PRACTICA.R

VICTOR RAUL MAYE MAMANI

2024-05-20

```r
library(TTR) # Método de Media móvil
```

```
## Warning: package 'TTR' was built under R version 4.3.3
```

```r
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.1
```

```r
library(forecast)
```

```
## Warning: package 'forecast' was built under R version 4.3.3
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```r
library(fastDummies)
```

```
## Warning: package 'fastDummies' was built under R version 4.3.3
```

```
## Thank you for using fastDummies!
```

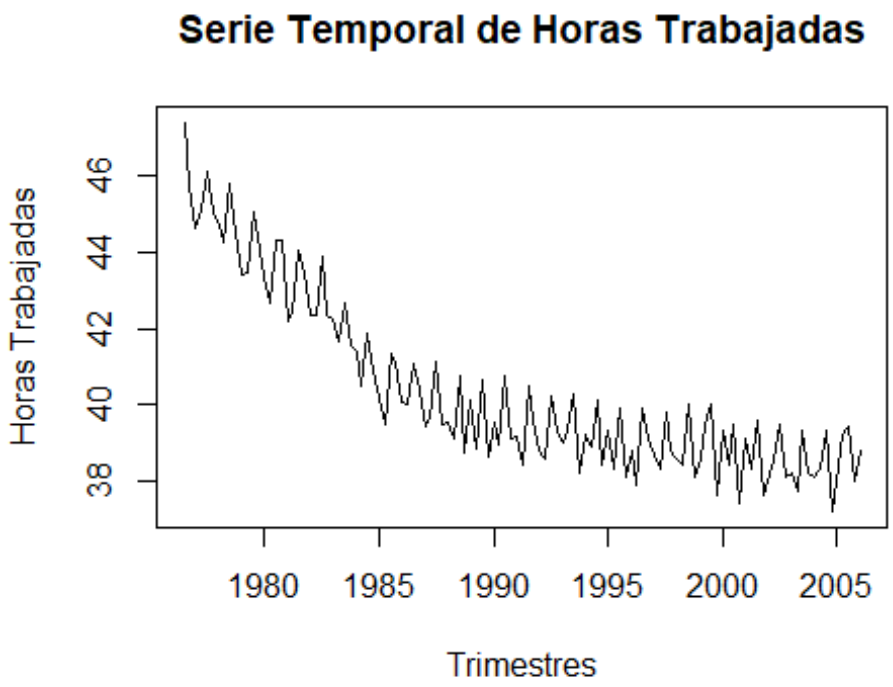
```
## To acknowledge our work, please cite the package:
```

```
## Kaplan, J. & Schlegel, B. (2023). fastDummies: Fast Creation of Dummy
(Binary) Columns and Rows from Categorical Variables. Version 1.7.1. URL:
https://github.com/jacobkap/fastDummies, https://jacobkap.github.io/fastD
ummies/.
```

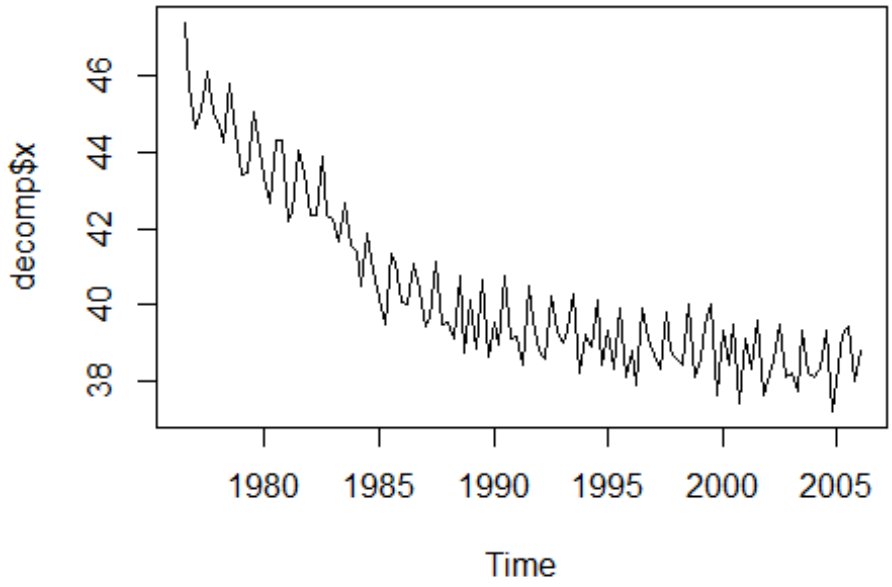
```r
library(mFilter)
```

```
## Warning: package 'mFilter' was built under R version 4.3.3
```

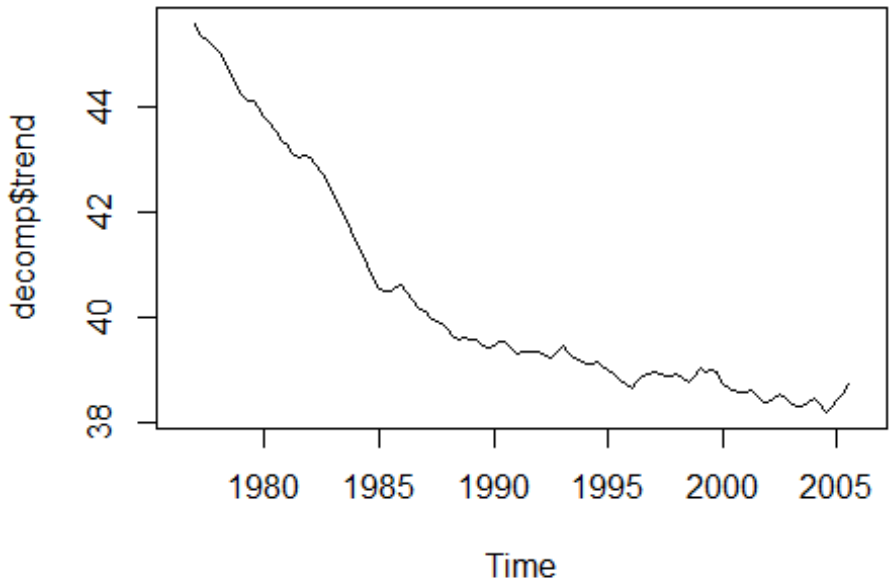
```r
ruta_archivo = 'C:\\Users\\LENOVO\\Downloads\\horastrabajadas.txt'
datos <- read.table(ruta_archivo, header = TRUE, sep = " ")

yts <- ts(datos$horastrabajadas, start = c(1976, 3), frequency = 4)

# a) Represente la serie e identifica las componentes de variación presen
tes en la misma.

plot(yts, main = "Serie Temporal de Horas Trabajadas", xlab = "Trimestres
", ylab = "Horas Trabajadas")
```



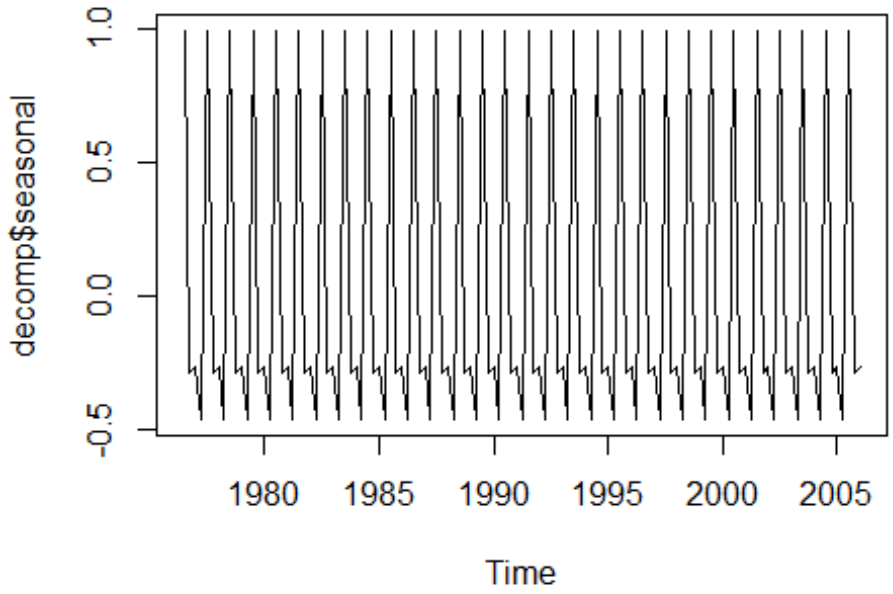

```
# LOS COMPONENTES DE VARIACION SE PUEDEN OBSERVAR EN LA PREGUNTA B INDIVIDUALMENTE

# b) Realice la descomposición de serie y represente de manera individual cada componente.

decomp = decompose(yts)
plot(decomp$x)
```



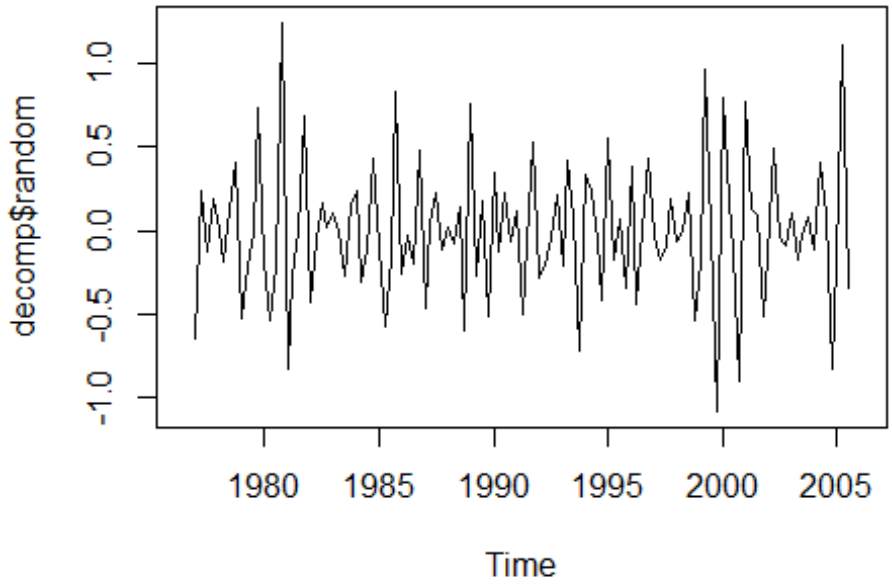

```r
# TENDENCIA
plot(decomp$trend)
```



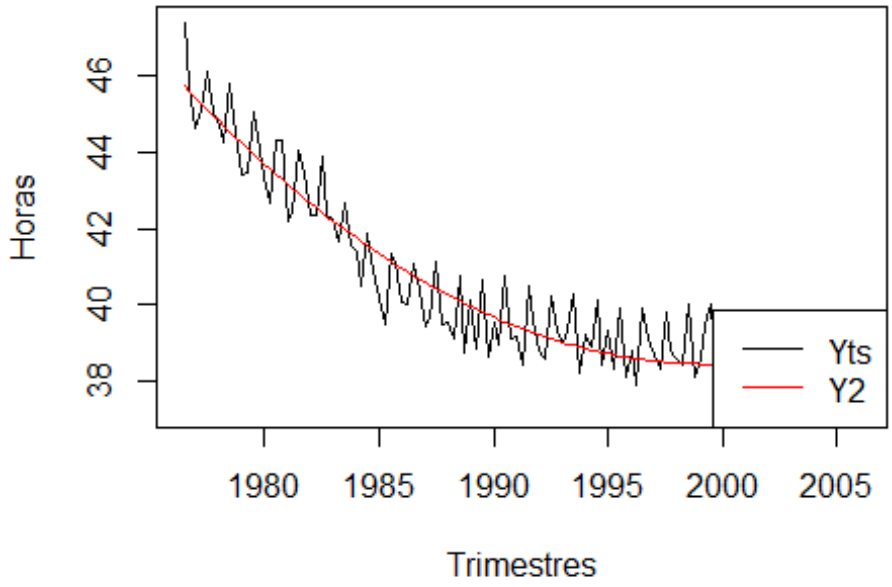
```r
# ESTACIONALIDAD
plot(decomp$seasonal)
```



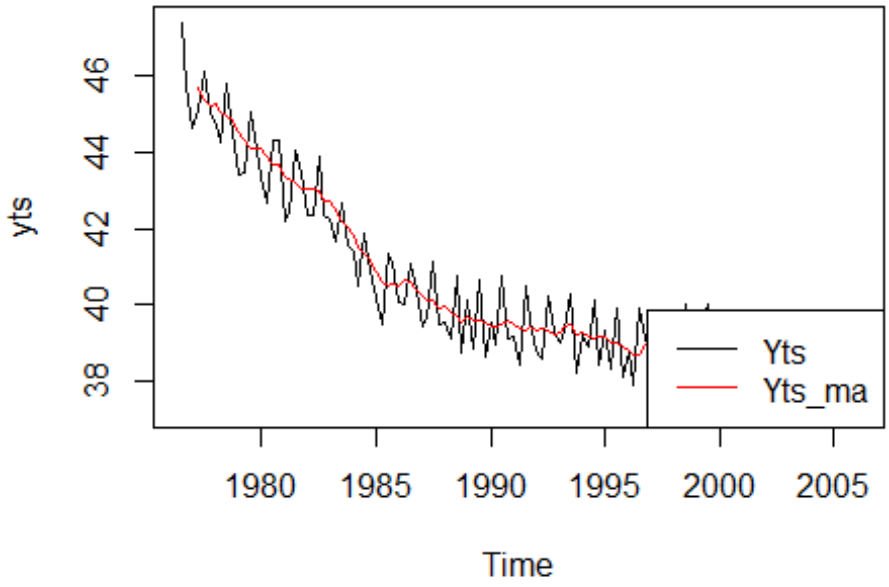

```r
# ALEATORIO
plot(decomp$random)
```



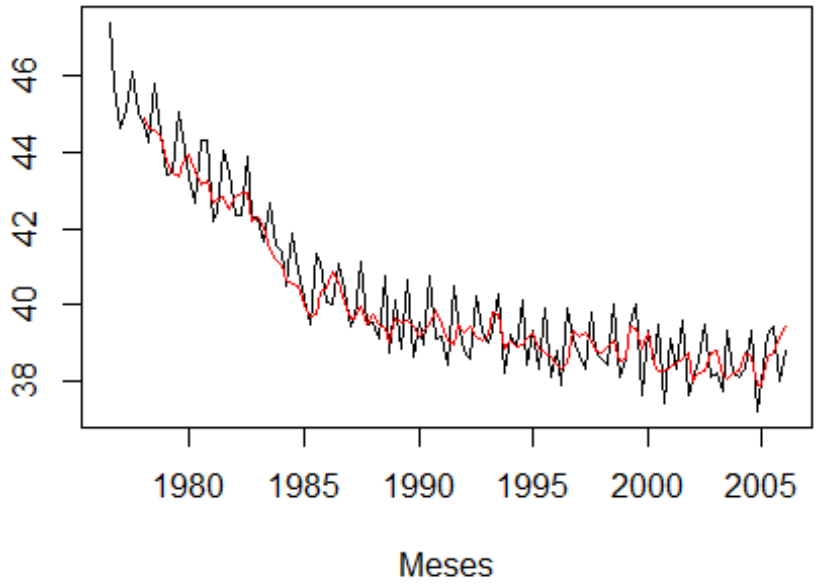

```r
# c) Estima la tendencia a través del ajuste de una función cuadrática, p
romedio móvil doble, el
# método de Holt-Winters y filtro de Hodrick - Prescott. ¿Qué método ajus
ta mejor la
# tendencia?, justifique su respuesta.


# Ajuste funcion cuadratica
t <- time(yts)
Y2 <- lm(yts ~ t + I(t^2))
pron = ts(predict(Y2, t), c(1976,3), frequency = 4)
plot(yts, xlab="Trimestres", ylab="Horas")
lines(pron, type = "l", col = "red")
legend(x = "bottomright", legend = c("Yts", "Y2"), col = c('black', 'red'
), lty = c(1, 1))
```



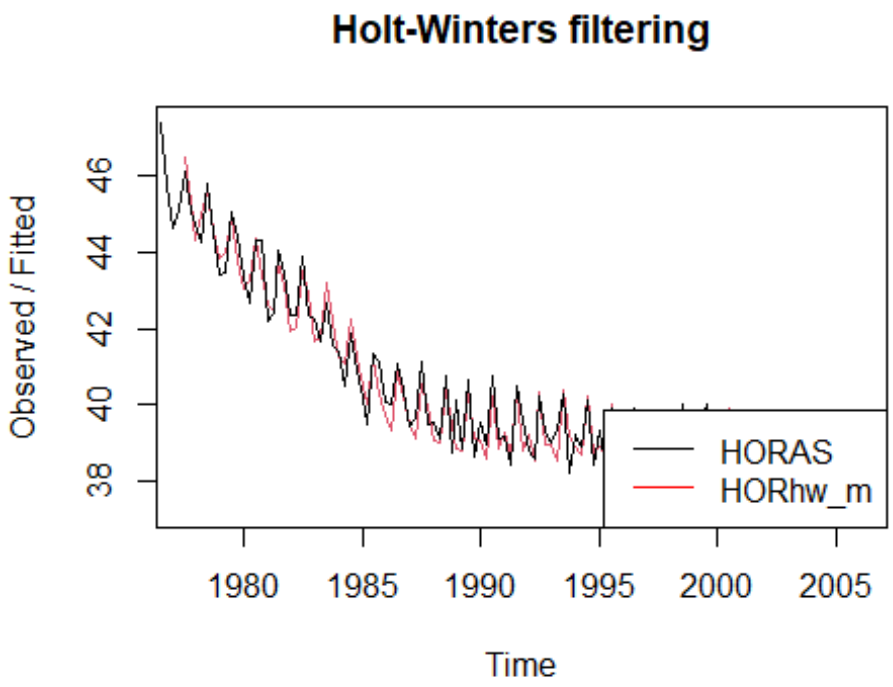

```
# Oservamos que la linea de ajuste de la funcion cuadratica se ajus
ta bien a la seri original


# Promedio móvil doble
k<-4 #Periodo
Yts_ma <- SMA(yts, k)
plot(yts)
lines(Yts_ma, type="l", col="red")
legend(x = "bottomright", legend =  c("Yts", "Yts_ma"), col = c("black","
red"), lty=c(1,1))
```




```
Yts_ma2 <- SMA(Yts_ma, k) # Media móvil de media móvil
a <- 2*Yts_ma - Yts_ma2
b <- (2/(k-1))*(Yts_ma - Yts_ma2)
p <- 2
Yma2 <- a + b*p
plot(yts, type='l', xlab="Meses", ylab=" ")
lines(Yma2, type='l', col = 'red')
```

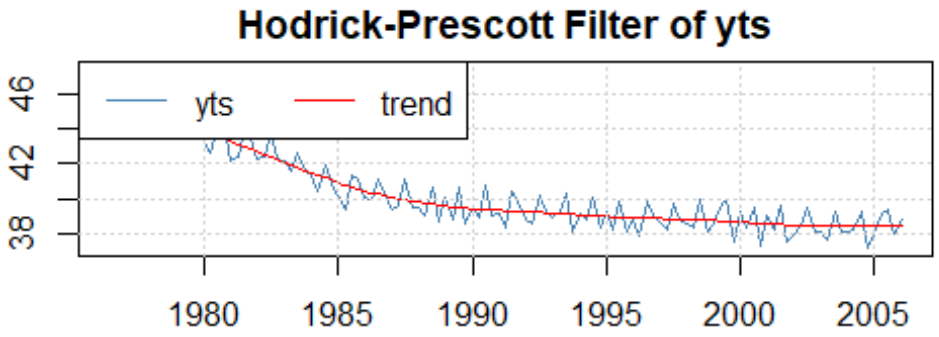


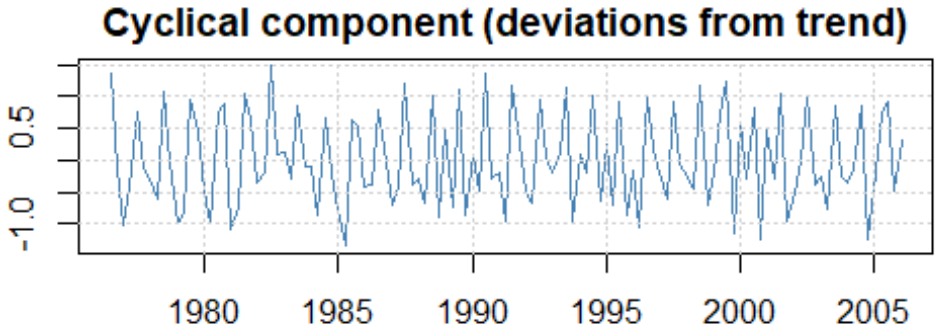

```
# El promedio movil doble tambien se ajusta, pero no podemos apreci
ar claramente la tendencia como en el funcion cuadratica

# Holt-Winters

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



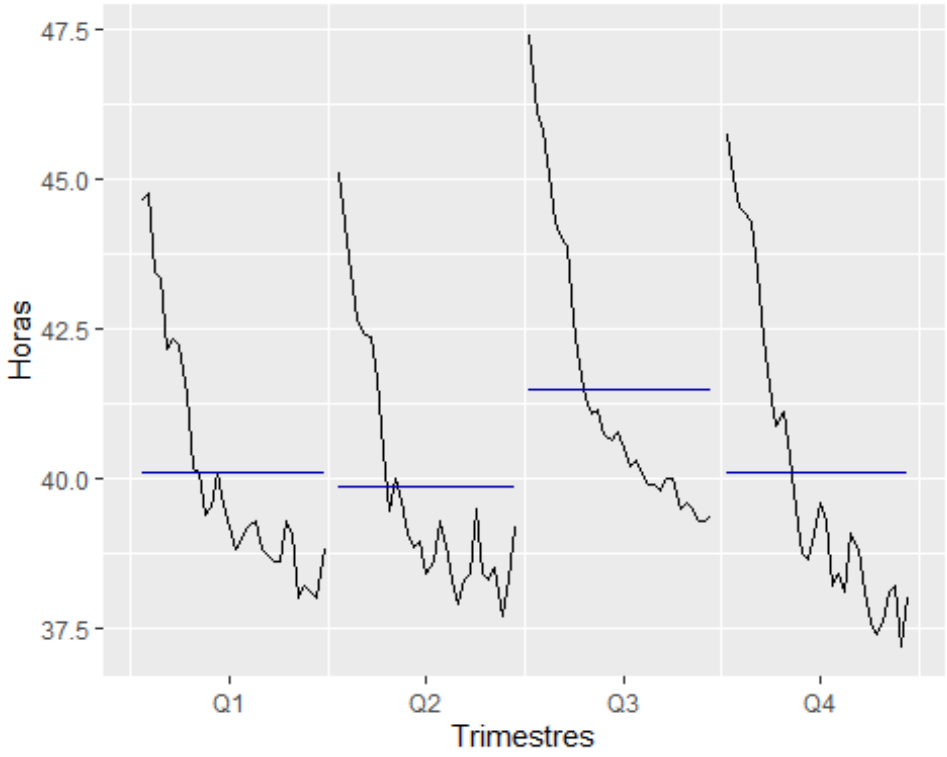
 # En este caso se observa que se ajusta demasiado a la serie original lo que dificulta la estimacion de la tendencia

```
# Filtro de Hodrick

lambda_hp <- 1600
Yhp <- hpfilter(yts, type="lambda", freq=lambda_hp)
plot(Yhp)
abline(h=0, col="green")
```

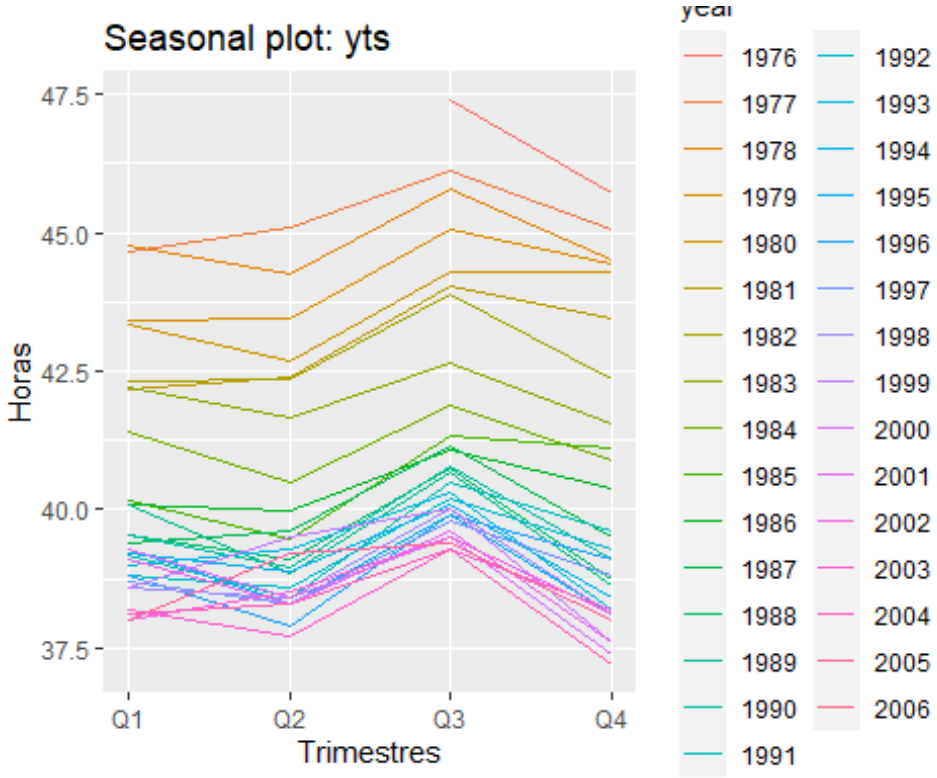





```r
# d) Detecte el componente estacional haciendo uso de los métodos gráfico
s. Para usted, ¿qué método gráfico es más efectivo?

ggsubseriesplot(yts, xlab = "Trimestres", ylab = "Horas")
```

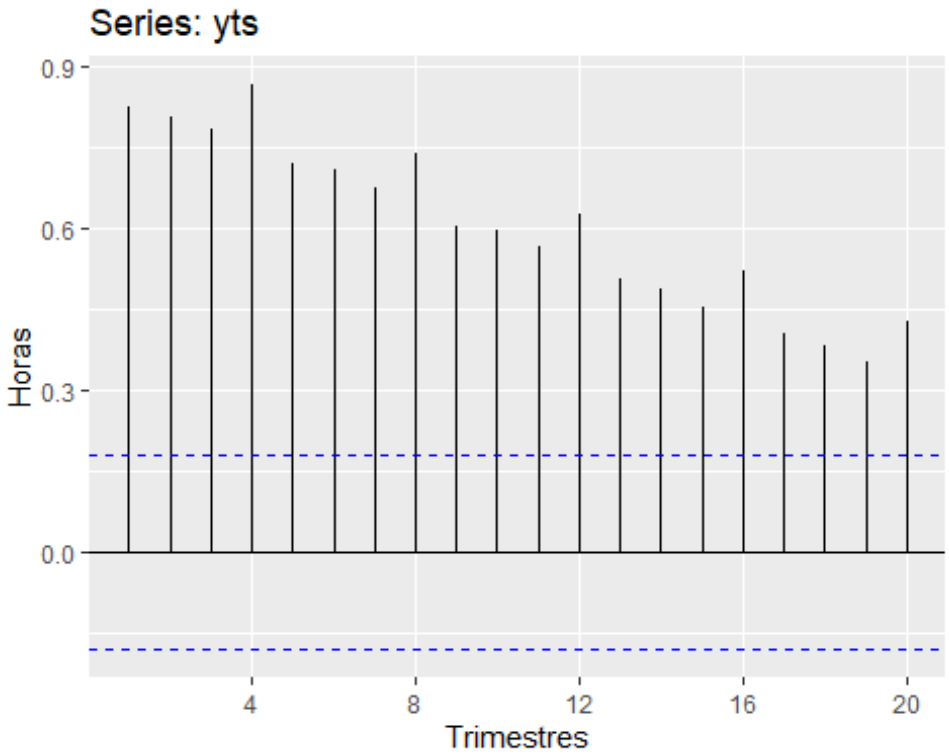



```
ggseasonplot(yts, xlab = "Trimestres", ylab = "Horas")
```




```
ggAcf(yts, xlab = "Trimestres", ylab = "Horas")
```

```
## Warning in ggplot2::geom_segment(lineend = "butt", ...): Ignoring unkn
own
## parameters: `xlab` and `ylab`
```




```
#Podemos observar que en los tres metodos no hay estacionalidad

# Siendo el mas efectivo el grafico de lineas ggseasonplot



# e) En caso de haber detectado el componente estacional, ratifique a tra
vés de la regresión de
# variables Dummy.
datos$trimestre = rep(1:4, length.out = 119)
datos$horas_1 <- lag(datos$horastrabajadas)
# Construcción de las Dummy
datos$dummy <- datos$trimestre
tabla_dummy <- dummy_columns(datos, select_columns ="dummy",remove_first_
dummy =TRUE)
head(tabla_dummy)

##   horastrabajadas trimestre horas_1 dummy dummy_2 dummy_3 dummy_4
## 1          47.395         1  47.395     1       0       0       0
## 2          45.730         2  45.730     2       1       0       0
## 3          44.647         3  44.647     3       0       1       0
## 4          45.087         4  45.087     4       0       0       1
```

```
## 5          46.100         1  46.100     1       0       0       0
## 6          45.045         2  45.045     2       1       0       0
```

```
# Dataframe para la regresion con Dummies
tabla_reg <- tabla_dummy[, c("horastrabajadas","horas_1","dummy_2","dummy
_3","dummy_4")]
# Regresion con Dummies
reg_1 <- lm(datos$horastrabajadas~., data = tabla_reg)
summary(reg_1)
```

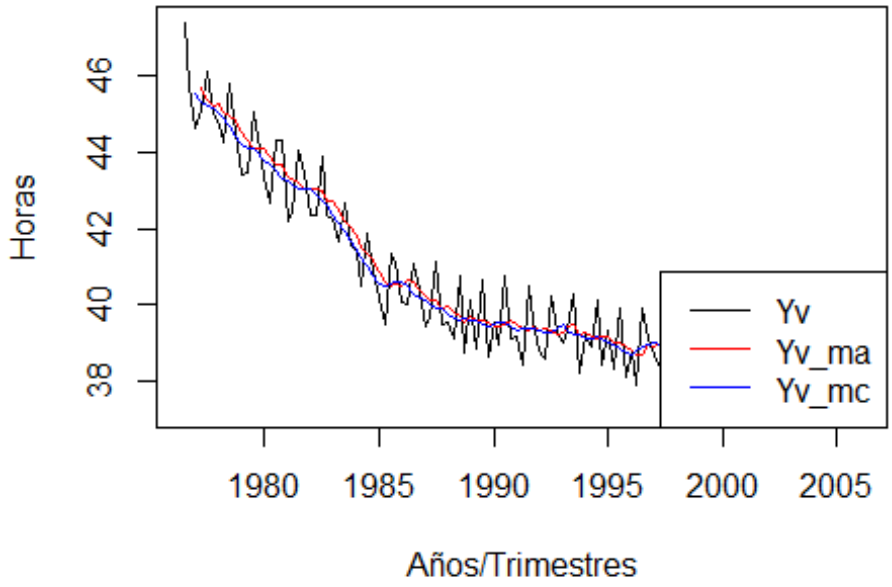
```
## Warning in summary.lm(reg_1): essentially perfect fit: summary may be
## unreliable
```

```
##
## Call:
## lm(formula = datos$horastrabajadas ~ ., data = tabla_reg)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -7.611e-16 -1.307e-16 -9.800e-18  6.070e-17  4.125e-15
##
## Coefficients:
##               Estimate Std. Error    t value Pr(>|t|)
## (Intercept) -1.042e-14  7.419e-16 -1.405e+01   <2e-16 ***
## horas_1      1.000e+00  1.779e-17  5.621e+16   <2e-16 ***
## dummy_2      2.697e-16  1.133e-16  2.381e+00   0.0189 *
## dummy_3      5.721e-17  1.134e-16  5.050e-01   0.6147
## dummy_4      6.170e-17  1.152e-16  5.360e-01   0.5933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.286e-16 on 114 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:      1
## F-statistic: 8.549e+32 on 4 and 114 DF,  p-value: < 2.2e-16
```

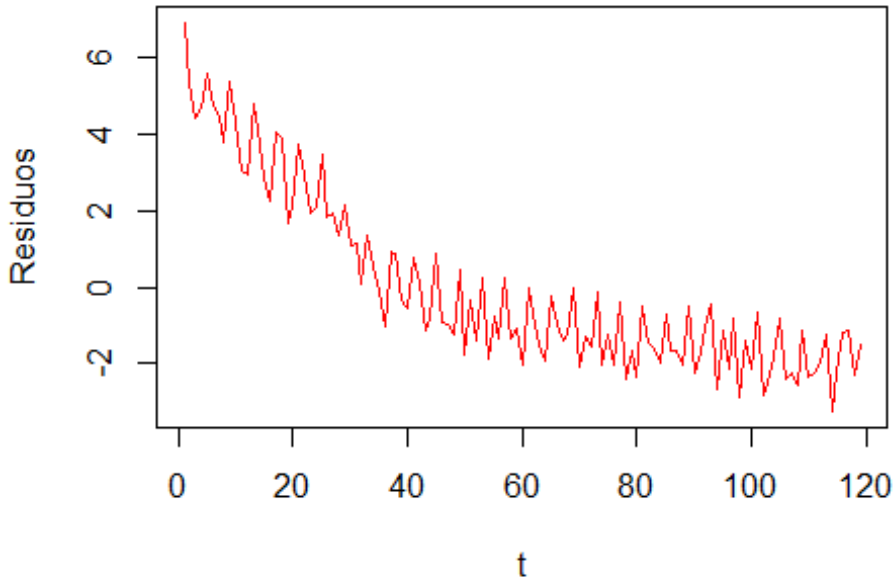
```
# Podemos observar que no todas las variables son significativas
```

```
# f) Desestacionalice la serie con el método de media móvil central.
```

```
Yv_ma <- SMA(yts, 4) # media móvil simple
Yv_mc <- ma(yts, 4) # media móvil centrada
plot(yts, xlab="Años/Trimestres", ylab="Horas")
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('bl
ack', 'red', 'blue'),lty=c(1,1,1))
```




```r
# g) Evalúe la existencia de las variaciones cíclicas mediante el ajuste regresión:

datos$t <- seq(1:NROW(datos))
# Construir series
cosP <- cos(2*pi/119*34*datos$t)
senP <- sin(2*pi/119*34*datos$t)
# Ajuste del modelo
ciclo <- lm(yts ~ cosP + senP)
plot(ciclo$residuals, type = "l", xlab="t", ylab="Residuos", col = "red")
```




```
# Podemos apreciar que si hay un componente ciclico con una amplitud de a
proximadamente 5 trimestres
```