

ECFG932 : Techniques avancées en analyse sensorielle

Guillaume Franchi

Cursus Ingénieur 3^{ème} année

JAR : Just About Right

Pré-requis

- Test de Student.
- Analyse en Composantes Principales (ACP).
- Analyse Factorielle des Correspondances (AFC)

Enseignement

- Cours : \approx 45min.
- Travaux pratiques : \approx 3h00.

1. Contexte et Objectifs

Exemple

- On a fait goûter 8 fromages à pâte pressée à un panel de 72 consommateurs :
 - Beaufort (B) ;
 - Cantal 1 (C1) ;
 - Cantal 2 (C2) ;
 - Comté (CE) ;
 - Emmental (E) ;
 - Morbier (M) ;
 - Reblochon (R) ;
 - St-Nectaire (S).
- Les consommateurs ont noté ces fromages par :
 - une **note hédonique** d'appréciation globale (échelle de 1 à 9) ;
 - une évaluation de 9 attributs sensoriels, selon une **échelle JAR** en 5 points :
 - ▷ Couleur ;
 - ▷ Consistance toucher ;
 - ▷ Intensité odeur ;
 - ▷ Intensité goût ;
 - ▷ Goût fruité ;
 - ▷ Goût salé ;
 - ▷ Fermeté texture ;
 - ▷ Crémeux texture ;
 - ▷ Intensité arrière-goût.

Exemple :

	Vraiment pas assez salé	Pas assez salé	Juste bien	Trop salé	Vraiment trop salé
Goût salé	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Exemple (suite)

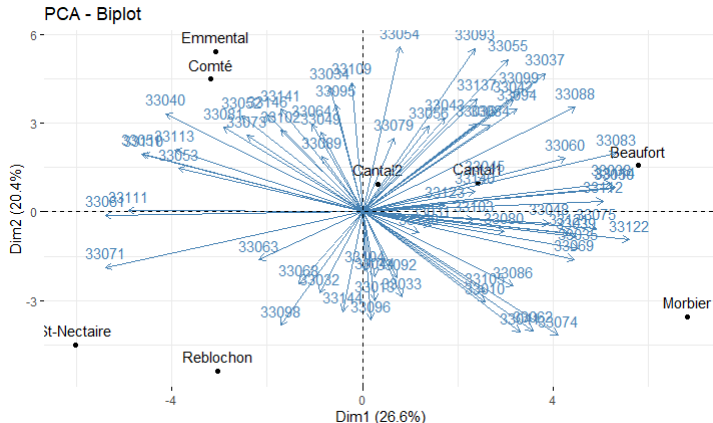
Conso	Produit	Liking	Couleur	Cons	Odeur	Int_g	Fr_g	Sel_g	Ferm	Crem	Int_ag
33001	Comté	3	2	4	2	2	2	3	2	2	2
33001	Morbier	4	2	2	2	4	2	4	2	4	4
33001	Beaufort	2	4	5	1	1	1	3	5	2	1
33001	Reblochon	3	3	1	3	2	2	3	2	3	2
33001	Cantal1	4	2	4	2	2	2	5	4	2	2
33001	Emmental	1	3	4	1	2	2	3	4	2	2
33001	Cantal2	7	2	4	1	3	3	4	3	3	3
33001	St-Nectaire	5	3	2	1	2	2	3	3	3	2
33010	Morbier	5	3	2	3	4	2	3	3	3	4

Source : Thèse CIFRE Oniris d'Alexiane Luc, 2022, *Evaluation de la perception des consommateurs à l'aide du protocole sensoriel Free JAR : une nouvelle méthodologie de collecte d'analyse statistique des données.*

🎯 Objectifs :

- ⚙️ Déterminer les liens entre les notes d'appréciation globale et les évaluations sensorielles JAR.
- ⚙️ Identifier les défauts d'un produit.
- ⚙️ Déterminer les relations entre les attributs JAR.
- ⚙️ Situer un produit par rapport aux autres.

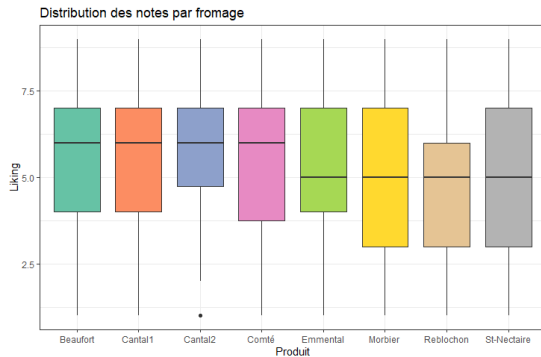
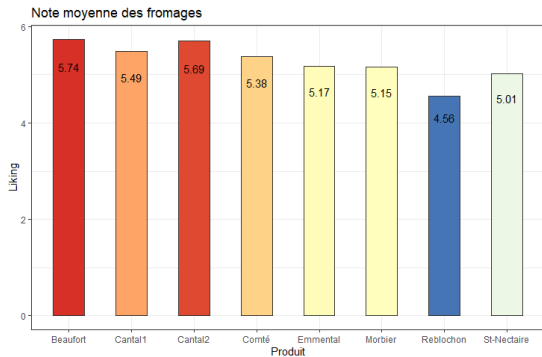
Cartographie interne des préférences



💡 Les goûts des consommateurs sont très hétérogènes.

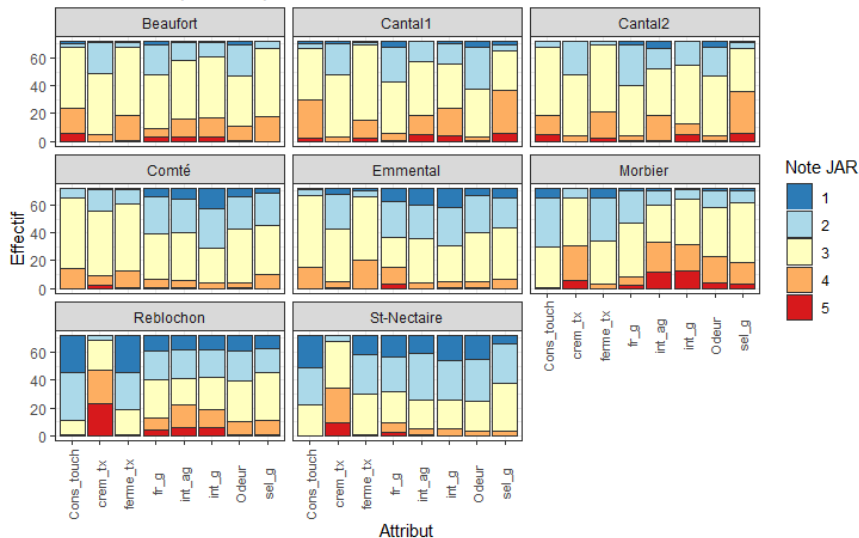
💡 Les produit « Comté » et « Emmental » d'une part, et « Morbier » d'autre part, sont des produits segmentants.

Représentations graphiques (1/3)



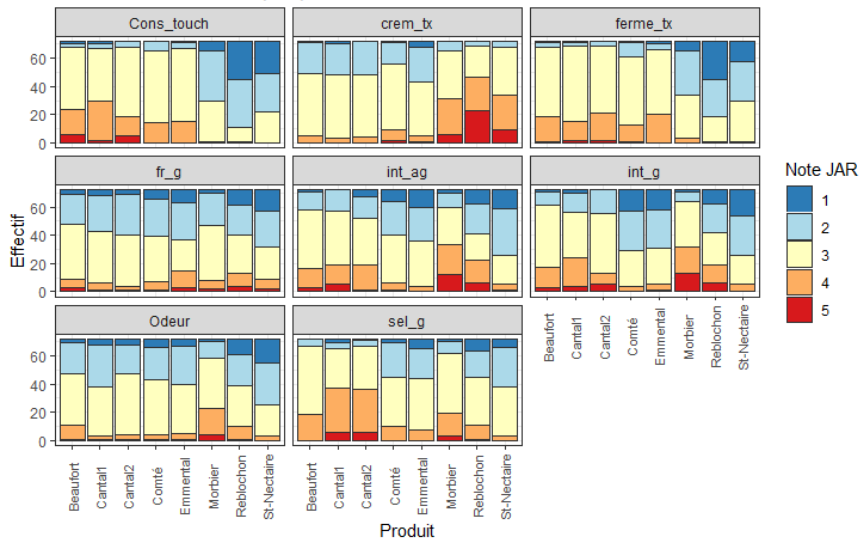
Représentations graphiques (2/3)

Résumé des produits par attribut



Représentations graphiques (3/3)

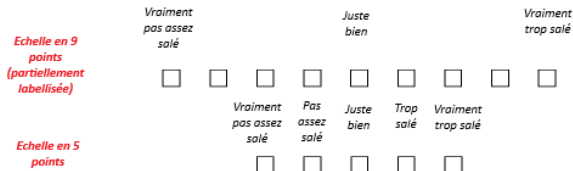
Résumé des attributs par produit



2. Analyses sur échelles JAR discontinues

Echelles JAR

- L'échelle JAR la plus utilisée pour noter un attribut du produit est une échelle de réponse discontinue structurée.
- Il est demandé à chaque individu de cocher la case correspondant le mieux à son avis sur l'attribut considéré.
- La valeur idéale est placée au centre de l'échelle, qui comporte en général 3,5,7 ou 9 points.



Remarques

- L'évaluation sensorielle des attributs du produit s'accompagne d'une note hédonique d'appréciation globale du produit. (*liking*).
- En général, chaque individu évalue plusieurs produits.
- La taille d'échantillon préconisée est d'au moins 75 consommateurs (*Source : SFAS, 2022*).

Périmètre d'analyse

- L'objectif principal est de mesurer l'impact d'un attribut sur l'appréciation générale.
- Deux solutions :
 - Conduire les analyses par produit.
 - Conduire les analyses tous produits confondus.

Remarque

- Les méthodes d'analyse sont identiques dans les deux cas, seules les interprétations diffèrent.
- Dans l'exemple qui suit, on ne fait que l'analyse du produit « Beaufort ».

⚙️ **Pré-traitement** : Les différents items de l'échelle sont regroupés en 3 catégories : « Pas assez », « JAR » ou « Trop ».

Vraiment pas assez salé <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	Juste bien <input type="checkbox"/>	Vraiment trop salé <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>		
Vraiment pas assez salé <input type="checkbox"/>	Pas assez salé <input type="checkbox"/>	Juste bien <input type="checkbox"/>	Trop salé <input type="checkbox"/>	Vraiment trop salé <input type="checkbox"/>

On calcule alors la fréquence de chacun de ces groupes :

- « Pas assez » salé : 0.0694
- « JAR » salé : 0.681
- « Trop » salé : 0.25

⚙️ Pénalité (*Mean drop*) :

- On calcule ensuite les moyennes des appréciations globales pour chaque groupe ainsi défini.
 - « Pas assez » salé : 3.20
 - « JAR » salé : 6.37
 - « Trop » salé : 4.72
- Les deux effets sur la moyenne d'un attribut sont alors
 - « Pas assez » salé : Moyenne « JAR » salé – Moyenne « Pas assez » salé

$$6.37 - 3.20 = 3.17.$$

- « Trop » salé : Moyenne « JAR » salé – Moyenne « Trop » salé

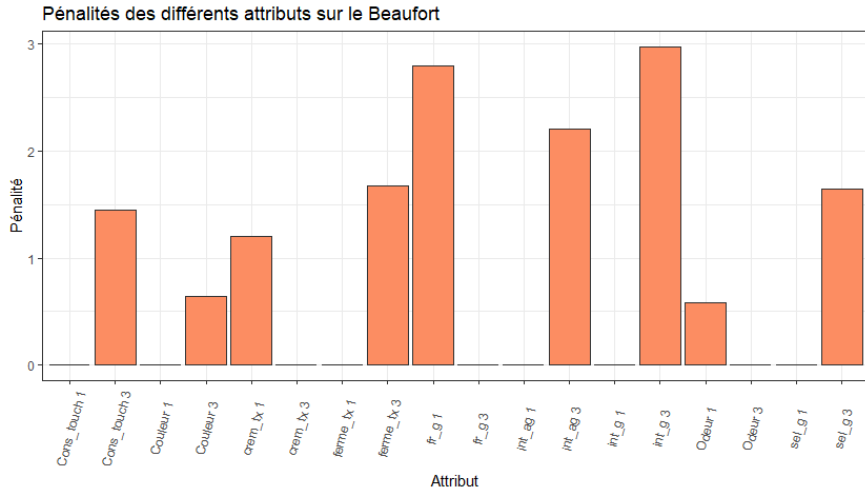
$$6.37 - 4.72 = 1.65.$$

Remarques

- Les catégories d'attribut JAR avec un pourcentage de consommateurs dans les catégories « Pas assez » ou « Trop » inférieur à 20% ne doivent pas être considérées (*même si les logiciels les présentent*).
- Il n'est pas souhaitable de conserver une catégorie avec moins de 15 sujets en absolu, même si le critère de 20% est respecté.

Source : SFAS, 2022

Représentation graphique :



⚙️ Significativité statistique :

Les effets sur la moyenne peuvent être statistiquement testés en effectuant un **test de Student**.

Exemple

Comparons la moyenne d'appréciation globale des consommateurs ayant trouvé le Beaufort « Trop salé » par rapport à la moyenne donnée par ceux l'ayant trouvé « JAR » salé.

Exemple (suite)

On note :

- m_3 et m_2 les moyennes théoriques des deux groupes ;
- $n_3 = 18$ et $n_2 = 49$ les effectifs des deux groupes ;
- $\bar{x}_3 = 4.72$ et $\bar{x}_2 = 6.37$ les moyennes des deux groupes ;
- $S_3^2 = \frac{1}{n_3} \sum_{i=1}^{n_3} (x_i^{(3)} - \bar{x}_3)^2$ et $S_2^2 = \frac{1}{n_2} \sum_{i=1}^{n_2} (x_i^{(2)} - \bar{x}_2)^2$ les variances empiriques des deux groupes.

On teste $H_0 : m_3 = m_2$ contre $H_1 : m_3 < m_2$. Sous les hypothèses adéquates, on a

$$T = \frac{\sqrt{n_2 + n_3 - 2} (\bar{x}_3 - \bar{x}_2)}{(n_3 S_3^2 + n_2 S_2^2) (1/n_3 + 1/n_2)} \sim \mathcal{T}_{n_3+n_2-2}.$$

On rejette H_0 au seuil de 5% si $T < q_{0.05}$, le quantile d'ordre 0.05 de la loi $\mathcal{T}_{n_3+n_2-2}$.

Exemple (suite)

Ici, on obtient $T \approx -2.87$ et $q_{0.05} \approx -1.67$ le quantile d'ordre 0.05 de la loi \mathcal{T}_{65} .

On rejette donc H_0 .

Par ailleurs, la p -value calculée ici vaut $p \approx 2.75 \times 10^{-3}$.

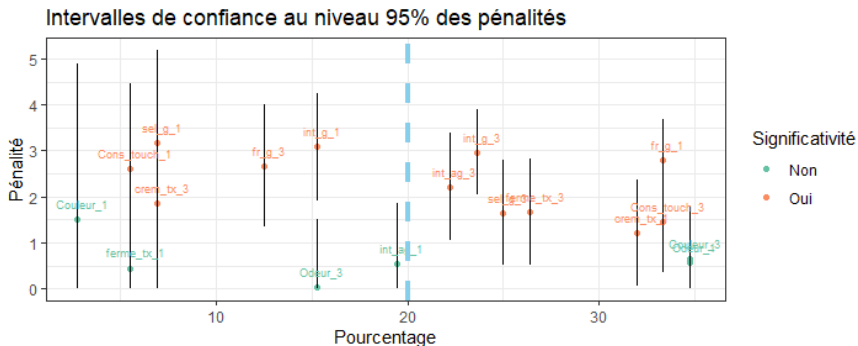
Remarque

La loi de Student permet également de construire des intervalles de confiance pour la pénalité théorique $m_2 - m_3$.

- Sur l'exemple du beaufort, et sur les catégories pertinentes, on obtient les résultats suivants.

Attribut	Effectif	Fréquence	Moyenne	Pénalité	p -value	Significativité à 5%
Cons_touch_3	24	33.33	4.92	1.45	5.06×10^{-3}	Oui
Couleur_3	25	34.72	5.36	0.64	0.13	Non
Odeur_1	25	34.72	5.36	0.58	0.17	Non
crem_tx_1	23	31.94	5.04	1.21	2.00×10^{-2}	Oui
ferme_tx_3	19	26.39	4.53	1.68	2.60×10^{-3}	Oui
fr_g_1	24	33.33	4.21	2.79	2.31×10^{-8}	Oui
int_ag_3	16	22.22	4.12	2.21	1.83×10^{-4}	Oui
int_g_3	17	23.61	3.94	2.97	1.2×10^{-8}	Oui
sel_g_3	18	25.00	4.72	1.65	2.75×10^{-3}	Oui

- On peut représenter l'ensemble des pénalités (*significatives ou non*) sur un graphique croisé ayant :
 - en abscisse le pourcentage de consommateurs de chaque catégorie ;
 - en ordonnée la pénalité de la catégorie.



- On peut également ajouter les intervalles de confiance pour les pénalités.

⚙️ **Interprétation :**

En règle générale :

- Une catégorie « Pas assez » ou « Trop » est un point d'amélioration si le pourcentage est supérieur à 30% ;
- Une catégorie « JAR » est un point fort si le pourcentage est supérieur à 60%, avec une répartition équilibrée des deux autres catégories.

Source : SFAS, 2022.

Exemple

La salinité du beaufort ne peut pas être considérée comme un point fort :

- « Pas assez » salé : 6.9% ;
- « JAR » salé : 68.1%
- « Trop » salé : 25%.

⚙️ La pénalité pondérée :

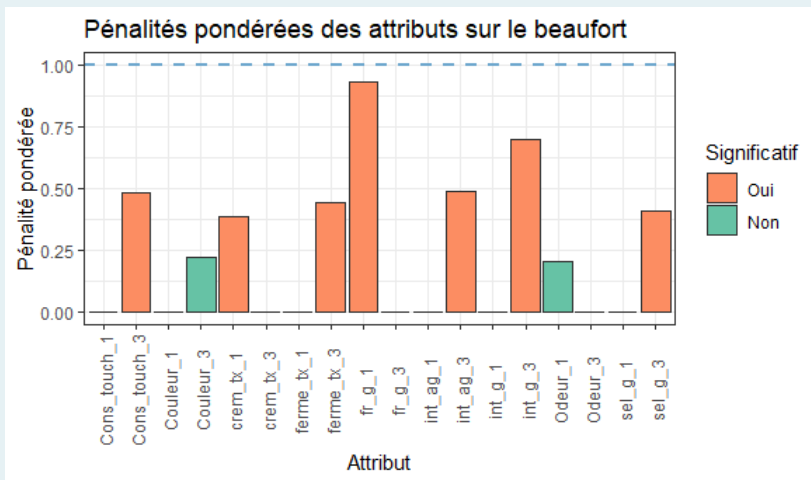
- Elle est obtenue en multipliant la pénalité d'une appréciation JAR « Trop » ou « Pas assez » par la fréquence des consommateurs ayant choisi cette appréciation.
- Dans notre exemple, on a comme pénalités pondérées
 - « Pas assez » salé : $3.17 \times 0.0694 \approx 0.22$.
 - « Trop » salé : $1.65 \times 0.25 \approx 0.41$.

A titre indicatif

- Sur une échelle d'appréciation globale graduée de 1 à 9 :
 - Si la pénalité pondérée est inférieure à 1 alors le défaut est considéré comme mineur.
 - Si la pénalité pondérée est supérieure à 1 alors le défaut est considéré comme majeur.
- Sur une échelle d'appréciation globale graduée de 1 à 6 :
 - Si la pénalité pondérée est inférieure à 0.5 alors le défaut est considéré comme mineur.
 - Si la pénalité pondérée est supérieure à 0.5 alors le défaut est considéré comme majeur.

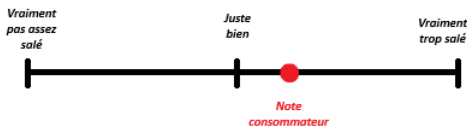
Exemple

Dans notre exemple sur le beaufort, il n'y a aucun défaut majeur.



3. Analyses sur échelles JAR continues

⚙️ Selon les études, il es parfois demandé au consommateur de noter chaque attribut sur une échelle **continue** (*mesurée, par exemple, de 0 à 10cm*).



⚙️ Chaque variable considérée est alors **quantitative** (*ou numérique*).

Exemple

Dans notre exemple des fromages, on pourrait avoir

Produit	Liking	Couleur	Cons_touch	Odeur	int_g	fr_g	sel_g	ferme_tx	crem_tx	int_ag
CE	2.90	4.20	6.20	3.90	3.90	3.90	5.30	3.90	4.20	4.10
M	4.00	4.10	4.10	4.10	6.30	4.10	6.40	3.80	6.30	6.00
B	2.30	5.80	7.40	3.40	3.50	3.20	5.10	7.00	3.90	3.30
R	3.00	5.10	3.10	4.80	4.40	4.30	5.20	3.90	4.80	3.90
C1	4.00	4.00	6.00	4.20	4.00	4.00	7.20	6.30	3.90	3.70
E	1.30	5.10	6.20	3.30	3.70	4.10	5.30	6.00	4.20	4.20

Modèle ANOVA

- On effectue une **analyse de variance** pour chaque attribut, dans le but de les trier du moins au plus discriminant, **tous produits confondus**.
- Le modèle associé est :

$$\mathbf{x}_{i,j}^{(k)} = \mu + \alpha_i^{(k)} + \beta_j^{(k)} + \varepsilon_{i,j}^{(k)}$$

où

- μ est une constante (*appréciation d'un produit de référence par un consommateur de référence*);
- $\mathbf{x}_{i,j}^{(k)}$ est la note du consommateur i donnée au produit j pour l'attribut k ;
- $\alpha_i^{(k)}$ est un effet aléatoire du niveau i du facteur *consommateur* pour l'attribut k ;
- $\beta_j^{(k)}$ est un effet fixe du niveau j du facteur produit pour l'attribut k ;
- $\varepsilon_{i,j}^{(k)}$ est une variable aléatoire résiduelle.

Remarques

Dans le modèle

$$\mathbf{x}_{i,j}^{(k)} = \mu + \alpha_i^{(k)} + \beta_j^{(k)} + \varepsilon_{i,j}^{(k)}$$

- ⚙️ Seuls l'effet fixe du produit et l'effet aléatoire du consommateur sont testés.
 - 💡 L'effet de l'interaction n'est pas testable, chaque consommateur n'ayant évalué un produit d'une seule fois.
- ⚙️ Cette analyse globale ne permet pas de discerner les différences individuelles entre produits... Il se peut qu'un seul produit soit responsable d'un effet produit.
 - 💡 Dans ce cas, on peut effectuer un test de comparaison multiple des moyennes (*comme le test de Duncan*).

Exemple

On a testé ci-dessous les effets produits des attributs « Goût salé », « Goût fruité » et « Crémeux ».

Produit	sel_g	fr_g	crem_tx
Beaufort (<i>référence</i>)	0.00	0.00	0.00
Cantal1	0.26	-0.19	-0.04
Cantal2	0.27	-0.21	-0.02
Comté	-0.50	-0.26	0.18
Emmental	-0.58	-0.13	-0.09
Morbier	-0.09	-0.03	0.70
Reblochon	-0.60	-0.13	1.22
St-Nectaire	-0.78	-0.43	0.79
Pr > F (Produit)	$< 2 \times 10^{-16}$	0.0541	$< 2 \times 10^{-16}$
F (Produit)	21.836	1.995	36.218

L'attribut « Crémeux » est ici le plus discriminant, avec une statistique de Fisher plus élevée.

L'effet produit n'est même pas significatif pour l'attribut « Goût fruité ».

Analyse par produit

Deux étapes sont nécessaires pour l'étude détaillée d'un attribut k d'un seul produit à la fois.

- Calcul de la moyenne sur l'échelle JAR de l'attribut k du produit.
- Comparaison de la moyenne au centre de l'échelle JAR avec un **test de Student** :

$$T_k = \frac{\sqrt{n-1} \times (\bar{x}_k - m)}{s_k}.$$

où

- \bar{x}_k est la moyenne du produit pour l'attribut k ;
- $s_k^2 = \frac{1}{n} \sum_{i=1}^n (x_i^{(k)} - \bar{x}_k)^2$ la variance empirique du produit pour l'attribut k ;
- n est le nombre de consommateur ;
- m est la moyenne à l'idéal (*centre de l'échelle*).

⚙ Si μ_k est la moyenne théorique du produit pour l'attribut k , on teste $H_0 : \mu_k = m$ contre $H_1 : \mu_k \neq m$.

⚙ Sous les hypothèses adéquates

$$T_k \sim \mathcal{T}_{n-1}.$$

⚙ On rejette H_0 au niveau 95% si $|T_k| > q_{0.975}$ le quantile d'ordre 0.975 de la loi \mathcal{T}_{n-1} .

Exemple

On a testé ci-dessous les différents attributs du beaufort.

Beaufort	Couleur	Cons_touch	Odeur	int_g	fr_g	sel_g	ferme_tx	crem_tx	int_ag
Moyenne	5.46	5.45	4.88	5.22	4.90	5.32	5.32	4.83	5.16
Ecart-type	0.65	0.85	0.79	0.72	0.84	0.56	0.60	0.65	0.80
T calculé	5.95	4.48	-1.25	2.54	-1.01	4.84	4.53	-2.28	1.73
$q_{0.975}$	1.99	1.99	1.99	1.99	1.99	1.99	1.99	1.99	1.99

Par exemple, le beaufort a été jugé « idéalement », fruité, mais trop salé et trop ferme.

4. Analyse exploratoire des données

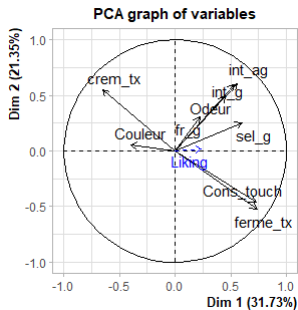
⚙️ On s'intéresse ici aux liens pouvant exister entre les différents attributs sensoriels, et comment un produit se positionne par rapport aux autres.

💡 En effet, certains attributs à améliorer peuvent être liées à d'autres...

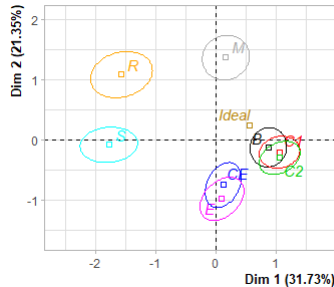
⚙️ Si on considère que les différentes notes sont de nature **numérique**, on peut effectuer une **Analyse en Composantes Principales** (ACP).

⚙️ On crée alors un produit fictif « Ideal », ainsi que des « produits moyens » correspondant aux moyennes obtenues par produit.

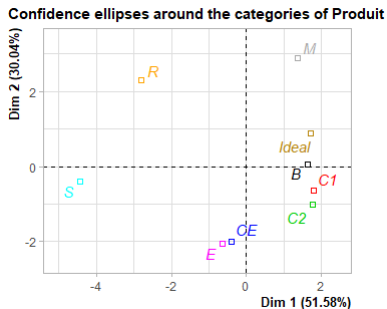
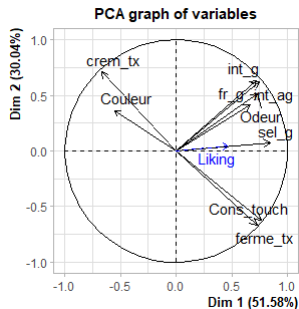
⚙️ Dans notre cas, l'ACP se fait sur les $8 \times 72 = 576$ données individuelles + le produit « Ideal ».



Confidence ellipses around the categories of Produit



- ⚙️ On peut également **agréger** les données individuelles.
- ⚙️ L'ACP est alors réalisée sur les « produits moyens ».
- ⚙️ (*Dans notre cas, sur les 8 « produits moyens » + le produit fictif « Ideal »*)



⚙️ Une telle ACP ne tient pas compte du caractère bipolaire des échelles JAR.

⚙️ Par exemple, si le comté semble avoir une intensité d'arrière-goût plus faible que la moyenne des tests effectués :

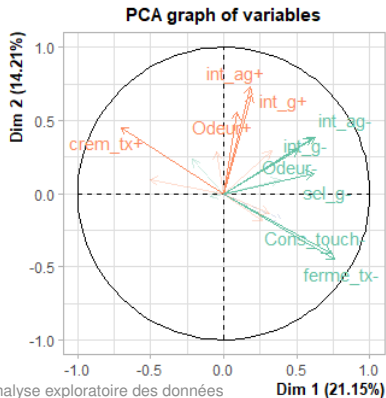
- Cette intensité est-elle « Trop faible » ?
- Est-elle « Juste bien », mais elle était trop forte pour l'ensemble des autres fromages ?

💡 On peut re-coder les données en séparant chaque attribut en deux (*dummy variables*) :

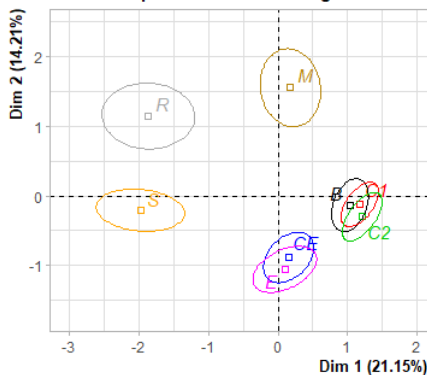
- le niveau JAR est codé 0 ;
- les valeurs 1 et 2 sont codées -1 et -2 ;
- les valeurs 4 et 5 sont codées 1 et 2.

⚙️ On obtient alors un tableau comme ci-dessous.

Produit	Liking	Couleur-	Couleur+	Cons_touch-	Cons_touch+
CE	3.00	-1.00	-0.00	0.00	1.00
M	4.00	-1.00	-0.00	-1.00	-0.00
B	2.00	0.00	1.00	0.00	2.00
R	3.00	0.00	0.00	-2.00	-0.00
C1	4.00	-1.00	-0.00	0.00	1.00
E	1.00	0.00	0.00	0.00	1.00



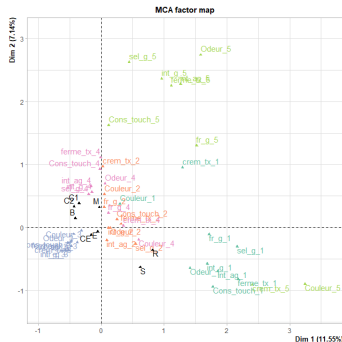
Confidence ellipses around the categories of Produit



⚙️ En considérant que les différentes notes sensorielles sont de nature purement **qualitative**, on peut effectuer une **Analyse des Correspondances Multiples** (ACM)

⚙️ On retire ici la note d'appréciation globale, et on réalise l'ACM sur les $8 \times 72 = 576$ données individuelles.

⚙️ La variable catégorielle « Produit » est comptée ici comme une variable qualitative supplémentaire.

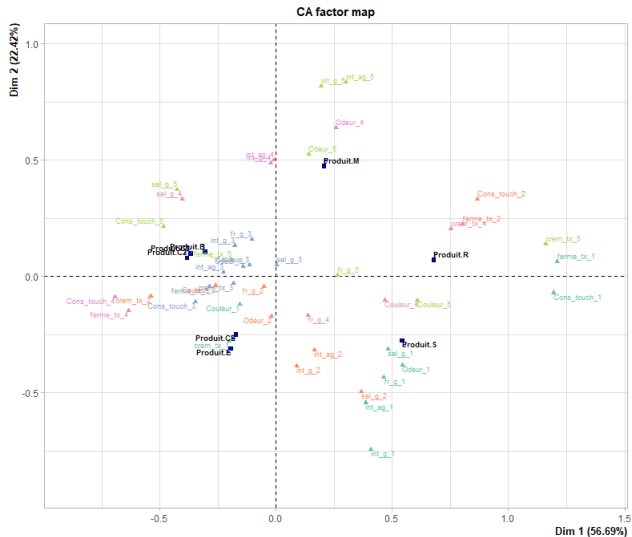


⚙ On peut également **agréger** les données individuelles.

⚙ On se retrouve alors avec une table de contingence de deux variables qualitatives : « Produit » et « Attribut_Note ».

Produit	Cons_touch_1	Cons_touch_2	Cons_touch_3	Cons_touch_4	Cons_touch_5	Couleur_1
B	2	2	44	18	6	1
C1	2	3	37	28	2	2
C2	0	4	49	14	5	1
CE	0	7	51	14	0	5
E	1	4	52	15	0	1
M	7	35	29	1	0	2
R	27	34	10	1	0	1
S	23	27	22	0	0	0

⚙️ On effectue ensuite une simple analyse des correspondances.



Résumé

- Echelle Jar = Echelle d'intensité bipolaire.
- Objectifs : identifier les défauts / points forts d'un produit, situer un produit par rapport aux autres.
- Cartographie interne.
- Analyse des pénalités (simple, pondérée).
- Analyse exploratoire : ACP, ACM, agrégation des données.
- Re-codage des attributs JAR en « dummy variables »