

Laboratorio 3:Análisis Factorial Múltiple (AFM)

KEVIN STEVEN GARCÍA^a, ALEJANDRO VARGAS^b, ALEJANDRO SOTO^c, NATALIA BUITRON^d

1. Introducción

En el presente informe veremos la aplicación del AFM a la base de datos data(orange) de la librería missMDA, la cual corresponde a la descripción sensorial de 12 jugos de naranja por 8 atributos, esta base presenta cerca del 20 % de datos faltantes, por lo cuál se hará primero un proceso de imputación para posteriormente realizar el método AFM. Se analizará e interpretará el porcentaje de Inercia explicado, la nube de individuos, la nube de variables, la nube de los grupos, los coeficientes Lg y Rv de Escoufier y se realizará el gráfico de representación Superpuesta y de los ejes parciales, todo esto luego de un debido análisis descriptivo de las variables del estudio.

La base de datos es la siguiente:

	Color.intensity	Odor.intensity	Attack.intensity	Sweet	Acid	Bitter	Pulp	Typicity
1	4.791667	5.291667	NA	NA	NA	2.833333	NA	5.208333
2	4.583333	6.041667	4.416667	5.458333	4.125000	3.541667	4.625000	4.458333
3	4.708333	5.333333	NA	NA	4.291667	3.166667	6.250000	5.166667
4	6.583333	6.000000	7.416667	4.166667	6.750000	NA	1.416667	3.416667
5	NA	6.166667	5.333333	4.083333	NA	4.375000	3.416667	4.416667
6	6.333333	5.000000	5.375000	5.000000	5.500000	3.625000	4.208333	4.875000
7	4.291667	4.916667	5.291667	5.541667	5.250000	NA	1.291667	4.333333
8	NA	4.541667	4.833333	NA	4.958333	2.916667	1.541667	3.958333
9	4.416667	NA	5.166667	4.625000	5.041667	3.666667	1.541667	3.958333
10	4.541667	4.291667	NA	5.791667	4.375000	NA	NA	5.000000
11	4.083333	5.125000	3.916667	NA	NA	NA	7.333333	5.250000
12	6.500000	5.875000	6.125000	4.875000	5.291667	4.166667	1.500000	3.500000

2. Análisis descriptivo

- Definición de variables:
 - Intensidad del color: Cuantitativa continua. Escala de intervalos.
 - Intensidad del olor: Cuantitativa continua. Escala de intervalos.
 - Intensidad del ataque(sabor inicial del jugo en la boca): Cuantitativa continua. Escala de intervalos.
 - Dulce: Cuantitativa continua. Escala de intervalos.
 - Ácido: Cuantitativa continua. Escala de intervalos.
 - Amargo: Cuantitativa continua. Escala de intervalos.
 - Pulpa: Cuantitativa continua. Escala de intervalos.
 - Tipicidad: Cuantitativa continua. Escala de intervalos.

^aCódigo: 1533173. E-mail: kevin.chica@correounivalle.edu.co

^bCódigo: 1525953. E-mail: jose.alejandro.vargas@correounivalle.edu.co

^cCódigo: 1532457. E-mail: asotomurillo@gmail.com

^dCódigo: . E-mail:

- Resumen Estadístico:

	Color.intensity	Odor.intensity	Attack.intensity	Sweet	Acid	Bitter	Pulp	Typicity
Mínimo	4.083	4.292	3.917	4.083	4.125	2.833	1.292	3.417
Cuartil 1	4.448	4.958	4.833	4.510	4.375	3.104	1.510	3.958
Mediana	4.646	5.292	5.292	4.938	5.042	3.583	2.479	4.438
Media	5.083	5.326	5.319	4.943	5.065	3.536	3.312	4.462
Cuartil 3	5.948	5.938	5.375	5.479	5.292	3.792	4.521	5.042
Máximo	6.583	6.167	7.417	5.792	6.750	4.375	7.333	5.250
Des. Estándar	0.9809135	0.630481	1.005842	0.6318818	0.7971605	0.5529671	2.22528	0.6521764
NA's	2	1	3	4	3	4	2	0

3. Análisis Factorial Múltiple

Para realizar el AFM, debemos dividir la tabla en subtablas dependiendo de los tipos de variables y su definición. En este caso, dado que todas las variables son cuantitativas continuas, nuestro criterio de división se basó en la definición de las variables. Decidimos dividir la tabla en dos subtablas, la primera consta de las tres primeras variables (columnas), Intensidad del color, intensidad del olor e intensidad del ataque, esta primera tabla fue denominada percepción previa, ya que las tres variables se miden antes de degustar el jugo. La segunda subtabla consta de las cinco últimas variables, dulce, ácido, amargo, pulpa y tipicidad, esta segunda tabla fue denominada percepción posterior, ya que todas estas variables involucradas son medibles solamente después de catar o degustar el jugo.

3.1. Imputación AFM

Dado que la base de datos tiene valores faltantes en una cantidad considerable (20 % de valores faltantes), se realizó una imputación o estimación de estos valores por el método EM.

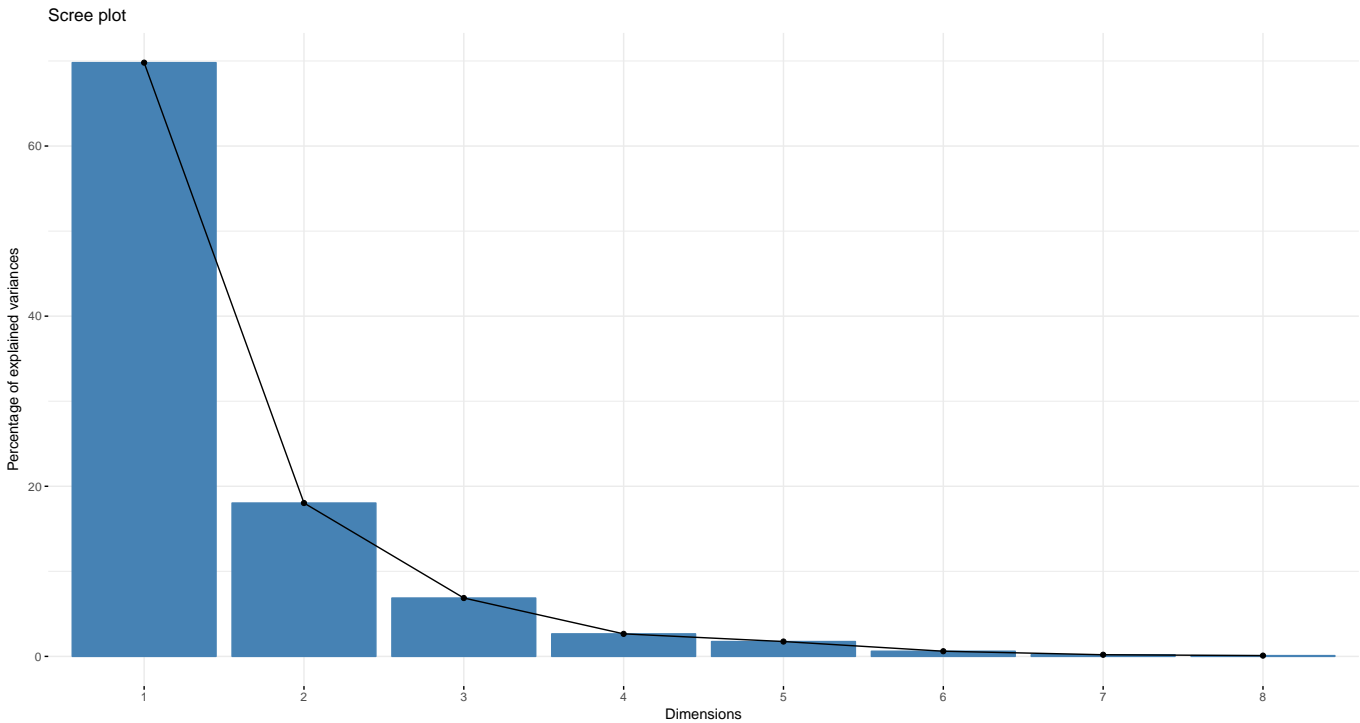
La matriz imputada es:

	Color.intensity	Odor.intensity	Attack.intensity	Sweet	Acid	Bitter	Pulp	Typicity
1	4.791667	5.291667	4.124561	5.501018	4.154464	2.833333	5.574109	5.208333
2	4.583333	6.041667	4.416667	5.458333	4.125000	3.541667	4.625000	4.458333
3	4.708333	5.333333	4.168132	5.449296	4.291667	3.166667	6.250000	5.166667
4	6.583333	6.000000	7.416667	4.166667	6.750000	4.748854	1.416667	3.416667
5	6.169035	6.166667	5.333333	4.083333	5.332292	4.375000	3.416667	4.416667
6	6.333333	5.000000	5.375000	5.000000	5.500000	3.625000	4.208333	4.875000
7	4.291667	4.916667	5.291667	5.541667	5.250000	3.213539	1.291667	4.333333
8	4.536606	4.541667	4.833333	5.430180	4.958333	2.916667	1.541667	3.958333
9	4.416667	4.957997	5.166667	4.625000	5.041667	3.666667	1.541667	3.958333
10	4.541667	4.291667	4.278448	5.791667	4.375000	2.724521	3.593447	5.000000
11	4.083333	5.125000	3.916667	5.712512	3.833623	2.790614	7.333333	5.250000
12	6.500000	5.875000	6.125000	4.875000	5.291667	4.166667	1.500000	3.500000

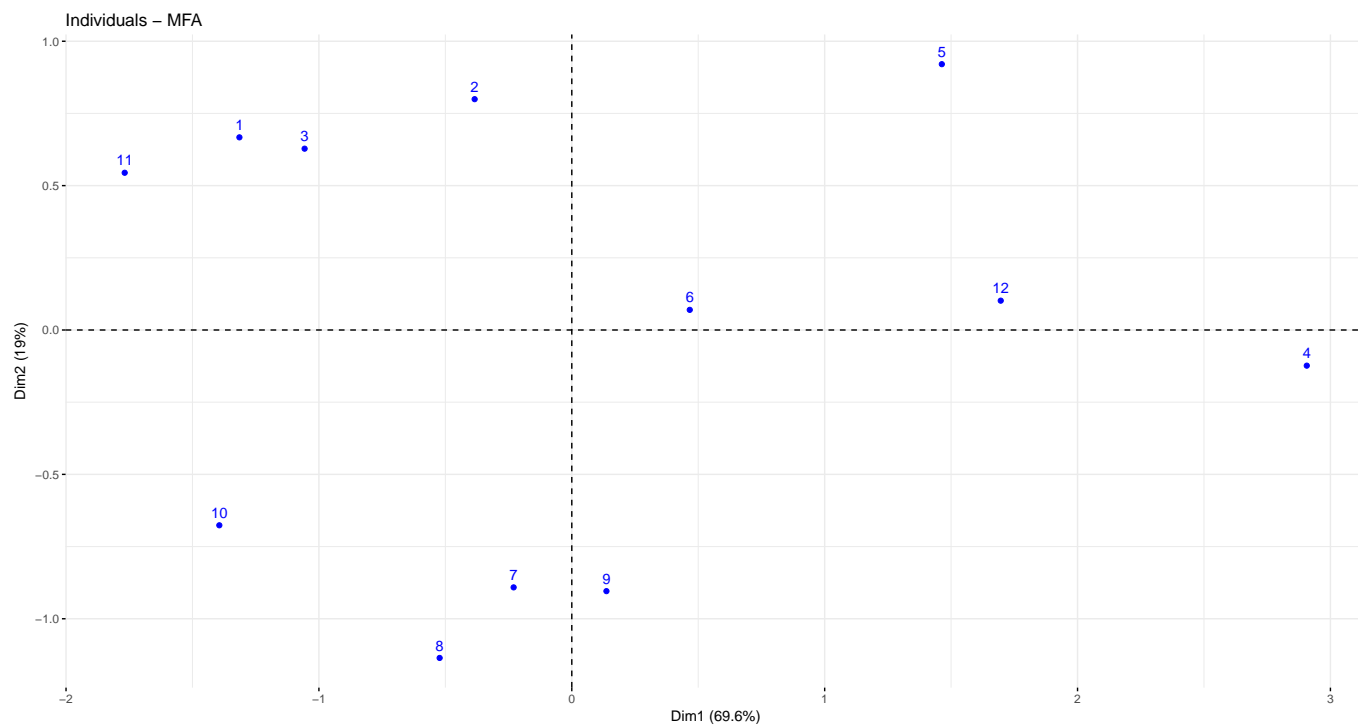
3.2. Porcentaje de inercia explicado

El porcentaje de varianza explicado por cada dimensión se puede ver en la siguiente tabla:

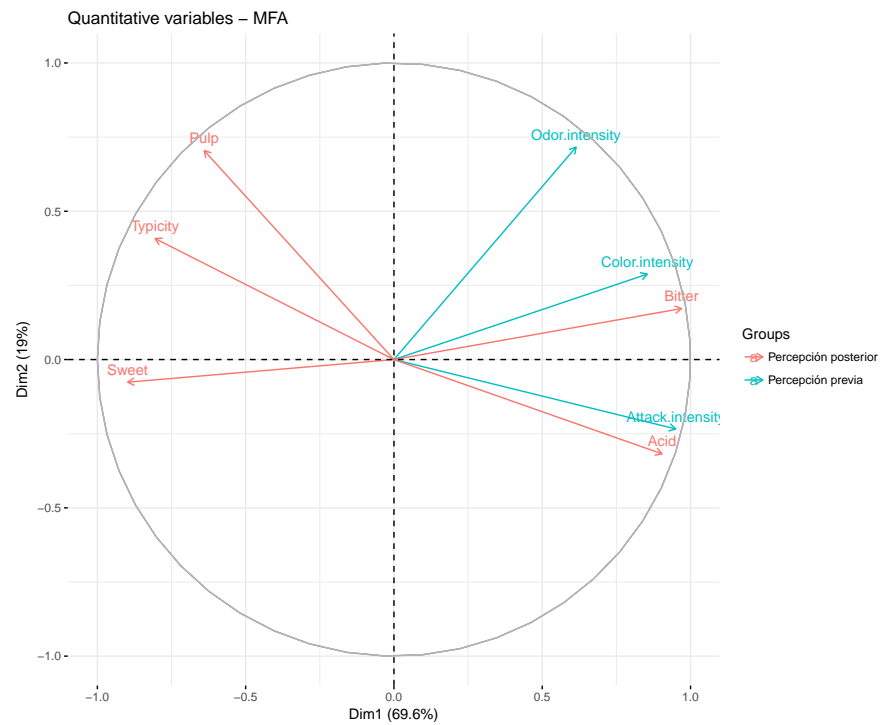
	Valor propio	Porcentaje de Inercia	Porcentaje de Inercia acumulado
Dim.1	1.840695280	69.6179879	69.61799
Dim.2	0.501038381	18.9500589	88.56805
Dim.3	0.169785432	6.4215518	94.98960
Dim.4	0.065450856	2.4754542	97.46505
Dim.5	0.044459815	1.6815401	99.14659
Dim.6	0.016131006	0.6101000	99.75669
Dim.7	0.004671617	0.1766879	99.93338
Dim.8	0.001761408	0.0666192	100.00000



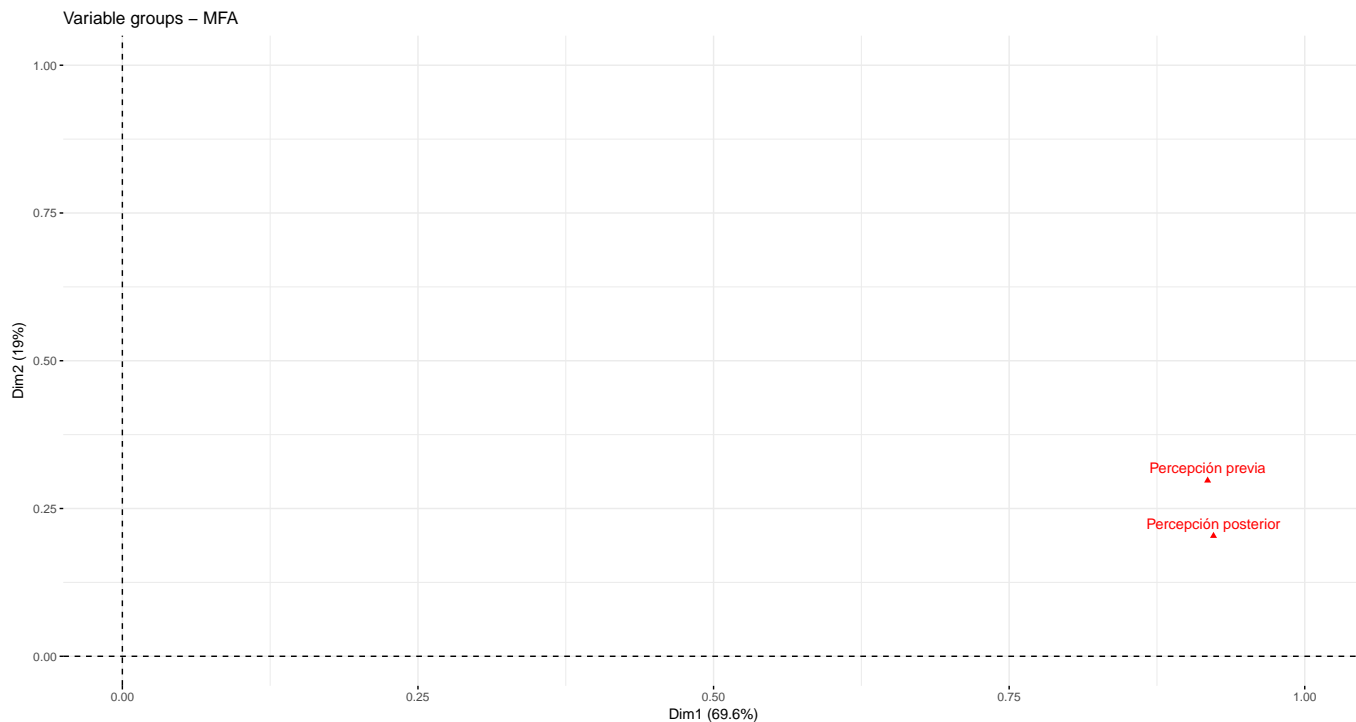
3.3. Nube de individuos



3.4. Nube de variables



3.5. Nube de los grupos



3.6. Coeficientes Lg y Rv de Escoufier

- Coeficiente Lg: Es un índice de unión general entre grupos de variables, que expresa de una manera “el número de dimensiones comunes (para ambos grupos) de inercia comparable a la inercia axial máxima de un grupo”

Los coeficientes Lg se pueden observar en la siguiente tabla:

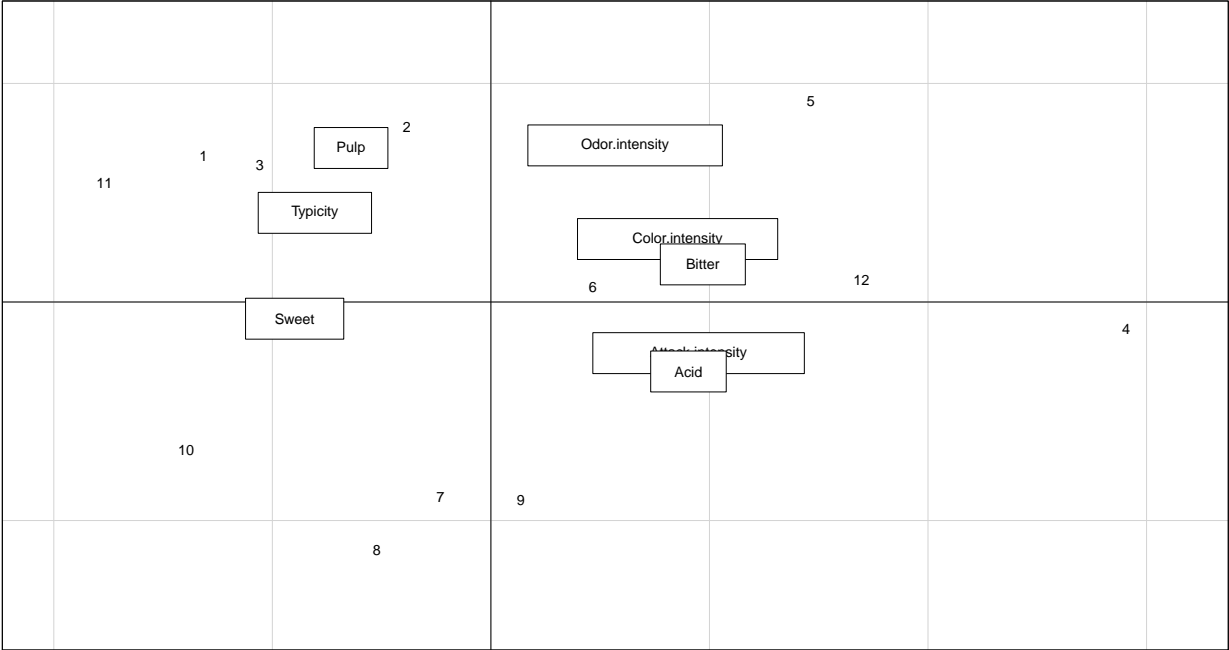
	Percepción previa	Percepción posterior	MFA
Percepción previa	1.0839432	0.7761085	1.0105158
Percepción posterior	0.7761085	1.0384112	0.9857795
MFA	1.0105158	0.9857795	1.0845333

- Coeficiente Rv de Escoufier: Es una generalización multivariada del coeficiente de correlación de Pearson al cuadrado. Este coeficiente mide el vínculo entre dos grupos o dos matrices de variables. Este coeficiente, al igual que el de correlación de Pearson, se encuentra entre 0 (todas las variables del primer grupo o matriz, son ortogonales a todas las variables del segundo grupo o matriz) y 1 (los dos grupos o matrices son homotéticos)

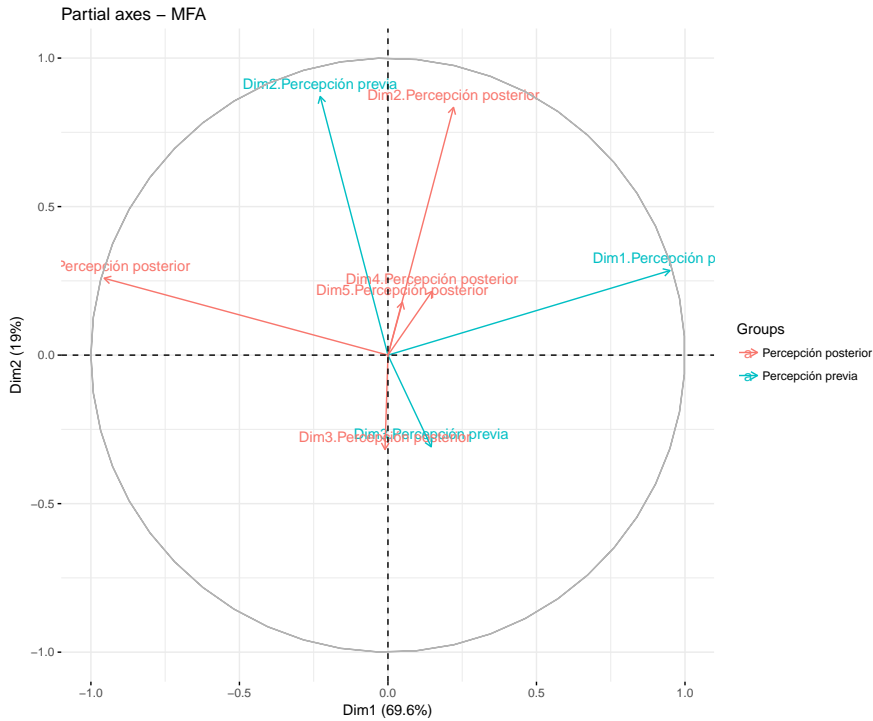
Los coeficientes Rv se pueden observar en la siguiente tabla:

	Percepción previa	Percepción posterior	MFA
Percepción previa	1.0000000	0.7315340	0.9320054
Percepción posterior	0.7315340	1.0000000	0.9289101
MFA	0.9320054	0.9289101	1.0000000

3.7. Representación superpuesta



3.8. Ejes parciales



? ? ? ? ?