

**Journée Annuelle de Conférences SFAS
01/04/16**

SensStat

**Plaire à tout le monde ?!
L'importance de la segmentation**

Christine Urbano et Eric Teillet

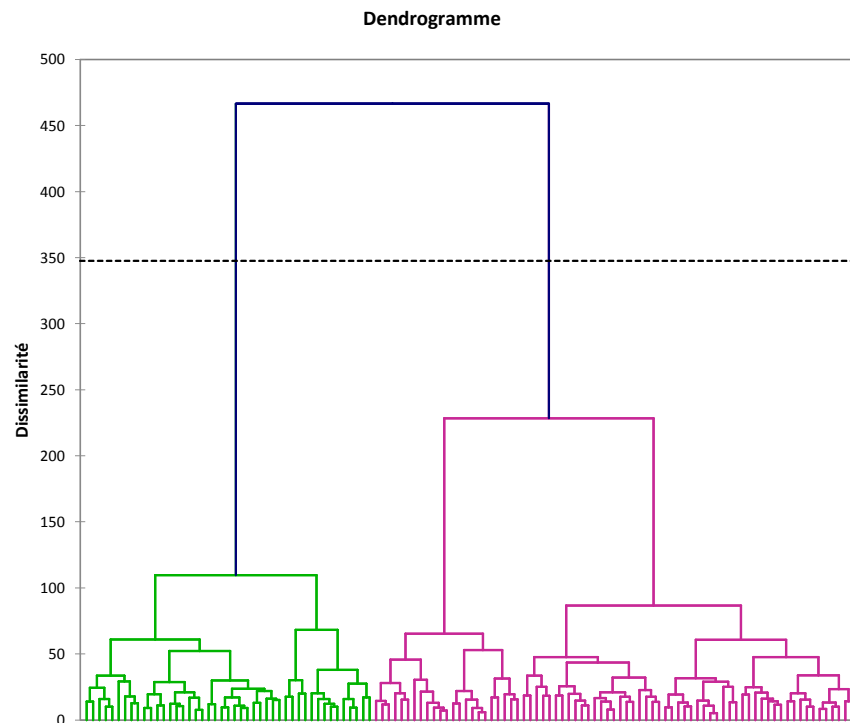


Quelques exemples

Classification des sujets



La classification ascendante hiérarchique (CAH) est une méthode statistique permettant de regrouper les consommateurs qui ont les mêmes préférences pour les mêmes produits (et donc de séparer les consommateurs ayant des préférences différentes).

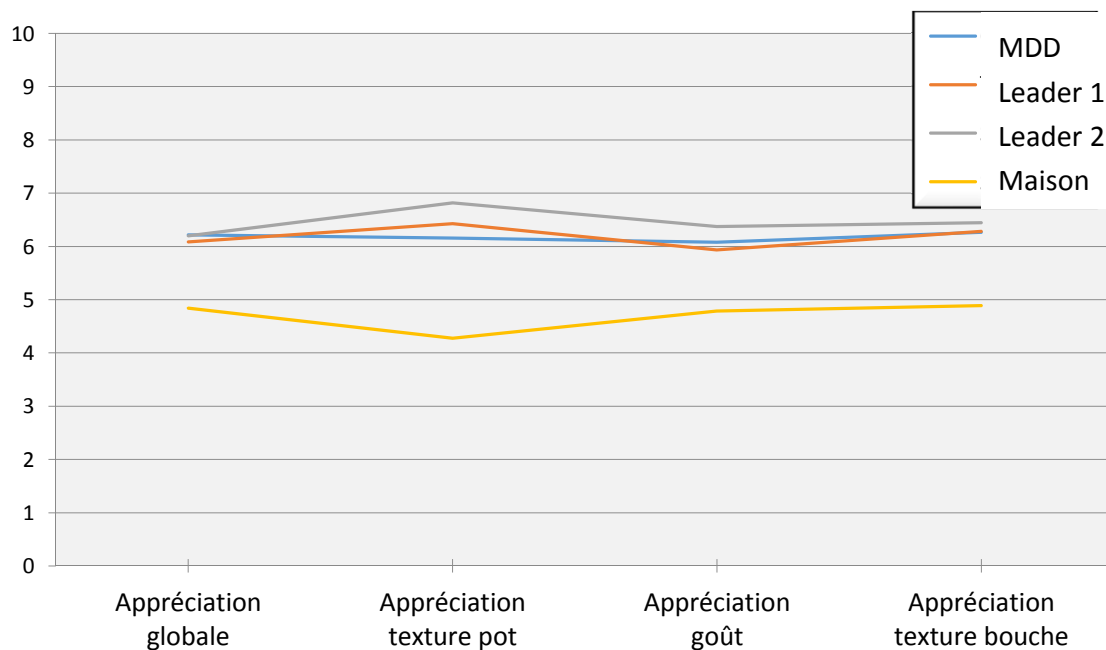


Ici les sujets recrutés semblent se dissocier en 2 classes principales.

Exemple 1 : comparaison de yaourts commerce vs maison



| | P value Prod | MDD | Leader 1 | Leader 2 | maison |
|----------------------|--------------|-------------|-------------|-------------|-------------|
| Appréciation globale | <.0001 | 6.2 ± 2.5 a | 6.1 ± 2.7 a | 6.2 ± 2.4 a | 4.8 ± 2.9 b |
| Texture dans le pot | <.0001 | 6.2 ± 2.3 a | 6.4 ± 2.5 a | 6.8 ± 2.2 a | 4.3 ± 2.9 b |
| Goût | <.0001 | 6.1 ± 2.6 a | 5.9 ± 2.7 a | 6.4 ± 2.2 a | 4.8 ± 2.9 b |
| Texture en bouche | <.0001 | 6.3 ± 2.5 a | 6.3 ± 2.6 a | 6.4 ± 2.4 a | 4.9 ± 3.0 b |



Moyenne ± écart-type. P value Prod = effet produit dans l'analyse de la variance.

La différence entre modalités reliées par une même lettre n'est pas significative. Comparaison de moyennes : Newman Keuls à 5%.

→ Globalement, le produit maison est moins apprécié que les trois autres produits.

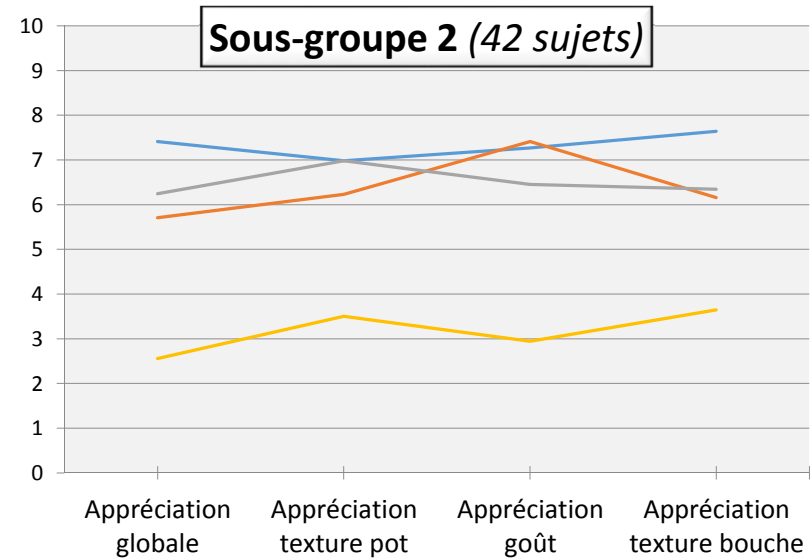
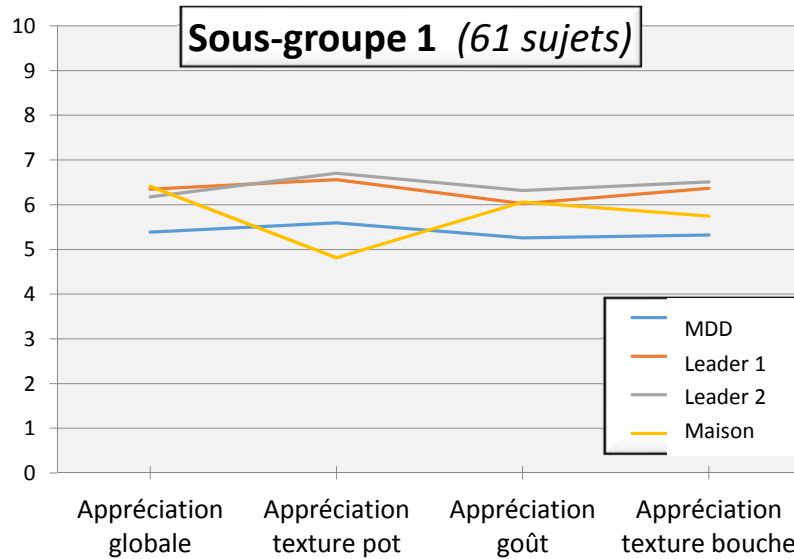
→ Sa texture dans le pot, sa texture en bouche et son goût sont moins appréciés que ceux des 3 autres produits.

→ Il n'y a pas de différence significative entre les 3 autres produits.

Exemple 1 : Segmentation des sujets



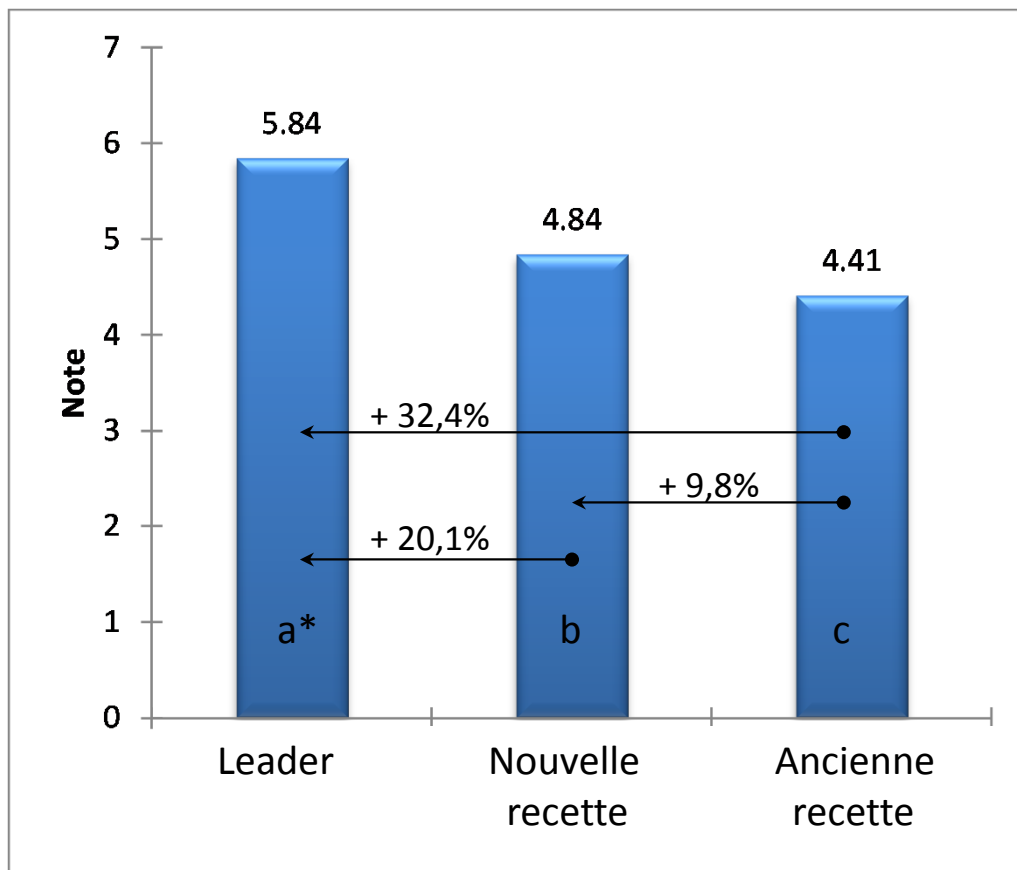
Segmentation par CAH avec indice de Ward, sujets centrés.



La classification ascendante hiérarchique a permis une segmentation des sujets en 2 sous-groupes, pour lesquels aucun lien direct avec les caractéristiques connues des sujets n'a pu être identifié.

- Le sous-groupe 1 (59% des sujets), apprécie moins la texture du produit maison dans le pot, mais cela n'a pas d'impact négatif sur l'appréciation globale.
- Le sous-groupe 2 (41% des sujets), apprécie nettement moins le produit maison que les 3 autres produits, sur l'ensemble des descripteurs.

Exemple 2 : ancienne et nouvelle recette vs leader



La nouvelle recette semble augmenter significativement les préférences des consommateurs, même si le leader reste préféré.

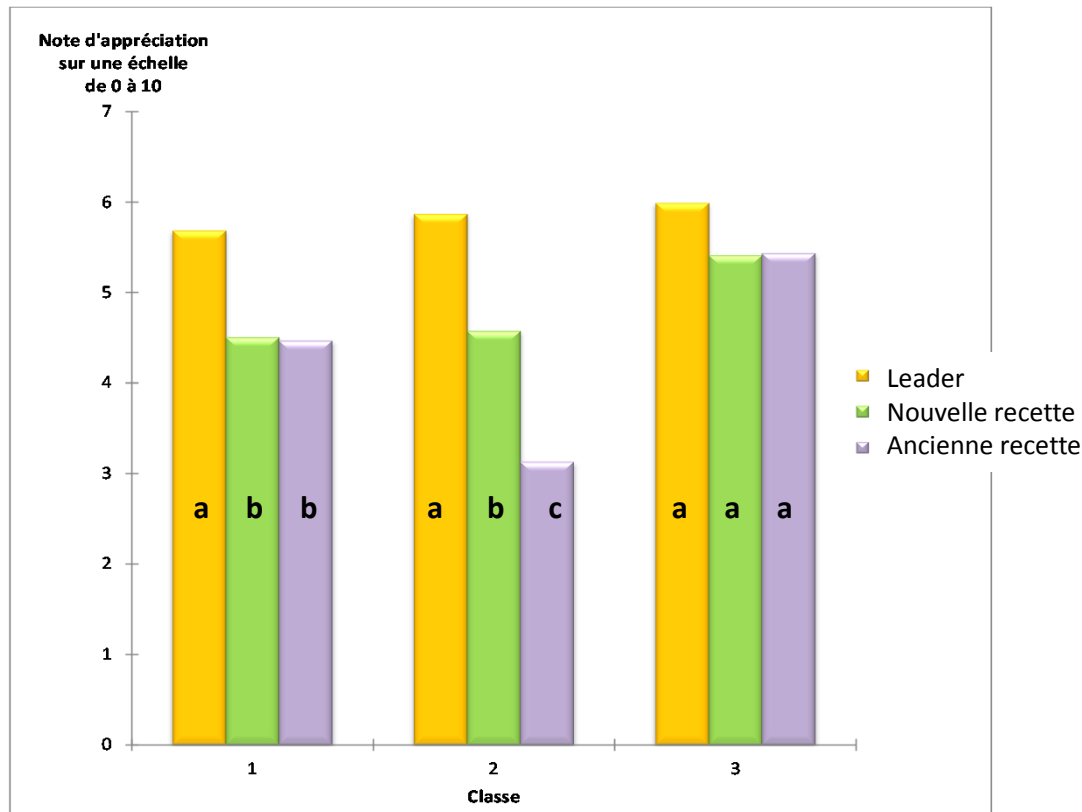
* La différence entre modalités reliées par une même lettre n'est pas significative.
Comparaison de moyennes : PPDS à 5%.

Exemple 2 : segmentation des résultats



La classification montre 3 groupes de consommateurs :

- classe 1 : 47 sujets (36.2%)
- classe 2 : 38 sujets (29.2%)
- classe 3 : 45 sujets (34.6%)



* La différence entre modalités reliées par une même lettre n'est pas significative.
Comparaison de moyennes : PPDS à 5%.

Conclusions :

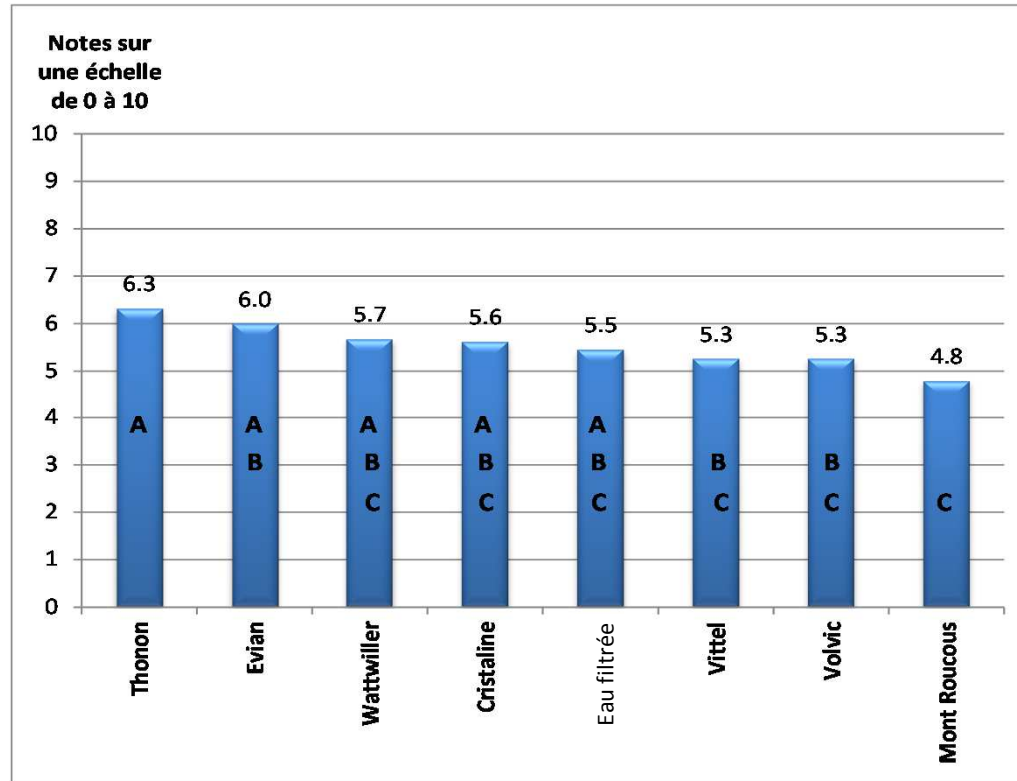
La classe 3 (34.6%) apprécie globalement de la même façon les 3 produits.

Les classes 1 et 2 (qui regroupent 65.4% des consommateurs) préfèrent significativement le leader.

La classe 1 ne différencie pas ancienne et nouvelle recette.

Seule la classe 2 préfère la nouvelle recette par rapport à l'ancienne.

Exemple 3 : comparaison d'une eau filtrée avec des eaux embouteillées



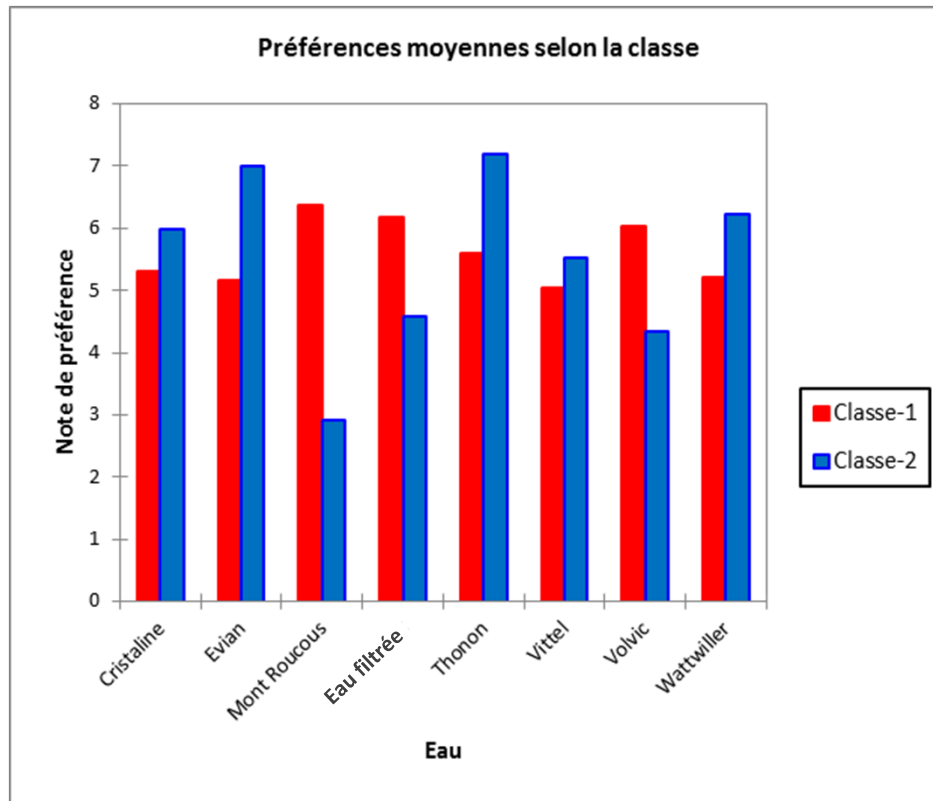
*Test de Newman-Keuls à 5%. La différence entre les produits reliés par une même lettre n'est pas statistiquement significative

Conclusions :

La moyenne obtenue par l'eau filtrée est moins élevée que celle de Thonon, Evian, Wattwiller et Cristaline (les eaux « neutres »).

Cependant, les tests statistiques montrent que les préférences sont globalement homogènes, et l'eau filtrée, bien que légèrement moins appréciée, se trouve dans le groupe de tête.

Exemple 3 : segmentation des résultats



Conclusions :

La classe 1 (54.3% des consommateurs) apprécie l'ensemble des eaux (l'eau du robinet filtrée arrive en 2^{ème} position).


La classe 2 (45.7%) rejette spécifiquement les eaux de faible minéralité : Volvic, eau du robinet filtrée et surtout Mont Roucous.

Exemple 3 : segmentation des résultats




- Lien avec les habitudes de consommation

Effectifs des fréquences de consommation d'eau embouteillée

|  Régulièrement | Occasionnellement | Jamais ou presque | |
|---|-------------------|-------------------|------------|
| Classe 1 | 27 | 23 | 7 |
| Classe 2 | 22 | 19 | 7 |
| Total | 49 (46.7%) | 42 (40%) | 14 (13.3%) |

Effectifs des fréquences de consommation d'eau du robinet

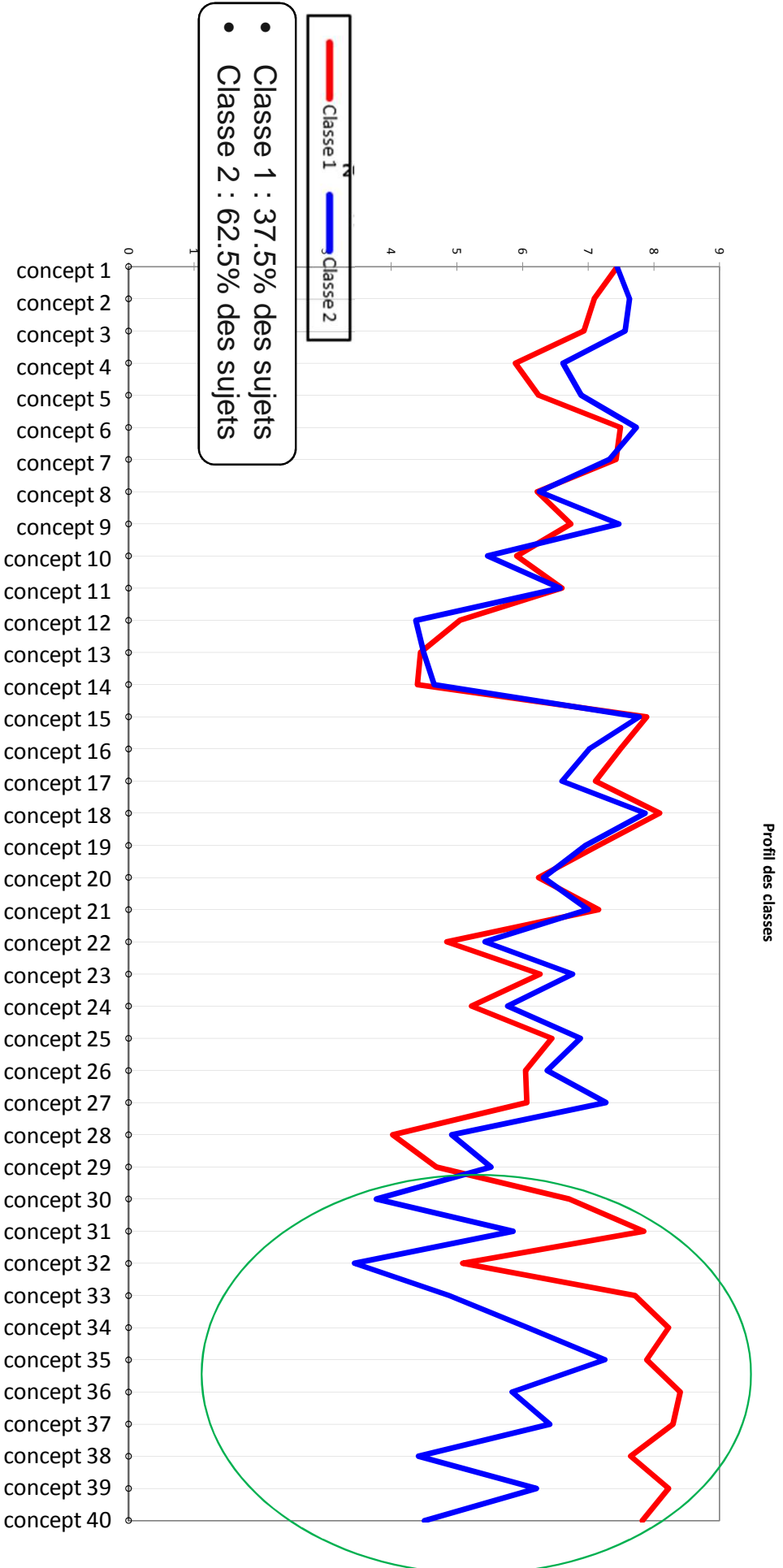
|  Régulièrement | Occasionnellement | Jamais ou presque | |
|--|-------------------|-------------------|----------|
| Classe 1 | 33 | 13 | 11 |
| Classe 2 | 30 | 7 | 11 |
| Total | 63 (60%) | 20 (19%) | 22 (21%) |

Conclusions :
aucun lien significatif (Chi²) n'a été trouvé entre les 2 classes de préférence et les variables du questionnaire.

En particulier, aucun lien n'a été montré entre ces classes et les fréquences de consommation d'eau en bouteille et d'eau du robinet.

Exemple 4 : test de différents concepts

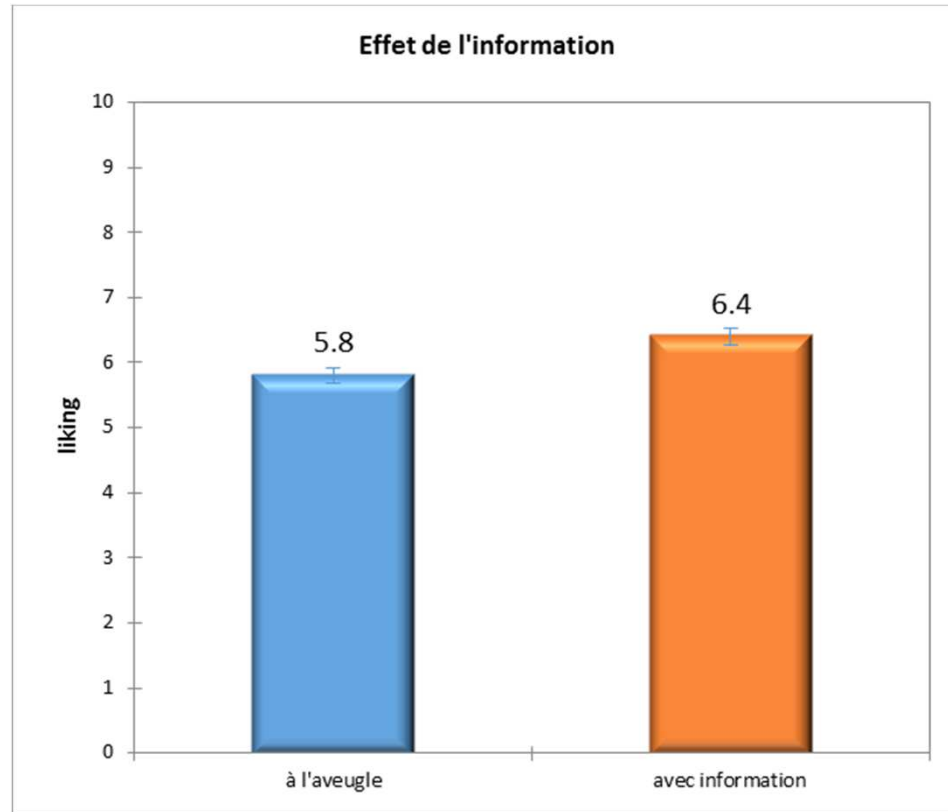
Classement de différents concepts, ici regroupés par thématique
La CAH a donné 2 classes principales.



Exemple 5 : effet d'une information



Un nouveau procédé de fabrication : faut-il communiquer ?



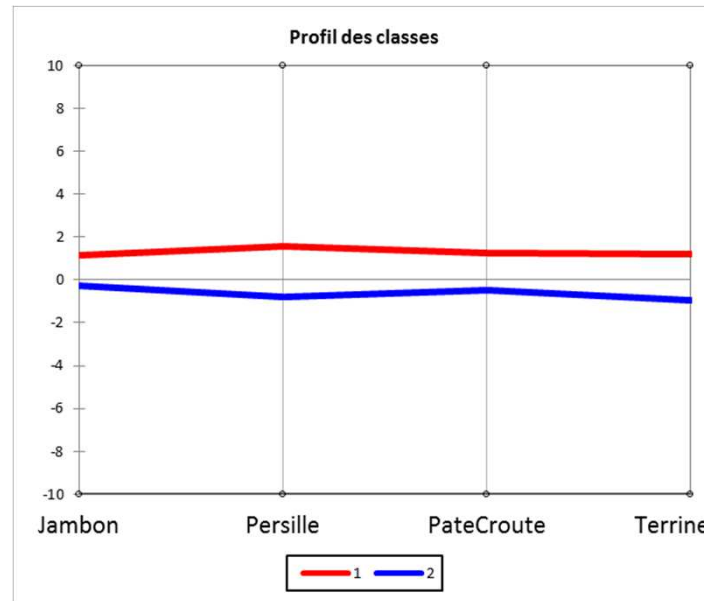
Tous produits confondus, l'information a une influence positive sur la moyenne globale des notes qu'elle a tendance à augmenter.

→ On aurait envie de conclure qu'on a intérêt à communiquer sur le nouveau procédé

Exemple 5 : segmentation des consommateurs



Une classification (CAH) a été effectuée sur les différences d'appréciation avec-sans information.



L'analyse a révélé deux classes de consommateurs.
Une première classe de consommateurs (45.05% des sujets) est influencée positivement par l'information (+1.30 en moyenne).
Une seconde classe (54.95%) est influencée plutôt négativement par l'information (-0.63 en moyenne).
Dans les deux cas, cet effet est tous produits confondus.

→ On conclut qu'il est préférable de ne pas communiquer sur le nouveau procédé

Exemple 6 : effet du type d'emballage

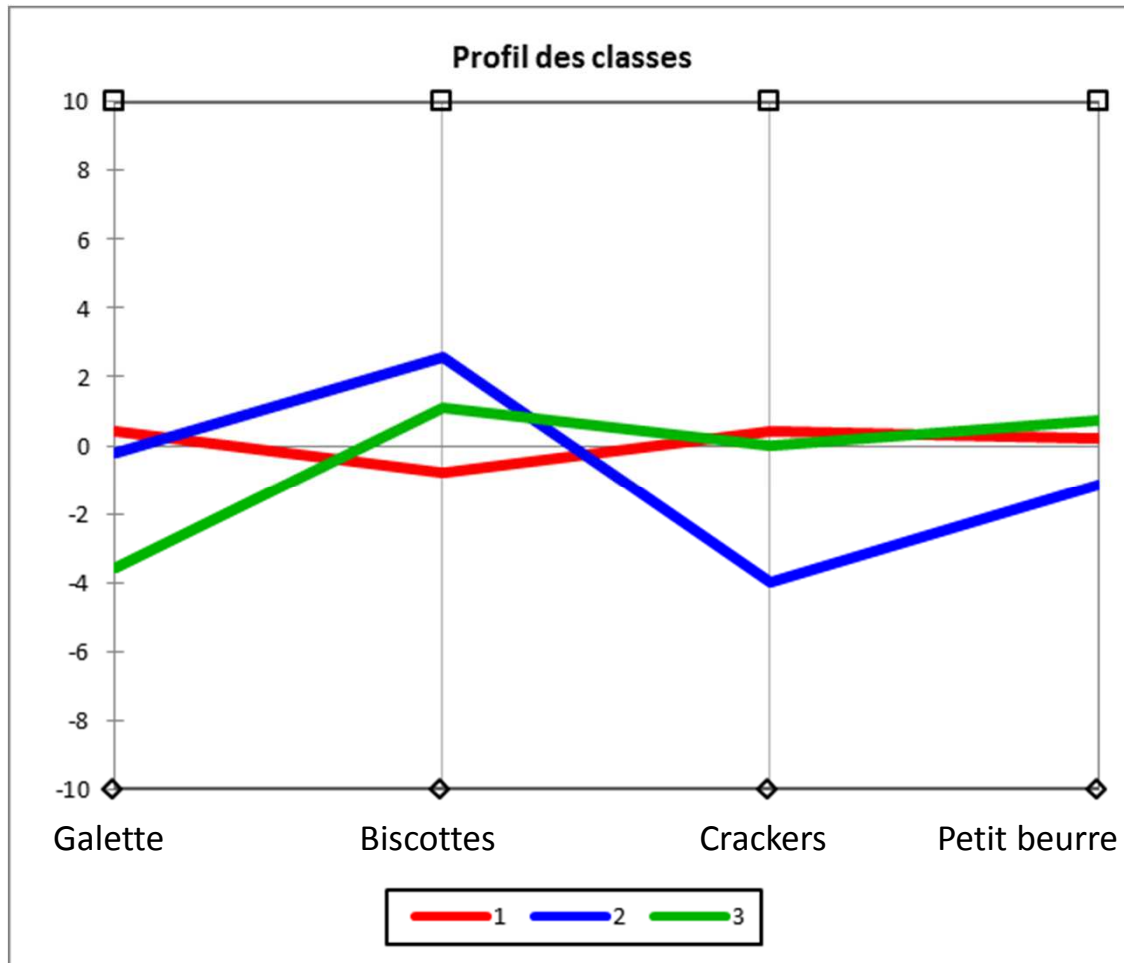


| | E. témoin | E. nouveau | P-value Produit |
|--------------|-----------|------------|-----------------|
| Biscotte | 6.15 | 6.35 | 0.4805 |
| Crackers | 6.28 | 5.83 | 0.0909 |
| Petit beurre | 5.81 | 5.88 | 0.7731 |
| Galette | 6.66 | 6.13 | 0.0210 * |

On observe une différence significative pour les galettes : le produit dans l'emballage témoin est légèrement plus apprécié (+0.5 points).

Exemple 6 : segmentation des consommateurs

Une classification (CAH) a été effectuée sur les différences d'appréciation Nouveau-Témoin



Il semble que l'analyse tous produits confondus révèle 3 classes de consommateurs.

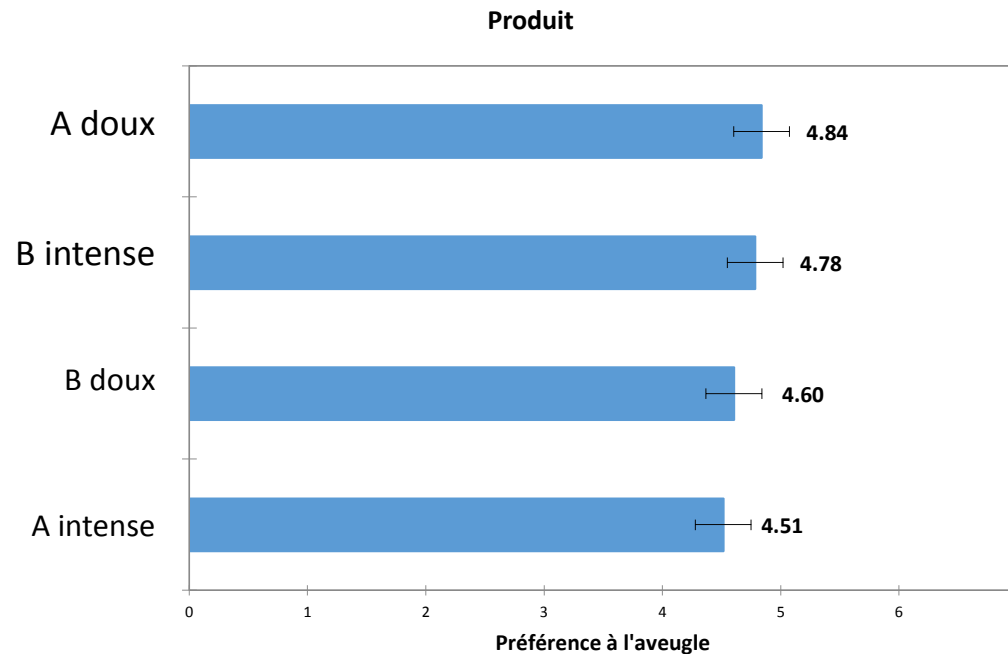
Une première classe de consommateurs, largement majoritaire (61.8% des sujets) ne fait globalement pas de différence d'appréciation entre les deux types d'emballage.

Il pourrait y avoir 2 autres classes de consommateurs qui apprécieraient moins un des types de produits emballés avec le nouvel emballage : 17.6% apprécieraient moins les crackers et 20.6% apprécieraient moins les galettes.

Exemple 7 : comparaison de cafés de différentes intensités



Une Anova « Sujet + Produit » a été effectuée sur les notes de préférence à l'aveugle.



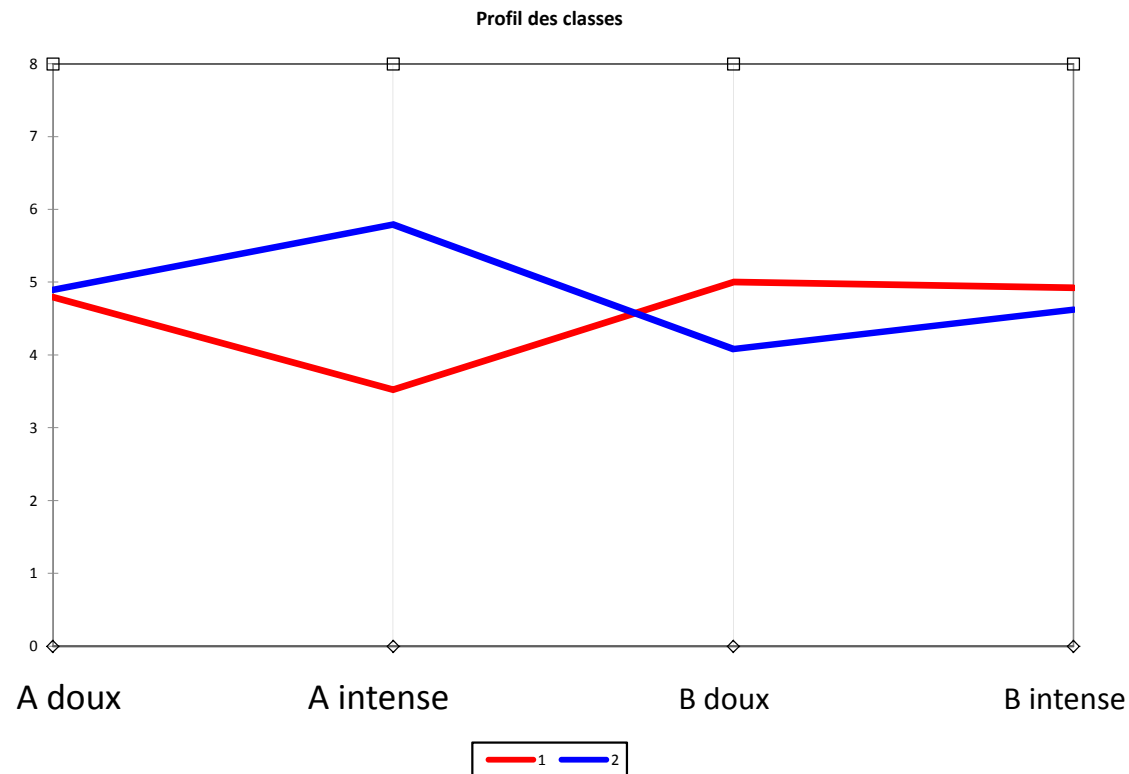
LSD = 0,33

Au niveau de l'ensemble du panel, il n'existe pas de préférence marquée envers l'un ou l'autre type de café.

Exemple 7 : segmentation des consommateurs



Une classification des consommateurs a été menée afin d'identifier de potentiels profils de préférences. Chaque consommateur a été centré mais non réduit.



L'analyse révèle 2 profils de consommateurs. Une première classe (en rouge, 43.2% des personnes interrogées) rejette particulièrement le produit A intense.

Une seconde classe (en bleu, 56.8% des personnes interrogées) apprécie particulièrement ce produit A intense.

La classe 1 apprécie par contre le produit B doux, contrairement à la classe 2 qui le déprécie.

Pour les 2 autres produits, les préférences moyennes sont homogènes entre les 2 classes.

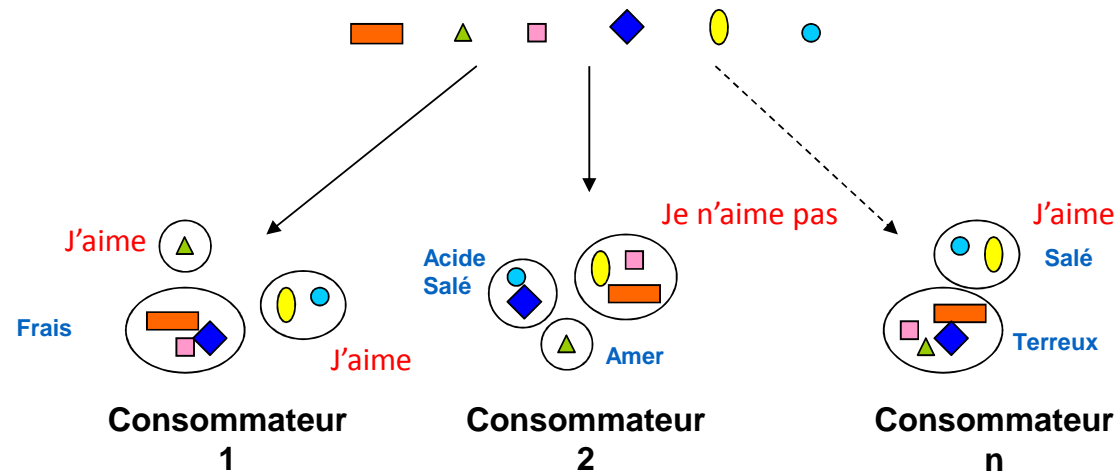
Ce que l'on fait moins souvent...

Classification des consommateurs sur la base de leurs perceptions

Tri libre



« Faites des groupes de produits qui se ressemblent »...



L'indice de Rand, un indice de proximité entre 2 tris :

$$R_a = (a+b)/(a+b+c+d)$$

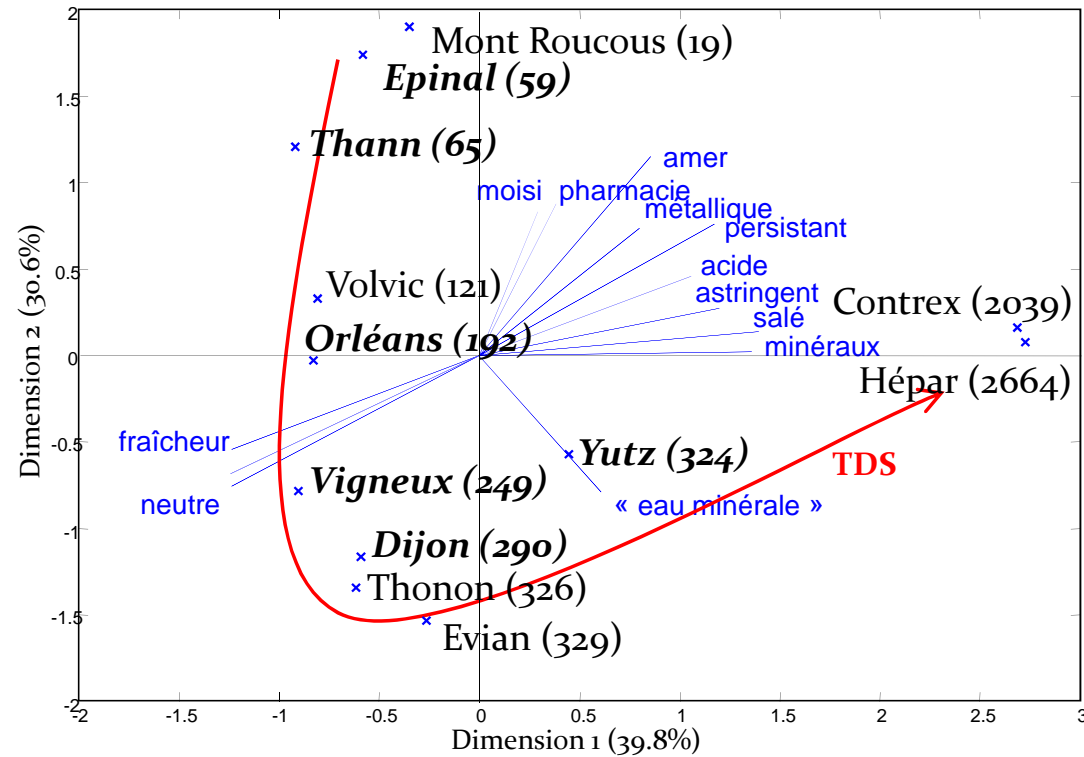
Nb : le Rand corrigé

Matrice 389*389 des proximités entre juges → Matrice 389*389 de dissimilarités entre juges → Classification

Exemple : tri libre auprès de 389 consommateurs



6 eaux du robinet, 6 eaux minérales, configuration générale

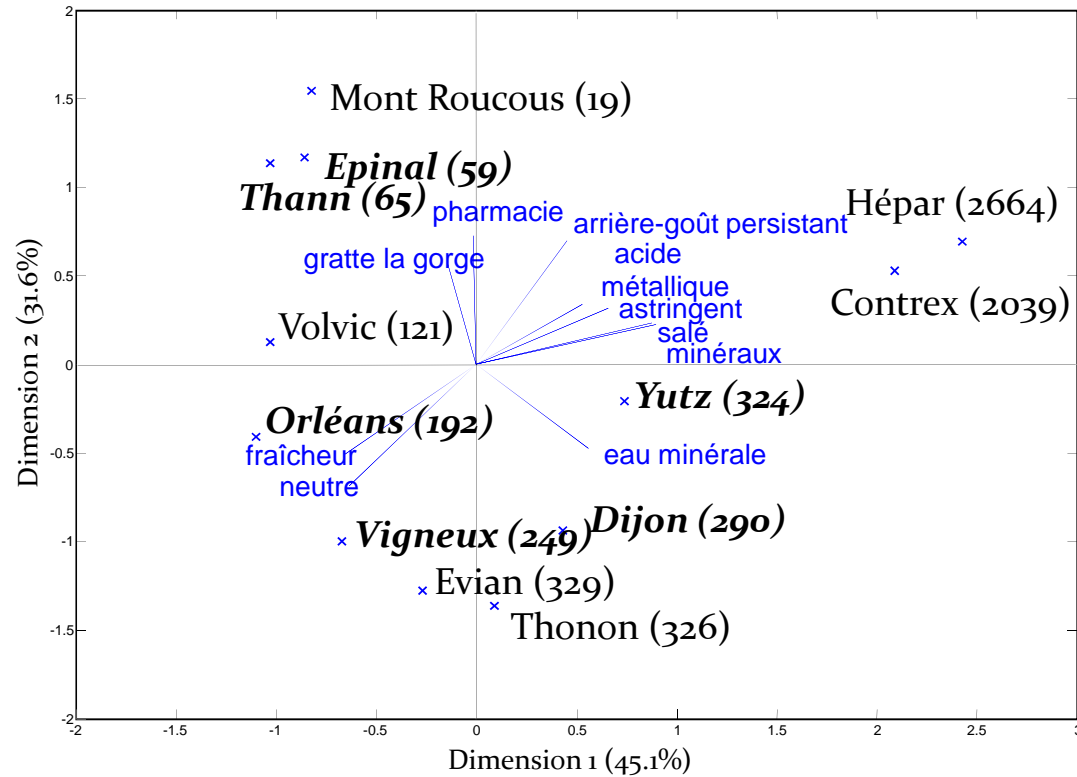


Espace sensoriel basé sur la minéralité, pas de discrimination eau du robinet / eau minérale

Exemple : tri libre auprès de 389 consommateurs



Après classification sur indice de Rand, 3 groupes. 1^{er} groupe (56%) :

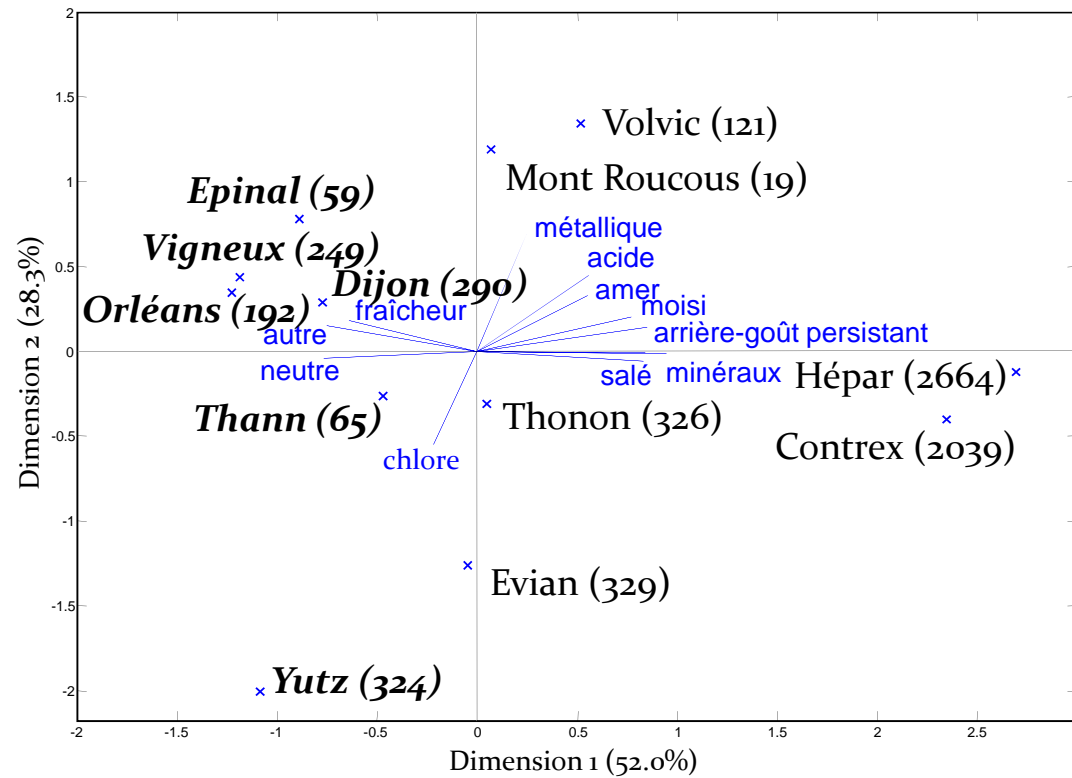


On retrouve bien la structure générale globale...

Exemple : tri libre auprès de 389 consommateurs



Après classification sur indice de Rand, 3 groupes. 2^{ème} groupe (36%) :



Eaux minérales et eaux du robinet discriminées !

En règle générale



- **Il ne faut pas se priver d'essayer de classer les consommateurs sur la base de leurs perceptions ! Si les perceptions ne sont pas homogènes, alors comment les préférences le seront-elles ?**
- **Ce genre de classification est faisable dès qu'un indice de proximité entre sujets peut-être calculé (Rand, RV, corrélation...) et peut donc être réalisée pour la plupart des tâches perceptives demandées aux consommateurs (tri, PSP, Napping... Mais aussi pour des méthodes « mixtes » de type échelles JAR...).**

Essayez ! On a parfois des surprises...!

Quelques considérations autour de la classification hédonique

Quel type de classification ?



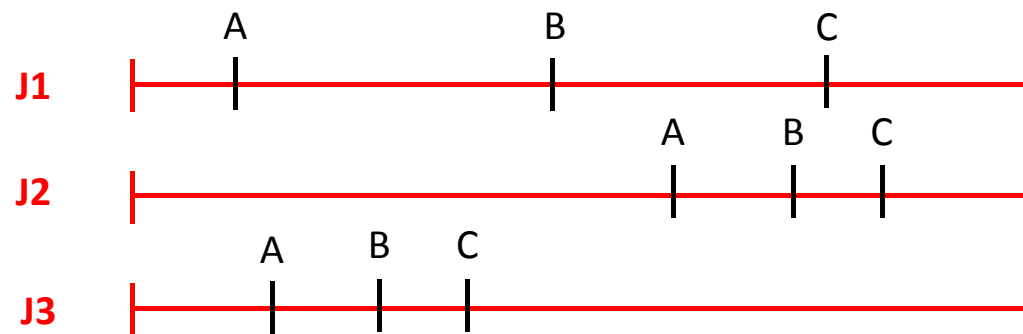
- **Classification ascendante hiérarchique ? K-means ? Varclus... ?**
 - **La CAH a le gros avantage de présenter un dendrogramme, on conserve une structure hiérarchique entre chaque niveau de coupure possible.
Quelles distances ? Quelles distances ultramétriques ?**
 - **Les K-means (et les nuées dynamiques) donnent parfois des résultats plus satisfaisants que la CAH à un nombre de classes donné. Ce sont des méthodes d'optimisation. Mais il faut choisir le nombre de classes « à priori » (pseudo F ?)**
 - **Les classifications de variables de type varclus donnent des résultats également intéressants. Elles peuvent être utilisées dans les approches de type cartographie interne ou externe.**

Pas de méthode « optimale » ou de règle absolue.

Centrage ? Réduction ?



- **Cela dépend des écoles.....**
- **Tout le monde s'accorde à autoriser voire imposer le centrage. Mais pour la réduction ?**



A vous de voir ! On notera tout de même que si l'on a une échelle verbale, la réduction perd de son sens...



Expert en analyse sensorielle

Expert en tests consommateurs

Expert en Statistiques Appliquées

Agréé Organisme de Formation

Agréé pour le Crédit Impôt Recherche

**Toute l'équipe SensoStat est à votre service !
Contactez-nous au 03 80 68 16 36**

Christine Urbano

Responsable Analyse Sensorielle et Tests
Consommateurs
christine.urbano@sensostat.fr
(+33) 6 87 74 76 67

Éric Teillet

Responsable Scientifique et Statistique
eric.teillet@sensostat.fr
(+33) 6 85 95 26 13