UNAP
FINESI

Series de Tiempo

VI Semestre
Grupo: B

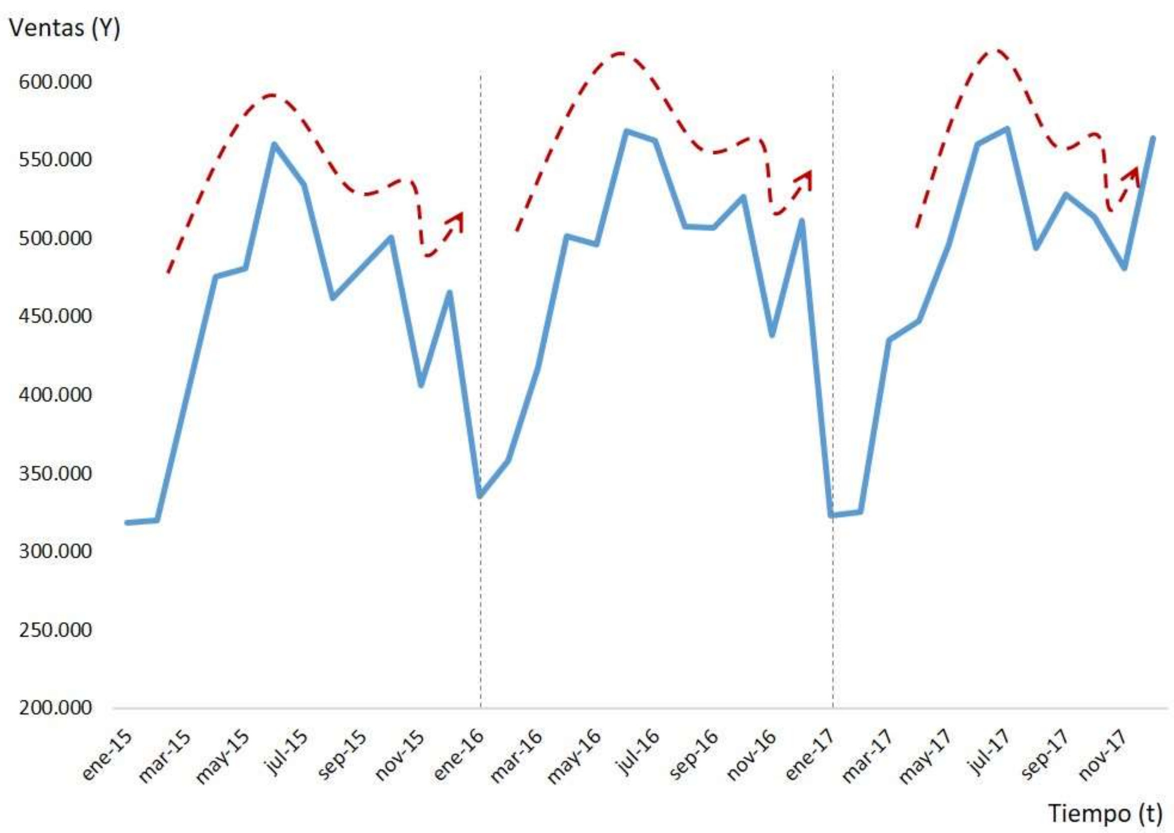
Mtr. Alcides Ramos Calcina
FINESI

DESCOMPOSICIÓN Y ANÁLISIS DE COMPONENTES - *Estacionalidad*

2. ANÁLISIS DE COMPONENTES

ESTACIONALIDAD

- Son los movimientos de la serie que se repiten de forma periódica, siendo la periodicidad inferior al año.
- Su identificación es importante, dado que una parte de las fluctuaciones de esta se manifiesta por el solo hecho de encontrarse en una época del año.
- Uno de los fines es su eliminación de la serie para visualizar otras componentes como tendencia y la irregularidad que se pueden confundir en las fluctuaciones estacionales.



ii) Estacionalidad

- Supondremos aquí que quitamos la componente de tendencia T_t de la serie y que el modelo aditivo es adecuado, esto es:

$$Y_t - T_t = E_t + a_t$$

Donde:

a_t : es ruido blanco

- Este fenómeno se puede detectar gráficamente, entre los más utilizados tenemos:
  - Gráfico de líneas
  - Líneas apiladas
  - Líneas separadas
  - Correlograma

- Los métodos para cuantificar la estacionalidad y desestacionalizar la serie son:
  - Regresión con variables Dummy
  - Métodos de suavizamiento
    - Promedio móvil
    - Alisado
  - Método X11
  - CENSUS X12-ARIMA
  - TRAMO/SSEATS

Ejemplo 4

La tabla 3.4 nos muestra las ventas mensuales del comercio al por mayor y menor y en la tabla 3.5 tenemos el valor de los Activos por acción trimestral de una compañía de administración de inversiones. Realice el análisis del componente estacional a través de los distintos métodos.

Tabla 3.4

Ventas mensuales del comercio al por mayor y menor periodo 2015 - 2018.

Mes	2015	2016	2017	2018
Enero	101684	101743	102661	108149
Febrero	82712	85799	82978	85679
Marzo	94665	92018	89527	93303
Abril	91333	90851	94226	96863
Mayo	97133	93858	96261	102144
Junio	98733	94407	94382	101962
Julio	104260	103603	105156	113154
Agosto	92082	90698	90961	95141
Setiembre	99904	96096	98168	103557
Octubre	98733	100762	102545	109875
Noviembre	95580	93643	94367	99527
Diciembre	126772	123248	129456	133554

Ejemplo 4

…

Tabla 3.5:
Valor de los activos por acción compañía General American Investors, periodo 2005-2016.

AÑO	TRIMESTRE			
	I	II	III	IV
2005	16.98	18.47	17.73	20.65
2006	21.95	23.85	20.44	19.29
2007	22.75	23.94	24.84	16.7
2008	18.04	19.19	18.97	17.03
2009	18.23	19.8	22.89	21.41
2010	21.5	25.05	20.33	20.6
2011	25.33	26.06	28.89	30.6
2012	27.44	26.69	28.71	28.56
2013	25.87	24.96	27.61	24.75
2014	23.32	22.61	24.08	22.31
2015	22.67	23.52	25.41	23.94
2016	25.68	-	-	-

1) Detección gráfica de la Estacionalidad

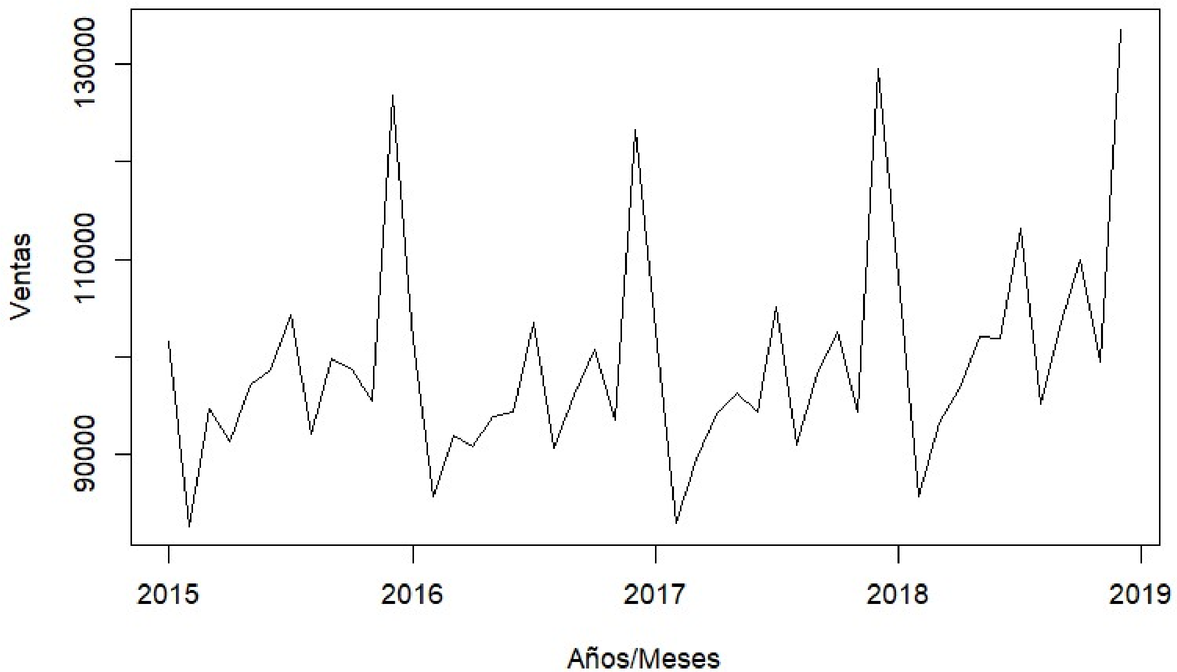
- **Gráfico de Líneas**

En la gráfica de serie original, nos interesa observar los picos que se dan y si estos picos se repiten todos los años, podemos concluir que existe estacionalidad.

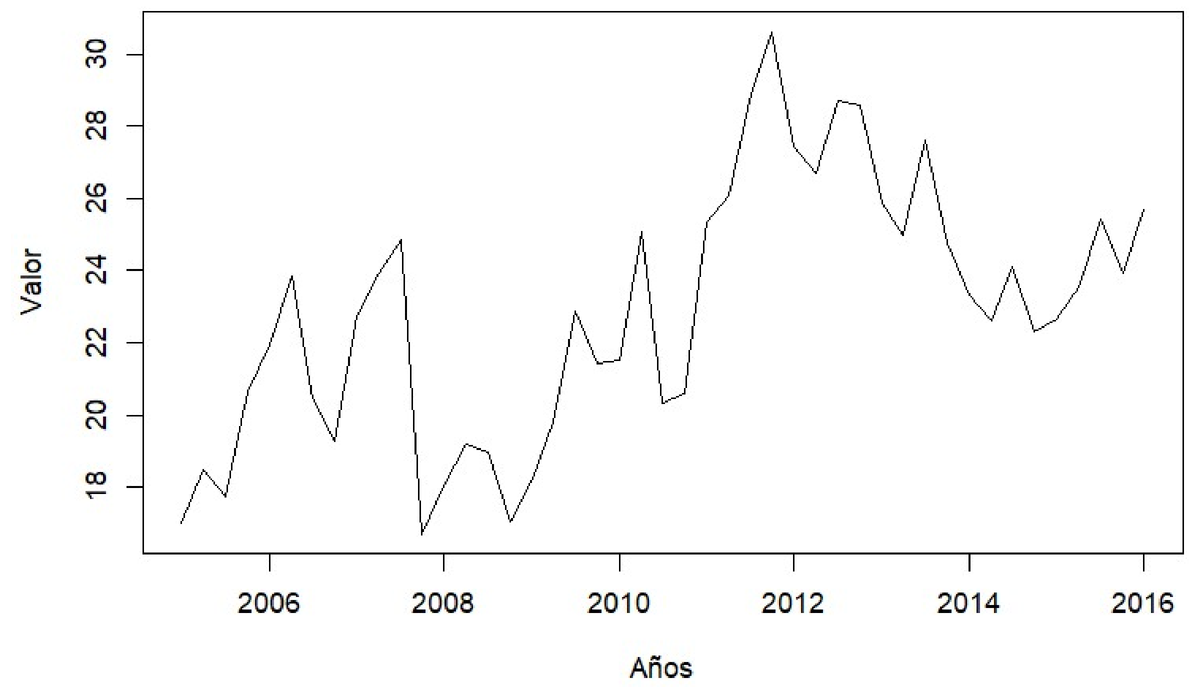
```
# Importar series
library(readxl)
serie_v <- read_excel("D:/.../Ejm_3_4.xlsx", + sheet = "Hoja1")
View(serie_v)

serie_a <- read_excel("D:/.../Ejm_3_4.xlsx", + sheet = "Hoja2")
View(serie_a)

# Convertir a una ST y gráfica
Yv <- ts(serie_v$Ventas, start = c(2015,1), frequency = 12)
plot(Yv, xlab="Años/Meses", ylab="Ventas")

Ya <- ts(serie_a$Activos, start = c(2005,1), frequency = 4)
plot(Ya, xlab="Años", ylab="Valor")
```

Ejemplos 4

Ventas mensuales (Yv) del comercio al por mayor y menor, 2015 - 2018.



Valor de los activos (Ya).



Detección gráfica de la Estacionalidad

- **Líneas apiladas**

Nos interesa observar el comportamiento de cada mes o trimestre, si el comportamiento es diferente, entonces existe estacionalidad.

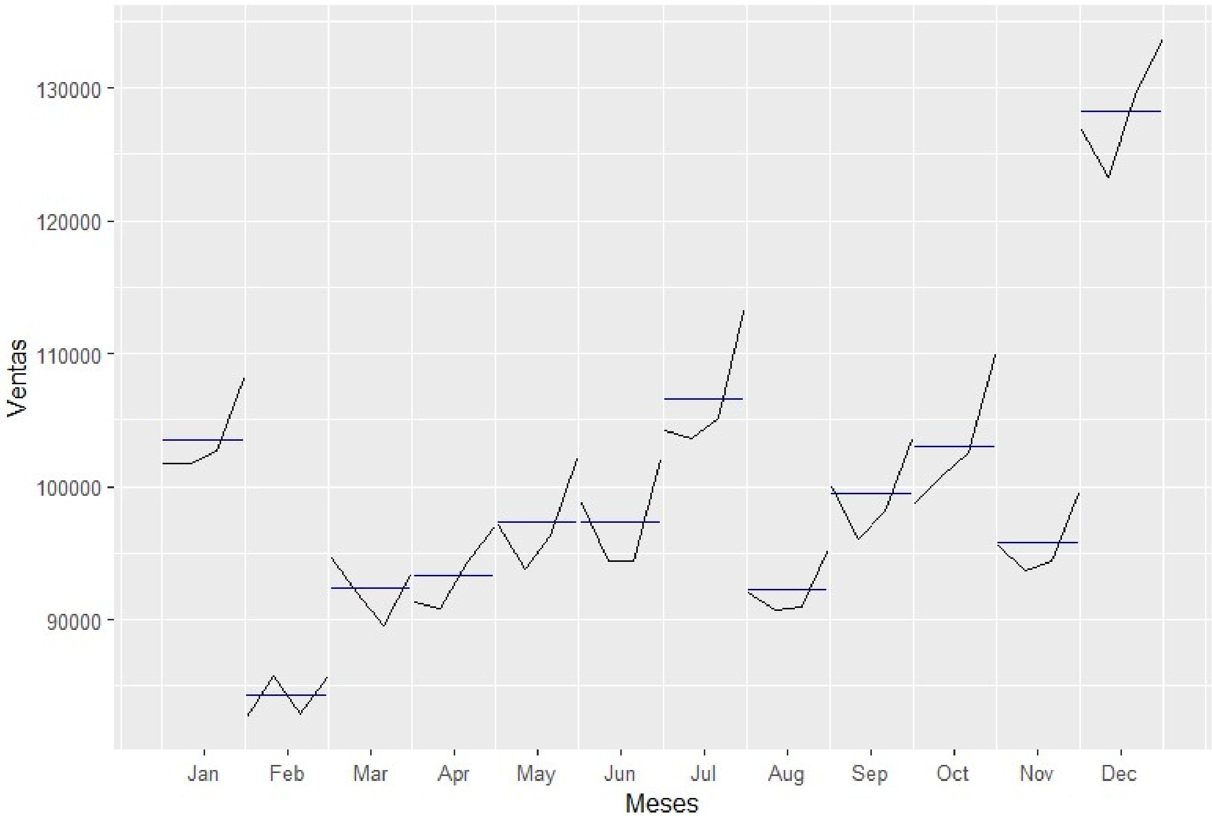
La estacionalidad de otros tipos, como los datos diarios, los datos por hora o los datos semanales, requieren un enfoque alternativo.

Script:

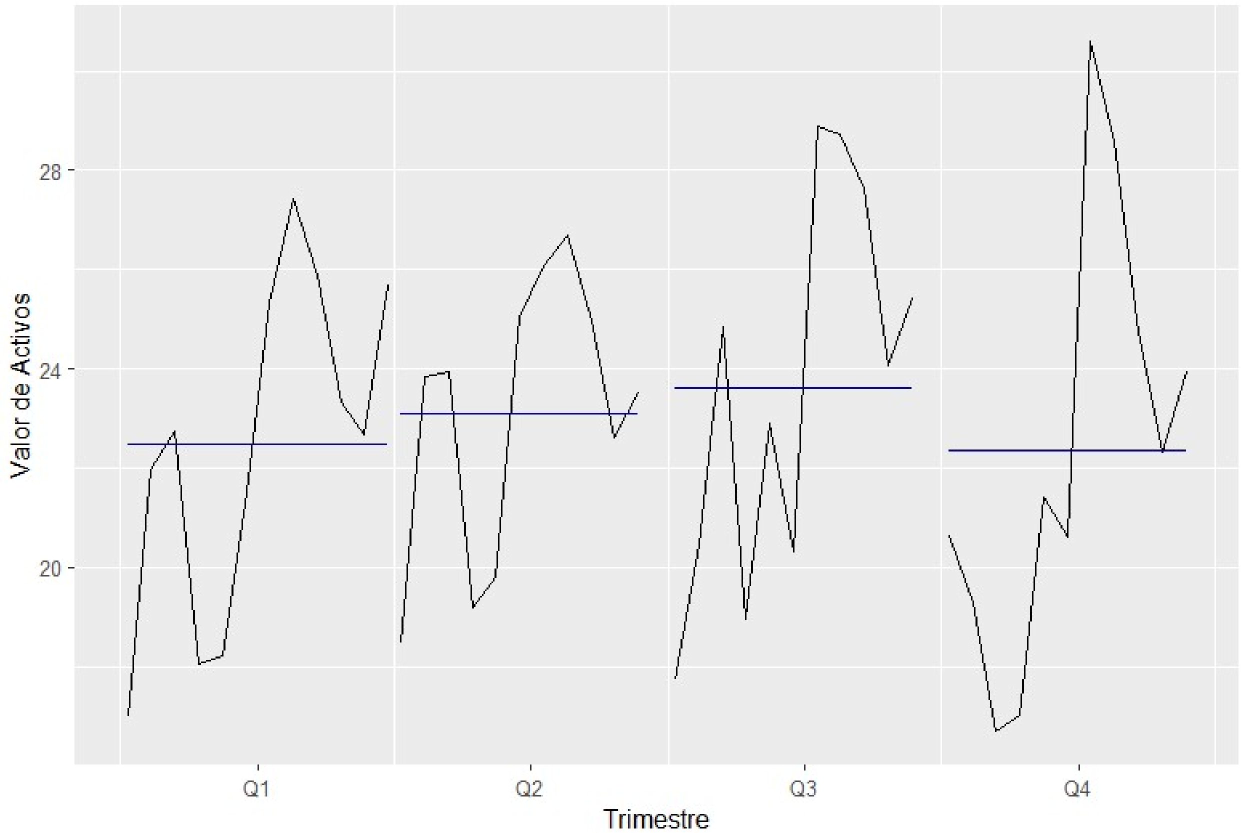
```
ggsubseriesplot(Yv, xlab = "Meses", ylab="Ventas")
ggsubseriesplot(Ya, xlab = "Trimestre", ylab="Valor de Activos")
```

Ejemplos 4

Líneas apiladas de las Ventas mensuales (Yv)



Líneas apiladas de los Valor de los activos (Ya).



Detección gráfica de la Estacionalidad

- **Líneas separadas**

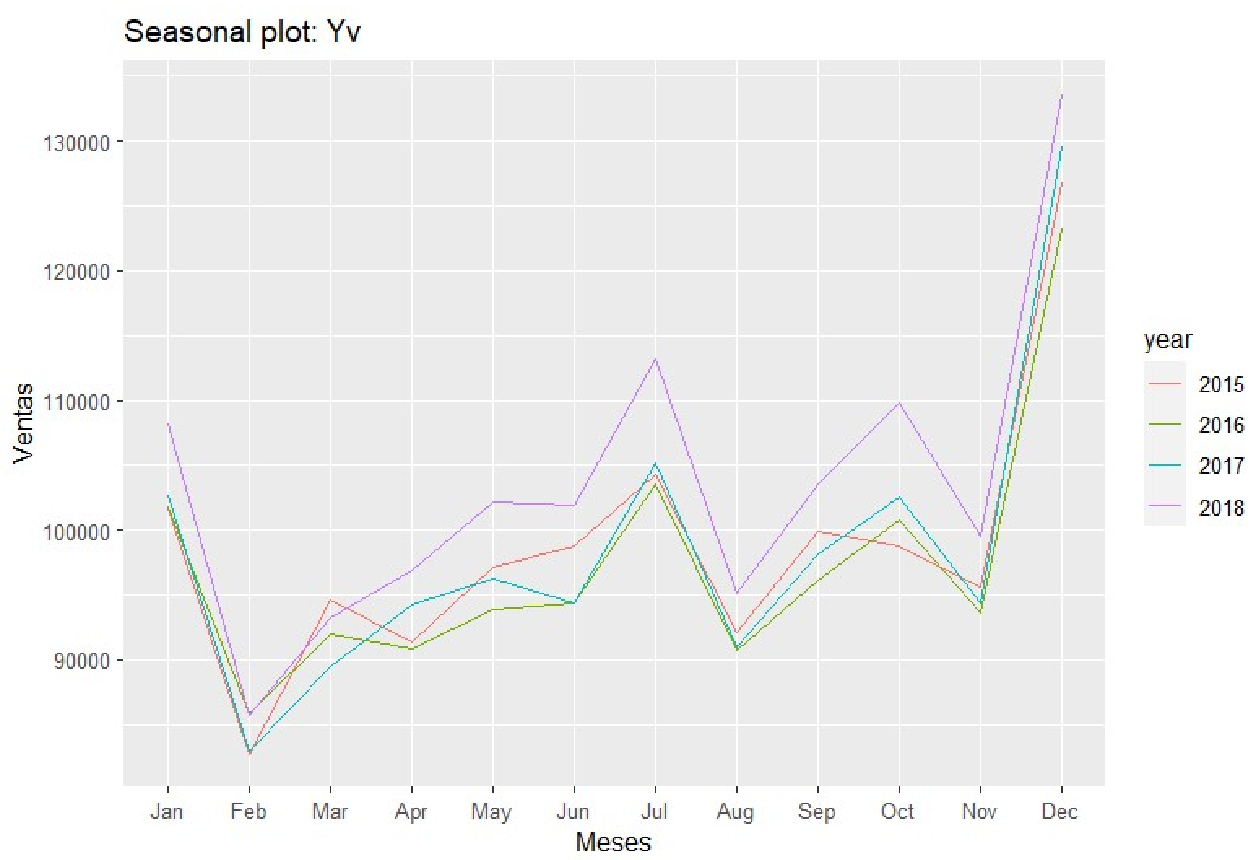
Nos interesa observar el comportamiento de cada año en los meses o trimestres, si el comportamiento no es diferente, entonces existe estacionalidad.

Graficando las variables Yv y Ya, tenemos:

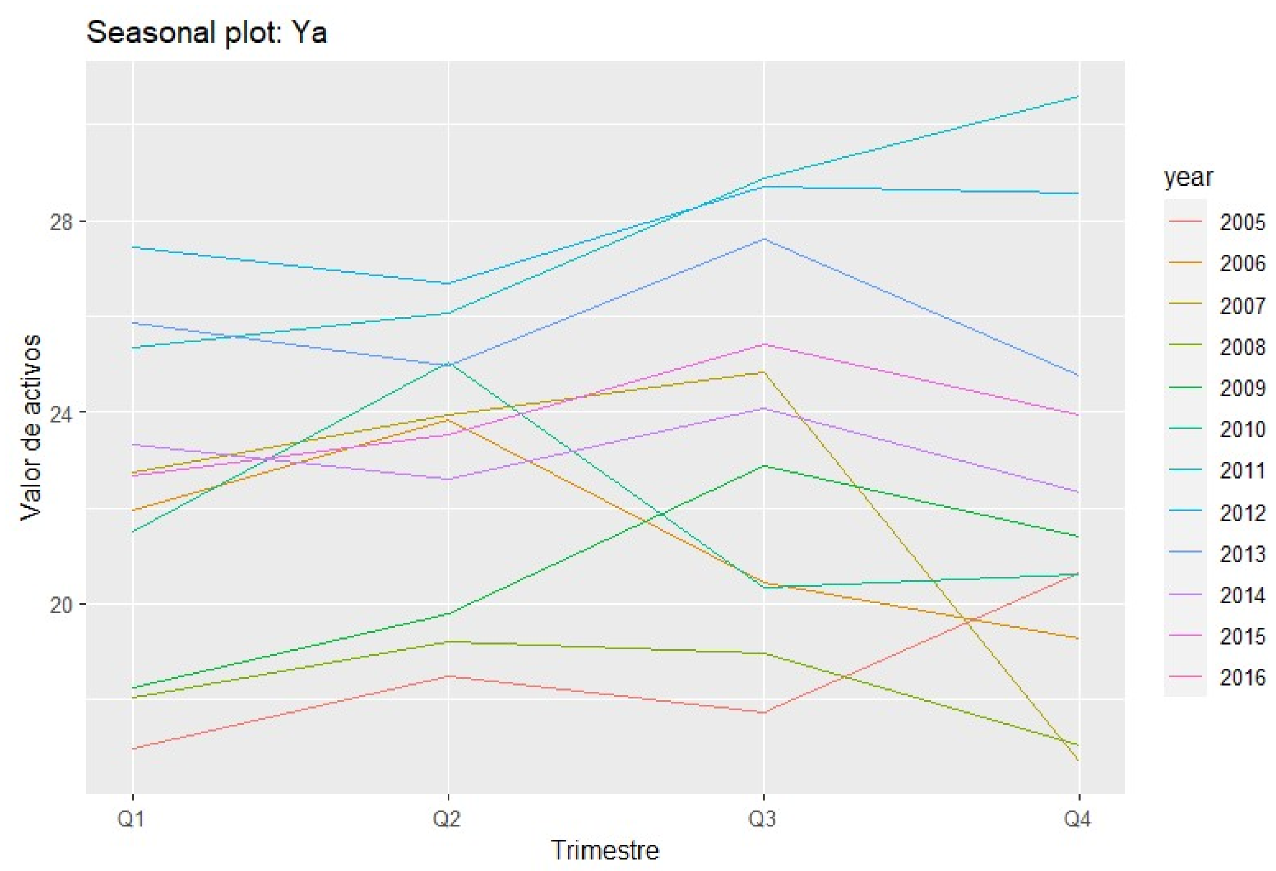
```
ggseasonplot(Yv, xlab = "Meses", ylab="Ventas")
ggseasonplot(Ya, xlab = "Trimestre", ylab="Valor de activos")
```

Ejemplos 4

Líneas separadas de las Ventas mensuales (Yv)



Líneas separadas de los Valor de los activos (Ya).



Detección gráfica de la Estacionalidad

- **Correlograma**

✓ La función de autocorrelación mide la correlación entre los valores de la serie distanciados un lapso de tiempo *k*.

✓ El coeficiente de autocorrelación para un retardo igual al periodo estacional debe ser significativamente diferente de 0.

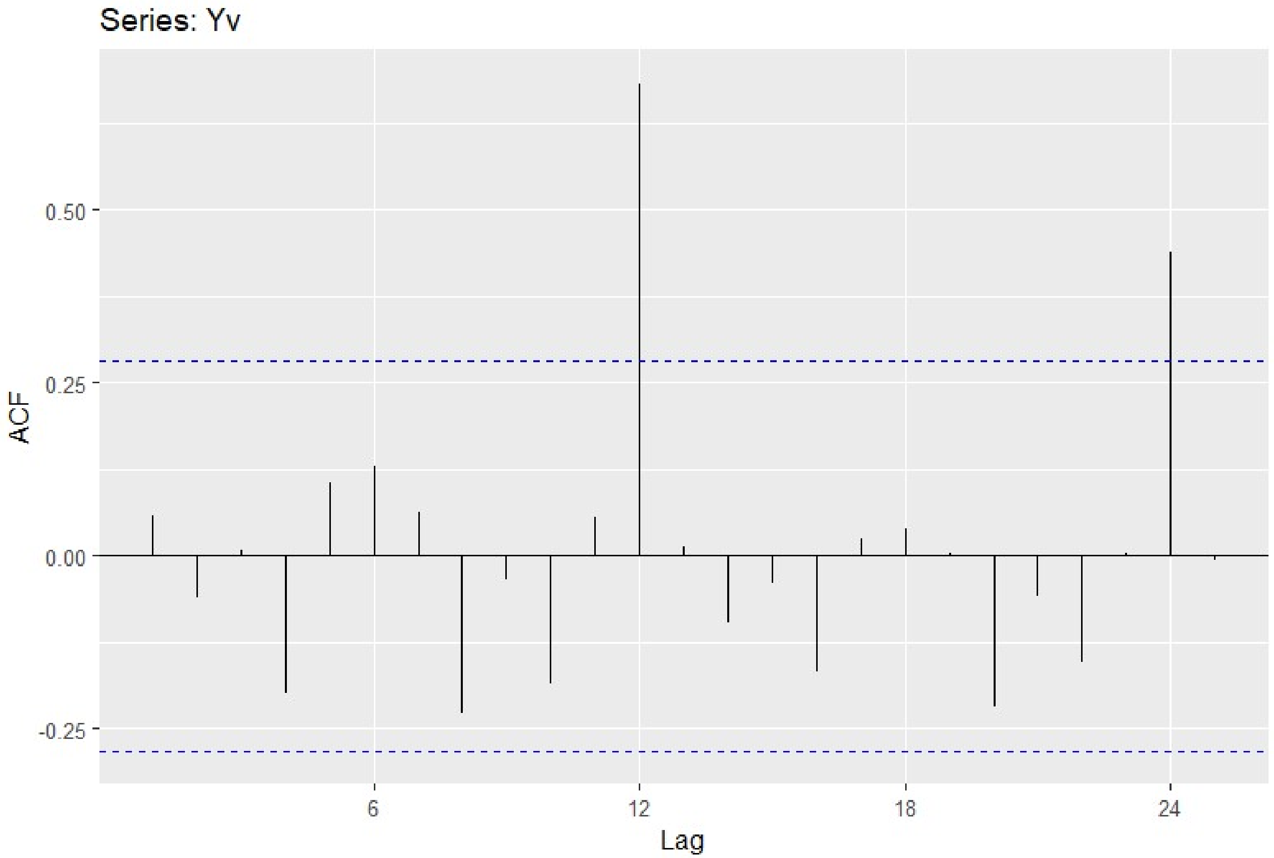
✓ Debemos obtener el correlograma considerando como mínimo 25 retardos si la serie es mensual, 9 retardos si la serie es trimestral, etc.

✓ Se tiene que observar los picos que se dan en el correlograma y si estos picos se repiten en el mismo periodo en los siguientes años, podemos concluir que existe estacionalidad.

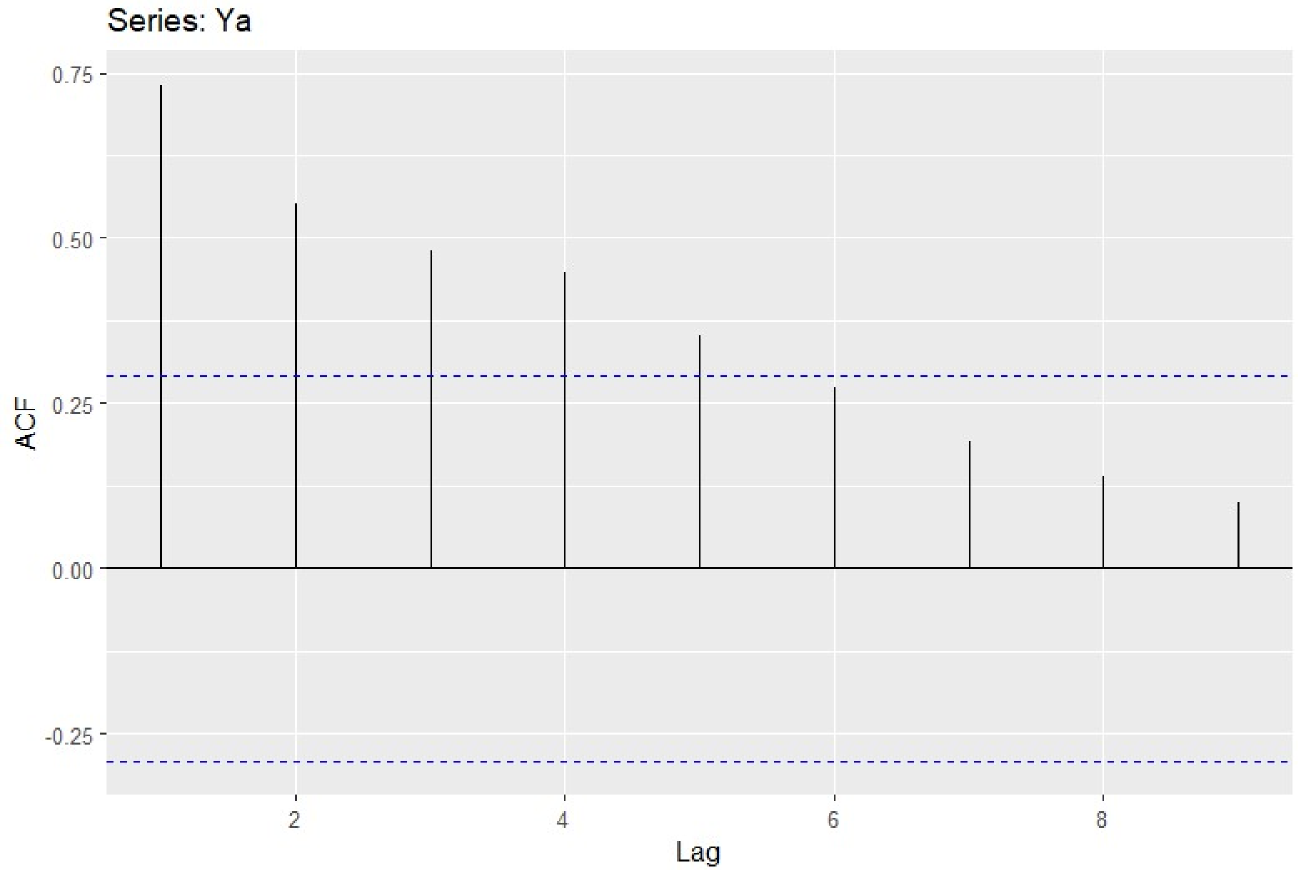
```
ggseasonplot(Yv, xlab = "Meses", ylab="Ventas")
ggseasonplot(Ya, xlab = "Trimestre", ylab="Valor de activos")
```

Ejemplos 4

Correlograma de las Ventas mensuales (Yv)



Correlograma de los Valor de los activos (Ya).



2) Cuantificación de la Estacionalidad

a) Cuantificación: Regresión con variables Dummy

- ✓ En el método de regresión se utiliza una serie de variables dicotómicas (Dummy) para capturar los efectos estacionales.

$$Y_t = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \cdots + \beta_k D_k + \gamma Y_{t-1} + \varepsilon_t$$

Donde, D_i toma valores 1 en el período que ésta representa (por ejemplo, enero de cada año) y cero en otro periodo.

- ✓ En la medida en que los coeficientes β_i sean estadísticamente significativos, habrá componentes estacionales.
- ✓ La serie desestacionalizada Y_t^{SE} se calcula como:

$$Y_t^{SE} = Y_t - \beta_0 - \beta_1 D_1 - \beta_2 D_2 - \cdots - \beta_k D_k$$

donde β_i son los parámetros estimados para cada componente estacional que resulten significativos.

Cuantificación: Regresión con variables Dummy

✓ El modelo lineal con variables Dummy para la variable Ventas es el siguiente:

$$Yv_t = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \cdots + \beta_{11} D_{11} + \gamma Yv_{t-1} + \varepsilon_t$$

```
# Incluir Ventas con retardo
serie_v$Ven_1 <- lag(serie_v$Ventas)

# Construcción de las Dummy
serie_v$dummy <- serie_v$Mes
tabla_dummy <- dummy_columns(serie_v, select_columns = c("dummy"),remove_first_dummy =
TRUE)
head(tabla_dummy)

# Dataframe para la regresion con Dummies
tabla_reg <- tabla_dummy[, c(4:5, 7:17)]

# Regresion con Dummies
reg_1 <- lm(serie_v$Ventas~., data = tabla_reg)
summary(reg_1)
```

Ejemplo 4

```
Call:
lm(formula = serie_v$Ventas ~ ., data = tabla_reg)
Residuals:
   Min      1Q Median      3Q     Max 
 -4586   -1729   -231    1675    3928 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) 13189.6431 16472.4490   0.801 0.428860    
Ven_1           0.7194     0.1297   5.545 3.36e-06 ***
dummy_2     -3395.1756  3525.7082  -0.963 0.342358    
dummy_3     18551.3774  5792.4138   3.203 0.002953 ** 
dummy_4     13674.3628  4813.0982   2.841 0.007544 ** 
dummy_5     17028.9040  4701.2087   3.622 0.000942 ***
dummy_6     14151.2989  4227.7360   3.347 0.002004 ** 
dummy_7     23307.7227  4225.1838   5.516 3.66e-06 ***
dummy_8      2386.7211  3205.7726   0.745 0.461685    
dummy_9     19900.8436  4831.9231   4.119 0.000230 ***
dummy_10    18261.1384  3987.9871   4.579 6.00e-05 ***
dummy_11     8509.6694  3589.4858   2.371 0.023564 *  
dummy_12    46167.0317  4410.7991  10.467 3.58e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2477 on 34 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.9628,	Adjusted R-squared:  0.9496 
F-statistic: 73.26 on 12 and 34 DF,  p-value: < 2.2e-16
```

En la regresión casi todas las variables Dummy son altamente significativas ($p < \alpha$), por lo tanto, las Ventas del Comercio presenta **ESTACIONALIDAD**.

Cuantificación: Regresión con variables Dummy

✓ Para la variable del Valor de los Activos, el modelo lineal con variables Dummy es:

$$ACT_t = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \gamma ACT_{t-1} + \varepsilon_t$$

```
# Incluir Activos con retardo
serie_a$Act_1 <- lag(serie_a$Activos)

# Construcción de las Dummy
serie_a$dummy <- serie_a$Trimestre
tabla2_dummy <- dummy_columns(serie_a, select_columns = c("dummy"),remove_first_dummy =
TRUE)
head(tabla2_dummy)

# Dataframe para la regresión con Dummies
tabla2_reg <- tabla2_dummy[, c(4:5, 7:9)]

# Regresión con Dummies
reg_2 <- lm(serie_a$Activos~., data = tabla2_reg)
summary(reg_2)
```

Ejemplo 4

```
Call:
lm(formula = serie_a$Activos ~ ., data = tabla2_reg)

Residuals:
    Min      1Q  Median      3Q     Max 
-6.5854 -0.8126 -0.2634  1.4639  4.1879 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  5.72560    2.25926   2.534   0.0154 *  
Act_1        0.77204    0.09650   8.001 9.45e-10 ***
dummy_2      0.24716    0.95205   0.260   0.7965    
dummy_3      0.06473    0.95470   0.068   0.9463    
dummy_4     -1.61772    0.95988  -1.685   0.0999 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.232 on 39 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.628,	Adjusted R-squared:  0.5898 
F-statistic: 16.46 on 4 and 39 DF,  p-value: 5.598e-08
```

En los resultados de la regresión, ninguna de las variables Dummy es significativa ($p > \alpha = 0.05$), por lo tanto , el Valor de los Activos **NO PRESENTA ESTACIONALIDAD.**

2) Cuantificación de la Estacionalidad

b) Desestacionalización

Medias Móviles

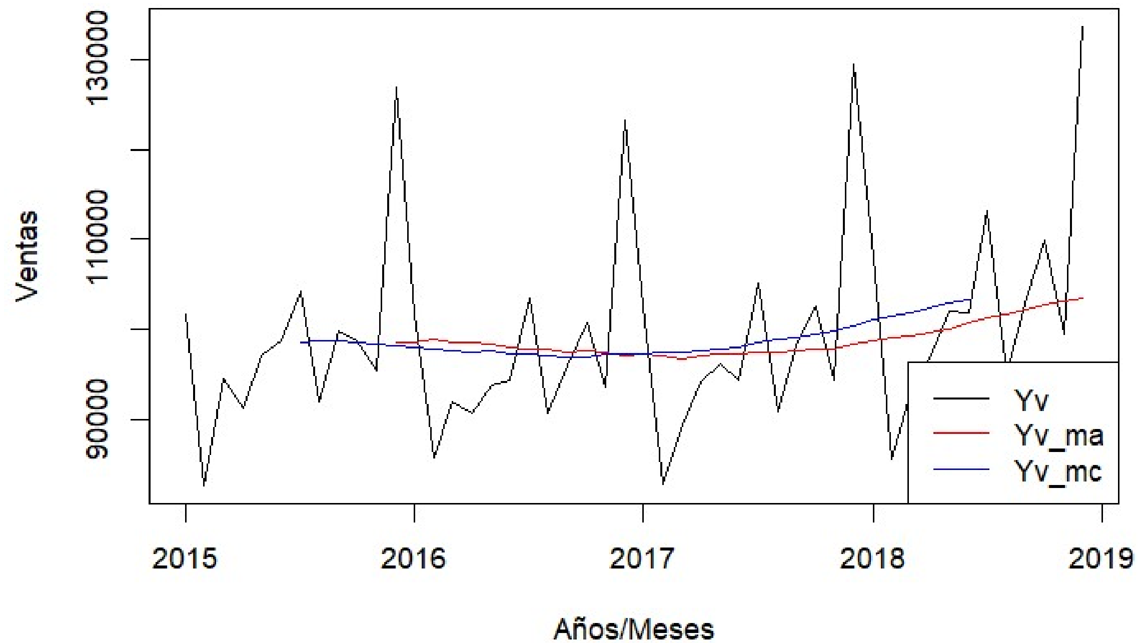
Se desestacionaliza transformando la serie original en la que las nuevas observaciones para cada periodo son un promedio de las observaciones originales.

Media móvil sin centrar	Medía móvil centrada
$\hat{Y}_{t+1} = \frac{1}{k}\sum_{r=1}^{k} Y_{t-r+1}$	$\hat{Y}_{t+1} = \frac{1}{k}\sum_{r=-k/2}^{k/2} \omega Y_{t-r}$

Desestacionalización

✓ Ventas del Comercio:

```
library(TTR) # Método de Media móvil
Yv_ma <- SMA(Yv, 12) # media móvil simple
Yv_mc <- ma(Yv, 12) # media móvil centrada
plot(Yv, xlab="Años/Meses", ylab="Ventas")
lines(Yv_ma, type = "l", col = "red")
lines(Yv_mc, type = "l", col = "blue")
legend(x = "bottomright", legend = c("Yv", "Yv_ma", "Yv_mc"), col = c('black', 'red', 'blue'),
lty=c(1,1,1))
```



2) Cuantificación de la Estacionalidad

<u>Alisado o Suavizamiento Exponencial</u>

- Se utilizan en situaciones en las que existen pocos datos (muestras pequeñas).
- Son muy simples en su aplicación, pero limitados.
- Entre estos tenemos:

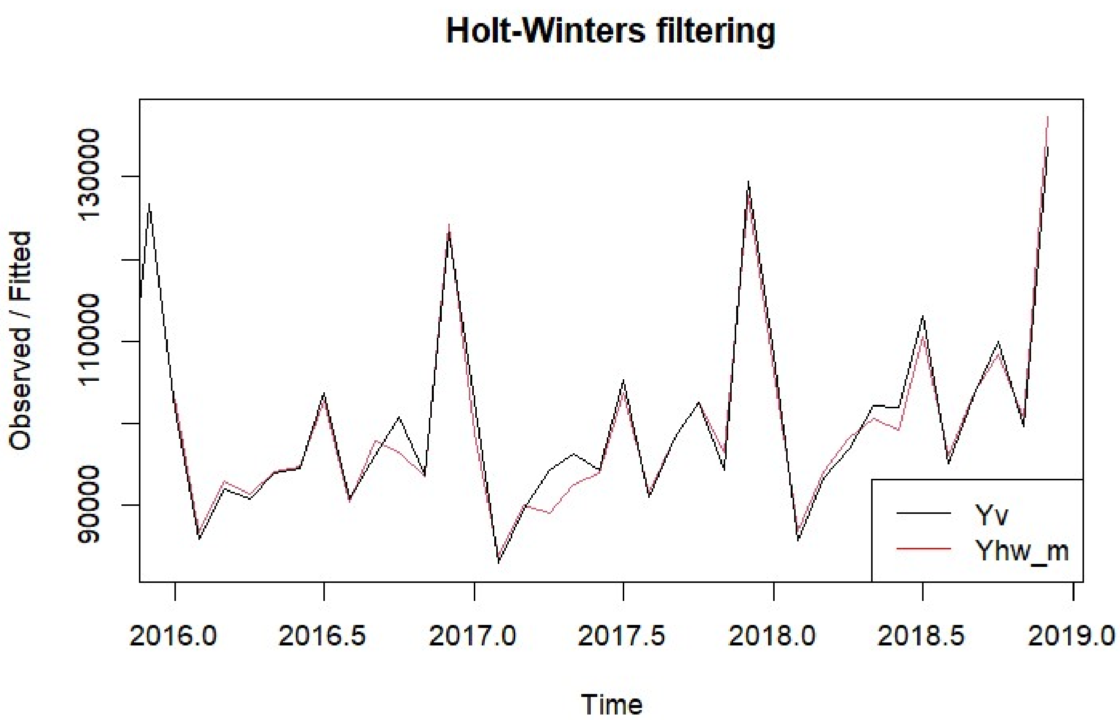
El modelo de Holt-Winters	Holt-Winters estacional multiplicativo
Para series con tendencia polinomial, ciclo estacional y componente aleatorio, tanto para tipo aditivo y multiplicativo.	El suavizamiento exponencial estacional con triple parámetro

Desestacionalización

✓ Ventas del Comercio:

```
Yhw_m <- HoltWinters(Yv, seasonal = "multiplicative")
plot(Yhw_m)
legend(x = "bottomright", legend = c("Yv", "Yhw_m"), col = c('black', 'red'), lty = c(1, 1))
Yhw_m
```

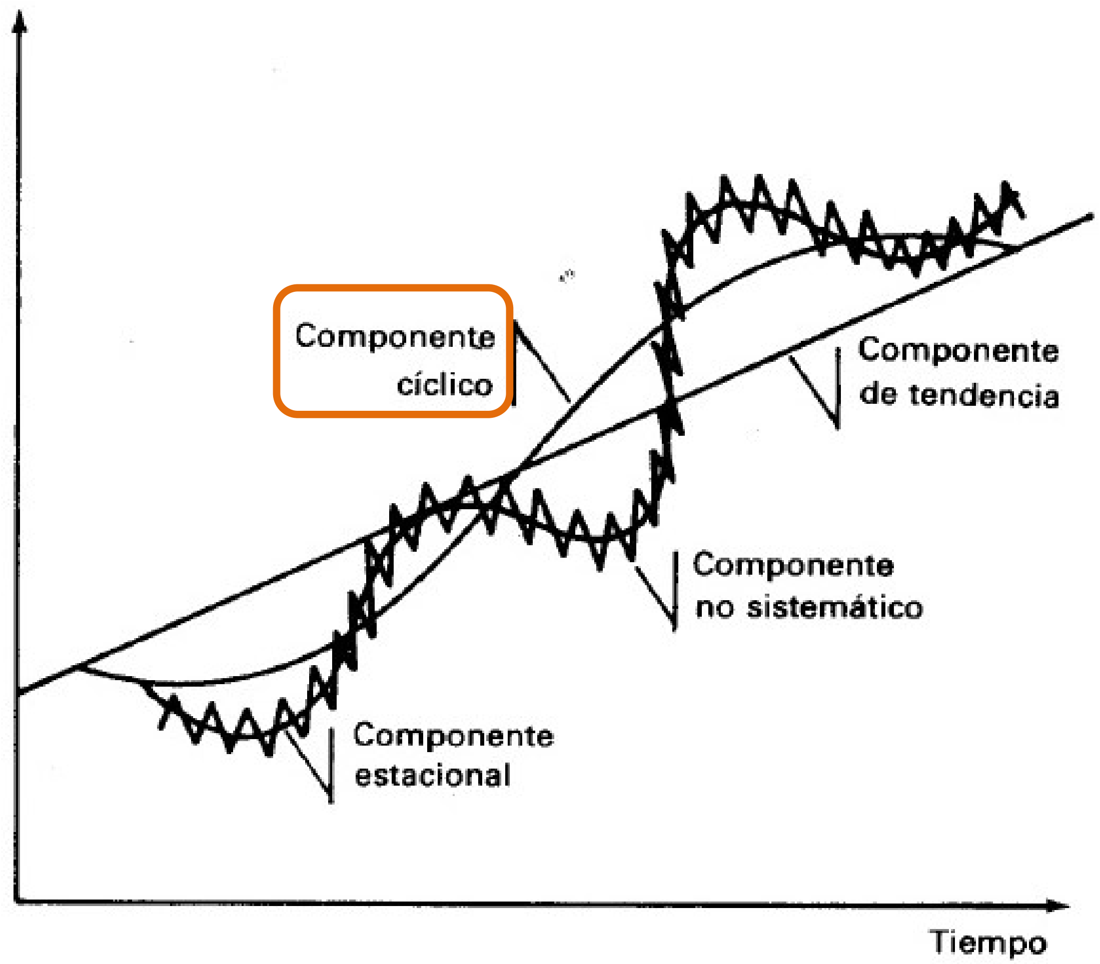
```
Holt-Winters exponential smoothing with trend and multiplicative seasonal component.

Call:
HoltWinters(x = Yv, seasonal = "multiplicative")

Smoothing parameters:
 alpha: 0.0464206
 beta : 1
 gamma: 0.8411737
```



2. ANÁLISIS DE COMPONENTES

VARIACIONES CÍCLICAS

- Es la que entraña más problemas. Esto por tres razones como mínimo.
  - Primero, no siempre existe.
  - Segundo hace falta mucha información, series muy largas, para que pueda detectarse la presencia de esta componente.
  - Tercero, es la menos sistemática de las tres, pues los ciclos, cuando existen, no siempre tienen la misma longitud y, además, se puede dar el caso de que se superpongan más de un ciclo de distintas longitudes de onda.



iii) Variaciones Cíclicas

- Para extraer y predecir el componente cíclico se utiliza el ajuste de funciones periódicas.
- Una función periódica es aquella que repite sus valores en el tiempo cada p periodos y puede venir expresada como:

$$Y_t = A.Cos\left(\frac{2\pi p}{N} + \theta\right)$$

Donde:

A : Amplitud de la oscilación

p : periodo

θ : desfase

N : número total de observaciones

- A efectos de ajustar y predecir series cíclicas podemos utilizar la expresión alternativa:

$$Y_t = \alpha.cos\left(\omega_0.p.t\right) + \beta.sen\left(\omega_0.p.t\right)$$

ω_0 : es lo que se denomina frecuencia básica y es igual a $\frac{2\pi}{N}$

iii) Variaciones Cíclicas

Paso 1: Identificar el número de máximos (mínimos) cíclicos p y construir las series:

$$COSP_t = cos\left(\frac{2\pi}{N}.p.t\right) \qquad SENP_t = sen\left(\frac{2\pi}{N}.p.t\right)$$

Paso 2: Ajustar mediante regresión el modelo:

$$Y_t = \beta_1 COSP_t + \beta_2 SENP_t$$

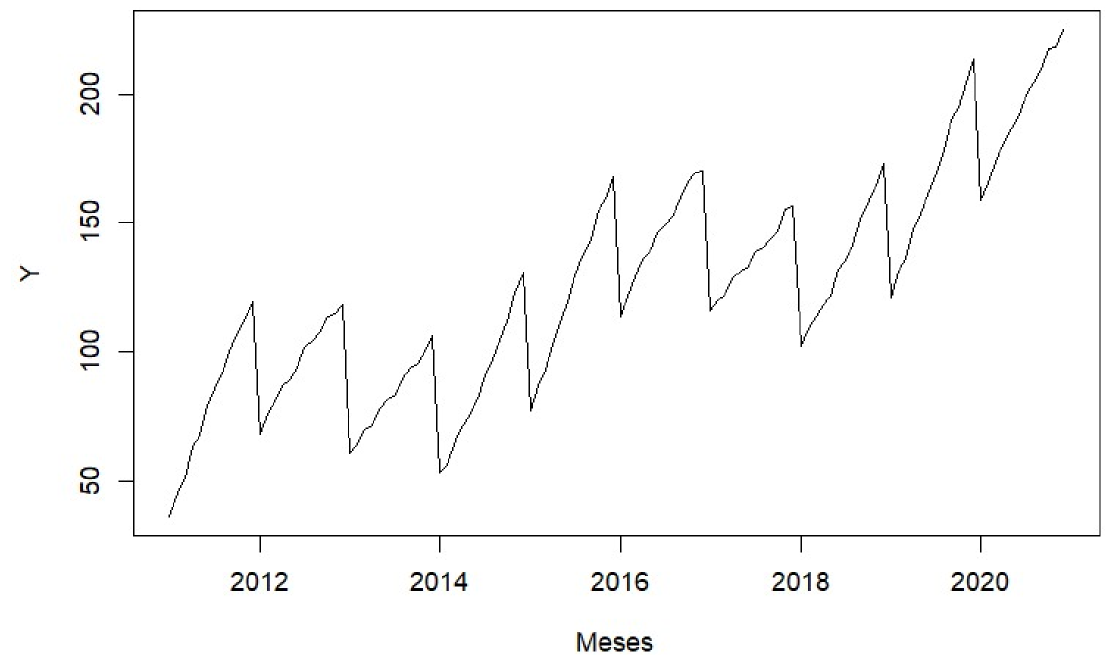
Paso 3: Calcular los errores (residuos) y si tienen comportamiento cíclico repetir el proceso añadiendo nuevos términos al modelo.

Para determinar el valor del periodo *p*, podríamos utilizar, como primera aproximación, el número de máximos (o mínimos) locales que presenta la serie a analizar.

Ejemplo 5

Para ejemplificar esta prueba trabajaremos con los datos del Ejemplo 2.1 del capítulo anterior, el cual presenta los valores de una serie de tiempo mensual en el periodo 2001: Enero- 2010: Diciembre, cuya gráfica de la serie se muestra en la Figura 2.1.

En los datos de esta serie debido a su longitud medianamente larga, existen sospechas de que pudiera existir la componente cíclica con un periodo aproximadamente de 5 años.

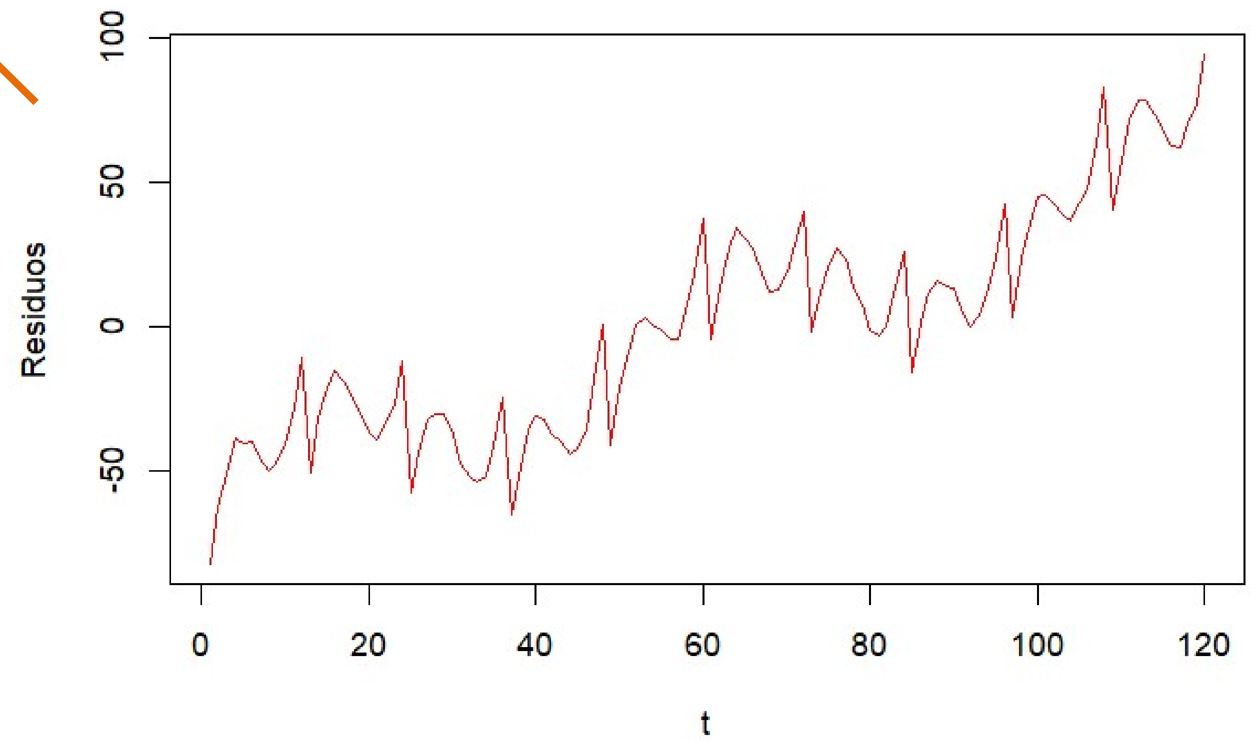


De la Figura podemos identificar 10 máximos cíclicos y generar las series que se indican en el paso 1.

Pero antes, debemos crear la serie T (la que indica el valor de t, toma valore de 1 a N = 120 datos) del siguiente modo:

Ejemplo 5

En esta figura podemos ver ahora claramente, que podría existir el componente cíclico con aproximadamente una amplitud de 4 años.

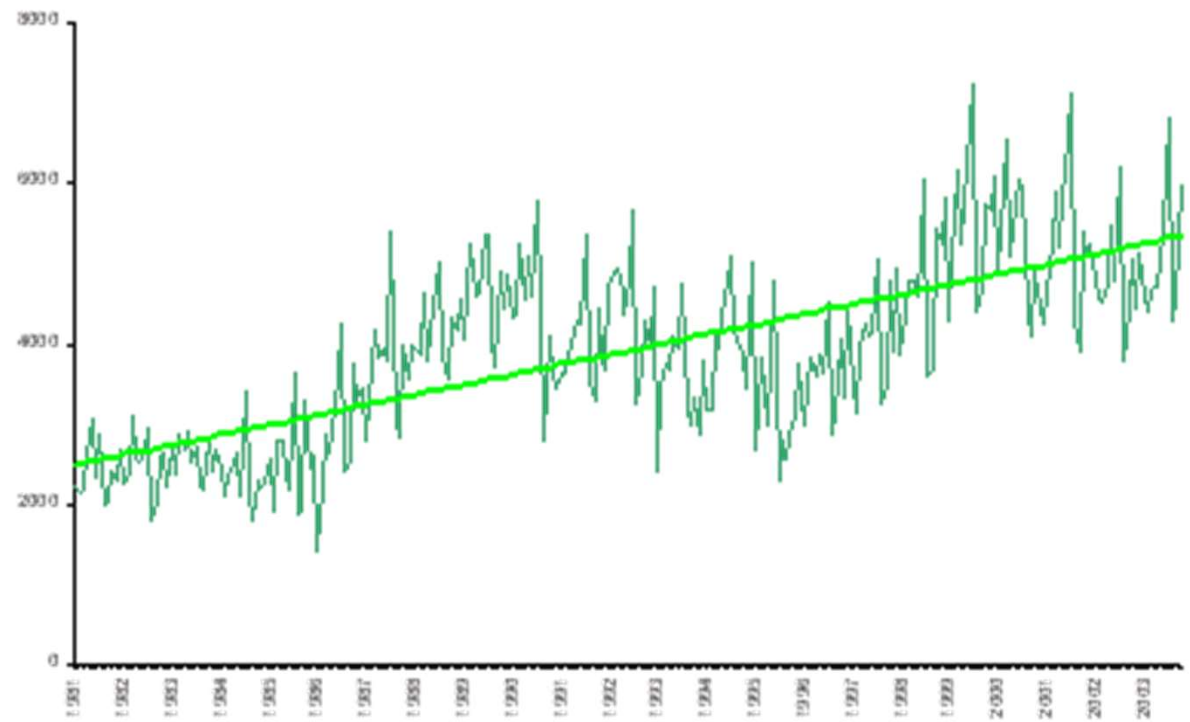



```
library(readxl)
serie <- read_excel("D:/.../Ejem_2_1.xlsx")
View(datos)
Yts<-ts(serie$Y, start = c(2011,1), frequency = 12)

# Generación de t
serie$t <- seq(1:NROW(serie))

# Construir series
cosP <- cos(2*pi/120*10*serie$t)
senP <- sin(2*pi/120*10*serie$t)

# Ajuste del modelo
ciclo <- lm(Yts ~ cosP + senP)
plot(ciclo$residuals, type = "l", xlab="t", ylab="Residuos", col = "red")
```

2. ANÁLISIS DE COMPONENTES

IRREGULARIDAD

- La última componente son las variaciones residuales.
- La misma tiene poco interés.
- La utilidad de esta última componente se basa en poder verificar si satisface ciertos supuestos o hipótesis, como el que sea realmente aleatorio.
- Mide la variabilidad de una serie cuando los demás componentes se han eliminado o no existen.

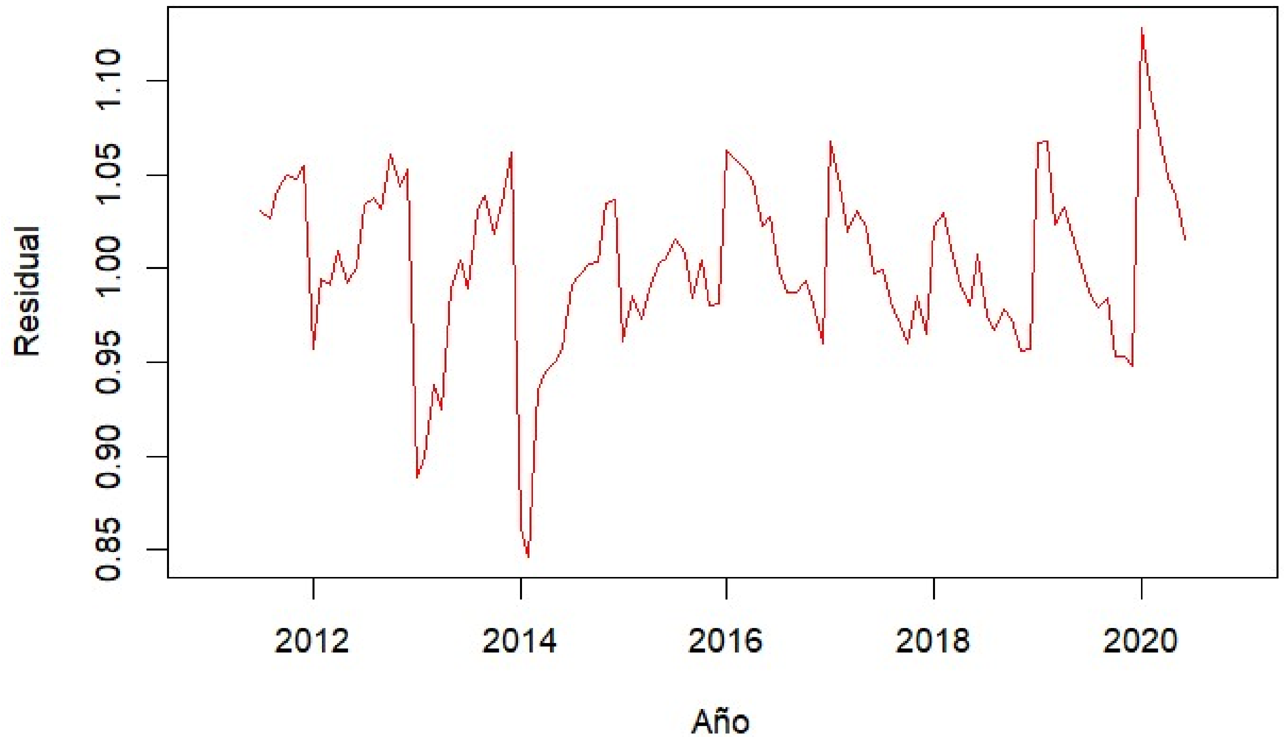


Ejemplo 6

Continuamos trabajando con los datos del Ejemplo 2.1 del capítulo anterior. Se puede graficar esta componente aleatoria y calcularse desde una serie de tiempo sin tratamiento previo.

Primero se descompone la serie original Yts en Ydesc, luego graficamos la componente aleatoria de la serie.

```
Ydesc <- decompose(Yts, type="multiplicative")
plot(Ydesc$random, type = "l", xlab="t", col = "red")
```




GRACIAS
https://aulavirtual2.unap.edu.pe/
2024-I