

AFC : l'exemple de la couleur des yeux et des cheveux

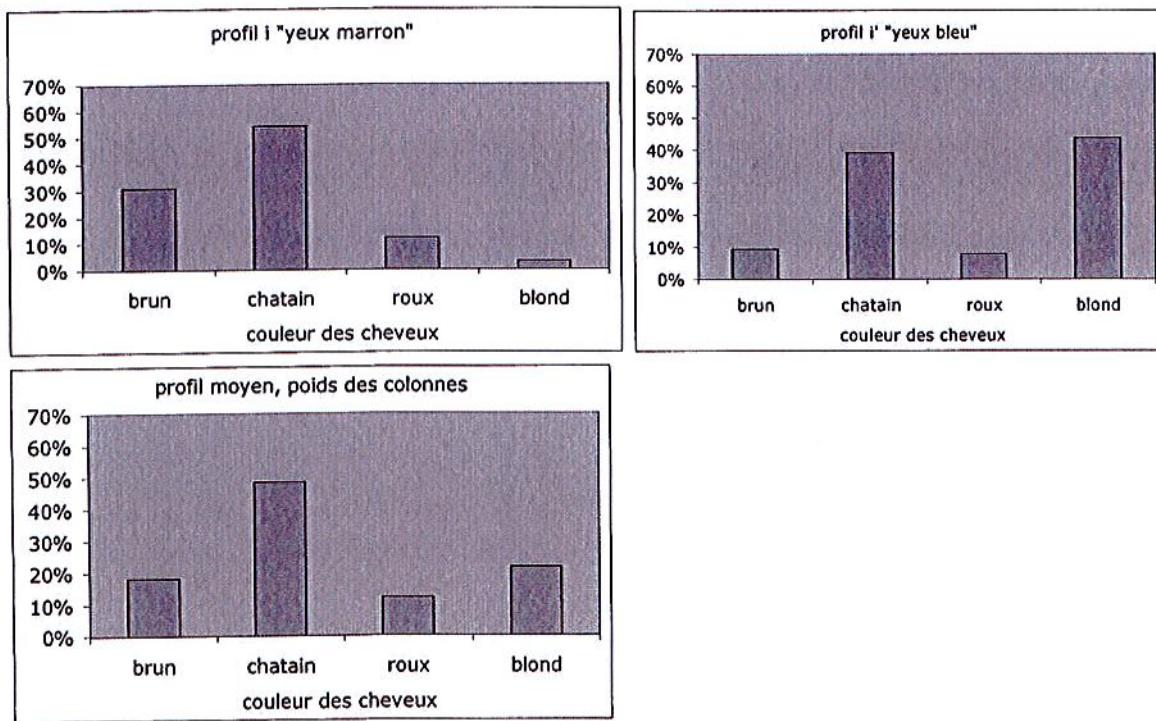
1 Analyse des profiles-lignes

Matrice L des quatre profiles-lignes de R^4 :

	brun	chatain	roux	blond
marron	31%	54%	12%	3%
noisette	16%	58%	15%	11%
vert	8%	45%	22%	25%
bleu	9%	39%	8%	44%
Profil moyen	18%	48%	12%	21%

Distances du chi-2:

$$d^2(\text{vert, noisette})=0,24 \text{ et } d^2(\text{marron, bleu})=1,08, \quad d^2(\text{marron, moyen})=0,2504$$

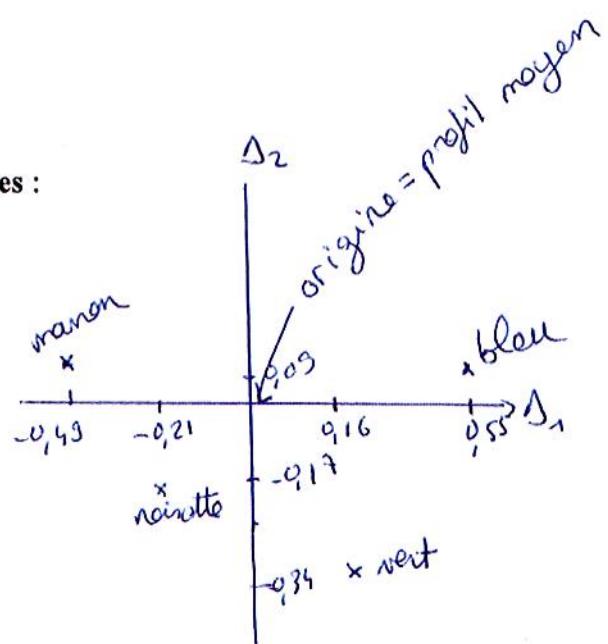
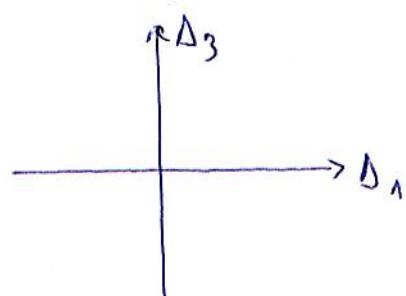


Matrice X des coordonnées factorielles des profils lignes :

	x^1	x^2	x^3
marron	-0,49	0,09	-0,02
noisette	-0,21	-0,17	0,10
vert	0,16	-0,34	-0,09
bleu	0,55	0,08	0,00

$$\bar{x}^1 = 0 \quad \bar{x}^2 = 0 \quad \bar{x}^3 = 0$$

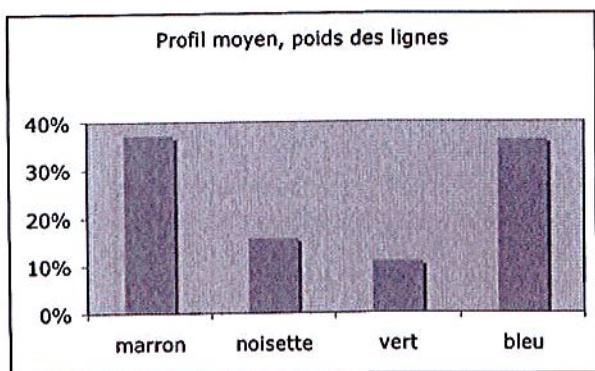
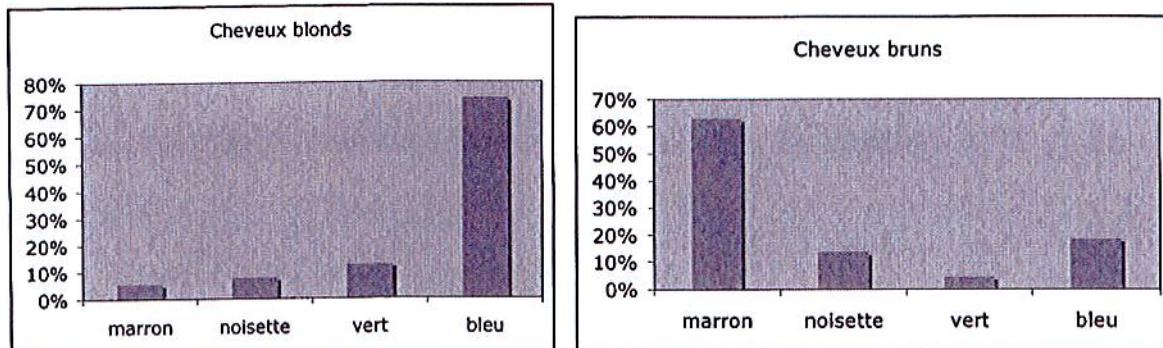
$$\text{var}(x^1) = 0,08 \quad \text{var}(x^2) = 0,022 \quad \text{var}(x^3) = 0,026$$



2 Analyse des profiles-colonnes

Matrice **C** des quatre profile-colonnes de R^4 :

	brun	chatain	roux	blond	Profil moyen
marron	63%	42%	37%	6%	37%
noisette	14%	19%	20%	8%	16%
vert	5%	10%	20%	13%	11%
bleu	19%	29%	24%	74%	36%

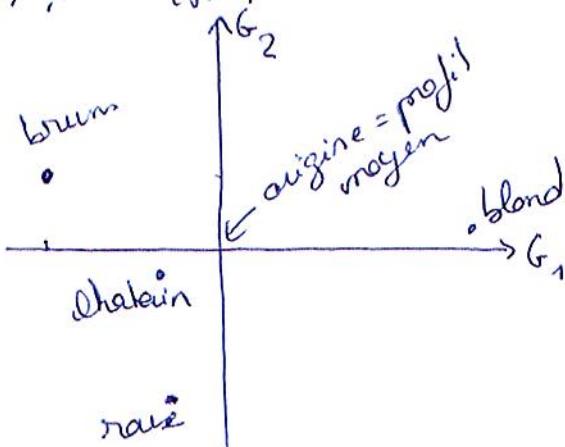


Matrice **Y** des coordonnées factorielles des profils colonnes :

	y¹	y²	y³
brun	-0,50	0,21	-0,06
chatain	-0,15	-0,03	0,05
roux	-0,13	-0,32	-0,08
blond	0,84	0,07	-0,02

$$\bar{y}^1 = 0 \quad \bar{y}^2 = 0 \quad \bar{y}^3 = 0$$

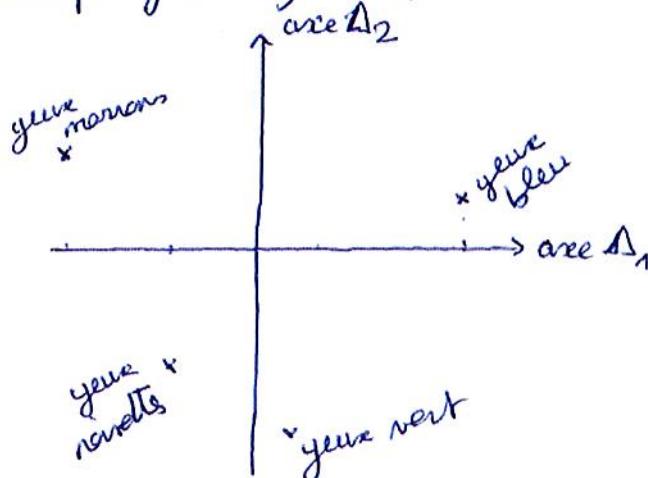
$$ren(y^1) = 0,208 \quad ren(y^2) = 0,022 \quad ren(y^3) = 0,026$$



1) Propriétés barycentriques: représentations simultanées des projets-yeux et des profils-colonnes: (3)

dernière représentation:

On dispose de la représentation en deux dimensions du nuage des profils-lignes (couleurs des yeux):



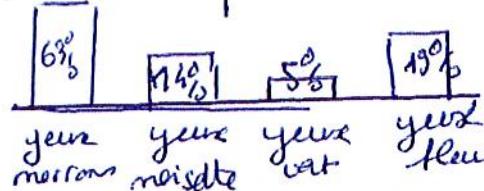
⇒ On veut utiliser la

relation barycentrique

$$y_{SD} = \frac{1}{f_{SD}} \sum_{i=1}^n \frac{f_i s_i}{f_{SD}} x_{i2}$$

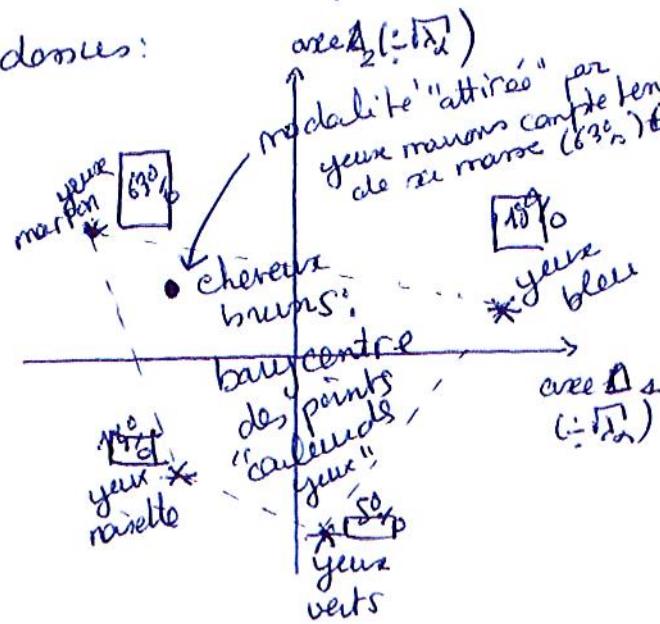
pour représenter en deux dimensions les points du nuage des profils-colonnes (couleur des cheveux) au centre de gravité des points du nuage des profils-lignes:

Par exemple, on veut positionner le point "cheveux bruns" ayant le profil suivant:

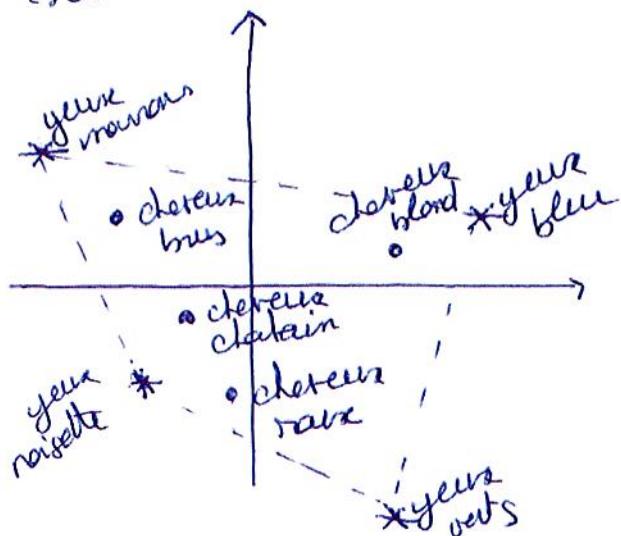


D'après la formule, on positionne "cheveux bruns" au barycentre des points i ($\hat{\equiv}$ par $\frac{1}{f_{SD}}$) représentant les couleurs des yeux, pondérés par les fréquences relatives $\frac{f_i s_i}{f_{SD}}$ du profil "cheveux brun".

ci-dessous:

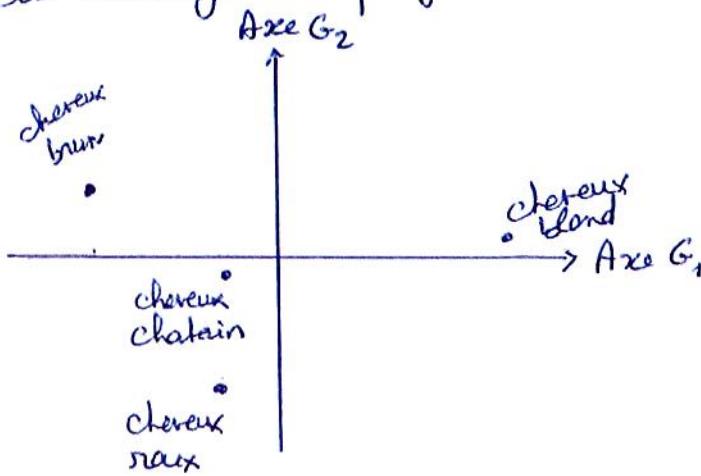


⇒ On construit ainsi la représentation simultanée suivante:



Seconde représentation

On considère maintenant la représentation en deux dimensions du nuage des profils colonnes (couleur des cheveux) :

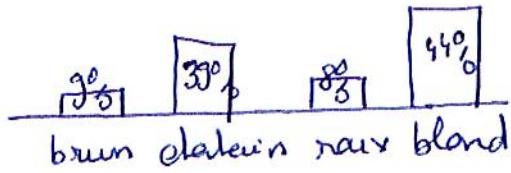


⇒ On utilise la relation重心
barycentrique

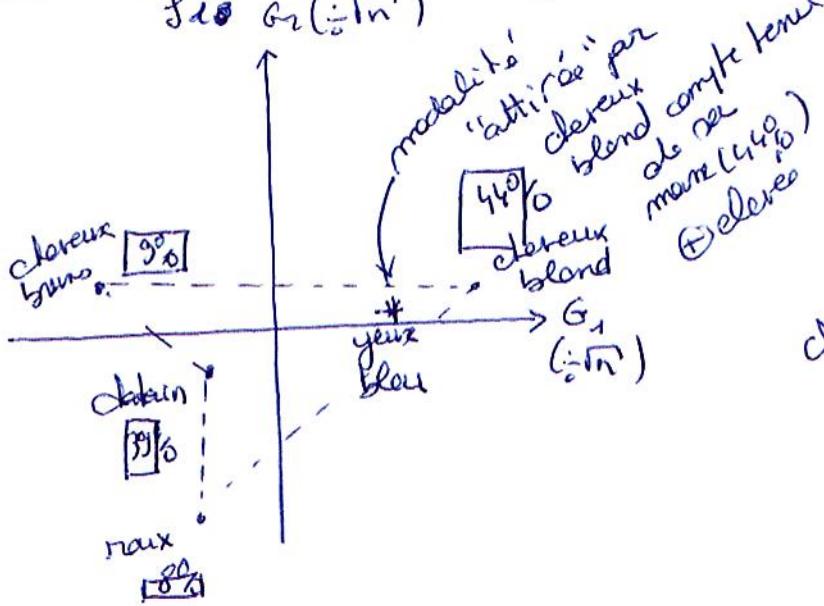
$$x_{\text{BD}} = \frac{1}{f_{\text{BD}}} \sum_{s=1}^q f_{is} y_{Gs}$$

pour représenter en deux dimensions les points du nuage des profils lignes (couleur des yeux) au centre de gravité des points du nuage des profils colonnes

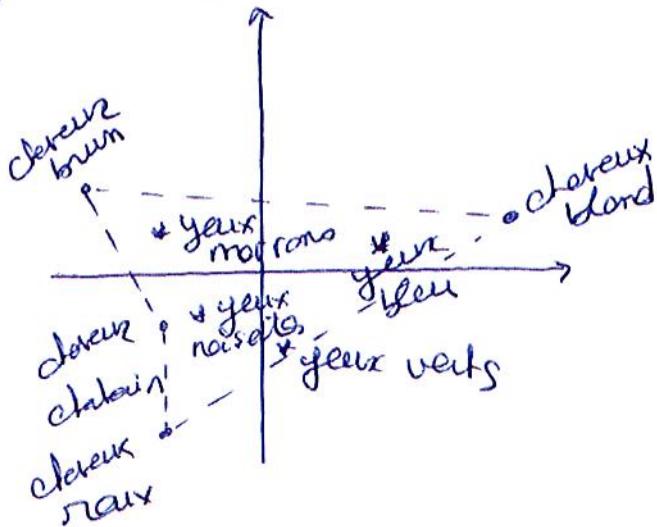
Par exemple, on veut positionner le point "yeux bleus" ayant le profil suivant :



On positionne donc "yeux bleus" au barycentre des points s (\vdots par \sqrt{n}) représentant les couleurs des cheveux, pondérés par les fréquences relatives f_{is} du profil "yeux bleus" ci-dessus :



⇒ On construit la représentation simultanée suivante :

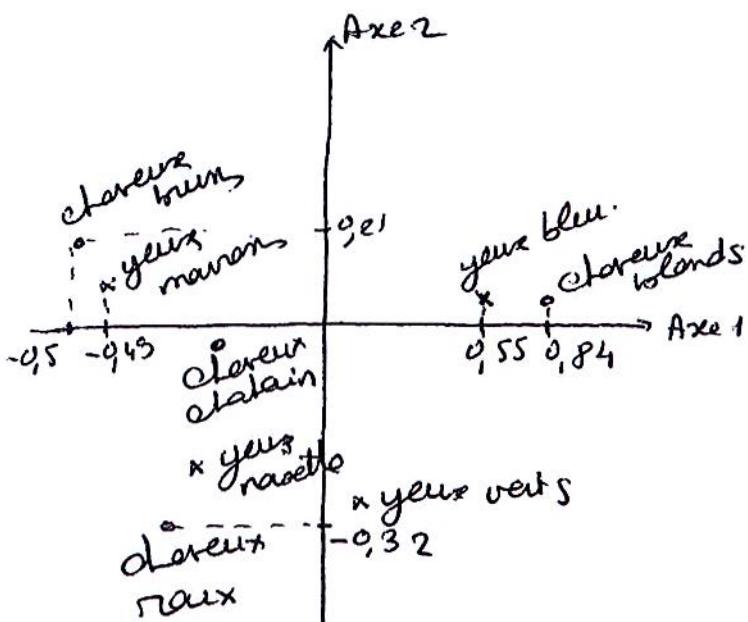


Troisième représentation : représentation simultanée des deux nuages de points

On a vu dans les deux représentations précédentes que l'on peut représenter les profils-colonnes au barycentre des profils-lignes. La première représentation par exemple nous indique que la modalité "éteveux bruns" est attirée par la modalité "yeux marrons" ce qui veut dire que parmi les personnes ayant les "éteveux bruns", il y en a une proportion plus grande que la moyenne (car on s'intéresse aux point excentrés, donc pas du point moyen) qui ont les yeux marrons. Cela correspond à la situation $\frac{f_{12}}{f_{11}} > f_{11}$

La seconde représentation quand à elle nous indique à l'inverse que la modalité "yeux marrons" est attirée par la modalité "éteveux bruns" ce qui veut dire que parmi les personnes ayant "les yeux marrons" il y en a une proportion plus grande que la moyenne qui ont les "éteveux bruns". Cela correspond à la situation $\frac{f_{12}}{f_{11}} > f_{12}$.

On voudrait maintenant avoir une seule représentation simultanée des deux nuages de points mais donnant une "idée" des liaisons entre les modalités des deux variables (celles pour lesquelles $f_{12} > f_{11} \cdot f_{22}$ c.a.d ayant un taux de liaison important). L'idée est alors de clouter sur chaque axe les centres de gravité par le coefficient $\frac{f_{12}}{f_{11}}$. On fait ainsi la représentation simultanée à colonnes qui présente sur un même graphique les deux séries étiquetées colonnes. Chaque point-ligne est à $\frac{1}{f_{11}}$ près au barycentre des profils colonnes et vice et versa.

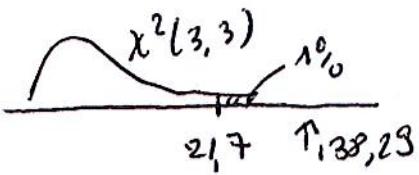


1) Interpretation des résultats
Inertie et tests d'indépendance

NO	VALEUR PROPRE	POUR-CENTAGE	POURCENT. CUMULE
1	.2088	89.37	89.37
2	.0222	9.51	98.89
3	.0026	1.11	100.00

Trace .2336 (= INERTIE TOTALE)

On a vu que l'inertie totale des nuages de profils ligne et colonne est égal à la somme des nuances de toutes ($k = \text{rang}(F)$), les composantes principales c'est à la somme des valeurs propres : Inertie = $0,233 = \frac{\chi^2}{m}$ Avec $n = 592$, $\Rightarrow \chi^2 = 138,29$.



→ rejet de H_0 ; couleur de cheveux et couleur des yeux sont indépendants

⇒ On peut avoir intérêt à faire un AFC pour d'évire cette dépendance

On va regarder l'interprétation que des deux premiers facteurs qui partent à eux deux 98,89% de l'inertie initiale
 ⇒ le premier plan factorial sautant que le premier axe domine!

Contributions et cosinus canons

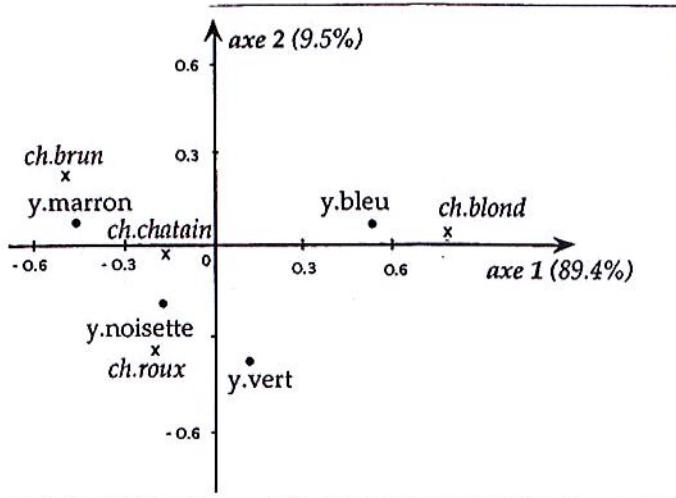
COLONNES	COORDONNEES			CONTRIBUTIONS			COSINUS CARRES		
	1	2	3	1	2	3	1	2	3
CHEVEUX									
Ch.Brun	-.50	.21	-.06	22.2	37.9	21.6	.84	.15	.01
Ch.chatain	-.15	-.03	.05	5.1	2.3	44.3	.86	.04	.09
Ch.roux	-.13	-.32	-.08	1.0	55.1	31.9	.13	.81	.05
Ch.blond	.84	.07	-.02	71.7	4.7	2.2	.99	.01	.00
LIGNES									
COORDONNEES			CONTRIBUTIONS			COSINUS CARRES			
YEUX	1	2	3	1	2	3	1	2	3
y.marron	-.49	.09	-.02	43.1	13.0	6.7	.97	.03	.00
y.noisette	-.21	-.17	.10	3.4	19.8	61.1	.54	.34	.12
y.vert	.16	-.34	-.09	1.4	55.9	31.9	.18	.77	.05
y.bleu	.55	.08	.00	52.1	11.2	.3	.98	.02	.00

$$ctr_2(i) = f_i \cdot \frac{x_{id}^2}{d^2(l_i, 2)}$$

yeux marrons:

$$ctr_2(\text{marrons}) = 0,37 \times \frac{(-0,45)^2}{0,2088} = 43,1$$

$$\cos^2 = \frac{x_{id}^2}{d^2(l_i, 2)} \Rightarrow \cos^2(\text{marron}) = \frac{-0,45^2}{0,2504} = 0,97$$



Plan factoriel des lignes:

- Les yeux bleu et marrons contribuent fortement à la construction de l'1^{er} axe (avec resp. 52,1 et 43,1%). Ils sont de plus bien représentés sur cet axe avec un \cos^2 resp. de 0,98 et 0,97.
- Les yeux vert contribuent à 55% au second axe et est relativement bien représenté sur cet axe avec un \cos^2 de 0,77.
- En terme de distance entre les modalités "couleurs des yeux", il semble que les deux profils "vert" et "noisette" soient relativement proche, mais la qualité de leur projection sur le plan est moyenne. Il faut donc prendre des précautions.

Le profil moyen des couleurs

- Les cheveux blond et bruns contribuent largement à l'inertie du premier axe (avec resp. 22% et 71%). Ils sont bien représentés sur cet axe (\cos^2 de 0,93 et 0,84) avec cheveux chatains qui joue contre moi contribue pas à l'inertie. En effet, il est proche de l'origine donc proche du profil moyen!
- Le second axe (avec seulement 9,5% d'inertie expliquée) est essentiellement construit par la couleur rousse (55%) qui n'appartient pas au brun (37,9%). Mais la couleur rousse est la seule bien représentée sur l'axe 2 (\cos^2 de 0,81)
- En ce qui concerne les proximités entre les couleurs de cheveux, on note que sur le premier axe, châtain est + proche de brun que de blond (qui sont les trois modalités bien représentées sur cet axe). On ne peut pas interpréter la proximité entre châtain et rousse sur cet axe car rousse est mal projeté (\cos^2 de 0,13)

Représentation simultanée

- Sur le premier axe, les yeux bleus s'associent au cheveux blonds, les yeux marrons aux cheveux bruns. Cela se vérifie sur les taux de liaison: cela vient du... de la catégorie cheveux châtain est proche de l'origine représentant le profil moyen et n'est caractéristique que d'aucune couleur des yeux
- Sur le second axe, les cheveux rousse n'associe aux yeux vert et dans une moindre mesure aux yeux marron (\cos^2 resp. 0,77 et 0,34). Cela se vérifie sur les

Taux de liaison:

	brun	châtain	rouge	blond
marron	(69%)	12%	-1%	-85%
mascotte	-12%	20%	26%	-50%
vert	-57%	-6%	(82%)	17%
bleu	-49%	-19%	-34%	(104%)

... et aussi... et aussi... et aussi...