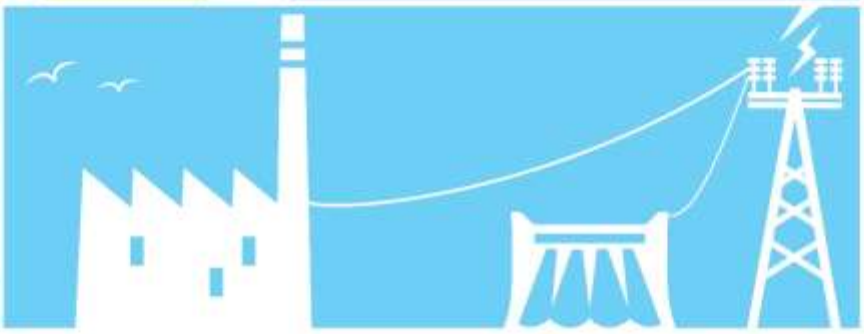
Chaire de gestion
du secteur de l'énergie
HEC MONTRÉAL

Rapport d'étude

Nº 01 | 2015

Portrait des différents profils de ménages canadiens en termes de dépenses d'énergie entre 1997 et 2009

Sylvain Brodier



Rapport d'étude n° 01 | 2015

Portrait des différents profils de ménages canadiens en termes de dépenses d'énergie entre 1997 et 2009

Sylvain Brodier

M.Sc Science de la gestion - intelligence d'affaires, HEC Montréal

Sous la supervision de Pierre-Olivier Pineau, professeur titulaire, Département des sciences de la décision, titulaire de la Chaire de gestion du secteur de l'énergie

Note aux lecteurs : Les rapports d'étude de la Chaire de gestion du secteur de l'énergie sont des publications aux fins d'information et de discussion. Ils ont été réalisés par des étudiants sous la supervision d'un professeur. Ils ne devraient pas être reproduits sans l'autorisation écrite du (des) auteur(s). Les commentaires et suggestions sont bienvenus, et devraient être adressés à (aux) auteur(s). Pour consulter les rapports d'étude et les cahiers de recherche de la Chaire de gestion du secteur de l'énergie à HEC Montréal, visitez le site http://energie.hec.ca.

À propos de la Chaire de gestion du secteur de l'énergie : Créée en 2013, la Chaire de gestion du secteur de l'énergie de HEC Montréal a pour mission d'augmenter les connaissances sur les enjeux liés à l'énergie, dans une perspective de développement durable, d'optimisation et d'adéquation entre les sources d'énergie et les besoins de la société. La création de cette chaire et de ce rapport est rendue possible grâce au soutien d'entreprises partenaires. Pour plus d'information ou pour consulter nos autres publications, visitez le site http://energie.hec.ca.

Avril, 2015

Chaire de gestion du secteur de l'énergie

HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada
H3T 2A7

Conception infographique : Émilie Parent et Johanne Whitmore

Copyright©2015 HEC Montréal. Tous droits réservés pour tous pays. Toute traduction et toute reproduction sous quelque forme que ce soit sont interdites. Les textes publiés dans la série des rapports d'étude n'engagent que la responsabilité de(s) auteur(s).

Remerciement

Je souhaite remercier Pierre-Olivier Pineau pour son encadrement et les nombreux échanges que nous avons pu avoir quant au projet.

Je tiens à remercier également Marc Fredette, Denis Larocque et Jean-François Plante avec qui j'ai pu échanger à plusieurs reprises ainsi que la Chaire de Gestion du Secteur de l'Énergie de HEC Montréal.

Table des matières

Remerciement 2

Liste des figures 10

Liste des tableaux 12

Sommaire exécutif 13

1 Introduction 19

1.1 La problématique 19

1.2 Les objectifs et le plan du rapport 24

1.3 Revue de littérature 26

1.3.1 Les déterminants/facteurs de la consommation (dépense) en énergie des ménages 27

1.3.2 Segmentation des ménages en tant que consommateur d'énergie 31

2 EDM et la préparation des données 41

2.1 L'enquête sur les dépenses des ménages (EDM) de Statistique Canada 41

2.1.1 Historique des enquêtes menées sur les dépenses des ménages au Canada 41

2.1.2 Les fichiers de microdonnées à grande diffusion (FMGD) 42

2.1.3 L'imputation des données manquantes : 43

2.1.4 Le masquage des données 43

2.1.5 La population cible d'EDM 44

2.1.6 Le questionnaire EDM en un clin d'oeil 45

2.1.7 La taille des échantillons de 1997 à 2009, EDM 46

2.2 La préparation des données 47

2.2.1 Changement de nom des variables par Statistique Canada 47

2.2.2 Ajout et suppression de variables d'une année à l'autre par Statistique Canada 47

2.2.3 Variation dans la codification / l'échelle des variables par Statistique Canada 47

2.2.4 Ajout de données externes pour fin d'analyse 48

2.3 Analyse descriptive des dépenses en énergie pour les ménages canadiens de 1997 à 2009 49

2.3.1 Préalable : point sur l'inférence statistique à partir des données EDM 49

2.3.2 EDM et les dépenses en énergie d'un ménage..... 49

2.4 La part des dépenses d'énergie dans le total des dépenses d'un ménage..... 56

2.4.1 Au niveau canadien..... 56

2.4.2 Au niveau des provinces..... 57

3 Méthodologie pour la construction et l'analyse du portrait des ménages québécois 59

3.1 L'analyse de regroupement (« clustering analysis »)..... 59

3.1.1 Le choix des variables : bruits et incertitudes..... 61

3.1.2 Le choix de l'échelle..... 62

3.1.3 Le choix de la mesure de distance (mesure de dissemblance) :..... 62

3.1.4 Première étape : une méthode hiérarchique..... 63

3.1.5 Représentation visuelle de la méthode hiérarchique ascendante : le dendogramme 63

3.1.6 Choix du nombre de groupes (analyse graphique et critères statistiques)..... 65

3.1.7 Deuxième étape : une méthode non-hiérarchique, k-means..... 67

3.1.8 Initialisation de k-means, optimalité et méta-heurisitque..... 67

3.1.9 Mesure de la qualité d'une partition donnée..... 67

3.2 Détection et sélection d'un regroupement (structure) à partir des données EDM pour le Québec..... 68

3.2.1 Pertinence du regroupement : interprétation des groupes et questionnement fertile 69

3.2.2 Pertinence du regroupement : possibilité de suivre ce regroupement dans le temps 74

3.2.3 Sélection d'un regroupement qui satisfait les deux critères : interprétable et valide dans le temps 80
3.3 Analyse du regroupement en 8 groupes, Québec, 2009 83
3.3.1 Méthode et validation statistique pour réaliser ces comparaisons 83
3.3.2 Étape 1 : analyse globale et sélection des variables pertinentes 85
3.3.3 Étape 2 : analyse plus fine, 2 à 2, par paire 87
3.3.4 Le cas des variables qui ont servi à construire le regroupement G024, G022 ainsi que HHINCTOT (qui n'a pas servi à construire le regroupement). 88
4 Résultats : 90
4.1 Québec, portrait des ménages en 2008 90
4.1.1 Le cas des variables numériques 90
4.1.2 Le cas des variables catégorielles 92
4.1.3 Analyse des caractéristiques des groupes et pertinence du regroupement 99
4.2 Québec, évolution du portrait des ménages de 1997 à 2009 103
4.2.1 L'évolution des centroïdes entre 1997 et 2009 103
4.2.2 L'évolution de la taille des groupes sur la période 1997 – 2009 105
4.3 Évolution de chaque groupe selon la dépense en électricité et la répartition des quintiles 106
4.3.1 Groupe 1 106
4.3.2 Groupe 2 107
4.3.3 Groupe 3 109
4.3.4 Groupe 4 110
4.3.5 Groupe 5 111
4.3.6 Groupe 6 113
4.3.7 Groupe 7 114
4.3.8 Groupe 8 116

5 Conclusion 118

6 Annexes 119

6.1 Annexe - La liste des variables complètes des données de 1997 à 2009, en ordre alphatbétique 119

6.2 Annexe - Le résumé chronologique des changements ("Résumé chronologique des changements - Enquête sur les dépenses des ménages (EDM)," n.d.) 127

6.2.1 Période de référence de changement -2012 127

6.2.2 Période de référence de changement -2011 127

6.2.3 Période de référence de changement -2010 127

6.2.4 Période de référence de changement -2009 128

6.2.5 Période de référence de changement -2007 128

6.2.6 Période de référence de changement -2006 128

6.2.7 Période de référence de changement -2004 129

6.2.8 Période de référence de changement -2000 129

6.2.9 Période de référence de changement -1999 130

6.3 Annexe - EDM et territoires 130

6.4 Annexes - Chiffres clés 131

6.4.1 Le nombre de ménage estimé par année au québec : 131

6.4.2 Le nombre de ménages dont le revenu est égal ou inférieure à zéro (Canada) 132

6.4.3 Évolution du prix moyen du Cent/kWh pour l'Ontario, l'Alberta, la Colombie-Britanique et le Québec (1997 – 2009) 133

6.4.4 Top 1997, 2009 et Top variation de la dépense en électricité 134

6.4.5 Évolution du prix de l'électricité (cent/kWh) dans les provinces canadiennes 134

6.5 Annexe - Inertie d'une partition (source : (Tufféry, 2012)) 136

6.6 Annexes - Québec, partition à 8 groupes 137

6.6.1 Premier regroupement (Québec, 2009) : HHINCTOT, G024, G022, EN_RE (avec standardisation) 137
6.6.2 Cas des valeurs extrêmes, Québec 2009 139
6.6.3 Représentation graphique du premier regroupement 140
6.6.4 Deuxième regroupement (Québec, 2008) : HHINCTOT, G024, G022, EN_RE (sans standardisation) 143
6.6.5 Représentation graphique du deuxième regroupement 145
6.7 Annexes - Analyse du regroupement 148
6.7.1 Légende pour la boîte à moustache. 148
6.7.2 Hypothèses de normalité pour G024, Québec 2008 148
6.7.3 Tests statistique formel pour la constance de la variance, variable G024, Québec 2008 149
6.7.4 Test de Welch pour l'égalité de la moyenne, G024, Québec 2008 (ANOVA) 150
6.7.5 Test de Fisher classique, qui ne prend pas en compte la constance de la variance, G024, Québec 2008. 151
6.7.6 Liste des variables utilisées pour la MANOVA 152
6.7.7 Analyse plus fine pour G024 : Proc Glimmix / Tukey-Kramer 154
6.7.8 Analyse plus fine pour G024 : Proc Glimmix / Bonferroni 155
6.8 Annexes - Portrait des ménages québecois, 2008 156
6.8.1 L'habitation 156
6.8.2 Les dépenses liées aux logements 158
6.8.3 Les variables catégorielles 164
6.9 Annexes – Québec, évolution du portrait des ménages de 1997 à 2009 166
6.9.1 Évolution des centroïdes du regroupement à 8 groupes, de 2008 à 2009 (G024 / HHINCTOT) 166

6.9.2 Évolution des centroïdes du regroupement à 8 groupes, de 1997 à 2009 (G024 / G022) 167

6.9.3 Évolution de la variable G024 par le groupe 1 entre 1997 et 2009 174

7 Bibliographie 175

Liste des figures

Figure 1-1 : Consommation d'énergie par secteur d'activité au Québec , 2012 (Source : État de l'énergie au Québec 2015)..... 20
Figure 1-2 : Consommation d'énergie par type d'utilisation dans le secteur résidentiel au Québec, 2011 (Source : État de l'énergie au Québec 2015)..... 20
Figure 1-3 : Dépenses d'énergie des ménages selon leur niveau de revenu (2009). Source : Statistiques Canada (EDM). Sur ce graphique, les pourcentages indiquent la part de la dépense d'électricité, de gaz et de transport sur le total de la dépense en énergie. 21
Figure 1-4 : Part des dépenses d'énergie des ménages selon leur niveau de revenu (2009). Source : Statistique Canada (EDM). 22
Figure 1-5 : Part des ménages qui n'ont pas de dépense en énergie selon le niveau de revenu (2009). Source : Statistique Canada (EDM). 23
Figure 1-6 : Médiane des dépenses en électricité selon le niveau de revenu (2009). Source : Statistique Canada (EDM). 24
Figure 1-7 : Caratéristiques du logement et du ménage qui influencent la consommation d'électricité résidentielle. (McLoughlin, Duffy, & Conlon, 2012) 27
Figure 1-8 : La part (moyenne) des dépenses en énergie par groupe de ménages sur la périodre 1982-2004 (Cashin & McGranahan, 2006). 35
Figure 1-9 : Méthodologie pour générer des profils spécifiques de consommation (Yao & Steemers, 2005)..... 37
Figure 2-1 : Les enquêtes sur les dépenses des ménages canadiens, avant 1997 jusqu'à nos jours. 41
Figure 2-2 : Le questionnaire EDM en un clin d'oeil 45
Figure 2-3 : La taille des échantillons pour chaque année (nombre d'observations). 46
Figure 3-1 : Dendogramme pour l'analyse de regroupement hiérarchique (Québec, 2009)..... 64
Figure 3-2 : R-squared, Québec, 2009. 65
Figure 3-3 : Semi-Partial R-Squared, Québec, 2009 65

Figure 3-4 : Québec, Comparaison du regroupement entre 2009 et 2008. La taille des bulles est proportionnelle à la taille estimée des groupes pour l'année en question. Voir aussi annexe 6.6.3. 77
Figure 3-5 : Québec, Comparaison du regroupement entre 2009 et 2008. Problématique des valeurs extrêmes pour En-Re..... 78
Figure 3-6 : Visualisation du regroupement sélectionné (8 groupes, Québec, 2008). 81
Figure 3-7 : Visualisation du regroupement sélectionné pour l'année 2008 et 2009. 82
Figure 3-8 : Québec, 2009, distribution de la dépense en électricité au sein des 8 groupes..... 85
Figure 3-9 : Distribution du revenu total au sein de chaque groupe, Québec 2008. 88
Figure 3-10 : Distribution de la dépense en combustibles au sein de chaque groupe, Québec 2008. 89
Figure 4-1 : Répartition des quintiles de revenu au sein des 8 groupes, Québec 2008 100
Figure 4-2 : Répartition des quintiles de revenu au sein des 8 groupes, Québec 1998 101
Figure 4-3 : Évolution des centroïdes, Québec, de 2008 à 2009. 103
Figure 4-4 : Évolution des centroïdes, Québec, de 1997 à 2009. 104
Figure 4-5 : Évolution de la taille des groupes, Québec. de 1997 à 2008. 105
Figure 6-1 : Légende de la boîte à moustache. 148

Liste des tableaux

Table 1 : Historique du regroupement, Québec, 2009 (Sortie SAS) 66
Table 2 : Québec, 2009, centroïde 74
Table 3 : Québec, 2009, centroïde ajusté pour l'année 2008 75
Table 4 : Québec, 2008, G024 moyenne et variance par groupe. 86
Table 5 : Québec, 2008, G024 test formel sur l'égalité de la variance entre groupes 86
Table 6 : Groupe 1 et 2, loyer moyen pour les locataires, Québec 2008 99

Sommaire exécutif

Contexte et problématique

Le secteur résidentiel compte pour près de 18% de la consommation d'énergie totale au Québec en 2014 (partie 1.1), ce qui en fait un champ d'étude opportun pour qui veut comprendre les déterminants de la demande en énergie, qu'il s'agisse des producteurs d'énergie ou bien des concepteurs de programme d'économie d'énergie par exemple.

L'analyse de la demande en énergie, et plus particulièrement celle des ménages, fait l'objet de nombreuses enquêtes. Différentes recherches ont été effectuées afin d'identifier quels sont les déterminants de la demande en énergie des ménages (partie 1.3.1). Aussi, certains articles de recherche proposent une segmentation des ménages en utilisant différents critères, comme le style de vie du ménage ou bien l'appartenance culturelle du ménage. Il est aussi courant d'utiliser un critère économique et, par exemple, de regrouper les ménages par quintile de revenu afin de comparer la consommation en énergie de chacun de ces groupes.

Il apparaît que ce regroupement des ménages par quintile de revenu présente plusieurs lacunes qui sont explicitées dans la partie 1.1 de ce rapport. Entre autres, les quintiles de revenu ne forment pas des groupes de ménages homogènes en matière de consommation/dépenses en énergie. Le présent rapport propose une segmentation des ménages québecois plus appropriée pour mettre en avant les différents profils de ménages en matière de dépenses en énergie.

Objectif du rapport

Dans ce rapport sont présentés des groupes de ménages homogènes en matière de dépenses en énergie et ainsi une typologie des ménages est mise à disposition pour fournir des clés de compréhension de la demande du secteur résidentiel.

Les données utilisées dans ce rapport

Le rapport repose sur les données de l'Enquête sur les Dépenses des Ménages (EDM). EDM est realisée chaque année depuis 1997 par Statistique Canada auprès de plus de 10.000 ménages canadiens quant à leurs dépenses. Près de 250 variables sont collectées pour recenser l'ensemble des dépenses des ménages dont celles en énergie comme l'électricité, le gaz et l'essence pour le transport. Fort de ces micro-données collectées par Statistique Canada de 1997 à 2009, il est

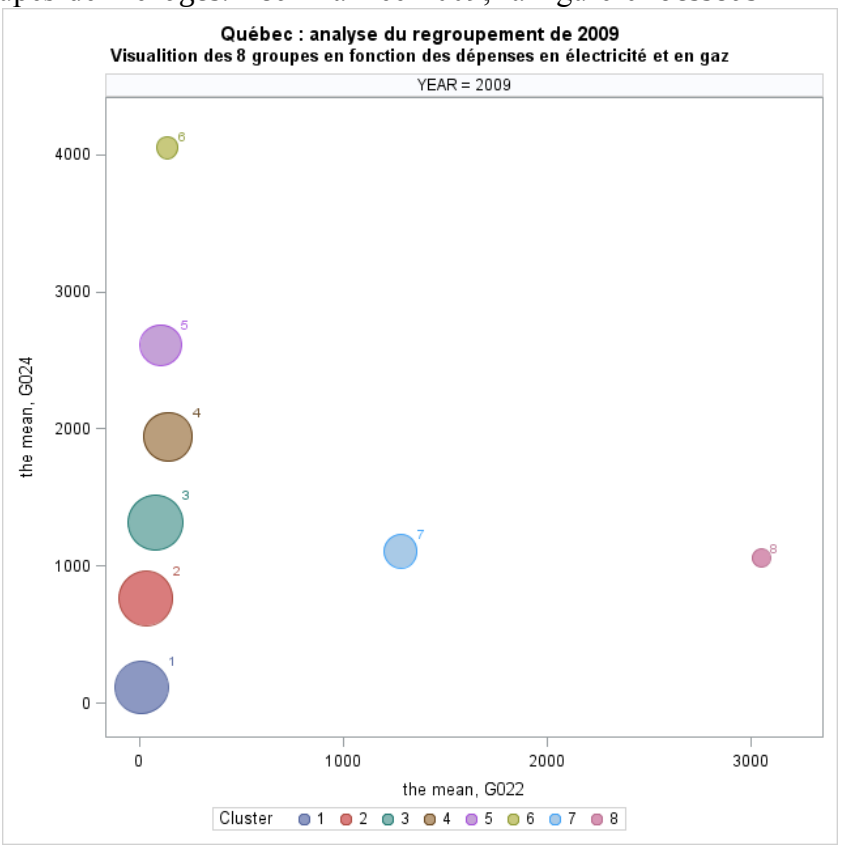
possible de réaliser une typologie des ménages en formant des groupes de ménages homogènes en matière de dépenses en énergie et de suivre l'évolution de ces groupes au cours de la période 1997-2009.

Méthodologie

Contrairement à de multiples études qui réalisent manuellement une segmentation des ménages puis analysent le profil de consommation en énergie de chaque segment, nous réalisons une segmentation automatique des ménages en fonction de leurs dépenses en énergie pour ensuite décrire les caractéristiques des ménages de chaque profil type identifié. La methodologie employée repose sur l'utilisation de techniques classiques de data mining (présentées dans la partie 3 du rapport) telles que les techniques de classification automatique ("clustering analysis"). Puis les groupes de ménages sont caractérisés et évalués selon deux critères principaux: les groupes doivent être interprétables et identifiables au cours du temps.

Résultats

L'analyse a permis d'identifier 8 groupes de ménages. Pour l'année 2009, la figure ci-dessous représente chacun des groupes de ménages relativement à leurs dépenses en électricté et en gaz. La taille de la bulle indique la taille que représente le groupe dans la population des ménages.

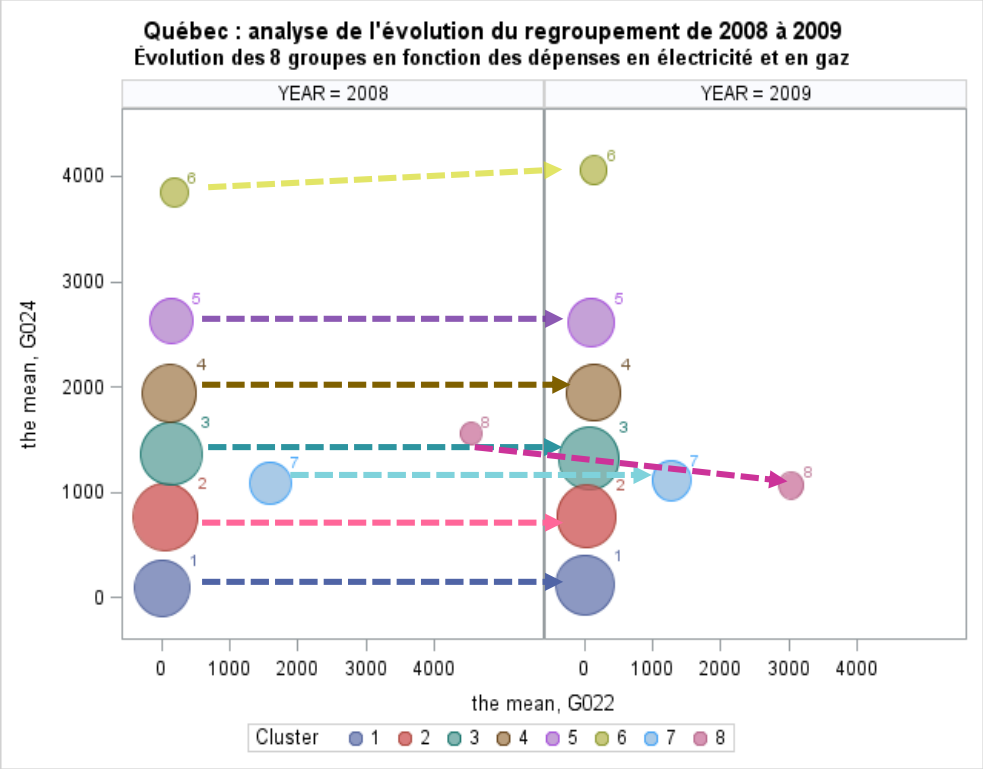


Les groupes de ménages 1 à 6 sont des profils de ménages qui ont presque exclusivement des dépenses en électricité. Le groupe 1, les « ***Électricité* - payeurs indirects** », sont des locataires qui ont les dépenses d'électricité incluses dans le loyer. Par conséquent les dépenses en électricité des ménages de ce

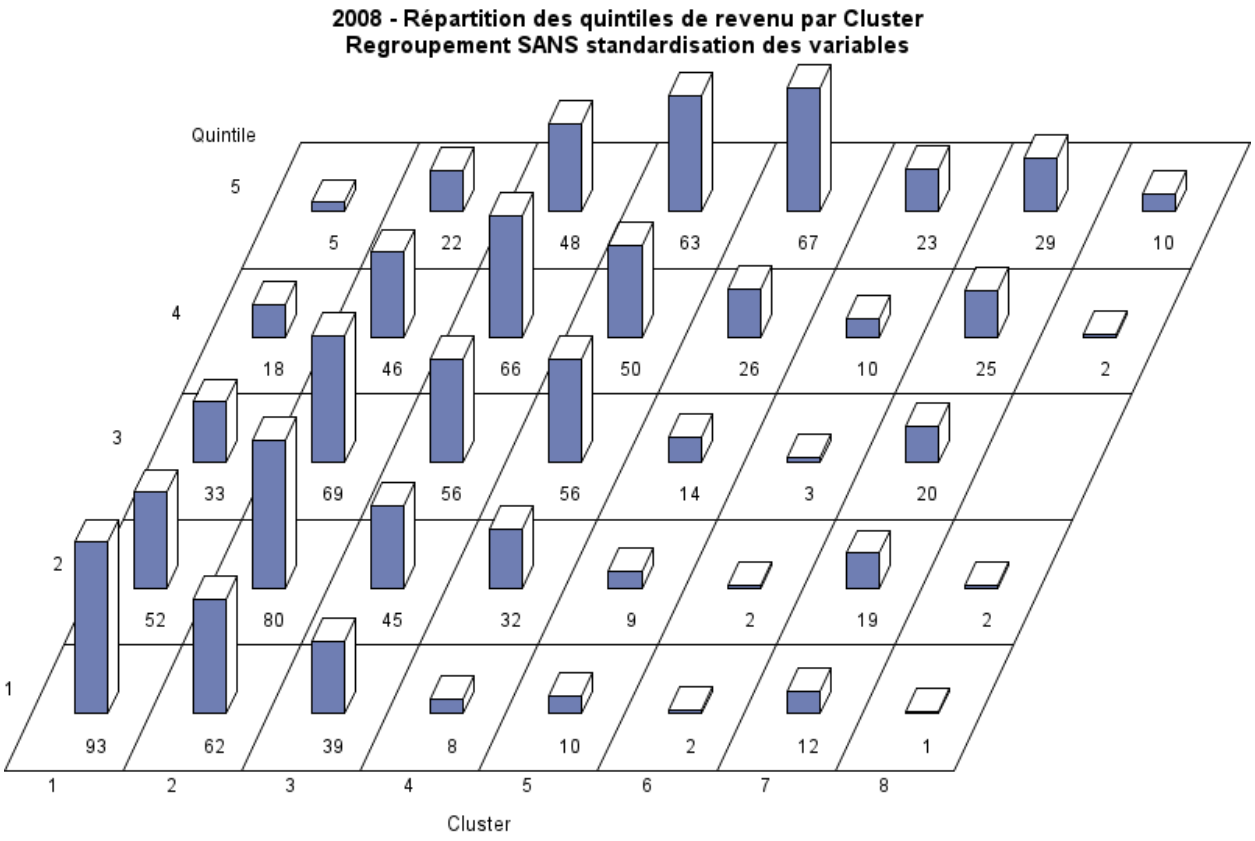
groupe sont proches de 0$. Puis, les groupes 2 à 5 ont successivement des dépenses en électricité de plus en plus importantes. Le groupe 2, les « ***Électricité* – poids mouches** » sont des ménages qui ont des dépenses en électricité inférieures à 1000$. Puis le groupe 3, les « ***Électricité* – poids plumes** », le groupe 4, les « ***Électricité* – poids moyens** » et enfin le groupe 5, les « ***Électricité* – poids lourds** » représentent des ménages qui ont des dépenses en électricité supérieures à 1000$ et inférieures à 3000$. Le groupe 6, quant à lui, représente des ménages dit « ***Électricité* – extrêmes** » qui ont des dépenses supérieures à 3000$.

Les groupes de ménages 7 et 8 représentent des ménages qui ont des dépenses en gaz plus importantes qu'en électricité. Le groupe 7, les « **E&Gaz – poids lourds** », ont des dépenses en gaz et en électricité supérieures à 1000$. Le groupe 8, les « **E&Gaz – extrêmes** » ont une dépense en gaz supérieure à 3000$.

Ces 8 groupes caractérisent 8 profils de dépenses en énergie et sont identifiables au cours du temps comme l'illustre la figure ci-dessous pour les années 2008 et 2009 :



Ce regroupement en 8 groupes permet de mettre en évidence que les groupes de quintiles ne permettent pas de rendre compte d'un profil de dépenses spécifiques en énergie. En effet, dans la figure ci-dessous, il est possible d'observer la distribution des cinq quintiles de revenu dans chacun des 8 groupes de profils de dépenses en énergie identifiés dans ce rapport pour l'année 2008.



Dans la partie 4.1 et 4.2 du rapport, ces groupes ainsi que leur évolution au cours de la période 1997-2009 sont présentés en détail. Le tableau de la page suivante présente néanmoins le ménage type pour chacun des groupes.

Le tableau ci-dessous caractérise les 8 profils types de dépenses en énergie des ménages pour la résidence principale de l'année 2008 au Québec, dans l'esprit des « personas[1] » en marketing.

N° groupe	**Taille du groupe**	**Zone urbaine/rurale**	**Type de logement**	**Nbr de pièces**	**Propriéataire / Locataire**	**Type de chauffage**	**Moyenne électricité**	**Moyenne gaz**	**Quintile de revenu majoritaire**
Le groupe	représente	habite en zone	dans un	de	il est	chauffage	dépense	dépense	appartient au
1, Payeurs Indirects	16%	Urbaine *(ruraux < 5%)*	Appart.	3.4	Locataire	Électri.	0$	0$	Q1, Q2
2, E-Poids mouches	**24%**	**Urbaine** *(ruraux < 15%)*	**Maison**	**4.4**	**Locataire**	**Électri.**	**760$**	**70$**	**Q2, Q1, Q3**
3, E-Poids plumes	**21%**	**Urbaine (ruraux =25%)**	**Maison**	**5.8**	**Propriétaire**	**Électri.**	**1360$**	**150$**	**Q3, Q4, Q2**
4, E-Poids moyens	16%	Urbaine *(ruraux < 20%)*	Maison	6.5	Propriétaire	Électri.	1930$	120$	Q5, Q4
5, E-Poids Lourds	**10%**	**Urbaine** *(ruraux < 10%)*	**Maison**	**7.3**	**Propriétaire**	**Électricité**	**2620$**	**150$**	**Q5**
6, E-Extrêmes	3%	Urbaine *(ruraux < 10%)*	Maison	8.0	Propriétaire	Électricité	3840$	190$	Q5
7, E&Gaz, Poids Lourds	**9%**	**Urbaine (ruraux =25%)**	**Maison**	**6.5**	**Propriétaire**	**Fournaise**	**1080$**	**1600$**	**Q5, Q4, Q3**
8, E&Gaz, Poids Extrêmes	1%	Urbaine *(ruraux < 10%)*	Maison	8.2	Propriétaire	Frounaise	1550$	4500$	Q5

[1] Un « persona » est une personne fictive qui représente un groupe cible, elle représente de façon simplifiée les caractéristiques du groupe. Dans ce rapport, le persona est un ménage fictif.

Notes quant à ce tableau synthétique :

- Il s'agit d'une description du profil type, « du profil moyen » du ménage de chacun des groupes. C'est-à-dire que si l'on tire au sort un ménage du groupe 1, les « Payeurs Indirects », il y a une forte probabilité (supérieure à 50%) qu'il s'agisse d'un ménage :
  - → vivant en zone urbaine
  - → dans un appartement de 3,4 pièces, se chauffant à l'électricité
  - → locataire du logement
  - → dépensant 0$ en facture d'électricité (incluse dans le loyer) et de gaz
  - → appartenant au premier quintile de revenu
- Les ménages du groupe 1 sont majoritairement des locataires et la dépense en électricité est incluse dans leur loyer ce qui fait qu'ils réalisent très peu de dépenses pour ce poste.
- Les ménages qui ont la plus forte probabilité d'habiter en zone rurale sont les ménages des groupes 3, 4 et 7.

1 Introduction

L'énergie est un enjeu primordial pour les sociétés humaines. De nos jours, les problématiques et questions liées à l'énergie sont nombreuses et font très souvent débats, notamment les surplus de production d'électricité au Québec ainsi que l'augmentation du prix de l'électricité.

Le débat sur ces questions peut s'établir à plusieurs niveaux : sur la véracité des faits, sur l'interprétation des faits, les conséquences prévisibles d'une situation donnée, et les choix ou compromis à faire quant à cette situation. Mais pour qu'une discussion sereine puisse avoir lieu sur une problématique donnée et que les esprits puissent arriver à un accord, il est essentiel que les différents acteurs, parties-prenantes au problème, s'entendent sur les faits, caractéristiques et les paramètres-clés de la situation. Dans bien des cas, la véritable difficulté ne se trouve pas dans la résolution du problème, mais dans l'identification et l'analyse de la problématique : le diagnostique juste, l'analyse fine du problème lui-même est le véritable défi. À la question de savoir ce qu'il ferait s'il lui restait seulement une heure pour sauver le monde, Albert Einstein est réputé avoir répondu « Je passerais 55 minutes à définir le problème, puis, seulement 5 minutes à le résoudre » (Basadur, Ellspermann, & Evans, 1994, p. 627).

En décembre 2013, au Québec, la chaire de gestion du secteur de l'énergie a été créée avec pour mission principale « d'augmenter les connaissances sur les enjeux liés à l'énergie dans une perspective de développement durable, d'optimisation et d'adéquation entre les sources d'énergie et les besoins de la société » ("Chaire de gestion du secteur de l'énergie," n.d.). Le travail présenté dans ce rapport a pour ambition de contribuer à cette mission, en approfondissant les connaissances qui ont trait aux dépenses en énergie des ménages québécois.

1.1 La problématique

Le **secteur de l'énergie** peut être représenté à travers le prisme du **marché** (dans une perspective économique), en analysant d'une part **l'offre** en énergie et ses déterminants, et d'autre part **la demande** en énergie et ses déterminants également. L'*une* des composantes principales de la **demande d'énergie** sur un territoire donné est **la consommation en énergie des « ménages ».**

La Figure 1-1 présente les principales composantes de la consommation d'énergie au Québec ("Chaire de gestion du secteur de l'énergie – État de l'énergie au Québec," n.d.), soit les secteurs : résidentiel, industriel, commercial et institutionnel, transports et agricole. Le secteur résidentiel compte pour près de 18% de la consommation totale au Québec en 2014.

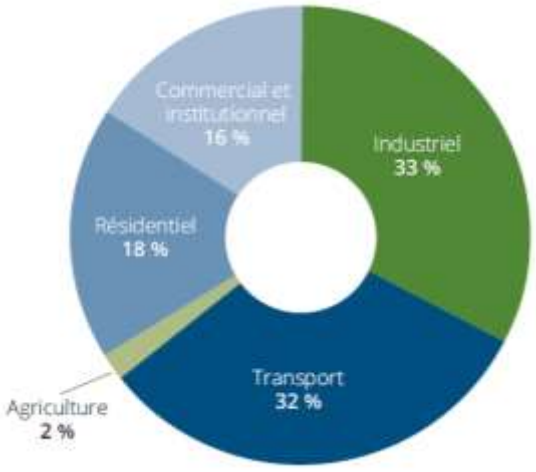


Figure 1-1 : Consommation d'énergie par secteur d'activité au Québec , 2012 (Source : État de l'énergie au Québec 2015)

La consommation du secteur résidentiel est le fait des ménages québecois pour faire chauffer leur logement (61% de la consommation en énergie), faire fonctionner leurs appareils ménagers (17%), faire chauffer l'eau ou encore pour l'éclairage et la climatisation en été comme l'illustre la Figure 1-2.

Le secteur résidentiel comptant pour près de 18% de la demande totale d'énergie au Québec (en 2014), il est indipsensable pour se représenter le secteur énergétique dans son ensemble d'avoir une représentation juste de cette composante de la demande. Depuis 1997, Statistique Canada réalise une enquête annuelle sur les dépenses des ménages canadiens (EDM), dont les dépenses en énergie, et produit un ensemble de statistiques quant à ces dernières. À partir de celles-ci, il est possible de produire plusieurs types d'analyse pour rendre compte de la réalité des ménages en termes de dépenses en énergie. Entre autres, une utilisation très pertinente de cette enquête est le regroupement des ménages par niveau de revenu. En regroupant les ménages en cinq (5) groupes, le premier groupe correspond au 20% des ménages qui ont les revenus les plus faibles et le cinquième groupe correspond au 20% des ménages qui ont les revenus les plus élevés, il est possible

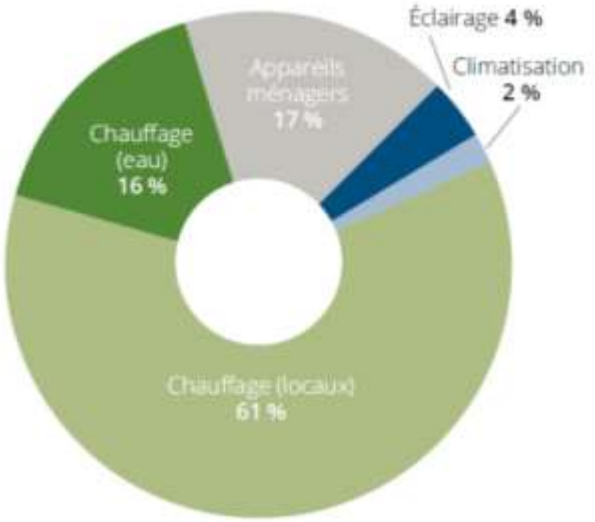


Figure 1-2 : Consommation d'énergie par type d'utilisation dans le secteur résidentiel au Québec, 2011 (Source : État de l'énergie au Québec 2015)

de décrire les dépenses en énergie de chaque groupe et d'effectuer des comparaisons entre eux. Ces groupes sont appelés des « quintiles ».

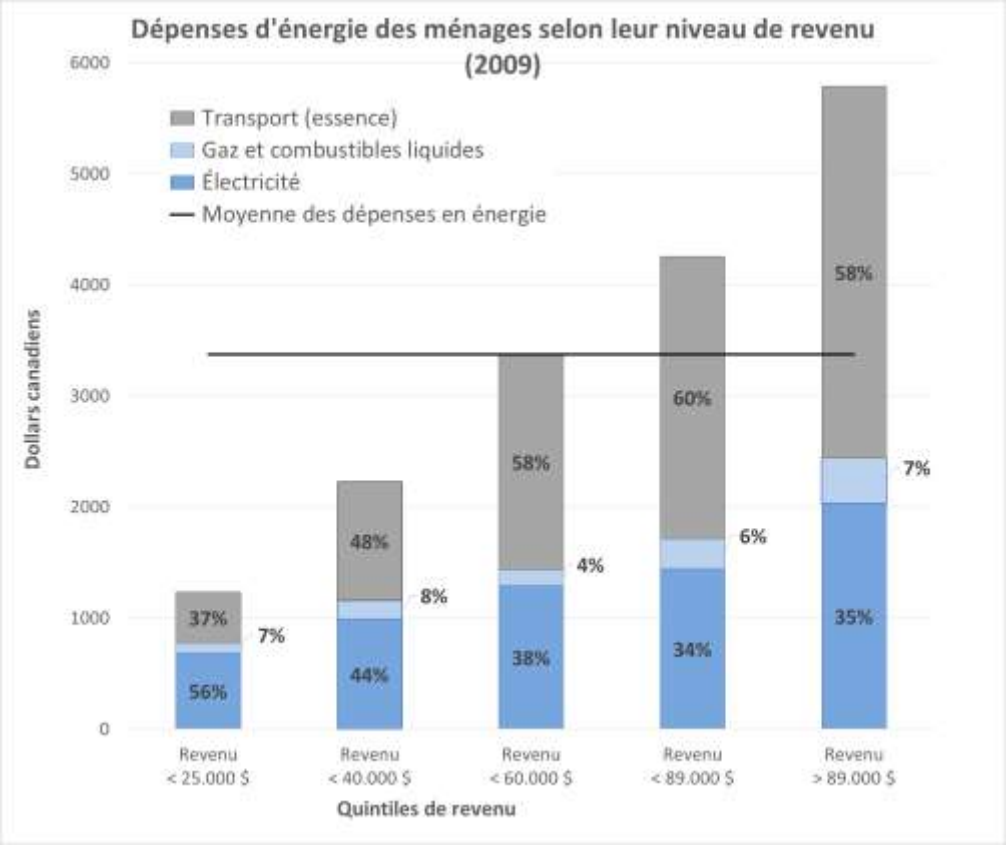


Figure 1-3 : Dépenses d'énergie des ménages selon leur niveau de revenu (2009). Source : Statistiques Canada (EDM). Sur ce graphique, les pourcentages indiquent la part de la dépense d'électricité, de gaz et de transport sur le total de la dépense en énergie.

La Figure 1-3 illustre ce type d'analyse. Pour l'année 2009, au Québec, le premier quintile (le groupe des ménages qui ont un revenu inférieur à 25 000 dollars canadiens) dépensait en moyenne 1230 dollars canadiens en énergie, dont 56% pour l'électricité, 7% pour le gaz et 37% en transport. En revanche, le cinquième quintile (le groupe des ménages qui ont un revenu supérieur à 89 000 dollars canadiens) dépensait en énergie 5785 dollars canadiens en moyenne, soit 35% pour l'électricité, 7% pour le gaz et 58% pour le transport. C'est une représentation de la réalité qui a le bénéfice d'être assez facilement intelligible : au lieu de parler des trois millions trois cent vingt deux milles[2] ménages québecois et de leurs dépenses, nous simplifions le réel en parlant de cinq (5) groupes de ménages de même taille et de la moyenne des dépenses en énergie de chacun d'eux.

[2] En 2009, le nombre estimé de ménages québecois à partir de l'enquête EDM de Statistiques Canada est 3.321.748 ménages.

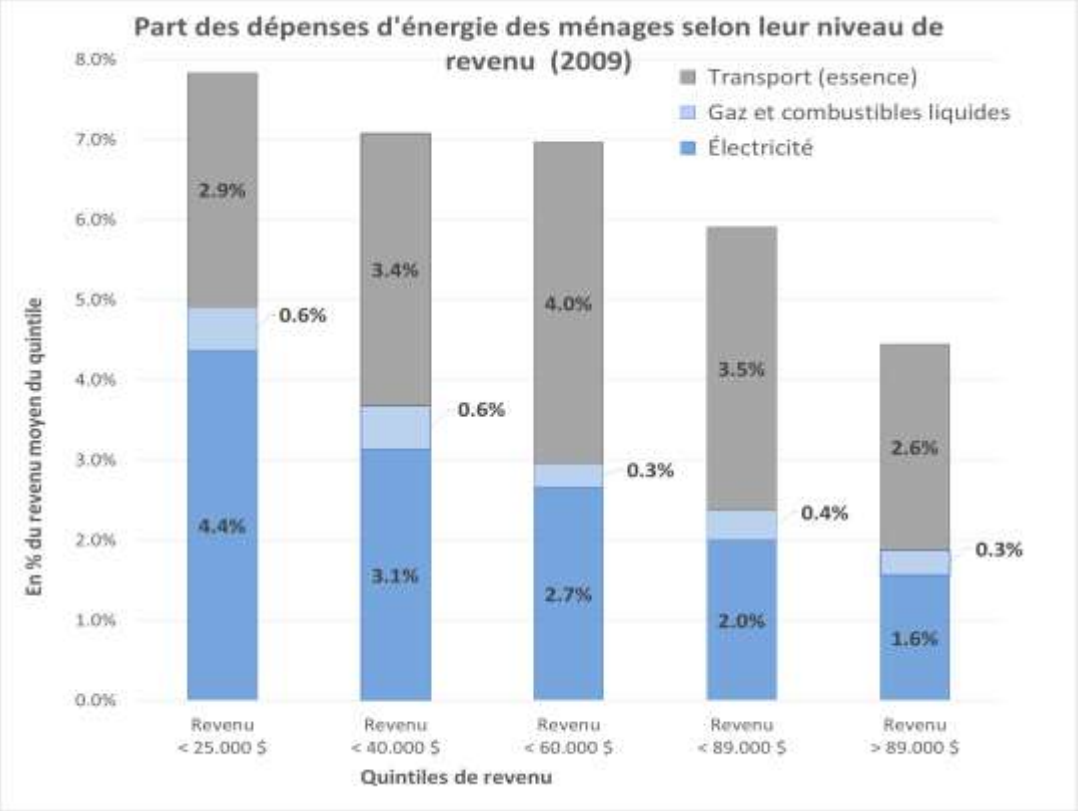


Figure 1-4 : Part des dépenses d'énergie des ménages selon leur niveau de revenu (2009). Source : Statistique Canada (EDM).

Ainsi, l'une des conclusions que l'on peut tirer de cette analyse est somme toute logique : en moyenne les ménages ont des dépenses en énergie d'autant plus importantes que leur revenu est important (c'est une tendance). Ce type d'analyse peut être très utile afin d'identifier quelles sont les catégories les plus exposées lors d'une hausse du prix de l'énergie, qu'il s'agisse du prix de l'essence, de l'électricité ou bien du gaz. En effet, en représentant la part occupée par les dépenses en énergie dans le revenu des ménages par quintile, il est possible d'identifier le groupe de ménages les plus sensibles à une augmentation du prix de l'énergie.

C'est ce qu'illustre la Figure 1-4. La part des dépenses d'énergie pèse proportionnellement plus sur les ménages à bas revenus que sur les ménages du cinquième quintile. De ce point de vue, les dépenses d'énergie des ménages peuvent être qualifiées de régressives. C'est-à-dire que les ménages à faible revenu souffriront particulièrement d'une hausse du prix de l'énergie (prix de l'essence, de l'électricité ou bien du gaz) ou de conditions climatiques extraordinaires qui exigent

des ménages une consommation d'énergie supérieure aux années classiques (cas d'un hiver particulièrement froid au Québec).

Cette représentation, si elle a le bénéfice d'être intelligible en simplifiant le réel, a aussi le défaut d'être simpliste et de ne pas rendre compte de la particularité de certains ménages (groupes de ménages). Deux critiques peuvent être formulées à cette représentation. La première est que cette représentation donne l'illusion d'une réalité qui n'existe que dans cette Figure 1-3. La seconde est qu'une partie de la réalité n'est pas retranscrite au travers de cette figure.

Par exemple, dans le premier quintile, près de 32% des ménages n'ont pas de dépense en électricité, soit plus d'un tiers des ménages (voir Figure 1-5). Or, la Figure 1-3 rend compte d'une dépense moyenne en électricité de 690 dollars canadiens pour le premier quintile. Au premier coup d'œil, le lecteur est loin de se douter qu'un tiers des ménages du premier quintile n'a en fait aucune dépense en élecricité. C'est l'une des limites de cette représentation des dépenses en énergie par quintile de revenu (première critique). La moyenne pour un quintile peut très bien masquer des réalités très diverses : au sein d'un même quintile, un premier ensemble de ménages pourraient dépenser très peu en électricité et un second ensemble de ménages pourrait quant à lui dépenser beaucoup plus (deux cas extrêmes). La moyenne des dépenses en électricité pour ces deux ensembles de ménages ne rend ni compte du premier, ni du second, mais d'un troisième ensemble de ménages qui aurait une dépense moyenne, en électricité, intermédiaire entre l'ensemble 1 et 2. Le seul problème, c'est que ce troisième ensemble n'existe pas. La représentation par quintile procure l'illusion d'une telle réalité.

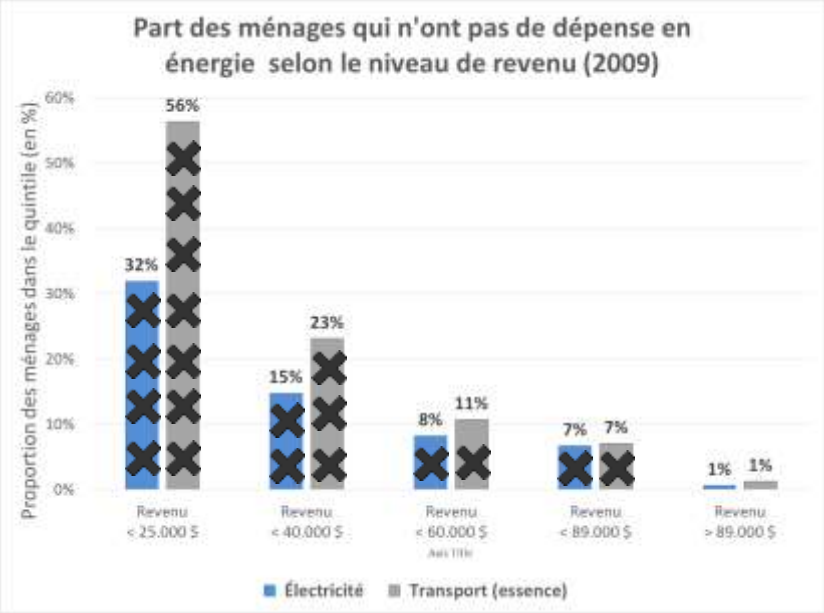


Figure 1-5 : Part des ménages qui n'ont pas de dépense en énergie selon le niveau de revenu (2009). Source : Statistique Canada (EDM).

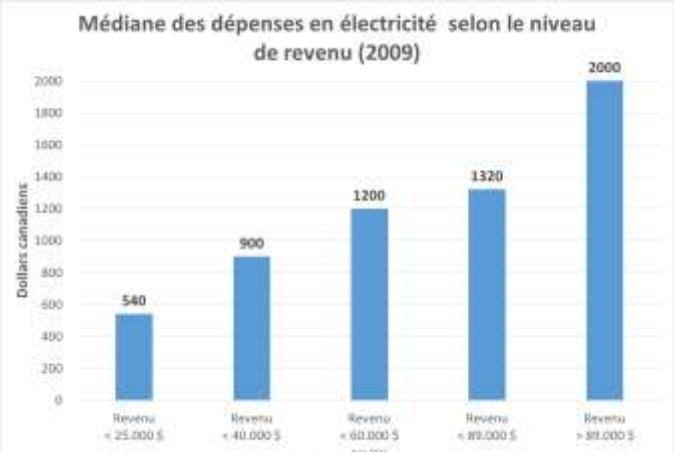


Figure 1-6 : Médiane des dépenses en électricité selon le niveau de revenu (2009). Source : Statistique Canada (EDM).

Une façon de pallier ce problème est d'analyser la médiane (voir Figure 1-6), plutôt que la moyenne, par quintile. Celle-ci nous apprend que pour le premier quintile, 50% des ménages dépensent moins de 540 dollars canadiens pour l'élecricité. Mais cette analyse ne laisse pas transparaître non plus que près d'un tiers des ménages du premier quintile n'ont aucune dépense en électricité.

De plus, le regroupement des ménages par quintile de revenu ne permet pas de rendre compte de groupes de ménages qui auraient des dépenses en énergie similaires et pourtant des revenus très différents (deuxième critique).

Comment dépasser ce problème ? Quels sont les groupes de ménages qui dépensent beaucoup en électricité au Québec, quelque soit leur revenu ? Et quels sont leurs caractéristiques ? Quels sont les profils types des ménages québecois en termes de dépenses en énergie ? Quelle est leur évolution au cours du temps ? Afin de contribuer au développement des connaissances dans le secteur de l'énergie, le travail présenté dans ce rapport propose une typologie des ménages canadiens, et plus particulièrement québecois, pour rendre compte de la variabilité des profils de ménages québecois en termes de dépenses en énergie.

1.2 Les objectifs et le plan du rapport

L'objectif de ce rapport est de construire une typologie des ménages québecois qui permettent de rendre compte de leurs dépenses en énergie. En utlisant l'enquête annuelle réalisée par Statistiques Canada sur les dépenses des ménages (EDM), il sera proposé un portrait des différents profils types de ménages québecois en termes de dépenses d'énergie entre 1997 et 2009. Les dépenses en énergie des ménages incluent aussi bien les dépenses pour le logement (énergie pour le chauffage du logement, de l'eau, pour l'utilisation des appareils électroménagers, la climatisation) que les dépenses pour le transport (dépense en essence).

La suite du présent rapport expose les données utilisées (EDM[3], de 1997 à 2009), la méthodologie suivie pour construire le portrait des ménages québecois et une analyse de l'évolution de ce portrait entre 1997 et 2009. Cette description des portraits des ménages québecois a pour ambition de permettre une meilleure définition des problèmes, d'initier une réflexion constructive des différents acteurs concernés par les enjeux énergétiques au Québec.

[3] EDM : Enquête sur les Dépenses des Ménages. « Le but principal de l'enquête est d'obtenir des renseignements détaillés sur les dépenses des ménages ainsi que de l'information limitée sur les caractéristiques du logement et l'équipement ménager » ("Enquête sur les dépenses des ménages (EDM)," n.d.).

1.3 Revue de littérature

Afin de présenter les articles qui proposent une étude des dépenses en énergie des ménages, il importe de présenter quelques paramètres afin de comprendre les différentes approches et motivations qui ont été recensées dans cette revue de littérature.

- **Dépense et consommation :** l'analyse peut porter soit sur la dépense ou bien sur la consommation d'énergie. Il est important d'indiquer que l'objet de ces études est soit la dépense en énergie des ménages (exprimée sous la forme de la valeur quantitative monétaire) ou bien la consommation d'énergie des ménages (exprimée en quantité d'énergie, en kWh par exmple). Dans chacun des cas, il s'agit de décrire, caractériser et de comprendre la dépense ou consommation d'énergie des ménages.
- **L'horizon temporel :** une étude peut porter aussi bien sur une année que sur une période. Par exemple, dans notre cas, l'étude menée porte sur la période 1997 à 2009, ce qui a pour conséquence d'inclure une anlayse comparative entre année ou du moins un intérêt pour les tendances et les évolutions au cours de la période.
- **Le grain des données :** la granularité des données conditionne les analyses possibles. Il peut s'agir de données aggrégées pour une population, de micro-données pour une étude précise (sondage ponctuel, représentatif d'un groupe précis), de micro-données de sondage annuel, visant à rendre compte d'une population d'un pays, de données de compteur intelligent qui donne la consommation d'électricité à chaque heure pour un ménage donné.
- **Le grain du regroupement** : certaines études dressent un portrait des ménages pour une région du monde en s'intéressant à un groupe de ménages précis comme celui des ménages des pays occidentaux dans son ensemble. D'autres articles s'intéressent plus particulièrement aux ménages d'un seul pays. Enfin, d'autres études portent sur un groupe de ménages particuliers, comme les ruraux ou bien les latinos.
- **Segmentation manuelle ou segmentation automatique :** il important de faire la distinction entre ces deux techniques de segmentation. Les différences entre ces deux segmentations sont présentées dans la partie 1.3.2.

Les articles sélectionnés pour cette revue de littérature permettent de dégager 2 catégories principales d'analyse. La première s'attache à identifier et quantifier les déterminants de la consommation/dépenses en énergie des ménages. La seconde réalise des segmentations de

ménages afin d'identifier des groupes de ménages correspondant à différents profils de consommation/dépenses en énergie (identification d'une structure).

1.3.1 Les déterminants/facteurs de la consommation (dépense) en énergie des ménages

Une première série d'articles s'intéressent aux déterminants de la consommation (dépense) en énergie des ménages

1.3.1.1 Déterminants de la consommation (dépenses en énergie)

Plusieurs recherches sont menées avec pour objectif principal d'identifier les déterminants de la consommation d'énergie. Ceux-ci peuvent être décrits comme comportementaux (le comportement des ménages), psychologiques (les croyances des ménages, l'attitude), physico-techniques (le type de logement, l'âge du logement, le type de chauffage…), socio-économiques (le revenu, la taille, l'âge du ménage…), l'équipement électroménager (sèche linge, climatiseur,

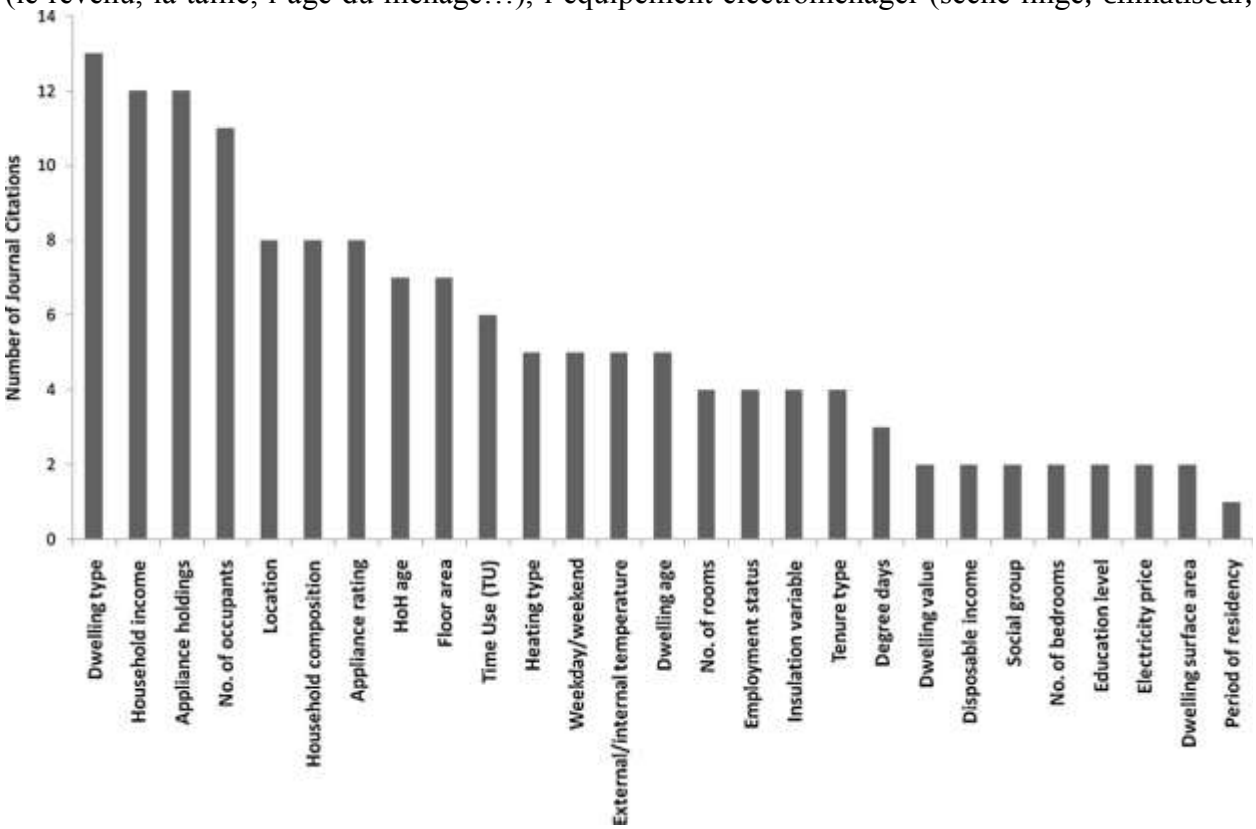


Figure 1-7 : Caratéristiques du logement et du ménage qui influencent la consommation d'électricité résidentielle. (McLoughlin, Duffy, & Conlon, 2012)

chauffe-eau), les conditions climatiques. À chaque fois, ces études tentent de mesurer (et de quantifier) l'effet de quelques-uns de ces facteurs sur le niveau de consommation d'énergie (*ou sur le changement du niveau de consommation*). Il s'agit de savoir quel facteur est significatif et, lorsqu'il l'est, de mesurer son effet. La méthodologie utilisée est dans la plupart des cas l'estimation des paramètres d'un modèle de regression linéaire. C'est de loin le type d'analyse le plus répandu.

Le tableau ci-dessous présente un échantillon d'articles qui illustrent cette approche.

Titre de l'article	Analyse des facteurs qui influencent :	Données de	Auteurs, date
« A study of domestic energy usage patterns in Hong Kong »	Le niveau de consommation d'énergie des ménages, aussi bien pour l'électricité que pour le gaz. Présentation du « electricity loading pattern ».	Hong Kong	(Tso & Yau, 2003)
« Behavioural, physical and socio-economic factors in household cooling energy consumption »	La demande d'énergie des ménages pour le refroidissement (climatisation).	E.U. (RECS[4], 2001).	(Yun & Steemers, 2011)
« Building and household X-factors and energy consumption at the residential sector. A structural equation analysis of the effects of household and building characteristics on the annual energy consumption of US residential buildings »	La consommation d'énergie des ménages.	E.U. (RECS, 2009)	(Estiri, 2014)
« Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: An Irish case study »	Le type (« pattern ») de consommation d'électricité des ménages.	Irlande (3941 ménages, 2009)	(McLoughlin, Duffy, & Conlon, 2012)
« U.S. household energy consumption and intensity trends: A decomposition approach »	Les changements dans le niveau de consommation d'énergie des ménages, en utilisant une approche par décomposition.	E.U. (RECS, 1980-2005)	(Hojjati & Wade, 2012)
« Demographic Determinants of Household Energy Use in the United States »	La consommation d'énergie des ménages, en se concentrant sur les facteurs démographiques.	E.U. (RECS et RTECS, 2000)	(O'Neill & Chen, 2002)
« Exploring variance in residential electricity consumption: Household features and building properties »	La consommation d'électricité des ménages. Pour les facteurs explicatifs, en plus des facteurs classiques, les auteurs	Suède, données de consommation et de sondage.	(Bartusch, Odlare, Wallin, & Wester, 2012)

[4] RECS pour « Residential Energy Consumption Survey ». Le sondage sur la consommation résidentiel d'énergie (RECS) est un sondage national sur l'énergie et les ménages mené tous les trois ans par l'administration de l'information en énergie (« Energy Information Administration », EIA) par le département américain de l'énergie. Le premier sondage a été mené en 1978.

Titre de l'article	Analyse des facteurs qui influencent :	Données de	Auteurs, date
	ont utilisé une variable géographique, distinguant quatre régions. Celle-ci a un rôle significatif.		
« How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings? »	L'utilisation de l'énergie et les changements dans l'utilisation de l'énergie. Les facteurs étudiés sont d'ordre sociodémographiques (influençant la consommation) et psychologiques (à l'origine d'un changement dans la consommation).	Pays-Bas, suivi de 189 ménages sur 5 mois.	(Abrahamse & Steg, 2009)
« Residential energy use and conservation: Economics and demographics »	Le niveau de consommation d'électricité et de gaz des ménages, en comparant l'influence des facteurs physiques (caractéristiques du logement) avec l'influence des caractéristiques propres du ménage selon le type d'énergie.	Pays-Bas, fusion de plusieurs fichiers de données (NVM pour les caractéristiques des logements – 2008-2009, CBS pour la consommation électrique de 2007 et les caractéristiques des ménages)	(Brounen, Kok, & Quigley, 2012)
« Residential energy use - Analysis of disaggregate data »	La consommation en énergie des ménages. Les facteurs principaux sont : le prix, l'année de la maison, la surface. et « heating degree days ».	E.U., National Interim Energy Consumptoin (1978, 1979)	(Hirst, Goeltz, & Carney, 1982)
« The impact of urban form on U.S. residential energy use »	Analyse de l'influence de l'urbanisation sur l'utilisation de l'énergie par les ménages en plus des autres facteurs. Utilisation d'une variable « urban form ».	E.U., (RECS 2001, Census 2000, et autres sources de données)	(Ewing & Rong, 2008)
« The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings »	Cet article propose une synthèse de tous les facteurs étudiés dans un grand nombre d'études.	« Littérature Review »	(Jones, Fuertes, & Lomas, 2015)
« Understanding the spectrum of residential energy consumption: A quantile regression approach »	La consommation d'énergie, par quartile (analyse des facteurs sur la distribution de la consommation d'énergie et non sur la moyenne)	E.U., (RECS, 2005)	(Kaza, 2010)

1.3.1.2 Revenu et consommation (dépense) en énergie

Parmi les facteurs étudiés, souvent le revenu du ménage fait controverse. Beaucoup d'articles s'intéressent aux relations entre le revenu et la consommation (dépense) en énergie des ménages. Par exemple, dans l'article intitulé « The role of income in energy consumption behaviour: Evidence from French households data », les auteurs s'intéressent non seulement à l'influence du revenu du ménage sur sa consommation d'énergie mais aussi sur son comportement (le ménage investi ou non pour un équipement de chauffage plus efficient, par exemple).

Ci-dessous, le tableau présente à titre d'illustration un échantillon d'articles qui a cet angle d'analyse.

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Household energy consumption versus income and relative standard of living: A panel approach »	La relation entre le revenu et la consommation d'électricité. D'après l'étude, le revenu et la consommation électrique ne sont pas cointégrés.	10 pays, données extraites du site internet de IEA, pour la période 1971-2001	(Joyeux & Ripple, 2007)
« How many times again will we examine the energy-income nexus using a limited range of traditional econometric tools? »	La relation entre consommation d'énergie et la croissance économique.	Critique	(Karanfil, 2009)
« The role of income in energy consumption behaviour: Evidence from French households data »	Mesurer de façon quantitative l'impact du revenu sur la consommation d'énergie dans le secteur résidentiel et le transport.	France, sondage sur mesure, encadré par SOFRES, 3000 ménages, 2009.	(Cayla, Maizi, & Marchand, 2011)

1.3.1.3 Locataire/propriétaire : impact sur la consommation d'énergie

Certaines études s'intéressent exclusivement aux différences de comportement de dépense/consommation en énergie entre les propriétaires et les locataires d'un logement. Ces derniers, les locataires, pour une part, ne payent pas directement l'énergie qu'ils utilisent car son coût est inclu dans le loyer de leur logement. C'est le cas notamment de l'étude intitulée « Split incentives and energy efficiency in Canadian multi-family dwellings » qui quantifie la différence d'intensité énergétique entre les payeurs directs et les payeurs indirects.

Ci-dessous, le tableau présente à titre d'illustration un échantillon d'articles qui focalisent leur analyse sur cette relation :

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Split incentives and energy efficiency in Canadian multi-family dwellings »	Différence d'intensité énergétique entre les payeurs directs et les payeurs indirects.	Canada, EDM, 2003	(Maruejols & Young, 2011)
« Energy use by apartment tenants when landlords pay for utilities »	Explication de la « sur-consommation » des locataires qui ne payent pas directement leur consommation d'énergie.	E.U., RECS et American Housing Survey.	(Levinson & Niemann, 2004)

1.3.2 Segmentation des ménages en tant que consommateur d'énergie

Une deuxième série d'article propose diverses segmentations de ménages selon différents critères. Les approches varient à plusieurs égards : choix des variables pour établir la segmentation, segmentation manuelle ou automatique, les motivations de la formation de groupes de ménages, la richesse du regroupement.

En majorité, les segmentations des ménages effectuées (dans les articles de recherche universitaire) quant à la question de la consommation/dépenses en énergie sont des segmentations dites manuelles. C'est-à-dire que les groupes sont choisis en fonction d'une variable, l'âge ou le revenu par exemple, et d'une règle pour segmenter la population des ménages en groupes … Par exemple, la règle suivante permet de diviser la population des ménages en deux groupes selon l'âge de la personne de référence du ménage : tous les ménages dont la personne de référence à moins de 50 ans appartiennent au groupe *A*, alors que les autres appartiennent au groupe *B*. De même, nous pouvons constituer des groupes de ménages suivant leur niveau de revenu (comme les quintiles par exemple), ou encore, regrouper les ménages en fonction du type d'habitation. Dans chacun de ces cas, une variable est sélectionnée et une règle permet de déterminer manuellement à quel groupe appartient le ménage. Un ménage appartient à un seul groupe, et la segmentation est relativement simple à faire. La difficulté se trouve dans le choix de la variable et de la règle. Pourquoi choisir l'âge pour segmenter le groupe des ménages, pourquoi choisir la règle du « 50 ans » ? Dans le choix des variables, la première catégorie d'articles explicitée ci-dessus, dans la partie 1.3.1, peut être très utile. En effet, il peut être pertinent de choisir les déterminants majeurs

de la consommation des ménages pour former des groupes de ménages semblables en termes de consommation/dépenses en énergie. Par exemple, d'après la partie 1.3.1, les variables telles que la taille du logement ou bien le nombre de personnes qui composent le ménage peuvent être très pertinents.

Dans le cas d'une segmentation automatique, le choix des variables est toujours cruciale, car c'est à partir de celles-ci que les groupes seront formés. En revanche, il ne s'agit plus de l'application d'une règle manuelle pour déterminer à quel groupe appartient le ménage. Il convient de choisir un critère, habituellement la ressemblance des ménages, puis d'une mesure pour quanitfier la ressemblance entre les ménages. Enfin, c'est un algorithme qui effectue le regroupement des ménages. Ces éléments sont présentés en détail dans la présentation de la méthodologie (partie 3). L'arrivée des compteurs intelligents dans le domicile des ménages offre des possibilités en termes de regroupement inenvisageables il y a encore 15 ans, avec des utilisations variées.

1.3.2.1 Segmentation manuelle des ménages en fonction du revenu du ménage

Plusieurs articles s'intéressent à la relation entre le niveau de revenu et niveau de consommation. Pour cela, ils constituent des groupes de ménages en fonction du revenu afin d'établir des comparaisons dans la consommation/dépenses en énergie entre ces différents groupes.

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Energy Consumption and Expenditures by Low-Income Customers »	Cet article pose la question de savoir si les dépenses en énergie sont régressives, à savoir si une hausse des coûts de l'énergie pourrait pénaliser (de façon non proportionnelle) plus fortement les faibles revenus.	E.U., RECS 1997 et Consumer Expenditure Survey (1999-2000)	(Colton, 2002)
« On the relation between the energy and social characteristics of the residential sector »	L'échantillon des ménages interrogés a été divisé en 7 groupes de revenus. Puis, les déterminants de la consommation en énergie sont analysés, et quantifiés par groupe de revenus.	Athènes, sondage portant sur 1110 ménages en 2004.	(Santamouris et al., 2007)

1.3.2.2 Segmentation manuelle des ménages en fonction du style de vie

Parmi les critères utilisés pour réaliser une segmentation des ménages, certaines études utilisent le style de vie. La première étape consiste à identifier des types de styles de vie, puis à segmenter les ménages en fonction de ce style de vie. Une méthode souvent utilisée pour identifier les styles de vie est l'utilisation d'un sondage AIO : « Activities, Interests and Opinions » (Plummer, 1974). L'analyse factorielle peut être utilisée pour identifier les styles de vies des ménages comme cela est le cas dans l'article « Lifestyle factors in U.S. residential electricity consumption » (Sanquist, Orr, Shui, & Bittner, 2012). La deuxième étape consiste alors à analyser les relations entre le style de vie et le niveau de consommation en énergie des ménages appartenant à ce groupe. C'est souvent sur cette deuxième étape, l'établissement d'un lien entre le style de vie (et donc les groupes de style de vie identifiés) et la consommation d'énergie que les analyses mériteraient d'être approfondies. Ce lien reste souvent vague et finalement peu significatif. Certains argumentent que c'est pour la bonne cause, car il n'existe pas de différence significative entre la consommation des différents groupes de styles de vie des ménages. Par contre, le style de vie d'un ménage pourrait expliquer les particularités de la consommation d'électricité au cours d'une journée « residential laod profiles of housholds », recherches possibles désormais que les logements des ménages sont dotés d'un compteur intelligent. Voici un échantillon d'articles qui effectue ce type d'anlayse :

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Green consumption or sustainable lifestyles? Identifying the sustainable consumer »	Cet article identifie des styles de vie des ménages et analyse dans quelle mesure ceux-ci ont un impact sur la consommation d'énergie ou du moins pourraient être qualifiés de style de vie « soutenable ».	Grande-Bretagne, Devon, sondage auprès de 1600 ménages, 2005	(Gilg, Barr, & Ford, 2005)
« On the relation between the energy and social characteristics of the residential sector »	Utilisation au sens large du concept de « Consumer Life Style, CLA » pour aborder les questions de la consommation d'énergie et des émissions de CO2. Une partie du rapport présente les « Heavy Energy Users » (voir p 49, partie 5.5) en effectuant des comparaisons entre des ménages qui ont un même niveau de consommation et pourtant des profils très différents (analyse la plus proche de ce qui est fait dans ce rapport).	E.U., RECS 1999.	(Santamouris et al., 2007)

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Lifestyle factors in U.S. residential electricity consumption »	Cet article identifie 5 styles de vie sur la base de critères sociaux et comportementaux qui sont associés à des utilisations type d'utilisation de l'énergie. Puis, le style de vie, avec d'autres variables sont utilisés comme variables explicatives (facteurs) du niveau de consommation d'énergie.	E.U., RECS 2005, soit 4382 ménages.	(Sanquist et al., 2012)
« Linking life-styles and energy use: a matter of time? »	En prenant pour point de départ de leur analyse que pour un groupe de ménages ayant des logements similaires, un lieu de vie identique… la seule explication des différences de niveau de consommation d'électricité observé ne peut résider exclusivement dans les disparités de revenu mais aussi dans le style de vie.	Données américaines, suédoises… (très variées).	(Schipper, Bartlett, Hawk, & Vine, 1989)

1.3.2.3 Segmentation manuelle selon des critères socio-démographiques (et économiques)

Plusieurs études proposent des comparaisons entre groupes de ménages suivant un critère socio-démographique. Ce choix est en fait guidé par le fait qu' un facteur socio-démographique donné est considéré comme ayant significativement un impact sur le niveau de consommation d'énergie des ménages. Pour citer une liste de quelques-uns de ces facteurs/critères de segmentation : nombre de personnes qui compose le ménage, la composition du ménage (nombre d'enfants, nombre d'adultes), le statut d'emploi (employé ou non), le niveau d'éducation, le sexe, ménages urbains ou ruraux, le statut marital.

Dans la première partie de la « Chicago Fed Letter », de juin 2006, David B. Cashin et Leslie McGranahan s'intéressent à l'évolution de la part que représente le poste énergie dans la dépense totale des ménages américains sur la période 1982-2004 (Cashin & McGranahan, 2006). Ils utilisent les données des sondages « Consumer Expenditure Survey », CES, de 1982 à 2004. Leur principale conclusion est la suivante : de 1990 à 2004, le poste de dépense énergie représentait 7% des dépenses des méanges américains, contrairement à 11% au début des années 1980 (9% en moyenne sur les années 80). En 2005, la part des dépenses en énergie des ménages américains dans le budget représente 8.5%. Ces conclusions sont à l'échelle de tous les ménages américains (analyse globale).

Dans la deuxième partie de cette lettre, les auteurs s'intéressent aux différences de dépenses en énergie entre différents groupes de ménages américains ainsi qu'à la sensibilité de ces groupes de ménages quant à l'évolution des prix des énergies (énergies pour l'habitaiton, pour la voiture…). La question principale à laquelle les auteurs souhaitent répondre est la suivante: « This allows us to ask which groups are more and less sensitive to changes in energy prices given their energy consumption patterns. We look at three groups ». Il est important de préciser que les groupes ont été construits manuellement, en utilisant des critères socio-démographiques comme l'âge, le niveau d'étude, des critères économiques comme les quartiles de revenus, niveau de pauvreté et enfin travailleur pauvre et non pauvre. Il s'agit donc d'une segmentation manuelle, où chaque « paire » de groupes de ménages est analysée selon le critère de regroupement choisi :

4. Average annual expenditure share, 1982–2004

	Total energy	Gasoline and motor oil	Electricity	Natural gas	Fuel oil
Urban population	8.0	3.8	2.8	1.1	0.3
Elderly	8.1	2.9	3.2	1.5	0.6
Non-elderly	7.9	3.9	2.7	1.1	0.3
Without high school diploma	9.6	3.9	3.6	1.6	0.5
High school graduate	9.1	4.2	3.2	1.3	0.4
Some college	8.1	4.1	2.7	1.1	0.3
College graduate	6.8	3.3	2.3	0.9	0.3
Bottom income quartile	9.2	3.8	3.5	1.5	0.4
Second income quartile	8.9	4.1	3.1	1.3	0.4
Third income quartile	8.2	4.1	2.7	1.1	0.3
Top income quartile	6.7	3.3	2.2	0.9	0.3
Nonpoor	7.8	3.8	2.7	1.1	0.3
Poor	9.1	3.7	3.6	1.5	0.3
Working poor	9.6	4.3	3.5	1.4	0.3
Nonworking poor	8.8	3.3	3.7	1.6	0.3

Note: All values are in percent.
Source: Authors' calculations based on data from the *Consumer Expenditure Survey*, 1982–2004.

Figure 1-8 : La part (moyenne) des dépenses en énergie par groupe de ménages sur la périodre 1982-2004 (Cashin & McGranahan, 2006).

Critères	Groupes
Age	• groupe de ménages âgés • groupe de ménages non âgés
Niveau d'études	• groupe de ménages sans diplôme d'étude supérieure • groupe de ménages avec diplôme d'étude supérieure • groupe de ménages – étudiant • groupe de ménages ayant atteint le niveau collégial
Revenu	• groupe de ménages du 1er quartile de revenu • groupe de ménages du 2ème quartile de revenu • groupe de ménages du 3ème quartile de revenu • groupe de ménages du 4ème quartile de revenu

Niveau de pauvreté	• groupe de ménages pauvres
	  ○ groupes des travailleurs pauvres
	  ○ groupe des non-travailleurs pauvres
	• groupe de ménages non pauvres

De cette analyse, les auteurs dégagent des caractéristiques de dépenses de chacun des groupes, constantes dans le temps. Ainsi, il est possible de déterminer les groupes les plus à risque par une hausse du prix d'une ou plusieurs sources d'énergie.

Ci-dessous, un tableau qui présente les articles produisant des analyse similaires :

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Household energy expenditures, 1982–2005 »	La dépense des ménages. Une brève présentation de cet article est proposée ci-dessus, étant un des rares articles à traiter de la dépense et non de la consommation d'énergie.	E.U., CES (Consumer Expenditure Survey), de 1982 à 2004.	(Cashin & McGranahan, 2006)

1.3.2.4 Segmentation manuelle sur la base de critère éthnique (minorité)

Certaines études s'intéressent aux différences de consommation entre des groupes de ménages formés sur des critères ethniques, comme par exemple les groupes de ménages latino et non latino, ou bien le groupe de ménages majoritaires dans un pays donné comparé aux groupes de ménages minoritaires. L'hypothèse sous jacente ainsi testée est que l'appartenance éthnique, ou du moins à un groupe social caractérisé comme « white, black, and Latino » pour reprendre les termes de l'article intitulé « Residential energy consumption across different population groups: comparative analysis for Latino and non-Latino households in USA » (Poyer, Henderson, & Teotia, 1997) permet d'expliquer la consommation d'énergie des ménages. Ainsi dans ces études, les ménages sont segmentés selon des critères « éthnique ».

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Residential energy consumption across different population groups: comparative analysis for Latino and non-Latino households in USA»	Élaboration d'un modèle pour expliquer la consommation et les dépenses d'essence, d'électricité et de gaz entre les ménages latinos et non latinos, et plus largement entre différents groupes sociaux.	E.U., RECS pour plusieurs années.	(Poyer et al., 1997)
« Residential energy demand : additional empirical evidence by minority household type »	Cette étude cherche à déterminer l'élasticité de la demande d'énergie pour des groupes de ménages minoritaires, comme les Noirs, les Hispaniques comparés aux groupes dit majoritaires.	E.U., RECS de 1980, 1982 et 1987.	(Poyer & Williams, 1993)

1.3.2.5 *Segmentation en fonction de l'équipement électroménager et de son utilisation pour la consommation en électricité du ménage*

D'après les auteurs de « Electricity load profiles in Europe: The importance of household segmentation », les critères socio-démographiques et le style de vie des ménages n'expliquent que partiellement la consommation en électricité des ménages. En revanche, d'après l'étude « Développement et application d'une méthodologie pour les profils de charge des consommateurs (Entwicklung und Anwendung einer Methodik zur Synthesezukünftiger Verbraucherlastgänge, Munich; 2013) » de Gobmaier TH, les caractéristiques de l'électroménager et son utilisation permettraient d'expliquer près de 96,9% de la variation de la consommation d'élecricité au cours d'une journée (profil de charge). C'est pourquoi l'électroménager (électrique) devrait être utilisé comme critère de segmentation.

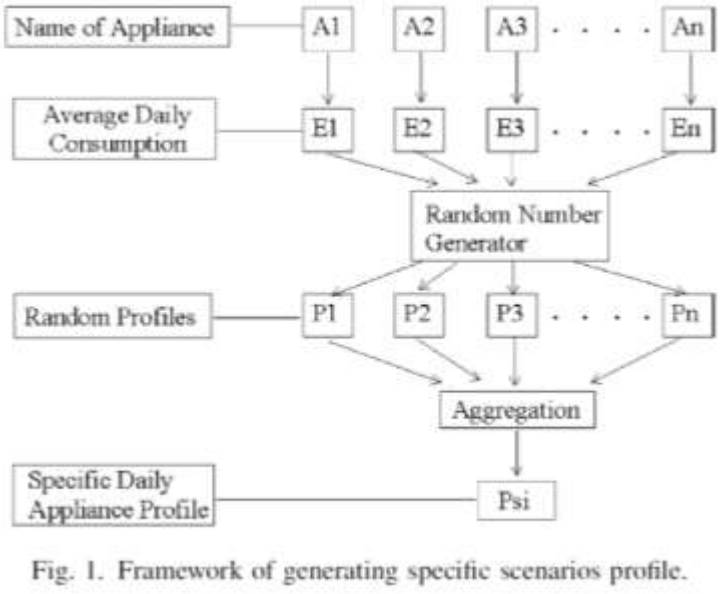


Figure 1-9 : Méthodologie pour générer des profils spécifiques de consommation (Yao & Steemers, 2005).

L'article « A method of formulating energy load profile for domestic buildings in the UK » propose une méthode (voir Figure 1-9) pour générer des profils de consommation d'électricité à partir des appareils électroménagers et de leurs utilisations puis compare les profils générés avec la consommation reélle de 1300 ménages britanniques.

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Electricity load profiles in Europe: The importance of household segmentation »	Un article synthétique sur les différentes approches pour la segmentation des ménages.	Allemagne. Notes : certains articles étaient seulement disponibles en allemand.	(Hayn, Bertsch, & Fichtner, 2014)
« A method of formulating energy load profile for domestic buildings in the UK »	L'article présente une méthode pour établir des profils de charge des consommateurs.	Grande-Bretagne, échantillon de 1300 clients, pour valider le modèle.	(Yao & Steemers, 2005)

1.3.2.6 Segmentation automatique des ménages

Jusqu'à présent toutes les études recensées dans les parties 1.3.2.1 à 1.3.2.5 utilisaient la segmentation manuelle pour établir des groupes de ménages, analyser la consommation spécifique du ou des groupes ou bien établir des comparaisons entre la consommation d'énergie de ces groupes de ménages. Dans l'étude intitulée "Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy-related behavioral characteristics" (Sütterlin, Brunner, & Siegrist, 2011), les auteurs poursuivent des objectifs semblables aux articles présentés jusqu'à présent. Il souhaite analyser différents segments de consommateurs d'énergie en fonction de critères comportementaux et psychologiques afin d'identifier les segments qui ont le plus grand potentiel en termes d'économie d'énergie. À la différence près que pour réaliser leur segmentation, ils n'utilisent pas une segmentation manuelle, mais une segmentation automatique. En effet, après avoir sélectionné les variables sur lesquelles réaliser l'analyse de regroupement, il utilisent un algorithme hiérarchique pour finalement aboutir à 6 groupes, de taille raisonnable. Puis, ils caractérisent les six-groupes relativement à leurs aptitudes aux économies d'énergie : « the idealistic energy-saver (15.6%),the selfless inconsequent energy-saver (26.4%), the thrifty energy-saver (14%),the materialistic energy consumer (25.1%),the convenience-oriented indifferent energy consumer (5.3%),and the problem-aware well-being-oriented energy consumer(13.6%) ».

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy related behavioral characteristics »	Réalisation d'une segmentation automatique des consommateurs pour caractériser les attitudes de chacun des groupes relativement aux économies d'énergie.	Suisse, un sondage conduit auprès de ménages suisses. Soit près de 1292 ménages composent l'échantillon final.	(Sütterlin et al., 2011)

1.3.2.7 Segmentation automatique des ménages et compteur intelligent (« Smart Grid »)

Enfin, une dernière catégorie d'articles s'attache à identifier des profils types de ménages en fonction des données de consommation d'électricité fournies par les nouveaux compteurs intelligents qui mesurent la consommation en temps réel, c'est-à-dire la consommation en kWh à chaque heure par exemple.

Cette innovation confère en effet un grand avantage. Si nous pouvions jusqu'à présent regrouper des ménages ayant le même niveau de consommation/dépense, annuel ou au mois, en électricité, cela n'indiquait en rien que pour un même groupe de ménages, un tiers des ménages de ce groupe consomment 50% du total journalier le matin, un autre tiers 80% entre 18h00 et 21h00 par exemple, et enfin pour le dernier tiers des ménages de ce groupe, leur consommation électrique serait également répartie sur la journée. De plus, les données sur la consommation des compteurs intelligents permettent d'identifier et de regrouper des ménages dont la consommation totale pourrait différer, mais la répartition (en valeur relative, pourcentage du total journalier) au cours de la journée serait semblable.

Les articles traitant de ce type de segmentation sont nombreux et la segmentation automatique reposant sur ce type de données doit faire face à un défi computationnel (du fait de la quantité de données à traiter).

Par exemple, l'article « Household Energy Consumption Segmentation Using Hourly Data » présente une méthode pour la segmentation des ménages sur la base des données horaires de consommation électrique permettant d'identifier des segments de consommateur qui ont un fort potentiel pour des politiques de gestion de la demande et d'économies d'énergie (« Demand Response, DR » et « Energy Efficiency » EE) (Jungsuk Kwac, Flora, & Rajagopal, 2014) du fait du consommation horaire stable par exemple. Dans cet article « energy consumers load shape information then is used to classify households according to extracted features such as entropy of shape code which measures the amount of variability in consumption. Load shape information enhances our ability to understand individual as well as groups of consumers. For example, time of day building occupancy and energy consuming activities can be interpreted from these shapes » (Jungsuk Kwac et al., 2014). Ce type d'analyse donne un plus grand moyen d'action aux pouvoirs publics, aux producteurs et distributeurs d'électricité. Elle a bien entendu son pendant juridique car il est possible d'inférer des données sociodémographiques précises du ménage comme l'illustre

l'article intitulé « Revealing household characteristics from smart meter data » (Beckel, Sadamori, Staake, & Santini, 2014).

Ci-dessous, le tableau présente à titre d'illustration un échantillon d'articles qui réalisent ce type d'analyse.

Titre de l'article	Analyse porte sur	Données de	Auteurs, date
« Household Energy Consumption Segmentation Using Hourly Data »	Segmentation des ménages en utilisant les données horaires de consommation d'électricité.	Californie, données de PG&E.	(Jungsuk Kwac et al., 2014)
« Customer Characterization Options for Improving the Tariff Offer »	Segmentation des consommateurs d'électricité en utilisant des données horaires de consommation. Puis, pour chaque groupe de consommateurs, la latitude du distributeur pour proposer des tarifs spécifiques est discutée.	Roumanie, 471 consommateurs non résidentiels.	(Chicco, Napoli, Postolache, Scutariu, & Toader, 2003)
« Revealing household characteristics from smart meter data »	Estimation des caractéristiques d'un ménage à partir de sa consommation horaire d'électricité.	Irlande, CER (Commission for Energy Regulation), 4232 ménages.	(Beckel et al., 2014)
« Smart meter data: Balancing consumer privacy concerns with legitimate applications »	Réflexion sur les applications possibles des données des compteurs intelligents et les enjeux juridiques.	Différents exemples.	(McKenna, Richardson, & Thomson, 2012)
« Smart Meter Driven Segmentation: What Your Consumption Says About You »	Segmentation des ménages à partir des données de consommation horaire et inférence des caractéristiques des ménages à partir de ces données afin de mieux cibler les politiques d'économie d'énergie.	E.U., données de 1100 ménages et sondage.	(McKenna et al., 2012)

2 EDM et la préparation des données

Cette section du rapport présente les données sur lesquelles repose l'analyse (construction d'un portrait des ménages québecois) du rapport.

2.1 L'enquête sur les dépenses des ménages (EDM) de Statistique Canada

2.1.1 Historique des enquêtes menées sur les dépenses des ménages au Canada

Avant 1997, Statistique Canada réalisait deux enquêtes relatives aux dépenses des ménages : l'enquête sur les dépenses des familles ("Enquête sur les dépenses des familles (Famex)," n.d.), effectuée tous les deux ans (voire tous les quatre ans[5]), et l'enquête annuelle sur l'équipement ménager des canadiens ("Enquête sur l'équipement ménager (EEM)," n.d.).

L'enquête sur les dépenses des familles, Famex, se consacrait aux conditions socio-économiques des ménages (caractéristiques du ménage, dépenses et épargne, revenu) alors que l'enquête sur l'équipement ménager, EEM[6], réalisait l'inventaire de l'équipement ménager des ménages canadiens et relevait les changements qui se produisaient dans les caractéristiques du logement.

En 1997, ces deux enquêtes, Famex et EEM, ont été intégrées à l'enquête sur les dépenses des ménages canadiens, EDM ("Enquête sur les dépenses des ménages (EDM)," n.d.). Celle-ci a été effectuée chaque année depuis 1997 avec quelques modifications d'année en année, certaines mineures (questions ajoutées ou supprimées au questionnaire d'enquête) et d'autres majeures (modification du mode de collecte des données en 2010 par exemple).

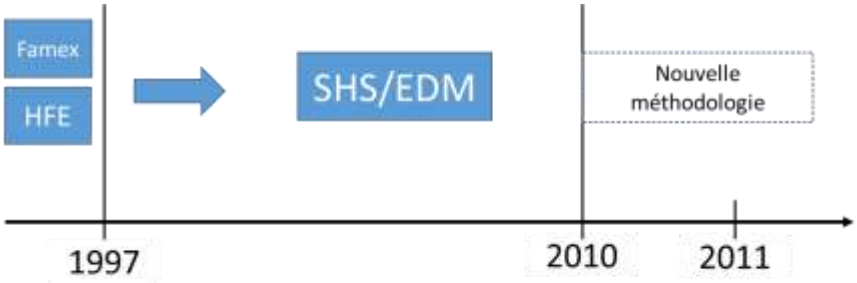


Figure 2-1 : Les enquêtes sur les dépenses des ménages canadiens, avant 1997 jusqu'à nos jours.

[5] Les fichiers de microdonnées à grande diffusion (FMGD) pour l'enquête sur les dépenses des familles (Famex) sont disponibles pour les années 1982, 1986, 1990, 1992, 1996 par exemple, ce qui semble indiquer une publication de cette enquête tous les quatres ans plus que tous les deux ans comme le mentionne Statistque Canada ("Enquête sur les dépenses des familles (Famex)," n.d.).

[6] En anglais, cette enquête portait le nom de *Household Facilities and Equipement Survey* (HFE).

Pour l'année donnée, une fois que celle-ci s'est écoulée, l'enquête sur les dépenses des ménages est menée. Le questionnaire est administré aux ménages sélectionnés pour l'enquête en demandant aux personnes de se souvenir des dépenses engagées durant l'année précédente et de fournir les caractéristiques du logement et de l'équipement ménager en date du 31 décembre de l'année qui vient de se terminer. Par exemple, l'enquête sur les dépenses des ménages de l'année 1997 a été menée entre janvier et mars 1998.

Depuis 2010, des changements ont été apportés à l'enquête sur les dépenses des ménages EDM. Désormais, pour la collecte des données, elle combine deux modes de collectes (le questionnaire, et un journal de dépenses). De plus, des changements ont été apportés aux méthodes de traitement et d'estimation des données, ce qui cause une rupture dans la série débutée en 1997. En conséquence, les comparaisons sur la série 1997 – 2009 sont valides, mais il faut faire preuve de prudence quant aux comparaisons d'année en année bien que les données recueillies soient similaires (voir annexe 6.1. pour une présentation des changements apportées à l'enquête EDM entre 1997).

2.1.2 Les fichiers de microdonnées à grande diffusion (FMGD)

Depuis 2001, les microdonnées (ou données brutes) de nombreuses enquêtes menées sur les ménages et la population par Statistique Canada sont mises à la disposition des chercheurs et des étudiants dans le réseau canadien des Centres de Données de Recherche (CDR). L'enquête sur les dépenses des ménages fait partie de cette Initiative de Démocratisation des Données (IDD). Dans un fichier de microdonnées à grande diffusion, les données de l'enquête ont seulement été traitées (aggrégation en catégories moins détaillées, plafonnement ou suppresion de certaines variables) pour éviter toute identification possible avec les répondants. Ainsi, les chercheurs et les étudiants disposent des microdonnées à grande diffusion de l'enquête sur les dépenses des ménages (EDM) pour les années 1997 à 2009. À compter de l'année 2010, pour des raisons budgétaires, Statistique Canada ne prépare plus les fichiers de microdonnées à grande diffusion pour l'enquête EDM[7].

[7] Il est possible d'accéder aux microdonnées confidentielles de l'enquête EDM pour les années 2010 et 2011 en suivant la procédure qui permet d'accéder aux microdonnées confidentielles d'un CDR.

Ce sont ces microdonnées à grande diffusion d'EDM qui seront utilisées pour le travail présenté dans ce rapport. De plus, les comparaisons d'année en année seront faites pour la série 1997 – 2009.

2.1.3 L'imputation des données manquantes :

Lorsque des informations sont manquantes ou invalides dans le questionnaire EDM, Statistique Canada a recours à une technique appelée imputation afin de compléter les données manquantes d'un répondant. Pour cela, Statistique Canada utilise la méthode des k plus proches voisins, c'est-à-dire « les données d'un autre répondant ayant des caractéristiques similaires (le donneur) sont donc utilisées pour imputer. L'imputation se fait sur un groupe de variables à la fois, ce groupe étant formé de sorte à tenir compte des relations entre ces variables. Les caractéristiques pour identifier le donneur sont choisies de sorte à être corrélées avec les variables à imputer. Le revenu du ménage, le type de logement et le nombre d'adultes et d'enfants sont des caractéristiques souvent utilisées » ("Enquête sur les dépenses des ménages (EDM)," n.d.). Cela veut dire que dans le fichier de micro-données lorsqu'un répondant a indiqué une dépense de 0 $ en électricité, il s'agit en effet d'un ménage dont la dépense en électricité est bien de 0 $. Comme cela sera présenté dans la partie 4.1.3.2, cette dépense de 0$ en électricité pour un certain nombre de ménages trouvent son explication dans les données.

2.1.4 Le masquage des données

Les fichiers de microdonnées à grande diffusion de l'EDM sont traités afin d'assurer la confidentialité des répondants. Voici deux exemples de « masquage » des données d'EDM :

La variable « Age » :

L'âge est traité comme une variable catégorielle. C'est-à-dire qu'à la place de l'âge précis d'un membre du ménage, il est indiqué le groupe d'age (ex de 20 à 30 ans) auquel appartient ce membre.

La variable « Revenu (HHINCTOT) » :

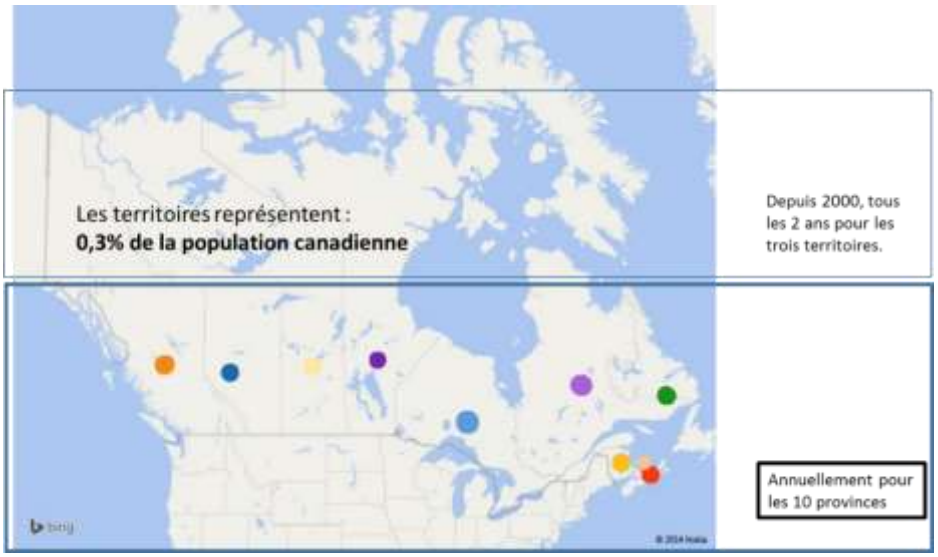
Le revenu de la personne de référence du ménage est arrondi. Par exemple, si le revenu est égal à 125 678$, alors celui-ci sera arrondi à 130 000$ (arrondi au dizaine de millier pour les revenus supérieurs à 100 000$, au millier pour les revenus supérieurs à 10 000$ et à la centaine pour les revenus inférieurs à 10 000$).

2.1.5 La population cible d'EDM

La population cible est décrite dans l'enquête EDM comme suit : « la population cible correspond à la population des 10 provinces canadiennes à l'exception des pensionnaires d'établissements institutionnels, des membres des Forces canadiennes vivant dans des camps militaires et des personnes vivant dans les réserves indiennes. Dans l'ensemble, ces exclusions représentent environ **2 % de la population des 10 provinces**. Pour des raisons opérationnelles, les personnes vivant dans des régions où le taux de logements vacants est très élevé et où les coûts de collecte seraient exorbitants sont exclues de la collecte. Sont également exclues, en plus des personnes vivant dans les institutions, les personnes vivant dans d'autres types de logements collectifs tels que :

1. les personnes qui vivent dans les résidences pour personnes âgées non autonomes; et
2. les personnes qui vivent de façon permanente dans les résidences scolaires, les hôtels et motels, les camps de travail, etc.; et
3. les membres d'ordre religieux et d'autres groupes vivant en communauté.

Les exclusions à la collecte représentent moins d'un demi-pourcent de la population cible. Ces personnes sont toutefois incluses dans les estimations démographiques auxquelles sont ajustées les estimations de l'EDM » ("Enquête sur les dépenses des ménages (EDM)," n.d.).



Ajoutons à cette information que de 1997 à 1999 EDM était menée chaque année dans les territoires[8] et à compter de l'année 2000, une fois sur deux seulement[9].

[8] Dans la suite de l'analyse, lorsqu'il sera fait mention des provinces canadiennes, cela exclura les territoires qui représentent moins de 3% de la population canadienne totale. Par conséquent, 6223 observations de l'échantillon sont éléminées pour les année 1997 à 2009, voir annexe 6.3.
[9] Pour une estimation du nombre de ménages dans les territoires, voir annexe 6.3.

2.1.6 Le questionnaire EDM en un clin d'oeil

Le questionnaire administré aux ménages canadiens chaque année vise à obtenir de l'information quant à la localisation (province, zone rurale/urbaine…) du ménage, aux caractéristiques du logement (type d'habitation, nombre de pièces…), la description du ménage (nombre de personnes vivant dans le ménage…), l'équipement ménager du ménage et les dépenses annuelles de celui-ci.

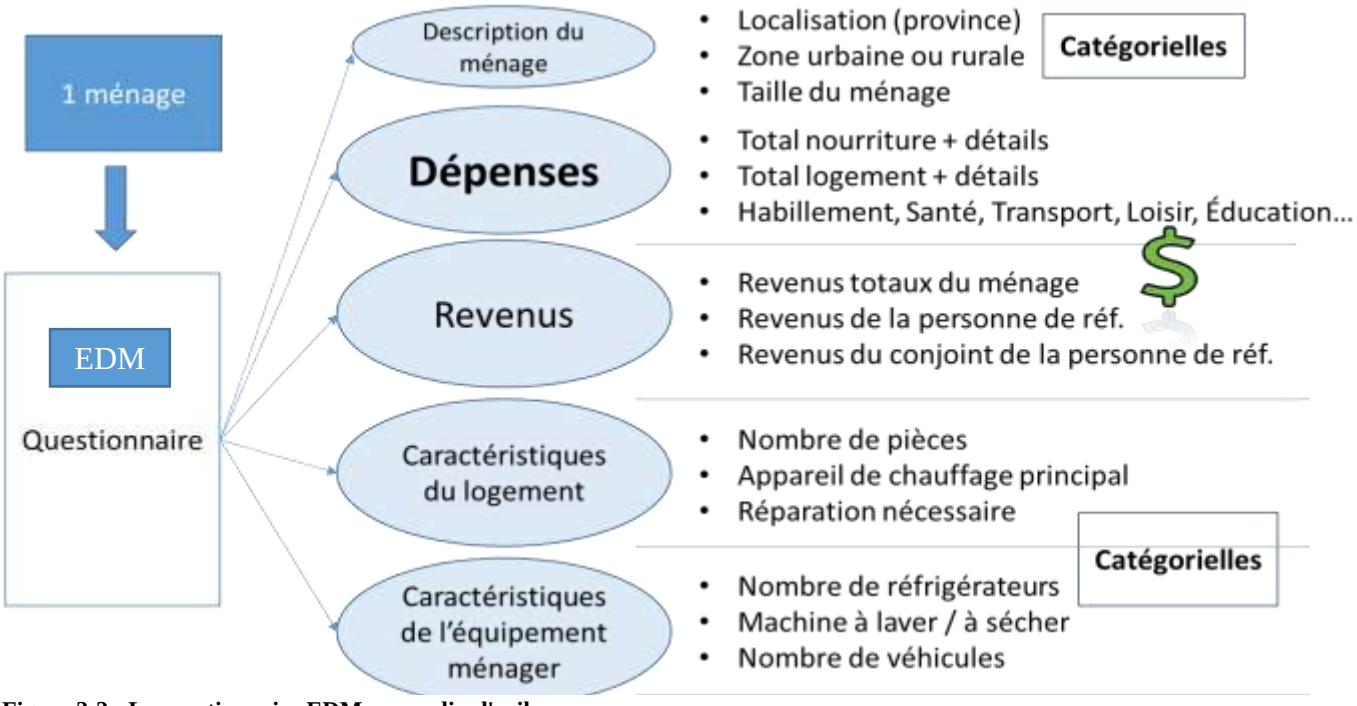


Figure 2-2 : Le questionnaire EDM en un clin d'oeil

Une liste complète des variables disponibles dans les fichiers de micro données est à disposition en annexe 6.1 .

2.1.7 La taille des échantillons de 1997 à 2009, EDM

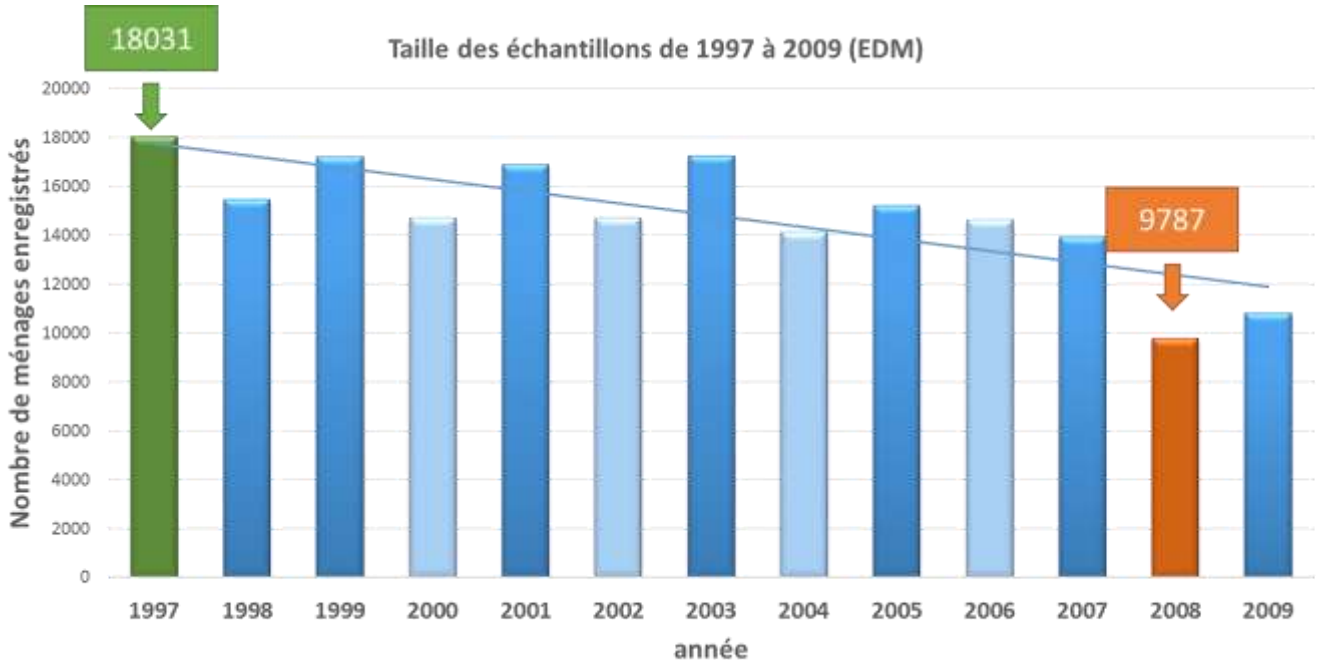


Figure 2-3 : La taille des échantillons pour chaque année (nombre d'observations).

À l'exception de l'année 2008, la taille des échantillons sont toutes supérieures à 10000 observations enregistrées. Statistique Canada mentionne que la taille des échantillons pour EDM en 2008 et 2009 est plus réduite en raison de la redéfinition d'EDM pour 2010 (nouvelle méthodologie qui combine l'utlisation d'un questoinnaire et d'un journal pour collecter les données, voire annexe 6.2.3 pour plus de détail) qui a occupé une partie de l'équipe habituellement en charge de l'administration d'EDM.

Jusqu'en 1999, EDM était menée dans les dix provinces ainsi que les trois territoires. À compter de 2000, EDM est menée une année sur deux dans les territoires en commençant à l'année 2001. Cela permet de comprendre pourquoi la taille des échantillons, observations enregistrées, pour les années 2000, 2002 et 2004 (années où EDM est menée uniquement dans les 10 provinces) est sensiblement inférieure aux années 2001, 2003 et 2005 (années où EDM est aussi menée dans les trois territoires).

2.2 La préparation des données

Même s'il s'agit d'une enquête réalisée par Statistique Canada, le format de l'enquête a connu d'année en année divers changements : d'un changement de nom de variable à la suppression pure et simple d'une variable en passant par la recodification d'une variable. Il faut donc d'abord s'assurer d'une nomenclature commune des données si l'on souhaite établir des comparaisons entre plusieurs années. Les trois cas présentés ci-dessous donnent un aperçu de la préparation de données. **Lorsque une préparation de données en particulier sera utile à la compréhension des analyses présentées dans le rapport elle sera alors explicitée.**

2.2.1 Changement de nom des variables par Statistique Canada

Par exemple, en 2008, la variable CH04TOI[10] a été renommée CH04TOIP (ajout d'un P). Dans le cadre de la préparation des données, celle-ci a été renommée CH04TOI pour 2008 et 2009. En 2009, la variable DSHWASH[11] a été renommée DSHWAHP (ajout d'un P). Celle-ci a été renommée DSHWASH dans tous les fichiers de EDM de 1997 à 2008.

2.2.2 Ajout et suppression de variables d'une année à l'autre par Statistique Canada

De 2002 à 2005, beaucoup de nouvelles variables sont apparues (comme CLOSEFAC[12], ESTHHLD[13], CHNGTEN[14], CHNGJOB[15] par exemple) dans l'enquête EDM. Mais celles-ci sont absentes des enquêtes à partir de 2006 (et manquent aux enquêtes qui précèdent 2002). Elles sont été éliminées des fichiers de données.

2.2.3 Variation dans la codification / l'échelle des variables par Statistique Canada

Par exemple, la variable WASHERP[16] pour laquelle la catégorie 3 a été remplacée par la catégorie 0. De même pour la variable AIRCONP[17], pour laquelle la catégorie 3 a été aussi remplacée par la catégorie 0. Les données ont donc été harmonisées en termes de codification pour les variables catégorielles.

[10] Voir la liste complète des variables disponibles pour la période 1997-2009 en annexe 6.1. Number of children 0 to 4 at time of interview.
[11] Id. Dishwasher.
[12] CLOSEFAC, Moved to be closer to facilities and services, 0 = no, 1 = yes.
[13] ESTHHLD, Moved to establish own household, 0 = no, 1 = yes.
[14] CHNGTEN, Moved to change from owner to renter or renter to owner, 0 = no, 1 = yes.
[15] CHNGJOB, Moved for a new job or job transfer, 0 = no, 1 = yes.
[16] WASHERP, Washing machine
[17] AIRCONP, Air conditioner

2.2.4 Ajout de données externes pour fin d'analyse

2.2.4.1 Prix de l'électricité par province, par année (de 1997 à 2009) :

Les caractéristiques de la variable « prix de l'électricité » sont les suivantes :

- Le prix moyen à la date de référence de l'année, soit mai ou avril,
- en Cent/kWh
- toutes taxes comprises
- pour le résidentiel
- prix pour le niveau de consommation 1000 kWh

Les prix sont issus des rapports produits annuellement par Hydro-Québec, *Comparaison des prix de l'électricité dans les grandes villes nord-américaines*, pour les années 1997 à 2009 (Hydro-Québec, n.d.).

2.2.4.2 Prix du gaz pour le Québec, par année (de 1997 à 2009) :

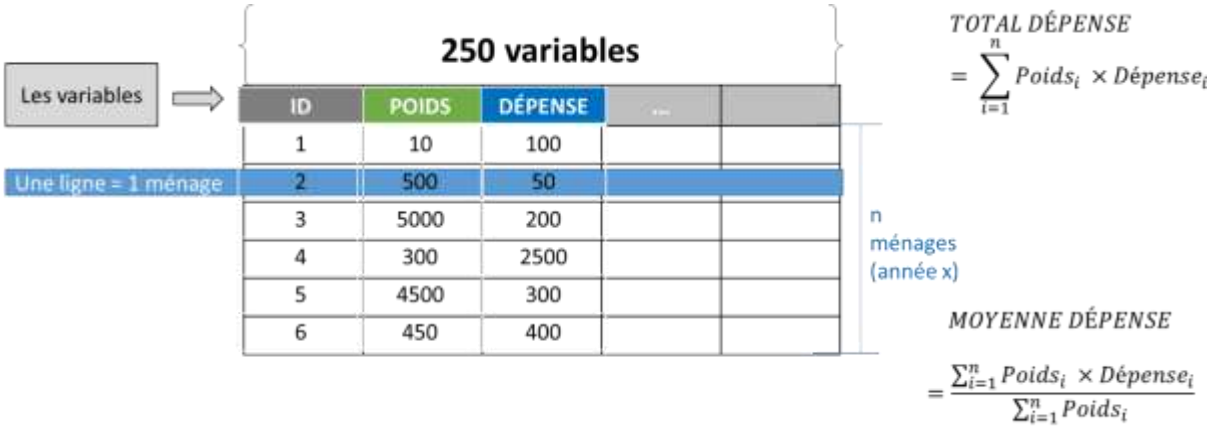
Les caractéristiques de la variable « prix du gaz » sont les suivantes :

- Le prix courants,
- en Cent par m^3
- ne comprend pas les taxes provinciales et fédérales
- pour le résidentiel résidentiel

Les prix du gaz ont pour source le ministère de l'Énergie et des Ressources naturelles et Statistique Canada ("MERN - Prix du gaz naturel," n.d.).

2.3 Analyse descriptive des dépenses en énergie pour les ménages canadiens de 1997 à 2009

2.3.1 Préalable : point sur l'inférence statistique à partir des données EDM



Il convient de distinguer la moyenne observée pour une variable quantitative dans l'échantillon des ménages d'une année donnée de EDM (le revenu des ménages par exemple) de l'estimation de la moyenne pour cette même variable à l'échelle de la population des ménages. À cette fin, Statistique Canada fournit une variable poids, « WEIGHT » qui a pour postulat qu'un ménage de l'échantillon représente en plus de lui-même aussi un certain nombre d'autres ménages. C'est ce que représente la variable poids, c'est-à-dire la part que ce ménage représente parmi l'ensemble des ménages de la population.

2.3.2 EDM et les dépenses en énergie d'un ménage

Dans l'enquête sur les dépenses des ménages canadiens (EDM), trois variables principales permettent de rendre compte des dépenses en énergie des ménages canadiens. Il s'agit des dépenses suivantes :

- l'électricité, variable nommée « G024 », pour la résidence principale
- les combustibles liquides, variable nommée « G022 », incluant le gaz, pour la résidence principale

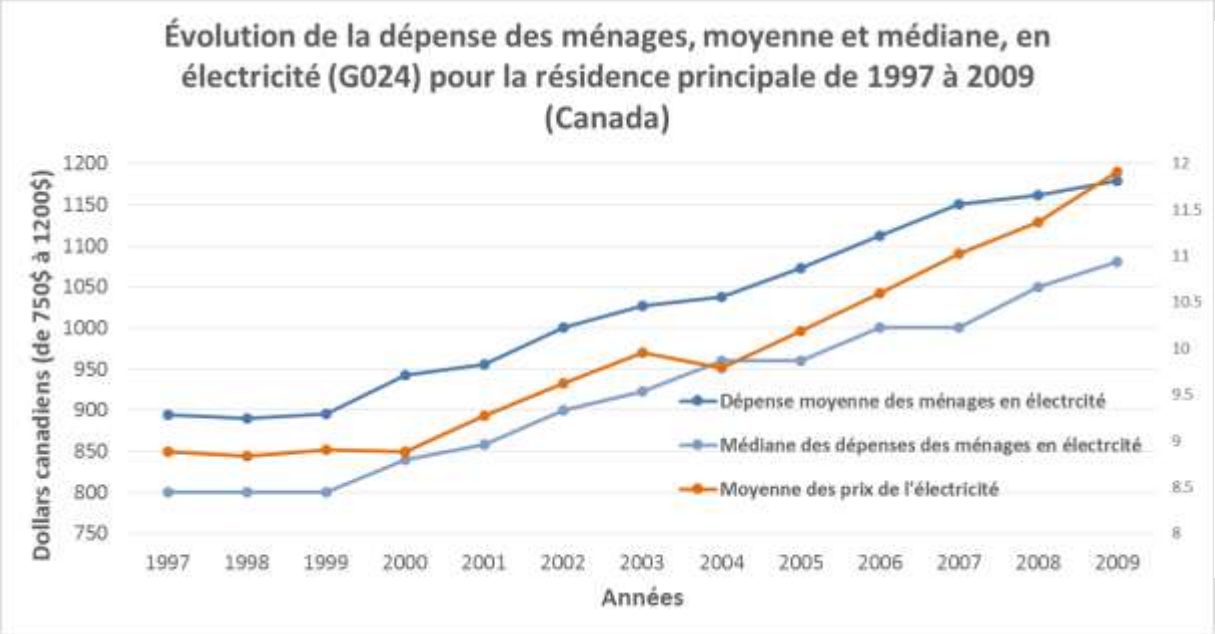
- des dépenses en essence et autres combustibles pour le transport, variable nommée « K020 »

L'analyse de ces trois postes de dépenses permettra de rendre compte de l'évolution des dépenses en énergie des ménages canadiens, de 1997 à 2009. Les graphiques ci-dessous présentent l'évolution des dépenses moyennes des ménages pour ces trois variables, pour le Canada dans son ensemble (hors territoires[18]).

L'objectif de cette courte analyse descriptive est d'identifier des tendances, des faits marquants afin d'aborder la phase de construction du portrait des ménages québecois avec une première connaissance des données.

2.3.2.1 G024 : Électricité pour la résidence principale

La dépense moyenne des ménages canadiens pour l'électricité est de 900 dollars canadiens en 1997. Celle-ci est en augmentation régulière chaque année et atteint environ 1200 dollars en 2009, soit une augmentation de 32% sur une période de 12 années. Sur cette même période, l'augmentation du prix de l'électricité moyen, à l'échelle du Canada est d'environ 34%.



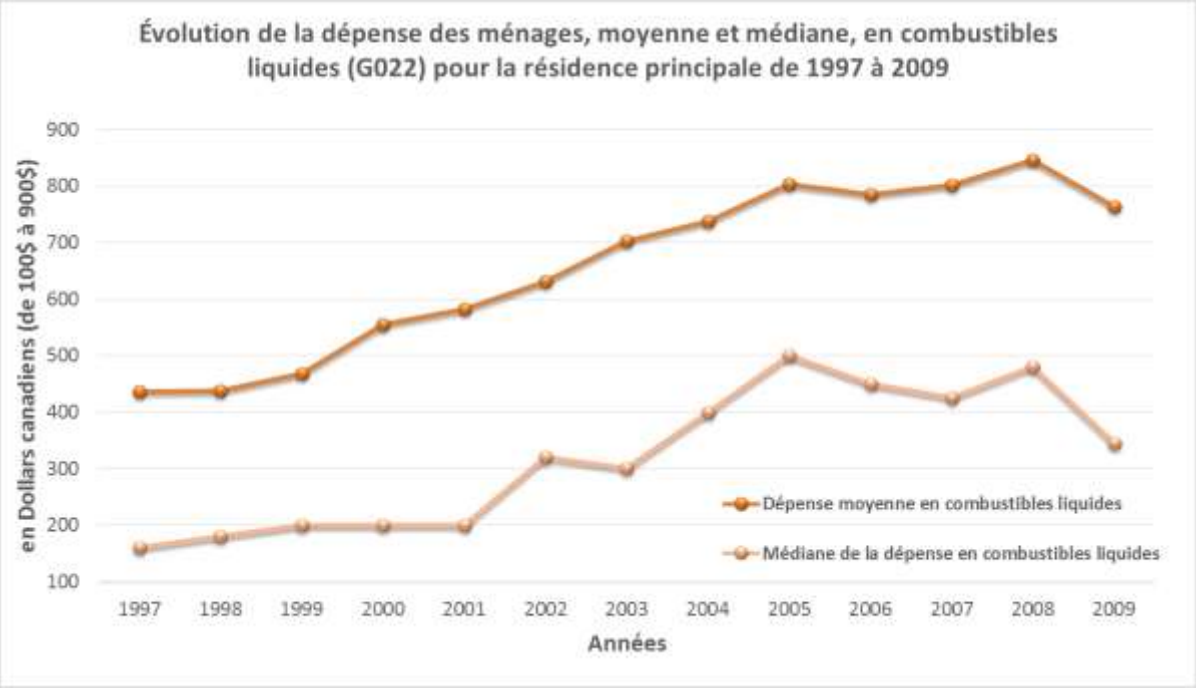
[18] À partir de 2001, EDM a été menée annuellement dans les 10 provinces et une fois sur deux dans les territoires. Par conséquent, les comparaisons sur la série 1997 à 2009 sont possibles pour les provinces uniquement.

Bien entendu, ces moyennes canadiennes masquent des réalités bien diverses, mais elles permettent d'indiquer une tendance entre 1997 et 2009. Par exemple, l'augmentation du prix de l'électricité pour l'Alberta est de 58% sur cette période alors qu'elle n'est que de 14% en Colombie-Britanique et de 15% au Québec. De même l'augmentation moyenne de la dépense en électricité d'un ménage en Alberta est de 86%, alors qu'elle n'est que de 20% et 7% respectivement pour le Québec et la Colombie-Britanique[19].

2.3.2.2 G022 : Combustibles liquides pour la résidence principale (essence, gaz, etc.)

De 1997 à 2003, Statistique Canada recueillait la valeur des dépenses pour G022. À partir de 2004, cette variable a été décomposée en deux : les dépenses pour le gaz pour la résidence principale sont agrégéees dans la variable G021 et les autres dépenses en combustibles (propane ou bois par exemple), toujours pour la résidence principale, sont aggrégées quant à elle dans la variable G023.

Afin d'établir la continuité entre 1997 et 2009, la variable G022 a été créée pour les fichiers de microdonnées de 2004 à 2009 en additionnant G022 et G023. Cette précision apportée, l'analyse de cette variable peut-être présentée.

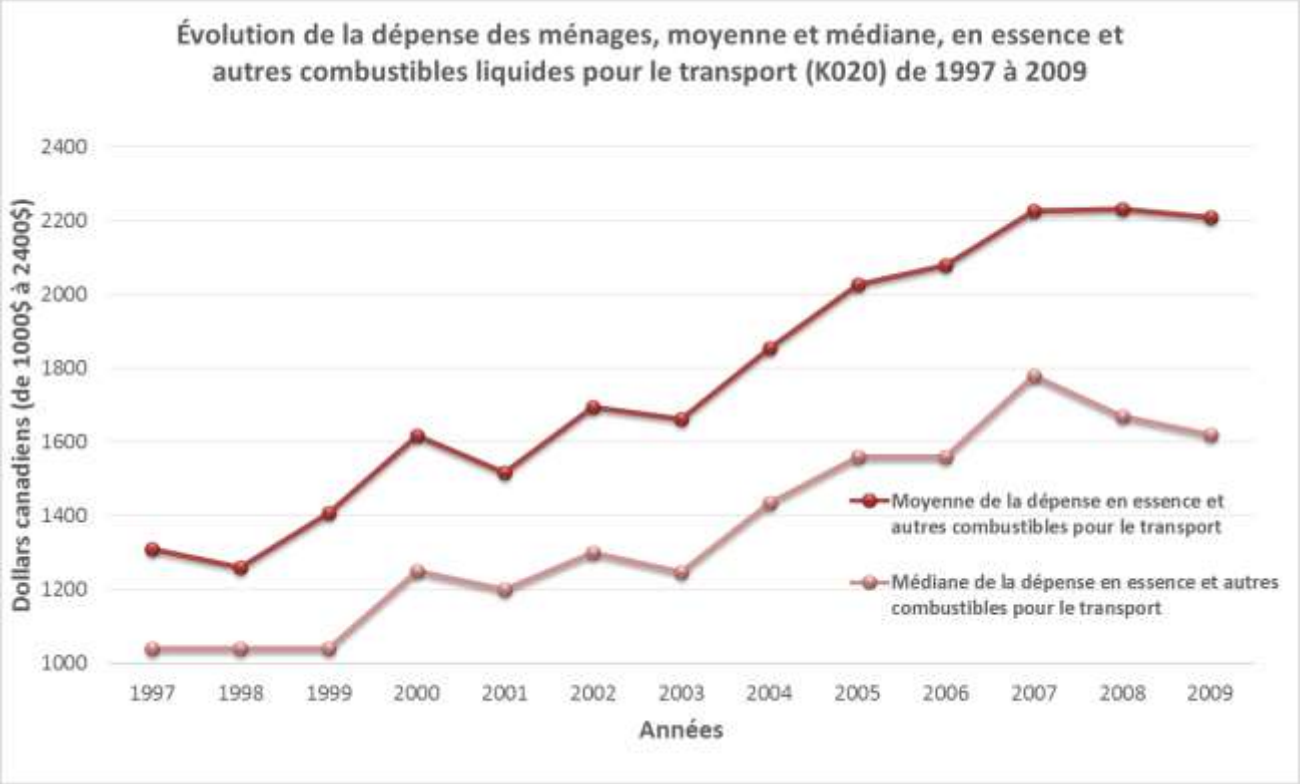


[19] Pour les estimations au niveau canadien, les observations pour lesquelles les provinces masquées ou manquantes sont conservées. En revanche, celles-ci sont éliminées pour les estimations au niveau des provinces.

La dépense moyenne d'un ménage canadien en combustibles liquides (dont le gaz) pour la résidence principale est elle aussi en augmentation. La moyenne en 1997 était de 436 dollars canadiens et elle atteint 763 dollars canadiens en 2009, soit une augmentation de 75% en 12 ans.

2.3.2.3 K020 : Essence et autre combustibles pour le transport

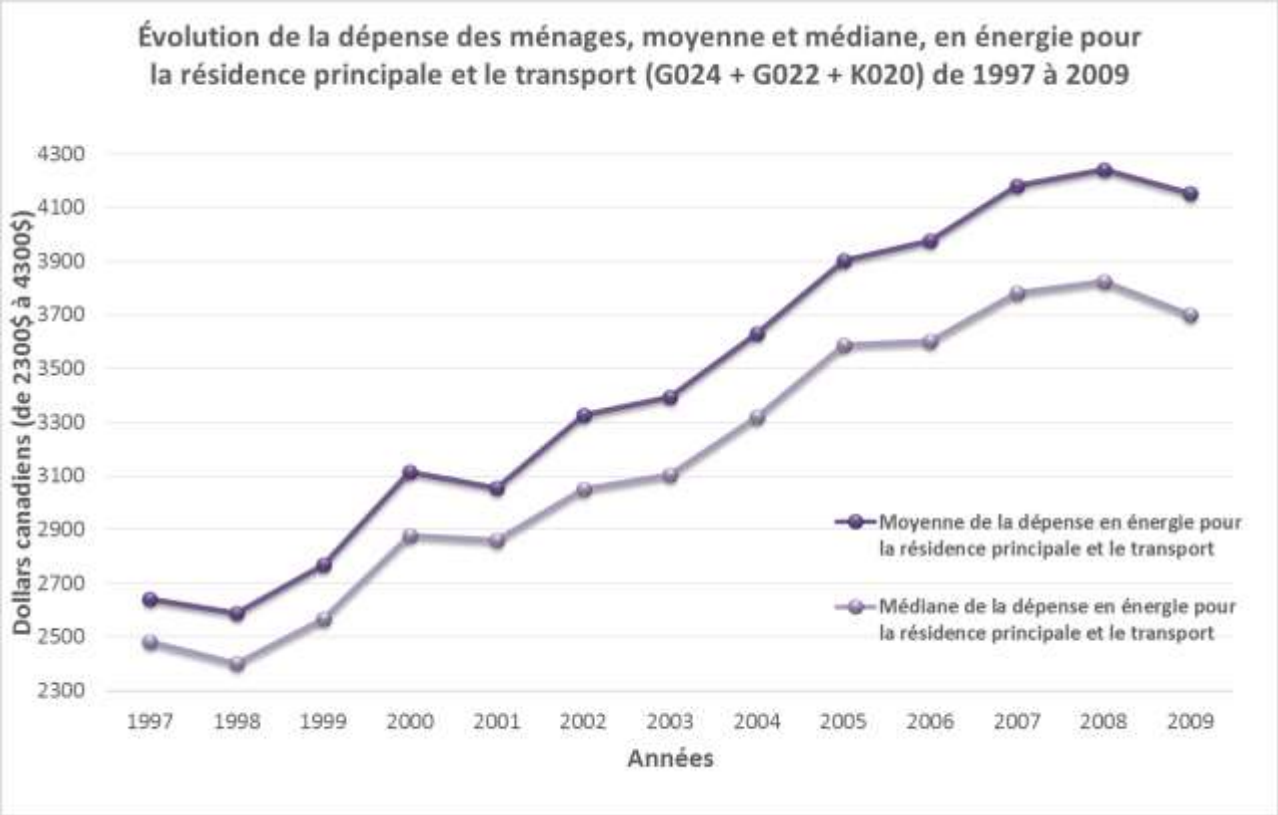
Cette variable aggrège l'ensemble des dépenses d'un ménage pour l'essence et autres combustibles (diesel, propane) utilisés pour les voitures privées ou bien louées (aussi bien les voitures que les camions et les vans).



En 1997, un ménage canadien dépense en moyenne 1310 dollars canadiens en énergie pour le transport. De 1997 à 2009, la moyenne a augmenté de près de 70% pour atteindre un montant de 2200 dollars canadiens.

2.3.2.4 *Energytot : G024 + G022 + K020*

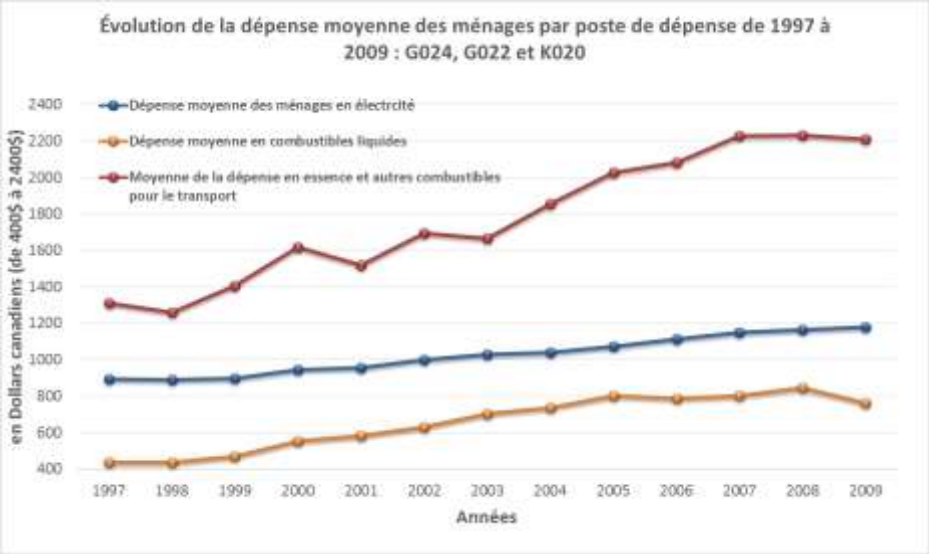
Afin de présenter une vue unique des dépenses des ménages canadiens en termes de dépenses énergétiques, la variable Energytot a été créée. Celle-ci correspond à la somme des dépenses en électricité (G024), en gaz (G022) pour la résidence principale et aux dépenses en énergie pour le transport (K020).



En 1997, un ménage canadien dépensait en moyenne 2650 dollars en énergie, alors qu'en 2009 celui-ci dépense 4150$, **soit une augmentation moyenne de 57%.**

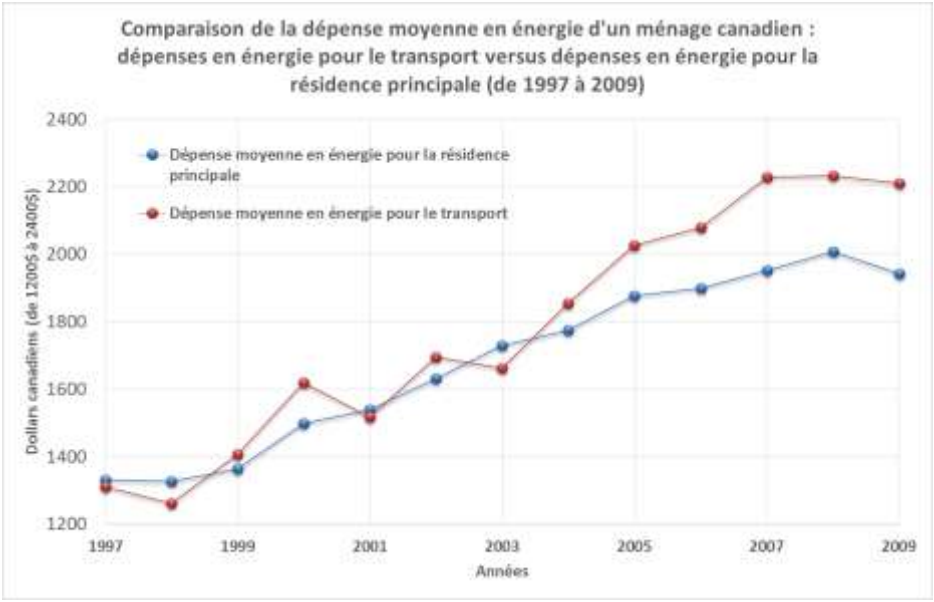
2.3.2.5 Comparaison des dépenses des ménages en énergie, de 1997 à 2009 :

2.3.2.5.1 Pour les postes G024, G022 et K020 :



La dépense d'énergie des ménages canadiens qui a le plus augmenté au cours de la période 1997 – 2009 est la dépense d'énergie pour le transport. Puis, dans une bien moins grande proportion, c'est la dépense en combustibles liquides. La dépense moyenne en électricité connaît un augmentation moyenne faible et régulière sur toute la période.

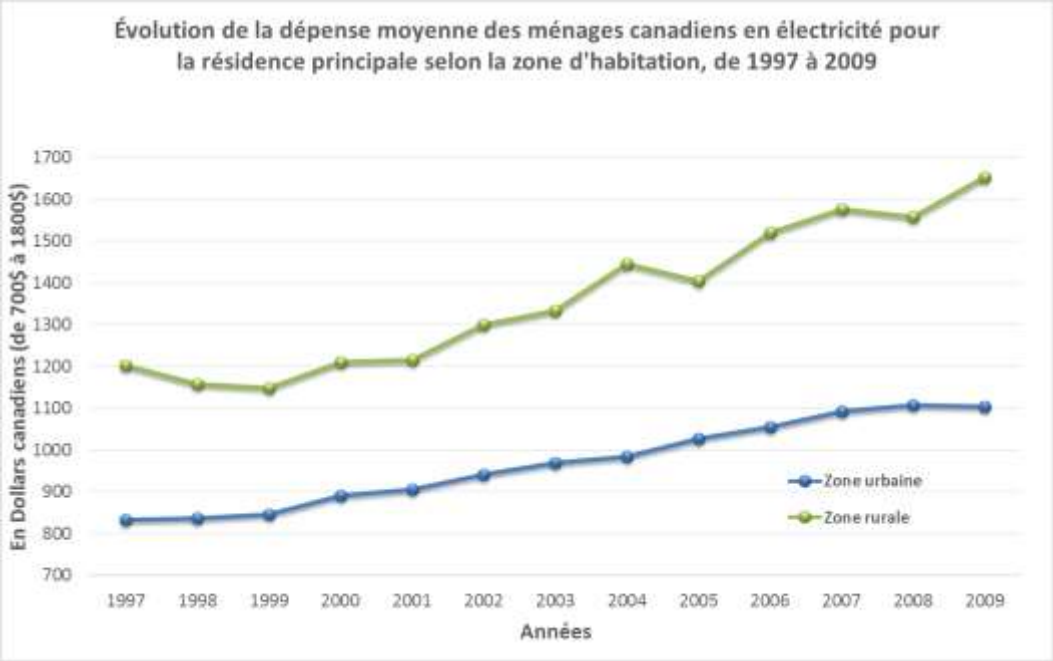
2.3.2.5.2 Énergie pour le transport versus énergie pour la résidence principale :



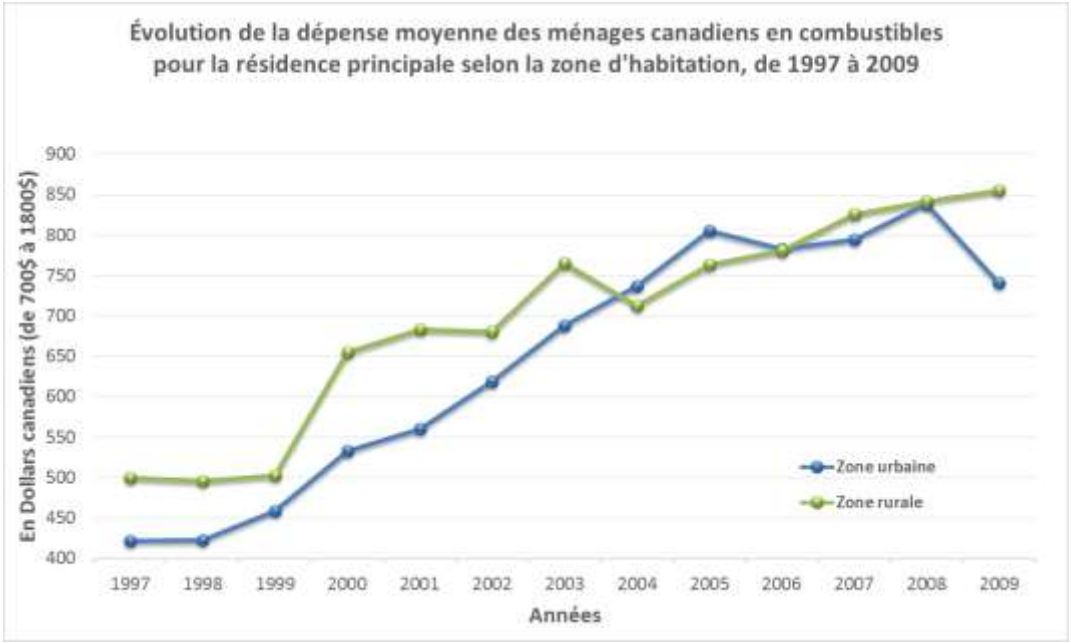
En additionnant les dépenses d'énergie pour la résidence principale (Électricité et Gaz), il est intéressant de constater que celle-ci est alors équivalente à la dépense moyenne en énergie pour le transport jusqu'en 2003. À partir de 2004, il y a un décrochage entre la dépense moyenne

d'énergie pour le transport et la dépense moyenne d'énergie pour la résidence principale, la dépense moyenne pour le transport augmentant plus vite chaque année.

2.3.2.5.3 Dépenses en énergie : ménage en zone urbaine versus ménage en zone rurale



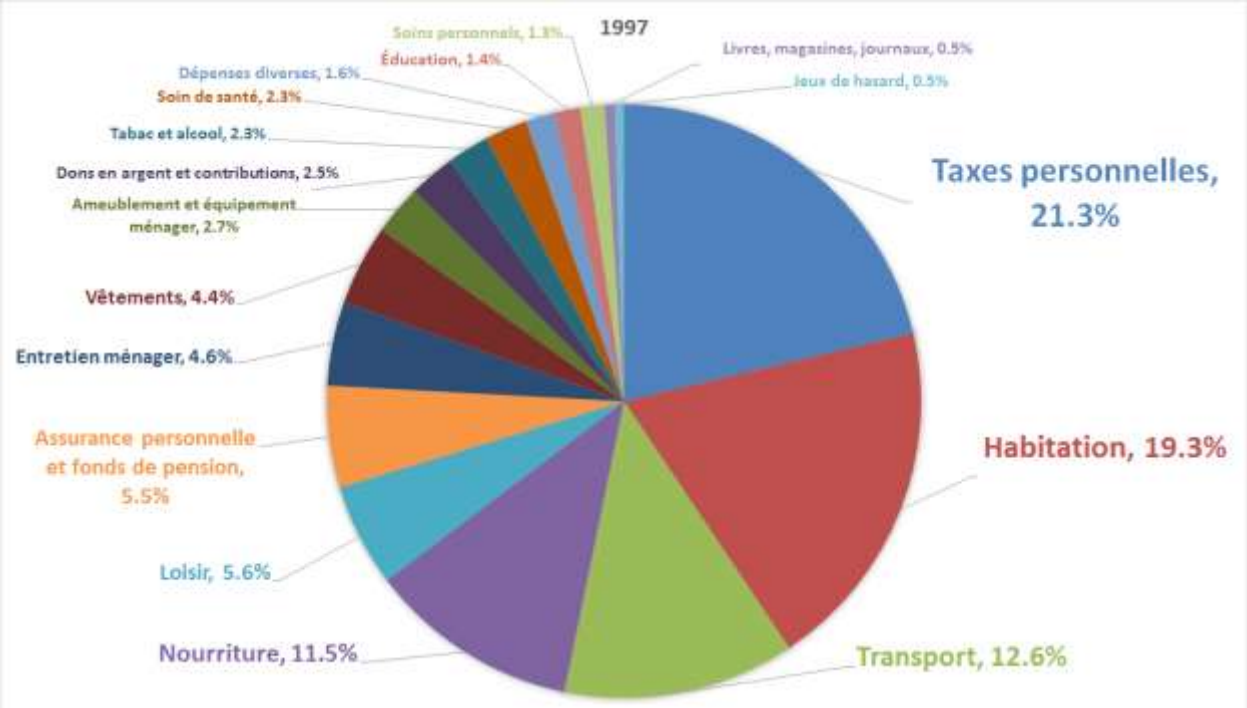
Au niveau canadien, les ménages habitants dans une zone rurale ont une dépense moyenne en électrcité pour la résidence principale plus importante que les ménages habitant en zone urbaine. Cette écart semble s'accroître avec le temps. La situation est moins nette quant aux dépenses pour l'essence :



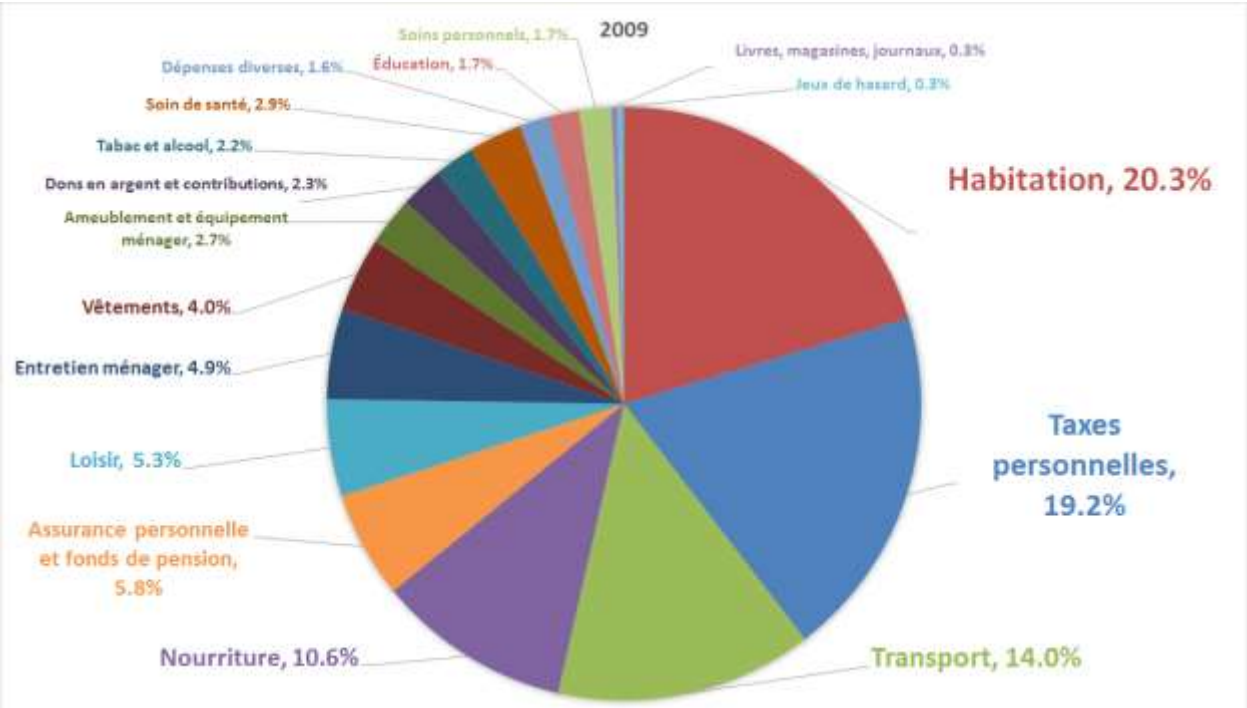
2.4 La part des dépenses d'énergie dans le total des dépenses d'un ménage

2.4.1 Au niveau canadien

2.4.1.1 1997, dépenses totales moyennes d'un ménage canadien / total = 48 976$



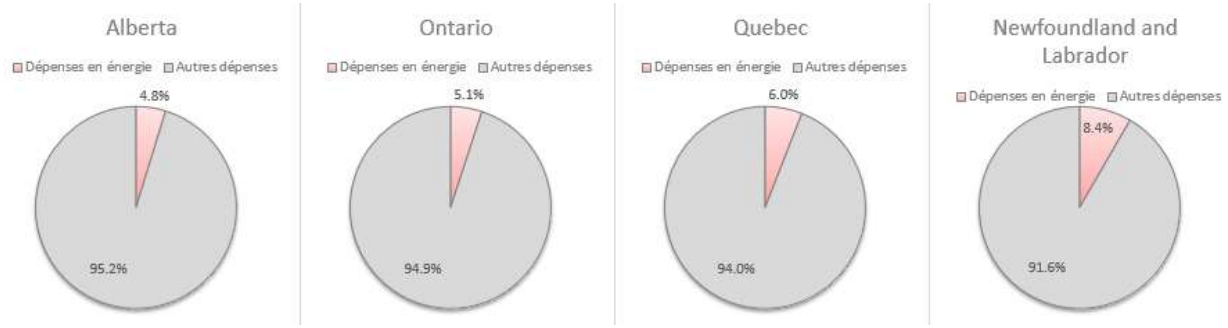
2.4.1.2 2009, dépenses totales moyennes d'un ménage canadien / total = 67 580$



Au niveau canadien, le trois premiers types de dépenses des ménages sont, dans l'ordre : les dépenses pour l'habitation (dont G024 et G022), celles liées aux taxes personnelles et en troisième position le transport pour l'année 2009. Le changement notable entre 1997 et 2009 est que les taxes personnelles étaient le premier poste de dépense en 1997. En 2009, ce sont les dépenses pour l'habitation qui forment le premier poste de dépense des ménages canadiens.

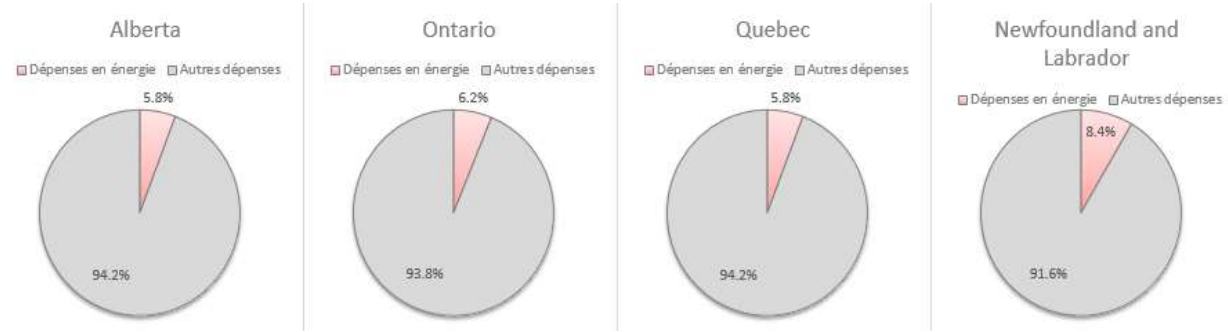
2.4.2 Au niveau des provinces

2.4.2.1 1997

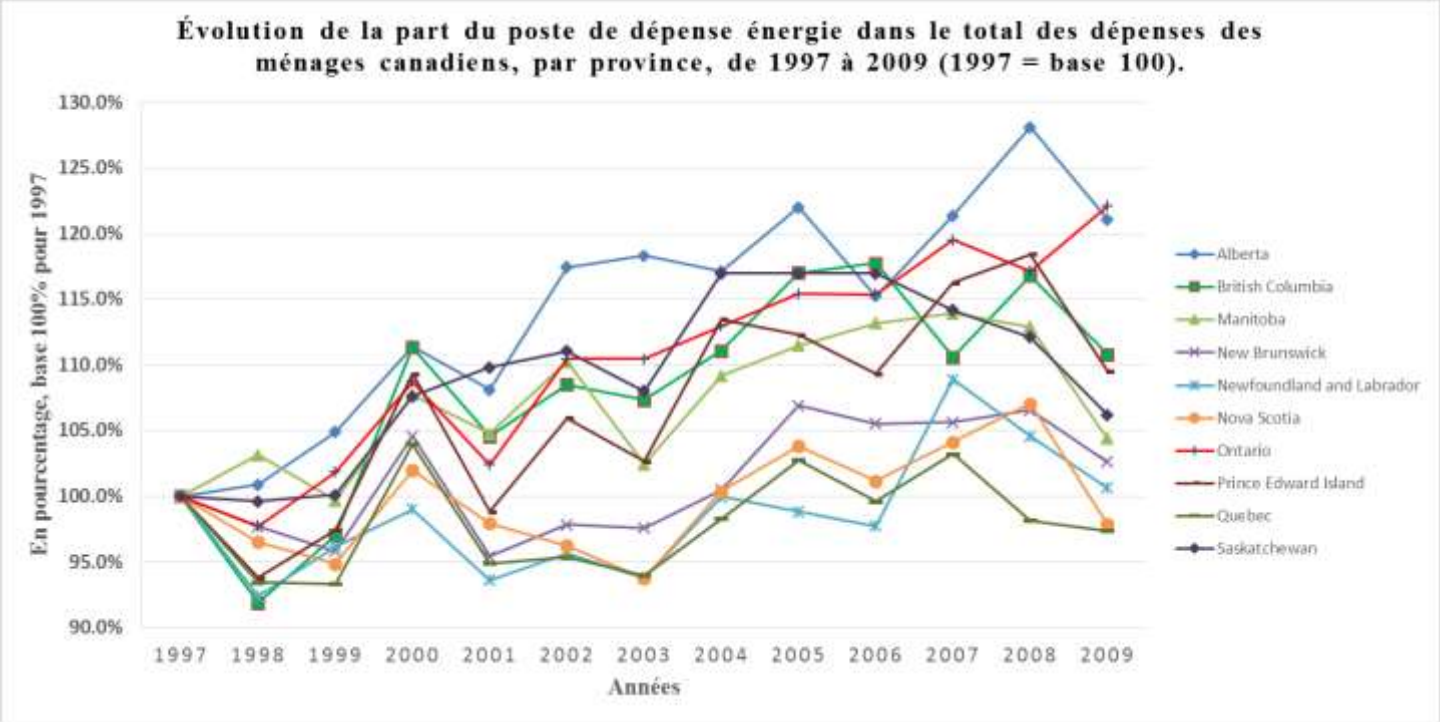


En 1997, la part du poste de dépense énergie dans le budget des ménages canadiens représentait entre 4.7% (la plus faible pour la Colombie-Britannique) et 8.4% (la plus élevée pour Terre-Neuve-et-Labrador) suivant les provinces.

2.4.2.2 2009



En 2009, c'est toujours à Terre-Neuve-et-Labrador que la part du budget des ménages consacrée à l'énergie est la plus importante, soit 8.4%, comme en 1997. L'évolution de la part occupée par le poste énergie dans le total des dépenses des ménages canadiens sur la période 1997 – 2009 connaît des variations diverses selon les provinces. Afin d'illustrer ce propos, le graphique ci-dessous met en lumière les particularités provinciales, en prenant comme année de référence 1997 (base 100%).



Les ménages[20] de l'Alberta, de l'Ontario et la Colombie-Britanique connaissent une hausse régulière de la part occupée par les dépenses en énergie dans leur budget alors que pour les ménages du Québec il s'agit d'une légère tendance à la baisse entre 1997 et 2009.

[20] Afin de mettre les choses en perpective, rappelons qu'en 2014, la population de l'Ontario représente 38% de la population canadienne, le Québec 23%, la Colombie- Britannique 13% et l'Alberta 12%. Toutes les autres provinces représentent moins de 5% de la population canadienne.

3 Méthodologie pour la construction et l'analyse du portrait des ménages québécois

La courte analyse descriptive présentée dans la partie 2.3.2 du rapport a permis d'analyser les trois variables clés pour comprendre les dépenses en énergie des ménages canadiens : **les dépenses en électricité et en gaz pour la résidence principale**, et les dépenses en énergie pour le transport. La revue de littérature présentée à la partie 1.3 a permis d'identifier plusieurs types de segmentation des ménages réalisées en fonction de leur consommation/dépenses en énergie. Celles-ci sont majoritairement manuelles et reposent sur des variables socio-démographiques, économiques ou bien culturo-éthniques à l'exception des études qui portent sur des données horaires de consommation électrique issuent de compteurs intelligents.

L'objectif de ce rapport est de réaliser un regroupement des ménages en fonction de leurs dépenses en énergie, un regroupement qui répondra à l'exigence suivante : **les groupes de ménages doivent être homogènes du point de vue de leurs dépenses en énergie**.

Puis, une analyse de ces groupes sera effectuée afin de décrire et de caractériser les différents profils types de ménages-acheteurs d'énergie (réalisation d'une typologie des ménages canadiens en fonction de leurs dépenses en énergie).

Par la suite, la question du choix du nombre de groupes de ménages du regroupement sera discutée en détail, mais il important de signaler, en terme d'objectif, que celui-ci doit être un astucieux compromis entre la simplification du réel afin de permettre sa resprésentation (humainement intelligible) et la préservation de la variabilité des profils types de ménages (richesse du réel).

3.1 L'analyse de regroupement[21] (« clustering analysis »)

L'analyse de regroupement est une méthode de type descriptif des données afin de réaliser un regroupement des observations dans le but d'identifier une structure dans les observations. Le terme anglais utilisé est « clustering analysis ». En français, plusieurs définitions sont utilisées : analyse de regroupement, classification automatique (dit plus simplement classification). Voici

[21] En marketing, l'analyse de regroupement est très souvent utilisée sous le nom de segmentation afin d'identifier des profils de consommation. « On lui donne aussi le nom de typologie ou analyse typologique. En médecine on parle de nosologie. En biologie et zoologie, on parle de taxinomie ou taxonomie (…). Enfin, les spécialistes de réseaux de neurones parlent de reconnaissance de forme non supervisée » (Tufféry, 2012, pp. 265, 266).

une définition proposée par Stéphen Tufféry, dans *Data Mining et statistque décisionnelle, l'intelligence des données* : « La classification est l'opération statistique qui consiste à regrouper des objets (individus ou variables) en un nombre limité de groupes, les classes (ou segments, ou clusters), qui ont deux propriétés. D'une part, ils ne sont pas prédéfinis par l'analyste mais découverts au cours de l'opération, contrairement aux classes de classement. D'autre part, les classes de classification regroupent les objets ayant les mêmes caractéristiques similaires et séparent les objets ayant les caractéristiques différentes (homogénéité interne et hétérogénéité externe), ce qui peut être mesuré par des critères tels que l'inertie inter-classe (inter-groupe, inter-cluster) » (Tufféry, 2012, p. 265).

Le regroupement peut se faire sur la base d'une probabilité. Dans ce cas, une observation peut être associée à plusieurs groupes. Dans le cas d'un regroupement par partition, une observation doit appartenir à un groupe uniquement. **C'est ce second type de regroupement qui sera utilisé dans cette analyse.**

Les critères pour effectuer les regroupements sont nombreux. Le plus souvent, c'est le critère de la ressemblance des observations qui est utilisée : l'objectif est de créer des groupes d'observations homogènes selon leurs caractéristiques, ou certaines caractéristiques sélectionnées en fonction de la problématique en question. Ce critère implique une mesure de dissemblance afin de quantifier la distance entre deux observations. Cette mesure de distance donne une grande importance à l'échelle utilisée et aux choix des variables utilisées pour décrire les observations.

Pour effectuer une analyse de regroupement, il existe deux grandes familles de techniques de regroupement. La première rassemble l'ensemble des techniques dites hiérarchiques qui procèdent par fusions, pour les méthodes ascendantes, ou bien par séparation pour les méthodes descendantes. La deuxième grande famille rassemble les techniques dites non-hiérarchiques où le nombre de groupes est choisi a priori et un algorithme optimise la répartition des observations au sein des groupes en optimisant un critère.

Enfin, dans ce type d'analyse la question de l'interprétation des groupes formés est essentielle et peut guider dans le choix des méthodes et des variables sélectionnées pour l'anlayse.

3.1.1 Le choix des variables : bruits et incertitudes

Le choix des variables à utiliser dans la construction des groupes est une étape cruciale. En effet, le choix des variables aura un impact certain sur les résultats obtenus. Deux aspects sont à prendre en compte.

Premièrement, l'utilisation des variables qui ne sont pas reliés à la problématique se traduira par une baisse de la qualité du regroupement. En effet, l'utilisation d'un premier groupe de variables pertinentes pour l'étude conjointement à un deuxième groupe de variables qui ne le sont pas aura pour conséquence que les groupes formés reflèteront auant les variations selon le premier groupe que le second groupe de variables. L'information fournie par le premier groupe de variables (pertinentes) pour construire les groupes sera « éparpillée » dans les informations sans intérêts fournies par le second groupe de variables. Ce point marque l'importance du choix des variables pour ne pas ajouter du bruit, de l'incertitude au modèle. Dans la présente analyse où plus de 250 variables sont disponibles il ne faut pas perdre de vue l'objectif : avoir des groupes de ménages **homogènes en termes de dépenses en énergie pour la résidence principale**.

Le critère utilisé pour définir l'homogénéité d'un groupe reposera sur les variables des dépenses en énergie des ménages qui composent le groupe, c'est-à-dire **les dépenses en électricité pour la résidence principale du ménage (G024) et les dépenses en combustibles liquides (G022)**, toujours pour les résidences principales du ménage. Aussi, afin d'affiner les groupes formés, une autre variable a été sélectionnée/créée (EN_RE) : la part que représente les dépenses en électricité pour la résidence principale sur le total des revenus du ménages avant taxes. L'objectif est de rajouter de l'information pour que des ménages qui ont des dépenses en électricité[22] semblables mais des revenus dissemblables soient séparés.

	Nom de la variable	**Description**
Variables choisies pour effectuer l'analyse de regroupement	G024	Électricité pour résidence principale
	G022	Combustibles liquides (dont gaz)
	EN_RE = G024 / HHINCTOT	Revenu du ménage avant taxe

[22] Une autre variable aurait pu être ajoutée également pour réaliser cette séparation entre les groupes de ménages qui ont des dépenses semblables en gaz pour la résidence principale et des revenus dissemblables, mais l'analyse souhaite se concentrer sur la question des dépenses en électricité.

Deuxièment, il faut éviter de sélectionner des variables qui mesurent la même information[23] car cette multiplicité exagèrera le poids de la même information fournie par ces variables et donc la mesure de dissemblance entre deux observations reflètera d'avantage la ressemblance ou la dissemblance relativement à cette information. Ce point fait écho à la notion d'échelle que nous allons présenter ci-dessous.

3.1.2 Le choix de l'échelle

L'échelle des variables est à considérer très sérieusement lors d'une analyse de regroupement. En effet, par exemple si une variable quantitative Var1 a varie de 0 à 150.000 et une variable Var2 varie de 1 à 5, inclure ces deux variables revient à faire une analyse de la Var1 et d'ignorer toute information issue de la variable Var2. Plus l'échelle d'une variable est grande, plus cette variable aura d'importance dans le regroupement. C'est le cas pour les variables G024, G022 et EN-RE. Sans standardisation[24] de ces variables préalables, une importance considérable sera accordée à G024 et G022[25] au détriment de EN_RE qui sera ignorée. Dans la suite de l'analyse, les variables G024, G022 et EN_RE sont standardisées pour la construction des groupes.

3.1.3 Le choix de la mesure de distance (mesure de dissemblance) :

Suivant le type de variable, binaire ou continue, des mesures de distances spécifiques seront préférées. Les variables G024, G022 et EN_RE sont continues. La mesure de distance la plus habituelle pour ce type de variable[26] est la distance Euclidienne. Celle-ci se calcule en prenant la racine carré de la somme des carrés des différences de valeurs entre chaque variable pour chaque observation.

La procédure suivie dans la suite de la construction du regroupement comporte deux étapes : pour la première étape, utilisation d'une méthode de regroupement hiérarchique afin d'identifier un nombre optimal de groupes, puis pour la deuxième étape, utilisation d'une méthode de regroupement non hiérarchique afin de raffiner le regroupement obtenue à l'étape 1.

[23] Dans ce cas, l'analyse en composante principale peut-être utile pour identifier ce type de biais.
[24] Standardiser une variable signifie transformer une variable de sorte que sa moyenne est égale à 0 et une variance de 1.
[25] Voir les différences d'échelle pour les variables G024 et G022 dans la partie descriptive du rapport dans les parties 2.3.2.1 et 2.3.2.2.
[26] Pour des variables binaires, la distance de Hamming sera préférée.

3.1.4 Première étape : une méthode hiérarchique

Dans un premier temps, une analayse de regroupement hiérarchique (descendante) sera réalisée afin d'identifier le nombre de groupe qui sera à la fois un bon compromis entre des groupes de ménages parfaitement homogènes en termes de dépenses en énergie (la solution parfaite étant un groupe par ménage) et un nombre limité de groupes (plus facile à analyser, à visualiser pour l'être humain).

Une méthode hiérarchique ascandante débute avec n groupes, n étant le nombre d'observations (nombre de ménages), puis regroupe les ménages ou groupes formés au préalable **un à un** en choissisant les deux ménages ou groupes dont la distance est la plus faible (c'est-à-dire les deux groupes ou ménages qui occasionnent la plus petite perte d'homogénéité globale, ou bien la plus petite hausse d'hétérogénéité au sein du groupe) jusqu'à ce que tous les ménages ne forment qu'un seul groupe. **Cette algorithme procède de manière récursive sans jamais de remise en question des choix précédents, ce qui peut avoir pour conséquence que la partition trouvée ne soit pas optimale selon un critère donné.** Cette méthode requiert que l'on calcule la distance entre deux groupes. Le critère de Ward sera retenu : critère d'homogénéité global des groupes où l'on cherche à minimiser la sommes des carrés des distances entre observations appartenant à un même groupe. Cette distance entre observations au sein d'un groupe est calculée par la somme des carrés des distances entre les observations et la moyenne (appelé aussi centroïde) du groupe. Plus cette somme des distances, pour tous les groupes de la partition est petite, plus les groupes sont dit homogènes.

3.1.5 Représentation visuelle de la méthode hiérarchique ascendante : le dendogramme

Le dendogramme illustre la méthode hiérarchique ascendante. Sur l'axe des abscisses, on représente les groupes. Sur l'axe des ordonnées on indique la valeur du critère (ici celui de ward, somme des distances au carrés pour tous les groupes) pour les groupes nouvellement formés jusqu'au dernier. Cette représentaion peut permettre d'identifier une partition donnée, associées à un nombre de groupes pour lequel si nous devions procéder à une fusion de deux groupes de cette partition cela se traduirait par une augmentation importante de la valeur du critère de ward.

Ainsi, après avoir réalisé une analyse hiérarchique ascendante avec **les variables G024, G022 et EN_RE**, le dendogramme de la Figure 3-1 permet d'identifer plusieurs palliers pour des hausses importantes du critère de ward.

2009, Québec : Analyse de regroupement hiérarchique

Dendogramme

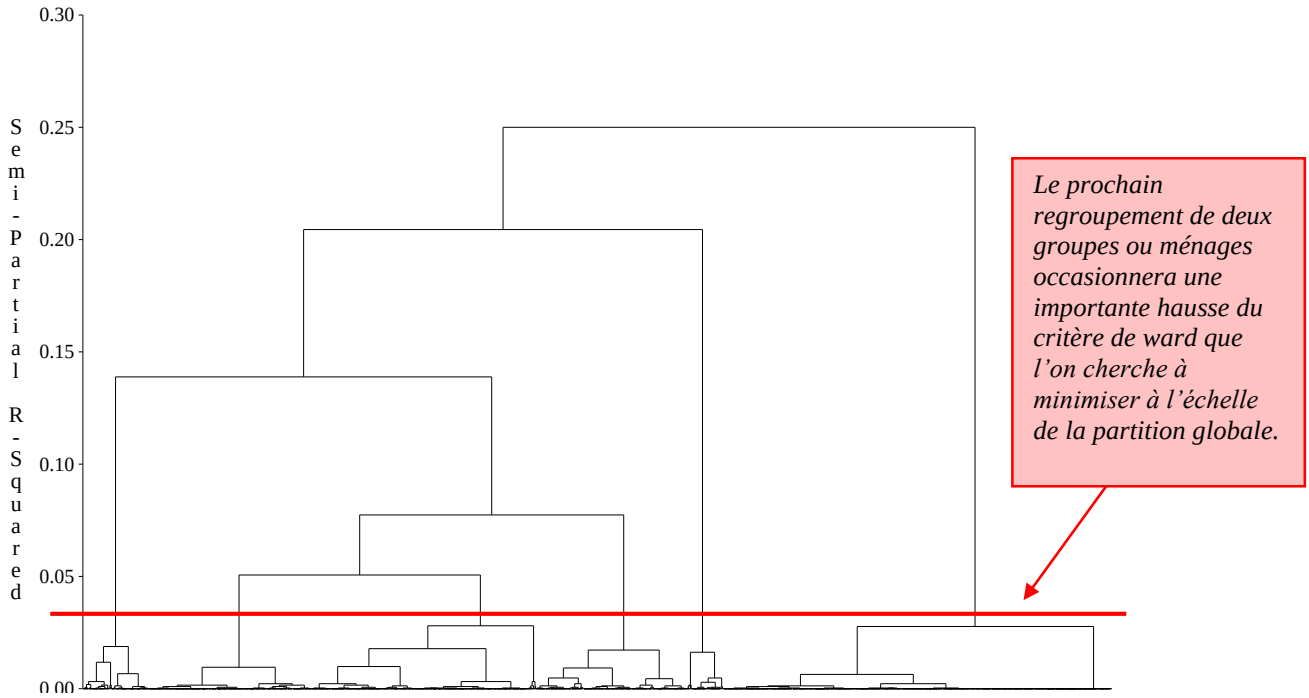


Figure 3-1 : Dendogramme pour l'analyse de regroupement hiérarchique (Québec, 2009)

3.1.6 Choix du nombre de groupes (analyse graphique et critères statistiques)

Afin de choisir le nombre optimal de groupes, le dendogramme peut servir de guide « visuel » comme l'explique la partie 3.1.5. Sur un plan plus quantitatif, le R-squared (RSQ) et le semi-partial R-Squared peuvent être des mesures statistiques utilisées dans la sélection du nombre optimal de groupes.

- **RSQ : mesure le degrés d'homogénéité des groupes globalement (proportion de la varaibilité, des variables sélectionnées pour l'analyse de regroupement, qui est expliquée par les groupes).**

Soit :

- i, un groupe
- j, une observation et
- k, la varaible k de l'observation j

alors :

- $SSE = \sum_i \sum_j \sum_k \left| x_{ijk} - \bar{x}_{i.k} \right|^2$
- $SST = \sum_i \sum_j \sum_k \left| x_{ijk} - \bar{x}_{..k} \right|^2$ et
- $R^2 = \frac{SST - SSE}{SST}$

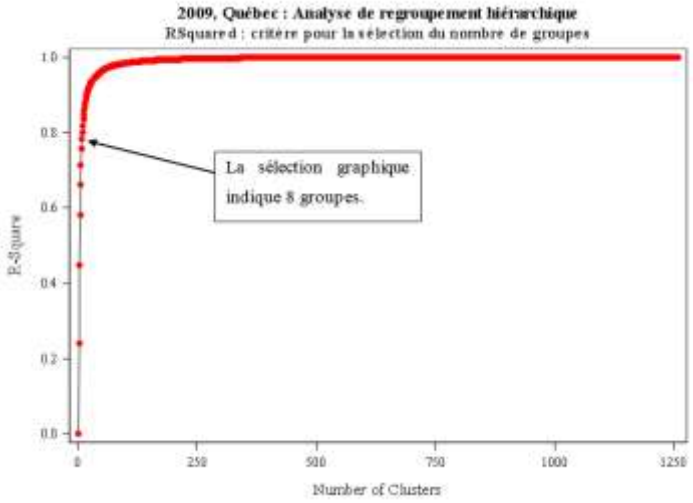


Figure 3-2 : R-squared, Québec, 2009.

- **SPRSQ : mesure la perte d'homogénéité résulant de la fusion de deux groupes.**

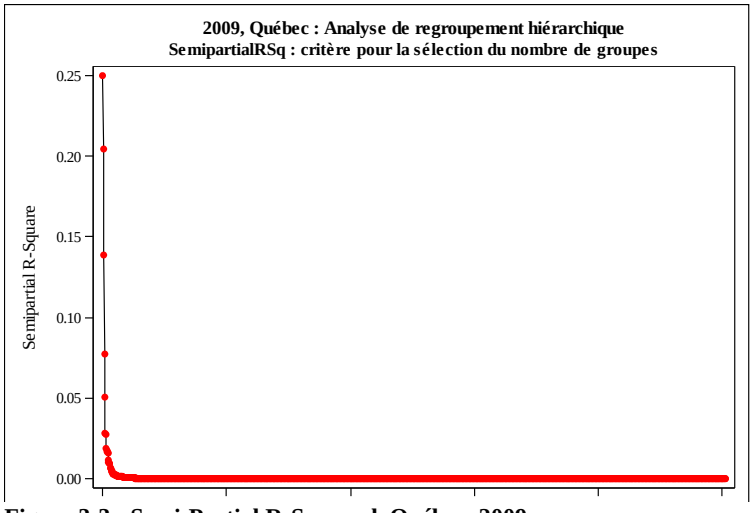


Figure 3-3 : Semi-Partial R-Squared, Québec, 2009

D’après la Table 1, nous pouvons observer une perte d’homogénéité importante lors du passage d’un regroupement de 8 à 7 groupes (ce passage de 8 à 7 groupes a pour conséquence une augmentation de 31% de de la perte d’homogénéité). L’analyse hiérarchique permet d’identifier ce regroupement précis à 8 groupes pour lequel le passage de 8 à 7 occasionne une perte conséquente d’homogénéité globale. C’est la raison pour laquelle il est opportun de déterminer le nombre de groupes à 8, plutôt qu’à 7 ou bien un nombre inférieur.

Table 1 : Historique du regroupement, Québec, 2009 (Sortie SAS)

	Cluster History								
Number of Clusters	**Clusters Joined**		**Freq**	**Semipartial R-Square**	**% augmen-tation**	**R-Square**	**Approximate Expected R-Square**	**Cubic Clustering Criterion**	**Tie**
1259	OB6	OB7	2	0.0000		0.0000	1.00	.	T
1258	CL1259	OB22	3	0.0000		0.0000	1.00	.	T
	….	….	….	….		….	….	….	….
	….	….	….	….		….	….	….	….
10	CL35	CL19	45	0.0163	16%	0.817	0.805	3.57	
9	CL14	CL20	184	0.017	4%	0.8	0.791	2.65	
8	**CL119**	**CL25**	**393**	**0.0176**	**3%**	**0.783**	**0.773**	**2.47**	
7	CL15	CL13	233	0.0254	**31%**	0.757	0.752	1.32	
6	CL12	CL11	135	0.0428	41%	0.714	0.725	-2.3	
5	CL16	CL9	454	0.0538	20%	0.661	0.688	-5.8	
4	CL5	CL7	687	0.0802	33%	0.581	0.638	-11	
3	CL6	CL4	822	0.1333	40%	0.447	0.558	-16	
2	CL3	CL10	867	0.2066	35%	0.241	0.37	-15	
1	CL2	CL8	1260	0.2406	14%	0	0	0	

En utilisant une technique de regroupement hiérarchique (algorithme de WALD), 8 groupes ont été identifiés et offre une image de la population des ménages Québécois pertinentes pour notre analyse selon les mesures statistiques du R-squared et du Semi-partial R-Squared. Cette solution permet d’expliquer 77% de la variabilité totale.

3.1.7 Deuxième étape : une méthode non-hiérarchique, k-means

Comme cela a été mentionné à la partie 3.1.4 un algorithme hiérarchique procède de manière récursive sans jamais de remise en question des choix précédents, ce qui peut avoir pour conséquence que la partition trouvée ne soit pas optimale selon un critère donné. Par conséquent, il peut être opportun de raffiner la solution à 8 groupes fournie lors de première étape en utilisant un algorithme non-hiérarchique tel que k-means par exemle.

L'algortihme k-mean procède en deux temps. Dans une première itération, il définit la valeur des centroïdes de chacun des 8 groupes (valeurs du ménage moyen pondérée en utilisant la variable WEIGHT[27]). Puis, dans une deuxième itération, il affecte chaque ménage au bon groupe, c'est-à-dire au groupe qui minimisera la somme des erreurs associeés aux différents groupes. L'algorithme[28] procède ainsi jusqu'à ce que plus aucun ménage ne soit changé de groupe.

3.1.8 Initialisation de k-means, optimalité et méta-heurisitque

Les considérations sur l'initialisation sont nombreuses. L'étape de l'initialisation est délicate[29]. Elle peut se faire aléatoirement ou bien en utilisant un algorithme hiérarchique par exemple. C'est le choix qui a été fait ici, car elle donne en général d'assez bon résultat. Elle garantit des solutions de qualités raisonnables, mais ne peuvent être considérées comme optimales. En effet, l'alogrithme k-means donne des solutions qui peuvent être des optimums locaux et par conséquent ceux-ci sont très fortement dépendants de la solution initiale utilisée. Une approche qui n'a pas été employée pour le travail présenté dans ce rapport est l'utilisation de métaheurisitques tels que « multistart » ou bien « VNS » afin de sortir des optimums locaux dans lesquels k-means pourrait faire impasse.

3.1.9 Mesure de la qualité d'une partition donnée

Afin de qualifer la qualité d'un regroupement, on peut faire appel à la notrion d'inertie[30], inertie totale, inertie intra-classe et inter-classe.

---

[27] Voir le point sur l'inférence statistique quant à l'utilisation de la variable WEIGHT pour calculer une moyenne à la partie 2.3.1. (noter que lors de la première étape, celle du regroupement hiérarchique, la pondération à l'aide de la variable WEIGHT n'était pas prise en compte, mais le regroupement final issue de k-means lui en tient compte).
[28] Pour une description de l'algorithme, voir annexe **Erreur ! Source du renvoi introuvable.**.
[29] Voir les différentes méthodes d'initilisation pour k-means dams *Initializing K-means Batch Clustering: A Critical Evaluation of Several Techniques*, (Steinley & Brusco, 2007).
[30] Voir annexe 6.5.

3.2 Détection et sélection d'un regroupement (structure) à partir des données EDM pour le Québec.

Visualisation[31] du regroupement en 8 groupes (obtenue avec la méthodologie décrite ci-dessus), pour Québec 2009, en utilisant 5 caractéristiques descriptives (qui donnent la fiche d'identité des groupes) :

- **N° :** correspond au numéro du groupe
- **Taille :** correspond au nombre de ménages qui composent le groupe
- **Revenu :** correspond au revenu moyen des ménages du groupe (HHINCTOT)
- **Électricité :** correspond à la dépense moyenne en électricité des ménages du groupe (G024)
- **Combustibles :** correspond à la dépense moyenne en combustibles pour les ménages du groupe (G022)
- **Électri./Revenu :** correspond au ratio[32] de la dépense moyenne en électricité sur le revenu moyen des ménages du groupe (EN_RE).

REGROUPEMENT en 8 groupes, Québec 2009

N°	Nom	Taille	Revenu	Électricité	Combustibles	Électricté /Rev.
1						
2						
3						
4						
5						
6						
7						
8						

Pour une présentation détaillée des moyennes des variables HHINCTOT (correspondant au revenu total avant taxe), G024, G022 et EN_RE de chaque groupe, voir annexe 6.6.

[31] Dans cette visualisation, lorsqu'une ligne pour une variable est entièrement coloriée, cela indique que ce groupe est le plus important (ce groupe a la dépense moyenne la plus importante) pour cette variable. La ligne de chacun des autres groupes est coloriée relativement à celle-ci.

[32] Ce ratio exprime dans quelle proportion la dépense en électricité pèse sur le revenu du ménage. Plus il est élevé, plus le ménage consacre une part importante de son revenu à ses dépenses en électricité.

3.2.1 Pertinence du regroupement : interprétation des groupes et questionnement fertile

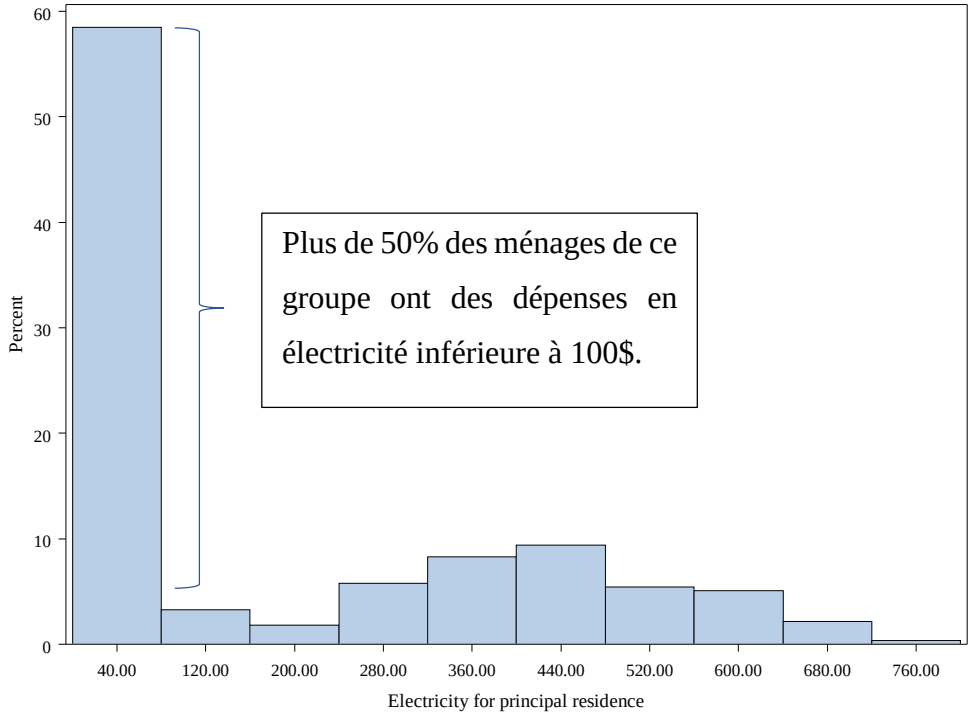
Ci-dessous, une description du regroupement en 8 groupes pour dresser la fiche d'identité des groupes formés et juger de leur pertinence selon deux critères : l'interprétation du regroupement et la richesse des questions issues de l'analyse. Est-ce que les groupes permettent de décrire la réalité de façon riche, d'initier une réflexion, de remettre en question des évidences, des faits et croyances ou encore à rendre palpable une réalité qui ne l'était pas ?

Québec, 2009 : **Groupe 1**						
N°	Nom	Taille	Revenu	Électricité	Combustibles	Électricté /Rev.
1	Électricté : 0…					

Le premier groupe a **une dépense en électricité très faible, voire inexistante** par rapport aux autres groupes. C'est ce qui le caractérise (en plus d'avoir un revenu moyen faible par rapport aux autres). Quelles caractéristiques de ce groupe permettraient d'expliquer ce profil de consommation d'énergie ?

	Revenu	Électricité
Groupe 1	⊖	✖

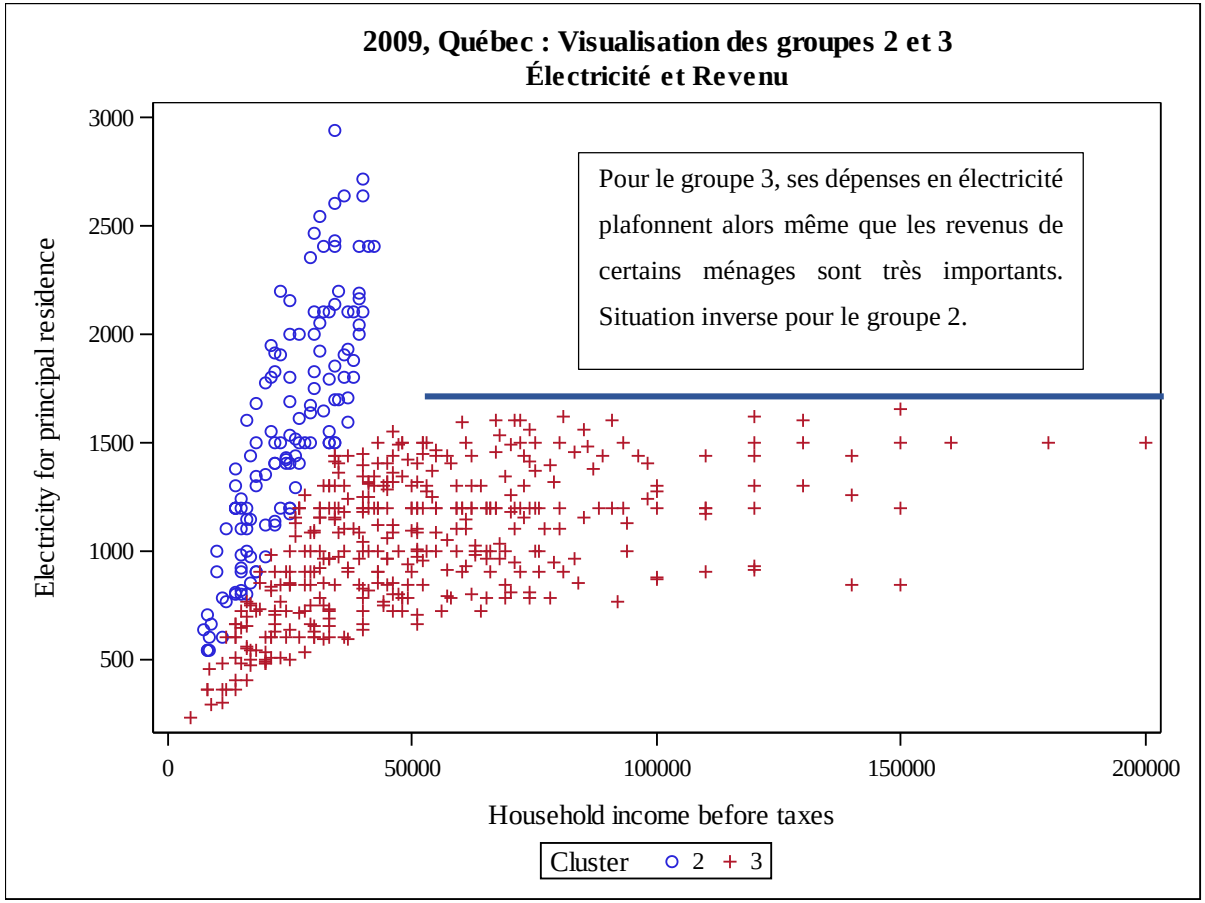
Québec, 2009, Distribution de la dépense en électricité



Québec, 2009 : **Groupe 2 et 3**						
N°	Nom	Taille	Revenu	Électricité	Combustibles	Électricté /Rev.
2						
3						

Les groupes 2 et 3 sont intéressants à mettre en vis-à-vis car ils s'opposent. Le groupe 2 a un revenu moyen plus faible que le groupe 3 mais a une dépense moyenne en électricité plus élevée (ce qui se traduit par un ratio Électricité/Revenu élevé pour le groupe 2).

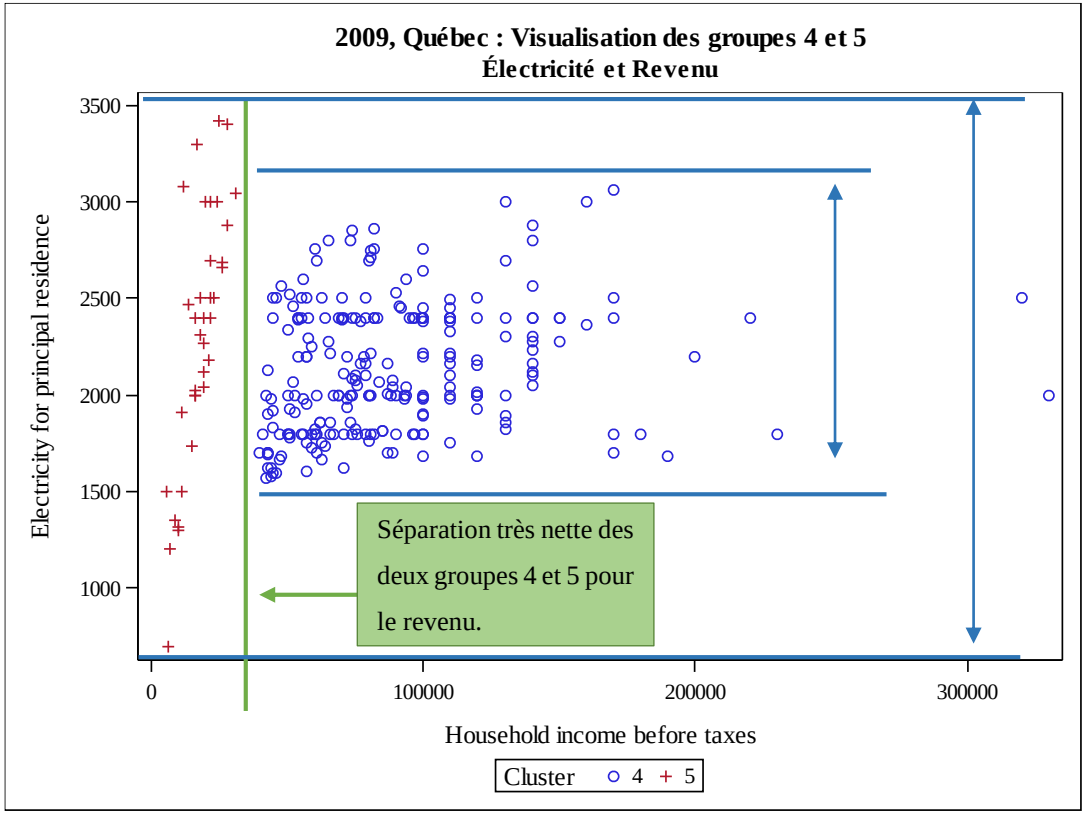
	Revenu	Électricité
Groupe 2	−	+
Groupe 3	+	−



Québec, 2009 : **Groupe 4 et 5**						
N°	Nom	Taille	Revenu	Électricité	Combustibles	Électricté /Rev.
4						
5						

Les groupes 4 et 5 sont similaires en termes de dépenses en électricité. Mais ils sont à l'opposé en termes de revenu. Quelles caractéristiques communes expliqueraient cette similitude du point de vue des dépenses en électricité en dépit de la différence abissale de revenu ?

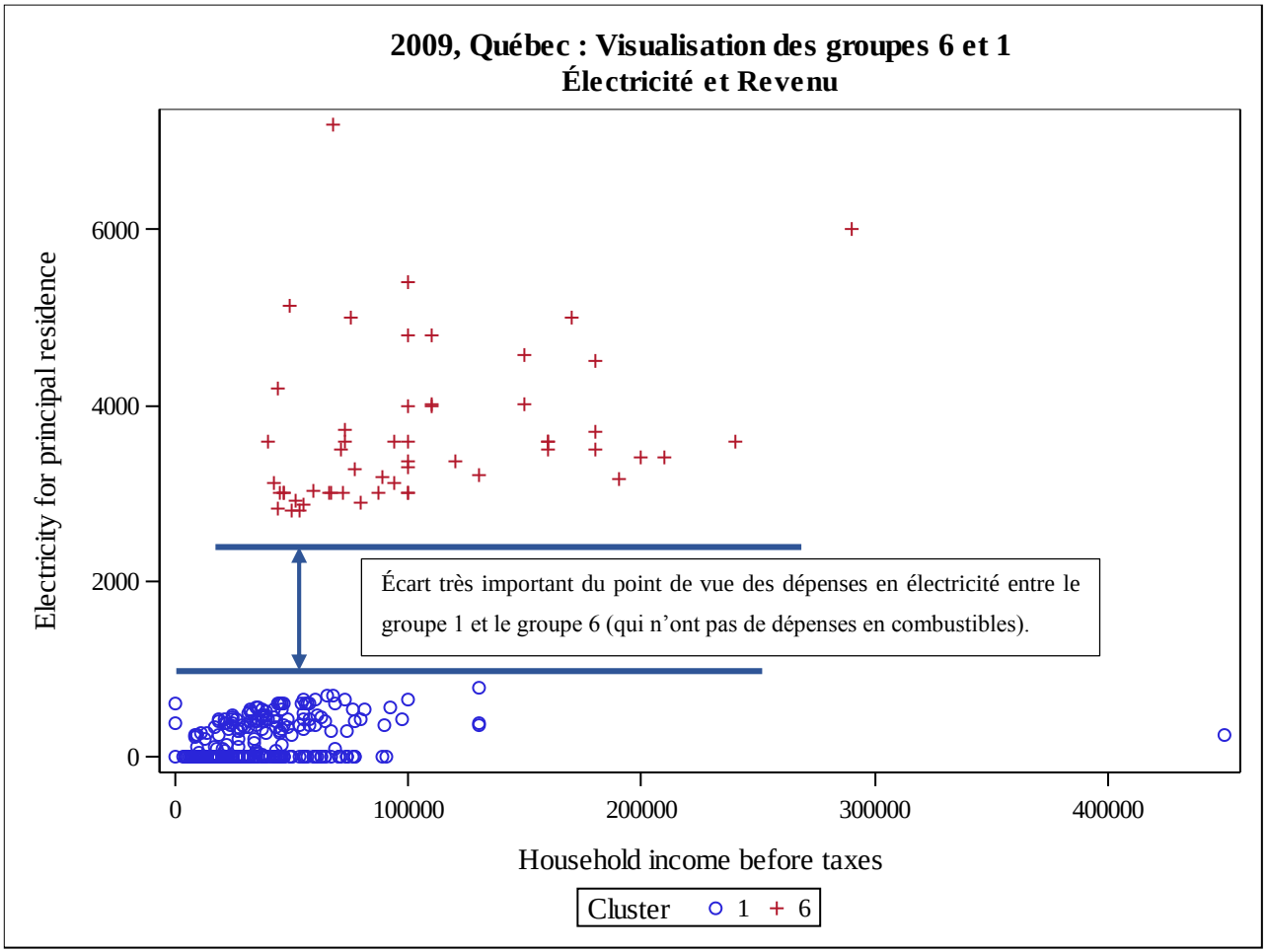
Nous pouvons observer une plus grande variabilité des dépenses en électricité pour le groupe 5. Mais l'analyse graphique illustre bien que les dépenses en électricité de ces deux groupes évoluent dans des intervalles similaires (il n'y a pas de séparation) alors que pour le revenu la séparation est nette.



REGROUPEMENT en 8 groupes, Québec 2009						
N°	Nom	Taille	Revenu	Électricité	Combustibles	Électricté /Rev.
6						

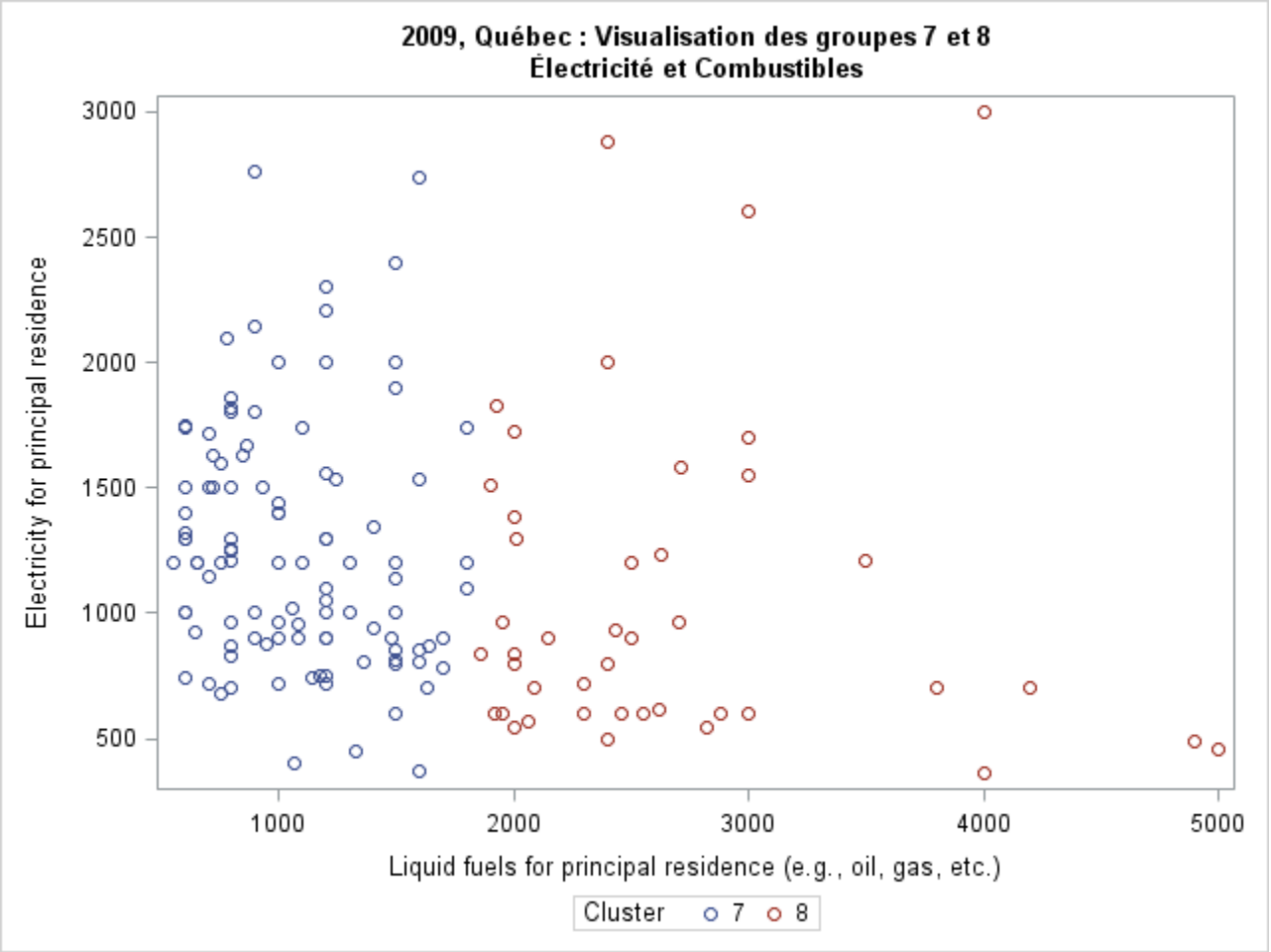
Le groupe 6 est un groupe « extrême » : le plus haut revenu moyen et la plus haute dépense moyenne en électricité entre tous les autres groupes. Il sera intéressant de le comparer avec le groupe 4. Le groupe 6 est l'opposé du groupe 1, comme l'illustre la représentation graphique.

	Revenu	Électricité
Groupe 6	⊕ ⊕	⊕ ⊕



REGROUPEMENT en 8 groupes, Québec 2009						
N°	Nom	Taille	Revenu	Électricité	Combustibles	Électricté /Rev.
7						
8						

Et enfin, les groupes 7 et 8, qui sont des groupes de ménages à haut revenu et qui consomment, en plus de l'électricité, du gaz pour la résidence principale. Nous pouvons observer une très nette différence entre le groupe 8 et le groupe 7 en termes du niveau de consommation de gaz. Pourquoi ces groupes de ménages ont-ils un profil consommateur de gaz (contrairement à tous les autres) ? Quelles caractéristiques les distinguent pour comprendre la différence du niveau de consommation en gaz ?



3.2.2 Pertinence du regroupement : possibilité de suivre ce regroupement dans le temps

Les données (les observations, les ménages) EDM sont indépendantes, aussi bien pour une année que pour les années entre elles. C'est-à-dire qu'il ne s'agit pas des mêmes ménages interrogés chaque année de 1997 à 2009 (ce qui serait le cas pour des données dites de « panel » ou bien des données longitudinales). L'objectif de pouvoir suivre l'évolution du regroupement choisi dans le temps requiert la possibilité de pouvoir déterminer, pour chaque année de 1997 à 2009, à quel groupe (un groupe correspondant à un profil type spécifique de dépenses en énergie) appartient chaque ménage de chaque année.

Pour pouvoir suivre l'évolution au cours de ces 13 années (de 1997 à 2009), le référent doit être le regroupement[33]. Ce qui identifie les 8 groupes détectés pour l'année 2009 sont les centroïdes de chacun de ces groupes. Un centroïde se définit comme la moyenne des coordonnées des observations du groupe en question. Pour faire une métaphore, le centroïde est en quelque sorte l'ADN du groupe. Par exemple, pour le groupe 8, son centroïde est donné dans le Table 2.

Table 2 : Québec, 2009, centroïde

Québec, 2009, groupe 8 (43 observations)	
Variable	**Moyenne**
G024 (Électricité)	1043
G022 (Gaz)	2676
En_Re (Élec/Revenu)	2

En moyenne, les ménages du groupe 8, pour l'année 2009 au Québec, dépensent 1043$ en électricité et 2676$ en gaz pour la résidence principale. La part de leur dépense en électricité représente en moyenne 2% du revenu total du ménage.

Afin d'identifier les ménages de l'année 2008 qui ont ce profil de dépenses en énergie, il convient d'ajuster le centroïde du groupe 8 pour que celui-ci reflète la valeur que devrait avoir le centroïde du groupe 8 de l'année 2008. C'est-à-dire, prendre en compte les hausses ou baisses des prix de l'électricité (G024 [34]), du gaz (G022[35]) et les hausses et baisses des revenus totaux (pour En_RE

[33] Si les données avaient été des données longitudinales, soit les mêmes ménages interrogés chaque année, alors le référent aurait pu être les ménages et l'analyse aurait pu porter sur l'évolution du profil de consommation de ces ménages.
[34] Voir annexe 6.4.5.
[35] Sources des données : le ministère de l'Énergie et des Ressources naturelles et Statistique Canada ("MERN - Prix du gaz naturel," n.d.).

= G024/HHINCTOT[36]) afin de calculer la valeur « théorique » du centroïde du groupe 8 pour l'année 2008. Puis, en applicant la méthode de construction du regroupement non-hiérarchique (k-means) décrite dans la partie 3.1.7 de ce rapport, chaque ménage de l'année 2008 est assigné à l'un des 8 groupes (détecter à partir des données de EDM, 2009). Ainsi, les ménages de l'année 2008 qui ont un profil de dépenses en énergie semblable au groupe 8 de l'année de l'année 2009 sont assignés eux aussi au groupe 8 de l'année 2008.

À titre d'illustration, voici le centroïde du groupe 8 de l'année 2009 ajusté pour l'année 2008, voir Table 3.

Table 3 : Québec, 2009, centroïde ajusté pour l'année 2008

Québec, 2009, groupe 8 (43 observations)		
Variable	**Moyenne**	**Note, en 2008 (par rapport à 2009)**
G024	1033	Le prix de l'électricité était 0.9% inférieur.
G022	2874	Le prix du gaz était 7% plus élevé.
En_Re	2	Le revenu total moyen des ménages du premier quintile était 6% inférieur alors que celui du 5ème quintile était 5% supérieur.

Plus précisement, afin de construire le regroupement de l'année 2008, la méthode est la suivante :

1. Initialisation de l'algorithme K-means à partir des centroïdes de l'année 2009 ajustés pour tenir compte de l'évolution des prix de l'électricité, du gaz et de l'évolution des salaires entre 2009 et 2008.
2. Lors de la première itération de K-means, chaque ménage de l'année 2008 est assigné au groupe de ménages dont il est le plus semblable.
3. Puis, le centroïde de chaque groupe est mis à jour pour tenir compte des ménages de 2008 qui désormais le composent.

[36] Pour le revenu total avant impôt du ménage (HHINCTOT) celui-ci est ajusté par quintile de revenu, c'est-à-dire en tenant compte de la hausse ou la baisse moyenne du revenu par quintile d'une année à l'autre calculée à partir des données EDM("Enquête sur les dépenses des ménages (EDM)," n.d.).

4. Lors des itérations subséquentes (jusqu'à ce que l'algorithme converge, c'est-à-dire que chaque ménage soit assigné à un groupe de tel sorte que le critère d'homogénéité est maximisé[37]) en plus de la réassignation d'un ménage à un nouveau groupe (dont il est désormais le plus proche), ce sont les centroïdes des groupes qui varient eux-aussi, permettant ainsi de suivre l'évolution d'un groupe sur la période considérée, soit de 1997 à 2009.

La difficulté tient alors dans le fait de détecter et de choisir une structure de regroupement qui persiste dans le temps. En plus des critères de l'interprétation ou bien de la richesse du questionnement quant au regroupement identifié, il est important de s'asssurer que ces groupes ne soient pas particuliers à une année donnée. Deux causes possibles pourraient expliquer l'identification d'une structure particulière à une année. La première est d'ordre algorithmique. Utilisant l'aglorithme k-means pour former le regroupement (sans utilisation d'une méta-heuristique de type VNS[38]), il se peut que celui-ci bute sur un optimum local, particulier et qui ne peut servir de référence pour des comparaisons avec les autres années. Le deuxième relève des « valeurs extrêmes ». Si certains ménages, un petit nombre, ont une particularité qui ne se retrouvent pas d'année en année, alors fixer cette particularité comme référence nuira aux comparaison d'année en année.

Les caractéristiques principales du regroupement en 8 groupes pour l'année 2009 ont été présentées à la partie 3.2.1 de ce rapport. Ci-dessous, c'est la comparaison entre 2009 et 2008 de ce regroupement qui est présenté et la validation ou non de ce regroupement pour faire le suivi sur la période 1999 – 2009.

[37] Le critère d'homogénéité des groupes du regroupement est la minimisation de la somme des carrés des distances des centroïdes aux observations, ici cette distance se calcule à partir des trois variables G024, G022 et En_Re (G024/HHINCTOT), avec standardisation de ces variables.
[38] VNS pour Variable Neighbourhood Search (Hansen, Mladenovic, & Moreno pérez, 2010).

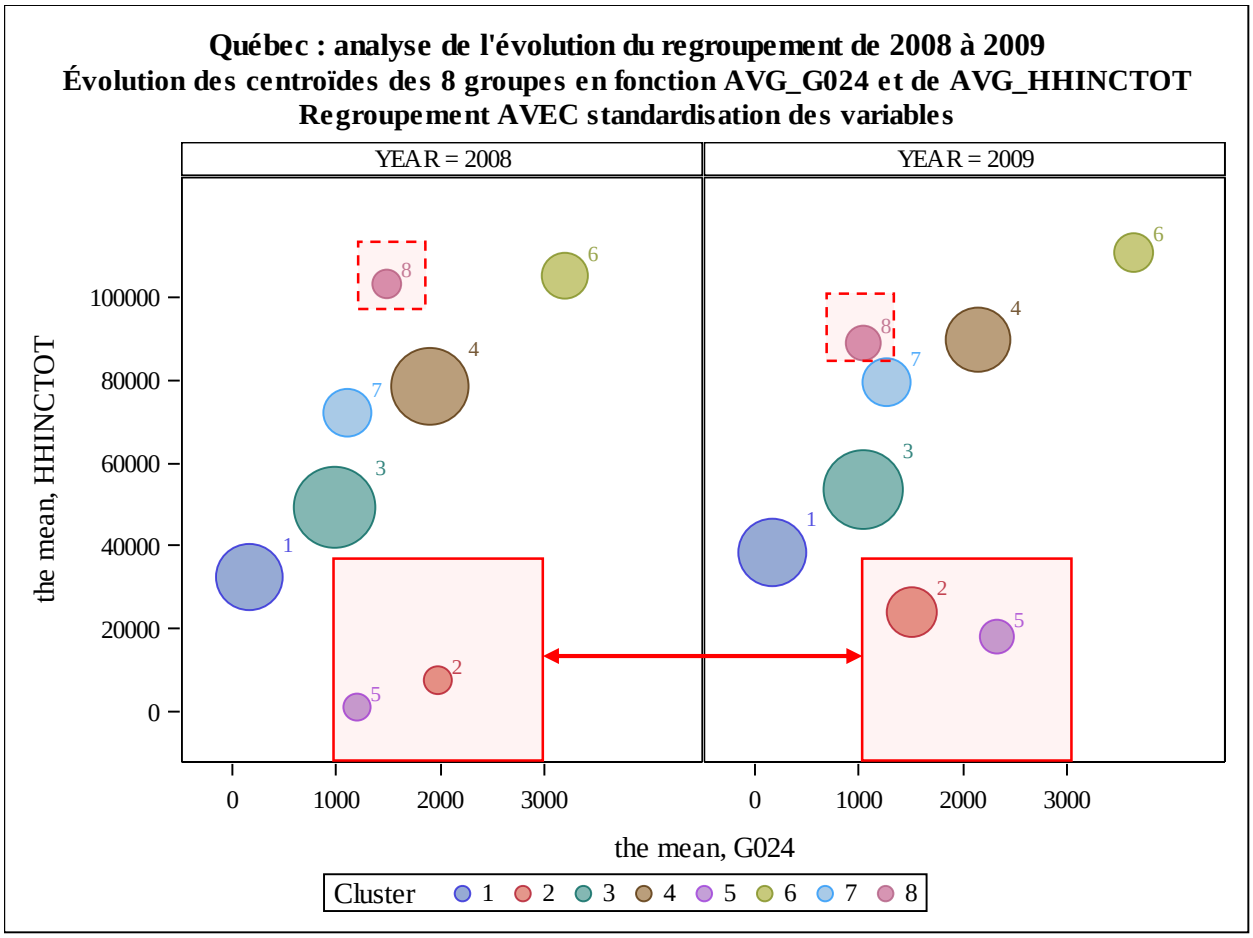
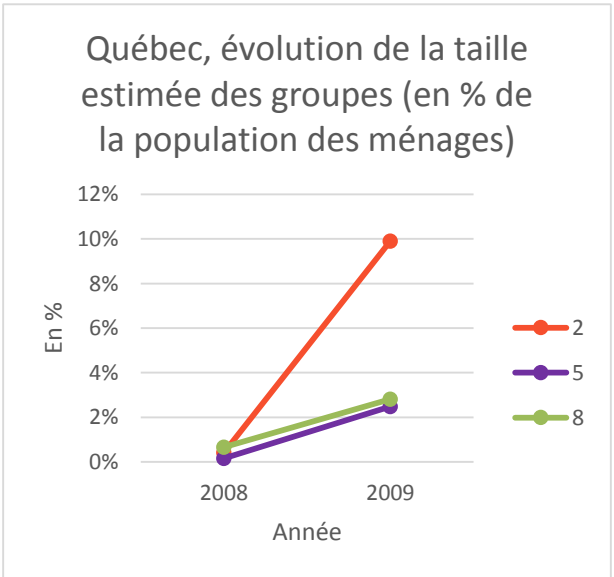


Figure 3-4 : Québec, Comparaison du regroupement entre 2009 et 2008. La taille des bulles est proportionnelle à la taille estimée des groupes pour l'année en question. Voir aussi annexe 6.6.3.



Tel que le présente la Figure 3-4, suivre les groupes 1, 3, 4, 6, 7 et 8 ne pose guère de difficulté. En revanche pour retrouver les groupes 2 et 5 en 2008, qui se caractérisent par de faibles revenus et une dépense en électricité moyenne supérieure à 1000$, cela s'avère délicat. La première difficulté tient au fait de la petite taille de ces groupes en 2008 : le groupe 2 qui représentait près de 10% des ménages et le groupe 5 près de 2% en 2009 sont insignifiants (non-représentatifs) en 2008. Notons que le groupe 8 est limite au point de sa taille (il représente 2% des ménages en 2009 et presque 0% en 2008). En revanche, son centroïde (ADN) caractérise le même profil en énergie de 2009 à 2008 contrairement aux

groupes 2 et 5. Les centroïdes respectifs de 2008 des groupes 2 et 5 rendent compte avant tout de valeur extrême pour En_Re (ratio entre les dépenses en électricité et le revenu du ménage), soit des ménages qui ont d'importantes dépenses en électricité (supérieures à 1200$ en moyenne) et de faibles revenus totaux (inférieurs à 10.000$ pour le groupe 2 et inférieur à 1200$ pour le groupe 5 en moyenne). Cela se traduit par un ratio En_Re (G024/HHINCTOT) très élevé, 111% pour le groupe 5 et 31% pour le groupe 2 en moyenne (tel que l'illustre la Figure 3-5).

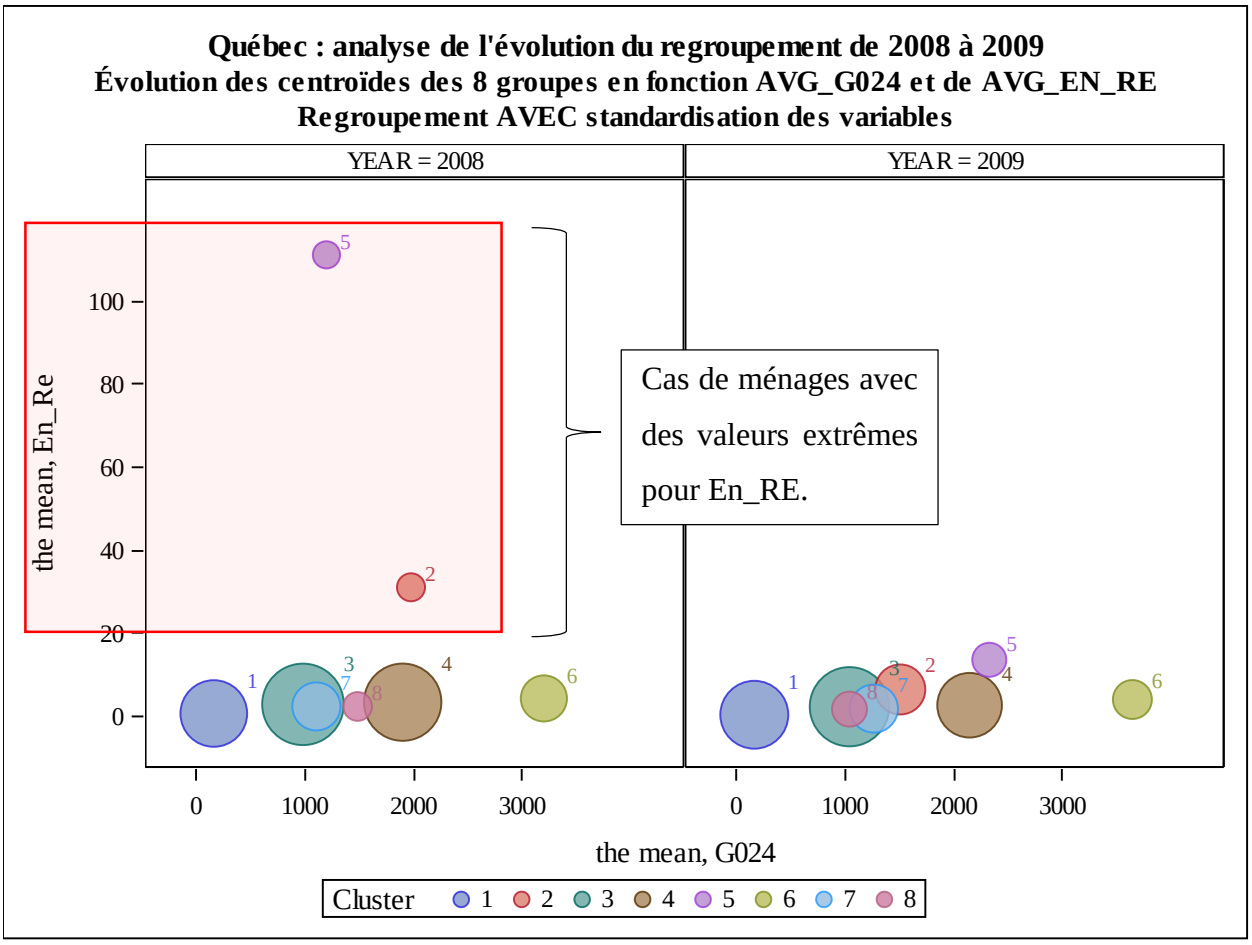


Figure 3-5 : Québec, Comparaison du regroupement entre 2009 et 2008. Problématique des valeurs extrêmes pour En-Re.

Afin de pallier cette difficulté, et de pouvoir suivre une structure de regroupement retrouvable dans le temps, les variables utilisées pour la construction du regroupement ne seront pas standardisées, ce qui aura pour effet d'éliminer l'effet de la variable En_Re étant donné son échelle (comprise entre 0 et 1, il s'agit d'un ratio), et de construire le regroupement seulement à partir des variables G024 et G022 qui ont la même échelle (en plus d'être toutes deux de même nature : une dépense en énergie). La distribution de la variable G024 (dépense en électricité) est présentée dans la partie

3.3.2, et la distribuiton de la variable G022 (dépense en combustible liquide, gaz) est présentée elle dans la partie 3.3.4. En n'utilisant plus la variable En-Re pour construire le regroupement, les « ménages extrêmes », an matière de En_Re, n'auront plus de poids dans la construction du regroupement.

Ainsi, toujours en utilisant les centroïdes du regroupement à 8 groupes de 2009, mais cette fois sans standardiser les variables, nous aboutissons à un regroupement à 8 groupes pour 2008 dont les caractéristiques sont pertinentes aussi bien pour 2008 que pour 2009[39] (et pour la période 1997 – 2009, comme cela sera présenté par la suite). Les caractéristiques de ce regroupement sont présentées ci-dessous, dans la partie 3.2.3.

[39] Le regroupement de 2009, sans standardisation des variables, a été obtenu en initialisant l'algorithme k-means avec les centroïdes de la partition finale de 2008 ajusté pour l'année 2009.

3.2.3 Sélection d'un regroupement qui satisfait les deux critères : interprétable et valide dans le temps

Visualisation[40] du regroupement en 8 groupes, pour Québec 2008, en utilisant les 5 caractéristiques descriptives (qui donnent la fiche d'identité des groupes) :

REGROUPEMENT en 8 groupes, Québec 2009						
N°	Nom	Taille	Revenu	Électricité	Combustibles	Électricté /Rev.
1						
2						
3						
4						
5						
6						
7						
8						

Ce regroupement répond aux deux critères mentionnés dans les parties 3.2.1 et 3.2.2, soit que le regroupement soit interprétable et aussi valide (retrouvable) sur la période 1997 – 2009 (possibilité de le suivre dans le temps). À partir de ce regroupement, nous pouvons distinguer trois catégories de groupes de ménages : le premier correspond au groupe 1 qui n'a pas de dépense en électricité ni en gaz pour la résidence principale. La seconde catégorie de groupes correspond aux groupes 2 à 6 pour lesquelles les ménages ont des dépenses

[40] Dans cette visualisation, lorsqu'une ligne pour une variable est entièrement coloriée, cela indique que ce groupe est le plus important (ce groupe a la dépense moyenne la plus importante) pour cette variable. La ligne de chacun des autres groupes est coloriée relativement à celle-ci. Donc il ne faut pas faire de comparaisons entre variables avec cette visualisation mais seulement entre les groupes par rapport aux variables.

en électricité (et des dépenses très faibles en gaz). Enfin, la troisième catégorie correspond aux groupes 7 et 8 qui ont des dépenses en gaz et en électricité. Ces trois catégories de groupes sont illustrées dans la Figure 3-6.

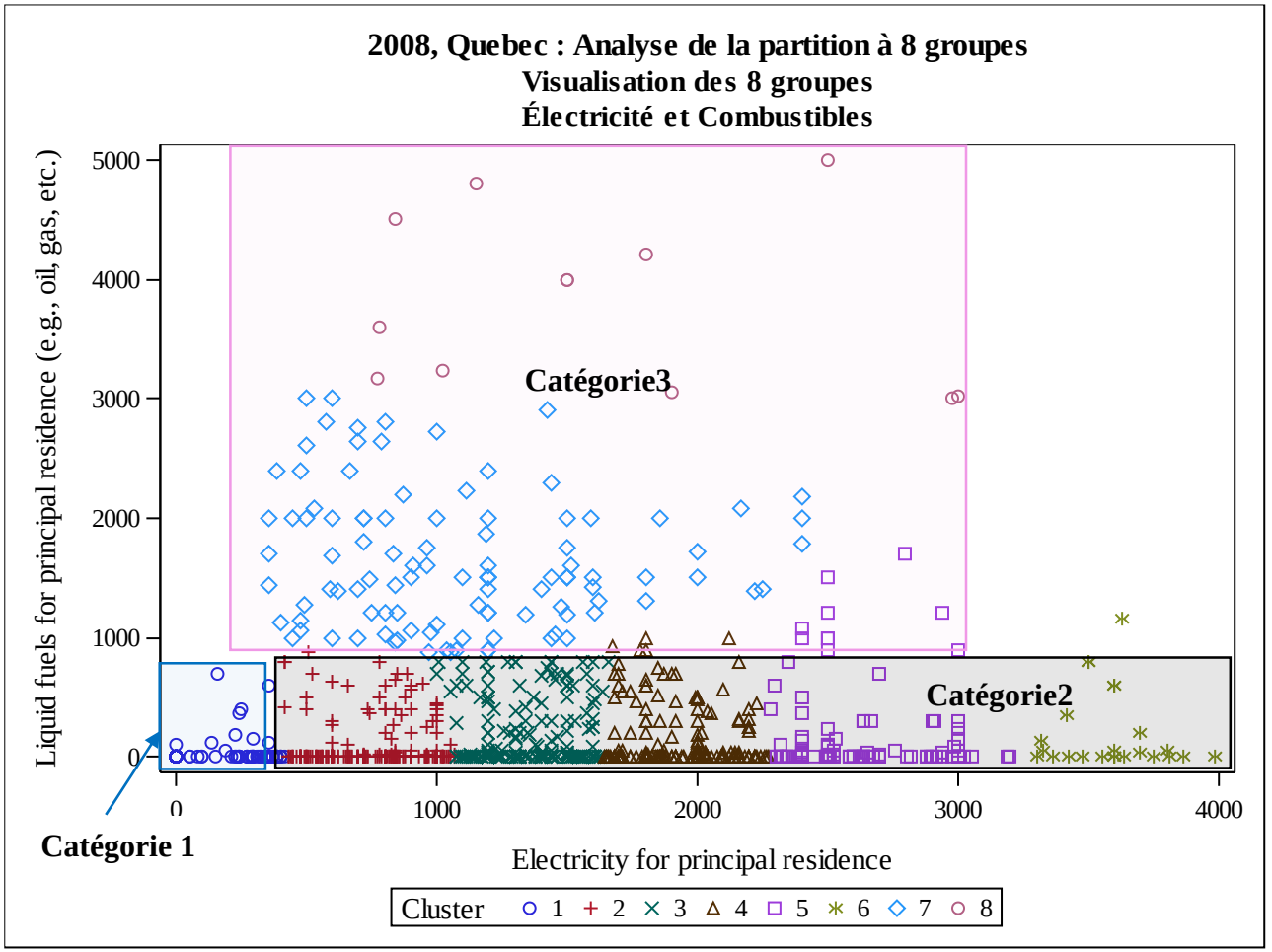


Figure 3-6 : Visualisation du regroupement sélectionné (8 groupes, Québec, 2008).

La Figure 3-7 ci-dessous permet d'illustrer qu'il est possible de suivre le regroupement sélectionné pour l'année 2008, d'année en année[41].

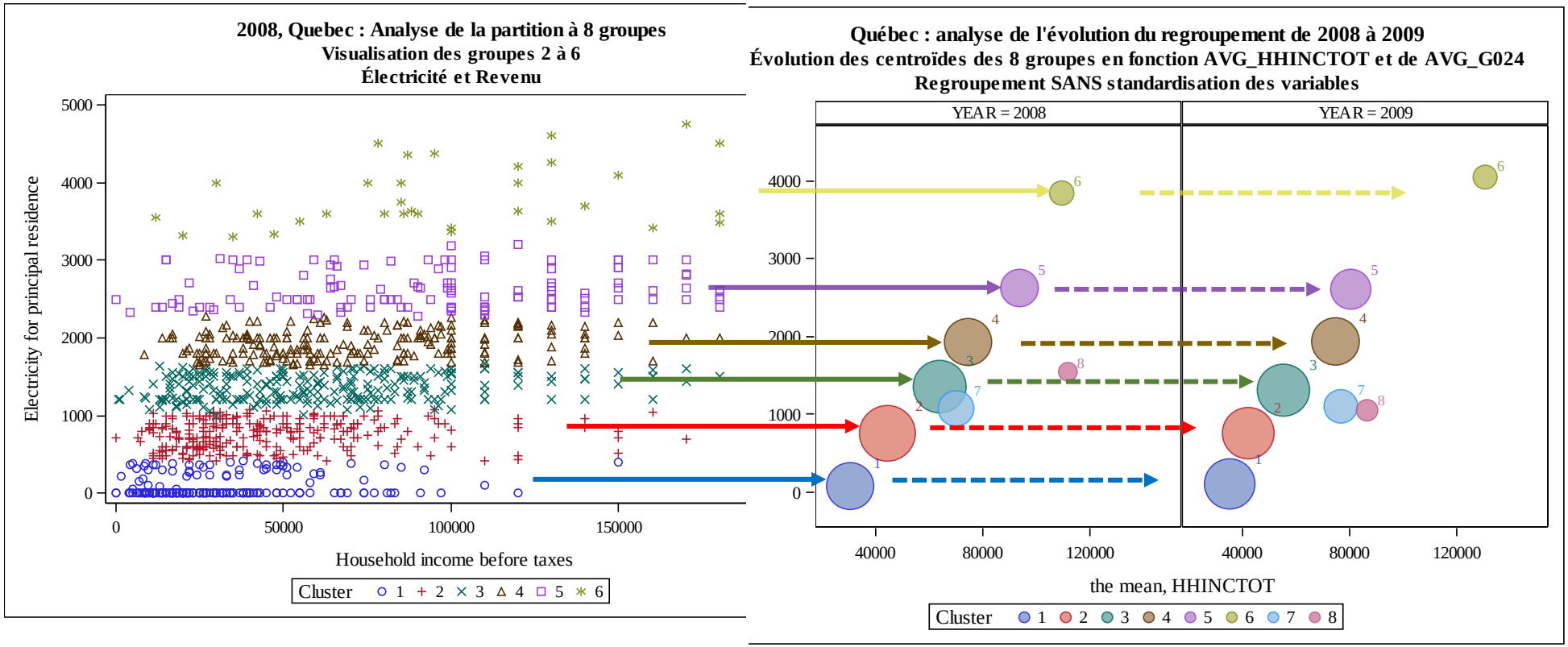


Figure 3-7 : Visualisation du regroupement sélectionné pour l'année 2008 et 2009.

[41] Le groupe 8 présente toujours une difficulté quant à sa taille, mais son centroïde définit toujours le même profil type de dépense.

3.3 Analyse du regroupement en 8 groupes, Québec, 2009

La construction du regroupement, présentée à la partie 3.1 et 3.2, a permis d'identifier un regroupement des ménages québecois en 8 groupes. Afin de présenter les caractéristiques de chacun d'eux, il est pertinent d'analyser si il existe une différence significative entre les moyennes de chacun des groupes relativement à certains postes de dépenses, ou bien d'analyser si la proportion de ménages avec certaines caractéristiques (zone d'habitation, type de logement) est plus importante dans un groupe donné que dans un autre. De plus, cette analyse doit permettre de quantifier cette différence à l'aide d'un intervalle de confiance (lorsque les modèles statistiques permettent cette inférence).

- Y-a-t'il une différence significative pour certaines variables quantitatives significatives entre les 8 groupes ?
- Par exemple, est-ce que le montant moyen dépensé par le groupe 1 pour le poste de dépenses X est significativement différent du montant moyen dépensé par le groupe 2 ?
- De même pour les variables catégorielles : peut-on estimer que la proportion de la catégorie « *c1* » de la variable Y est significativement différente entre les groupes ?

3.3.1 Méthode et validation statistique pour réaliser ces comparaisons

Afin de déterminer si des groupes de ménages ont une moyenne qui diffère pour une variable donnée, il convient de s'assurer qu'il y a suffisament de preuves dans l'échantillon pour pouvoir affirmer raisonnablement que la moyenne entre deux groupes est significativmenent différente à l'échelle de la population (inférence statistique valide[42]).

Pour une analyse de ce type dite « analyse de la variance » (ou ANOVA[43]), il y a trois hypothèses cruciales à vérifier pour garantir la validité des tests et donc la véracité de l'inférence:

1. La première, les observations (ici, les ménages) doivent être indépendantes
2. La deuxième, la variable d'intérêt doit suivre une loi normale (variable d'intérêt pourrait être un poste de dépense x).

[42] Les hypothèses statistiques testées sont les suivantes :

$$H_0 : toutes\ les\ moyennes\ sont\ égales$$
$$H_1 : au\ moins\ deux\ moyennes\ ne\ sont\ pas\ égales$$

[43] ANOVA, en anglais, signifie *Analysis Of Variance*

3. Enfin, la troisième, s'assurer que la variance entre les groupes pour la variable d'intérêt est constante (homoscedasticié).

Le mode de collecte de Statistique Canada pour l'enquête sur les dépenses des ménages (EDM) garantit l'indépendance des observations pour chaque année, de 1997 à 2009. De plus, il est d'usage, lorsque la taille des échantillons est assez grande, de supposer la condition de normalité satisfaite[44]. En revanche, pour plusieurs variables d'intérêt (ici poste de dépenses par exemple), la variance entre les groupes (comme G024) n'est pas constante. Par conséquent, l'utilisation de modèles qui permettent une variance propre à chaque groupe sera privilégiée surtout pour les analyses fines (c'est-à-dire, les comparaisons deux à deux entre groupes).

Cette analyse est effectuée en deux étapes :

1) Étape 1[45] : analyse globale et sélection des variables pertinentes
2) Étape 2[46] : analyse fine et comparaisons multiples

[44] « It is usually not critical that the distribution of the response be precisely normal: the Central Limit Theorem states that estimated group means are approximately normally distributed even if the observations have non-normal distributions. This happy fact provides approximate large-sample justification for the methods described in this chapter, as long as the other assumptions are valid. However, if the sample sizes are small and the distributions are not even close to normal, then the Central Limit Theorem may not apply » (Westfall, Tobias, & Russell, 2011, p. 102).
[45] MANOVA / PROC GLM / hovtest=levene welch
[46] PROC GLIMMIX

3.3.2 Étape 1 : analyse globale et sélection des variables pertinentes

À titre illustratif, nous présentons en détail dans cette partie la démarche suivie pour tous les résultats présentés dans la partie 4 en utilisant le cas « singulier » de l'analyse de la variable G024 : *Dépense en électicité des ménages*. Suite au regroupement en 8 groupes, différentes questions peuvent être soulevées, notamment « existe-il une différence significative entre la dépense moyenne en électricité du groupe 1 et la dépense moyenne en électicité du groupe 6 ? », par exemple.

Une analyse graphique (voire Figure 3-8) de la distribution[47] de la variable G024 entre les 8 groupes permet d'observer que les moyennes entre ces 8 groupes recouvrent des réalités diverses (c'est le but initial de la construction de ce regroupement, donc aucune surprise à ce stade !) et pour beaucoup de groupes, cette analyse graphique laisse supposer des différences qui pourraient (devraient) être significatives à l'échelle de la population des ménages. Comme le mentionnait le point méthodologique, les hypothèses de normalité pour G024[48], ainsi que l'indépendance des

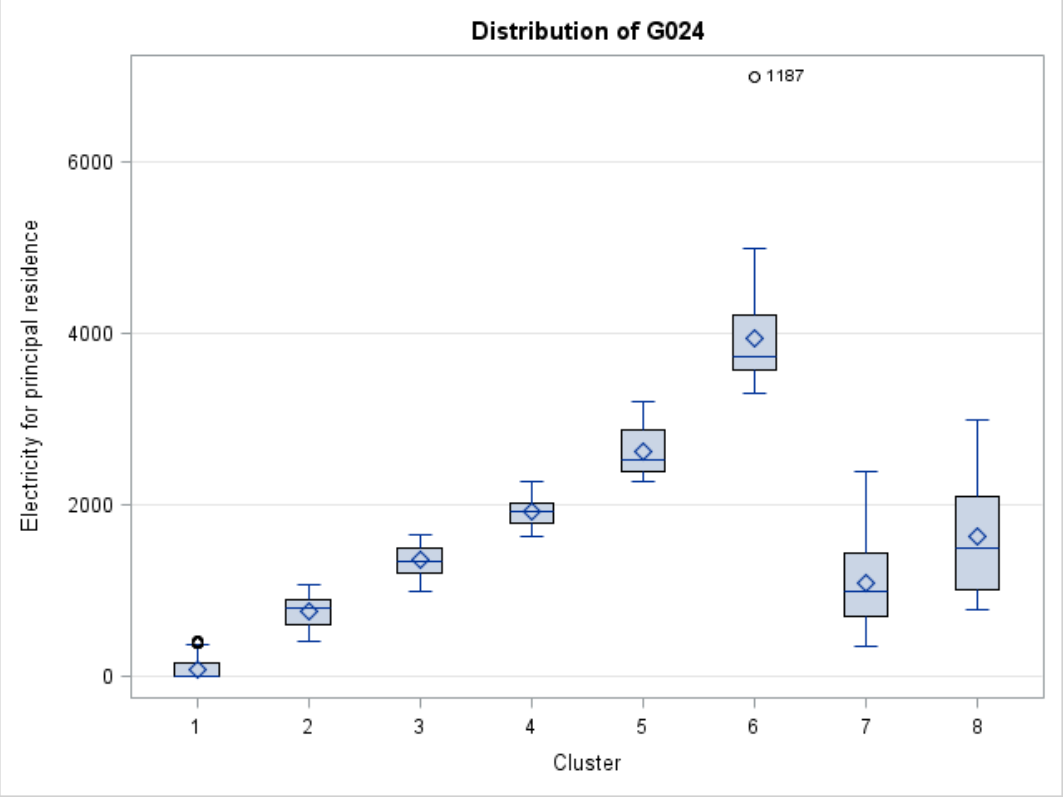


Figure 3-8 : Québec, 2009, distribution de la dépense en électricité au sein des 8 groupes

[47] Voir annexe 6.7.1, pour une description de la légende de la boîte à moustache (box-and-whiskers plot).
[48] Pour l'hypothèse de normalité, celle-ci n'est pas vérifiée (voire annexe 0) mais nous supposons la taille des groupes assez grande.

observations sont considérées vérifiées pour nos analyses EDM. En revanche, afin de poursuivre la comparaison des moyennes de dépense en électricité, il convient de prendre en compte le fait que la variance des dépenses des ménages en électricité entre les groupes n'est constante. Nous pouvons vérifier cela à l'aide d'une preuve numérique, tel que le présente le tableau ci-dessous (Table 4) ainsi que d'un test statistique.

Table 5 : Québec, 2008, G024 test formel sur l'égalité de la variance entre groupes

Levene's Test for Homogeneity of G024 Variance ANOVA of Squared Deviations from Group Means					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
CLUSTER	7	1.314E13	1.876E12	22.29	<.0001
Error	1221	1.028E14	8.417E10		

Table 4 : Québec, 2008, G024 moyenne et variance par groupe.

Level of CLUSTER	N	G024	
		Mean	Std Dev
1	201	81.96020	139.960703
2	279	756.25090	177.345215
3	254	1356.09843	162.975259
4	209	1927.04306	169.579834
5	126	2625.31746	241.677294
6	40	3943.82500	658.778625
7	105	1092.30476	502.689921
8	15	1632.13333	741.794048

La variance de la dépense en électricité pour le groupe 1 est de 140$ alors que celle-ci est de 658$ pour le groupe 6. Cette différence confirme que la variance n'est pas constante entre les groupes. De plus, un test formel confirme cette analyse (ci-dessus, Table 5). Le test statistique formel corrobore l'analyse graphique en réussissant à rejeter l'hypothèse d'homoscedasticité : par conséquent nous sommes dans un cas où la variance des dépenses en électricité des ménages n'a pas une variance constante entre les groupes (hestéroscedasticité). C'est une situation qui sera souvent rencontrée et prise en compte pour l'anlayse des différences entre le 8 groupes de ménages en utilisant des modèles appropriés[49] (tel que le test de Welch avec Proc GLM[50]). En assumant ces différentes variances entre les groupes et en utilisant l'approximation de Satterthwaite[51], nous

[49] La problématique est que le test de Fisher classique suppose une variance constante entre les groupes. Comme nous pouvons le constater à l'annexe 10.7.7, la variance moyenne (Root MSE) est estimée à 261.62 dollars, or celle-ci est trop petite pour juger de la comparaison entre les groupes 6, 7, 8, et trop grande pour juger de la comparaison entre les groupes 1, 2, 3 et 4. Une facon de dépasser ce problème est d'utiliser un modèle qui tient compte de l'hétéroscedasticité afin de permettre différentes variances entre groupes. Nous sommes dans un cas d'ANOVA en présence d'hestéroscedasticité.

[50] Le test de Welch dans Proc GLM a l'inconvénient de ne pas tenir compte du poids (variable Weight) du ménage, et donc de sa représentativité. En revanche, la procédure de SAS Proc Glimmix permet de tenir compte de la variable Weight tout en réalisant une approximation des degrés de liberté selon la formule de Satterthwaite.

[51] Cette approximation permet d'utiliser une distribution t pour chaque comparaison avec différents degrés de libertés.

pouvons légitimement comparer les moyennes des dépenses en électricité (et pour d'autres variables) entre les différents groupes de ménages.

À ce stade de l'analyse, c'est-à-dire à la fin de l'étape 1, il est possible d'affirmer qu'au moins deux groupes ont une dépense moyenne en électicité qui n'est pas égal (d'après le test Welch's ANOVA, voir annexe 6.7.4.). Cette première étape est menée au cours d'une analyse multivariée d'ANOVA (MANOVA) avec toutes les variables susceptibles d'intérêt pour juger de la différence entres les groupes[52] et permettant de caractériser ces différents groupes. Les variables sélectionnées pour réaliser la MANOVA sont présentées en annexe 6.7.6.

3.3.3 Étape 2 : analyse plus fine, 2 à 2, par paire

L'analyse MANOVA permet d'identifier un ensemble de variables d'intérêt pour lesquelles au moins deux groupes du regroupement ont des moyennes significativement différentes. Afin d'analyser plus finement ces différences entre groupes, il convient de poursuivre avec une analyse plus fine, c'est-à-dire en analysant les différences par paire de groupes (2 à 2) toujours en tenant compte du fait de la non égalité de la variance entre groupes[53] lorsque celle-ci a lieu.

Pour réaliser l'analyse fine, c'est-à-dire la comparaison des moyennes pour chaque paire de groupes, il est important d'ajuster les p-values pour prendre en compte le fait des comparaisons multiples qui peuvent fausser l'estimation des p-values. En effet, certaines différences entre groupes pourraient être jugées significatives alors qu'elles ne le sont pas. Deux options sont possibles : soit l'option de Tukey-Kramer[54] (voir annexe 6.7.7, pour le cas de la varaible G024 et le regroupement de 2008) ou bien celle de Bonferroni[55] (voir annexe 6.7.8, pour le cas de la varaible G024 et le regroupement de 2008) qui toutes deux ajustent les p-values selon leur procédure propre. C'est l'option de Tukey-Kramer qui sera retenue pour l'étape 2 (moins restrictive).

Dans le cas de l'analyse des différences entre chaque paire de groupes quant à la moyenne des dépenses en électricité, celle-ci permet de conclure que chaque groupe de ménages du

[52] Voire liste complète en annexe 6.7.6.

[54] Après plusieurs tests, il semble que l'utilisation de la variable Weight fausse l'estimation des p-value, aussi bien pour la procédure de Tukey-Kramer que pour la procédure de Bonferroni.

[55] Id.

regroupement à 8 groupes de 2008 a une moyenne significativement différente des autres à l'exception des paires de groupes suivantes : (3 et 8), (4 et 8) et (7 et 8).

3.3.4 Le cas des variables qui ont servi à construire le regroupement G024, G022 ainsi que HHINCTOT (qui n'a pas servi à construire le regroupement).

Aussi bien pour les dépenses moyennes en électricité que pour les dépenses moyennes en combustibles liquides, les analyses de l'étape 1 indiquent que ces moyennes entres les groupes ne sont pas égales (au moins pour un couple de moyenne[56]). Il en est de même pour le revenu moyen des ménages, où les moyennes des revenus entre les groupes ne sont pas égales. Les figures ci-dessous rendent compte de la distribution des variables HHINCTOT et G022 illustrent ce point.

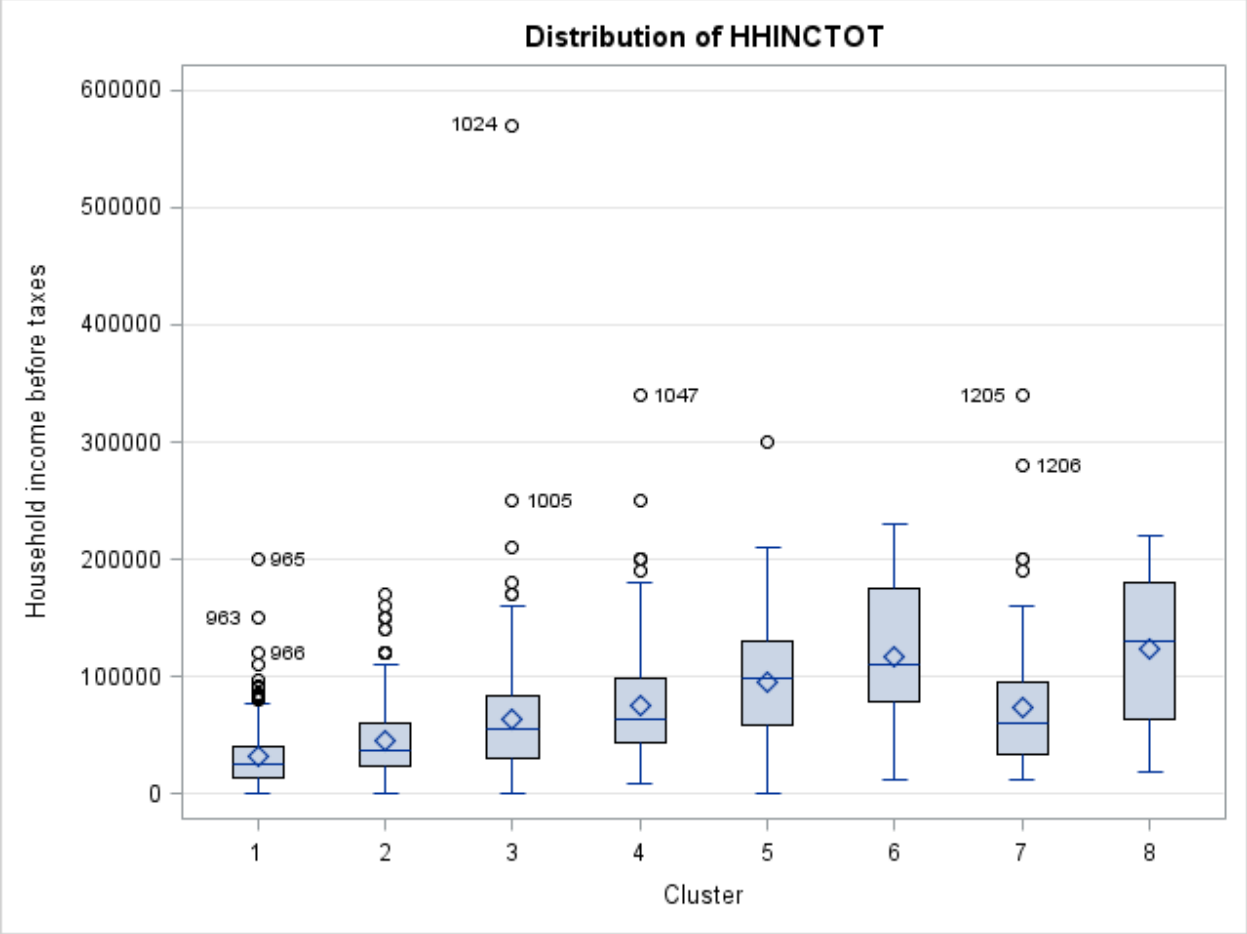


Figure 3-9 : Distribution du revenu total au sein de chaque groupe, Québec 2008.

[56] L'hypothèse nulle est que toutes les moyennes sont égales alors que la contre-hypothèse est qu'au moins un couple de moyenne diffère entre deux groupes.

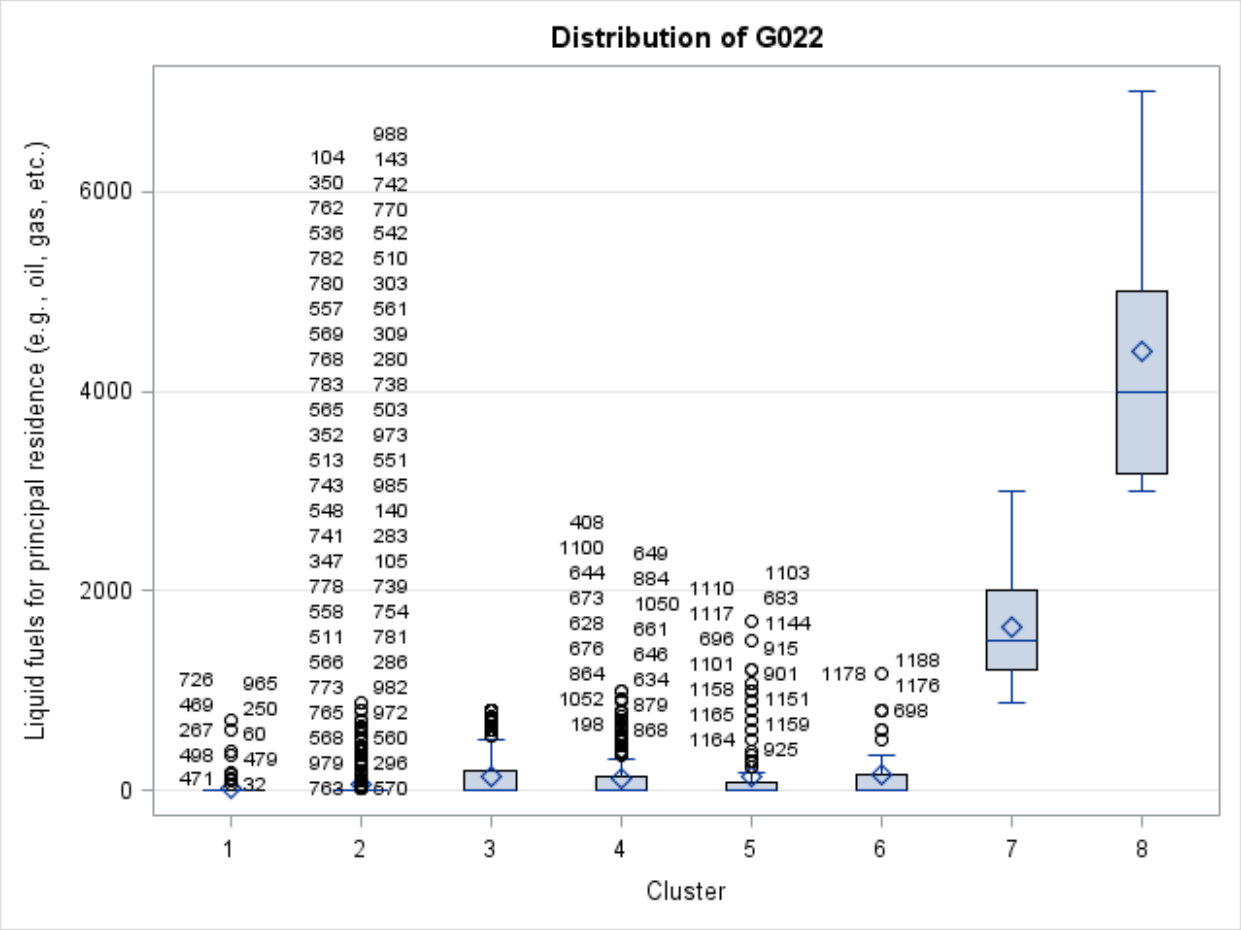


Figure 3-10 : Distribution de la dépense en combustibles au sein de chaque groupe, Québec 2008.

En revanche, lors d'une analyse plus fine (comparison des moyennes, groupe 2 à 2) et en ajustant les p-values pour comparaisons multiples, les choses sont un peu plus contrastées. Aucune différence de moyenne n'est jugée significative quant au revenu total des ménages (HHINCTOT). En revanche, pour la dépense moyenne en combustibles (gaz), les moyennes sont toutes significativement différentes entre toutes les paires de groupes à l'exception des paires constituées exclusivement des groupes 2, 3, 4, 5 et 6.

4 Résultats :

4.1 Québec, portrait des ménages en 2008

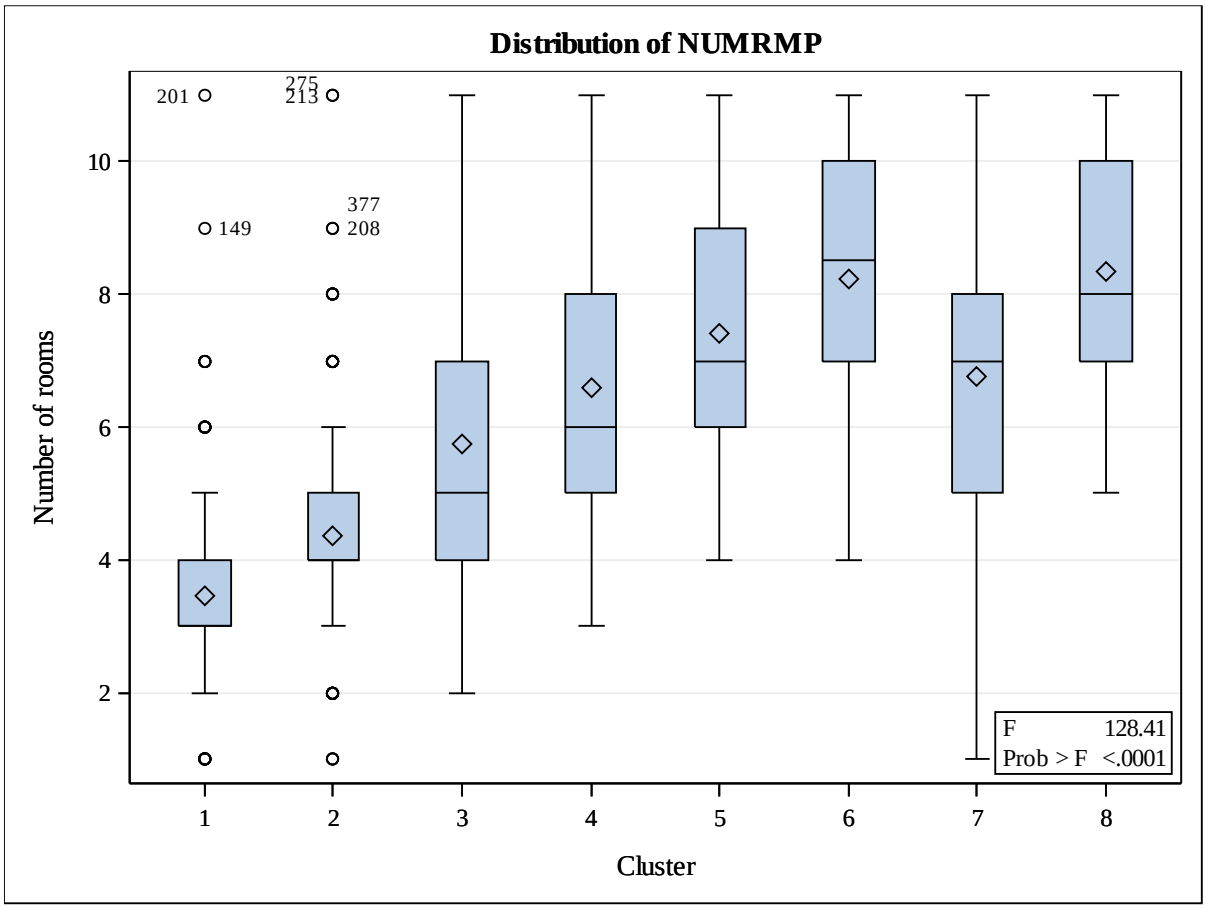
Cette partie du rapport présente le portrait des ménages québecois, en fonction de leurs dépenses en énergie, en 2008. Les caractéristiques et faits saillants, distinctifs, du portrait sont mis en avant et illustrés. De plus, ces caractéristiques sont comparés avec le regroupement de 1998, soit 10 années plus tôt afin de les mettre en perspective.

4.1.1 Le cas des variables numériques

La partie méthodologique précédente, partie 3.3, illustrait le cas des variables numériques qui ont servi à construire le regroupement, soit G024 (dépenses en électricité) et G022 (dépenses en combustibles, gaz). La conclusion était la suivante : plusieurs groupes de ménages ont des moyennes significativment différentes à l'échelle de la population des ménages québecois quant à leurs dépenses en électricité et en gaz pour la résidence principale. En revanche, bien qu'il soit possible d'observer des différences (mêmes importantes) quant au revenu total moyen entre les groupes, celles-ci ne sont pas jugées significatives à l'échelle de la population des ménages (cela peut-être lié aussi à la présence de tous les quintiles de revenu, dans des proportions variables, dans chaque groupe comme le présente la partie 4.3 du rapport).

L'analyse a été menée avec toutes les variables numériques présentées en annexe 6.7.6 (par exemple, G004 pour le montant du loyer, ou encore L101 pour les dépenses totales de santé et bien d'autres encore). Si il est possible d'observer pour beaucoup de variables numériques des différences quant à la moyenne entre les groupes au sein de l'échantillon (les données de l'enquête), ces différences observées ne sont jamais significatives à l'échelle de la population des ménages à part pour de rares exceptions, par exemple, le nombre de pièces des logements des ménages. Cette dernière appuit une observation déjà connue qui est que la consommation d'énergie, et donc dans notre cas la dépense d'énergie, pour la résidence principale, est fortement corrélée avec la superficie du logement du ménage en question. Le graphique ci-dessous illustre cette différence :

- Le nombre de pièces moyen du logement du ménage. Les moyennes de chaque paire de groupes sont significatives à l'exception des paires suivantes : 4 et 7 ; 5 et 8 ; 6 et 8.



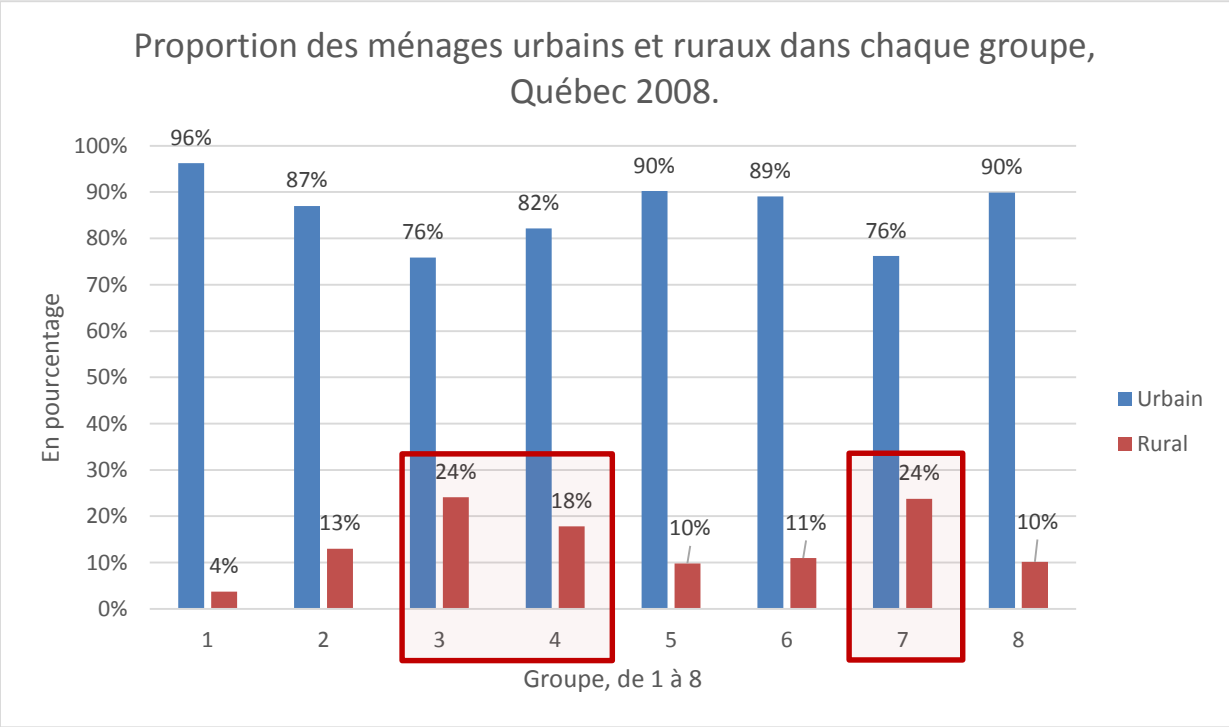
Cl.	1998 Moy.	2008 Moy.
1	3.6	3.4
2	4.5	4.4
3	5.6	5.8
4	6.5	6.5
5	7.4	7.3
6	8.6	8.0
7	6.5	6.5
8	7.2	8.2

Les groupes 1, 2 et 3 concentrent les ménages dont le logement a le plus petit nombre de pièces, soit moins de 4 pièces pour le groupe 1, moins de 5 pièces pour le groupe 2 et moins de 6 pièces pour le groupe 3. Les groupes de 4 à 6 sont composés de ménages dont le nombre de pièces est en moyenne supérieure à 6. Ces moyennes sont constantes dans le temps et très semblables à celles de 1998.

4.1.2 Le cas des variables catégorielles

Cette analyse permet de mettre en avant plusieurs différences significatives entre les groupes quant aux caractéristiques des ménages qui sont présentées ci-dessous par une série de graphiques pour les variables suivantes : le lieu d'habitation (Urbain-Rural), le type de logement (Maison, Appartement…), le type du ménage (propriétaire, locataire…), le type d'appareil de chauffage (Fournaise, chauffage à électricité…), l'âge du système de chauffage… Une description complète de ces variables est fournie en annexe 6.8.3. Pour tous les graphiques ci-dessous, les proportions sont significativement différentes entre les groupes[57] (sauf exceptions qui sont mentionnées).

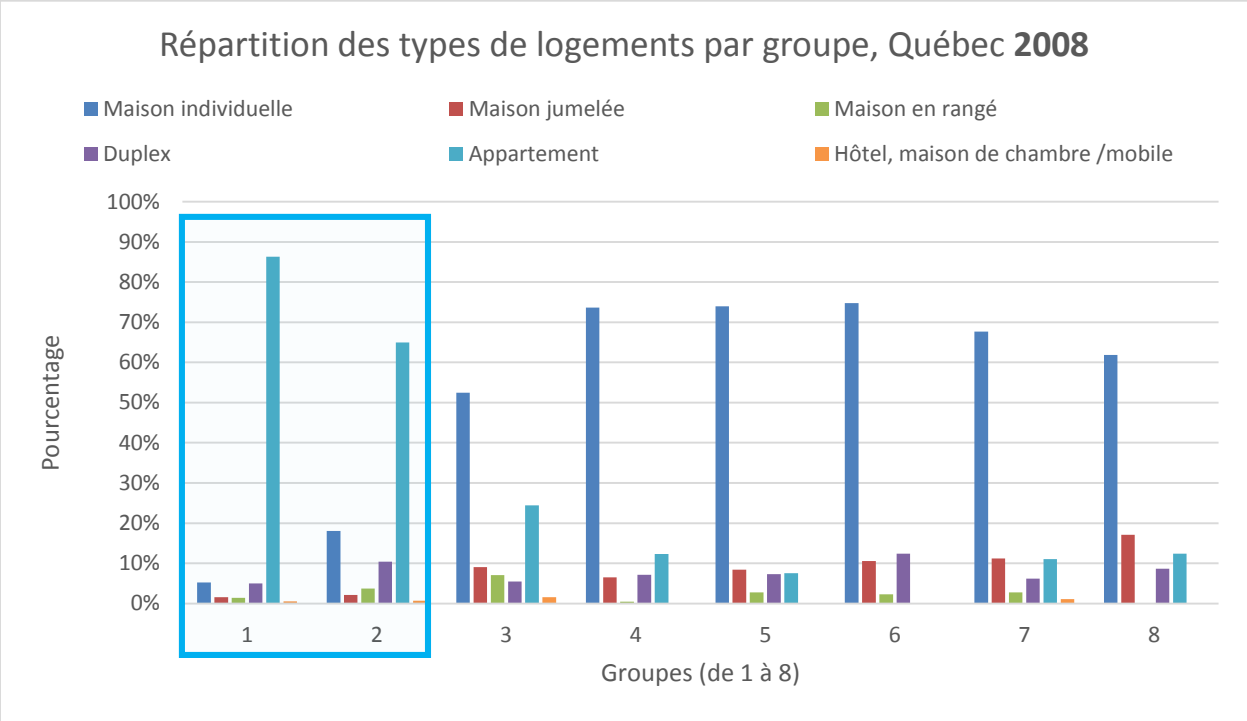
- La proportion de ménages urbains/ruraux qui composent le groupe (différence non significative uniquement entre le groupe 5 et 8).



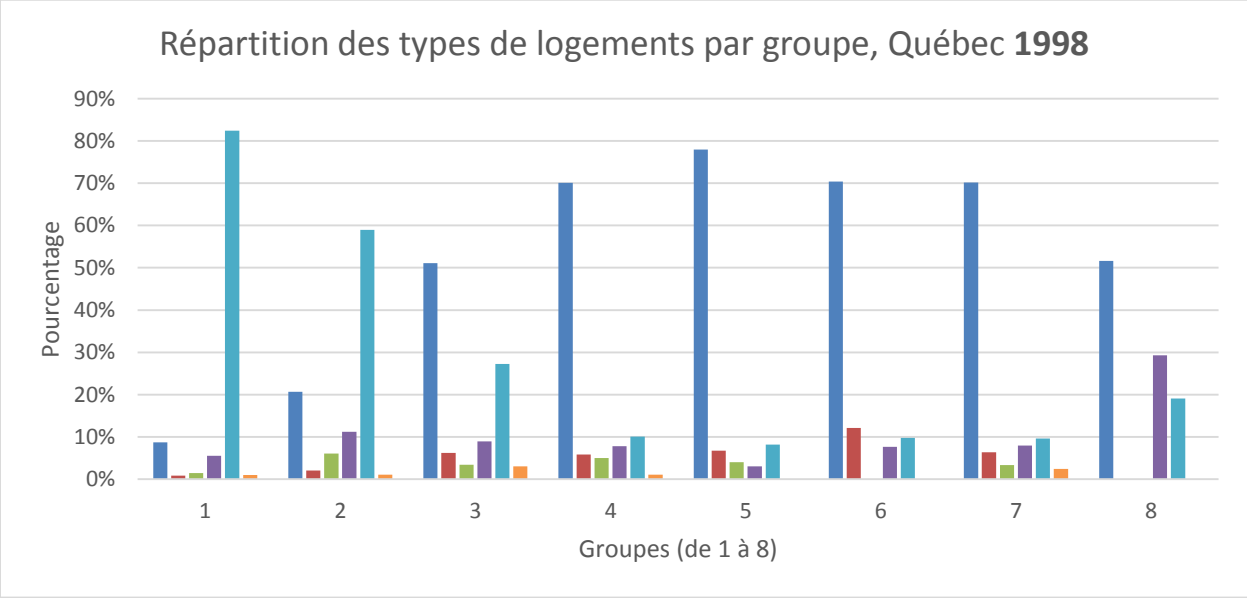
Les groupes 3, 4 et 7 sont ceux qui ont, en proportion, le plus de ménages vivant en zone rurale. En revanche, le groupe 1 est celui pour lequel cette proportion est la plus faible. Quelque soit le groupe, la proportion de ménage vivant en zone urbaine est majoritaire (la moyenne est de 86%). **Les données sont semblables en 1998 (en pourcentage).**

[57] Note importante : lors des analyses et comparaisons, les conclusions ou analyses concernant les groupes 6 et 8 sont à relativiser étant donné que ces groupes connaissent d'assez grande variation dans le temps et que leur représentativité peut être discutée (ils regroupent des ménages limites).

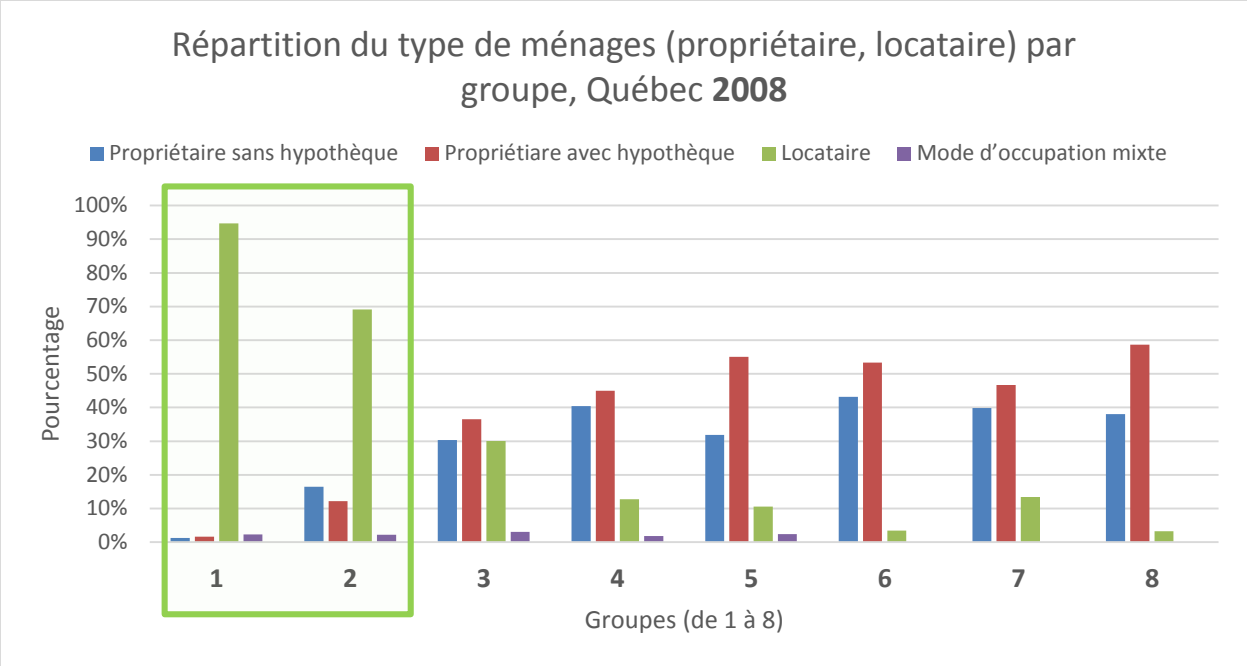
- Le type de logement par groupe diffère de manière significative entre tous les groupes.



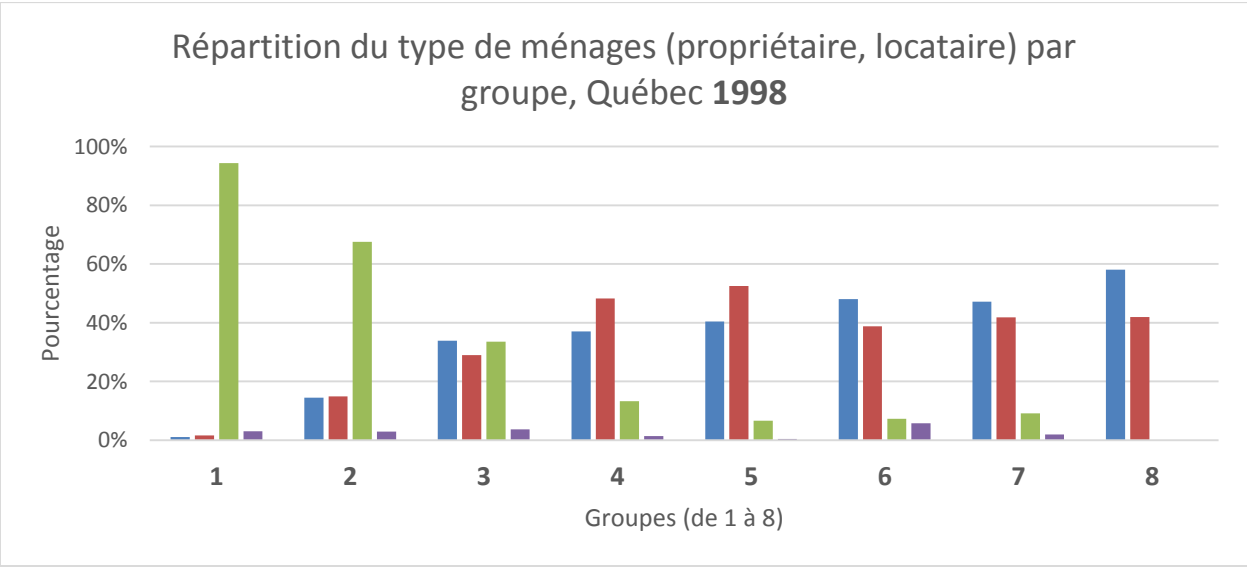
Le groupe 1 et 2 sont ceux pour lesquels la proportion de ménages vivant en appartement représentent plus de 50% : près de 90% pour le groupe 1 et environ 65% pour le groupe 2. Pour tous les autres groupes (de 2 à 8), plus de 50% des ménages de ces groupes vivent dans une maison individuelle. Les données sont globalement semblables à celles datant de 1998.



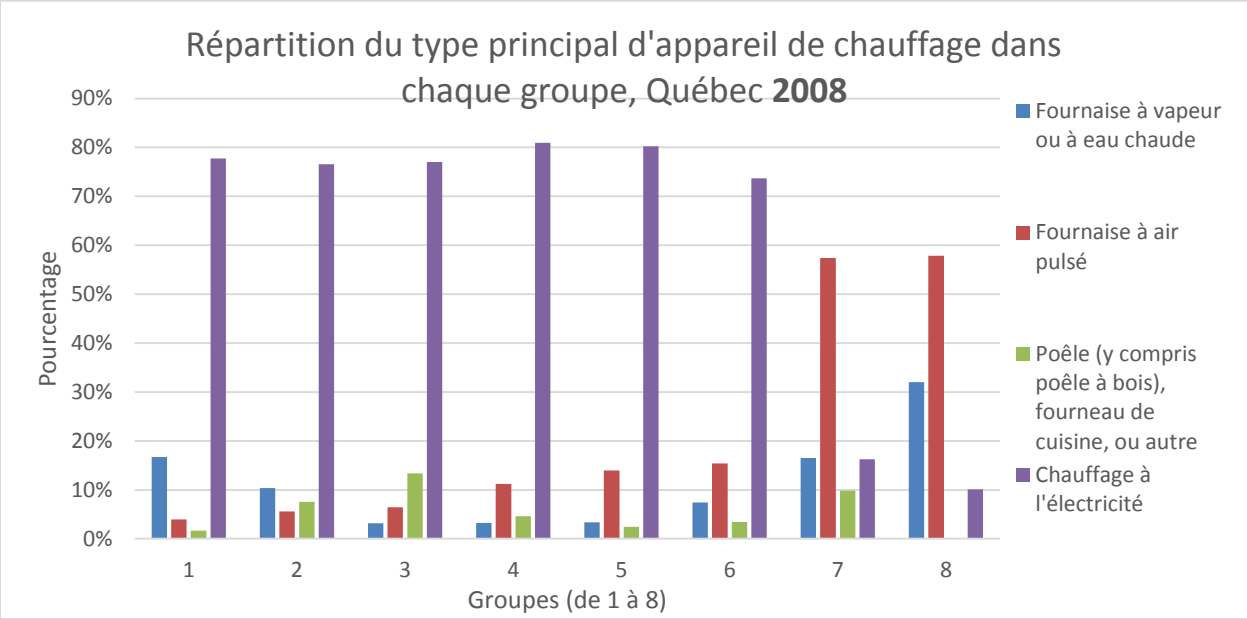
- Le type du ménage, soit propriétaire ou locataire, diffère de manière significative entre les groupes.



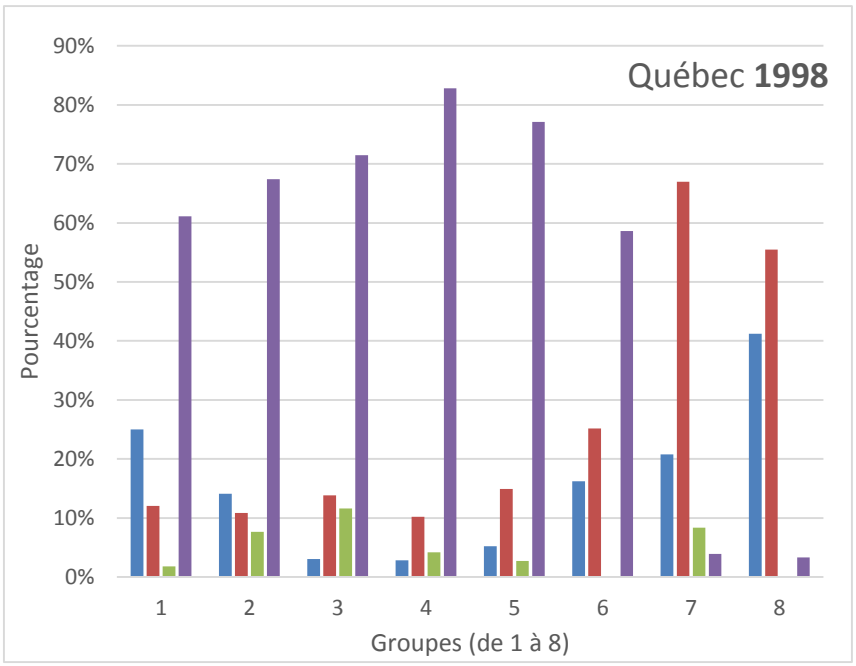
Les locataires sont majoritairement concentrés dans les groupes 1 et 2 où ils représentent respectivement 95% et 65% des ménages de ces groupes. Pour tous les autres groupes, la proportion des locataires est inférieure à 30% : plus de 60% des ménages de chacun de ces groupes (groupes 3 à 8) sont des propriétaires. Les données étaient semblables en 1998 à l'exception des groupes 6 à 8 pour lesquelles les ménages propriétaires sans hypothèque étaient majoritaires.



- La répartition des types de chauffage au sein des groupes.

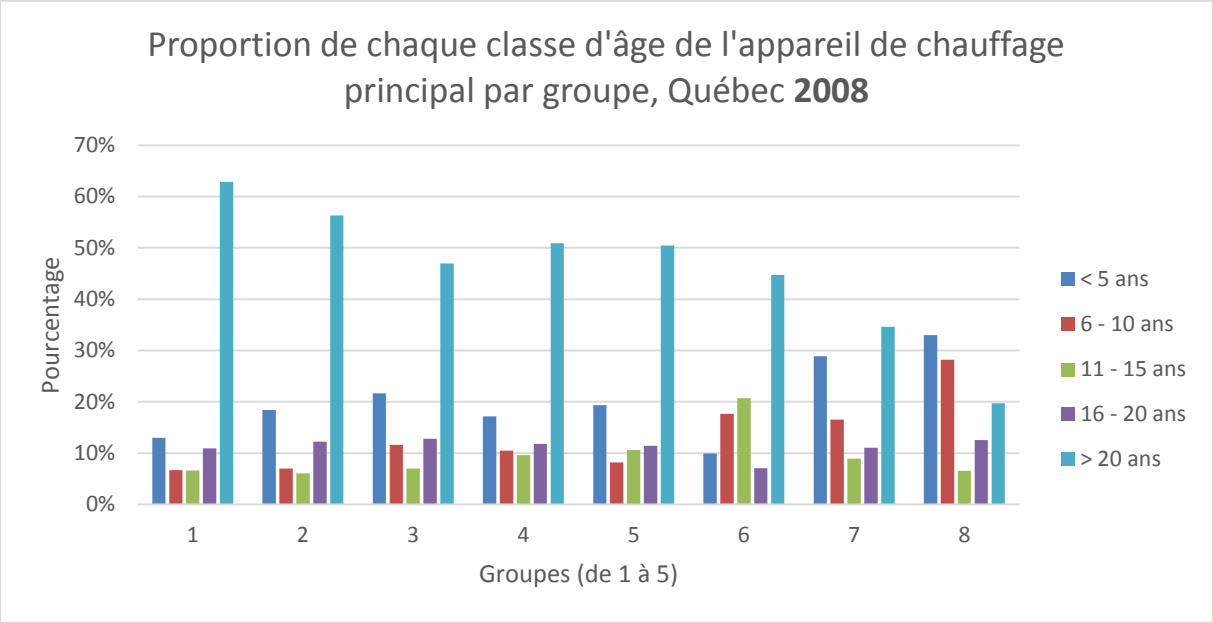


Plus de 70% des ménages des groupes 1 à 6 ont un type d'appareil de chauffage principal à électricité. 60 % des ménages de chacun des groupes 7 et 8 ont comme type d'appareil de chauffage principal une fournaise à air pulsé. Il est intéressant de noter que c'est l'une des caractéristiques (le type de chauffage principal) des ménages, et donc des 8 groupes, qui a le plus évolué entre 1998 et 2008, comme l'illustre le graphique ci-dessous.

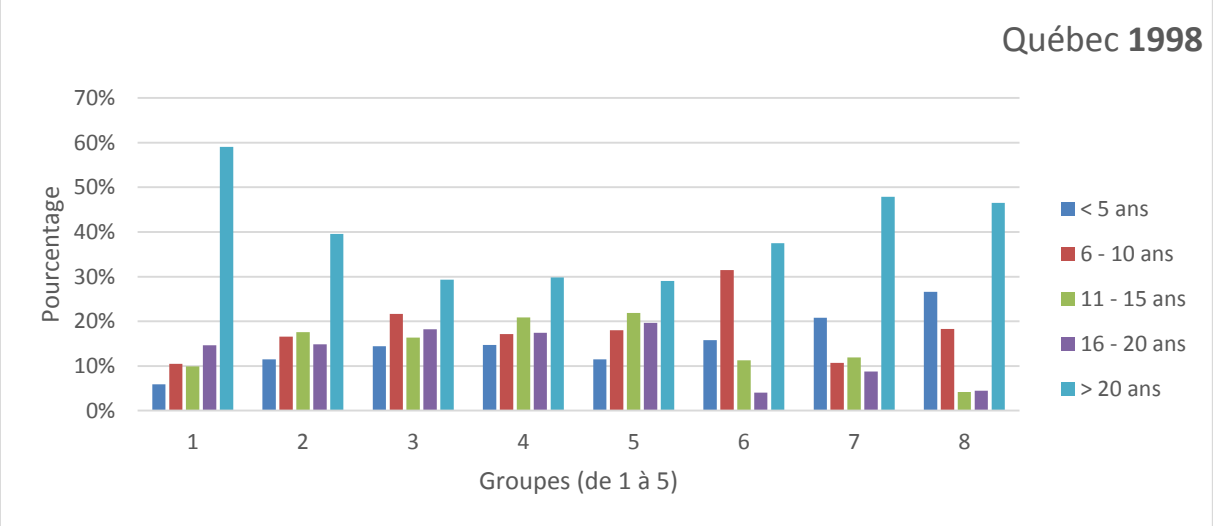


De 1998 à 2008, pour tous les groupes (à l'exception du groupe 4), la part occupée par le chauffage électrique dans les ménages est en augmentation.

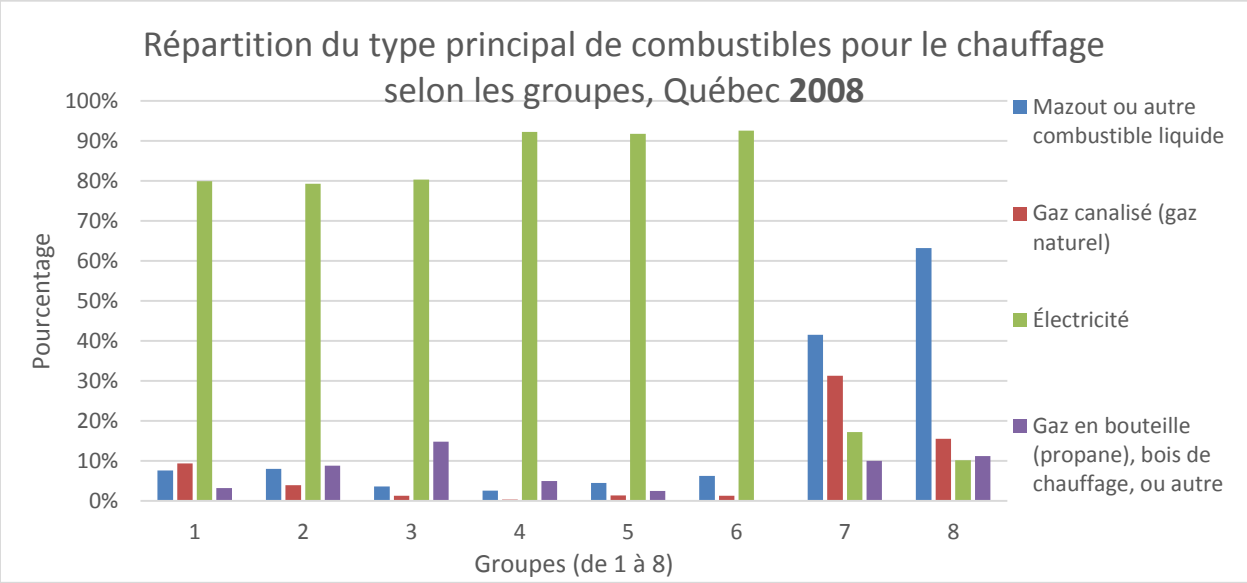
- La répartition des classes d'âge d'équipement de chauffage au sein des groupes.



Les groupes 1 et 2 ont une part très importante de l'équipement de chauffage dont l'âge moyen est supérieur à 20 ans. Ce fait est vrai aussi bien en 1998 qu'en 2008. Le plus grand changement entre 1998 et 2008 s'observe chez les groupes 7 et 8 pour lesquels le parc d'appareil de chauffage rajeunit, fait concomittant avec une augmentation importante du nombre d'appareil de chauffage électrique chez ces deux groupes. Notons un vieillissement du parc d'appareil de chauffage pour pour les groupes 3, 4, 5 et 6.

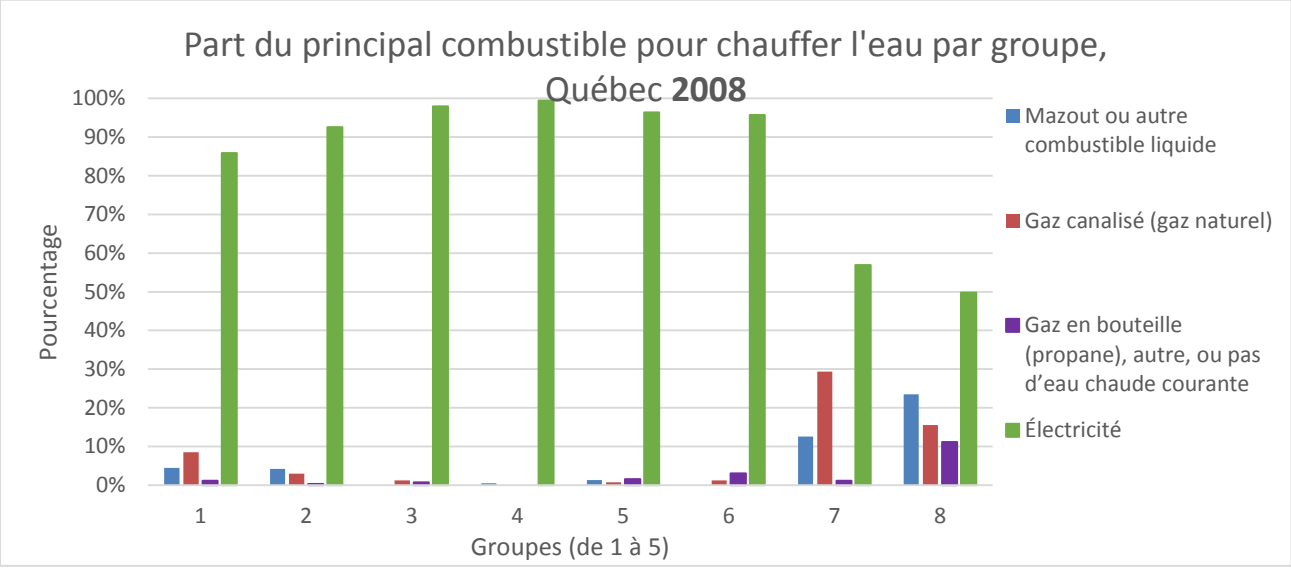


- La part de chaque type de combustibles pour le chauffage selon les groupes.

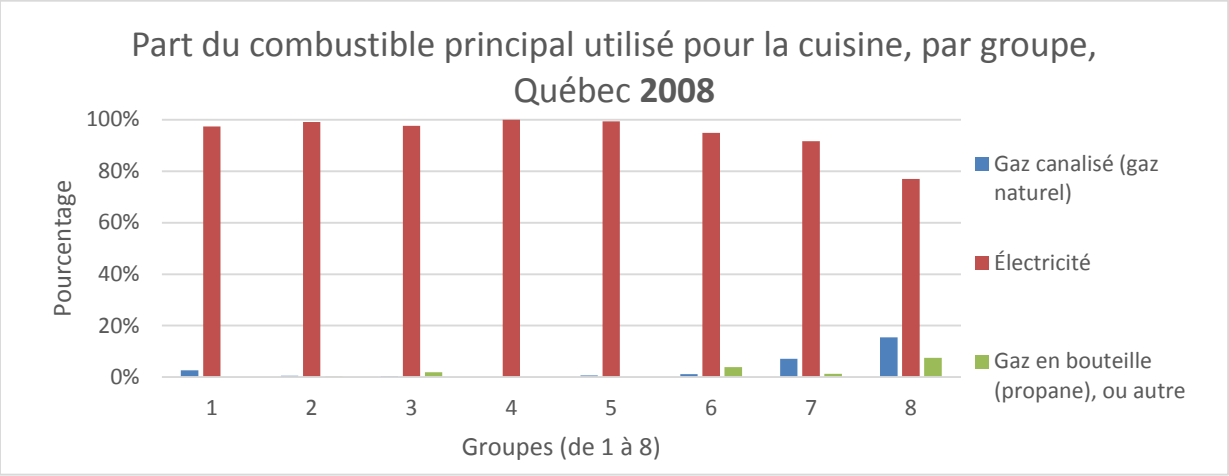


Le type d'appareil de chauffage principal des groupes 1 à 6 étant l'appareil à électricité, constater que le combustible principal de ces mêmes groupes est l'électricité est sans surprise. Aussi, pour les groupes 7 et 8, c'est le mazout (ou autre combustibles) et le gaz naturel qui sont les principaux combustibles, là aussi sans surprise puisque l'appareil principal de chauffage de ces groupes est la fournaise à air pulsé. Ces caractérisques se retrouvent aussi bien pour le type de combustibles utilisés pour le chauffage de l'eau (voir graphique ci-dessous).

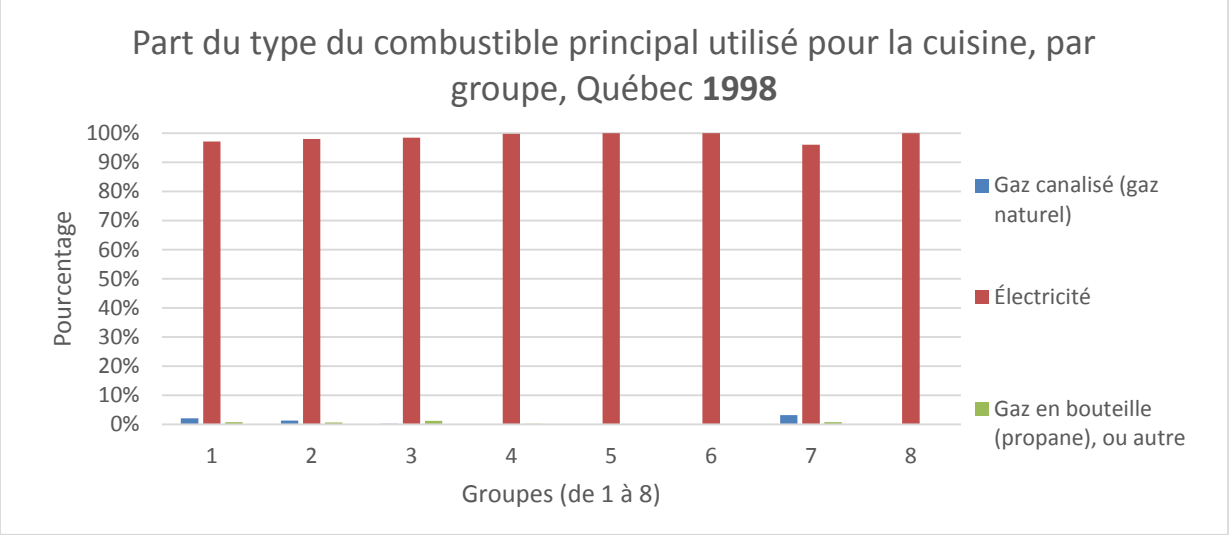
- Part du type d'énergie principal pour chauffer l'eau au sein des groupes.



- La répartition du principal combustible utilisé pour la cuisine dans chaque groupe



Ce qu'il peut être intéressant de noter hormis la prédominance de l'utilisation de l'électricité comme combustible principal pour la cuisine dans chaque groupe, c'est le retour (de 1998 à 2008) du gaz naturel comme combustible pour la cuisine dans les groupes 7 et 8, avec 7% et respectivement 16% des ménages de ces groupes qui utilisent du gas naturel pour la cuisine.



4.1.3 Analyse des caractéristiques des groupes et pertinence du regroupement

Ces différentes analyses (graphiques) permettent de confirmer les profils de dépenses en énergie proposés dans la partie 3.2.3 du regroupement en 8 groupes.

4.1.3.1 Les trois catégories de profil de dépenses en énergie

Les catégories 1 (constituée du groupe 1) et 2 (constituée du groupe 2 à 6) ont un chauffage principal à l'électricité (à plus de 70%) et utilisent donc comme combustible l'électricité (à plus de 80%). C'est pourquoi ce sont les groupes qui ont des dépenses exclusivement en électricité et non en gaz. La catégorie 3 (constituée des groupes 7 et 8) a un profil plus mixte dont l'appareil de chauffage principal est une fournaise à air pulsé. Les groupes de la catégorie 3 consomment principalement du mazout, du gaz naturel et aussi, mais dans une moindre mesure, de l'électricité. Ces analyses permettent d'expliquer les profils de dépenses en énergie mis en avant par le regroupement en 8 groupes.

4.1.3.2 La spécificité de la catégorie 1 (le groupe 1)

Le cas du groupe 1, dont près de 72% des ménages qui le composent, n'ont pas de dépense en électricité, et le reste (soit environ 28%) ont une dépense annuelle en électricité inférieure à 420$ peut s'expliquer par le fait que plus de 90% de ces ménages sont des locataires dont l'électricité doit être inclus dans le loyer. Pour corroborer cette analyse, nous pouvons comparer la moyenne des dépenses pour le loyer entre le groupe 1 et 2, pour les ménages locataires de ces groupes sachant qu'en moyenne, les logements du groupe 2 ont un nombre de pièces plus important que les logements du groupe 1. Nous pourrions nous attendre, de ce point de vue, à un loyer moyen plus élevé pour les ménages du groupe 2.

Level of CLUSTER	N	G004	
		Mean	Std Dev
1	**190**	8096.97368	5928.80548
2	**192**	6296.13542	2486.98005

Table 6 : Groupe 1 et 2, loyer moyen pour les locataires, Québec 2008

En 2008, la moyenne du loyer (annuel) pour les locataires du groupe 1 est de 8097$ alors que celle-ci est de 6296$ pour les locataires du groupe 2. Toutefois, en prenant en compte le fait de comparaisons multiples, cette différence n'est pas jugée significative. Elle l'est toutefois en réalisant une ANOVA (sans comparison multiple) où l'on compare le loyer moyen des locataires du groupe 1 et 2. Le test de Welch[58] estime la différence significative. À ce stade de l'analyse, ces

[58] Utilisation du test de Welch prend en compte la non égalité des variances entre ces deux groupes, comme l'illustre la Table 6

données permettent de dire que probablement les locataires du groupe 1 ont des dépenses d'électricité liées à leur logement (chauffage principalement) inclues dans le loyer contrairement au groupe 2. Ils ont des dépenses en électricité indirectement à travers leur loyer. Le groupe 1 est composé en majorité de locataires ayant les dépenses d'électricité inclues dans le loyer, ce qui pourrait expliquer pourquoi plus de 70% des ménages qui composent le groupe 1 n'ont aucune dépense en électricité. Ce point est aussi souligné dans la revue de littérature, voir la partie 1.3.1.3 et l'article de la « Chicago Fed Letter », de juin 2006 où les auteurs expliquent qu'il y a trois types de dépenses indirectes en énergie qui peuvent influencer les résultats d'une analyse qui portent sur les dépenses des méanges dont les dépenses en gaz et en électricité qui sont parfois inclues dans le loyer des appartements ou maisons louées (Cashin & McGranahan, 2006).

4.1.3.3 Les profils de dépenses en énergie (3 catégories, 8 groupes) et les quintiles de revenu

La problématique évoquée en début de rapport dans l'introduction, partie 1.1 du rapport, faisait la critique d'une représentation des ménages en groupe de quintile. Deux critiques étaient adressées à cette représentaiton, soit le fait de ne pas rendre compte de ménages qui n'avaient pas de dépense en électricité par exemple, car masqués par la moyenne du quintile auquel ils appartiennent

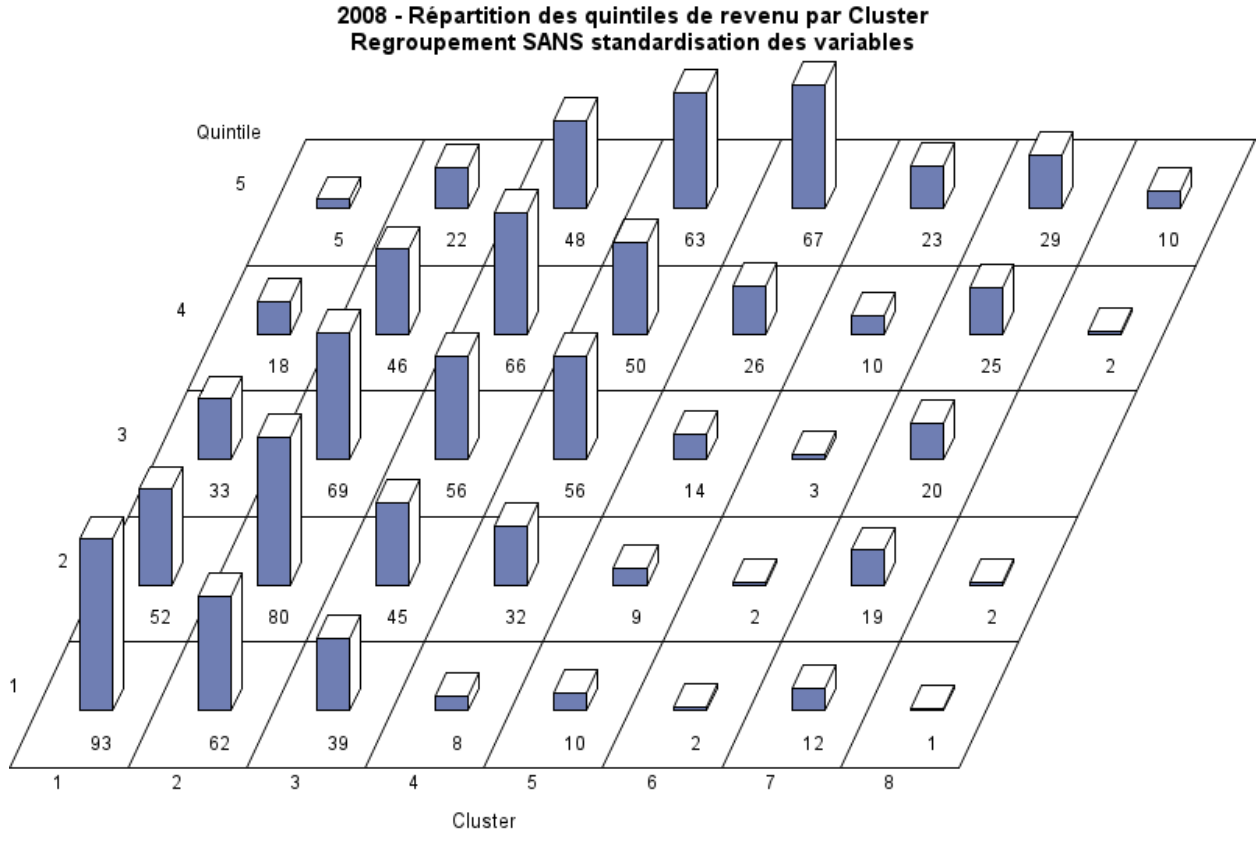


Figure 4-1 : Répartition des quintiles de revenu au sein des 8 groupes, Québec 2008

(critique 1), et le fait de ne pas rendre compte de ménages qui auraient un profil de dépenses en énergie similaire indépendemment de leur niveau de revenu. La figure ci-dessous, Figure 4-1, présente la distribution des quintiles au sein des 8 groupes du regroupement sélectionné pour représenter les différents profils de dépenses en énergie des ménages québecois pour la résidence principale en 2008.

Comme l'illustre cette Figure 4-1, les 8 profils de dépenses en énergie identifiés ne concordent pas avec les quintiles. Ceux-ci se distribuent dans chaque groupe. Bien entendu, il est possible d'établir que certains quintiles sont concentrés dans certains groupes. Par exemple, le premier quintile est avant tout présent dans le groupe 1, puis le groupe 2 et enfin, dans une moindre mesure, le groupe 3. Alors que le quintile 5 est surtout présent dans le groupe 5, puis le groupe 4 et enfin, dans une moindre mesure, dans le groupe 3. Une analyse de l'évolution de la part de chaque quintile au sein de chaque groupe entre 1997 et 2009 est présentée dans la partie 4.3 du rapport.

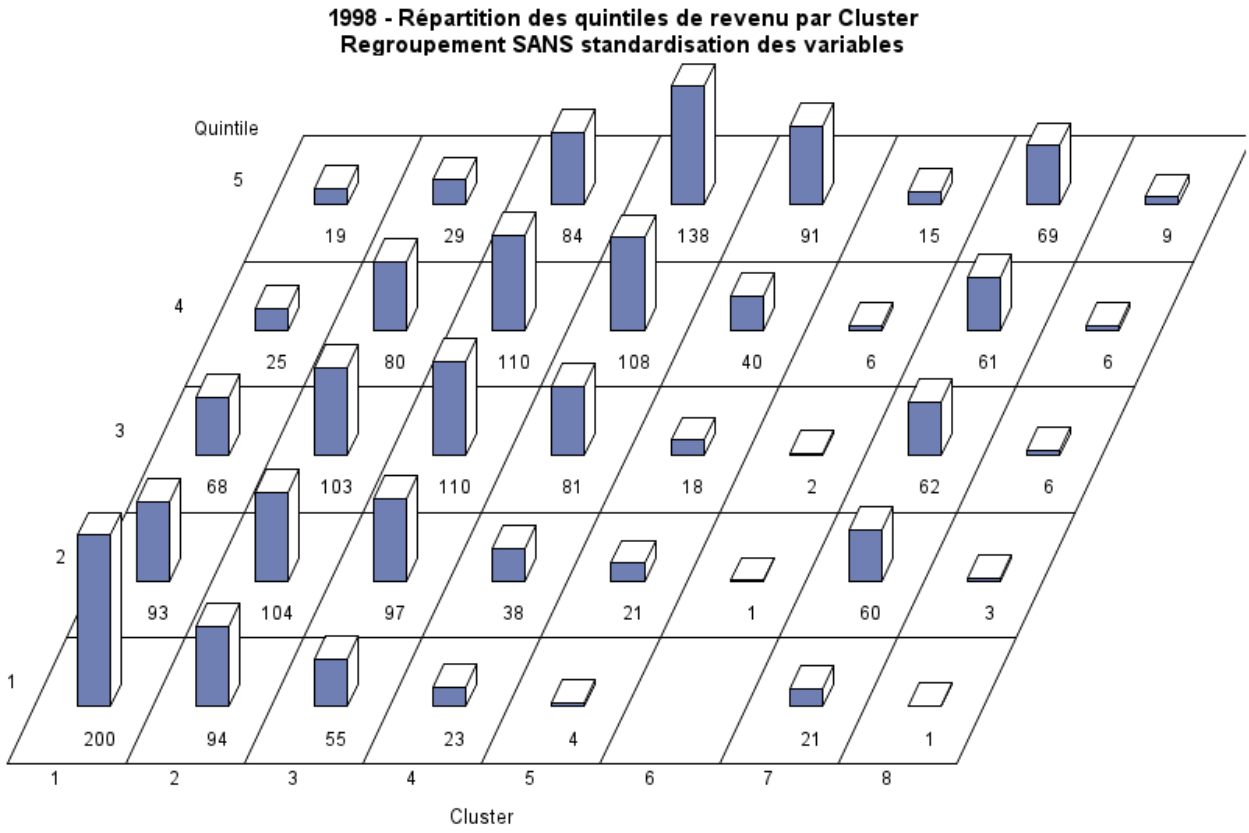


Figure 4-2 : Répartition des quintiles de revenu au sein des 8 groupes, Québec 1998

En 1998, la répartition des quintiles de revenu était similaire à la seule différence notable que les ménages du cinquième quintile étaient concentrés dans le groupe 4 en premier lieu et non dans le groupe 5 comme cela est le cas en 2008.

4.2 Québec, évolution du portrait des ménages de 1997 à 2009

4.2.1 L'évolution des centroïdes entre 1997 et 2009

Il est possible de suivre le regroupement à 8 groupes au cours du temps. La Figure 4-3 illustre l'évolution des centroïdes de 2008 à 2009. La taille des bulles qui représentent chaque groupe est proportionnelle à la taille du groupe.

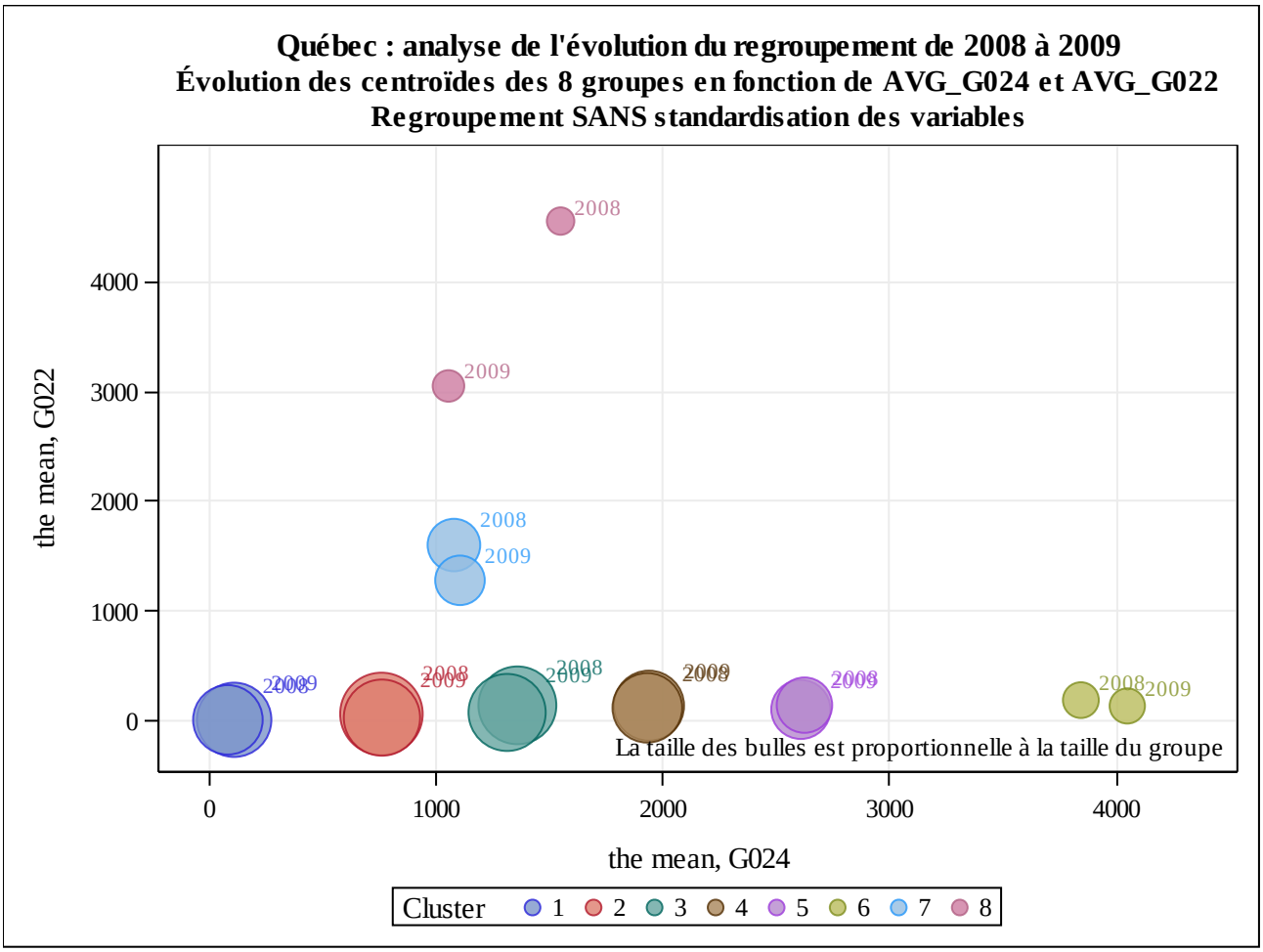


Figure 4-3 : Évolution des centroïdes, Québec, de 2008 à 2009.

Les groupes 1 à 5 (et 7) sont les plus constants dans le temps. Ce sont aussi les plus importants en taille. Par contre les groupes 6 et 8, qui représentent chacun moins de 5% de la population des ménages québecois enregistre la plus grande variation (en terme de centroïde) entre 2008 et 2009. Mais c'est le cas pour chaque année comme l'illustre la Figure 4-4 (voir aussi annexe 6.9.2).

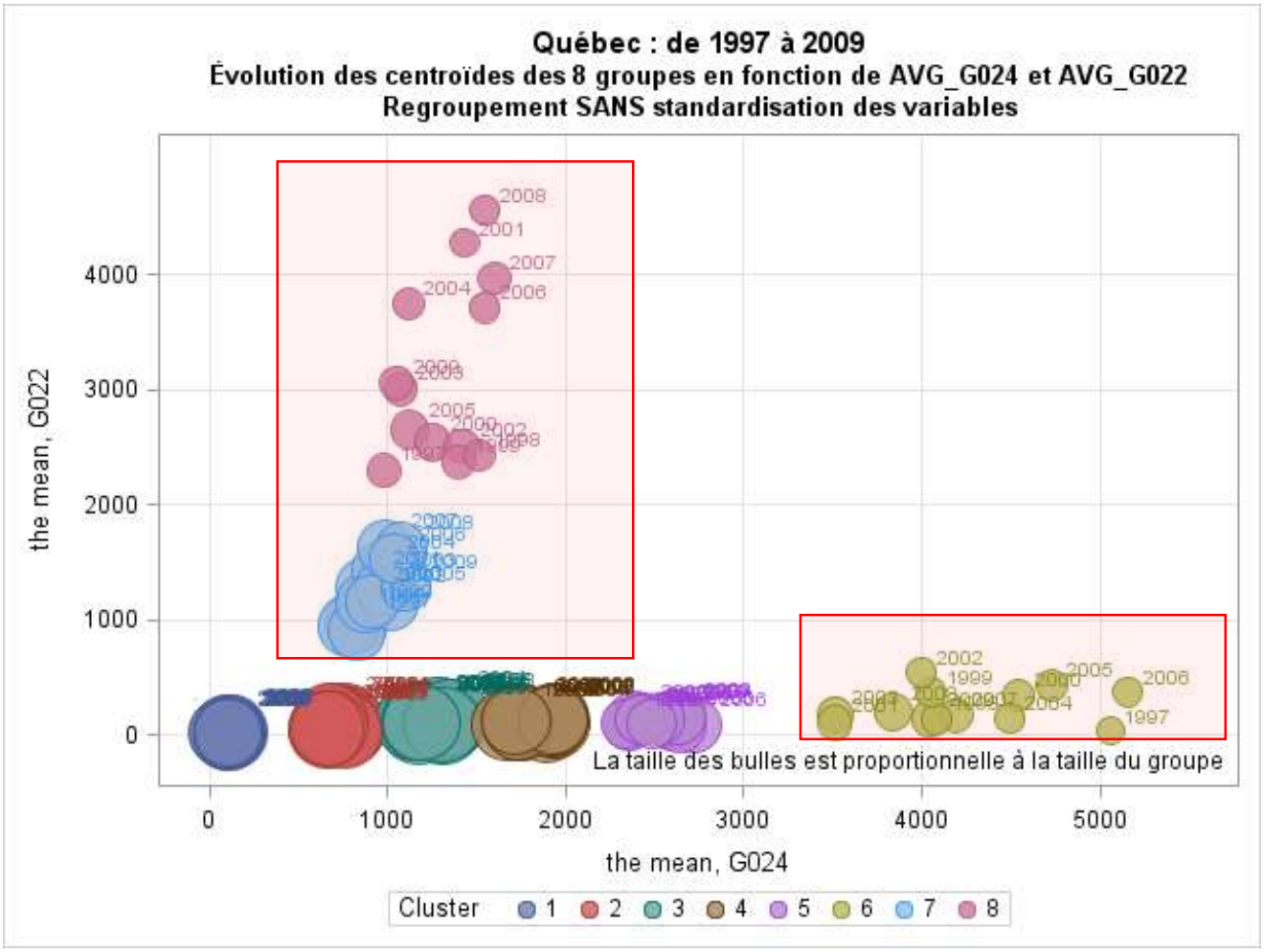


Figure 4-4 : Évolution des centroïdes, Québec, de 1997 à 2009.

Sur cette figure, ce sont les groupes 6 et 8 (dans une moindre mesure le groupe 7) qui connaissent le plus de variance entre 1997 et 2008 en termes de dépense moyenne pour le gaz (G022, en ordonnée sur la Figure 4-4) et de dépense moyenne pour l'électricité (G024, en abscisse sur la Figure 4-4).

4.2.2 L'évolution de la taille des groupes sur la période 1997 – 2009

La taille des groupes, représentée ci-dessous (Figure 4-5) en pourcentage de la population totale des ménages, est assez stable sur la période 1997 – 2009.

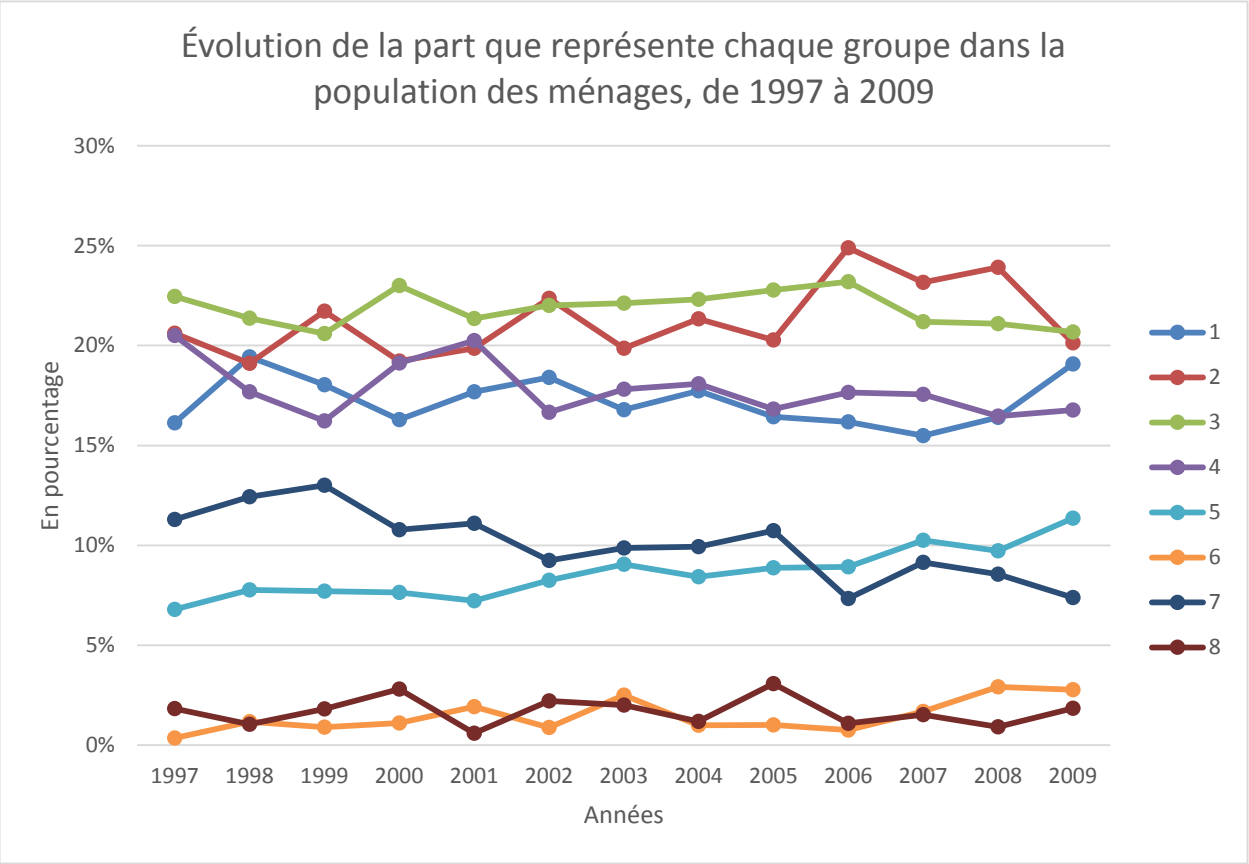


Figure 4-5 : Évolution de la taille des groupes, Québec. de 1997 à 2008.

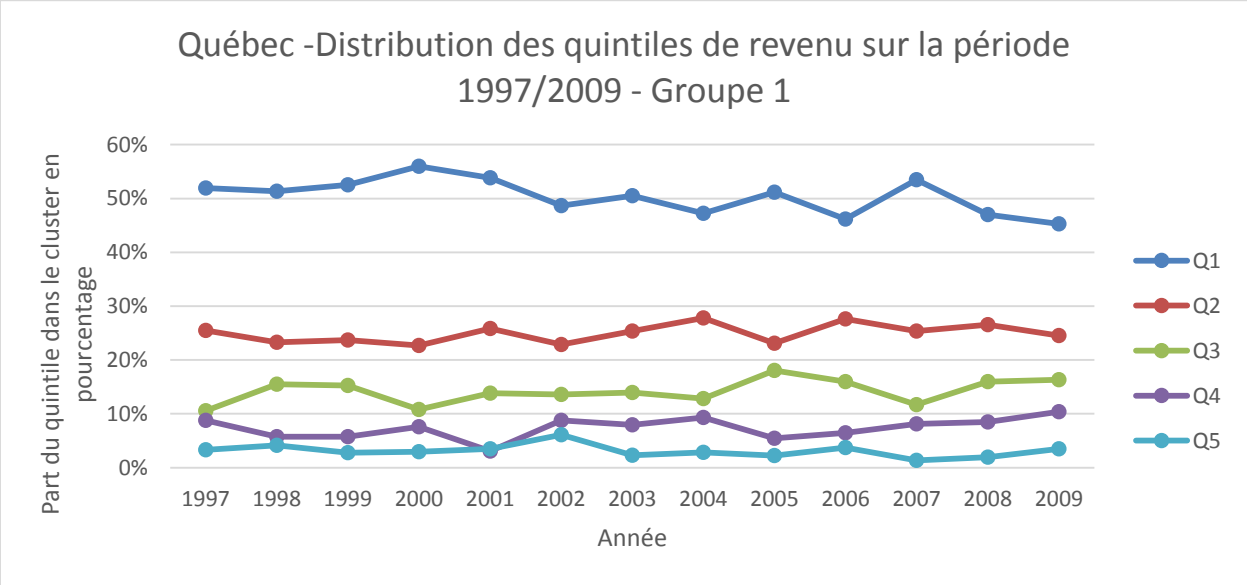
Il est intéressant de garder à l'esprit que les groupes 2 et 3 sont les plus importants et ils représentent, chacun, entre 20 et 30% de la population totale des ménages québecois sur la période 1997 – 2009. À l'opposé, les groupes 6 et 8 sont les plus petits (et connaissent les plus grandes variations, comme cela a été présenté dans la partie 4.2.1), ils représentent chacun moins de 5% de la population des ménages. Entre les deux, les groupes 4 et 1 représentent 15 à 20% des ménages. Les groupes 5 et 7 sont compris dans une fourchette de 5 à 15%. Ce sont les seules groupes pour lesquels il est possible d'observer une tendance en terme de taille : le groupe 5 est en hausse constante de 1997 à 2009, alors que le groupe 7 est en baisse constante.

4.3 Évolution de chaque groupe selon la dépense en électricité et la répartition des quintiles

Note : au Québec, entre 1997 et 2009, la moyenne des dépenses des ménages en électricité (pour les ménages qui ont des dépenses en électricité) est passsée de 1192$ à 1452$, soit une augmentation de 22%.

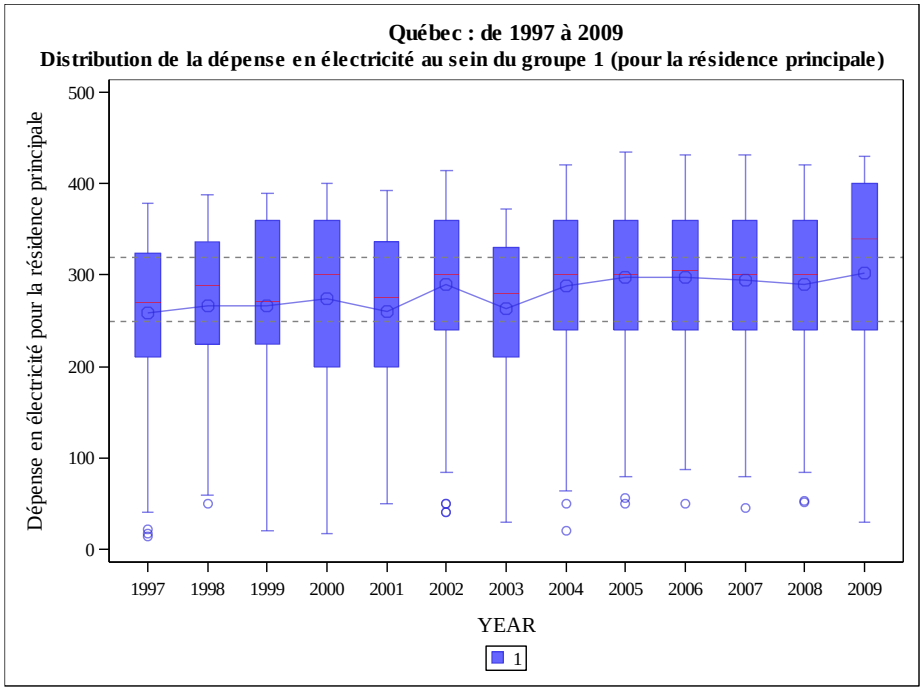
4.3.1 Groupe 1

Le groupe 1 est majoritairement composé des ménages qui appartiennent au premier quintile de revenu, soit près de 50% des ménages du groupe 1 appartiennent au premier quintile de revenu. Cette tendance semble en légère baisse depuis les années 2000 sans jamais vraiment s'éloigner de la barre des 50%. Puis, ce sont des ménages du deuxième et troisième quintiles de revenu qui composent le groupe 1. Enfin, représentant près de 10% (10% en 1997 et 10% en 2009, mais toujours en dessous de 10% entre ces deux dates) sans jamais dépasser cette limite, le groupe 1 se compose aussi avec des ménages du quatrième quintile. Enfin, toujours en dessous du seuil des 5%, le cinquième quintile de revenu n'est pas asbent du groupe 1.

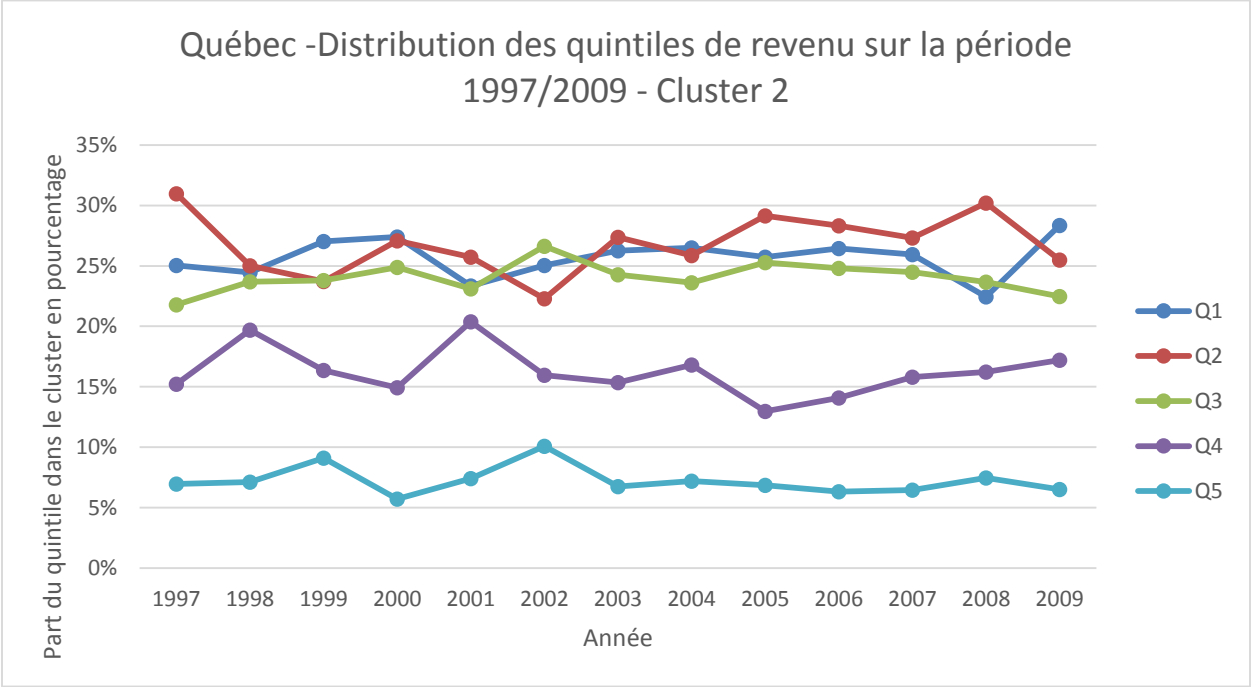


En ce qui concerne les caractéristiques du groupe 1 quant aux dépenses en électricité, celles-ci n'ont guère changé entre 1997 et 2009. Plus de 50% de ces ménages n'ont aucune dépense directe en électricité (car celles-ci sont comprises dans le loyer qu'ils payent en tant que locataire). **Pour les ménages de ce groupe qui ont une dépense en électricité**, comme l'illustre la figure ci-

dessous (voir annexe 6.9.3) , celle-ci ont légèrement augmenté entre 1997 et 2009. En moyenne, elles sont comprises entre 250$ et 320$ sur la période. Sur cette période, la dépense moyenne des ménages qui ont une dépense en électricité supérieure à 0 est passée de 260$ à 303$, soit une augmentation de 17% en 13 ans. L'augemtation du prix de l'électricité sur cette période a été de 15%.

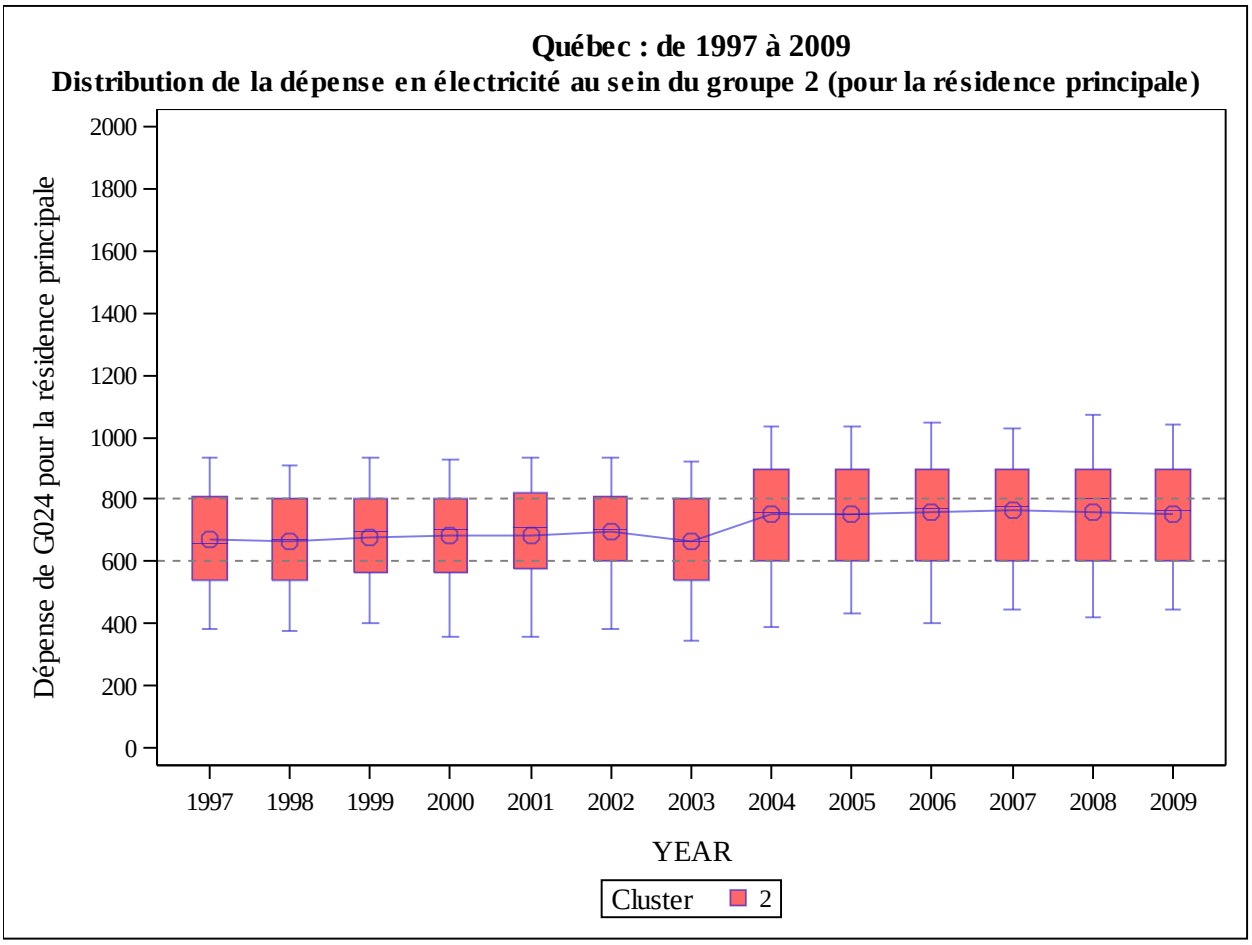


4.3.2 Groupe 2



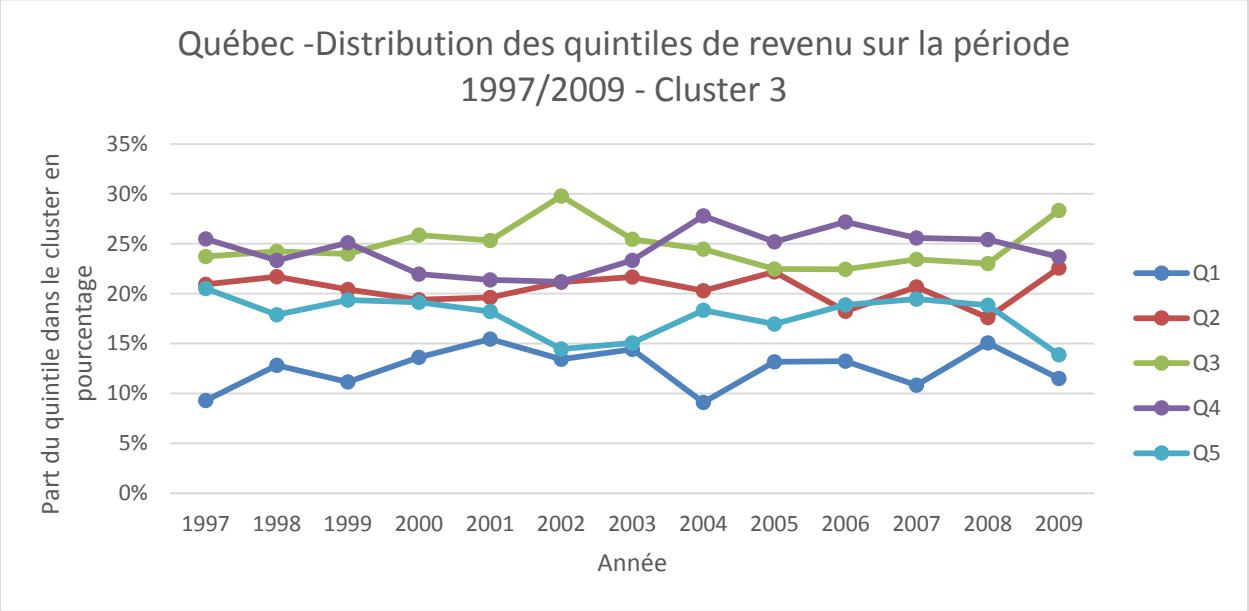
Le groupe 2 est composé en majorité des ménages appartenant au 1er, 2ème et 3ème quintiles de revenu, tous les trois représentant chacun entre 20% et 30% des ménages du groupe 2. Puis, ce

sont les ménages du 4ème quintile qui représentent entre 15 et 20% des ménages du groupe 2. Enfin, de 5 à 10% des ménages du groupe 2 appartiennent au 5ème quintile sur la période 1997 – 2009. Ces chiffres restent relativement constant sur la période.

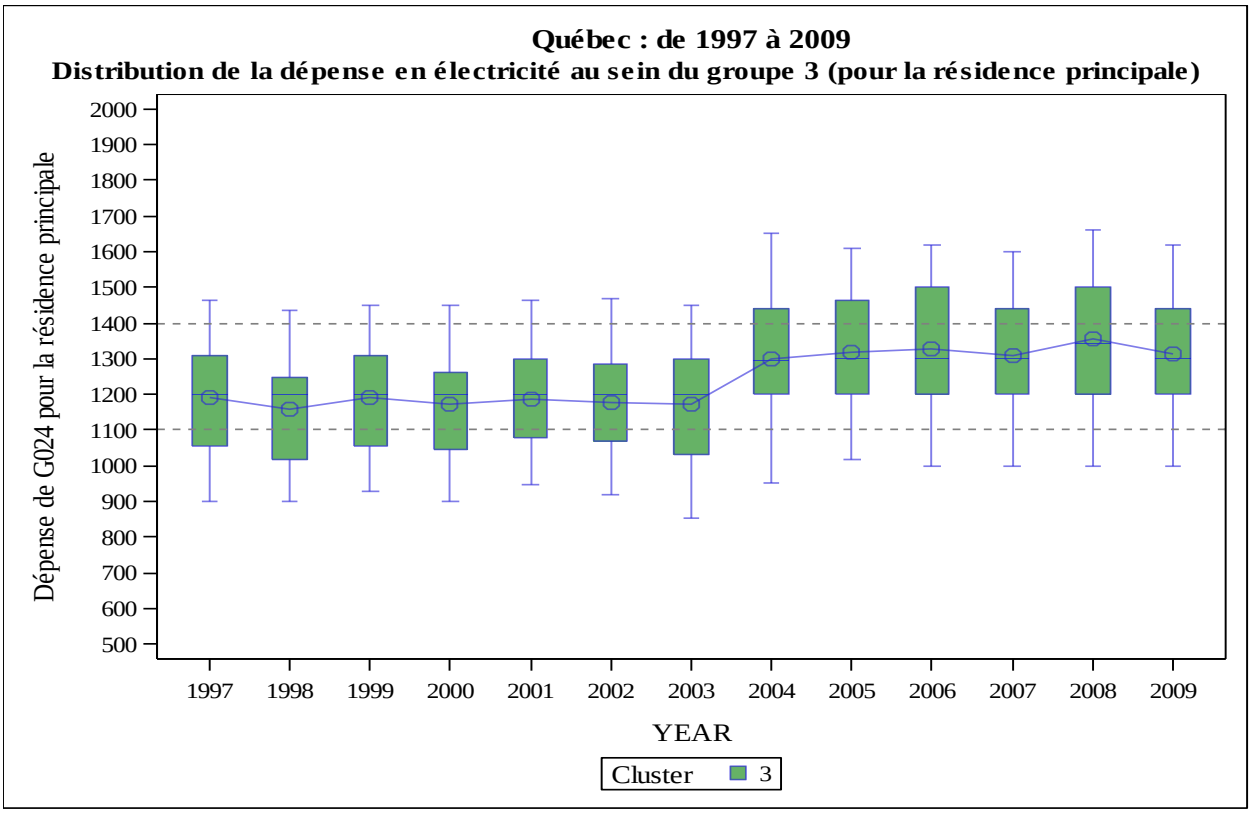


En ce qui concerne les caractéristiques du groupe 2 quant aux dépenses en électricité, celles-ci sont comprisent entre 600$ et 800$ sur la période. Entre 1997 et 2009, la moyenne passe de 671$ à 752$, soit une augmentation de 12% (le prix de l'électricité a augmenté lui de 15% sur cette même période).

4.3.3 Groupe 3



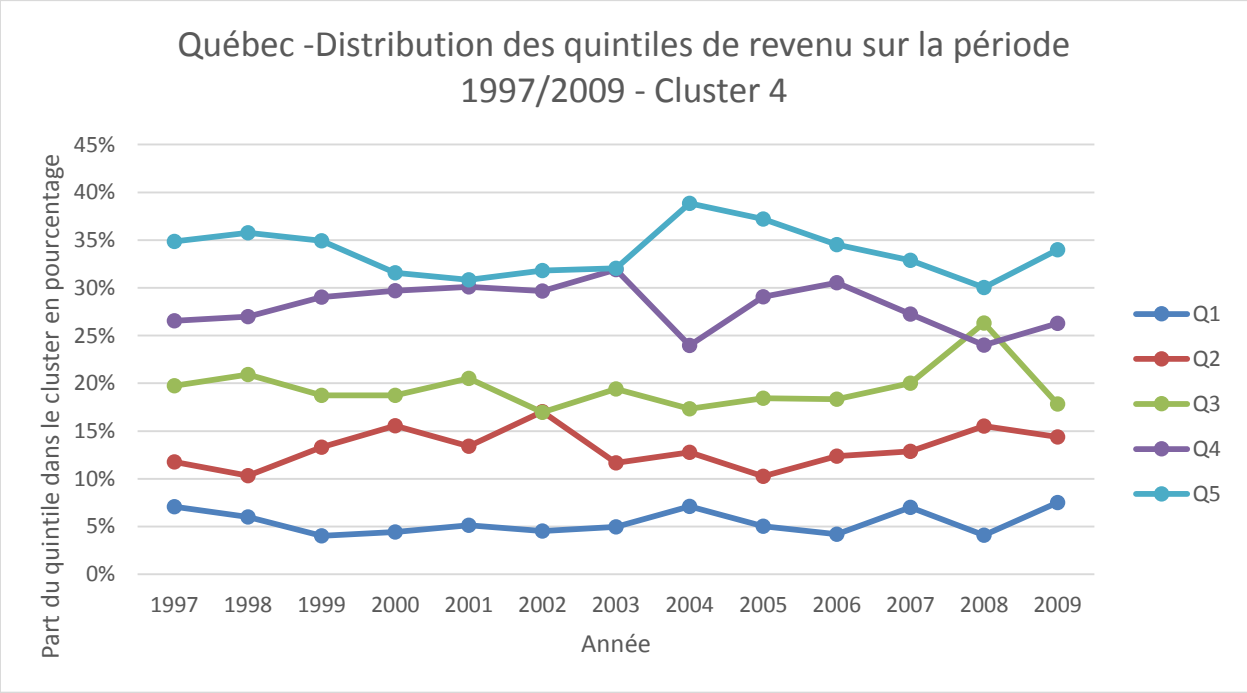
Le groupe 3 est composé de ménages qui appartiennent principalement au 3ème et 4ème quintiles de revenu (près de 50% en tout), puis ce sont des ménages du 2ème quintile qui forment ce groupe. Les ménages du 5ème quintile et du 1er quintile représentent moins de 20% et respectivement 15% des ménages de ce groupe.



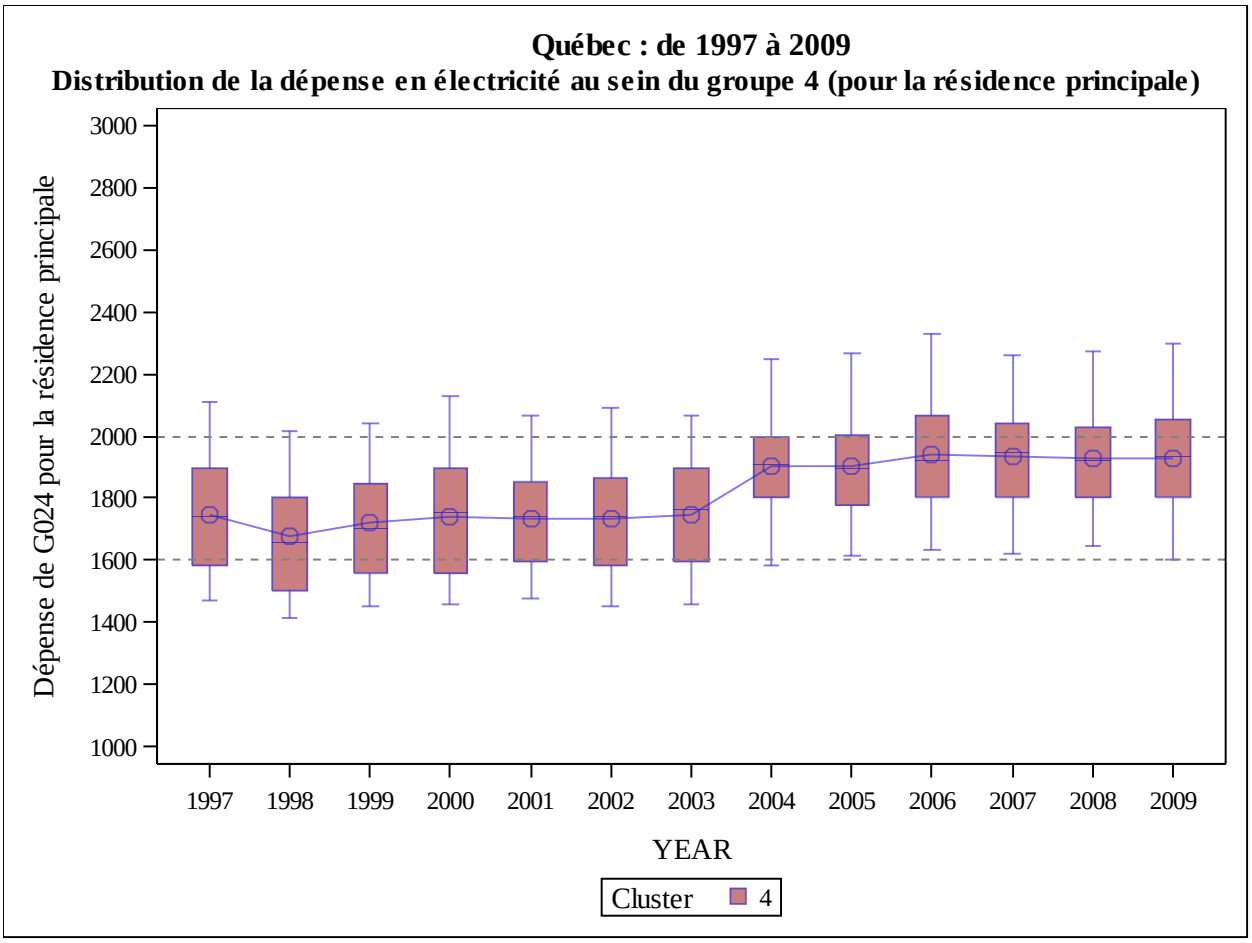
La dépense en électricité du groupe 2 varie entre 1100 et 1400$ sur la période de 1997 à 2009. La moyenne passe de 1191$ en 1997 à 1311$ en 2009, soit une augmentation de 10% (le prix de l'électricité a augmenté lui de 15% sur cette même période).

4.3.4 Groupe 4

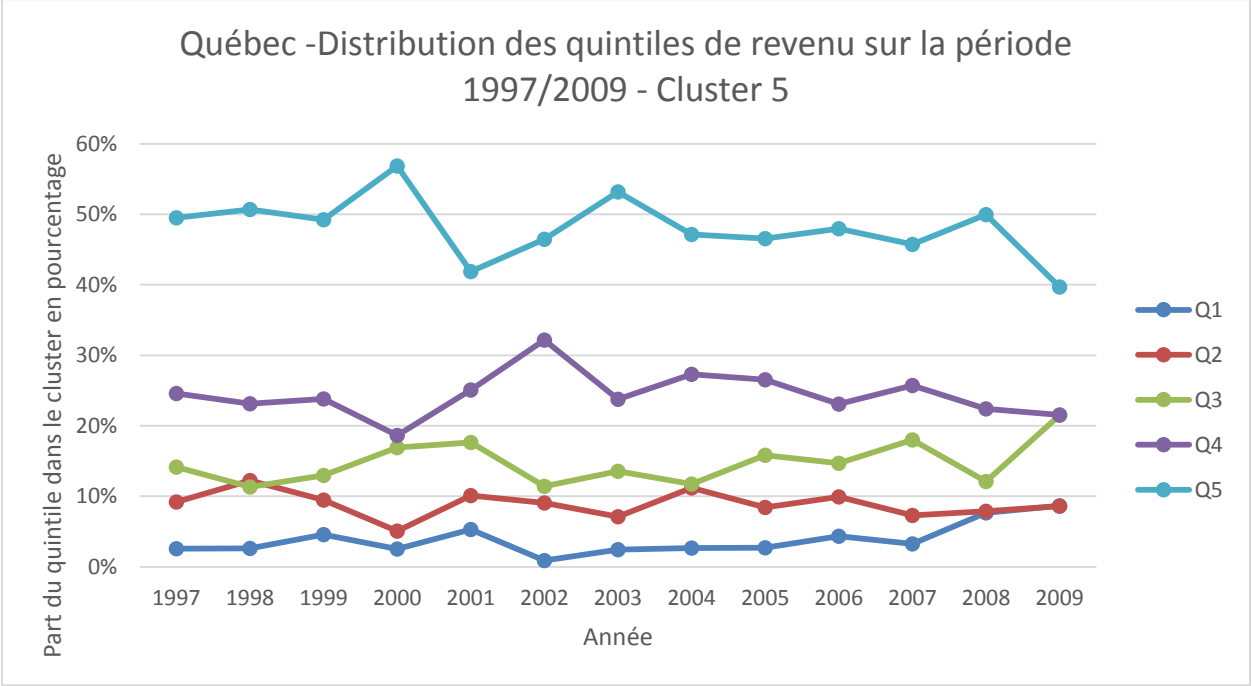
Les ménages du groupe 4 appartiennent pour la majorité au 5ème et 4ème quintiles, représentant près de 60% des ménages de ce groupe. Un autre 20% appartient au 3ème quintile. Entre 10% et 15% appartiennent au deuxième quintile. Et enfin, environ 5% des ménages sont du premier quintile.



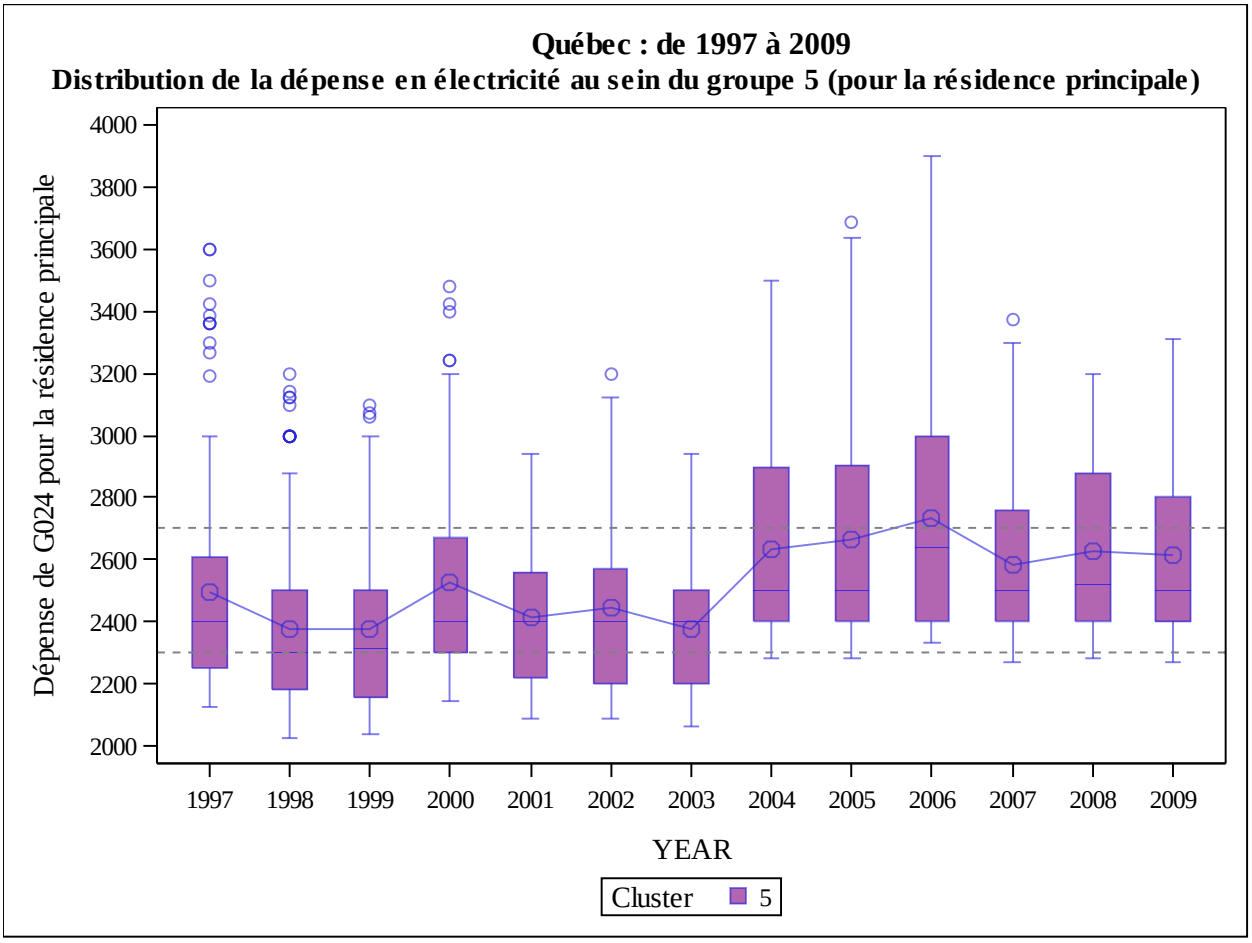
Pour les dépenses moyennes en électricité du groupe 4, elles sont comprises entre 1600 et 2000$ sur la période 1997-2009, la moyenne passant de 1748$ en 1997 à 1929$ en 2009, soit une augmentation de 10% (le prix de l'électricité a augmenté lui de 15% sur cette même période).


Québec : de 1997 à 2009
Distribution de la dépense en électricité au sein du groupe 4 (pour la résidence principale)
Dépense de G024 pour la résidence principale
3000
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
YEAR
Cluster
4

4.3.5 Groupe 5

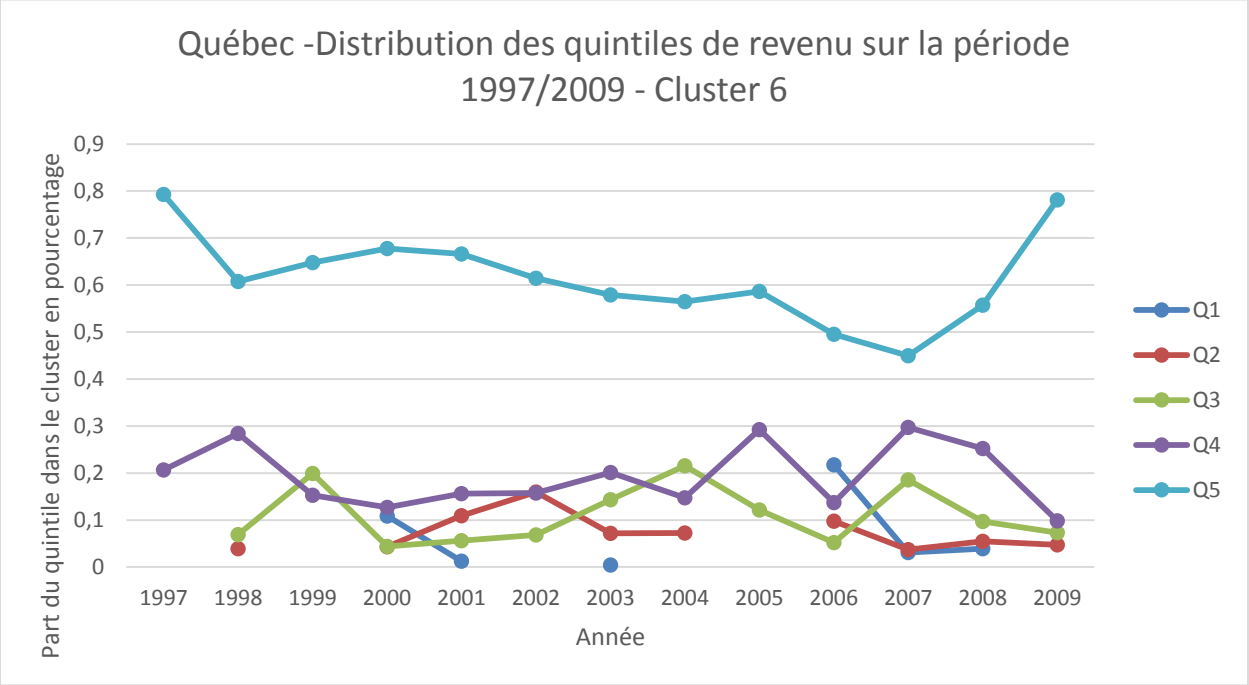

Québec -Distribution des quintiles de revenu sur la période 1997/2009 - Cluster 5
Part du quintile dans le cluster en pourcentage
60%
50%
40%
30%
20%
10%
0%
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
Année
Q1
Q2
Q3
Q4
Q5

Le groupe 5 est quant à lui composé à près de 50% de ménages du 5ème quintile. Puis, entre 20% et 30% des ménages composent ce groupe. Entre 10% et 20% des ménages sont du 3ème quintile. Les ménages appartenant au 2ème et 3ème quintiles représentent moins de 10% (chacun) du groupe 5.



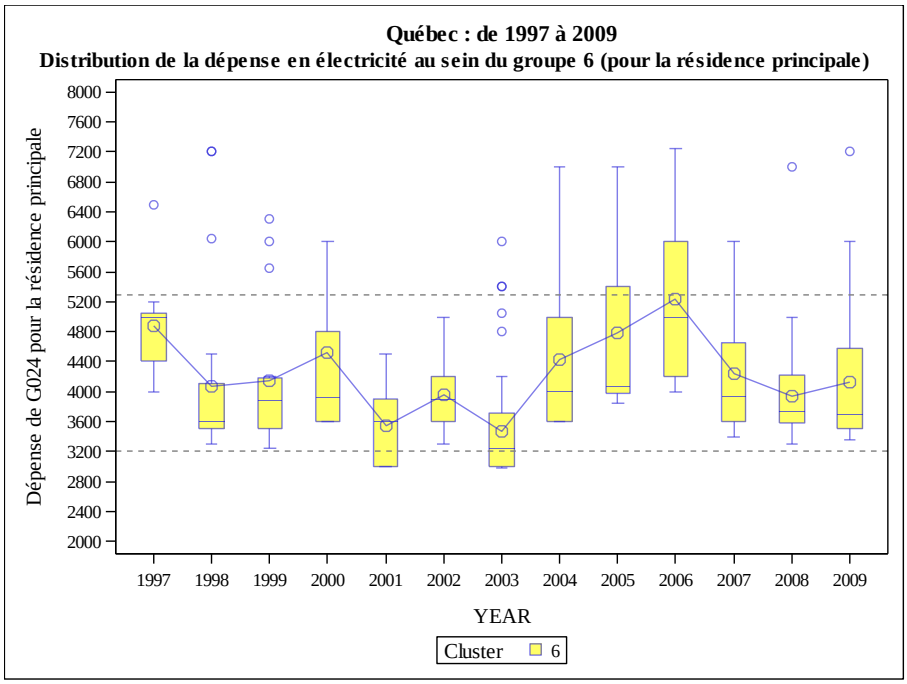
Les dépenses moyennes en électricité du groupe 5 sont comprises entre 2300$ et 2600$ sur la période 1997 – 2009. Elles passent de 2489$ en 1997 à 2608$ 2009, soit une augmentation de 5% en 13 ans.

4.3.6 Groupe 6

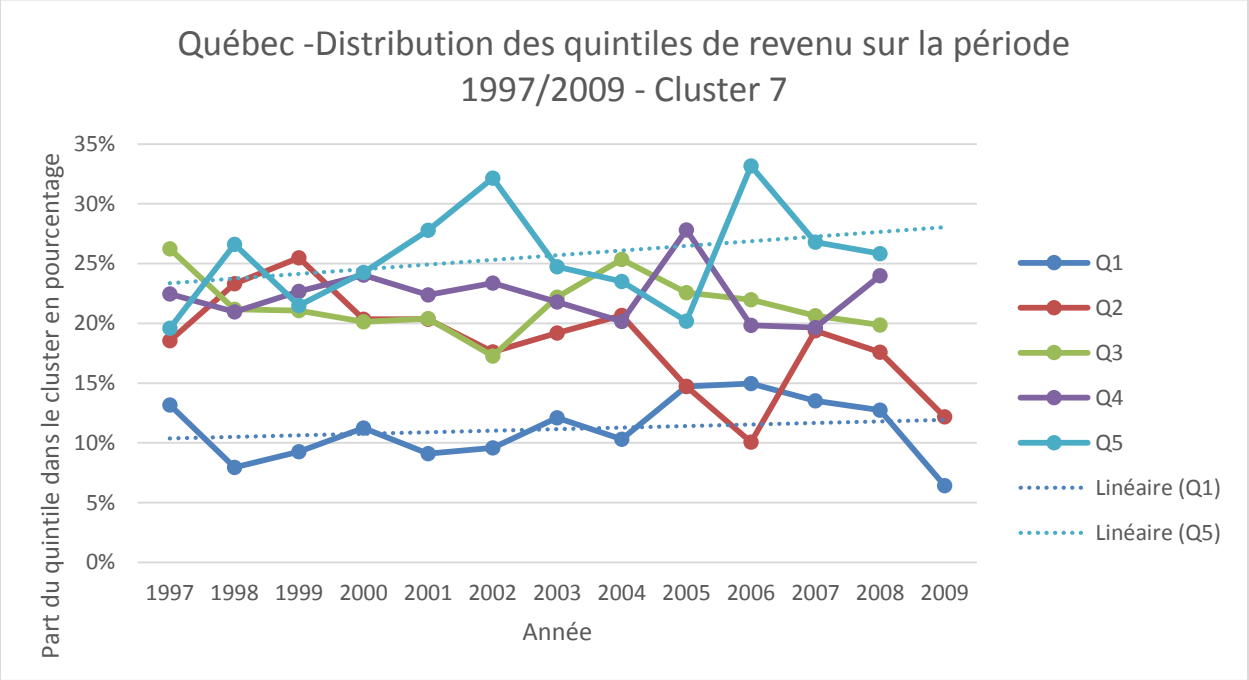


Les ménages du groupe 6 sont pour près de 60% du 5ème quintile. Les autres quintiles représentent moins de 30% des ménages du groupe 6 et parfois sont mêmes absents comme pour le 1er et 2ème quintile.

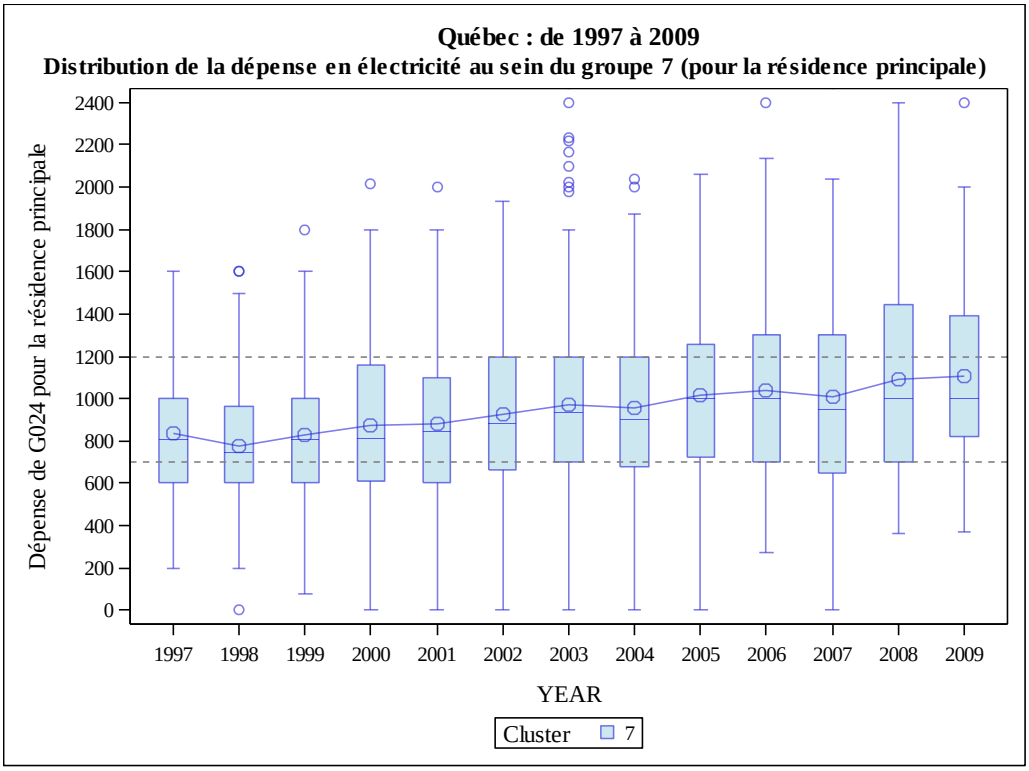
Pour ce qui a trait aux dépenses en électricité, on peut observer une grande variabilité sur la période. Le groupe 6 représente dans une certaine mesure des valeurs extrêmes. La moyenne des dépenses en électricité sur la période oscille entre 3200$ et 5200$.



4.3.7 Groupe 7

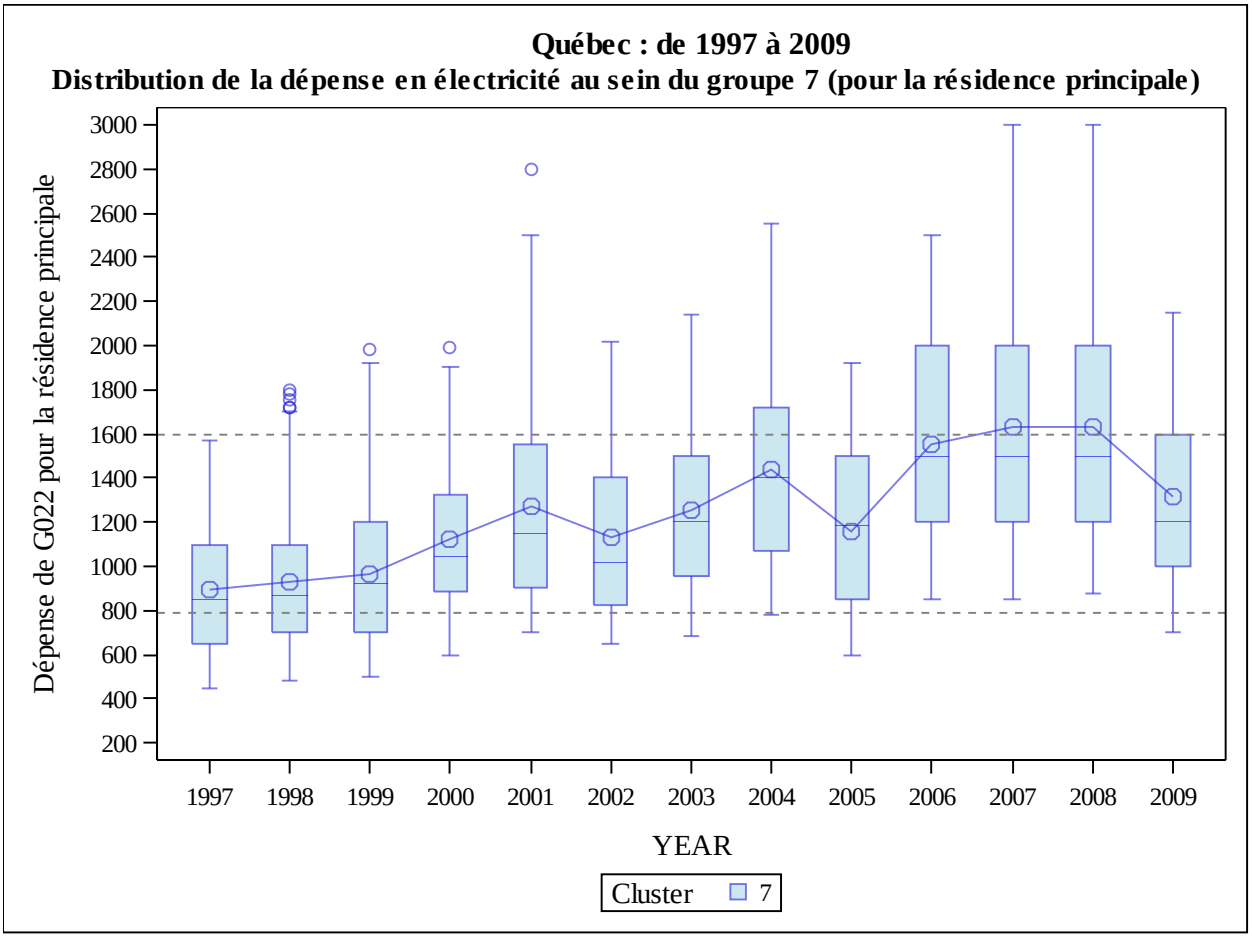


Le groupe 7 est principalement composé de ménages appartenant au 5ème quintile, ceux-ci représentant entre 20 et 30% des ménage du groupe 7. Puis, à part égale, le 4ème quintile et le 3ème quintile représentent entre 20 et 25% des ménages de ce groupe. Les ménages du 2ème quintile sont de moins en moins présents dans ce groupe (ce qui pourrait aussi expliquer la diminution de la taille du groupe 7 observée dans la partie 4.2.2. Enfin, les ménages appartenant au premier quintile représentent environ



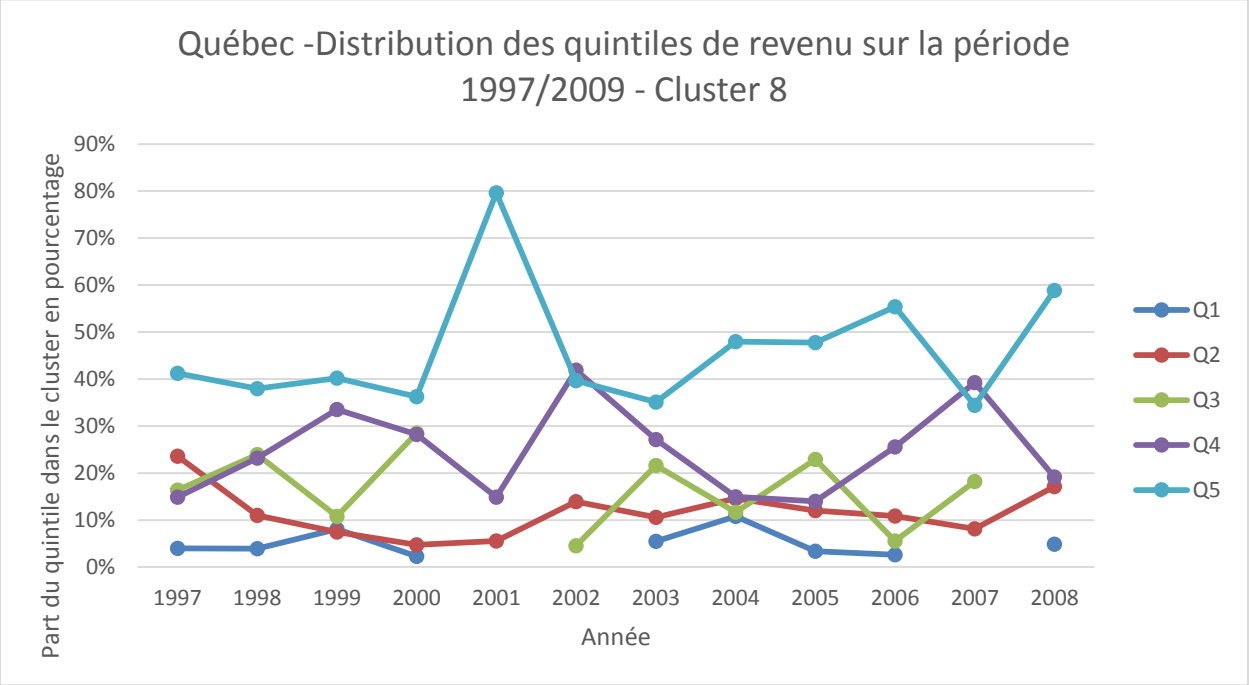
10% des ménages de groupe.

Les dépenses moyennes en électricité de ce groupe varient entre 600$ et 1200$ sur la période 1997 – 2009. Cette moyenne passe alors de 830$ en 1997 à 1105$ en 2009, soit une augmentation de 33%. Cette augmentation plus forte que chez tous les groupes précédents illustre l'équipement progressif en appareil de chauffage à électricité de ce groupe.

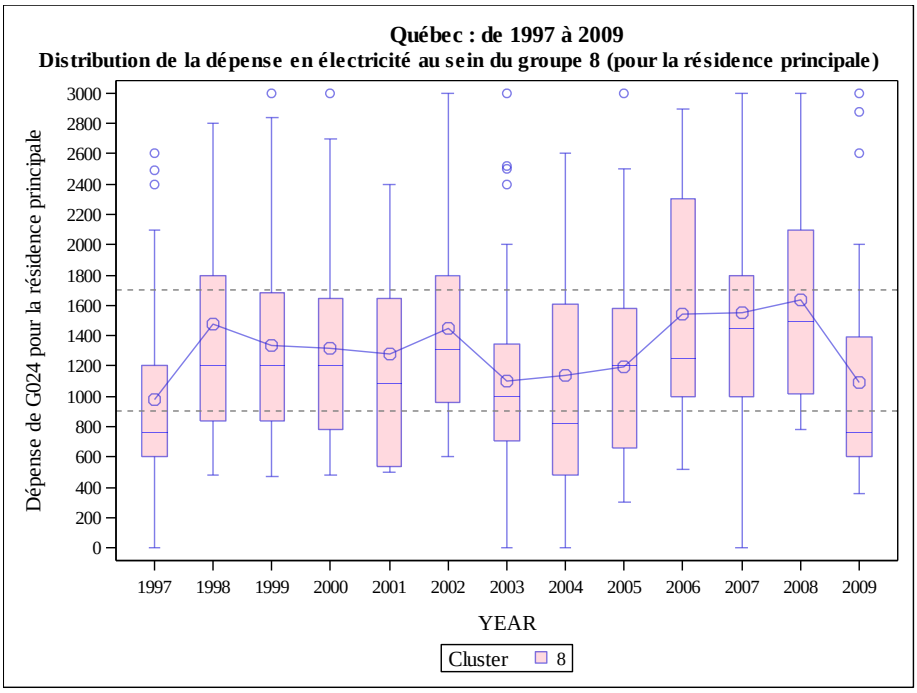


La dépense moyenne des ménages du groupe 7 en gaz est comprise entre 800$ et 1600$ sur la période 1997 – 2009. En 1997, la moyenne était de 893$ et celle-ci atteint 1282$ en 2009, soit une augmentation de 43% (sur cette même période le prix du gaz a augmenté de 76%).

4.3.8 Groupe 8

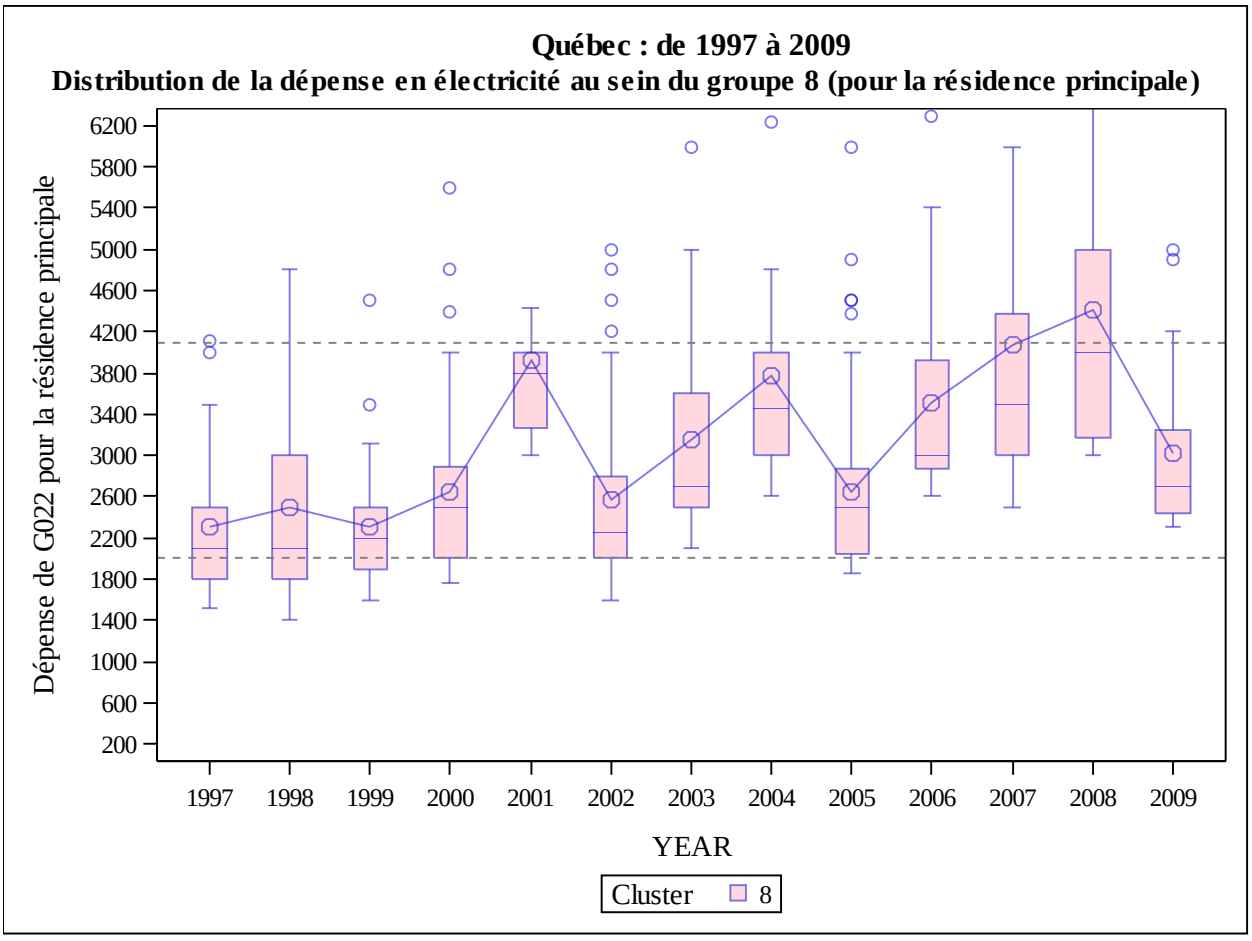


Le groupe 8 est constitué principalement des ménages appartenant au cinquième quintile (représentant plus de 40% des ménages du groupe). Puis, les ménages appartenant au 4ème quintile de revenu représentent entre 20 et 40% des ménages du groupe 8. Les ménages appartenant au 3ème et 2ème quintile représentent entre 10% et 20% des ménages du groupe. Enfin, les ménages du 1er quintile représentent moins de 10% et sont parfois absents du groupe 8 sur la période 1997 – 2009.



La dépense moyenne des ménages du groupe 8 en électricité est comprise entre 900$ et 1700$ sur la période 1997 – 2009. En 1997, la moyenne était de 983$ et celle-ci atteint 1549$ en **2008**, soit une augmentation de 53%. Cette

augmentation plus forte que chez tous les groupes précédents illustre l'équipement progressif en appareil de chauffage à électricité de ce groupe. Ce trait est caractéristique des groupes 7 et 8 qui s'équipent progressivement en appareil de chauffage à électricité.



En ce qui concerne la dépense moyenne du groupe 8 en gaz, celle-ci oscille entre 1800$ et 4200$ sur la période. Cette variation traduit la nature extrême des ménages qui constituent le groupe 8.

5 Conclusion

La méthodologie décrite dans ce document a permis de construire un regroupement en 8 groupes pour le Québec qu'il est possible de suivre sur la période 1997 à 2009. Suite à l'analyse, aucun fait ou caractéristique des groupes n'est une surprise ce qui est rassurant pour l'intelligibilité des groupes. Si les 8 groupes permettent de rendre compte de 8 profils de dépenses en énergie des ménages québecois pour la résidence principale, il sera intéressant de confronter cette structure avec les autres provinces canadiennes et, éventuellement, d'établir des comparaisons. Aussi, il pourrait être intéressant d'intégrer les dépenses en énergie pour le transport dans la construction du regroupement qui sont aujourd'hui plus importantes que les dépenses en énergie pour la résidence principale (en moyenne) afin de rendre compte d'un profil de dépenses en énergie des ménages complet. Enfin, sur un plan méthodologique, il serait intéressant de mieux prendre en compte les valeurs extrêmes dans la construction du regroupement. Enfin, les données horaires de consommation électrique, disponibles depuis la mise en place des compteurs dit « intelligents », offrent des possibilités de segmentations du secteur résidentiel très interessantes quant à la mise en place de stratégies d'économie d'énergie.

6 Annexes

6.1 Annexe - La liste des variables complètes des données de 1997 à 2009, en ordre alphatbétique

#	Nom de la variable	Description de la variable
17	ADRENALT	Additions, renovations and alterations
52	AGEYOUNP	Age of youngest child (member sometime in 1997)
71	AIRCONP	Air conditioning
20	BUYHMFLG	Purchased dwelling in x year (year of survey) indicator
76	CABLETV	Cablevision
1	CASEID	Identification number
75	CDPLYER	Compact disc player
244	CH04TOI	Number of children aged 0 to 4 years in the household at December 31
245	CH517TOI	Number of children aged 5 to 17 years in the household at December 31
78	COMPUTER	Home computer
25	COOKINGP	Principal cooking fuel
66	DRYERP	Clothes dryer
67	DSHWASHP	Dishwasher
235	DVD	DVD player
46	ECFAM	Number of economic families in the household
64	EIFLAG	Employment insurance benefit indicator
84	F001	Total food
85	F002	Food purchased from stores
86	F008	Food purchased from restaurants
69	FREEZER	Freezer
3	FYPYFLAG	Full-year/part-year household indicator
87	G001	Total shelter
88	G002	Principal accommodation
89	G003	Rented living quarters
90	G004	Rent (amount paid)
91	G007	Owned living quarters
94	G009	Condominium charges on principal residence
95	G010	Property taxes on principal residence

#	Nom de la variable	Description de la variable
96	G011	Homeowners' insurance premiums on principal residence
97	G019	Water, fuel and electricity for principal residence
98	G020	Water and sewage for principal residence
237	G021	Natural gas
99	G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)
238	G023	Other fuel
100	G024	Electricity for principal residence
101	G025	Other accomodation
102	G026	Owned vacation home
103	G032	Traveller accommodation
92	G040	Regular mortgage payments
104	H001	Total household operation
105	H002	Communications
106	H003	Telephone
107	H004	Purchase of telephones
108	H005	Telephone services
109	H008	Cellular services
110	H009	Internet services
111	H010	Postal other communication services
112	H011	Child care expenses
113	H016	Domestic and other custodial services (excluding child care)
114	H017	Pet expenses
115	H022	Household cleaning supplies
116	H023	Paper, plastic foil household supplies
117	H026	Garden supplies and services (excluding tools)
239	H070	On-line services
22	HEATAGE	Age of principal heating equipment
23	HEATFUEP	Principal heating fuel
21	HEATMAIP	Principal heating equipment
57	HHINCEAR	Household income from earnings
58	HHINCINV	Household income from investment
60	HHINCOTH	Household income from other sources
56	HHINCTOT	Household income before taxes

#	Nom de la variable	Description de la variable
59	HHINCTRA	Household income from government transfer payments
51	HHSZTOTP	Household size (number of persons a member sometime in 1997)
47	HHTYPEP	Household type
24	HOTWATEP	Principal heating fuel for hot water
118	I001	Total household furnishings and equipment
119	I002	Household furnishings
120	I003	Furniture
121	I004	Rugs, mats and underpadding
122	I005	Window coverings and household textiles
123	I006	Art, antiques and decorative ware
124	I010	Household equipment
125	I011	Household appliances
127	I042	Maintenance and repair of furniture and equipment
128	I046	Services related to furnishings and equipment
126	I027TOT	Other household tools, equipment and accessories
80	INTERNET	Internet use from home
129	J001	Total clothing
130	J002	Women's and girls' wear including gifts (4 years and over)
131	J007	Clothing gifts to non-household members - Women and girls (4 years and over)
132	J008	Men's and boys' wear including gifts (4 years and over)
133	J013	Clothing gifts to non-household members - Men and boys (4 years and over)
134	J014	Children's wear including gifts (under 4 years)
135	J017	Clothing gifts to non-household members - Children (under 4 years)
137	J021	Clothing services
136	J041	Clothing material and notions
138	J023TOT	Laundry and dry-cleaning
139	K001	Total transportation
140	K002	Private transportation
141	K003	Purchase of automobiles and trucks
142	K007	Accessories and attachments for automobiles and trucks
143	K008	Rented and leased automobiles and trucks
144	K019	Operation of owned and leased automobiles and trucks
145	K020	Gasoline and other fuels

#	Nom de la variable	Description de la variable
146	K022	Automotive maintenance and repair jobs
147	K023	Garage rent and parking
148	K028	Private and public vehicle insurance premiums
149	K031	Public transportation
152	K037	Other passenger transportation
153	K038	Household movers, storage and delivery services
150	K032TOT	Local and commuter transportation
151	K034TOT	Inter-city transportation
154	L101	Total health care
155	L102	Direct health care costs to household
156	L103	Health care supplies
157	L104	Medicinal and pharmaceutical products
158	L105	Prescription drugs
159	L106	Other medicines, drugs and pharmaceutical products
160	L107	Physicians' care
161	L108	Eye care goods and services
162	L112	Dental care
163	L114	Hospital care
164	L116	Other health care practitioners
165	L117	Other medical services
166	L118	Health insurance premiums
167	L119	Public hospital, medical and drug insurance plans
168	L120	Private health care insurance plans
169	L201	Total personal care
170	L202	Personal care supplies and equipment
171	L207	Hair grooming and other personal care services
172	M101	Total recreation
173	M102	Recreation equipment and associated services
174	M103	Sports and athletic equipment
175	M110	Computer equipment and supplies
176	M116	Photographic goods and services
177	M126	Recreation vehicles and associated services
178	M127	Purchase of recreation vehicles

#	Nom de la variable	Description de la variable
179	M128	Bicycles, parts and accessories
180	M129	Purchase of other vehicles and outboard motors
181	M139	Operation of recreational vehicles
182	M140	Bicycle maintenance and repairs
185	M148	Home entertainment equipment and services
186	M150	Audio equipment (e.g., radio, CD player, speaker, etc.)
187	M151	Compact discs, audio tapes, videotapes and video-discs
190	M156	Rental of videotapes and video-discs
191	M157	Rental of home entertainment equipment and other services
192	M159	Recreation services
193	M160	Entertainment
194	M161	Movie theatres
195	M162	Live sports events (as a spectator)
196	M163	Live performing arts
201	M164	Admission to museums and other activities and venues
197	M165	Rental of cablevision and satellite services
198	M166	Use of recreation facilities
199	M167	Membership fees and dues for sports and recreation facilities
200	M171	Children's camps
202	M172	Package travel tours
203	M173	Other cultural and recreational services
184	M180	Expenses for rented and leased recreational vehicles
189	M186	Television sets, VCRs, camcorders and other television/video components
188	M187	Blank audio and videotapes
204	M201	Total reading materials and other printed matter
205	M202	Newspapers
206	M203	Magazines and periodicals
207	M204	Books and pamphlets (excluding school books)
208	M301	Total education
210	M308	Tuition fees
240	M408	Cablevision services
241	M409	Satellite services
246	M106TOT	Toys, elecronic games and arts/hobby materials

#	Nom de la variable	Description de la variable
183	M142TOT	Other recreation vehicle operation
209	M302TOT	Education supplies and textbooks
55	MAINCSCP	Household major source of income
70	MICROWV	Microwave oven
79	MODEM	Modem
62	MONFLOWS	Money flows of household (assets, loans and other debts)
211	N101	Total tobacco products and alcoholic beverages
212	N102	Tobacco products and smokers' supplies
214	N105	Alcoholic beverages
215	N106	Alcoholic beverages served on licensed premises
216	N107	Alcoholic beverages purchased from stores
217	N108	Self-made alcoholic beverages
213	N109	Cigarettes, cigars and similar products
221	N201	Games of chance expenses (net)
236	NETCONEP	Internet connection
18	NEWINST	New installations equipment and fixtures
81	NMCOLTVP	Number of colour televisions
68	NMFRIDGP	Number of refrigerators
73	NMPHNUMP	Number of telephone numbers
72	NMPHONP	Number of telephones
77	NMVCRP	Number of video cassette recorders
82	NMVEHONP	Number of vehicles owned at December 31
10	NUMBEDRP	Number of bedrooms
11	NUMBTHRP	Number of bathrooms
74	NUMCELLP	Cellular telephone
53	NUMFTP	Number of full-time earners
54	NUMPTP	Number of part-time earners
9	NUMRMP	Number of rooms
218	O101	Miscellaneous expenses
219	O104	Financial services
220	O109	Dues to unions and professional associations
222	O201	Personal taxes
223	O301	Personal insurance payments and pension contributions

#	Nom de la variable	Description de la variable
225	O304	Employment insurance premiums (EI)
226	O305	Retirement and pension fund payments
227	O306	Canada and Quebec Pension Plan (CPP/QPP)
228	O307	Other government retirement pension
229	O308	Other retirement or pension funds (excluding RRSPs)
230	O401	Gifts of money and contributions
242	O405	Charitable contributions
224	O302TOT	Life insurance premiums
243	O403TOT	Gifts of money and other support payments to persons living inside and outside Canada
93	P140	Maintenance, repairs and replacements
49	P2564TOI	Number of adults aged 25 to 64 years in the household at December 31
4	PROVINCP	Province/Territory
14	PURPRICP	Purchase price of home bought in 1997
26	REDURENT	Reduced Rent Paid
16	REPAIRS	Dwelling repairs needed
33	RPAGEGRP	Age group of reference person
28	RPINCEAR	Income of reference person from earnings
29	RPINCINV	Income of reference person from investment
31	RPINCOTH	Income of reference person from other sources
27	RPINCTOT	Income of reference person before taxes
30	RPINCTRA	Income of reference person from government transfers
32	RPMARP	État matrimonial de la personne de référence
34	RPSEX	Sex of reference person
35	RPWEEKFT	Number of weeks worked full-time
36	RPWEEKPT	Number of weeks worked part-time
63	RRSPCHNG	Household Registered Retirement Savings Plan (RRSP) change
234	SATELLIT	Satellite dish
50	SE65TOIP	Number of seniors aged 65 and over in the household at December 31
15	SELPRICE	Selling price of home sold in 1997
42	SPAGEGRP	Age group of spouse
38	SPINCEAR	Income of spouse from earnings
39	SPINCINV	Income of spouse from investment

#	Nom de la variable	Description de la variable
41	SPINCOTH	Income of spouse from other sources
37	SPINCTOT	Income of spouse before taxes
40	SPINCTRA	Income of spouse from government transfer payments
43	SPSEXP	Sex of spouse
44	SPWEEKFT	Number of weeks worked full-time by spouse
45	SPWEEKPT	Number of weeks worked part-time by spouse
13	TENTOIP	Dwelling tenure at December 31
12	TENURYRP	Dwelling tenure during reference year
231	TOTCUCON	Total current consumption
232	TOTEXPEN	Total expenditures
61	TOTHMONR	Total other money receipts of household (Includes net winnings from games of chance)
7	TYPDWELP	Type of dwelling occupied at December 31
5	URBRUR	Urban/rural code
6	URBSIZEP	Size of area of residence
83	VEHLEASP	Vehicles leased at December 31, 1997
65	WASHERP	Washing machine
2	WEIGHT	Public-use microdata weight
48	Y1824TOI	Number of youths aged 18 to 24 years in the household at December 31
233	YEAR	Year of the interview
8	YEARBUIP	Period of construction
19	YRMOVEDP	Year of move code

6.2 Annexe - Le résumé chronologique des changements (“Résumé chronologique des changements - Enquête sur les dépenses des ménages (EDM),” n.d.).

« Ce programme statistique a débuté : 1997

6.2.1 Période de référence de changement -2012

En 2012, la taille d'échantillon pour le journal des dépenses a été réduite à 50% de l'échantillon total étant donné que la base de référence des dépenses du journal a été établie durant les premières années de sa collecte. Des modifications correspondantes ont été apportées à la méthode de pondération pour les données collectées à l'aide du journal.

Des données ont également été collectées dans les territoires en 2012 selon le modèle précédent de l'EDM (questionnaire avec période de rappel de 12 mois uniquement, aucun journal de dépenses). De ce fait, les estimations des territoires ne sont pas combinées avec les estimations des provinces mais présentées séparément (voir section 6 du « Guide de l'utilisateur, Enquête sur les dépenses des ménages, 2012 »).

6.2.2 Période de référence de changement -2011

La méthode d'ajustement des journaux incomplets a été améliorée lors de l'EDM 2011. De même, l'âge des membres du ménage est maintenant défini à la date de l'entrevue plutôt qu'au 31 décembre de l'année de l'enquête.

Dans le but d'assurer la comparabilité des données, les données de 2010 ont également été révisées pour prendre en compte ces changements. L'estimation de 2010 révisée des dépenses moyennes par ménage en biens et services de toutes sortes a crû de 1,3% lorsque comparée à l'estimation de 2010 déjà publiée en avril 2012.

6.2.3 Période de référence de changement -2010

Une nouvelle méthodologie qui combine l'utilisation d'un questionnaire et d'un journal pour collecter les dépenses des ménages a été instaurée pour l'enquête de 2010. Les périodes de référence ont été réduites pour plusieurs items de dépenses et la collecte est effectuée en continue au cours de l'année. Bien que les dépenses recueillies depuis l'EDM 2010 soient similaires à celles des années antérieures, **les changements apportés aux méthodes de collecte des données, de traitement et d'estimation ont causé une rupture dans les séries de données.** Par conséquent,

il faut faire preuve de prudence au moment de comparer les données de l'EDM depuis 2010 à celles d'années antérieures, à moins d'indication contraire.

Le contenu de l'enquête a également fait l'objet d'une revue en 2010 dans le but de réduire le temps d'entrevue. Par conséquent, plusieurs composantes de l'équipement ménager et des caractéristiques du logement ont été exclues ainsi que la plupart des questions relatives à la variation de l'actif et du passif des ménages. Des modifications ont également été apportées à certaines définitions. De plus, débutant avec l'enquête de 2010, les données relatives au revenu des ménages et à l'impôt sur le revenu proviennent principalement d'une source de données administratives.

6.2.4 Période de référence de changement -2009

Pour l'année de référence 2009, un questionnaire par interview téléphonique assistée par ordinateur intitulé « Enquête sur les rénovations et les réparations du logement » a été utilisé pour recueiller des informations au sujet des dépenses en rénovations et réparations du logement. Le questionnaire est basé sur un sous-ensemble de questions demandées durant l'IPAO. Les données ainsi recueillies ont été combinées avec les mêmes données de l'IPAO permettant de produire des estimations de meilleure qualité pour le Système de comptabilité nationale du Canada.

6.2.5 Période de référence de changement -2007

À cause du passage à un questionnaire électronique IPAO, on a dû modifier les mesures de traitement des données et de contrôle de la qualité. En plus des règles de validation automatiques dans le questionnaire électronique, des règles d'équilibrage comparant le revenu total, les dépenses et les changements des avoirs et des dettes appliquées par les intervieweurs ont été réinstaurées en 2007.

6.2.6 Période de référence de changement -2006

Population cible - À compter de 2006, la distinction entre les membres et les ménages pour l'année entière et pour une partie de l'année n'existe plus. Les données portant sur les dépenses sont recueillies relativement à l'année de référence pour tous les membres du ménage au moment de l'interview.

Cette collecte s'est faite par interviews sur place à l'aide de questionnaires électroniques et d'ordinateurs portatifs (il s'agit du système d'interviews sur place assistées par ordinateur ou IPAO) par opposition aux questionnaires papier utilisés avant l'année de référence 2006.

Détection des erreurs - À cause du passage à un questionnaire électronique IPAO, on a dû modifier les mesures de traitement des données et de contrôle de la qualité. Des vérifications automatiques intégrées à ce questionnaire ont remplacé les contrôles d'équilibre et les vérifications dans les bureaux régionaux qui se faisaient auparavant.

Comme les données de la période 1997-2005 ne se rapportaient qu'aux membres pour l'année entière, afin de maintenir la comparabilité avec les années antérieures, les données de la période 2002-2005 ont été révisées de façon à inclure à la fois les ménages pour l'année entière et les ménages pour une partie de l'année. Les données de la période 1997-2001 seront révisées ultérieurement.

6.2.7 Période de référence de changement -2004

À compter de l'année de référence 2003, l'Enquête sur les réparations et les rénovations effectuées par les propriétaires-occupants (ERRP -- numéro d'enregistrement 3886) fait partie de l'Enquête sur les dépenses des ménages.

Estimation - De 1997 à 2003, l'EDM s'est appuyée sur des repères provenant du Recensement de 1996. À partir de l'année de référence 2004, l'EDM utilise des poids qui tiennent compte des nouvelles projections démographiques découlant du Recensement de 2001.

Révisions et désaisonnalisation - Toutes les estimations rétroactives à 1997 ont été rajustées en vue d'utiliser les nouvelles estimations démographiques annuelles fondées sur les données du Recensement de 2001. La méthode de calcul des poids d'enquête a été modifiée et uniformisée pour toutes les années à partir de 1997.

6.2.8 Période de référence de changement -2000

Sources des données - Pour les années 1997, 1998, 1999 et par la suite, a tous les deux ans (commençant avec 2001) les données sont recueillies pour les territoires; pour les autres années, les données sont recueillies seulement des 10 provinces.

6.2.9 Période de référence de changement -1999

Menée depuis 1997, l'Enquête sur les dépenses des ménages tente de reprendre une bonne part du contenu de l'Enquête sur les dépenses des familles (numéro d'enregistrement 3504) ainsi que du contenu de l'Enquête sur l'équipement ménager (numéro d'enregistrement 3505). La première année de référence pour laquelle nous avons des metadonnées c'est 1999 ».

6.3 Annexe - EDM et territoires

Année	Nombre de ménages des Territoires dans l'échantillon par année	Nombre de ménages des Territoires estimé par année
1997	884	23124
1998	743	25360
1999	832	26586
2001	799	29673
2003	850	31537
2005	760	31311
2007	698	33935
2009	657	33276

Les Territoires font références aux trois territoires fédéraux canadiens qui sont :

1. Les Territoires du Nord-Ouest
2. Le Nunavut
3. Yukon

6.4 Annexes - Chiffres clés

6.4.1 Le nombre de ménage estimé par année au québec :

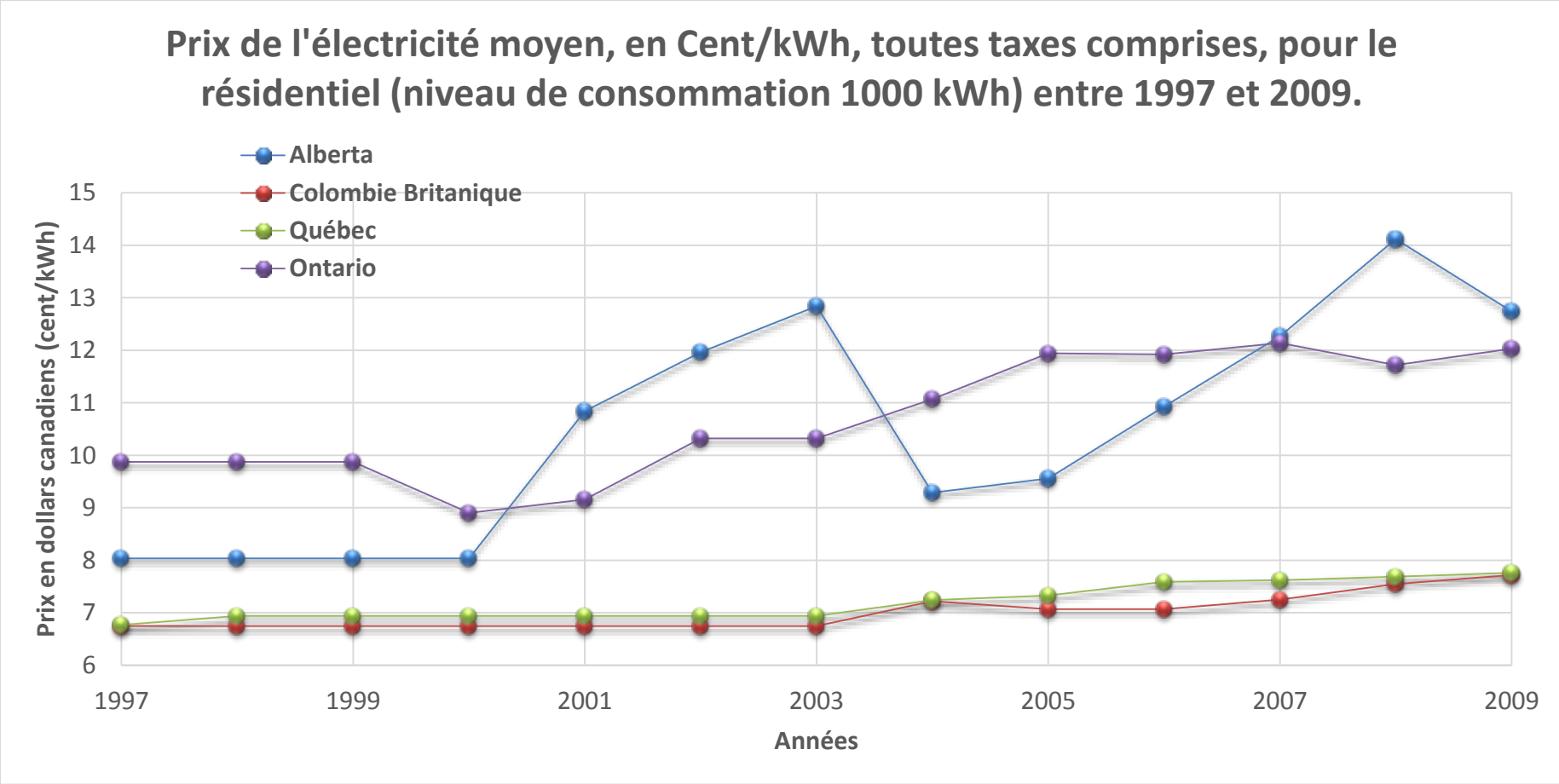
Année	Estimation
1997	2881751
1998	2904228
1999	2972520
2000	3005688
2001	3022909
2002	3063937
2003	3111087
2004	3150956
2005	3187997
2006	3242957
2007	3276637
2008	3295831
2009	3321748

6.4.2 Le nombre de ménages dont le revenu est égal ou inférieure à zéro (Canada)

YEAR	Nbre_Ménage	Estimation
1997	21	18037
1998	26	23847
1999	23	22462
2000	33	33461
2001	32	25737
2002	34	27261
2003	47	51071
2004	34	41156
2005	33	40396
2006	14	17055
2007	25	27920
2008	12	32500
2009	6	6249

Il est possible de faire l'hypothèse que ces ménages sont des travailleurs autonmes ou bien qu'ils possèdent une entreprise individuelle qui enregistre des pertes ou bien aucun revenu pour l'année en question.

6.4.3 Évolution du prix moyen du Cent/kWh pour l'Ontario, l'Alberta, la Colombie-Britanique et le Québec (1997 – 2009)p



6.4.4 Top 1997, 2009 et Top variation de la dépense en électricité

	1997	2009	% d'augmenation entre 1997 et 2009	Top 1997	Top 2009	Top % 1997/2009
Alberta	**602**	**1121**	**86%**	**10**	**7**	**1**
Prince Edward Island	822	1347	**64%**	7	3	2
New Brunswick	1377	2010	**46%**	1	1	3
Newfoundland and Labrador	1344	1935	44%	2	2	4
Saskatchewan	851	1173	38%	5	6	5
Ontario	803	1096	36%	8	8	6
Nova Scotia	1030	1336	30%	4	4	7
Manitoba	837	1079	29%	6	9	8
Quebec	**1080**	**1295**	**20%**	**3**	**5**	**9**
British Columbia	**802**	**856**	**7%**	**9**	**10**	**10**

6.4.5 Évolution du prix de l'électricité (cent/kWh) dans les provinces canadiennes

	PROVINCES									
Années	**AB**	**BC**	**MB**	**NB**	**NL**	**NS**	**ON**	**PE**	**QC**	**SK**
1997	8.038	6.747	6.867	9.491	9.409	10.81	9.876	11.198	6.766	9.626
1998	8.038	6.747	6.867	9.847	9.523	10.81	9.876	10.075	6.939	9.626
1999	8.038	6.747	6.867	10.204	9.462	10.8	9.876	10.439	6.939	9.626
2000	8.038	6.747	6.867	10.51	9.624	10.8	8.902	10.761	6.939	9.648
2001	10.842	6.747	6.867	10.51	9.516	10.8	9.159	11.241	6.939	10.142
2002	11.96	6.75	6.87	10.66	9.45	10.8	10.32	11.58	6.94	10.91
2003	12.84	6.75	6.87	10.83	9.78	11.2	10.32	13.1	6.94	10.91
2004	9.29	7.22	6.87	11.01	9.93	11.2	11.07	13.1	7.24	10.91
2005	9.56	7.07	7.34	11.33	10.82	11.85	11.94	13.2	7.33	11.47
2006	10.93	7.07	7.34	11.67	11.36	12.9	11.92	13	7.59	12.2
2007	12.27	7.25	7.44	12.49	11.78	12.46	12.14	14.18	7.62	12.66
2008	14.12	7.55	7.38	13.01	11.79	12.34	11.72	15.55	7.69	12.55
2009	12.74	7.72	7.95	13.18	12.45	14.55	12.03	18.15	7.76	12.55
% 1997/2009	**58%**	**14%**	**16%**	**39%**	**32%**	**35%**	**22%**	**62%**	15%	**30%**

Provinces	Top prix 1997	Top prix 1999	Top augmentation de 1997 à 2009	Augmentation (en %) du prix entre 1997 et 2009	
île-du-Prince-Édouard	**1**	**1**	**1**	**62%**	**Groupe 1**
Alberta	**7**	**4**	**2**	**58%**	
Nouveau-Brunswick	5	3	3	39%	
Terre-neuve-et-Labrador	6	6	5	35%	
Nouvelle-Écosse	2	2	4	32%	
Saskatchewan	4	5	6	30%	Groupe 2
Ontario	3	7	7	22%	
Manitoba	**8**	**8**	8	16%	
Québec	**9**	**9**	10	15%	**Groupe 3**
Colombie-Britanique	**10**	**10**	9	14%	

6.5 Annexe - Inertie d'une partition (source : (Tufféry, 2012))

« Inertie totale I : c'est la moyenne pondérée (généralement par l'inverse de l'effectif total) des carrés des distances des individus au barycentre (centre de gravité) de la population ($\bar{x}$ *est la moyenne des* x_i).

$$\sum_{i\epsilon\, I} p_i \,(\, x_i - \,\bar{x}\,)^2$$

Inertie intraclasse : I_A

L'inertie d'une classe est calculée de la même façon par rapport à son barycentre :

$$\sum_{i\epsilon\, I_j} p_i \,(\, x_i - \,\bar{x}_j\,)^2$$

Si la population est segmentée en k classes, d'inerties $I_1, \dots, I_k$ l'inertie intraclasse est, par définition : la somme des inerties totales de chaque classe.

$$I_A = \sum_{j=1}^{k} I_j$$

Une classe est d'autant plus homogène que son inertie est plus faible, et la classification est d'autant meilleure que I_A est petite.

Inertie interclasse : I_R

L'inertie interclasse de la classification est définie comme étant la moyenne (pondérée par la somme des poids de chaque classe $p_i = \sum_{i\epsilon\, I_j} p_i$) des carrés des distances des distances des barycentres de chaque classe au barycentre global. Elle peut s'écrire :

$$\sum_{i\epsilon\ classes} \left(\sum_{i\epsilon\, I_j} p_i \right) (\bar{x}_j - \,\bar{x})^2$$

Plus I_R est grande, et plus les classes sont séparées les uns des autres, ce qui indique une bonne classification.

Nous avons deux critères de bonnes classifications : grande I_R, petite I_A **»**

6.6 Annexes - Québec, partition à 8 groupes

6.6.1 Premier regroupement (Québec, 2009) : HHINCTOT, G024, G022, EN_RE (avec standardisation)

Cluster=1

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	277	38448	41286	0	450000
G024	Electricity for principal residence	277	168	221	0	780
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	277	14	74	0	600
En_Re		277	0	1	0	3

Cluster=2

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	134	24053	9437	7200	42000
G024	Electricity for principal residence	134	1510	550	540	2940
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	134	71	188	0	900
En_Re		134	7	1	4	10

Cluster=3

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	379	53621	31082	4600	200000
G024	Electricity for principal residence	379	1042	309	228	1650
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	379	45	121	0	600
En_Re		379	2	1	1	5

Cluster=4

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	233	89783	38781	40000	330000
G024	Electricity for principal residence	233	2147	324	1570	3060
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	233	76	158	0	700
En_Re		233	3	1	1	6

Cluster=5

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	36	18141	6077	5600	31000
G024	Electricity for principal residence	36	2328	607	696	3420
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	36	112	270	0	1000
En_Re		36	14	4	10	27

Cluster=6

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	54	110803	56622	40000	290000
G024	Electricity for principal residence	54	3643	830	2800	7200
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	54	116	247	0	1000
En_Re		54	4	2	2	11

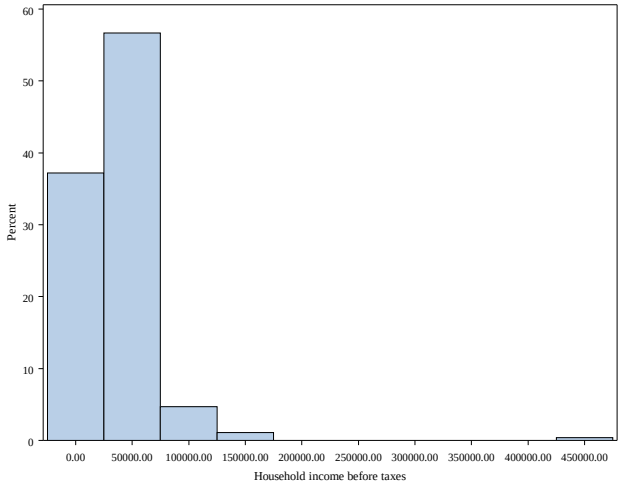
Cluster=7

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	104	79530	41429	13000	280000
G024	Electricity for principal residence	104	1267	481	372	2760
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	104	1059	333	550	1800
En_Re		104	2	1	0	7

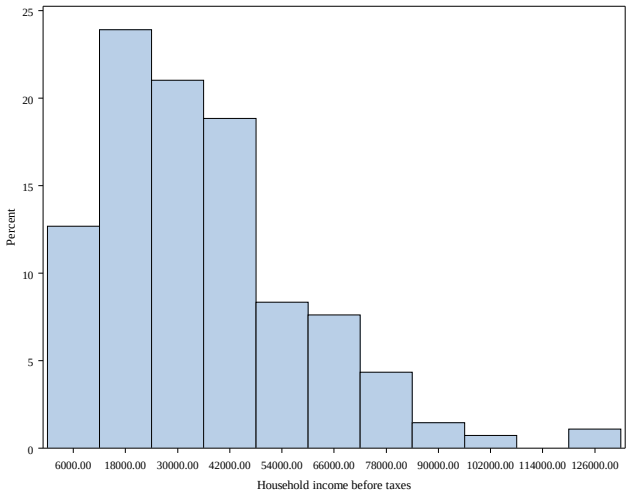
Cluster=8

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	43	88981	57247	12000	230000
G024	Electricity for principal residence	43	1043	586	360	3000
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	43	2676	810	1860	4996
En_Re		43	2	2	0	10

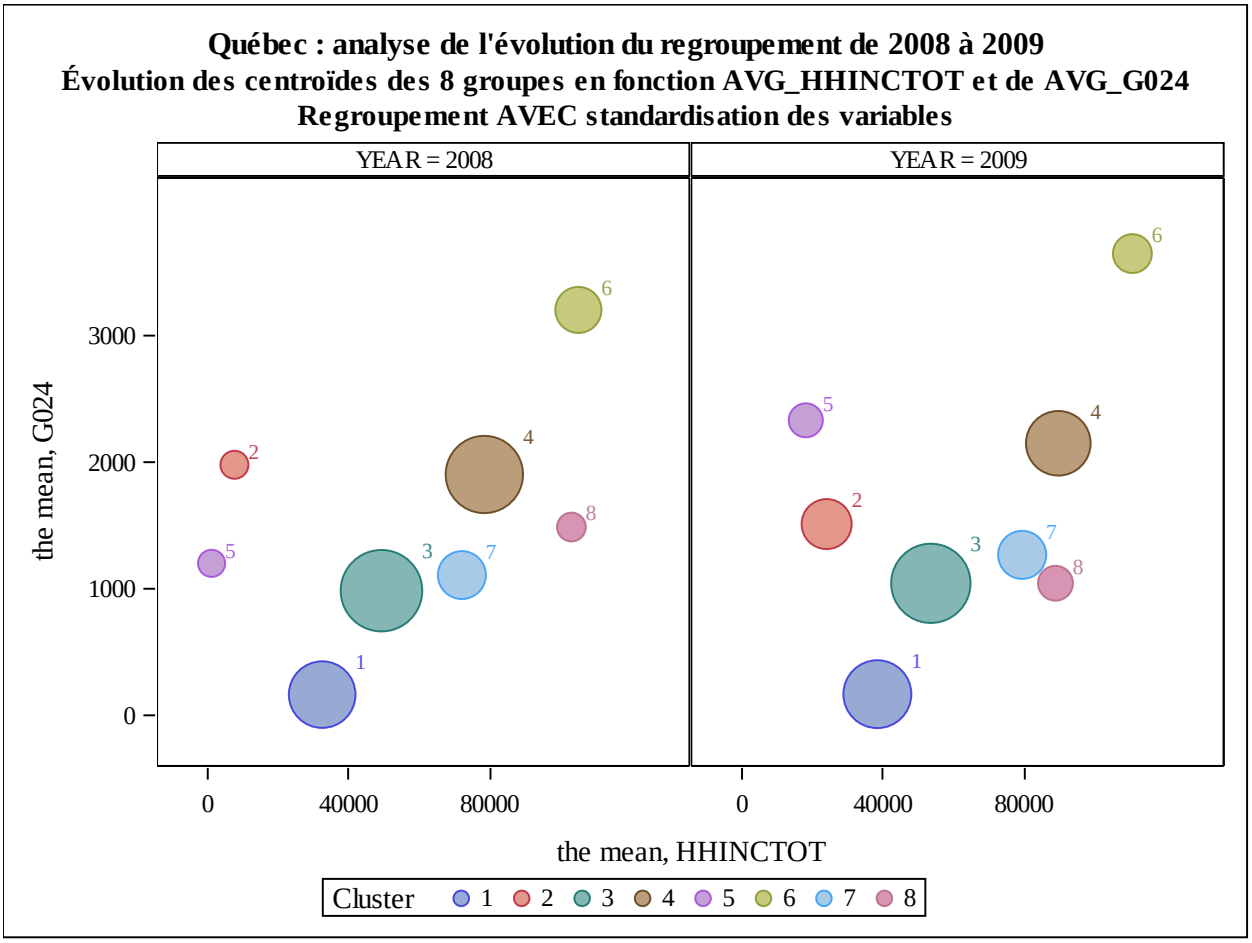
6.6.2 Cas des valeurs extrêmes, Québec 2009

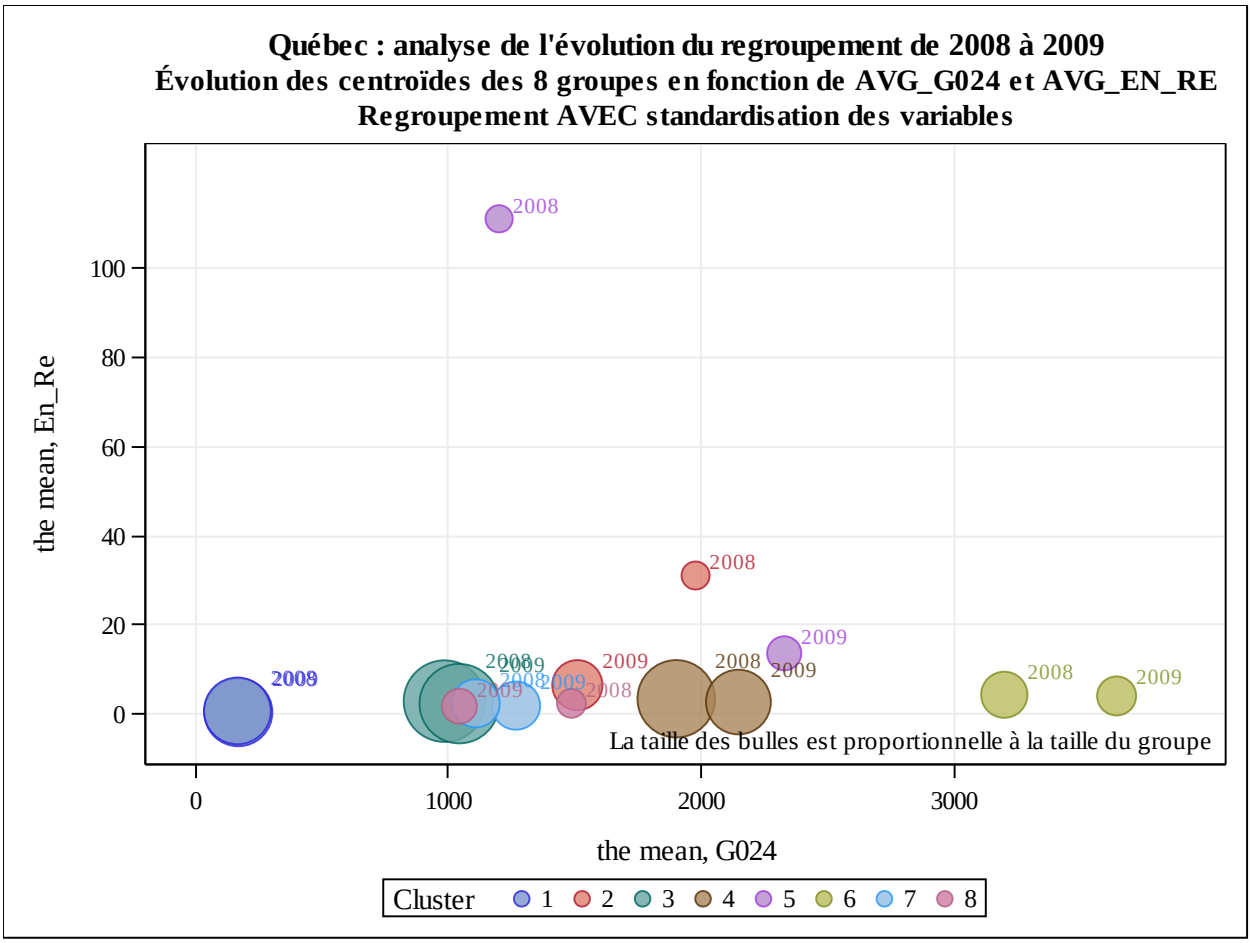


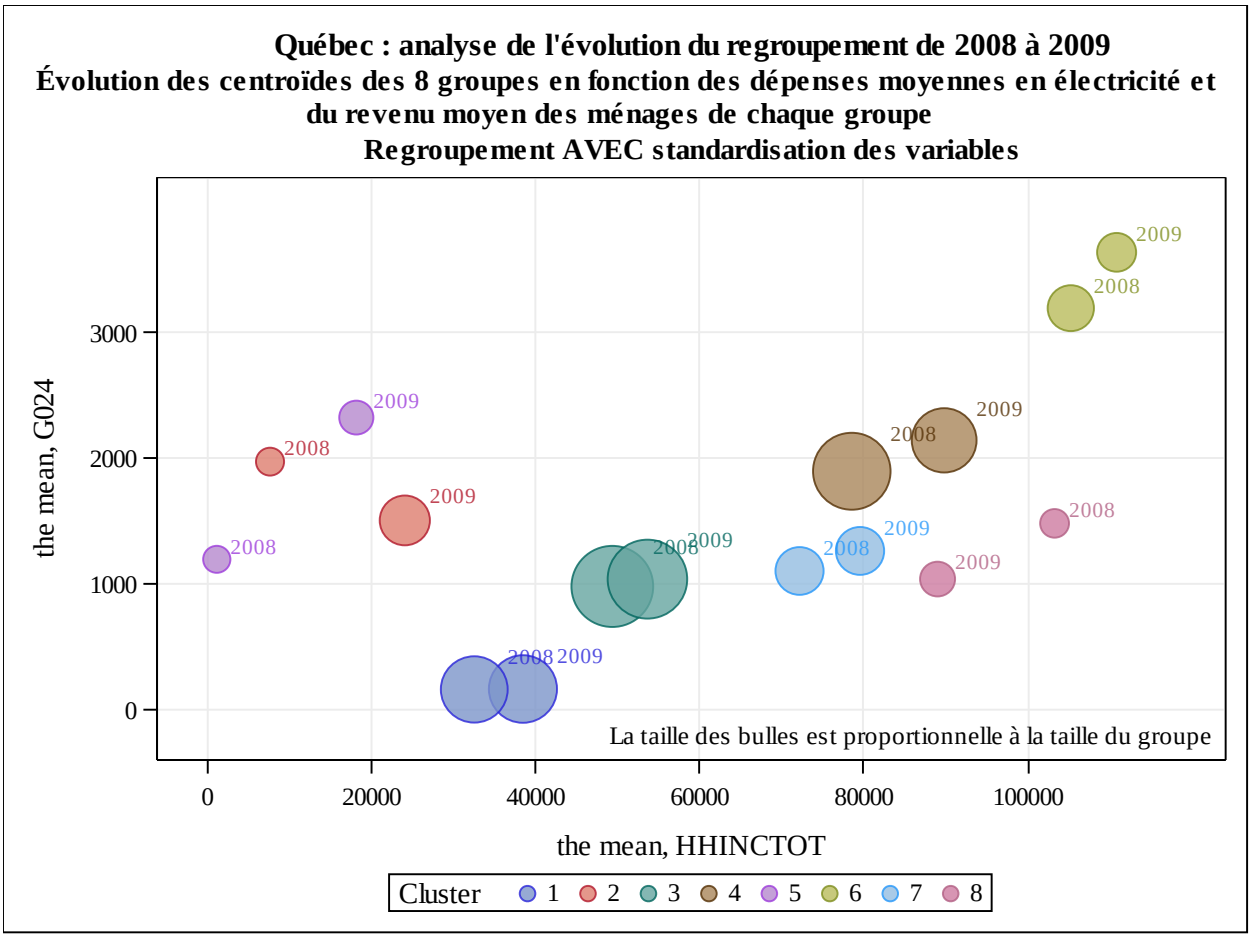
Notons une valeur extrême pour la variable revenu qui pourrait contrarier l'interprétation du groupe 1.



6.6.3 Représentation graphique du premier regroupement




Québec : analyse de l'évolution du regroupement de 2008 à 2009
Évolution des centroïdes des 8 groupes en fonction de AVG_G024 et AVG_EN_RE
Regroupement AVEC standardisation des variables
the mean, En_Re
100
80
60
40
20
0
2008
2008
2009
2008
2009
2008
2009
2008
2009
2008
2009
2008
2009
2008
2009
La taille des bulles est proportionnelle à la taille du groupe
0
1000
2000
3000
the mean, G024
Cluster 1 2 3 4 5 6 7 8


Québec : analyse de l'évolution du regroupement de 2008 à 2009
Évolution des centroïdes des 8 groupes en fonction des dépenses moyennes en électricité et du revenu moyen des ménages de chaque groupe
Regroupement AVEC standardisation des variables
the mean, G024
3000
2000
1000
0
2009
2008
2009
2008
2009
2008
2009
2008
2009
2008
2009
2008
2009
2008
La taille des bulles est proportionnelle à la taille du groupe
0
20000
40000
60000
80000
100000
the mean, HHINCTOT
Cluster 1 2 3 4 5 6 7 8

6.6.4 Deuxième regroupement (Québec, 2008) : HHINCTOT, G024, G022, EN_RE (sans standardisation)

Cluster=1

Variable	Label	N	Mean	50th Pctl	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	201	30686	25000	24354	0	200000
G024	Electricity for principal residence	201	82	0	139	0	420
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	201	13	0	74	0	700
En_Re		201	0	0	1	0	14

Cluster=2

Variable	Label	N	Mean	50th Pctl	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	279	44566	36000	30190	0	170000
G024	Electricity for principal residence	279	759	800	177	420	1070
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	279	67	0	176	0	872
En_Re		279	3	2	2	0	11

Cluster=3

Variable	Label	N	Mean	50th Pctl	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	254	64032	56000	53916	800	570000
G024	Electricity for principal residence	254	1358	1344	163	1000	1660
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	254	145	0	252	0	800
En_Re		254	4	2	10	0	150

Cluster=4

Variable	Label	N	Mean	50th Pctl	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	209	74595	64000	43661	8500	340000
G024	Electricity for principal residence	209	1931	1980	169	1644	2270
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	209	122	0	229	0	1000
En_Re		209	4	3	3	1	21

Cluster=5

Variable	Label	N	Mean	50th Pctl	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	126	93765	91000	50317	0	300000
G024	Electricity for principal residence	126	2624	2580	229	2280	3200
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	126	146	0	335	0	1700
En_Re		126	5	3	6	0	53

Cluster=6

Variable	Label	N	Mean	50th Pctl	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	40	109499	100000	51763	12000	230000
G024	Electricity for principal residence	40	3843	3636	528	3305	7000
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	40	192	0	320	0	1164
En_Re		40	5	4	4	2	30

Cluster=7

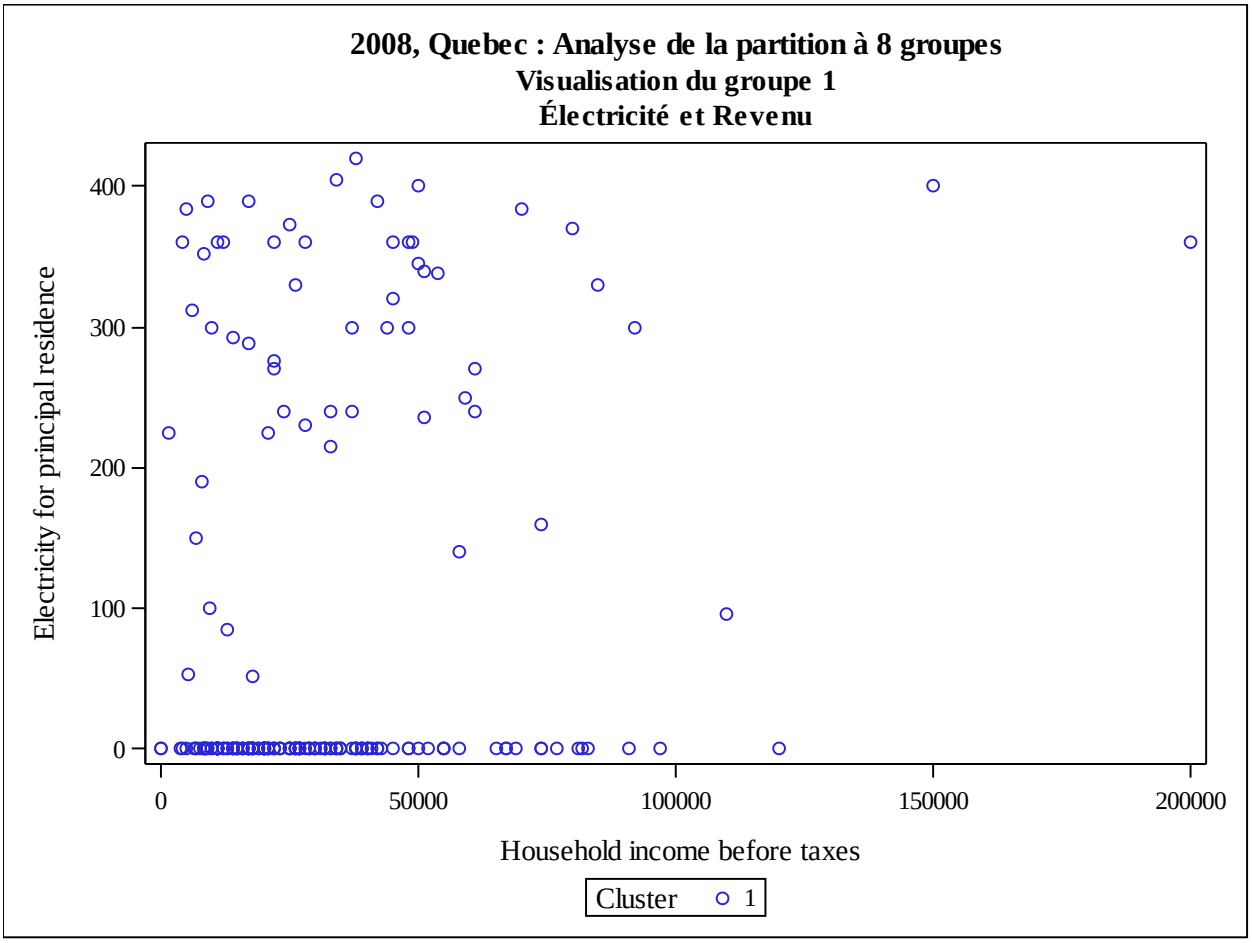
Variable	Label	N	Mean	50th Pctl	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	105	70132	60000	50073	12000	340000
G024	Electricity for principal residence	105	1078	1000	494	360	2400
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	105	1603	1500	530	875	3000
En_Re		105	2	2	2	0	10

Cluster=8

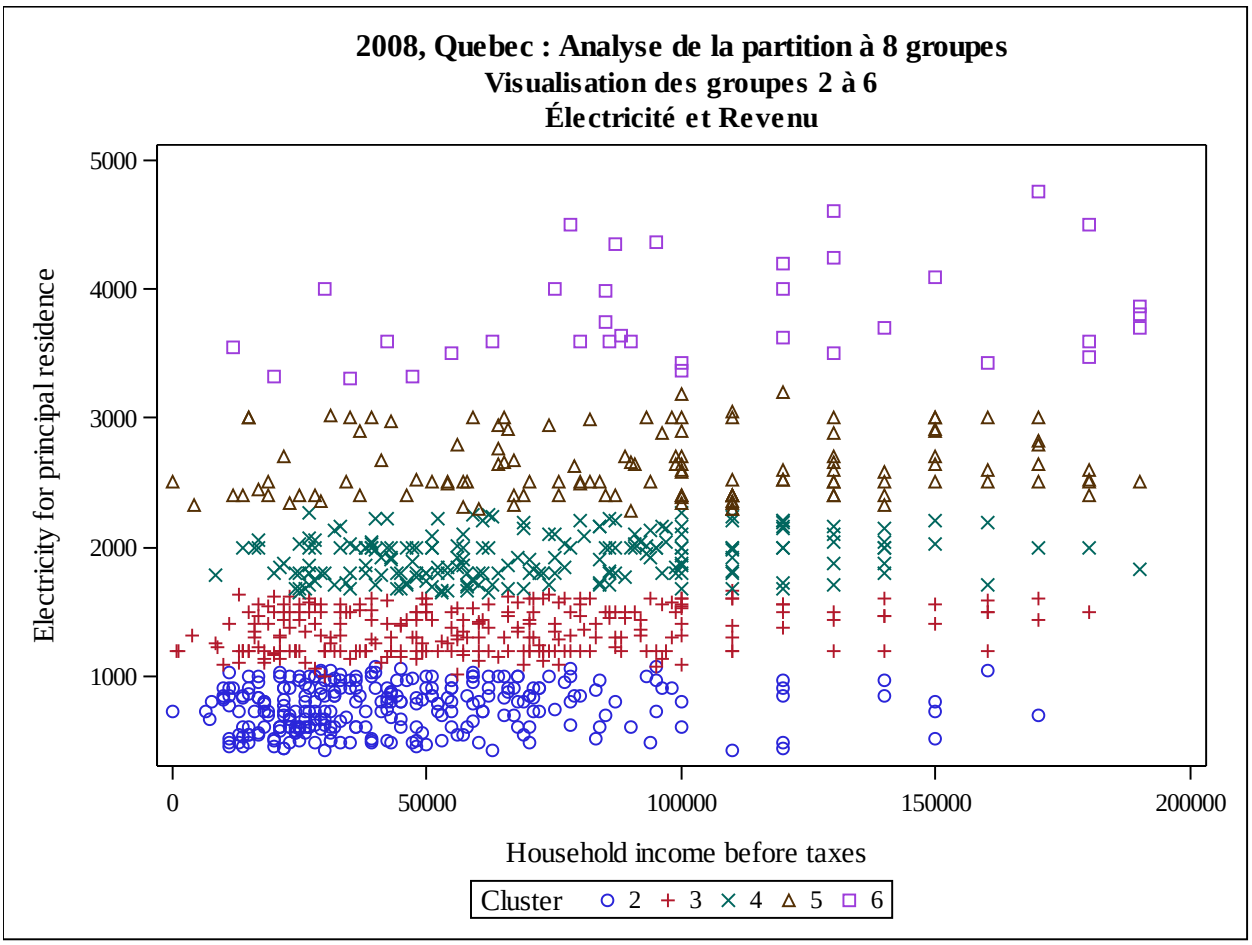
Variable	Label	N	Mean	50th Pctl	Std Dev	Minimum	Maximum
HHINCTOT	Household income before taxes	15	111784	120000	60664	19000	220000
G024	Electricity for principal residence	15	1549	1440	653	777	3000
G022	Liquid fuels for principal residence (e.g., oil, gas, etc.)	15	4552	4200	1381	3000	7011
En_Re		15	2	1	2	0	8

6.6.5 Représentation graphique du deuxième regroupement

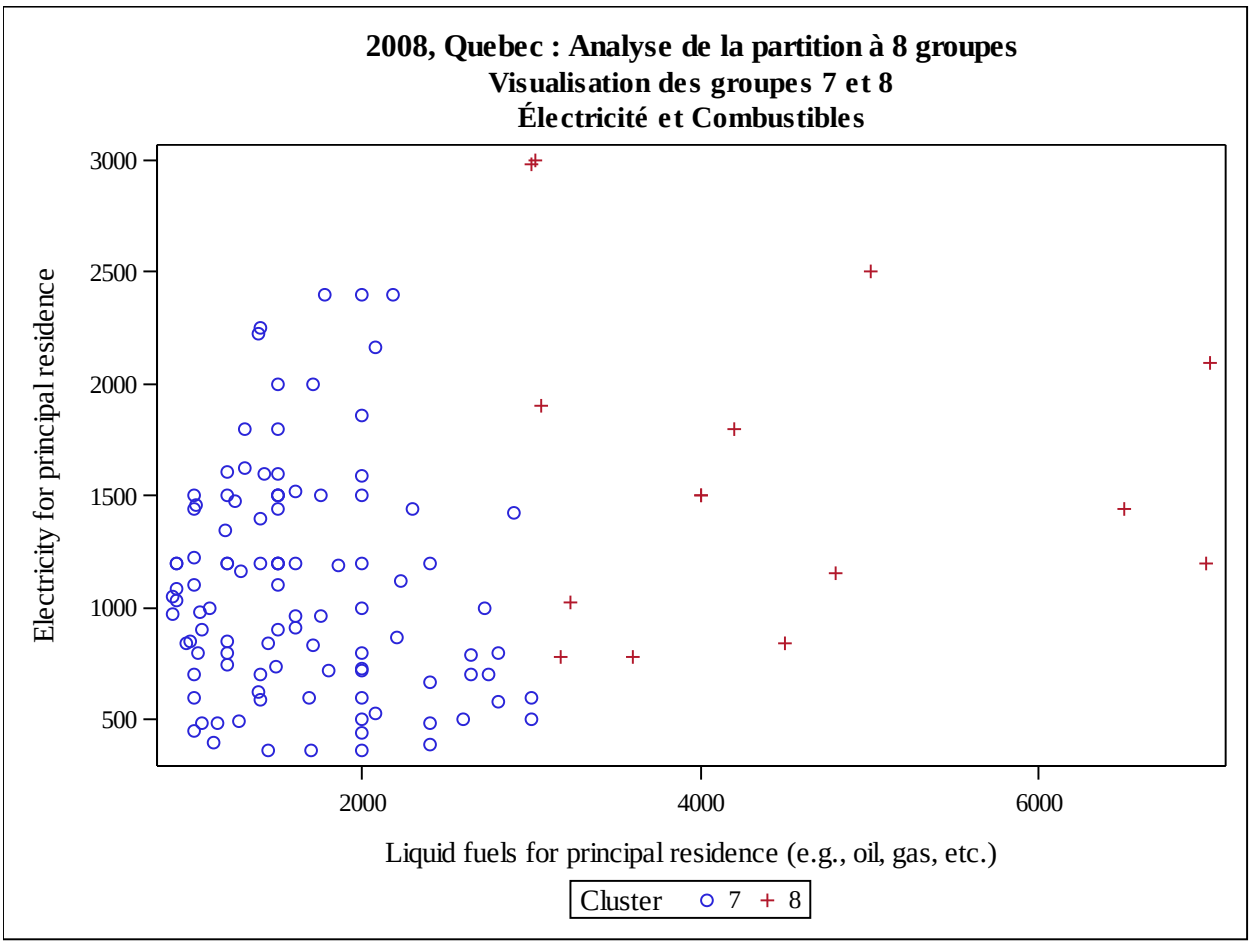
6.6.5.1 La catégorie 1 : les ménages qui n'ont pas de dépense, ni électricité, ni en gaz.



6.6.5.2 La catégorie 2 : les ménages qui ont des dépenses en électricité.



6.6.5.3 La catégorie 3 : consommateur de gaz (et d'électricité)



6.7 Annexes - Analyse du regroupement

6.7.1 Légende pour la boîte à moustache.

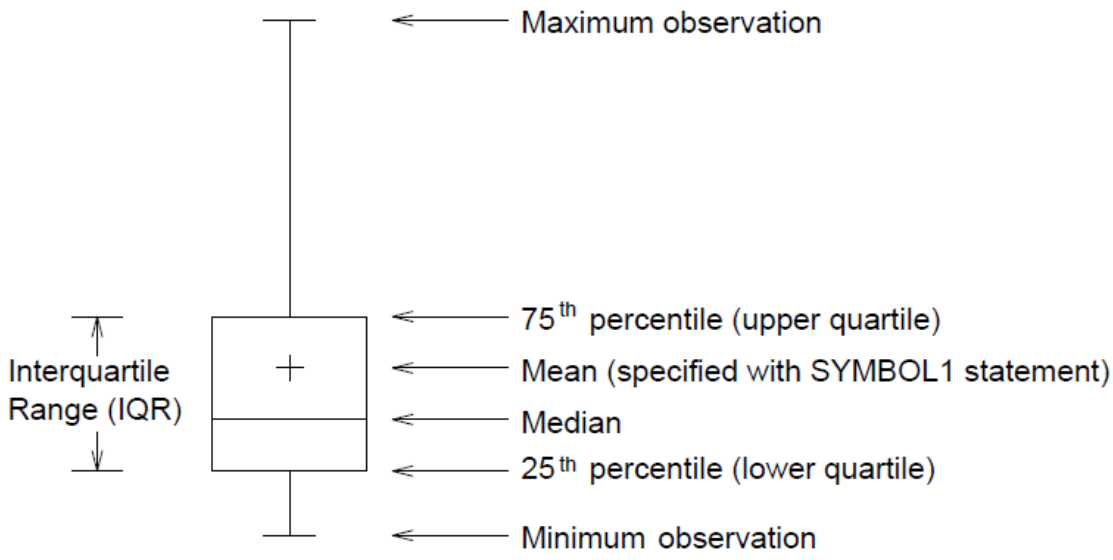
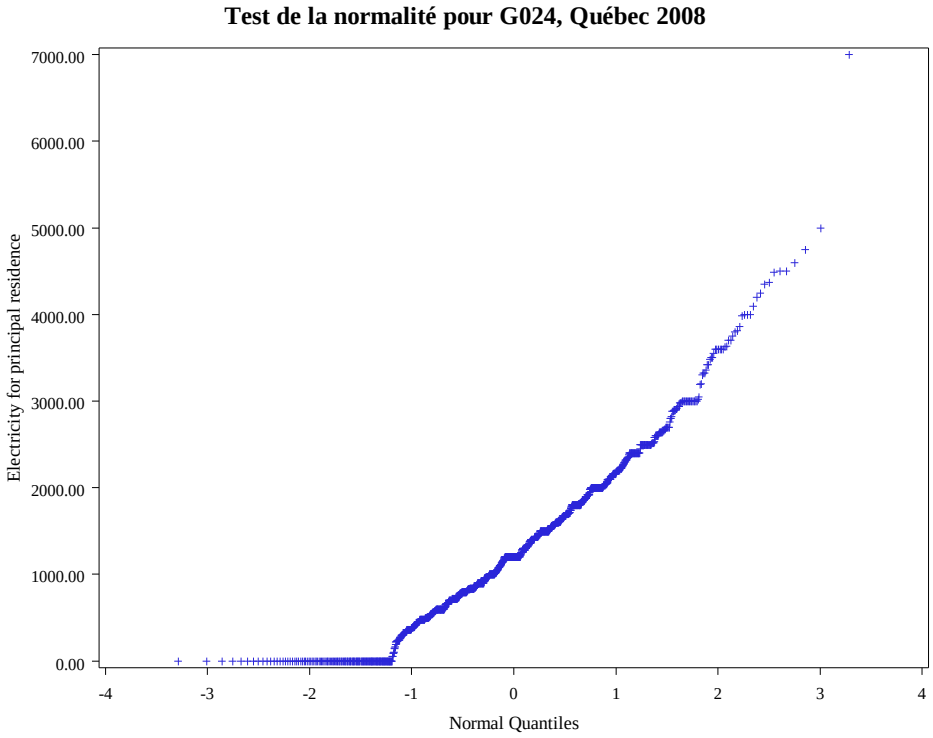


Figure 6-1 : Légende de la boîte à moustache.

Source : ("SAS/STAT(R) 9.2 User's Guide, Second Edition," n.d.).

6.7.2 Hypothèses de normalité pour G024, Québec 2008



6.7.3 Tests statistique formel pour la constance de la variance, variable G024, Québec 2008

Level of CLUSTER	N	G024	
		Mean	Std Dev
1	201	81.96020	139.960703
2	279	756.25090	177.345215
3	254	1356.09843	162.975259
4	209	1927.04306	169.579834
5	126	2625.31746	241.677294
6	40	3943.82500	658.778625
7	105	1092.30476	502.689921
8	15	1632.13333	741.794048

Levene's Test for Homogeneity of G024 Variance ANOVA of Squared Deviations from Group Means					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
CLUSTER	7	1.314E13	1.876E12	22.29	<.0001
Error	1221	1.028E14	8.417E10		

6.7.4 Test de Welch pour l'égalité de la moyenne, G024, Québec 2008 (ANOVA)

Levene's Test for Homogeneity of G024 Variance ANOVA of Squared Deviations from Group Means					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
CLUSTER	7	1.314E13	1.876E12	22.29	<.0001
Error	1221	1.028E14	8.417E10		

Welch's ANOVA for G024			
Source	**DF**	**F Value**	**Pr > F**
CLUSTER	7.0000	3209.71	<.0001
Error	169.9		

6.7.5 Test de Fisher classique, qui ne prend pas en compte la constance de la variance, G024, Québec 2008.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	970745835	138677976	2026.07	<.0001
Error	1221	83573400	68447		
Corrected Total	1228	1054319235			

R-Square	Coeff Var	Root MSE	G024 Mean
0.920732	20.06601	261.6232	1303.813

Source	DF	Type I SS	Mean Square	F Value	Pr > F
CLUSTER	7	970745834.6	138677976.4	2026.07	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
CLUSTER	7	970745834.6	138677976.4	2026.07	<.0001

6.7.6 Liste des variables utilisées pour la MANOVA

Nom de la variable	Description	Type	Catégorie
NUMRMP	Number of rooms	Ordinal	Dwelling
NUMBEDRP	Number of bedrooms	Ordinal	Dwelling
NUMBTHRP	Number of bathrooms	Ordinal	Dwelling
ADRENALT	Additions, renovations and alterations	Numerical	Dwelling
NEWINST	New installations equipment and fixtures	Numerical	Dwelling
ECFAM		Ordinal	Household description
HHSZTOTP	Household size (number of persons a member sometime in 1997)	Ordinal	Household description
HHINCTOT	Household income before taxes	Numerical	Household description
NMFRIDGP	Number of refrigerators	Ordinal	Household equipment
NMVEHONP	Number of vehicles owned at December 31	Ordinal	Household equipment
F001	Total food	Numerical	Food
F002	Food purchased from stores	Numerical	Food
F008	Food purchased from restaurants	Numerical	Food
G001	Total shelter	Numerical	Shelter
G002	Principal accommodation	Numerical	Shelter
G003	Rented living quarters	Numerical	Shelter
G004	Rent (amount paid)	Numerical	Shelter
G007		Numerical	Shelter
G040	Regular mortgage payments	Numerical	Shelter
P140	Maintenance, repairs and replacements	Numerical	Shelter
G009		Numerical	Shelter
G010	Property taxes on principal residence	Numerical	Shelter
G011		Numerical	Shelter
G022		Numerical	Shelter
ENERGYTOT	somme de G024 + G022 + K020 ;	Numerical	Shelter
G024	Electricity for principal residence	Numerical	Shelter
G025	Other accomodation	Numerical	Shelter
G026	Owned vacation home	Numerical	Shelter
H001	Total household operation	Numerical	Household operation
I001	Total household furnishings and equipment	Numerical	Household furnishings and equipment
J001	Total clothing	Numerical	Clothing
K001	Total transportation	Numerical	Transportation
K002	Private transportation	Numerical	Transportation
K003	Purchase of automobiles and trucks	Numerical	Transportation
K020	Gasoline and other fuels	Numerical	Transportation

Nom de la variable	Description	Type	Catégorie
K028	Numerical		Transportation
K031	Public transportation	Numerical	Transportation
K032TOT	Local and commuter transportation	Numerical	Transportation
L101	Total health care	Numerical	Health Care
M101	Total recreation	Numerical	Recreation
M103	Sports and athletic equipment	Numerical	Recreation
M127	Purchase of recreation vehicles	Numerical	Recreation
M128	Bicycles, parts and accessories	Numerical	Recreation
M129	Numerical		Recreation
M139	Operation of recreational vehicles	Numerical	Recreation
M140	Bicycle maintenance and repairs	Numerical	Recreation
M142TOT	Other recreation vehicle operation	Numerical	Recreation
M201		Numerical	Reading materials and other printed matter
M202	Newspapers	Numerical	Reading materials and other printed matter
M203	Magazines and periodicals	Numerical	Reading materials and other printed matter
M204		Numerical	Reading materials and other printed matter
M301	Total education	Numerical	Education
M302TOT	Education supplies and textbooks	Numerical	Education
M308	Tuition fees	Numerical	Education
N101		Numerical	Tabacco products and alcoholic beverages
N201	Games of chance expenses (net)	Numerical	Other expenses
O403TOT	Numerical		Other expenses
O401	Gifts of money and contributions	Numerical	Other expenses
TOTCUCON	Total current consumption	Numerical	Expenditures
TOTEXPEN	Total expenditures	Numerical	Expenditures

6.7.7 Analyse plus fine pour G024 : Proc Glimmix / Tukey-Kramer

Differences of CLUSTER Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer												
Cluster	**Cluster**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adj P**	**Alpha**	**Lower**	**Upper**	**Adj Lower**	**Adj Upper**
1	2	-674.29	14.4978	474	-46.51	<.0001	<.0001	0.05	-702.78	-645.80	-718.43	-630.15
1	3	-1274.14	14.2137	449.9	-89.64	<.0001	<.0001	0.05	-1302.07	-1246.20	-1317.42	-1230.85
1	4	-1845.08	15.3314	398.9	-120.35	<.0001	<.0001	0.05	-1875.22	-1814.94	-1891.80	-1798.37
1	5	-2543.36	23.6857	178.2	-107.38	<.0001	<.0001	0.05	-2590.10	-2496.62	-2616.01	-2470.71
1	6	-3861.86	104.63	1221	-36.91	<.0001	<.0001	0.05	-4067.14	-3656.59	-4179.54	-3544.19
1	7	-1010.34	50.0410	1221	-20.19	<.0001	<.0001	0.05	-1108.52	-912.17	-1162.28	-858.41
1	8	-1550.17	191.78	1221	-8.08	<.0001	<.0001	0.05	-1926.44	-1173.91	-2132.46	-967.88
2	3	-599.85	14.7411	531	-40.69	<.0001	<.0001	0.05	-628.81	-570.89	-644.71	-554.99
2	4	-1170.79	15.8216	458.3	-74.00	<.0001	<.0001	0.05	-1201.88	-1139.70	-1218.97	-1122.62
2	5	-1869.07	24.0059	188.2	-77.86	<.0001	<.0001	0.05	-1916.42	-1821.71	-1942.65	-1795.48
2	6	-3187.57	104.70	1221	-30.44	<.0001	<.0001	0.05	-3392.99	-2982.16	-3505.47	-2869.68
2	7	-336.05	50.1933	1221	-6.70	<.0001	<.0001	0.05	-434.53	-237.58	-488.45	-183.66
2	8	-875.88	191.82	1221	-4.57	<.0001	0.0001	0.05	-1252.22	-499.54	-1458.29	-293.47
3	4	-570.94	15.5617	436.9	-36.69	<.0001	<.0001	0.05	-601.53	-540.36	-618.34	-523.55
3	5	-1269.22	23.8354	183.2	-53.25	<.0001	<.0001	0.05	-1316.25	-1222.19	-1342.31	-1196.13
3	6	-2587.73	104.66	1221	-24.72	<.0001	<.0001	0.05	-2793.07	-2382.39	-2905.50	-2269.95
3	7	263.79	50.1120	1221	5.26	<.0001	<.0001	0.05	165.48	362.11	111.65	415.94
3	8	-276.03	191.80	1221	-1.44	0.1504	0.8389	0.05	-652.34	100.27	-858.38	306.31
4	5	-698.27	24.5184	199.6	-28.48	<.0001	<.0001	0.05	-746.62	-649.93	-773.38	-623.17
4	6	-2016.78	104.82	1221	-19.24	<.0001	<.0001	0.05	-2222.43	-1811.13	-2335.03	-1698.53
4	7	834.74	50.4404	1221	16.55	<.0001	<.0001	0.05	735.78	933.70	681.59	987.88
4	8	294.91	191.89	1221	1.54	0.1246	0.7873	0.05	-81.5595	671.38	-287.70	877.52
5	6	-1318.51	106.36	1221	-12.40	<.0001	<.0001	0.05	-1527.18	-1109.83	-1641.45	-995.57
5	7	1533.01	53.5742	1221	28.61	<.0001	<.0001	0.05	1427.91	1638.12	1370.35	1695.67
5	8	993.18	192.74	1221	5.15	<.0001	<.0001	0.05	615.05	1371.32	408.00	1578.37
6	7	2851.52	115.14	1221	24.77	<.0001	<.0001	0.05	2625.63	3077.41	2501.95	3201.09
6	8	2311.69	218.02	1221	10.60	<.0001	<.0001	0.05	1883.95	2739.43	1649.74	2973.64
7	8	-539.83	197.71	1221	-2.73	0.0064	0.1145	0.05	-927.72	-151.93	-1140.12	60.4619

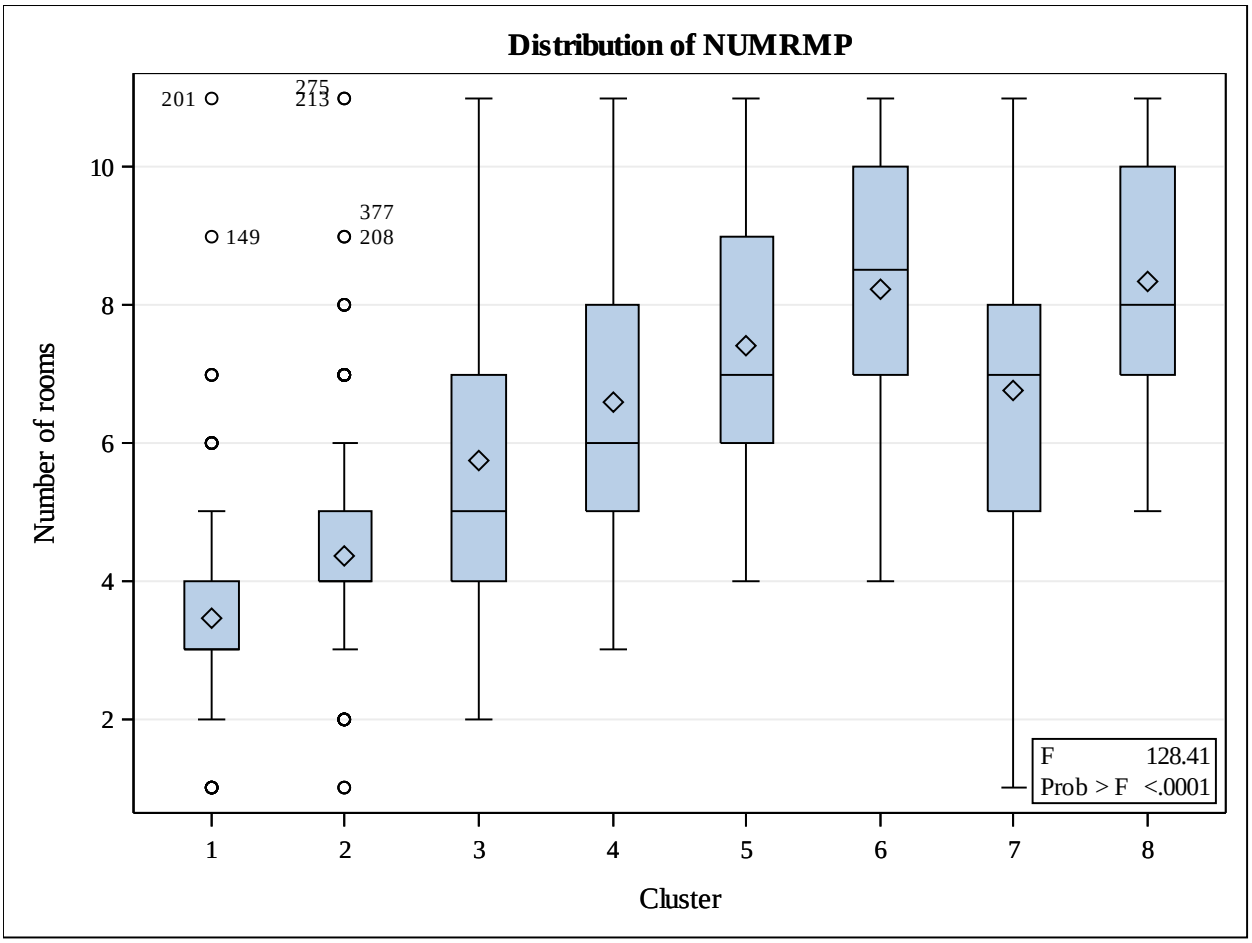
6.7.8 Analyse plus fine pour G024 : Proc Glimmix / Bonferroni

Differences of CLUSTER Least Squares Means Adjustment for Multiple Comparisons: Bonferroni												
Cluster	**Cluster**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adj P**	**Alpha**	**Lower**	**Upper**	**Adj Lower**	**Adj Upper**
1	2	-674.29	14.4978	474	-46.51	<.0001	<.0001	0.05	-702.78	-645.80	-719.84	-628.75
1	3	-1274.14	14.2137	449.9	-89.64	<.0001	<.0001	0.05	-1302.07	-1246.20	-1318.80	-1229.47
1	4	-1845.08	15.3314	398.9	-120.35	<.0001	<.0001	0.05	-1875.22	-1814.94	-1893.30	-1796.87
1	5	-2543.36	23.6857	178.2	-107.38	<.0001	<.0001	0.05	-2590.10	-2496.62	-2618.48	-2468.24
1	6	-3861.86	104.63	1221	-36.91	<.0001	<.0001	0.05	-4067.14	-3656.59	-4189.42	-3534.31
1	7	-1010.34	50.0410	1221	-20.19	<.0001	<.0001	0.05	-1108.52	-912.17	-1167.00	-853.68
1	8	-1550.17	191.78	1221	-8.08	<.0001	<.0001	0.05	-1926.44	-1173.91	-2150.58	-949.77
2	3	-599.85	14.7411	531	-40.69	<.0001	<.0001	0.05	-628.81	-570.89	-646.13	-553.57
2	4	-1170.79	15.8216	458.3	-74.00	<.0001	<.0001	0.05	-1201.88	-1139.70	-1220.51	-1121.08
2	5	-1869.07	24.0059	188.2	-77.86	<.0001	<.0001	0.05	-1916.42	-1821.71	-1945.14	-1792.99
2	6	-3187.57	104.70	1221	-30.44	<.0001	<.0001	0.05	-3392.99	-2982.16	-3515.36	-2859.79
2	7	-336.05	50.1933	1221	-6.70	<.0001	<.0001	0.05	-434.53	-237.58	-493.19	-178.92
2	8	-875.88	191.82	1221	-4.57	<.0001	0.0002	0.05	-1252.22	-499.54	-1476.41	-275.35
3	4	-570.94	15.5617	436.9	-36.69	<.0001	<.0001	0.05	-601.53	-540.36	-619.86	-522.03
3	5	-1269.22	23.8354	183.2	-53.25	<.0001	<.0001	0.05	-1316.25	-1222.19	-1344.78	-1193.66
3	6	-2587.73	104.66	1221	-24.72	<.0001	<.0001	0.05	-2793.07	-2382.39	-2915.39	-2260.07
3	7	263.79	50.1120	1221	5.26	<.0001	<.0001	0.05	165.48	362.11	106.91	420.68
3	8	-276.03	191.80	1221	-1.44	0.1504	1.0000	0.05	-652.34	100.27	-876.50	324.43
4	5	-698.27	24.5184	199.6	-28.48	<.0001	<.0001	0.05	-746.62	-649.93	-775.91	-620.64
4	6	-2016.78	104.82	1221	-19.24	<.0001	<.0001	0.05	-2222.43	-1811.13	-2344.94	-1688.63
4	7	834.74	50.4404	1221	16.55	<.0001	<.0001	0.05	735.78	933.70	676.83	992.65
4	8	294.91	191.89	1221	1.54	0.1246	1.0000	0.05	-81.5595	671.38	-305.82	895.64
5	6	-1318.51	106.36	1221	-12.40	<.0001	<.0001	0.05	-1527.18	-1109.83	-1651.49	-985.52
5	7	1533.01	53.5742	1221	28.61	<.0001	<.0001	0.05	1427.91	1638.12	1365.29	1700.73
5	8	993.18	192.74	1221	5.15	<.0001	<.0001	0.05	615.05	1371.32	389.80	1596.57
6	7	2851.52	115.14	1221	24.77	<.0001	<.0001	0.05	2625.63	3077.41	2491.07	3211.97
6	8	2311.69	218.02	1221	10.60	<.0001	<.0001	0.05	1883.95	2739.43	1629.15	2994.24
7	8	-539.83	197.71	1221	-2.73	0.0064	0.1797	0.05	-927.72	-151.93	-1158.80	79.1383

6.8 Annexes - Portrait des ménages québecois, 2008

6.8.1 L'habitation

6.8.1.1 Distribution du nombre de pièces : variable NUMRMP

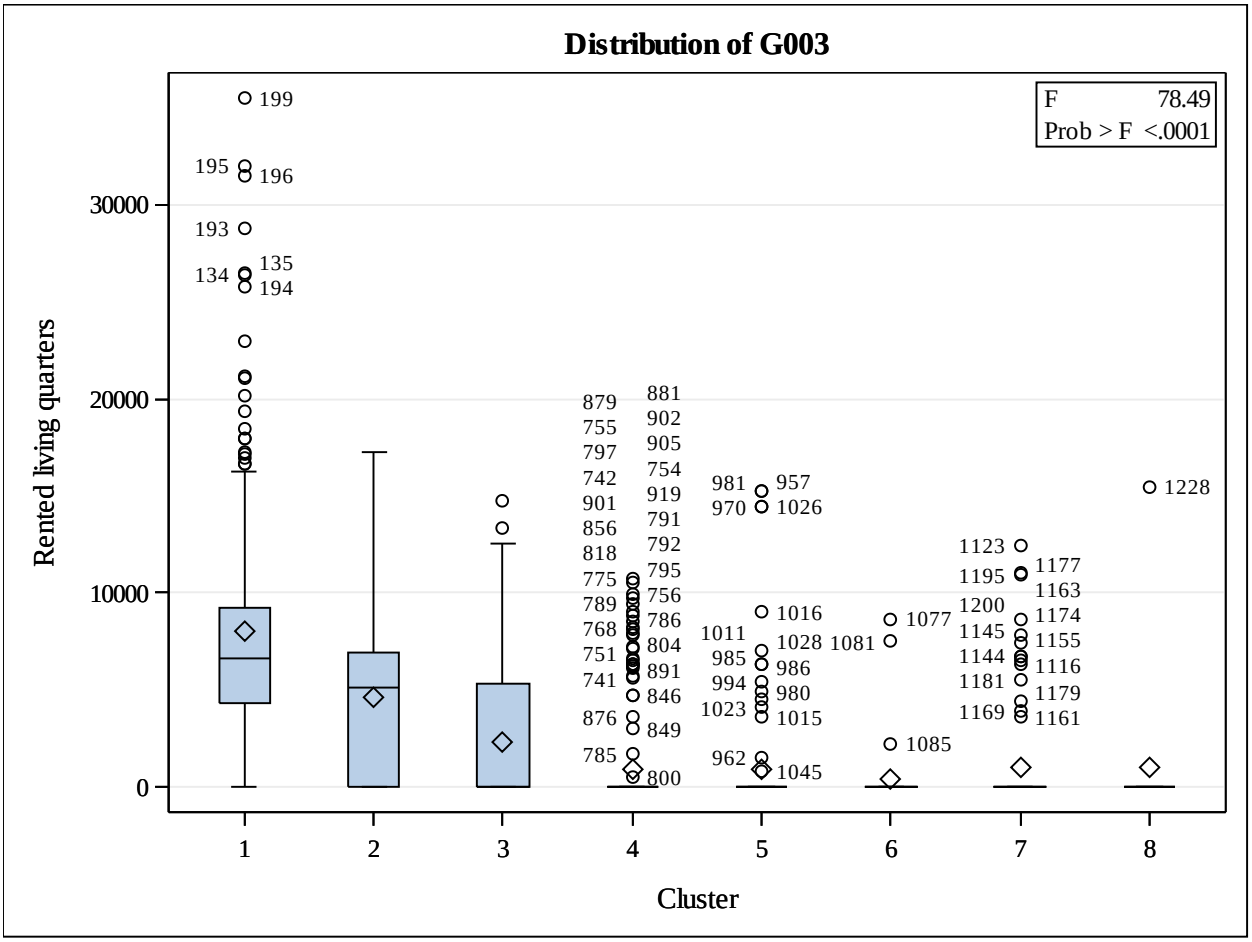


6.8.1.2 Analyse plus fine, groupes 2 à 2 : variable NUMPR

Differences of CLUSTER Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer												
Cluster	**Cluster**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adj P**	**Alpha**	**Lower**	**Upper**	**Adj Lower**	**Adj Upper**
1	2	-0.9115	0.1316	431.8	-6.92	<.0001	<.0001	0.05	-1.1702	-0.6527	-1.3124	-0.5105
1	3	-2.2864	0.1502	453	-15.22	<.0001	<.0001	0.05	-2.5816	-1.9912	-2.7438	-1.8290
1	4	-3.1404	0.1558	398.7	-20.16	<.0001	<.0001	0.05	-3.4466	-2.8341	-3.6151	-2.6657
1	5	-3.9550	0.1900	219.5	-20.81	<.0001	<.0001	0.05	-4.3295	-3.5805	-4.5366	-3.3734
1	6	-4.7673	0.3288	47.3	-14.50	<.0001	<.0001	0.05	-5.4286	-4.1059	-5.8097	-3.7249
1	7	-3.3137	0.2162	162.4	-15.33	<.0001	<.0001	0.05	-3.7406	-2.8869	-3.9775	-2.6499
1	8	-4.8756	0.4753	15.33	-10.26	<.0001	<.0001	0.05	-5.8868	-3.8644	-6.5308	-3.2204
2	3	-1.3749	0.1407	484	-9.77	<.0001	<.0001	0.05	-1.6514	-1.0984	-1.8033	-0.9465
2	4	-2.2289	0.1466	397.6	-15.20	<.0001	<.0001	0.05	-2.5172	-1.9406	-2.6758	-1.7821
2	5	-3.0435	0.1826	197.7	-16.67	<.0001	<.0001	0.05	-3.4037	-2.6834	-3.6030	-2.4840
2	6	-3.8558	0.3246	44.97	-11.88	<.0001	<.0001	0.05	-4.5096	-3.2021	-4.8872	-2.8244
2	7	-2.4023	0.2097	147.2	-11.46	<.0001	<.0001	0.05	-2.8166	-1.9879	-3.0470	-1.7575
2	8	-3.9642	0.4724	14.96	-8.39	<.0001	<.0001	0.05	-4.9713	-2.9570	-5.6149	-2.3135
3	4	-0.8540	0.1635	449.2	-5.22	<.0001	<.0001	0.05	-1.1753	-0.5326	-1.3520	-0.3560
3	5	-1.6686	0.1964	245.9	-8.49	<.0001	<.0001	0.05	-2.0555	-1.2817	-2.2691	-1.0681
3	6	-2.4809	0.3325	49.47	-7.46	<.0001	<.0001	0.05	-3.1490	-1.8128	-3.5331	-1.4287
3	7	-1.0273	0.2218	178.5	-4.63	<.0001	0.0002	0.05	-1.4650	-0.5897	-1.7076	-0.3470
3	8	-2.5892	0.4779	15.67	-5.42	<.0001	0.0012	0.05	-3.6041	-1.5744	-4.2485	-0.9300
4	5	-0.8146	0.2007	253.3	-4.06	<.0001	0.0017	0.05	-1.2099	-0.4193	-1.4281	-0.2011
4	6	-1.6269	0.3351	50.93	-4.86	<.0001	0.0003	0.05	-2.2996	-0.9542	-2.6859	-0.5680
4	7	-0.1733	0.2256	186.3	-0.77	0.4433	0.9945	0.05	-0.6184	0.2717	-0.8650	0.5183
4	8	-1.7352	0.4797	15.9	-3.62	0.0023	0.0375	0.05	-2.7526	-0.7179	-3.3974	-0.07314
5	6	-0.8123	0.3523	61.14	-2.31	0.0245	0.3075	0.05	-1.5168	-0.1078	-1.9180	0.2934
5	7	0.6413	0.2505	214.3	2.56	0.0112	0.1767	0.05	0.1475	1.1350	-0.1255	1.4080
5	8	-0.9206	0.4919	17.56	-1.87	0.0780	0.5849	0.05	-1.9559	0.1146	-2.6036	0.7623
6	7	1.4536	0.3671	69.96	3.96	0.0002	0.0042	0.05	0.7215	2.1857	0.3067	2.6005
6	8	-0.1083	0.5603	27.57	-0.19	0.8481	1.0000	0.05	-1.2569	1.0402	-1.9428	1.7262
7	8	-1.5619	0.5025	19.09	-3.11	0.0058	0.0866	0.05	-2.6134	-0.5104	-3.2647	0.1409

6.8.2 Les dépenses liées aux logements

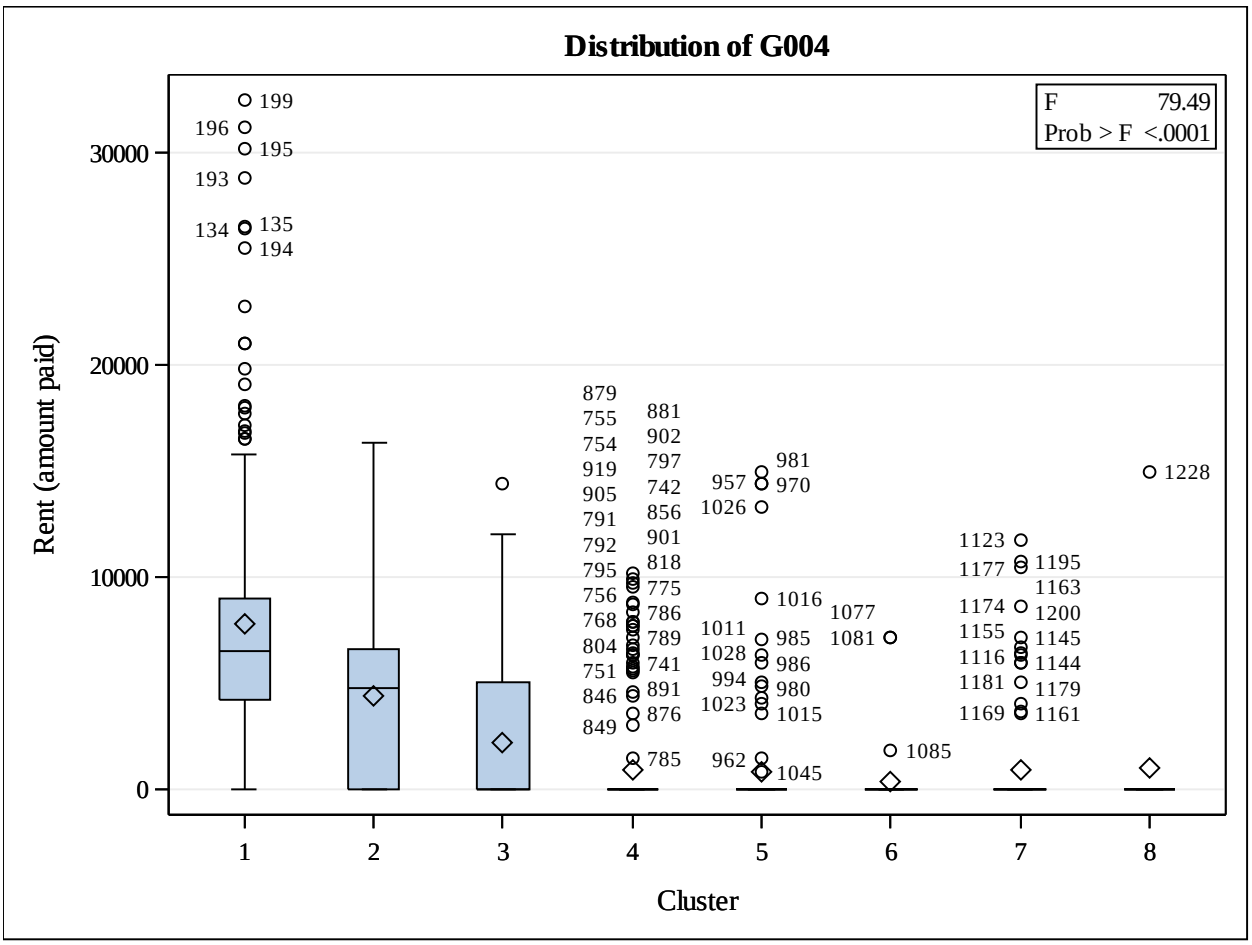
6.8.2.1 Distribution de G003, les dépenses totales pour la résidence principale louée.



6.8.2.2 Analyse plus fine, groupes 2 à 2 : variable G003

Differences of CLUSTER Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer												
Cluster	**Cluster**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adj P**	**Alpha**	**Lower**	**Upper**	**Adj Lower**	**Adj Upper**
1	2	3410.98	483.34	1	7.06	0.0896	0.2238	0.05	-2730.41	9552.37	-12105	18926
1	3	5697.25	484.21	1	11.77	0.0540	0.1357	0.05	-455.16	11850	-9846.09	21241
1	4	7071.47	464.92	1	15.21	0.0418	0.1052	0.05	1164.05	12979	-7852.93	21996
1	5	7118.01	505.23	1	14.09	0.0451	0.1135	0.05	698.41	13538	-9100.34	23336
1	6	7553.88	517.26	1	14.60	0.0435	0.1096	0.05	981.46	14126	-9050.56	24158
1	7	7042.08	503.29	1	13.99	0.0454	0.1143	0.05	647.21	13437	-9113.81	23198
1	8	6982.50	1118.29	1	6.24	0.1011	0.2518	0.05	-7226.76	21192	-28916	42881
2	3	2286.27	311.38	1	7.34	0.0862	0.2154	0.05	-1670.21	6242.75	-7709.31	12282
2	4	3660.50	280.46	1	13.05	0.0487	0.1225	0.05	96.9003	7224.10	-5342.51	12664
2	5	3707.03	343.17	1	10.80	0.0588	0.1476	0.05	-653.33	8067.40	-7308.92	14723
2	6	4142.90	360.64	1	11.49	0.0553	0.1389	0.05	-439.48	8725.29	-7433.95	15720
2	7	3631.11	340.30	1	10.67	0.0595	0.1494	0.05	-692.78	7955.00	-7292.69	14555
2	8	3571.53	1055.03	1	3.39	0.1829	0.4421	0.05	-9833.87	16977	-30296	37439
3	4	1374.23	281.95	1	4.87	0.1288	0.3183	0.05	-2208.33	4956.78	-7676.68	10425
3	5	1420.76	344.39	1	4.13	0.1514	0.3710	0.05	-2955.11	5796.64	-9634.37	12476
3	6	1856.63	361.80	1	5.13	0.1225	0.3033	0.05	-2740.51	6453.78	-9757.51	13471
3	7	1344.84	341.53	1	3.94	0.1583	0.3869	0.05	-2994.69	5684.36	-9618.47	12308
3	8	1285.26	1055.43	1	1.22	0.4377	0.8663	0.05	-12125	14696	-32595	35165
4	5	46.5366	316.71	1	0.15	0.9071	1.0000	0.05	-3977.61	4070.69	-10120	10213
4	6	482.40	335.56	1	1.44	0.3869	0.8069	0.05	-3781.31	4746.11	-10289	11254
4	7	-29.3903	313.59	1	-0.09	0.9405	1.0000	0.05	-4013.98	3955.20	-10096	10037
4	8	-88.9713	1046.72	1	-0.09	0.9460	1.0000	0.05	-13389	13211	-33689	33512
5	6	435.87	389.50	1	1.12	0.4643	0.8922	0.05	-4513.15	5384.89	-12067	12939
5	7	-75.9270	370.74	1	-0.20	0.8714	1.0000	0.05	-4786.61	4634.76	-11977	11825
5	8	-135.51	1065.24	1	-0.13	0.9194	1.0000	0.05	-13671	13400	-34330	34059
6	7	-511.79	386.97	1	-1.32	0.4121	0.8381	0.05	-5428.71	4405.12	-12934	11910
6	8	-571.37	1070.99	1	-0.53	0.6880	0.9934	0.05	-14180	13037	-34951	33808
7	8	-59.5810	1064.32	1	-0.06	0.9644	1.0000	0.05	-13583	13464	-34225	34106

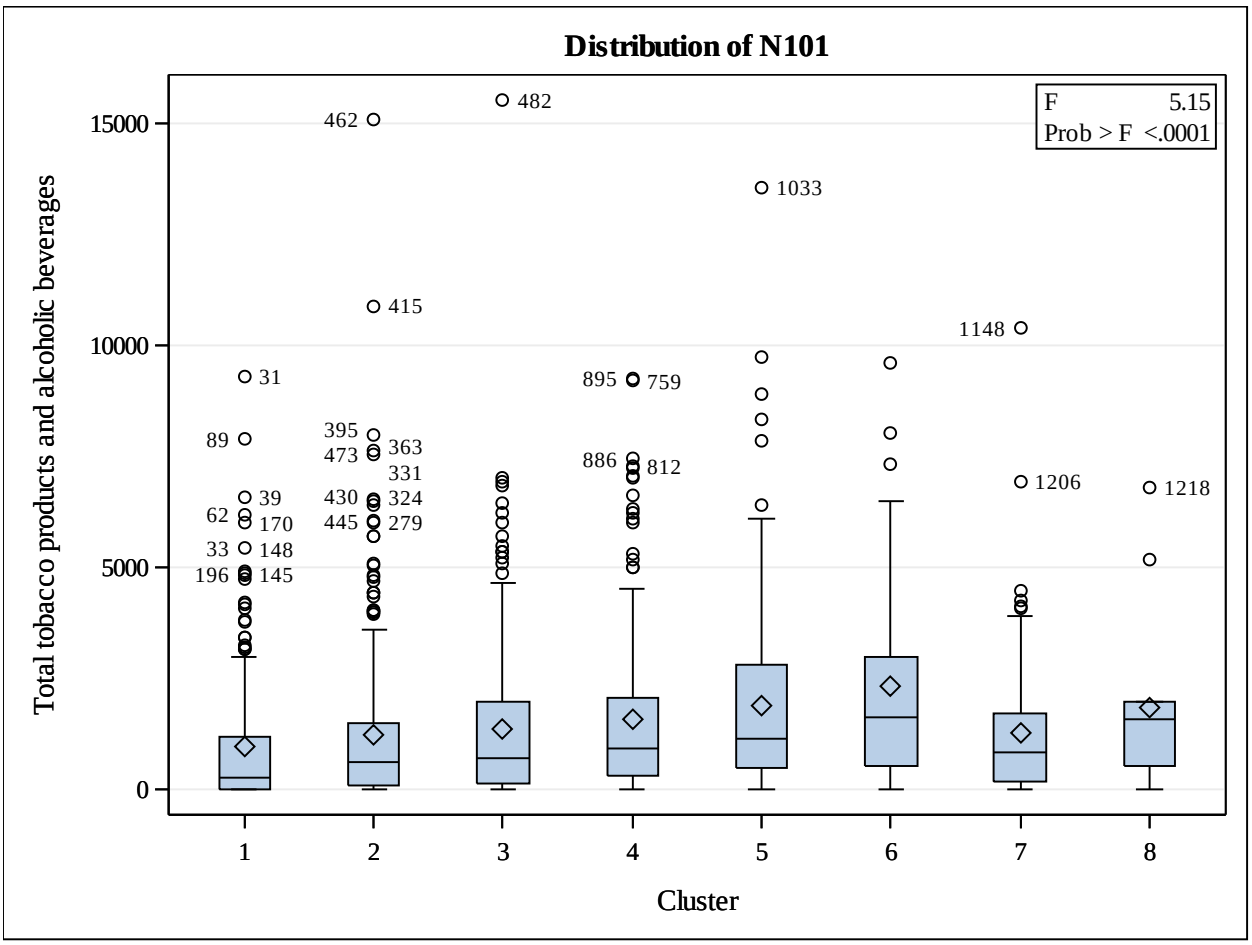
6.8.2.3 Distribution de G004, le loyer.



6.8.2.4 Analyse plus fine, groupes 2 à 2 : variable G003

Differences of CLUSTER Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer												
Cluster	**Cluster**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adj P**	**Alpha**	**Lower**	**Upper**	**Adj Lower**	**Adj Upper**
1	2	3015.46	422.67	1	7.13	0.0887	0.2215	0.05	-2355.11	8386.02	-10553	16584
1	3	5216.06	426.12	1	12.24	0.0519	0.1305	0.05	-198.36	10630	-8462.83	18895
1	4	6393.43	408.28	1	15.66	0.0406	0.1022	0.05	1205.70	11581	-6712.76	19500
1	5	6392.42	460.68	1	13.88	0.0458	0.1153	0.05	538.98	12246	-8395.60	21180
1	6	6863.20	466.96	1	14.70	0.0432	0.1089	0.05	929.96	12796	-8126.42	21853
1	7	6448.88	445.92	1	14.46	0.0440	0.1106	0.05	782.92	12115	-7865.49	20763
1	8	6896.11	801.20	1	8.61	0.0736	0.1845	0.05	-3284.09	17076	-18823	32615
2	3	2200.60	295.51	1	7.45	0.0850	0.2125	0.05	-1554.26	5955.47	-7285.62	11687
2	4	3377.97	269.15	1	12.55	0.0506	0.1273	0.05	-41.9121	6797.85	-5261.96	12018
2	5	3376.97	343.47	1	9.83	0.0645	0.1620	0.05	-987.19	7741.12	-7648.55	14402
2	6	3847.74	351.84	1	10.94	0.0581	0.1459	0.05	-622.87	8318.35	-7446.73	15142
2	7	3433.42	323.41	1	10.62	0.0598	0.1502	0.05	-675.85	7542.69	-6948.17	13815
2	8	3880.65	740.04	1	5.24	0.1200	0.2972	0.05	-5522.51	13284	-19875	27637
3	4	1177.37	274.54	1	4.29	0.1458	0.3581	0.05	-2310.98	4665.71	-7635.53	9990.26
3	5	1176.36	347.70	1	3.38	0.1830	0.4423	0.05	-3241.64	5594.37	-9985.21	12338
3	6	1647.14	355.98	1	4.63	0.1355	0.3340	0.05	-2876.06	6170.33	-9780.19	13074
3	7	1232.82	327.90	1	3.76	0.1655	0.4032	0.05	-2933.60	5399.24	-9293.16	11759
3	8	1680.05	742.02	1	2.26	0.2648	0.6109	0.05	-7748.23	11108	-22139	25499
4	5	-1.0051	325.59	1	-0.00	0.9980	1.0000	0.05	-4138.08	4136.07	-10453	10451
4	6	469.77	334.42	1	1.40	0.3939	0.8158	0.05	-3779.46	4718.99	-10265	11205
4	7	55.4480	304.36	1	0.18	0.8853	1.0000	0.05	-3811.81	3922.70	-9714.71	9825.61
4	8	502.68	731.92	1	0.69	0.6169	0.9795	0.05	-8797.26	9802.63	-22993	23998
5	6	470.77	396.69	1	1.19	0.4458	0.8746	0.05	-4569.68	5511.22	-12263	13205
5	7	56.4531	371.70	1	0.15	0.9040	1.0000	0.05	-4666.46	4779.36	-11875	11988
5	8	503.69	762.39	1	0.66	0.6283	0.9826	0.05	-9183.38	10191	-23970	24977
6	7	-414.32	379.46	1	-1.09	0.4721	0.8991	0.05	-5235.77	4407.13	-12595	11767
6	8	32.9153	766.20	1	0.04	0.9727	1.0000	0.05	-9702.58	9768.41	-24563	24628
7	8	447.23	753.56	1	0.59	0.6590	0.9891	0.05	-9127.71	10022	-23743	24637

6.8.2.5 Distribution de N101, dépenses pour le tabac et alcohol.



6.8.2.6 Analyse plus fine, groupes 2 à 2 : variable N101

Differences of CLUSTER Least Squares Means Adjustment for Multiple Comparisons: Tukey-Kramer												
Cluster	**Cluster**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adj P**	**Alpha**	**Lower**	**Upper**	**Adj Lower**	**Adj Upper**
1	2	-281.09	152.59	1	-1.84	0.3166	0.7029	0.05	-2219.87	1657.70	-5179.19	4617.02
1	3	-469.40	167.21	1	-2.81	0.2179	0.5174	0.05	-2594.04	1655.25	-5837.06	4898.27
1	4	-592.33	167.59	1	-3.53	0.1755	0.4258	0.05	-2721.76	1537.11	-5972.10	4787.45
1	5	-966.03	227.31	1	-4.25	0.1471	0.3611	0.05	-3854.22	1922.16	-8262.69	6330.64
1	6	-1550.51	430.32	1	-3.60	0.1723	0.4187	0.05	-7018.29	3917.26	-15364	12263
1	7	-298.99	189.26	1	-1.58	0.3593	0.7690	0.05	-2703.80	2105.81	-6374.44	5776.45
1	8	-1128.02	528.87	1	-2.13	0.2791	0.6376	0.05	-7847.99	5591.96	-18105	15849
2	3	-188.31	163.00	1	-1.16	0.4542	0.8828	0.05	-2259.40	1882.78	-5420.67	5044.05
2	4	-311.24	163.39	1	-1.90	0.3077	0.6881	0.05	-2387.24	1764.77	-5556.02	4933.54
2	5	-684.94	224.22	1	-3.05	0.2014	0.4825	0.05	-3533.96	2164.08	-7882.65	6512.77
2	6	-1269.42	428.70	1	-2.96	0.2073	0.4952	0.05	-6716.61	4177.76	-15031	12492
2	7	-17.9066	185.55	1	-0.10	0.9388	1.0000	0.05	-2375.52	2339.71	-5974.15	5938.33
2	8	-846.93	527.56	1	-1.61	0.3547	0.7623	0.05	-7550.17	5856.30	-17782	16088
3	4	-122.93	177.12	1	-0.69	0.6138	0.9786	0.05	-2373.49	2127.63	-5808.71	5562.85
3	5	-496.63	234.42	1	-2.12	0.2808	0.6406	0.05	-3475.25	2481.99	-8021.75	7028.49
3	6	-1081.11	434.12	1	-2.49	0.2431	0.5688	0.05	-6597.19	4434.96	-15017	12855
3	7	170.40	197.75	1	0.86	0.5472	0.9514	0.05	-2342.28	2683.09	-6177.60	6518.41
3	8	-658.62	531.97	1	-1.24	0.4325	0.8609	0.05	-7417.95	6100.71	-17735	16418
4	5	-373.70	234.69	1	-1.59	0.3570	0.7657	0.05	-3355.74	2608.33	-7907.46	7160.06
4	6	-958.19	434.27	1	-2.21	0.2709	0.6224	0.05	-6476.11	4559.74	-14899	12982
4	7	293.33	198.07	1	1.48	0.3781	0.7952	0.05	-2223.41	2810.07	-6064.92	6651.58
4	8	-535.69	532.09	1	-1.01	0.4979	0.9200	0.05	-7296.53	6225.15	-17616	16545
5	6	-584.48	460.62	1	-1.27	0.4249	0.8526	0.05	-6437.23	5268.26	-15371	14202
5	7	667.03	250.63	1	2.66	0.2288	0.5400	0.05	-2517.48	3851.54	-7378.25	8712.32
5	8	-161.99	553.81	1	-0.29	0.8188	0.9998	0.05	-7198.76	6874.78	-17940	17616
6	7	1251.52	443.08	1	2.82	0.2166	0.5148	0.05	-4378.41	6881.44	-12972	15475
6	8	422.49	663.48	1	0.64	0.6390	0.9851	0.05	-8007.87	8852.86	-20876	21721
7	8	-829.02	539.31	1	-1.54	0.3672	0.7802	0.05	-7681.58	6023.53	-18141	16483

6.8.3 Les variables catégorielles

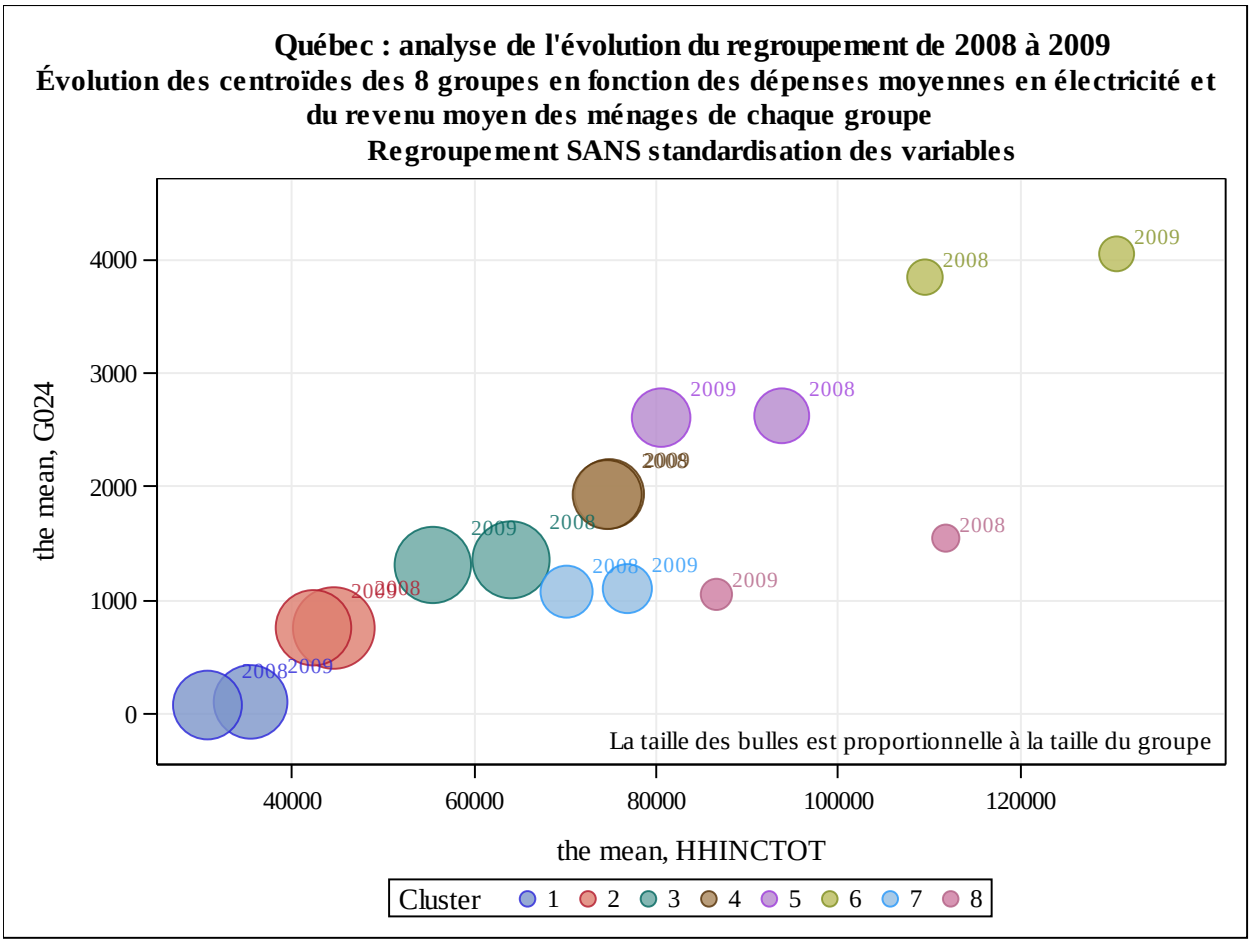
Variables	Catégorie de la variable
Le lieu d'habitation[59]	Urbain
	Rural
Le type de logement	Maison individuelle non attenante
	Maison jumelée ou maison double
	Maison en rangé
	Duplex
	Appartement
	Hôtel, maison de chambre ou pension, maison mobile
Le type du ménage (mode d'occupation)	Propriétaire sans hypothèque
	Propriétiare avec hypothèque
	Locataire (ou occupé sans frais)
	Mode d'occupation mixte
Le type d'appareil de chauffage principal	Fournaise à vapeur ou à eau chaude
	Fournaise à air pulsé
	Poêle (y compris poêle à bois), fourneau de cuisine, ou autre
	Chauffage à l'électricité
L'âge du système de chauffage principal	5 ans ou moins
	6 à 10 ans
	11 à 15 ans
	16 à 20 ans
	Plus de 20 ans
Le type du principal combustible utilisé pour le système de chauffage principal	Mazout ou autre combustible liquide
	Gaz canalisé (gaz naturel)
	Électricité
	Gaz en bouteille (propane), bois de chauffage, ou autre

[59] Lorsqu'une observation était marquée comme « Enregistrement masqués », celle-ci a été éliminée de notre analyse.

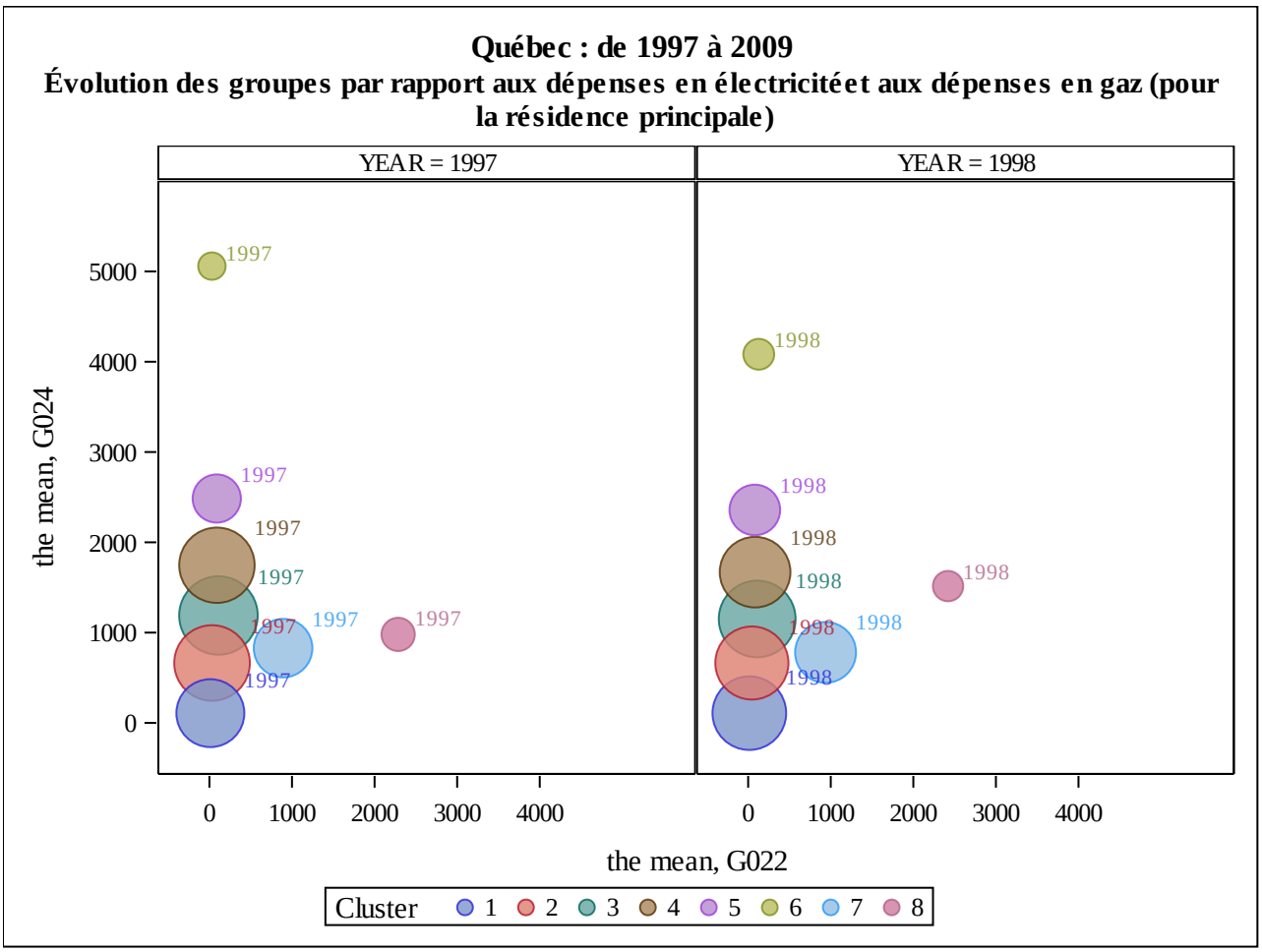
Le type principal du combustible de chauffage pour l'eau chaude	Mazout ou autre combustible liquide
	Gaz canalisé (gaz naturel)
	Gaz en bouteille (propane), autre, ou pas d'eau chaude courante
	Électricité
Le type principal du combustible pour cuisiner	Gaz canalisé (gaz naturel)
	Électricité
	Gaz en bouteille (propane), ou autre

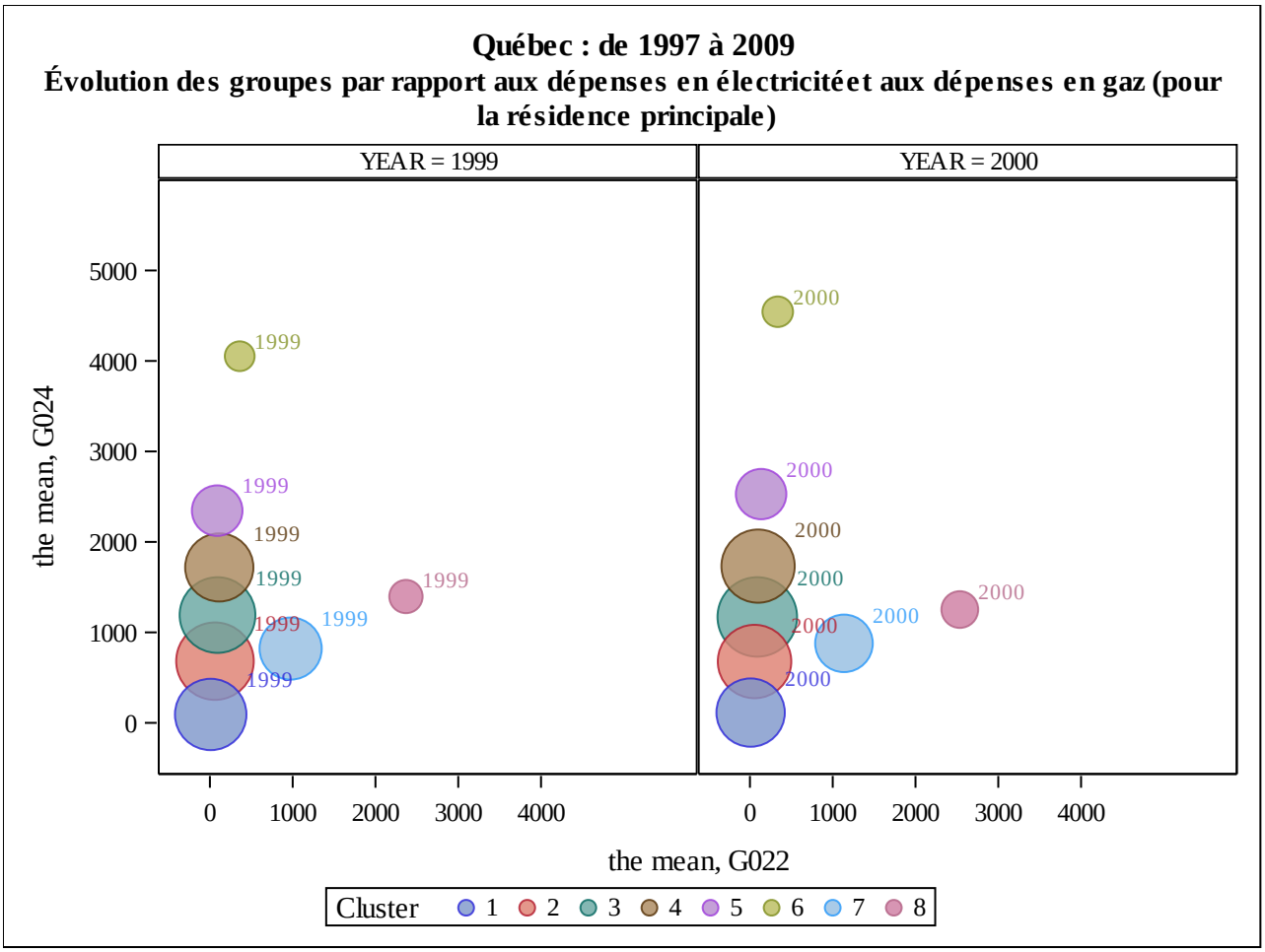
6.9 Annexes – Québec, évolution du portrait des ménages de 1997 à 2009

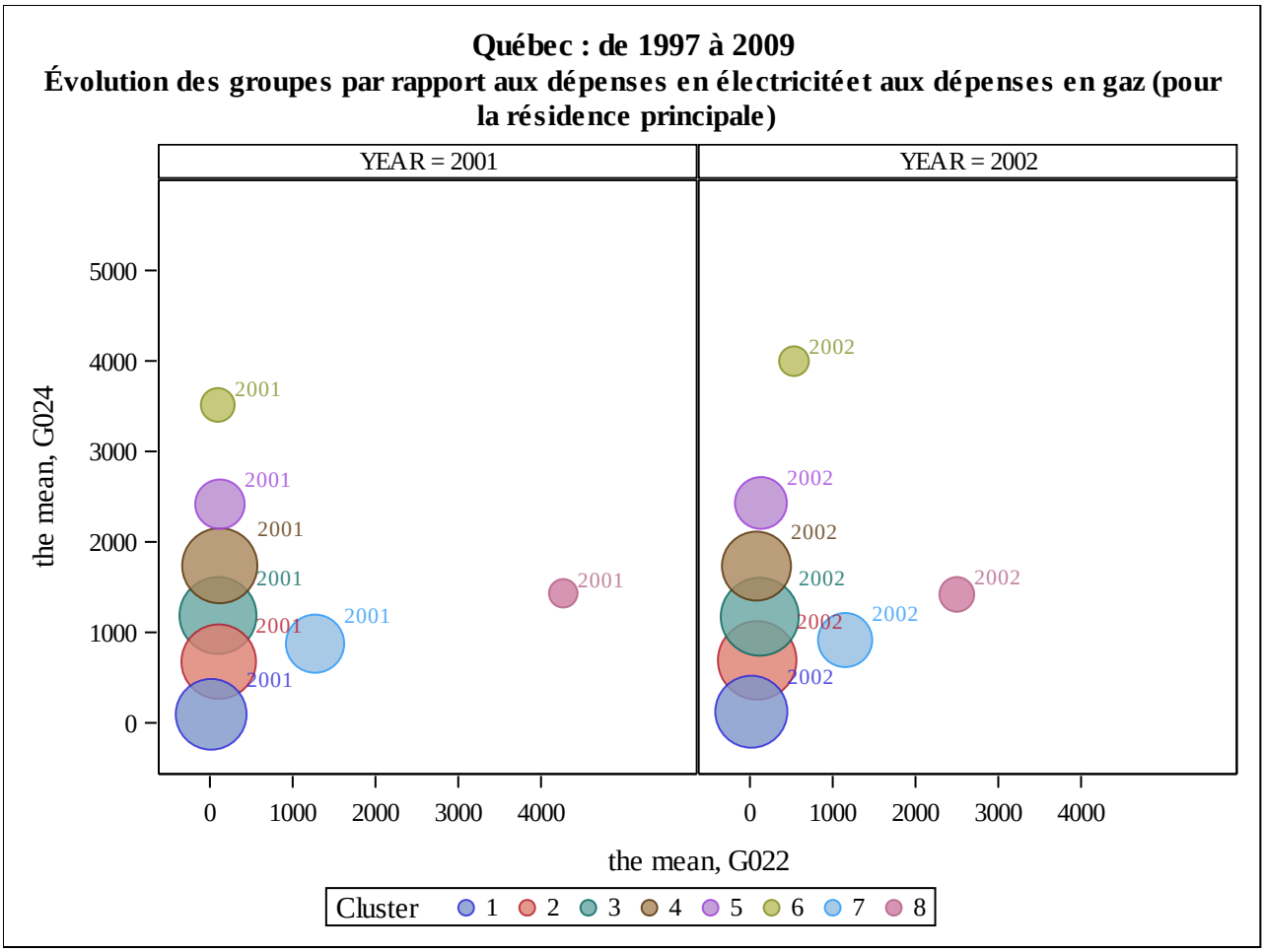
6.9.1 Évolution des centroïdes du regroupement à 8 groupes, de 2008 à 2009 (G024 / HHINCTOT)

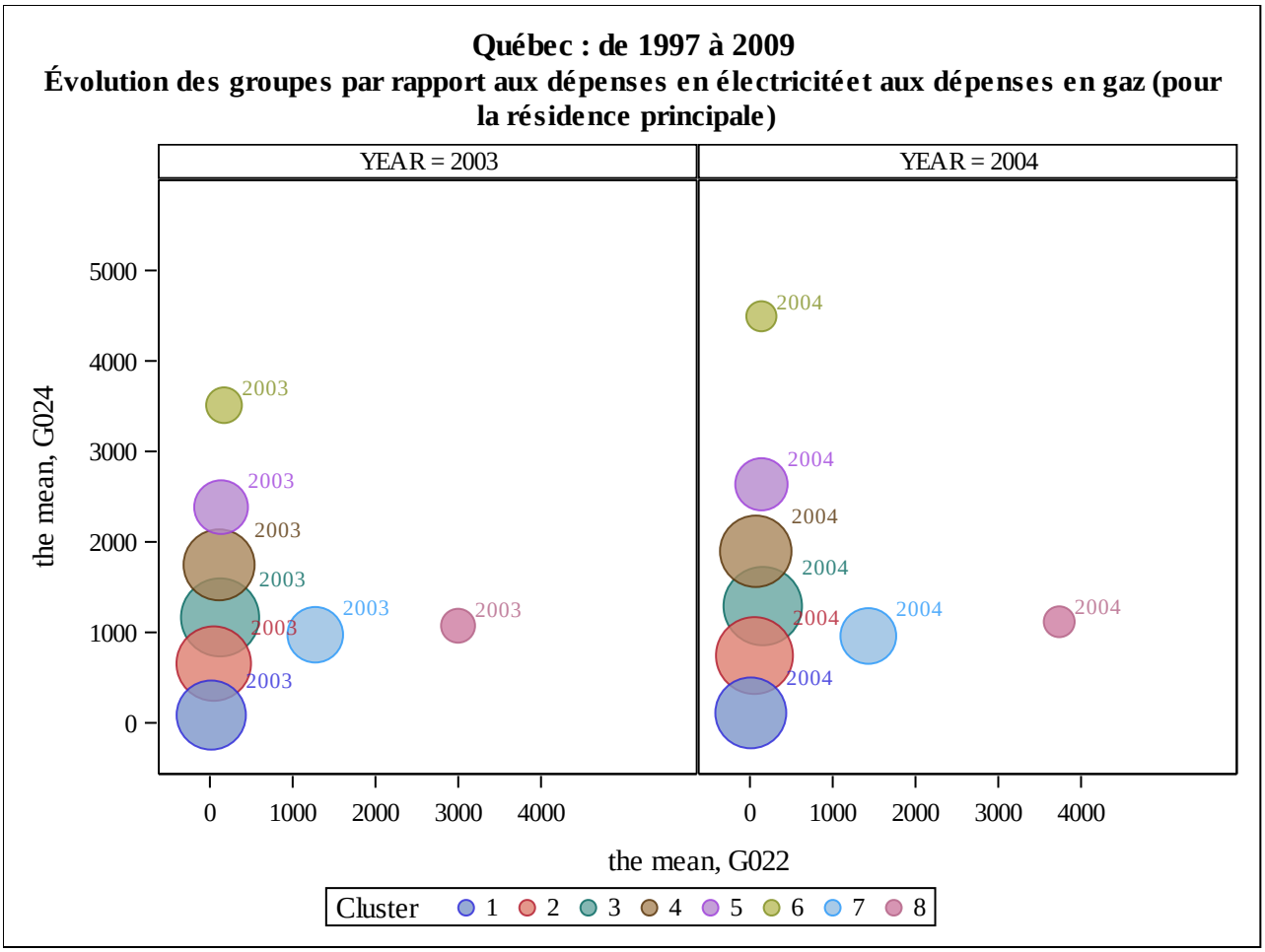


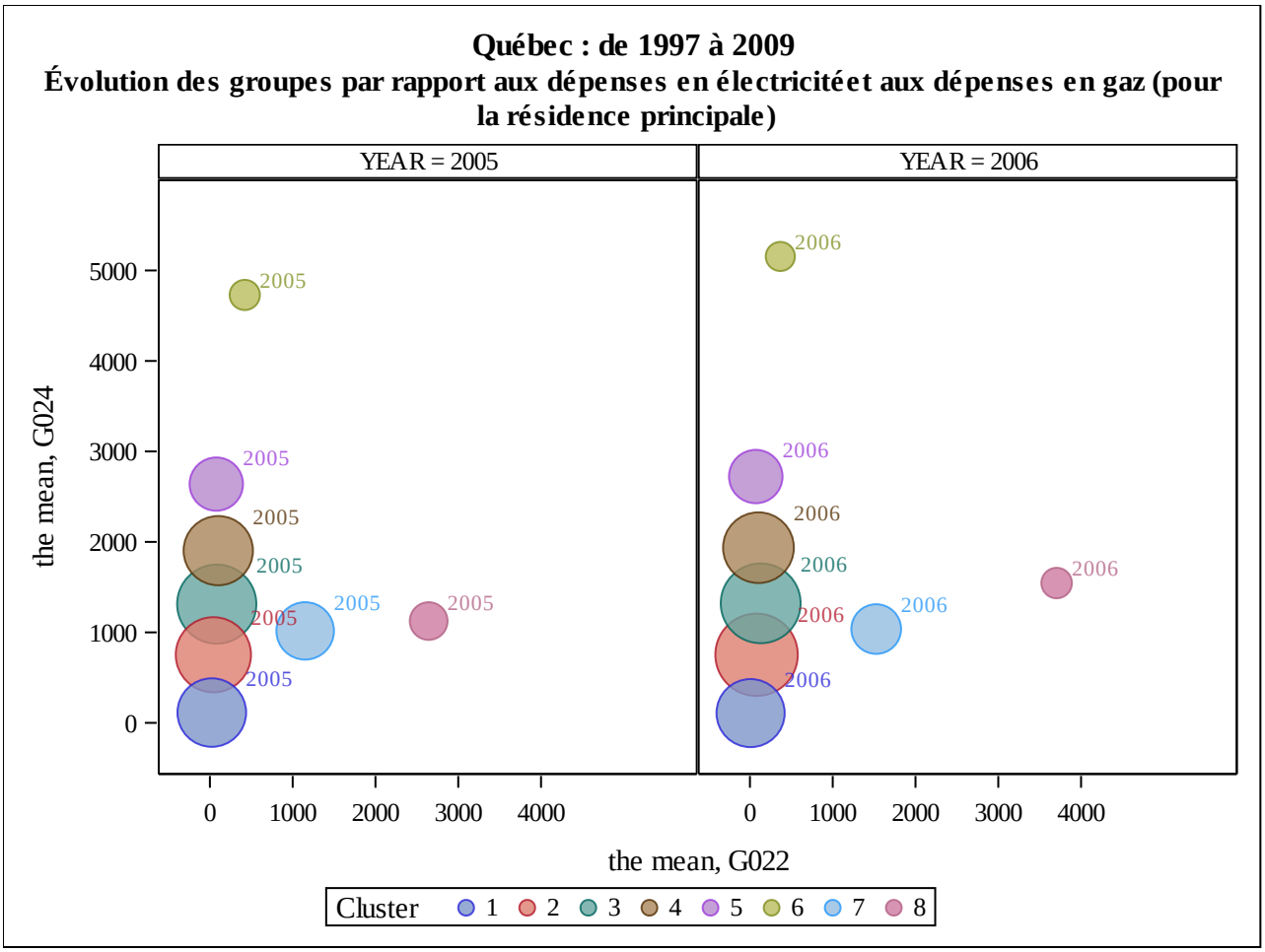
6.9.2 Évolution des centroïdes du regroupement à 8 groupes, de 1997 à 2009 (G024 / G022)

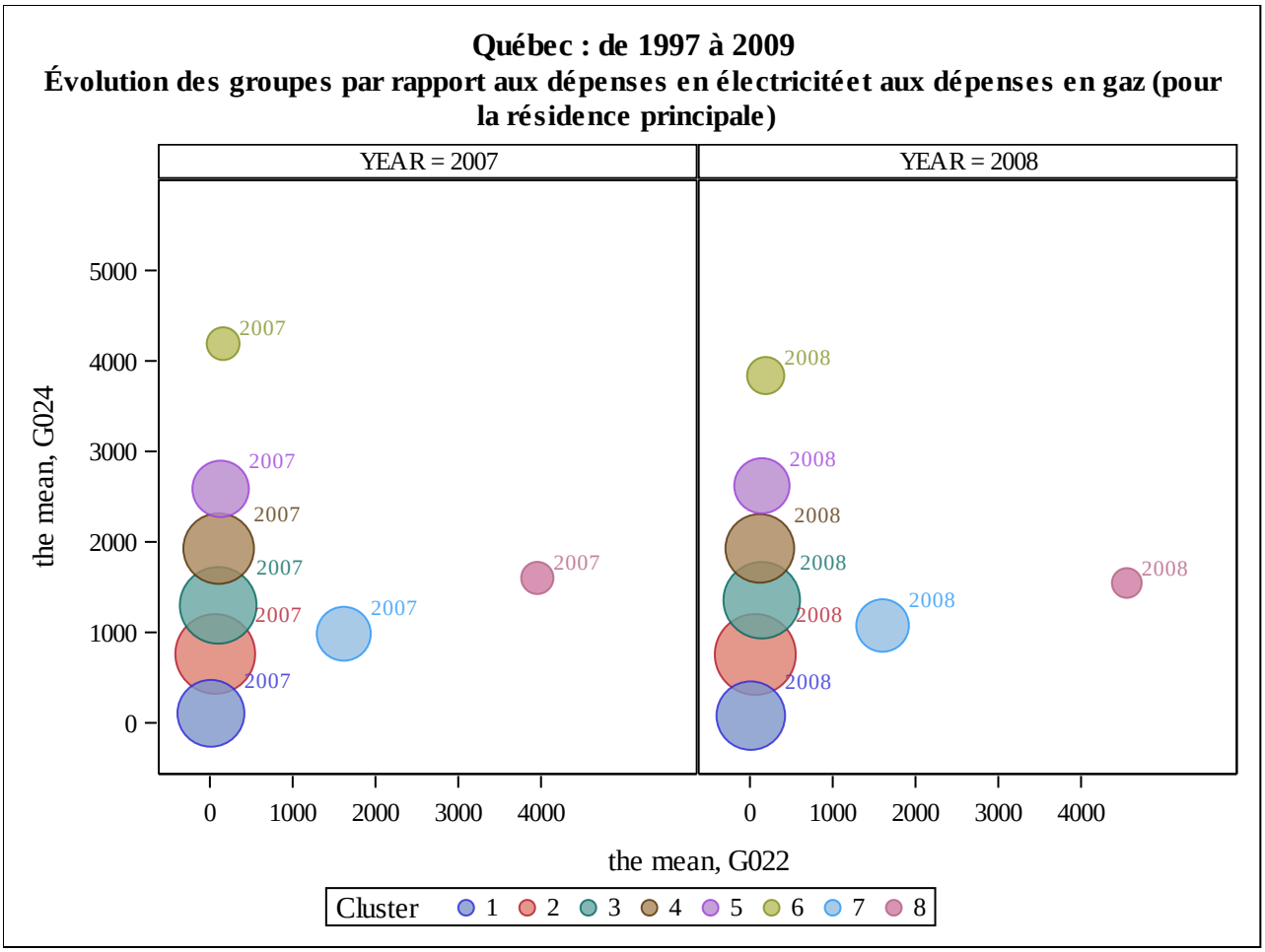


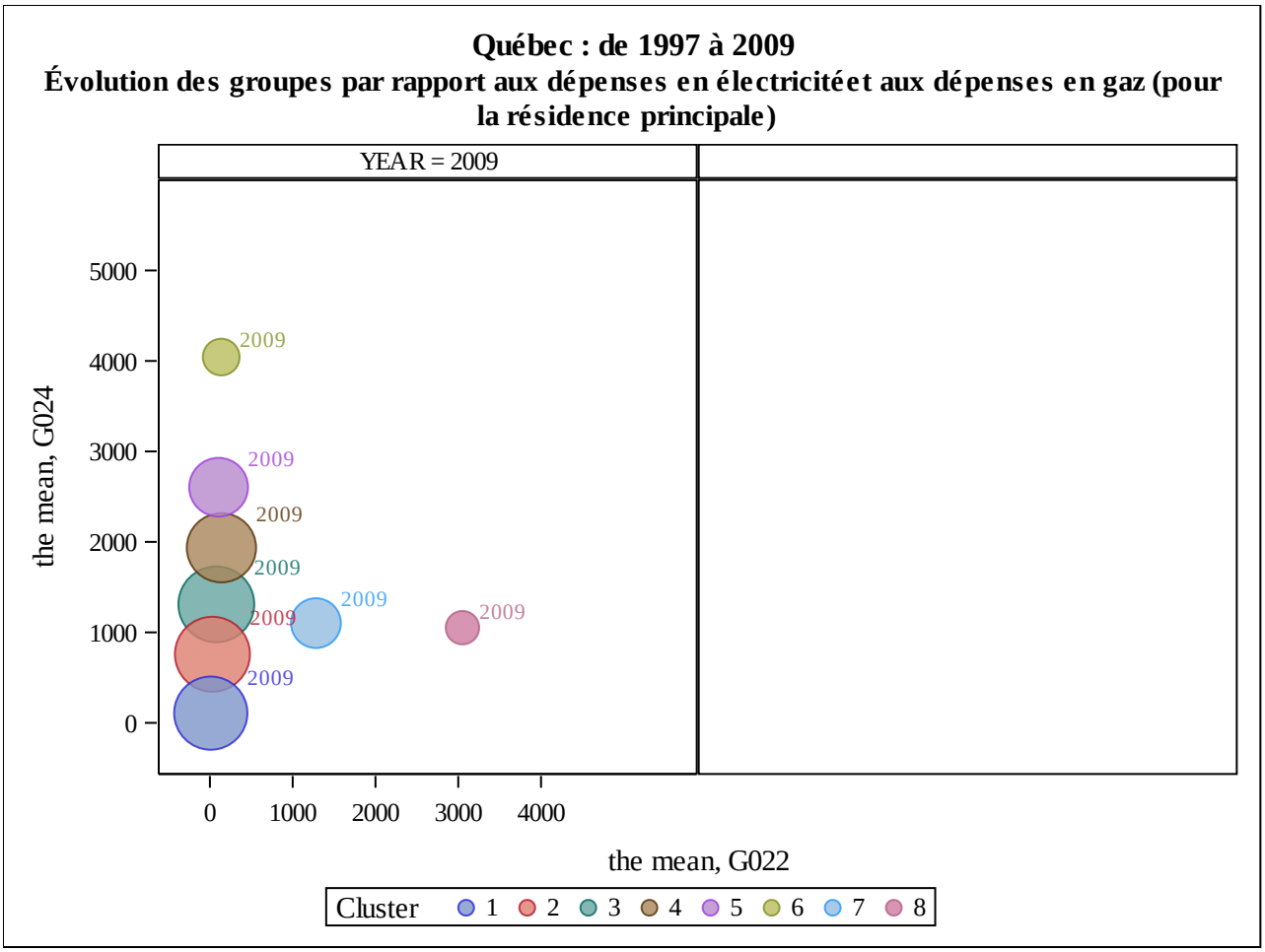

Québec : de 1997 à 2009
Évolution des groupes par rapport aux dépenses en électricitéet aux dépenses en gaz (pour la résidence principale)
YEAR = 1999
YEAR = 2000
the mean, G024
the mean, G022
0
1000
2000
3000
4000
5000
1999
2000
Cluster 1 2 3 4 5 6 7 8


Québec : de 1997 à 2009
Évolution des groupes par rapport aux dépenses en électricitéet aux dépenses en gaz (pour la résidence principale)
YEAR = 2001
YEAR = 2002
the mean, G024
the mean, G022
0
1000
2000
3000
4000
5000
2001
2002
Cluster 1 2 3 4 5 6 7 8

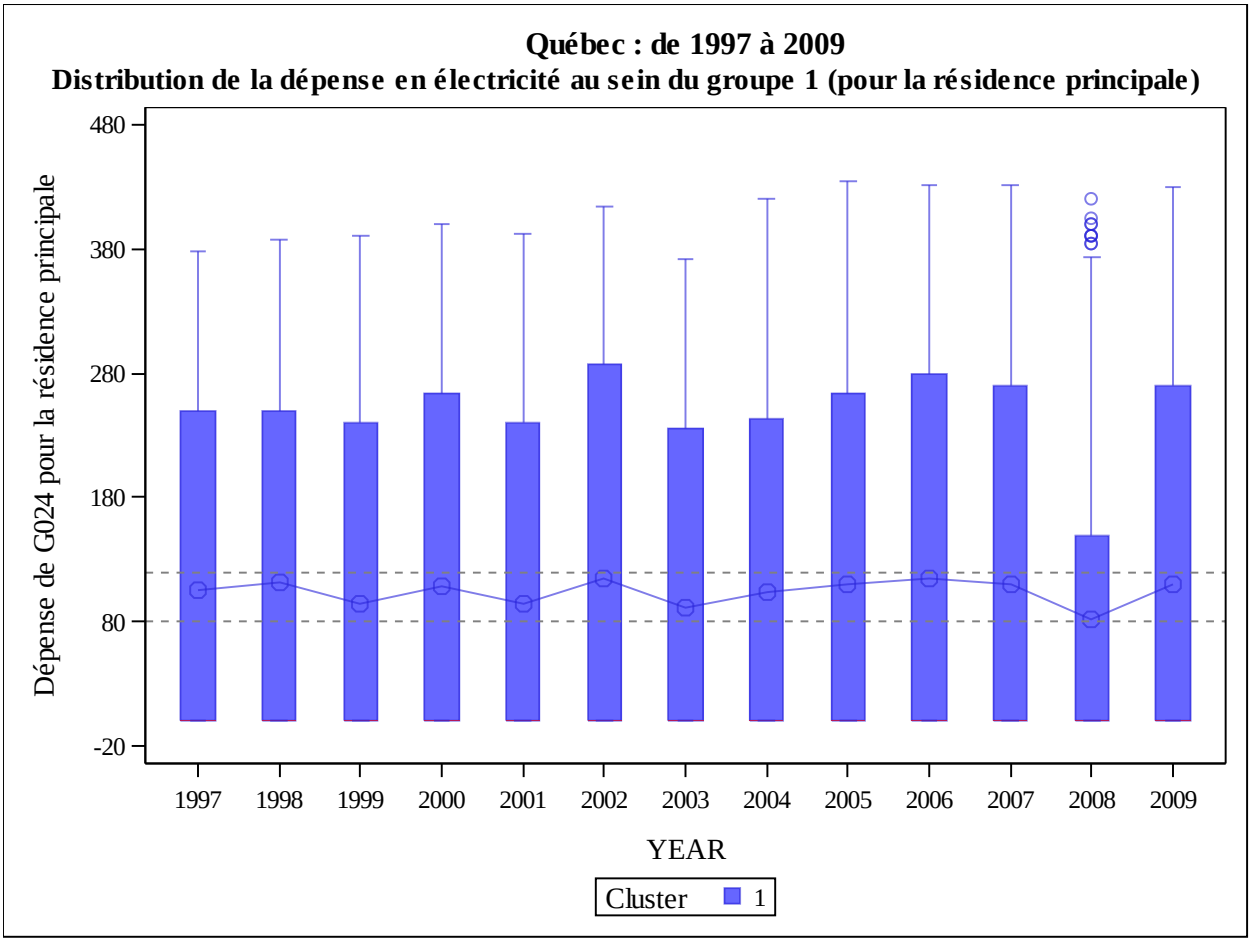

Québec : de 1997 à 2009
Évolution des groupes par rapport aux dépenses en électricitéet aux dépenses en gaz (pour la résidence principale)
YEAR = 2003
YEAR = 2004
the mean, G024
the mean, G022
0
1000
2000
3000
4000
5000
2003
2004
Cluster 1 2 3 4 5 6 7 8


Québec : de 1997 à 2009
Évolution des groupes par rapport aux dépenses en électricitéet aux dépenses en gaz (pour la résidence principale)
YEAR = 2005
YEAR = 2006
the mean, G024
the mean, G022
0
1000
2000
3000
4000
5000
2005
2006
Cluster 1 2 3 4 5 6 7 8


Québec : de 1997 à 2009
Évolution des groupes par rapport aux dépenses en électricitéet aux dépenses en gaz (pour la résidence principale)
YEAR = 2007
YEAR = 2008
the mean, G024
the mean, G022
0
1000
2000
3000
4000
5000
2007
2008
Cluster 1 2 3 4 5 6 7 8


Québec : de 1997 à 2009
Évolution des groupes par rapport aux dépenses en électricitéet aux dépenses en gaz (pour la résidence principale)
YEAR = 2009
the mean, G024
5000
4000
3000
2000
1000
0
2009
2009
2009
2009
2009
2009
2009
0
1000
2000
3000
4000
the mean, G022
Cluster
1
2
3
4
5
6
7
8

6.9.3 Évolution de la variable G024 par le groupe 1[60] entre 1997 et 2009



Tous les ménages du groupe 1 sont inclus dans la figure ci-dessus.

[60] La barre rouge dans le graphique indique la médiane.

7 Bibliographie

Abrahamse, W., & Steg, L. (2009). How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings? *Journal of Economic Psychology*, *30*(5), 711–720. http://doi.org/10.1016/j.joep.2009.05.006

Bartusch, C., Odlare, M., Wallin, F., & Wester, L. (2012). Exploring variance in residential electricity consumption: Household features and building properties. *Applied Energy*, *92*(0), 637–643. http://doi.org/10.1016/j.apenergy.2011.04.034

Basadur, M., Ellspermann, S., & Evans, G. (1994). A new methodology for formulating ill-structured problems. *Omega*, 627 – 645.

Beckel, C., Sadamori, L., Staake, T., & Santini, S. (2014). Revealing household characteristics from smart meter data. *Energy*, *78*(0), 397–410. http://doi.org/10.1016/j.energy.2014.10.025

Brounen, D., Kok, N., & Quigley, J. M. (2012). Residential energy use and conservation: Economics and demographics. *Green Building, the Economy, and Public Policy*, *56*(5), 931–945. http://doi.org/10.1016/j.euroecorev.2012.02.007

Cashin, D. B., & McGranahan, L. (2006). Household energy expenditures, 1982 to 2005. *Chicago Fed Letter*, (Jun). Retrieved from http://ideas.repec.org/a/fip/fedhle/y2006ijunn227.html

Cayla, J.-M., Maizi, N., & Marchand, C. (2011). The role of income in energy consumption behaviour: Evidence from French households data. *Clean Cooking Fuels and Technologies in Developing Economies*, *39*(12), 7874–7883. http://doi.org/10.1016/j.enpol.2011.09.036

Chaire de gestion du secteur de l'énergie. (n.d.). Retrieved December 13, 2014, from http://energie.hec.ca/

Chaire de gestion du secteur de l'énergie – État de l'énergie au Québec. (n.d.). Retrieved December 13, 2014, from http://energie.hec.ca/publications/etat-de-lenergie-au-quebec/

Chicco, G., Napoli, R., Postolache, P., Scutariu, M., & Toader, C. (2003). Customer characterization options for improving the tariff offer. *Power Systems, IEEE Transactions on*, *18*(1), 381–387. http://doi.org/10.1109/TPWRS.2002.807085

Colton, R. D. (2002). Energy Consumption and Expenditures by Low-Income Customers. *The Electricity Journal*, *15*(3), 70–75. http://doi.org/10.1016/S1040-6190(02)00279-8

Enquête sur l'équipement ménager (EEM). (n.d.). Retrieved December 14, 2014, from http://www23.statcan.gc.ca/imdb/p2SV_f.pl?Function=getSurvey&SDDS=3505

Enquête sur les dépenses des familles (Famex). (n.d.). Retrieved December 14, 2014, from http://www23.statcan.gc.ca/imdb/p2SV_f.pl?Function=getSurvey&SDDS=3504

Enquête sur les dépenses des ménages (EDM). (n.d.). Retrieved December 14, 2014, from http://www23.statcan.gc.ca/imdb/p2SV_f.pl?Function=getSurvey&SDDS=3508

Estiri, H. (2014). Building and household X-factors and energy consumption at the residential sector: A structural equation analysis of the effects of household and building characteristics on the annual energy consumption of US residential buildings. *Energy Economics*, *43*(0), 178–184. http://doi.org/10.1016/j.eneco.2014.02.013

Ewing, R., & Rong, F. (2008). The impact of urban form on U.S. residential energy use. *Housing Policy Debate*, *19*(1), 1–30. http://doi.org/10.1080/10511482.2008.9521624

Gilg, A., Barr, S., & Ford, N. (2005). Green consumption or sustainable lifestyles? Identifying the sustainable consumer. *Futures*, *37*(6), 481–504. http://doi.org/10.1016/j.futures.2004.10.016

Hansen, P., Mladenovic, N., & Moreno pérez, J. A. (2010). Variable neighbourhood search: methods and applications. *Annals of Operations Research*, *175*(1), 367–407. http://doi.org/10.1007/s10479-009-0657-6

Hayn, M., Bertsch, V., & Fichtner, W. (2014). Electricity load profiles in Europe: The importance of household segmentation. *Energy Research & Social Science*, *3*(0), 30–45. http://doi.org/10.1016/j.erss.2014.07.002

Hirst, E., Goeltz, R., & Carney, J. (1982). Residential energy use: Analysis of disaggregate data. *Energy Economics*, *4*(2), 74–82. http://doi.org/10.1016/0140-9883(82)90024-X

Hojjati, B., & Wade, S. H. (2012). U.S. household energy consumption and intensity trends: A decomposition approach. *Special Section: Frontiers of Sustainability*, *48*(0), 304–314. http://doi.org/10.1016/j.enpol.2012.05.024

Hydro-Québec. (n.d.). *Comparaison des prix de l'électricité dans les grandes villes nord-américaines.* (No. Rapports de 1997 à 2009).

Jones, R. V., Fuertes, A., & Lomas, K. J. (2015). The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings. *Renewable and Sustainable Energy Reviews*, *43*(0), 901–917. http://doi.org/10.1016/j.rser.2014.11.084

Joyeux, R., & Ripple, R. D. (2007). Household energy consumption versus income and relative standard of living: A panel approach. *Energy Policy*, *35*(1), 50–60. http://doi.org/10.1016/j.enpol.2005.10.012

Jungsuk Kwac, Flora, J., & Rajagopal, R. (2014). Household Energy Consumption Segmentation Using Hourly Data. *Smart Grid, IEEE Transactions on*, *5*(1), 420–430. http://doi.org/10.1109/TSG.2013.2278477

Karanfil, F. (2009). How many times again will we examine the energy-income nexus using a limited range of traditional econometric tools? *Energy Policy*, *37*(4), 1191–1194. http://doi.org/10.1016/j.enpol.2008.11.029

Kaza, N. (2010). Understanding the spectrum of residential energy consumption: A quantile regression approach. *Energy Efficiency Policies and Strategies with Regular Papers.*, *38*(11), 6574–6585. http://doi.org/10.1016/j.enpol.2010.06.028

Levinson, A., & Niemann, S. (2004). Energy use by apartment tenants when landlords pay for utilities. *Resource and Energy Economics*, *26*(1), 51–75. http://doi.org/10.1016/S0928-7655(03)00047-2

Maruejols, L., & Young, D. (2011). Split incentives and energy efficiency in Canadian multi-family dwellings. *Energy Policy*, *39*(6), 3655–3668. http://doi.org/10.1016/j.enpol.2011.03.072

McKenna, E., Richardson, I., & Thomson, M. (2012). Smart meter data: Balancing consumer privacy concerns with legitimate applications. *Modeling Transport (Energy) Demand and Policies*, *41*(0), 807–814. http://doi.org/10.1016/j.enpol.2011.11.049

McLoughlin, F., Duffy, A., & Conlon, M. (2012). Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: An Irish case study. *Energy and Buildings*, *48*(0), 240–248. http://doi.org/10.1016/j.enbuild.2012.01.037

MERN - Prix du gaz naturel. (n.d.). Retrieved January 11, 2015, from http://www.mern.gouv.qc.ca/energie/statistiques/statistiques-energie-prix-gaz.jsp

O'Neill, B. C., & Chen, B. S. (2002). Demographic Determinants of Household Energy Use in the United States. *Population and Development Review*, *28*(ArticleType: research-article / Issue Title: Supplement: Population and Environment: Methods of Analysis / Full

publication date: 2002 / Copyright © 2002 Population Council), 53–88. http://doi.org/10.2307/3115268

Plummer, J. T. (1974). The Concept and Application of Life Style Segmentation. *Journal of Marketing*, *38*(1), 33.

Poyer, D. A., Henderson, L., & Teotia, A. P. S. (1997). Residential energy consumption across different population groups: comparative analysis for Latino and non-Latino households in USA. *Energy Economics*, *19*(4), 445–463. http://doi.org/10.1016/S0140-9883(97)01024-4

Poyer, D. A., & Williams, M. (1993). Residential energy demand: additional empirical evidence by minority household type. *Energy Economics*, *15*(2), 93–100. http://doi.org/10.1016/0140-9883(93)90027-O

Résumé chronologique des changements - Enquête sur les dépenses des ménages (EDM). (n.d.). Retrieved December 14, 2014, from http://www23.statcan.gc.ca/imdb/p2SV_f.pl?Function=getMainChange&InstaId=15492&SurvId=3508&SurvVer=2&SDDS=3508

Sanquist, T. F., Orr, H., Shui, B., & Bittner, A. C. (2012). Lifestyle factors in U.S. residential electricity consumption. *Energy Policy*, *42*(0), 354–364. http://doi.org/10.1016/j.enpol.2011.11.092

Santamouris, M., Kapsis, K., Korres, D., Livada, I., Pavlou, C., & Assimakopoulos, M. N. (2007). On the relation between the energy and social characteristics of the residential sector. *Energy and Buildings*, *39*(8), 893–905. http://doi.org/10.1016/j.enbuild.2006.11.001

SAS/STAT(R) 9.2 User's Guide, Second Edition. (n.d.). Retrieved January 12, 2015, from http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_boxplot_sect012.htm

Schipper, L., Bartlett, S., Hawk, D., & Vine, E. (1989). Linking Life-Styles and Energy Use: A Matter of Time? *Annual Review of Energy*, *14*(1), 273–320. http://doi.org/10.1146/annurev.eg.14.110189.001421

Steinley, D., & Brusco, M. J. (2007). Initializing K-means Batch Clustering: A Critical Evaluation of Several Techniques. *Journal of Classification*, *24*(1), 99–121.

Sütterlin, B., Brunner, T. A., & Siegrist, M. (2011). Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy-related behavioral characteristics. *Clean Cooking Fuels and Technologies in Developing Economies*, *39*(12), 8137–8152. http://doi.org/10.1016/j.enpol.2011.10.008

Tufféry, S. (2012). *Data Mining et statistique décisionnelle, l'intelligence des données.* (Éditions TECHNIP).

Westfall, P., Tobias, R., & Russell, W. (2011). *Multiple Comparisons and Multiple Tests Using SAS* (Second Edition). SAS Institute.

Yao, R., & Steemers, K. (2005). A method of formulating energy load profile for domestic buildings in the UK. *Energy and Buildings*, *37*(6), 663–671. http://doi.org/10.1016/j.enbuild.2004.09.007

Yun, G. Y., & Steemers, K. (2011). Behavioural, physical and socio-economic factors in household cooling energy consumption. *Applied Energy*, *88*(6), 2191–2200. http://doi.org/10.1016/j.apenergy.2011.01.010