# Tendencia

Federico Medina

2023-11-14

## Problema 1

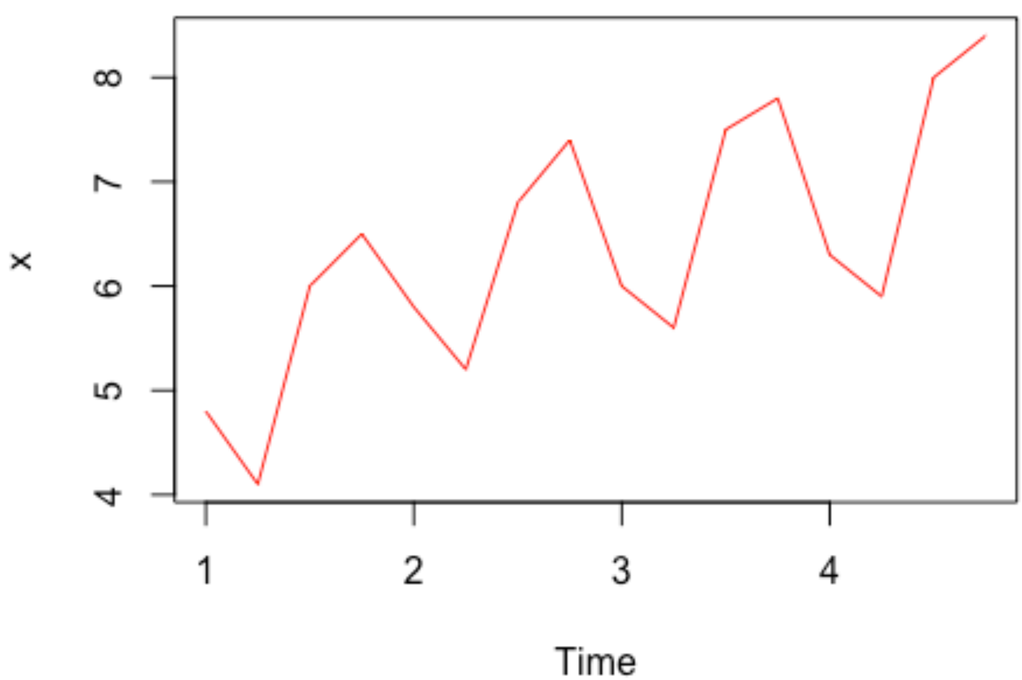
### Usa los datos de las ventas de televisores para familiarizarte con el análisis de tendencia de una serie de tiempo:

```
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)
```

### Serie de Tiempo

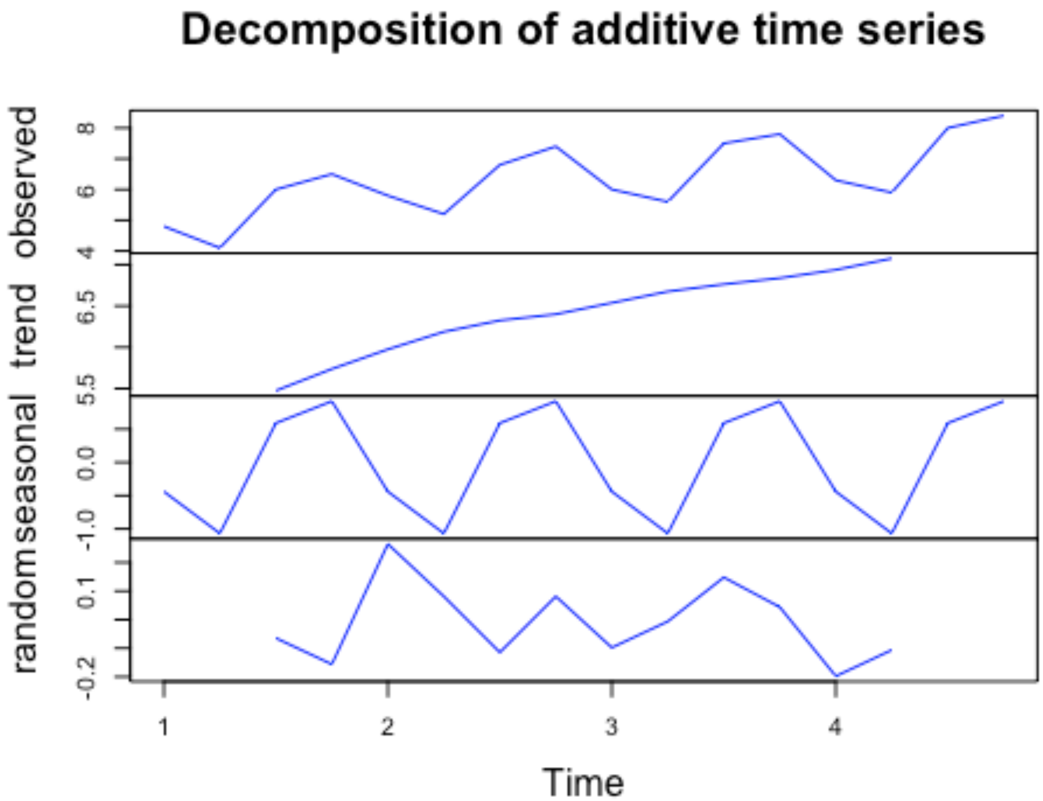
#### Realiza el gráfico de dispersión. Observa la tendencia y los ciclos.

```
x= ts(ventas, frequency = 4, start(c(2016,1)))
plot.ts(x, col = "red")
```



### Realiza el análisis de tendencia y estacionalidad. Descompón la serie en sus 3 componentes e interprétalos

```
T = decompose(x)
plot(T, col ="blue")
```




En la descomposición anterior observamos primeramente lo que son los datos por año. Esto tanto en la primer gráfica como en la segunda en la que se llama 'observed'. Vemos claramente que hay una estacionalidad en cada uno de estos años, lo cual se confirma en la 3er linea de la segunda gráfica impresa, la que se llama 'seasonal'. Después, vemos que la tendencia va incrementando positivamente, lo cual se muestra en la 2da linea de esta gráfica. Finalmente, observamos que la 'random' muestra los residuos de los datos.

## Analiza el modelo lineal de la tendencia:

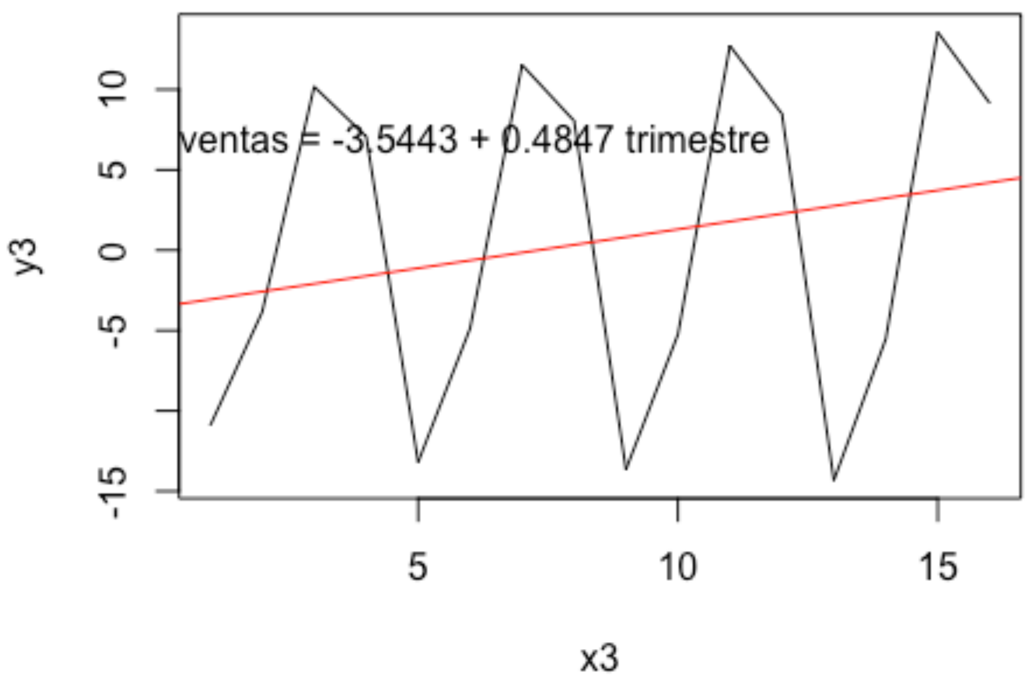
### Realiza la regresión lineal de la tendencia (ventas desestacionalizadas vs tiempo)

```
ventas_desestacionalizadas = (T$x)/(T$seasonal)
x3 = 1:16
y3 = ventas_desestacionalizadas
N3 = lm(y3~x3)
N3

##
## Call:
```

```
## lm(formula = y3 ~ x3)
##
## Coefficients:
## (Intercept)           x3
##     -3.5443       0.4847
```

**Dibuja la recta junto con las ventas desestacionalizadas.**

```r
plot(x3, y3, type = "l")
abline(N3, col = "red")
text(6, 7, " ventas = -3.5443 + 0.4847 trimestre")
```




**Analiza la pertinencia del modelo lineal:**

```r
summary(N3)

##
## Call:
## lm(formula = y3 ~ x3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -17.088  -8.085   1.836   8.971  12.267
##
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -3.5443     5.5166  -0.642    0.531
## x3            0.4847     0.5705   0.850    0.410
##
## Residual standard error: 10.52 on 14 degrees of freedom
## Multiple R-squared:  0.04902,    Adjusted R-squared:  -0.0189
## F-statistic: 0.7217 on 1 and 14 DF,  p-value: 0.4099
```

Podemos observar que el modelo anterior es un modelo que no tiene variables significativas, por lo que no es un buen modelo para explicar la serie de tiempo. Con esto, sabemos que $\beta_1$ no es significante. La variabilidad explicada por el modelo es de 0.04902, siendo muy bajo. Para los residuos observamos que el error es de 10.52 en 14 grados de libertad, lo cual indica que el modelo explica una gran parte de la variabilidad, lo cual es algo bueno. Para sacar la normalidad hay que hacer una prueba de shapiro para poder ver si es normal o no:

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

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(N3)
## W = 0.90397, p-value = 0.09308
```

Viendo que la prueba de shapiro dio un valor de 0.90 y el valor p es de 0.09, vemos que los residuos no son normales, por lo tanto se rechaza Ho.

### Cálculo de CME y MAPE

```
CME2=mean(residuals(N3)^2,na.rm='TRUE')
cat('CME2 = ', CME2)
```

```
## CME2 =  96.83152
```

```
cat(' \n')
```

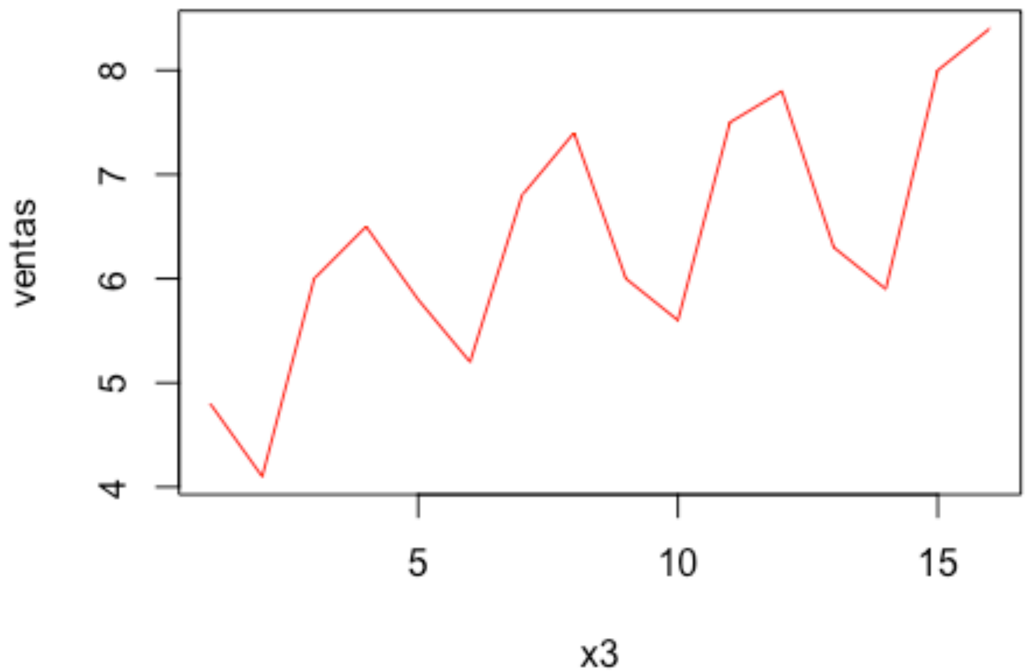
```
porcentual <- abs(residuals(N3))*100
MAPE <- mean(porcentual, na.rm = TRUE)
cat('MAPE = ', MAPE)
```

```
## MAPE =  903.5255
```

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

```
e = NA
g = NA
f = function(x) -3.5443 + 0.4847*x

for(i in  1:16){
  g[i] = f(i)*T$seasonal[i]
  e[i] = ventas[i] - g[i]
}
```

```
plot(x3, ventas, type='l', col='red')
lines(x3, g, col='blue', lty = 2)
```

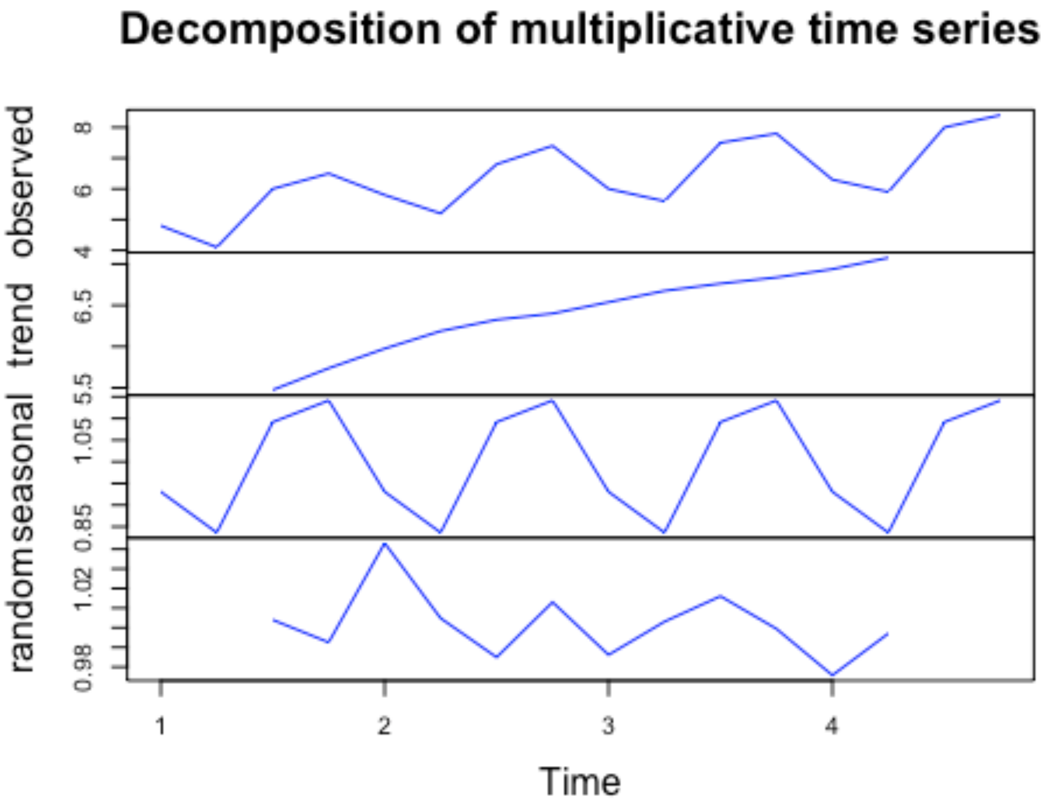



**Concluye sobre el modelo: de acuerdo al análisis de verificación de los supuestos, ¿es el mejor modelo que puedes obtener?**

Definitivamente el modelo no es el mejor que se puede obtener. Vemos que en dicho modelo ninguna variable es significativa, lo cual nos hace entender que el modelo no es 'afectado' por ninguna variable. Con esto y con la significanncia del modelo mostrado en el summary() podemos observar que el modelo no sirve para nada.

**Propón un posible mejor modelo para la tendencia de los datos.**

```
T = decompose(x, type='m')
plot(T, col ="blue")
```

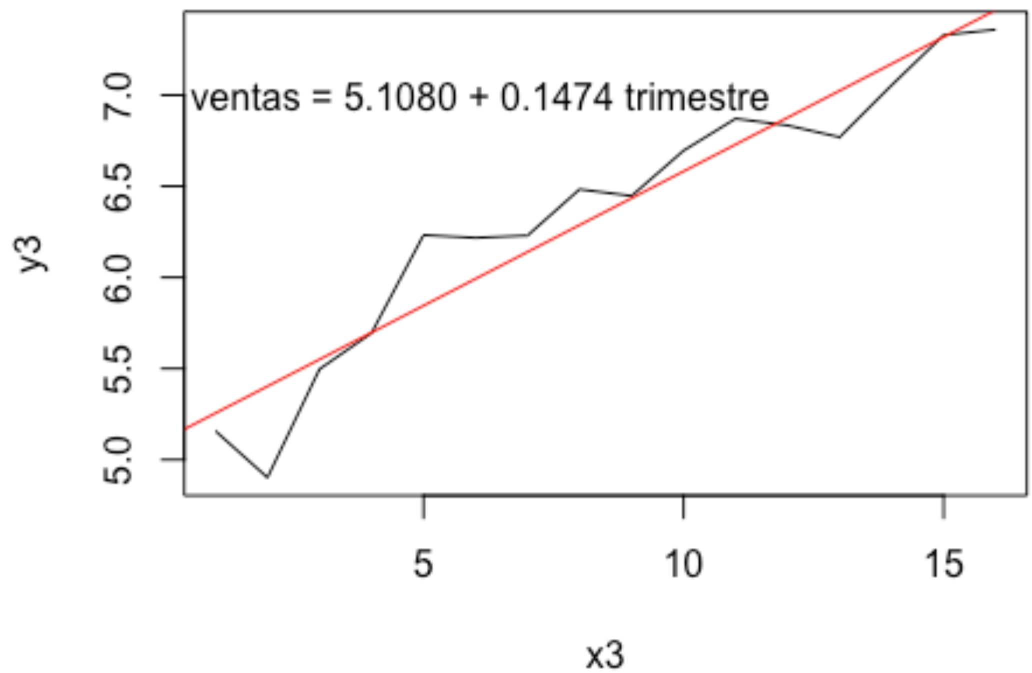



### Realiza la regresión lineal de la tendencia (ventas desestacionalizadas vs tiempo)

```
ventas_desestacionalizadas = (T$x)/(T$seasonal)
x3 = 1:16
y3 = ventas_desestacionalizadas
N3 = lm(y3~x3)
N3

##
## Call:
## lm(formula = y3 ~ x3)
##
## Coefficients:
## (Intercept)           x3
##      5.1080       0.1474
```

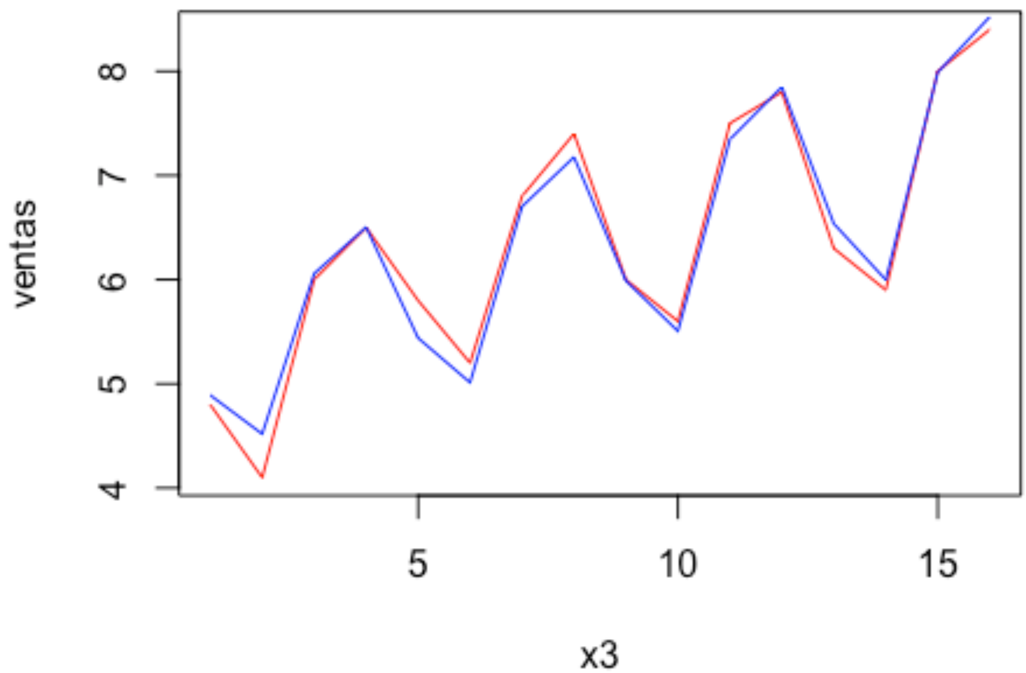
### Dibuja la recta junto con las ventas desestacionalizadas.

```
plot(x3, y3, type = "l")
abline(N3, col = "red")
text(6, 7, " ventas = 5.1080 + 0.1474 trimestre")
```




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

```
e = NA
g = NA
f = function(x) 5.1080 + 0.1474*x

for(i in  1:16){
  g[i] = f(i)*T$seasonal[i]
  e[i] = ventas[i] - g[i]
}

CME_tendencia = mean(residuals(N3)^2, na.rm = TRUE)
cat('CME del modelo: ', CME_tendencia)

## CME del modelo:  0.0397064

plot(x3, ventas, type='l', col='red')
lines(x3, g, type='l', col='blue')
```




### Summary

```
summary(N3)

##
## Call:
## lm(formula = y3 ~ x3)
##
## 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 ***
## x3           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
```

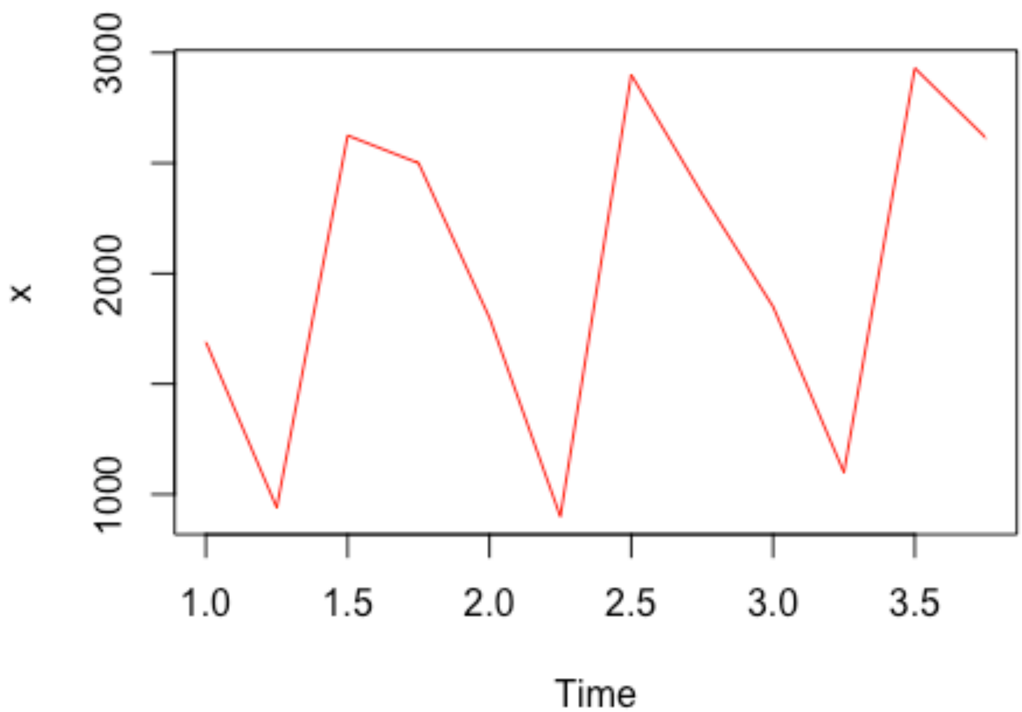
Observando los datos anteriores, vemos que ahora si ambas variables son significantes al igual que el modelo, lo cual evidentemente significa que es mejor que el otro modelo.

### Realiza el pronóstico para el siguiente año

```
f = function(x) {5.1080 + 0.1474*x}

a1 = T$seasonal[1]
a2 = T$seasonal[2]
a3 = T$seasonal[3]
a4 = T$seasonal[4]


cat('el pronostico siguiente es: ', f(17)*a1*1000, '\n')

## el pronostico siguiente es:  7085.872

cat('el pronostico siguiente es: ', f(18)*a2*1000, '\n')

## el pronostico siguiente es:  6491.284

cat('el pronostico siguiente es: ', f(19)*a3*1000, '\n')

## el pronostico siguiente es:  8632.585

cat('el pronostico siguiente es: ', f(20)*a4*1000, '\n')

## el pronostico siguiente es:  9195.263
```

## Problema 2

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

x= ts(ventas, frequency = 4, start(c(2016,1)))
plot.ts(x, col = "red")
```




### Promedios Moviles

```
movil = NA
for(i in 4:12){
  movil[i-1] <- c((ventas[i-3] + ventas[i-2] + ventas[i-1] + ventas[i])/4)
}
movil[12] = NA
movil
```

```
##  [1]      NA      NA 1938.75 1966.25 1956.25 2025.00 1990.00 2002.50
2052.50
## [10] 2060.00 2123.75      NA
```

### Promedios Moviles Centralizados

```
centralizado = NA
for(i in 4:12){
  centralizado[i-1] <- c((movil[i] + movil[i+1])/2)
}
centralizado[12] = NA
```

### Valor estacional Irregular

```
estacional = NA
estacional <- ventas/centralizado
```

```r
df <- data.frame('Ventas' = ventas, 'Promedio Movil' = movil, 'Promedio Movil
Centralizado' = centralizado, 'Valor Estacional Irregular' = estacional)
df
```

```
##    Ventas Promedio.Movil Promedio.Movil.Centralizado
Valor.Estacional.Irregular
## 1    1690             NA                          NA
NA
## 2     940             NA                          NA
NA
## 3    2625        1938.75                    1961.250
1.3384321
## 4    2500        1966.25                    1990.625
1.2558870
## 5    1800        1956.25                    2007.500
0.8966376
## 6     900        2025.00                    1996.250
0.4508453
## 7    2900        1990.00                    2027.500
1.4303329
## 8    2360        2002.50                    2056.250
1.1477204
## 9    1850        2052.50                    2091.875
0.8843741
## 10   1100        2060.00                          NA
NA
## 11   2930        2123.75                          NA
NA
## 12   2615             NA                          NA
NA
```

### Indices Estacionales

```r
indice = NA
estacional1 <- c(0, 0, 1.3384321, 1.2558870, 0.8966376, 0.4508453, 1.4303329,
1.1477204, 0.8843741, 0, 0, 0)

indice[1] <- c((estacional1[1]+estacional1[5]+estacional1[9]/3))
indice[2] <- c((estacional1[2]+estacional1[6]+estacional1[10]/3))
indice[3] <- c((estacional1[3]+estacional1[7]+estacional1[11]/3))
indice[4] <- c((estacional1[4]+estacional1[8]+estacional1[12]/3))

df <- data.frame('Ventas' = ventas, 'Promedio Movil' = movil, 'Promedio Movil
Centralizado' = centralizado, 'Valor Estacional Irregular' = estacional,
'Indice Estacional' = indice)
df
```

```
##    Ventas Promedio.Movil Promedio.Movil.Centralizado
Valor.Estacional.Irregular
## 1    1690             NA                          NA
NA
```

```
## 2     940            NA                            NA
NA
## 3    2625       1938.75                      1961.250
1.3384321
## 4    2500       1966.25                      1990.625
1.2558870
## 5    1800       1956.25                      2007.500
0.8966376
## 6     900       2025.00                      1996.250
0.4508453
## 7    2900       1990.00                      2027.500
1.4303329
## 8    2360       2002.50                      2056.250
1.1477204
## 9    1850       2052.50                      2091.875
0.8843741
## 10   1100       2060.00                            NA
NA
## 11   2930       2123.75                            NA
NA
## 12   2615            NA                            NA
NA
##    Indice.Estacional
## 1          1.1914290
## 2          0.4508453
## 3          2.7687650
## 4          2.4036074
## 5          1.1914290
## 6          0.4508453
## 7          2.7687650
## 8          2.4036074
## 9          1.1914290
## 10         0.4508453
## 11         2.7687650
## 12         2.4036074
```

La editorial recibe el mayor indice estacional en el 3er trimestre. Esto tiene sentido en base a los datos ya que en el 3er trimestre recibe la mayor cantidad de ventas, lo cual significa que el crecimiento mas alto sera en ese trimestre. Igualmente, sabiendo que la editorial vende libros de universidad, en el tercer trimestre es cuando empiezan las clases de agosto, lo cual tiene mucho sentido que en ese trimestre haya una mayor cantidad de ventas.