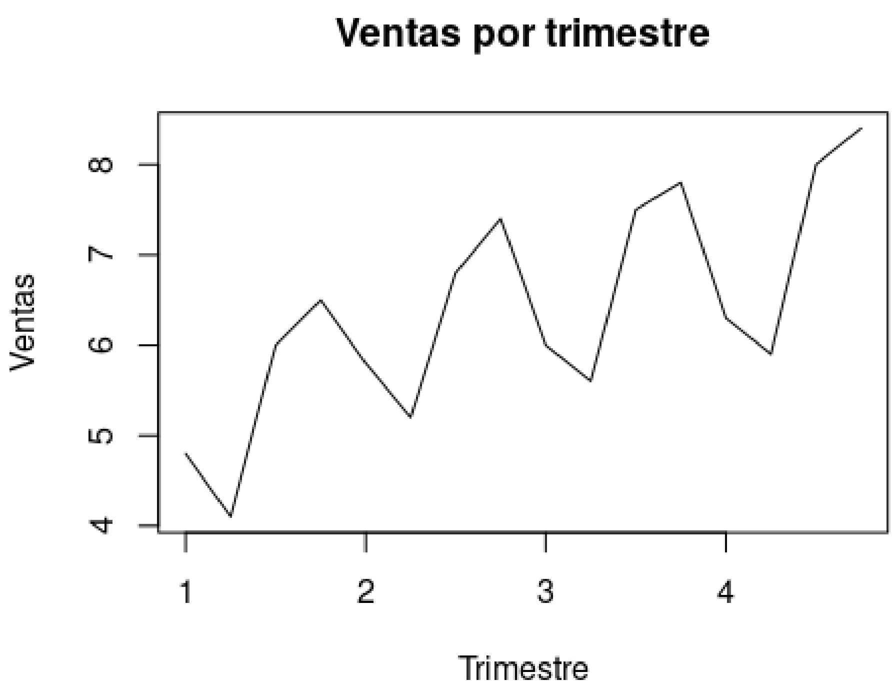
A8-Series de tiempo no estacionarias. Tendencia

2023-11-15

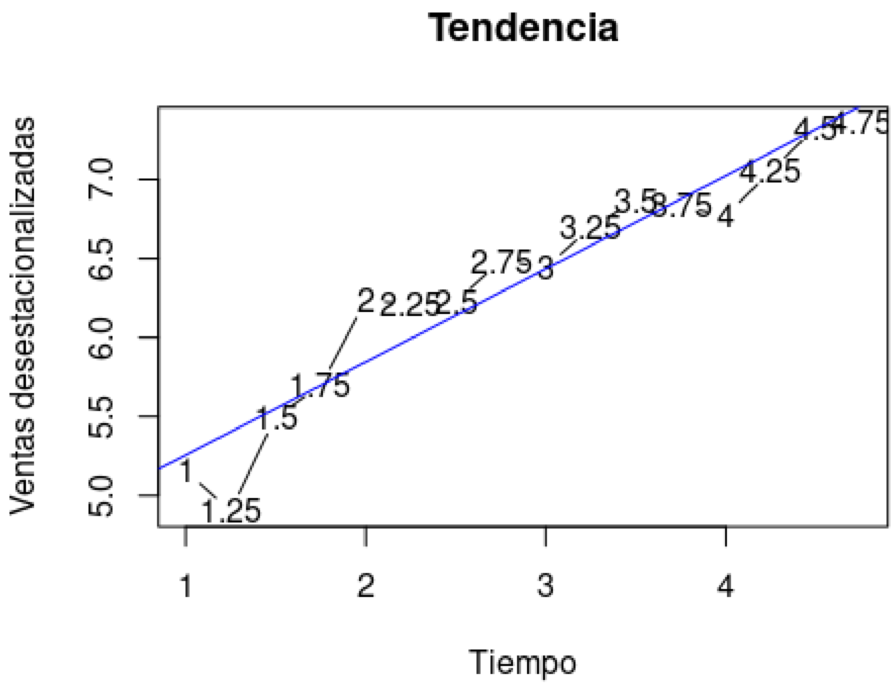
R markdown

```
# Creamos el data set de series de tiempo
ventas_tv <- 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)

T <- ts(ventas_tv, frequency = 4, start = c(1, 1))

# plot de dispersion
plot(T, main = "Ventas por trimestre", xlab = "Trimestre", ylab = "Ventas")
```




```
# Decomposicion
D <- decompose(T, type = "multiplicative")

# Componentes de tendencia
trend <- D$trend
seasonal <- D$seasonal
residuals <- D$residuals
```

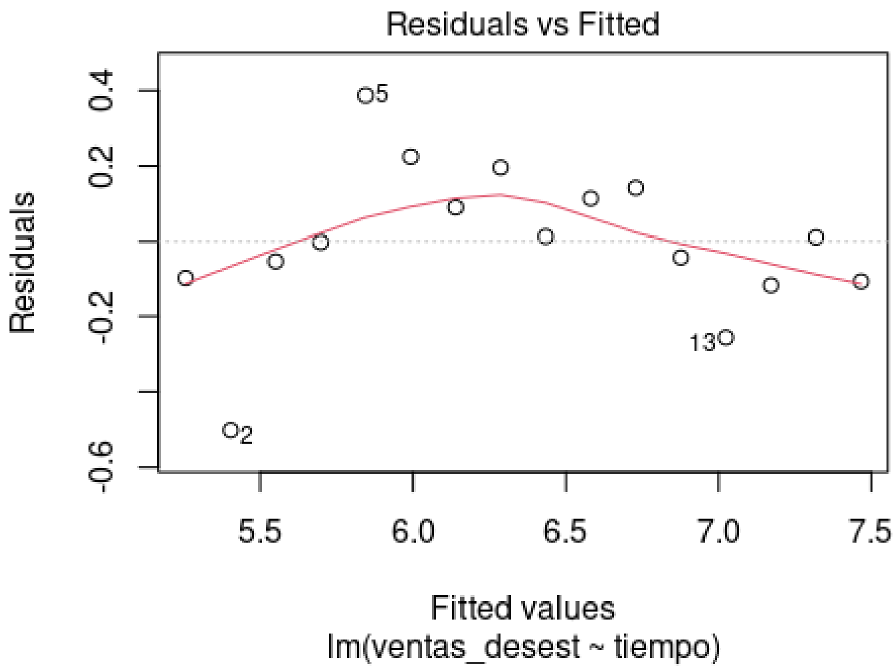
```r
# Analisis de regresion
datos <- data.frame(tiempo = time(T), ventas_desest = ventas_tv / D$seasonal)
modelo <- lm(ventas_desest ~ tiempo, data = datos)

# Mapeo de regresion
plot(datos$tiempo, datos$ventas_desest, main = "Tendencia", xlab = "Tiempo",
ylab = "Ventas desestacionalizadas")
abline(modelo, col = "blue")
```



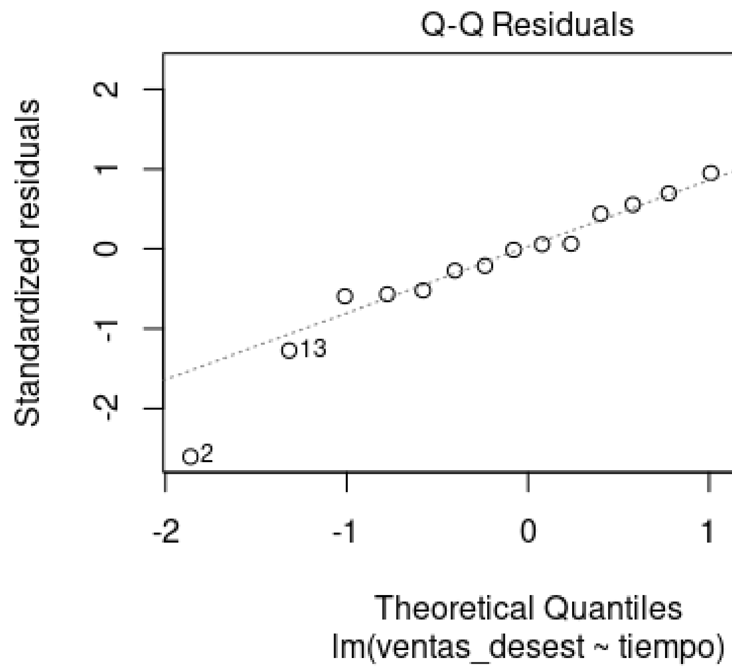

```r
# Analisis del modelo lineal
summary(modelo)

##
## Call:
## lm(formula = ventas_desest ~ tiempo, data = datos)
##
## 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)  4.66590    0.14313   32.60 1.32e-14 ***
## tiempo       0.58953    0.04621   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
```

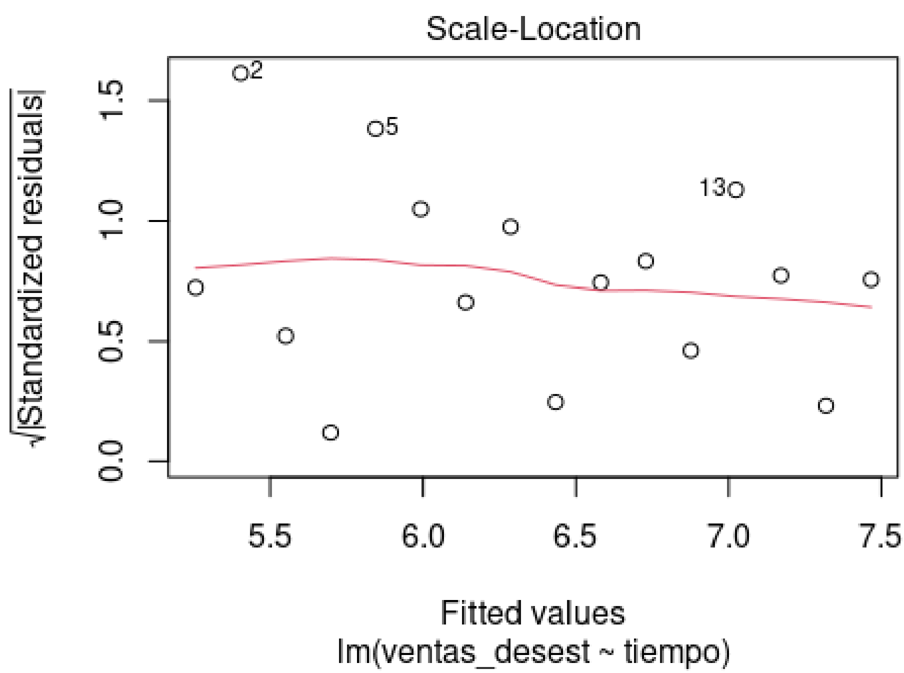
```r
# Analisis del residuo
plot(modelo, which = 1)
```



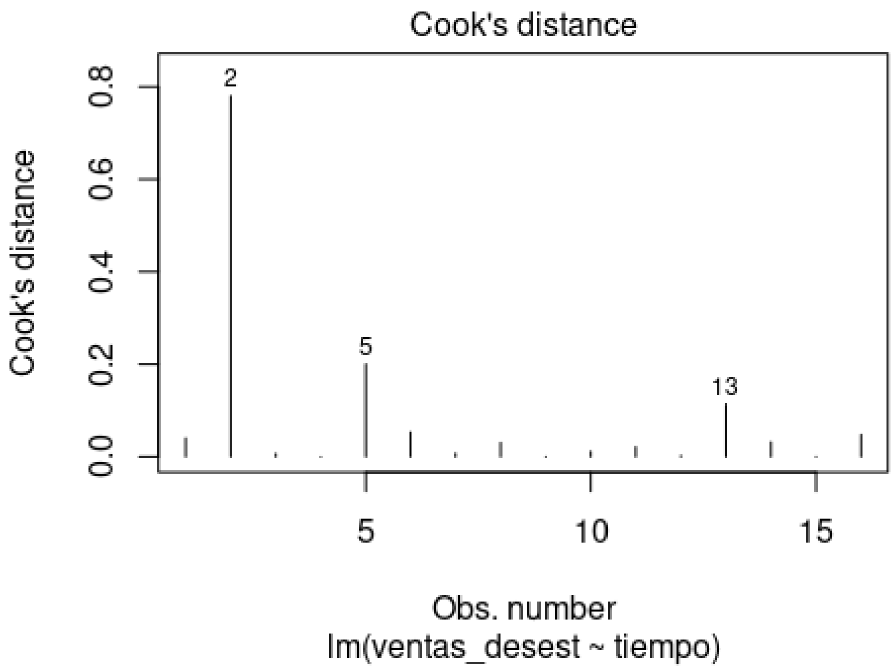

```r
plot(modelo, which = 2)
```



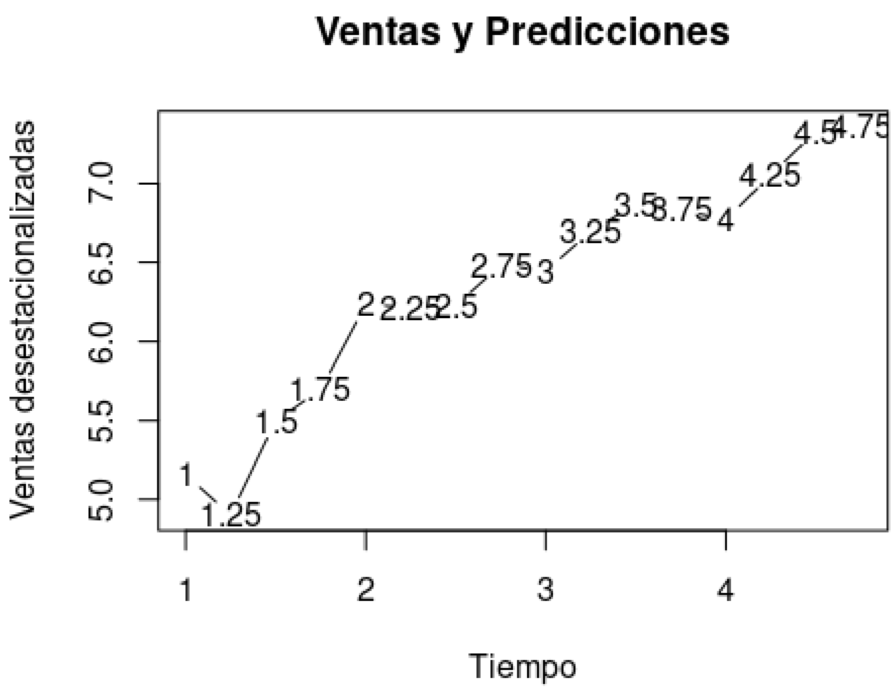

```r
plot(modelo, which = 3)
```




```r
plot(modelo, which = 4)
```



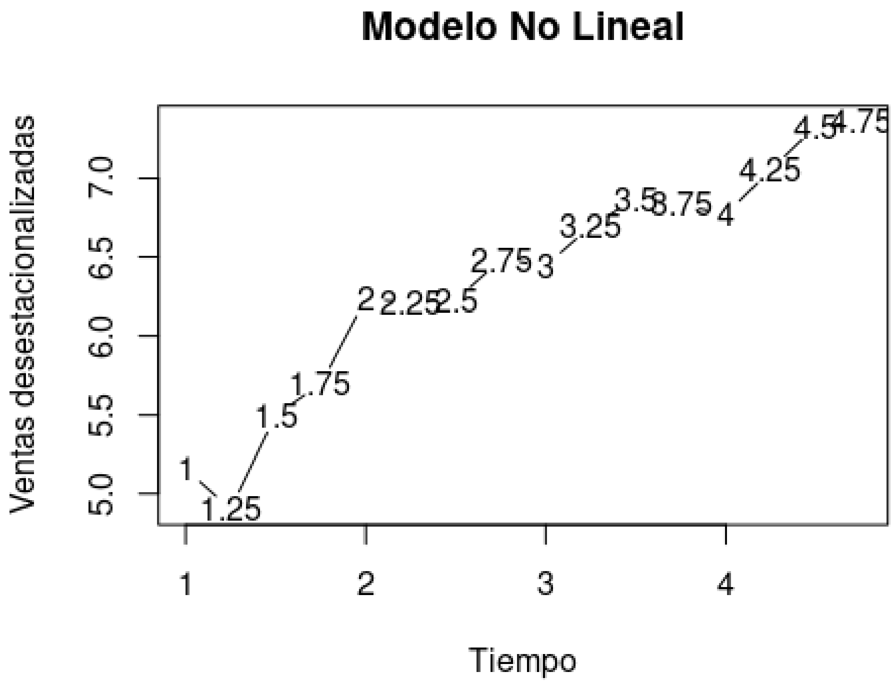

```
# EMC del resultado
CME <- mean(modelo$residuals^2)

# Error de porcentaje
EPAM <- mean(abs((modelo$residuals / datos$ventas_desest) * 100), na.rm =
TRUE)

# Predicciones de las ventas
predicciones <- predict(modelo, data.frame(tiempo = datos$tiempo))

# Plot ventas_desest and predictions
plot(datos$tiempo, datos$ventas_desest, main = "Ventas y Predicciones", xlab
= "Tiempo", ylab = "Ventas desestacionalizadas", type = "l")

# Add the regression line using predicted values
lines(datos$tiempo, predicciones, col = "blue", type = "l")
```



Conclusion sobre el modelo

El coeficiente de la variable tiempo es 0.58953. Esto sugiere que por cada cambio de unidad a lo largo del tiempo, se espera que las ventas ajustadas estacionalmente aumenten en aproximadamente esta cantidad de unidades. Este coeficiente es significativamente diferente de cero, lo que da como resultado un valor p muy bajo. La interseccion es 4.66590. Representa el valor esperado de las ventas desestacionalizadas cuando el tiempo es cero. El valor p asociado con la interseccion es muy bajo, esto indica que hay una evidencia significativa para rechazar la hipotesis nula de que la interseccion es cero. Tambien con de acuerdo a los residuales podemos concluir que el modelo captura bien la variabilidad. Podemos aplicar un modelo no lineal para hacer las comparacion. En este caso estaremos aplicando el modelo en segundo grado para observar los resultados y decidir si se mejoran los resultados

```r
# Ajuste de un modelo no lineal de segundo grado
modelo_no_lineal <- lm(ventas_desest ~ tiempo + I(tiempo^2), data = datos)

# Visualizamos el ajuste
plot(datos$tiempo, datos$ventas_desest, main = "Modelo No Lineal", xlab =
"Tiempo", ylab = "Ventas desestacionalizadas")
lines(datos$tiempo, predict(modelo_no_lineal), col = "blue", type = "l")
```




```
summary(modelo_no_lineal)

##
## Call:
## lm(formula = ventas_desest ~ tiempo + I(tiempo^2), data = datos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.36986 -0.07058 -0.00100  0.11345  0.33110
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.97430    0.28807  13.796 3.86e-09 ***
## tiempo       1.16274    0.22054   5.272 0.000151 ***
## I(tiempo^2) -0.09969    0.03776  -2.640 0.020392 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1784 on 13 degrees of freedom
## Multiple R-squared:  0.9484, Adjusted R-squared:  0.9405
## F-statistic: 119.6 on 2 and 13 DF,  p-value: 4.268e-09
```

Comparacion

Parece ser que este modelo tiene un mejor ajuste que el modelo anterior, el estadistico F es mas alto. El R^2 es mas alto, y el error estandar residual es menor. Tal que el modelo no lineal explica una mayor proporcion de la variabilidad en las ventas desestacionalizadas, y que el modelo no lineal tiene un mejor ajuste a los datos. Usaremos este modelo para predecir las ventas en el siguiente tiempo.

```
# Crear un nuevo dataframe con el año siguiente
new_year <- data.frame(tiempo = seq(max(datos$tiempo) + 1, max(datos$tiempo)
+ 4, by = 1))

# Realizamos la prediccion
prediccion <- predict(modelo_no_lineal, newdata = new_year)

# Visualizamos la prediccion
print(prediccion)
```

```
##        1        2        3        4
## 7.364089 7.280715 6.997963 6.515833
```