Assignment\_4

2022-10-29

DF <- read.csv("~/Downloads/Pharmaceuticals.csv")

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──  
## ✔ ggplot2 3.3.5 ✔ purrr 0.3.4   
## ✔ tibble 3.1.6 ✔ dplyr 1.0.10  
## ✔ tidyr 1.2.0 ✔ stringr 1.4.0   
## ✔ readr 2.1.2 ✔ forcats 0.5.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(factoextra)

## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa

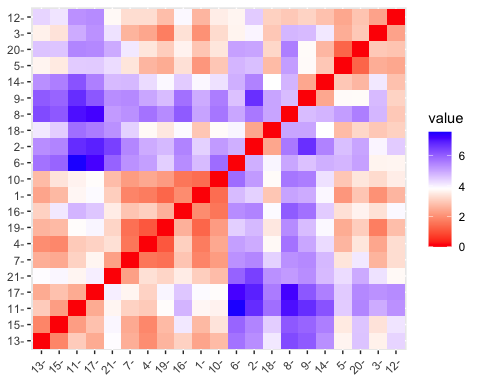
library(ISLR)  
library(flexclust)

## Loading required package: grid  
## Loading required package: lattice  
## Loading required package: modeltools  
## Loading required package: stats4

df <- DF[,c(3:11)]  
summary(df)

## Market\_Cap Beta PE\_Ratio ROE   
## Min. : 0.41 Min. :0.1800 Min. : 3.60 Min. : 3.9   
## 1st Qu.: 6.30 1st Qu.:0.3500 1st Qu.:18.90 1st Qu.:14.9   
## Median : 48.19 Median :0.4600 Median :21.50 Median :22.6   
## Mean : 57.65 Mean :0.5257 Mean :25.46 Mean :25.8   
## 3rd Qu.: 73.84 3rd Qu.:0.6500 3rd Qu.:27.90 3rd Qu.:31.0   
## Max. :199.47 Max. :1.1100 Max. :82.50 Max. :62.9   
## ROA Asset\_Turnover Leverage Rev\_Growth   
## Min. : 1.40 Min. :0.3 Min. :0.0000 Min. :-3.17   
## 1st Qu.: 5.70 1st Qu.:0.6 1st Qu.:0.1600 1st Qu.: 6.38   
## Median :11.20 Median :0.6 Median :0.3400 Median : 9.37   
## Mean :10.51 Mean :0.7 Mean :0.5857 Mean :13.37   
## 3rd Qu.:15.00 3rd Qu.:0.9 3rd Qu.:0.6000 3rd Qu.:21.87   
## Max. :20.30 Max. :1.1 Max. :3.5100 Max. :34.21   
## Net\_Profit\_Margin  
## Min. : 2.6   
## 1st Qu.:11.2   
## Median :16.1   
## Mean :15.7   
## 3rd Qu.:21.1   
## Max. :25.5

df <- scale(df)  
distance <- get\_dist(df)  
fviz\_dist(distance)



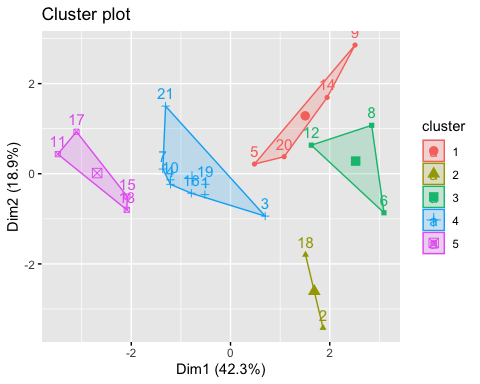
k4 <-kmeans(df, centers = 5, nstart = 10)  
k4$center

## Market\_Cap Beta PE\_Ratio ROE ROA Asset\_Turnover  
## 1 -0.76022489 0.2796041 -0.47742380 -0.7438022 -0.8107428 -1.2684804  
## 2 -0.43925134 -0.4701800 2.70002464 -0.8349525 -0.9234951 0.2306328  
## 3 -0.87051511 1.3409869 -0.05284434 -0.6184015 -1.1928478 -0.4612656  
## 4 -0.03142211 -0.4360989 -0.31724852 0.1950459 0.4083915 0.1729746  
## 5 1.69558112 -0.1780563 -0.19845823 1.2349879 1.3503431 1.1531640  
## Leverage Rev\_Growth Net\_Profit\_Margin  
## 1 0.06308085 1.5180158 -0.006893899  
## 2 -0.14170336 -0.1168459 -1.416514761  
## 3 1.36644699 -0.6912914 -1.320000179  
## 4 -0.27449312 -0.7041516 0.556954446  
## 5 -0.46807818 0.4671788 0.591242521

k4$size

## [1] 4 2 3 8 4

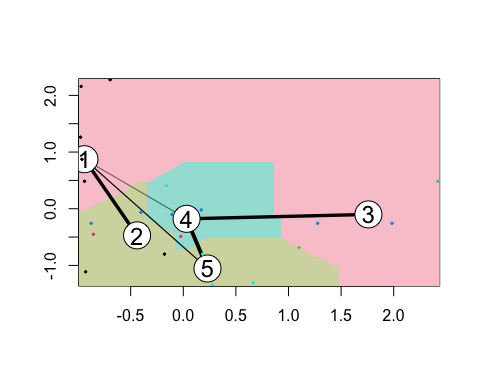
fviz\_cluster(k4, data = df)



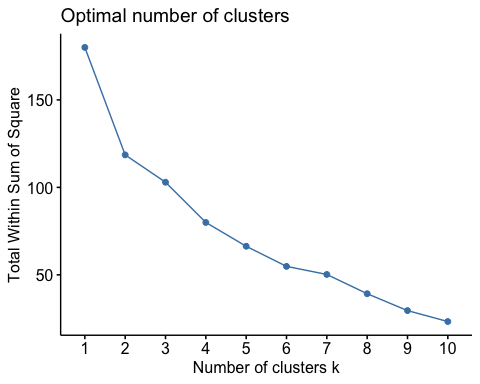
k4 = kcca(df, k=5, kccaFamily("kmedians"))  
k4

## kcca object of family 'kmedians'   
##   
## call:  
## kcca(x = df, k = 5, family = kccaFamily("kmedians"))  
##   
## cluster sizes:  
##   
## 1 2 3 4 5   
## 7 2 2 6 4

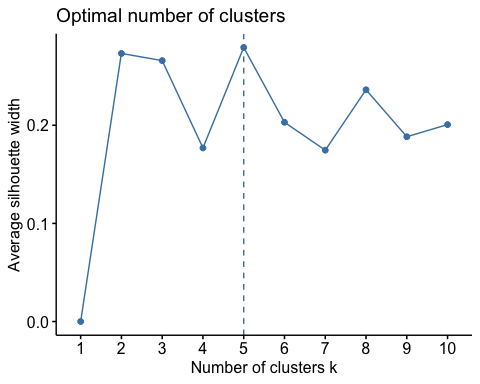
clusters\_index <- predict(k4)  
image(k4)  
points(df, col=clusters\_index, pch=19, cex=0.3)



fviz\_nbclust(df, kmeans, method = "wss")



fviz\_nbclust(df, kmeans, method = "silhouette")

 B. Cluster 1 has lowest Leverage.

Cluster 2 has highest PE ratio, lowest ROE, lowest ROA, lowest Asset Turnover, lowest Net Profit Margin

Cluster 3 has highest Market Cap, highest ROE, highest ROA, highest Asset Turnover

Cluster 4 has highest Beta, highest Leverage, highest Rev growth and lowest Market Cap.

Cluster 5 has highest Net Profit Margin, lowest Beta, lowest PE Ratio, lowest Rev growth.

C. There is pattern in the clusters with respect to Media recommendation variable. Cluster 5 with highest Net Profit Margin, lowest Beta, lowest PE Ratio, lowest Rev growth have mostly Hold recommendation.

D. We can name various clusters based on their dependence on the quantitative variables.

Cluster 1 - Lowest Leverage cluster

Cluster 2 - High PE ratio, Low ROE, ROA, Asset Turnover, Net Profit Margin Cluster

Cluster 3 - High Market Cap, ROE, ROA, Asset Turnover cluster

Cluster 4 - High Beta, Leverage, Rev growth and Low Market Cap cluster

Cluster 5 - High Net Profit Margin, Low Beta, PE Ratio, Rev growth cluster