A8_series de tiempo_fer

Fernanda Pérez

2024-11-14

Analizar los pronósticos en series de tiempo

Para los datos de las ventas de televisores analiza la serie de tiempo más apropiada:

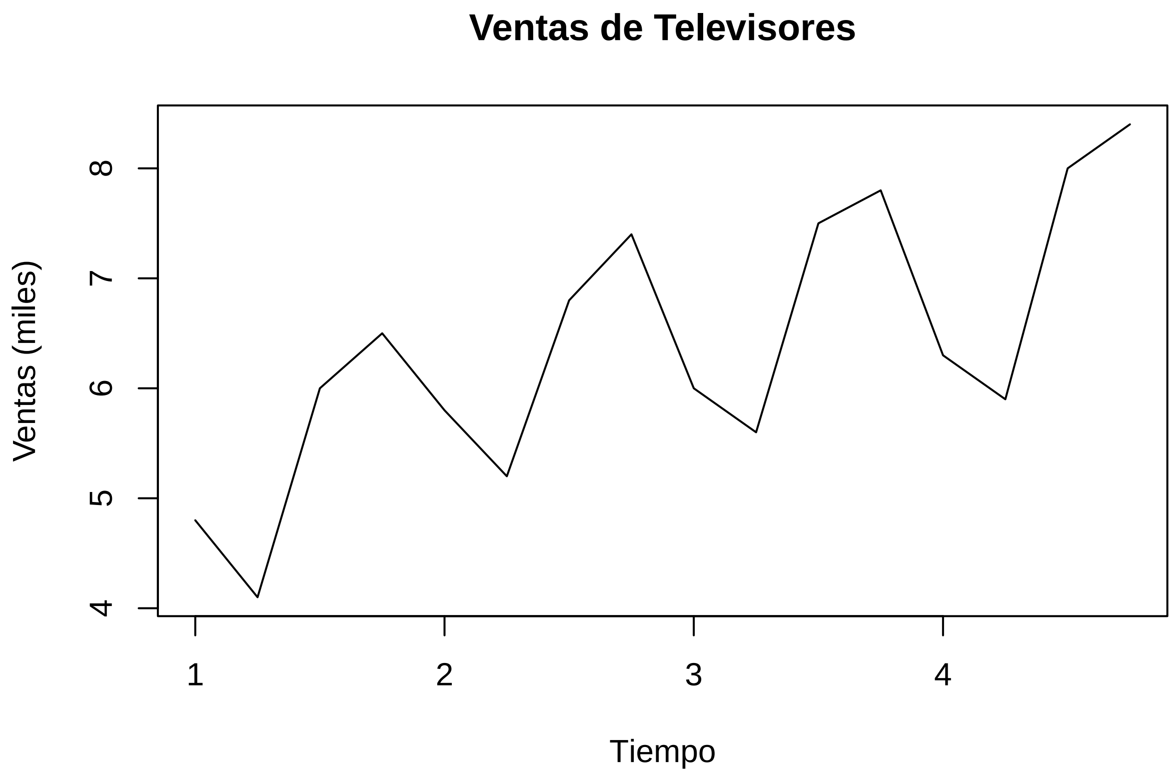
A) Realiza el análisis de tendencia y estacionalidad:

Identifica si es una serie estacionaria

La serie de ventas de televisores con la que estamos trabajando en esta actividad no pareciera ser estacionaria. Recordando brevemente que una serie estacionaria debe cumplir con: mantener una media y varianza constantes a lo largo del tiempo, sin mostrar tendencias o patrones cíclicos. Pero en este ejercicio podemos ver un aumento gradual en las ventas a lo largo de los trimestres, indicandonos que hay una tendencia ascendente. ´Podemos deducir que la serie tiene un componente de tendencia y posiblemente estacionalidad, características que no son propias de una serie estacionaria.

Grafica la serie para verificar su tendencia y estacionalidad

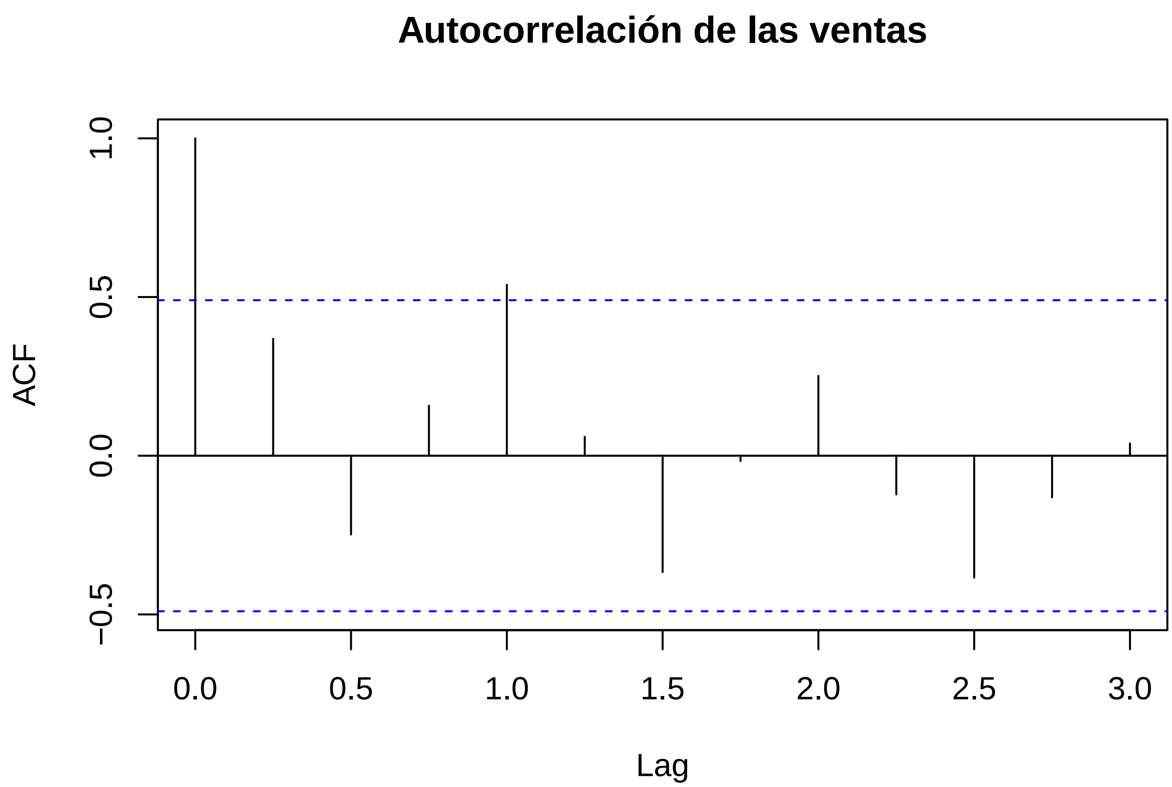
```r
ventas <- c(4.8, 4.1, 6.0, 6.5, 5.8, 5.2, 6.8, 7.4, 6.0, 5.6, 7.5, 7.8, 6.3, 5.9, 8.0, 8.4)
tiempo <- ts(ventas, start = c(1, 1), frequency = 4)

plot(tiempo, main = "Ventas de Televisores", ylab = "Ventas (miles)", xlab = "Tiempo")
```



En la gráfica de la serie de tiempo de ventas de televisores que estamos observando, podemos ver que hay una tendencia ascendente conforme avanzan los trimestres. Los valores de ventas van creciendo, indicandonos una tendencia positiva en las ventas a lo largo del tiempo. Tambien, observamos picos y valles en intervalos regulares, lo cual sugiere la presencia de estacionalidad, con fluctuaciones recurrentes posiblemente relacionadas con la variación estacional de la demanda.

Analiza su gráfico de autocorrelación

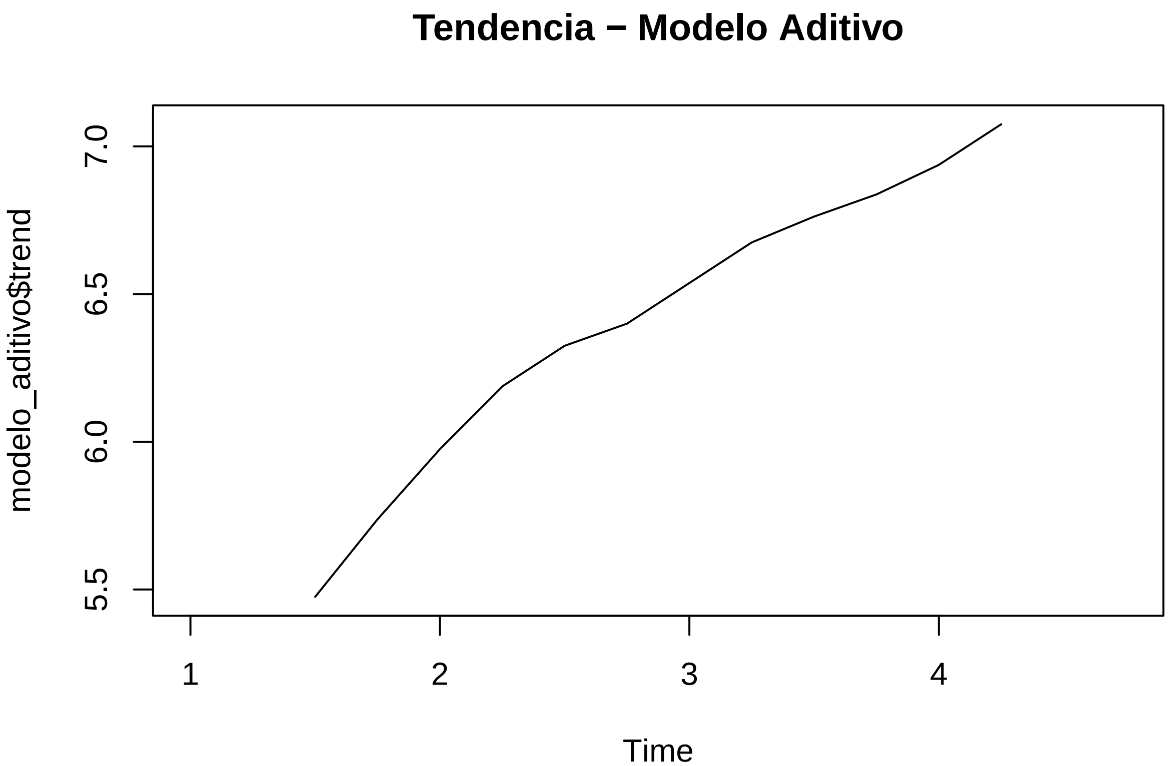
```
acf(tiempo, main = "Autocorrelación de las ventas")
```



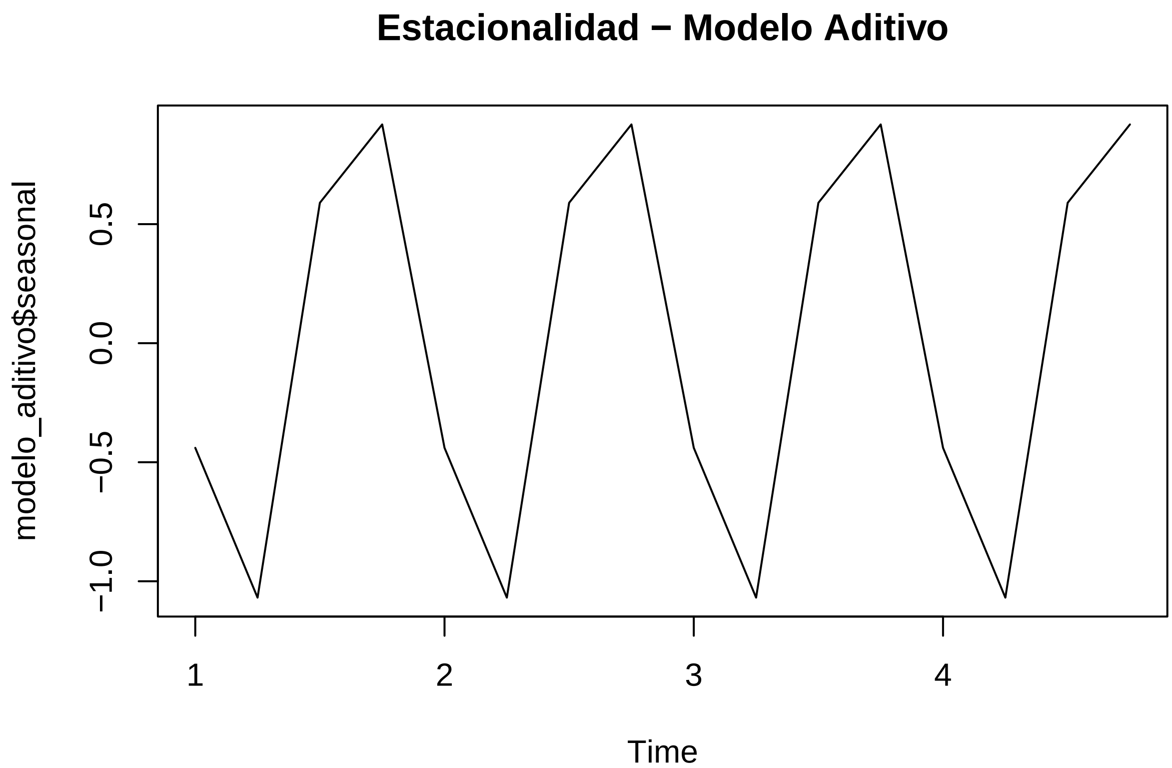
En el gráfico de autocorrelación que estamos viendo podemos observar valores significativos en los primeros rezagos, lo que respalda la presencia de dependencia temporal en la serie. La autocorrelación en el rezago 1 es muy cercana a 1, lo cual nos indica quw hay una fuerte relación entre las ventas de trimestres consecutivos. Y si seguimos observando podemos identificar que algunos de los rezagos posteriores también presentan correlaciones significativas, lo cual refuerza la idea de que la serie presenta una estructura de dependencia temporal y posiblemente estacionalidad.

Identifica si el modelo puede ser sumativo o multiplicativo (puedes probar con ambos para ver con cuál es mejor el modelo)

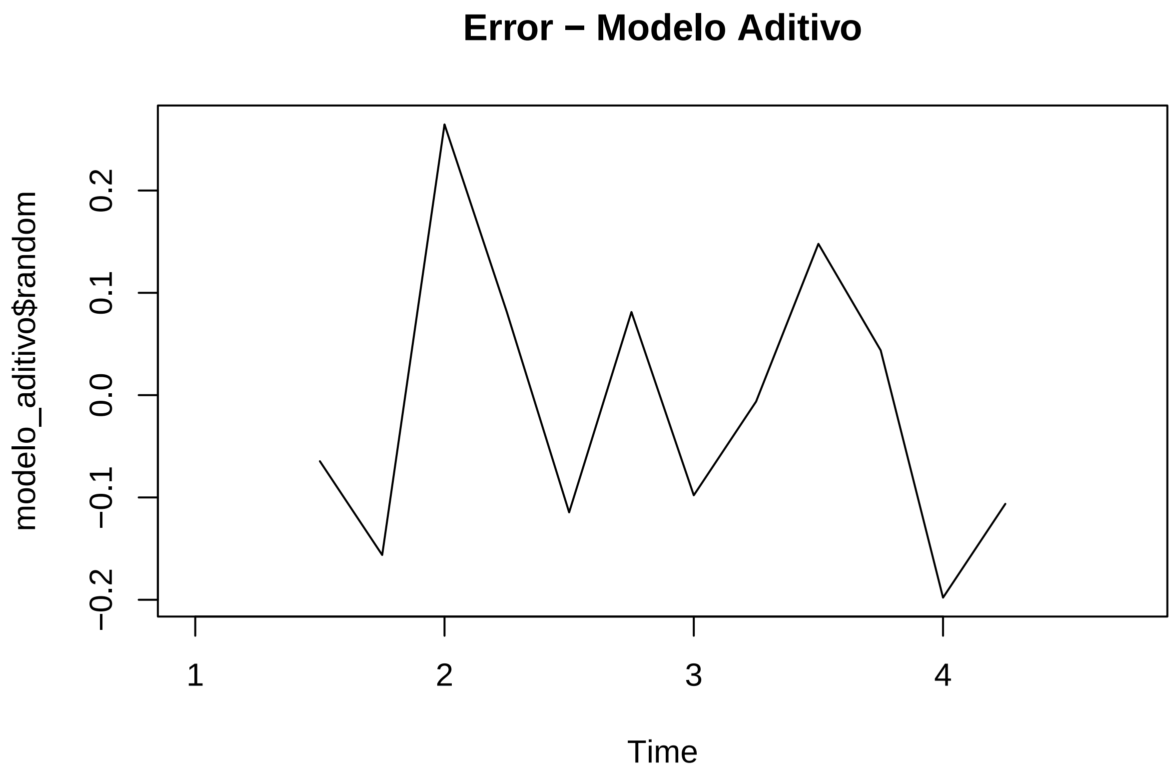
```
modelo_aditivo <- decompose(tiempo, type = "additive")
modelo_multiplicativo <- decompose(tiempo, type = "multiplicative")

plot(modelo_aditivo$trend, main = "Tendencia - Modelo Aditivo")
```



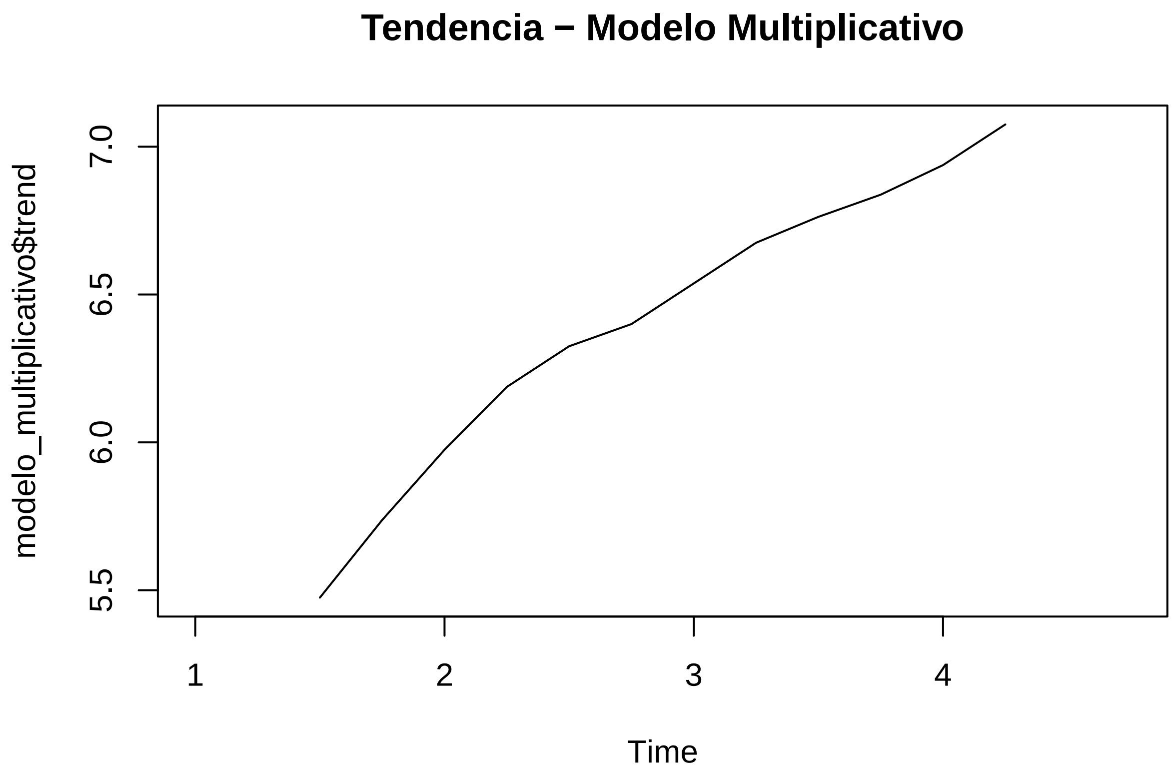

```r
plot(modelo_aditivo$seasonal, main = "Estacionalidad - Modelo Aditivo")
```



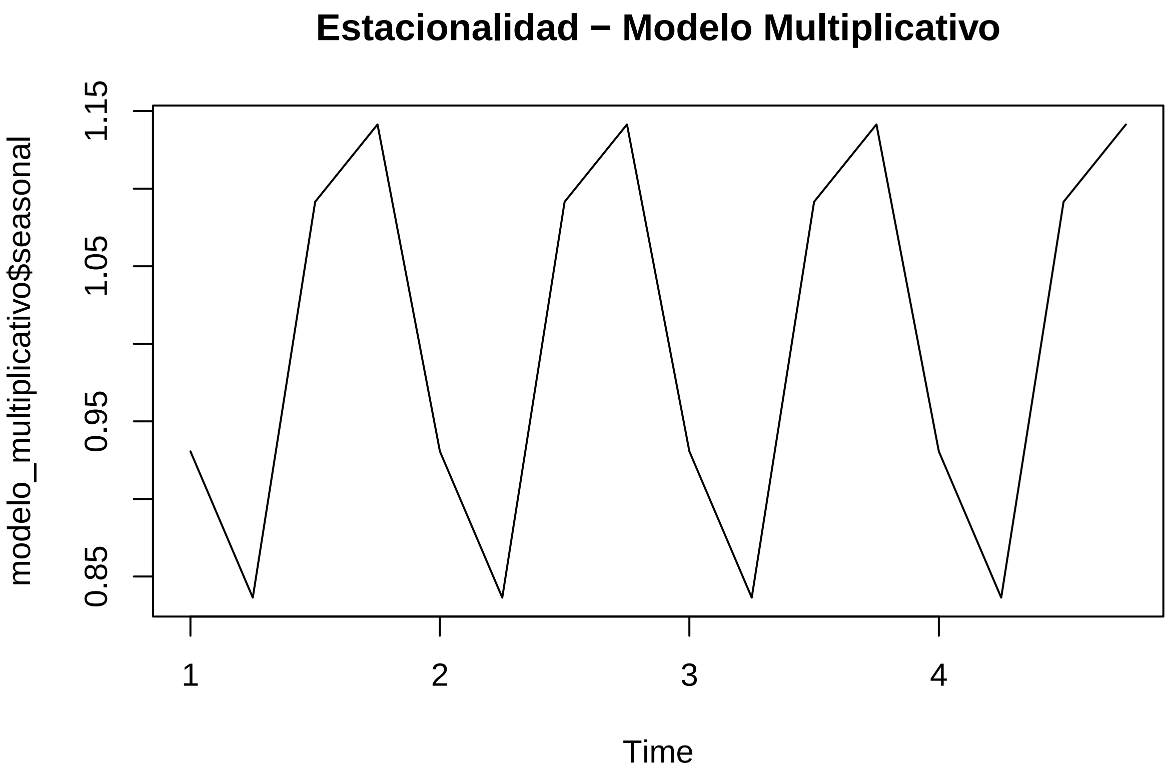

```r
plot(modelo_aditivo$random, main = "Error - Modelo Aditivo")
```




```r
plot(modelo_multiplicativo$trend, main = "Tendencia - Modelo Multiplicativo")
```




```r
plot(modelo_multiplicativo$seasonal, main = "Estacionalidad - Modelo Multiplicativo")
```




```r
plot(modelo_multiplicativo$random, main = "Error - Modelo Multiplicativo")
```

Se hicieron las graficas para analizar los 2 modelos (sumativo y multiplicativo) para ver cual es mejor, vemos que los dos logran capturar la tendencia ascendente y la estacionalidad de la serie. Pero en el modelo multiplicativo, la estacionalidad y el error pareciera que se adaptan un poco mejor al comportamiento de los datos, dado que la amplitud de las fluctuaciones estacionales aumenta con el tiempo, esto nos indica que el modelo multiplicativo es el que será más adecuado para representar esta serie de ventas con las que estamos trabajando, dado que refleja mejor las variaciones proporcionales de los datos.

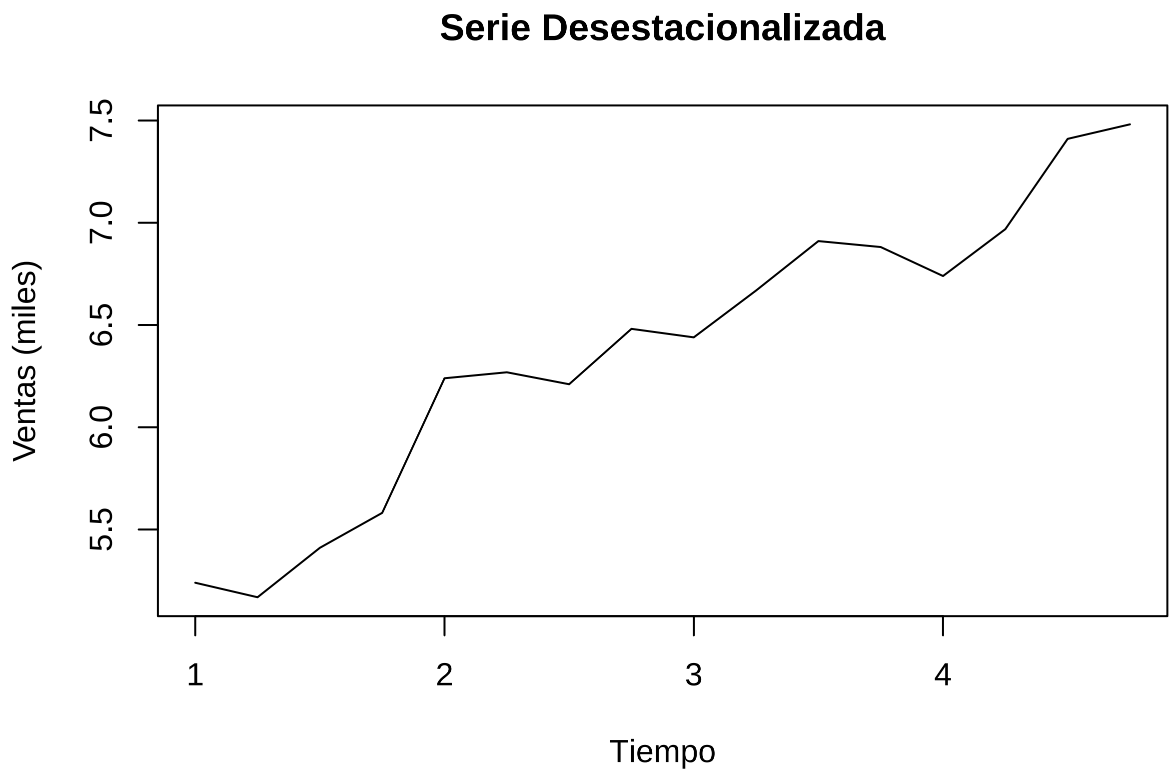
B) Calcula los índices estacionales

```
indices_estacionales <- modelo_aditivo$seasonal
print(indices_estacionales)
```

```
##         Qtr1       Qtr2       Qtr3       Qtr4
## 1 -0.4395833 -1.0687500  0.5895833  0.9187500
## 2 -0.4395833 -1.0687500  0.5895833  0.9187500
## 3 -0.4395833 -1.0687500  0.5895833  0.9187500
## 4 -0.4395833 -1.0687500  0.5895833  0.9187500
```

y grafica la serie desestacionalizada

```
serie_desestacionalizada <- tiempo - indices_estacionales
plot(serie_desestacionalizada, main = "Serie Desestacionalizada", ylab = "Ventas (miles)", xlab = "Tiem
```



Ahora que hemos calculado los índices estacionales estos nos muestran los ajustes que se deben realizar en cada trimestre para eliminar el efecto de la estacionalidad en la serie. Vemos que hay valores negativos en el primer y segundo trimestre, indicandonos que las ventas han disminuido, por otro lado los valores positivos en el tercer y cuarto trimestre nos sugieren un aumento. La serie desestacionalizada que vemos en la gráfica, nos muestra la tendencia subyacente sin las fluctuaciones estacionales, con esto vemos un crecimiento mucho más estable y continuo en las ventas a lo largo del tiempo. Lo cualñ facilita el análisis de la tendencia real de la serie ya que no entra la influencia de los patrones estacionales.

C) Analiza el modelo lineal de la tendencia

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

```r
datos <- data.frame(Tiempo = 1:length(serie_desestacionalizada), Ventas = as.numeric(serie_desestacional

modelo_lineal <- lm(Ventas ~ Tiempo, data = datos)
summary(modelo_lineal)
```

```
##
## Call:
## lm(formula = Ventas ~ Tiempo, data = datos)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.2992 -0.1486 -0.0037  0.1005  0.3698
```

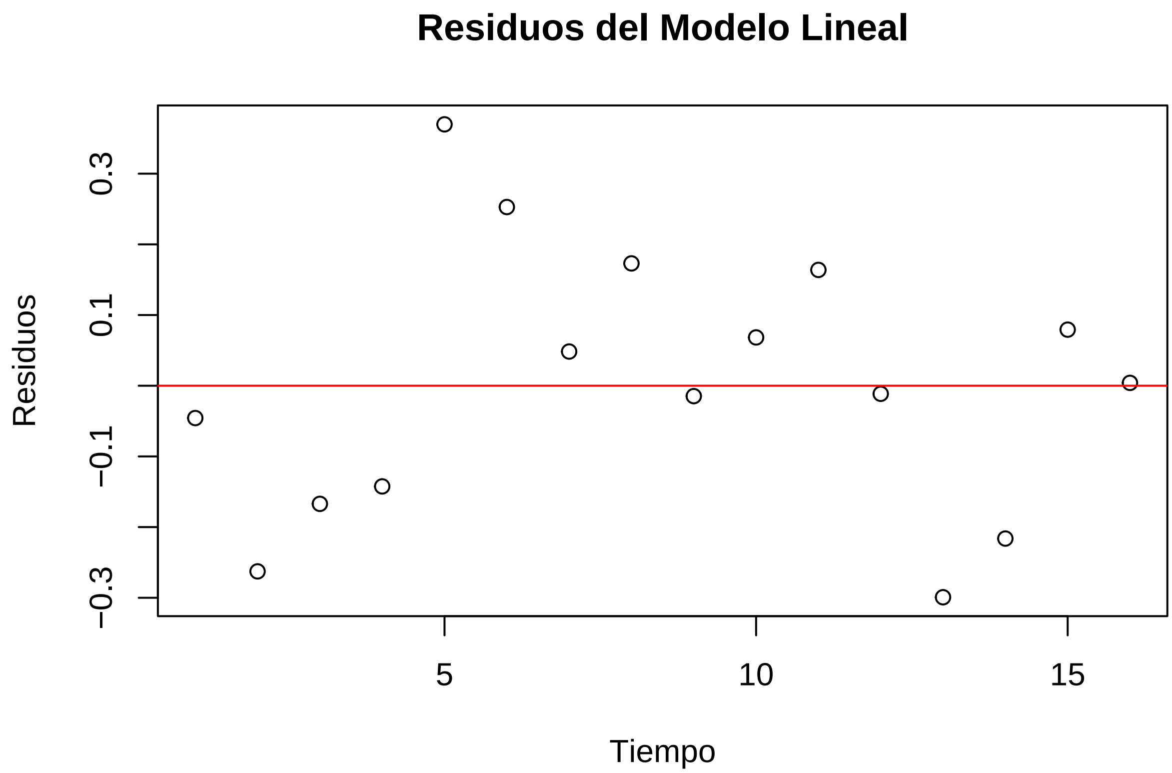
```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.13917    0.10172   50.52  < 2e-16 ***
## Tiempo       0.14613    0.01052   13.89  1.4e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.194 on 14 degrees of freedom
## Multiple R-squared:  0.9324, Adjusted R-squared:  0.9275
## F-statistic:   193 on 1 and 14 DF,  p-value: 1.399e-09
```

Analiza la significancia del modelo lineal, global e individual

Vemos que el análisis de la regresión lineal que se hizo nos muestra que el modelo es significativamente alto, tanto a nivel global como individual. A nivel global vemos que el valor p del modelo (1.399e-09) es muchisimo más pequeño que el nivel de significancia usual de 0.05, indicandonos que el modelo en su conjunto explica de manera significativa la variabilidad en las ventas desestacionalizadas. Y ahora analizandoolo a nivel individual, los coeficientes de la constante (intercepto) y de la variable Tiempo son también altamente significativos, ya que tiene valores p menores a 0.001. Como el coeficiente de Tiempo es positivo (0.14613), nos indica que hay una tendencia ascendente en las ventas desestacionalizadas. Y como tambien el R-cuadrado ajustado de 0.9275 sugiere que el modelo explica aproximadamente casi el 93% de la variabilidad de las ventas, lo cual es un muy porcentaje ya qye indica un buen ajuste.

Haz el análisis de residuos

```r
plot(residuals(modelo_lineal), main = "Residuos del Modelo Lineal", ylab = "Residuos", xlab = "Tiempo")
abline(h = 0, col = "red")
```



En este grafico de residuos del modelo lineal vemos que los residuos se distribuyen de forma aleatoria alrededor de la línea cero y no vemos ningun patron evidente, esto indica que el modelo lineal si es adecuado y que las suposiciones de homocedasticidad , o sea, de varianza constante y linealidad se cumplen razonablemente. El que no haya patrones claros nos dice que los errores no están correlacionados con el tiempo.

D) Calcula el CME y el EPAM de la predicción de la serie de tiempo

```r
errores <- residuals(modelo_lineal)
CME <- mean(errores^2, na.rm = TRUE)
print(paste("CME:", CME))
```

```
## [1] "CME: 0.0329191687091504"
```

y el EPAM de la predicción de la serie de tiempo

```r
EPAM <- mean(abs(errores / datos$Ventas), na.rm = TRUE) * 100
print(paste("EPAM:", EPAM))
```

```
## [1] "EPAM: 2.34131898056551"
```

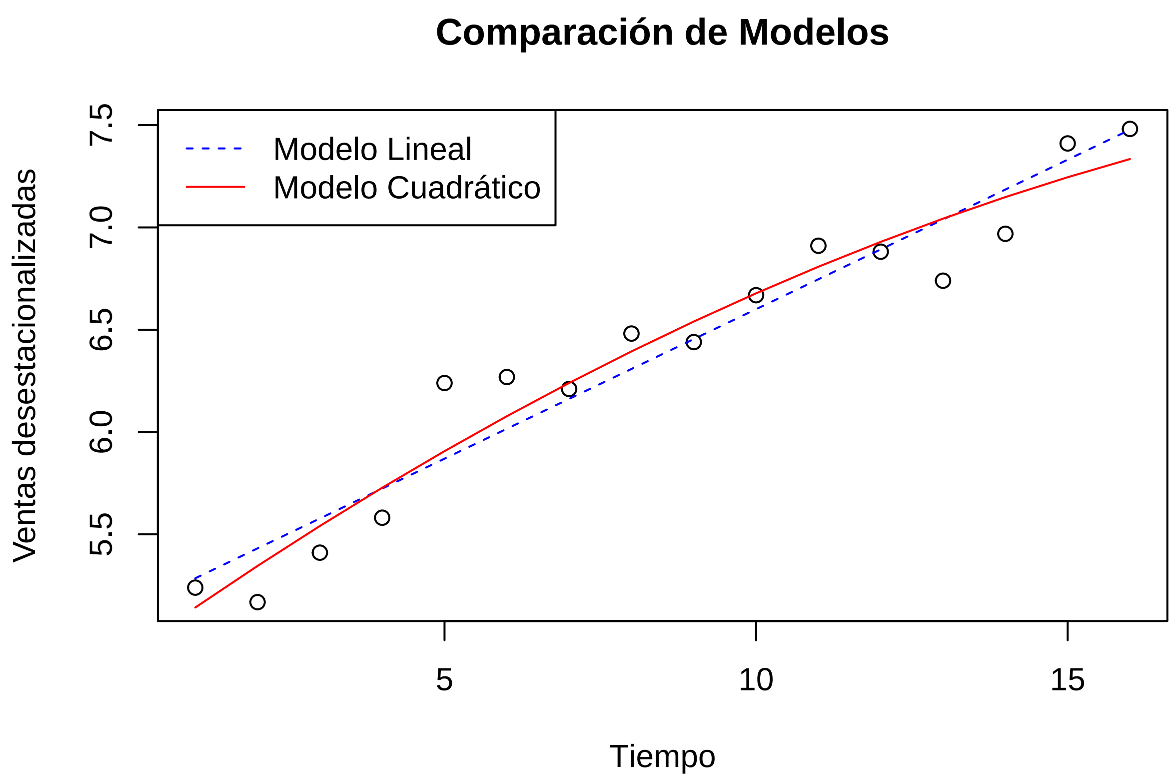
Se obtuvo que el cuadrado medio del error (CME) tuvo un valor de 0.0329, lo cual nos indica que el promedio de los errores al cuadrado es bajo, lo cual es un indicio de un buen ajuste del modelo lineal a los datos. Y como el error porcentual absoluto medio (EPAM) tuvo un valor de 2.34%, lo cual significa que el error promedio de las predicciones es del 2.34% respecto a las ventas reales. Los dos indicadores reflejan que el modelo tiene una buena precisión en la predicción de las ventas desestacionalizadas.

E) Explora un mejor modelo, por ejemplo un modelo cuadrático: y = B__0 + B__1 x + B__2 x^2. Para ello transforma la variable ventas (recuerda que la regresión no lineal es una regresión lineal con una tranformación).

```r
# modelo cuadratico
modelo_cuadratico <- lm(Ventas ~ Tiempo + I(Tiempo^2), data = datos)
summary(modelo_cuadratico)
```

```
##
## Call:
## lm(formula = Ventas ~ Tiempo + I(Tiempo^2), data = datos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.30333 -0.13440 -0.01928  0.11368  0.33301
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.930833   0.155679  31.673 1.08e-13 ***
## Tiempo       0.215572   0.042149   5.115 0.000199 ***
## I(Tiempo^2) -0.004085   0.002410  -1.695 0.113918
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1822 on 13 degrees of freedom
## Multiple R-squared:  0.9446, Adjusted R-squared:  0.9361
## F-statistic: 110.8 on 2 and 13 DF,  p-value: 6.805e-09
```

```r
plot(datos$Tiempo, datos$Ventas, main = "Comparación de Modelos", ylab = "Ventas desestacionalizadas", 
lines(datos$Tiempo, predict(modelo_lineal), col = "blue", lty = 2)
lines(datos$Tiempo, predict(modelo_cuadratico), col = "red", lty = 1)
legend("topleft", legend = c("Modelo Lineal", "Modelo Cuadrático"), col = c("blue", "red"), lty = c(2, 
```



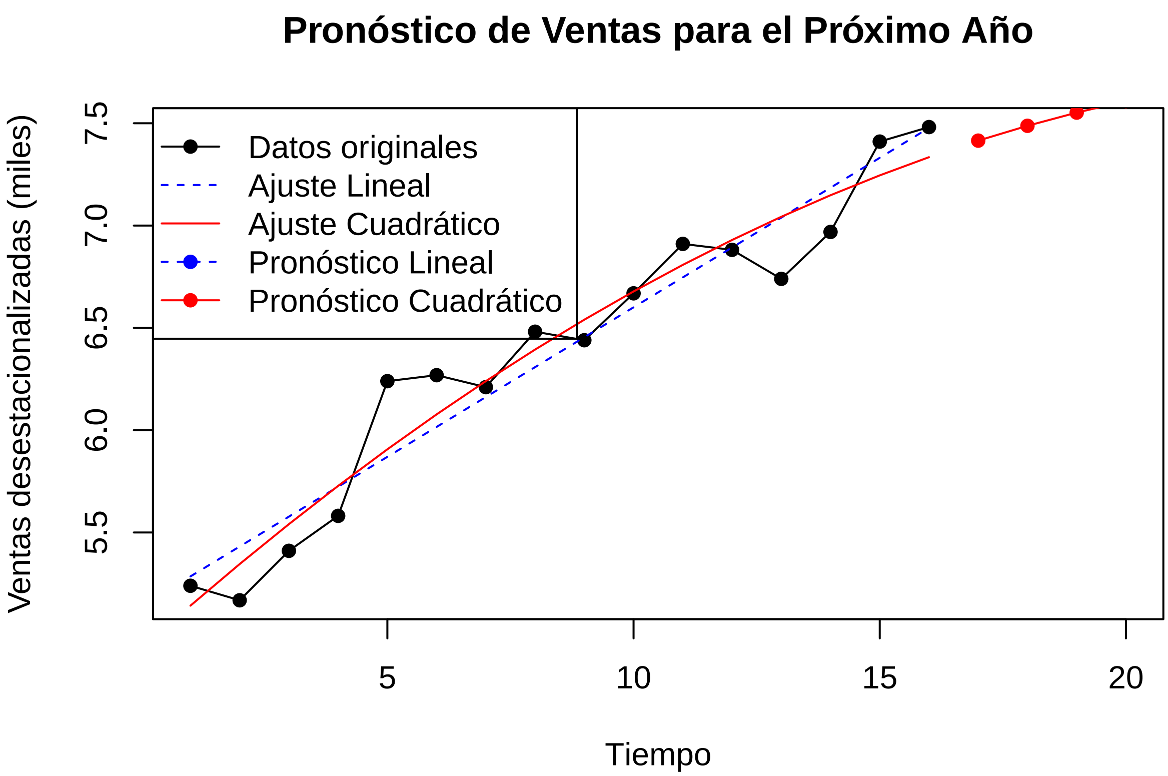
F) Concluye sobre el mejor modelo

Despues del analisis realizado al comparar el modelo lineal con el modelo cuadrático, observamos que ambos modelos presentan un buen ajuste a los datos, con altos valores de R-cuadrado. Pero vemos que el modelo cuadrático tiene un R-cuadrado un poco mayor (0.9446) y un error residual estándar menor, lo cualnos indica un mejor ajuste, sin embargo el término cuadrático no es estadísticamente significativo ($p > 0.05$), lo que sugiere que la diferencia en ajuste no es lo suficientemente relevante. Por tal motivo el modelo lineal es suficiente para capturar la tendencia de la serie de tiempo, siendo más simple y adecuado en este caso.

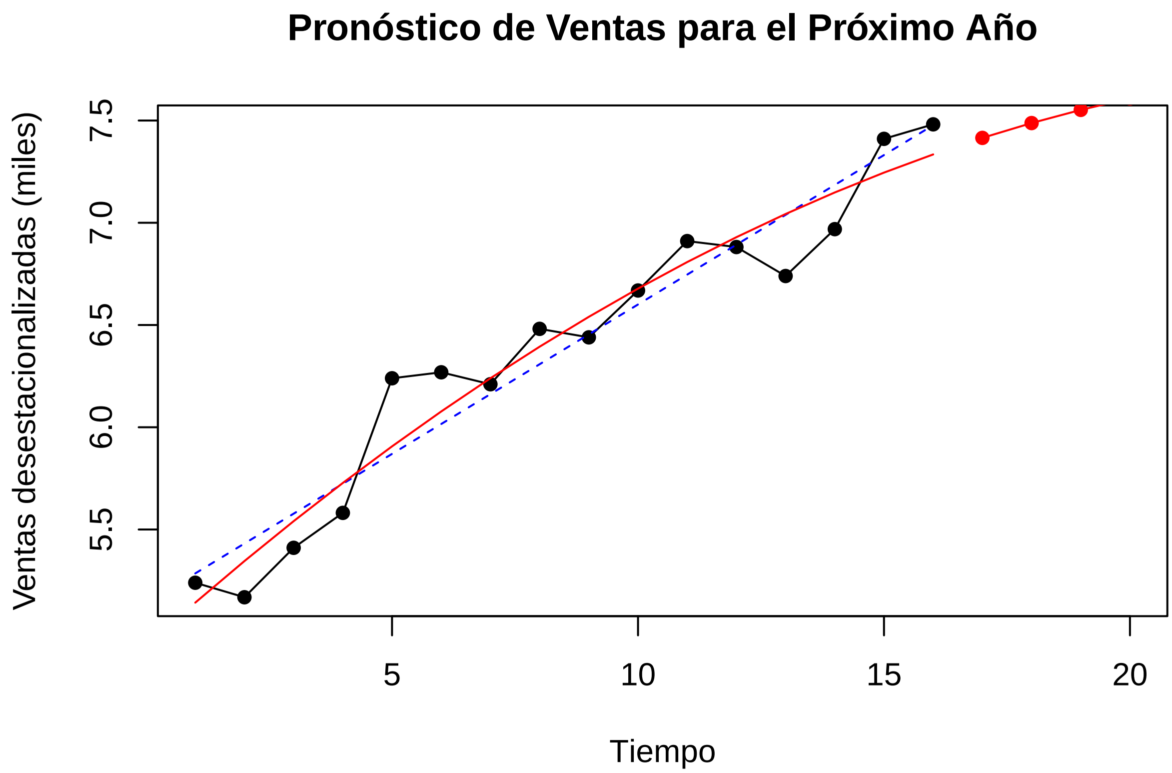
G) Realiza el pronóstico para el siguiente año y grafícalo junto con los pronósticos previos y los datos originales.

```r
tiempo_futuro <- data.frame(Tiempo = (length(datos$Tiempo) + 1):(length(datos$Tiempo) + 4))
pronostico_lineal <- predict(modelo_lineal, tiempo_futuro)
pronostico_cuadratico <- predict(modelo_cuadratico, tiempo_futuro)
plot(datos$Tiempo, datos$Ventas, type = "o", col = "black", pch = 16,
     main = "Pronóstico de Ventas para el Próximo Año",
     ylab = "Ventas desestacionalizadas (miles)", xlab = "Tiempo", xlim = c(1, length(datos$Tiempo) + 4)
lines(datos$Tiempo, predict(modelo_lineal), col = "blue", lty = 2)
lines(datos$Tiempo, predict(modelo_cuadratico), col = "red", lty = 1)
lines(tiempo_futuro$Tiempo, pronostico_lineal, col = "blue", lty = 2, type = "o", pch = 16)
lines(tiempo_futuro$Tiempo, pronostico_cuadratico, col = "red", lty = 1, type = "o", pch = 16)
legend("topleft", legend = c("Datos originales", "Ajuste Lineal", "Ajuste Cuadrático",
```

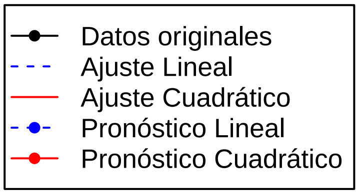
```
                                    "Pronóstico Lineal", "Pronóstico Cuadrático"),
       col = c("black", "blue", "red", "blue", "red"), lty = c(1, 2, 1, 2, 1), pch = c(16, NA, NA, 16,
```



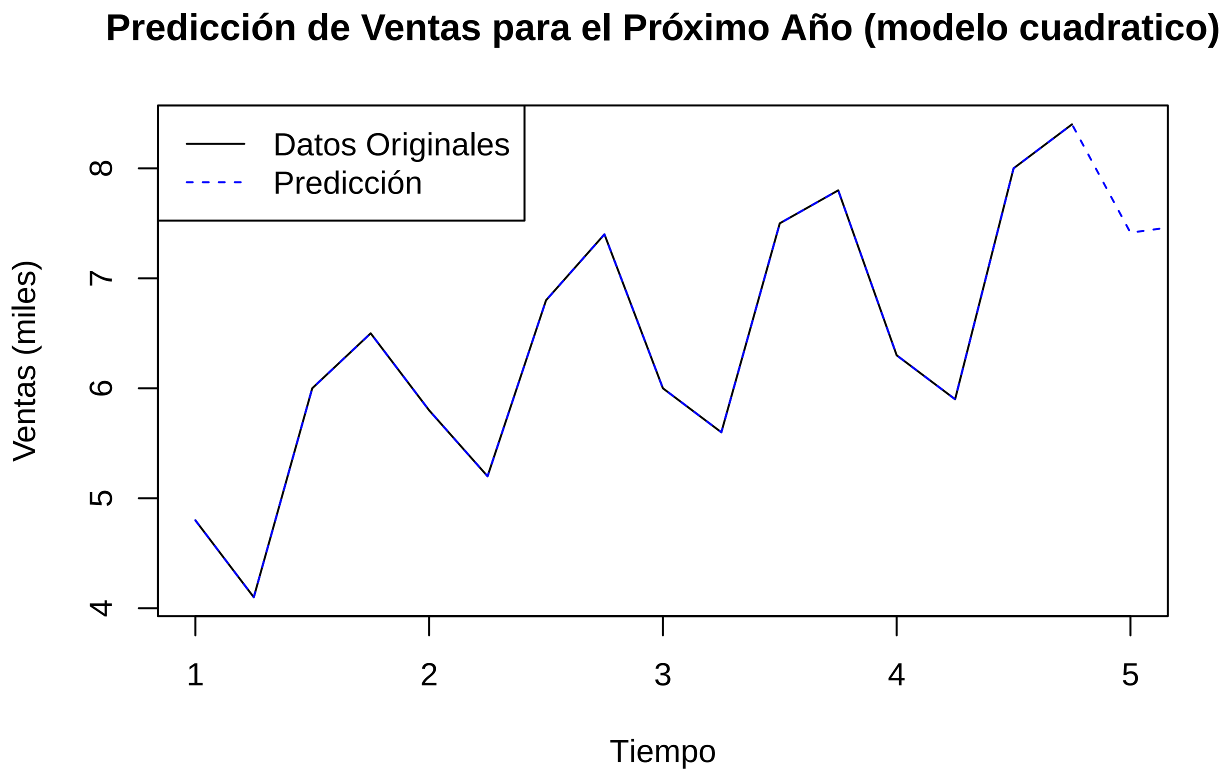

```
par(mar = c(5, 4, 4, 2) + 0.1)
plot(datos$Tiempo, datos$Ventas, type = "o", col = "black", pch = 16,
     main = "Pronóstico de Ventas para el Próximo Año",
     ylab = "Ventas desestacionalizadas (miles)", xlab = "Tiempo", xlim = c(1, length(datos$Tiempo) + 4
lines(datos$Tiempo, predict(modelo_lineal), col = "blue", lty = 2)
lines(datos$Tiempo, predict(modelo_cuadratico), col = "red", lty = 1)
lines(tiempo_futuro$Tiempo, pronostico_lineal, col = "blue", lty = 2, type = "o", pch = 16)
lines(tiempo_futuro$Tiempo, pronostico_cuadratico, col = "red", lty = 1, type = "o", pch = 16)
```



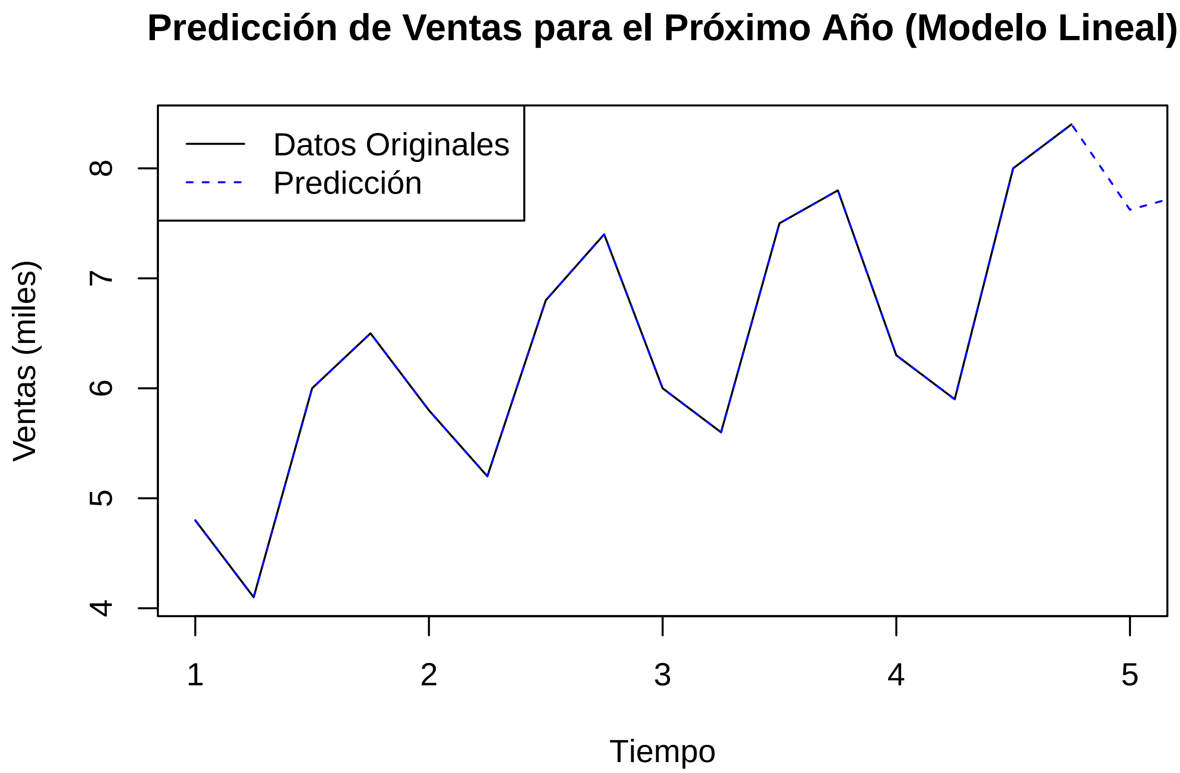

```
plot(0, type = "n", axes = FALSE, xlab = "", ylab = "")
legend("center", legend = c("Datos originales", "Ajuste Lineal", "Ajuste Cuadrático",
                            "Pronóstico Lineal", "Pronóstico Cuadrático"),
       col = c("black", "blue", "red", "blue", "red"), lty = c(1, 2, 1, 2, 1), pch = c(16, NA, NA, 16, 
```




```r
nuevos_tiempos <- (length(ventas) + 1):(length(ventas) + 4)
predicciones <- predict(modelo_cuadratico, newdata = data.frame(Tiempo = nuevos_tiempos))
ventas_ts <- ts(ventas, start = c(1, 1), frequency = 4)
plot(ventas_ts, xlim = c(1, 5), ylim = c(min(ventas_ts), max(c(ventas, predicciones))),
     main = "Predicción de Ventas para el Próximo Año (modelo cuadratico)", ylab = "Ventas (miles)", xla
lines(ts(c(ventas, predicciones), start = c(1, 1), frequency = 4), col = "blue", lty = 2)
legend("topleft", legend = c("Datos Originales", "Predicción"),
       col = c("black", "blue"), lty = c(1, 2))
```




```r
nuevos_tiempos <- (length(ventas) + 1):(length(ventas) + 4)
predicciones <- predict(modelo_lineal, newdata = data.frame(Tiempo = nuevos_tiempos))
ventas_ts <- ts(ventas, start = c(1, 1), frequency = 4)
plot(ventas_ts, xlim = c(1, 5), ylim = c(min(ventas_ts), max(c(ventas, predicciones))),
     main = "Predicción de Ventas para el Próximo Año (Modelo Lineal)", ylab = "Ventas (miles)", xlab =
lines(ts(c(ventas, predicciones), start = c(1, 1), frequency = 4), col = "blue", lty = 2)
legend("topleft", legend = c("Datos Originales", "Predicción"),
       col = c("black", "blue"), lty = c(1, 2))
```


Predicción de Ventas para el Próximo Año (Modelo Lineal)
Datos Originales
Predicción
Ventas (miles)
4
5
6
7
8
1
2
3
4
5
Tiempo