

## Ejercicio 2

Los datos del archivo QUINOA corresponden a 24 accesiones de quinoa nativa del Noroeste Argentino conservadas en el Banco de Germoplasma, caracterizadas a través de 10 variables cuantitativas y 8 variables cualitativas. En el identificador de cada accesión se encuentra la indicación de la procedencia del mismo: AL (altiplano), VA (Valles de altura), VS (Valles secos), VH (Valles húmedos orientales).

A) Realice un ACP con las variables cuantitativas.

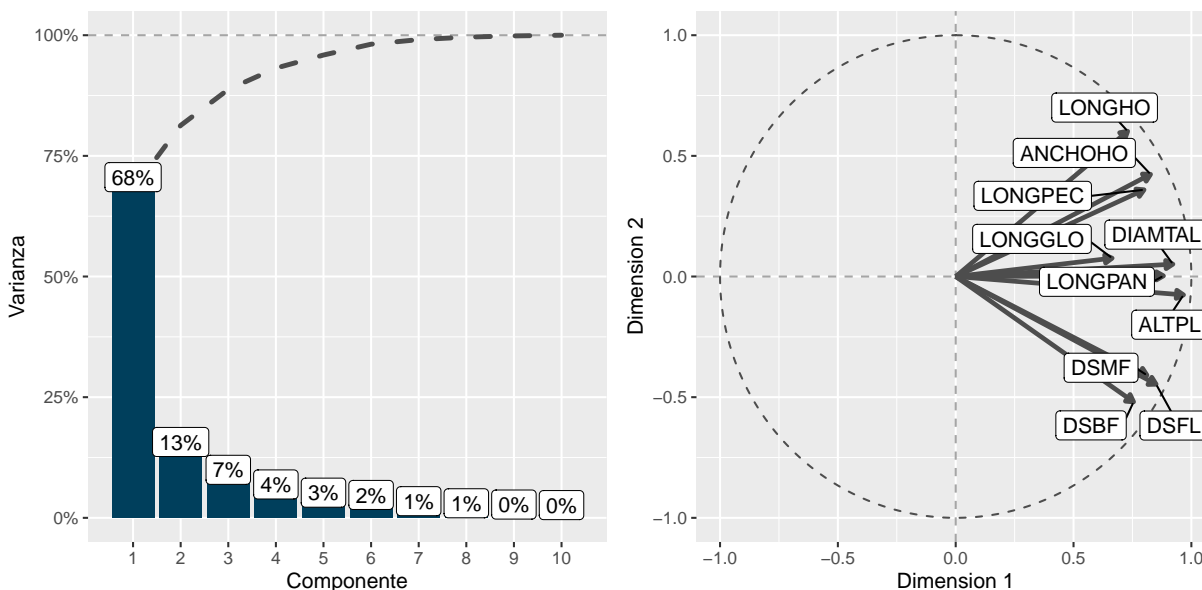


Figure 1: Izquierda: porcentaje de varianza retenido por cada componente. Derecha: Correlacion de las variables con los componentes.

Luego de realizar el analisis de componentes principales, podemos observar que el porcentaje de variabilidad total explicada por las dos primeras componentes es de 81.31%; el 68.09% corresponde a la primera componente, mientras que el 13.22% restante corresponde a la segunda componente. Esto indica que la configuracion en el plano principal conserva en muy buena manera la configuracion de los datos en el espacio original.

En relacion a las variables que mas aportan a la formacion de las componentes, podemos decir que todas las variables contribuyen de forma positiva a la primer componente, y que longitud del glomérulo es quien lo hace en menor medida. Con respecto a la segunda componente, se puede notar que las variables que mas aportan son Longitud y Ancho de la hoja (de forma positiva) y días desde siembra a botón floral y a floración (de forma negativa).

El panel derecho de la **Figura 1** tambien nos permite identificar grupos de variables segun su correlacion. Podemos identificar a los siguientes tres grupos que presentan correlacion muy alta entre sus variables:

- Grupo 1: Longitud del glomérulo, diámetro del tallo, longitud de la panoja y altura de planta (medidas morfológicas de la planta y su fruto).

- Grupo 2: Longitud de la hoja, ancho de la hoja y longitud del peciolo (medidas morfológicas asociadas a la hoja).
- Grupo 3: Días desde siembra a botón floral, días desde siembra a floración y días desde siembra a madurez fisiológica (medidas temporales).

Se puede notar que el Grupo 1 presenta una correlacion moderada con los grupos 2 y 3, mientras que estos dos ultimos estan practicamente no correlacionados entre si ya que forman un angulo aproximado de 90 grados.

A partir de la **Figura 2** podemos hacer una caracterizacion de accesiones de quinoa segun la procedencia.

En primer lugar, se puede notar que las accesiones del Altiplano poseen valores bajos en la primer componente, y en torno al promedio o mas bien altos en la segunda. Esto nos permite decir que estas accesiones presentan valores pequeños en cuanto a las características morfológicas de la hoja y de la planta, pero que sus variables temporales tambien son menor que el promedio. Es decir, la aparicion de boton floral, la floracion y la madurez fisiologica se dan mas tempranamente que en otras regiones.

Por otro lado, se tiene que las accesiones de Valles Humedo y Seco presentan similares valores en cuanto a la primer dimension, pero difieren en cuanto a la segunda. Las accesiones del Valle Humedo suelen presentar valores menor al promedio y mientras que las del Valle Seco son mayores. Asi, podemos decir que en ambas procedencias se dan similares características en cuanto a las medidas morfológicas de la planta (de gran tamaño en ambos casos), pero que las accesiones del Valle Seco suelen asociarse a mayor tamaño de hoja y peciolo, mientras que las del Valle Humedo se asocian a mayor tiempo hasta la aparicion del boton floral, floracion y madurez fisiológica.

Con respecto a Valles de Altura, podemos decir que presentan un comportamiento cercano al promedio en terminos de la primera componente y estan en torno al promedio, o mas bien bajo el promedio, en cuanto a la segunda. En otras palabras, se asocian a accesiones de tamaño de planta promedio, tamaño de hoja tirando a bajo, y tiempos de floracion y madurez tardios.

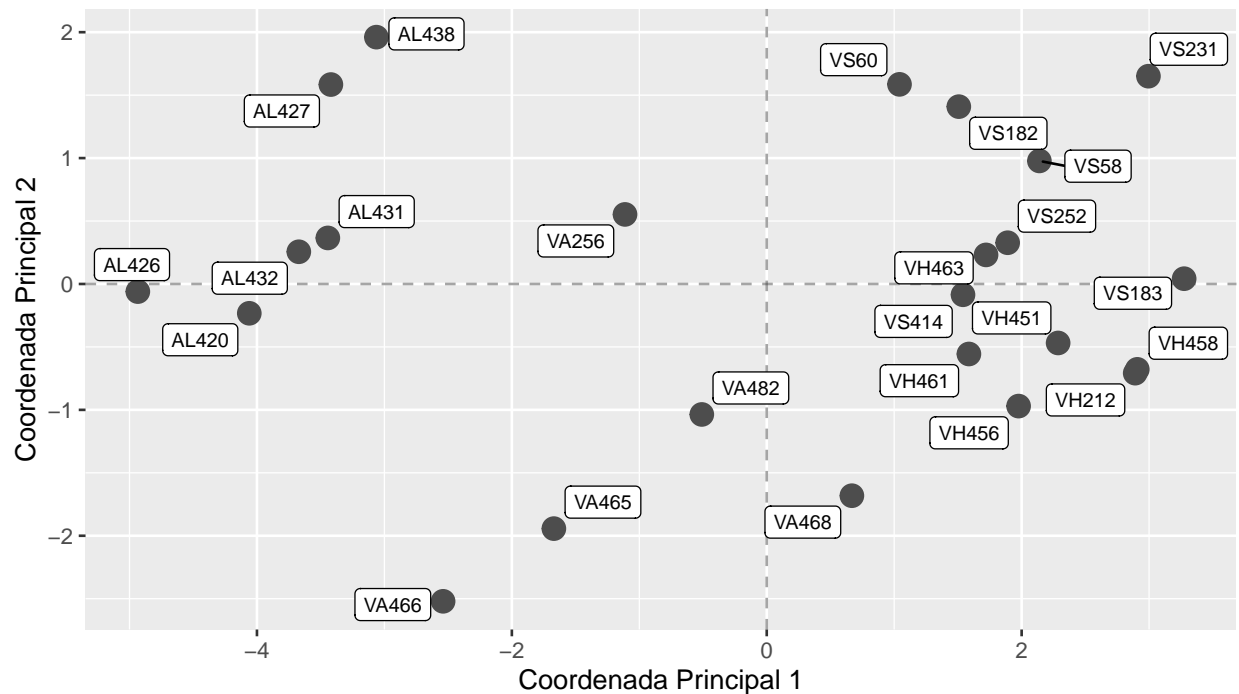
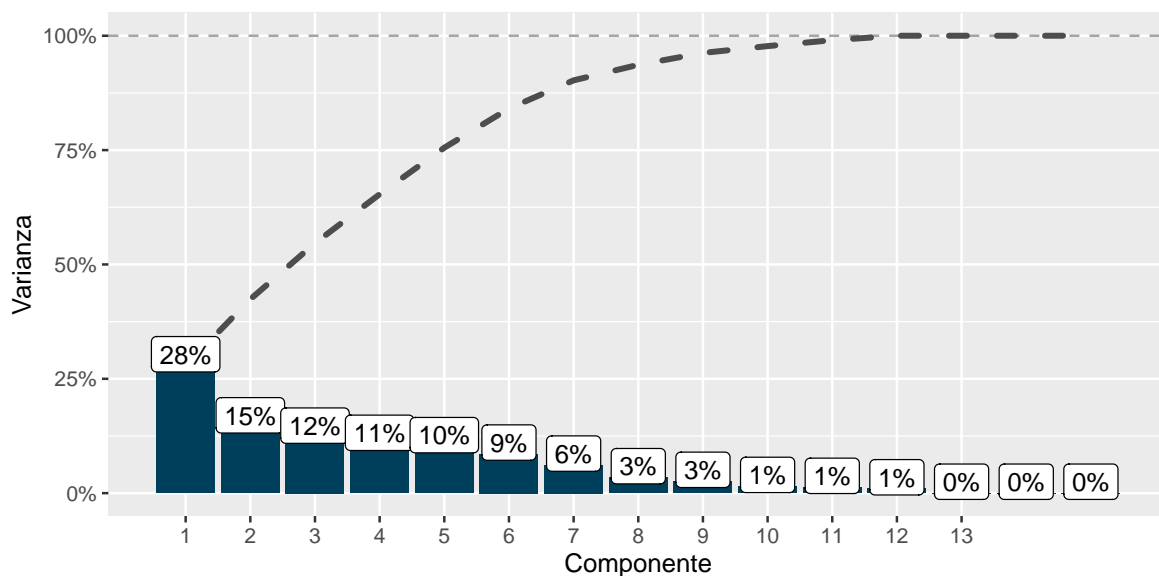


Figure 2: Accesiones de Quinoa en el plano principal del ACP.

B) Realice un ACM con las variables cualitativas y compare las configuraciones de individuos provistas por

ambas técnicas en el plano principal.



El porcentaje de inercia explicado es por las primeras dos componentes es de 42.39%, bastante menor que lo obtenido con ACP. Esto indica que la representacion que obtenemos en el plano principal del analisis obtenido con ACM no conserva la configuracion original de los datos de la misma manera que lo hace el analisis ACP con las variables cuantitativas.

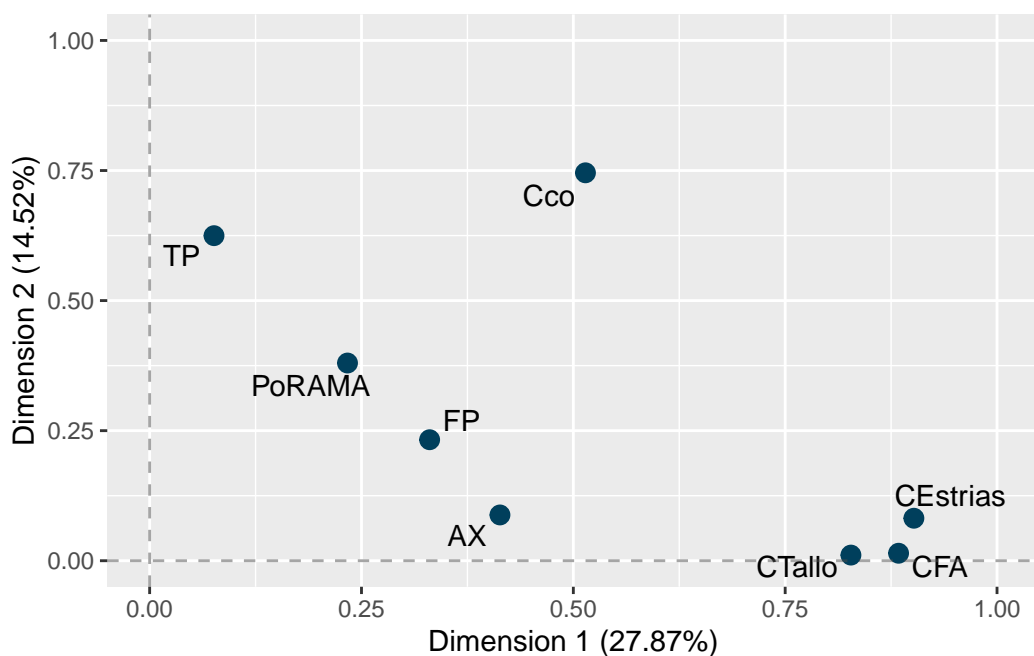


Figure 3: Contribucion de las variables a los dos primeros ejes.

Las variables que mas aportan a la formacion de la Dimension 1 son Color de las estrias, Color de panoja a fin de antesis y Color de tallo. Esta contribucion se da de manera practicamente exclusiva, ya que no se observa contribucion en la segunda dimension. Por otro lado, tipo de panojy el color de la

panoja a la cosecha y a son las que mas contribuyen a la Dimension 2. Esta ultima variable tambien contribuye moderadamente a la primera dimension.

Finalmente notamos que la presencia de ramas, la forma de la panoja, y la presencia de axilas tienen muy poca contribucion a la formacion de cualquiera de los dos ejes.

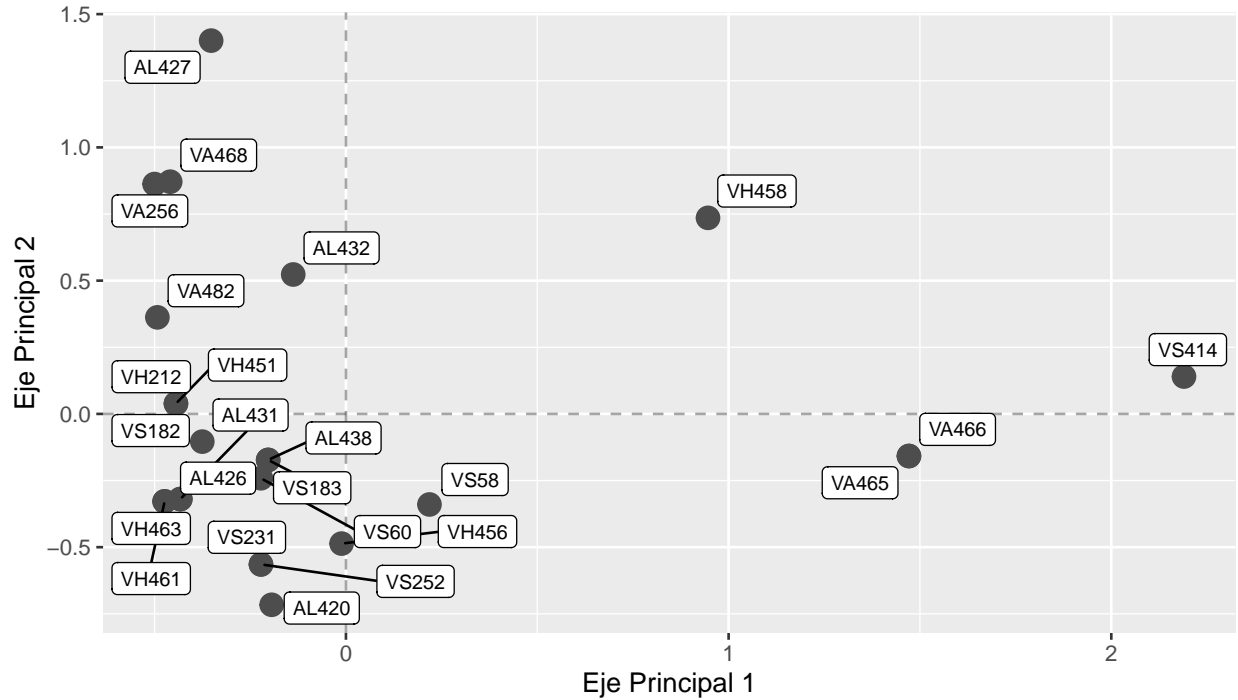


Figure 4: Configuración de los individuos en el plano principal de ACM.

Se nota en la **Figura 4** que en el análisis AFM basado en las variables cualitativas no permite formar agrupamientos en términos del ambiente en el plano principal, a diferencia de lo que sucedió con el análisis ACP basado en variables cuantitativas.

Luego, si comparamos las configuraciones de individuos provistas por las dos técnicas realizadas, podemos concluir que las mismas parecen no ser homogéneas, ya que, como se pudo observar, una técnica logra una diferenciación por ambientes mientras que la otra no.

Con alguna técnica numérica podríamos realizar esa comparación: Para poder comparar, siempre en el plano principal, en acp, se podría buscar una matriz de distancias euclídeas entre individuos, lo mismo puedo tmb hacer para acm. Luego, puedo comparar esas matrices de distancia, medir el “acuerdo”, a través de una corr de matrices de distancias (tiene que dar un valor bajo)

Realice un AFM considerando como grupos a ambos tipos de variables:

C) Realice el análisis de la interestructura. ¿Qué puede decir de la relación que existe entre los grupos?

Table 1: Coeficiente Lg

	Cuantitativas	Cualitativas	MFA
Cuantitativas	1.06	0.23	1.11
Cualitativas	0.23	1.92	1.85
MFA	1.11	1.85	2.55

Al obtener el Coeficiente  $L_g$ , el cual nos brinda una medida de estructura comun entre las variables Cualitativas y Cuantitativas, observamos que el mismo es de 0.23. Dado que se obtuvo un valor bajo, podemos decir que las variables del grupo “Cuantitativas” no estan correlacionadas con las del grupo “Cualitativas”. Es decir, las configuraciones no tienen estructura en comun.

Table 2: Coeficiente Ng

	Cuantitativas	Cualitativas	MFA
Cuantitativas	1.03	0.48	1.05
Cualitativas	0.48	1.39	1.36
MFA	1.05	1.36	1.60

Al obtener una medida de dimensionalidad de cada grupo (coeficiente  $N_g$ ), observamos que el grupo “Cualitativas” tiene una dimensionalidad mayor que “Cuantitativas”

Table 3: Coeficiente RV

	Cuantitativas	Cualitativas	MFA
Cuantitativas	1.00	0.16	0.68
Cualitativas	0.16	1.00	0.84
MFA	0.68	0.84	1.00

Tal como se intuyo al realizar el analisis por separado, las configuraciones cualitativas y cuantitativas de las quinoas no son “homoteticas” entre si. Luego, al ser un valor bajo indica que los grupos de variables brindan informacion complementaria.

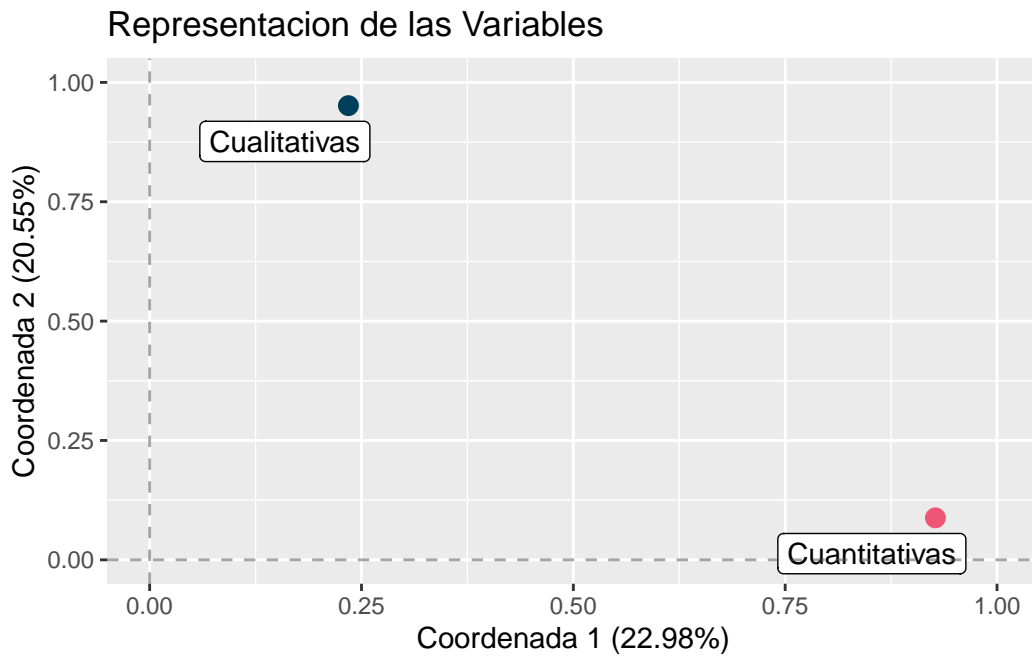


Figure 5: Caracterizacion de los grupos de variables en el plano principal AFM.

Si observamos el grafico anterior, podemos ver que las variables Cuantitativas contribuyen unicamente

a la formacion de la Coordenada 1, mientras que las Cualitativas son las que contribuyen a la formacion de la Coordenada 2 del Analisis Factorial Multiple. Estas ultimas, presentan una leve contribucion en la primer coordenada, que se corresponde con el coeficiente  $N_g$  un poco mayor a 1 como es el 1.39 antes visto.

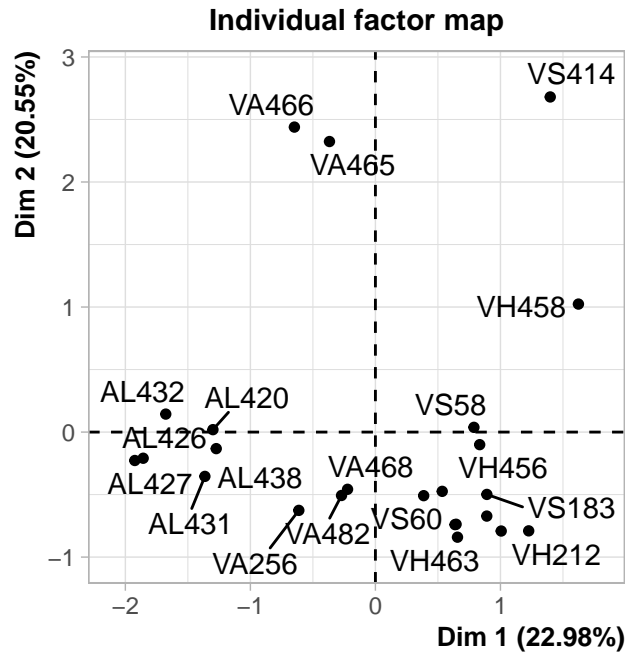
Table 4: Contribuciones a los Ejes - Cualitativas

	Dim.1	Dim.2
AxNo	0.07	3.73
AxSi	0.12	6.22
CEstAm	0.14	0.50
CEstNo	1.26	2.64
CEstPu	0.74	16.54
CEstRo	2.95	0.85
CTaPu	0.38	10.47
CTaRo	1.44	6.63
CTaVe	0.01	2.43
PoRaBa	0.65	0.64
PoRaNo	0.39	1.56
PoRaOb	0.00	3.61
CFAbI	0.15	3.31
CFAPu	0.74	16.54
CcoAm	0.00	0.28
CcoBl	0.26	0.54
CcoGr	2.73	0.05
CcoMar	1.44	6.63
CcoPu	1.95	0.97
TPDif	0.86	0.14
TPNoDif	3.27	0.52
FPAmA	0.42	4.51
FPGlom	0.21	2.26

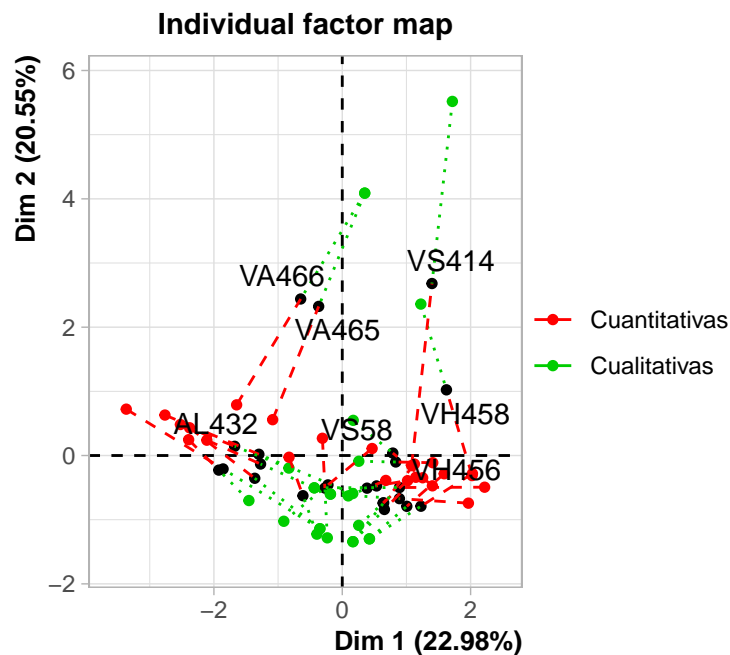
Table 5: Contribuciones a los Ejes - Cuantitativas

	Dim.1	Dim.2
DIAMTAL	9.41	0.77
ALTPL	10.98	0.59
LONGHO	5.48	1.79
ANCHOHO	7.47	2.85
LONGPEC	6.04	1.61
LONGPAN	8.59	0.32
LONGGLO	6.09	0.02
DSBF	8.47	0.43
DSFL	9.97	0.04
DSMF	7.31	0.04

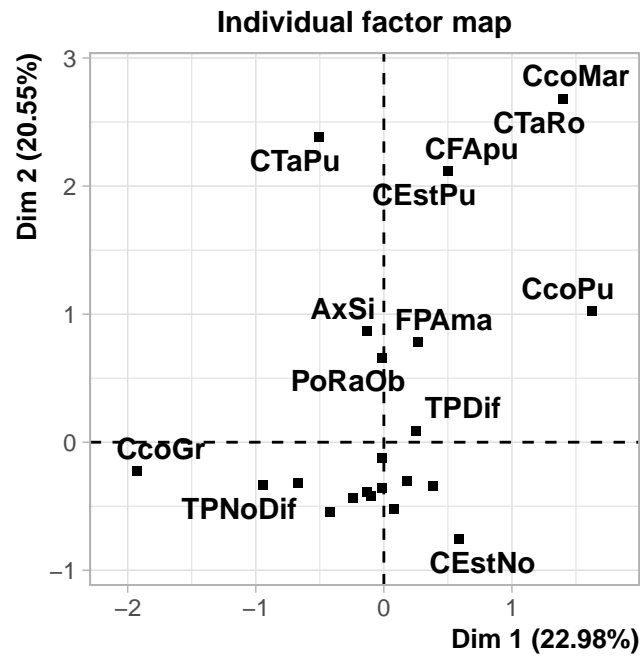
D) ¿Existe un agrupamiento de individuos por zona?



Vemos que abajo a la izq estan las Altiplano (similar a como habia quedado en las cuali), luego vemos valles humedos y secos juntos. Las de altura estan mas dispersas, y forman dos grupos sobre el eje vertical, lo que estaria indicando que difieren en sus vbles cualitativas, ya q estas vbles son las q contribuyen a la formacion del eje vertical. Podemos ver que hay dos “individuos” (VS414 Y VH458) qson atipicos. Seguramente esto se deba a que tengan valores mas altos de cualitativas, distinto al resto (ya q el eje 2 es cuali) VER!!



Con este grafico confirmamos lo anterior, VS414 y VH458 poseen valores mas altos de cualitativas q el resto, es por ello q se apartan del grupo. REDACTAR MEJOR!



E) Caracterice los grupos de individuos a través de todas las variables.

Con los graficos anteriores caracterizar los grupos y decir cuales quinoas tienen hojas mas grandes y bla bla bla