



République Tunisienne
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université de Carthage - Ecole Supérieure de la Statistique et de l'Analyse de
l'Information



Projet d'analyse de données



Réalisé par
IKBEL GHRAB

La consommation de café en Tunisie

Encadrant
M.BELMUFTI Ghazi

La consommation de café en Tunisie

Introduction

La Tunisie a connu ces derniers mois une pénurie de plusieurs produits alimentaires importants, dont le café, ce qui a suscité des préoccupations quant à l'approvisionnement et la disponibilité des produits de base dans le pays. Cette enquête sur la consommation de café en Tunisie aidera donc à comprendre l'impact de la pénurie de café sur les habitudes de consommation des Tunisiens, ainsi que sur les éventuelles modifications de ces habitudes. Les résultats de cette enquête seront donc utiles pour mieux comprendre les tendances de consommation de café en Tunisie, ainsi que pour identifier les opportunités émergentes pour les producteurs et les distributeurs de café.

Statistique descriptive

Importation de la base

```
library(readxl)
Formulaire<- read_excel("C:/Users/USER/Desktop/projetFormulaire2023/Consommation-du-café-réponses.xlsx")
```

Présentation de la population ayant répondu à l'enquête:

Répartition par genre:

Camembert de sexe :

```
## Présentation de la population ayant répondu à l'enquête:
```

```
### Répartition par genre:
```

```
#### Camembert de sexe :
```

```
library(dplyr)
```

```
## Warning: le package 'dplyr' a été compilé avec la version R 4.2.3
```

```
## Attachement du package : 'dplyr'
```

```
## Les objets suivants sont masqués depuis 'package:stats':
```

```
##      filter, lag
```

```
## Les objets suivants sont masqués depuis 'package:base':
```

```
##      intersect, setdiff, setequal, union
```

```
library(lessR)
```

```
## lessR 4.2.8                                     feedback: gerbing@pdx.edu
```

```
## -----
```

```
## > d <- Read("")  Read text, Excel, SPSS, SAS, or R data file
```

```
##   d is default data frame, data= in analysis routines optional
```

```
## Learn about reading, writing, and manipulating data, graphics,
```

```
## testing means and proportions, regression, factor analysis,
```

```
## customization, and descriptive statistics from pivot tables.
```

```
##   Enter:  browseVignettes("lessR")
```

```
## View changes in this and recent versions of lessR.
##   Enter: news(package="lessR")
## Interactive data analysis.
##   Enter: interact()
## Attachement du package : 'lessR'

## Les objets suivants sont masqués depuis 'package:dplyr':
##   recode, rename

library(FactoMineR)
library(factoextra)

## Le chargement a nécessité le package : ggplot2

## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa

Homme=sum(Formulaire$sexe== "Homme")
Femme=sum(Formulaire$sexe == "Femme")

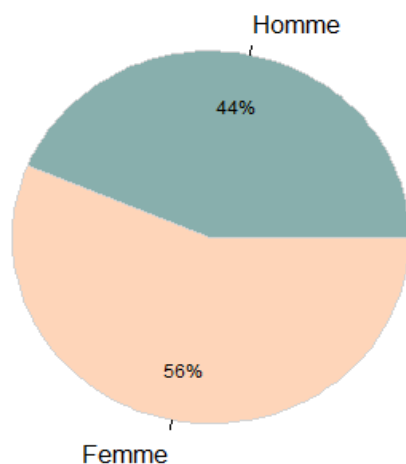
n=c(Homme,Femme)
Sexe=as.factor(c("Homme", "Femme"))

gender=as.data.frame(cbind(Sexe,n))
PieChart(x=Sexe,y=n,gender,fill =c("#89b0ae", "#ffd6ba"),
         values_color="Black",hole=0,main=" Répartition par genre")

## >>> Note: Sexe is not in a data frame (table)
## >>> Note: Sexe is not in a data frame (table)
## >>> Note: n is not in a data frame (table)

## >>> suggestions
## PieChart(Sexe, hole=0) # traditional pie chart
## PieChart(Sexe, values="%") # display %'s on the chart
## PieChart(Sexe) # bar chart
## Plot(Sexe) # bubble plot
## Plot(Sexe, values="count") # lollipop plot
## --- Sexe ---
##           Homme  Femme    Total
## Frequencies:   45    58     103
## Proportions:  0.437 0.563    1.000
## Chi-squared test of null hypothesis of equal probabilities
##   Chisq = 1.641, df = 1, p-value = 0.200
```

Répartition par genre



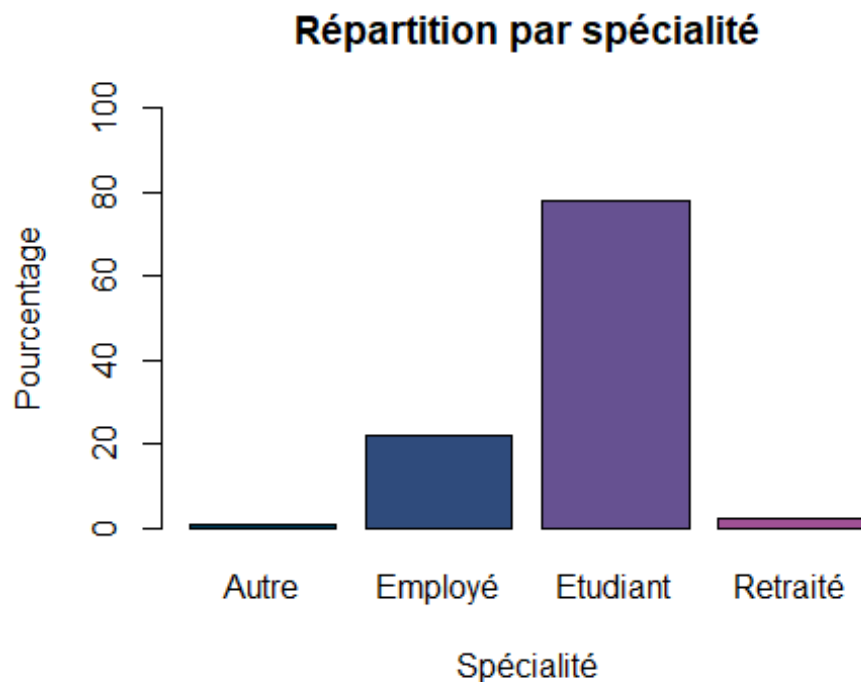
Interpretation :

Selon les résultats de l'enquête, sur un total de 103 personnes interrogées, 56% sont des femmes. Cela peut indiquer une légère majorité de femmes parmi la population interrogée. Mais nous ne pouvons tirer aucune interprétation définitive à partir de cette donnée seule car il est également important de considérer d'autres facteurs tels que l'âge les intérêts... qui pourraient influencer la consommation de café des tunisiens..

Répartition par professions:

Diagramme en battons des professions:

```
colors <- c("#003f5c", "#2f4b7c", "#665191", "#a05195")
barplot(table(Formulaire$profession), col=colors, ylim=c(0, 100),
        main="Répartition par spécialité", xlab="Spécialité", ylab="Pourcentage")
```



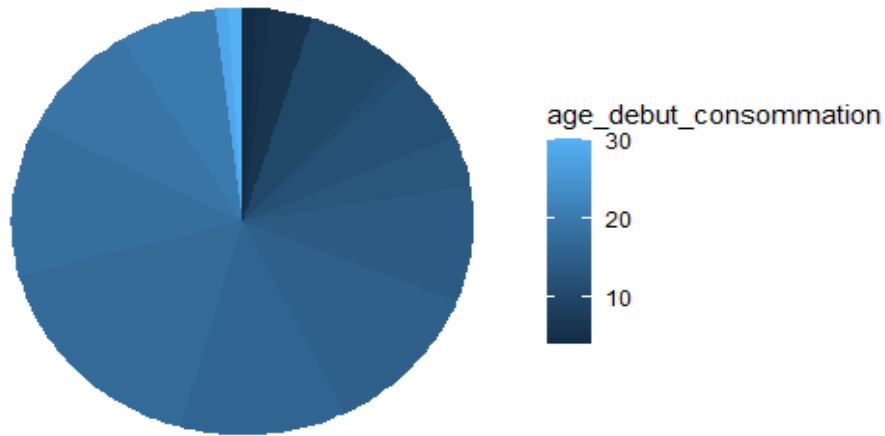
Conclusion: La plupart des personnes interrogées étaient des étudiants (environ 76% des personnes interrogées), suivis des employés (environ 21% des personnes interrogées), tandis que les retraités et les personnes dans la catégorie "Autre" étaient des minorités (environ 2% et moins de 1% des personnes interrogées respectivement).

Age de début de consommation du café:

```
age_debut <- count(Formulaire, age_debut_consommation)

ggplot(age_debut, aes(x="", y=n, fill=age_debut_consommation)) +
  geom_bar(stat="identity", width=1) +
  coord_polar("y", start=0)+
  theme_void() +
  ggtitle("Répartition selon l'âge de début de consommation du café")
```

Répartition selon l'âge de début de consommation du café



Conclusion

Ces résultats indiquent que la consommation de café est plus fréquente chez les jeunes adultes et les adultes. Il est intéressant de noter que la plupart des personnes ont commencé à consommer du café lorsqu'elles étaient mineures, ce qui peut avoir des implications pour la réglementation de la consommation de café chez les jeunes.

Analyse en composantes principales :

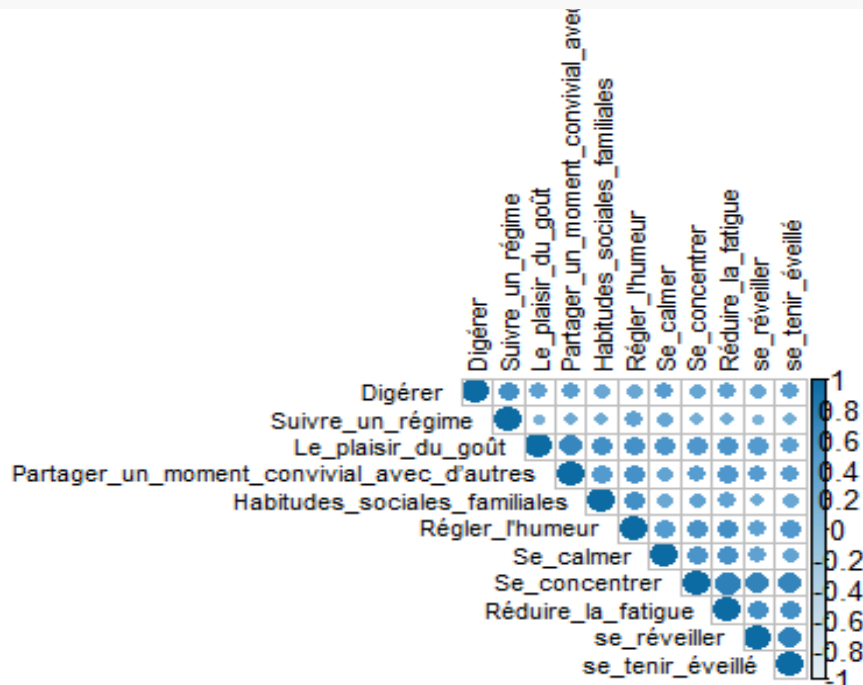
1. Matrice de corrélation:

```
library(corrplot)

## corrplot 0.92 loaded

X=as.matrix(Formulaire[,c(17:27)])

M <- cor(X)
library(RColorBrewer)
col <- colorRampPalette(c("#E6F0F3", "#BDD7E7", "#8CBEDB", "#539ACF", "#0D6BA4"))(100)
corrplot(M, type = "upper", order = "hclust", col = col, tl.col = "black", tl.cex = 0.7)
```



Interpretation:

La matrice de corrélation suggère que “se réveiller”, “se tenir éveillé”, “se concentrer” et “le plaisir du goût” sont des variables étroitement liées, tandis que “habitudes sociales familiales” et “suivre un régime” sont moins liées aux autres variables.

Choix de nombre d'axes:

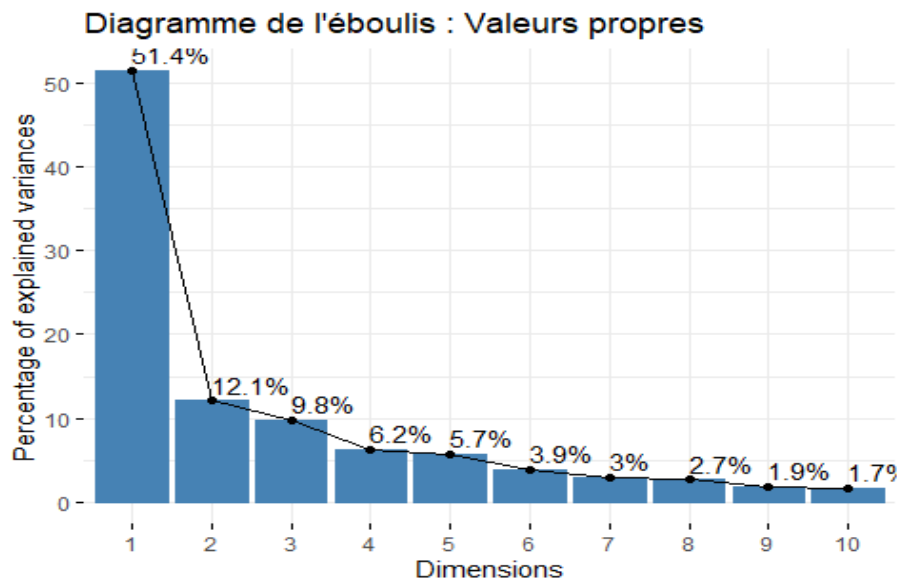
```
library(FactoMineR)
library(factoextra)
res.pca=PCA(X,ncp = 4,graph= F)
head(res.pca$eig)
```

##		eigenvalue	percentage of variance	cumulative percentage of variance
##	comp 1	5.6520589	51.382353	51.38235
##	comp 2	1.3288016	12.080015	63.46237
##	comp 3	1.0800724	9.818840	73.28121
##	comp 4	0.6852311	6.229373	79.51058
##	comp 5	0.6268186	5.698351	85.20893
##	comp 6	0.4301370	3.910337	89.11927

Interprétation: Les trois premières composantes principales expliquent environ 73% de la variance totale des données, tandis que les quatre premières composantes expliquent environ 80% de la variance totale.

```
library(ggplot2)
library(factoextra)
```

```
fviz_eig(res.pca, addlabels = TRUE) +
  ggtitle("Diagramme de l'éboulis : Valeurs propres")
```



Interpretation :

1-Critère de Kaiser : On retiendrait les composantes principales dont l'eigenvalue est supérieure à 1, ce qui signifie que la première, la deuxième et la troisième composante principale seraient retenues.

2-Critère du taux d'inertie cumulée : Selon le critère du taux d'inertie cumulée, on pourrait retenir les trois premières composantes principales qui expliquent 73.28% de l'inertie totale.

3-Critère du coude : On peut voir que le coude se situe après la deuxième composante principale, ce qui indique que l'on peut se contenter de retenir les deux premières composantes principales.

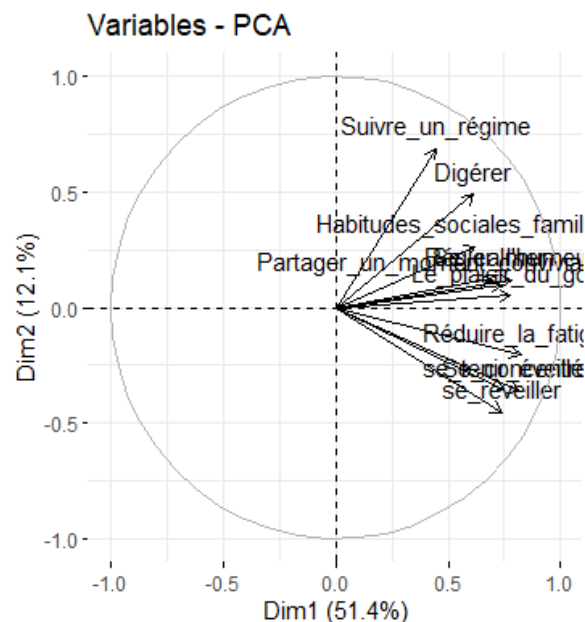
=> Nous pourrions retenir les deux premières composantes principales qui expliquent environ 63.46% de l'inertie totale. .

Interpretation de la carte des variables :

```
res.pca$var$coord
```

	Dim.1	Dim.2	Dim.3
## se_réveiller	0.7364349	-0.45318155	0.20941634
## se_tenir_éveillé	0.7470688	-0.35783509	0.26618498
## Se_concentrer	0.8179580	-0.36026915	0.13994091
## Se_calmer	0.6907580	0.12413276	0.15769424
## Le_plaisir_du_goût	0.7784446	0.05427431	-0.38988520
## Réduire_la_fatigue	0.8227466	-0.20609433	0.08414905
## Digérer	0.6142355	0.48771188	0.39988861
## Partager_un_moment_convivial_avec_d'autres	0.7429677	0.10011089	-0.28987011
## Habitudes_sociales_familiales	0.6211144	0.26441321	-0.54085513
## Suivre_un_régime	0.4431741	0.68725900	0.39140170
## Régler_l'humeur	0.7816958	0.12097289	-0.26863580
##	Dim.4		
## se_réveiller	0.18663812		
## se_tenir_éveillé	0.30068953		
## Se_concentrer	-0.04473490		
## Se_calmer	-0.63808683		
## Le_plaisir_du_goût	-0.14641283		
## Réduire_la_fatigue	-0.17686657		
## Digérer	0.07612212		
## Partager_un_moment_convivial_avec_d'autres	0.15599562		
## Habitudes_sociales_familiales	0.19115957		
## Suivre_un_régime	0.17622570		
## Régler_l'humeur	0.01954168		

```
fviz_pca_var(res.pca, shadow=TRUE)
```



Interpretation:

Sur la première composante principale, on peut observer une forte contribution des variables "Se_concentrer", "Réduire_la_fatigue" et "Régler_l'humeur". Ces variables sont donc fortement corrélées à la première composante principale.

Sur la deuxième composante principale, les variables "Suivre_un_régime", "Digérer" et "Habitudes_sociales_familiales" sont les plus corrélées.

Enfin, sur la troisième composante principale, la variable "Le_plaisir_du_goût" est la plus corrélée.

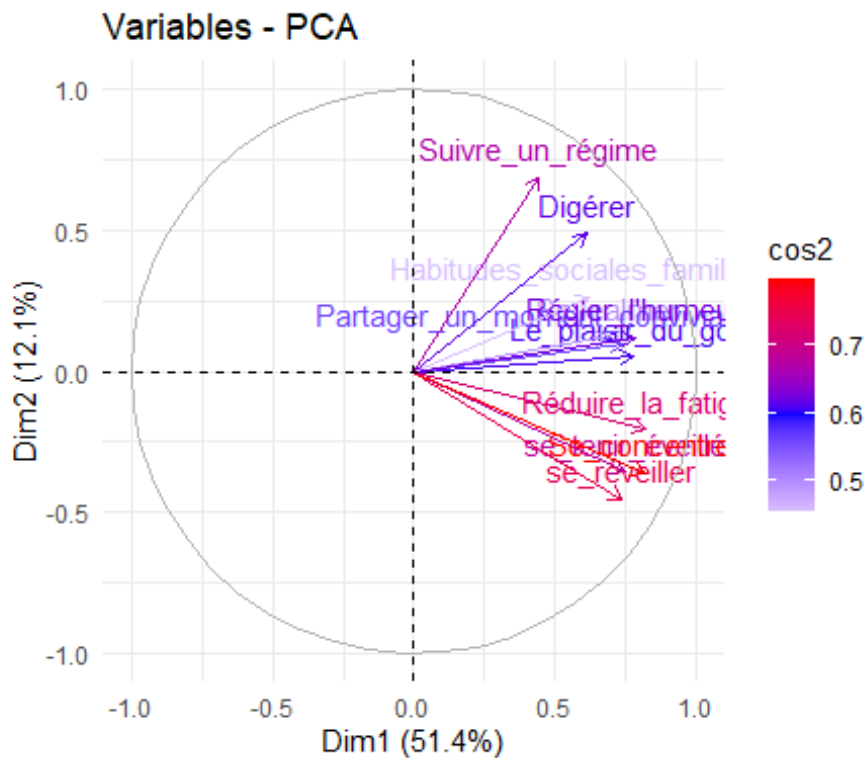
Ces observations indiquent que les variables qui contribuent le plus à la première composante principale sont celles qui sont liées à l'énergie mentale, à la relaxation et à l'humeur. Les variables qui contribuent le plus à la deuxième composante principale sont celles qui sont liées aux habitudes alimentaires et sociales. Enfin, la troisième composante principale est liée aux plaisirs gustatifs.

Tableau des cosinus carrés des variables sur les axes de l'ACP. :

```
res.pca$var$cos2

##                Dim.1        Dim.2        Dim.3
## se_réveiller      0.5423364  0.205373516  0.043855205
## se_tenir_éveillé  0.5581117  0.128045951  0.070854445
## Se_concentrer     0.6690552  0.129793863  0.019583458
## Se_calmer         0.4771466  0.015408943  0.024867473
## Le_plaisir_du_goût 0.6059760  0.002945701  0.152010468
## Réduire_la_fatigue 0.6769120  0.042474872  0.007081063
## Digérer           0.3772852  0.237862876  0.159910898
## Partager_un_moment_convivial_avec_d'autres 0.5520010  0.010022190  0.084024683
## Habitudes_sociales_familiales 0.3857831  0.069914347  0.292524271
## Suivre_un_régime  0.1964033  0.472324938  0.153195288
## Régler_l'humeur   0.6110483  0.014634441  0.072165193
##                Dim.4
## se_réveiller      0.0348337884
## se_tenir_éveillé  0.0904141948
## Se_concentrer     0.0020012111
## Se_calmer         0.4071547984
## Le_plaisir_du_goût 0.0214367159
## Réduire_la_fatigue 0.0312817841
## Digérer           0.0057945774
## Partager_un_moment_convivial_avec_d'autres 0.0243346346
## Habitudes_sociales_familiales 0.0365419821
## Suivre_un_régime  0.0310554979
## Régler_l'humeur   0.0003818773

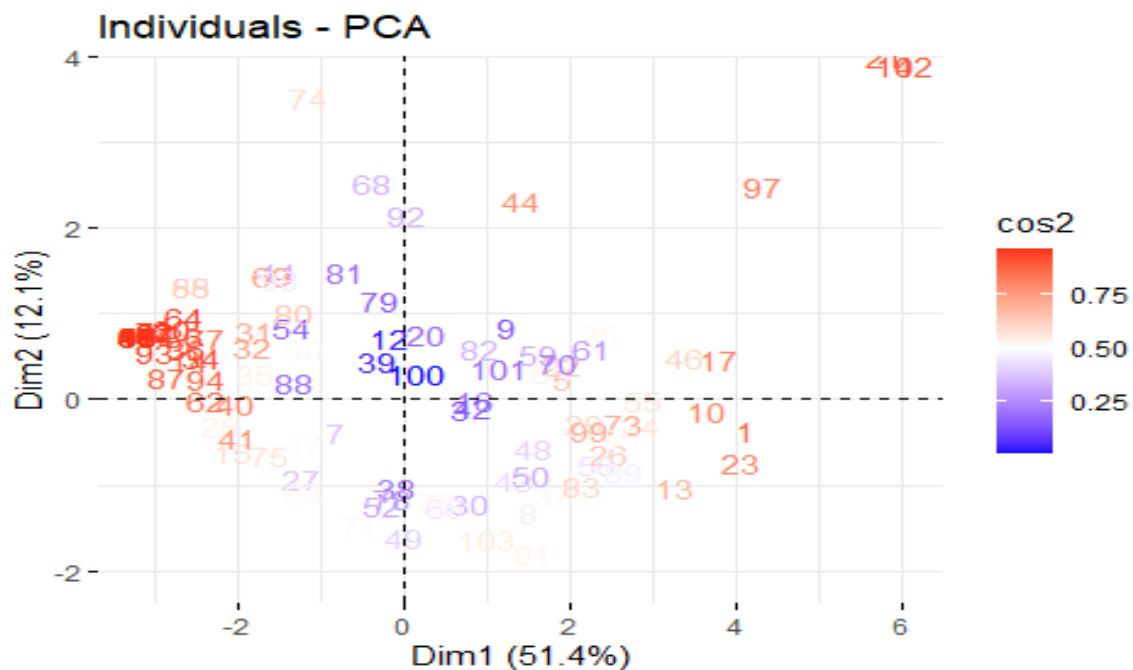
fviz_pca_var(res.pca, col.var = "cos2")+
  scale_color_gradient2(low="white", mid="blue",
                        high="red", midpoint = 0.6)+
  theme_minimal()
```

Interpretation : On peut voir que les variables “Se_concentrer”, “Réduire_la_fatigue” et “Digérer” ont des cosinus élevés pour la première dimension de l’ACP, ce qui indique qu’elles contribuent fortement à la formation de cette dimension. De même, les variables “Le_plaisir_du_goût” et “Habitudes_sociales_familiales” ont des cosinus élevés pour la deuxième dimension de l’ACP. Enfin, la variable “Suivre_un_régime” a un cosinus élevé pour la troisième dimension de l’ACP.

La carte des individus :

```
fviz_pca_ind(res.pca,geom = "text",col.ind="cos2")+
scale_color_gradient2(low="blue", mid="white",
high="red", midpoint=0.5)
```



ACM :

```
library(FactoMineR)
library(factoextra)
l.active<-Formulaire[,42:49]
res.mca <- MCA (l.active, graph = FALSE)
```

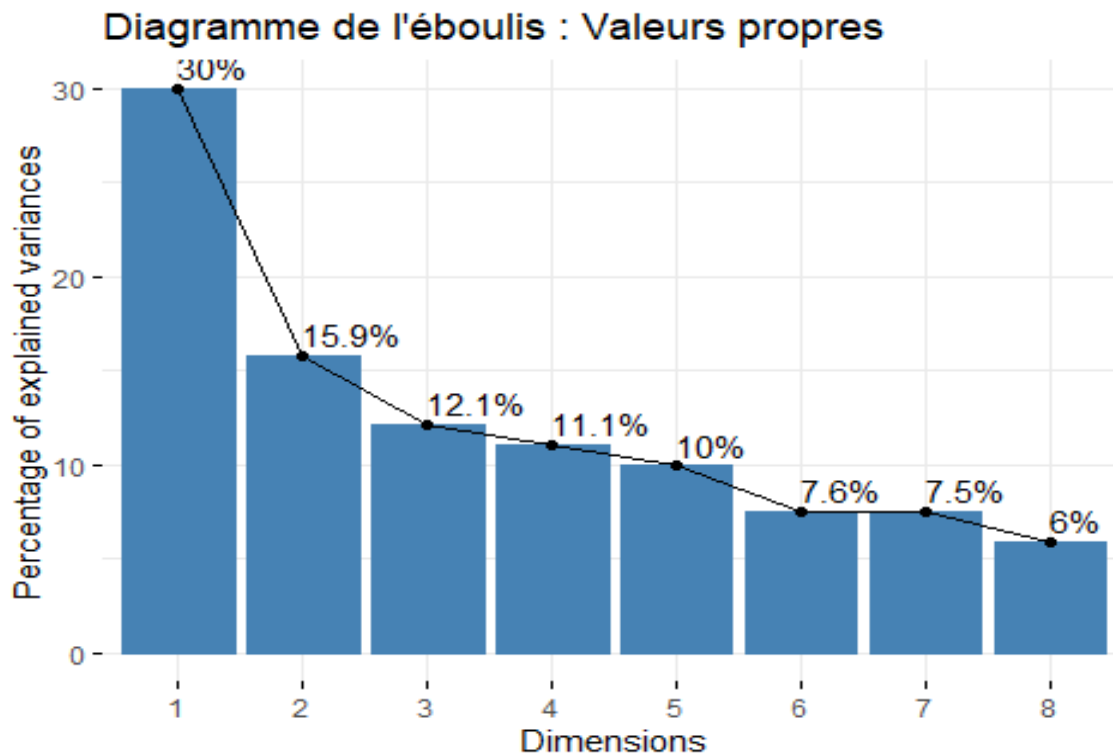
1.Visualisation et interpretation :

```
res.mca$eig

##      eigenvalue percentage of variance cumulative percentage of variance
## dim 1 1.29981271          29.981271          29.98127
## dim 2 0.65850093          15.850093          45.83136
## dim 3 0.12132032          12.132032          57.96340
## dim 4 0.11065621          11.065621          69.02902
## dim 5 0.09963522           9.963522          78.99254
## dim 6 0.07558721           7.558721          86.55126
## dim 7 0.07494959           7.494959          94.04622
## dim 8 0.05953781           5.953781         100.00000

library(ggplot2)
library(factoextra)

fviz_eig(res.mca, addlabels = TRUE) +
  ggtitle("Diagramme de l'éboulis : Valeurs propres")
```



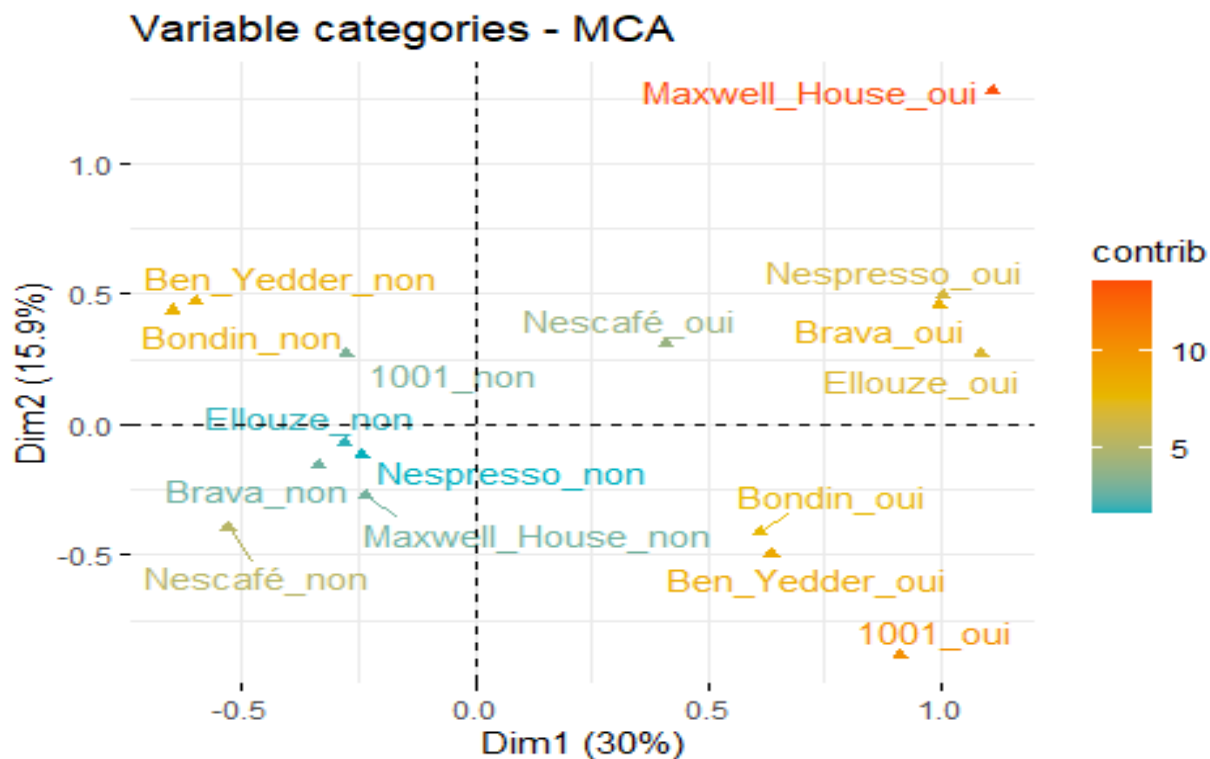
Interprétation: L'ACM réalisée sur les données l.active a permis d'obtenir les résultats suivants :

- seule la première dimension répond à ce critère, ce qui signifie qu'elle est la plus importante pour expliquer les données.

- Le critère du taux d'inertie cumulée les deux premières dimensions cumulent une proportion de variance d'environ 45.8%, ce qui suggère que ces deux dimensions sont importantes pour expliquer les données.
- Le critère du coude : On peut voir que le coude se situe après la deuxième composante principale, ce qui indique que l'on peut se contenter de retenir les deux premières composantes principales 3. Interprétation de la première carte des modalités.

#Représentation de la première carte des modalités

```
fviz_mca_var(res.mca,
  col.var = "contrib", # Couleur en fonction de la contribution
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"), # Choix de couleurs
  repel = TRUE, # Évite le chevauchement des labels
  ggtheme = theme_minimal() # Style du graphique
)
```



```
dimdesc(res.mca, axes=1:2, proba=0.05)$`Dim 1`
## Link between the variable and the categorical variable (1-way anova)
##
##      R2      p.value
## Bondin    0.3985819 0.0000000000008723857
## Ben_Yedder 0.3823707 0.0000000000034092025
## Brava      0.3362696 0.0000000001373972383
## Ellouze    0.3031851 0.0000000016829478644
## Maxwell_House 0.2624013 0.0000000318412362472
## 1001       0.2534035 0.0000000597014130797
## Nespresso  0.2437924 0.0000001159669848023
## Nescafé    0.2184972 0.0000006426891668188
## Link between variable and the categories of the categorical variables
##
##      Estimate      p.value
## Bondin=Bondin_oui 0.3458343 0.0000000000008723857
```

```
## Ben_Yedder=Ben_Yedder_oui      0.3387284 0.0000000000034092025
## Brava=Brava_oui                 0.3654635 0.00000000001373972383
## Ellouze=Ellouze_oui             0.3741705 0.00000000016829478644
## Maxwell_House=Maxwell_House_oui 0.3692917 0.00000000318412362472
## 1001=1001_oui                   0.3260013 0.0000000597014130797
## Nespresso=Nespresso_oui         0.3417341 0.0000001159669848023
## Nescafé=Nescafé_oui             0.2580090 0.0000006426891668188
## Nescafé=Nescafé_non             -0.2580090 0.0000006426891668188
## Nespresso=Nespresso_non         -0.3417341 0.0000001159669848023
## 1001=1001_non                   -0.3260013 0.0000000597014130797
## Maxwell_House=Maxwell_House_non -0.3692917 0.00000000318412362472
## Ellouze=Ellouze_non             -0.3741705 0.00000000016829478644
## Brava=Brava_non                 -0.3654635 0.00000000001373972383
## Ben_Yedder=Ben_Yedder_non       -0.3387284 0.0000000000034092025
## Bondin=Bondin_non               -0.3458343 0.000000000008723857
```

Interpretation: =>l'axe1 dans cette analyse ACM semble être un axe de consommation de café.

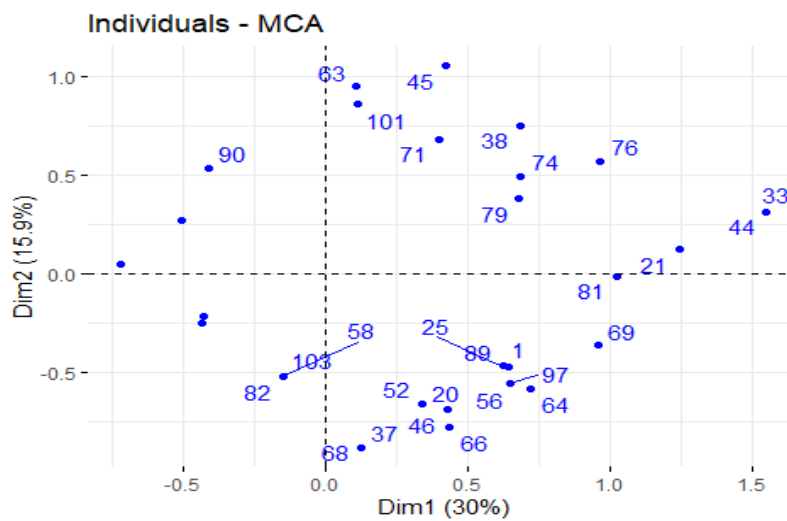
```
dimdesc(res.mca, axes=1:2, proba=0.05)$`Dim 2`
## Link between the variable and the categorical variable (1-way anova)
##                               R2                p.value
## Maxwell_House 0.34581316 0.00000000006524969
## 1001           0.23806060 0.00000017169049789
## Ben_Yedder    0.23427024 0.00000022223552217
## Bondin        0.17995101 0.00000799024039682
## Nescafé       0.12266214 0.00028711768603295
## Brava         0.06999252 0.00692293226043614
## Nespresso     0.05897042 0.01345187486632598
##
## Link between variable and the categories of the categorical variables
##                               Estimate                p.value
## Maxwell_House=Maxwell_House_oui 0.3082465 0.00000000006524969
## 1001=1001_non                    0.2297459 0.00000017169049789
## Ben_Yedder=Ben_Yedder_non        0.1927786 0.00000022223552217
## Bondin=Bondin_non                0.1689574 0.00000799024039682
## Nescafé=Nescafé_oui              0.1405588 0.00028711768603296
## Brava=Brava_oui                  0.1212319 0.00692293226043613
## Nespresso=Nespresso_oui          0.1222042 0.01345187486632595
## Nespresso=Nespresso_non          -0.1222042 0.01345187486632595
## Brava=Brava_non                  -0.1212319 0.00692293226043610
## Nescafé=Nescafé_non              -0.1405588 0.00028711768603295
## Bondin=Bondin_oui                -0.1689574 0.00000799024039682
## Ben_Yedder=Ben_Yedder_oui        -0.1927786 0.00000022223552217
## 1001=1001_oui                   -0.2297459 0.00000017169049789
## Maxwell_House=Maxwell_House_non -0.3082465 0.00000000006524969
```

Interpretation: =>l'axe2 dans cette analyse ACM semble être un axe qui indique la préférence globale pour différentes marques de café:ceux qui aiment le café arabesque et ceux non.

4. Interprétation de la première carte des individus.

#Représentation de la première carte des individus

```
fviz_mca_ind (res.mca,select.ind = list(cos2 = 0.4),
              repel = TRUE,
              ggtheme = theme_minimal ())
```

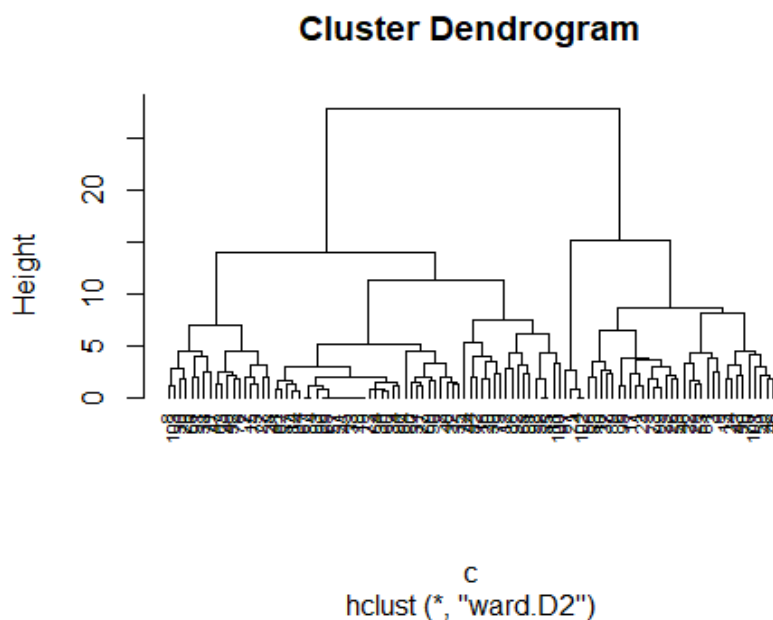


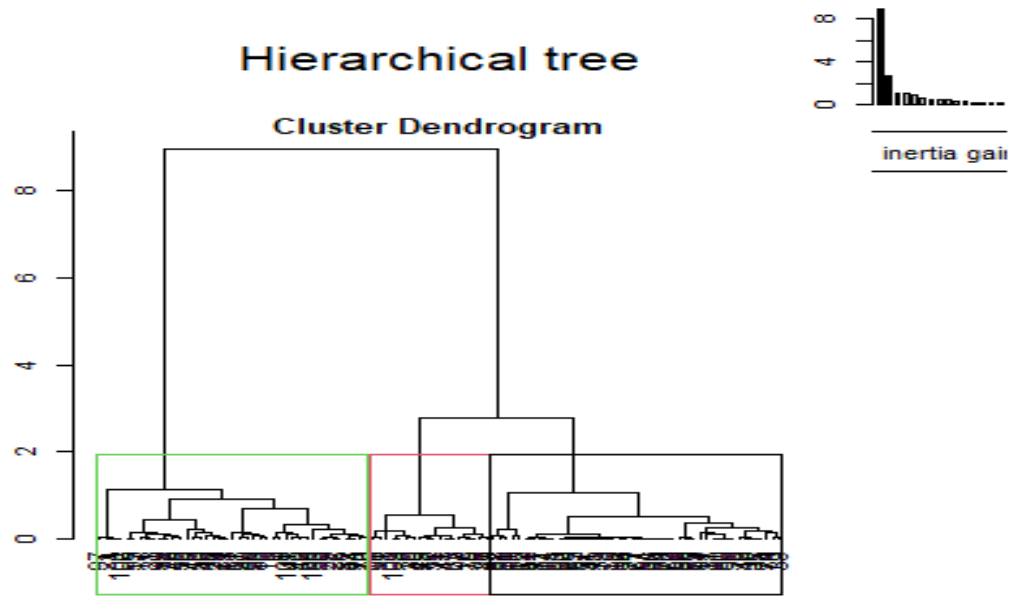
Interpretation: Les individus qui se situent à droite de l'axe 1 sont associées à une consommation de café élevée, tandis que ceux situant à gauche sont associées à une consommation de café faible. Les individus qui se situent en bas de l'axe 2 sont associées à une préférence pour les marques de café arabesque, tandis que celles situées en bas sont associées à une préférence pour les autres marques de café.

Classification :

```
new<-Formulaire[17:27]
c<-dist(scale(new),method="euclidean")
h<- hclust(c, method="ward.D2")
plot(h, hang = -1, cex =0.6)

library(FactoMineR)
res.HCPC<-HCPC(new, consol=TRUE, graph=F)
plot.HCPC(res.HCPC,choice='tree', title ='Hierarchical tree')
```

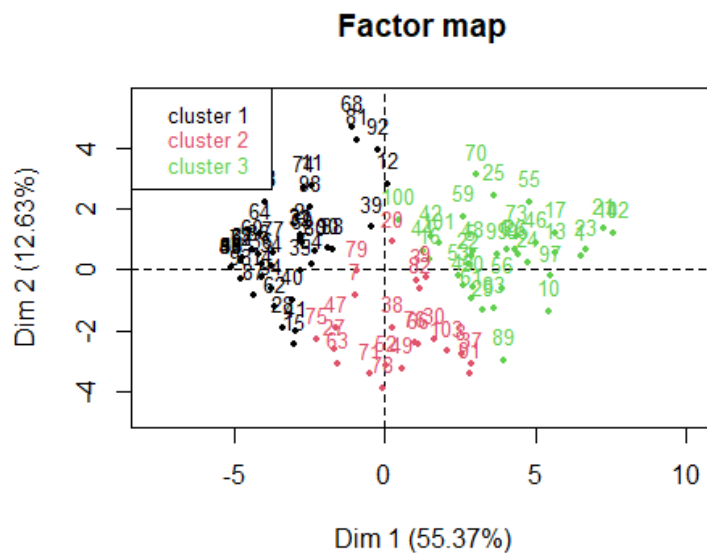




on peut choisir 3 classes.

Arbre hiérarchique :

```
plot.HCPC(res.HCPC, choice = 'map', draw.tree=FALSE, title = 'Factor map')
```



```
plot.HCPC(res.HCPC, choice = '3D.map', ind.names=FALSE, centres.plot=FALSE, angles=60, title='Hierarchical tree on the factor map')
```

A 3D scatter plot showing the results of a Principal Component Analysis (PCA) for three clusters. The axes are labeled: 'height' (vertical axis, ranging from 0 to 10), 'Dim 1 (55.37%)' (horizontal axis, ranging from -6 to 8), and 'Dim 2 (12.63%)' (depth axis, ranging from -4 to 4). A legend in the top left corner identifies the clusters: 'cluster 1' (black dots), 'cluster 2' (red dots), and 'cluster 3' (green dots). The plot shows three distinct groups of points, with cluster 1 being the largest and most spread out, cluster 2 being smaller and more compact, and cluster 3 being the smallest and most compact. A dendrogram is visible in the background, showing the hierarchical clustering of the data points.

```
res.HCPC$desc.var
```

```
##
## Link between the cluster variable and the quantitative variables
## =====
##                               Eta2
## Se_concentrer                0.69701145
## se_réveiller                 0.65101324
## se_tenir_éveillé             0.62591391
## Réduire_la_fatigue           0.57077927
## Le_plaisir_du_goût           0.52767693
## Régler_l'humeur              0.52117592
## Habitudes_sociales_familiales 0.38998622
## Se_calmer                     0.38237513
## Partager_un_moment_convivial_avec_d'autres 0.37856604
## Digérer                      0.19933353
## Suivre_un_régime             0.06051602
##
##                               P-value
## Se_concentrer                0.0000000000000000000000000117845
## se_réveiller                 0.000000000000000000000000138180772
## se_tenir_éveillé             0.0000000000000000000000004452235514
## Réduire_la_fatigue           0.0000000000000000000000004305605695121
## Le_plaisir_du_goût           0.000000000000000000000000515178709345847
## Régler_l'humeur              0.0000000000000000000000001020467371410654
## Habitudes_sociales_familiales 0.000000000000184919244517927782410
## Se_calmer                     0.000000000000343746850432199704730
## Partager_un_moment_convivial_avec_d'autres 0.000000000000467465088838480103848
## Digérer                      0.0000148792838954689459396094064
## Suivre_un_régime             0.0441030725907842480193465917182
##
## Description of each cluster by quantitative variables
## =====
## $`1`
##                               v.test Mean in category
## Habitudes_sociales_familiales -3.503782      1.933333
## Digérer                       -3.848393      1.088889
## Se_calmer                     -4.658711      1.311111
## Partager_un_moment_convivial_avec_d'autres -5.445833      1.933333
## Le_plaisir_du_goût            -5.447499      2.000000
```

```

## Régler_l'humeur -5.505350 1.377778
## Réduire_la_fatigue -7.024306 1.177778
## se_tenir_éveillé -7.973171 1.533333
## se_réveiller -8.147530 1.711111
## Se_concentrer -8.388368 1.200000
## Overall mean sd in category
## Habitudes_sociales_familiales 2.563107 1.388844
## Digérer 1.514563 0.354163
## Se_calmer 2.058252 0.811643
## Partager_un_moment_convivial_avec_d'autres 2.864078 1.236482
## Le_plaisir_du_goût 2.970874 1.316561
## Régler_l'humeur 2.310680 0.824321
## Réduire_la_fatigue 2.378641 0.436597
## se_tenir_éveillé 2.922330 0.858939
## se_réveiller 3.233010 1.067129
## Se_concentrer 2.757282 0.400000
## Overall sd p.value
## Habitudes_sociales_familiales 1.5989675 0.00045870094612697996650
## Digérer 0.9839891 0.00011889522451890459972
## Se_calmer 1.4266890 0.00000318195463410064843
## Partager_un_moment_convivial_avec_d'autres 1.5204020 0.00000005156349813649501
## Le_plaisir_du_goût 1.5854699 0.00000005108313361026193
## Régler_l'humeur 1.5074516 0.00000003684357668924066
## Réduire_la_fatigue 1.5208359 0.00000000000215132096423
## se_tenir_éveillé 1.5497530 0.00000000000000154653625
## se_réveiller 1.6616981 0.00000000000000037143388
## Se_concentrer 1.6515131 0.0000000000000004929382
##
## $`2`
## v.test Mean in category Overall mean
## se_réveiller 3.869216 4.454545 3.233010
## se_tenir_éveillé 3.196961 3.863636 2.922330
## Se_concentrer 3.091396 3.727273 2.757282
## Habitudes_sociales_familiales -3.050624 1.636364 2.563107
## sd in category Overall sd p.value
## se_réveiller 0.7215685 1.661698 0.0001091858
## se_tenir_éveillé 1.1791929 1.549753 0.0013888356
## Se_concentrer 1.2497934 1.651513 0.0019921759
## Habitudes_sociales_familiales 0.7713892 1.598968 0.0022836647
##
## $`3`
## v.test Mean in category
## Le_plaisir_du_goût 7.269721 4.527778
## Régler_l'humeur 7.204930 3.777778
## Réduire_la_fatigue 6.810701 3.777778
## Habitudes_sociales_familiales 6.266886 3.916667
## Se_calmer 6.183992 3.250000
## Se_concentrer 6.068703 4.111111
## Partager_un_moment_convivial_avec_d'autres 5.801516 4.055556
## se_tenir_éveillé 5.546064 4.083333
## se_réveiller 5.149609 4.388889
## Digérer 4.279180 2.083333
## Suivre_un_régime 2.482695 1.527778
## Overall mean sd in category
## Le_plaisir_du_goût 2.970874 0.7988224
## Régler_l'humeur 2.310680 1.2272623
## Réduire_la_fatigue 2.378641 1.1331154
## Habitudes_sociales_familiales 2.563107 1.2990381
## Se_calmer 2.058252 1.5343294
## Se_concentrer 2.757282 1.0999439
## Partager_un_moment_convivial_avec_d'autres 2.864078 1.0526275

```


## se_tenir_éveillé	2.922330	0.8936504
## se_réveiller	3.233010	1.0076865
## Digérer	1.514563	1.3202483
## Suivre_un_régime	1.252427	1.2130336
##	Overall sd	p.value
## Le_plaisir_du_goût	1.5854699	0.0000000000003602296
## Régler_l'humeur	1.5074516	0.0000000000005807384
## Réduire_la_fatigue	1.5208359	0.00000000000097124244
## Habitudes_sociales_familiales	1.5989675	0.000000000003683407175
## Se_calmer	1.4266890	0.000000000006250071625
## Se_concentrer	1.6515131	0.00000000012894713381
## Partager_un_moment_convivial_avec_d'autres	1.5204020	0.00000000065718036914
## se_tenir_éveillé	1.5497530	0.0000000292171891830
## se_réveiller	1.6616981	0.0000002610299256182
## Digérer	0.9839891	0.0000187583504782157
## Suivre_un_régime	0.8210631	0.0130392566089234160

Interpretation:

* Le cluster 1 privilégie les aspects sociaux et conviviaux ainsi que la satisfaction gustative lors de la consommation de café. Ils cherchent également à réguler leur humeur et à réduire leur fatigue.

* Le cluster 2 se caractérise par des individus qui ont tendance à privilégier les aspects pratiques et fonctionnels du café plutôt que les aspects sociaux.

* Le cluster 3 privilégie principalement les aspects liés au plaisir gustatif, à la régulation de l'humeur et à la réduction de la fatigue. Ils attachent également de l'importance à la socialisation et aux moments conviviaux, ainsi qu'à la concentration et à la digestion.

```
res.HCPC$desc.axes
## Link between the cluster variable and the quantitative variables
## =====
##          Eta2          P-value
## Dim.1 0.8131807 3.724691e-37
## Dim.2 0.3954355 1.180683e-11
## Description of each cluster by quantitative variables
## =====
## $`1`
##          v.test Mean in category          Overall mean sd in category
## Dim.2  3.339528          0.6472348 -0.0000000000000007626047          1.499813
## Dim.1 -8.354337          -3.3900271  0.000000000000000233793567          1.405823
##          Overall sd          p.value
## Dim.2  1.724126 0.00083920896785608936012
## Dim.1  3.609803 0.000000000000000006580112
## $`2`
##          v.test Mean in category          Overall mean sd in category
## Dim.9 -2.197267          -0.2938968  0.00000000000000004151547          0.9878720
## Dim.3 -2.341366          -0.5668491 -0.00000000000000017370814          0.9767467
## Dim.2 -6.332234          -2.0742312 -0.00000000000000007626047          1.2758674
##          Overall sd          p.value
## Dim.9  0.7040136 0.0280013956674155
## Dim.3  1.2742864 0.0192133213378341
## Dim.2  1.7241264 0.00000000002416362
## $`3`
##          v.test Mean in category          Overall mean sd in category
## Dim.1 8.163152          3.9804068  0.000000000000000233793567          1.775836
## Dim.2 1.968902          0.4585423 -0.00000000000000007626047          1.155516
##          Overall sd          p.value
## Dim.1  3.609803 0.000000000000000003263937
## Dim.2  1.724126 0.0489642958136675340852
```

```
res.km <- eclust(Formulaire[8:16], "kmeans", nstart = 25,k=3)
```

