

Series_NoEstacionarias

Arturo

2023-11-14

Nombre: Arturo Garza Campuzano

Matrícula: A00828096

Series No Estacionarias

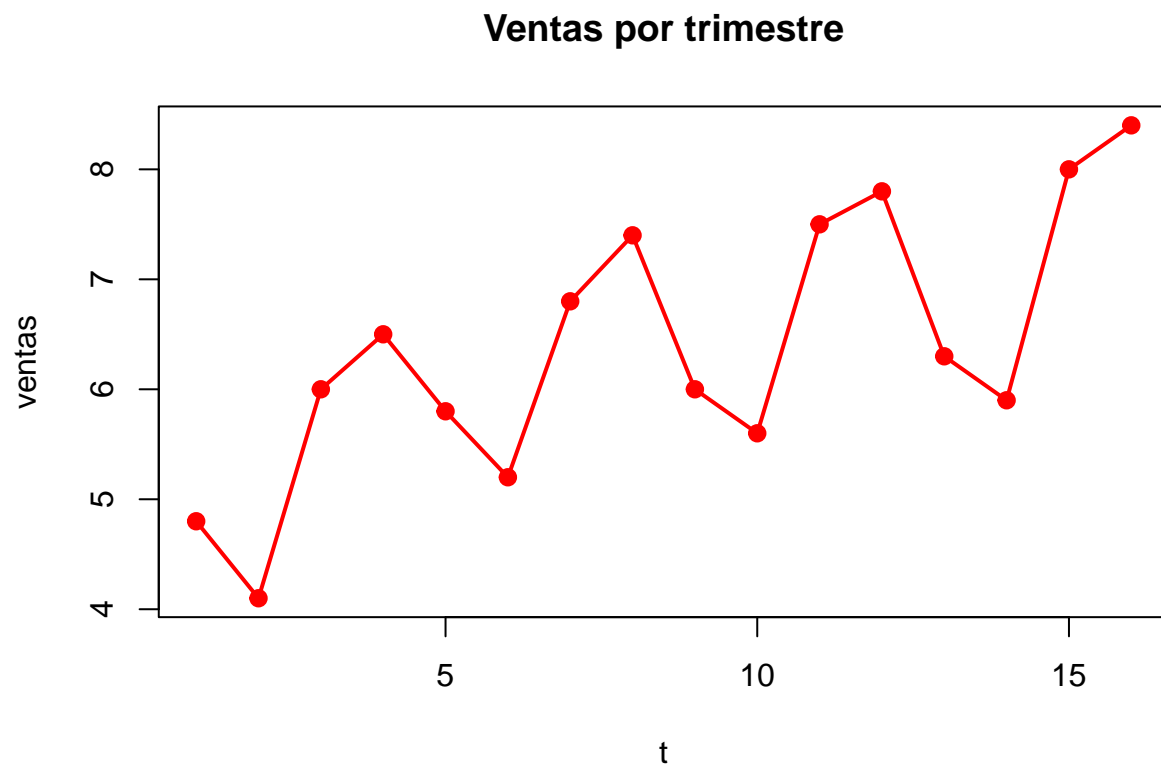
Se utilizan los datos de las ventas de televisores para familiarizarse con el análisis de tendencia de una serie de tiempo:

```
# Datos
t = 1:16
ventas = c(4.8, 4.1, 6.0, 6.5, 5.8, 5.2, 6.8, 7.4, 6.0, 5.6, 7.5, 7.8, 6.3, 5.9, 8.0, 8.4)

# Serie de tiempo
x = ts(ventas, frequency = 4, start=c(2016,1))
```

1. Gráfica de dispersión

```
# Alternativa: plot.ts(x, col = "red")
plot(t, ventas, type = "o", col = "red", main = "Ventas por trimestre", lwd = 2, pch = 19)
```

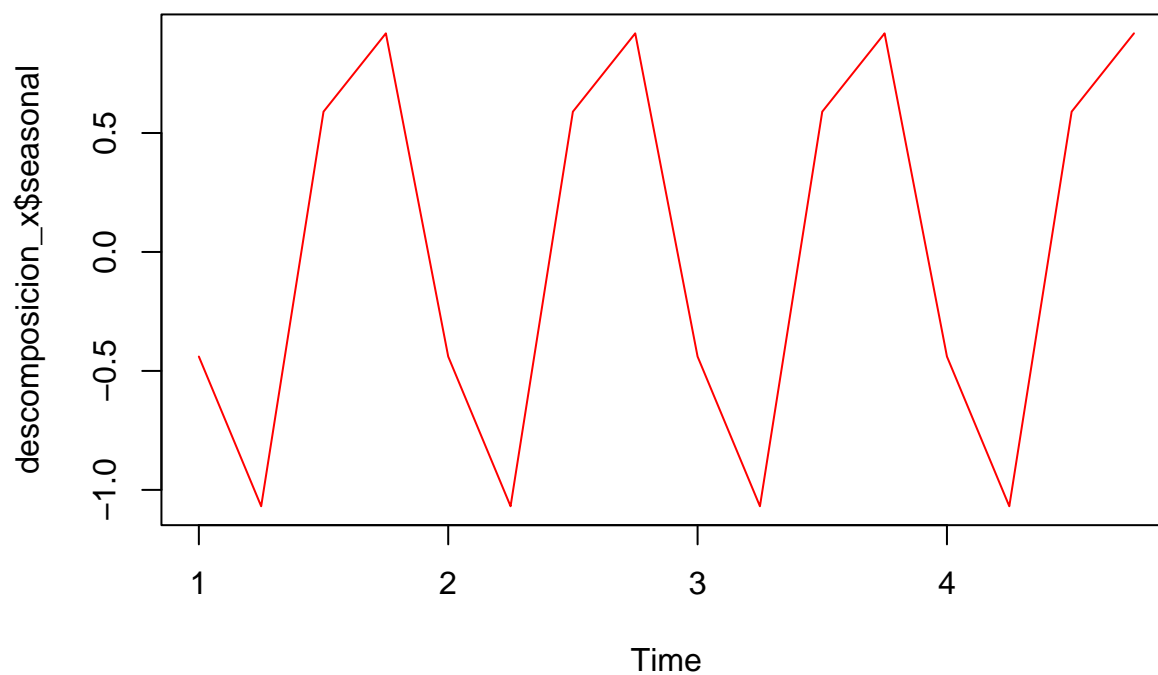


2. Análisis de tendencia y estacionalidad

```
# Descomposicion
descomposicion_x = decompose(x)

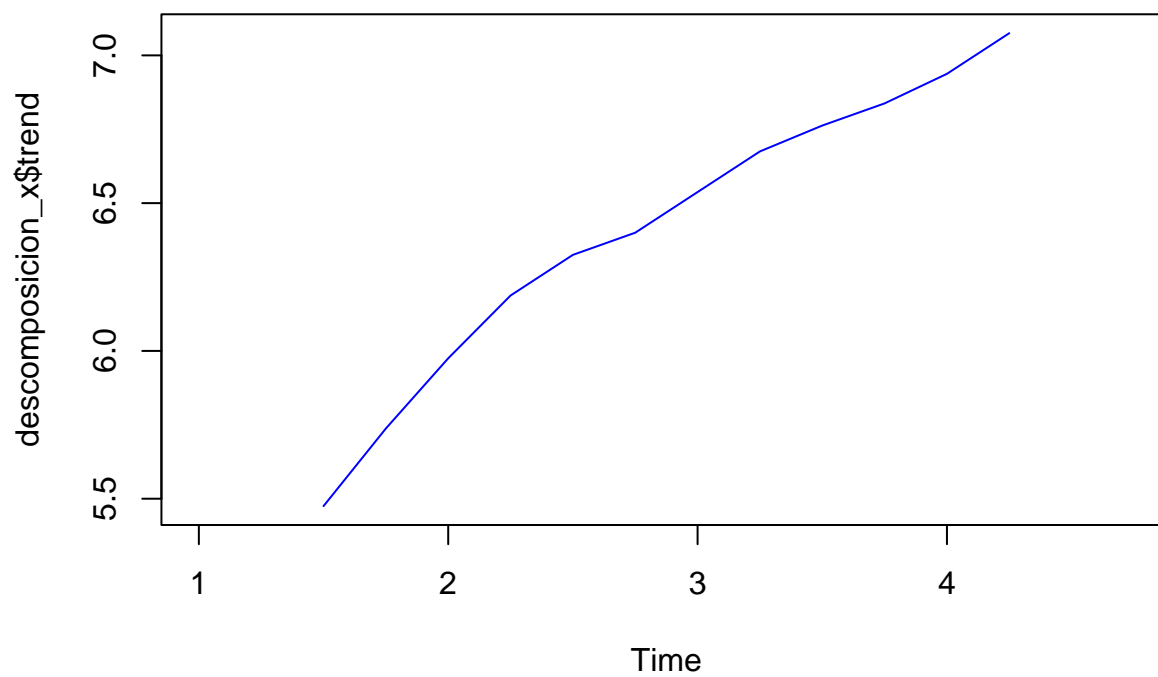
# Grafico de las componentes descompuestas
plot(descomposicion_x$seasonal, col = "red", main = "Estacionalidad")
```

Estacionalidad



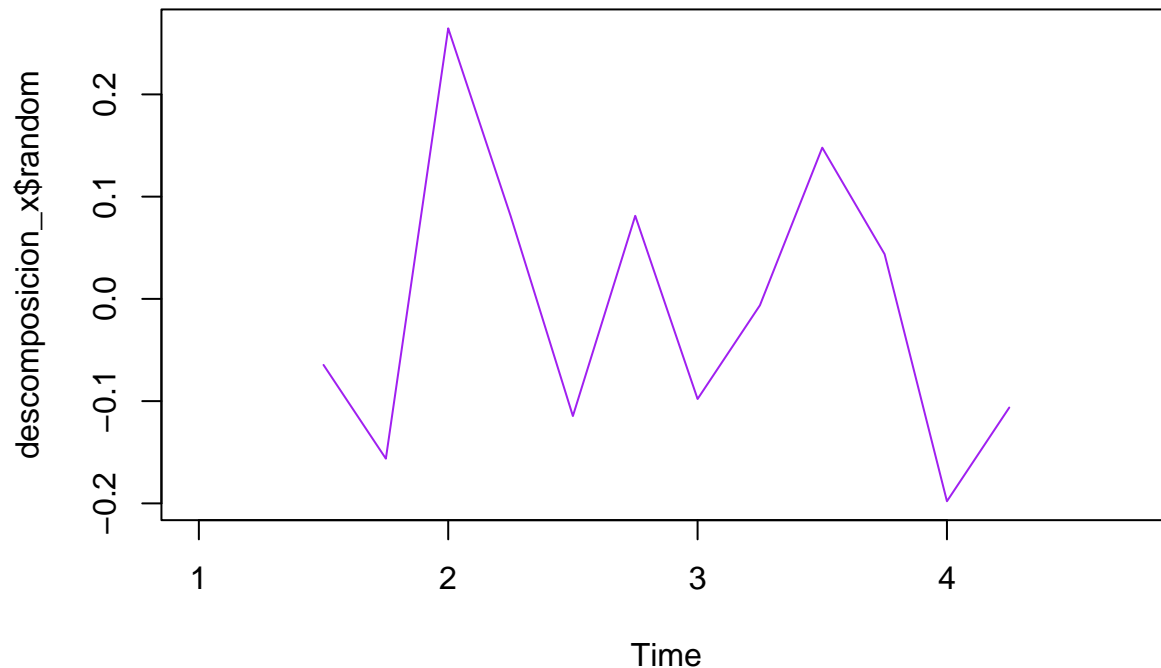
```
plot(descomposicion_x$trend, col = "blue", main = "Tendencia")
```

Tendencia



```
plot(descomposicion_x$random, col = "purple", main = "Residuos")
```

Residuos



Interpretación:

- Estacionalidad:
 - Los primeros dos trimestres indican una influencia estacional negativa.
 - Los últimos dos trimestres indican una influencia estacional positiva.
 - La secuencia de valores sugiere un patrón de variación que se repite cada cuatro elementos.
- Tendencia: En este caso, parece haber un aumento general en los valores de la serie temporal a lo largo del tiempo. Esto parece seguir un comportamiento lineal.
- Residuos: Los datos muestran una variabilidad moderada, con una concentración leve de valores negativos.

3. Análisis del modelo lineal

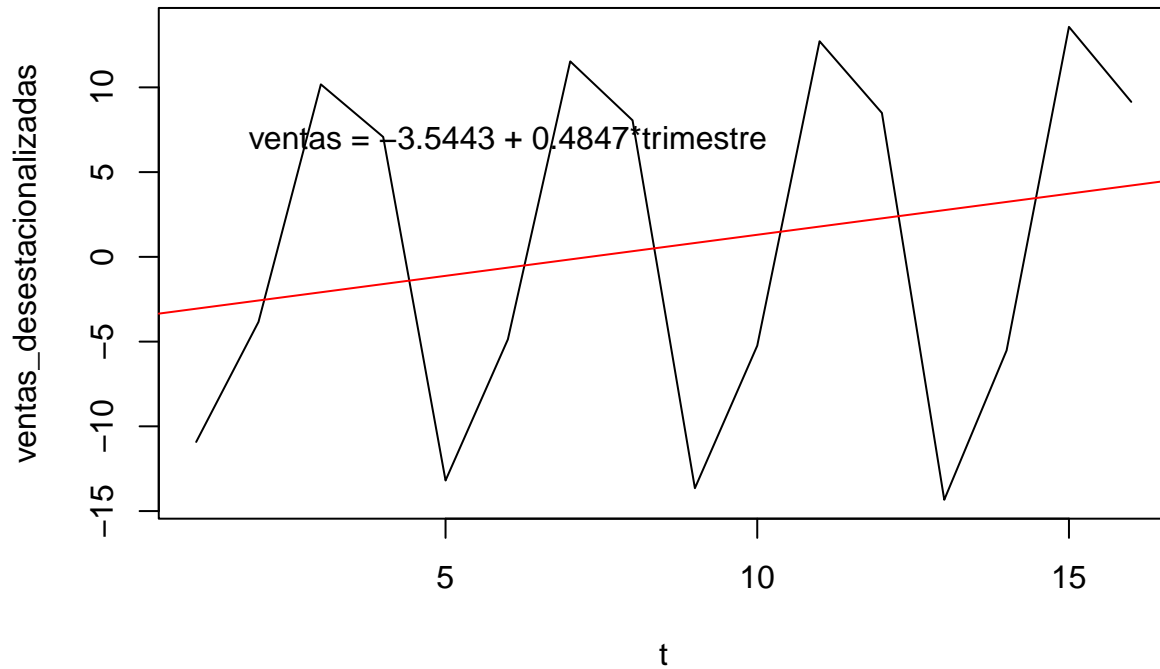
3.1 Regresión lineal de la tendencia

```
# Regresion lineal de la tendencia
ventas_desestacionalizadas = (descomposicion_x$x)/(descomposicion_x$seasonal)
tendencia_rl = lm(ventas_desestacionalizadas~t)
tendencia_rl
```

```
##
## Call:
## lm(formula = ventas_desestacionalizadas ~ t)
##
## Coefficients:
## (Intercept)          t
##      -3.5443       0.4847
```

```
# Grafico de ventas desestacionalizadas vs tiempo
plot(t, ventas_desestacionalizadas, type = "l")
```

```
abline(tendencia_rl, col = "red")
text(6, 7, "ventas = -3.5443 + 0.4847*trimestre")
```



3.2 Recta junto con las ventas deestacionalizadas

```
summary_tendencia_rl <- summary(tendencia_rl)
summary_tendencia_rl
```

3.2.1 Análisis de la pertinencia del modelo lineal

```
##
## Call:
## lm(formula = ventas_deestacionalizadas ~ t)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -17.088  -8.085   1.836   8.971  12.267
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -3.5443     5.5166  -0.642   0.531
## t              0.4847     0.5705   0.850   0.410
##
## Residual standard error: 10.52 on 14 degrees of freedom
## Multiple R-squared:  0.04902,    Adjusted R-squared:  -0.0189
## F-statistic: 0.7217 on 1 and 14 DF,  p-value: 0.4099
```

3.2.2 Significancia de beta 1

- Hipótesis nula (H_0): La pendiente es igual a cero.
- Hipótesis alternativa (H_1): La pendiente no es igual a cero.

```
# P-valor asociado a la pendiente
cat("P-valor para la pendiente (beta 1):", summary_tendencia_rl$coefficients["t", "Pr(>|t|)"])
```

```
## P-valor para la pendiente (beta 1): 0.4098838
```

Dado que el p-valor (0.41) es mayor que un nivel de significancia de 0.05 no se tiene evidencia suficiente para rechazar la hipótesis nula. En otras palabras, no hay suficiente evidencia estadística para afirmar que hay una relación lineal significativa entre el tiempo y las ventas desestacionalizadas.

```
# Coeficiente de determinación
cat("Coeficiente de Determinación (R^2):", summary_tendencia_rl$r.squared)
```

3.2.3 Variabilidad explicada por el modelo

```
## Coeficiente de Determinación (R^2): 0.04902201
```

En este caso, el coeficiente de determinación indica que aproximadamente el 4.9% de la variabilidad en las ventas desestacionalizadas es explicada por el modelo de regresión lineal con respecto al tiempo. Esto sugiere que el modelo de regresión lineal no es muy efectivo para explicar la variabilidad en las ventas desestacionalizadas.

```
# Residuos del modelo
residuos <- residuals(tendencia_rl)

# Grafico de dispersión de los residuos
plot(residuos, ylab = "Residuos", xlab = "Índice de Observación", main = "Gráfico de Residuos")
abline(h = 0, col = "red")
```



3.2.4 Análisis de los residuos

3.2.5 Prueba de normalidad

- Hipótesis nula (H_0): Los residuos siguen una distribución normal.
- Hipótesis alternativa (H_1): Los residuos no siguen una distribución normal.

```
# Prueba de Shapiro-Wilk
shapiro_test <- shapiro.test(residuos)

# Resultado de la prueba de normalidad
print("Prueba de Shapiro-Wilk para la normalidad de los residuos:")

## [1] "Prueba de Shapiro-Wilk para la normalidad de los residuos:"
print(shapiro_test)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  residuos
## W = 0.90397, p-value = 0.09308
```

Dado que el valor de p (0.09) es mayor que un nivel de significancia de 0.05 no se tiene suficiente evidencia para rechazar la hipótesis nula. En otras palabras, no hay suficiente evidencia estadística para afirmar que los residuos no siguen una distribución normal.

4. Cálculo del CME y EPAM de la predicción de la serie de tiempo

```
# Realizar las predicciones con el modelo
predicciones <- predict(tendencia_rl)

# Calcular los residuos
residuos <- residuals(tendencia_rl)

# Calcular el Error Cuadrático Medio (CME)
cme <- mean(residuos^2)

# Calcular el Error Porcentual Absoluto Medio (EPAM)
epam <- mean(abs(residuos / ventas_desestacionalizadas) * 100)

# Mostrar los resultados
print("Error Cuadrático Medio (CME):")

## [1] "Error Cuadrático Medio (CME):"
print(cme)

## [1] 96.83152
print("Error Porcentual Absoluto Medio (EPAM):")

## [1] "Error Porcentual Absoluto Medio (EPAM):"
print(epam)

## [1] 94.88578
```

5. Gráfico de los valores de las ventas y predicciones vs el tiempo

```
# Datos de ventas
ventas <- c(4.8, 4.1, 6.0, 6.5, 5.8, 5.2, 6.8, 7.4, 6.0, 5.6, 7.5, 7.8, 6.3, 5.9, 8.0, 8.4)
```

```

# Crear un marco de datos con los valores originales y el tiempo
datos <- data.frame(Tiempo = seq_along(ventas), Ventas = ventas)

# Ajustar el modelo de regresión lineal
tendencia_rl <- lm(Ventas ~ Tiempo, data = datos)

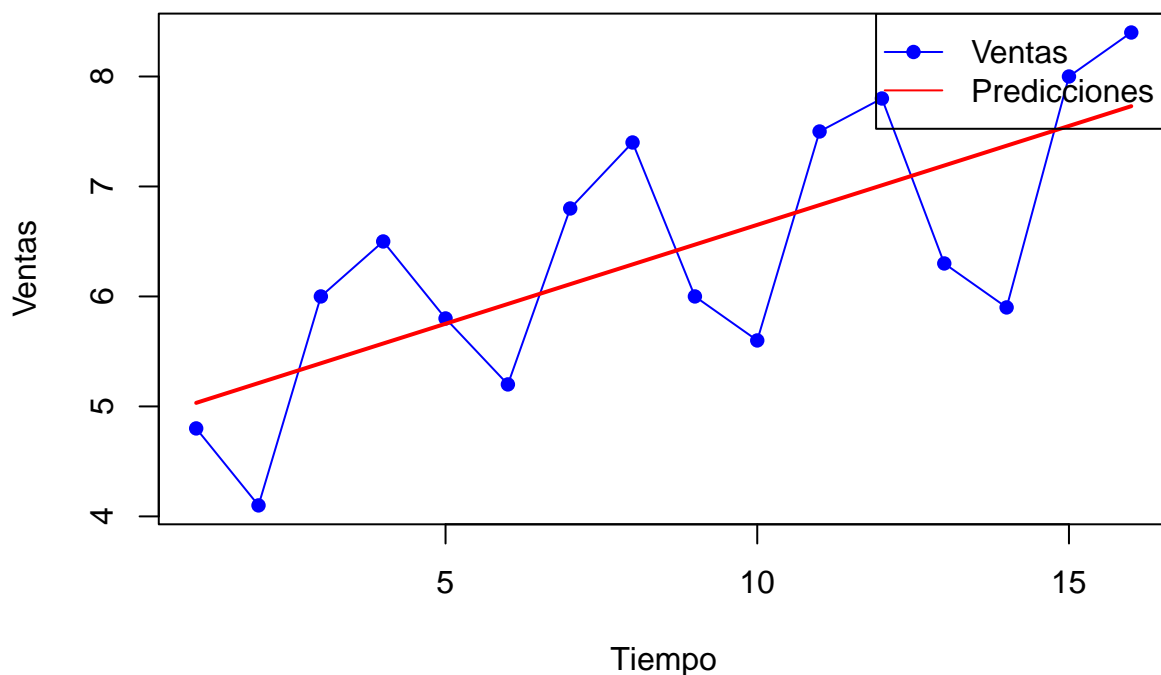
# Resumen del modelo
summary_tendencia_rl <- summary(tendencia_rl)

# Obtener las predicciones del modelo
predicciones <- predict(tendencia_rl, data.frame(Tiempo = seq_along(ventas)))

# Graficar los valores originales y las predicciones
plot(datos$Tiempo, datos$Ventas, type = "o", col = "blue", pch = 16, main = "Ventas vs Predicciones", xlab = "Tiempo", ylab = "Ventas")
lines(seq_along(ventas), predicciones, col = "red", type = "l", lwd = 2)
legend("topright", legend = c("Ventas", "Predicciones"), col = c("blue", "red"), lty = c(1, 1), pch = c(16, 1))

```

Ventas vs Predicciones



6. Conclusión sobre el modelo

El modelo de regresión lineal parece tener limitaciones para explicar las ventas desestacionalizadas en función del tiempo. Sería recomendable explorar modelos más complejos o considerar otras variables que puedan mejorar la capacidad predictiva del modelo.

7. Posible mejor modelo para la tendencia de los datos

Para llegar a un mejor modelo se podrían probar las siguientes opciones:

- Modelo Polinómico
- Modelo No Lineal

- Modelo de Suavización
- Modelo de Regresión con Variables Adicionales

8. Pronóstico para el siguiente año

```
f = function(x) {-3.5443 + 0.4847*x}
```

```
# Indices estacionales
```

```
a1 = descomposicion_x$seasonal[1]
```

```
a2 = descomposicion_x$seasonal[2]
```

```
a3 = descomposicion_x$seasonal[3]
```

```
a4 = descomposicion_x$seasonal[4]
```

```
# Pronosticos
```

```
f(17)*a1*1000
```

```
## [1] -2064.108
```

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

```
## [1] -5536.446
```

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

```
## [1] 3339.99
```

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

```
## [1] 5650.037
```

9. Un problemilla más

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.

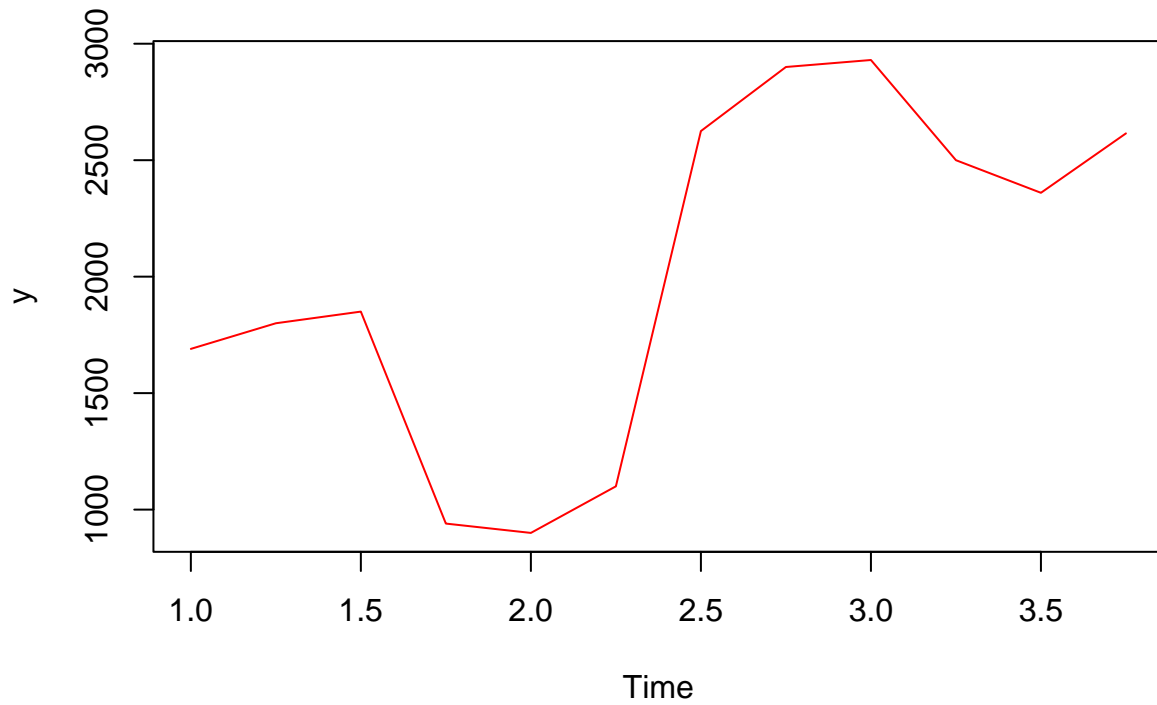
```
t_1 = 1:12
```

```
ventas_libros = c(1690, 1800, 1850, 940, 900, 1100, 2625, 2900, 2930, 2500, 2360, 2615)
```

```
y = ts(ventas_libros, frequency = 4, c(1, 1))
```

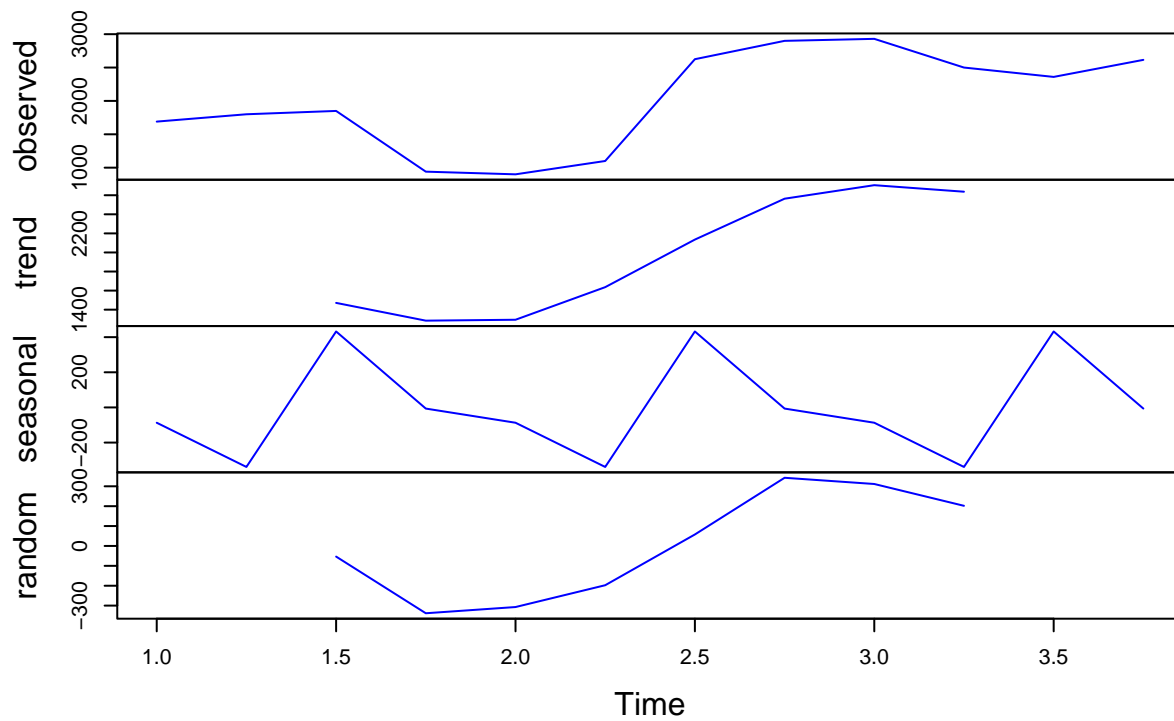
```
# Grafica de dispersion
```

```
plot.ts(y, col = "red")
```



```
# Descomposicion
descomposicion_y = decompose(y)
plot(descomposicion_y, col = "blue")
```

Decomposition of additive time series



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

```
# Promedios moviles de cuatro trimestres
ma_four_quarters <- filter(y, rep(1, 4) / 4, sides = 1)

# Promedios moviles centrados
ma_centered <- filter(y, rep(1, 4) / 4)

# Muestra los resultados
print("Promedios móviles de cuatro trimestres:")
```

```
## [1] "Promedios móviles de cuatro trimestres:"
```

```
print(ma_four_quarters)
```

```
##      Qtr1    Qtr2    Qtr3    Qtr4
## 1      NA      NA      NA 1570.00
## 2 1372.50 1197.50 1391.25 1881.25
## 3 2388.75 2738.75 2672.50 2601.25
```

```
print("Promedios móviles centrados:")
```

```
## [1] "Promedios móviles centrados:"
```

```
print(ma_centered)
```

```
##      Qtr1    Qtr2    Qtr3    Qtr4
## 1      NA 1570.00 1372.50 1197.50
## 2 1391.25 1881.25 2388.75 2738.75
## 3 2672.50 2601.25      NA      NA
```

9.2 Calcule los índices estacionales de los cuatro trimestres.

```
# Calcula los indices estacionales
seasonal_indices <- decompose(y)$seasonal

# Muestra los resultados
print("Índices estacionales:")
```

```
## [1] "Índices estacionales:"
```

```
print(seasonal_indices)
```

```
##      Qtr1    Qtr2    Qtr3    Qtr4
## 1 -87.109375 -338.671875 432.265625 -6.484375
## 2 -87.109375 -338.671875 432.265625 -6.484375
## 3 -87.109375 -338.671875 432.265625 -6.484375
```

9.3 ¿Cuándo obtiene la editorial el mayor índice estacioanl?, ¿parece razonable este resultado, ¿por qué?

En este caso, parece que el mayor índice estacional corresponde al tercer trimestre, donde se tiene un valor de 432.27. Tomando en cuenta el contexto del problema, esto sugiere que el negocio no cuenta con ganancias; lo cual parece que no es razonable.