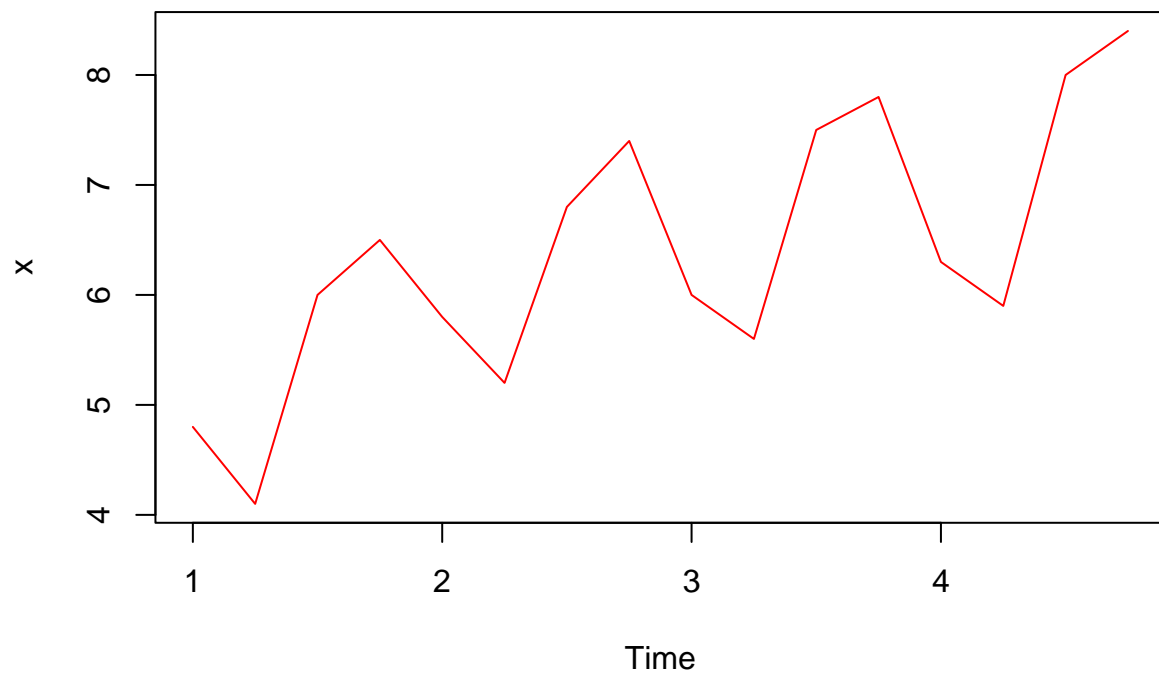


A8_Blanquita

Francisco Castorena, A00827756

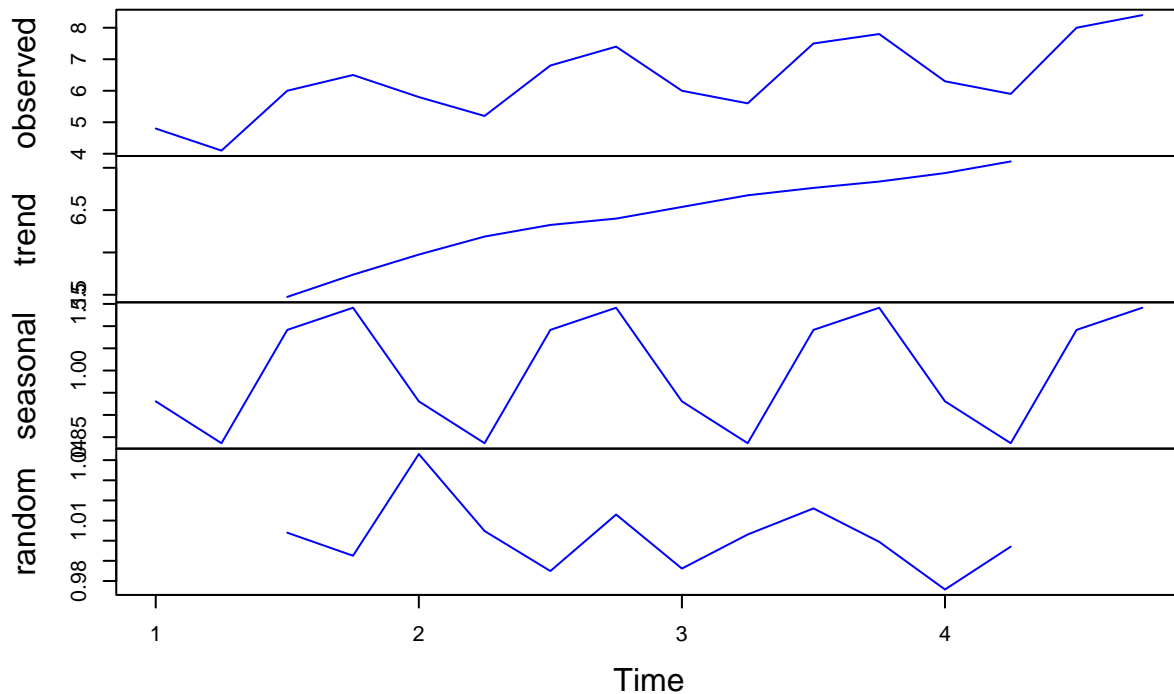
2023-11-18

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



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

Decomposition of multiplicative time series



Análisis de tendencia y estacionalidad

```
#tendencia <- D$trend
#estacionalidad <- D$seasonal
#print('Tendencia')
#print(tendencia)
```

Se puede observar una tendencia positiva por trimestre de los datos, en la gráfica se observa la tendencia positiva a largo plazo lo que sugiere que esta debe continuar, no hay indicios en estos datos que puedan indicar un decaimiento significativo a largo plazo.

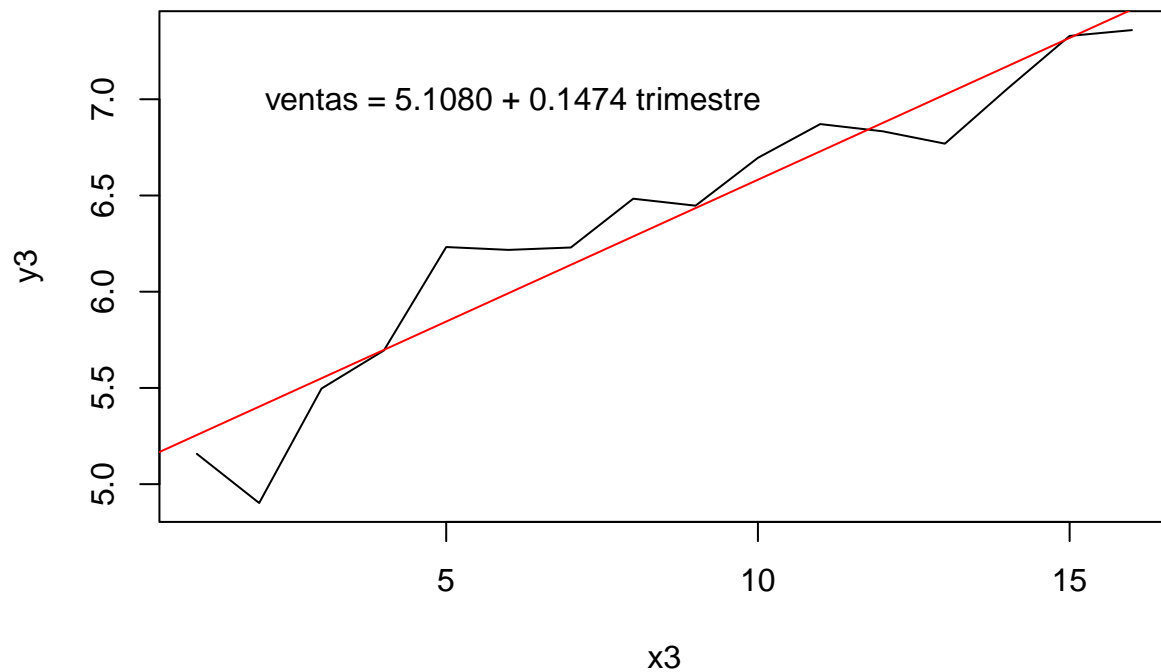
```
#print('Estacionalidad')
#print(estacionalidad)
```

Se puede observar un patrón cíclico de los datos analizando los trimestres por año, se puede observar como en los dos primeros trimestres los valores bajan y después en los otros dos, los valores suben, por lo que se puede esperar con estos datos que los primeros trimestres representen las ventas más bajas y los siguientes dos las más altas.

```
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
```

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



```
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
```

Significancia B1

En base a este modelo decimos que hay una relación significativa entre y_3 y x_3 , Beta1 (0.14738) sugiere la magnitud de cambio en y_3 por unidad de cambio en x_3

Variabilidad explicada por el modelo

El valor de r^2 de este modelo es de 0.9208, lo que indica que puede explicar un 92.08% de la variabilidad de los datos, este es un porcentaje muy alto, por lo cual se considera un muy buen modelo.

Análisis de residuos

Podemos observar como la mediana de los residuos es cercana a 0 lo cual es un indicio de que los datos se comportan con normalidad, por otro lado, hay una distribución equitativa de distancias entre cuartiles, podemos ver que entre el min y la mediana hay una distancia cercana a 0.5 y de la mediana a max también hay una distancia (0.38) relativamente parecida a 0.5, por lo que se comportan de manera normal.

Prueba de normalidad

```
shapiro.test(resid(N3))
```

```
##
## Shapiro-Wilk normality test
##
## data:  resid(N3)
## W = 0.96379, p-value = 0.7307
```

Para la prueba de normalidad se analizó a los residuos del modelo generado, El valor del estadístico W (0.96379) sugiere que los datos se asemejan a una distribución normal, por otra para el p-value es alto, lo que indica que no hay evidencia suficiente para rechazar la hipótesis nula de normalidad.

CME

```
mean(resid(N3)^2)
```

```
## [1] 0.0397064
```

EPAM

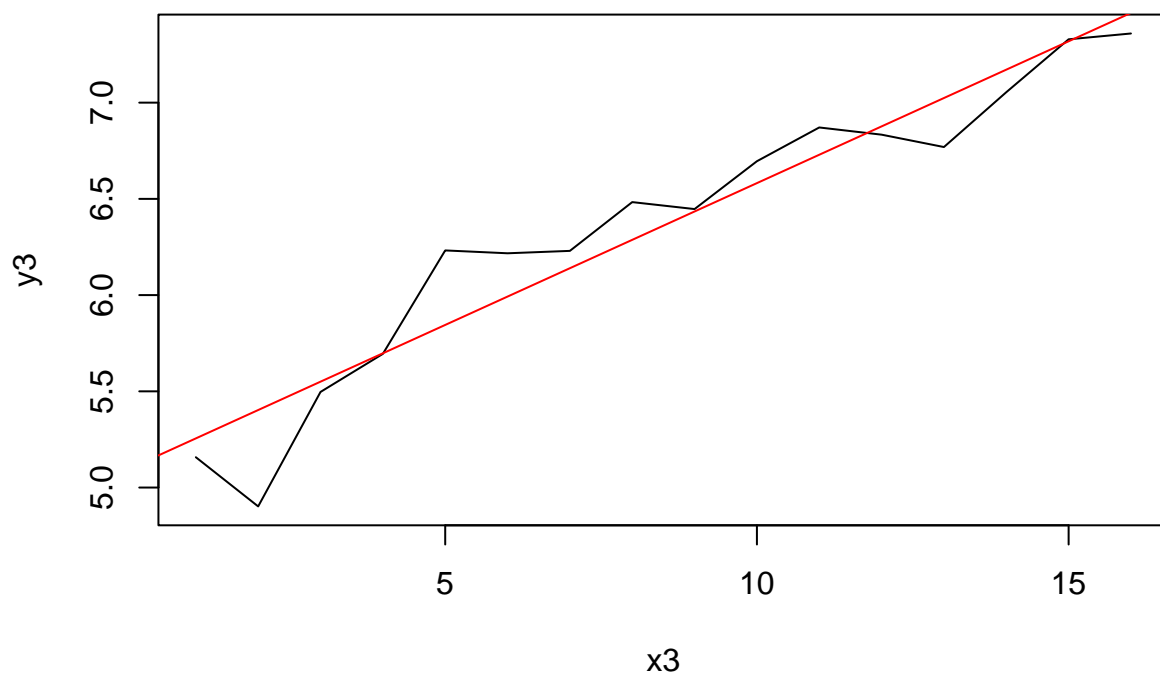
```
mean(abs(resid(N3)/ventas_desestacionalizadas)) * 100
```

```
## [1] 2.439533
```

SE busca que el CME y el EPAM sean bajos, podemos observar que ambos valores son bajos lo cual indica que el modelo se ajusta bien a los datos.

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

```
plot(x3, y3, type = "l")  
abline(N3, col = "red")
```



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

Se puede observar una relación significativa entre el trimestre y el valor de ventas, por otro lado la variabilidad explicada por el modelo es alta, el análisis de los residuos para dar indicios de que los datos se comportan de forma normal, viendo la gráfica se observa que la recta se ajusta bien a los datos, debido a que no se observa gran variabilidad de estos, por lo que se considera que este modelo es bueno para realizar predicciones sobre los datos, claro que siempre es posible garantizar la robustez del modelo cuando se toman más datos en cuenta, tomando únicamente a consideración los datos de la table, el modelo es bueno.

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

Aún y que el modelo presentado aquí ya es considerado bueno, se podrían implementar modelos no lineales, que sean capaz de seguir patrones no lineales sobre los datos, se podrían utilizar también modelos de predicción por técnicas vistas anteriormente como suavizamiento exponencial a regresión polinómica.

Realiza el pronóstico para el siguiente año.

```
# Supongamos que 'datos_nuevos' es un marco de datos con los valores de x3 para el próximo año
datos_nuevos <- data.frame(x3 = 17:20) # Aquí, asumo 4 trimestres para el próximo año

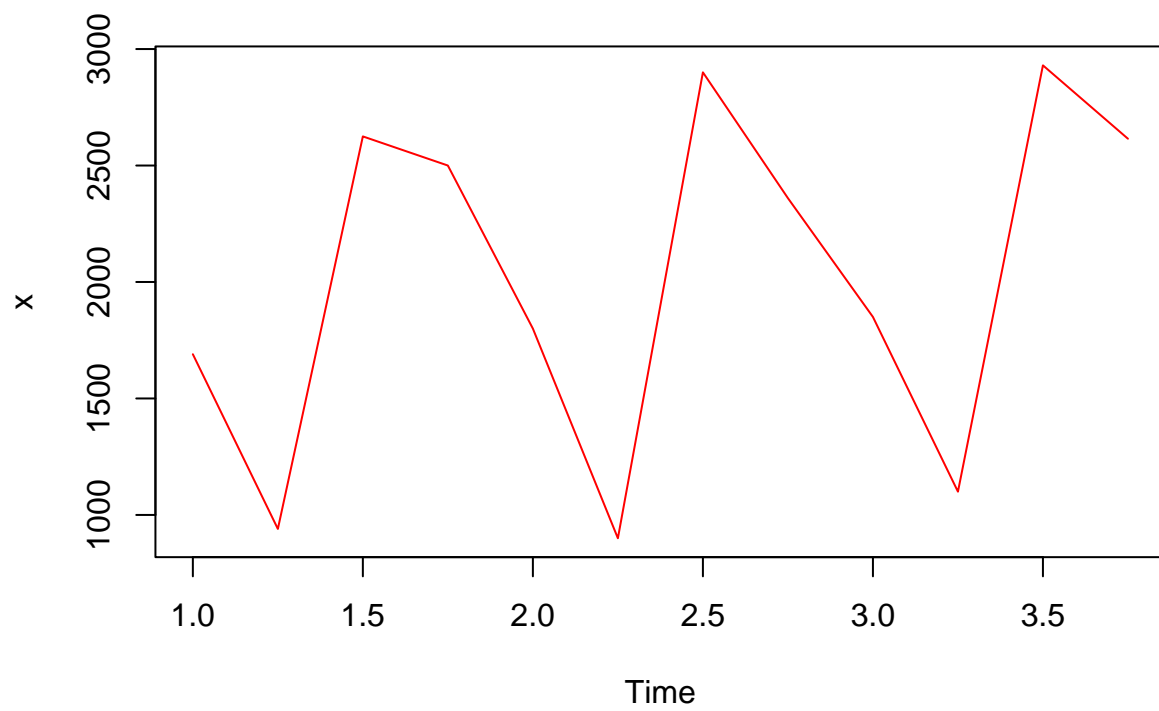
# Realiza el pronóstico utilizando el modelo de regresión lineal
pronostico <- predict(N3, newdata = datos_nuevos)

# Muestra el pronóstico
print(pronostico)
```

```
##           1           2           3           4
## 7.613536 7.760918 7.908300 8.055682
```

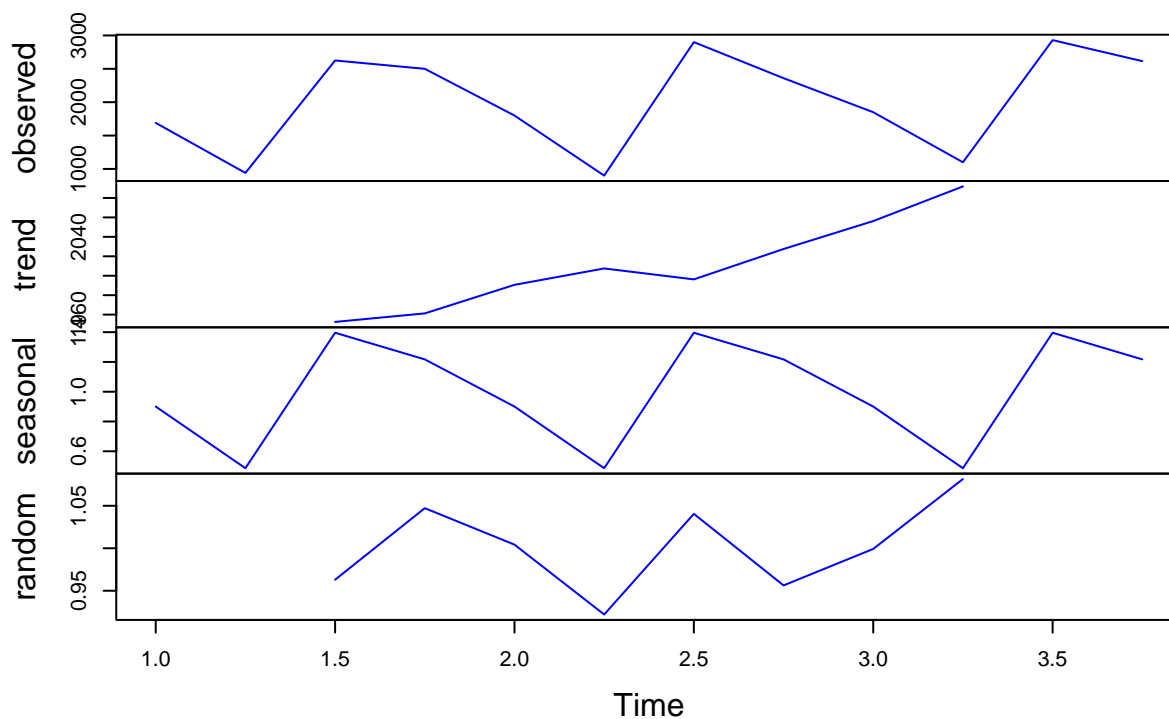
Realiza el problema de “Un problemilla más” sobre las ventas trimestrales de libros de texto universitarios.

```
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")
```



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

Decomposition of multiplicative time series



a) Encuentre los promedios móviles de cuatro trimestres

```
prom_4trim <- filter(x, rep(1/4, 4), sides = 1)
prom_4trim
```

```
##      Qtr1    Qtr2    Qtr3    Qtr4
## 1      NA      NA      NA 1938.75
## 2 1966.25 1956.25 2025.00 1990.00
## 3 2002.50 2052.50 2060.00 2123.75
```

Promedio movil centrado

```
prom_cent <- filter(x, rep(1/4, 4), sides = 2)
prom_cent
```

```
##      Qtr1    Qtr2    Qtr3    Qtr4
## 1      NA 1938.75 1966.25 1956.25
## 2 2025.00 1990.00 2002.50 2052.50
## 3 2060.00 2123.75      NA      NA
```


Indices estacionales de los cuatro trimestres

```
# Dividir los datos en trimestres
trimestres <- split(x, cycle(x))

# Calcular el promedio para cada trimestre
indices_estacionales <- sapply(trimestres, mean)

# Muestra los resultados
print(indices_estacionales)
```

```
##           1           2           3           4
## 1780.000  980.000 2818.333 2491.667
```

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

El mayor índice estacional se obtiene en el tercer trimestre, si parece razonable este resultado debido a que muchas escuelas (o todas) suelen empezar el ciclo escolar en este trimestre, lo cual si indica un aumento en la tendencia de venta de libros para esta época.