

GEM

Alfonso Camarillo Núñez

05 de marzo de 2024

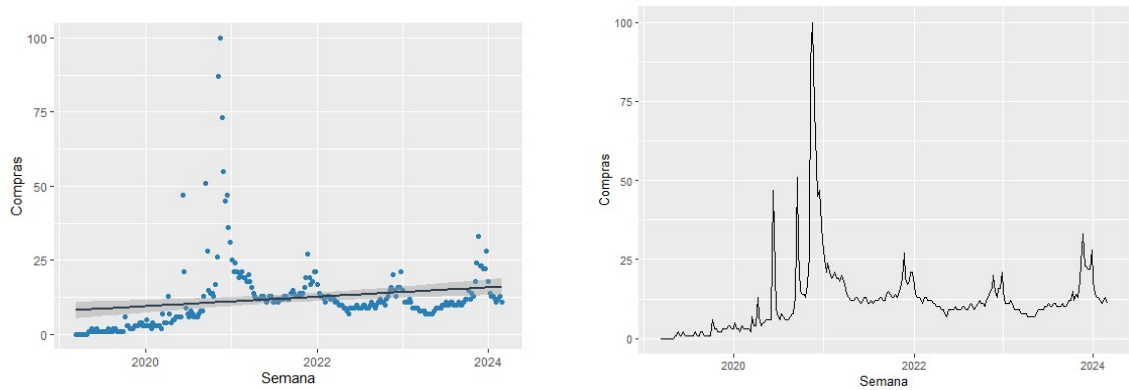
MÓDULO III: MÉTODOS ESTADÍSTICOS MÚLTIPLES - SERIES DE TIEMPO

Analís con Modelo Holt Winters

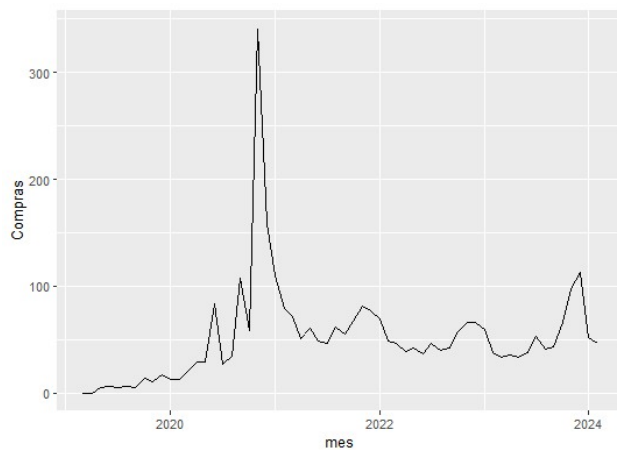
Los datos para evaluar son obtenidos de Google trends de la compra en miles por semana de la play station 5 en el mundo en los últimos 5 años en búsquedas en la web, los datos son los siguientes:

Semana	Compras	Semana	Compras	Semana	Compras	Semana	Compras	Semana	Compras	Semana	Compras	Semana	Compras	Semana	Compras	Semana	Compras	Semana	Compras
10/03/2019	0	09/09/2019	1	08/03/2020	2	06/09/2020	13	07/03/2021	18	05/09/2021	12	06/03/2022	12	04/09/2022	9	05/03/2023	9	03/09/2023	10
17/03/2019	0	15/09/2019	1	15/03/2020	7	13/09/2020	51	14/03/2021	20	12/09/2021	14	13/03/2022	12	11/09/2022	10	12/03/2023	8	10/09/2023	10
24/03/2019	0	22/09/2019	1	22/03/2020	4	20/09/2020	28	21/03/2021	18	19/09/2021	15	20/03/2022	11	18/09/2022	11	19/03/2023	8	17/09/2023	11
31/03/2019	0	29/09/2019	1	29/03/2020	4	27/09/2020	15	28/03/2021	16	26/09/2021	14	27/03/2022	11	25/09/2022	12	26/03/2023	8	24/09/2023	12
07/04/2019	0	06/10/2019	6	05/04/2020	13	04/10/2020	14	04/04/2021	14	03/10/2021	13	03/04/2022	10	02/10/2022	11	02/04/2023	8	01/10/2023	12
14/04/2019	0	13/10/2019	3	12/04/2020	7	11/10/2020	14	11/04/2021	13	10/10/2021	13	10/04/2022	10	09/10/2022	11	09/04/2023	7	08/10/2023	15
21/04/2019	0	20/10/2019	3	19/04/2020	4	18/10/2020	13	18/04/2021	12	17/10/2021	14	17/04/2022	10	16/10/2022	10	16/04/2023	7	15/10/2023	12
28/04/2019	0	27/10/2019	2	26/04/2020	5	25/10/2020	17	25/04/2021	12	24/10/2021	13	24/04/2022	9	23/10/2022	12	23/04/2023	7	22/10/2023	14
05/05/2019	1	03/11/2019	2	03/05/2020	5	01/11/2020	26	02/05/2021	12	31/10/2021	14	01/05/2022	9	30/10/2022	13	30/04/2023	7	29/10/2023	13
12/05/2019	1	10/11/2019	2	10/05/2020	6	08/11/2020	87	09/05/2021	13	07/11/2021	16	08/05/2022	8	06/11/2022	15	07/05/2023	7	05/11/2023	18
19/05/2019	2	17/11/2019	3	17/05/2020	6	15/11/2020	100	16/05/2021	13	14/11/2021	19	15/05/2022	7	13/11/2022	16	14/05/2023	8	12/11/2023	24
26/05/2019	1	24/11/2019	3	24/05/2020	6	22/11/2020	73	23/05/2021	12	21/11/2021	27	22/05/2022	9	20/11/2022	20	21/05/2023	9	19/11/2023	33
02/06/2019	1	01/12/2019	3	31/05/2020	6	29/11/2020	55	30/05/2021	11	28/11/2021	19	29/05/2022	9	27/11/2022	15	28/05/2023	9	26/11/2023	23
09/06/2019	2	08/12/2019	4	07/06/2020	47	06/12/2020	45	06/06/2021	12	05/12/2021	17	05/06/2022	9	04/12/2022	13	04/06/2023	9	03/12/2023	23
16/06/2019	1	15/12/2019	4	14/06/2020	21	13/12/2020	47	13/06/2021	13	12/12/2021	18	12/06/2022	9	11/12/2022	16	11/06/2023	9	10/12/2023	22
23/06/2019	1	22/12/2019	3	21/06/2020	9	20/12/2020	36	20/06/2021	13	19/12/2021	21	19/06/2022	10	18/12/2022	16	18/06/2023	10	17/12/2023	22
30/06/2019	1	29/12/2019	3	28/06/2020	7	27/12/2020	31	27/06/2021	11	26/12/2021	21	26/06/2022	9	25/12/2022	21	25/06/2023	10	24/12/2023	28
07/07/2019	1	05/01/2020	5	05/07/2020	6	03/01/2021	25	04/07/2021	11	02/01/2022	17	03/07/2022	9	01/01/2023	15	02/07/2023	11	31/12/2023	18
14/07/2019	1	12/01/2020	3	12/07/2020	8	10/01/2021	21	11/07/2021	12	09/01/2022	14	10/07/2022	9	08/01/2023	11	09/07/2023	10	07/01/2024	14
21/07/2019	1	19/01/2020	3	19/07/2020	7	17/01/2021	24	18/07/2021	11	16/01/2022	13	17/07/2022	10	15/01/2023	11	16/07/2023	10	14/01/2024	13
28/07/2019	2	26/01/2020	2	26/07/2020	6	24/01/2021	21	25/07/2021	12	23/01/2022	13	24/07/2022	10	22/01/2023	11	23/07/2023	11	21/01/2024	13
04/08/2019	1	02/02/2020	4	02/08/2020	6	31/01/2021	19	01/08/2021	12	30/01/2022	12	31/07/2022	9	29/01/2023	12	30/07/2023	11	28/01/2024	12
11/08/2019	1	09/02/2020	3	09/08/2020	6	07/02/2021	20	08/08/2021	12	06/02/2022	11	07/08/2022	9	05/02/2023	11	06/08/2023	10	04/02/2024	11
18/08/2019	2	16/02/2020	3	16/08/2020	7	14/02/2021	21	15/08/2021	13	13/02/2022	13	14/08/2022	10	12/02/2023	9	13/08/2023	10	11/02/2024	12
25/08/2019	2	23/02/2020	3	23/08/2020	8	21/02/2021	19	22/08/2021	13	20/02/2022	13	21/08/2022	11	19/02/2023	9	20/08/2023	10	18/02/2024	13
01/09/2019	1	01/03/2020	3	30/08/2020	8	28/02/2021	19	29/08/2021	12	27/02/2022	12	28/08/2022	10	26/02/2023	9	27/08/2023	11	25/02/2024	11

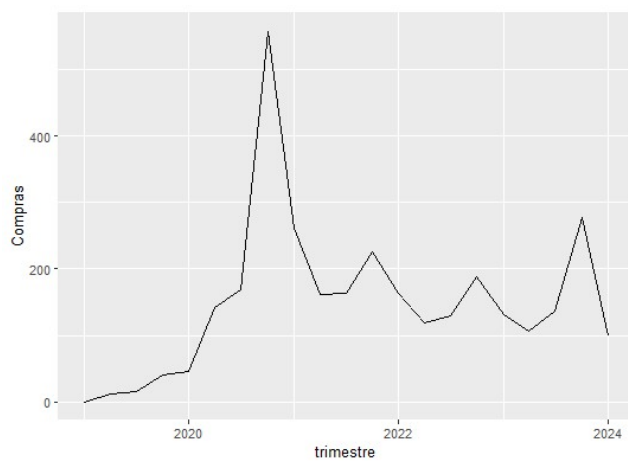
Al graficar tenemos la siguiente imagen, mostrando una línea de tendencia creciente:



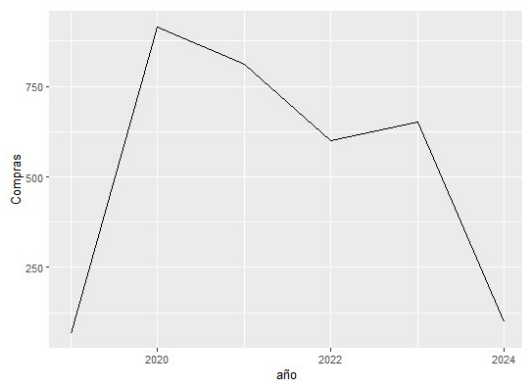
Donde realizando un análisis por mes tenemos la siguiente gráfica:



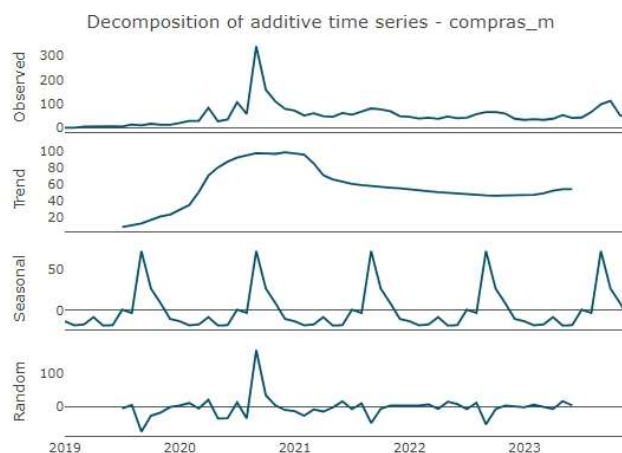
Al hacerlo por trimestre vemos lo siguiente



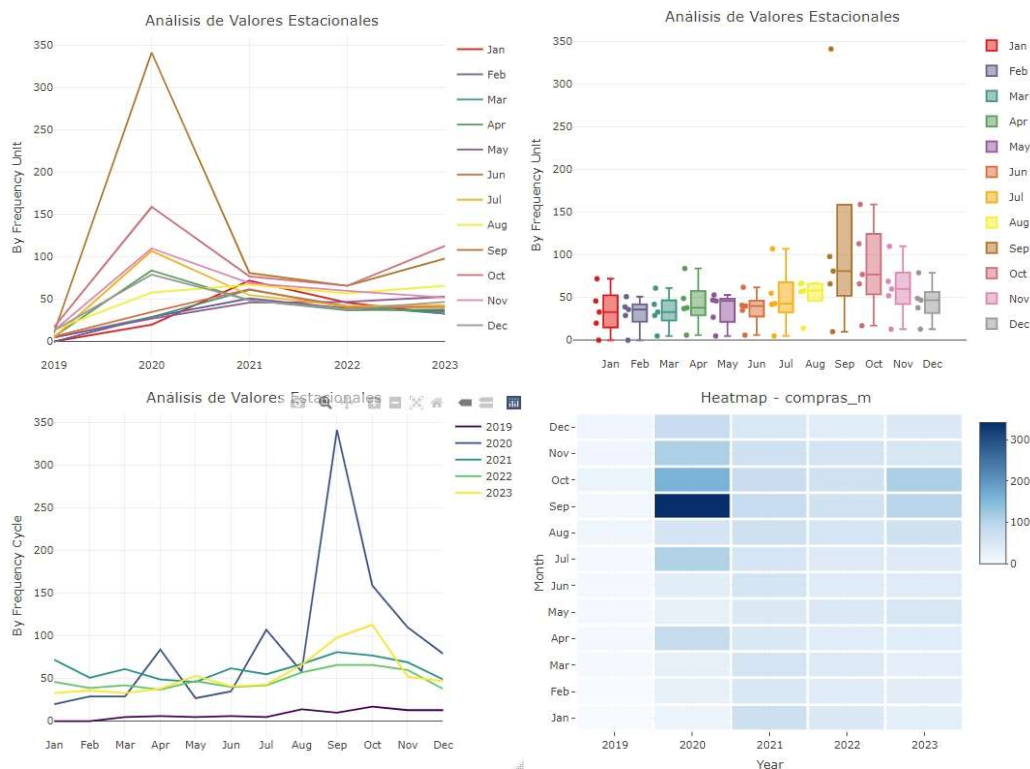
Por año vemos el patrón de tendencia el cual va disminuyendo en los últimos años

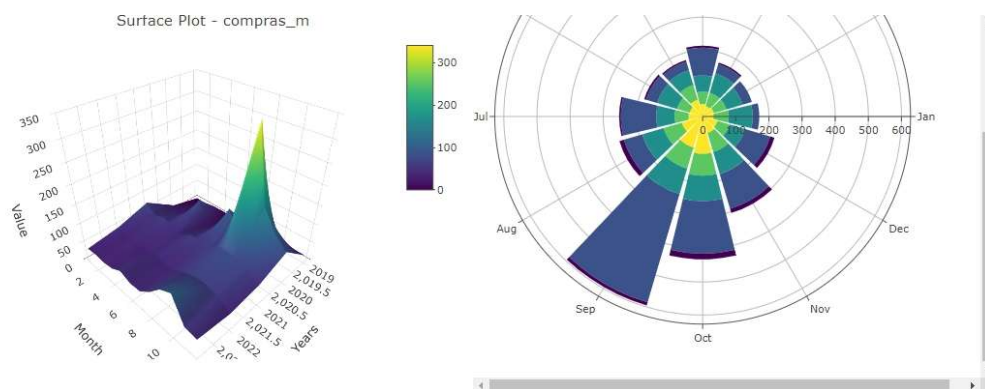


Al descomponer la serie mensual vemos la siguiente grafica donde vemos la serie original su tendencia que primero fue creciente y después de mediados de 2021 decrecio pero después se ha mantenido con un ligero crecimiento, la estacionalidad donde se ve que si hay este componente y en algunos meses la compra está por debajo del promedio y en otros se tiene un pico, y al fina el componente aleatorio.

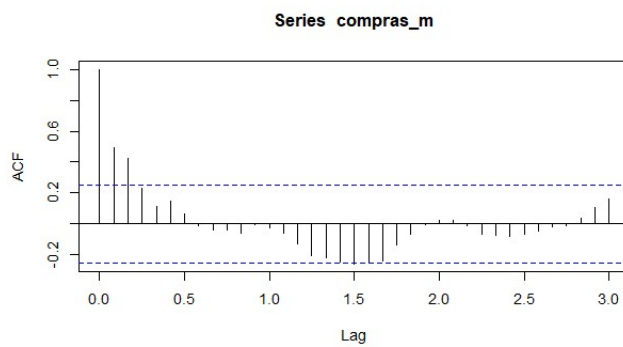


Analizando la estacionalidad a nivel mensual vemos en las siguientes graficas que hay cierto paralelismo en las gráficas de cada mes y se nota que el mes donde mas se adquiere es entre septiembre y octubre

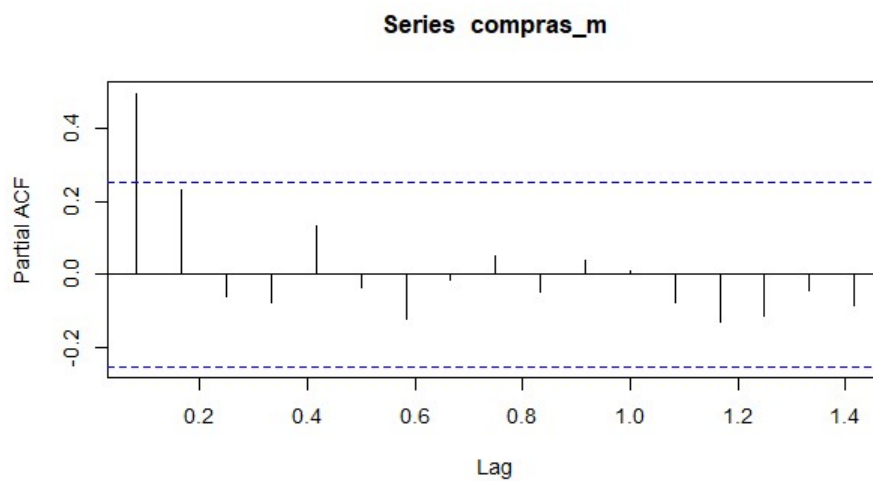




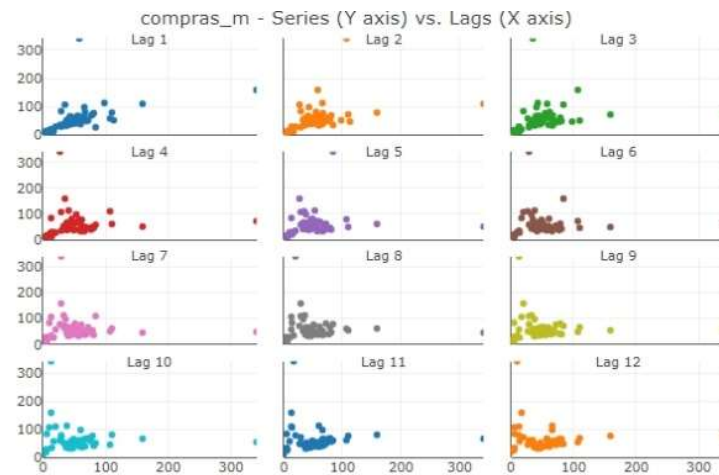
Para el análisis de autocorrelación de primer grado tenemos hasta el segundo orden que si se presenta estacionalidad, donde después ya no es tan significativa la autocorrelación con la serie



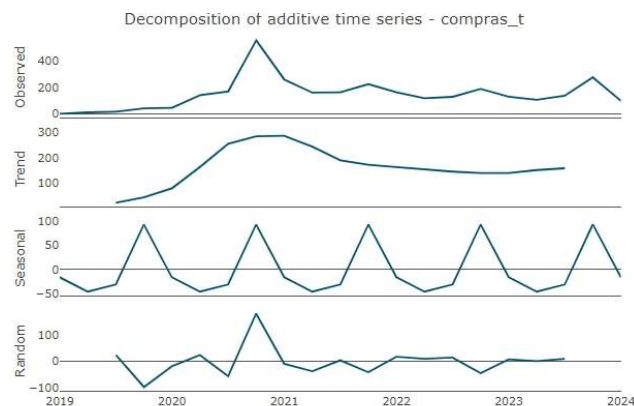
Para la autocorrelación parcial tenemos



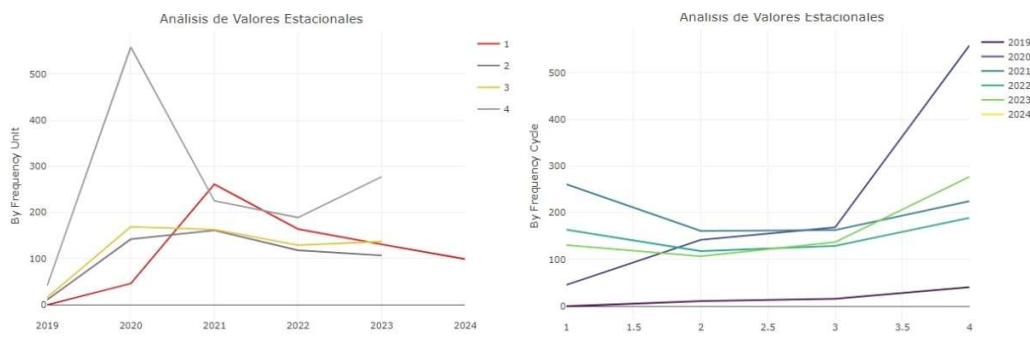
Y los desfases serían los siguientes

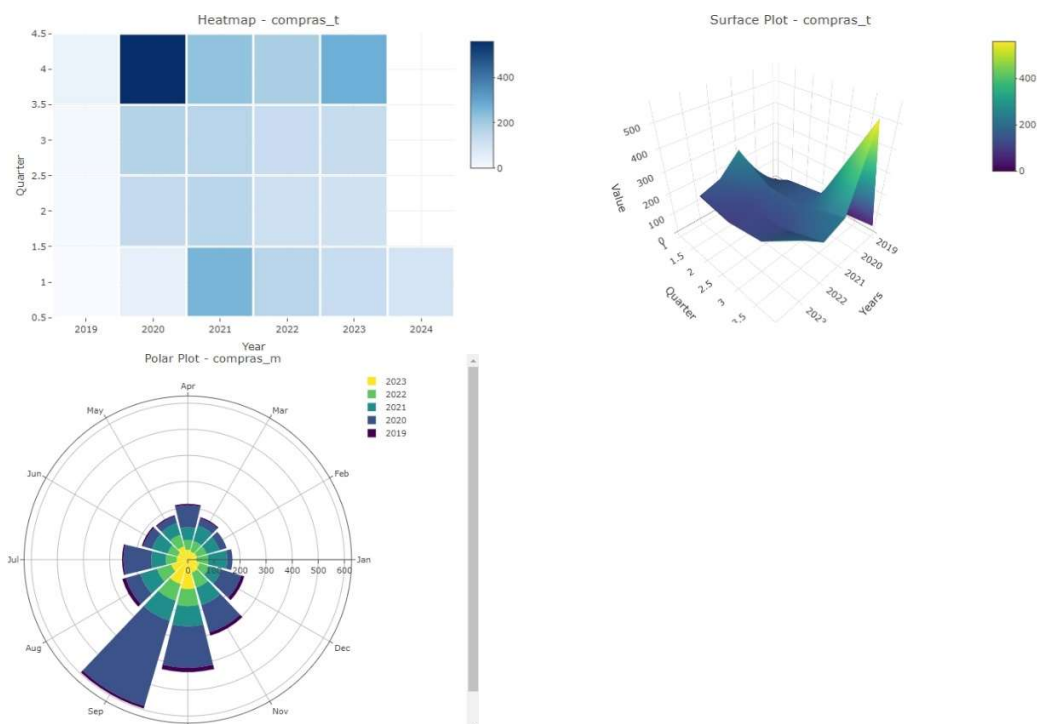


Al descomponer la serie trimestral vemos la siguiente grafica donde vemos la serie original su tendencia que primero fue creciente y después de mediados de 2021 decreció, pero después se ha mantenido con un ligero crecimiento, la estacionalidad donde se ve que si hay este componente y en algunos meses la compra está por debajo del promedio y en otros se tiene un pico, y al fina el componente aleatorio.

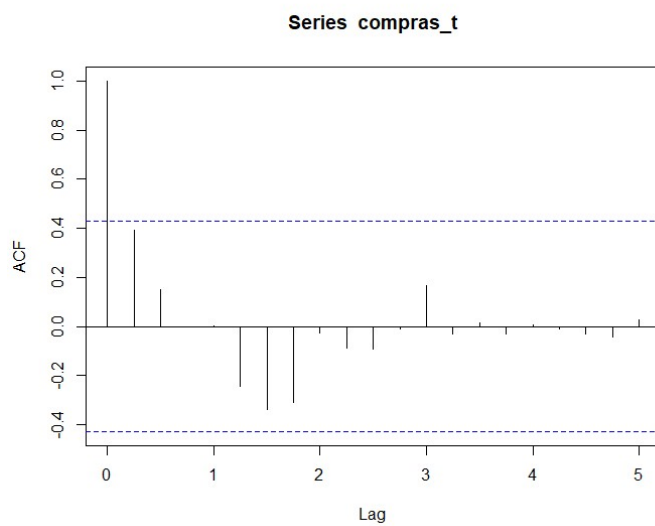


Analizando la estacionalidad a nivel trimestral vemos en las siguientes graficas que hay cierto paralelismo en las gráficas de cada mes y se nota que el mes donde más se adquiere es entre septiembre y octubre

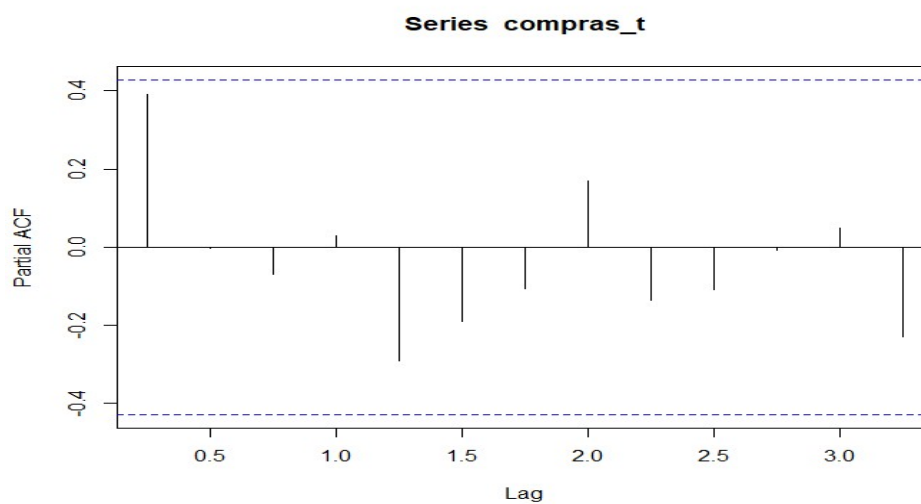




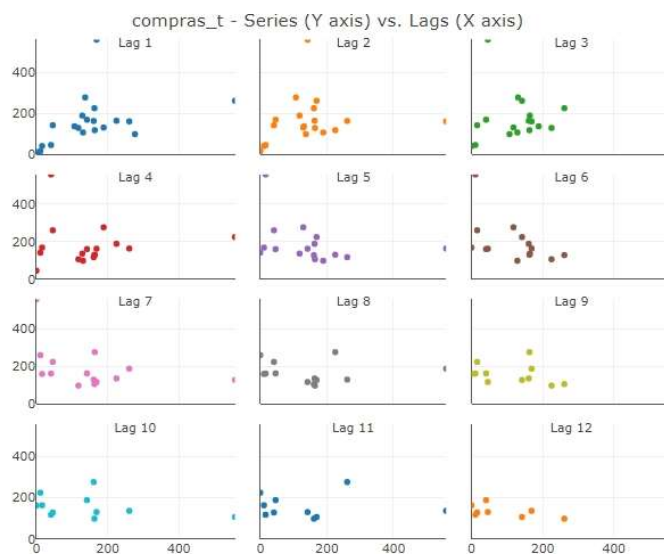
Para el análisis de autocorrelación de primer grado tenemos hasta el segundo orden que si se presenta estacionalidad, donde después ya no es tan significativa la autocorrelación con la serie



Para la autocorrelación parcial tenemos



Y los desfases serían los siguientes



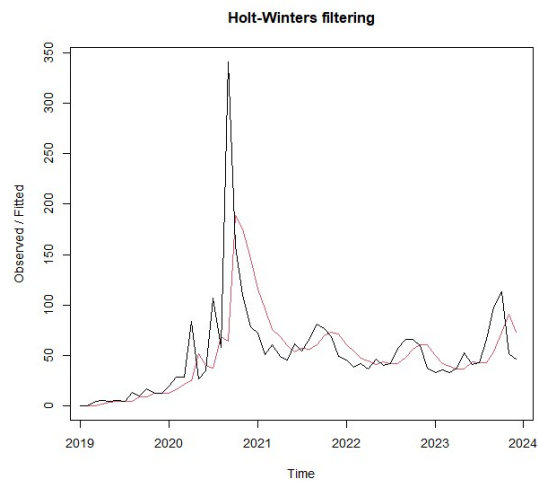
Aplicando el modelo de suavización simple de Holt Winters en las compras mensuales tenemos:

```
Call:
HoltWinters(x = compras_m, beta = FALSE, gamma = FALSE)

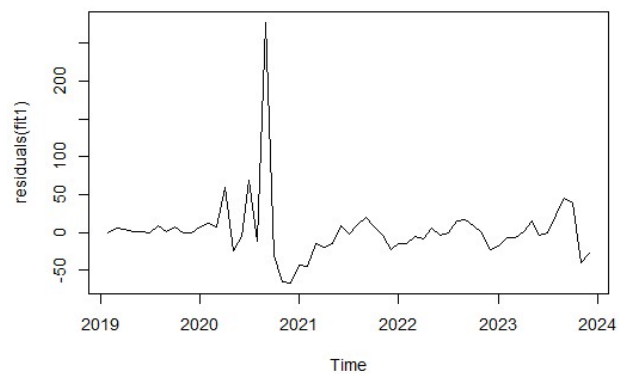
Smoothing parameters:
alpha: 0.451346
beta : FALSE
gamma: FALSE

Coefficients:
[,1]
a 61.58069
```

Con este ajuste observamos que el modelo no se ajusta bien a lo esperado hay variación, no es tan preciso.



Al graficar los residuales vemos que son aleatorios ya que no se ve algún patrón por lo que el modelo cumple la independencia de observaciones



Al calcular las medias de precisión tenemos los siguientes valores, y al evaluar los residuales vemos además que no hay autocorrelación y los valores se ajustan una distribución normal.

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	2.312506	43.25182	20.66961	-1.09283	33.43832	0.5100983	0.02684347

Ljung-Box test

```
data: Residuals from HoltWinters
Q* = 6.4381, df = 12, p-value = 0.8924

Model df: 0. Total lags used: 12
```

Aplicando el modelo de suavización doble de Holt Winters en las compras mensuales tenemos:

Holt-Winters exponential smoothing with trend and without seasonal component.

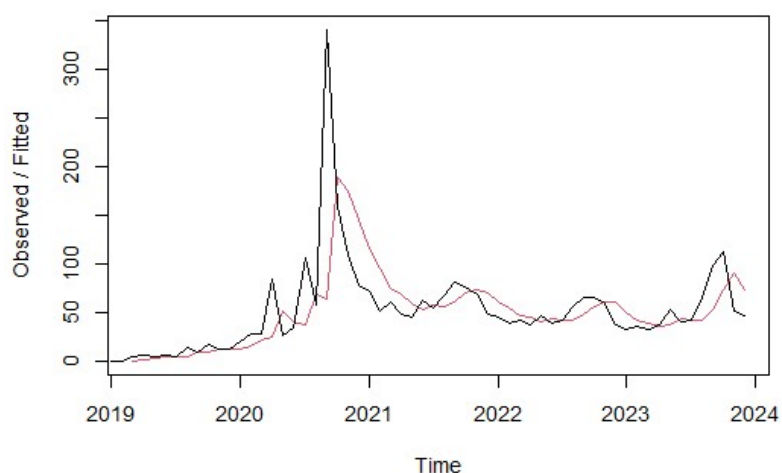
```
Call:
HoltWinters(x = compras_m, gamma = FALSE)
```

```
Smoothing parameters:
alpha: 0.4513559
beta : 0
gamma: FALSE
```

```
Coefficients:
      [,1]
a 61.58043
b  0.00000
```

Con este ajuste observamos que el modelo no se ajusta bien a lo esperado hay variación, no es tan preciso.

Holt-Winters filtering



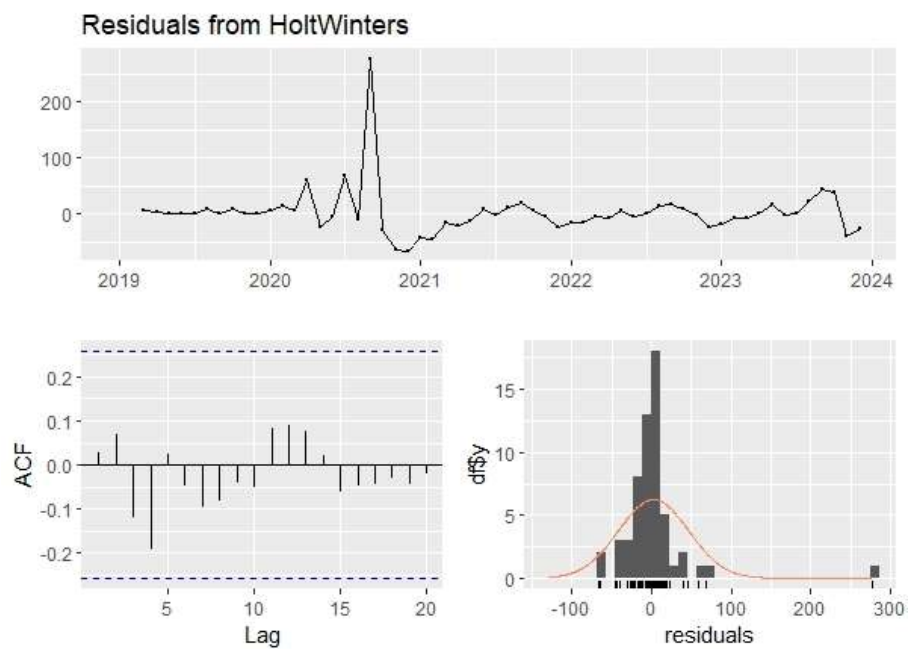
Al graficar los residuales vemos que son aleatorios ya que no se ve algún patrón por lo que el modelo cumple la independencia de observaciones

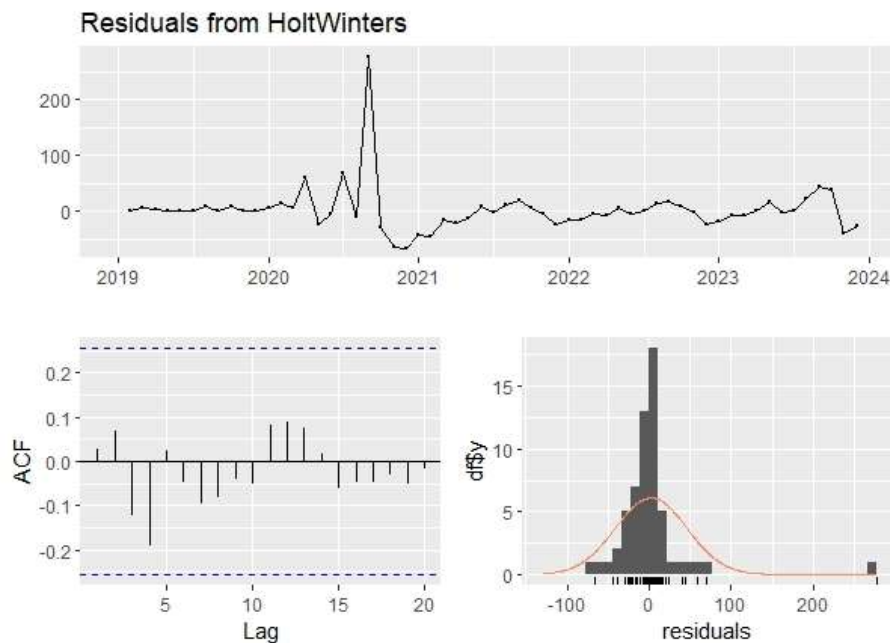
Al calcular las medias de precisión tenemos los siguientes valores, y al evaluar los residuales vemos además que no hay autocorrelación y los valores se ajustan una distribución normal.

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	2.352315	43.62308	21.02598	-1.092941	33.4382	0.5188931	0.02688175

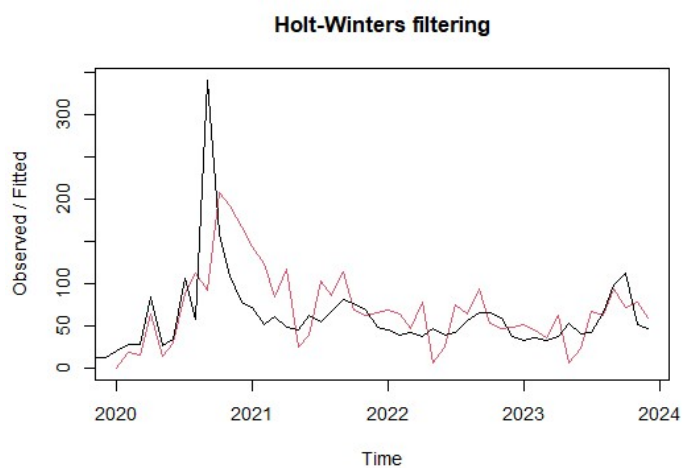
Ljung-Box test

```
data: Residuals from HoltWinters
Q* = 6.3497, df = 12, p-value = 0.8974
Model df: 0. Total lags used: 12
```





Aplicando el modelo de suavización triple aditivo de Holt Winters en las compras mensuales tenemos que con este ajuste el modelo no se ajusta bien a lo esperado hay mucha variación, no es tan preciso.



Smoothing parameters:
alpha: 0.4064143
beta : 0.04506472
gamma: 0.2178292

Coefficients:

```
[,1]
a    67.737131
b     1.367314
s1   -8.433819
s2   -6.125295
s3  -11.611777
s4   11.575699
s5  -25.239869
s6  -31.815878
s7   -2.457431
s8    5.709046
s9   34.220124
s10  13.906822
s11  -7.428481
s12 -14.595093
```

Al graficar los residuales vemos que son aleatorios ya que no se ve algún patrón por lo que el modelo cumple la independencia de observaciones

Al calcular las medias de precisión tenemos los siguientes valores, y al evaluar los residuales vemos además que no hay autocorrelación y los valores se ajustan una distribución normal.

```

> accuracy(rFesb)

```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-6.814877	49.5774	31.65858	-17.20647	50.21122	0.7812915	-0.03105801

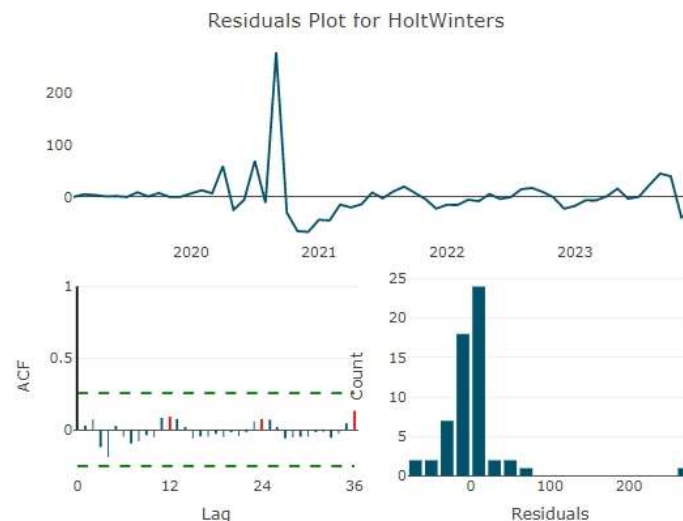
```

Ljung-Box test

data: Residuals from HoltWinters
Q* = 3.5385, df = 12, p-value = 0.9904

Model df: 0. Total lags used: 12

```



Al evaluar los tres modelos vemos que los valores son muy parecidos por lo que podemos elegir el modelo simple.

Simple

```

> accuracy(rFesb)

```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	2.312506	43.25182	20.66961	-1.09283	33.43832	0.5100983	0.02684347

Doble

```

> accuracy(rFesb)

```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	2.352315	43.62308	21.02598	-1.092941	33.4382	0.5188931	0.02688175

Triple

```

> accuracy(rFesb)

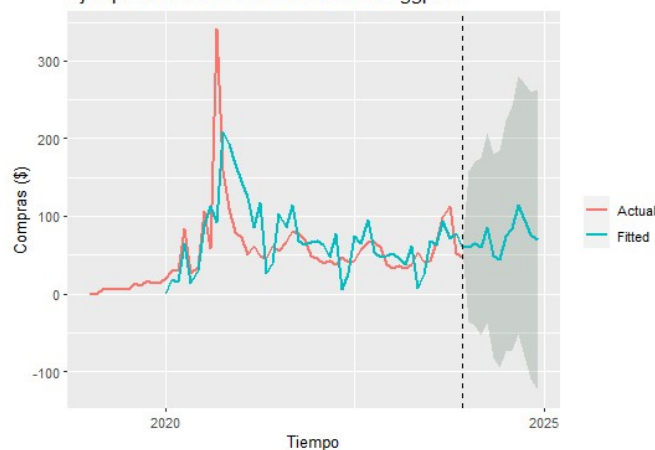
```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-6.814877	49.5774	31.65858	-17.20647	50.21122	0.7812915	-0.03105801

Prediciendo el siguiente año tenemos:

	fit	upr	lwr	> forecast(fit1,12)					
				Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2024	60.67063	157.9367	-36.59543	Jan 2024	60.67063	-2.928227	124.2695	-36.59543	157.9367
Feb 2024	64.34646	170.0221	-41.32919	Feb 2024	64.34646	-4.751134	133.4441	-41.32919	170.0221
Mar 2024	60.22730	174.3516	-53.89704	Mar 2024	60.22730	-14.394598	134.8492	-53.89704	174.3516
Apr 2024	84.78209	207.4120	-37.84781	Apr 2024	84.78209	4.598704	164.9655	-37.84781	207.4120
May 2024	49.33383	180.5393	-81.87164	May 2024	49.33383	-36.456813	135.1245	-81.87164	180.5393
Jun 2024	44.12514	183.9860	-95.73572	Jun 2024	44.12514	-47.324962	135.5752	-95.73572	183.9860
Jul 2024	74.85090	223.4544	-73.75257	Jul 2024	74.85090	-22.315689	172.0175	-73.75257	223.4544
Aug 2024	84.38469	241.8236	-73.05425	Aug 2024	84.38469	-18.559101	187.3285	-73.05425	241.8236
Sep 2024	114.26308	280.6346	-52.10846	Sep 2024	114.26308	5.478576	223.0476	-52.10846	280.6346
Oct 2024	95.31709	270.7216	-80.08744	Oct 2024	95.31709	-19.373767	210.0080	-80.08744	270.7216
Nov 2024	75.34910	259.8895	-109.19126	Nov 2024	75.34910	-45.315355	196.0136	-109.19126	259.8895
Dec 2024	69.54981	263.3307	-124.23106	Dec 2024	69.54981	-57.156694	196.2563	-124.23106	263.3307

Ejemplo de Predicciones Holt-Winters ggplot2



Para validar

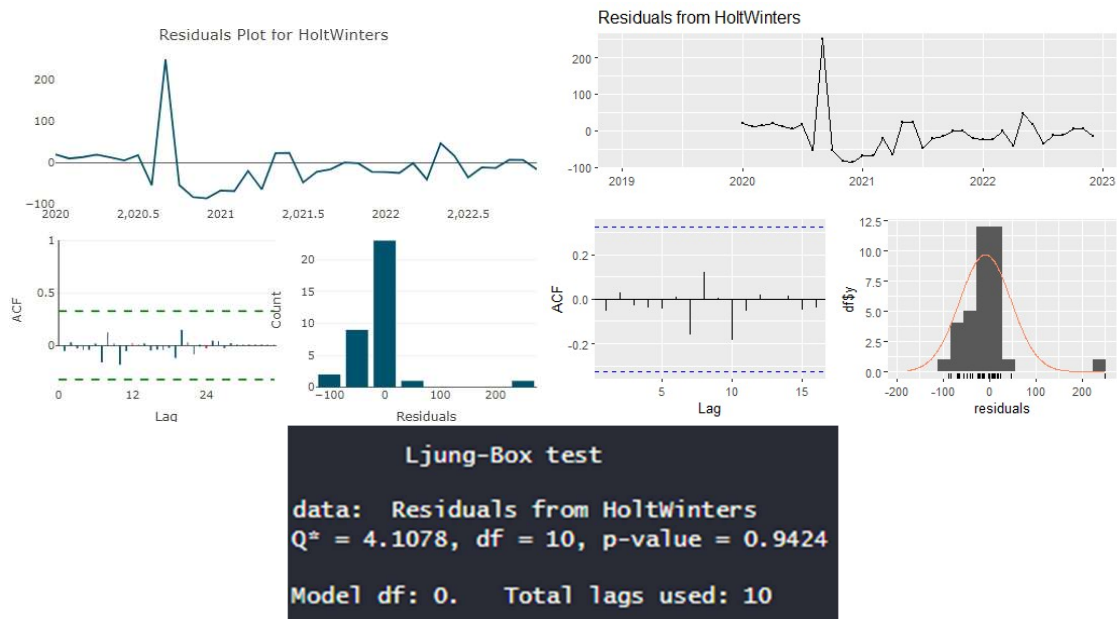
Particionamos la información

```

The train series is a ts object with 1 variable and 48 observations
Frequency: 12
Start time: 2019 1
End time: 2022 12
> ts_info(test)
The test series is a ts object with 1 variable and 12 observations
Frequency: 12
Start time: 2023 1
End time: 2023 12

> accuracy(fc1, test)
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1 Theil's U
Training set -8.760712 55.39876 34.76839 -19.60096 52.77009 0.6930575 -0.05348417 NA
Test set     -6.897862 23.23256 21.16365 -24.27179 45.71275 0.4218667 0.00201613 1.790693

```

Al evaluar los residuales vemos además que no hay autocorrelación y los valores se ajustan una distribución normal.

Regresión en la serie de tiempo

```

Coefficients:
(Intercept) 19.7354 28.3537 0.696 0.4910
trend       0.7771  0.5647 1.376 0.1775
season2     -5.5271 37.1207 -0.149 0.8825
season3     -1.8042 37.1336 -0.049 0.9615
season4      7.1687 37.1550  0.193 0.8481
season5     -6.3583 37.1851 -0.171 0.8652
season6     -2.6354 37.2236 -0.071 0.9440
season7     13.0875 37.2707  0.351 0.7276
season8      9.0604 37.3263  0.243 0.8096
season9     83.7833 37.3903  2.241 0.0315 *
season10    38.2563 37.4628  1.021 0.3142
season11    20.7292 37.5435  0.552 0.5844
season12     1.7021 37.6326  0.045 0.9642
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 52.49 on 35 degrees of freedom
Multiple R-squared:  0.2803,    Adjusted R-squared:  0.03351
F-statistic: 1.136 on 12 and 35 DF,  p-value: 0.3645

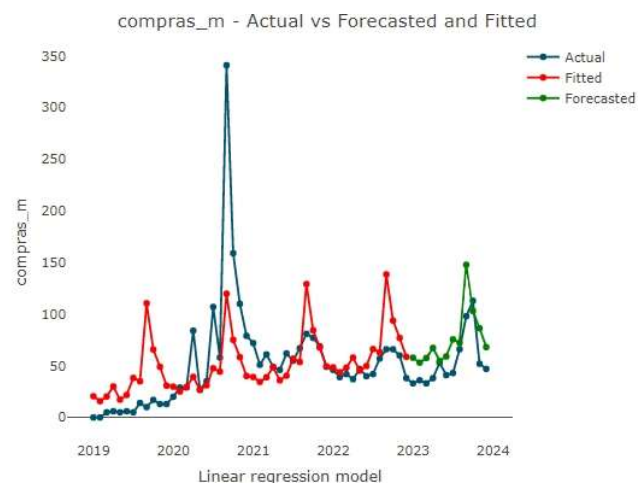
```

Al analizar la regresión vemos que el modelo no estima tan bien ya que el r^2 no es muy alto

Al analizar las medidas de error tenemos

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	5.181764e-16	44.82231	26.73281	-Inf	Inf	0.5328800	0.4870784	NA
Test set	-2.079167e+01	25.81860	22.44792	-46.60696	48.07267	0.4474668	-0.7236439	1.691838

Realizando un pronóstico obtenemos



Al checar los residuos de la regresión vemos que no están dentro de lo esperado por lo que hay autocorrelación, aunque los residuales se ajustan a una distribución normal.

```
Breusch-Godfrey test for serial correlation of order up to 16  
data: Residuals from Linear regression model  
LM test = 24.302, df = 16, p-value = 0.08312
```

