

# T4. Taller encuestas. Análisis factorial (PCA)\*

VIII Jornadas R Albacete 2016\*\*

Álvaro Hernández Vicente, Elvira Ferre Jaén, Antonio José Perán Orcajada, Ana Belén  
Marín Valverde, Antonio Maurandi López\*\*\*

17 de noviembre de 2016

## Índice

1. Lectura de datos	1
2. Análisis de Componentes Principales	1
Referencias y bibliografía	3

## 1. Lectura de datos

Se leen los datos `saeraq.Rdata`, que provienen del fichero `raq.dat` utilizado en (A. Field, Miles, and Field 2012).

```
load( "saeraq.RData" )
```

## 2. Análisis de Componentes Principales

```
indicesPreg <- grep( "^Q", colnames( df ) )  
df[ , indicesPreg ] <- as.data.frame( lapply( df[ , indicesPreg ], as.numeric ) )
```

```
# library( FactoMineR )  
pca <- PCA( df, quali.sup = c( 1:3, 5:6 ), quanti.sup = 4,  
           scale.unit = TRUE, graph = FALSE )
```

```
# library( factoextra )  
fviz_screplot( pca, addlabels = TRUE ) +  
  ggtitle("") +  
  scale_x_discrete( name = "Componentes principales" ) +  
  scale_y_continuous( name = "Autovalor y varianza explicada" ) +  
  theme_minimal()
```

---

\* *doc:T4\_analisisPCA.Rmd*

\*\* <http://r-es.org/8jornadasR/>

\*\*\* Servicio de Apoyo Estadístico; alvarohv@um.es, elvira@um.es, antoniojose.peran@um.es, anabelen.marin4@um.es, amaurandi@um.es

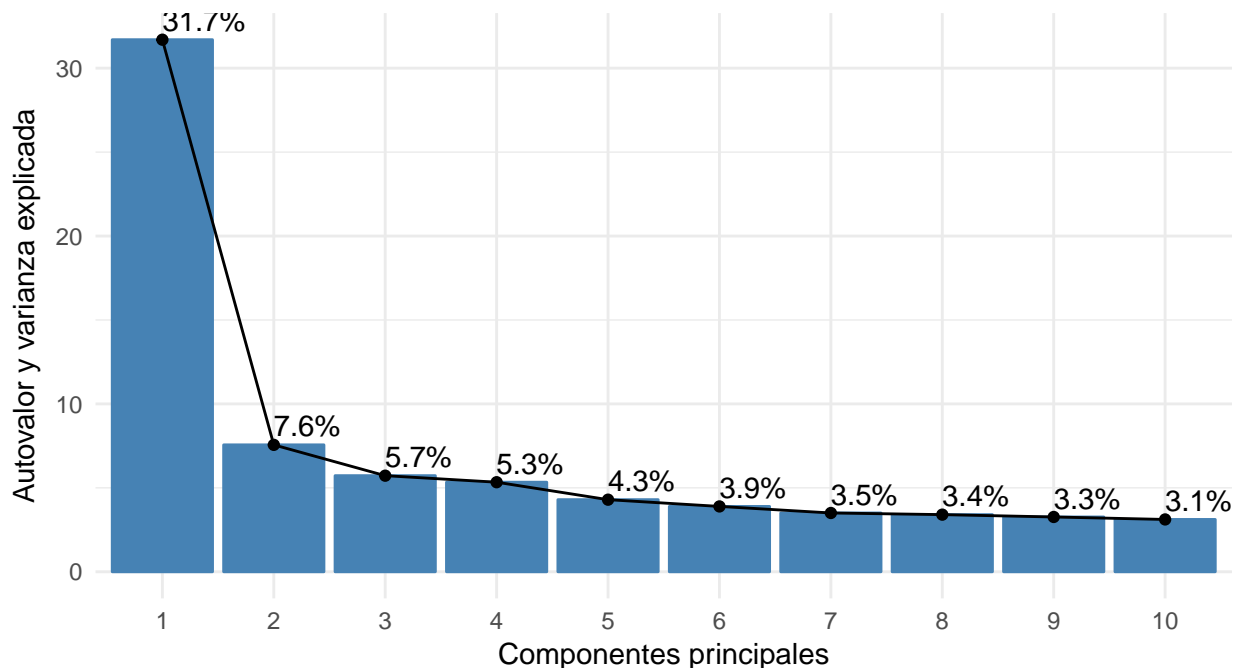


Figura 1: PCA. Varianza explicada por cada componente principal

```
fviz_pca_var( pca, axes = c( 1, 2 ), geom = "text", alpha.var = "contrib",
              labelsize = 2 ) +
ggtitle("") +
theme_minimal()
```

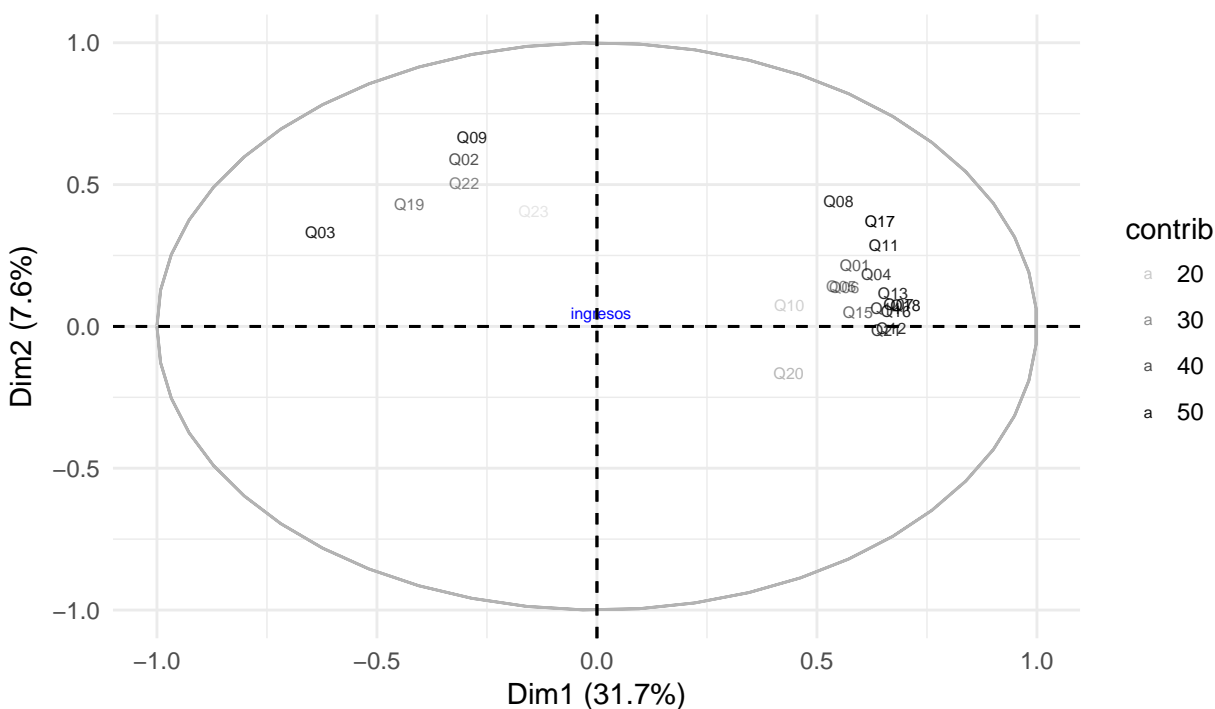


Figura 2: PCA. Gráfico de las variables con las etiquetas utilizadas para las proteínas

```
fviz_pca_biplot( pca, pointsize = 1.5, labelsize = 2.5, #alpha.var="contrib",
  habillage = df$sexo, addEllipses = TRUE,
  ellipse.level = 0.95, select.var = list( contrib = 25 ) ) +
  ggtitle("") +
  theme_minimal()
```

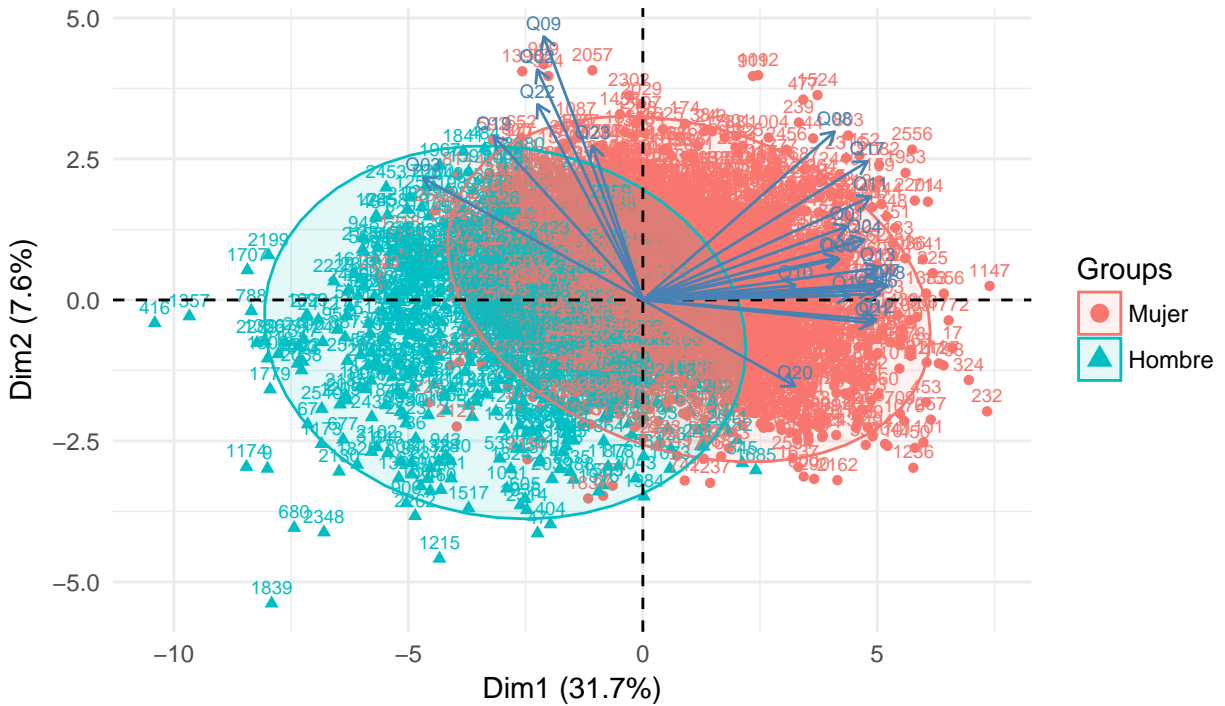


Figura 3: PCA Biplot. Gráfico de variables e individuos según las dos primeras componentes principales. Elipses de confianza al 95 % y color según fracción

## Referencias y bibliografía

Field, Andy, Jeremy Miles, and Zoe Field. 2012. *Discovering Statistics Using R*. 1st edition. Sage Publications Ltd.