Series de tiempo no estacionarias. Tendencia

Luis Ángel Guzmán Iribe - A01741757

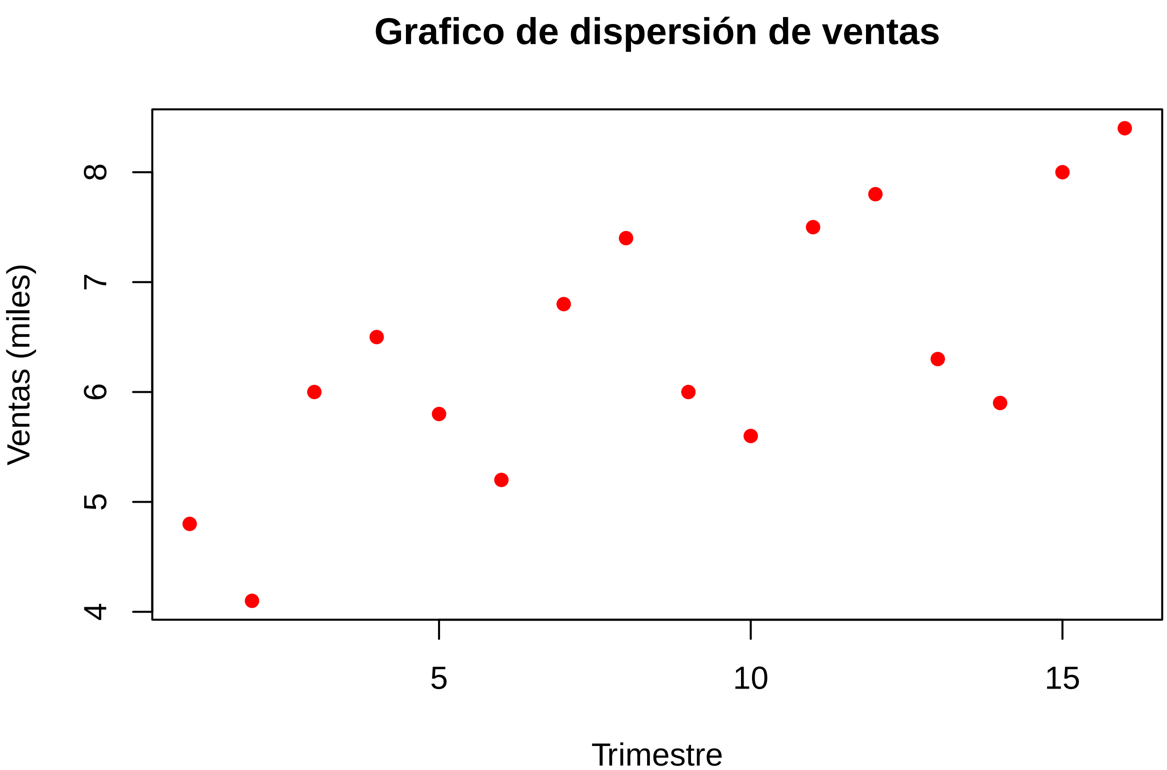
2023-11-14

Ventas por trimestre

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

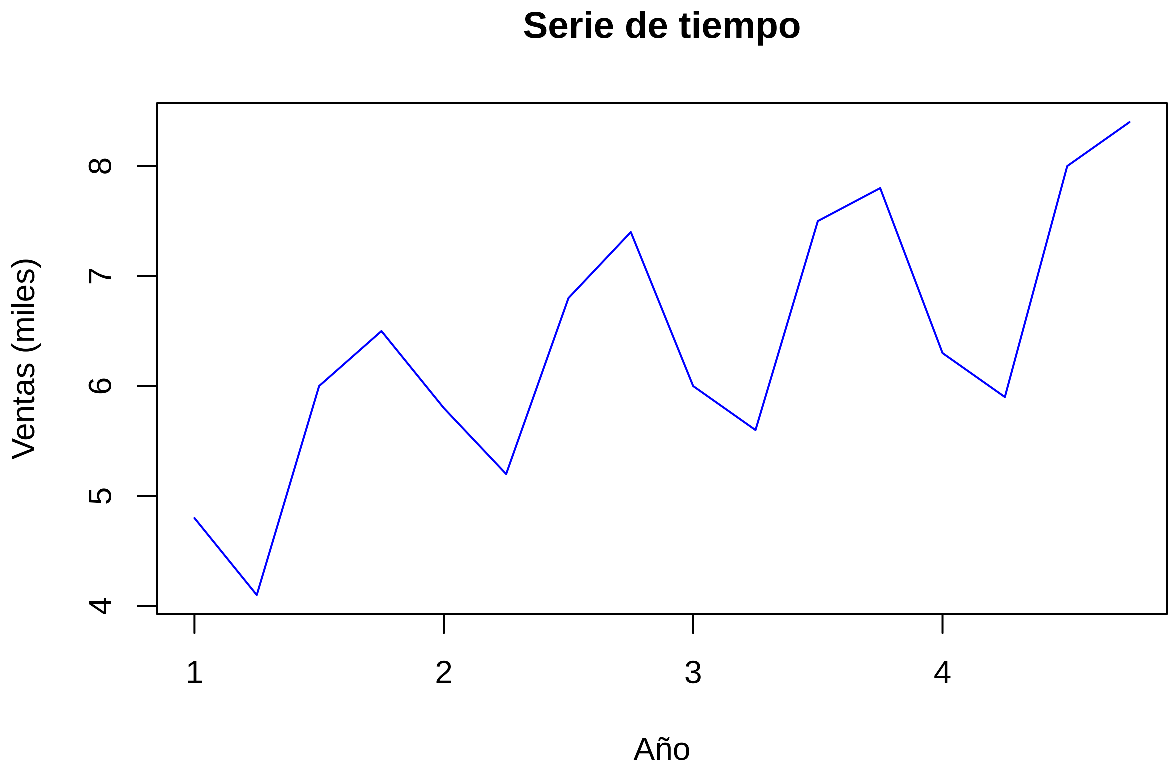
```r
t = 1:16
ventas <- c(4.8, 4.1, 6.0, 6.5, 5.8, 5.2, 6.8, 7.4, 6.0, 5.6, 7.5, 7.8, 6.3, 5.9, 8.0, 8.4)

plot(t, ventas, main = "Grafico de dispersión de ventas", xlab = "Trimestre", ylab = "Ventas (miles)", p
```



Realiza el análisis de tendencia y estacionalidad

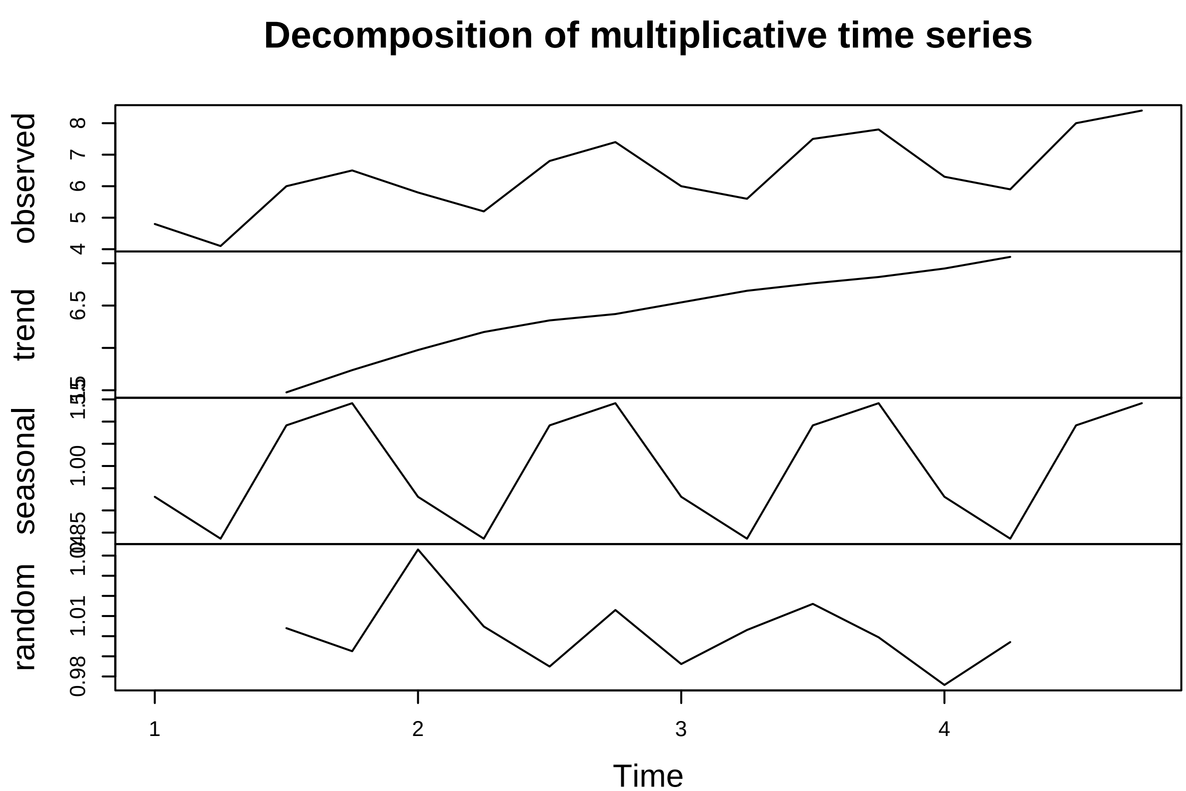
```r
T = ts(ventas, frequency = 4, start(c(2016, 1)))
plot(T, main = "Serie de tiempo", xlab = "Año", ylab = "Ventas (miles)", pch = 16, col = "blue")
```



###

Descompón la serie en sus 3 componentes e interprétalos

```r
D = decompose(T, type = "m")
plot(D)
```



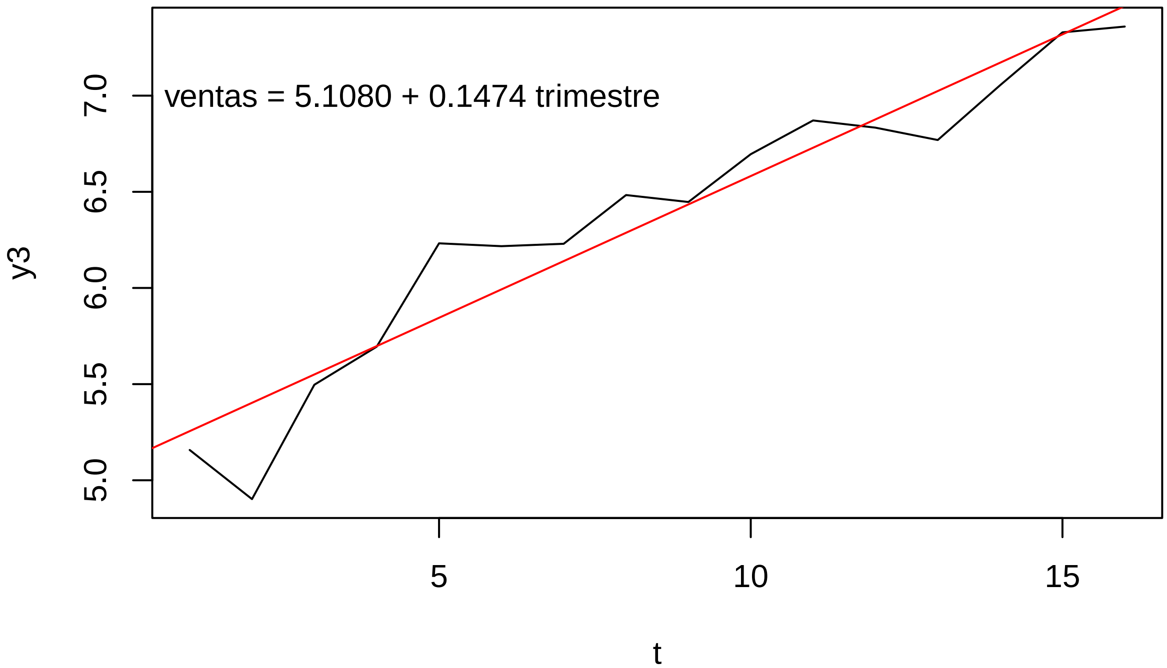
Observamos una serie de 4 estaciones con una tendencia a la alza año tras año. Al dividir la serie de tiempo en sus componentes **estacional**, **tendencia** y **aleatorio**, podemos apreciar que el trimestre con menos ventas es el segundo trimestre del año, este valor sube considerablemente hacia el tercer trimestre y vuelve a incrementar ligeramente hacia el cuarto trimestre del año, para bajar considerablemente el primer trimestre del año, y volver a tocar fondo en el segundo trimestre.

La tendencia indica un incremente considerable el primer año, y un incremento más bien modesto en los años siguientes.

El factor aleatorio parece no mostrar ningún comportamiento en particular.

Modelo lineal

```
ventas_desestacionalizadas = (D$x)/(D$seasonal)
y3 = ventas_desestacionalizadas
N3 = lm(y3~t)
plot(t, y3, type = "l")
abline(N3, col = "red")
text(4.5, 7, " ventas = 5.1080 + 0.1474 trimestre")
```




```
summary(N3)
```

```
##
## Call:
## lm(formula = y3 ~ t)
##
## 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 ***
## t            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
```

Significancia de β_1

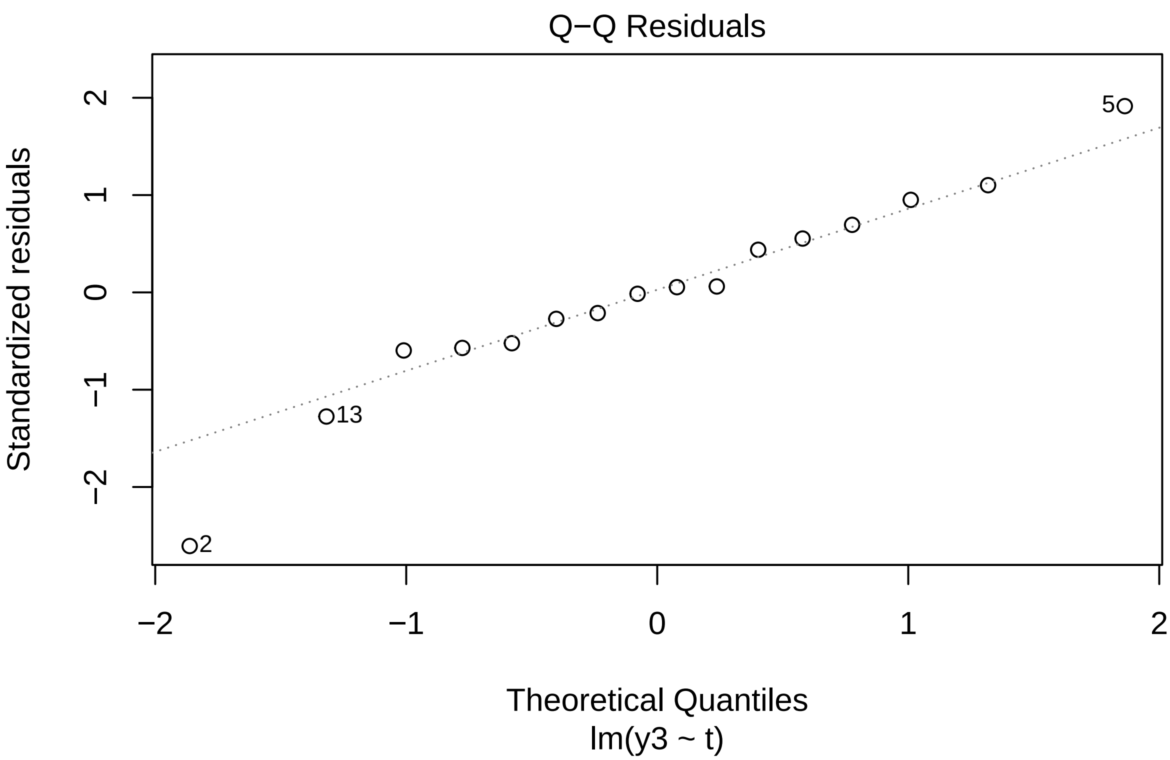
Tomando en cuenta el valor P cercano a 0, podemos rechazar la hipotesis nula de $\beta_1 = 0$, además de ser significativamente diferente de 0.

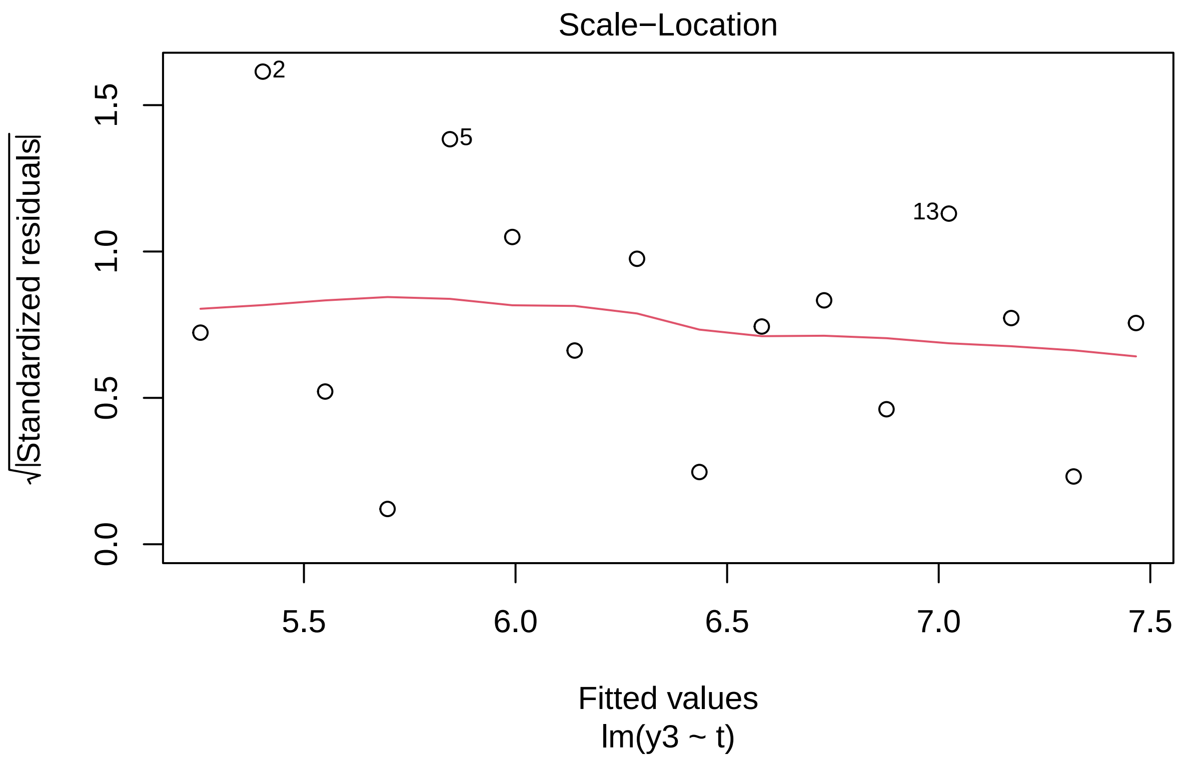
Variabilidad explicada por el modelo

Además de contar con una significancia elevada, contamos con un coeficiente de determinación (R^2) de 0.92, lo que indica que el 92% de la variabilidad del modelo puede ser explicada por la variable t, lo cual es un resultado satisfactorio, y que apoya la idea de la funcionalidad del modelo.

Análisis de residuos

```r
plot(N3, which = c(2,3))
```





En estas graficas, podemos apreciar 2 cosas, la distribución de los resuduos, y la normalidad de los mismos. La qqplot sugieren que los residuos tienen una distribución normal, alrededor de las predicciones efectuadas por el modelo, y la grafica de disperción sugiere homocedasticidad en los residuos, lo que quiere decir que no existen tendencias notables en la distribución de los residuos alrededor del modelo.

Prueba de normalidad

```
shapiro.test(N3$residuals)
```

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

H_0 los datos muestran una distribución normal H_1 los no datos muestran una distribución normal $\alpha = 0.05$

p value $= 0.7303$

Se rechaza la hipotesis nula, los datos no muestran una distribución normal.

CME y EMAP

```
predictions <- predict(N3)
CME <- mean((T - predictions)^2, na.rm = TRUE)
EPAM <- mean(abs((T - predictions) / T) * 100, na.rm = TRUE)
cat("CME:", CME, "\n")
```

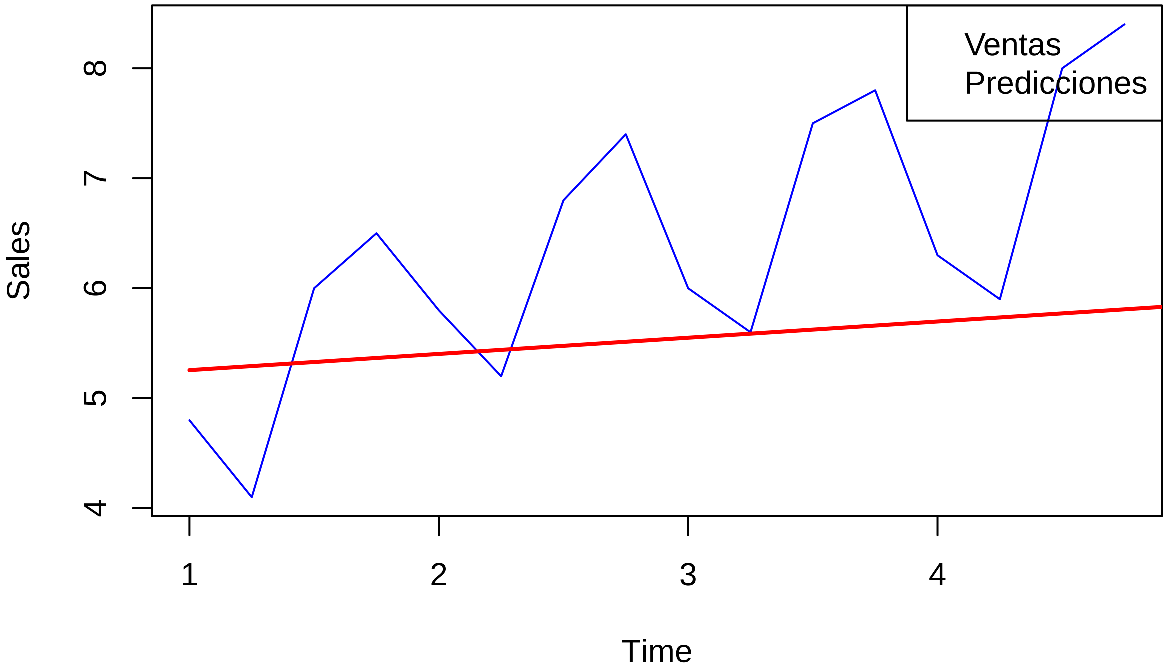
```
## CME: 0.6957218
```

```
cat("EMAP:", EPAM, "\n")
```

```
## EMAP: 12.5894
```

Graficación

```r
plot(T, col = "blue", ylab = "Sales")
lines(predictions, col = "red", lwd = 2)
legend("topright", legend = c("Ventas", "Predicciones"), col = c("blue", "red"))
```



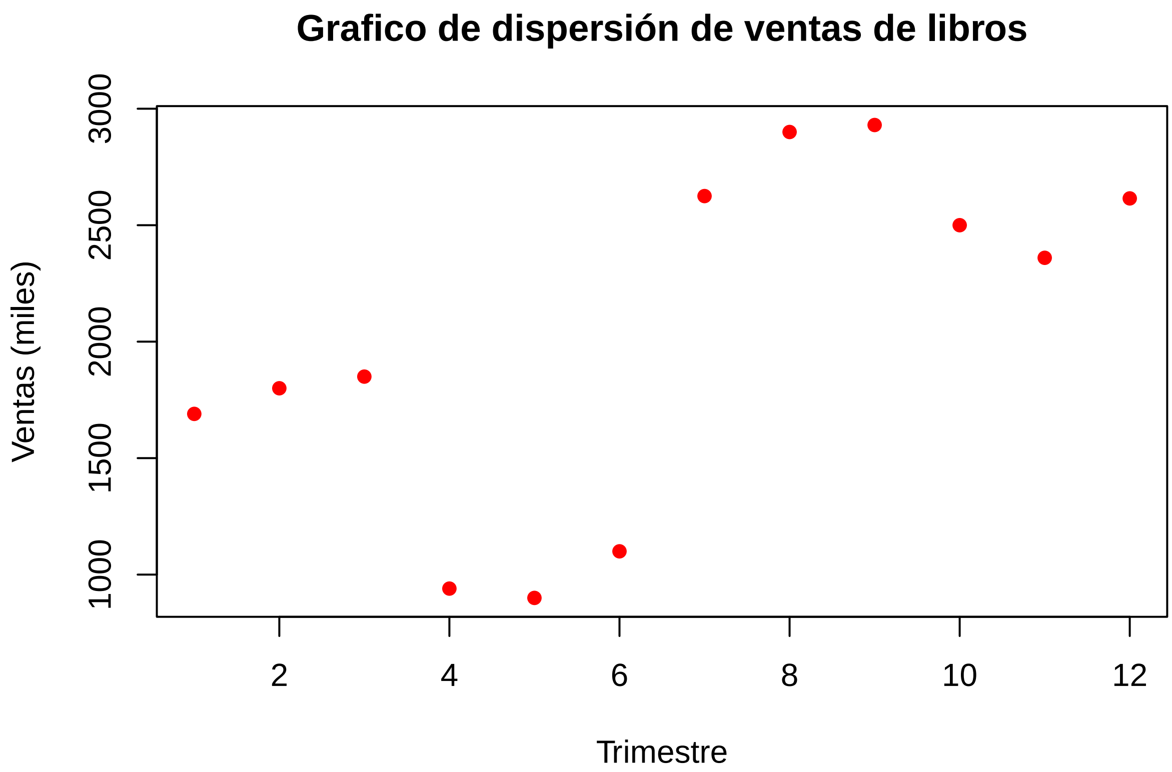
El modelo lineal se queda corto en este caso para describir la naturaleza estacional de nuestro conjunto de datos. Si bien el modelo lineal predice de manera apropiada la tendencia sin estaciones (como se demostró con las pruebas de significancia, residuos y normalidad), podría no ser lo más apropiado tratar de utilizar este modelo para describir datos que varían de esta forma. Sin embargo, puede utilizarse para construir un modelo que tome en cuenta las variaciones estacionales.

Ventas de libros

```r
ventas2 = c(1690, 1800, 1850, 940, 900, 1100, 2625, 2900, 2930, 2500, 2360, 2615)

T2 <- ts(ventas2, frequency = 4, start = c(1, 1))

plot(1:12, ventas2, main = "Grafico de dispersión de ventas de libros", xlab = "Trimestre", ylab = "Ven
```



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

```
promedios_moviles_anuales <- filter(T2, rep(1/4, 4), sides = 1)
promedios_moviles_centrados <- filter(T2, rep(1/4, 4), sides = 2)

print("Promedios móviles de cuatro trimestres:")
```

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

```
print(promedios_moviles_anuales)
```

```
##      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(promedios_moviles_centrados)
```

```
##      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
```

Calcule los índices estacionales de los cuatro trimestres

```
promedio_trimestre <- tapply(ventas2, cycle(ventas2), mean)

indices_estacionales <- (ventas2 / promedio_trimestre) * 100
```

```
## Warning in ventas2/promedio_trimestre: Recycling array of length 1 in vector-array arithmetic is dep
##   Use c() or as.vector() instead.
```

```
indices_estacionales
```

```
##  [1]  83.76704  89.21933  91.69765  46.59232  44.60967  54.52292 130.11152
##  [8] 143.74226 145.22924 123.91574 116.97646 129.61586
```

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

resultado? ¿Por qué?

La editorial obtiene el máximo indice estacional en el primer trimestre del tercer año, lo cual parece un resultado razonable porque este es el año que presentó una mayor variación en ventas en general, a su vez, también es el valor más alto alcanzado por la gráfica, lo cual solidifica la idea de que esta temporada en particular presentó mayor variabilidad.