
FAST, a novel “Factorial Approach” for Sorting Task data

Marine Cadoret, Sébastien Lê, Jérôme Pagès
AGROCAMPUS OUEST, France



9th Sensometrics conference
July 20-23 2008, Canada

Introduction

- ✱ Sorting task (or categorization) consists in grouping objects in function of their resemblances.
- ✱ Following this task, a verbalization task can also be asked to describe the groups (“qualified” categorization).

Data

- ★ 98 consumers carried out a “qualified” categorization on 12 luxury perfumes:



Angel



Lolita
Lempicka



L'Instant



Cinéma



J'adore
(ET)



J'adore
(EP)



Shalimar



Aromatics
Elixir



Coco
Mademoiselle



Chanel n°5



Pure
Poison



Pleasures

« gourmand,
vanilla, woody »



« spicy, aldehyde »



« white flower,
vanilla, orange »



« flower,
floral,
green »

« oriental,
showy,
woody,
Patchouli oil »



Data

produit	juge 12	juge 13	juge 14	juge 15	juge 16
Angel	1	4	1	5	2
Aromatic Elixir	3	3	5	2	1
Chanel n°5	4	3	4	1	3
Cinéma	2	5	6	4	2
Coco Mademoiselle	1	5	2	4	3
J'adore (EP)	1	6	2	3	3
J'adore (ET)	2	6	2	3	3
L'instant	1	4	6	2	4
Lolita Lempicka	1	5	1	5	2
Pleasures	3	4	6	3	4
Pure Poison	1	1	2	4	4
Shalimar	2	2	3	2	1

For example, Consumer 12 put in the group 1 Angel, Coco Mademoiselle, J'adore (EP), L'instant, Lolita Lempicka and Pure Poison

Data

produit	juge 12	juge 13	juge 14	juge 15	juge 16
Angel	fleuri doux	fruité fort	vanillé épicé esprit des îles	à manger sucré	nourriture épice
Aromatic Elixir	fort homme	capiteux grand-mère	rude fort	le vieux	ménager cire
Chanel n°5	Gr 4	capiteux grand-mère	toilettes	savon	connu classique
Cinéma	fleuri artificiel herbe	fruité moyen	sucré	doux	nourriture épice
Coco Mademoiselle	fleuri doux	fruité moyen	douceur fleuri	doux	connu classique
J'adore (EP)	fleuri doux	sucré faible	douceur fleuri	fleuri	connu classique
J'adore (ET)	fleuri artificiel herbe	sucré faible	douceur fleuri	fleuri	connu classique
L'instant	fleuri doux	fruité fort	sucré	le vieux	fleuri
Lolita Lempicka	fleuri doux	fruité moyen	vanillé épicé esprit des îles	à manger sucré	nourriture épice
Pleasures	fort homme	fruité fort	sucré	fleuri	fleuri
Pure Poison	fleuri doux	acidulé désodorisant	douceur fleuri	doux	fleuri
Shalimar	fleuri artificiel herbe	fort lavande eau de cologne	renfermé agressif	le vieux	ménager cire

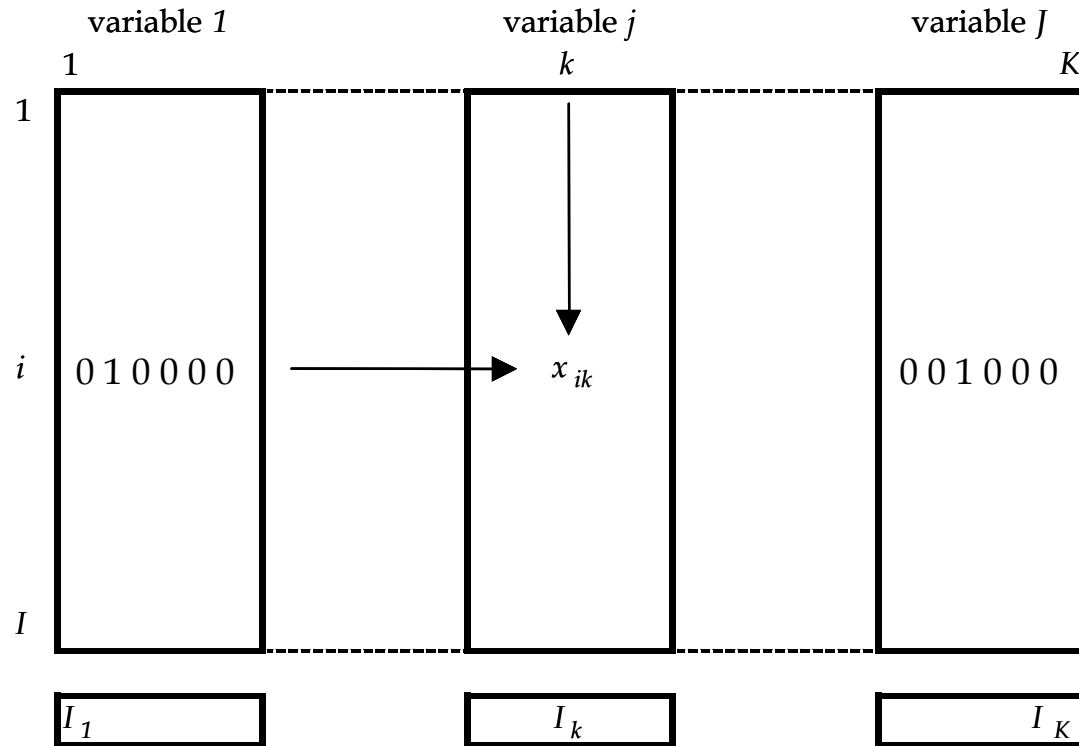
Each consumer can be considered as a categorical variable

Let's run MCA on this data table!

The approach

Why does it work?

Disjunctive table



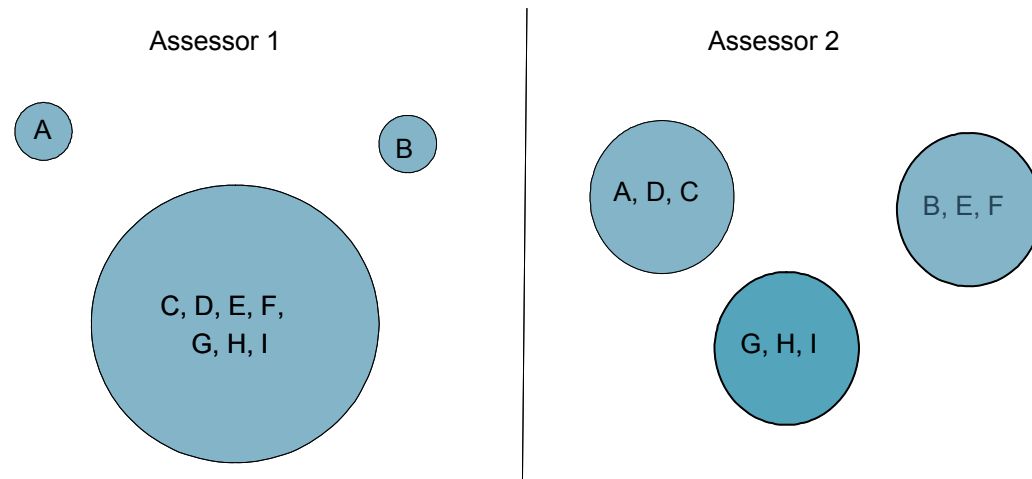
x_{ik} is equal to 1 if perfume i belongs to group k ; I_k is the number of perfumes in the group k

Distance between individuals

$$d_{i,i'}^2 = \frac{I}{J} \sum_{k=1}^K \frac{(x_{ik} - x_{i'k})^2}{I_k}$$

- * The distance between two products is null if they were put systematically together.
- * Two products are all the more close (resp. distant) that they were put together by a great (resp. few) number of consumers.

Objects categorized by two assessors



- * They have made three groups each (a disc represents a group).
- * The two assessors distinguished between products A and B, but in a more remarkable way for assessor 1

Distance between categories

$$d_{k,k'}^2 = I \sum_{i=1}^I \left(\frac{x_{ik}}{I_k} - \frac{x_{ik'}}{I_{k'}} \right)^2$$

- * Two categories/descriptions are all the more distant that they have few common individuals. In other words, that the number of individuals that were put either in k , either in k' is big.

$$I_{k \neq k'}$$

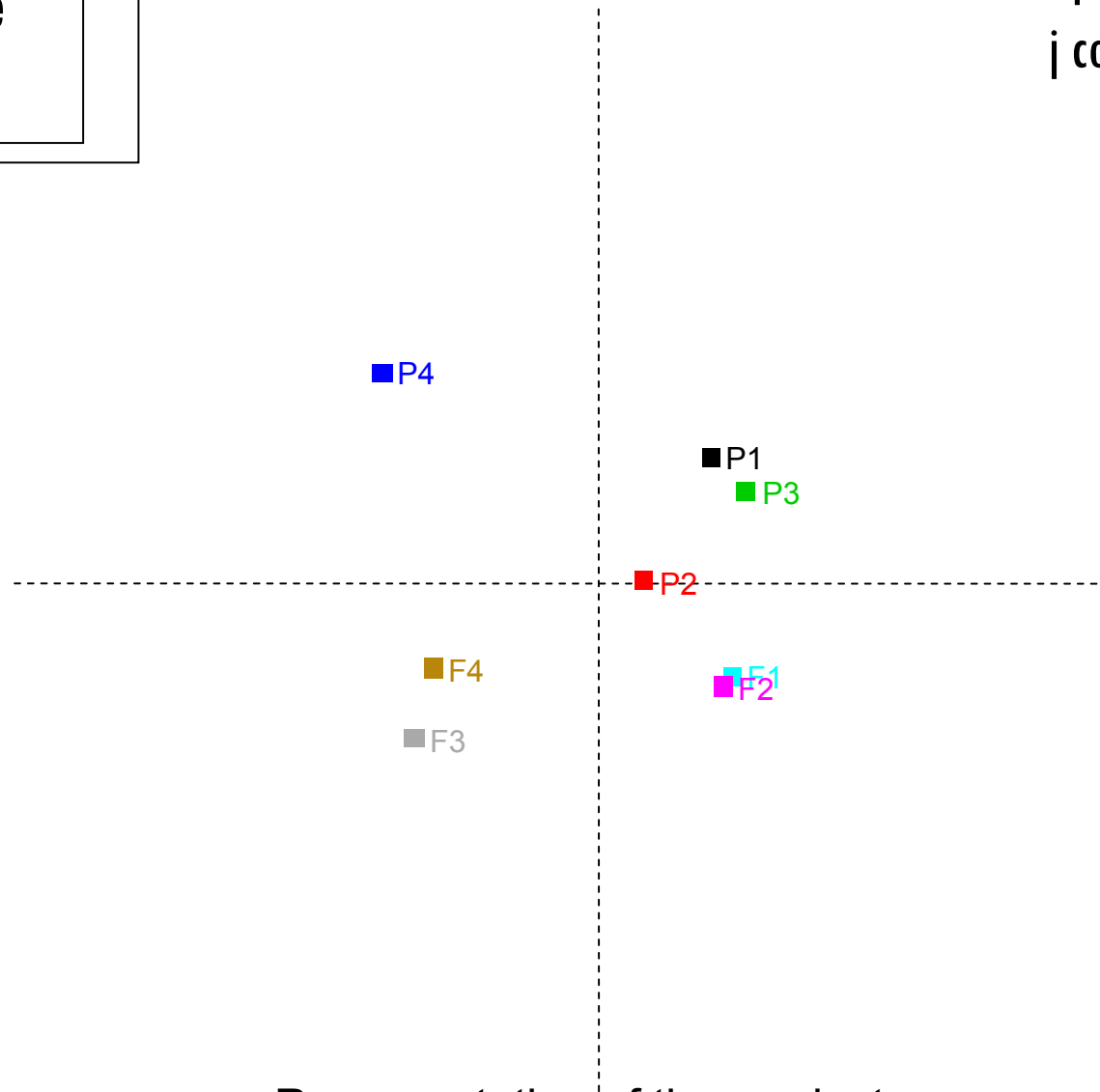
The approach

In how can it be specific to sensory analysis?

Data table

$i \times j$

i products
 j consumers

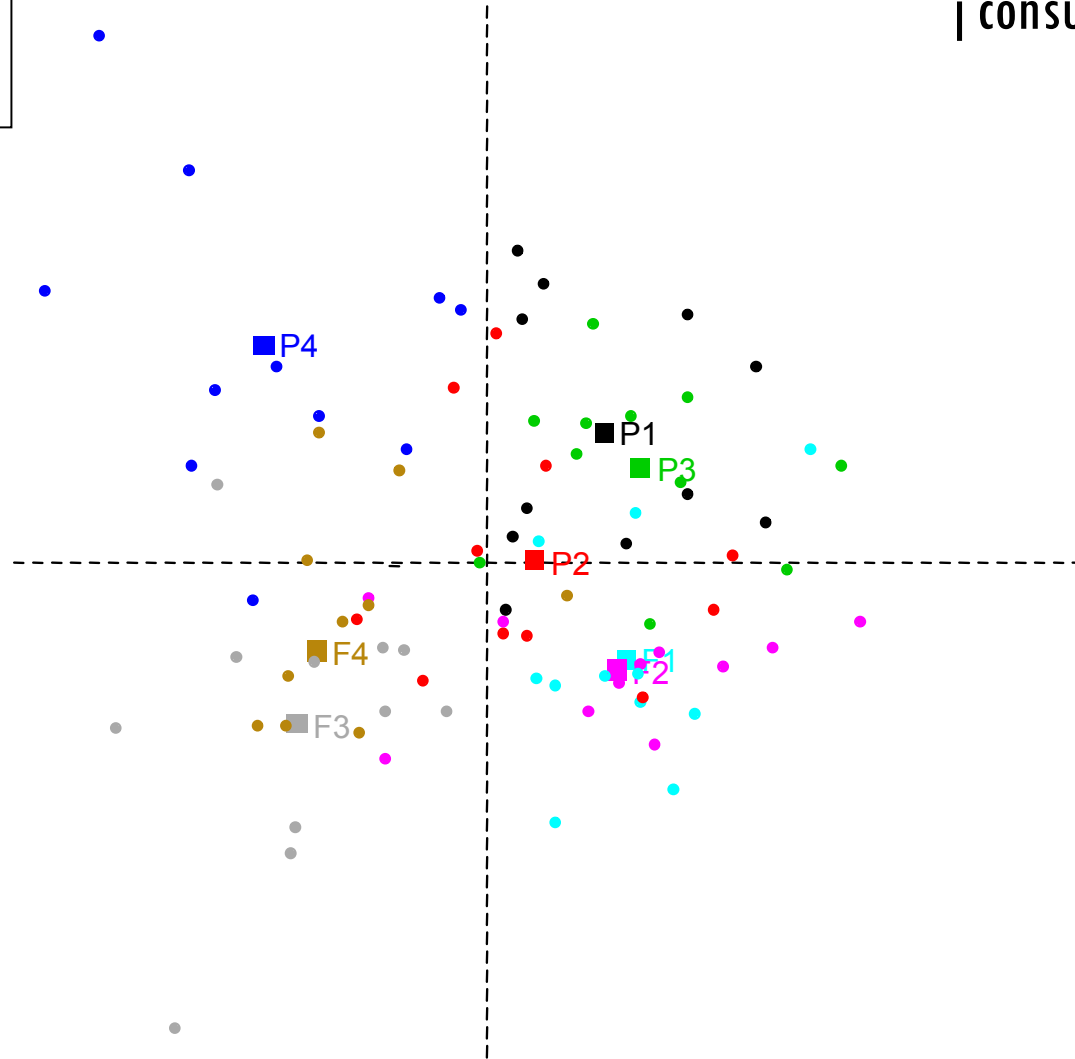


Representation of the products

Data table

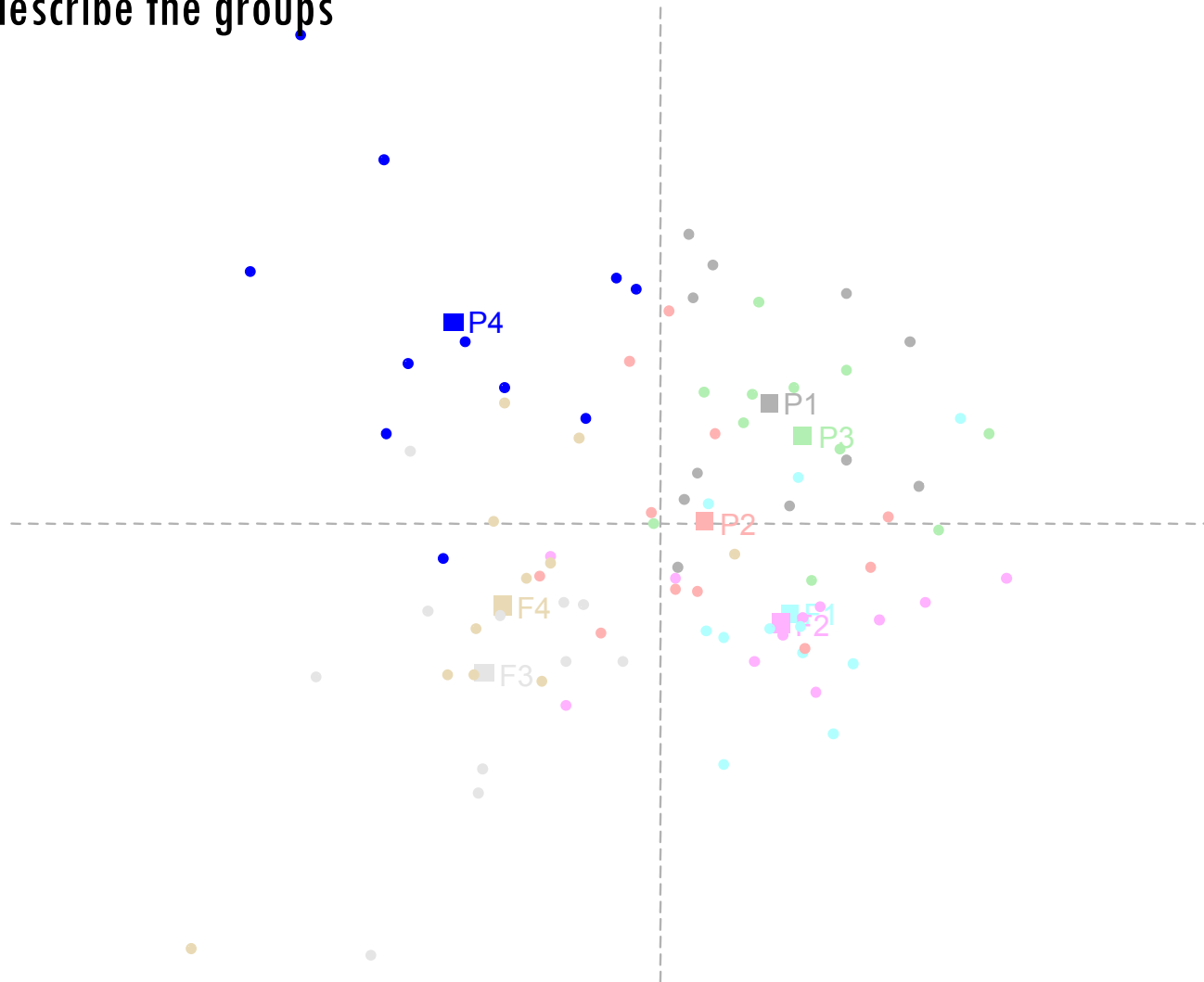
$i \times j$

i products
 j consumers



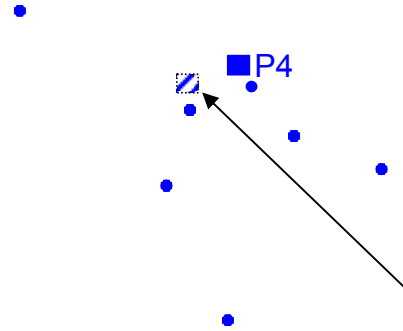
Superimposed representation of the products and their descriptions

P4 is at the barycentre of the words used to describe the groups



Data table
 $i \times j$

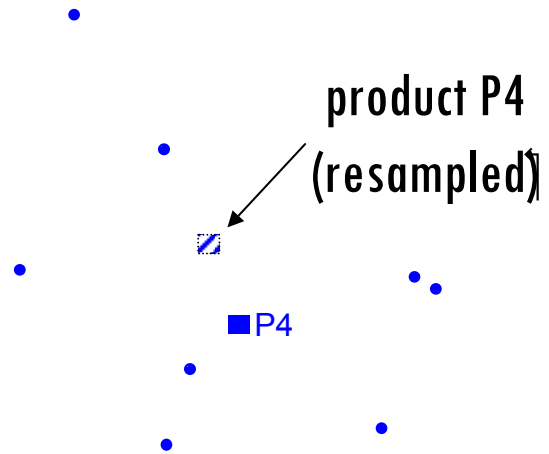
Panelist's words
(resampled)



product P4
(resampled)

Data table

$i \times j$

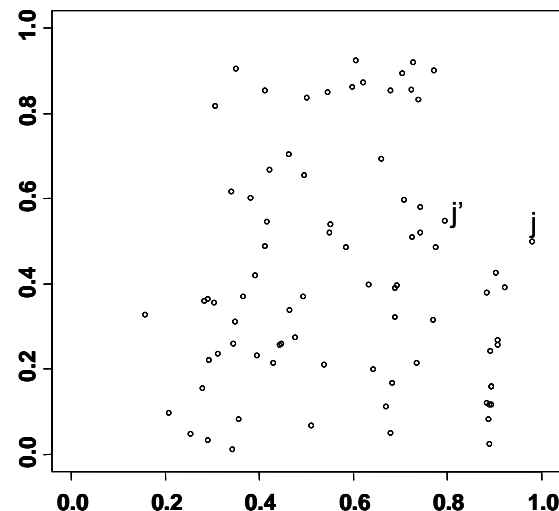


Confidence ellipses around products



Representation of the consumers

- ✱ Using Multiple Factor Analysis
- ✱ 2 consumers are all the more close as they carried out similar categorizations
- ✱ Consumers representation linked to the ones of products and words



Results

« FAST » function implemented
in **SensoMineR**

Co-occurrences among the perfumes

	Shalimar	Aromatics Elixir	Chanel n°5	Angel	Lolita Lempicka	Cinéma	L'instant	Pure Poison	Coco Mademoiselle	Pleasures	J'adore (EP)	J'adore (ET)	Alone
Shalimar	98	42	30	21	9	10	13	11	9	6	6	7	24
Aromatics Elixir	42	98	51	27	6	8	13	12	12	11	12	7	6
Chanel n°5	30	51	98	15	8	9	10	21	11	14	12	14	17
Angel	21	27	15	98	36	18	14	10	10	11	11	12	13
Lolita Lempicka	9	6	8	36	98	42	22	18	21	18	18	18	7
Cinéma	10	8	9	18	42	98	26	28	30	22	23	24	5
L'instant	13	13	10	14	22	26	98	25	20	23	28	22	9
Pure Poison	11	12	21	10	18	28	25	98	33	30	29	28	7
Coco Mademoiselle	9	12	11	10	21	30	20	33	98	28	28	38	8
Pleasures	6	11	14	11	18	22	23	30	28	98	38	48	8
J'adore (EP)	6	12	12	11	18	23	28	29	28	38	98	56	2
J'adore (ET)	7	7	14	12	18	24	22	28	38	48	56	98	2

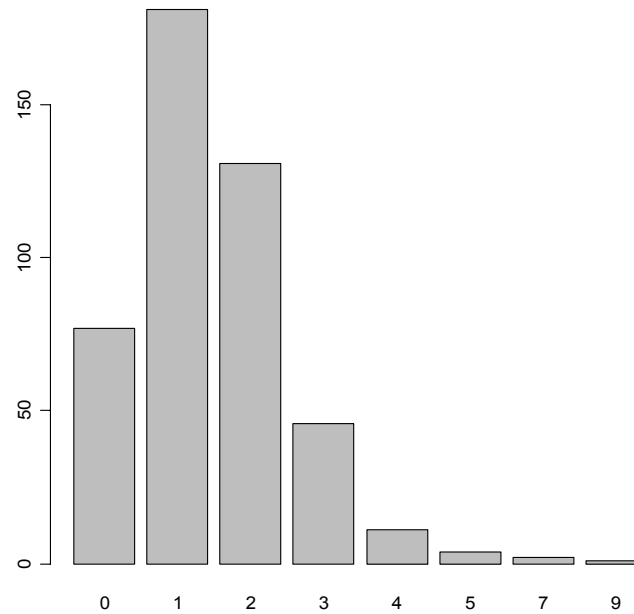
Some textual analysis

- ★ Description of the products Angel and Shalimar sorted by descending order of significativity

Angel	intern %	global %	p-value
vanillé	2,58%	0,55%	0,007
épicé	3,23%	0,94%	0,011
sucré	9,03%	5,18%	0,025
fort	10,32%	6,45%	0,036

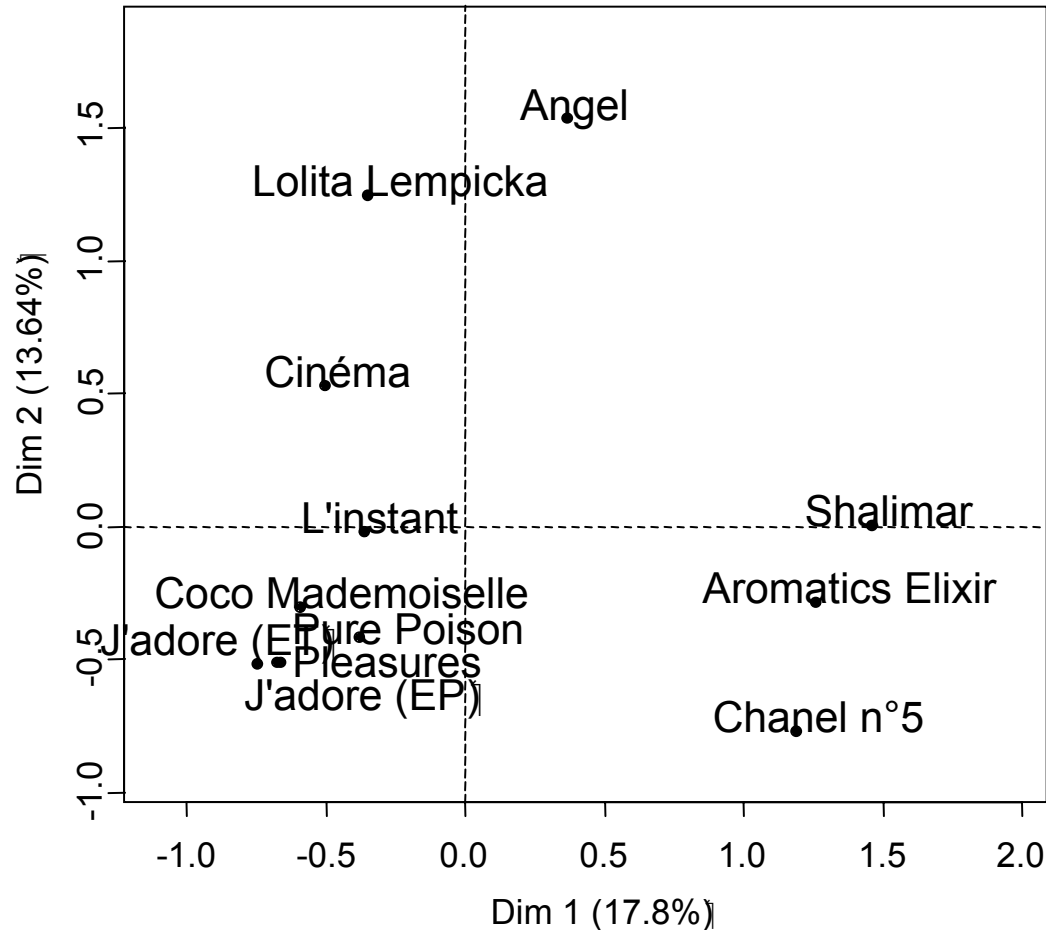
Shalimar	intern %	global %	p-value
fort	12,42%	6,45%	0,003
agressif	3,92%	1,05%	0,004
mentholé	1,31%	0,11%	0,007
oriental	1,31%	0,17%	0,020
vieux	2,61%	0,77%	0,025
médicament	1,31%	0,22%	0,038
poivré	1,96%	0,55%	0,045
masculin	1,96%	0,55%	0,045

Number of words per group



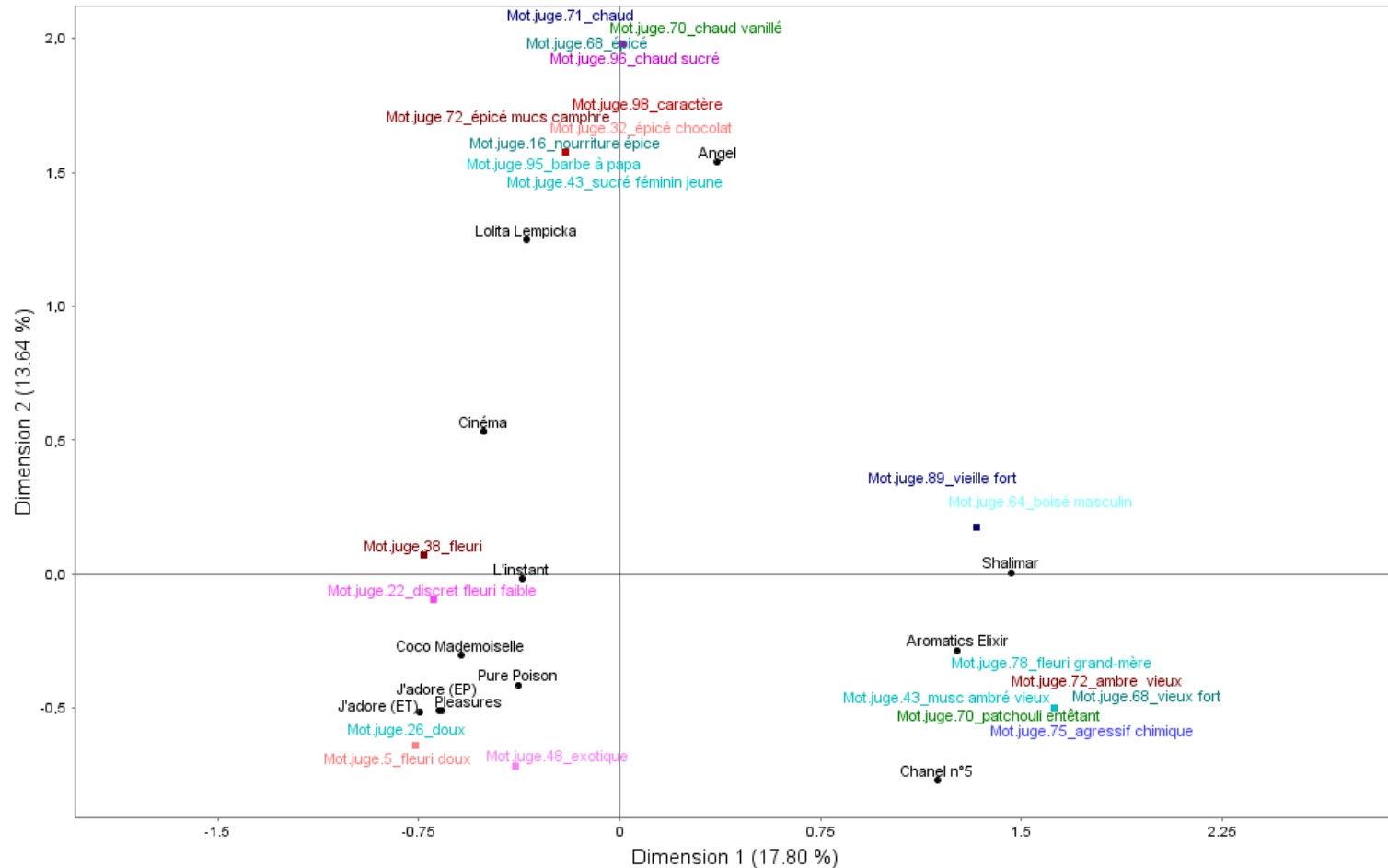
Representation of the perfumes

MCA factor map



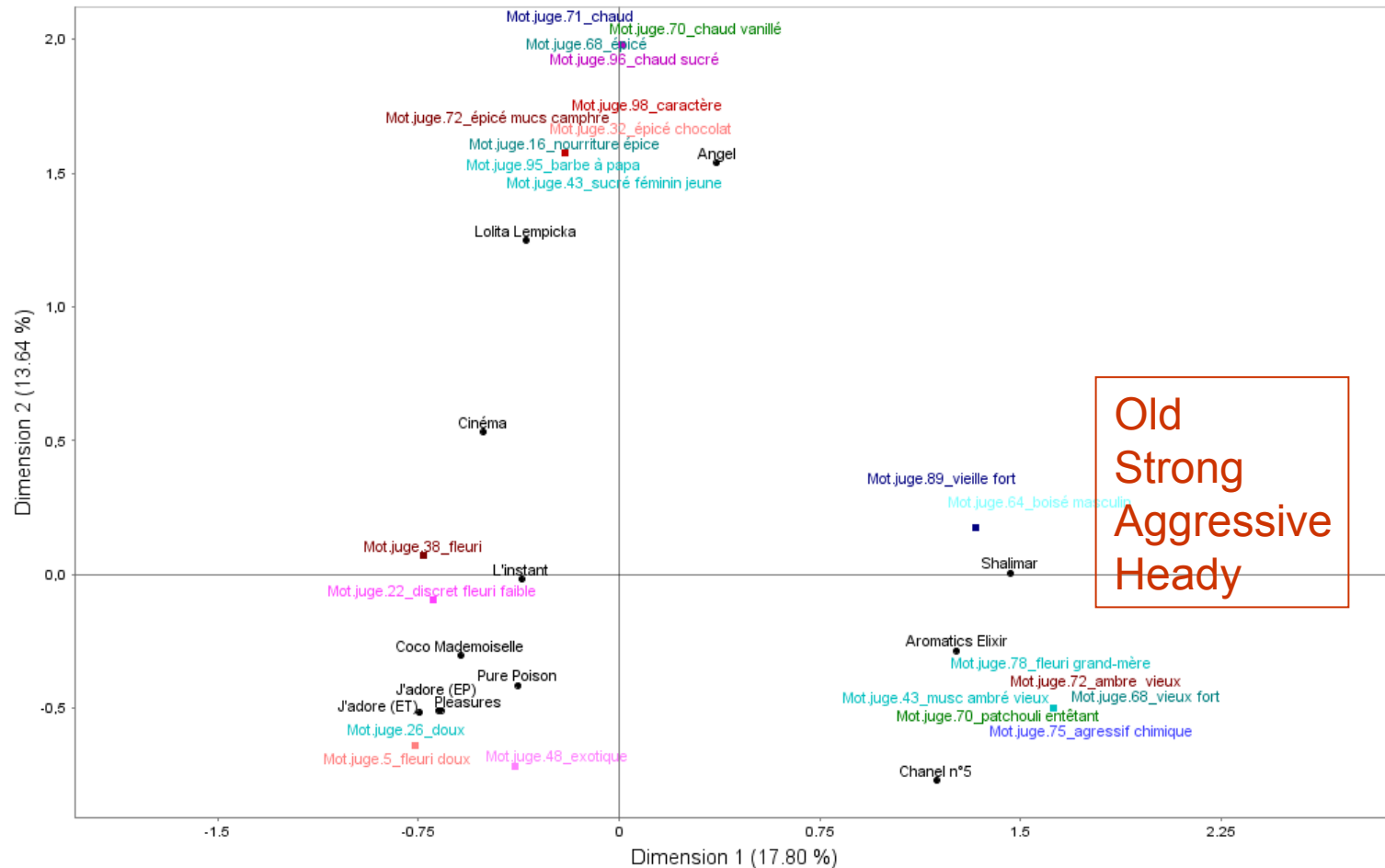
- * Plan defined by dimensions 1 and 2 of MCA

Representation of the words



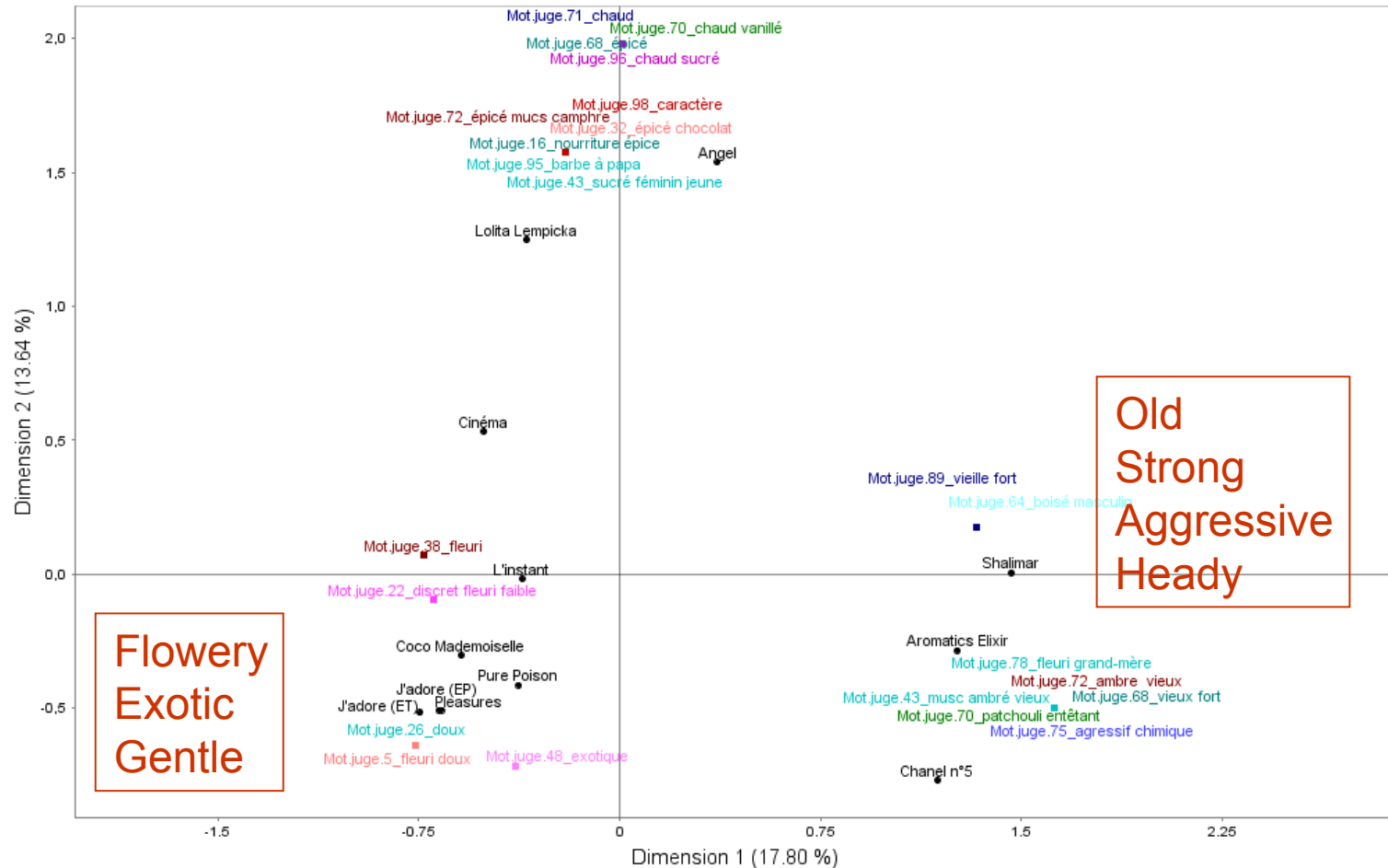
* Plan defined by dimensions 1 and 2 of MCA

Representation of the words



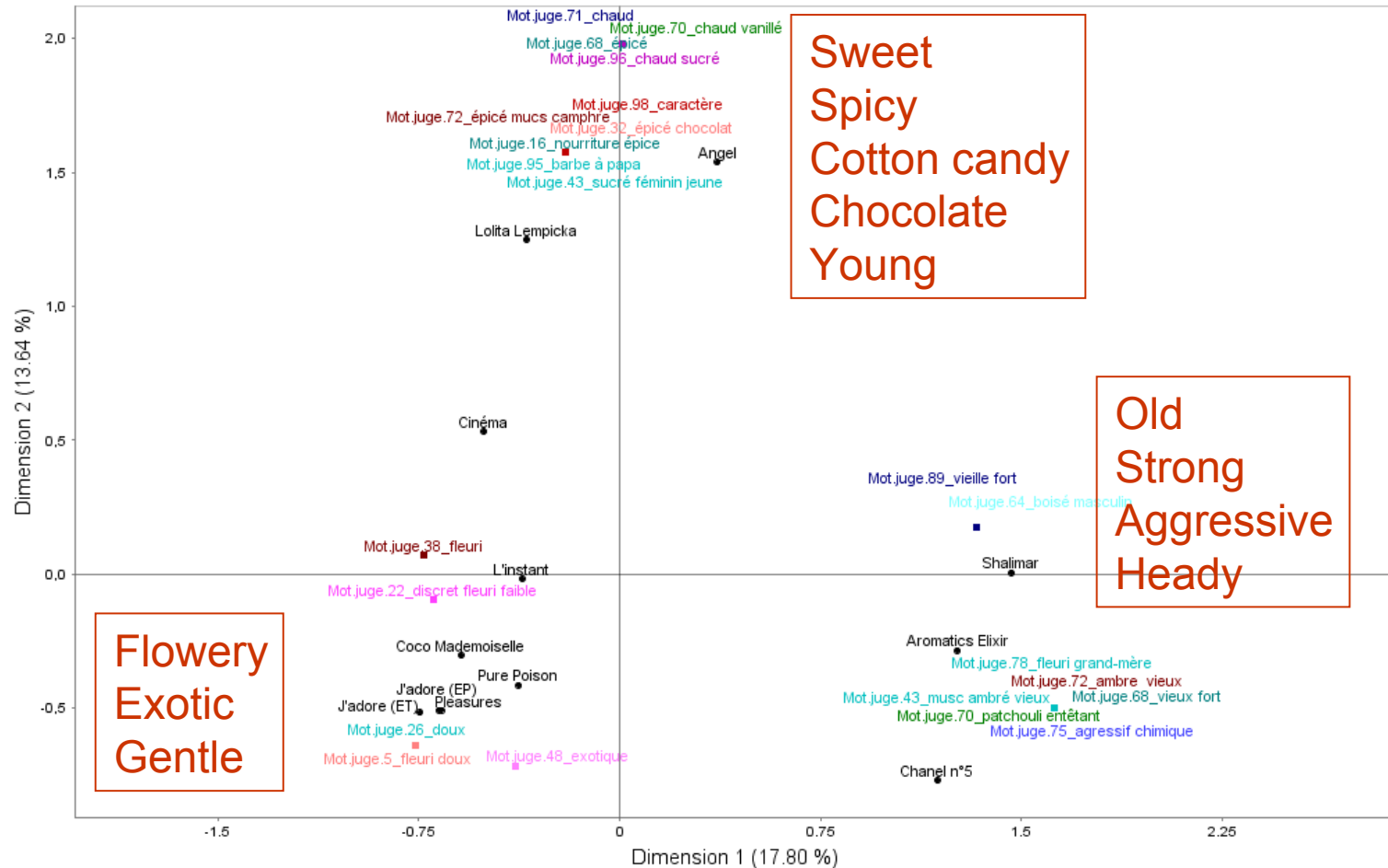
* Plan defined by dimensions 1 and 2 of MCA

Representation of the words



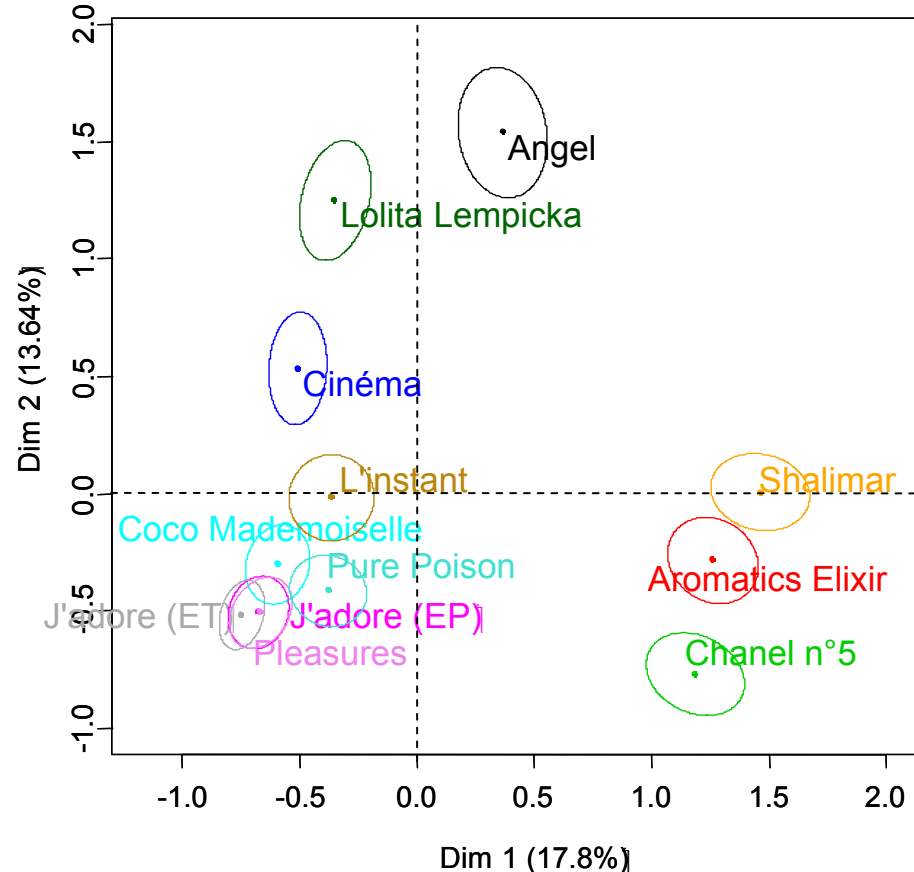
* Plan defined by dimensions 1 and 2 of MCA

Representation of the words



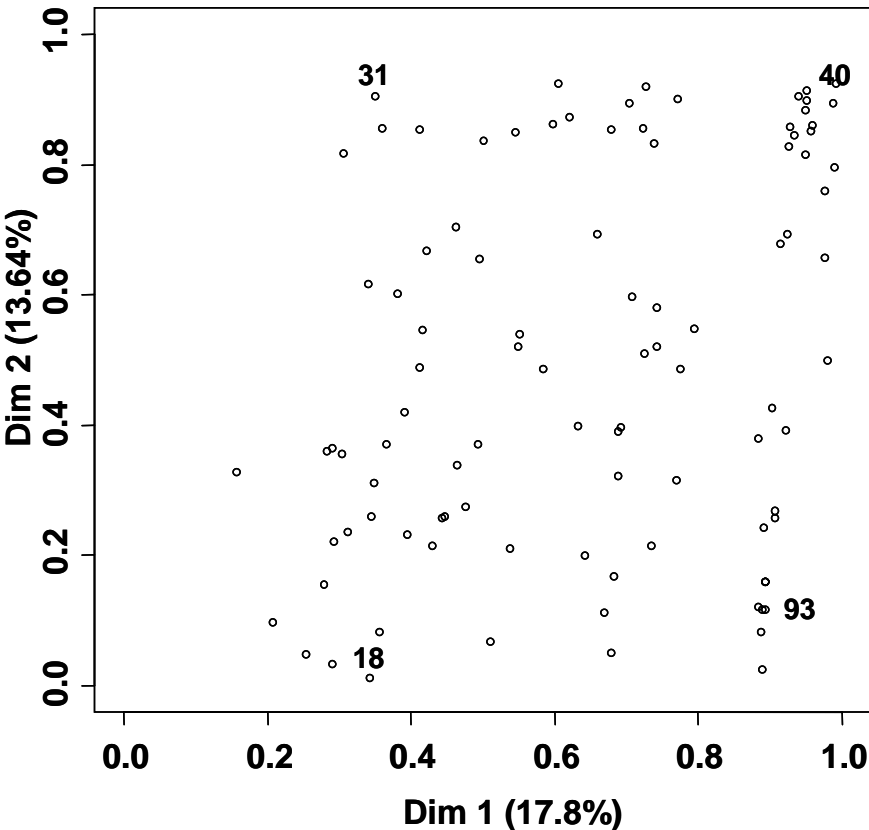
* Plan defined by dimensions 1 and 2 of MCA

Representation of the perfumes and their respective confidence ellipses



* Plan defined by dimensions 1 and 2 of MCA

Representation of the consumers



	juge 18	juge 31	juge 40	juge 93
Shalimar	2	1	4	4
Aromatics Elixir	2	2	5	3
Chanel n°5	3	4	5	3
Coco Mademoiselle	3	1	2	2
J'adore (EP)	1	1	1	1
J'adore (ET)	3	1	1	2
L'instant	2	1	2	1
Pleasures	3	1	1	1
Pure Poison	1	2	2	2
Angel	3	5	6	1
Cinéma	3	3	3	2
Lolita Lempicka	1	3	3	2

* Plan defined by dimensions 1 and 2 of MFA

The logo for SensoMineR features the word "SensoMineR" in a stylized, 3D font. The letters "Senso" are yellow with a blue outline, "Mine" is blue with a yellow outline, and "R" is blue with a yellow outline. A grey, metallic-looking ring encircles the "R".

<http://sensominer.free.fr>

Journal of sensory studies **SensoMineR** a package for sensory data analysis

The logo for FactoMineR features the word "FACTOMINE" in a blue, 3D font with a grey outline, followed by a large "R" in the same style. A grey, metallic-looking ring encircles the "R".

<http://factominer.free.fr>

Journal of statistical software **FactoMineR**: an R package for multivariate analysis