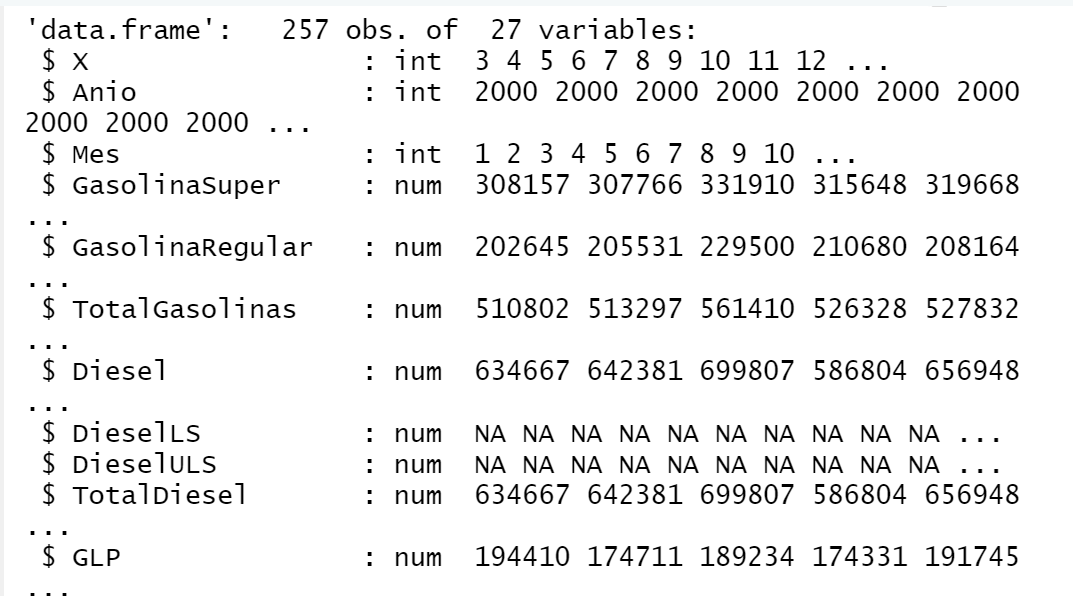
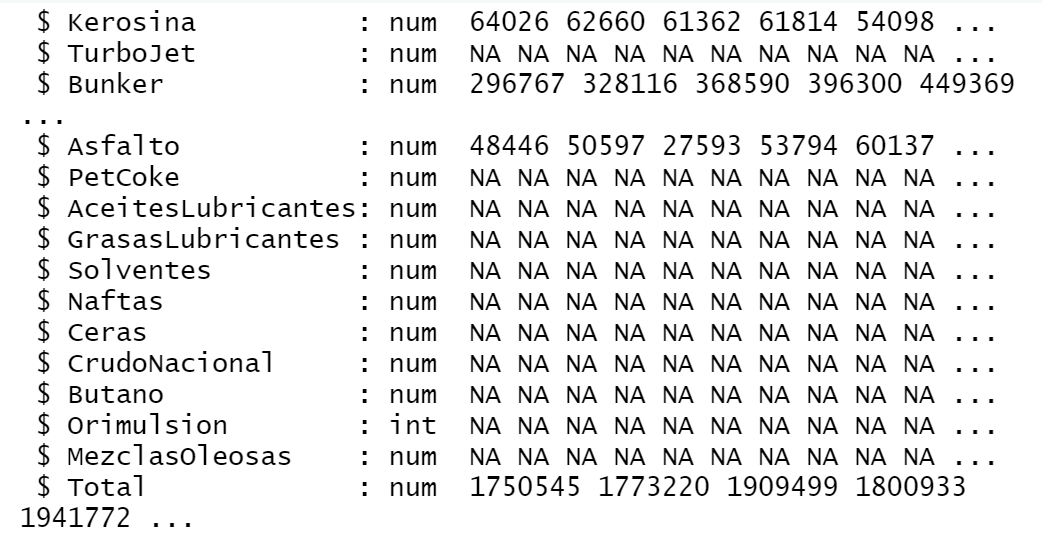
## Análisis Exploratorio

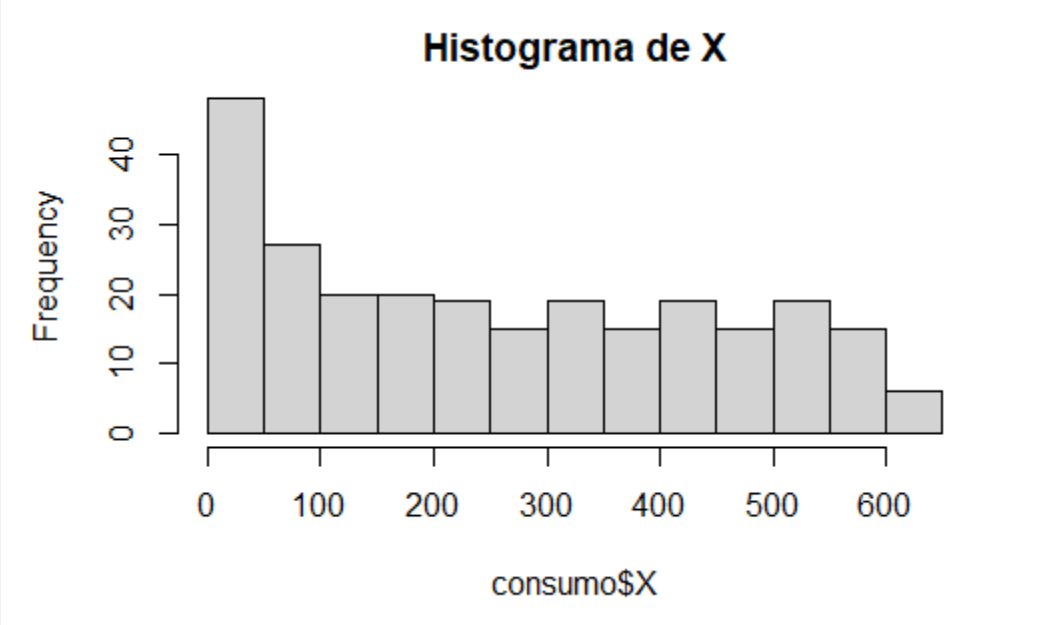
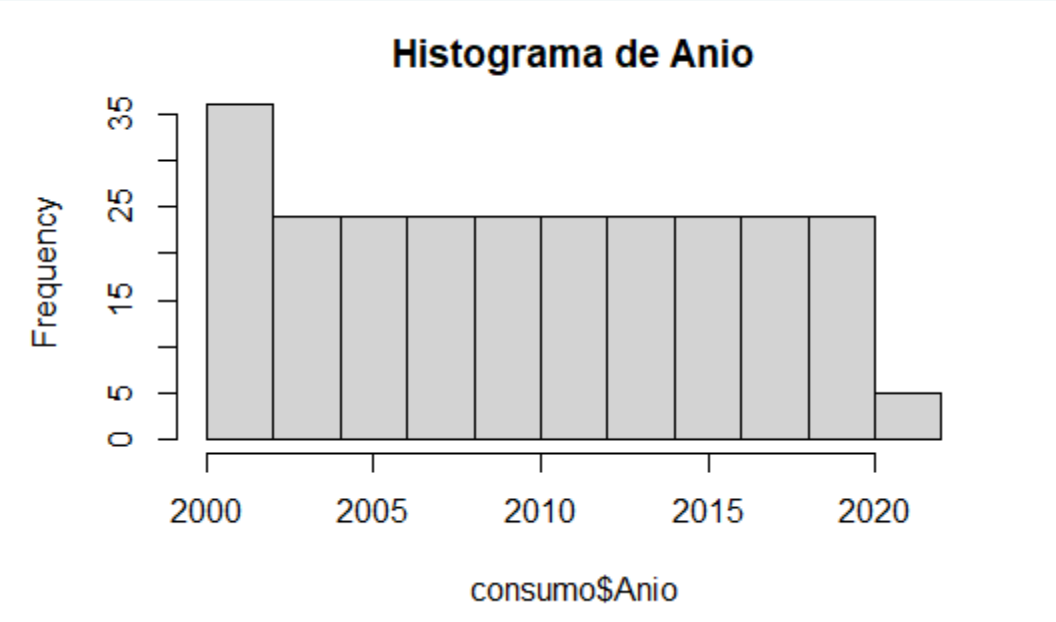
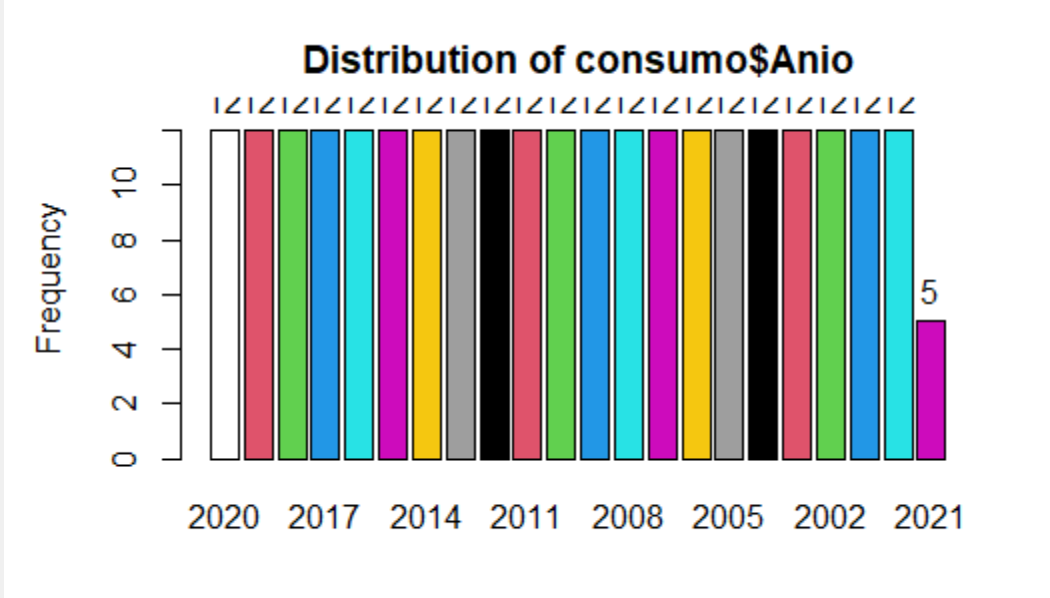
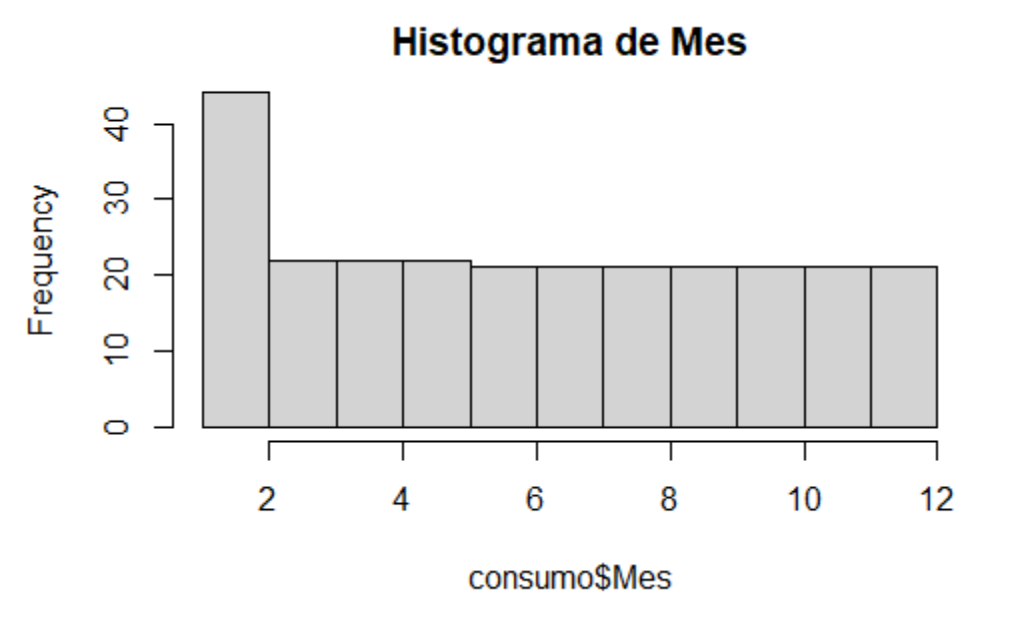
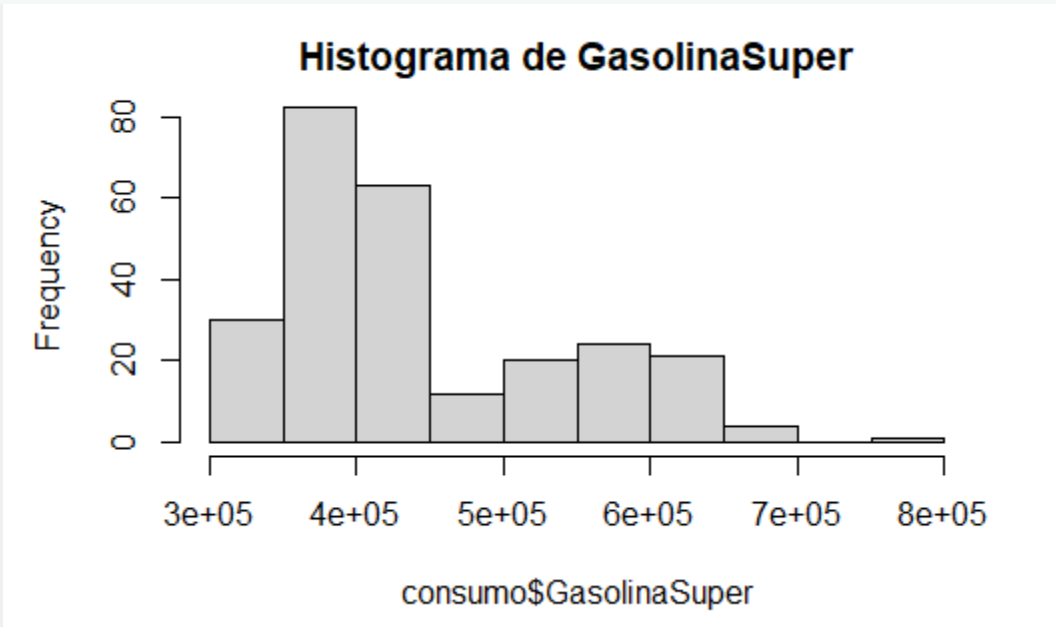
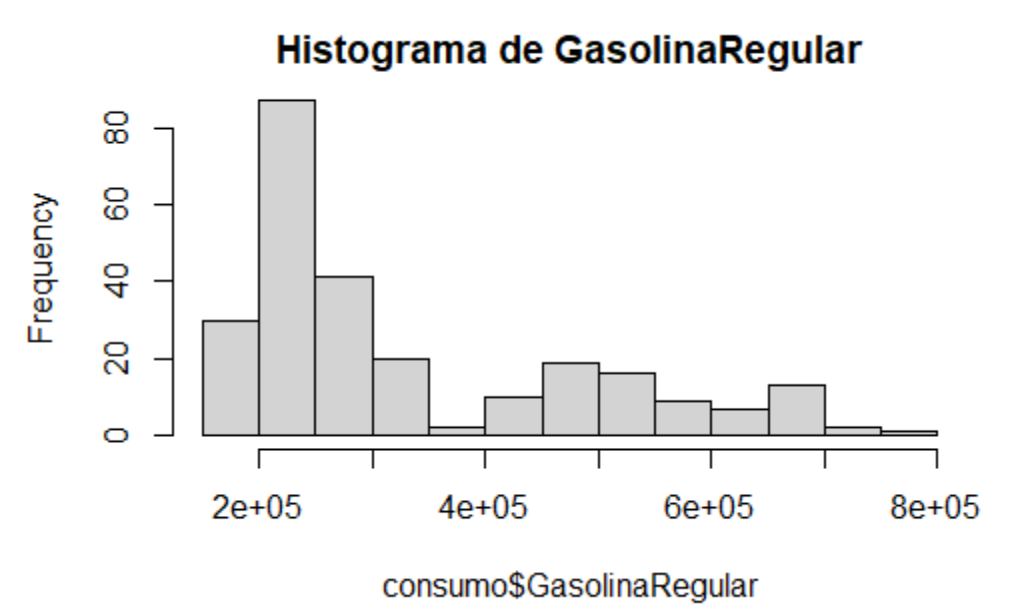
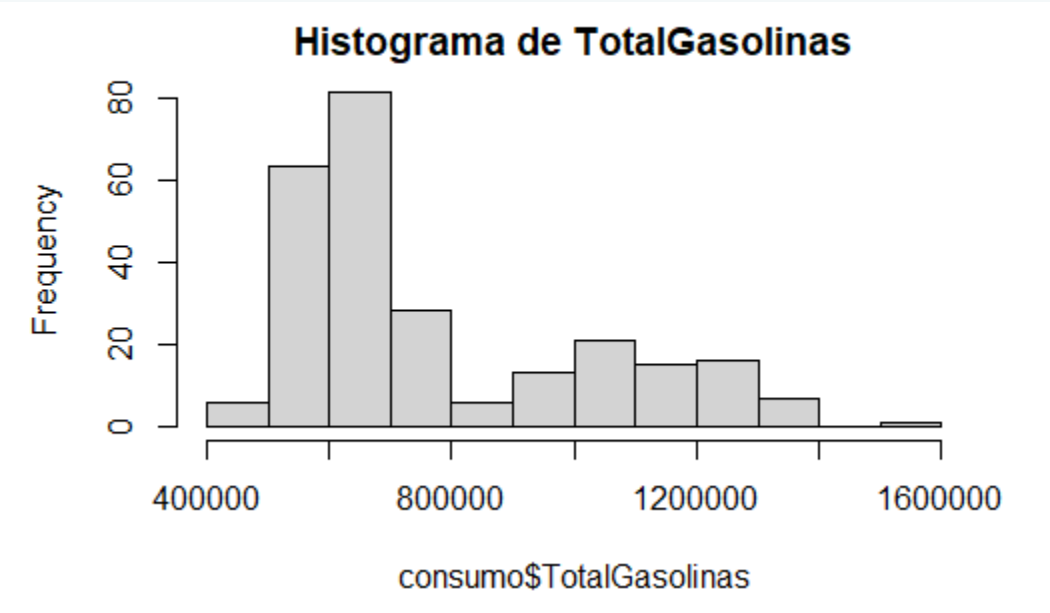
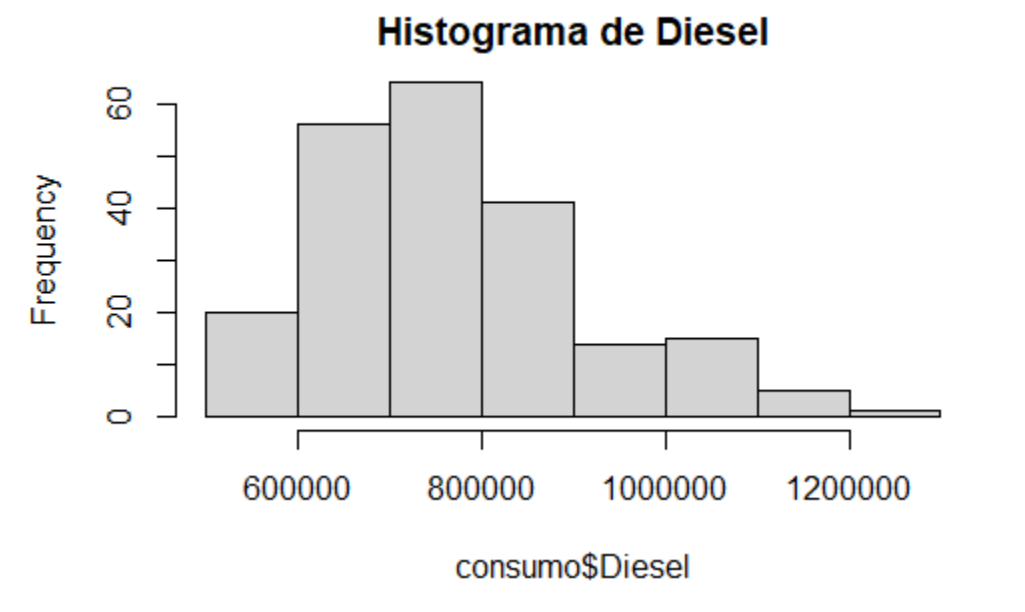
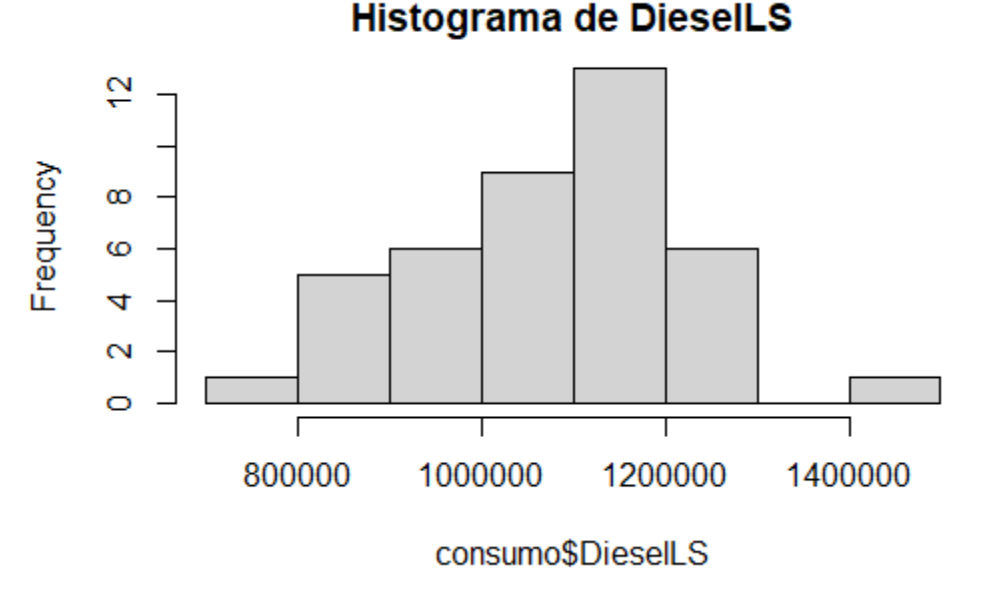
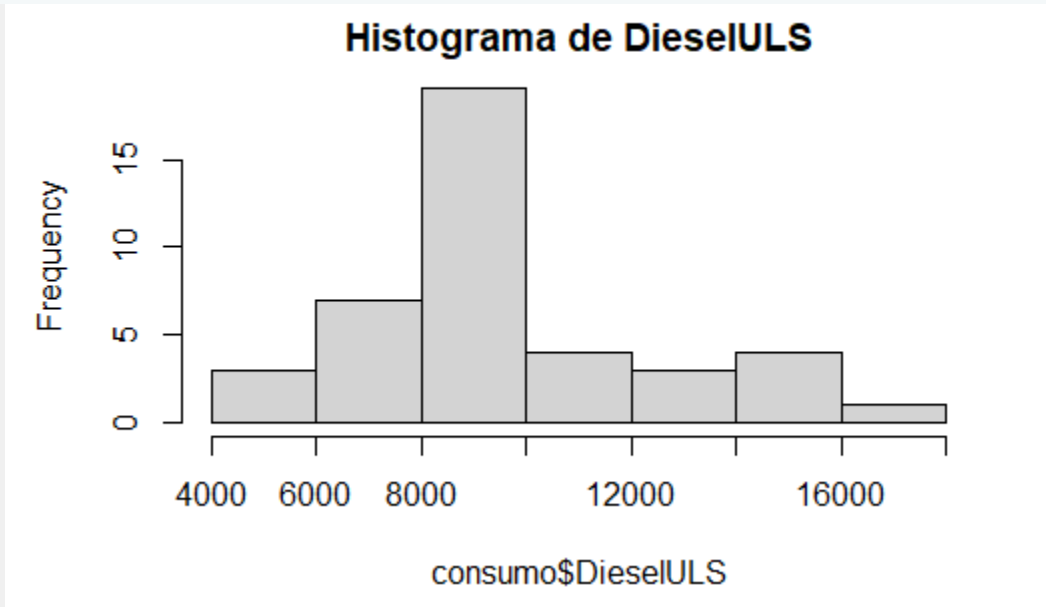
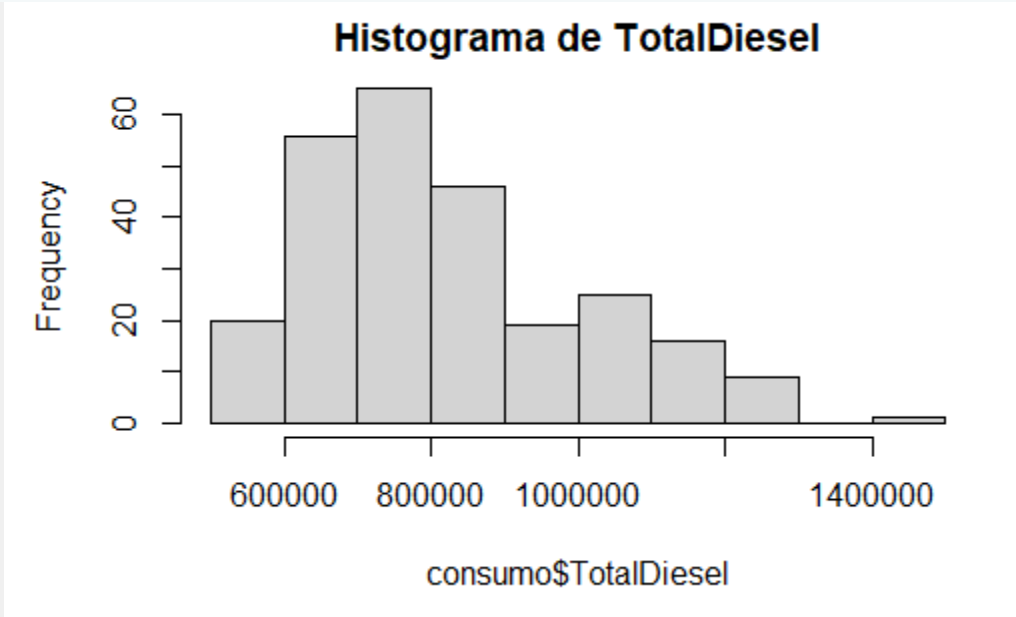
### Dataset Consumo

Tipo de variables y cuales son

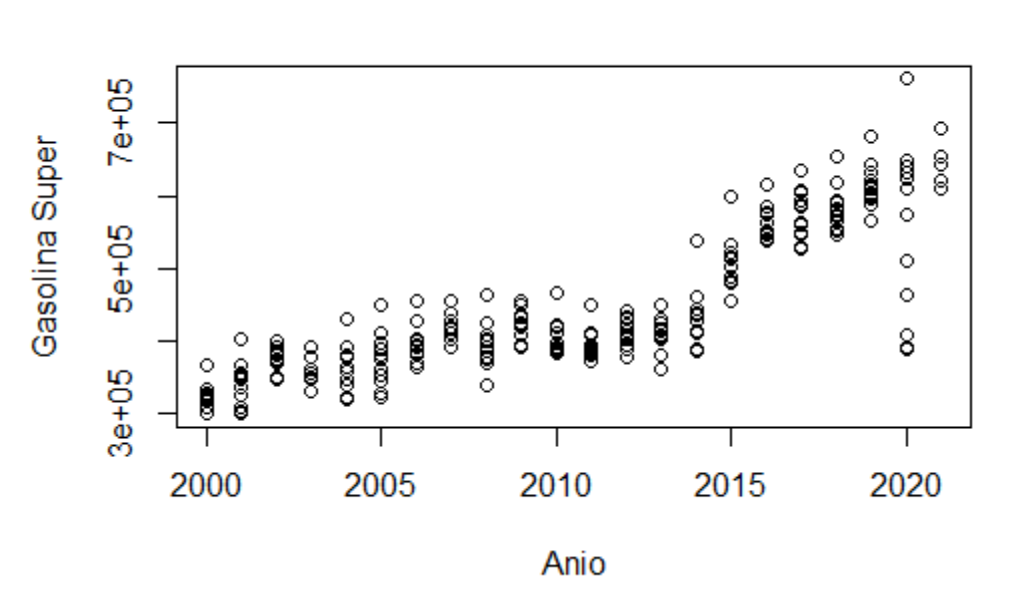




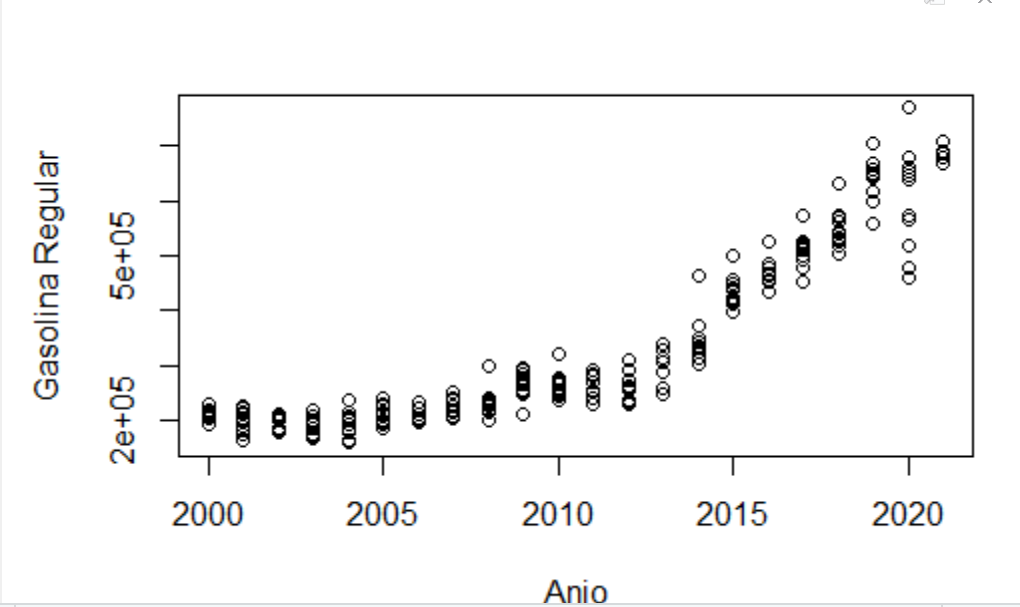
Podemos ver que todas las variables tratadas en las datasets son variables cuantitativas, por lo que haremos histogramas de las variables.

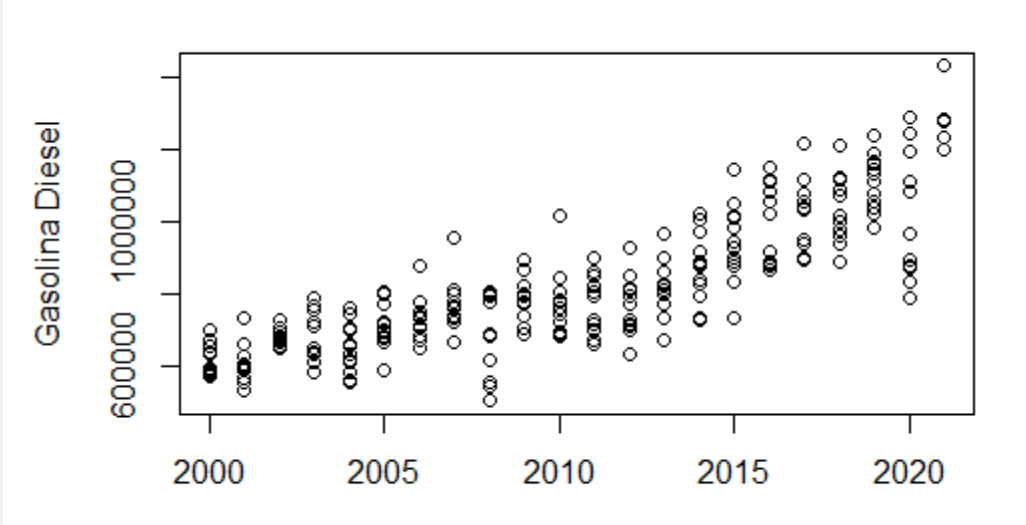
De estas graficas podemos observar que la mayoría no sigue una distribucion normal (las primeras 6 graficas) pero al momento con toparnos con la gasolina Diesel, y sus variaciones, tienen una tendencia a ser normales. Sin embargo, son tambien estas variables las que presentan cantidades de NA en sus filas, por lo que no son muy confiables estas graficas, pero si puede ser un indicador que en estar variable se comprotara diferente a los otros tipos de gasolina.



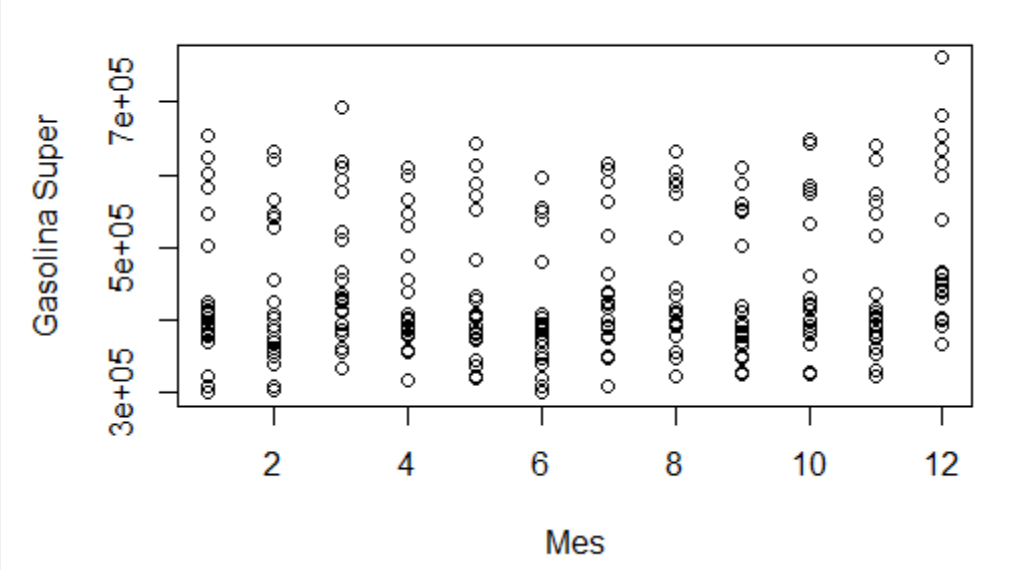
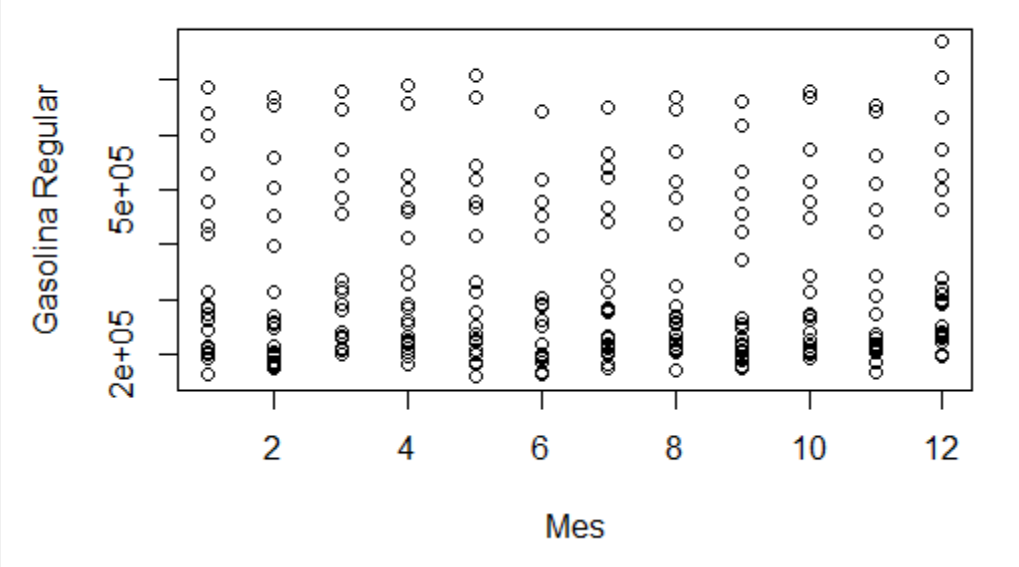
De este cruce de variables podemos ver que puede que exista un tipo de tendencia, que con los años, el consumo de gasolina super creció con los años.

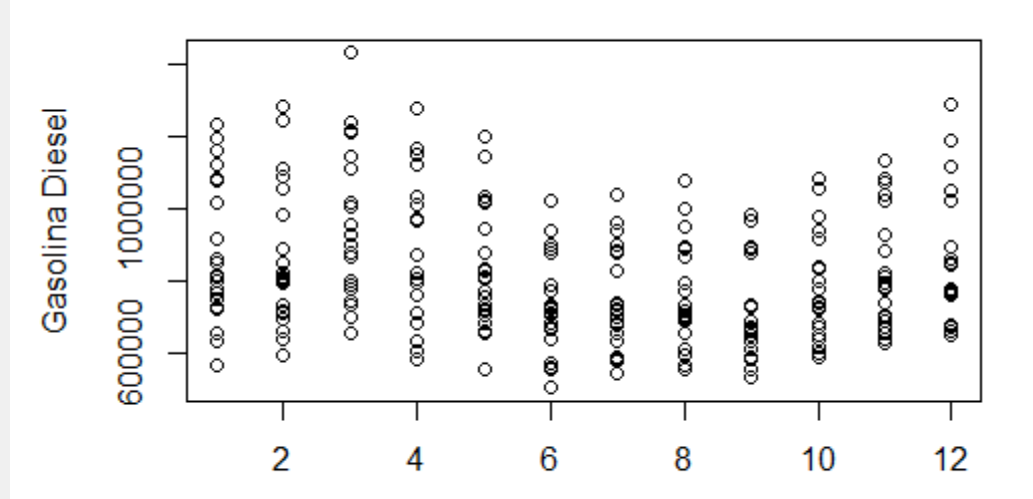


De este cruce de variables podemos ver que puede que exista un tipo de tendencia, que con los años, el consumo de gasolina regular creció con los años pero su crecimiento toma forma de una curva exponencial.



De este cruce de variables podemos ver que, como fue prediche, en comparacion con los otros dos tipos de gasolina, se ve un comportamiento más aleatorio.

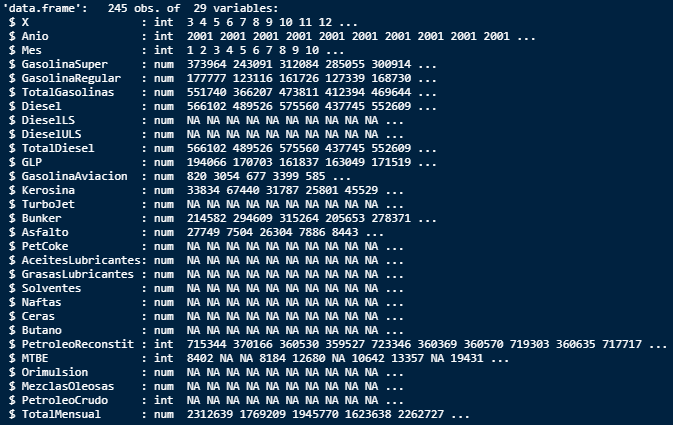
 



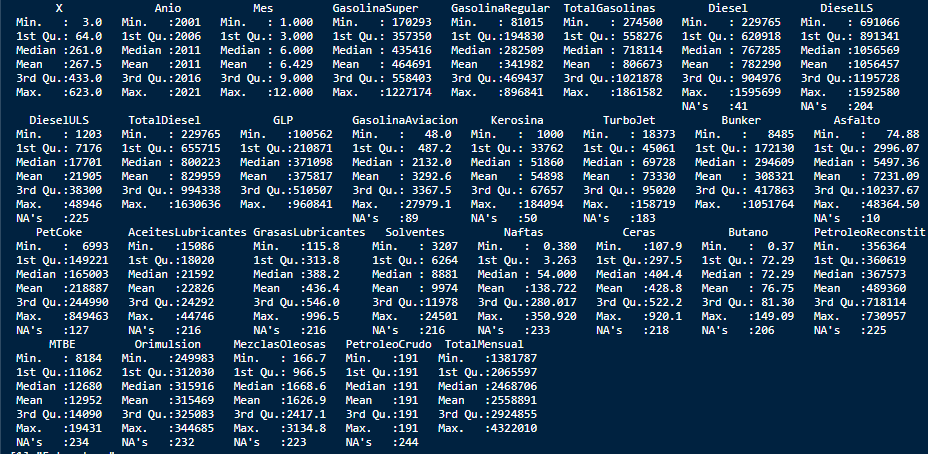
Ahora vemos que las variables se comportan de una manera aleatoria conforme a los meses del año. Ya sea cualquier tipo de las gasolinas tratadas.

Data set importaciones

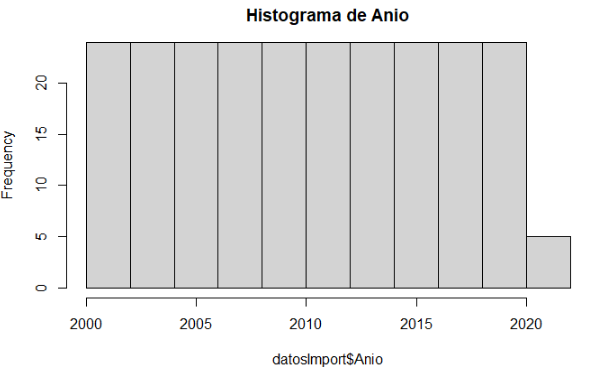
**Breve resumen del data set Importación de Combustibles**

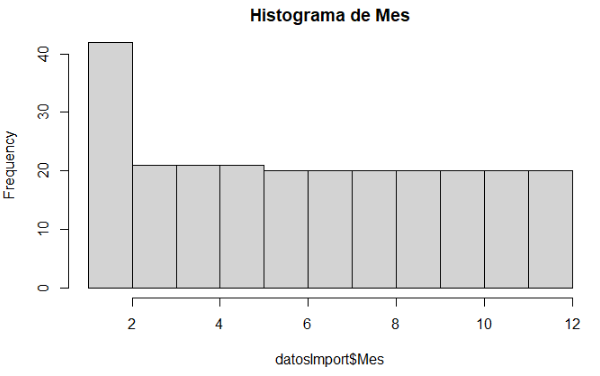


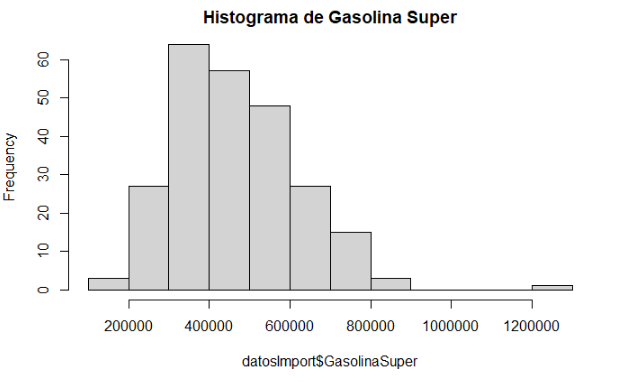
En la imagen de arriba, se puede observar la estructura de los datos, el tipo de cada uno y algunos de sus valores. También se sacaron sus dimensiones, y se tiene un data set de 29 columnas de 245

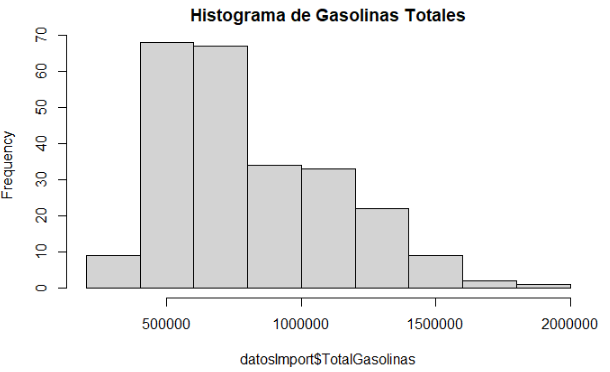
También se hizo un resumen estadístico de los datos.

Se hicieron histogramas con las variables cuantitativas a trabajar para la exploración de datos

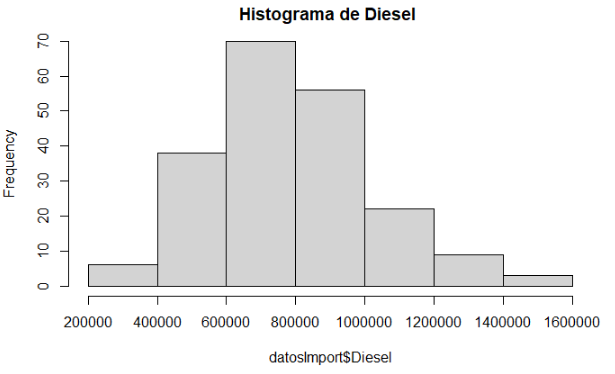




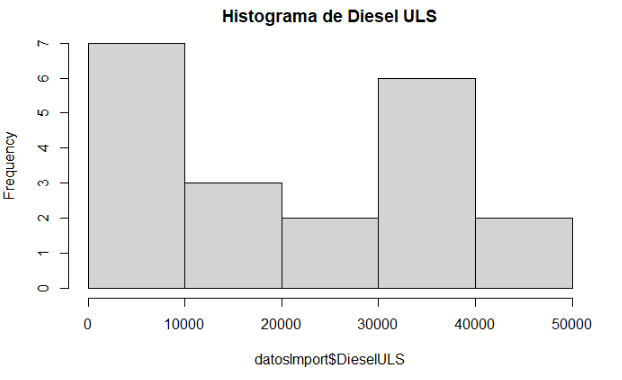
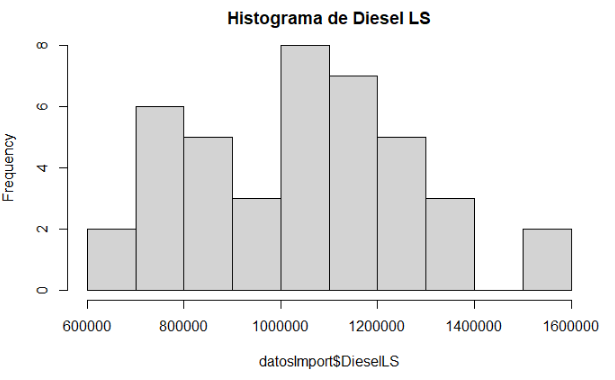
 Podemos observar que la variable “GasolinaSuper” sigue una distribución normal.

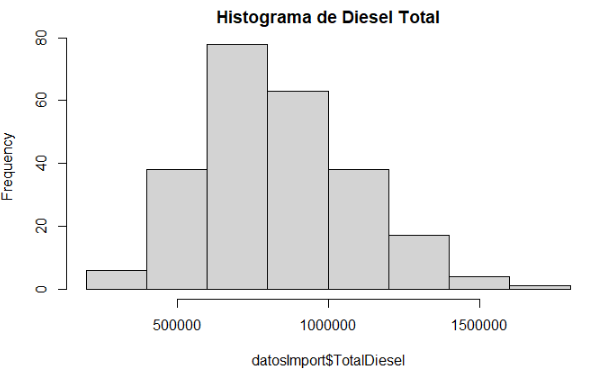


Gasolinas totales también aparenta tener una distribución normal.



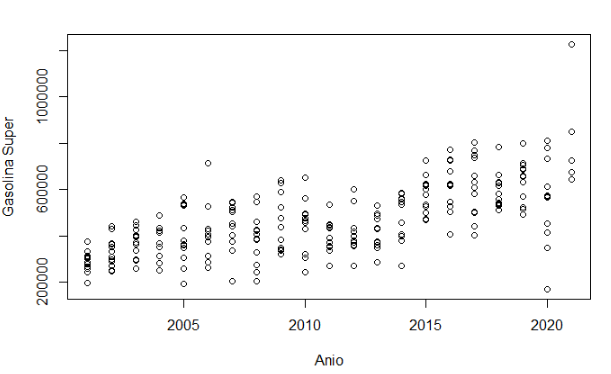
Diesel total también sigue una distribución normal

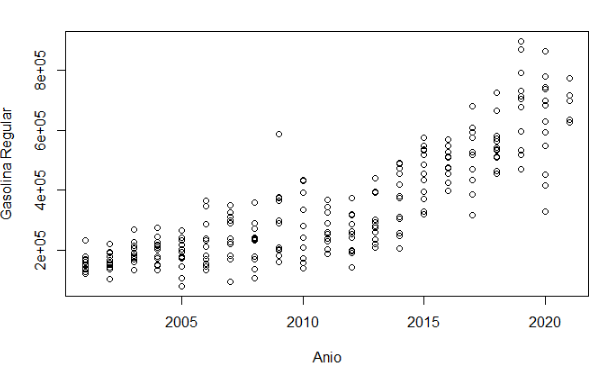


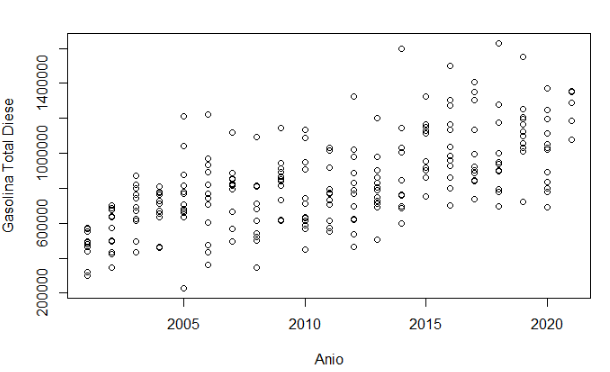


Diesel total también tiene una distribución normal.

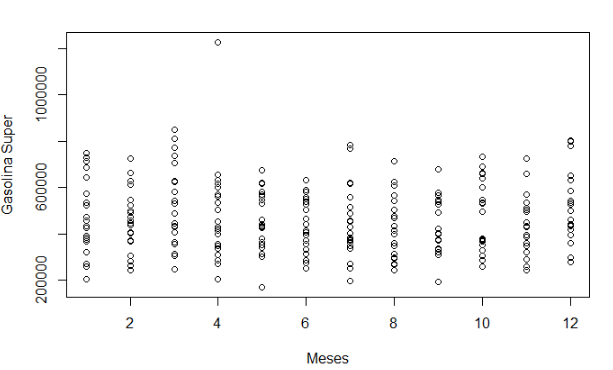
También se hizo un cruce de variables de los tipos de gasolinas con los años y los meses para ver si se encuentra algún patrón

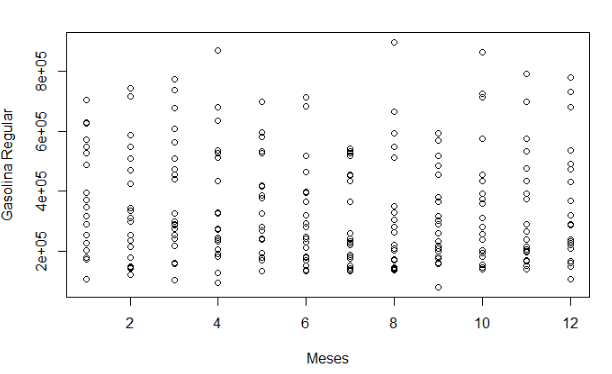


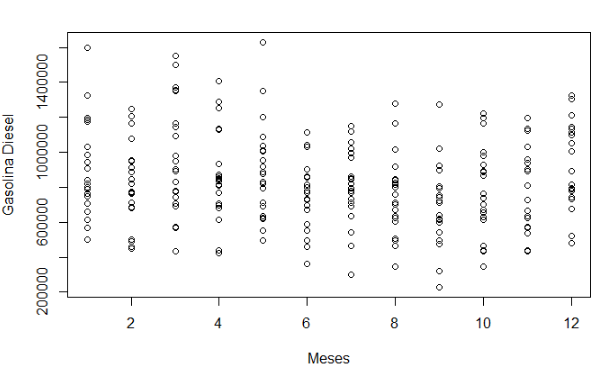




En las tres gráficas anteriores del cruce de variables de las importaciones de cualquier tipo de gasolina y el año, se puede observar que hay una tendencia a crecer. Esto implica que conforme pasa el tiempo, cada vez se importa más gasolina sin importar su tipo.







Por otro lado, al observar las gráficas del cruce de variables de cualquier tipo de gasolina con los meses, no se puede observar alguna tendencia. Esto implica que sin importar el mes, siempre va a haber una importación de combustible más o menos constante.

## Lineas de tiempo

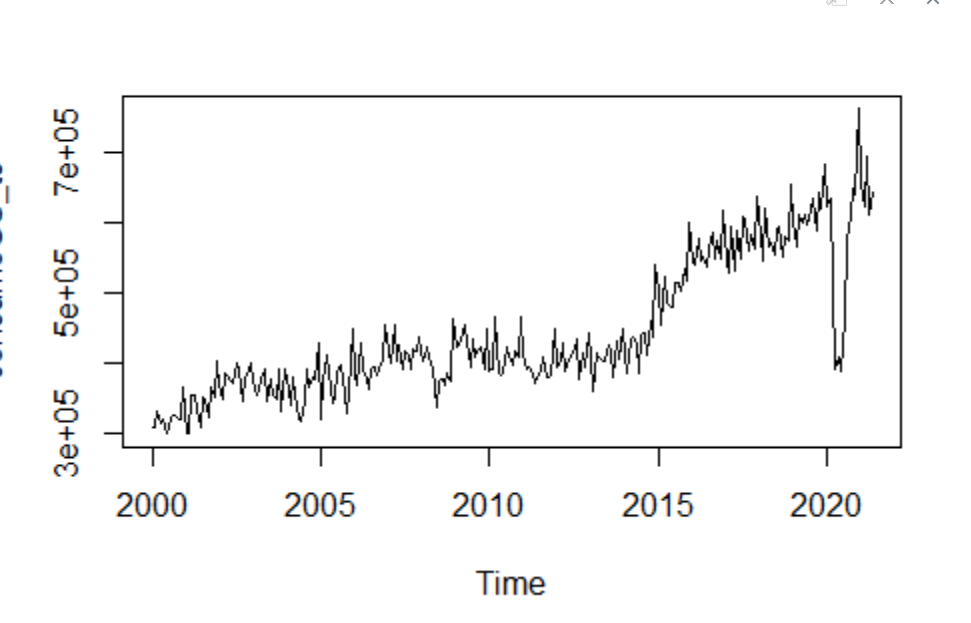
### Data set consumo

**Linea del tiempo de gasolina Super**

Inicio: 2000-01

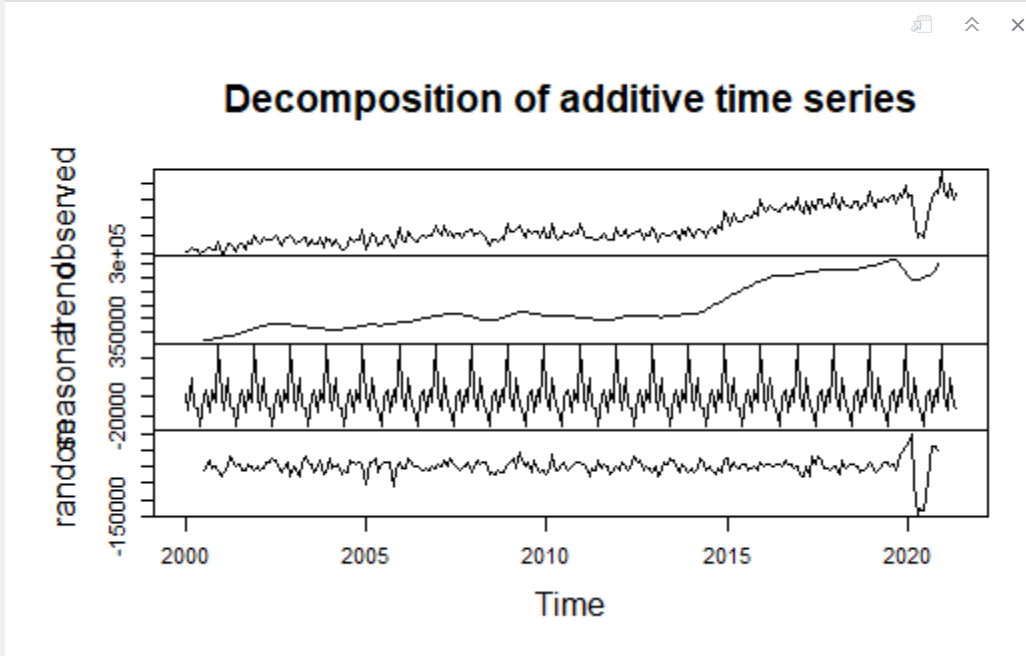
Fin: 2021-05

Frecuencia: 12



De esta linea podemos ver a primera vista que posee una tendencia a crecer, y que hay momentos en la linea de tiempo que hace que las varianzas tambien se vean afectadas y sean diferentes entre si. Podemos ver que se trata de una serie con frecuencia anual donde hay datos de todos los meses desde enero de 2000 hasta mayo del 2021.Como se puede ver la cantidad de galones vendidos va aumentando a medida que pasan los años.

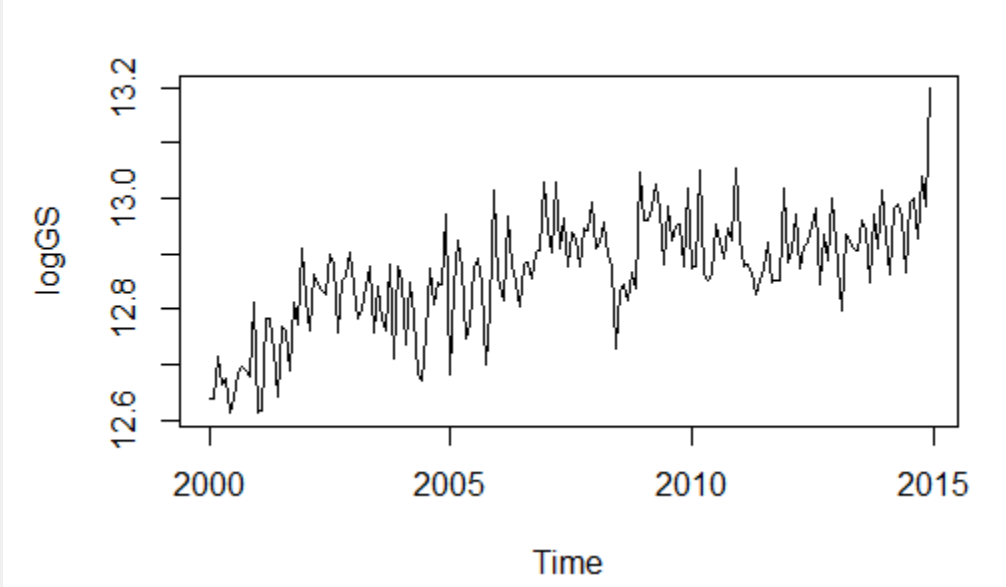
**Descomposición de la serie**



Podemos observar una serie con tendencia a aumentar, que no es estacionaria en varianza, y además tiene estacionalidad.

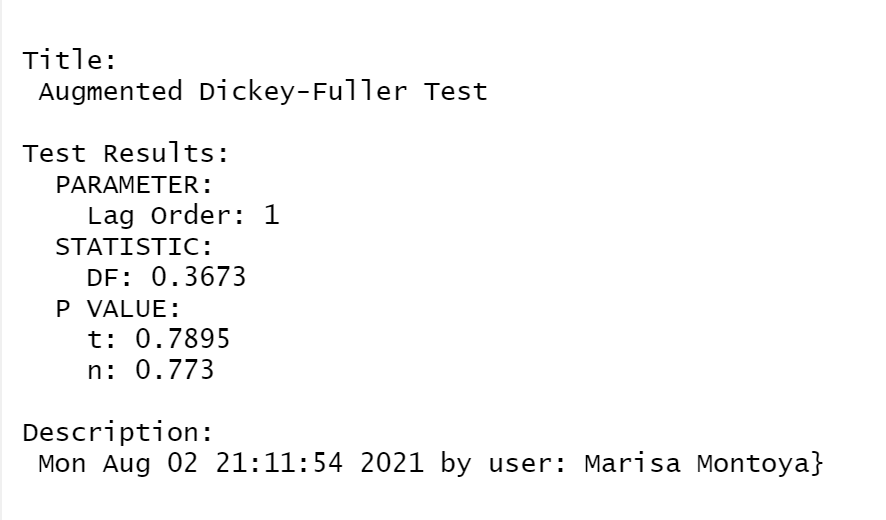
**Estimar los parámetros del modelo.**

Cómo no es estacionaria en varianza le haremos una transformación logaritmica para hacerla constante en varianza.



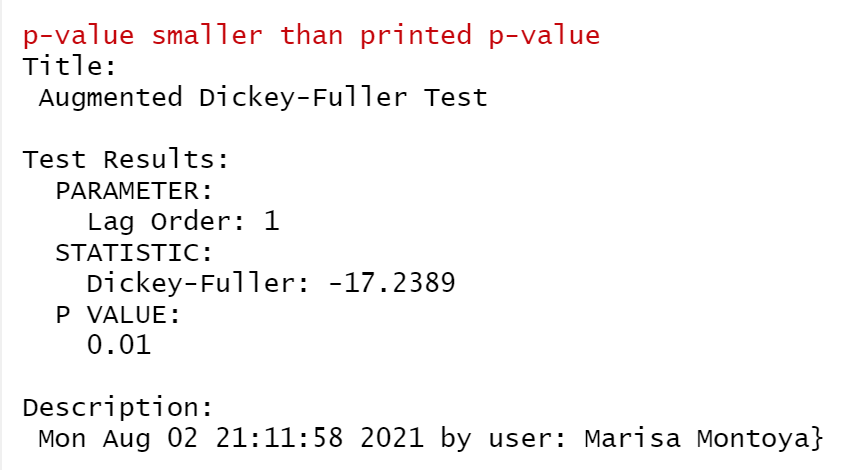
Podemos observar que las varianzas si se pudieron hacer más constantes, sin embargo, se pueden observar picos de varianza desiguales, lo cual puede ser un problema y concreta que no se pudo estacionarizar en varianza en su totalidad.

Por lo que ahora trataremos de quitar la tendencia y estacionarizar en normalidad.



La prueba de Dickey-Fuller Test nos demuestra que p es mayor a 0.05 por lo que no podemos rechazar Ho, y por ende, asumir que no hay raíces unitarias. Por lo que la serie de tiempo no es estacionaria en media.

Por lo que debemos de hacer lo mismo, pero con una diferenciación en la serie de tiempo.



Como se puede ver el valor de p ahora sí está por debajo de 0.05 por lo que se puede descartar la hipótesis nula de que existan raíces unitarias.

Solo se hizo uso de una diferenciación, por lo que d es 1, pero para conocer p y q se debe de realizar un gráfico de autocorrelación y autocorrelación parcial.