

INTRODUCTION AUX ANALYSES MULTIDIMENSIONNELLES

PREMIERE SESSION 2007/2008- Durée 3h

N.B. Calculatrices et résumés de cours autorisés

Exercice I

Pour un échantillon de 40 sujets, on dispose des informations suivantes :

- Le sexe (0 = masculin, 1 = féminin)
- Le nombre d'années d'étude (de 8 à 17), codé ED
- L'âge (de 30 à 63 ans), codé AGE
- Le salaire en Euros (de 837 à 6250 Euros), codé SAL.

On analyse ces données en réalisant une ACP normée sur les variables ED, AGE et SAL ; le sexe est utilisé comme variable illustrative. On obtient pour les variables, les résultats suivants :

Corrélations

Variable	Corrélations (salaire-sexe-bis)		
	SAL	ED	AGE
SAL	1,0000	0,4603	-0,0427
ED	0,4603	1,0000	-0,4246
AGE	-0,0427	-0,4246	1,0000

Valeurs propres

Valeur numéro	Val. Propres (matrice de corrél.) & stat. associées Variables actives seules			
	Val Propre	% Total variance	Cumul Val Propre	Cumul %
1	1,6478	54,9267	1,6478	54,9
2	0,9575	31,9161	2,6053	86,8
3	0,3947	13,1572	3,0000	100,0

Saturations, contributions et qualités de représentation des variables sur les deux premiers axes

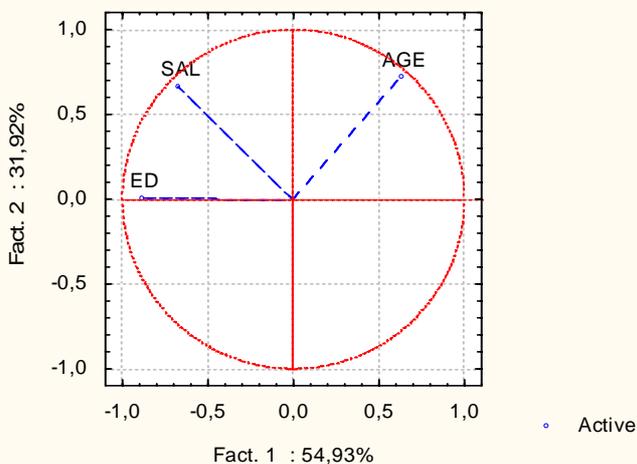
Variable	Coord. factorielles des var	
	Fact. 1	Fact. 2
SAL	-0,6753	0,6632
ED	-0,8921	0,0054
AGE	0,6292	0,7195

Variable	Contributions des var	
	Fact. 1	Fact. 2
SAL	0,2767	0,4593
ED	0,4830	0,0000
AGE	0,2402	0,5406

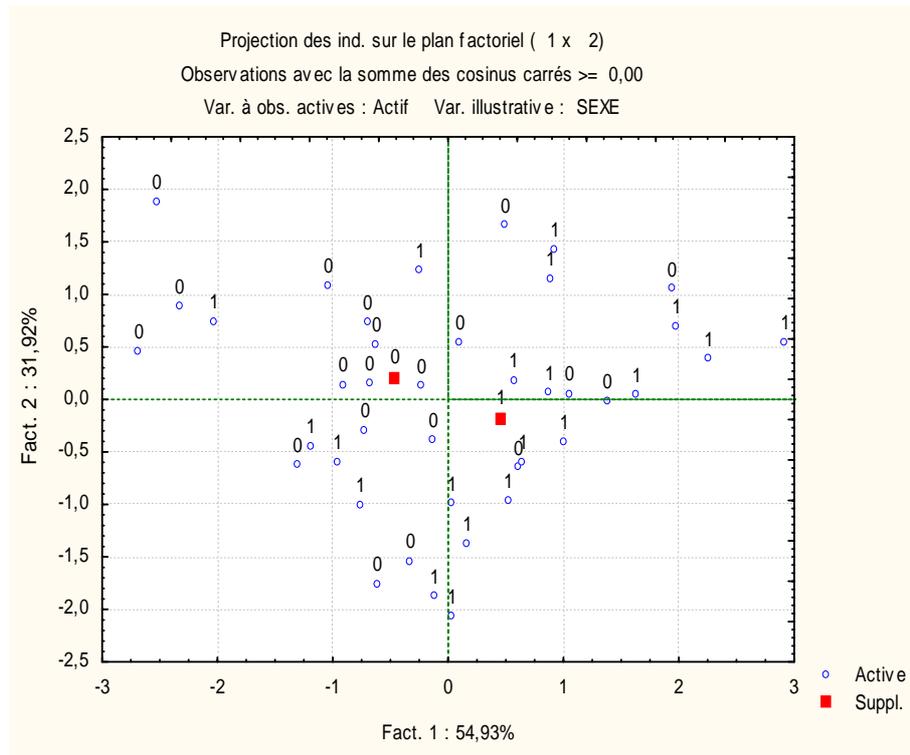
Variable	Communautés	
	Avec 1 facteur	Avec 2 facteurs
SAL	0,4560	0,8958
ED	0,7959	0,7960
AGE	0,3959	0,9135

Représentation des variables dans le premier plan factoriel

Projection des variables sur le plan factoriel (1 x 2)



Représentation des individus dans le premier plan factoriel, avec les moyennes par sexe placées en individus supplémentaires



1) Etude du tableau des corrélations

Commenter le tableau des corrélations entre les 3 variables SAL, ED et AGE.

2) Etude du tableau des valeurs propres

a) A quoi correspond la somme des valeurs propres ?

b) On choisit de n'étudier que les deux premières composantes principales. Justifier ce choix en analysant le tableau des valeurs propres.

3) Etude du nuage des variables

a) La représentation graphique des variables montre qu'elles sont toutes très bien représentées dans le plan (CP1, CP2). Justifier cette affirmation.

b) Quelle est la variable le plus fortement corrélée à la première variable principale ?

c) Question analogue pour la deuxième variable principale.

d) L'une des variables est pratiquement indépendante de la 2^e variable principale. Laquelle ?

4) On a placé les moyennes par sexe comme individus supplémentaires dans l'analyse.

a) Quel rôle jouent de tels individus dans le déroulement des calculs nécessaires à l'exécution de l'ACP ?

b) Commenter les valeurs numériques obtenues et la position de ces individus sur le graphique.

N.B. On n'a aucune information sur la population d'où est issu l'échantillon observé ici. On se gardera donc de toute généralisation des résultats observés sur cet exemple.

Exercice II

Ci-dessous figurent des extraits de l'article "LA DÉSIGNATION DES MAGHRÉBINS: EFFET DU CHOIX LEXICAL SUR LES ÉLÉMENTS ACTIVÉS DANS LA REPRÉSENTATION SOCIALE" publié en 2003 par Edith Salès-Wuillemin, Philippe Castel et Marie-Françoise Lacassagne (Laboratoire de Psychologie Clinique et Sociale - Université de Bourgogne (Dijon, France)).

Cet article concerne la représentation sociale des Maghrébins, il montre l'incidence du choix du mot inducteur dans une épreuve d'associations verbales. Six cent cinquante participants (répartis en treize groupes indépendants) ont réalisé cette tâche à partir du mot inducteur MAGHREBIN ou de l'un de ses équivalents. Les sujets devaient produire 5 verbes, 5 adjectifs et 5 substantifs (l'ordre a été randomisé). L'analyse des résultats fait ressortir que des termes différents, même s'ils sont synonymes, mobilisent des zones et des dimensions différentes (contrastées) de la représentation sociale.

Six cent cinquante sujets, tous volontaires, étudiants de première année en sciences humaines, de langue maternelle française (hommes et femmes, ayant entre 17 et 21 ans, âge médian 19 ans) ont participé à cette étude. Dans une première phase, la tâche des sujets consistait à donner, sans limite de temps, les mots qu'ils estimaient être synonymes du vocable MAGHREBIN. Après élimination des termes les plus péjoratifs (pour des raisons déontologiques) et ceux désignant des nationalités précises (Turcs), nous en avons retenu 12: ARABE, BEUR, BOUGNOUL, CLANDESTIN, ETRANGER, GRIS, IMMIGRE, INTEGRISTE, ISLAMISTE, MUSULMAN, NORD-AFRICAINE, REBEU. Durant la deuxième phase, chacun des 13 mots (MAGHREBIN et ses 12 synonymes) a servi de mot inducteur dans une tâche d'associations verbales contraintes. La tâche demandée aux sujets consistait à écrire les premiers mots qui leur venaient à l'esprit à partir du mot stimulus présenté, et pour qu'il n'y ait pas d'ambiguïté, le mot inducteur était inscrit à la suite, en gras. Afin de couvrir différentes dimensions de la représentation, chaque sujet devait donner les cinq premiers substantifs, les cinq premiers adjectifs, et les cinq premiers verbes qui lui venaient à l'esprit dès qu'il entendait le mot inducteur.

Nous avons d'abord éliminé les réponses correspondant à des désignations de Maghrébins, c'est-à-dire les mots fournis dans la phase de préparation (immigré, islamiste, ...). Puis, pour accéder à la dimension sociale, nous avons opéré une sélection sur les mots restants. Concrètement, nous avons retenu ceux cités par plus de 10 % de l'effectif dans au moins une des conditions (à savoir six sujets sur 50) ce qui nous a permis de retenir 75 mots.

Nous avons pris en compte le nombre de sujets ayant produit chacun des mots induits. Ceci a été réalisé pour chacun des mots inducteurs présentés. Les données ainsi produites sont les suivantes :

	ARA	BEU	BOU	CLA	ETR	GRI	IMM	ISL	INT	MAG	MUS	NOR	REB	TOTAL
s'intégrer	18	16	10	8	21	11	22	12	4	13	7	5	11	158
intégrer	21	17	10	8	1	12	23	13	4	16	11	5	13	154
Religion	21	4	2	1	7	4	0	36	21	6	34	4	9	149
Racisme	11	17	21	4	14	13	12	5	6	15	4	8	14	144
prier	17	2	3	4	6	5	3	5	23	6	35	8	4	121
immigré	13	10	5	14	5	6	7	6	4	13	1	9	9	102
différents	9	6	4	2	22	9	11	6	4	7	6	2	8	96
travailler	9	9	5	7	9	6	9	11	3	11	2	12	3	96
voler	7	11	9	4	2	13	1	9	2	13	4	3	11	89
voyager	10	0	3	6	2	2	14	7	3	8	3	5	15	78
rejeter	5	7	5	2	0	13	8	4	3	6	4	4	11	72
exclus	4	8	7	10	11	4	7	1	0	5	4	4	6	71
agressifs	5	5	6	1	6	10	4	3	2	7	4	4	12	69
Culture	9	4	2	2	11	3	5	3	3	5	4	8	3	62
croyants	5	3	0	0	3	3	1	3	14	6	14	6	4	62
violents	4	9	6	0	4	6	1	5	7	8	1	3	6	60
exclure	2	7	3	2	9	5	4	4	4	5	2	3	6	56

pauvres	2	1	4	17	4	3	5	0	1	6	3	3	3	52
racistes	8	8	8	0	6	5	6	4	1	0	2	0	4	52
Couleur	2	2	2	1	10	15	5	0	1	3	1	5	1	48
integration	7	3	0	1	10	1	10	1	1	7	3	2	0	46
tuer	3	3	1	0	9	1	0	6	12	3	3	3	2	46
Cité	7	4	8	0	2	3	2	2	2	1	1	0	13	45
banlieue	1	11	3	1	3	4	5	1	0	2	2	1	10	44
religieux	5	0	1	1	2	2	2	5	5	3	8	5	4	43
agresser	5	5	4	4	3	5	0	0	2	4	3	3	4	42
insulter	0	5	7	4	5	8	0	2	1	2	1	3	3	41
noirs	1	1	0	6	5	8	2	0	1	2	1	11	1	39
s'adapter	5	6	1	3	2	1	9	2	0	4	0	3	2	38
Pays	1	1	1	3	11	4	6	3	0	1	1	5	0	37
rejeté	5	4	7	4	2	0	7	0	3	1	2	0	2	37
croire	3	2	1	1	3	0	0	2	10	2	7	3	0	34
accepter	3	3	2	4	0	3	8	2	1	3	1	2	1	33
respecter	1	3	1	0	1	1	1	1	7	3	6	5	3	33
Ramadan	5	1	1	0	6	1	0	3	1	1	13	0	0	32
partir	0	0	2	3	0	1	11	4	0	3	0	0	6	30
Violence	2	3	1	0	1	1	0	6	8	1	0	1	6	30
jeûner	5	0	2	2	0	2	1	0	4	3	8	1	2	30
manger	1	1	3	1	4	2	1	2	1	0	6	5	2	29
cacher	0	0	0	9	0	1	0	12	1	3	3	0	0	29
Coran	4	0	2	0	0	0	1	7	1	0	10	1	2	28
regrouper	2	1	4	0	8	4	4	0	1	2	1	0	0	27
dangereux	0	1	2	2	3	1	0	8	5	1	1	2	1	27
fuir	1	2	1	2	0	0	6	8	0	1	1	2	2	26
immigration	2	2	1	4	3	0	2	0	2	1	0	7	1	25
Mosquée	0	1	0	0	0	0	1	8	0	1	12	0	1	24
malheureux	0	1	1	8	3	1	1	0	1	2	2	2	2	24
Différence	0	3	1	0	1	6	2	1	4	2	0	1	3	24
sympathiques	7	0	1	1	2	0	1	1	0	3	0	1	4	21
fanatiques	1	1	0	0	0	1	0	6	8	0	3	1	0	21
extrémistes	1	0	0	0	0	1	0	6	10	0	1	1	1	21
nationalité	0	0	0	7	5	0	1	0	0	3	1	0	2	19
seuls	1	0	0	6	3	1	4	0	0	0	1	1	2	19
Noirs	0	1	1	3	2	3	0	0	0	1	1	7	0	19
islam	2	0	1	0	2	0	0	6	0	1	4	0	3	19
insérer	4	0	0	0	6	1	0	1	0	2	0	3	1	18
apprendre	4	1	0	1	6	1	1	1	0	0	0	0	3	18
Sans-papiers	0	0	0	14	1	0	1	0	0	0	0	1	0	17
guerre	0	0	1	1	0	0	2	6	3	1	0	2	1	17
Attentat	1	0	1	0	0	0	0	7	5	1	0	0	1	16
péjoratif	0	4	4	1	0	6	0	0	0	0	0	0	1	16
parler	2	0	0	0	0	0	0	0	2	3	0	4	4	15
quartiers	0	0	0	0	0	1	1	0	0	1	0	6	6	15
illégaux	0	0	0	13	0	0	0	0	0	0	0	0	1	14
intègres	0	0	0	0	0	7	7	0	0	0	0	0	0	14
croyance	2	0	0	0	0	0	0	6	1	0	3	0	2	14
voilée	2	1	0	0	0	0	0	1	1	0	6	1	0	12
Bateau	0	0	0	6	5	0	0	0	0	0	0	0	0	11
chaleur	1	0	0	0	0	0	0	0	0	2	1	6	0	10
arriver	0	0	0	0	2	0	6	1	0	0	0	0	0	9
dominer	0	0	0	1	0	0	0	0	6	0	1	0	0	8
prière	0	0	0	0	0	0	0	0	0	0	6	0	1	7
cachés	0	0	0	7	0	0	0	0	0	0	0	0	0	7
désespérés	0	0	0	6	1	0	0	0	0	0	0	0	0	7
tchador	1	0	0	0	0	0	0	6	0	0	0	0	0	7
Total	280	221	182	223	275	231	254	271	220	236	270	203	259	3125

Une Analyse Factorielle des Correspondances (AFC) a ensuite été réalisée. Les principaux résultats de cette analyse sont donnés ci-dessous.

Taux de liaison

Le tableau ci-dessous donne certains taux de liaison. Plus précisément, on s'est limité aux lignes pour lesquelles le taux de liaison est maximum dans une colonne donnée.

	ARA	BEU	BOU	CLA	ETR	GRI	IMM	ISL	INT	MAG	MUS	NOR	REB
banlieue	-0,75	2,54	0,17	-0,68	-0,23	0,23	0,40	-0,74	-1,00	-0,40	-0,47	-0,65	1,74
sympathiques	2,72	-1,00	-0,18	-0,33	0,08	-1,00	-0,41	-0,45	-1,00	0,89	-1,00	-0,27	1,30
péjoratif	-1,00	2,54	3,29	-0,12	-1,00	4,07	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-0,25
quartiers	-1,00	-1,00	-1,00	-1,00	-1,00	-0,10	-0,18	-1,00	-1,00	-0,12	-1,00	5,16	3,83
intègres	-1,00	-1,00	-1,00	-1,00	-1,00	5,76	5,15	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00
Bateau	-1,00	-1,00	-1,00	6,64	4,17	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00
chaleur	0,12	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	1,65	0,16	8,24	-1,00
arriver	-1,00	-1,00	-1,00	-1,00	1,53	-1,00	7,20	0,28	-1,00	-1,00	-1,00	-1,00	-1,00
dominer	-1,00	-1,00	-1,00	0,75	-1,00	-1,00	-1,00	-1,00	9,65	-1,00	0,45	-1,00	-1,00
prière	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	8,92	-1,00	0,72
cachés	-1,00	-1,00	-1,00	13,01	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00
tchador	0,59	-1,00	-1,00	-1,00	-1,00	-1,00	-1,00	8,88	-1,00	-1,00	-1,00	-1,00	-1,00

Valeurs propres

Nombre de Dims.	Valeurs Propres et Inertie de toutes les Dimensions Table d'Entrée (Lignes x Colonnes) : 75 x 13 Inertie Totale = 1,0578 Chi ² = 3305,8 dl = 888 p = 0,000				
	ValSing.	ValProp.	%age Inertie	%age Cumulé	Chi ²
1	0,5332	0,2842	26,87	26,87	888,28
2	0,4590	0,2107	19,91	46,78	658,32
3	0,3308	0,1094	10,35	57,13	341,99
4	0,3099	0,0960	9,08	66,21	300,12
5	0,2819	0,0795	7,51	73,72	248,34
6	0,2670	0,0713	6,74	80,46	222,79
7	0,2472	0,0611	5,78	86,24	191,01
8	0,2373	0,0563	5,32	91,56	175,99
9	0,1785	0,0319	3,01	94,58	99,62
10	0,1582	0,0250	2,37	96,94	78,18
11	0,1368	0,0187	1,77	98,71	58,45
12	0,1169	0,0137	1,29	100,00	42,67

Résultats relatifs aux lignes (cf. page suivante)

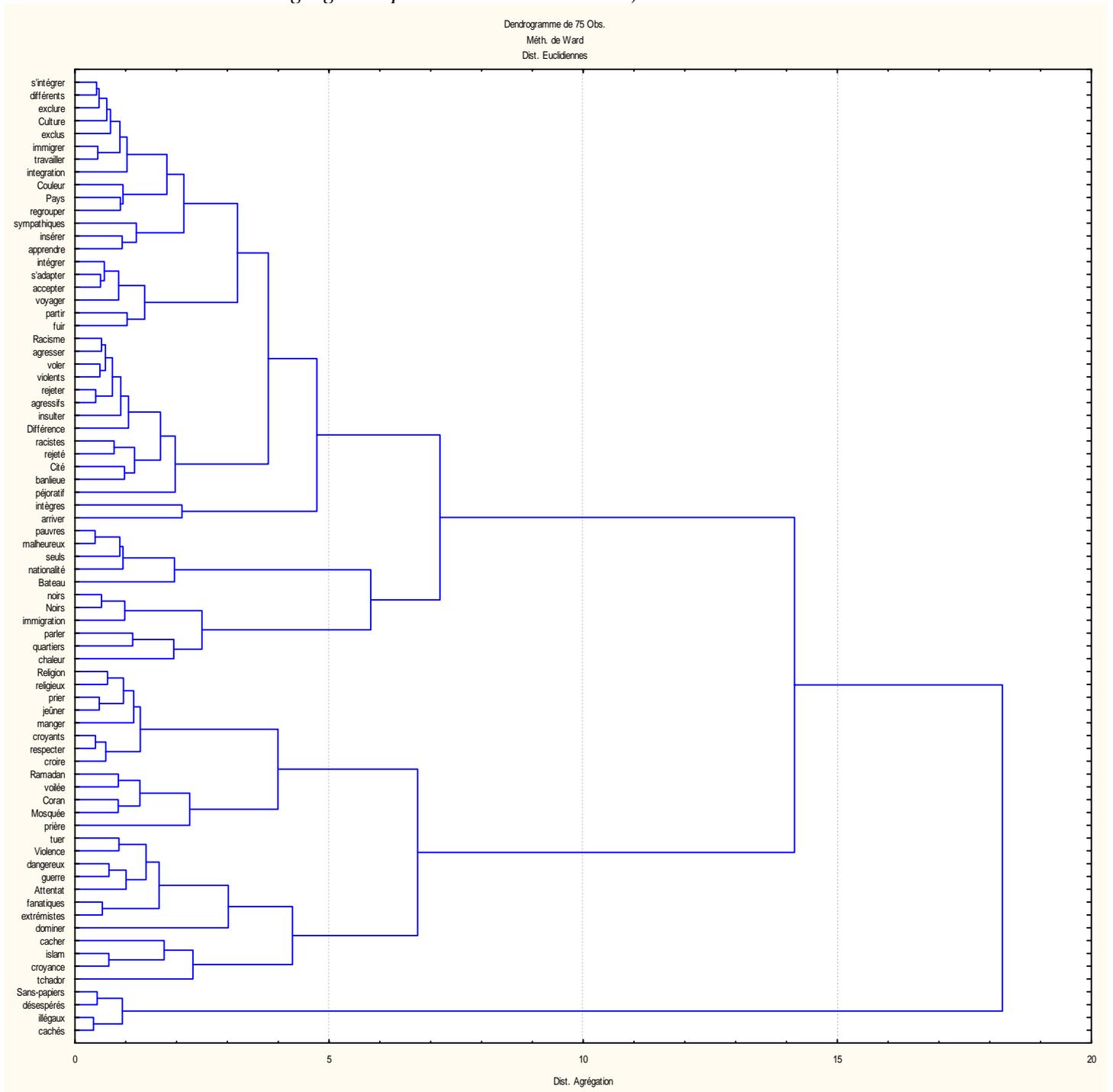
Coordonnées Ligne et Contributions à l'Inertie (Racisme-sales dans Racisme-sales.stw)										
Table d'Entrée (Lignes x Colonnes) : 75 x 13										
Standardisation : Profils ligne et colonne										
NomLigne	Ligne Numéro	Coord. Dim.1	Coord. Dim.2	Masse	Qualité	Inertie Relative	Inertie Dim.1	Cosinus ² Dim.1	Inertie Dim.2	Cosinus ² Dim.2
s'intégrer	1	0,1466	0,2437	0,0506	0,5014	0,0077	0,0038	0,1333	0,0142	0,3681
intégrer	2	0,0472	0,2105	0,0493	0,1914	0,0113	0,0004	0,0092	0,0104	0,1822
Religion	3	-0,8047	-0,4194	0,0477	0,9099	0,0408	0,1086	0,7156	0,0398	0,1944
Racisme	4	0,1397	0,4062	0,0461	0,6275	0,0128	0,0032	0,0664	0,0361	0,5611
prier	5	-0,6447	-0,4135	0,0387	0,6356	0,0338	0,0566	0,4504	0,0314	0,1853
immigrer	6	0,3260	-0,0232	0,0326	0,4337	0,0076	0,0122	0,4315	0,0001	0,0022
différents	7	0,0644	0,2738	0,0307	0,2312	0,0099	0,0004	0,0121	0,0109	0,2191
travailler	8	0,1529	0,0899	0,0307	0,1567	0,0058	0,0025	0,1165	0,0012	0,0403
voler	9	0,0387	0,2739	0,0285	0,1832	0,0112	0,0002	0,0036	0,0101	0,1796
voyager	10	0,1219	0,0805	0,0250	0,0440	0,0114	0,0013	0,0307	0,0008	0,0134
rejeter	11	0,0573	0,3840	0,0230	0,3854	0,0085	0,0003	0,0084	0,0161	0,3770
exclus	12	0,4742	0,0664	0,0227	0,6846	0,0072	0,0180	0,6714	0,0005	0,0132
agressifs	13	0,0471	0,3912	0,0221	0,5130	0,0063	0,0002	0,0073	0,0160	0,5057
Culture	14	0,0262	0,1442	0,0198	0,0781	0,0052	0,0000	0,0025	0,0020	0,0756
croyants	15	-0,6617	-0,2662	0,0198	0,6151	0,0155	0,0306	0,5294	0,0067	0,0857
violents	16	-0,1366	0,3182	0,0192	0,3000	0,0073	0,0013	0,0467	0,0092	0,2533
exclure	17	0,0692	0,2435	0,0179	0,3191	0,0034	0,0003	0,0239	0,0050	0,2953
pauvres	18	0,8676	-0,5517	0,0166	0,9124	0,0182	0,0441	0,6497	0,0240	0,2627
racistes	19	0,0286	0,4971	0,0166	0,4082	0,0096	0,0000	0,0013	0,0195	0,4068
Couleur	20	0,3242	0,5451	0,0154	0,3185	0,0183	0,0057	0,0832	0,0217	0,2353
integration	21	0,1571	0,2889	0,0147	0,1277	0,0118	0,0013	0,0291	0,0058	0,0985
tuer	22	-0,5148	-0,1736	0,0147	0,3230	0,0127	0,0137	0,2900	0,0021	0,0330
Cité	23	-0,0398	0,4907	0,0144	0,2200	0,0150	0,0001	0,0014	0,0165	0,2186
banlieue	24	0,2000	0,5597	0,0141	0,3573	0,0132	0,0020	0,0405	0,0209	0,3169
religieux	25	-0,4540	-0,2148	0,0138	0,6748	0,0049	0,0100	0,5514	0,0030	0,1234
agresser	26	0,1847	0,1151	0,0134	0,1630	0,0037	0,0016	0,1174	0,0008	0,0456
insulter	27	0,3525	0,2616	0,0131	0,2571	0,0093	0,0057	0,1658	0,0043	0,0913
noirs	28	0,5388	-0,0390	0,0125	0,2054	0,0168	0,0127	0,2043	0,0001	0,0011
s'adapter	29	0,3661	0,3024	0,0122	0,3296	0,0079	0,0057	0,1959	0,0053	0,1337
Pays	30	0,3998	0,1483	0,0118	0,1848	0,0110	0,0067	0,1624	0,0012	0,0223
rejeté	31	0,2606	0,1315	0,0118	0,1074	0,0089	0,0028	0,0856	0,0010	0,0218
croire	32	-0,6878	-0,4593	0,0109	0,5816	0,0121	0,0181	0,4022	0,0109	0,1794
accepter	33	0,3786	0,1250	0,0106	0,2859	0,0056	0,0053	0,2578	0,0008	0,0281
respecter	34	-0,5151	-0,1289	0,0106	0,3653	0,0077	0,0099	0,3438	0,0008	0,0215
Ramadan	35	-0,6920	-0,3465	0,0102	0,3414	0,0170	0,0173	0,2730	0,0058	0,0684
partir	36	0,4090	0,2343	0,0096	0,1303	0,0155	0,0057	0,0981	0,0025	0,0322
Violence	37	-0,5727	-0,0816	0,0096	0,2735	0,0111	0,0111	0,2681	0,0003	0,0054
jeûner	38	-0,3989	-0,3052	0,0096	0,3149	0,0073	0,0054	0,1986	0,0042	0,1162
manger	39	-0,1781	-0,0557	0,0093	0,0573	0,0053	0,0010	0,0522	0,0001	0,0051
caché	40	0,1248	-1,1754	0,0093	0,5327	0,0230	0,0005	0,0059	0,0609	0,5268
Coran	41	-0,8681	-0,4448	0,0090	0,5846	0,0138	0,0238	0,4631	0,0084	0,1216
regrouper	42	0,2326	0,5399	0,0086	0,3058	0,0092	0,0016	0,0479	0,0120	0,2579
dangereux	43	-0,3688	-0,4063	0,0086	0,3060	0,0080	0,0041	0,1383	0,0068	0,1678
fuir	44	-0,0187	-0,0809	0,0083	0,0060	0,0091	0,0000	0,0003	0,0003	0,0057
immigration	45	0,4459	-0,1740	0,0080	0,2030	0,0085	0,0056	0,1762	0,0011	0,0268
Mosquée	46	-1,0785	-0,6904	0,0077	0,5023	0,0237	0,0314	0,3563	0,0174	0,1460
malheureux	47	0,7984	-0,6826	0,0077	0,9033	0,0089	0,0172	0,5218	0,0170	0,3815
Différence	48	-0,0530	0,4424	0,0077	0,2150	0,0067	0,0001	0,0030	0,0071	0,2120
sympathiques	49	0,0776	0,2061	0,0067	0,0401	0,0077	0,0001	0,0050	0,0014	0,0351
fanatiques	50	-1,0923	-0,6220	0,0067	0,6690	0,0150	0,0282	0,5052	0,0123	0,1638

NomLigne	Coordonnées Ligne et Contributions à l'Inertie (Racisme-sales dans Racisme-sales.stw) Table d'Entrée (Lignes x Colonnes) : 75 x 13 Standardisation : Profils ligne et colonne									
	Ligne Numéro	Coord. Dim.1	Coord. Dim.2	Masse	Qualité	Inertie Relative	Inertie Dim.1	Cosinus ² Dim.1	Inertie Dim.2	Cosinus ² Dim.2
extrémistes	51	-1,1028	-0,6226	0,0067	0,4850	0,0210	0,0288	0,3678	0,0124	0,1172
nationalité	52	1,0295	-0,7307	0,0061	0,7182	0,0128	0,0227	0,4776	0,0154	0,2406
seuls	53	0,9416	-0,4749	0,0061	0,7401	0,0086	0,0190	0,5900	0,0065	0,1501
Noirs	54	0,5261	-0,1008	0,0061	0,1395	0,0118	0,0059	0,1345	0,0003	0,0049
islam	55	-0,6295	-0,2550	0,0061	0,3555	0,0075	0,0085	0,3054	0,0019	0,0501
insérer	56	0,0898	0,2685	0,0058	0,0528	0,0083	0,0002	0,0053	0,0020	0,0475
apprendre	57	0,2355	0,2417	0,0058	0,0843	0,0074	0,0011	0,0410	0,0016	0,0432
Sans-papiers	58	2,1207	-2,0069	0,0054	0,9868	0,0444	0,0861	0,5206	0,1040	0,4662
guerre	59	-0,3903	-0,3577	0,0054	0,1923	0,0075	0,0029	0,1045	0,0033	0,0878
Attentat	60	-0,9596	-0,4977	0,0051	0,4167	0,0136	0,0166	0,3284	0,0060	0,0883
péjoratif	61	0,4395	0,7564	0,0051	0,2584	0,0143	0,0035	0,0652	0,0139	0,1932
parler	62	-0,1592	0,1604	0,0048	0,0264	0,0088	0,0004	0,0131	0,0006	0,0133
quartiers	63	0,1676	0,4295	0,0048	0,0596	0,0162	0,0005	0,0079	0,0042	0,0517
illégaux	64	2,3048	-2,3093	0,0045	0,9552	0,0472	0,0837	0,4767	0,1134	0,4785
intègres	65	0,5493	1,0057	0,0045	0,2406	0,0231	0,0048	0,0553	0,0215	0,1853
croyance	66	-0,9605	-0,5182	0,0045	0,5424	0,0093	0,0145	0,4201	0,0057	0,1223
voilée	67	-0,9945	-0,5709	0,0038	0,5082	0,0094	0,0134	0,3822	0,0059	0,1260
Bateau	68	1,5712	-1,2261	0,0035	0,7199	0,0184	0,0306	0,4474	0,0251	0,2725
chaleur	69	-0,0494	0,0359	0,0032	0,0007	0,0160	0,0000	0,0005	0,0000	0,0002
arriver	70	0,4431	0,6009	0,0029	0,1078	0,0141	0,0020	0,0380	0,0049	0,0698
dominer	71	-1,0097	-1,1529	0,0026	0,3178	0,0179	0,0092	0,1380	0,0162	0,1799
prière	72	-1,2751	-0,8586	0,0022	0,3049	0,0164	0,0128	0,2098	0,0078	0,0951
cachés	73	2,4791	-2,5326	0,0022	0,9652	0,0276	0,0484	0,4723	0,0682	0,4929
désespérés	74	2,1938	-2,1220	0,0022	0,9777	0,0202	0,0379	0,5051	0,0479	0,4726
tchador	75	-1,0247	-0,6324	0,0022	0,1883	0,0163	0,0083	0,1364	0,0043	0,0519

Résultats relatifs aux colonnes

Nom Col.	Coordonnées Colonne et Contributions à l'Inertie (Racisme-sales dans Racisme-sales.stw) Table d'Entrée (Lignes x Colonnes) : 75 x 13 Standardisation : Profils ligne et colonne									
	Colonne Numéro	Coord. Dim.1	Coord. Dim.2	Masse	Qualité	Inertie Relative	Inertie Dim.1	Cosinus ² Dim.1	Inertie Dim.2	Cosinus ² Dim.2
ARA	1	-0,1912	0,0996	0,0896	0,1237	0,0318	0,0115	0,0973	0,0042	0,0264
BEU	2	0,1204	0,4330	0,0707	0,3616	0,0373	0,0036	0,0259	0,0629	0,3357
BOU	3	0,1573	0,4242	0,0582	0,2965	0,0380	0,0051	0,0359	0,0497	0,2606
CLA	4	1,3217	-1,1624	0,0714	0,9893	0,2113	0,4386	0,5578	0,4577	0,4315
ETR	5	0,2568	0,1569	0,0880	0,1095	0,0688	0,0204	0,0798	0,0103	0,0298
GRI	6	0,2160	0,5026	0,0739	0,3472	0,0602	0,0121	0,0542	0,0886	0,2930
IMM	7	0,3697	0,4206	0,0813	0,3072	0,0784	0,0391	0,1339	0,0683	0,1733
ISL	8	-0,6055	-0,3552	0,0867	0,3864	0,1045	0,1119	0,2875	0,0520	0,0990
INT	9	-0,8053	-0,4276	0,0704	0,5019	0,1102	0,1606	0,3915	0,0611	0,1104
MAG	10	0,0928	0,1733	0,0755	0,1268	0,0218	0,0023	0,0283	0,0108	0,0986
MUS	11	-0,7966	-0,5052	0,0864	0,5897	0,1232	0,1929	0,4205	0,1047	0,1692
NOR	12	0,0898	0,0373	0,0650	0,0084	0,0694	0,0018	0,0071	0,0004	0,0012
REB	13	0,0206	0,2728	0,0829	0,1304	0,0450	0,0001	0,0007	0,0293	0,1296

Dendrogramme relatif aux individus lignes (distance euclidienne dans l'espace multidimensionnel de dimension 12 et calcul des distances d'agrégation par la méthode de Ward)



1) Etude du tableau des taux de liaison

- a) Quelle est, dans chacune des colonnes, la valeur minimale des taux de liaison ? Quelle est la signification de cette valeur ?
- b) Quel est le taux de liaison le plus élevé ? Quelle autre interprétation peut-on donner de cette valeur ?
- c) Pour chacun des mots inducteurs, quel est le mot réponse qui lui est le plus fortement associé ?

2) Au vu du tableau des valeurs propres, combien d'axes factoriels faudrait-il étudier ? Justifier.

N.B. L'étude ci-dessous ne portera que sur les deux premiers axes.

3) On a calculé le carré de la distance euclidienne des points lignes à l'origine des axes, dans l'espace factoriel de dimension 12. Les valeurs extrêmes obtenues sont :

$$d^2(\text{s'intégrer}, O) = 0,16 \quad ; \quad d^2(\text{cachés}, O) = 13,01$$

Comment pouvait-on prévoir (qualitativement) ces résultats en examinant le tableau de contingence ?

4) Etude des qualités de représentation

- a) Parmi les mots inducteurs, quel est celui qui est le mieux représenté dans le premier plan factoriel ? Quel est celui qui est le plus mal représenté ? Justifier.
- b) De même, quelles sont les 7 modalités lignes les mieux représentées dans le premier plan factoriel ?

5) Etude de la première variable factorielle

- a) On considère le nuage des mots inducteurs. Quels sont les individus dont la contribution est supérieure à la moyenne ? Pour chacun d'eux, préciser le signe de la coordonnée correspondante. Comment peut-on interpréter cet axe en termes d'opposition entre mots inducteurs ?
- b) On considère le nuage des mots réponses. Quels sont les individus dont la contribution relative à la formation de l'axe est supérieure à 2% ? Pour chacun d'eux, préciser le signe de la coordonnée correspondante.

6) Etude de la deuxième variable factorielle : mener une étude analogue pour la deuxième variable factorielle.

7) Faire une synthèse des deux études précédentes en décrivant les résultats obtenus dans le premier plan factoriel.

8) Le graphique ci-dessous donne la position de tous les individus lignes et colonnes dans le premier plan factoriel, sans indiquer les étiquettes des points. Indiquer sur ce graphique une douzaine d'étiquettes d'individus lignes et 5 ou 6 étiquettes d'individus colonnes choisies de manière à illustrer les résultats de l'étude précédente. Le graphique annoté devra être rendu avec la copie.

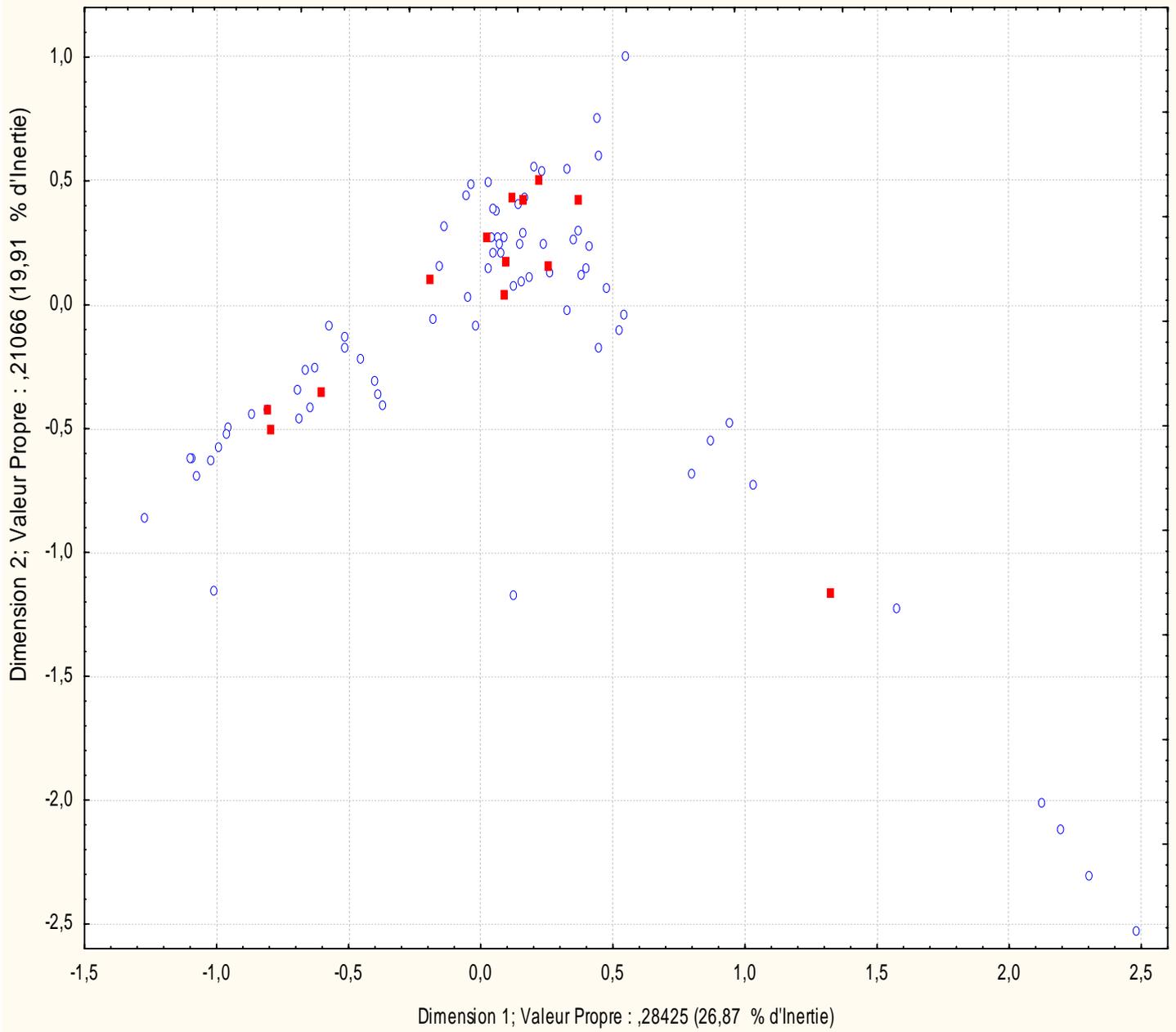
9) On a réalisé le dendrogramme ci-dessus en utilisant la distance euclidienne dans l'espace multidimensionnel de dimension 12 pour évaluer les distances entre individus lignes et la méthode de Ward pour calculer les distances d'agrégation entre classes.

- a) Comment appelle-t-on la méthode d'analyse des données utilisée ici ?
- b) Combien de classes semble-t-il pertinent de conserver pour réaliser une partition des individus lignes ? Décrire en extension la classe qui a le plus faible effectif.

Tracé 2D des Coordonnées Ligne & Colonne ; Dimension : 1 x 2

Table d'Entrée (Lignes x Colonnes) : 75 x 13

Standardisation : Profils ligne et colonne



Exercice I

I-1) Les variables SAL et ED sont corrélées positivement alors que AGE et ED sont corrélées négativement, avec des coefficients de corrélation assez modérés (0,4 en valeur absolue). En revanche, les variables AGE et SAL sont pratiquement indépendantes ($r=-0,04$, proche de 0). Ainsi, les salaires élevés sont plutôt liés à un nombre d'années d'étude élevé, les âges élevés correspondent à un nombre d'années d'étude plus faible, tandis que le niveau du salaire ne semble pas lié à l'âge.

I-2) a) On a réalisé ici une ACP normée. La somme des valeurs propres correspond au nombre de variables, ainsi qu'à l'inertie totale du nuage de points dans l'espace multidimensionnel.

I-2) b) L'inertie totale se répartit selon les 3 axes principaux. Cela correspond à une inertie moyenne par axe de $100\%/3 = 33\%$. On pourrait donc ne garder que les valeurs propres correspondant à plus de 33% d'inertie, ou de manière équivalente, celles qui sont supérieures à 1. Cependant, cela conduirait à ne garder qu'un seul axe. Et on peut remarquer que la 2^{ème} valeur propre est très proche de 1, alors que la 3^{ème} est nettement inférieure à 1. On peut donc choisir d'étudier les deux premiers axes.

I-3) a) Graphiquement, on constate que les extrémités des vecteurs représentant les projections des variables sont très proches du cercle unité, ce qui montre que la projection n'a provoqué qu'une faible déformation. On peut aussi se référer au tableau des communautés : les valeurs des \cos^2 (qualités de représentation) dans le premier plan factoriel sont comprises entre 0,80 et 0,91, ce sont donc des valeurs élevées.

I-3) b) Le résultat peut être lu sur le graphique. On peut aussi se référer à la colonne "Fact. 1" du tableau des saturations (les saturations sont les corrélations entre les variables d'origine et les variables principales). On voit que la corrélation la plus élevée (en valeur absolue) est celle qui existe entre ED et CP1. Il faut toutefois noter que cette corrélation est négative : ces deux variables sont anti-corrélées.

I-3) c) Pour la deuxième variable principale, c'est la variable AGE qui lui est le mieux corrélée ($r=0,72$), mais cette corrélation est assez proche de celle liant SAL et CP2. Dans les deux cas, il s'agit de corrélations positives.

I-3) d) La variable ED est pratiquement indépendante de CP2 ($r=0,0054$). Cela est confirmé par la représentation graphique, où la variable ED est pratiquement orthogonale à l'axe des ordonnées.

I-4) a) Les valeurs observées des variables sur les individus supplémentaires subissent les mêmes transformations que celles observées sur les individus actifs. En revanche, les individus supplémentaires n'interviennent ni dans le calcul de l'inertie totale, ni dans la détermination des composantes principales.

I-4) b) Les sujets de sexe masculin (ainsi que l'individu "moyenne des hommes") se trouvent majoritairement au-dessus de la première bissectrice (droite d'équation $Y=X$) tandis que ceux de sexe féminin se trouvent majoritairement en dessous. Si on associe cette représentation avec celle des variables, on constate :
- que les hommes sont associés à des salaires et des durées d'étude plus élevés que les femmes ;
- que la répartition des sujets le long de la première bissectrice correspond dans une large mesure à un classement par âge.

Exercice II

II-1) a) Dans chaque colonne, la valeur minimum est -1. Cette valeur correspond à des mots qui n'ont pas été associés au mot inducteur concerné.

II-1) b) Le taux le plus élevé est 13,01 (liaison CLANDESTIN - cachés). L'effectif observé est donc de 1300% supérieur à l'effectif théorique que l'on aurait observé sous hypothèse d'indépendance (ou encore, 14 fois plus élevé que l'effectif théorique).

II-1) c)

Mot inducteur	Mot réponse le plus fortement associé
Arabe	sympathiques
Beur	banlieue, péjoratif
Bougnoul	péjoratif
Clandestin	cachés
Etranger	bateau
Gris	intègres
Immigré	arriver
Islamiste	tchador
Intégriste	dominer
Maghrébin	chaleur
Musulman	prière
Nord-africain	chaleur
Rebeu	quartiers

II-2) L'ensemble de l'inertie est ici répartie sur 12 axes principaux, soit en moyenne 8,33% d'inertie par axe. Les axes représentant une inertie supérieure à la moyenne sont donc les 4 premiers.

II-3) "s'intégrer" a été choisi 158 fois au total ; c'est l'effectif le plus élevé du tableau de contingence, et de plus, il est cité en association avec tous les mots inducteurs. Il a donc fortement influencé le profil moyen, et son profil ne s'écarte que peu de ce profil moyen. En revanche, "cachés" est concentré exclusivement sur la colonne "Clandestin", et c'est le seul mot qui ne soit cité que dans une seule colonne. Son profil est donc très éloigné du profil moyen.

II-4) a) En observant les valeurs contenues dans la colonne "Qualité" du tableau des résultats relatifs aux colonnes, on voit que le mot inducteur le mieux représenté est "Clandestin" (qualité de représentation : 0,99) alors que le plus mal représenté est "Nord-africain" (qualité de représentation : 0,01).

II-4) b) De même, pour les modalités lignes, on obtient, dans l'ordre : "sans-papiers" (0,98), "désespérés" (0,97), "cachés" (0,96), "illégaux" (0,95), "religion" (0,91), "pauvres" (0,91), "malheureux" (0,90).

II-5) a) On retient ici les mots inducteurs dont la contribution à l'inertie du premier axe est supérieure à $100\%/13=7,7\%$.

-	+
Musulman Islamiste Intégriste	Clandestin

II-5) b)

-	+
religion prier croyants Coran	pauves désespérés nationalité sans-papiers

mosquée fanatiques	illégaux bateau cachés
-----------------------	------------------------------

II-6) a) On retient ici les mots inducteurs dont la contribution à l'inertie du deuxième axe est supérieure à $100\%/13=7,7\%$.

-	+
Clandestin Musulman	Gris

II-6) b)

-	+
religion prier pauvres cachés sans-papiers illégaux bateau désespérés	racisme couleur banlieue intègres

II-7) Le premier axe oppose les termes relatifs à la religion, voire à l'extrémisme religieux aux termes relatifs aux difficultés relatives à l'immigration : immigration clandestine, difficulté à obtenir des papiers, etc. Sur le second axe, la religion et la clandestinité se retrouvent sur la partie négative de l'axe, tandis que la partie positive concerne plutôt les immigrés "légaux", présents dans les banlieues et victimes du racisme. Les résultats obtenus suggèrent donc trois pôles : religion, immigration clandestine, racisme et banlieues.

II-9) a) On a utilisé ici une CAH (classification ascendante hiérarchique).

II-9) b) On observe sur le dendrogramme une augmentation importante de la distance d'agrégation entre la partition en 3 classes (distance d'agrégation de l'ordre de 7) et la partition en 2 classes (distance d'agrégation de l'ordre de 14). On peut donc choisir une partition en 3 classes. La classe la moins nombreuse est alors {sans-papiers, désespérés, illégaux, cachés};

II-8)

Tracé 2D des Coordonnées Ligne & Colonne ; Dimension : 1 x 2

Table d'Entrée (Lignes x Colonnes) : 75 x 13

Standardisation : Profils ligne et colonne

