Analyse Factorielle Multiple

AFM

AUTEUR·RICE

Benjamin Charlier, Adrien Mazoyer

1 Objectifs

1.1 Description des données

On travaille sur des tableaux de données où

- les **lignes** correspondent aux **individus**
- les **colonnes** correspondent à des **variables quantitatives** regroupées

	Groupe 1			$\cdots$	**Groupe K**		
	var. 1	$\cdots$	**var. J_1**	$\cdots$	**var. 1**	$\cdots$	**var. J_K**
Individu 1	$t_{1,(1,1)}$	$\cdots$	$t_{1,(1,J_1)}$	$\cdots$	$t_{1,(K,1)}$	$\cdots$	$t_{1,(K,J_K)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Individu n	$t_{n,(1,1)}$	$\cdots$	$t_{n,(1,J_1)}$	$\cdots$	$t_{n,(K,1)}$	$\cdots$	$t_{n,(K,J_K)}$

- $x_{i,(k,j)}$: mesure de la variable j du groupe k pour l'individu i.
- les individus $i = 1, \ldots, n$
- les groupes $k = 1, \ldots, K$
- les variables $j = 1, \ldots, J_K$

1.2 Objectifs généraux

Étude sur les individus

- Identifier des profils d'individus
- Étudier la variabilité des profils
- Visualiser les individus

Étude sur les groupes

- Étudier les ressemblances entre les groupes
- Comparer avec les analyses des groupes (type ACP)
- Construction de variables synthétiques

2 Une ACP globale pondérée

2.1 Approche naïve

▸ **Les données**

Effectuer directement une ACP classique sur le tableau, en ignorant la structure en groupe des variables ?

- **Exemple jouet:** On dispose de 11 variables réparties dans 3 groupes distincts :
  - **Groupe A :** $\{A_1, A_2, A_3, A_4, A_5\}$: variables très fortement corrélées entre elles.
  - **Groupe B :** $\{B_1, B_2, B_3\}$: variables très faiblement corrélées entre elles.
  - **Groupe C :** $\{C_1, C_2, C_3\}$: variables très faiblement corrélées entre elles.
- **Conséquence:** Les groupes sont déséquilibrés du point de vue des liaisons. Le groupe A va très fortememt contribuer au premier Axe factoriel. La structur de dépendance des autres groupes ne sera pas visible sur le premier plan factoriel.

2.2 Matrice d'intérêt

Comme pour l'ACP/ACM, il faut identifier la "bonne matrice" à décomposer.

ⓘ Rappels

- **ACP :** Analyse de la matrice de corrélation (variables centrées-réduites) afin d'égaliser la variance de chaque colonne.
- **AFC :** Analyse du tableau de contingence transformé en profils-lignes / profils-colonnes avec la métrique du χ^2 (pondération par les marges).
- **ACM :** Analyse du Tableau Disjonctif Complet (TDC) pondéré afin de donner un poids plus fort aux modalités rares.

ⓘ Objectifs pour l'AFM

- Équilibrer l'influence des groupes : Éviter que la première direction principale globale ne soit portée par un unique groupe dominant.
- Respecter la structure interne : Conserver sans altération les corrélations et poids relatifs des variables au sein de chaque groupe.
- Préserver la multidimensionnalité : Garantir qu'un groupe riche en dimensions le reste dans l'analyse globale.

2.3 Choix de la pondération

- **Distance entre individus:** i et i' intègre la métrique de chaque groupe. Chaque variable j du groupe k est pondérée par l'inverse de la première valeur propre $\lambda_1^{(k)}$ issue de l'ACP individuelle de ce groupe :

$$d^2(i,i') = \sum_{k=1}^{K} \frac{1}{\lambda_1^{(k)}} \sum_{j=1}^{J_k} \left(x_{i,(k,j)} - x_{i',(k,j)}\right)^2$$

Remarque : Les variables $x_{.,(k,j)}$ sont supposées centrées et réduites.

- **Inertie totale de l'AFM :** se décompose comme la somme pondérée des inerties de chaque groupe :

$$\text{Inertie}_{\text{AFM}} = \sum_{k=1}^{K} \frac{1}{\lambda_1^{(k)}} \text{Inertie}_k$$

Conséquence géométrique

- L'inertie du groupe k dans l'AFM devient $\frac{\text{Inertie}_k}{\lambda_1^{(k)}}$
- La **première composante principale** est à inertie **1** dans l'analyse globale

2.4 Effets de la pondération en AFM

Avant pondération ($\lambda_m^{(k)}$)

	Groupe_A	Groupe_B	Groupe_C
comp 1	4.699	1.167	1.169
comp 2	0.105	0.991	0.962
comp 3	0.075	0.842	0.869

Après pondération ($\lambda_m^{(k)}/\lambda_1^{(k)}$)

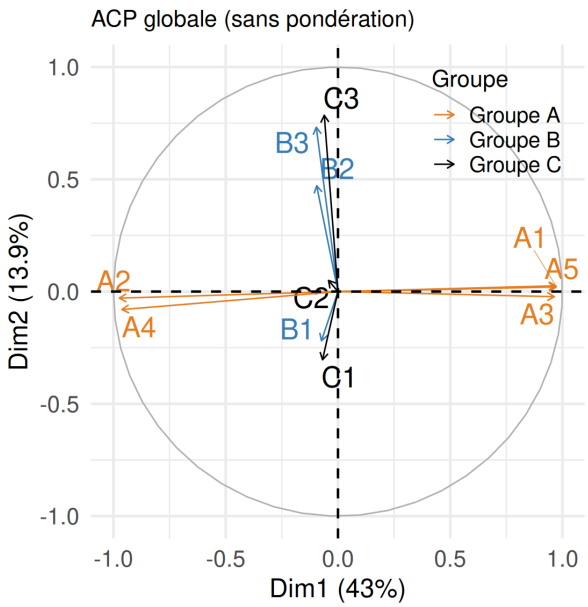
	Groupe_A	Groupe_B	Groupe_C
comp 1	1.000	1.000	1.000
comp 2	0.022	0.850	0.823
comp 3	0.016	0.721	0.744

Lecture directe depuis `FactoMineR`

- `$separate.analyses[[k]]$eig` stocke les valeurs propres brutes de l'ACP du groupe k.
- `$global.pca$eig` stocke les valeurs propres brutes de l'ACP pondérée par block.

2.5 Conséquence de la pondération

- **ACP globale non pondérée (domination du Groupe A)**

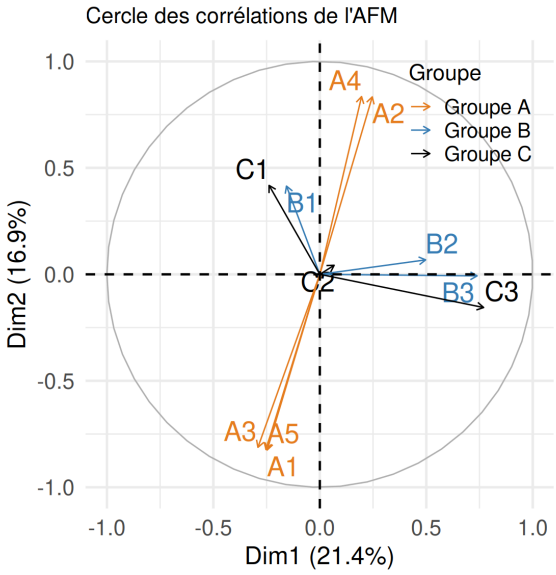


Valeurs propres et inertie expliquée (ACP non pondérée)

```
> print_table(acp_globale$eig[1:5, 1:2])
```

	eigenvalue	percentage of variance
comp 1	4.725	42.953
comp 2	1.526	13.870
comp 3	1.100	9.997
comp 4	0.996	9.055
comp 5	0.923	8.386

- **ACP globale pondérée de l'AFM (rééquilibrage des groupes)**



Valeurs propres et inertie expliquée (ACP globale AFM)

```
> print_table(toy.afm$global.pca$eig[1:5,
      1:2])
```

	eigenvalue	percentage of variance
comp 1	1.325	21.368
comp 2	1.048	16.906
comp 3	0.938	15.126
comp 4	0.818	13.197
comp 5	0.778	12.548

- Le premier plan capture un pourcentage d'inertie plus faible
- La structure de dépendance de chaque groupe est mieux préservée et représentée

2.6 Interprétation de base

- L'AFM est une ACP pondérée:
  - Représentations des individus et des variables
  - Variables synthétiques : cos2, contribution
  - Ajout post-ACP de variables ou d'individus supplémentaire pour aider à l'interprétation.
- Spécificités de l'AFM: Structure en groupe des variables
- On peut obtenir un résumé des principaux résultats en utilisant la fonction `summary`.

```
> summary(toy.afm)
```

```
Call:
MFA(base = df_toy, group = c(5, 3, 3), type = c("s", "s", "s"),
     name.group = c("Groupe A", "Groupe B", "Groupe C"), graph = FALSE)


Eigenvalues
                       Dim.1  Dim.2  Dim.3  Dim.4  Dim.5
Variance               1.325  1.048  0.938  0.818  0.778
% of var.             21.368 16.906 15.126 13.197 12.548
Cumulative % of var. 21.368 38.274 53.399 66.596 79.144

Groups
             Dim.1    ctr   cos2    Dim.2    ctr   cos2    Dim.3    ctr   cos2
Groupe A |  0.066  4.958  0.004 |  0.726 69.213  0.526 |  0.015  1.650  0.000 |
Groupe B |  0.700 52.851  0.219 |  0.151 14.422  0.010 |  0.189 20.108  0.016 |
Groupe C |  0.559 42.190  0.140 |  0.172 16.365  0.013 |  0.734 78.241  0.242 |

Individuals (the 10 first)
             Dim.1    ctr   cos2    Dim.2    ctr   cos2    Dim.3    ctr   cos2
1         |  1.094  0.903  0.117 |  2.113  4.258  0.437 | -1.689  3.040  0.279 |
2         |  0.755  0.430  0.170 | -0.073  0.005  0.002 | -1.193  1.516  0.424 |
3         | -0.724  0.396  0.184 |  0.271  0.070  0.026 | -0.532  0.301  0.099 |
4         |  1.581  1.887  0.555 |  0.218  0.045  0.011 | -0.767  0.627  0.130 |
5         |  0.440  0.146  0.045 | -0.346  0.114  0.028 |  0.171  0.031  0.007 |
6         | -0.549  0.228  0.073 | -0.551  0.290  0.074 |  0.963  0.988  0.226 |
7         | -1.106  0.924  0.161 |  1.174  1.315  0.182 |  0.031  0.001  0.000 |
8         |  1.414  1.508  0.201 | -1.635  2.550  0.269 |  0.498  0.264  0.025 |
9         |  1.348  1.372  0.270 |  1.532  2.239  0.348 |  1.450  2.241  0.312 |
10        | -0.574  0.248  0.135 |  0.525  0.263  0.113 |  0.730  0.568  0.218 |

Continuous variables (the 10 first)
             Dim.1    ctr   cos2    Dim.2    ctr   cos2    Dim.3    ctr   cos2
A1        | -0.247  0.984  0.061 | -0.825 13.823  0.681 | -0.136  0.422  0.019 |
A2        |  0.246  0.975  0.061 |  0.833 14.102  0.695 |  0.111  0.281  0.012 |
A3        | -0.291  1.356  0.084 | -0.812 13.397  0.660 | -0.119  0.319  0.014 |
A4        |  0.196  0.617  0.038 |  0.835 14.151  0.697 |  0.129  0.376  0.017 |
A5        | -0.253  1.027  0.064 | -0.823 13.741  0.677 | -0.106  0.253  0.011 |
B1        | -0.156  1.584  0.024 |  0.414 14.039  0.172 | -0.004  0.002  0.000 |
B2        |  0.497 15.983  0.247 |  0.068  0.378  0.005 |  0.258  6.066  0.066 |
B3        |  0.739 35.285  0.546 | -0.007  0.004  0.000 | -0.392 14.040  0.154 |
C1        | -0.238  3.668  0.057 |  0.418 14.236  0.174 | -0.440 17.653  0.194 |
C2        |  0.067  0.293  0.005 |  0.043  0.148  0.002 |  0.814 60.396  0.662 |
```

2.7 Exemple: le jeu de donnée des vins

▸ Descriptions

- $n = 10$ vins différents sont soumis à des analyses sensorielles
- $K = 3$ groupes : expert, étudiant, consommateur ;
- $J_1 = 27$ critères de description pour les experts, et $J_2 = J_3 = 15$ critères pour les étudiants et les consommateurs.

	Groupe Expert	**Groupe Étudiant**	**Groupe Consommateur**
	27 colonnes	**15 colonnes**	**15 colonnes**
Vin 1	⋯	⋯	⋯
⋮	⋮	⋮	⋮
Vin 10	⋯	⋯	⋯

L'AFM va nous permettre de :

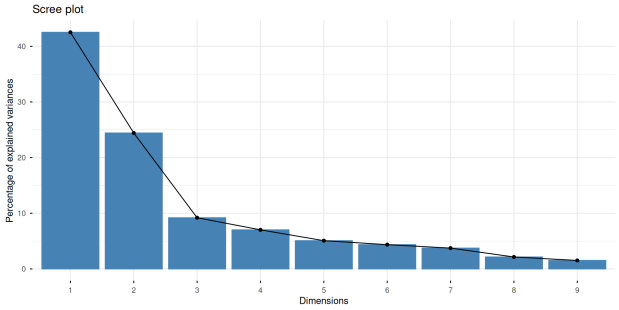
- dresser des profils de vins selon les critères d'analyse,
- identifier des différences de description entre les 3 groupes.

2.8 Exemple (jury de vins)

L'AFM avec groupe supplémentaire

```
> library(FactoMineR)
> vins.mfa <- MFA(vins, group=c(1,27,15,15,60), type=c("n",rep("s",4)),
       num.group.sup=c(1,5),
+            name.group=c("Cépage", "Expert", "Étudiant", "Consommateur", "Préférence"),
       ncp=9, graph=FALSE)
```

Valeurs propres et inertie expliquée (ACP globale AFM)



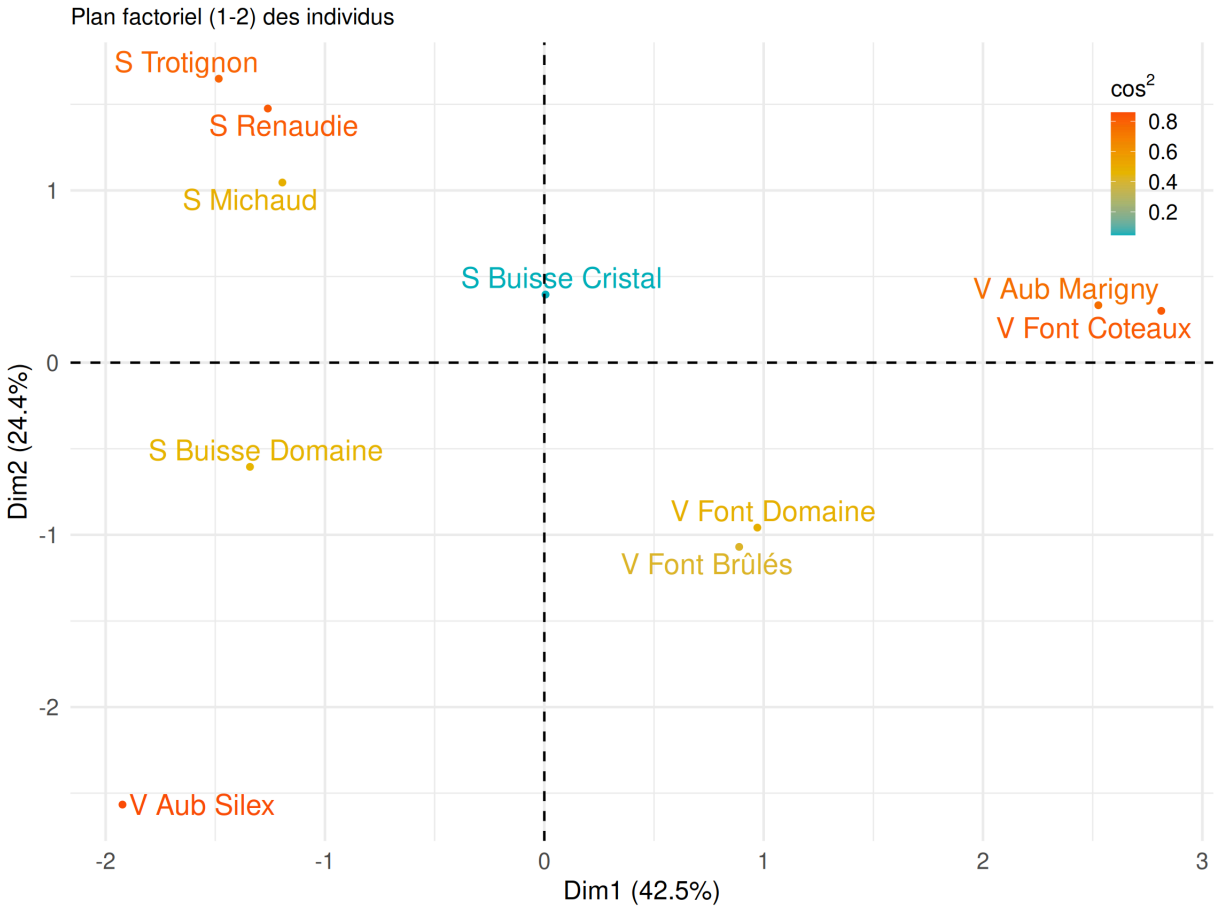

```
> print_table(vins.mfa $global.pca$eig[,
       c(1,3)])
```

	eigenvalue	cumulative percentage of variance
comp 1	2.674	42.515
comp 2	1.536	66.939
comp 3	0.578	76.135
comp 4	0.442	83.155
comp 5	0.319	88.230
comp 6	0.274	92.592
comp 7	0.235	96.331
comp 8	0.135	98.477
comp 9	0.096	100.000

- Règle du coude : on ne garderait que les 3 premières directions
- On pourrait rajouter la 3ème (et 4ème) pour arriver à une inertie projetée $> 70\%$ ($> 80\%$) de l'inertie totale.

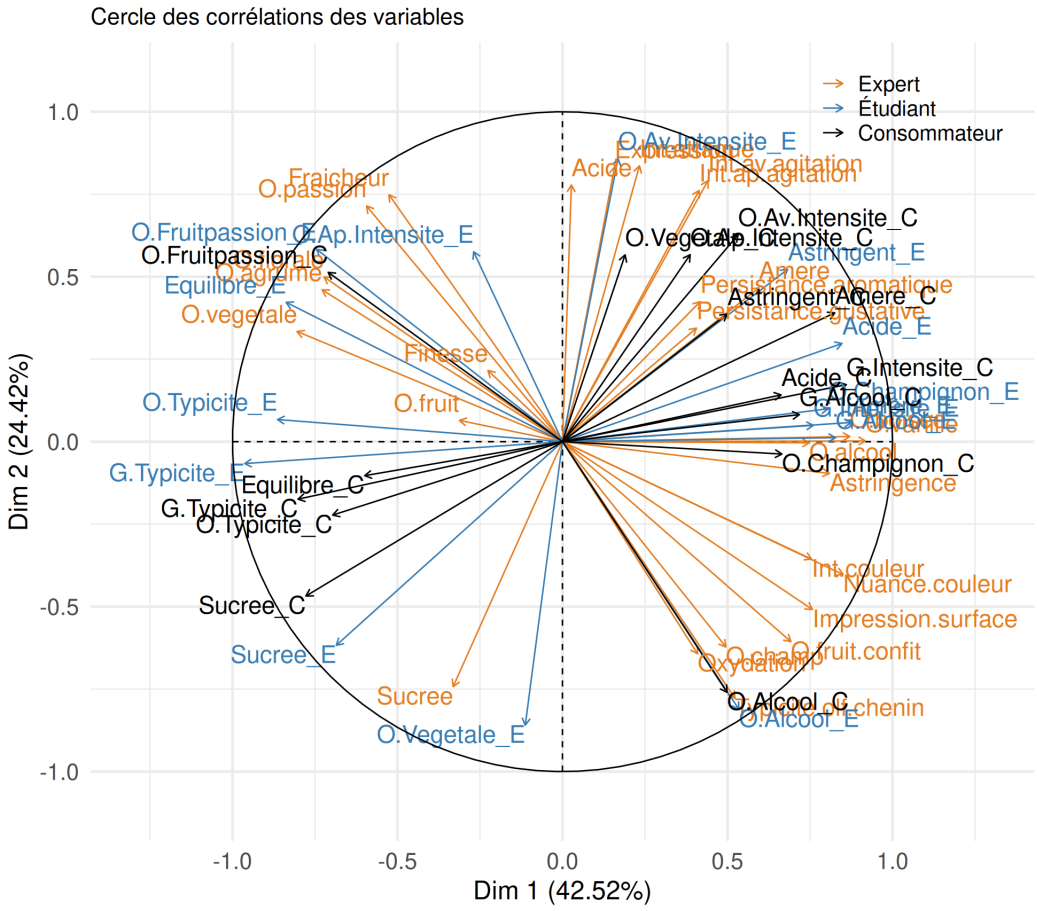
2.9 Étude des individus

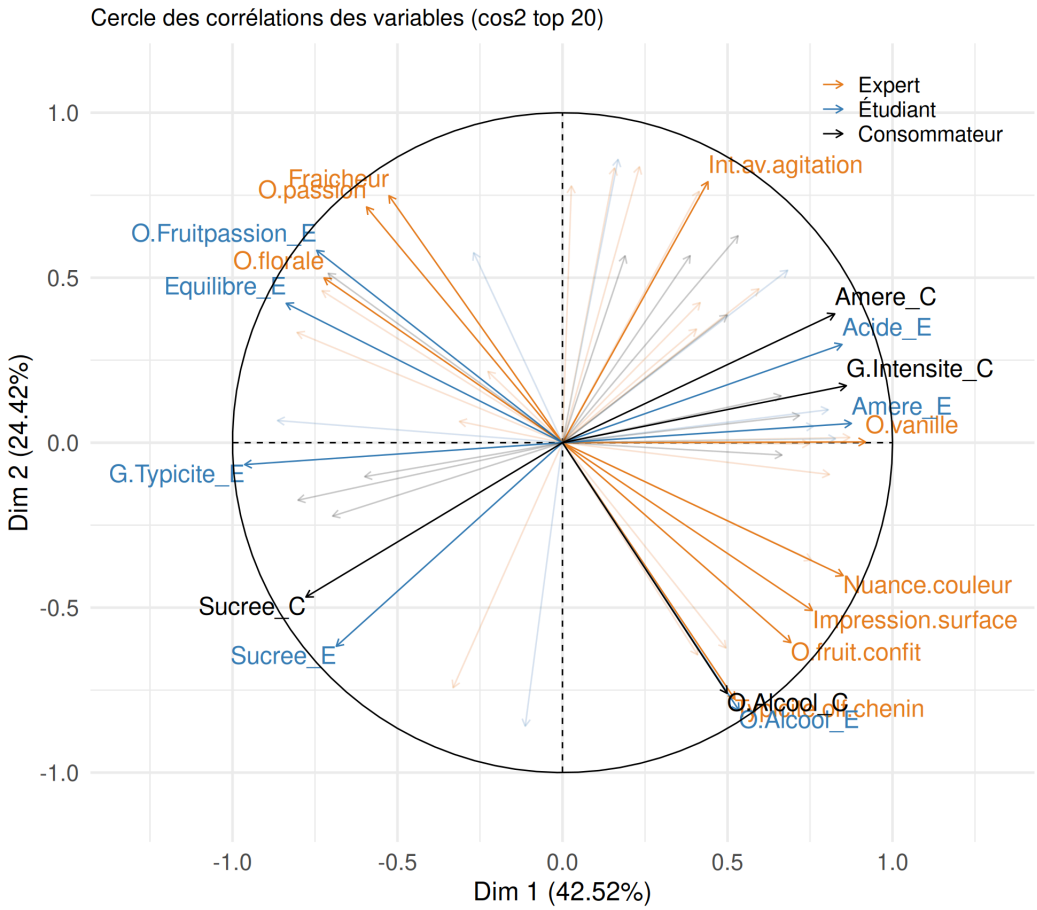
▶ **Plot**

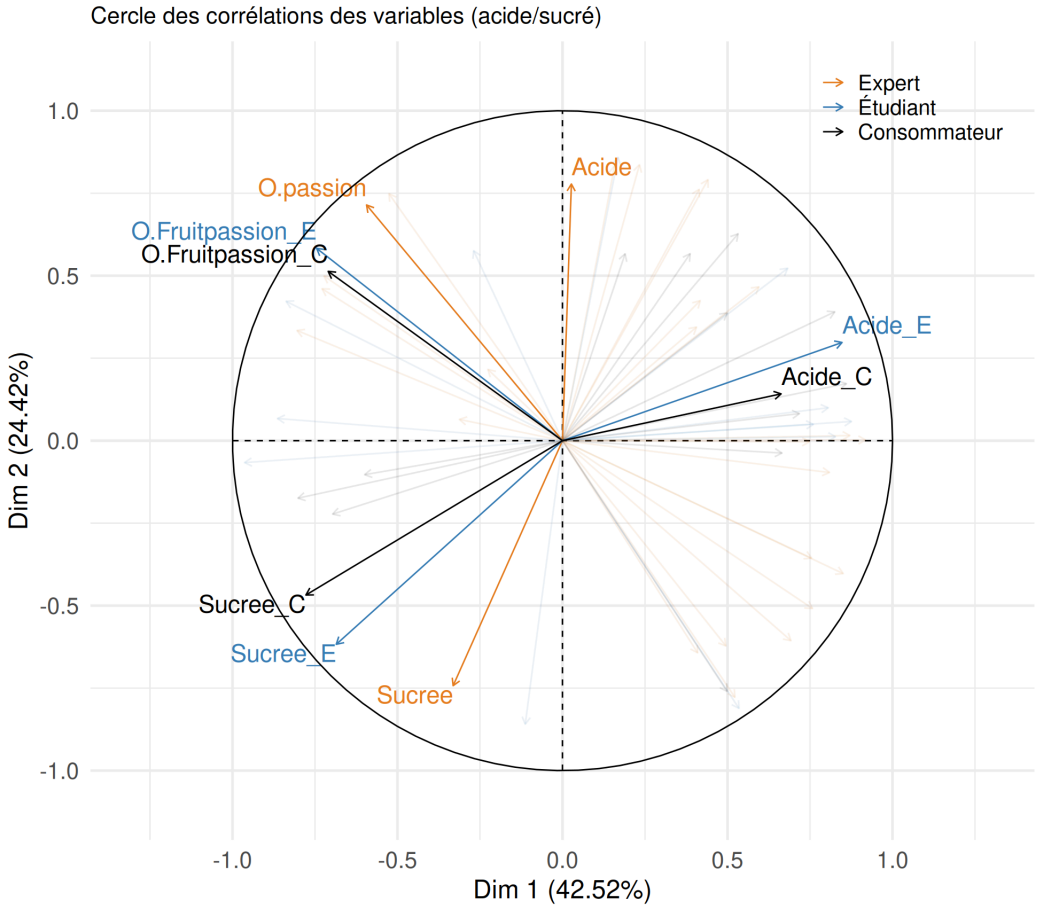


2.10 Étude des variables

▸ Graphique (par groupe)







3 Représentation des groupes

3.1 Caractérisation de la première composante

Formulation mathématique

- **En ACP classique** (centrée-réduite), la première composante factorielle v_1 (de norme 1) cherche à maximiser la liaison globale avec l'ensemble des p variables :

$$v_1 = \underset{\|v\|=1}{\operatorname{argmax}} \sum_{j=1}^{p} \operatorname{cov}^2(v, x_{\cdot,j})$$

- **En AFM (ACP pondérée)**, chaque variable du groupe k est pondérée par $1/\sqrt{\lambda_1^{(k)}}$. La première composante v_1 de l'AFM satisfait donc :

$$\begin{aligned} v_1 &= \underset{\|v\|=1}{\operatorname{argmax}} \sum_{k=1}^{K} \sum_{j=1}^{J_k} \operatorname{cov}^2\left(v, \frac{x_{:,(k,j)}}{\sqrt{\lambda_1^{(k)}}}\right) \\ &= \underset{\|v\|=1}{\operatorname{argmax}} \sum_{k=1}^{K} \frac{1}{\lambda_1^{(k)}} \sum_{j=1}^{J_k} \operatorname{cov}^2(v, x_{:,(k,j)}) \end{aligned}$$

3.2 Interprétation géométrique

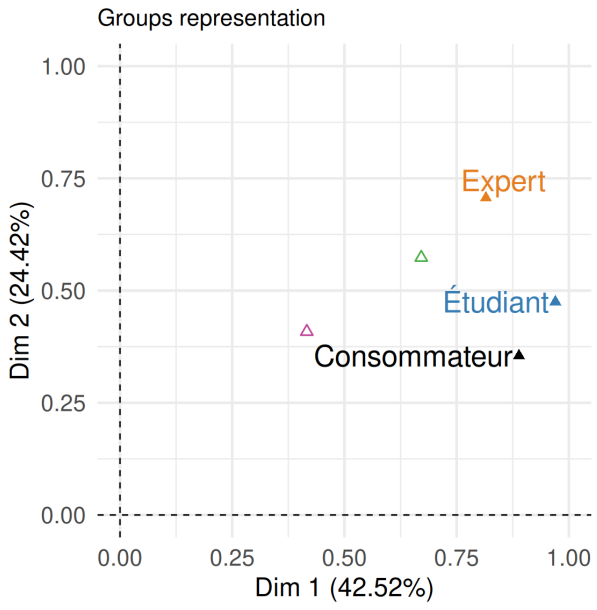
Propriété de borne supérieure (Équilibre inter-groupes)

Soit $I_k(v) = \sum_{j=1}^{J_k} \text{cov}^2(v, x_{:,(k,j)})$ définie pour tout $v \in \mathbb{R}^J$ tel que $\|v\| = 1$. L'inertie du groupe k projetée sur la direction v, est

$$0 \leq \mathcal{L}_g(k, v) = \frac{I_k(v)}{\lambda_1^{(k)}} \leq 1$$

- Interprétations,
  - $\mathcal{L}_g(k, v) = 0$ si toutes les variables du groupe k sont non-corrélées à v
  - $\mathcal{L}_g(k, v) = 1$ si v correspond à la 1ère composante principale du groupe k.
- Conséquences :
  - **Contribution maximale égalisée :** Quel que soit le nombre de variables J_k ou la structure du groupe k, sa contribution maximale à la première dimension de l'AFM vaut **au plus 1**.
  - **Consensus :** La composante v_1 s'ajuste pour être la plus proche possible des premières composantes de l'ensemble des K groupes.

3.3 Représentation des groupes



Les inerties projetées

```
> inertie_groupe_dim12 <-
      vins.mfa$group$coord[, 1:2]
```

	Dim1	Dim2
Expert	0.815	0.706
Étudiant	0.970	0.475
Consommateur	0.889	0.355

La somme des inerties des groupes actifs = inertie de l'axe 1 de l'AFM

```
> cat("Somme des coord des groupes (Axes 1-
      2) :",
      colSums(inertie_groupe_dim12), "\n",
+     "Valeurs propres de l'Axe 1-2 (AFM)
      :", vins.mfa$eig[1:2, 1])
```

```
Somme des coord des groupes (Axes 1-2) :
2.674248 1.536237
 Valeurs propres de l'Axe 1-2 (AFM)     :
2.674248 1.536237
```

- 1ère dimension commune aux 3 groupes (abscisses proches de 1)
- 2ème dimension liée surtout au groupe Expert

3.4 Similarité entre groupes

▸ **Formulations mathématiques**

- **Liaison globale :** Le coefficient $\mathcal{L}_g$ mesure la covariance globale croisée entre deux groupes de variables k et k' :

$$\mathcal{L}_g(k,k') = \sum_{j=1}^{J_k}\sum_{j'=1}^{J_{k'}} \mathrm{cov}^2\left(\frac{x_{:,(k,j)}}{\sqrt{\lambda_1^{(k)}}}, \frac{x_{:,(k',j')}}{\sqrt{\lambda_1^{(k')}}}\right) = \frac{1}{\lambda_1^{(k)}\lambda_1^{(k')}}\sum_{j=1}^{J_k}\sum_{j'=1}^{J_{k'}} \mathrm{cov}^2\left(x_{:,(k,j)}, x_{:,(k',j')}\right)$$

- **Correlation généralisée :** Le coefficient RV est la version normalisée du coefficient $\mathcal{L}_g$

$$RV(k,k') = \frac{\mathcal{L}_g(k,k')}{\sqrt{\mathcal{L}_g(k,k)\,\mathcal{L}_g(k',k')}}$$

Interprétation des bornes du $RV \in [0;1]$

- $RV(k,k') = 0$ **:** Indépendance linéaire totale. Toutes les variables du groupe k sont orthogonales à toutes les variables du groupe k'.
- $RV(k,k') = 1$ **:** Structure identique. Les nuages d'individus générés séparément par l'ACP du groupe k et l'ACP du groupe k' sont rigoureusement identiques (à un facteur d'échelle/homothétie et une rotation près).

3.5 Indice de multidimensionnalité

▸ **Formulations mathématiques**

Alors que les termes hors-diagonale $\mathcal{L}_g(k,k')$ mesurent la liaison entre deux groupes distincts, les termes diagonaux $\mathcal{L}_g(k,k)$ caractérisent la structure interne du groupe k :

$$\mathcal{L}_g(k,k) = \frac{1}{\left(\lambda_1^{(k)}\right)^2}\sum_{j=1}^{J_k}\left(\lambda_j^{(k)}\right)^2 = 1 + \sum_{j=2}^{J_k}\left(\frac{\lambda_j^{(k)}}{\lambda_1^{(k)}}\right)^2$$

Propriétés clés de $\mathcal{L}_g(k,k)$

- Borne inférieure : $\mathcal{L}_g(k,k) \geq 1$ pour tout groupe.
- Groupe unidimensionnel ($\mathcal{L}_g(k,k) = 1$) : Toutes les variables du groupe sont parfaitement colinéaires (une seule dimension suffit à tout résumer).
- Groupe multidimensionnel ($\mathcal{L}_g(k,k) > 1$) : Le groupe porte plusieurs dimensions d'information indépendantes non négligeables. Plus $\mathcal{L}_g(k,k)$ est élevé, plus l'information du groupe est complexe/riche.

4 Individus partiels, ACPs séparées

Questions d'intérêt

- Les profils d'individus sont-ils similaires entre les différents groupes ?
- Y a-t-il des individus remarquables pour certains groupes de variables ?

Pour que ce soit pertinent, on doit comparer les profils d'individus des groupes dans un **espace commun**. On va les comparer dans l'espace obtenu par l'AFM.

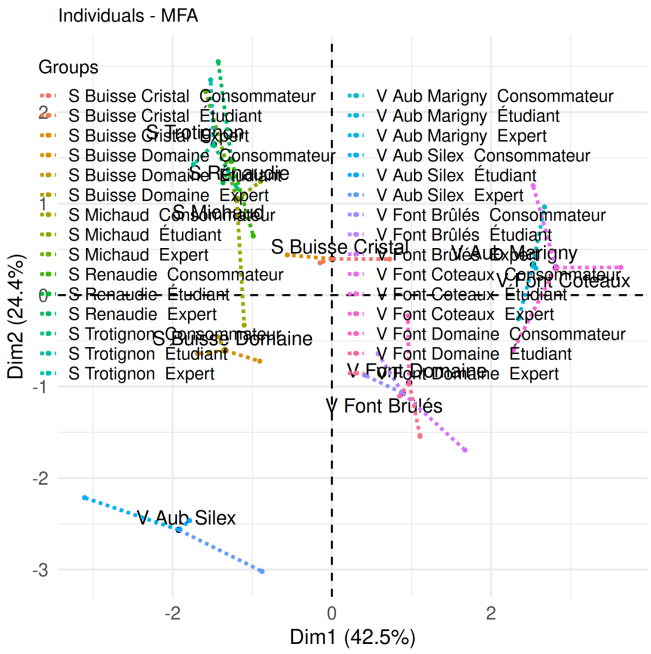
Définition

Pour chaque groupe κ de variables, on note $X^{(\kappa)}$ la matrice des individus partiels, construite à partir de X comme suit :

$$X^{(\kappa)}_{i,(k,j)} = \begin{cases} x_{i,(k,j)} & \text{si } k = \kappa, \\ 0 & \text{sinon.} \end{cases}$$

	Groupe A			Groupe B	
	var 1	**var 2**	**var 3**	**var 1**	**var 2**
Individu 1	$x_{1,(A,1)}$	$x_{1,(A,2)}$	$x_{1,(A,3)}$	$x_{1,(B,1)}$	$x_{1,(B,2)}$
Individu partiel A	$x_{1,(A,1)}$	$x_{1,(A,2)}$	$x_{1,(A,3)}$	0	0
Individu partiel B	0	0	0	$x_{1,(B,1)}$	$x_{1,(B,2)}$

4.1 Projection des points partiels



On projette les individus partiels supplémentaires[1] dans le plan principal, et on observe leur position par rapportaux individus associés :

[1] Un facteur d'échelle de K doit être appliqué pour garantir que l'individu est le barycentre de ses individus partiels

- `V Aub Silex` a été jugé bien plus extrêmes par `Conso` que par `Expert`, relativement à la première direction ;
- `S Michaud` a été jugé plus extrêmes par `Expert` que par `Conso`, relativement à la deuxième direction

Il y a peu d'écart pour les points partiels de `Etudiant`.

4.2 Étude des inerties

Décomposition

$$\text{Inertie}_{\text{tot},r} = \text{Inertie}_{\text{inter},r} + \text{Inertie}_{\text{intra},r}$$

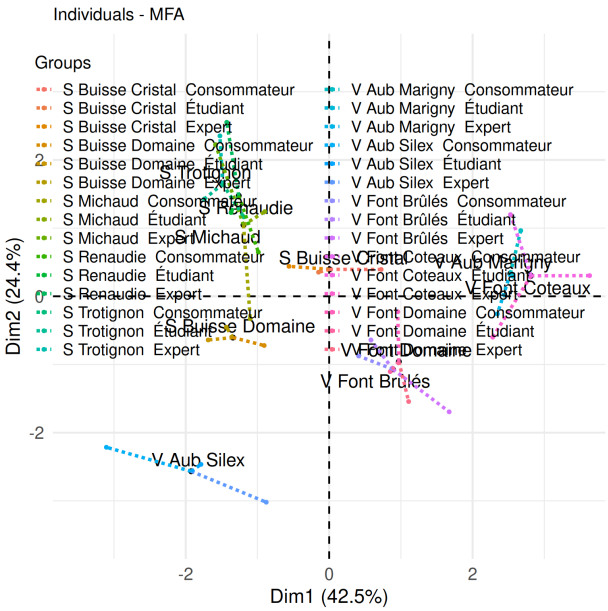
$$\sum_{i=1}^{n}\sum_{k=1}^{K}\left(C_r^{(\text{ind},k)}(i)\right)^2 = K\sum_{i=1}^{n}\left(C_r^{(\text{ind})}(i)\right)^2 + \sum_{i=1}^{n}\sum_{k=1}^{K}\left(C_r^{(\text{ind},k)}(i) - C_r^{(\text{ind})}(i)\right)^2$$

avec :

- $C_r^{(\text{ind},k)}(i)$: coordonnée sur l'axe r de l'individu partiel i selon le groupe k.
- $C_r^{(\text{ind})}(i)$: coordonnée sur l'axe r de l'individu moyen (consensus) i, telle que $C_r^{(\text{ind})}(i) = \frac{1}{K}\sum_{k=1}^{K} C_r^{(\text{ind},k)}(i)$.

À partir de ces inerties, on peut étudier les écarts entre les points partiels et les points moyens associés.

4.3 Étude des inerties



$$\frac{\text{Inertie}_{\text{inter},1}}{\text{Inertie}_{\text{tot},1}}$$

	data
Dim.1	0.928
Dim.2	0.818
Dim.3	0.777

- Les coordonnées des points partiels selon la 1ère direction sont globalement proches de celles de leur point moyen.
- On peut s'intéresser aux individus avec de grandes inerties intra pour avoir des informations plus précises sur les différences entre les groupes.

4.4 Inertie intra individus




```
> id <- grep("V Aub Silex", rownames(vins.mfa$ind$coord))
> print_table(vins.mfa$ind$within.inertia[id, 1:2, drop =
      FALSE])
```

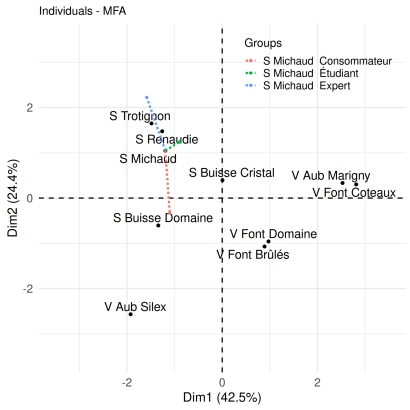
	Dim.1	Dim.2
V Aub Silex	40.078	3.314

- L'inertie intra de `V Aub Silex` pour la 1ère direction est très élevée: plus fortes différences entre les groupes selon cette direction.

```
> ids <- grep("V Aub Silex",
      rownames(vins.mfa$ind$coord.partiel))
> print_table(vins.mfa$ind$within.partial.inertia[ids,
      1:2])
```

	Dim.1	Dim.2
V Aub Silex .Expert	17.470	2.005
V Aub Silex .Étudiant	0.296	0.101
V Aub Silex .Consommateur	22.312	1.208

- Cette inertie intra selon la 1ère direction vient surtout des groupes `Expert` et `Conso`.




```
> id <- grep("S Michaud", rownames(vins.mfa$ind$coord))
> print_table(vins.mfa$ind$within.inertia[id, 1:2, drop =
      FALSE])
```

	Dim.1	Dim.2
S Michaud	3.979	32.333

- L'inertie intra de `S Michaud` pour la 2ème direction est très élevée: plus fortes différences entre les groupes selon cette direction.

```
> ids <- grep("S Michaud",
      rownames(vins.mfa$ind$coord.partiel))
> print_table(vins.mfa$ind$within.partial.inertia[ids,
      1:2])
```

	Dim.1	Dim.2
S Michaud .Expert	2.432	13.542
S Michaud .Étudiant	1.408	0.374
S Michaud .Consommateur	0.139	18.417

- Cette inertie intra selon la 1ère direction vient surtout des groupes `Expert` et `Conso`.

4.5 AFM vs ACPs séparées

Comment comparer l'AFM (= ACP globale pondérée) avec les ACPs de chacun des groupes ?

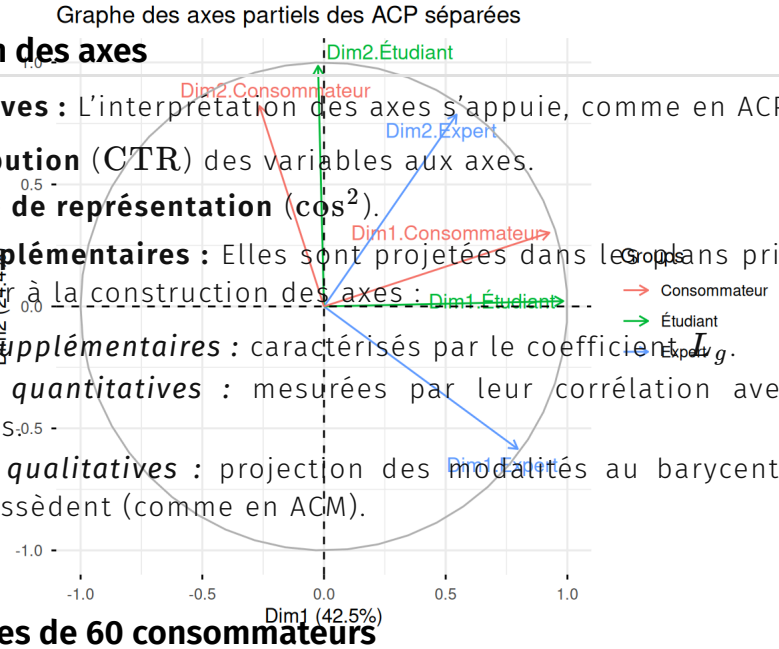
Projection des composantes principales : On met en œuvre une méthode similaire à celle des individus partiels : - On effectue pour chaque groupe une ACP classique ; - On extrait les (premières) composantes principales des ACPs ; - On projette ces composantes en tant que variable supplémentaire

5 Aide à l'interprétation

~~4.6 Exemple~~

5.1 Interprétation des axes

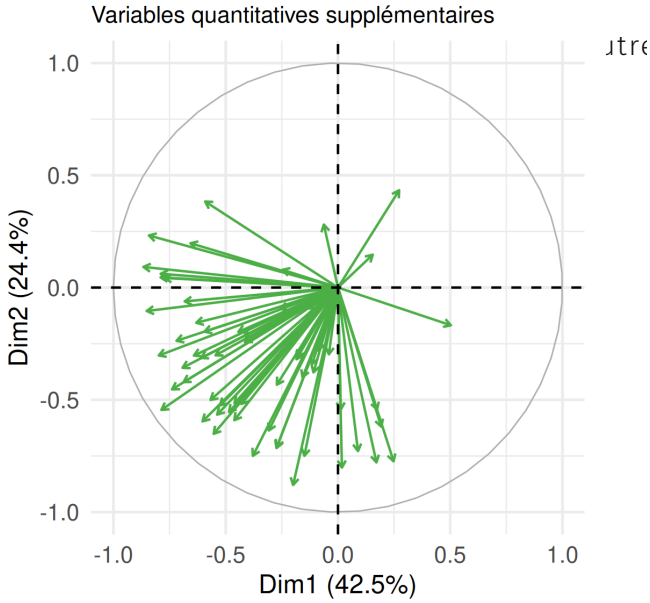
- **Variables actives :** L'interprétation des axes s'appuie, comme en ACP/ACM, sur :
  - La **contribution** (CTR) des variables aux axes.
  - La **qualité de représentation** ($\cos^2$).
- **Variables supplémentaires :** Elles sont projetées dans les plans principaux de l'AFM sans participer à la construction des axes :
  - *Groupes supplémentaires :* caractérisés par le coefficient I_g.
  - *Variables quantitatives :* mesurées par leur corrélation avec les directions principales.
  - *Variables qualitatives :* projection des modalités au barycentre des individus qui les possèdent (comme en ACM).

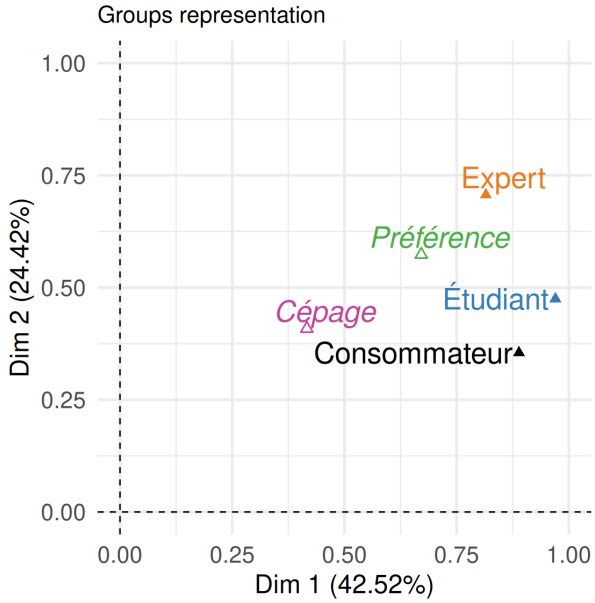


5.2 Exemple: Notes de 60 consommateurs

Variables quantitatives: note pour chaque vin

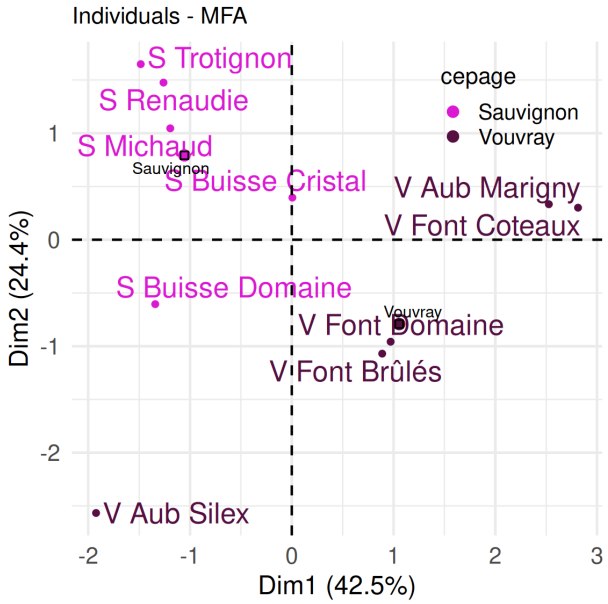
- Les 2 premières directions du groupe `Etudiant` coincident avec celle de l'AFM (voir

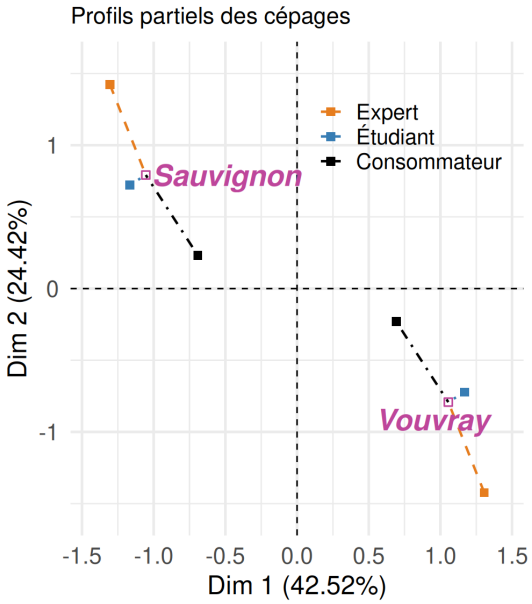




5.3 Exemple: Cépage de chacun des vins

Variable qualitative: *Sauvignon* ou *Vouvray*





5.4 Données qualitatives et mixtes

Données qualitatives : L'AFM s'étend naturellement aux tableaux composés de variables qualitatives : - L'**ACM** (Analyse Factorielle des Correspondances Multiples) remplace l'ACP dans l'étape préalable d'analyse groupe par groupe. - La pondération du groupe j reste définie par $w_j = \frac{1}{\lambda_1^{(j)}}$, où $\lambda_1^{(j)}$ est la ***première valeur propre de l'ACM*** de ce groupe.

Données mixtes : L'AFM permet de traiter simultanément des groupes d'actives *quantitatives et qualitatives* :

- **Pondération unifiée** $w_j = \frac{1}{\lambda_1^{(j)}}$:
  - Groupe ***quantitatif*** $\rightarrow \lambda_1^{(j)}$ issu d'une ***ACP*** (centrée-réduite ou non).
  - Groupe ***qualitatif*** $\rightarrow \lambda_1^{(j)}$ issu d'une ***ACM***.
- ***Analyse globale*** : ACP pondérée appliquée au tableau juxtaposé, où les variables qualitatives sont codées sous forme de ***tableau disjonctif complet*** (TDC).