TAREA3

María Pallares Diez

2025-01-23

Introducción

Este documento presenta un análisis exhaustivo de clúster basado en los datos del archivo `datosp1_b.RData`. El objetivo principal es identificar los clústeres bien definidos siguiendo los métodos explicados en las clases de teoría.

Carga y Exploración de Datos

```
load("datosp1_b.RData")
datos<-datos_b
str(datos)
```

```
## 'data.frame':    100 obs. of  12 variables:
##  $ ID                  : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Edad                : num  17 16.4 17.1 18 16.3 16.3 18 17.3 16 17 ...
##  $ Horas_estudio       : num  9.56 12.65 9.99 6.93 7.5 ...
##  $ Promedio_matematicas: num  50 54.7 55.5 56.3 58.9 59.8 60.6 60.7 61.2 62.4 ...
##  $ Promedio_ciencias   : num  55 53.3 56.2 51.3 61.7 ...
##  $ Promedio_lectura    : num  61.8 56 56.1 55.9 57 ...
##  $ Asistencia          : num  72.5 62.8 57.6 56.6 54.9 ...
##  $ Horas_sueño         : num  4.0762 6.302 0.0622 5.0344 4.5811 ...
##  $ Nivel_estres        : num  2.6 6.4 4.5 4.3 6.4 5.9 4.3 7.3 2.8 6.2 ...
##  $ Uso_dispositivos    : num  7.96 7.87 8.11 7.02 6.31 ...
##  $ Condicion_fisica    : num  143 120 218 114 265 ...
##  $ Centro              : Factor w/ 9 levels "1","2","3","4",..: 1 1 1 1 1 1 1 1 1 1 ...
```

```
summary(datos)
```

```
##        ID              Edad       Horas_estudio    Promedio_matematicas
##  Min.   :  1.00   Min.   :15.00   Min.   : 6.816   Min.   : 50.00
##  1st Qu.: 25.75   1st Qu.:15.90   1st Qu.:16.229   1st Qu.: 68.47
##  Median : 50.50   Median :16.40   Median :19.137   Median : 75.95
##  Mean   : 50.50   Mean   :16.42   Mean   :19.452   Mean   : 75.59
##  3rd Qu.: 75.25   3rd Qu.:16.93   3rd Qu.:23.204   3rd Qu.: 82.08
##  Max.   :100.00   Max.   :18.00   Max.   :32.598   Max.   :100.00
##
##  Promedio_ciencias Promedio_lectura   Asistencia      Horas_sueño
##  Min.   : 51.28    Min.   : 50.23   Min.   : 48.47   Min.   : 0.06225
```

```
##  1st Qu.: 68.18    1st Qu.: 65.61   1st Qu.: 62.77   1st Qu.: 6.66938
##  Median : 77.32    Median : 75.81   Median : 71.32   Median : 9.40755
##  Mean   : 75.51    Mean   : 75.53   Mean   : 74.38   Mean   : 9.72000
##  3rd Qu.: 83.34    3rd Qu.: 86.52   3rd Qu.: 86.30   3rd Qu.:13.12855
##  Max.   :100.33    Max.   :102.70   Max.   :106.49   Max.   :19.65346
##
##   Nivel_estres    Uso_dispositivos   Condicion_fisica     Centro
##  Min.   : 1.000   Min.   :-0.04651   Min.   :-69.6    1      :12
##  1st Qu.: 3.575   1st Qu.: 2.40814   1st Qu.:110.3    2      :11
##  Median : 5.050   Median : 3.45247   Median :170.4    3      :11
##  Mean   : 4.960   Mean   : 3.74339   Mean   :158.9    4      :11
##  3rd Qu.: 6.200   3rd Qu.: 4.84802   3rd Qu.:209.2    5      :11
##  Max.   :10.000   Max.   : 8.11393   Max.   :478.3    6      :11
##                                                      (Other):33
```

Preprocesamiento

Estandarización y Normalización

Dado que las variables pueden tener diferentes escalas, realizamos estandarización para garantizar la comparabilidad entre ellas. En el proceso de análisis de clustering, se eliminaron las variables `ID` y `Centro` debido a las siguientes razones:

- La variable `ID` es únicamente un identificador único para las observaciones y no contiene información útil para agrupar datos basados en similitudes.
- La variable `Centro`, aunque categórica, no aporta información que diferencie significativamente las observaciones desde una perspectiva de clustering, ya que está relacionada con una clasificación previa.

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
datos <- datos %>% select(-ID, -Centro)
datos_numericos <- datos %>% select(where(is.numeric))
datos_estandarizados <- as.data.frame(scale(datos_numericos))
summary(datos_estandarizados)
```

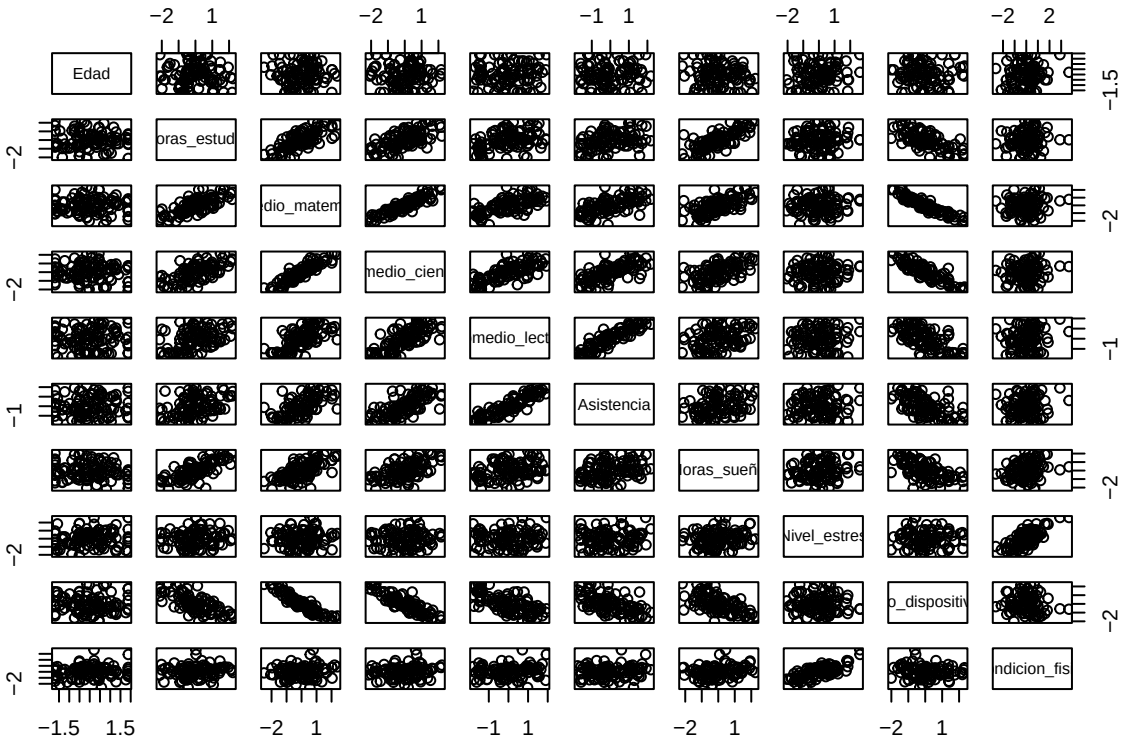
```
##       Edad          Horas_estudio      Promedio_matematicas Promedio_ciencias
##  Min.   :-1.69629   Min.   :-2.13684   Min.   :-2.50141     Min.   :-2.1531
##  1st Qu.:-0.62269   1st Qu.:-0.54510   1st Qu.:-0.69520     1st Qu.:-0.6514
##  Median :-0.02624   Median :-0.05333   Median : 0.03559     Median : 0.1605
```

```
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000     Mean   : 0.0000
##  3rd Qu.: 0.60003   3rd Qu.: 0.63444   3rd Qu.: 0.63440     3rd Qu.: 0.6959
##  Max.   : 1.88239   Max.   : 2.22304   Max.   : 2.38683     Max.   : 2.2052
##  Promedio_lectura     Asistencia       Horas_sueño        Nivel_estres
##  Min.   :-1.79993   Min.   :-1.7660   Min.   :-2.18980   Min.   :-2.0590
##  1st Qu.:-0.70626   1st Qu.:-0.7912   1st Qu.:-0.69170   1st Qu.:-0.7201
##  Median : 0.01945   Median :-0.2081   Median :-0.07084   Median : 0.0468
##  Mean   : 0.00000   Mean   : 0.0000   Mean   : 0.00000   Mean   : 0.0000
##  3rd Qu.: 0.78160   3rd Qu.: 0.8131   3rd Qu.: 0.77286   3rd Qu.: 0.6447
##  Max.   : 1.93235   Max.   : 2.1891   Max.   : 2.25231   Max.   : 2.6205
##  Uso_dispositivos  Condicion_fisica
##  Min.   :-2.0109   Min.   :-2.6282
##  1st Qu.:-0.7085   1st Qu.:-0.5590
##  Median :-0.1544   Median : 0.1327
##  Mean   : 0.0000   Mean   : 0.0000
##  3rd Qu.: 0.5861   3rd Qu.: 0.5791
##  Max.   : 2.3190   Max.   : 3.6755
```

Análisis Exploratorio

Antes de proceder con el análisis, se verifica la posible presencia de outliers y patrones generales.

```
library(ggplot2)
pairs(datos_estandarizados)
```



Análisis del Gráfico de Pares

El gráfico de pares muestra las relaciones entre las diferentes variables numéricas del conjunto de datos estandarizados, permitiendo identificar patrones, correlaciones, posibles agrupaciones y valores atípicos que podrían influir en el análisis de clustering.

Correlaciones entre Variables:

- Algunas variables, como `Promedio_matematicas` y `Promedio_ciencias`, presentan una fuerte correlación positiva, indicando que los estudiantes con mejor desempeño en matemáticas tienden a destacar también en ciencias.
- De manera similar, `Promedio_lectura` y `Asistencia` muestran una relación positiva, lo que sugiere que la asistencia regular puede estar asociada con mejores resultados en lectura.
- Estas correlaciones sugieren que estas variables podrían contribuir juntas a la formación de clusters, simplificando el problema al reducir la dimensionalidad efectiva.

Dispersión de los Datos:

- Variables como `Nivel_estres` y `Horas_sueño` exhiben distribuciones más dispersas, sin patrones evidentes de relación con otras variables, lo que indica que podrían aportar información complementaria para diferenciar observaciones en el espacio multidimensional.
- De forma similar, `Uso_dispositivos` muestra cierta dispersión, aunque su influencia en la segmentación dependerá de cómo interactúe con otras variables.

Outliers (Valores Atípicos):

- Se detectan valores atípicos en algunas variables donde se observan individuos con valores significativamente alejados del resto.
- Estos outliers podrían distorsionar las métricas de distancia, como la euclidiana, y afectar los resultados del clustering si no se tratan adecuadamente.

Clusters Potenciales:

- Aunque no se identifican agrupaciones claras directamente desde el gráfico, algunos patrones de alineación, como los observados entre `Promedio_matematicas` y `Promedio_ciencias`, sugieren la existencia de clusters naturales en el espacio multidimensional.
- Esto puede ser confirmado y detallado aplicando métodos de clustering como K-means o jerárquico.

Encontrar la mejor combinación de distancias y métodos jerárquicos

La función 'find_best_clustering' está diseñada para probar todas las combinaciones posibles de métricas de distancia y métodos de clustering jerárquico con el objetivo de encontrar la combinación que maximiza la correlación cofenética. La correlación cofenética mide qué tan bien el dendrograma generado refleja las distancias originales entre las observaciones.

```
find_best_clustering <- function(datos_estandarizados) {
  library(cluster)
  library(stats)
  distance_methods <- c("euclidean", "maximum", "manhattan", "canberra", "binary", "minkowski")
  hclust_methods <- c("average", "single", "complete", "ward.D", "ward.D2", "mcquitty", "median", "centi
```

```r
  results <- data.frame(
    Distance = character(),
    Method = character(),
    CopheneticCorrelation = numeric(),
    stringsAsFactors = FALSE
  )
  for (distance in distance_methods) {
    dist_matrix <- tryCatch(
      dist(datos_estandarizados, method = distance),
      error = function(e) {
        message(paste("Error with distance:", distance, "-", e$message))
        return(NULL)
      }
    )
    if (!is.null(dist_matrix)) {
      for (method in hclust_methods) {
        # Calcular el modelo jerárquico
        hclust_model <- tryCatch(
          hclust(dist_matrix, method = method),
          error = function(e) {
            message(paste("Error with method:", method, "-", e$message))
            return(NULL)
          }
        )
        if (!is.null(hclust_model)) {
          if (sd(as.vector(dist_matrix)) > 0 && sd(cophenetic(hclust_model)) > 0) {
            cophenetic_corr <- cor(cophenetic(hclust_model), dist_matrix)
          } else {
            cophenetic_corr <- NA
          }
          results <- rbind(results, data.frame(
            Distance = distance,
            Method = method,
            CopheneticCorrelation = cophenetic_corr
          ))
        }
      }
    }
  }
  results <- results[!is.na(results$CopheneticCorrelation), ]
  best_result <- results[which.max(results$CopheneticCorrelation), ]
  list(
    BestCombination = best_result,
    AllResults = results
  )
}
resultado <- find_best_clustering(datos_estandarizados)
print(resultado$BestCombination)
```

```
##    Distance  Method CopheneticCorrelation
## 25 canberra average            0.7701103
```

```r
View(resultado$AllResults)
```

Vemos que la combinación que ofrece una mayor correlación cofenética es la distancia de **Canberra** con el método jerárquico **Average**

Cálculo de Distancias

Se han probado diferentes métricas de distancia para determinar cuál se adapta mejor a las características de los datos y ofrece una representación más fiel de las relaciones entre las observaciones. Finalmente, hemos seleccionado la **distancia Canberra**, ya que obtuvo la correlación cofenética más alta (**0.7701**), lo que indica que el dendrograma generado por esta métrica refleja mejor las distancias originales.

Propiedades de la Distancia Canberra

1. **Definición Matemática**:
   - La distancia Canberra calcula las diferencias absolutas entre las coordenadas de dos puntos, normalizadas por la suma de sus valores absolutos:

$$d(x, y) = \sum_{i=1}^{n} \frac{|x_i - y_i|}{|x_i| + |y_i|}$$

   - Esto significa que cada componente contribuye proporcionalmente a la distancia total, dependiendo de su magnitud.
2. **Sensibilidad a Diferencias Relativas**:
   - A diferencia de otras métricas como la distancia euclidiana, la distancia Canberra presta especial atención a las diferencias relativas entre los valores, lo que la hace útil cuando las variables tienen diferentes escalas o rangos.
3. **Manejo de Valores Pequeños**:

- Las dimensiones con valores pequeños tienen un mayor peso relativo en el cálculo, permitiendo que pequeñas diferencias en estas dimensiones influyan en la distancia total.

Contexto del Análisis

En este análisis, los datos incluyen variables numéricas estandarizadas que presentan posibles outliers y diferentes escalas. Tras probar varias métricas de distancia, se observó que la distancia Canberra generó la correlación cofenética más alta (**0.7701**), indicando que proporciona la mejor representación de las relaciones entre las observaciones. Esto asegura que las diferencias relativas entre las observaciones sean capturadas de manera precisa, ofreciendo un análisis más robusto y representativo de la estructura de los datos.

```r
library(stats)
distancias <- dist(datos_estandarizados, method = "canberra", diag = FALSE, upper =
FALSE, p = 2)
```

Análisis de Clúster

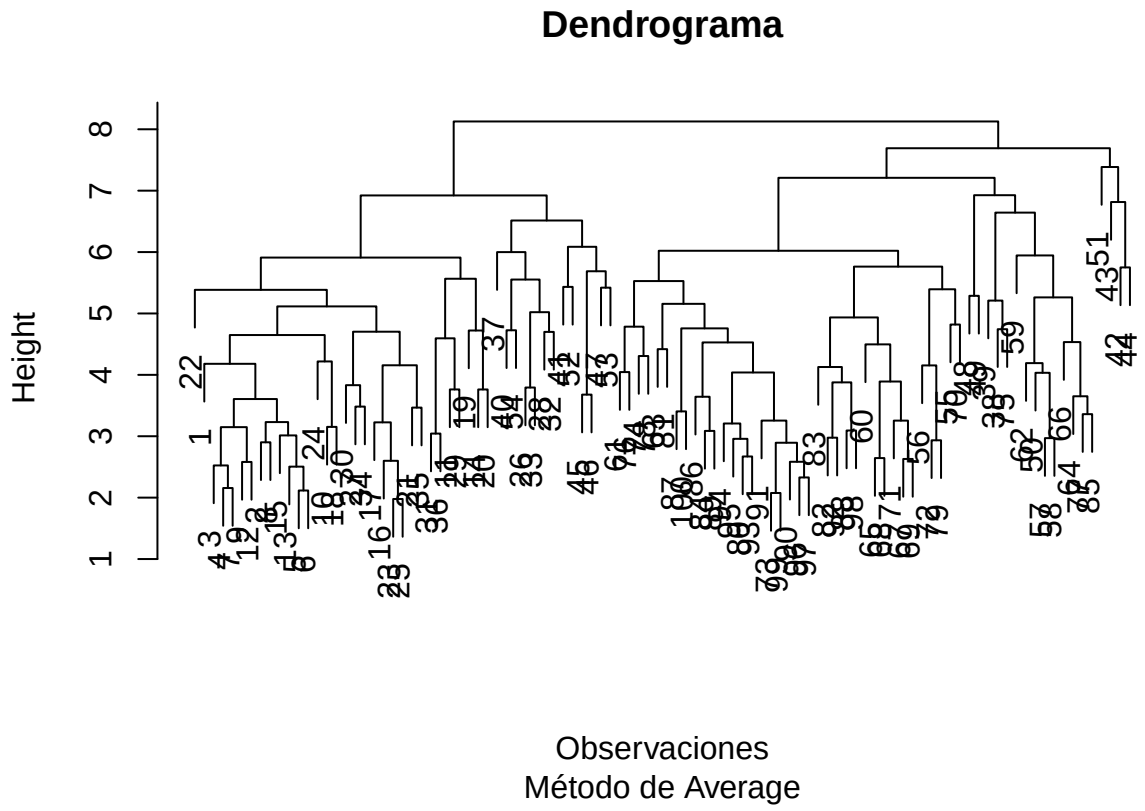
Método Jerárquico

El método **Average** fue seleccionado para el análisis de clustering jerárquico porque ofrece un equilibrio entre cohesión interna y separación externa. Calcula las distancias entre clusters como el promedio de las distancias entre todas las observaciones de cada grupo, asegurando que todos los puntos sean considerados en la formación de los clusters.

Ventajas del Método Average

1. **Equilibrio entre Cohesión y Separación**:
   - Forma clusters balanceados, evitando los problemas de `single-linkage` (clusters alargados) y `complete-linkage` (clusters muy compactos).
2. **Robustez frente a Outliers**:
   - Es menos sensible a valores atípicos, ya que utiliza promedios en lugar de depender de un único par de puntos.
3. **Resultados Consistentes**:
   - Captura estructuras jerárquicas de forma representativa, incluso en datos complejos.

```
library(cluster)
hclust_model <- hclust(distancias, method = "average")
plot(hclust_model, main = "Dendrograma", xlab = "Observaciones", sub = "Método de Average")
```

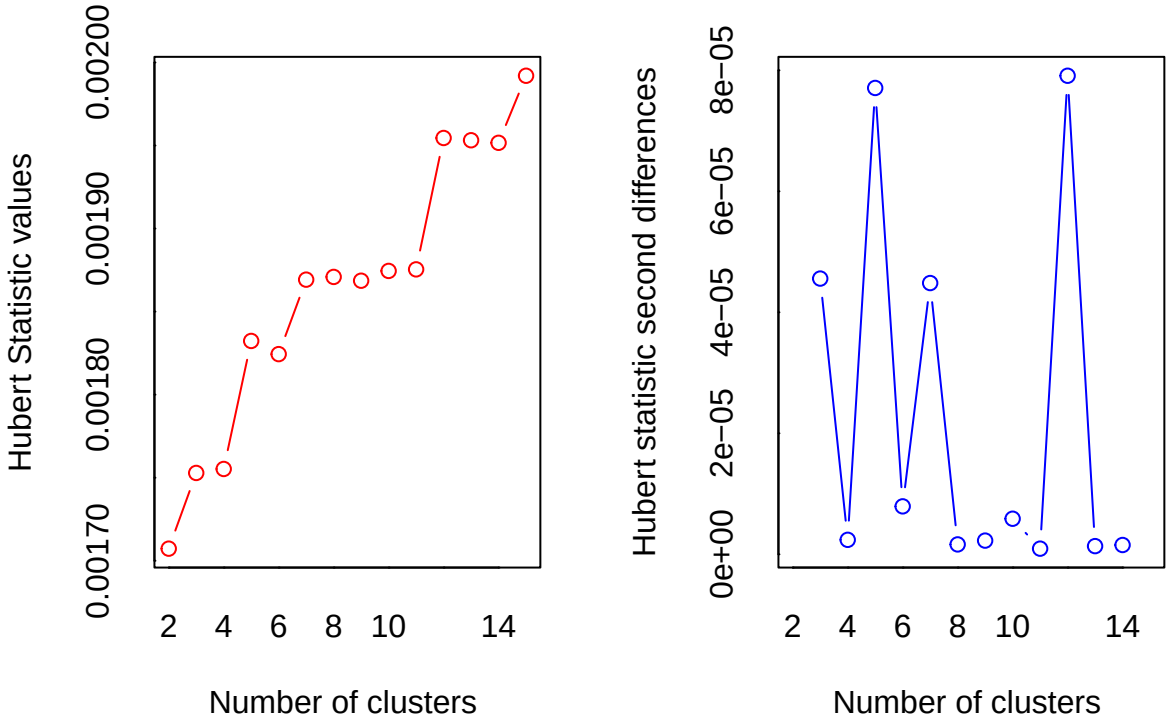


Podemos observar claramente 2 grupos.

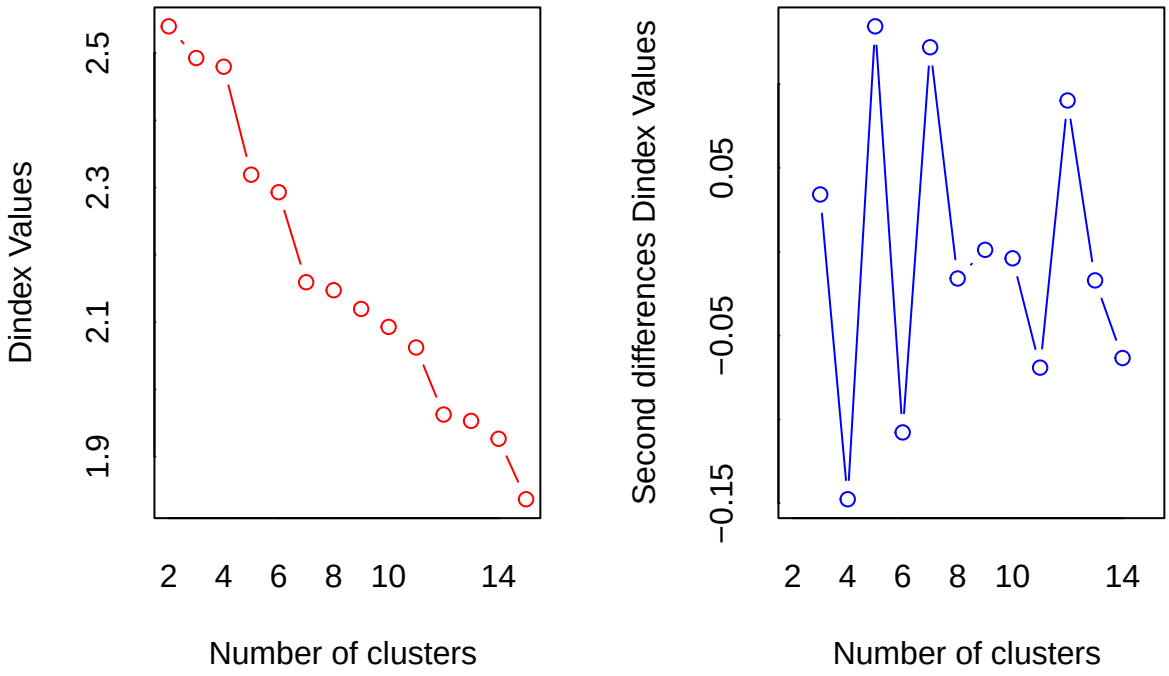
Método K-means

Se determina el número óptimo de clústeres mediante `NbClust` y se aplica k-means con 2 grupos.

```
library(NbClust)
nbclust.average <- NbClust(data = datos_estandarizados,
                           diss = NULL ,
                           distance="canberra",
                           method ="average")
```




```
## *** : The Hubert index is a graphical method of determining the number of clusters.
##                 In the plot of Hubert index, we seek a significant knee that corresponds to a
##                 significant increase of the value of the measure i.e the significant peak in Hubert
##                 index second differences plot.
##
```




```
## *** : The D index is a graphical method of determining the number of clusters.
##                 In the plot of D index, we seek a significant knee (the significant peak in Dindex
##                 second differences plot) that corresponds to a significant increase of the value of
##                 the measure.
##
## *******************************************************************
## * Among all indices:
## * 9 proposed 2 as the best number of clusters
## * 1 proposed 3 as the best number of clusters
## * 1 proposed 4 as the best number of clusters
## * 2 proposed 5 as the best number of clusters
## * 3 proposed 7 as the best number of clusters
## * 1 proposed 9 as the best number of clusters
## * 2 proposed 11 as the best number of clusters
## * 1 proposed 13 as the best number of clusters
## * 3 proposed 15 as the best number of clusters
##
##                    ***** Conclusion *****
##
## * According to the majority rule, the best number of clusters is  2
##
##
## *******************************************************************
```

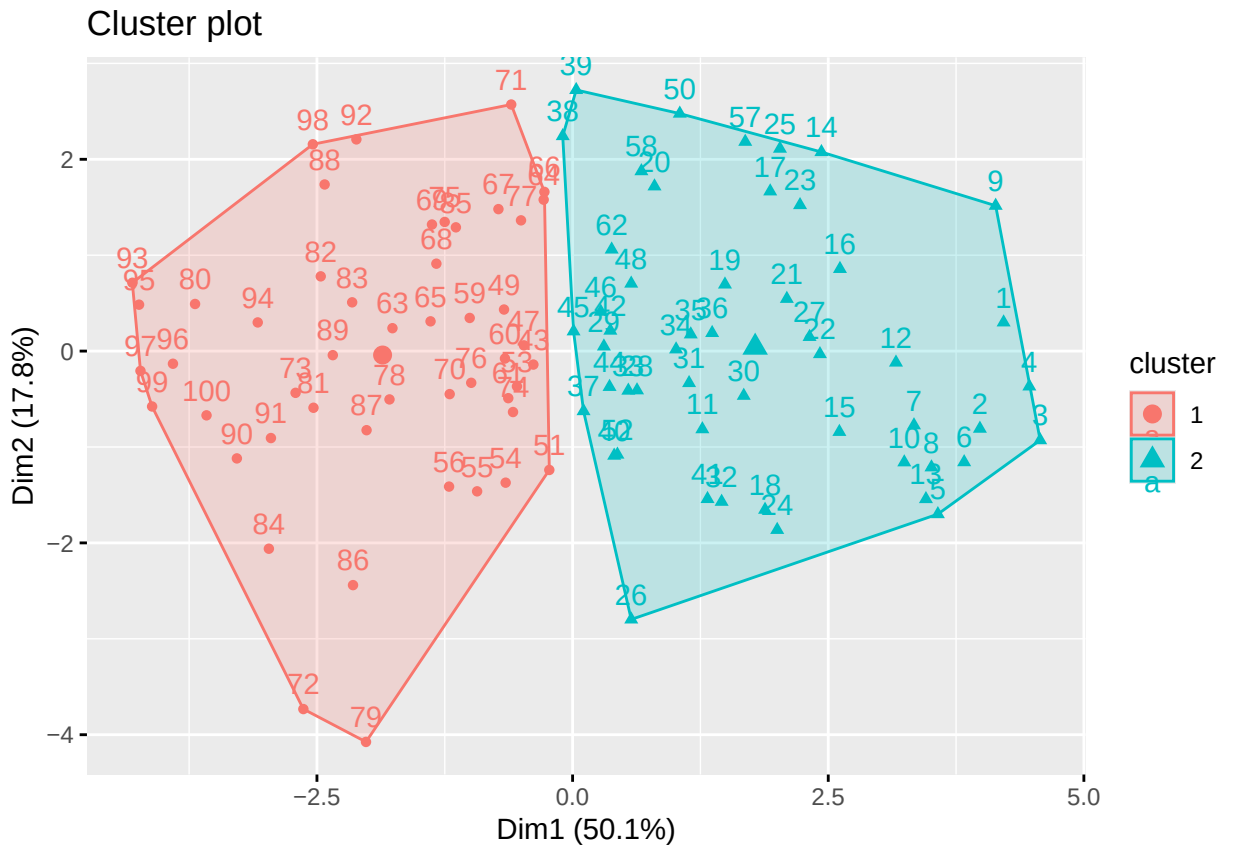
```
print(nbclust.average$Best.nc)
```

```
##                     KL      CH Hartigan     CCC   Scott      Marriot  TrCovW
## Number_clusters 13.0000  2.0000   4.0000  2.0000 15.0000 7.000000e+00    5.00
## Value_Index     32.0747 42.7442  13.5771 -2.7031 64.6884 1.213982e+16 1066.62
##                   TraceW Friedman   Rubin Cindex      DB Silhouette   Duda
## Number_clusters  5.0000   7.0000  7.0000 2.0000 11.0000     2.0000 2.0000
## Value_Index     73.3208   5.1129 -0.1931 0.5283  1.4142     0.2319 0.9615
##                 PseudoT2 Beale Ratkowsky    Ball PtBiserial Frey McClain
## Number_clusters     2.00 2.000    2.0000   3.000     9.0000    1  2.0000
## Value_Index         2.12 0.264    0.3276 120.078     0.6736   NA  0.7649
##                     Dunn Hubert SDindex Dindex    SDbw
## Number_clusters 15.0000      0 11.0000      0 15.0000
## Value_Index      0.3942      0  1.9674      0  0.2859
```

```
kmeans_model <- kmeans(datos_estandarizados, centers = 2, nstart = 50)
library(factoextra)
```

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

```
fviz_cluster(kmeans_model, data = datos_estandarizados)
```



Interpretación del Gráfico de Clusters

El gráfico mostrado representa los resultados del análisis de clustering utilizando el método **K-means** con 2 clusters, seleccionados como el número óptimo tras un análisis con el paquete `NbClust`.

Proceso para Determinar el Número de Clusters

1. **Uso de `NbClust`**:
    - Se utilizó la función `NbClust` para evaluar el número óptimo de clusters en un rango de 2 a 15.
    - Esta herramienta aplica múltiples índices de validación y selecciona el número de clusters más adecuado basado en criterios de cohesión interna y separación entre clusters.
2. **Selección de 2 Clusters**:
    - La mayoría de los índices reportados por `NbClust` recomendaron **2 clusters**, lo que indica que esta es la partición más natural para los datos.

Aplicación del Método average

- Se aplicó el algoritmo average con 2 clusters utilizando 50 inicializaciones aleatorias para garantizar la estabilidad del resultado.
- El algoritmo asigna las observaciones a los clusters basándose en la distancia Canberra, minimizando la suma de las distancias al centroide dentro de cada cluster.

Interpretación del Gráfico

1. **Dimensiones**:
    - El gráfico proyecta las observaciones en un espacio bidimensional utilizando técnicas de reducción de dimensionalidad, como PCA, para facilitar la visualización.
    - La primera dimensión (Dim1) explica el 50.1% de la varianza y la segunda dimensión (Dim2) explica el 17.8%, lo que indica que estas dos dimensiones capturan la mayoría de la información relevante de los datos.
2. **Clusters Identificados**:
    - Los puntos se agrupan en dos clusters bien diferenciados:
        - **Cluster 1 (rojo):** Representa observaciones que comparten características comunes y están más cercanas a su centroide.
        - **Cluster 2 (azul):** Contiene observaciones con características distintas a las del Cluster 1.
    - Las áreas sombreadas muestran la región aproximada ocupada por cada cluster, proporcionando una clara separación visual entre ellos.
3. **Separación y Cohesión**:
    - La distancia entre los centroides indica una buena separación entre los clusters.
    - La cohesión dentro de los clusters es alta, como se observa en la proximidad de los puntos a sus respectivos centroides.

Evaluación y Comparación de Resultados

Resumen de Clústeres

```r
aggregate(datos, by = list(Cluster = kmeans_model$cluster), mean)
```

```
##   Cluster     Edad Horas_estudio Promedio_matematicas Promedio_ciencias
## 1       1 16.58163      23.24765             83.77755          84.37757
## 2       2 16.26863      15.80520             67.71569          66.99262
##   Promedio_lectura Asistencia Horas_sueño Nivel_estres Uso_dispositivos
## 1         85.09392   83.99953   12.015352     5.336735         2.324095
## 2         66.34946   65.13211    7.514656     4.598039         5.107033
##   Condicion_fisica
## 1         171.7043
## 2         146.5058
```

```r
table(kmeans_model$cluster)
```

```
##
##  1  2
## 49 51
```

Vemos como se forman los dos grupos o cluster uno con 51 datos y el otro con 49.

Validación

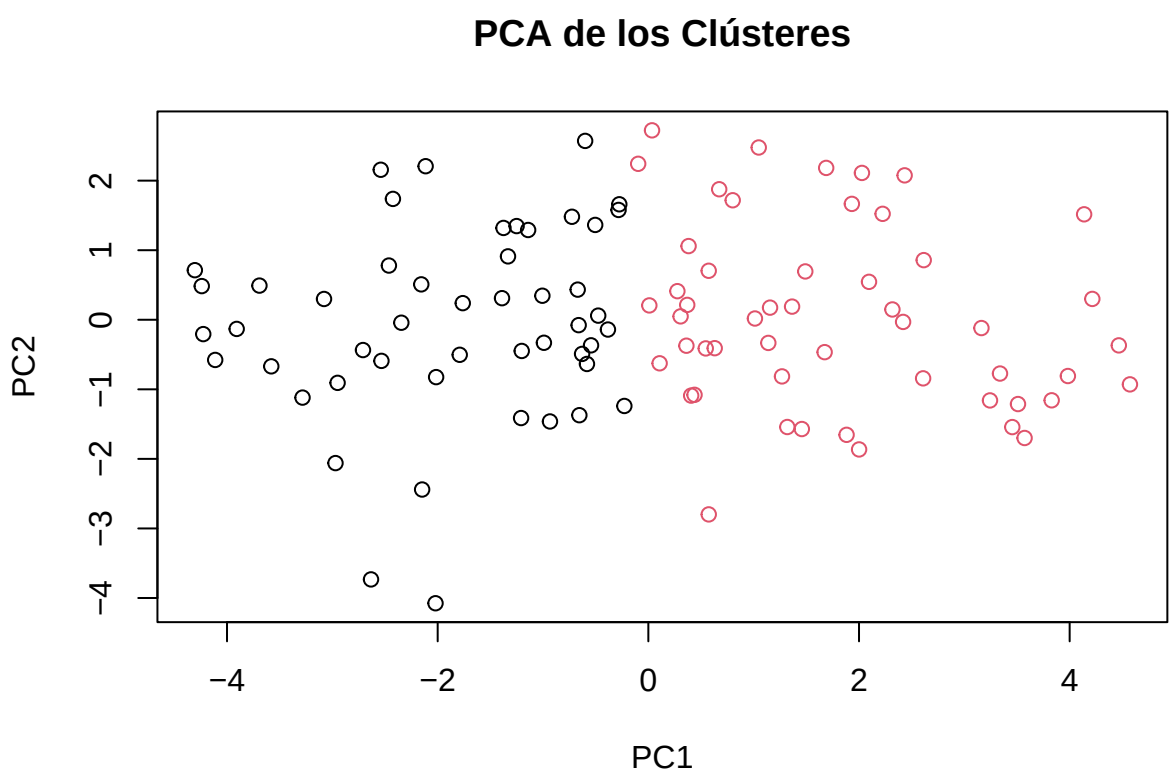
```r
library(cluster)
cophenetic_corr <- cor(distancias, cophenetic(hclust_model))
print(paste("Correlación cofenética:", cophenetic_corr))
```

```
## [1] "Correlación cofenética: 0.770110330494069"
```

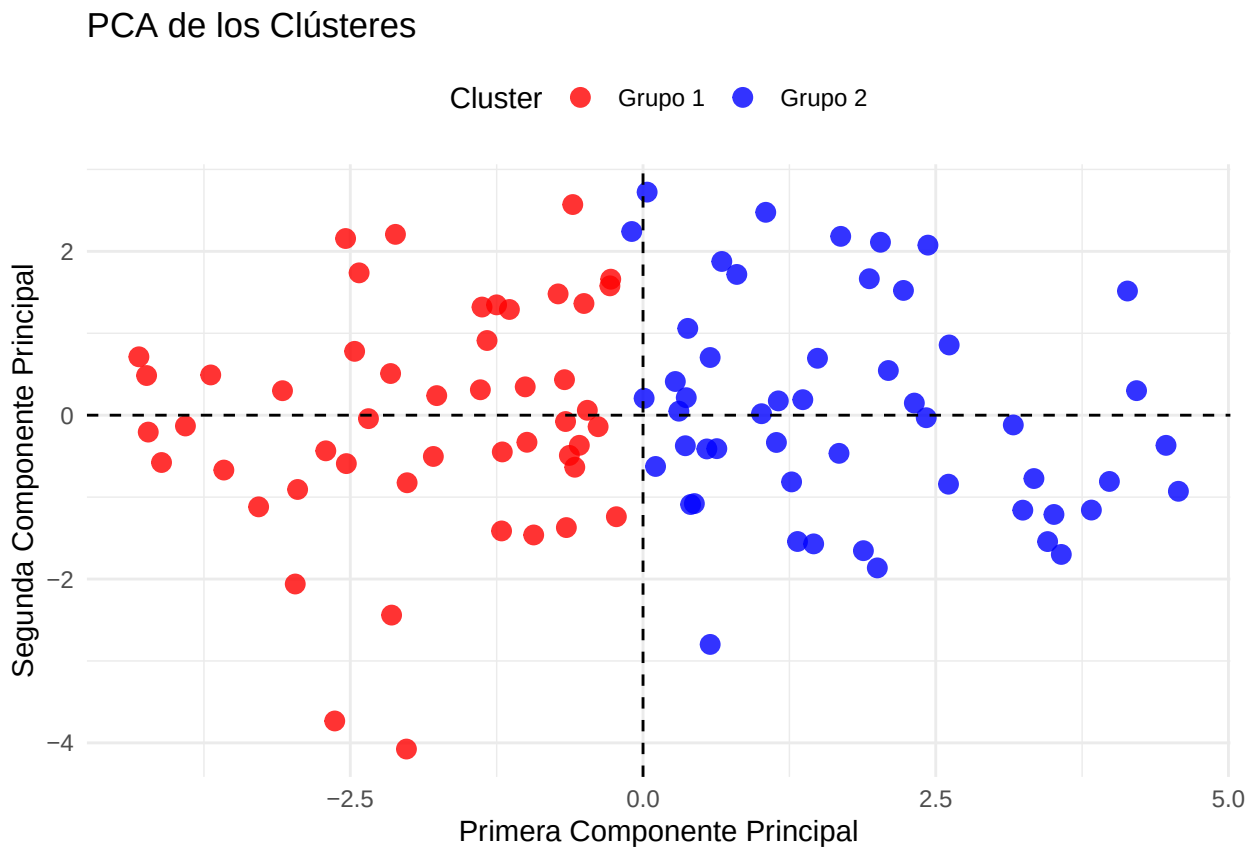
La correlación cofenética mide qué tan bien el dendrograma refleja las distancias originales entre las observaciones. Un valor de 0.7701 indica una correspondencia alta, lo que significa que el dendrograma representa de manera fiel las relaciones de proximidad en los datos. Mientras más cerca esté de 1, mejor es la representación.

Visualización Adicional con PCA

```r
pca_result <- prcomp(datos_estandarizados)
plot(pca_result$x[,1:2], col = kmeans_model$cluster, main = "PCA de los Clústeres", xlab = "PC1", ylab =
```




```r
pca_model <- prcomp(datos_estandarizados, scale. = TRUE)
pca_data <- as.data.frame(pca_model$x[, 1:2])
colnames(pca_data) <- c("PC1", "PC2")
pca_data$Cluster <- factor(kmeans_model$cluster, levels = c(1, 2), labels = c("Grupo 1", "Grupo 2"))
colores <- c("Grupo 1" = "red", "Grupo 2" = "blue")
ggplot(pca_data, aes(x = PC1, y = PC2, color = Cluster)) +
  geom_point(size = 3, alpha = 0.8) +
  geom_vline(xintercept = 0, linetype = "dashed", color = "black") +
  geom_hline(yintercept = 0, linetype = "dashed", color = "black") +
  scale_color_manual(values = colores) +
  labs(title = "PCA de los Clústeres",
       x = "Primera Componente Principal",
       y = "Segunda Componente Principal") +
  theme_minimal() +
  theme(legend.position = "top")
```



Interpretación del PCA de los Clústeres

El gráfico muestra los resultados del análisis de componentes principales (PCA) aplicado a los datos, visualizando cómo las observaciones se agrupan en **dos clústeres** generados mediante un modelo de clustering (K-means).

Descripción del Gráfico

1. **Ejes del PCA**:
   - El eje horizontal (Primera Componente Principal) explica la mayor parte de la variación en los datos.
   - El eje vertical (Segunda Componente Principal) captura una cantidad adicional de la variación, aunque menor que la primera componente.
2. **Distribución de los Clústeres**:
   - **Clúster 1 (rojo)**: Observaciones mayormente distribuidas en la parte derecha del gráfico (valores positivos de PC1). Este grupo representa un conjunto de individuos con características similares que los diferencian claramente del Clúster 2.
   - **Clúster 2 (azul)**: Observaciones localizadas principalmente en la parte izquierda del gráfico (valores negativos de PC1), indicando otro grupo homogéneo pero diferenciado del primero.
3. **Separación de Clústeres**:
   - Existe una separación visible entre los dos clústeres a lo largo del eje PC1, lo que sugiere que la Primera Componente Principal es clave para diferenciar estos grupos.
4. **Líneas de Referencia**:

- Las líneas punteadas (vertical y horizontal) marcan el origen del espacio PCA, ayudando a interpretar las posiciones relativas de los puntos respecto a las componentes principales.

Conclusión

El análisis de clustering realizado permitió identificar dos clústeres bien diferenciados en el conjunto de datos proporcionado. A continuación, se destacan los puntos más relevantes:

1. **Selección de Métrica y Método**:
   - Se probaron distintas métricas de distancia y métodos jerárquicos para determinar la mejor combinación utilizando la correlación cofenética como criterio de evaluación.
   - La combinación óptima fue la métrica de distancia **Canberra** junto con el método jerárquico **Average**, alcanzando una correlación cofenética de **0.7701**, lo que indica que el dendrograma refleja de manera fiel las distancias originales entre las observaciones.
2. **Resultados del Clustering**:
   - Los métodos jerárquicos y de partición (K-means) identificaron consistentemente **dos clústeres** principales en los datos.
   - El análisis de componentes principales (PCA) confirmó una clara separación entre estos dos grupos, mostrando diferencias significativas en las dimensiones clave.
3. **Características de los Clústeres**:
   - **Clúster 1**: Observaciones con menores valores promedio en las variables relacionadas con el rendimiento académico y la condición física.
   - **Clúster 2**: Observaciones con mayores valores promedio en las mismas variables, indicando un mejor desempeño general.
4. **Robustez del Análisis**:
   - La validación cruzada utilizando diferentes métodos y métricas confirmó la estabilidad y robustez de los resultados.
   - La correlación cofenética y los gráficos de PCA respaldan la calidad del agrupamiento.
5. **Impacto del Preprocesamiento**:
   - La estandarización de las variables numéricas fue esencial para asegurar que todas las dimensiones contribuyeran de manera equitativa al análisis.
   - La exclusión de variables no informativas, como `ID` y `Centro`, permitió mejorar la interpretabilidad de los resultados.