



Observatoire de la qualité de
l'air intérieur



Rapport intermédiaire

Juin 2010

Programme OQAI Logements

Erreur ! Source du renvoi introuvable.
Rapport intermédiaire

Convention xxx

Responsable scientifique : S. THiria

Auteurs

LOCEAN, Université Pierre et Marie Curie

Juin 2010

Diffusion restreinte

Toute reproduction ou représentation intégrale ou partielle, par quelque procédé que ce soit, des pages publiées dans le présent document, faite sans l'autorisation du CSTB est illicite et constitue une contrefaçon. Seules sont autorisées, d'une part, les reproductions strictement réservées à l'usage du copiste et non destinées à une utilisation collective et, d'autre part, les analyses et courtes citations justifiées par le caractère scientifique ou d'information de l'œuvre dans laquelle elles sont incorporées (Loi du 1er juillet 1992 - art. L 122-4 et L 122-5 et Code Pénal art. 425).

*Toute citation d'extraits ou reproduction doit obligatoirement faire apparaître la référence de ce document sous la forme : **Insérer ici la référence bibliographique du rapport sous la forme : Nom_auteur1 Initiale du prénom. Nom_auteur2 Initiale du Prénom. Titre, CSTB, date de publication, nombre de pages [n° du rapport]© 2010 CSTB***

Résumé

L'Observatoire de la Qualité de l'Air Intérieur (OQAI) a engagé entre 2003 et 2005 une campagne nationale dans les logements sur un échantillon de 567 logements représentatifs du parc des 24 millions de résidences principales de la France continentale métropolitaine. Cette campagne vise à dresser un état de la pollution de l'air dans l'habitat afin de donner les éléments utiles pour l'estimation de l'exposition des populations, la quantification et la hiérarchisation des risques sanitaires associés, ainsi que l'identification des facteurs prédictifs de la qualité de l'air intérieur.

Les données utilisées ont été préalablement validées par l'OQAI et les données manquantes complétées par l'AFSSET. Étant donné le grand nombre et la diversité des questions qui constituent l'enquête, il a été décidé de les regrouper en trois ensembles cohérents en fonction de trois critères tenant compte respectivement de la structure technique du logement, de la structure des ménages et des habitudes des ménages. Ces trois ensembles ont permis de procéder à trois études séparées, chacune permettant de faire apparaître des groupes cohérents qui permettent de décrire l'échantillon analysé en respectivement six, sept et neuf modalités.

L'analyse descriptive de la multipollution réalisée par l'AFSSET sur les polluants dans les logements a permis de dégager quatre groupes de pollution permettant de décrire la pollution dans les logements. A partir de ces derniers, quatorze sous groupes de pollution ont été caractérisés.

La dernière étude qui fait l'objet de ce rapport s'intéresse aux liens susceptibles d'être établis entre la typologies de logements, la topologie des ménages et la typologie des habitudes. Une analyse statistique multivariée prenant en compte ces typologies a permis de proposer une classification en quatorze classes à partir des 567 logements représentant l'échantillon initial. Ces quatorze classes ont été analysées en termes de pollution en prenant en compte :

- d'une part, les typologies réalisées au préalable sur les polluants (cf. rapport 2 : analyse descriptive de la multipollution)
- d'autre part, chaque polluant prioritaire séparément (liste OQAI des polluants prioritaires).

Ce rapport détaille les différentes analyses et présente une caractérisation de ces quatorze classes. Une description simplifiée des classes versus l'échantillon initiale est donnée ci-après avec les principaux résultats concernant la pollution.

- **La classe 1** (56 logements) est constituée de personnes jeunes (41.80 ans vs 48.97 ans) vivants seules (1.67 vs 2.83) dans les petits logements collectifs de type 5. Ils ont des revenus très inférieurs à la moyenne de l'échantillon. Par rapport à l'échantillon initial les individus de cette classe sont moyennement fumeurs à l'intérieur du logement (74% vs 68 %), ils bricolent et jardinent très peu, ils ont une activité très faible dans leur logement.

Cette classe contient une forte proportion de logements moyennement et positivement multi-pollués en particulier par les hydrocarbures aromatiques (24.07 % vs 14.7 %, $p_v=0.02$). Par ailleurs elle est liée fortement et négativement au formaldéhyde ($p_v=0.02$) et au 1-métoxy-2-propanol ($p_v<0.02$).

- **La classe 2** (40 logements) est constituée de jeunes couples (39.8% vs 48.97%) vivants généralement dans des anciens logements collectifs de type 5 ou dans les maisons individuelles avec des revenus très inférieurs à la moyenne. Par rapport à l'échantillon initial ils sont non fumeurs pour la plupart. Avec peu d'activité dans leur logement ils bricolent et jardinent peu.

Cette classe est fortement dominée par les logements moyennement et positivement multipollués en particulier par les aldéhydes ($p<0.01$). Par ailleurs, elle est liée fortement et négativement au 2-butoxyethanol ($p_v<0.01$).

- **La classe 3** (93 logements) est constituée de personnes vivants majoritairement dans les petites maisons collectives avec des revenus moyens. Par rapport à l'échantillon complet ils ont peu d'activités dans leur logement, ils sont un peu plus fumeurs à l'intérieur de leur logement, ils bricolent et jardinent rarement ou jamais (17 % vs 51 % respectivement 16 % vs 40%) pour la plupart. Cette classe n'est caractérisée ni par un groupe particulier de pollution, ni par un polluant prioritaire.

- **La classe 4** (36 logements) n'est pas caractérisée par aucun type de logements. Les individus de cette classe sont des couples avec grands enfants. Par rapport à l'échantillon ils ont une activité assez importante (6.26 vs 4.56), ils fument moyennement dans leur logement (42 % vs 32 %) et ils bricolent peu. Cette classe est fortement et positivement liée aux logements monopollués par le 2-butoxyethanol (8.82 % vs 1.9 %).

Par ailleurs elle est liée fortement et positivement à plusieurs COV. On y retrouve des liens significatifs et positifs avec les aldéhydes notamment, l'acétaldéhyde ($p_v=0.03$), l'hexaldéhyde ($p_v=0.004$) l'hydrocarbure aromatique 1-2-4-triméthylbenzène ($p_v=0.003$) et le décane ($p_v=0.02$). Enfin, elle est liée dans une moindre mesure au formaldéhyde ($p_v=0.06$) et à l'acroléine ($p_v=0.08$). On retrouve une forte multipollution qui est significativement différente des multipollutions considérées dans la classification établie par AFSSET dans les premières études.

- **La classe 5** (41 logements) est constituée uniquement des retraités vivants dans les petites maisons individuelles séparées du garage et de l'endroit pour bricoler avec absence de cheminée. Par rapport à l'échantillon initial, ils ont une activité faible chez eux, ils sont majoritairement non-fumeur à l'intérieur de leur logement et ont des revenus inférieurs à la moyenne (2525 % vs 2013 %). Cette classe n'est pas caractérisée par un groupe particulier de pollution, en particulier elle est fortement et négativement liée au benzène ($p=0.002$).
- **La classe 6** (67 logements) est constituée essentiellement de grandes maisons individuelles « tout en un » (type 2 47.7%, type 3 50 %) sans cheminée (1.59 % vs 61.71%). Par rapport à l'échantillon les occupants de cette classe ont des revenus légèrement supérieurs à la moyenne avec une faible activité dans leur logement. Cependant, ils bricolent et jardinent beaucoup. Cette classe regroupe les logements fortement monopollués par le n-un décane ($p<0.03$) et le styrène ($p<0.001$). Par ailleurs elle est liée positivement et fortement au Tétrachloroéthylène ($p=0.001$), au benzène ($p=0.02$) et 2-Butoxyéthanol ($p=0.02$).
- **La classe 7** (24 logements) est constituée des maisons individuelles « tout en un » avec une forte présence de cheminée (96 % vs 4%) occupées par des couples âgés qui ont peu d'activité chez eux. A part une plus importante activité de bricolage et de jardinage ils sont majoritairement non fumeurs. Cette classe est majoritairement constituée par les logements fortement multipollués en particulier par les hydrocarbures aromatiques (28.57 % vs 6.2 %). Par ailleurs elle est liée fortement et positivement au 1-méthoxy-2-propanol ($p=0.02$).
- **La classe 8** (31 logements) est composée uniquement de tous les logements ayant la modalité 2 de la typologie habitude. Ce sont des couples avec enfants en bas âges vivant avec des revenus moyens. Ils ont une forte activité dans leur logement (5.74 vs 4.56), ils sont tous non fumeurs, ils bricolent et jardinent très peu. Ils possèdent majoritairement des animaux domestiques surtout les chiens, ils sortent pratiquement tous les jours les ordures ménagères. Cette classe est constituée des logements fortement et positivement monopollués par le n-undecane (7.49 % vs 1.5 %).
- **La classe 9** (38 logements) est constituée d'individus jeunes ayant une situation financière moyenne et une activité assez importante (5.50 vs 4.56) chez eux. Ils sont majoritairement fumeurs (86% vs 33%) et ils utilisent peu d'encens. Cette classe est dans une moindre mesure dominée par les logements moyennement multipollué par les hydrocarbures aromatiques (25 % vs 14.7 %).
- **La classe 10** (35 logements) : elle est caractérisée par les couples avec enfants en bas âges vivant dans des petites maisons individuelles anciennes avec des revenus supérieurs à la moyenne. Ils ont une activité assez forte dans leur logement, et sont très majoritairement non fumeur, ils bricolent tous au moins une fois par semaine. Cette classe est liée fortement et négativement au benzène. Globalement la pollution à l'intérieur de cette classe est semblable à celle de l'échantillon.
- **La classe 11** (21 logements) représente les familles nombreuses habitant des grandes maisons individuelles avec des revenus très nettement supérieurs à la moyenne et ayant une forte activité chez eux. Ils sont majoritairement non fumeurs chez eux, ils bricolent et jardinent moyennement. Cette classe est constituée d'une forte proportion

de maisons individuelles fortement multi-polluées par les hydrocarbures aromatiques (21.87 % vs 6.2 %). Par ailleurs, elle est liée positivement et significativement au Tétrachloroéthylène ($p=0.002$) et au Trichloroéthylène ($p=0.02$). On y retrouve dans des proportions moindres un lien positive avec le 1-métoxy-2-propanol ($p=0.04$).

- **La classe 12** (32 logements) : elle est constituée des grandes maisons individuelles habitées généralement par les familles nombreuses avec des revenus légèrement supérieures à la moyenne et avec de grands enfants. Ils bricolent plus que la moyenne avec une forte activité dans leurs logements, ils sont nombreux à fumer à l'intérieur du logement (94 % vs 68 %) et utilisent beaucoup d'encens. Cette classe est constituée majoritairement par les logements moyennement multipollués en particulier par les aldéhydes (21.87 % vs 10.2 %). Par ailleurs, elle est liée fortement et négativement au Styrène ($p<0.01$), à l'hexaldéhyde ($p<0.01$) et au Trichloroéthylène ($p<0.01$).
- **La classe 13** (22 logements) : Elle est constituée des couples avec enfants en bas âges, ayant une activité forte dans leur logement. Ils bricolent beaucoup (91 % vs 51 %), et sont majoritairement non fumeurs, avec une importante utilisation de désodorisants comme les encens. Cette classe est formée par une forte proportion de logements pollués par le formaldéhyde (15% vs 3.6%). Par ailleurs, elle est liée fortement et négativement au 1-métoxy-2-propanol ($p=0.02$).
- **La classe 14** (28 logements) : La classe 14 est caractérisée par les grands logements individuels occupés généralement par des couples aisés avec grands enfants. Ils ont une activité assez importante dans leur logement, majoritairement non fumeurs, ils bricolent et jardinent beaucoup, ils utilisent beaucoup les encens. Cette classe n'est caractérisée par aucun groupe de pollution. Cependant, elle présente des pollutions assez fortes avec l'hexaldéhyde ($p=0.007$), le Tétrachloroéthylène ($p=0.009$) et le 2-butoxyethanol ($p=0$).

Mots clés : Analyse statistique multivariée, qualité de l'air, polluant, pollution multiple, classification hiérarchique ascendante, test d'indépendance.

Table des matières

INTRODUCTION.....	11
1. DONNÉES ET MÉTHODES	12
1.1 Variables de l'étude	12
1.2 Méthodes statistiques appliquées.....	13
a. Test du KHI2	13
b. Test de comparaison des moyennes de wilcoxon	13
c. Analyse factorielle Multiples des Correspondances (ACM)	14
d. Description d'une sous population par une variable qualitative (valeur.test)	14
e. Classification Ascendante Hiérarchique (CAH) sur les composantes principales issues de l'ACM	15
2. PRÉSENTATION DES RÉSULTATS	15
2.1 Test de KHI2	15
2.2 Analyse Factorielle des Correspondances Multiples (ACM)..	16
2.3 Détermination des classes d'individus	18
3. LIEN AVEC LA POLLUTION	25
3.1 Lien avec les groupes de pollution	26
3.2 Lien avec les polluants prioritaires	29
RÉFÉRENCES BIBLIOGRAPHIQUES	36
ANNEXES	38
ANNEXE 1 :	39
ANNEXE 2 :	42
ANNEXE 3 : LES VARIABLES PRISES EN COMPTES POUR DÉCRIRE LES 14 CLASSES	44

Liste des tableaux

Tableau 1- résultats du test de kHI2 entre les trois typologies.	15
Tableau 2: les variables importantes définies lors des études précédentes	19
Tableau 3 : résumé de la description des groupes de pollutions	25
Tableau 5 : Caractéristique des 14 classes avec les groupes de pollutions	28
Tableau 6 : table descriptive des COV et autres polluants prioritaire établis par L'OQAI.....	31
Tableau 7 : lien entre les classes 1 à 7 et les polluants prioritaires.	33
Tableau 8 : lien entre les classes 8 à 14 et les polluants prioritaires.	34
Tableau 9 : lien entre la pollution intérieur et la variable Fumeur.	35

Liste des figures

Figure 1: pourcentage pour chaque typologie et chaque modalité des 567 logements de l'enquête OQAI	12
Figure 2 : Rapport de corrélation entre les trois typologies et le premier plan factoriel.	16
Figure 3 : diagramme des inerties associées aux axes factoriels	17
Figure 4 : représentation des modalités sur le plan factoriel 1 et 2.	18
Figure 5: représentation box plot des 14 polluants	30

Introduction

L'observatoire de la qualité de l'air intérieur (OQAI) a engagé une campagne nationale dans les logements sur un échantillon de 567 logements représentatifs du parc des 24 millions de résidences principales de la France continentale métropolitaine. Cette campagne vise à dresser un état de la pollution de l'air dans les habitats afin de donner des éléments utiles pour l'estimation de l'exposition des populations, la quantification et la hiérarchisation des risques sanitaires associés, l'identification des facteurs prédictifs de la qualité de l'air intérieur.

Plus de 30 paramètres (chimiques, biologiques, physiques) de pollution ont été mesurés, sur une durée d'une semaine, à plusieurs emplacements à l'intérieur des logements. Dans le même temps des informations détaillées ont été collectées sur les caractéristiques techniques des logements, sur leur environnement ainsi que sur les ménages et leurs activités à travers un questionnaire.

Des études précédentes ont permis d'établir trois typologies suivant les critères, de logements, de ménages et des habitudes des ménages :

- La typologie des logements est liée entre autre au statut d'occupation, à la structure technique du bâtiment, au revêtement et à l'aménagement intérieur du logement.
- La typologie des ménages est liée à la constitution des ménages, elle est caractérisée entre autre par le nombre de personnes vivants dans le logement, l'âge des personnes, les revenus du ménage.
- La typologie habitudes des ménages est liée entre autre au comportement des personnes vivant dans le logement. Elle tient compte de l'utilisation des produits cosmétiques, du traitement des ordures ménagères, des activités de bricolage et de jardinage.

L'objectif de cette étude est de définir des classes homogènes et significatives d'individus en croisant les trois typologies (logements, ménages et d'habitudes) à l'aide d'analyses statistiques multivariées. Chaque classe sera alors caractérisée par rapport à la pollution à l'intérieur des logements (polluants). Le rapport commence par une présentation des méthodes statistiques utilisées et donne les résultats obtenus sur l'enquête.

1. Données et méthodes

1.1 Variables de l'étude

Les variables utilisées dans cette étude concernent globalement toutes les variables issues de la base OQAI et sont indicées sur les logements.

Une étude préalable a permis de définir les variables de l'étude, les données manquantes ont été complétées par l'AFSSET et l'OQAI. L'ensemble des variables a été repartitionné en trois sous-ensembles de questions cohérentes relatives aux logements, aux habitudes des habitants et à la composition des ménages. Pour chaque sous-ensemble une typologie répartissant les 567 logements en différents types a été réalisée et analysée (cf. rapport 1 sur la typologie logement, rapport 3 sur la typologie des habitudes et rapport 4 sur la typologie des ménages). On retrouve en annexe 1 la description de chaque type pour chacune des trois typologies. La figure 1 redonne les effectifs relatifs à chaque modalité.

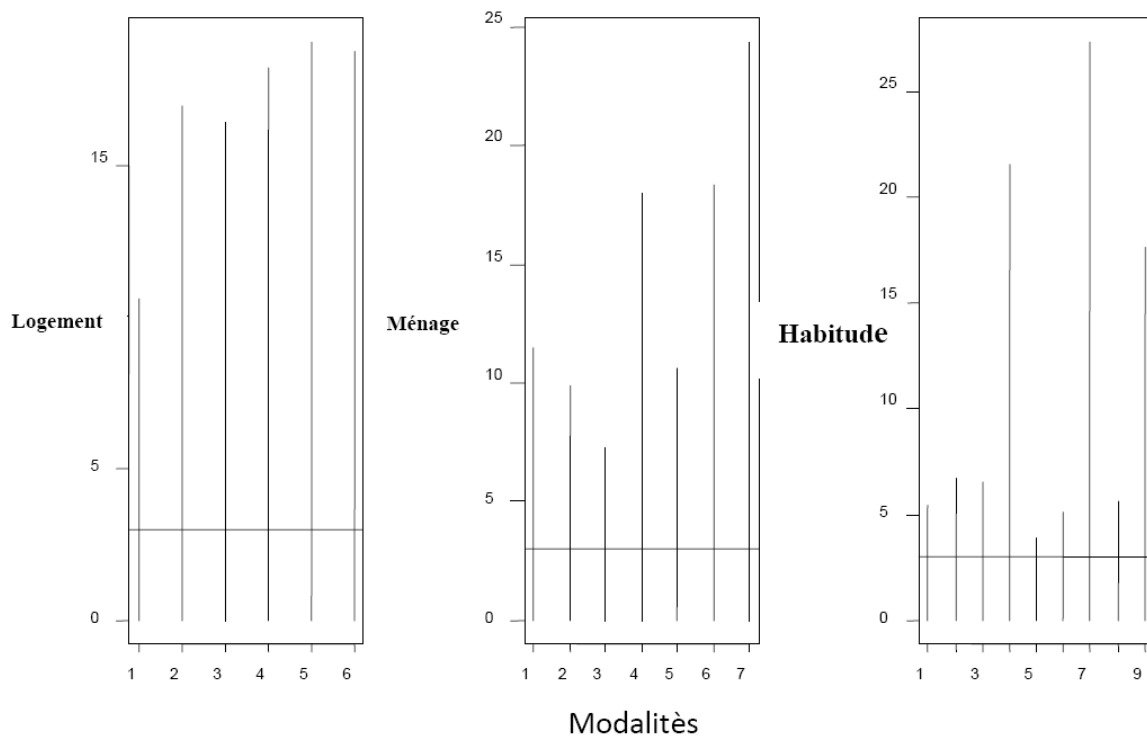


Figure 1: pourcentage pour chaque typologie et chaque modalité des 567 logements de l'enquête OQAI

1.2 Méthodes statistiques appliquées.

a. Test du KHI2

L'analyse de la dépendance entre deux variables qualitatives X et Y, peut être réalisée à l'aide du test du KHI2.

Sous l'hypothèse (H0) : *Les variables X et Y sont indépendantes.*

Contre alternative (H1) : *Les variables X et Y sont liées.*

On définit la statistique du KHI2 comme suit (équation 1) :

$$\alpha = n \sum_{i=1}^n \sum_{j=1}^p \frac{(f_{ij} - f_{i.}f_{.j})^2}{(f_{i.}f_{.j})^2} \quad (1)$$

Où

- f_{ij} désigne la proportion d'individus ayant simultanément la modalité i de la variable X et la modalité j de la variable Y.
- $f_{i.}$ Désigne la proportion d'individus ayant choisi la modalité i de la variable X par exemple.

Sous hypothèse d'indépendance (H0) $f_{ij}=f_{i.}f_{.j}$. Si la statistique de KHI2 (α) est grande par rapport au quantile d'ordre 0.95 d'une variable aléatoire suivant une loi de KHI2 à $(n-1)(p-1)$ degrés de libertés H0 est rejeté. Autrement dit, pour seuil de 0.05 H0 est rejeté si $P(\chi^2_{(n-1)(p-1)} > \alpha) < 0.05$.

b. Test de comparaison des moyennes de wilcoxon

Ce test à l'avantage de tenir compte des échantillons de loi non normale et de faibles effectifs. Le second avantage de ce test est que les valeurs extrêmes ont peu d'influences sur les résultats du test. Il se présente comme suit :

Pour deux échantillons Z1 et Z2, on définit un nouvel échantillon Z représentant l'union de Z1 et Z2. En considérant le rang des données lorsqu'elles sont rangées par ordre croissant dans Z, si les individus des groupes Z1 et Z2 proviennent de la même population, les rangs des individus du groupe Z1 sont tirés au hasard parmi les $n+m$ premiers entiers. Si les moyennes des rangs deux échantillons ne sont pas égales les rangs du premier groupe auront tendance à être trop grands ou trop petits. La statistique utilisée est donc la somme des rangs. Évidemment le raisonnement est symétrique sur les deux groupes. On raisonne sur le premier groupe. On appelle m l'effectif du premier groupe et n l'effectif total. Dans l'espace $\binom{n}{m}$, si $m \geq 10$ et $n - m \geq 10$, la somme des rangs appelée statistique de Wilcoxon suit approximativement une loi normale de moyenne et de variance :

$$E(SR) = \frac{m(n+1)}{2} \quad V(SR) = \frac{m(n-m)(n+1)}{12} \quad (2)$$

On utilisera pour les tests la statistique SR^* suivante :

$$E(SR^*) = \frac{m(n-m)}{2} \quad V(SR^*) = \frac{m(n-m)(n+1)}{12} \quad (3)$$

Finalement pour $T1 = \frac{SR^* - E(SR^*)}{V(SR^*)}$, soit X une variable aléatoire de loi normale centrée et réduite, si $2 * P(X < T1) < 0.05$, on rejette l'idée que les deux moyennes sont égales : la moyenne du premier échantillon est significativement différente de la moyenne du second échantillon.

c. Analyse factorielle Multiples des Correspondances (ACM)

L'analyse factorielle des correspondances multiples permet d'analyser un tableau d'individus \times variables lorsque les variables sont qualitatives. On construit tout d'abord un tableau disjonctif complet, ce tableau est obtenu en remplaçant dans le tableau d'origine chaque variable par l'ensemble des variables indicatrices des différentes modalités de cette variable. Soit Z ce tableau. L'élément Z_{ik} de ce tableau vaut **1** si l'individu i possède la modalité k et **0** sinon, il est de dimension $(n \times K)$ avec $K = \sum_1^J K_j$ où K_j désigne le nombre de modalités associées à la variable j et n désigne le nombre total d'individus (567). L'analyse factorielle des correspondances multiples consiste donc à effectuer une analyse factorielle des correspondances sur ce tableau disjonctif complet. Il permet donc de décrire de vastes tableaux, dont les fichiers d'enquêtes socio-économiques constituent des exemples privilégiés. L'ACM propose un premier regroupement assez « fin » des modalités proches au sens de la métrique du KHI2. La proximité entre les modalités des différentes variables nominales se traduit par une ressemblance des individus qui les auront choisies. [cf. *statistique exploratoire multidimensionnelle, visualisation et inférences fouilles des données* Ludovic Lebart, Marie Piron & Alain Morineau page 132-246]

d. Description d'une sous population par une variable qualitative (valeur-test)

Pour chaque modalité de la typologie (j) et pour chaque composante factorielle (notée X), on calcule la valeur test définie par l'équation (4):

$$valeur\ test = \frac{\bar{X}_k - \bar{x}}{\sqrt{\frac{s^2}{I_k} \left(\frac{I - I_k}{I - 1} \right)}} \quad (4)$$

Avec \bar{x}_k est la moyenne de la variable X pour les individus de la modalité k , \bar{x} représente la moyenne de X sur l'ensemble de l'échantillon, et I_k le nombre d'individus ayant pris la modalité k . Cette valeur permet de tester l'hypothèse nulle:

Les valeurs de x pour les individus prenant la modalité k sont tirées au hasard parmi l'ensemble des valeurs possibles de X .

On considère \bar{X}_k , la variable aléatoire représentant la moyenne des individus possédant la modalité k . Son espérance et sa variance sont définies respectivement par \bar{x} et $\frac{s^2}{I_k} \left(\frac{I - I_k}{I - 1} \right)$.

La valeur test peut être vue comme l'écart normalisé entre la moyenne des individus possédant la modalité k et celle de l'échantillon pour l'axe X . Lorsque celle-ci est supérieure en valeur absolue à 2, la modalité k sera bien représentée sur l'axe X . L'intérêt de la probabilité critique est qu'elle fournit une indication quant à la « significativité » d'un écart. Il est donc possible de trier les variables quantitatives par valeurs-tests décroissantes en se limitant aux probabilités critiques inférieures à 5%.

e. Classification Ascendante Hiérarchique (CAH) sur les composantes principales issues de l'ACM

La CAH permet de regrouper de façon hiérarchique les individus. Les typologies étant qualitatives, l'ACM effectuée en amont de la classification permet d'utiliser les coordonnées factorielles comme variables quantitatives. Cette méthode nécessite un critère de similarité. Selon ce critère on agrège les deux groupes les plus proches pour former un nouveau groupe. Ce groupe peut à nouveau être agrégé. On procède ainsi de façon itérative jusqu'à n'avoir qu'un seul groupe. Cette méthode produit un arbre qui résume la succession de ces regroupements appelés dendrogramme. Le niveau auquel cet arbre est coupé permet de déterminer le nombre de classes.

Nous avons choisi comme critère de similarité, le critère de Ward. Pour deux classes p (de centre de gravité g_p et de masse m_p) et q (de centre de gravité g_q et de masse m_q), il est défini par la relation(5) :

$$\Delta(p, q) = \frac{m_p m_q}{m_p + m_q} d(g_p, g_q)^2 \quad (5)$$

On choisit de regrouper les classes p et q telle que $\Delta(p, q)$ soit minimum, cela revient à choisir :

- Des classes dont les centres gravités sont proches ($d(g_p, g_q)^2$ petit).
- Des classes d'effectifs faibles ($\frac{m_p m_q}{m_p + m_q}$ petit).

2. Présentation des résultats

2.1 Test de KHI2

Une première analyse des liens entre les trois typologies a été réalisée en effectuant le test de KHI2. Le tableau 1 présente les résultats obtenus.

	α	p.value
Logement-ménage	114.4779	$(8.462)^{-12}$
Logement-habitude	172.6236	$(2.2)^{-16}$
Ménage-habitude	164.3131	$(1.214)^{-14}$

Tableau 1- résultats du test de KHI2 entre les trois typologies.

Les « p.value » du test de KHI2 d'indépendance (voir Tableau 1) entre les trois typologies sont très petites (par rapport au seuil fixé à 0.05). Il existe donc une relation de forte dépendance entre ces typologies. Par ailleurs, à travers les p-values, la typologie habitude semble la plus discriminante dans la caractérisation des liens entre ces trois typologies.

2.2 Analyse Factorielle des Correspondances Multiples (ACM)

L'Analyse Factorielle des Correspondances Multiple (ACM) permet de regrouper dans des plans factoriels, les modalités qui permettent de distinguer au mieux les individus ayant les mêmes caractéristiques en termes de typologies (habitude du ménage, structure du logement et de constitution des ménages).

La force de la liaison entre une composante factorielle Y et les typologies, peut être estimée par le ratio ρ défini par l'équation (6) :

$$\rho = \frac{\sum_1^p n_i (\bar{Y}_i - \bar{Y})^2}{\sum_1^p \sum_1^{p_i} (Y_{ij} - \bar{Y})^2} = \frac{\text{somme des carrés interclasses}}{\text{somme des carrés totale}} \quad (6)$$

La figure 2, présente la projection des trois typologies sur le premier plan factoriel (axe 1, axe 2) à l'aide du ratio ρ .

Cette figure montre que pour un individu quelconque la connaissance des modalités des ménages ou d'habitudes qui lui sont associées, permet de définir sa position dans le premier plan factoriel (axe 1,2). La modalité de logement associée à chaque individu permettra seulement de prévoir sa position sur la première composante factorielle (Axe 1).

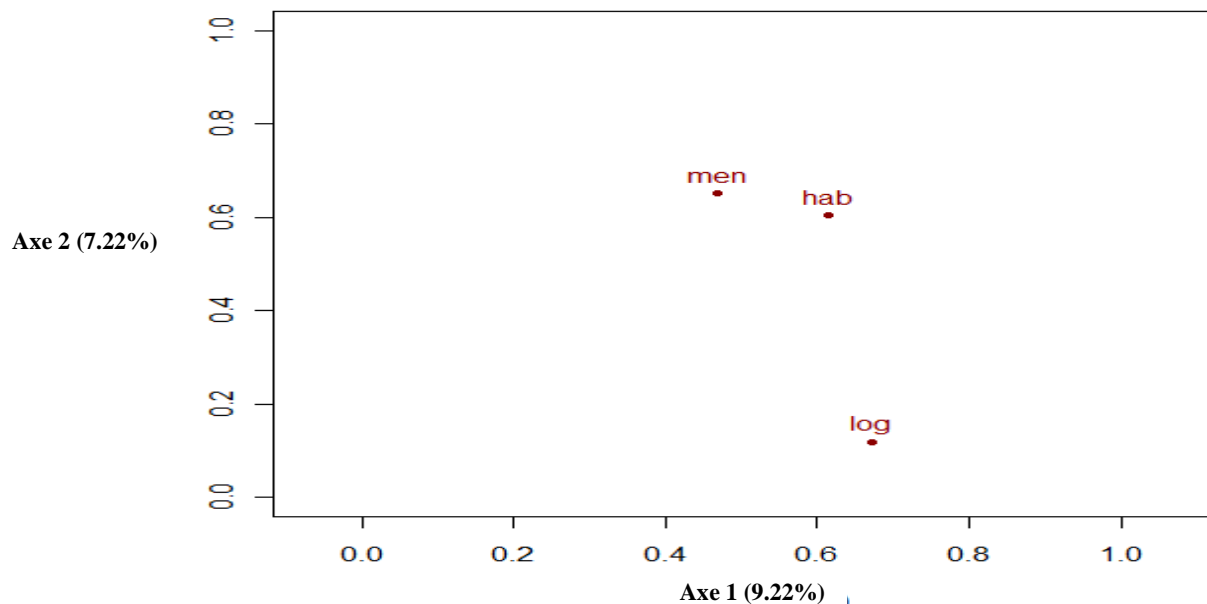


Figure 2 : Rapport de corrélation entre les trois typologies et le premier plan factoriel.

La figure 3, présente l'histogramme des inerties des composantes factorielles obtenues. Comme on peut le voir, les pourcentages d'inertie associés aux premiers axes sont faibles. De ce fait on a été amené à interpréter un plus grand nombre d'axes. Les deux premiers axes représentent 16.44% de l'inertie totale. Cela est une particularité de l'analyse factorielle des correspondances multiples. En effet, les liens ne sont plus linéaires comme en ACP mais multilinéaire. D'où la nécessité d'interpréter un plus grand nombre d'axes.

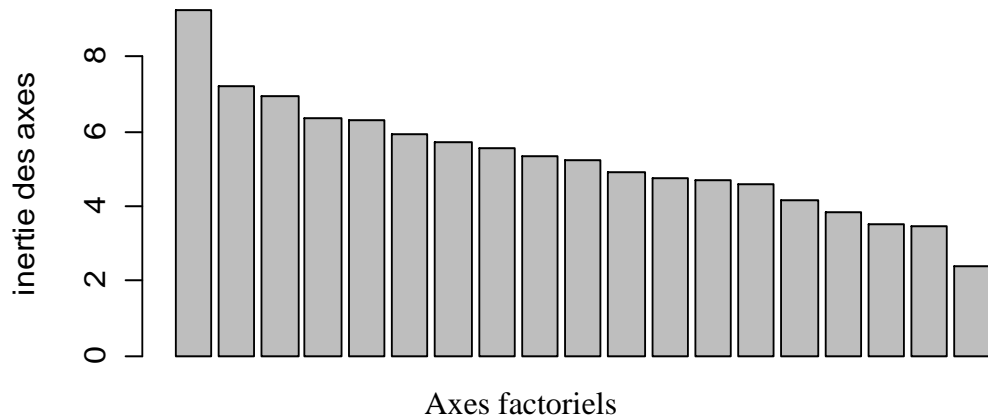


Figure 3 : diagramme des inerties associées aux axes factoriels

La Figure 4, présente le nuage de modalités sur le premier plan factoriel. Sachant que si deux modalités à fortes contributions sont proches, alors elles sont probablement choisies par les mêmes individus. Les combinaisons de modalités à forte contribution observées sur le premier plan factoriel (Figure 4) semblent les suivantes :

- (men_2, log_5, hab_7 ou hab_4), caractérisé par les petits logements collectifs, habité par de jeunes personnes seules ayant des revenus inférieurs à la moyenne et ont une activité moyenne ou faible dans leur logement. Ils bricolent peu.
- (men_5, log_6, hab_4 ou hab_7), caractérisé par de grands logements collectifs avec généralement de grands enfants. Ils ont une activité moyenne ou bien faible dans leur logement, bricolent peu.
- (men_7, hab_9, log_2), caractérisé par des logements individuels « tout en un » moins récents, habités généralement par des retraités ayant une activité assez faible, ils bricolent et jardinent beaucoup.

Nous rappelons que ces combinaisons de modalités sont obtenues uniquement dans le premier plan factoriel (axes 1, 2) représentant moins de 16.44 % de l'inertie du nuage des individus. En théorie, il faut analyser tous les autres plans factoriels afin de définir toutes les combinaisons de modalités résultantes de l'ACM. Cette première étude indique qu'il peut être possible de trouver des corrélations permettant de définir des classes d'habitants appartenant aux mêmes types de logement, d'habitude et de ménage. Cependant, la faible inertie obtenue après projection ne permet pas d'aller loin dans les interprétations. Dans la suite, une classification hiérarchique ascendante (CAH) sur les quinze premiers axes factoriels, représentant une inertie totale de 86.82 % a été effectuée afin, d'identifier la presque totalité des combinaisons de modalités.

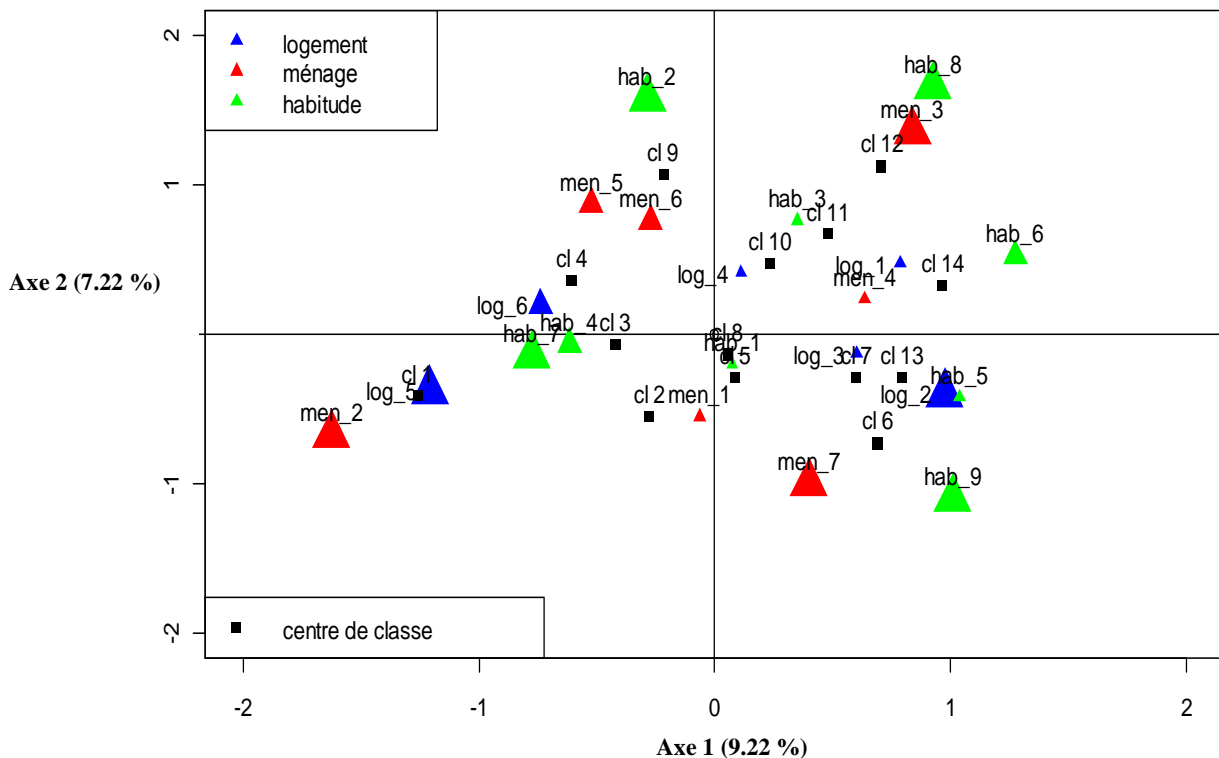


Figure 4 : représentation des modalités sur le plan factoriel 1 et 2.

2.3 Détermination des classes d'individus

L'analyse factorielle des correspondances multiple effectuée en amont permet d'utiliser les coordonnées factorielles comme variables quantitatives. Nous allons utiliser les quinze premiers axes factoriels qui résument 86.82 % de l'inertie totale pour construire la classification hiérarchique ascendante (CAH) en utilisant le critère de Ward.

La figure 5 montre l'arbre hiérarchique ainsi que le diagramme des inerties associé à la classification obtenue.

Cette classification a permis d'établir quatorze classes homogènes, que nous allons décrire par rapport aux variables jugées importantes lors des précédentes études sur les typologies [cf. rapport 1, 2, 3, 4 respectivement sur typologie des logements, typologies de la multi pollution, typologie des habitudes, typologie des ménages]. Le tableau 2 suivant résume les variables considérer pour décrire les classes [cf. annexe 3 pour plus de détails].

typologie	Catégorie	Variables associées
Logement		NIACe1, HSRF, DCA3e1, DGG2be1, CHEM1, FC3
Ménage	Personne	Age, NOCCUA, FC10, FC11, FC11b
	Revenus	REV1, REV3
habitude	Divers	DGG3n, Fumeurn, Activité, QPD2b
	Plantes et animaux	QPV, ANTCPn
	Bricolage et jardinage	TMG6n, TMG7n
	Ordures	QOM1

Tableau 2: les variables importantes définies lors des études précédentes

Par ailleurs, les classes sont également décrites par rapport aux modalités de chaque typologie.

Pour cela, la statistique valeur-test permet d'évaluer l'écart entre les valeurs moyennes relatives à la classe et celle relatives à l'échantillon complet des individus. Cette approche permet de définir le profil type de chaque classe, sachant qu'une modalité d'une variable qualitative est considérée comme caractéristique d'une classe si son abondance dans la classe est significative par rapport à sa présence dans l'échantillon.

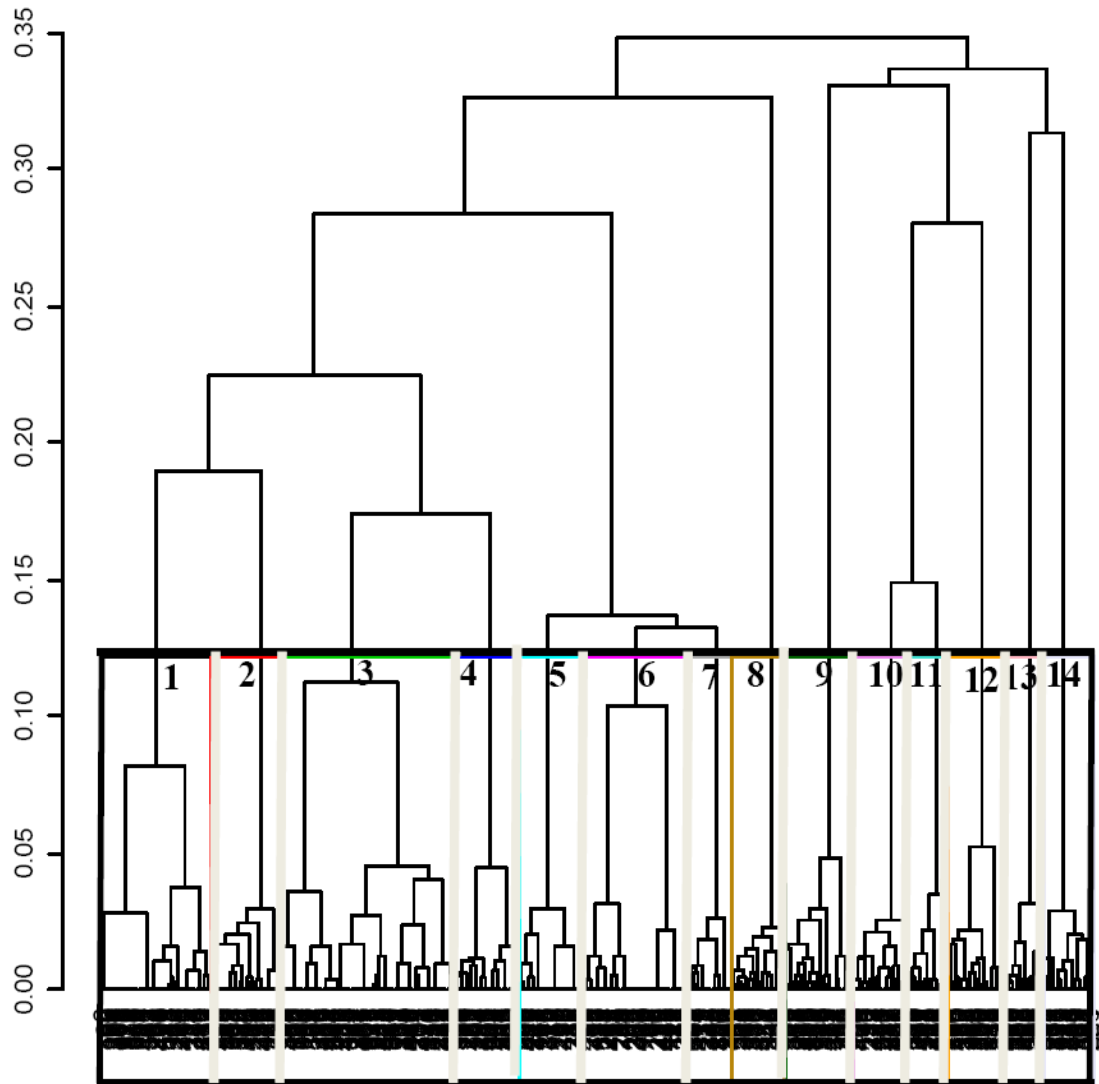


Figure 5 : dendrogramme de la CAH

Les descriptions des trois typologies (logement, ménage, habitude) proposée dans les précédents rapports sont rappelés en annexe 1. Nous donnons maintenant une description des 14 classes en fonction des typologies et des variables statistiquement significatives établies lors des définitions des trois typologies. A chaque fois que le test de la moyenne a été accepté avec une p.value inférieure à 0.05 les pourcentages sont indiqués. Quand la p-value est légèrement supérieure mais que la question semble importante les pourcentages ne sont pas indiqués.

- **Classe1 : 56 logements (9.89%)**

Cette classe est majoritairement constituée de logements collectifs de type 5 (83 % vs 19.22 %) caractérisés par une petite surface (110 m² vs 61 m²) sans caves ni garages attenant, situés généralement en zones urbaines. Les ménages de cette classe se distinguent par la modalité 2 (67.85 % vs 9.87 %). Ils sont caractérisés par des individus jeunes (41.80 ans vs 48.97 ans) vivant généralement seuls (1.67 vs 2.83) avec des revenus très nettement inférieurs à la moyenne (1734 vs 2522).

85.71 % des logements de cette classe sont représentés par la modalité 7 de la typologie habitude. Ils sont caractérisés par une très faible activité (2.64 vs 4.565) dans leur logement, par rapport à l'échantillon complet ils fument moyennement à l'intérieur du logement, ils bricolent rarement ou jamais (91% vs 49%, $p < 0.0001$), ils jardinent rarement ou jamais (94.64% vs 60%, $p < 0.0001$), ils sont moins nombreux à avoir des animaux domestiques et ne sortent pas tous les jours les ordures ménagères (86 % vs 48%). Enfin ils sont plus nombreux à n'avoir aucune plante à l'intérieur du logement (48 % vs 19 %, $p < 0.0001$)

- **Classe 2 : 40 logements (7.07%)**

Elle est majoritairement constituée d'anciens logements collectifs de type 5 (40 % vs 19.22 %) ayant un taux important de meubles en bois agglomérés. Elle contient aussi une proportion non négligeable de maisons individuelles (47.5 %). Tous les logements de cette classe appartiennent à la modalité 1 des ménages. Ce sont généralement les jeunes couples (2.12 vs 2.83) vivants avec des revenus très inférieurs (2114 vs 2522) à la moyenne.

Ils ont généralement peu d'activité (3.80 vs 4.56). Par rapport à l'échantillon complet, ils fument moyennement (36 % vs 32 %), ils jardinent très peu (29 % vs 40 %). Par rapport à la classe 1 ils sont plus nombreux à posséder plus d'une plantes (92 % classe 2 vs 52 % classe 1) et ils utilisent moyennement les encens. Bien que les proportions soient faibles par rapport à échantillon total, ils sont plus nombreux à sortir tous les jours des ordures ménagères à l'extérieur (23% pour la classe 2 vs 14 % pour la classe 1).

- **Classe 3 : 96 logement (16.96%)**

Cette classe est dominée par les petits logements collectifs (71 %) dont une forte proportion de logement de type 6 (55.20 % vs 18 %). Elle contient majoritairement les petits appartements (90 m² vs 110 m²) sans cave ni garage attenant (94 % vs 83 %).

Elle est constituée de couples avec enfants en bas âge ayant des revenus moyens. Dans ces logements, on retrouve des individus ayant peu d'activité (3.73 vs 4.56) dans leurs logements, bricolant et jardinant rarement ou jamais (17 % vs 51 % respectivement 16 % vs 40 %). Ils sont majoritairement non fumeurs à l'intérieur du logement (59 % vs 68 %), ils utilisent très peu d'encens (2.66 vs 4.28) et sortent majoritairement tous les jours les ordures à l'extérieur du logement (74 % vs 52 %).

- **Classe 4 : 36 logement (6.36%)**

Cette classe n'est pas caractérisée par rapport aux logements. Cependant, globalement on y retrouve 58 % de maisons individuelles et de 41.66 % d'appartements dans les grands immeubles avec cave non attenant (92 % vs 84 %) et garage non attenant (25 % vs 19 %).

Elle est constituée majoritairement de couples avec de grands enfants (1.50 enfants vs 0.58 enfants) ayant des revenus moyens.

Les individus de cette classe ont dans l'ensemble une très forte activité dans leur logement (6.26 vs 4.56), ils sont un peu plus fumeurs à l'intérieur du logement par rapport à l'échantillon (42 % vs 32 %), ils bricolent très peu (11 % vs 51 %) et jardinent peu (33 % vs 40 %). Ils utilisent moyennement les encens (4.88 vs 4.28). Enfin ils sont plus nombreux à rentrer une ou plusieurs voitures dans le garage (22 % vs 18 %) et ils ne sortent pas tous les jours les ordures ménagères à l'extérieur du logement. (56 % vs 52 %)

- **Classe 5 : 41 logements (7.24%)**

Elle est constituée des maisons individuelles avec jardin séparés du garage et de l'endroit pour bricoler. Ils se distinguent des autres logements par une forte utilisation de gaz butane et une présence de cheminée (type 4).

Ce sont des personnes âgées (57 % vs 48.97 %), (majoritairement retraités) vivant en couple sans enfant et ayant des revenus nettement inférieurs à la moyenne (2013 vs 2522). Dans l'ensemble, ils ont une faible activité (3.10 vs 5.56). Majoritairement non-fumeurs à l'intérieur des logements (78 % vs 68 %), ils bricolent et jardinent moyennement pour la plupart. Ils utilisent très peu d'encens.

- **Classe 6 : 67 logements (11.82 %)**

Elle est constituée à 97 % par les grandes maisons de type 2 et 3 (47.74 % de type 2 et 50 % de type 3), ce sont majoritairement de grandes maisons individuelles « tout en un » habitées par des couples âgés (58.8 vs 48) majoritairement retraités et ayant des revenus légèrement supérieurs à la moyenne.

Ils sont majoritairement non-fumeurs (84 % vs 68 %) avec une activité faible (3.77 vs 4.5) dans leur logement, ils bricolent et jardinent majoritairement au moins une fois par semaine (84 % vs 51 % respectivement 82 % vs 40 %). Ils ont plus d'une plante à l'intérieur du logement (87 % vs 81 %) et utilisent peu d'encens. Enfin ils sont moins nombreux à sortir tous les jours des ordures à l'extérieur du logement (15 % vs 47.66 %).

- **Classe 7: 24 logement (4.24 %)**

Elle est constituée uniquement des maisons individuelles récentes « tout en un » avec une cheminée et équipées d'un système de chauffage individuel (type 1), ils possèdent majoritairement un garage attenant. Ils sont plus nombreux à rentrer par jours une ou plusieurs voitures dans leur garage.

Les individus de cette classe sont des personnes plus âgées avec des revenus moyens légèrement inférieurs à la moyenne. Ils sont majoritairement non fumeurs à l'intérieur du logement. Ils ont peu d'activité dans leur logement à part une plus importante activité de

bricolage (92 % vs 51 %) et de jardinage (71% vs 60%). Ils ont plus d'une plante à l'intérieur du logement (96 % vs 78 %) ils utilisent peu d'encens et ne sortent pas tous les jours ordures ménagères à l'extérieur du logement.

- **Classe 8 : 31 logements (5.48)**

Cette classe est composée uniquement de logements ayant des habitudes de type 2. Ce sont des couples avec enfants en bas âges et des revenus supérieurs à la moyenne, avec une forte activité dans leurs logements (4.56 vs 5.74). Ils sont tous non fumeurs ils bricolent très peu (17 % vs 51%) de même qu'ils jardinent très peu (23% vs 40%). Ils sont plus nombreux à rentrer des voitures dans le garage. Ils sortent majoritairement tous les jours les ordures ménagères à l'extérieur du logement

- **Classe 9 : 38 logement (6.7 %)**

On retrouve un mélange de logements collectifs et de maisons individuelles. Les logements ne sont pas très caractéristiques pour cette classe.

Les individus de cette classe sont plutôt les jeunes couples avec grands enfants (41 ans vs 48 ans) ayant des revenus inférieurs à la moyenne.

Cette classe est constituée de ménages ayant une activité assez importante dans leur logement (5.50 vs 4.56), majoritairement fumeurs à l'intérieur du logement (87 % vs 34 %) et utilisant peu de parfums d'ambiance. Ils ont aussi des animaux domestiques, ils ont quasiment tous des chiens et ont plus de petits animaux de compagnies (hamsters) que les autres ménages.

- **Classe 10 : 36 logements (7 %)**

Cette classe est caractérisée par une sur-représentés des logements récents de type 2 (32.43 % vs 16.9 %), elle est caractérisée par les maisons individuels qui représentent près de (2/3) de l'effectif de la classe et par les ménages de type 6 (35.13 % vs 18.34 %).

Les individus de cette classe sont des familles nombreuses avec enfants en bas âges et ayant des revenus nettement supérieurs à la moyenne (3003 vs 2522).

Elle est constituée de ménages ayant une activité assez importante (5.40 vs 4.56) dans leur logement, ils sont très majoritairement non-fumeurs à l'intérieur du logement (97 % vs 68 %). Pour la plupart, ils bricolent tous fréquemment (100% vs 51%). Ils sont moins nombreux à posséder un animal domestique et moins nombreux à sortir au moins une fois par jour les ordures ménagères.

- **Classe 11 : 21 logements (3.7 %)**

Cette classe représente les grandes maisons individuelles avec une dominance des maisons de type 1 (38.05 % de la classe vs 10.58 %). Ce sont les grandes maisons individuelles avec garages et caves attenant « tout en un ». Les individus de cette classe sont des familles nombreuses aisées (3451 vs 2522) avec de grands enfants. Ils sont plutôt non-fumeur (76.20 % vs 68 %), avec une activité très fortement supérieure à la moyenne (6.72 vs 4.56). Ils bricolent peu (48% vs 51%) et jardinent peu (48 % vs 40 %). Ils possèdent plus de plantes par contre ils sont moins nombreux à sortir les ordures ménagères tous les jours (33% vs 48%).

- **Classe 12 : 41 logements (7.2%)**

Dans cette classe on retrouve majoritairement les grandes maisons individuelles de type 3 (31.25 % vs 16.40 %), et une sur-représentation des ménages de type 3 (28.12 % vs 16.40 %) séparées du garage et des endroits pour bricoler.

Ce sont des familles nombreuses avec de grands enfants et des revenus légèrement supérieurs à la moyenne. Les individus de cette classe sont les jeunes couples avec enfant en bas et ayant une forte activité dans leurs logements (7.32 vs 4.56), ils bricolent beaucoup plus que la moyenne (91 % vs 51 %). Ils sont majoritairement fumeurs à l'intérieur du logement (94 % vs 68 %) et n'ont pas d'animaux, avec une utilisation plus importante que la moyenne d'encens (9.14 vs 4.2). Ils sont nettement plus nombreux à rentrer des voitures dans leur garage et à sortir les ordures ménagères tous les jours (65 % vs 48 %).

- **Classe 13 : 22 logements (4%)**

Cette classe est caractérisée par une sur-représentation des logements individuels de type 1 (31.81% vs 10.58%) et de type 2 (36.36% vs 16.93%). Ce sont des maisons individuelles anciennes « tout en un », disposant d'un garage et d'un endroit pour bricoler presque systématiquement.

Ce sont des jeunes couples avec des revenus supérieurs à la moyenne (2893 vs 2522). Ils ont une activité supérieure à la moyenne dans leur logement et sont majoritairement non fumeurs (86% vs 68%) à l'intérieur du logement avec une utilisation importante de désodorisants comme d'encens (7.33 vs 4.28%). Ils sont plus nombreux à rentrer des voitures leur garage. Ils bricolent majoritairement au moins une fois par semaine (90% vs 51%) et jardinent majoritairement au moins une fois par semaine (90 % vs 40%). Ils sont plus nombreux à sortir tous les jours les ordures ménagères (95% % vs 47%) et à rentrer une ou plusieurs voitures (50% vs 18%).

- **Classe 14 : 29 logements (5%)**

Cette classe est constituée à 93 % des grandes maisons individuelles tout en un, habitée par les couples aisés (3132 vs 2522) avec grands enfants et ayant une activité assez importante dans les logements (5.87 vs 4.56). Ils sont nombreux à bricoler au moins une fois par semaine (86% vs 51%) et à jardiner (82% vs 40%). Ils possèdent en moyenne beaucoup de plantes et d'animaux domestiques. Ils sont majoritairement non fumeurs à l'intérieur du logement (93% vs 68%) et utilisent beaucoup d'encens (6.67 vs 4.2).

3. Lien avec la pollution

Nous allons à présent, caractériser les quatorze classes obtenues par les groupes de pollution. L'analyse qui suit, utilise les groupes de pollutions (a, b, ab, c, d, cd, e, e2, f, g, h, i, j, k, l, m, fghijklm) déterminés dans l'étude « **analyses descriptives de la multi pollution** ». Le tableau 2, résume ces groupes de pollution (une définition plus large en annexe 2). Rappelons que ces quatorze groupes de pollutions sont scindés en quatre grands types décrit comme suit : les logements fortement « multipollués » (groupe a et b), les logements fortement « monopollués » (de i à m), les logements moyennement « multipollués » (groupe c et d), les logements « faiblement pollués » (e et e2). Cette étude a adopté un point de vue relatif, puisqu'elle positionne, vis-à-vis de la pollution par les COV, les logements les uns par rapport aux autres et qu'aucune notion de niveau de toxicité des polluants n'a été introduite. Les terminologies « fortement », « moyennement » et « faiblement » pollué ne peuvent être interprétées directement d'un point de vue sanitaire. De la même manière l'appellation « monopollués » ne signifie pas qu'il y a qu'un seul COV présent dans les logements en question, mais que l'un des COV y est très prépondérant relativement aux autres en terme de niveaux de concentration comparativement aux autres logements.

Rappelons aussi que, la typologie analyse descriptive de la multipollution a été obtenue grâce aux composés organiques volatiles COV. Ces groupes sont par la suite testés avec les polluants autres que les COV (les allergènes de poussière, le radon, les microparticules).

Nom	Fréquence dans l'échantillon	Définition
Fortement multipollué		
a	6.2%	Dominé par les hydrocarbures aromatiques
b	2.4%	Dominé par les hydrocarbures aliphatiques
Moyennement multipollué		
c	14.7%	Dominé par les hydrocarbures aromatiques
d	10.2%	Dominé par les aldéhydes
Faiblement pollué		
e	30.2%	
e2	13.5%	
Fortement monopollué		
f	3.8%	1,4-di chlorobenzène
g	1.5%	n-un décane
h	1.9%	Styrène
i	2.8%	Trichloréthylène
j	2.6%	1-méthoxy-2-propanol
k	4.9%	Tétrachloroéthylène
l	1.9%	2-butoxyethanol
m	3.6%	Formaldéhyde

Tableau 3 : résumé de la description des groupes de pollution

3.1 Lien avec les groupes de pollution

Il est possible de caractériser les quatorze classes par niveau et type de pollution. L'ensemble des pollutions retenues est donnée dans le tableau 4. On voit apparaître quelques caractéristiques communes qui permettent de regrouper certaines classes.

On retrouve un ensemble de classes qui ne se détachent pas de la pollution moyenne par rapport aux groupes de pollutions établis par AFSSET; il s'agit des classes C3, C5, C14 qui regroupent des individus avec de faibles revenus qui ont peu d'activité dans leur logement (C3 et C5) ou des familles aisées qui à l'opposée ont beaucoup d'activité (C14).

Les classes C4, C6, C8, C10, C13 présentent une proportion plus forte que dans la moyenne de l'échantillon de logements monopollués. On retrouve différents polluants notamment le formaldéhyde, le 2-butoxy éthanol, le styrène et le n-un décane.

Les classes C1, C9, C2, C12, généralement occupées par les jeunes, sont fortement et positivement liées aux logements moyennement multipollués. Les classes C1 et C9 le sont plus particulièrement par les hydrocarbures aromatiques et les classes C2 et C12 par les aldéhydes.

Enfin les classes C7 et C11, constituées majoritairement de maisons « tout en un », sont fortement multipolluées par les hydrocarbures aromatiques et aliphatiques.

Dans ce qui suit les quatorze classes seront regroupées en quatre grandes familles qui diffèrent en termes de niveau et de groupe de pollution.

- La famille noté zéro (0) : elle contient les classes 3, 5, et 14 dont la pollution est semblable à celle de l'échantillon initiale par rapport au groupes de pollution.
- La famille noté un (1) : elle contient les classes 4, 6, 8, 10, 13, liées aux groupe de logements fortement monopollués par rapport à un ou plusieurs polluants.
- La famille noté deux (2) : contient les classes 1, 2, 9, 12 liées aux groupes de logements moyennement multipollués par les hydrocarbures.
- La famille noté trois (3) : contient les classes 7 et 11 liées aux groupes de logements fortement multipollués par les hydrocarbure en général.

Classe des individus issus de la CAH après ACM	Groupes de pollutions associés	Pourcentage du groupe de pollution dans la classe vs l'échantillon	p.value
C1(56 logements) est constituée de personnes jeunes vivant seules dans les logements collectifs (Log5) avec des revenus inférieurs à la moyenne de l'échantillon, ils fument moyennement, ils bricolent et jardinent peu et ont une activité très faible.	c+	24.07% vs 14.7%	0.02
C2(40 logements) est constituée de jeunes couples vivants généralement dans des logements collectifs ou dans les maisons individuelles avec de faibles revenus. Ils sont non fumeurs pour la plupart, et ont dans l'ensemble peu d'activité dans leur logement.	d+	21.05% vs 10.2%	0.009
C3 (93 logements) est constituée de personnes vivants majoritairement dans les grands logements collectifs et ayant une situation financière moyenne. Ils ont peu d'activités dans leur logement, légèrement plus fumeurs (à l'intérieur du logement), ils bricolent et jardinent rarement ou jamais pour la plupart et ils utilisent très peu d'encens.			
C4 (logements) n'est pas caractérisées par les logements. Les individus de cette classe sont des couples avec grands enfants. Ils ont une activité assez importante et fument moyennement (40% vs 32%) et utilisent moyennement les encens dans leurs logements. Ils bricolent très peu.	l+	8.82% vs 1.9%	0.001
C5 (41 logements) est constituée uniquement des retraités vivant dans des maisons individuelles de type 4. Ils ont une activité faible dans leur maison, sont majoritairement non-fumeurs, et ont des revenus inférieurs à la moyenne. Ils bricolent et jardinent moyennement.			
C6 (67 logement) est constituée essentiellement de grandes maisons individuelles (log2 ,47.7%, log3, 50%). Ils ont des revenus légèrement supérieurs à la moyenne. Ils ont une faible activité dans leur logement. Cependant, ils bricolent et jardinent beaucoup. Ils sont moins nombreux à sortir les ordures ménagères tous les jours.	d-, g+,	0% vs 10.2% 6.35% vs. 1.5%	0.002, 0.0003
	cd-	12.70% vs. 24.9%	0.009
C7 (24 logement) est constituée des petites maisons individuelles avec des individus qui ont peu d'activité dans leur maison à part une plus importante activité de bricolage et de jardinage sont moyennement non fumeurs.	a +	28.57% vs. 6.2%	<0.0001
	ab+, cd-	33.33% vs 8.6%, 4.76% vs 24.9%	<0.0001 0.01

C8 (31 logement) est composée uniquement de tous les logements ayant la modalité 2 de la typologie habitude. Ce sont des couples avec enfants en bas âges et des revenus autour de la moyenne (2618 vs 2522), avec une forte activité dans leurs logements (5.74 vs 4.56). Ils sont tous non fumeurs, ils bricolent et jardinent peu.	h+	7.49% vs 1.5%	0.03
C9 (38 logement) est constituée des jeunes en situation financière inférieure à la moyenne ayant une activité assez importante dans le logement. Ils sont majoritairement fumeurs, et utilisent peu d'encens. Ils ont quasiment tous des animaux domestiques. Ils bricolent moyennement et jardinent beaucoup.	c+	25% vs 14.7%	0.03
C10 (36 logement) est caractérisée par les couples avec enfants en bas âges vivant dans des maisons individuelles « tout en un ». Ils ont une activité assez forte dans leur logement, et sont non fumeurs. Ils bricolent beaucoup à l'intérieur du logement.	l+	6.66% vs 1.9%	0.020
C11 (logement) représente les personnes aisées vivant des grandes maisons individuelles avec une forte activité. Ils sont majoritairement non fumeurs, bricolent peu (39% vs 51%) et jardinent peu (43% vs 40%).	a+	21.87% vs 6.2%	0.002
	ab+	21.05% vs 8.6%	0.02
C12 (41 Logements) est constituée des grandes maisons individuelles habitées généralement par les familles nombreuses avec grands enfants ayant des revenus légèrement supérieurs à la moyenne avec de grands enfants. Ils bricolent plus que la moyenne avec une forte activité (dans leurs logements) et utilisent beaucoup d'encens.	d+	21.87% vs 10.2%	0.0118
C13 (22 logements) est constituée des jeunes couples avec enfants en bas âges, ayant une activité forte dans leur logement. Ils bricolent beaucoup, et sont majoritairement non fumeurs, avec une importante utilisation de désodorisants.	m+	15% vs 3.6%	0.002
C14 (29 logements) est caractérisée par logement individuels occupés généralement par des couples aisés avec grands enfants. Ils ont une activité assez importante dans leurs logements, majoritairement non fumeurs, bricolent et utilisent beaucoup d'encens.			

Tableau 4 : Caractéristique des 14 classes avec les groupes de pollutions

3.2 Lien avec les polluants prioritaires

Parmi les polluants mesurés au cours de la campagne logement seuls les polluants prioritaires établis par l'OQAI sont retenus pour l'étude statistique effectuée ici. Parmi les paramètres dits de confort (température, humidité, CO₂), seule l'humidité relative a été prise en compte (valeur moyenne sur la semaine), en raison de ses liens à priori avec la présence d'acariens (tableau 6 pour les COV et autres paramètres). Les mesures prises en compte sont celles effectuées dans la chambre pour les COV, les allergènes de poussières, l'humidité et celles effectuées dans le séjour pour les autres paramètres. Pour le radon, deux paramètres ont été considérés, l'un correspondant aux mesures réalisées dans la chambre et l'autre, dans le séjour. Enfin, une table de données de 567 lignes (une ligne correspondant à un logement) et 24 colonnes (une colonne correspondant à un paramètre) a été constituée. Au sein de la base de données élaborée de 567 logements pour 24 paramètres, 1411 valeurs manquantes ont été observées, ce qui représente 10% des données de cette table. Ces valeurs n'ont pas été complétées.

Tous les paramètres sont concernés dans des quantités allant de 17 au minimum pour le formaldéhyde à 277 au maximum pour les PM et de 104 pour le radon à 129 pour les allergènes de poussières. Quatre cent vingt logements (soit 74 %) admettent au moins une valeur manquante sur un paramètre et pour 28 d'entre eux, il manque pratiquement l'ensemble des mesures des COV. Ainsi, seuls 147 logements disposent des données complètes sur les paramètres retenus, mais 532 logements, soit 94 % de l'ensemble des logements enquêtés, disposent des données complètes pour les quatorze COV prioritaires. La base de données constituée de 532 logements pour quatorze COV prioritaires, les allergènes de poussières, le radon mesuré dans la chambre et dans le séjour a donc été extraite de la base précédente. La représentation en Boxplot des polluants permet de voir et de gommer l'influence des valeurs extrêmes. En effet pour chaque représentation Boxplot (cf. figure 5) des polluants, les valeurs des concentrations dites extrêmes sont celles qui sont supérieures à la valeur adjacente de la moustache supérieure. C'est à dire pour chaque polluant, les concentrations supérieures à $X_{max} = Q3 + 1.5 * (Q3 - Q1)$ où Q1 et Q3 représentent respectivement les quantiles d'ordre 25 % et 75 % sont remplacées par X_{max}.

Cela a l'avantage d'atténuer l'influence des valeurs extrêmes dans le calcul de la moyenne.

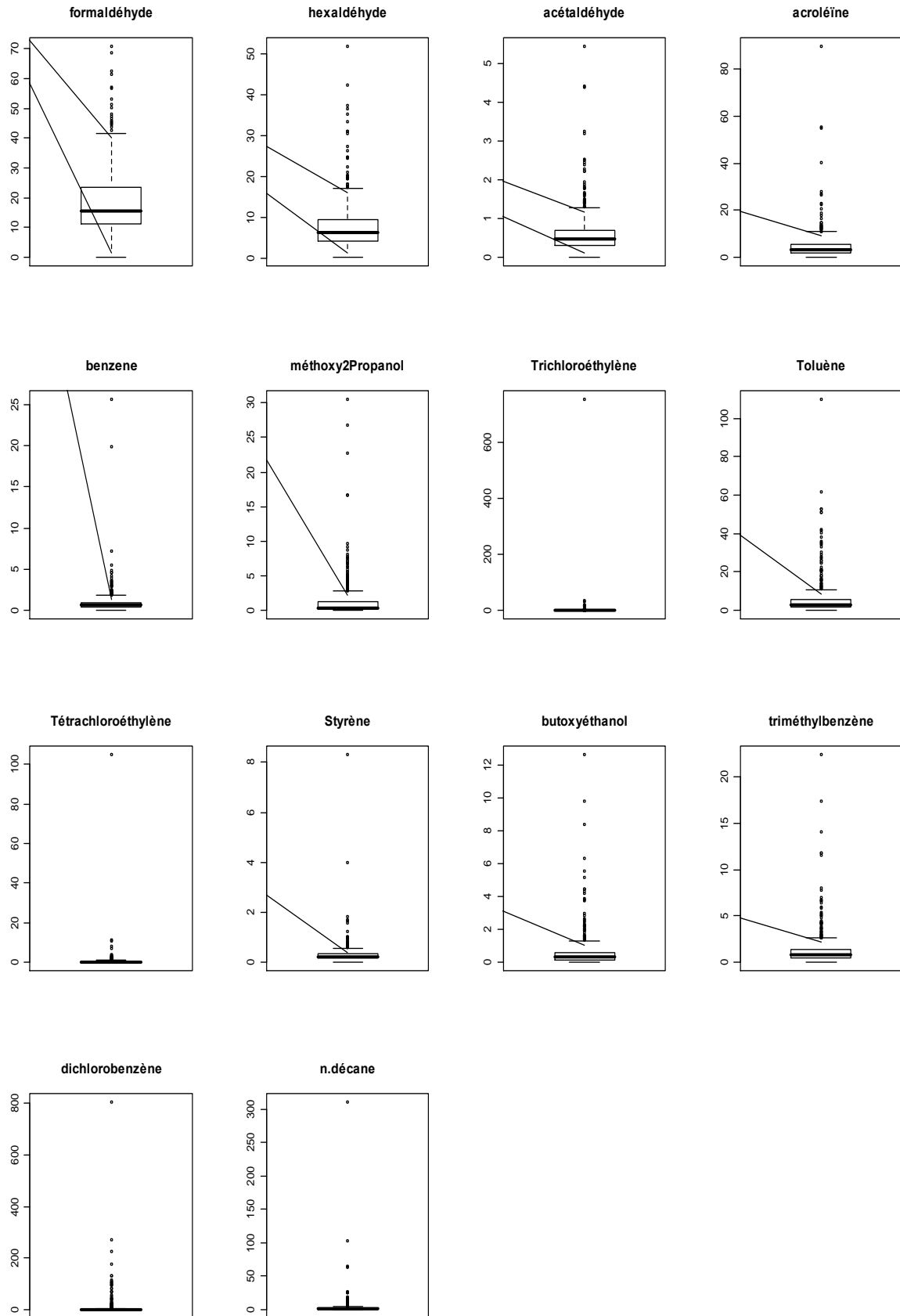


Figure 5: représentation box plot des 14 polluants

Substance	codes	médiane	moyenne	max	Xmax	NA
Formaldéhyde	a21	15.55	17.90	70	60.21	15
Acétaldéhyde	a22	6.18	7.50	51	25	15
Acroléine	a23	0.47	0.58	5.45	1.85	15
Hexaldéhyde	a24	3.22	4.70	89.45	16.07	15
Benzène	c41	0.63	0.90	25.59	2.71	15
Toluène	c44	3.06	5.66	109.80	16.41	15
124-triméthylbenzène	c52	0.80	1.32	22.38	3.90	15
Styrène	c49	0.22	0.29	8.30	0.82	26
n-décane	c54	0.92	3.08	310	6.54	26
Trichloroéthylène	c43	0.19	1.93	751.7	0.97	26
Tétrachloroéthylène	c45	0.21	0.58	104.8	1.29	26
1,4-dichlorobenzène	c53	0.67	8.44	804.8	6.42	26
1-méthoxy-2-Propanol	c42	0.25	1.21	30.46	4.27	15
2-butoxy éthanol	c51	0.31	0.55	12.60	1.97	26
Derf1	x11	1.6	0.558	608	50	129
Derp1	x12	1.6	7.15	129	29.89	129
Radon82sej	X82sej	37	69.73	2161	186.5	109
Radon82chb	X82chb	34	62.79	1215	217	103

Tableau 5 : table descriptive des COV et autres polluants prioritaire établis par L'OQAI

Sont données en rouge les aldéhydes, en vert les hydrocarbures aromatiques, en violet les hydrocarbures halogénés, en gris les éthers de glycol, en bleu les hydrocarbures aliphatiques et en noir les polluant autre que les COV à l'occurrence les allergènes de poussières et d'acariens et le radon mesurer dans la chambre et le séjour.

Aux polluants du tableau 5 sont associées les variables d'humidité que sont :

- HumiPPAL, l'humidité dans la pièce principale : 1 si présence d'humidité dans la pièce principale, 2 sinon.
- HumiPSERV, l'humidité dans la pièce de service : 1 si présence d'humidité dans la pièce principale, 2 sinon
- HumiLOG, l'humidité dans le logement : 1 si présence d'humidité dans la pièce Principal, 2 si présence d'humidité dans la pièce de service, 3 si présence d'humidité dans la pièce de principale ou de service, 4 si absence d'humidité dans la pièce principale et de service.

Le test wilcoxon basé sur les rangs [cf. section 1-2] a permis d'identifier en plus des COV d'autres paramètres significativement associés aux 14 classes.

Si l'on considère les polluants séparément, on voit apparaitre des multipollutions qui sont différentes de celles de l'AFSSET mais qui semblent importantes.

La famille 0 constituée des classes liées aux logements dont la pollution par rapport aux groupes est semblable à celui de l'échantillon ne présente pas de pollutions spécifiques par rapport aux polluants prioritaires. Ce qui confirme le classement précédent avec les groupes de pollutions (pour les classe C3 et C5). Par ailleurs on voit apparaître une pollution spécifique dans la classe C14 qui peut être due à la différence de revenus bien supérieure dans la classe 14. Les logements de la classe C14 sont occupés par les personnes qui bricolent et jardinent plus cela peu expliquer la forte mono-pollution au tétrachloroéthylène à l'hexaldéhyde et au 2-butoxyéthanol.

La famille 1 caractérisée des familles avec enfants présentant de fortes mono-pollutions est liée positivement et significativement à un ou plusieurs polluants individuels. Cependant dans la classes 10 la pollution apparaît moins significative ($p=0.08$ pour le 1-métoxy-2propanol).

La famille 2 qui regroupe les logements moyennement multipollués apparaît moins pollué par les polluants individuels.

Le famille 3 qui regroupe les logements fortement multipollués est liée fortement et positivement aux éthers de glycols et aux hydrocarbures aromatiques.

Les tableau 7 ci-dessous montre les résultats du test de la moyenne de wilcoxon entre les classes et les polluants prioritaires.

Classes issues du croisement des typologies	Polluants	Médiane et Moyenne : polluant dans la classe vs échantillon	p.value
C1 (56 logements) est constituée de personnes jeunes vivant seules dans les logements collectifs (Log5) avec des revenus inférieurs à la moyenne de l'échantillon, ils fument moyennement, ils bricolent et jardinent peu et ont une activité très faible.	Formaldéhyde(ppb) 1-métoxy2propanol(ppb)	13.41 vs 16.02, 15.18 vs 17.9 0.24 vs 0.25, 1.14 vs 1.21	0.01 0.019
C2 (40 logement) est constituée de jeunes couples vivants généralement dans des logements collectifs ou dans les maisons individuelles avec de faibles revenus. Ils sont non fumeurs pour la plupart, et ont dans l'ensemble peu d'activité dans leur logement.	2-butoxyethanol (ppb)	0.15 vs 0.31, 0.36 vs 0.51	0.01
C3 (93 logement) est constituée de personnes vivants majoritairement dans les grands logements collectifs et ayant une situation financière moyenne. Ils ont peu d'activités dans leur logement, légèrement plus fumeurs (à l'intérieur du logement), ils bricolent et jardinent rarement ou jamais pour la plupart et ils utilisent très peu d'encens.	Derp1 (µg/g) HumPPAL HumPLOG	2.63 vs 1.63, 10 vs 7.67 36.45% vs 22% 17.70% vs 10.24%	0.004 0.0001 0.004
C4 (36 logements) n'est pas caractérisées par les logements. Les individus de cette classe sont des couples avec grands enfants. Ils ont une activité assez importante et fument moyennement (40% vs 32%) et utilisent moyennement les encens dans leurs logements. Ils bricolent très peu	Derp1(µg/g) Radon82chb(Bq/m ³) acétaldéhyde(ppb) Hexaldéhyde(ppb) 1,2,4triméthylbenzène(ppb) décane(ppb)	3.53vs 3.56, 12.13 vs 7.61 62 vs 34, 79 vs 62 6.91 vs 6.18, 9.80 vs 7.49 4.14 vs 3.22, 5.04 vs 4.06 1.31 vs 0.8, 1.45 vs 1.05 1.46 vs 0.92, 2 vs 1.5	0.009 0.003 0.03 0.004 0.003 0.02
C5 (41 logements) est constituée uniquement des retraités vivant dans des maisons individuelles de type 4. Ils ont une activité faible dans leur maison, sont majoritairement non-fumeurs, et ont des revenus inférieurs à la moyenne. Ils bricolent et jardinent moyennement.			
C6 (67 logements) est constituée essentiellement de grandes maisons individuelles (log2 ,47.7%, log3, 50%). Ils ont des revenus légèrement supérieurs à la moyenne. Ils ont une faible activité dans leur logement. Cependant, ils bricolent et jardinent beaucoup. Ils sont moins nombreux à sortir les ordures ménagères tous les jours.	Derf1(µg/g) radon82chb(Bq/m ³) radon82sej(Bq/m ³) Tétrachloroéthylène(ppb) benzène(ppb) 2butoxyéthanol (ppb)	0.35 vs 1.62, 5.45 vs 7.61 24 vs 34, 52 vs 62.79 23 vs 37, 58.46 vs 69.73 0.30 vs 0.21, 0.37vs 0.27 0.74 vs 0.63, 0.85 vs 0.75 0.45 vs 0.31, 0.52 vs 0.41	0.006 0.003 0.008 0.001 0.02 0.02
C7 (24 logements) est constituée des petites maisons individuelles avec des individus qui ont peu d'activité dans leur maison à part une plus importante activité de bricolage et de jardinage sont moyennement non fumeurs..	1-métoxy-2-propanol (ppb) HumiPSERV	1.02 vs 0.25, 1.1 vs 0.82 8.33% vs 27.70%	0.02 0.015

Tableau 6 : lien entre les classes 1 à 7 et les polluants prioritaires.

C8 (31 logements) est composée uniquement de tous les logements ayant la modalité 2 de la typologie habitude. Ce sont des couples avec enfants en bas âges et des revenus autour de la moyenne (2618 vs 2522), avec une forte activité dans leurs logements (5.74 vs 4.56). Ils sont tous non fumeurs, ils bricolent et jardinent peu.	Tétrachloroéthylène (ppb) derp1(µg/g) radon82chb(Bq/m³) radon82sej(Bq/m³)	0.09 vs 0.21, 0.16 vs 0.58 0.4 vs 1.6, 1.88 vs 16 63 vs 34, 86 vs 62 66 vs 37, 98 vs 69	0.02 0.003 0.004 0.005
C9 (38 logements) est constituée des jeunes en situation financière inférieure à la moyenne ayant une activité assez importante dans le logement. Ils sont majoritairement fumeurs, et utilisent peu d'encens. Ils ont quasiment tous des animaux domestiques. Ils bricolent moyennement et jardinent beaucoup	Radon82sej(Bq/m³)	50 vs 37, 104 vs 69	0.01
C10 (35 logements) est caractérisée par les couples avec enfants en bas âges vivant dans des maisons individuelles « tout en un ». Ils ont une activité assez forte dans leur logement, et sont non fumeurs. Ils bricolent beaucoup à l'intérieur du logement			
C11 (21 logements) représente les personnes aisées vivant des grandes maisons individuelles avec une forte activité. Ils sont majoritairement non fumeurs, bricolent peu (39% vs 51%) et jardinent peu (43% vs 40%).	Radon82chb(Bq/m³) Radon82sej(Bq/m³) Tétrachloroéthylène(ppb) Trichloroéthylène(ppb) 1-métoxy-2-propanol 2-butoxy-éthanol	23 vs 34, 25 vs 62 18 vs 37, 25 vs 69 0.38 vs 0.21, 0.39 vs 0.26 0.23 vs 0.19, 0.32 vs 0.21 0.81 vs 0.25, 1.17 vs 0.83 0.44 vs 0.31, 0.40 vs 0.27	0.005 0.0009 0.003 0.01 0.02 0.02
C12 (32 logements) est constituée des grandes maisons individuelles habitées généralement par les familles nombreuses avec grands enfants ayant des revenus légèrement supérieurs à la moyenne avec de grands enfants. Ils bricolent plus que la moyenne avec une forte activité (dans leurs logements) et utilisent beaucoup d'encens	Hexaldéhyde(ppb) Trichloréthylène(ppb) Styrène(ppb) humiPPAL	2.33 vs 3.22, 3.06 vs 4.06 0.09 vs 0.19, 0.14 vs 0.21 0.18 vs 0.22, 0.20 vs 0.26 6.22% vs 22%	0.01 0.007 0.008 0.012
C13 (22 logement) est constituée des jeunes couples avec enfants en bas âges, ayant une activité forte dans leur logement. Ils bricolent beaucoup, et sont majoritairement non fumeurs, avec une importante utilisation de désodorisants.	1-métoxy-2-propanol (ppb)	0.97 vs 0.25, 1.06 vs 0.82	0.03
C14 (28 logement) est caractérisée par logement individuels occupés généralement par des couples aisés avec grands enfants. Ils ont une activité assez importante dans leurs logements, majoritairement non fumeurs, bricolent et utilisent beaucoup d'encens.	Hexaldéhyde(ppb) Tétrachloroéthylène(ppb) 2-butoxyethanol(ppb) Derf1(µg/g) Radon82chb(Bq/m³) Radon82sej(Bq/m³)	4.84 vs 3.22, 5.07 vs 4.06 0.38 vs 0.21, 0.40 vs 0.27 0.76 vs 0.31, 0.66 vs 0.40 0.25 vs 1.62, 1.32 vs 7.165 18 vs 34, 27 vs 62 20 vs 37, 29 vs 69	0.007 0.009 0 0.001 0.002 0.003

Tableau 7 : lien entre les classes 8 à 14 et les polluants prioritaires.

Enfin, étant donnée le nombre élevé de valeur manquantes pour les PM, nous avons décidé d'étudier le croisement des PM2.5 et PM10 avec les classe à travers la variable fumeur. Le lien entre la proportion de fumeurs à l'intérieur de chaque classe et la pollution dans la classe ne semble pas lié. En effet dans le tableau 9 on ne voit pas de relation évidente entre, la pollution intérieur des logements et la variable FUMEUR.

classe	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
effectifs	56	40	93	36	41	67	24	31	38	36	29	41	22	29
Proportion générale de fumeurs	32% (18%)	48%(19)	55%(51)	58%(21)	29%(12)	19%(13)	41%(10)	0%	100%	23%(8)	43%(12)	100%	23%(5)	25%(7)
Proportion de fumeur à l'intérieur du logement	27% (15)	36%(14)	39%(36)	41%(15)	22%(9)	16%(10)	33%(8)	0%	87%(32)	3%(8)	24%(7)	94%(37)	14%(3)	7%(2)
Pollution par rapport aux groupes de pollution	c+	d+		l+		g+	a+	h+	c+	j+	a+	d+	m+	
Famille	2	2	0	1	0	1	3	1	2	1	3	2	1	0

Tableau 8 : lien entre la pollution intérieur et la variable Fumeur.

Conclusion

Cette étude fait suite aux 3 études qui avaient permis de définir 3 typologies caractérisant les logements, les ménages et les habitudes en utilisant l'ensemble des 125 questions de l'enquête. Elle a permis à travers l'utilisation de méthodes statistiques multivariées la description de 14 classes qui sont décrites d'une part, en croisant les typologies logements, ménages, habitudes et en utilisant les variables caractéristiques qui en sont issues. Ces deux descriptions montrent à travers le croisement des quatorze classes avec les données de la pollution intérieure et en particulier avec la typologie de la multi pollution chimique élaborée par ailleurs qu'il existe une relation significative entre certains types de logements, de ménages et surtout d'habitudes de ménage et la pollution de l'air intérieur dans les logements. En effet les classes regroupant les logements multipollués significativement et positivement par les groupes de pollutions et les polluants prioritaires sont généralement dominées par les maisons individuelles « tout en un » ayant beaucoup d'activités de bricolages et de jardinage. A l'inverse les classes dominées par les logements collectifs, les maisons individuelles séparée du garage et des endroit pour bricoler semble sont généralement moyennement polluées à la fois par les groupes de pollution et les polluants prioritaires. L'enquête ne nous permet pas de comprendre la pollution liée aux fumeurs qui se retrouvent répartis dans toutes les classes. D'autre part les mesures de PM qui permettent d'évaluer la pollution liée aux fumeurs ne sont pas suffisante (problèmes de valeurs manquantes) et ne permettent pas de finaliser l'étude.

Références bibliographiques

H.BÉNARIBI ET C.DUBOUDIN, JUIN 2008, TYPOLOGIES DES LOGEMENTS ENQUETÉS DANS LE CADRE DE LA CAMPAGNE NATIONALE LOGEMENTS : RAPPORT OQAI-AFFSET

C. DUBOUDIN, JUILLET 2008 : ANALYSE DESCRIPTIVE DES POLLUANTS

RUBEN RODRIGUEZ HERRERA, DANIEL SALLES – LE GAC. INITIATION À L'ANALYSE FACTORIELLE DES DONNÉES : FONDEMENT MATHÉMATIQUE : FONDEMENTS MATHÉMATIQUES ET INTERPRÉTATIONS

MICHEL TENENHAUS. STATISTIQUE : MÉTHODES POUR DÉCRIRE, EXPLIQUER ET PRÉVOIRE.DUNOD

LUDOVIC LEBART, MARIE PIRON, ALAIN MORINEAU. STATISTIQUE EXPLORATOIRE MULTIDIMENSIONNELLE : VISUALISATION ET INFÉRENCE EN FOUILLES DE DONNÉES. DUNOD

FRANÇOIS HUSSON, SÉBASTIEN LÊ ET JÉRÔME PAGÈS. ANALYSE DES DONNÉES AVEC R.PRESSE UNIVERSITAIRE DE RENNES

PIERRE-ANDRÉ CORNILLON ET AUTRES. STATISTIQUE AVEC R. PRESSE UNIVERSITAIRE DE RENNE

JEAN-PIERRE NAKACHE, JOSIANE CONFAIS APPROCHE PRAGMATIQUE DE LA CLASSIFICATION, ARBRE DE HIÉRARCHIQUE ET PARTINNEMENTS. EDITION TECHNIP

Webographie

matisse.univ-paris1.fr/leguen/leguen2001b.pdf La boîte à moustaches de TUKEY un outil pour initier à la Statistique
<http://pbil.univ-lyon1.fr/R/enseignement.html>

Annexes

Annexe 1 :

Typologie logements (6 modalités)

logement I correspond aux maisons individuelles récentes « tout en un » plutôt petites, propriété des occupants, avec un jardin. Garage, endroit pour bricoler, présents pratiquement systématiquement, sont donc attenants et communicant à la partie habitée. Ces maisons ne disposent pas de cheminée. Elles sont équipées de système de chauffage.

logement II correspond aux grandes maisons individuelles « tout en un », moins récentes que celles du type I, avec un jardin. Elles disposent d'un garage et d'un endroit pour bricoler attenant et communicant. Une cave communicante est également présente pratiquement dans un cas sur deux. Elles sont équipées de système de chauffage central individuel, ou de chauffages électriques dans un tiers des cas, et d'un système de ventilation mécanique générale simple flux dans un cas sur deux, comme les logements de type I. En revanche, elles s'en différencient par la présence de cheminées, d'appareils à combustion indépendants raccordés à un conduit de fumée, et de meubles en bois massif. Elles sont plutôt situées en milieu rural.

logement III correspond aux grandes maisons individuelles, anciennes, avec un jardin. Elles disposent d'un garage non attenant pour un cas sur deux, d'un endroit pour bricoler non attenant. Le système de chauffage individuel est omniprésent, en plus de la cheminée et des appareils à combustion indépendants raccordés à un conduit de fumée, présents systématiquement. A la différence avec les autres types de logements individuels, le tiers de ces logements ne disposent d'aucun système d'aération. Ces maisons se différencient également par un taux élevé de murs en peinture à la place du papier peint, de parquets en bois massif et de meubles en bois massif

logement IV correspond aux petites maisons individuelles avec un jardin, séparées du garage, et de l'endroit où bricoler. Comme les logements de type plutôt individuel, ils ont généralement un système de chauffage individuel, mais ce qui les distinguent, c'est une utilisation plus importante des combustibles butanes ou propane pour cette installation. On note l'absence de la cheminée. Ils se différencient également par un taux plus élevé de sols en plastique, de menuiserie en PVC et de meubles en bois aggloméré.

logement V correspond aux petits logements collectifs dans de petits immeubles sans caves et sans garages. Ces logements se distinguent par un taux de meubles en bois aggloméré très important, que se soit par rapport à l'échantillon complet ou aux logements 5 de type plutôt collectif. Ils se différencient des logements du type collectifs par la présence du système de chauffage individuel (1/3). Ces logements se situent dans des zones urbaines de grandes tailles, l'unité urbaine de Paris et des unités urbaines de tailles comprises entre 0,2 et 2 millions d'habitants.

logement VI correspond aux grands logements collectifs dans de grands immeubles en béton, avec des caves non attenantes. Ces logements sont en général chauffés au moyen de systèmes de chauffage collectif. Ils se localisent généralement dans les grandes unités urbaines.

Typologies des ménages (7 modalités)

ménage 1 est constituée de ménages tous en couple, relativement précaires avec peu d'enfants. Elle inclut une proportion non négligeable de retraités.
ménage 2 est majoritairement constituée de personnes seules relativement jeunes qui ont des revenus inférieurs à la moyenne.
ménage 3 est constituée de familles nombreuses, aisées et ayant des grands enfants.
ménage 4 contient beaucoup de ménages en couple plutôt aisés avec un unique enfant de plus de 10 ans.
ménage 5 structure les ménages avec de grands enfants. Elle inclut l'essentiel des familles monoparentales de l'échantillon.
ménage 6 regroupe, en moyenne, les plus jeunes ménages qui sont en couple avec des enfants en bas âge.
Ménage 7 est constituée en majorité des retraités. Elle regroupe l'essentiel des retraités de l'échantillon

Typologie d'habitude (9 modalités)

habitude 1 est constituée de ménages qui entretiennent beaucoup leurs logements, possèdent des animaux domestiques (surtout des chiens) mais ne fument pas.
habitude 2 est constituée de ménages ayant une activité assez importante dans leur logement, majoritairement fumeurs et utilisant des parfums d'ambiance. Ils ont aussi des animaux domestiques.
habitude 3 est constituée de ménages ayant une activité assez importante dans leur logement, majoritairement non-fumeurs. Pour la plupart, ils bricolent fréquemment. Ils sont moins nombreux à posséder un animal domestique.
habitude 4 est constituée de ménages ayant une activité moyenne dans leur logement, ils sont un peu plus nombreux à être fumeurs. Ils sont peu nombreux à avoir des animaux domestiques, bricolent et jardinent peu pour la plupart.
habitude 5 est constitués de ménages ayant une activité moyenne dans leur logement, ils jardinent et bricolent beaucoup. Ils sont majoritairement non fumeurs et utilisé des désodorisants comme de l'encens
habitude 6 est constituée de ménages ayant une activité assez importante dans leur logement, sont nombreux à bricoler, possèdent en moyenne beaucoup de plante et des animaux domestiques. Ils sont majoritairement non fumeurs, et utilisent beaucoup de désodorisant comme l'encens
habitude 7 est constitués de ménages ayant une activité assez faible dans leur logement, ils bricolent et jardinent peu. Ils ont peu de plantes, sont moins nombreux à avoir des animaux. Ils sont majoritairement non fumeurs
habitude 8 est constituée de ménages ayant une forte activité dans leur logement à tous les niveaux (cosmétiques et nettoyages). Ils bricolent beaucoup, possèdent des animaux domestiques et utilisent beaucoup d'encens
habitude 9 est constituée de ménages ayant une activité assez faible dans leur logement. Ils bricolent et jardinent beaucoup. Ils sont majoritairement non-fumeurs et utilisent peu d'encens

Description des quatre grands groupes de pollutions

Groupe 1 : il contient les logements les plus fortement pollués par plusieurs COV simultanément représentant 8.5% de l'échantillon analysé (45 logements sur 532). Ils se caractérisent par un niveau de concentration 2 à 20 fois supérieur à celui de l'échantillon. Deux groupes peuvent être distingués l'un plutôt pollué par les hydrocarbures aromatiques (a) l'autre plutôt pollué par hydrocarbures aliphatiques (b).

Groupe 2 : il est constitué des logements moyennement pollués par plusieurs COV simultanément représentant 25% de l'échantillon analysé. Ils se caractérisent par un niveau de concentration médian compris entre 1.5 et 2.5 fois supérieur à celui de l'échantillon. On distingue de groupes : l'un concerné par les hydrocarbures aromatiques (c) et l'autre par les aldéhydes(d).

Groupe 3 : il contient Les logements faiblement pollués, ils représentent 44% de l'échantillon analysée soit (233 logements sur 532). Ils se caractérisent par des niveaux de concentration médians inférieurs ou égaux à ceux de l'échantillon et se divisent en deux groupes e et e2

Groupe 4 : il contient les logements fortement monopollué, ils représentent environ 23% de l'échantillon. Avec des niveaux de concentration médians entre 5 et 400 fois supérieurs à ceux de l'échantillon complet pour les COV. Huit sous groupes peuvent être distingués correspondant à 8 COV

Annexe 2 :

Description des classes par rapport aux groupes de pollutions. Le tableau ci-dessous si lit comme sui : La colonne Cla/Mod donne pour les modalités, la proportion exacte d'individus de la classe dans la modalité alors que la colonne Mod/cia donne la proportion de la modalité dans la classe.

la classe 1 est caractérisée par une surreprésentation des modalités 2 des ménages, 5 des logements et 7 des habitudes. 67.85% (Mod/cia) des individus de la classe sont de la modalité 2 des ménages contre 9.87% dans échantillon. 83.92% (Mod/cia) des logements de la classe sont de type 5 contre 19.22 % dans l'échantillon. Et 85.71 % (Mod/cia) des habitudes sont de type 7 contre 27.33% dans l'échantillon. Elle se distingue également par une sous-représentation des maisons individuelles. V.test négligeable.

Classe 1	Cla/Mod	Mod/Cla	Global	p,value	v,test
men_2	67,86	67,86	9,89	0	11,65
log_5	43,52	83,93	19,08	0	11,2
hab_7	30,97	85,71	27,39	0	9,57
log_1	1,67	1,79	10,6	0,02	-2,28
men_7	4,35	10,71	24,38	0,01	-2,49
men_4	2,94	5,36	18,02	0,01	-2,65
log_3	1,08	1,79	16,43	0	-3,41
log_4	0,97	1,79	18,2	0	-3,71

classe 2	Cla/Mod	Mod/Cla	Global	p,value	v,test
men_1	60	100	11,48	0	13,73
hab_7	12,26	48,72	27,39	0,01	2,79
log_5	13,89	38,46	19,08	0,01	2,78

classe 3	Cla/Mod	Mod/Cla	Global	p,value	v,test
hab_4	62,3	79,17	21,55	0	13,68
log_6	50	55,21	18,73	0	8,98
men_6	28,85	31,25	18,37	0	3,27
men_7	24,64	35,42	24,38	0,01	2,56
men_4	24,51	26,04	18,02	0,04	2,04
log_2	6,25	6,25	16,96	0	-3,16
hab_9	1	1,04	17,67	0	-5,42

classe 4	Cla/Mod	Mod/Cla	Global	p,value	v.test
men_5	60	100	10.6	0	13.39
hab_4	14.75	50	21.55	0	3.75
hab_7	10.97	47.22	27.39	0.01	2.47
hab_9	1	2.78	17.67	0.01	-2.48

classe 5	Cla/Mod	Mod/Cla	Global	p,value	v,test
log_4	39,81	100	18,2	0	12,18
men_7	18,12	60,98	24,38	0	5,05
hab_9	16	39,02	17,67	0	3,22
hab_7	13,55	51,22	27,39	0	3,21
classe 6	Cla/Mod	Mod/Cla	Global	p,value	v,test
hab=hab_9	50	74,63	17,67	0	11,14
log=log_3	36,56	50,75	16,43	0	6,96
log=log_2	33,33	47,76	16,96	0	6,23
men=men_7	26,09	53,73	24,38	0	5,43
men=men_5	1,67	1,49	10,6	0,01	-2,71
log=log_5	0,93	1,49	19,08	0	-4,42
classe 7	Cla/Mod	Mod/Cla	Global	p.value	v.test
log_1	40	100	10.6	0	10.53
hab_9	17	70.83	17.67	0	5.71
classe 8	Cla/Mod	Mod/Cla	Global	p.value	v.test
hab=hab_1	100	100	5.48	0	15.1
classe 9	Cla/Mod	Mod/Cla	Global	p,value	v,test
hab_2	100	100	6,71	0	16,3
men_5	16,67	26,32	10,6	0,01	2,68
men_7	0,72	2,63	24,38	0	-3,5
classe 10	Cla/Mod	Mod/Cla	Global	p.value	v.test
hab_3	94.59	100	6.54	0	15.39
men_6	12.5	37.14	18.37	0.01	2.55
log_2	12.5	34.29	16.96	0.02	2.41
classe 11	Cla/Mod	Mod/Cla	Global	p,value	v,test
men_3	51,22	100	7,24	0	10,75
log_1	13,33	38,1	10,6	0	3,2
men_4	0	0	18,02	0,03	-2,19
men_6	0	0	18,37	0,03	-2,23
men_7	0	0	24,38	0	-2,81
classe 12	Cla/Mod	Mod/Cla	Global	p.value	v.test
hab_8	100	100	5,65	0	15,28
men_3	21,95	28,12	7,24	0	3,56
log_3	10,75	31,25	16,43	0,05	1,97
log_5	0,93	3,12	19,08	0,02	-2,39
men_7	1,45	6,25	24,38	0,01	-2,46
classe 13	Cla/Mod	Mod/Cla	Global	p.value	v.test
hab_5	100	100	3.89	0	13.2
log_1	11.67	31.82	10.6	0.01	2.57
log_2	8.33	36.36	16.96	0.04	2.04

classe 14	Cla/Mod	Mod/Cla	Global	p.value	v.test
hab_6	96.55	100	5.12	0	14.28
men_4	15.69	57.14	18.02	0	4.61
log_6	0.94	3.57	18.73	0.04	-2.05
log_5	0.93	3.57	19.08	0.04	-2.09

Annexe 3 : LES VARIABLES PRISES EN COMPTES POUR DÉCRIRE LES 14 CLASSES

Type	Code question	Libellé question	Réponses possibles	Regroupement
numeric	HSRF	Quelle est la surface totale de votre logement, en dehors de celle des pièces professionnelles et des pièces annexes ? Tenir compte de toutes les pièces situées dans le logement.		
numeric	NIACe1	Ancienneté de l'immeuble ou du logement	date numérique	
factor	CHEM1	Votre logement est-il équipé d'une cheminée ?	1=Oui;2=non	
factor	DCA3	Existence d'une cave	1=oui communicante ; 2=oui non-communicante; 3=non	
factor	DGG2b	Existence d'un garage ?	1=oui attenant et communiquant avec le logement; 2=oui non attenant; 3=non	
factor	FC3	Votre logement est :	1=maison individuelle;2=appart dans immeuble collectif;3=studio dans immeuble collectif; 4= pièce indépendante;5=logement-foyer personnes âgées;6=ferme,bât d'exploitation agricole;7=chambre hôtel;8=construction provisoire;9=logement dans un immeuble à usage professionnel	16;2345789
factor	REV1	Quelle est la ressource principale de votre ménage	1 Salaires ou traitements 2 Revenus d'une activité indépendante 3 Prérétraite, retraite, pensions et rentes diverses 4 Indemnités de chômage / RMI et allocations sociales	0,64 0,05 0,24 0,04 0,03

			5 Revenus des actifs fonciers ou financiers	
factor	STRMEN	structure du ménage	1 personne seule 2 familles monoparentales 3 couple 4 autre ménage sans famille	0,17 0,08 0,74 0,01
factor	DGG3n	Rentrez-vous une ou plusieurs voitures dans ce garage attenant?	1 tous les jours 2 Parfois 3 Rarement ou jamais 4 question pas posée	1 : 0.18 2 : 0.03 3 : 0.04 4 : 0.75
factor	FUMEURn	Des personnes du ménage fument-elles à l'intérieur de votre logement	1 Non, pas de fumeur 2 Il y a des fumeurs mais personne ne fume à l'intérieur du logement 3 Un fumeur 4 Deux fumeurs ou Trois fumeurs et plus	1 : 0.57 2 : 0.11 3 : 0.22 4 : 0.11
factor	QPV	Combien de plantes avez-vous à l'intérieur de votre logement ?	1 aucune plante 2 de 1 à 4 plantes 3 de 5 à 14 plantes 4 15 plantes et plus	1 : 0,19 2 : 0,45 3 : 0,29 4 : 0,07
factor	QOM1	A cette époque de l'année, vous sortez vos ordures ménagères à l'extérieur du logement :	1 tous les jours 2 2 à 3 fois par semaines 3 tous les autres cas	1 : 0,48 2 : 0,46 3 : 0,07
factor	TMG6n	Au cours des 4 dernières semaines, vous avez bricolé :	1 Rarement ou jamais 2 Moins d'1 fois/semaine 3 Une fois par semaine 4 plusieurs fois/semaine 5 Tous les jours	0,49 2 : 0,12 3 : 0,11 4 : 0,18 5 : 0,09
factor	TMG7n	Au cours des 4 dernières semaines, vous avez jardiné :	1 Rarement ou jamais 2 Moins d'1 fois/semaine 3 Une fois par semaine 4 plusieurs fois/semaine 5 Tous les jours	1 : 0,60 2 : 0,11 3 : 0,12 4 : 0,10 5 : 0,06

Id	Description	Nom	Min	Max	Médiane	Moyenne	Quantile (25%)	Quantile (75%)	Ecart-Type
1	Age de la personne (en années)	AGE	18	89	48	48,96	38	59	14,95
2	Revenu du ménage ? (variable quantitative) 5 tranches	REV3	535	7600	2349	2522,1	1299	2899	1500,5
3	Combien y a-t-il de personnes constituant le ménage ?	FC10	1	8	2	2,8307	2	4	0,014
4	Combien d'enfants de moins de 10 ans ?	FC11	0	3	0	0,418	0	1	0,7565
5	Combien d'enfants de plus de 10 ans ?	FC11b	0	3	0	0,5873	0	1	0,9029