

la revue de Modulad

**essai de structuration de la qualité de la vie à
l'aide de techniques d'analyses des données**

Roland Chifflet

**Unité Pharmacologie Clinique 162 avenue Lacassagne et
SCRIPT INSERM Hopital ST-Jean de Dieu
LYON**

Résumé

Pendant longtemps la préoccupation des cliniciens a davantage été de soigner et de maintenir en vie que de préoccuper des effets secondaires des traitements qu'ils donnaient. De nos jours on s'intéresse davantage à la préservation de ce que l'on appelle "la qualité de vie" des patients. La définition de la qualité de vie n'est pas universelle et dépend des disciplines où elle intervient (médecine, sociologie etc...). Sa prise en compte se fait à partir d'un questionnaire préalablement validé. Un problème est alors de dégager des "structures de qualité de vie" à partir des réponses des personnes interrogées. Le but assigné à ce travail est de montrer que des techniques d'Analyse des données à la fois simples et classiques telles que l'analyse en composantes principales, les Nuées Dynamiques, la Classification Hiérarchique, couplées à des statistiques élémentaires permettent de dégager de telles structures. On montre en particulier que les études prenant en compte l'ensemble de la population isolent certaines de ses structures mais pas toutes. Pour les mettre toutes en évidence on est amené à faire l'étude des classes provenant de son éclatement en divers découpages optimaux

mots clés

questionnaire, Qualité de Vie, structures de qualité de vie, structures axiales structures hiérarchiques, classification, partition, analyse en composantes principales, nuées dynamiques

1. INTRODUCTION

1.1 Le contexte de l'étude effectuée

Il s'agissait dans un cadre bien précis (celui de l'étude des données de qualité de vie) de montrer que les techniques d'analyse des données jointes à des méthodes de description statistique élémentaires aident à la description et à la structuration de données. En particulier si les données sont hétérogènes ou de taille trop élevée pour les logiciels utilisés lors de leur traitement, des méthodes de partitionnement du genre "nuées dynamiques" les éclatent en sous groupes homogènes, de "tailles raisonnables" qui pourront alors être étudiés. La démarche suivie s'appuie sur l'Analyse en composantes principales, sur les nuées-dynamiques et sur des heuristiques d'interprétation de classes et de sélection d'une ou plusieurs partitions optimales au sens d'un critère mesurant la cohésion de ses ou de leurs diverses classes. Il nous fallait aussi définir une méthode de présentation des résultats proche des données brutes traitées pour susciter un intérêt à ce travail par des psychiatres.

1.2 Le problème posé

On doit définir décrire et structurer la qualité de vie subjective de patients interrogés à l'aide du questionnaire PQVS [Garin et al 1989]. On appellera structure de qualité de vie tout assemblage cohérent de divers aspects de qualité de vie ou de groupes de patients homogènes par rapport aux items de qualité de vie. L'assemblage se veut cohérent du point de vue logico-mathématique. Par exemple si l'on utilise des distances spécifiques sur les données et si l'on privilégie certains de leurs éléments lors de l'analyse, il faut s'assurer que les résultats obtenus ne dépendent pas uniquement des ces choix. L'assemblage se veut aussi cohérent d'un certain point de vue psychologique. On s'assurera que les résultats trouvés n'heurtent pas "le bon sens". Par exemple on se refusera de qualifier "d'ouverte" une classe d'individus ne s'intéressant qu'à leur fonctionnement psycho-somatique. On devra enfin être prudent dans l'énoncé des résultats en ayant à l'esprit qu'ils dépendent a priori du matériel utilisé, que ce soit du questionnaire PQVS ou de la population enquêtée.

1.3 Les idées directrices de son traitement

Selon que la population étudiée forme un continuum où se décompose en sous groupes distincts deux attitudes sont a priori à envisager. Dans le premier cas on s'intéressera aux zones de forte densité du continuum et dans le second cas on étudiera la décomposition optimale de la population en sous groupes homogènes. Les zones à forte densité sont des sous ensembles des données à forte concentration, ce ne sont a priori pas des sous-classes des classes de la partition optimale. Un exemple géographique intuitif simple peut être pris pour étayer cette affirmation : si on considère le Rhône et le massif central un découpage géographique a priori optimal en termes de distances est le découpage en départements. Nous rappelons que les départements ont été constitués sous la révolution de telle sorte que tout individu d'un département quelconq puisse se rendre au chef-lieu et en revenir à cheval en une journée. Par contre les zones à forte densité (LYON et sa banlieu) ne se répartissent pas de façon homogène sur ce découpage. Elles sont en effet regroupées au sein d'un même département (le Rhône).

Pour éviter d'avoir à déterminer si la population que nous étudions forme un continuum ou non nous utiliserons pour son étude des techniques ne faisant pas appel à l'une ou à l'autre de ces deux hypothèses.

1.4 Quelques considérations à caractère philosophique

La qualité de la vie est abordé par l'étude d'un certain discours des patients. La statistique se veut une science de l'étude d'ensembles de chiffres. Notre travail a du tenir compte de ces deux aspects. Nous avons aussi opté pour les techniques d'analyse des données qui cherchent à préserver leur aspect multidimensionnel et avons donc utilisé une approche "systémique" (Rosnay 1973) des données. Une autre attitude possible aurait été d'avoir une approche "analytique" (Rosnay 1973) de ce traitement par la définition, le calcul et l'interprétation d'indices sur nos données. Nous pouvons justifier le choix de notre démarche en disant que nous considérons que le comportement des enquêtés est global et ne doit pas être défini par un nombre restreint de facteurs. Le choix de cette attitude nous a permis d'obtenir quelques résultats que nous développons dans les paragraphes suivants

2. LES DONNEES

2.1 Les données brutes

Ce sont les réponses au questionnaire PQVS d'échantillons d'individus pris en compte dans l'essai thérapeutique OCAPL. Les questions portaient sur l'importance accordée au changement de certains aspects de la qualité de la vie après administration de traitements. Les aspects étudiés étaient au nombre de 16: Les gestes courants G, les sorties et déplacements à l'extérieur S, la Fatigue F, la digestion D, le sommeil So, la sexualité Sx, l'attention et la concentration A, les relations en général Rg, les activités quotidiennes Aq, la préoccupation par la santé S, l'intérêt porté à ce qui se passe dans le monde IM, les activités en groupe AG, les rencontres avec des amis RA, les relations avec les proches RP, les activités domestiques ou professionnelles AD, les loisirs L. Les modalités de réponses aux diverses questions étaient au nombre de 3: pas important PI, assez important Ai et très important TI. Une étude antérieure avait montré l'existence de 2 groupes d'items. Le premier, désigné sous le vocable fonctionnel comportait, les gestes courants, les sorties et déplacements, la fatigue, la digestion, le sommeil, l'attention, la sexualité, l'attention. Le second désigné sous le terme relationnel comportait les items de relation: Relations en général, relations avec les proches, rencontres avec des amis. Le reste des items est alors mis dans un troisième groupe que nous désignons sous l'appellation: "le reste".

2.2 Le codage des données

Les divers paliers de réponses: "Pas important", "Assez important", "Très important" sont ordonnés de façon naturelle. Nous considérons comme égale la distance entre 2 paliers de réponses consécutifs. Nous avons alors pris comme codage de ces paliers les valeurs 0,1 et 2. Nous avons défini comme distance entre 2 individus quelconques la distance euclidienne classique entre les 2 vecteurs de leurs modalités de réponses codées. Sous ces hypothèses les techniques classiques d'analyse en composantes principales, d'analyse discriminante, de partitionnement par une méthode basée sur la distance euclidienne, sont permises.

2.3 La classification a priori des variables

Les variables ont été classées initialement en 3 groupes (fonctionnel, relationnel, et le reste). Nous avons éclaté "le reste" en 2 sous-groupes: "les activités" et "le temps libre". Tous les items ont pu être placé sans ambiguïté dans l'une de ces classes. Seul l'item "sexualité" pouvait théoriquement être placé dans plusieurs catégories. Les analyses que nous avons faites par la suite ont confirmé cette hypothèse. Elles ont montré que cet item pouvait se placer dans l'une des 3 catégories suivantes "fonctionnel, relationnel, temps libre". Nous l'avions initialement placé dans le "fonctionnel", puis par la suite nous l'avons mis de façon "provocatrice" dans le "temps libre" pour bien montrer que cet item nous posait problème.

3. METHODES ET METHODOLOGIE

3.1 Introduction

Elles consistent en des statistiques élémentaires, des techniques multidimensionnelles et des méthodes d'interprétation de classes ou de détermination du nombre de classes de partitions optimales. Parmi les statistiques élémentaires, citons les tracés d'histogrammes par variable faites sur la population totale ou sur des sous-populations et divers calculs et représentations graphiques. Parmi les calculs citons les évaluations des distances entre les éléments de la population et les centres des classes mises en évidence. Parmi les représentations graphiques citons celles positionnant les moyennes des variables par classe et par item. Parmi les méthodes multidimensionnelles mises en oeuvre citons l'analyse en composantes principales, l'analyse factorielle des correspondances, l'analyse factorielle des correspondances multiples, l'analyse discriminante

3.2 Rappel sur les Nuées Dynamiques

On considère une population d'individus caractérisée par des variables à valeurs réelles. Nous dirons qu'une partition P_i en un nombre i de classes de cette population est optimale relativement à ce nombre i si la somme des inerties de ses classes est minimale. La méthode des Nuées dynamiques est une méthode "noyaux-partitions" Diday et al 1982 visant pour tout entier i à trouver une telle partition P_i . La qualité $W(P_i)$ de cette partition s'exprime comme le rapport de la variance inter-classe par la variance totale de la population.

Un des problèmes qui se pose est de savoir s'il existe une "structure classificatoire" (ou des "structures classificatoires") sous-jacente(s) aux données définie(s) par une(des) partition(s) de la population en classes bien disjointes. Si les nombres de classes de ces partitions "optimales" sont connus on emploiera les nuées dynamiques pour les mettre en évidence. Nous allons donner une heuristique visant à trouver ces nombres.

3.3 Recherche des nombres de classes des partitions "optimales"

On considère les partitions optimales P_1, \dots, P_i en $1, \dots, i$ classes et on désigne par W_1, \dots, W_i leurs qualités respectives. On définit :

le taux d'accroissement θ du critère W ceci pour tout entier $i > 2$ par $\theta_i = (W_i - W_{i-1}) / W_{i-1}$
et pour $i=1$ par $\theta_1 = 0$ et pour $i=2$ par $\theta_2 = (W_2 - W_1) / ((W_2 + W_1) / 2)$. Comme $\theta_1 = 0$ on a :
 $\theta_1 = 2$

la vitesse V de ce taux pour tout entier $i > 1$ par $V_i = (\theta_{i+1} - \theta_{i-1}) / 2$ et on pose $V_1 = 0$.

l'accélération f de ce taux pour tout entier $i > 1$ par $f_i = (V_{i+1} - V_{i-1}) / 2$

On dira que les nombres de classes des partitions "optimales" seront pris parmi les entiers i pour lesquels le signe de f_i est différent de celui de f_{i+1} ou de celui de f_{i-1} , ceci par analogie avec la mécanique où l'on considère comme instants privilégiés dans la trajectoire rectiligne d'un mobile ceux où son accélération change de signe. Parmi ces partitions on retiendra comme partitions "optimales" celles pour lesquelles le signe de l'accélération de leur taux d'accroissement est le signe de f_i et on retiendra la première partition dont le signe de l'accélération de son taux d'accroissement est distinct du signe de f_i . Ceci peut encore s'expliquer par analogie avec la mécanique lorsque les conditions initiales relatives au déplacement d'un mobile caractérisent ce déplacement et que l'on cherche à les retrouver dans son parcours.

Ce n'est bien sûr qu'une heuristique que nous avons établie en fait après avoir examiné un certain nombre de partitions et retenu celles qui pour nous, s'interprétaient bien.

3.4 Interprétation et caractérisation des classes

3.4.1 Interprétation des classes

structure associée à une classe donnée

Pour une classe donnée on dresse le tableau des pourcentages de réponses obtenues pour chacune des modalités (PI, AI, TI) des variables. On détermine un premier seuil s_1 dépassé par le pourcentage des réponses obtenues pour une modalité donnée par environ 1/3 des variables. Ces variables seront dites variables caractéristiques de la classe. On définit alors un seuil s_2 tel que environ 1/3 des variables non encore sélectionnées aient 2 modalités de réponses contigues (à savoir PI-AI ou AI-TI) dont la somme des pourcentages de réponses le dépasse. Ces deux ensembles de variables forment la structure associée à la classe. L'interprétation de la classe consiste à lui donner un qualificatif reflétant sa structure et différent de ceux des autres classes de la partition dont elle fait partie.

exemple d'interprétation d'une classe

Lorsque l'on voit la structure de la classe 5 décrite dans le tableau 2 on peut imaginer (et il faudrait alors le vérifier) qu'elle est constituée de gens pour lesquels le fonctionnel propre: digestion, sommeil, fatigue: que l'on associe ici aux activités quotidiennes et professionnelles garde globalement une certaine importance. Le fonctionnel liés à l'extérieur: gestes courants, sortie, attention, santé, est considéré comme peu important ou n'apparaît pas dans la structure. Dans le relationnel ce sont les relations en général qui dominent et l'emportent notamment sur les rencontres avec les amis. Il y a un désintérêt total pour ce qui se passe dans le monde et les activités en groupe. Par contre la sexualité semble être quelque chose d'important. On peut intuitivement la lier au relationnel en général. On a tendance à qualifier cette classe "intervertie individualiste". Ce qualificatif est cohérent avec ce que l'on vient de dire. Bien sur on ne peut dire que c'est la vérité. Cela doit servir d'hypothèse de base.

3.4.2 rattachement d'une classe à un palier de réponse.

Pour une question et une classe on définit leur palier de réponse associé comme étant le palier de réponse le plus proche de la valeur moyenne des réponses codées à cette question des individus de cette classe. Pour chaque palier de réponse on détermine le nombre de questions dont il est le palier associé. On appelle palier associé à la classe le palier de réponses caractéristique du plus grand nombre de questions. Une classe sera caractérisée par son palier caractéristique s'il est lié à la quasi-totalité des questions.

3.4.3 Conséquence de ces 2 manières de caractériser les classes

Deux classes peuvent différer pour une caractérisation et pas pour l'autre. En particulier 2 classes peuvent être associées au même palier de réponse et néanmoins s'opposer par leurs structures respectives. Ce sera le cas des classes 4 et 5 de la partition optimale en 7 classes que nous avons dégagée et que nous verrons par la suite.

3.5 Stabilité d'une partition

La méthode des nuées dynamiques que nous avons utilisée nous procure pour un nombre k une partition en k classes. Pour une quelconque A de ces classes et pour un quelconque de ses éléments désigné par x , la distance euclidienne classique de x au centre de gravité de A est inférieure à la distance de x au centre de gravité de toute autre classe. La vérification de la stabilité consiste à vérifier cette propriété lorsqu'on remplace les centres de gravité des classes par certains éléments des classes et la distance euclidienne par une autre distance. Cette vérification de la stabilité a le mérite d'être simple: elle permet surtout de déceler des partitions très instables.

3.6 La méthodologie suivie

Pour dégrossir le problème on a commencé par établir des statistiques élémentaires sur la population totale (tracés d'histogrammes, calculs de dispersions) et une analyse globale (ACP). L'analyse des correspondances multiples n'a pu être effectuée du fait de la taille de nos données et de la version du logiciel utilisé qui n'était pas la bibliothèque MODULAD. Nous sommes alors passé aux analyses des partitions optimales obtenus par les nuées dynamiques en utilisant le critère de sélection du nombre de classes défini plus haut. Nous avons cherché un lien d'une part entre ces diverses partitions, et d'autre part entre les résultats que nous en avons dégagés et ceux que nous avons trouvés par l'ACP. Des tests de cohérence des classes étudiées ont été pratiqués. Parmi eux citons la vérification de la stabilité des classes d'une partition étudiée. Cette vérification se fait, tout d'abord, d'une part en remplaçant leurs noyaux assimilés à leurs centres de gravité par leurs éléments les plus proches de ces noyaux et d'autre part en prenant comme distance entre 2 individus la somme des valeurs absolues des différences de leurs modalités codées de réponses. Elle se fait ensuite en vérifiant que les classes sont formés globalement d'individus dont la nouvelle distance à l'élément remplaçant son centre de gravité est minimale.

4. RESULTATS

4.1 Résultats liés aux analyses globales

Les histogrammes (que l'on peut reconstituer grâce au tableau 2) faits sur chacune des variables montrent que les modalités de réponses se répartissent de façon dissymétrique autour de la modalité AI. La modalité AI est la plus sélectionnée, vient ensuite la modalité PI pour les items : "intérêt à ce qui se passe dans le monde", "activités en groupe", "sexualité"; ou la modalité TI pour les autres items. A la vue de ces résultats grossiers on aurait tendance à placer l'item sexualité avec les items "intérêt porté à ce qui se passe dans le monde" et "activité de groupe".

L'Analyse en Composantes Principales fait ressortir 2 axes représentant 45% et 5% d'inertie du nuage de la population (voir figure 1). Le premier axe est un axe de taille. Toutes les variables se trouvent regroupées d'un même côté par rapport à lui et contribuent à sa création. Cet axe peut s'interpréter comme un axe "d'humeur" opposant "les blasés" pour qui rien n'est important aux "enthousiastes" pour qui tout est important. Le deuxième axe quant à lui oppose "le fonctionnel" (incluant ici la sexualité) au reste des items. Il regroupe en fait 2 oppositions : l'opposition fonctionnel-relacionnel et l'opposition soi-non soi dont nous parlerons plus loin.

4.2 Résultats liés aux analyses locales

4.2.1 introduction

Le tableau 1 montre que l'accélération du taux d'accroissement de la variance inter-classe sur la variance totale change de signe au passage de 3 à 4 classes et de 6 à 7 classes. Nous avons décidé, conformément au paragraphe 3.3 de retenir les partitions en 3, 4 et 7 classes pour les étudier. La partition en 7 classes a été baptisée Partition Optimale comme étant la meilleure au sens du rapport de l'inertie inter-classe sur l'inertie totale du nuage.

4.2.2 Phénomènes liés à la partition en 3 classes

Le phénomène humeur est caractéristique de cette partition (voir figure 2). Chaque classe est attachée à l'une des modalités: PI, AI, TI. Toutes les classes sont attachées à des modalités différentes. Ce dernier point est important car il met bien en évidence un phénomène humeur: ce serait différent si 2 classes étaient attachées à la même modalité. De façon grossière on peut dire qu'un individu a tendance à répondre de la même façon pour un grand nombre de questions. Ceci conforte l'interprétation du premier axe de l'ACP.

4.2.3 Phénomènes liés à la partition en 4 classes

On voit (figure 2) se superposer 2 phénomènes : le phénomène humeur avec trois classes attachées chacune à l'une des 3 modalités PI, AI, TI et le phénomène fonctionnel opposé au reste des items qui est visible sur la classe restante. On retrouve les résultats de l'analyse en composantes principales.

4.2.4 Phénomènes liés à la partition en 7 classes

On voit (figures 2 et 3) se superposer 3 phénomènes: le phénomène humeur, l'opposition fonctionnel-relacionnel et l'opposition "Soi non-Soi" marquée par une opposition entre les deux groupes de classes (1 2 5 6 7) et (3 4). Les patients du premier groupe privilégient le fonctionnel au relacionnel dans leur investissement et se désintéressent des activités de groupe et de se qui se passe dans le monde. On retrouve et on complète les résultats donnés par l'ACP.

4.3 Liens entre les partitions optimales en 3 et 7 classes

On a regroupé les 7 classes de la partition optimale par un processus itératif de regroupement de classes 2 à 2. On a commencé par regrouper les 2 classes d'inertie minimale ce qui nous a donné une nouvelle classe. On a considéré cette nouvelle classe et les classes qui n'avaient pas servi au regroupement. On a alors regroupé les 2 classes dont la réunion était d'inertie minimale. On a poursuivi ce processus jusqu'à obtenir 3 classes ((1,2,3) (4,5) (6,7)) que l'on a comparé aux 3 classes (G1, G2, G3) de la partition optimale obtenue préalablement par les "Nuées Dynamiques". On a obtenu le tableau suivant:

classes	(1,2,3)	G1	(1,2,3) G1	(4,5)	G2	(4,5) G2	(6,7)	G3	(6,7) G3
cardinaux	119	88	88	98	135	88	66	60	50

Ce tableau se déduit du tableau de contingence croisant les partitions optimales en 3 et 7 classes, se trouvant après la figure 2.

On a alors identifié la partition ((1,2,3) (4,5) (6,7)) à la partition (G1 G2 G3).

La réunion des classes 4 et 5 est naturelle en effet ce sont des classes centrales attachées à la modalité de réponse AI. Les classes (1,2,3) sont quant à elles associées à la modalité de réponse TI et les classes (6,7) à la modalité de réponse PI.

4.4 Interprétation de la partition optimale

Classe 1 (N=44) sujets mettant l'accent sur le monde extérieur proche

Les items caractérisant cette classe sont :

Les relations (en général, avec des amis, avec des proches) les activités (quotidiennes à la maison ou professionnelles ou de loisirs) ainsi que l'attention et la digestion

Classe 2 (N=42) sujets mettant l'accent sur leur fonctionnement somatique et leur santé

Les items caractérisant cette classe sont

Les gestes courants, les sorties et déplacements, la fatigue, la digestion, le sommeil, l'attention et la santé.

Classe 3 (N=33) sujets conscients de l'importance de leurs vies fonctionnelle et relationnelle

Les sujets de cette classe privilégient les items de vie relationnelle et plus particulièrement les activités, les loisirs et l'attention aux items de vie fonctionnelle comprenant les gestes courants, les sorties et déplacements, la fatigue, la digestion, la santé.

Classe 4 (N=75) classe de la "vie relationnelle ouverte"

Ses sujets sont intéressés davantage par ce qui se passe autour d'eux que par leur propre santé. Cette vie relationnelle ouverte est caractérisée par l'investissement d'un cercle extérieur élargi (les gens en général et les amis: les activités en groupe et le monde extérieur en général: mais aussi les activités quotidiennes, les activités domestiques ou professionnelles et les loisirs.

Classe 5 (N=23) "intervertie ou individualise"

Cette classe s'oppose à la précédente par sa structure. Ses sujets ne s'intéressent et d'ailleurs assez modérément qu'à ce qui concerne leur fonctionnement somatique (gestes courants, déplacements, fatigue, digestion, sexualité); seules les relations en général sont mentionnées (assez distantes sans doute car les relations avec des amis ou des proches n'apparaissent pas). C'est à partir de cette caractérisation qu'on a été tenté de rapprocher l'item sexualité des items du relationnel.

Classe 6 (N=43) classe des désabusés ou classe du "train-train" journalier

Ses sujets se caractérisent par la persistance d'un minimum d'intérêt pour quelques items assurant un minimum de vie en société (activités quotidiennes, relations avec les proches, loisirs) et pour les items les plus généraux de la Santé (fatigue, santé en général) et par un désintérêt total pour le "cercle extérieur ouvert" (intérêt pour le monde et activités en groupe).

Classe 7 (N=22) démission ou fatalisme à l'égard des items du questionnaire PQVS

Ses sujets considèrent globalement que les valeurs qu'ils accordent aux items du questionnaire sont sans importance. Il s'en dégage un refus du relationnel, des activités et de ce qui se passe dans le monde

5. DISCUSSION

5.1 Justification de la sélection des partitions étudiées

Les indices pour lesquels "l'accélération" du critère de qualité des partitions optimales change de signe (voir tableau 1) sont au nombre de 4 à savoir: 3,4 et 6,7. D'autre part il y a un fort accroissement de ce critère dans le passage de la partition optimale en 2 classes à celle en 3 classes. Cet accroissement est constant et peu important dans le passage entre les partitions optimales en 4 et 5 classes et entre les partitions optimales en 5 et 6 classes. L'accroissement de ce critère est faible à partir de la partition optimale en 7 classes.

Nous avons estimé qu'il y avait 3 partitions importantes: les 2 partitions optimales extrêmes en 3 et 7 classes et la partition optimale en 4 classes qui nous a semblé être un démarquage par rapport à la structure définie par la partition optimale en 3 classes. Cette sélection se justifie par ailleurs par l'heuristique que nous avons décrite en 3.3

5.2 Structures dégagées de qualité de vie (figure 4)

5.2.1 Structures axiales

L'analyse en composantes principales nous a montré 2 structures définies par 2 axes factoriels : un axe d'humeur opposant les "enthousiastes" aux "blasés" et un axe fonctionnel opposant la préoccupation par le fonctionnel aux autres préoccupations. Les analyses locales que nous avons menées ont montré l'existence de "phénomènes" humeur, fonctionnel, d'attitude(Soi:Non-Soi), opposition fonctionnel-relacionnel. Nous avons été alors amené à émettre l'hypothèse de l'existence de 3 structures : un axe d'humeur, un axe d'attitude à l'égard de l'environnement(axe Soi:Non-soi) et un axe de comportement(axe fonctionnel-relacionnel).

5.2.2 Structure hiérarchique

On a vu au paragraphe 4.3 que la partition optimale en 3 classes pouvait être considéré comme un regroupement optimal des classes de la partition optimale en 7 classes. On a donc une "hiérarchie naturelle" sur la population étudiée. Cette hiérarchie se compose de cette population, des classes de la partition optimale en 3 classes et de celles de la partition optimale en 7 classes.

5.2.3 Rapports entre ces structures

On peut associer à chacun des paliers de la hiérarchie certains des phénomènes mis à jour. Au palier relatif à la population totale on dira qu'on associe le phénomène "la modalité de réponse choisie le plus souvent est la modalité "assez important" et ceci quel que soit l'item considéré". au palier relatif à la partition optimale en 3 classes, on associe le facteur humeur et au palier relatif à la partition optimale en 7 classes on associe les 3 facteurs d'humeur,d'attitude et de fonctionnement.

5.3 Remarques relatives aux hypothèses initiales

Comme nous l'avons déjà mentionné au paragraphe 4.1 les histogrammes (que l'on peut reconstituer grâce au tableau 2) faits sur chacune des variables montrent que les modalités de réponses se répartissent de façon dissymétrique autour de la modalité AI. La modalité AI est la plus sélectionnée, vient ensuite la modalité PI pour les items ("intérêt à ce qui se passe dans le monde", "activités en groupe", "sexualité") et enfin la modalité TI pour les autres items. On pourrait donc dire que l'écart séparant la modalité "pas-important" de la modalité "assez-important" est supérieur à celui séparant la modalité "assez-important" de la modalité "très-important" pour les items "intérêt à ce qui se passe dans le monde", "activités en groupe", "sexualité" et que ce phénomène est inversé pour les autres items.

Nous avons fait toute notre étude en supposant égaux ces écarts. Il aurait peut être été souhaitable de déterminer des coefficients de pondération pour ces écarts et ceci pour chacun des items.

6. CONCLUSION

6.1 Les résultats obtenus vis à vis du problème posé

- On peut dire que le problème initial avait plusieurs aspects. On désirait
- 1° savoir si la population étudiée était hétérogène ou si elle formait un continuum.
 - 2° dégager des structures de la qualité de la vie en précisant le sens donné au terme structure.
 - 3° tester la fiabilité des résultats obtenus.
 - 4° montrer l'utilité des méthodes d'Analyses des données dans ce genre d'études.

Pour tenter de résoudre le premier point, nous avons repris le critère classique en classification d'inertie inter-classe et à partir des valeurs de ce critère sur des partitions optimales en 2, 3, ..., l classes avons défini une mesure de son amélioration entre sa valeur sur la partition optimale en $i+1$ classe et sa valeur sur la partition optimale en i classes. L'étude de l'accélération de cette mesure a montré que la population ne formait pas un continuum homogène.

Bien que le premier point n'ait pu être définitivement résolu nous avons donné un sens au terme structure de la qualité de la vie en l'associant aux termes axes factoriels et hiérarchie de partitions. Une étude globale des données au moyen de l'ACP et l'étude des partitions optimales que nous avons dégagées nous ont permis de mettre en évidence des structures hiérarchiques et axiales de la qualité de la vie que nous avons reliées entre elles.

Les résultats obtenus ont été testés sur 2 autres échantillons (l'un de 283 individus et l'autre de 119 individus). On y retrouve les structures hiérarchiques et axiales dégagées ainsi qu'une partition optimale extrême en 3 classes. Par contre le nombre de classes de la seconde partition optimale extrême est de 5 pour l'échantillon de 119 éléments et de 6 pour le second.

Les méthodes d'analyse des données ont montré leur intérêt dans ce genre d'étude en permettant notamment la scission d'une population à étudier en sous groupes homogènes, en permettant des représentation visuelles d'aspects saillants des données et en dégageant des structures en leur sein.

6.2 Perspectives

Plusieurs voies de développement s'ouvrent à partir de cette étude. La première consiste dans la poursuite d'analyses descriptives sur d'autres échantillons de données: une seconde réside dans le traitement d'autres problèmes tels que la détection d'effets traitement à partir du questionnaire PQVS, enfin une troisième consiste dans l'élaboration d'un outil informatique convivial intégrant les divers logiciels et permettant le suivi et la modification conviviales d'applications informatiques axées sur l'utilisation des techniques d'analyse des données

7 BIBLIOGRAPHIE

Diday E, Pouget J, Lemaire J, TESTU E. 1982 éléments d'analyse des données Dunod 1982

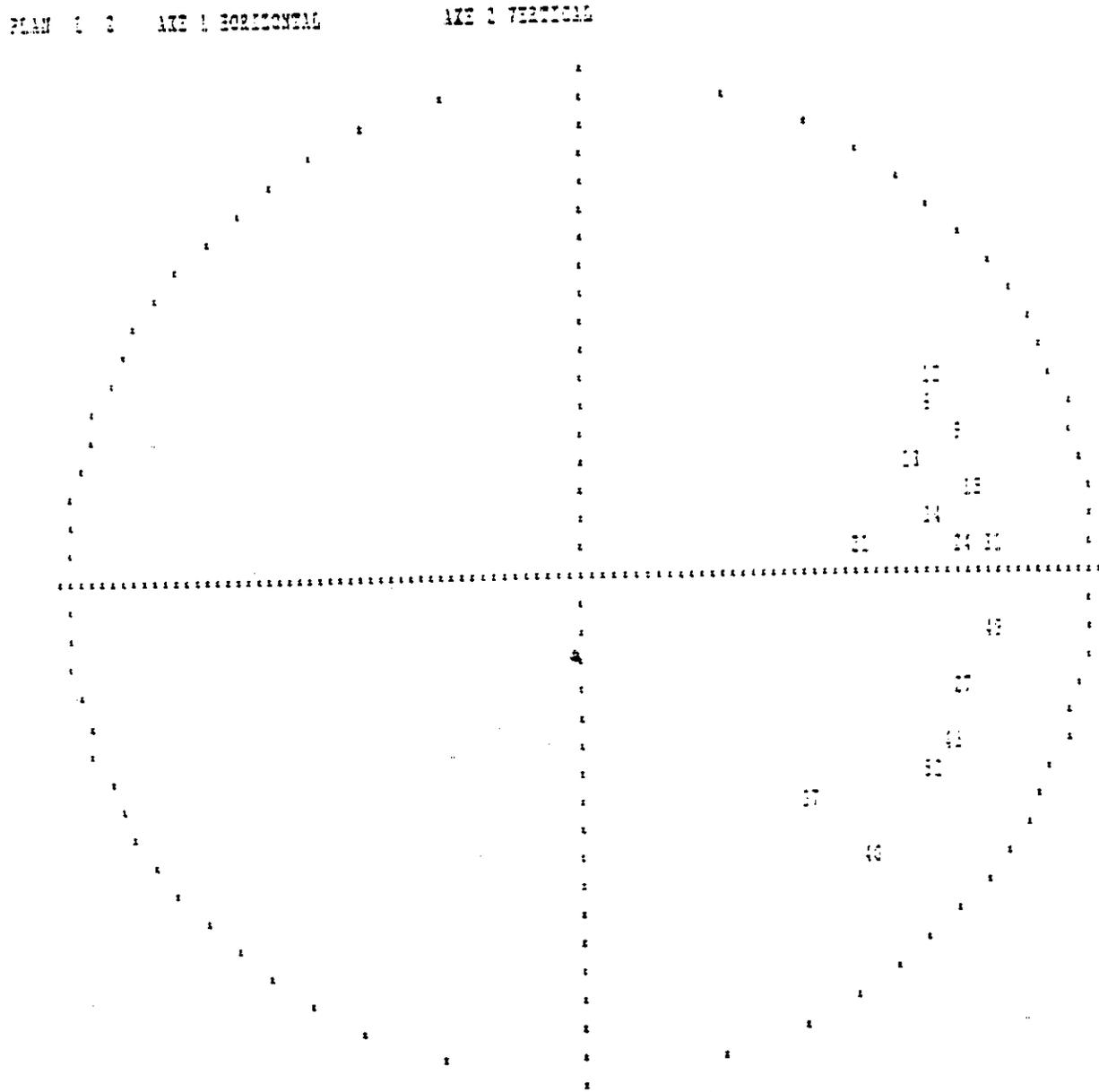
Gérin P, Dazord A, Boissel JP, Hanauer M.TE, Holey P, Chauvin F. 1989
l'évaluation de la qualité de la vie dans les essais thérapeutiques. Pharmacologie clinique, actualités et perspectives III G Strauch coll INSERM vol 185 1989 159-181

Rosnay P. 1973 Le macroscope éditions du seuil.

Structure de la Qualité de la vie

(CCAP! N=488)

Analyse en Composantes Principales



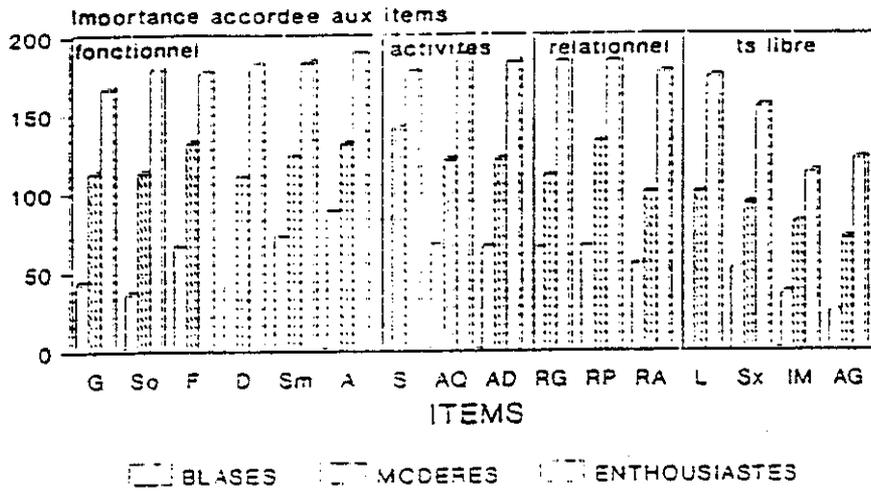
Code des variables

- 6 Gestes courants
- 9 Déplacements
- 12 Fatigue
- 15 Digestion
- 18 Sommeil
- 21 Sexualité
- 24 Attention, concentration
- 27 Relations en général
- 31 Activités Quotidiennes
- 34 Préoccupation par la Santé
- 37 Intérêt pour le monde
- 40 Activités en Groupe
- 43 Relations avec des Amis
- 46 Relations avec les Proches (superposé avec 27)
- 49 Activités
- 52 Loisirs

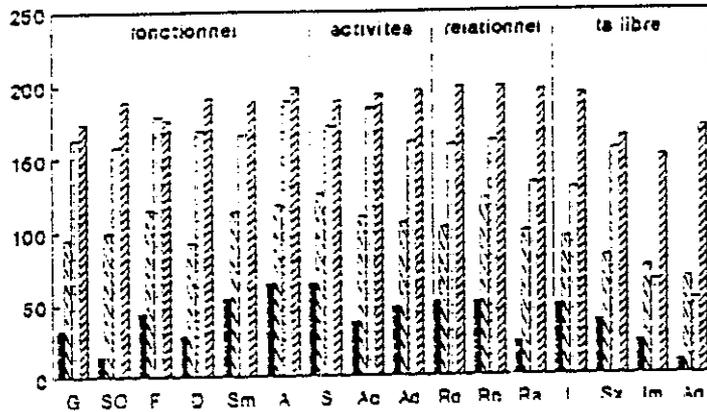
figure 1

TYPOLOGIE SUR LES ITEMS

Partition optimale en 3 classes



partition optimale en 4 classes



partition optimale en 7 classes

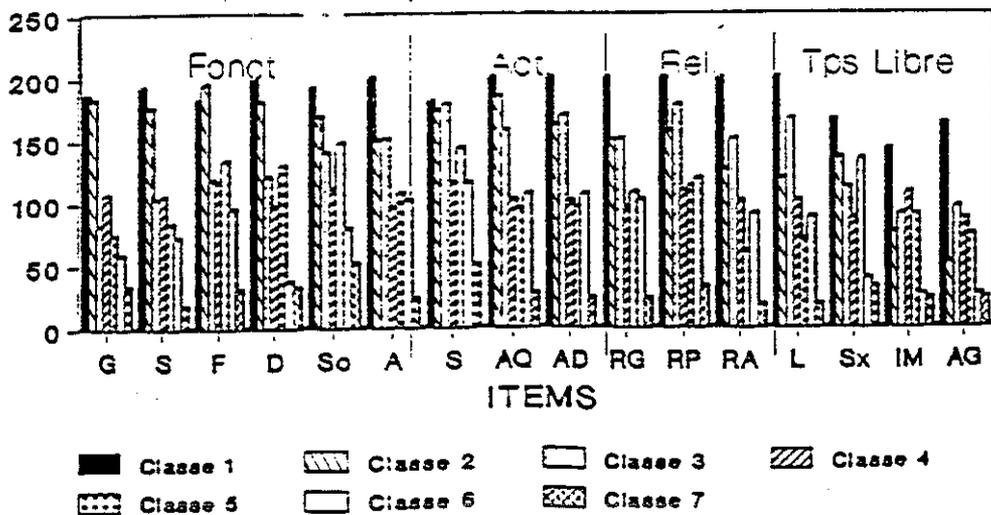
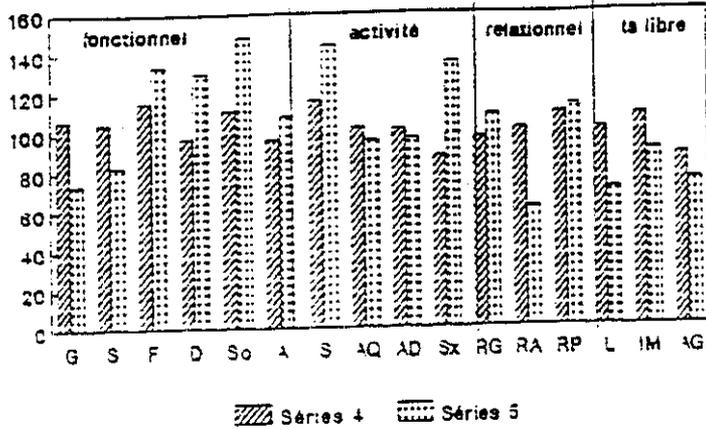


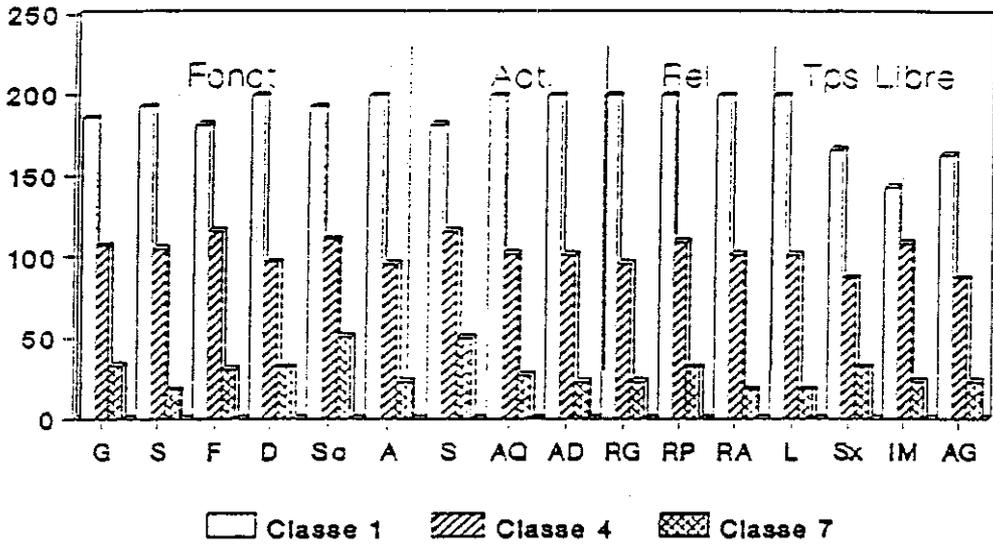
figure 2

partition optimale en 7 classes

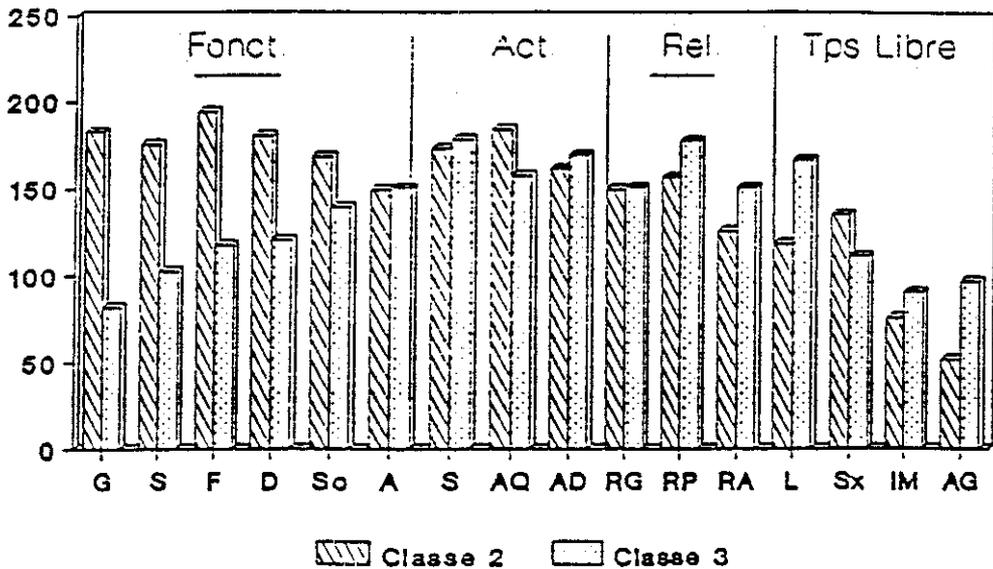
Sci Non-sci



Des "enthousiastes"(1) aux "biaisés"(7)

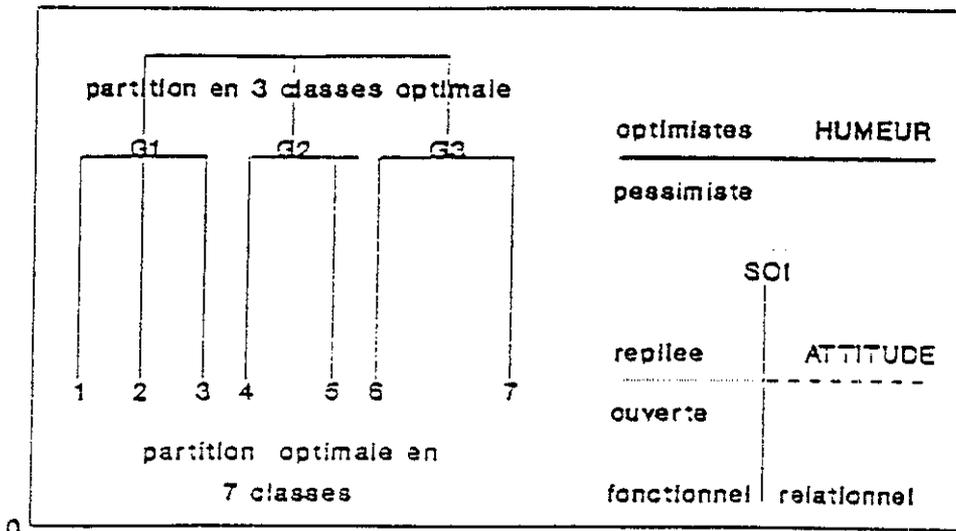


Balance inverse entre Fonct. et Rel.



Qualite de la vie (ocapi)

Liens entre structures axiale et hierar.



Qualité de la vie

Structure axiale échantillon A

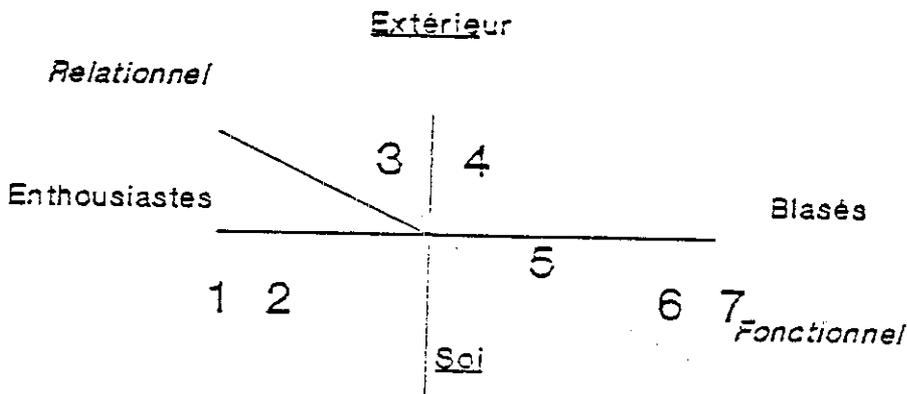


figure 4

tableau de contingence associé aux partitions en 3 et 7 classes

	C1	C2	C3	C4	C5	C6	C7
G1	45	28	15	0	0	0	0
G2	0	14	17	71	17	15	1
G3	0	0	0	5	5	28	22

Structure de la qualité de la vie

(OCAPI N=588)

Classification par les nuées dynamiques

Nbre de classes	Variance interclasse	Var.InterCl / Var.Tot	..Ecart
2	2603	0.319	
.....	0.075
3	3294	0.404	
.....	0.044
4	3660	0.448	
.....	0.026
5	3870	0.474	
.....	0.028
6	4099	0.502	
.....	0.025
7	4299	0.527	
.....	0.012
8	4404	0.539	

TABLEAU RELATIF A LA SELECTION DES VARIABLES OPTIMALES

nb	1	2	3	4	5	6	7	8	9
Θ	0	2	0.235	0.108	0.058	0.059	0.050	0.022	0.02
V	0	0.117	-1.4	-0.088	-0.03	-0.04	-0.02	-0.03	-
Γ	-	-0.7	-0.1	0.7	0.04	0.03	-0.01	-	-
S			+	+		+	+		

nb = nb de classes de la partition optimale
 Θ = taux d'accroissement de critère mesurant la qualité des partitions
 V = vitesse de ce taux Γ = taux d'accélération de ce taux
 S(i) = + indique que la partition optimale en i classes est une partition Optimale

tableau 1

PARTITION OPTIMALE EN 7 CLASSES

(C1, C2, C3, C4, C5, C6, C7) désigne l'ensemble ordonné de ses classes

TABLEAU DES POURCENTAGES PAR ITEM DES MODALITES DE REPONSES SELECTIONNEES

	GESTES COURANTS			SORTIES			FATIGUE			DIGESTION			SOMMEIL			SEXUALITE			ATTENTION			RENCONTRE AVEC GENS		
	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI
C1	0	14	86	0	7	93	5	9	86	0	5	95	0	7	93	14	7	79	0	1	99	0	2	98
C2	0	19	81	0	24	76	0	7	93	0	19	81	2	29	69	17	33	50	0	12	88	2	48	50
C3	19	71	9	9	78	13	6	66	28	3	72	25	5	50	47	19	50	31	0	34	66	0	47	53
C4	10	72	17	9	76	15	3	79	18	9	84	7	3	78	17	22	69	9	1	88	11	9	86	5
C5	74	17	9	30	52	18	4	57	39	4	61	35	4	48	48	4	57	39	57	43	0	13	65	22
C6	48	45	7	39	50	11	16	70	14	62	36	22	32	54	14	62	37	2	9	66	25	14	70	7
C7	74	17	9	87	13	0	78	13	9	78	17	5	61	22	17	70	26	4	70	26	4	78	18	4

	ACTIVITE QUOTIDIENNE			SANTÉ			INTERET MONDE			ACTIVITE DE GROUPE			RECONTRE AVEC AMI			RELATION PROCHES			ACTIVITE DOMESTIQ			LOISIRS		
	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI	PI	AI	TI
C1	0	5	95	0	18	82	11	35	55	2	32	66	0	5	95	0	5	95	0	5	95	0	5	95
C2	0	12	88	0	29	71	33	60	7	0	45	55	7	60	33	0	40	60	0	40	60	12	52	36
C3	0	44	56	0	19	81	28	53	19	21	59	19	0	47	53	3	16	81	0	28	72	3	28	69
C4	3	92	5	3	79	18	5	82	13	15	84	1	7	85	8	4	83	13	0	99	1	5	95	0
C5	18	65	17	5	52	43	91	9	0	78	22	0	43	52	5	17	52	31	13	78	9	39	44	17
C6	11	71	18	2	77	21	75	23	2	75	25	0	21	68	11	4	73	23	9	73	18	10	77	5
C7	78	22	0	57	39	4	74	17	9	91	4	5	87	13	0	74	22	4	78	18	4	87	13	0

Exemple d'interprétation de classe
interprétation de la classe 5

variables sélectionnées

	premier niveau			deuxième niveau		
	PI	AI	TI	PI	AI	TI
gestes courants	74	--	--	sommeil	--	48 48
digestion	--	61	--	activités professionnelles	--	52 43
activités quotidiennes	--	65	--	rencontres avec des amis	43	52 --
intérêt au monde	91	--	--	sexualité	--	57 39
activité de groupe	78	--	--	attention	57	43 --
rencontres avec les gens	--	65	--	fatigue	--	57 39

A un premier niveau les individus de cette classe marquent leur désintérêt pour le relationnel ouvert (monde et activité de groupe). Lorsque l'on regarde les variables du premier et du 2^{me} niveau on peut émettre l'hypothèse qu'ils sont intéressés par leur fonctionnement somatique, que pour eux les relations en générales sont plus importantes que les relations avec des amis. Ils mettent de l'intérêt dans les activités quotidiennes et professionnelles et attachent de l'importance à la sexualité que nous rattachons faute de plus amples information, dans ce cas précis aux relations avec les gens et donc au relationnel

