#Revue française de science politique, Vol. 69, N° 4, 2019

#"Devenir(s) militants. Proposition de méthode pour une exploration des conséquences biographiques de l'engagement des soixante-huitard.e.s français.e.s"

#Code pour obtenir les ACM et l'AFM

#Thierry Rossier

#Olivier Fillieule

#Remerciements en particulier à Olivier Robette et Matthias Studer pour leur aide précieuse. Et à Davide Morselli pour avoir passé en revue ce code.

####Lancer le fichier "Rossier\_et\_Fillieule\_RFSP\_Data.RDATA"

####Pour toute question écrire à thierry.rossier.elites@gmail.com

####Installation packages

#install.packages("soc.ca")

#install.packages("GDAtools")

#install.packages("FactoMineR")

#install.packages("cluster")

#install.packages("WeightedCluster")

#install.packages("dplyr")

####Chargement des packages

library(soc.ca) #ACMs spécifiques

library(GDAtools) #ACMs spécifiques et AFM

library(FactoMineR) #Caractérisation des clusters

library(cluster) #Clustering de Ward

library(WeightedCluster) #Recalcul de partitionning around medoids

library(dplyr) #Recodage des clusters

####################################################

####"Annexe 5. Analyse des correspondances multiples spécifiques pour les quatre périodes"

#Nous utilisons le package soc.ca. Utilisant le moteur graphique de ggplot2, nous le préférons à d'autres packages d'ACM.

#Modalité passive pour les quatre ACM

options(passive = c("Milit: Non"))

####ACM 60\_74

#Objet ACM

soc.mca\_Sombrero\_60\_74 <- soc.mca(active\_60\_74[2:9])

#Contrôle modalité passive

soc.mca\_Sombrero\_60\_74$names.passive

#Graphique axes 1 et 2

map.active(soc.mca\_Sombrero\_60\_74, dim = c(1, 2), point.shape = "variable", point.alpha = 1, point.fill = "black", point.color = "black", point.size = "freq", label = TRUE, label.repel = TRUE, label.alpha = 1, label.color = "black", label.size = 4, label.fill = NULL, map.title = "60\_74 Axes 1-2", labelx = "default", labely = "default", legend = NULL)

####ACM 75\_83

#Objet ACM

soc.mca\_Sombrero\_75\_83 <- soc.mca(active\_75\_83[2:11])

#Contrôle modalité passive

soc.mca\_Sombrero\_75\_83$names.passive

#Graphique axes 1 et 2

map.active(soc.mca\_Sombrero\_75\_83, dim = c(1, 2), point.shape = "variable", point.alpha = 1, point.fill = "black", point.color = "black", point.size = "freq", label = TRUE, label.repel = TRUE, label.alpha = 1, label.color = "black", label.size = 4, label.fill = NULL, map.title = "75\_83 Axes 1-2", labelx = "default", labely = "default", legend = NULL)

####ACM 84\_94

#Objet ACM

soc.mca\_Sombrero\_84\_94 <- soc.mca(active\_84\_94[2:11])

#Contrôle modalité passive

soc.mca\_Sombrero\_84\_94$names.passive

#Graphique axes 1 et 2

map.active(soc.mca\_Sombrero\_84\_94, dim = c(1, 2), point.shape = "variable", point.alpha = 1, point.fill = "black", point.color = "black", point.size = "freq", label = TRUE, label.repel = TRUE, label.alpha = 1, label.color = "black", label.size = 4, label.fill = NULL, map.title = "84\_94 Axes 1-2", labelx = "default", labely = "default", legend = NULL)

####ACM 95\_15

#Objet ACM

soc.mca\_Sombrero\_95\_15 <- soc.mca(active\_95\_15[2:11])

#Contrôle modalité passive

soc.mca\_Sombrero\_95\_15$names.passive

#Graphique axes 1 et 2

map.active(soc.mca\_Sombrero\_95\_15, dim = c(1, 2), point.shape = "variable", point.alpha = 1, point.fill = "black", point.color = "black", point.size = "freq", label = TRUE, label.repel = TRUE, label.alpha = 1, label.color = "black", label.size = 4, label.fill = NULL, map.title = "95\_15 Axes 1-2", labelx = "default", labely = "default", legend = NULL)

####################################################

####Refaire les ACM avec GDAtools, en vue de l'AFM. Les variables actives sont renommées pour les graphiques de l'AFM

active\_60\_74\_names\_AFM <- active\_60\_74\_names[2:9]

getindexcat(active\_60\_74\_names\_AFM)

speACM\_Sombrero\_60\_74\_names <- speMCA(active\_60\_74\_names\_AFM, ncp=10, excl=26)

active\_75\_83\_names\_AFM <- active\_75\_83\_names[2:11]

getindexcat(active\_75\_83\_names\_AFM)

speACM\_Sombrero\_75\_83\_names <- speMCA(active\_75\_83\_names\_AFM, ncp=10, excl=28)

active\_84\_94\_names\_AFM <- active\_84\_94\_names[2:11]

getindexcat(active\_84\_94\_names\_AFM)

speACM\_Sombrero\_84\_94\_names <- speMCA(active\_84\_94\_names\_AFM, ncp=10, excl=24)

active\_95\_15\_names\_AFM <- active\_95\_15\_names[2:11]

getindexcat(active\_95\_15\_names\_AFM)

speACM\_Sombrero\_95\_15\_names <- speMCA(active\_95\_15\_names\_AFM, ncp=10, excl=26)

####################################################

####AFM

MCA\_4\_periodes\_names <- list(speACM\_Sombrero\_60\_74\_names, speACM\_Sombrero\_75\_83\_names, speACM\_Sombrero\_84\_94\_names, speACM\_Sombrero\_95\_15\_names)

MultiMCA\_names <- multiMCA(MCA\_4\_periodes\_names, ncp = 10, compute.rv = TRUE)

####"Tableau 2. Contribution des quatre ACM aux trois premiers axes de l'AFM"

round(MultiMCA\_names$group$contrib[, 1:3],1)

####"Figure 3. Analyse factorielle multiple 1960-2015, nuage des modalités, axes 1 et 2"

plot(MultiMCA\_names, type = "v", col=c('black', 'gray20', 'gray40', 'gray60'), axes = 1:2, points = "all", app = 2)

####Modalités notables (Merci à Nicolas Robette et Olivier Roueff pour cette fonction. Voir: "L’espace contemporain des goûts culturels", in Sociologie, Vol. 8, N° 4, 2017, pp. 369 à 394)

temp <- do.call('rbind.data.frame', lapply(MultiMCA\_names$VAR, function(x) x$coord))

aides <- list()

for(i in 1:ncol(temp)) {

x <- temp[, i]

names(x) <- substr(rownames(temp), 6, nchar(rownames(temp)))

x <- x[order(x)]

aides[[i]] <- x[abs(x)>=0.5]

}

####"Tableau 3. Modalités notablement et très associées (en gras) à l'Axe 1 de l'AFM, par période"

aides[[1]]

####Variables supplémentaires

####"Tableau 4. Pourcentage de la variance des trois premiers axes de l'AFM expliquée par les variables supplémentaires (eta2)"

l <- list(sup$Sexe, sup$Pere, sup$Cohorte, sup$Mobilite, sup$MobiliteThelot)

n <- c("Sexe", "Pere", "Cohorte", "Mobilite", "Mobilite Thelot")

dimeta2(MultiMCA\_names, l, n, dim=1:3)

####Création d'objets en vue de produire des statistiques sur les variables supplémentaires

Varsup\_Sexe <- varsup(MultiMCA\_names, sup$Sexe)

Varsup\_Pere <- varsup(MultiMCA\_names, sup$Pere)

Varsup\_Cohorte <- varsup(MultiMCA\_names, sup$Cohorte)

Varsup\_MobiliteThelot <- varsup(MultiMCA\_names, sup$MobiliteThelot)

Varsup\_Mobilite <- varsup(MultiMCA\_names, sup$Mobilite)

####"Tableau 5. Coordonnées des modalités des variables supplémentaires sur les 3 premiers axes de l'AFM

#Coordonnées sur les trois axes"

Varsup\_Sexe$coord[,1:3]

Varsup\_Pere$coord[,1:3]

Varsup\_Cohorte$coord[,1:3]

Varsup\_Mobilite$coord[,1:3]

Varsup\_MobiliteThelot$coord[,1:3]

#Fréquences des variables

table(sup$Sexe)

table(sup$Pere)

table(sup$Cohorte)

table(sup$MobiliteThelot)

table(sup$Mobilite)

####"Tableau 6. Modalités notablement et très associées (en gras) à l'Axe 2 de l'AFM, par période"

aides[[2]]

####"Figure 4. Analyse factorielle multiple 1960-2015, nuage des modalités, axes 1 et 3"

plot(MultiMCA\_names, type = "v", col=c('black', 'gray20', 'gray40', 'gray60'), axes = c(1,3), points = "all", app = 2)

####"Tableau 7. Modalités notablement et très associées (en gras) à l'Axe 3 de l'AFM, par période"

aides[[3]]

####################################################

####Clustering

#Procédure de Ward sur les dix premiers axes

coords <- MultiMCA\_names$ind$coord[, 1:10]

clust <- agnes(coords, method="ward")

#Partition en 5

clusters\_5 <- cutree(clust, k = 5)

clusters\_5 <- factor(clusters\_5)

table(clusters\_5)

#Procédure de Partitioning around medoids (PAM)

diss <- dist(MultiMCA\_names$ind$coord[,1:10], method = "euclidean")

PAM\_5 <- wcKMedoids(diss, 5, initialclust = clust)

table(PAM\_5$clustering)

#Export des 5 clusters PAM dans un data.frame

Clusters <- data.frame(PAM\_5$clustering)

#Recodage des 5 clusters PAM

Clusters$PAM\_5 <- factor(Clusters$PAM\_5.clustering)

Clusters$PAM\_5 <- recode(Clusters$PAM\_5,

"8" = "Cluster 1",

"21" = "Cluster 2",

"140" = "Cluster 3",

"223" = "Cluster 4",

"271" = "Cluster 5")

table(Clusters$PAM\_5)

#Création d'un data.frame avec toutes les variables actives et supplémentaires utilisées pour qualifier les 5 clusters

Clusters\_Qualification\_5 <- data.frame(active\_60\_74\_names[,2:9], active\_75\_83\_names[,2:11], active\_84\_94\_names[,2:11], active\_95\_15\_names[,2:11], sup[,2:6], sup\_autres[,2:6], Clusters$PAM\_5)

names(Clusters\_Qualification\_5)

####"Annexe 6. Caractérisation des cinq clusters de l'AFM. Modalités asociées"

catdes(Clusters\_Qualification\_5,49,proba = 0.05)