Assignment 4

Jared Bartee

2023-10-25

# Load data source and create dataframe

Pharm <- read.csv("Pharmaceuticals.csv")  
head(Pharm)

## 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

#Collect the quantitative variables (1-9) to cluster the 21 firms

Pharm1 <- Pharm[,3:11]  
head(Pharm1)

## 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  
## Net\_Profit\_Margin  
## 1 16.1  
## 2 5.5  
## 3 11.2  
## 4 18.0  
## 5 12.9  
## 6 2.6

#Scale all quantitative variables in the dataframe

PharmS <- scale(Pharm1)  
head(PharmS)

## Market\_Cap Beta PE\_Ratio ROE ROA Asset\_Turnover  
## [1,] 0.1840960 -0.80125356 -0.04671323 0.04009035 0.2416121 0.0000000  
## [2,] -0.8544181 -0.45070513 3.49706911 -0.85483986 -0.9422871 0.9225312  
## [3,] -0.8762600 -0.25595600 -0.29195768 -0.72225761 -0.5100700 0.9225312  
## [4,] 0.1702742 -0.02225704 -0.24290879 0.10638147 0.9181259 0.9225312  
## [5,] -0.1790256 -0.80125356 -0.32874435 -0.26484883 -0.5664461 -0.4612656  
## [6,] -0.6953818 2.27578267 0.14948233 -1.45146000 -1.7127612 -0.4612656  
## Leverage Rev\_Growth Net\_Profit\_Margin  
## [1,] -0.2120979 -0.5277675 0.06168225  
## [2,] 0.0182843 -0.3811391 -1.55366706  
## [3,] -0.4040831 -0.5721181 -0.68503583  
## [4,] -0.7496565 0.1474473 0.35122600  
## [5,] -0.3144900 1.2163867 -0.42597037  
## [6,] -0.7496565 -1.4971443 -1.99560225

#K-Means Cluster Analysis - Fit the data with 5 clusters

fit <- kmeans(PharmS,5)

#Below command gives the mean value of all quantitative variables for each cluster

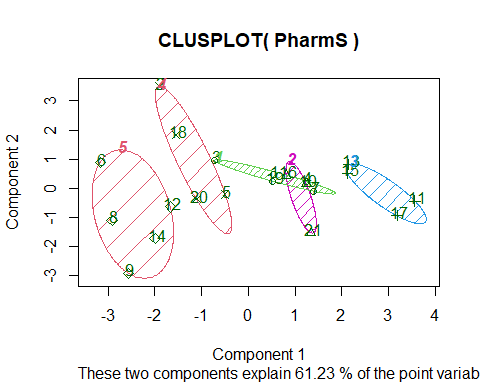
aggregate(PharmS, by = list(fit$cluster), FUN = mean)

## Group.1 Market\_Cap Beta PE\_Ratio ROE ROA  
## 1 1 -0.2063280 -0.2481660 -0.3385541 -0.03813318 0.4069821  
## 2 2 0.2600876 -0.7493205 -0.2817392 0.58367759 0.4107405  
## 3 3 1.6955811 -0.1780563 -0.1984582 1.23498791 1.3503431  
## 4 4 -0.4964157 -0.7136164 1.1595829 -0.74214495 -0.7778567  
## 5 5 -0.9090570 1.4110965 -0.2613021 -0.70634774 -1.1114156  
## Asset\_Turnover Leverage Rev\_Growth Net\_Profit\_Margin  
## 1 0.6457718 -0.42712134 -0.4707453 0.1531171  
## 2 -0.6150208 -0.02011273 -1.0931619 1.2300167  
## 3 1.1531640 -0.46807818 0.4671788 0.5912425  
## 4 -0.2306328 -0.27289324 0.6033984 -0.8374273  
## 5 -1.0147843 1.03196612 0.2701808 -0.6941793

PharmS1 <- data.frame(PharmS, fit$cluster)  
PharmS1

## Market\_Cap Beta PE\_Ratio ROE ROA Asset\_Turnover  
## 1 0.1840960 -0.80125356 -0.04671323 0.04009035 0.2416121 0.0000000  
## 2 -0.8544181 -0.45070513 3.49706911 -0.85483986 -0.9422871 0.9225312  
## 3 -0.8762600 -0.25595600 -0.29195768 -0.72225761 -0.5100700 0.9225312  
## 4 0.1702742 -0.02225704 -0.24290879 0.10638147 0.9181259 0.9225312  
## 5 -0.1790256 -0.80125356 -0.32874435 -0.26484883 -0.5664461 -0.4612656  
## 6 -0.6953818 2.27578267 0.14948233 -1.45146000 -1.7127612 -0.4612656  
## 7 -0.1078688 -0.10015669 -0.70887325 0.59693581 0.8617498 0.9225312  
## 8 -0.9767669 1.26308721 0.03299122 -0.11237924 -1.1677918 -0.4612656  
## 9 -0.9704532 2.15893320 -1.34037772 -0.70899938 -1.0174553 -1.8450624  
## 10 0.2762415 -1.34655112 0.14948233 0.34502953 0.5610770 -0.4612656  
## 11 1.0999201 -0.68440408 -0.45749769 2.45971647 1.8389364 1.3837968  
## 12 -0.9393967 0.48409069 -0.34100657 -0.29136529 -0.6979905 -0.4612656  
## 13 1.9841758 -0.25595600 0.18013789 0.18593083 1.0872544 0.9225312  
## 14 -0.9632863 0.87358895 0.19240011 -0.96753478 -0.9610792 -1.8450624  
## 15 1.2782387 -0.25595600 -0.40231769 0.98142435 0.8429577 1.8450624  
## 16 0.6654710 -1.30760129 -0.23677768 -0.52338423 0.1288598 -0.9225312  
## 17 2.4199899 0.48409069 -0.11415545 1.31287998 1.6322239 0.4612656  
## 18 -0.0240846 -0.48965495 1.90298017 -0.81506519 -0.9047030 -0.4612656  
## 19 -0.4018812 -0.06120687 -0.40231769 -0.21181593 0.5234929 0.4612656  
## 20 -0.9281345 -1.11285216 -0.43297324 -1.03382590 -0.6979905 -0.9225312  
## 21 -0.1614497 0.40619104 -0.75792214 1.92938746 0.5422849 -0.4612656  
## Leverage Rev\_Growth Net\_Profit\_Margin fit.cluster  
## 1 -0.21209793 -0.52776752 0.06168225 1  
## 2 0.01828430 -0.38113909 -1.55366706 4  
## 3 -0.40408312 -0.57211809 -0.68503583 1  
## 4 -0.74965647 0.14744734 0.35122600 1  
## 5 -0.31449003 1.21638667 -0.42597037 4  
## 6 -0.74965647 -1.49714434 -1.99560225 5  
## 7 -0.02011273 -0.96584257 0.74744375 1  
## 8 3.74279705 -0.63276071 -1.24888417 5  
## 9 0.61983791 1.88617085 -0.36501379 5  
## 10 -0.07130879 -0.64814764 1.17413980 2  
## 11 -0.31449003 0.76926048 0.82363947 3  
## 12 1.10620040 0.05603085 -0.71551412 5  
## 13 -0.62166634 -0.36213170 0.33598685 3  
## 14 0.44065173 1.53860717 0.85411776 5  
## 15 -0.39128411 0.36014907 -0.24310064 3  
## 16 -0.67286239 -1.45369888 1.02174835 2  
## 17 -0.54487226 1.10143723 1.44844440 3  
## 18 -0.30169102 0.14744734 -1.27936246 4  
## 19 -0.74965647 -0.43544591 0.29026942 1  
## 20 -0.49367621 1.43089863 -0.09070919 4  
## 21 0.68383297 -1.17763919 1.49416183 2

library(cluster)  
clusplot(PharmS, fit$cluster, color = TRUE, shade = TRUE, labels = 2, lines = 0)



#**Answers**

1. Cluster 1: Rows 1, 3, 4, 5, 19, 20 Cluster 2: Rows 2, 6, 18 Cluster 3: Rows 11, 13, 15, 17 Cluster 4: Rows 7, 10, 16, 21 Cluster 5: Rows 8, 9, 12, 14; By the output of the mean value of all quantitative variables for each cluster, we see that Cluster 1 has the lowest Leverage; Cluster 2 has the highest PE ratio, lowest ROE, lowest ROA, lowest Asset Turnover, & lowest Net Profit Margin; Cluster 3 has the highest Market Cap, highest ROE, highest ROA, & highest Asset Turnover; Cluster 4 has the highest Net Profit Margin, lowest Beta, lowest PE Ratio, & lowest Rev growth; Cluster 5 has the highest Beta, highest Leverage, highest Rev Growth, & lowest Market Cap.
2. There appears to be a pattern in the clusters regarding the Media recommendation variable. Cluster 3 does not have any sell media recommendations, this cluster is mostly buy recommendation with one strong buy recommendation. Cluster 2 has mostly hold recommendations. Cluster 4 also has mostly hold recommendations.
3. 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 Net Profit Margin, Low Beta, PE Ratio, Rev Growth Cluster Cluster 5: High Beta, Leverage, Rev Growth and Low Market Cap Cluster