

Venta de Televisores

Andrea Piñeiro

2022-12-01

Resumen

El análisis de series de tiempo es importante y nos puede ayudar a realizar predicciones. En esta actividad se analizan las ventas de televisores en una tendencia de serie de tiempo.

Para analizar las ventas a lo largo del tiempo y durante los diferentes ciclos en un año.

Introducción

Para analizar la venta de televisores, se observan y analizan las tendencias y estacionalidad, así como el modelo lineal de la misma para poder explicar las ventas a lo largo del tiempo y de los trimestres del año.

Posteriormente realizamos predicciones de la venta de televisores con el modelo, verificando que el modelo sea relevante.

Análisis de los resultados

Carga de Datos

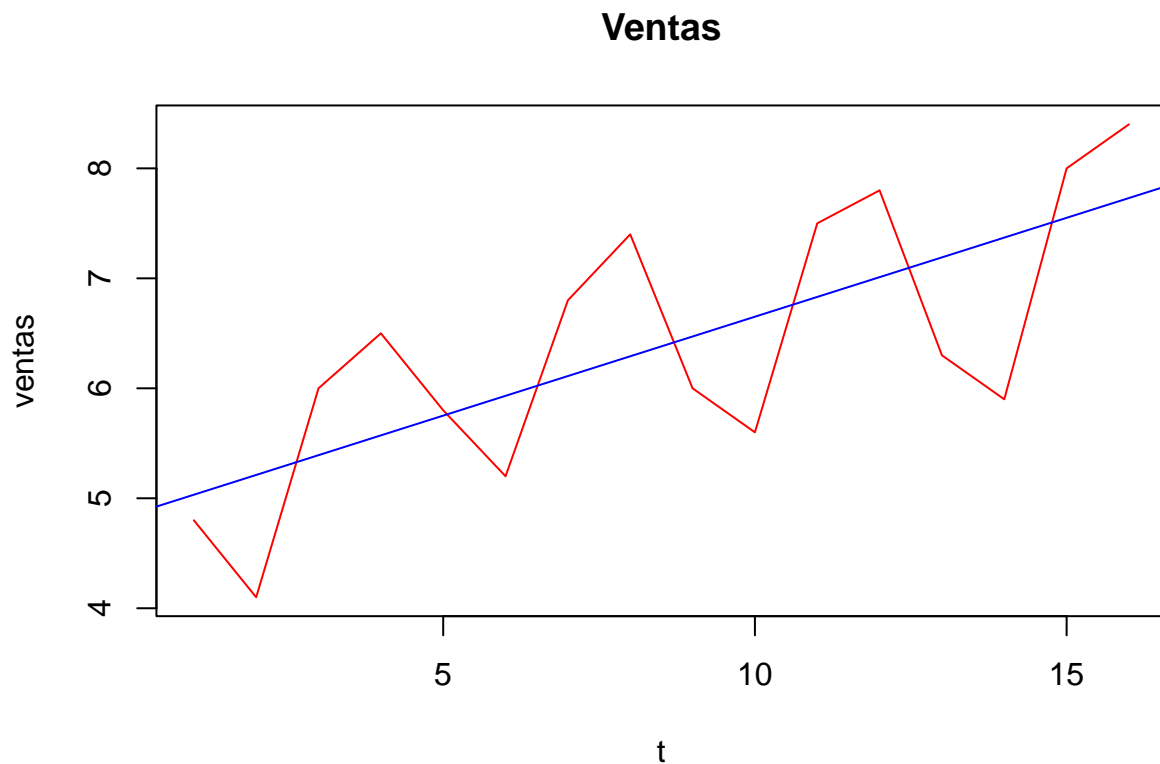
##	año	trimestre	ventas
## 1	1	1	4.8
## 2	1	2	4.1
## 3	1	3	6.0
## 4	1	4	6.5
## 5	2	1	5.8
## 6	2	2	5.2
## 7	2	3	6.8
## 8	2	4	7.4
## 9	3	1	6.0
## 10	3	2	5.6
## 11	3	3	7.5
## 12	3	4	7.8
## 13	4	1	6.3
## 14	4	2	5.9
## 15	4	3	8.0
## 16	4	4	8.4

Creamos la serie de tiempo.

##	Qtr1	Qtr2	Qtr3	Qtr4
## 1	4.8	4.1	6.0	6.5
## 2	5.8	5.2	6.8	7.4
## 3	6.0	5.6	7.5	7.8
## 4	6.3	5.9	8.0	8.4

Gráfico de dispersión. Observa la tendencia y los ciclos.

A continuación mostramos el gráfico de dispersión de las ventas a lo largo del tiempo. En este gráfico se muestra la relación que hay entre el tiempo y las ventas.



Como se puede observar la tendencia general de las ventas de televisores va en aumento a lo largo del tiempo y de los años.

Existen ciclos por cada uno de los Trimestres en el año. 1er Trimestre: Baja un poco a comparación de las ventas del anterior Trimestre 2° Trimestre: Baja aún más las ventas 3° Trimestre: Las ventas suben mucho más de los anteriores trimestres 4° Trimestre: Las ventas suben un poco más del trimestre anterior, es el punto más alto de ventas en el Año.

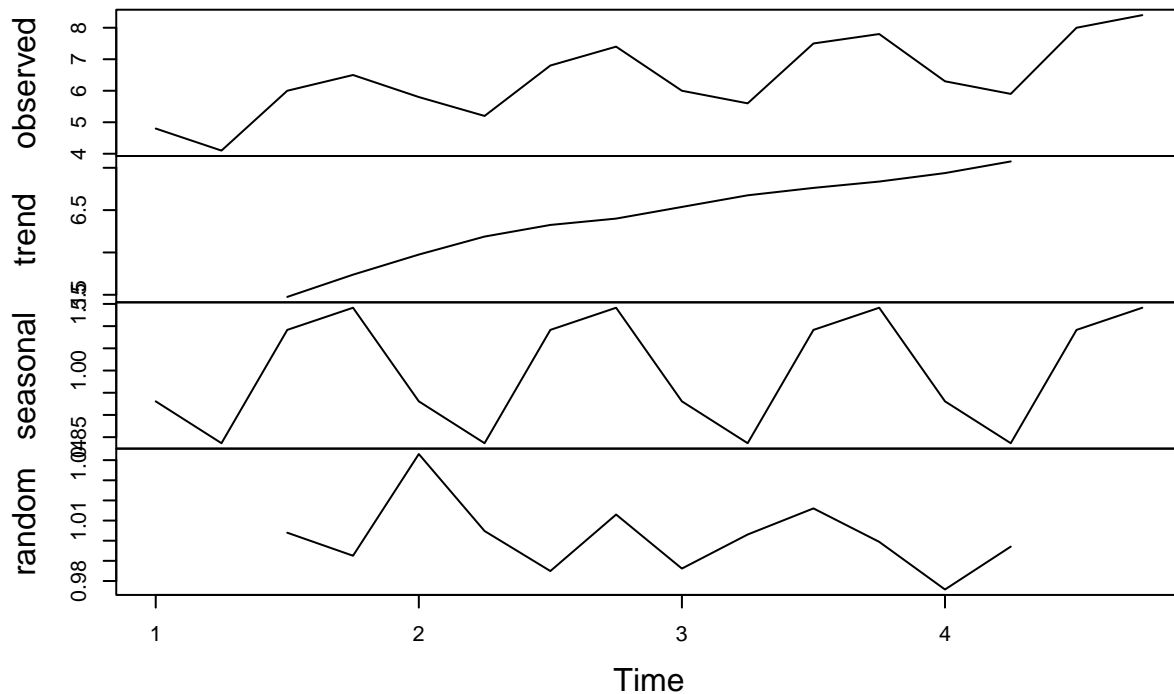
Análisis de tendencia y estacionalidad.

Descompón la serie en sus 3 componentes e interprétalos.

Como se puede observar los 3 componentes son: * seasonal * trend * random

A continuación graficaremos los componentes para interpretarlos de manera más sencilla.

Decomposition of multiplicative time series



* TREND: Como podemos observar la tendencia general a lo largo del tiempo va hacia la alza, por lo que la venta de televisores cada año aumenta.

- SEASONAL: Como podemos observar hay un patrón que se repite cada año. Son 4 puntos (uno para cada trimestre) que se repiten, lo que quiere decir que hay un patrón específico para cada trimestre.
- RANDOM: Son los residuos, la variabilidad que no se puede explicar.

Analiza el modelo lineal de la tendencia

Realiza la regresión lineal de la tendencia (ventas desestacionalizadas vs tiempo)

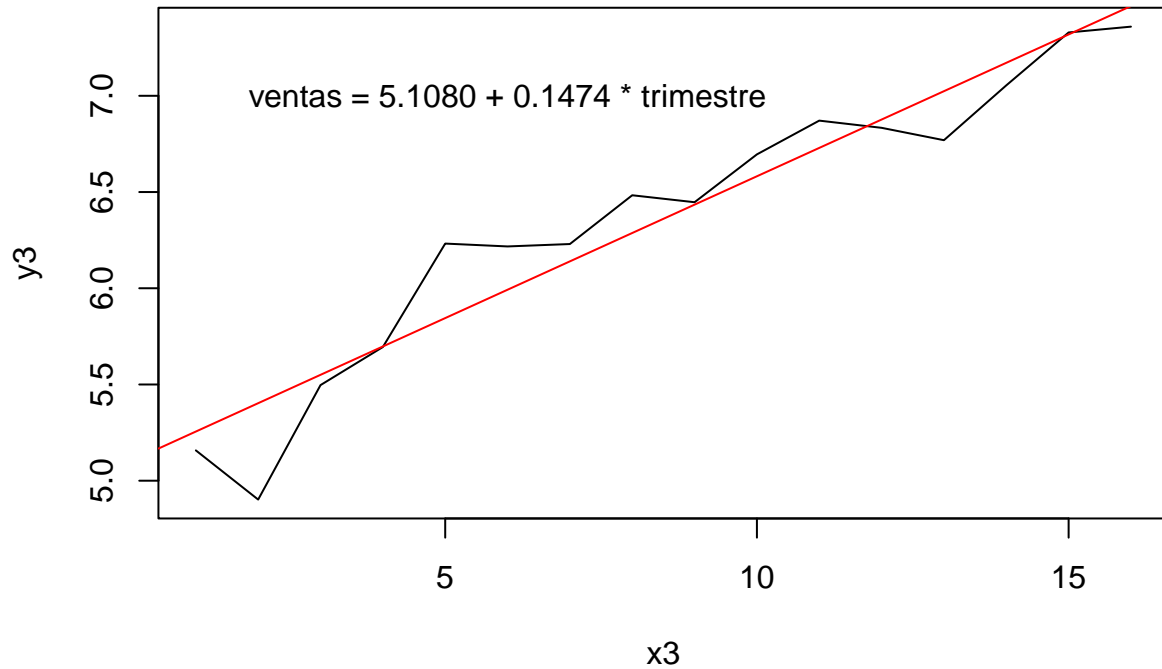
A continuación obtenemos el modelo lineal para explicar las ventas de televisores desestacionalizadas.

```
##
## Call:
## lm(formula = y3 ~ x3)
##
## Coefficients:
## (Intercept)      x3
##      5.1080      0.1474
```

Con esto obtenemos la fórmula de la regresión lineal. Ventas = 5.1080 + 0.1474 * trimestre

Dibuja la recta junto con las ventas desestacionalizadas

A continuación podemos ver una gráfica con la regresión.



Significancia de B1 Tomando un α de 0.05 y las siguientes hipótesis: * H_0 : $B1 = 0$ No hay relación entre las variables. * H_1 : $B1 \neq 0$ hay relación entre las variables.

```
##
## Call:
## lm(formula = y3 ~ x3)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.5007 -0.1001  0.0037  0.1207  0.3872
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.10804    0.11171   45.73  < 2e-16 ***
## x3           0.14738    0.01155   12.76 4.25e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.213 on 14 degrees of freedom
## Multiple R-squared:  0.9208, Adjusted R-squared:  0.9151
## F-statistic: 162.7 on 1 and 14 DF, p-value: 4.248e-09
```

Como se puede observar obtenemos un p-value menor a α para B1, por lo que si existe una relación significativa entre las variables en la regresión lineal.

Variabilidad explicada por el modelo La variabilidad del modelo se puede observar en el punto anterior. Tiene una R cuadrada de 0.92 y una R Caudrada Ajustada de 0.91

Lo que quiere decir que el modelo explica el 92% de la variabilidad, por lo que podemos concluir que este es un buen modelo.

Análisis de los residuos Normalidad de los residuos: H_0 : Los datos provienen de una población normal. H_1 : Los datos no provienen de una población normal.

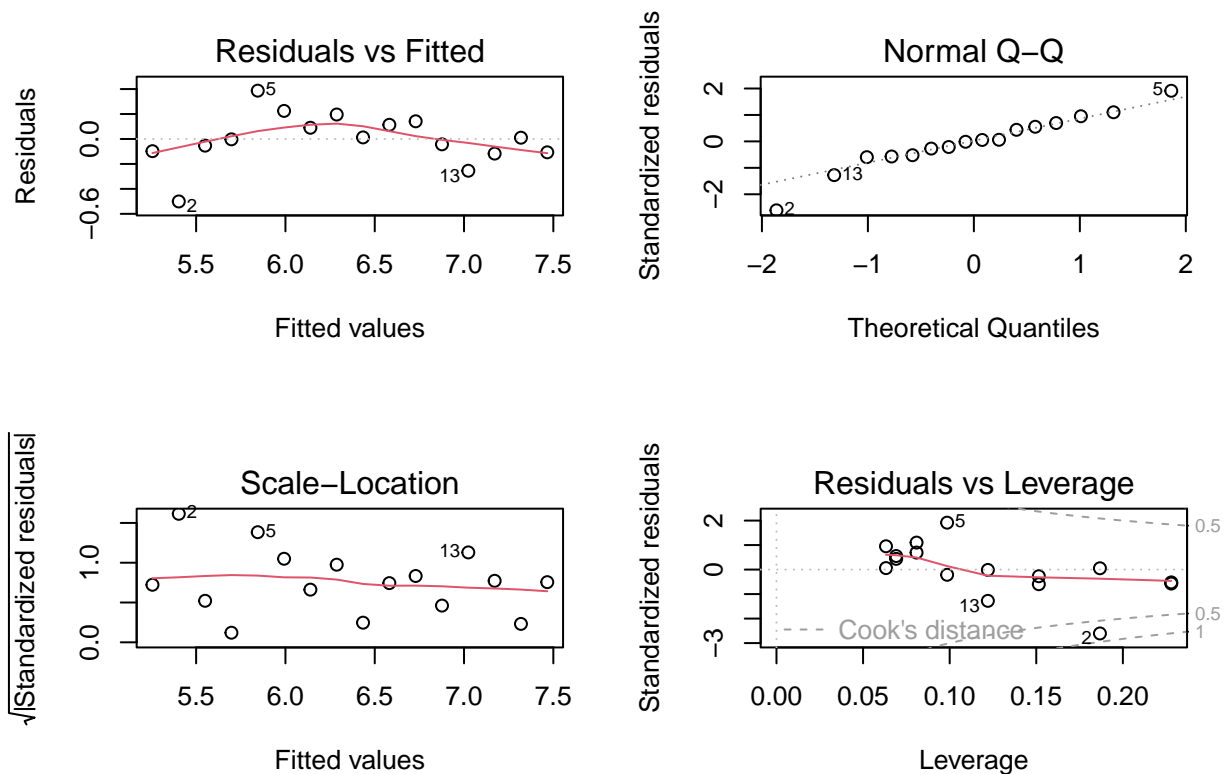
La prueba de Shaprio Wilk test nos ayuda a testear la normalidad para saber si vienen de distribución normal.

```
##
## Shapiro-Wilk normality test
##
## data: N3$residuals
## W = 0.96379, p-value = 0.7307
```

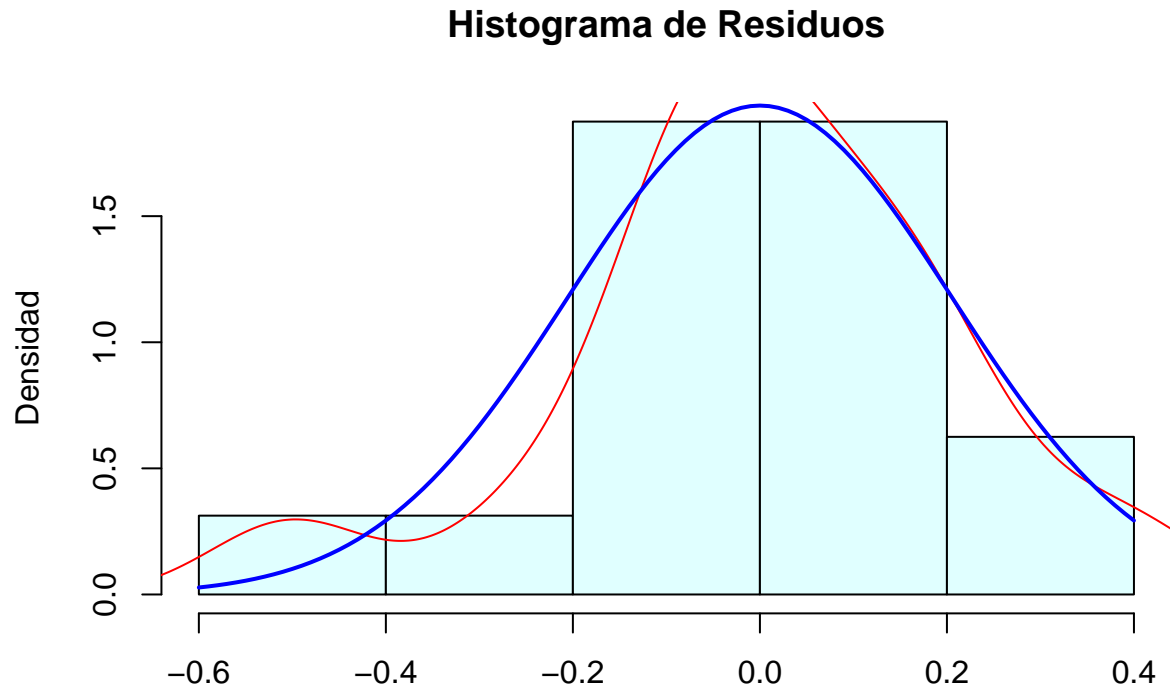
Dado que el p-value no es menor a 0.05 por lo que podemos asumir que los datos provienen de una distribución normal.

Como podemos ver los residuos muestran una *normalidad* casi ideal, por lo que no se rechaza que los datos provienen de una población normal.

homocedasticidad e Independencia



Además se puede observar en la gráfica, los residuos no muestran ninguna estructura evidente. Por lo que el modelo de regresión simple es adecuado y tiene *homocedasticidad*. Y tienen *independencia* pues no tienen un patrón específico.



Los residuos está distribuido bastante bien, tiene un poco de sesgo a la izquierda pero es mínimo.

Prueba de la normalidad

##	Test	Statistic	p value	Result
## 1	Mardia Skewness	5.81205826154813	0.830798314776995	YES
## 2	Mardia Kurtosis	-1.91138314522295	0.0559553610876324	YES
## 3	MVN	<NA>	<NA>	YES

Según Mardia hay normalidad en los datos, debido a que el p value es menor a nuestra α no rechazamos la hipótesis nula.

Por lo que concluimos que si hay normalidad multivariada.

Calcula el CME y el EPAM (promedio de los errores porcentuales) de la predicción de la serie de tiempo.

El CME mide la diferencia cuadrática media entre los valores estimados y los valores reales de la distribución.

El CME es de 40.96092

Como podemos observar el valor de CME es alto, lo que quiere decir que si hay diferencia entre la línea de regresión y el cuadrado entre los puntos.

Concluye sobre el modelo

El modelo es una aproximación buena para pronosticar las ganancias que se tendrá en los siguientes años. Como se pudo observar con la R cuadrada se ajusta correctamente a los valores y además las variables se comportan como normales.

Con el análisis de los residuos pudimos concluir que el modelo es correcto y puede ser usado para este problema ya que los residuos tienen Homocedasticidad, normalidad e independencia.

Realiza el pronóstico para el siguiente año.

A continuación se realizará el pronóstico para el siguiente año por cada uno de los cuatrimestres.

```
## [1] 7085.872
```

```
## [1] 6491.284
```

```
## [1] 8632.585
```

```
## [1] 9195.263
```

Con esto podemos ver que en los trimestres que va a haber más ventas será en el último y el 3° como regularmente.

Conclusión

Al final de la actividad podemos concluir que el modelo obtenido por regresión lineal de ventas decentralizadas es bueno ya que explica la variabilidad de los datos. Podemos usar este modelo para pronosticar las ventas en los siguientes años y en los siguientes trimestres.

Este análisis también nos ayudó a ver que durante los último dos trimestres tenemos mayor venta de televisores, mientras que en el segundo trimestre es el punto más bajo. A partir de esta información se pueden tomar decisiones para beneficiar al negocio, como aumentar o disminuir la producción en estas fechas para no quedarnos con pérdidas.

Se podría buscar otros modelos no lineales para verificar si obtenemos mejores aproximaciones con ellos; sin embargo, este es un muy buen modelo.

Referencias Bibliográficas

Avijeet Biswal. (2020, August 16). A Guide to Time Series Forecasting in R You Should Know. Simplilearn.com; Simplilearn. <https://www.simplilearn.com/tutorials/data-science-tutorial/time-series-forecasting-in-r>

François St-Amant. (2020, June 13). Time Series Forecasting in R - Towards Data Science. Medium; Towards Data Science. <https://towardsdatascience.com/a-guide-to-forecasting-in-r-6b0c9638c261>

Anexos

https://drive.google.com/drive/folders/1G-RTJbDTmjHiZog6dUs6CHwjo7NOes5V?usp=share_link