# Assignment_04_FML

Yash_Patel

2022-11-03

Loading and installing the required packages

```
#install.packages("tidyverse")
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
#install.packages("factoextra")
library(factoextra)
```

```
## Warning: package 'factoextra' was built under R version 4.2.2
```

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

```
library(cluster)
```

**Loading the dataset**

```
Pharma<- read.csv("C:/Users/YASH/Downloads/Pharmaceuticals.csv")
summary(Pharma)
```

```
##     Symbol              Name             Market_Cap          Beta
##  Length:21          Length:21          Min.   :  0.41   Min.   :0.1800
##  Class :character   Class :character   1st Qu.:  6.30   1st Qu.:0.3500
##  Mode  :character   Mode  :character   Median : 48.19   Median :0.4600
##                                        Mean   : 57.65   Mean   :0.5257
##                                        3rd Qu.: 73.84   3rd Qu.:0.6500
##                                        Max.   :199.47   Max.   :1.1100
##     PE_Ratio          ROE             ROA        Asset_Turnover    Leverage
##  Min.   : 3.60   Min.   : 3.9   Min.   : 1.40   Min.   :0.3    Min.   :0.0000
```

```
##  1st Qu.:18.90   1st Qu.:14.9   1st Qu.: 5.70   1st Qu.:0.6    1st Qu.:0.1600
##  Median :21.50   Median :22.6   Median :11.20   Median :0.6    Median :0.3400
##  Mean   :25.46   Mean   :25.8   Mean   :10.51   Mean   :0.7    Mean   :0.5857
##  3rd Qu.:27.90   3rd Qu.:31.0   3rd Qu.:15.00   3rd Qu.:0.9    3rd Qu.:0.6000
##  Max.   :82.50   Max.   :62.9   Max.   :20.30   Max.   :1.1    Max.   :3.5100
##    Rev_Growth    Net_Profit_Margin Median_Recommendation   Location
##  Min.   :-3.17   Min.   : 2.6      Length:21             Length:21
##  1st Qu.: 6.38   1st Qu.:11.2      Class :character      Class :character
##  Median : 9.37   Median :16.1      Mode  :character      Mode  :character
##  Mean   :13.37   Mean   :15.7
##  3rd Qu.:21.87   3rd Qu.:21.1
##  Max.   :34.21   Max.   :25.5
##    Exchange
##  Length:21
##  Class :character
##  Mode  :character
##
##
##
```

## A.

**for initial analysis seperating columns with numerical values**

```
Pharma[,1] <- row.names(Pharma)
Pharma.01 <- Pharma[,3:11]
head(Pharma.01)
```

```
##   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
```

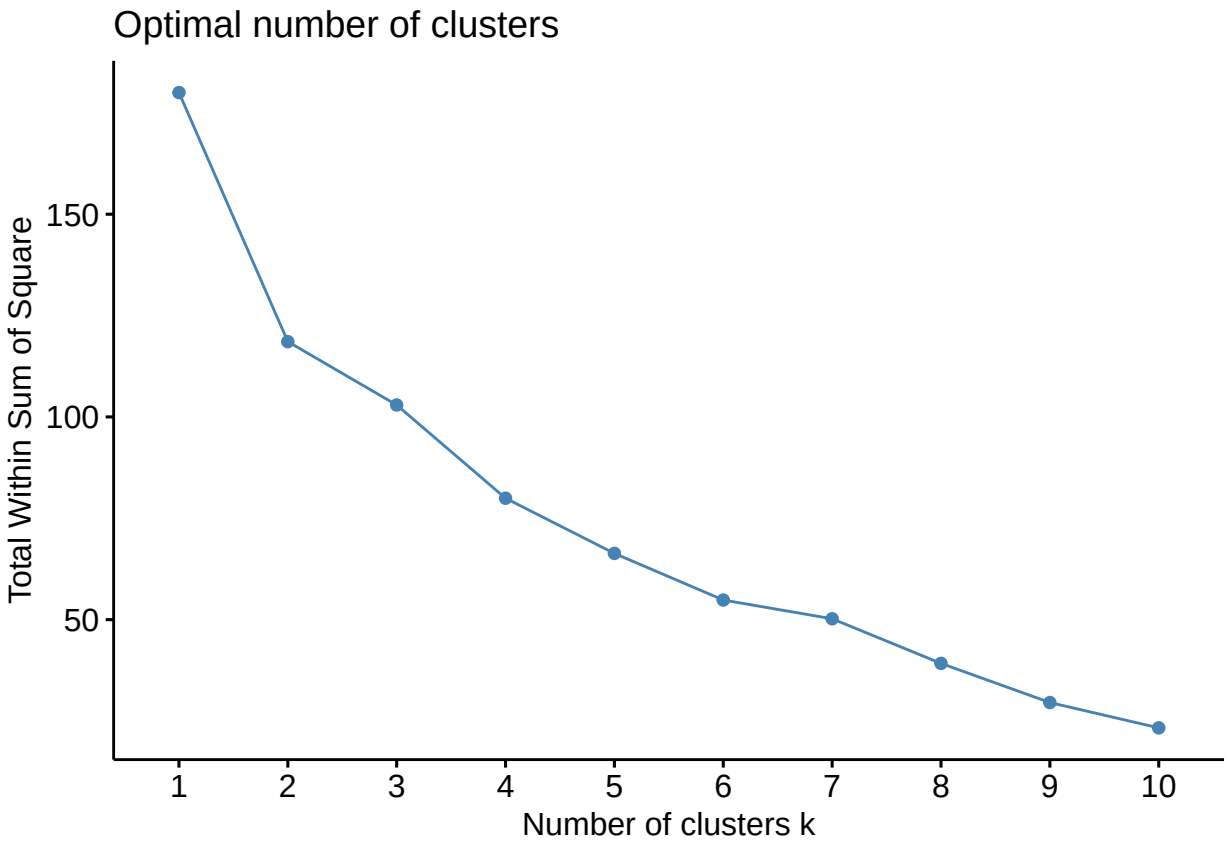
```
#Scaling the data
Scale.pharma <- scale(Pharma.01)
head(Scale.pharma)
```

```
##      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
```

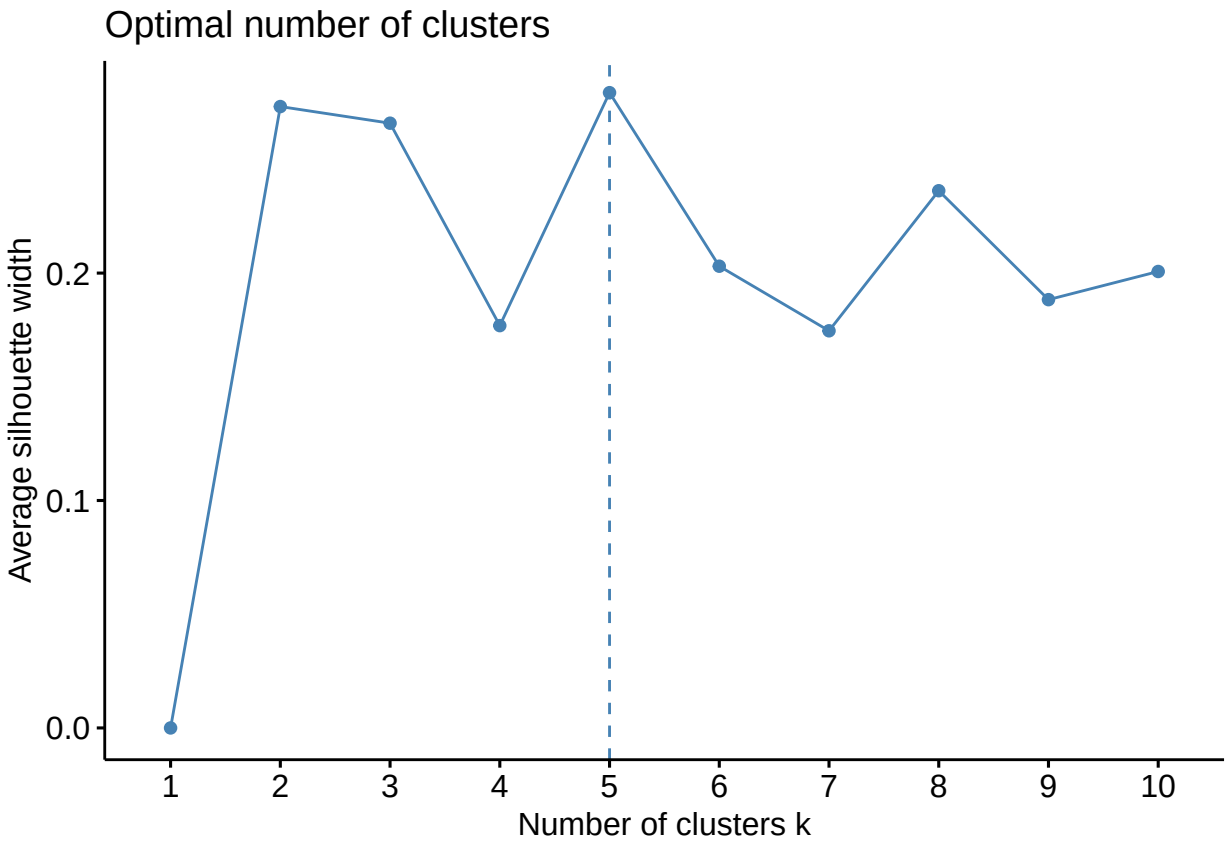
```
## [4,]  0.1702742 -0.02225704 -0.24290879  0.10638147  0.9181259      0.9225312
## [5,] -0.1790256 -0.80125356 -0.32874435 -0.26484883 -0.5664461     -0.4612656
## [6,] -0.6953818  2.27578267  0.14948233 -1.45146000 -1.7127612     -0.4612656
##        Leverage Rev_Growth Net_Profit_Margin
## [1,] -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
```

**Using the Elbow method and the silhouette method, we can determine the number of clusters.**

```
fviz_nbclust(Scale.pharma, kmeans, method = "wss")
```




```
fviz_nbclust(Scale.pharma, kmeans, method = "silhouette")
```

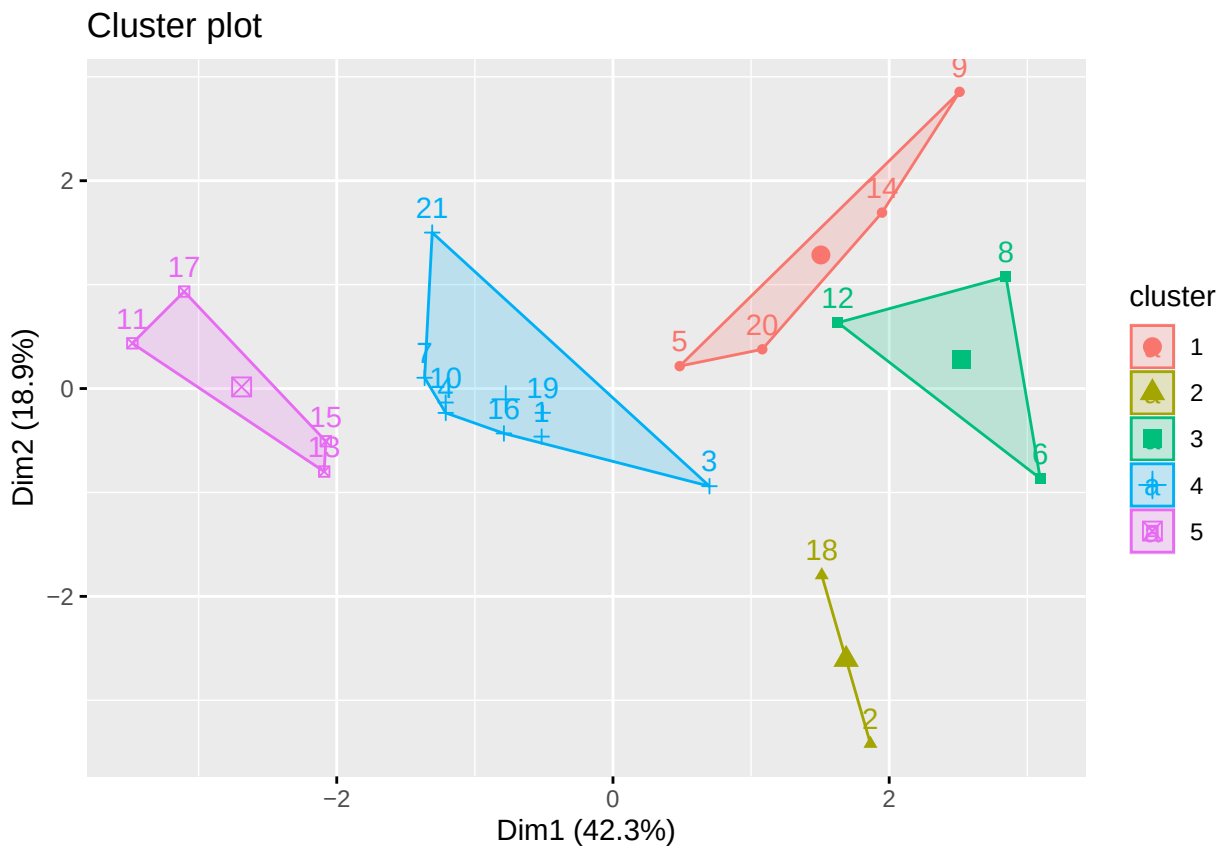



In order to determine the right number of clusters, I used the Elbow and Silhouette methods. so here we will use the silhouette method because it is thought to be better compared to the elbow method.

```
set.seed(4560)
Pharma.5.clust<- kmeans(Scale.pharma, centers = 5, nstart = 25)
Pharma.5.clust$centers
```

```
##    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
```

```
fviz_cluster(Pharma.5.clust, data = Scale.pharma)
```




**According to the above clusters, the sizes are 8, 2, 4, 4, and 3.**

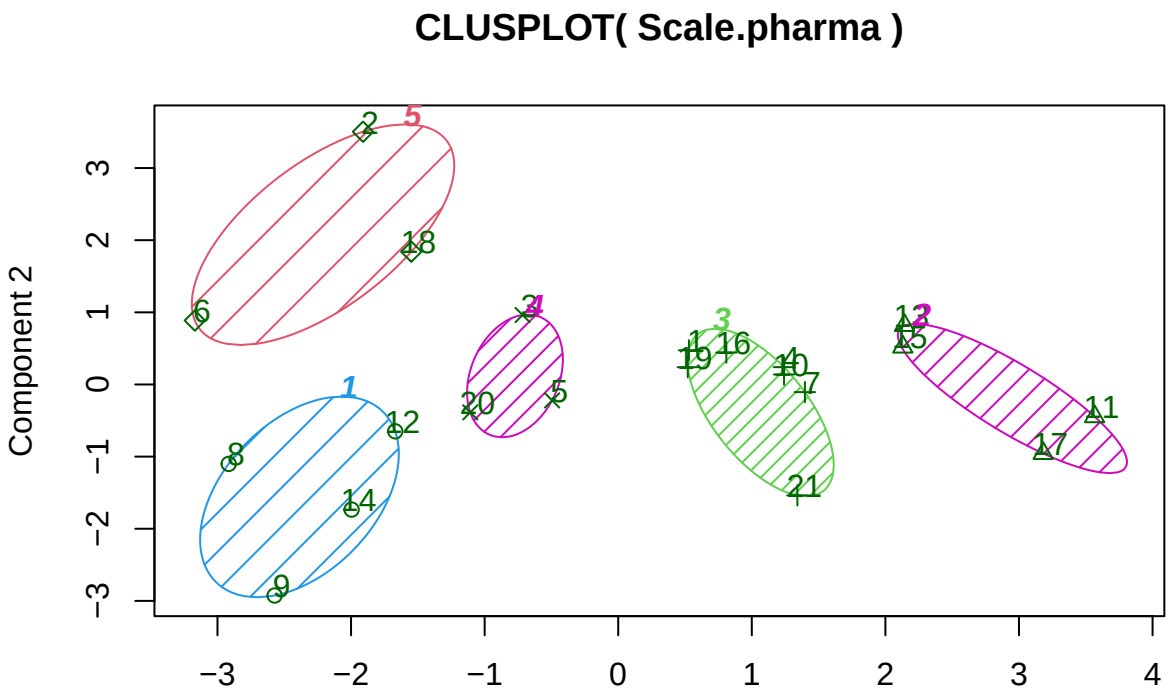
```
pharma.fit <- kmeans(Scale.pharma, 5)


Pharma.02 <- data.frame(Scale.pharma, pharma.fit$cluster)
Pharma.02
```

```
##    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 pharma.fit.cluster
## 1  -0.21209793 -0.52776752        0.06168225                 3
## 2   0.01828430 -0.38113909       -1.55366706                 5
## 3  -0.40408312 -0.57211809       -0.68503583                 4
## 4  -0.74965647  0.14744734        0.35122600                 3
## 5  -0.31449003  1.21638667       -0.42597037                 4
## 6  -0.74965647 -1.49714434       -1.99560225                 5
## 7  -0.02011273 -0.96584257        0.74744375                 3
## 8   3.74279705 -0.63276071       -1.24888417                 1
## 9   0.61983791  1.88617085       -0.36501379                 1
## 10 -0.07130879 -0.64814764        1.17413980                 3
## 11 -0.31449003  0.76926048        0.82363947                 2
## 12  1.10620040  0.05603085       -0.71551412                 1
## 13 -0.62166634 -0.36213170        0.33598685                 2
## 14  0.44065173  1.53860717        0.85411776                 1
## 15 -0.39128411  0.36014907       -0.24310064                 2
## 16 -0.67286239 -1.45369888        1.02174835                 3
## 17 -0.54487226  1.10143723        1.44844440                 2
## 18 -0.30169102  0.14744734       -1.27936246                 5
## 19 -0.74965647 -0.43544591        0.29026942                 3
## 20 -0.49367621  1.43089863       -0.09070919                 4
## 21  0.68383297 -1.17763919        1.49416183                 3
```

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




### B. Interpret the clusters with respect to the numerical variables used in forming the clusters.

observing the mean values of all quantitative variables for each cluster

```
aggregate(Scale.pharma, by = list(pharma.fit$cluster), FUN=mean)
```

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

Cluster01 - JNJ, MRK, GSK, and PFE are the companies with the largest market capitalizations, and they manage their companies by adequately funding them (leverage below 0.47)

Cluster02 - AHM, AVE, and WPI have the lowest asset turnover and beta, respectively, indicating that their stocks are underperforming the current market performance index.

Cluster03 - IVX, MRX, ELN, and CHTT have the lowest market capitalization, the company does not use debt to fund its operations, and they all have the highest revenue growth. These company stocks also provide good returns because their beta value is greater than 1.

Cluster04 - With the highest price to earnings ratio, AGN, PHA, and BAY are the least profitable. They also have a Return on Equity below 1, demonstrating that investing in these stocks will not be as profitable as investing in other stocks.

Cluster05- The highest asset turnover, slowest revenue growth, and highest net profit margin are found in ABT, NVS, AZN, LLY, BMY, WYE, and SGP. Due to their rapid growth, these businesses are profitable.

## C. Is there a pattern in the clusters with respect to the numerical variables (10 to 12)? (those not used in forming the clusters)

To find the patterns relating to media recommendations, location, and exchange, we manually filtered each cluster.

For cluster01: The nature of the stocks is moderate, i.e., neither they are weak stocks nor have they recently produced strong returns.

For cluster02: In terms of location, the stocks are diverse. Their fundamentals are technically sound, and media recommendations are extremely favorable.

For cluster03: Their leverage ratio is high, but they are moderately recommended due to their financial stability.

For cluster04: According to media recommendations, these are the stocks that should be held because they will eventually become profitable investments.

For Cluster05: Due to their high net profit margin, the stocks in the cluster are recommended to be held for a longer period of time.

## D. Provide an appropriate name for each cluster using any or all of the variables in the dataset.

Cluster01: Growth Cluster - Because these are stable stocks Cluster02: Multibagger cluster - Despite their low beta, market recommendations are very positive. Cluster03: Fundamental Cluster - Stocks that have good financial and other fundamental stability. Cluster04: Hold cluster - These stocks have decent stats. Cluster05: Long term Cluster - Stocks that have a high net profit margin are considered to be profitable, so holding them in your portfolio is advisable.