



## Reporte “Venta de televisores”

Inteligencia artificial avanzada para ciencia de datos II

Modulo 5: Estadística Avanzada para ciencia de datos

Grupo 502

Luis Gabriel Martínez Rentería

A01651812

Fecha de entrega: 04 / 12 / 22

## **Resumen**

El objetivo de este análisis es el elaborar un modelo que permita describir las propiedades de una serie atemporal, de manera que las observaciones que se hicieron coincidan con las pautas muestrales. Se describirá la evolución de las ventas de televisores observadas, se realizarán pronósticos de las ventas en el futuro y podrán contrastarse las variables que refieren a los componentes de dicha serie.

## **Introducción**

El análisis de las series de tiempo nos permite analizar observaciones en un periodo de tiempo, ordenadas cronológicamente y espaciadas entre si a la misma distancia. El objetivo principal de este tipo de análisis es hacer pronósticos basándonos en el comportamiento observado, ya que se espera que las variables se comporten de manera dependiente entre si.

El objetivo principal es encontrar patrones de estacionalidad y efectuar predicciones a un futuro determinado.

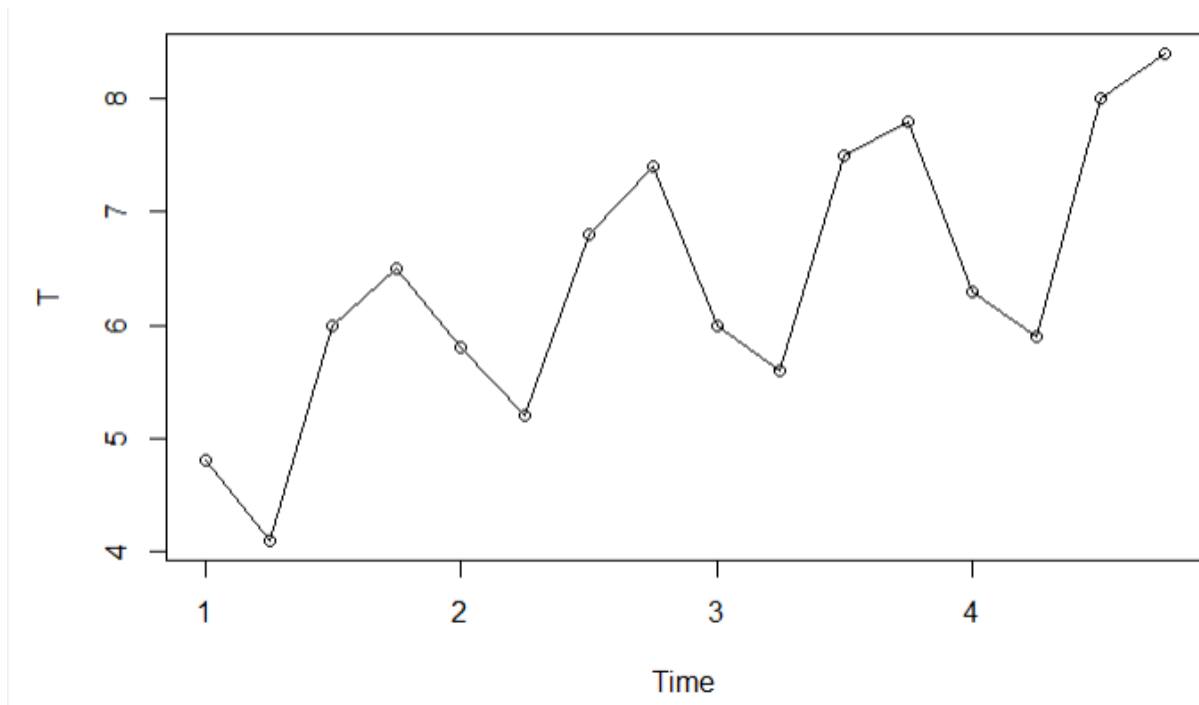
Se basa en la suposición de que los valores que toma la variable observada es la consecuencia de tres componentes:

- Tendencia: Define el cambio a largo plazo de la media
- Estacional: Mide efectos que muestran periodicidad o comportamientos similares en un periodo
- Aleatoria: NO responde a patrones de comportamiento, es el resultado de factores aleatorios

Se considera para este caso que el comportamiento es de una serie no estacionaria, ya que la variabilidad cambia en el tiempo. Los cambios en la media determinan una tendencia a crecer o decrecer a lo largo de la serie de tiempo, por lo que esta no oscila alrededor de un valor constante.

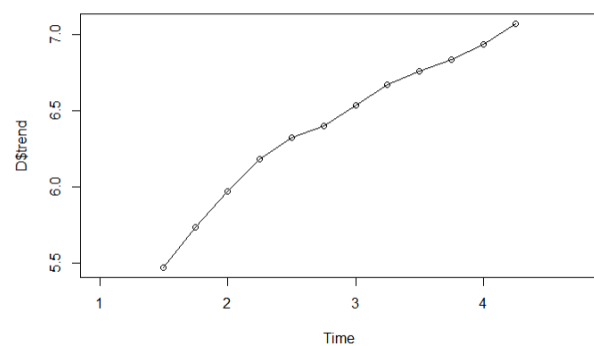
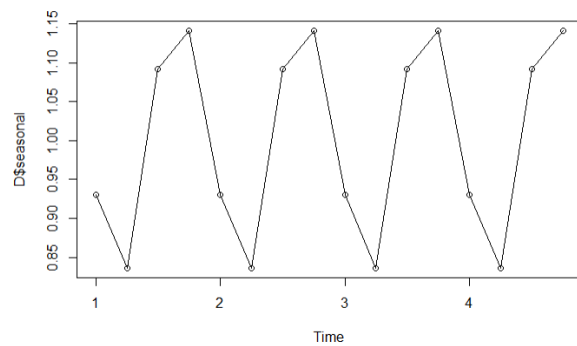
## Análisis de resultados

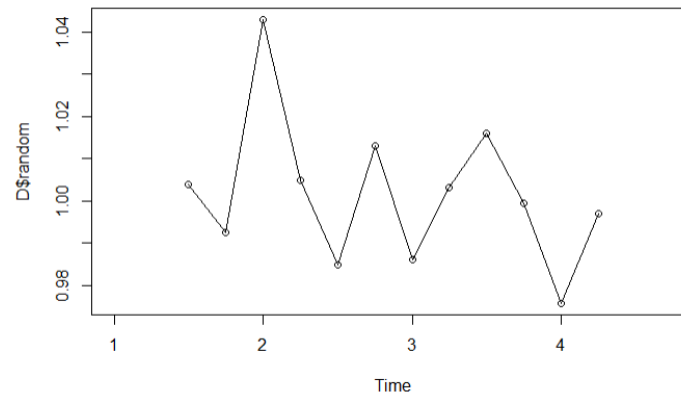
Realiza el gráfico de dispersión. Observa la tendencia y los ciclos



Con esta gráfica se puede observar una tendencia de subida, por lo que podemos asumir que por trimestre va a ir subiendo el número de ventas.

Realiza el análisis de tendencia y estacionalidad





Podemos reafirmar con todas las gráficas que se está trabajando con una serie no estacionaria, ya que no se encuentra un valor por el que se oscile, sino una tendencia activa de los valores a lo largo del tiempo.

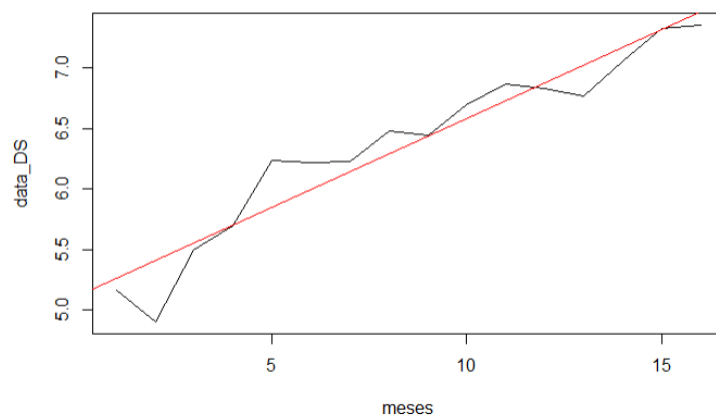
### Analiza el modelo lineal de la tendencia

```
Call:
lm(formula = DS ~ meses)

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 ***
meses        0.14738    0.01155   12.76 4.25e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

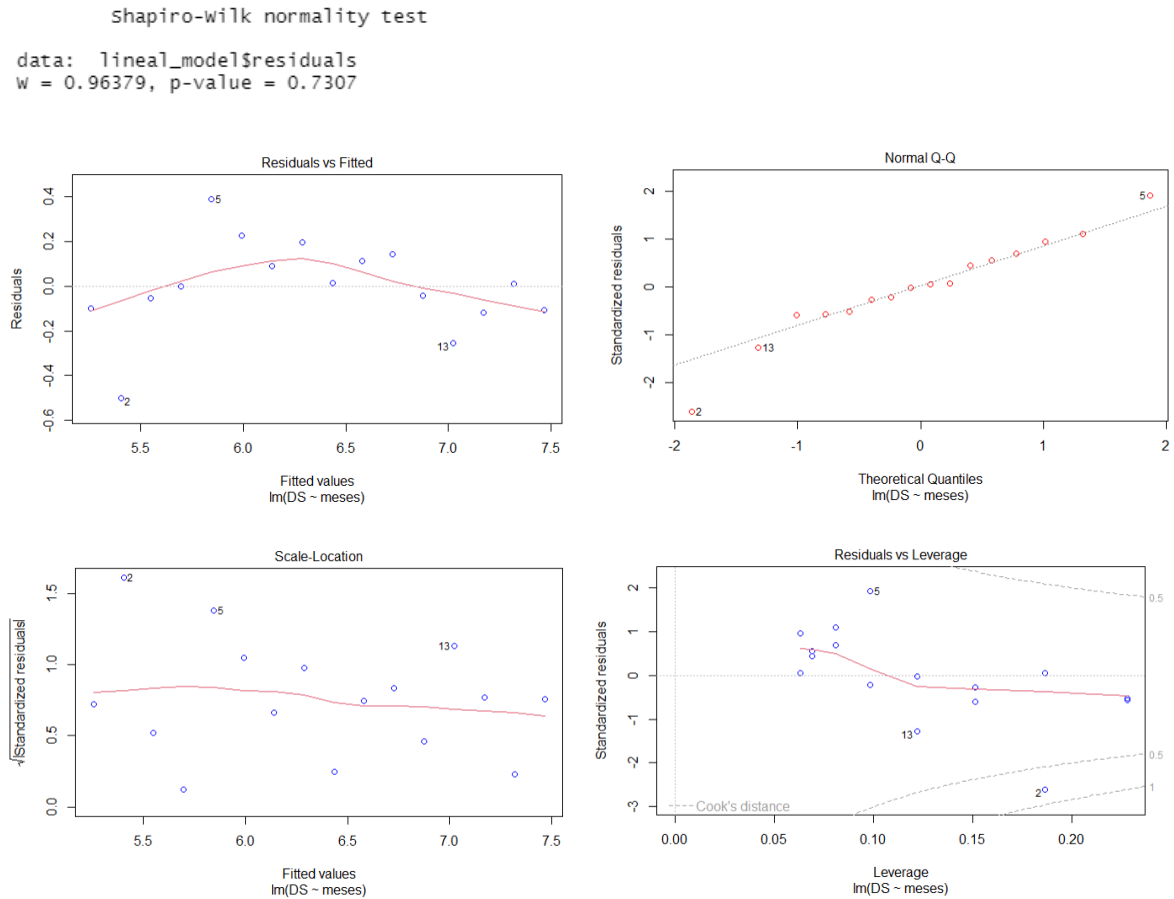
Residual standard error: 0.213 on 14 degrees of freedom
Multiple R-squared:  0.9208,    Adjusted R-squared:  0.9151
F-statistic: 162.7 on 1 and 14 DF,  p-value: 4.248e-09
```



Como se puede observar en la gráfica superior, el modelo lineal se ajusta relativamente bien a los datos desestacionalizados, por lo que se podría inferir que se puede usar para generar predicciones, aun así, se va a hacer un análisis de pertinencia

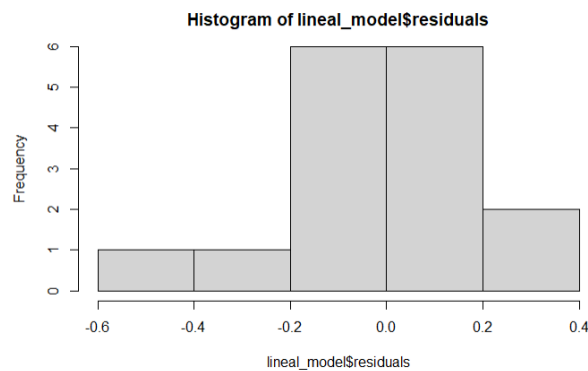
## Analiza la pertinencia del modelo lineal

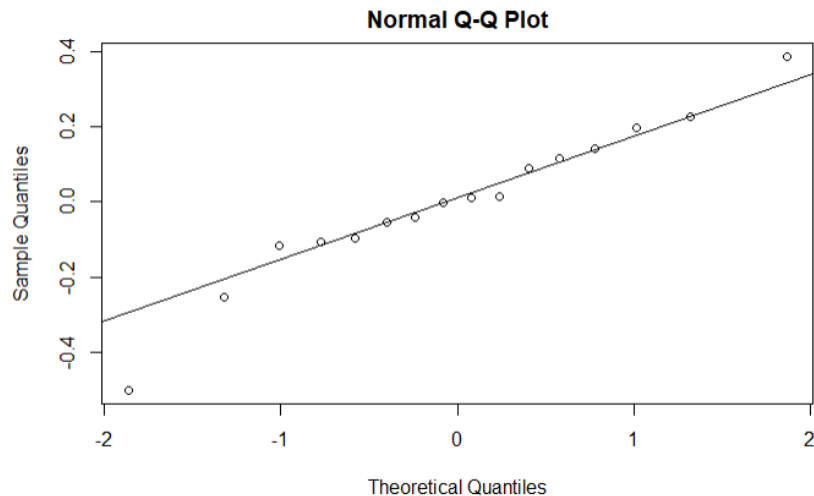
### Análisis de los residuos:



Según el análisis de residuos, se nos permite validar el uso de un modelo lineal, ya que cumple con normalidad en los residuos, lo que nos ayuda a que se pueda usar una regresión lineal como método principal.

## Prueba de normalidad





$H_0$ : Los residuos tienen un comportamiento normal

$H_1$ : los residuos no tienen un comportamiento normal

$\alpha = 0.05$

Se rechaza  $H_0$  si  $p > \alpha$

```
call:
lm(formula = DS ~ meses)

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 ***
meses        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 podemos ver, p-value del residuo es de 4.248e-09, por lo que se puede decir que los residuos se comportan de manera normal, por lo que es óptimo usar la regresión lineal.

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

CME: 0.04537875

EPAM: 15.07404

## **Conclusión**

Podemos ver que el modelo propuesto tiene mucha similitud con las ventas reales, por lo que se puede decir que el modelo se ajusta correctamente. También se puede observar un grado de normalidad en los datos, por lo que se puede explicar un alto grado de variabilidad, mientras que se tiene dependencia directa entre las variables.

El modelo nos puede ayudar a hacer aproximaciones a predicciones a futuro, como también nos ayuda a analizar las ventas producidas durante los primeros periodos de tiempo.

## **Anexo:**

[https://drive.google.com/file/d/16diDuacjr2Ir69-J7EESMCXjF5Q0wODv/view?usp=share\\_link](https://drive.google.com/file/d/16diDuacjr2Ir69-J7EESMCXjF5Q0wODv/view?usp=share_link)