

apprentissage d'ACM (STHDA)

Je note très bien que la totalité de ce TP se trouve sur le lien suivant : <http://www.sthda.com/french/articles/38-methodes-des-composantes-principales-dans-r-guide-pratique/75-acm-analyse-des-correspondances-multiples-avec-r-l-essen>

L'objectif c'est de faire apprendre l'ACF en utilisant R notebook. Rq: J'ai changé quelque codes avec des commentaires de plus. ##Introduction : L'Analyse des Correspondances Multiples (ACM ou MCA pour multiple correspondence analysis) est une extension de l'analyse factorielle des correspondances pour résumer et visualiser un tableau de données contenant plus de deux variables catégorielles. On peut aussi la considérer comme une généralisation de l'analyse en composantes principales lorsque les variables à analyser sont catégorielles plutôt que quantitatives (Abdi and Williams 2010).

L'ACM est généralement utilisée pour analyser des données d'enquête ou de sondage.

L'objectif est d'identifier:

Un groupe de personnes ayant un profil similaire dans leurs réponses aux questions Les associations entre les catégories des variables. [source STHDA]

Calcul :

Pour le calcul nous allons utiliser la fonction MCA() incluse dans le package FactoMiner, et pour les visualization nous allons opter pour factoextra. ##Charger les librararies :

```
library(FactoMiner)
```

```
## Warning: package 'FactoMiner' was built under R version 3.4.2
```

```
library(factoextra)
```

```
## Warning: package 'factoextra' was built under R version 3.4.2
```

```
## Loading required package: ggplot2
```

```
## Welcome! Related Books: `Practical Guide To Cluster Analysis in R` at https://goo.gl/13EFCZ
```

Format des données :

Nous allons utiliser les données : dataset_suivi ###Charger les données :

```
# xlsx files
```

```
library(xlsx)
```

```
## Warning: package 'xlsx' was built under R version 3.4.2
```

```
## Loading required package: rJava
```

```
## Warning: package 'rJava' was built under R version 3.4.2
```

```
## Loading required package: xlsxjars
```

```
## Warning: package 'xlsxjars' was built under R version 3.4.2
```

```
my_data <- read.xlsx("dataset.xlsx",sheetIndex = 1,as.data.frame = T,encoding = 'UTF-8')
```

nommer les colonnes :


```
## $ performer_travail_projet : Factor w/ 5 levels "D'accord","Désaccord",...: 4 4 1 1 5 3 4 1 4 1 ...
## $ motivation_certificat : Factor w/ 5 levels "D'accord","Désaccord",...: 1 4 1 2 5 1 2 1 2 4 ...
## $ trouver_emploi : Factor w/ 5 levels "D'accord","Désaccord",...: 3 4 1 5 5 2 3 1 2 2 ...
## $ nbr_certificat : num 1 2 1 1 1 1 1 1 3 1 ...
## $ satisfait_gestion_temps : Factor w/ 5 levels "D'accord","Désaccord",...: 1 4 1 3 4 4 2 1 4 3 ...
## $ objectifs_atteints : Factor w/ 5 levels "D'accord","Désaccord",...: 4 4 1 1 3 1 5 1 4 1 ...
## $ satisfaction : num 7 10 9 6 1 7 3 6 7 6 ...
## $ vacances_etude : Factor w/ 3 levels "Les deux","Les vacances",...: 1 1 1 3 3 3 1 3 1 1 ...
## $ temps_alloue_cours : Factor w/ 18 levels "0","0.01","0.1",...: 14 9 13 5 9 11 9 7 13 5 ...
## $ mooc_inclu_etude : Factor w/ 2 levels "Non","Oui": 1 1 2 2 2 2 2 2 1 ...
## $ edt_chargee : Factor w/ 2 levels "Non","Oui": 1 1 2 2 2 2 2 1 1 2 ...
## $ plateformes_utilisees : Factor w/ 30 levels "Canvas, OpenClassrooms, FUN",...: 14 2 21 23 23 2 ...
## $ videos_cours : Factor w/ 5 levels "Bien","Mauvais",...: 4 4 1 3 1 3 1 4 4 1 ...
## $ transcription : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 1 1 3 3 4 1 4 ...
## $ diapositives_cours : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 1 3 1 4 1 4 1 ...
## $ forum_discussion : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 4 3 3 5 1 4 1 ...
## $ examens_exercices : Factor w/ 5 levels "Bien","Mauvais",...: 4 4 1 3 2 1 1 3 1 1 ...
## $ accompagnement_pedagogique : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 3 1 2 1 1 1 3 ...
## $ aides_financiere : Factor w/ 2 levels "Non","Oui": 2 2 2 1 1 1 2 1 1 1 ...
## $ prix : num NA 10 20 0 0 0 0 100 50 30 ...
## $ outils_exercice : Factor w/ 2 levels "Non","Oui": 2 2 2 1 2 2 2 2 1 2 ...
## $ sciences_donnees : Factor w/ 3 levels "Inintéressant",...: 2 2 2 2 2 3 2 2 2 2 ...
## $ informatique : Factor w/ 3 levels "Inintéressant",...: 2 2 3 2 3 2 2 2 2 2 ...
## $ sciences_physiques : Factor w/ 3 levels "Inintéressant",...: 2 3 2 2 2 2 3 2 3 3 ...
## $ business : Factor w/ 3 levels "Inintéressant",...: 2 3 3 3 1 2 3 1 1 1 ...
## $ arts : Factor w/ 3 levels "Inintéressant",...: 1 3 2 1 1 3 2 1 3 3 ...
## $ langues : Factor w/ 3 levels "Inintéressant",...: 1 3 2 1 1 2 1 1 3 3 ...
## $ sciences_sociales : Factor w/ 3 levels "Inintéressant",...: 1 3 3 1 1 2 1 1 1 1 ...
## $ sexe : Factor w/ 2 levels "Femme","Homme": 2 2 2 2 2 1 2 2 2 2 ...
## $ pays : Factor w/ 13 levels "Autre","Burkinafaso",...: 8 8 5 5 5 5 8 5 5 5 ...
## $ age : num 22 22 22 22 20 22 22 22 23 23 ...
## $ est_etudiant : Factor w/ 2 levels "Non","Oui": 2 2 2 2 2 2 2 2 2 2 ...
## $ niveau_actuel : Factor w/ 5 levels "BAC +2","BAC +3",...: 4 4 2 4 2 3 4 4 3 4 ...
## $ formation : Factor w/ 6 levels "FIG","FIL","FIP",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ background : Factor w/ 4 levels "Autre école",...: 3 3 4 3 3 3 3 3 3 3 ...
```

```
#change from factor to numeric the feature : temps_alloue_cours
my_data2$temps_alloue_cours<-as.numeric(my_data2$temps_alloue_cours)
```

suivi terminé et non terminé :

```
# Choisir que les personnes qui ont suivi des moocs et qu'ils l'ont terminé ou pas
data_suivi<-filter(my_data2,a_suivi %in% c("Oui, je l'ai suivi mais je ne l'ai pas terminé", "Oui, je l'ai terminé"))
#Supprimer chaque colonne qui est totalement des NaN ( en fait ce sont celle de la partie: jamais suivi)
data_suivi<-data_suivi[ , ! apply( data_suivi , 2 , function(x) all(is.na(x)) ) ]
```

```
str(data_suivi
)
```

```
## 'data.frame': 115 obs. of 36 variables:
## $ a_suivi : Factor w/ 3 levels "Jamais","Oui, je l'ai suivi et j'ai réussi",...: 2 2 2 2 2 2 2 2 2 2 ...
## $ performer_travail_projet : Factor w/ 5 levels "D'accord","Désaccord",...: 4 4 1 1 5 3 4 1 4 1 ...
## $ motivation_certificat : Factor w/ 5 levels "D'accord","Désaccord",...: 1 4 1 2 5 1 2 1 2 4 ...
## $ trouver_emploi : Factor w/ 5 levels "D'accord","Désaccord",...: 3 4 1 5 5 2 3 1 2 2 ...
## $ nbr_certificat : num 1 2 1 1 1 1 1 1 3 1 ...
```

```
## $ satisfait_gestion_temps : Factor w/ 5 levels "D'accord","Désaccord",...: 1 4 1 3 4 4 2 1 4 3 ...
## $ objectifs_atteints : Factor w/ 5 levels "D'accord","Désaccord",...: 4 4 1 1 3 1 5 1 4 1 ...
## $ satisfaction : num 7 10 9 6 1 7 3 6 7 6 ...
## $ vacances_etude : Factor w/ 3 levels "Les deux","Les vacances",...: 1 1 1 3 3 3 1 3 1 1 ...
## $ temps_alloue_cours : num 14 9 13 5 9 11 9 7 13 5 ...
## $ mooc_inclu_etude : Factor w/ 2 levels "Non","Oui": 1 1 2 2 2 2 2 2 1 ...
## $ edt_chargee : Factor w/ 2 levels "Non","Oui": 1 1 2 2 2 2 2 1 1 2 ...
## $ plateformes_utilisees : Factor w/ 30 levels "Canvas, OpenClassrooms, FUN",...: 14 2 21 23 23 2 ...
## $ videos_cours : Factor w/ 5 levels "Bien","Mauvais",...: 4 4 1 3 1 3 1 4 4 1 ...
## $ transcription : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 1 1 3 3 4 1 4 ...
## $ diapositives_cours : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 1 3 1 4 1 4 1 ...
## $ forum_discussion : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 4 3 3 5 1 4 1 ...
## $ examens_exercices : Factor w/ 5 levels "Bien","Mauvais",...: 4 4 1 3 2 1 1 3 1 1 ...
## $ accompagnement_pedagogique : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 3 1 2 1 1 1 3 ...
## $ aides_financiere : Factor w/ 2 levels "Non","Oui": 2 2 2 1 1 1 2 1 1 1 ...
## $ prix : num NA 10 20 0 0 0 0 100 50 30 ...
## $ outils_exercice : Factor w/ 2 levels "Non","Oui": 2 2 2 1 2 2 2 2 1 2 ...
## $ sciences_donnees : Factor w/ 3 levels "Inintéressant",...: 2 2 2 2 2 3 2 2 2 2 ...
## $ informatique : Factor w/ 3 levels "Inintéressant",...: 2 2 3 2 3 2 2 2 2 2 ...
## $ sciences_physiques : Factor w/ 3 levels "Inintéressant",...: 2 3 2 2 2 2 3 2 3 3 ...
## $ business : Factor w/ 3 levels "Inintéressant",...: 2 3 3 3 1 2 3 1 1 1 ...
## $ arts : Factor w/ 3 levels "Inintéressant",...: 1 3 2 1 1 3 2 1 3 3 ...
## $ langues : Factor w/ 3 levels "Inintéressant",...: 1 3 2 1 1 2 1 1 3 3 ...
## $ sciences_sociales : Factor w/ 3 levels "Inintéressant",...: 1 3 3 1 1 2 1 1 1 1 ...
## $ sexe : Factor w/ 2 levels "Femme","Homme": 2 2 2 2 2 1 2 2 2 2 ...
## $ pays : Factor w/ 13 levels "Autre","Burkinafaso",...: 8 8 5 5 5 5 8 5 5 5 ...
## $ age : num 22 22 22 22 20 22 22 22 23 23 ...
## $ est_etudiant : Factor w/ 2 levels "Non","Oui": 2 2 2 2 2 2 2 2 2 2 ...
## $ niveau_actuel : Factor w/ 5 levels "BAC +2","BAC +3",...: 4 4 2 4 2 3 4 4 3 4 ...
## $ formation : Factor w/ 6 levels "FIG","FIL","FIP",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ background : Factor w/ 4 levels "Autre école",...: 3 3 4 3 3 3 3 3 3 3 ...
```

Ces données proviennent d'une enquête menée auprès des étudiants de l'IMT Atlantique à propos des moocs qu'ils ont suivi et terminé ou non.

Les données contiennent 115 lignes (individus) et 36 colonnes (variables). Nous n'utiliserons que certains des individus (étudiants) et variables pour effectuer l'ACM. Les coordonnées des individus et des variables restantes seront prédites.

Nos données contiennent donc des:

Individus actifs (lignes 1:115): individus qui sont utilisés dans l'ACM. Variables actives (toutes les colonnes sauf les colonnes numériques+ plateforme_utilisee+ a_suivi+sexe): variables utilisées dans l'ACM. Variables supplémentaires(sexe, age, satisfaction): elles ne participent pas à l'ACM. Les coordonnées de ces variables seront prédites.

Nous commençons par extraire les individus actifs et les variables actives pour l'ACM:

```
data_suivi.active <- data_suivi[, c( "motivation_certificat" , "trouver_emploi" ,
    "vacances_etude" , "mooc_inclu_etude" , "edt_chargee" ,
    "videos_cours" , "transcription" , "diapositives_cours" ,
    "forum_discussion" , "examens_exercices" , "accompagnement_pedagogique" , "aides_financiere" ,
    "outils_exercice" , "sciences_donnees" , "informatique" ,
    "sciences_physiques" , "business" , "arts" , "langues" ,
    "sciences_sociales" ,
    "est_etudiant" , "niveau_actuel" , "formation" , "background" )]
```

```
str(data_suivi.active)
```

```
## 'data.frame': 115 obs. of 24 variables:
## $ motivation_certificat : Factor w/ 5 levels "D'accord","Désaccord",...: 1 4 1 2 5 1 2 1 2 4 ...
## $ trouver_emploi : Factor w/ 5 levels "D'accord","Désaccord",...: 3 4 1 5 5 2 3 1 2 2 ...
## $ vacances_etude : Factor w/ 3 levels "Les deux","Les vacances",...: 1 1 1 3 3 3 1 3 1 1 ...
## $ mooc_inclu_etude : Factor w/ 2 levels "Non","Oui": 1 1 2 2 2 2 2 2 1 ...
## $ edt_chargee : Factor w/ 2 levels "Non","Oui": 1 1 2 2 2 2 2 1 1 2 ...
## $ videos_cours : Factor w/ 5 levels "Bien","Mauvais",...: 4 4 1 3 1 3 1 4 4 1 ...
## $ transcription : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 1 1 3 3 4 1 4 ...
## $ diapositives_cours : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 1 3 1 4 1 4 1 ...
## $ forum_discussion : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 4 3 3 5 1 4 1 ...
## $ examens_exercices : Factor w/ 5 levels "Bien","Mauvais",...: 4 4 1 3 2 1 1 3 1 1 ...
## $ accompagnement_pedagogique: Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 3 1 2 1 1 1 3 ...
## $ aides_financiere : Factor w/ 2 levels "Non","Oui": 2 2 2 1 1 1 2 1 1 1 ...
## $ outils_exercice : Factor w/ 2 levels "Non","Oui": 2 2 2 1 2 2 2 2 1 2 ...
## $ sciences_donnees : Factor w/ 3 levels "Inintéressant",...: 2 2 2 2 2 3 2 2 2 2 ...
## $ informatique : Factor w/ 3 levels "Inintéressant",...: 2 2 3 2 3 2 2 2 2 2 ...
## $ sciences_physiques : Factor w/ 3 levels "Inintéressant",...: 2 3 2 2 2 2 3 2 3 3 ...
## $ business : Factor w/ 3 levels "Inintéressant",...: 2 3 3 3 1 2 3 1 1 1 ...
## $ arts : Factor w/ 3 levels "Inintéressant",...: 1 3 2 1 1 3 2 1 3 3 ...
## $ langues : Factor w/ 3 levels "Inintéressant",...: 1 3 2 1 1 2 1 1 3 3 ...
## $ sciences_sociales : Factor w/ 3 levels "Inintéressant",...: 1 3 3 1 1 2 1 1 1 1 ...
## $ est_etudiant : Factor w/ 2 levels "Non","Oui": 2 2 2 2 2 2 2 2 2 2 ...
## $ niveau_actuel : Factor w/ 5 levels "BAC +2","BAC +3",...: 4 4 2 4 2 3 4 4 3 4 ...
## $ formation : Factor w/ 6 levels "FIG","FIL","FIP",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ background : Factor w/ 4 levels "Autre école",...: 3 3 4 3 3 3 3 3 3 3 ...
```

Résumé des données :

La fonction `summary()` peut être utilisée pour calculer la fréquence des catégories des variables.

Résumés statistiques:

```
summary(data_suivi.active)
```

```
##          motivation_certificat          trouver_emploi
## D'accord          :27      D'accord          :23
## Désaccord         :25      Désaccord         :30
## Neutre            :17      Neutre            :32
## Tout à fait d'acord :30      Tout à fait d'acord : 6
## Tout à fait en désaccord:16      Tout à fait en désaccord:24
##
##          vacances_etude mooc_inclu_etude edt_chargee      videos_cours
## Les deux          :58      Non:25          Non:47      Bien          :52
## Les vacances: 9      Oui:90          Oui:68      Mauvais          : 3
## Vos études       :48
##
##
##
##          transcription      diapositives_cours      forum_discussion
## Bien          :52      Bien          :56      Bien          :42
## Mauvais        : 6      Mauvais        : 4      Mauvais        : 9
## Moyen          :21      Moyen          :31      Moyen          :44
```

```

## Très bien :35 Très bien :23 Très bien :14
## Très mauvais: 1 Très mauvais: 1 Très mauvais: 6
##
## examens_exercices accompagnement_pedagogique aides_financiere
## Bien :59 Bien :41 Non:91
## Mauvais : 2 Mauvais :15 Oui:24
## Moyen :31 Moyen :36
## Très bien :22 Très bien :15
## Très mauvais: 1 Très mauvais: 8
##
## outils_exercice sciences_donnees informatique
## Non:20 Inintéressant : 8 Inintéressant : 4
## Oui:95 Intéressant :87 Intéressant :91
## Pas très intéressant:20 Pas très intéressant:20
##
##
## sciences_physiques business
## Inintéressant :21 Inintéressant :21
## Intéressant :49 Intéressant :56
## Pas très intéressant:45 Pas très intéressant:38
##
##
## arts langues
## Inintéressant :37 Inintéressant :35
## Intéressant :48 Intéressant :50
## Pas très intéressant:30 Pas très intéressant:30
##
##
## sciences_sociales est_etudiant niveau_actuel
## Inintéressant :38 Non: 2 BAC +2 : 7
## Intéressant :38 Oui:113 BAC +3 :20
## Pas très intéressant:39 BAC +4 :36
## BAC +5 :45
## BAC +6 ou plus: 4
## NA's : 3
##
## formation background
## FIG :97 Autre école: 4
## FIL : 4 DUT : 4
## FIP : 3 Prépa :95
## Master de recherche: 1 Université :12
## Master of science : 7
## Master spécialisé : 3

```

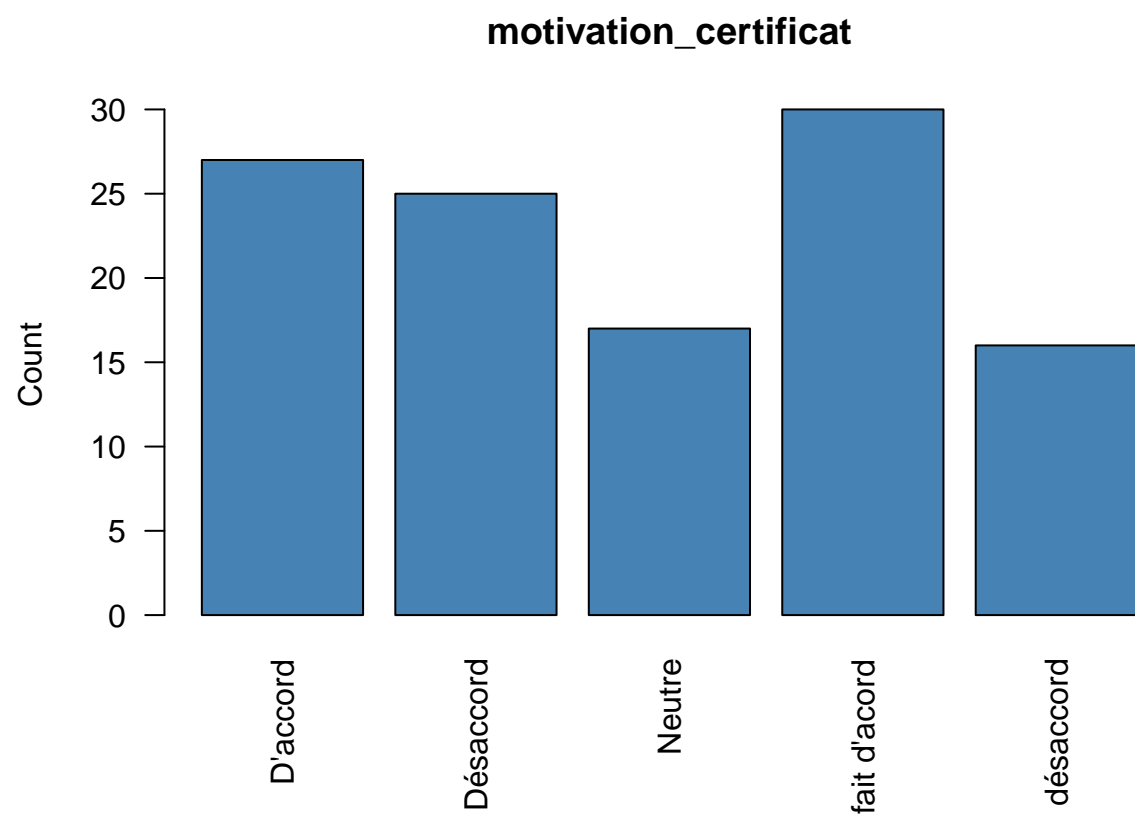
La fonctions `summary()` renvoient la taille des catégories des variables.

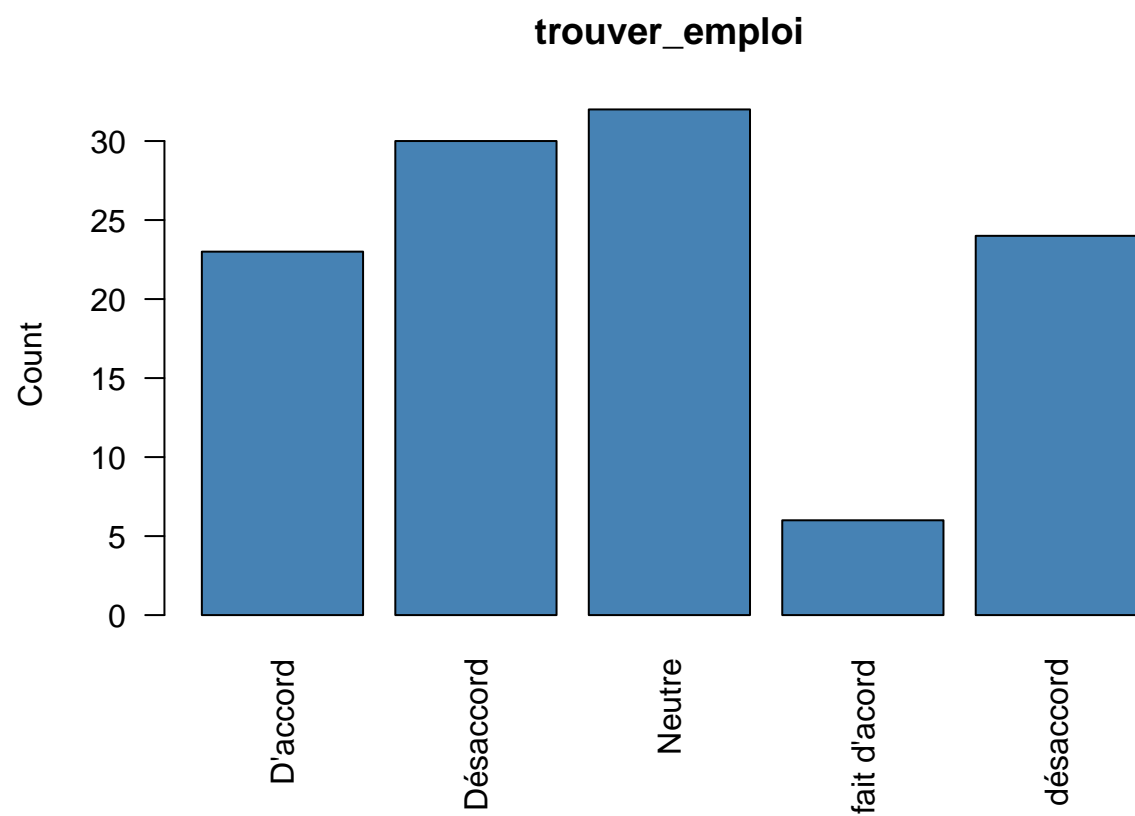
Il est également possible de visualiser la fréquence des catégories des variables. Le code R ci-dessous, montre toutes les colonnes:

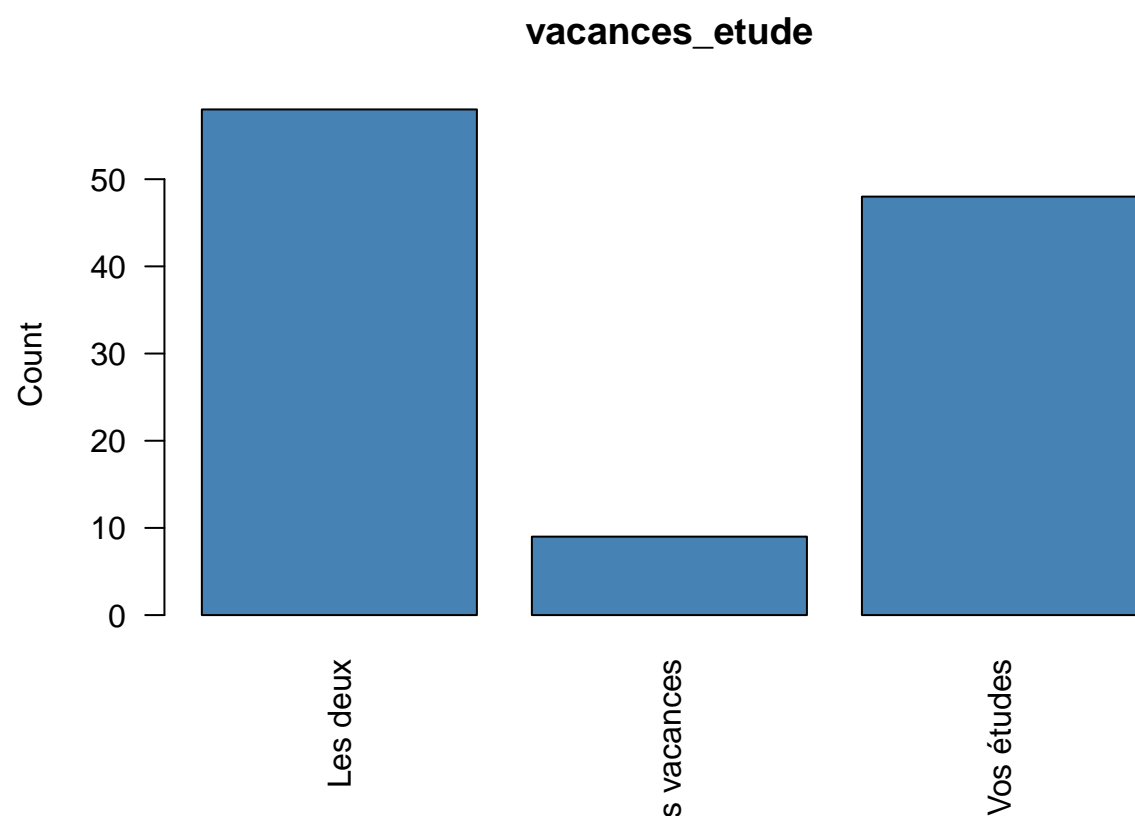
```

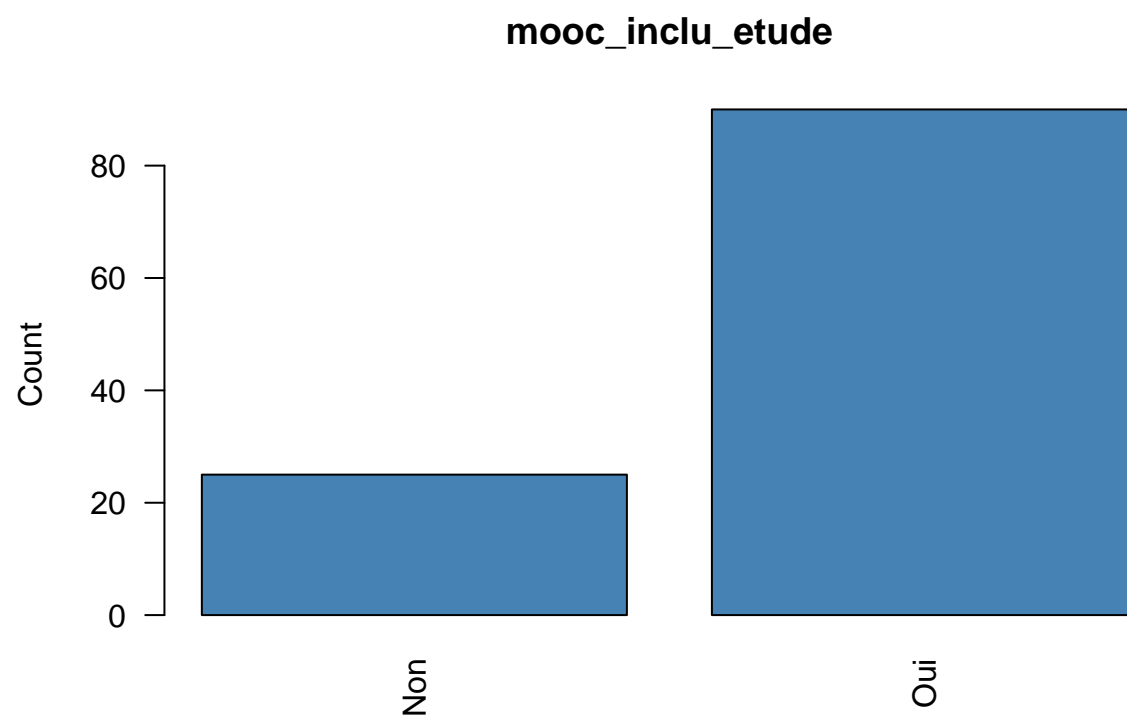
for (i in 1:24) {
  plot(data_suivi.active[,i], main = colnames(data_suivi.active)[i],
       ylab = "Count", col="steelblue", las = 2)
}

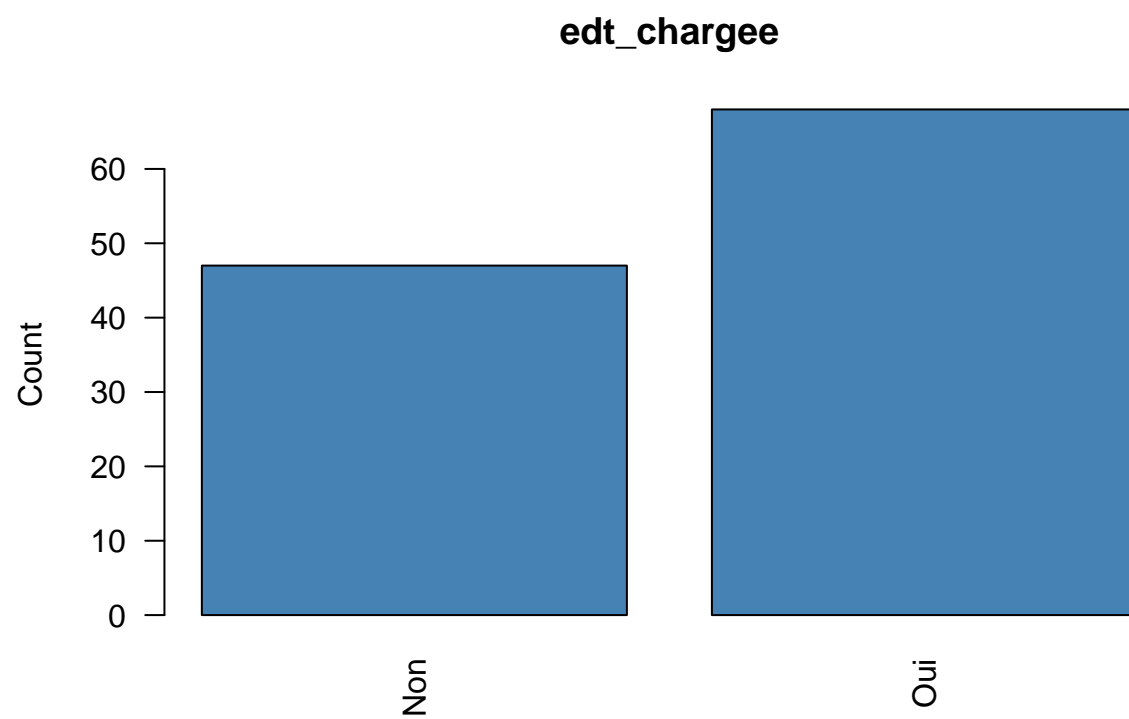
```

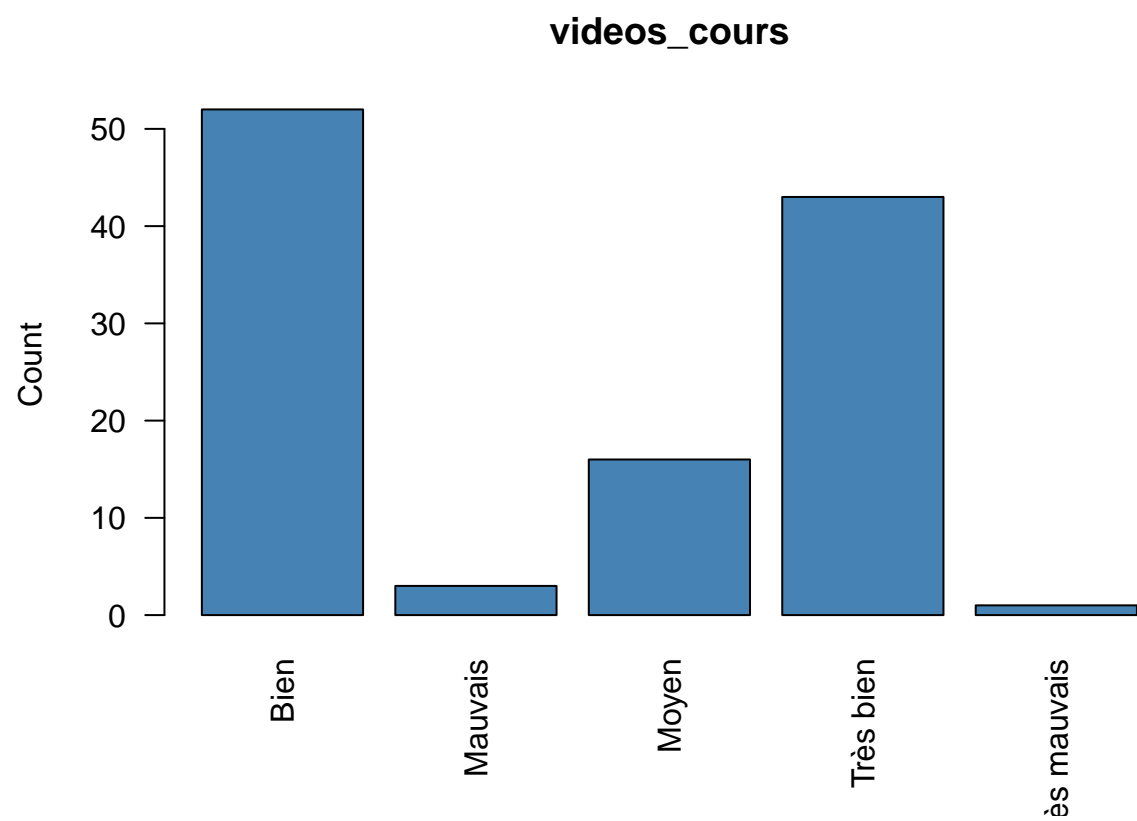


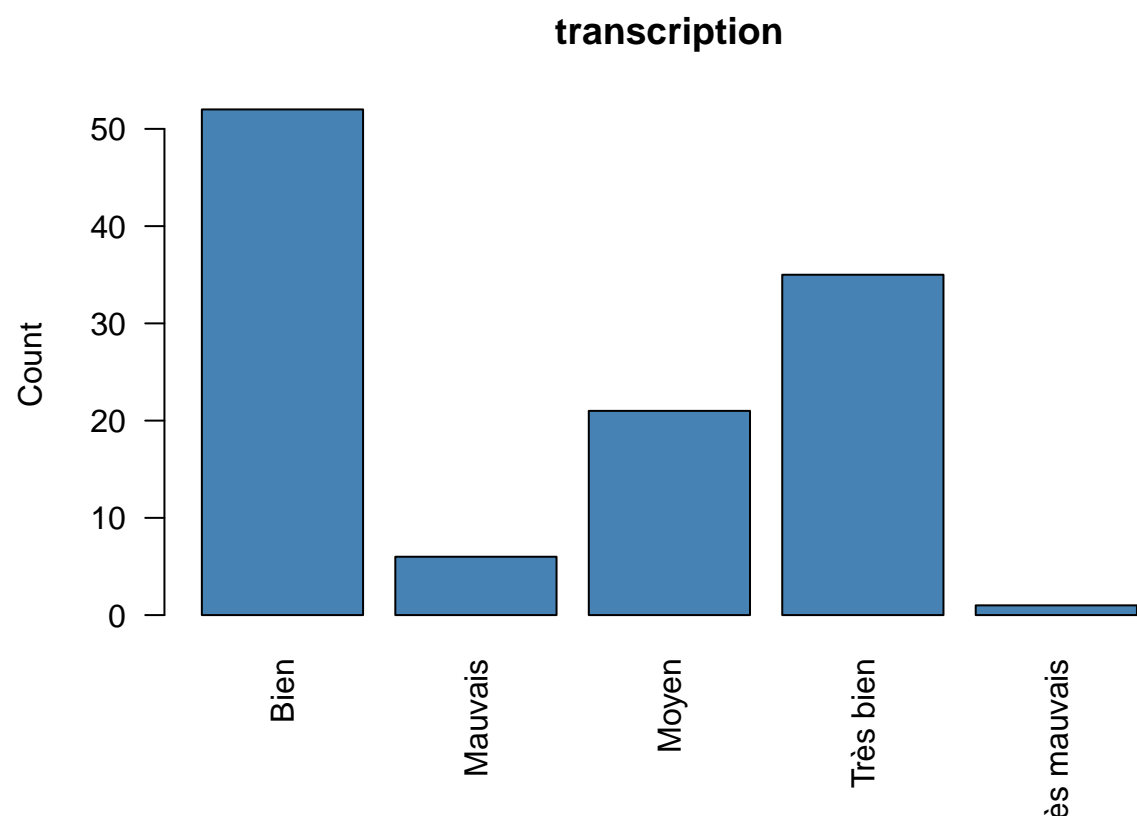


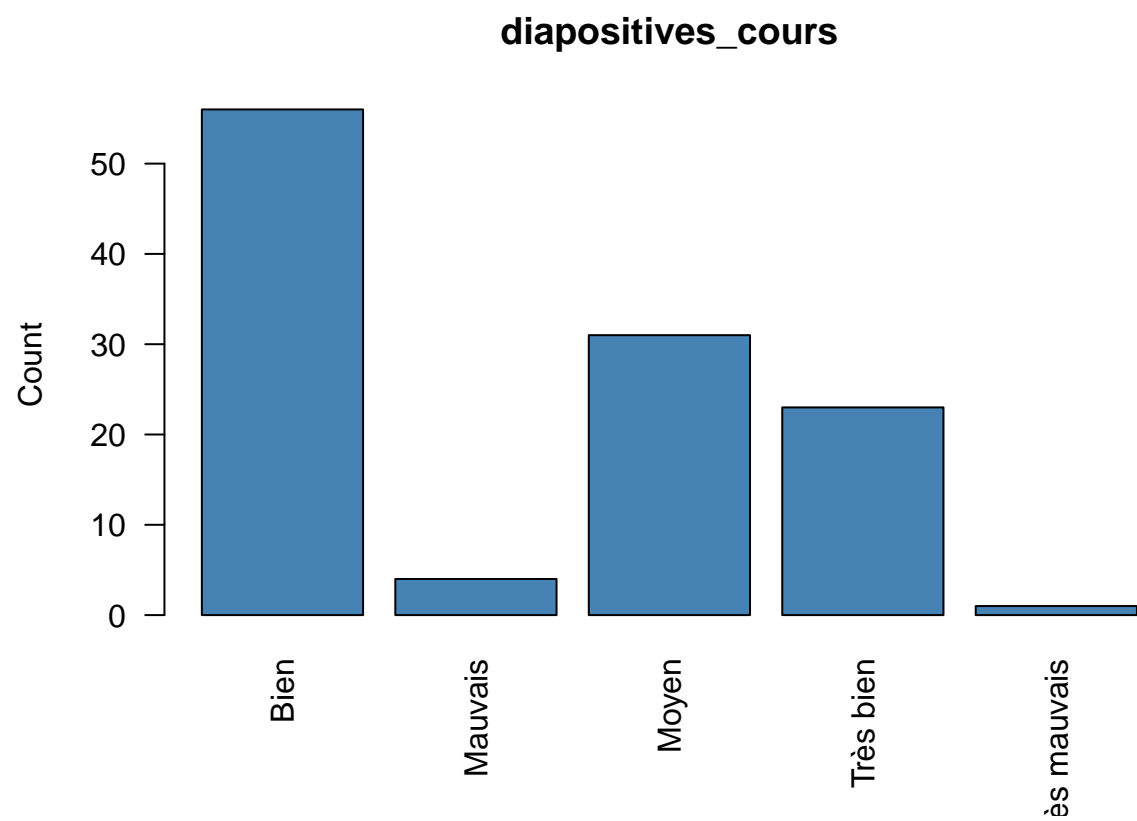


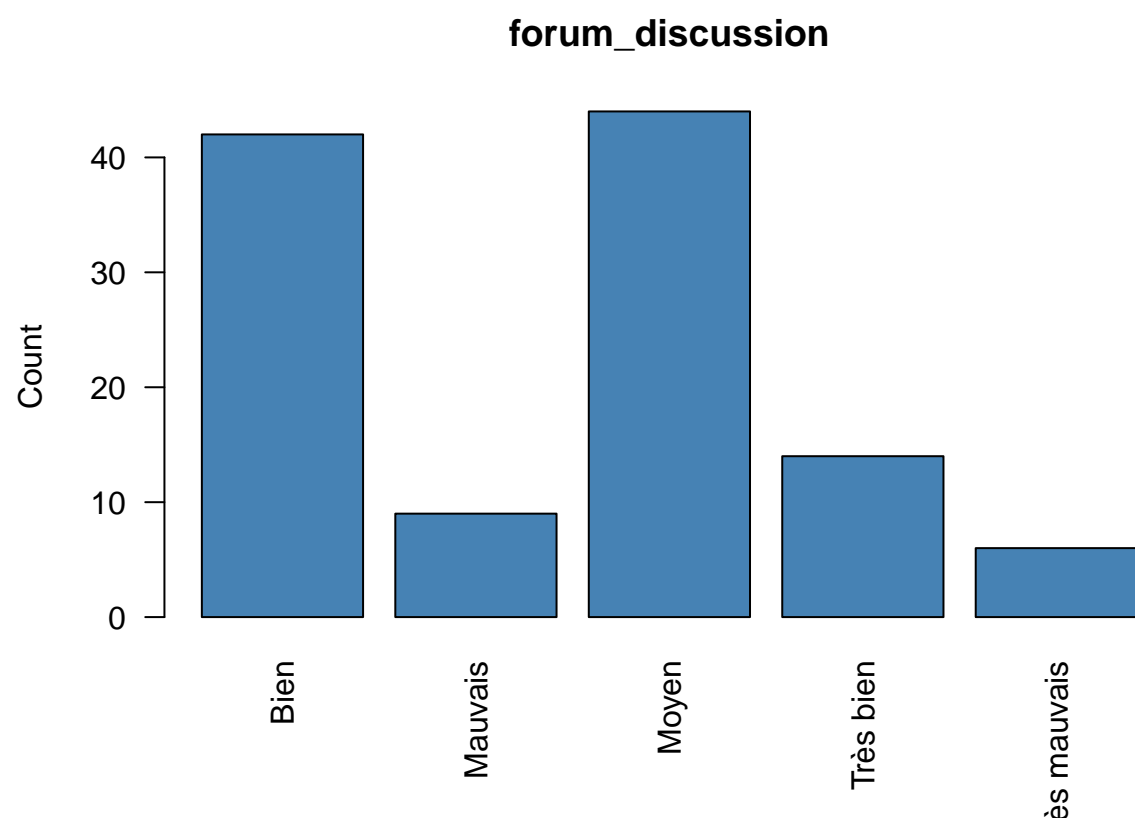


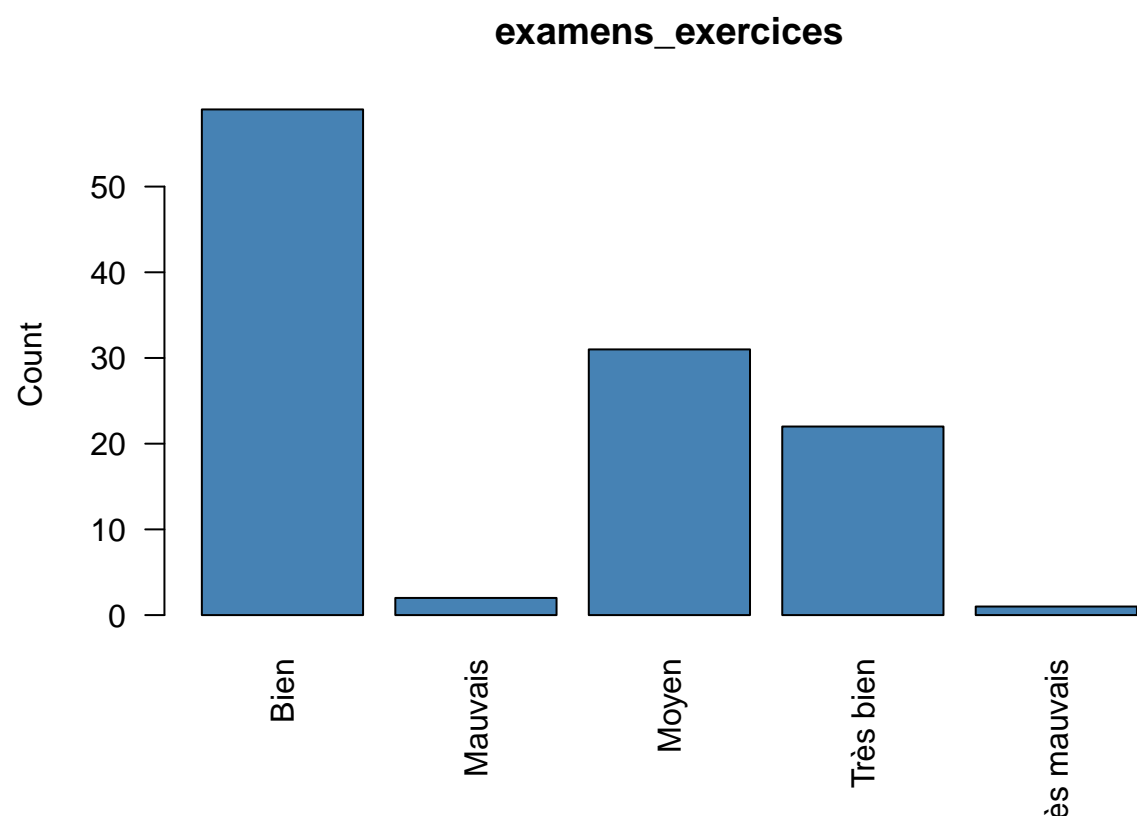


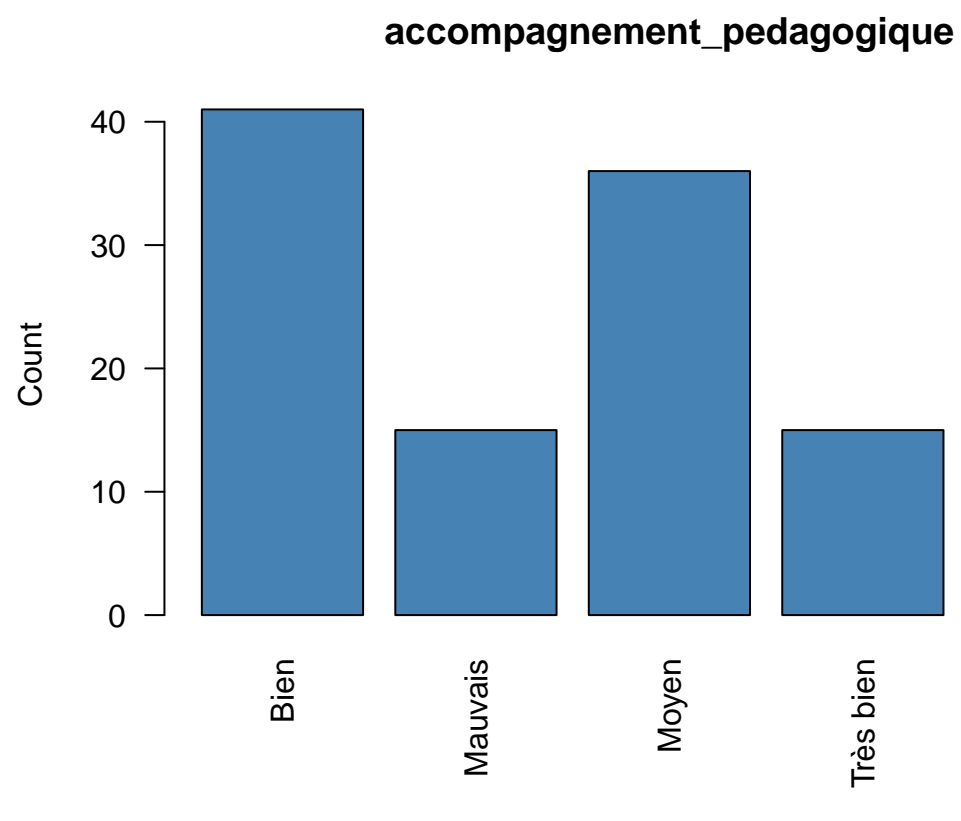


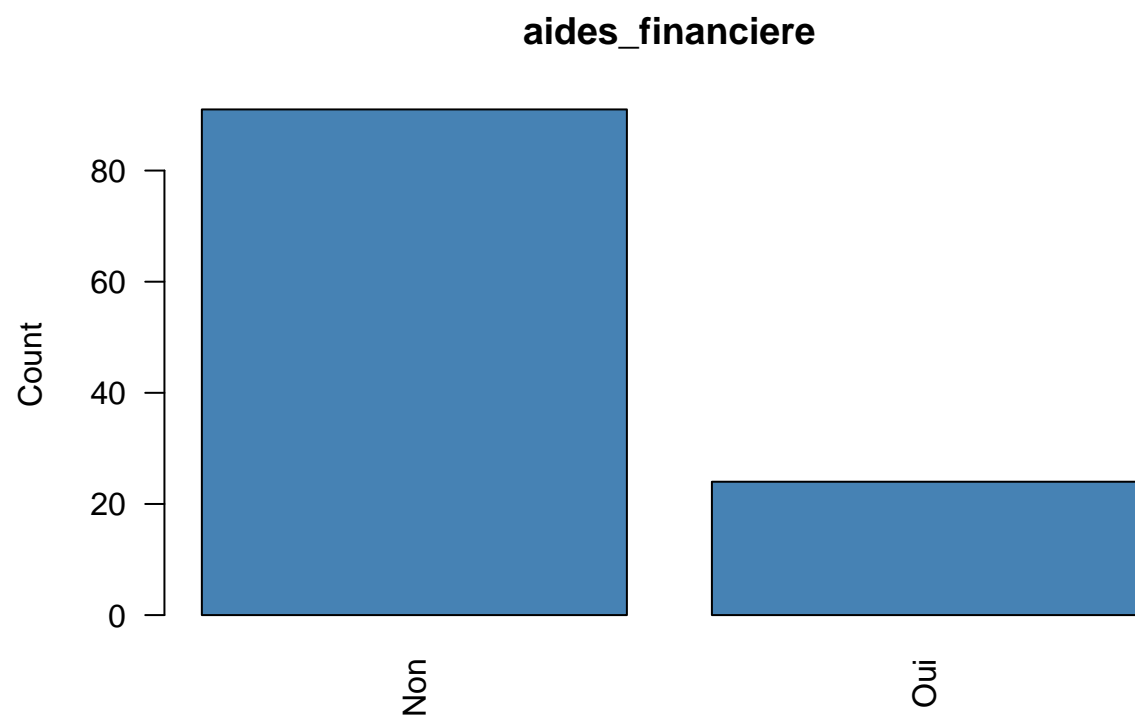


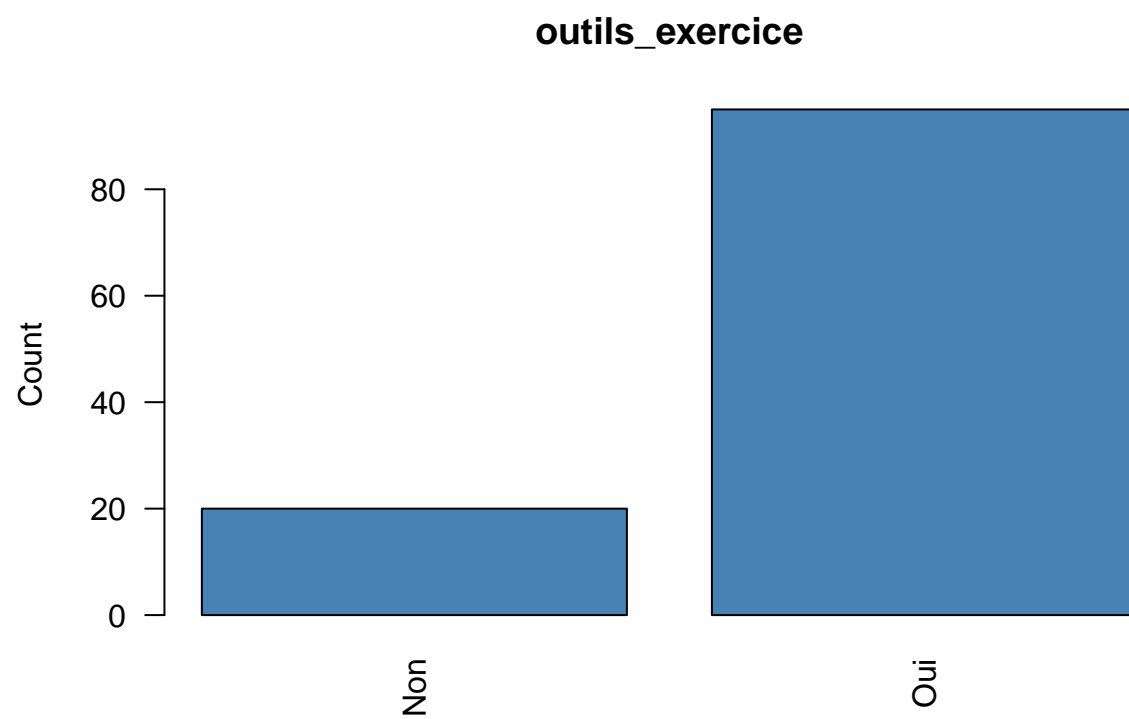


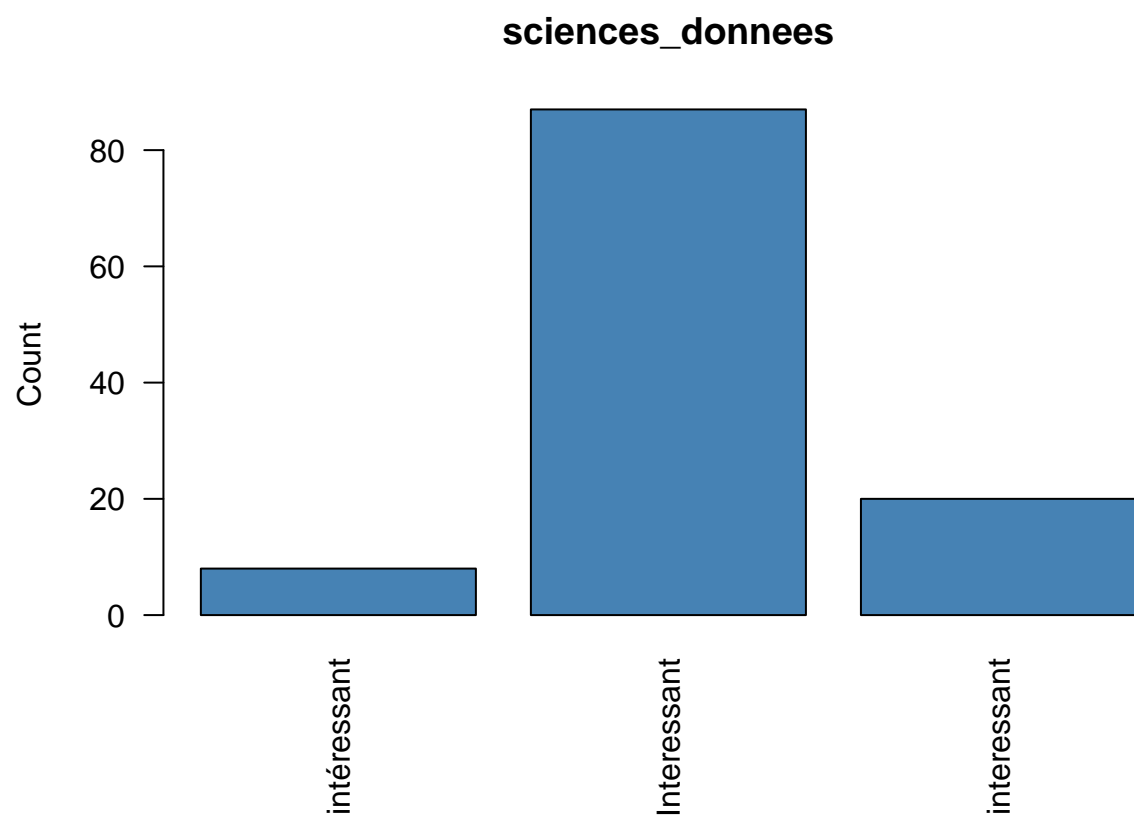


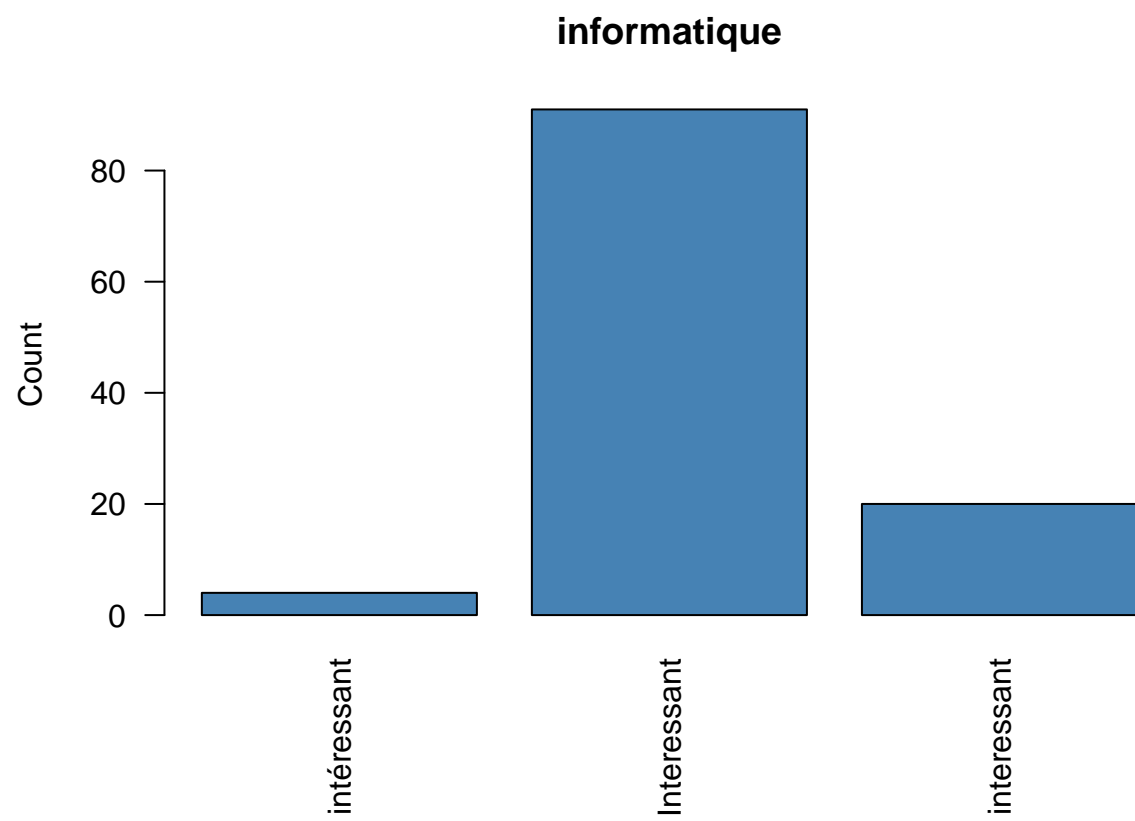


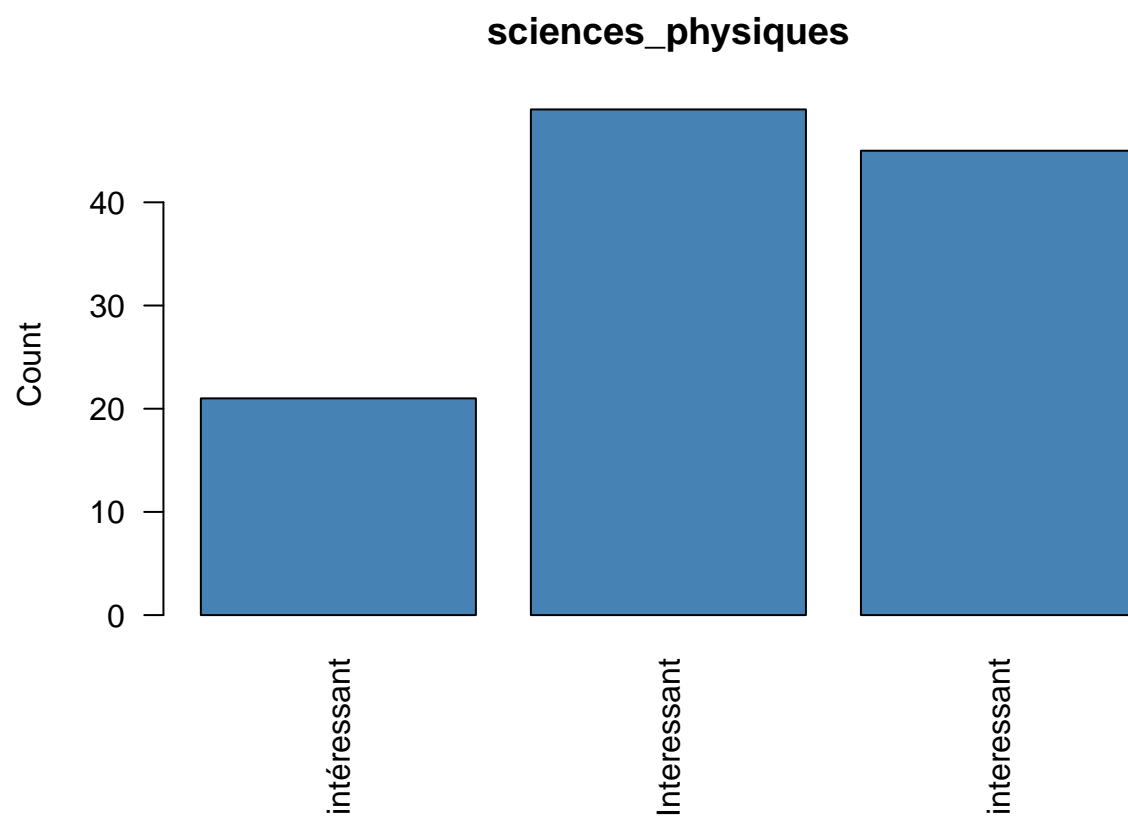


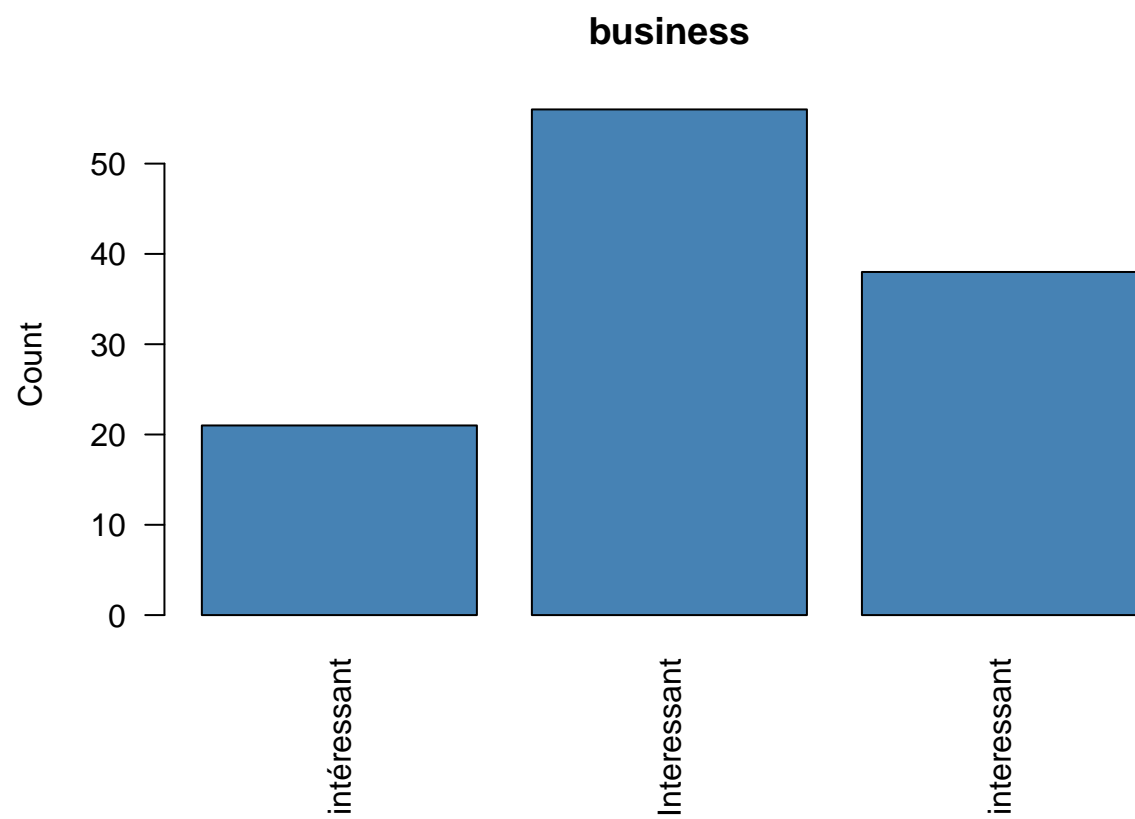


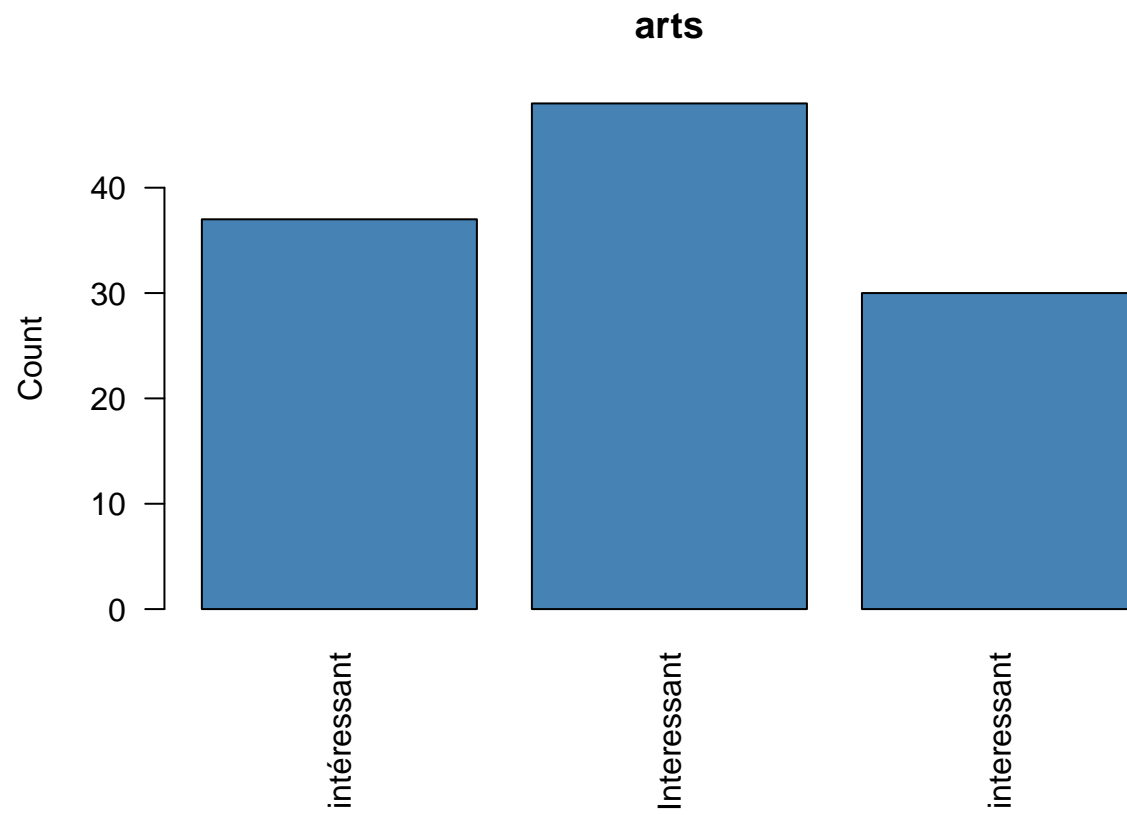


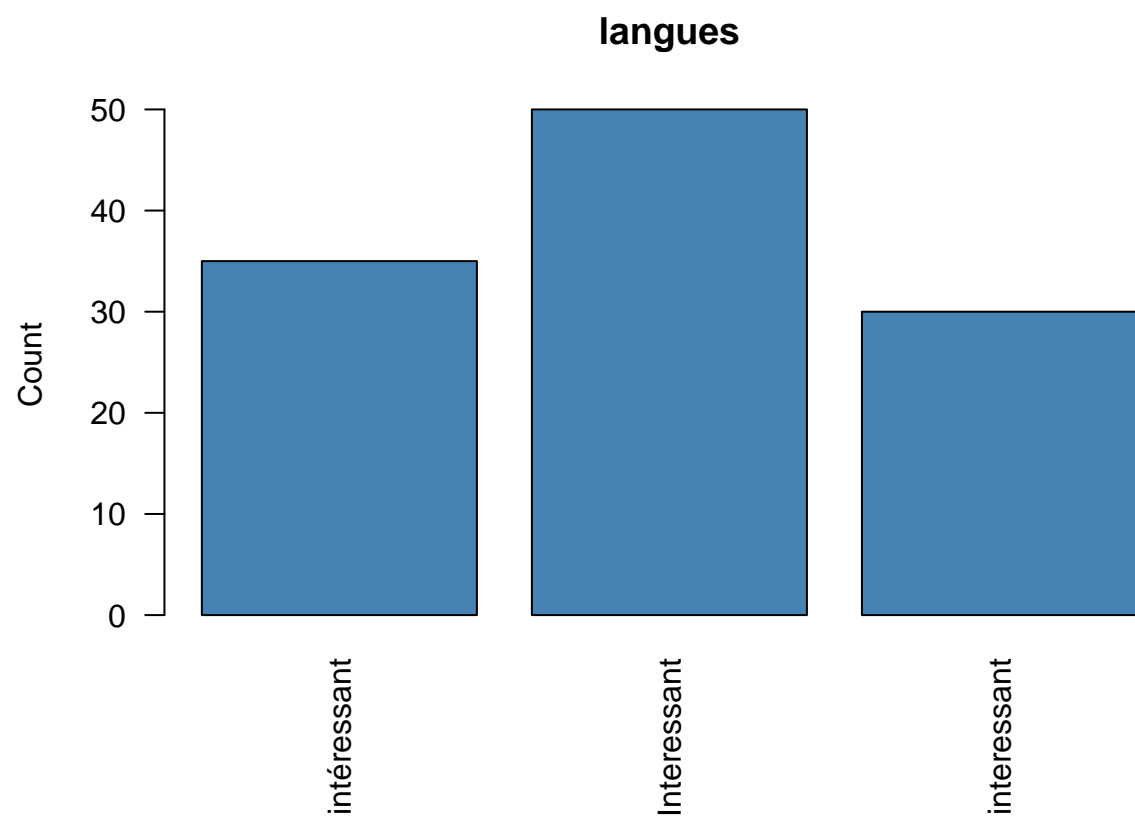


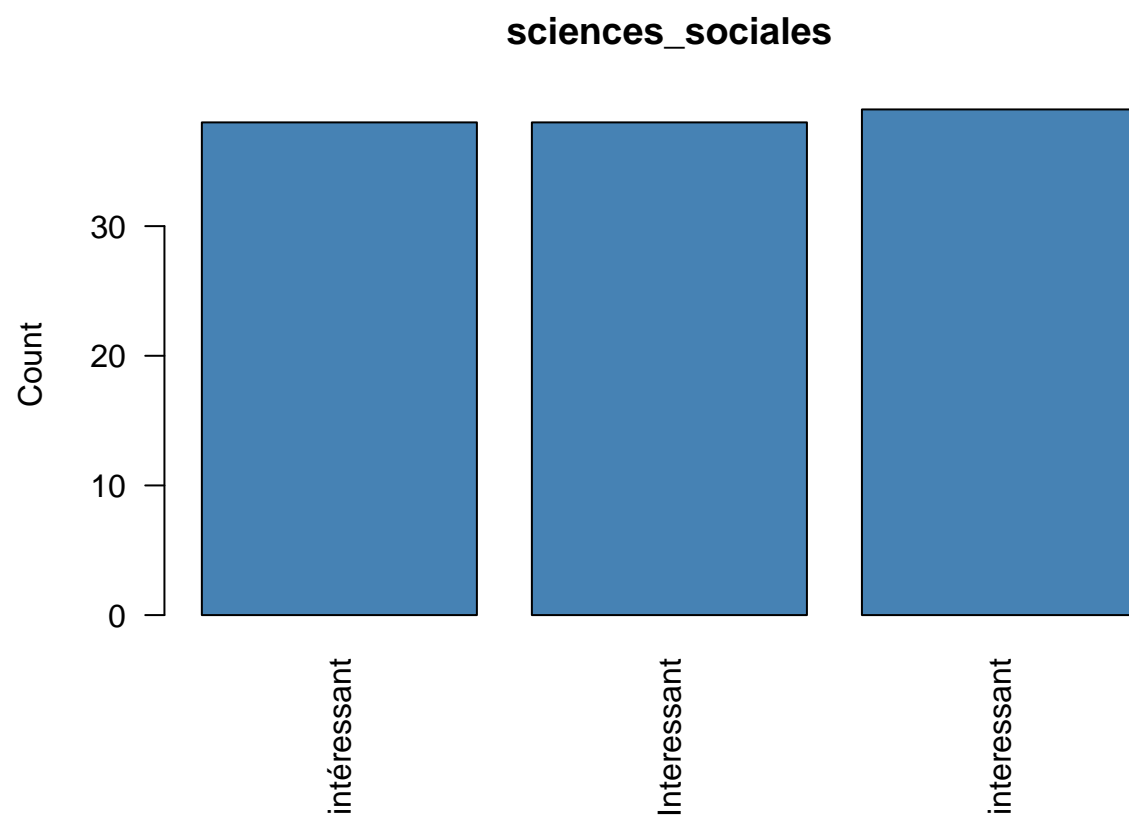


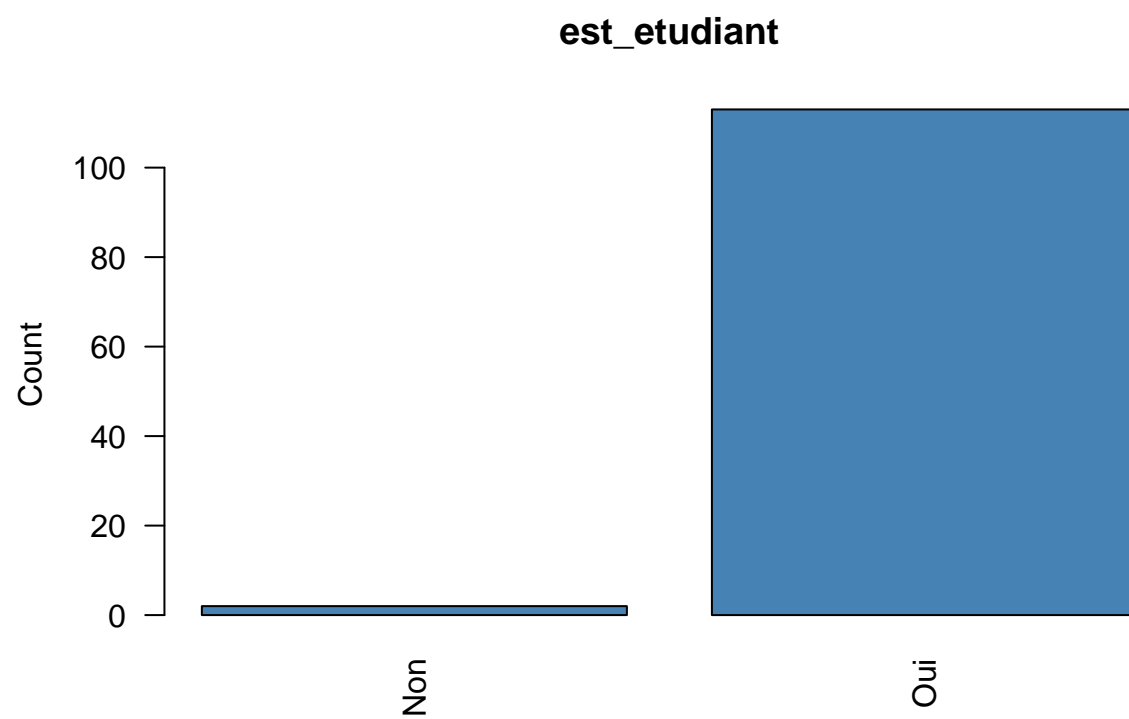


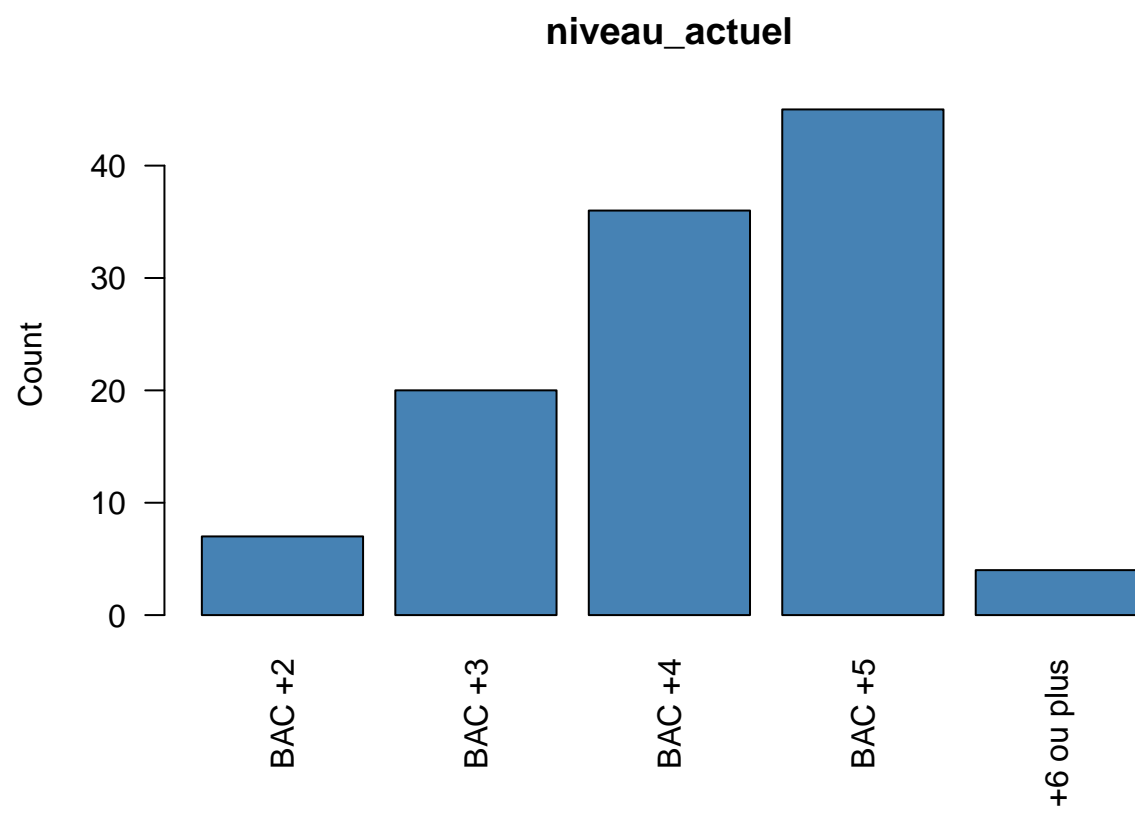


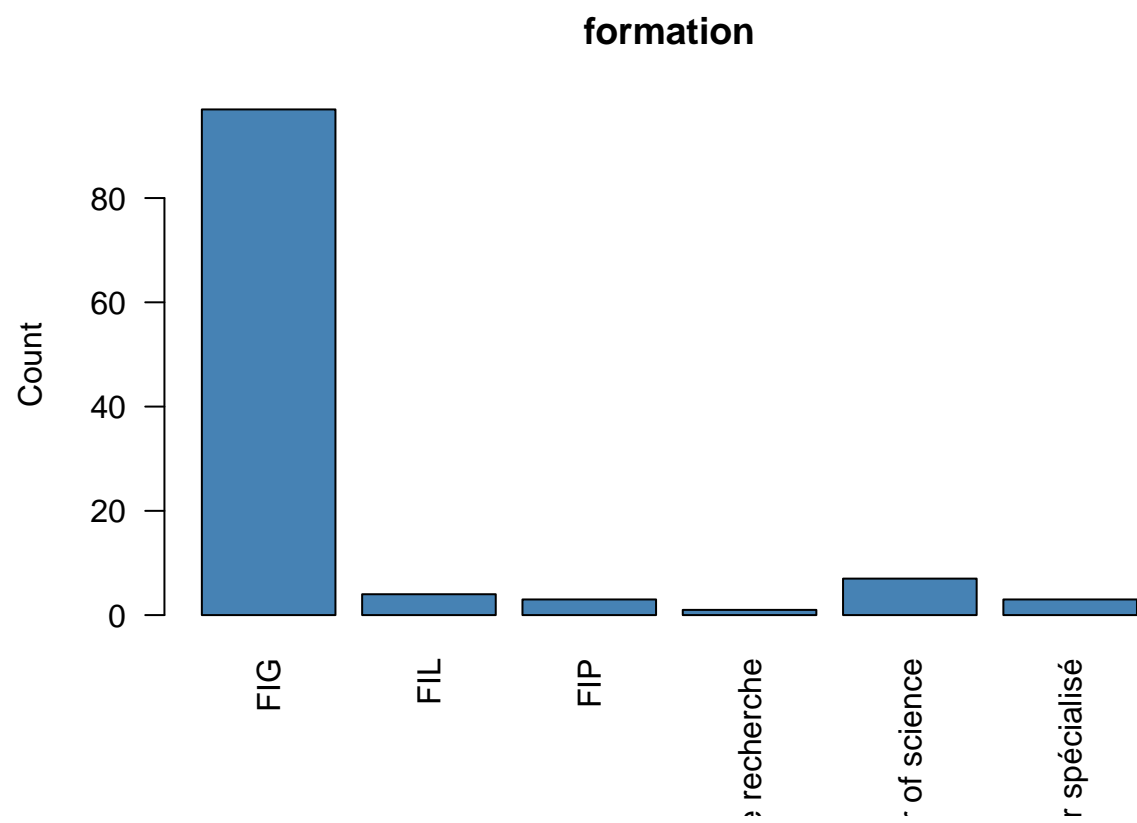


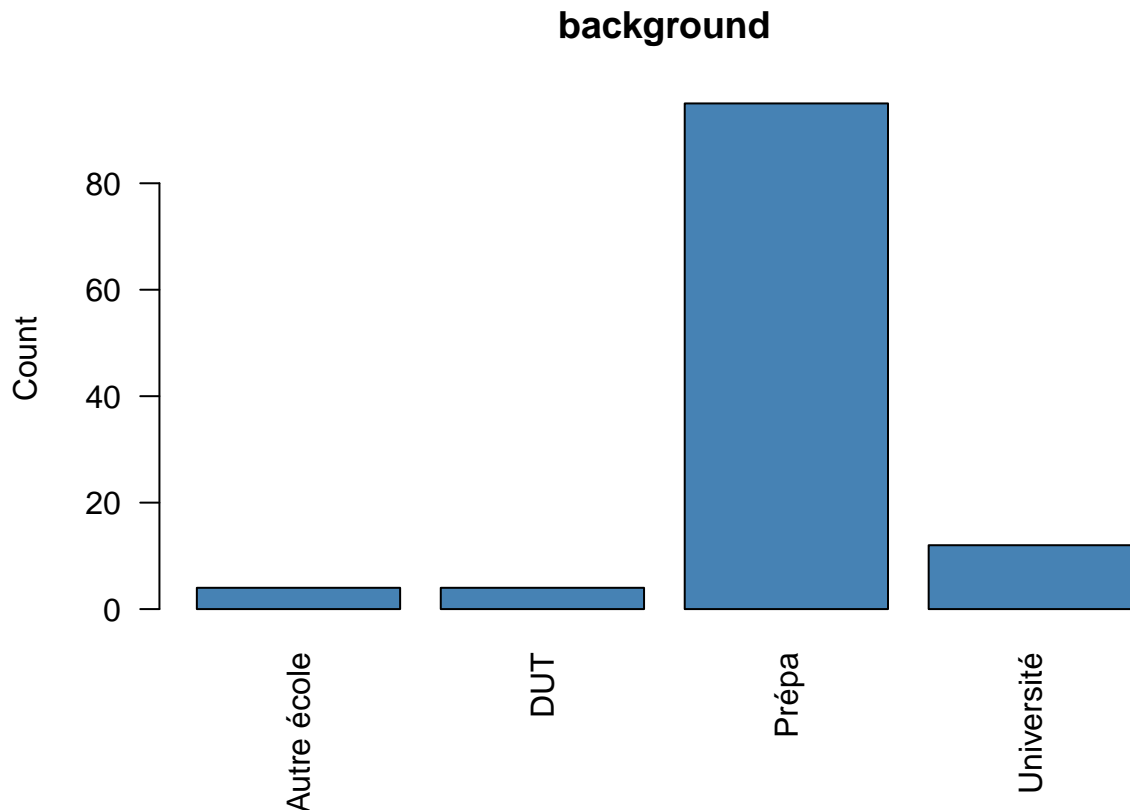












On va éliminer les variables suivantes car elles présentent des catégories à très faible fréquence, chose qui peut fausser l'analyse : `est_etudiant` et `formation`.

```
data_suivi.active <- data_suivi[, c( "motivation_certificat" , "trouver_emploi",
  "vacances_etude" , "mooc_inclu_etude" , "edt_chargee" ,
  "videos_cours" , "transcription" , "diapositives_cours" ,
  "forum_discussion" , "examens_exercices" , "accompagnement_pedagogique", "aides_financieres",
  "outils_exercice" , "sciences_donnees" , "informatique" ,
  "sciences_physiques" , "business" , "arts" , "langues" ,
  "sciences_sociales" ,
  "niveau_actuel" , "background" )]
```

code R pour l'ACM :

Dans le code R ci-dessous, l'ACM est effectuée uniquement sur les individus / variables actifs:

```
res.mca <- MCA (data_suivi.active, graph = FALSE,)
```

Le résultat de l'ACM est une liste comprenant:

```
print(res.mca)
```

```
## **Results of the Multiple Correspondence Analysis (MCA)**
## The analysis was performed on 115 individuals, described by 22 variables
## *The results are available in the following objects:
##
##   name          description
```

```
## 1 "$eig"          "eigenvalues"
## 2 "$var"          "results for the variables"
## 3 "$var$coord"    "coord. of the categories"
## 4 "$var$cos2"     "cos2 for the categories"
## 5 "$var$contrib"  "contributions of the categories"
## 6 "$var$v.test"   "v-test for the categories"
## 7 "$ind"          "results for the individuals"
## 8 "$ind$coord"    "coord. for the individuals"
## 9 "$ind$cos2"     "cos2 for the individuals"
## 10 "$ind$contrib" "contributions of the individuals"
## 11 "$call"         "intermediate results"
## 12 "$call$marge.col" "weights of columns"
## 13 "$call$marge.li" "weights of rows"
```

L'objet créé avec la fonction `MCA()` contient de nombreuses informations trouvées dans de nombreuses listes et matrices différentes. Ces valeurs sont décrites dans la section suivante.

Visualisation et interprétation :

La proportion des variances retenues par les différentes dimensions (axes) peut être extraite à l'aide de la fonction `get_eigenvalue()` [factoextra package] comme suit:

```
eig.val <- get_eigenvalue(res.mca)
eig.val
```

##		eigenvalue	variance.percent	cumulative.variance.percent
##	Dim.1	2.244256e-01	8.228939e+00	8.228939
##	Dim.2	1.845199e-01	6.765729e+00	14.994668
##	Dim.3	1.469717e-01	5.388964e+00	20.383632
##	Dim.4	1.202204e-01	4.408082e+00	24.791714
##	Dim.5	1.135142e-01	4.162189e+00	28.953903
##	Dim.6	9.913870e-02	3.635086e+00	32.588988
##	Dim.7	9.358436e-02	3.431426e+00	36.020415
##	Dim.8	9.175658e-02	3.364408e+00	39.384823
##	Dim.9	8.949087e-02	3.281332e+00	42.666155
##	Dim.10	8.504986e-02	3.118495e+00	45.784649
##	Dim.11	8.049434e-02	2.951459e+00	48.736109
##	Dim.12	7.510969e-02	2.754022e+00	51.490131
##	Dim.13	7.095637e-02	2.601733e+00	54.091864
##	Dim.14	6.901848e-02	2.530678e+00	56.622542
##	Dim.15	6.779297e-02	2.485742e+00	59.108284
##	Dim.16	6.147105e-02	2.253938e+00	61.362222
##	Dim.17	5.942392e-02	2.178877e+00	63.541099
##	Dim.18	5.705544e-02	2.092033e+00	65.633132
##	Dim.19	5.458862e-02	2.001583e+00	67.634715
##	Dim.20	5.086812e-02	1.865165e+00	69.499880
##	Dim.21	4.901492e-02	1.797214e+00	71.297094
##	Dim.22	4.711864e-02	1.727684e+00	73.024777
##	Dim.23	4.587909e-02	1.682233e+00	74.707010
##	Dim.24	4.285540e-02	1.571365e+00	76.278375
##	Dim.25	4.207825e-02	1.542869e+00	77.821244
##	Dim.26	4.084900e-02	1.497797e+00	79.319041
##	Dim.27	3.896778e-02	1.428818e+00	80.747859
##	Dim.28	3.734587e-02	1.369349e+00	82.117208

## Dim.29	3.390000e-02	1.243000e+00	83.360208
## Dim.30	3.129339e-02	1.147424e+00	84.507632
## Dim.31	2.989750e-02	1.096242e+00	85.603874
## Dim.32	2.882906e-02	1.057065e+00	86.660939
## Dim.33	2.810839e-02	1.030641e+00	87.691580
## Dim.34	2.543766e-02	9.327140e-01	88.624294
## Dim.35	2.482764e-02	9.103467e-01	89.534641
## Dim.36	2.282658e-02	8.369747e-01	90.371615
## Dim.37	2.259302e-02	8.284107e-01	91.200026
## Dim.38	2.131029e-02	7.813772e-01	91.981403
## Dim.39	1.985776e-02	7.281179e-01	92.709521
## Dim.40	1.955573e-02	7.170433e-01	93.426565
## Dim.41	1.906460e-02	6.990352e-01	94.125600
## Dim.42	1.782241e-02	6.534885e-01	94.779088
## Dim.43	1.622303e-02	5.948444e-01	95.373933
## Dim.44	1.516499e-02	5.560498e-01	95.929982
## Dim.45	1.450214e-02	5.317451e-01	96.461728
## Dim.46	1.309846e-02	4.802770e-01	96.942005
## Dim.47	1.241891e-02	4.553600e-01	97.397365
## Dim.48	1.145991e-02	4.201967e-01	97.817561
## Dim.49	1.057649e-02	3.878045e-01	98.205366
## Dim.50	9.633116e-03	3.532142e-01	98.558580
## Dim.51	8.598828e-03	3.152904e-01	98.873870
## Dim.52	7.378790e-03	2.705556e-01	99.144426
## Dim.53	7.149648e-03	2.621538e-01	99.406580
## Dim.54	5.142360e-03	1.885532e-01	99.595133
## Dim.55	4.229992e-03	1.550997e-01	99.750233
## Dim.56	3.859123e-03	1.415012e-01	99.891734
## Dim.57	2.952714e-03	1.082662e-01	100.000000
## Dim.58	4.566735e-32	1.674470e-30	100.000000
## Dim.59	8.584285e-33	3.147571e-31	100.000000
## Dim.60	8.073256e-33	2.960194e-31	100.000000

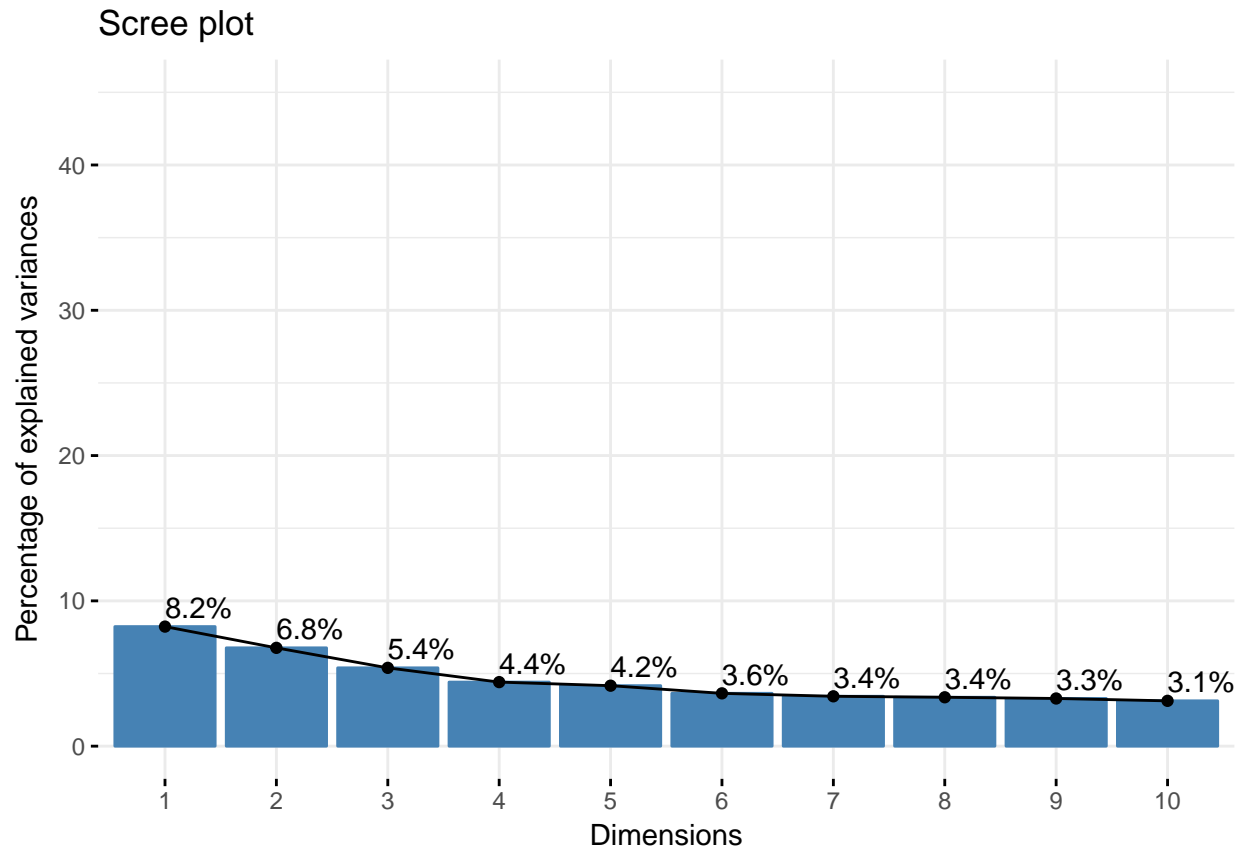
Les dimensions sont ordonnées de manière décroissante et listées en fonction de la quantité de variance expliquée. La dimension 1 explique la plus grande variance, suivie de la dimension 2 et ainsi de suite.

Le pourcentage cumulé expliqué est obtenu en ajoutant les proportions successives de variances expliquées pour obtenir le total courant. Par exemple, 8.22% plus 6.76% est égal à 14.99%, et ainsi de suite. Environ 45.78% de la variance totale est expliquée par les dix premières dimensions.

Les valeurs propres peuvent être utilisées pour déterminer le nombre d'axes à retenir. Il n'y a pas de «règle générale» pour choisir le nombre de dimensions à conserver pour l'interprétation des données. Par exemple, Nous optons pour 50.16% de la variance totale expliquée, soit dix Composantes principales. D'ailleurs «la part de variance expliquée dans une ACM est très souvent faible, il n'est pas nécessaire d'avoir la même démarche de conservation des dimensions que pour une ACP. Ceci est purement calculatoire, à cause de la nature qualitative des variables : les contributions sont diluées entre les différentes modalités ce qui entraîne des parts d'inertie faibles.»(Romain Billot)

Une autre méthode pour déterminer le nombre de dimensions est de regarder le graphique des valeurs propres (scree plot), ordonnées de la plus grande à la plus petite valeur. Le nombre d'axes est déterminé par le point point, au-delà duquel les valeurs propres restantes sont toutes relativement petites et de tailles comparables. Pour visualiser les pourcentages de variances expliquées par chaque dimension de l'ACM, utilisez la fonction `fviz_eig()` ou `fviz_screplot()` [package `factoextra`]:

```
fviz_screplot (res.mca, addlabels = TRUE, ylim = c (0, 45))
```

Il est également possible de calculer une valeur propre moyenne au-dessus de laquelle l'axe doit être conservé dans le résultat.

Si les données étaient aléatoires, la valeur attendue de la valeur propre pour chaque axe serait $1 / (\text{nrow}(\text{data_suivi.active}) - 1) = 1/114 = 0.87\%$ en termes de lignes.

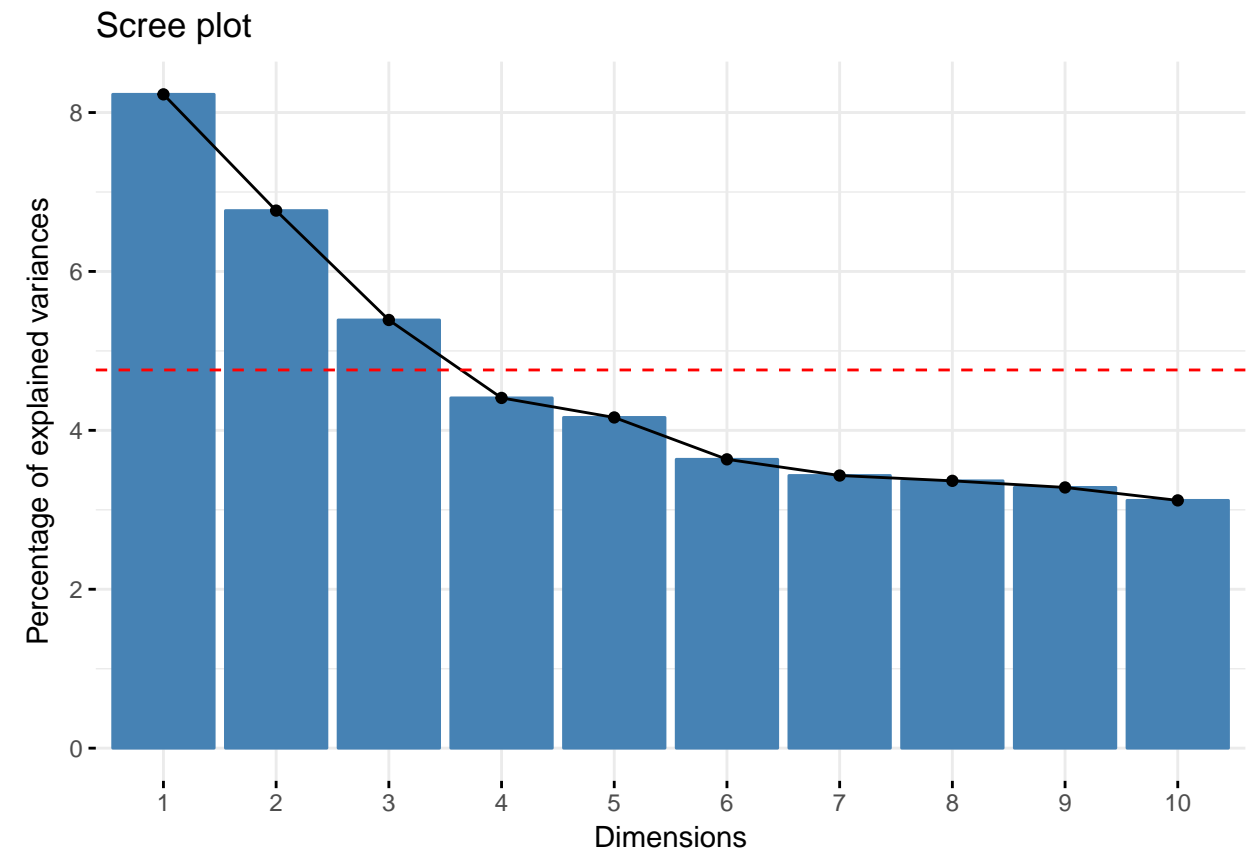
De même, l'axe moyen devrait représenter $1 / (\text{ncol}(\text{data_suivi.active}) - 1) = 1/21 = 4.76\%$ en termes de 22 colonnes.

Tout axe avec une contribution supérieure au maximum de ces deux pourcentages devrait être considéré comme important et inclus dans la solution pour l'interprétation des données.

Donc on doit retenir selon M. T. Bendixen 3 axes principales. Ces 3 axes ont pour inertie totale : 20.38% (une valeur pas vraiment acceptable). On note très bien que l'essentiel est de pouvoir jouer sur toutes les dimensions et avoir si possible des interprétation logique relative à l'étude effectuée.

Le code R ci-dessous, montre le graphique des valeurs propres avec une droite en pointillée rouge spécifiant la valeur propre moyenne:

```
fviz_screplot (res.mca) +
  geom_hline (yintercept = 4.76, linetype = 2, color = "red")
```



Graphique des variables :

Résultats :

La fonction `get_mca_var()` [factoextra] sert à extraire les résultats pour les catégories des variables. Cette fonction renvoie une liste contenant les coordonnées, les cos2 et les contributions des catégories:

```
var <- get_mca_var(res.mca)
var
```

```
## Multiple Correspondence Analysis Results for variables
## =====
##   Name      Description
## 1 "$coord"   "Coordinates for categories"
## 2 "$cos2"    "Cos2 for categories"
## 3 "$contrib" "contributions of categories"
```

Les composants de `get_mca_var()` peuvent être utilisés dans le graphique des variables comme suit:

`varcoord` : coordonnées des variables pour créer un nuage de points
`varcos2` : qualité de représentation des variables.
`var$contrib` : contributions (en pourcentage) des variables à la définition des dimensions.

Les différents composants peuvent être consultés comme suit:

```
# Coordonnées
print('les coordonnées :')
```

```
## [1] "les coordonnées :"
```

```
head(var$coord)
```

```
##                               Dim 1      Dim 2
## motivation_certificat_D'accord -0.2923891  0.1297577
## motivation_certificat_Désaccord  0.1944805 -0.4869087
## motivation_certificat_Neutre     -0.2328676  0.1111734
## motivation_certificat_Tout à fait d'acord  0.1619016  0.3849359
## motivation_certificat_Tout à fait en désaccord 0.1333872 -0.2980479
## trouver_emploi_D'accord         -0.4642665  0.7551930
##                               Dim 3      Dim 4
## motivation_certificat_D'accord  0.007642434  0.02648756
## motivation_certificat_Désaccord  0.158729347  0.53364470
## motivation_certificat_Neutre     -0.092547391 -0.88369553
## motivation_certificat_Tout à fait d'acord -0.147882569  0.14465643
## motivation_certificat_Tout à fait en désaccord 0.114700207 -0.21082190
## trouver_emploi_D'accord         0.138958130  0.27311295
##                               Dim 5
## motivation_certificat_D'accord -0.14103830
## motivation_certificat_Désaccord -0.09786537
## motivation_certificat_Neutre     -0.35473690
## motivation_certificat_Tout à fait d'acord  0.72074115
## motivation_certificat_Tout à fait en désaccord -0.58356494
## trouver_emploi_D'accord         0.53466642
```

```
# Cos2: qualité de représentation
```

```
print('la qualité de représentation :')
```

```
## [1] "la qualité de représentation :"
```

```
head(var$cos2)
```

```
##                               Dim 1      Dim 2
## motivation_certificat_D'accord  0.026230315  0.005165919
## motivation_certificat_Désaccord  0.010506291  0.065855582
## motivation_certificat_Neutre     0.009406782  0.002143998
## motivation_certificat_Tout à fait d'acord  0.009251344  0.052297293
## motivation_certificat_Tout à fait en désaccord 0.002875498  0.014356772
## trouver_emploi_D'accord         0.053885855  0.142579100
##                               Dim 3      Dim 4
## motivation_certificat_D'accord  1.792027e-05  0.0002152608
## motivation_certificat_Désaccord  6.998613e-03  0.0791046281
## motivation_certificat_Neutre     1.485769e-03  0.1354653319
## motivation_certificat_Tout à fait d'acord  7.718560e-03  0.0073854646
## motivation_certificat_Tout à fait en désaccord 2.126244e-03  0.0071831715
## trouver_emploi_D'accord         4.827340e-03  0.0186476705
##                               Dim 5
## motivation_certificat_D'accord  0.006103167
## motivation_certificat_Désaccord  0.002660453
## motivation_certificat_Neutre     0.021829087
## motivation_certificat_Tout à fait d'acord  0.183341580
## motivation_certificat_Tout à fait en désaccord 0.055038067
## trouver_emploi_D'accord         0.071467044
```

```
# Contributions aux axes
```

```
print('la contribution des variables :')
```

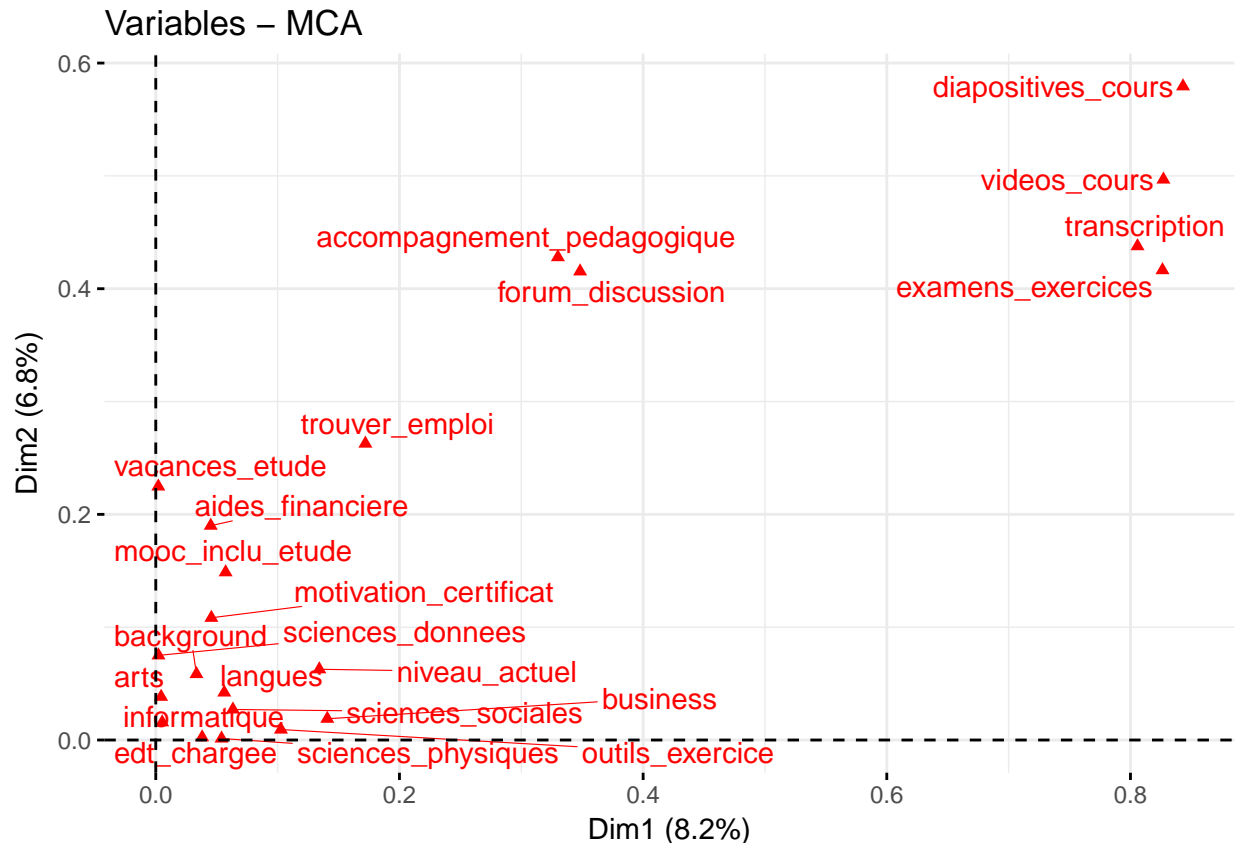
```
## [1] "la contribution des variables :"
```

```
head(var$contrib)
```

```
##
##               Dim 1      Dim 2
## motivation_certificat_D'accord 0.40653058 0.09737929
## motivation_certificat_Désaccord 0.16653249 1.26961319
## motivation_certificat_Neutre    0.16235820 0.04500769
## motivation_certificat_Tout à fait d'acord 0.13849396 0.95221378
## motivation_certificat_Tout à fait en désaccord 0.05013665 0.30445866
## trouver_emploi_D'accord        0.87311141 2.80982982
##
##               Dim 3      Dim 4
## motivation_certificat_D'accord 0.0004241044 0.006228005
## motivation_certificat_Désaccord 0.1693948176 2.340699396
## motivation_certificat_Neutre    0.0391582188 4.364710887
## motivation_certificat_Tout à fait d'acord 0.1764415549 0.206394438
## motivation_certificat_Tout à fait en désaccord 0.0566101146 0.233804394
## trouver_emploi_D'accord        0.1194376928 0.564044871
##
##               Dim 5
## motivation_certificat_D'accord 0.18701094
## motivation_certificat_Désaccord 0.08337334
## motivation_certificat_Neutre    0.74488763
## motivation_certificat_Tout à fait d'acord 5.42636516
## motivation_certificat_Tout à fait en désaccord 1.89726278
## trouver_emploi_D'accord        2.28940575
```

Dans cette section, nous décrirons comment visualiser uniquement les catégories des variables. Ensuite, nous mettrons en évidence les catégories en fonction soit de leurs qualités de représentation, soit de leurs contributions aux dimensions. ###Corrélation entre les variables et les axes principaux : Pour visualiser la corrélation entre les variables et les axes principaux de l'ACM, tapez ceci:

```
fviz_mca_var (res.mca, choice = "mca.cor",
              repel = TRUE,
              ggtheme = theme_minimal ())
```



+ Le graphique ci-dessus permet d'identifier les variables les plus corrélées avec chaque axe. Les corrélations au carré entre les variables et les axes sont utilisées comme coordonnées.

- On constate que les variables Diapositives_cours, videos_cours, transcription_examens_exercices sont les plus corrélées avec la dimension 1. De même, les variables accompagnement_pedagogique et forum_discussion sont les plus corrélées avec la dimension 2.

Coordonnées des catégories des variables :

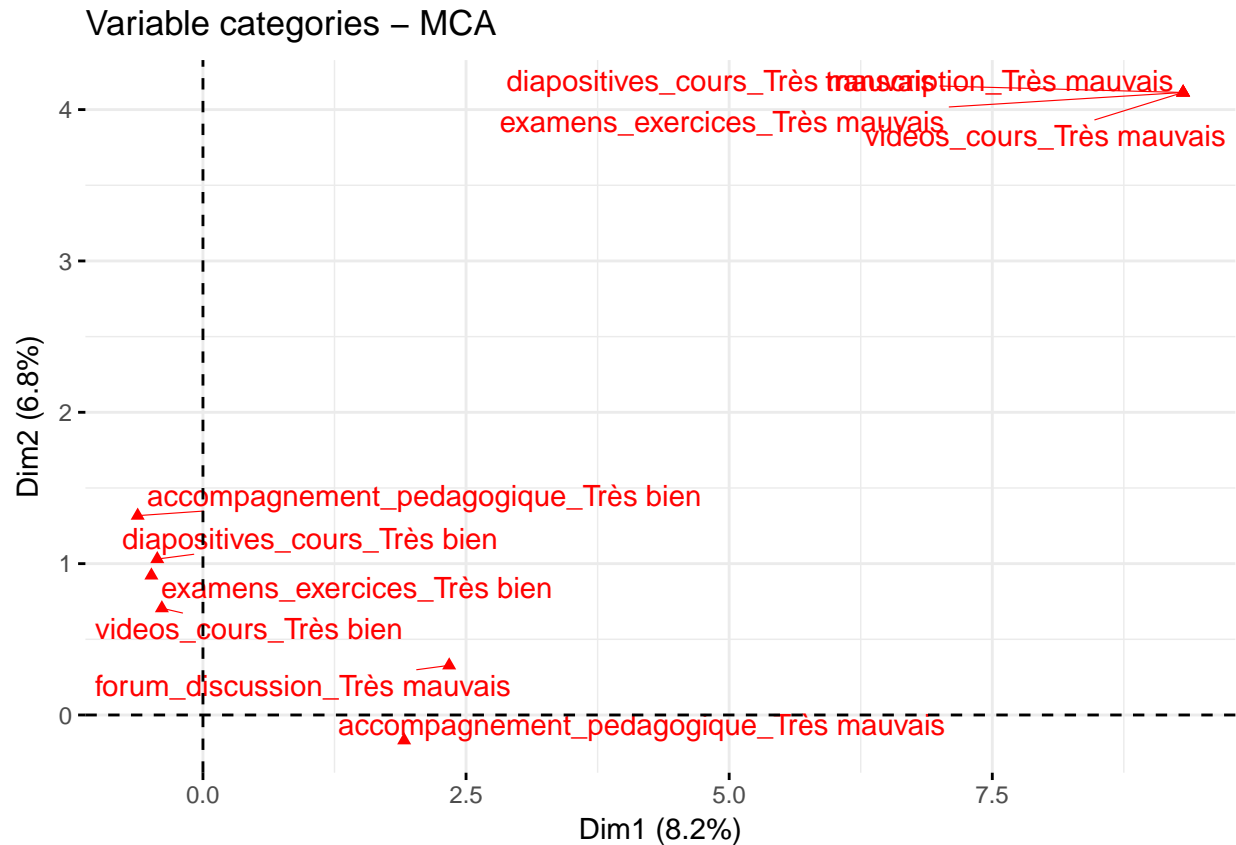
Le code R ci-dessous affiche les coordonnées de chacune des catégories des variables dans les dimensions (1, 2, 3, 4, et 5):

```
head(round(var$coord, 2), 4)
```

```
##               Dim 1 Dim 2 Dim 3 Dim 4 Dim 5
## motivation_certificat_D'accord -0.29  0.13  0.01  0.03 -0.14
## motivation_certificat_Désaccord  0.19 -0.49  0.16  0.53 -0.10
## motivation_certificat_Neutre -0.23  0.11 -0.09 -0.88 -0.35
## motivation_certificat_Tout à fait d'accord 0.16  0.38 -0.15  0.14  0.72
```

Utilisez la fonction `fviz_mca_var()` [factoextra] pour visualiser uniquement les catégories des variables:

```
fviz_mca_var(res.mca,
  repel = TRUE,
  axes=c(1,2),
  ggtheme = theme_minimal(), select.var = list(cos2=0.25))
```



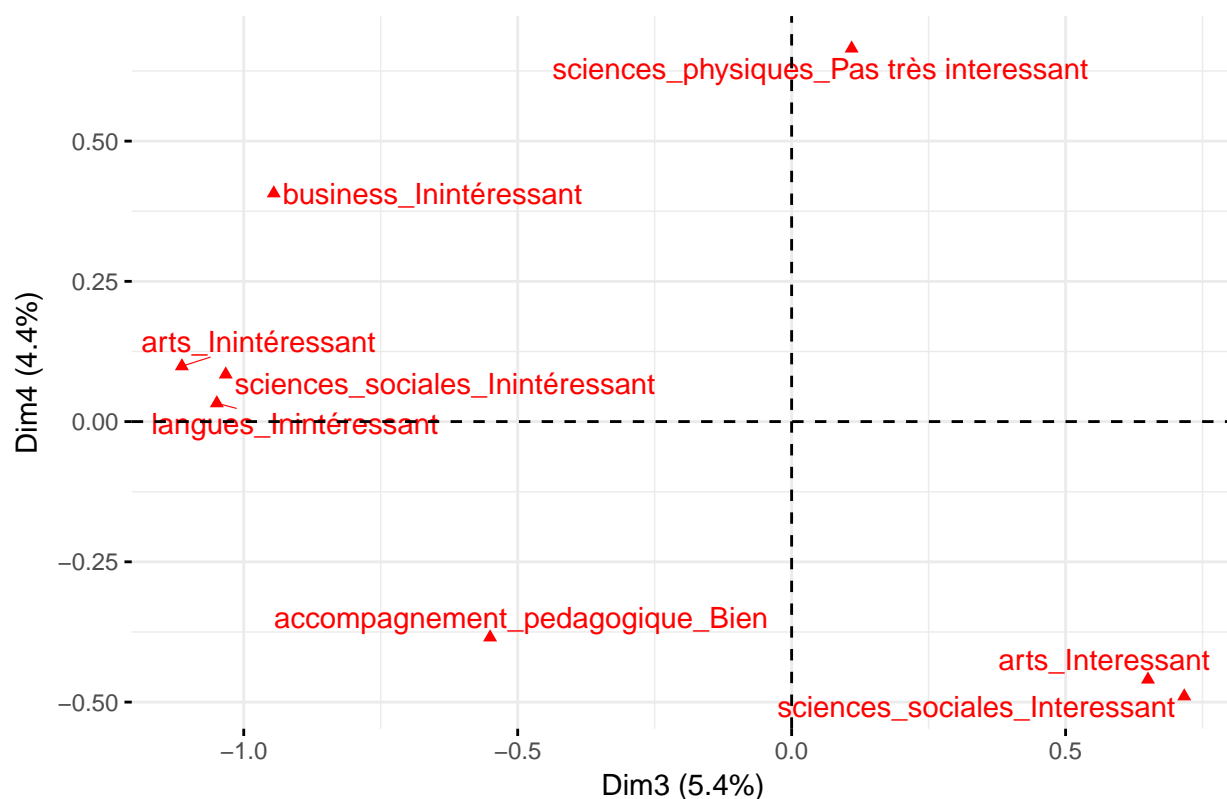
Le graphique ci-dessus montre les relations entre les catégories des variables. Il peut être interprété comme suit:

- Les catégories avec un profil similaire sont regroupées (exemple : diapositive_cours_Très mauvais, videos_cours_Très_mauvais).
- La distance entre les catégories et l'origine mesure la qualité des catégories. Les points qui sont loin de l'origine sont bien représentés par l'ACM(exemple : diapositive_cours_Très mauvais, accompagnement_pedagogiques_Très bien). ==> Les 1 et 2 expliquent les outils des cours (exemens_exercices, forum_discussion, ...).

On va faire de même pour les dimensions 3 et 4 :

```
fviz_mca_var (res.mca,
              repel = TRUE,
              axes=c(3,4),
              ggtheme = theme_minimal (),select.var = list(cos2=0.22))
```

Variable categories – MCA



==> Ces dimensions expliquent les domaines choisis par les sondés. (ajouter commentaires) - les groupes se trouvant dans des côtés opposés à l'origine sont corrélés négativement.

Qualité de représentation des catégories des variables :

Les deux dimensions 1 et 2 capturent 14.99% de l'inertie totale (variation) contenue dans les données. Tous les points ne sont pas aussi bien représentés par les deux dimensions.

La qualité de représentation, appelée cosinus carré (\cos^2), mesure le degré d'association entre les catégories des variables et les dimensions. Le \cos^2 peut être extrait comme suit :

```
head(var$cos2, 4)
```

```
##               Dim 1      Dim 2
## motivation_certificat_D'accord  0.026230315  0.005165919
## motivation_certificat_Désaccord  0.010506291  0.065855582
## motivation_certificat_Neutre     0.009406782  0.002143998
## motivation_certificat_Tout à fait d'accord 0.009251344  0.052297293
##               Dim 3      Dim 4
## motivation_certificat_D'accord  1.792027e-05  0.0002152608
## motivation_certificat_Désaccord  6.998613e-03  0.0791046281
## motivation_certificat_Neutre     1.485769e-03  0.1354653319
## motivation_certificat_Tout à fait d'accord 7.718560e-03  0.0073854646
##               Dim 5
## motivation_certificat_D'accord  0.006103167
## motivation_certificat_Désaccord  0.002660453
## motivation_certificat_Neutre     0.021829087
```

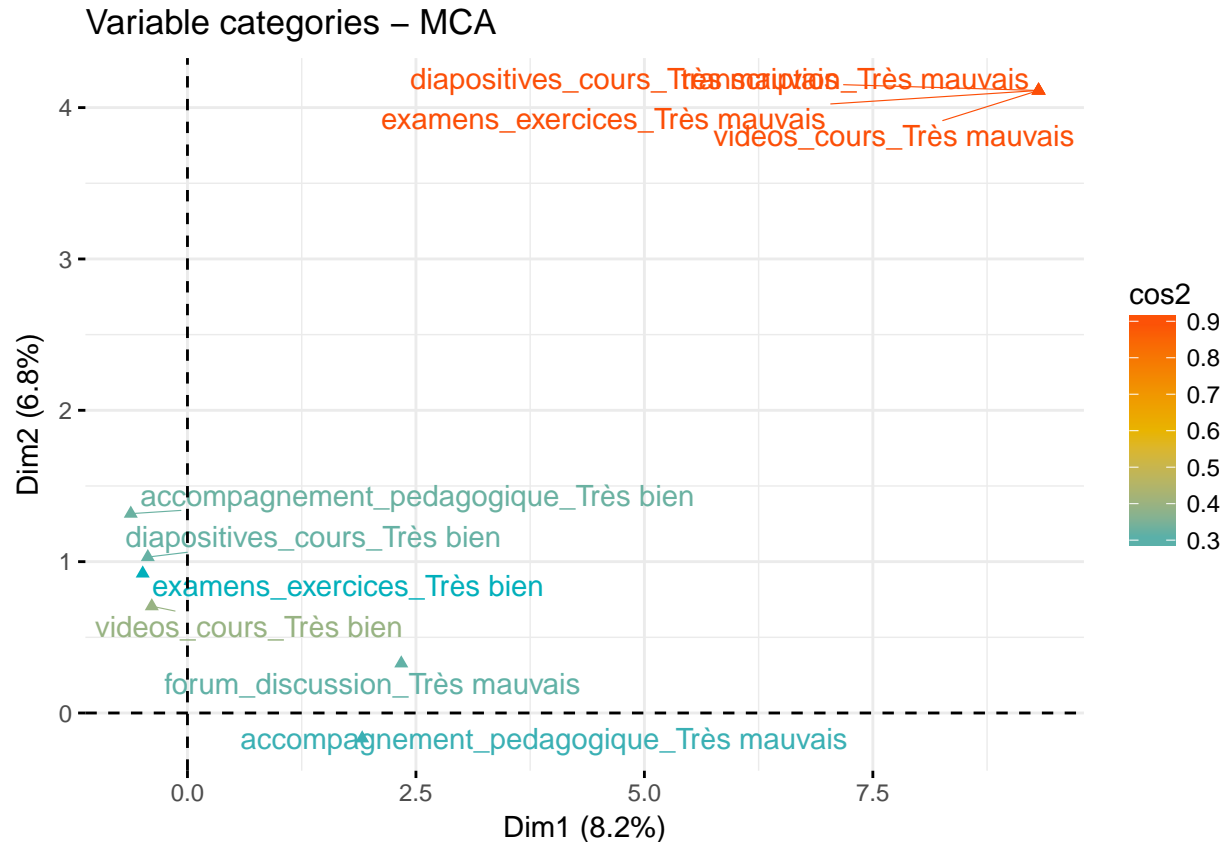
```
## motivation_certificat_Tout à fait d'acord 0.183341580
```

Si une catégorie d'une variable donnée est bien représentée par deux dimensions, la somme des cos2 est proche de 1. Pour certains éléments, plus de 2 dimensions sont nécessaires pour représenter parfaitement les données.

Il est possible de colorer les variables en fonction de la valeur de leur cos2 à l'aide de l'argument `col.var = "cos2"`. Cela produit un gradient de couleurs. Dans ce cas, l'argument `gradient.cols` peut être utilisé pour spécifier une palette de couleur personnalisée. Par exemple, `gradient.cols = c("white", "blue", "red")` signifie que:

- les variables à faible valeur de cos2 seront colorées en "white" (blanc)
- les variables avec des valeurs moyennes de cos2 seront colorées en "blue" (bleu)
- les variables avec des valeurs élevées de cos2 seront colorées en "red" (rouge)

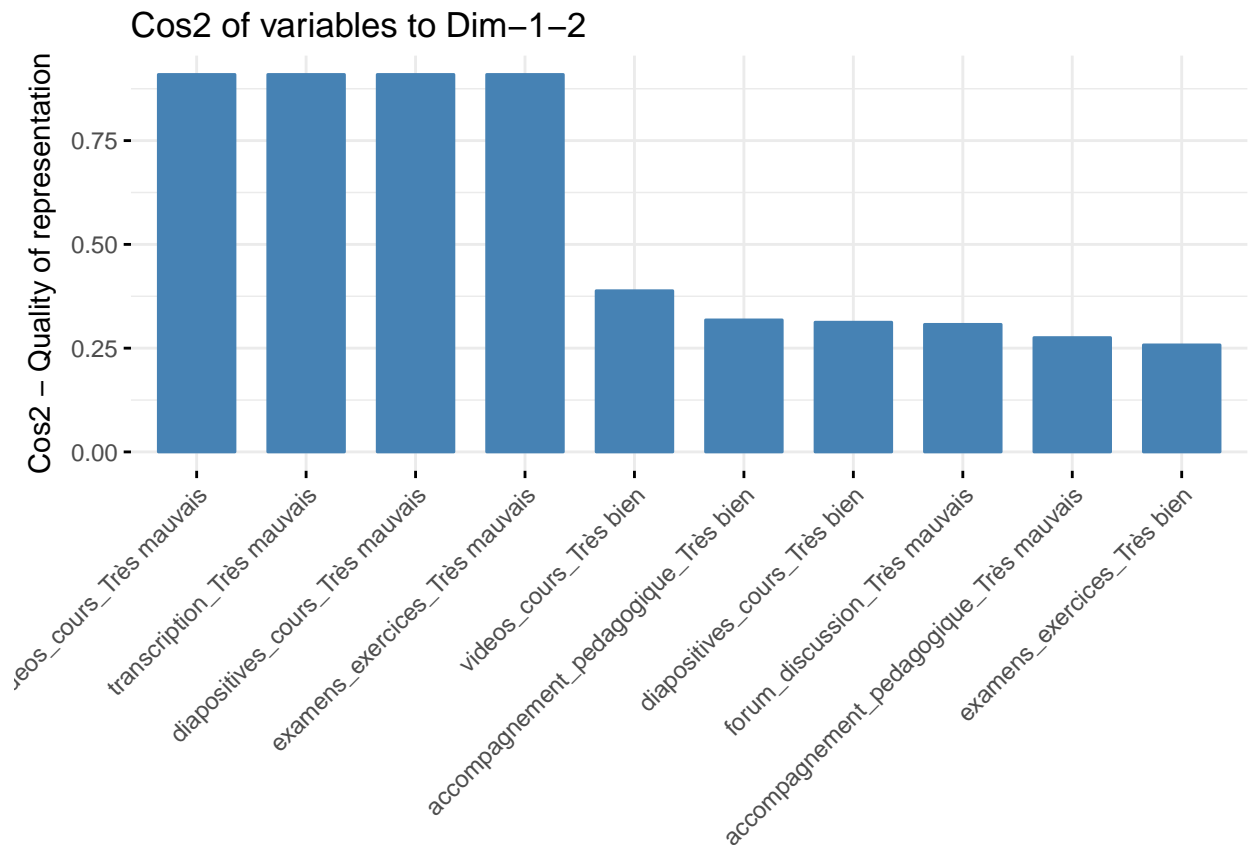
```
# Colorer en fonction du cos2
fviz_mca_var(res.mca, col.var = "cos2",
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
  repel = TRUE,
  select.var = list(cos2=0.25),
  ggtheme = theme_minimal())
```



Ainsi les catégories des variables en relation avec la gestion des cours en lignes (examens, ..) sont très bien représentées sur les dimensions 1 et 2.

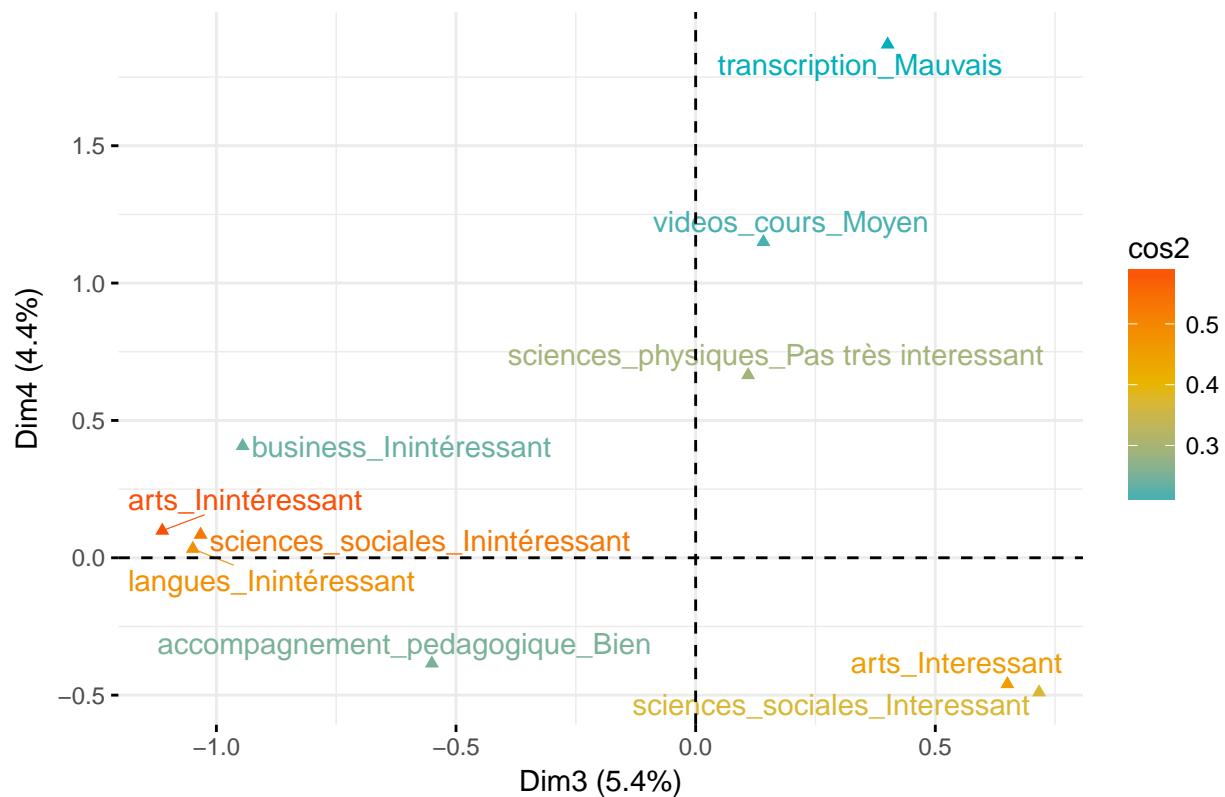
Il est également possible de créer un barplot du cos2 des variables avec la fonction `fviz_cos2()` [factoextra]:

```
# Cos2 des variable sur Dim.1 et Dim.2
fviz_cos2(res.mca, choice = "var", axes = 1:2, top = 10)
```

```
# Colorer en fonction du cos2
fviz_mca_var(res.mca, col.var = "cos2",
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
  repel = TRUE,
  axes=c(3,4),
  select.var = list(cos2=0.20),
  ggtheme = theme_minimal())
```

Variable categories – MCA



==> ainsi les domaines des moocs sont très bien représentées dans les dimensions 3 et 4.

Contribution des variables aux dimensions :

La contribution des variables (en %) à la définition des dimensions peut être extraite comme suit:

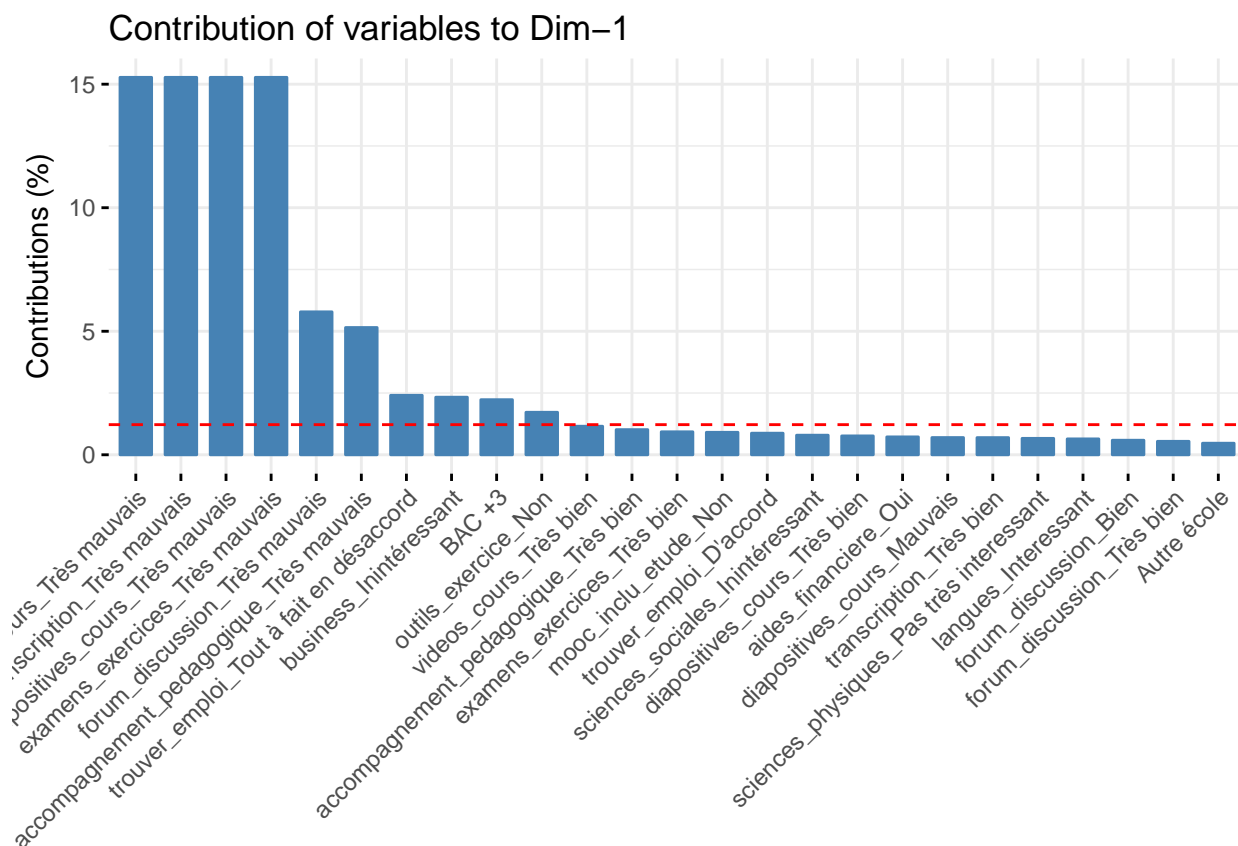
```
head(round(var$contrib,2), 4)
```

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
## motivation_certificat_D'accord	0.41	0.10	0.00	0.01	0.19
## motivation_certificat_Désaccord	0.17	1.27	0.17	2.34	0.08
## motivation_certificat_Neutre	0.16	0.05	0.04	4.36	0.74
## motivation_certificat_Tout à fait d'accord	0.14	0.95	0.18	0.21	5.43

Les variables avec les plus grandes valeurs, contribuent le mieux à la définition des dimensions. Les catégories qui contribuent le plus à Dim.1 et Dim.2 sont les plus importantes pour expliquer la variabilité dans le jeu de données.

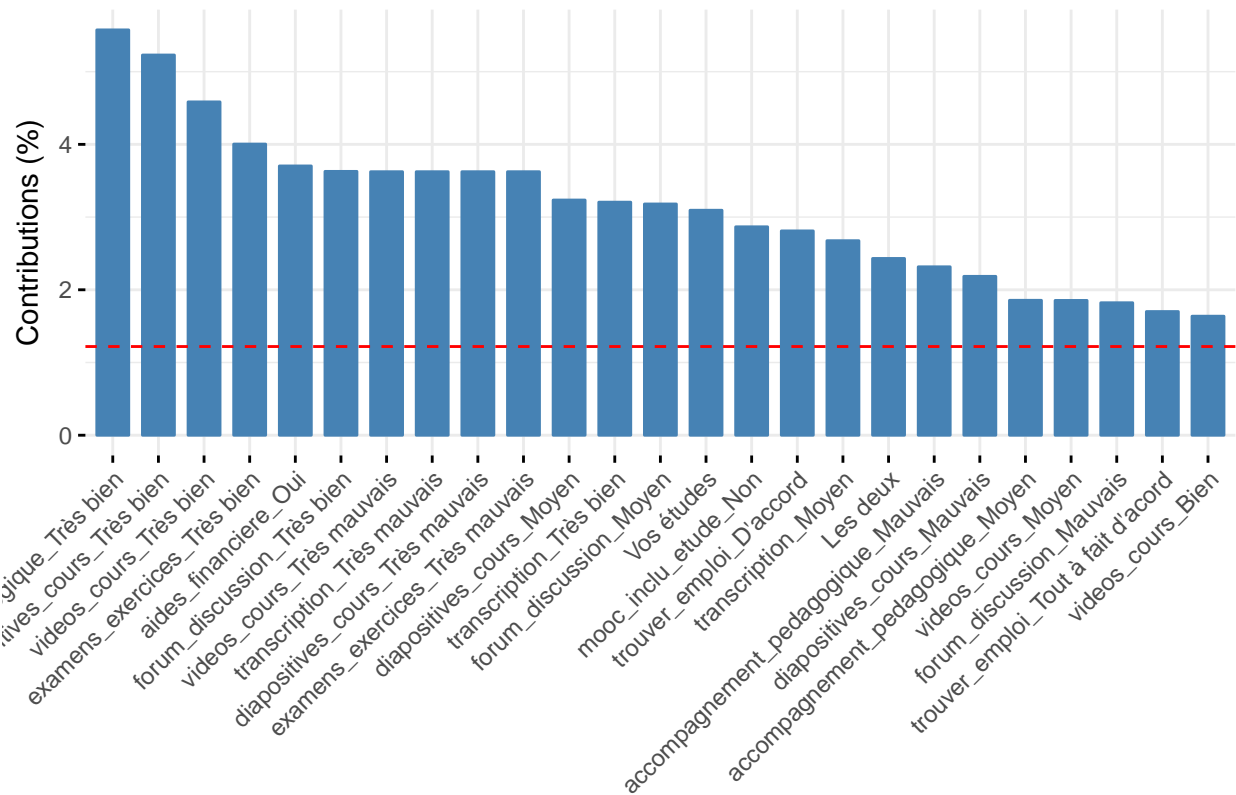
La fonction `fviz_contrib()` [factoextra] peut être utilisée pour faire un barplot de la contribution des catégories des variables. Le code R ci-dessous montre le top 15 des catégories contribuant aux dimensions:

```
# Contributions des variables à la dimension 1
fviz_contrib(res.mca, choice = "var", axes = 1, top = 25)
```



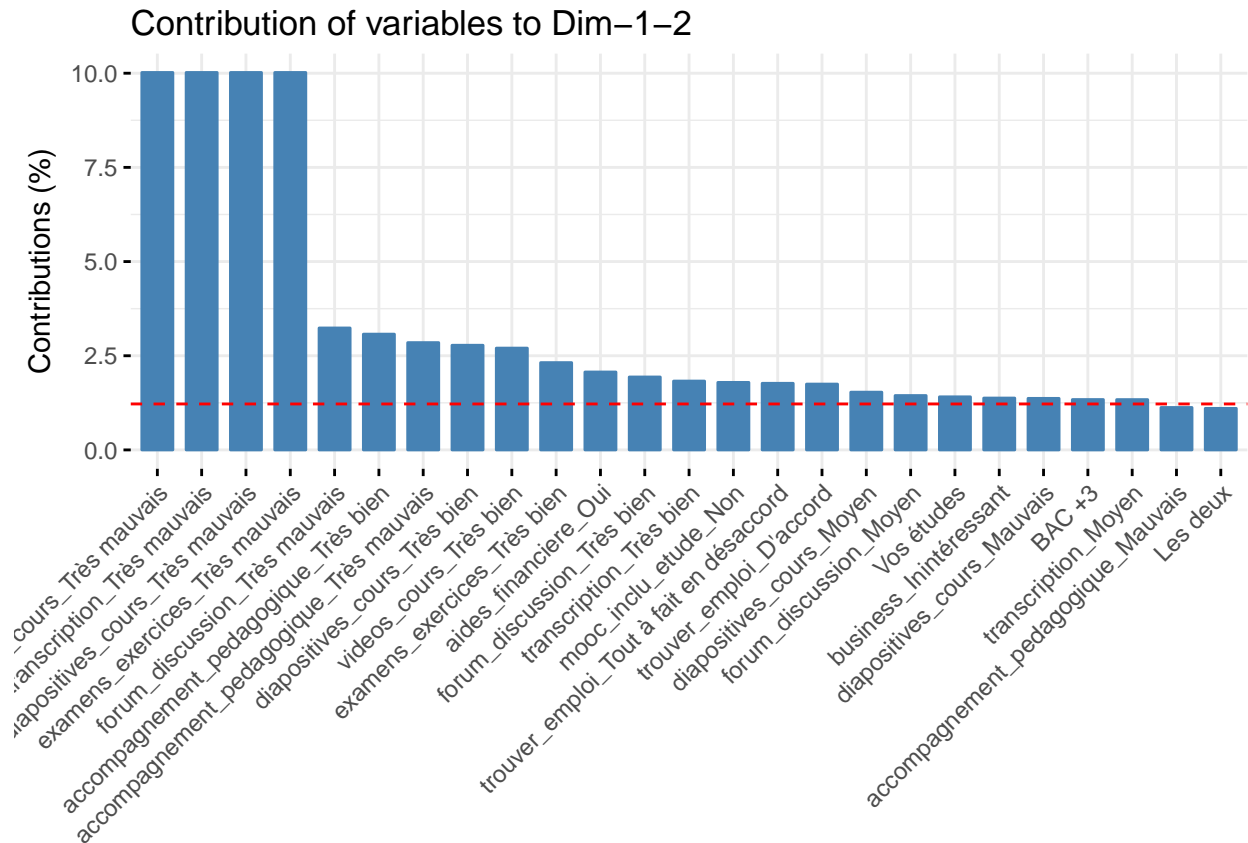
```
# Contributions des variables à la dimension 2
fviz_contrib(res.mca, choice = "var", axes = 2, top = 25)
```

Contribution of variables to Dim-2



Les contributions totales aux dimensions 1 et 2 sont obtenues comme suit:

```
# Contribution totale aux dimensions 1 et 2
fviz_contrib(res.mca, choice = "var", axes = 1:2, top = 25)
```

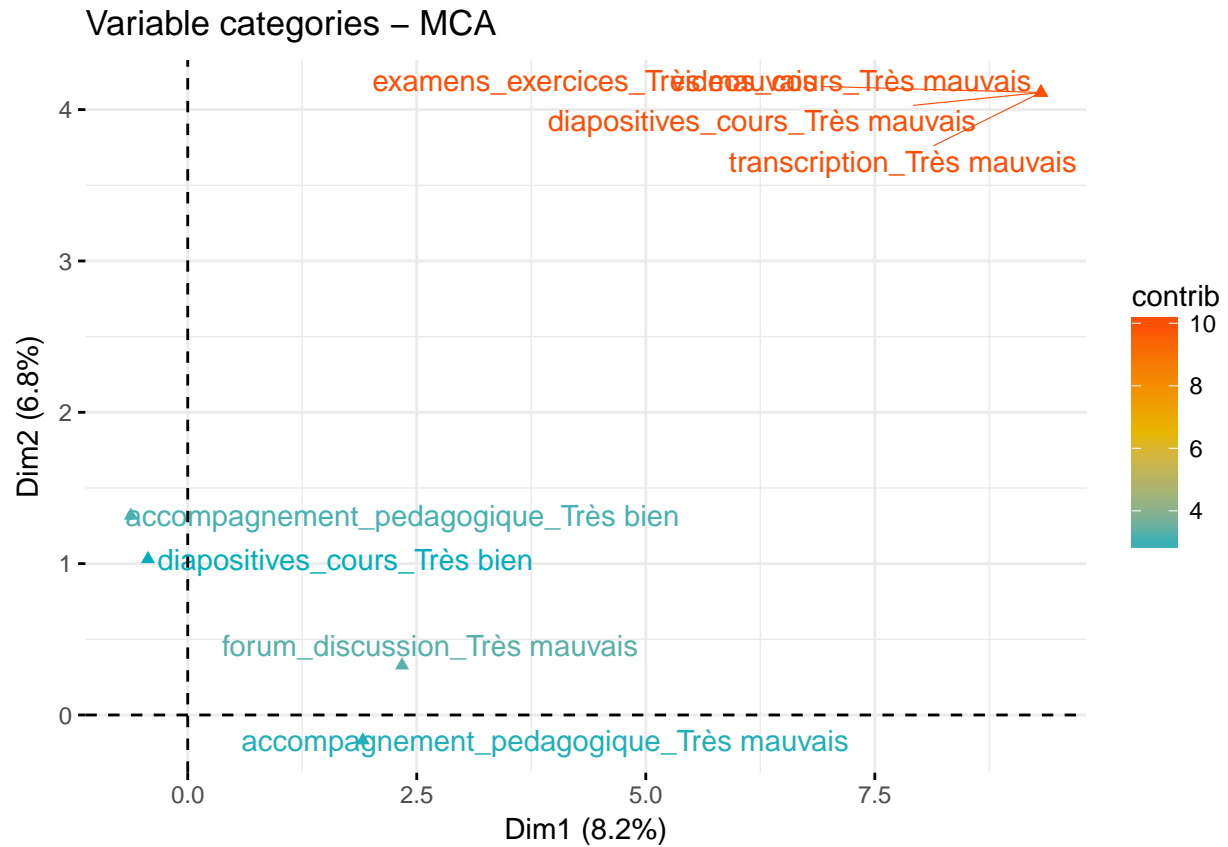


La ligne en pointillé rouge, sur le graphique ci-dessus, indique la valeur moyenne attendue sous l'hypothèse nulle.

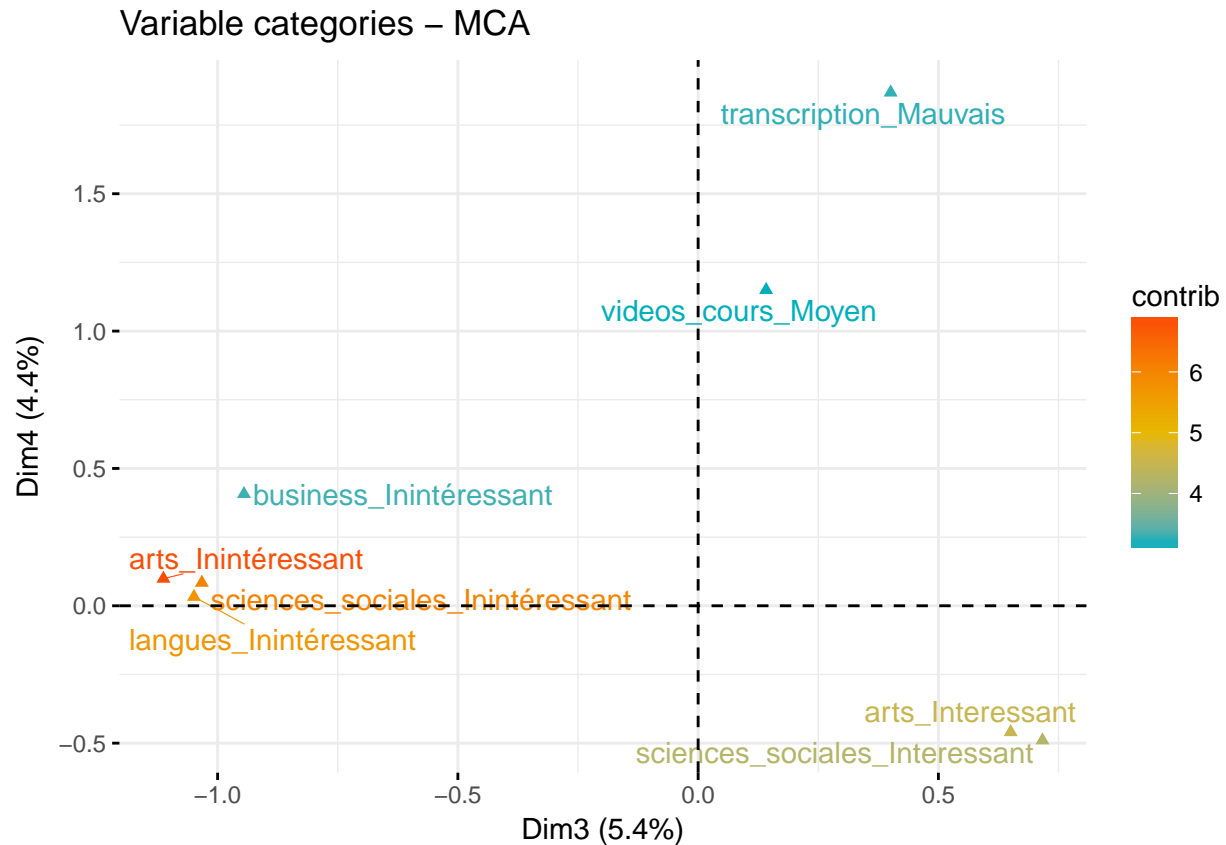
On peut voir que:

les catégories diapositive_cours_Très mauvais, transcription_Très mauvais, diapositives_cours_très mauvais sont les plus importantes dans la définition de la première dimension. Les accompagnement_pédagogiques Très bien, diapositives_cours_très Bien, vidéos_cours_TrèsBien contribuent le plus à la dimension 2 Les catégories les plus importantes peuvent être mises en évidence sur le graphique comme suit:

```
fviz_mca_var(res.mca, col.var = "contrib",
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
  repel = TRUE,
  select.var = list(contrib=7.5),
  ggtheme = theme_minimal()
)
```



```
fviz_mca_var(res.mca, col.var = "contrib",
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
  repel = TRUE,
  axes=c(3,4),
  select.var = list(contrib=7.5),
  ggtheme = theme_minimal()
)
```



***** a changer !!! le commentaire : Le graphique ci-dessus donne une idée du pôle des dimensions auquel les catégories contribuent réellement.

Il est évident que les catégories Abdo_n, Diarrhea_n, Fever_n et Mayo_n ont une contribution importante au pôle positif de la première dimension, tandis que les catégories Fever_y et Diarrhea_y ont une contribution majeure au pôle négatif de la première dimension; etc,

Graphique des individus :

Résultats

La fonction `get_mca_ind()` [factoextra] sert à extraire les résultats pour les individus. Cette fonction renvoie une liste contenant les coordonnées, la cos2 et les contributions des individus:

```
ind <- get_mca_ind (res.mca)
ind

## Multiple Correspondence Analysis Results for individuals
## =====
##   Name      Description
## 1 "$coord"   "Coordinates for the individuals"
## 2 "$cos2"    "Cos2 for the individuals"
## 3 "$contrib" "contributions of the individuals"
```

Pour accéder aux différents composants, utilisez ceci:

```
print(' Coordonnées')
```

```
## [1] " Coordonnées"
head(ind$coord)

##           Dim 1      Dim 2      Dim 3      Dim 4      Dim 5
## 1 -0.29371140  0.4752218 -0.58105187 -0.06175800 -0.09296914
## 2 -0.37720771  1.0313020  0.09724377  0.75880670  0.14015673
## 3 -0.20873787  0.2464032 -0.14139764 -0.40091258  0.52805405
## 4  0.20657258 -0.2722099 -0.31760656  0.23985738 -0.32867303
## 5  0.37540679 -0.5536708 -0.70535956 -0.06536711 -0.13503426
## 6 -0.03723415 -0.5467696  0.45204681  0.11576866  0.08114459

print(' Qualité de representation')

## [1] " Qualité de representation"
head(ind$cos2)

##           Dim 1      Dim 2      Dim 3      Dim 4      Dim 5
## 1 0.0489557573 0.12816068 0.191598446 0.0021644554 0.004905008
## 2 0.0443075364 0.33119835 0.002944693 0.1792993791 0.006117084
## 3 0.0201360077 0.02805845 0.009239643 0.0742797151 0.128862819
## 4 0.0197696994 0.03432911 0.046734075 0.0266539092 0.050047556
## 5 0.0309597283 0.06734357 0.109298453 0.0009386668 0.004005722
## 6 0.0006651164 0.14342460 0.098035126 0.0064297914 0.003158887

print('Contributions')

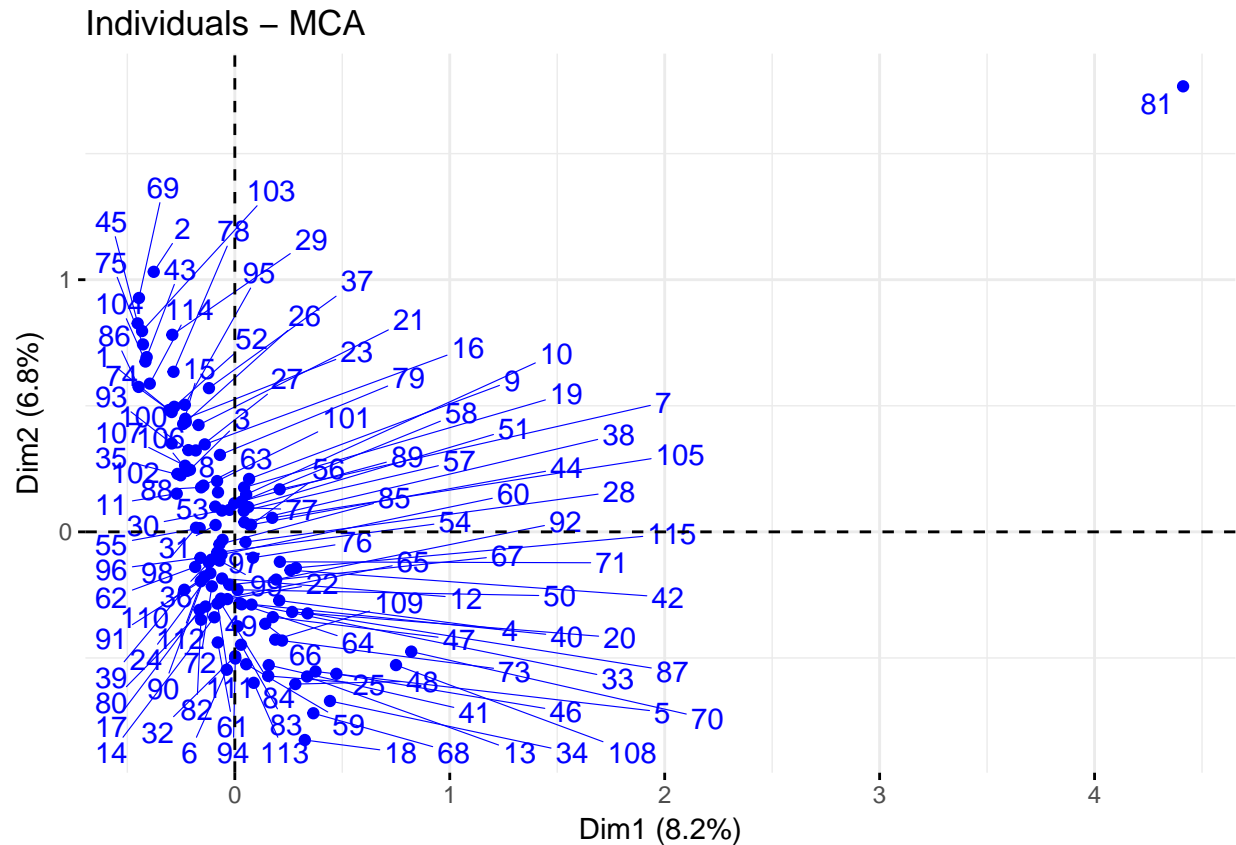
## [1] "Contributions"
head(ind$contrib)

##           Dim 1      Dim 2      Dim 3      Dim 4      Dim 5
## 1 0.334249937 1.0642696 1.99755222 0.02758738 0.06621088
## 2 0.551303648 5.0122265 0.05594894 4.16472412 0.15048033
## 3 0.168823243 0.2861224 0.11829126 1.16258118 2.13603599
## 4 0.165338908 0.3491941 0.59682533 0.41613103 0.82752275
## 5 0.546051969 1.4446488 2.94367221 0.03090598 0.13968178
## 6 0.005371709 1.4088597 1.20902463 0.09694065 0.05043952
```

graphique: qualité et contribution

Utilisez la fonction `fviz_ca_row()` [dans `factoextra`] pour visualiser uniquement les points lignes:

```
fviz_mca_ind(res.mca, repel = TRUE)
```

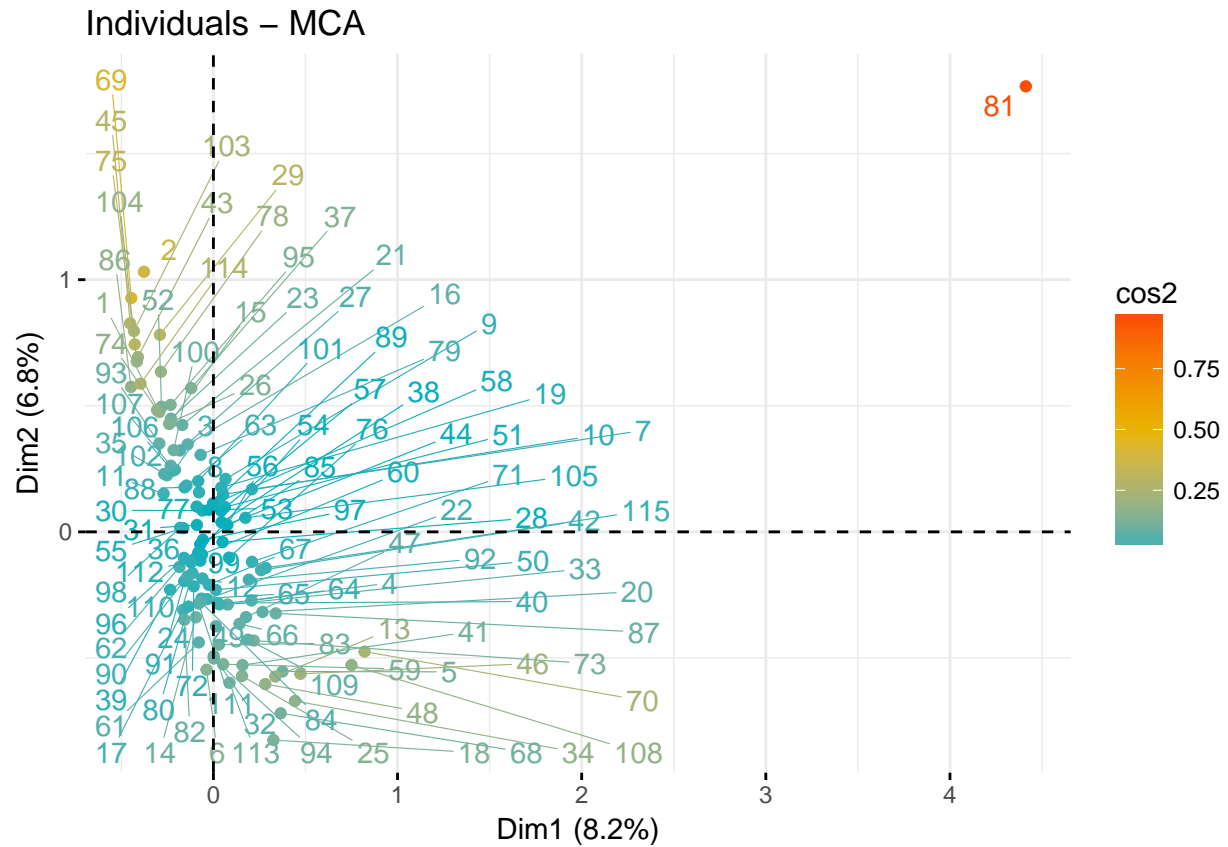



Le graphique ci-dessus montre les relations entre les points lignes:

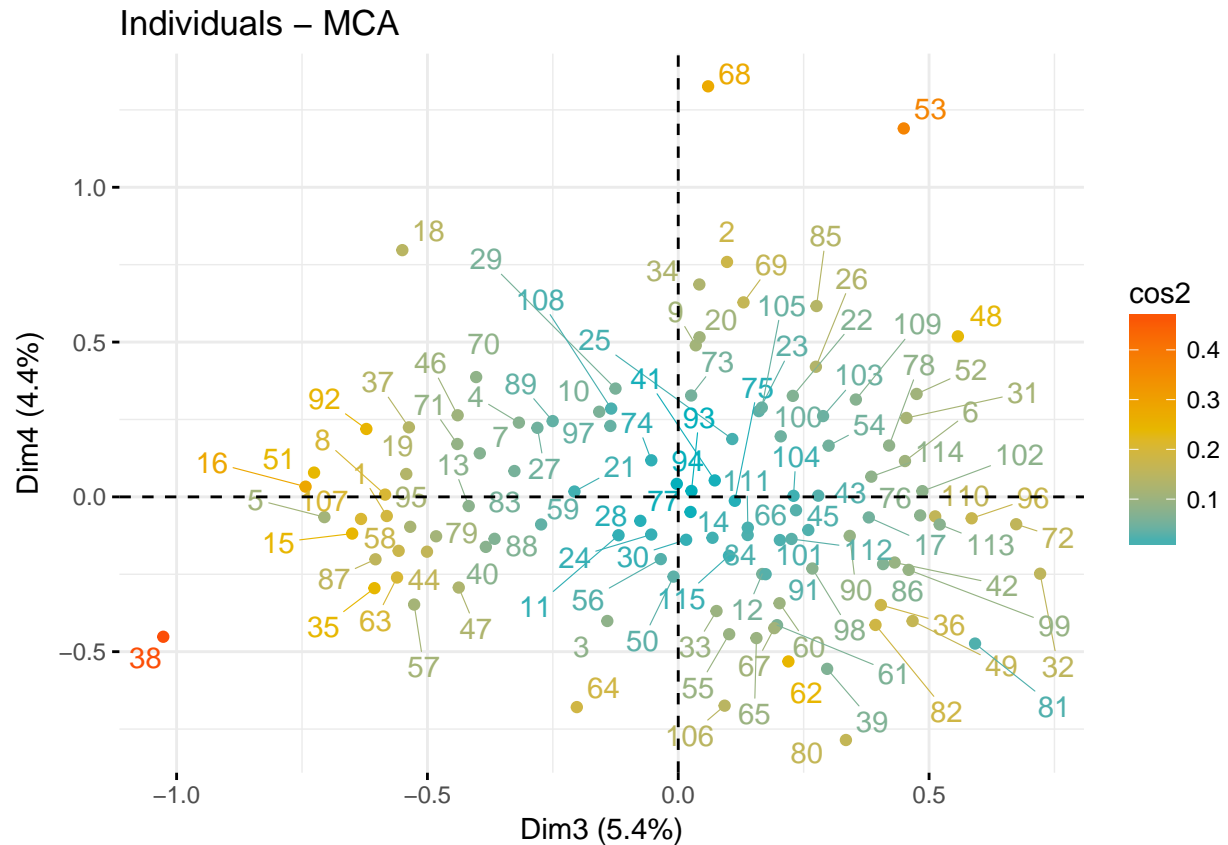
Les lignes avec un profil similaire sont regroupées. Les lignes corrélées négativement sont positionnées sur des côtés opposés de l'origine de du graphique (quadrants opposés). La distance entre les points lignes et l'origine mesure la qualité des points lignes sur le graphique. Les points lignes qui sont loin de l'origine sont bien représentés sur le graphique.

La fonction `fviz_mca_ind()` [factoextra] sert à visualiser uniquement des individus. Comme les variables, il est également possible de colorer les individus en fonction de leurs `cos2`:

```
fviz_mca_ind(res.mca, col.ind = "cos2",
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
  repel = TRUE,
  ggtheme = theme_minimal())
```



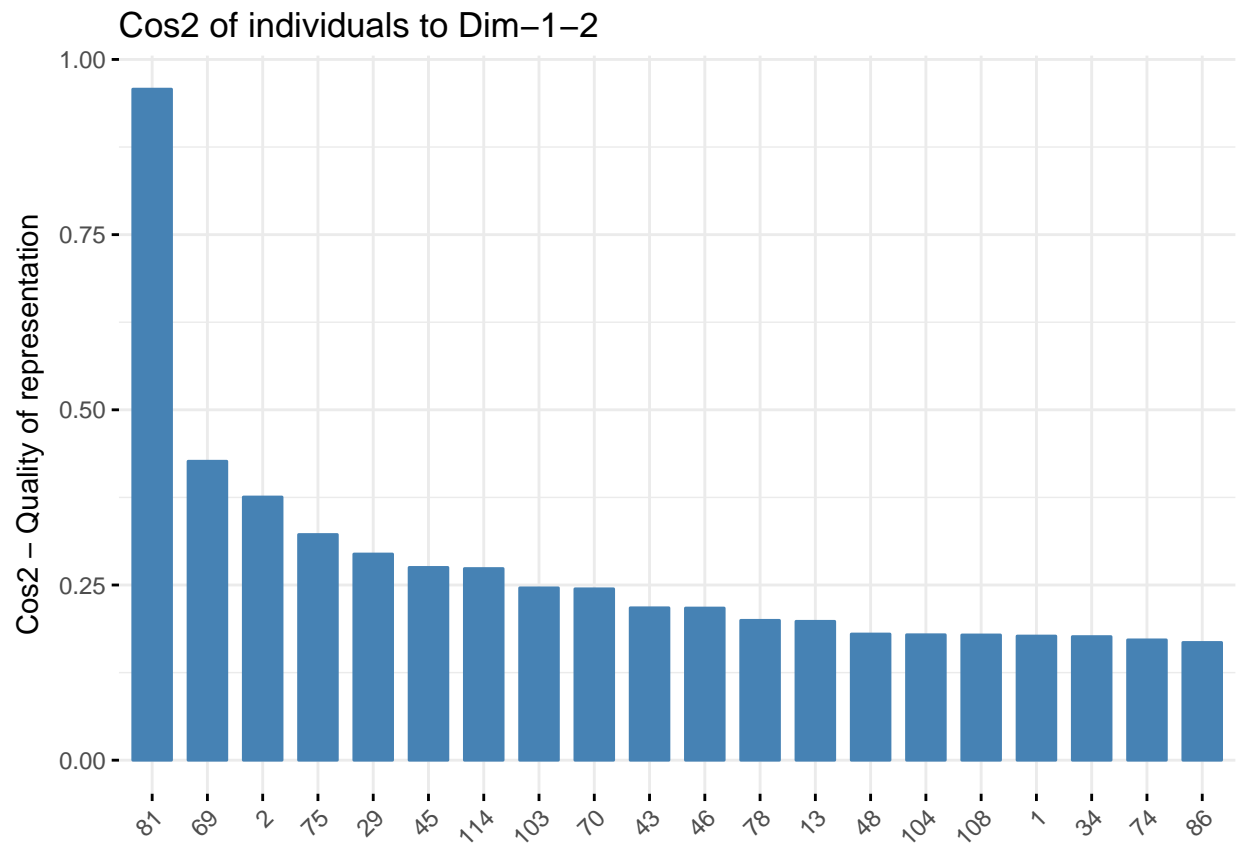
```
fviz_mca_ind(res.mca, col.ind = "cos2",
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
  repel = TRUE,
  axes=c(3,4),
  ggtheme = theme_minimal())
```



Le graphique ci-dessus montre les relations entre les points lignes:

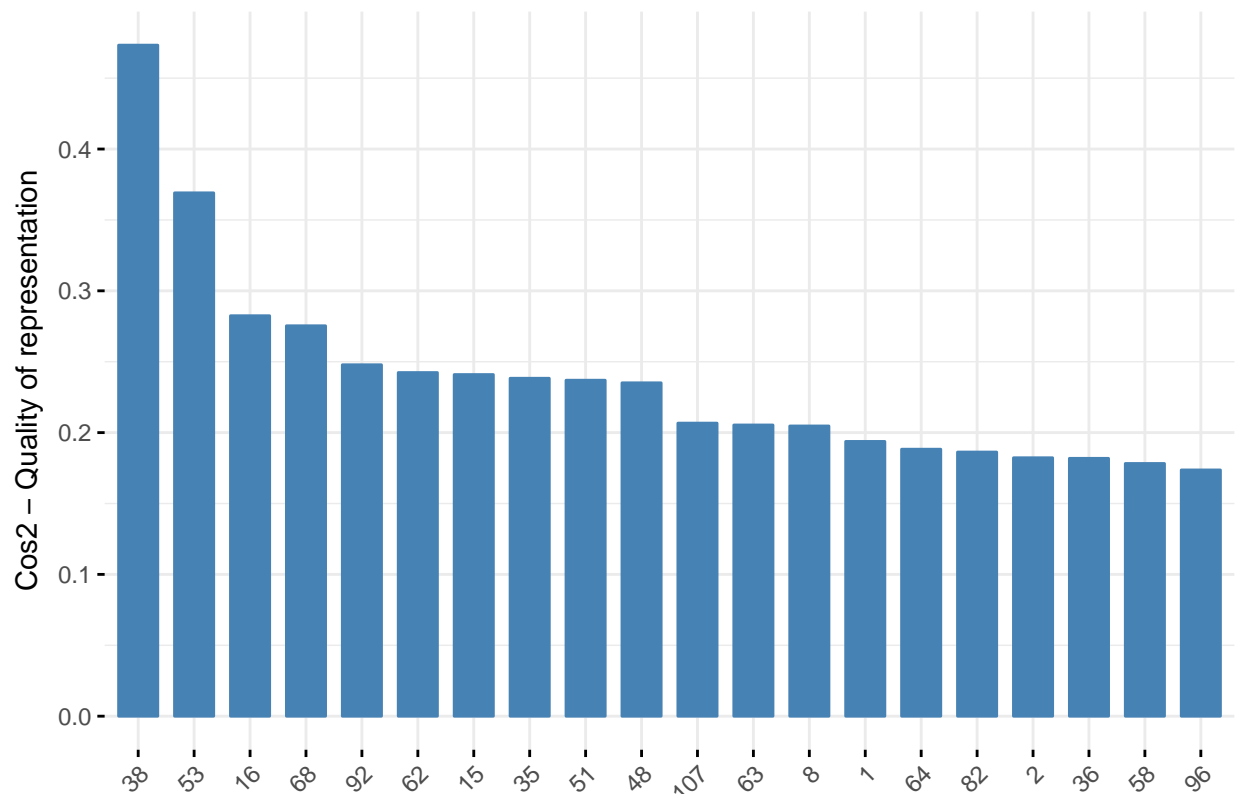
Les lignes avec un profil similaire sont regroupées. Les lignes corrélées négativement sont positionnées sur des côtés opposés de l'origine de du graphique (quadrants opposés). La distance entre les points lignes et l'origine mesure la qualité des points lignes sur le graphique. Les points lignes qui sont loin de l'origine sont bien représentés sur le graphique. Il est également possible de créer un bar plot du \cos^2 des lignes en utilisant la fonction `fviz_cos2()` [factoextra]:

```
# Cos2 des lignes sur Dim.1 et Dim.2
fviz_cos2(res.mca, choice = "ind", axes = 1:2, las=2, top = 20)
```



```
# Cos2 des lignes sur Dim.1 et Dim.2
fviz_cos2(res.mca, choice = "ind", axes = 3:4, las=2, top = 20)
```

Cos2 of individuals to Dim-3-4



Colorer les individus par groupes : Le code R ci-dessous colore les individus par groupes en utilisant la variable `outils_exercice`. L'argument `habillage` sert à spécifier la variable à utiliser pour colorer les individus par groupes. Une ellipse de concentration peut également être ajoutée autour de chaque groupe en utilisant l'argument `addEllipses = TRUE`. Si vous voulez une ellipse de confiance autour du point moyen (centre de gravité) des groupes, utilisez `ellipse.type = "confidence"`. L'argument `palette` permet de modifier les couleurs du groupe.

```
str(data_suivi
)
```

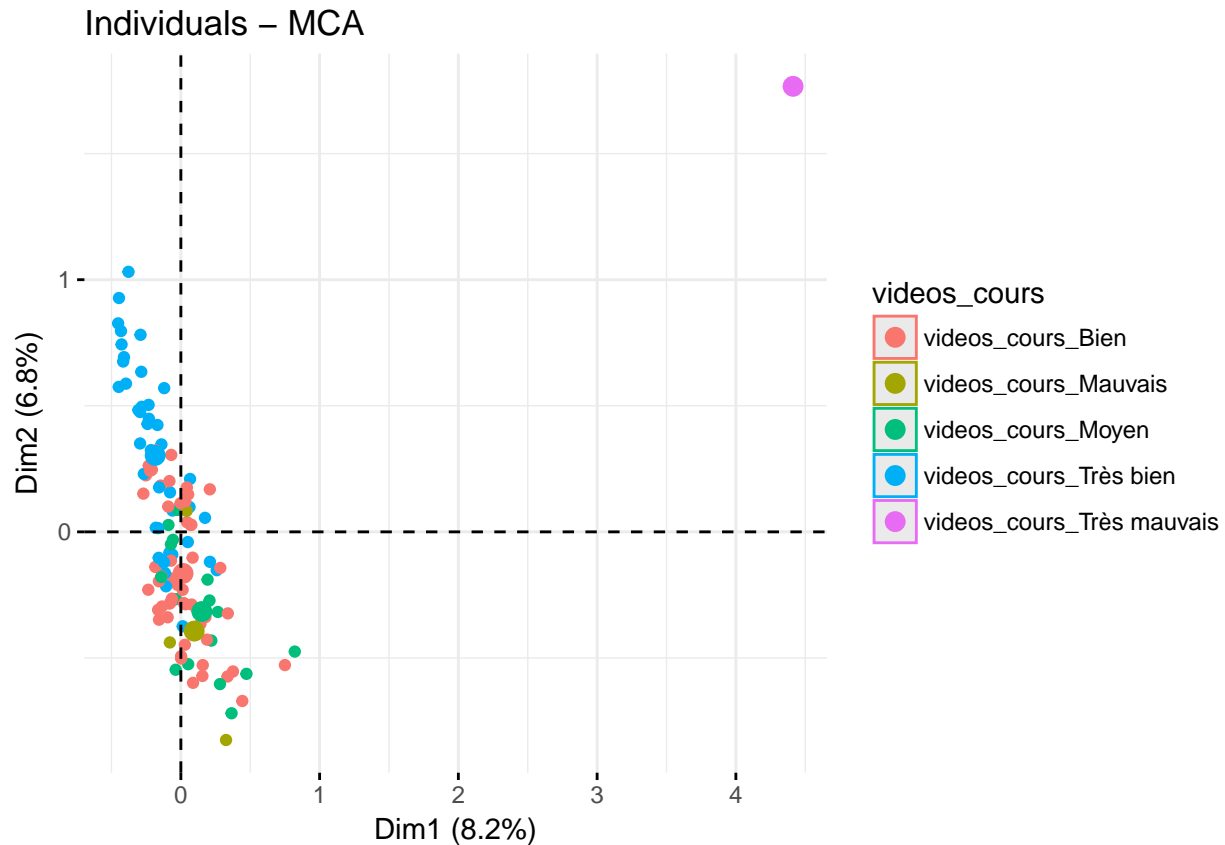
```
## 'data.frame': 115 obs. of 36 variables:
## $ a_suivi : Factor w/ 3 levels "Jamais","Oui, je l'ai suivi et j'ai réussi",...: 2
## $ performer_travail_projet : Factor w/ 5 levels "D'accord","Désaccord",...: 4 4 1 1 5 3 4 1 4 1 ...
## $ motivation_certificat : Factor w/ 5 levels "D'accord","Désaccord",...: 1 4 1 2 5 1 2 1 2 4 ...
## $ trouver_emploi : Factor w/ 5 levels "D'accord","Désaccord",...: 3 4 1 5 5 2 3 1 2 2 ...
## $ nbr_certificat : num 1 2 1 1 1 1 1 1 3 1 ...
## $ satisfait_gestion_temps : Factor w/ 5 levels "D'accord","Désaccord",...: 1 4 1 3 4 4 2 1 4 3 ...
## $ objectifs_atteints : Factor w/ 5 levels "D'accord","Désaccord",...: 4 4 1 1 3 1 5 1 4 1 ...
## $ satisfaction : num 7 10 9 6 1 7 3 6 7 6 ...
## $ vacances_etude : Factor w/ 3 levels "Les deux","Les vacances",...: 1 1 1 3 3 3 1 3 1 1 ...
## $ temps_alloue_cours : num 14 9 13 5 9 11 9 7 13 5 ...
## $ mooc_inclu_etude : Factor w/ 2 levels "Non","Oui": 1 1 2 2 2 2 2 2 1 ...
## $ edt_chargee : Factor w/ 2 levels "Non","Oui": 1 1 2 2 2 2 2 1 1 2 ...
## $ plateformes_utilisees : Factor w/ 30 levels "Canvas, OpenClassrooms, FUN",...: 14 2 21 23 23 2
## $ videos_cours : Factor w/ 5 levels "Bien","Mauvais",...: 4 4 1 3 1 3 1 4 4 1 ...
## $ transcription : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 1 1 3 3 4 1 4 ...
## $ diapositives_cours : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 1 3 1 4 1 4 1 ...
```

```
## $ forum_discussion      : Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 4 3 3 5 1 4 1 ...
## $ examens_exercices     : Factor w/ 5 levels "Bien","Mauvais",...: 4 4 1 3 2 1 1 3 1 1 ...
## $ accompagnement_pedagogique: Factor w/ 5 levels "Bien","Mauvais",...: 1 4 1 3 1 2 1 1 1 3 ...
## $ aides_financiere      : Factor w/ 2 levels "Non","Oui": 2 2 2 1 1 1 2 1 1 1 ...
## $ prix                  : num NA 10 20 0 0 0 0 100 50 30 ...
## $ outils_exercice       : Factor w/ 2 levels "Non","Oui": 2 2 2 1 2 2 2 2 1 2 ...
## $ sciences_donnees      : Factor w/ 3 levels "Inintéressant",...: 2 2 2 2 2 3 2 2 2 2 ...
## $ informatique          : Factor w/ 3 levels "Inintéressant",...: 2 2 3 2 3 2 2 2 2 2 ...
## $ sciences_physiques    : Factor w/ 3 levels "Inintéressant",...: 2 3 2 2 2 2 3 2 3 3 ...
## $ business              : Factor w/ 3 levels "Inintéressant",...: 2 3 3 3 1 2 3 1 1 1 ...
## $ arts                  : Factor w/ 3 levels "Inintéressant",...: 1 3 2 1 1 3 2 1 3 3 ...
## $ langues                : Factor w/ 3 levels "Inintéressant",...: 1 3 2 1 1 2 1 1 3 3 ...
## $ sciences_sociales     : Factor w/ 3 levels "Inintéressant",...: 1 3 3 1 1 2 1 1 1 1 ...
## $ sexe                  : Factor w/ 2 levels "Femme","Homme": 2 2 2 2 2 1 2 2 2 2 ...
## $ pays                  : Factor w/ 13 levels "Autre","Burkinafaso",...: 8 8 5 5 5 5 8 5 5 5 ..
## $ age                   : num 22 22 22 22 20 22 22 22 23 23 ...
## $ est_etudiant          : Factor w/ 2 levels "Non","Oui": 2 2 2 2 2 2 2 2 2 2 ...
## $ niveau_actuel         : Factor w/ 5 levels "BAC +2","BAC +3",...: 4 4 2 4 2 3 4 4 3 4 ...
## $ formation             : Factor w/ 6 levels "FIG","FIL","FIP",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ background            : Factor w/ 4 levels "Autre école",...: 3 3 4 3 3 3 3 3 3 3 ...
```

```
fviz_mca_ind(res.mca,
  label = "none", # masquer le texte des individus
  habillage = "videos_cours", # colorer par groupes

  addEllipses = TRUE, ellipse.type = "confidence",
  ggtheme = theme_minimal())
```

```
## Warning: Computation failed in `stat_conf_ellipse()`:
## missing value where TRUE/FALSE needed
```



il y a un outlier qui peut fausser les résultats, donc on peut refaire l'étude en supprimant la ligne 81.

Description des dimensions :

La fonction `dimdesc()` [FactoMineR] peut être utilisée pour identifier les variables les plus corrélées avec une dimension donnée:

```
res.desc <- dimdesc (res.mca, axes = c(1,2))
print(' Description de la dimension 1')
```

```
## [1] " Description de la dimension 1"
```

```
res.desc[[1]]
```

```
## $quali
##
##          R2      p.value
## diapositives_cours 0.84313824 2.685140e-43
## videos_cours        0.82704527 5.670494e-41
## examens_exercices   0.82628382 7.213719e-41
## transcription       0.80578313 3.251274e-38
## forum_discussion    0.34829032 1.195042e-09
## accompagnement_pedagogique 0.33006014 5.178616e-09
## business            0.14078126 2.040843e-04
## trouver_emploi      0.17203359 3.238232e-04
## outils_exercice     0.10274073 4.787202e-04
## niveau_actuel       0.13423792 7.063558e-03
## mooc_inclu_etude    0.05732141 9.965385e-03
```

```
## aides_financiere          0.04511321 2.267256e-02
## sciences_sociales         0.06332074 2.565020e-02
## edt_chargee               0.03805973 3.667513e-02
## langues                    0.05620508 3.918749e-02
## sciences_physiques        0.05398220 4.470552e-02
##
## $category
##
## Estimate      p.value
## examens_exercices_Très mauvais 3.49343746 6.720722e-37
## diapositives_cours_Très mauvais 3.47554830 6.720722e-37
## transcription_Très mauvais 3.52345825 6.720722e-37
## videos_cours_Très mauvais 3.51390855 6.720722e-37
## forum_discussion_Très mauvais 0.92744894 2.130549e-10
## accompagnement_pedagogique_Très mauvais 0.78305690 2.064327e-09
## trouver_emploi_Tout à fait en désaccord 0.37543480 1.876305e-05
## business_Inintéressant 0.30709528 3.610169e-05
## BAC +3 0.39149687 6.053095e-05
## outils_exercice_Non 0.20030813 4.787202e-04
## sciences_sociales_Inintéressant 0.16263635 9.259715e-03
## mooc_inclu_etude_Oui 0.13749004 9.965385e-03
## sciences_physiques_Pas très interessant 0.15225237 1.258501e-02
## aides_financiere_Non 0.12380255 2.267256e-02
## edt_chargee_Oui 0.09400123 3.667513e-02
## diapositives_cours_Mauvais -0.46552584 4.339933e-02
## edt_chargee_Non -0.09400123 3.667513e-02
## aides_financiere_Oui -0.12380255 2.267256e-02
## forum_discussion_Bien -0.31455971 2.166185e-02
## diapositives_cours_Très bien -1.14206754 1.984289e-02
## transcription_Très bien -1.04752339 1.713647e-02
## trouver_emploi_D'accord -0.20201116 1.254764e-02
## langues_Interessant -0.15123400 1.083710e-02
## examens_exercices_Très bien -1.15065994 1.034224e-02
## mooc_inclu_etude_Non -0.13749004 9.965385e-03
## accompagnement_pedagogique_Très bien -0.41645734 9.706851e-03
## videos_cours_Très bien -1.08308162 1.049077e-03
## outils_exercice_Oui -0.20030813 4.787202e-04
```

```
print('Description de la dimension 2')
```

```
## [1] "Description de la dimension 2"
```

```
res.desc[[2]]
```

```
## $quali
##
## R2      p.value
## diapositives_cours 0.57937583 6.775245e-20
## videos_cours 0.49679120 1.117634e-15
## transcription 0.43768590 4.446772e-13
## accompagnement_pedagogique 0.42798480 1.115003e-12
## examens_exercices 0.41645621 3.254846e-12
## forum_discussion 0.41536611 3.597652e-12
## vacances_etude 0.22471371 6.454728e-07
## trouver_emploi 0.26271281 8.108191e-07
## aides_financiere 0.18998003 1.122656e-06
## mooc_inclu_etude 0.14867534 2.085636e-05
```



```

## motivation_certificat      0.10833309 1.269551e-02
## sciences_donnees           0.07500396 1.270086e-02
##
## $category
##
##               Estimate      p.value
## videos_cours_Très bien      0.06388185 2.982591e-10
## diapositives_cours_Très bien 0.19746063 3.864968e-09
## accompagnement_pedagogique_Très bien 0.55699282 5.762571e-09
## examens_exercices_Très bien 0.14151321 5.039995e-07
## Les deux                    0.19445408 5.803971e-07
## aides_financiere_Oui        0.23036531 1.122656e-06
## transcription_Très bien      0.02797286 1.389276e-06
## forum_discussion_Très bien   0.44857623 5.505964e-06
## mooc_inclu_etude_Non        0.20077870 2.085636e-05
## examens_exercices_Très mauvais 1.51241156 2.124520e-05
## diapositives_cours_Très mauvais 1.52180463 2.124520e-05
## transcription_Très mauvais   1.51398811 2.124520e-05
## videos_cours_Très mauvais    1.52771489 2.124520e-05
## trouver_emploi_D'accord      0.23732795 3.180767e-05
## forum_discussion_Bien        0.14934944 8.838873e-04
## sciences_donnees_Interessant 0.18203573 3.061049e-03
## trouver_emploi_Tout à fait d'acord 0.40707230 3.534026e-03
## Autre école                  0.37258411 1.282711e-02
## motivation_certificat_Tout à fait d'acord 0.17901985 1.396112e-02
## accompagnement_pedagogique_Bien 0.11666931 1.959796e-02
## arts_Interessant            0.11144650 4.392838e-02
## examens_exercices_Mauvais    -0.89096986 3.471232e-02
## sciences_donnees_Pas très interessant -0.09345263 1.660736e-02
## transcription_Mauvais        -0.66303039 1.609505e-02
## trouver_emploi_Tout à fait en désaccord -0.27558531 1.544688e-02
## trouver_emploi_Désaccord     -0.25098922 1.483563e-02
## motivation_certificat_Désaccord -0.19548779 5.632583e-03
## forum_discussion_Mauvais     -0.44148062 2.158780e-03
## examens_exercices_Moyen      -0.46099852 1.537344e-03
## videos_cours_Moyen           -0.55505356 1.337447e-03
## diapositives_cours_Mauvais    -0.93099059 9.912576e-04
## accompagnement_pedagogique_Mauvais -0.37348267 3.333851e-04
## accompagnement_pedagogique_Moyen -0.21950299 3.006960e-04
## videos_cours_Bien            -0.40381397 1.367759e-04
## transcription_Moyen          -0.58413603 6.209733e-05
## mooc_inclu_etude_Oui         -0.20077870 2.085636e-05
## diapositives_cours_Moyen     -0.54486111 2.338580e-06
## aides_financiere_Non         -0.23036531 1.122656e-06
## forum_discussion_Moyen       -0.27342738 2.784813e-07
## Vos études                   -0.23118342 1.725165e-07

```

Eléments supplémentaires :