

Projet Musculation

Aicha Ettriki

Sujet:

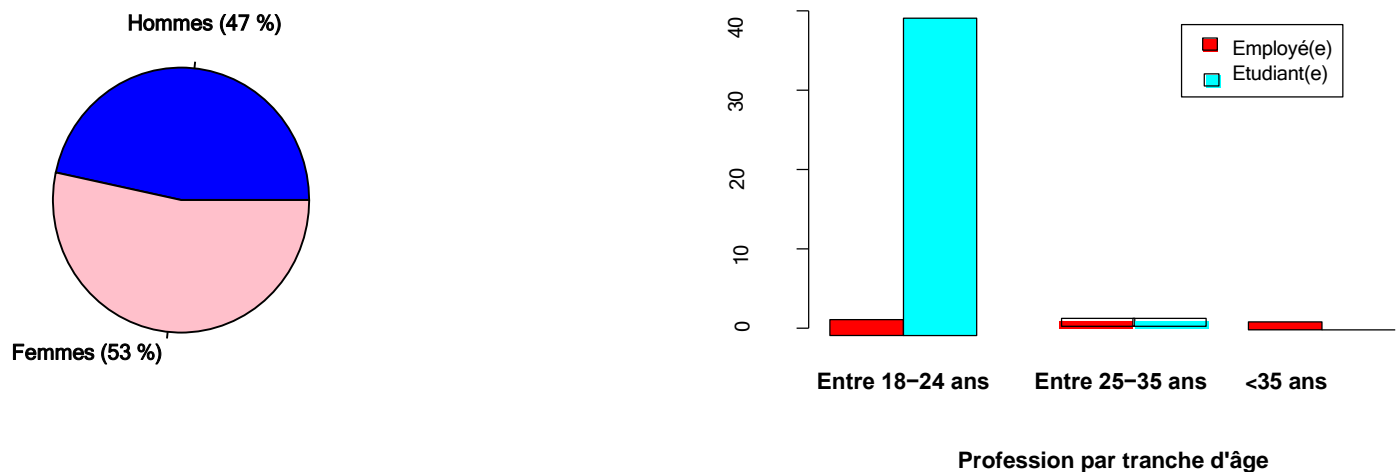
Cette enquête vise à comprendre le niveau de satisfaction globale des individus concernant leur progression physique, leur alimentation et leur santé mentale (1ere ACP) ainsi que la relation entre la pratique de la musculation et l'impact de la musique pendant l'entraînement (2eme ACP).

Nous devons tout d'abord charger les librairies dont nous aurons besoin

```
library(FactoMineR)
library(factoextra)
library(readxl)
library(corrplot)
library(ggplot2)
library(RColorBrewer)
```

```
Formulaire_Musculation_reponses_ <- read_excel(
  "/cloud/project/Formulaire Musculation (réponses).xlsx")
```

Nous pouvons voir un petit aperçu sur les personnes qui ont répondu à notre formulaire :



Nous remarquons que 53% des réponses sont des femmes et que plus de 40 personnes sont des étudiants.

Premiere ACP

1- Traitement des données

Nous allons changer les noms des colonnes du premier bloc pour qu'ils soient plus lisibles.

```
variables_quantitatives1 <- Formulaire_Musculation_reponses_[, (11:20)]

colonnes_a_modifier <- colnames(Formulaire_Musculation_reponses_)[11:20]

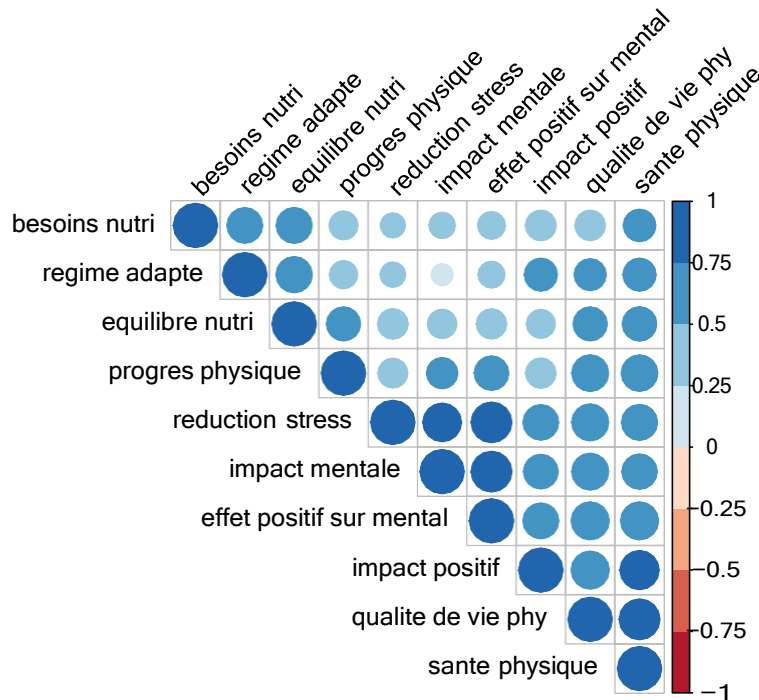
nouveaux_noms_colonnes <- c("regime adapte", "equilibre nutri", "besoins nutri",
  "impact mentale", "reduction stress", "effet positif sur mental",
  "qualite de vie phy", "progres physique",
  "sante physique", "impact positif")

colnames(Formulaire_Musculation_reponses_)[11:20] <- nouveaux_noms_colonnes
```

2-Pertinence de l'ACP

Pour mieux comprendre les relations entre les différentes variables de notre ensemble de données, nous avons construit une matrice de corrélation:

```
colonnes_selectionnees <- Formulaire_Musculation_reponses[, 11:20]
matrice_cor <- cor(colonnes_selectionnees)
corrplot(matrice_cor, method = "circle", type="upper", order="hclust",
          col=brewer.pal(n=8, name="RdBu"), tl.col = "black", tl.srt = 45)
```



Il semble y avoir une forte corrélation entre certaines variables : “impact mental”, “réduction de stress” et “effet positif sur le mental”. De même, une corrélation similaire est observée entre “amélioration de la qualité de vie physique”, “progrès physique”, “impact sur d’autres aspects” et “amélioration de la santé physique”. De plus, “régime adapté”, “équilibre nutritionnel” et “compréhension des besoins nutritionnels” semblent être étroitement corrélés.

Exécution de la fonction PCA.

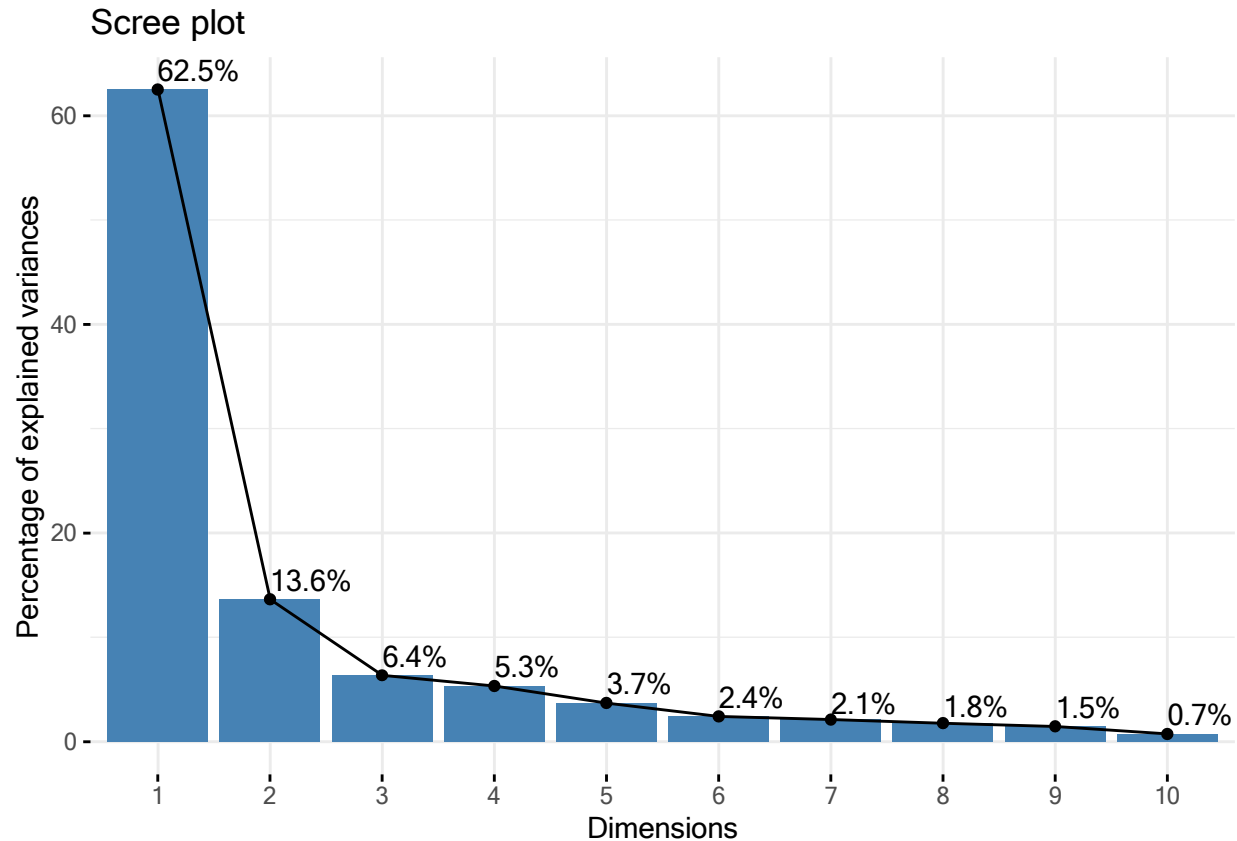
```
pca1 <- PCA(colonnes_selectionnees, graph = FALSE)
```

3-Choix du nombre d’axes à retenir

##	eigenvalue	percentage of variance	cumulative percentage of variance
## comp 1	6.2508764	62.508764	62.50876
## comp 2	1.3634248	13.634248	76.14301
## comp 3	0.6354348	6.354348	82.49736
## comp 4	0.5333059	5.333059	87.83042
## comp 5	0.3689709	3.689709	91.52013
## comp 6	0.2415493	2.415493	93.93562

- Critère de kaiser : on remarque qu’il y a 2 axes dont les valeurs propres sont supérieures à 1 donc on retient 2 axes d’après ce critère.
- Critère du taux d’inertie cumulée : les 2 premiers axes capturent une proportion considérable de l’inertie totale, ce qui signifie qu’ils sont essentiels.

```
fviz_screplot(pca1,addlabels=TRUE, ncp=10)
```



c. Critère du coude : On remarque que le coude se trouve au niveau du troisième axe, mais nous choisissons de retenir les 2 premiers axes.

Affichage des deux premières dimensions

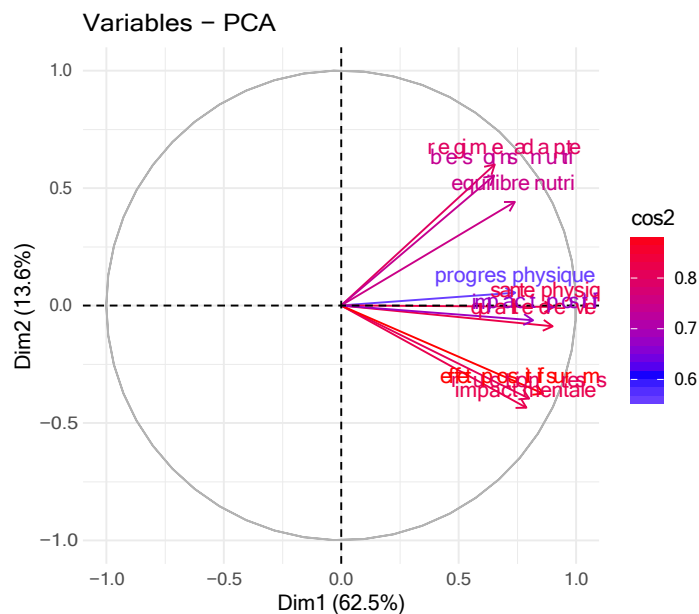
##	Dim.1	Dim.2
## regime adapte	0.6546184	0.600745111
## equilibre nutri	0.7395600	0.440738464
## besoins nutri	0.6510208	0.553260967
## impact mentale	0.7885157	-0.435441076
## reduction stress	0.8006376	-0.396634298
## effet positif sur mental	0.8595874	-0.374901149
## qualite de vie phy	0.9004468	-0.088378672
## progres physique	0.7413260	0.054232990
## sante physique	0.9058257	-0.006280071
## impact positif	0.8179347	-0.062545863

Nous pouvons nommer la dimension 1 “Amélioration physique et mentale” car elle dépend de toutes les variables en particulier “Amélioration de la qualité de vie physique” et ” Amélioration de la santé physique”. La Dimension 2 “Nutrition” car elle dépend surtout du “régime adapté” et des “besoins nutritionnels”.

Nous allons confirmer ce que nous avons dit avec le cercle ci-dessous

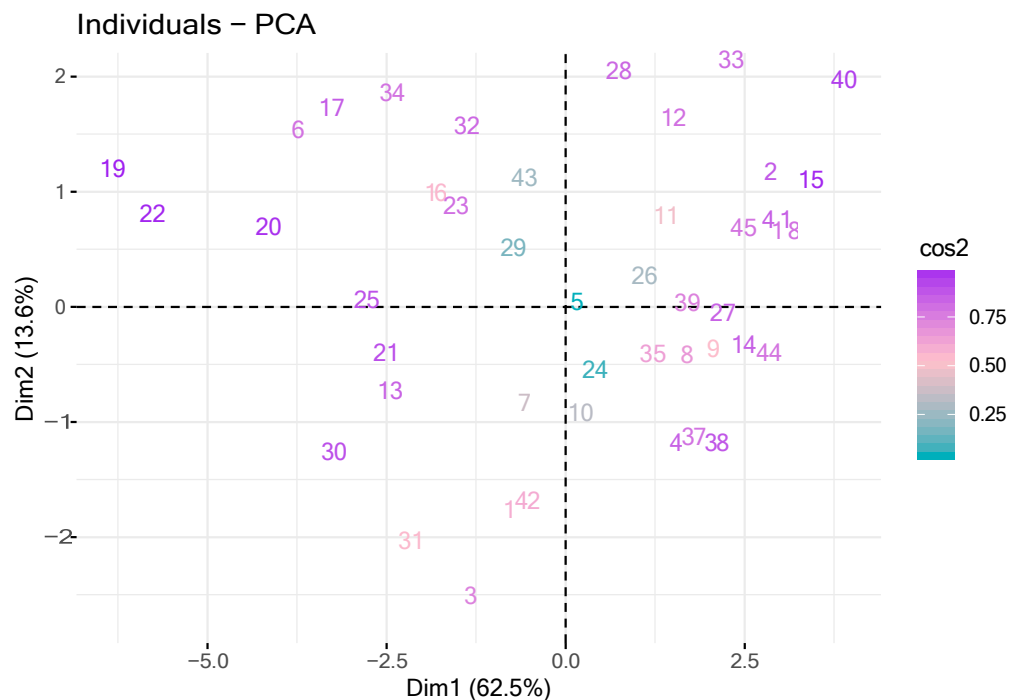
4-Interprétation du cercle

```
fviz_pca_var(pca1, col.var = "cos2") + scale_color_gradient2(low = "white", mid = "blue",
  high = "red", midpoint = 0.6) + theme_minimal()
```



5-Interprétation de la carte des individus

```
fviz_pca_ind(pca1, geom = "text", col.ind="cos2", gradient.cols=c("#00AFBB","pink","purple"))
```



D'après le graphique ci-dessous nous pouvons observer que nos données sont éparpillées et qu'il y a 3 groupes. Le premier est en haut à droite (2, 15, 45, 41,18) avec une forte contribution de cos2, il fait attention à leurs besoins nutritionnels et voient une amélioration physique et mentale. Le groupe (8, 9, 14, 26, 27, 30,35, 36, 39, 44) qui est neutre par rapport à la nutrition et qui voit une amélioration physique et mentale. Un autre groupe en bas à droite (4 ,37 ,38) ne fait pas attention à la nutrition mais qui voit quand même une amélioration.

Deuxième ACP

1-Traitement des données

Nous allons procéder comme la première ACP mais cette fois-ci avec le deuxième bloc et ensuite changer les noms des colonnes.

```
variables_quantitatives2 <- Formulaire_Musculation_reponses[, (27:37)]

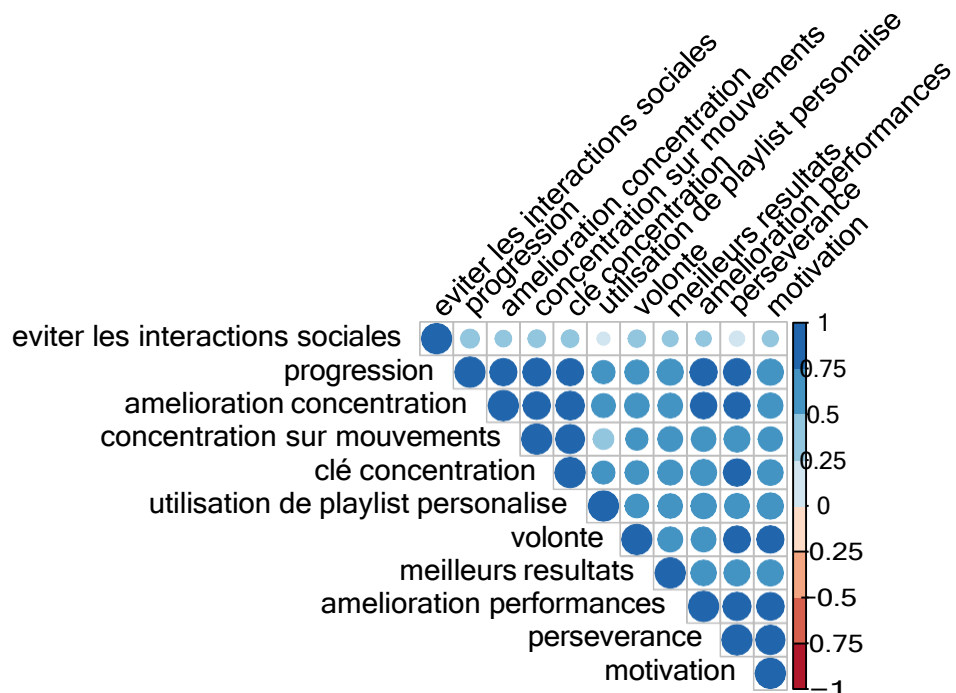
colonnes_a_modifier <- colnames(Formulaire_Musculation_reponses)[27:37]
nouveau_noms_colonnes2 <- c("amelioration concentration", "concentration sur mouvements",
"clé concentration", "amelioration performances",
"eviter les interactions sociales", "progression",
"meilleurs resultats", "perseverance", "motivation", "volonte",
"utilisation de playlist personalise")

colnames(Formulaire_Musculation_reponses)[27:37] <- nouveau_noms_colonnes2
```

2. Pertinence de l'ACP

La matrice de corrélation:

```
colonnes_selectionnees2 <- Formulaire_Musculation_reponses[, 27:37]
matrice_cor2 <- cor(colonnes_selectionnees2)
```

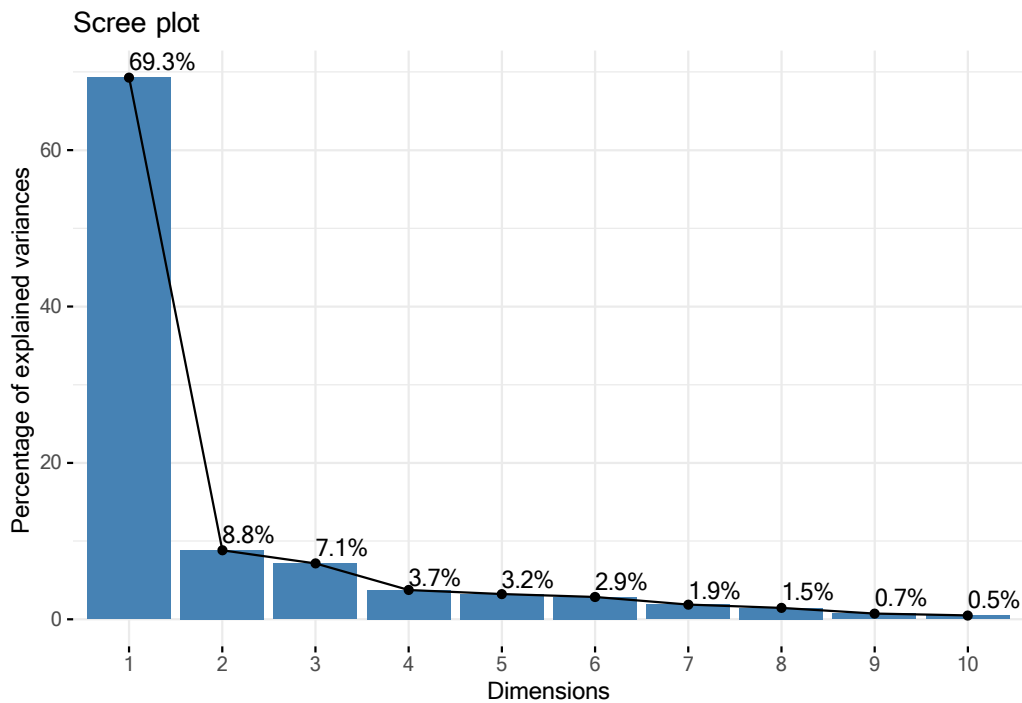


Il y a une forte corrélation entre certaines variables “amélioration de la concentration”, “concentration sur mouvements ”et “rôle essentiel pour la concentration”. De même, une corrélation similaire est observée “amélioration de la réalisation des exos”, “atteinte de meilleurs résultats” et “influence positive sur performances». De plus, ”persévérance”, “motivation” et “volonté” semblent être fortement corrélées.

Exécution de la fonction PCA.

```
pca2 <- PCA(colonnes_selectionnees2, graph = FALSE)
```

3-Choix du nombre d'axes à retenir



- c. Critère du coude : On remarque que le coude se trouve au niveau du deuxième axe, de plus les deux premiers axes capturent une proportion considérable de l'inertie totale donc nous devons retenir les 2 premiers axes.

```
pca2$var$coord[, 1:2]
```

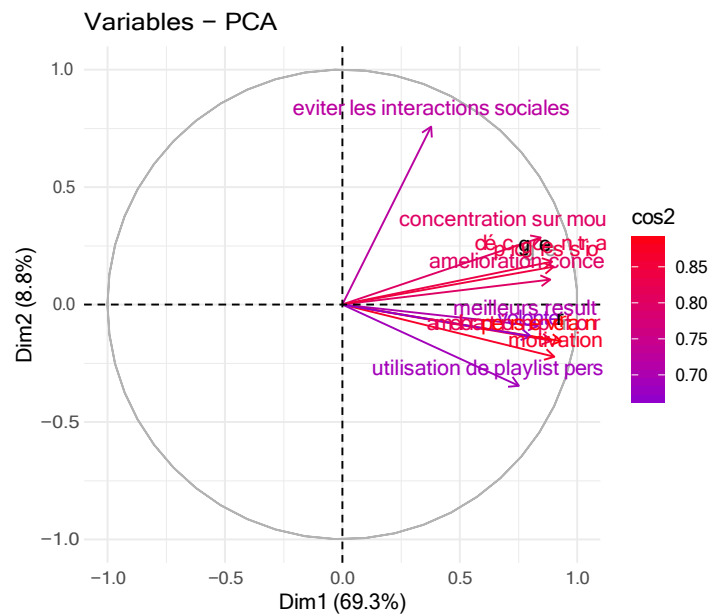
##	Dim.1	Dim.2
## amelioration concentration	0.8844576	0.10596895
## concentration sur mouvements	0.8441227	0.28517964
## clé concentration	0.8925741	0.18335577
## amelioration performances	0.8907963	-0.15403491
## eviter les interactions sociales	0.3777398	0.75795812
## progression	0.9001720	0.16122459
## meilleurs resultats	0.8335404	-0.09237356
## perseverance	0.9310723	-0.15636871
## motivation	0.9023034	-0.22223262
## volonte	0.8022901	-0.13589583
## utilisation de playlist personalise	0.7508572	-0.34653804

Nous pouvons nommer la dimension 1 “Améliorations et concentration” puisque “la persévérance», «motivation” et “Progression” sont très élevées et la dimension 2 “Eviter les interactions sociales” car éviter les interactions sociales est de 0,75 qui est la plus forte.

Cela va être confirmé avec le cercle ci-dessous, nous aurons presque tous les vecteurs proches l’un de l’autre donc ils seront fortement corrélés entre eux

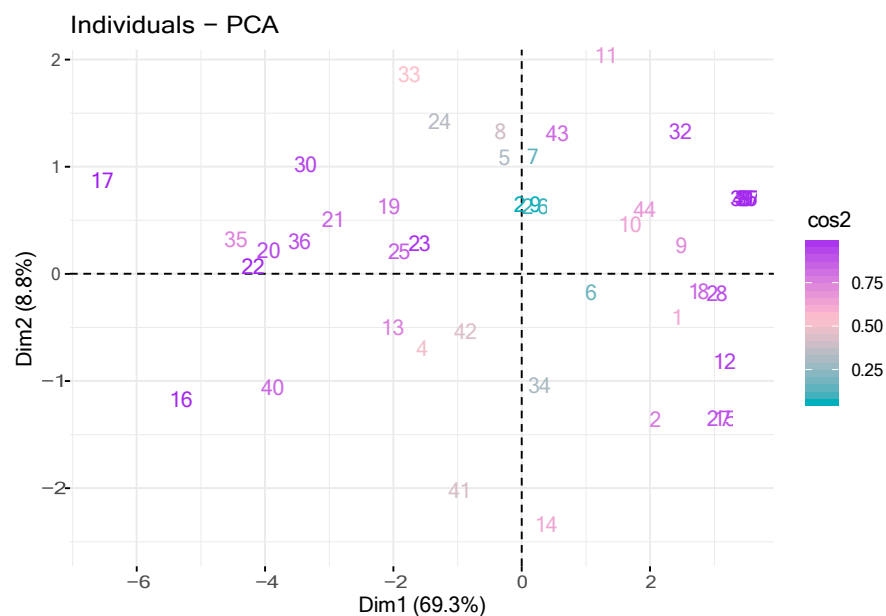
4-Interprétation du cercle

```
fviz_pca_var(pca2, col.var = "cos2",cex=2 ) + scale_color_gradient2(low = "white", mid = "blue", high =
```



5-Interprétation de la carte des individus

```
fviz_pca_ind(pca2,geom = "text",col.ind="cos2",gradient.cols=c("#00AFBB","pink","purple"))
```



D'après le graphique ci-dessus nous pouvons observer que nos données sont éparpillées. Nous avons un groupe en haut à gauche (19,20,21,22,23,25,35,36) qui écoute de la musique pour se concentrer davantage et améliorer ses performances mais qui ne se concentre pas sur les interactions sociales. Un deuxième groupe (5,7,8,43,29,26), les positions des individus sont proches de l'axe de la Dim1 indiquent leurs neutralités, de plus ces individus ont un faible cos2. Nous remarquons également un groupe (1,2,12,18,25,27,28) qui est opposé au premier ses individus écoutent la musique pour éviter les interactions sociales. Enfin, nous pouvons observer un petit groupe (9,10,44,des individus qui se superposent) qui montre un certain équilibre avec les deux dimensions.

ACM

1- Traitement des données

Nous avons deux parties de questions Oui/Non mais tout d'abord nous allons changer les noms des colonnes pour que ce soit plus clair.

```
colonnes_partie1 <- colnames(Formulaire_Musculation_reponses_)[7:10]
nouveaux_noms_colonnes_partie1 <- c("programme specifique", "element essentiel dans routine",
                                     "supp nutritionnel", "competition muscu")
colnames(Formulaire_Musculation_reponses_)[7:10] <- nouveaux_noms_colonnes_partie1
col_partie2 <- colnames(Formulaire_Musculation_reponses_[c((24:26),(38:39))])
nouveaux_noms_colonnes_partie2 <- c("Adaptation musique entraînement",
                                     "Rythme musical entraînement",
                                     "Entraînement sans musique", "Influence musique sur humeur", "Expérience améliorée par musique")
colnames(Formulaire_Musculation_reponses_[c((24:26),(38:39))]) <- nouveaux_noms_colonnes_partie2

tab1 <- apply(Formulaire_Musculation_reponses_[7:10], 2, table)
tab2 <- apply(Formulaire_Musculation_reponses_[c((24:26),(38:39))], 2, table)
tab_mat2 <- matrix(c(5, 40, 18, 27, 32, 13, 6, 39, 2, 43), nrow = 2, byrow = TRUE)
```

Voilà la sortie de la deuxième table :

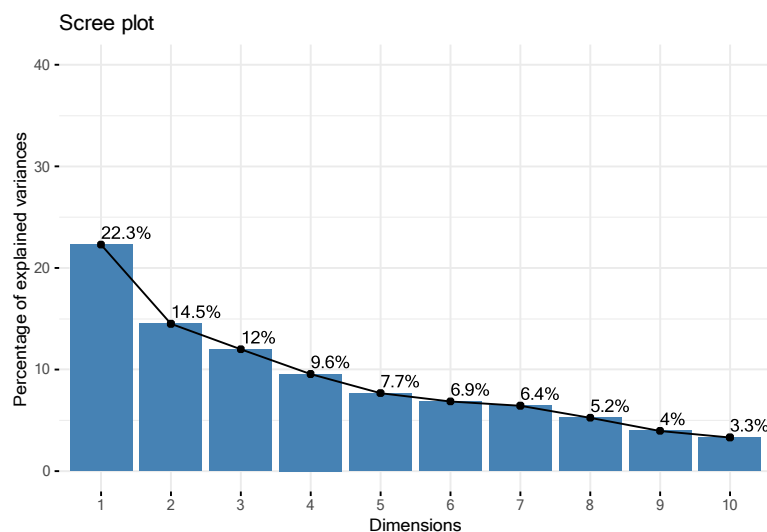
##	Non	Oui
## programme specifique	20	25
## element essentiel dans routine	17	28
## supp nutritionnel	31	14
## competition muscu	43	2

2- Exécution de l'ACM

```
res.MCA <- MCA(Formulaire_Musculation_reponses_[c((1:5), (7:10), (38:39))], graph=FALSE)
```

Un aperçu des axes à retenir :

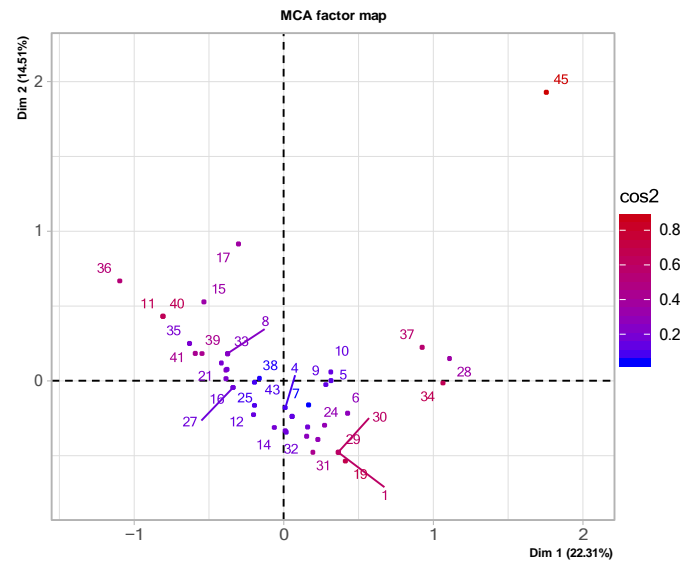
```
fviz_eig(res.MCA, addlabels = TRUE, ylim = c(0, 40))
```



D'après le graphique ci-dessus et le taux de l'inertie, Nous choisissons les deux premiers axes ou les 3 premiers.

3-Visualisation et Interprétation des graphes

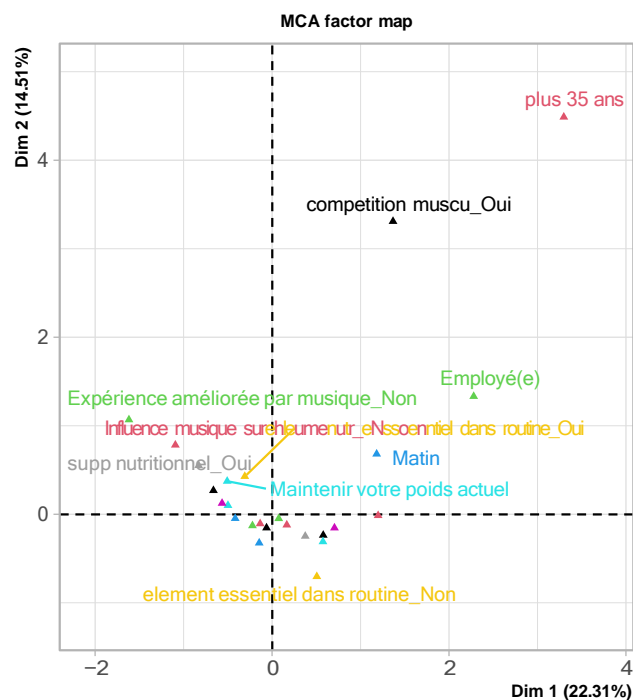
```
plot.MCA(res.MCA,invisible= 'var',habillage='cos2',cex=0.7,cex.main=0.7,cex.axis=0.7,label =c('ind'))
```



Nous remarquons que la majorité des individus se concentrent au centre du graph avec un faible cos2. Ils ne sont pas bien représentés par les axes. Les individus qui sont plus loin du centre sont divisés en deux groupes et ont un cos2 plus fort. Ces groupes peuvent être caractérisés par des profils similaires ou des comportements communs par rapport aux variables.

L'individu 45 peut être considéré comme un cas particulier, ayant des caractéristiques très différentes des autres individus.

```
plot.MCA(res.MCA,invisible= 'ind',col.var=c(1,1,2,2,2,3,3,4,4,4,5,5,5,6,6,7,7,8,8,9,9,10,10,11,11),cex=
```



Nous pouvons remarquer que les variables ne sont pas fortement associées aux dimensions principales mais qu'elles ont une corrélation forte entre eux et peuvent apporter des informations similaires. Nous pouvons constater que l'individu 45 du 1er graph soit associé à l'âge de plus de 35 ans.

```
plot.MCA(res.MCA, choix='var',invisible='quanti.sup',col.var='#1A0066',col.quali.sup='black',cex=0.75,c
```



Nous pouvons voir que la majorité des points sont proches les uns des autres donc les variables sont associées entre elles et sont fréquemment observées ensemble. La variable “catégorie socio-professionnelle” a une corrélation avec la Dimension 1. Nous remarquons que les variables “Compétition muscu” et “âge” sont éloignées par rapport aux autres et les deux variables sont corrélées avec la Dimension 2 (0,5).