

Actividad1_10

Ricardo Kaleb Flores Alfonso

2024-10-10

1) Lea los datos con `read.csv` y cree una nueva matriz sin las variables fecha y hora, finalmente, cree una nueva matriz estandarizada.

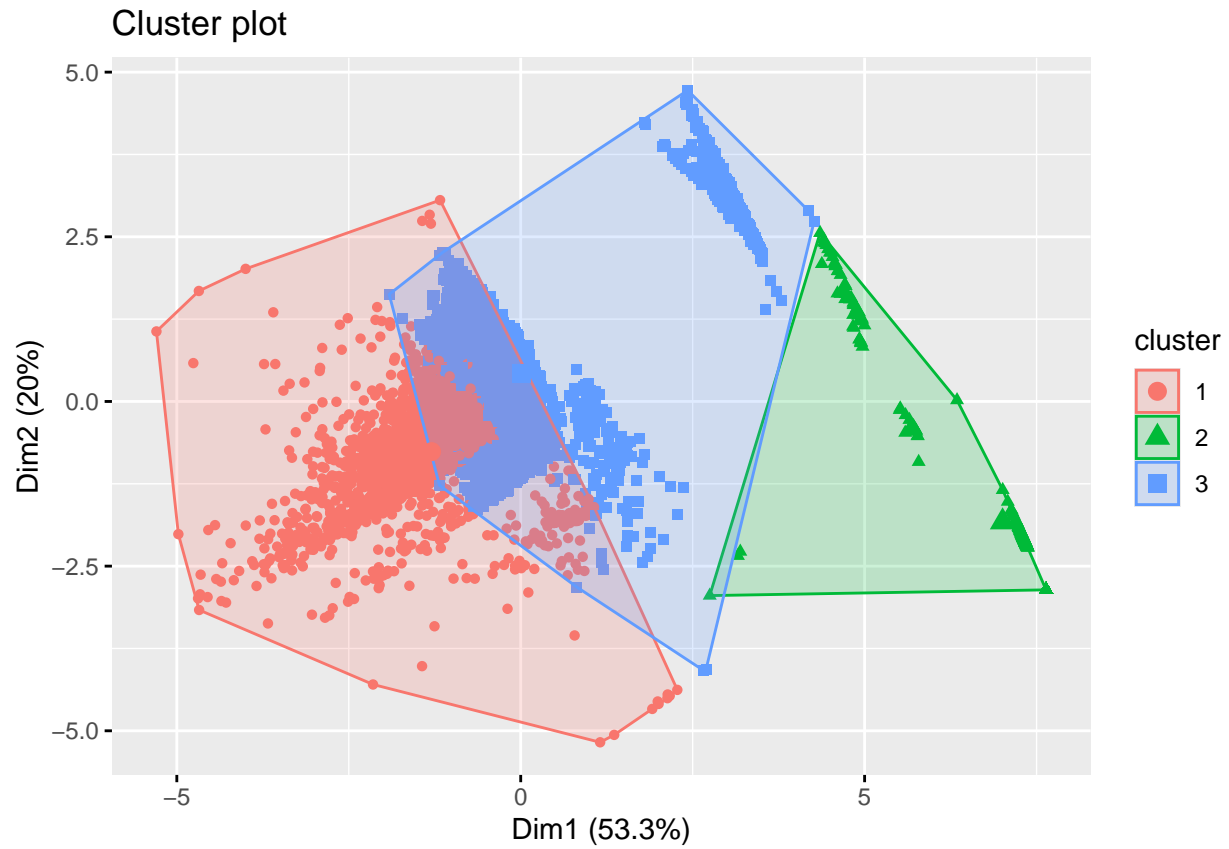
```
M = read.csv("TLA2021.csv")
names(M)
```

```
## [1] "FECHA" "CO"      "NO"      "NO2"     "NOX"     "O3"      "PM10"    "SO2"
```

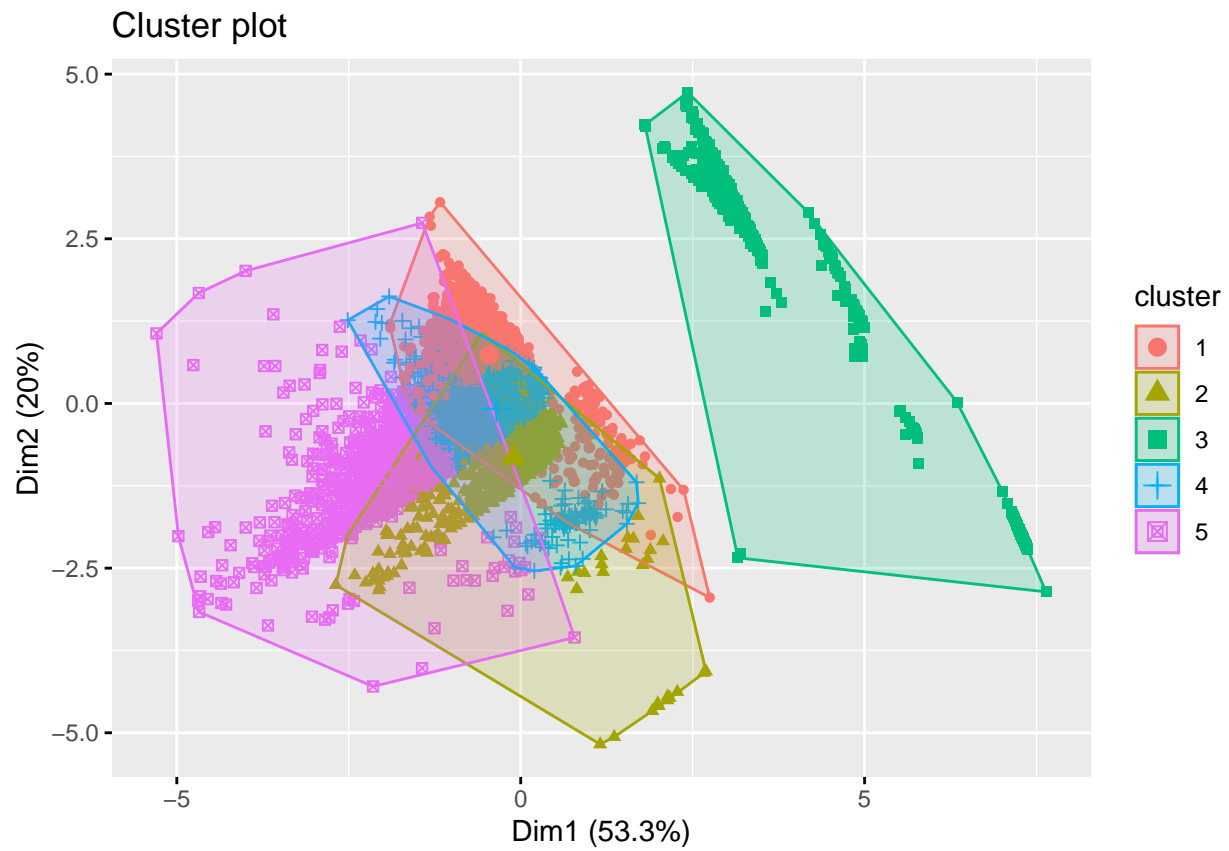
```
M1 = M[, -1] # quitando la columna de fecha
Mstand = scale(x = M1, center = TRUE, scale = TRUE)
```

2. Aplique el algoritmo de k-means y su gráfico con las funciones `kmeans` y `fviz_cluster`. Podría serles útil:

```
M1f = data.frame(Mstand) # Se convierte la matriz Mstand a un data.frame porque así lo requiere la fun
km_clusters = kmeans(M1f, centers = 3)
fviz_cluster(object = km_clusters, data = M1f, show.clust.cent = TRUE, ellipse.type = "convex", star.pl
```

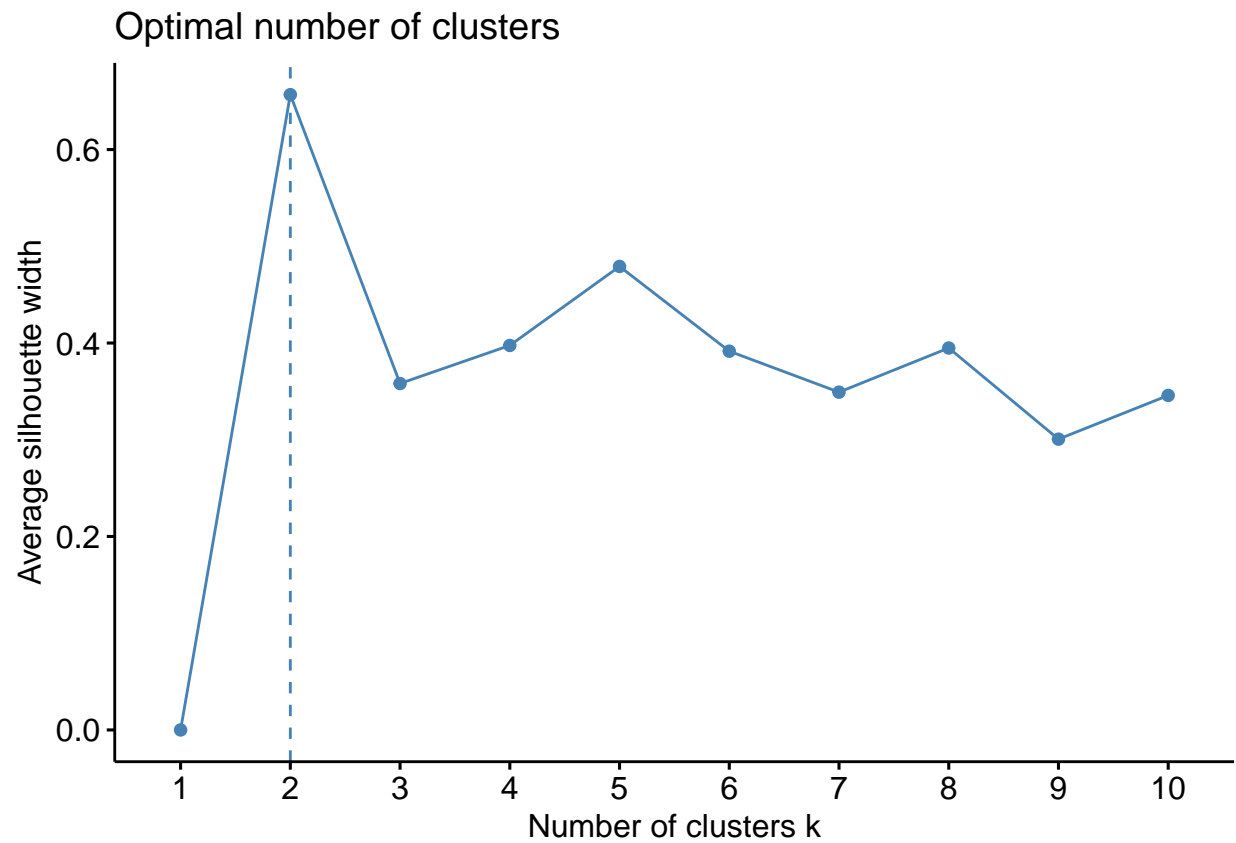


```
M1f = data.frame(Mstand) # Se convierte la matriz Mstand a un data.frame porque así lo requiere la fun
km_clusters = kmeans(M1f, centers = 5)
fviz_cluster(object = km_clusters, data = M1f, show.clust.cent = TRUE, ellipse.type = "convex", star.pl
```

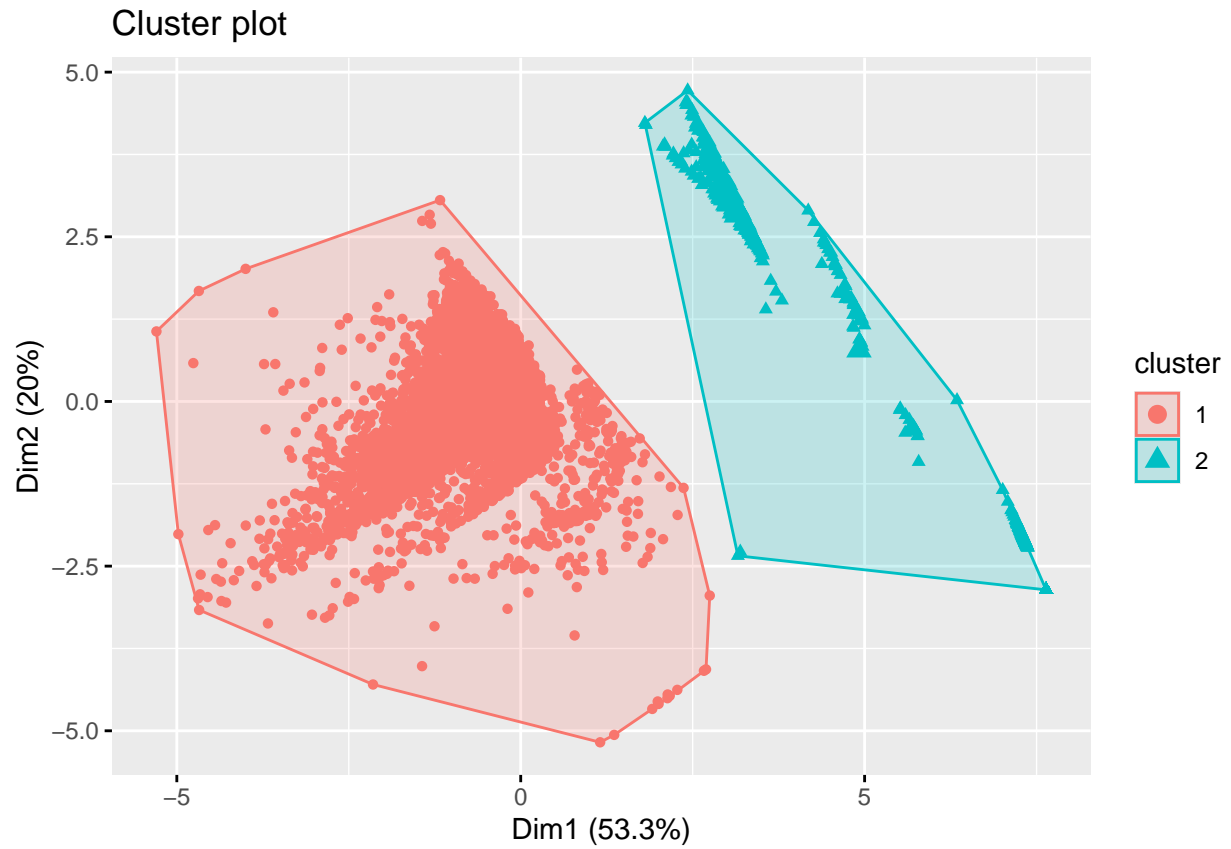


3) Definan el número de agrupaciones óptimas.

```
fviz_nbclust(M1f, kmeans)
```



```
M1f = data.frame(Mstand) # Se convierte la matriz Mstand a un data.frame porque así lo requiere la fun  
km_clusters = kmeans(M1f, centers = 2)  
fviz_cluster(object = km_clusters, data = M1f, show.clust.cent = TRUE, ellipse.type = "convex", star.pl
```



Se observa que la cantidad máxima recomendada de clusters son 2, sin embargo es posible segmentarlos en la cantidad que se requiera dependiendo de cual sea la necesidad a cumplir cuando se realiza este metodo.