

# Análisis Multivariado

## Taller No. 7: Clasificación de razas de perros

Laura Elizabeth Holguín  
Eduardo Contreras Bohórquez  
Diana Catalina Martínez

1. Numere los nodos del árbol con los números de la descripción de los nodos (histograma de índices).

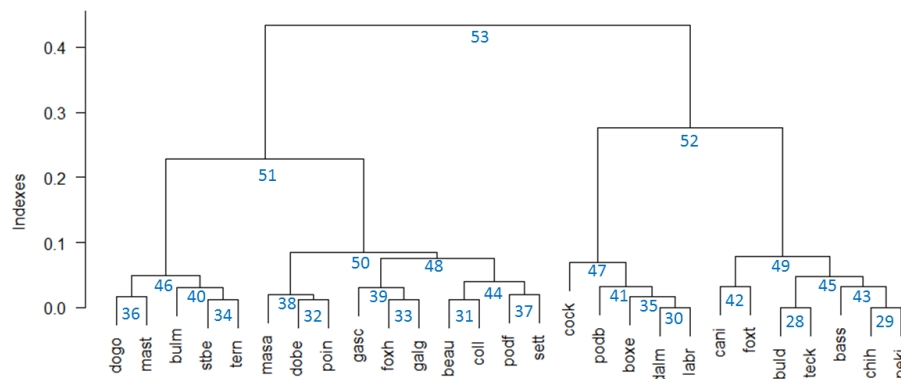


Figura 1: Árbol de clasificación por el método de Ward de las razas de perros según las coordenadas de todos los ejes factoriales

Los nodos correspondientes a la figura 1 corresponden al cuadro 1.

2. Describa las tres primeras uniones en la clasificación jerárquica.

Los primeros tres grupos presentan distancia 0 entre sus individuos; corresponden a los duplos de las razas Chihuahua-Pequinés, Buld-Tec y Dalmata-Labrador. Entre los 6 individuos comparten tener afectividad alta, agresividad baja y velocidad baja; y a nivel grupal son separados por peso, talla e inteligencia, de modo que Chihuahua-Pequinés y Buld-Teck son perros de tamaño pequeño, que difieren en su inteligencia baja y media respectivamente, mientras que los Labrador-Dálmata son de tamaño mediano y con inteligencia media.

3. Justifique la selección de 4 clases o cambie la decisión.

De acuerdo al histograma de índices de nivel para el ACM de las razas de perros (figura 2), se puede observar que la mejor altura para el corte del árbol es cuando el valor del índice es 0.1. Es decir, el mayor

	Nodo	Prim	Benj	Indice
1	28	4	26	0.0000
2	29	7	20	0.0000
3	30	10	17	0.0000
4	31	9	24	0.0124
5	32	11	18	0.0124
6	33	25	27	0.0124
7	34	13	15	0.0124
8	35	3	30	0.0165
9	36	12	19	0.0176
10	37	2	31	0.0206
11	38	23	32	0.0206
12	39	16	22	0.0292
13	40	5	33	0.0312
14	41	21	35	0.0325
15	42	6	14	0.0327
16	43	1	29	0.0330
17	44	28	43	0.0470
18	45	34	39	0.0478
19	46	36	40	0.0494
20	47	37	38	0.0691
21	48	8	41	0.0693
22	49	42	44	0.0790
23	50	45	47	0.0852
24	51	46	50	0.2278
25	52	48	49	0.2757
26	53	51	52	0.4331

Cuadro 1: Nodos e índices correspondientes a la clasificación jerárquica

salto en el histograma de los índices de nivel se observa cuando la inercia-entre está alrededor de 0.1, a esta altura del árbol (figura 1) se pueden definir claramente 4 clases para agrupar las razas de perros, la cual sería la partición de mejor calidad. Lo anterior indica que en la agrupación anterior las clases estaban próximas, y luego de esta partición los grupos se encuentran necesariamente alejados (se incrementa la inercia-entre), y se pueden por tanto, diferenciar claramente las agrupaciones para las razas de perros.

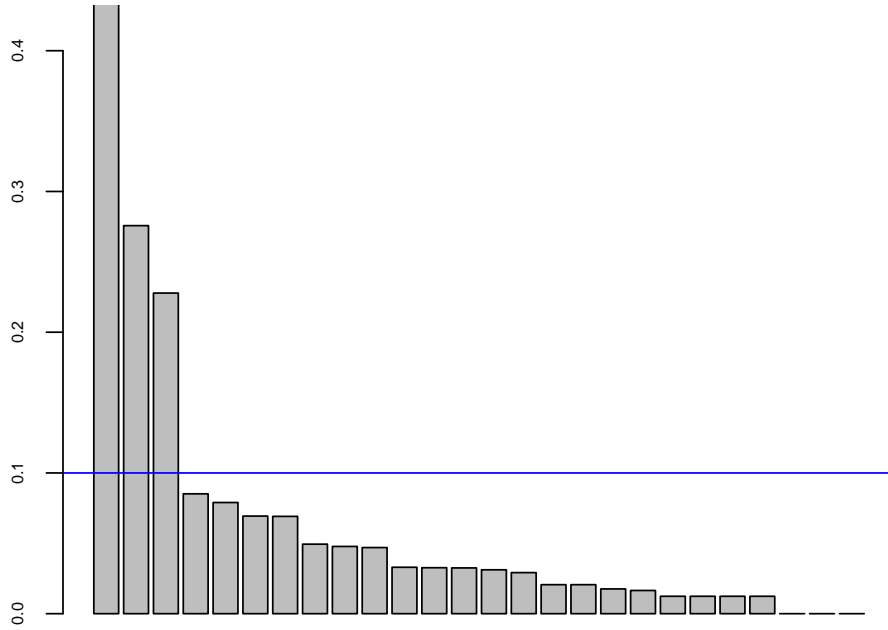


Figura 2: Histograma de índices

4. Para la partición en cuatro clases deduzca la inercia entre clases a partir de los Índices de nivel.

$$\begin{aligned} \text{Inercia entre} &= \text{Inercia total} - \text{inercia intra} \\ \text{Inercia entre} &= 1.6667 - 0.073 \\ \text{Inercia entre} &= 0.9366 \end{aligned}$$

5. A partir del árbol determine las razas de cada clase.

Las clases de acuerdo al dendograma (figura 1) serían:

Clase	Razas
Clase 1	Dogo, Mast, Bulm, Stbe, Tern
Clase 2	Foxh, Galg, Gasc, Podf, Beau, Coll, Sett, Poin, Dobe, Masa
Clase 3	Cock, Podb, Boxe, Dalm, Labr
Clase 4	Cani, Foxt, Buld, Teck, Bass, Chih, Peki

6. ¿Qué porcentaje de inercia explica la clasificación?

El porcentaje de inercia que explica la clasificación en 4 clases sería:

$$\frac{\text{Inercia-entre}}{\text{Inercia total}} \cdot 100 = \frac{0.9376}{1.6667} \cdot 100 = 56,25487$$

El porcentaje de inercia total que explica la partición en 4 clases corresponde al 56.3 % de la inercia total.

## 7. Resume las características de cada una de las clases.

	clasif	Test.Value	p.Value	Class.Cat	Cat.Class	Global	Weight
SIZE.sma	1	4.87	0.00	100.00	100.00	25.90	7.00
WEIG.lig	1	4.44	0.00	87.50	100.00	29.60	8.00
SPEE.low	1	2.03	0.04	50.00	71.40	37.00	10.00
SPEE.hig	1	-2.10	0.04	0.00	0.00	33.30	9.00
WEIG.med	1	-3.10	0.00	0.00	0.00	51.90	14.00
SIZE.lar	1	-3.32	0.00	0.00	0.00	55.60	15.00
WEIG.med1	2	3.85	0.00	71.40	100.00	51.90	14.00
SPEE.hig1	2	3.77	0.00	88.90	80.00	33.30	9.00
SIZE.lar1	2	3.57	0.00	66.70	100.00	55.60	15.00
SIZE.sma1	2	-2.29	0.02	0.00	0.00	25.90	7.00
WEIG.lig1	2	-2.54	0.01	0.00	0.00	29.60	8.00
SPEE.low1	2	-3.05	0.00	0.00	0.00	37.00	10.00
SIZE.med	3	4.37	0.00	100.00	100.00	18.50	5.00
SPEE.med	3	2.38	0.02	50.00	80.00	29.60	8.00
AFFE.hig	3	2.25	0.02	35.70	100.00	51.90	14.00
AFFE.low	3	-2.25	0.02	0.00	0.00	48.10	13.00
SIZE.lar2	3	-2.58	0.01	0.00	0.00	55.60	15.00
WEIG.heal	4	4.37	0.00	100.00	100.00	18.50	5.00
AFFE.low1	4	2.41	0.02	38.50	100.00	48.10	13.00
SIZE.lar3	4	2.08	0.04	33.30	100.00	55.60	15.00
WEIG.med2	4	-2.41	0.02	0.00	0.00	51.90	14.00
AFFE.hig1	4	-2.41	0.02	0.00	0.00	51.90	14.00

Cuadro 2: Caracterización de las razas de perros por clase según variables cualitativas

En la figura 3 se observa la asociación entre cada una de las variables cualitativas y la partición de las razas en cuatro clases, es decir, los perfiles fila de las razas de perros para cada variable, según los grupos formados.

En el grupo 1 se encuentran las razas de tamaño pequeño y peso ligero, poco veloces, muy afectivos y poco agresivos, y de inteligencia mediana a baja.

El grupo 2 se encuentra conformado por las razas de mayor tamaño y peso mediano, de gran agilidad y velocidad, pero afectividad baja. Así mismo, se encuentran razas de inteligencia alta, media y baja, y de agresividad tanto alta como baja.

En el grupo 3 las razas son de tamaño mediano, peso mediano, velocidad media, inteligencia media, afectividad alta y agresividad mixta.

En el grupo 4 están las razas de tamaño grande, peso alto, poco veloces, inteligencia mixta, afectividad baja y alta agresividad.

Adicionalmente, de acuerdo al gráfico de asociación entre clases y variables cualitativas de las razas (Fig. 3) así como al ordenamiento de valores test de las categorías de dichas variables (Cuadro 2) se obtiene:

- La clase 1 contiene al 26 % de las razas, los perros son de tamaño pequeño y peso ligero, con tendencia a poca velocidad.
- La clase 2 representa al 37 % de las razas de perros, se encuentra conformado por organismos de mayor tamaño y peso mediano, de gran agilidad y velocidad.
- En la clase 3 se encuentran el 18 % de las razas, las cuales son de tamaño mediano, velocidad media y afectividad alta.
- La clase 4 contiene el restante 18 % de razas de perros los cuales se caracterizan por tener tamaño grande, peso alto y afectividad baja.

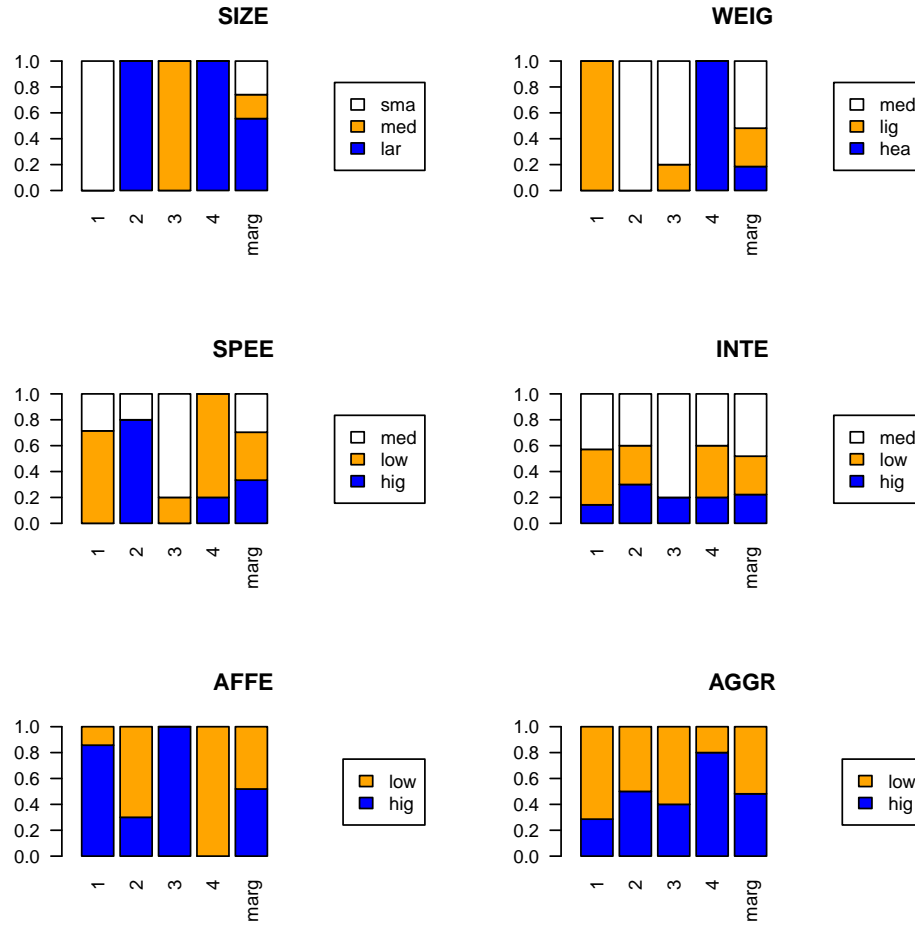


Figura 3: Perfiles fila para las razas de perros según sus características y los grupos formados

En general, la mayoría de las razas analizadas son de gran tamaño (55.6 %) y peso medio (51.9 %).

A continuación se presenta el índice de asociación entre la clasificación y las variables cualitativas:

	chi2	dfr	pval	tval	phi2
SIZE	54.00	6.00	0.00	6.05	2.00
WEIG	49.76	6.00	0.00	5.72	1.84
SPEE	25.70	6.00	0.00	3.48	0.95
AFPE	15.15	3.00	0.00	2.93	0.56

Las variables con mayor índice de asociación son tamaño y peso, indicando que son las mejores discriminadoras para clasificar las razas en los grupos. Aunque la inteligencia y la agresividad a menudo se pueden asociar más con unas razas de perros que con otras, según este índice no son determinantes de la agrupación propuesta.

8. Comente el primer plano factorial del ACM incluyendo las clases obtenidas (centros de gravedad y distinción de las razas de cada clase).

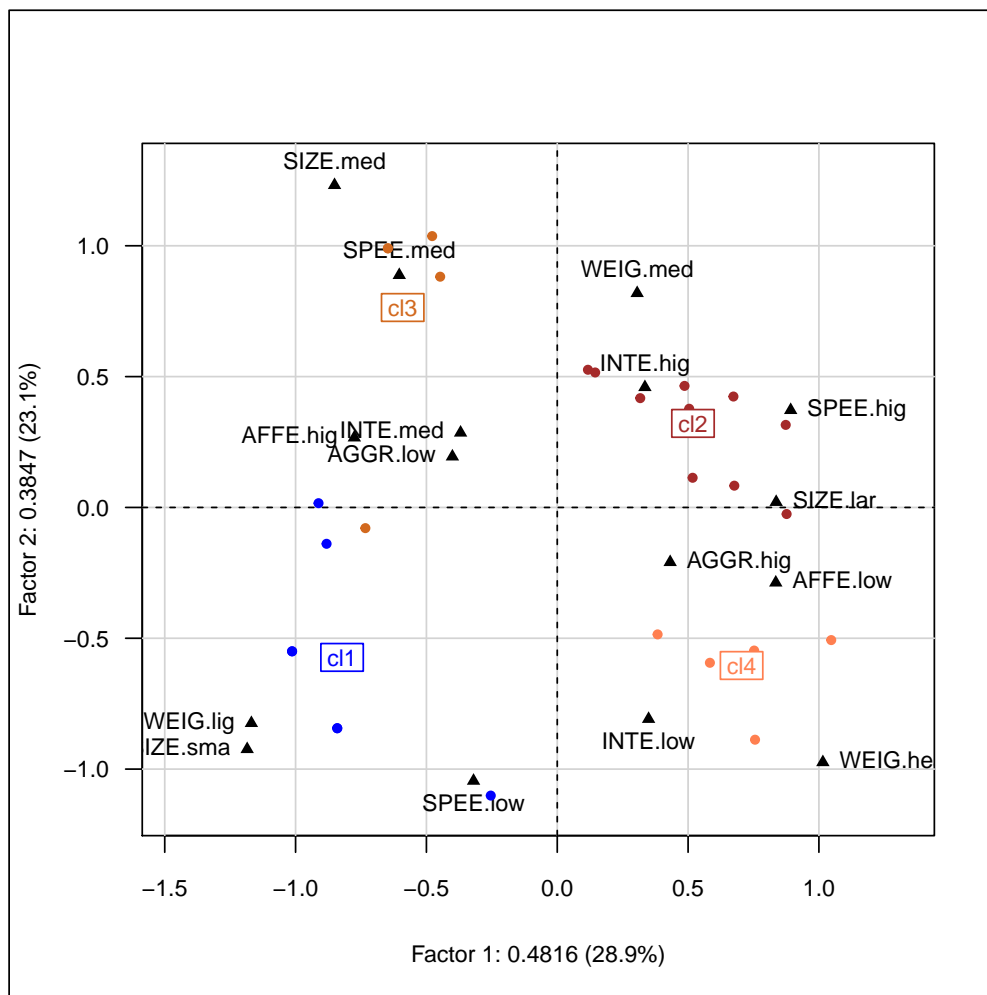


Figura 4: Proyección de clases sobre el primer plano factorial del ACM de las razas de perros

Como se observa en la figura correspondiente al primer plano factorial (figura 4), el centro de gravedad de la clase 2 es el más cercano al primer eje en el primer plano factorial, comparte con la clase 4 talla, velocidad y agresividad altas y afectividad baja, difieren en la inteligencia y el peso; las características que comparten son totalmente contrastantes a las de la clase 1. Los razas asociadas a las clases 2 y 4 se observan fuertemente atraídas por sus respectivos centros de gravedad, lo cual también se refleja en el valor de su inercia (0.085 y 0.079 respectivamente), mientras que las razas asociadas a las clases 1 y 2 se observan más dispersas de su centro de gravedad y presentan los valores más bajos de inercia (0.0494 y 0.0693 respectivamente).

9. Haga un resumen del análisis que responda a los objetivos del ejercicio.

De entrada el ACM en su histograma de inercias permite conocer que existen tres puntos de divergencia entre las razas, y posteriormente la clasificación obtenida presenta a los cuatro grupos que retienen al 83 %

de la variabilidad. Esto significa que las características evaluadas logran definir grupos de razas, principalmente con respecto a las variables de talla, peso y velocidad en una muy buena aproximación estadística. De acuerdo a estos rasgos, validados por tener una alta asociación con las razas de perros y su relación con la variable indicativa función, es posible seleccionar razas de perros de acuerdo con su función:

- De compañía: razas pequeñas y de baja velocidad.
- De utilidad: razas medianas, de gran peso y gran velocidad.
- De caza: razas grandes, de peso medio y gran velocidad.