

# TD5

## Enoncé du TD5

Les données sont extraites de l'enquête *Trajectoires et Origines (2019-2020)* s'intéressant en particulier aux immigrés (nommés ici « G1 ») et aux descendants d'immigrés (enfants d'un ou deux immigrés, nommés « G2 ») résidant en France métropolitaine. Les personnes originaires ou descendantes des régions d'Outre-Mer ont également été surreprésentées dans l'échantillon de l'enquête. Enfin, l'enquête a également interrogé des personnes sans ascendance migratoire (« population majoritaire »), la composition de l'échantillon permettant ainsi de comparer le poids des origines par rapport à la population qui n'a pas d'ascendance migratoire sur au moins deux générations.

Dans les données présentées (fichier « tabl.csv »), on a croisé le lien à la migration et l'origine géographique (en lignes) avec la profession, codée suivant la nomenclature des Professions et Catégories Socioprofessionnelles (en colonne).

On se pose la question de l'association entre origines migratoires et situation professionnelle :

- Les immigrés et leurs descendants, suivant leur origine géographique, ont-ils plus de chances d'appartenir à certaines professions ?
- En particulier, sont-ils plus souvent en bas de la hiérarchie des professions ou au contraire en haut ?
- Quelles distinctions existe-t-il au sein du monde des professions qui rendent compte de la segmentation du marché du travail suivant l'origine migratoire ?

Le script R « TD5.R » propose quelques lignes de commande permettant de répondre aux questions du TD et vous pouvez également vous appuyer sur les scripts de la séance précédente pour compléter ce travail.

*Ceci est un corrigé rapide qui peut être complété de manière plus détaillée en examinant davantage les plans factoriels et en interprétant sociologiquement la segmentation du marché du travail. Il vise simplement à donner quelques pistes d'analyses.*

## Question 1

Après avoir ouvert la base (« tab1.csv ») et fait tourner les premières lignes de commande, vous utiliserez les fonctions `dim`, `rownames` et `colnames` permettant de décrire les données. Vous justifierez l'adéquation entre la méthode de l'Analyse Factorielle des Correspondances (AFC) et les données ici étudiées.

```
library(readr)
library(tibble)
library(FactoMineR)
library(explor)
library(factoextra)

setwd("/Users/mathieuferry/Documents/Enseignements/Année 2023-2024/L3 Sociologie Analyse d

tab1<-read_csv2("tab1.csv")
tab1<-column_to_rownames(tab1,var="...1")

dim(tab1)
```

```
[1] 30 25
```

```
rownames(tab1)
```

```
[1] "Pop majoritaire"           "G1 Outre-Mer"
[3] "G2 Outre-Mer"             "G1 Algérie"
[5] "G2 Algérie"               "G1 Maroc/Tunisie"
[7] "G2 Maroc ou de Tunisie"   "G1 Afrique sahélienne"
[9] "G2 Afrique sahélienne"    "G1 Afrique guinéenne ou centrale"
[11] "G2 Afrique guinéenne ou centrale" "G1 autres pays d'Afrique"
[13] "G2 autres pays d'Afrique"  "G1 Asie du Sud-est"
[15] "G2 Asie du Sud-est"        "G1 Turquie ou du Moyen-Orient"
[17] "G2 Turquie ou du Moyen-Orient" "G1 Chine"
[19] "G1 autres pays d'Asie"     "G2 autres pays d'Asie"
[21] "G1 Portugal"              "G2 Portugal"
[23] "G1 Espagne ou d'Italie"    "G2 Espagne ou d'Italie"
[25] "G1 autres pays de l'UE27"  "G2 autres pays de l'UE27"
[27] "G1 autres pays d'Europe"   "G2 autres pays d'Europe"
[29] "G1 autres pays"           "G2 autres pays"
```

```
colnames(tab1)
```

```
[1] "Agriculteurs"  
[2] "Artisan"  
[3] "Commerçant"  
[4] "Chefs d'entreprise (>=10 salariés)"  
[5] "Professions libérales"  
[6] "Cadres fonction publique"  
[7] "Professions scientifiques"  
[8] "Professions info/arts/spectacles"  
[9] "Cadres d'entreprise"  
[10] "Ingénieurs/cadres techniques d'entreprise"  
[11] "Professeurs des écoles"  
[12] "Prof intermédiaires santé"  
[13] "Professions intermédiaires fonction publique"  
[14] "Prof intermédiaires entreprises"  
[15] "Techniciens"  
[16] "Contremaîtres"  
[17] "Employés fonction publique"  
[18] "Policiers et militaires"  
[19] "Employés d'entreprise"  
[20] "Employés de commerce"  
[21] "Personnels aux particuliers"  
[22] "Ouvriers qualif"  
[23] "Chauffeurs"  
[24] "Ouvriers non qualif"  
[25] "Ouvriers agricoles"
```

Le fichier tab1 correspond à un tableau de contingence croisant en ligne le lien à la migration (immigré G1, descendant d'immigré G2 et population majoritaire avec l'origine migratoire) avec 30 modalités différentes et en colonnes les PCS avec 25 modalités différentes. Dans les cellules se trouvent les effectifs d'individus.

Ce tableau de contingence, de taille importante, est difficilement analysable facilement par le recours à des pourcentages en ligne / colonne, ou même par des représentations de type diagrammes en bâtons.

En revanche, il se prête bien à une analyse factorielle des correspondances (voir cours magistral pour plus de détail).

## Question 2

Réalisez l'AFC.

Avant toute chose, on peut déjà regarder s'il y a un lien entre l'origine migratoire et la position sur le marché du travail. On peut conduire un test statistique du chi-2 d'indépendance. Dans ce test, l'hypothèse H0 est que les deux variables sont indépendantes et l'hypothèse alternative (H1) est que les deux variables sont liées. Si la p-value associée au test statistique est faible ( $<0.05$ ), alors on rejette H0 et donc on rejette l'hypothèse d'indépendance entre les deux variables. Pour plus de détails sur ce test, on peut regarder cette vidéo : <https://www.youtube.com/watch?v=-Xn1nmHjnHU>.

```
chisq.test(tab1)
```

```
Warning in chisq.test(tab1): Chi-squared approximation may be incorrect
```

```
Pearson's Chi-squared test
```

```
data: tab1  
X-squared = 1746.6, df = 696, p-value < 2.2e-16
```

Ici, on voit que la statistique du khi-deux est de 1747, avec 696 degrés de liberté, R nous indique que la probabilité associée que sous l'hypothèse d'indépendance on obtienne cette statistique est extrêmement faible. Donc, on rejette l'hypothèse d'indépendance, et on peut conclure qu'il y a un lien statistique entre nos deux variables étudiées.

Note : la conduite du test renvoie un "Warning" : R nous prévient que le calcul du test n'est pas optimal car normalement on ne calcule pas la statistique du khi-deux quand il y a des cellules vides, mais nous ignorons ce potentiel problème ici.

Quelle est l'intensité de ce lien ? On peut l'étudier grâce au V de Cramer (voir le cours magistral !). C'est la racine carrée du chi2 divisée par divisée par le nombre minimum entre les colonnes et les lignes moins 1.

```
sqrt(1746.6/(sum(tab1)*min(nrow(tab1)-1,ncol(tab1)-1)))
```

```
[1] 0.05372636
```

On peut aussi directement utiliser la fonction `cramer.v` qui est dans le package `questionr` :

```
install.packages("questionr", repos = "http://cran.us.r-project.org")
```

The downloaded binary packages are in  
/var/folders/xs/xxbdqg252p5f7\_s35g2mw9nc0000gq/T//Rtmp8Sd89W/downloaded\_packages

```
library(questionr)  
cramer.v(tab1)
```

```
[1] 0.05372675
```

On a bien le même résultat V-Cramer = 0.05. Au maximum il peut être de 1, au minimum il est de 0. Comme il est  $< 0.2$ , on peut dire que la liaison ici est plutôt faible même si elle existe.

On réalise l'AFC :

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

Les résultats sont stockés dans l'objet `res.afc` qu'on va exploiter à présent.

*En examinant les données, peut-on dire qu'il y a des associations exclusives ici qui ressortent sur l'un des axes ?*

S'il y avait une partition exclusive des lignes et des colonnes en deux classes (dans le tableau, il y aurait beaucoup de 0 pour certaines cellules), la première valeur propre serait de 1 (ou proche).

```
res.afc$eig[1,]
```

	eigenvalue	percentage of variance
	0.02418766	34.91416148
cumulative percentage of variance		
	34.91416148	

Ici, la première valeur propre est égale à 0.02, donc il n'y a pas de partition exclusive dans les données qui ressort dans l'AFC.

*Combien d'axes peut-on interpréter dans cette analyse ? Vous comparerez le critère du coude à la règle de Kaiser.*

Pour le critère de Kaiser, on regarde les valeurs propres qui ont une valeur supérieure à la valeur propre moyenne :

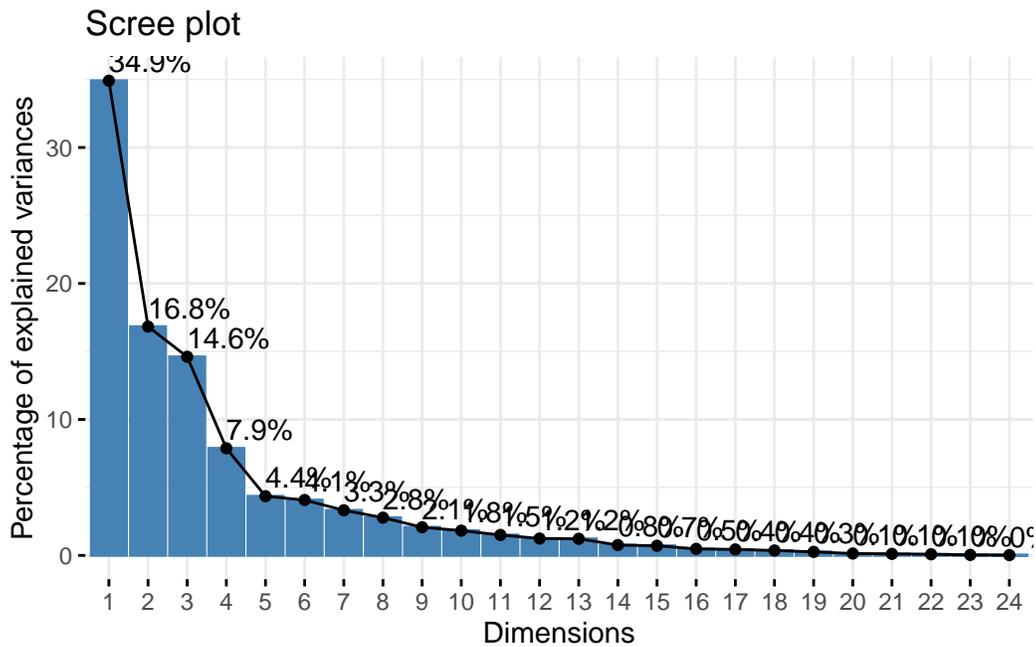
```
vp<-res.afc$eig
# On regarde lesquelles sont supérieures à l'inertie moyenne
vp[,1]>mean(vp[,1])
```

```
dim 1  dim 2  dim 3  dim 4  dim 5  dim 6  dim 7  dim 8  dim 9  dim 10  dim 11
  TRUE   TRUE   TRUE   TRUE   TRUE  FALSE  FALSE  FALSE  FALSE  FALSE  FALSE
dim 12  dim 13  dim 14  dim 15  dim 16  dim 17  dim 18  dim 19  dim 20  dim 21  dim 22
  FALSE  FALSE
dim 23  dim 24
  FALSE  FALSE
```

Cela nous pousserait à examiner 5 axes.

Pour le critère du coude, on peut regarder l'historique des valeurs propres :

```
##Le scree plot des valeurs propres
fviz_eig(res.afc,addlabels=T,ncp =24)
```



On a un premier décrochage entre la première et seconde dimension de l'AFC : la première est prépondérante. Ensuite un décrochage entre la troisième et quatrième : il semble que trois axes aient une valeur propre importante. Par facilité, on va ici se concentrer sur ces trois premiers axes.

```
sum(vp[1:3,2])
```

```
[1] 66.34759
```

```
sum(vp[1:4,2])
```

```
[1] 74.22648
```

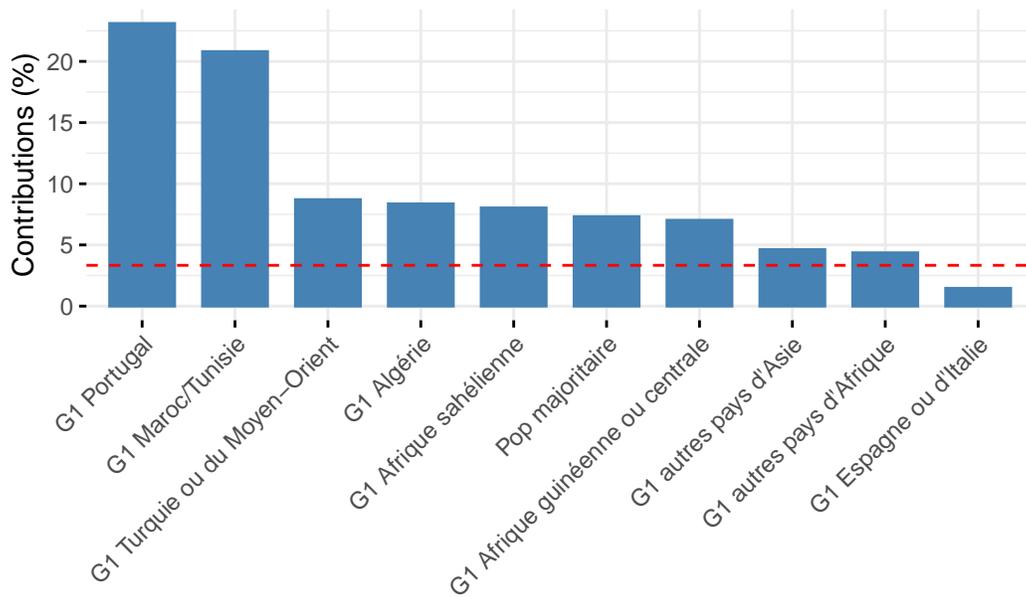
Ces trois premiers axes synthétisent les deux tiers des associations entre l'origine migratoire et la position sur le marché du travail, ce qui semble vraiment pas mal ! Si on poussait jusqu'à 4 axes, on synthétiserait les trois-quarts des associations. On regarde les trois premiers, et on verra si le quatrième est interprétable... Ensuite, l'information est plus négligeable.

### Question 3

*Après avoir déterminé le nombre d'axes à interpréter, vous les interprétez.*

```
##Regarder les profils lignes et les profils colonnes qui contribuent le plus à l'axe 1  
#On n'affiche que les 10 premiers pour plus de lisibilité  
fviz_contrib(res.afc,choice="row",axes =1,top=10)
```

Contribution of rows to Dim-1



```
fviz_contrib(res.afc,choice="col",axes =1,top=10)
```

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbsToSbcs': dot substituted  
for <e2>

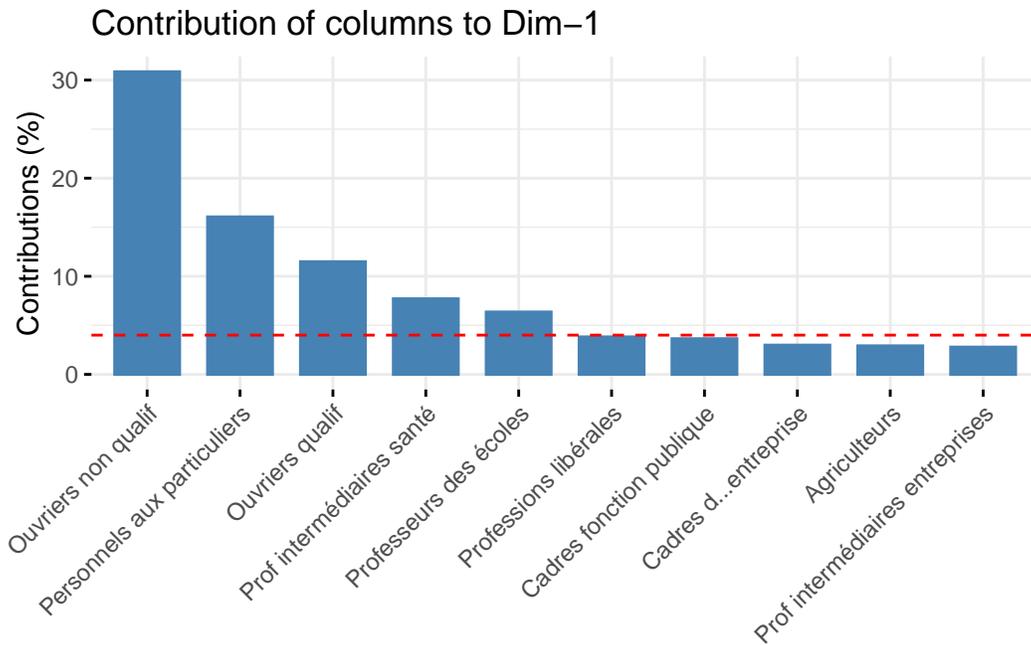
Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbsToSbcs': dot substituted  
for <80>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbsToSbcs': dot substituted  
for <99>

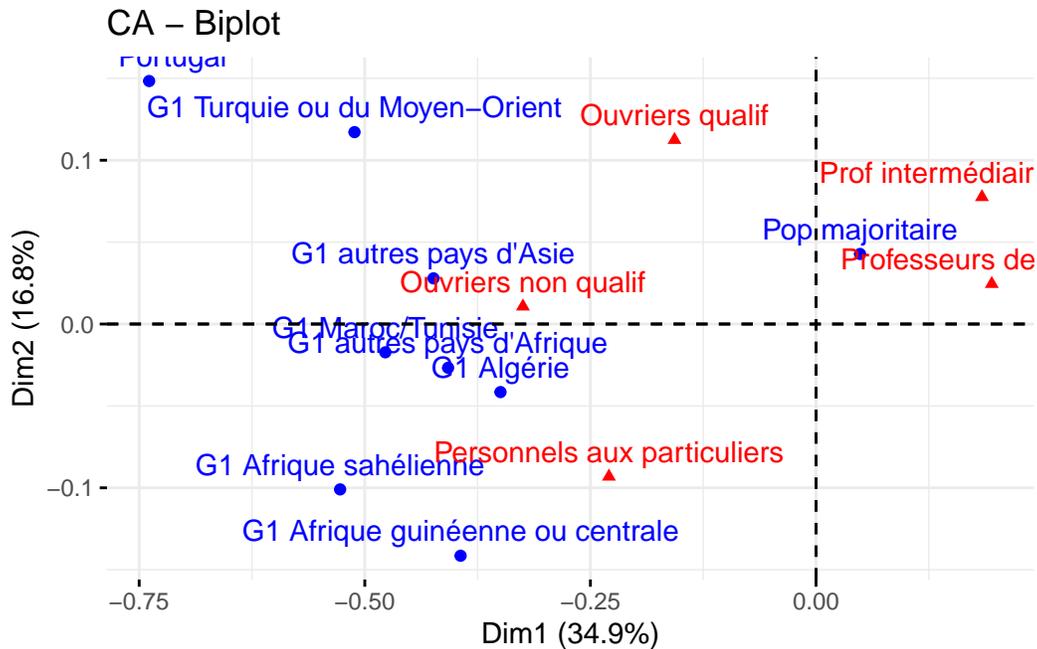
Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbsToSbcs': dot substituted  
for <e2>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbsToSbcs': dot substituted  
for <80>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted  
for <99>

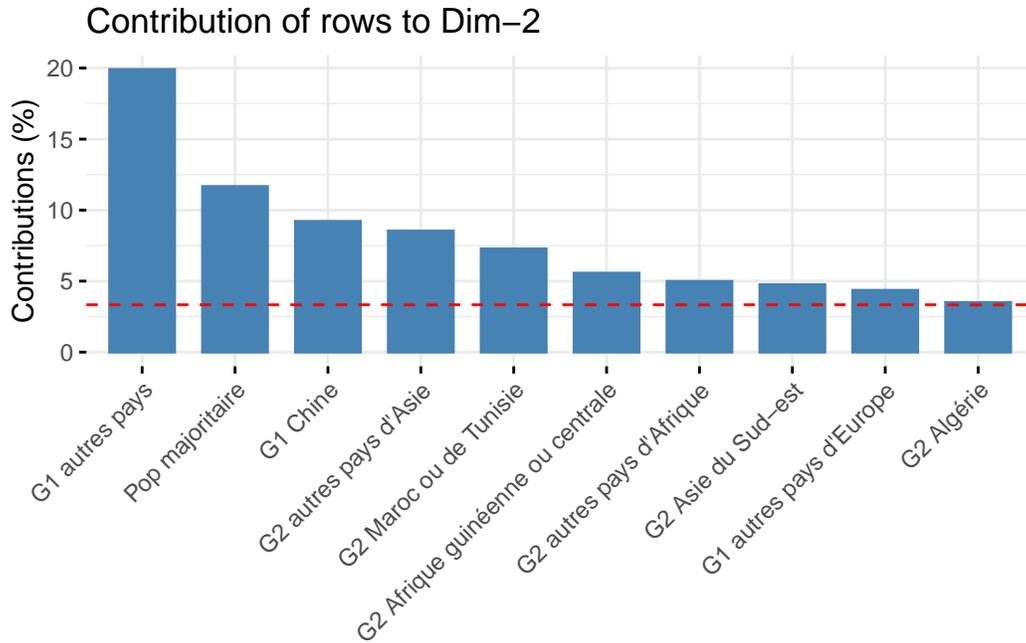


```
##Visualiser la contribution des lignes et colonnes qui ont une contribution supérieure à  
rowcontrib1<-rownames(res.afc$row$contrib)[res.afc$row$contrib[,1]>100/nrow(tab1)]  
colcontrib1<-rownames(res.afc$col$contrib)[res.afc$col$contrib[,1]>100/ncol(tab1)]  
fviz_ca_biplot(res.afc,axes=c(1,2),  
               select.row = list(name = rowcontrib1),  
               select.col=list(name=colcontrib1))
```



L'axe 1 (35 % de l'inertie totale) recouvre principalement à gauche les ouvriers (qualifiés et non qualifiés) et les personnels aux particuliers ainsi que les immigrés du Portugal, du Maghreb et d'Afrique. Ces profils s'opposent à droite aux professions intermédiaires et aux professeurs des écoles qui sont associés à la population majoritaire. Cet axe dénote la position subalterne des immigrés de quasiment toutes les origines (mais pas européennes à part pour les Portugais), qu'on retrouve davantage dans les emplois du bas de la hiérarchie sociale, par rapport à la population majoritaire qui ont une position plus élevée dans la hiérarchie sociale (cela étant, notons que la population majoritaire n'est pas particulièrement associée ici aux positions les plus élevées dans la hiérarchie des PCS, notamment les cadres).

```
##Regarder les profils lignes et les profils colonnes qui contribuent le plus à l'axe 2
#On n'affiche que les 10 premiers pour plus de lisibilité
fviz_contrib(res.afc,choice="row",axes =2,top=10)
```



```
fviz_contrib(res.afc,choice="col",axes =2,top=10)
```

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbsToSbcs': dot substituted  
for <e2>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbsToSbcs': dot substituted  
for <80>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbsToSbcs': dot substituted  
for <99>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Employés d'entreprise' in 'mbsToSbcs': dot substituted  
for <e2>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Employés d'entreprise' in 'mbsToSbcs': dot substituted  
for <80>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Employés d'entreprise' in 'mbcsToSbcs': dot substituted  
for <99>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted  
for <e2>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted  
for <80>

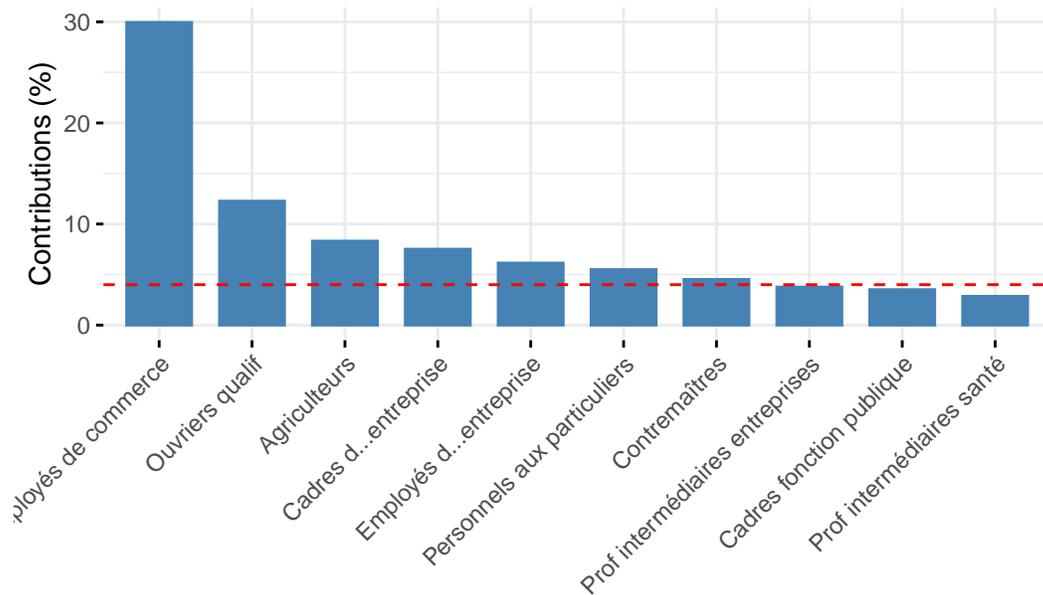
Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted  
for <99>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Employés d'entreprise' in 'mbcsToSbcs': dot substituted  
for <e2>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Employés d'entreprise' in 'mbcsToSbcs': dot substituted  
for <80>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Employés d'entreprise' in 'mbcsToSbcs': dot substituted  
for <99>

## Contribution of columns to Dim-2



```
##Visualiser la contribution des lignes et colonnes qui ont une contribution supérieure à
rowcontrib2<-rownames(res.afc$row$contrib)[res.afc$row$contrib[,2]>100/nrow(tab1)]
colcontrib2<-rownames(res.afc$col$contrib)[res.afc$col$contrib[,2]>100/ncol(tab1)]
fviz_ca_biplot(res.afc,axes=c(1,2),
               select.row = list(name = rowcontrib2),
               select.col=list(name=colcontrib2))
```

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbscsToSbcs': dot substituted  
for <e2>

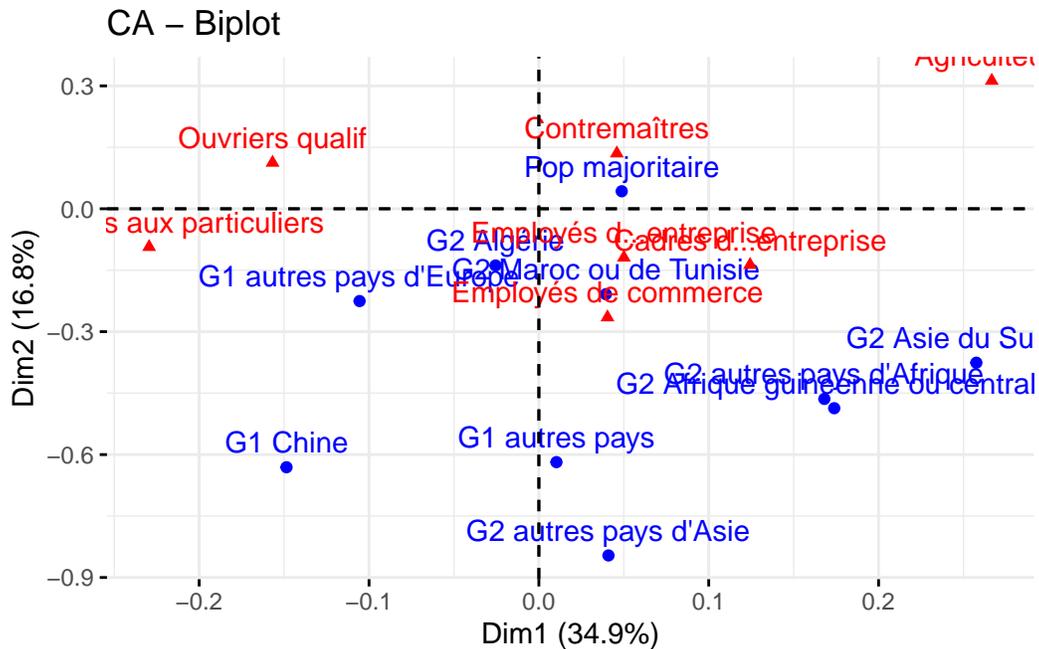
Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbscsToSbcs': dot substituted  
for <80>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbscsToSbcs': dot substituted  
for <99>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Employés d'entreprise' in 'mbscsToSbcs': dot substituted  
for <e2>

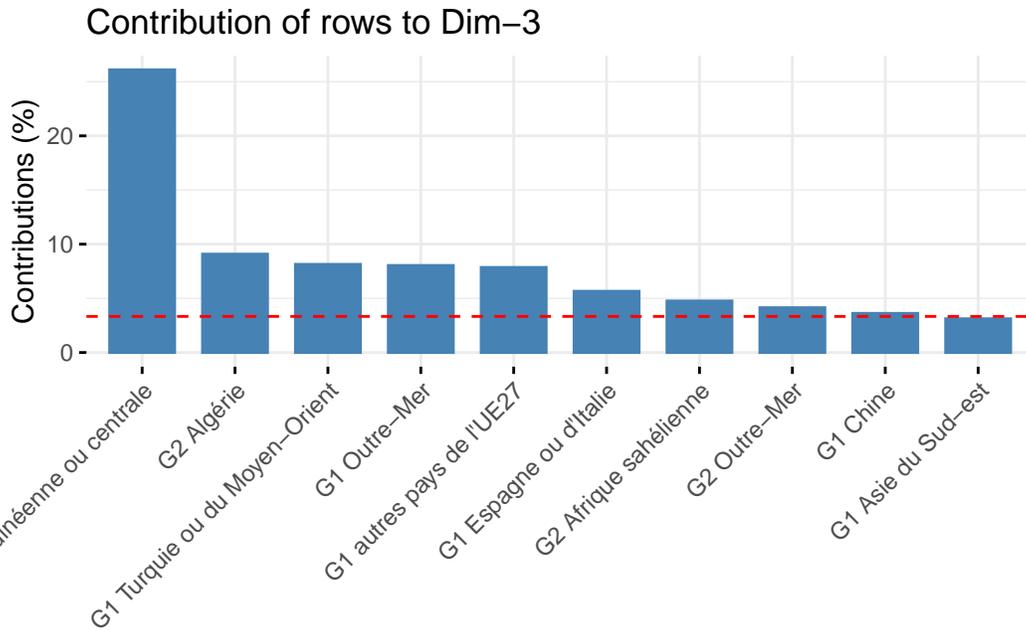
Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Employés d'entreprise' in 'mbscsToSbcs': dot substituted  
for <80>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Employés d'entreprise' in 'mbscsToSbcs': dot substituted  
for <99>



L'axe 2 (vertical) oppose en haut les ouvriers qualifiés, contremaîtres, et agriculteurs, associées à la population majoritaire, par rapport à en bas les employés d'entreprise, de commerce et cadres d'entreprises associés aux immigrés des autres pays (européens notamment) et aux descendants d'immigrés de plusieurs origines. Ici, le bas de la hiérarchie sociale semble plutôt associé à la population majoritaire : il ne faut en effet pas oublier que sans ascendance migratoire, on peut aussi être dans les catégories socioprofessionnelles les plus basses, tandis que les immigrés (non racialisés ?) et les enfants d'immigrés ont tendance à avoir en moyenne une position plus typiquement associée à une position un peu plus élevée dans la hiérarchie sociale, mais dans le pôle "privé" : on ne retrouve en effet pas d'employés du public en bas de l'axe. On peut se demander dans quelle mesure cet axe ne recouvre pas une opposition entre les populations rurales (en haut) et urbaines (en bas).

```
##Regarder les profils lignes et les profils colonnes qui contribuent le plus à l'axe 3
#On n'affiche que les 10 premiers pour plus de lisibilité
fviz_contrib(res.afc,choice="row",axes =3,top=10)
```



```
fviz_contrib(res.afc,choice="col",axes =3,top=10)
```

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbsToSbs': dot substituted for <e2>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbsToSbs': dot substituted for <80>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbsToSbs': dot substituted for <99>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbsToSbs': dot substituted  
for <e2>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted  
for <80>

Warning in grid.Call(C\_textBounds, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted  
for <99>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbcsToSbcs': dot substituted for <e2>

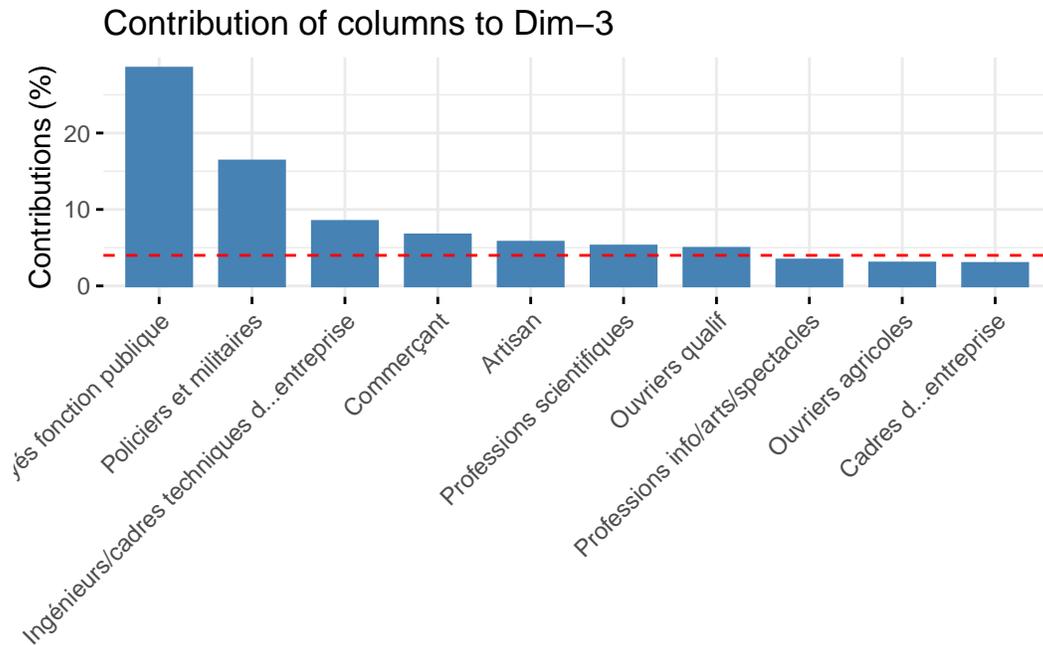
Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbcsToSbcs': dot substituted for <80>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbcsToSbcs': dot substituted for <99>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted  
for <e2>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted  
for <80>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted  
for <99>

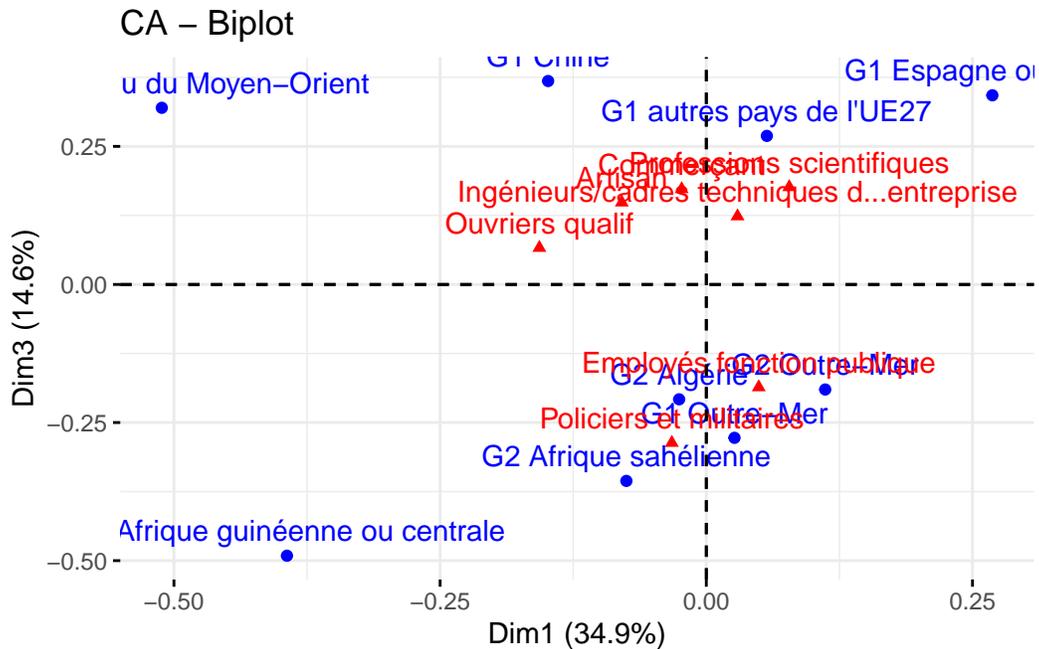


```
##Visualiser la contribution des lignes et colonnes qui ont une contribution supérieure à
rowcontrib3<-rownames(res.afc$row$contrib)[res.afc$row$contrib[,3]>100/nrow(tab1)]
colcontrib3<-rownames(res.afc$col$contrib)[res.afc$col$contrib[,3]>100/ncol(tab1)]
fviz_ca_biplot(res.afc,axes=c(1,3),
               select.row = list(name = rowcontrib3),
               select.col=list(name=colcontrib3))
```

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbscsToSbcs': dot substituted for <e2>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbscsToSbcs': dot substituted for <80>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbscsToSbcs': dot substituted for <99>



L'axe 3 toujours représenté de manière verticale dénote ici une particularité du marché du travail en bas pour les immigrés d'Afrique guinéenne/centrale, les ultra-marins et leurs descendants, et les descendants d'immigrés algériens et d'Afrique sahélienne qui ont des positions d'emploi spécifiquement associés à la "petite fonction publique" ou à la fonction publique policière.

On laisse de côté les autres axes (axe 4 : spécificité des immigrés d'autres pays dans les emplois d'ingénieurs => immigration du "brain drain" possiblement).

L'AFC conduit donc à identifier des spécificités de la position des immigrés et des descendants d'immigrés sur le marché du travail par rapport à la population majoritaire : position plus subalterne pour les immigrés, davantage dans le pôle privé (urbain ?) pour les descendants et dans la petite fonction publique ou policière pour les "racisés". Notons qu'on ne retrouve pas forcément de proximité entre immigrés et descendants d'une même origine : il semble qu'il y ait des effets liés à l'immigration de la position sur le marché du travail, mais pas forcément de "niche" identifiée où une même origine (immigrés comme descendants) seraient cantonnés à un segment particulier du marché du travail (on peut nuancer cette remarque par rapport à l'axe 3 en ne regardant pas seulement les profils-lignes les plus contributeurs, par exemple avec la fonction explor).

Bien sur, cette analyse est ici conduite rapidement et on pourrait creuser l'analyse de manière plus attentive !

## Question 4

A votre avis, pourquoi les profils-lignes sont-ils plus éloignés de l'origine par rapport aux profils-colonnes ? Piste : vous conduirez la même AFC en mettant en ligne supplémentaire le profil-ligne de la « population majoritaire ».

Cette question appelle plusieurs observations :

- Les profils-lignes correspondent à l'origine migratoire et la modalité de la population majoritaire est souvent plus proche du centre des axes par rapport aux immigrés et descendants d'immigrés qui sont davantage éclatés sur les axes.
- Numériquement, la population majoritaire est bien plus importante que celle des immigrés et des descendants : on peut le calculer avec les lignes de code ci-dessous, cette population correspond aux trois-quarts de la population.
- La centralité de la population majoritaire et la dispersion des autres groupes est donc en fait avant tout le résultat de cette configuration démographique !
- Egalement, la population majoritaire est représentée dans à peu près tous les groupes de PCS : ce n'est pas le cas des immigrés et des descendants qui numériquement plus faibles ressortent dans des segments spécifiques du marché du travail.

```
mytable<-as.table(as.matrix(tab1))
```

```
data.frame(round(prop.table(margin.table(mytable, 1))*100,1)) # A frequencies (summed over
```

	Var1	Freq
1	Pop majoritaire	74.3
2	G1 Outre-Mer	1.1
3	G2 Outre-Mer	1.2
4	G1 Algérie	1.7
5	G2 Algérie	2.1
6	G1 Maroc/Tunisie	2.2
7	G2 Maroc ou de Tunisie	2.0
8	G1 Afrique sahélienne	0.7
9	G2 Afrique sahélienne	0.4
10	G1 Afrique guinéenne ou centrale	1.1
11	G2 Afrique guinéenne ou centrale	0.3
12	G1 autres pays d'Afrique	0.6
13	G2 autres pays d'Afrique	0.3
14	G1 Asie du Sud-est	0.3
15	G2 Asie du Sud-est	0.4

16	G1 Turquie ou du Moyen-Orient	0.8
17	G2 Turquie ou du Moyen-Orient	0.4
18	G1 Chine	0.3
19	G1 autres pays d'Asie	0.6
20	G2 autres pays d'Asie	0.1
21	G1 Portugal	1.0
22	G2 Portugal	1.7
23	G1 Espagne ou d'Italie	0.5
24	G2 Espagne ou d'Italie	2.1
25	G1 autres pays de l'UE27	1.1
26	G2 autres pays de l'UE27	0.8
27	G1 autres pays d'Europe	1.0
28	G2 autres pays d'Europe	0.2
29	G1 autres pays	0.6
30	G2 autres pays	0.1

```
data.frame(round(prop.table(margin.table(mytable, 2))*100,1)) # B frequencies (summed over
```

	Var1	Freq
1	Agriculteurs	1.0
2	Artisan	2.6
3	Commerçant	2.3
4	Chefs d'entreprise (>=10 salariés)	0.6
5	Professions libérales	1.9
6	Cadres fonction publique	2.0
7	Professions scientifiques	1.7
8	Professions info/arts/spectacles	1.2
9	Cadres d'entreprise	4.7
10	Ingénieurs/cadres techniques d'entreprise	5.6
11	Professeurs des écoles	4.1
12	Prof intermédiaires santé	5.5
13	Professions intermédiaires fonction publique	1.6
14	Prof intermédiaires entreprises	7.9
15	Techniciens	4.6
16	Contremaîtres	2.9
17	Employés fonction publique	8.3
18	Policiers et militaires	2.0
19	Employés d'entreprise	5.0
20	Employés de commerce	5.0
21	Personnels aux particuliers	7.4
22	Ouvriers qualif	11.3

23	Chauffeurs	2.9
24	Ouvriers non qualif	7.1
25	Ouvriers agricoles	1.0

On peut creuser cette analyse en retirant la population majoritaire de l'analyse, par exemple en mettant ce profil-ligne en supplémentaire. En ce cas, on étudie la segmentation du marché du travail en regardant les immigrés et les descendants les uns par rapport aux autres et non par rapport à la population majoritaire.

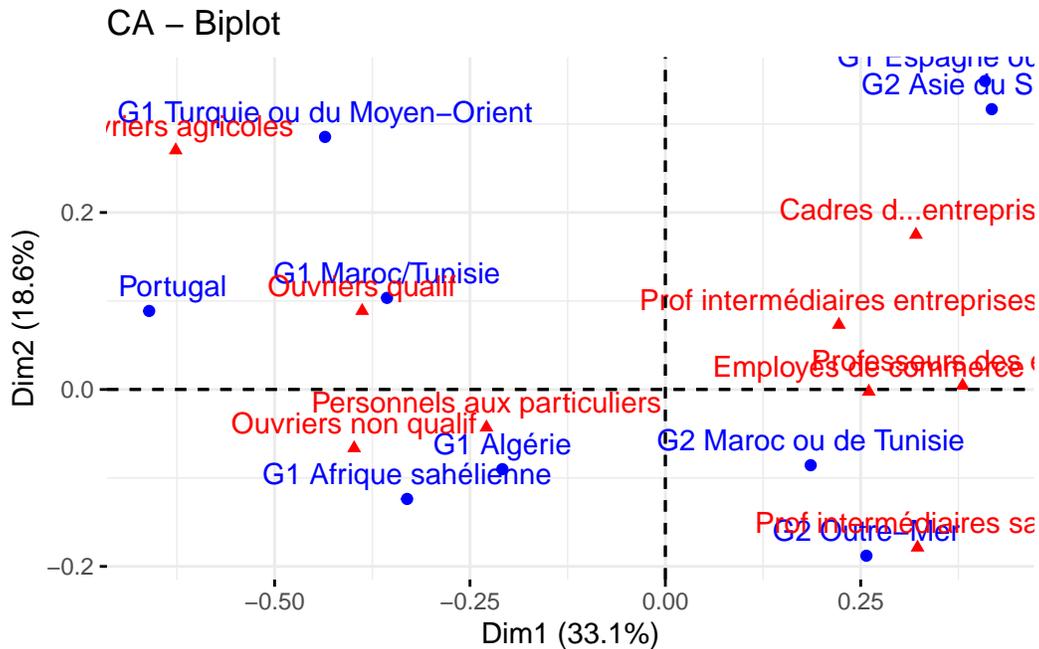
```
res.afc2<-CA(tab1,row.sup=1,graph=F)

rowcontrib1<-rownames(res.afc2$row$contrib)[res.afc2$row$contrib[,1]>100/nrow(tab1)]
colcontrib1<-rownames(res.afc2$col$contrib)[res.afc2$col$contrib[,1]>100/ncol(tab1)]
fviz_ca_biplot(res.afc2,axes=c(1,2),
               select.row = list(name = rowcontrib1),
               select.col=list(name=colcontrib1))
```

```
Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted
for <e2>
```

```
Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted
for <80>
```

```
Warning in grid.Call.graphics(C_text, as.graphicsAnnot(x$label), x$x, x$y, :
conversion failure on 'Cadres d'entreprise' in 'mbcsToSbcs': dot substituted
for <99>
```



```

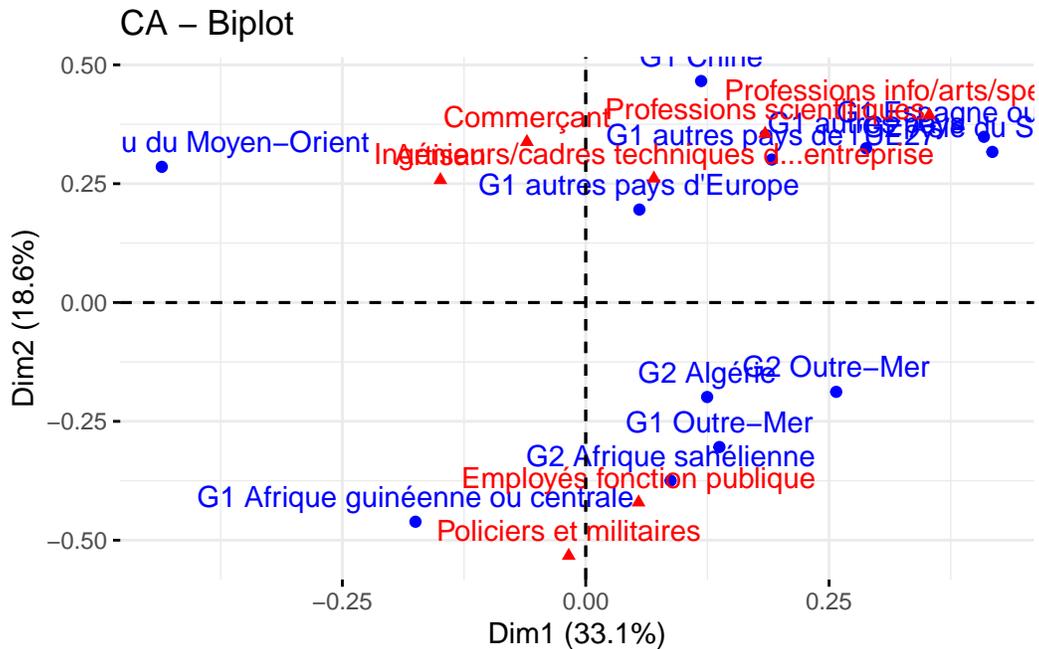
rowcontrib2<-rownames(res.afc2$row$contrib)[res.afc2$row$contrib[,2]>100/nrow(tab1)]
colcontrib2<-rownames(res.afc2$col$contrib)[res.afc2$col$contrib[,2]>100/ncol(tab1)]
fviz_ca_biplot(res.afc2,axes=c(1,2),
               select.row = list(name = rowcontrib2),
               select.col=list(name=colcontrib2))

```

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbscsToSbcs': dot substituted for <e>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbscsToSbcs': dot substituted for <80>

Warning in grid.Call.graphics(C\_text, as.graphicsAnnot(x\$label), x\$x, x\$y, :  
conversion failure on 'Ingénieurs/cadres techniques d'entreprise' in  
'mbscsToSbcs': dot substituted for <99>



## Question 5

Ouvrez le fichier « tab2.csv ». Ce fichier contient des profils-colonnes supplémentaires renseignant le niveau de diplôme des profils-lignes. Vous conduirez une AFC où ces colonnes sont mises en « profils-lignes supplémentaires ». L'une des dimensions de segmentation du marché du travail suivant le lien à la migration est-elle associée au niveau de diplôme ?

Les colonnes supplémentaires sont entre les positions 26 et 33, ce qu'il faut indiquer dans la fonction CA.

Note : ici, on inclut la population majoritaire comme profil-ligne actif.

Avec la fonction `fviz_ca_col`, on voit bien que le premier axe oppose très clairement les non-diplômés aux autres. Sur le deuxième et troisième axe, le diplôme est moins structurant mais les plus diplômés ressortent quand même (en bas sur le 2e et en haut sur le 3e).

```
###Intégrer des colonnes supplémentaires à l'analyse
setwd("/Users/mathieuferry/Documents/Enseignements/Année 2023-2024/L3 Sociologie Analyse de

tab2<-read_csv2("tab2.csv")
```

i Using ",'" as decimal and "'.'" as grouping mark. Use `read_delim()` for more control.

New names:

Rows: 30 Columns: 34

-- Column specification

```
----- Delimiter: ";" chr
(1): ...1 dbl (33): Agriculteurs, Artisan, Commerçant, Chefs d'entreprise (>=10
salari...
i Use `spec()` to retrieve the full column specification for this data. i
Specify the column types or set `show_col_types = FALSE` to quiet this message.
* `` -> `...1`
```

```
tab2<-column_to_rownames(tab2,var="...1")
```

```
rownames(tab2)
```

```
[1] "Pop majoritaire"           "G1 Outre-Mer"
[3] "G2 Outre-Mer"             "G1 Algérie"
[5] "G2 Algérie"               "G1 Maroc/Tunisie"
[7] "G2 Maroc ou de Tunisie"   "G1 Afrique sahélienne"
[9] "G2 Afrique sahélienne"    "G1 Afrique guinéenne ou centrale"
[11] "G2 Afrique guinéenne ou centrale" "G1 autres pays d'Afrique"
[13] "G2 autres pays d'Afrique"  "G1 Asie du Sud-est"
[15] "G2 Asie du Sud-est"       "G1 Turquie ou du Moyen-Orient"
[17] "G2 Turquie ou du Moyen-Orient" "G1 Chine"
[19] "G1 autres pays d'Asie"    "G2 autres pays d'Asie"
[21] "G1 Portugal"              "G2 Portugal"
[23] "G1 Espagne ou d'Italie"   "G2 Espagne ou d'Italie"
[25] "G1 autres pays de l'UE27" "G2 autres pays de l'UE27"
[27] "G1 autres pays d'Europe"  "G2 autres pays d'Europe"
[29] "G1 autres pays"          "G2 autres pays"
```

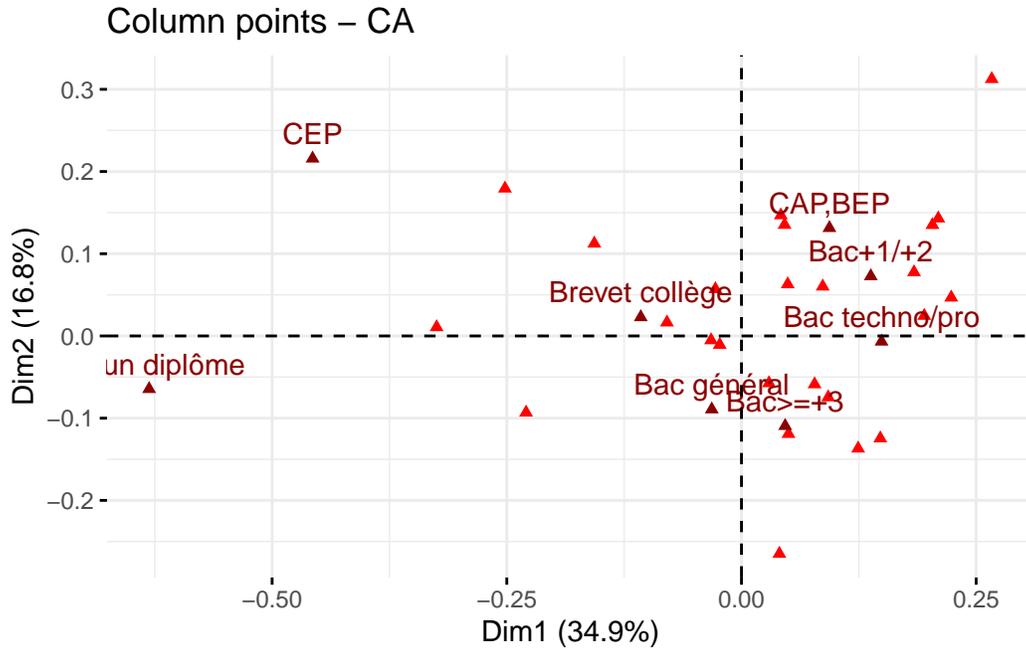
```
colnames(tab2)
```

```
[1] "Agriculteurs"
[2] "Artisan"
[3] "Commerçant"
[4] "Chefs d'entreprise (>=10 salariés)"
[5] "Professions libérales"
[6] "Cadres fonction publique"
[7] "Professions scientifiques"
[8] "Professions info/arts/spectacles"
```

- [9] "Cadres d'entreprise"
- [10] "Ingénieurs/cadres techniques d'entreprise"
- [11] "Professeurs des écoles"
- [12] "Prof intermédiaires santé"
- [13] "Professions intermédiaires fonction publique"
- [14] "Prof intermédiaires entreprises"
- [15] "Techniciens"
- [16] "Contremaîtres"
- [17] "Employés fonction publique"
- [18] "Policiers et militaires"
- [19] "Employés d'entreprise"
- [20] "Employés de commerce"
- [21] "Personnels aux particuliers"
- [22] "Ouvriers qualif"
- [23] "Chauffeurs"
- [24] "Ouvriers non qualif"
- [25] "Ouvriers agricoles"
- [26] "Aucun diplôme"
- [27] "CEP"
- [28] "Brevet collègue"
- [29] "CAP,BEP"
- [30] "Bac techno/pro"
- [31] "Bac général"
- [32] "Bac+1/+2"
- [33] "Bac>=+3"

```
res.afc3<-CA(tab2,col.sup=26:33,graph=F)
```

```
fviz_ca_col(res.afc3,axes=c(1,2),label="col.sup")
```



```
fviz_ca_col(res.afc3, axes=c(1,3), label="col.sup")
```

