Series de tiempo no estacionarias

Franco Quintanilla

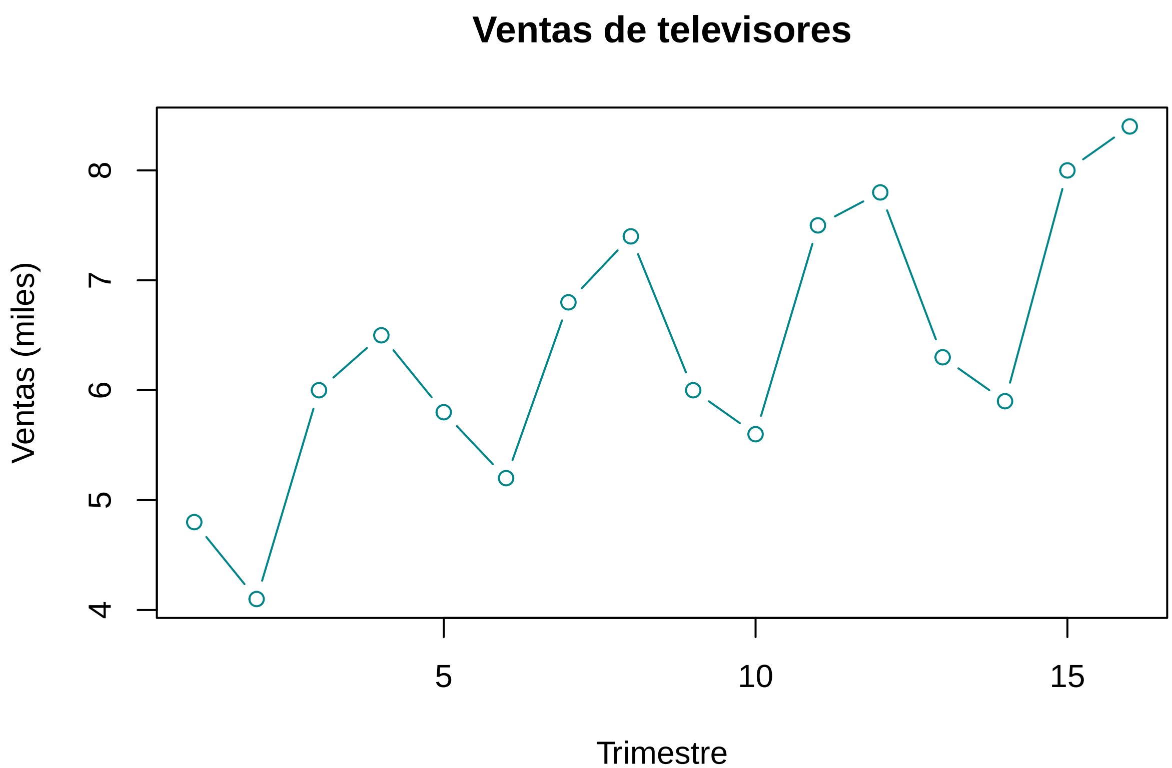
2022-11-14

1. Cargamos los datos

```
t = seq(1, 16, 1)
v = 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)
```

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

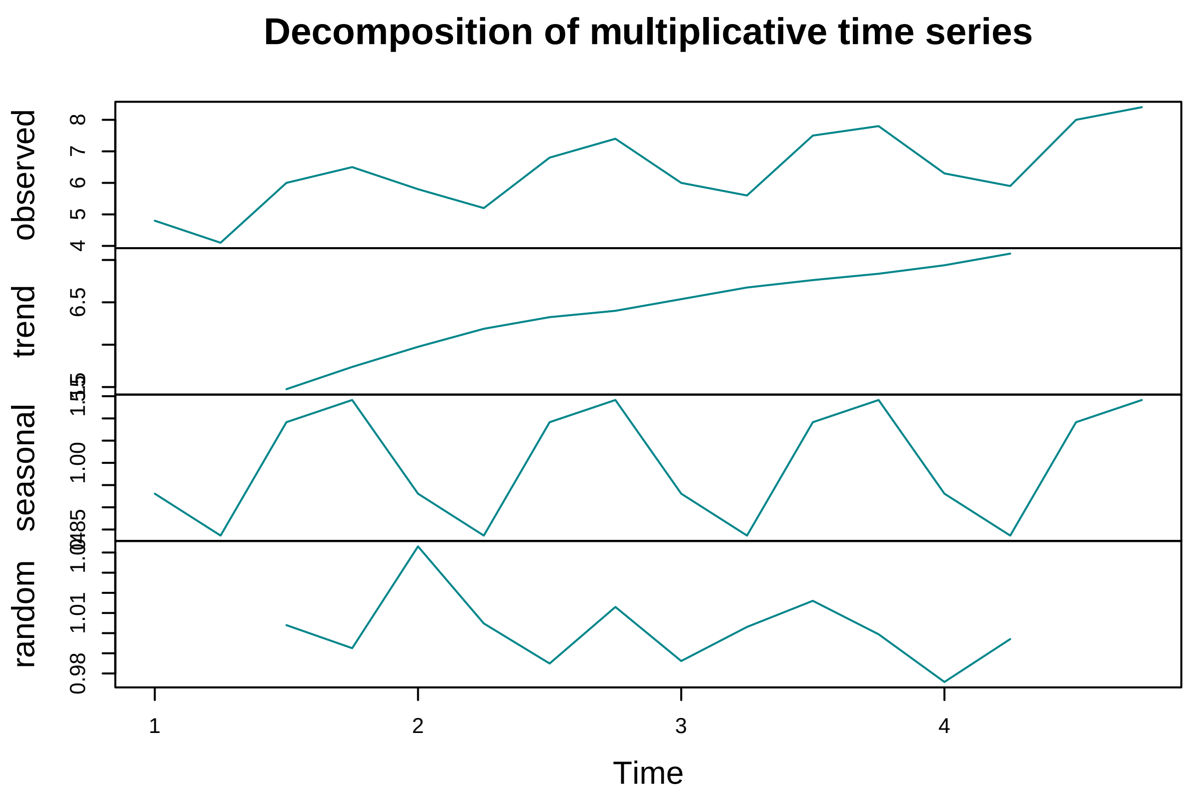
```
plot(t, v, main="Ventas de televisores", xlab="Trimestre", ylab="Ventas (miles)",
     type="b", col="#00868B")
```



3. Realiza el análisis de tendencia y estacionalidad

- Descompón la serie en sus 3 componentes e interprétalos

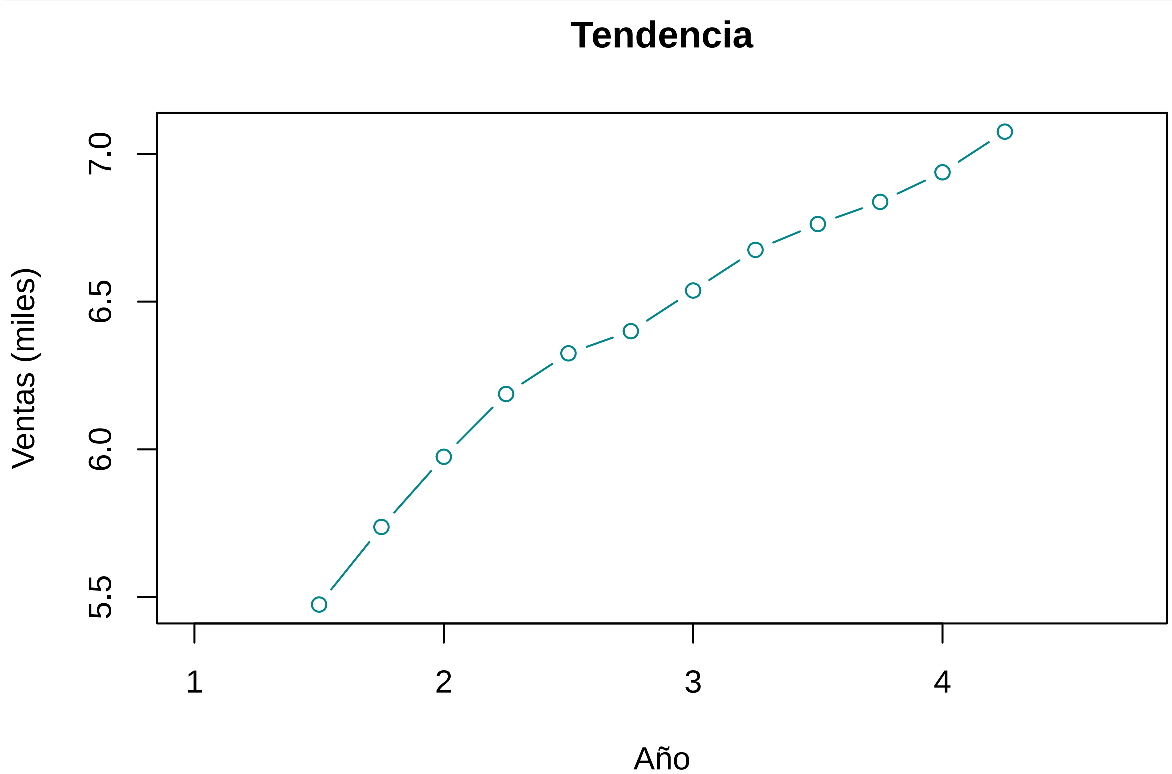
```
ts = ts(v, frequency=4)
decom = decompose(ts, type="multiplicative")
plot(decom, col="#00868B")
```



4. Analiza el modelo lineal de la tendencia:

- Realiza el gráfico de dispersión de la tendencia

```
plot(decom$trend, main="Tendencia", xlab="Año", ylab="Ventas (miles)",
     type="b", col="#00868B")
```

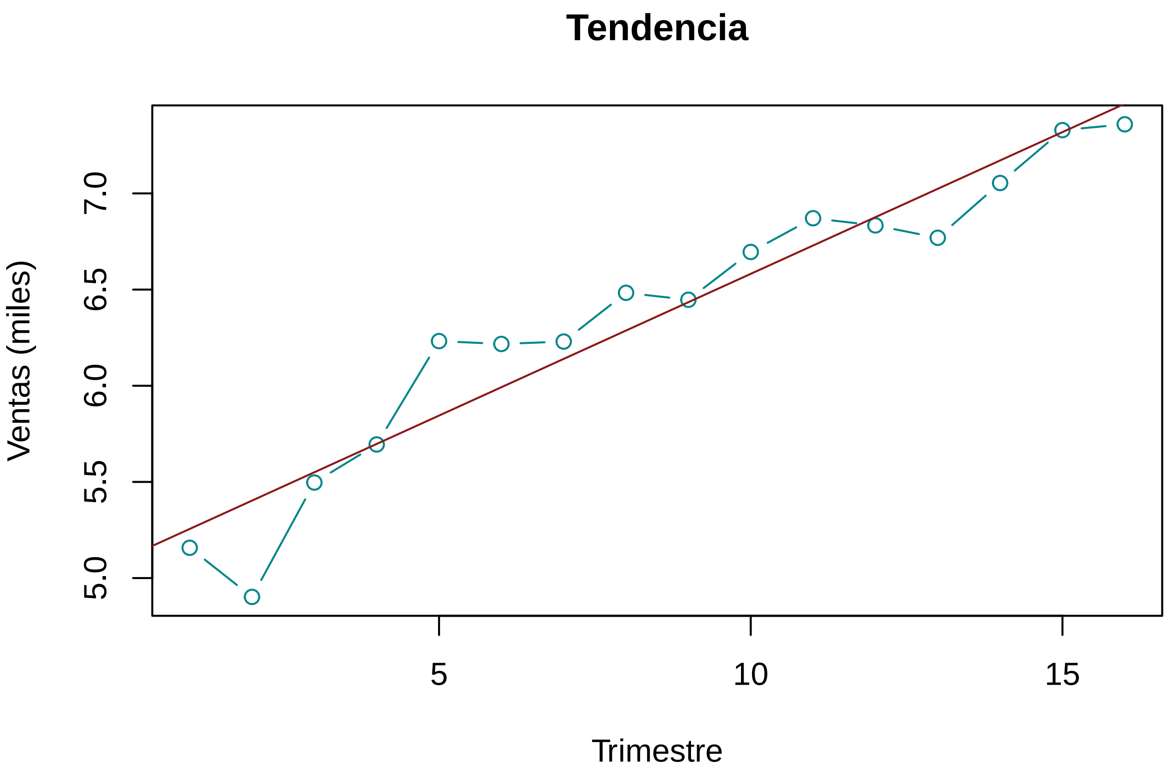


- Dibuja la recta junto con las ventas desestacionalizadas.

```r
# Definimos parametros
y = decom$x/decom$seasonal
x = seq(1, 16, 1)

# Hacemos el modelo lineal
linmod = lm(y ~ x)
```

```r
plot(x, y, main="Tendencia", xlab="Trimestre", ylab="Ventas (miles)",
     type="b", col="#00868B")
abline(linmod, col="#8B1A1A")
```



- Analiza la pertinencia del modelo lineal:

```r
sumlm <- summary(linmod)
sumlm
```

```
##
## Call:
## lm(formula = y ~ x)
##
## 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 ***
## x            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
```

Como observamos, nuestro modelo es muy bueno, ya que explica el 92% de la información y de la variabilidad del probelma, ya que cuenta con un valor de R^2 de 0.9208.

- Significancia de β_1

```
sumlm$coefficients[2,4]
```

```
## [1] 4.247717e-09
```

Como observamos que $p_{value} < 0.05$ rechazamos la hipotesis nula, la cual nos dice que $\beta_1 = 0$, por lo que podemos concluir que la pendiente es significativa.

- Variabilidad explicada por el modelo

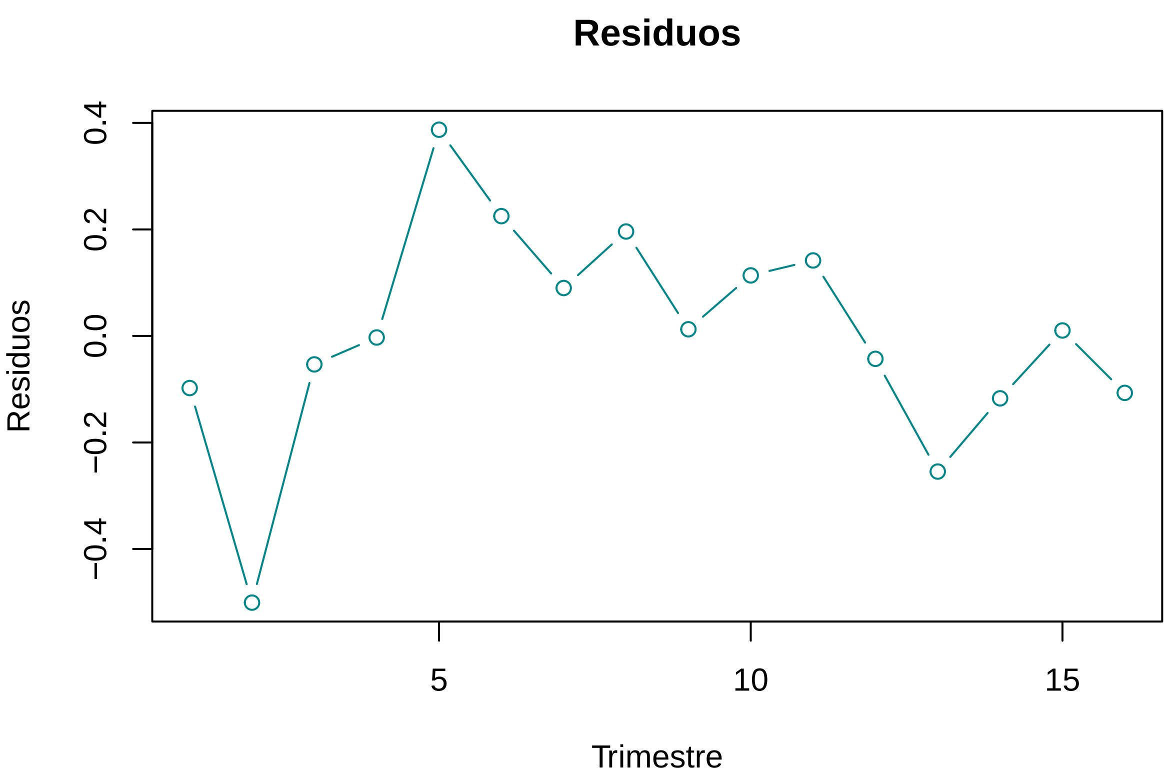
```
sumlm$r.squared
```

```
## [1] 0.9207911
```

Como mencionamos anteriormente, el modelo explica el 92% de la información y de la variabilidad de los datos.

- Análisis de los residuos

```
plot(linmod$residuals, main="Residuos", xlab="Trimestre", ylab="Residuos",
     type="b", col="#00868B")
```



Podemos ver que los residuos parecen comportarse de manera aleatoria, ya que no hay una tendencia en el comportamiento de los mismos.

- Prueba de normalidad (Shapiro)

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

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

Ahora, como el $p_{value} > 0.05$ no podemos rechazar la hipotesis nula de que los residuos siguen una distribucion normal.

5. Calcula el CME y el EPAM (promedio de los errores porcentuales) de la predicción de la serie de tiempo.

- CME

```
fun = function(x){
  linmod$coefficients[1] + linmod$coefficients[2]*x
  }
```

```
pv = fun(t)
cme = mean((pv - v)^2, na.rm=TRUE)
cme
```

```
## [1] 0.6957218
```

- EPAM

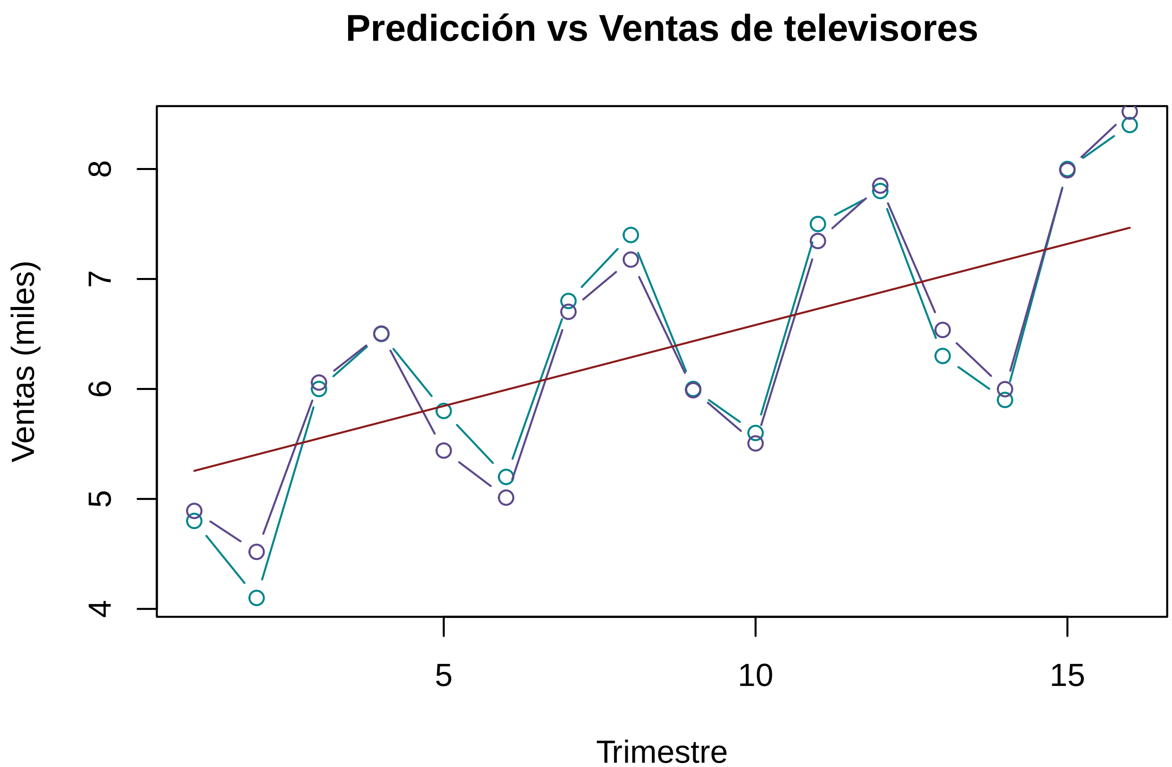
```
epam = mean(abs((pv - v)/v), na.rm=TRUE)
epam
```

```
## [1] 0.125894
```

6. Dibuja el gráfico de los valores de las ventas y las predicciones vs el tiempo

```
y2 = predict(linmod)*decom$seasonal
```

```
plot(t, v, xlab="Trimestre", ylab="Ventas (miles)", main="Predicción vs Ventas de televisores",
     type="b", col="#00868B")
lines(t, y2, type="b", col="#5D478B")
lines(t, predict(linmod), col="#8B1A1A")
```



7. Concluye sobre el modelo

El modelo lineal se ajusta muy bien a los datos, ya que:

- Se explica un 92.08% de la variabilidad de los datos
- $p_{value} < 0.05$ por lo que rechazamos que $\beta_1 = 0$
- Los residuos se comportan de manera aleatoria, no tendencia
- Los errores porcentuales son bajos, por lo que el modelo es preciso

8. Realiza el pronóstico para el siguiente año.

- Pronóstico para el siguiente año

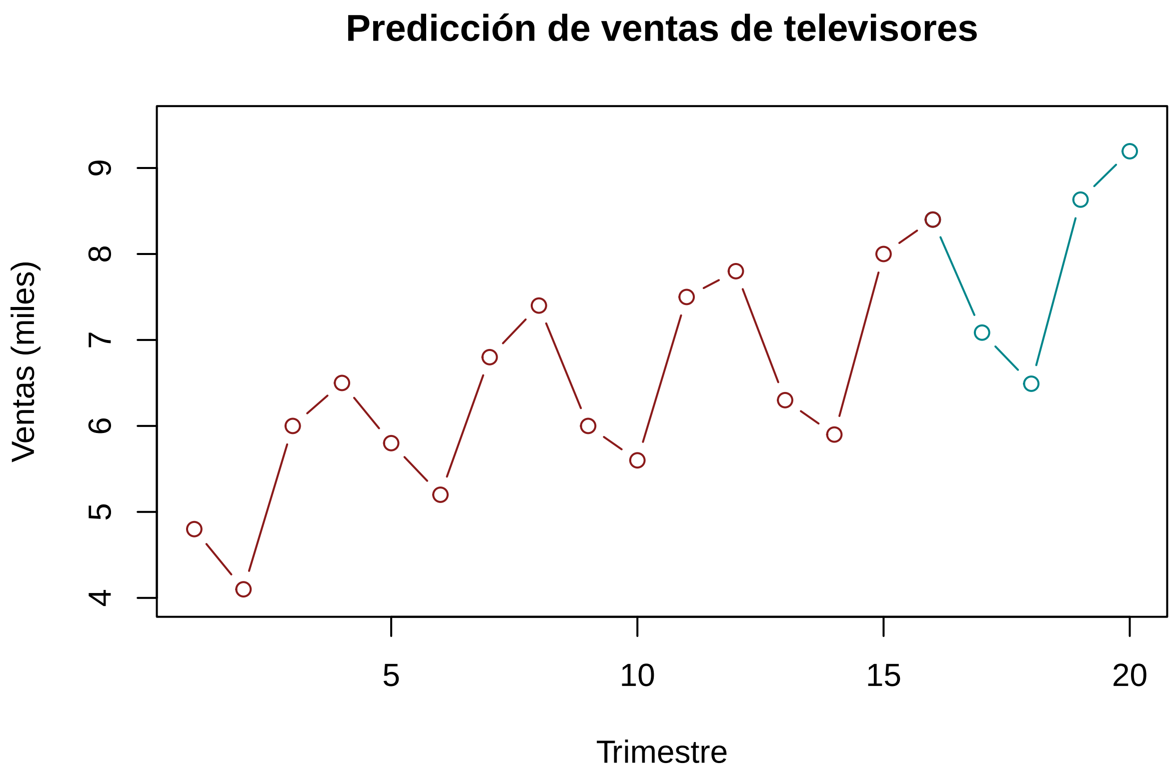
```
fun2 = function(x){
  linmod$coefficients[1] + linmod$coefficients[2] * x
  }
```

```
x2 = seq(17, 20, 1)
y2 = fun2(x2)*decom$figure
y2
```

```
## [1] 7.085626 6.491048 8.632258 9.194899
```

Gráfico de las predicciones

```
x3 = seq(16, 20, 1)
p2 = c(v[16], y2)
plot(x3, p2, xlab="Trimestre", ylab="Ventas (miles)", main="Predicción de ventas de televisores",
     type="b", col="#00868B", xlim = c(1,20), ylim = c(4, 9.5))
lines(t, v, type="b", col="#8B1A1A")
```



Un Problemilla mas

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.

- Cargamos los datos

```
vl = c(1690, 940, 2625, 2500, 1800, 900, 2900, 2360, 1850, 1100, 2930, 2615)
tl =seq(1, 12, 1)
```

- Hacemos la serie de tiempo

```
ts_l = ts(vl, frequency=4)
decom_l = decompose(ts_l, type="multiplicative")
```

- Promedios móviles de cuatro trimestres

```
pm_trim = filter(vl, rep(1/4, 4), sides=2)
pm_trim
```

```
## Time Series:
## Start = 1
## End = 12
## Frequency = 1
##  [1]      NA 1938.75 1966.25 1956.25 2025.00 1990.00 2002.50 2052.50 2060.00
## [10] 2123.75      NA      NA
```

- Promedios móviles centrados

```
pmc = filter(vl, rep(1/3, 3), sides=2)
pmc
```

```
## Time Series:
## Start = 1
## End = 12
## Frequency = 1
##  [1]       NA 1751.667 2021.667 2308.333 1733.333 1866.667 2053.333 2370.000
##  [9] 1770.000 1960.000 2215.000       NA
```

- Calcule los índices estacionales de los cuatro trimestres

```
iest = decom_l$figure
iest
```

```
## [1] 0.9003979 0.4862339 1.3961436 1.2172246
```

- ¿Cuándo obtiene la editorial el mayor índice estacional? ¿Parece razonable este resultado? ¿Por qué?

El mayor índice estacional se obtiene en el tercer trimestre, ya que es el trimestre con mayor venta de libros. Este resultado si es es razonable, ya que al final del tercer trimestre es cuando se empieza el ciclo escolar en Mexico.