

Assignment 4

Jacob Fabian

2022-10-29

```
library(readr)
library(cluster)
```

```
Pharmaceuticals_1_ <- read_csv("Pharmaceuticals (1).csv")
```

```
## Rows: 21 Columns: 14
## -- Column specification -----
## Delimiter: ","
## chr (5): Symbol, Name, Median_Recommendation, Location, Exchange
## dbl (9): Market_Cap, Beta, PE_Ratio, ROE, ROA, Asset_Turnover, Leverage, Rev...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
View(Pharmaceuticals_1_)
head(Pharmaceuticals_1_)
```

```
## # A tibble: 6 x 14
##   Symbol Name Marke~1 Beta PE_Ra~2 ROE ROA Asset~3 Lever~4 Rev_G~5 Net_P~6
##   <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 ABT Abbo~ 68.4 0.32 24.7 26.4 11.8 0.7 0.42 7.54 16.1
## 2 AGN Alle~ 7.58 0.41 82.5 12.9 5.5 0.9 0.6 9.16 5.5
## 3 AHM Amer~ 6.3 0.46 20.7 14.9 7.8 0.9 0.27 7.05 11.2
## 4 AZN Astr~ 67.6 0.52 21.5 27.4 15.4 0.9 0 15 18
## 5 AVE Aven~ 47.2 0.32 20.1 21.8 7.5 0.6 0.34 26.8 12.9
## 6 BAY Baye~ 16.9 1.11 27.9 3.9 1.4 0.6 0 -3.17 2.6
## # ... with 3 more variables: Median_Recommendation <chr>, Location <chr>,
## # Exchange <chr>, and abbreviated variable names 1: Market_Cap, 2: PE_Ratio,
## # 3: Asset_Turnover, 4: Leverage, 5: Rev_Growth, 6: Net_Profit_Margin
```

```
p1 <- Pharmaceuticals_1_[,3:11]
head(p1)
```

```
## # A tibble: 6 x 9
##   Market_Cap Beta PE_Ratio ROE ROA Asset_Turnover Leverage Rev_Gr~1 Net_P~2
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 68.4 0.32 24.7 26.4 11.8 0.7 0.42 7.54 16.1
## 2 7.58 0.41 82.5 12.9 5.5 0.9 0.6 9.16 5.5
## 3 6.3 0.46 20.7 14.9 7.8 0.9 0.27 7.05 11.2
```

```
## 4      67.6  0.52    21.5  27.4  15.4          0.9    0      15      18
## 5      47.2  0.32    20.1  21.8   7.5          0.6    0.34    26.8    12.9
## 6      16.9  1.11    27.9   3.9   1.4          0.6    0      -3.17    2.6
## # ... with abbreviated variable names 1: Rev_Growth, 2: Net_Profit_Margin
```

```
ps <- scale(p1)
head(ps)
```

```
##      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
```

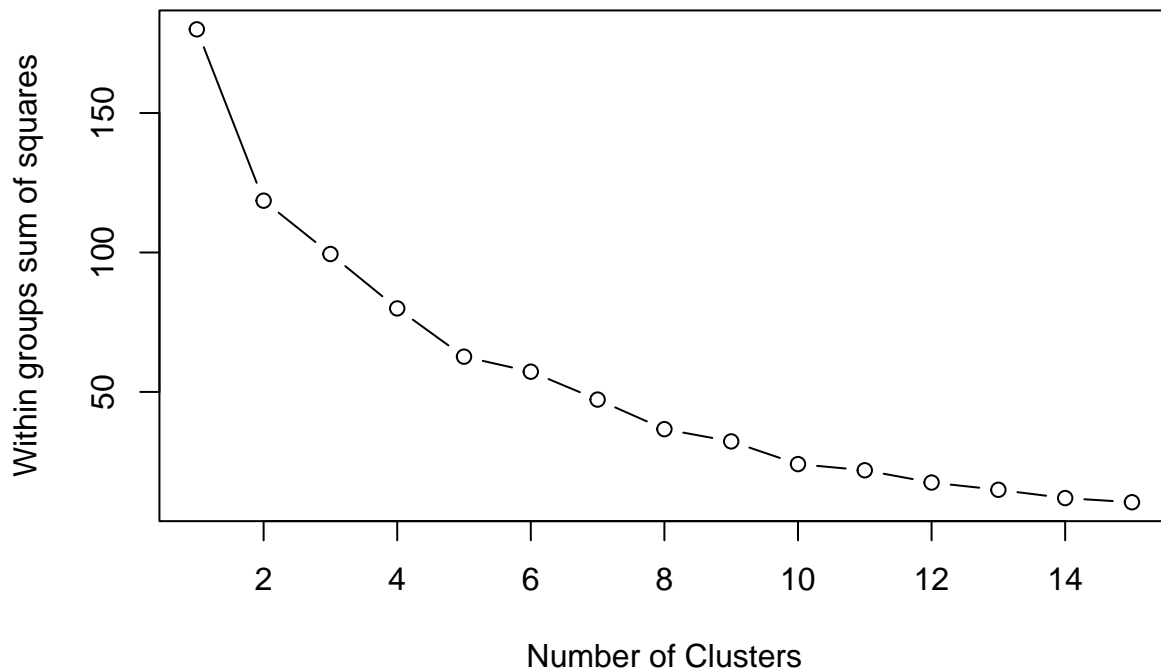
```
wss <- (nrow(Pharmaceuticals_1_)-1)*sum(apply(ps,2,var))
wss
```

```
## [1] 180
```

```
library(stats)
for (i in 2:15) wss[i] <- sum(kmeans(ps,centers=i)$withinss)
wss
```

```
## [1] 180.00000 118.56934 99.45504 79.95035 62.64525 57.26029 47.25303
## [8] 36.65309 32.26758 24.12543 21.90915 17.51296 14.90811 11.94864
## [15] 10.44947
```

```
plot(1:15, wss, type="b", xlab="Number of Clusters",ylab="Within groups sum of squares")
```



```
fit <- kmeans(ps, 5)
aggregate(ps, by=list (fit$cluster), FUN=mean)
```

##	Group.1	Market_Cap	Beta	PE_Ratio	ROE	ROA
## 1	1	-0.52462814	0.4451409	1.8498439	-1.0404550	-1.1865838
## 2	2	-0.66114002	-0.7233539	-0.3512251	-0.6736441	-0.5915022
## 3	3	-0.96247577	1.1949250	-0.3639982	-0.5200697	-0.9610792
## 4	4	1.69558112	-0.1780563	-0.1984582	1.2349879	1.3503431
## 5	5	0.08926902	-0.4618336	-0.3208615	0.3260892	0.5396003
##	Asset_Turnover	Leverage	Rev_Growth	Net_Profit_Margin		
## 1	1.480297e-16	-0.3443544	-0.5769454	-1.6095439		
## 2	-1.537552e-01	-0.4040831	0.6917224	-0.4005718		
## 3	-1.153164e+00	1.4773718	0.7120120	-0.3688236		
## 4	1.153164e+00	-0.4680782	0.4671788	0.5912425		
## 5	6.589509e-02	-0.2559803	-0.7230135	0.7343816		

```
ps1 <- data.frame(ps, fit$cluster)
ps1
```

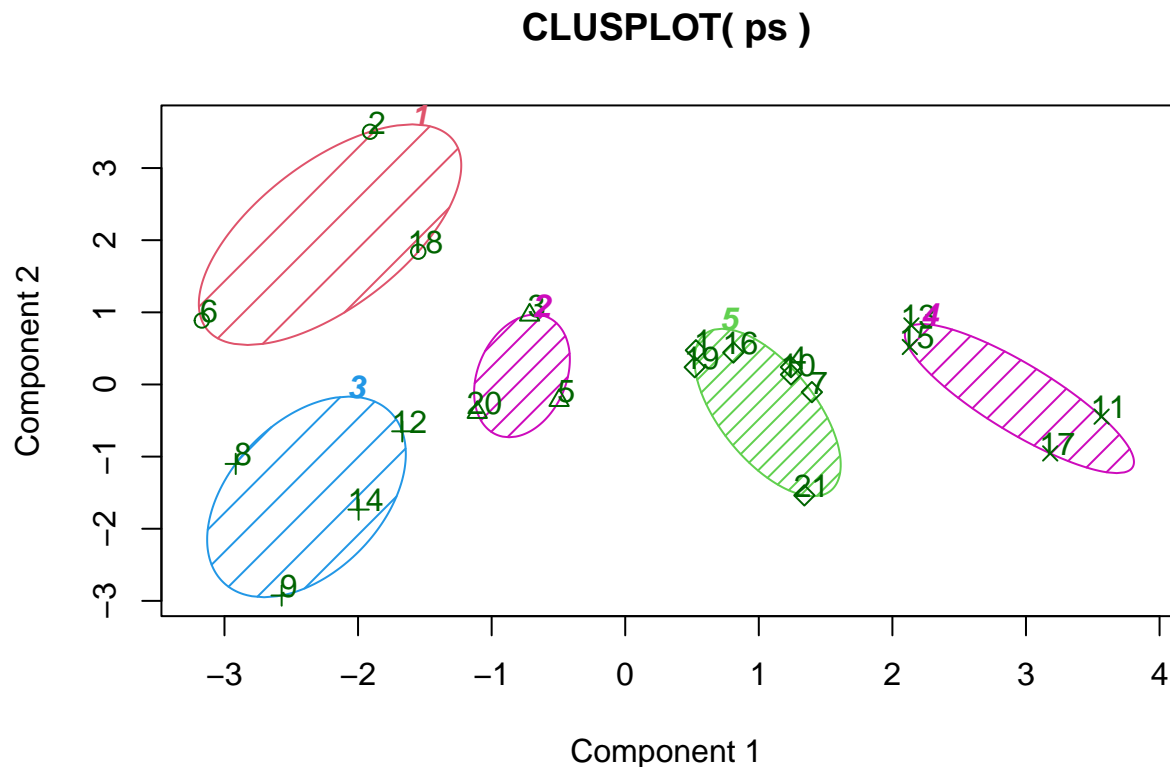
##	Market_Cap	Beta	PE_Ratio	ROE	ROA	Asset_Turnover
## 1	0.1840960	-0.80125356	-0.04671323	0.04009035	0.2416121	0.0000000
## 2	-0.8544181	-0.45070513	3.49706911	-0.85483986	-0.9422871	0.9225312
## 3	-0.8762600	-0.25595600	-0.29195768	-0.72225761	-0.5100700	0.9225312
## 4	0.1702742	-0.02225704	-0.24290879	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      5
## 2  0.01828430 -0.38113909     -1.55366706      1
## 3 -0.40408312 -0.57211809     -0.68503583      2
## 4 -0.74965647  0.14744734      0.35122600      5
## 5 -0.31449003  1.21638667     -0.42597037      2
## 6 -0.74965647 -1.49714434     -1.99560225      1
## 7 -0.02011273 -0.96584257      0.74744375      5
## 8  3.74279705 -0.63276071     -1.24888417      3
## 9  0.61983791  1.88617085     -0.36501379      3
## 10 -0.07130879 -0.64814764      1.17413980      5
## 11 -0.31449003  0.76926048      0.82363947      4
## 12 1.10620040  0.05603085     -0.71551412      3
## 13 -0.62166634 -0.36213170      0.33598685      4
## 14 0.44065173  1.53860717      0.85411776      3
## 15 -0.39128411  0.36014907     -0.24310064      4
## 16 -0.67286239 -1.45369888      1.02174835      5
## 17 -0.54487226  1.10143723      1.44844440      4
## 18 -0.30169102  0.14744734     -1.27936246      1
## 19 -0.74965647 -0.43544591      0.29026942      5
## 20 -0.49367621  1.43089863     -0.09070919      2
## 21 0.68383297 -1.17763919      1.49416183      5

```

```
clusplot(ps,fit$cluster, color=TRUE, shade=TRUE, labels=2, lines=0)
```



These two components explain 61.23 % of the point variability.

###Cluster 1 - Row 11,17

###Cluster 2 - Row 2,6,18

###Cluster 3 - Row 1,3,7,9,16,21

###Cluster 4 - Row 5,8,9,12,14,20

###Cluster 5 - Row 4,13,15

###Due to the aggregate(ps,by=list(fit\$cluster),FUN=mean)

###Cluster 1 has highest Net Profit Margin, lowest Beta, PE Ratio, and Rev growth.

###Cluster 2 has highest Beta, Leverage, Rev growth and the lowest Market Cap

###Cluster 3 has lowest Leverage.

###Cluster 4 has highest PE ratio, lowest ROE, ROA, Asset Turnover, and Net Profit Margin

###Cluster 5 has highest Market Cap, ROE, ROA, and Asset Turnover

###There is pattern in the clusters. Cluster 4 with highest PE ratio, lowest ROE, ROA, Asset Turnover, Net Profit Margin has a Hold recommendation.Cluster 5 with highest Market Cap, ROE, ROA, and Asset Turnover has a buy recommendation.

###We can name various clusters ###Cluster 1 - High Net Profit Margin, Low Beta, PE Ratio, Rev growth cluster

###Cluster 2 - High Beta, Leverage, Rev growth and Low Market Cap cluster

###Cluster 3 - Lowest Leverage cluster

###Cluster 4 - High PE ratio, Low ROE, ROA, Asset Turnover, Net Profit Margin Cluster

###Cluster 5 - High Market Cap, ROE, ROA, Asset Turnover cluster