A8_Series de tiempo no estacionarias

Héctor San Román Caraza

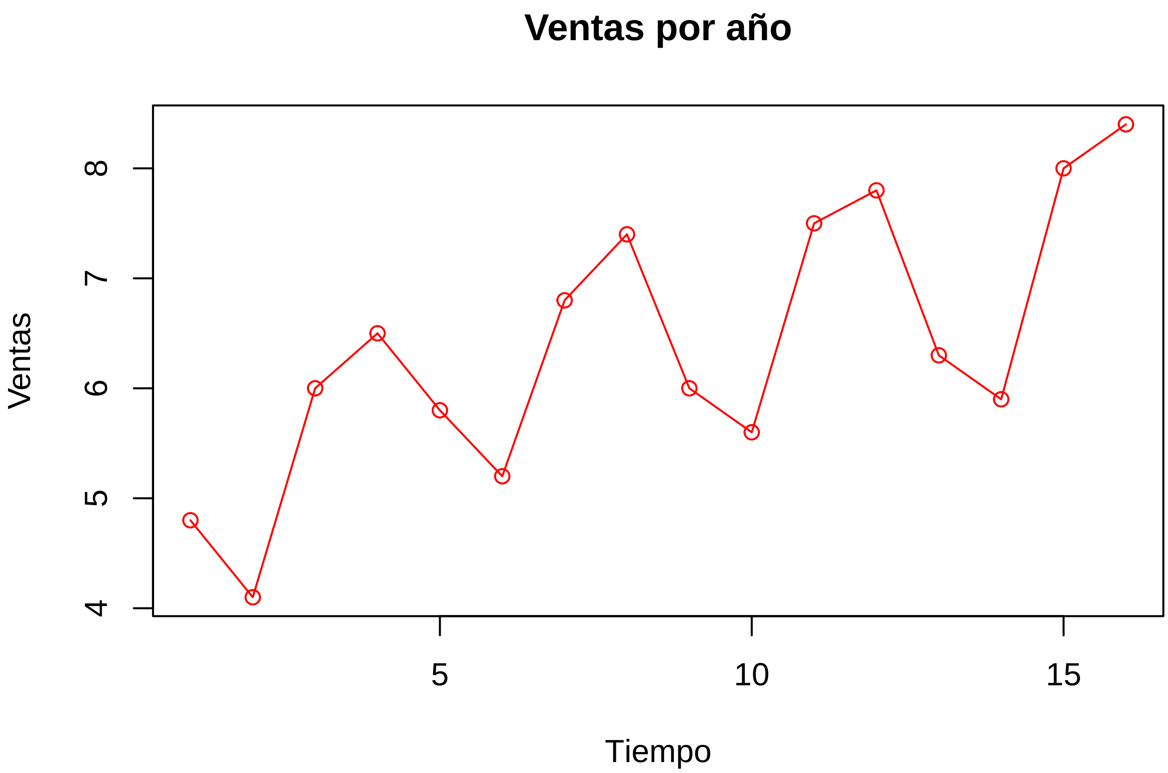
21/11/2022

```r
data = data.frame(
  "Año"=c(1,NA,NA,NA,2,NA,NA,NA,3,NA,NA,NA,4,NA,NA,NA),
  "Trimestre"=c(1,2,3,4,1,2,3,4,1,2,3,4,1,2,3,4),
  "Ventas(miles)"=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)
)
head(data,n=16)
```

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

Grafico de dispersion de los datos

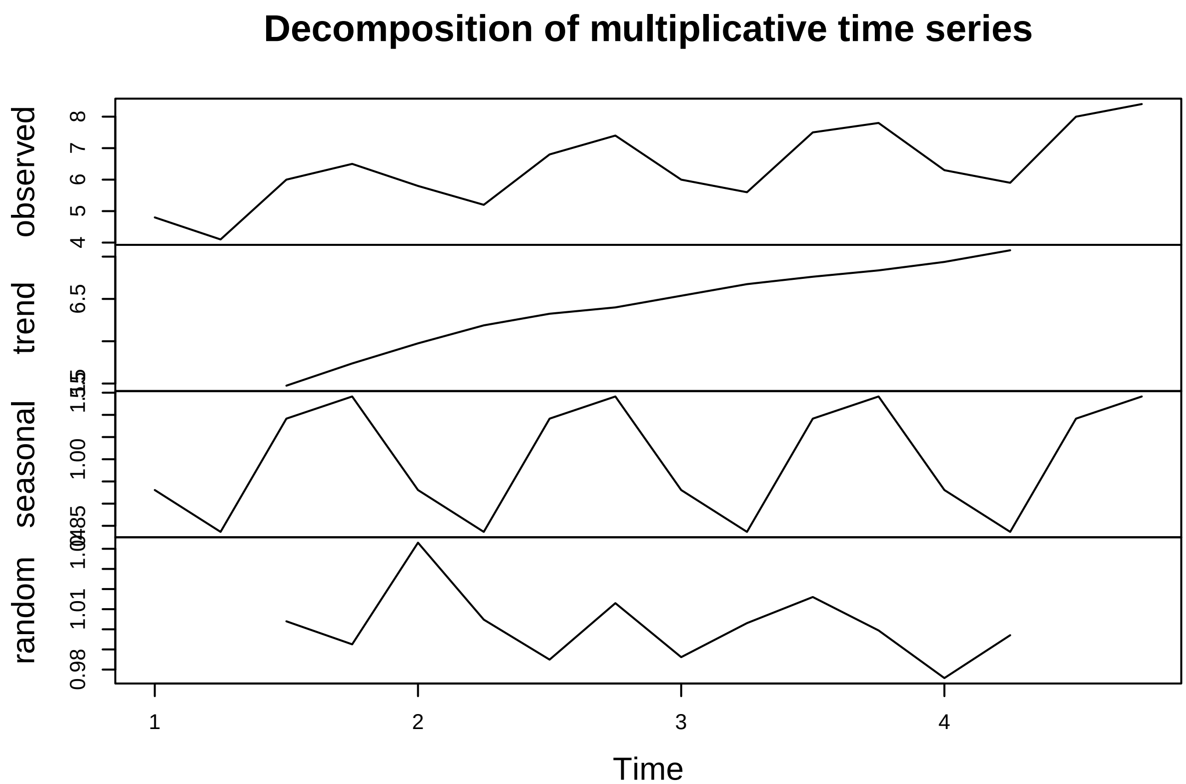
```r
t = 1:16
plot(t, data$Ventas.miles., type ="o", col ="red",ylab = "Ventas",xlab="Tiempo", main = "Ventas por año
```



Haciendo un análisis básico de lo que vemos en la gráfica de dispersión podemos ver que existen ciertas tendencias en los trimestres de acuerdo al año. En los trimestres 1 y 2 de cada año vemos que tendremos mínimas en el año en cuanto a ventas. El trimestre 3 suele ser mucho mayor al trimestre 2 cada año.Vemos que pasa algo similar de trimestre 4 a 1, existe una baja cuando se acaba el 4 trimestre y e llega al 1 del próximo año.

Descomposición

```
Tiempo = ts(data$Ventas.miles., frequency = 4, start(c(2016, 1)))
D =decompose(Tiempo, type = "m")
plot(D)
```




```
D
```

```
## $x
##   Qtr1 Qtr2 Qtr3 Qtr4
## 1  4.8  4.1  6.0  6.5
## 2  5.8  5.2  6.8  7.4
## 3  6.0  5.6  7.5  7.8
## 4  6.3  5.9  8.0  8.4
##
## $seasonal
##        Qtr1      Qtr2      Qtr3      Qtr4
## 1 0.9306617 0.8363763 1.0915441 1.1414179
## 2 0.9306617 0.8363763 1.0915441 1.1414179
## 3 0.9306617 0.8363763 1.0915441 1.1414179
## 4 0.9306617 0.8363763 1.0915441 1.1414179
##
## $trend
##     Qtr1   Qtr2   Qtr3   Qtr4
## 1     NA     NA 5.4750 5.7375
## 2 5.9750 6.1875 6.3250 6.4000
## 3 6.5375 6.6750 6.7625 6.8375
## 4 6.9375 7.0750     NA     NA
##
## $random
##        Qtr1      Qtr2      Qtr3      Qtr4
## 1        NA        NA 1.0039818 0.9925353
## 2 1.0430335 1.0048157 0.9849340 1.0129944
```

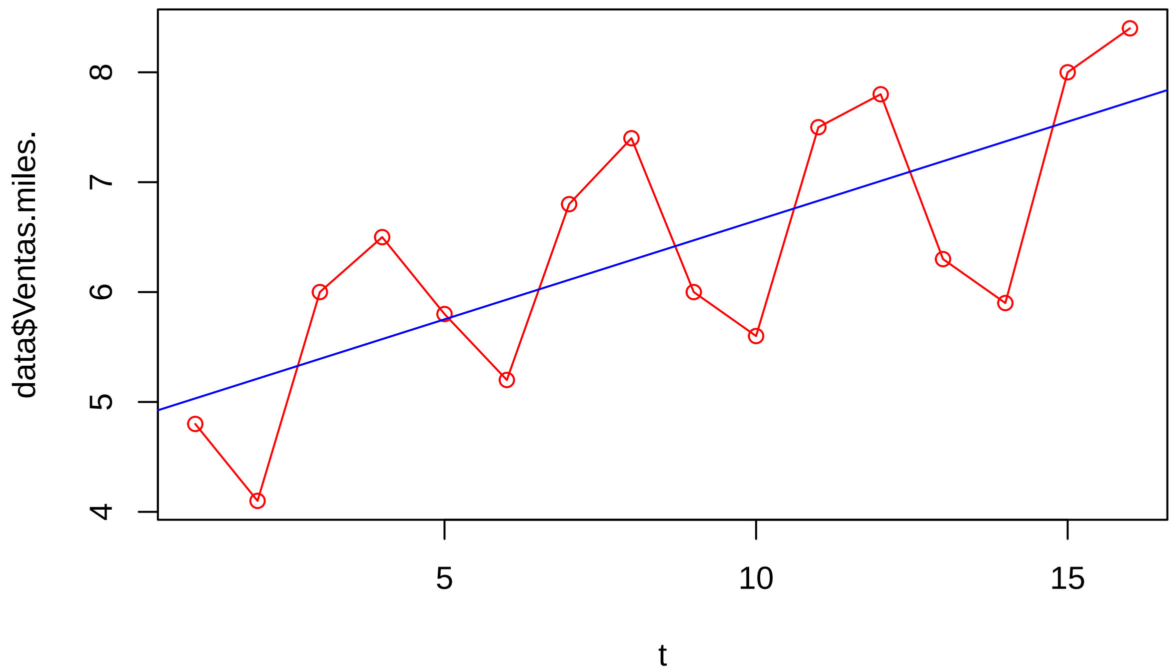
```
## 3 0.9861607 1.0030787 1.0160445 0.9994305
## 4 0.9757661 0.9970658        NA        NA
##
## $figure
## [1] 0.9306617 0.8363763 1.0915441 1.1414179
##
## $type
## [1] "multiplicative"
##
## attr(,"class")
## [1] "decomposed.ts"
```

Tendencia Lineal

```
N = lm(data$Ventas.miles.~t)
plot(t, data$Ventas.miles., type= "o", col = "red")
abline(N, col = "blue")
N
```

```
##
## Call:
## lm(formula = data$Ventas.miles. ~ t)
##
## Coefficients:
## (Intercept)            t
##      4.8525       0.1799
```

```
text(4,28,"ventas= 4.8525 + 0.1799 año")
```

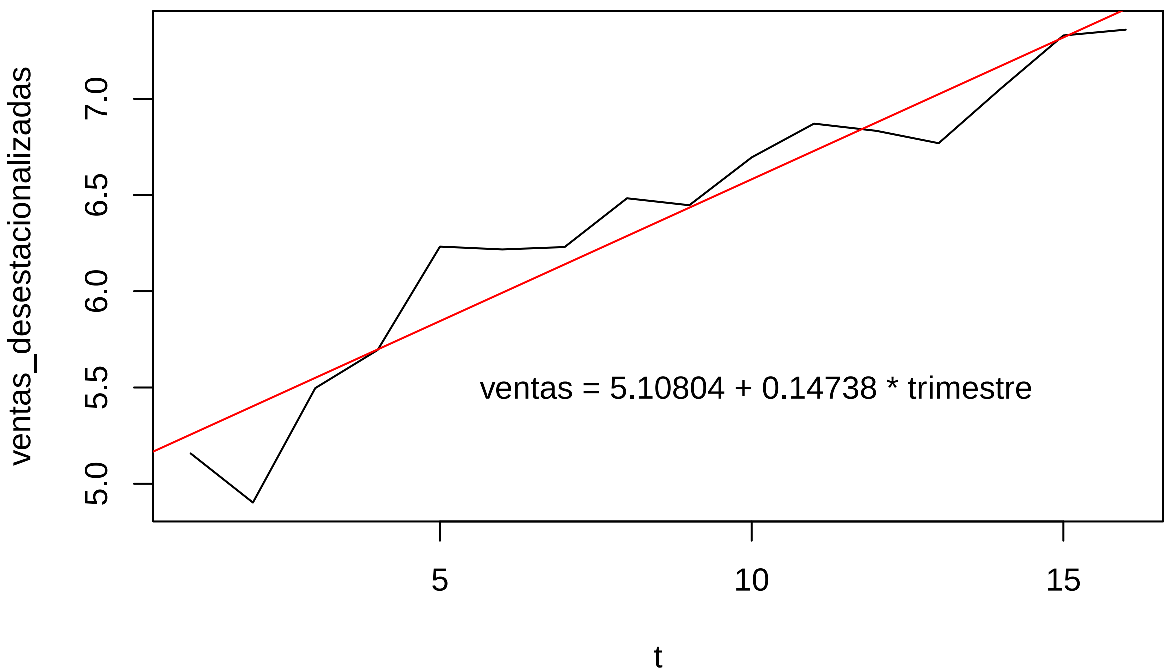


Ventas Desestacionalizadas vs Tiempo

```
ventas_desestacionalizadas = (D$x)/(D$seasonal)
N2 = lm(ventas_desestacionalizadas~t)
print(summary(N2))
```

```
##
## Call:
## lm(formula = ventas_desestacionalizadas ~ t)
##
## 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 ***
## t            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
```

```r
plot(t, ventas_desestacionalizadas, type = "l")
abline(N2, col = "red")
text(10, 5.5, " ventas = 5.10804 + 0.14738 * trimestre")
```



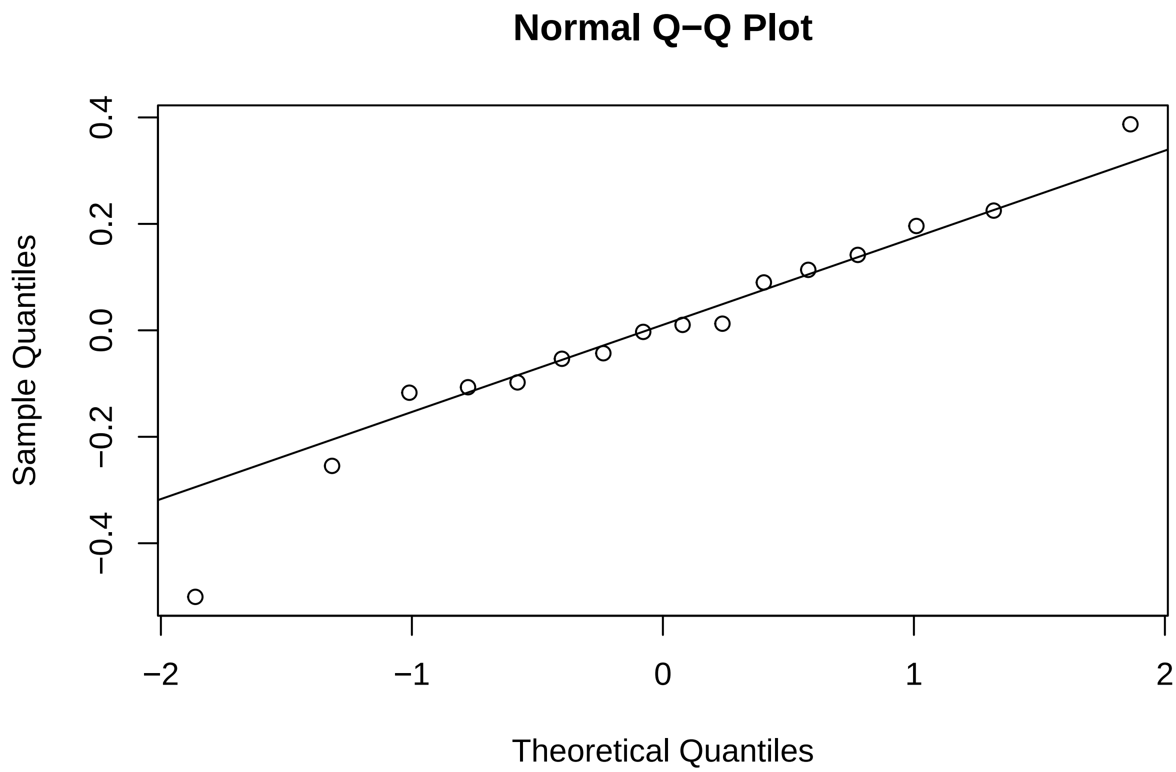
El modelo se ajusta bien a los datos con una $r^2 = 0.9151$, lo que nos habla bien de la variabilidad explicada.

Significancia de B1

La $\hat{}B_1$ nos dicta la pendiente, o la inclinación, de nuestra ecuación de ajuste. Vemos que el P-value de este coeficiente es menor a 0.05, lo que nos dice que este valor es significativo.

Análisis de los residuos

```r
qqnorm(N2$residuals)
qqline(N2$residuals)
```



Vemos que la línea de nuestro QQ PLOT nos dice que existe noralidad en nuestros residuos.

Prueba normalidad (Shapiro)

```
shapiro.test(N2$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  N2$residuals
## W = 0.96379, p-value = 0.7307
```

Vemos que nuestro p-value no es menor a nuestro alfa de 0.05, con lo que podemos decir que existe normalidad en los datos.

Pronósticos para próximo año

```
f = function(x) {-3.5443 + 0.4847*x}
# Los índices estacionales son:
a1 = D$seasonal[1]
a2 = D$seasonal[2]
a3 = D$seasonal[3]
a4 = D$seasonal[4];
f(17)*a1
```

```
## [1] 4.370015
```

```r
f(18)*a2
```

```
## [1] 4.33268
```

```r
f(19)*a3
```

```
## [1] 6.183597
```

```r
f(20)*a4
```

```
## [1] 7.019378
```

```r
CME=mean(exp(2),na.rm=TRUE)
CME
```

```
## [1] 7.389056
```

Problema : "Un problemilla más"

```r
df = data.frame(
  "Trimestre"=c(1,2,3,4),
  "Año 1"=c(1690,940,2625,2500),
  "Año 2"=c(1800,900,2900,2360),
  "Año 3"=c(1850,1100,2930,2615)
)
head(df, n=4)
```

```
##   Trimestre Año.1 Año.2 Año.3
## 1         1  1690  1800  1850
## 2         2   940   900  1100
## 3         3  2625  2900  2930
## 4         4  2500  2360  2615
```

a) Encuentre los promedios móviles de cuatro trimestres y los promedios móviles centrados

```r
prom1 = sum(df$Año.1)/4
prom2 = sum(df$Año.2)/4
prom3 = sum(df$Año.3)/4

cat("Promedios móviles","\n")
```

```
## Promedios móviles
```

```r
prom1
```

```
## [1] 1938.75
```

```r
prom2
```

```
## [1] 1990
```

```r
prom3
```

```
## [1] 2123.75
```

```r
cat("Promedios móviles centrados","\n")
```

```
## Promedios móviles centrados
```

```r
prom1c = mean(c(prom1,prom2))
prom2c = mean(c(prom2,prom3))
```

```
prom3c = mean(c(prom1c,prom2c))

prom1c
```

```
## [1] 1964.375
```

```
prom2c
```

```
## [1] 2056.875
```

```
prom3c
```

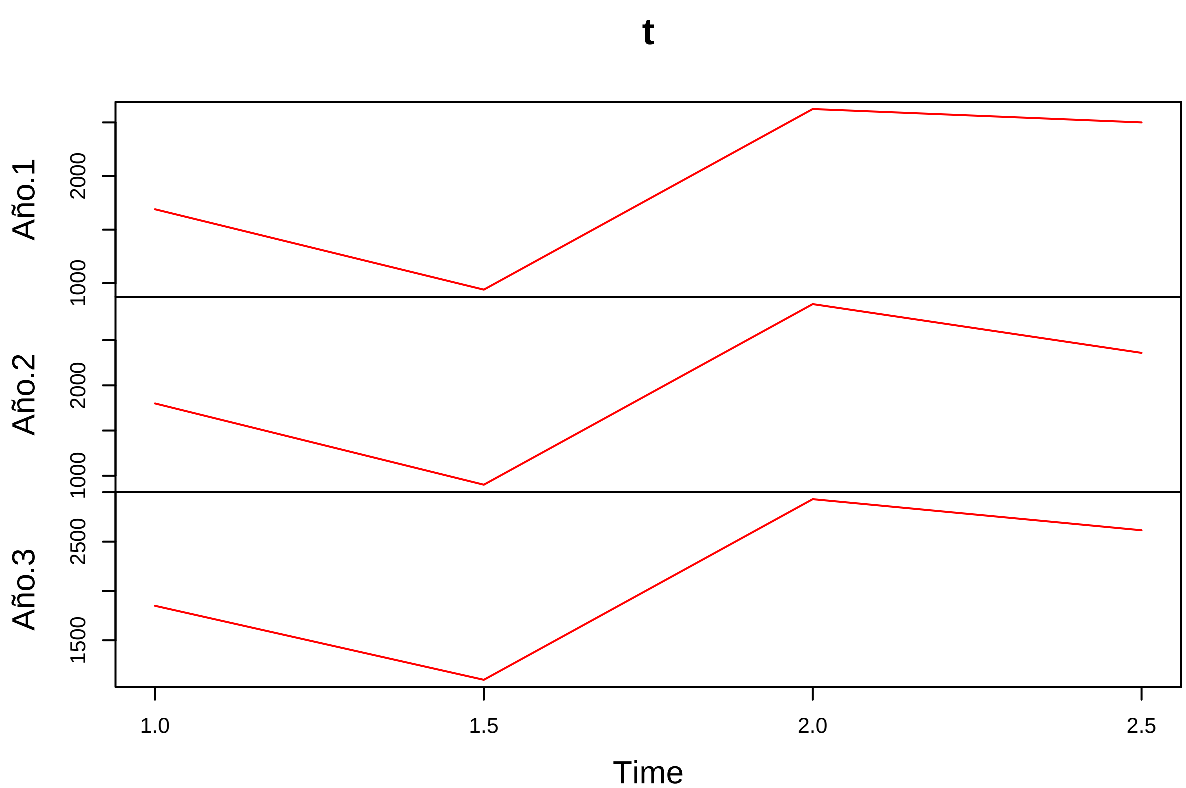
```
## [1] 2010.625
```

b) Calcule los índices estacionales de los cuatro trimestres

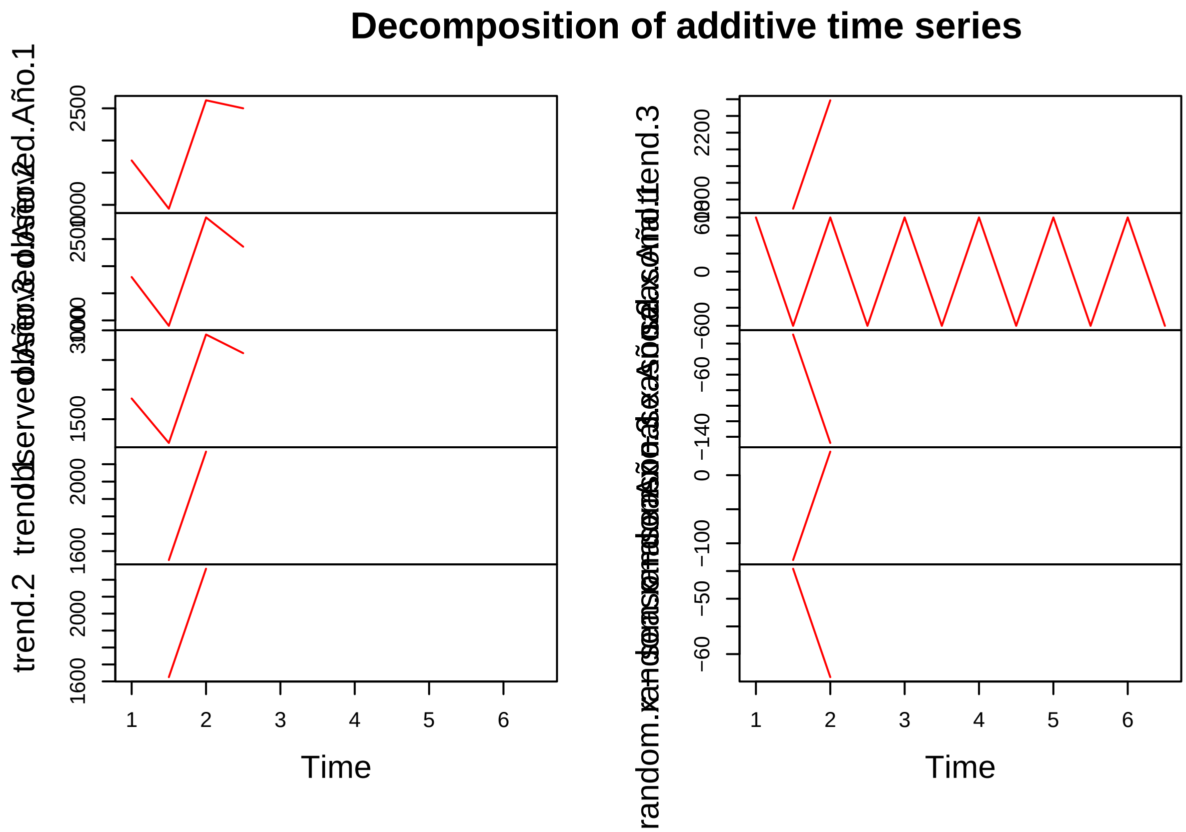
```
df2 = subset(df, select = -c(1) )
head(df2)
```

```
##   Año.1 Año.2 Año.3
## 1  1690  1800  1850
## 2   940   900  1100
## 3  2625  2900  2930
## 4  2500  2360  2615
```

```
t= ts(df2, frequency = 2, start(c(2016,1)))
plot.ts(t, col = "red")
```




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

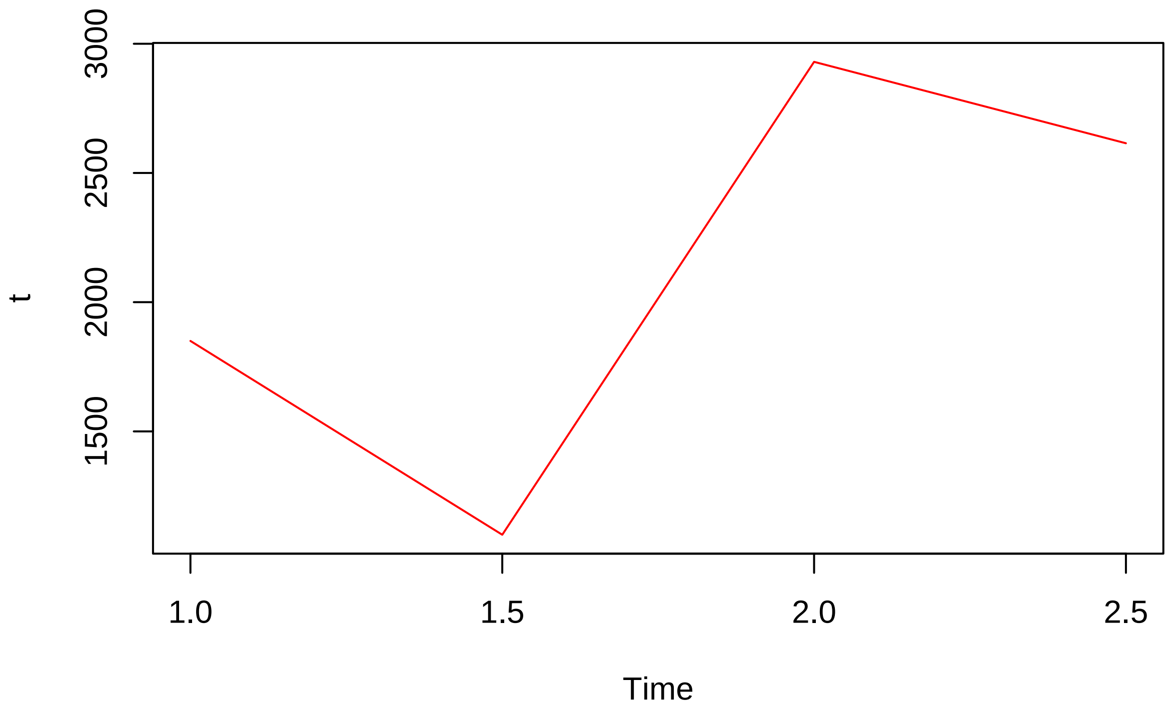



```
D$seasonal
```

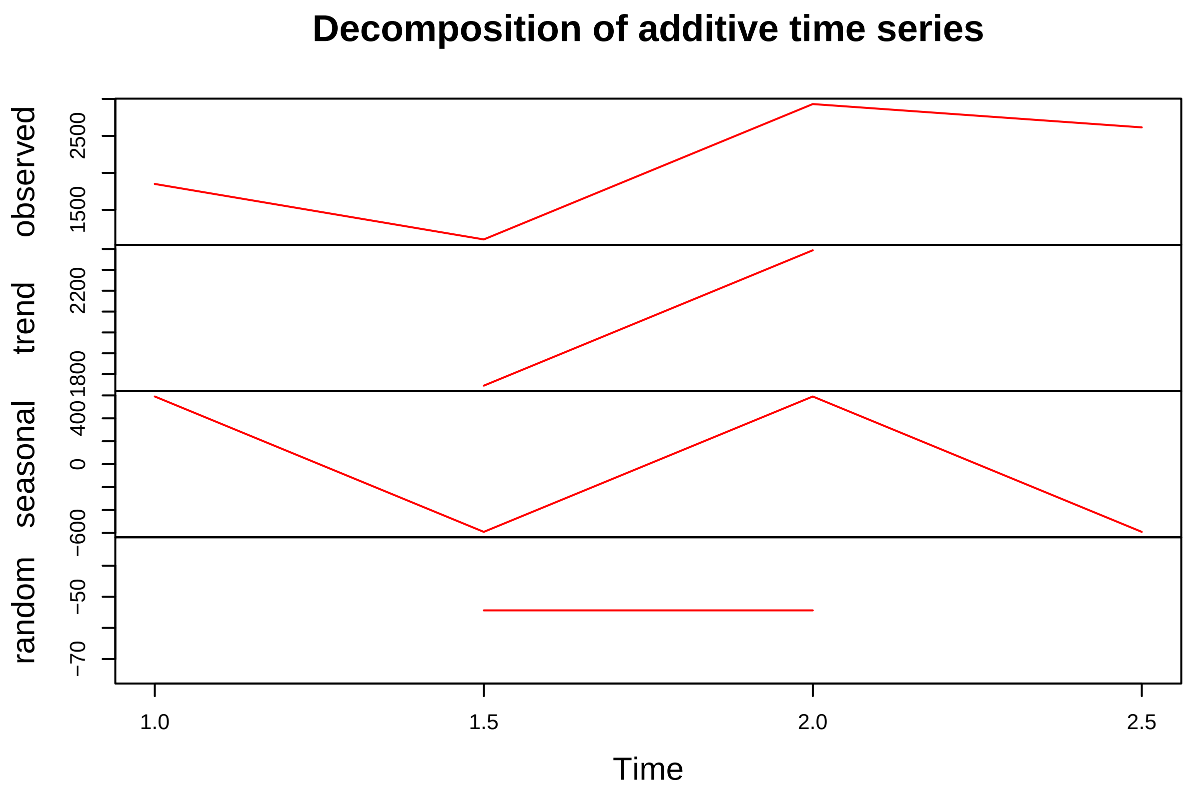
```
## Time Series:
## Start = c(1, 1)
## End = c(6, 2)
## Frequency = 2
##  [1]  600.4167 -600.4167  600.4167 -600.4167  600.4167 -600.4167  600.4167
##  [8] -600.4167  600.4167 -600.4167  600.4167 -600.4167
```

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

```
t= ts(df$Año.3, frequency = 2, start(c(2016,1)))
plot.ts(t, col = "red")
```




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




```
D$seasonal
```

```
## Time Series:
## Start = c(1, 1)
## End = c(2, 2)
## Frequency = 2
## [1]  590.625 -590.625  590.625 -590.625
```

El mayor índice estacionario se obtiene en el 1ero y 3er trimestre del año. Estos índices lucen extraños, pues entre estos trimestres no se muestra variación alguna.Al hacer un análisis de vista en los datos vemos que en los trimestres 3 y 4 tenemos mayor venta que en los otros trimestres; sin embargo, del trimestre 2 al 3 vemos que existe un salto muy grande en ventas. Quizá estos índices nos dicen algo, pero ciertamente su interpretación hay que trabajarla más.