

SECTION: Psychologie - Master 2ème année

N.B. Calculatrice et documents autorisés

ANALYSE MULTIDIMENSIONNELLE DES DONNEES
Sujet donné en 2008-2009 et indications de correction

Une équipe de chercheurs s'est intéressée à l'effet de la culture sur la catégorisation des odeurs¹. Pour l'expérience décrite ci-dessous, trois groupes de sujets ont été recrutés : 30 étudiants de l'Université de Bourgogne (Dijon - France), 30 étudiants de l'Université de Dallas (Texas - USA) et 30 étudiants de l'Institut Polytechnique de Danang (Vietnam).

Le matériel était constitué de 40 odorants fournis par une société spécialisée. Chacun de ces odorants était présenté aux sujets, qui devaient évaluer la typicalité de son odeur par rapport à 11 catégories (animal, pâtisserie, sucrerie, nettoyant, cosmétique, fleur, fruit, pharmacie, moisissure, nature, épice).

Les sujets donnaient leurs réponses sur des échelles de Likert en 7 points (1 = non typique, 7 = tout à fait typique).

On construit le protocole des moyennes observées pour le score de typicalité de chaque odorant, pour l'ensemble des participants. Les valeurs observées sont les suivantes :

	animal	pâtisserie	sucrerie	nettoyant	cosmétique	fleur	fruit	pharmacie	moisissure	nature	épice
ambre	1,36	1,84	2,19	2,36	4,04	2,65	1,96	2,66	2,32	2,27	2,54
anis	1,23	2,34	5,20	1,76	1,77	2,22	2,23	2,65	1,51	2,02	3,26
abricot	1,32	2,55	4,70	2,10	3,00	3,41	4,01	2,28	1,30	2,33	1,66
cassis	1,28	2,40	5,49	1,86	2,04	2,73	4,42	2,08	1,52	2,15	1,57
beurre	2,42	2,65	2,01	1,92	1,97	1,77	1,71	1,85	2,39	2,16	1,86
urine de chat	3,59	1,26	1,37	2,16	1,76	1,53	1,42	2,12	3,32	2,45	1,73
cannelle	1,34	3,27	3,59	1,69	2,29	2,23	1,89	2,07	1,49	2,63	4,83
civette	4,52	1,21	1,24	2,09	1,29	1,50	1,29	1,73	3,39	2,80	1,34
clou de girofle	1,62	2,05	1,93	1,91	1,77	2,00	1,62	3,15	2,04	2,23	4,28
biscuits	1,12	4,70	4,81	1,62	3,07	2,43	2,61	1,99	1,34	1,71	2,35
détergent	1,70	1,39	1,42	4,38	2,93	2,03	1,53	2,34	1,92	2,13	1,56
eucalyptus	1,26	1,26	2,69	3,19	1,89	1,71	1,50	5,10	1,68	2,24	1,65
gingembre	1,51	2,16	2,61	2,89	2,11	2,19	2,42	2,37	1,87	2,91	3,41
noisette	2,23	3,27	2,94	1,45	1,67	1,47	1,82	1,68	2,76	2,01	2,44
miel	2,34	2,22	2,59	2,25	1,53	2,35	1,77	2,24	2,22	2,83	2,12
jasmin	1,97	1,57	1,87	2,99	2,93	4,28	1,94	2,29	2,42	2,94	1,71
lavande	1,31	1,33	1,74	4,02	2,39	3,59	1,84	2,39	2,11	3,15	2,07
cuir	2,32	1,49	1,40	3,09	2,06	1,44	1,41	2,71	2,75	2,25	2,17
mangue	1,49	1,73	2,53	2,91	3,09	3,33	3,28	2,31	2,16	2,62	1,91
melon	1,19	2,57	5,30	1,70	2,51	2,87	5,81	1,78	1,28	2,61	1,48
lait	1,48	3,51	3,77	1,74	2,41	2,44	2,72	2,04	1,67	2,32	2,55
anti-mites	1,78	1,57	1,53	4,45	2,25	1,76	1,52	2,64	2,43	1,85	1,48

¹ C. Chrea, D. Valentin, C. Sulmont-Rossé, D. Hoang-Nguyen, H. Abdi, Semantic, Typicality and Odor Representation : A Cross-cultural Study, Chem. Senses V. 30 no. 1, pp 37-49, 2005
 C. Chrea, D. Valentin, C. Sulmont-Rossé, D. Hoang-Nguyen, H. Abdi, Odeurs et catégorisation : notre représentation mentale des odeurs est-elle universelle ou dépendante de notre culture ? in Actes de la 3^{ème} journée du Sensolier, Paris, France, 2005.

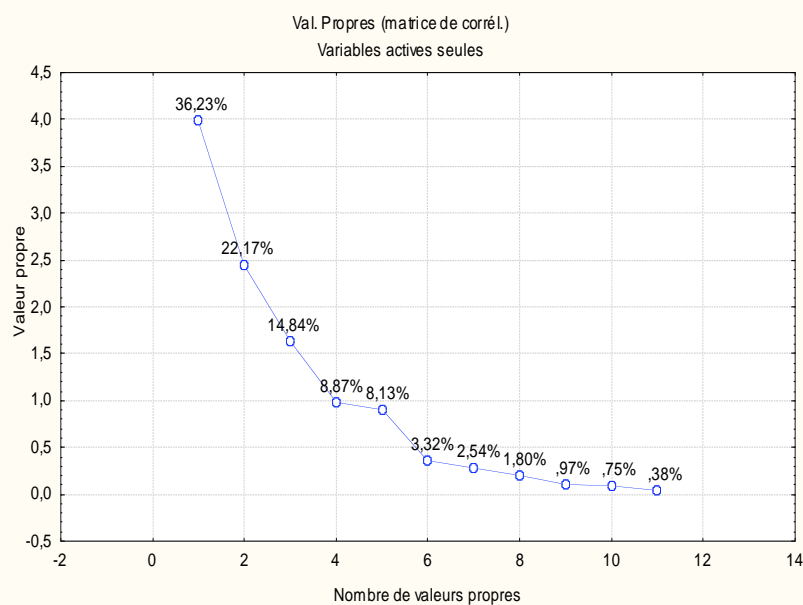
moisi	2,10	1,17	1,21	2,36	1,26	1,75	1,32	1,96	5,10	4,22	1,74
champignon	2,34	1,17	1,45	2,31	1,44	1,53	1,35	2,21	4,46	3,36	1,67
musc	2,31	1,41	1,73	2,85	3,54	2,56	1,75	2,20	2,00	2,01	1,71
noix de muscade	1,80	1,95	1,79	2,31	1,63	2,10	1,63	2,58	2,94	2,26	3,82
fleur d'oranger	1,40	1,85	1,92	3,46	2,95	3,89	2,14	2,44	1,94	2,68	1,68
arachide	1,73	3,15	2,65	1,90	1,72	1,78	1,95	2,02	2,58	2,39	2,53
pin	1,44	1,47	1,68	3,09	2,89	2,28	1,58	2,81	2,70	2,85	1,98
ananas	1,20	2,58	4,83	2,03	2,26	2,67	5,40	1,85	1,53	2,54	1,75
rose	1,30	1,86	2,66	3,21	3,87	3,82	2,13	1,86	1,49	2,46	1,34
savon	1,21	1,35	1,50	4,70	4,04	2,27	1,54	2,07	1,36	1,90	1,40
fraise	1,09	3,36	5,75	1,66	2,58	2,70	4,84	2,20	1,17	2,26	1,65
thé	2,07	1,73	1,46	1,89	1,86	2,28	1,43	2,14	2,58	2,85	2,40
truffe	3,28	1,74	1,56	1,81	1,61	1,59	1,47	2,00	3,29	2,57	1,73
vanille	1,12	3,46	3,81	2,09	3,41	2,96	2,40	1,85	1,42	2,28	2,90
violette	1,53	1,66	3,06	2,83	3,14	3,26	2,26	2,10	2,28	2,37	1,53
noix	1,58	3,21	2,31	1,45	1,50	1,60	1,92	2,43	2,42	2,54	3,72
pyrole	1,17	1,47	2,79	2,43	2,14	1,93	1,49	5,23	1,41	1,67	1,88
boisé	2,11	1,66	1,74	2,35	1,78	2,27	1,52	2,24	3,06	3,39	2,15

Les données rassemblées dans le tableau ci-dessus ont été traitées à l'aide d'une ACP normée, dont les principaux résultats sont rassemblés ci-dessous.

Résultats de l'analyse en composantes principales :

Val. Propres & stat. associées.

	Val Propre	% Total variance	Cumul Val Propre	Cumul %
1	3,99	36,23	3,99	36,23
2	2,44	22,17	6,42	58,40
3	1,63	14,84	8,06	73,24
4	0,98	8,87	9,03	82,11
5	0,89	8,13	9,93	90,24
6	0,36	3,32	10,29	93,56
7	0,28	2,54	10,57	96,10
8	0,20	1,80	10,77	97,90
9	0,11	0,97	10,88	98,87
10	0,08	0,75	10,96	99,62
11	0,04	0,38	11,00	100,00



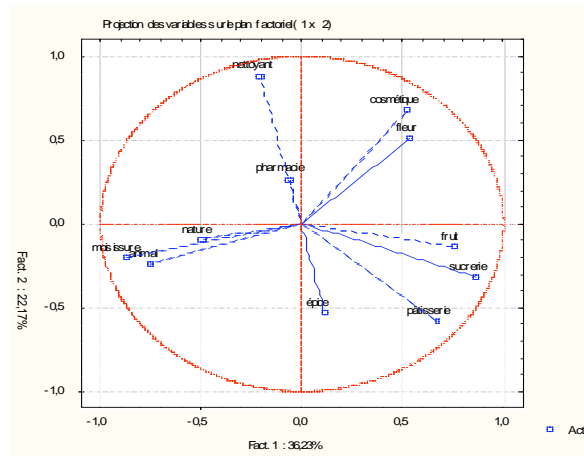
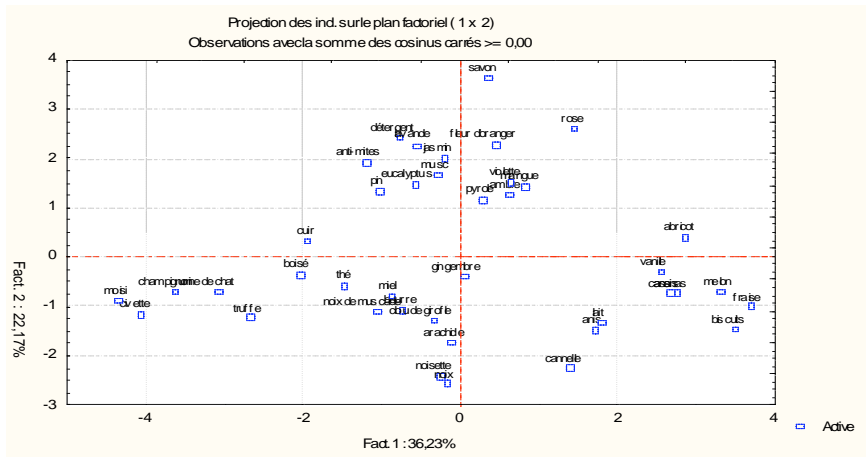
Scores, contributions et qualités de représentation des individus

	Coord. 1	Ctr. 1	Cos ² 1	Coord. 2	Ctr.2	Cos ²	Cos ² 1&2
ambre	0,64	0,26	0,06	1,26	1,62	0,25	0,31
anis	1,71	1,84	0,32	-1,50	2,29	0,24	0,56
abricot	2,87	5,18	0,80	0,38	0,14	0,01	0,82
cassis	2,67	4,46	0,60	-0,75	0,58	0,05	0,64
beurre	-0,71	0,32	0,12	-1,09	1,22	0,28	0,40
urine de chat	-3,06	5,88	0,71	-0,72	0,53	0,04	0,75
cannelle	1,41	1,24	0,13	-2,27	5,27	0,34	0,47
civette	-4,04	10,26	0,64	-1,18	1,43	0,05	0,69
clou de girofle	-0,33	0,07	0,01	-1,30	1,73	0,17	0,18
biscuits	3,50	7,68	0,63	-1,48	2,26	0,11	0,75
détergent	-0,76	0,36	0,06	2,41	5,98	0,62	0,68
eucalyptus	-0,56	0,19	0,02	1,46	2,20	0,11	0,12
gingembre	0,05	0,00	0,00	-0,42	0,18	0,05	0,05
noisette	-0,24	0,04	0,01	-2,44	6,11	0,68	0,68
miel	-0,87	0,48	0,29	-0,81	0,67	0,25	0,54
jasmin	-0,20	0,03	0,00	1,98	4,04	0,39	0,39
lavande	-0,53	0,17	0,03	2,24	5,16	0,50	0,52
cuir	-1,92	2,32	0,64	0,29	0,09	0,01	0,65
mangue	0,82	0,42	0,15	1,42	2,06	0,45	0,60
melon	3,32	6,93	0,57	-0,71	0,52	0,03	0,60
lait	1,81	2,05	0,57	-1,35	1,86	0,32	0,89
anti-mites	-1,18	0,87	0,13	1,90	3,71	0,34	0,47
moisi	-4,33	11,74	0,60	-0,91	0,84	0,03	0,63
champignon	-3,62	8,21	0,78	-0,73	0,55	0,03	0,81
musc	-0,27	0,04	0,01	1,65	2,80	0,45	0,46
noix de muscade	-1,05	0,70	0,16	-1,12	1,28	0,18	0,35
fleur d'oranger	0,48	0,15	0,03	2,27	5,26	0,66	0,69
arachide	-0,11	0,01	0,00	-1,75	3,14	0,73	0,74
pin	-1,01	0,64	0,25	1,33	1,82	0,43	0,68
ananas	2,75	4,75	0,55	-0,75	0,58	0,04	0,59
rose	1,45	1,32	0,18	2,59	6,90	0,58	0,76
savon	0,36	0,08	0,01	3,61	13,38	0,71	0,72
fraise	3,71	8,64	0,78	-1,02	1,07	0,06	0,84
thé	-1,45	1,32	0,59	-0,61	0,38	0,10	0,69
truffe	-2,64	4,38	0,65	-1,24	1,59	0,14	0,79
vanille	2,55	4,09	0,69	-0,33	0,11	0,01	0,70
violette	0,65	0,27	0,10	1,50	2,32	0,55	0,66
noix	-0,15	0,01	0,00	-2,58	6,85	0,70	0,71
pyrole	0,31	0,06	0,00	1,14	1,33	0,06	0,06
boisé	-2,02	2,55	0,63	-0,39	0,15	0,02	0,65

Saturations, contributions et qualités de représentation des variables

	Fact. 1	Contrib. 1	Cos ²	Fact. 2	Contrib. 2	Cos ² 1&2
animal	-0,76	0,144	0,57	-0,24	0,024	0,63
pâtisserie	0,67	0,114	0,45	-0,59	0,142	0,80
sucrerie	0,87	0,189	0,75	-0,32	0,042	0,85
nettoyant	-0,20	0,010	0,04	0,89	0,322	0,83
cosmétique	0,52	0,068	0,27	0,68	0,190	0,73
fleur	0,53	0,071	0,28	0,51	0,105	0,54
fruit	0,76	0,144	0,58	-0,14	0,008	0,59
pharmacie	-0,06	0,001	0,00	0,27	0,029	0,07
moisissure	-0,88	0,194	0,77	-0,21	0,017	0,82
nature	-0,50	0,062	0,25	-0,09	0,003	0,25
épice	0,11	0,003	0,01	-0,53	0,117	0,30

Représentations des observations et des variables dans le premier plan factoriel



1) Examiner le tableau des valeurs propres et la représentation graphique correspondante. Selon le choix du critère utilisé, de 3 à 5 composantes principales pourraient être retenues. Justifier.

N.B. Dans les questions qui suivent, on n'étudie que les deux premières composantes principales.

2) a) Quels sont les 8 individus (odorants) qui contribuent le plus à la formation du premier axe ? Quel est le signe de la coordonnée correspondante ?

b) De même, quelles sont les 5 catégories qui contribuent le plus à la formation du premier axe ? Quel est le signe des saturations correspondantes ?

c) Faire une synthèse des résultats trouvés concernant le premier axe factoriel.

3) Etudier de même le second axe.

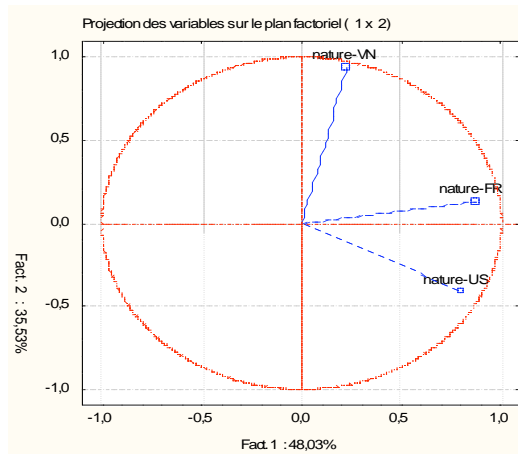
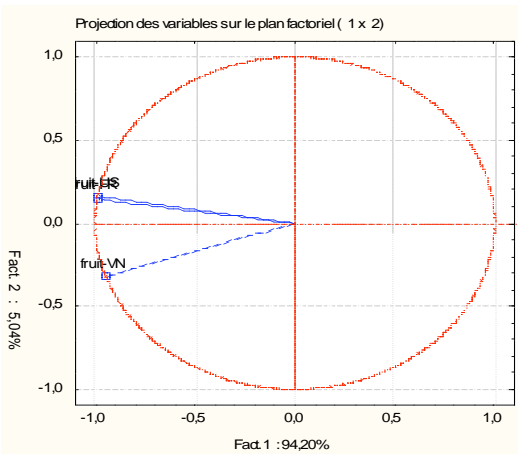
4) Quelle est la variable (catégorie) la plus mal représentée dans le premier plan factoriel ? Justifier à partir des tableaux des résultats et à partir de la représentation graphique.

5) On poursuit l'étude en comparant les scores de typicalité observés dans les 3 cultures, pour chacune des catégories. Comme précédemment, les données sont traitées par des ACP. Les individus sont les 40 odorants et les variables sont les 3 scores de la catégorie dans les 3 cultures.

Pour les catégories "fruit" et "nature", les scores de typicalité observés sur les premiers odorants sont les suivants :

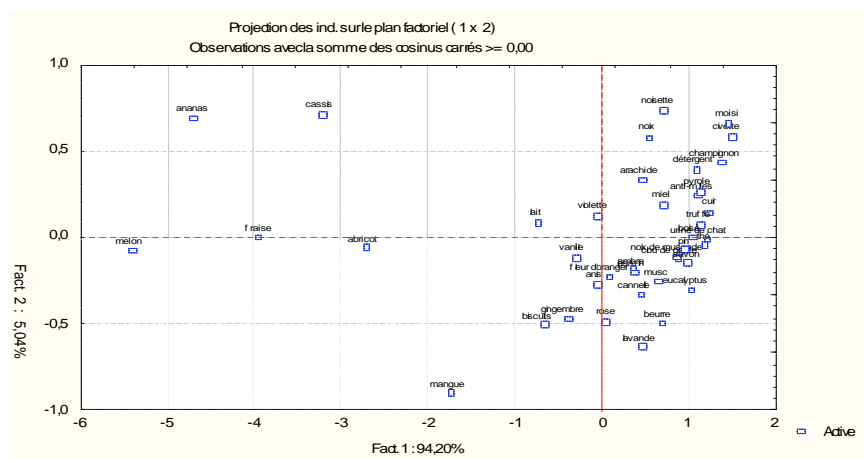
	fruit-FR	fruit-US	fruit-VN	nature-FR	nature-US	nature-VN
ambre	1,63	1,58	2,67	2,5	1,58	2,73
anis	2,13	1,63	2,93	2,6	1,26	2,2
abricot	4	3,95	4,07	2,8	1,53	2,67
cassis	4,5	5	3,77	2,57	1,68	2,2
beurre	1,3	1,11	2,73	2,1	2,26	2,13
urine de chat	1,07	1,05	2,13	1,93	3,26	2,17
cannelle	1,57	1,37	2,73	2,6	1,53	3,77
civette	1,03	1,26	1,57	2,43	3,89	2,07
clou de girofle	1,23	1,26	2,37	2,77	1,68	2,23

Pour les deux ACP correspondant à ces deux catégories, les représentations graphiques des variables selon les 2 premiers axes factoriels sont les suivantes :



a) Pour laquelle de ces deux catégories semble-t-il y avoir le plus de différences selon la culture ? Justifier.

Pour la catégorie "fruit", le diagramme des individus est le suivant :



b) Du point de vue des odorants, quelles sont les différences qui semblent apparaître selon les cultures, pour cette catégorie ?

Indications de réponses

1) Seules les trois premières valeurs propres sont supérieures à 1 et donc, seules les trois premières composantes principales représentent une inertie relative supérieure à la moyenne ($100\%/11=9,09\%$). Cette règle nous conduirait donc à n'étudier que trois composantes. Cependant, on observe une brusque décroissance des valeurs propres entre la 5^e et la 6^e valeurs propres (méthode du coude). Il paraît donc intéressant d'étudier les cinq premières composantes principales. Ainsi, 3 ou 5 composantes pourront être étudiées. Notons qu'il n'y a en revanche aucune raison de s'arrêter à la 4^{ème} valeur propre.

2) Sur le premier axe, on obtient pour les individus :

-	+
moisi (12%)	fraise (9%)
civette (10%)	biscuit (8%)
champignon (8%)	melon (7%)
urine de chat (6%)	abricot (5%)

De même, on obtient pour les variables :

-	+
moisissure (19%)	sucrerie (19%)
animal (14%)	fruit (14%)
	pâtisserie (11%)

On voit que cet axe oppose des odeurs naturelles et animales (moisissures, urine de chat) à des odeurs de fruits ou de desserts (fruits, pâtisseries). Les premières sont souvent désagréables, mais ce n'est pas systématique (la civette, par exemple, est utilisée en parfumerie). On notera qu'il s'agit, aux deux extrémités de l'axe, d'odeurs naturelles, traditionnelles (par opposition aux odeurs de détergents que nous retrouverons sur l'axe 2).

3) Sur le deuxième axe, on obtient pour les individus :

-	+
noix (7%)	savon (13%)
noisette (6%)	rose (7%)
cannelle (5%)	détergent (6%)
	fleur d'orange (5%)
	lavande (5%)

De même, on obtient pour les variables :

-	+
pâtisserie (14%) épice (12%)	nettoyant (32%) cosmétique (19%) fleur (10%)

On voit que les deux catégories "nettoyant" et "cosmétique", toutes deux associées à la partie positive de l'axe, contribuent à plus de 50% de son inertie. Dans une certaine mesure, cet axe peut être qualifié d'unidimensionnel. Sa partie positive est associée à des odeurs de produits chimiques domestiques et à des odeurs artificielles couramment utilisées dans de tels produits (lavande, fleur d'oranger, etc). Il s'agit là plutôt d'odeurs obtenues artificiellement par des procédés chimiques. A l'opposé, sur la partie négative, on retrouve des odeurs très marquées (cannelle, noix), entrant dans la catégorie "épices", mais qui ne sont pas utilisées dans la chimie à usage domestique.

4) La réponse est trouvée en examinant la colonne " $\text{Cos}^2 1\&2$ " du tableau des "Saturation, contributions et qualités de représentation des variables". La variable la plus mal représentée est la catégorie "Pharmacie" (qualité de représentation de 0,07 dans le premier plan factoriel). On peut aussi remarquer qu'il s'agit du vecteur dont l'extrémité est la plus proche de l'origine sur le graphique des variables.

5) a) Pour la catégorie "fruit", on constate que le premier axe représente une très forte part de la dispersion (94% de la variance) et les deux premiers axes la quasi-totalité (plus de 99%). Les trois variables fruit-US, fruit-FR et fruit-VN sont très fortement corrélées au premier axe et fortement corrélées positivement entre elles. On constate malgré tout une certaine différence entre fruit-VN et les deux autres variables, fruit-VN s'opposant aux deux autres sur le deuxième axe.

Pour la catégorie "nature", au contraire, le premier axe ne représente plus que 48% de l'inertie ; une part importante de l'inertie se retrouve sur le 2^{ème} axe (35%), ce qui montre que la situation est nettement plus complexe. Les qualités de représentation des variables sont moins bonnes que dans le cas de la catégorie "fruit". On voit cependant que nature-US et nature-FR sont assez fortement corrélées positivement, alors que nature-VN est sans doute faiblement corrélée aux deux précédentes (angle proche de l'angle droit), et il est même vraisemblable que nature-VN et nature-US soient négativement corrélées.

Il semble donc que la catégorie "fruit" diffère peu selon la culture, alors que des différences importantes apparaissent pour la catégorie "nature", particulièrement entre les USA et le Vietnam.

b) On retrouve dans la partie positive de l'axe 1 les individus peu liés à la catégorie "fruit". Les individus les plus caractéristiques de la catégorie se projettent sur la partie négative de l'axe : ananas, cassis, melon, fraise, abricot, mangue, etc.

Leurs projections selon le 2^{ème} axe nous renseignent alors sur les différences entre les cultures : melon, fraise, abricot ont une coordonnée proche de 0 sur le 2^{ème} axe, et leur catégorisation comme "fruit" est sans doute équivalente dans les trois cultures. En revanche, cassis et ananas sont plus caractéristiques de "fruit" en France et aux USA qu'au Vietnam, et au contraire, la mangue (peut-être aussi le gingembre) est plus caractéristique de "fruit" au Vietnam qu'en France ou aux USA.