

# Assignment\_4

Nithin Kumar Chinthakindi

2023-11-19

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.3      v readr      2.1.4
## v forcats    1.0.0      v stringr    1.5.0
## v ggplot2    3.4.3      v tibble     3.2.1
## v lubridate  1.9.2      v tidyr      1.3.0
## v purrr      1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(factoextra)
```

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

```
library(dplyr)
library(ggplot2)
library(cluster)
```

```
Pharmaceuticals <- read.csv("/Users/nithinkumarch/Downloads/Pharmaceuticals.csv")
```

```
# Task 1: Cluster the 21 firms using only the numerical variables (1 to 9).
```

```
# Eliminate missing data and rescale variables for comparability before clustering the data.
```

```
Pharma_data <- na.omit(Pharmaceuticals)
Pharma_data
```

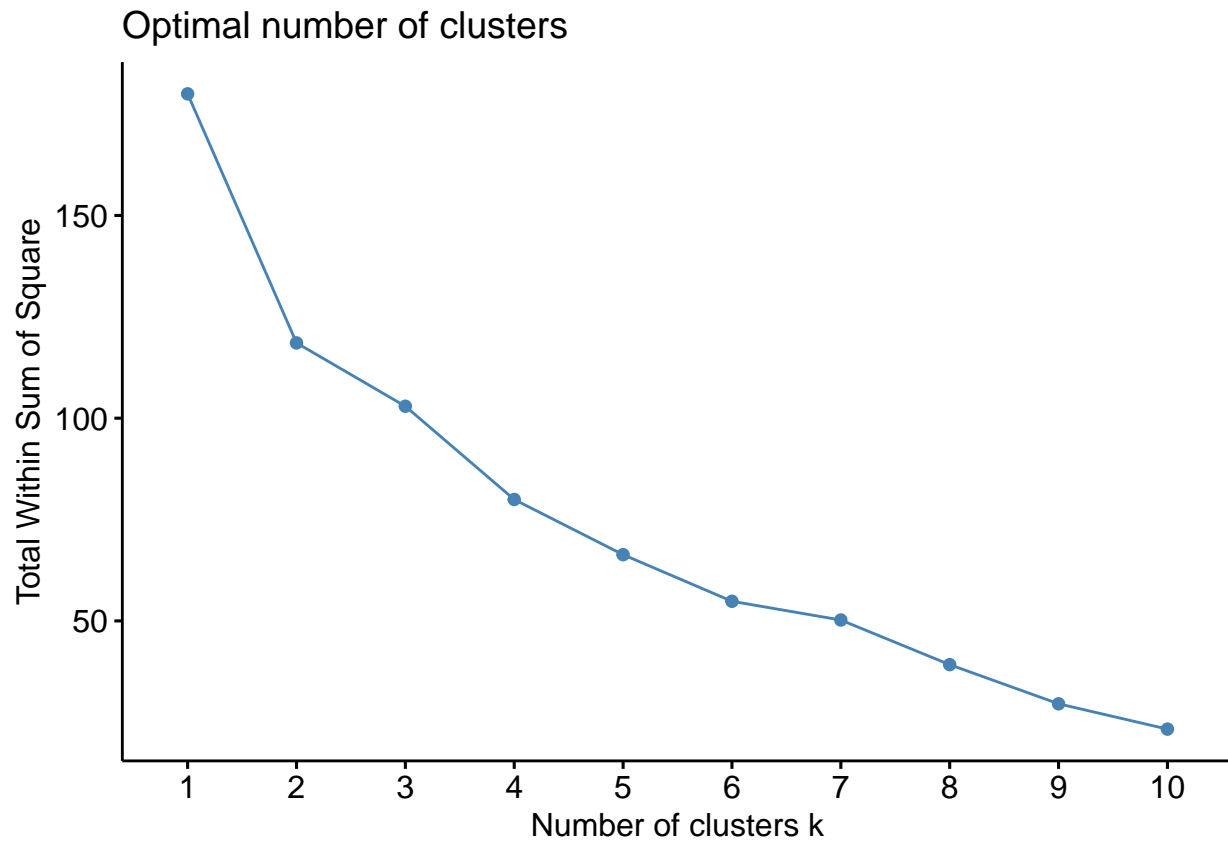
| ##    | Symbol | Name                               | Market_Cap | Beta | PE_Ratio | ROE  | ROA  |
|-------|--------|------------------------------------|------------|------|----------|------|------|
| ## 1  | ABT    | Abbott Laboratories                | 68.44      | 0.32 | 24.7     | 26.4 | 11.8 |
| ## 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  | BMJ    | 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                         |            |                   |                       |      |      |

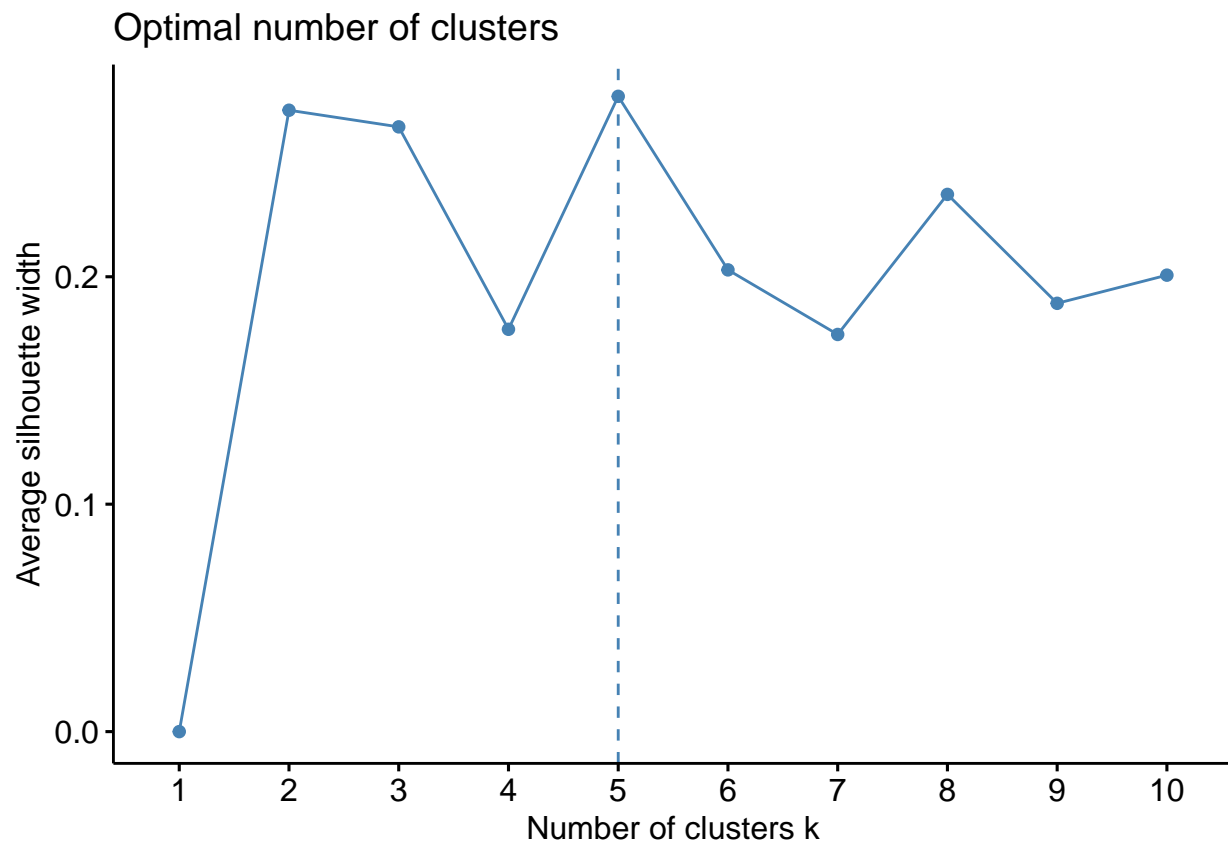
```
# Clustering the 21 enterprises using the quantitative variables (1-9)
row.names(Pharma_data) <- Pharma_data[, 1]
Pharma_data1 <- Pharma_data[, 3:11] # Considering only numerical values (columns 3-11)
```

```
# Applying the scale method to normalize the data frame
Pharma_data2 <- scale(Pharma_data1)
```

```
# Use the Elbow Method to calculate the number of clusters.
fviz_nbclust(Pharma_data2, kmeans, method = "wss")
```

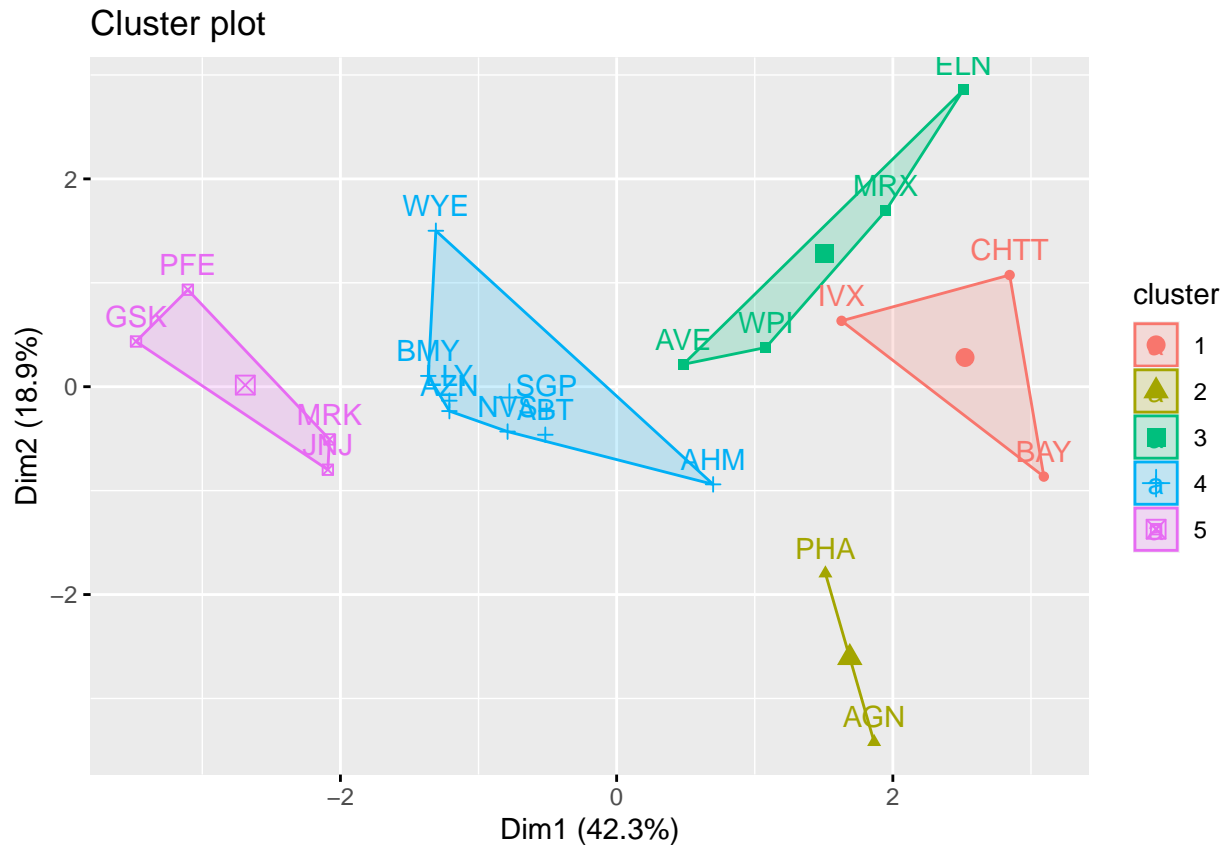


```
# Use the Silhouette method to calculate the number of clusters.
fviz_nbclust(Pharma_data2, kmeans, method = "silhouette")
```



```
# Using k=5 and K-means
set.seed(64060)
k_5 <- kmeans(Pharma_data2, centers = 5, nstart = 25)

# Putting the clusters to paper
fviz_cluster(k_5, data = Pharma_data2)
```



*# Task 2: Analyze the clusters in relation to the numerical variables that were utilized to create them*  
*# By taking note of each cluster's mean values for all quantitative variables #... (Cluster interpretation)*

*# Task 3: In relation to the numerical variables (10 to 12), are there any patterns in the clusters?*  
*# (those not utilized in cluster formation)*

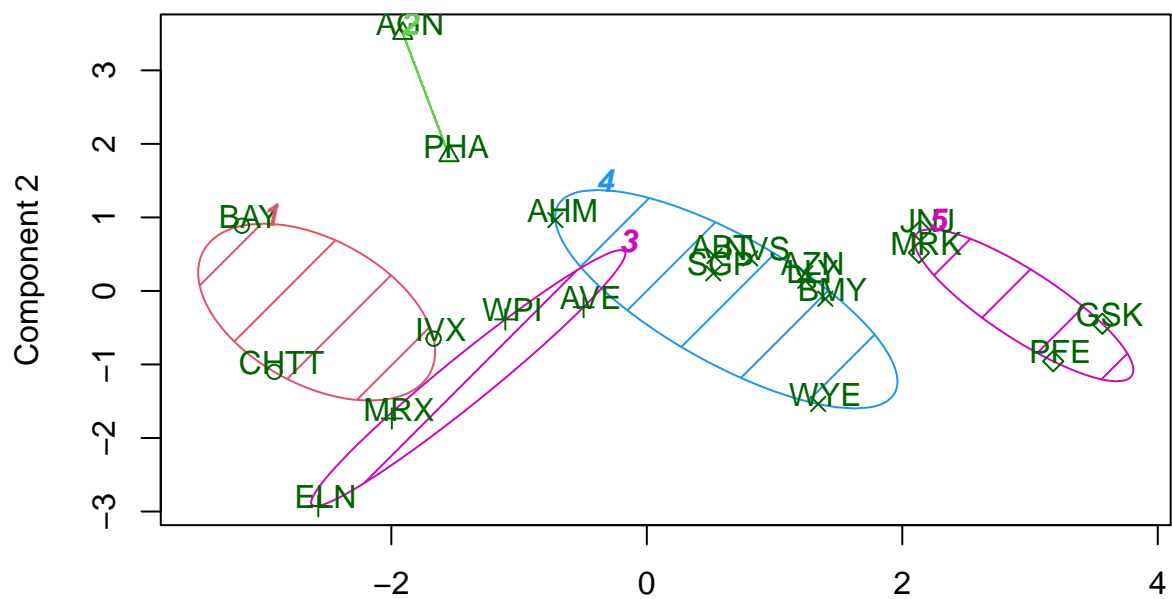
*#Note: It appears that the code provided lacks any analysis pertaining to variables 10 through 12.*

*# Task 4: Using any or all of the variables in the dataset, give each cluster a suitable name.*  
*#... (Comments on Cluster Naming)*

*# Viewing the cluster plot*

```
clusplot(Pharma_data2, k_5$cluster, color = TRUE, shade = TRUE, labels = 2, lines = 0)
```

# CLUSPLOT( Pharma\_data2 )



These two components explain 61.23 % of the point variability.