

## Ejercicio 2

Con el objetivo de tipificar los establecimientos frutícolas del Alto Valle de Río Negro en función de factores socio-económicos, tecnológicos y productivos, se seleccionó una muestra aleatoria de 86 chacras de la región. Recolectándose información acerca a 40 puntos de una encuesta frutícola que corresponden tanto a variables dicotómicas, cualitativas como cuantitativas. Además, se presenta el inconveniente de contar con información faltante, hecho que complica la utilización de la mayoría de las estrategias clásicas de caracterización. Por ello se decidió aplicar el coeficiente de similitud de Gower que por otro lado permite asignar diferente importancia a los caracteres a través de ponderaciones, alternativa válida para eliminar pesos espurios en encuestas donde sobre una misma temática se emplea un mayor número de preguntas respecto a otras.

a) Aplique el coeficiente de similitud de Gower para cuantificar la semejanza entre unidades productivas considerando que se debe asignar igual importancia ( $1/6$ ) a cada grupo de variables. Tenga en cuenta además que las variables 5.2 y 5.3 miden un único concepto que es el tipo de conducción, las variables 5.4, 5.5, 5.6 y 5.7 también cuantifican una sola característica que es el tipo de defensa contra heladas, y por último las variables 6.2, 6.3 y 6.4 miden el estado de la plantación (se podría haber puesto una multinomial con niveles Bueno, Regular y Malo, pero se perdería información)

```
pesos <- c(
  rep(1 / 54, 9),
  rep(1 / 18, 3),
  rep(1 / 42, 7),
  rep(1 / 48, 8),
  1 / 30,
  1 / 60,
  1 / 60,
  rep(1 / 120, 4),
  1 / 30,
  1 / 30,
  1 / 12,
  rep(1 / 36, 3)
)

datos_fct <- as.data.frame(unclass(datos), stringsAsFactors = TRUE)
distancia <- daisy(datos_fct, metric = "gower", weights = pesos)
similitud <- 1 - distancia
```

b) Sobre la matriz de similitud hallada aplique Análisis de Coordenadas Principales y conforme grupos de explotaciones agrícolas según la primer coordenada y subgrupos en función de la segunda coordenada.

En la **Figura 1** se observa la configuración de los individuos en el plano principal de ACP y sus agrupamientos.

- Grupo A: Conformados por los sub-grupos 1 y 2.
  - Se caracterizan por tener valores muy bajos en la primer coordenada.
  - Se diferencian en la segunda coordenada.
- Grupo B: Conformado por los sub-grupos 3 y 4.
  - Se caracterizan por tener valores altos en la primer coordenada.
  - Se diferencian en la segunda coordenada.
- Grupo C: Conformado el sub-grupo 5 solamente.
  - Valores de bajos a promedio en la primer coordenada.
  - Valores promedio en la segunda coordenada.

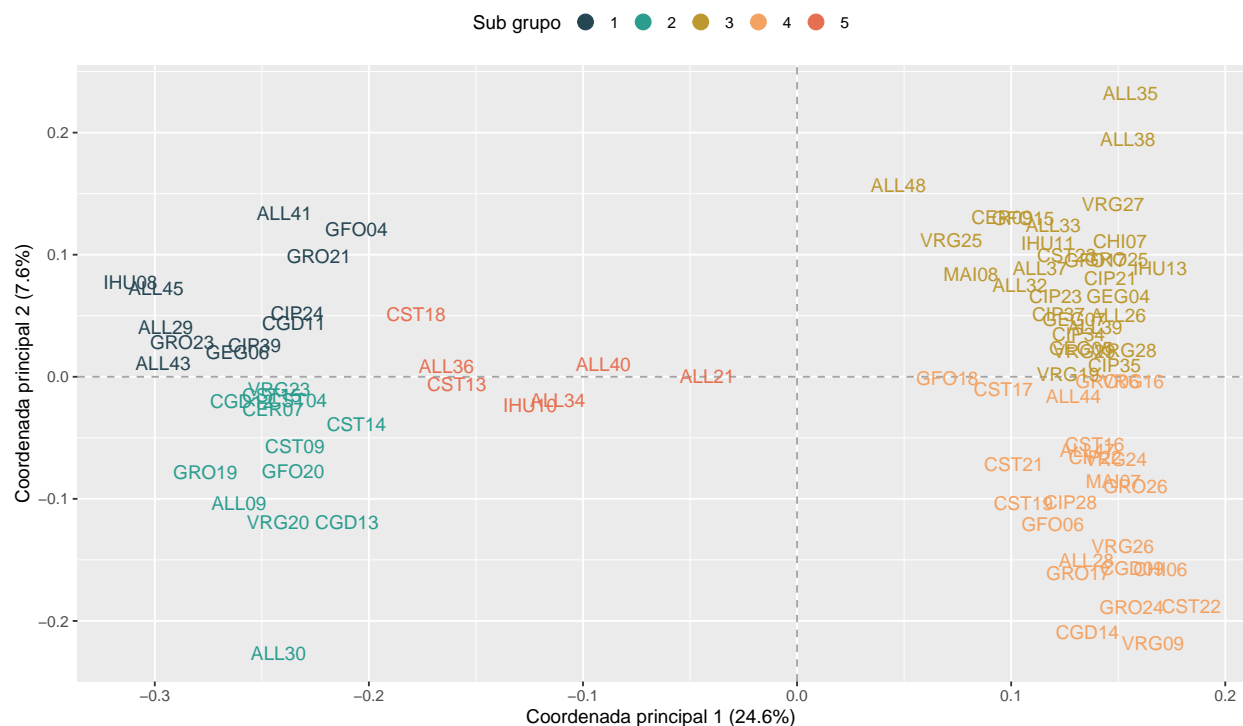


Figure 1: Caracterizacion los establecimientos frutícolas en el plano principal.

c) Caracterice grupos y subgrupos en forma descriptiva y somera en función de las variables originales

#### Grupo A: Tipo empresarial (25 establecimientos)

El 29% de los establecimientos pertenece a este grupo. Las explotaciones tienen un tamaño entre 7 y 98.5 has, con un promedio de 35 ha (44 ha sub-grupo 1 y 27 ha sub-grupo 2). De tenencia propia en el 92% de los casos. Todos los establecimientos de este grupo poseen empaquetacion propia y comercializan en el mercado interno y externo. A su vez, el 36% poseen ademas industrializacion propia (41% sub-grupo 1 y 30% sub-grupo 2). En el 72% de los casos se observa que no trabaja ni el productor ni la familia del mismo (83% sub-grupo 1 y 61% sub-grupo 2). El productor dedica un promedio de 1.36 meses-hombre al año al trabajo en la chacra (1.5 sub-grupo 1 y 1.23 sub-grupo 2). En general tienen mas cantidad de maquinaria que los otros establecimientos. Por ejemplo, estos establecimientos tienen en promedio mas de dos tractores, mientras que los otros menos. El 98% de la superficie se encuentra en estado bueno. En conclusion, se puede inferir que estos establecimientos son de tipo empresarial, y se puede notar que los establecimientos del sub-grupo 1 tienden a ser mas grandes y mas desarrollados que los establecimientos del sub-grupo 2.

##### Sub-grupo 1 - Modernos (12 establecimientos)

- Montes frutales 19 años promedio.
- Densidad media de 544 arboles por ha.
- 66% bajo el sistema de conduccion en espaldera.
- 30.8% de montes juvenes.
- 11.2% variedades nuevas.
- Rendimiento promedio de los montes 38.4 tn/ha.

##### Sub-grupo 2 - Tradicionales (13 establecimientos)

- Montes frutales 26.8 años promedio.
- Densidad media de 363 arboles por ha.
- 77% bajo el sistema tradicional o libre.

- 8.7% de montes jóvenes.
- 6.6% variedades nuevas.
- Rendimiento promedio de los montes 36.8 tn/ha.

#### **Grupo B: Tipo familiar** (54 establecimientos)

Agrupar al 62% de los establecimientos. La superficie va entre 5 y 30 ha, con un promedio de 12 ha. De tenencia propia en el 88% de los casos. Ninguno de estos establecimientos industrializa, empaqueta, ni realiza la comercialización de sus productos. Estos establecimientos se caracterizan por su componente familiar, ya que en el 85% de los casos trabaja el productor o el productor junto a su familia. Además, el productor dedica un promedio de 6.98 meses-hombre al año al trabajo en el establecimiento. El promedio de tractores por establecimiento es 1.52, por debajo de la cantidad del grupo A.

##### Sub-grupo 3 - Modernos (30 establecimientos)

- Montes frutales 16.8 años promedio.
- Densidad media de 525.6 árboles por ha.
- 52% bajo el sistema de conducción en espaldera.
- 31.8% de montes jóvenes.
- 13.53% variedades nuevas
- Rendimiento promedio de los montes 41.85 tn/ha.

##### Sub-grupo 4 - Tradicionales (24 establecimientos)

- Montes frutales 28.9 años promedio.
- Densidad media de 318 árboles por ha.
- 79% bajo el sistema tradicional o libre
- 9.97% de montes jóvenes.
- 3.1% variedades nuevas
- Rendimiento promedio de los montes 34.8 tn/ha.

#### **Grupo C: Tipo mixto** (7 establecimientos)

La superficie está entre 11.85 y 52.65 ha, promediando en 22.5 ha. De tenencia propia en el 85% de los casos. Montes frutales de edad 20 años promedio (situado entre los tradicionales y modernos). En el 71% de los casos trabaja el productor o el productor junto a su familia, dedicando el productor un promedio de 6 meses-hombre al año. A diferencia del grupo familiar, el 85% de estos establecimientos posee empaque propio y el 28.5% industrializa sus productos. El rendimiento promedio es de 38.5 tn/ha. Presenta un 26% de montes jóvenes y un 7.6% de variedades nuevas.

De la comparación de los datos mencionados, con los datos observados para los otros dos grupos claramente diferenciados, vemos que el Grupo C se sitúa entre el medio de los otros grupos. Tiene un componente familiar importante, pero también tiene un componente industrial.

**d)** Verifique si encuentra esos grupos en un cluster UPGMA realizado a partir de la matriz de similitud de Gower

Todos los establecimientos del grupo Grupo A (Empresarial) se corresponden con el cluster marcado en azul. Además, todos los establecimientos del Grupo B (Familiar) se corresponden con el cluster marcado en rosa. Luego, de los 7 establecimientos del grupo C, 2 pertenecen al cluster violeta (se parecen más al familiar), 3 pertenecen al cluster amarillo (se parecen más al empresarial) y uno solo, el CST18, se mezcla con el grupo A en el cluster azul.

En conclusión, es posible encontrar a estos grupos en el cluster UPGMA.

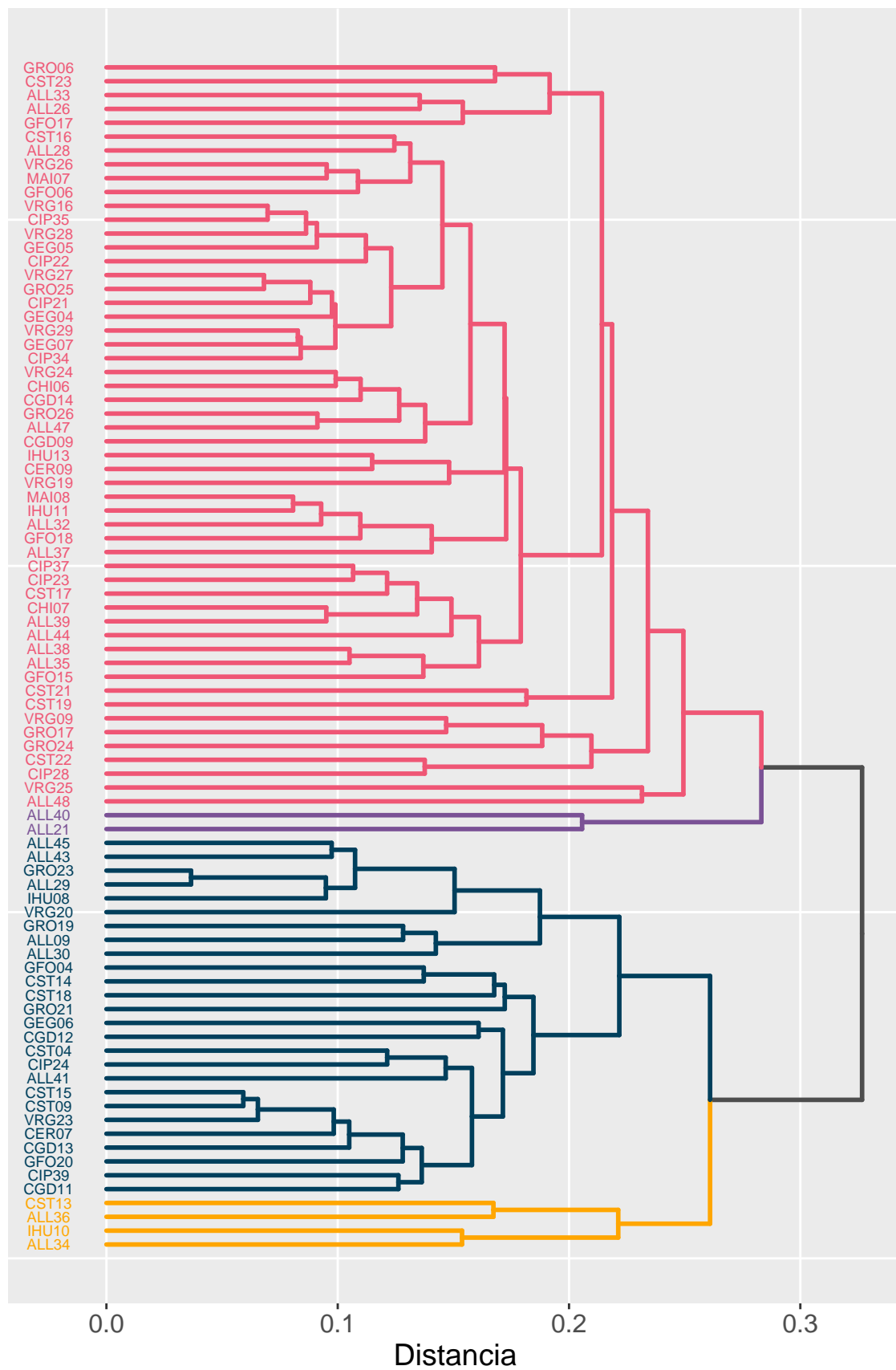


Figure 2: Dendrograma ultrametrico con ligamiento UPGMA.