

Ejercicio 3

Los datos del archivo banana16 corresponden a 40 poblaciones de banana (*Musa spp.*) de las cuales 20 son clones de origen genético desconocido (recolectados en campos de productores) y 20 son variedades de uso internacional Jaffa, Gal, Gran Enanao y Williams, respectivamente (testigos). Los datos son un extracto de la base original perteneciente al Centro Regional Chaco – Formosa de INTA.

A) Analice las correlaciones entre variables en ambos grupos.

A continuacion se presentan tres matrices de correlacion. La primera esta calculada utilizando todos los datos, la segunda utiliza solo a los clones y la tercera utiliza a los datos de testeo.

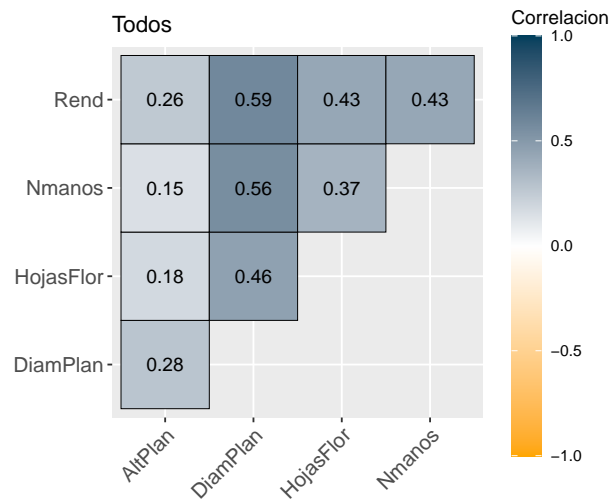


Figure 1: Correlacion entre las variables para el conjunto de datos completo

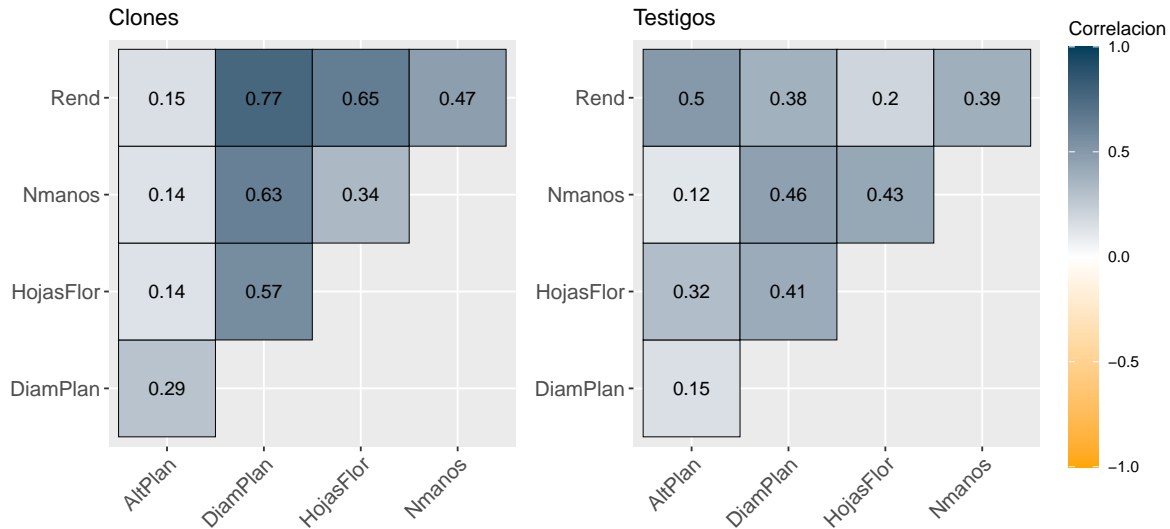


Figure 2: Correlacion entre las variables segun origen genetico.

Al obtener las matrices de correlaciones entre las variables podemos ver que las mismas son todas positivas, tanto en el grupo “Clones” como en el grupo “Testigos”. Cabe destacar que es notable la diferencia en las magnitudes de las correlaciones en los grupos.

Si observamos la matriz correspondiente al grupo “Clones” podemos ver que la variable Rendimiento presenta una correlacion alta con Diametro de Planta y, en menor medida con Hojas a Floracion. Ademas se observan correlaciones, de moderadas a altas, para las variables Diametro de Planta con Cantidad de Manos y Hojas a Floracion.

Por otro lado, observando la matriz del grupo “Testigo” vemos que la variable Rendimiento, a diferencia de lo observado en la matriz anterior, esta mas correlacionada con Altura de planta.

Otra diferencia a destacar, es que la variable Diametro de planta presenta correlaciones mas altas en el grupo “Clones” que en el grupo “Testigos”.

Realice un Análisis Factorial Múltiple Dual y responda las siguientes cuestiones

A) ¿Cuál es el porcentaje de explicación del plano principal?

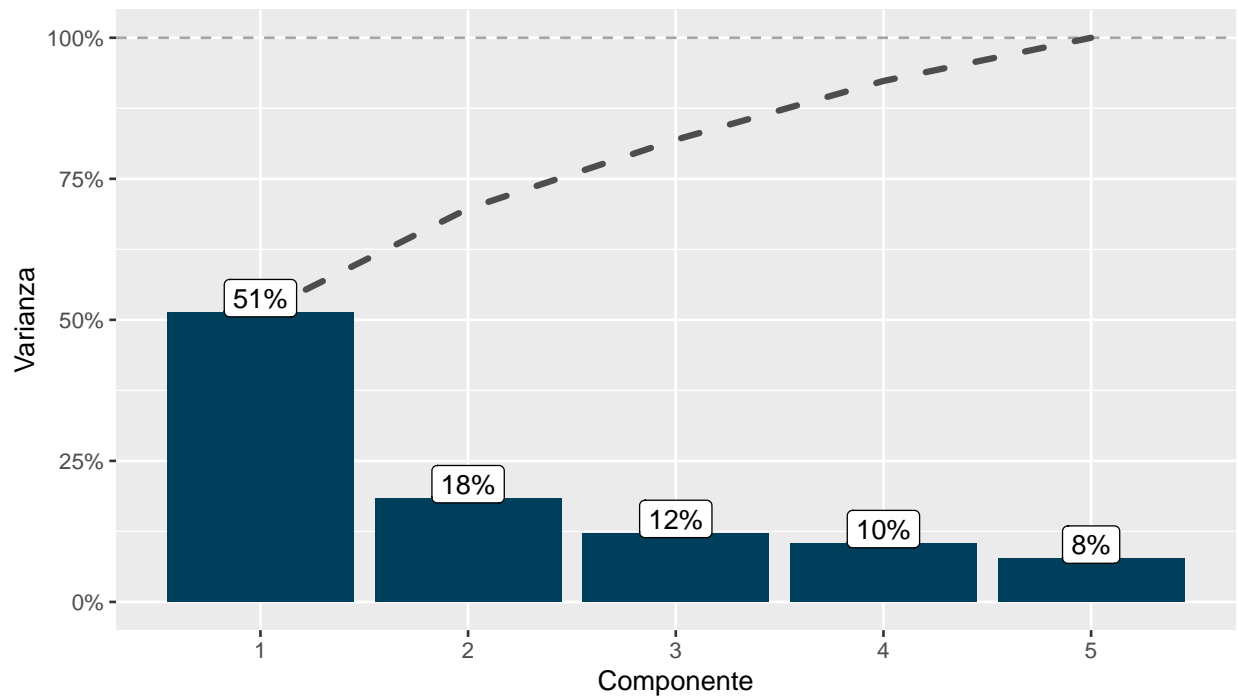


Figure 3: Variabilidad explicada por componente.

El porcentaje de variabilidad explicada sobre el plano principal es del 69.72%. Como puede observarse en la **Figura 3**, el primer autovalor explica un 51.35% mientras que el segundo un 18.37%

B) ¿Qué puede decir a partir del gráfico de las condiciones?

Al representar los grupos sobre el plano principal podemos observar que ambos tienen una alta relación con la primera dimensión, aportando en igual medida a la misma; pero si se proyectan los puntos sobre la segunda dimensión, puede observarse que el grupo “Testigo” tiene un aporte levemente mayor en la formación de esta.

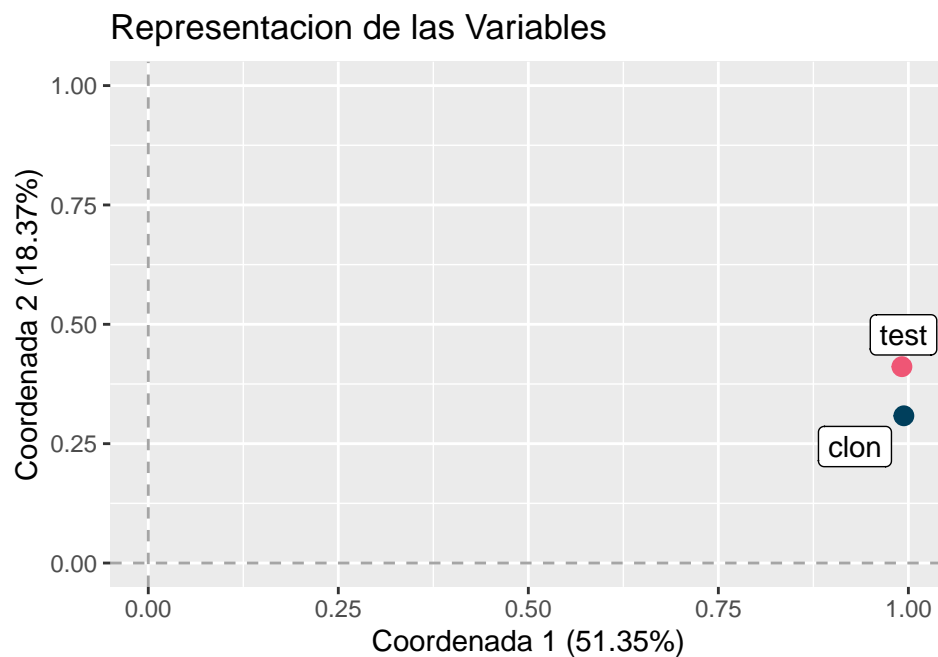


Figure 4: Caracterizacion de los grupos en el plano principal AFM.

C) ¿Cómo se ven representadas en los gráficos del DMFA las conclusiones del inciso a)?

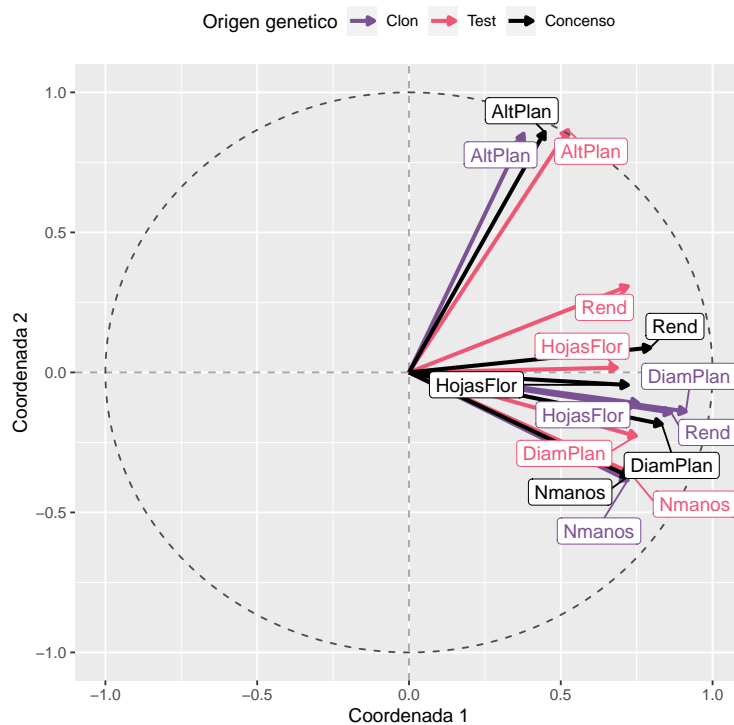


Figure 5: Variables consenso vs Variables parciales

En la **Figura 5** se representa el comportamiento de las variables consenso y parciales correspondientes

a cada grupo. Esta representación nos va a permitir observar los principales cambios que introdujo el agrupamiento de individuos.

Al considerar el grupo “Clones”, vemos que, luego del agrupamiento de individuos, Altura de planta y Cantidad de manos se mantienen prácticamente igual; pero observando la variable Rendimiento vemos que la misma tiene ahora una coordenada negativa sobre el Eje 2. Como vimos en la matriz de correlación correspondiente a este grupo (Figura 2), la variable Rendimiento presenta mayores correlaciones con Diámetro de Planta, Cantidad de manos y Hojas a Floración que con la variable Altura de Planta. Esto explica la transición de la variable Rendimiento sobre el eje 2.

Luego, si consideramos el grupo “Testigos”, podemos concluir que, a diferencia de lo observado en el grupo “Clones”, luego del agrupamiento de individuos la variable Rendimiento tiene una coordenada mayor sobre el eje 2. Este cambio se debe a que, como vimos en la matriz de correlacion correspondiente a este grupo (Figura 2), la variable Rendimiento presenta una correlacion positiva alta con la variable Altura de Planta.