

Forecastig de Serie de Tiempo de Jugadores de CS:GO

Ricardo García Espinosa

2024-11-20

Introducción

Este documento presenta un análisis de los jugadores diarios del juego Counter Strike: Global Offensive, utilizando datos desde junio de 2012 hasta agosto de 2024.

```
## [1] ".....Players....."
```

| ## | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | NA's |
|----|------|---------|--------|--------|---------|---------|------|
| ## | 204 | 515136 | 652278 | 691169 | 939625 | 1818773 | 163 |

```
## [1] ".....Players.av....."
```

| ## | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | NA's |
|----|--------|---------|--------|--------|---------|---------|------|
| ## | 540140 | 809913 | 864647 | 860693 | 929433 | 1306150 | 4790 |

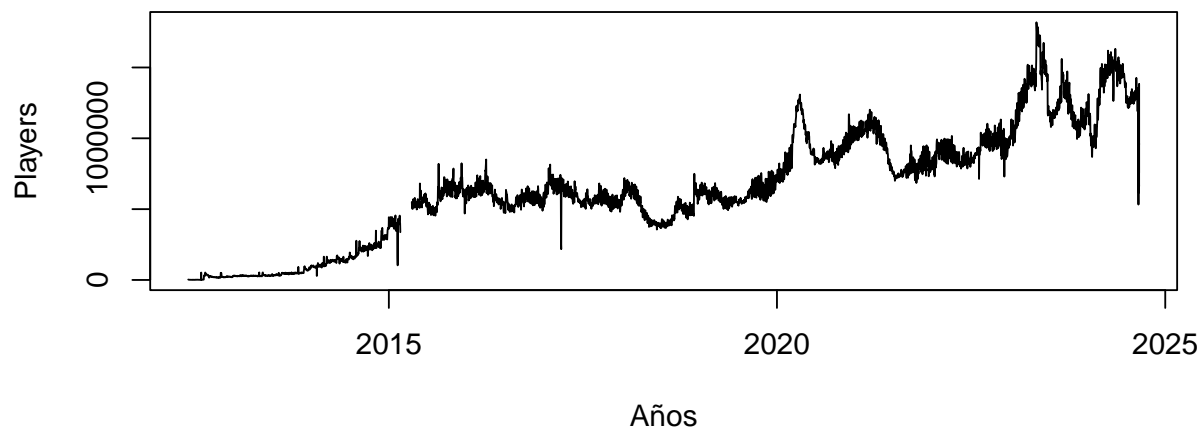
```
## [1] "....."
```

```
## Inicio de la serie: 2011-11-30 00:00:00
```

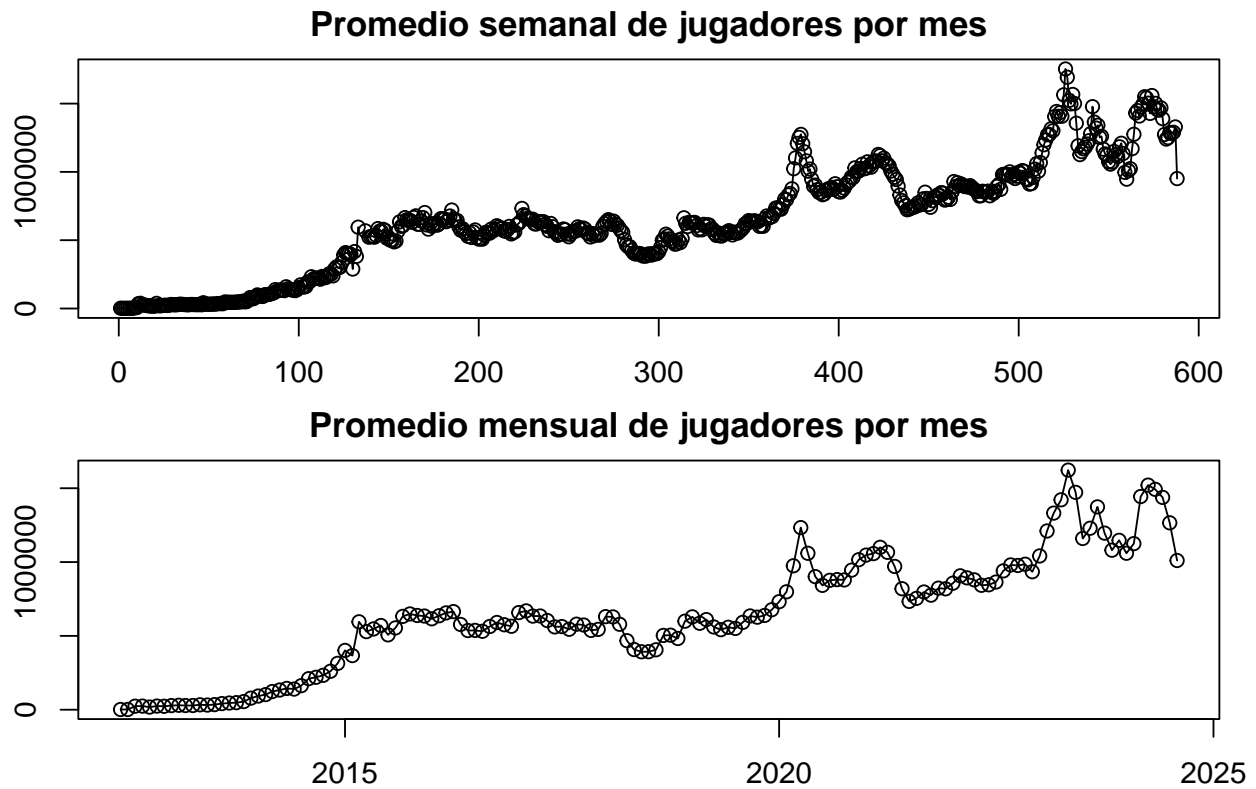
```
## Fin de la serie: 2024-09-27 17:40:00
```

Se analiza la relación entre el la fecha y el promedio de jugadores. Aunque la fecha oficial de lanzamiento fue en Agosto del 2012, existió acceso anticipado desde diciembre del 2011.

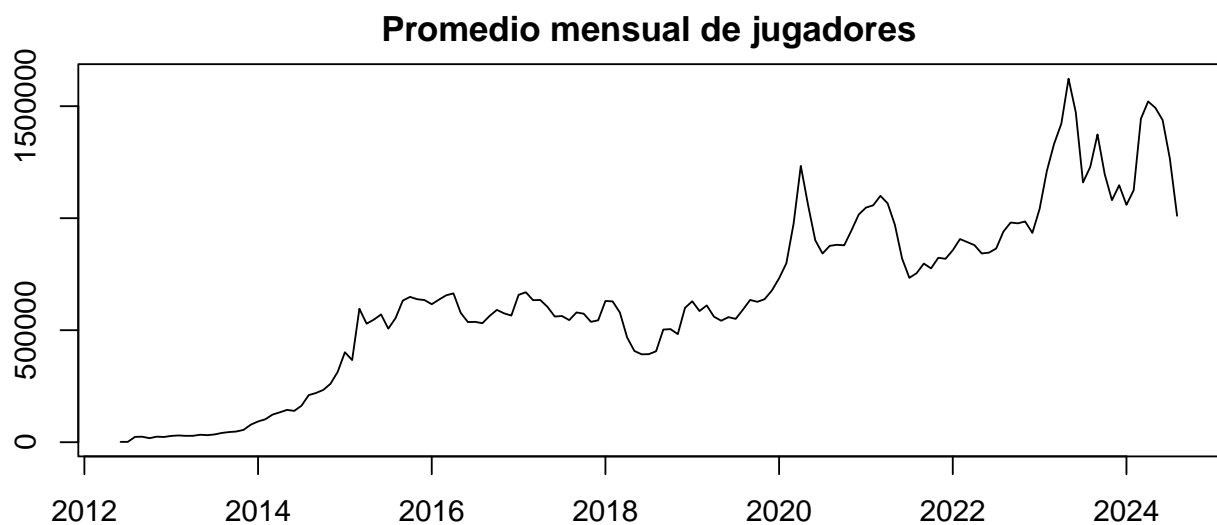
Observaciones diarias



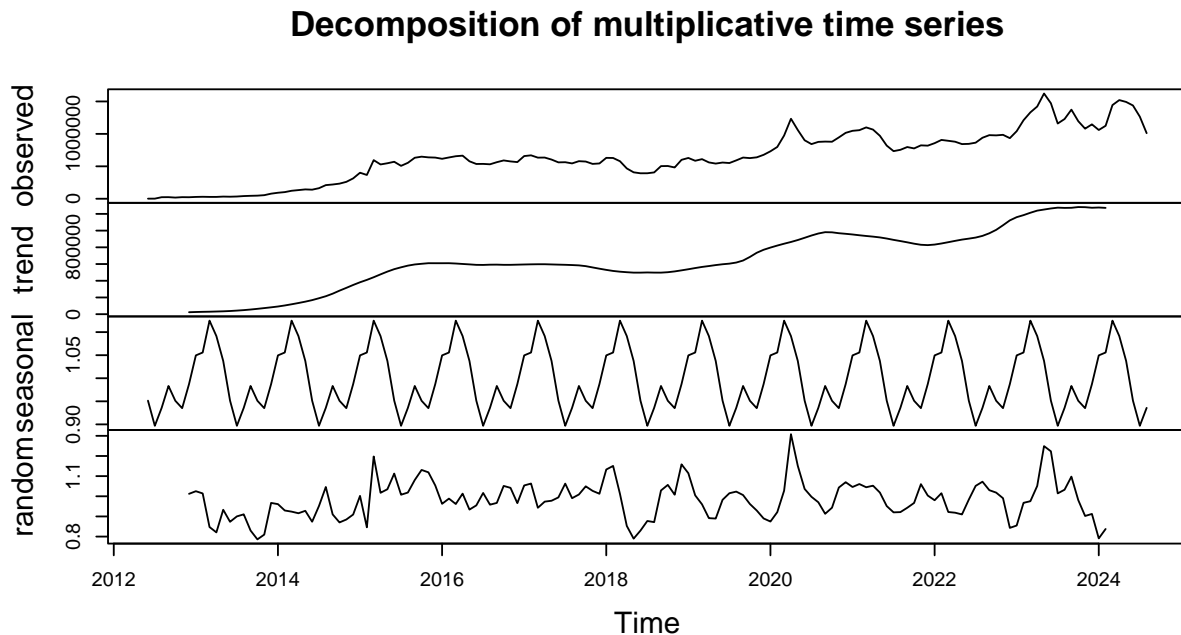
Se puede notar que, hay partes en donde no hay datos, para facilidad del análisis, se calculan los promedios semanales y mensuales por año, como el comportamiento de la serie semanal no se ve afectado en el promedio mensual será plausible trabajar con los promedios mensuales.



Analizando el promedio mensual de jugadores



Descomposición de la Serie:



Esta serie consta de 12 años de observaciones mensuales y parece existir una tendencia creciente, Se observa una posible varianza no constante y no aparenta estacionariedad. El efecto estacional parece aumentar con el nivel de la serie, lo que sugiere una posible relación multiplicativa entre la tendencia y la estacionalidad.

A continuación se verifica estadísticamente.

Para la estacionalidad usar la prueba de Kruskal-Wallis, con $\alpha = 0.05$, buscamos rechazar H_0 , p-value = 0.00017. Entonces:

```
## ### Resultados de la Prueba de Kruskall Wallis
##
## - **Prueba**: Kruskall-Wallis (kw)
## - **Resultado**: Hay evidencia de estacionalidad
```

Con *bptest*, *ncvTest* y con $\alpha = 0.05$ buscamos no rechazar H_0 para que sea plausible la Homocedasticidad. Obtenemos P-values menores a α . Por lo tanto :

Bajo la prueba de *bptest* se tiene que

```
## ### Resultados de la Prueba de Breusch-Pagan
##
## - **Prueba**: Breusch-Pagan (BP)
## - **Resultado**: Es heterocedástica
```

Bajo la prueba de *ncvTest* se tiene que

```
## ### Resultados de la Prueba de Non-Constant Variance
##
## - **Prueba**: Non-Constant Variance (ncv)
## - **Resultado**: Es heterocedástica
```

Hay evidencia en contra de la homocedasticidad en ambas pruebas.

Para ver si hay estacionariedad se usa la Prueba Dickey-Fuller Aumentada (ADF), buscamos rechazar H_0 para asegurar estacionariedad. Esta es robusta para decirnos si cuenta con varianza y media constantes.

```
## ### Resultados de la Prueba de Dickey-Fuller Aumentada
##
## - **Prueba**: Dickey-Fuller Aumentada (ADF)
## - **Resultado**: No es estacionaria
```

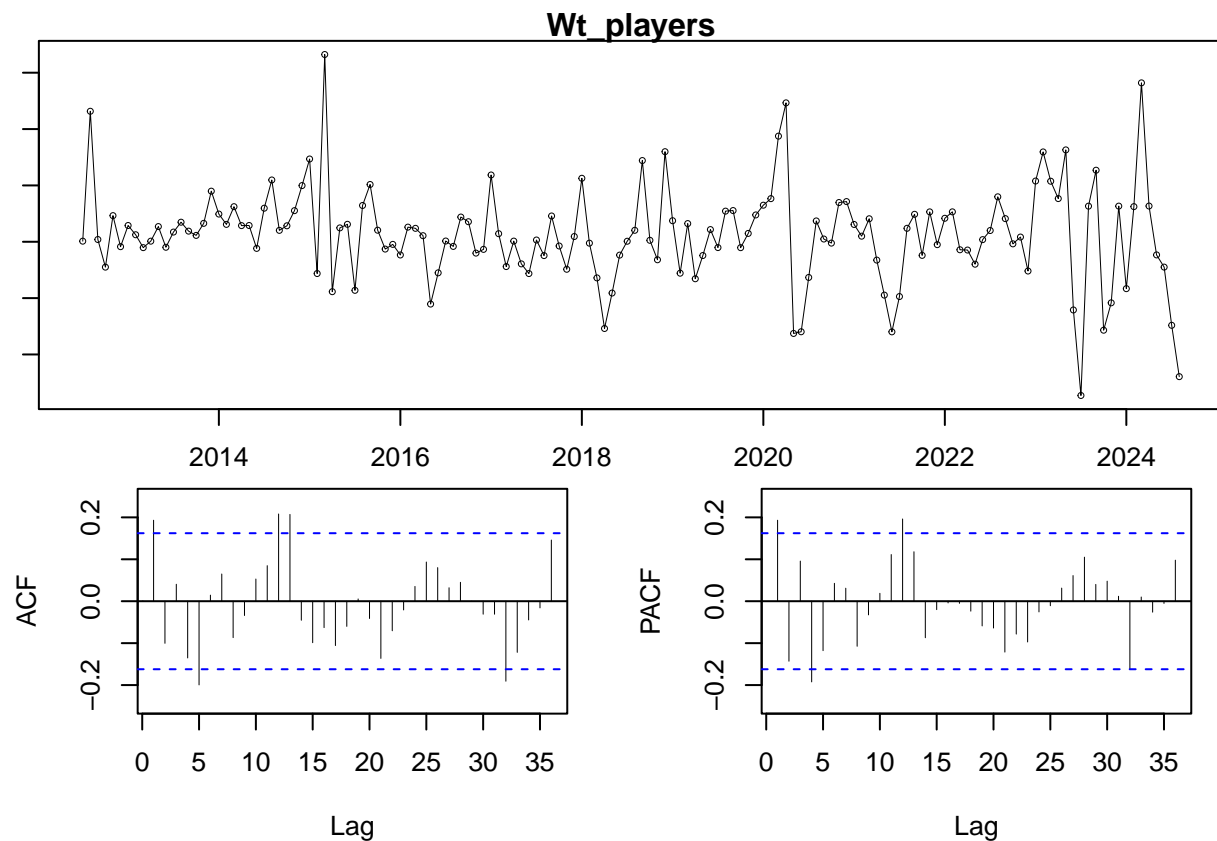
Hay que realizar transformación a los datos. Mediante una transformación de Box-cox, veremos cual es la conveniente.

Box Cox da un valor $\lambda = 0.3418106$, por lo que aplicar raíz a los datos es la transformación adecuada. Con los datos transformados, comprobamos si ahora se existe estacionariedad, con la misma prueba de antes, llegamos a que

```
## ### Resultados de la Prueba de Breusch-Pagan
##
## - **Prueba**: Breusch-Pagan (BP)
## - **Resultado**: Es heterocedástica
##
## ### Resultados de la Prueba de Non-Constant Variance
##
## - **Prueba**: Non-Constant Variance (ncv)
## - **Resultado**: Es heterocedástica
##
## ### Resultados de la Prueba de Dickey-Fuller Aumentada
##
## - **Prueba**: Dickey-Fuller Aumentada (ADF)
## - **Resultado**: No es estacionaria
```

Ahora aplicaremos las diferenciaciones necesarias para estabilizar la varianza.

| d | D | varianza | |
|---|---|----------|----------|
| 6 | 1 | 0 | 1839.812 |



```
## [1] "Hay evidencia de estacionariedad"
```

La serie transformada (con raíz cuadrada y una diferenciación) parece haberse estabilizado bien, eliminando cualquier tendencia. No se observa un patrón claro de estacionalidad.

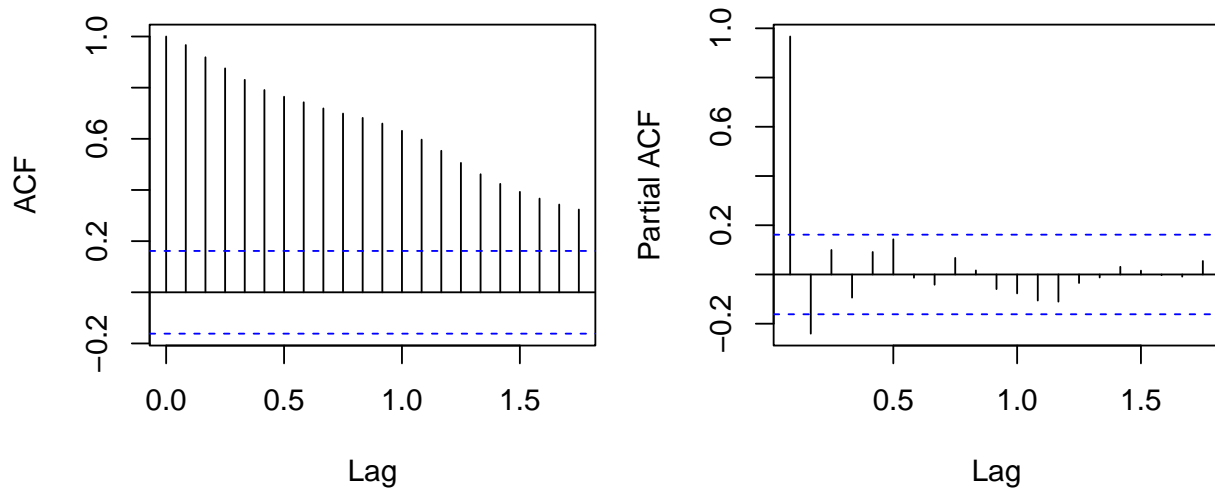
En el ACF, solo unos pocos lags cruzan los límites de significancia, lo que indica que la autocorrelación no es fuerte y la serie parece suficientemente diferenciada. Además de un pico en el primer lag, posiblemente de un valor $q = 1$.

El PACF muestra picos en los primeros lags, sugiriendo un componente autoregresivo con un pequeño valor de p (posiblemente $p = 1$ o $p = 2$).

La serie podría ser modelada con un ARIMA básico, ya que $d = 1$ parece suficiente.

Modelación

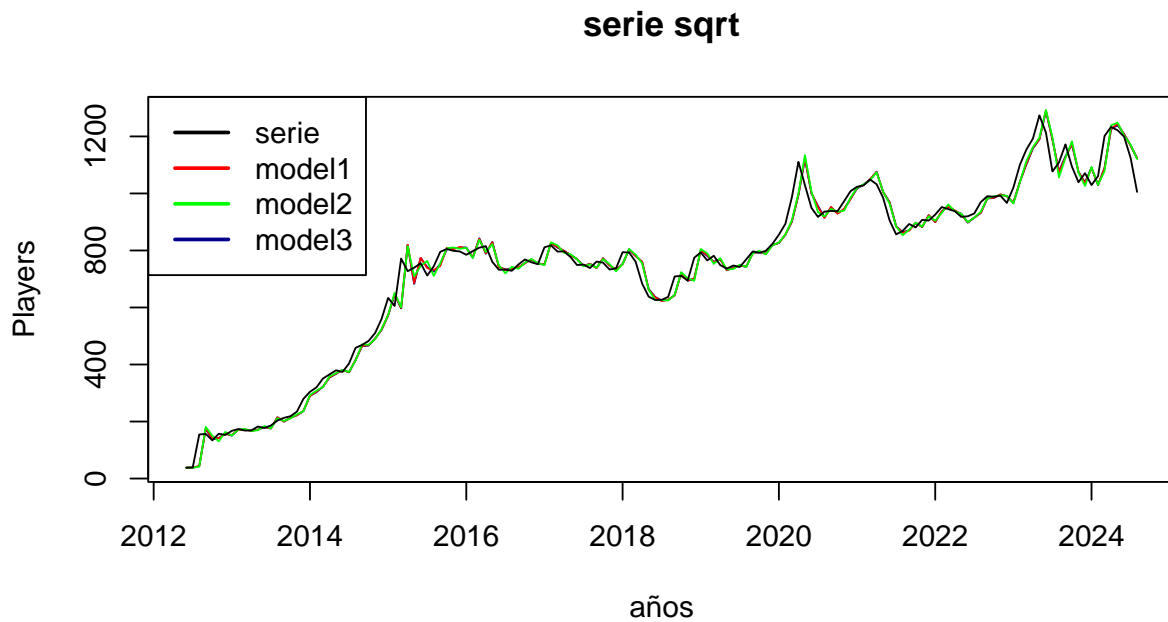
Primero veamos los correlogramas:



Se ajustan distintos modelos:

1. model1: $SARIMA(1, 1, 1)X(1, 0, 0)_{12}$
2. model2: $SARIMA(1, 1, 1)X(1, 0, 0)_{12}$
3. model3: $SARIMA(2, 1, 1)X(1, 0, 0)_{12}$

Gráfico de los modelos.



BIC por modelo

```
## [1] "BIC model1 1516.13214288903"
```

```
## [1] "BIC model2 1515.03578342066"
```

```
## [1] "BIC model3 1520.99188195147"
```

El modelo con menor BIC es el primer modelo, veamos como se comporta el RMSE.

RMSE de cada modelo

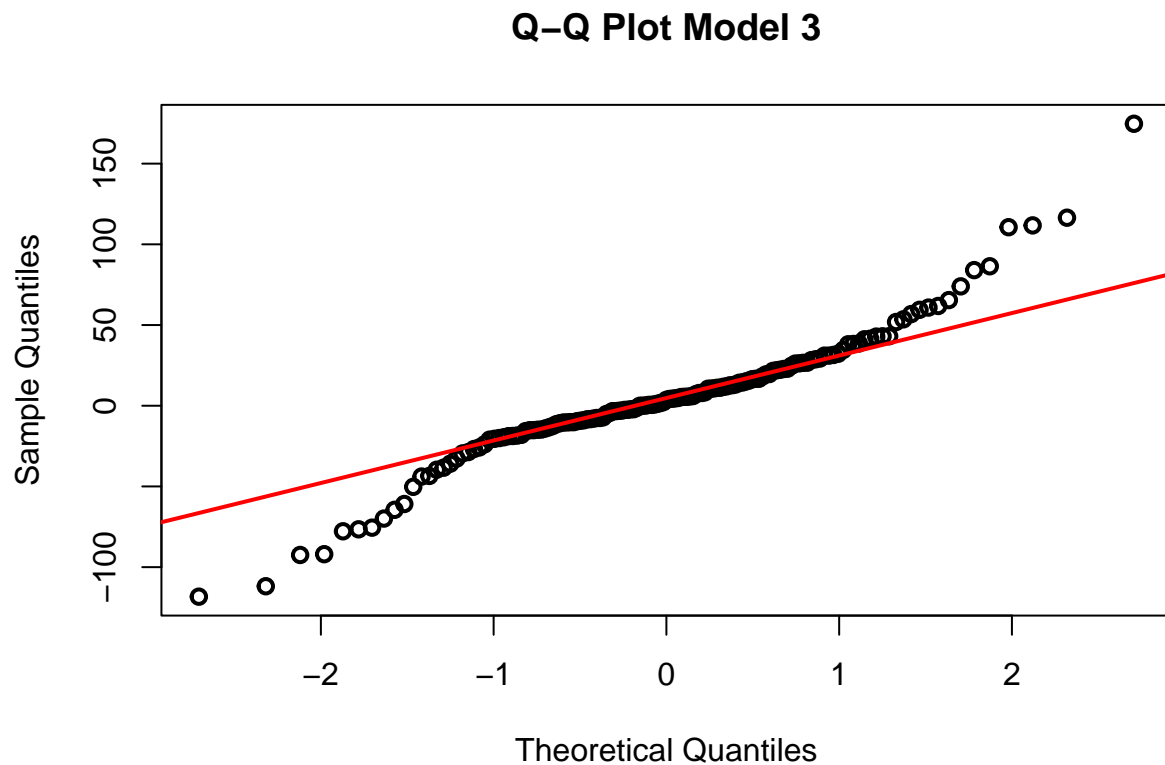
```
## [1] "RMSE model1 40.4109627483548"
```

```
## [1] "RMSE model2 40.977675618795"
```

```
## [1] "RMSE model3 40.3936707898227"
```

Usaremos model3 para las predicciones, veamos primero si cumple los supuestos.

Verificacion de supuestos:



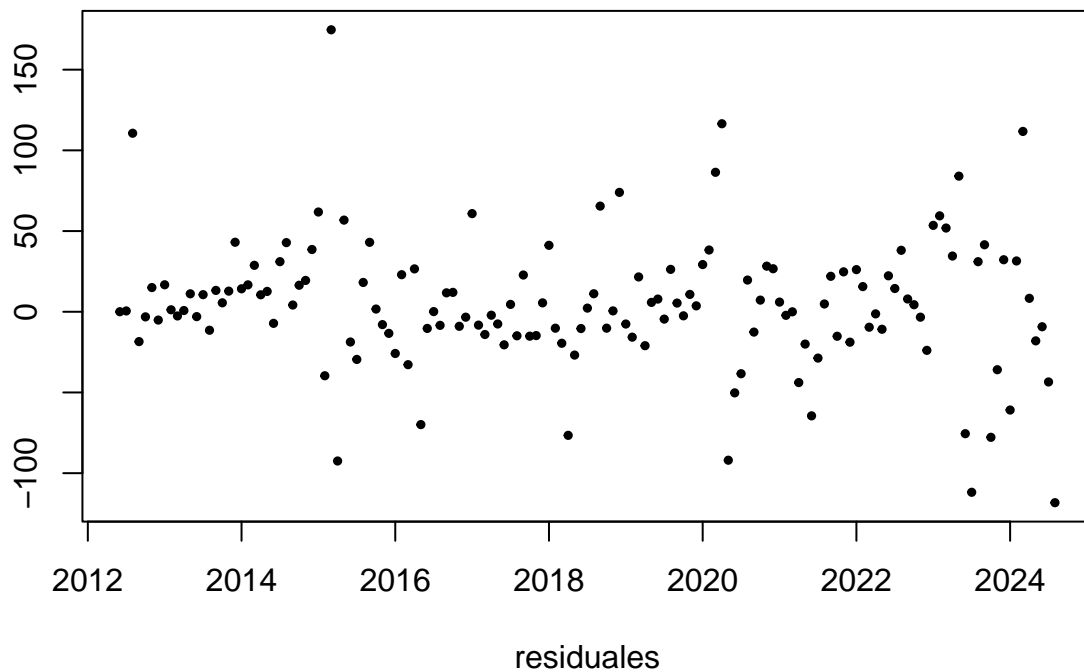
```
##
## Anderson-Darling normality test
##
## data:  model3$residuals
## A = 2.8962, p-value = 2.583e-07

##
## Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  model3$residuals
## D = 0.11409, p-value = 7.427e-05
```

El modelo no cumple la normalidad, sin embargo por teorema de límite central no importa mucho.

Homocedasticidad

Gáfica de residuales model 3



```
##
## studentized Breusch-Pagan test
##
## data:  Y3 ~ X3
## BP = 2.8521, df = 1, p-value = 0.09125
```

El supuesto de homocedasticidad se cumple.

Independencia

```
##
## Box-Pierce test
##
## data: model3$residuals
## X-squared = 0.022092, df = 1, p-value = 0.8818

##
## Box-Pierce test
##
## data: model3$residuals
## X-squared = 0.043491, df = 2, p-value = 0.9785

##
## Box-Pierce test
##
## data: model3$residuals
## X-squared = 0.1378, df = 3, p-value = 0.9869

##
## Box-Pierce test
##
## data: model3$residuals
## X-squared = 0.82952, df = 4, p-value = 0.9344

##
## Box-Pierce test
##
## data: model3$residuals
## X-squared = 9.3202, df = 12, p-value = 0.6754

##
## Box-Pierce test
##
## data: model3$residuals
## X-squared = 14.558, df = 13, p-value = 0.3357

##
## Box-Pierce test
##
## data: model3$residuals
## X-squared = 20.479, df = 24, p-value = 0.6692

##
## Box-Pierce test
##
## data: model3$residuals
## X-squared = 20.541, df = 25, p-value = 0.7179

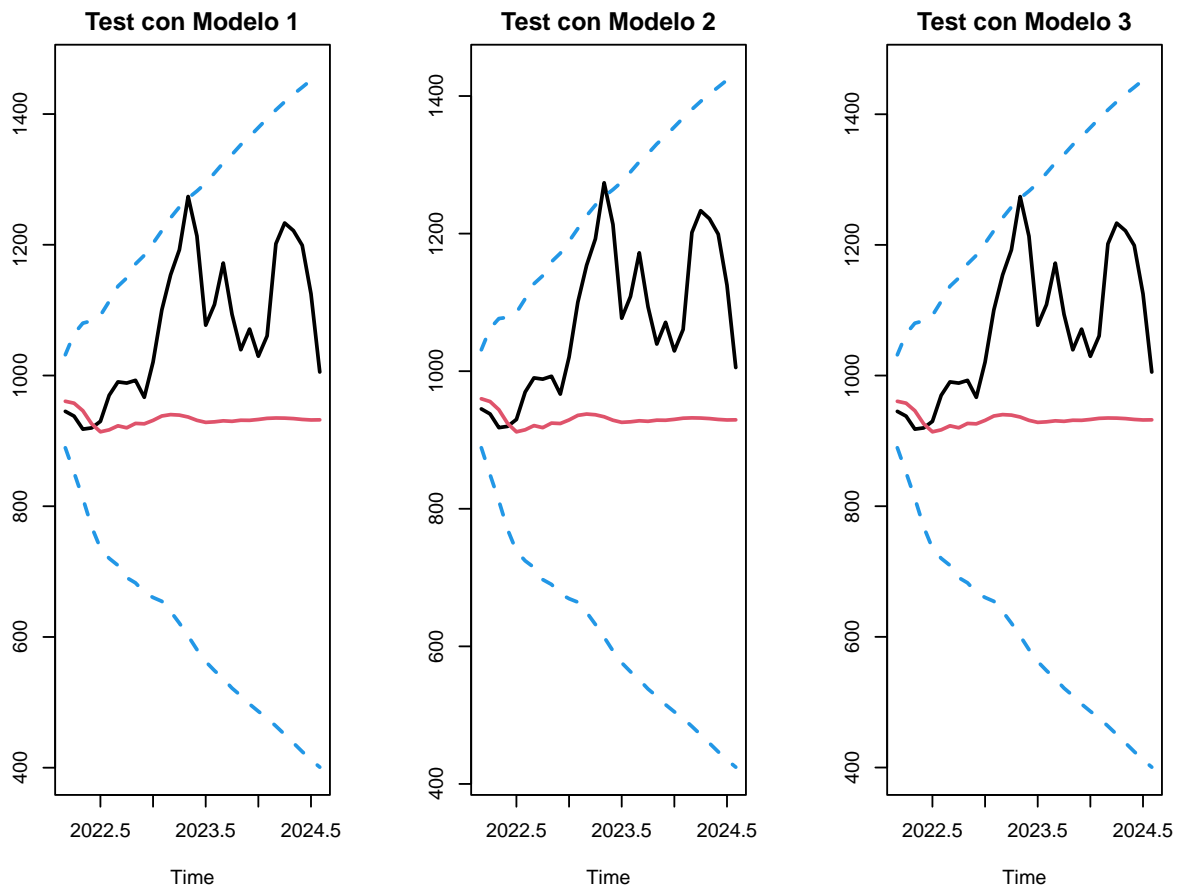
##
## Box-Pierce test
```

```
##
## data: model3$residuals
## X-squared = 21.17, df = 26, p-value = 0.7331
```

Se verifica que se cumple la independencia.

Predicciones

Aunque los tres modelos ajustados no cumplen la normalidad, si cumplen los demás supuestos. Se usa model3 pues tiene menor RMSE en las predicciones.

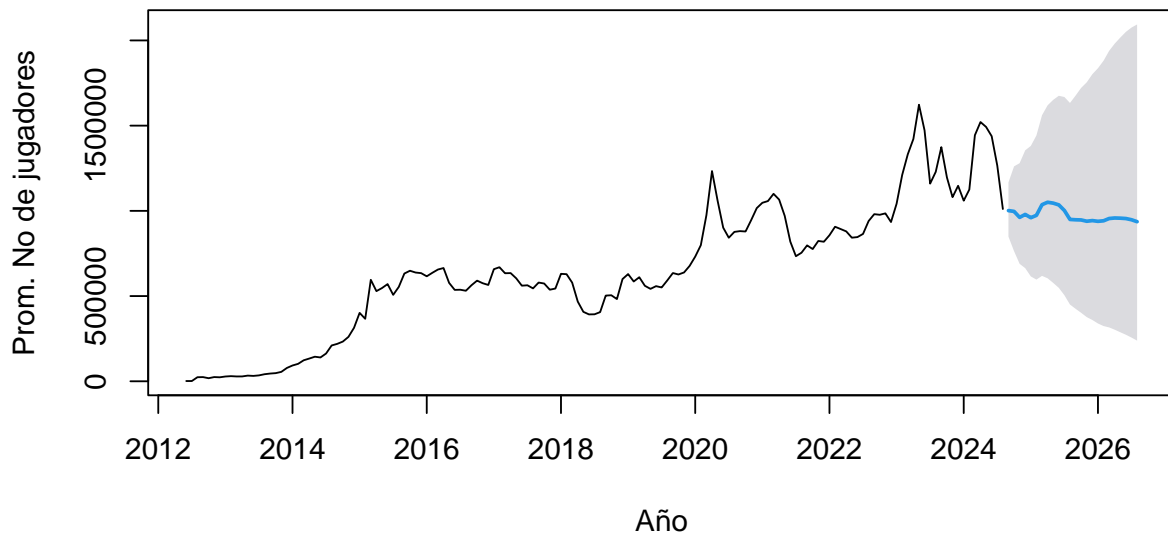


```
## [1] "RMSE model1t 174.092730385252"
```

```
## [1] "RMSE model2t 145.586598017503"
```

```
## [1] "RMSE model3t 143.555328704296"
```

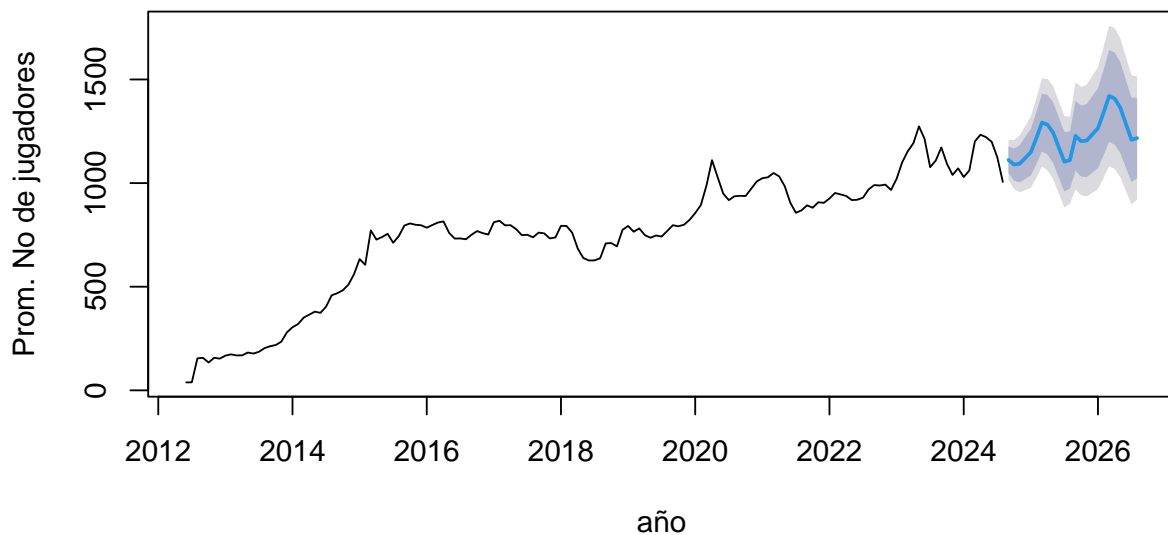
Predicción con el modelo3 a 2 años



En conclusión vemos que nuestras predicciones tenderán a ser decrecientes, pero gracias a los intervalos de confianza podemos ver que no es el mejor modelo que le podemos ajustar

Finalmente al tener intervalos de confianza más grandes se decide aplicar un modelo más sencillo, el cual será el modelo Holt-Winter multiplicativo aplicado a la serie del promedio de jugadores mensuales $Xt_players$

Pronóstico jugadores Holt-Winters multiplicativo



Concluimos que gracias a los intervalos, el modelo Holt-Winters multiplicativo nos sirve mejor para poder hacer las predicciones en este caso a dos años