

## Final exam q6

Mehrab Atighi

1/12/2022

```
# multivariate 2 final exam codes: qestion :6
```

```
library(MASS)
```

```
data = read.table("F:/lessons/Multi countios Variate2/data/gandom.txt")
```

```
data = t(data)
```

```
colnames(data) = c("Cambridge" , "Cockle Park" , "Harpers Adams" , "Headley  
Hall" , "Morley" , "Myerscough" , "Rosemaund" , "Seale-Hayne" , "Seale-Hayne" ,  
"Sutton Bonington" , "Terrington" , "Wye")
```

ماتریس فاصله‌ها را محاسبه میکنیم تا به کمک آن بتوانیم خوشه بندی خود را انجام بدهیم نکته اینجا هستش که این ماتریس با توجه به فاصله‌ی اقلیدسی محاسبه شده است. و همانطور که میبینیم فاصله‌ی V2 , V4 از همه کمتر هست پس اول این دو یک خوشه میشوند و به همین ترتیب ....

```
(Dist1 = dist(data , method = "euclidean",  
diag = TRUE , upper = TRUE))
```

```
##          V1          V2          V3          V4  
## V1  0.00000 34.52965 39.34903 40.20390  
## V2 34.52965  0.00000 43.91066 32.64609  
## V3 39.34903 43.91066  0.00000 49.59618  
## V4 40.20390 32.64609 49.59618  0.00000
```

```
model1 = hclust(Dist1 , method = "ward.D")  
model1
```

```
##
```

```
## Call:
```

```
## hclust(d = Dist1, method = "ward.D")
```

```
##
```

```
## Cluster method    : ward.D
```

```
## Distance          : euclidean
```

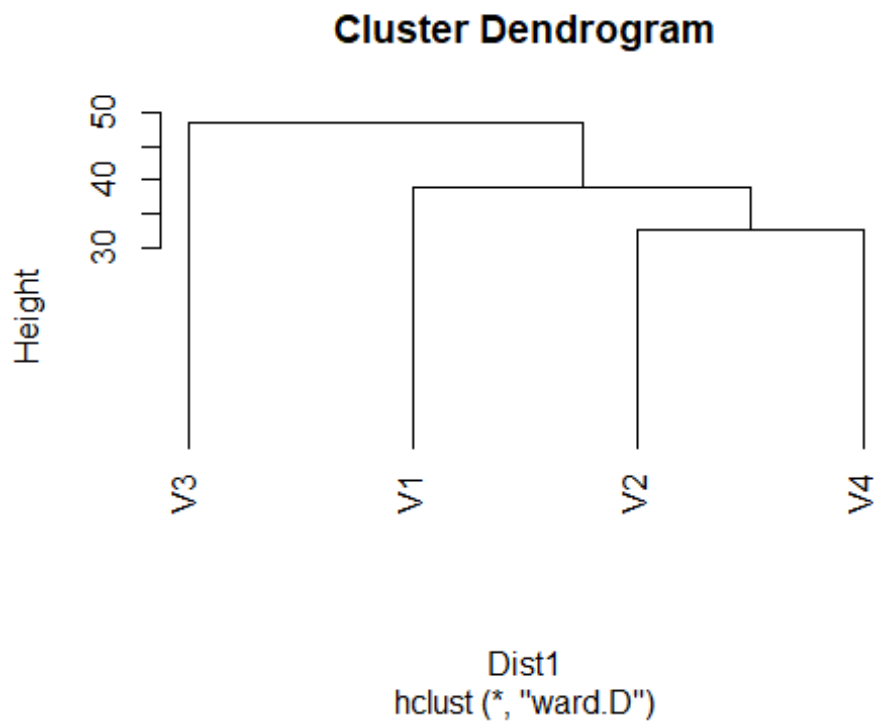
```
## Number of objects: 4
```

```
plot( model1 , hang = -1 )
```

```
library(factoextra)
```

```
## Loading required package: ggplot2
```

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

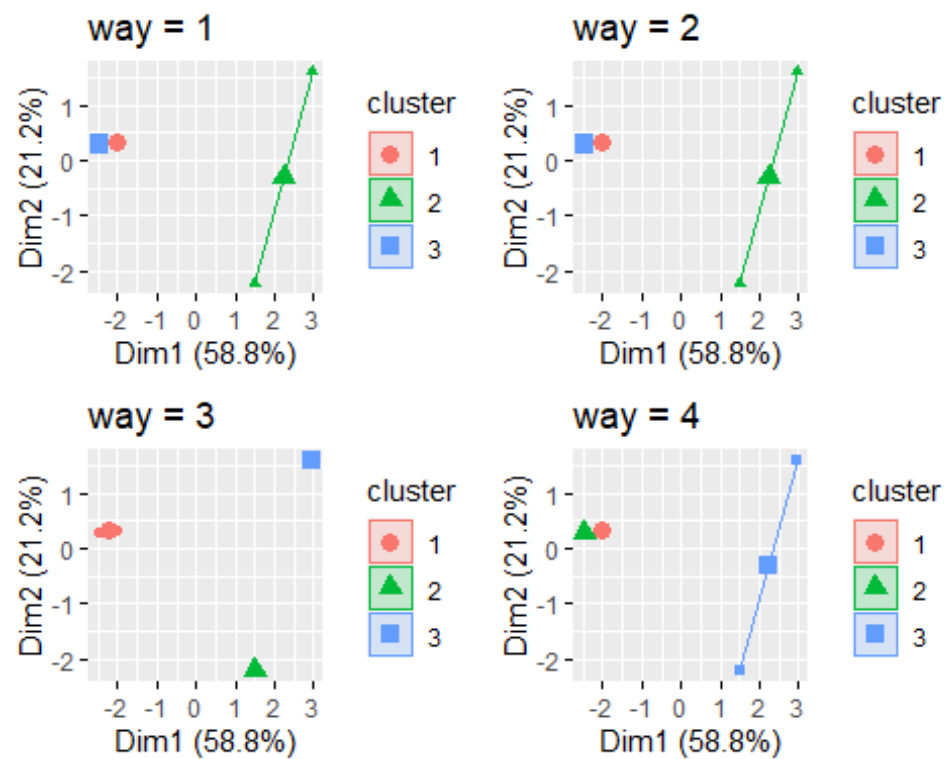


ابتدا سال‌های 1971 و 1973 با یکدیگر خوشه شده‌اند و سپس آن دو با سال 1970 و سپس این 3 با 1972 و در نهایت یک خوشه شده‌اند.

```
library(gridExtra)

model2 = kmeans(data , 3 , algorithm = "Hartigan-Wong")
model3 = kmeans(data , 3 , algorithm = "Lloyd")
model4 = kmeans(data , 3 , algorithm = "Forgy")
model5 = kmeans(data , 3 , algorithm = "MacQueen")

p1 <- fviz_cluster(model2, geom = "point", data = data) + ggtitle("way = 1")
p2 <- fviz_cluster(model3, geom = "point", data = data) + ggtitle("way = 2")
p3 <- fviz_cluster(model4, geom = "point", data = data) + ggtitle("way = 3")
p4 <- fviz_cluster(model5, geom = "point", data = data) + ggtitle("way = 4")
grid.arrange(p1, p2, p3, p4, nrow = 2)
```



به زیبایی هرچه تمام تر و کمک گرفتن از مولفه های اصلی، برای رستم میبینیم که چقدر زیبا و قشنگ توانسته که این داده ها را از یکدیگر جدا و خوشه بندی بکند.