


Nama : Ibnu Fajar Setiawan NIM : 065002000006	 UNIVERSITAS TRISAKTI Praktikum Data Analitik	Modul 8 Nama Dosen: Syandra Sari, S.Kom, M.Kom
Hari/Tanggal : Kamis, 17 November 2022		Nama Aslab : 1. Azzahra Nuranisa (065001900044) 2. Ida Jubaidah (065001900037)

Praktikum 8 – Data Analitik

ANALISIS KOMPONEN UTAMA (PRINCIPAL COMPONENT ANALYSIS (PCA)) DAN ANALISIS KLASSTER

DESKRIPSI MODUL : Melihat hubungan antara variabel Katagorik dengan Variabel Numerik.

No	Elemen Kompetensi	Indikator Kinerja	Jml Jam	hlm
1	Mampu melakukan analisis data terhadap data harga komoditas dengan principal component analysis dan cluster analysis	Dapat melakukan analisis data terhadap data harga komoditas dengan principal component analysis dan cluster analysis	2	

TEORI SINGKAT

Analisis komponen utama (principal component analysis) dapat digunakan untuk membuat visualisasi atau peta posisi dari berbagai obyek dalam dimensi yang lebih rendah. Hal ini dapat terjadi karena beberapa variabel asal dari matriks data saling berkorelasi atau terjadinya multikolinieritas. Sedangkan analisis klaster digunakan (cluster analysis) dapat digunakan untuk mengelompokkan obyek berdasarkan nilai kemiripannya (similarity) berdasarkan konsep jarak antara dua titik.

LAB SETUP

Untuk dapat menjalankan praktikum ini maka yang harus disiapkan adalah :

1. Aplikasi RStudio
2. Xampp

ELEMEN KOMPETENSI I

Deskripsi : Dapat melakukan analisis data terhadap data harga komoditas dengan principal component analysis

Kompetensi Dasar : Mampu melakukan analisis data terhadap data harga komoditas dengan principal component analysis

PRAKTIKUM

Dalam praktikum ini akan dipelajari dan dipraktekkan bagaimana melakukan analisis data terhadap data harga komoditas dengan principal component analysis menggunakan data harga komoditas di berbagai kota di Amerika Serikat

```
dataku=read.csv("D:/File Kuliah Semester 5/Data Analitik/Prak-8/komoditas.csv")  
  
View(dataku)  
  
str(dataku)
```

```
dataku.pca=dataku[,2:6]
komoditas.pca <- prcomp(dataku.pca, center = TRUE, scale. = TRUE)
print(komoditas.pca)
plot(komoditas.pca, type = "l")

summary(komoditas.pca)
predict(komoditas.pca, newdata=tail(dataku))
hasil=predict(komoditas.pca, newdata=tail(dataku))
View(hasil)

install_github("vqv/ggbiplot")
komoditas.kota <- dataku[,1]
library(devtools)
library(ggbiplot)
g <- ggbiplot(komoditas.pca,ellipse = TRUE, circle = TRUE)
g <- g + scale_color_discrete(name = "")
g <- g + theme(legend.direction = 'horizontal', legend.position = 'top')
print(g)

g <- ggbiplot(komoditas.pca, obs.scale = 1,var.scale = 1,groups = komoditas.kota, circle =
TRUE)
g <- g + scale_color_discrete(name = "")
g <- g + theme(legend.direction = 'horizontal', legend.position = 'top')
print(g)
```

Output :

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

FilekerjaPrak8.R* x hasil x dataku x

Filter

	City	Bread	Burger	Milk	Oranges	Tomatoes
1	Arlanta	24.5	94.5	73.9	80.1	41.6
2	Baltimo	26.6	91.0	67.5	74.6	53.3
3	Boston	29.7	100.8	61.4	104.0	59.6
4	Buffalo	22.8	86.6	65.3	118.4	51.2
5	Chicago	26.7	86.7	62.7	105.9	51.2
6	Cincinnati	25.3	102.5	63.3	99.3	45.6
7	Cleveland	22.8	88.8	52.4	110.9	46.8
8	Dallas	23.3	85.5	62.5	117.9	41.8
9	Detroit	24.1	93.7	51.5	109.7	52.4
10	Honolulu	29.3	105.9	80.2	133.2	61.7
11	Houston	22.3	83.6	67.8	108.6	42.4
12	Kansas	26.1	88.9	65.4	100.9	43.2
13	Los Angeles	26.9	89.3	56.2	82.7	38.4
14	Milwaukee	20.3	89.6	53.8	111.8	53.9
15	Minneapolis	24.6	92.2	51.9	106.0	50.7
16	New York	30.8	110.7	66.0	107.3	62.6
17	Philadelphia	24.5	92.3	66.7	98.0	61.7
18	Pittsburgh	26.2	95.4	60.2	117.1	49.3
19	St. Louis	26.5	92.4	60.8	115.1	46.2

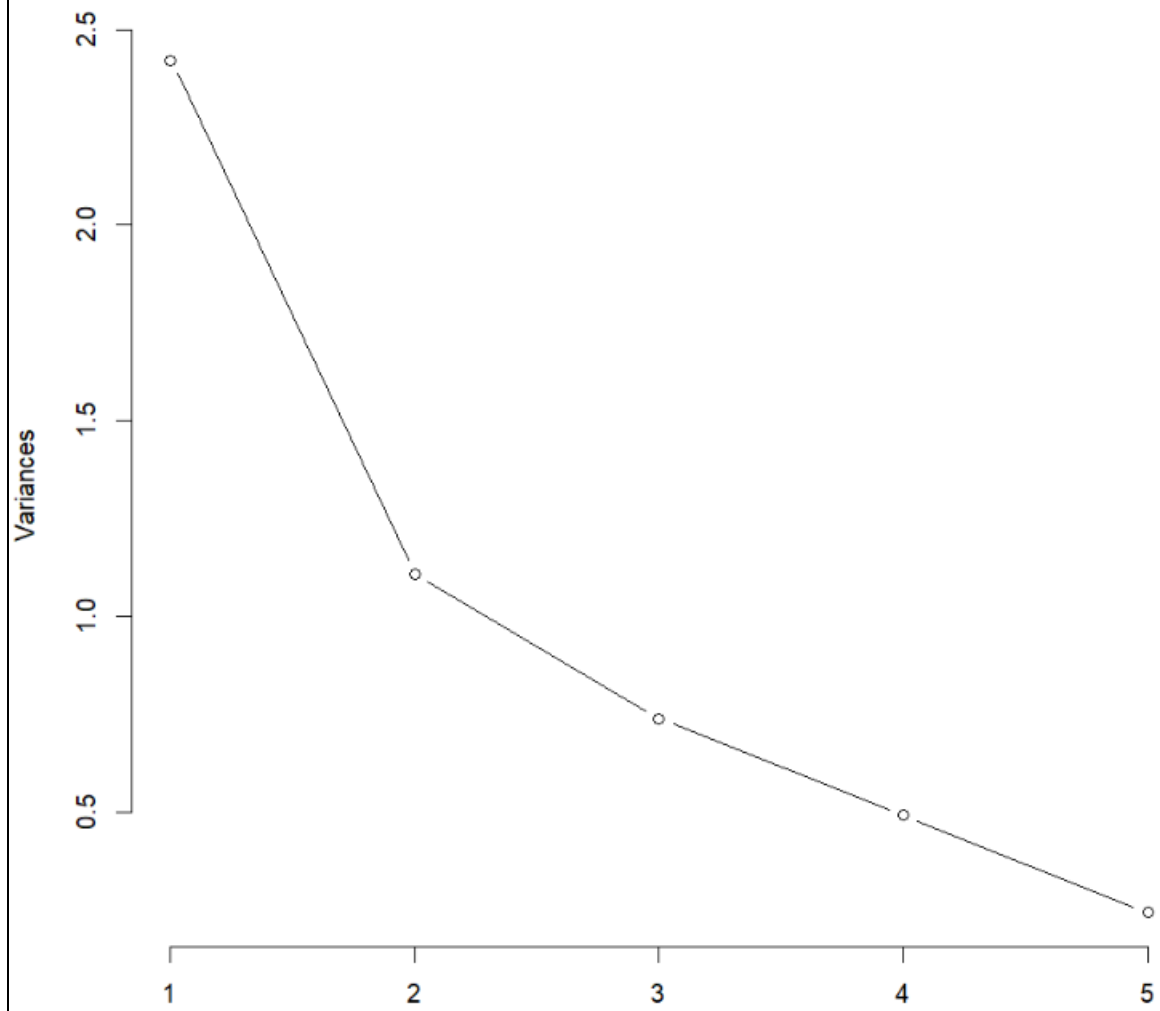
Showing 1 to 19 of 23 entries, 6 total columns

Console Terminal x Background Jobs x

```
R 4.2.1 ~ /
> dataku=read.csv("D:/File Kuliah Semester 5/Data Analitik/Prak-8/komoditas.csv")
> View(dataku)
> |
```

```
> dataku=read.csv("D:/File Kuliah Semester 5/Data Analitik/Prak-8/komoditas.csv")
> View(dataku)
> str(dataku)
'data.frame': 23 obs. of 6 variables:
 $ City : chr "Arlanta" "Baltimo" "Boston" "Buffalo" ...
 $ Bread : num 24.5 26.6 29.7 22.8 26.7 25.3 22.8 23.3 24.1 29.3 ...
 $ Burger : num 94.5 91 100.8 86.6 86.7 ...
 $ Milk : num 73.9 67.5 61.4 65.3 62.7 63.3 52.4 62.5 51.5 80.2 ...
 $ Oranges : num 80.1 74.6 104 118.4 105.9 ...
 $ Tomatoes: num 41.6 53.3 59.6 51.2 51.2 45.6 46.8 41.8 52.4 61.7 ...
> dataku.pca=dataku[,2:6]
> komoditas.pca <- prcomp(dataku.pca, center = TRUE, scale. = TRUE)
> print(komoditas.pca)
Standard deviations (1, .., p=5):
[1] 1.5562491 1.0519527 0.8593360 0.7009800 0.4915821

Rotation (n x k) = (5 x 5):
      PC1      PC2      PC3      PC4      PC5
Bread 0.4960921 -0.31116324 0.38400354 0.50926221 -0.500263314
Burger 0.5756736 -0.04237556 0.26199524 -0.02488239 0.773009430
Milk 0.3400964 -0.42779987 -0.83615624 0.04580486 0.008144644
Oranges 0.2238587 0.79723159 -0.29066455 0.47931306 -0.009064873
Tomatoes 0.5066698 0.28773392 0.01601972 -0.71287711 -0.389928880
> plot(komoditas.pca, type = "l")
> |
```

komoditas.pca

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

FilekjerPrak8.R* x hasil x dataku x

Filter

	PC1	PC2	PC3	PC4	PC5
18	0.6037072	0.8071284	0.2265320	0.5828393	0.14339325
19	0.2257479	0.5205205	0.1305183	0.8809165	-0.06240393
20	-1.8911706	-0.7300545	0.5653872	0.9435969	-0.18954093
21	-0.8622959	-0.1935930	0.4784665	0.8343109	-0.31822364
22	-2.0904434	-0.3677658	-0.6486172	-0.5605815	-0.68562580
23	-0.3945226	-1.3977846	-0.1146839	-0.6838917	0.56682383

Showing 1 to 6 of 6 entries, 5 total columns

Console Terminal Background Jobs

```
R 4.2.1 ~/  
> summary(komoditas.pca)  
Importance of components:  
Standard deviation      PC1      PC2      PC3      PC4      PC5  
Proportion of Variance  0.4844  0.2213  0.1477  0.09827 0.04833  
Cumulative Proportion   0.4844  0.7057  0.8534  0.95167 1.00000  
> predict(komoditas.pca, newdata=tail(dataku))  
      PC1      PC2      PC3      PC4      PC5  
18  0.6037072  0.8071284  0.2265320  0.5828393  0.14339325  
19  0.2257479  0.5205205  0.1305183  0.8809165 -0.06240393  
20 -1.8911706 -0.7300545  0.5653872  0.9435969 -0.18954093  
21 -0.8622959 -0.1935930  0.4784665  0.8343109 -0.31822364  
22 -2.0904434 -0.3677658 -0.6486172 -0.5605815 -0.68562580  
23 -0.3945226 -1.3977846 -0.1146839 -0.6838917  0.56682383  
> hasil=predict(komoditas.pca, newdata=tail(dataku))  
> View(hasil)  
> |
```

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

FilekjerPrak8.R* x hasil x dataku x

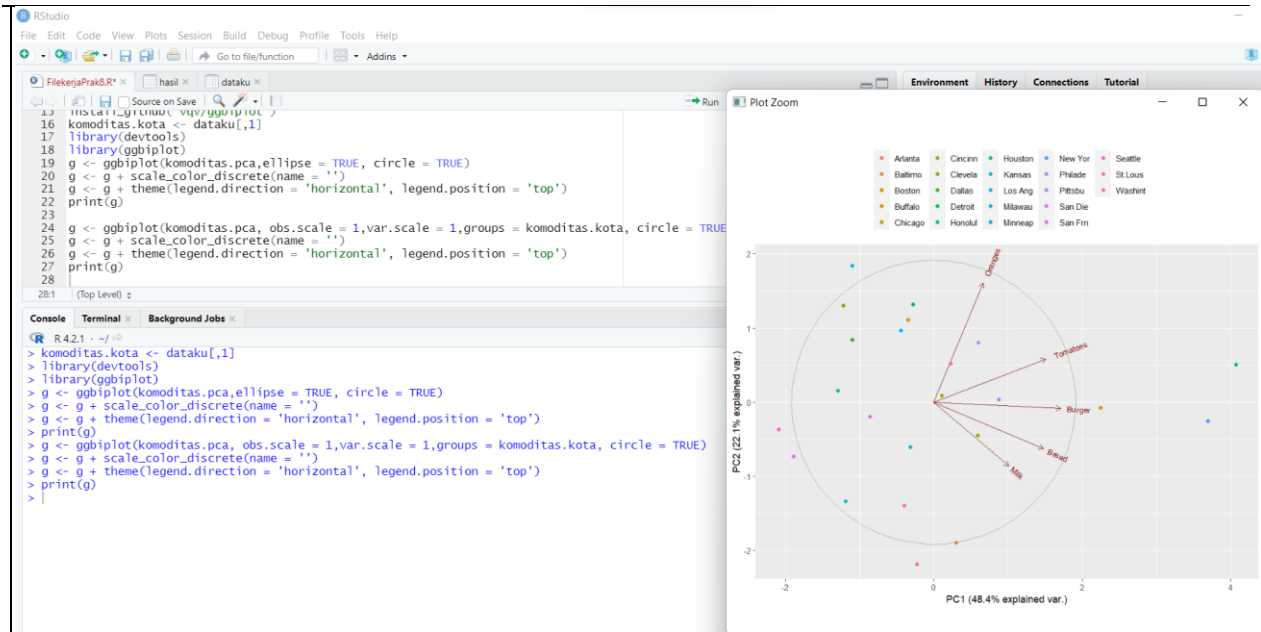
Source on Save

```
14  
15 install_github("vqv/ggbiplot")  
16 komoditas.kota <- dataku[,1]  
17 library(devtools)  
18 library(ggbiplot)  
19 g <- ggbiplot(komoditas.pca, ellipse = TRUE, circle = TRUE)  
20 g <- g + scale_color_discrete(name = '')  
21 g <- g + theme(legend.direction = 'horizontal', legend.position = 'top')  
22 print(g)  
23  
24 g <- ggbiplot(komoditas.pca, obs.scale = 1, var.scale = 1, groups = komoditas.kota)  
25 g <- g + scale_color_discrete(name = '')  
26 g <- g + theme(legend.direction = 'horizontal', legend.position = 'top')  
27 print(g)  
28  
24:1 (Top Level) >
```

Console Terminal Background Jobs

```
R 4.2.1 ~/  
> komoditas.kota <- dataku[,1]  
> library(devtools)  
> library(ggbiplot)  
> g <- ggbiplot(komoditas.pca, ellipse = TRUE, circle = TRUE)  
> g <- g + scale_color_discrete(name = '')  
> g <- g + theme(legend.direction = 'horizontal', legend.position = 'top')  
> print(g)  
> |
```

Plot Zoom



ELEMEN KOMPETENSI II

Deskripsi : Dapat melakukan analisis data terhadap data harga komoditas dengan principal cluster analysis

Kompetensi Dasar : Mampu melakukan analisis data terhadap data harga komoditas dengan principal cluster analysis

Dalam praktikum ini akan dipelajari dan dipraktekkan bagaimana melakukan analisis data terhadap data harga komoditas dengan cluster analysis menggunakan data harga komoditas di berbagai kota di Amerika Serikat

```
##ek 2
install.packages("NbClust")
library(factoextra)
library(NbClust)
str(dataku)
ibnu.pca=dataku[,2:6]
View(ibnu.pca)
nb <- NbClust(ibnu.pca, distance = "euclidean",min.nc = 2, max.nc = 10, method = "complete",
index = "all")
km.res=kmeans(ibnu.pca,3,nstart = 25)
fviz_cluster(km.res, data = ibnu.pca, geom = "point", stand = FALSE, frame.type = "norm")
fviz_cluster(km.res, data = ibnu.pca)
```

Output:

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

+ Go to file/function Addins

FileKerjaPrak8.R* Ibnu.pca hasil dataku

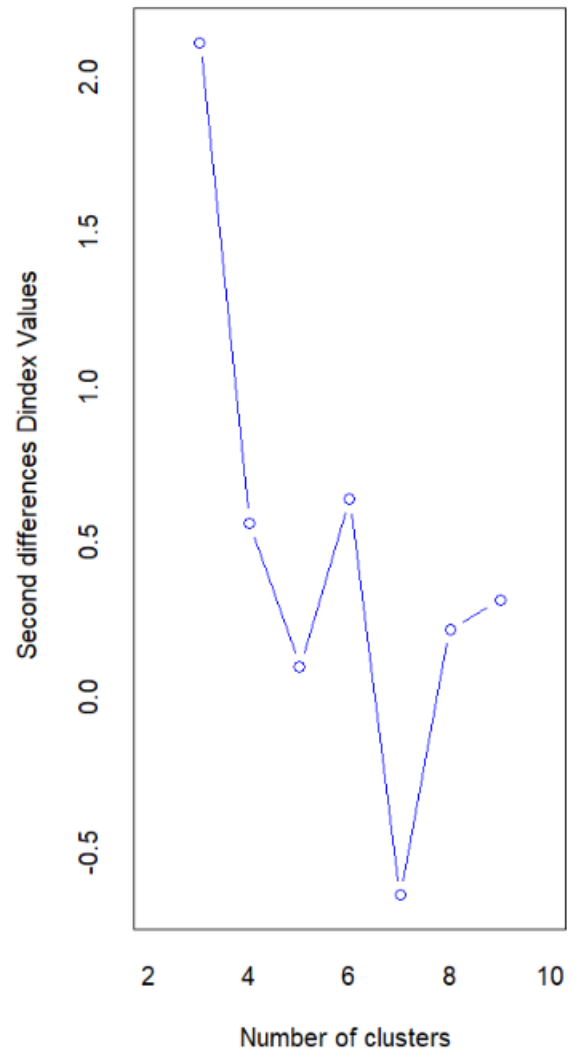
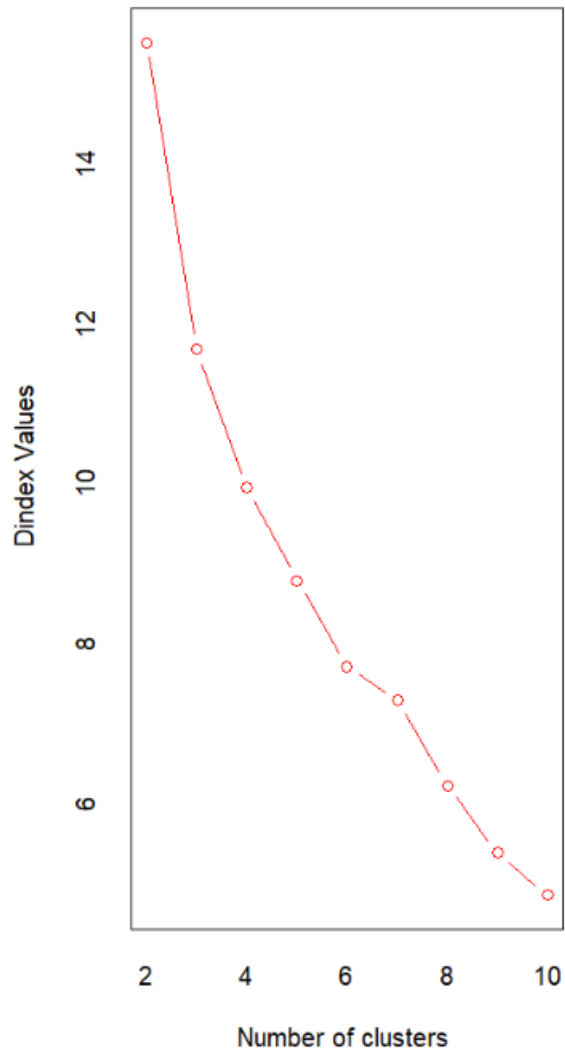
	Bread	Burger	Milk	Oranges	Tomatoes
1	24.5	94.5	73.9	80.1	41.6
2	26.6	91.0	67.5	74.6	53.3
3	29.7	100.8	61.4	104.0	59.6
4	22.8	86.6	65.3	118.4	51.2
5	26.7	86.7	62.7	105.9	51.2
6	25.3	102.5	63.3	99.3	45.6
7	22.8	88.8	52.4	110.9	46.8
8	23.3	85.5	62.5	117.9	41.8
9	24.1	93.7	51.5	109.7	52.4
10	29.3	105.9	80.2	133.2	61.7
11	22.3	83.6	67.8	108.6	42.4
12	26.1	88.9	65.4	100.9	43.2
13	26.9	89.3	56.2	82.7	38.4
14	20.3	89.6	53.8	111.8	53.9
15	24.6	92.2	51.9	106.0	50.7
16	30.8	110.7	66.0	107.3	62.6
17	24.5	94.5	73.9	80.1	41.6

Showing 1 to 17 of 23 entries, 5 total columns

Console Terminal Background Jobs

R 4.2.1 ~/

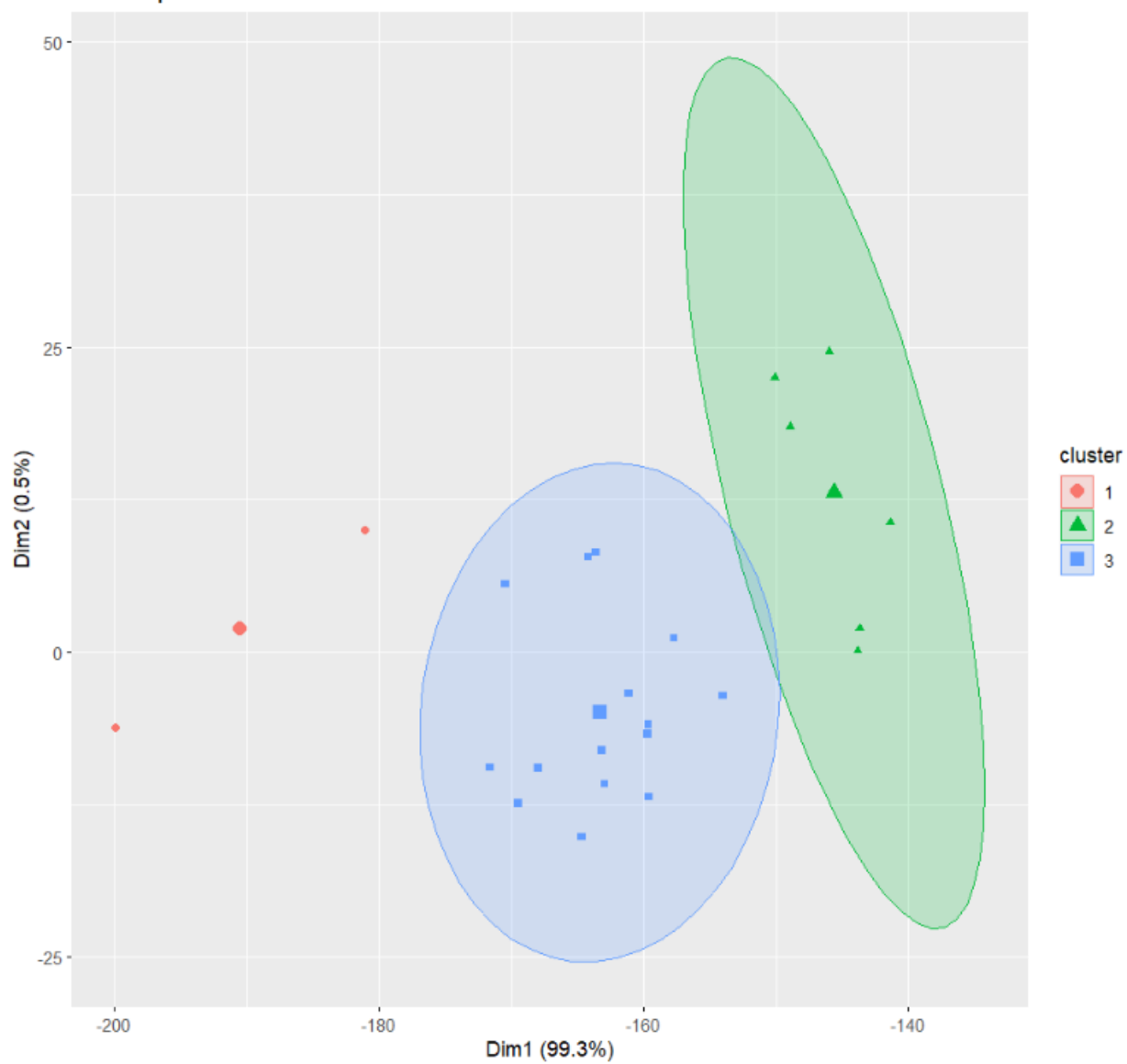
```
> str(dataku)
'data.frame': 23 obs. of 6 variables:
 $ City : chr "Arlanta" "Baltimo" "Boston" "Buffalo" ...
 $ Bread : num 24.5 26.6 29.7 22.8 26.7 25.3 22.8 23.3 24.1 29.3 ...
 $ Burger : num 94.5 91 100.8 86.6 86.7 ...
 $ Milk : num 73.9 67.5 61.4 65.3 62.7 63.3 52.4 62.5 51.5 80.2 ...
 $ Oranges : num 80.1 74.6 104 118.4 105.9 ...
 $ Tomatoes: num 41.6 53.3 59.6 51.2 51.2 45.6 46.8 41.8 52.4 61.7 ...
> Ibnu.pca=dataku[,2:6]
> view(Ibnu.pca)
> |
```



Plot Zoom



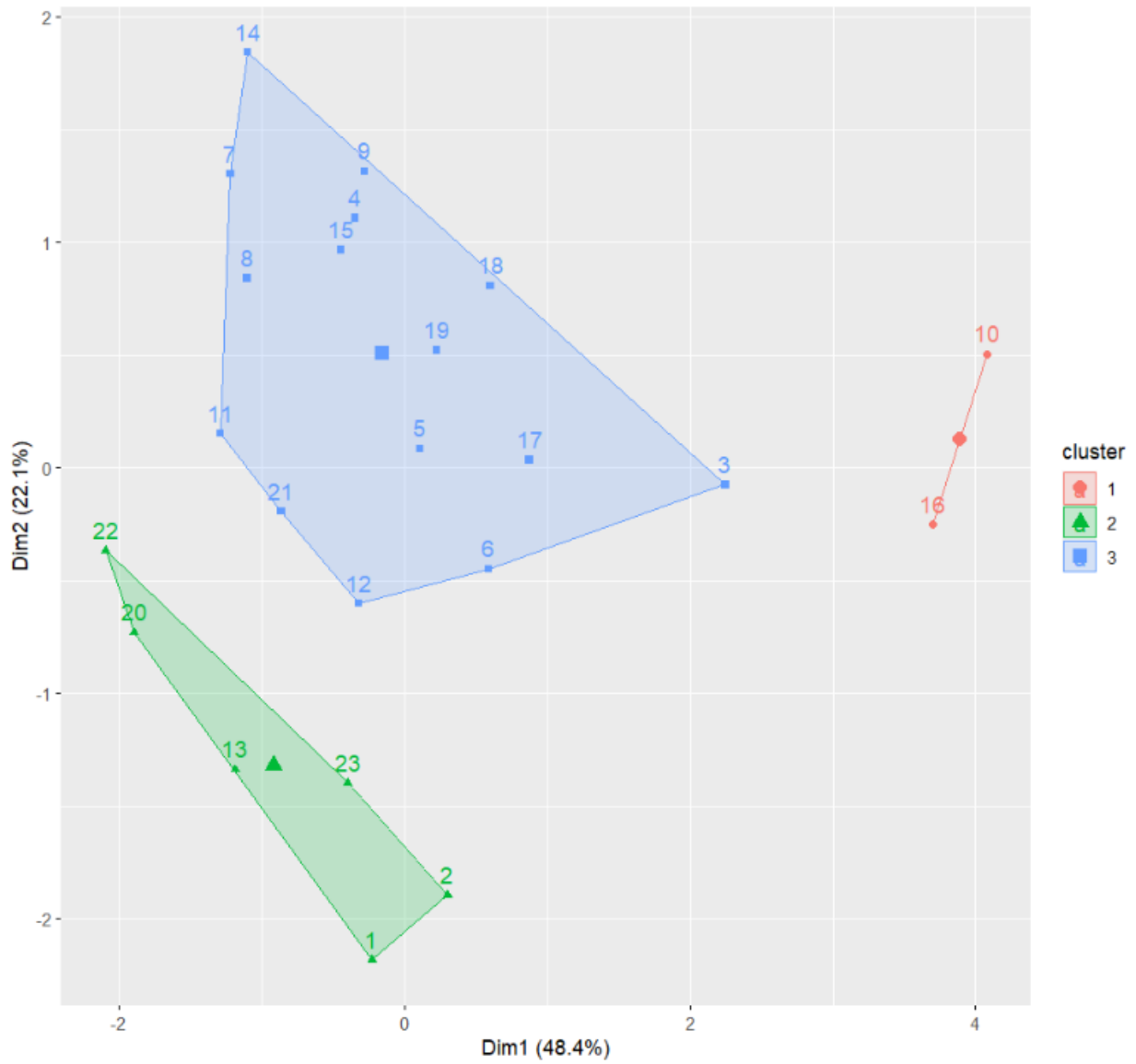
Cluster plot



Plot Zoom



Cluster plot



1. Cek List

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak
1	Elemen Kompetensi I Dapat melakukan analisis data terhadap data harga komoditas dengan principal component analysis	✓	
2	Elemen Kompetensi II Dapat melakukan analisis data terhadap data harga komoditas dengan principal cluster analysis	✓	

2. Form Umpan Balik

Elemen Kompetensi	Waktu Pengerjaan	Kriteria
Elemen Kompetensi I Dapat melakukan analisis data terhadap data harga komoditas dengan principal component analysis	60 menit	1
Elemen Kompetensi II Dapat melakukan analisis data terhadap data harga komoditas dengan principal cluster analysis	60 menit	1

Kriteria

- 1.Sangat Menarik
- 2.Cukup Menarik
- 3.Kurang Menarik
- 4.Sangat Kurang Menarik

Sumber : Sharma, S. (1996) Applied Multivariate Technique.

<https://tgmstat.wordpress.com/2013/11/28/computing-and-visualizing-pca-in-r/>

Sumber : <http://www.sthda.com/english/wiki/print.php?id=239#nbclust-r-function>