Actividad 1.11 Series de tiempo estacionarias

Franco Mendoza Muraira A01383399

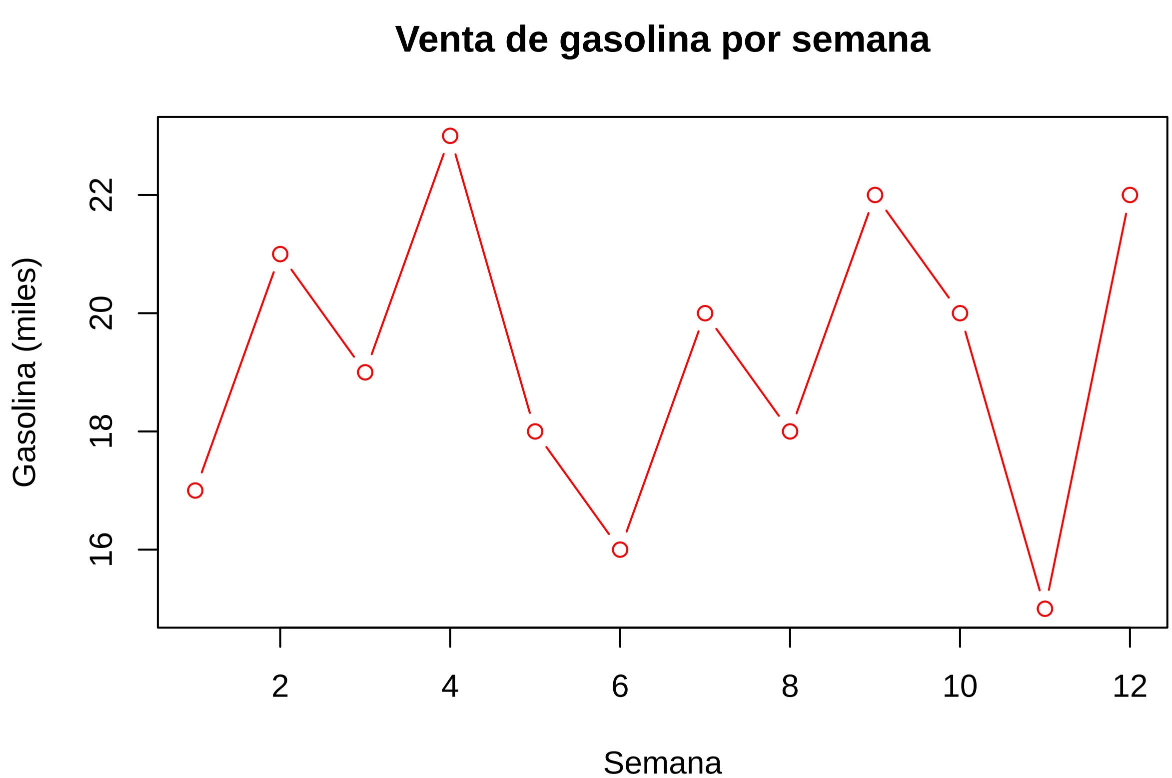
2023-11-26

Problema 1

Usa los datos de las ventas de gasolina en una estación de servicio para analizar modelos de pronósticos de la serie de tiempo:

```
semana = seq(1,12)
ventas_gasolina <- c(17, 21, 19, 23, 18, 16, 20, 18, 22, 20, 15, 22)

df1 = data.frame(semana,ventas_gasolina)
df2=data.frame(semana,ventas_gasolina)
df3 = data.frame(semana,ventas_gasolina)
plot(df1,main = "Venta de gasolina por semana",ylab = "Gasolina (miles)",xlab = "Semana",type='b',col="
```



Promedios móviles

```
p = 3
movil = c()

for (i in 1:length(semana)){
  if (i<=p){
    movil = c(movil,NA)
  }
  else {
    movil = c(movil,mean(ventas_gasolina[(i-p):(i-1)]))
  }

}
df1$Promedio_Movil = movil
```

Errores

```
n =12
error = df1$ventas_gasolina-df1$Promedio_Movil
error_cuadrado = error^2
error_porcentual = abs((error/ventas_gasolina)*100)

df1$Error = error
df1$Error_Cuadrado = error_cuadrado
df1$Error_Porcentual = error_porcentual

CME <- sum(error_cuadrado[(p+1):n])/(n-p)
EPAM <- sum(error_porcentual[(p+1):n])/(n-p)
cat("El promedio de los cuadrados de los errores es:",CME)
```

```
## El promedio de los cuadrados de los errores es: 10.22222
```

```
cat("\n\nEl promedio de los errores porcentuales es:",EPAM,"%")
```

```
##
##
## El promedio de los errores porcentuales es: 14.35661 %
```

```
df1
```

```
##   semana ventas_gasolina Promedio_Movil Error Error_Cuadrado Error_Porcentual
## 1      1              17             NA    NA             NA               NA
## 2      2              21             NA    NA             NA               NA
## 3      3              19             NA    NA             NA               NA
## 4      4              23             19     4             16         17.39130
## 5      5              18             21    -3              9         16.66667
## 6      6              16             20    -4             16         25.00000
## 7      7              20             19     1              1          5.00000
```

```
## 8        8                18                18      0              0           0.00000
## 9        9                22                18      4             16          18.18182
## 10      10                20                20      0              0           0.00000
## 11      11                15                20     -5             25          33.33333
## 12      12                22                19      3              9          13.63636
```

Promedios móviles ponderados

```
pesos <- c(1/6, 2/6, 3/6)
promedio_pond <- c()

for (i in 1:length(semana)){
  if (i<=p){
    promedio_pond = c(promedio_pond,NA)
  }
  else{
    promedio_pond = c(promedio_pond,sum(ventas_gasolina[(i-p):(i-1)]*pesos))
  }
}
df2$Promedio_Ponderado = promedio_pond
```

Errores

```
n =12
error = df2$ventas_gasolina-df2$Promedio_Ponderado
error_cuadrado = error^2
error_porcentual = abs((error/ventas_gasolina)*100)

df2$Error = error
df2$Error_Cuadrado = error_cuadrado
df2$Error_Porcentual = error_porcentual

CME <- sum(error_cuadrado[(p+1):n])/(n-p)
EPAM <- sum(error_porcentual[(p+1):n])/(n-p)
cat("El promedio de los cuadrados de los errores es:",CME)
```

```
## El promedio de los cuadrados de los errores es: 11.49074
```

```
cat("\n\nEl promedio de los errores porcentuales es:",EPAM,"%")
```

```
##
##
## El promedio de los errores porcentuales es: 15.99248 %
```

```
df2
```

```
##    semana ventas_gasolina Promedio_Ponderado      Error Error_Cuadrado
## 1       1              17                 NA         NA             NA
```

```
## 2       2              21                 NA        NA            NA
## 3       3              19                 NA        NA            NA
## 4       4              23           19.33333  3.6666667    13.4444444
## 5       5              18           21.33333 -3.3333333    11.1111111
## 6       6              16           19.83333 -3.8333333    14.6944444
## 7       7              20           17.83333  2.1666667     4.6944444
## 8       8              18           18.33333 -0.3333333     0.1111111
## 9       9              22           18.33333  3.6666667    13.4444444
## 10     10              20           20.33333 -0.3333333     0.1111111
## 11     11              15           20.33333 -5.3333333    28.4444444
## 12     12              22           17.83333  4.1666667    17.3611111
##    Error_Porcentual
## 1                NA
## 2                NA
## 3                NA
## 4         15.942029
## 5         18.518519
## 6         23.958333
## 7         10.833333
## 8          1.851852
## 9         16.666667
## 10         1.666667
## 11        35.555556
## 12        18.939394
```

Metodo de suavizamiento exponencial

Alfa Optimo

```
alpha_values <- seq(0.1, 0.9, by = 0.1)
cme_values <- numeric(length(alpha_values))

for (i in 1:length(alpha_values)) {
  alpha <- alpha_values[i]
  smooth_exp <- numeric(length = n)
  smooth_exp[1] <- ventas_gasolina[1]
  for (j in 2:n) {
    smooth_exp[j] <- alpha * ventas_gasolina[j - 1] + (1 - alpha) * smooth_exp[j - 1]
  }
  residuals <- ventas_gasolina - smooth_exp
  cme_values[i] <- mean(residuals^2)
}

best_alpha <- alpha_values[which.min(cme_values)]
best_alpha
```

```
## [1] 0.2
```

Suavizamiento con el mejor alfa

```
alpha <- 0.2

smooth_exp <- numeric(length = length(ventas_gasolina))
smooth_exp[1] <- ventas_gasolina[1]

pronosticadas <- c()
errores <- c()
error_cuadrado <- c()
error_porcentual <- c()

for (i in 2:length(ventas_gasolina)) {
  smooth_exp[i] <- alpha * ventas_gasolina[i - 1] + (1 - alpha) * smooth_exp[i - 1]
  pronosticadas[i] <- smooth_exp[i]
  errores[i] <- ventas_gasolina[i] - pronosticadas[i]
  error_cuadrado[i] <- errores[i]^2
  error_porcentual[i] <- abs(errores[i] / ventas_gasolina[i]) * 100
}



df3$Pronostico = pronosticadas
df3$Error = errores
df3$Error_Cuadrado = error_cuadrado
df3$Error_Porcentual = error_porcentual

CME <- mean(error_cuadrado[-1])
EPAM <- mean(error_porcentual[-1])
cat("El promedio de los cuadrados de los errores es:",CME)
```

```
## El promedio de los cuadrados de los errores es: 8.982231
```

```
cat("\n\nEl promedio de los errores porcentuales es:",EPAM,"%")
```

```
##
##
## El promedio de los errores porcentuales es: 13.40243 %
```

```
df3
```

```
##    semana ventas_gasolina Pronostico      Error Error_Cuadrado Error_Porcentual
## 1       1              17         NA         NA             NA               NA
## 2       2              21   17.00000  4.0000000     16.0000000        19.047619
## 3       3              19   17.80000  1.2000000      1.4400000         6.315789
## 4       4              23   18.04000  4.9600000     24.6016000        21.565217
## 5       5              18   19.03200 -1.0320000      1.0650240         5.733333
## 6       6              16   18.82560 -2.8256000      7.9840154        17.660000
## 7       7              20   18.26048  1.7395200      3.0259298         8.697600
## 8       8              18   18.60838 -0.6083840      0.3701311         3.379911
## 9       9              22   18.48671  3.5132928     12.3432263        15.969513
```

```
## 10     10               20   19.18937  0.8106342      0.6571279         4.053171
## 11     11               15   19.35149 -4.3514926     18.9354879        29.009951
## 12     12               22   18.48119  3.5188059     12.3819951        15.994572
```

Comparando los resultados, un menor valor de CME indica un mejor ajuste del modelo a los datos. En este caso, el suavizado exponencial con un valor de alfa de 0.2 tiene el CME más bajo (8.982231) en comparación con los promedios móviles y ponderados, lo que sugiere que este modelo tiene un mejor ajuste a los datos de venta de gasolina.

Además, el error porcentual promedio (EPAM) también es menor en el modelo de suavizado exponencial, lo que indica que este modelo tiene un menor error porcentual en la predicción de las ventas.

Por lo tanto, según los resultados, el modelo de suavizado exponencial con alfa de 0.2 parece ser el más efectivo para predecir la venta de gasolina en comparación con los promedios móviles y ponderados.

Prediccion para la semana 13

```
## Usando el metodo de suavizamiento, el valor predicho para miles de galones
##  de gasolina para la semana 13 es: 19.18496
```

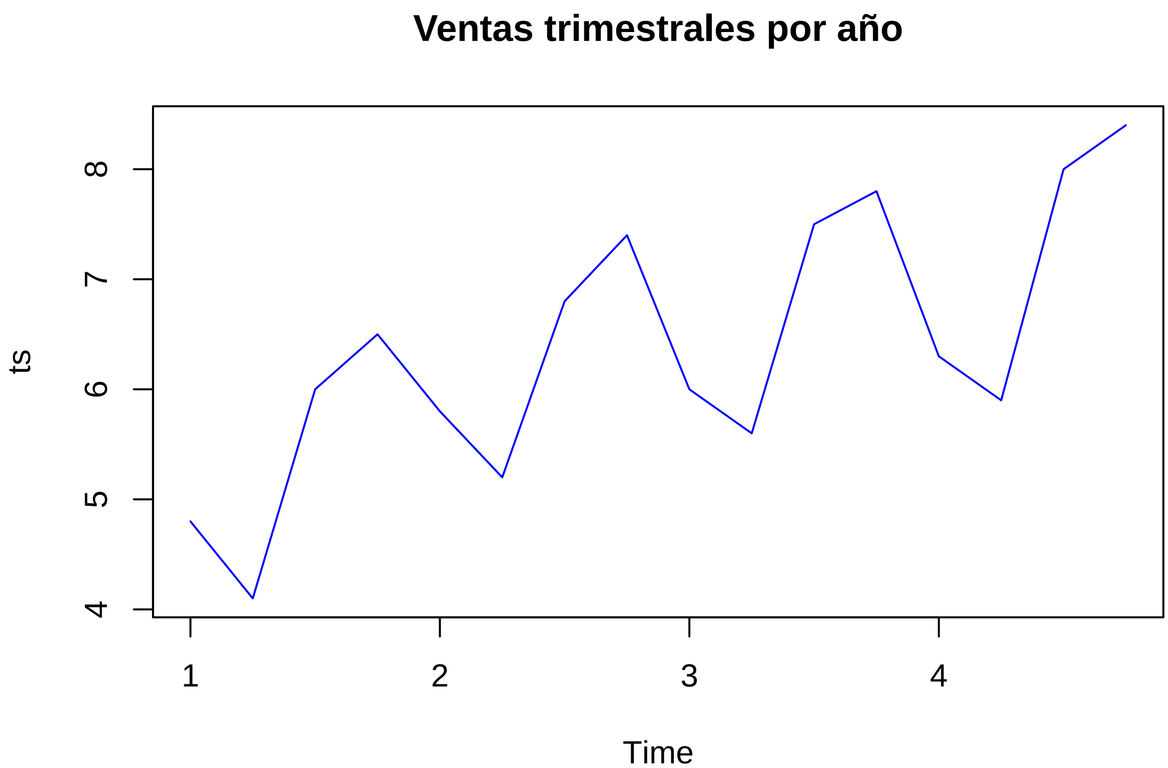
Problema 2

```r
anio = rep(seq(1,4),each=4)
tri = rep(seq(1,4),times=4)
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)
df= data.frame(
  Año = anio,
  Trimestre = tri,
  Ventas = ventas
)
df
```

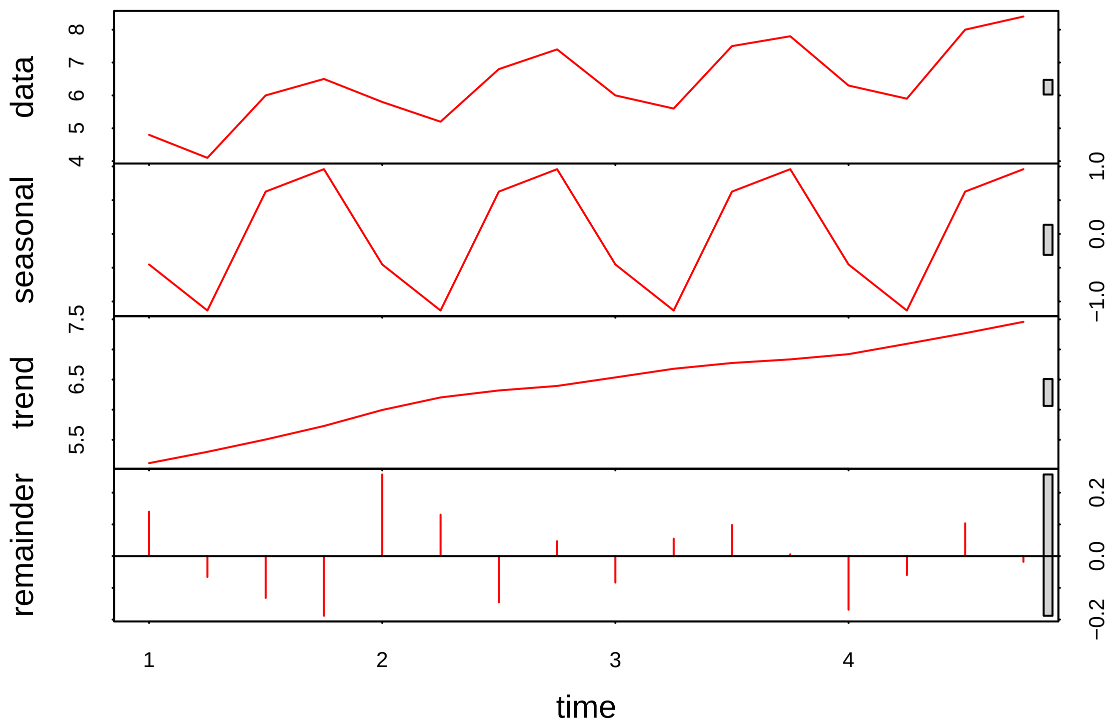
```
##    Año Trimestre Ventas
## 1    1         1    4.8
## 2    1         2    4.1
## 3    1         3    6.0
## 4    1         4    6.5
## 5    2         1    5.8
## 6    2         2    5.2
## 7    2         3    6.8
## 8    2         4    7.4
## 9    3         1    6.0
## 10   3         2    5.6
## 11   3         3    7.5
## 12   3         4    7.8
## 13   4         1    6.3
## 14   4         2    5.9
## 15   4         3    8.0
## 16   4         4    8.4
```

Analisis de tendencia y estacionalidad

```
ts = ts(ventas,frequency = 4)
decompose = stl(ts,s.window='periodic')
plot.ts(ts,col ="blue",main = "Ventas trimestrales por año")
```




```
plot(decompose,col="red")
```



Podemos ver la estacionalidad en el modelo, pero tambien como la tendencia es lineal.

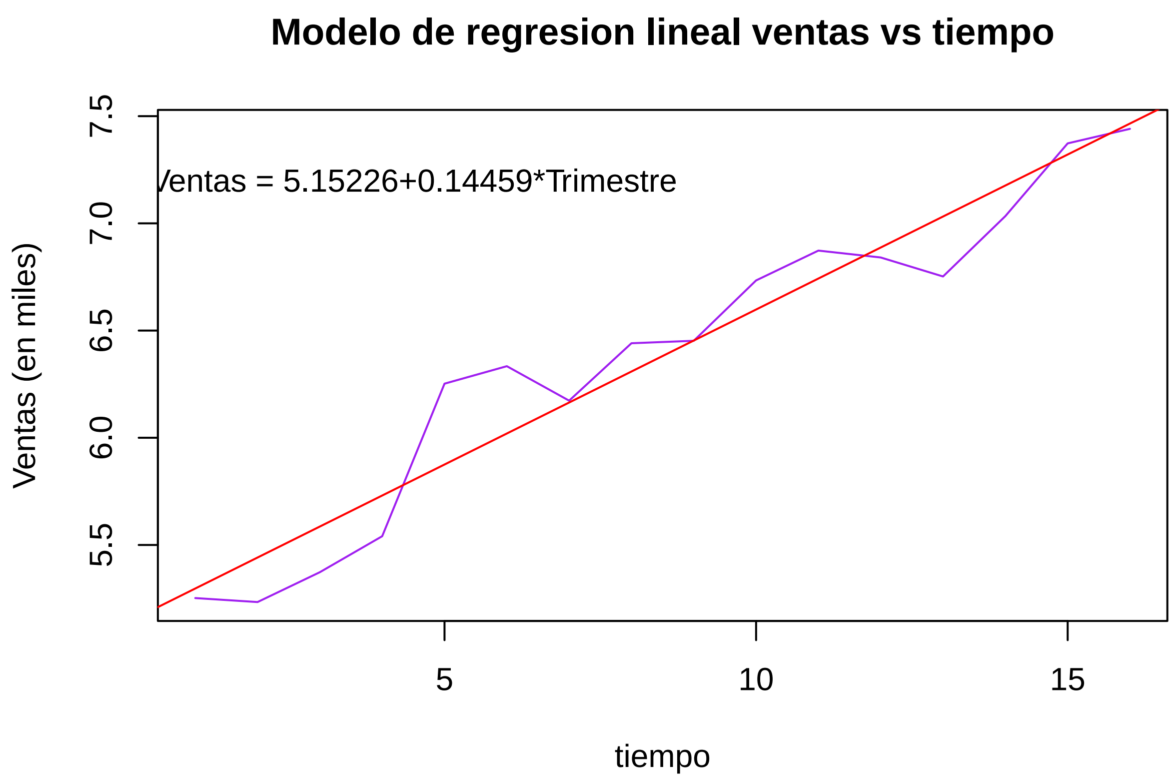
Modelo de regresión lineal con ventas desestacionalizadas

```
deestacional = decompose$time.series[,"trend"]+decompose$time.series[,"remainder"]
tiempo=1:length(deestacional)
y = deestacional
A = lm(y~tiempo)
summary(A)
```

```
##
## Call:
## lm(formula = y ~ tiempo)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.27957 -0.15444 -0.01302  0.13070  0.37712
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.15226    0.10147   50.78  < 2e-16 ***
## tiempo       0.14459    0.01049   13.78 1.56e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.1935 on 14 degrees of freedom
## Multiple R-squared:  0.9313, Adjusted R-squared:  0.9264
## F-statistic: 189.8 on 1 and 14 DF,  p-value: 1.557e-09
```

```
plot(tiempo,y,col="purple",main ="Modelo de regresion lineal ventas vs tiempo",type="l",ylab = "Ventas
abline(A,col ="red")
text(4.5,7.2,"Ventas = 5.15226+0.14459*Trimestre")
```



El modelo de regresión lineal revela una relación significativa entre el tiempo y las ventas, con un R-cuadrado del 93.13%, indicando que el 93.13% de la variabilidad en las ventas se explica por el tiempo. Cada unidad de aumento en el tiempo está asociada, en promedio, con un incremento de aproximadamente 0.14459 miles en las ventas. Ambos coeficientes son estadísticamente significativos, respaldando la validez del modelo para predecir las ventas en función del tiempo. Este alto nivel de ajuste y la significancia de las variables señalan la efectividad del modelo para explicar y predecir las fluctuaciones en las ventas a lo largo del tiempo.

Calculo de CME y EPAM de la predicción de la serie de tiempo

```
df$Predicciones <- predict(A,newdata = data.frame(tiempo=tiempo))

df$Errores <- df$Ventas - df$Predicciones

CME <- mean(df$Errores^2)
```

```r
EPAM <- mean(abs(df$Errores / df$Ventas)) * 100
cat("El promedio de los cuadrados de los errores es:",CME)
```

```
## El promedio de los cuadrados de los errores es: 0.6993267
```

```r
cat("\n\nEl promedio de los errores porcentuales es:",EPAM,"%")
```

```
##
##
## El promedio de los errores porcentuales es: 12.68761 %
```

```r
df
```

```
##    Año Trimestre Ventas Predicciones     Errores
## 1    1         1    4.8     5.296848 -0.49684834
## 2    1         2    4.1     5.441435 -1.34143511
## 3    1         3    6.0     5.586022  0.41397813
## 4    1         4    6.5     5.730609  0.76939136
## 5    2         1    5.8     5.875195 -0.07519541
## 6    2         2    5.2     6.019782 -0.81978217
## 7    2         3    6.8     6.164369  0.63563106
## 8    2         4    7.4     6.308956  1.09104430
## 9    3         1    6.0     6.453542 -0.45354247
## 10   3         2    5.6     6.598129 -0.99812923
## 11   3         3    7.5     6.742716  0.75728400
## 12   3         4    7.8     6.887303  0.91269724
## 13   4         1    6.3     7.031890 -0.73188953
## 14   4         2    5.9     7.176476 -1.27647629
## 15   4         3    8.0     7.321063  0.67893694
## 16   4         4    8.4     7.465650  0.93435018
```

El modelo exhibe un CME (Error Cuadrático Medio) de 0.6993267 y un EPAM (Error Porcentual Absoluto Medio) del 12.68761%. Estos valores sugieren que el modelo tiene un rendimiento razonable al predecir las ventas. El CME indica que, en promedio, los errores de predicción tienen una magnitud de 0.6993267 miles, mientras que el EPAM muestra que los errores promedio representan aproximadamente el 12.68761% de las ventas reales. Aunque estos errores son relativamente bajos, es crucial considerar el contexto específico y la relevancia del nivel de precisión necesario para el análisis de ventas, así como la comparación con otros modelos alternativos o el ajuste potencial del modelo actual para mejorar su rendimiento predictivo.

Pronóstico para el siguiente año

```
## [1] "Predicciones de ventas para el año 5, sus cuatro trimestres:"
```

```
## [1] 7610.29
```

```
## [1] 7754.88
```

```
## [1] 7899.47
```

```
## [1] 8044.06
```