

L'Analyse des données

Quantifier en sociologie. Séance 12

Joanie Cayouette

Stéphane Beaud

Françoise Oeuvrard

Introduction

- Méthode statistique conceptualisée par Benzécri dans les années 1970.
- Dite « à la française » (en réalité, influence internationale)
- Analyse factorielle des correspondances : sur des variables numériques.
- Analyse des correspondances multiples : sur des variables dites catégorisées.
- Point de départ : un tableau individus/ variables
- On recherche les similitudes entre les lignes et les similitudes de distribution entre colonnes.

identifiant	Migration	Diplôme père	Diplôme mère	Situation financière	Type de logement	Emp père	Emp mère
0750977E1	Africains anciens	BREVET	Sans diplôme	Vous y arrivez difficilement (fin)	Locataire HLM	EMPLOYÉS COMMERCE/SERVICE	employées service
0752840E1	Mixtes	*	Dip. du supérieur	Ça va (fin)	Locataire HLM	ARTISANS, COMMERÇANTS	Employées commerce / commerçantes
0750977E2	Maghrébins anciens	BAC	cap/bep	Vous y arrivez difficilement (fin)	Locataire HLM	TECHNICIENS/AGENTS DE MAITRICE	Employées de service sans emploi
0752840E4	Français	BAC	Bac	Vous êtes à l'aise (fin)	Locataire HLM	EMPLOYÉS ADMINISTRATIFS / FONCTION PUBLIQUE	Employées administratif + fonction publique
0752840E5	Mixtes	SANS DIPLÔME	Sans diplôme	Ça va (fin)	Locataire HLM	EMPLOYÉS ADMINISTRATIFS / FONCTION PUBLIQUE	sans profession
0750977E6	Asiatiques	SANS DIPLÔME	Bac	Ça va (fin)	Locataire d'un particulier	OUVRIERS SANS EMPLOI	Autres employées / ouvrières sans emploi
0751138E6	Mixtes	DIP. DU SUPÉRIEUR	Dip. du supérieur	Vous êtes à l'aise (fin)	*	PROF. INTERM. NON INDUSTRIELLES	Employées administratif + fonction publique
0752840E6	Maghrébins anciens	DIP. DU SUPÉRIEUR	Dip. du supérieur	Ça va (fin)	Locataire HLM	OUVRIERS SANS EMPLOI	employées service
0750977E7	Français	BAC	Dip. du supérieur	Ça va (fin)	Locataire HLM	EMPLOYÉS COMMERCE/SERVICE	employées service
0752840E7	Asiatiques	*	*	C'est juste, il faut faire attention (fin)	Locataire HLM	OUVRIERS	sans profession
0750977E10	Maghrébins récents	BAC	*	Vous y arrivez difficilement (fin)	Locataire HLM	OUVRIERS SANS EMPLOI	sans profession

Définition de la population

- 1) On s'intéresse aux classes populaires.
 - Je supprime donc tous les individus dont un des parents est cadre ou chef d'entreprise ou professeur des écoles.
 - On passe de $n=2509$ à $n=1635$ individus.
- 2) Pour des raisons pratiques, je supprime tous les individus dont on ne peut dire la profession des parents (99).
 - $n=1574$
- 3) On supprime ceux avec trop de non réponses.
 - $n=1509$

Les variables d'emploi (situation d'emploi et CSP)

	Sans réponse	En emploi	Chômage	Retraite	Ne travaille plus mais a travaillé	N'a jamais travaillé	TOTAL
Professions intermédiaires administratives	3 3,4%	74 84,1%	6 6,8%	1 1,1%	1 1,1%	0	88
Techniciens	1 2,1%	45 93,8%	1 2,1%	0	1 2,1%	0	48
Agents de maîtrise	2 6,5%	25 80,6%	0	1 3,2%	2 6,5%	0	31
Ouvriers qualifiés industriels	1 3,4%	15 51,7%	7 24,1%	4 13,8%	1 3,4%	0	29
Ouvriers qualifiés artisanaux	28 10,8%	169 65,3%	28 10,8%	7 2,7%	18 6,9%	0	259
Ouvriers qualifiés magasinage	8 10,0%	57 71,3%	8 10,0%	2 2,5%	2 2,5%	0	80
Chauffeurs	1 5,0%	16 80,0%	1 5,0%	0	1 5,0%	0	20
Ouvriers non qualifiés industriels	2 6,1%	23 69,7%	4 12,1%	3 9,1%	1 3,0%	0	33
Ouvriers non qualifiés artisanaux	5 7,0%	51 71,8%	6 8,5%	5 7,0%	4 5,6%	0	71

Emploi du père (12 modalités) :

1. Non réponse (n=20)
2. Il travaille, profession inconnue. (n=74)
3. Absence de père (n=206)
4. Sans profession déclarée (n=102)
5. Artisans, commerçants (n=89)
6. Professions intermédiaires non industrielles (santé, travail social, administratif) (n=114)

Emploi du père (12 modalités) :

7. Techniciens et agents de maîtrise (n=79)

8. Employés sans emploi (n=65)

9. Employés administratifs, fonction publique (n=138)

10. Employés des commerces et services (n=117)

11. Ouvriers sans emploi (n=163)

12. Ouvriers (n=341)

Emploi de la mère (10 modalités) :

1. Non réponse (ou mère absente) (n=47)
2. Elle travaille, profession inconnue (n=63)
4. Sans profession déclarée (n=269)
5. Employées de services (n=262)
6. Professions intermédiaires non industrielles (n=201)

Emploi de la mère (10 modalités) :

7. Employées de services sans emploi (n=116)

8. Autres employées et ouvrières sans emploi (n=113)

9. Employées administratives + fonction publique (n=241)

10. Employées de commerce + commerçantes (n=99)

12. Ouvrières, techniciennes, agents de maîtrise, artisans (n=97)

Niveau de diplôme du père

C10	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Non réponse	122	8.09	122	8.09
Absence de père	191	12.67	313	20.76
Sans diplôme	429	28.45	742	49.20
Brevet	125	8.29	867	57.49
CAP/BEP	251	16.64	1118	74.14
Bac	256	16.98	1374	91.11
Diplôme du supérieur	134	8.89	1508	100.00

Homogamie des diplômes

Niveau de diplôme du père	Niveau de diplôme de la mère						
	Non réponse	Sans diplôme	Brevet	CAP/BEP	Bac	Supérieur	Total
Non réponse	85 5,64 69.67 46.70	7 0.46 5.74 1,59	7 0,46 5.74 4.64	7 0,46 5,74 2,77	9 0.60 7.38 2,88	7 0,46 5.74 4,09	122 8,09
Pas de père	29 1,92 15.18 15.93	48 3.18 25.13 10.93	19 1.26 9.95 12.58	43 2.85 22.51 17.00	34 2.25 17.80 10.90	18 1.19 9.42 10.53	191 12.67
Sans diplôme	15 0.99 3.50 8.24	297 19.69 69,23 67.65	20 1.33 4.66 13.25	35 2.32 8.16 13.83	45 2.98 10.49 14.42	17 1.13 3.96 9.94	429 28.45
Brevet	6 0.40 4.80 3.30	18 1.19 14.40 4.10	53 3.51 42.40 35.10	14 0.93 11.20 5.53	21 1.39 16.80 6.73	13 0.86 10.40 7.60	125 8.29
CAP/BEP	21 1.39 8.37 11.54	21 1.39 8.20 13.91	45 2.98 17.58 17.79	107 7.10 41.80 34.29	107 7.10 41.80 34.29	44 2.92 17.19 25.73	256 16.98
Bac	18 1.19 7.03 9.89	21 1.39 8.20 4.78	21 1.39 8.20 13.91	45 2.98 17.58 17.79	107 7.10 41.80 34.29	44 2.92 17.19 25.73	256 16.98
Supérieur	8 0.53 5.97 4.40	13 0.86 9.70 2.96	7 0.46 5.22 4.64	22 1.46 16.42 8.70	34 2.25 25.37 10.90	50 3.32 37.31 29.24	134 8.89
Total	182 12.07	439 29.11	151 10.01	253 16.78	312 20.69	171 11.34	1508 100.00

Pays de naissance des deux parents

Fréquences les plus importantes :

- Deux France : 309
- Couple mixte (un France, un autre) : 274
- Deux Autres pays asiatiques : 173
- Deux Algériens : 110
- 2 Afrique ex Afrique Occ Française : 85

Il avait 144 possibilités (12x12). Or, en ne prenant que 26 possibilités (18% des possibilités), on a 71,09% des combinaisons.

Les deux pays de naissance des parents s'articulent très étroitement.

Pays de naissance des deux parents

Variable « origine » (7 modalités)

1. Deux parents nés en France (n=308)
2. Un parent né à l'étranger (n=266)
3. Deux parents nés dans un pays asiatique (n=219)
4. Deux parents nés dans un autre pays occidental (Europe, Amérique, Océanie) (n=67)
5. Deux parents nés au Maghreb ou en Turquie (n=285)
6. Deux parents nés en Afrique (n=285)
7. Pays de naissance non renseignés ou couples hétéroclites (n=78)

Trajectoire migratoire des deux parents

Variable « année d'arrivée » (7 modalités)

1. Le père est arrivé avant 1974 (n=94)
2. Les deux sont arrivés entre 1974 et 1989 (n=211)
3. Un des deux est arrivé avant 1989 (n=258)
4. Les deux sont arrivés dans les années 1990 (n=272)
5. Les deux sont arrivés dans les années 2000 ou l'un des deux est arrivé dans les années 1990, l'autre dans les années 2000.
(n=151)
6. Non renseigné (n=279)

Variable synthétique de la trajectoire migratoire :

1. Deux parents nés en France (n=308)
2. Un parent né en France, l'autre né à l'étranger (couple mixte) (n=266)
3. Deux parents nés dans un pays asiatique (en sachant qu'il s'agit généralement d'une immigration récente) (n=219)
4. Deux parents nés dans un autre pays occidental (n=67)
5. Deux parents nés au Maghreb ou en Turquie dont au moins un des deux est arrivé en France avant 1990 (n=179)
6. Deux parents nés au Maghreb ou en Turquie qui sont arrivés après 1989 (n=106)
7. Deux parents nés en Afrique (sauf Maghreb) dont au moins un des deux est arrivé en France avant 1990 (n=131)
8. Deux parents nés en Afrique (sauf Maghreb) qui sont arrivés après 1989 (n=154)
9. Non renseigné ou couples hétéroclites (n=78)

7 variables actives (construction de l'espace social des classes populaires et moyennes) (50 modalités) :

1. Emploi du père (12)
2. Emploi de la mère (10)
3. Diplôme du père (7)
4. Diplôme de la mère (6)
5. Trajectoire migratoire (7)
6. Estimation de la situation financière (4)
7. Type de logement (4)

Variables supplémentaires (rapport au choix du collège) :

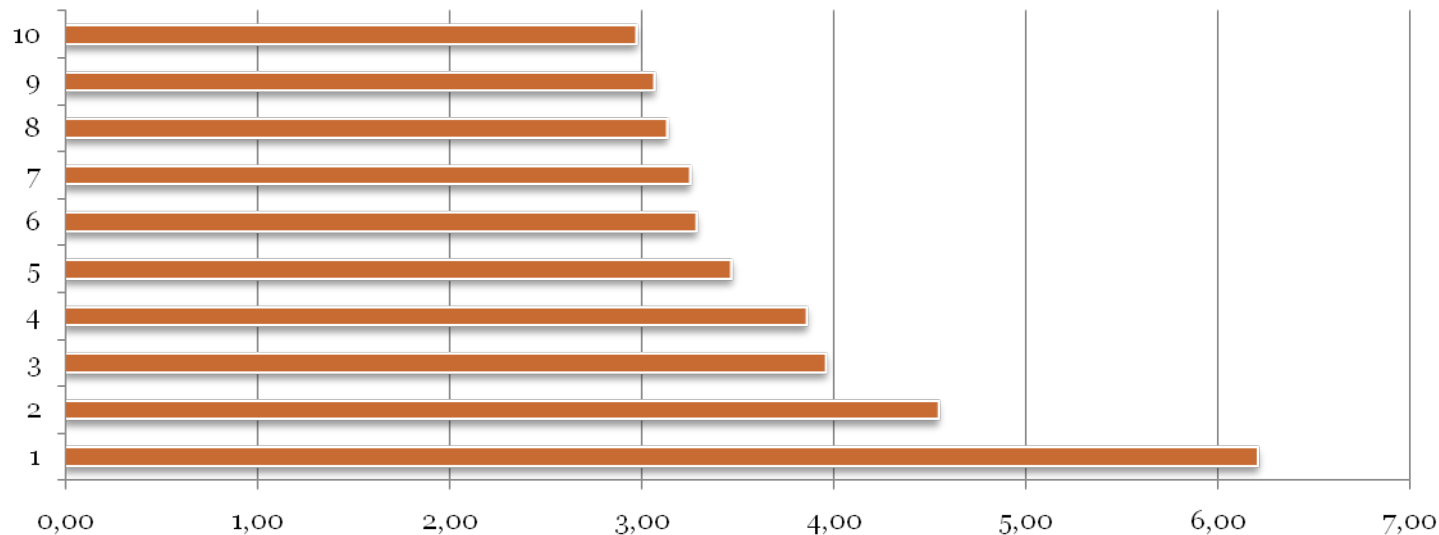
1. Type de collège demandé
2. S'est renseigné par Internet
3. S'est renseigné en parlant à des parents d'élèves
4. S'est renseigné en parlant à des voisins
5. Est allé à une journée portes ouvertes
6. Est allé à la sortie des classes

Profession du père et Trajectoire migratoire

Quelques surprises...

- Les Maghrébins « anciens » sont surreprésentés parmi les artisans / commerçants, ce qui n'est pas le cas des Maghrébins « récents »
- Surreprésentation des Français parmi les professions intermédiaires non industrielles.
- Maghrébins « anciens » et « récents » sont bien représentés parmi les techniciens et agents de maîtrise. Les Français y sont eux surreprésentés.
- Les employés sans emploi : beaucoup d'Asiatiques et d'Africains (encore plus de « récents »).
- Les Asiatiques sont surreprésentés comme employés des commerces et service ; les Maghrébins « récents » sont quasi absents de ce secteur.
- Les Maghrébins « anciens » sont beaucoup plus souvent ouvriers sans emploi que les Maghrébins « récents ».
- Les Asiatiques sont souvent ouvriers sans emploi.
- Surreprésentation des Asiatiques et des Occidentaux parmi les ouvriers en emploi.

Part d'inertie expliquée par les axes



L'ACM nous fournit un espace multi-dimensionnel.

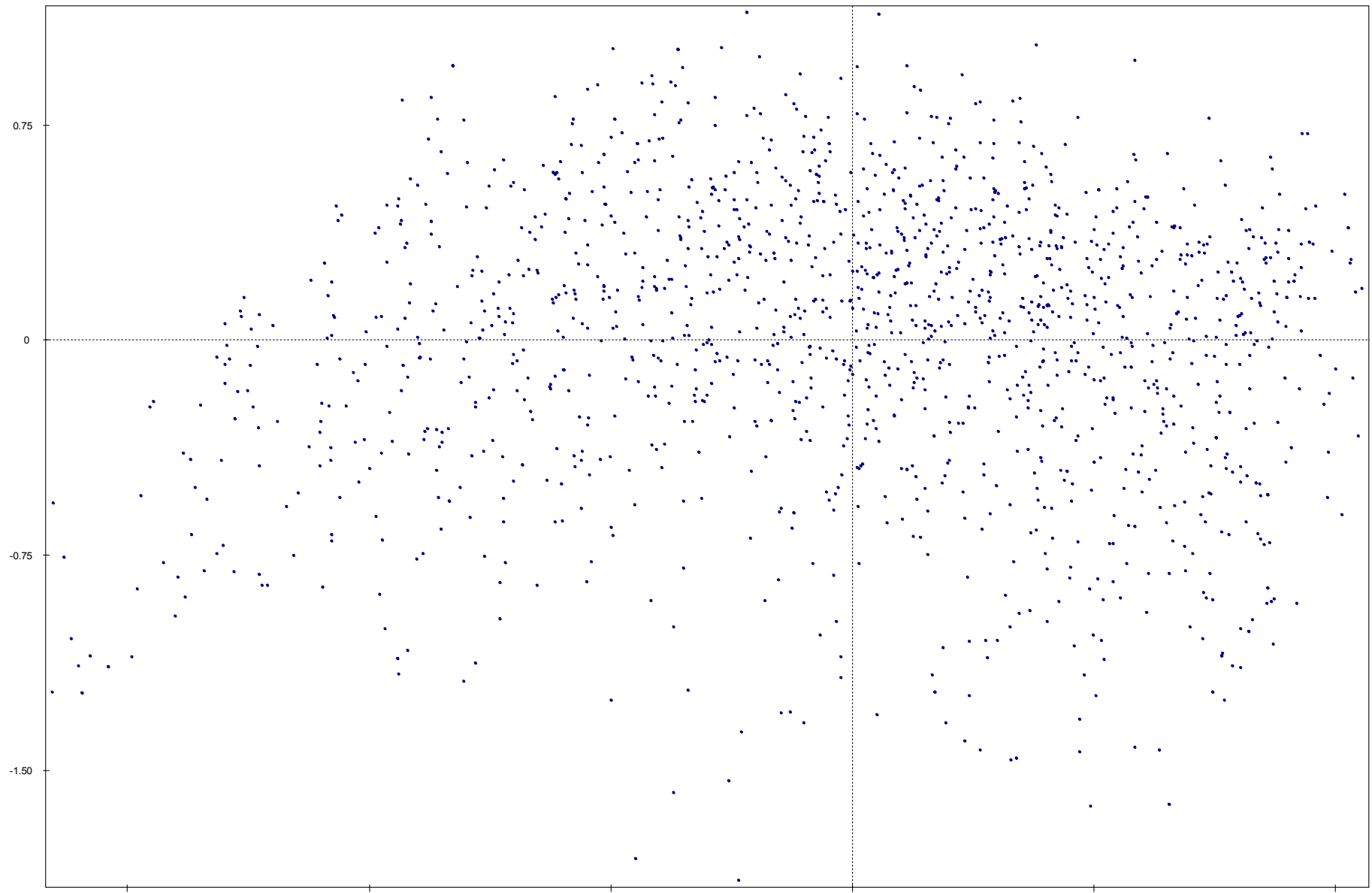
On cherche à interpréter chaque dimension.

On dit que les derniers axes sont du « bruit », de l'aléatoire (du résiduel, dû au codage, aux erreurs....)

Le nombre d'axes : nombre de modalités - 1 = 46.

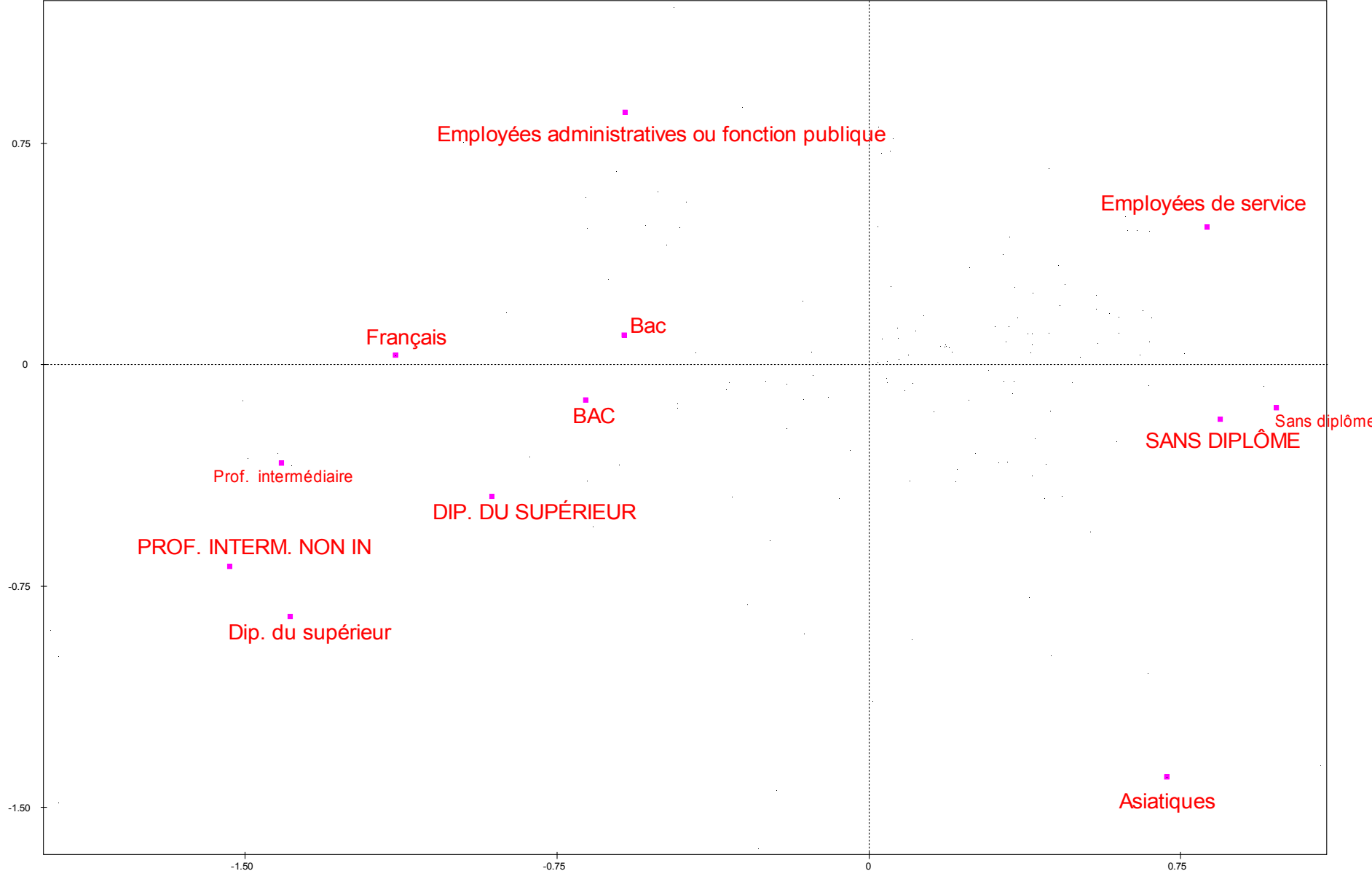
Règle du coude : on interprète les premiers axes, on n'a « pas le droit » de s'arrêter lorsque deux axes sont près (par exemple, on ne pourrait interpréter l'axe 6 sans interpréter l'axe 7).

Axe 2 - 4.54 %



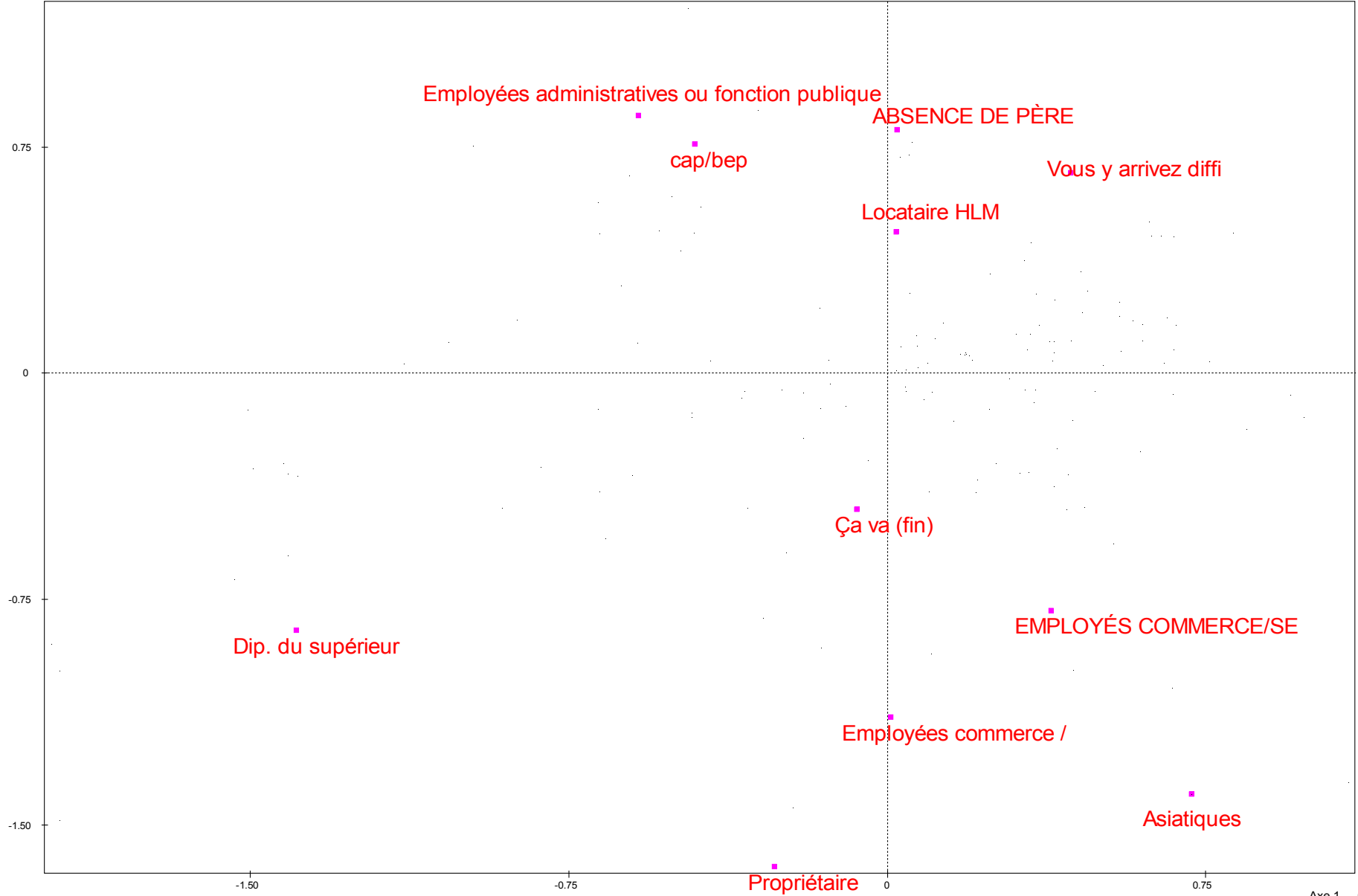
Axe 1 - 6.21 %

Axe 2 - 4.54 %



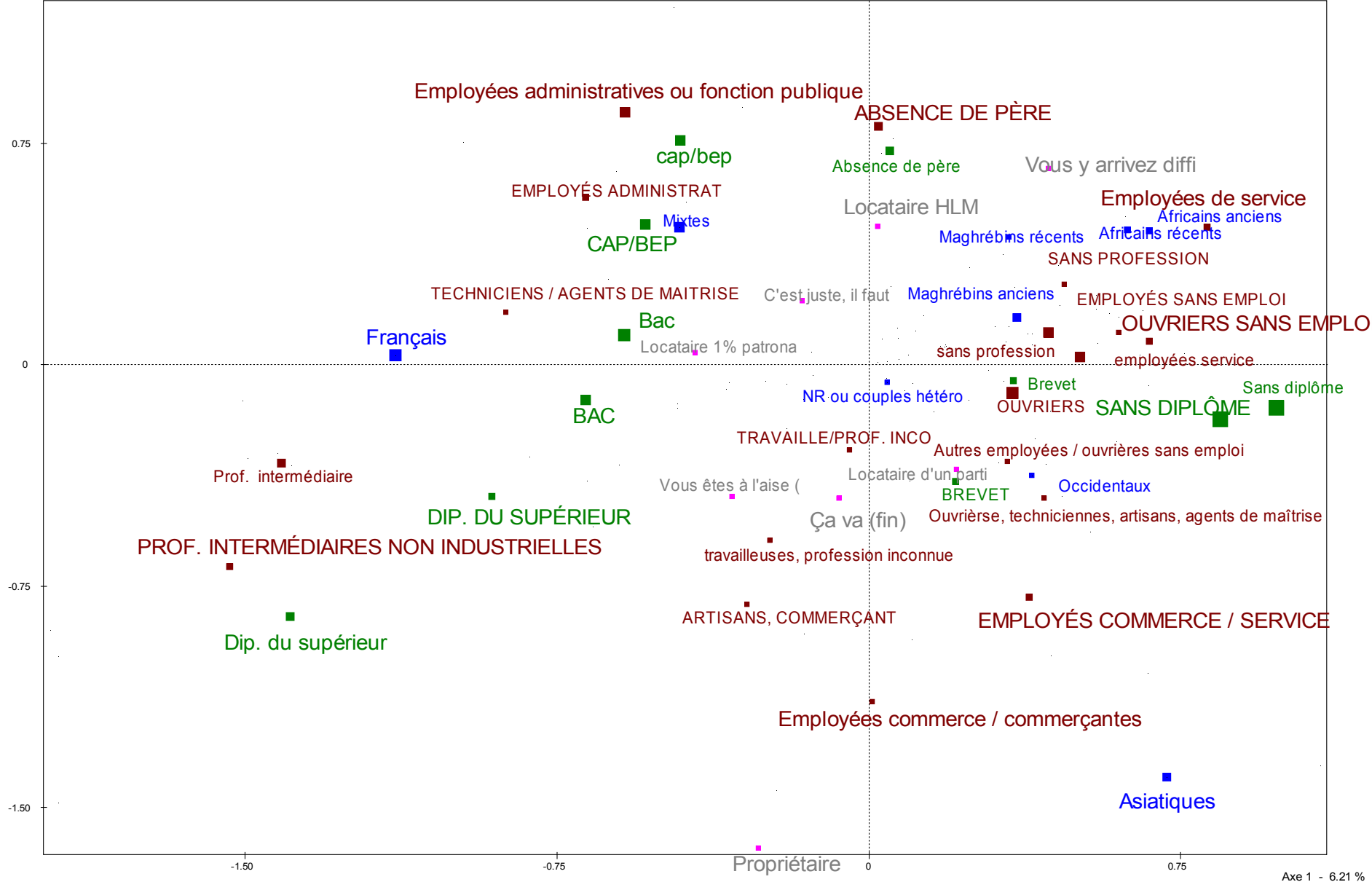
Axe 1 - 6.21 %

Axe 2 - 4.54 %

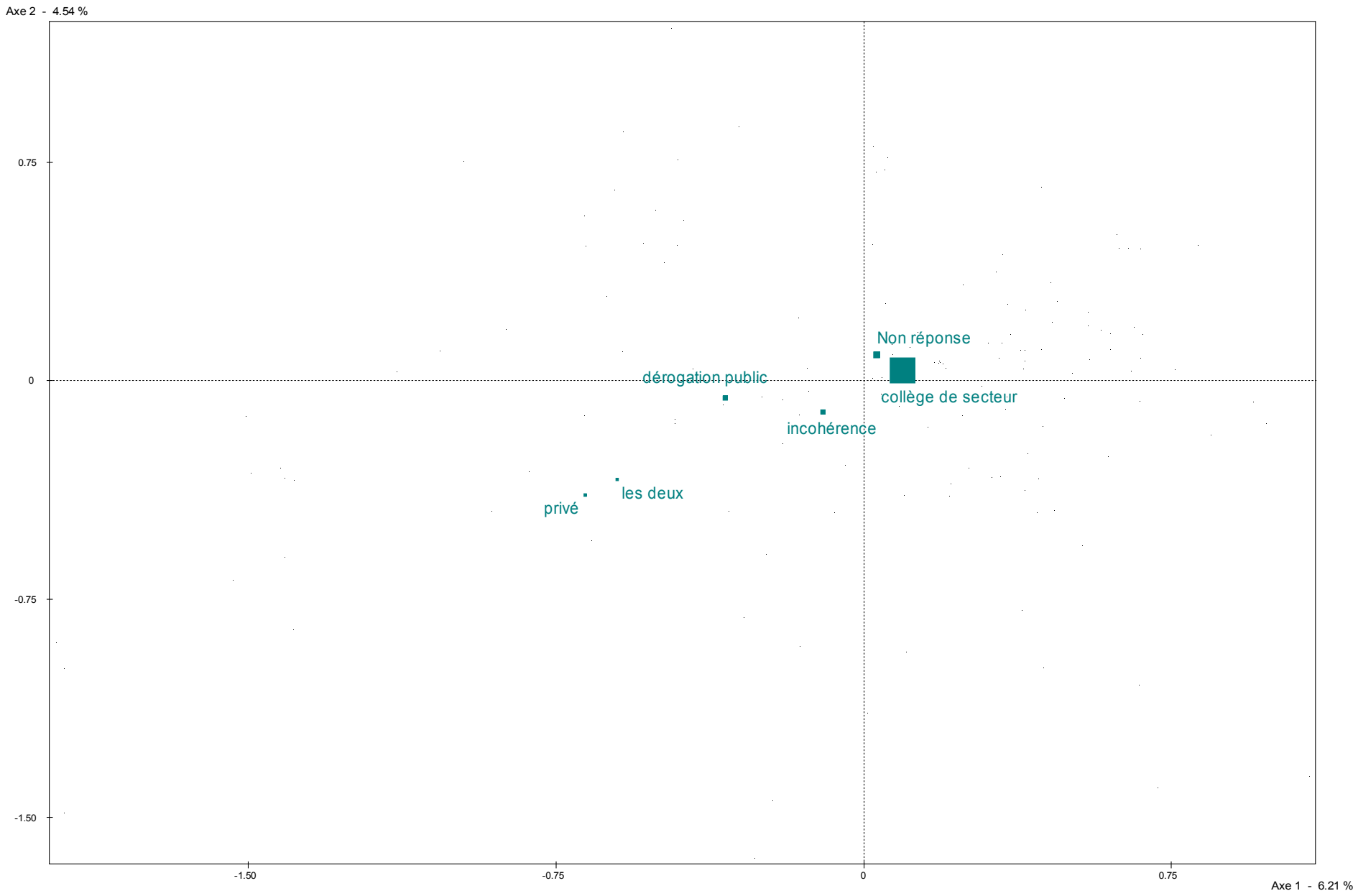


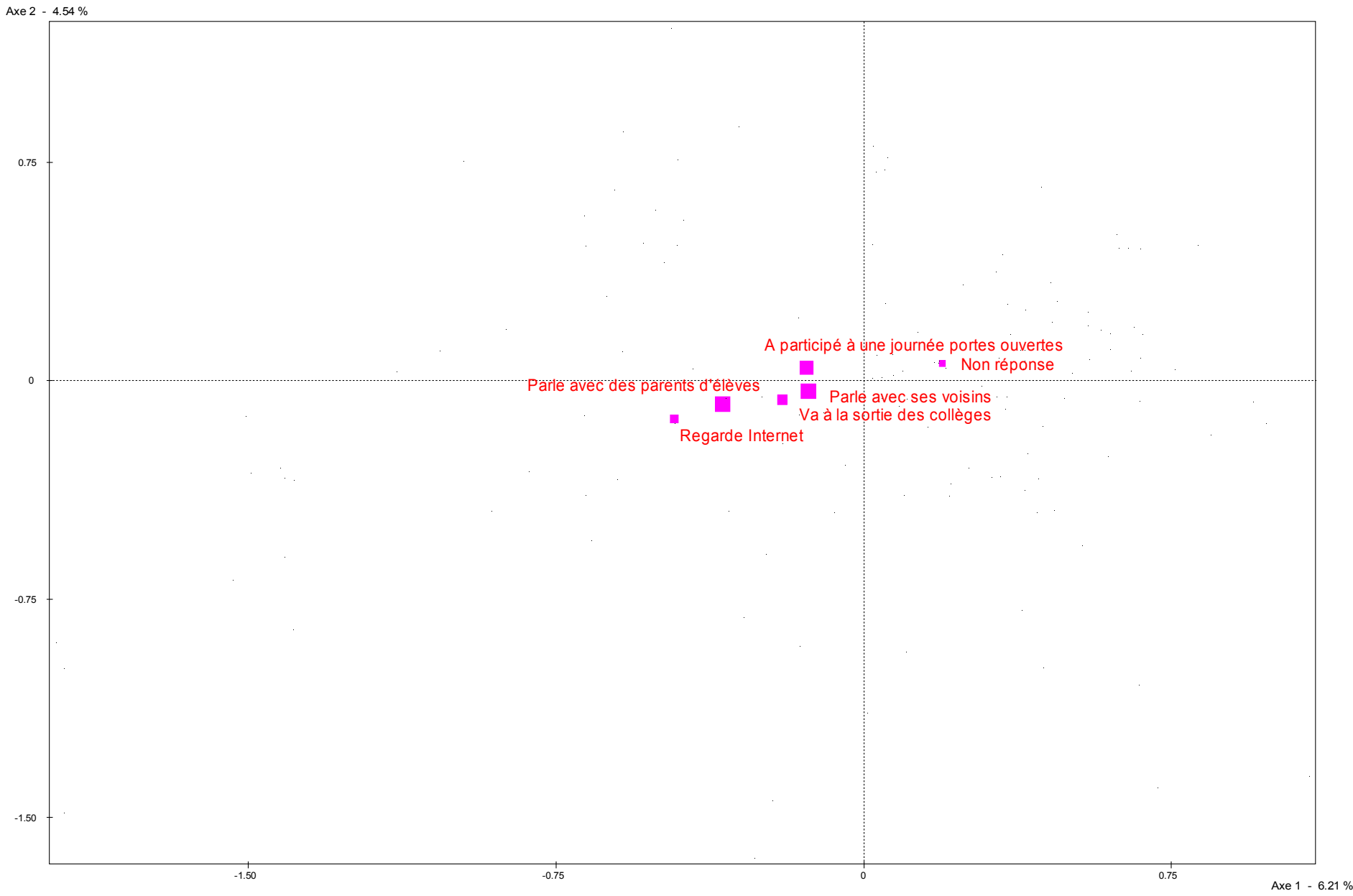
Axe 1 - 6.21 %

Axe 2 - 4.54 %



Axe 1 - 6.21 %





Bibliographie (quelques manuels)

COMBESSIE Jean-Claude (2007) *La méthode en sociologie*, La découverte, p. 103-107.

Le Roux B. et Rouanet H.: *Multiple Correspondence Analysis*, SAGE, Series: *Quantitative Applications in the Social Sciences*, CA:Thousand Oaks Paris (2010).

Voir également site personnel de B. Le Roux (en ligne : <http://www.math-info.univ-paris5.fr/~lerb/index.html>)

CIBOIS Philippe, *Les méthodes d'analyse d'enquête*, disponible en ligne : <http://cibois.pagesperso-orange.fr/SitePhCibois.htm>