

1. ACP **Muestra:** base de los que han constestado en castellano en el estudio “Shangai” (401)

Se ha realizado un ACP tomando como activas las variables marcadas en rojo/bold y como suplementarias las marcadas en normal. Se indican las categóricas con (categ):

[1] "Satisfacción(categ)" **"conf.general"** **"conf.cabina"** **"espai.cabin"** **"conf.cama"**

**[6] "conf.seat"** **"climat"** **"limp.espcom"** **"limp.cabin"** **"limp.aseos"**

[11] "trato.LL" "trato.TR" "aspect.gener" "idiomas" "AMI2 (categ) “

[16] "mean.SAT"

**1.1.** Los valores propios se listan en la Tabla 1. ¿Cuántos ejes tienen una inercia superior a la inercia media? **1 pto**

	eigenvalue	percentage of variance	cumulative percentage of variance
comp 1	4.86	60.69	60.69
comp 2	1.03	12.84	73.53
comp 3	0.54	6.74	80.27
comp 4	0.46	5.70	85.97
comp 5	0.40	5.04	91.01
comp 6	0.32	4.04	95.06
comp 7	0.21	2.65	97.71
comp 8	0.18	2.29	100.00

Tabla 1. Los valores propios

Se trata de un ACP estandarizado, dado que la suma de valores propios es igual al número de variables.

En un ACP estandarizado, la inercia media por eje es igual a 1. En este análisis, dos ejes tienen una inercia (igual al valor propio) superior a la inercia media.

Los principales resultados del ACP se presentan en la Figura 1. A partir de dicha figura, contesten a las siguientes preguntas

**1.2.** A partir de la representación de los individuos y de las variables sobre el primer plano factorial, qué se puede decir de los scores de los individuos 768 y 611? qué se puede decir de los scores de los individuos 768 y 643? **1 pto**

Los individuos 768 y 611 han valorado de forma similar los diferentes aspectos del viaje, ligeramente por debajo de la media, aunque 611 está, en media, ligeramente más satisfecho.

Los individuos 768 y 643 tienen, en media, una satisfacción similar (coordinada sobre el primer eje similar) pero 768 puntúa más alto los aspectos de limpieza que los aspectos de confort y espacio, mientras que 643 expresa más satisfacción relativamente a los aspectos de confort-espacio que a los aspectos de limpieza.

1.3. Hagan una interpretación de los dos primeros ejes a partir de los gráficos y de la descripción de los ejes **3 ptos**

El primer eje es un eje de “tamaño de satisfacción”. Está correlacionado positivamente, y de forma significativa, con todas las variables de satisfacción activas. Este eje opone los individuos satisfechos (a la derecha) a los individuos insatisfechos (a la izquierda).

Las variables suplementarias (cuantitativas y categóricas) ratifican esta interpretación.

En efecto este primer eje está correlacionado positivamente, y de forma significativa, con todas las variables cuantitativas suplementarias. Se puede notar la alta correlación entre la variable “mean.SAT”, media aritmética de las variables de satisfacción, y el primer eje.

Las categorías “Muy satisfecho” y “Satisfecho” se posicionan en la parte derecha del eje, mientras las categorías “Poco satisfecho” y “Nada Satisfecho” se posicionan en la parte izquierda del eje. Además, las categorías de la variable “Satisfacción” se posicionan sobre este eje de menor a mayor satisfacción, respetando así su orden.

La categoría “Ami2-SI” (recomendaría a un amigo) se posiciona en la parte derecha del eje, mientras Ami2-No” (no recomendaría a un amigo) se posiciona en la parte izquierda del eje.

El segundo eje está correlacionado positivamente con las variables activas relativas a la limpieza y negativamente a las variables activas relativas al confort y espacio. Este eje diferencia a los individuos entre los que son más satisfechos por la limpieza (posicionados en la parte positiva del eje) y los que son más satisfechos por el confort-espacio (posicionados en la parte negativa del eje).

La forma de la nube de individuos muestra que hay más variabilidad sobre el segundo eje en el caso de los individuos poco satisfechos (posicionados a la izquierda del primer eje) que en el caso de los individuos muy satisfechos (posicionados a la derecha del primer eje).

Las variables categóricas suplementarias, con un grado de asociación muy baja con el segundo eje ( $R^2$  muy débiles), ratifican que este eje no mide la intensidad de la satisfacción sino su “forma”.

## 2. Clasificación de los individuos

Después del ACP, se han clasificado a los individuos a partir de 2 ejes. Los principales resultados se presentan en la Figura 2.

### 2.1. Justifiquen la elección del número de ejes **1 pto**

Se escogen los 2 ejes cuya inercia es superior a la inercia media (ver apartado anterior)

### 2.2. Indiquen cuál(es) es(son) la(s) variable(s) categórica(s) que caracterizan la partición. Comenten el test efectuado. **1 pto**

La partición viene caracterizada por las variables categóricas suplementarias “Satisfacción” y AMI2

Para cada variable categórica, se construye la tabla cruzando la partición y la variable categórica. Se efectúa el clásico test de independencia para contrastar las siguientes hipótesis:

H0): en la población, existe independencia entre la variable categórica examinada y la partición

H1) en la población, no hay independencia entre la variable categórica examinada y la partición.

Se calcula el chi-2 para calcular el  $p$ -value.

En el caso de de “Satisfacción” y AMI2, se rechaza la independencia entre la partición y la variable con un  $p$ -value muy pequeño. Por lo tanto existe una relación entre la partición y el nivel de satisfacción

### 2.3. Indiquen cómo se interpreta la caracterización de la clase 1 reproducida en la figura 2. Detallen los tests efectuados. **3 ptos**

Se caracteriza a la clase 1 por a) Las categorías de las variables categóricas suplementarias b) Las variables cuantitativas activas y suplementarias

#### a. Caracterización por las categorías

Todas las categorías listadas en los resultados correspondientes a una clase caracterizan dicha clase, aunque unas por estar sobrerrepresentadas en la clase ( $\text{Mod/Cla} > \text{Global}$ ) y otras por estar infrarepresentadas en la clase ( $\text{Mod/Cla} < \text{Global}$ ). Las no significativas no se indican.

El test efectuado para una categoría dada compara estas dos proporciones ( $\text{Mod/Cla}$  y  $\text{Global}$ ):

H0) en la población,  $\text{Mod/Cla} = \text{Global}$

H1) en la población,  $\text{Mod/Cla} \neq \text{Global}$

Menor es el  $p$ -value, mayor es la sobre o infra representación. Que una categoría esté sobre representada/infra-representada se ve porque  $\text{Mod/Cla} > \text{Global}$ /  $\text{Mod/Cla} < \text{Global}$

#### b. Caracterización por las variables cuantitativas

Las variables cuantitativas listadas caracterizan a la clase 1. Todas presentan en la clase 1 una media superior a la media en toda la muestra. El test efectuado corresponde a la comparación de la media en la clase con la media global.

H0) media poblacional de la variable examinada en la clase = media global poblacional de la variable examinada

H0) media poblacional de la variable examinada en la clase  $\neq$  media global poblacional de la variable examinada

Si el  $p$ -value es inferior a 0.05, se rechaza H0.



<pre> &gt; \$ \$Dim.1 \$Dim.1\$quanti       correlation p.value mean.SAT      0.8534377      0 conf.general  0.8408318      0 conf.cabina   0.8322009      0 limp.cabin    0.7856992      0 limp.espcom   0.7731482      0 conf.seat     0.7644485      0 conf.cama     0.7603838      0 espai.cabin   0.7581301      0 limp.aseos    0.7093787      0 aspect.gener  0.6197342      0 climat        0.5761579      0 trato.TR      0.4792135      0 trato.LL      0.4641923      0 idiomas       0.4201582      0 </pre>	<pre> \$Dim.1\$quali       R2      p.value Satisfacción 0.2480229 9.072725e-88 AMI2         0.1062672 1.294219e-37  \$Dim.1\$category       Estimate      p.value Muy satisfecho 3.0227371 4.258646e-61 SI-ami         0.7186964 1.294219e-37 Satisfecho     1.3024428 7.667881e-27 Satis-MISS     0.8707582 1.332391e-04 PocoSatisfecho -1.6005845 2.735001e-14 NadaSatisfecho -3.3341086 2.034497e-14 NO-ami         -0.7186964 1.294219e-37 </pre>
<pre> \$Dim.2 \$Dim.2\$quanti       correlation      p.value limp.espcom  0.50284561 0.000000e+00 limp.cabin   0.45868989 0.000000e+00 limp.aseos   0.41932859 0.000000e+00 aspect.gener 0.14752223 1.405609e-08 trato.TR     0.09111186 4.800950e-04 idiomas      0.09033047 5.370644e-04 trato.LL     0.08089894 1.942604e-03 conf.cama    -0.14924827 9.477455e-09 conf.seat    -0.21356086 1.424928e-16 conf.general -0.22221616 7.568527e-18 conf.cabina  -0.35081611 1.122580e-43 espai.cabin  -0.38395934 1.159413e-52 </pre>	<pre> \$Dim.2\$quali       R2      p.value Satisfacción 0.016371958 0.0002068783 AMI2         0.007003915 0.0013453334  \$Dim.2\$category       Estimate      p.value NO-ami      0.08485282 0.001345333 Satis-MISS  0.33215444 0.005537684 Satisfecho  -0.15270124 0.014835010 Muy satisfecho -0.24995238 0.006630095 SI-ami      -0.08485282 0.001345333 </pre>

Figura 1. Principales resultados gráficos y numéricos del ACP

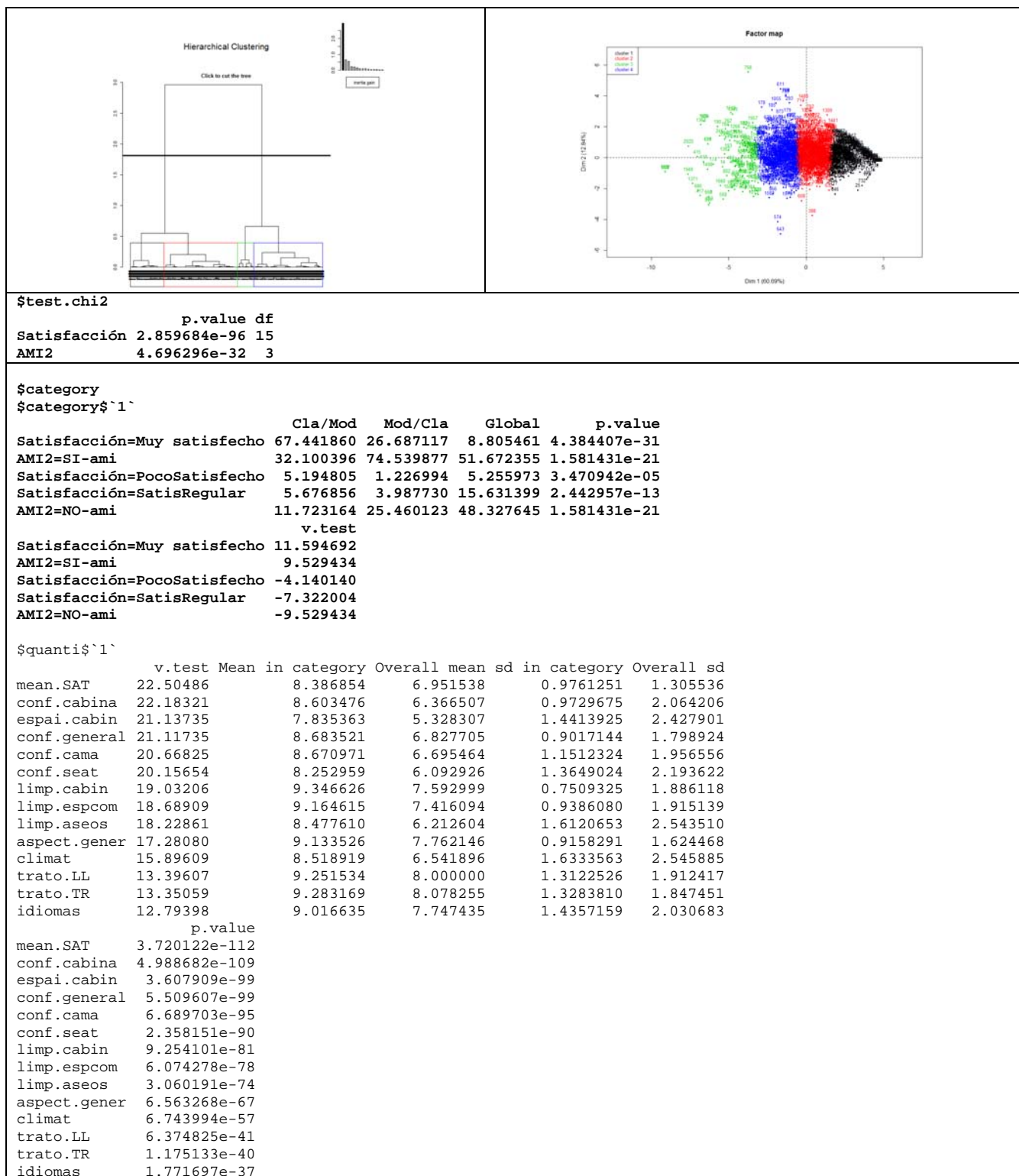


Figura 2. Principales resultados gráficos y numéricos de la clasificación automática