Cardsort analysis

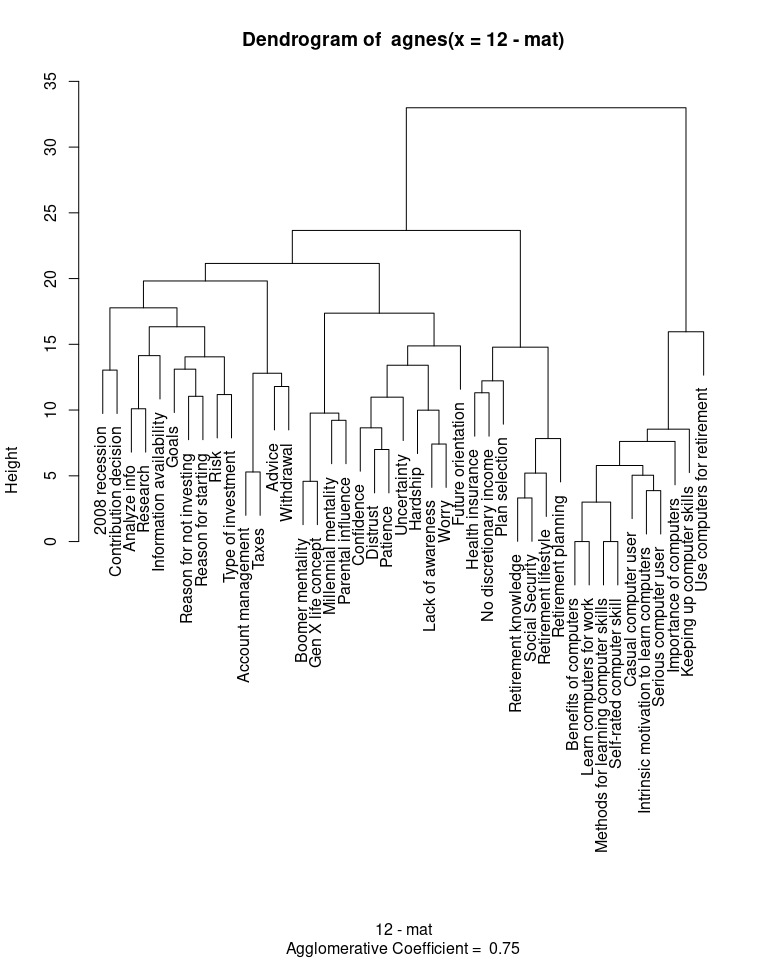
Shane Mueller

4/14/2021

library(MASS)  
data <- read.csv("natasha\_data.csv")  
tab1 <- table(data$participant,data$card.label,data$category.label)  
  
mat <- matrix(0,dim(tab1)[2],dim(tab1)[2])  
rownames(mat) <- colnames(mat) <- rownames(tab1[1,,])  
for(sub in 1:dim(tab1)[1])  
{  
   
 t1 <- tab1[sub,,]  
 byindex <- apply(t1,1,function(x){which(x==1)})  
 submat <- outer(byindex,byindex,"==")  
 mat <- mat + submat  
  
}

## look at a clustering

library(cluster)  
a <- agnes(12-mat)  
plot(a,which=2)



## K-means cluster

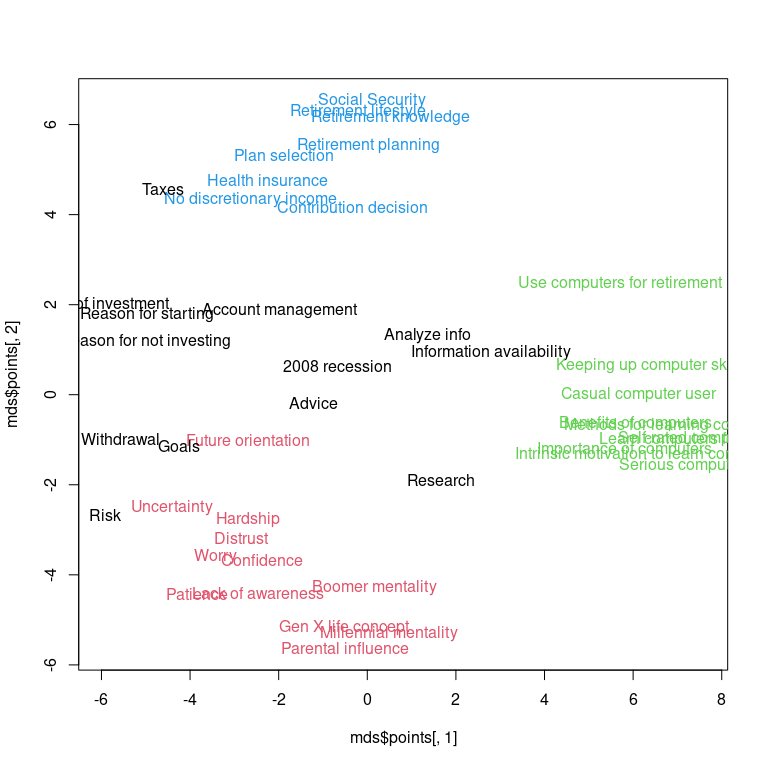
km <-kmeans(12-mat,centers=4)  
km$cluster

## 2008 recession Account management   
## 1 1   
## Advice Analyze info   
## 1 1   
## Benefits of computers Boomer mentality   
## 3 2   
## Casual computer user Confidence   
## 3 2   
## Contribution decision Distrust   
## 4 2   
## Future orientation Gen X life concept   
## 2 2   
## Goals Hardship   
## 1 2   
## Health insurance Importance of computers   
## 4 3   
## Information availability Intrinsic motivation to learn computers   
## 1 3   
## Keeping up computer skills Lack of awareness   
## 3 2   
## Learn computers for work Methods for learning computer skills   
## 3 3   
## Millennial mentality No discretionary income   
## 2 4   
## Parental influence Patience   
## 2 2   
## Plan selection Reason for not investing   
## 4 1   
## Reason for starting Research   
## 1 1   
## Retirement knowledge Retirement lifestyle   
## 4 4   
## Retirement planning Risk   
## 4 1   
## Self-rated computer skill Serious computer user   
## 3 3   
## Social Security Taxes   
## 4 1   
## Type of investment Uncertainty   
## 1 2   
## Use computers for retirement Withdrawal   
## 3 1   
## Worry   
## 2

mds <- isoMDS(12-mat,k=2)

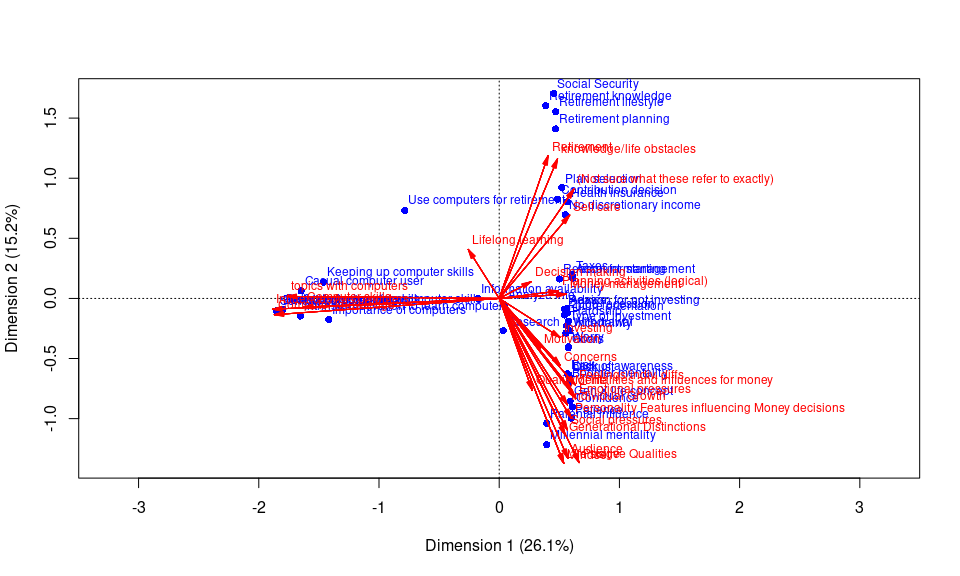
## initial value 31.345692   
## iter 5 value 26.590630  
## iter 10 value 25.497491  
## iter 15 value 24.966145  
## iter 20 value 24.547825  
## iter 25 value 23.887370  
## iter 30 value 23.520014  
## final value 23.346122   
## converged

plot(mds$points[,1],mds$points[,2],type="n")  
text(mds$points[,1],mds$points[,2],rownames(mat),col=km$cluster)



## Correspondence analysis

library(ca)  
ca <- ca(table(as.factor(data$card.label),as.factor(data$category.label)))  
plot(ca,arrows=c(F,T),xlim=c(-2,2))



par(mfrow=c(1,2))  
plot(ca,dim=c(1,2), what=c("none","all"),main="Cluster names",xlim=c(-2,2))  
plot(ca,dim=c(1,2), what=c("all","none"),main="Sorted topics",xlim=c(-2,2))

