

161

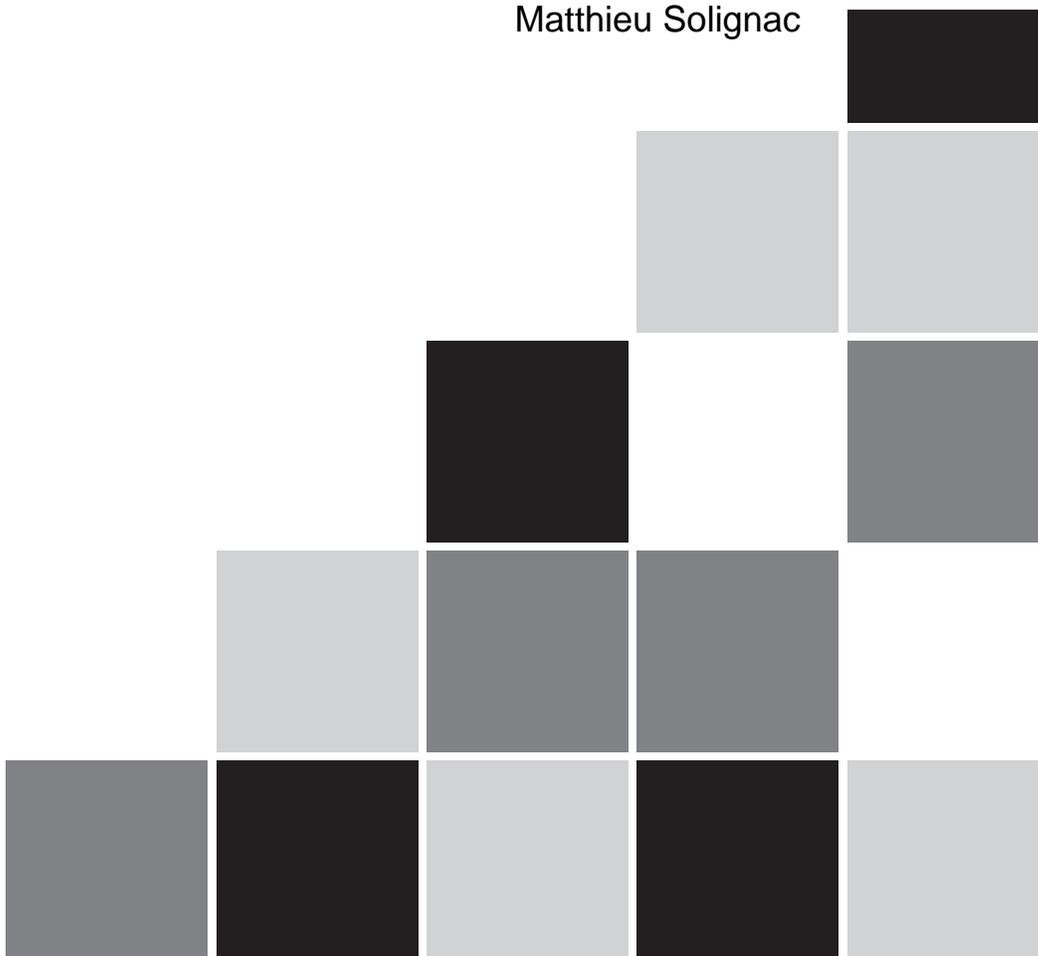
2009

DOCUMENTS DE TRAVAIL

LES POLITIQUES DE CONCILIATION VIE PROFESSIONNELLE / VIE FAMILIALE MENEES PAR LES EMPLOYEURS :

Elaboration d'une typologie des établissements
de l'Enquête Familles-Employeurs

Matthieu Solignac



LES POLITIQUES DE CONCILIATION
VIE PROFESSIONNELLE / VIE FAMILIALE
MENÉES PAR LES EMPLOYEURS :
Elaboration d'une typologie des établissements
de l'Enquête Familles-Employeurs
*Traitement de la non-réponse partielle
et méthodes de classification.*

Matthieu SOLIGNAC

2009

MOTS-CLEF : TRAITEMENT DES DONNEES, NON-REPONSE, FAMILLE, ENTREPRISE,
CONDITIONS DE TRAVAIL

INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES
133, boulevard Davout 75980 Paris Cedex 20
Tél. : (33) 01 56 06 20 00 Fax : (33) 01 56 06 21 99

Ce document de travail résulte du stage d'application de l'École Nationale de la Statistique et de l'Administration Économique (ENSAE) réalisé été 2007 à l'INED sous la direction d'Ariane Pailhé et Anne Solaz. Pour toute correspondance : matthieu.solignac@ined.fr. Je remercie Ariane Pailhé et Anne Solaz pour l'opportunité qu'elle m'ont offerte de mener ce travail, leurs conseils et leurs relectures. Les erreurs restantes demeurent bien entendu de ma seule responsabilité.

Table des matières

Introduction	5
I Traitement de la non-réponse	7
1 Première approche de la non-réponse partielle	11
1.1 La non-réponse : modalité à prendre en compte ou absence d'information?	11
1.2 Un coût du traitement de la non-réponse indispensable à supporter	15
2 La non-réponse dans le volet Employeurs	17
2.1 État de la non-réponse sur le champ d'intérêt	17
2.2 Les limites des différentes procédures de traitement simple de la non-réponse	21
2.3 Les méthodes de hot-deck	24
3 Mise en oeuvre d'un hot-deck métrique	27
3.1 Présentation des variables auxiliaires utilisées	27
3.2 Définition des donneurs	27
3.3 Calcul des pondérations affectées aux différentes variables auxiliaires	28
3.4 Prise en compte de l'information fournie par les variables potentiellement imputables . .	29
3.5 Bilan du traitement	30
II Classification	33
4 Éléments de méthode	35
4.1 Choix de la procédure de classification	35
4.2 Mise en oeuvre	36
5 Réalisation préalable d'une ACM	39
5.1 L'axe 1 des droits à des dispositifs de conciliation	39
5.2 L'axe 2 des aménagements horaires	40
5.3 L'axe 3 des mesures moins directement orientées vers les parents	40
6 Une classification en 7 classes	43
6.1 Présentation générale	43
6.2 Description des classes obtenues	44
6.3 Synthèse des résultats	48
Éléments de conclusion	59
Bibliographie	63

III	Annexe	65
A	Questionnaire et libellés des modalités dans les sorties	67
B	Annexe non-réponse	77
	B.1 La non-réponse simultanée à deux variables	77
	B.2 La constitution des 8 grandes régions françaises	78
	B.3 Code SAS	78
C	Annexe ACM	101
	C.1 Histogramme des valeurs-propres et coordonnées des projections des modalités des variables actives	101
	C.2 Plans 1-3 et 2-3	103
	C.3 ACM en utilisant les pondérations	104
	C.4 Projection de quelques variables supplémentaires	108
D	Annexe classification	125
	D.1 Classification ascendante hiérarchique sur les dix premiers axes	125
	D.2 Classification ascendante hiérarchique sur les vingt-huit premiers axes	132
E	Passage de SAS à SPAD	141
	E.1 Méthode de transfert proposée par l'aide de SPAD	141
	E.2 Solution alternative	142

Table des figures

1.1	Résultat obtenu pour une variable dichotomique quelconque	12
1.2	Résultat potentiellement utilisé dans le cadre d'une analyse de type "plancher"	12
2.1	Variabiles utilisées	19
2.2	Nombre de non-réponses parmi les non-répondants (partiels)	20
2.3	La non-réponse pour les 22 variables	21
2.4	Répartition des poids dans le volet employeurs (avec valeurs extrêmes tronquées d'ordre 2%)	23
2.5	Tests de la prédiction obtenue par le modèle de régression	26
3.1	Calcul des pondérations affectées aux 4 variables auxiliaires	29
3.2	Pondérations pour le calcul de la distance	30
3.3	Nombre d'utilisations des donneurs	30
3.4	Répartition des effectifs dans les différentes modalités suivant le traitement de la non-réponse retenu	32
5.1	Plan 1-2 de l'ACM	42
6.1	Structure de la classification	45
6.2	Récapitulatif des classes	49
6.3	Barycentres des 7 classes dans le plan 1-2	50
6.4	Projection des établissements des différentes classes	51
6.5	Dispersion des effectifs (75% des effectifs) autour des barycentres de chacune des 7 classes dans le plan 1-2	53
6.6	Caractéristiques des classes 1 et 2	55
6.7	Caractéristiques des classes 3 et 4	56
6.8	Caractéristiques des classes 5 et 6	57
6.9	Caractéristiques de la classe 7	58
A.1	Libellés des modalités dans les sorties	68
B.1	Part (en %) des établissements non-répondants pour deux variables	77
C.1	Histogramme des valeurs propres pour l'ACM	101
C.2	Coordonnées des projections pour l'ACM	102
C.3	Histogramme des valeurs propres pour l'ACM pondérée	104
C.4	Coordonnées des projections pour l'ACM pondérée	105
C.5	Plan 1-2 de l'ACM pondérée	106
C.6	Plan 1-3 de l'ACM pondérée	106
C.7	Plan 2-3 de l'ACM pondérée	107
C.8	Type d'établissement (plan 1-2)	108
C.9	Type d'établissement (plan 1-3)	109

C.10	Type d'établissement (plan 2-3)	109
C.11	Secteurs d'activité (plan 1-2)	110
C.12	Secteurs d'activité (plan 1-3)	110
C.13	Effectif de l'établissement (plan 1-2)	111
C.14	Effectif de l'établissement (plan 1-3)	112
C.15	Effectif de l'établissement (plan 2-3)	112
C.16	Proportion de femmes (plan 1-2)	113
C.17	Proportion de femmes (plan 1-3)	114
C.18	Proportion de femmes (plan 2-3)	114
C.19	Proportion de CDI (plan 1-2)	115
C.20	Proportion de CDI (plan 1-3)	116
C.21	Proportion de CDI (plan 2-3)	116
C.22	Age des salariés (plan 1-2)	117
C.23	Proportion de cadres (plan 2-3)	118
C.24	Régions (plan 1-2)	119
C.25	Régions (plan 1-3)	119
C.26	Caractéristiques du directeur (plan 1-2)	120
C.27	Caractéristiques du directeur (plan 1-3)	121
C.28	Plan 2-3	121
C.29	Opinion sur le rôle de l'entreprise (plan 1-2)	122
C.30	Opinion sur le rôle de l'entreprise (plan 1-3)	123
C.31	Opinion sur le rôle de l'entreprise (plan 2-3)	123
D.1	Dendrogramme de la classification réalisée sur les dix premiers axes	125
D.2	Classification sur les dix premiers axes	126
D.3	Classification sur les dix premiers axes	127
D.4	Classification en 7 classes sur les dix premiers axes	128
D.5	Barycentres des 7 classes dans le plan 1-3	129
D.6	Barycentres des 7 classes dans le plan 2-3	129
D.7	Classification en 6 classes sur les dix premiers axes	130
D.8	Barycentre des 6 classes	131
D.9	Dendrogramme de la classification sur les vingt-huit premiers axes	132
D.10	Classification sur les vingt-huit premiers axes	133
D.11	Classification sur les vingt-huit premiers axes	134
D.12	Classification en 7 classes sur les vingt-huit premiers axes	135
D.13	Classification en 6 classes sur les vingt-huit premiers axes	136
D.14	Barycentre des 7 classes	137
D.15	Projection des établissements des 7 classes	139

Introduction

Jusqu'à la fin des années 60, une majorité de femmes françaises n'occupaient pas d'emploi entre 25 et 49 ans. Depuis la fin du siècle dernier, le taux d'activité des femmes de cette classe d'âge est d'environ 80%. Cette évolution traduit principalement un nouveau rapport des femmes à l'emploi et à la famille, le modèle dominant étant désormais celui d'un cumul des activités professionnelles et domestiques. Dans le même temps, les hommes tendraient à accorder à la vie familiale une importance accrue. Désormais, pour les hommes comme pour les femmes, le rapport entre emploi et famille serait donc majoritairement vécu sur le mode de la conciliation.

Les méthodes d'enquête généralement retenues pour étudier la façon de mener en parallèle vie professionnelle et vie familiale consistent à interroger les salariés. Mais autant les données recueillies sur l'environnement familial peuvent être développées, autant la description du milieu professionnel reste souvent pauvre. Elle se limite généralement au secteur d'activité et aux effectifs. En dépit de leur importance, ces variables ne permettent de comprendre qu'une faible partie des différences des comportements individuels : la variance expliquée reste généralement assez faible. Les caractéristiques individuelles non prises en compte sont certes susceptibles d'être à l'origine de cette situation. Mais c'est avant tout le manque de données sur l'environnement de travail qui paraît être la source d'une part importante d'hétérogénéité inobservée. Un des principaux objectifs de l'enquête couplée Familles-Employeurs (2004-2005) est justement de remédier à ce manque d'information.

Réalisée par l'INED en partenariat avec l'INSEE, cette enquête sur la conciliation entre vie professionnelle et vie familiale a été menée à la fois auprès des ménages et de leurs établissements employeurs¹, qu'ils soient publics ou privés². L'enquête administrée en face à face à deux individus de 20-49 ans du ménage a été suivie de l'envoi d'un questionnaire spécifique aux établissements les employant lorsque les effectifs de ceux-ci étaient supérieurs ou égaux à vingt. Cette enquête a finalement permis de recueillir les réponses de 9547 individus (volet Salariés) dont 2852 couples et de 2672 établissements d'au moins vingt salariés les employant (volet Employeurs).

Le volet Employeurs contient de nombreuses informations permettant de caractériser la prise en compte des contraintes familiales par l'établissement de travail. Mais cette masse d'information est d'autant plus difficile à prendre en compte que les dimensions de la conciliation vie professionnelle/vie familiale sont multiples et souvent non hiérarchisables : certains établissements peuvent choisir de mettre à disposition des infrastructures, d'autres de laisser aux salariés une grande souplesse dans l'aménagement de leur temps travail ou encore de leur verser des aides financières. Ainsi un travail de synthèse de l'ensemble de l'information recueillie auprès des établissements sur les mesures de conciliation est nécessaire.

¹Il s'agit bien de l'établissement et non de l'entreprise. La consigne donnée aux enquêteurs pour définir l'établissement a été la suivante : "L'établissement pour lequel la personne travaille coïncide en général avec son lieu de travail. Si elle travaille dans une antenne, une annexe, on parle de l'établissement dont cette antenne dépend (par exemple on parle de la mairie pour le gardien d'un stade municipal). S'il y a un doute, par exemple pour les prestataires de service extérieur ou les intérimaires, prendre l'établissement qui gère les horaires quotidiens" (d'après [17] p.27).

²De ce fait, les résultats obtenus à partir de cette enquête peuvent difficilement être confrontés à d'autres sources. Le niveau établissement est en effet rarement retenu dans la collecte d'information portant sur le secteur public français.

L'objectif d'un tel travail est notamment d'élaborer une variable synthétique qui permette de profiter des potentialités offertes par le caractère couplé de l'enquête³. Faisant suite à une première tentative de classement des établissements par une méthode a priori ([13]), ce sont les méthodes de classification automatique qui ont été ici retenues. Ce travail a finalement débouché sur l'élaboration d'une typologie en 7 classes. Chacune d'elle est associée à un mode particulier de prise en compte par l'établissement Employeur des contraintes principalement parentales de ses salariés. Les phénomènes de faibles taux de réponse et modalités aux fréquences trop faibles n'ont pas permis de tenir compte de toutes les dimensions des contraintes familiales. La question de la prise en charge des personnes âgées n'a ainsi pas pu être retenue. La variable synthétique élaborée devrait néanmoins permettre de tirer plus facilement partie du caractère couplé de l'enquête.

La prise en compte de l'utilisation future de la typologie dans le cadre d'analyses couplées est en effet une des spécificités de ce travail. Ainsi elle a vocation à être utilisée comme variable synthétique caractérisant l'environnement de travail lors d'analyses menées au niveau des salariés. D'autre part, cette typologie doit être utilisable par toute personne s'intéressant à l'enquête Familles-Employeurs. L'homogénéité des typologies qui pourront être faites à la suite de celle-ci a donc été un souci constant. Cela a conduit à sélectionner un mode de traitement de la non-réponse particulier (hot-deck métrique) et à privilégier les procédures déterministes (classification ascendante hiérarchique) par rapport à celles faisant intervenir de l'aléa.

Ce document de travail a vocation à servir de guide dans l'utilisation de la typologie obtenue. Celle-ci est le résultat de nombreux choix dont la connaissance est indispensable pour une utilisation convenable de celle-ci. Mais ce travail d'explicitation de l'élaboration de cette typologie doit également permettre à ceux qui le souhaitent de construire plus facilement des typologies adaptées à leurs besoins particuliers. Une première partie est consacrée au traitement de la non-réponse partielle. Souvent passée sous silence, cette étape est cependant indispensable tant d'un point de vue pratique que méthodologique. La méthode de correction finalement retenue est le hot-deck métrique élaboré par M. Vanderschelden. La seconde partie est consacrée à la classification proprement dite. Ce travail, effectué sur la base corrigée, conduit à l'obtention de la partition des établissements en 7 classes.

³On pourra consulter à ce sujet le recueil d'études menées à partir de données appareillant informations sur les salariés et sur leurs entreprises [6]. Les numéros entre crochets renvoient à la bibliographie finale

Première partie

Traitement de la non-réponse

Toute enquête statistique est confrontée à des phénomènes de non-réponse. D'origines multiples, ils varient notamment de façon non négligeable suivant le mode d'enquête. Pour les questionnaires auto-administrés envoyés par courrier (cas du questionnaire Employeurs), ils sont généralement importants. Mais encore faut-il bien distinguer deux types de non-réponse, la non-réponse totale et la non-réponse partielle.

La **non-réponse totale** désigne les observations qui entrent dans le champ de l'enquête (ici établissement de plus de 20 salariés) mais qui sont jugées inexploitable du fait de l'absence de réponse (non retour du questionnaire) ou du trop grand nombre de non-réponses. Le principal problème consiste à distinguer parmi les entreprises qui n'ont pas renvoyé le questionnaire celles qui appartiennent au champ de l'enquête (et qui constituent la "véritable" non-réponse) de celles qui en sont exclues (ici les établissements de moins de 20 salariés)⁴.

La non-réponse totale ayant déjà fait l'objet d'un traitement⁵, nous nous concentrons ici sur la non-réponse partielle. Par **non-réponse partielle** on entend ici le fait pour une observation du "volet Employeurs"⁶ d'avoir certaines des questions se rapportant à notre champ d'étude non renseignées à l'issue de l'enquête. Ce phénomène se révèle parfois important pour les variables correspondant précisément au thème de la conciliation vie familiale/vie professionnelle dans les établissements. Combinée à l'effectif limité enquêté (moins de 3000 établissements) et aux faibles effectifs pour certaines modalités pourtant essentielles (par exemple l'existence de places en crèche), cette situation est susceptible d'entamer la qualité de l'analyse : le comportement de l'utilisateur de la base de données face à cette non-réponse peut influencer considérablement sur le résultat final obtenu.

Les choix du traitement de la non-réponse ne peuvent être dissociés de l'utilisation qu'on veut faire de la base de données corrigée (voir [3]), à savoir réaliser une classification. Les procédures de classification ne peuvent opérer qu'à partir du moment où la non-réponse partielle a été traitée. La modification de ce traitement peut influencer de façon non négligeable sur la classification selon le critère d'agrégation retenu. Il convient ainsi d'adapter le traitement de la non-réponse en fonction de l'usage des variables qui va être fait par la suite. **Le traitement de la non-réponse dans le cadre d'une classification conduit à nous intéresser moins à l'impact du traitement de la non-réponse sur la variance de chaque variable qu'à celui sur la corrélation entre les variables.** Il ne s'agit donc pas de travailler directement sur la précision d'estimateurs de variables quantitatives continues mais sur les relations entre variables qualitatives (ou rendue comme telles). Le traitement de la non-réponse ne sera appliqué que sur les variables actives retenues pour la classification (22 variables). La définition du champ de travail de manière précise permet de limiter le travail sur la non-réponse, toujours délicat, au strict nécessaire⁷. Enfin, d'un point de vue plus pratique, notons que c'est souvent le manque d'information sur le traitement de la non-réponse choisi qui rend impossible d'aboutir à une typologie précédemment réalisée par autrui. Cette situation tend alors à rendre caduc l'objectif même de la classification à savoir constituer un repère utilisable par tous.

Il ne s'agit pas de tomber dans l'excès inverse d'un relativisme total qui consisterait à proposer une multitude de typologies. L'objectif est de choisir une typologie en fournissant les éléments nécessaires à sa reconstitution à partir des données initiales. Les choix effectués ne sont pas incontestables. Tout comme le traitement de la non-réponse ne parviendra jamais à remplacer une réponse, on n'aboutit finalement qu'à *une* typologie parmi d'autres possibilités. Mais dès lors que ces choix sont explicités, ils ont le mérite de pouvoir être discutés.

⁴Sur le traitement de la non-réponse totale pour les enquêtes par courrier auprès des entreprises, voir [2] et [3].

⁵Voir [17] p.40 et surtout [7] p. 24.

⁶Plus justement il s'agit de cette base de données à laquelle ont été retranchées 21 observations. Nous verrons plus loin que nous avons traité celles-ci comme des non-réponses totales en raison du trop grand manque d'information sur les 22 variables qui nous intéressent.

⁷Pour autant, une analyse exploratoire, à l'aide d'un traitement de la non-réponse rudimentaire, est souvent un préalable nécessaire.

Cette première partie a vocation à souligner l'importance de la phase de traitement de la non-réponse et à présenter quelques méthodes pouvant être envisagées dans le cadre de la réalisation d'une classification. Le programme SAS de traitement de la non-réponse est fourni en annexe B.3 p.78 : il retrace l'ensemble des modifications effectuées sur la base de données initiale pour aboutir à la base utilisée pour la classification. La méthode du hot-deck métrique finalement retenue pour traiter la non-réponse reprend le travail réalisé par M. Vanderschelden ([20]). Ce document ainsi que la synthèse de N. Caron sur la correction de la non-réponse ont largement inspiré ce travail tant au niveau théorique qu'au niveau de la mise en oeuvre pratique.

Chapitre 1

Première approche de la non-réponse partielle

1.1 La non-réponse : modalité à prendre en compte ou absence d'information ?

Dans le cadre d'une analyse des mesures prises par les entreprises, les réponses "Ne sais pas" (NSP) ou "Non-réponse" (NR)¹ sont porteuses de sens sur l'effectivité des mesures au sein de l'entreprise ou sur le rapport à celles-ci : mauvaise connaissance de l'individu interrogé sur les mesures en place au sein de l'entreprise, volonté de l'établissement de ne pas communiquer sur l'absence de celles-ci sont autant d'interprétations possibles de la non-réponse. La non-réponse (NR) et l'impossibilité de répondre déclarée (NSP) peuvent ainsi être considérées comme des modalités comme les autres.

Cependant, une telle approche relève d'autres types d'analyses. L'étude de la diffusion de l'information (Dans les établissements dans lesquels des dispositifs sont en place, sont-ils bien connus des employés ? Par quel type d'employés ?) ou du comportement des acteurs, du répondant au codeur en passant par l'enquêteur (Comment se comportent-ils en situation d'ignorance ? Y-a-t-il une stratégie de dissimulation qui conduit l'individu interrogé à choisir ces modalités ?) ne peuvent être l'objet que d'une étude ultérieure plus fine qui nécessite d'avoir au préalable des informations sur la situation objective au sein de l'entreprise : la mesure sur laquelle porte la question est-elle en place ou non ? Or c'est justement cette dernière interrogation qui est l'objet du travail présent. Dans ce cadre, la non-réponse est donc assimilée à une absence d'information. **Le fait de ne pas répondre ou de déclarer une impossibilité de répondre définit l'appartenance à une catégorie commune que nous nommerons non-réponse.**

Certes la non-réponse pourrait être en réalité liée à une modalité particulière. Par exemple, la non-réponse quant à l'existence d'une politique d'entreprise pourrait seulement masquer son absence. Sans forcément voir une volonté de dissimulation derrière la non-réponse, on pourrait également considérer qu'une politique d'entreprise n'existe que si elle est connue par ses membres. Dès lors toute non-réponse serait assimilée à la modalité correspondant à l'absence de mesures favorisant la conciliation vie familiale/vie professionnelle.

Cette solution de traitement de la non-réponse est cependant peu acceptable. D'une part, certaines politiques familiales sont très ciblées suivant l'âge et la situation familiale et peuvent être mise en

¹Parmi les variables utilisées dans le cadre de notre travail, toutes proposent la modalité NR. Neuf d'entre elles, indiquées en italique dans la liste 2.1 p. 19, proposent en plus la modalité NSP.

place à différents niveaux (établissement, entreprise...). Il paraît probable que la personne répondant au nom de l'établissement ne soit pas toujours au fait de celles-ci et n'ait pas forcément le temps de se renseigner². Surtout, l'idée d'attribuer par défaut à l'établissement non-répondant une absence de mesures de conciliation (ou des mesures défavorables à celles-ci) est difficile à tenir du point de vue de la neutralité de la démarche. Si les résultats conduisaient à mettre en évidence un recours important à des politiques familiales par les entreprises, ceux-ci n'en seraient que plus renforcés par l'hypothèse initiale³. Cependant les premiers résultats tendent à montrer que la mise en place de politiques familiales au sein des établissements est limitée (notamment si on considère le secteur privé). Par conséquent, pour avoir un résultat potentiellement interprétable, il faudrait plutôt remplacer les données manquantes selon un principe opposé : attribuer l'existence de mesures de conciliation les plus développées en cas de non-réponse⁴.

Cette méthode et son interprétation en terme de "plancher" ne peut cependant pas être utilisée aisément. La principale raison est qu'en cas de proportion de non-réponse élevée et de modalité positive peu fréquente, le caractère théorique de la démarche n'a plus rien à voir avec la réalité. Supposons que parmi les 80% de répondants à une question dichotomique, 5% optent pour la réponse positive (voir figure 1.1 p.12). Le traitement de la non-réponse proposé précédemment, en allouant les effectifs des modalités "NR" et "NSP" à la modalité positive, conduit à raisonner sur un quart de réponses positives soit une part cinq fois plus importante que celle effectivement observée parmi les répondants (voir figure 1.2 p.12). D'autre part, ce traitement de la non-réponse affecte les corrélations entre modalités de différentes variables, corrélations sur lesquelles reposent les méthodes d'analyse de données. C'est justement l'objet de la suite de ce propos.

Modalité	part (%)
Oui	5
Non	75
Non-réponse (NR)	14
Ne sait pas (NSP)	6

FIG. 1.1 – Résultat obtenu pour une variable dichotomique quelconque

Modalité	Part (%)
Oui	25
Non	75

FIG. 1.2 – Résultat potentiellement utilisé dans le cadre d'une analyse de type "plancher"

²Ce dernier point pourrait être conforté par les premiers éléments dont nous disposons sur le moment choisi pour répondre au questionnaire : juste avant les dates d'expiration du délai et juste après les relances, le vendredi après-midi (pour le questionnaire en ligne) et plus généralement la fin de semaine semblent avoir été des moments privilégiés (voir [7] p.17 à 22 et [17] p.37 à 40).

³On analyserait alors une situation de type "plancher" : "Même en postulant que les non-répondants mettent en place les mesures les plus défavorables à la famille, on réussit à mettre en évidence que les établissements de travail ont des politiques familiales développées".

⁴Cette solution permettrait de développer des interprétations du type : "Même en postulant que les non-répondants mettent en place les mesures les développées en faveur de la vie familiale, on réussit à mettre en évidence que les établissements de travail ont des politiques familiales peu développées".

Le coût de l'absence d'information due à la non-réponse

La non-réponse pose plusieurs problèmes. Si on utilise les formes habituelles d'estimateurs en les restreignant aux seuls répondants⁵, elle est génératrice de biais et elle conduit également à une précision plus faible dès lors que cette non-réponse est importante (puisque la taille de l'échantillon exploitable devient plus petite). Mais cette situation affecte également les relations entre les variables, relations qui sont justement à la base des méthodes d'analyse de données.

L'introduction d'un biais

Dans le cadre de la théorie des sondages, l'estimateur d'Horvitz-Thompson⁶ est sans biais à condition que tous les individus sélectionnés dans la base de sondage répondent⁷. Or toute enquête est confrontée à de la non-réponse et le fait de ne pas répondre à une question est rarement le fruit du hasard (Voir [4] p.6,7.)

La non-réponse introduit ainsi un biais dans les estimations réalisées à partir des seuls répondants, un biais d'autant plus important que les non-répondants ont des caractéristiques ou un comportement différents des répondants. Au niveau pratique, les logiciels (SAS comme SPAD) n'admettent pas la présence de données manquantes dans les procédures d'analyse de données et opèrent par eux-même une suppression des observations pour lesquelles sont repérées des données manquantes. Pourtant cette solution consistant à ne garder que les individus ayant répondu à l'ensemble des questions n'est valable qu'à condition que la non-réponse soit le fruit du hasard, ce qui est peu probable.

Le recours à des estimateurs spécifiques tel que celui d'Horvitz-Thompson s'explique par la nature même de la constitution de la base de données : celle-ci n'est pas exhaustive et a été élaborée à partir d'un sondage. Le travail s'effectue ainsi (comme souvent) sur un échantillon issu de celui-ci. Chaque observation (chaque établissement) présente dans l'échantillon final qu'est le volet Employeurs représente en réalité un nombre particulier d'établissements, souvent différent de un. Ainsi dans l'enquête qui nous occupe, le plan de sondage est issu des individus : un établissement est interrogé dès lors qu'un de ses employés l'a été. Par conséquent, un établissement avait d'autant plus de chance d'être sélectionné qu'il avait un nombre important de salariés. Pour faire simple, si un seul établissement de 25 salariés et un unique établissement de 100 salariés sont présents dans l'échantillon, on peut s'attendre à ce qu'il y ait 4 fois plus d'établissements de 25 salariés que d'établissements de 100 salariés dans la réalité. On attribuera un poids à chaque observation, tenant compte de ce caractère inégal de probabilité d'inclusion dans l'échantillon : on pourra ainsi donner un poids 4 fois plus important à l'observation de l'établissement comptant 25 salariés. C'est le rôle de la variable de pondération présente dans la base de données.

Nous pouvons maintenant préciser notre propos. Pour estimer un total Y à partir de l'échantillon sur lequel on travaille (par exemple, pour connaître le nombre "réel" d'entreprises mettant à disposition des places en crèche), on utilise l'estimateur sans biais d'Horvitz-Thompson (dit aussi estimateur par valeurs dilatées). Soit U la population de taille N (l'ensemble des établissements dans la réalité) et s l'échantillon de taille n (l'ensemble des établissements du volet Employeurs). On définit l'estimateur d'Horvitz-Thompson \hat{Y}_π du total Y comme :

$$\hat{Y}_\pi = \sum_{k \in s} \frac{y_k}{\pi_k}$$

⁵L'expression "répondants" désignent les observations pour lesquelles toutes les variables sont renseignées. Par extension et selon le contexte, cette expression désignera également les observations pour lesquelles une variable ou un groupe de variables (alors précisées) ne sont pas affectées par de la non-réponse.

⁶La présentation formalisée de cet estimateur donnée par la suite s'appuie sur le cours de sondage dispensé par O. Sautory en 2e année à l'ENSAE [19].

⁷Voir [4] p.5.

avec π_k la probabilité d'inclusion, y_k la modalité prise par l'observation k pour la variable qui nous intéresse (ici l'existence (1) ou l'absence (0) d'une crèche). La probabilité d'inclusion π_k d'une observation k est l'inverse de son poids W_k : $W_k = \frac{1}{\pi_k}$. On vérifie bien que plus un établissement compte de salariés, plus sa probabilité d'inclusion est forte.

L'absence de biais est conditionné au fait que la méthode de tirage utilisée assure effectivement des probabilités égales aux π_k . Supprimer les observations manquantes revient à leur attribuer une probabilité d'inclusion nulle. Or si certaines probabilités d'inclusion sont nulles alors⁸ :

$$E(\widehat{Y}_\pi) = E\left(\sum_{k \in s} \frac{y_k}{\pi_k}\right) = E\left(\sum_{k \in U, \pi_k > 0} \delta_k \frac{Y_k}{\pi_k}\right) = \frac{1}{\pi_k} E(\delta_k) \sum_{k \in U, \pi_k > 0} Y_k = \sum_{k \in U, \pi_k > 0} Y_k \neq Y$$

La perte de précision

L'analyse portant sur les seuls répondants, en réduisant la population étudiée, affecte la précision des estimateurs. En effet la variance est directement liée à la taille de l'échantillon. Comme il est relativement fréquent qu'un individu interrogé ne réponde pas à au moins une question, la précision peut être affectée de façon importante.

Pour analyser plus précisément cet impact, analysons la perte de précision sur l'estimateur d'une moyenne \bar{Y} (par exemple la part des établissements possédant une crèche). Nous nous plaçons dans le cadre simple⁹ d'un sondage aléatoire simple sans remise de taille n . Il s'agit alors d'analyser l'impact d'une variation de n sur l'intervalle de confiance d'un estimateur sans biais \bar{y} d'une moyenne \bar{Y} définis tels que :

$$\bar{y} = \frac{1}{n} \sum_{k \in s} y_k \text{ et } \bar{Y} = \frac{1}{N} \sum_{k \in U} y_k \text{ avec } U \text{ la population de taille } N \text{ et } s \text{ l'échantillon de taille } n.$$

En admettant que n est suffisamment grand pour utiliser la loi asymptotique de l'estimateur (ie pour admettre que \bar{y} suit une loi normale $N(\bar{Y}, V(\bar{y}))$), on obtient l'intervalle de confiance estimé à 95% suivant :

$$IC = \left[\bar{y} - 1.96 \sqrt{(1-f) \frac{s^2}{n}}, \bar{y} + 1.96 \sqrt{(1-f) \frac{s^2}{n}} \right]$$

avec :

- $f = \frac{n}{N}$ le taux de sondage
- $s^2 = \frac{1}{n-1} \sum_{k \in s} (y_k - \bar{y})^2$ l'estimateur de la variance

$$\text{d'où } IC = \left[\bar{y} - 1.96 \sqrt{\left(1 - \frac{n}{N}\right) \frac{\sum_{k \in s} (y_k - \sum_{k \in s} \frac{y_k}{n})^2}{n(n-1)}}, \bar{y} + 1.96 \sqrt{\left(1 - \frac{n}{N}\right) \frac{\sum_{k \in s} (y_k - \sum_{k \in s} \frac{y_k}{n})^2}{n(n-1)}} \right]$$

Par conséquent, si n diminue d'une proportion λ , l'intervalle de confiance augmente : il s'accroît d'une proportion λ sans compter son augmentation par l'intermédiaire du taux de sondage.

⁸On utilise les variables de Cornfield définies pour tout $k \in U$ par

$$\delta_k = \begin{cases} 1 & \text{si } \{k\} \in s \text{ (probabilité } \pi_k) \\ 0 & \text{sinon.} \end{cases}$$

⁹Cette situation, ne correspondant pas au mode de recueil de notre base de donnée, a une vocation seulement illustratrice.

Sachant que ce sont principalement les petits établissements qui sont concernés par la non-réponse, et que ce sont justement ceux-ci qui possèdent les poids les plus élevés, le phénomène d'augmentation de l'intervalle de confiance est d'autant plus important.

La perturbation des corrélations entre les variables

Dès lors que le comportement et les caractéristiques des non-répondants sont différents de ceux des répondants, les relations observées entre les modalités des différentes variables sont également affectées. Les difficultés posées par la non-réponse dans le cadre d'une classification résultent principalement de cet effet. Le coefficient de corrélation ρ_{xy} entre deux variables x et y est donné par :

$$\hat{\rho}_{xy} = \frac{1}{N-1} \frac{\sum_{k \in U} x_k y_k - N \bar{X} \bar{Y}}{s_x s_y}$$

Par conséquent, le coefficient de corrélation est sensible à l'introduction de biais (par l'intermédiaire des composantes $\bar{X}\bar{Y}$), à la perte de précision (par l'intermédiaire des composantes s_x et s_y), à la modification de la taille de l'échantillon (à travers N). Mais il est également touché par l'évolution de la relation entre les variables x et y par le paramètre $\sum_{k \in U} x_k y_k$. Cette sensibilité des coefficients de corrélation est à l'origine des résultats différents qu'on peut obtenir à l'issue d'une classification suivant le mode de traitement de la non-réponse retenu.

1.2 Un coût du traitement de la non-réponse indispensable à supporter

Si les coûts de l'absence de traitement de la non-réponse sont nombreux tant en terme de biais, de précision que de modification des relations entre variables, le traitement de la non-réponse est l'est également, notamment en temps. De plus, son efficacité reste limitée. Le traitement de la non-réponse ne sera jamais parfait. Il n'est qu'un palliatif qui ne remplacera jamais une réponse indiquant la situation réelle de l'entreprise. Enfin, le traitement de la non-réponse conduit à un certain fétichisme de la réponse. Le traitement de la non-réponse peut ainsi suggérer à tort l'idée que les réponses¹⁰ sont parfaites. Or il convient de garder en mémoire que celles-ci sont toujours d'une qualité limitée. Elles sont en effet affectées par les réponses erronées voire plus ou moins sincères des individus, le comportement de l'enquêteur, les règles ou erreurs dans le processus de codage etc. Par conséquent, on peut s'interroger sur l'opportunité de procéder à un traitement de la non-réponse.

Cependant le fait de ne conserver que les répondants est déjà une forme de traitement de la non-réponse. Un traitement fondé sur l'hypothèse d'une identité de comportement et de caractéristiques entre les répondants et des non-répondants. Par conséquent, c'est moins la question du traitement ou non de la non-réponse que celle de la manière de réaliser ce traitement qui se pose¹¹. Le fait que le travail sur la non-réponse n'atteigne jamais la perfection n'implique pas qu'il ne doit pas être engagé. Plutôt que de laisser l'illusion d'une objectivité des données et de passer sous silence l'intervention de l'utilisateur sur celles-ci, mieux vaut afficher clairement les choix réalisés et leurs éventuelles conséquences sur les résultats obtenus. Enfin, tout travail a vocation à être vérifié et réutilisé par d'autres. Le silence sur les méthodes utilisées peut, au mieux, susciter la suspicion sur les résultats obtenus.

¹⁰Par réponse, on entend l'utilisation de modalités autres que NR ou NSP.

¹¹Rappelons que les logiciels SAS et SPAD choisissent, par défaut, de sélectionner les seuls répondants pour les procédures d'analyse de données.

Chapitre 2

La non-réponse dans le volet Employeurs de l'enquête "Familles et employeurs"

2.1 État de la non-réponse sur le champ d'intérêt

Variables retenues

Le travail est mené à partir de 22 variables (voir figure 2.1 p. 19). La réalisation de la typologie est ainsi le résultat de la confrontation d'une problématique (Quelles sont les formes de conciliation vie professionnelle/vie familiale mises en place dans les établissements?) et de la base de données qu'est le volet Employeurs de l'enquête Familles-Employeurs¹. La notion de "conciliation" est ici directement liée aux variables disponibles dans l'enquête.

Encore faut-il noter que certaines d'entre elles n'ont pu être retenues. Ainsi avant même de commencer à analyser l'enquête, la question de la prise en charge des personnes âgées avait été retenue comme une dimension potentiellement intéressante à étudier. Cependant ce thème n'est abordé que de façon indirecte par l'intermédiaire d'une unique question ("Un salarié de cet établissement peut-il bénéficier d'aménagements ou d'arrangements d'horaires pour accompagner un proche en fin de vie?") qui présente un taux de non-réponse particulièrement élevé (plus de 30%). De façon générale, l'absence parmi les variables retenue d'une variable potentiellement intéressante présente dans l'enquête s'explique par deux raisons : un taux de non-réponse trop élevé, une modalité présentant un effectif trop faible.

On a essayé dans la mesure du possible de contourner ces deux obstacles, notamment pour les mesures qui semblaient présenter un intérêt majeur (par exemple pour les variables se rapportant à l'existence d'un mode de garde des enfants au sein de l'entreprise). Pour cela, on a procédé à des regroupement de modalités ou de variables. Si nous reprenons directement certaines variables de l'enquête, d'autre sont construites par agrégation de plusieurs variables. Ainsi le nombre d'établissement ayant un dispositif de crèche (148 établissements) est relativement faible. La même situation se présente pour la mise à disposition d'une garderie (100 établissements sur 2673). La création d'une variable d'existence d'un dispositif de garde au sein de l'établissement par fusion de ces deux variables permet de les prendre en compte tout en limitant les problèmes de distortion engendrés par les modalités à faibles effectifs.

¹le questionnaire d'enquête est reproduit en annexe A p.67

D'autres variables sont des synthèses de variables filtrées et servent à préciser les situations. Ainsi en est-il de la variable portant sur le nombre de jours de congés rémunérés en cas de maladie de l'enfant. Elle est élaborée à partir d'informations provenant de deux questions : la première interrogeant sur la possibilité ou non d'avoir des jours de congé rémunérés en cas de maladie de l'enfant, l'autre permettant de préciser, en cas de réponse positive, la durée d'un tel congé. On en tire une variable à 4 modalités :

- absence de congé rémunéré
- 1 à 3 jours de congé
- 4 à 6 jours de congé
- plus de 6 jours de congé

On procède de même pour obtenir les variables de gestion du temps de travail (RTTdd est la synthèse de 3 variables) et du nombre de jours de congé maternité supplémentaires accordés pour un premier enfant (CONGDURD).

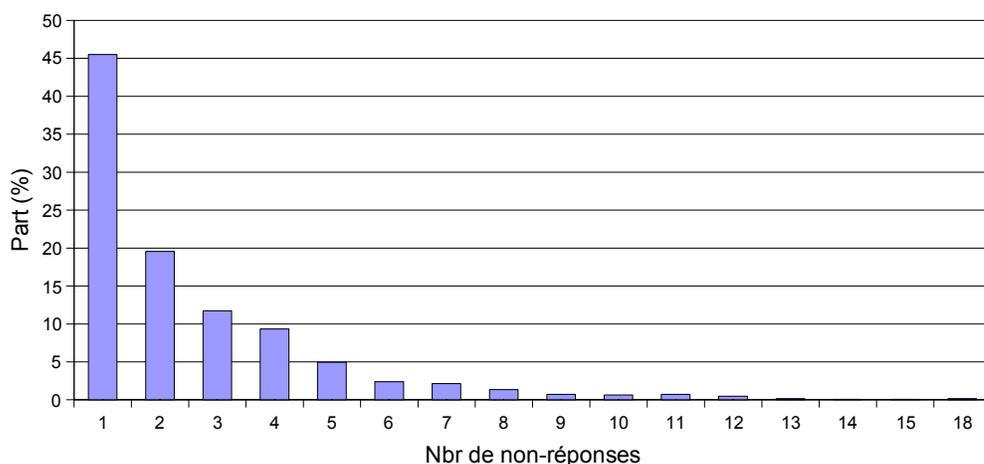
Cette méthode est indirectement une forme de traitement de la non-réponse, puisqu'un établissement ayant déclaré une garderie ou une crèche mais n'ayant pas répondu à l'autre question est considéré comme répondant. Cependant ce recours à des variables de synthèse voire à des indicateurs ne peut se faire qu'avec parcimonie. L'objectif est en effet de réaliser une classification automatique. Celle-ci n'a plus lieu d'être dès lors qu'on résume l'ensemble des variables en un nombre assez réduit d'indicateurs construits a priori. Or l'objectif de notre travail est justement d'analyser empiriquement les relations entre les variables et non pas de les postuler.

	Variable	Signification	Modalités	Variables utilisées
1	ENFMALDD	nombre de jours de congés en cas de maladie d'un enfant	0 1-3 4-6 >6	<i>ENFMAL</i> <i>ENFMALNB</i>
2	CONGDURD	nbr de jours de congés maternité accordés en plus pour un premier enfant	0 1-3 >3	<i>CONGDUR</i> <i>CONGDURNB</i>
3	GARDEDTT	possibilité de garde des enfants fournie par l'entreprise	oui non	<i>GARDE</i> <i>CRECHE</i>
4	ICRECHED	intention de créer une crèche	oui non	<i>ICRECHE</i>
5	PRESTAI	accès à des centres ou colonies de vacances	oui non	<i>PRESTAI</i>
6	PRESTAAD	existence d'une prime de naissance	oui non	<i>PRESTAA</i>
7	PRESTADD	aide financière aux frais de garde des enfants	oui non	<i>PRESTAD</i>
8	PRESTAED	aide financière aux frais d'études des enfants	oui non	<i>PRESTAE</i>
9	PRESTAGD	prestation pour enfants handicapés	oui non	<i>PRESTAG</i>
10	AMENAD	aménagement d'horaire possible le jour de la rentrée scolaire	oui non	<i>AMENA</i>
11	AMENBD	aménagement d'horaire possible par rapport aux horaires d'école ou de crèche	oui non	<i>AMENB</i>
12	AMENCD	aménagement d'horaire possible en cas de maladie de l'enfant	oui non	<i>AMENC</i>
13	AMENDD	aménagement d'horaire possible en cas de trajets domicile-travail importants	oui non	<i>AMEND</i>
14	DOMBD	possibilité de travailler parfois à domicile pour convenance personnelle	oui non	<i>DOMB</i>
15	PARTEMBD	existence de postes à temps partiel à l'embauche	oui non	<i>PARTEMB</i>
16	PARTACCD	temps partiel accordé lorsqu'il est demandé par le salarié	toujours parfois non	<i>PARTACC</i>
17	RTTDD	possibilité de gérer son temps libre supplémentaire issu de la RTT	libre accord contraint abs de RTT	<i>RTTSO</i> <i>RTTPRISA</i> <i>RTTPRIB</i>
18	PROMOCD	effet de la disponibilité sur les possibilités de promotion	toujours souvent parfois jamais	<i>PROMOC</i>
19	CONGD	maintien du salaire (ou complément) lors du congé paternité ou maternité en cas de dépassement du plafond SECU	oui non	<i>CONGSALB</i>
20	PRESTAHD	existence de chèques vacances	oui non	<i>PRESTAH</i>
21	FMUTUD	financement d'une mutuelle	tous cadre non	<i>FMUTU</i>
22	MUTUD	accès aux services mutualistes	oui non	<i>MUTU</i>

État de la non-réponse

1420 individus (sur 2673) répondent aux 22 variables retenues, soit seulement 53.1% de l'échantillon². Cette non-réponse apparaît relativement diffuse parmi les établissements. 45.5% des non-répondants ont donné 21 réponses sur les 22 demandées (graphique 2.2 p.20)³. Et dans plus de 75% des cas de non-réponse, moins de trois réponses sont manquantes.

Répartition (en %) des observations touchées par la non-réponse suivant le nombre de non-réponses



Note de lecture : Pour 45,49 % des 1253 observations concernées par la non-réponse, une seule variable est manquante.

Source : Enquête « Familles-Employeurs », volet Employeurs.

FIG. 2.2 – Nombre de non-réponses parmi les non-répondants (partiels)

La non-réponse pour chaque variable (voir figure 2.3 p.21) peut paraître limitée en proportion. Elle n'est supérieure à 10% que dans quatre cas (dont trois des quatre questions portant les possibilités d'aménagement du temps de travail). Cependant dans la mesure où la classification a vocation à être utilisée au niveau individuel, c'est plutôt le nombre de non-réponses et le degré de concentration de celles-ci sur quelques individus qui constituent les paramètres à prendre en compte. Une non-réponse de 5% pour une variable donnée signifie que cette seule variable va engendrer un problème de classification pour au moins 130 individus (voire plus si plusieurs individus travaillent dans ces établissements). Les taux de non-réponse par variable non négligeable combinés au caractère diffus de la non-réponse parmi les établissements sont ainsi à l'origine, pour la classification à venir, d'une difficulté qu'il convient de résoudre.

21 observations sont jugées inexploitable : il s'agit de celles ayant plus de 50% de non-réponse sur les 22 variables actives. On peut noter qu'il s'agit de petits établissements (leur distribution des poids est plus élevée). Si on s'intéressait vraiment aux établissements, on pourrait mettre en place une repondération pour tenir compte des caractéristiques spécifiques de cette non-réponse qualifiée ici de totale. Cependant l'objectif est de construire une typologie pour analyser les individus. Par conséquent, on se contentera seulement de renormaliser les pondérations. Ainsi les effectifs de salariés concernés seront d'autant plus limités que les 21 établissements semblent de faible taille.

²A noter qu'une partie de la non-réponse a été traitée par combinaison de plusieurs variables.

³Voir également annexe B.1 p.77 qui croise la non-réponse des variables deux à deux.

Variable	Part (%) de NR-NSP	Part (%) pondérée
1) enfmaldd	5.31	9.96
2)congdudd	2.28	3.88
3)gardentrd	3.59	3.65
4) icreched	5.54	3.89
5) prestaidd	4.68	6.21
6) prestaad	3.93	4.45
7) prestadd	5.20	5.74
8) prestaed	4.75	5.30
9) prestagd	6.47	7.94
10) amenad	5.09	5.94
11) amenbdd	13.69	16.62

12) amencc	10.29	12.02
13) amendd	13.77	17.71
14) dombd	2.81	2.99
15) partaccd	3.55	4.90
16) partemdb	1.16	1.11
17) rttdd	0.49	4.35
18) promocd	6.06	6.91
19) congdd	11.75	16.57
20) prestahd	3.78	4.5
21) fmutud	1.72	1.98
22) mutud	4.08	4.17

FIG. 2.3 – La non-réponse pour les 22 variables

2.2 Les limites des différentes procédures de traitement simple de la non-réponse

Élimination des non-répondants

Nous avons précédemment abordé la solution consistant à ignorer la non-réponse pour travailler uniquement à partir des répondants. La justification de cette méthode repose sur l'hypothèse (forte) de comportement identique des répondants et des non-répondants. Elle ne peut donc être que rarement sinon jamais utilisée.

Regard sur une procédure courante de réallocation

Présentation

Une méthode de traitement de la non-réponse souvent utilisée en première analyse consiste à réallouer les non-répondants dans les différentes modalités, proportionnellement à la part des répondants prenant ces modalités⁴.

Cette méthode pose plusieurs problèmes. D'une part, elle repose sur l'hypothèse forte que les non-répondants ont pour chaque réponse, en moyenne, la même structure de réponse que les répondants, une hypothèse finalement proche de la précédente. En effet, les réponses de l'ensemble des répondants à cette question, y compris les non-répondants pour d'autres questions, sont prises en compte. Mais pour chaque question, les non-répondants se voient finalement imputés une réponse de façon à ce que la structure des réponses des répondants à cette question soit conservée. Il s'agit donc d'utiliser l'ensemble de l'information disponible et de traiter la non-réponse de façon "neutre" relativement aux répondants.

Le poids des individus non-répondants

L'objectif de préservation de la structure de réponse des répondants n'est bien souvent même pas respecté dans les faits. Le caractère pondéré des observations rend en effet difficile de traduire correc-

⁴Le procédé encore plus frustrant consistant à réallouer les non-répondants à la modalité la plus fréquente procède du même type de "stratégie".

tement au niveau individuel cette exigence globale. Analysons la démarche suivie.

On travaille sur des individus et non directement sur les parts globales. Le principe d'imputation des données manquantes est que chaque individu a une chance d'appartenir à une modalité égale à la part représentée par celle-ci pour l'ensemble des répondants. Il faut garder en mémoire que le travail s'effectue (comme souvent) sur un échantillon issu d'un sondage. Le calcul des estimateurs doit ainsi prendre en compte le fait que les données sont issues d'un sondage à probabilités inégales. Le recours aux estimateurs de type Horvitz-Thompson est donc nécessaire. On les utilise par l'intermédiaire des procédures "surveymeans" et "surveyfreq" désormais disponibles sous SAS (voir [10]). La procédure d'enquête a été réalisée sans qu'on connaisse le nombre d'établissements de plus de 20 salariés existants en France. Pour le calcul des moyennes (par exemple la proportion d'établissements mettant à disposition un crèche), c'est l'**estimateur d'Hàjek** (biaisé) et non celui d'Horvitz-Thompson qui est utilisé par les procédures SAS⁵. En effet, contrairement à l'estimateur d'Horvitz-Thompson, l'estimateur d'Hàjek ne nécessite de connaissance préalable de l'effectif total de la population étudiée : cet effectif N est lui-même approximé par son estimateur d'Horvitz-Thompson :

$$\bar{Y}_\pi = \frac{1}{N} \sum_{k \in s} \frac{y_k}{\pi_k} = \frac{1}{\sum_{k \in s} \frac{1}{\pi_k}} \sum_{k \in s} \frac{y_k}{\pi_k}$$

Supposons que pour une variable dichotomique, la répartition obtenue pour les répondants d'après l'estimateur d'Hàjek soit 75% pour la modalité A et 25% pour la modalité B. En suivant la méthode de traitement de la non-réponse précédemment décrite, chaque individu non-répondant a 75% de chance de se classer dans la catégorie A et 25% de se classer dans la catégorie B. En pratique, on peut par exemple utiliser pour chaque individu le tirage d'une valeur issue d'une loi uniforme. Il suffit ensuite de classer les observations selon l'ordre croissant de cette valeur et attribuer la modalité A aux 75% d'individus classés le plus haut, la modalité B aux autres individus.

Cependant cette méthode ne tient pas compte du fait que chaque individu non-répondant a également un poids qui lui est propre. Ainsi, un classement de ce type peut finalement aboutir à une modification des proportions d'effectifs des différentes modalités (par exemple 80% et 20% si les individus affectés à A ont en moyenne un poids supérieur à ceux affectés à B). Il convient donc de tenir compte des poids des non-répondants.

Tenir compte du poids des non-répondant est plus difficile, surtout si les écarts de poids sont importants. Ainsi, si on veut attribuer 75% des poids à la catégorie A, on obtiendra rarement le poids exact. Par exemple, si on lui a déjà attribué 73% du poids et que l'individu suivant en représente 2,1%. Une solution pourrait consister à prendre le premier poids suivant qui convient. Cependant le poids est directement lié à l'effectif de l'établissement. Or la taille de l'établissement apparaît comme un déterminants majeur des politiques familiales. Ainsi les modalités réunissant une faible part de répondants ne pourront accueillir que les non-répondants ayant des poids limités, donc des établissements d'effectifs élevés. D'où des conséquences potentiellement non négligeables sur les résultats obtenus. Or le volet employeurs se caractérise justement par un nombre important de modalités de variables dichotomiques aux parts faibles et par des écarts de poids très importants. Ainsi **les rapports de poids sont de 1 à 433 entre les observations aux pondérations extrêmes** et encore ont-elles été préalablement tronquées d'ordre 2% (voir figure 2.4 p.23).

Une solution serait de créer des observations artefact, c'est à dire de répliquer l'individu de manière à pouvoir affecter à celui-ci les différentes modalités et la part du poids associée. Ainsi pour notre variable dichotomique, on pourrait imaginer de créer deux individus identiques, l'un prenant la modalité A et étant doté de 75% du poids initial, l'autre prenant la modalité B et ayant 25% du poids initial.

⁵L'utilisation d'un estimateur biaisé plutôt que celui d'Horvitz-Thompson, non biaisé, peut paraître paradoxal au premier abord. Mais le calcul de l'estimation d'une variable constante ou l'exemple plus imagé du cirque de D. Basu [5] montrent que l'absence de biais est loin d'être un signe de perfection.

Quantile	Valeur
Maximum	7.5996043
90%	2.4794606
75%	1.3129650
Médiane	0.5047289
25%	0.1681377
10%	0.0547646
5%	0.0297747
Minimum	0.0175513

FIG. 2.4 – Répartition des poids dans le volet employeurs (avec valeurs extrêmes tronquées d’ordre 2%)

Cependant une telle solution se heurte rapidement à une contrainte de taille par la multiplication rapide du nombre d’observations sans compter la perturbation qu’elle peut engendrer en terme de calcul de la variance par les logiciels. Enfin ce raisonnement sur la non-réponse repose sur une hypothèse forte : pour chaque question, les non-répondants ont un comportement similaire au comportement moyen des répondants à cette question.

La constitution de sous-catégories : un raffinement qui peut résoudre les problèmes de poids

Il est cependant possible de raffiner la procédure de traitement de la non-réponse en créant des sous-catégories. L’imputation d’une information manquante pour un individu donné ne se fera plus en fonction du comportement moyen de l’ensemble des répondants mais seulement des répondants lui ressemblant le plus (suivant un ensemble de critères préalablement définis). Le problème de l’hypothèse de comportement similaire entre répondants et non-répondants se posera également mais cette fois-ci au sein de sous-catégories rassemblant les établissements les plus proches selon l’ensemble de variables retenues.

Les informations utilisées pour la constitution des sous-catégories doivent être disponible pour l’ensemble des répondants et des non-répondants. Cette contrainte est forte, notamment pour cette enquête comprenant les établissements publics. Il s’agit donc de tirer au mieux parti des variables à la fois déterminantes pour les réponses et disponibles pour tous les individus.

On peut dresser l’inventaire des variables exhaustives ou de celles qu’on peut constituer comme telles à l’aide d’informations auxiliaires :

- le poids
- le type d’établissement : privé, public, groupe ou non, français ou non
- l’activité en 14 et 32 positions (NES16, NES32)
- la région
- l’ancienneté
- le nombre de jour d’ouverture

Des analyses préalables ont suggéré que l’effectif était un déterminant majeur de l’existence de politiques d’entreprise. L’information recueillie n’est pas exhaustive pour cette variable. On peut utiliser les informations disponibles dans la base de données du Système Informatique des Entreprises et des Établissements (SIRENE) : l’établissement y est répertorié par son numéro Siret (SIREN + NIC) : effectif de l’établissement, son secteur d’activité (comparable aux modalités de la variable PUBPRIV), l’activité principale (correspondant à NES16 dans l’enquête), date de création. Cependant celles-ci sont

affectées par un manque d'actualisation particulièrement problématique pour les effectifs (selon [3], [2]).

Nous avons noté précédemment que la non prise en compte des poids des non-répondants peut déstabiliser les résultats de l'analyse. Nous proposons donc d'utiliser le poids des individus comme variable première dans l'élaboration des sous-catégories. Il dépend à la fois de l'effectif et des comportements de réponse (non-réponse totale) des établissements. Ne reste alors qu'à déterminer les tranches.

Si le manque d'actualisation du fichier SIRENE pose problème sur les données d'effectif, il n'en est pas de même pour les informations sur la création de l'établissement (variable DCRET). Son utilisation permettrait ainsi de compenser les informations manquantes dans le volet Employeurs sur la date de création de l'établissement. Notre travail à partir des données SIRENE dont nous disposons n'a pas permis de compléter suffisamment de réponses manquantes pour pouvoir utiliser l'ancienneté comme variable auxiliaire. Initialement cette observation est manquante pour 204 observations⁶. En utilisant le fichier SIRENE, nous sommes parvenus à obtenir la date de création pour seulement 84 observations supplémentaires. Cette information complémentaire ne sera pas utilisée.

Une fois les variables utilisables déterminées, encore faut-il vérifier qu'elles expliquent bien les variables pour lesquelles on désire traiter la non-réponse. Pour cela, on peut mettre en place des modèles logit, probit ou polytomiques emboîtés prenant pour variables expliquées chacune des 22 variables et pour variables explicatives les variables auxiliaires retenues. On vérifie ainsi une significativité globales des variables auxiliaires comme variables explicatives.

2.3 Les méthodes de hot-deck

Présentation

Le principe de la méthode finalement retenue pour le traitement de la non-réponse utilise l'idée des sous-catégories présentée précédemment. Cependant elle se distingue des solutions déjà envisagées en ne s'occupant plus du résultat global obtenu pour chaque réponse. Le non-répondant, plutôt que d'être confronté à l'ensemble des répondants, l'est à un individu proche de lui suivant les informations disponibles. Il s'agit de remplacer la donnée absente par une donnée "plausible" issue d'une autre observation relativement semblable au niveau des critères déterminant la façon de répondre à cette question. Ce type de procédure est désigné sous le nom de **hot-deck par classe**. L'imputation par hot-deck regroupe plusieurs méthodes basées sur le concept de "donneur". La donnée manquante est remplacée par la valeur observée pour un individu répondant choisi au hasard. L'idée est de choisir une valeur à partir de la distribution empirique des répondants. Gestion de l'aléa et possibilité de prendre en compte les corrélations entre variables constituent les apports principaux de cette famille de méthode.

La dimension aléatoire de ce type de méthodes

Les différentes méthodes hot-deck que nous envisageons initialement sont en effet des versions stochastiques de la méthode de prédiction par la moyenne par classe. La valeur de l'aléa correspond à la déviation du répondant choisi comme donneur par rapport à la moyenne de sa classe. Dans ce cas particulier, cette méthode revient à repondérer les individus répondants (ou plus exactement à leur attribuer un poids aléatoire) ce qui permet de retrouver les formules de variance par l'intermédiaire des résultats obtenus dans le cadre de la repondération⁷.

⁶Encore faut-il noter qu'on impute une durée maximale à la fonction publique et à l'administration.

⁷Pour plus de détails, voir [4].

La prise en compte de la corrélation entre variables

Imputations à partir d'un donneur unique choisi au hasard

Pour tenir compte des relations entre les différentes variables à imputer, le même donneur peut être utilisé pour imputer simultanément toutes les variables non renseignées d'une même observation. On utilise ainsi la méthode du hot-deck aléatoire (voir [20] p.18). On impute simultanément les valeurs manquantes d'une observation par les valeurs prises par un unique donneur choisi au hasard dans la classe du receveur.

Imputation à partir des valeurs prises par le donneur le plus proche du receveur

La méthode précédente permet de tenir compte des éventuelles corrélations entre les données manquantes. L'utilisation d'un donneur unique permet d'attribuer une configuration observée de modalités à la place des données manquantes. Cependant cette méthode ne permet pas de prendre en compte les corrélations éventuelles entre données manquantes et données déjà disponibles pour l'individu non-répondant.

Il est envisageable d'utiliser un **hot-deck hiérarchisé**. La donnée manquante est remplacée par la valeur observée pour un individu répondant ayant les mêmes caractéristiques pour les variables $X_1 X_2 \dots X_k$. S'il n'en existe pas, elle est remplacée par une unité ayant les mêmes valeurs pour $X_1 X_2 \dots X_{k-1}$ etc. Se pose cependant une question d'ordre d'importance des valeurs (Quelle est la variable pour laquelle j'accepterais que les individus soient différents?). Cette méthode peut cependant être généralisée en n'attribuant pas forcément le même poids à l'ensemble des observations : on utilise alors la méthode du hot-deck métrique.

Le **hot-deck métrique** (voir [20] p.24) est une méthode d'imputation qui consiste à remplacer une valeur manquante pour un receveur par la valeur observée pour le donneur le plus proche, au sens d'une distance (plus exactement une mesure de similarité). La "distance" est calculée à partir des variables auxiliaires renseignées pour les répondants et les non-répondants. Cette distance doit être définie judicieusement de façon à "respecter" la corrélation entre les variables auxiliaires et la variable d'intérêt (en particulier en accordant plus d'importance aux variables les plus liées à la variable d'intérêt). Il s'agit de tirer parti de l'ensemble de l'information auxiliaire disponible. Les autres méthodes de hot-deck par classe restreignent souvent beaucoup l'utilisation de celles-ci car les classes doivent être de taille suffisamment importante (on doit alors limiter le nombre de modalités utilisées).

L'introduction de l'aléa est possible en retenant comme donneurs potentiels tous les individus dont la distance avec l'individu non-répondant considérée est inférieure à un certain seuil et en procédant à un tirage au sort. Mais encore faut-il que l'individu retenu comme donneur ne soit pas défaillant lui-même pour les mêmes variables. Il est possible de prévoir de recommencer la procédure de sélection dans ce cas. Nous envisagerons une mise en place pratique du hot-deck métrique qui contournera cette difficulté par une définition préalable des donneurs potentiels.

Conclusion

Le fait de pouvoir tenir compte des relations entre variables apparaît être un élément important. Cependant l'élaboration d'une typologie est déterminée par les relations de corrélations entre variables. En renforçant les corrélations existantes, ce mode de traitement de la non-réponse peut finalement fausser les analyses dès lors que les corrélations chez les non-répondants sont différentes de celles des répondants. Une façon de vérifier l'influence de ce type d'imputation sur le résultat de l'analyse peut-être d'utiliser à titre de comparaison une méthode d'imputation ne tenant justement pas compte des

corrélations entre variables.

L'utilisation du hot-deck métrique n'est pas la seule technique qui est utilisée pour traiter la non-réponse. Nous avons déjà vu que le mode de constitution des variables utilisées jouaient un rôle. Pour les variables filtrées, il arrive que les observations, malgré une information à la première question, ne fournissent pas de précisions par l'intermédiaire des suivantes. Dans ce cas, le traitement de cette non-réponse est particulier et fait intervenir la prédiction par un modèle de régression. C'est notamment le cas pour la variable ENFMALDD.

Certains individus ont répondu qu'ils bénéficiaient de jours de congés rémunérés en cas de maladie de l'enfant (ENFMAL=1 ou 2) mais n'ont pas précisé leur nombre (ENFMALNB=88, 88 indiquant l'absence de réponse⁸). Nous avons donc construit à partir des répondants à la seconde question un modèle linéaire permettant de prévoir le nombre de jours de congés rémunérés en cas de maladie de l'enfant⁹. Les variables explicatives sont le poids (la pondération est une approximation exhaustive de l'effectif), une variable à 8 modalités synthétisant des informations sur l'entreprise (public/privé, groupe ou non, nationalité), la durée du congé en cas de mariage, le maintien ou non du salaire lors du congé maternel ou paternel en cas de dépassement du plafond SECU. Le test effectué sur environ 7% de l'effectif répondant montre (voir figure 2.5 p. 26) que la prévision est correcte (appartenance à la bonne tranche de durée sur les 3 existantes) dans près de 60% des cas (le test est effectué plusieurs fois -5 fois ici- car la sélection de l'échantillon test est aléatoire).

Essai	R2 ajusté	Prévision exacte (part)
1	0.35	57%
2	0.36	69%
3	0.30	53%
4	0.30	53%
5	0.29	63%
Ensemble	0.31	59%

FIG. 2.5 – Tests de la prédiction obtenue par le modèle de régression

⁸Choisir de coder l'absence de réponse avec un autre nombre serait probablement préférable car les réponses sont comprises entre 1 et 94. Cependant on peut penser que cela affecte très peu l'ensemble puisque la fréquence pour des durées aussi élevées est au plus de 2.

⁹Il s'agit avant tout d'un modèle statistique, la valeur de chacun des coefficients ne nous occupant pas ici.

Chapitre 3

Mise en oeuvre d'un hot-deck métrique

3.1 Présentation des variables auxiliaires utilisées

On utilise 4 variables auxiliaires :

1. *classe* : le poids (en classes de déciles)
2. *tgroupe* : le type d'activité (public/privé, groupe, nationalité)
3. *nes16d* : le secteur d'activité
4. *rrrr* : la région (on a divisé à cet effet la France en 8 grandes régions¹)

3.2 Définition des donneurs

On utilise un unique donneur pour imputer les différentes valeurs manquantes d'une observation. Ne peut être donneur qu'une observation pour laquelle la variable en question est renseignée. Pour une observation donnée, la liste des configurations possibles pour l'imputation est une combinaison de 1 à 10 variables sur les 22 retenues (on a supprimé les observations ayant plus de non-réponses). Travailler sur chacun des cas possibles peut s'avérer non seulement rapidement fastidieux mais surtout impossible à gérer². Une solution simple serait de définir comme donneurs potentiels les observations pour lesquelles l'ensemble des informations sont renseignées (aucune non-réponse sur les 22 variables). On aurait ainsi 1420 donneurs pour 1232 receveurs. On peut cependant distinguer différents cas pour récupérer comme donneurs des observations n'étant pas totalement complètes. On a ainsi constitué trois groupes de receveur en se référant à deux séries de variables de même type : d'une part 5 variables de prestations financières (ou assimilables) PRESTAI, PRESTAA, PRESTAD, PRESTAE et PRESTAG, d'autre part les 4 variables d'aménagement d'horaire AMENA, AMENB, AMENC et AMEND. Ces trois groupes de receveurs sont définis selon les critères suivants :

- les deux séries de variables ne sont pas à imputer (492 observations).
- seules les deux séries de variables sont à imputer (363 observations)
- les observations à traiter restantes³ (377 observations)

¹Voir annexe B.2 p.78. Conserver toutes les régions conduit à obtenir de trop nombreuses classes vides.

²Le nombre de configurations possibles est en effet $C_{22}^1 + C_{22}^2 + C_{22}^3 + \dots + C_{22}^{10}$.

³Celles qui n'appartiennent pas aux deux premières catégories

Le premier groupe a pour donneurs les 1783 observations. Ceux-ci n'ont de non-réponse que parmi les deux séries de variables. Les donneurs du deuxième groupe sont les observations ne comportant pas de non-réponse pour les deux séries de variables (1912). Seules les observations ayant les 22 variables renseignées sont les donneurs du troisième groupe (1420).

3.3 Calcul des pondérations affectées aux différentes variables auxiliaires

Les pondérations affectées aux variables auxiliaires sont d'autant plus fortes que leurs corrélations avec les variables à imputer sont fortes. **Leur pondération est la somme normalisée des V de Cramer calculés avec l'ensemble des variables à imputer** (3.1 p. 29.). Notons que cette méthode présente l'inconvénient de ne pas prendre en considération les corrélations possibles entre les variables auxiliaires ce qui peut conduire à surestimer le poids attribué à celles-ci. Cependant les corrélations entre variables auxiliaires retenues sont très limitées. Mais si l'effet régional ne s'expliquait que par l'inégale répartition des secteurs d'activité sur le territoire⁴, alors l'effet de secteur d'activité serait non seulement pris en compte par la pondération attribuée à cette variable mais également à la pondération de la variable région. Par conséquent il convient d'être prudent dans le choix des variables auxiliaires, même si les marges de liberté sont des plus limitées en raison de leur faible nombre.

V de Cramer

Le V de Cramer est une mesure de la liaison entre deux variables qualitatives dérivée du χ^2 . Pour mesurer la liaison entre les variables X et Y ayant respectivement p et q modalités, on peut mesurer la distance entre les effectifs observés (N individus au total) et les effectifs attendus (notés N^t) sous l'hypothèse d'indépendance entre X et Y. On considère alors la mesure :

$$\chi_{emp}^2 = \sum_{i=1}^p \sum_{j=1}^q \frac{(N_{ij} - N_{ij}^t)^2}{N_{ij}^t}$$

avec N_{ij} nombre d'observation prenant les modalités i pour la variable X et j pour la variable Y; N_{ij}^t nombre d'observation théorique en cas d'indépendance entre les deux variables.

Cette mesure dépend du nombre d'observations et du nombre de modalités, ce qui ne permet pas d'effectuer des comparaisons de la liaison dans des tableaux de contingence différents. Sachant que $0 \leq \chi_{emp}^2 \leq N \min(p-1, q-1)$, on construit l'indicateur dit V de Cramer de la façon suivante :

$$V = \left(\frac{\chi_{emp}^2}{N \min(p-1, q-1)} \right)^{\frac{1}{2}}$$

⁴Ces deux variables sont en réalité relativement indépendantes, le V de Cramer ne s'élevant qu'à 0.14.

V de Cramer	classe	tgroupe	nes16d	rrrr
enfmaldd	0.1909	0.4267	0.4150	0.0978
congdudd	0.1001	0.1501	0.4340	0.0812
gardentrd	0.3715	0.3646	0.2344	0.0790
icreched	0.2160	0.1422	0.1296	0.0680
prestaidd	0.3089	0.4388	0.2972	0.1185
prestaad	0.2238	0.4288	0.3329	0.0983
prestaadd	0.2751	0.6024	0.4773	0.1013
prestaed	0.2236	0.4291	0.3041	0.0981
prestagd	0.2772	0.6472	0.4439	0.0429
amenad	0.0804	0.1629	0.1662	0.0898
amenbd	0.0890	0.1452	0.1810	0.0908
amencd	0.0738	0.1395	0.1659	0.0816
amendd	0.1146	0.1521	0.2005	0.1226
dombd	0.0843	0.2266	0.2430	0.0988
partaccd	0.1354	0.1854	0.1951	0.0811
partembd	0.0872	0.3410	0.4628	0.1236
rttddd	0.1768	0.2077	0.2207	0.1106
promocd	0.0999	0.1752	0.1882	0.0593
congdd	0.2168	0.5037	0.4638	0.1422
prestaidd	0.2869	0.3655	0.3154	0.0659
fmutud	0.1066	0.4972	0.4319	0.1070
mutud	0.1852	0.4137	0.3002	0.0857
TOTAL	3.924	7.1456	6.7822	2.0441

FIG. 3.1 – Calcul des pondérations affectées aux 4 variables auxiliaires

3.4 Prise en compte de l'information fournie par les variables potentiellement imputables

Parmi les 22 variables potentiellement imputables, les réponses à un certain nombre d'entre elles sont connues. Or si le hot-deck réalisé à partir d'un seul donneur permet de conserver les corrélations entre réponses manquantes, il ne prend pas en compte la corrélation qui peut exister entre variables manquantes et variables auxquelles l'individu a répondu. Pour prendre en compte ce type de corrélation, on introduit dans le calcul de la distance les 22 variables pondérées selon le même principe que les variables auxiliaires (somme normalisée des V de Cramer avec les 21 autres variables).

On peut alors calculer la distance entre un donneur potentiel et un receveur. Pour chaque variable, on crée une variable dichotomique qui prend la valeur 1 si les modalités des deux observations pour cette variable sont différentes, 0 sinon. Cette variable est pondérée par l'indice issu du calcul des V de Cramer (voir figure 3.2 p.30). La somme des 26 variables dichotomiques (4 correspondant aux variables auxiliaires, 22 se rapportant aux variables retenues comme actives pour la suite de l'analyse) est enfin normalisée par la somme des indices.

Variable	Indice		
classe	3.924	amenad	3.1260
groupe	7.1456	amenbd	3.1354
nes16	6.7822	amencd	2.7588
region	2.0441	amendd	2.7652
enfmaldd	5.5375	dombd	2.4427
congdudd	2.5136	partacd	3.2440
gardentrd	3.5398	partembd	2.4472
icreched	2.2278	rttddd	3.5968
prestaidd	5.0041	promocd	3.0060
prestaad	2.9040	congdd	4.5136
prestaidd	5.8510	prestahd	4.5398
prestaed	4.3356	fmutud	4.5552
prestaidd	5.7706	mutud	3.9129
		Total	101.6135

FIG. 3.2 – Pondérations pour le calcul de la distance

3.5 Bilan du traitement

Un donneur n'est utilisé qu'une seule fois dans plus de 60% des cas (voir 3.3 p.30). Lorsqu'il est utilisé plusieurs fois, dans 60% des cas il n'est utilisé que deux fois et dans 22% des cas trois fois. Seuls 13 donneurs sont utilisés plus de 5 fois.

Nombre d'utilisations	Effectif	Part (en %)
1	483	63.6
2	167	21.97
3	70	8.2
4	22	3.7
5	6	1.46
6	4	0.5
7	5	0.7
8	1	0.1
9	1	0.1
11	1	0.1

FIG. 3.3 – Nombre d'utilisations des donneurs

La base de données ainsi traitée présente des proportions différentes pour les différentes modalités par rapport à la base initiale ou par rapport à la base composée des seuls répondants (voir 3.4 p.32). Ainsi 36.90% des établissements répondants acceptent les aménagements d'horaire en fonction de ceux de la crèche ou de l'école (variable AMENBD). En procédant à une réallocation proportionnelle des répondants, cette part passe à 37.14% tandis qu'à l'issue de la procédure de hot-deck métrique elle est de 39.06%. Les écarts sont limités en proportion mais il convient de ne pas oublier que c'est le traitement de chaque établissement qui nous importe en raison de l'utilisation finale de la classification. Ainsi l'attention apportée dans ce tableau aux proportions globales et à des traitements de la non-réponse inappropriés pour l'objectif recherché (en particulier la méthode consistant à supprimer les non-répondants) n'a ici qu'une vocation illustrative. Pour une variable donnée, une différence de proportion de 2% conduit à une situation différente pour 53 établissements. Les corrélations entre variables en sont affectées. Ainsi pour les variables ENFMALDD et AMENDD, le V de Cramer est 0.09 à la

suite de la suppression des non-répondants contre 0.28 à la suite du traitement par hot-deck métrique⁵.

A l'issue de cette étape préliminaire de traitement de la non-réponse partielle, on dispose d'une base de données comprenant 22 variables complètement renseignées pour 2652 observations. Par rapport à la base initiale, seules 21 observations ont été perdues. La méthode de traitement de la non-réponse utilisée laisse la possibilité de les conserver. Ce n'est que dans un soucis de prudence que seules les observations dont la moitié des informations étaient renseignées ont été conservées. C'est à partir de cette base que va être réalisée la procédure de classification proprement dite.

⁵Les deux tableaux de contingence n'ayant pas le même nombre d'observation, on n'utilise pas directement le χ^2 mais un de ses dérivés.

Variable	Modalité	Procédure de hot-deck métrique (en %)	Réallocation proportionnelle (écart)	Elimination des non-répondants (écart)
ENFMALDD	0	32.01	- 1.41	- 3.98
	1-3	13.35	+ 0.16	+ 0.95
	4-6	22.59	+ 0.07	+ 1.14
	>6	32.05	+ 1.18	+ 1.89
CONGDURD	oui	8.18	+ 0.05	+ 0.2
GARDEDTT	oui	7.84	+ 0.06	-0.73
ICRECHED	oui	4.15	-0.29	+ 0.22
PRESTAID	oui	40.16	-0.03	+ 2.3
PRESTAAD	oui	65.84	- 0.04	+ 2.4
PRESTADD	oui	27.64	+ 0.2	+ 1.16
PRESTAED	oui	32.01	- 0.3	+ 0.67
PRESTAGD	oui	26.51	- 0.01	+ 1.52
CONGD	oui	70.59	+ 0.72	+ 0.04
AMENAD	oui	85.90	+ 0.05	- 0.48
AMENBD	oui	39.06	- 1.92	- 2.16
AMENCD	oui	71.68	+ 0.19	- 2.17
AMENDD	oui	23.64	- 2.53	- 2.02
DOMBD	oui	11.92	- 0.03	+ 0.97
PARTEMBD	oui	49.25	- 0.22	- 0.66
PARTACCD	non	8.33	+ 0.14	- 0.58
	parfois	42.12	+ 0.15	- 0.15
	toujours	49.55	- 0.29	+ 0.73
RTTDD	non	19.46	- 0.19	- 2.56
	contraint	14.71	+ 0.06	+ 0.78
	accord	53.73	+ 0.15	+ 0.57
	libre	12.10	- 0.02	+ 1.21
PROMOCD	jamais	25.19	- 0.57	+ 0.87
	parfois	34.01	+ 0.56	- 0.35
	souvent	31.37	+ 0.04	+ 0.32
	toujours	9.43	- 0.05	- 0.84
PRESTAHD	oui	50.04	+ 0.14	+ 3.06
FMUTUD	tous	55.02	- 0.12	- 0.09
	cadre	7.96	+ 0.14	0
	pas	37.03	- 0.03	+ 0.08
MUTUD	oui	31.64	- 0.17	+ 0.26

FIG. 3.4 – Répartition des effectifs dans les différentes modalités suivant le traitement de la non-réponse retenu

Deuxième partie

Classification

Chapitre 4

Éléments de méthode

4.1 Choix de la procédure de classification

Les méthodes de classification ont pour objectif de regrouper les unités étudiées dans un nombre limité de classes homogènes. Formellement ces méthodes ont pour objet d'opérer une partition de l'ensemble des unités étudiées à l'aide d'une mesure de dissimilarité définie sur cet ensemble¹. Loin d'être une procédure automatique, de nombreux choix sont laissés à l'initiative de l'utilisateur lors de la réalisation d'une classification². Le nombre de partitions possible d'un ensemble (nombre de Bell) est ainsi trop élevé pour étudier l'ensemble des différents regroupements possibles des 2652 établissements afin de sélectionner celui qui apparaîtrait optimal au regard d'un critère préalablement défini³. La recherche d'une partition optimale se fait donc de façon approximative à partir de différents algorithmes. On rappelle ici les principes des deux principales familles d'algorithmes et de deux procédures combinant celles-ci utilisées dans le cadre de ce travail. Les choix effectués pour leur mise en oeuvre sont ensuite présentée.⁴

Principe des méthodes de CAH

Les méthodes de classification ascendantes hiérarchiques procèdent à un regroupement successif de classes par agrégation, à chaque étape, des deux éléments de la partition précédente les plus proches. Initialement, chaque singleton de l'ensemble étudié (les établissements) constitue une classe (il y a ainsi 2652 classes). A la première étape, l'agrégation des deux classes les plus proches conduit à une partition en 2651 classes, à la deuxième étape on obtient 2650 classes etc. L'algorithme se termine avec la constitution d'une seule classe à la 2651e étape. Cet algorithme nécessite le choix d'un indice de dissimilarité entre individus et d'une stratégie d'agrégation basée sur la définition d'une dissimilarité entre parties disjointes de l'ensemble. A partir de la partition précédente, cette procédure permet de déterminer la partition optimale comprenant une classe de moins. Elle ne permet cependant pas d'obtenir une partition qui soit un optimum global.

¹Une mesure de dissimilarité est une application δ de $I * I$ dans \mathfrak{R}^+ caractérisant les dissemblances entre les individus vérifiant les propriétés : $\delta(i, i) = 0$ et $\delta(i, j) = \delta(j, i)$.

²Voir [1].

³Dès qu'un ensemble est composé de 15 éléments, le nombre de partitions possibles de celui-ci dépasse le milliard.

⁴Les questions de mise en place pratique d'une classification (choix des variables, nombre d'axes retenus, nombre de classes) qui vont être abordées à partir de maintenant rejoignent celles qui l'ont été par R. Le Lan ([11]) lors des *Journées de Méthodologie statistique 2005*. Nous n'avons cependant pu prendre connaissance de son travail qu'après la fin de notre étude. Celui-ci met en évidence à partir d'un exemple l'influence du nombre d'axes retenu pour l'analyse et il suggère des idées d'outils d'aide à la décision.

Méthode des centres mobiles (méthode de partitionnement)

Les méthodes de classification par agrégation autour des centres mobiles opèrent par réallocation des observations à des centres de classes recalculés à chaque itération. L'initialisation de l'algorithme consiste à choisir aléatoirement des individus, dont le nombre, fixé préalablement, correspond au nombre de classes qui seront obtenues au final. On affecte alors à chaque classe l'individu le plus proche de son centre (à la première étape, la classe n'est constituée que de l'individu sélectionné). Le nouveau centre de la classe qui résulte de cette agrégation est calculé et l'individu le plus proche de celui-ci est affecté à cette classe, ce qui modifie à nouveau le centre de la classe. Chaque individu est alors affecté à la classe dont le centre est le plus proche, ce centre étant recalculé etc. La solution obtenue est un optimum local : elle dépend du choix initial des centres de classes. La variabilité des résultats obtenus peut être limitée par la recherche des **formes fortes**. Cela consiste à s'assurer de l'obtention de façon stable de classes ou portions de classes quel que soit le choix des centres initiaux. Une forme forte est ainsi constituée des observations se retrouvant systématiquement dans le même groupe lorsqu'on répète l'algorithme (en général deux ou trois fois).

La combinaison des méthodes : classification mixte et consolidation

Combinaison des deux principes de classification précédents, **la classification mixte** consiste à procéder à la méthode des centres mobiles à partir d'un nombre important de centres initiaux, puis, à partir des classes obtenues, à procéder à une classification ascendante hiérarchique.

L'utilisation de **la consolidation** à l'issue d'une CAH est également une forme de combinaison des méthodes qui permet d'accroître la qualité de la classification en améliorant l'inertie inter-classes⁵. Elle consiste à procéder à une réallocation des individus à partir des classes obtenues, en reprenant le principe des centres mobiles. La procédure de réaffectation et de détermination de nouveaux centres de gravité est répétée jusqu'à ce que la redistribution des individus n'entraîne qu'un gain d'inertie inter-classes négligeable.

4.2 Mise en oeuvre

Nous résumons dans l'énumération suivante les caractéristiques de la procédure de classification utilisée. Certaines justifications sont présentées directement à la suite. Les autres sont données dans les deux chapitres de présentation des résultats qui suivent.

Caractéristiques de la classification

- *le choix des variables et de leurs modalités : **21 variables, 50 modalités.***
- *le format des variables : **variables continues** (à l'issue d'une ACM préalable)*
- *le nombre d'axes de l'ACM préalable retenus : **10 axes.***
- *une mesure d'éloignement (dissemblance, dissimilarité ou distance) entre les individus : **distance euclidienne.***
- *la stratégie d'agrégation : **méthode de Wald.***
- *la méthode de classification : **classification ascendante hiérarchique suivie d'une consolidation.***
- *On a également utilisé la procédure de **classification mixte à titre de comparaison.** Dans ce cas, il faut choisir le nombre de classes retenu (on a généralement retenu les nombres de 20 ou 30) et le nombre d'itération utilisées pour la définition des formes fortes (généralement 5).*

⁵Distance entre les différentes classes.

Utilisation des résultats de l'ACM

Nous travaillons à partir de variables qualitatives. Il serait donc tout à fait envisageable de procéder à une classification à partir de celle-ci en choisissant des indices de dissimilarité de type χ^2 . Cependant on peut également utiliser les projections des individus sur les axes de l'ACM et travailler ainsi à partir de variables quantitatives. Il est en effet équivalent d'effectuer une classification à partir de l'ensemble des variables ou à partir de l'ensemble des facteurs issus de l'analyse factorielle. On ne perd aucune information : il ne s'agit que d'un changement de repère. Il est cependant envisageable de ne retenir qu'une partie de ces facteurs pour l'établissement de la classification⁶. Cela permet en particulier d'éliminer les effets dus aux "bruits" : on peut en effet s'attendre à ce que les conséquences dans la détermination de l'axe des erreurs commises lors de l'enquête sont d'autant plus grandes que le numéro de celui-ci est élevé. Contrairement au logiciel SAS, SPAD procède systématiquement à cette ACM préalable.

Choix du critère de dissemblance et de la méthode d'agrégation

Nous procédons à la **classification à partir de données rendues de type quantitatives**. La dissemblance entre les individus est calculée à partir de la distance euclidienne usuelle. La dissemblance entre deux sous-ensemble disjoint est calculée à partir du **critère de Ward** de perte d'inertie minimum. Il s'agit de minimiser la perte d'inertie intra-classe⁷ lors de l'agrégation.

On se place dans un espace euclidien. Soit deux groupes A et B de poids p_A et p_B et de centre de gravité g_a et g_b tels que $p_A = \sum_{i \in A} p_i$ et $g_A = \frac{1}{p_A} \sum_{i \in A} p_i x_i$. On définit alors la distance $d^2(A, B)$ par :

$$d^2(A, B) = \frac{p_A p_B}{p_A + p_B} \|g_a - g_b\|^2$$

Les méthodes de classification utilisées ici sont la CAH et la classification mixte. La première méthode aura notre préférence pour son caractère déterministe. Cependant les résultats obtenus à partir de la seconde influenceront dans le choix du nombre d'axes et de classes retenus.

Le choix du logiciel utilisé pour la classification

Le choix est souvent lié à ceux qui précèdent quand il ne les conditionnent pas... Il convient cependant d'être conscient de l'influence du choix du logiciel sur les résultats obtenus (options par défaut mais aussi algorithmes potentiellement différents⁸). On utilisera ici les logiciels SAS et SPAD. Ce dernier présente l'avantage de fournir des sorties graphiques plus agréables à lire. L'ensemble du traitement des données est cependant effectué sous SAS : création des variables, définition des formats (ie nom des modalités) etc. L'analyse de données réalisée en parallèle sous SAS permet de confronter les résultats obtenus par les deux logiciels. On utilise à cet effet les macros d'analyse de données mises à disposition par l'INSEE⁹. Cependant la transmission d'une base de données de SAS à SPAD n'est pas forcément aisée dès lors qu'on souhaite conserver les propriétés définies sous SAS (en particulier les formats)¹⁰.

⁶Voir [16] p.8.

⁷Distance entre les éléments d'une même classe.

⁸Voir à ce propos [9].

⁹Voir [8].

¹⁰Pour le passage de SAS à SPAD, voir annexe E p.141

Chapitre 5

Réalisation préalable d'une ACM

La base de données a donné lieu à la réalisation de différentes ACM suivies de classifications. A la suite d'une procédure d'ensemble à vocation exploratoire, il a été procédé à quelques modifications en raison des trop faibles effectifs de certaines modalités. En effet, l'inertie d'une modalité par rapport au centre de gravité est d'autant plus élevée que cette modalité est rare. Ainsi l'inertie I d'une modalité X_j^m par rapport au centre de gravité (avec M le nombre de variables et f_j^m la fréquence de la modalité X_j^m) est $I(X_j^m) = \frac{1}{M}(1 - f_j^m)$. Et le faible poids affecté à une modalité rare ne suffit pas toujours à compenser son éloignement du centre de gravité¹.

La variable codant l'intention de créer une crèche a été éliminée de l'analyse. Ses effectifs faibles pour la modalité positive conduisait à lui faire jouer un rôle trop important dans l'analyse par rapport à ce qu'elle recouvre : une intention déclarée alors que toutes les autres variables se rapportent à des mesures effectives. D'autre part, les faibles effectifs pour les deux modalités codant la durée du congé maternité supplémentaire ont finalement été fusionnées : on n'utilise ainsi plus qu'une variable binaire².

Ces changements ont un impact quasi-nul sur l'ACM (modification d'un dixième des coordonnées de projection des autres modalités). Par contre, le premier a un impact important sur la typologie obtenue : une classe se caractérisait par la seule réponse positive à la question.

5.1 L'axe 1 des droits à des dispositifs de conciliation

L'analyse de l'histogramme des valeurs propres (voir annexe C.1 p.101) conduit à ne retenir pour la présente analyse que **les trois premiers axes qui expliquent à eux seuls près de 30% de l'inertie**. L'inertie totale du nuage³ est égale au nombre moyen de modalités diminué d'une unité, soit 1.3810.

Le premier axe y contribue pour moitié (14,53%)⁴. Pour cet axe, quatre variables sont à l'origine de plus de 50% de l'inertie expliquée. Il s'agit (dans l'ordre décroissant) de l'existence d'aide aux frais de garde, d'aide en cas de handicap de l'enfant, de la durée de congé rémunéré en cas de maladie de l'enfant et de l'accès aux colonies de vacances. Leurs projections sont orientées dans le même sens.

¹Voir chapitre 4[18].

²Ce changement conduit à une diminution de l'inertie de cette variable par rapport au centre de gravité car l'inertie diminue avec le nombre de modalités de la variable : $I(X^m) = \sum_{j=1}^{p_m} I(X_j^m) = \frac{1}{M}(p_m - 1)$.

³Elle ne dépend pas des liaisons entre les variables et n'a pas d'interprétation statistique (voir [18]).

⁴La part d'inertie expliquée par le premier axe ne peut dépasser $1/I$ (avec I l'inertie totale du nuage) soit ici 72%. Ceci explique que les parts d'inertie expliquées sont toujours faibles pour les ACM et donnent une vision "pessimiste" de la quantité d'information extraite (voir [18]).

La possibilité de garde des enfants est un aspect important dans la détermination de l'axe. L'existence d'un tel dispositif au sein même de l'entreprise a d'ailleurs la coordonnée la plus extrême. La possibilité de travailler à domicile en cas de contraintes personnelles, l'acceptation du passage à temps partiel dès lors qu'il est demandé, la gestion personnelle des congés pour RTT apparaissent du même ordre. Cette aide aux enfants à une acception assez large, de la mise à disposition de structures aux aides financières. Le complément en cas de congé parental, les chèques vacances, l'aide aux études, le financement des mutuelles sont projetés de façon similaire.

Un point commun à ces différentes mesures est qu'il s'agit de dispositifs automatiques, de droits du salarié et pas simplement d'arrangements ou de faveurs accordées au cas par cas selon le bon vouloir de la direction⁵. L'existence d'arrangements pour imprévu ou situation particulière ne se projette pas sur l'axe et les quatre variables correspondantes ont même tendance à se comporter de façon inverse à l'exception de la mesure "jour de rentrée". Avec les services mutualistes (à la différence du financement de la mutuelle), ce sont les seules mesures orientées dans ce sens.

Par conséquent le premier axe discriminerait selon l'existence ou non de droits au sein de l'entreprise relevant de la conciliation vie professionnelle-vie familiale. Les mesures de conciliation dont l'attribution semble se faire de façon discrétionnaire sont des mesures de portée limitée (quelques heures d'ajustement et non pas la possibilité de travailler chez soi, recours possible aux services de la mutuelles et non aide en cas de handicap de l'enfant) qui ne se projettent pas sur l'axe voire tendent à se projeter en sens inverse.

5.2 L'axe 2 des aménagements horaires

L'axe 2 est déterminé justement par la possibilité d'avoir recours à des aménagements horaires : les quatre variables caractérisant ces dispositifs sont à l'origine de 80% de l'inertie. A celles-ci s'ajoute la projection de la variable relative aux horaires de travail : plus la liberté du salarié augmente, plus il tend à se projeter du côté des aménagements d'horaire. Mais c'est l'absence de RTT qui se projette le mieux de ce côté là, tout comme la possibilité de travailler à domicile.

Aucune autre mesure ne se projette nettement du même côté que les quatre mesures d'aménagement horaire à deux exceptions près : l'existence de postes à temps partiel et la possibilité de passer à ce type de contrat dès lors que le salarié en fait la demande. Ainsi le second axe discriminerait les établissements selon l'existence ou non d'aménagements d'horaire pour des cas particuliers. On peut noter que les mesures de type temps partiel se projettent à proximité des arrangements de type temporaire et conditionnel.

5.3 L'axe 3 des mesures moins directement orientées vers les parents

Sur l'axe 3, trois variables sont à l'origine de plus de 50% de l'inertie expliquée : le financement d'une mutuelle, la gestion libre du temps issu de la RTT et l'existence d'une prime de naissance se projettent dans le même sens. L'existence de postes à temps partiel se projette en sens inverse. Mais chèques vacances, aménagements horaires, aide au financement des études, accès aux colonies de vacances, et congé maternité supplémentaire sont orientés dans le même sens que les

⁵Notons cependant que la formulation des questions concernant les mesures financières ou assimilables (variable ayant pour préfixe "Presta") ne permet pas de vérifier cette idée. Elles portent sur la possibilité pour certains salariés (et non tous) de bénéficier de telles mesures ("Certains salariés peuvent-ils bénéficier etc").

premières variables.

Ce troisième axe semble déterminé par des variables moins directement liées à la gestion des enfants. Il s'agit de mesures de type plus générales, moins orientées en direction des seuls parents.

On trouvera en annexe le détail des résultats de cette ACM ainsi que les projections de diverses variables supplémentaires sur les trois axes précédemment décrits (annexe C.4 p.108).



FIG. 5.1 – Plan 1-2 de l'ACM

Chapitre 6

Une classification en 7 classes

6.1 Présentation générale

On a procédé à deux types de classification. La première par ascendante hiérarchique, la seconde en utilisant une méthode mixte¹. La seconde méthode de classification n'est utilisée qu'à titre de comparaison. Dans les deux cas, on procède à une consolidation. La stratégie d'agrégation retenue est celle de Wald.

L'utilisation dans l'analyse de variables dont certaines modalités ont des effectifs très faibles affecte fortement la classification. Une classification effectuée sur l'ensemble des axes de l'analyse tend à conduire à deux types de typologie. Une première en 3 classes, caractérisée par une classe de très faible effectif correspondant à la prise en charge de la garde des enfants au sein de l'entreprise, une classe intermédiaire qui regroupe la majorité des effectifs et une classe marquée par l'absence de toute mesure à l'exception du financement de la mutuelle. La seconde proposition est une classification en 6 ou 7 classes caractérisées par deux classes intermédiaires d'effectifs élevés tandis que les autres classes regroupent chacune 5 ou 6% des effectifs autour d'une unique mesure. Ces modalités partagées par tous les établissements de la classe sans exception sont respectivement l'existence de dispositifs de garde au sein des établissements, l'existence d'un congé maternité supplémentaire, le financement de la mutuelle pour les cadres, l'absence de postes à temps partiel. Par conséquent, tandis que la première classification reste d'un intérêt limité, la seconde est principalement due à l'inertie élevée de certaines modalités engendrée par leur rareté. Il est généralement recommandé de ne pas utiliser de variables ayant des écarts d'effectif entre les modalités trop importants afin d'éviter ces effets de "distortion". Cependant dans le cadre de notre analyse, la majorité des modalités correspondant à l'existence d'une mesure sont dans cette situation. Par conséquent, on a choisi de garder les 21 variables de l'analyse mais de ne procéder à la classification qu'à partir d'un nombre limité d'axes.

Le choix courant de ne retenir qu'un nombre limité d'axes de l'ACM pour réaliser une classification est généralement justifié par l'accroissement de l'importance prise dans l'analyse par le "bruit"² avec l'augmentation du nombre d'axes. Dans le cas présent, on pu également remarquer qu'au delà d'un certain nombre d'axes, chacun d'entre eux tend à s'organiser autour d'une seule variable présentant généralement une modalité de faible effectif. Limitation du "bruit" et du poids acquis par certaines modalités du fait de leur faible effectif ont conduit à ne pas conserver tous les axes de l'ACM³.

¹On récupère les 20 à 30 formes fortes issues de 3 à 5 tirages des points de départ initiaux.

²Impact des erreurs, qu'elles proviennent des répondants, des enquêteurs, des codeurs, de ceux qui ont manipulé la base etc

³Un autre argument en faveur de la limitation du nombre d'axes retenus serait la diminution des différences de sauts d'inertie entre les partitions (Voir [11]). Ce phénomène rend moins évident le choix du nombre de classes qui peut avoir devenir grand

La réalisation de classifications successives sur un nombre décroissant d'axes permet d'observer certaines régularités. Ainsi on peut noter que les nombres de 6 ou 7 classes apparaissent le plus souvent (notamment avec la classification mixte). Ces 6 ou 7 classes se répartissent en trois grands groupes : un premier caractérisé par l'absence de mesure comprenant 2 ou 3 classes, un second intermédiaire (2 classes), un troisième caractérisé par l'existence de nombreuses mesures (2 classes). On observe que plus le nombre d'axes est élevé, plus ces classes non intermédiaires se recomposent autour d'une variables : les classes extrêmes ont alors des effectifs faibles (6 à 7%) marqués par une modalité partagée par tous sans exception. Les deux classes intermédiaires voient au contraire leurs effectifs augmenter dans ce cas.

Nous avons finalement retenu le nombre de 10 axes. Ce nombre peut paraître limité mais ces dix axes suffisent à expliquer plus de la moitié de l'inertie tandis que chaque axe supplémentaire explique moins de 3.4% de l'inertie. Les critères ayant présidés à ce choix sont multiples. En particulier, on a cherché à obtenir une structure de classes peu sensible au choix du nombre d'axes (structure "constante") comme au choix de la méthode de classification⁴. En retenant 10 axes, une classification mixte (20 formes fortes issues de 5 tirages) propose 3,6,7 classes tandis que la classification hiérarchique permet d'envisager la constitution de 6 ou 7 classes. **Nous retenons finalement 7 classes au regard du dendogramme** (voir figure D.1 p.125) mais aussi de l'analyse des classifications envisageables avec un nombre différent d'axes⁵.

Les résultats présentés et qui seront utilisés sont ceux issus de la classification ascendante hiérarchique : cette méthode présente en effet l'avantage d'être déterministe (on obtient toujours le même résultat).

Les projections des barycentres des sept classes (voir figures 6.3 p.50 et annexes D.5 et D.6 129) ainsi que les représentations graphiques de la dispersion de celles-ci sur les premiers axes (voir figures 6.4 p.51 et 6.5 p.53) ont été réalisées.

6.2 Description des classes obtenues

La structure emboîtée en 3, 6 et 7 groupes est relativement visible avec 10 axes. Ayant retenu la classification en 7 classes, nous ferons donc la présentation de celles-ci tout en les organisant en trois groupes correspondant globalement aux trois grandes classes. Les 6 classes sont issues de la fusion des classes 1 et 2 au sein de l'ensemble caractérisé par l'absence de mesures de conciliation.

⁴A titre de comparaison, l'annexe D.2 p.132 présente les résultats obtenus en retenant 28 axes.

⁵Pour 11 axes, la CAH propose 3 ou 7 classes. Entre 8 et 13 axes la classification mixte propose presque systématiquement 7 classes, cette structure en 7 classes se retrouvant également pour un nombre d'axes retenus plus élevé (notamment pour 28 axes). A titre de comparaison, les figures D.7 p.130 et D.8 p.131 présentent en annexe des résultats obtenus avec une partition en 6 classes

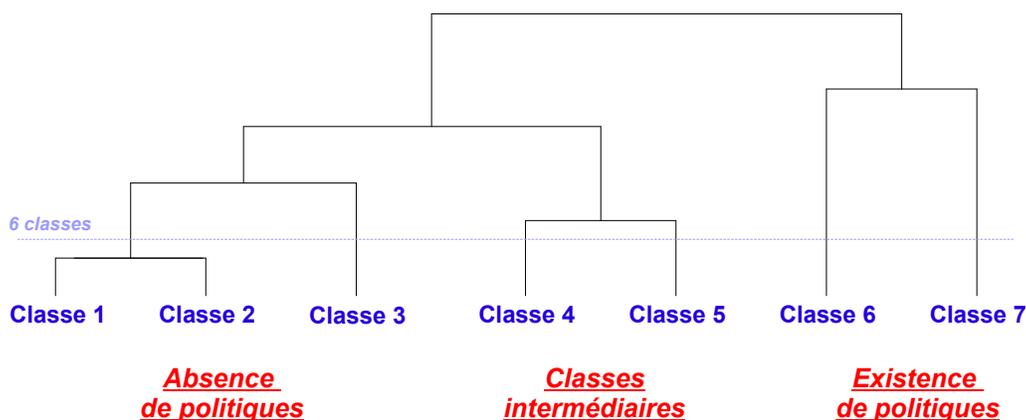


FIG. 6.1 – Structure de la classification

L'existence de politiques de conciliation : les classes 6 et 7

La classe 7 du paternalisme bienveillant

La classe 7 paraît regrouper les établissements dans lesquels tout est fait pour permettre la présence des salariés à leur poste de travail. Cette classe se caractérise ainsi par la présence plus fréquente de l'ensemble des mesures répertoriées à l'exception notable des aménagements horaires (hors rentrée scolaire).

L'absence relative des aménagements horaires paraît compensée par d'autres dispositifs (voir figure 6.9 p.58), de la mise à disposition de structure de garde à l'attribution d'aides financières. Les possibilités de garde dans l'établissement y sont ainsi 5 fois plus présentes qu'en moyenne (dans 30% des cas contre 8%). Les aides financières, en particulier les aides aux frais de garde et prestation enfants-handicapés, sont quasi-généralisées (plus de 90% des établissements contre 30%). L'attribution du complément de salaire en cas de congé parental est quasi-systématique en cas de dépassement du plafond de sécurité sociale. Mais l'absence de congé maternité supplémentaire est tout aussi général. Ainsi les mesures en cas de naissance d'un enfant semblent résumer la politique en place dans ces établissements à travers le maintien du salaire quel que soit son niveau et non l'attribution de congés supplémentaires.

Cette classe paraît se distinguer par une politique de conciliation qui permet une présence du salarié au travail sauf cas de force majeure comme la maladie d'un enfant : le congé possible est alors de plus de 6 jours. Ces mesures privilégient prestations monétaires et mise à la disposition de structures de façon à assurer la présence du salarié à son poste de travail.

On peut interpréter cette situation comme une forme de paternalisme moderne conciliant à la fois la mise à disposition de structures et le versement de prestations. Dès lors que le temps de travail est fixé (passage à temps partiel est possible sur simple demande dans plus de 60% des établissements, prise de congé RTT négociée), tout est fait pour permettre au salarié de s'y tenir par l'intermédiaire de mesures établies par avance.

Les variables supplémentaires tendent à confirmer ce premier constat. Dans cette classe constituée pour plus d'un tiers par des établissements relevant de la fonction publique hospitalière, les plages d'ac-

tivité sont relativement étendues. Le travail le dimanche (25% des établissements), la nuit, le samedi sont des situations fréquentes dans cette classe. La seule exception à cette tendance visant à favoriser la présence du salarié sur le lieu de son travail est, outre le congé en cas de maladie de l'enfant, le congé attribué pour mariage relativement long (plus de 5 jours dans 70% des cas).

La classe 6 de la présence au travail modulable

Dans la classe 6, la conciliation prend la forme d'une adaptation du temps et des horaires de travail. Moins institutionnalisées, les mesures de conciliation paraissent plus élaborées au cas par cas.

Cette classe se distingue avant tout (voir figure 6.8 p.57) par la possibilité de bénéficier de congés prolongés. Ainsi dans plus de la moitié des cas (contre 8% en moyenne), un congé supplémentaire est accordé pour la naissance du premier enfant. Le choix dans la prise des congés RTT est libre dans plus de la moitié des cas (contre 12%). De même, les aménagements d'horaire pour trajet, pour problème de garde ainsi que la possibilité de travailler à domicile sont un peu plus fréquents. Il est plus fréquemment possible de passer à temps partiel sur simple demande (tandis que la mise à disposition de poste à temps partiel à l'embauche est plus rare).

L'aide financière est présente dans plus de la moitié des cas aussi bien au niveau des frais de garde que des frais d'étude. Prime de naissance, financement de la mutuelle, chèques vacances et accès aux colonies sont de même présents dans plus de 70% des établissements.

La conciliation ne se traduit donc pas par la mise à disposition de structures particulières mais par des possibilités d'aménagements de temps de travail en cas de difficultés et par des prestations financières relativement développées. La conciliation passe avant tout par le caractère modulable des horaires de travail. Elle se distingue également de la précédente classe par le caractère moins établi des mesures. Il s'agit ici plus de traitement des difficultés au cas par cas, de souplesse d'horaire tolérée par l'établissement en cas de contraintes familiales particulières.

Les variables supplémentaires révèlent que le travail se concentre sur 5 jours dans 80% des cas. C'est la classe des établissements financiers, des administrations centrales et des sièges sociaux. 80% des établissements financiers sont dans cette classe. Ils représentent plus de 20% des établissements de la classe contre moins de 3% en moyenne. Le secteur "public ou nationalisés" est deux fois plus fréquents qu'en moyenne. Les établissements ayant les effectifs les plus élevés sont sur-représentés : 50% ont plus de 500 salariés et 30% plus de 1000 salariés. La proportion de jeunes est faible mais la part de cadre est importante : plus d'un quart des effectifs dans 50% des cas et plus de 50% des effectifs dans 30% des cas. La proportion de femmes ayant le statut de cadres l'est dans une moindre mesure alors que la part de femmes est plutôt élevée (plus de 50% des effectifs dans 40% des cas). Enfin les établissements d'Île de France sont sur-représentés : 35% des établissements de cette classe sont situés dans cette seule région.

L'absence de politique de conciliation : les classes 1, 2 et 3

La classe 3 des établissements masculins ignorant la vie familiale

L'existence de contraintes familiales est totalement ignorée dans ces petits établissements dans lesquels les femmes sont très peu présentes.

A ce titre il est particulièrement significatif qu'aucun de ces établissements ne propose de postes à temps partiels (voir figure 6.7 p.56). Mais elle se distingue également par une absence plus fréquente de toutes les mesures. Ainsi aucun des établissements n'a un dispositif de garde des enfants. L'avancement est également plus fréquemment dépendant de la disponibilité. La conciliation apparaît ainsi

hors du champs de préoccupation de l'entreprise et est au contraire valorisée l'absence d'impact de la vie familiale sur le travail.

Cela ne signifie pas pour autant l'absence de prestations au delà du salaire mais celles-ci sont d'un autre type. Les cadres bénéficient deux fois plus fréquemment d'un financement de leur mutuelle. L'accès au logement social (variable supplémentaire) est facilité dans 80% des établissements de cette classe.

Le secteur de la construction est près de 3 fois plus présent dans cette classe. Les secteurs de la métallurgie, du commerce de gros sont également sur-représentés. L'ignorance de la situation familiale est quasi-généralisée. Il s'agit généralement de petites structures : 50% des établissements ont entre 20 et 49 salariés. Les syndicats sont absents dans plus de 50% des cas. La relative absence des femmes est marquée : dans 50% des cas, la part de femmes est inférieure à 25% de femmes. La présence de DRH est rare. Le travail se concentre sur 5 jours. Le directeur est plutôt jeune, les cadres sont peu nombreux, le CDI est la règle.

La classe 2 : les établissements ignorant les dispositifs de conciliation

L'ensemble des dispositifs de conciliation sont généralement absents de ces établissements. Ils se distinguent de la première classe par leur caractère plus diversifié et lié à aux marchés étrangers.

La classe 2 se distingue de la troisième au niveau de plusieurs dispositifs (voir figure 6.6 p.55). L'absence de postes à temps partiel reste sur-représentée, mais dans une bien moindre mesure. Dans le même temps, le passage à temps partiel est possible dans plus de deux tiers des cas. D'autre part, des dispositifs de financement de la mutuelle (pour tous ou seulement les cadres) existent dans plus de 80% des établissements. L'avancement y est seulement "parfois" lié à la disponibilité. Par contre, même une mesure quasi-généralisée tel que l'aménagement d'horaire le jour de la rentrée est absente dans la moitié des cas.

Comme dans la classe précédente, les secteurs de la construction, de la métallurgie sont sur-représentés ainsi que ceux de l'industrie agro-alimentaire, de la chimie, des services opérateurs. La situation familiale des salariés n'est pas connue de la direction, les femmes très peu présentes. Les syndicats sont généralement absents. Cette classe se distingue notamment par le marché auquel elle s'adresse. Dans cette classe, les établissements sont plus fréquemment possédés par des groupes étrangers et s'adressent plus au marché européen. Finalement cette classe se caractérise par le caractère plus diversifié des établissements ignorant la vie familiale qu'elle regroupe et l'ouverture de ceux-ci vers des marchés plus étendus.

La classe 1 des dispositifs réservés aux situations exceptionnelles

La classe 1 est elle aussi marquée par l'absence de mesures mais se distingue par l'existence de dispositifs permettant au salariés de faire face à certaines situations particulières telles que la maladie d'un enfant ou la rentrée scolaire.

Dans la quasi-totalité des cas, des jours de congés rémunérés ainsi que des aménagements horaires sont accordés en cas de maladie de l'enfant (voir figure 6.6 p.55). La gestion des RTT fait l'objet de négociation avec l'employeur, la prime de naissance et le financement de la mutuelle sont accordés dans 75% des cas.

Commerce de détail, chimie, industrie d'équipement des foyers sont les secteurs sur-représentés dans cette classe. On y retrouve plus fréquemment les groupes français ou les établissements privés non lucratifs. Le travail de nuit, le travail le dimanche (25% des établissements est assez fréquent. Syndicats et drh sont présents plus fréquemment.

Les classes médianes : les classes 4 et 5

La classe 4 des aménagements horaires possibles mais pénalisants

Tout en permettant une forme de conciliation par l'intermédiaire des aménagements des horaires, ces établissements semblent limiter le recours à ces dispositifs par le caractère pénalisant de leur usage voire discrétionnaire de leur attribution.

La classe 4 (voir figure 6.7 p.56) se caractérise par les possibilités importantes de bénéficier d'aménagements horaires. En cas de maladie de l'enfant, alors qu'aucun jour de congé n'est rémunéré, l'aménagement horaire est possible dans plus de 97% des cas. Cependant l'avancement est plus fréquemment lié à la disponibilité : dans plus de 50% des cas, il est déclaré dépendre "souvent" voire "toujours" de la disponibilité. Financement de la mutuelle et possibilité de passer à temps partiel sont des dispositifs un peu plus fréquents. Les autres mesures, notamment monétaires, sont très fréquemment absentes. Ainsi les établissements de cette classe, tout en permettant dans une certaine mesure la conciliation par l'intermédiaire de possibilités d'aménagements horaires, paraissent la limiter : ces aménagements sont attribués au cas par cas et le recours à ceux-ci nuit directement à la carrière.

Métallurgie, commerce de gros et de détail, construction : autant de secteurs relativement présents dans cette classe. L'absence de RTT est fréquente tout comme le travail de nuit (contrairement au travail le dimanche). Les syndicats ont tendance à être absents de ces établissements composés de nombreux jeunes : les moins de 20 ans (plus de 5% dans 16% des cas), 20-29 ans (plus de 20% dans plus de 50% des cas), peu d'anciens. La région bourguignonne est légèrement sur-représentée.

La classe 5 des aménagements sur un modèle "Education nationale"

Cette classe regroupe les établissements procurant aux salariés des prolongations de congés, des possibilités de gestion de leur temps de travail et un accès aux services mutualistes. Il s'agit avant tout des établissements de la fonction publique et plus spécifiquement du secteur de l'éducation.

Présentant de nombreuses similitude avec la précédente classe, la cinquième (voir figure 6.8 p.57) s'en distingue par l'existence de mesures supplémentaires. La durée de congé en cas de maladie de l'enfant est plus longue (dans 70% des cas, elle est supérieure à 6 jours), dans près de 70% des cas, le passage à temps partiel est accordé sur simple demande, tandis que des postes de ce type existent plus fréquemment à l'embauche. Le travail à domicile est possible pour convenance personnelle trois fois plus fréquemment tandis qu'un accès aux services des mutuelles concerne la moitié des établissements. Enfin le complément pour congé parental est systématiquement accordé en cas de dépassement du plafond de sécurité sociale.

Fonction publique d'État (60% contre 13% en moyenne) et notamment le secteur de l'éducation (45% contre 8%) sont fortement présents dans cette classe. Elles regroupe des établissements composés d'une forte proportion de cadres, travaillant plus fréquemment moins de 35h. La présence féminine est importante notamment aux postes de direction. La proportion de CDI est plus limitée mais il y a plus rarement un recours à la sous-traitance et les effectifs sont plus fréquemment stables.

6.3 Synthèse des résultats

Récapitulatif des différentes classes

A l'issue de cette classification, les 2652 établissements sont répartis dans les 7 classes⁶.

⁶On rappelle que 22 établissements ont été supprimé lors du traitement de la non-réponse. La technique utilisée laissait la possibilité de conserver l'ensemble des observations mais il a été choisi de ne pas retenir les observations dont plus de la moitié des variables étaient non-renseignées

Catégorie	Numéro	Dénomination	Part	Pondérée
L'existence de politiques de conciliation	7	Le paternalisme bienveillant	20,40%	12.43%
	6	Le temps de travail modulable	10,75%	6.01%
Les situations intermédiaires	4	Les aménagements horaires possibles mais pénalisant	17,04%	23.48%
	5	Les aménagements sur un modèle "Education nationale"	10,37%	10.86%
L'absence de politiques de conciliation	1	Les dispositifs pour les seules situations exceptionnelles	18,44%	16.67%
	2	Les établissements ignorant les dispositifs de conciliation	16,55%	19.55%
	3	Les petits établissements masculins ignorant la vie familiale	6,45%	11.00%

FIG. 6.2 – Récapitulatif des classes

Éléments sur la robustesse des classes

Pour tester la robustesse, on met en place un modèle polytomique non ordonné prenant pour variable expliquée la classe et pour variables explicatives l'ensemble des modalités actives de la régression⁷.

Le modèle est construit à partir de 90% des établissements classés : on conserve 250 établissements sur lesquels tester la validité du pouvoir prédictif du modèle. A ces 250 établissements est attribuée la classe la plus probable d'après le modèle et on confronte celle-ci à la classe obtenue par la procédure de classification. Pour les 10 tests effectués (la sélection des 250 établissements est aléatoire), le plus bas taux de concordance obtenu est 94,8% (13 établissements classés différemment selon les deux procédures). On peut cependant noter que ces 250 établissements ne sont pas totalement neutres dans cette procédure de test. S'ils n'interviennent pas dans l'étalonnage du modèle, ils participent à la constitution des classes par analyse de données.

On a donc procédé à une classification sur les établissements à l'exception de 250 d'entre eux tirés au sort. La classification hiérarchique sur 10 axes conduit à retenir 4,7 ou 10 classes⁸. La plupart présentent des caractéristiques similaires aux classes retenues. Cependant les deux classes intermédiaires présentent des caractéristiques différentes de celles obtenues initialement. En jouant sur 6 classes, on obtient un taux de concordance de 77,20% entre la prévision sur 90% de l'effectif et l'ADD initiale.

⁷Il s'agit d'un modèle statistique.

⁸Ce résultat conforte le choix de retenir 7 classes.

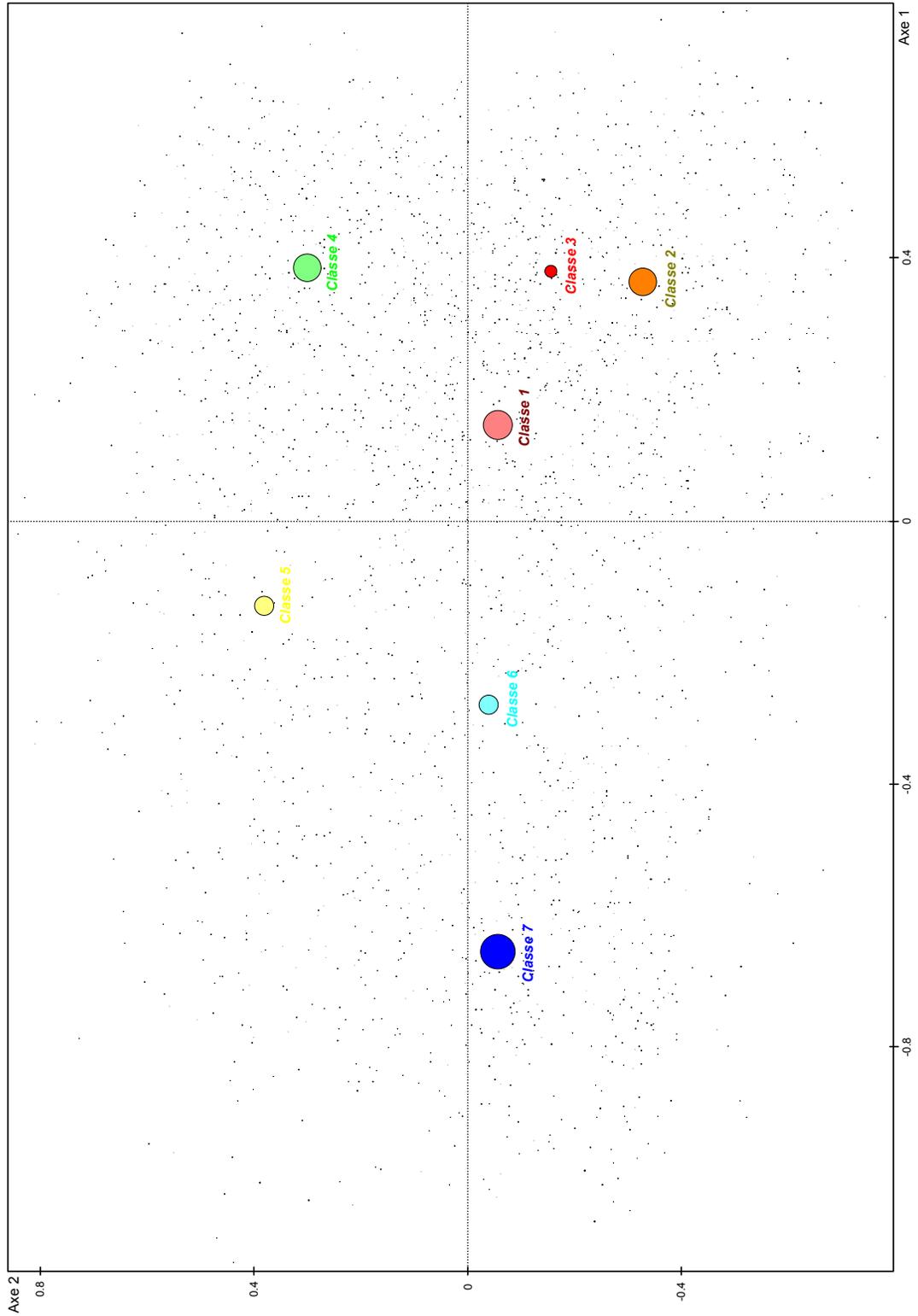


FIG. 6.3 – Barycentres des 7 classes dans le plan 1-2

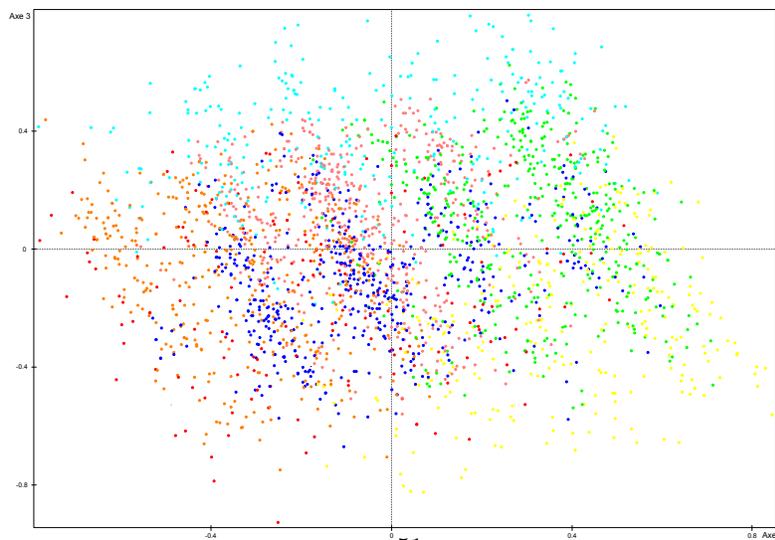
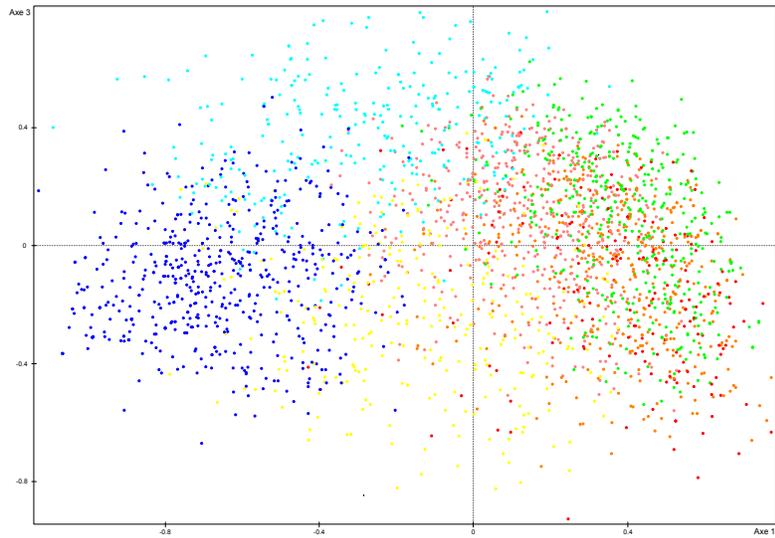
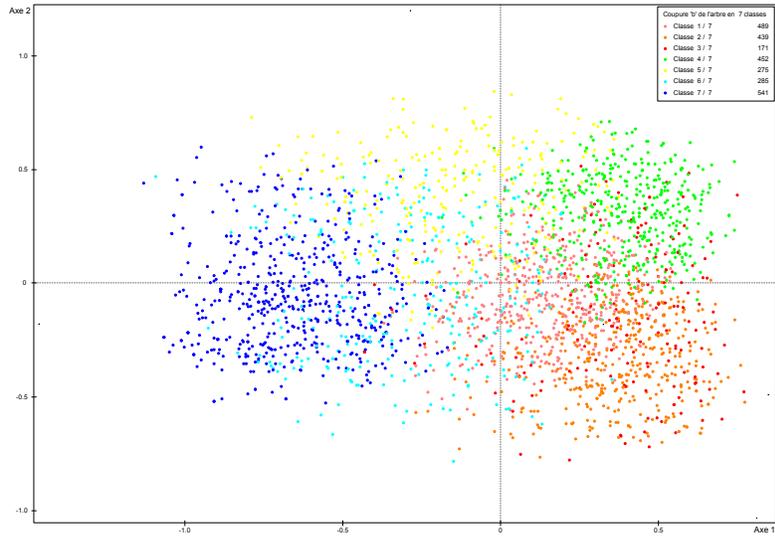


FIG. 6.4 – Projection des établissements des différentes classes

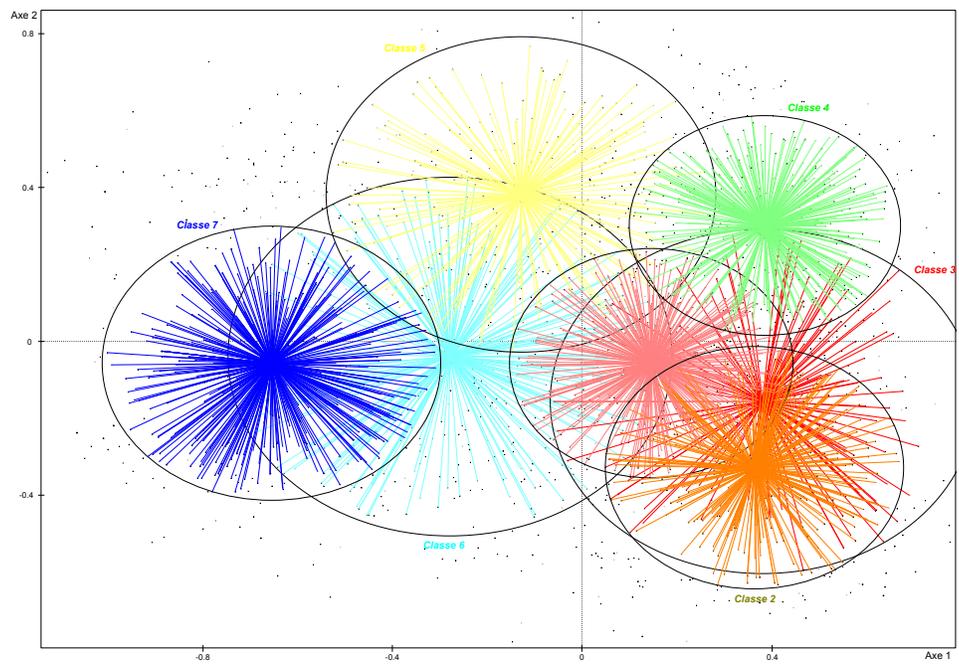


FIG. 6.5 – Dispersion des effectifs (75% des effectifs) autour des barycentres de chacune des 7 classes dans le plan 1-2

De la classe : Coupeure 'b' de l'arbre en 7 classes - Classe 1 / 7 Effectif = 489 Pourcentage = 18,44

Histogramme des : % de la classe dans la modalité

Libellés des variables	Modalités caractéristiques	% de la classe dans la modalité	% de la modalité dans la classe	% de la modalité dans l'échantillon	Valeurs-test	Probabilités	Poids
ENFALdd	1_3_enfmal	41,52	30,06	13,34	10,95	0,0000	354,00
ENFALdd	4_sj_enfmal	39,06	47,85	22,58	13,84	0,0000	599,00
fnutud	atousFinMutu	25,08	74,84	55,01	9,93	0,0000	1459,00
prestadd	pas_aide_fraisgarde	24,23	95,09	72,36	14,00	0,0000	1919,00
rtidd	2RTT_accord	24,14	70,34	53,73	8,22	0,0000	1425,00
amenibd	pas_h_garde	23,76	76,52	60,93	9,06	0,0000	1616,00
prestagd	pas_presthandic	23,60	94,06	73,49	12,70	0,0000	1949,00
prestaid	pas_colo	22,74	73,82	59,84	7,08	0,0000	1587,00
amenidd	pas_h_trajet	22,66	93,86	76,35	11,16	0,0000	2025,00
rnutud	bpasServMutu	21,89	81,18	68,36	6,94	0,0000	1813,00
partaccd	2lrs_tpspartiel	21,84	58,69	49,54	4,43	0,0000	1314,00
partembd	apostespartiel	21,66	57,87	49,24	4,18	0,0000	1306,00
prestaad	aprime_naiss	21,07	75,25	65,83	4,90	0,0000	1746,00
congdd	pas_comp1_congepatma	21,02	33,53	29,41	2,14	0,0159	780,00
amenccd	h_enfmal	20,93	80,98	71,68	5,15	0,0000	1901,00
prestaed	pas_aideetude	20,79	76,68	67,98	4,60	0,0000	1803,00
amenad	h_entrée	19,88	92,63	85,89	4,97	0,0000	2278,00
congdudd	pas_maternsup	19,87	98,97	91,81	7,50	0,0000	2435,00
gardecdd	pas_enf_entr	19,80	98,97	92,15	7,27	0,0000	2444,00
dombcd	pas_l_domicil	19,73	94,27	88,08	4,94	0,0000	2336,00
congdd	comp1_congepatmit	17,36	66,46	70,58	-2,14	0,0159	1872,00
promocdd	jamais_avancem	15,71	21,47	25,18	-2,05	0,0197	868,00
partembd	pas_postespartiel	15,30	42,12	50,75	-4,18	0,0000	1348,00
prestaed	aideetude	13,42	23,31	32,01	-4,80	0,0000	849,00
prestaad	pas_prime_naiss	13,35	24,74	34,16	-4,90	0,0000	906,00
amenccd	pas_h_enfmal	12,38	19,01	28,31	-5,15	0,0000	751,00
prestaid	colo	12,01	26,17	40,15	-7,08	0,0000	1065,00
rtidd	0pas_RTT	11,43	12,06	19,45	-4,70	0,0000	516,00
rnutud	aServMutu	10,96	18,81	31,63	-6,94	0,0000	839,00
amenabd	h_garde	10,13	21,47	39,06	-9,06	0,0000	1036,00
amenad	pas_h_entrée	9,62	7,36	14,10	-4,97	0,0000	374,00
fnutud	cpasFinMutu	9,26	18,60	37,02	-9,67	0,0000	982,00
dombcd	L_domicil_poss	8,86	5,72	11,91	-4,94	0,0000	316,00
ENFALdd	>sj_enfmal	7,52	13,08	32,05	-10,53	0,0000	850,00
rtidd	3RTT_libre	6,54	4,29	12,10	-6,38	0,0000	321,00
ENFALdd	0_enfmal	5,18	8,99	32,01	-13,11	0,0000	849,00
amenidd	h_trajet	4,78	6,13	23,64	-11,16	0,0000	627,00
prestagd	apresthandi	4,12	5,93	26,50	-12,70	0,0000	703,00
prestadd	aide_fraisgarde	3,27	4,90	27,63	-14,00	0,0000	733,00
gardecdd	enf_entr	2,40	1,02	7,84	-7,27	0,0000	208,00
congdudd	maternsup	2,30	1,02	8,18	-7,50	0,0000	217,00
partaccd	0pas_tpspartiel	0,90	0,40	8,33	-8,58	0,0000	221,00

De la classe : Coupeure 'b' de l'arbre en 7 classes - Classe 2 / 7 Effectif = 439 Pourcentage = 16,55

Histogramme des : % de la classe dans la modalité

Libellés des variables	Modalités caractéristiques	% de la classe dans la modalité	% de la modalité dans la classe	% de la modalité dans l'échantillon	Valeurs-test	Probabilités	Poids
amenad	pas_h_entrée	58,02	49,43	14,10	20,33	0,0000	374,00
ENFALdd	0_enfmal	41,22	79,72	32,01	22,75	0,0000	849,00
amenccd	pas_h_enfmal	38,61	66,05	28,31	19,20	0,0000	751,00
fnutud	atousFinMutu	30,33	14,57	7,95	5,12	0,0000	211,00
congdd	pas_comp1_congepatma	30,12	53,53	29,41	11,61	0,0000	780,00
rtidd	1RTT_contraint	30,00	26,65	14,70	7,18	0,0000	390,00
partaccd	1parfois_tpspartiel	28,37	72,20	42,11	13,93	0,0000	1117,00
amenibd	pas_h_garde	25,86	95,21	60,93	18,00	0,0000	1616,00
prestaidd	pas_chèque_vac	23,62	71,29	49,96	9,86	0,0000	1325,00
prestaid	pas_colo	23,18	83,82	59,84	11,74	0,0000	1587,00
prestadd	pas_aide_fraisgarde	22,40	97,94	72,36	15,41	0,0000	1919,00
prestagd	pas_presthandic	21,90	97,26	73,49	14,36	0,0000	1949,00
amenidd	pas_h_trajet	21,03	97,03	76,35	12,93	0,0000	2025,00
prestaed	pas_aideetude	21,02	86,33	67,98	9,52	0,0000	1803,00
fnutud	atousFinMutu	20,28	67,42	55,01	5,73	0,0000	1459,00
rnutud	bpasServMutu	20,13	83,14	68,36	7,57	0,0000	1813,00
partembd	pas_postespartiel	18,87	57,85	50,75	3,21	0,0006	1346,00
promocdd	parfois_avancem	18,73	38,49	34,01	2,10	0,0176	902,00
gardecdd	pas_enf_entr	17,79	99,08	92,15	6,97	0,0000	2444,00
dombcd	pas_l_domicil	17,72	94,30	88,08	4,65	0,0000	2336,00
congdudd	pas_maternsup	17,57	97,49	91,81	5,22	0,0000	2435,00
rtidd	2RTT_accord	15,15	49,20	53,73	-2,02	0,0211	1425,00
partembd	apostespartiel	14,16	42,14	49,24	-3,21	0,0006	1306,00
promocdd	jamais_avancem	12,72	19,36	25,18	-3,08	0,0010	668,00
congdd	comp1_congepatmit	10,89	46,46	70,58	-11,81	0,0000	1872,00
amenabd	h_entrée	9,74	50,56	85,89	-20,33	0,0000	2278,00
prestaidd	chèque_vac	9,49	28,70	50,03	-9,86	0,0000	1327,00
partaccd	2lrs_tpspartiel	9,28	27,79	49,54	-10,07	0,0000	1314,00
rnutud	aServMutu	8,82	16,85	31,63	-7,57	0,0000	839,00
fnutud	cpasFinMutu	8,04	17,99	37,02	-9,39	0,0000	982,00
dombcd	L_domicil_poss	7,91	5,69	11,91	-4,65	0,0000	316,00
amenccd	h_enfmal	7,83	33,94	71,68	-18,20	0,0000	1901,00
rtidd	3RTT_libre	7,76	5,89	12,10	-4,77	0,0000	321,00
ENFALdd	1_3_enfmal	7,34	5,92	13,34	-5,24	0,0000	354,00
prestaed	aideetude	7,06	13,66	32,01	-9,52	0,0000	849,00
prestaid	colo	6,66	16,17	40,15	-11,74	0,0000	1065,00
ENFALdd	4_sj_enfmal	6,34	8,65	22,58	-8,23	0,0000	599,00
congdudd	maternsup	5,06	2,50	8,18	-5,22	0,0000	217,00
ENFALdd	>sj_enfmal	2,94	5,69	32,05	-14,47	0,0000	850,00
amenidd	h_trajet	2,07	2,96	23,64	-12,93	0,0000	627,00
amenabd	h_garde	2,02	4,78	39,06	-18,00	0,0000	1036,00
gardecdd	enf_entr	1,92	0,91	7,84	-6,97	0,0000	208,00
prestagd	apresthandi	1,70	2,73	26,50	-14,36	0,0000	703,00
prestadd	aide_fraisgarde	1,22	2,05	27,63	-15,41	0,0000	733,00
partaccd	0pas_tpspartiel	0,00	0,00	8,33	-8,82	0,0000	221,00

FIG. 6.6 – Caractéristiques des classes 1 et 2

De la classe : Coupeure 'b' de l'arbre en 7 classes - Classe 3 / 7 Effectif = 171 Pourcentage = 6,45

Histogramme des : % de la classe dans la modalité

Libellés des variables	Modalités caractéristiques	% de la classe dans la modalité	% de la modalité dans la classe	% de la modalité dans l'échantillon	Valeurs-test	Probabilités	Poids
partacc	0pas_tpspartiel	76,32	99,41	8,33	31,66	0,0000	221,00
fnmutud	bcadreFinMutu	13,27	16,37	7,95	3,67	0,0001	211,00
rttd	0pas_RTT	12,01	36,25	19,45	5,25	0,0000	516,00
ENFMALdd	0j_enfmal	12,01	59,64	32,01	7,61	0,0000	849,00
amenad	pas_h_entrée	11,22	24,56	14,10	3,68	0,0001	374,00
partembd	pas_postespartiel	10,69	84,21	50,75	9,37	0,0000	1346,00
cong	pas_compl_congepatma	10,64	48,53	29,41	5,37	0,0000	780,00
prestahd	pas_chèque_vac	10,18	78,94	49,96	7,98	0,0000	1325,00
presta	pas_colo	8,56	79,53	59,84	5,56	0,0000	1587,00
prestagd	pas_presthandic	8,56	97,68	73,49	8,71	0,0000	1949,00
presta	pas_aide_fraisgarde	8,44	96,73	72,36	7,62	0,0000	1919,00
amenoc	pas_h_enfmal	8,36	39,84	26,31	2,42	0,0077	751,00
promoc	svt_avancem	8,29	40,35	31,37	2,48	0,0063	832,00
presta	pas_prime_naiss	8,16	43,27	34,16	2,48	0,0065	906,00
fnmutud	bpasServMutu	8,16	86,54	68,36	5,56	0,0000	1813,00
amenbd	pas_h_garde	7,85	74,26	60,93	3,70	0,0001	1616,00
presta	pas_aideetude	7,76	81,87	67,98	4,11	0,0000	1803,00
amenoc	pas_h_trajet	7,55	89,47	76,35	4,39	0,0000	2025,00
gardectt	pas_ent_entr	6,99	100,00	92,15	4,89	0,0000	2444,00
congurd	pas_maternsup	6,85	97,68	91,81	3,09	0,0010	2435,00
dombd	pas_L_domicil	6,80	92,98	88,08	2,01	0,0220	2336,00
amenoc	h_enfmal	5,68	63,15	71,68	-2,42	0,0077	1901,00
amenad	h_entrée	5,66	75,43	85,89	-3,68	0,0001	2278,00
presta	apime_naiss	5,55	56,72	65,83	-2,48	0,0065	1746,00
rttd	2RTT_accord	5,19	43,27	53,73	-2,75	0,0029	1425,00
promoc	parfois_avancem	4,76	25,14	34,01	-2,49	0,0063	902,00
cong	pas_compl_congepatmat	4,70	51,46	70,58	-5,37	0,0000	1872,00
fnmutud	cpasFinMutu	4,48	25,73	37,02	-3,15	0,0008	962,00
ENFMALdd	4_sj_enfmal	4,34	15,20	22,58	-2,36	0,0088	599,00
amenbd	h_garde	4,24	25,73	39,06	-3,70	0,0001	1036,00
dombd	L_domicil_poss	3,79	7,01	11,91	-2,01	0,0220	316,00
presta	aideetude	3,65	18,12	32,01	-4,11	0,0000	849,00
presta	colo	3,28	20,46	40,15	-5,56	0,0000	1065,00
amenad	h_trajet	2,87	10,52	23,64	-4,39	0,0000	627,00
mutud	aServMutu	2,74	13,45	31,63	-5,56	0,0000	839,00
prestahd	chèque_vac	2,71	21,05	50,03	-7,98	0,0000	1327,00
rttd	3RTT_libre	2,49	4,67	12,10	-3,25	0,0005	321,00
partembd	apostespartiel	2,06	15,78	49,24	-9,37	0,0000	1306,00
congurd	maternsup	1,84	2,33	8,18	-3,09	0,0010	217,00
ENFMALdd	>6j_enfmal	1,76	8,77	32,05	-7,35	0,0000	850,00
presta	aide_fraisgarde	1,22	5,26	27,63	-7,62	0,0000	733,00
prestagd	apresthandic	0,56	2,33	26,50	-8,71	0,0000	703,00
partacc	2trs_tpspartiel	0,07	0,58	49,54	-15,09	0,0000	1314,00
partacc	1parfois_tpspartiel	0,00	0,00	42,11	-13,73	0,0000	1117,00
gardectt	ent_entr	0,00	0,00	7,84	-4,89	0,0000	208,00

De la classe : Coupeure 'b' de l'arbre en 7 classes - Classe 4 / 7 Effectif = 452 Pourcentage = 17,04

Histogramme des : % de la classe dans la modalité

Libellés des variables	Modalités caractéristiques	% de la classe dans la modalité	% de la modalité dans la classe	% de la modalité dans l'échantillon	Valeurs-test	Probabilités	Poids
amenoc	h_trajet	44,49	61,72	23,64	19,39	0,0000	627,00
amenbd	h_garde	39,67	90,92	39,06	25,38	0,0000	1036,00
ENFMALdd	0j_enfmal	35,21	66,15	32,01	16,45	0,0000	849,00
promoc	trs_avancem	34,00	18,80	9,42	6,82	0,0000	250,00
cong	pas_compl_congepatma	33,84	58,40	29,41	14,18	0,0000	780,00
fnmutud	bcadreFinMutu	33,64	15,70	7,95	6,07	0,0000	211,00
rttd	0pas_RTT	28,68	32,74	19,45	7,38	0,0000	516,00
prestahd	pas_chèque_vac	26,79	78,53	49,96	13,62	0,0000	1325,00
presta	pas_colo	24,63	86,50	59,84	13,43	0,0000	1587,00
partacc	1parfois_tpspartiel	23,54	58,18	42,11	7,49	0,0000	1117,00
presta	pas_aide_fraisgarde	23,24	96,67	72,36	16,35	0,0000	1919,00
amenoc	h_enfmal	23,09	97,12	71,68	15,25	0,0000	1901,00
prestagd	pas_presthandic	22,72	98,00	73,49	15,25	0,0000	1949,00
promoc	svt_avancem	22,47	41,37	31,37	4,89	0,0000	832,00
presta	pas_aideetude	22,01	97,68	67,98	10,57	0,0000	1803,00
mutud	bpasServMutu	20,95	84,07	68,36	8,22	0,0000	1813,00
fnmutud	atousFinMutu	20,42	65,92	55,01	5,11	0,0000	1459,00
presta	pas_prime_naiss	20,30	40,70	34,16	3,13	0,0008	906,00
amenad	h_entrée	19,71	99,33	85,89	11,02	0,0000	2278,00
congurd	pas_maternsup	18,27	98,45	91,81	6,46	0,0000	2435,00
gardectt	pas_ent_entr	18,24	98,67	92,15	6,52	0,0000	2444,00
dombd	pas_L_domicil	18,02	93,14	86,08	3,76	0,0000	2336,00
presta	apime_naiss	15,34	59,29	65,83	-3,13	0,0008	1746,00
partacc	2trs_tpspartiel	14,00	40,70	49,54	-4,08	0,0000	1314,00
rttd	2RTT_accord	13,61	42,92	53,73	-5,00	0,0000	1425,00
promoc	parfois_avancem	13,41	26,76	34,01	-3,56	0,0001	902,00
cong	compl_congepatmat	10,04	41,59	70,58	-14,18	0,0000	1872,00
ENFMALdd	4_sj_enfmal	10,01	13,27	22,58	-5,38	0,0000	599,00
dombd	L_domicil_poss	9,81	6,95	11,91	-3,76	0,0000	316,00
rttd	3RTT_libre	9,34	6,63	12,10	-4,06	0,0000	321,00
promoc	jamaiss_avancem	8,83	13,05	25,18	-6,83	0,0000	668,00
mutud	aServMutu	8,58	15,92	31,63	-8,22	0,0000	839,00
amenoc	pas_h_trajet	8,54	38,27	76,35	-19,39	0,0000	2025,00
fnmutud	cpasFinMutu	8,45	18,36	37,02	-9,36	0,0000	962,00
prestahd	chèque_vac	7,30	21,46	50,03	-13,62	0,0000	1327,00
presta	aideetude	6,47	12,16	32,01	-10,57	0,0000	849,00
presta	colo	5,72	13,49	40,15	-13,43	0,0000	1065,00
congurd	maternsup	3,22	1,54	8,18	-6,46	0,0000	217,00
ENFMALdd	>6j_enfmal	3,17	5,97	32,05	-14,53	0,0000	850,00
gardectt	ent_entr	2,88	1,32	7,84	-6,52	0,0000	208,00
amenbd	pas_h_garde	2,53	9,07	60,93	-25,38	0,0000	1616,00
partacc	0pas_tpspartiel	2,26	1,10	8,33	-7,15	0,0000	221,00
amenoc	pas_h_enfmal	1,73	2,87	28,31	-15,25	0,0000	751,00
prestagd	apresthandic	1,28	1,99	26,50	-15,25	0,0000	703,00
presta	aide_fraisgarde	0,81	1,32	27,63	-16,35	0,0000	733,00
amenad	pas_h_entrée	0,80	0,66	14,10	-11,02	0,0000	374,00

FIG. 6.7 – Caractéristiques des classes 3 et 4

De la classe : Coupure 'b' de l'arbre en 7 classes - Classe 5 / 7 Effectif = 275 Pourcentage = 10,37
 Histogramme des : % de la classe dans la modalité

Libellés des variables	Modalités caractéristiques	% de la classe dans la modalité	% de la modalité dans la classe	% de la modalité dans l'échantillon	Valeurs-test	Probabilités	Poids
dombd	L_domicil_poss	34,17	39,27	11,91	12,51	0,0000	316,00
fnutud	cpasFinMutu	23,62	84,36	37,02	17,06	0,0000	982,00
prestaad	pas_prime_naiss	23,50	77,45	34,16	15,52	0,0000	906,00
amenad	h_trajet	22,32	50,90	23,64	10,38	0,0000	627,00
rttd	0pas_RTT	21,70	40,72	19,45	8,60	0,0000	516,00
ENFMALdd	>61_enfmal	21,64	66,90	32,05	12,50	0,0000	850,00
promocd	jamais_avancem	21,10	51,27	25,18	9,80	0,0000	668,00
amenbd	h_garde	19,11	72,00	39,06	11,64	0,0000	1036,00
fnutud	aServMutu	15,85	48,36	31,63	6,05	0,0000	839,00
cong	compl_congepatmat	14,26	97,09	70,58	11,86	0,0000	1872,00
partacc	2lrs_tpspartiel	14,00	66,90	49,54	6,06	0,0000	1314,00
partembd	apostespartiel	13,78	65,45	49,24	5,65	0,0000	1306,00
amenad	h_enfmal	13,46	93,09	71,68	9,23	0,0000	1901,00
prestaed	pas_aideetude	13,42	88,00	67,98	8,02	0,0000	1803,00
amenad	h_rentrée	11,10	92,00	85,89	3,15	0,0008	2278,00
cong	pas_maternsup	10,96	97,09	91,81	3,61	0,0001	2435,00
promocd	parfois_avancem	8,53	28,00	34,01	-2,17	0,0146	902,00
rttd	RTT_accord	8,14	42,18	53,73	-3,98	0,0000	1425,00
fnutud	0pas_ServMutu	7,93	51,63	66,36	-6,05	0,0000	1813,00
partacc	1parfois_tpspartiel	7,16	29,09	42,11	-4,63	0,0000	1117,00
dombd	pas_L_domicil	7,14	60,72	88,08	-12,51	0,0000	2336,00
partembd	pas_postespartiel	7,05	34,54	50,75	-5,65	0,0000	1346,00
amenad	pas_h_trajet	6,66	49,09	76,35	-10,38	0,0000	2025,00
amenad	pas_h_rentrée	5,88	8,00	14,10	-3,15	0,0008	374,00
promocd	tirs_avancem	5,60	5,09	9,42	-2,84	0,0040	250,00
promocd	svt_avancem	5,16	15,63	31,37	-6,20	0,0000	832,00
partacc	0pas_tpspartiel	4,97	4,00	8,33	-2,83	0,0023	221,00
ENFMALdd	1_31_enfmal	4,80	6,18	13,34	-3,89	0,0000	354,00
amenbd	pas_h_garde	4,76	28,00	60,93	-11,64	0,0000	1616,00
prestaed	aideetude	3,88	12,00	32,01	-8,02	0,0000	849,00
cong	maternsup	3,88	2,90	8,18	-3,61	0,0001	217,00
rttd	1RTT_contraint	3,58	5,09	14,70	-5,18	0,0000	390,00
prestaad	aprime_naiss	3,55	22,54	85,63	-15,52	0,0000	1746,00
fnutud	atousFinMutu	2,87	14,18	55,01	-14,78	0,0000	1459,00
amenad	pas_h_enfmal	2,52	6,90	28,31	-9,23	0,0000	751,00
ENFMALdd	01_enfmal	2,00	6,18	32,01	-10,64	0,0000	849,00
fnutud	bcadreFinMutu	1,89	1,45	7,95	-4,80	0,0000	211,00
cong	pas_compl_congepatma	1,02	2,90	29,41	-11,86	0,0000	780,00

De la classe : Coupure 'b' de l'arbre en 7 classes - Classe 6 / 7 Effectif = 285 Pourcentage = 10,75
 Histogramme des : % de la classe dans la modalité

Libellés des variables	Modalités caractéristiques	% de la classe dans la modalité	% de la modalité dans la classe	% de la modalité dans l'échantillon	Valeurs-test	Probabilités	Poids
cong	maternsup	74,85	56,84	8,18	24,20	0,0000	217,00
rttd	3RTT_libre	48,90	55,08	12,10	19,24	0,0000	321,00
prestaad	colo	19,34	72,28	40,15	11,57	0,0000	1065,00
prestaad	aide_fraisgarde	18,82	48,42	27,63	7,86	0,0000	733,00
prestaad	aideetude	18,72	55,78	32,01	8,72	0,0000	849,00
ENFMALdd	1_31_enfmal	17,51	21,75	13,34	4,08	0,0000	354,00
prestaad	chèque_vac	16,42	76,49	50,03	9,80	0,0000	1327,00
ENFMALdd	4_51_enfmal	16,36	34,38	22,58	4,76	0,0000	599,00
fnutud	atousFinMutu	16,17	82,80	55,01	10,38	0,0000	1459,00
dombd	L_domicil_poss	15,82	17,54	11,91	2,88	0,0019	316,00
partembd	pas_postespartiel	15,23	71,92	50,75	7,62	0,0000	1346,00
partacc	2lrs_tpspartiel	14,99	69,12	49,54	7,00	0,0000	1314,00
promocd	parfois_avancem	14,41	45,61	34,01	4,23	0,0000	902,00
cong	compl_congepatmat	14,15	92,98	70,58	9,73	0,0000	1872,00
prestaad	aprime_naiss	13,97	85,61	65,63	7,86	0,0000	1746,00
prestaad	apresthnaiss	13,79	34,03	26,50	2,81	0,0017	703,00
amenad	h_trajet	13,23	29,12	23,64	2,19	0,0140	627,00
promocd	svt_avancem	12,62	36,84	31,37	2,02	0,0216	832,00
amenbd	h_garde	12,35	44,91	39,06	2,06	0,0193	1036,00
dombd	pas_L_domicil	10,05	82,45	88,08	-2,88	0,0019	2336,00
amenbd	pas_h_trajet	9,97	70,87	76,35	-2,19	0,0140	2025,00
amenbd	pas_h_garde	9,71	55,08	60,93	-2,06	0,0193	1616,00
prestaad	pas_presthnaiss	9,64	65,96	73,49	-2,91	0,0017	1949,00
prestaad	pas_aide_fraisgarde	7,66	51,57	72,36	-7,86	0,0000	1919,00
rttd	2RTT_accord	7,57	37,89	53,73	-5,61	0,0000	1425,00
partacc	1parfois_tpspartiel	7,34	26,77	42,11	-4,85	0,0000	1117,00
prestaed	pas_aideetude	6,98	44,21	67,98	-8,72	0,0000	1803,00
promocd	jamais_avancem	6,13	14,38	25,18	-4,59	0,0000	668,00
partembd	apostespartiel	6,12	26,07	49,24	-7,62	0,0000	1306,00
prestaad	pas_chèque_vac	5,05	23,50	49,96	-9,60	0,0000	1325,00
cong	pas_maternsup	5,05	43,15	91,81	-24,20	0,0000	2435,00
prestaad	pas_colo	4,97	27,71	59,84	-11,57	0,0000	1587,00
prestaad	pas_prime_naiss	4,52	14,38	34,16	-7,86	0,0000	906,00
fnutud	cpasFinMutu	4,27	14,73	37,02	-8,70	0,0000	982,00
promocd	tirs_avancem	3,60	3,15	9,42	-4,16	0,0000	250,00
fnutud	bcadreFinMutu	3,31	2,45	7,95	-3,95	0,0000	211,00
ENFMALdd	01_enfmal	3,29	9,62	32,01	-9,21	0,0000	849,00
rttd	1RTT_contraint	2,82	3,85	14,70	-6,12	0,0000	390,00
partacc	0pas_tpspartiel	2,71	2,10	8,33	-4,48	0,0000	221,00
cong	pas_compl_congepatma	2,56	7,01	29,41	-9,73	0,0000	780,00
rttd	0pas_RTT	1,74	3,15	19,45	-8,50	0,0000	516,00

FIG. 6.8 – Caractéristiques des classes 5 et 6

De la classe : Coupure 'b' de l'arbre en 7 classes - Classe 7 / 7 Effectif = 541 Pourcentage = 20,40
 Histogramme des : % de la classe dans la modalité

Libellés des variables	Modalités caractéristiques	% de la classe dans la modalité	% de la modalité dans la classe	% de la modalité dans l'échantillon	Valeurs-test	Probabilités	Poids
gardectt	ent_entr	71,15	27,35	7,84	16,59	0,0000	208,00
prestagd	apresthandi	68,56	89,09	26,50	35,70	0,0000	703,00
prestadd	aide_fraisgarde	65,07	86,17	27,63	34,19	0,0000	733,00
ENFMALdd	>6j_entfmal	51,52	80,96	32,05	26,70	0,0000	850,00
prestaed	aideetude	46,76	73,38	32,01	22,43	0,0000	849,00
rmutud	aServMutu	42,31	65,81	31,63	18,41	0,0000	839,00
fmfutud	cpasFmMutu	41,85	75,97	37,02	20,79	0,0000	982,00
prestaidd	colo	41,78	82,25	40,15	22,55	0,0000	1065,00
prestaehd	chèque_vac	35,11	86,13	50,03	19,60	0,0000	1327,00
promocdd	jamais_avancem	29,94	36,96	25,18	6,81	0,0000	668,00
congdd	compl_congepatmat	28,57	98,89	70,58	19,19	0,0000	1872,00
rtidd	2RTT_accordi	26,17	68,94	53,73	8,00	0,0000	1425,00
partaccdd	2firs_tpspartiel	25,79	62,06	49,54	6,81	0,0000	1314,00
amenadd	pas_h_garde	25,49	76,15	60,93	8,29	0,0000	1616,00
amenocdd	pas_h_entfmal	25,29	35,12	28,31	3,82	0,0000	751,00
prestaad	aprime_naiss	24,74	79,85	65,83	7,91	0,0000	1746,00
partembdd	apostespartiel	24,11	58,22	49,24	4,64	0,0000	1306,00
amenocdd	pas_h_trajet	23,55	88,17	76,35	7,61	0,0000	2025,00
amenadd	h_rentrée	22,73	95,74	85,89	8,13	0,0000	2278,00
congduard	pas_maternsup	21,39	96,30	91,81	4,51	0,0000	2435,00
amenocdd	h_entfmal	18,46	64,87	71,68	-3,82	0,0000	1901,00
partembdd	pas_postespartiel	16,79	41,77	50,75	-4,64	0,0000	1346,00
gardectt	pas_ent_entr	16,08	72,64	92,15	-16,59	0,0000	2444,00
partaccdd	1parfois_tpspartiel	15,66	32,34	42,11	-5,16	0,0000	1117,00
promocdd	firs_avancem	15,20	7,02	9,42	-2,11	0,0173	250,00
rtidd	3RTT_libre	14,64	8,68	12,10	-2,73	0,0031	321,00
promocdd	svt_avancem	14,54	22,36	31,37	-5,12	0,0000	832,00
ENFMALdd	4_6j_entfmal	14,35	15,89	22,58	-4,23	0,0000	589,00
amenadd	h_garde	12,45	23,94	39,06	-8,29	0,0000	1036,00
partaccdd	0pas_tpspartiel	12,21	4,99	8,33	-3,21	0,0006	221,00
prestaad	pas_prime_naiss	12,03	20,14	34,16	-7,91	0,0000	906,00
rmutud	bpasServMutu	10,25	34,38	68,36	-18,41	0,0000	1813,00
amenocdd	h_trajet	10,20	11,82	23,64	-7,61	0,0000	627,00
congduard	maternsup	9,21	3,69	8,18	-4,51	0,0000	217,00
rtidd	0pas_RTT	8,72	8,31	19,45	-7,83	0,0000	516,00
fmfutud	elouFmMutu	8,56	23,10	55,01	-16,89	0,0000	1459,00
prestaed	pas_aideetude	7,98	26,61	67,98	-22,43	0,0000	1803,00
amenadd	pas_h_rentrée	6,14	4,25	14,10	-8,13	0,0000	374,00
prestaidd	pas_colo	6,04	17,74	59,84	-22,55	0,0000	1587,00
prestaehd	pas_chèque_vac	5,66	13,86	49,96	-19,60	0,0000	1325,00
prestadd	pas_aide_fraisgarde	3,33	11,82	72,36	-34,19	0,0000	1919,00
prestagd	pas_presthandic	3,02	10,90	73,49	-35,70	0,0000	1949,00
fmfutud	bcaDefmMutu	2,36	0,92	7,95	-7,96	0,0000	211,00
ENFMALdd	1_3_6j_entfmal	2,25	1,47	13,34	-10,71	0,0000	354,00
ENFMALdd	0j_entfmal	1,06	1,66	32,01	-19,80	0,0000	849,00
congdd	pas_compl_congepatma	0,76	1,10	29,41	-19,19	0,0000	780,00

FIG. 6.9 – Caractéristiques de la classe 7

Éléments de conclusion

Bilan

L'objectif de ce travail était de forger un instrument permettant de résumer les différentes mesures des établissements employeurs se rapportant à la conciliation de la vie professionnelle et de la vie familiale des salariés. Le recours aux méthodes de classification automatique a permis d'obtenir une typologie de 7 classes d'établissements caractérisées par des "politiques familiales" différentes. Cette classification a vocation à faciliter l'utilisation du caractère couplé de l'Enquête Familles-Employeurs. La volonté de forger une classification de référence a conduit à rechercher un mode de traitement de la non-réponse partielle adapté. La présentation en détail de la construction de la typologie dans ce document de travail, tant sur le plan théorique que pratique, laisse la possibilité de reprendre la méthode utilisée pour forger d'autres typologies. A l'issue de ce travail et de la mise à disposition d'une typologie des établissements, nous espérons que celle-ci constituera un outil d'analyse intéressant et que ce document sera un guide utile non seulement pour son utilisation mais également pour la constitution d'autres classifications.

Cinq questions sur la classification

"Combien de classes doit-on retenir ?"

La classification en 7 classes a retenu notre attention dans la présentation. Il paraît préférable d'utiliser ce niveau d'agrégation qui permet de prendre en compte la diversité des situations. Les 7 classes ne constituent pas les degrés ordonnés d'une échelle de la conciliation vie familiale/vie professionnelle allant de l'absence de conciliation à une situation de conciliation parfaite. Cette conciliation peut prendre de multiples formes, difficilement traduisible de façon unidimensionnelle.

Cependant les contraintes d'effectifs peuvent rendre difficile la manipulation d'une variable dotée d'un nombre aussi élevé de modalités. Dans ce cas, il est tout à fait possible d'utiliser cette partition au niveau plus agrégé, celui en 3 classes. A la différence des 7 classes, ce niveau de concentration caractérise trois degrés de conciliation : son absence, sa présence et une classe intermédiaire.

"Pourquoi n'a t-on pas tenu compte des pondérations pour réaliser la classification ?"

L'opportunité de prendre en compte la pondération dans l'exploitation d'une enquête est une question qui fait souvent débat. Les avis sont souvent partagés⁹ : faut-il privilégier la "véritable" observation et occulter les effets procédure de sondage ? Ou faut-il corriger la sur-représentation des établissements à fort effectifs au risque de "créer" des tendances de comportements à partir de quelques observations

⁹On peut consulter à ce titre dans [12] le paragraphe au titre significatif "Pondérer ou ne pas pondérer : that's the question!".

fortement pondérées ?

Dans le cas qui nous occupe, les poids sont extrêmement dispersés (voir tableau 2.4 p.23) : les rapports de poids sont de 1 à 433 entre les observations aux pondérations extrêmes (et encore ont-elles été préalablement tronquées d'ordre 2%). En utilisant les pondérations, on court le risque de ne faire porter l'analyse que sur les établissements dotés d'un poids élevé. Ainsi seuls les établissements à faibles effectifs et ayant des caractéristiques corrélées avec un taux de non-réponse totale élevé joueront un rôle dans la construction. De plus l'analyse menée n'est pas un travail d'estimation d'un indicateur mais une mise en évidence des corrélations entre variables qualitatives. Enfin, rappelons que l'objectif premier de cette typologie n'est pas de servir à l'analyse de l'ensemble des établissements mais de caractériser chacun de ceux qui ont été interrogés afin de mener des analyses au niveau de leurs salariés.

Par conséquent, nous avons choisi de ne pas faire intervenir les poids dans les analyses de données (ACM, typologie). Cela ne signifie pas pour autant que le processus de construction de l'échantillon est oublié. Le poids est en effet utilisé dans la phase de correction de la non-réponse partielle. Par conséquent, c'est moins l'absence d'utilisation des poids que l'attention portée aux observations individuelles qui caractérise la méthode retenue.

A titre de comparaison, une annexe ¹⁰ présente les résultats de l'ACM prenant en compte les pondérations.

"Pourquoi cette typologie ne s'applique-t-elle pas à tous les établissements (notamment ceux de moins de 20 salariés) ?"

Le travail de construction de la typologie a été réalisée à partir du volet Employeurs de l'enquête. Celui-ci rassemble les réponses des 2673 établissements d'au moins vingt salariés, situés en France métropolitaine, employant au moins un salarié né entre 1955 et 1984 (âgé de 20 à 49 ans en 2004). Dans cette enquête couplée, ce sont en effet les salariés qui sont d'abord interrogés. Leurs établissements ne le sont qu'ensuite, s'ils répondent aux critères précédemment énoncés.

Pour les établissements de moins de 20 salariés, on dispose d'informations données par le ou les salariés. Cependant celles-ci ne sont pas aussi précises que le questionnaire envoyé aux établissements de plus de 20 salariés. On pourrait envisager de créer une typologie à partir des questions identiques pour les deux types d'établissements. Cependant outre le caractère plus réduit de la caractérisation de la conciliation, cette idée se heurte à la difficulté d'assimiler les comportements de réponse entre des salariés et des établissements. Les travaux de comparaison menés sur les établissements de plus de vingt salariés montrent des différences importantes dans leurs réponses.

"La prise en compte de l'existence de postes à temps partiels comme mesure de conciliation n'est-il pas fortement contestable ?"

Les études sur le temps partiel¹¹ montrent qu'il s'agit d'un dispositif à destination des seules femmes (on le vérifie ici avec la classe 2) et que la réflexion sur son caractère contraint ou non dépasse le simple cadre de la relation salarié-employeur et s'avère souvent peu pertinente (voir M.Marvani [14], p.90). Ainsi si la possibilité de passer à temps partiel peut à la limite être considérée, avec les réserves formulées précédemment, comme une forme de conciliation, la question de l'existence de postes de ce type ne doit-elle pas être éliminée de l'analyse ?

¹⁰(l'annexe C.3 p.104)

¹¹Voir notamment la synthèse réalisée par M. Marvani dans le chapitre 4 de *Travail et emploi des femmes*, La Découverte, 2003 [14]. Pour des données plus complètes, on pourra consulter deux publications *Premières synthèses* (DARES) de V.Ulrich et S.Zilberman datées de septembre 2007. On pourra également se reporter au paragraphe intitulé "Les dangers du temps partiel" p.11 dans *Le deuxième âge de l'émancipation* de D. Méda et H. Périer, 2007 [15].

Il convient cependant de ne pas confondre la modalité caractérisant l'existence de postes de ce type et la variable à laquelle elle se rapporte. Le fait de retenir la variable ne préjuge pas du sens que l'on donne à l'existence de postes à temps partiel. C'est justement ce qui distingue la classification automatique d'une analyse des établissements sous la forme d'un score a priori (attribution d'un certain nombre de "points" de conciliation à chaque mesure). Dans la réalisation d'un tel score, le concepteur doit caractériser le fait pour un établissement de proposer des postes à temps partiel. Il doit ainsi commencer par déterminer si cela constitue une mesure de conciliation ou au contraire une contrainte. On raisonne alors de façon unidimensionnelle et a priori. Dans le cadre de la classification automatique, le fait de retenir la variable "existence de postes à temps partiel" ne préjuge pas du caractère de la mesure. Ses deux modalités apparaîtront en fonction des corrélations qu'elles entretiennent avec les autres mesures. Ainsi l'utilisateur n'a pas à juger le dispositif et la méthode utilisée laisse la possibilité à ce dernier d'apparaître de façon plus nuancée que dans une opposition unidimensionnelle conciliation - absence de conciliation.

"Je souhaiterais enlever (rajouter) une ou plusieurs variables dans la détermination de la typologie."

La typologie présentée dans ce document résulte de la prise en compte de précédentes observations réalisées l'enquête et de différents essais. Elle ne constitue cependant qu'une proposition parmi d'autres typologie potentiellement possibles. Son statut est avant tout celle d'une typologie par défaut. Un des objectifs de ce document est d'expliquer aux utilisateurs de cette typologie le processus qui a permis d'aboutir à celle-ci. Mais ce document a également pour mission de fournir la structure nécessaire pour élaborer facilement une typologie alternative. Il suffit de modifier à la marge le code SAS pour procéder à la correction de la non-réponse partielle puis de passer à la phase de classification sous SPAD. Cette dernière étape est cependant moins automatique que la première, notamment pour le choix du nombre d'axes retenus, de la méthode de classification retenue et du nombre de classe finalement obtenu. La construction de sa propre typologie en suivant ce document présente l'avantage de donner une certaine homogénéité dans les procédures utilisées, notamment en matière de traitement de la non-réponse. Mais l'utilisateur reste libre en dernier recours d'élaborer son propre traitement de la non-réponse.

Bibliographie

- [1] BACCINI, A., AND BESSE, P. Data mining I. Exploration statistique. *Publications du laboratoire de statistique et probabilités, Université Paul Sabatier, Toulouse III* (version septembre 2005).
- [2] BRION, P. L'impact de différentes stratégies du traitement des non-réponses dans les enquêtes auprès des entreprises. *INSEE, Actes des journées de méthodologie statistique 2005*.
- [3] BRION, P., CARON, N., AND PIETRI-BESSY, P. Redresser la non réponse totale dans les enquêtes auprès des entreprises : les pièges à éviter. Illustration avec l'enquête innovation.
- [4] CARON, N. La correction de la non-réponse par repondération et par imputation. *Document de travail de la série Méthodologie Statistique, INSEE, M0502*. Version actualisée du document 9604 disponible dans la même série intitulé "Les principales techniques de correction de la non-réponse et les modèles associés".
- [5] GODAMBE, V., AND SPROTT, D., Eds. *Foundations of statistical inference*. 1971, ch. An essay on the logical foundations of survey sampling, Part one, pp. 203–242.
- [6] HALTIWANGER, J., LANE, J., SPLETZER, J., THEEUWES, J., AND TROSKE, K., Eds. *The creation and analysis of employer-employee matched data*. North Holland, 1999.
- [7] INED INSEE. *Documentation du volet employeurs de l'enquête "Famille et employeurs"*, version 2. mars 2007 ed.
- [8] ISNARD, M., AND SAUTORY, O. Les macros SAS d'analyse de données. *Documents de travail de la direction des statistiques démographiques et sociales (INSEE), F9405* (mars 1994).
- [9] LADIRAY, D., AND QUENNEVILLE, B. Statisticiens : attention, logiciels! *Actes des journées de méthodologie statistique* (2002).
- [10] LE GUENNEC, J. La régression sur échantillon avec SAS. *Actes des journées de méthodologie statistique 2005*.
- [11] LE LAN, R. Analyse de données et classification sur données d'enquête. Choix sur les variables, le nombre de classes et le nombre d'axes. *Actes des journées de méthodologie statistique 2005*.
- [12] LEBLANC, D., LOLLIVIER, S., MERPSAT, M., AND VERGER, D. L'économétrie et l'étude des comportements. Présentation et mise en oeuvre de modèles de régression qualitatifs. Les modèles univariés à résidus logistiques ou normaux (logit, probit). *Série des documents de travail Méthodologie statistique, direction des statistiques démographiques et sociales (INSEE), 0001* (janvier 2000).
- [13] LEFEVRE, C., PAILHÉ, A., AND SOLAZ, A. Politiques de conciliation vie familiale-vie professionnelle des entreprises en France. *Conférence de l'Association internationale des démographes de langue française (AIDELF)* (Aveiro, septembre 2006).
- [14] MARUANI, M. *Travail et emploi des femmes*. La Découverte, 2003.
- [15] MÉDA, D., AND PÉRIVIER, H. *Le deuxième âge de l'émancipation*. La République des idées, Seuil, 2007.
- [16] NAKACHE, P., AND CONFAIS, J. *Méthodes de classification*. CISIA-CERESTA.

- [17] PAILHÉ, A., AND SOLAZ, A. L'enquête familles et employeurs. Protocole d'une double enquête et bilan de collecte. *Document de travail INED 143* (2007).
- [18] SAUTORY, O. *Analyse des données*. Polycopié ENSAE.
- [19] SAUTORY, O. Polycopiés du cours de sondage de 2^e année ENSAE, 2006-2007.
- [20] VANDERSCHULDEN, M. Homogamie et choix du conjoint. Traitement de la non réponse, imputation de variables qualitatives corrélées. *Document de travail de la direction des statistiques démographiques et sociales*, F0505 (mai 2005).

Troisième partie

Annexe

Annexe A

Questionnaire et libellés des modalités dans les sorties

	Variable	Modalités	Libellés
1	ENFMALDD	0 1-3 4-6 >6	<i>0j_enfmal</i> <i>1_3j_enfmal</i> <i>4_6j_enfmal</i> <i>> 6j_enfmal</i>
2	CONGDURD	0 >0	<i>0_maternsup</i> <i>maternsup</i>
3	GARDEDTT	oui non	<i>enf_entr</i> <i>pas_enf_entr</i>
5	PRESTAID	oui non	<i>colo</i> <i>pas_colo</i>
6	PRESTAAD	oui non	<i>aprime_naiss</i> <i>pas_prime_naiss</i>
7	PRESTADD	oui non	<i>aide_fraisgarde</i> <i>pas_aide_fraisgarde</i>
8	PRESTAED	oui non	<i>aideetude</i> <i>pas_aideetude</i>
9	PRESTAGD	oui non	<i>aprestandi</i> <i>pas_prestandic</i>
10	CONGD	oui non	<i>compl_congepatmat</i> <i>pas_compl_congepatma</i>
11	AMENAD	oui non	<i>h_rentree</i> <i>pas_h_rentree</i>
12	AMENBD	oui non	<i>h_garde</i> <i>pas_h_garde</i>
13	AMENCD	oui non	<i>h_enfmal</i> <i>pas_h_enfmal</i>
14	AMENDD	oui non	<i>h_trajet</i> <i>pas_h_trajet</i>
15	DOMBD	oui non	<i>L_domicil_poss</i> <i>pas_L_domicil</i>
16	PARTEMBD	oui non	<i>apostespartiel</i> <i>pas_postespartiel</i>
17	PARTACCD	toujours parfois non	<i>0pas_tpspartiel</i> <i>1parfois_tpspartiel</i> <i>2tjrs_tpspartiel</i>
18	RTTDD	libre accord contraint abs de RTT	<i>3RTT_libre</i> <i>2RTT_accord</i> <i>1RTT_contraint</i> <i>0pas_RTT</i>
19	PROMOCD	toujours souvent parfois jamais	<i>tjrs_avancem</i> <i>svt_avancem</i> <i>parfois_avancem</i> <i>jamais_avancem</i>
20	PRESTAHD	oui non	<i>chque_vac</i> <i>pas_chque_vac</i>
21	FMUTUD	tous cadre non	<i>atousFinMutu</i> <i>bcadreFinMutu</i> <i>cpasFinMutu</i>
22	MUTUD	oui non	<i>aServMutu</i> <i>bpasServMutu</i>

A5 Par quel moyen principal sont contrôlés les horaires des... (1 seule réponse par ligne)

	Pointeuse, badge	Signature	Contrôleur, responsable hiérarchique	Autre	Pas de contrôle
- cadres, catégories A ?	<input type="checkbox"/>	<input type="checkbox"/>	CONTR0A-A	<input type="checkbox"/>	<input type="checkbox"/>
- employés et agents administratifs ?	<input type="checkbox"/>	<input type="checkbox"/>	CONTR0A-B	<input type="checkbox"/>	<input type="checkbox"/>
- ouvriers et agents de production ?	<input type="checkbox"/>	<input type="checkbox"/>	CONTR0A-C	<input type="checkbox"/>	<input type="checkbox"/>

A6 Dans cet établissement, certaines catégories de salariés se succèdent-elles sur un même poste de travail ? (horaires aménagés, travail posté ou autres formules) SUCPOSTE

- Oui
 Non

A7 Certaines catégories de personnel sont-elles soumises à...

(A l'exception du personnel de gardiennage et de maintenance du lieu)

a. des permanences pour répondre à la clientèle, aux usagers ? PERMA

- Oui
 Non

b. des permanences pour surveiller un processus de production ou une activité de l'établissement ? PERMB

- Oui
 Non

c. des astreintes en dehors du lieu habituel de travail ? PERMC

- Oui
 Non

A8 Certains salariés travaillent-ils parfois à leur domicile ?

(en dehors des travailleurs à domicile proprement dits)

a. En cas de surcroît d'activité DOMA

- Oui
 Non

b. En cas de contraintes personnelles DOMB

- Oui
 Non

A9 A l'embauche, certains types de postes sont-ils prévus avec des horaires flexibles ? FLEX

(variables et définis d'une semaine à l'autre)

- Oui
 Non

B - Gestion du personnel

B1 A l'embauche, certains types de postes sont-ils prévus comme des postes à temps partiel ?

- Oui PARTEMB
 Non

B2 Lorsqu'il est demandé par le salarié, le temps partiel est-il accordé ? PARTACC

- Oui, à toute demande
 Seulement dans certains services ou à certaines catégories de personnel
 Non

B3 Dans cet établissement, quel est le nombre de jours de congés payés annuels ? (hors RTT, en jours ouvrés pour un salarié à temps complet, maximum et minimum si plusieurs formules)

- a. Minimum CONGNBMI jours
 b. Maximum CONGNBMA jours

B4 Dans cet établissement, quel a été le nombre de jours de congés RTT en 2004 ? (maximum et minimum si plusieurs formules)

- a. Minimum RTTNBMI jours
 b. Maximum RTTNBMA jours
 Pas de RTT ⇒ B6

B5 Les jours RTT peuvent-ils être...

a. ajoutés aux congés ordinaires à n'importe quel moment de l'année ? RTTPRISA

- Oui
 Oui, sauf période particulière
 Non

b. pris systématiquement le même jour de la semaine (par exemple les mercredis) ? RTTPRISB

- Oui
 Non

B6 En cas d'enfant malade, les parents ont-ils droit à des jours de congé rémunérés spécifiques ? ENFMAL

- Oui, sans condition sur les droits du conjoint
 Oui, selon les droits du conjoint
 Non
 Ne sait pas

Si oui, combien au maximum par an ? jours
 ENFMALNB

B7 Un salarié de cet établissement peut-il bénéficier d'aménagements ou d'arrangements d'horaires ? (en dehors des obligations légales, de dérogations)

	Oui	Non	Ne sait pas
- Le jour de la rentrée scolaire	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> AMENA
- En raison des horaires d'école, de crèche, etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> AMENB
- En cas d'enfant malade	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> AMENC
- Si ses trajets entre son domicile et le travail sont particulièrement longs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> AMEND
- Pour problèmes de santé ou un handicap personnel	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> AMENE
- Pour accompagner un proche en fin de vie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> AMENF
- Autre préciser AMENG			

B8 La situation familiale des salariés de l'établissement est-elle connue par le service du personnel (ou le directeur de l'établissement) ?

	Oui, pour tous les salariés	Oui, pour certains	Non
- Leur situation conjugale	<input type="checkbox"/> SPCONJ	<input type="checkbox"/>	<input type="checkbox"/>
- Le nombre d'enfants	<input type="checkbox"/> SPNBENF	<input type="checkbox"/>	<input type="checkbox"/>
- L'âge du dernier enfant	<input type="checkbox"/> SPAGEDER	<input type="checkbox"/>	<input type="checkbox"/>

B9 Au cours de l'année 2004, combien de salariées ont pris un congé de maternité ? **MATER**
(Si ne sait pas noter NSP)

B10 Au cours de l'année 2004, combien de salariés hommes ont pris pour une naissance...
(Si ne sait pas noter NSP) **PATERA**

a. seulement le congé de 3 jours ?

b. le congé de 3 jours plus un congé de paternité (voir notice) ? **PATERB**

➤ Si établissement de la Fonction publique ⇒ aller en B12

B11 Dans cet établissement, lors de son congé de maternité ou paternité, un(e) salarié(e) dont le salaire dépasse le plafond de la sécurité sociale bénéficie-t-il(elle), de votre part, d'un maintien de salaire ou d'un complément d'indemnisation ? (voir notice) **CONGSAL**

Oui

Non

Ne sait pas

BIIM

BIIP

B12 Dans cet établissement, la durée normale du congé de maternité accordé pour un premier enfant est-elle supérieure au congé légal (16 semaines) ? (hors congés pathologiques, naissances multiples et congé parental) **CONGDUR**

Oui

Non

Si oui, nombre total de semaines ?
(si ne sait pas noter NSP) **CONGDURNB**

B13 Dans cet établissement, combien de jours de congés supplémentaires sont accordés au maximum à un salarié... (Si ne sait pas noter NSP)

a. qui se marie ? **CONGMAR** jours

b. qui déménage pour raisons personnelles ? jours
CONGDEM

B14 La vie familiale des salariés est-elle prise en compte sur les points suivants ?

	Oui, pour tous les salariés	Oui, dans certains cas	Non	Sans objet
- Organisation des horaires de travail	<input type="checkbox"/>	VIEFAMA	<input type="checkbox"/>	<input type="checkbox"/>
- Missions et déplacements	<input type="checkbox"/>	VIEFAMB	<input type="checkbox"/>	<input type="checkbox"/>
- Changements de postes, mobilités internes	<input type="checkbox"/>	VIEFAMC	<input type="checkbox"/>	<input type="checkbox"/>
- Attribution du temps partiel	<input type="checkbox"/>	VIEFAMD	<input type="checkbox"/>	<input type="checkbox"/>
- Planification des congés	<input type="checkbox"/>	VIEFAME	<input type="checkbox"/>	<input type="checkbox"/>

B15 Lors d'un recrutement, cet établissement souhaite-t-il disposer des informations complémentaires suivantes ?

	Oui	Non
- Situation conjugale de la personne RECRUTA	<input type="checkbox"/>	<input type="checkbox"/>
- Nombre de ses enfants RECRUTB	<input type="checkbox"/>	<input type="checkbox"/>
- Age de ses enfants RECRUTC	<input type="checkbox"/>	<input type="checkbox"/>
- Activité de son conjoint RECRUTD	<input type="checkbox"/>	<input type="checkbox"/>

B16 Dans cet établissement, ces critères sont-ils déterminants pour la promotion ou l'avancement des salariés ?

	Toujours	Souvent	Parfois	Jamais
- L'ancienneté PROMOA	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- La formation continue PROMOB	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- La disponibilité (horaires, déplacements) PROMOC	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- La mobilité géographique PROMOD	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- L'efficacité, la tenue des objectifs PROMOE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

B17 Un entretien annuel d'évaluation est-il organisé entre le service du personnel et les salariés ? **ENTRETI**

Oui, pour tous les salariés

Oui, seulement pour les cadres

Non

B18 Y-a-t-il dans l'établissement, un service du personnel et/ou une direction des ressources humaines ? **DRH**

Oui

Non

C - Environnement de travail

C1 Y a-t-il dans cet établissement...

	Oui	Non
- un comité d'entreprise ? CENT	<input type="checkbox"/>	<input type="checkbox"/>
- un comité d'établissement ? CETAB	<input type="checkbox"/>	<input type="checkbox"/>
- une délégation du personnel ? DELEGP	<input type="checkbox"/>	<input type="checkbox"/>
- des délégués syndicaux de salariés ? DELEGSYN	<input type="checkbox"/>	<input type="checkbox"/>
- une association pour les oeuvres sociales ? ASSOC	<input type="checkbox"/>	<input type="checkbox"/>
- un comité d'hygiène et de sécurité ? COMHS	<input type="checkbox"/>	<input type="checkbox"/>

C2 Quel a été le pourcentage des sommes versées au comité d'entreprise ou son équivalent pour ses oeuvres sociales par rapport à votre masse salariale en 2004 ? **PCTCE**

Moins de 1 %	<input type="checkbox"/>
De 1 % à moins de 2 %	<input type="checkbox"/>
De 2 % à moins de 3 %	<input type="checkbox"/>
Plus de 3 %	<input type="checkbox"/>
Pas de C.E. ou équivalent	<input type="checkbox"/>
Ne sait pas	<input type="checkbox"/>

C3 a. Quel a été le pourcentage des sommes versées pour la formation continue par rapport à votre masse salariale en 2004 ? PCTCEAP

- Moins de 1 %
- De 1 % à moins de 2 %
- De 2 % à moins de 3 %
- Plus de 3 %
- Pas de formation continue
- Ne sait pas

b. Quel a été le temps total de formation en 2004 ? (si ne sait pas noter NSP)

en jours [] ou en heures []
TPSFORMJ TPSFORMH

C4 Cet établissement fait-il bénéficier des salariés...

	Oui	Non
- d'un logement de fonction ? LOGE	<input type="checkbox"/>	<input type="checkbox"/>
- d'une voiture de fonction ? VOITU	<input type="checkbox"/>	<input type="checkbox"/>
- d'un ordinateur portable ou de bureau pour leur domicile ? ORDI	<input type="checkbox"/>	<input type="checkbox"/>
- d'un téléphone portable ? PORTA	<input type="checkbox"/>	<input type="checkbox"/>
- d'un accès internet payé pour leur domicile ? INTERNET	<input type="checkbox"/>	<input type="checkbox"/>

C5 Existe-t-il dans cet établissement ou à proximité...

	Oui	Non
- une cantine ou un restaurant d'entreprise ? CANTI	<input type="checkbox"/>	<input type="checkbox"/>
- un distributeur de boissons ? DISTRI	<input type="checkbox"/>	<input type="checkbox"/>
- une salle de détente ? SALLE	<input type="checkbox"/>	<input type="checkbox"/>
- une pièce repas avec micro-ondes ? PIECE	<input type="checkbox"/>	<input type="checkbox"/>

C6 Existe-t-il dans l'établissement une participation pour les frais de repas de midi ? (plusieurs réponses possibles)

- Une subvention à un restaurant d'entreprise .FMIDI-1
- Un panier repasFMIDI-2
- Des chèques restaurantFMIDI-3
- Autre, préciser FMIDI-4
- NonFMIDI-5

C7 Cet établissement participe-t-il au financement d'une mutuelle pour ses salariés ? FMUTU

- Oui, pour l'ensemble des salariés
- Oui, seulement pour les cadres
- Non

C8 Existe-t-il, relié à cet établissement, pour les besoins des salariés...

	Oui	Non
- des services mutualistes ? MUTU	<input type="checkbox"/>	<input type="checkbox"/>
- une coopérative ? COOP	<input type="checkbox"/>	<input type="checkbox"/>
- une association sportive ou culturelle ? SPORT	<input type="checkbox"/>	<input type="checkbox"/>
- une garderie, un centre aéré ? GARDE	<input type="checkbox"/>	<input type="checkbox"/>
- une crèche ou des places en crèche ? CRECHE	<input type="checkbox"/>	<input type="checkbox"/>
⇒ Nombre de places [] CRECHENB		

C9 Certains salariés peuvent-ils bénéficier par l'employeur, le comité d'entreprise, la mutuelle ou un autre organisme de gestion des oeuvres sociales...

	Oui	Non
- de primes de naissance ? PRESTAA	<input type="checkbox"/>	<input type="checkbox"/>
- de primes de mariage ? PRESTAB	<input type="checkbox"/>	<input type="checkbox"/>
- de primes d'habillement ? PRESTAC	<input type="checkbox"/>	<input type="checkbox"/>
- d'une aide financière aux frais de garde des enfants ? PRESTAD	<input type="checkbox"/>	<input type="checkbox"/>
- d'une aide financière aux études des enfants ? PRESTAE	<input type="checkbox"/>	<input type="checkbox"/>
- de services domestiques (blanchisserie, ménage, etc.) ? PRESTAF	<input type="checkbox"/>	<input type="checkbox"/>
- de prestations pour un enfant handicapé ? PRESTAG	<input type="checkbox"/>	<input type="checkbox"/>
- de chèques vacances ? PRESTAH	<input type="checkbox"/>	<input type="checkbox"/>
- d'un accès à des centres ou colonies de vacances ? PRESTAI	<input type="checkbox"/>	<input type="checkbox"/>
- d'un accès à des logements sociaux ? PRESTAJ	<input type="checkbox"/>	<input type="checkbox"/>
- de contrats d'assurance-vie, de prévoyance ? PRESTAK	<input type="checkbox"/>	<input type="checkbox"/>
- de plans d'épargne spécifiques ? PRESTAL	<input type="checkbox"/>	<input type="checkbox"/>
- de plans de retraite ? PRESTAM	<input type="checkbox"/>	<input type="checkbox"/>

C10 Au cours de l'année 2004, cet établissement a-t-il participé à l'organisation pour ses salariés...

	Oui	Non
- de rencontres sportives ? ORGA	<input type="checkbox"/>	<input type="checkbox"/>
- d'événements culturels ? ORGB	<input type="checkbox"/>	<input type="checkbox"/>
- de spectacles ou de fêtes de Noël pour les enfants ? ORGC	<input type="checkbox"/>	<input type="checkbox"/>
- d'autres événements (voyages, loisirs, soirées...) ? ORGD	<input type="checkbox"/>	<input type="checkbox"/>

C11 La création d'une crèche ou de places en crèche est-elle prévue pour cet établissement ? CRECHECR

(et éventuellement places supplémentaires s'il y en a déjà)

- Oui
- Non
- Ne sait pas

Si oui : a. Dans combien d'années ? CRECHEAN []

b. Combien de places ? CRECHEPL []

(si ne sait pas noter NSP)

➤ Si établissement de la Fonction publique ⇒ aller en D

C12 Cet établissement bénéficie-t-il d'une déduction fiscale... (voir notice)

a. pour augmenter l'offre de garde ? (crèches d'entreprise/ inter-entreprises, halte-garderies, financements conjoints avec les communes ou les CAF) DEDUCA

- Oui
- Non
- Ne sait pas

b. pour favoriser des actions à caractère familial ?

(compléments financiers versés pendant les congés parentaux, actions de formation pour favoriser un retour à l'emploi des personnes après leur congé parental, etc.) DEDUCB

- Oui
- Non
- Ne sait pas

D - Caractéristiques de l'établissement

➤ Si cet établissement est le seul établissement de l'entreprise ou de l'organisme de la fonction publique (voir notice) ⇒ aller en D2

D1 L'autonomie de décision de cet établissement par rapport au service de tutelle, au siège, au groupe ou à un autre établissement de l'entreprise, est-elle très, assez ou pas importante en matière...

Autonomie de l'établissement en matière :	Très importante	Assez importante	Peu ou pas importante
- d'investissement ?	<input type="checkbox"/>	AUTONA <input type="checkbox"/>	<input type="checkbox"/>
- de gestion des emplois (embauches, nombre de postes à pourvoir, licenciements) ?	<input type="checkbox"/>	AUTONB <input type="checkbox"/>	<input type="checkbox"/>
- d'évolution des salaires et de distribution de primes ?	<input type="checkbox"/>	AUTONC <input type="checkbox"/>	<input type="checkbox"/>
- de mobilité interne des salariés ?	<input type="checkbox"/>	AUTOND <input type="checkbox"/>	<input type="checkbox"/>
- d'organisation du temps de travail ?	<input type="checkbox"/>	AUTONE <input type="checkbox"/>	<input type="checkbox"/>
- de gestion des oeuvres sociales ?	<input type="checkbox"/>	AUTONF <input type="checkbox"/>	<input type="checkbox"/>

D2 Le directeur actuel de l'établissement est... DIRSEXE

- un homme ?
 une femme ?

a. Il ou elle a... DIRAGE

- moins de 30 ans ?
 30 à 40 ans ?
 40 à 50 ans ?
 plus de 50 ans ?

b. Quelle est son ancienneté dans ce poste ?

DIRANC ans

D3 Dans son activité principale, diriez-vous que le marché ou le public de cet établissement est plutôt... MARCHE

- régional ou local ?
 national ?
 européen ?
 international ?
 Ne sait pas

D4 L'activité de cet établissement connaît-elle des fluctuations importantes ?

a. Annuelles FLUCAN

- Oui
 Non

b. Saisonnières FLUCSAI

- Oui
 Non

c. Hebdomadaires FLUCHEB

- Oui
 Non

D5 Cet établissement utilise-t-il les dispositifs organisationnels suivants ?

a. Certification ou démarche de qualité totale (certification ISO, EAQF, etc.) DISPOSIA

- Oui, depuis plus de 3 ans
 Oui, depuis moins de 3 ans
 Non
 Ne sait pas

b. Système de livraison ou de production de type "juste à temps" DISPOSIB

- Oui, depuis plus de 3 ans
 Oui, depuis moins de 3 ans
 Non
 Ne sait pas

c. Groupes de travail ou de projets autonomes, équipes de résolution de problèmes DISPOIC

- Oui, depuis plus de 3 ans
 Oui, depuis moins de 3 ans
 Non
 Ne sait pas

D6 Dans cet établissement, a-t-on...

a. développé la polyvalence et la rotation des postes de travail ? POLYVA

- Oui, depuis plus de 3 ans
 Oui, depuis moins de 3 ans
 Non

b. mis en place ou changé les systèmes d'informatisation ? INFOR

(grand système informatique centralisé, Internet, intranet ou messagerie électronique interne, mise en réseau des micro-ordinateurs)

- Oui, depuis plus de 3 ans
 Oui, depuis moins de 3 ans
 Non

➤ Si établissement de la Fonction publique ⇒ aller en D8

D7 Au cours des trois dernières années, cet établissement a-t-il connu un plan de licenciement ou un plan social ? PLANLIC1

- Oui
 Non

D8 a. Cet établissement a-t-il déménagé dans les trois dernières années ? DEMEN

- Oui
 Non

Si non ⇒ aller en E

b. Y a-t-il eu une politique d'accompagnement pour les salariés (transports, modes de garde, horaires, prime de déménagement...) ? DEMENPOL

- Oui
 Non

Si non ⇒ aller en E

c. Quel a été son coût financier total ? DEMENCT (si ne sait pas noter NSP)

euros

d. Quel a été le nombre de salariés concernés ?

DEMENN salariés

E - Effectifs de l'établissement enquêté

➤ On s'intéresse au nombre total de personnes et non à leur équivalent temps plein au 31 décembre 2004 (voir notice pour la définition des effectifs demandés). Si l'effectif est nul, noter 0.
 Vous pouvez indiquer le nombre ou le % (même approximatif)
 Les questions E2 à E9 portent sur l'effectif déclaré en E1a.

E1 a. Dans cet établissement, quel est le nombre de salariés ou d'agents publics (au 31 décembre 2004) ?

b. dont combien de femmes ? ou %

E2 Combien de salariés sont âgés de...
 moins de 20 ans ou %
 20 à 29 ans ou %
 30 à 49 ans ou %
 50 ans et plus ou %

E3 a. Combien y a-t-il de cadres EFCAD ou catégories A ? (voir notice)

b. dont combien de femmes ? ou %

E4 Combien y a-t-il d'apprentis, de contrats de qualification ou de professionnalisation ?

E5 Combien de salariés ont plus de 5 ans d'ancienneté dans cet établissement ? ou %

E6 a. Combien de salariés ont un contrat à durée indéterminée (CDI) ou sont titulaires de la fonction publique ? (y compris fonctionnaires stagiaires)

b. dont combien de femmes ? ou %

E7 a. Combien de salariés ont un contrat à durée déterminée (CDD) ?

b. dont combien de femmes ? ou %

E8 a. Combien de salariés sont à temps partiel ?

b. dont combien de femmes ? ou %

E9 a. Combien de salariés sont en congé parental ?
 - A temps plein EFCPPL
 - A temps partiel EFCPPA
b. dont combien de femmes ?
 - A temps plein EFCPPLF
 - A temps partiel EFCPPAF

E10 a. Au cours de l'année 2004, cet établissement a-t-il fait appel à des intérimaires ou des vacataires ? VAC

Oui
 Non

Si oui, combien ? VACNB

b. Certains de ces intérimaires ou vacataires ont-ils été embauchés ? VACEMB

Oui
 Non

Si oui, combien ? VACEMBNB

E11 Au cours de l'année 2004, certains CDD ont-ils été transformés en CDI ou titularisés dans cet établissement ?

Oui CDDCDI
 Non CDDCDINB

Si oui, combien ?

E12 Au cours de l'année 2004, cet établissement a-t-il fait appel à des sociétés prestataires de services extérieurs sur le site ?

Oui
 Non

Si oui, combien de sociétés prestataires ? PRESTA
 PRESTANB

E13 Au cours de l'année 2004, cet établissement a-t-il eu recours à...

	Oui	Non
- des heures supplémentaires ou complémentaires ?	<input type="checkbox"/>	<input type="checkbox"/> RECA
- l'embauche de saisonniers ?	<input type="checkbox"/>	<input type="checkbox"/> RECB
- la modulation de l'activité et des horaires (alternance de périodes de haute et basse activité) ?	<input type="checkbox"/>	<input type="checkbox"/> RECC
- du chômage technique ?	<input type="checkbox"/>	<input type="checkbox"/> RECD
- des licenciements (personnels ou économiques) ?	<input type="checkbox"/>	<input type="checkbox"/> RECE
- la sous-traitance ?	<input type="checkbox"/>	<input type="checkbox"/> RECF

E14 Au cours de l'année 2004, combien y a-t-il eu de recrutements ou d'arrivées dans cet établissement ? (Y compris mutations)

Dont combien : a. de femmes ?

b. de moins de 30 ans ?

E15 Les personnes recrutées reçoivent-elles un livret d'accueil ? LIVRET

Oui
 Non

E16 Au cours de l'année 2004, combien y a-t-il eu de départs dans cet établissement ?

dont :
 a. Départs en retraite et pré-retraite
 b. Démissions
 c. Licenciements
 d. Fins de contrat
 e. Mutations, mobilité vers un autre établissement
 f. Autres

E17 D'une année à l'autre, prévoir l'évolution des besoins de main-d'oeuvre de cet établissement est-il ? *EVOMDO*

Très facile

Plutôt facile

Plutôt difficile

Très difficile

Ne sait pas

E18 Au cours des trois dernières années, dans cet établissement, quelles ont été les évolutions des effectifs ? *(voir notice)*

	En hausse	Stable	En baisse
- De l'ensemble des salariés ? <i>EVOEFA</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- De cadres ou catégories A ? <i>EVOEFB</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
- De non cadres ou catégories B et C ? <i>EVOEFC</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

E19 Quelle est la rémunération des salariés à temps plein de cet établissement ? *(y compris indemnités, primes)*

a. Rémunération brute mensuelle minimum *(hors stagiaires, vacataires, apprentis et jeunes en formation, intérimaires) :*

SMIC SMIC ou euros

b. Rémunération brute mensuelle maximum *(hors cadres dirigeants) :*

euros

E20 Certains salariés peuvent-ils bénéficier d'un intéressement ou d'une participation aux résultats ?

Oui

Non *PARTICIP*

F - Renseignements concernant l'entreprise ou l'organisme de la fonction publique

Si cet établissement est le seul établissement de l'entreprise (voir notice) ⇒ aller en F5

F1 Quel est le code NAF (ou APE) de l'entreprise ou organisme ? *ENTNAF*

CODE-NAF identique à celui de l'établissement ou

F2 Quel est l'intitulé en clair de son activité principale *(si différent de l'établissement) ? ENTACT*

.....

.....

F3 Quel était le nombre de salariés total de l'entreprise ou organisme en France au 31 décembre 2004 ? *ENTELF*

Si vous ne savez pas exactement ENTFCFL

de 20 à 49 salariés

de 50 à 199 salariés

de 200 à 499 salariés

de 500 à 1999 salariés

2000 salariés et plus

F4 L'entreprise ou organisme de la fonction publique a-t-elle/il d'autres établissements... *ENTETDEP*

a. dans le département ou les départements voisins ?

Oui *ENTETDEP*

Non

b. ailleurs en France ?

Oui *ENTETFRA*

Non

c. à l'étranger ?

Oui *ENTETETR*

Non

F5 Votre entreprise ou organisme de la fonction publique pratique-t-elle/il une politique systématique d'aides financières ou en nature pour favoriser une mobilité géographique ? *ENTETMOB*

Oui, pour tous les salariés

Oui, seulement pour les cadres

Non, pas d'aides

Pas de mobilité

Ne sait pas

Si fonction publique ⇒ aller en G

F6 Cette entreprise... *(plusieurs réponses possibles)*

exerce en franchise *ENTSITUA-A*

est liée à un groupement (GIE, association d'entreprises, centrale d'achat, concession, agent de marques...) ... *ENTSITUA-B*

appartient à un groupe français *ENTSITUA-C*

appartient à un groupe international *ENTSITUA-D*

est totalement indépendante *ENTSITUA-E*

Ne sait pas *ENTSITUA-F*

F7 Depuis combien d'années l'entreprise a-t-elle été créée ? *(création historique) ENTDOUR*

3 ans ou moins

4 à 10 ans

11 à 20 ans

plus de 20 ans

Ne sait pas

F8 Quel était le chiffre d'affaires de l'entreprise en 2004 *(en millions d'Euros HT) ? ENTCA*

Moins de 0,2

De 0,2 à moins de 0,5

De 0,5 à moins de 1

De 1 à moins de 5

De 5 à moins de 20

De 20 à moins de 50

De 50 à moins de 200

200 ou plus

Sans objet

F9 Quelle était la masse salariale brute de l'entreprise au 31 décembre 2004 ? *ENTMASAL*

euros

G - Conclusion

G1 Un employeur doit-il aider les salariés à coordonner leur travail avec leur vie familiale ? **COOR**
 Oui, absolument
 Oui, dans certains cas
 Non, ce n'est pas son rôle
 Autre réponse

Si oui, pourquoi ? (plusieurs réponses possibles)
 Cela attire et/ou fidélise les salariés **COORRASA**
 Cela améliore productivité et performance **COORRASB**
 Cela réduit l'absentéisme **COORRASC**
 Cela sert l'image de l'entreprise ou de l'organisme **COORRASD**
 Cela favorise le bien-être des salariés, réduit leur stress **COORRASE**
 Autre, préciser **COORRASF**

Dans un établissement comme le vôtre, certaines mesures pourraient-elles aider les salariés à mieux concilier travail et vie familiale ? **COORMES**
 Oui
 Non

Si oui, lesquelles ?
COORMES2

Personne répondant à ce questionnaire

<p>1. Nom _____</p> <p>2. Prénom _____</p> <p>3. Téléphone <input type="text"/> <input type="text"/></p> <p>4. E-mail ou courriel REPMAIL _____</p> <p>5. Fonction actuelle dans l'entreprise ou l'établissement REPFON _____</p>	<p>➤ Si vous êtes le directeur de l'établissement ⇒ aller en 7</p> <p>6. En quelle année êtes-vous entré(e) dans cet établissement ? REPANC <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/></p> <p>7. Accepteriez-vous de rencontrer un chercheur pour un entretien sur le thème des interactions entre vie professionnelle et vie familiale ? Oui, éventuellement <input type="checkbox"/> Non <input type="checkbox"/> REPSUITE</p>
--	--

< **Nous vous remercions vivement d'avoir répondu à ce questionnaire** >

Si vous avez des remarques, merci de les formuler ci-dessous :

REMARQUES



Ce questionnaire doit être retourné, dûment rempli, dans l'enveloppe réponse ci-jointe à :

INED
 Service des enquêtes
 133, boulevard Davout - 75680 PARIS Cedex 20
 Contacts : 01 56 06 22 94 Martine Quaglia, 01 56 06 22 74 Ariane Pailhé, 01 56 06 21 30 Anne Solaz
 e.mail : enquete.efe@ined.fr

Annexe B

Annexe non-réponse

B.1 La non-réponse simultanée à deux variables

	1)	2)	3)	4)	5)	6)	7)	8)	9)	10)	11)	12)	13)	14)	15)	16)	17)	18)	19)	20)	21)	22)
1)																						
2)	0.64																					
3)	0.52	0.22																				
4)	0.41	0.26	0.60																			
5)	0.49	0.37	0.94	0.71																		
6)	0.41	0.41	1.08	0.71	2.21																	
7)	0.49	0.37	1.27	0.75	2.88	2.47																
8)	0.49	0.37	1.16	0.67	2.81	2.66	3.25															
9)	0.56	0.56	1.27	0.86	3.25	2.54	3.59	3.29														
10)	0.79	0.49	0.34	0.26	0.64	0.67	0.75	0.64	0.79													
11)	1.57	0.90	0.97	1.27	1.35	1.31	1.72	1.53	1.95	4.04												
12)	1.31	0.67	0.75	0.79	1.1	0.90	1.20	0.94	1.50	3.29	6.62											
13)	1.80	0.90	1.23	1.16	1.31	1.12	1.53	1.35	1.68	3.67	9.09	6.21										
14)	0.49	0.15	0.26	0.19	0.22	0.30	0.41	0.34	0.26	0.3	0.71	0.75	0.71									
15)	0.41	0.19	0.30	0.19	0.30	0.30	0.37	0.26	0.49	0.60	1.16	1.01	1.27	2.81								
16)	0.22	0.15	0.15	0.07	0.22	0.22	0.30	0.22	0.34	0.22	0.60	0.45	0.56	0.34	0.30							
17)	0	0.04	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.07	0.04	0.07	0.11	0.19	0.04	0						
18)	0.64	0.56	0.71	0.67	0.82	0.75	1.01	0.86	1.1	0.97	1.95	1.50	2.13	0.07	0.79	0.19	0.11					
19)	2.21	1.12	0.45	1.01	0.97	0.75	0.90	0.79	1.12	1.61	2.51	2.24	2.81	0.41	0.94	0.26	0	1.35				
20)	0.49	0.41	0.94	0.64	2.36	0.22	2.39	2.32	2.66	0.60	1.12	0.97	1.01	0.37	0.22	0.19	0.11	0.79	0.75			
21)	0.22	0.19	0.37	0.37	0.45	0.45	0.34	0.45	0.37	0.11	0.41	0.26	0.52	0.11	0.19	0.07	0.07	0.41	0.45	0.34		
22)	0.60	0.30	1.95	0.75	1.20	1.01	1.08	1.05	1.23	0.37	1.01	0.79	1.16	0.34	0.37	0.19	0.07	0.82	0.41	0.97	0.52	

FIG. B.1 – Part (en %) des établissements non-répondants pour deux variables

Lecture : 2.21% des établissements sont non-répondants à la fois pour l'existence d'une prime de naissance (6) et pour la prise en charge des frais de garde des enfants (7)).

Ces résultats confirment le caractère diffus de la non-réponse parmi les établissements.

B.2 La constitution des 8 grandes régions françaises

L'utilisation de ces 8 grandes régions est limitée à l'étape de traitement de la non-réponse. Il s'agit de prendre en compte un aspect régional tout en limitant le nombre de modalités. Le seul cas de non-réponse hors-champ a été affecté à la région regroupant le plus grand effectif (groupe 1).

Région 1	Région 2	Région 3	Région 4
Île de France	Champagne-Ardennes Picardie Haute-Normandie Nord-Pas-de-Calais	Lorraine Alsace Franche-Comté	Centre Bourgogne
Région 5	Région 6	Région 7	Région 8
Basse-Normandie Pays de la Loire Bretagne Poitou-Charentes	Aquitaine Midi-Pyrénées Limousin	Rhône-Alpes Auvergne	Languedoc-Roussillon Provence Alpes Côte d'Azur

B.3 Code SAS

Le code SAS fournit correspond à l'ensemble des opérations permettant d'obtenir à partir de la table initiale la base de données corrigée de la non-réponse utilisée pour réaliser la classification. L'algorithme de traitement de la non réponse par hot-deck métrique a été élaboré par M.Vanderschelden[20]. Elle fournit en annexe une partie de son code SAS. Le programme qui suit est une application de celui-ci au cas particulier qui nous intéresse. Nous n'avons pas reproduit ici le code correspondant aux variables supplémentaires.

```

*=====;
*création des variables actives;
*=====;

* 1) variable ENFMALDD;
*=====;
data entrepr.empl1;
set entrepr.employeur;
ENFMALdd=1*(ENFMAL=3 or ENFMALNB=0) + 2*((ENFMAL lt 3) and (1 le ENFMALNB le
3)) + 3*(ENFMAL
lt 3 ) and (4 le ENFMALNB le 6))+ 4*((ENFMAL lt 3) and ( 7 le ENFMALNB ) and
(ENFMALNB ne 88))
+ 8*((ENFMAL lt 3) and (ENFMALNB = 88)) + 9*(ENFMAL=8 or ENFMAL=9);
run;

* 2) congurd : construction congé supp maternité;
*=====;
data entrepr.empl1;
set entrepr.empl1;
congurd=1*(CONGDUR=2 or CONGDURNB=0) +
2*(CONGDUR =1 and (16 le CONGDURNB le 19)) + 2*((CONGDUR =1) and (CONGDURNB =
888))
+ 2* ((CONGDUR =1) and (CONGDURNB = 999))
+ 3*((CONGDUR =1 ) and (20 le CONGDURNB) and (CONGDURNB ne 888) and (CONGDURNB
ne 999))
+ 9*((CONGDUR =8) or (CONGDUR =9));
run;

* 3) gardentrd;
*=====;
*garde;
data entrepr.empl1;
set entrepr.empl1;
garded=0*(garde=2) + 1*(garde=1) + 9*(garde=8);
run;

*creche;
data entrepr.empl1;
set entrepr.empl1;
creched=0*(creche=2) + 1*(creche=1) + 9*(creche=8);
run;

*on impute la nr à la plus grande valeur soit pas de garde;
*effectifs trop petits de oui on va combiner les 2 variables;
data entrepr.empl1;
set entrepr.empl1;
if (garded=1 or creched=1) then do gardentrd=1;
end;
if (garded=9 and creched=9) then do gardentrd=9;
end;
if gardentrd=. then do gardentrd=0;
end;
run;

* 4) icreched : intention creation creche;
*=====;
data entrepr.empl1;
set entrepr.empl1;
icreched=0*(creche=2) + 1*(creche=1) + 9*(creche=8 or creche=9);
run;

```

```

* 5) prestaïd : colonies;
*=====;
data entrepr.empl1;
set entrepr.empl1;
prestaïd=0*(prestaï=2) + 1*(prestaï=1) + 9*(prestaï=8);
run;

*6) prestaad : prime naissance;
*=====;
data entrepr.empl1;
set entrepr.empl1;
prestaad=0*(presta=2) + 1*(presta=1) + 9*(presta=8);
run;

* 7) prestadd : frais de garde;
*=====;
data entrepr.empl1;
set entrepr.empl1;
prestadd=0*(prestad=2) + 1*(prestad=1) + 9*(prestad=8);
run;

* 8) prestaed : aide aux études;
*=====;
data entrepr.empl1;
set entrepr.empl1;
prestaed=0*(presta=2) + 1*(presta=1) + 9*(presta=8);
run;

* 9) prestagd : prestation enfant handicap;
*=====;
data entrepr.empl1;
set entrepr.empl1;
prestagd=0*(prestag=2) + 1*(prestag=1) + 9*(prestag=8);
run;

*10) congsalbd;
*=====;
data entrepr.empl1;
set entrepr.empl1;
congdb=0*(congsalb=2)+ 1*(congsalb=1) + 9*(congsalb=8) + 9*(congsalb=9);
run;

* 11) rentrée;
*=====;
data entrepr.empl1;
set entrepr.empl1;
amenad=0*(amena=2) + 1*(amena=1) + 9*(amena=8) + 9*(amena=9);
run;

* 12) creche;
*=====;
data entrepr.empl1;
set entrepr.empl1;
amenbd=0*(amenb=2) + 1*(amenb=1) + 9*(amenb=8) + 9*(amenb=9);
run;

* 13) enfant malade;
*=====;
data entrepr.empl1;
set entrepr.empl1;
amencd=0*(amenc=2) + 1*(amenc=1) + 9*(amenc=8) + 9*(amenc=9);
run;

```

```

* 14) trajets;
*=====;
data entrepr.empl1;
set entrepr.empl1;
amendd=0*(amend=2) + 1*(amend=1) + 9*(amend=8) + 9*(amend=9);
run;

* 15) possibilité de travailler à domicile;
*=====;
data entrepr.empl1;
set entrepr.empl1;
dombd=0*(domb=2) + 1*(domb=1) + 9*(domb=8);
run;

* 16) 17) tps partiel accordé NR sera répartie aléatoirement;
*=====;
data entrepr.empl1;
set entrepr.empl1;
partacc=1*(partacc=3)+2*(partacc=2) +3*(partacc=1) + 9*(partacc=8);
run;

data entrepr.empl1;
set entrepr.empl1;
partemb=0*(partemb=2)+1*(partemb=1) +9*(partemb=8);
run;

* 18) RTTdd;
*=====;
data entrepr.empl1;
set entrepr.empl1;
rttd=1*(rttso=1) + 2*(rttso=2) + 9*(rttso=8);
run;

data entrepr.empl1;
set entrepr.empl1;
if rttd=9 then rttprisa=-1; else rttprisa=rttprisa;
if rttd=9 then rttprisb=-1; else rttprisb=rttprisb;
if rttd=1 then rttprisa=-1; else rttprisa=rttprisa;
if rttd=1 then rttprisb=-1; else rttprisb=rttprisb;
if 1 lt rttprisb then rttprisbis=1; else rttprisbis=0;
if 1 lt rttprisa then rttprisais=1; else rttprisais=0;
run;

data entrepr.empl1;
set entrepr.empl1;
rttddd=1*(rttd=1) + 1*((rttprisa=8) and (rttprisb=2)) + 2*((rttprisa=3) and
(rttprisb=2)) + 2*((rttprisa=3) and (rttprisb=8)) + 3*((rttprisa=2) and
(rttprisbis=1))
+ 3*((rttprisa=1) and (rttprisbis=1)) + 3*((rttprisais=1) and (rttprisb=1)) +
3*((rttprisa=8) and (rttprisb=8))
+ 4*((rttprisa=1) and (rttprisb=1)) + 9*(rttd=9);
run;

* 19) avancement dispo;
*=====;
data entrepr.empl1;
set entrepr.empl1;
promocd=1*(promoc=4) + 2*(promoc=3) + 3*(promoc=2) + 4*(promoc=1) + 9*(promoc=8
or promoc=9);
run;

```

```

* 20)cheques vacances;
*=====;
data entrepr.empl1;
set entrepr.empl1;
prestahd=0*(prestah=2) + 1*(prestah=1) + 9*(prestah=8);
run;

* 21) 22) mutuelle (différentiation cadre et les autres);
data entrepr.empl1;
set entrepr.empl1;
fmutud=1*(fmutu=3)+ 2*(fmutu=2) + 3*(fmutu=1)+ 9*(fmutu=8);
run;

data entrepr.empl1;
set entrepr.empl1;
mutud=0*(mutu=2)+ 1*(mutu=1)+ 9*(mutu=8);
run;

*=====;
*=====;

```

```

*=====;
*Elimination des non ré pondant : création de la table score0;
*=====;

*score;
data entrepr.empl1;
set entrepr.empl1;
if prestaed=9 then s5=1;
else s5=0;
if prestaad=9 then s6=1;
else s6=0;
if prestadd=9 then s7=1;
else s7=0;
if prestaed=9 then s8=1;
else s8=0;
if prestagd=9 then s9=1;
else s9=0;
if enfmaldd=9 then s1=1;
else s1=0;
if congduard=9 then s2=1;
else s2=0;
if gardentrd=9 then s3=1;
else s3=0;
if icreched=9 then s4=1;
else s4=0;
if amenad=9 then s10=1;
else s10=0;
if amenbd=9 then s11=1;
else s11=0;
if amencd=9 then s12=1;
else s12=0;
if amendd=9 then s13=1;
else s13=0;
if dombd=9 then s14=1;
else s14=0;
if partaccd=9 then s15=1;
else s15=0;
if partembd=9 then s16=1;
else s16=0;
if rttddd=9 then s17=1;
else s17=0;
if promocd=9 then s18=1;
else s18=0;
if congdd=9 then s19=1;
else s19=0;
if prestahd=9 then s20=1;
else s20=0;
if fmutud=9 then s21=1;
else s21=0;
if mutud=9 then s22=1;
else s22=0;
run;

data entrepr.empl1;
set entrepr.empl1;
score1=s5 + s6 + s7 + s8 + s9;
score2=s10 + s11 + s12 + s13;
scoretot= s1 + s2 + s3 + s4 + s5 + s6 + s7 + s8 + s9 + s10
+ s11 + s12 + s13 + s14 + s15 + s16 + s17 + s18 + s19+ s20+ s21 + s22;
run;

```

```

data entrepr.empl1;
set entrepr.empl1;
if scoretot lt 11 then selection=1;
else selection=0;
run;

data entrepr.score0;
set entrepr.empl1;
if selection=1;
run;

*repondération. Sum=2643.69;
*=====;

proc means sum data=entrepr.score0;
var poidset_no;
run;

data entrepr.score0;
set entrepr.score0;
poids=(poidset_no/2643.69)*2652;
run;

*distinction des obs touchées par la NR;
*=====;
data entrepr.score0;
set entrepr.score0;
if 1 le scoretot then correct=1;
else correct=0;
run;

*=====;
*=====;

```

```

*=====;
*=variables auxiliaires=;
*=====;

*1)poids (par déciles);
*=====;
proc univariate data=entrepr.score0;
var poids;
output out=t pctlpts=10 20 30 40 50 60 70 80 90 pctlpre=P;
run;

data entspr.score0;
set entspr.score0;
if 0 le poids le 0.0549151761 then classe=1;
if 0.0549151761 lt poids le 0.1235974242 then classe=2;
if 0.1235974242 lt poids le 0.2145841102 then classe=3;
if 0.2145841102 lt poids le 0.3439116016 then classe=4;
if 0.3439116016 lt poids le 0.5032951617 then classe=5;
if 0.5032951617 lt poids le 0.7552501945 then classe=6;
if 0.7552501945 lt poids le 1.0756930628 then classe=7;
if 1.0756930628 lt poids le 1.5953774917 then classe=8;
if 1.5953774917 lt poids le 2.4798287083 then classe=9;
if 2.4798287083 lt poids then classe=10;
run;

*2) var. tgroupe;
*=====;
data entspr.score0;
set entspr.score0;
if entsitua_c=1 & pubpriv=2 then tgroupe=3;/*groupe international*/
else if entsitua_d=1 & pubpriv=2 then tgroupe=2;/*groupe français*/
else if (entsitua_a=1 or entsitua_b or entsitua_e) & pubpriv=2 then
tgroupe=1;/*autre
prive lucratif*/
else if pubpriv=1 then tgroupe=4;/*prive non lucratif*/
else if pubpriv=3 then tgroupe=8;/*fonction publique etat*/
else if pubpriv=4 then tgroupe=6;/*hospitalière*/
else if pubpriv=5 then tgroupe=7;/*territoriale*/
else if pubpriv=6 then tgroupe=5;/*nationalise*/
run;

```

```
*3) var. NES16d (14 modalités) ;
*=====;
```

```
data entrepr.score0;
set entrepr.score0;
if nes16='EA' then nes16d= 'EA';
if nes16='EB' then nes16d= 'EB';
if nes16='EC' then nes16d= 'EC';
if nes16='ED' then nes16d='ED';
if nes16='EE' then nes16d='EE';
if nes16='EF' then nes16d='EF';
if nes16='EG' then nes16d= 'EG';
if nes16='EH' then nes16d= 'EH';
if nes16='EJ' then nes16d= 'EJ';
if nes16='EK' then nes16d='EK';
if nes16='EL' then nes16d='EL';
if nes16='EM' then nes16d='EM';
if nes16='EN' then nes16d='EN';
if nes16='EP' then nes16d='EP';
if nes36='Q1' then nes16d= 'Q1';
if nes36='Q2' then nes16d= 'Q2';
if nes16='ER' then nes16d='ER';
run;
```

```
*4) region;
data entrepr.score0;
set entrepr.score0;
if region=11 then rrrr=1;
if region=. then rrrr=1;
if region=21 then rrrr=2;
if region=22 then rrrr=2;
if region=23 then rrrr=2;
if region=24 then rrrr=4;
if region=25 then rrrr=5;
if region=26 then rrrr=4;
if region=31 then rrrr=2;
if region=41 then rrrr=3;
if region=42 then rrrr=3;
if region=43 then rrrr=3;
if region=52 then rrrr=5;
if region=53 then rrrr=5;
if region=54 then rrrr=5;
if region=72 then rrrr=6;
if region=73 then rrrr=6;
if region=74 then rrrr=6;
if region=82 then rrrr=7;
if region=83 then rrrr=7;
if region=91 then rrrr=8;
if region=93 then rrrr=8;
run;
```

```
*=====;
*=====;
```

```

*=====;
*Traitement partiel de la NR pour la variable agrégée ENFMALDD;
*=====;

*dichotomisation des var. explicatives;
*=====;

data entrepr.score0;
set entrepr.score0;
if tgrroupe=1 then tgr1=1; else tgr1=0;
if tgrroupe=3 then tgr3=1; else tgr3=0;
if tgrroupe=4 then tgr4=1; else tgr4=0;
if tgrroupe=5 then tgr5=1; else tgr5=0;
if tgrroupe=6 then tgr6=1; else tgr6=0;
if tgrroupe=7 then tgr7=1; else tgr7=0;
if tgrroupe=8 then tgr8=1; else tgr8=0;
run;

data entrepr.score0;
set entrepr.score0;
if amenc=1 then amen1=1; else amen1=0;
if amenc=8 then amen8=1; else amen8=0;
if amenc=9 then amen9=1; else amen9=0;
run;

data entrepr.score0;
set entrepr.score0;
if congmar=888 then c=1;
if congmar=999 then c=1;
run;

data entrepr.score0;
set entrepr.score0;
if c=1 then congmar=0; else congmar=congmar;
drop c;
run;

data entrepr.score0;
set entrepr.score0;
if congsalb=1 then cong1=1; else cong1=0;
if congsalb=8 then cong8=1; else cong8=0;
if congsalb=9 then cong9=1; else cong9=0;
run;

*estimation globale;
*=====;

data entrepr.reg0;
set entrepr.score0;
if 2 le enfmaldd le 4;
run;

proc reg data=entrepr.reg0;
model enfmaln= poids
tgr1 tgr3 tgr4 tgr5 tgr6 tgr7 tgr8
cong1 cong8 cong9 congmar;
run;

```

```

*Calcul de la prévision;
*=====;

data entrepr.score1;
set entrepr.score0;
estim= 0.06038*poids
-0.15633*tgr1 + 0.85205*tgr3 + 2.16043*tgr4 + 3.62496*tgr5 + 6.38891*tgr6 +
5.35778*tgr7 + 6.39772*tgr8
+ 1.04906*cong1 + 1.15733*cong8 + 0.28363*cong9 +0.25382*congmard + 2.58891;
run;

data entrepr.score1;
set entrepr.score1;
estimtr= 2*(estim le 3) + 3*(3 lt estim le 6)+ 4*(6 lt estim );
run;

*Affectation de la prévision aux NR;
*=====;

data entrepr.score1;
set entrepr.score1;
if enfmaldd=8 then change=1; else change=0;
run;

data entrepr.score1;
set entrepr.score1;
if change=1 then enfmaldd=estimtr; else enfmaldd=enfmaldd;
run;

*=====;
*=====;

```

```

*=====;
*Traitement de la NR par hot-deck metrique ;
(utilisation de la méthode développée par M. Vanderschelden);
*=====;

*Constitution de 3 cas;
*=====;

data entrepr.score1;
set entrepr.score1;
cas= 1*(correct=1 and score1=score2=0) + 2*(correct=1 and score1 + score2 =
scoretot);
run;

data entrepr.score1;
set entrepr.score1;
receveur=correct+ cas;
run;

data entrepr.score1;
set entrepr.score1;
if scoretot=0 then donneur1=1; else donneur1=0;
if score1+score2=scoretot then donneur2=1; else donneur2=0;
if score1=score2=0 then donneur3=1; else donneur3=0;
run;

data entrepr.prep;
set entrepr.score1;
num=numident;
keep numident num receveur donneur1 donneur2 donneur3
ENFMAldd congduurd gardentrd icreched prestaidd
prestaad prestadd prestaed prestagd amenad
amenbd amencd amendd dombd partaccd
partembd rttddd promocd congdu prestahd
fmutud mutud
classe tgroupe nes16d rrrr poids;
run;

data entrepr.cas2;
set entrepr.prep;
if (receveur=2 or donneur2=1);
run;

data entrepr.cas3;
set entrepr.prep;
if (receveur=3 or donneur3=1);
run;

data entrepr.cas1;
set entrepr.prep;
if (receveur=1 or donneur1=1);
run;

```

```

*Traitement du cas 1;
*=====;;

proc sort data=entrepr.cas1;
by num;
run;

data tabdon1;
set entspr.cas1;
rename num=dnum ENFMALdd=dENFMALdd congduurd=dcongduurd gardentrd=dgardentrd
icreched=dicreched prestaidd=dprestaidd prestaad=dprestaad prestagd=dprestagd
amenad=damenad
amenbd=damenbd amencd=damencd amendd=damendd dombd=ddombd partaccd=dpartaccd
partembd=dpartembd rttddd=drttddd promocd=dpromocd congdd=dcongdd
prestaidd=dprestaidd
fmutud=dfmutud mutud=dmutud
classe=dclasse tgroupe=dtgroupe nes16d=dnes16d rrrr=drrrr;
if donneur1=1;
drop donneur1 donneur2 donneur3 numident poids;
run;

data tabrec1;
set entspr.cas1;
if receveur=1;
drop receveur donneur1 donneur2 donneur3;
run;

data tabrec1;
set tabrec1;
nb=_N_;
run;

%macro gene (nbrec= , tabrec= , tabdon= );
%do i=1 %to &nbrec;

data receveur;
set &tabrec.;
if nb=&i.;
run;

proc sql;
create table fusion as select * from receveur , &tabdon.;
quit;

* produit cartésien de la table du receveur et de la table des donneurs;

```

```

*Calcul de la distance;
*sur les 4 var auxiliaires et les 22 variables à imputer;
data result&i.;
set fusion end=fin;
if classe=dclasse then d1=0; else d1=1;
if tgroupe=dtgroupe then d2=0; else d2=1;
if nes16d=dnes16d then d3=0; else d3=1;
if rrrr=drrrr then d4=0; else d4=1;
if (ENFMALdd ne 9 and ENFMALdd=dENFMALdd) or ENFMALdd=9 then ds1=0; else ds1=1;
if (congurd ne 9 and congurd=dcongurd) or congurd=9 then ds2=0; else ds2=1;
if (gardentrd ne 9 and gardentrd=dgardentrd) or gardentrd=9 then ds3=0; else
ds3=1;
if (icreched ne 9 and icreched=dicreched) or icreched=9 then ds4=0; else ds4=1;
if (prestaed ne 9 and prestaed=dprestaed) or prestaed=9 then ds5=0; else ds5=1;
if (prestaad ne 9 and prestaad=dprestaad) or prestaad=9 then ds6=0; else ds6=1;
if (prestadd ne 9 and prestadd=dprestadd) or prestadd=9 then ds7=0; else ds7=1;
if (prestaed ne 9 and prestaed=dprestaed) or prestaed=9 then ds8=0; else ds8=1;
if (prestagd ne 9 and prestagd=dprestagd) or prestagd=9 then ds9=0; else ds9=1;
if (amenad ne 9 and amenad=damenad) or amenad=9 then ds10=0; else ds10=1;
if (amenbd ne 9 and amenbd=damenbd) or amenbd=9 then ds11=0; else ds11=1;
if (amencd ne 9 and amencd=damencd) or amencd=9 then ds12=0; else ds12=1;
if (amendd ne 9 and amendd=damendd) or amendd=9 then ds13=0; else ds13=1;
if (dombd ne 9 and dombd=ddombd) or dombd=9 then ds14=0; else ds14=1;
if (partaccd ne 9 and partaccd=dpartaccd) or partaccd=9 then ds15=0; else
ds15=1;
if (partembd ne 9 and partembd=dpartembd) or partembd=9 then ds16=0; else
ds16=1;
if (rttddd ne 9 and rttddd=drttddd) or rttddd=9 then ds17=0; else ds17=1;
if (promocd ne 9 and promocd=dpromocd) or promocd=9 then ds18=0; else ds18=1;
if (congnd ne 9 and congnd=dcongnd) or congnd=9 then ds19=0; else ds19=1;
if (prestaed ne 9 and prestaed=dprestaed) or prestaed=9 then ds20=0; else
ds20=1;
if (fmutud ne 9 and fmutud=dfmutud) or fmutud=9 then ds21=0; else ds21=1;
if (mutud ne 9 and mutud=dmutud) or mutud=9 then ds22=0; else ds22=1;

* Calcul de la distance totale pondérée et normalisée;
dist=( 3.924*d1 + 7.1456*d2 + 6.7822*d3 + 2.0441*d4
+ 5.5375*ds1 + 2.5136*ds2 + 3.5398*ds3 + 2.2278*ds4 + 5.0041*ds5 + 2.9040*ds6
+ 5.8510*ds7 + 4.3356*ds8 + 5.7706*ds9 + 3.1260*ds10 + 3.1354*ds11 +
2.7588*ds12
+ 2.7652*ds13 + 2.4427*ds14 + 3.2440*ds15 + 2.4472*ds16 + 3.5968*ds17 +
3.0060*ds18 + 4.5136*ds19 + 4.5398*ds20 + 4.5552*ds21 + 3.9129*ds22)/101.6135
;

*calcul du nbr de caractéristiques communes entre le donneur et le receveur sur
les variables auxiliaires;
nbcc=d1+d2+d3+d4;

*calcul du nbr de caractéristiques communes entre le donneur et le receveur sur
les variables à imputer;
nbcci=ds1 + ds2 + ds3 + ds4
+ ds5 + ds6 + ds7 + ds8 + ds9 + ds10 + ds11 + ds12 + ds13 + ds14 + ds15
+ ds16 + ds17 + ds18 + ds19 + ds20 + ds21 + ds22;

*initialisation des valeurs;
retain min 2 prox 5 don 0 imp 0 ;

```

```

*si le donneur potentiel correspondant à l'observation courante a une distance
plus petite
que celui qui était sélectionné jusqu'à présent,
alors il est sélectionné comme donneur à la place du précédent sinon le donneur
précédent est conservé;

if dist lt min then do;

distmin=dist;
distance=nbcc;
donneur=dnum;
distimp=nbcci;

min=dist;
prox=nbcc;
don=dnum;
imp=nbcci;
end;

else do;
distmin=min;
distance=prox;
donneur=don;
distimp=imp;
end;

if fin then output result&i.;
run;

data result1;
update result1 result&i.;
by num;
run;
%end;
%mend;

%gene(nbrec=377, tabrec=tabrec1 , tabdon=tabdon1);
option nonotes;

data result1;
set result1;
dnum=donneur;
drop d1 d2 d3 d4 ds1 ds2 ds3 ds4
ds5 ds6 ds7 ds8 ds9 ds10 ds11 ds12 ds13 ds14 ds15
ds16 ds17 ds18 ds19 ds20 ds21 ds22;
run;

proc sort data=result1;
by dnum;
run;

data entrepr.resultat1;
merge result1 tabdon1;
by dnum;
if numident ne .;
run;

```

```

data entrepr.resultat1;
set entrepr.resultat1;
if ENFMALdd=9 then ENFMALdd=dENFMALdd; else ENFMALdd=ENFMALdd;
if congdudd=9 then congdudd=dcongdudd; else congdudd=congdudd;
if gardentrd=9 then gardentrd=dgardentrd; else gardentrd=gardentrd;
if icreched=9 then icreched=dicreched; else icreched=icreched;
if prestaidd=9 then prestaidd=dprestaidd; else prestaidd=prestaidd;
if prestaad=9 then prestaad=dprestaad; else prestaad=prestaad;
if prestadd=9 then prestadd=dprestadd; else prestadd=prestadd;
if prestaed=9 then prestaed=dprestaed; else prestaed=prestaed;
if prestagd=9 then prestagd=dprestagd; else prestagd=prestagd;
if amenad=9 then amenad=damenad; else amenad=amenad;
if amenbd=9 then amenbd=damenbd; else amenbd=amenbd;
if amencd=9 then amencd=damencd; else amencd=amencd;
if amendd=9 then amendd=damendd; else amendd=amendd;
if dombd=9 then dombd=ddombd; else dombd=dombd;
if partaccd=9 then partaccd=dpartaccd; else partaccd=partaccd;
if partembd=9 then partembd=dpartembd; else partembd=partembd;
if rttddd=9 then rttddd=drttddd; else rttddd=rttddd;
if promocd=9 then promocd=dpromocd; else promocd=promocd;
if congdd=9 then congdd=dcongdd; else congdd=congdd;
if prestahd=9 then prestahd=dprestahd; else prestahd=prestahd;
if fmutud=9 then fmutud=dfmutud; else fmutud=fmutud;
if mutud=9 then mutud=dmutud; else mutud=mutud;
run;

*=====;

*Traitement du cas 2;
*=====;

proc sort data=entrepr.cas2;
by num;
run;

data tabdon2;
set entrepr.cas2;
rename num=dnum ENFMALdd=dENFMALdd congdudd=dcongdudd gardentrd=dgardentrd
icreched=dicreched prestaidd=dprestaidd
prestaad=dprestaad prestadd=dprestadd prestaed=dprestaed prestagd=dprestagd
amenad=damenad
amenbd=damenbd amencd=damencd amendd=damendd dombd=ddombd partaccd=dpartaccd
partembd=dpartembd rttddd=drttddd promocd=dpromocd congdd=dcongdd
prestahd=dprestahd
fmutud=dfmutud mutud=dmutud
classe=dclasse tgroupe=dtgroupe nes16d=dnes16d rrrr=drrrr;
if donneur2=1;
drop donneur1 donneur2 donneur3 numident poids;
run;

data tabrec2;
set entrepr.cas2;
if receveur=2;
drop receveur donneur1 donneur2 donneur3;
run;

data tabrec2;
set tabrec2;
nb=_N_;
run;

```

```

%gene(nbrec=492, tabrec=tabrec2 , tabdon=tabdon2);
option nonotes;

data result1;
set result1;
dnum=donneur;
drop d1 d2 d3 d4 ds1 ds2 ds3 ds4
ds5 ds6 ds7 ds8 ds9 ds10 ds11 ds12 ds13 ds14 ds15
ds16 ds17 ds18 ds19 ds20 ds21 ds22;
run;

proc sort data=result1;
by dnum;
run;

data entrepr.resultat2;
merge result1 tabdon2;
by dnum;
if numident ne .;
run;

data entrepr.resultat2;
set entrepr.resultat2;
if ENFMALdd=9 then ENFMALdd=dENFMALdd; else ENFMALdd=ENFMALdd;
if congdudd=9 then congdudd=dcongdudd; else congdudd=congdudd;
if gardentrd=9 then gardentrd=dgardentrd; else gardentrd=gardentrd;
if icreched=9 then icreched=dicreched; else icreched=icreched;
if prestaidd=9 then prestaidd=dprestaidd; else prestaidd=prestaidd;
if prestaadd=9 then prestaadd=dprestaadd; else prestaadd=prestaadd;
if prestaedd=9 then prestaedd=dprestaedd; else prestaedd=prestaedd;
if prestagdd=9 then prestagdd=dprestagdd; else prestagdd=prestagdd;
if amenadd=9 then amenadd=damenadd; else amenadd=amenadd;
if amenbdd=9 then amenbdd=damenbdd; else amenbdd=amenbdd;
if amenccd=9 then amenccd=damenccd; else amenccd=amenccd;
if amenddd=9 then amenddd=damenddd; else amenddd=amenddd;
if dombdd=9 then dombdd=ddombdd; else dombdd=dombdd;
if partacdd=9 then partacdd=dpartacdd; else partacdd=partacdd;
if partembdd=9 then partembdd=dpartembdd; else partembdd=partembdd;
if rttddd=9 then rttddd=drttddd; else rttddd=rttddd;
if promocdd=9 then promocdd=dpromocdd; else promocdd=promocdd;
if congdd=9 then congdd=dcongdd; else congdd=congdd;
if prestahdd=9 then prestahdd=dprestahdd; else prestahdd=prestahdd;
if fmutudd=9 then fmutudd=dfmutudd; else fmutudd=fmutudd;
if mutudd=9 then mutudd=dmutudd; else mutudd=mutudd;
run;

*=====;

```

```

*Traitement du cas3;
*=====;

proc sort data=entrepr.cas3;
by num;
run;

data tabdon3;
set entspr.cas3;
rename num=dnum ENFMALdd=dENFMALdd congduurd=dcongduurd gardentrd=dgardentrd
icreched=dicreched prestaidd=dprestaidd
prestaad=dprestaad prestadd=dprestadd prestaed=dprestaed prestagdd=dprestagdd
amenad=damenad
amenbd=damenbd amencd=damencd amendd=damendd dombd=ddombd partaccdd=dpartaccdd
partembd=dpartembd rttddd=drttddd promocdd=dpromocdd congdd=dcongdd
prestaidd=dprestaidd
fmutud=dfmutud mutud=dmutud
classe=dclasse tgrouped=dtgrouped nes16d=dnes16d rrrr=drrrr;
if donneur3=1;
drop donneur1 donneur2 donneur3 numident poids;
run;

data tabrec3;
set entspr.cas3;
if receveur=3;
drop receveur donneur1 donneur2 donneur3;
run;

data tabrec3;
set tabrec3;
nb=_N_;
run;

%gene(nbrec=377, tabrec=tabrec3 , tabdon=tabdon3);
option nonotes;

data result1;
set result1;
dnum=donneur;
drop d1 d2 d3 d4 ds1 ds2 ds3 ds4
ds5 ds6 ds7 ds8 ds9 ds10 ds11 ds12 ds13 ds14 ds15
ds16 ds17 ds18 ds19 ds20 ds21 ds22;
run;

proc sort data=result1;
by dnum;
run;

data entspr.resultat3;
merge result1 tabdon3;
by dnum;
if numident ne .;
run;

```

```

data entrepr.resultat3;
set entrepr.resultat3;
if ENFMALdd=9 then ENFMALdd=dENFMALdd; else ENFMALdd=ENFMALdd;
if congdudd=9 then congdudd=dcongdudd; else congdudd=congdudd;
if gardentrd=9 then gardentrd=dgardentrd; else gardentrd=gardentrd;
if icreched=9 then icreched=dicreched; else icreched=icreched;
if prestaidd=9 then prestaidd=dprestaidd; else prestaidd=prestaidd;
if prestaad=9 then prestaad=dprestaad; else prestaad=prestaad;
if prestadd=9 then prestadd=dprestadd; else prestadd=prestadd;
if prestaed=9 then prestaed=dprestaed; else prestaed=prestaed;
if prestagd=9 then prestagd=dprestagd; else prestagd=prestagd;
if amenad=9 then amenad=damenad; else amenad=amenad;
if amenbd=9 then amenbd=damenbd; else amenbd=amenbd;
if amencd=9 then amencd=damencd; else amencd=amencd;
if amendd=9 then amendd=damendd; else amendd=amendd;
if dombd=9 then dombd=ddombd; else dombd=dombd;
if partaccd=9 then partaccd=dpartaccd; else partaccd=partaccd;
if partembd=9 then partembd=dpartembd; else partembd=partembd;
if rttddd=9 then rttddd=drttddd; else rttddd=rttddd;
if promocd=9 then promocd=dpromocd; else promocd=promocd;
if congdd=9 then congdd=dcongdd; else congdd=congdd;
if prestahd=9 then prestahd=dprestahd; else prestahd=prestahd;
if fmutud=9 then fmutud=dfmutud; else fmutud=fmutud;
if mutud=9 then mutud=dmutud; else mutud=mutud;
run;

data all;
set entrepr.prep;
if receveur=0;
run;

proc sort data=entrepr.resultat1;
by numident;
run;

proc sort data=entrepr.resultat2;
by numident;
run;

proc sort data=entrepr.resultat3;
by numident;
run;

data entrepr.impute;
merge all entrepr.resultat1 entrepr.resultat2 entrepr.resultat3;
by numident;
run;

*=====;
*=====;

```

```

=====;
*preparation des formats;
=====;

proc format;

value ENFMALdd 1="0j_enfmal"
               2="1_3j_enfmal"
               3="4_6j_enfmal"
               4=">6j_enfmal"
               ;

value congdudd 1="pas_maternsup"
               2="maternsup"
               3="maternsup"
               ;

value gardentrd 0="pas_enf_entr"
                1="enf_entr"
                ;

value icreched 0="pas_icreche"
               1="icreche"
               ;

value prestaidd 0="pas_colo"
                1="colo"
                ;

value prestaad 0="pas_prime_naiss"
               1="aprice_naiss"
               ;

value prestaadd 0="pas_aide_fraisgarde"
                1="aide_fraisgarde"
                ;

value prestaed 0="pas_aideetude"
               1="aideetude"
               ;

value prestagd 0="pas_presthandic"
               1="apresthandi"
               ;

value congsalbd 0="pas_compl_congepatmat"
                1="compl_congepatmat"
                ;

value amenad 0="pas_h_rentree "
             1="h_rentree"
             ;

value amenbd 0="pas_h_garde "
             1="h_garde"
             ;

value amencd 0="pas_h_enfmal "
             1="h_enfmal"
             ;

```

```

value amendd 0="pas_h_trajet "
              1="h_trajet"
              ;

value dombd 0="pas_L_domicil"
            1="L_domicil_poss";

value partembd 0="pas_postespartiel"
              1="apostespartiel";

value partaccd 1="0pas_tpspartiel"
               2="1parfois_tpspartiel"
               3="2tjrs_tpspartiel"
               ;

value rttdd 1="0pas_RTT"
            2="1RTT_contraint"
            3="2RTT_accord"
            4="3RTT_libre";

value promocd 1="jamais_avancem"
              2="parfois_avancem"
              3="svt_avancem"
              4="tjrs_avancem"
              ;

value prestahd 0="pas_chèque_vac"
               1="chèque_vac"
               ;

value fmutud
1="cpasFinMutu"
2="bcadreFinMutu"
3="atousFinMutu";

value mutud
0="bpasServMutu"
1="aServMutu";

```

```

data entrepr.affectfini;
set entrepr.impute;
cENFMALdd=put (ENFMALdd,ENFMALdd.);
ccongdurd =put (congduard,congduard.);
cgardentrd=put (gardentrd,gardentrd.);
cicreched =put (icreched,icreched.);
cprestaaid =put (prestaaid,prestaaid.);
cprestaad =put (prestaad,prestaad.);
cprestadd =put (prestadd,prestadd.);
cprestaed =put (prestaed,prestaed.);
cprestagd =put (prestagd,prestagd.);
ccongd =put (congduard,congduard.);
camenad =put (amenad,amenad.);
camenbd =put (amenbd,amenbd.);
camencd =put (amencd,amencd.);
camendd =put (amendd,amendd.);
cdombd =put (dombd,dombd.);
cpartembd=put (partembd,partembd.);
cpartaccd =put (partaccd,partaccd.);
crttdd =put (rttddd,rttdd.);
cpromocd =put (promocd,promocd.);
cprestahd =put (prestahd,prestahd.);
cfmutud=put (fmutud,fmutud.);
cmutud=put (mutud,mutud.);
run;

data entrepr.affectfini;
set entrepr.affectfini;
drop ENFMALdd congduard gardentrd
icreched prestaaid prestaad prestadd prestaed prestagd
congduard amenad amenbd amencd amendd
dombd partembd partaccd rttddd promocd prestahd fmutud mutud;
run;

data entrepr.affectfini;
set entrepr.affectfini;
rename cENFMALdd=ENFMALdd;
rename ccongduard =congduard;
rename cgardentrd=gardentrd;
rename cicreched=icreched;
rename cprestaaid=prestaaid;
rename cprestaad=prestaad;
rename cprestadd=prestadd;
rename cprestaed=prestaed;
rename cprestagd=prestagd;
rename ccongd=congduard;
rename camenad=amenad;
rename camenbd=amenbd;
rename camencd=amencd;
rename camendd=amendd;
rename cdombd=dombd;
rename cpartembd=partembd;
rename cpartaccd=partaccd;
rename crttd=rttd;
rename cpromocd=promocd;
rename cprestahd=prestahd;
rename cfmutud=fmutud;
rename cmutud=mutud;
run;

```

```
data entrepr.affectfini;  
set entrepr.affectfini;  
keep ENFMALdd congduurd gardedtt icreched prestaidd  
prestaad prestadd prestaed prestagd  
congdamenamdamenbd amencd amendd  
dombd partembd partaccd rttd d promocd prestahd fmutud mutud numident poidd ;  
run;
```

Annexe C

Annexe ACM

C.1 Histogramme des valeurs-propres et coordonnées des projections des modalités des variables actives

SOMME DES VALEURS PROPRES 1.3810

HISTOGRAMME DES 29 PREMIERES VALEURS PROPRES

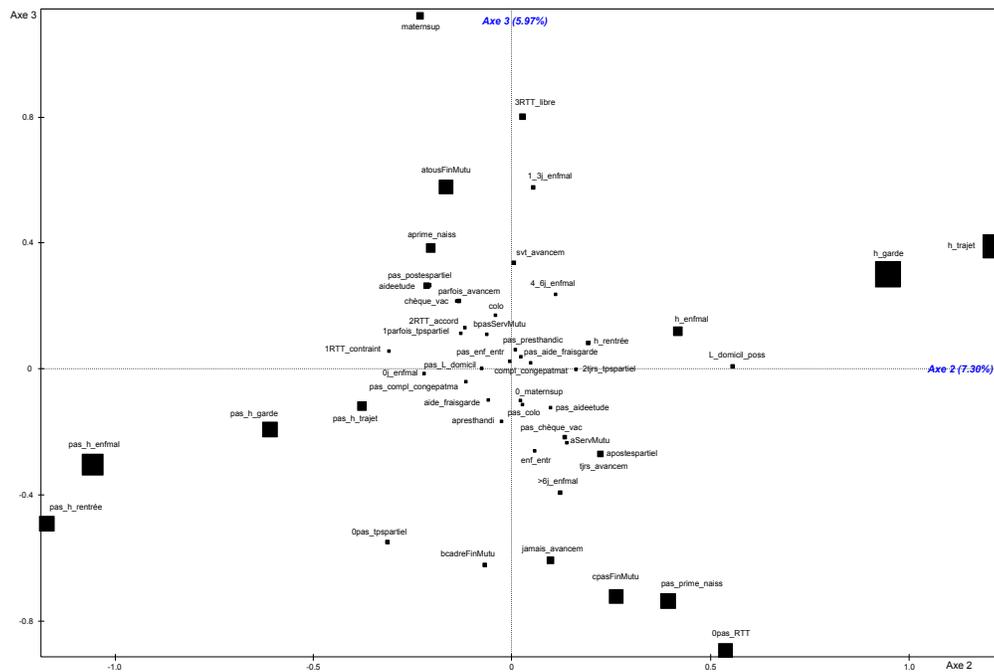
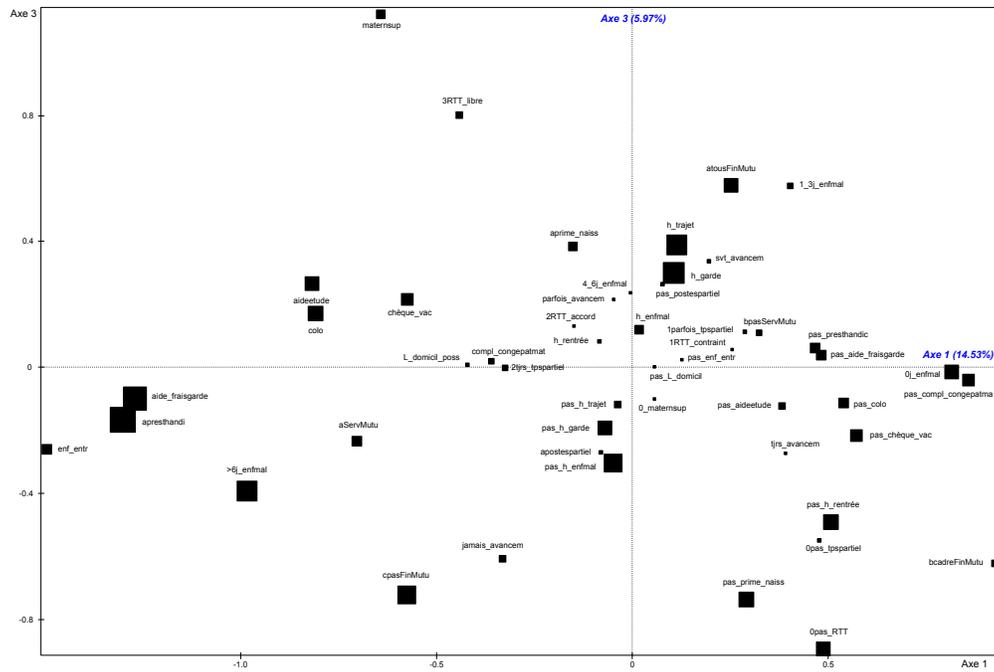
NUMERO	VALEUR PROPRE	POURCENTAGE	POURCENTAGE CUMULE	
1	0.2006	14.53	14.53	*****
2	0.1008	7.30	21.83	*****
3	0.0824	5.97	27.80	*****
4	0.0590	4.27	32.07	*****
5	0.0560	4.05	36.12	*****
6	0.0525	3.80	39.92	*****
7	0.0501	3.63	43.55	*****
8	0.0496	3.59	47.14	*****
9	0.0480	3.47	50.62	*****
10	0.0472	3.42	54.03	*****
11	0.0463	3.35	57.38	*****
12	0.0460	3.33	60.72	*****
13	0.0432	3.13	63.84	*****
14	0.0428	3.10	66.94	*****
15	0.0416	3.01	69.95	*****
16	0.0408	2.95	72.91	*****
17	0.0386	2.80	75.70	*****
18	0.0382	2.76	78.47	*****
19	0.0353	2.55	81.02	*****
20	0.0335	2.43	83.45	*****
21	0.0329	2.38	85.83	*****
22	0.0319	2.31	88.14	*****
23	0.0300	2.17	90.31	*****
24	0.0273	1.98	92.29	*****
25	0.0261	1.89	94.18	*****
26	0.0237	1.71	95.90	*****
27	0.0204	1.48	97.37	*****
28	0.0196	1.42	98.79	*****
29	0.0167	1.21	100.00	*****

FIG. C.1 – Histogramme des valeurs propres pour l'ACM

MODALITES			COORDONNEES					CONTRIBUTIONS					COSINUS CARRES					
IDEN	LIBELLE	P.REL	DISTO	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
2 . ENFMALdd																		
AA_1	- 0j_enfmal	1.52	2.12	0.82	-0.22	-0.01	0.48	0.44	5.1	0.7	0.0	6.0	5.3	0.31	0.02	0.00	0.11	0.09
AA_2	- 1_3j_enfmal	0.64	6.49	0.40	0.05	0.58	-0.56	-0.21	0.5	0.0	2.6	3.4	0.5	0.03	0.00	0.05	0.05	0.01
AA_3	- 4_6j_enfmal	1.08	3.43	0.00	0.11	0.24	-0.39	-0.83	0.0	0.1	0.7	2.8	13.1	0.00	0.00	0.02	0.04	0.20
AA_4	- >6j_enfmal	1.53	2.12	-0.98	0.12	-0.39	0.02	0.23	7.3	0.2	2.8	0.0	1.4	0.45	0.01	0.07	0.00	0.02
CONTRIBUTION CUMULEE =									12.9	1.1	6.2	12.2	20.3					
3 . congurd																		
AB_1	- 0_maternsup	4.37	0.09	0.06	0.02	-0.10	0.09	-0.05	0.1	0.0	0.5	0.6	0.2	0.04	0.00	0.11	0.09	0.03
AB_2	- maternsup	0.39	11.22	-0.64	-0.23	1.12	-1.02	0.60	0.8	0.2	6.0	6.8	2.5	0.04	0.00	0.11	0.09	0.03
CONTRIBUTION CUMULEE =									0.9	0.2	6.5	7.5	2.8					
4 . gardedtt																		
AC_1	- enf_entr	0.37	11.75	-1.49	0.06	-0.26	0.63	0.09	4.2	0.0	0.3	2.5	0.1	0.19	0.00	0.01	0.03	0.00
AC_2	- pas_enf_entr	4.39	0.09	0.13	0.00	0.02	-0.05	-0.01	0.4	0.0	0.0	0.2	0.0	0.19	0.00	0.01	0.03	0.00
CONTRIBUTION CUMULEE =									4.5	0.0	0.3	2.7	0.1					
6 . prestaidd																		
AE_1	- colo	1.91	1.49	-0.81	-0.04	0.17	0.08	0.17	6.2	0.0	0.7	0.2	1.0	0.44	0.00	0.02	0.00	0.02
AE_2	- pas_colo	2.85	0.67	0.54	0.03	-0.11	-0.05	-0.12	4.2	0.0	0.4	0.1	0.7	0.44	0.00	0.02	0.00	0.02
CONTRIBUTION CUMULEE =									10.4	0.1	1.1	0.3	1.7					
7 . prestaad																		
AF_1	- aprime_naiss	3.14	0.52	-0.15	-0.20	0.38	0.16	-0.08	0.4	1.3	5.6	1.4	0.3	0.04	0.08	0.28	0.05	0.01
AF_2	- pas_prime_naiss	1.63	1.93	0.29	0.39	-0.74	-0.32	0.15	0.7	2.5	10.7	2.8	0.7	0.04	0.08	0.28	0.05	0.01
CONTRIBUTION CUMULEE =									1.0	3.8	16.3	4.2	1.0					
8 . prestadd																		
AG_1	- aide_fraisgarde	1.32	2.62	-1.27	-0.06	-0.10	0.18	0.09	10.6	0.0	0.2	0.7	0.2	0.61	0.00	0.00	0.01	0.00
AG_2	- pas_aide_fraisgarde	3.45	0.38	0.48	0.02	0.04	-0.07	-0.04	4.0	0.0	0.1	0.3	0.1	0.61	0.00	0.00	0.01	0.00
CONTRIBUTION CUMULEE =									14.6	0.1	0.2	0.9	0.3					
9 . prestaed																		
AH_1	- aideetude	1.52	2.12	-0.82	-0.21	0.27	0.43	-0.11	5.1	0.7	1.3	4.8	0.3	0.31	0.02	0.03	0.09	0.01
AH_2	- pas_aideetude	3.24	0.47	0.38	0.10	-0.12	-0.20	0.05	2.4	0.3	0.6	2.3	0.1	0.31	0.02	0.03	0.09	0.01
CONTRIBUTION CUMULEE =									7.4	1.0	1.9	7.1	0.5					
10 . prestagd																		
AI_1	- aprestandi	1.26	2.77	-1.30	-0.02	-0.17	0.31	0.05	10.6	0.0	0.4	2.1	0.1	0.61	0.00	0.01	0.04	0.00
AI_2	- pas_prestandi	3.50	0.36	0.47	0.01	0.06	-0.11	-0.02	3.8	0.0	0.2	0.8	0.0	0.61	0.00	0.01	0.04	0.00
CONTRIBUTION CUMULEE =									14.4	0.0	0.6	2.8	0.1					
11 . congdd																		
AJ_1	- compl_congepatmat	3.36	0.42	-0.36	0.05	0.02	-0.19	0.03	2.1	0.1	0.0	2.1	0.1	0.31	0.01	0.00	0.09	0.00
AJ_2	- pas_compl_congepatma	1.40	2.40	0.86	-0.12	-0.04	0.46	-0.08	5.2	0.2	0.0	5.1	0.1	0.31	0.01	0.00	0.09	0.00
CONTRIBUTION CUMULEE =									7.3	0.3	0.0	7.2	0.2					
12 . amenad																		
AK_1	- h_rentree	4.09	0.16	-0.08	0.19	0.08	0.11	-0.04	0.1	1.5	0.3	0.8	0.1	0.04	0.23	0.04	0.07	0.01
AK_2	- pas_h_rentree	0.67	6.09	0.51	-1.17	-0.49	-0.66	0.26	0.9	9.2	2.0	4.9	0.8	0.04	0.23	0.04	0.07	0.01
CONTRIBUTION CUMULEE =									1.0	10.7	2.3	5.7	0.9					
13 . amenbd																		
AL_1	- h_garde	1.86	1.56	0.11	0.95	0.30	0.13	0.15	0.1	16.6	2.0	0.6	0.8	0.01	0.58	0.06	0.01	0.02
AL_2	- pas_h_garde	2.90	0.64	-0.07	-0.61	-0.19	-0.09	-0.10	0.1	10.7	1.3	0.4	0.5	0.01	0.58	0.06	0.01	0.02
CONTRIBUTION CUMULEE =									0.2	27.3	3.3	0.9	1.3					
14 . amencc																		
AM_1	- h_enfmal	3.41	0.40	0.02	0.42	0.12	0.08	-0.05	0.0	5.9	0.6	0.4	0.2	0.00	0.44	0.04	0.02	0.01
AM_2	- pas_h_enfmal	1.35	2.53	-0.05	-1.06	-0.30	-0.20	0.14	0.0	14.9	1.5	1.0	0.5	0.00	0.44	0.04	0.02	0.01
CONTRIBUTION CUMULEE =									0.0	20.8	2.1	1.3	0.6					
15 . amendd																		
AN_1	- h_trajet	1.13	3.23	0.11	1.22	0.39	0.02	0.29	0.1	16.5	2.1	0.0	1.7	0.00	0.46	0.05	0.00	0.03
AN_2	- pas_h_trajet	3.64	0.31	-0.04	-0.38	-0.12	-0.01	-0.09	0.0	5.1	0.6	0.0	0.5	0.00	0.46	0.05	0.00	0.03
CONTRIBUTION CUMULEE =									0.1	21.7	2.7	0.0	2.2					
16 . dombd																		
AO_1	- L_domicil_oss	0.57	7.39	-0.42	0.56	0.01	-0.75	0.58	0.5	1.7	0.0	5.4	3.4	0.02	0.04	0.00	0.08	0.05
AO_2	- pas_L_domicil	4.19	0.14	0.06	-0.08	0.00	0.10	-0.08	0.1	0.2	0.0	0.7	0.5	0.02	0.04	0.00	0.08	0.05
CONTRIBUTION CUMULEE =									0.6	2.0	0.0	6.1	3.9					
17 . partembd																		
AP_1	- apostespartiel	2.35	1.03	-0.08	0.22	-0.27	-0.09	-0.64	0.1	1.1	2.1	0.3	17.4	0.01	0.05	0.07	0.01	0.40
AP_2	- pas_postespartiel	2.42	0.97	0.08	-0.22	0.26	0.08	0.62	0.1	1.1	2.0	0.3	16.9	0.01	0.05	0.07	0.01	0.40
CONTRIBUTION CUMULEE =									0.1	2.3	4.2	0.6	34.3					
18 . partaccd																		
AQ_1	- Opas_tpspartiel	0.40	11.00	0.48	-0.31	-0.55	0.41	1.64	0.5	0.4	1.5	1.1	19.1	0.02	0.01	0.03	0.02	0.24
AQ_2	- lparfois_tpspartiel	2.01	1.37	0.29	-0.13	0.11	0.32	-0.15	0.8	0.3	0.3	3.5	0.8	0.06	0.01	0.01	0.07	0.02
AQ_3	- 2tjrs_tpspartiel	2.36	1.02	-0.32	0.16	0.00	-0.34	-0.15	1.2	0.6	0.0	4.6	0.9	0.10	0.03	0.00	0.11	0.02
CONTRIBUTION CUMULEE =									2.5	1.3	1.8	9.2	20.8					
19 . rttd																		
AR_1	- Opas_RTT	0.93	4.14	0.49	0.54	-0.90	-0.07	0.29	1.1	2.7	9.0	0.1	1.4	0.06	0.07	0.19	0.00	0.02
AR_2	- 1RTT_contraint	0.70	5.80	0.25	-0.31	0.05	0.90	-0.20	0.2	0.7	0.0	9.5	0.5	0.01	0.02	0.00	0.14	0.01
AR_3	- 2RTT_accord	2.56	0.86	-0.15	-0.12	0.13	-0.01	-0.19	0.3	0.3	0.5	0.0	1.7	0.03	0.02	0.02	0.00	0.04
AR_4	- 3RTT_libre	0.58	7.26	-0.44	0.03	0.80	-0.92	0.64	0.6	0.0	4.5	8.3	4.2	0.03	0.00	0.09	0.12	0.06
CONTRIBUTION CUMULEE =									2.2	3.7	14.1	17.9	7.8					
20 . promood																		
AS_1	- jamais_avancem	1.20	2.97	-0.33	0.10	-0.61	-0.36	-0.08	0.7	0.1	5.4	2.7	0.1	0.04	0.00	0.12	0.04	0.00
AS_2	- parfois_avancem	1.62	1.94	-0.05	-0.14	0.22	-0.20	0.12	0.0	0.3	0.9	1.1	0.4	0.00	0.01	0.02	0.02	0.01
AS_3	- svt_avancem	1.49	2.19	0.20	0.00	0.34	0.24	-0.06	0.3	0.0	2.1	1.4	0.1	0.02	0.00	0.05	0.03	0.00
AS_4	- tjrs_avancem	0.45	9.61	0.39	0.22	-0.27	0.90	-0.04	0.3	0.2	0.4	6.1	0.0	0.02	0.01	0.01	0.08	0.00
CONTRIBUTION CUMULEE =									1.3	0.6	8.8	11.3	0.7					
21 . prestahd																		
AT_1	- cheque_vac	2.38	1.00	-0.57	-0.13	0.22	-0.04	-0.07	3.9	0.4	1.4	0.1	0.2	0.33	0.02	0.05	0.00	0.00
AT_2	- pas_cheque_vac	2.38	1.00	0.57	0.13	-0.22	0.04	0.07	3.9	0.4	1.4	0.1	0.2	0.33	0.02	0.05	0.00	0.00
CONTRIBUTION CUMULEE =									7.8	0.9	2.7	0.1	0.4					
22 . fmutud																		
AU_1	- atousFinMutu	2.62	0.82	0.25	-0.17	0.58	-0.05	-0.03	0.8	0.7	10.6	0.1	0.0	0.08	0.03	0.41	0.00	0.00
AU_2	- bcadreFinMutu	0.38	11.57	0.93	-0.07	-0.62	0.29	0.01	1.6	0.0	1.8	0.5	0.0	0.07	0.00	0.03	0.01	0.00
AU_3	- cpasFinMutu	1.76	1.70	-0.58	0.26	-0.72	0.02	0.04	2.9	1.2	11.2	0.0	0.0	0.19	0.04	0.31	0.00	0.00
CONTRIBUTION CUMULEE =									5.4	1.9	23.6	0.7	0.1					
23 . mutud																		
AV_1	- aServMutu	1.51	2.16	-0.70	0.14	-0.23	0.17	0.05	3.7	0.3	1.0	0.7	0.1	0.23	0.01	0.03	0.01	0.00
AV_2	- bpasServMutu	3.26	0.46	0.32	-0.06	0.11	-0.08	-0.02	1.7	0.1	0.5	0.3	0.0	0.23	0.01	0.03	0.01	0.00
CONTRIBUTION CUMULEE =									5.4	0.4	1.5	1.1	0.1					

FIG. C.2 – Coordonnées des projections pour l'ACM

C.2 Plans 1-3 et 2-3



C.3 ACM en utilisant les pondérations

SOMME DES VALEURS PROPRES 1.3810

HISTOGRAMME DES 29 PREMIERES VALEURS PROPRES

NUMERO	VALEUR PROPRE	POURCENTAGE	POURCENTAGE CUMULE	
1	0.1881	13.62	13.62	*****
2	0.1015	7.35	20.96	*****
3	0.0808	5.85	26.81	*****
4	0.0643	4.66	31.47	*****
5	0.0561	4.06	35.53	*****
6	0.0550	3.98	39.51	*****
7	0.0537	3.89	43.40	*****
8	0.0518	3.75	47.16	*****
9	0.0498	3.61	50.76	*****
10	0.0494	3.58	54.34	*****
11	0.0477	3.46	57.80	*****
12	0.0462	3.35	61.15	*****
13	0.0448	3.24	64.39	*****
14	0.0431	3.12	67.51	*****
15	0.0406	2.94	70.45	*****
16	0.0396	2.87	73.32	*****
17	0.0387	2.80	76.12	*****
18	0.0379	2.74	78.86	*****
19	0.0356	2.58	81.44	*****
20	0.0323	2.34	83.78	*****
21	0.0319	2.31	86.09	*****
22	0.0298	2.16	88.25	*****
23	0.0289	2.10	90.34	*****
24	0.0264	1.91	92.26	*****
25	0.0257	1.86	94.11	*****
26	0.0234	1.69	95.81	*****
27	0.0214	1.55	97.36	*****
28	0.0196	1.42	98.78	*****
29	0.0169	1.22	100.00	*****

FIG. C.3 – Histogramme des valeurs propres pour l'ACM pondérée

MODALITES			COORDONNEES					CONTRIBUTIONS					COSINUS CARRES					
IDEN	LIBELLE	P.REL	DISTO	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
2 . ENFMALdd																		
AA_1	- 0j_enfmal	2.10	1.27	0.65	0.24	0.12	0.20	-0.04	4.7	1.2	0.4	1.3	0.1	0.33	0.04	0.01	0.03	0.00
AA_2	- 1_3j_enfmal	0.60	6.97	0.30	-0.27	0.48	-0.18	0.36	0.3	0.4	1.7	0.3	1.4	0.01	0.01	0.03	0.00	0.02
AA_3	- 4_6j_enfmal	0.93	4.13	-0.28	-0.19	-0.16	-0.34	-0.38	0.4	0.3	0.3	1.6	2.4	0.02	0.01	0.01	0.03	0.03
AA_4	- >6j_enfmal	1.14	3.18	-1.13	-0.14	-0.35	0.00	0.19	7.7	0.2	1.7	0.0	0.8	0.40	0.01	0.04	0.00	0.01
CONTRIBUTION CUMULEE =									13.1	2.1	4.1	3.2	4.5					
3 . congurd																		
AB_1	- maternsup	0.28	16.26	-0.76	-0.13	1.15	-0.83	1.68	0.8	0.0	4.5	2.9	13.9	0.04	0.00	0.08	0.04	0.17
AB_2	- pas_maternsup	4.49	0.06	0.05	0.01	-0.07	0.05	-0.10	0.1	0.0	0.3	0.2	0.9	0.04	0.00	0.08	0.04	0.17
CONTRIBUTION CUMULEE =									0.9	0.0	4.8	3.1	14.8					
4 . gardedtt																		
AC_1	- enf_entr	0.14	32.17	-1.35	-0.34	-0.76	0.19	1.51	1.4	0.2	1.0	0.1	5.8	0.06	0.00	0.02	0.00	0.07
AC_2	- pas_enf_entr	4.62	0.03	0.04	0.01	0.02	-0.01	-0.05	0.0	0.0	0.0	0.0	0.2	0.06	0.00	0.02	0.00	0.07
CONTRIBUTION CUMULEE =									1.4	0.2	1.1	0.1	6.0					
6 . prestaidd																		
AE_1	- colo	1.32	2.60	-1.08	0.07	0.25	0.39	0.04	8.2	0.1	1.0	3.2	0.0	0.45	0.00	0.02	0.06	0.00
AE_2	- pas_colo	3.44	0.38	0.42	-0.03	-0.09	-0.15	-0.02	3.2	0.0	0.4	1.2	0.0	0.45	0.00	0.02	0.06	0.00
CONTRIBUTION CUMULEE =									11.4	0.1	1.4	4.4	0.1					
7 . prestaad																		
AF_1	- aprime_naiss	2.64	0.81	-0.17	0.16	0.53	-0.04	-0.17	0.4	0.7	9.1	0.1	1.3	0.04	0.03	0.35	0.00	0.04
AF_2	- pas_prime_naiss	2.12	1.24	0.21	-0.20	-0.66	0.05	0.21	0.5	0.9	11.3	0.1	1.7	0.04	0.03	0.35	0.00	0.04
CONTRIBUTION CUMULEE =									0.9	1.6	20.4	0.2	3.0					
8 . prestadd																		
AG_1	- aide_fraisgarde	0.86	4.55	-1.62	0.11	-0.17	0.46	-0.01	11.9	0.1	0.3	2.9	0.0	0.57	0.00	0.01	0.05	0.00
AG_2	- pas_aide_fraisgarde	3.90	0.22	0.36	-0.02	0.04	-0.10	0.00	2.6	0.0	0.1	0.6	0.0	0.57	0.00	0.01	0.05	0.00
CONTRIBUTION CUMULEE =									14.5	0.1	0.4	3.5	0.0					
9 . prestaed																		
AH_1	- aideetude	1.00	3.76	-1.01	0.22	0.45	0.52	-0.26	5.4	0.5	2.5	4.2	1.2	0.27	0.01	0.05	0.07	0.02
AH_2	- pas_aideetude	3.76	0.27	0.27	-0.06	-0.12	-0.14	0.07	1.4	0.1	0.7	1.1	0.3	0.27	0.01	0.05	0.07	0.02
CONTRIBUTION CUMULEE =									6.8	0.6	3.2	5.3	1.5					
10 . prestagd																		
AI_1	- aprestandi	0.76	5.29	-1.70	0.01	-0.25	0.53	-0.26	11.6	0.0	0.6	3.3	0.9	0.54	0.00	0.01	0.05	0.01
AI_2	- pas_prestandi	4.01	0.19	0.32	0.00	0.05	-0.10	0.05	2.2	0.0	0.1	0.6	0.2	0.54	0.00	0.01	0.05	0.01
CONTRIBUTION CUMULEE =									13.8	0.0	0.7	4.0	1.1					
11 . congdd																		
AJ_1	- compl_congepatmat	3.04	0.57	-0.43	-0.04	-0.01	-0.18	0.04	3.0	0.1	0.0	1.6	0.1	0.33	0.00	0.00	0.06	0.00
AJ_2	- pas_compl_congepatma	1.73	1.76	0.76	0.08	0.01	0.32	-0.07	5.3	0.1	0.0	2.8	0.2	0.33	0.00	0.00	0.06	0.00
CONTRIBUTION CUMULEE =									8.3	0.2	0.0	4.4	0.3					
12 . amenad																		
AK_1	- h_rentrée	4.07	0.17	-0.07	-0.22	0.09	0.10	-0.10	0.1	2.0	0.4	0.6	0.7	0.03	0.29	0.04	0.06	0.06
AK_2	- pas_h_rentrée	0.69	5.92	0.40	1.31	-0.51	-0.58	0.57	0.6	11.7	2.2	3.6	4.0	0.03	0.29	0.04	0.06	0.06
CONTRIBUTION CUMULEE =									0.7	13.7	2.6	4.2	4.7					
13 . amenbd																		
AL_1	- h_garde	2.00	1.38	0.17	-0.88	0.20	0.10	0.05	0.3	15.4	1.0	0.3	0.1	0.02	0.57	0.03	0.01	0.00
AL_2	- pas_h_garde	2.76	0.72	-0.12	0.64	-0.15	-0.07	-0.04	0.2	11.1	0.7	0.2	0.1	0.02	0.57	0.03	0.01	0.00
CONTRIBUTION CUMULEE =									0.5	26.5	1.7	0.5	0.1					
14 . amencc																		
AM_1	- h_enfmal	3.61	0.32	0.01	-0.39	0.08	0.06	-0.08	0.0	5.4	0.3	0.2	0.4	0.00	0.48	0.02	0.01	0.02
AM_2	- pas_h_enfmal	1.16	3.12	-0.03	1.22	-0.24	-0.18	0.24	0.0	17.0	0.8	0.6	1.2	0.00	0.48	0.02	0.01	0.02
CONTRIBUTION CUMULEE =									0.0	22.4	1.1	0.7	1.6					
15 . amendd																		
AN_1	- h_trajet	1.22	2.89	0.15	-1.14	0.26	-0.02	0.21	0.1	15.8	1.0	0.0	0.9	0.01	0.45	0.02	0.00	0.01
AN_2	- pas_h_trajet	3.54	0.35	-0.05	0.40	-0.09	0.01	-0.07	0.1	5.5	0.4	0.0	0.3	0.01	0.45	0.02	0.00	0.01
CONTRIBUTION CUMULEE =									0.2	21.2	1.4	0.0	1.3					
16 . dombd																		
AO_1	- L_domicil_poss	0.48	8.94	-0.54	-0.43	-0.44	-0.07	0.72	0.7	0.9	1.1	0.0	4.4	0.03	0.02	0.02	0.00	0.06
AO_2	- pas_L_domicil	4.28	0.11	0.06	0.05	0.05	0.01	-0.08	0.1	0.1	0.1	0.0	0.5	0.03	0.02	0.02	0.00	0.06
CONTRIBUTION CUMULEE =									0.8	1.0	1.3	0.0	4.9					
17 . partembd																		
AP_1	- apostespartiel	2.29	1.08	0.02	-0.26	-0.32	-0.47	-0.41	0.0	1.6	3.0	7.8	6.9	0.00	0.06	0.10	0.20	0.16
AP_2	- pas_postespartiel	2.47	0.92	-0.01	0.24	0.30	0.43	0.38	0.0	1.4	2.7	7.2	6.4	0.00	0.06	0.10	0.20	0.16
CONTRIBUTION CUMULEE =									0.0	3.0	5.7	14.9	13.2					
18 . partaccd																		
AQ_1	- Opas_tpspartiel	0.57	7.34	0.51	0.45	0.01	1.22	0.46	0.8	1.1	0.0	13.2	2.2	0.04	0.03	0.00	0.20	0.03
AQ_2	- lparfois_tpspartiel	2.21	1.16	0.25	0.11	0.15	0.05	-0.21	0.8	0.3	0.6	0.1	1.7	0.06	0.01	0.02	0.00	0.04
AQ_3	- 2tjrs_tpspartiel	1.98	1.40	-0.43	-0.25	-0.17	-0.40	0.10	2.0	1.2	0.7	5.0	0.4	0.13	0.05	0.02	0.12	0.01
CONTRIBUTION CUMULEE =									3.5	2.6	1.3	18.3	4.3					
19 . rttdd																		
AR_1	- Opas_RTT	1.43	2.33	0.41	-0.23	-0.64	0.53	0.23	1.3	0.7	7.3	6.3	1.3	0.07	0.02	0.18	0.12	0.02
AR_2	- 1RTT_contraint	0.72	5.57	0.05	0.39	0.16	0.10	-1.21	0.0	1.1	0.2	0.1	18.8	0.00	0.03	0.00	0.00	0.26
AR_3	- 2RTT_accord	2.24	1.13	-0.19	0.04	0.23	-0.24	0.02	0.4	0.0	1.4	2.0	0.0	0.03	0.00	0.05	0.05	0.00
AR_4	- 3RTT_libre	0.37	11.91	-0.54	-0.10	0.80	-0.79	1.34	0.6	0.0	2.9	3.6	11.9	0.02	0.00	0.05	0.05	0.15
CONTRIBUTION CUMULEE =									2.3	1.9	11.9	12.0	32.0					
20 . promood																		
AS_1	- jamais_avancem	1.19	3.01	-0.29	0.04	-0.59	-0.64	-0.12	0.5	0.0	5.1	7.6	0.3	0.03	0.00	0.12	0.14	0.00
AS_2	- parfois_avancem	1.48	2.22	-0.13	0.15	0.17	0.06	0.09	0.1	0.3	0.5	0.1	0.2	0.01	0.01	0.01	0.00	0.00
AS_3	- svt_avancem	1.56	2.06	0.18	-0.08	0.34	0.15	0.18	0.3	0.1	2.2	0.6	0.9	0.02	0.00	0.06	0.01	0.02
AS_4	- tjrs_avancem	0.54	7.81	0.48	-0.26	-0.14	0.80	-0.48	0.7	0.4	0.1	5.4	2.2	0.03	0.01	0.00	0.08	0.03
CONTRIBUTION CUMULEE =									1.6	0.8	8.0	13.6	3.6					
21 . prestahd																		
AT_1	- chèque_vac	1.79	1.66	-0.83	0.11	0.24	-0.07	-0.10	6.5	0.2	1.3	0.1	0.3	0.41	0.01	0.04	0.00	0.01
AT_2	- pas_chèque_vac	2.97	0.60	0.50	-0.06	-0.15	0.04	0.06	3.9	0.1	0.8	0.1	0.2	0.41	0.01	0.04	0.00	0.01
CONTRIBUTION CUMULEE =									10.4	0.3	2.1	0.2	0.5					
22 . fmutud																		
AU_1	- atousFinMutu	2.57	0.85	0.12	0.11	0.61	-0.23	-0.08	0.2	0.3	11.7	2.1	0.3	0.02	0.01	0.43	0.06	0.01
AU_2	- bcadreFinMutu	0.51	8.26	0.85	0.16	-0.44	0.77	0.49	2.0	0.1	1.3	4.7	2.2	0.09	0.00	0.02	0.07	0.03
AU_3	- cpasFinMutu	1.68	1.84	-0.44	-0.22	-0.79	0.11	-0.03	1.7	0.8	13.0	0.3	0.0	0.11	0.03	0.34	0.01	0.00
CONTRIBUTION CUMULEE =									3.9	1.2	26.0	7.1	2.6					
23 . mutud																		
AV_1	- aServMutu	1.20	2.96	-0.75	-0.16	-0.23	0.10	0.02	3.6	0.3	0.8	0.2	0.0	0.19	0.01	0.02	0.00	0.00
AV_2	- bpasServMutu	3.56	0.34	0.25	0.05	0.08	-0.04	-0.01	1.2	0.1	0.3	0.1	0.0	0.19	0.01	0.02	0.00	0.00
CONTRIBUTION CUMULEE =									4.8	0.4	1.1	0.3	0.0					

FIG. C.4 – Coordonnées des projections pour l'ACM pondérée

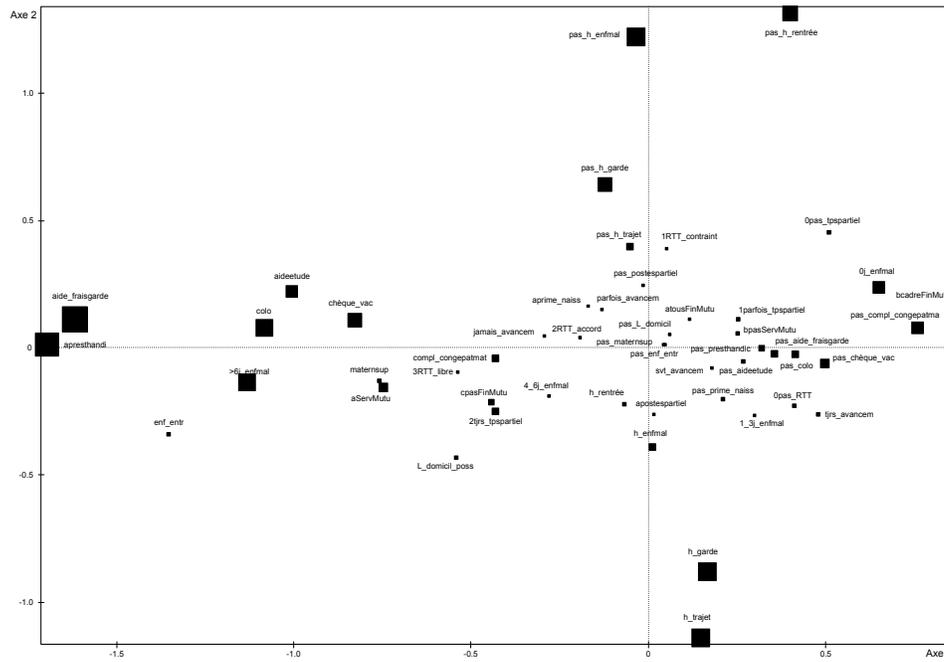


FIG. C.5 – Plan 1-2 de l'ACM pondérée

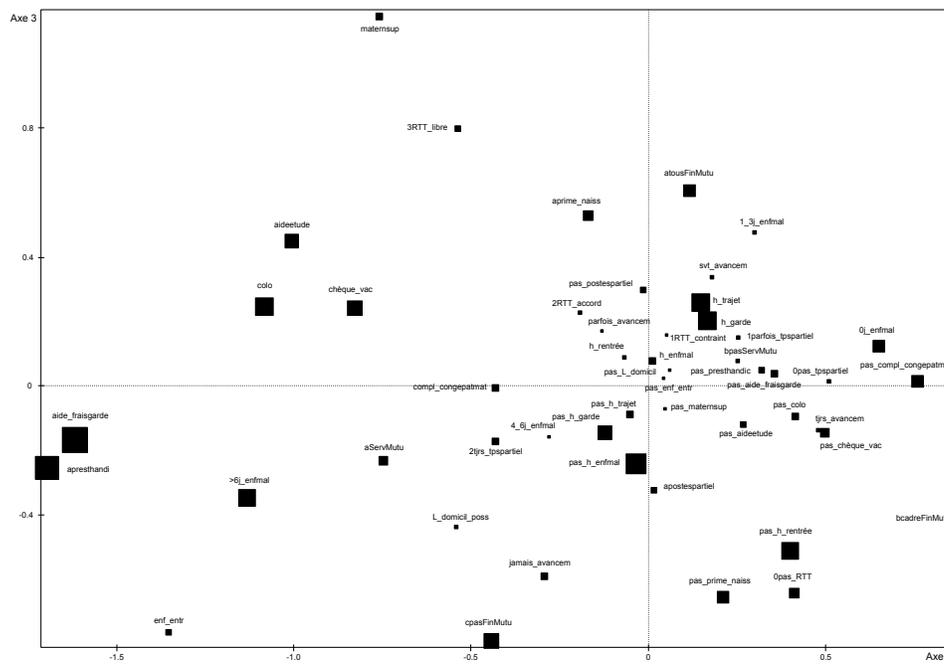


FIG. C.6 – Plan 1-3 de l'ACM pondérée

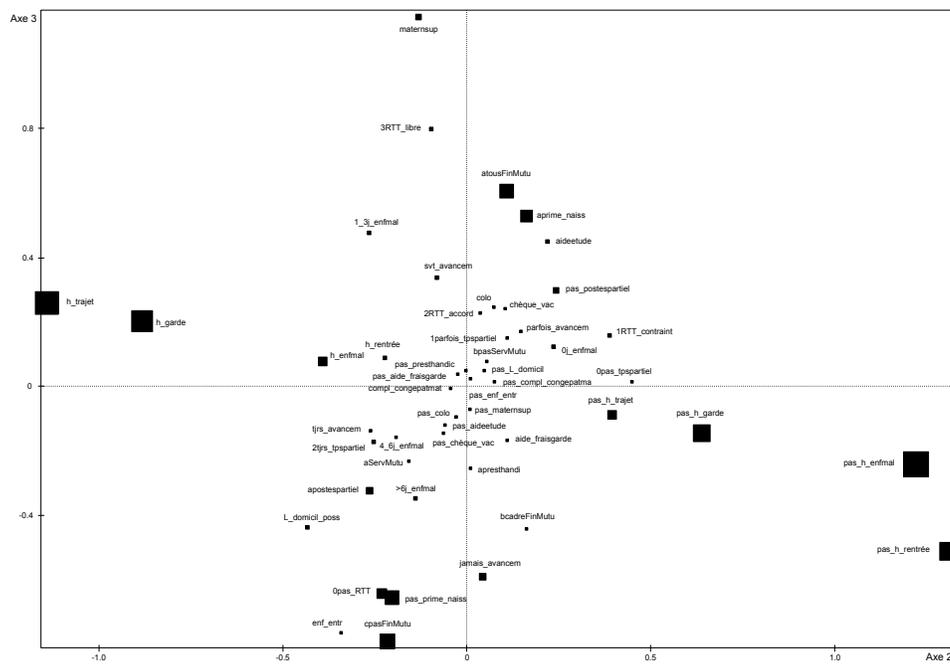


FIG. C.7 – Plan 2-3 de l'ACM pondérée

C.4 Projection de quelques variables supplémentaires

Pour l'ensemble des projections de variables supplémentaire un code couleur est utilisé : plus la couleur de nom de la modalité est écrit est couleur claire, moins la position de celle-ci est significative dans le plan. Plus précisément l'écriture en noire correspond à la significativité de la projection de la modalité sur les deux axes. Le gris foncé est utilisé lorsqu'elle n'est significative que sur un seul des deux axes. En cas d'absence de significativité de la projection sur ces deux axes, c'est le gris clair qui est utilisé.

Le test de significativité est fondé sur le calcul d'une statistique de test dite **valeur-test**. L'hypothèse testée est H_0 : les n_j observations prenant la modalité Y_j sont tirés au hasard dans la population. Soit ξ_j la coordonnée de la projection de la modalité sur l'axe étudié. On utilise la statistique (valeur-test) : $U = \sqrt{n_j \frac{n-1}{n-n_j} \xi_j}$.

Si H_0 est vraie, U suit une loi normale centrée réduite. $|U|$ est alors petite. Au contraire, plus U est élevée, plus la coordonnée ξ_j est significativement différente de 0. Le logiciel SPAD édite cette valeur-test pour les seules variables supplémentaires (SAS l'indique aussi pour les variables actives : le test a alors au mieux une valeur indicative puisque ces variables ont contribué à la déterminations des axes). On rejette H_0 ie la projection de la variable sur l'axe étudié est jugée significative dès lors que $|U| > 1.96$.

Type d'établissement

On observe une très nette séparation entre secteur privé et secteur public sur le premier axe. A l'exception du secteur privé non lucratif, les établissements privés tendent à ne proposer aucune des mesures organisant les deux premiers axes ([15] p.62).

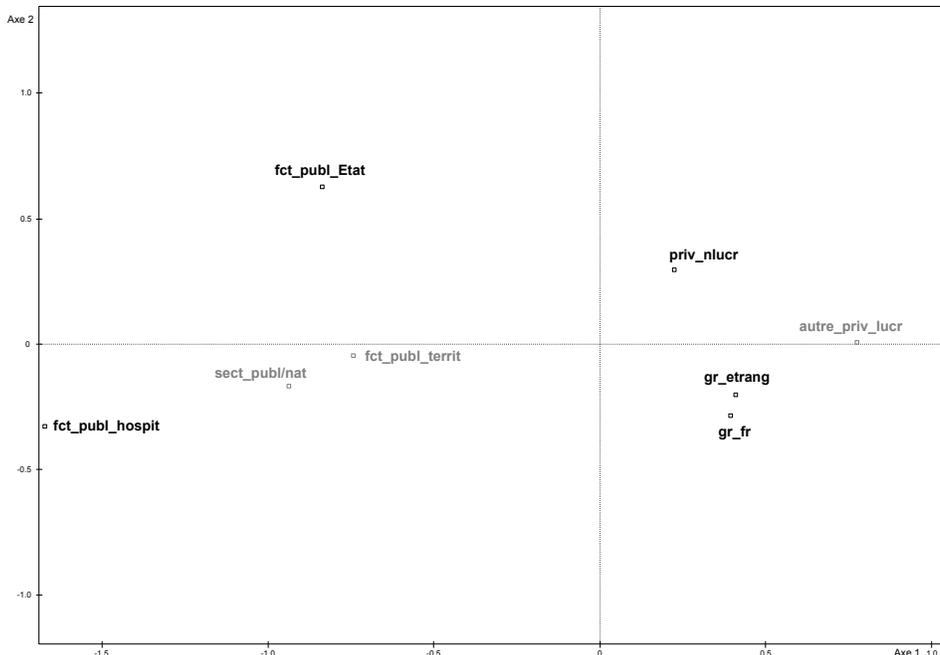


FIG. C.8 – Type d'établissement (plan 1-2)

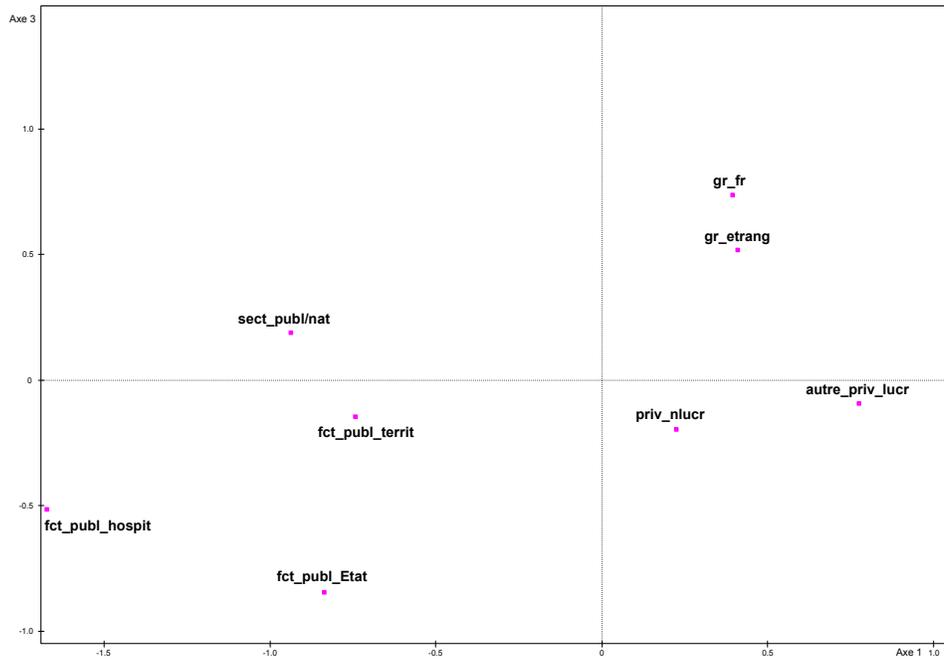


FIG. C.9 – Type d'établissement (plan 1-3)

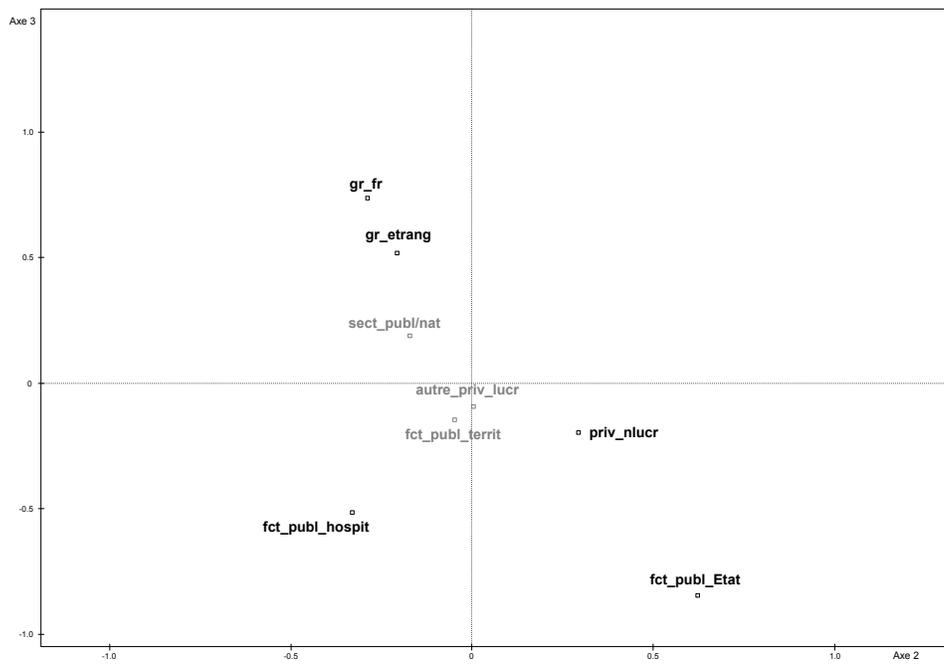


FIG. C.10 – Type d'établissement (plan 2-3)

Secteurs d'activité

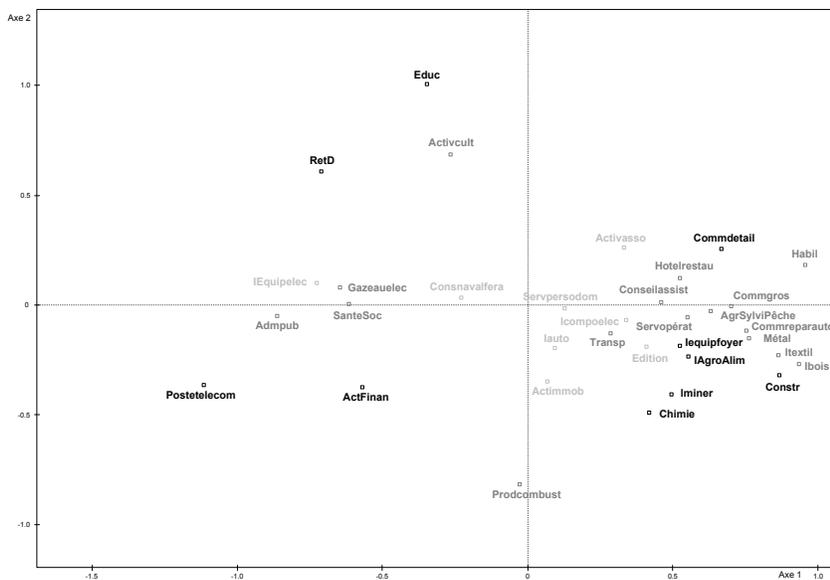


FIG. C.11 – Secteurs d'activité (plan 1-2)

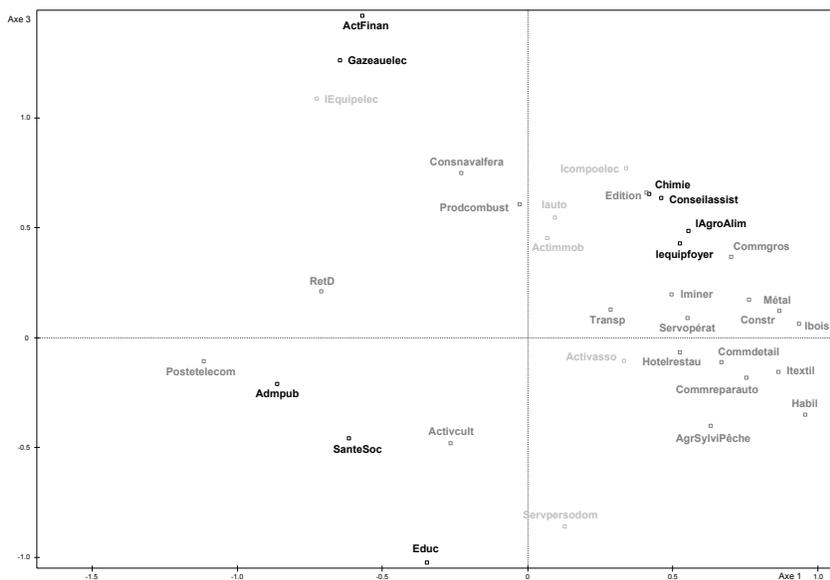


FIG. C.12 – Secteurs d'activité (plan 1-3)

Effectif de l'établissement

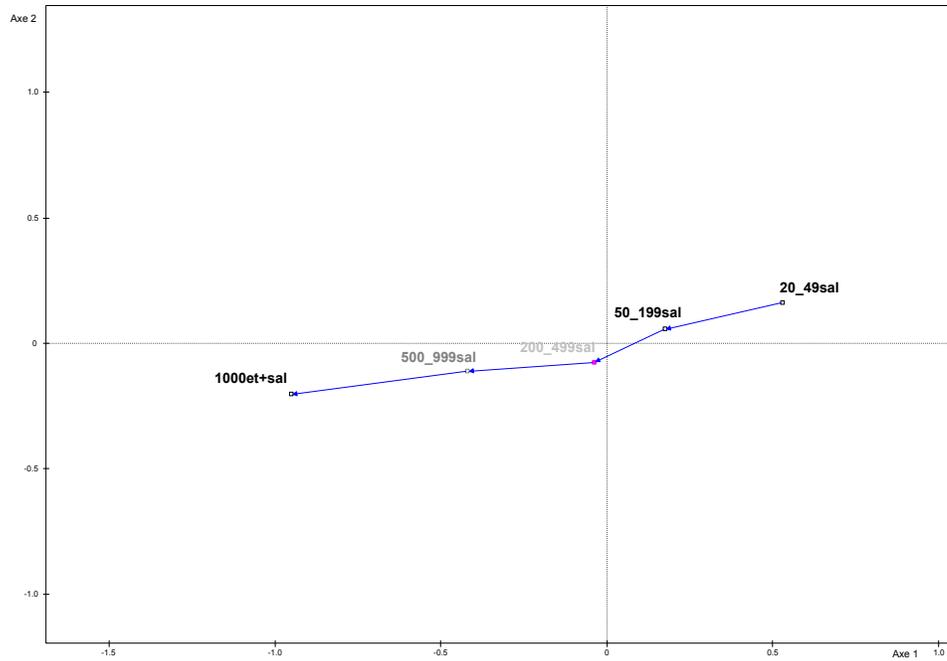


FIG. C.13 – Effectif de l'établissement (plan 1-2)

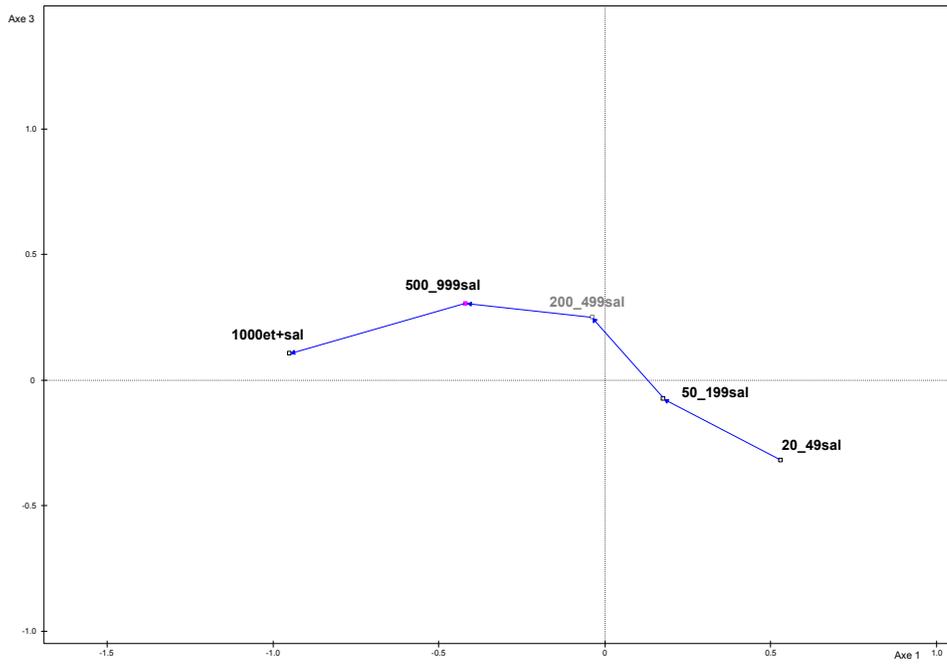


FIG. C.14 – Effectif de l'établissement (plan 1-3)

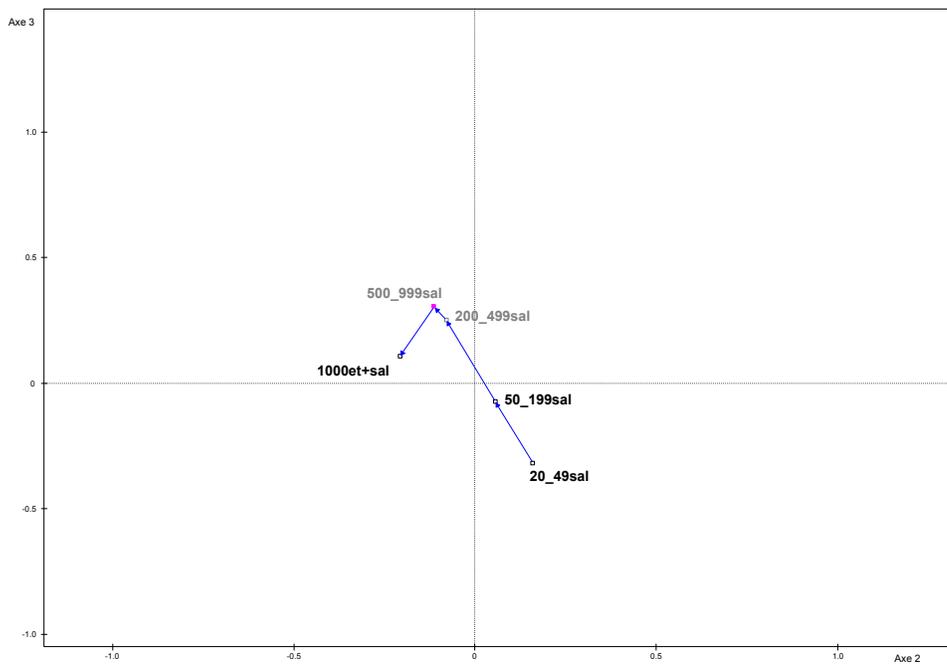


FIG. C.15 – Effectif de l'établissement (plan 2-3)

Proportion de femmes

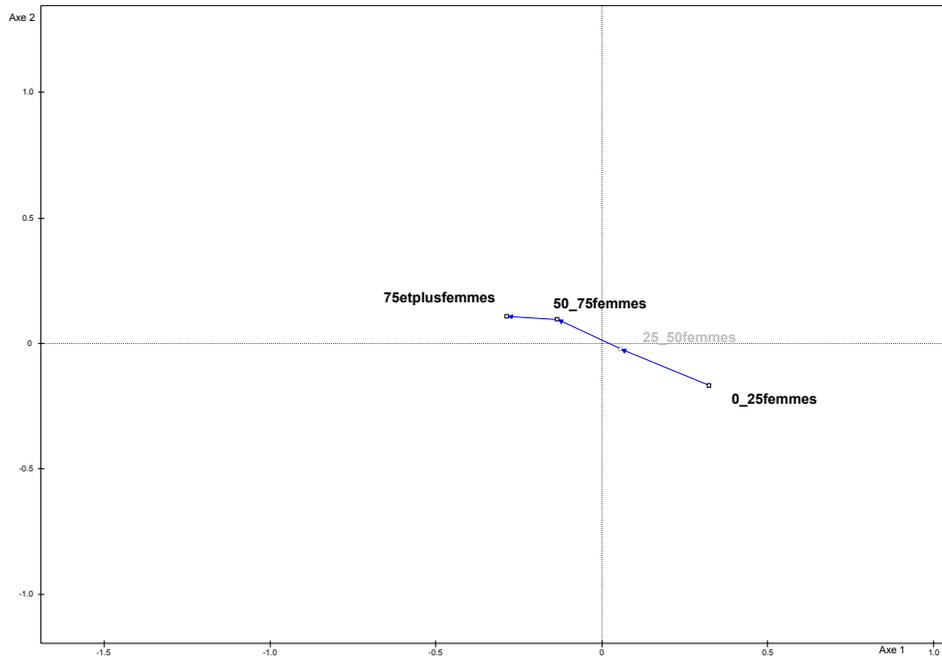


FIG. C.16 – Proportion de femmes (plan 1-2)

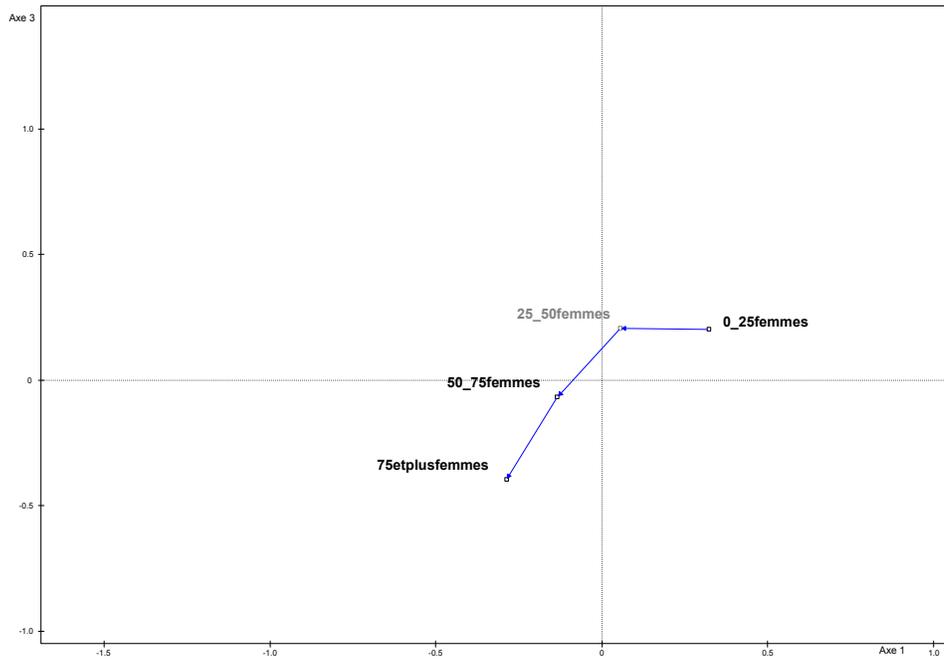


FIG. C.17 – Proportion de femmes (plan 1-3)

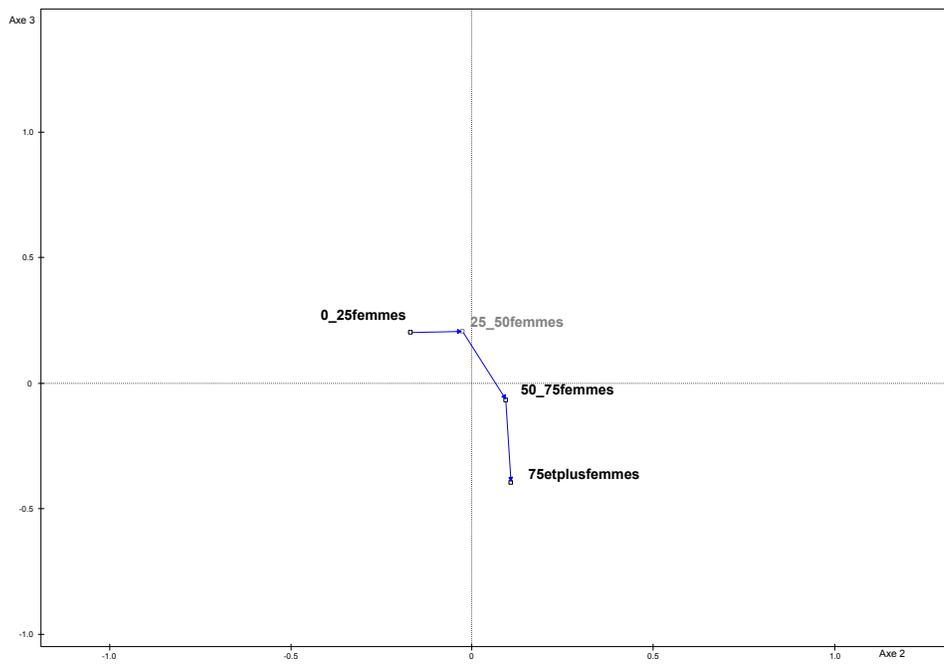


FIG. C.18 – Proportion de femmes (plan 2-3)

Proportion de CDI

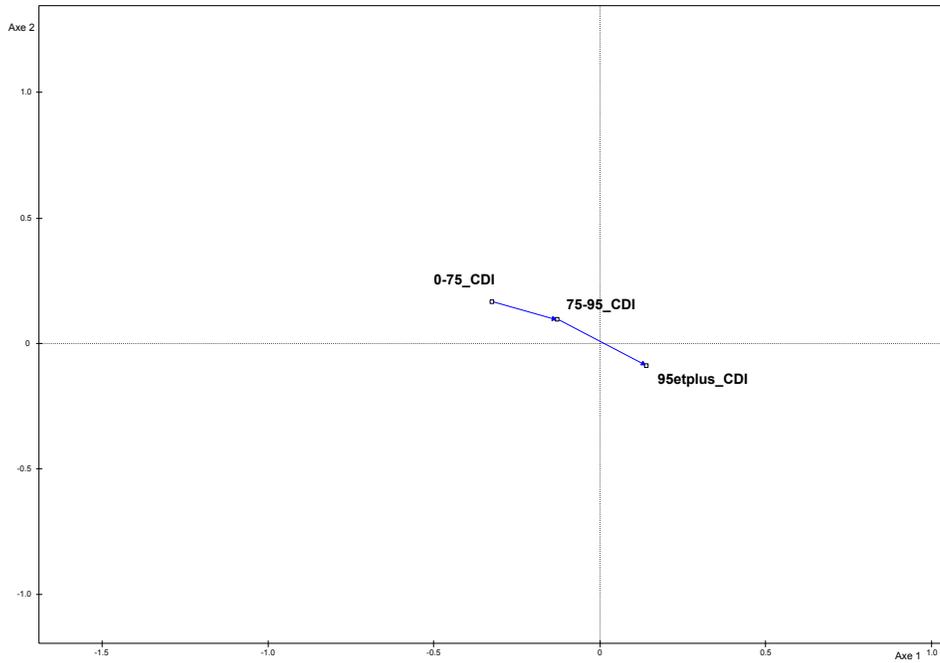


FIG. C.19 – Proportion de CDI (plan 1-2)

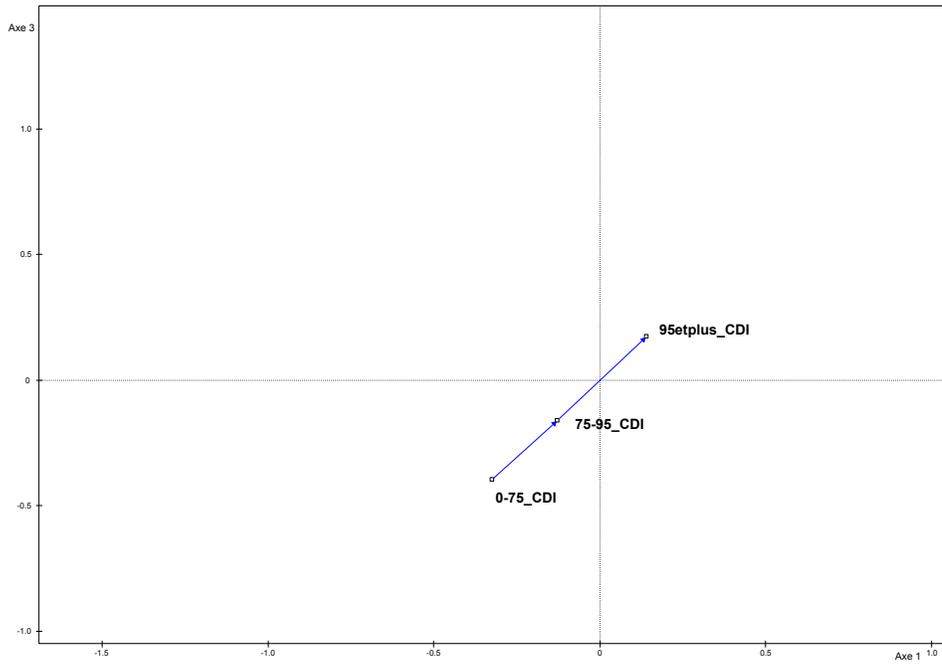


FIG. C.20 – Proportion de CDI (plan 1-3)

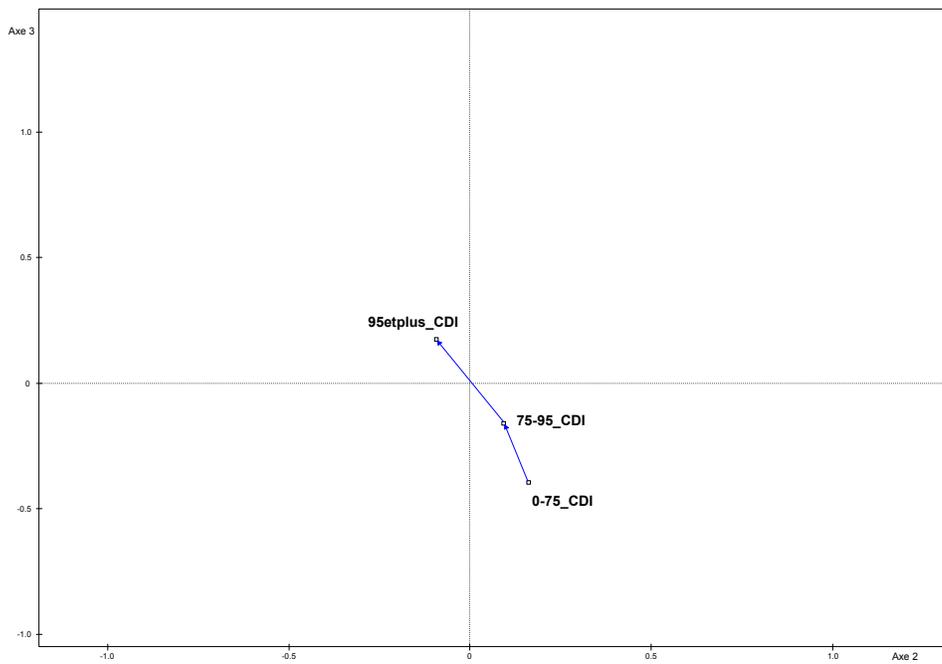


FIG. C.21 – Proportion de CDI (plan 2-3)

Age des salariés

La proportion des différentes classes d'âge des salariés ne se projette de façon significative que sur le premier axe.

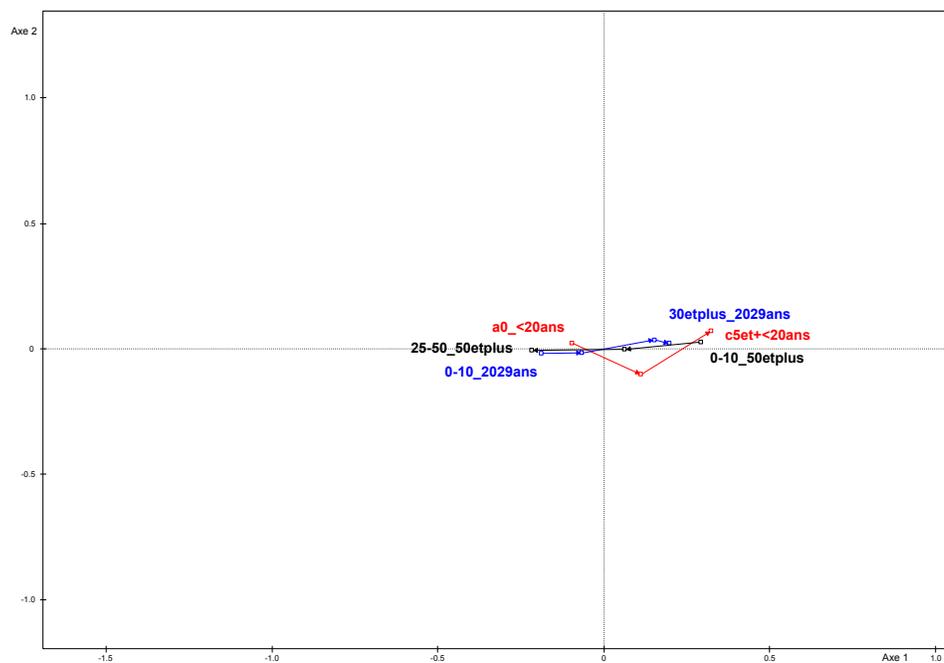


FIG. C.22 – Age des salariés (plan 1-2)

Proportion de cadres

L'axe 2 distingue les établissements ayant plus de 50% des cadres des autres. Sur l'axe 3, disposition hiérarchique des établissements de ceux qui ont le moins de cadres à ceux qui en ont le plus.

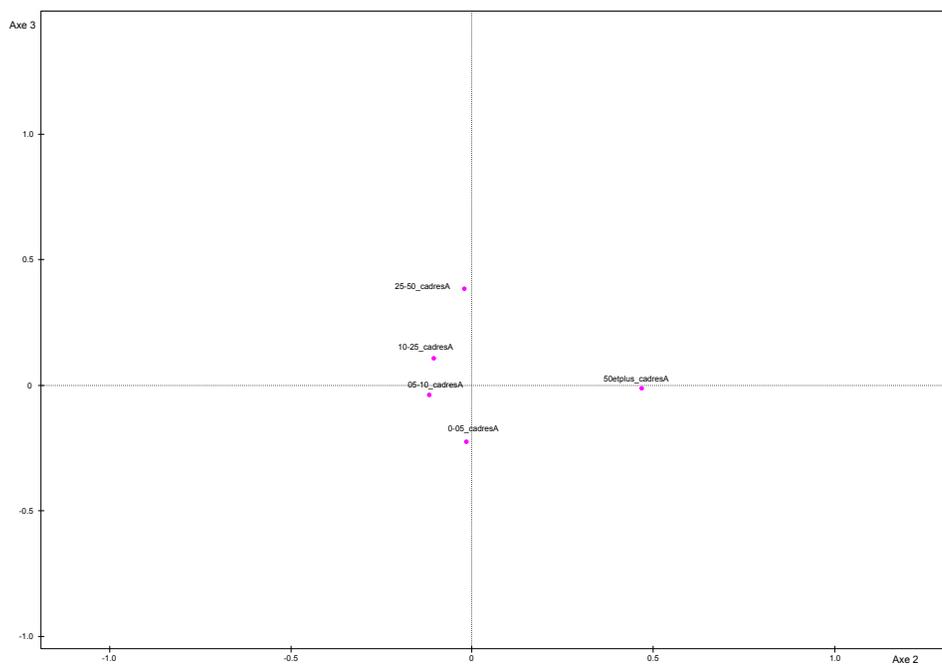


FIG. C.23 – Proportion de cadres (plan 2-3)

Régions

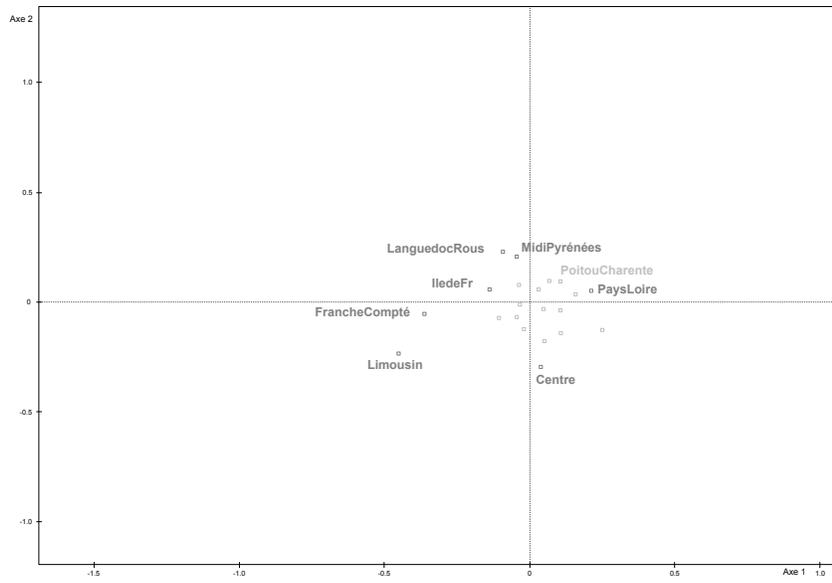


FIG. C.24 – Régions (plan 1-2)

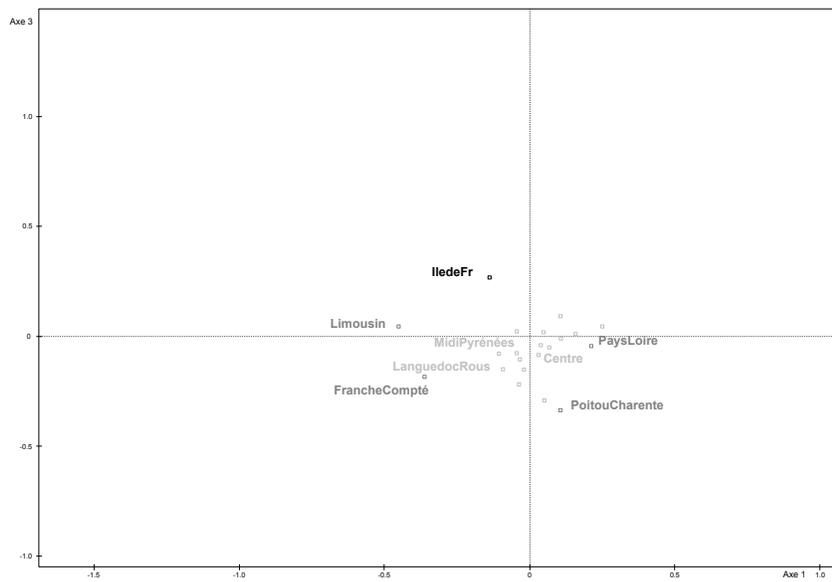


FIG. C.25 – Régions (plan 1-3)

Caractéristiques du directeur (âge et sexe)

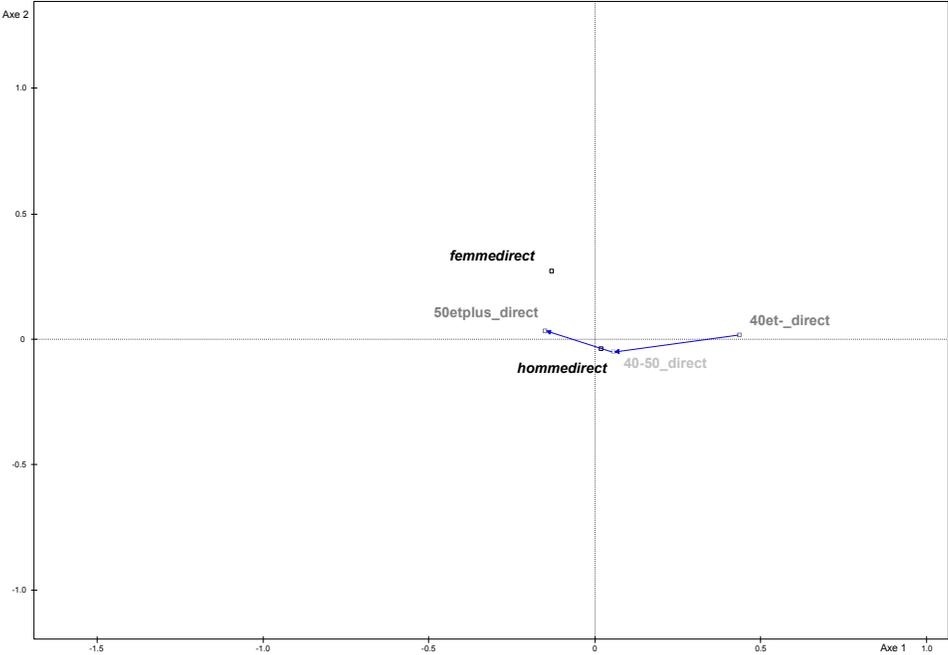


FIG. C.26 – Caractéristiques du directeur (plan 1-2)

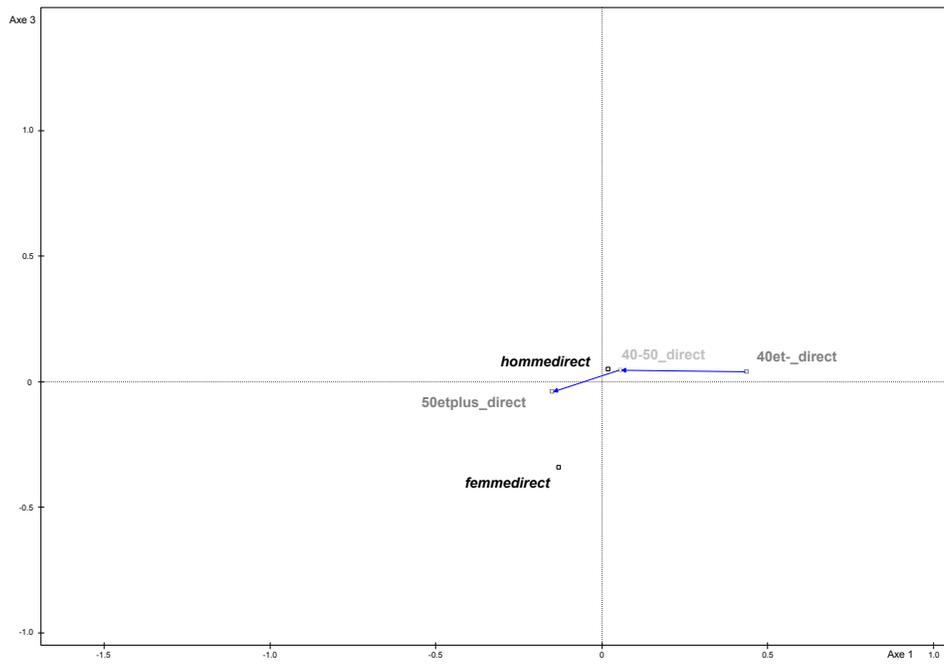


FIG. C.27 – Caractéristiques du directeur (plan 1-3)

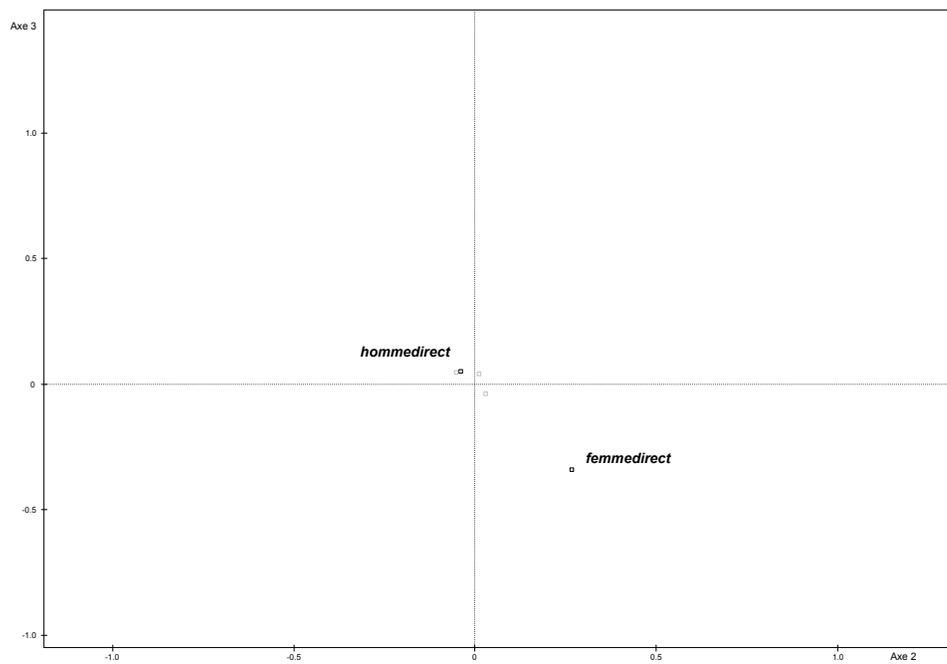


FIG. C.28 – Plan 2-3

Opinion sur le rôle de l'entreprise

Réponse à la question "Un employeur doit-il aider les salariés à coordonner leur travail avec leur vie de famille?"

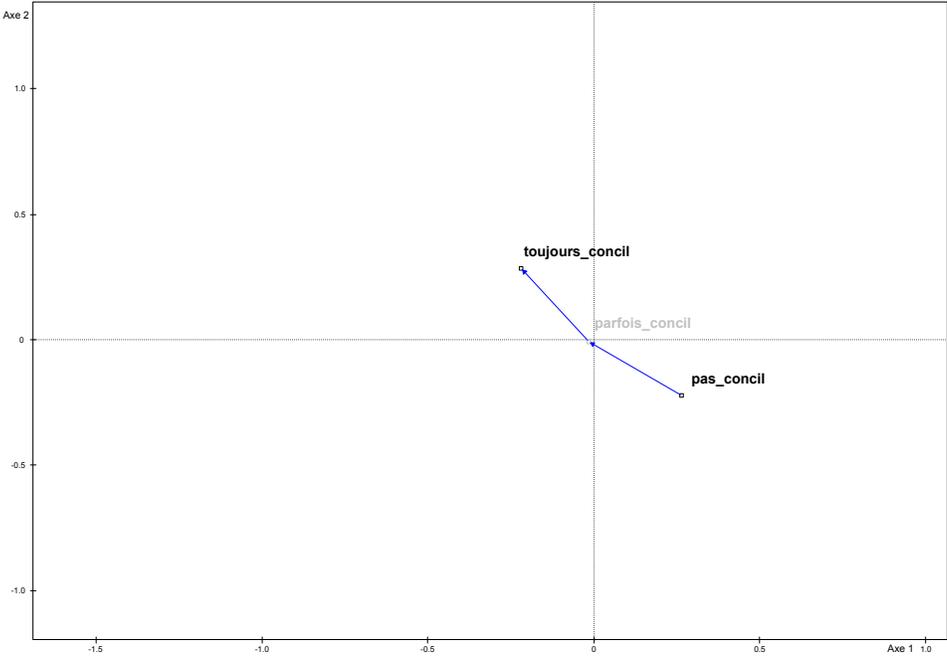


FIG. C.29 – Opinion sur le rôle de l'entreprise (plan 1-2)

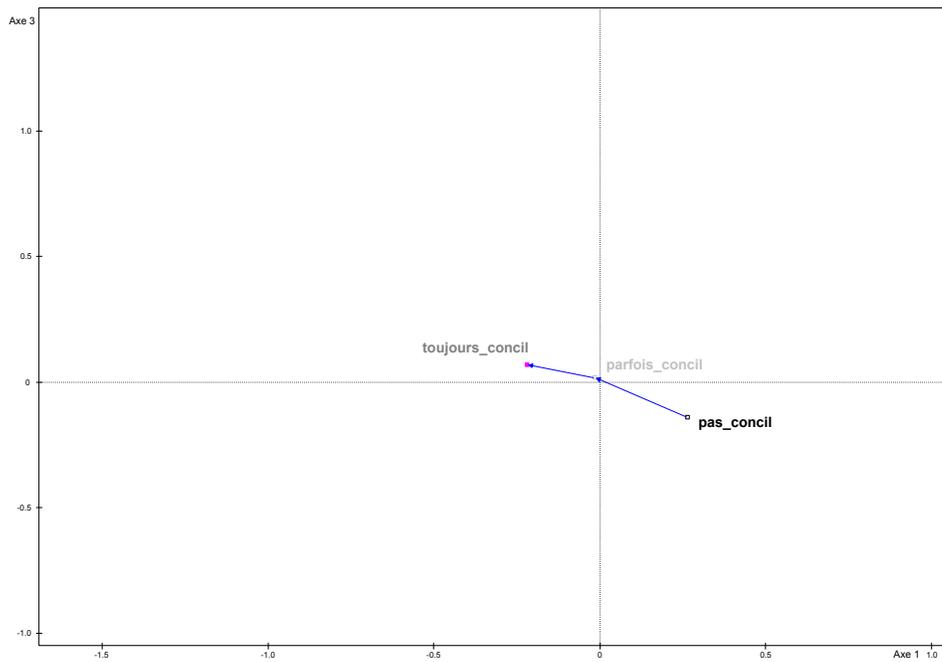


FIG. C.30 – Opinion sur le rôle de l’entreprise (plan 1-3)

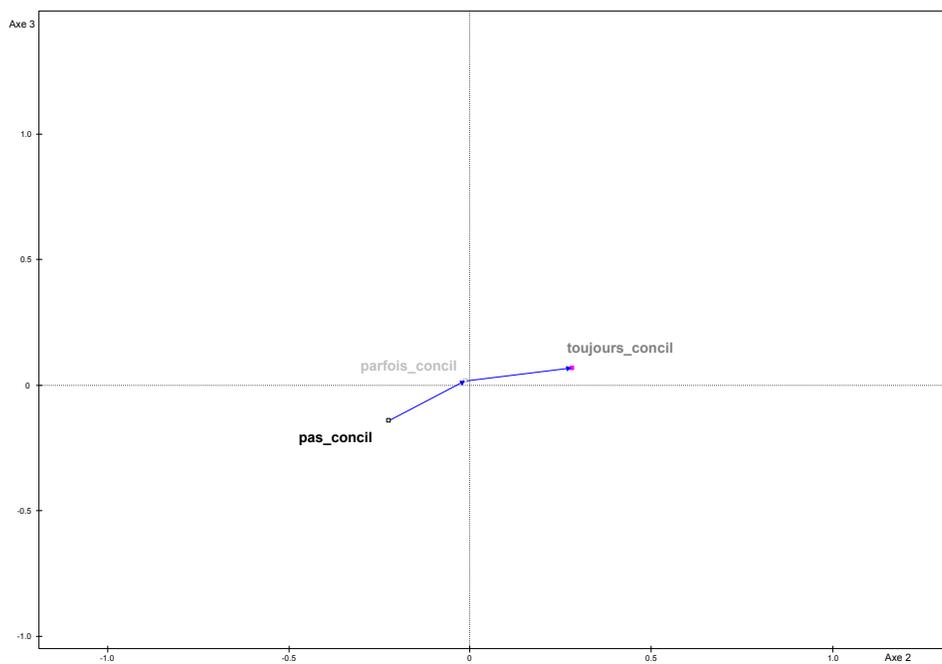


FIG. C.31 – Opinion sur le rôle de l’entreprise (plan 2-3)

Annexe D

Annexe classification

D.1 Classification ascendante hiérarchique sur les dix premiers axes

```
CLASSIFICATION HIERARCHIQUE (VOISINS RECIPROQUES)
SUR LES 10 PREMIERS AXES FACTORIELS
DESCRIPTION DES 50 NOEUDS D'INDICES LES PLUS ELEVES
NUM. AINE BENJ EFF. POIDS INDICE HISTOGRAMME DES INDICES DE NIVEAU
5254 5210 5160 91 91.00 0.00198 **
5255 5226 5240 125 125.00 0.00198 **
5256 5232 5231 81 81.00 0.00204 **
5257 5249 5207 117 117.00 0.00213 **
5258 5214 5217 71 71.00 0.00219 **
5259 5256 5170 107 107.00 0.00224 **
5260 5218 5247 104 104.00 0.00229 **
5261 5228 5255 184 184.00 0.00232 **
5262 5246 5215 97 97.00 0.00235 **
5263 5184 5227 115 115.00 0.00239 **
5264 5253 5194 120 120.00 0.00242 **
5265 5230 5189 110 110.00 0.00249 **
5266 5243 5155 75 75.00 0.00250 **
5267 5201 5219 72 72.00 0.00251 **
5268 5229 5191 70 70.00 0.00254 **
5269 5244 5234 99 99.00 0.00266 **
5270 5224 5172 84 84.00 0.00270 **
5271 5254 5204 149 149.00 0.00277 **
5272 5267 5205 105 105.00 0.00278 **
5273 5206 5220 110 110.00 0.00308 ***
5274 5263 5186 151 151.00 0.00322 ***
5275 5250 5209 113 113.00 0.00326 ***
5276 5245 5222 93 93.00 0.00328 ***
5277 5212 5251 158 158.00 0.00340 ***
5278 5233 5187 71 71.00 0.00351 ***
5279 5241 5235 117 117.00 0.00368 ***
5280 5239 5265 163 163.00 0.00388 ***
5281 5258 5252 159 159.00 0.00405 ***
5282 5278 5269 170 170.00 0.00431 ***
5283 5268 5272 175 175.00 0.00460 ***
5284 5248 5271 247 247.00 0.00481 ****
5285 5242 5274 259 259.00 0.00487 ****
5286 5262 5273 207 207.00 0.00541 ****
5287 5277 5280 321 321.00 0.00608 *****
5288 5261 5260 288 288.00 0.00629 *****
5289 5276 5266 168 168.00 0.00642 *****
5290 5279 5259 224 224.00 0.00707 *****
5291 5285 5286 466 466.00 0.00838 *****
5292 5275 5270 197 197.00 0.00861 *****
5293 5288 5257 405 405.00 0.01159 *****
5294 5291 5289 634 634.00 0.01191 *****
5295 5284 5264 367 367.00 0.01274 *****
5296 5293 5281 564 564.00 0.01326 *****
5297 5292 5283 372 372.00 0.01353 *****
5298 5287 5295 688 688.00 0.01828 *****
5299 5290 5296 788 788.00 0.02454 *****
5300 5282 5298 858 858.00 0.02883 *****
5301 5299 5300 1646 1646.00 0.03420 *****
5302 5294 5297 1006 1006.00 0.04042 *****
5303 5302 5301 2652 2652.00 0.12266 *****
SOMME DES INDICES DE NIVEAU = 0.74616
```

FIG. D.1 – Dendrogramme de la classification réalisée sur les dix premiers axes

COORDONNEES ET VALEURS-TEST
 AXES 1 A 5

ELEMENTS				VALEURS-TEST					COORDONNEES				
NUM .	IDENT	POIDS	EFF	1	2	3	4	5	1	2	3	4	5
1 .	5194	49.00	49	2.80	-2.11	8.10	-2.12	-4.55	0.40	-0.30	1.15	-0.30	-0.64
2 .	5253	71.00	71	5.47	-4.06	-1.94	-8.54	-5.69	0.64	-0.48	-0.23	-1.00	-0.67
3 .	5204	58.00	58	2.43	-2.07	5.83	1.17	-9.65	0.32	-0.27	0.76	0.15	-1.25
4 .	5254	91.00	91	3.19	-5.72	3.63	-5.96	-10.35	0.33	-0.59	0.37	-0.61	-1.07
5 .	5248	98.00	98	5.67	0.82	-6.66	-6.41	-12.26	0.56	0.08	-0.66	-0.64	-1.22
6 .	5189	71.00	71	7.08	-3.12	-3.93	0.09	-3.15	0.83	-0.37	-0.46	0.01	-0.37
7 .	5230	39.00	39	5.09	-6.28	-6.63	-6.80	-0.43	0.81	-1.00	-1.05	-1.08	-0.07
8 .	5239	53.00	53	8.84	-9.72	-7.58	1.17	2.89	1.20	-1.32	-1.03	0.16	0.39
9 .	5251	91.00	91	5.88	-11.73	3.97	2.02	2.50	0.61	-1.21	0.41	0.21	0.26
10 .	5212	67.00	67	6.57	-11.27	1.57	8.56	-0.65	0.79	-1.36	0.19	1.03	-0.08
11 .	5234	36.00	36	2.83	-4.05	0.21	-1.63	8.96	0.47	-0.67	0.04	-0.27	1.48
12 .	5244	63.00	63	6.59	1.36	-0.90	5.48	11.84	0.82	0.17	-0.11	0.68	1.47
13 .	5187	33.00	33	5.36	-8.81	-4.52	-0.25	12.95	0.93	-1.53	-0.78	-0.04	2.24
14 .	5233	38.00	38	8.01	-1.03	-7.73	5.45	12.15	1.29	-0.17	-1.24	0.88	1.96
15 .	5252	88.00	88	5.84	-1.91	2.10	10.12	-5.07	0.61	-0.20	0.22	1.06	-0.53
16 .	5217	37.00	37	4.70	6.34	4.07	6.76	-0.49	0.77	1.04	0.67	1.10	-0.08
17 .	5214	34.00	34	5.18	7.31	-5.74	0.86	-1.95	0.88	1.25	-0.98	0.15	-0.33
18 .	5207	38.00	38	6.33	9.57	-2.91	1.24	-1.18	1.02	1.54	-0.47	0.20	-0.19
19 .	5249	79.00	79	11.03	4.86	-7.29	4.07	0.32	1.22	0.54	-0.81	0.45	0.04
20 .	5247	52.00	52	4.33	8.66	3.54	-2.26	-2.36	0.60	1.19	0.49	-0.31	-0.32
21 .	5218	52.00	52	3.55	10.18	2.64	-3.56	-6.34	0.49	1.40	0.36	-0.49	-0.87
22 .	5240	61.00	61	6.38	7.26	7.74	1.76	3.80	0.81	0.92	0.98	0.22	0.48
23 .	5226	64.00	64	5.76	4.18	7.23	6.66	1.46	0.71	0.52	0.89	0.82	0.18
24 .	5228	59.00	59	6.98	1.12	4.19	9.70	-2.37	0.90	0.14	0.54	1.25	-0.31
25 .	5170	26.00	26	-3.27	2.71	-0.76	-5.19	4.84	-0.64	0.53	-0.15	-1.01	0.94
26 .	5231	34.00	34	-0.29	1.52	-5.00	-4.51	3.24	-0.05	0.26	-0.85	-0.77	0.55
27 .	5232	47.00	47	-3.18	9.16	-0.47	1.29	4.22	-0.46	1.32	-0.07	0.19	0.61
28 .	5235	67.00	67	0.26	14.85	-10.14	-5.17	1.77	0.03	1.79	-1.22	-0.62	0.21
29 .	5241	50.00	50	-0.45	0.81	-11.27	-7.11	-3.43	-0.06	0.11	-1.58	-1.00	-0.48
30 .	5205	33.00	33	-0.04	5.66	1.59	-7.44	0.46	-0.01	0.98	0.28	-1.29	0.08
31 .	5219	34.00	34	-0.53	3.95	10.99	-5.13	1.16	-0.09	0.67	1.87	-0.87	0.20
32 .	5201	38.00	38	0.26	2.44	9.78	-1.16	6.80	0.04	0.39	1.57	-0.19	1.09
33 .	5191	34.00	34	-1.56	5.56	6.10	-7.50	-1.56	-0.27	0.95	1.04	-1.28	-0.27
34 .	5229	36.00	36	-5.00	3.99	8.71	-1.97	-0.74	-0.83	0.66	1.44	-0.33	-0.12
35 .	5172	35.00	35	-1.37	-3.87	9.76	-8.59	0.79	-0.23	-0.65	1.64	-1.44	0.13
36 .	5224	49.00	49	1.90	-9.76	5.98	-10.33	4.61	0.27	-1.38	0.85	-1.46	0.65
37 .	5209	46.00	46	-10.87	-3.42	0.76	-0.95	4.52	-1.59	-0.50	0.11	-0.14	0.66
38 .	5250	67.00	67	-9.62	-7.37	8.47	-9.56	5.65	-1.16	-0.89	1.02	-1.15	0.68
39 .	5155	24.00	24	-6.53	-0.98	-5.67	3.86	8.25	-1.33	-0.20	-1.15	0.78	1.68
40 .	5243	51.00	51	-7.97	6.39	-7.86	0.65	2.27	-1.11	0.89	-1.09	0.09	0.31
41 .	5222	43.00	43	-10.39	7.50	5.80	-1.71	8.42	-1.57	1.13	0.88	-0.26	1.27
42 .	5245	50.00	50	-10.39	6.59	1.33	6.78	1.74	-1.46	0.92	0.19	0.95	0.24
43 .	5220	54.00	54	-2.54	-3.43	2.58	-0.24	0.47	-0.34	-0.46	0.35	-0.03	0.06
44 .	5206	56.00	56	-6.65	-5.80	-5.08	-2.62	-3.57	-0.88	-0.77	-0.67	-0.35	-0.47
45 .	5215	40.00	40	-3.43	0.01	0.58	6.04	-4.07	-0.54	0.00	0.09	0.95	-0.64
46 .	5246	57.00	57	-10.10	-3.27	-6.26	9.38	-2.85	-1.32	-0.43	-0.82	1.23	-0.37
47 .	5186	36.00	36	-7.36	-2.54	1.53	2.66	-6.10	-1.22	-0.42	0.25	0.44	-1.01
48 .	5227	75.00	75	-12.49	-0.56	-5.54	0.61	2.01	-1.42	-0.06	-0.63	0.07	0.23
49 .	5184	40.00	40	-11.26	-2.69	-1.42	6.74	-0.93	-1.77	-0.42	-0.22	1.06	-0.15
50 .	5242	108.00	108	-18.55	-4.95	-6.67	5.85	-7.19	-1.75	-0.47	-0.63	0.55	-0.68

FIG. D.2 – Classification sur les dix premiers axes

DESCRIPTION DES NOEUDS DE LA HIERACHIE
(INDICES EN POURCENTAGE DE LA SOMME DES INDICES : 0.47346)

NOEUD NUMERO	INDICE	SUCESSEURS		EFFECT.	POIDS	COMPOSITION	
		AINE	BENJ			PREMIER	DERNIER
51	0.42	23	22	125	125.00	22	23
52	0.43	27	26	81	81.00	26	27
53	0.45	19	18	117	117.00	18	19
54	0.46	17	16	71	71.00	16	17
55	0.47	52	25	107	107.00	25	27
56	0.48	21	20	104	104.00	20	21
57	0.49	24	51	184	184.00	22	24
58	0.50	46	45	97	97.00	45	46
59	0.51	49	48	115	115.00	48	49
60	0.51	2	1	120	120.00	1	2
61	0.53	7	6	110	110.00	6	7
62	0.53	40	39	75	75.00	39	40
63	0.53	32	31	72	72.00	31	32
64	0.54	34	33	70	70.00	33	34
65	0.56	12	11	99	99.00	11	12
66	0.57	36	35	84	84.00	35	36
67	0.58	4	3	149	149.00	3	4
68	0.59	63	30	105	105.00	30	32
69	0.65	44	43	110	110.00	43	44
70	0.68	59	47	151	151.00	47	49
71	0.69	38	37	113	113.00	37	38
72	0.69	42	41	93	93.00	41	42
73	0.72	10	9	158	158.00	9	10
74	0.74	14	13	71	71.00	13	14
75	0.78	29	28	117	117.00	28	29
76	0.82	8	61	163	163.00	6	8
77	0.85	54	15	159	159.00	15	17
78	0.91	74	65	170	170.00	11	14
79	0.97	64	68	175	175.00	30	34
80	1.02	5	67	247	247.00	3	5
81	1.03	50	70	259	259.00	47	50
82	1.14	58	69	207	207.00	43	46
83	1.28	73	76	321	321.00	6	10
84	1.33	57	56	288	288.00	20	24
85	1.36	72	62	168	168.00	39	42
86	1.49	75	55	224	224.00	25	29
87	1.77	81	82	466	466.00	43	50
88	1.82	71	66	197	197.00	35	38
89	2.45	84	53	405	405.00	18	24
90	2.52	87	85	634	634.00	39	50
91	2.69	80	60	367	367.00	1	5
92	2.80	89	77	564	564.00	15	24
93	2.86	88	79	372	372.00	30	38
94	3.86	83	91	688	688.00	1	10
95	5.18	86	92	788	788.00	15	29
96	6.09	78	94	858	858.00	1	14
97	7.22	95	96	1646	1646.00	1	29
98	8.54	90	93	1006	1006.00	30	50
99	25.91	98	97	2652	2652.00	1	50

FIG. D.3 – Classification sur les dix premiers axes

PARTITION PAR COUPURE D'UN ARBRE HIERARCHIQUE
 RECHERCHE DES MEILLEURES PARTITIONS
 RECHERCHE DES PALIERS

PALIER	VALEUR DU	ENTRE	PALIER
5297--	5298	-4.72	*****
5298--	5299	-3.05	*****

LISTE DES 2 MEILLEURE(S) PARTITION(S) ENTRE 3 ET 10 CLASSES
 1 - PARTITION EN 7 CLASSES
 2 - PARTITION EN 6 CLASSES

Coupure 'b' de l'arbre en 7 classes
 FORMATION DES CLASSES (INDIVIDUS ACTIFS)
 DESCRIPTION SOMMAIRE

CLASSE	EFFECTIF	POIDS	CONTENU
bb1b	367	367.00	1 A 5
bb2b	321	321.00	6 A 10
bb3b	170	170.00	11 A 14
bb4b	564	564.00	15 A 24
bb5b	224	224.00	25 A 29
bb6b	372	372.00	30 A 38
bb7b	634	634.00	39 A 50

COORDONNEES ET VALEURS-TEST AVANT CONSOLIDATION
 AXES 1 A 5

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST					COORDONNEES					DISTO.
			1	2	3	4	5	1	2	3	4	5	
Coupure 'b' de l'arbre en 7 classes													
bb1b - Classe 1 / 7	367	367.00	9.5	-6.2	3.0	-11.0	-20.7	0.21	-0.09	0.04	-0.13	-0.24	0.15
bb2b - Classe 2 / 7	321	321.00	15.6	-20.0	-4.7	3.3	0.6	0.37	-0.33	-0.07	0.04	0.01	0.29
bb3b - Classe 3 / 7	170	170.00	11.7	-5.6	-6.3	5.2	23.4	0.39	-0.13	-0.13	0.09	0.41	0.57
bb4b - Classe 4 / 7	564	564.00	21.4	18.8	5.8	13.3	-5.0	0.36	0.22	0.06	0.12	-0.04	0.20
bb5b - Classe 5 / 7	224	224.00	-2.9	14.7	-13.7	-9.4	4.4	-0.08	0.30	-0.25	-0.15	0.07	0.20
bb6b - Classe 6 / 7	372	372.00	-10.4	-2.6	21.9	-19.0	8.4	-0.22	-0.04	0.30	-0.22	0.10	0.24
bb7b - Classe 7 / 7	634	634.00	-36.6	-2.0	-9.7	12.3	-2.0	-0.57	-0.02	-0.10	0.10	-0.02	0.35

AXES 6 A 10

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST					COORDONNEES					DISTO.
			6	7	8	9	10	6	7	8	9	10	
Coupure 'b' de l'arbre en 7 classes													
bb1b - Classe 1 / 7	367	367.00	-1.0	3.1	-9.0	7.2	-5.3	-0.01	0.03	-0.10	0.08	-0.06	0.15
bb2b - Classe 2 / 7	321	321.00	-5.1	-1.0	15.4	0.6	5.5	-0.06	-0.01	0.18	0.01	0.06	0.29
bb3b - Classe 3 / 7	170	170.00	10.7	-4.6	-19.5	11.6	-9.6	0.18	-0.08	-0.32	0.19	-0.15	0.57
bb4b - Classe 4 / 7	564	564.00	-3.6	-1.8	1.7	-7.2	4.1	-0.03	-0.02	0.01	-0.06	0.03	0.20
bb5b - Classe 5 / 7	224	224.00	-3.9	-0.2	8.2	2.9	-1.1	-0.06	0.00	0.12	0.04	-0.01	0.20
bb6b - Classe 6 / 7	372	372.00	7.9	8.6	0.6	-14.5	3.1	0.09	0.09	0.01	-0.15	0.03	0.24
bb7b - Classe 7 / 7	634	634.00	-2.0	-4.3	-0.9	3.9	-0.2	-0.02	-0.03	-0.01	0.03	0.00	0.35

CONSOLIDATION DE LA PARTITION
 AUTOUR DES 7 CENTRES DE CLASSES, REALISEE PAR 10 ITERATIONS A CENTRES MOBILES
 PROGRESSION DE L'INERTIE INTER-CLASSES

ITERATION	I.TOTALE	I.INTER	QUOTIENT
0	0.74616	0.26894	0.36043
1	0.74616	0.30921	0.41307
2	0.74616	0.31047	0.41609
3	0.74616	0.31110	0.41694
4	0.74616	0.31124	0.41713
5	0.74616	0.31133	0.41724

ARRET APRES L'ITERATION 5 L'ACCROISSEMENT DE L'INERTIE INTER-CLASSES
 PAR RAPPORT A L'ITERATION PRECEDENTE N'EST QUE DE 0.02 %.

DECOMPOSITION DE L'INERTIE
 CALCULEE SUR 10 AXES.

	INERTIES		EFFECTIFS		POIDS		DISTANCES	
	AVANT	APRES	AVANT	APRES	AVANT	APRES	AVANT	APRES
INTER-CLASSES	0.2689	0.3113						
INTRA-CLASSE								
CLASSE 1 / 7	0.0588	0.0737	367	489	367.00	489.00	0.1452	0.1112
CLASSE 2 / 7	0.0949	0.0670	321	439	321.00	439.00	0.2911	0.2813
CLASSE 3 / 7	0.0342	0.0337	170	171	170.00	171.00	0.5654	0.5673
CLASSE 4 / 7	0.1010	0.0725	564	452	564.00	452.00	0.2049	0.2737
CLASSE 5 / 7	0.0394	0.0467	224	275	224.00	275.00	0.2041	0.2713
CLASSE 6 / 7	0.0853	0.0582	372	285	372.00	285.00	0.2427	0.3534
CLASSE 7 / 7	0.1145	0.0830	634	541	634.00	541.00	0.3451	0.4653
TOTALE	0.7462	0.7462						

QUOTIENT (INERTIE INTER / INERTIE TOTALE) : AVANT ... 0.3604
 APRES ... 0.4172

COORDONNEES ET VALEURS-TEST APRES CONSOLIDATION
 AXES 1 A 5

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST					COORDONNEES					DISTO.
			1	2	3	4	5	1	2	3	4	5	
Coupure 'b' de l'arbre en 7 classes													
bb1b - Classe 1 / 7	489	489.00	8.0	-4.4	7.5	-7.6	-22.9	0.15	-0.06	0.09	-0.08	-0.22	0.11
bb2b - Classe 2 / 7	439	439.00	18.6	-23.7	-6.6	1.2	-0.1	0.36	-0.33	-0.08	0.01	0.00	0.28
bb3b - Classe 3 / 7	171	171.00	11.5	-6.6	-7.1	5.2	23.1	0.38	-0.16	-0.15	0.09	-0.40	0.57
bb4b - Classe 4 / 7	452	452.00	20.1	22.1	6.9	12.6	0.3	0.38	0.30	0.09	0.13	0.00	0.27
bb5b - Classe 5 / 7	275	275.00	-5.0	21.0	-14.8	-13.0	4.2	-0.13	0.38	-0.24	-0.18	0.06	0.27
bb6b - Classe 6 / 7	285	285.00	-11.1	-2.2	23.8	-17.1	9.9	-0.28	-0.04	0.38	-0.23	0.13	0.35
bb7b - Classe 7 / 7	541	541.00	-38.1	-4.6	-10.3	14.2	-3.1	-0.66	-0.06	-0.11	0.13	-0.03	0.47

AXES 6 A 10

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST					COORDONNEES					DISTO.
			6	7	8	9	10	6	7	8	9	10	
Coupure 'b' de l'arbre en 7 classes													
bb1b - Classe 1 / 7	489	489.00	-4.1	-4.0	-11.1	4.9	-10.8	-0.04	-0.04	-0.10	0.04	-0.10	0.11
bb2b - Classe 2 / 7	439	439.00	-2.9	2.1	17.3	-1.7	-7.4	-0.03	0.02	0.17	-0.02	0.07	0.28
bb3b - Classe 3 / 7	171	171.00	11.0	-4.2	-20.2	10.9	-9.3	0.19	-0.07	-0.33	0.18	-0.15	0.57
bb4b - Classe 4 / 7	452	452.00	-2.0	-0.1	5.2	15.2	8.0	-0.02	0.00	0.05	-0.05	0.07	0.27
bb5b - Classe 5 / 7	275	275.00	-0.9	-3.0	7.6	-3.3	-3.3	-0.01	-0.04	0.10	0.05	-0.04	0.27
bb6b - Classe 6 / 7	285	285.00	5.9	10.9	-3.6	-13.5	4.4	0.08	0.14	-0.04	-0.17	0.05	0.35
bb7b - Classe 7 / 7	541	541.00	-2.1	-1.6	-0.9	2.7	0.9	-0.02	-0.01	-0.01	0.02	0.01	0.47

FIG. D.4 – Classification en 7 classes sur les dix premiers axes

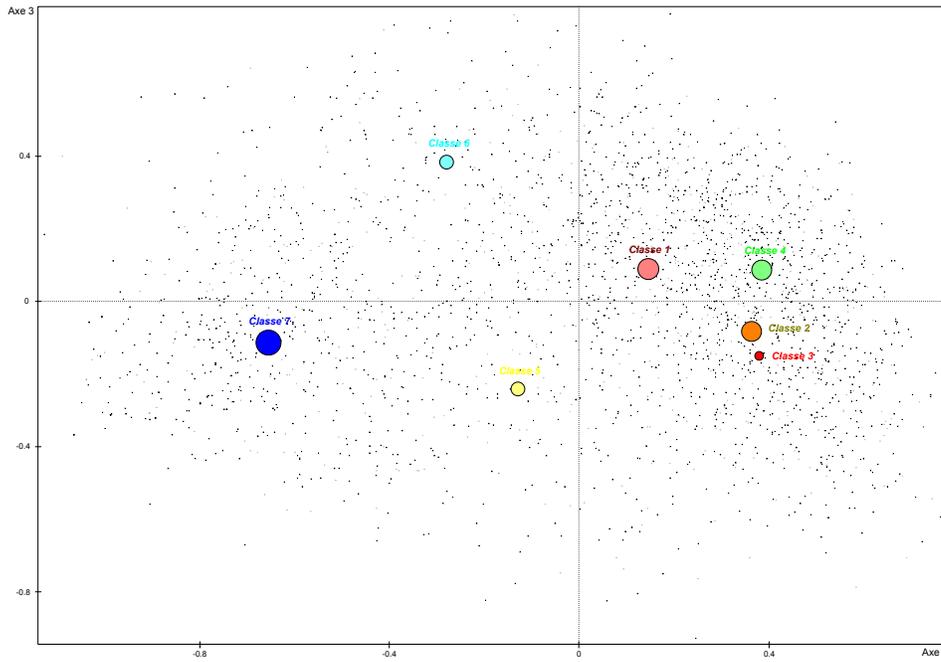


FIG. D.5 – Barycentres des 7 classes dans le plan 1-3

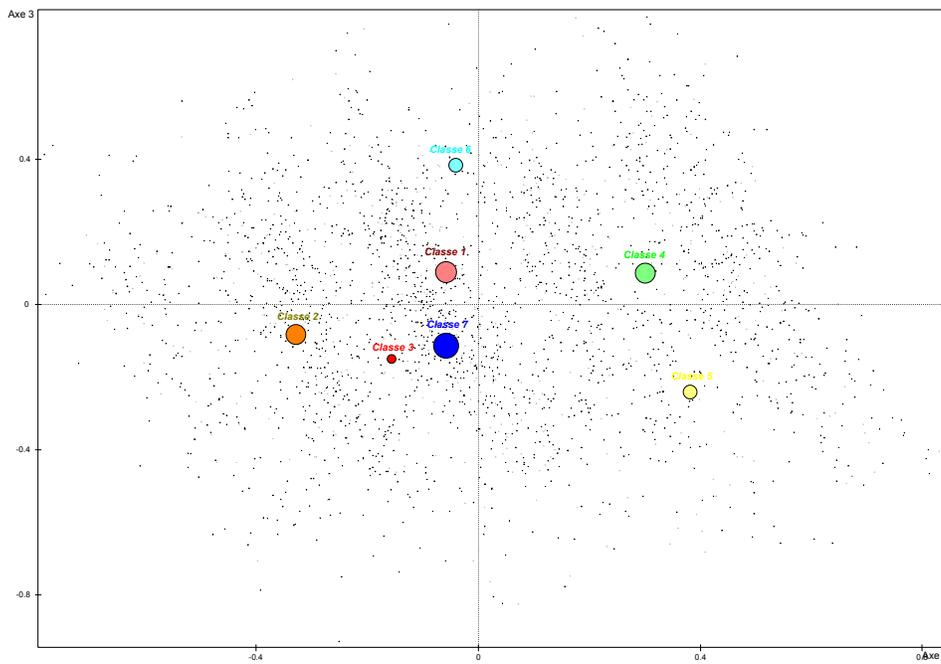


FIG. D.6 – Barycentres des 7 classes dans le plan 2-3

2 - PARTITION EN 6 CLASSES
Coupure 'a' de l'arbre en 6 classes
FORMATION DES CLASSES (INDIVIDUS ACTIFS)
DESCRIPTION SOMMAIRE

CLASSE	EFFECTIF	POIDS	CONTENU
aa1a	688	688.00	1 A 10
aa2a	170	170.00	11 A 14
aa3a	564	564.00	15 A 24
aa4a	224	224.00	25 A 29
aa5a	372	372.00	30 A 38
aa6a	634	634.00	39 A 50

COORDONNEES ET VALEURS-TEST AVANT CONSOLIDATION
AXES 1 A 5

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST					COORDONNEES					DISTO.
			1	2	3	4	5	1	2	3	4	5	
Coupure 'a' de l'arbre en 6 classes													
aa1a - Classe 1 / 6	688	688.00	19.1	-19.7	-1.1	-6.2	-15.8	0.28	-0.21	-0.01	-0.05	-0.12	0.14
aa2a - Classe 2 / 6	170	170.00	11.7	-5.6	-6.3	5.2	23.4	0.39	-0.13	-0.13	0.09	0.41	0.57
aa3a - Classe 3 / 6	564	564.00	21.4	18.8	5.8	13.3	-5.0	0.36	0.22	0.06	0.12	-0.04	0.20
aa4a - Classe 4 / 6	224	224.00	-2.9	14.7	-13.7	-9.4	4.4	-0.08	0.30	-0.25	-0.15	0.07	0.20
aa5a - Classe 5 / 6	372	372.00	-10.4	-2.6	21.9	-19.0	8.4	-0.22	-0.04	0.30	-0.22	0.10	0.24
aa6a - Classe 6 / 6	634	634.00	-36.6	-2.0	-9.7	12.3	-2.0	-0.57	-0.02	-0.10	0.10	-0.02	0.35

AXES 6 A 10

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST						COORDONNEES						DISTO.
			6	7	8	9	10	6	7	8	9	10			
Coupure 'a' de l'arbre en 6 classes															
aa1a - Classe 1 / 6	688	688.00	-4.5	1.7	4.4	6.1	-0.1	-0.03	0.01	0.03	0.04	0.00	0.14		
aa2a - Classe 2 / 6	170	170.00	10.7	-4.6	-19.5	11.6	-9.6	0.18	-0.08	-0.32	0.19	-0.15	0.57		
aa3a - Classe 3 / 6	564	564.00	-3.6	-1.8	-1.7	-7.2	4.3	-0.03	-0.02	0.01	-0.06	0.03	0.20		
aa4a - Classe 4 / 6	224	224.00	-3.9	-0.2	8.2	2.9	-1.1	-0.06	0.00	0.12	0.04	-0.01	0.20		
aa5a - Classe 5 / 6	372	372.00	7.9	8.6	0.6	-14.5	3.1	0.09	0.09	0.01	-0.15	0.03	0.24		
aa6a - Classe 6 / 6	634	634.00	-2.0	-4.3	-0.9	3.9	-0.2	-0.02	-0.03	-0.01	0.03	0.00	0.35		

CONSOLIDATION DE LA PARTITION
AUJOURD'HUI 6 CENTRES DE CLASSES, REALISEE PAR 10 ITERATIONS A CENTRES MOBILES
PROGRESSION DE L'INERTIE INTER-CLASSES

ITERATION	I.TOTALE	I.INTER	QUOTIENT
0	0.74616	0.25066	0.33593
1	0.74616	0.28697	0.38459
2	0.74616	0.28857	0.38674
3	0.74616	0.28892	0.38721
4	0.74616	0.28904	0.38737
5	0.74616	0.28908	0.38743

ARRET APRES L'ITERATION 5 L'ACCROISSEMENT DE L'INERTIE INTER-CLASSES
PAR RAPPORT A L'ITERATION PRECEDENTE N'EST QUE DE 0.014 %.

DECOMPOSITION DE L'INERTIE
CALCULEE SUR 10 AXES.

INERTIES	INERTIES		EFFECTIFS		POIDS		DISTANCES	
	AVANT	APRES	AVANT	APRES	AVANT	APRES	AVANT	APRES
INTER-CLASSES	0.2507	0.2891						
INTRA-CLASSE								
CLASSE 1 / 6	0.1210	0.1242	688	735	688.00	735.00	0.1428	0.1640
CLASSE 2 / 6	0.0342	0.0344	170	174	170.00	174.00	0.5654	0.5598
CLASSE 3 / 6	0.1010	0.0888	564	538	564.00	538.00	0.2049	0.2296
CLASSE 4 / 6	0.0394	0.0544	224	317	224.00	317.00	0.2041	0.2273
CLASSE 5 / 6	0.0853	0.0678	372	329	372.00	329.00	0.2427	0.2987
CLASSE 6 / 6	0.1145	0.0876	634	559	634.00	559.00	0.3451	0.4560
TOTALE	0.7462	0.7462						

QUOTIENT (INERTIE INTER / INERTIE TOTALE) : AVANT ... 0.3359
APRES ... 0.3874

COORDONNEES ET VALEURS-TEST APRES CONSOLIDATION
AXES 1 A 5

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST					COORDONNEES					DISTO.
			1	2	3	4	5	1	2	3	4	5	
Coupure 'a' de l'arbre en 6 classes													
aa1a - Classe 1 / 6	735	735.00	20.4	-26.0	-1.8	-1.9	-12.6	0.29	-0.26	-0.02	-0.01	-0.09	0.16
aa2a - Classe 2 / 6	174	174.00	11.7	-6.7	-7.2	5.2	22.9	0.38	-0.16	-0.15	0.09	0.40	0.56
aa3a - Classe 3 / 6	538	538.00	20.3	22.8	10.2	10.6	-3.8	0.35	0.28	0.11	0.10	-0.03	0.23
aa4a - Classe 4 / 6	317	317.00	-2.5	20.0	-17.3	-14.4	-0.1	-0.06	0.33	-0.26	-0.18	0.00	0.23
aa5a - Classe 5 / 6	329	329.00	-10.9	-2.4	25.1	-17.2	8.0	-0.25	-0.04	0.37	-0.22	0.10	0.30
aa6a - Classe 6 / 6	559	559.00	-38.7	-3.8	-10.2	13.7	-2.8	-0.65	-0.05	-0.11	0.13	-0.03	0.46

AXES 6 A 10

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST						COORDONNEES						DISTO.
			6	7	8	9	10	6	7	8	9	10			
Coupure 'a' de l'arbre en 6 classes															
aa1a - Classe 1 / 6	735	735.00	-5.0	-0.6	9.3	0.5	-0.5	-0.04	0.00	0.06	0.00	0.00	0.16		
aa2a - Classe 2 / 6	174	174.00	10.9	-4.2	-20.2	10.9	-9.2	0.18	-0.07	-0.33	0.17	-0.15	0.56		
aa3a - Classe 3 / 6	538	538.00	-3.2	-1.3	2.6	-2.9	6.9	-0.03	-0.01	0.02	-0.02	0.06	0.23		
aa4a - Classe 4 / 6	317	317.00	-0.6	-3.0	5.0	3.3	-4.7	-0.01	-0.04	0.06	0.04	-0.05	0.23		
aa5a - Classe 5 / 6	329	329.00	5.7	9.9	-4.7	-11.8	3.3	0.07	0.11	-0.05	-0.13	0.04	0.30		
aa6a - Classe 6 / 6	559	559.00	-2.2	-1.2	-0.7	2.6	0.4	-0.02	-0.01	-0.01	0.02	0.00	0.46		

FIG. D.7 – Classification en 6 classes sur les dix premiers axes

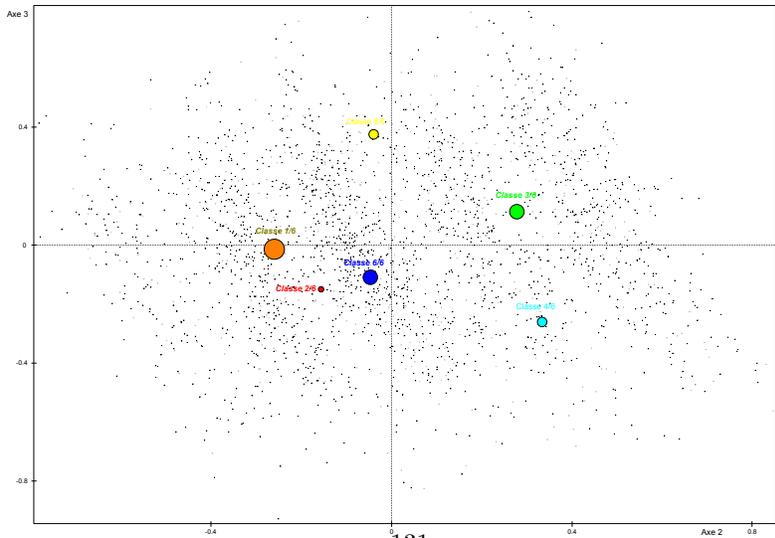
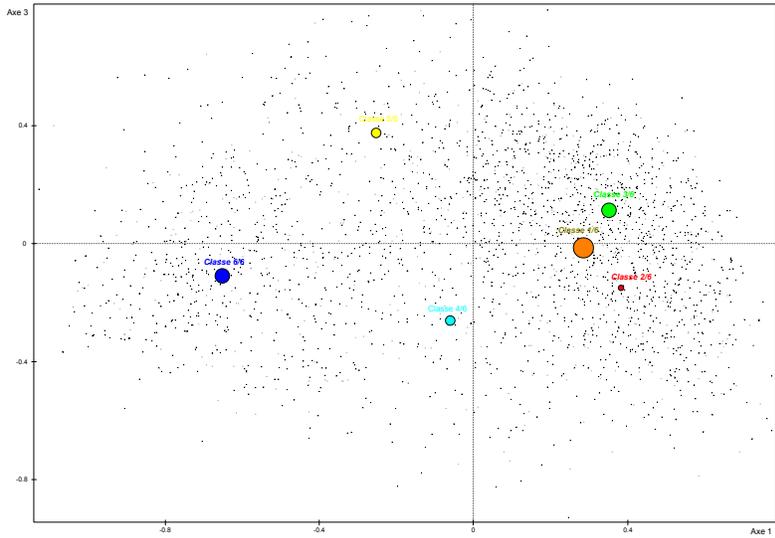
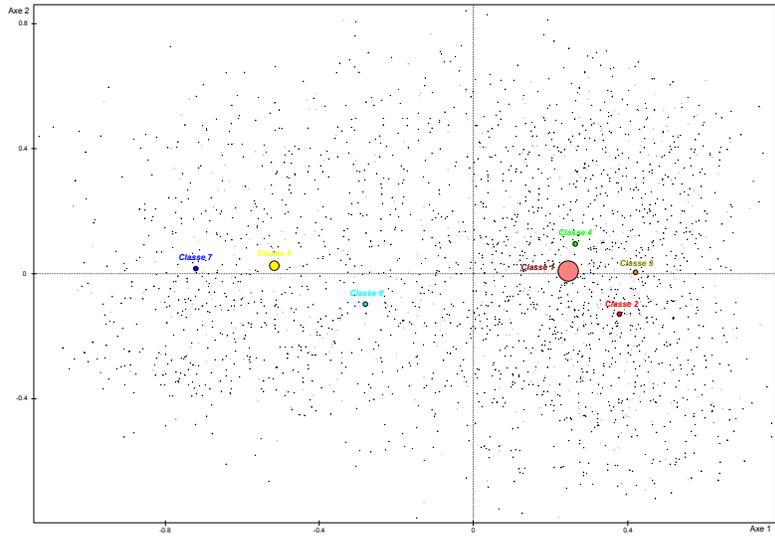


FIG. D.8 – Barycentre des 6 classes

D.2 Classification ascendante hiérarchique sur les vingt-huit premiers axes

```

CLASSIFICATION HIERARCHIQUE (VOISINS RECIPROQUES)
SUR LES 28 PREMIERS AXES FACTORIELS
DESCRIPTION DES 50 NOEUDS D'INDICES LES PLUS ELEVES
NUM. AINE BENJ EFF. POIDS INDICE HISTOGRAMME DES INDICES DE NIVEAU
5254 5239 5097 89 89.00 0.00310 **
5255 5169 5243 99 99.00 0.00312 **
5256 5250 5122 110 110.00 0.00313 **
5257 5159 5209 61 61.00 0.00318 **
5258 5198 5170 92 92.00 0.00323 **
5259 5238 5125 119 119.00 0.00324 **
5260 5168 5041 38 38.00 0.00326 **
5261 5232 5221 94 94.00 0.00331 **
5262 5203 5162 83 83.00 0.00338 **
5263 5256 5201 134 134.00 0.00342 **
5264 5242 5193 105 105.00 0.00353 ***
5265 5236 5152 58 58.00 0.00364 ***
5266 5247 5244 139 139.00 0.00366 ***
5267 5240 5262 147 147.00 0.00368 ***
5268 5223 5143 106 106.00 0.00368 ***
5269 5253 5186 149 149.00 0.00368 ***
5270 5249 5233 104 104.00 0.00374 ***
5271 5190 5225 98 98.00 0.00395 ***
5272 5264 5220 160 160.00 0.00429 ***
5273 5214 5251 96 96.00 0.00436 ***
5274 5248 5255 196 196.00 0.00436 ***
5275 5270 5245 163 163.00 0.00444 ***
5276 5228 5261 155 155.00 0.00447 ***
5277 5267 5096 176 176.00 0.00480 ***
5278 5254 5271 187 187.00 0.00497 ***
5279 5268 5173 130 130.00 0.00512 ***
5280 5224 5265 167 167.00 0.00546 ****
5281 5263 5260 172 172.00 0.00575 ****
5282 5246 5279 239 239.00 0.00580 ****
5283 5280 5147 191 191.00 0.00624 ****
5284 5274 5231 283 283.00 0.00651 ****
5285 5226 5208 80 80.00 0.00688 ****
5286 5282 5273 335 335.00 0.00709 *****
5287 5286 5235 386 386.00 0.00746 *****
5288 5285 5241 140 140.00 0.00822 *****
5289 5269 5258 241 241.00 0.00844 *****
5290 5289 5257 302 302.00 0.00877 *****
5291 5287 5288 526 526.00 0.01027 *****
5292 5284 5252 362 362.00 0.01279 *****
5293 5259 5292 481 481.00 0.01413 *****
5294 5290 5293 783 783.00 0.01727 *****
5295 5272 5294 943 943.00 0.02090 *****
5296 5266 5278 326 326.00 0.02263 *****
5297 5295 5296 1269 1269.00 0.02402 *****
5298 5276 5297 1424 1424.00 0.02934 *****
5299 5283 5291 717 717.00 0.03523 *****
5300 5281 5277 348 348.00 0.03708 *****
5301 5299 5275 880 880.00 0.04089 *****
5302 5300 5298 1772 1772.00 0.04221 *****
5303 5301 5302 2652 2652.00 0.13692 *****
SOMME DES INDICES DE NIVEAU = 1.36427

```

FIG. D.9 – Dendrogramme de la classification sur les vingt-huit premiers axes

COORDONNEES ET VALEURS-TEST
 AXES 1 A 5

ELEMENTS			VALEURS-TEST					COORDONNEES					
NUM .	IDENT	POIDS	EFF	1	2	3	4	5	1	2	3	4	5
1 .	5225	64.00	64	4.40	-7.28	4.96	-6.50	-5.66	0.54	-0.90	0.61	-0.80	-0.70
2 .	5190	34.00	34	4.77	1.64	-4.04	-2.95	-2.90	0.81	0.28	-0.69	-0.50	-0.49
3 .	5254	89.00	89	3.63	6.57	8.00	-2.90	-3.76	0.38	0.69	0.83	-0.30	-0.39
4 .	5244	79.00	79	2.07	5.65	10.80	-6.72	4.20	0.23	0.63	1.20	-0.74	0.47
5 .	5247	60.00	60	0.02	-4.55	4.41	-9.32	0.17	0.00	-0.58	0.56	-1.19	0.02
6 .	5252	79.00	79	2.01	3.87	5.67	-8.57	0.51	0.22	0.43	0.63	-0.95	0.06
7 .	5231	87.00	87	2.79	-1.35	5.05	-2.48	-11.72	0.29	-0.14	0.53	-0.26	-1.24
8 .	5243	51.00	51	-0.80	-1.44	5.29	1.07	-3.30	-0.11	-0.20	0.73	0.15	-0.46
9 .	5169	48.00	48	3.93	-3.39	2.67	-0.84	-0.83	0.56	-0.49	0.38	-0.12	-0.12
10 .	5248	97.00	97	7.69	-3.68	-0.37	4.30	-3.79	0.77	-0.37	-0.04	0.43	-0.38
11 .	5125	28.00	28	1.81	-1.38	-0.21	1.70	-7.49	0.34	-0.26	-0.04	0.32	-1.41
12 .	5238	91.00	91	7.49	-3.11	4.08	12.47	-0.88	0.77	-0.32	0.42	1.28	-0.09
13 .	5209	37.00	37	-0.42	7.32	-7.37	-2.86	-0.19	-0.07	1.20	-1.20	-0.47	-0.03
14 .	5159	24.00	24	-0.32	7.50	-8.21	-7.31	2.15	-0.06	1.52	-1.67	-1.49	0.44
15 .	5170	36.00	36	2.91	1.94	-3.73	-3.16	-6.23	0.48	0.32	-0.62	-0.52	-1.03
16 .	5198	56.00	56	7.78	2.65	-4.08	3.13	2.73	1.03	0.35	-0.54	0.41	0.36
17 .	5186	55.00	55	6.95	7.31	5.44	5.23	4.12	0.93	0.98	0.73	0.70	0.55
18 .	5253	94.00	94	5.17	12.71	0.68	-2.07	-5.76	0.52	1.29	0.07	-0.21	-0.58
19 .	5220	55.00	55	6.07	-13.04	-0.47	0.71	0.58	0.81	-1.74	-0.06	0.09	0.08
20 .	5193	27.00	27	4.17	-1.26	-9.11	-6.14	-0.33	0.80	-0.24	-1.74	-1.18	-0.06
21 .	5242	78.00	78	3.60	-9.36	-3.20	-6.56	-3.51	0.40	-1.04	-0.36	-0.73	-0.39
22 .	5221	60.00	60	3.52	0.45	-0.74	2.81	-6.28	0.45	0.06	-0.09	0.36	-0.80
23 .	5232	34.00	34	4.56	-4.87	-0.93	7.88	-1.68	0.78	-0.83	-0.16	1.34	-0.29
24 .	5228	61.00	61	6.81	8.34	0.18	7.95	2.59	0.86	1.06	0.02	1.01	0.33
25 .	5096	29.00	29	6.77	-7.00	-5.78	0.65	1.98	1.25	-1.29	-1.07	0.12	0.37
26 .	5162	24.00	24	4.94	0.25	-1.69	-0.45	-3.64	1.00	0.05	-0.34	-0.09	-0.74
27 .	5203	59.00	59	8.99	4.54	-7.01	3.15	-3.05	1.16	0.58	-0.90	0.41	-0.39
28 .	5240	64.00	64	3.88	0.70	-0.54	2.36	-1.41	0.48	0.09	-0.07	0.29	-0.17
29 .	5041	18.00	18	5.35	-2.74	-4.27	2.59	7.81	1.26	-0.64	-1.00	0.61	1.83
30 .	5168	20.00	20	5.45	0.00	-5.06	6.26	8.88	1.22	0.00	-1.13	1.39	1.98
31 .	5201	24.00	24	4.24	-3.86	-8.20	-1.74	9.75	0.86	-0.78	-1.67	-0.35	1.98
32 .	5122	21.00	21	3.13	-1.49	0.19	0.17	4.71	0.68	-0.32	0.04	0.04	1.02
33 .	5250	89.00	89	7.43	-2.90	-1.08	4.87	14.84	0.77	-0.30	-0.11	0.51	1.55
34 .	5245	59.00	59	-9.25	-0.34	6.44	-6.73	6.02	-1.19	-0.04	0.83	-0.87	0.78
35 .	5233	35.00	35	-0.91	1.84	4.17	-9.30	3.53	-0.15	0.31	0.70	-1.56	0.59
36 .	5249	69.00	69	-3.79	-7.14	14.40	-11.09	4.41	-0.45	-0.85	1.71	-1.32	0.52
37 .	5241	60.00	60	-10.08	1.58	3.43	-5.16	7.72	-1.29	0.20	0.44	-0.66	0.99
38 .	5208	48.00	48	-8.92	3.19	-5.06	-3.29	2.22	-1.28	0.46	-0.72	-0.47	0.32
39 .	5226	32.00	32	-7.22	0.23	-5.43	4.98	8.89	-1.27	0.04	-0.95	0.88	1.56
40 .	5235	51.00	51	-6.25	0.79	1.45	7.73	-2.93	-0.87	0.11	0.20	1.07	-0.41
41 .	5251	47.00	47	-6.26	5.47	-7.19	-1.48	0.65	-0.90	0.79	-1.04	-0.21	0.09
42 .	5214	49.00	49	-7.66	-1.27	1.68	2.15	-6.46	-1.08	-0.18	0.24	0.30	-0.92
43 .	5173	24.00	24	-4.86	-0.09	-3.61	5.92	-1.53	-0.99	-0.02	-0.73	1.20	-0.31
44 .	5143	31.00	31	-3.79	-3.26	-4.94	-1.30	-1.66	-0.68	-0.58	-0.88	-0.23	-0.30
45 .	5223	75.00	75	-9.67	1.39	-4.24	-0.16	4.14	-1.10	0.16	-0.48	-0.02	0.47
46 .	5246	109.00	109	-15.99	-4.74	-4.25	5.19	-5.92	-1.50	-0.44	-0.40	0.49	-0.56
47 .	5147	24.00	24	-7.13	0.99	5.82	-0.11	3.34	-1.45	0.20	1.18	-0.02	0.68
48 .	5152	23.00	23	-6.55	0.30	1.09	-0.49	4.28	-1.36	0.06	0.23	-0.10	0.89
49 .	5236	35.00	35	-6.87	5.40	-3.92	0.93	2.28	-1.15	0.91	-0.66	0.16	0.38
50 .	5224	109.00	109	-19.05	-2.71	-4.89	11.17	-3.84	-1.79	-0.25	-0.46	1.05	-0.36

FIG. D.10 – Classification sur les vingt-huit premiers axes

DESCRIPTION DES NOEUDS DE LA HIERACHIE
(INDICES EN POURCENTAGE DE LA SOMME DES INDICES : 0.61593)

NOEUD NUMERO	INDICE	SUCESSEURS		EFFECT.	POIDS	COMPOSITION	
		AINE	BENJ			PREMIER	DERNIER
51	0.51	9	8	99	99.00	8	9
52	0.51	33	32	110	110.00	32	33
53	0.52	14	13	61	61.00	13	14
54	0.52	16	15	92	92.00	15	16
55	0.53	12	11	119	119.00	11	12
56	0.53	30	29	38	38.00	29	30
57	0.54	23	22	94	94.00	22	23
58	0.55	27	26	83	83.00	26	27
59	0.56	52	31	134	134.00	31	33
60	0.57	21	20	105	105.00	20	21
61	0.59	49	48	58	58.00	48	49
62	0.59	5	4	139	139.00	4	5
63	0.60	28	58	147	147.00	26	28
64	0.60	45	44	106	106.00	44	45
65	0.60	18	17	149	149.00	17	18
66	0.61	36	35	104	104.00	35	36
67	0.64	2	1	98	98.00	1	2
68	0.70	60	19	160	160.00	19	21
69	0.71	42	41	96	96.00	41	42
70	0.71	10	51	196	196.00	8	10
71	0.72	66	34	163	163.00	34	36
72	0.73	24	57	155	155.00	22	24
73	0.78	63	25	176	176.00	25	28
74	0.81	3	67	187	187.00	1	3
75	0.83	64	43	130	130.00	43	45
76	0.89	50	61	167	167.00	48	50
77	0.93	59	56	172	172.00	29	33
78	0.94	46	75	239	239.00	43	46
79	1.01	76	47	191	191.00	47	50
80	1.06	70	7	283	283.00	7	10
81	1.12	39	38	80	80.00	38	39
82	1.15	78	69	335	335.00	41	46
83	1.21	82	40	386	386.00	40	46
84	1.33	81	37	140	140.00	37	39
85	1.37	65	54	241	241.00	15	18
86	1.42	85	53	302	302.00	13	18
87	1.67	83	84	526	526.00	37	46
88	2.08	80	6	362	362.00	6	10
89	2.29	55	88	481	481.00	6	12
90	2.80	86	89	783	783.00	6	18
91	3.39	68	90	943	943.00	6	21
92	3.67	62	74	326	326.00	1	5
93	3.90	91	92	1269	1269.00	1	21
94	4.76	72	93	1424	1424.00	1	24
95	5.72	79	87	717	717.00	37	50
96	6.02	77	73	348	348.00	25	33
97	6.64	95	71	880	880.00	34	50
98	6.85	96	94	1772	1772.00	1	33
99	22.23	97	98	2652	2652.00	1	50

FIG. D.11 – Classification sur les vingt-huit premiers axes

PARTITION PAR COUPEURE D'UN ARBRE HIERARCHIQUE
 RECHERCHE DES MEILLEURES PARTITIONS
 RECHERCHE DES PALIERS

	PALIER	VALEUR DU	
	ENTREE	PALIER	
	5298--	5299	-5.99 *****
	5297--	5298	-4.26 *****
	5294--	5295	-1.58 *****

LISTE DES 3 MEILLEURES PARTITION(S) ENTRE 3 ET 10 CLASSES
 1 - PARTITION EN 6 CLASSES
 2 - PARTITION EN 7 CLASSES
 3 - PARTITION EN 10 CLASSES

Coupeure 'b' de l'arbre en 7 classes
 FORMATION DES CLASSES (INDIVIDUS ACTIFS)
 DESCRIPTION SOMMAIRE

CLASSE	EFFECTIF	POIDS	CONTENU
bb1b	1269	1269.00	1 A 21
bb2b	155	155.00	22 A 24
bb3b	176	176.00	25 A 28
bb4b	172	172.00	29 A 33
bb5b	163	163.00	34 A 36
bb6b	526	526.00	37 A 46
bb7b	191	191.00	47 A 50

COORDONNEES ET VALEURS-TEST AVANT CONSOLIDATION
 AXES 1 A 5

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST					COORDONNEES					DISTO.
			1	2	3	4	5	1	2	3	4	5	
Coupeure 'b' de l'arbre en 7 classes													
bb1b - Classe 1 / 7	1269	1269.00	23.4	1.8	9.1	-10.4	-13.0	0.21	0.01	0.05	-0.05	-0.06	0.07
bb2b - Classe 2 / 7	155	155.00	8.8	3.3	-0.8	10.6	-3.1	0.31	0.08	-0.02	0.20	-0.06	0.55
bb3b - Classe 3 / 7	176	176.00	12.4	0.3	-7.5	3.4	-3.2	0.41	0.01	-0.16	0.06	-0.06	0.67
bb4b - Classe 4 / 7	172	172.00	11.9	-5.1	-7.1	6.0	22.0	0.39	-0.12	-0.15	0.11	0.38	0.67
bb5b - Classe 5 / 7	163	163.00	-8.6	-4.1	15.5	-15.9	8.3	-0.29	-0.10	0.34	-0.29	0.15	0.65
bb6b - Classe 6 / 7	526	526.00	-29.7	0.7	-9.3	4.8	1.0	-0.52	0.01	-0.10	0.05	0.01	0.32
bb7b - Classe 7 / 7	191	191.00	-22.6	0.8	-3.0	8.8	0.8	-0.71	0.02	-0.06	0.15	0.01	0.88

CONSOLIDATION DE LA PARTITION
 AUTOUR DES 7 CENTRES DE CLASSES, REALISEE PAR 10 ITERATIONS A CENTRES MOBILES
 PROGRESSION DE L'INERTIE INTER-CLASSES

ITERATION	I.TOTALE	I.INTER	QUOTIENT
0	1.36427	0.32167	0.23578
1	1.36427	0.34404	0.25218
2	1.36427	0.34433	0.25239
3	1.36427	0.34439	0.25244

ARRET APRES L'ITERATION 3 L'ACCROISSEMENT DE L'INERTIE INTER-CLASSES
 PAR RAPPORT A L'ITERATION PRECEDENTE N'EST QUE DE 0.017 %.
 DECOMPOSITION DE L'INERTIE
 CALCULEE SUR 28 AXES.

INERTIES	INERTIES		EFFECTIFS		POIDS		DISTANCES	
	AVANT	APRES	AVANT	APRES	AVANT	APRES	AVANT	APRES
INTER-CLASSES	0.3217	0.3444						
INTRA-CLASSE								
CLASSE 1 / 7	0.5114	0.4552	1269	1179	1269.00	1179.00	0.0706	0.0897
CLASSE 2 / 7	0.0546	0.0652	155	177	155.00	177.00	0.5534	0.5302
CLASSE 3 / 7	0.0677	0.0671	176	177	176.00	177.00	0.6684	0.6792
CLASSE 4 / 7	0.0667	0.0708	172	180	172.00	180.00	0.6722	0.6627
CLASSE 5 / 7	0.0691	0.0835	163	191	163.00	191.00	0.6465	0.6328
CLASSE 6 / 7	0.1995	0.2065	526	558	526.00	558.00	0.3244	0.3243
CLASSE 7 / 7	0.0737	0.0717	191	190	191.00	190.00	0.8821	0.9072
TOTALE	1.3643	1.3643						

QUOTIENT (INERTIE INTER / INERTIE TOTALE) : AVANT ... 0.2358
 APRES ... 0.2524

COORDONNEES ET VALEURS-TEST APRES CONSOLIDATION
 AXES 1 A 5

IDEN - LIBELLE	EFF.	P.ABS	VALEURS-TEST					COORDONNEES					DISTO.
			1	2	3	4	5	1	2	3	4	5	
Coupeure 'b' de l'arbre en 7 classes													
bb1b - Classe 1 / 7	1179	1179.00	25.3	1.0	9.2	-7.7	-12.8	0.25	0.01	0.06	-0.04	-0.07	0.09
bb2b - Classe 2 / 7	177	177.00	8.1	4.1	-1.4	10.4	-3.5	0.26	0.09	-0.03	0.18	-0.06	0.53
bb3b - Classe 3 / 7	177	177.00	12.9	0.2	-8.1	3.3	-3.6	0.42	0.00	-0.17	0.06	-0.06	0.68
bb4b - Classe 4 / 7	180	180.00	11.7	-5.7	-7.2	5.2	22.7	0.38	-0.13	-0.15	0.09	0.39	0.66
bb5b - Classe 5 / 7	191	191.00	-9.0	-4.5	16.9	-16.7	8.3	-0.28	-0.10	0.34	-0.28	0.14	0.63
bb6b - Classe 6 / 7	558	558.00	-30.7	2.0	-8.9	2.2	0.2	-0.52	0.02	-0.10	0.02	0.00	0.32
bb7b - Classe 7 / 7	190	190.00	-23.0	0.7	-4.4	9.9	0.8	-0.72	0.02	-0.09	0.17	0.01	0.91

FIG. D.12 – Classification en 7 classes sur les vingt-huit premiers axes

Coupure 'a' de l'arbre en 6 classes
FORMATION DES CLASSES (INDIVIDUS ACTIFS)
DESCRIPTION SOMMAIRE

CLASSE	EFFECTIF	POIDS	CONTENU
aa1a	1424	1424.00	1 A 24
aa2a	176	176.00	25 A 28
aa3a	172	172.00	29 A 33
aa4a	163	163.00	34 A 36
aa5a	526	526.00	37 A 46
aa6a	191	191.00	47 A 50

COORDONNEES ET VALEURS-TEST AVANT CONSOLIDATION
AXES 1 A 5

IDEN - LIBELLE	CLASSES			VALEURS-TEST					COORDONNEES					DISTO.
	EFF.	P.ABS		1	2	3	4	5	1	2	3	4	5	
Coupure 'a' de l'arbre en 6 classes														
aa1a - Classe 1 / 6	1424	1424.00	27.6	3.4	8.8	-5.4	-14.5	0.22	0.02	0.05	-0.02	-0.06	0.07	0.07
aa2a - Classe 2 / 6	176	176.00	12.4	0.3	-7.5	3.4	-3.2	0.41	0.01	-0.16	0.06	-0.06	0.67	0.67
aa3a - Classe 3 / 6	172	172.00	11.9	-5.1	-7.1	6.0	22.0	0.39	-0.12	-0.15	0.11	0.38	0.67	0.67
aa4a - Classe 4 / 6	163	163.00	-8.6	-4.1	15.5	-15.9	8.3	-0.29	-0.10	0.34	-0.29	0.15	0.65	0.65
aa5a - Classe 5 / 6	526	526.00	-29.7	0.7	-9.3	4.8	1.0	-0.52	0.01	-0.10	0.05	0.01	0.32	0.32
aa6a - Classe 6 / 6	191	191.00	-22.6	0.8	-3.0	8.8	0.8	-0.71	0.02	-0.06	0.15	0.01	0.88	0.88

CONSOLIDATION DE LA PARTITION
AUTOUR DES 6 CENTRES DE CLASSES, REALISEE PAR 10 ITERATIONS A CENTRES MOBILES
PROGRESSION DE L'INERTIE INTER-CLASSES

ITERATION	I.TOTALE	I.INTER	QUOTIENT
0	1.36427	0.29232	0.21427
1	1.36427	0.31273	0.22923
2	1.36427	0.31311	0.22951
3	1.36427	0.31321	0.22958
4	1.36427	0.31322	0.22959

DECOMPOSITION DE L'INERTIE
CALCULEE SUR 28 AXES.

INERTIES	INERTIES		EFFECTIFS		POIDS		DISTANCES	
	AVANT	APRES	AVANT	APRES	AVANT	APRES	AVANT	APRES
INTER-CLASSES	0.2923	0.3132						
INTRA-CLASSE								
CLASSE 1 / 6	0.5953	0.5329	1424	1322	1424.00	1322.00	0.0685	0.0885
CLASSE 2 / 6	0.0677	0.0671	176	177	176.00	177.00	0.6684	0.6792
CLASSE 3 / 6	0.0667	0.0714	172	181	172.00	181.00	0.6722	0.6619
CLASSE 4 / 6	0.0691	0.0851	163	193	163.00	193.00	0.6465	0.6284
CLASSE 5 / 6	0.1995	0.2228	526	589	526.00	589.00	0.3244	0.3056
CLASSE 6 / 6	0.0737	0.0717	191	190	191.00	190.00	0.8821	0.9072
TOTALE	1.3643	1.3643						

QUOTIENT (INERTIE INTER / INERTIE TOTALE) : AVANT ... 0.2143
APRES ... 0.2296

COORDONNEES ET VALEURS-TEST APRES CONSOLIDATION
AXES 1 A 5

IDEN - LIBELLE	CLASSES			VALEURS-TEST					COORDONNEES					DISTO.
	EFF.	P.ABS		1	2	3	4	5	1	2	3	4	5	
Coupure 'a' de l'arbre en 6 classes														
aa1a - Classe 1 / 6	1322	1322.00	29.7	2.8	8.4	-2.7	-14.3	0.26	0.02	0.05	-0.01	-0.07	0.09	0.09
aa2a - Classe 2 / 6	177	177.00	12.9	0.2	-8.1	3.3	-3.6	0.42	0.00	-0.17	0.06	-0.06	0.68	0.68
aa3a - Classe 3 / 6	181	181.00	11.8	-5.5	-7.1	5.2	22.8	0.38	-0.13	-0.15	0.09	0.39	0.66	0.66
aa4a - Classe 4 / 6	193	193.00	-8.9	-4.2	16.9	-16.6	8.4	-0.28	-0.09	0.34	-0.28	0.14	0.63	0.63
aa5a - Classe 5 / 6	589	589.00	-30.8	2.0	-8.8	2.4	-0.1	-0.50	0.02	-0.09	0.02	0.00	0.31	0.31
aa6a - Classe 6 / 6	190	190.00	-23.0	0.7	-4.4	9.9	0.8	-0.72	0.02	-0.09	0.17	0.01	0.91	0.91

FIG. D.13 – Classification en 6 classes sur les vingt-huit premiers axes

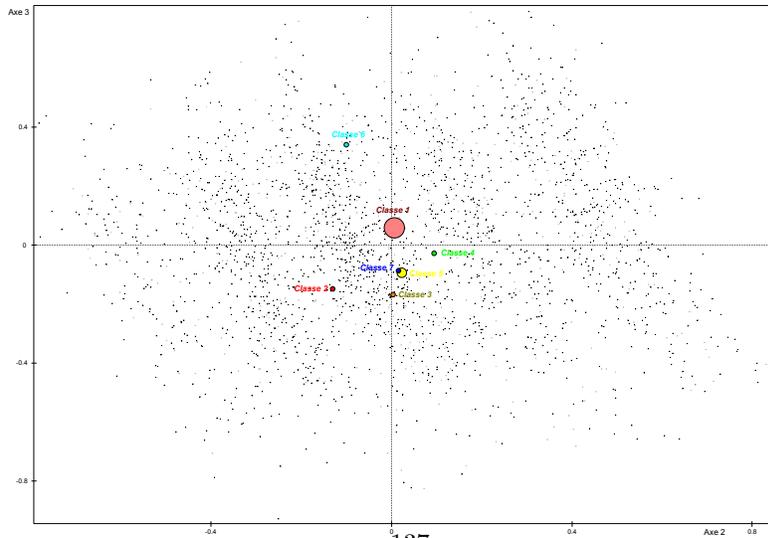
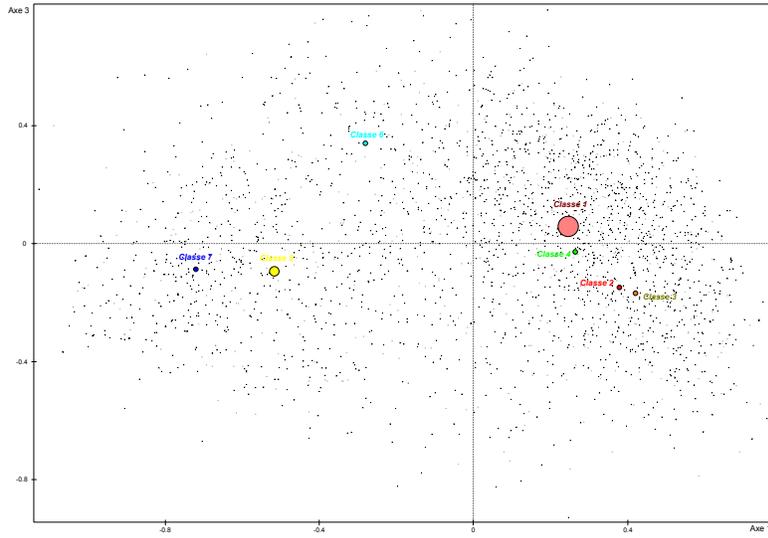
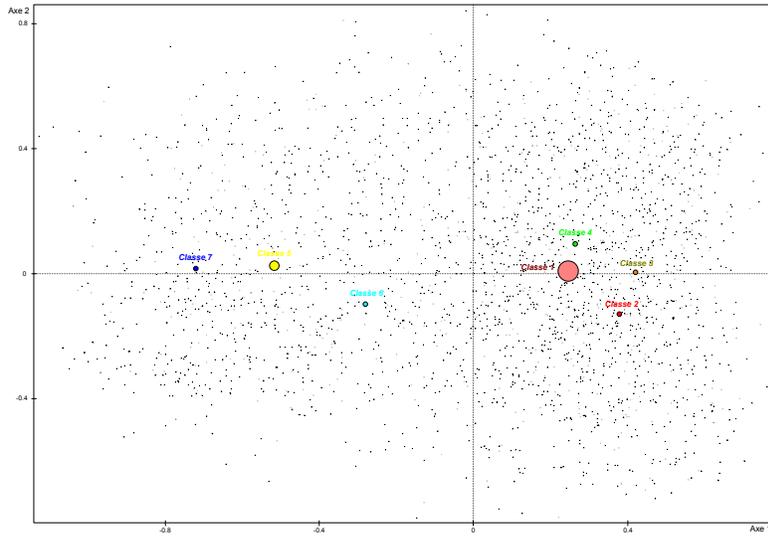


FIG. D.14 – Barycentre des 7 classes

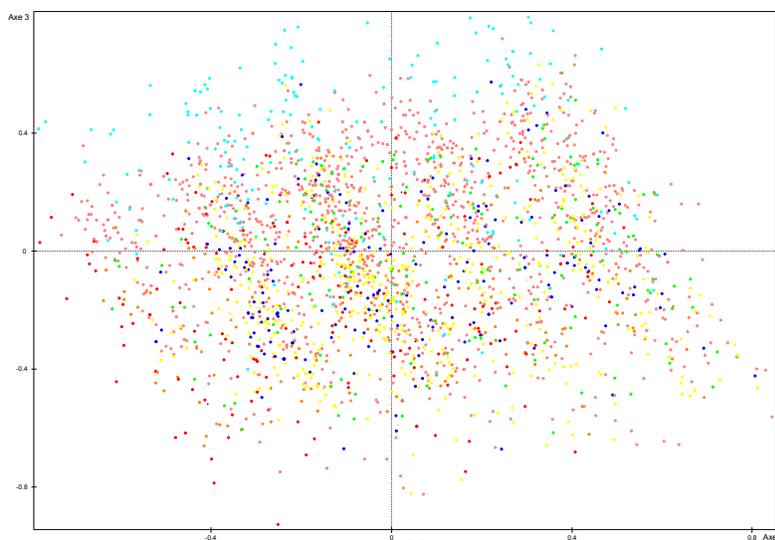
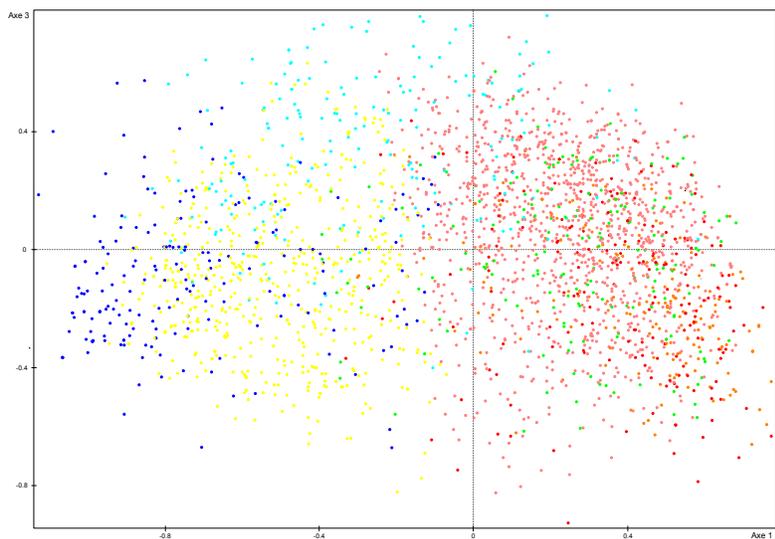
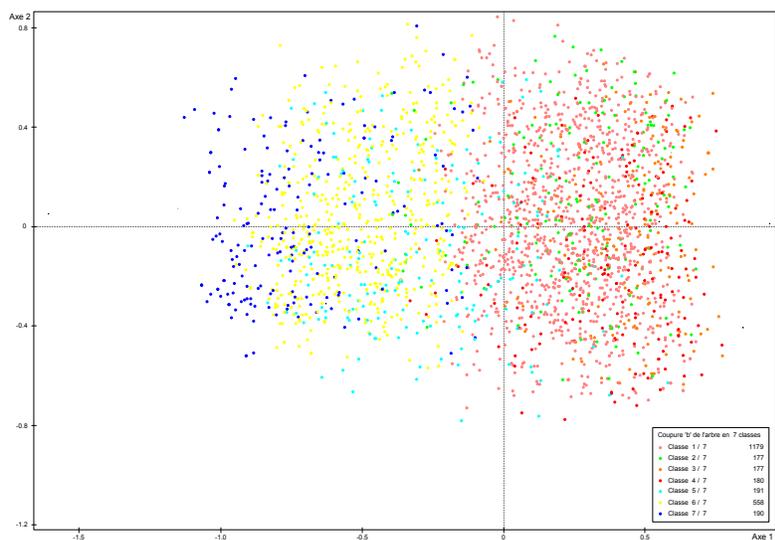


FIG. D.15 – Projection des établissements des 7 classes

Annexe E

Passage de SAS à SPAD

E.1 Méthode de transfert proposée par l'aide de SPAD

Conversion de la table SAS au format XPORT

SPAD présente une option d'importation des tables SAS. Cependant il ne reconnaît ces tables que sous le format XPORT. Le programme suivant, récupérable dans l'aide de SPAD, permet la transformation de la table SAS dans ce format de façon à permettre ensuite l'importation sous SPAD.

Supposons qu'on souhaite transférer sous SPAD la table *spadtest*. Il suffit d'indiquer nom de la nouvelle table (*nom*) précédé du chemin où on souhaite la placer, ici dans un dossier *transfert* (C:\bureau\transfert).

```
options nodate nonumber ;
/*****
/* Définition de la library SAS contenant la table TOTO , (à définir entre quotes) */
/*****
libname entree '.' ;

libname sortie XPORT 'C:\bureau\transfert\nom.xpt' ;

/*****/
```

Cette procédure a abouti à la création de 2 bibliothèques portant les noms *entree* et *sortie*. On copie alors notre table *spadtest* dans la bibliothèque *entree* (utiliser l'explorateur et faire directement un copier-coller). Il suffit alors de faire tourner le programme suivant :

```
/* Copie en format XPORT de la table spadtest dans le fichier nom.xpt */
/*****/
proc copy in= entree out= sortie ;
select spadtest;
run;
```

On obtient dans le dossier *transfert* (C:\bureau \transfert) le fichier *nom.xpt*. Attention, ce fichier est présent même si le dossier peut continuer d'être indiqué comme vide.

Importation de la table XPORT dans SPAD

- Dans SPAD, on utilise la procédure "importation d'un fichier SAS" de la barre d'outil. En cas d'existence d'un dossier spécifié par défaut pour la récupération des tables, on aura préalablement copié la table *nom.xpt* dans ce dossier.
- Dans la fenêtre qui apparaît, on indique l'intégralité du chemin conduisant à la base ou son nom si elle est a été placée dans le dossier par défaut.
- Spécifier le type des variable ("identificateur", "continue" pour la pondération, "alphabétique" pour les autres.

Problème : ce type d'importation ne permet pas de conserver les formats. Une solution alternative permet de passer outre cette contrainte.

E.2 Solution alternative

Transformation des variables en mode caractère

Sous SAS, on met les données sous forme caractère en renommant les modalités (Attention à ne pas utiliser la modalité 0 est interprétée comme une non-réponse par SPAD). On peut utiliser la démarche suivante. Dans cet exemple, on part de la variable ENFMALDD à 4 modalités numériques.

Définition des classes

```
*/ ENMALdd=0 : pas de congé  
ENFMALdd=1 : de 1 à 3 jours ENFMALdd=2 : de 4 à 6 jours ENFMALdd=3 :  
plus de 6 jours ;
```

Définition du format des modalités

```
proc format; value ENFMALdd  
    1="pas_conge_enfmal"  
    2="1_3j_enfmal"  
    3="4_6j_enfmal"  
    4=">6j_enfmal";
```

Conversion

```
data entrepr.solution; set entrepr.empl1;  
cENFMALdd=put(ENFMALdd,ENFMALdd.); run;
```

```
data entrepr.solution; set entrepr.solution; drop ENFMALdd; run;
```

```
data entrepr.solution; set entrepr.solution; rename  
cENFMALdd=ENFMALdd; run;
```

*on indique les variables que l'on souhaite conserver dans la base (identifiant, pondération etc). Dans le code suivant, on n'a retenu que la variable ENFMALdd;

```
data entrepr.solution; set entrepr.solution; keep ENFMALdd; run;
```

Exportation au format CVS (effectuée depuis SAS)

- Dans SAS, "exporter au format CVS" (procédure de la barre d'outils).
- Vérifier dans un logiciel de tableur (Excel) que les modalités s'affichent bien sous la forme caractère.

Importation sous format ASCII (effectuée depuis SPAD)

- Dans SPAD, "importer le fichier sous forme ASCII" (procédure de la barre d'outils). En cas d'existence d'un dossier spécifié par défaut pour la récupération des tables, on aura préalablement copié la table au format CVS dans ce fichier.
- Dans la fenêtre qui apparaît, on indique l'intégralité du chemin conduisant à la base ou simplement son nom si elle a été placée dans le dossier par défaut.
- Spécifier la forme du séparateur entre les colonnes (en général, la virgule)
- Spécifier le type des variables ("identificateur", "continue" pour la pondération, "alphabétique" pour les autres).

Documents de Travail

Ces fascicules vous seront adressés sur simple demande à l'auteur :
Institut national d'études démographiques, 133, bd Davout, 75980 PARIS Cedex 20
Tél : (33) 01 56 06 20 86 Fax : (33) 01 56 06 21 99

- N° 161. – Matthieu SOLIGNAC, *Les politiques de conciliation vie professionnelle/vie familiale menées par les employeurs : élaboration d'une typologie des établissements de l'Enquête Familles-Employeurs*, 2009, 143 p.
- N° 160. – Géraldine DUTHÉ, Raphaël LAURENT, Gilles PISON, *Vivre et mourir après 60 ans en milieu rural africain. Isolement, recours aux soins et mortalité des personnes âgées à Mlomp*, 2009, 26 p.
- N° 159. – Nathalie DONZEAU et Jean-Louis PAN KÉ SON, *La mobilité résidentielle depuis la fin des Trente Glorieuses*, 2009, 34 p.
- N° 158. – Olivier THÉVENON, *The costs of raising children and the effectiveness of policies to support parenthood in European countries: a Literature Review* 2009, 612 p.
- N° 157. – Jean-Louis PAN KÉ SON, *L'émergence du sentiment d'insécurité en quartiers défavorisés. Dépassement du seuil de tolérance... aux étrangers ou à la misère ?*, 2009, 20 p.
- N° 156. – Maryse Marpsat, *The Ined Research on Homelessness, 1993-2008*, 2008, 218 p.
- N° 155. – Éva BEAUJOUAN, Anne SOLAZ, *Childbearing after separation: Do second unions make up for earlier missing births? Evidence from France*, 2008, 24 p.
- N° 154. – Carole BONNET, Laurent GOBILLON, Anne LAFERRÈRE, *The effect of widowhood on housing and location choices*, 2008, 40 p.
- N° 153. – Louise MARIE DIOP-MAES, *La population ancienne de l'Afrique subsaharienne. Les éléments d'évaluation*, 2008, 20 p.
- N° 152. – Traduction en Russe du N° 121.
- N° 151. – P. FESTY, J. ACCARDO, D. DEMAILLY, L. PROKOFIEVA, I. KORTCHAGINA, A. SZUKIELOJC-BIENKUNSKA, L. NIVOROZHKINA, L. OVTCHAROVA, M. SEBTI, A. PATERNO, S. STROZZA, I. ELISEEVA, A. SHEVYAKOV, *Mesures, formes et facteurs de la pauvreté. Approches comparative*, 2008, 196 p.
- N° 150. – Géraldine DUTHÉ, Serge H. D. FAYE, Emmanuelle GUYAVARCH, Pascal ARDUIN, Malick A. KANTE, Aldiouma DIALLO, Raphaël LAURENT, Adama MARRA, Gilles PISON, *La détermination des causes de décès par autopsie verbale : étude de la mortalité palustre en zone rurale sénégalaise*, 2008, 42 p.
- N° 149. – Maryse MARPSAT, *Services for the Homeless in France. Description, official statistics, client recording of information. A report for the European Commission*, 2007, 84 p.
- N° 148. – Olivier THÉVENON, *L'activité féminine après l'arrivée d'enfants : disparités et évolutions en Europe à partir des enquêtes sur les Forces de travail, 1992-2005*, 2007, 56 p.
- N° 147. – Magali BARBIERI, *Population en transition. Dix communications présentées au XXV^e Congrès général de la population, Tours, France, 18-23 juillet 2005*, 2007, 201 p.
- N° 146. – François CHAPIREAU, *La mortalité des malades mentaux hospitalisés en France pendant la deuxième guerre mondiale*, 2007, 36 p.
- N° 145. – Maryse MARPSAT, *Explorer les frontières. Recherches sur des catégories « en marge »*, Mémoire présenté en vue de l'habilitation à diriger des recherches en sociologie, 2007, 274 p.
- N° 144. – Arnaud RÉGNIER-LOILIER et Pascal SEBILLE, *Modifications to the Generations and Gender Surveys questionnaire in France (wave 1)*, 192 p.
- N° 143. – Ariane PAILHÉ et Anne SOLAZ, *L'enquête Familles et employeurs. Protocole d'une double enquête et bilan de collecte*, 180 p.

- N° 142. – Annie BACHELOT et Jacques de MOUZON, *Données de l'enquête « Caractéristiques des couples demandant une fécondation in vitro en France »*, 2007, 44 p.
- N° 141. –Olivia EKERT-JAFFE, Shoshana GROSSBARD et Rémi MOUGIN, *Economic Analysis of the Childbearing Decision*, 2007, 108 p.
- N° 140. – Véronique HERTRICH and Marie LESCLINGAND, *Transition to adulthood and gender: changes in rural Mali*
- N° 139. – Patrick SIMON et Martin CLÉMENT, *Rapport de l'enquête « Mesure de la diversité ». Une enquête expérimentale pour caractériser l'origine*, 2006, 86 p.
- N° 138. – Magali BARBIERI, Alfred NIZARD et Laurent TOULEMON, *Écart de température et mortalité en France*, 2006, 80 p.
- N° 137. – Jean-Louis PAN KE SHON, *Mobilités internes différentielles en quartiers sensibles et ségrégation*, 2006, 42 p.
- N° 136. – Francisco MUNOZ-PEREZ, Sophie PENNEC, avec la collaboration de Geneviève Houriet Segard, *Évolution future de la population des magistrats et perspectives de carrière, 2001-2040*, 2006, XXX + 114 p.
- N° 135. – Alexandre DJIRIKIAN et Valérie LAFLAMME, sous la direction de Maryse MARPSAT, *Les formes marginales de logement. Étude bibliographique et méthodologique de la prise en compte du logement non ordinaire*, 2006, 240 p.
- N° 134. – Catherine BONVALET et Éva LELIÈVRE, *Publications choisies autour de l'enquête « Biographies et entourage »*, 2006, 134 p.
- N° 133. – Arnaud RÉGNIER-LOILIER, *Présentation, questionnaire et documentation de l'« Étude des relations familiales et intergénérationnelles » (Erfi). Version française de l'enquête « Generations and Gender Survey » (GGS)*, 2006, 238 p.
- N° 132. – Lucie BONNET et Louis BERTRAND (sous la direction de), *Mobilités, habitat et identités*, Actes de la journée d'étude « Jeunes chercheurs ». Le logement et l'habitat comme objet de recherche. Atelier 3, 2005, 92 p.
- N° 131. – Isabelle FRECHON et Catherine Villeneuve-Gokalp, *Étude sur l'adoption*, 2005, 64 p.
- N° 130. – Dominique MEURS, Ariane PAIHLÉ et Patrick SIMON, *Mobilité intergénérationnelle et persistance des inégalités. L'accès à l'emploi des immigrés et de leurs descendants en France*, 2005, 36 p.
- N° 129. – Magali MAZUY, Nicolas RAZAFINDRATSIMA, Élise de LA ROCHEBROCHARD, *Dépêrdition dans l'enquête« Intentions de fécondité »*, 2005, 36 p.
- N° 128. – Laure MOGUEROU et Magali BARBIERI, *Population et pauvreté en Afrique. Neuf communications présentées à la IV^e Conférence africaine sur la population*, Tunis, Tunisie, 8-12 décembre 2003, 2005, 184 p.
- N° 127. – Jean-Louis PAN KÉ SHON, *Les sources de la mobilité résidentielle. Modifications intervenues sur les grandes sources de données dans l'étude des migrations*, 2005, 30 p.
- N° 126. – Thierry DEBRAND et Anne-Gisèle PRIVAT, *L'impact des réformes de 1993 et de 2003 sur les retraites. Une analyse à l'aide du modèle de microsimulation Artémis*, 2005, 28 p.
- N° 125. – Kees WAALDIJK (ed), *More or less together: levels of legal consequences of marriage, cohabitation and registered partnership for different-sex and same-sex partners: a comparative study of nine European countries*, 2005, 192 p. (s'adresser à Marie DIGOIX)
- N° 124. – Marie DIGOIX et Patrick FESTY (eds), *Same-sex couples, same-sex partnerships, and homosexual marriages: A Focus on cross-national differentials*, 2004, 304 p.
- N° 123. – Marie DIGOIX et Patrick FESTY (sous la dir.), *Séminaire « Comparaisons européennes », années 2001-2002*, 2004, 220 p.
- N° 122. – Emmanuelle GUYAVARCH et Gilles PISON, *Les balbutiements de la contraception en Afrique au Sud du Sahara*, septembre 2004, 48 p.
- N° 121. – Maryse JASPARD et Stéphanie CONDON, *Genre, violences sexuelles et justice*. Actes de la journée-séminaire du 20 juin 2003, 2004, 135p.

- N° 120. – Laurent TOULEMON et Magali MAZUY, *Comment prendre en compte l'âge à l'arrivée et la durée de séjour en France dans la mesure de la fécondité des immigrants ?*, 2004, 34 p.
- N° 119. – Céline CLÉMENT et Bénédicte GASTINEAU (coord.), *Démographie et sociétés. Colloque international « Jeunes Chercheurs »*, Cerpos-Université Paris X-Nanterre, 1^{er} et 2 octobre 2002, 2003, 350 p.
- N° 118. – Monique BERTRAND, Véronique DUPONT et France GUERIN-PACE (sous la dir.), *Espaces de vie. Une revue des concepts et des applications*, 2003, 188 p.
- N° 117. – Stephanie CONDON et Armelle ANDRO, *Questions de genre en démographie. Actes de la journée du 22 juin 2001*, 2003, 128 p.
- N° 116. – Maryse JASPARD et l'équipe Enveff, *Le questionnaire de l'enquête Enveff. Enquête nationale sur les violences envers les femmes en France*, 2003, 10 + 88 p.
- N° 115. – Zahia OUADAH-BEDIDI et Jacques VALLIN, *Disparités régionales de l'écart d'âge entre conjoints en Algérie. Évolution depuis 1966*, 2003, 32 p.
- N° 114. – Magali MAZUY, *Situations familiales et fécondité selon le milieu social. Résultats à partir de l'enquête EHF de 1999*, 2002, 60 p.
- N° 113. – Jean-Paul SARDON, *Fécondité et transition en Europe centrale et orientale*, 2002, 38 p.
- N° 112. – Thérèse LOCOH, *Deux études sur la fécondité en Afrique : 1) Structures familiales et évolutions de la fécondité dans les pays à fécondité intermédiaire d'Afrique de l'Ouest ; 2) Baisse de la fécondité et mutations familiales en Afrique sub-saharienne*, 2002, 24 p. et 30 p.
- N° 111. – Thierry DEBRAND et Anne-Gisèle PRIVAT, *Individual real wages over business cycle: The impact of macroeconomic variations on individual careers and implications concerning retirement pensions*, 2002, 38 p.
- N° 110. – Recueil préparé par Amandine LEBUGLE et Jacques VALLIN, *Sur le chemin de la transition. Onze communications présentées au XXIV^e Congrès général de la population à Salvador de Bahia, Brésil, août 2001*, 2002, 234 p.
- N° 109. – Éric BRIAN, Jean-Marc ROHRBASSER, Christine THÉRÉ, Jacques VÉRON (intervenants et organisateurs), *La durée de vie : histoire et calcul. Séminaire de la valorisation de la recherche*, 7 février 2000, 2002, 70 p.
- N° 108. – France MESLÉ et Jacques VALLIN, *Montée de l'espérance de vie et concentration des âges au décès*, 2002, 20 p.
- N° 107. – Alexandre AVDEEV, *La mortalité infantile en Russie et en URSS: éléments pour un état des recherches*, 2002, 48 p.
- N° 106. – Isabelle ATTANÉ (organisatrice), *La Chine en transition : questions de population, questions de société. Séminaire de la valorisation de la recherche*, 31 janvier et 1^{er} février 2001 (s'adresser à Céline PERREL), 2002, 46 p.
- N° 105. – A. AVDEEV, J. BELLENGER, A. BLUM, P. FESTY, A. PAILHE, C. GOUSSEFF, C. LEFÈVRE, A. MONNIER, J.-C. SEBAG, J. VALLIN (intervenants et organisateurs), *La société russe depuis la perestroïka : rupture, crise ou continuité? Séminaire de la valorisation de la recherche*, 1^{er} mars 2001 (s'adresser à Céline PERREL), 2001, 124 p.
- N° 104. – Jacques VÉRON, Sophie PENNEC, Jacques LÉGARÉ, Marie DIGOIX (éds), *Le contrat social à l'épreuve des changements démographiques ~ The Social Contract in the Face of Demographic Change*, Actes des 2^e Rencontres Sauvy, 2001, 386 p.
- N° 103. – Gilles PISON, Alexis GABADINHO, Catherine ENEL, *Mlomp (Sénégal). Niveaux et tendances démographiques; 1985-2000*, 2001, 182 p.
- N° 102. – *La famille en AOF et la condition de la femme. Rapport présenté au Gouverneur général de l'AOF. par Denise SAVINEAU (1938). Introduction de Pascale Barthélémy*, 2001, XXII-222 p.
- N° 101. – Jean-Paul SARDON, *La fécondité dans les Balkans*, 2001, 88 p.
- N° 100. – Jean-Paul SARDON, *L'évolution récente de la fécondité en Europe du Sud*, 26 p.

- N° 99.– S. JUSTEAU, J.H. KALTENBACH, D. LAPEYRONNIE, S. ROCHÉ, J.C. SEBAG, X. THIERRY ET M. TRIBALAT (intervenants et organisateurs), *L'immigration et ses amalgames*. Séminaire de la valorisation de la recherche, 24 mai 2000, 2001, 94 p.
- N° 98.– Juliette HALIFAX, *L'insertion sociale des enfants adoptés. Résultats de l'enquête « Adoption internationale et insertion sociale », 2000 (Ined – Les Amis des enfants du monde)*, 2001, 58 p.
- N° 97.– Michèle TRIBALAT, *Modéliser, pour quoi faire?*, 2001, 10 p.
- N° 96.– O. EKERT-JAFFÉ, H. LERIDON, S. PENNEC, I. THÉRY, L. TOULEMON et J.-C. SEBAG (intervenants et organisateurs), *Évolution de la structure familiale*. Séminaire de la valorisation de la recherche, 28 juin 2000, 2001, 110 p.
- N° 95.– A. ANDRO, A. LEBUGLE, M. LESCLINGAND, T. LOCOH, M. MOUVAGHA-SOW, Z. OUADAH-BEDIDI, J. VALLIN, C. VANDERMEERSCH, J. VÉRON, *Genre et développement. Huit communications présentées à la Chaire Quetelet 2000*, 2001, 158 p.
- N° 94.– C. BONVALET, C. CLÉMENT, D. MAISON, L. ORTALDA et T. VICHNEVSKAIA, *Réseaux de sociabilité et d'entraide au sein de la parenté : Six contributions*, 2001, 110 p.
- N° 93.– Magali MAZUY et Laurent TOULEMON, *Étude de l'histoire familiale. Premiers résultats de l'enquête en ménages*, 2001, 100 p.
- N° 92.– *Politiques sociales en France et en Russie*, INED/IPSEP, 2001, 246 p.
- N° 91.– Françoise MOREAU, *Commerce des données sur la population et libertés individuelles*, 2001, 20 p. + Annexes.
- N° 90.– Youssef COURBAGE, Sergio DELLAPERGOLA, Alain DIECKHOFF, Philippe FARGUES, Emile MALET, Elias SANBAR et Jean-Claude SEBAG (intervenants et organisateurs), *L'arrière-plan démographique de l'explosion de violence en Israël-Palestine*. Séminaire de la valorisation de la recherche, 30 novembre 2000, 2000, 106 p.
- N° 89.– Bénédicte GASTINEAU et Elisabete de CARVALHO (coordonné par), *Démographie: nouveaux champs, nouvelles recherches*, 2000, 380 p.
- N° 88.– Gil BELLIS, Jean-Noël BIRABEN, Marie-Hélène CAZES et Marc de BRAEKELEER (modérateur et intervenants), *Génétique et populations*. Séminaire de la valorisation de la recherche, 26 janvier 2000, 2000, 96 p.
- N° 87.– Jean-Marie FIRDION, Maryse MARPSAT et Gérard MAUGER (intervenants), *Étude des sans-domicile: le cas de Paris et de l'Ile-de-France*. Séminaire de la valorisation de la recherche, 19 avril 2000, 2000, 90 p.
- N° 86.– François HÉRAN et Jean-Claude SEBAG (responsables modérateurs), *L'utilisation des sources administratives en démographie, sociologie et statistique sociale*. Séminaire de la valorisation de la recherche, 20 septembre 2000, 2000, 170 p.
- N° 85.– Michel BOZON et Thérèse LOCOH (sous la dir.), *Rapports de genre et questions de population. II. Genre, population et développement*, 2000, 200 p.
- N° 84.– Michel BOZON et Thérèse LOCOH (sous la dir.), *Rapports de genre et questions de population. I. Genre et population, France 2000*, 2000, 260 p.
- N° 83.– Stéphanie CONDON, Michel BOZON et Thérèse LOCOH, *Démographie, sexe et genre: bilan et perspectives*, 2000, 100 p.
- N° 82.– Olivia EKERT-JAFFE et Anne SOLAZ, *Unemployment and family formation in France*, 2000, 26 p.
- N° 81.– Jean-Marie FIRDION, *L'étude des jeunes sans domicile dans les pays occidentaux : état des lieux*, 1999, 28 p.
- N° 80.– *Age, génération et activité : vers un nouveau contrat social ? / Age, cohort and activity: A new "social contract"?*, Actes des 1^{ères} rencontres Sauvy (s'adresser à Marie DIGOIX), 1999, 314 p.
- N° 79.– Maryse MARPSAT, *Les apports réciproques des méthodes quantitatives et qualitatives : le cas particulier des enquêtes sur les personnes sans domicile*, 1999, 24 p.

- N° 78.– *Les populations du monde, le monde des populations. La place de l'expert en sciences sociales dans le débat public*, Actes de la Table ronde pour l'inauguration de l'Ined, 1999, 54 p.
- N° 77.– Isabelle SÉGUY, Fabienne LE SAGER, *Enquête Louis Henry. Notice descriptive des données informatiques*, 1999, 156 p.
- N° 76.– I. SÉGUY, H. COLENÇON et C. MÉRIC, *Enquête Louis Henry. Notice descriptive de la partie nominative*, 1999, 120 p.
- N° 75.– Anne-Claude LE VOYER (s'adresser à H. LERIDON), *Les processus menant au désir d'enfant en France*, 1999, 200 p.
- N° 74.– Jacques VALLIN et France MESLÉ, *Le rôle des vaccinations dans la baisse de la mortalité*, 1999, 20 p.
- N° 73.– Bernard ZARCA, *Comment passer d'un échantillon de ménages à un échantillon de fratries ? Les enquêtes «Réseaux familiaux» de 1976, «Proches et parents» de 1990 et le calcul d'un coefficient de pondération*, 1999, 20 p.
- N° 72.– Catherine BONVALET, *Famille-logement. Identité statistique ou enjeu politique?* 1998, 262 p.
- N° 71.– Denise ARBONVILLE, *Normalisation de l'habitat et accès au logement. Une étude statistique de l'évolution du parc "social de fait" de 1984 à 1992*, 1998, 36 p.
- N° 70.– *Famille, activité, vieillissement : générations et solidarités*. Bibliographie préparée par le Centre de Documentation de l'Ined, 1998, 44 p.
- N° 69.– XXIII^e Congrès général de la population, Beijing, Chine, 11-17 octobre 1997:
 A) *Contribution des chercheurs de l'Ined au Congrès*, 1997, 178 p.
 B) *Participation of Ined Researchers in the Conference*, 1997, 180 p.
- N° 68.– France MESLÉ et Jacques VALLIN, *Évolution de la mortalité aux âges élevés en France depuis 1950*, 1998, 42 p.
- N° 67.– Isabelle SEGUY, *Enquête Jean-Noël Biraben «La population de la France de 1500 à 1700». Répertoire des sources numériques*, 1998, 36 p.
- N° 66.– Alain BLUM, *I. Statistique, démographie et politique. II. Deux études sur l'histoire de la statistique et de la statistique démographique en URSS (1920-1939)*, 1998, 92 p.
- N° 65.– Annie LABOURIE-RACAPÉ et Thérèse LOCOH, *Genre et démographie : nouvelles problématiques ou effet de mode ?* 1998, 27 p.
- N° 64.– C. BONVALET, A. GOTMAN et Y. GRAFMEYER (éds), et I. Bertaux-Viame, D. Maison et L. Ortalda, *Proches et parents : l'aménagement des territoires*, 1997.
- N° 63.– Corinne BENVENISTE et Benoît RIANDEY, *Les exclus du logement : connaître et agir*, 1997, 20 p.
- N° 62.– Sylvia T. WARGON, *La démographie au Canada, 1945-1995*, 1997, 40 p.
- N° 61.– Claude RENARD, *Enquête Louis Henry. Bibliographie de l'enquête*, 1997, 82 p.
- N° 60.– H. AGHA, J.C. CHASTELAND, Y. COURBAGE, M. LADIER-FOULADI, A.H. MEHRYAR, *Famille et fécondité à Shiraz (1996)*, 1997, 60 p.
- N° 59.– Catherine BONVALET, Dominique MAISON et Laurent ORTALDA, *Analyse textuelle des entretiens «Proches et Parents»*, 1997, 32 p.
- N° 58.– B. BACCAÏNI, M. BARBIERI, S. CONDON et M. DIGOIX (éds),
Questions de population. Actes du Colloque Jeunes Chercheurs:
 I. *Mesures démographiques dans des petites populations*, 1997, 50 p.
 II. *Nuptialité – fécondité – reproduction*, 1997, 120 p.
 III. *Histoire des populations*, 1997, 90 p.
 IV. *Économie et emploi*, 1997, 50 p.
 V. *Vieillesse – retraite*, 1997, 66 p.
 VI. *Famille*, 1997, 128 p.
 VII. *Santé – mortalité*, 1997, 136 p.
 VIII. *Population et espace*, 1997, 120 p.
 IX. *Migration – intégration*, 1997, 96 p.

- N° 57.– Isabelle SÉGUY et Corinne MÉRIC, *Enquête Louis Henry. Notice descriptive non nominative*, 1997, 106 p.
- N° 56.– Máire Ní BHROLCHÁIN and Laurent TOULEMON, *Exploratory analysis of demographic data using graphical methods*, 1996, 50 p.
- N° 55.– Laurent TOULEMON et Catherine de GUIBERT-LANTOINE, *Enquêtes sur la fécondité et la famille dans les pays de l'Europe (régions ECE des Nations unies). Résultats de l'enquête française*, 1996, 84 p.
- N° 54.– G. BALLAND, G. BELLIS, M. DE BRAEKELEER, F. DEPOID, M. LEFEBVRE, I. SEGUY, *Généalogies et reconstitutions de familles. Analyse des besoins*, 1996, 44 p.
- N° 53.– Jacques VALLIN et France MESLÉ, *Comment suivre l'évolution de la mortalité par cause malgré les discontinuités de la statistique ? Le cas de la France de 1925 à 1993*, 1996, 46p .
- N° 52.– Catherine BONVALET et Eva LELIÈVRE, *La notion d'entourage, un outil pour l'analyse de l'évolution des réseaux individuels*, 1996, 18 p.
- N° 51.– Alexandre AVDEEV, Alain BLUM et Serge ZAKHAROV, *La mortalité a-t-elle vraiment augmenté brutalement entre 1991 et 1995?*, 1996, 80 p.
- N° 50.– France MESLÉ, Vladimir SHKOLNIKOV, Véronique HERTRICH et Jacques VALLIN, *Tendances récentes de la mortalité par cause en Russie, 1965-1993*, 1995, 70 p.
Avec, en supplément, 1 volume d'Annexes de 384 p.
- N° 49.– Jacques VALLIN, *Espérance de vie : quelle quantité pour quelle qualité de vie ?*, 1995, 24 p.
- N° 48.– François HÉRAN, *Figures et légendes de la parenté:*
I. *Variations sur les figures élémentaires*, 1995, 114 p.
II. *La modélisation de l'écart d'âge et la relation groupe/individu*, 1995, 84 p.
III. *Trois études de cas sur l'écart d'âge: Touaregs, Alyawara, Warlpiri*, 1995, 102 p.
IV. *Le roulement des alliances*, 1995, 60 p.
V. *Petite géométrie fractale de la parenté*, 1995, 42 p.
VI. *Arbor juris. Logique des figures de parenté au Moyen Age*, 1996, 62 p.
VII. *De Granet à Lévi-Strauss*, 1996, 162 p.
VIII. *Les vies parallèles. Une analyse de la co-alliance chez les Etoro de Nouvelle-Guinée*, 1996, 80 p.
IX. *Ambrym ou l'énigme de la symétrie oblique : histoire d'une controverse*, 1996, 136 p.
- N° 47.– Olivia EKERT-JAFFÉ, Denise ARBONVILLE et Jérôme WITTEWER, *Ce que coûtent les jeunes de 18 à 25 ans*, 1995, 122 p.
- N° 46.– Laurent TOULEMON, *Régression logistique et régression sur les risques. Deux supports de cours*, 1995, 56 p.
- N° 45.– Graziella CASELLI, France MESLÉ et Jacques VALLIN, *Le triomphe de la médecine. Évolution de la mortalité en Europe depuis le début de siècle*, 1995, 60 p.
- N° 44.– Magali BARBIERI, Alain BLUM, Elena DOLGIKH, Amon ERGASHEV, *La transition de fécondité en Ouzbékistan*, 1994, 76 p.
- N° 43.– Marc De BRAEKELEER et Gil BELLIS, *Généalogies et reconstitutions de familles en génétique humaine*, 1994, 66 p.
- N° 42.– Serge ADAMETS, Alain BLUM et Serge ZAKHAROV, *Disparités et variabilités des catastrophes démographiques en URSS*, 1994, 100 p.
- N° 41.– Alexandre AVDEEV, Alain BLUM et Irina TROITSKAJA, *L'avortement et la contraception en Russie et dans l'ex-URSS : histoire et présent*, 1993, 74 p.
- N° 40.– Gilles PISON et Annabel DESGREES DU LOU, *Bandafassi (Sénégal) : niveaux et tendances démographiques 1971-1991*, 1993, 40 p.
- N° 39.– Michel Louis LÉVY, *La dynamique des populations humaines*, 1993, 20 p.
- N° 38.– Alain BLUM, *Systèmes démographiques soviétiques*, 1992, 14 + X p.

- N° 37.– Emmanuel LAGARDE, Gilles PISON, Bernard LE GUENNO, Catherine ENEL et Cheikh SECK, *Les facteurs de risque de l'infection à VIH2 dans une région rurale du Sénégal*, 1992, 72 p.
- N° 36.– Annabel DESGREES DU LOU et Gilles PISON, *Les obstacles à la vaccination universelle des enfants des pays en développement. Une étude de cas en zone rurale au Sénégal*, 1992, 26 p.
- N° 35.– France MESLÉ, Vladimir SHKOLNIKOV et Jacques VALLIN, *La mortalité par causes en URSS de 1970 à 1987 : reconstruction de séries statistiques cohérentes*, 1992, 36 p.
- N° 34.– France MESLÉ et Jacques VALLIN, *Évolution de la mortalité par cancer et par maladies cardio-vasculaires en Europe depuis 1950*, 1992, 48 p.
- N° 33.– Didier BLANCHET, *Vieillesse et perspectives des retraites : analyses démographiques*, 1991, 120 p.
- N° 32.– Noël BONNEUIL, *Démographie de la nuptialité au XIX^e siècle*, 1990, 32 p.
- N° 31.– Jean-Paul SARDON, *L'évolution de la fécondité en France depuis un demi-siècle*, 1990, 102 p.
- N° 30.– Benoît RIANDEY, *Répertoire des enquêtes démographiques : bilan pour la France métropolitaine*, 1989, 24 p.
- N° 29.– Thérèse LOCOH, *Changement social et situations matrimoniales : les nouvelles formes d'union à Lomé*, 1989, 44 p.
- N° 28.– Catherine ENEL, Gilles PISON, et Monique LEFEBVRE, *Migrations et évolution de la nuptialité. L'exemple d'un village joola du sud du Sénégal, Mlomp*, 1989, 26 p. (Sénégal) depuis 50 ans, 1^{ère} édition : 1989, 36 p. ; 2^{ème} édition revue et augmentée : 1990, 48 p.
- N° 27.– Nicolas BROUARD, *L'extinction des noms de famille en France : une approche*, 1989, 22 p.
- N° 26.– Gilles PISON, Monique LEFEBVRE, Catherine ENEL et Jean-François TRAPE, *L'influence des changements sanitaires sur l'évolution de la mortalité : le cas de Mlomp*, 1989, 36 p.
- N° 25.– Alain BLUM et Philippe FARGUES, *Estimation de la mortalité maternelle dans les pays à données incomplètes. Une application à Bamako (1974-1985) et à d'autres pays en développement*, 1989, 36 p.
- N° 24.– Jacques VALLIN et Graziella CASELLI, *Mortalité et vieillissement de la population*, 1989, 30 p.
- N° 23.– Georges TAPINOS, Didier BLANCHET et Olivia EKERT-JAFFÉ, *Population et demande de changements démographiques, demande et structure de consommation*, 1989, 46 p.
- N° 22.– Benoît RIANDEY, *Un échantillon probabiliste de A à Z : l'exemple de l'enquête Peuplement et dépeuplement de Paris. INED (1986)*, 1989, 12 p.
- N° 21.– Noël BONNEUIL et Philippe FARGUES, *Prévoir les « caprices » de la mortalité. Chronique des causes de décès à Bamako de 1964 à 1985*, 1989, 44 p.
- N° 20.– France MESLÉ, *Morbidité et causes de décès chez les personnes âgées*, 1988, 18 p.
- N° 19.– Henri LERIDON, *Analyse des biographies matrimoniales dans l'enquête sur les situations familiales*, 1988, 64 p.
- N° 18.– Jacques VALLIN, *La mortalité en Europe de 1720 à 1914 : tendances à long terme et changements de structure par âge et par sexe*, 1988, 40 p.
- N° 17.– Jacques VALLIN, *Évolution sociale et baisse de la mortalité : conquête ou reconquête d'un avantage féminin ?*, 1988, 36 p.
- N° 16.– Gérard CALOT et Graziella CASELLI, *La mortalité en Chine d'après le recensement de 1982:*
 I.– *Analyse selon le sexe et l'âge au niveau national et provincial*, 1988, 72 p.
 II.– *Tables de mortalité par province*, 1988, 112 p.

- N° 15.– Peter AABY (s'adresser à J. VALLIN), *Le surpeuplement, un facteur déterminant de la mortalité par rougeole en Afrique*, 1987, 52 p.
- N° 14.– Jacques VALLIN, *Théorie(s) de la baisse de la mortalité et situation africaine*, 1987, 44 p.
- N° 13.– Kuakuvi GBENYON et Thérèse LOCOH, *Différences de mortalité selon le sexe, dans l'enfance en Afrique au Sud du Sahara*, 1987, 30 p.
- N° 12.– Philippe FARGUES, *Les saisons et la mortalité urbaine en Afrique. Les décès à Bamako de 1974 à 1985*, 1987, 38 p.
- N° 11.– Gilles PISON, *Les jumeaux en Afrique au Sud du Sahara : fréquence, statut social et mortalité*, 1987, 48 p.
- N° 10.– Philippe FARGUES, *La migration obéit-elle à la conjoncture pétrolière dans le Golfe ? L'exemple du Koweït*, 1987, 30 p.
- N° 9.– Didier BLANCHET, *Deux études sur les relations entre démographie et systèmes de retraite*, 1986, 26 p.
- N° 8.– Didier BLANCHET, *Équilibre malthusien et liaison entre croissances économique et démographique dans les pays en développement : un modèle*, 1986, 20 p.
- N° 7.– Jacques VALLIN, France MESLÉ et Alfred NIZARD, *Reclassement des rubriques de la 8ème révision de la Classification internationale des maladies selon l'étiologie et l'anatomie*, 1986, 56 p.
- N° 6.– Philippe FARGUES, *Un apport potentiel des formations sanitaires pour mesurer la mortalité dans l'enfance en Afrique*, 1986, 34 p.
- N° 5.– Jacques VALLIN et France MESLÉ, *Les causes de décès en France de 1925 à 1978*, 1986, 36 p.
- N° 4.– Graziella CASELLI, Jacques VALLIN, J. VAUPEL et A. YASHIN, *L'évolution de la structure par âge de la mortalité en Italie et en France depuis 1900*, 1986, 28 p.
- N° 3.– Paul PAILLAT, *Le vécu du vieillissement en 1979*, 1981, 114 p.
- N° 2.– Claude LÉVY, *Aspects socio-politiques et démographiques de la planification familiale en France, en Hongrie et en Roumanie*, 1977, 248 p.
- N° 1.– Georges TAPINOS, *Les méthodes d'analyse en démographie économique*, 1976, 288 p.