

Ejercicio 1

Datos

Inspeccionamos con `glimpse(datos)`

```
## Rows: 18
## Columns: 14
## $ VAR      <chr> "BLANCO", "BULIDA", "CURROT.T", "CANINO", "CANINO.T", "CHIRIV~
## $ TAMFLOR   <dbl> 33.26, 26.84, 28.84, 27.52, 33.88, 26.32, 30.63, 27.60, 28.80~
## $ LONGPET   <dbl> 13.90, 13.44, 12.68, 12.32, 14.40, 13.36, 13.79, 13.68, 12.32~
## $ ANCHOPET  <dbl> 18.32, 15.88, 15.44, 14.28, 16.40, 15.04, 17.04, 17.40, 15.24~
## $ SUPHOJA   <dbl> 44.09, 36.67, 51.72, 36.04, 38.97, 41.28, 42.06, 53.33, 50.01~
## $ LONANCHO  <dbl> 1.053, 1.112, 0.912, 1.005, 1.058, 0.988, 0.991, 1.078, 0.880~
## $ PECLIMBO  <dbl> 0.415, 0.395, 0.389, 0.452, 0.311, 0.411, 0.385, 0.344, 0.412~
## $ PESOF     <dbl> 49.29, 49.00, 43.04, 66.79, 63.11, 43.80, 39.63, 44.98, 40.77~
## $ LONGF     <dbl> 46.34, 44.58, 40.06, 51.30, 50.70, 40.29, 39.71, 40.10, 42.54~
## $ ANCHOF    <dbl> 44.93, 45.09, 43.54, 50.80, 49.22, 42.51, 41.06, 42.25, 39.94~
## $ ESPESORF  <dbl> 45.40, 43.38, 42.97, 49.10, 46.82, 44.55, 40.79, 46.89, 41.77~
## $ PESOEND   <dbl> 4.31, 2.57, 1.62, 3.42, 3.34, 3.46, 1.88, 2.58, 1.52, 1.20, 1~
## $ LONGEND   <dbl> 31.07, 25.26, 20.13, 23.77, 29.09, 23.90, 19.66, 20.76, 17.97~
## $ ANCHOEND  <dbl> 24.18, 20.34, 17.62, 21.21, 22.86, 22.52, 20.31, 19.13, 17.57~
```

A. Cuáles son los valores de n y p ? Cuanto vale y que indica el valor x_{32} ? Y el vector \mathbf{x}_6 ?

El valor de n es 18 y el valor de p es 14. El valor de x_{32} es 28.84 e indica el tamaño de la flor de la observacion 3. El vector \mathbf{x}_6 indica la relacion entre el ancho y largo de la hoja.

B. Cómo clasificaría las variables sobre las que se está trabajando?

Las variables con las que se esta trabajando son de tipo continuo en todos los casos. Mas aun, todas estan medidas en escala de intervalo.

C. Encuentre el vector de medias y matriz de varianzas-covariancias asociados a la tabla de datos.

Las medias son:

La matriz de covarianza es:

D. Podría decir cuál y cuáles variables son las más dispersas?

Utilizando el coeficiente de variación, podemos decir que las variables más dispersas, en orden decreciente, son:

E. Estandarice las variables por media y desvío. Ahora puede responder al inciso (d)?

Como son variables estandarizadas, el coeficiente de variacion no existe, pero todas tienen desvio estandar 1 (ya que la media es 0)

```
## # A tibble: 1 x 13
##   TAMFLOR LONGPET ANCHOPET SUPHOJA LONANCHO PECLIMBO PESOF LONGF ANCHOF
```

TAMFLOR	LONGPET	ANCHOPET	SUPHOJA	LONANCHO	PECLIMBO	PESOF	LONGF	ANCHOF
28.9	12.9	15.5	41	1	0.4	46.7		

	TAMFLOR	LONGPET	ANCHOPET	SUPHOJA	LONANCHO	PECLIMBO	I
TAMFLOR	9.9	2.2	4.1	2.2	0.0	0	
LONGPET	2.2	1.0	1.6	1.0	0.0	0	
ANCHOPET	4.1	1.6	3.1	3.5	0.0	0	
SUPHOJA	2.2	1.0	3.5	32.6	-0.1	0	
LONANCHO	0.0	0.0	0.0	-0.1	0.0	0	
PECLIMBO	0.0	0.0	0.0	0.0	0.0	0	
PESOF	15.2	4.9	7.9	-14.3	0.4	0	
LONGF	10.0	2.9	4.9	-5.2	0.1	0	
ANCHOF	6.1	1.9	3.2	-5.8	0.2	0	
ESPESORF	4.3	1.8	3.2	-1.7	0.1	0	
PESOEND	1.7	0.7	1.1	-0.7	0.0	0	
LONGEND	9.1	3.3	5.3	-3.2	0.1	0	
ANCHOEND	3.7	1.8	2.7	-2.1	0.1	0	

VAR	CV
PESOEND	0.4
PESOF	0.3
LONGEND	0.2
LONGF	0.1
SUPHOJA	0.1
ANCHOEND	0.1
ANCHOPET	0.1
ANCHOF	0.1
TAMFLOR	0.1
ESPESORF	0.1
PECLIMBO	0.1
LONGPET	0.1
LONANCHO	0.1

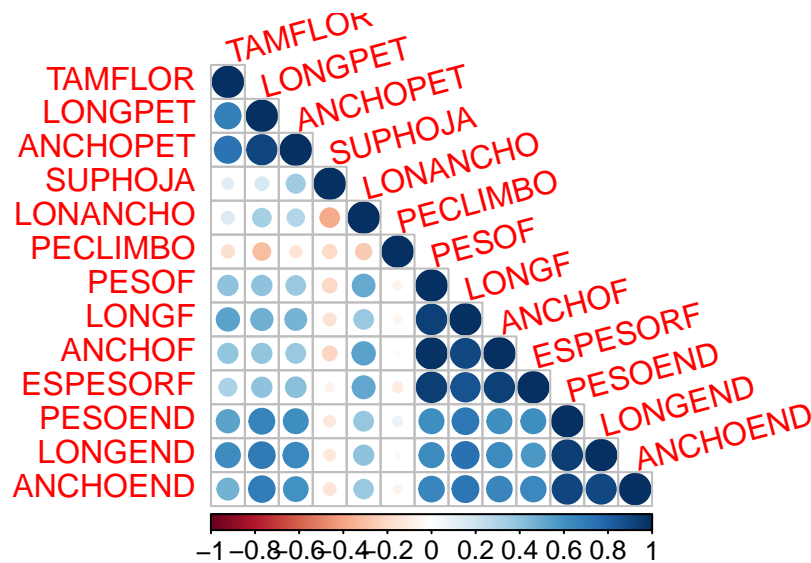
	TAMFLOR	LONGPET	ANCHOPET	SUPHOJA	LONANCHO	PECLIMBO	
TAMFLOR	1.0	0.7	0.7	0.1	0.2	-0.2	
LONGPET	0.7	1.0	0.9	0.2	0.3	-0.3	
ANCHOPET	0.7	0.9	1.0	0.4	0.3	-0.1	
SUPHOJA	0.1	0.2	0.4	1.0	-0.4	-0.2	
LONANCHO	0.2	0.3	0.3	-0.4	1.0	-0.3	
PECLIMBO	-0.2	-0.3	-0.1	-0.2	-0.3	1.0	
PESOF	0.4	0.4	0.4	-0.2	0.5	-0.1	
LONGF	0.5	0.5	0.5	-0.2	0.4	-0.1	
ANCHOF	0.4	0.4	0.4	-0.2	0.5	0.0	
ESPESORF	0.3	0.4	0.4	-0.1	0.5	-0.1	
PESOEND	0.5	0.7	0.6	-0.1	0.4	0.1	
LONGEND	0.6	0.7	0.6	-0.1	0.4	0.0	
ANCHOEND	0.5	0.7	0.6	-0.1	0.4	-0.1	

```
##      <dbl>      <dbl>      <dbl>      <dbl>      <dbl>      <dbl>      <dbl>      <dbl>      <dbl>
## 1 2.22e15 -1.30e15 -5.65e15 3.13e15 -3.61e15 1.41e15 5.53e16 4.25e15 -2.52e15
## # ... with 4 more variables: ESPESORF <dbl>, PESOEND <dbl>, LONGEND <dbl>,
## #   ANCHOEND <dbl>
```

F. Halle la matriz de correlación. Que variables son las más relacionadas?

La matriz de correlación es:

A continuación una representación gráfica de la matriz de correlación, la cual permite identificar de forma más sencilla las variables más relacionadas:



G. Pueden dividirse las variables en subgrupos, de modo que las variables dentro de un mismo subgrupo tengan elevadas correlaciones entre sí y que las que se encuentren en subgrupos diferentes tengan bajas correlaciones? Si es así, cuáles variables quedan en cada uno de los subgrupos?

- Los subgrupos de variables con altas correlaciones son los siguientes:
 1. Peso, longitud, ancho y espesor del fruto (características del fruto)
 2. Peso, longitud y ancho del endocarpio (características del endocarpio)

Variedad 1	Variedad 2	Distancia
PALAU	CORBATO	4.794812
MANRI	GINESTA	5.712183
CRISTALI	CURROT.T	5.778856

3. Tamaño de la flor, longitu y ancho del petalo (caracteristicas de la flor)

- El subgrupo de variables con bajas correlaciones son los siguientes:

1. Superficie de la hoja, relacion entre peciolo-limbo y relacion entre longitud y ancho de la hoja (caracteristicas de la hoja); tamaño de la flor, longitud y ancho del petalo (caracteristicas de la flor)

Cabe destacar que la superficie de la hoja y relacion peciolo-limbo presentan correlacion muy baja o nula con cualquiera de las otras variables.

H. Encuentre la matriz que mide el grado de similaridad entre las variedades en función de la distancia euclídea calculada sobre los datos originales.

	BLANCO	BULIDA	CURROT.T	CANINO	CANINO.T	CHIRIVEL	CORBATO	CRISTALI
BLANCO	0.0	12.7	18.5	23.6	16.4	14.0	18.4	17.9
BULIDA	12.7	0.0	18.0	20.9	18.5	9.1	14.5	18.9
CURROT.T	18.5	18.0	0.0	32.5	29.4	12.7	11.4	5.8
CANINO	23.6	20.9	32.5	0.0	10.5	27.8	33.3	31.7
CANINO.T	16.4	18.5	29.4	10.5	0.0	25.0	29.9	28.6
CHIRIVEL	14.0	9.1	12.7	27.8	25.0	0.0	9.1	13.5
CORBATO	18.4	14.5	11.4	33.3	29.9	9.1	0.0	14.4
CRISTALI	17.9	18.9	5.8	31.7	28.6	13.5	14.4	0.0
CURROT	20.4	18.6	5.8	34.3	31.4	13.2	9.7	9.3
GABACHET	41.9	37.4	30.9	57.0	54.1	31.6	25.1	34.5
GANDIA	24.5	16.7	15.1	35.1	33.9	11.9	9.8	18.5
GINESTA	23.6	13.9	19.3	32.0	31.4	11.4	12.1	21.8
MANRI	23.0	14.8	16.8	34.1	32.7	10.0	9.3	19.1
MARTINET	10.2	13.9	26.1	17.4	9.5	19.8	24.4	25.7
PALABRAS	21.3	11.2	18.4	27.4	27.1	11.7	12.0	20.4
PALAU	21.1	15.7	11.1	34.1	31.6	10.1	4.8	14.6
R.CARLET	13.7	9.1	11.6	25.6	22.5	6.9	8.8	13.0
TADEO	30.9	30.7	39.7	11.8	16.3	36.9	41.9	38.2

I. Podría decir cuales son los tres pares de variedades que presentan mayor semejanza?

Los pares de variedades que presentan mayor semejanza son:

J. Repita lo realizado en el inciso (h) pero sobre las variables estandarizadas por media y desvío estándar. Son las mismas las tres variedades más parecidas? Comente al respecto

Con los datos estandarizados, los pares de variedades mas parecidas son:

Podemos ver que los tres pares de variedades mas parecidas son distintos a los que vimos en el inciso anterior donde utilizamos los datos sin estandarizar. Esto sucede porque las variables estan medidas en diferentes unidades de medicion, y al utilizar las variables sin estandarizar se le da mayor peso a las que tienen una variabilidad mayor valor en la escala de medida original.

K. Mida el grado de concordancia entre ambas matrices de distancia.

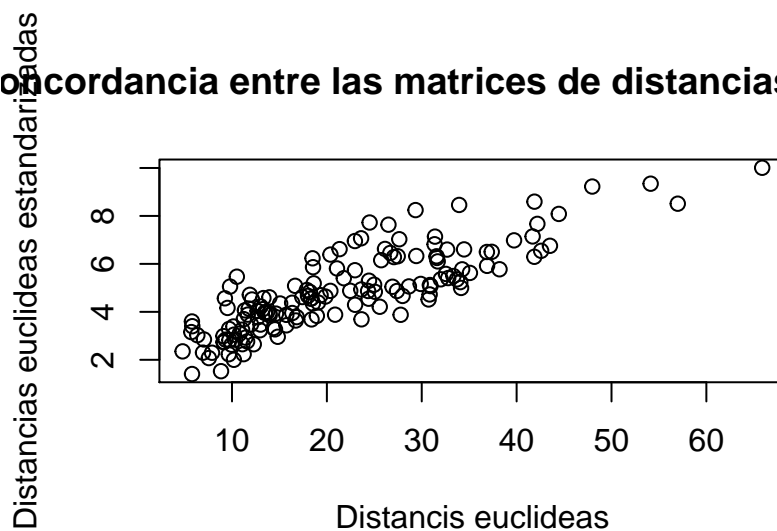
El grado de concordancia entre las matrices de distancias es: 0.85.

A continuacion se realiza una representacion grafica:

	BLANCO	BULIDA	CURROT.T	CANINO	CANINO.T	CHIRIVEL	CORBATO	CRISTALI
BLANCO	0.0	4.0	5.9	4.9	4.0	3.9	4.6	4.9
BULIDA	4.0	0.0	4.7	3.9	4.3	2.7	3.3	3.8
CURROT.T	5.9	4.7	0.0	5.6	6.3	3.8	2.9	3.6
CANINO	4.9	3.9	5.6	0.0	5.5	3.9	5.5	6.1
CANINO.T	4.0	4.3	6.3	5.5	0.0	5.1	5.2	5.1
CHIRIVEL	3.9	2.7	3.8	3.9	5.1	0.0	3.0	4.0
CORBATO	4.6	3.3	2.9	5.5	5.2	3.0	0.0	3.3
CRISTALI	4.9	3.8	3.6	6.1	5.1	4.0	3.3	0.0
CURROT	6.4	5.2	1.4	5.8	7.1	4.1	3.3	4.6
GABACHET	8.6	6.5	5.1	8.5	9.3	6.2	4.8	6.6
GANDIA	7.7	5.1	4.3	5.6	8.5	4.7	5.1	6.2
GINESTA	7.1	3.9	4.7	5.4	6.8	4.1	4.5	5.4
MANRI	5.7	3.0	3.8	5.0	6.6	2.6	2.8	4.4
MARTINET	2.0	4.6	6.6	4.6	4.2	4.6	5.3	6.2
PALABRAS	6.6	3.7	3.7	4.9	6.3	3.9	3.5	4.9
PALAU	5.8	3.4	2.7	5.2	6.3	3.4	2.4	3.9
R.CARLET	4.0	2.7	2.8	4.2	4.9	2.3	1.5	3.5
TADEO	5.1	4.5	7.0	4.1	4.4	5.9	6.3	5.8

Variedad 1	Variedad 2	Distancia
CURROT	CURROT.T	1.407096
R.CARLET	CORBATO	1.524706
MARTINET	BLANCO	1.999641

Concordancia entre las matrices de distancias euclídeas



L. Realice un Análisis de Componentes Principales utilizando de la matriz de correlaciones.

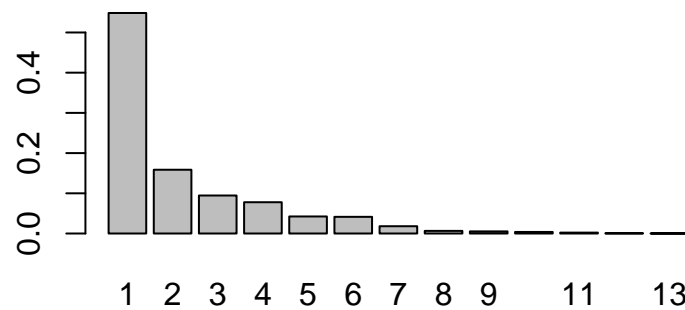
Los autovalores obtenidos son los siguientes:

También pueden observarse gráficamente:

	Autovalores	% Varianza	% Varianza acumulada
comp 1	7.1299316	54.8456276	54.84563
comp 2	2.0614209	15.8570840	70.70271
comp 3	1.2284033	9.4492561	80.15197
comp 4	1.0113575	7.7796727	87.93164
comp 5	0.5510668	4.2389756	92.17062
comp 6	0.5390785	4.1467574	96.31737
comp 7	0.2349837	1.8075671	98.12494
comp 8	0.0841759	0.6475072	98.77245
comp 9	0.0684849	0.5268066	99.29925
comp 10	0.0501440	0.3857230	99.68498
comp 11	0.0286852	0.2206554	99.90563
comp 12	0.0093799	0.0721532	99.97779
comp 13	0.0028878	0.0222141	100.00000

	Dim.1	Dim.2
TAMFLOR	0.6512996	0.4741214
LONGPET	0.7486536	0.5549196
ANCHOPET	0.7128111	0.6131597
SUPHOJA	-0.1005814	0.6917032
LONANCHO	0.5411750	-0.3094653
PECLIMBO	-0.1276622	-0.2842706
PESOF	0.8613521	-0.3914137
LONGF	0.8960725	-0.2319444
ANCHOF	0.8585841	-0.3992579
ESPESORF	0.8342138	-0.3324540
PESOEND	0.8731975	0.0621437
LONGEND	0.8946818	0.1208855
ANCHOEND	0.8814714	0.0478205

AUTOVALORES



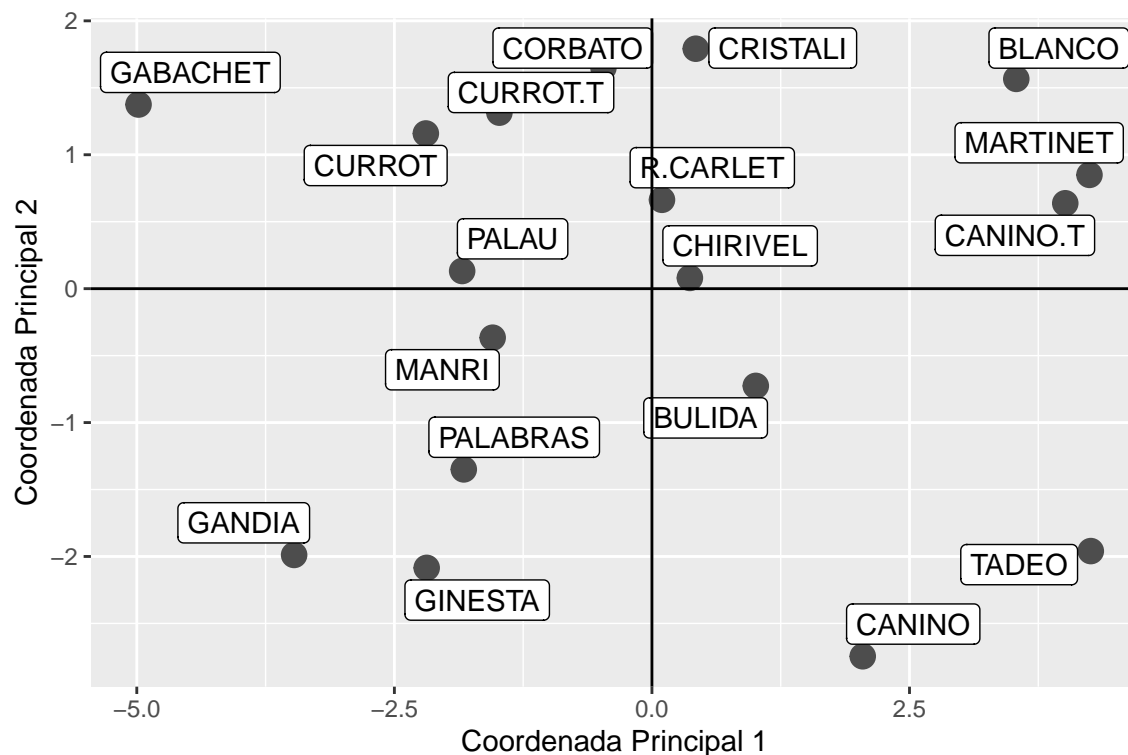
Y los autovectores (cargas asociadas a cada componente) son:

M. Analice los porcentaje de variabilidad explicada por los primeros ejes principales.

Observando los autovalores y el porcentaje de varianza explicada de cada uno, podemos decir que la primer componente explica el 54.85%, mientras que la segunda componente explica 15.86%. Luego, la varianza total explicada por estas dos componentes es 70.7%.

N. Establezca intuitivamente grupos de variedades similares según su cercanía en el plano principal.

A continuación se presenta la representación gráfica de las variedades en el plano principal:



Tras observar el gráfico podemos decir que encontramos cuatro grupos. El primer grupo contiene a MARTINET, CANINO.T y BLANCO. El segundo está conformado por CANINO y TADEO. El tercer grupo está conformado solamente por GABACHET, que se diferencia de todas las variedades. Y el cuarto grupo que se compone por el resto de las variedades, que están ubicadas alrededor del comportamiento promedio, es decir, el origen del plano.

O. Encuentre e interprete gradientes de las variables originales en el plano principal en función de sus cargas sobre las dos primeras componentes.

Observando nuevamente las coordenadas de las variables (autovectores):

Considerando la primera componente, los damascos que tengan flores, frutos y endocarpio grandes se van a ubicar a la derecha del gráfico.

Con respecto a la segunda componente, los damascos cuyas hojas sean grandes estarán ubicados en la parte superior del gráfico, lo mismo ocurre con aquellos damascos con flores grandes. Los damascos que tengan frutos pequeños, estarán ubicados en la parte inferior del gráfico.

Luego:

- Damascos con hojas grandes estarán ubicados en el segundo cuadrante del gráfico.
- Damascos con endocarpio y frutos grandes estarán ubicados en el cuarto cuadrante del gráfico.
- Damascos con flores grandes estarán ubicados en el primer cuadrante del gráfico.

P. Caracterice los grupos determinados en el inciso (N) según los gradientes descriptos en (O).

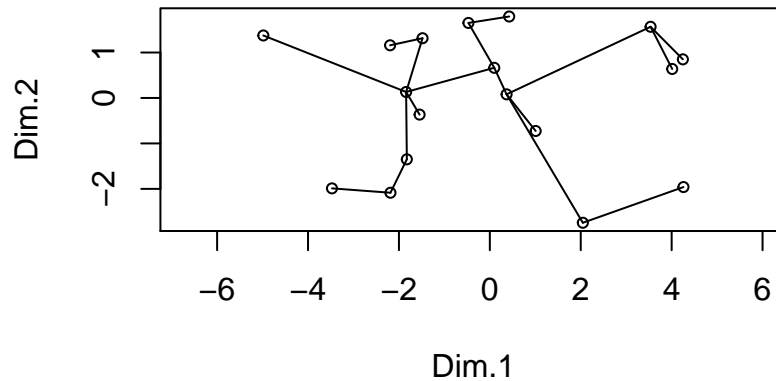
	Dim.1	Dim.2
TAMFLOR	0.6512996	0.4741214
LONGPET	0.7486536	0.5549196
ANCHOPET	0.7128111	0.6131597
SUPHOJA	-0.1005814	0.6917032
LONANCHO	0.5411750	-0.3094653
PECLIMBO	-0.1276622	-0.2842706
PESOF	0.8613521	-0.3914137
LONGF	0.8960725	-0.2319444
ANCHOF	0.8585841	-0.3992579
ESPESORF	0.8342138	-0.3324540
PESOEND	0.8731975	0.0621437
LONGEND	0.8946818	0.1208855
ANCHOEND	0.8814714	0.0478205

Grupo 1: es un grupo de variedades caracterizadas por tener flores, endocarpio y hojas grandes y frutos medianos.

Grupo 2: es un grupo caracterizado por tener endocarpio grande, frutos grandes, hojas pequeñas y flores medianas

Grupo 3: es un fruto caracterizado por tener hojas grandes, frutos y endocarpio pequeños y flores medianas.

Q. Superponga en la representación del plano principal un MST. Comente al respecto, haría algún reagrupamiento ?



R. Con el software R realice el ACP recurriendo a operaciones con matrices (decomposición espectral)

Realizando la descomposicion espectral, los autovalores son los siguientes:

Y los autovectores:

Las nuevas coordenadas de las variedades sobre las componentes halladas:

S. Verifique que con el enfoque Biplot (DVS) llega a los mismos resultados

Valores singulares (d):

Vectores singulares por izquierda (u):

AUTOVALORES
7.1
2.1
1.2
1.0
0.6
0.5
0.2
0.1
0.1
0.1
0.0
0.0
0.0

AUTOVECTOR 1	AUTOVECTOR 2
-0.2	-0.3
-0.3	-0.4
-0.3	-0.4
0.0	-0.5
-0.2	0.2
0.0	0.2
-0.3	0.3
-0.3	0.2
-0.3	0.3
-0.3	0.2
-0.3	0.0
-0.3	-0.1
-0.3	0.0

	TAMFLOR	LONGPET	ANCHOPET	SUPHOJA	LONANCHO	PECLIMBO	PESOF	LONC
BLANCO	-3.4	-1.5	-1.3	0.5	-0.5	0.0	-0.8	-0.1
BULIDA	-1.0	0.7	0.3	1.3	-0.5	-0.4	0.1	0.0
CURROT.T	1.4	-1.3	0.4	-2.0	0.0	0.4	-0.1	0.0
CANINO	-2.0	2.7	-1.3	-1.2	0.0	0.2	0.1	0.0
CANINO.T	-3.9	-0.6	1.7	0.7	0.9	1.6	-0.1	0.0
CHIRIVEL	-0.4	-0.1	-1.0	0.3	-1.3	0.8	0.3	-0.1
CORBATO	0.5	-1.6	0.2	0.4	0.4	-0.1	0.8	-0.1
CRISTALI	-0.4	-1.7	2.1	-0.4	-1.5	-0.5	-0.4	-0.1
CURROT	2.1	-1.1	-0.4	-2.1	0.3	0.5	0.1	0.0
GABACHET	4.8	-1.3	-0.3	1.3	1.1	-0.8	-0.4	0.0
GANDIA	3.4	1.9	-0.8	-0.4	-0.2	-0.2	-0.8	0.0
GINESTA	2.1	2.0	0.8	1.1	-0.2	1.0	-0.5	-0.1
MANRI	1.5	0.4	-0.5	0.9	-0.9	-0.4	1.0	0.0
MARTINET	-4.1	-0.8	-2.0	0.4	0.7	-0.2	-0.1	0.0
PALABRAS	1.8	1.3	0.8	0.1	0.8	0.6	0.5	-0.1
PALAU	1.8	-0.1	0.3	0.1	0.4	-0.5	-0.2	-0.1
R.CARLET	-0.1	-0.6	-0.3	-0.2	0.3	-0.3	0.4	-0.1
TADEO	-4.1	1.9	1.5	-0.7	0.5	-1.6	0.0	-0.1

d
11.0
5.9
4.6
4.1
3.1
3.0
2.0
1.2
1.1
0.9
0.7
0.4
0.2

TAMFLOR	LONGPET	ANCHOPET	SUPHOJA	LONANCHO	PECLIMBO	PESOF	LONGF	ANCHOF
-0.3	0.3	0.3	0.1	0.1	0.0	0.4	0.1	0.0
-0.1	-0.1	-0.1	0.3	0.2	0.1	0.0	-0.7	0.0
0.1	0.2	-0.1	-0.5	0.0	-0.1	0.0	-0.2	0.3
-0.2	-0.5	0.3	-0.3	0.0	-0.1	-0.1	-0.1	-0.1
-0.4	0.1	-0.4	0.2	-0.3	-0.5	0.1	-0.2	0.2
0.0	0.0	0.2	0.1	0.4	-0.3	-0.1	0.3	0.0
0.0	0.3	0.0	0.1	-0.1	0.0	-0.4	0.1	0.2
0.0	0.3	-0.5	-0.1	0.5	0.2	0.2	0.1	-0.4
0.2	0.2	0.1	-0.5	-0.1	-0.2	-0.1	-0.2	-0.2
0.4	0.2	0.1	0.3	-0.3	0.3	0.2	-0.1	-0.2
0.3	-0.3	0.2	-0.1	0.1	0.1	0.4	-0.1	0.2
0.2	-0.3	-0.2	0.3	0.1	-0.3	0.2	0.3	0.0
0.1	-0.1	0.1	0.2	0.3	0.1	-0.5	-0.1	0.1
-0.4	0.1	0.4	0.1	-0.2	0.1	0.0	0.0	-0.3
0.2	-0.2	-0.2	0.0	-0.3	-0.2	-0.2	0.1	-0.5
0.2	0.0	-0.1	0.0	-0.1	0.2	0.1	0.1	0.2
0.0	0.1	0.1	0.0	-0.1	0.1	-0.2	0.3	0.4
-0.4	-0.3	-0.3	-0.2	-0.2	0.5	0.0	0.2	0.1

TAMFLOR	LONGPET	ANCHOPET	SUPHOJA	LONANCHO	PECLIMBO	PESOF	LONGF	ANCHOF
-0.2	0.3	0.1	0.0	-0.7	0.1	0.3	0.3	0.1
-0.3	0.4	0.0	0.2	0.0	0.1	-0.5	-0.2	0.0
-0.3	0.4	0.0	0.0	0.0	0.4	-0.3	0.0	-0.2
0.0	0.5	-0.2	-0.6	0.4	0.1	0.4	-0.1	0.1
-0.2	-0.2	-0.4	0.5	0.2	0.6	0.4	-0.1	0.0
0.0	-0.2	0.8	-0.2	0.0	0.5	0.0	0.0	0.1
-0.3	-0.3	-0.1	-0.3	-0.1	0.0	-0.1	-0.1	0.1
-0.3	-0.2	0.0	-0.3	-0.2	-0.2	0.0	-0.5	-0.4
-0.3	-0.3	-0.1	-0.2	-0.1	0.1	0.0	0.0	0.5
-0.3	-0.2	-0.2	-0.3	0.2	0.1	-0.1	0.5	-0.3
-0.3	0.0	0.3	0.1	0.2	-0.2	0.3	0.2	-0.4
-0.3	0.1	0.2	0.2	0.1	-0.2	0.3	-0.4	0.2
-0.3	0.0	0.2	0.2	0.3	-0.3	-0.1	0.4	0.4

	TAMFLOR	LONGPET	ANCHOPET	SUPHOJA	LONANCHO	PECLIMBO	PESOF	LONC
BLANCO	-3.4	1.5	1.3	0.5	0.5	0.0	0.8	0
BULIDA	-1.0	-0.7	-0.3	1.3	0.5	0.4	-0.1	-1
CURROT.T	1.4	1.3	-0.4	-2.0	0.0	-0.4	0.1	-1
CANINO	-2.0	-2.7	1.3	-1.2	0.0	-0.2	-0.1	-1
CANINO.T	-3.9	0.6	-1.7	0.7	-0.9	-1.6	0.1	-1
CHIRIVEL	-0.4	0.1	1.0	0.3	1.3	-0.8	-0.3	0
CORBATO	0.5	1.6	-0.2	0.4	-0.4	0.1	-0.8	0
CRISTALI	-0.4	1.7	-2.1	-0.4	1.5	0.5	0.4	0
CURROT	2.1	1.1	0.4	-2.1	-0.3	-0.5	-0.1	-1
GABACHET	4.8	1.3	0.3	1.3	-1.1	0.8	0.4	-1
GANDIA	3.4	-1.9	0.8	-0.4	0.2	0.2	0.8	-1
GINESTA	2.1	-2.0	-0.8	1.1	0.2	-1.0	0.5	0
MANRI	1.5	-0.4	0.5	0.9	0.9	0.4	-1.0	-1
MARTINET	-4.1	0.8	2.0	0.4	-0.7	0.2	0.1	0
PALABRAS	1.8	-1.3	-0.8	0.1	-0.8	-0.6	-0.5	0
PALAU	1.8	0.1	-0.3	0.1	-0.4	0.5	0.2	0
R.CARLET	-0.1	0.6	0.3	-0.2	-0.3	0.3	-0.4	0
TADEO	-4.1	-1.9	-1.5	-0.7	-0.5	1.6	0.0	0

Vectores singulares por derecha (v):

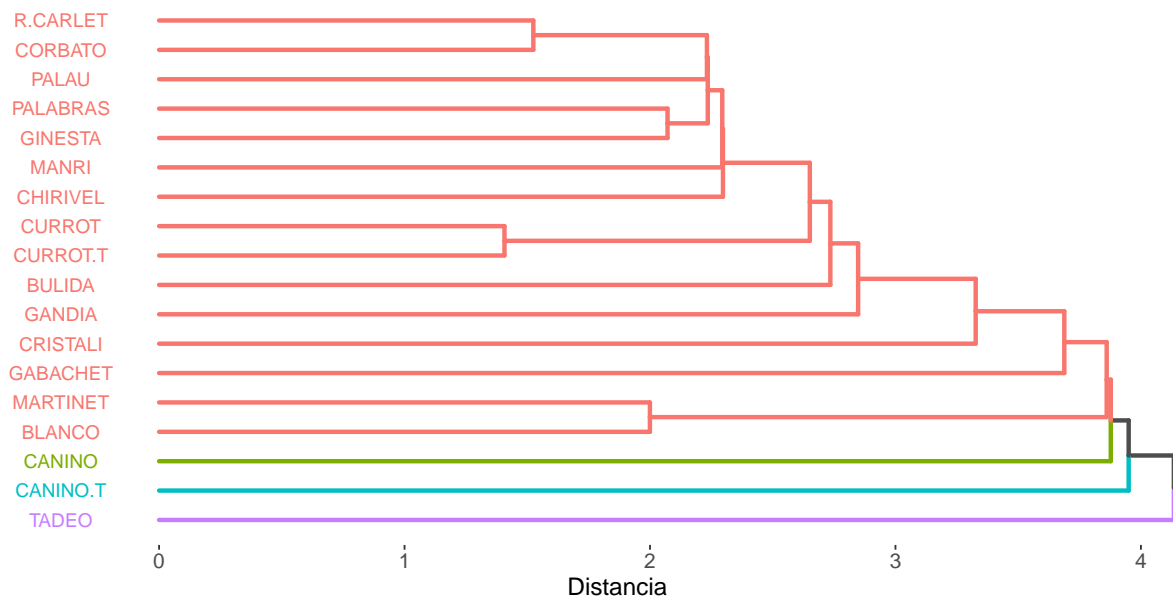
Las nuevas cordenadas de las variedades halladas a traves de DVS:

FALTA VERIFICAR QUE SE LLEGA AL MISMO RESULTADO

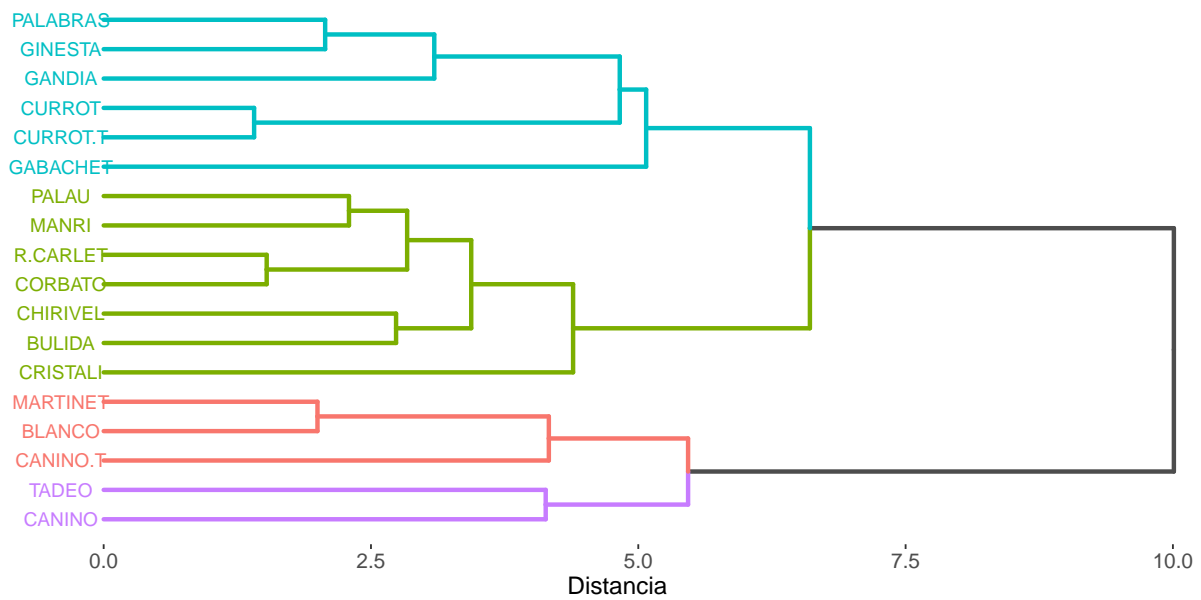
Graficos de variedades en plano de 2 componentes encimados. Como hago para encimar en ggplot2??

T. Obtenga 4 dendogramas ultramétricos según diferentes criterios de encadenamiento (SIMPLE, COM-PUESTO, UPGMA y WARD).

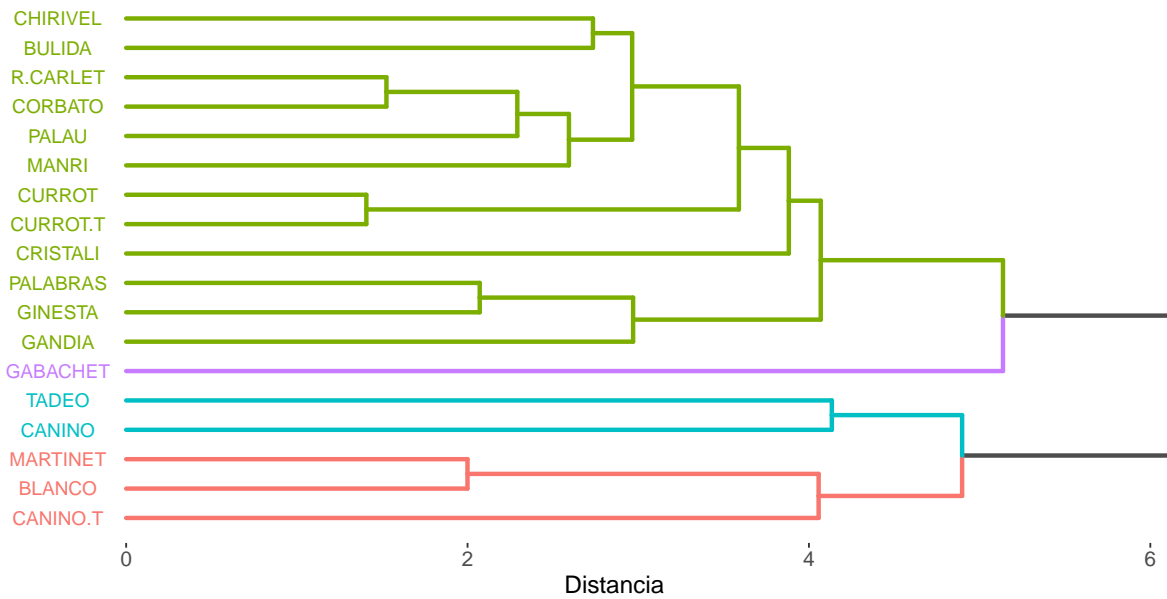
Criterio de encadenamiento Simple:



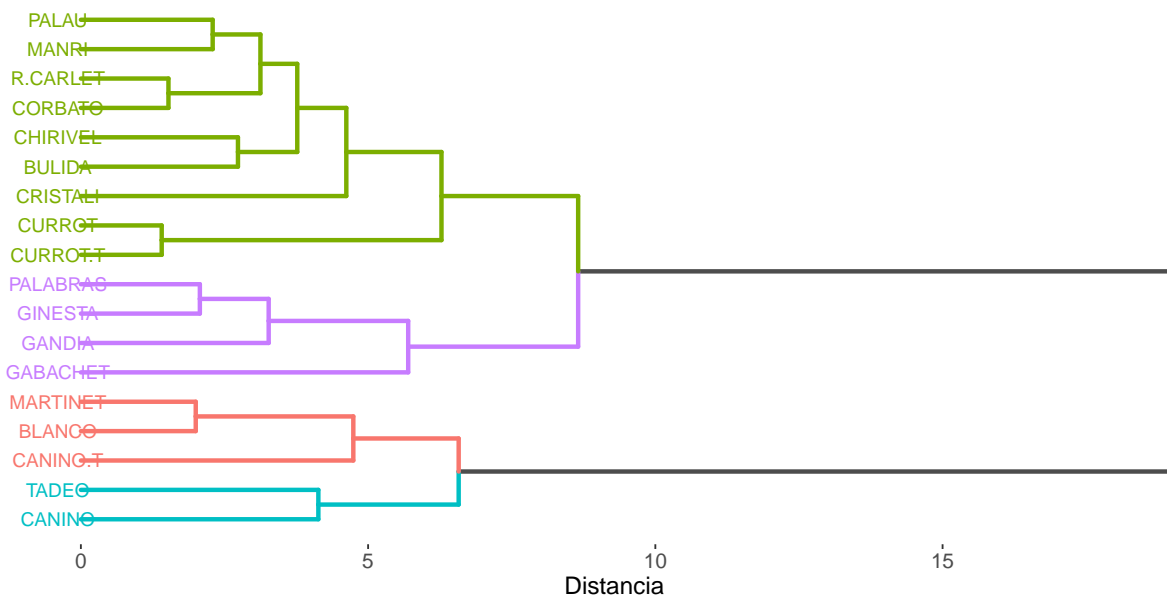
Criterio de encadenamiento Compuesto:



Criterio de encadenamiento UPGMA

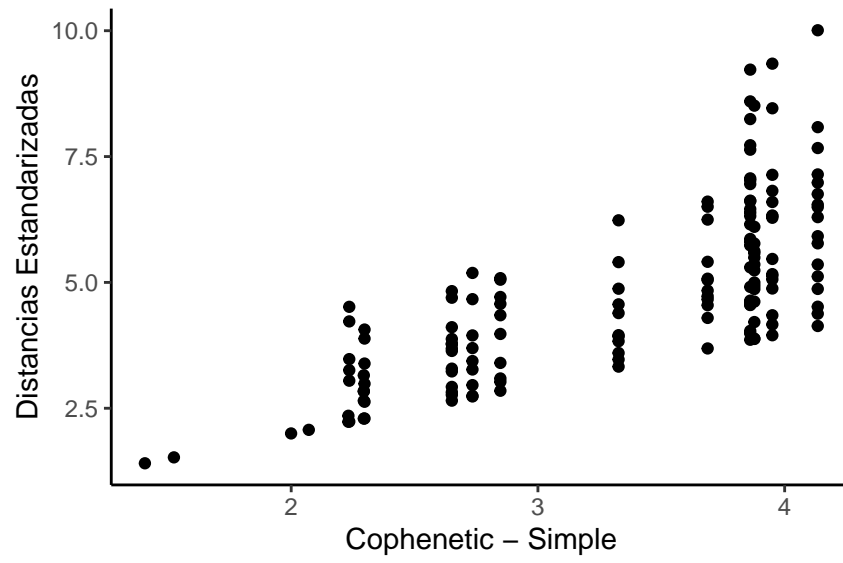


Criterio de encadenamiento WARD

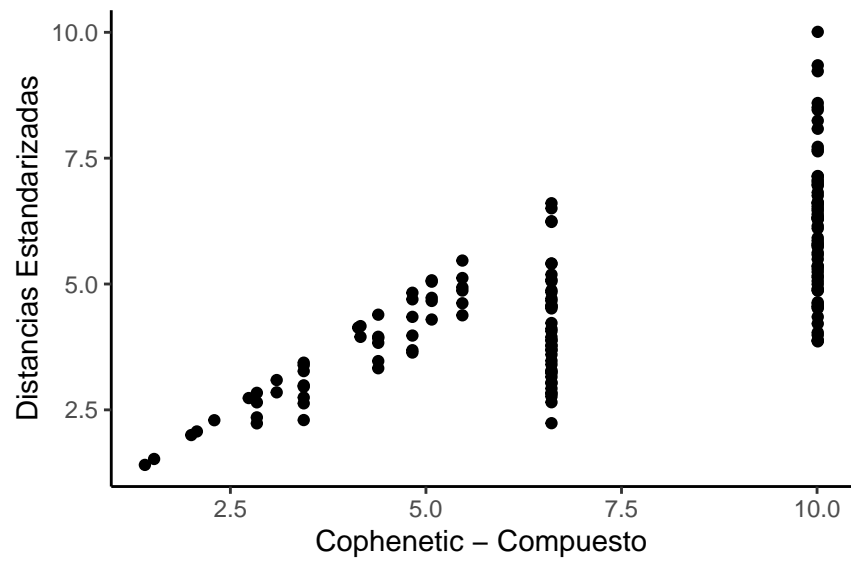


U. Asocie a cada uno de los árboles obtenidos en el inciso anterior la matriz cofenética correspondiente. Que miden los elementos de estas matrices ?

Encadenamiento Simple:

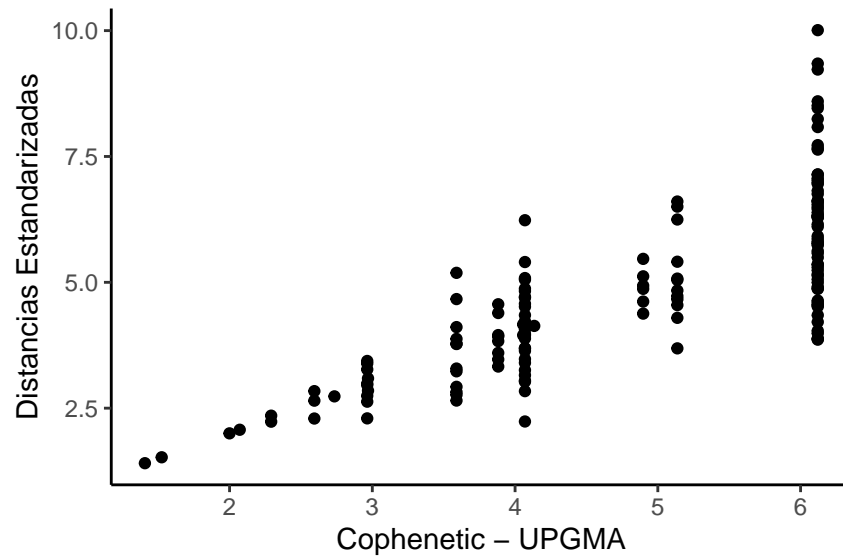


Encadenamiento Compuesto:

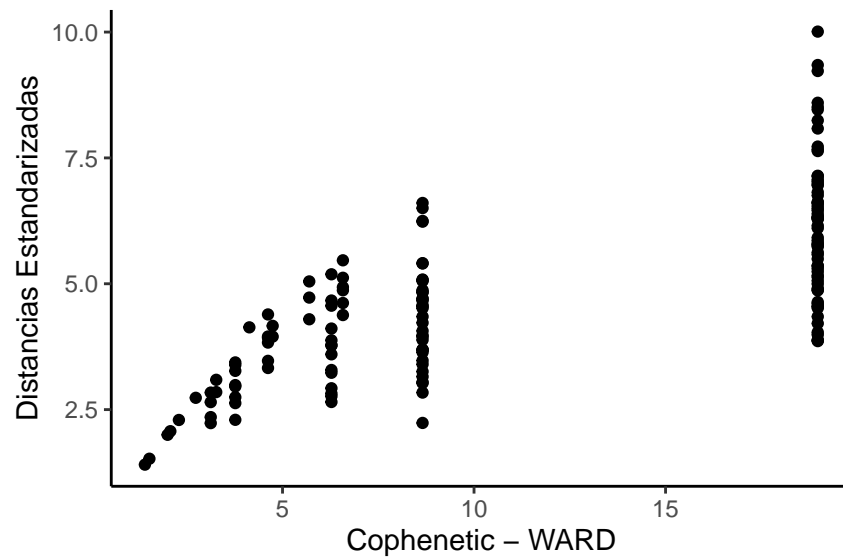


Encadenamiento UPGMA:

Correlacion
0.73
Correlacion
0.72



Encadenamiento WARD:



Los elementos de las matrices anteriores permiten tener una medida de la coherencia del criterio de agrupamiento jerárquico

V. Cuantifique la concordancia entre la matriz de distancia que dio origen a los dendogramas y las 4 matrices copenéticas. A que conclusión llega ?

Encadenamiento simple

Encadenamiento Compuesto

Correlacion
0.77

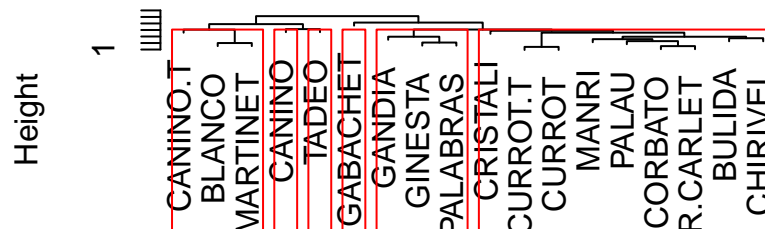
Correlacion
0.73

Encadenamiento UPGMA

Encadenamiento WARD

Observando las correlaciones obtenidas entre la matriz de distancia y las diferentes matrices cofeneticas puede decirse que la obtenida a traves del metodo de encadenamiento UPGMA es la mas alta QUE MAS??

W. Existe algún punto de corte sobre el índice de jerarquización del dendograma UPGMA que origine los mismos agrupamiento de variedades obtenidos en Análisis de Componentes Principales ?



matriz_distancia_std_dist
hclust (*, "average")

X. Mida el grado de concordancia entre los resultados obtenidos por Componentes y por Cluster UPGMA
????

Y. Halle el dendograma aditivo Neighbor Joining, representa mejor la configuración de variedades que el Cluster UPGMA ?

x
0.76

