Assignment\_4 (FML)

“ADITI”

2024-03-17

#read the csv file  
pharma\_data <- read.csv("C:/Users/tiwar/OneDrive/Desktop/FML ASSIGNMENT/Pharmaceuticals.csv")  
  
#read the first five rows of pharma\_data  
head(pharma\_data)

## Symbol Name Market\_Cap Beta PE\_Ratio ROE ROA Asset\_Turnover  
## 1 ABT Abbott Laboratories 68.44 0.32 24.7 26.4 11.8 0.7  
## 2 AGN Allergan, Inc. 7.58 0.41 82.5 12.9 5.5 0.9  
## 3 AHM Amersham plc 6.30 0.46 20.7 14.9 7.8 0.9  
## 4 AZN AstraZeneca PLC 67.63 0.52 21.5 27.4 15.4 0.9  
## 5 AVE Aventis 47.16 0.32 20.1 21.8 7.5 0.6  
## 6 BAY Bayer AG 16.90 1.11 27.9 3.9 1.4 0.6  
## Leverage Rev\_Growth Net\_Profit\_Margin Median\_Recommendation Location Exchange  
## 1 0.42 7.54 16.1 Moderate Buy US NYSE  
## 2 0.60 9.16 5.5 Moderate Buy CANADA NYSE  
## 3 0.27 7.05 11.2 Strong Buy UK NYSE  
## 4 0.00 15.00 18.0 Moderate Sell UK NYSE  
## 5 0.34 26.81 12.9 Moderate Buy FRANCE NYSE  
## 6 0.00 -3.17 2.6 Hold GERMANY NYSE

#read the last five rows of pharma\_data  
tail(pharma\_data)

## Symbol Name Market\_Cap Beta PE\_Ratio ROE ROA  
## 16 NVS Novartis AG 96.65 0.19 21.6 17.9 11.2  
## 17 PFE Pfizer Inc 199.47 0.65 23.6 45.6 19.2  
## 18 PHA Pharmacia Corporation 56.24 0.40 56.5 13.5 5.7  
## 19 SGP Schering-Plough Corporation 34.10 0.51 18.9 22.6 13.3  
## 20 WPI Watson Pharmaceuticals, Inc. 3.26 0.24 18.4 10.2 6.8  
## 21 WYE Wyeth 48.19 0.63 13.1 54.9 13.4  
## Asset\_Turnover Leverage Rev\_Growth Net\_Profit\_Margin Median\_Recommendation  
## 16 0.5 0.06 -2.69 22.4 Hold  
## 17 0.8 0.16 25.54 25.2 Moderate Buy  
## 18 0.6 0.35 15.00 7.3 Hold  
## 19 0.8 0.00 8.56 17.6 Hold  
## 20 0.5 0.20 29.18 15.1 Moderate Sell  
## 21 0.6 1.12 0.36 25.5 Hold  
## Location Exchange  
## 16 SWITZERLAND NYSE  
## 17 US NYSE  
## 18 US NYSE  
## 19 US NYSE  
## 20 US NYSE  
## 21 US NYSE

#nrow and ncol will return both no of rows and no of columns in pharma\_data  
nrow(pharma\_data)

## [1] 21

ncol(pharma\_data)

## [1] 14

#selecting all the columns -  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

pharma\_data1 <- select\_if(pharma\_data,is.numeric)  
pharma\_data1 <- pharma\_data1[1:9]  
pharma\_data1

## Market\_Cap Beta PE\_Ratio ROE ROA Asset\_Turnover Leverage Rev\_Growth  
## 1 68.44 0.32 24.7 26.4 11.8 0.7 0.42 7.54  
## 2 7.58 0.41 82.5 12.9 5.5 0.9 0.60 9.16  
## 3 6.30 0.46 20.7 14.9 7.8 0.9 0.27 7.05  
## 4 67.63 0.52 21.5 27.4 15.4 0.9 0.00 15.00  
## 5 47.16 0.32 20.1 21.8 7.5 0.6 0.34 26.81  
## 6 16.90 1.11 27.9 3.9 1.4 0.6 0.00 -3.17  
## 7 51.33 0.50 13.9 34.8 15.1 0.9 0.57 2.70  
## 8 0.41 0.85 26.0 24.1 4.3 0.6 3.51 6.38  
## 9 0.78 1.08 3.6 15.1 5.1 0.3 1.07 34.21  
## 10 73.84 0.18 27.9 31.0 13.5 0.6 0.53 6.21  
## 11 122.11 0.35 18.0 62.9 20.3 1.0 0.34 21.87  
## 12 2.60 0.65 19.9 21.4 6.8 0.6 1.45 13.99  
## 13 173.93 0.46 28.4 28.6 16.3 0.9 0.10 9.37  
## 14 1.20 0.75 28.6 11.2 5.4 0.3 0.93 30.37  
## 15 132.56 0.46 18.9 40.6 15.0 1.1 0.28 17.35  
## 16 96.65 0.19 21.6 17.9 11.2 0.5 0.06 -2.69  
## 17 199.47 0.65 23.6 45.6 19.2 0.8 0.16 25.54  
## 18 56.24 0.40 56.5 13.5 5.7 0.6 0.35 15.00  
## 19 34.10 0.51 18.9 22.6 13.3 0.8 0.00 8.56  
## 20 3.26 0.24 18.4 10.2 6.8 0.5 0.20 29.18  
## 21 48.19 0.63 13.1 54.9 13.4 0.6 1.12 0.36  
## Net\_Profit\_Margin  
## 1 16.1  
## 2 5.5  
## 3 11.2  
## 4 18.0  
## 5 12.9  
## 6 2.6  
## 7 20.6  
## 8 7.5  
## 9 13.3  
## 10 23.4  
## 11 21.1  
## 12 11.0  
## 13 17.9  
## 14 21.3  
## 15 14.1  
## 16 22.4  
## 17 25.2  
## 18 7.3  
## 19 17.6  
## 20 15.1  
## 21 25.5

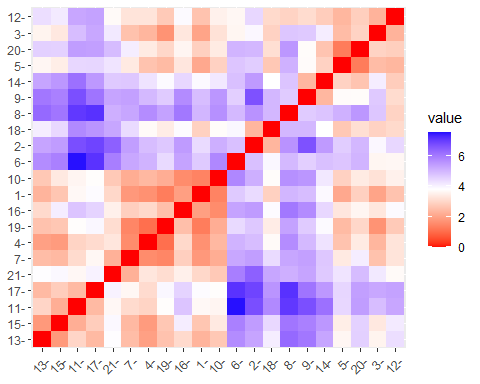
set.seed(222)  
library(factoextra)

## Warning: package 'factoextra' was built under R version 4.3.3

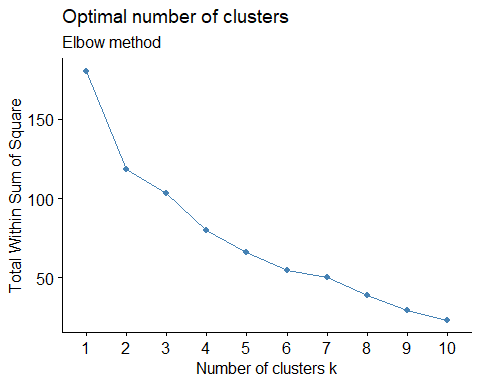
## Loading required package: ggplot2

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

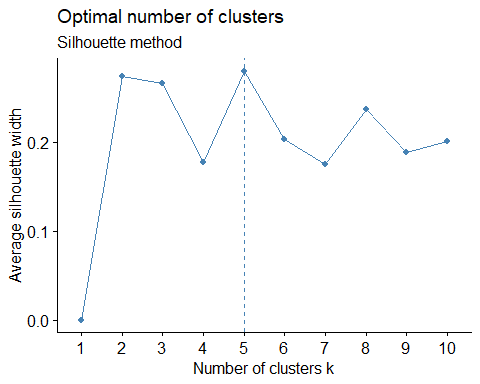
library(ISLR)  
library(cluster)  
scaled\_pharmadt <- scale(pharma\_data1)  
distance <- get\_dist(scaled\_pharmadt)  
fviz\_dist(distance)



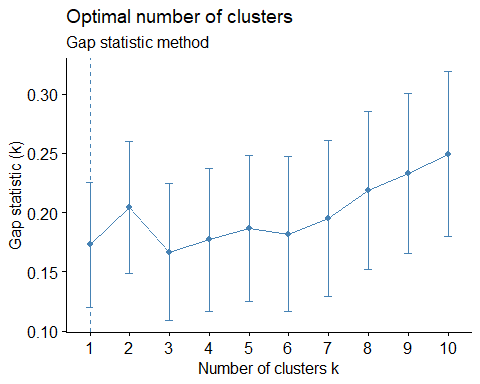
#Determine the optimal number of clusters using the elbow method (WSS)  
wss\_plt <- fviz\_nbclust(scaled\_pharmadt, kmeans, method = "wss")+labs(subtitle = "Elbow method")  
print(wss\_plt)



set.seed(333)  
# Determine the optimal number of clusters using the silhouette method  
sil\_plt <- fviz\_nbclust(scaled\_pharmadt, kmeans, method = "silhouette")+labs(subtitle = "Silhouette method")  
print(sil\_plt)



set.seed(111)  
fviz\_nbclust(scaled\_pharmadt, kmeans, nstart = 10, method = "gap\_stat", nboot = 50)+labs(subtitle = "Gap statistic method")



#K-means clusters  
k4 <- kmeans(scaled\_pharmadt,centers = 4, nstart = 25)  
k4$centers

## Market\_Cap Beta PE\_Ratio ROE ROA Asset\_Turnover  
## 1 -0.52462814 0.4451409 1.8498439 -1.0404550 -1.1865838 1.480297e-16  
## 2 1.69558112 -0.1780563 -0.1984582 1.2349879 1.3503431 1.153164e+00  
## 3 -0.03142211 -0.4360989 -0.3172485 0.1950459 0.4083915 1.729746e-01  
## 4 -0.82617719 0.4775991 -0.3696184 -0.5631589 -0.8514589 -9.994088e-01  
## Leverage Rev\_Growth Net\_Profit\_Margin  
## 1 -0.3443544 -0.5769454 -1.6095439  
## 2 -0.4680782 0.4671788 0.5912425  
## 3 -0.2744931 -0.7041516 0.5569544  
## 4 0.8502201 0.9158889 -0.3319956

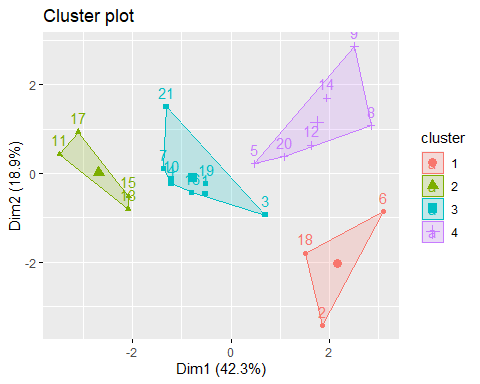
k4$size

## [1] 3 4 8 6

k4$cluster[21]

## [1] 3

#cluster plot  
fviz\_cluster(k4 , data = scaled\_pharmadt) #so there are 4 clusters formed.



library(flexclust)

## Warning: package 'flexclust' was built under R version 4.3.3

## Loading required package: grid

## Loading required package: lattice

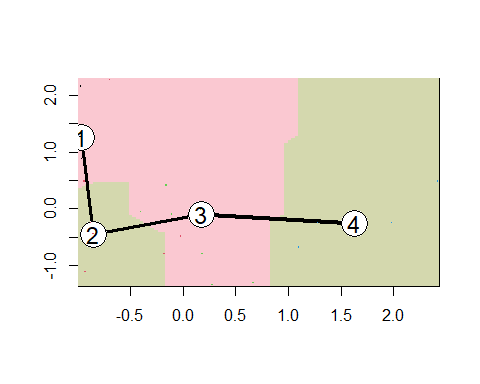
## Loading required package: modeltools

## Loading required package: stats4

set.seed(111)  
k4 <- kcca(scaled\_pharmadt, k = 4 , kccaFamily("kmedians"))  
k4

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

#apply the predict functon  
cluster\_index <- predict(k4)  
image(k4)  
points(scaled\_pharmadt, col = cluster\_index, pch = 15 , cex = 0.2)



# Appylyig knn for other numeric data from the pharma\_data.  
pharma\_data

## Symbol Name Market\_Cap Beta PE\_Ratio ROE ROA  
## 1 ABT Abbott Laboratories 68.44 0.32 24.7 26.4 11.8  
## 2 AGN Allergan, Inc. 7.58 0.41 82.5 12.9 5.5  
## 3 AHM Amersham plc 6.30 0.46 20.7 14.9 7.8  
## 4 AZN AstraZeneca PLC 67.63 0.52 21.5 27.4 15.4  
## 5 AVE Aventis 47.16 0.32 20.1 21.8 7.5  
## 6 BAY Bayer AG 16.90 1.11 27.9 3.9 1.4  
## 7 BMY Bristol-Myers Squibb Company 51.33 0.50 13.9 34.8 15.1  
## 8 CHTT Chattem, Inc 0.41 0.85 26.0 24.1 4.3  
## 9 ELN Elan Corporation, plc 0.78 1.08 3.6 15.1 5.1  
## 10 LLY Eli Lilly and Company 73.84 0.18 27.9 31.0 13.5  
## 11 GSK GlaxoSmithKline plc 122.11 0.35 18.0 62.9 20.3  
## 12 IVX IVAX Corporation 2.60 0.65 19.9 21.4 6.8  
## 13 JNJ Johnson & Johnson 173.93 0.46 28.4 28.6 16.3  
## 14 MRX Medicis Pharmaceutical Corporation 1.20 0.75 28.6 11.2 5.4  
## 15 MRK Merck & Co., Inc. 132.56 0.46 18.9 40.6 15.0  
## 16 NVS Novartis AG 96.65 0.19 21.6 17.9 11.2  
## 17 PFE Pfizer Inc 199.47 0.65 23.6 45.6 19.2  
## 18 PHA Pharmacia Corporation 56.24 0.40 56.5 13.5 5.7  
## 19 SGP Schering-Plough Corporation 34.10 0.51 18.9 22.6 13.3  
## 20 WPI Watson Pharmaceuticals, Inc. 3.26 0.24 18.4 10.2 6.8  
## 21 WYE Wyeth 48.19 0.63 13.1 54.9 13.4  
## Asset\_Turnover Leverage Rev\_Growth Net\_Profit\_Margin Median\_Recommendation  
## 1 0.7 0.42 7.54 16.1 Moderate Buy  
## 2 0.9 0.60 9.16 5.5 Moderate Buy  
## 3 0.9 0.27 7.05 11.2 Strong Buy  
## 4 0.9 0.00 15.00 18.0 Moderate Sell  
## 5 0.6 0.34 26.81 12.9 Moderate Buy  
## 6 0.6 0.00 -3.17 2.6 Hold  
## 7 0.9 0.57 2.70 20.6 Moderate Sell  
## 8 0.6 3.51 6.38 7.5 Moderate Buy  
## 9 0.3 1.07 34.21 13.3 Moderate Sell  
## 10 0.6 0.53 6.21 23.4 Hold  
## 11 1.0 0.34 21.87 21.1 Hold  
## 12 0.6 1.45 13.99 11.0 Hold  
## 13 0.9 0.10 9.37 17.9 Moderate Buy  
## 14 0.3 0.93 30.37 21.3 Moderate Buy  
## 15 1.1 0.28 17.35 14.1 Hold  
## 16 0.5 0.06 -2.69 22.4 Hold  
## 17 0.8 0.16 25.54 25.2 Moderate Buy  
## 18 0.6 0.35 15.00 7.3 Hold  
## 19 0.8 0.00 8.56 17.6 Hold  
## 20 0.5 0.20 29.18 15.1 Moderate Sell  
## 21 0.6 1.12 0.36 25.5 Hold  
## Location Exchange  
## 1 US NYSE  
## 2 CANADA NYSE  
## 3 UK NYSE  
## 4 UK NYSE  
## 5 FRANCE NYSE  
## 6 GERMANY NYSE  
## 7 US NYSE  
## 8 US NASDAQ  
## 9 IRELAND NYSE  
## 10 US NYSE  
## 11 UK NYSE  
## 12 US AMEX  
## 13 US NYSE  
## 14 US NYSE  
## 15 US NYSE  
## 16 SWITZERLAND NYSE  
## 17 US NYSE  
## 18 US NYSE  
## 19 US NYSE  
## 20 US NYSE  
## 21 US NYSE

pharma\_data2 <- pharma\_data[12:14]  
pharma\_data2

## Median\_Recommendation Location Exchange  
## 1 Moderate Buy US NYSE  
## 2 Moderate Buy CANADA NYSE  
## 3 Strong Buy UK NYSE  
## 4 Moderate Sell UK NYSE  
## 5 Moderate Buy FRANCE NYSE  
## 6 Hold GERMANY NYSE  
## 7 Moderate Sell US NYSE  
## 8 Moderate Buy US NASDAQ  
## 9 Moderate Sell IRELAND NYSE  
## 10 Hold US NYSE  
## 11 Hold UK NYSE  
## 12 Hold US AMEX  
## 13 Moderate Buy US NYSE  
## 14 Moderate Buy US NYSE  
## 15 Hold US NYSE  
## 16 Hold SWITZERLAND NYSE  
## 17 Moderate Buy US NYSE  
## 18 Hold US NYSE  
## 19 Hold US NYSE  
## 20 Moderate Sell US NYSE  
## 21 Hold US NYSE