5-44 présente la proportion de ces activités observées qui ont été générées par l'ensemble des processus génératifs utilisant les quinze modèles. Cette proportion est présentée en termes de nombre d'activités, mais également en termes de durée d'activité.

Tableau 5-44 Proportion d'activités générées parmi celles observées

Modèle	Nombre d'activité			Durée d'activité		
	V1	V2	V3	V1	V2	V3
1	99.23%	98.49%	98.63%	99.25%	99.81%	99.83%
2	101.84%	99.92%	99.98%	100.42%	100.93%	100.57%
3	98.90%	98.73%	99.23%	97.82%	98.31%	98.24%
4	101.14%	99.38%	99.56%	100.05%	100.23%	100.48%
5	99.51%	99.73%	99.50%	101.14%	101.02%	100.92%

Parmi les 183 310 activités observées, 99,23 % d'entre elles sont générés grâce au processus utilisant le modèle 1 avec le vecteur V1. Ce taux est également similaire pour l'ensemble des autres modèles, dont les valeurs varient entre 97,82 % et 101,84 %. Il est également possible d'approfondir l'analyse des activités générés en calculant ce taux selon l'âge et le genre, ainsi que le type d'activité. Ces taux sont indiqués dans le Tableau 5-45. Les détails du calcul de ces taux sont présentés en Annexe D dans le Tableau D-38.

Tableau 5-45 Proportion d'activités générées par rapport à celles observées par segment de population et par type d'activité – Segmentation V1 – Modèle 3

Âge	Genre	Travail	Études	Loisir	Magasinage	Autre	Total
5 - 9	H	0.00	0.99	1.38	1.09	0.90	1.02
	F	0.00	1.00	1.36	1.14	0.92	1.02
10 - 19	H	3.22	0.73	1.10	2.65	1.69	1.02
	F	1.04	1.01	1.12	0.77	1.12	1.02
20 - 29	H	1.07	0.37	0.83	1.73	1.46	1.00
	F	1.15	0.89	1.17	0.71	1.04	1.03
30 - 39	H	0.87	0.72	1.11	1.56	1.01	1.00
	F	1.22	0.86	0.94	0.73	0.93	1.01
40 - 49	H	1.01	0.98	1.07	1.07	1.04	1.03
	F	0.95	0.87	1.04	1.15	0.90	0.98
50 - 59	H	0.62	2.47	1.21	1.54	0.76	0.87
	F	1.19	0.54	0.90	0.77	1.18	1.05
60 - 69	H	0.23	1.75	1.10	1.07	0.86	0.82
	F	1.73	1.29	0.97	0.81	1.14	1.06
70 +	H	0.44	1.41	0.75	0.75	0.69	0.72
	F	1.95	0.41	1.26	1.09	1.27	1.21
Total		1.01	0.87	1.05	1.04	1.01	0.99

Il existe des différences entre les hommes et les femmes concernant la proportion d'activités générées dans le cas des activités de type *magasinage*. En effet, il y a plus d'activités générées que d'observées dans le cas des hommes, alors que c'est l'inverse pour les femmes. Pour les activités de type *loisir*, le modèle surestime le nombre d'activités pour les moins de 19 ans, contrairement aux autres segments de population. Le nombre d'activités générées de type *études* est similaire à ce qui est observé pour les personnes de 5 à 19 ans, à l'exception des hommes de plus de 10 ans. Finalement, le nombre d'activités générées de type travail semble similaire au nombre d'activités observées pour les individus de 20 à 49 ans.

Bref, de façon générale, l'ensemble des modèles ayant pour but de prédire l'appartenance à chacun des trois vecteurs ont la capacité de générer des activités qui sont semblables à ce qui est observé dans l'enquête OD. Toutefois, un autre critère indispensable à l'évaluation des trois typologies développées dans ce mémoire est leur habilité à capter des comportements au sein de la population qui connaissent une certaine évolution temporelle. C'est cette évolution qui fait l'objet du chapitre suivant.

CHAPITRE 6 ANALYSE DES TRANSFORMATIONS DES SYSTÈMES D'ACTIVITÉS

Les trois typologies développées dans le cadre de ce mémoire ont comme principal objectif de détecter des comportements typiques en matière de système d'activités au sein de la population et ce, parmi la multitude de comportements qui existent. Comme mentionné dans le Chapitre 4, la pertinence de ces typologies repose sur leurs habilités à capturer l'évolution des comportements qui ont une influence sur les décisions des individus par rapport au choix de mobilité auxquels ils sont confrontés. Le but de ce chapitre est d'évaluer cette pertinence. Il s'agit donc de comprendre, d'analyser, de quantifier et de modéliser les tendances concernant l'évolution des patrons composant ces typologies afin de dégager des constats sur la transformation des systèmes d'activité. Ces constats permettront ainsi de s'assurer de la pertinence des résultats découlant de ce mémoire en vue de leur intégration à la procédure de prévision de la demande en transport.

L'analyse de ces transformations se fait en deux temps. Tout d'abord, pour chaque segmentation, l'évolution temporelle des patrons est analysée grâce à l'observation des proportions de chacun des groupes pour chacune des quatre enquêtes. Ces proportions sont par la suite exposées pour différents segments de la population.

Le deuxième élément de ce chapitre a pour but de modéliser les transformations des systèmes d'activités. Afin d'y parvenir, un des modèles développés dans le chapitre précédent est modifié afin d'y intégrer une composante prévisionnelle.

6.1 Évolution temporelle des patrons

L'analyse des transformations découlant des trois typologies est présentée dans cette section. Cette analyse est possible grâce à la fusion des données provenant des quatre enquêtes Origine-Destination, soit celles de 1998, 2003, 2008 et 2013, lors de la segmentation de la population.

6.1.1 Première segmentation

La Figure 6-1 présente l'évolution des proportions d'individus dans chacun des groupes pour les quatre enquêtes. Ces proportions sont calculées avec les facteurs de pondération provenant des données de l'enquête, tout comme l'ensemble des analyses dans cette section.

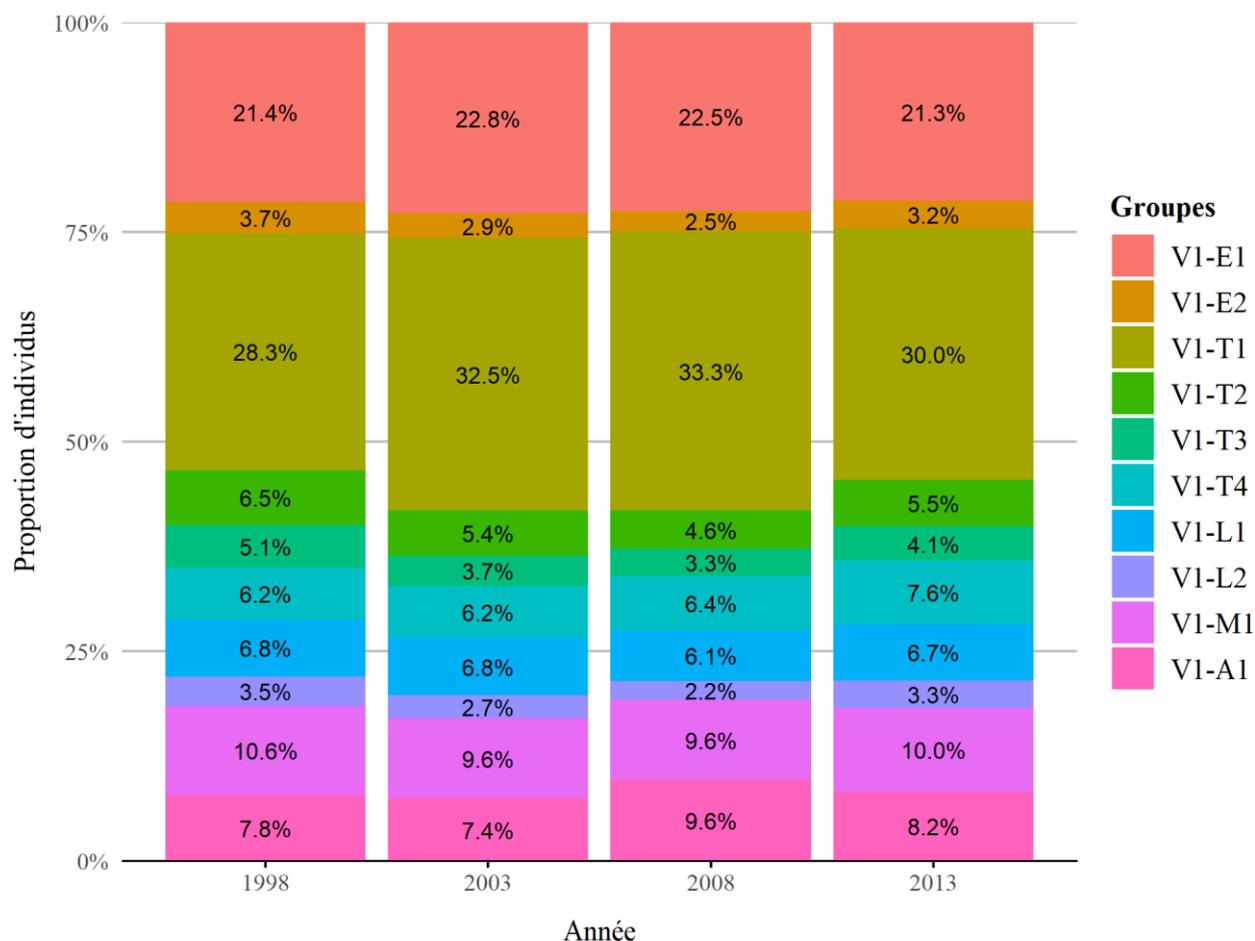


Figure 6-1 Évolution des proportions dans chacun des groupes avec la segmentation 1

L'évolution observée ci-haut peut être la combinaison de plus plusieurs tendances. La variation de la proportion d'un certain segment de la population peut également s'accompagner d'un changement de comportement en matière de système d'activités pour ce segment. Par exemple, le vieillissement de la population peut engendrer une diminution des patrons typiques réalisés par des étudiants, mais cette diminution peut également découler d'un changement de comportement. Cette combinaison des effets peut être détectée en observant des transferts entre les groupes au fil des années.

Par exemple, la proportion de personnes dans V1-E1 et V1-E2 diminue de 2003 à 2013, passant de 25,7 à 24,6 %, après avoir augmenté de 1998 à 2003. Les proportions dans chacun des deux groupes connaissent également la même évolution, à l'exception des membres de V1-E2, dont la proportion augmente de 2008 à 2013. Ainsi, alors que la proportion totale d'étudiants est en baisse, ceux dont

le système d'activités intègre une activité de type *loisir* sont en hausse à partir de 2008, ce qui peut laisser présager un changement de comportement chez les étudiants.

La proportion de personnes dans l'ensemble des groupes où le *travail* est l'activité dominante est relativement stable, avec une augmentation de 1998 à 2003 et une diminution de 0,5 % dans les années subséquentes. Toutefois, les proportions de chacun des groupes connaissent d'importantes variations. Alors que la proportion de personnes dans V1-T1 augmente significativement de 1998 à 2008, contrairement aux autres dont les proportions diminuent ou augmentent très peu, c'est l'inverse qui est observé quant à l'évolution de 2008 à 2013. En effet, la proportion de personnes dont la seule activité réalisée est le *travail* diminue alors que les trois patrons caractérisés par la réalisation d'activités secondaires gagnent en popularité.

Les quatre derniers groupes, dans l'ensemble, voient leurs proportions augmenter de 2003 à 2013, mais les proportions observées pour chacun d'entre eux ne permettent pas de dégager des tendances qui mèneraient à des constats sur les changements de comportements. La pertinence de la typologie repose sur sa capacité à capter des patrons dont l'évolution permet d'observer les changements de comportements dans la population, ce qui n'est pas le cas de ces quatre derniers patrons. En revanche, une analyse approfondie des tendances pourrait éclaircir les constats découlant des proportions observées pour chacun des groupes. Cette analyse approfondie consiste à disséquer les proportions par segment de population, pour ainsi déterminer si les tendances observées à l'intérieur de chacun des groupes sont les mêmes.

Les proportions par segment de population pour V1-E1 et V1-E2 sont présentées dans la Figure 6-2. Les proportions pour l'ensemble de la population, qui sont les mêmes que ceux présentés à la Figure 6-1, sont illustrées par la courbe noire.

Dans le cas de V1-E1, les proportions concernant les personnes de 5 à 19 ans suivent la même évolution que l'ensemble du groupe, contrairement aux personnes de 20 à 29 ans. En effet, parmi l'ensemble des personnes appartenant à cette catégorie d'âge, 17,6 % d'entre eux font partie du groupe V1-E1 en 1998. Ce chiffre monte continuellement jusqu'à 25 % en 2013. Dans le cas de V1-E2, cette proportion semble subir les mêmes variations que l'ensemble du groupe. Ainsi, la forte augmentation pour V1-E1 n'est pas le résultat d'un changement de comportement en tant qu'étudiant, mais plutôt un plus grand nombre de personnes de 20 à 29 ayant comme principale occupation les *études*.

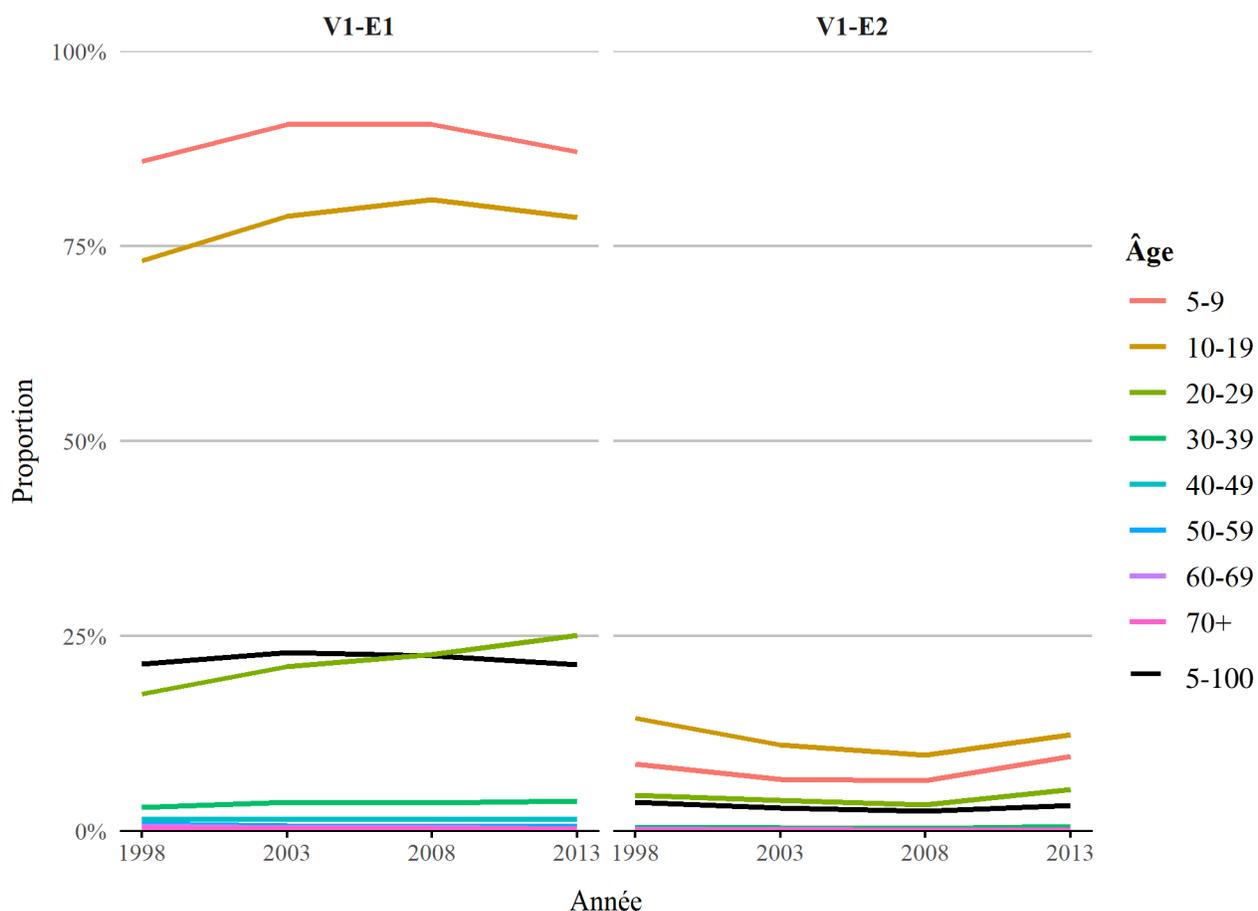


Figure 6-2 Évolution de la proportion de personnes dans V1-E1 et V1-E2 pour chaque segment de la population

La Figure 6-3 présente également les proportions par segment de population, mais pour les membres des groupes V1-T1 à V1-T4. Plusieurs segments de population se distinguent dans le cas des quatre groupes où le *travail* est la principale activité. Tout d'abord, la proportion de personnes de 50 à 69 ans dans V1-T1 varie différemment des autres segments de la population, ce qui indique qu'ils sont de plus en plus nombreux à travailler. De plus, la proportion de personnes de 60 à 69 ans semble stable dans le cas de V1-T2 et V1-T3, contrairement aux autres segments dont l'évolution est sensiblement la même. Finalement, la proportion de personnes dans le groupe V1-T4 varie peu, à l'exception des personnes entre 30 et 49 ans, qui sont de plus en plus nombreuses à réaliser ce patron, particulièrement entre 2008 et 2013.

La Figure 6-4 présente la variation de la proportion de personnes dans les quatre derniers groupes et ce, dans chaque segment de la population.

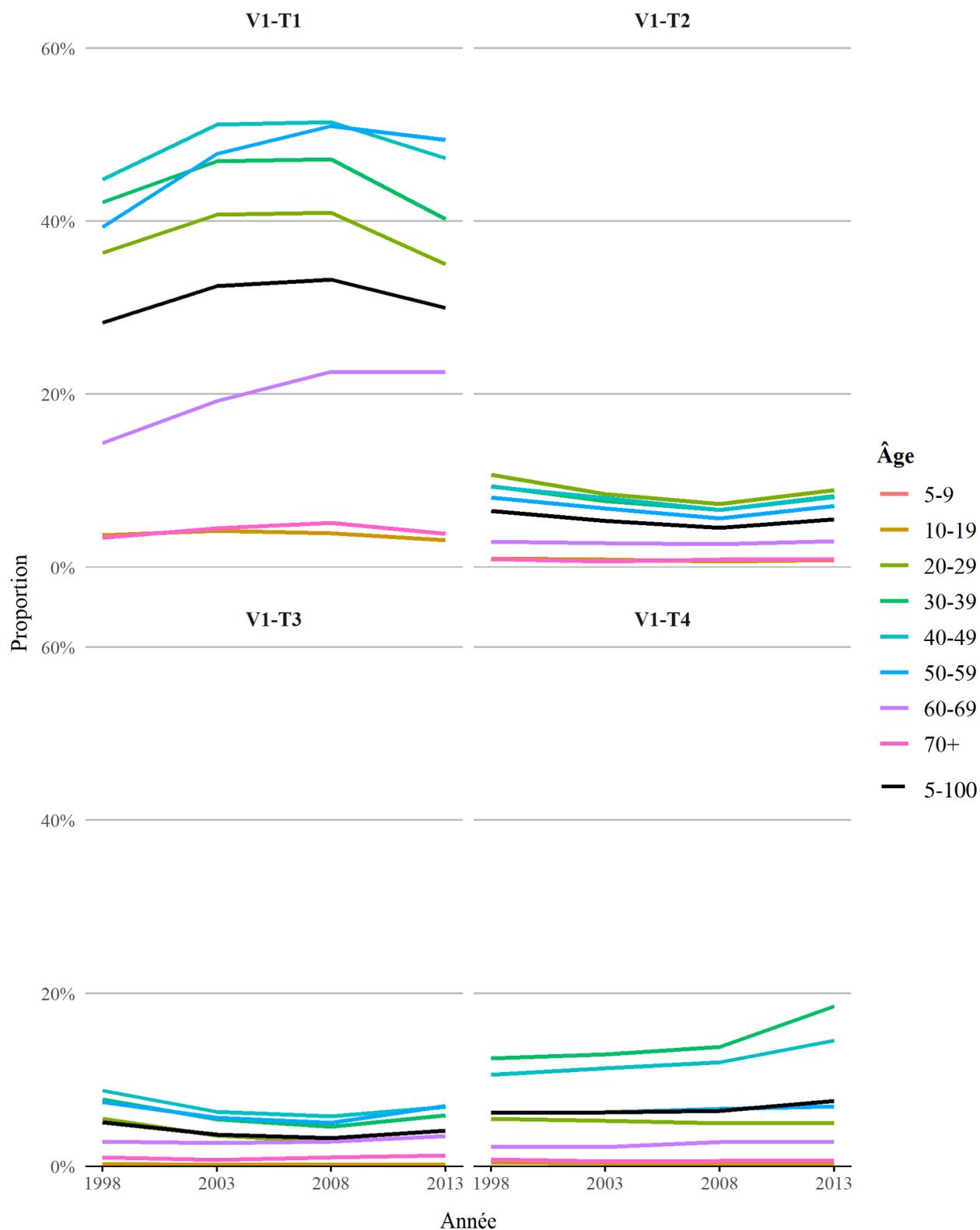


Figure 6-3 Évolution de la proportion de personnes dans V1-T1, V1-T2, V1-T3 et V1-T4

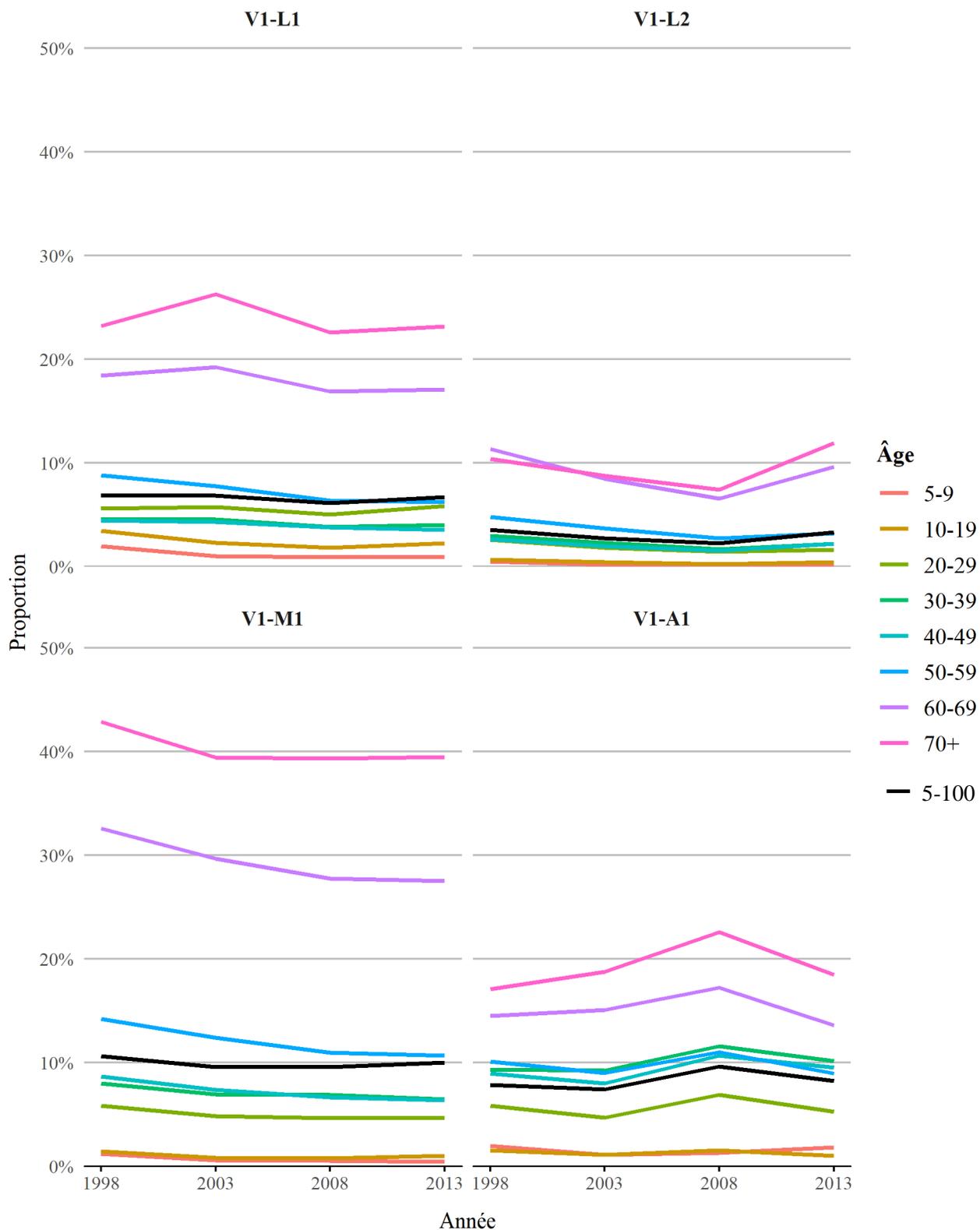


Figure 6-4 Évolution de la proportion de personnes dans V1-L1, V1-L2, V1-M1 et V1-A1

Alors que les variations observées pour l'ensemble de la population sont relativement faibles, la décomposition des proportions permet de constater que les variations sont plus élevées pour les personnes de 60 ans et plus dans le cas des quatre derniers groupes. En effet, de 1998 à 2008, les personnes âgées sont généralement moins nombreuses à réaliser les patrons V1-L1, V1-L2 et V1-M1, contrairement à V1-A1. En revanche, l'inverse est observé entre 2008 et 2013, où les proportions de personnes âgées sont stables pour V1-L1 et V1-M1, mais en augmentation pour V1-L2 et en diminution pour V1-A1.

6.1.2 Deuxième segmentation

La Figure 6-5 présente la proportion de personnes dans chacun des groupes issus de la segmentation avec le deuxième vecteur (V2). Les proportions sont présentées pour chacune des quatre enquêtes.

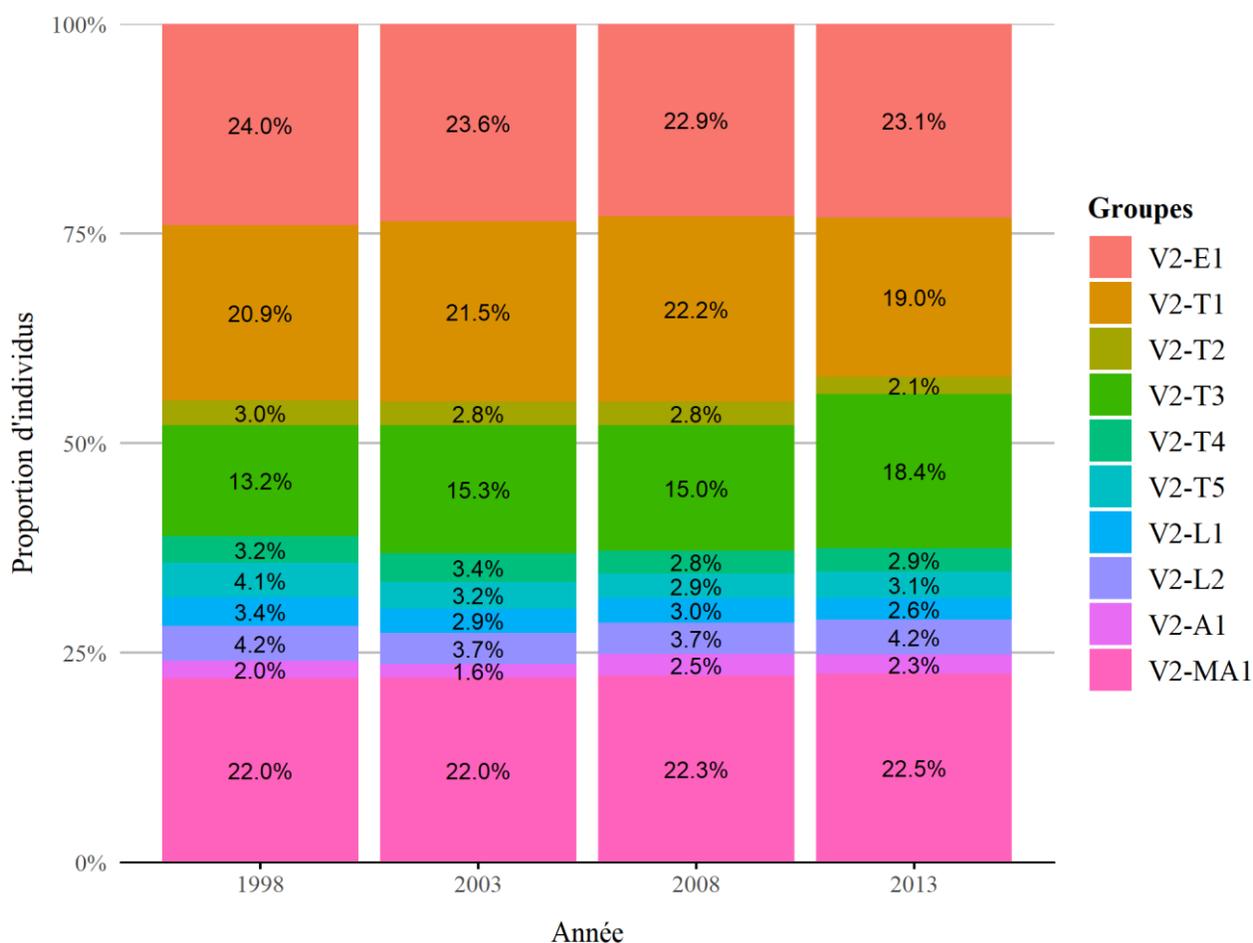


Figure 6-5 Évolution temporelle des proportions d'individus dans chacun des groupes

Les proportions observées dans le groupe où l'*étude* est la principale activité, de même que dans les quatre derniers groupes, ne semblent pas connaître de variation significative. En effet, la proportion d'étudiants ne varie pas au-delà de 1 % entre chacune des années, tout comme la proportion de membres dans V2-L1, V2-L1, V2-A1 et V2-MA1.

La proportion de personnes dans les quatre groupes où le *travail* est l'activité dominante est relativement stable, alors que cette proportion est de 44,4, 46,1, 45,6 et 45,4 % pour chacune des quatre années d'enquête, soit 1998, 2003, 2008 et 2013 respectivement. Alors que cette proportion augmente de 1998 à 2003, elle diminue légèrement par la suite. À l'intérieur de cette proportion, l'évolution la plus importante concerne le groupe V2-T3. Malgré la stabilité de la proportion totale de travailleurs, et même une légère baisse entre 2003 et 2013, la proportion de personnes dans ce groupe augmente de façon significative. Cette hausse s'accompagne également d'une baisse concernant la proportion de personnes dans le groupe V2-T1. À titre de rappel, la principale différence entre ces deux groupes est la durée moyenne de l'activité *travail*, ainsi que la distribution de l'heure de début de l'activité dominante, dont la pointe est plus tôt dans le cas de V2-T3 par rapport à V2-T1. L'évolution des proportions permet donc de constater que le nombre de travailleurs à quitter leur domicile plus tôt dans la matinée est en augmentation.

6.1.3 Troisième segmentation

L'évolution temporelle des proportions de personnes dans les groupes de la dernière segmentation est présentée dans la Figure 6-6.

Tout comme les proportions observées pour les groupes issus de la deuxième segmentation, le nombre d'étudiants demeure relativement stable au fil des années, avec une baisse de 1,3 % de 2003 à 2013. Les proportions dans les groupes V3-L1 et V3-A1 sont également en baisse, passant de 2,7 à 2 % et de 2,3 à 1,4 % respectivement. Dans le cas des deux derniers groupes, de fortes variations sont observées en ce qui concerne leurs proportions, particulièrement dans le cas de V3-MA1, sans qu'une tendance ne soit identifiable.

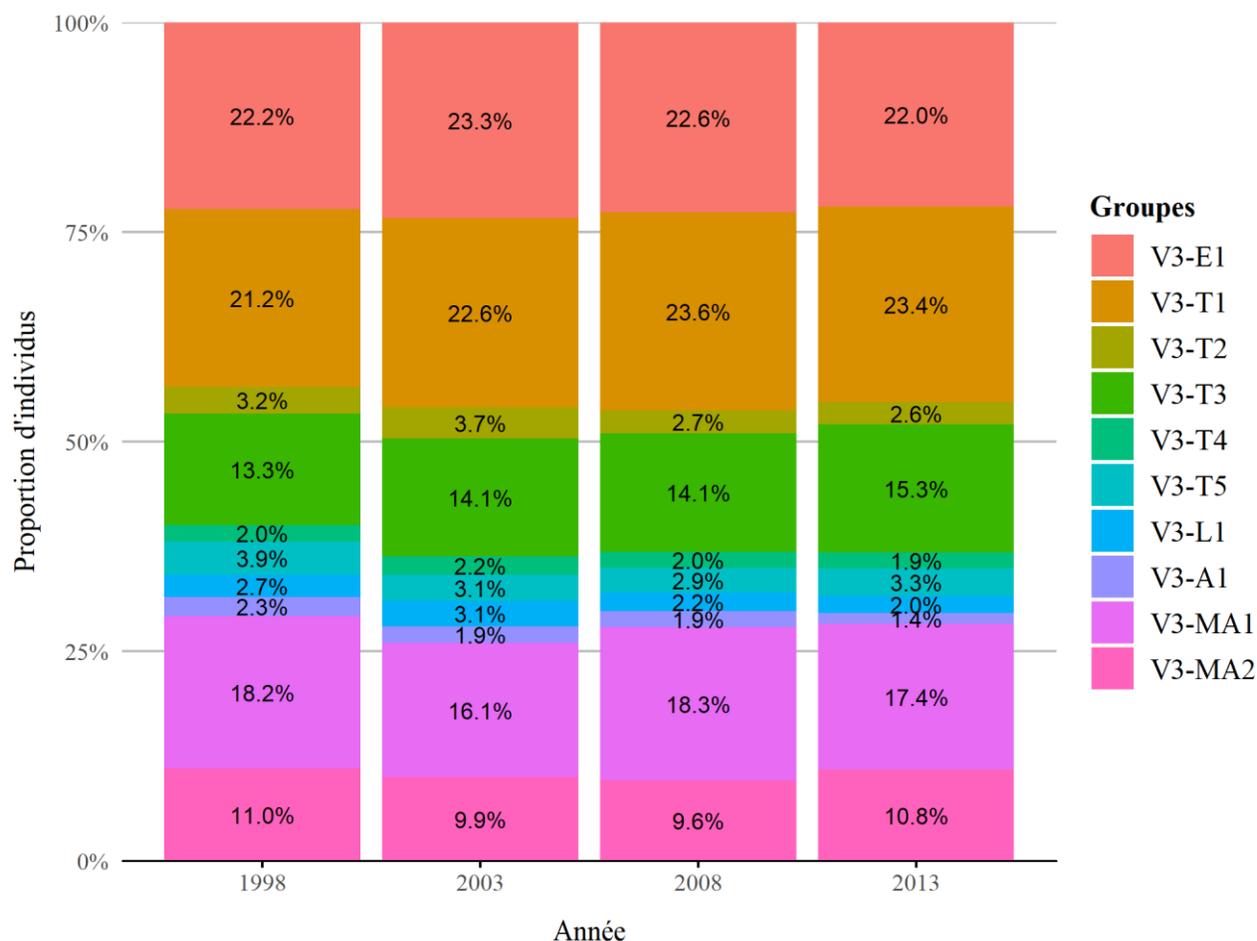


Figure 6-6 Évolution temporelle des proportions d'individus dans chacun des groupes

Les groupes où le *travail* est la principale activité connaissent globalement une hausse quant à leur poids au sein de la population, variant de 43,6 % à 46,5 %. Alors que de 1998 à 2003, l'ensemble des groupes est en hausse, à l'exception de V3-T5, les proportions observées pour chacun des groupes dans les années subséquentes ne suivent pas la même évolution. En effet, V3-T1, V3-T3 et V3-T5 sont en hausse de 2003 à 2013 contrairement à V3-T2 et V3-T4. La plus grande variation concerne le groupe V3-T3, alors que sa proportion varie de 14,1 à 15,3 %. Le groupe V2-T3 de la deuxième segmentation connaît également une hausse importante, ce qui confirme que les individus quittant leur domicile plus tôt pour aller travailler sont en augmentation. Toutefois, cette augmentation n'est pas la même pour chaque segment de la population, comme en témoigne la Figure 6-7, qui indique la proportion de personnes dans le groupe V3-T3, représentée par la courbe noire et également les proportions pour chaque groupe d'âge.

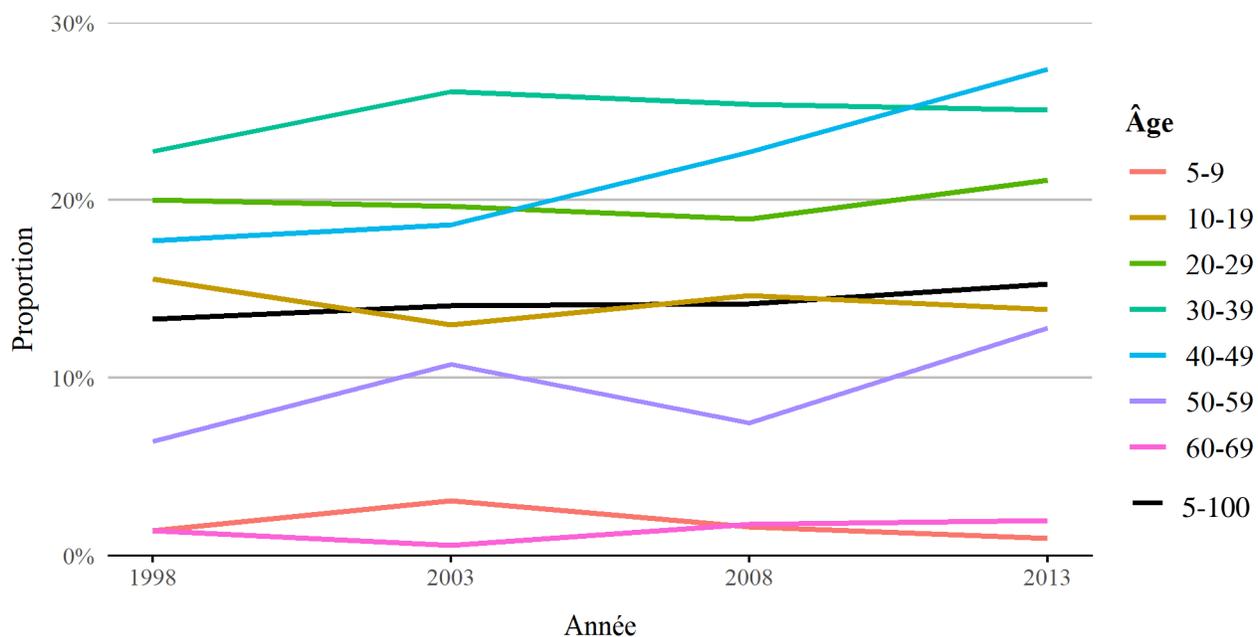


Figure 6-7 Proportion de personnes dans chacun des groupes d'âge parmi les membres du groupe V3-T3

Il est possible de constater que l'évolution n'est pas la même pour chaque segment de la population. En effet, le poids des membres composant V3-T3 parmi les individus de 40 à 49 ans est en forte augmentation, alors que ce poids pour les autres groupes d'âge connaît des variations moins importantes.

6.2 Modélisation de l'évolution de l'appartenance aux groupes

Afin de modéliser l'appartenance aux groupes, les modèles développés lors du chapitre précédent sont modifiés. Cette modification consiste à ajouter une variable numérique à un des modèles qui indique le nombre d'années depuis 1998. Puisque les données des quatre enquêtes sont fusionnées lors de la segmentation de la population et l'estimation des modèles, une variable additionnelle est ajoutée aux données, soit l'année de l'enquête auquel chaque individu appartient. Selon la valeur de cette variable, il est donc possible de calculer le nombre d'années qui s'est écoulé depuis la première enquête, soit celle de 1998.

Puisque quatre enquêtes sont utilisées, seules les valeurs de 0, 5, 10 et 15 sont présentes dans les données. Toutefois, puisque la variable de l'année est numérique, elle peut prendre d'autres valeurs. C'est de cette façon que le modèle peut être utilisé dans un modèle de prévision. Ainsi,

avec une certaine population projetée dans le futur, par exemple à l'horizon 2028, les individus composant cette population auraient une variable « année » dont la valeur serait de 30. Il serait donc possible d'appliquer l'ensemble de la méthode visant à générer des activités.

L'analyse de la variable « année » est réalisée pour chacun des modèles estimés avec les trois différentes segmentations. Parmi les cinq modèles développés au chapitre précédent, c'est au modèle 2 que la variable « année » est ajoutée. Cette variable s'ajoute donc au genre, à l'âge et au type de ménage, tous trois des variables catégorielles à la suite de la transformation de la variable de l'âge. L'analyse de corrélation entre ces trois variables et celle du nombre d'années depuis 1998 démontre que cette dernière n'est pas fortement corrélée avec les trois variables originales.

L'estimation du modèle 2, incluant la variable de l'année, permet d'obtenir le coefficient de chacune des variables, ainsi que l'erreur qui s'y rattache. Ce coefficient permet par la suite d'obtenir le rapport de probabilité. L'ensemble de ces résultats est présenté dans le Tableau 6-1. Les coefficients sont significatifs à $p < 0,01$, à l'exception du coefficient correspondant au groupe V1-T3. Les rapports de probabilités des autres variables sont à l'Annexe E, dans les Tableau E-39 et Tableau E-40.

Tableau 6-1 Résultats du modèle 2 – variable « année »

	Rapport de probabilités	Coefficient	Erreur standard	Significativité
V1-E1	1.037	0.036	0.001	0.00E+00 ***
V1-E2	1.021	0.020	0.002	0.00E+00 ***
V1-T1	1.010	0.010	0.001	0.00E+00 ***
V1-T2	1.007	0.007	0.001	9.68E-08 ***
V1-T3	0.998	-0.002	0.002	1.49E-01
V1-T4	1.028	0.028	0.001	0.00E+00 ***
V1-L1	0.992	-0.008	0.001	1.70E-10 ***
V1-L2	0.941	-0.061	0.002	0.00E+00 ***
V1-M1	0.987	-0.013	0.001	0.00E+00 ***
V1-A1	1.006	0.006	0.001	3.24E-08 ***
V1-I	-	-	-	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Les chances que l'individu de référence adopte les patrons V1-E1 et V1-E2 sont 3,7 et 2,1 % plus élevées par rapport à l'adoption de V1-I avec l'augmentation d'une année. Ainsi, les chances d'adopter ces patrons augmentent avec le temps. À l'exception de V1-T3, les chances d'adopter les patrons où le *travail* est la principale activité par rapport à V1-I augmentent également avec le

temps, au rythme de 1 % pour V1-T1, 0,7 % pour V1-T2 et 2,8 % pour V1-T4. Les chances de magasiner après le *travail*, soit le patron V1-T3, par rapport à rester à la maison ne peuvent être interprétées, puisque le coefficient n'est pas significatif. À l'inverse de V1-A1, dont les chances d'être adopté augmentent de 0,6 % année après année, les patrons dominés par les *loisirs* et le *magasinage* ont moins de chances d'être adoptés par rapport à V1-I. Cette situation pourrait être attribuable au fait que les personnes âgées ont tendance à rester à la maison plutôt que de réaliser des activités.

Grâce aux coefficients, il est possible de calculer la probabilité d'adopter un patron selon différentes valeurs des variables explicatives. De plus, puisque la variable de l'année est numérique, il est possible d'observer l'évolution de la probabilité au fil du temps, tel qu'illustrée sur la Figure 6-8, dont les probabilités sont évaluées de 1998 à 2028, pour les hommes de 30 à 34 ans concernant les patrons dominés par le *travail*.

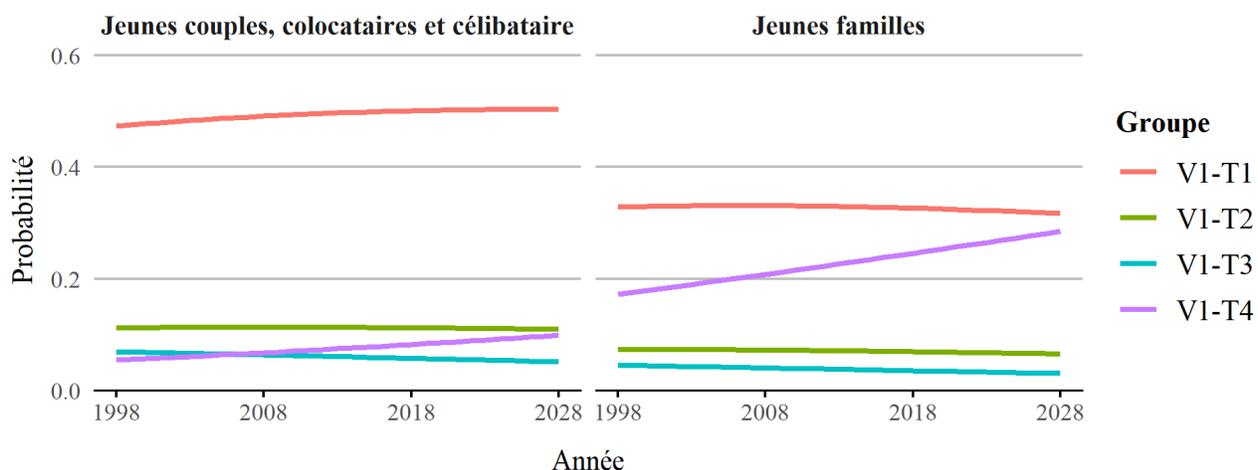


Figure 6-8 Évolution de la probabilité d'appartenir aux groupes dont le *travail* est l'activité dominante pour les hommes âgés entre 30 et 34 ans

Les hommes composant les « Jeunes couples, colocataires et célibataires » ont une probabilité de plus en plus élevée d'adopter V1-T1. De plus, la probabilité d'adopter V1-T4 augmente également pour l'ensemble des hommes de 30 à 34 ans et ce, peu importe le type de ménage. Cependant, cette augmentation est beaucoup plus importante pour les membres des « Jeunes familles ».

CHAPITRE 7 CONCLUSION

La modélisation et la prévision de la demande en transport sont des tâches fondamentales pour la planification des réseaux de transport urbains ainsi que plus largement, pour le développement cohérent des espaces urbanisés. Alors que la majorité des modèles qui sont apparus dans les années 1950 ont comme principale assise le déplacement, une autre génération de modèle a émergé au cours des dernières décennies afin de se baser sur ce qui motive les personnes à se déplacer, soit les activités.

Ce mémoire s'inscrit dans un projet de recherche plus large qui vise à améliorer le processus de prévision de la demande en transport au MTQ. Le but de ce projet de recherche, réalisé en collaboration avec des professionnels du MTQ, est donc d'intégrer les pratiques les plus récentes en matière de modèles d'activités dans le processus de prévision de la demande de transport au Québec.

Une étape essentielle des modèles basés sur les activités est la génération d'une série d'activités pour chaque individu présent dans une certaine population, généralement projetée dans le futur. L'objectif ultime de ce mémoire est donc de générer une séquence d'activités ayant chacune un type, une heure de début et une durée. À travers le développement de cette méthode, il devient donc possible de prendre en compte le système d'activités des individus.

7.1 Synthèse

La revue de la littérature a permis de comprendre les enjeux concernant l'amélioration des modèles de prévision de la demande en transport en étudiant leurs évolutions depuis leur toute première apparition. Le fonctionnement de la PSC a donc été décortiqué, tout comme le modèle de prévision de la demande de transport au Québec auquel ce mémoire vise à faire des contributions. Les limites de ces derniers ont par la suite été exposées, pour ainsi comprendre ce que les modèles d'activités peuvent leur apporter. En effet, la prise en compte de l'ensemble du système d'activités d'un individu est essentielle afin d'améliorer les modèles présentement utilisés dans la pratique.

L'influence des caractéristiques sociodémographiques sur la structure des systèmes d'activités de même que les transformations de ces systèmes dans le temps ont également été présentées, confirmant que l'analyse et la modélisation de l'évolution des systèmes d'activités sont des tâches

indispensables à l'amélioration du processus de prévision de la demande de transport et des modèles d'activités en général.

L'amélioration du processus de génération des activités est au centre de plusieurs projets de recherche depuis quelques années. Une majorité de ces tentatives impliquent l'utilisation de patrons d'activités typiques, dont la détection a fait l'objet d'un grand nombre d'études. Les caractéristiques des données visant à alimenter ces études varient grandement d'une tentative à l'autre, alors que le nombre d'observations peut aller de quelques centaines à quelques milliers. De plus, la diversité des types d'activités peut inclure plus d'une dizaine de types et la fenêtre temporelle peut être d'une seule journée ou d'une semaine complète. De plus, les méthodes permettant d'identifier ces patrons, prenant souvent la forme d'un vecteur, varient également dans la littérature. En effet, ces vecteurs nécessitent une division temporelle de la journée ou de la semaine d'un individu. Cette division peut varier d'une minute à une heure. En outre, les algorithmes permettant de comparer ces vecteurs sont également multiples.

La principale contribution de ce mémoire concerne également la création de patrons typiques d'activités. Ainsi, parmi l'ensemble des comportements observés dans l'enquête OD, soit la principale source de données de ce projet, le défi est d'identifier les comportements les plus typiques en matière de système d'activités. Contrairement aux différentes études recensées, le nombre d'observations est beaucoup plus élevé, soit plus de 150 000 personnes pour une seule enquête. Cependant, seuls cinq types d'activités sont considérés (en raison des enjeux de comparabilité à travers les enquêtes). Ainsi, contrairement aux autres tentatives, la force de la méthode de détection des patrons présentée dans ce mémoire concerne sa capacité à représenter l'ensemble de la population d'une région métropolitaine. Le nombre minimal de types d'activité permet également de détecter un nombre limité de patrons, permettant d'identifier les principaux types de comportements en matière de systèmes d'activités. Ce nombre limité de patrons permet de concentrer l'effort de détection sur les comportements les plus pertinents dans un contexte de prévision de la demande en transport. Cette recherche des comportements les plus pertinents est également renforcée par la mise en commun d'enquêtes réalisées à différents moments sur une période de 15 ans. Cette fusion de différents ensembles de données constitue également un élément inédit dans l'identification de patrons typiques. Bref, l'échelle spatiale et temporelle de la méthode présentée dans ce mémoire est sans précédent.

Trois vecteurs différents, soit V1, V2 et V3, sont créés pour chaque individu présent dans l'échantillon de l'enquête afin de détecter ces patrons. Des méthodes de segmentation différentes sont appliquées à chacun d'entre eux selon leurs natures. La création de trois vecteurs permet de détecter des comportements typiques complètement différents. Les forces et les faiblesses de chacun d'entre eux sont ainsi révélées. V1 permet principalement de partitionner les individus selon la réalisation d'une combinaison de différents types d'activités et d'ainsi regrouper les individus réalisant une activité dominante suivie d'une activité secondaire par exemple. Quant à V2 et V3, leur nature permet essentiellement de détecter des comportements typiques selon la structure temporelle du système d'activité des individus. Ainsi, les principales différences entre les patrons concernent principalement l'heure de début et la durée des activités présentes dans le système de chaque personne. Bref, l'essai de plusieurs méthodes afin d'identifier ces patrons constitue une autre force de ce mémoire.

Le deuxième composant de ce mémoire concerne l'exploration des liens entre les caractéristiques des personnes et des ménages et la structure du système d'activités des individus. Les principaux constats concernent l'influence de l'âge sur les activités réalisées, alors que les personnes de moins de 24 ans réalisent en grande majorité des activités de type *études*, les personnes de plus de 65 ans sont plus enclines à réaliser des activités de type *magasinage*, *loisir* ou *autres*, et les travailleurs sont essentiellement âgés entre 25 et 64 ans. Le lien le plus fort concerne celui entre le type de ménage « Jeunes familles » et le patron V1-T4, ce qui permet de croire que les membres de ce type de ménage sont plus enclins à réaliser des activités de type *autres* avant et après le *travail*. Il pourrait donc s'agir de parents reconduisant leurs enfants à l'école.

L'analyse des liens entre les types de ménage et les patrons d'activités permet également de conclure que de façon générale, les membres des « Jeunes couples, colocataires et célibataires » et des « Cinquantenaires » ont des comportements similaires malgré la différence d'âge entre les membres des deux types. Cette similarité peut être attribuable à la dynamique présente dans les ménages sans enfants, ce qui confirme que la présence de ces derniers peut avoir une influence considérable sur les comportements de mobilité. En outre, les membres des « Jeunes retraités » sont généralement plus enclins à réaliser des activités de type *loisir* plutôt que de rester à la maison, contrairement aux membres des « Vieux retraités », confirmant la baisse du nombre d'activités réalisé avec l'augmentation de l'âge.

L'introduction de variables d'interaction dans les modèles d'allocation des patrons permet également d'observer une variation importante de l'influence de l'âge et du type de ménage sur l'adoption des patrons selon le genre d'une personne.

Finalement, l'analyse et la modélisation de l'évolution temporelle des patrons d'activités permettent de tirer des constats importants concernant la transformation des systèmes d'activités. Tout d'abord, les systèmes d'activités composés d'activités de type *autre* avant et après le *travail*, soit le patron V1-T4, sont en augmentation, ce qui pourrait indiquer que les parents sont de plus en plus nombreux à reconduire leurs enfants à l'école. Un autre élément marquant de cette analyse concerne les patrons V2-T3 et V3-T3, confirmant que les personnes quittent leur domicile de plus en plus hâtivement, pour possiblement éviter la congestion présente sur les réseaux de transport lors de la pointe du matin.

7.2 Limites

Quelques limites accompagnent le développement de la méthode de génération des activités présentée dans ce mémoire. Tout d'abord, l'algorithme de segmentation utilisé avec les vecteurs V2 et V3 ne peut être appliqué à l'ensemble des individus interrogés lors des quatre enquêtes OD. Bien que l'algorithme ait été appliqué à plusieurs sous-échantillons et que les patrons obtenus soient similaires, le fait de ne pas inclure l'ensemble des observations lors de l'application de l'algorithme met à mal la force de la méthode présentée dans ce mémoire, qui vise à détecter des patrons représentatifs de l'ensemble de la population sur plusieurs années. Les autres études utilisant des vecteurs similaires à V2 et V3 avaient pour but d'identifier des patrons typiques à partir de données contenant moins de 10 000 personnes. Dans le cas de ces études, des algorithmes de segmentation hiérarchique sont donc appropriés, contrairement à la méthode présentée dans ce mémoire.

Une seconde limite importante concerne la modélisation de l'évolution temporelle des patrons d'activités. Cette modélisation est réalisée avec une régression logistique et la variable permettant de modéliser la transformation des patrons, soit le nombre d'années depuis 1998, est numérique. Ainsi, l'évolution de l'appartenance aux groupes issus des différentes segmentations est considérée comme linéaire. Les probabilités de réaliser un certain patron varient d'année en année à un rythme constant. De plus, la variation captée par le modèle évolue dans une seule direction, soit en

augmentant ou en diminuant. Cette situation n'est pas représentative de la réalité, comme peut laisser le croire l'analyse temporelle des proportions de personnes dans chacun des groupes, présentée avant les résultats de l'estimation du modèle avec une composante temporelle.

De plus, la méthode de génération des activités développée dans ce mémoire nécessite non seulement l'application du modèle, mais également la construction de distributions des attributs des activités. Alors que le modèle intègre une composante prévisionnelle avec la variable « année », les distributions, construites avec les données des enquêtes antérieures, n'évoluent pas avec le temps. Bien que l'utilisation d'un modèle avec un volet tendanciel demeure une amélioration par rapport aux modèles d'activités existants qui ne prennent pas en compte les transformations des systèmes d'activités, la nature statique des distributions constitue malgré tout une limite de la méthode de génération. En effet, la revue de littérature a démontré que les distributions des attributs des activités évoluent avec le temps.

Outre la méthodologie, la modélisation des transformations des systèmes d'activités pourrait être améliorée grâce à l'ajout d'un plus grand nombre d'enquêtes. En ajoutant l'enquête de 2018, il serait possible de dégager des constats qui sont davantage pertinents et ainsi développer une méthode de prévision plus robuste.

La principale limite de la méthode de génération développée dans ce mémoire concerne la performance des modèles visant à expliquer l'appartenance aux groupes. Les taux de réussite de ceux-ci s'élèvent à environ 30 %, ce qui signifie que pour la majorité des individus présents dans les jeux de validation, le modèle ne parvient pas à prédire le bon groupe. Toutefois, les modèles arrivent généralement bien à prédire l'activité principale de chaque individu. Autrement dit, le groupe prédit pour les individus dont la principale activité est l'*étude* est également un groupe où l'activité dominante est l'*étude*. Le même phénomène est observé pour les travailleurs. De plus, les personnes âgées, dont la majorité reste à la maison ou réalise des activités de type *magasinage*, *loisir* ou *autres*, se voient également attribuer un groupe par le modèle où la principale activité est parmi les trois mentionnés. Ainsi, avec l'âge, le genre et le type de ménage dans lequel un individu vit, le modèle réussit à deviner si une personne réalise une activité dominante de type *études*, *travail* ou autre (*loisir*, *magasinage* et *autre*), mais ne parvient pas à prédire la réalisation d'activités secondaires, ou les particularités de la structure temporelle de la réalisation de l'activité dominante.

7.3 Perspectives

Afin de pallier la principale limite de ce mémoire, une piste de solution qui pourrait être explorée concerne la construction de semaines typiques en utilisant les patrons typiques d'activité. Alors que de multiples tentatives visant à modéliser le système d'activités des individus sur une semaine complète ont été recensées dans la littérature, il peut être pertinent de reconstruire la semaine typique d'un individu avec les observations récoltées lors d'une seule journée moyenne de semaine. Les tentatives identifiées dans la littérature se basent sur des ensembles de données à petite échelle et n'ont pas comme objectif de tirer des conclusions représentatives de l'ensemble d'une population présent dans une région métropolitaine. Ces tentatives, à défaut d'être à grande échelle, récoltent des informations sur l'ensemble d'une semaine, contrairement à l'enquête Origine-Destination, qui se limite à une seule journée. Il devient alors pertinent de reconstruire des semaines typiques avec des données représentatives de l'ensemble de la population de la GRM.

Les patrons typiques observés dans ce mémoire peuvent donc potentiellement être considérés comme divers fragments composant une semaine complète pour un individu. Autrement dit, la semaine typique d'un étudiant pourrait par exemple être composée de trois patrons V1-E1, soit lundi, mardi et mercredi, et de deux patrons V1-E2 le jeudi et le vendredi. Des semaines similaires pourraient être construites pour les travailleurs ou les retraités, dont les semaines pourraient être composées de jours sans aucune activité et d'autres avec une activité de type *loisir*. Afin d'y parvenir, le but serait donc de construire des patrons typiques hebdomadaires d'activités à partir des patrons typiques journaliers. Cette construction pourrait se faire grâce à la recherche de personnes similaires composant différents ménages qui ont été sondés lors de différents jours de la semaine (Verreault & Morency, 2011).

Un autre élément à explorer serait l'intégration d'une composante spatiale à la recherche de patrons typiques. Alors que seule la structure temporelle des systèmes d'activités est considérée dans ce mémoire, l'ajout d'une notion d'espace géographique nécessaire à la réalisation des activités ou, plus précisément, d'une distance totale entre le domicile et les lieux d'activités pourrait enrichir le processus de génération. Ceci est d'autant plus important qu'un élément essentiel des méthodes de génération des activités est la détermination de l'emplacement du lieu de chaque activité générée. De plus, un des objectifs poursuivis par le MTQ en améliorant son modèle de prévision est de prendre davantage en compte les liens entre l'aménagement du territoire et l'utilisation des réseaux

de transport. Intégrer la localisation des activités dans la détection des patrons typiques pourrait ainsi s'avérer utile non seulement pour prédire la destination du déplacement effectué par les individus, mais également pour comprendre et quantifier les liens qui existent entre le développement des villes et des réseaux de transport.

Finalement, dans le cadre de ce mémoire, la définition d'une activité implique qu'une personne doit quitter son domicile et se déplacer vers le lieu de son activité. Cependant, une activité peut également se dérouler à la maison. Plusieurs études ont également tenté d'intégrer ce genre d'activités dans le processus de modélisation des systèmes d'activités. Ainsi, en les intégrant, les activités virtuelles comme le télétravail ou la livraison de biens au domicile, peuvent être considérées dans le modèle de prévision. Les activités qui auparavant nécessitaient un déplacement et qui se transforment en activité virtuelle peuvent donc s'intégrer au modèle si les tendances indiquent qu'il y a un plus grand nombre de personnes qui pratiquent le télétravail par exemple.

RÉFÉRENCES

- Allahviranloo, M., & Priol, E. L. (2018, 4-7 Nov. 2018). *Mobility pattern recognition method: Segmentation and geographic projection*. Paper presented at the 2018 21st International Conference on Intelligent Transportation Systems (ITSC).
- Allard, M. D., & Janes, M. (2008). Time use of working parents: a visual essay. *Monthly Lab. Rev.*, *131*, 3.
- Alsnih, R., & Hensher, D. A. (2003). The mobility and accessibility expectations of seniors in an aging population. *Transportation Research Part A: Policy and Practice*, *37*(10), 903-916.
- Arentze, T. A., Ettema, D., & Timmermans, H. J. P. (2011). Estimating a model of dynamic activity generation based on one-day observations: Method and results. *Transportation Research Part B: Methodological*, *45*(2), 447-460. doi:[10.1016/j.trb.2010.07.005](https://doi.org/10.1016/j.trb.2010.07.005)
- Arentze, T. A., & Timmermans, H. J. (2004). A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological*, *38*(7), 613-633.
- Bernardo, C., Paleti, R., Hoklas, M., & Bhat, C. (2015). An empirical investigation into the time-use and activity patterns of dual-earner couples with and without young children. *Transportation Research Part A: Policy and Practice*, *76*, 71-91. doi:<https://doi.org/10.1016/j.tra.2014.12.006>
- Bhat, C. R., & Koppelman, F. S. (1993). A conceptual framework of individual activity program generation. *Transportation Research Part A: Policy and Practice*, *27*(6), 433-446. doi:[https://doi.org/10.1016/0965-8564\(93\)90050-U](https://doi.org/10.1016/0965-8564(93)90050-U)
- Bhat, C. R., & Koppelman, F. S. (1993). A conceptual framework of individual activity program generation. *Transportation Research, Part A (Policy and Practice)*, *27A*(6), 433-446.
- Bonnel, P. (2004). *Prévoir la demande en transport*: Presses de l'école nationale des ponts et chaussées.
- Bowman, J. L., & Ben-Akiva, M. E. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice*, *35*(1), 1-28. doi:[https://doi.org/10.1016/S0965-8564\(99\)00043-9](https://doi.org/10.1016/S0965-8564(99)00043-9)
- Brownson, R. C., & Boehmer, T. K. (2004). Patterns and trends in physical activity, occupation, transportation, land use and sedentary behaviors. *Does the built environment influence physical activity? Examining the evidence—Special Report 282, Washington, DC. ed.*
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research*, *33*(2), 261-304.
- Chapin, F. S. (1974). *Human Activity Patterns in the City: Things People Do in Time and Space*. In: New-York, Wiley-Interscience.
- Chen, J., Shaw, S.-L., Yu, H., Lu, F., Chai, Y., & Jia, Q. (2011). Exploratory data analysis of activity diary data: a space-time GIS approach. *Journal of Transport Geography*, *19*(3), 394-404. doi:<https://doi.org/10.1016/j.jtrangeo.2010.11.002>

- Desgagnés, P., & Bédard, S. (2005). *Aménagement et transport : Évaluation de scénarios d'aménagement de la Ville de Québec avec les outils prévisionnels du MTQ*. Récupéré de <http://www.bv.transports.gouv.qc.ca/mono/1142451.pdf>
- Desgraupes, B. (2014). Notes de cours Méthodes statistiques. Récupéré de http://bdesgraupes.pagesperso-orange.fr/UPX/L2/MethStats_seance_11_doc.pdf
- Ettema, D., Borgers, A., & Timmermans, H. (1993). Simulation model of activity scheduling behavior. *Transportation Research Record*, 1-1.
- Fabre, L. (2018). *Développement d'une typologie de ménages pour enrichir la prévision de la demande de transport*. (Masters thesis). École Polytechnique de Montréal, Récupéré de <https://publications.polymtl.ca/3192/>
- Fortin, P., Godbout, L., & St-Cerny, S. (2012). Impact of Quebec's universal low fee childcare program on female labour force participation, domestic income, and government budgets. In.
- Hafezi, M. H., Liu, L., & Millward, H. (2017). Identification of Representative Patterns of Time Use Activity Through Fuzzy C-Means Clustering. *Transportation Research Record*, 2668(1), 38-50. doi:10.3141/2668-05
- Hafezi, M. H., Liu, L., & Millward, H. (2019). A time-use activity-pattern recognition model for activity-based travel demand modeling. *Transportation*, 46(4), 1369-1394. doi:10.1007/s11116-017-9840-9
- Hägerstrand, T. (1970). WHAT ABOUT PEOPLE IN REGIONAL SCIENCE? *Papers in Regional Science*, 24(1), 7-24. doi:10.1111/j.1435-5597.1970.tb01464.x
- Hannes, E., Liu, F., Vanhulsel, M., Janssens, D., Bellemans, T., Vanhoof, K., & Wets, G. (2012). Tracking household routines using scheduling hypothesis embedded in skeletons. *Transportmetrica*, 8(3), 225-241.
- Hirsh, M., Prashke, J. N., & Ben-Akiva, M. (1986). Dynamic Model of Weekly Activity Pattern. *Transportation Science*, 20(1), 24-36. doi:10.1287/trsc.20.1.24
- Hurlin, C. (2003). Économétrie des Variables Qualitatives - Chapitre 2 - Modèles Multinomiaux - Modèles Logit Multinomiaux Ordonnés et non Ordonnés. Récupéré de https://www.univ-orleans.fr/deg/masters/ESA/CH/Qualitatif_Chapitre2.pdf
- Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3), 264-323.
- Jiang, S., Ferreira, J., & González, M. C. (2012). Clustering daily patterns of human activities in the city. *Data Mining and Knowledge Discovery*, 25(3), 478-510. doi:10.1007/s10618-012-0264-z
- Joh, C.-H., Arentze, T., Hofman, F., & Timmermans, H. (2002). Activity pattern similarity: a multidimensional sequence alignment method. *Transportation Research Part B: Methodological*, 36(5), 385-403. doi:[https://doi.org/10.1016/S0191-2615\(01\)00009-1](https://doi.org/10.1016/S0191-2615(01)00009-1)
- Joh, C.-H., Arentze, T., & Timmermans, H. (2001). Pattern recognition in complex activity travel patterns: comparison of Euclidean distance, signal-processing theoretical, and multidimensional sequence alignment methods. *Transportation Research Record*, 1752(1), 16-22.

- Joly, I., Crozet, Y., Bonnel, P., & Raux, C. (2002). *La "Loi" de Zahavi, quelle pertinence pour comprendre la contraction ou la dilatation des espaces-temps de la ville ?* Rapport de recherche. France. Récupéré de <https://halshs.archives-ouvertes.fr/halshs-00088507>
- Jones, P. M., Dix, M. C., Clarke, M. I., & Heggie, I. G. (1983). *Understanding travel behaviour*.
- Kitamura, R., Lula, C. V., & Pas, E. I. *AMOS: An activity-based, flexible and truly behavioral tool for evaluation of TDM measures*. Paper presented at the Proceedings of the 21st Summer Annual Meeting: Transportation Planning Methods, England.
- Le Gauffre. (2016). Simulation stochastique et méthode de Monte-Carlo, cours d'Aide à la Décision 2 (AD2). In: Institut National des Sciences Appliquées de Lyon.
- Lin, H.-Z., Lo, H.-P., & Chen, X.-J. (2009). Lifestyle classifications with and without activity-travel patterns. *Transportation Research Part A: Policy and Practice*, 43(6), 626-638.
- Liu, F., Janssens, D., Cui, J., Wets, G., & Cools, M. (2015). Characterizing activity sequences using profile Hidden Markov Models. *Expert Systems with Applications*, 42(13), 5705-5722.
- Madhulatha, T. S. (2012). An overview on clustering methods. *arXiv preprint arXiv:1205.1117*.
- Masson, S. (1998). Interactions between transport system and land use system. From traditional models heritage to the contribution of integrated land use and transport models. *Les Cahiers scientifiques du transport*(33), 79-108. Récupéré de <https://halshs.archives-ouvertes.fr/halshs-01359553>
- McBride, E. C., Davis, A. W., & Goulias, K. G. (2019). Fragmentation in Daily Schedule of Activities using Activity Sequences. *Transportation Research Record*, 2673(4), 844-854. doi:10.1177/0361198119837501
- McNally Michael, G., & Rindt Craig, R. (2007). The Activity-Based Approach. In A. H. David & J. B. Kenneth (Eds.), *Handbook of Transport Modelling* (Vol. 1, pp. 55-73): Emerald Group Publishing Limited.
- Mohamad, I. B., & Usman, D. (2013). Standardization and its effects on K-means clustering algorithm. *Research Journal of Applied Sciences, Engineering and Technology*, 6(17), 3299-3303.
- Morency, C., & Trépanier, M. (2013). Enrichissement du processus de modélisation de la demande de transport à l'aide des modèles d'activités. In Montréal, QC: Polytechnique Montréal.
- MTQ. (2002). Survol technique du modèle de transport de la région de Montréal. In: Ministère des transports du Québec.
- MTQ. (2007). *Déplacements des personnes dans la grande région de Montréal - Scénario prévisionnel 2026 tendanciel*. Récupéré de <http://numerique.banq.qc.ca/patrimoine/details/52327/46807>
- MTQ. (2017a). *Concepts et méthodologie des enquêtes origine-destination, Méthodologie générale des enquêtes-ménages*. Récupéré de <https://www.transports.gouv.qc.ca/fr/ministere/Planification-transports/enquetes-origine-destination/Documents/concept-methodologie.pdf>
- MTQ. (2017b). *Scénario de prévision de la demande en transport dans la région de Montréal à l'ori on 2036 : prise en compte de nouvelles dimensions tendanciennes*. Récupéré de

https://aqtr.com/system/files/file_manager/4_avril_518b_jarouche_fadi_et_tiomodieune_dort.pdf

- Nation Unies. (2017). *World Population Prospects: The 2017 Revision*. Récupéré de <https://www.un.org/development/desa/publications/world-population-prospects-the-2017-revision.html>
- Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48(3), 443-453. doi:[https://doi.org/10.1016/0022-2836\(70\)90057-4](https://doi.org/10.1016/0022-2836(70)90057-4)
- Newman, P. W. G., & Kenworthy, J. R. (1996). The land use—transport connection: An overview. *Land Use Policy*, 13(1), 1-22. doi:[https://doi.org/10.1016/0264-8377\(95\)00027-5](https://doi.org/10.1016/0264-8377(95)00027-5)
- Pas, E. I. (1984). The Effect of Selected Sociodemographic Characteristics on Daily Travel-Activity Behavior. *Environment and Planning A: Economy and Space*, 16(5), 571-581. doi:10.1068/a160571
- Pendyala, R. M., Kitamura, R., Kikuchi, A., Yamamoto, T., & Fujii, S. (2005). Florida activity mobility simulator: overview and preliminary validation results. *Transportation Research Record*, 1921(1), 123-130.
- Québec, I. d. l. s. d. (2018). Ménages privés selon la taille, Québec, 1951-2016. Récupéré de https://www.stat.gouv.qc.ca/statistiques/population-demographie/familles-menages/tableau_03.htm
- Roorda, M. J. (2005). *Activity-based modelling of household travel*. (Ph.D.). University of Toronto (Canada), ProQuest Dissertations & Theses Global database. (NR07596)
- Saneinejad, S., & Roorda, M. (2009). Application of sequence alignment methods in clustering and analysis of routine weekly activity schedules. *Transportation Letters*, 1(3), 197-211. doi:10.3328/TL.2009.01.03.197-211
- Saunier, N. (2017). Gestion de données en transport CIV8760.
- Shen, J., & Cheng, T. (2016). A framework for identifying activity groups from individual space-time profiles. *International journal of geographical information science*, 30(9), 1785-1805.
- Steinley, D. (2006). K-means clustering: a half-century synthesis. *British Journal of Mathematical and Statistical Psychology*, 59(1), 1-34.
- Sweet, M., & Kanaroglou, P. (2016). Gender differences: The role of travel and time use in subjective well-being. *Transportation Research Part F: Traffic Psychology and Behaviour*, 40, 23-34. doi:<https://doi.org/10.1016/j.trf.2016.03.006>
- Verreault, H. (2018). Cours de Planification des transports CIV8750.
- Verreault, H., & Morency, C. (2011). Transcending the Typical Weekday with Large-Scale Single-Day Survey Samples. *Transportation Research Record*, 2230(1), 38-47. doi:10.3141/2230-05
- W Hosmer, D., & Lemeshow, S. (1989). *Applied Logistic Regression*. New York.
- Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301), 236-244.

- Wilson, C. (1998). Analysis of Travel Behavior Using Sequence Alignment Methods. *Transportation Research Record*, 1645(1), 52-59. doi:10.3141/1645-07
- Wongchavalidkul, N., & Piantanakulchai, M. (2015). The integration of classification tree and Sequence Alignment Method for exploring groups of population based on daily time use data. *Applied Soft Computing*, 34, 106-119. doi:<https://doi.org/10.1016/j.asoc.2015.04.035>
- Yasmin, F. (2016). *Enhancing the Modelling of Travel Demand Using an Activity-based Approach*. (Ph.D.). École Polytechnique de Montréal (Canada), Dissertations & Theses @ Ecole Polytechnique de Montreal; ProQuest Dissertations & Theses Global database. (10609454)
- Zhang, W., & Thill, J.-C. (2017). Detecting and visualizing cohesive activity-travel patterns: A network analysis approach. *Computers, Environment and Urban Systems*, 66, 117-129. doi:<https://doi.org/10.1016/j.compenvurbsys.2017.08.004>

ANNEXE A RÉSULTAT DES TESTS DE KOLMOGOROV-SMIRNOV

Tableau A-1 V1 – Heure de début – E1, E1, T1, T2 et T3

Groupes comparés	<i>Travail</i>	<i>Études</i>	<i>Loisir</i>	<i>Magasinage</i>	<i>Autres</i>
V1-E1 & V1-E2	D	D	N	D	D
V1-E1 & V1-T1	D	D	N	D	D
V1-E1 & V1-T2	D	D	N	D	D
V1-E1 & V1-T3	D	D	I	D	D
V1-E1 & V1-T4	D	D	N	D	D
V1-E1 & V1-L1	D	N	D	D	D
V1-E1 & V1-L2	D	D	D	D	D
V1-E1 & V1-M1	D	D	D	D	D
V1-E1 & V1-A1	D	D	N	D	D
V1-E2 & V1-T1	D	D	N	D	D
V1-E2 & V1-T2	D	D	D	D	D
V1-E2 & V1-T3	D	D	D	D	D
V1-E2 & V1-T4	D	D	D	N	D
V1-E2 & V1-L1	D	N	D	D	D
V1-E2 & V1-L2	D	D	D	D	D
V1-E2 & V1-M1	D	D	D	D	D
V1-E2 & V1-A1	D	D	D	D	D
V1-T1 & V1-T2	D	D	N	D	D
V1-T1 & V1-T3	D	D	N	D	D
V1-T1 & V1-T4	D	D	N	D	D
V1-T1 & V1-L1	D	D	D	N	D
V1-T1 & V1-L2	D	D	D	D	D
V1-T1 & V1-M1	D	D	D	D	D
V1-T1 & V1-A1	D	D	D	D	D
V1-T2 & V1-T3	D	N	D	D	D
V1-T2 & V1-T4	D	N	D	D	D
V1-T2 & V1-L1	D	D	D	N	D
V1-T2 & V1-L2	D	D	D	D	D
V1-T2 & V1-M1	D	D	D	D	D
V1-T2 & V1-A1	D	D	D	D	D
V1-T3 & V1-T4	D	D	N	D	D
V1-T3 & V1-L1	D	N	D	D	D
V1-T3 & V1-L2	D	N	D	D	D
V1-T3 & V1-M1	D	D	D	D	D
V1-T3 & V1-A1	D	N	D	D	D

D	: $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I	: $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N	: $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau A-2 V1 – Heure de début T4, L2, L1 et M1

Groupes comparés	Travail	Études	Loisir	Magasinage	Autres
V1-T4 & V1-L1	D	D	D	D	D
V1-T4 & V1-L2	D	D	D	D	D
V1-T4 & V1-M1	D	D	D	D	D
V1-T4 & V1-A1	D	D	D	D	D
V1-L1 & V1-L2	N	D	D	D	D
V1-L1 & V1-M1	N	D	N	D	D
V1-L1 & V1-A1	D	D	D	D	D
V1-L2 & V1-M1	N	N	N	D	D
V1-L2 & V1-A1	D	N	D	D	D
V1-M1 & V1-A1	D	N	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau A-3 V1 – Durée E1 et E2

Groupes comparés	Travail	Études	Loisir	Magasinage	Autres
V1-E1 & V1-E2	D	D	D	D	D
V1-E1 & V1-T1	D	D	I	D	D
V1-E1 & V1-T2	D	D	D	D	D
V1-E1 & V1-T3	D	D	D	N	D
V1-E1 & V1-T4	D	D	D	D	D
V1-E1 & V1-L1	D	D	D	D	D
V1-E1 & V1-L2	D	D	D	D	D
V1-E1 & V1-M1	D	D	I	D	D
V1-E1 & V1-A1	D	D	D	D	D
V1-E2 & V1-T1	D	D	D	D	D
V1-E2 & V1-T2	D	D	D	D	D
V1-E2 & V1-T3	D	D	D	D	D
V1-E2 & V1-T4	D	D	D	D	D
V1-E2 & V1-L1	D	D	D	D	D
V1-E2 & V1-L2	D	D	D	D	D
V1-E2 & V1-M1	D	D	D	D	D
V1-E2 & V1-A1	D	D	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau A-4 V1 – Durée T1, T2, T3, T4, L1, L2 et M1

Groupes comparés	Travail	Études	Loisir	Magasinage	Autres
V1-T1 & V1-T2	D	D	D	D	D
V1-T1 & V1-T3	D	D	D	D	D
V1-T1 & V1-T4	D	D	D	D	D
V1-T1 & V1-L1	D	D	D	I	D
V1-T1 & V1-L2	D	D	D	D	D
V1-T1 & V1-M1	D	D	I	D	D
V1-T1 & V1-A1	D	D	D	D	D
V1-T2 & V1-T3	D	D	D	D	D
V1-T2 & V1-T4	D	D	D	D	D
V1-T2 & V1-L1	D	D	D	D	D
V1-T2 & V1-L2	D	D	D	D	D
V1-T2 & V1-M1	D	D	D	D	D
V1-T2 & V1-A1	D	D	D	D	D
V1-T3 & V1-T4	D	N	D	D	D
V1-T3 & V1-L1	D	D	D	D	D
V1-T3 & V1-L2	D	D	D	D	D
V1-T3 & V1-M1	D	D	D	D	D
V1-T3 & V1-A1	D	D	D	D	D
V1-T4 & V1-L1	D	D	D	D	D
V1-T4 & V1-L2	D	D	D	D	D
V1-T4 & V1-M1	D	D	D	D	D
V1-T4 & V1-A1	D	D	D	D	D
V1-L1 & V1-L2	D	D	D	D	D
V1-L1 & V1-M1	N	D	D	D	D
V1-L1 & V1-A1	I	D	D	D	D
V1-L2 & V1-M1	D	D	D	D	D
V1-L2 & V1-A1	D	N	D	D	D
V1-M1 & V1-A1	I	D	D	D	D

D	: $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I	: $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N	: $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau A-5 V2 – Heure de début E1, T1, T2, T3 et T4

Groupes comparés	Travail	Études	Loisir	Magasinage	Autres
V2-E1 & V2-T1	D	D	D	D	D
V2-E1 & V2-T2	D	D	D	D	D
V2-E1 & V2-T3	D	D	D	D	D
V2-E1 & V2-T4	D	N	D	D	D
V2-E1 & V2-T5	D	N	D	N	D
V2-E1 & V2-L1	D	D	N	D	D
V2-E1 & V2-L2	D	N	D	D	D
V2-E1 & V2-A1	D	N	N	N	D
V2-E1 & V2-MA1	D	D	D	D	D
V2-T1 & V2-T2	D	D	D	D	D
V2-T1 & V2-T3	D	D	D	D	D
V2-T1 & V2-T4	D	N	D	D	D
V2-T1 & V2-T5	D	N	D	D	D
V2-T1 & V2-L1	D	N	D	D	D
V2-T1 & V2-L2	D	N	D	D	D
V2-T1 & V2-A1	D	I	D	D	D
V2-T1 & V2-MA1	D	D	D	D	D
V2-T2 & V2-T3	D	D	D	D	D
V2-T2 & V2-T4	D	N	N	D	N
V2-T2 & V2-T5	D	N	D	D	D
V2-T2 & V2-L1	D	I	D	D	D
V2-T2 & V2-L2	D	I	D	D	D
V2-T2 & V2-A1	D	N	D	D	D
V2-T2 & V2-MA1	D	N	D	D	N
V2-T3 & V2-T4	D	N	D	D	D
V2-T3 & V2-T5	D	N	D	I	D
V2-T3 & V2-L1	D	N	D	D	D
V2-T3 & V2-L2	D	N	D	D	D
V2-T3 & V2-A1	N	N	D	N	D
V2-T3 & V2-MA1	D	D	D	D	D
V2-T4 & V2-T5	D	I	D	D	D
V2-T4 & V2-L1	N	N	D	D	D
V2-T4 & V2-L2	N	I	D	D	D
V2-T4 & V2-A1	D	I	D	D	D
V2-T4 & V2-MA1	D	N	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau A-6 V2 – Heure de début T5, L1, L2 et A1

Groupes comparés	<i>Travail</i>	<i>Études</i>	<i>Loisir</i>	<i>Magasinage</i>	<i>Autres</i>
V2-T5 & V2-L1	D	I	D	D	D
V2-T5 & V2-L2	D	I	D	D	D
V2-T5 & V2-A1	D	I	D	N	D
V2-T5 & V2-MA1	D	N	D	D	D
V2-L1 & V2-L2	N	I	D	D	N
V2-L1 & V2-A1	D	I	N	D	D
V2-L1 & V2-MA1	D	N	D	N	D
V2-L2 & V2-A1	D	I	D	D	D
V2-L2 & V2-MA1	D	N	D	D	D
V2-A1 & V2-MA1	D	N	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau A-7 V2 – Durée E1 et T1

Groupes comparés	<i>Travail</i>	<i>Études</i>	<i>Loisir</i>	<i>Magasinage</i>	<i>Autres</i>
V2-E1 & V2-T1	D	D	D	N	D
V2-E1 & V2-T2	D	D	D	N	I
V2-E1 & V2-T3	D	D	D	N	N
V2-E1 & V2-T4	D	N	D	N	I
V2-E1 & V2-T5	D	N	D	I	N
V2-E1 & V2-L1	D	D	D	D	D
V2-E1 & V2-L2	D	N	D	I	N
V2-E1 & V2-A1	D	N	I	I	D
V2-E1 & V2-MA1	D	D	D	D	D
V2-T1 & V2-T2	D	N	I	N	N
V2-T1 & V2-T3	D	D	D	N	D
V2-T1 & V2-T4	D	N	N	N	N
V2-T1 & V2-T5	D	N	D	I	I
V2-T1 & V2-L1	D	D	D	D	D
V2-T1 & V2-L2	D	N	D	N	D
V2-T1 & V2-A1	D	N	D	I	D
V2-T1 & V2-MA1	D	D	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau A-8 V2 – Durée T2, T3, T4, T5, L1, L2 et A1

Groupes comparés	<i>Travail</i>	<i>Études</i>	<i>Loisir</i>	<i>Magasinage</i>	<i>Autres</i>
V2-T2 & V2-T3	D	I	N	N	I
V2-T2 & V2-T4	D	I	N	N	N
V2-T2 & V2-T5	D	N	D	N	N
V2-T2 & V2-L1	D	N	D	N	N
V2-T2 & V2-L2	D	N	D	N	N
V2-T2 & V2-A1	D	I	D	N	D
V2-T2 & V2-MA1	D	D	N	D	D
V2-T3 & V2-T4	D	I	D	N	N
V2-T3 & V2-T5	D	N	D	N	D
V2-T3 & V2-L1	D	N	D	D	D
V2-T3 & V2-L2	D	N	D	N	N
V2-T3 & V2-A1	D	N	D	I	D
V2-T3 & V2-MA1	D	D	N	D	D
V2-T4 & V2-T5	D	I	D	N	N
V2-T4 & V2-L1	D	I	D	D	D
V2-T4 & V2-L2	D	I	D	N	N
V2-T4 & V2-A1	D	I	D	N	D
V2-T4 & V2-MA1	D	N	D	D	D
V2-T5 & V2-L1	D	N	D	D	D
V2-T5 & V2-L2	D	I	D	N	D
V2-T5 & V2-A1	D	I	D	I	D
V2-T5 & V2-MA1	D	I	D	D	D
V2-L1 & V2-L2	N	N	D	D	N
V2-L1 & V2-A1	N	N	D	D	D
V2-L1 & V2-MA1	D	D	D	N	D
V2-L2 & V2-A1	D	I	D	I	D
V2-L2 & V2-MA1	D	N	D	D	D
V2-A1 & V2-MA1	D	N	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes

I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques

N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau A-9 V3 – Heure de début E1, T1, T2, T3 et T4

Groupes comparés	Travail	Études	Loisir	Magasinage	Autres
V3-E1 & V3-T1	D	D	N	D	D
V3-E1 & V3-T2	D	D	D	D	D
V3-E1 & V3-T3	D	D	D	D	D
V3-E1 & V3-T4	D	N	D	D	D
V3-E1 & V3-T5	D	D	D	I	D
V3-E1 & V3-L1	D	D	D	D	D
V3-E1 & V3-A1	D	N	D	D	D
V3-E1 & V3-MA1	D	D	D	D	D
V3-E1 & V3-MA2	D	D	D	D	D
V3-T1 & V3-T2	D	D	D	D	D
V3-T1 & V3-T3	D	D	D	D	D
V3-T1 & V3-T4	D	N	D	D	D
V3-T1 & V3-T5	D	N	D	D	D
V3-T1 & V3-L1	D	D	D	D	D
V3-T1 & V3-A1	D	I	N	D	D
V3-T1 & V3-MA1	D	D	D	D	D
V3-T1 & V3-MA2	D	D	D	D	D
V3-T2 & V3-T3	D	N	D	D	D
V3-T2 & V3-T4	D	N	N	D	N
V3-T2 & V3-T5	D	N	D	D	D
V3-T2 & V3-L1	D	N	D	N	D
V3-T2 & V3-A1	D	N	D	D	D
V3-T2 & V3-MA1	D	N	D	D	N
V3-T2 & V3-MA2	D	D	D	D	N
V3-T3 & V3-T4	D	N	D	D	D
V3-T3 & V3-T5	D	N	D	N	D
V3-T3 & V3-L1	D	D	D	D	D
V3-T3 & V3-A1	N	N	D	N	D
V3-T3 & V3-MA1	D	D	D	D	D
V3-T3 & V3-MA2	D	D	D	D	D
V3-T4 & V3-T5	D	N	D	D	D
V3-T4 & V3-L1	N	N	D	D	D
V3-T4 & V3-A1	D	I	D	D	D
V3-T4 & V3-MA1	D	N	D	D	D
V3-T4 & V3-MA2	D	N	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes
I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques
N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

Tableau A-10 V3 – Heure de début T5, L1, A1 et MA1

Groupes comparés	Travail	Études	Loisir	Magasinage	Autres
V3-T5 & V3-L1	D	N	D	D	D
V3-T5 & V3-A1	D	I	D	N	D
V3-T5 & V3-MA1	D	N	D	D	D
V3-T5 & V3-MA2	D	N	D	D	D
V3-L1 & V3-A1	D	N	N	D	D
V3-L1 & V3-MA1	D	I	D	D	D
V3-L1 & V3-MA2	N	D	D	D	D
V3-A1 & V3-MA1	D	N	D	D	D
V3-A1 & V3-MA2	D	N	D	D	D
V3-MA1 & V3-MA2	D	D	D	D	D
D	: $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes				
I	: $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques				
N	: $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant				

Tableau A-11 V3 – Durée – E1 et T1

Groupes comparés	Travail	Études	Loisir	Magasinage	Autres
V3-E1 & V3-T1	D	D	D	N	I
V3-E1 & V3-T2	D	D	D	N	N
V3-E1 & V3-T3	D	D	D	N	D
V3-E1 & V3-T4	D	N	D	D	N
V3-E1 & V3-T5	D	D	D	I	N
V3-E1 & V3-L1	D	D	D	D	D
V3-E1 & V3-A1	D	N	N	N	D
V3-E1 & V3-MA1	D	D	D	N	D
V3-E1 & V3-MA2	D	D	D	D	D
V3-T1 & V3-T2	D	D	I	N	N
V3-T1 & V3-T3	D	D	I	N	D
V3-T1 & V3-T4	D	N	N	N	N
V3-T1 & V3-T5	D	N	D	I	N
V3-T1 & V3-L1	D	N	D	D	D
V3-T1 & V3-A1	D	N	D	N	D
V3-T1 & V3-MA1	D	D	D	D	D
V3-T1 & V3-MA2	D	I	D	D	D
D	: $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes				
I	: $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques				
N	: $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant				

Tableau A-12 V3 – Durée T2, T3, T4, T5, L1, A1 et MA1

Groupes comparés	Travail	Études	Loisir	Magasinage	Autres
V3-T2 & V3-T3	D	I	I	N	N
V3-T2 & V3-T4	D	I	I	D	I
V3-T2 & V3-T5	D	I	D	N	I
V3-T2 & V3-L1	D	N	D	N	D
V3-T2 & V3-A1	D	N	D	N	D
V3-T2 & V3-MA1	D	D	N	N	D
V3-T2 & V3-MA2	D	D	D	D	D
V3-T3 & V3-T4	D	I	N	N	N
V3-T3 & V3-T5	D	I	D	N	N
V3-T3 & V3-L1	D	N	D	D	D
V3-T3 & V3-A1	D	N	D	D	D
V3-T3 & V3-MA1	D	D	D	D	D
V3-T3 & V3-MA2	D	D	D	D	D
V3-T4 & V3-T5	D	I	D	N	N
V3-T4 & V3-L1	D	N	D	D	D
V3-T4 & V3-A1	D	I	D	D	D
V3-T4 & V3-MA1	D	N	D	D	D
V3-T4 & V3-MA2	D	N	D	D	D
V3-T5 & V3-L1	D	N	D	N	D
V3-T5 & V3-A1	D	N	D	N	D
V3-T5 & V3-MA1	D	D	D	I	D
V3-T5 & V3-MA2	D	D	D	D	D
V3-L1 & V3-A1	N	I	D	N	D
V3-L1 & V3-MA1	D	D	D	D	D
V3-L1 & V3-MA2	I	N	D	D	D
V3-A1 & V3-MA1	D	N	N	N	D
V3-A1 & V3-MA2	N	N	D	D	D
V3-MA1 & V3-MA2	D	D	D	D	D

D : $p < 0,1$, l'hypothèse nulle peut être rejetée, les distributions sont différentes

I : $p > 0,9$, l'hypothèse nulle peut être acceptée, les distributions sont identiques

N : $0,1 < p < 0,9$, l'hypothèse nulle ne peut être rejetée ou acceptée, le test est non-concluant

ANNEXE B ANALYSE DE CORRÉLATION DES MODÈLES

Tableau B-13 Matrice de corrélation du modèle 3 – Variable ménage

		Jeunes couples, colocataires et célibataire	Familles traditionnelles	Adultes et personnes âgées	Cinquantennaires	Jeunes retraités	Vieux retraités
5-14		-0.13	-0.12	-0.08	-0.13	-0.14	-0.09
15-19	F	-0.04	0.11	-0.04	-0.07	-0.07	-0.05
	H	-0.05	0.12	-0.04	-0.07	-0.07	-0.05
20-24	F	0.13	0.12	-0.03	-0.06	-0.07	-0.04
	H	0.09	0.16	-0.03	-0.06	-0.07	-0.04
25-29	F	0.21	-0.02	-0.02	-0.06	-0.07	-0.04
	H	0.23	0.02	-0.01	-0.06	-0.06	-0.04
30-34	F	0.13	-0.07	-0.02	-0.07	-0.07	-0.05
	H	0.19	-0.06	0	-0.07	-0.07	-0.05
35-39	F	0.06	-0.08	-0.02	-0.06	-0.08	-0.05
	H	0.12	-0.08	0	-0.06	-0.08	-0.05
40-44	F	-0.02	-0.01	-0.01	0.02	-0.08	-0.05
	H	0.02	-0.05	0	0.01	-0.08	-0.05
45-49	F	-0.07	0.13	-0.01	0.12	-0.08	-0.05
	H	-0.06	0.08	-0.01	0.12	-0.08	-0.05
50-54	F						
	H	-0.07	0.13	-0.01	0.2	-0.08	-0.05
55-59	F	-0.06	0.01	0.02	0.19	0.08	-0.05
	H	-0.06	0.05	0	0.23	-0.01	-0.05
60-64	F	-0.06	-0.06	0.05	-0.05	0.36	-0.04
	H	-0.06	-0.04	0.04	-0.01	0.28	-0.04
65-69	F	-0.06	-0.07	0.05	-0.06	0.36	-0.03
	H	-0.05	-0.07	0.05	-0.05	0.33	-0.04
70-74	F	-0.05	-0.07	0.04	-0.05	0.17	0.2
	H	-0.05	-0.06	0.04	-0.05	0.23	0.07
75-79	F	-0.04	-0.06	0.04	-0.04	-0.04	0.41
	H	-0.04	-0.05	0.03	-0.04	0	0.32
80-84	F	-0.03	-0.05	0.05	-0.04	-0.04	0.34
	H	-0.03	-0.04	0.03	-0.03	-0.03	0.3
85-89	F	-0.02	-0.03	0.05	-0.02	-0.03	0.21
	H	-0.02	-0.03	0.03	-0.02	-0.02	0.18
90-95	F	-0.01	-0.02	0.04	-0.02	-0.02	0.12
	H	-0.01	-0.01	0.02	-0.01	-0.01	0.09
95+	F	-0.01	-0.01	0.02	-0.01	-0.01	0.05
	H	0	-0.01	0.01	0	0	0.03

Tableau B-14 Matrice de corrélation du modèle 3 – Variable d'interaction – 54 ans et moins

15-19	F	-0.07																
	H	-0.07	-0.04															
20-24	F	-0.07	-0.03	-0.03														
	H	-0.07	-0.03	-0.03	-0.03													
25-29	F	-0.07	-0.03	-0.03	-0.03	-0.03												
	H	-0.06	-0.03	-0.03	-0.03	-0.03	-0.03											
30-34	F	-0.07	-0.04	-0.04	-0.03	-0.03	-0.03	-0.03										
	H	-0.07	-0.04	-0.04	-0.03	-0.03	-0.03	-0.03	-0.03									
35-39	F	-0.08	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04								
	H	-0.08	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.03	-0.04	-0.04							
40-44	F	-0.08	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04						
	H	-0.08	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04					
45-49	F	-0.08	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.05				
	H	-0.08	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04			
50-54	H	-0.08	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04		
55-59	F	-0.07	-0.04	-0.04	-0.03	-0.03	-0.03	-0.03	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	
	H	-0.07	-0.04	-0.04	-0.03	-0.03	-0.03	-0.03	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04
60-64	F	-0.07	-0.03	-0.04	-0.03	-0.03	-0.03	-0.03	-0.03	-0.04	-0.03	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04
	H	-0.07	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04
65-69	F	-0.06	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.04	-0.03	-0.04	-0.03	-0.03
	H	-0.06	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
70-74	F	-0.06	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
	H	-0.05	-0.03	-0.03	-0.02	-0.02	-0.02	-0.02	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
75-79	F	-0.05	-0.02															
	H	-0.04	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
80-84	F	-0.04	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
	H	-0.03	-0.02	-0.02	-0.02	-0.02	-0.02	-0.01	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
85-89	F	-0.03	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	H	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
90-95	F	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	H	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
95+	F	-0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	H	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
			F	H	F	H	F	H	F	H	F	H	F	H	F	H	F	H
		5-9																
			15-19		20-24		25-29		30-34		35-39		40-44		45-49		50-54	

ANNEXE C RAPPORTS DE PROBABILITÉ DES MODÈLES 1 À 5 POUR V2 ET V3

Tableau C-18 V2 – Modèle 1 – *Études et travail*

	V2-E1	V2-T1	V2-T2	V2-T3	V2-T4	V2-T5	V2-I
Constante	13048.62***	0.1***	0.08***	0.09***	0.07***	0.03***	-
Âge	0.79***	1	0.97***	1	0.98***	0.99**	-
Jeunes couples, colocataires et célibataire	0.03***	15.83***	6.47***	10.96***	7.7***	20.51***	-
Jeunes familles	0.06***	12.65***	3.66***	12.22***	4.7***	8.79***	-
Familles traditionnelles	0.05***	10.54***	5.18***	9.43***	5.42***	10.97***	-
Adultes et personnes âgées	0.03***	2.15***	2.44**	2.12***	1.12	2.54**	-
Cinquantennaires	0.25*	10.42***	4.88***	9.74***	3.58***	11***	-
Retraités	-	-	-	-	-	-	-
Familles multigénérationnelles	0.02***	4.15***	2.94***	3.13***	2.77***	3.56***	-
Homme	1.06	1.71***	1.93***	1.39***	1.89***	1.84***	-
Femme	-	-	-	-	-	-	-

Tableau C-19 V2 – Modèle 1 – *Loisir, magasinage et autres*

	V2-L1	V2-L2	V2-A1	V2-MA1	V2-I
Constante	0.38**	0.48**	0.06***	1.12	-
Âge	0.99**	0.99**	0.99	1	-
Jeunes couples, colocataires et célibataire	1	0.44***	1.77	1.05	-
Jeunes familles	0.74	0.37***	2.64***	1.19	-
Familles traditionnelles	0.61**	0.37***	1.7**	0.99	-
Adultes et personnes âgées	0.47***	0.5***	0.93	0.67***	-
Cinquantennaires	0.81	0.66**	1.56*	1.09	-
Retraités	-	-	-	-	-
Familles multigénérationnelles	0.33***	0.37***	1.38	0.71**	-
Homme	0.94	1.17	2.45***	1.15**	-
Femme	-	-	-	-	-

Tableau C-20 V2 – Modèle 3 – Études et travail – Variable âge

		V2-E1	V2-T1	V2-T2	V2-T3	V2-T4	V2-T5	V2-I
Constante		0.02**	0.07***	0.03***	0.06***	0.02***	0.02*	-
5-14		723.16***	0.04	0***	0***	0***	0***	-
15-19	F	272.42***	2.38	1.65	2.63	2.61	0	-
	H	267.28***	4.6	5.92	5.05*	4.47	3.97	-
20-24	F	88.21***	9.19**	9.7*	14.24***	8.66*	9.7	-
	H	74.06***	17.65***	10.5**	17.65***	11.71**	15.89	-
25-29	F	25.76**	19.59***	7.53*	19.43***	13.81**	23.01	-
	H	28.88**	32.89***	9.18*	34.3***	22.61***	28.83	-
30-34	F	15.91*	26.82***	4.16	31.91***	7.82*	13.31	-
	H	10.3	54.12***	14.18**	46.57***	26.63***	41.1	-
35-39	F	8.46	24.7***	3.77	27.15***	9.39*	8.42	-
	H	4.5	39.72***	5.68	32.76***	25.21***	25.82	-
40-44	F	8.72	32.97***	6.83	39***	10.15*	18.93	-
	H	5.21	58.36***	12.92**	57.51***	20.3**	33.55	-
45-49	F	2.2	28.85***	3.68	38.2***	7.87*	15.49	-
	H	3.85	48.39***	11.34**	40.26***	13.18**	26.47	-
50-54	F	1.52	24.16***	5.77	25.22***	4.04	11.04	-
	H	2.09	46.83***	9.72*	52.32***	13.28**	20.77	-
55-59	F	1.09	12.66***	1.14	18.35***	5.17	7.16	-
	H	1.02	31.43***	5.66	30.65***	8.31*	8	-
60-64	F	0.85	8.22**	0.96	8.09**	4.11	1.8	-
	H	0.38	11.94**	2.96	16.46***	6.05	6.59	-
65-69	F	0.94	1.29	0.25	2.22	1.31	0.66	-
	H	0.37	3.54	1.15	4.31	1.53	1.99	-
70-74	F	0.38	0.6	0	0.83	1.66	0	-
	H	1.45	2.21	0.39	1.9	0.01	1.18	-
75-79	F	-	-	-	-	-	-	-
	H	0.47	2.16	0.56	2.35	0.02	0.05	-
80-84	F	0.4	0.05	0.03	1.56	0.03	0.08	-
	H	0.55	0.05	0.05	0	1.73	0.15	-
85-89	F	0.3	0.03	0.03	0	0.04	0.14	-
	H	0.41	0.05	0.08	0.01	0.1	0.27	-
90+	F	0.3	0.04	0.08	0.02	0.1	0.29	-
	H	0.28	0.09	0.22	0.11	0.29	0.56	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Tableau C-21 V2 – Modèle 3 – *Études et travail* – Variable type de ménage

	V2-E1	V2-T1	V2-T2	V2-T3	V2-T4	V2-T5	V2-I
Jeunes couples, colocataires et célibataire	0.72**	1.12	1.35	0.8*	1.26	1.76***	-
Jeunes familles	-	-	-	-	-	-	-
Familles traditionnelles	0.94	0.92	1.07	0.79**	1.32	1.29	-
Adultes et personnes âgées	0.57	0.44***	1.03	0.43***	0.44**	0.59	-
Cinquantennaires	0.68	0.73**	0.85	0.66***	0.76	1.07	-
Jeunes retraités	0.42	0.62**	0.92	0.54***	0.7	0.75	-
Vieux retraités	0.23	0.1***	0.33	0.23**	0.59	0***	-
Familles multigénérationnelles	0.62***	0.67**	1.03	0.48***	0.95	0.69	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Tableau C-22 V2 – Modèle 3 – *Loisir, magasinage et autres* – Variable type de ménage

	V2-L1	V2-L2	V2-A1	V2-MA1	V2-I
Jeunes couples, colocataires et célibataire	1.34	1.12	0.69	0.73**	-
Jeunes familles	-	-	-	-	-
Familles traditionnelles	0.61**	0.96	0.59**	0.76**	-
Adultes et personnes âgées	0.56	1.08	0.5*	0.57***	-
Cinquantennaires	0.63*	1.61*	0.42***	0.75**	-
Jeunes retraités	1.04	2.25***	0.63	0.9	-
Vieux retraités	0.97	1.27	0.71	0.82	-
Familles multigénérationnelles	0.38**	0.86	0.73	0.61***	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Tableau C-23 V2 – Modèle 3 – *Loisir, magasinage et autres* – Variable âge

		V2-L1	V2-L2	V2-A1	V2-MA1	V2-I
Constante		0.13***	0.11***	0.05***	0.97	-
5-14		1.01	1.55	2.26	0.3***	-
15-19	F	1.6	0.63	0.77	0.9	-
	H	0.4	1.6	1.8	0.73	-
20-24	F	1.89	1	1.35	1.63*	-
	H	1.5	0.85	5.64**	1.5	-
25-29	F	1.4	2.18	3.06	1.38	-
	H	0.93	2	4.68*	1.57	-
30-34	F	1.24	1.01	0.58	2.07***	-
	H	1.13	1.91	8.72***	2**	-
35-39	F	1.43	0.87	1.64	1.85**	-
	H	0.9	0.77	8.32***	1.2	-
40-44	F	2.1	1.26	2.2	1.71**	-
	H	1.08	0.79	10.39***	1.79**	-
45-49	F	0.84	1.15	4.11*	1.27	-
	H	1.52	0.99	10.27***	1.7*	-
50-54	F	1.79	0.71	3.03	1.47	-
	H	2.93**	1.23	8.2***	2.13***	-
55-59	F	1.64	1.37	1.9	1.31	-
	H	1.96	1.75	8.14***	1.79**	-
60-64	F	0.91	1.33	2.97	1.36	-
	H	1.39	1.46	3.17	1.46	-
65-69	F	1.42	1.14	2.31	1.26	-
	H	1.21	1.02	0.8	1.62**	-
70-74	F	1.41	1.21	1.52	0.96	-
	H	1.05	1.46	1.33	1.47*	-
75-79	F	-	-	-	-	-
	H	0.67	1.69	1.92	1.08	-
80-84	F	0.11**	1.19	1.23	0.72	-
	H	0	1.67	0.74	1.52	-
85-89	F	0.14*	0.6	0.5	0.29***	-
	H	0.77	1.63	1.39	0.77	-
90+	F	0.35	0.3	0.02	0.27**	-
	H	0.04	0.03	0.11	0.38	-

*** = $p < 0,01$ | ** = $p < 0,05$ | * = $p < 0,1$

Tableau C-24 V2 – Modèle 4 – *Études et travail*

		V2-E1	V2-T1	V2-T2	V2-T3	V2-T4	V2-T5	V2-I
Constante		0.01***	0.03***	0.05***	0.05***	0.04***	0.02***	-
5-14		434.46***	0***	0***	0***	0***	0***	-
15-24		121.41***	0.31***	1.35	0.33***	0.92	0.6**	-
25-44		8.34***	1.29***	1.57**	1.21*	2.1***	1.76***	-
45-64		-	-	-	-	-	-	-
65+		0.55	0.11***	0.15***	0.12***	0.17***	0.12***	-
Jeunes couples, colocataires et célibataire	F	2.82	56.07***	4.72*	22.46***	5.02*	14.4*	-
	H	2.52	87.37***	4.46*	28.73***	6.26**	27.86**	-
Jeunes familles	F	6.86	44.55***	1.43	26.44***	1.77	5.98	-
	H	6.28	98.31***	5.37*	42.8***	7.5**	16.53*	-
Familles traditionnelles	F	4.97	49.72***	2.7	22.82***	4.2*	13.64*	-
	H	5.48	76.69***	5.33*	33.28***	7.23**	15.57*	-
Adultes et personnes âgées	F	2.16	14.87**	2.25	6.35**	1.11	1.41	-
	H	2.72	28.17***	2.58	15.67***	1.64	8.75	-
Cinquantennaires	F	2.09	33.98***	2.05	19.45***	2.38	7.54	-
	H	3.22	73.01***	4.53*	28.04***	3.93	18.01**	-
Jeunes retraités	F	1.72	14.75**	0.67	7.21**	1.97	3.25	-
	H	0.8	26.84***	2.52	12.68***	2.53	3.76	-
Vieux retraités	F	-	-	-	-	-	-	-
	H	1.14	2.3	0.01	3.6	1.11	0.01	-
Familles multigénérationnelles	F	4.36	32.63***	2.12	14.07***	2.71	3.94	-
	H	2.42	38.22***	3.68	12.23***	4.37	9.05	-

Tableau C-25 V2 – Modèle 4 – *Loisir, magasinage et autres*

		V2-L1	V2-L2	V2-A1	V2-MA1	V2-I
Constante		0.11***	0.18***	0.07***	0.78*	-
5-14		0.77	1.44	0.39**	0.23***	-
15-24		1	0.85	0.47***	0.82	-
25-44		0.93	1.04	0.95	1.14	-
45-64		-	-	-	-	-
65+		0.88	0.81	0.41***	0.82**	-
Jeunes couples, colocataires et célibataire	F	2.27**	0.66	0.83	1.24	-
	H	2.55**	1.23	3.88**	1.6**	-
Jeunes familles	F	2.08**	0.78	1.49	1.93***	-
	H	1.33	0.55	5.75***	1.62***	-
Familles traditionnelles	F	1.13	0.7	1.1	1.36*	-
	H	1.33	0.71	3**	1.5**	-
Adultes et personnes âgées	F	0.57	0.58	0.98	0.8	-
	H	1.1	1.28	1.5	1.2	-
Cinquantennaires	F	1.15	0.86	0.7	1.15	-
	H	1.85	1.54	2.46*	1.93***	-
Jeunes retraités	F	1.92**	1.94***	1.6	1.56***	-
	H	1.68*	2.02***	1.2	1.9***	-
Vieux retraités	F	-	-	-	-	-
	H	0.72	1.3	1.75	1.31*	-
Familles multigénérationnelles	F	0.95	0.69	1.03	1	-
	H	0.3	0.65	2.79*	1.16	-

Tableau C-26 V2 – Modèle 5 – *Études et travail*

		V2-E1	V2-T1	V2-T2	V2-T3	V2-T4	V2-T5	V2-I
	Constante	0.2***	1.94***	0.13***	1.74***	0.24***	0.23***	-
	Homme	0.96	1.84***	1.99***	1.51***	2***	1.93***	-
	Femme	-	-	-	-	-	-	-
5-14	Jeunes couples, colocataires et célibataire	18654.3	0.12	0.77	0.14	0.63	0.65	-
	Jeunes familles	85.08***	0***	0	0***	0*	0	-
	Familles traditionnelles	137.67***	0	0	0	0	0	-
	Familles multigénérationnelles	125.87***	0	0	0	0	0	-
15-24	Jeunes couples, colocataires et célibataire	6.25***	0.29***	0.86	0.35***	0.49*	0.64	-
	Jeunes familles	31.99***	0.13***	1.28	0.25***	0.45**	0.24***	-
	Familles traditionnelles	21.71***	0.27***	1.28	0.16***	0.7	0.54**	-
	Familles multigénérationnelles	10.76***	0.14***	1.73	0.1***	0.24*	0.39	-
25-44	Jeunes couples, colocataires et célibataire	2.18***	1.07	2.1***	0.7**	1.53**	2.11***	-
	Jeunes familles	-	-	-	-	-	-	-
	Familles traditionnelles	1.67*	0.68**	1.17	0.65**	1.15	1.12	-
	Adultes et personnes âgées	0.8	0.47***	1.36	0.36***	0.66	0.94	-
	Cinquantennaires	0.5	0.78	1.51	0.69	0.43	1.23	-
	Familles multigénérationnelles	2.56**	0.69	1.05	0.65	1.21	0.63	-
45-64	Jeunes couples, colocataires et célibataire	108.68	13.63	0.13	17.72	0.07	55.37	-
	Jeunes familles	0.34*	1.07	1.52	0.88	0.74	0.99	-
	Familles traditionnelles	0.14***	0.72***	0.99	0.71***	0.61**	0.85	-
	Adultes et personnes âgées	0.09**	0.21***	0.6	0.26***	0.11***	0.06***	-
	Cinquantennaires	0.12***	0.57***	0.7	0.55***	0.4***	0.64**	-
	Jeunes retraités	0.01***	0.21***	0.32***	0.21***	0.24***	0.19***	-
	Familles multigénérationnelles	0	0.43***	0.24*	0.31***	0.54	0.29**	-
65+	Familles traditionnelles	0.04	0.49	0.03	0	1.22	0.01	-
	Adultes et personnes âgées	0.04***	0.02***	0.04***	0.03***	0	0.02***	-
	Cinquantennaires	0.18	0	0.12	0.45	0.06	0.06	-
	Jeunes retraités	0.01***	0.03***	0.06***	0.04***	0.06***	0.03***	-
	Vieux retraités	0	0***	0.03***	0.01***	0.02***	0	-
	Familles multigénérationnelles	0	0.01***	0.18**	0	0	0.05***	-

Tableau C-27 V2 – Modèle 5 – *Loisir, magasinage et autres*

		V2-L1	V2-L2	V2-A1	V2-MA1	V2-I
Constante		0.17***	0.13***	0.16***	1.62***	-
Homme		0.93	1.16	2.56***	1.18***	-
Femme		-	-	-	-	-
5-14	Jeunes couples, colocataires et célibataire	0.76	0.79	0.67	0.17	-
	Jeunes familles	0.87	1.31	0.33***	0.15***	-
	Familles traditionnelles	0	1.5	0	0.1**	-
	Familles multigénérationnelles	0.01	0.02	1.9	0.43	-
15-24	Jeunes couples, colocataires et célibataire	1.32	0.29*	0.49	0.61**	-
	Jeunes familles	1.27	0.9	0.06***	0.5***	-
	Familles traditionnelles	0.88	1.09	0.47**	0.56***	-
	Familles multigénérationnelles	0	0.63	0	0.49**	-
25-44	Jeunes couples, colocataires et célibataire	1.49	1.6*	0.5**	0.68***	-
	Jeunes familles	-	-	-	-	-
	Familles traditionnelles	0.58	0.8	0.19***	0.49***	-
	Adultes et personnes âgées	1.48	1.47	0.43	0.49**	-
	Cinquantennaires	0.41	0.5	0.14*	0.52**	-
	Familles multigénérationnelles	1.14	0.94	0.81	0.91	-
45-64	Jeunes couples, colocataires et célibataire	130.81	0.2	64.11	11.39	-
	Jeunes familles	1.37	0.49	1.12	0.79	-
	Familles traditionnelles	0.76	0.87	0.53**	0.66***	-
	Adultes et personnes âgées	0.55	0.81	0.38**	0.51***	-
	Cinquantennaires	0.97	1.64**	0.43***	0.67***	-
	Jeunes retraités	1.12	2.78***	0.42***	0.72***	-
	Familles multigénérationnelles	0.14*	0.66	0.53	0.33***	-
65+	Familles traditionnelles	5.91*	0.04	1.54	1.6	-
	Adultes et personnes âgées	0.24***	1.06	0.11***	0.33***	-
	Cinquantennaires	0.22	0.23	0.07	0.01	-
	Jeunes retraités	1.08	2.12***	0.18***	0.66***	-
	Vieux retraités	0.52***	1.24	0.16***	0.42***	-
	Familles multigénérationnelles	0.54	1.16	0.13***	0.33***	-

Tableau C-28 V3 – Modèle 1 – *Études et travail*

	V3-E1	V3-T1	V3-T2	V3-T3	V3-T4	V3-T5	V3-I
Constante	17638.53***	0.09***	0.07***	0.09***	0.07***	0.03***	-
Âge	0.77***	1	0.98***	1	0.98***	0.99**	-
Jeunes couples, colocataires et célibataire	0.03***	19.14***	6.77***	10.16***	7.92***	15.43***	-
Jeunes familles	0.05***	16***	3.87***	10.87***	4.48***	6.49***	-
Familles traditionnelles	0.05***	12.62***	5.45***	8.76***	5.16***	9.49***	-
Adultes et personnes âgées	0.02***	2.43***	2.91***	2.1***	1.22	1.72	-
Cinquantennaires	0.29	11.69***	4.8***	9.61***	4.08***	9.18***	-
Retraités	-	-	-	-	-	-	-
Familles multigénérationnelles	0.02***	5.22***	3.03***	2.7***	2.01*	3.17***	-
Homme	1.05	1.63***	2.01***	1.44***	1.97***	1.91***	-
Femme	-	-	-	-	-	-	-

Tableau C-29 V3 – Modèle 1 – *Loisir, magasinage et autres*

	V3-L1	V3-A1	V3-MA1	V3-MA2	V3-I
Constante	0.53*	0.05***	1.17	0.77	-
Âge	0.99***	1	0.99***	0.99**	-
Jeunes couples, colocataires et célibataire	0.99	1.84*	1.06	0.51***	-
Jeunes familles	0.7	2.8***	1.04	0.78	-
Familles traditionnelles	0.61**	1.79**	0.95	0.68***	-
Adultes et personnes âgées	0.52***	0.66	0.65***	0.57***	-
Cinquantennaires	0.6**	1.71**	1.09	0.83	-
Retraités	-	-	-	-	-
Familles multigénérationnelles	0.26***	1.34	0.77*	0.42***	-
Homme	0.92	2.45***	1.29***	1	-
Femme	-	-	-	-	-

Tableau C-30 V3 – Modèle 3 – Études et travail – Variable âge

		V3-E1	V3-T1	V3-T2	V3-T3	V3-T4	V3-T5	V2-I
Constante		0.02***	0.05***	0.02***	0.04***	0.03***	0.02*	-
5-14		715.27***	0.09	0***	0***	0***	0***	-
15-19	F	269.96***	2.97	2.25	3.76	1.36	0.01	-
	H	257.36***	5.23	11.67**	9.66**	2.94	3.28	-
20-24	F	72.47***	14.07**	15.99**	18.57***	5.99	7.3	-
	H	66.9***	25.7***	20.28**	21.59***	9.26*	9.13	-
25-29	F	23.08**	31.18***	14.08**	25.88***	11.11*	13.42	-
	H	25.28**	44.57***	16.58**	48.9***	17.24**	18.78	-
30-34	F	11.82*	40.28***	7.78*	43.94***	5.34	9.95	-
	H	10.05	82.29***	28.81***	63.51***	18.23**	27.54	-
35-39	F	7.38	37.33***	6.62	38.45***	6.1	5.48	-
	H	4.47	60.84***	10.98**	45.65***	15.15**	17.14	-
40-44	F	6.58	50.49***	11.96**	50.85***	7.03	11.19	-
	H	4.65	85.28***	25.52***	85.61***	13.21**	20.2	-
45-49	F	1.5	44.64***	6.88	51.65***	5.04	8.18	-
	H	2.75	65***	22.49***	61.38***	8.75*	16.28	-
50-54	F	1.75	33.41***	11.63**	41.24***	1.96	6.78	-
	H	2.3	66.48***	18.72**	75.45***	8.51	13.42	-
55-59	F	1.4	19.76***	1.96	25***	3.49	3.14	-
	H	1.46	44.19***	11.77**	45.79***	5.51	6.72	-
60-64	F	1	11.03**	1.68	11.84***	2.39	1.89	-
	H	0.63	18.57***	4.79	20.43***	4.19	3.51	-
65-69	F	0.53	2.1	0.43	3.09	0.72	0.4	-
	H	0.43	4.37	1.93	7.48**	1.14	1.31	-
70-74	F	0.44	0.98	0.35	1.23	0.81	0.01	-
	H	0.46	1.99	0.59	3.42	0.01	0.78	-
75-79	F	-	-	-	-	-	-	-
	H	0.59	1.06	0.01	3.51	0.03	3.49	-
80-84	F	0.45	0.17	0.01	2.28	0.03	0.08	-
	H	0.58	0.24	0.03	0***	2.17	0.14	-
85-89	F	0.31	0.15	0.02	0***	0.05	0.12	-
	H	0.41	0.18	0.05	0	0.11	0.23	-
90+	F	0.28	0.1	0.05	0.01	0.11	0.25	-
	H	0.29	0.12	0.17	0.14	0.3	0.5	-

Tableau C-31 V3 – Modèle 3 – *Études et travail* – Variable type de ménage

	V3-E1	V3-T1	V3-T2	V3-T3	V3-T4	V3-T5	V2-I
Jeunes couples, colocataires et célibataire	0.71**	1.06	1.34	0.85	1.27	1.86***	-
Jeunes familles	-	-	-	-	-	-	-
Familles traditionnelles	0.9	0.88	1.09	0.81*	1.23	1.58**	-
Adultes et personnes âgées	0.52	0.41***	1.18	0.49***	0.46*	0.51	-
Cinquantennaires	0.7	0.66***	0.74	0.71**	0.84	1.28	-
Jeunes retraités	0.59	0.54***	0.99	0.63**	0.71	1.09	-
Vieux retraités	0.37	0.11**	1.59	0.38*	0.29	0	-
Familles multigénérationnelles	0.6***	0.68**	1.02	0.49***	0.68	0.86	-

Tableau C-32 V3 – Modèle 3 – *Loisir, magasinage et autres* – Variable type de ménage

	V3-L1	V3-A1	V3-MA1	V3-MA2	V2-I
Jeunes couples, colocataires et célibataire	1.34	0.63*	0.84	0.65**	-
Jeunes familles	-	-	-	-	-
Familles traditionnelles	0.64**	0.55***	0.78**	0.81	-
Adultes et personnes âgées	0.57*	0.35***	0.61***	0.64**	-
Cinquantennaires	0.49***	0.41***	0.84	0.82	-
Jeunes retraités	1.07	0.6	1.06	0.96	-
Vieux retraités	0.83	0.6	0.75	1.13	-
Familles multigénérationnelles	0.3***	0.6	0.76	0.49***	-

Tableau C-33 V3 – Modèle 3 – *Loisir, magasinage et autres* – Variable âge

		V3-L1	V3-A1	V3-MA1	V3-MA2	V2-I
Constante		0.17***	0.06***	0.7	0.35***	-
5-14		0.82	2.04	0.37***	1.01	-
15-19	F	1.82	0.44	1.14	1.08	-
	H	0.45	1.87	1.31	1.33	-
20-24	F	1.85	1.08	1.95**	1.69	-
	H	1.53	5.3**	2**	1.02	-
25-29	F	1.24	3.79*	1.67*	1.19	-
	H	1.69	4.56*	1.93**	0.84	-
30-34	F	1.33	1.58	1.86**	2.14**	-
	H	1.08	8.24***	2.36***	1.4	-
35-39	F	1.95	1.51	1.61	1.69	-
	H	0.88	8.75***	1.36	0.84	-
40-44	F	1.74	2.33	1.78**	1.61	-
	H	1.39	7.69***	2.08**	1.13	-
45-49	F	1.01	3.58*	1.32	1.28	-
	H	1.51	9.95***	1.8**	1.25	-
50-54	F	2.17	3.41	1.31	1.38	-
	H	2.5*	10.66***	2.23***	1.37	-
55-59	F	1.54	2.59	1.22	1.66	-
	H	1.46	8.94***	1.85**	2.05**	-
60-64	F	1.43	2.77	1.16	1.72*	-
	H	1.84	3.17	1.41	1.5	-
65-69	F	1.43	1.95	0.96	1.87**	-
	H	1.3	1.67	1.28	1.87**	-
70-74	F	1.45	2.27	0.97	0.83	-
	H	0.96	1.85	1.54*	1.45	-
75-79	F	-	-	-	-	-
	H	0.68	1.52	1.12	1.34	-
80-84	F	0.49	1.1	0.72	0.8	-
	H	0.18	0.74	1.63*	1.29	-
85-89	F	0.12**	0.51	0.28***	0.38**	-
	H	1.62	1.44	0.84	0.73	-
90+	F	0.32	0.03	0.2**	0.32*	-
	H	0.03	0.11	0.52	0.01	-

Tableau C-34 V3 – Modèle 4 – *Études et travail*

		V3-E1	V3-T1	V3-T2	V3-T3	V3-T4	V3-T5	V3-I
Constante		0.01***	0.05***	0.09***	0.1***	0.03***	0.01***	-
5-14		571.88***	0***			0***	0***	-
15-24		150.91***	0.32***	1.21	0.32***	1.01	0.67*	-
25-44		8.76***	1.36***	1.47**	1.16	2.37***	1.87***	-
45-64		-	-	-	-	-	-	-
65+		0.39	0.09***	0.12***	0.14***	0.17***	0.15***	-
Jeunes couples, colocataires et célibataire	F	1.47	34.94***	2.56	12.11***	7.43*	15.15*	-
	H	1.48	58.12***	2.6	15***	8.77**	25.59**	-
Jeunes familles	F	4.01	31.75***	0.78	13.02***	2.4	5.09	-
	H	3.69	63.87***	3.04	22.59***	9.3**	16.83*	-
Familles traditionnelles	F	3.04	34.05***	1.36	11.5***	5.31	14.95*	-
	H	2.92	49.1***	3.06	17.99***	10.32**	18.68**	-
Adultes et personnes âgées	F	1.17	8.75**	1.23	3.87**	1.42	0.88	-
	H	1.43	19.61***	1.82	8.09***	3.23	8.05	-
Cinquantennaires	F	1.81	22.38***	1.12	11.02***	3.58	7.99	-
	H	2.27	44.71***	2.18	15.75***	6.82*	21.01**	-
Jeunes retraités	F	1.54	8.77**	0.34	3.93**	2.65	3.74	-
	H	1.2	15.71***	1.25	6.75***	3.54	6.16	-
Vieux retraités	F	-	-	-	-	-	-	-
	H	1.2	1.87	0.01	1.65	1.95	0.02	-
Familles multigénérationnelles	F	2.41	24.09***	1.2	6.56***	3.43	5.01	-
	H	1.45	26.54***	2.18	6.75***	3.83	10.14	-

Tableau C-35 V3 – Modèle 4 – *Loisir, magasinage et autres*

		V3-L1	V3-A1	V3-MA1	V3-MA2	V3-I
Constante		0.15***	0.07***	0.54***	0.4***	-
5-14		0.61	0.34***	0.26***	0.73	-
15-24		0.96	0.41***	1.1	0.92	-
25-44		0.97	0.89	1.24*	0.97	-
45-64		-	-	-	-	-
65+		0.74*	0.47***	0.79**	0.88	-
Jeunes couples, colocataires et célibataire	F	2.11**	0.88	1.38	0.91	-
	H	2.52**	4.69***	2.07***	0.72	-
Jeunes familles	F	1.79*	2.05	1.9***	1.59**	-
	H	1.35	7.02***	1.8***	0.91	-
Familles traditionnelles	F	1.27	1.24	1.41*	0.97	-
	H	1.01	3.8***	1.76***	1.09	-
Adultes et personnes âgées	F	0.51	0.81	0.85	0.72	-
	H	1.48	1.28	1.35	0.89	-
Cinquantennaires	F	0.85	0.94	1.27	0.99	-
	H	1.19	3.14**	2.41***	1.42	-
Jeunes retraités	F	2.22***	1.83	1.68***	1.49**	-
	H	1.72**	1.62	2.24***	1.52**	-
Vieux retraités	F	-	-	-	-	-
	H	0.74	1.62	1.42*	1.23	-
Familles multigénérationnelles	F	0.62	1.48	1.35	0.54*	-
	H	0.42	3.11**	1.38	0.77	-

Tableau C-36 V3 – Modèle 5 – *Études et travail*

		V3-E1	V3-T1	V3-T2	V3-T3	V3-T4	V3-T5	V3-I
Constante		0.15***	2.2***	0.13***	1.55***	0.2***	0.17***	-
Homme		0.97	1.77***	2.07***	1.57***	2.09***	2.02***	-
Femme		-	-	-	-	-	-	-
5-14	Jeunes couples, colocataires et célibataire	8186.16	0.12	0.78	0.18	0.7	0.76	-
	Jeunes familles	111.52***	0***	0	0***	0*	0	-
	Familles traditionnelles	178.26***	0.01	0.02	0***	0.01	0.01	-
	Familles multigénérationnelles	184.29***	0	0.03	0***	0.02	0.02	-
15-24	Jeunes couples, colocataires et célibataire	6.89***	0.28***	0.87	0.37***	0.57	0.58	-
	Jeunes familles	39.84***	0.13***	1.26	0.23***	0.52*	0.31**	-
	Familles traditionnelles	26.5***	0.25***	1.25	0.19***	0.69	0.66	-
	Familles multigénérationnelles	13.59***	0.13***	1.67	0.09***	0.28*	0.51	-
25-44	Jeunes couples, colocataires et célibataire	2.63***	1	2.07***	0.75**	1.7***	2.16***	-
	Jeunes familles	-	-	-	-	-	-	-
	Familles traditionnelles	1.7	0.66**	1.16	0.72*	1.27	1.28	-
	Adultes et personnes âgées	0.86	0.47***	1.66	0.37***	0.75	0.89	-
	Cinquantennaires	0.46	0.68	1.12	0.79	0.99	1.29	-
	Familles multigénérationnelles	2.96**	0.77	1.05	0.64	0.68	0.55	-
45-64	Jeunes couples, colocataires et célibataire	68.8	6.02	0.2	9.35	0.14	33.97	-
	Jeunes familles	0.16*	0.94	1.49	0.99	0.82	0.85	-
	Familles traditionnelles	0.14***	0.67***	1.03	0.75**	0.59**	0.98	-
	Adultes et personnes âgées	0.13**	0.17***	0.6	0.35***	0.13***	0	-
	Cinquantennaires	0.19***	0.49***	0.67	0.63***	0.41***	0.72	-
	Jeunes retraités	0.04***	0.17***	0.3***	0.23***	0.23***	0.24***	-
	Familles multigénérationnelles	0.01	0.41***	0.36*	0.32***	0.47*	0.37*	-
65+	Familles traditionnelles	0.11	0.47	0.05	0	1.47	0.04	-
	Adultes et personnes âgées	0	0.01***	0.04***	0.03***	0	0.03***	-
	Cinquantennaires	0.34	0.01	0.18	0.6	0.12	0.15	-
	Jeunes retraités	0**	0.02***	0.06***	0.05***	0.05***	0.05***	-
	Vieux retraités	0**	0***	0.04***	0.01***	0.02***	0	-
	Familles multigénérationnelles	0.01	0.01***	0.09**	0.01***	0	0.07***	-

Tableau C-37 V3 – Modèle 5 – *Loisir, magasinage et autres*

		V3-L1	V3-A1	V3-MA1	V3-MA2	V3-I
Constante		0.24***	0.17***	1.17*	0.55***	-
Homme		0.92	2.58***	1.32***	1.01	-
Femme		-	-	-	-	-
5-14	Jeunes couples, colocataires et célibataire	0.76	0.71	0.27	0.52	-
	Jeunes familles	0.6	0.29***	0.17***	0.57**	-
	Familles traditionnelles	0.01	0.01	0.14*	0.77	-
	Familles multigénérationnelles	0.02	1.79	0.62	1.5	-
15-24	Jeunes couples, colocataires et célibataire	1.08	0.51	0.77	0.58*	-
	Jeunes familles	0.93	0.05***	0.66**	0.84	-
	Familles traditionnelles	0.92	0.36***	0.77	0.71	-
	Familles multigénérationnelles	0	0	0.73	0.16**	-
25-44	Jeunes couples, colocataires et célibataire	1.49*	0.5**	0.84	0.57***	-
	Jeunes familles	-	-	-	-	-
	Familles traditionnelles	0.41**	0.23***	0.58***	0.54**	-
	Adultes et personnes âgées	1.03	0.25*	0.56*	0.68	-
	Cinquantennaires	0.42	0.13**	0.54**	0.48*	-
	Familles multigénérationnelles	0.78	0.97	1.03	0.74	-
45-64	Jeunes couples, colocataires et célibataire	44.18	27.63	0.02	18.62	-
	Jeunes familles	1.22	1.18	0.79	0.67	-
	Familles traditionnelles	0.7	0.56**	0.65***	0.74*	-
	Adultes et personnes âgées	0.55	0.36**	0.41***	0.75	-
	Cinquantennaires	0.66*	0.48***	0.73***	0.87	-
	Jeunes retraités	1.24	0.41***	0.72**	1.08	-
	Familles multigénérationnelles	0.19**	0.48	0.46***	0.16***	-
65+	Familles traditionnelles	4.39	1.47	0.72	3.75	-
	Adultes et personnes âgées	0.37***	0.08***	0.35***	0.41***	-
	Cinquantennaires	0.23	0.11	0.03	0.09	-
	Jeunes retraités	0.95	0.21***	0.64***	0.96	-
	Vieux retraités	0.44***	0.15***	0.38***	0.69***	-
	Familles multigénérationnelles	0.4*	0.11***	0.31***	0.62*	-

ANNEXE D COMPARAISON DU NOMBRE D'ACTIVITÉS OBSERVÉES ET GÉNÉRÉES AVEC V1

Tableau D-38 Nombre d'activités observées et générées et ratio entre ces nombres avec V1 – Modèle 3

Âge	Genre	<i>Travail</i>			<i>Études</i>			<i>Loisir</i>			<i>Magasinage</i>			<i>Autres</i>		
		Générées	Observées	Ratio	Générées	Observées	Ratio	Générées	Observées	Ratio	Générées	Observées	Ratio	Générées	Observées	Ratio
5 - 9	H	25	0	0.00	94639	93832	0.99	9606	13225	1.38	3855	4183	1.09	8941	8049	0.90
	F	15	0	0.00	84949	84606	1.00	7998	10874	1.36	3389	3877	1.14	8345	7647	0.92
10 - 19	H	14081	45327	3.22	175732	128523	0.73	34313	37597	1.10	6193	16398	2.65	11232	18962	1.69
	F	11605	12114	1.04	166284	167153	1.01	28822	32328	1.12	10727	8307	0.77	11323	12684	1.12
20 - 29	H	116650	124979	1.07	49546	18220	0.37	46338	38299	0.83	21820	37715	1.73	34913	51116	1.46
	F	104936	120833	1.15	52175	46247	0.89	40858	47846	1.17	34076	24082	0.71	46037	47696	1.04
30 - 39	H	177527	155013	0.87	9352	6721	0.72	37840	41992	1.11	33709	52660	1.56	88145	88870	1.01
	F	143049	174619	1.22	10484	9010	0.86	38803	36613	0.94	54117	39757	0.73	127441	118323	0.93
40 - 49	H	192560	193663	1.01	4353	4271	0.98	35401	38015	1.07	40394	43288	1.07	96850	101005	1.04
	F	174490	166498	0.95	5634	4915	0.87	38103	39544	1.04	57802	66198	1.15	98411	88477	0.90
50 - 59	H	129547	80233	0.62	1172	2899	2.47	31519	38021	1.21	38672	59617	1.54	56027	42539	0.76
	F	113008	134986	1.19	2363	1280	0.54	39787	35950	0.90	57859	44382	0.77	46614	55220	1.18
60 - 69	H	41247	9581	0.23	837	1468	1.75	33889	37257	1.10	49025	52230	1.07	36504	31557	0.86
	F	24545	42542	1.73	775	1003	1.29	38665	37491	0.97	61896	50221	0.81	31766	36267	1.14
70 +	H	7289	3198	0.44	219	309	1.41	27000	20238	0.75	39727	29880	0.75	24884	17249	0.69
	F	4569	8917	1.95	443	183	0.41	31952	40133	1.26	55034	60180	1.09	27566	35078	1.27

ANNEXE E RAPPORTS DE PROBABILITÉ DU MODÈLE 6 – V1

Tableau E-39 Rapport de probabilités – *Études et travail*

	V1-E1	V1-E2	V1-T1	V1-T2	V1-T3	V1-T4	V1-I
Constante	0.08***	0.03***	0.42***	0.05***	0.24***	0.05***	-
5-9	122.61***	25.6***	0.06***	3.16***	0.23***	1.19*	-
10-14	157.87***	48.66***	0.04***	2.38***	0.39***	1.65***	-
15-19	58.45***	25.87***	0.8***	2.68***	0.28***	1.83***	-
20-24	17.91***	8.85***	2.19***	5.55***	0.57***	6.22***	-
25-29	4.69***	1.74***	2.44***	4.9***	0.54***	9.46***	-
30-34	2.44***	0.56***	3.29***	6.18***	1.07	14.56***	-
35-39	2.78***	1.6***	4.8***	7.95***	1.69***	19.98***	-
40-44	2.97***	2.05***	6.45***	10.13***	2.33***	20.09***	-
45-49	2.28***	2.59***	6.77***	9.81***	2.23***	24.67***	-
50-54	2.13***	2.68***	6.44***	8***	2.06***	18.92***	-
55-59	1.64***	2.48***	4.35***	4.08***	1.92***	7.81***	-
60+	-	-	-	-	-	-	-
Jeunes couples, colocataires et célibataire	1.19***	2.38***	2.16***	2.28***	2.3***	0.47***	-
Jeunes familles	-	-	-	-	-	-	-
Familles traditionnelles	0.93***	1.05*	1.07***	1.05*	0.79***	0.43***	-
Adultes et personnes âgées	0.48***	0.46***	0.46***	0.45***	0.34***	0.15***	-
Cinquantennaires	0.48***	0.46***	0.64***	0.83***	0.58***	0.23***	-
Jeunes retraités	0.7***	0.4***	0.63***	1	0.35***	0.48***	-
Vieux retraités	0.43***	0.36***	0.06***	0.55***	0.11***	0.28***	-
Familles multigénérationnelles	0.69***	0.5***	0.54***	0.43***	0.26***	0.35***	-
Homme	1.2***	1.35***	1.78***	1.85***	1.33***	1.8***	-
Femme	-	-	-	-	-	-	-
Nombre d'années	1.04***	1.02***	1.01***	1.01***	1	1.03***	-

Tableau E-40 Rapport de probabilités – *Loisir, magasinage et autres*

	V1-L1	V1-L2	V1-M1	V1-A1	V1-I
Constante	1.57***	28.9***	0.93**	1.52***	-
5-9	0.09***	0***	0.17***	0.15***	-
10-14	0.19***	0***	0.28***	0.11***	-
15-19	0.22***	0***	0.17***	0.15***	-
20-24	0.2***	0***	0.22***	0.23***	-
25-29	0.13***	0***	0.28***	0.24***	-
30-34	0.13***	0***	0.51***	0.35***	-
35-39	0.17***	0.01***	0.77***	0.41***	-
40-44	0.24***	0.01***	0.66***	0.78***	-
45-49	0.26***	0.04***	0.71***	0.69***	-
50-54	0.42***	0.12***	0.91***	1	-
55-59	0.67***	0.32***	1.18***	1.16***	-
60+	-	-	-	-	-
Jeunes couples, colocataires et célibataire	1.62***	2.85***	1.24***	0.77***	-
Jeunes familles	-	-	-	-	-
Familles traditionnelles	0.63***	0.08***	0.68***	0.39***	-
Adultes et personnes âgées	0.21***	0.01***	0.67***	0.2***	-
Cinquantennaires	0.48***	0.07***	0.73***	0.22***	-
Jeunes retraités	0.31***	0.02***	0.79***	0.19***	-
Vieux retraités	0.15***	0.01***	0.59***	0.12***	-
Familles multigénérationnelles	0.25***	0.02***	0.56***	0.26***	-
Homme	1.22***	1.02	0.92***	1.45***	-
Femme	-	-	-	-	-
Nombre d'années	0.99***	0.94***	0.99***	1.01***	-