A8-Series de tiempo no estacionarias. Tendencia

Diego Alberto Baños Lopez
A01275100

2023-11-17

Ejercicio 1

Usa los datos de las ventas de televisores para familiarizarte con el análisis de tendencia de una serie de tiempo:

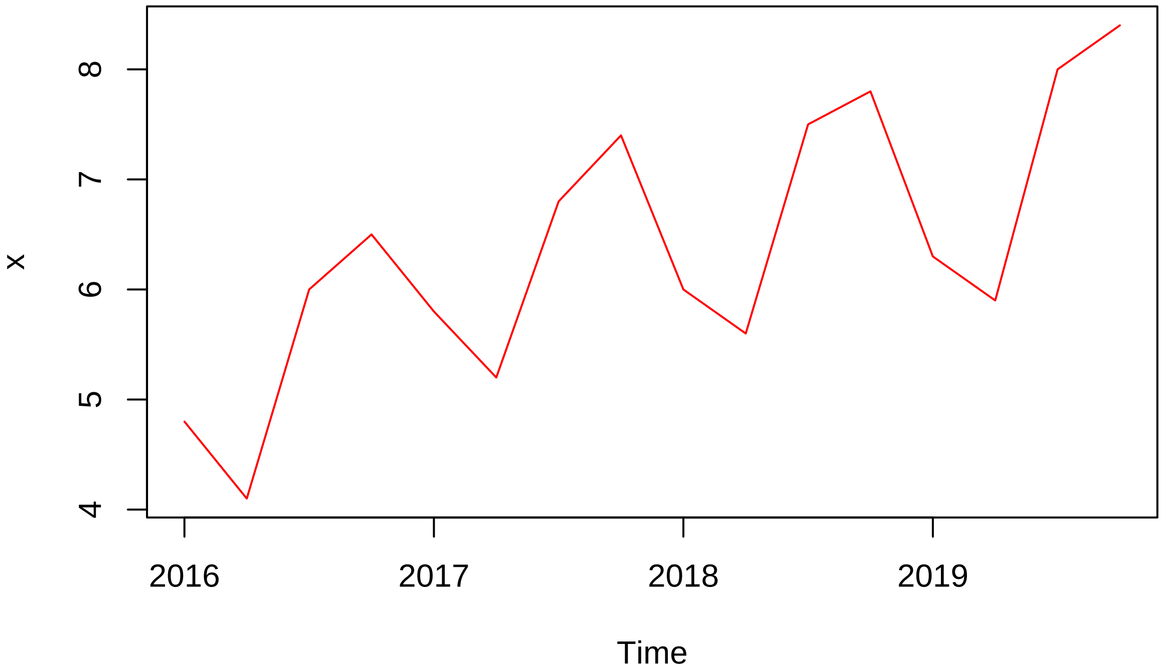
```r
# Datos de ventas
ventas <- c(
  4.8, 4.1, 6, 6.5, 5.8,
  5.2, 6.8, 7.4, 6, 5.6,
  7.5, 7.8, 6.3, 5.9, 8, 8.4
)

# Crear serie de tiempo
x <- ts(ventas, frequency = 4, start = c(2016, 1))

# Crear un data frame con los datos
datos_tabla <- data.frame(
  Año = rep(seq(2016, 2019), each = 4),
  Trimestre = rep(1:4, times = 4),
  Ventas = ventas
)

# Mostrar la tabla
print(datos_tabla)
```

```
##     Año Trimestre Ventas
## 1  2016         1    4.8
## 2  2016         2    4.1
## 3  2016         3    6.0
## 4  2016         4    6.5
## 5  2017         1    5.8
## 6  2017         2    5.2
## 7  2017         3    6.8
## 8  2017         4    7.4
## 9  2018         1    6.0
## 10 2018         2    5.6
## 11 2018         3    7.5
## 12 2018         4    7.8
## 13 2019         1    6.3
## 14 2019         2    5.9
## 15 2019         3    8.0
```

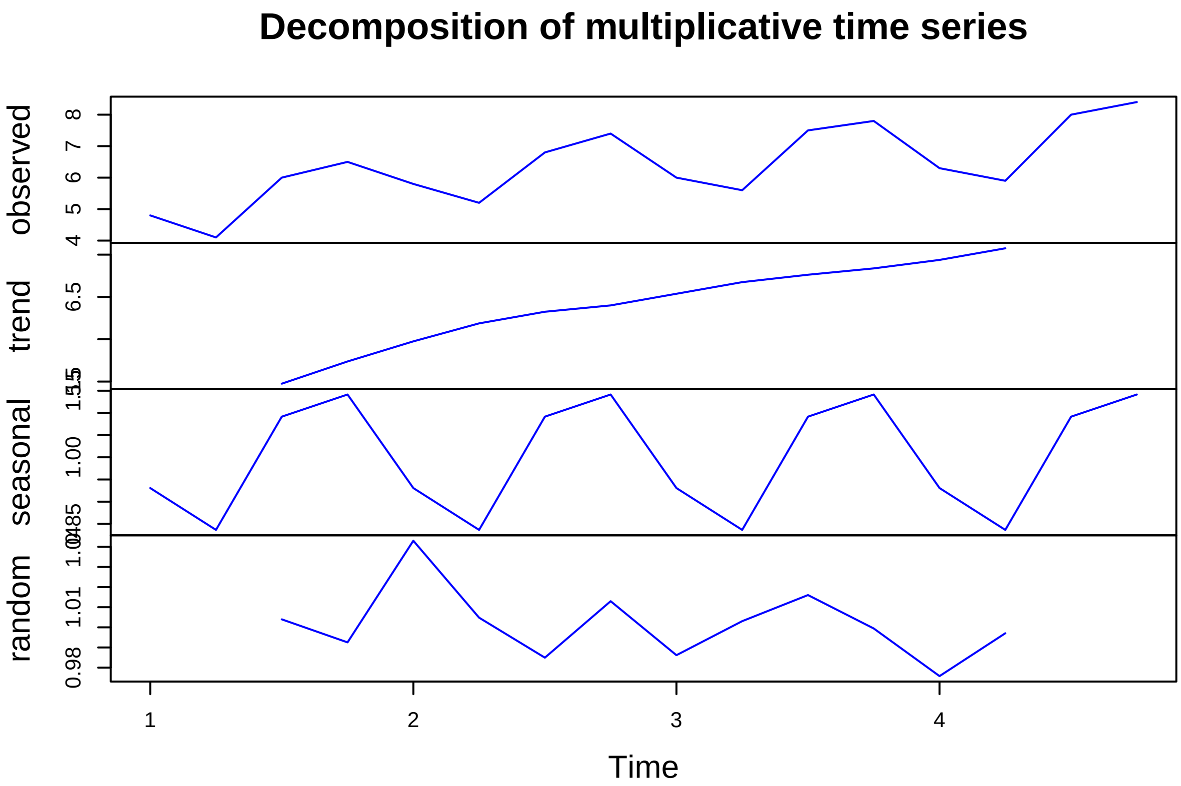
```
## 16 2019         4    8.4
```

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

```r
ventas <- c(
  4.8, 4.1, 6, 6.5, 5.8, 5.2,
  6.8, 7.4, 6, 5.6, 7.5, 7.8,
  6.3, 5.9, 8, 8.4
)
T <- ts(ventas, frequency = 4, start(c(2016, 1)))
D <- decompose(T, type = "m")
plot.ts(x, col = "red")
```




```r
plot(D, col = "blue")
```



Realiza el análisis de tendencia y estacionalidad:

Descompón la serie en sus 3 componentes e interprétalos

```
# Análisis de tendencia y estacionalidad
tendencia <- D$trend
estacionalidad <- D$seasonal
print("Tendencia")
```

```
## [1] "Tendencia"
```

```
print(tendencia)
```

```
##     Qtr1   Qtr2   Qtr3   Qtr4
## 1     NA     NA 5.4750 5.7375
## 2 5.9750 6.1875 6.3250 6.4000
## 3 6.5375 6.6750 6.7625 6.8375
## 4 6.9375 7.0750     NA     NA
```

```
print("Estacionalidad")
```

```
## [1] "Estacionalidad"
```

```
print(estacionalidad)
```

```
##        Qtr1      Qtr2      Qtr3      Qtr4
## 1 0.9306617 0.8363763 1.0915441 1.1414179
## 2 0.9306617 0.8363763 1.0915441 1.1414179
## 3 0.9306617 0.8363763 1.0915441 1.1414179
## 4 0.9306617 0.8363763 1.0915441 1.1414179
```

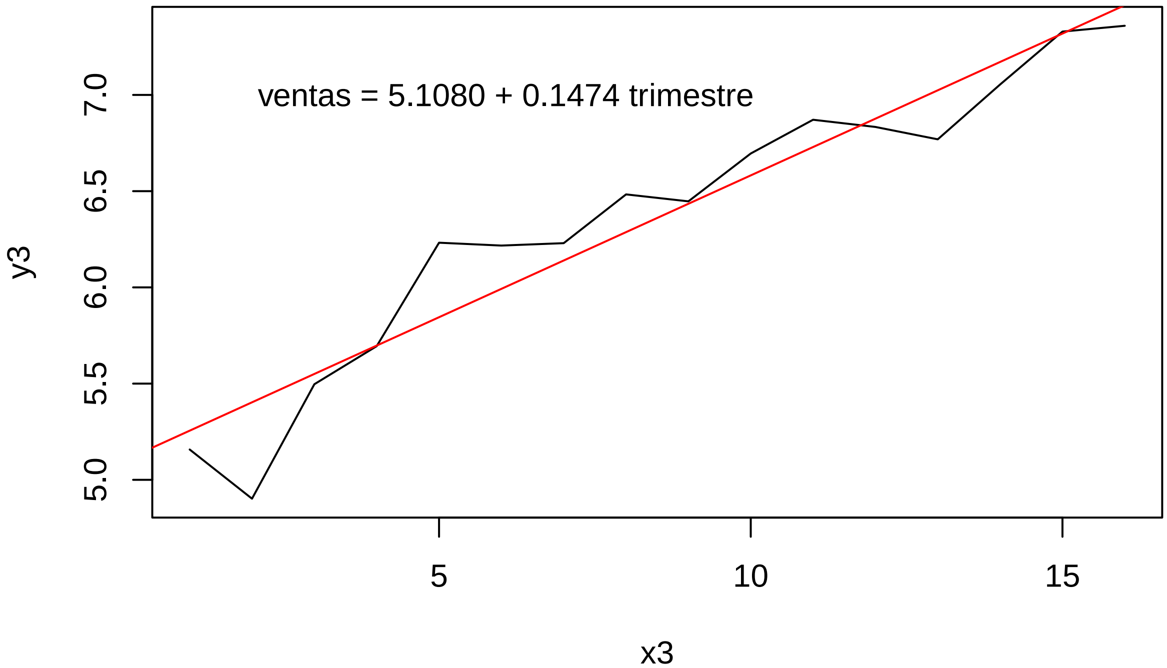
Analiza el modelo lineal de la tendencia:

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

```
ventas_desestacionalizadas <- (D$x) / (D$seasonal)
x3 <- 1:16
y3 <- ventas_desestacionalizadas
N3 <- lm(y3 ~ x3)
N3
```

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

```
plot(x3, y3, type = "l")
abline(N3, col = "red")
text(6, 7, " ventas = 5.1080 + 0.1474 trimestre")
```



Dibuja la recta junto con las ventas desestacionalizadas.

Analiza la pertinencia del modelo lineal:

```
summary(N3)
```

Significancia de β_1

```
##
## 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
```

Variabilidad explicada por el modelo (c.d)

```r
summary(N3)$r.squared
```

```
## [1] 0.9207911
```

Análisis de los residuos

```r
resid(N3)
```

```
##            1            2            3            4            5            6
## -0.097803257 -0.500706332 -0.053387333 -0.002898368  0.387173048  0.224963115
##            7            8            9           10           11           12
##  0.089991415  0.196066619  0.012545918  0.113688739  0.141756821 -0.043020048
##           13           14           15           16
## -0.254630784 -0.117149041  0.010295542 -0.106886054
```

Prueba de normalidad

```r
shapiro.test(resid(N3))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  resid(N3)
## W = 0.96379, p-value = 0.7307
```

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

```r
print("CME")
```

CME

```
## [1] "CME"
```

```r
mean(resid(N3)^2)
```

```
## [1] 0.0397064
```

```r
print("EPAM")
```

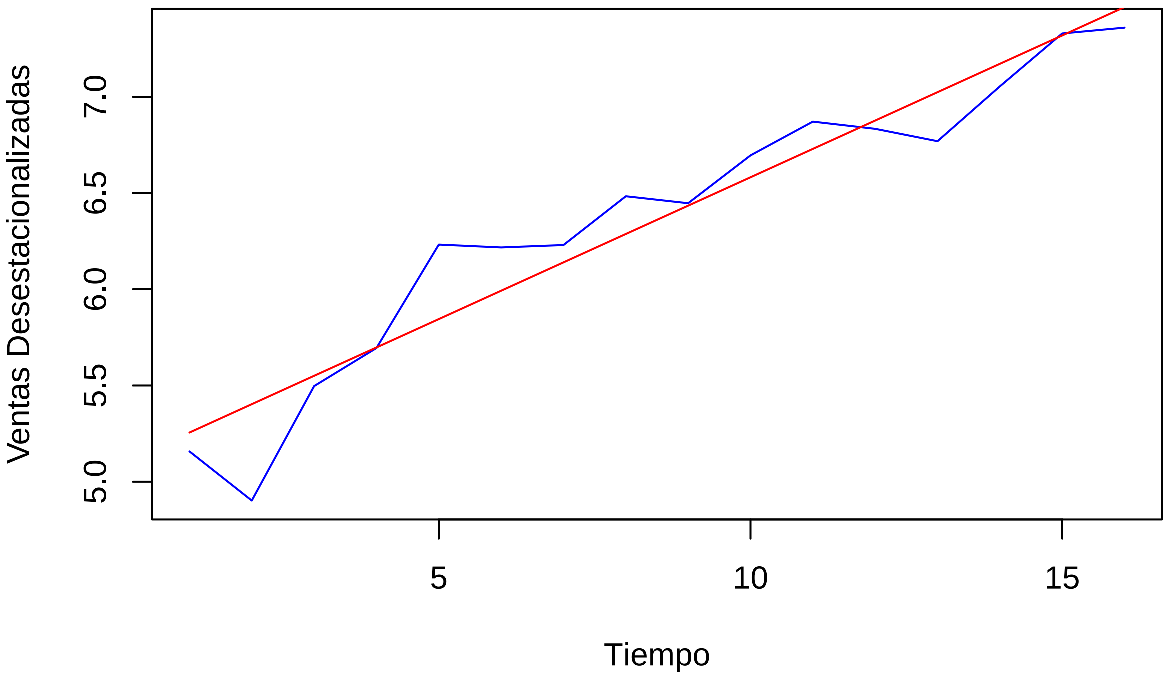
```
## [1] "EPAM"
```

```r
mean(abs(resid(N3) / ventas_desestacionalizadas)) * 100
```

```
## [1] 2.439533
```

Dibuja el gráfico de los valores de las ventas y las predicciones vs el tiempo

```r
predicciones <- predict(N3)
plot(x3, ventas_desestacionalizadas,
  type = "l", col = "blue", ylab = "Ventas Desestacionalizadas",
  xlab = "Tiempo"
)
lines(x3, predicciones, col = "red")
```



###

Concluye sobre el modelo: de acuerdo al análisis de verificación de los supuestos, ¿es el mejor modelo que puedes obtener?

- En base al alto coeficiente de determinación ($R^2 = 0.921$), la normalidad de los residuos (p-valor de la prueba de Shapiro-Wilk = 0.731), y los bajos valores de CME y EPAM, podemos concluir que el modelo lineal proporciona un ajuste significativo y adecuado para la serie de tiempo de ventas de televisores.

Propón un posible mejor modelo para la tendencia de los datos.

Por el momento solo se podria proponer u8nmodelo ARIMA si hay autocorrelación en los residuos del modelo.

Realiza el pronóstico para el siguiente año.

```r
f <- function(x) {
  5.1080 + 0.1474 * x
}

a1 <- D$seasonal[1]
a2 <- D$seasonal[2]
a3 <- D$seasonal[3]
```

```r
a4 <- D$seasonal[4]
f(17) * a1 * 1000
```

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

```r
f(18) * a2 * 1000
```

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

```r
f(19) * a3 * 1000
```

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

```r
f(20) * a4 * 1000
```

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

Ejercicio 2

A continuación, se presentan los datos correspondientes a los últimos tres años de ventas trimestrales (número de ejemplares vendidos) de un libro de texto universitario

```r
ventas <- data.frame(
  trimestre = c(1, 2, 3, 4),
  anio1 = c(1690, 940, 2625, 2500),
  anio2 = c(1800, 900, 2900, 2360),
  anio3 = c(1850, 1100, 2930, 2615)
)
```

```r
ventas$promedio_movil4 <- rowMeans(ventas[, c("anio1", "anio2", "anio3")])
ventas$promedio_movil_centrado <- c(
  NA,
  head(ventas$promedio_movil4, -1) + tail(ventas$promedio_movil4, -1)
) / 2
print(ventas[, c("trimestre", "promedio_movil4", "promedio_movil_centrado")])
```

Encuentre los promedios móviles de cuatro trimestres y los promedios móviles centrados

```
##   trimestre promedio_movil4 promedio_movil_centrado
## 1         1        1780.000                      NA
## 2         2         980.000                1380.000
## 3         3        2818.333                1899.167
## 4         4        2491.667                2655.000
```

```r
promedio_anual <- colMeans(ventas[, c("anio1", "anio2", "anio3")])
indices_estacionales <- t(t(ventas[, c("anio1", "anio2", "anio3")]) / promedio_anual) * 100
print(indices_estacionales)
```

Calcule los índices estacionales de los cuatro trimestres

```
##          anio1     anio2     anio3
## [1,]  87.16957  90.45226  87.11006
## [2,]  48.48485  45.22613  51.79517
## [3,] 135.39652 145.72864 137.96351
## [4,] 128.94907 118.59296 123.13125
```

```r
trimestre_max_indice <- which.max(apply(indices_estacionales, 2, max))
cat(
  "El mayor índice estacional se encuentra en el trimestre",
  trimestre_max_indice
)
```

¿Cuándo obtiene la editorial el mayor índice estacional? ¿Parece razonable este resultado? ¿Por qué?

```
## El mayor índice estacional se encuentra en el trimestre 2
```

Los resultados parecen ser razonable acorde a lo que se obtiene, ya que parecen ser congruentes con patrones de demanda que por lo general estas instituciones tienen, marcando cuando inicia un ciclo escolar, y por lo tanto marcando una etapa de mayor demanda