Module 5 assignment

Benjan

11/8/2022

# 1.

Lets read the data first.

urlRemote <- "https://raw.githubusercontent.com/"  
pathGithub <- "EricBrownTTU/ISQS6350/main/"  
filename <- "crime.csv"  
crime <- read.csv(paste0(urlRemote, pathGithub, filename))

Now, we calculate the distance.

d <- round(dist(crime,diag=TRUE,upper=TRUE),2)

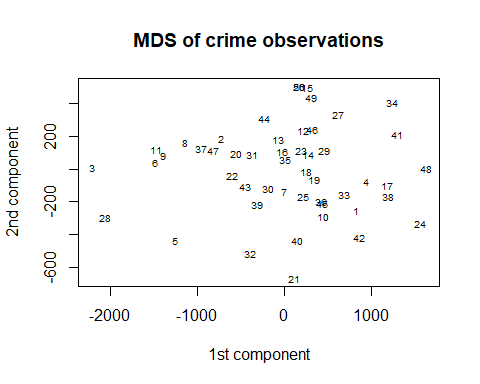
## Warning in dist(crime, diag = TRUE, upper = TRUE): NAs introduced by coercion

We will use the distance ‘d’ to perform MDS analysis.

crime\_mds <- cmdscale(d)

Let us plot the cmd data.

plot(crime\_mds, type = "n",  
 xlab = "1st component", ylab = "2nd component",  
 main = "MDS of crime observations")  
text(crime\_mds, labels = rownames(crime),cex = 0.6)



# 2

Lets perform the MDS on correlation data:

crime\_p <- cor(crime[-1])  
crime\_p

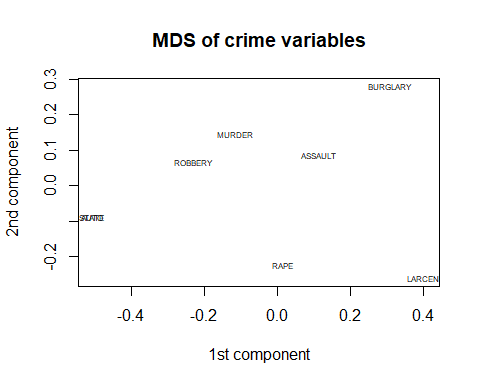
## MURDER RAPE ROBBERY ASSAULT BURGLARY LARCENY  
## MURDER 1.00000000 0.6012205 0.4837076 0.6485505 0.3858168 0.1019198  
## RAPE 0.60122047 1.0000000 0.5918793 0.7402595 0.7121301 0.6139882  
## ROBBERY 0.48370757 0.5918793 1.0000000 0.5570782 0.6372420 0.4467399  
## ASSAULT 0.64855048 0.7402595 0.5570782 1.0000000 0.6229085 0.4043633  
## BURGLARY 0.38581683 0.7121301 0.6372420 0.6229085 1.0000000 0.7921210  
## LARCENY 0.10191983 0.6139882 0.4467399 0.4043633 0.7921210 1.0000000  
## AUTO 0.06881448 0.3489015 0.5906795 0.2758426 0.5579533 0.4441799  
## AUTO  
## MURDER 0.06881448  
## RAPE 0.34890153  
## ROBBERY 0.59067951  
## ASSAULT 0.27584265  
## BURGLARY 0.55795326  
## LARCENY 0.44417992  
## AUTO 1.00000000

crime.mds <- cmdscale(1-abs(crime\_p)) #mds distance  
crime.mds

## [,1] [,2]  
## MURDER -0.50659827 -0.08906733  
## RAPE -0.11345197 0.14540154  
## ROBBERY 0.01683486 -0.22426319  
## ASSAULT -0.22884495 0.06494758  
## BURGLARY 0.11510613 0.08538362  
## LARCENY 0.31018514 0.27927327  
## AUTO 0.40676907 -0.26167550

We will plot the data now.

plot(crime.mds,type="n",xlab = "1st component", ylab = "2nd component",main = "MDS of crime variables")  
text(crime.mds,labels=colnames(crime),cex=0.5)



# 3.

The data gives us a visualization of similar and dissimilar variables. As *Murder*, *Robbery* and *Assault* are closer to each other, they are similar in terms of correlation.