

ProjetAD

Lassoued Achraf

05/05/2023

## 1/-Introduction :

### \*Définition du l’enquete :

Cette étude est faite dans le but de dresser l’évolution de regarder les films en Tunisie.

Pour cela un questionnaire contenant 15 questions a été posé et nous avons enregistré les différentes réponses .

Ensuite nous avons procedé aux étapes ci-dessous afin d’analyser les données recueillies.

### \*Importation du base :

library(readxl)  
ques <- read\_excel("Titre de la fiche d'exercices (réponses)2.xlsx")

### \*Résumé du base :

Parmi les questions proposées dans le questionnaire Nous avons demandé aux gens qui ont répondus au questionnaire d’attribuer un cause pour regarder les films: Presse spécialisée, Presse généraliste , Site Internet …. Le type des films: action, fiction animations ….

1 :Toujours. 2 : Souvent. 3 : Parfois. 4 : Rarement. 5 : Jamais. Pour analyser ces facteurs on va réaliser une analyse de la composantes principale (ACP) dans le but de donner les liens pouvant exister entre les différentes données et à en tirer une information statistique qui permet de décrire de façon plus succincte les principales informations contenues dans ces données. On peut également chercher à classer les données en différents sous-groupes plus homogènes.

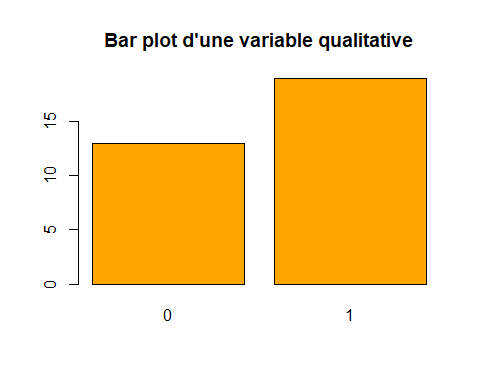
summary(ques[,39:49])

## 9 [Presse spécialisée] 9 [Presse généraliste] 9 [Site Internet]  
## Min. :1.000 Min. :1.000 Min. :1.000   
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:1.000   
## Median :2.000 Median :2.000 Median :2.000   
## Mean :2.469 Mean :2.375 Mean :2.094   
## 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:2.250   
## Max. :5.000 Max. :5.000 Max. :4.000   
## 9[Emission télévisuelle] Bande annonce 9 [Conseils d'un ami]  
## Min. :1.000 Min. :1.000 Min. :1.000   
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:2.000   
## Median :2.000 Median :2.000 Median :2.000   
## Mean :2.188 Mean :2.125 Mean :2.062   
## 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:2.000   
## Max. :5.000 Max. :5.000 Max. :5.000   
## 9[Conseils d'un membre de la famille] Critiques de films 9 [Bandes Annonces]  
## Min. :1.000 Min. :1.000 Min. :1.000   
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:1.000   
## Median :2.000 Median :2.000 Median :1.000   
## Mean :2.375 Mean :2.156 Mean :1.812   
## 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:2.000   
## Max. :5.000 Max. :5.000 Max. :5.000   
## 9 [L'affiche] 9 [Le hasard]   
## Min. :1.000 Min. :1.00   
## 1st Qu.:1.000 1st Qu.:2.00   
## Median :2.000 Median :2.00   
## Mean :2.094 Mean :2.25   
## 3rd Qu.:3.000 3rd Qu.:3.00   
## Max. :5.000 Max. :4.00

## 2/-Statistique descriptive :

### i :

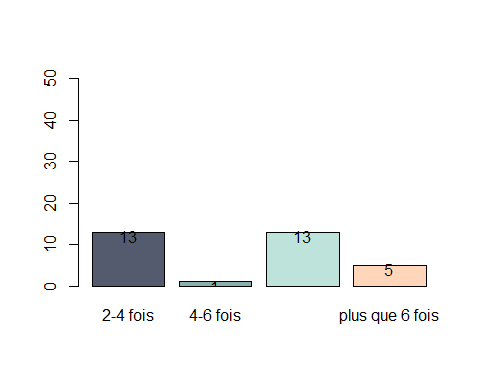
barplot(table(ques$genre),col="orange", main = "Bar plot d'une variable qualitative")



Cette enquête a été effectué sur 30 personnes .Parmis les personnes interrogées, il y en a 36.7% de femmes et 63.3% d’hommes.

### ii :

colors=c("#555b6e","#89b0ae","#bee3db","#ffd6ba")  
b <- barplot(table(ques$`3-Combien de fois par semaine regardez-vous les films/séries?`),col=colors,ylim=c(0,50))  
text(b[,1],table(ques$`3-Combien de fois par semaine regardez-vous les films/séries?`)-1,table(ques$`3-Combien de fois par semaine regardez-vous les films/séries?`))



Interprétation: on remarque que la majorité des repondants regarde entre 1 et 4 film dans la semaine

## 3/- Analyse en Composantes Principales :

Pour avoir les principaux raisons pour regarder un film,nous avouns effectuer une ACP sur 11 variables quantitatives (Presse spécialisée, Presse généraliste , Site Internet , Emission télévisuelle, Bande annonce, Conseils d’un ami, Conseils d’un membre de la famille, Critiques de films ,Bandes Annonces,L’affiche ,Le hasard) Cette méthode synthetise toutes ces informations en seulement quelques nouvelles variables appelées composantes principales( une combinaison linéaire des variables d’origine. Elle regroupe également les individus ayant les meme raisons de regarder les films.

### \*Pertinence de l’ACP:

Le corrgram donné ci-dessous permet d’étudier les corrélations entre les variables quantitatives : il est clair qu’il existe des corrélations importantes entre des groupes de variables ce qui suggère la pertinence de cette ACP.

X=as.matrix(ques[,39:49])  
p=ncol(X)  
M=diag(rep(1,p))   
M<-cor(X)

COL2(diverging = c("RdBu", "BrBG", "PiYG", "PRGn", "PuOr", "RdYlBu"), n = 200)

corrplot(M, type="upper", order="hclust",   
 col=COL2('PuOr', 10))



### \*Choix des axes :

library(FactoMineR)

X=ques[,39:49]  
res.pca=PCA(X,ncp = 5,graph = F)  
head(res.pca$eig)

## eigenvalue percentage of variance cumulative percentage of variance  
## comp 1 6.7935040 61.759128 61.75913  
## comp 2 1.1769713 10.699739 72.45887  
## comp 3 0.7299692 6.636084 79.09495  
## comp 4 0.6271021 5.700929 84.79588  
## comp 5 0.5613177 5.102888 89.89877  
## comp 6 0.3796950 3.451773 93.35054

#### \*\*critére de kaiseur :

Le critère de Kaiser nous conduit à sélectionner 2 axes (dont la valeur propre est supérieure à 1).

#### \*\*Taux d’inertie :

Les deux premiers axes retiennent environ 72.45887 % > 60% de l’inertie totale donc on peut retenir ces 2 axes selon ce critère.

### \*Les coordonnées :

res.pca$var$coord

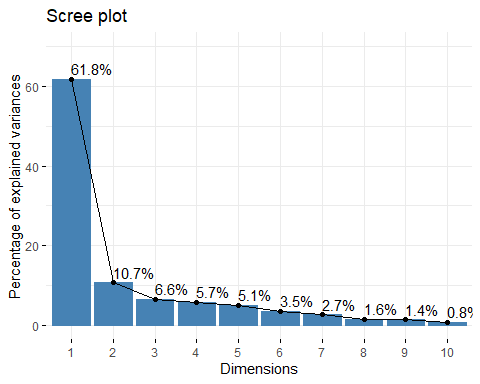
res.pca$var$cos2

## Dim.1 Dim.2 Dim.3  
## 9 [Presse spécialisée] 0.8141510 4.545945e-02 0.0404694246  
## 9 [Presse généraliste] 0.8371014 3.711057e-02 0.0548922134  
## 9 [Site Internet] 0.4663137 9.313348e-04 0.2301383301  
## 9[Emission télévisuelle] 0.7466353 4.269405e-02 0.0007380943  
## Bande annonce 0.5841523 1.540944e-02 0.1655455194  
## 9 [Conseils d'un ami] 0.5993294 1.953473e-01 0.0432247293  
## 9[Conseils d'un membre de la famille] 0.6133992 1.682963e-02 0.0349138301  
## Critiques de films 0.5771250 2.455068e-01 0.0040886395  
## 9 [Bandes Annonces] 0.3029363 4.975399e-01 0.0191265354  
## 9 [L'affiche] 0.6261784 8.006376e-02 0.0091720539  
## 9 [Le hasard] 0.6261820 7.896905e-05 0.1276598377  
## Dim.4 Dim.5  
## 9 [Presse spécialisée] 0.0044021311 0.005211801  
## 9 [Presse généraliste] 0.0003689796 0.001904090  
## 9 [Site Internet] 0.0792467899 0.189500639  
## 9[Emission télévisuelle] 0.0348917597 0.089838149  
## Bande annonce 0.0191083955 0.148682370  
## 9 [Conseils d'un ami] 0.0014507219 0.022460589  
## 9[Conseils d'un membre de la famille] 0.2020950451 0.004849832  
## Critiques de films 0.0410157280 0.005888910  
## 9 [Bandes Annonces] 0.0680128288 0.042414740  
## 9 [L'affiche] 0.1732655113 0.022726621  
## 9 [Le hasard] 0.0032442465 0.027839955

#### \*\*Critère de Coude :

library(factoextra)

fviz\_screeplot(res.pca, addlabels = TRUE, ylim = c(0, 70), ncp=10)



Conclusion :: Le critère de coude conduit aussi à retenir les 2 premiers axes. ==>Après la réalisation de ces 3 critères on va retenir les 2 premiers axes.

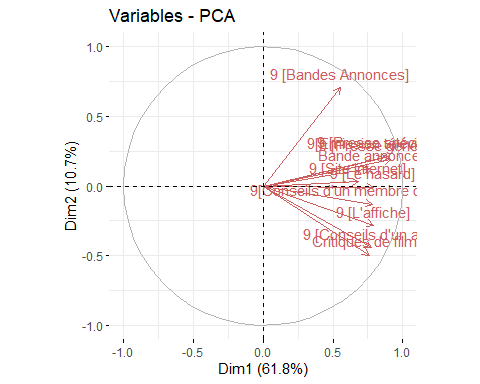
### \*Cercle de corrélation :

La corrélation entre une variable et une composante principale (PC) est utilisée comme coordonnées de la variable sur la composante principale. La représentation des variables diffère de celle des observations : les observations sont représentées par leurs projections, mais les variables sont représentées par leurs corrélations.

#### Qualité de représentation :

##### méthode1:

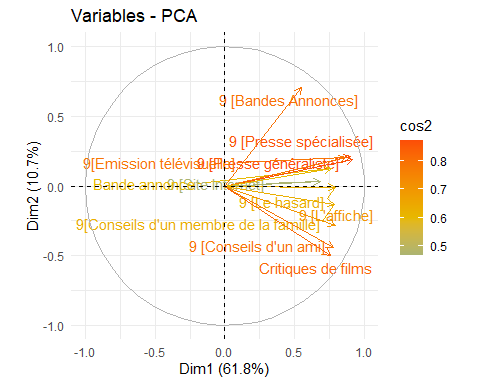
fviz\_pca\_var(res.pca, col.var = "indianred",col.quanti.sup="darkblue" )



Le graphique ci-dessus montre que les variables les plus corrélées à l’axe 1 sont : Emission télevisuelle , les presses..(Réseaux sociaux) On peut voir aussi que les variables les plus corrélées à l’axe 2 sont : Bandes Annonces , Critiques de films,conseil d’un amis.

##### méthode2:

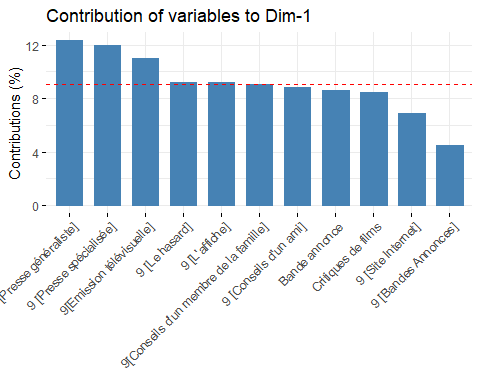
fviz\_pca\_var(res.pca, col.var="cos2",gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"), repel = TRUE) +  
 scale\_color\_gradient2(low="#00AFBB", mid="#E7B800",   
 high="#FC4E07", midpoint=0.6) +   
 theme\_minimal()

 Interpretation : La variable {conseils d’amis} est positivement corrélée a la variable {critique de films } et {affiche} . L’individus(es) qui est influée par la l’affiche est plus influée par la conseils d’amis et critique de films lors du choix du film

### \*Contribution :

#### Premier axe :

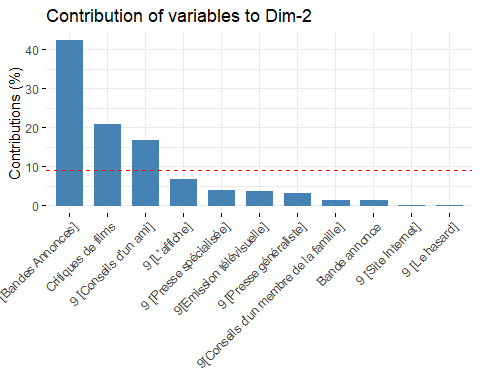
fviz\_contrib(res.pca, choice = "var", axes = 1, top = 30)



Presse généraliste, Presse spécialisée , Emission télévisuelle sont les variables qui ont un rôle dominant dans la formation du premier axe factoriel.

#### Deuxiéme axe :

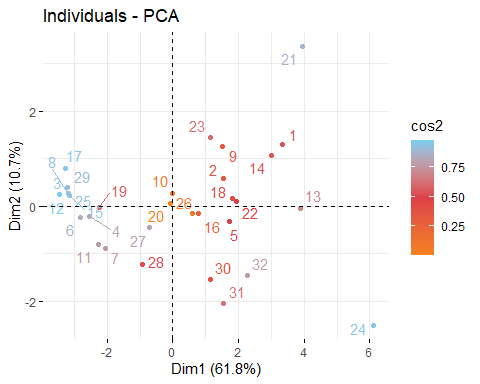
fviz\_contrib(res.pca, choice = "var", axes = 2, top = 30)

 Bandes Annonces, Critiques de films et Conseils d’un ami sont les variables qui ont un rôle dominant dans la formation du deuxiéme axe factoriel.

### \*Graphique des individus :

#### carte des individus:

fviz\_pca\_ind (res.pca, col.ind = "cos2",  
gradient.cols = c("#f6821f", "#dc4048", "#7acdf1"),  
repel = TRUE)



Le prémier axe oppose des individus tels que 14 1 9 23 21 9 (à droite du graphe, caractérisés par une coordonnée fortement positive sur l’axe) à des individus comme 28 11 3 7 11 (à gauche du graphe , caractérisés par une coordonnée fortement négative sur l’axe).

\*Le groupe auquel les individus 14 1 21 appartiennent (caractérisés par une coordonnée positive sur l’axe 1) partage : De fortes valeurs pour les variables : les presses, emission televisuelle…

\*Le groupe auquel les individus 17 8 4 12 4 11 appartiennent (caractérisés par une coordonnée négative sur l’axe) partage : bande d’annance …

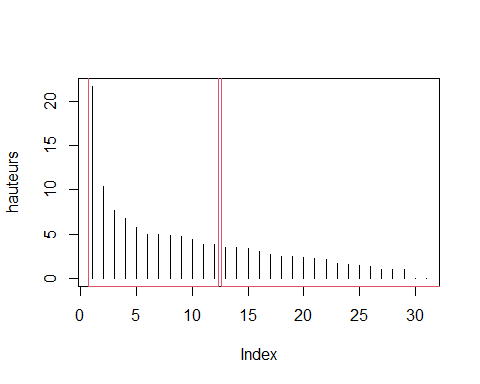
## 4/-Classification :

### 1/Visualisation des individus et les grouper :

library(NbClust)  
library(cluster)  
  
cl3 = kmeans(X, centers = 2)   
cl3$cluster

## [1] 1 1 2 2 1 2 2 2 1 1 2 2 1 1 2 1 2 1 2 1 1 1 1 1 2 1 2 2 2 1 1 1

classif<-agnes(X, method="ward")  
d=dist(scale(X),method = "euclidian")  
hc1=hclust(d,method = "ward.D") # or complete or average  
classif2<-as.hclust(classif)  
plot(rev(classif2$height), type="h", ylab="hauteurs")  
rect.hclust(hc1,k=2)



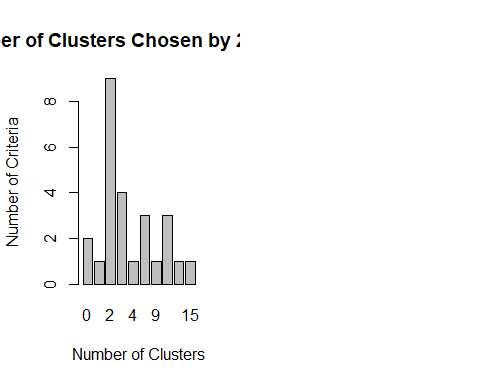
classes<-cutree(classif2,k=2)  
classes

## [1] 1 1 2 2 1 2 2 2 1 1 2 2 1 1 2 1 2 1 2 1 1 1 1 1 2 1 1 2 2 1 1 1

#2. Déterminer la meilleure partition  
set.seed(1234)  
nc\_km <- NbClust(X, min.nc = 2, max.nc = 15, method = "kmeans" )

## \* Among all indices:   
## \* 9 proposed 2 as the best number of clusters   
## \* 4 proposed 3 as the best number of clusters   
## \* 1 proposed 4 as the best number of clusters   
## \* 3 proposed 5 as the best number of clusters   
## \* 1 proposed 9 as the best number of clusters   
## \* 3 proposed 12 as the best number of clusters   
## \* 1 proposed 13 as the best number of clusters   
## \* 1 proposed 15 as the best number of clusters   
##   
## \*\*\*\*\* Conclusion \*\*\*\*\*   
##   
## \* According to the majority rule, the best number of clusters is 2

# Afficher le graphique des nombres de classes  
barplot(table(nc\_km$Best.n[1, ]), xlab = "Number of Clusters", ylab = "Number of Criteria", main = "Number of Clusters Chosen by 26 Criteria")



# Application de la méthode k-means avec le nombre de clusters optimal  
k <- kmeans(X, centers = 2, nstart = 25)  
  
k

## K-means clustering with 2 clusters of sizes 14, 18  
##   
## Cluster means:  
## 9 [Presse spécialisée] 9 [Presse généraliste] 9 [Site Internet]  
## 1 1.071429 1.071429 1.357143  
## 2 3.555556 3.388889 2.666667  
## 9[Emission télévisuelle] Bande annonce 9 [Conseils d'un ami]  
## 1 1.142857 1.428571 1.5  
## 2 3.000000 2.666667 2.5  
## 9[Conseils d'un membre de la famille] Critiques de films 9 [Bandes Annonces]  
## 1 1.285714 1.285714 1.071429  
## 2 3.222222 2.833333 2.388889  
## 9 [L'affiche] 9 [Le hasard]  
## 1 1.357143 1.500000  
## 2 2.666667 2.833333  
##   
## Clustering vector:  
## [1] 2 2 1 1 2 1 1 1 2 2 1 1 2 2 1 2 1 2 1 2 2 2 2 2 1 2 1 1 1 2 2 2  
##   
## Within cluster sum of squares by cluster:  
## [1] 33.07143 201.61111  
## (between\_SS / total\_SS = 50.7 %)  
##   
## Available components:  
##   
## [1] "cluster" "centers" "totss" "withinss" "tot.withinss"  
## [6] "betweenss" "size" "iter" "ifault"

#3. Rajouter la classe d'affectation de chaque individu en tant que variable  
  
X.comp<-cbind.data.frame(X, as.factor(classes))

#4. Description des classes  
# La fonction catdes permet de trier les variables quanti de la plus caractérisante à la moins caractérisante en positif  
  
res.cat=catdes(X.comp, num.var=12)  
res.cat

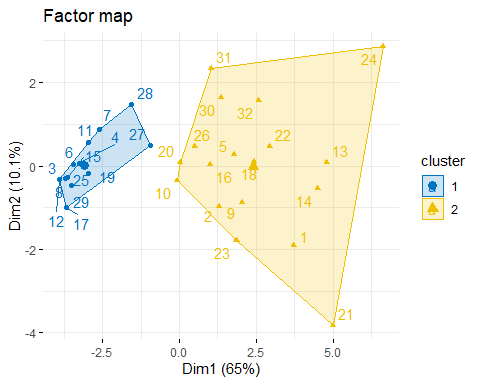
##   
## Link between the cluster variable and the quantitative variables  
## ================================================================  
## Eta2 P-value  
## 9.[Presse.généraliste] 0.6957099 3.034640e-09  
## 9.[Presse.spécialisée] 0.6949062 3.158776e-09  
## 9[Emission.télévisuelle] 0.5107244 4.332846e-06  
## 9.[Site.Internet] 0.4710674 1.445676e-05  
## Critiques.de.films 0.4662791 1.662423e-05  
## 9[Conseils.d'un.membre.de.la.famille] 0.4661137 1.670431e-05  
## 9.[Le.hasard] 0.4537112 2.385623e-05  
## 9.[L'affiche] 0.4403977 3.468597e-05  
## Bande.annonce 0.3380852 4.827707e-04  
## 9.[Conseils.d'un.ami] 0.3311995 5.688090e-04  
## 9.[Bandes.Annonces] 0.2639924 2.630696e-03  
##   
## Description of each cluster by quantitative variables  
## =====================================================  
## $`1`  
## v.test Mean in category Overall mean  
## 9.[Presse.généraliste] 4.644029 3.315789 2.37500  
## 9.[Presse.spécialisée] 4.641346 3.473684 2.46875  
## 9[Emission.télévisuelle] 3.979002 2.947368 2.18750  
## 9.[Site.Internet] 3.821399 2.631579 2.09375  
## Critiques.de.films 3.801927 2.789474 2.15625  
## 9[Conseils.d'un.membre.de.la.famille] 3.801253 3.157895 2.37500  
## 9.[Le.hasard] 3.750340 2.789474 2.25000  
## 9.[L'affiche] 3.694906 2.631579 2.09375  
## Bande.annonce 3.237382 2.631579 2.12500  
## 9.[Conseils.d'un.ami] 3.204245 2.473684 2.06250  
## 9.[Bandes.Annonces] 2.860728 2.315789 1.81250  
## sd in category Overall sd p.value  
## 9.[Presse.généraliste] 0.9761704 1.3635890 3.416795e-06  
## 9.[Presse.spécialisée] 1.0447070 1.4574030 3.461469e-06  
## 9[Emission.télévisuelle] 1.1458706 1.2854352 6.920513e-05  
## 9.[Site.Internet] 0.8085417 0.9473441 1.326968e-04  
## Critiques.de.films 1.0041465 1.1210869 1.435748e-04  
## 9[Conseils.d'un.membre.de.la.famille] 1.2253102 1.3863171 1.439662e-04  
## 9.[Le.hasard] 0.8321783 0.9682458 1.765952e-04  
## 9.[L'affiche] 0.8085417 0.9797760 2.199684e-04  
## Bande.annonce 0.9846467 1.0532687 1.206319e-03  
## 9.[Conseils.d'un.ami] 0.8187552 0.8637672 1.354172e-03  
## 9.[Bandes.Annonces] 1.2999041 1.1842060 4.226694e-03  
##   
## $`2`  
## v.test Mean in category Overall mean  
## 9.[Bandes.Annonces] -2.860728 1.076923 1.81250  
## 9.[Conseils.d'un.ami] -3.204245 1.461538 2.06250  
## Bande.annonce -3.237382 1.384615 2.12500  
## 9.[L'affiche] -3.694906 1.307692 2.09375  
## 9.[Le.hasard] -3.750340 1.461538 2.25000  
## 9[Conseils.d'un.membre.de.la.famille] -3.801253 1.230769 2.37500  
## Critiques.de.films -3.801927 1.230769 2.15625  
## 9.[Site.Internet] -3.821399 1.307692 2.09375  
## 9[Emission.télévisuelle] -3.979002 1.076923 2.18750  
## 9.[Presse.spécialisée] -4.641346 1.000000 2.46875  
## 9.[Presse.généraliste] -4.644029 1.000000 2.37500  
## sd in category Overall sd p.value  
## 9.[Bandes.Annonces] 0.2664694 1.1842060 4.226694e-03  
## 9.[Conseils.d'un.ami] 0.4985185 0.8637672 1.354172e-03  
## Bande.annonce 0.6249260 1.0532687 1.206319e-03  
## 9.[L'affiche] 0.6056929 0.9797760 2.199684e-04  
## 9.[Le.hasard] 0.4985185 0.9682458 1.765952e-04  
## 9[Conseils.d'un.membre.de.la.famille] 0.5756396 1.3863171 1.439662e-04  
## Critiques.de.films 0.4213250 1.1210869 1.435748e-04  
## 9.[Site.Internet] 0.4615385 0.9473441 1.326968e-04  
## 9[Emission.télévisuelle] 0.2664694 1.2854352 6.920513e-05  
## 9.[Presse.spécialisée] 0.0000000 1.4574030 3.461469e-06  
## 9.[Presse.généraliste] 0.0000000 1.3635890 3.416795e-06

# #proba: the significance threshold considered to characterize the category (by default 0.05)

res.hcpc <- HCPC(X, graph = FALSE , nb.clust = 2)

## Warning: Unknown or uninitialised column: `call`.  
## Unknown or uninitialised column: `call`.

fviz\_cluster(res.hcpc,  
 repel = TRUE,   
 show.clust.cent = TRUE,   
 palette = "jco",   
 ggtheme = theme\_minimal(),  
 main = "Factor map"  
 )



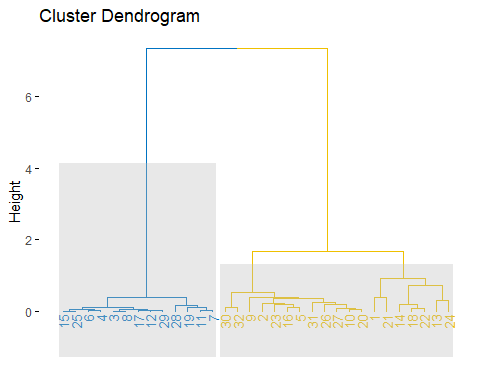
# Extract cluster membership information  
cluster\_labels <- res.hcpc$data.clustering

La classification réalisée sur les individus fait apparaître classes. La classe 1 est composée d’individus tels que 8 25 6 17 28 3 15 11 12 4 Ce groupe est caractérisé par : De fortes valeurs pour les variables : presses

la classe 2 est composée d’individus tels que 20 26 23 5 16 27 30 32 31 20 2 14 13 11 Ce groupe est caractérisé par : De fortes valeurs pour les variables Développer des relations sociales

### les différents classes :

fviz\_dend(res.hcpc,   
 cex = 0.7,   
 palette = "jco",   
 rect = TRUE, rect\_fill = TRUE,   
 labels\_track\_height = 0.8,  
 color\_labels\_by\_k = TRUE # Set colors based on clustering results  
 )



## 5/- AFC :

### Tableaus des profils :

library(questionr)  
library("factoextra")  
library(FactoMineR)  
  
variables <- ques[, c(5,46)]  
  
variables <- lapply(variables, factor)  
  
tab1<-table(variables)  
tab1

## Critiques de films  
## 4-Votre statut? 1 2 3 4 5  
## Autre 1 1 1 0 0  
## Cadre ou profession intellectuelle 0 0 0 0 1  
## Demandeur d’emploi 0 0 1 0 0  
## Employé 1 2 0 1 0  
## Etudiant, lycéen 7 10 3 0 1  
## Profession libérale 1 0 0 1 0

#Tableau des profils-lignes  
lprop(tab1)

## Critiques de films  
## 4-Votre statut? 1 2 3 4 5 Total  
## Autre 33.3 33.3 33.3 0.0 0.0 100.0  
## Cadre ou profession intellectuelle 0.0 0.0 0.0 0.0 100.0 100.0  
## Demandeur d’emploi 0.0 0.0 100.0 0.0 0.0 100.0  
## Employé 25.0 50.0 0.0 25.0 0.0 100.0  
## Etudiant, lycéen 33.3 47.6 14.3 0.0 4.8 100.0  
## Profession libérale 50.0 0.0 0.0 50.0 0.0 100.0  
## All 31.2 40.6 15.6 6.2 6.2 100.0

#Tableau des profils-colonnes  
tab1

## Critiques de films  
## 4-Votre statut? 1 2 3 4 5  
## Autre 1 1 1 0 0  
## Cadre ou profession intellectuelle 0 0 0 0 1  
## Demandeur d’emploi 0 0 1 0 0  
## Employé 1 2 0 1 0  
## Etudiant, lycéen 7 10 3 0 1  
## Profession libérale 1 0 0 1 0

cprop(tab1)

## Critiques de films  
## 4-Votre statut? 1 2 3 4 5 All   
## Autre 10.0 7.7 20.0 0.0 0.0 9.4  
## Cadre ou profession intellectuelle 0.0 0.0 0.0 0.0 50.0 3.1  
## Demandeur d’emploi 0.0 0.0 20.0 0.0 0.0 3.1  
## Employé 10.0 15.4 0.0 50.0 0.0 12.5  
## Etudiant, lycéen 70.0 76.9 60.0 0.0 50.0 65.6  
## Profession libérale 10.0 0.0 0.0 50.0 0.0 6.2  
## Total 100.0 100.0 100.0 100.0 100.0 100.0

### Test du chi-2

test<-chisq.test(tab1)

test  
## Pearson's Chi-squared test  
## X-squared = 33.975, df = 20, p-value = 0.02629

Interpretation :

La valeur du χ2 est de 33.975 . Le degré de liberté du χ2 est de (p−1)∗(q−1)=(5−1)∗(6−1)=20 En se basant sur la valeur p de 0.02629 , qui est inférieure au seuil conventionnel de 0,05, nous pouvons rejeter l’hypothèse nulle et conclure qu’il y a des preuves d’une association significative entre les deux variables. En d’autres termes, les résultats suggèrent que les variables “Votre statut” et “Dans quelle tranche d’âge vous situez vous” ne sont pas indépendantes et sont liées entre elles.

l’AFC est donc pertinente.

### Exécution de l’AFC

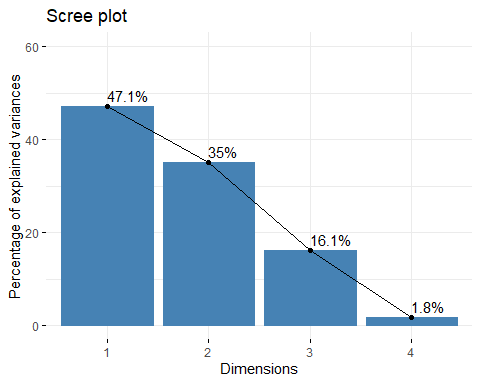
res.ca<-CA(tab1,graph=F)

### Choix des axes :

res.ca$eig

## eigenvalue percentage of variance cumulative percentage of variance  
## dim 1 0.49973668 47.068523 47.06852  
## dim 2 0.37170116 35.009286 82.07781  
## dim 3 0.17123392 16.127949 98.20576  
## dim 4 0.01904986 1.794242 100.00000

fviz\_screeplot (res.ca, addlabels = TRUE, ylim = c (0, 60))

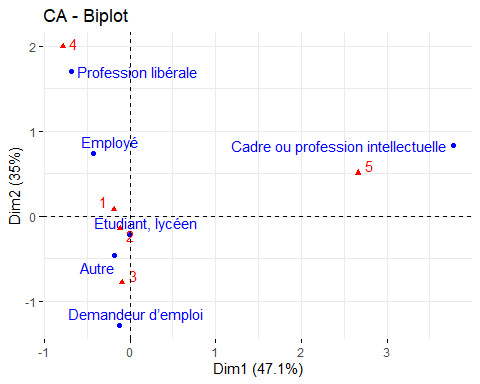
Les 2 premiers axes de l’AFC expriment 72.1% de l’inertie totale du jeu de données ; cela signifie que la majoritée la variabilité totale du nuage des individus (ou des variables) est représentée dans ce plan

.

### Graphique des variables et des individus :

1 :Toujours. 2 : Souvent. 3 : Parfois. 4 : Rarement. 5 : Jamais.

fviz\_ca\_biplot (res.ca, repel = TRUE,   
 ggtheme = theme\_minimal())

▶ Quelques associations : profession libérale 4(rarement ) : les employées rearement influsence par la Critiques de films ,et demandeur d’emploi sont souvent influence par la Critiques de films ,autres(chomeur ou retretaire) 5(toujours). ▶ “statut employé” est plus excentre que “souvent” sur le premier axe car ´ les statut employé sont beaucoup mieux carcterisés par le modalités souvent ´que l’inverse.

## 5/- ACM :

Cette partie est spécifiée pour comprendre la position de chaque modalité dans notre jeu de données c’est-à-dire on va faire un ACP sur les variables qualitatives. Nous allons étudier les modalités pour mieux comprendre la consommation et le moment de consommation des sucreries d’un individu .

active<-ques[,50:54]  
sapply(active, class)

## [Sur l'ordinateur] [A la télévision] [Au cinéma]   
## "character" "character" "character"   
## [La location de films] [Le streaming]   
## "character" "character"

res.mca <- MCA (active, graph = FALSE)  
fviz\_screeplot(res.mca) A picture containing text, diagram, line, screenshot

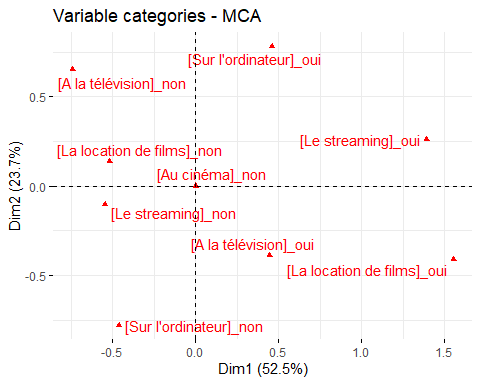
Description automatically generated

res.mca$eig

## eigenvalue percentage of variance cumulative percentage of variance  
## dim 1 0.42002787 52.503484 52.50348  
## dim 2 0.18920702 23.650878 76.15436  
## dim 3 0.15809878 19.762347 95.91671  
## dim 4 0.03266632 4.083291 100.00000

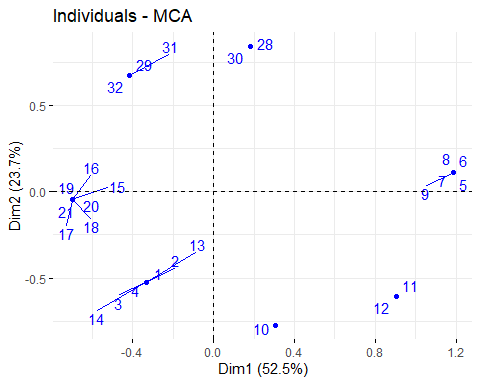
Critère 1 : Retenir les axes dont les valeurs propres sont supérieures à la moyenne des valeurs propres ( m= 4) est 1/m=0.25 . Donc d’aprés ce critère on choisit un seul axe . Critère 2: Critère du coude : On remarque que le coude se trouve au niveau du deuxième axe (voir la figure Scree plot), d’après ce critère, on devrait retenir les 2 premiers axes. Critère 3 : ce critère est specifiée pour l’inertie cumulée , si on prend un seul axe l’inertie cumulée est 49.29619% mais si on prend les trois premiers axes on obtient 68.35856% Conclusion: on retient les deux premiers axes

fviz\_mca\_var (res.mca,  
 repel = TRUE,   
 ggtheme = theme\_minimal ())



G1=[location de films\_non,Au cinéma\_oui,sur l’ordinateur\_non,A la télévision\_non] /// G2=[sur l’ordinateur\_oui,location de films\_non,télévision\_oui,La streaming\_oui] Le groupe G1 semble préférer l’expérience de visionnage de films au cinéma, ce qui peut être considéré comme une activité de divertissement saine et socialement engageante. D’un autre côté, le groupe G2 a tendance à privilégier la visualisation sur ordinateur et à la télévision, ce qui peut entraîner une sédentarité accrue et des effets négatifs sur la santé.

fviz\_mca\_ind (res.mca,select.ind = list(cos2 = 0.4),  
 repel = TRUE,   
 ggtheme = theme\_minimal ())

 Le premier groupe est composé de 1 14 2 13 3 personnes, qui ont tendance à regarder des films généralement au cinéma. De même, le groupe G2 est constitué de 12 7 10 6 30 individus qui préfèrent utiliser des ordinateurs et des télévisions.