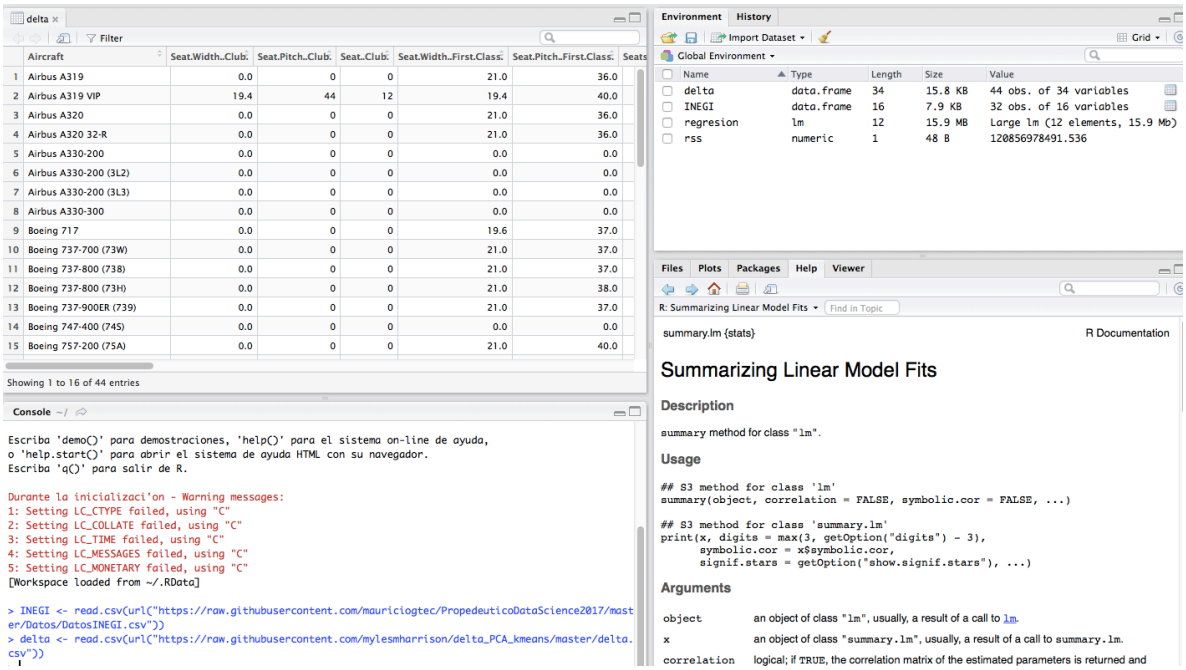
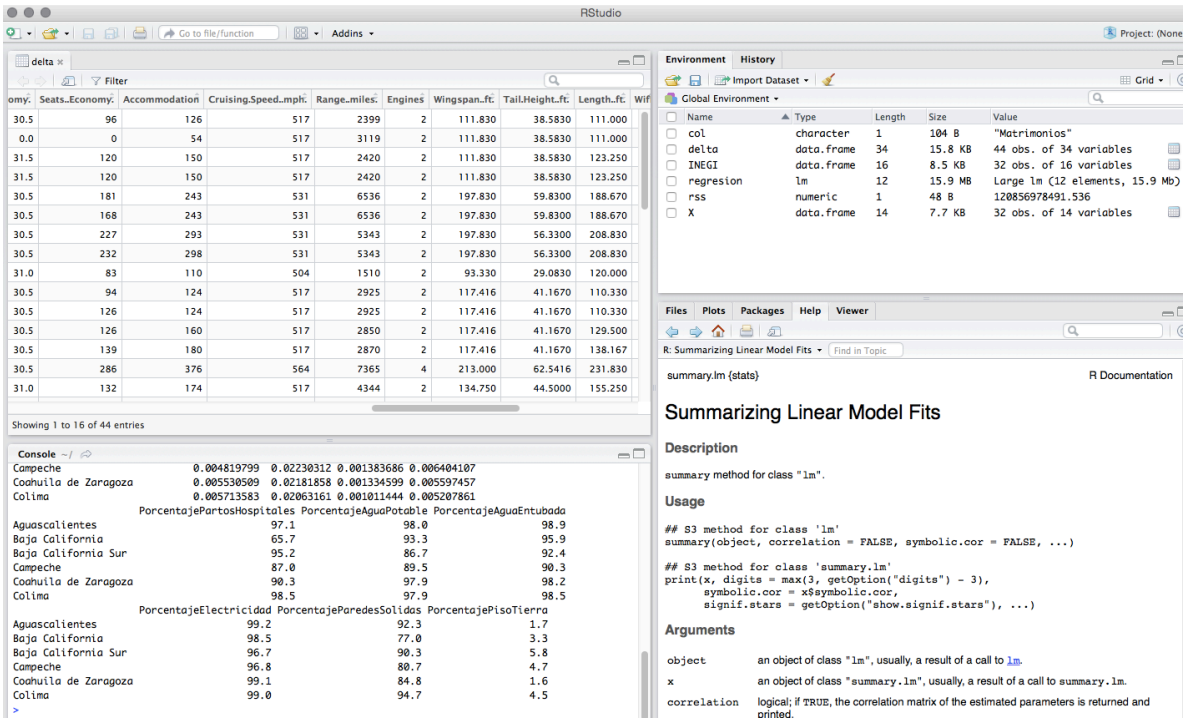
Para la realización del ejercicio de componentes principales (PCA) vamos a usar dos bases de datos. La primera es la que se usó en clases bajo el nombre de INEGI. La segunda corresponde a la variable “delta” (base de datos conseguida para este ejercicio), que incluye información (número de asientos totales, de primera y clase turista, velocidad crucero del avión, rango máximo de la unidad, etcétera) de todos los aviones de la compañía DELTA. El objetivo principal es hacer un ejercicio de componentes principales con la paquetería FactoMineR.

#Se cargan ambas bases de datos desde las respectivas direcciones en línea en las variables INEGI y delta.



Se va a hacer el tratamiento hecho durante clases para hacer ciertas variables per cápita y sea más fácil su manipulación. Además se creará una matriz con el nombre de “X” que será la utilizada en el análisis PCA.



Ahora se instalará la paquetería FactoMineR:

Para realizar el análisis PCA se requiere llamar la paquetería instalada:

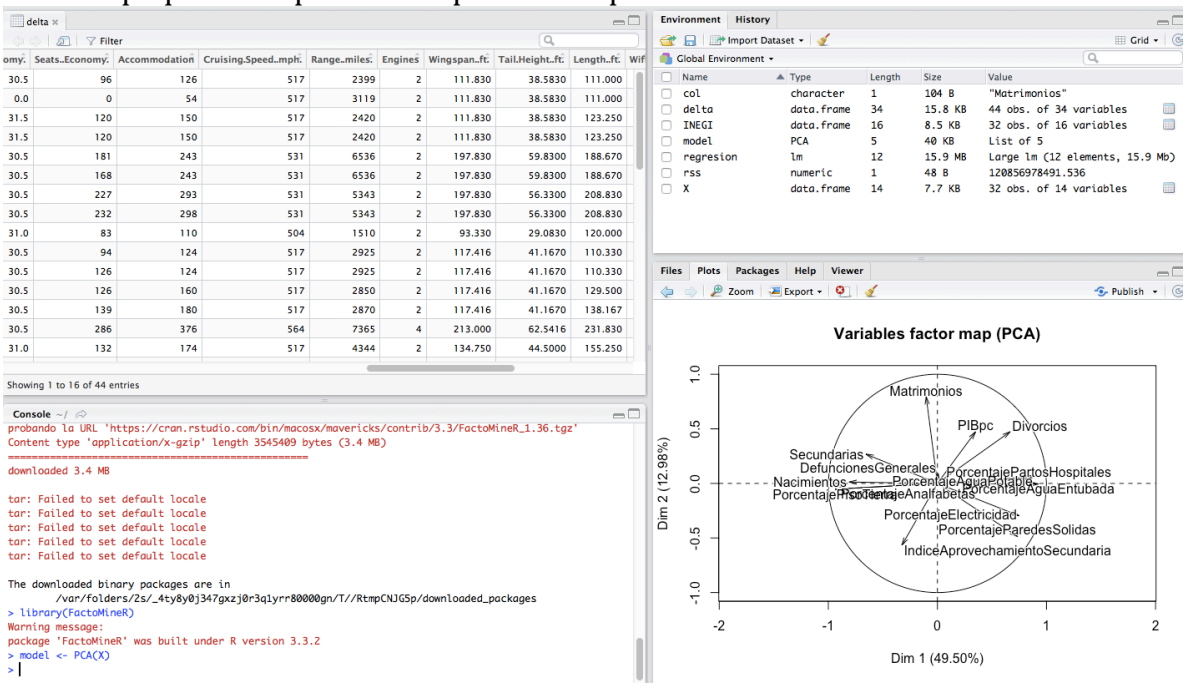
```
Console ~/

probando la URL 'https://cran.rstudio.com/bin/macosx/mavericks/contrib/3.3/FactoMineR_1.36.tgz'
Content type 'application/x-gzip' length 3545409 bytes (3.4 MB)
==================================================
downloaded 3.4 MB

tar: Failed to set default locale
tar: Failed to set default locale
tar: Failed to set default locale
tar: Failed to set default locale
tar: Failed to set default locale

The downloaded binary packages are in
        /var/folders/2s/_4ty8y0j347gxzj0r3q1yrr80000gn/T//RtmpCNJG5p/downloaded_packages
> library(FactoMineR)
Warning message:
package 'FactoMineR' was built under R version 3.3.2
>
```

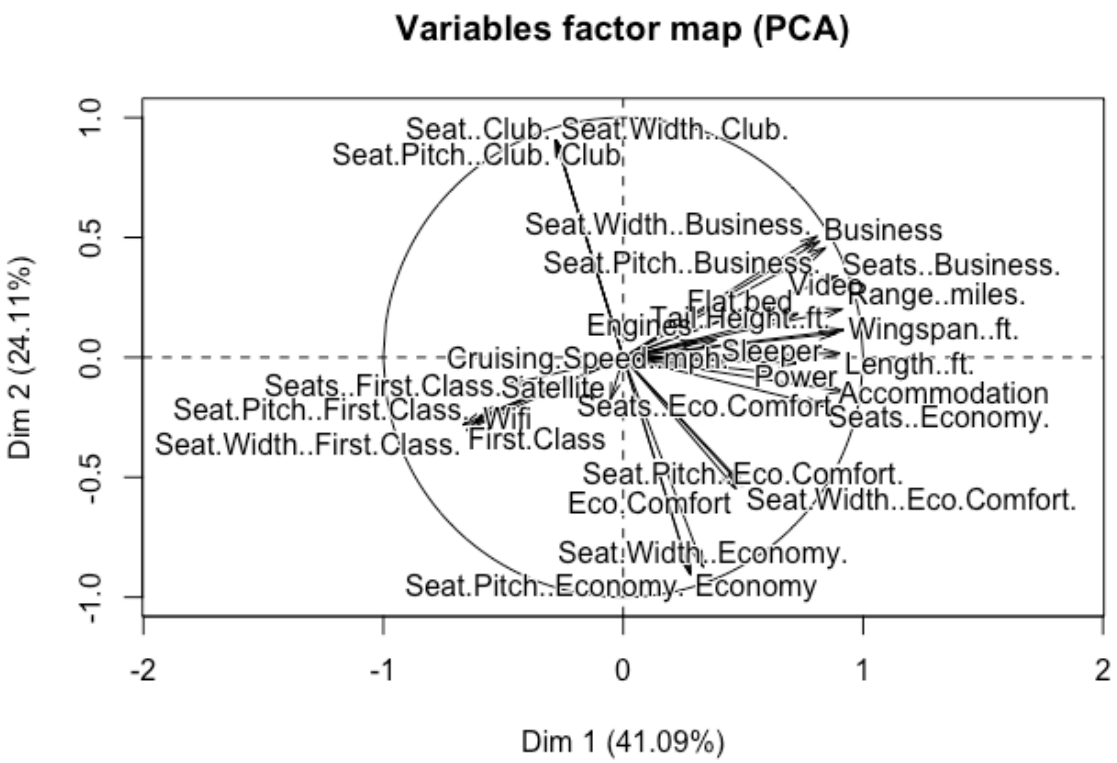
Con la paquetería operando se procede a aplicar el análisis PCA a la variable “X”:



En la parte inferior derecha podemos observar el resultado del análisis. La dimensión 1 (horizontal) captura el 49.5% de la variación total de todas las variables juntas, mientras que la dimensión 2 (vertical) explica el 12.98% de la variación total. A continuación llamaremos el mismo análisis PCA, pero en esta ocasión con la base de datos propia que llamamos “delta”. Dada la estructura de la base de datos “delta”, podemos notar que la primer columna se trata de todos los tipos de aviones que tiene la compañía. En ese sentido, se debe excluir la primer columna para hacer el análisis PCA. En el siguiente código se puede ver como se llama la función PCA sobre la base de datos “delta” excluyendo la primer columna.

```
116
117
118  #PCA
119  library(FactoMineR)
120  model <- PCA(X)
121
122  #PCA DELTA
123  summary(delta)
124
125  model2 <- PCA(delta[-1])
126  model2
126:7  (Untitled)                                    R Script
```

El resultado es el siguiente:



La dimensión 1 (horizontal) captura el 41.09% de la variación total de todas las variables juntas, mientras que la dimensión 2 (vertical) explica el 24.11% de la variación total. Con eso concluye el ejercicio de PCA con la base de datos de INEGI y la propia de “delta”.