

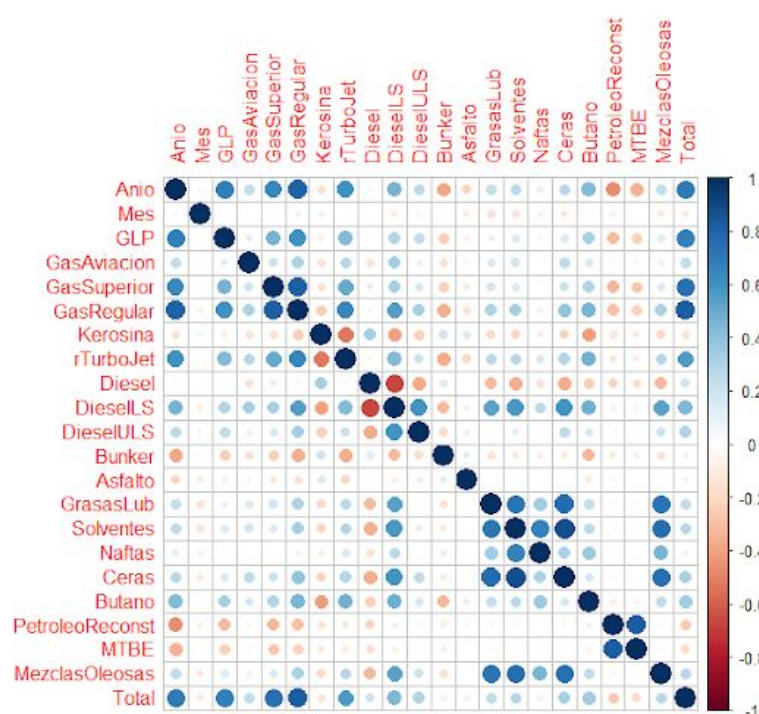
## Laboratorio 3

### Serie de tiempo

#### Análisis exploratorio

Primero se observó que hay muchas observaciones en las cuales, para diferentes variables, tenían un NA. Por ello, se sustituyeron los NA por el valor 0, para poder analizar todas estas variables cualitativas. Asimismo, habían un par de columnas las cuales estaban llenas de NA, se procedió con eliminar dichas variables ya que se consideró que no daban un aporte significativo al dataset.

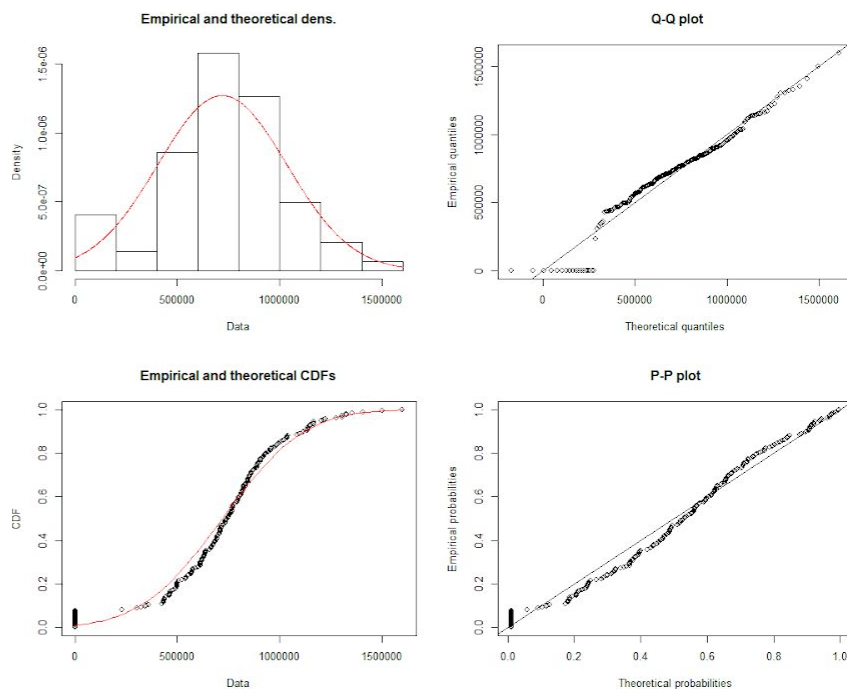
Seguido de esto, se realizó una matriz de correlación para determinar la correlación entre todas las variables.



**Figura 1.** Matriz de correlación.

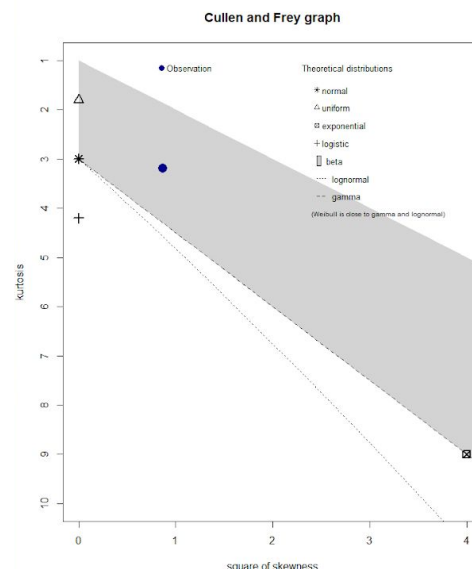
Se encontró que las variables GasSuperior y GasRegular se encuentran altamente correlacionadas con la variable Anio, asimismo se observó que la variable Ceras presenta un coeficiente de correlación alto con las variables GrasasLub y Solventes.

A continuación, se determinó si las variables estaban distribuidas normalmente. Para ello se utilizó el paquete fitdistrplus de R y la función descdist() para poder determinar el tipo de distribución que tenía la variable estudiada.



**Figura 2.** Prueba de normalidad para la variable Diésel.

En este caso, se determinó que la variable diesel si tiene una distribución normal. Luego, analizando la variable de Gas Regular, se tiene que no es una distribución normal, sino más bien beta.



**Figura 3.** Prueba para determinar tipo de distribución.

Por último, se determinó que la variable de gas super tiene una distribución normal.