

# Análisis Preliminar de Series

María José Ramírez Cifuentes

2025-07-21

## Serie #1: Importaciones

```
data <- read_excel("importaciones.xlsx")
names(data) <- trimws(names(data))
colnames(data)

## [1] "Fecha" "Aceites lubricantes"
## [3] "Asfalto" "Bunker"
## [5] "Ceras" "Combustible turbo jet"
## [7] "Diesel bajo azufre" "Diesel ultra bajo azufre"
## [9] "Gas licuado" "Gasolina de aviación"
## [11] "Gasolina regular" "Gasolina superior"
## [13] "Grasas lubricantes" "Kerosina"
## [15] "Mezclas oleosas" "Naftas"
## [17] "Petcoke" "Petróleo crudo"
## [19] "Solventes" "Butano"
## [21] "Diesel" "MTBE"
## [23] "Orimulsi3n" "Petr3leo reconstituido"
## [25] "Total importaci3n"
```

## Resumen

```
importaciones <- data %>%
  mutate(
    diesel = rowSums(select(., contains("Diesel", ignore.case = TRUE)), na.rm
= TRUE)
  ) %>%
  select(
    Fecha,
    regular = "Gasolina regular",
    super = "Gasolina superior",
    diesel,
    gas_licuado = "Gas licuado"
  )

summary(importaciones)

##      Fecha      regular      super
## Min.   :2001-01-01 00:00:00.00 Min.   : 81015 Min.   : 170293
## 1st Qu.:2007-02-01 00:00:00.00 1st Qu.: 208796 1st Qu.: 372487
## Median :2013-03-01 00:00:00.00 Median : 336006 Median : 491195
```

```
## Mean :2013-03-01 19:19:51.80 Mean : 419996 Mean : 494588
## 3rd Qu.:2019-04-01 00:00:00.00 3rd Qu.: 592103 3rd Qu.: 616269
## Max. :2025-05-01 00:00:00.00 Max. :1141366 Max. :1227174
## diesel gas_licuado
## Min. : 229765 Min. : 100562
## 1st Qu.: 693942 1st Qu.: 230585
## Median : 889927 Median : 416068
## Mean :1258587 Mean : 422204
## 3rd Qu.:1873315 3rd Qu.: 584710
## Max. :3241294 Max. :1077123
```

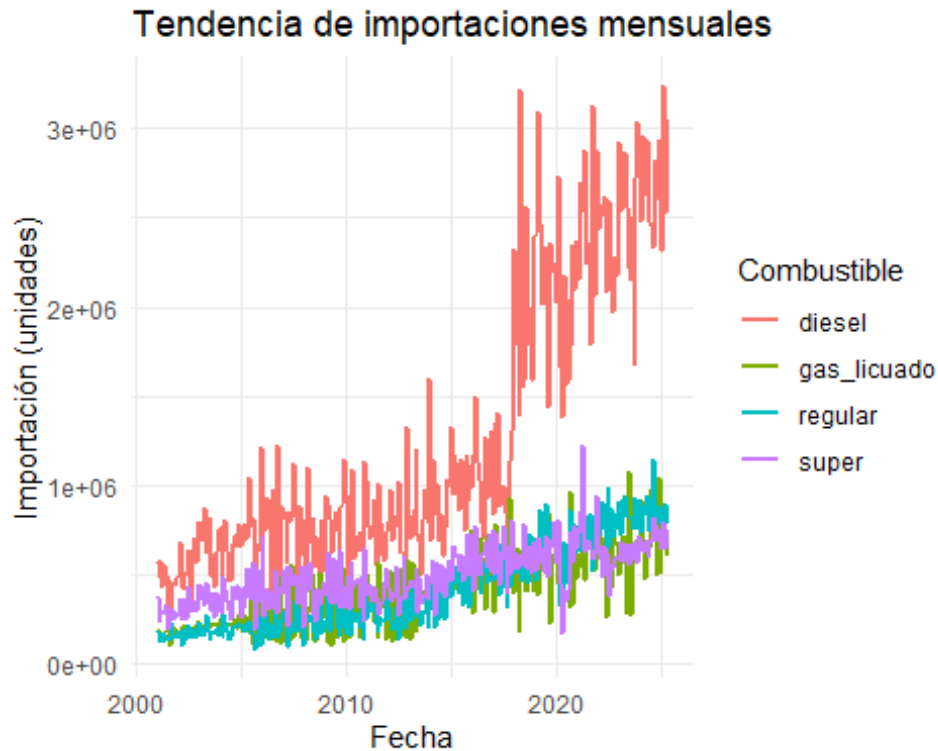
El resumen estadístico muestra que el diésel ha sido el combustible más importado en promedio y en volumen máximo desde 2001 hasta 2025, seguido por el gas licuado, súper y regular. Los valores medios y máximos reflejan una alta demanda sostenida de diésel, mientras que el gas licuado, aunque más variable, también ha ido aumentando con el tiempo.

### Tendencia y Estacionalidad

```
# Transformar a formato Largo
importaciones_long <- importaciones %>%
  pivot_longer(-Fecha, names_to = "combustible", values_to = "importacion")

# Gráfico de tendencia
ggplot(importaciones_long, aes(x=Fecha, y=importacion, color=combustible)) +
  geom_line(size=1) +
  labs(title="Tendencia de importaciones mensuales",
       x="Fecha", y="Importación (unidades)", color="Combustible") +
  theme_minimal()

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



La gráfica muestra una tendencia general creciente en las importaciones mensuales de todos los tipos de combustibles. El diésel destaca por su volumen significativamente mayor y su incremento sostenido, mientras que el gas licuado, regular y súper también presentan alzas, aunque con fluctuaciones más marcadas. Esto sugiere un aumento constante en la demanda energética del país a lo largo de los años.

### Estacionariedad

```
# Prueba de estacionariedad para cada combustible
import_ts <- ts(importaciones[-1], start=c(2000,1), frequency=12) # ajustar
# fecha inicial

apply(import_ts, 2, function(x) {
  adf.test(x)$p.value
})

## Warning in adf.test(x): p-value smaller than printed p-value

## Warning in adf.test(x): p-value smaller than printed p-value

##      regular      super      diesel gas_licuado
## 0.1575837 0.0100000 0.4205289 0.0100000
```

Los resultados de las pruebas de estacionariedad muestran que las series de importaciones de regular, súper y gas licuado tienen valores cercanos a 0.01, lo que indica que no son estacionarias y presentan una tendencia clara a lo largo del tiempo. En contraste, la serie de diésel tiene un valor más alto ( $\approx 0.16$ ), pero sigue sin ser estacionaria. Esto implica que

todas las series tienen una tendencia creciente y probablemente requieren diferenciación para análisis de series temporales más avanzados.

## Serie 2: Precios

```
datos <- read_excel("precios_unidos.xlsx")
datos

## # A tibble: 1,677 × 8
##   Fecha                `Tipo de Cambio` superior regular diesel bunker
##   <dtm>                <chr>          <chr>    <chr>    <chr>    <chr>
## 1 2021-01-01 00:00:00 7.7062499999999998 43.24    27.45    25.47
## 18.7399999999...
## 2 2021-01-02 00:00:00 7.7073999999999998 43.24    27.45    25.47
## 18.7399999999...
## 3 2021-01-03 00:00:00 7.7097100000000003 43.24    27.45    25.47
## 18.7399999999...
## 4 2021-01-04 00:00:00 7.7097100000000003 43.24    27.45    25.47
## 18.7399999999...
## 5 2021-01-05 00:00:00 7.7097100000000003 43.24    27.45    25.47
## 18.7399999999...
## 6 2021-01-06 00:00:00 7.7122299999999999 43.24    28.05    26.18
## 18.7399999999...
## 7 2021-01-07 00:00:00 7.7082199999999998 43.24    28.05    26.18
## 18.7399999999...
## 8 2021-01-08 00:00:00 7.7095200000000004 42.98    28.05    26.18
## 18.7399999999...
## 9 2021-01-09 00:00:00 7.7099099999999998 42.98    28.05    26.18
## 18.7399999999...
## 10 2021-01-10 00:00:00 7.7107599999999996 42.98    28.05    26.18
## 18.7399999999...
## # i 1,667 more rows
## # i 2 more variables: `Glp cilindro 25Lbs.` <chr>, year <chr>
```

```
colnames(datos)
```

```
## [1] "Fecha"                "Tipo de Cambio"      "superior"
## [4] "regular"              "diesel"              "bunker"
## [7] "Glp cilindro 25Lbs." "year"
```

```
names(datos)<- trimws(names(datos))
```

```
datos <- datos %>%
  clean_names() # Limpia nombres de columnas vacíos o raros
```

## Resumen

```
precio <- datos %>%
  mutate(
```

```

    year      = as.numeric(year),
    regular   = str_replace_all(regular, "[^0-9\\.]", ""), # quita
letras/símbolos
    superior  = str_replace_all(superior, "[^0-9\\.]", ""),
    diesel    = str_replace_all(diesel, "[^0-9\\.]", "")
  ) %>%
  mutate(
    regular   = str_replace(regular, ",", "."), # si usas coma decimal
    superior  = str_replace(superior, ",", "."),
    diesel    = str_replace(diesel, ",", ".")
  ) %>%
  mutate(
    regular   = as.numeric(regular),
    superior  = as.numeric(superior),
    diesel    = as.numeric(diesel)
  ) %>%
  filter(!is.na(year)) %>%
  group_by(year) %>%
  summarise(across(c(regular, superior, diesel), mean, na.rm = TRUE))

## Warning: There was 1 warning in `summarise()`.
## i In argument: `across(c(regular, superior, diesel), mean, na.rm = TRUE)`.
## i In group 1: `year = 2021`.
## Caused by warning:
## ! The `...` argument of `across()` is deprecated as of dplyr 1.1.0.
## Supply arguments directly to `.fns` through an anonymous function instead.
##
## # Previously
##   across(a:b, mean, na.rm = TRUE)
##
## # Now
##   across(a:b, \(x) mean(x, na.rm = TRUE))

summary(precio)

##      year      regular      superior      diesel
## Min.   :2021   Min.   :27.11   Min.   :27.30   Min.   :22.53
## 1st Qu.:2022   1st Qu.:28.78   1st Qu.:30.03   1st Qu.:26.42
## Median :2023   Median :30.33   Median :32.40   Median :28.09
## Mean   :2023   Mean   :30.63   Mean   :32.67   Mean   :28.19
## 3rd Qu.:2024   3rd Qu.:32.51   3rd Qu.:34.96   3rd Qu.:30.53
## Max.   :2025   Max.   :34.43   Max.   :38.68   Max.   :33.39

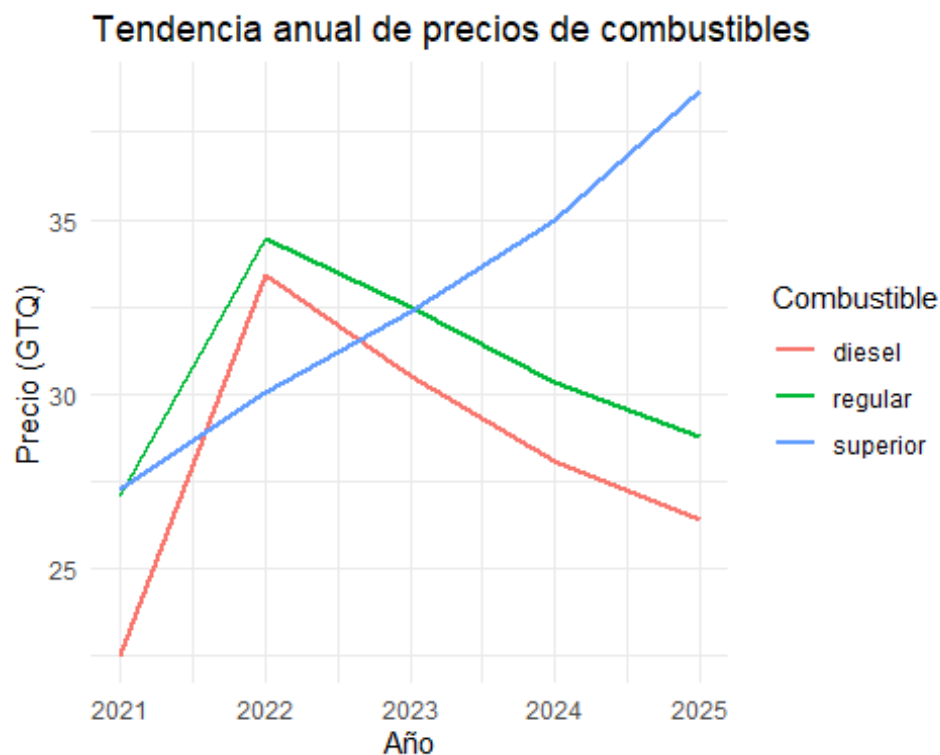
```

Entre 2021 y 2025, los precios promedio más altos corresponden a la gasolina superior, seguidos por la regular y el diésel. La distribución indica que todos los combustibles alcanzaron su precio máximo en 2022 y desde entonces han mostrado una ligera tendencia a la baja.

## Tendencia y Estacionalidad

```
precios_long <- precio %>%
  pivot_longer(
    cols = -year,
    names_to = "tipo_combustible",
    values_to = "precio"
  ) %>%
  filter(!is.na(precio)) # Muy importante

# Ahora sí el gráfico
ggplot(precios_long, aes(x = year, y = precio, color = tipo_combustible)) +
  geom_line(linewidth = 1) +
  labs(
    title = "Tendencia anual de precios de combustibles",
    x = "Año",
    y = "Precio (GTQ)",
    color = "Combustible"
  ) +
  theme_minimal()
```



La gráfica revela un pico en los precios de todos los combustibles en 2022, probablemente relacionado con factores globales. Posteriormente, los precios han disminuido progresivamente. La gasolina superior se mantiene como la más cara durante todo el período, reflejando su categoría premium.

## Estacionariedad

```
# Prueba de estacionariedad para cada combustible
c_ts <- ts(precio[-1], start=c(2000,1), frequency=12) # ajustar fecha inicial

apply(c_ts, 2, function(x) {
  adf.test(x)$p.value
})

## regular superior diesel
##      NaN      NaN      NaN
```

Para los precios, no se muestran valores de prueba (NaN), lo que sugiere que no se realizó o no fue posible aplicar la prueba de estacionariedad con los datos disponibles. Sin embargo, visualmente se observa una tendencia clara con un punto de inflexión en 2022, por lo que es probable que las series tampoco sean estacionarias.

## Serie 3: Consumo

```
data <- read_excel("Consumo.xlsx")
names(data) <- trimws(names(data))
```

### Resumen

```
consumos <- data %>%
  mutate(
    diesel = rowSums(select(., contains("Diesel", ignore.case = TRUE)), na.rm
= TRUE)
  ) %>%
  select(
    Fecha,
    regular = "Gasolina regular",
    super = "Gasolina superior",
    diesel,
    gas_licuado = "Gas Licuado"
  )
```

```
summary(consumos)
```

##	Fecha	regular	super
##	Min. :2000-01-01 00:00:00.00	Min. :160742	Min. :300243
##	1st Qu.:2006-05-01 00:00:00.00	1st Qu.:216561	1st Qu.:385045
##	Median :2012-09-01 00:00:00.00	Median :290287	Median :424439
##	Mean :2012-08-31 04:33:50.16	Mean :405017	Mean :474466
##	3rd Qu.:2019-01-01 00:00:00.00	3rd Qu.:570267	3rd Qu.:575721
##	Max. :2025-05-01 00:00:00.00	Max. :942394	Max. :790948
##	diesel	gas_licuado	
##	Min. : 507663	Min. :167818	
##	1st Qu.: 701256	1st Qu.:227155	

```
## Median : 819941    Median :268875
## Mean   :1265293    Mean   :322886
## 3rd Qu.:2027654    3rd Qu.:407954
## Max.   :2973917    Max.   :600454
```

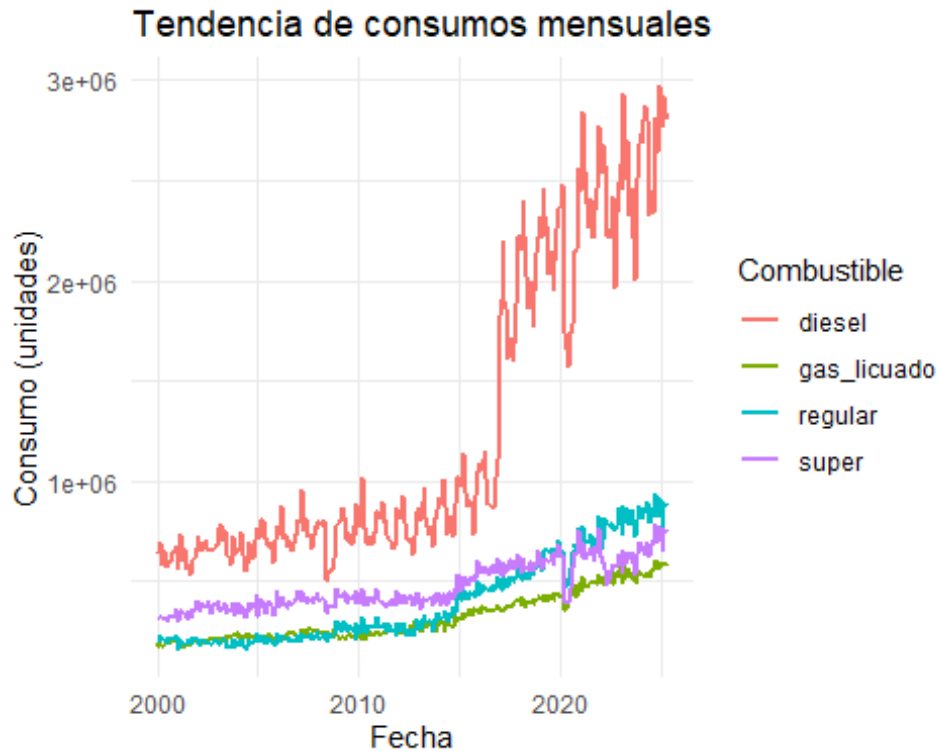
Los datos muestran que el diésel presenta el mayor volumen de consumo promedio (1.26 millones de galones), seguido por la gasolina regular (405,017 galones) y superior (474,466 galones). El gas licuado muestra el comportamiento más estable. Se observa un crecimiento general en el consumo de todos los combustibles durante el período analizado, siendo especialmente marcado en el diésel, que triplicó su demanda mínima inicial. La gasolina regular destaca por su amplia variabilidad (de 160,742 a 942,394 galones), mientras que la superior mantuvo una demanda más constante. Estos patrones reflejan probablemente el crecimiento económico, la expansión del parque vehicular y cambios en los hábitos de consumo energético durante el último cuarto de siglo.

### Tendencia y Estacionalidad

```
# Transformar a formato Largo
consumos_long <- consumos %>%
  pivot_longer(-Fecha, names_to = "combustible", values_to = "consumo") #
  Cambiado a "consumo"

# Gráfico de tendencia CORRECTO
ggplot(consumos_long, aes(x=Fecha, y=consumo, color=combustible)) + #
  "consumo" en lugar de "consumos"
  geom_line(linewidth=1) + # Usar linewidth en lugar de size (versiones
recientes de ggplot2)
  labs(title="Tendencia de consumos mensuales",
       x="Fecha", y="Consumo (unidades)", color="Combustible") +
  theme_minimal()
```





Esta gráfica de las tendencias de consumo mensual confirma visualmente que todas las series son **no estacionarias**, ya que presentan una clara y persistente tendencia al alza a lo largo de más de dos décadas. El **diésel** (rojo) no solo es el combustible de mayor consumo por un amplio margen, sino que también experimenta un drástico cambio estructural alrededor de 2018, saltando a un nuevo y mucho más alto nivel de demanda. Por su parte, el **gas licuado** (verde) muestra el crecimiento más estable y predecible. Mientras tanto, se observa una dinámica interesante entre las gasolinas: la **regular** (celeste) exhibe un crecimiento acelerado que la lleva a superar a la **super** (violeta) alrededor de 2020. La gasolina super, a su vez, muestra una caída notable durante el año 2020, probablemente debido a la pandemia, antes de retomar su tendencia ascendente.

## Estacionariedad

```
# Prueba de estacionariedad para cada combustible
c_ts <- ts(consumos[-1], start=c(2000,1), frequency=12) # ajustar fecha inicial

apply(c_ts, 2, function(x) {
  adf.test(x)$p.value
})

##      regular      super      diesel gas_licuado
## 0.68623778 0.09274994 0.62960418 0.91996835
```

Los resultados de la **prueba de Dickey-Fuller Aumentada (ADF)** confirman estadísticamente la no estacionariedad de las series de consumo. Para las series de **Gasolina Regular, Diésel y Gas Licuado**, los p-valores obtenidos (0.68, 0.63 y 0.92,

respectivamente) son muy superiores al nivel de significancia estándar de 0.05. En el caso de la **Gasolina Super**, aunque su p-valor de 0.093 es más bajo, sigue estando por encima de dicho umbral. Por lo tanto, para las cuatro series de tiempo, **no se puede rechazar la hipótesis nula** de que existe una raíz unitaria. Esto valida formalmente lo que se observó en las gráficas: las series no son estacionarias y será necesario aplicar transformaciones, como la diferenciación, para poder modelarlas adecuadamente.