

Resultados Preliminares

José Eduardo Zárate Aranda

3/20/2021

Inicialización del código

```
shell("cls")
rm(list = ls(all.names = TRUE))

#Imports Dataset
library(readxl)
DatosDeVinos <- read_excel("~/R/ProyectoAspectosAnaliticos/DatosDeVinos.xlsx", sheet=2)
DatosDeVinos <- as.data.frame(DatosDeVinos)
#Creates rownames for each row with variable Type
str(DatosDeVinos)
```

```
## 'data.frame': 30 obs. of 18 variables:
## $ Type : chr "Vino 1" "Vino 2" "Vino 3" "Vino 4" ...
## $ Malic_Acid : num 1.659 2.256 3.118 0.071 0.429 ...
## $ Ethanol : num 13.4 11.8 12 13.1 14 ...
## $ Total_Acid : num 3.24 3.78 3.91 3.47 3.67 ...
## $ Volatile_Acid : num 0.202 0.228 0.381 0.556 0.311 ...
## $ Lactic_Acid : num 0.221 0.302 0.264 1.664 1.11 ...
## $ Tartaric_Acid : num 1.93 1.57 1.14 2.31 2.48 ...
## $ Glucose_Fructose : num 9.802 0.403 0.396 -0.111 0.381 ...
## $ Density : num 0.996 0.993 0.992 0.995 0.995 ...
## $ Folin_Ciocalteu_Index: num 3.45 10.41 11.73 60.37 55.78 ...
## $ Glycerol : num 7.19 7.12 5.73 8.93 9.03 ...
## $ Gluconic_Acid : num 0.1251 0.3452 -0.0372 0.5514 -1.1448 ...
## $ Sorbic_Acid : num -1.23 37.37 84.05 -8.16 23.88 ...
## $ CO2 : num 615 583 766 392 243 ...
## $ Citric_Acid : num 0.663 0.323 0.368 0.114 0.241 ...
## $ Methanol : num 0.0591 0.0534 0.024 0.1803 0.2028 ...
## $ Ethylacetate : num 39.6 470.1 569.1 -165.9 43.7 ...
## $ pH : num 3.41 3.24 3.36 3.69 3.39 ...
```

```
rownames(DatosDeVinos) <- c(DatosDeVinos$Type)
head(DatosDeVinos)
```

```
##      Type Malic_Acid Ethanol Total_Acid Volatile_Acid Lactic_Acid
## Vino 1 Vino 1  1.659100  13.362    3.2351    0.20197    0.22088
## Vino 2 Vino 2  2.256100  11.848    3.7846    0.22775    0.30157
## Vino 3 Vino 3  3.117500  12.018    3.9126    0.38062    0.26350
## Vino 4 Vino 4  0.070986  13.050    3.4736    0.55592    1.66410
## Vino 5 Vino 5  0.429280  13.990    3.6727    0.31076    1.11050
## Vino 6 Vino 6  0.280540  12.111    3.3800    0.41086    1.50460
##      Tartaric_Acid Glucose_Fructose Density Folin_Ciocalteu_Index Glycerol
```

```
## Vino 1      1.9346      9.802000 0.9961000      3.4546 7.188799
## Vino 2      1.5710      0.403360 0.9929101      10.4120 7.121000
## Vino 3      1.1353      0.396450 0.9921400      11.7330 5.728500
## Vino 4      2.3082     -0.111440 0.9953400      60.3720 8.930700
## Vino 5      2.4753      0.380610 0.9946800      55.7770 9.032600
## Vino 6      2.2342     -0.044855 0.9943800      50.7710 7.971600
##      Gluconic_Acid Sorbic_Acid      CO2 Citric_Acid Methanol Ethylacetate
## Vino 1      0.125110     -1.2281 615.3300      0.66330 0.059090      39.568
## Vino 2      0.345230     37.3650 582.9600      0.32346 0.053365      470.080
## Vino 3     -0.037167     84.0530 766.1199      0.36837 0.024013      569.080
## Vino 4      0.551400     -8.1593 391.7100      0.11382 0.180300     -165.890
## Vino 5     -1.144800     23.8830 242.9900      0.24131 0.202810      43.715
## Vino 6     -0.192260    -43.5150 448.7600      0.14329 0.119020      340.950
##      pH
## Vino 1 3.41
## Vino 2 3.24
## Vino 3 3.36
## Vino 4 3.69
## Vino 5 3.39
## Vino 6 3.48
```

```
#Removes Type variable
```

```
DatosDeVinos <- DatosDeVinos[,-1]
head(DatosDeVinos)
```

```
##      Malic_Acid Ethanol Total_Acid Volatile_Acid Lactic_Acid Tartaric_Acid
## Vino 1  1.659100  13.362    3.2351    0.20197    0.22088    1.9346
## Vino 2  2.256100  11.848    3.7846    0.22775    0.30157    1.5710
## Vino 3  3.117500  12.018    3.9126    0.38062    0.26350    1.1353
## Vino 4  0.070986  13.050    3.4736    0.55592    1.66410    2.3082
## Vino 5  0.429280  13.990    3.6727    0.31076    1.11050    2.4753
## Vino 6  0.280540  12.111    3.3800    0.41086    1.50460    2.2342
##      Glucose_Fructose Density Folin_Ciocalteu_Index Glycerol Gluconic_Acid
## Vino 1      9.802000 0.9961000      3.4546 7.188799      0.125110
## Vino 2      0.403360 0.9929101      10.4120 7.121000      0.345230
## Vino 3      0.396450 0.9921400      11.7330 5.728500     -0.037167
## Vino 4     -0.111440 0.9953400      60.3720 8.930700      0.551400
## Vino 5      0.380610 0.9946800      55.7770 9.032600     -1.144800
## Vino 6     -0.044855 0.9943800      50.7710 7.971600     -0.192260
##      Sorbic_Acid      CO2 Citric_Acid Methanol Ethylacetate      pH
## Vino 1     -1.2281 615.3300      0.66330 0.059090      39.568 3.41
## Vino 2     37.3650 582.9600      0.32346 0.053365      470.080 3.24
## Vino 3     84.0530 766.1199      0.36837 0.024013      569.080 3.36
## Vino 4     -8.1593 391.7100      0.11382 0.180300     -165.890 3.69
## Vino 5     23.8830 242.9900      0.24131 0.202810      43.715 3.39
## Vino 6    -43.5150 448.7600      0.14329 0.119020      340.950 3.48
```

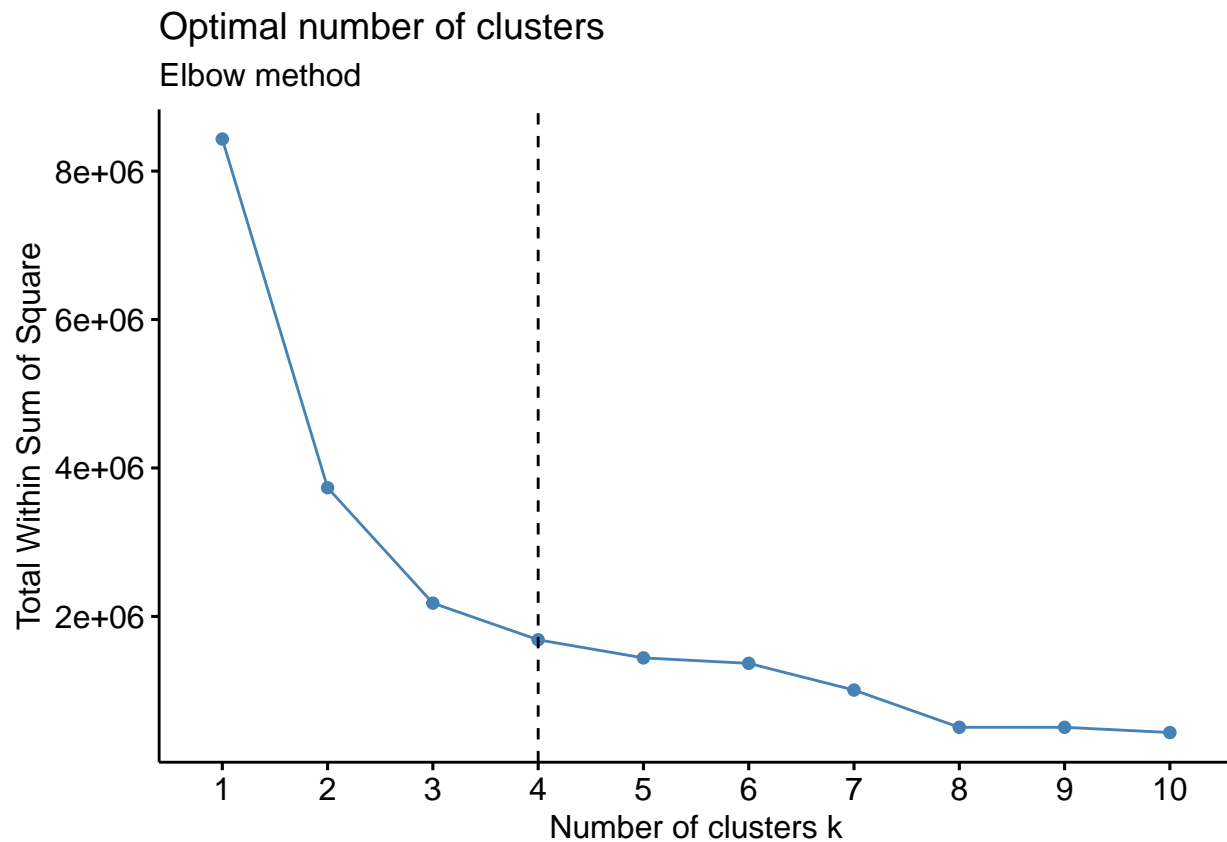
```
#Reviews the k value
```

```
library(factoextra)
```

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

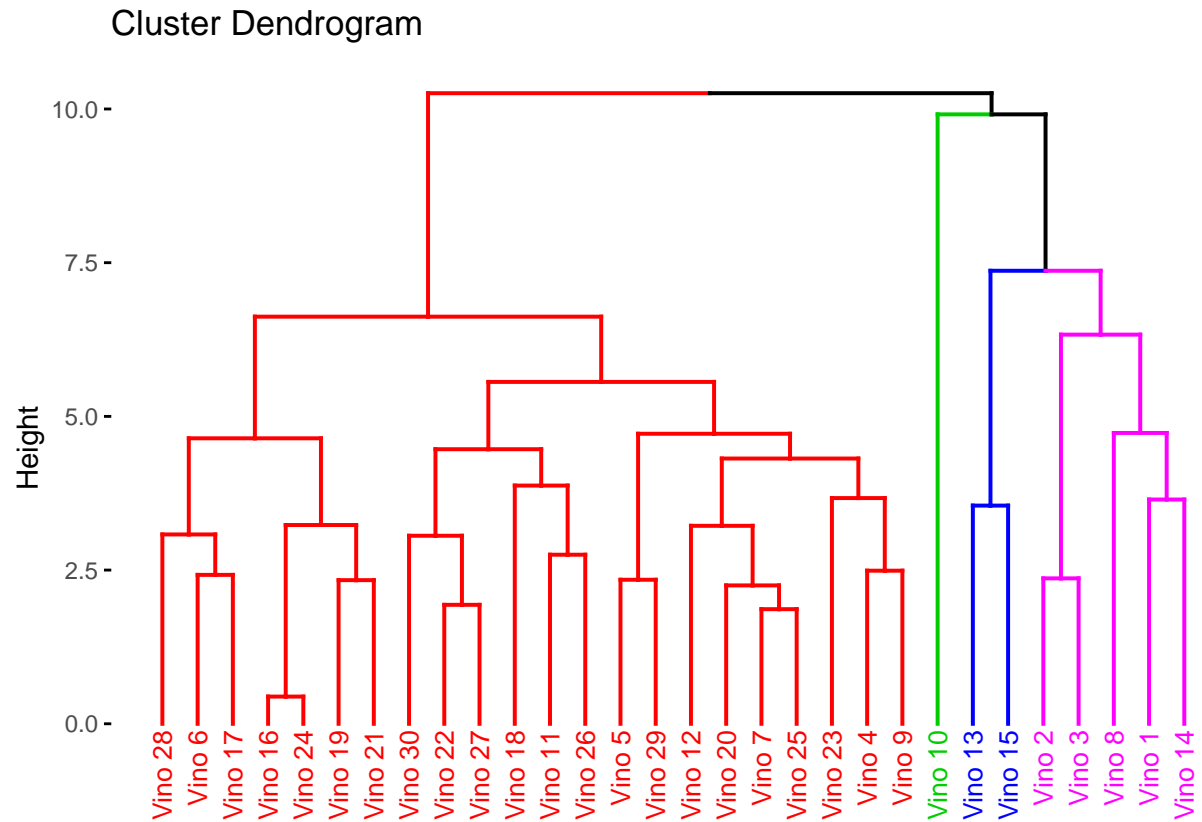
```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
fviz_nbclust(DatosDeVinos,kmeans,method = "wss")+geom_vline(xintercept = 4, linetype=2)+labs(subtitle="Elbow method")
```



```
#Scales the dataset by taking the mean deviations of each column and by dividing the centered columns o
DatosDeVinos <- scale(DatosDeVinos)
#Computes and returns the distance matrix computed by using the specified distance measure to compute t
library(stats)
res.dist <- dist(x=DatosDeVinos, method="euclidean")
output <- as.matrix(res.dist)
#Applies the hierarchical clustering
res.hc <- hclust(d=res.dist,method = "complete")

#Formats and plots the dendrogram:
#Add colors to clusters
fviz_dend(x=res.hc, cex=0.7, lwd=0.7, k=4, k_colors=c("red","green3","blue","magenta"))
```



```
#Delimits clusters by squares and fills them
fviz_dend(res.hc, cex=0.8,lwd=0.8,k=4,rect=TRUE,k_colors="jco",rect_border = "jco",rect_fill = TRUE,, y
```

Dendograma de los 30 tipos de vino

