Series_de_Tiempo_Primer_Parcial_Ibarra_Sergio

Sergibar

2023-04-15

Universidad Nacional Autónoma de México

Posgrado en Ingeniería de Sistemas

Series de tiempo Dr. Wulfrano Gomez Gallardo

Primer Parcial: Aplicaicón de métodos de suavizamiento esáticos y de HoltWinters para el tratamiento y pronóstico de series de tiempo (3 casos aplicativos)

Serie de tiempo 1 SIN tendencia y SIN estacionalidad: Historico anual Accidentes terrestres en México

Tenemos data anual desde 2005 a 2021 sobre el número de accidentes terrestres que ha habido en México, son 22 datos en total. La idea es usar 19 de esos 22 como una especie de "data de entrenamiento" para aplicar los modelos A) De suavizamiento "estático" y B) De "suavizamiento dinámico" o Holt-Winters para posteriormente predecir con cada modelo los valores de los 3 años restantes y poder comparar contra los reales e identificar el mejor modelo entre los dos mencionados para modelar los accidentes terrestres en México

Primero importamos la data de Accidentes terrestres en MEX Data obtenida de la página oficial del INEGI: https://www.inegi.org.mx/temas/accidentes/

```
library("readxl")
```

```
## Warning: package 'readxl' was built under R version 4.2.3
```

```
Accidentes_transito_MEX <-read_excel("Accidentes_transito_pp_.xlsx")
```

Revisamos que la data de nacimientos MEX sehay importado correctamente

```
Accidentes_transito_MEX
```

```
## # A tibble: 19 x 2
##    Date                  Num
```

```
##    <dttm>               <dbl>
##  1 2000-01-12 00:00:00 331938
##  2 2001-01-12 00:00:00 364869
##  3 2002-01-12 00:00:00 399002
##  4 2003-01-12 00:00:00 403940
##  5 2004-01-12 00:00:00 424940
##  6 2005-01-12 00:00:00 434940
##  7 2006-01-12 00:00:00 444940
##  8 2007-01-12 00:00:00 454940
##  9 2008-01-12 00:00:00 436435
## 10 2009-01-12 00:00:00 418467
## 11 2010-01-12 00:00:00 410267
## 12 2011-01-12 00:00:00 397185
## 13 2012-01-12 00:00:00 380411
## 14 2013-01-12 00:00:00 365772
## 15 2014-01-12 00:00:00 362574
## 16 2015-01-12 00:00:00 360573
## 17 2016-01-12 00:00:00 360051
## 18 2017-01-12 00:00:00 357789
## 19 2018-01-12 00:00:00 355281
```

```r
summary(Accidentes_transito_MEX)
```

```
##       Date                              Num
##  Min.   :2000-01-12 00:00:00.00   Min.   :331938
##  1st Qu.:2004-07-13 00:00:00.00   1st Qu.:361574
##  Median :2009-01-12 00:00:00.00   Median :397185
##  Mean   :2009-01-11 15:09:28.42   Mean   :392859
##  3rd Qu.:2013-07-13 12:00:00.00   3rd Qu.:421704
##  Max.   :2018-01-12 00:00:00.00   Max.   :454940
```

```r
typeof(Accidentes_transito_MEX)
```

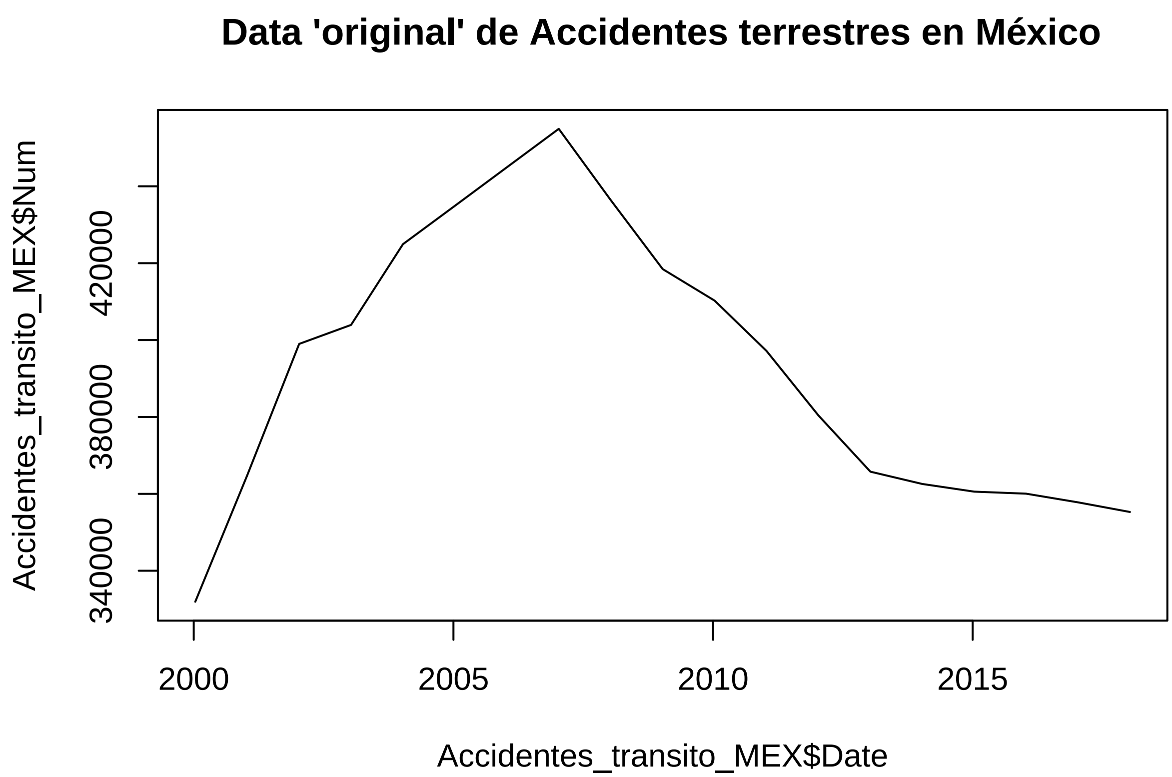
```
## [1] "list"
```

```r
dim(Accidentes_transito_MEX)
```

```
## [1] 19  2
```

Graficamos la "data original" de Accidentes terrestres en México de los últimos años De 2000 a 2018)

```r
plot(Accidentes_transito_MEX$Date, Accidentes_transito_MEX$Num, type="l", main="Data 'original' de Acci
```



A. Método de "suavizamiento estático"

Comenzaremos tratando la serie con el método estático aplicando el siguiente flujo lógico de trabajo y considerando que estamos en el caso de serie CON TENDECIA y SIN ESTACIONALIDAD:

1. Se llevará a cabo el "pronóstico puntual" de los siguientes 3 años como la media historica
2. Se calcularán los intervalos de pronóstico al 95% de confianza
3. Se calculará el error tipo MAPE entre los valores pronósticados con el "método estático" y los valores reales de la serie de los últimos 6 trimestres

A.1 Tratamiendo del ruido blanco en la serie y pronostico puntual

Considerese el siguiente razonamiento acerca de las series del tiempo: Serie total = Componente de tendencia + Componente de estacionalidad + Componente estocástico

En este caso el Componente de estacionalidad = 0 y el Componente de tendencia = 0, por lo tanto tendríamos

Serie total Accidentes MEX = Componente estocástico

Calculamos el "ruido blanco de la serie" como la media de los datos hisóricos

```
Accidentes_transito_ruido_blanco <- mean(Accidentes_transito_MEX$Num)
Accidentes_transito_ruido_blanco
```

```
## [1] 392858.6
```

Una vez teniendo únicamente el componente estocástico de nuestra serie, es posible "extraer los errores" de la serie y conocer así el error estimado (varianza) y la desviación estandar. La idea es que extrayendo de la serie del ruido blanco su media histórica estaríamos determinando la varianza, de la siguiente manera:

Varianza (errores al cuadrado) = Serie original - media de componete estocástico.

Errores= Valores originales de la serie - Media del componente estocástico o ruido blanco

Se determinan la varianza (errores al cuadrado) de la serie de PEA_MEX y se calcula también la desviación estandar de esos errores

```
Error_est_accidentes<-Accidentes_transito_MEX$Num-Accidentes_transito_ruido_blanco
head(Error_est_accidentes)
```

```
## [1] -60920.632 -27989.632   6143.368  11081.368  32081.368  42081.368
```

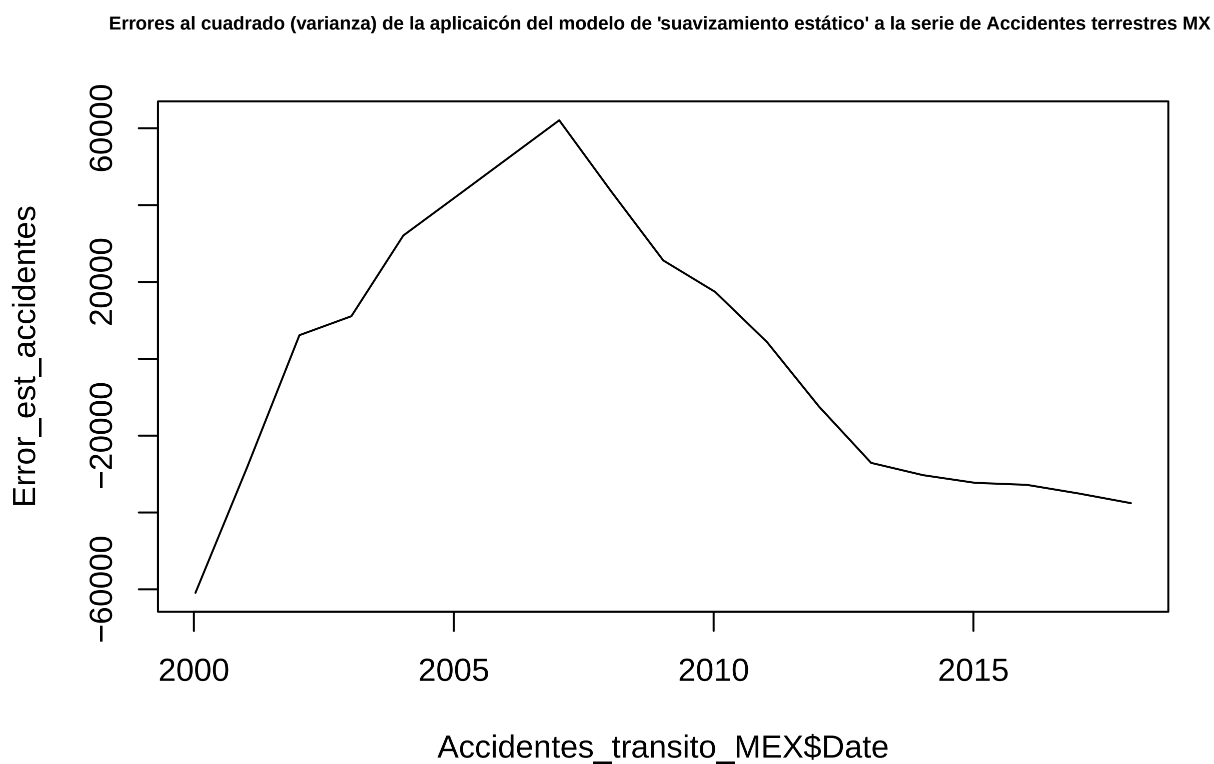
```
tail(Error_est_accidentes)
```

```
## [1] -27086.63 -30284.63 -32285.63 -32807.63 -35069.63 -37577.63
```

```
##Claculando la desviaicón estandar del error

sd_est_accidentes<-sd(Error_est_accidentes)
sd_est_accidentes
```

```
## [1] 36105.41
```

```
plot(Accidentes_transito_MEX$Date, Error_est_accidentes, type="l", main="Errores al cuadrado (varianza)
```



Se calcula la media de los errores

```r
mean(Error_est_accidentes)
```

```
## [1] 9.188813e-12
```

Se calcula la varianza y desviación estandar de los errores

```r
## Se calcula la varianza
var(Error_est_accidentes)
```

```
## [1] 1303600384
```

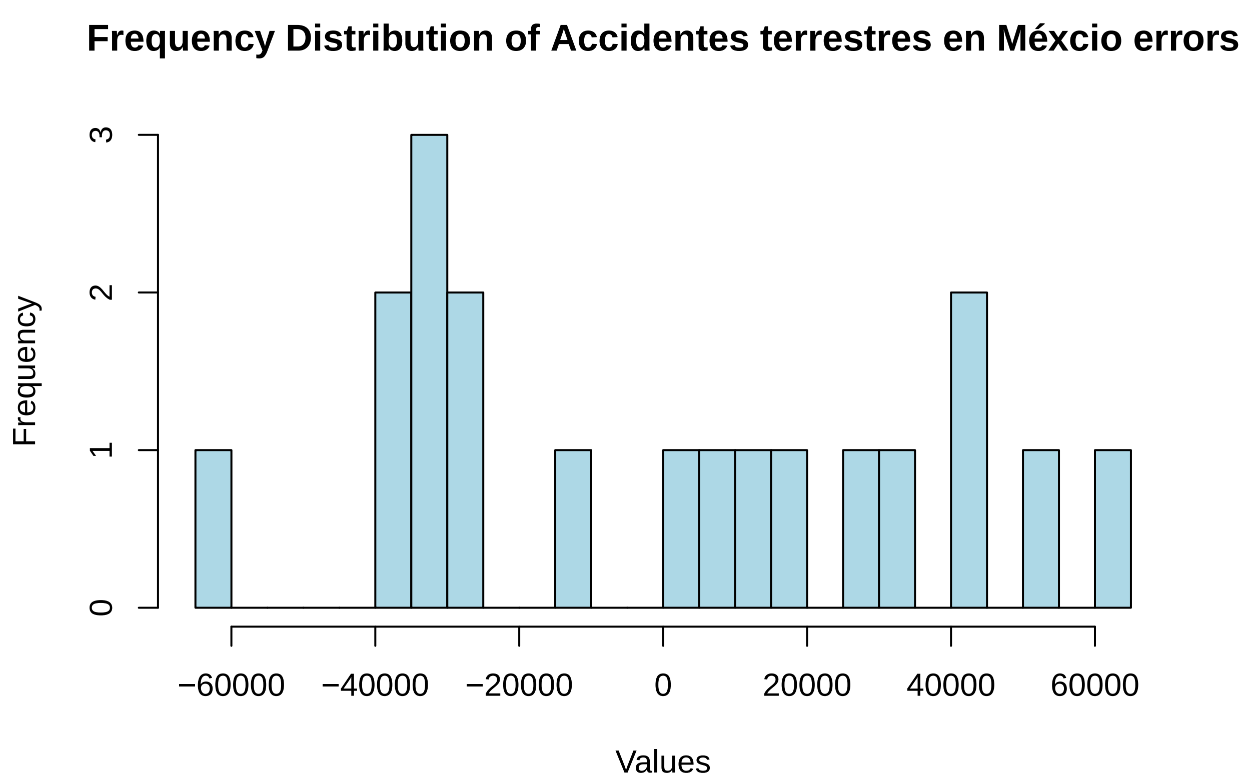
```r
## Se calcula la desviación estandar
sd_est_accidentes
```

```
## [1] 36105.41
```

Se tiene una desviación estandar de ~37 mil y una media muy cercana a cero para los errores lo cual es totalmente concordante con le hecho de que "el ruido blanco" o componente estocástico tiende a tener una distribuión normal y que si a ese ruido blanco se le extrae su media tambien tiende a tener una distribución normal con media cero y en este caso una desviación estandar de ~37 mil que representa aproximadamente el 10% del valor de la variable a pronosticar y: Número de accidentes terrestres anuales en México.

Procedamos a observar como se comporta la Distribución de dichos errores:

```r
hist(Error_est_accidentes, breaks = 20, col = "lightblue", main = "Frequency Distribution of Accidentes
```



Se observa como los errores de la serie se acercan a una distribución normal con media cero y desviación estandar de ~35 mil, a excepción de los valores atípicos del año 2020 derivados de la pandemia de Covid-19.

Pronóstico puntual Recordemos que para este caso donde no hay componente de tendencia ni estacionalidad y el valor de la serie es directamente el ruido blanco. Entonces, solo por cuestión de concordancia definamos a nuestra "mejor estimaicón puntual" como la media histórica de mi serie

```
estimacion_puntual_accidentes <- mean(Accidentes_transito_MEX$Num)
estimacion_puntual_accidentes
```

```
## [1] 392858.6
```

Vamos a hacer el pronóstico de la serie de PEA MEX para los próximos 3 años

```
##Creamos un eje x que incluya a los 66 datos + 6 que se pronosticarán
Date_accidentes2<-c(1:22)


## Creamos el vector de datos que incluirá los  datos históricos  sin tendencia + 6 valores NA a pronos
accidentes_para_pron<-c(Accidentes_transito_MEX$Num, rep(NA, times=3))


head(accidentes_para_pron)
```

```
## [1] 331938 364869 399002 403940 424940 434940
```

```
tail(accidentes_para_pron)
```

```
## [1] 360051 357789 355281     NA     NA     NA
```

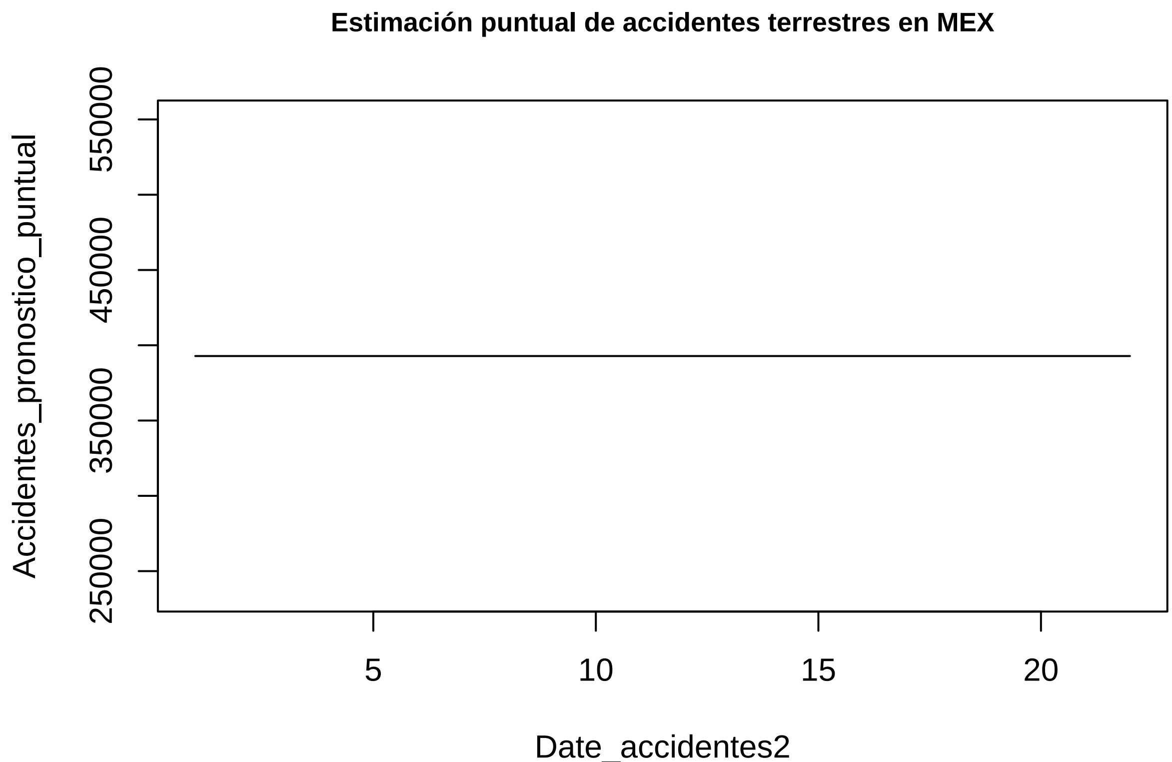
```
## Entonces el mejor pronóstico que tenemos para esta técnica es la media de la serie historica


Accidentes_pronostico_puntual <- rep(estimacion_puntual_accidentes, times=22)
head(Accidentes_pronostico_puntual)
```

```
## [1] 392858.6 392858.6 392858.6 392858.6 392858.6 392858.6
```

```
tail(Accidentes_pronostico_puntual)
```

```
## [1] 392858.6 392858.6 392858.6 392858.6 392858.6 392858.6
```

```
plot(Date_accidentes2,Accidentes_pronostico_puntual, type="l", main = "Estimación puntual de accidentes
```



Comentario sobre el pronóstico puntual sin tendencia ni estacionalidad, que para el caso de la serie de esta serie tiene un valor de:

```r
estimacion_puntual_accidentes <- mean(Accidentes_transito_MEX$Num)
estimacion_puntual_accidentes
```

```
## [1] 392858.6
```

Este valor es bastante concordante con los datos historicos, ya que en general el rango de accidentes se encuentra entre 330 mil y 440 mil, por lo que ese valor de "estimación puntual" sueva bastante razonable

A.2 Calculo de intervalos de confianza para el pronóstico

Calculamos los límites de intervalos de confianza al 95% pata el pronostico Se considera 95% como +- 2 desviaciones estandar de la media

```r
tao<-c(1:3)
Lim_accidentes<-2*sd_est_accidentes*tao^.5
Lim_accidentes
```

```
## [1]  72210.81 102121.51 125072.80
```

Calculamos los limites inferior y superiror como NA para los valores historicos y media +- 2 veces desciación estandar para los valores pronósticados

Graficamos los valores pronósticadoos t los limites de confianza

```r
LI_accidentes<-c(rep(NA, times=19), Accidentes_pronostico_puntual[1]-Lim_accidentes)
head(LI_accidentes)
```

```
## [1] NA NA NA NA NA NA
```

```r
tail(LI_accidentes)
```

```
## [1]       NA       NA       NA 320647.8 290737.1 267785.8
```

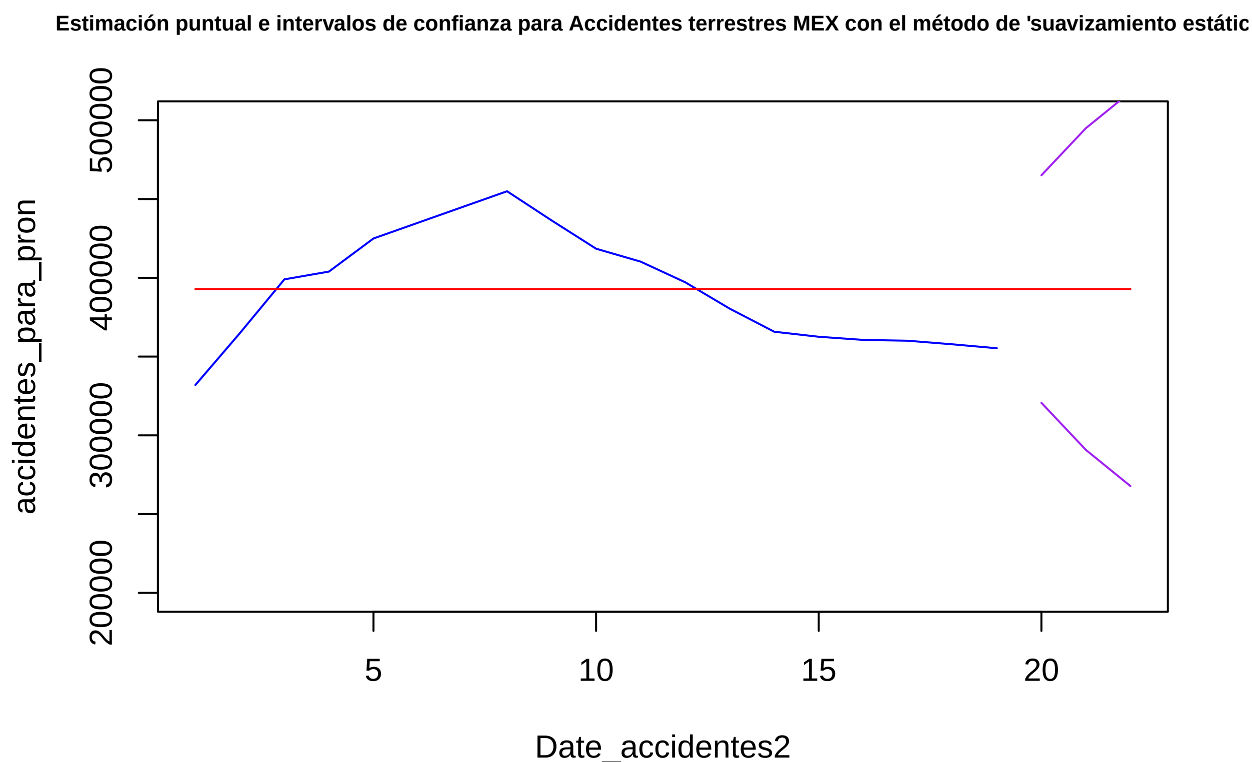
```r
LS_accidentes<-c(rep(NA, times=19), Accidentes_pronostico_puntual[1]+Lim_accidentes)
head(LS_accidentes)
```

```
## [1] NA NA NA NA NA NA
```

```r
tail(LS_accidentes)
```

```
## [1]       NA       NA       NA 465069.4 494980.1 517931.4
```

```r
plot(Date_accidentes2, accidentes_para_pron, type="l" ,main="Estimación puntual e intervalos de confian

lines(Accidentes_pronostico_puntual, col="Red")
lines(LI_accidentes, col="Purple")
lines(LS_accidentes, col="Purple")
```



```r
library("readxl")


Accidentes_transito_MEX_completo <-read_excel("Accidentes_transito_full_.xlsx")
```

Visualizamos el Data.Frame con los valores históricos la "data original" + los resultados del pronóstico puntual y los intervalos de confianza al 95%

```r
df_Accidentes_MX_pronostico<-data.frame(Accidentes_transito_MEX_completo$Num,
                                                Accidentes_pronostico_puntual, LI_accidentes, LS_acciden
View(df_Accidentes_MX_pronostico)

tail(df_Accidentes_MX_pronostico)
```

```
##    Accidentes_transito_MEX_completo.Num Accidentes_pronostico_puntual
## 17                               360051                      392858.6
## 18                               357789                      392858.6
## 19                               355281                      392858.6
## 20                               352729                      392858.6
## 21                               315068                      392858.6
## 22                               340415                      392858.6
##    LI_accidentes LS_accidentes
## 17            NA            NA
## 18            NA            NA
## 19            NA            NA
## 20      320647.8      465069.4
## 21      290737.1      494980.1
## 22      267785.8      517931.4
```

Se observa que los pronósticos puntuales vs los reales parecen no estar muy alejados unos de otros, sin embargo los intervalos de confianza si "sufren DEMASIADO a medida que nos alejamos del último valor real" pues son cada vez más grandes hasta llegar a abarcar un rango de [~260 mil a 517 mil] para el 95% de intervalo de confianza apenas en el 3er valor pronósticado.

A.5 Calculo del error tipo MAPE para los pronósticos hechos con los modelos estáticos de suavizamiento

Calculemos el error tipo MAPE entre los valores pronosticados con el "método de suavizamiento estático" y los valores reales de la serie

Para ello importamos el Data Frame con la data historica completa, es decir, la que contiene los 66 trimestres propuestos aquí como "train data" + los 6 trimestres que se pronósticaran

```r
##Primero formemos un vector con los valores reales
accidentes_MEX_completa_reales3 <- Accidentes_transito_MEX_completo$Num[19:22]
accidentes_MEX_completa_reales3
```

```
## [1] 355281 352729 315068 340415
```

```r
## Despues formemos un vector con únicamente los valores pronósticados
accidentes_MEX_forecast3 <-  Accidentes_pronostico_puntual[19:22]
accidentes_MEX_forecast3
```

```
## [1] 392858.6 392858.6 392858.6 392858.6
```

```r
##Calculemos el MAPE entre estos dos vectores

# First, calculate the absolute percentage error (APE) for each element in the vectors
APE_accidentes_estatico <- abs((accidentes_MEX_completa_reales3 - accidentes_MEX_forecast3) / accidente

# Calculate the mean APE across all elements in the vectors
MAPE_accidentes_estatico <- mean(APE_accidentes_estatico) * 100

# Print the MAPE value
cat("El error tipo MAPE  para el forecast calculado con el método de 'suavizamiento estático'de la seri
```

```
## El error tipo MAPE  para el forecast calculado con el método de 'suavizamiento estático'de la serie
```

B. Método de Holt como "modelo de suaviazamiento y pronóstico NO ESTÁTICO" que si considera los cambios de nivel a tráves del tiempo.

Vamos a usar ahora el método de HoltWinters como "modelo de suaviazamiento y pronóstico NO ESTÁTICO" que si considera los cambios de nivel a tráves del tiempo.

Para ello debemos tener nuestra data a suavizar en el tipo de dato serie de tiempo ts en R

```r
acidentes_MEX_ts <-ts(Accidentes_transito_MEX$Num, frequency =1, start =c(2000,1))
head(acidentes_MEX_ts)
```

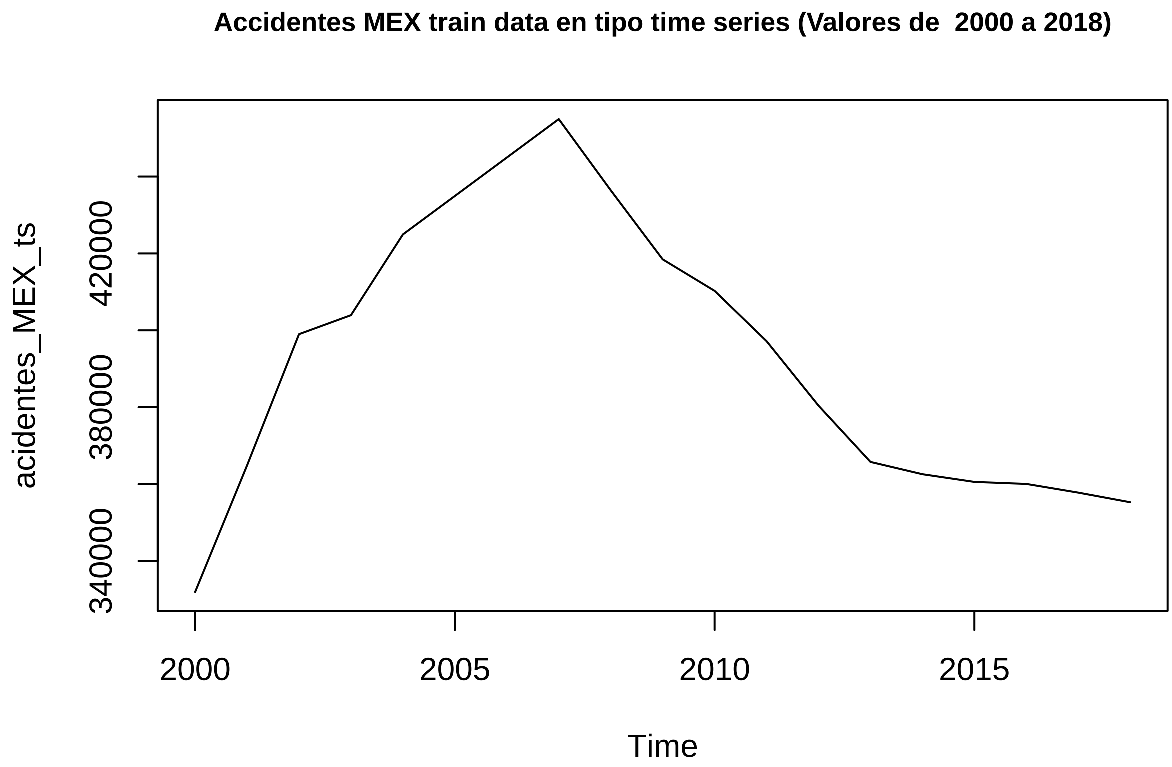
```
## [1] 331938 364869 399002 403940 424940 434940
```

```r
tail(acidentes_MEX_ts)
```

```
## [1] 365772 362574 360573 360051 357789 355281
```

Graficamos la "data original de la PEA_MEX" que se usará como "train para el modelo" y que fue convertida a tipo de dato serie de tiempo

```r
plot(acidentes_MEX_ts, main="Accidentes MEX train data en tipo time series (Valores de  2000 a 2018)",
```



Aplicación del método de Holt Winters" a la serie SIN TENDENCIA y SIN ESTACIONALIDAD de Accidentes terrestres en MX

Como se observará este primer cálculo dejará como grados de libertad los valores de alpha, así como los valores de arranque, de manera que el "el algoritmo de HoltWinters de R" determine la "mejor combinaión posible". Como en este caso no tenemos componente de seasonalidad ni ne tendencia se indica que el parámetro beta y gamma tendrán un valor "false"

```
accidentes_MEX_ts_h1 <-HoltWinters(acidentes_MEX_ts, beta=FALSE, gamma=FALSE)
accidentes_MEX_ts_h1
```

```
## Holt-Winters exponential smoothing without trend and without seasonal component.
##
## Call:
## HoltWinters(x = acidentes_MEX_ts, beta = FALSE, gamma = FALSE)
##
## Smoothing parameters:
##  alpha: 0.9999403
##  beta : FALSE
##  gamma: FALSE
##
## Coefficients:
##       [,1]
## a 355281.1
```

Interpretación de los valores de los coeficientes del método de Holt Winters para el caso de la serie de accidentes terrestres México Interpretación de los valores de coficientes α y β para el caso del pronóstico usando modelo de HoltWinters SIN ESPECIFICAR DATOS DE ARRANQUE

$$0 < \alpha < 1$$

Recordemos que: Valores de α cercano a 1 reflejan caminata aleatoria Valores de α cercano a 0 reflejan una "gran influencia" de los valores pasado, es decir, que "el ruido no es un componente de gran relevancia"

En nuestro caso el valor de α es:

```r
accidentes_MEX_ts_h1$alpha
```

```
## [1] 0.9999403
```

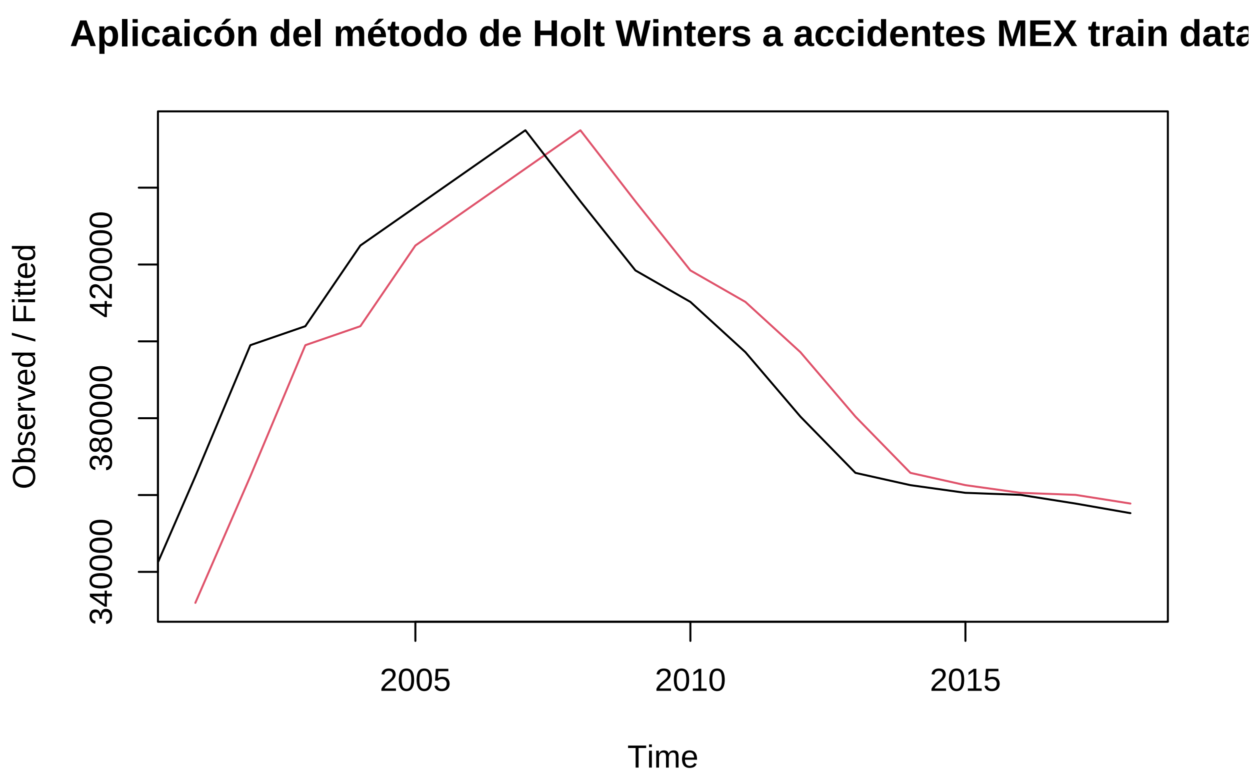
Lo que indica que nuestra data es prácticamente una caminata aleatoria

Pronóstico de la serie de PEA en MEX con el método de Holt Winters

Ahora pasemos a realizar los pronósticos de la serie pero con el método de HoltWinters que considera cambios dinámicos de nivel

Graficamos los valores originales vs suavizados con el método de Holt-Winters

```r
plot(accidentes_MEX_ts_h1, main="Aplicaicón del método de Holt Winters a accidentes MEX train data")
```



```r
# Load the forecast package
library(forecast)
```

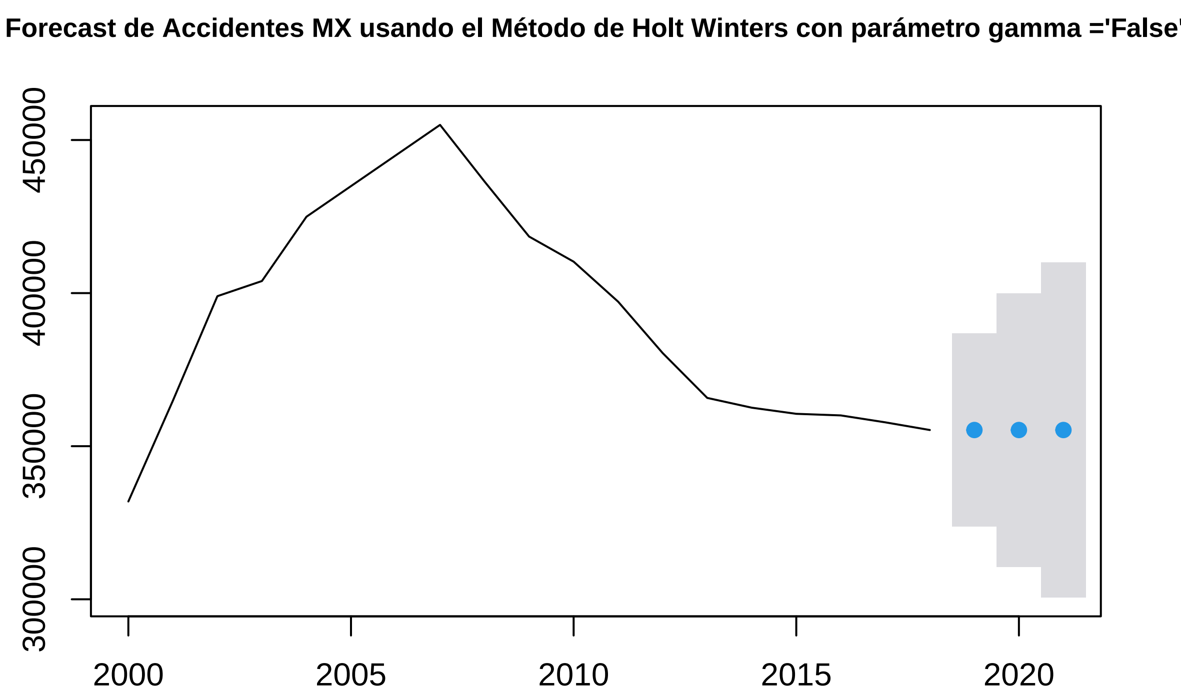
```
## Warning: package 'forecast' was built under R version 4.2.3

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

Ahora se llevará a cabo el pronóstico con la librería de forecast y basados en el método de Holt Winters. De igual manera se usará un intervalo de confianza del 95%

```r
accidentes_MEX_ts_h1_fc <- forecast(accidentes_MEX_ts_h1, h=3, level=0.95)
accidentes_MEX_ts_h1_fc
```

```
##      Point Forecast    Lo 95    Hi 95
## 2019       355281.1 323712.8 386849.5
## 2020       355281.1 310638.1 399924.2
## 2021       355281.1 300605.3 409957.0
```

```r
plot(accidentes_MEX_ts_h1_fc, main="Forecast de Accidentes MX usando el Método de Holt Winters con parám
```



Se observa como el prronostico puntual es "relativamente bueno" y es parecido en valor al obtenido previamente con el método estático, sin embargo los Intervalos de confianza obtenidos con este método son MEJORES en comparación con los del anterior

Veamos el valor la Suma de Errores al Cuadrado del pronóstico de PEA en MEX usando el Método de suaviazamiento dinámico de HoltWinters

```
accidentes_MEX_ts_h1$SSE
```

```
## [1] 4440463515
```

Calculemos el error tipo MAPE entre los valores pronosticados con el "método de Holt Winters" y los valores reales de la serie

```
accidentes_MEX_completa_forecast3_HW  <- accidentes_MEX_ts_h1_fc$mean[1:3]
accidentes_MEX_completa_forecast3_HW
```

```
## [1] 355281.1 355281.1 355281.1
```

```
##Calculemos el MAPE entre estos dos vectores

# First, calculate the absolute percentage error (APE) for each element in the vectors
APE_accidentes_HW <- abs((accidentes_MEX_completa_reales3 - accidentes_MEX_completa_forecast3_HW) / acc
```

```
## Warning in accidentes_MEX_completa_reales3 -
## accidentes_MEX_completa_forecast3_HW: longer object length is not a multiple of
## shorter object length
```

```
# Calculate the mean APE across all elements in the vectors
MAPE_accidentes_HW <- mean(APE_accidentes_HW) * 100

# Print the MAPE value
cat("MAPE error para el forecast calculado con el método de 'HoltWinters'es:", MAPE_accidentes_HW, "%")
```

```
## MAPE error para el forecast calculado con el método de 'HoltWinters'es: 4.463494 %
```

Comparación del MAPE del pronostico hecho con métodos estáticos vs el pronóstico hecho con el método " de suavizamiento dinámico" de HoltWinters para el pronostico puntual

En este caso y dado un cálculo con el métodO de HolttWinters sin valores de arranque especificados los errores de ambos métodos (estático vs HoltWinters) fueron los sigientes:

```
cat("MAPE error para el forecast calculado con el método de 'suavizamiento estático'es:", MAPE_accident
```

```
## MAPE error para el forecast calculado con el método de 'suavizamiento estático'es: 15.51242 %
```

```
cat("MAPE error para el forecast calculado con el método de 'HoltWinters'es:", MAPE_accidentes_HW, "%")
```

```
## MAPE error para el forecast calculado con el método de 'HoltWinters'es: 4.463494 %
```

Lo que indica que para este caso el Método de HoltWinters fue ligeramente mejor para el cálculo de las estimaciones puntuales

Comparación de los intervalos del pronostico hecho con métodos estáticos vs el pronóstico hecho con el método " de suavizamiento dinámico" de HoltWinters

Como se sabe, la parte "más sustancial" o esencial de los pronósticos es el intervalo y no el pronóstico puntual, por lo que en nuestro caso se considerará en general como "un mejor método de pronóstico" no aquel que otorge el menor error tipo MAPE en las estimaciones puntuales, sino aquel cuyos intervalos de confianza al 95% sean menores incluso para valores pronósticados relativamente lejanos del último valor real.

En este caso se observa claramente como el método de HoltWinters otorga mejores intervalos de confianza para pronósticos incluso para aquellos 'más alejados' del último valor real vs aquellos generados con el "método de suavizamiento estático"

Serie de tiempo 2 CON tendencia y SIN estacionalidad: Historico trimestral de Población Económicamente Activa (PEA) en México

Tenemos data trimestral de Enero de 2005 a Agosto de 2022 sobre la Población Económicamente activa en México, son 72 datos en total. La idea es usar 66 de esos 72 como una especie de "data de entrenamiento" para aplicar los modelos A) De suavizamiento "estático" y B) De "suavizamiento dinámico" o Holt-Winters y posteriormente predecir con cada modelo los valores de los 6 trimestres restantes y poder comparar contra los reales e identificar el mejor modelo entre los dos mencionados para modelar la Población Económicamente Activa en México

Primero importamos la data de PEA en MEX Data obtenida de la página oficial del INEGI: https://www.inegi.org.mx/temas/empleo/

```
library("readxl")

PEA_mex <-read_excel("PEA_MEX_pp.xlsx")
```

Revisamos que la data de nacimientos MEX sehay importado correctamente

```
head(PEA_mex)
```

```
## # A tibble: 6 x 2
##   Date                      PEA
##   <dttm>                  <dbl>
## 1 2005-01-01 00:00:00 43099847
## 2 2005-04-01 00:00:00 43180433
## 3 2005-07-01 00:00:00 44000204
## 4 2005-10-01 00:00:00 44245519
## 5 2006-01-01 00:00:00 44306012
## 6 2006-04-01 00:00:00 44611672
```

```
summary(PEA_mex)
```

```
##       Date                             PEA
##  Min.   :2005-01-01 00:00:00.00   Min.   :43099847
##  1st Qu.:2009-01-23 12:00:00.00   1st Qu.:46978140
##  Median :2013-02-15 00:00:00.00   Median :51090386
##  Mean   :2013-02-14 12:21:49.08   Mean   :50515648
##  3rd Qu.:2017-03-09 12:00:00.00   3rd Qu.:53357321
##  Max.   :2021-04-01 00:00:00.00   Max.   :57668254
```

```r
typeof(PEA_mex)
```

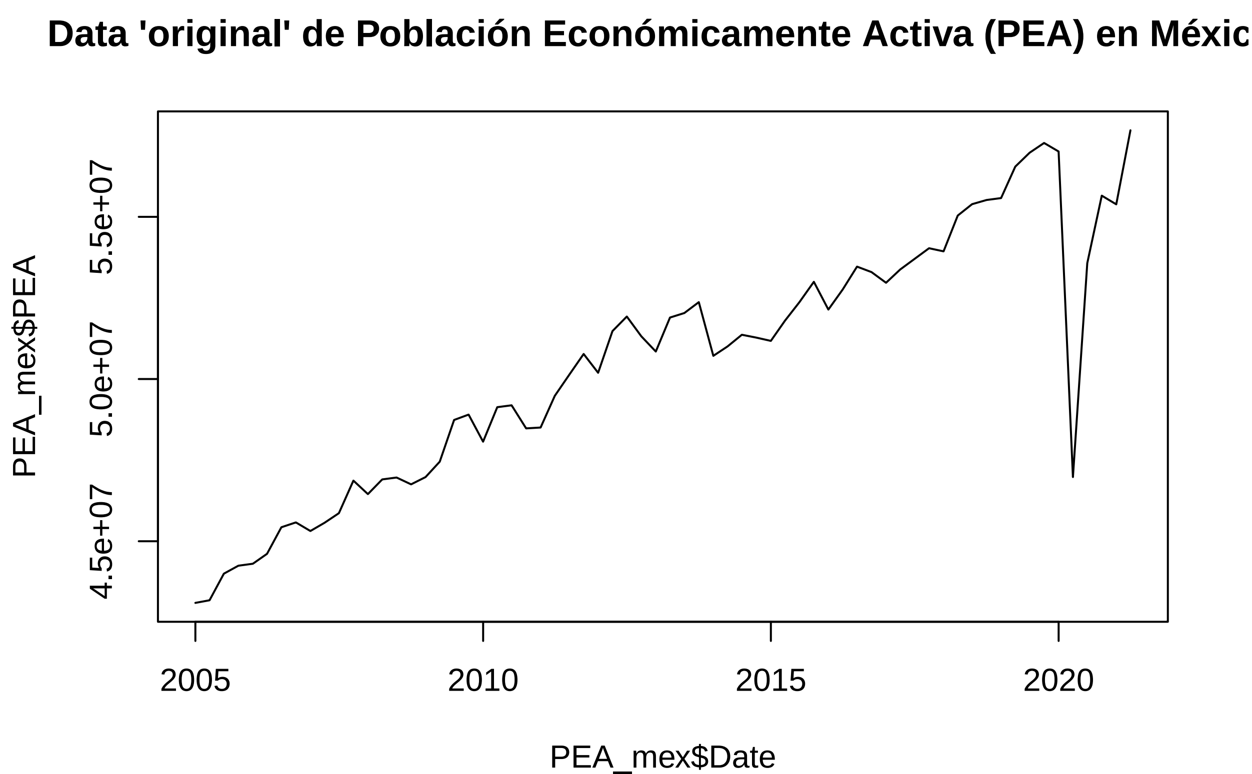
```
## [1] "list"
```

```r
dim(PEA_mex)
```

```
## [1] 66  2
```

Graficamos la "data original" de PEA en México de los últimos años (Enero 2005 a Oct 2021)

```r
plot(PEA_mex$Date, PEA_mex$PEA,type="l", main="Data 'original' de Población Económicamente Activa (PEA)
```



A. Método de "suavizamiento estático"

Comenzaremos tratando la serie con el método estático aplicando el siguiente flujo lógico de trabajo y considerando que estamos ene el caso de serie CON TENDECIA y SIN ESTACIONALIDAD:

1. Se removerá el factor de tendencia a la serie original obteniendo una serie sin tendencia
2. Será sobre dicha serie sin tendecia que se llevará a cabo el "pronóstico puntual" de los siguientes 6 trimestre como la media historica de dicho componente estocástico
3. Se calcularán los intervalos de pronóstico al 95% de confianza
4. Se "devolverán" los efectos de tendencia de la serie con el fin de llevar los pronósticos a las dimensiones correctas
5. Se calculará el error tipo MAPE entre los valores pronósticados con el "método estático" y los valores reales de la serie de los últimos 6 trimestres

A.1 Remoción de tendencia de serie original

Primero se calculará la linea "general" de tendencia de nuestra serie de (que representa la tendencia de la "serie de datos originales") y se removerá dicha tendencia de la 'data original' de PEA_mex

Calculemos la linea de tendencia

```
linea_PEA<-lm(PEA~Date, data=PEA_mex)
linea_PEA
```

```
##
## Call:
## lm(formula = PEA ~ Date, data = PEA_mex)
##
## Coefficients:
## (Intercept)         Date
##   1.789e+07    2.397e-02
```

Se puede observar una "ecuación general de linea de tendencia" en este caso es: y = 17,890,000 + 0.02397 * (Date) Es muy razonable que "la ordenada al origen" sea de alrededor de 20 millones, pues todos los datos de Población Económicamente Activa son del orden de 40 a 55 millones. También resulta lógica que el valor de la pendiente sea positivo pero no muy cercano a 1, pues en general para cada aumento en el valor del tiempo, no existe un incremento o decremento realmente signifiativo en el valor del PEA (excepto en el año 2020)

Ahora para graficar la "serie original" de PEA_MEX vs el filtro1 (Linea de tendencia) primero debemos "crear/simular" datos de Enero 2005 a Oct 2021 a partir de nuestra linea de previamente computada de nombre: linea_PEA

```
linea_PEA_aplicada <- predict(linea_PEA, newdata = data.frame(Date = PEA_mex$Date))

head(linea_PEA_aplicada)
```

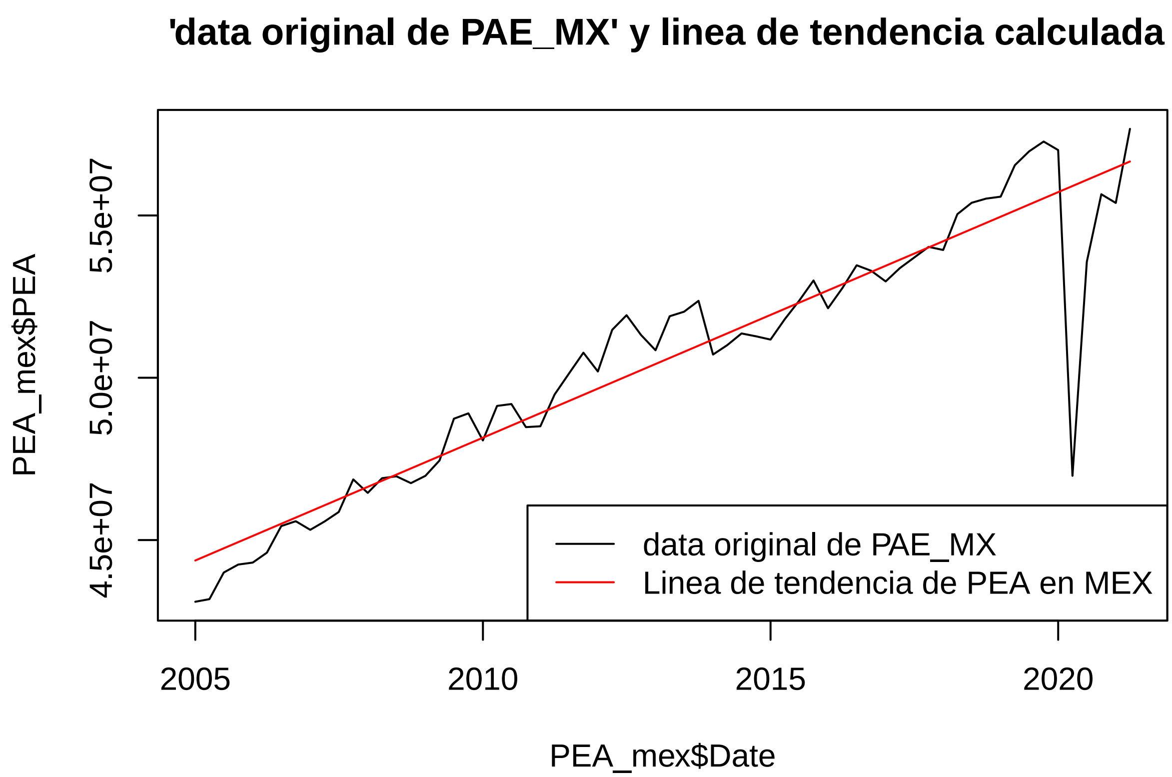
```
##        1        2        3        4        5        6
## 44370784 44557211 44745709 44936278 45126848 45313275
```

```
tail(linea_PEA_aplicada)
```

```
##       61       62       63       64       65       66
## 55717958 55906456 56094954 56285523 56476093 56662520
```

Graficamos entonces a la "data original de PAE_MX" y a la data calculada a partir de la linea tendencia

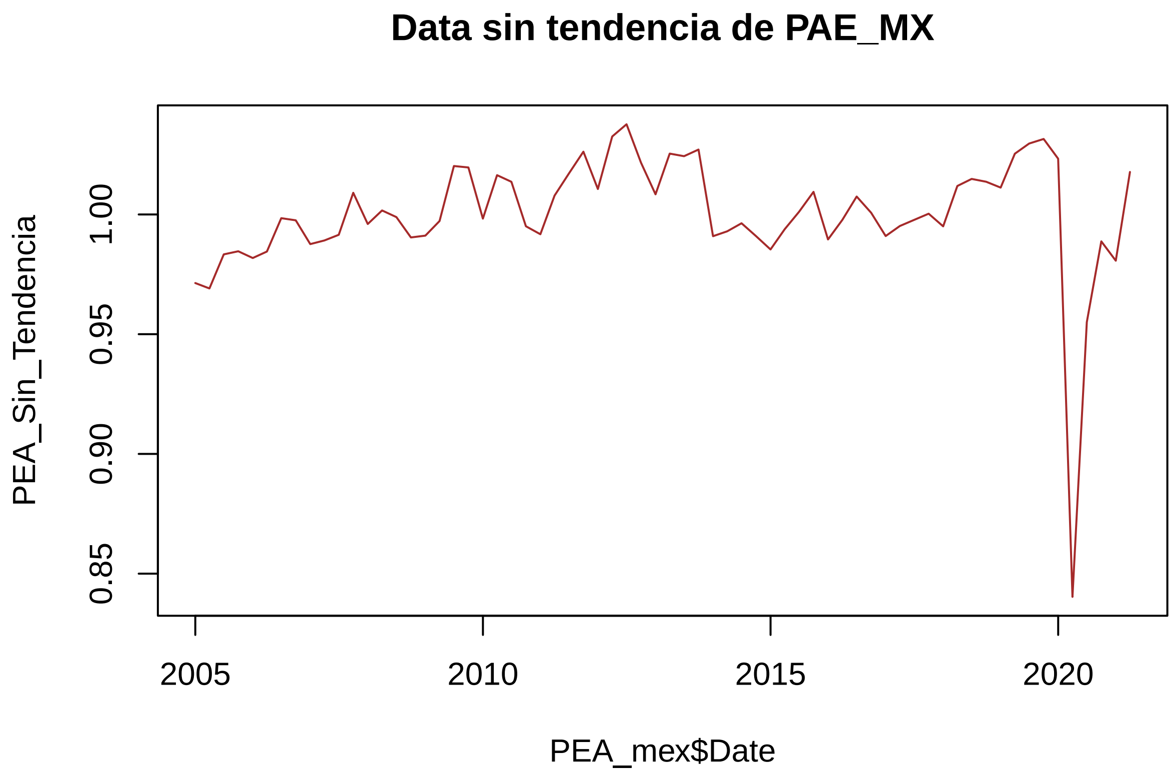
```
plot(PEA_mex$Date, PEA_mex$PEA, type="l", main=" 'data original de PAE_MX' y linea de tendencia calcula
lines(x = PEA_mex$Date, y = linea_PEA_aplicada, col = "red")

legend("bottomright", legend = c("data original de PAE_MX", "Linea de tendencia de PEA en MEX"),
       col = c("black", "red"), lty = 1)
```



Ahora vamos a "quitar el efecto de la tendencia" de nuestra "serie original de datos" En este caso se asumirá un "efecto multiplicativo" y por lo tanto se removerá dividiendo cada elemento de la serie original / valor de la linea de tendnecia en ese valor de x

Se grafica la serie sin tendencia:

```
PEA_Sin_Tendencia<-PEA_mex$PEA/linea_PEA_aplicada

plot(PEA_mex$Date, PEA_Sin_Tendencia, type="l", main=" Data sin tendencia de PAE_MX ", col="brown")
```



A.2 Tratamiendo del ruido blanco en la serie y pronostico puntual

Considerese el siguiente razonamiento acerca de las series del tiempo: Serie total = Componente de tendencia + Componente de estacionalidad + Componente estocástico

En este caso el Componente de estacionalidad = 0, por lo tanto tendríamos

Serie total PAEA MEX = Componente de tendencia + Componente estocástico, por lo tanto al haber previamente removido el componente estocástico de la serie lo que hemos hecho en realidad es obtener el componente estocástico o ruido blanco para este caso

Por simplicidad para cálculos posteriores, nombremos de esa manera a la variable

```
PEA_MX_ruido_blanco <- PEA_Sin_Tendencia
head(PEA_MX_ruido_blanco)
```

```
##         1         2         3         4         5         6
## 0.9713564 0.9691009 0.9833391 0.9846280 0.9818105 0.9845166
```

```
tail(PEA_MX_ruido_blanco)
```

```
##        61        62        63        64        65        66
## 1.0232781 0.8403117 0.9550198 0.9887701 0.9806828 1.0177496
```

Una vez teniendo únicamente el componente estocástico de nuestra serie, es posible "extraer los errores" de la serie y conocer así el error estimado (varianza) y la desviación estandar. La idea es que extrayendo de la serie del ruido blanco su media histórica estaríamos determinando la varianza, de la siguiente manera:

Varianza (errores al cuadrado) = Serie original - tendecia - media de componete estocástico. En nuestro caso, dichos errores será nuestra serie sin tendencia - mean(PEA_MX_ruido_blanco). Nosotros hemos ya calculado el resultado de Serie original - tendecia y lo hemos llamada justamente ruido blanco y lo lo tanto el calculo del error podría resumirse a

Errores= Componente estocástico o ruido blanco - Media del componente estocástico o ruido blanco

Se determinan la varianza (errores al cuadrado) de la serie de PEA_MEX y se calcula también la desviación estandar de esos errores

```
Error_est_PEA<-PEA_Sin_Tendencia-mean(PEA_MX_ruido_blanco)
head(Error_est_PEA)
```

```
##           1           2           3           4           5           6
## -0.02859805 -0.03085360 -0.01661542 -0.01532647 -0.01814402 -0.01543788
```

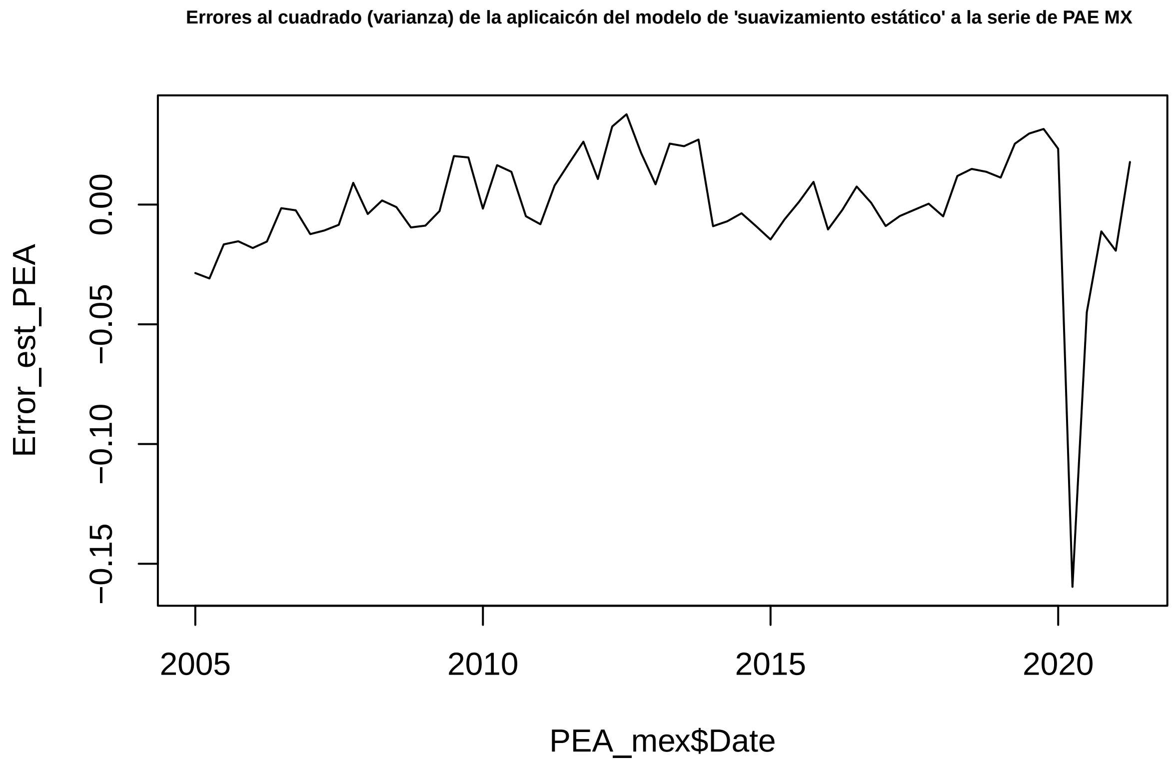
```
tail(Error_est_PEA)
```

```
##          61          62          63          64          65          66
##  0.02332363 -0.15964281 -0.04493471 -0.01118444 -0.01927170  0.01779506
```

```
##Claculando la desviaicón estandar del error

sd_est_PEA<-sd(Error_est_PEA)
sd_est_PEA
```

```
## [1] 0.02599512
```

```
plot(PEA_mex$Date, Error_est_PEA, type="l", main="Errores al cuadrado (varianza) de la aplicaicón del mo
```



Se calcula la media de los errores

```r
mean(Error_est_PEA)
```

```
## [1] 5.047007e-18
```

Se calcula la varianza y desviación estandar de los errores

```r
## Se calcula la varianza
var(Error_est_PEA)
```

```
## [1] 0.0006757464
```

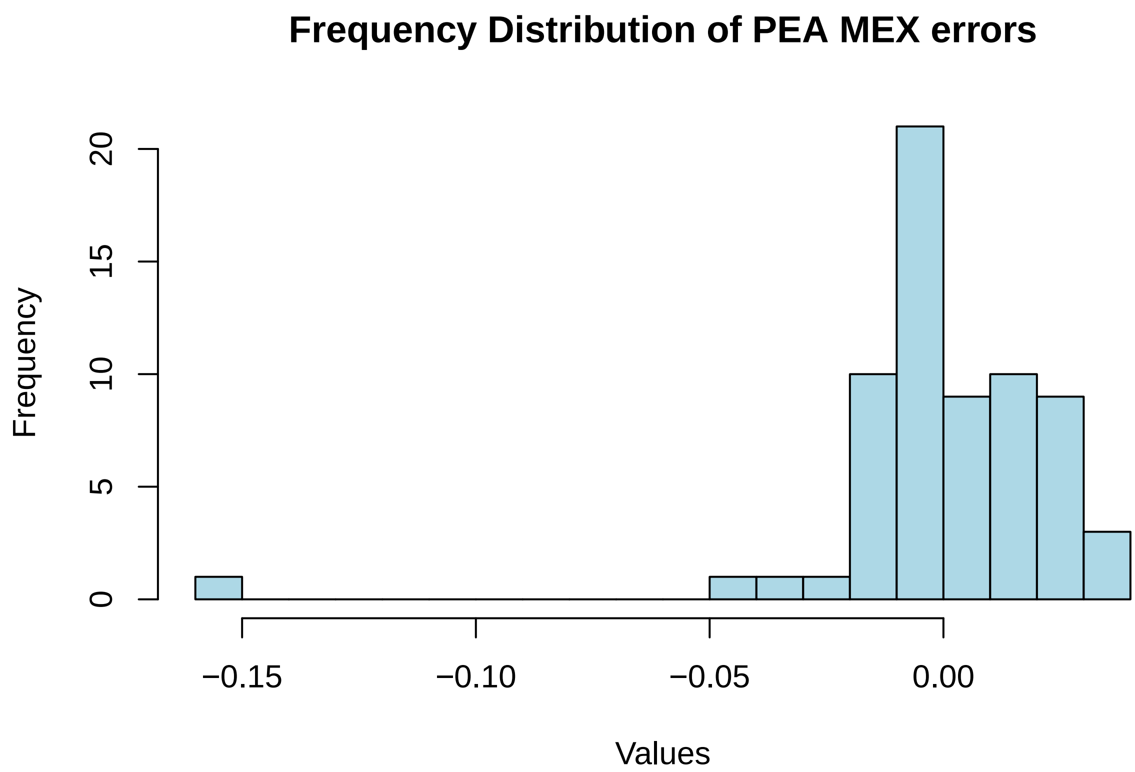
```r
## Se calcula la desviación estandar
sd_est_PEA
```

```
## [1] 0.02599512
```

Se tiene una desviación estandar de 0.025 y una media muy cercana a cero para los errores lo cual es totalmente concordante con le hecho de que tras haber removido el componente de tendencia, "el ruido blanco" o componente estocástico resultante tiende a tener una distribuión normal. Y que si a ese ruido blanco se le extrae su media tambien tiende a tener una distribución normal con media cero y en este caso una desviación estandar de 0.025

Procedamos a observar como se comporta la Distribución de dichos errores:

```r
hist(Error_est_PEA, breaks = 20, col = "lightblue", main = "Frequency Distribution of PEA MEX errors", 
```



Se observa como los errores de la serie siguen prácticamente todos una distribución normal con media cero y desviación estandar de 0.025, a excepción de los valores atípicos del año 2020 derivados de la pandemia de Covid-19.

Pronóstico puntual Una vez teniendo nuestra seríe SIN TENDENCIA lo que como ya se mencionó, implica que, "unicamente tenemos el componente estocástico o de ruido blanco en la serie" podemos proceder a llevar a cabo el pronóstico de 6 trimestres.

Como ha sido estudiado en la clase, nuestro mejor pronóstico puntual dadas estas condiciones en la serie es la media de los datos históricos del ruido blanco

```r
estimacion_puntual_PEA_ST<-mean(PEA_Sin_Tendencia)
estimacion_puntual_PEA_ST
```

```
## [1] 0.9999545
```

Ahora vamos a hacer el pronóstico de la serie de PEA MEX para los próximos 6 trimestres

```r
##Creamos un eje x que incluya a los 66 datos + 6 que se pronosticarán
Date_PAE2<-c(1:72)


## Creamos el vector de datos que incluirá los  datos históricos  sin tendencia + 6 valores NA a pronos
```

```r
PAE_Sin_Tendencia_para_pron<-c(PEA_Sin_Tendencia, rep(NA, times=6))
head(PAE_Sin_Tendencia_para_pron)
```

```
##         1         2         3         4         5         6
## 0.9713564 0.9691009 0.9833391 0.9846280 0.9818105 0.9845166
```

```r
tail(PAE_Sin_Tendencia_para_pron)
```

```
##
## NA NA NA NA NA NA
```

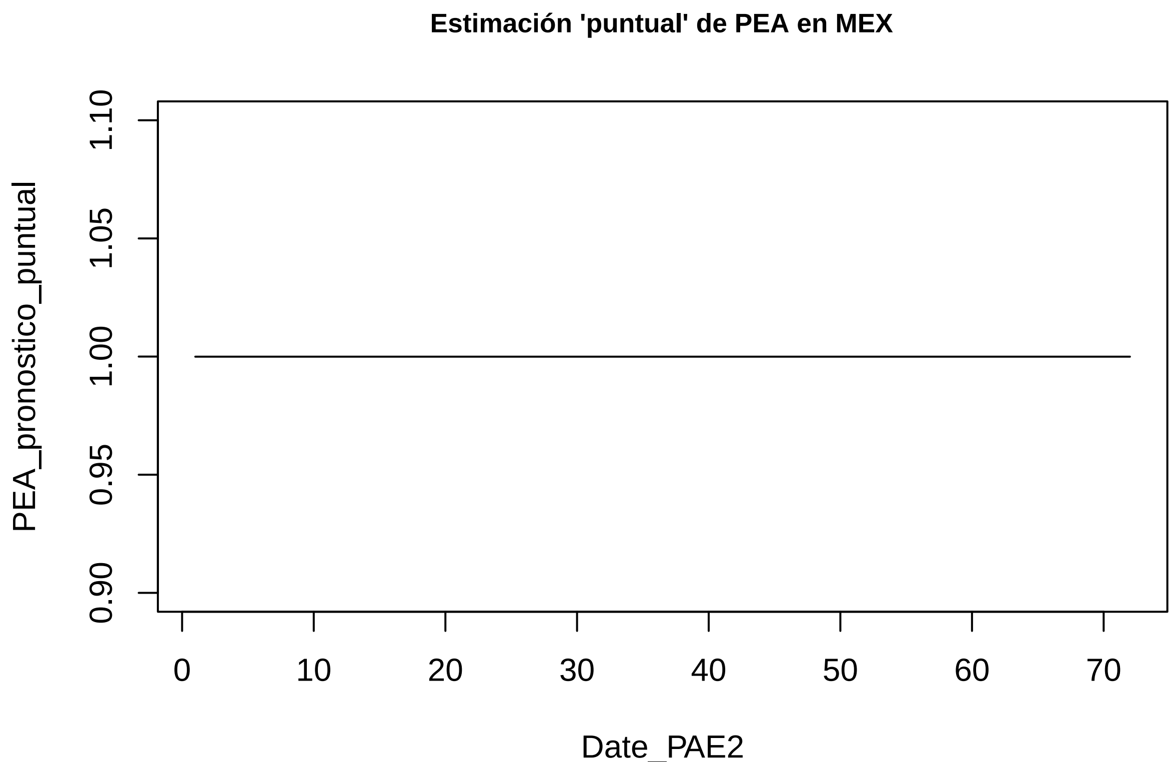
```r
## Entonces el mejor pronóstico que tenemos para esta técnica es la media de la serie sin tendencia


PEA_pronostico_puntual <- rep(estimacion_puntual_PEA_ST, times=72)
head(PEA_pronostico_puntual)
```

```
## [1] 0.9999545 0.9999545 0.9999545 0.9999545 0.9999545 0.9999545
```

```r
tail(PEA_pronostico_puntual)
```

```
## [1] 0.9999545 0.9999545 0.9999545 0.9999545 0.9999545 0.9999545
```

```r
plot(Date_PAE2,PEA_pronostico_puntual, type="l", main = "Estimación 'puntual' de PEA en MEX", cex.main=(
```



Comentario sobre el pronóstico puntual sin tendencia ni estacionalidad, que para el caso de la serie de esta serie tiene un valor de:

```
mean(PEA_MX_ruido_blanco)
```

```
## [1] 0.9999545
```

Este valor es bastante concordante con el hecho que. Al haber "quitado los componentes de tendencia de la serie original", es decir obtenido el ruido blanco de la serie los valores de nuestra variable y: PEA en Mex están en este momento del algoritmo en escala de la unidad: 1. Y por lo tanto la media de dichos valores será muy cercana a 1.

En general se podría decir dados los resultados del presente trabajo, así como los resultados de los ejercicios vistos en clase, que el valor de la media estimada o pronóstico puntual para la serie sin componente de tendencia tiende al valor de 1.

A.3 Calculo de intervalos de confianza para el pronóstico

Calculamos los límites de intervalos de confianza al 95% pata el pronostico Se considera 95% como +- 2 desviaciones estandar de la media

```
tao<-c(1:6)
Lim_PEA<-2*sd_est_PEA*tao^.5
Lim_PEA
```

```
## [1] 0.05199024 0.07352531 0.09004974 0.10398049 0.11625372 0.12734957
```

Calculamos los limites inferior y superior

```
LI_PEA<-c(rep(NA, times=66), estimacion_puntual_PEA_ST-Lim_PEA)
head(LI_PEA)
```

```
## [1] NA NA NA NA NA NA
```

```
tail(LI_PEA)
```

```
## [1] 0.9479643 0.9264292 0.9099048 0.8959740 0.8837008 0.8726049
```

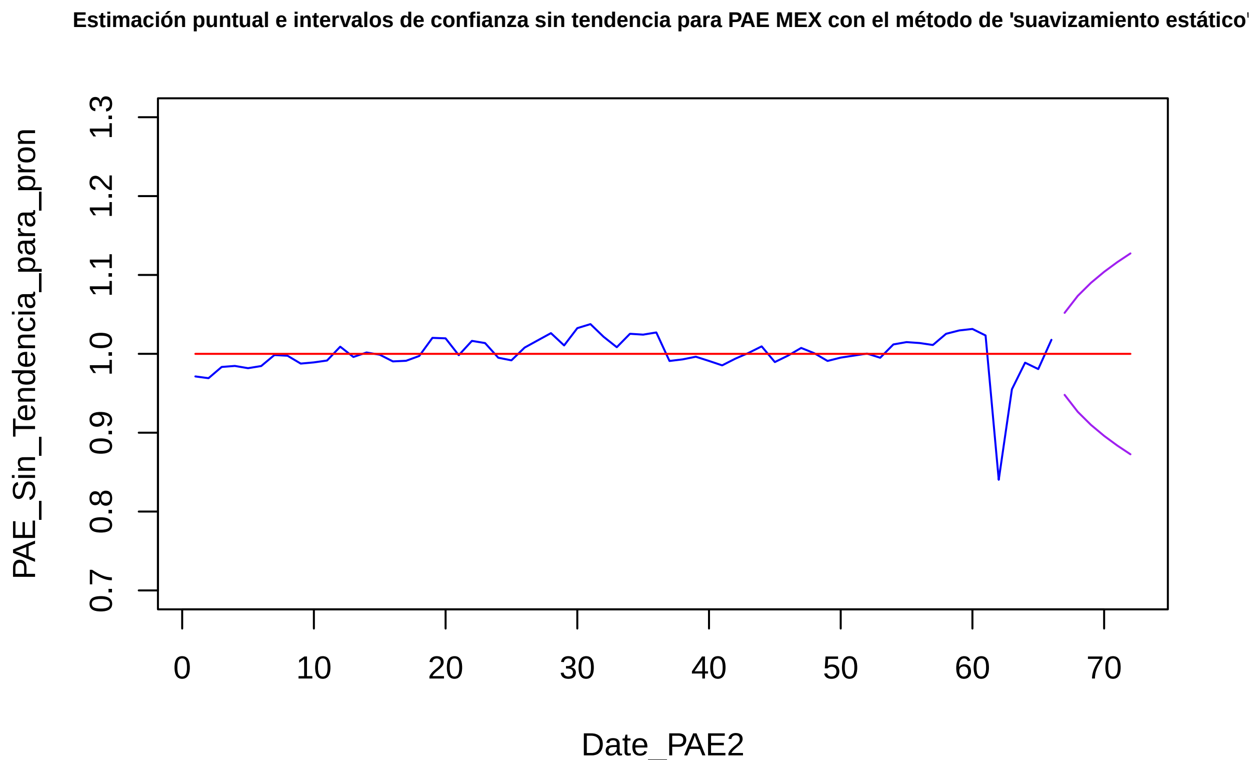
```
LS_PEA<-c(rep(NA, times=66), estimacion_puntual_PEA_ST+Lim_PEA)
head(LS_PEA)
```

```
## [1] NA NA NA NA NA NA
```

```
tail(LS_PEA)
```

```
## [1] 1.051945 1.073480 1.090004 1.103935 1.116208 1.127304
```

```r
plot(Date_PAE2, PAE_Sin_Tendencia_para_pron, type="l", ylim = c(0.7,1.3) ,main="Estimación puntual e in

lines(PEA_pronostico_puntual, col="Red")
lines(LI_PEA, col="Purple")
lines(LS_PEA, col="Purple")
```



A.4 "Devolución de efectos de tendencia y estacionalidad a los pronósticos e intervalos calculados"

Ahora vamos a "llevar nuestro pornóstico a las dimensiones reales" Para ello primero vamos a "devolver la tendencia a la serie" de nombre PEA_Sin_Tendencia

Para eso primero necesitamos "crear un data frame que contenga la información histórica + 6 de pronóstico con un valor NA

```r
# create a new data set with 72 rows
extended_dates <- seq(as.Date("2005-01-01"), as.Date("2022-10-01"), by = "quarter")
extended_trimester_PEA <- data.frame(Date = extended_dates,
                                     PEA = rep(NA, length(extended_dates)))

head(extended_trimester_PEA)
```

```
##         Date PEA
## 1 2005-01-01  NA
```

```
## 2 2005-04-01  NA
## 3 2005-07-01  NA
## 4 2005-10-01  NA
## 5 2006-01-01  NA
## 6 2006-04-01  NA
```

```
tail(extended_trimester_PEA)
```

```
##          Date PEA
## 67 2021-07-01  NA
## 68 2021-10-01  NA
## 69 2022-01-01  NA
## 70 2022-04-01  NA
## 71 2022-07-01  NA
## 72 2022-10-01  NA
```

```
# copy the original data to the new data set for the first 66 rows
for(i in 1:66){
  extended_trimester_PEA$PEA[i] <- PEA_mex$PEA[i]
}

extended_trimester_PEA
```

```
##          Date      PEA
## 1  2005-01-01 43099847
## 2  2005-04-01 43180433
## 3  2005-07-01 44000204
## 4  2005-10-01 44245519
## 5  2006-01-01 44306012
## 6  2006-04-01 44611672
## 7  2006-07-01 45431392
## 8  2006-10-01 45580994
## 9  2007-01-01 45314888
## 10 2007-04-01 45569395
## 11 2007-07-01 45864926
## 12 2007-10-01 46868952
## 13 2008-01-01 46453196
## 14 2008-04-01 46905921
## 15 2008-07-01 46964082
## 16 2008-10-01 46753657
## 17 2009-01-01 46977904
## 18 2009-04-01 47453163
## 19 2009-07-01 48738589
## 20 2009-10-01 48903792
## 21 2010-01-01 48069274
## 22 2010-04-01 49133132
## 23 2010-07-01 49190032
## 24 2010-10-01 48478718
## 25 2011-01-01 48505168
## 26 2011-04-01 49482112
## 27 2011-07-01 50127032
## 28 2011-10-01 50772496
## 29 2012-01-01 50192842
```

```
## 30 2012-04-01 51477178
## 31 2012-07-01 51927050
## 32 2012-10-01 51317999
## 33 2013-01-01 50847242
## 34 2013-04-01 51895865
## 35 2013-07-01 52034353
## 36 2013-10-01 52370886
## 37 2014-01-01 50715329
## 38 2014-04-01 51004605
## 39 2014-07-01 51364782
## 40 2014-10-01 51277056
## 41 2015-01-01 51176166
## 42 2015-04-01 51800807
## 43 2015-07-01 52368409
## 44 2015-10-01 52997084
## 45 2016-01-01 52141197
## 46 2016-04-01 52760481
## 47 2016-07-01 53465906
## 48 2016-10-01 53296175
## 49 2017-01-01 52967544
## 50 2017-04-01 53377703
## 51 2017-07-01 53702713
## 52 2017-10-01 54032400
## 53 2018-01-01 53936667
## 54 2018-04-01 55039393
## 55 2018-07-01 55390625
## 56 2018-10-01 55519394
## 57 2019-01-01 55578352
## 58 2019-04-01 56547664
## 59 2019-07-01 56976075
## 60 2019-10-01 57277858
## 61 2020-01-01 57014967
## 62 2020-04-01 46978848
## 63 2020-07-01 53571791
## 64 2020-10-01 55653440
## 65 2021-01-01 55385133
## 66 2021-04-01 57668254
## 67 2021-07-01       NA
## 68 2021-10-01       NA
## 69 2022-01-01       NA
## 70 2022-04-01       NA
## 71 2022-07-01       NA
## 72 2022-10-01       NA
```

```r
# fit the model with the extended data set
linea_PEA_extended <- lm(PEA ~ Date, data = extended_trimester_PEA)
linea_PEA_extended
```

```
##
## Call:
## lm(formula = PEA ~ Date, data = extended_trimester_PEA)
##
## Coefficients:
## (Intercept)         Date
```

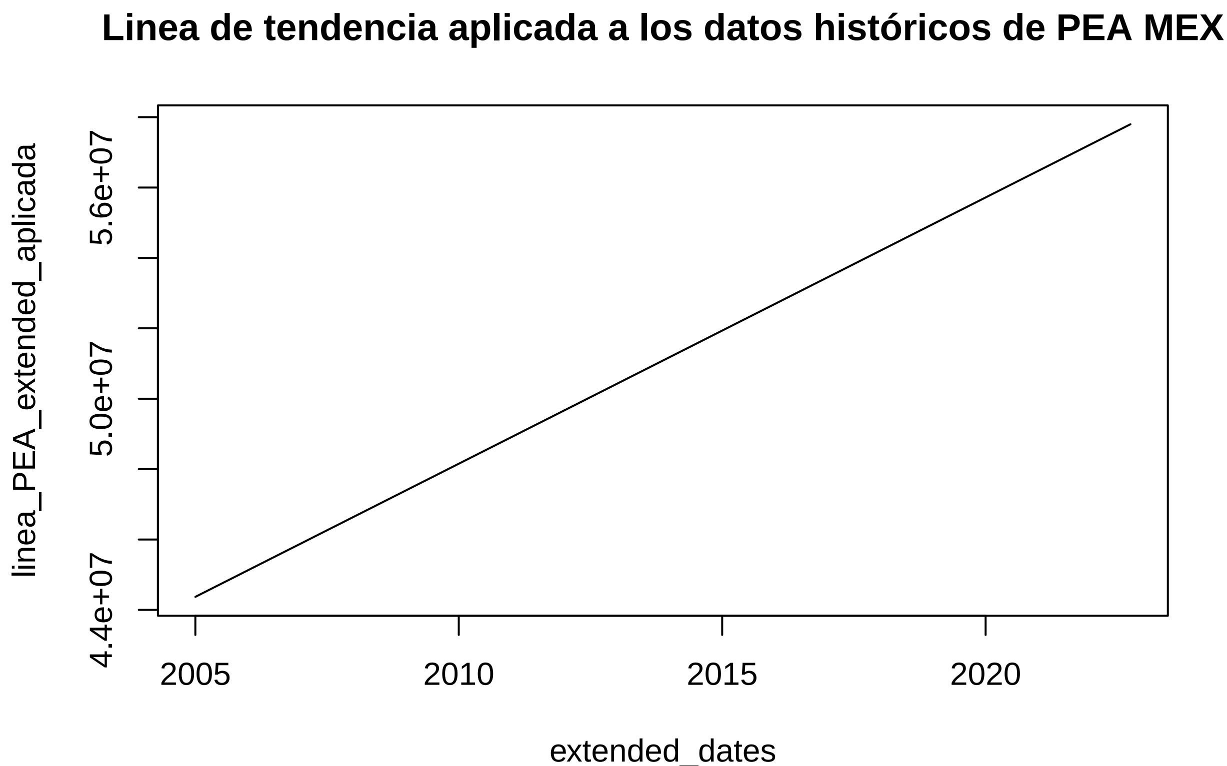
```
##    17889903          2071
```

Posteriormente calculamos los valores historicos + 6 pronosticados a partir de la linea de tendencia que previamente se habia calculado de nombre linea_PEA_extended

```r
# predict values for the extended date range
linea_PEA_extended_aplicada <- predict(linea_PEA_extended,
                                       newdata = data.frame(Date = extended_dates))
linea_PEA_extended_aplicada
```

```
##        1        2        3        4        5        6        7        8
## 44370784 44557211 44745709 44936278 45126848 45313275 45501773 45692342
##        9       10       11       12       13       14       15       16
## 45882912 46069339 46257837 46448406 46638976 46827474 47015972 47206542
##       17       18       19       20       21       22       23       24
## 47397111 47583538 47772036 47962606 48153175 48339602 48528100 48718670
##       25       26       27       28       29       30       31       32
## 48909239 49095666 49284164 49474734 49665303 49853801 50042299 50232869
##       33       34       35       36       37       38       39       40
## 50423438 50609865 50798363 50988933 51179502 51365929 51554427 51744997
##       41       42       43       44       45       46       47       48
## 51935566 52121993 52310491 52501061 52691630 52880128 53068627 53259196
##       49       50       51       52       53       54       55       56
## 53449766 53636192 53824691 54015260 54205830 54392256 54580754 54771324
##       57       58       59       60       61       62       63       64
## 54961894 55148320 55336818 55527388 55717958 55906456 56094954 56285523
##       65       66       67       68       69       70       71       72
## 56476093 56662520 56851018 57041587 57232157 57418584 57607082 57797651
```

```r
# plot the predicted values
plot(extended_dates, linea_PEA_extended_aplicada, type = "l", main="Linea de tendencia aplicada a los da
```



Entonces para finalmente "devolver el efecto de la tendnecia a los datos de media_nacimientos_ST" se multiplica cada valor de la serie por el valor calculado a partir de la linea de tencia que se llama: linea_PEA_extended_aplicada

```
##Creando el vector original + los valores a predecir como NAs
PEA_original_para_pron<-c(PEA_mex$PEA, rep(NA, times=6))
head(PEA_original_para_pron)
```

```
## [1] 43099847 43180433 44000204 44245519 44306012 44611672
```

```
Est_punt_PEA_ajustada<-c(rep(NA, times=66), mean(PEA_pronostico_puntual)*linea_PEA_extended_aplicada[67
head(Est_punt_PEA_ajustada)
```

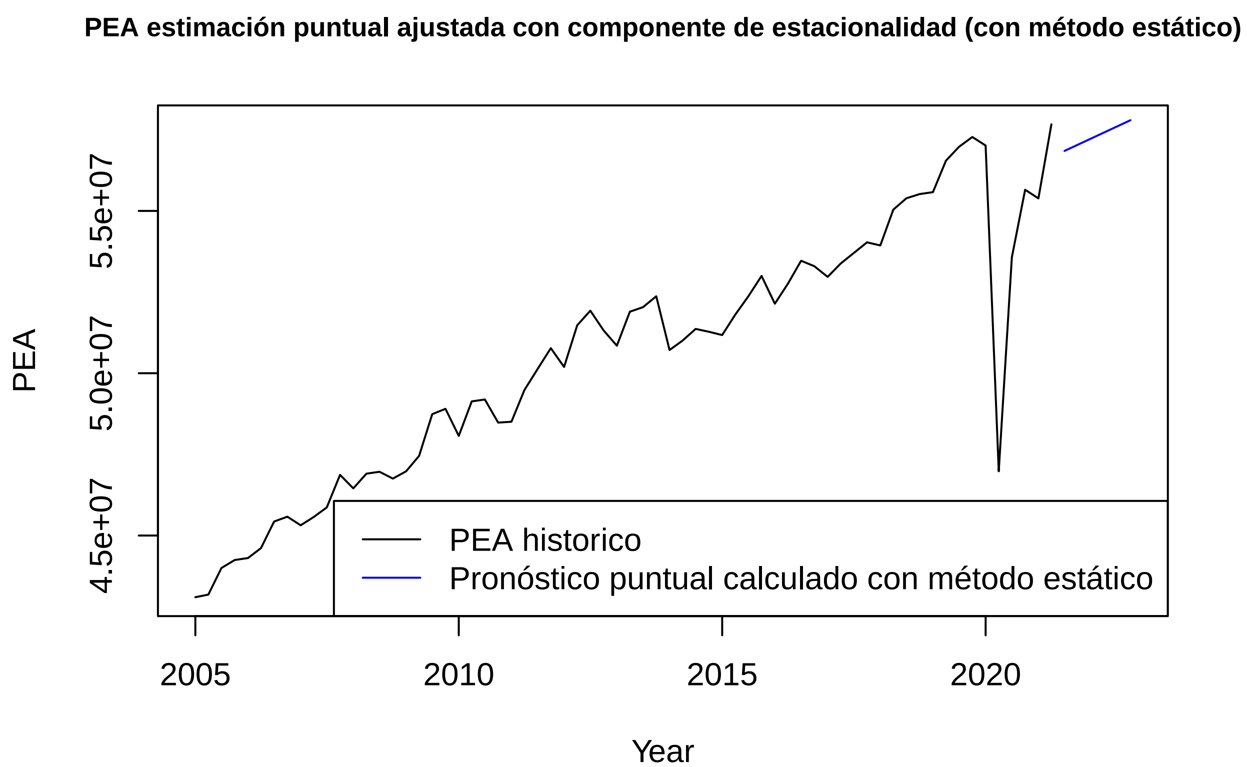
```
##
## NA NA NA NA NA NA
```

```
tail(Est_punt_PEA_ajustada)
```

```
##       67       68       69       70       71       72
## 56848431 57038992 57229553 57415971 57604460 57795021
```

```
# plot the original data and the predicted values together
plot(extended_dates, PEA_original_para_pron, main="PEA estimación puntual ajustada con componente de es
lines(extended_dates, Est_punt_PEA_ajustada, col="blue")
```

```r
legend("bottomright", legend = c("PEA historico", "Pronóstico puntual calculado con método estático"),
       col = c("black", "blue"), lty = 1)
```



Calculamos ahora los límites de estimación devolviendo el efecto de tendencia

Para ello importamos el Data Frame con la data historica completa, es decir, la que contiene los 66 trimestres propuestos aquí como "train data" + los 6 trimestres que se pronósticaran

```r
library("readxl")

PEA_MEX_completa <-read_excel("PEA_MEX_full.xlsx")
head(PEA_MEX_completa)
```

```
## # A tibble: 6 x 2
##   Date                     PEA
##   <dttm>                 <dbl>
## 1 2005-01-01 00:00:00 43099847
## 2 2005-04-01 00:00:00 43180433
## 3 2005-07-01 00:00:00 44000204
## 4 2005-10-01 00:00:00 44245519
## 5 2006-01-01 00:00:00 44306012
## 6 2006-04-01 00:00:00 44611672
```

```r
tail(PEA_MEX_completa)
```

```
## # A tibble: 6 x 2
##   Date                     PEA
##   <dttm>                 <dbl>
## 1 2021-07-01 00:00:00 58307446
## 2 2021-10-01 00:00:00 58761793
## 3 2022-01-01 00:00:00 58085314
## 4 2022-04-01 00:00:00 59338419
## 5 2022-07-01 00:00:00 59480471
## 6 2022-10-01 00:00:00 60145456
```

```r
LI_PEA_ajustado<-c(rep(NA, times=66), LI_PEA[67:72]*linea_PEA_extended_aplicada[67:72])
LI_PEA_ajustado
```

```
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                            67       68       69       70       71       72 
##       NA       NA 53892733 52844991 52075812 51445559 50907423 50434515
```

```r
LS_PEA_ajustado<-c(rep(NA, times=66), LS_PEA[67:72]*linea_PEA_extended_aplicada[67:72])
LS_PEA_ajustado
```

```
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
##                                                                         
##       NA       NA       NA       NA       NA       NA       NA       NA 
```

```
##
##       NA       NA       NA       NA       NA       NA       NA       NA
##                         67       68       69       70       71       72
##       NA       NA 59804129 61232992 62383294 63386383 64301498 65155527
```

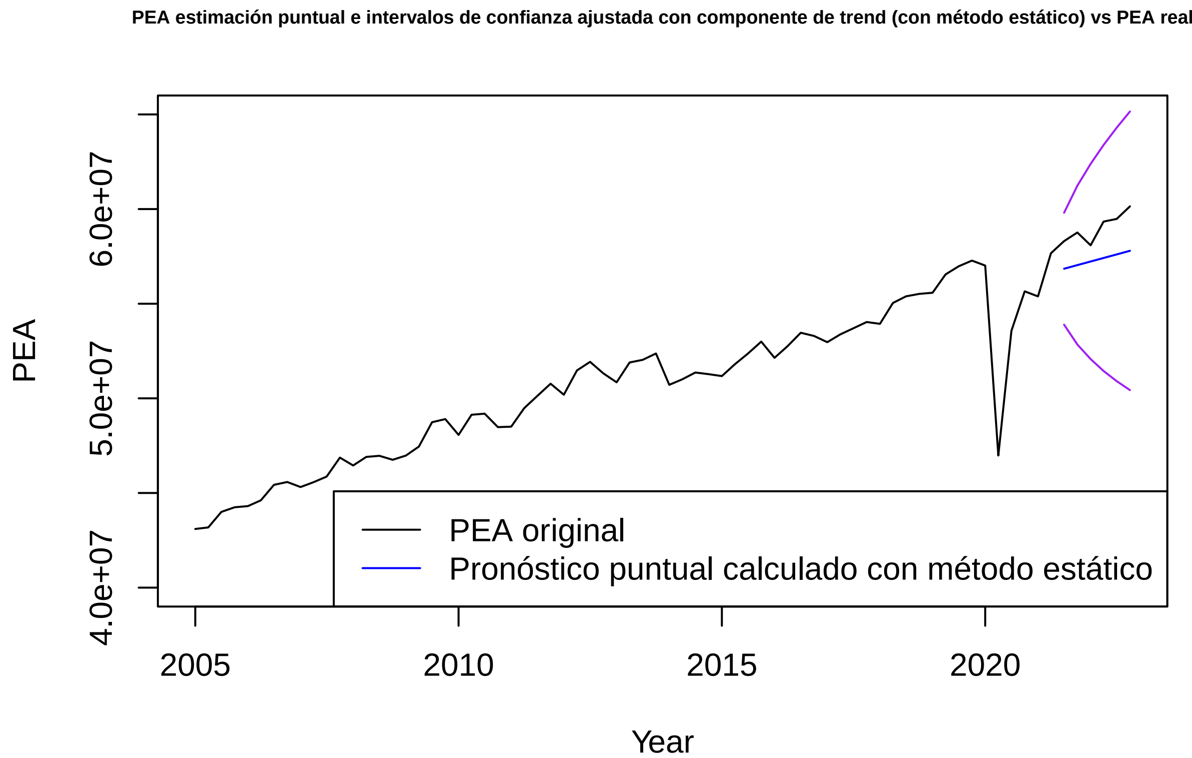
```r
# Plot only non-NA values
# plot the original data and the predicted values together


PEA_original_para_pron<-c(PEA_mex$PEA, rep(NA, times=6))
PEA_original_para_pron
```

```
##  [1] 43099847 43180433 44000204 44245519 44306012 44611672 45431392 45580994
##  [9] 45314888 45569395 45864926 46868952 46453196 46905921 46964082 46753657
## [17] 46977904 47453163 48738589 48903792 48069274 49133132 49190032 48478718
## [25] 48505168 49482112 50127032 50772496 50192842 51477178 51927050 51317999
## [33] 50847242 51895865 52034353 52370886 50715329 51004605 51364782 51277056
## [41] 51176166 51800807 52368409 52997084 52141197 52760481 53465906 53296175
## [49] 52967544 53377703 53702713 54032400 53936667 55039393 55390625 55519394
## [57] 55578352 56547664 56976075 57277858 57014967 46978848 53571791 55653440
## [65] 55385133 57668254       NA       NA       NA       NA       NA       NA
```

```r
dim(PEA_original_para_pron)
```

```
## NULL
```

```r
# plot the original data and the predicted values together
plot(extended_dates,PEA_MEX_completa$PEA, main="PEA estimación puntual e intervalos de confianza ajusta

lines(extended_dates, Est_punt_PEA_ajustada, col="blue")
x_vals <- c(rep(NA, times = 66), extended_dates[67:72])

lines(x_vals, LI_PEA_ajustado, col = "purple")
lines(x_vals, LS_PEA_ajustado, col = "purple")


legend("bottomright", legend = c("PEA original", "Pronóstico puntual calculado con método estático"),
       col = c("black", "blue"), lty = 1)
```



Visualizamos el Data.Frame con los valores históricos la "data original" + los resultados del pronóstico puntual y los intervalos de confianza al 95%

```
df_demanda_PEA_MX_pronostico<-data.frame(PEA_MEX_completa$PEA,
                                                Est_punt_PEA_ajustada, LI_PEA_ajustado, LS_PEA_ajustado)
View(df_demanda_PEA_MX_pronostico)

tail(df_demanda_PEA_MX_pronostico)
```

```
##    PEA_MEX_completa.PEA Est_punt_PEA_ajustada LI_PEA_ajustado LS_PEA_ajustado
## 67             58307446              56848431        53892733        59804129
## 68             58761793              57038992        52844991        61232992
## 69             58085314              57229553        52075812        62383294
## 70             59338419              57415971        51445559        63386383
## 71             59480471              57604460        50907423        64301498
## 72             60145456              57795021        50434515        65155527
```

Se observa que los pronósticos puntuales vs los reales parecen no estar muy alejados unos de otros, sin embargo los intervalos de confianza si "sufren a medida que nos alejamos del último valor real" pues son cada vez más grandes hasta llegar a abarcar un rango de [50,000000 a 65,000000] para el 95% de intervalo de confianza en el 6t0 valor pronósticado. Lo que representa alrededor del 25% del rango de nuestra variable y.

A.5 Calculo del error tipo MAPE para los pronósticos hechos con los modelos estáticos de suavizamiento

Calculemos el error tipo MAPE entre los valores pronosticados con el "método de suavizamiento estático" y los valores reales de la serie

```r
##Primero formemos un vector con los valores reales
PEA_MEX_completa_reales6 <- PEA_MEX_completa$PEA[67:72]
PEA_MEX_completa_reales6
```

```
## [1] 58307446 58761793 58085314 59338419 59480471 60145456
```

```r
## Despues formemos un vector con únicamente los valores pronósticados
PEA_MEX_forecast6 <- Est_punt_PEA_ajustada[67:72]
PEA_MEX_forecast6
```

```
##       67       68       69       70       71       72
## 56848431 57038992 57229553 57415971 57604460 57795021
```

```r
##Calculemos el MAPE entre estos dos vectores

# First, calculate the absolute percentage error (APE) for each element in the vectors
APE_PEA_estatico <- abs((PEA_MEX_completa_reales6 - PEA_MEX_forecast6) / PEA_MEX_completa_reales6)

# Calculate the mean APE across all elements in the vectors
MAPE_PEA_estatico <- mean(APE_PEA_estatico) * 100

# Print the MAPE value
cat("El error tipo MAPE  para el forecast calculado con el método de 'suavizamiento estático'de la seri
```

```
## El error tipo MAPE  para el forecast calculado con el método de 'suavizamiento estático'de la serie
```

B. Método de Holt como "modelo de suaviazamiento y pronóstico NO ESTÁTICO" que si considera los cambios de nivel a tráves del tiempo.

Vamos a usar ahora el método de HoltWinters como "modelo de suaviazamiento y pronóstico NO ESTÁTICO" que si considera los cambios de nivel a tráves del tiempo.

Para ello debemos tener nuestra data a suavizar en el tipo de dato serie de tiempo ts en R

```r
PEA_MEX_ts <-ts(PEA_mex$PEA, frequency =4, start =c(2005,1))
head(PEA_MEX_ts)
```

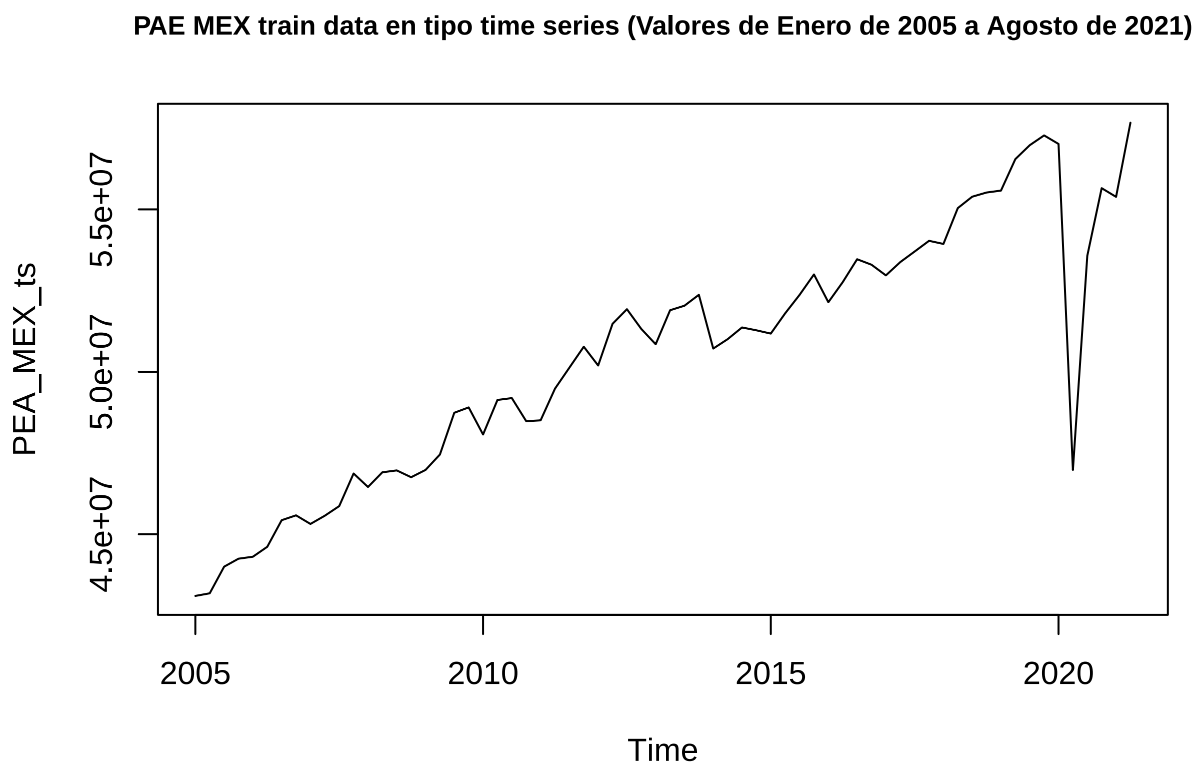
```
##          Qtr1     Qtr2     Qtr3     Qtr4
## 2005 43099847 43180433 44000204 44245519
## 2006 44306012 44611672
```

```r
tail(PEA_MEX_ts)
```

```
##          Qtr1     Qtr2     Qtr3     Qtr4
## 2020 57014967 46978848 53571791 55653440
## 2021 55385133 57668254
```

Graficamos la “data original de la PEA__MEX” que se usará como “train para el modelo” y que fue convertida a tipo de dato serie de tiempo

```
plot(PEA_MEX_ts, main="PAE MEX train data en tipo time series (Valores de Enero de 2005 a Agosto de 202
```



Aplicación del método de Holt Winters” a la serie CON TENDENCIA y SIN ESTACIONALIDAD de PEA MEX

Como se observará este primer cálculo dejará como grados de libertad los valores de alpha, gamma y betha, así como los valores de arranque, de manera que el “el algoritmo de HoltWinters de R” determine la “mejor combinaión posible”. Como en este caso no tenemos componente de seasonalidad se indica que el parámetro gamma tendrá un valor “false”

```
PEA_MEX_ts_h1 <-HoltWinters(PEA_MEX_ts, gamma=FALSE)
PEA_MEX_ts_h1
```

```
## Holt-Winters exponential smoothing with trend and without seasonal component.
##
## Call:
## HoltWinters(x = PEA_MEX_ts, gamma = FALSE)
##
## Smoothing parameters:
##  alpha: 0.2764468
##  beta : 0.0237071
```

```
##  gamma: FALSE
##
## Coefficients:
##         [,1]
## a 55825043.5
## b   154038.1
```

Interpretación de los valores de los coeficientes del método de Holt Winters para el caso de la serie de PEA en México Interpretación de los valores de coficientes α y β para el caso del pronóstico usando modelo de HoltWinters SIN ESPECIFICAR DATOS DE ARRANQUE

$$0 < \alpha < 1$$

Recordemos que: Valores de α cercano a 1 reflejan caminata aleatoria Valores de α cercano a 0 reflejan una "gran influencia" de los valores pasado, es decir, que "el ruido no es un componente de gran relevancia"

En nuestro caso el valor de α es:

```
PEA_MEX_ts_h1$alpha
```

```
##     alpha
## 0.2764468
```

Lo que indica que "La data historica SI tiene UNA RELEVANCIA SIGNIFICATIVA vs el componente meramente estocástico"

$$0 < \beta < 1$$

Recordemos que: Valores de β cercano a 1 reflejan cambios CONSTANTES de nivel a lo largo de la serie Valores de β cercano a 0 reflejan POCOS cambios de nivel a lo largo de la serie

En nuestro caso el valor de β es:

```
PEA_MEX_ts_h1$beta
```

```
##      beta
## 0.0237071
```

Lo que indica muy pocos cambios de nivel a lo largo de la serie o una especie de "pendiente constante"

Y entonces para el caso de valores de α cercanos a 0 y de β cercanos a cero (como es nuestra serie de demanda_electrico) indica una serie que otorga mucha relevancia a los valores históricos y con CAMBIOS DE NIVEL MUY BAJOS

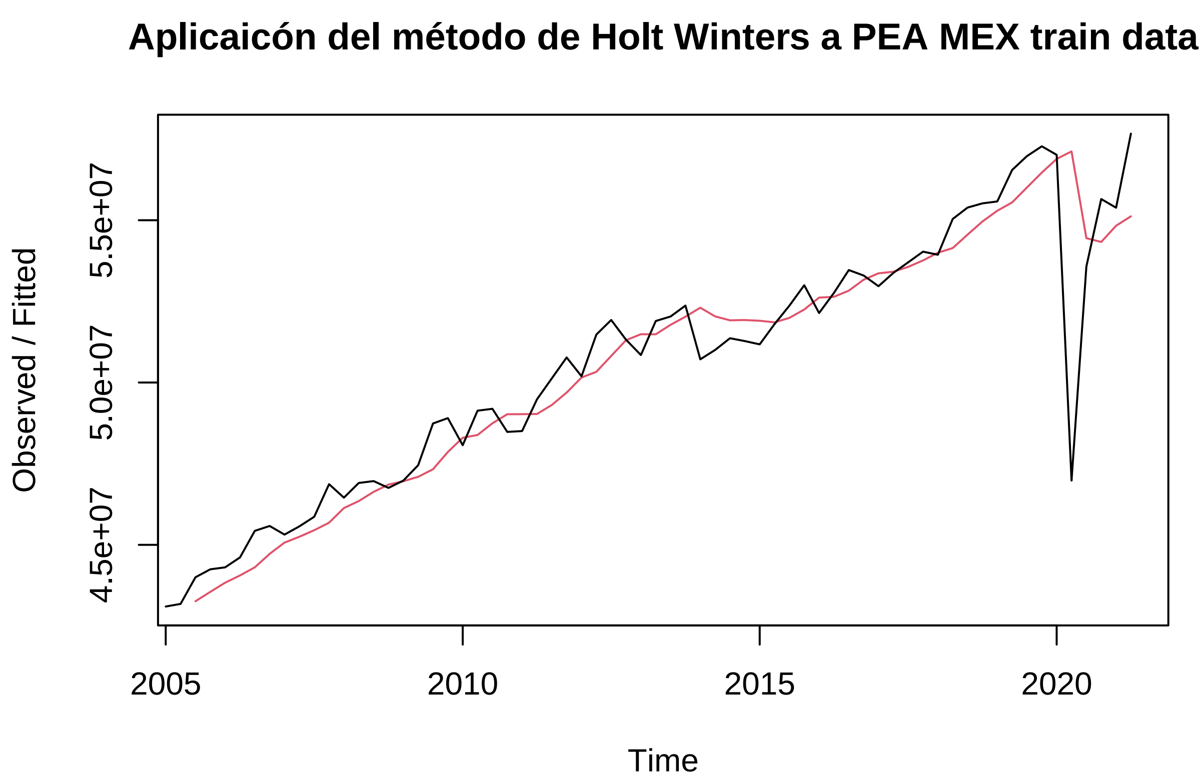
Afortunadamente no obtuvimos simultaneamente valores deα cercanos a 1 y de β cercanos a 1 pues eso reflejaría una CAMINATA ALEATORIA con tendencia estocástica que sería prácticamente imposible de pronósticar.

Pronóstico de la serie de PEA en MEX con el método de Holt Winters

Ahora pasemos a realizar los pronósticos de la serie pero con el método de HoltWinters que considera cambios dinámicos de nivel

Graficamos los valores originales vs suavizados con el método de Holt-Winters

```r
plot(PEA_MEX_ts_h1, main="Aplicaicón del método de Holt Winters a PEA MEX train data")
```




```r
# Load the forecast package
library(forecast)
```

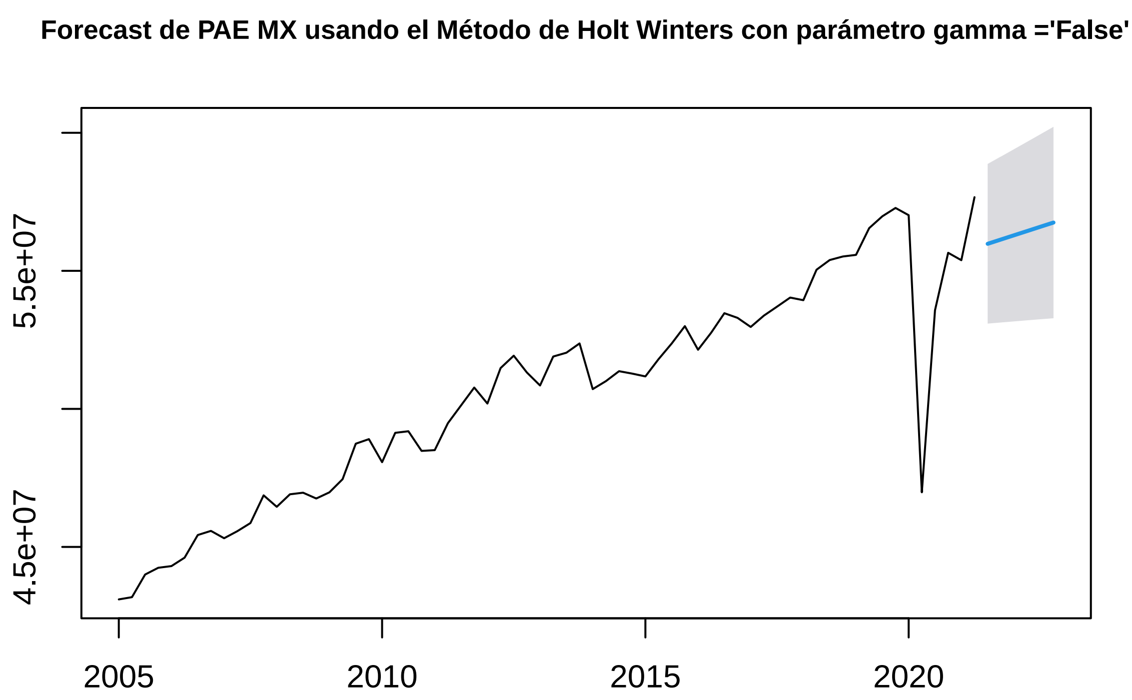
Ahora se llevará a cabo el pronóstico con la librería de forecast y basados en el método de Holt Winters. De igual manera se usará un intervalo de confianza del 95%

```r
PEA_MEX_ts_h1_fc <- forecast(PEA_MEX_ts_h1, h=6, level=0.95)
PEA_MEX_ts_h1_fc
```

```
##         Point Forecast    Lo 95    Hi 95
## 2021 Q3       55979082 53089359 58868804
## 2021 Q4       56133120 53129908 59136331
## 2022 Q1       56287158 53169562 59404753
## 2022 Q2       56441196 53208307 59674085
## 2022 Q3       56595234 53246128 59944340
## 2022 Q4       56749272 53283015 60215529
```

Se observa como el pronóstico puntual es “relativamente bueno” y es muy parecido en valor al obtenido previamente con el método estático, sin embargo los Intervalos de confianza obtenidos con este método son BASTANTE MEJORES en comparación con los del anterior

```
plot(PEA_MEX_ts_h1_fc, main="Forecast de PAE MX usando el Método de Holt Winters con parámetro gamma ='I
```



Veamos el valor la Suma de Errores al Cuadrado del pronóstico de PEA en MEX usando el Método de suaviazamiento dinámico de HoltWinters

```
PEA_MEX_ts_h1$SSE
```

```
## [1] 1.38911e+14
```

Calculemos el error tipo MAPE entre los valores pronosticados con el "método de Holt Winters" y los valores reales de la serie

```
## Formemos un vector con únicamente los valores reales de los ultimos 6 trimestres
PEA_completo_realest6 <- PEA_MEX_completa$PEA[67:72]
PEA_completo_realest6
```

```
## [1] 58307446 58761793 58085314 59338419 59480471 60145456
```

```
## Formemos un vector con únicamente los valores pronósticados
PEA_HW_forecast6 <- PEA_MEX_ts_h1_fc$mean
PEA_HW_forecast6
```

```
##          Qtr1     Qtr2     Qtr3     Qtr4
## 2021                   55979082 56133120
## 2022 56287158 56441196 56595234 56749272
```

```r
##Calculemos el MAPE entre estos dos vectores

# First, calculate the absolute percentage error (APE) for each element in the vectors
APE_PEA_HW <- abs((PEA_completo_realest6 - PEA_HW_forecast6) / PEA_completo_realest6)

# Calculate the mean APE across all elements in the vectors
MAPE_PEA_HW <- mean(APE_PEA_HW) * 100

# Print the MAPE value
cat("MAPE error para el forecast calculado con el método de 'HoltWinters'es:", MAPE_PEA_HW, "%")
```

```
## MAPE error para el forecast calculado con el método de 'HoltWinters'es: 4.490383 %
```

Error cuadrado calculado con el Método de HoltWinters vs el calculado con el método de suafizamiento estático

```r
PEA_MEX_ts_h1$SSE
```

```
## [1] 1.38911e+14
```

Entonces el error asociado al pronóstico del cálculo de pronóstico PUNTUAL para el caso Demanda_eléctirco es:

```r
sse_PEA_hw <- ((PEA_MEX_ts_h1$SSE)/(length(PEA_MEX_ts)))^0.5
sse_PEA_hw
```

```
## [1] 1450762
```

Recordar que la SSE de la demanda_electrico para el caso del método de suavizamiento estático era:

```r
SSE_PEA_E <- sum(PEA_MX_ruido_blanco^2)
SSE_PEA_E
```

```
## [1] 66.03792
```

Lo que indica que el método de suavizamiento estático resultó bastante mejor en nuestro caso para el PRONOSTICO PUNTUAL de la demanda de gas natural en el sector eléctrico en México vs el método de HoltWinters

Comparación del MAPE del pronostico hecho con métodos estáticos vs el pronóstico hecho con el método " de suavizamiento dinámico" de HoltWinters para el pronostico puntual

En este caso y dado un cálculo con el métodO de HolttWinters sin valores de arranque especificados los errores de ambos métodos (estático vs HoltWinters) fueron los sigientes:

```r
cat("MAPE error para el forecast calculado con el método de 'suavizamiento estático'es:", MAPE_PEA_esta
```

```
## MAPE error para el forecast calculado con el método de 'suavizamiento estático'es: 2.868186 %
```

```r
cat("MAPE error para el forecast calculado con el método de 'HoltWinters'es:", MAPE_PEA_HW, "%")
```

```
## MAPE error para el forecast calculado con el método de 'HoltWinters'es: 4.490383 %
```

Lo que confirma que el método de suavizamiento estático resultó mejor en nuestro caso para el PRONOSTICO PUNTUAL de la PEA en México.

Comparación de los intervalos del pronostico hecho con métodos estáticos vs el pronóstico hecho con el método " de suavizamiento dinámico" de HoltWinters

Intervalos calculados con el método de suavizamiento estático

```r
LI_PEA_ajustado[67:72]
```

```
##       67       68       69       70       71       72
## 53892733 52844991 52075812 51445559 50907423 50434515
```

```r
LS_PEA_ajustado[67:72]
```

```
##       67       68       69       70       71       72
## 59804129 61232992 62383294 63386383 64301498 65155527
```

```r
intervalos_metodo_E<- data.frame(LI_PEA_ajustado_E =LI_PEA_ajustado[67:72], LS_PEA_ajustado_E= LS_PEA_a

intervalos_metodo_E
```

```
##    LI_PEA_ajustado_E LS_PEA_ajustado_E
## 67          53892733          59804129
## 68          52844991          61232992
## 69          52075812          62383294
## 70          51445559          63386383
## 71          50907423          64301498
## 72          50434515          65155527
```

Intervalos calculados con el método de suavizamiento dinámico Holt-Witers

```r
PEA_MEX_ts_h1_fc
```

```
##         Point Forecast    Lo 95    Hi 95
## 2021 Q3       55979082 53089359 58868804
## 2021 Q4       56133120 53129908 59136331
## 2022 Q1       56287158 53169562 59404753
## 2022 Q2       56441196 53208307 59674085
## 2022 Q3       56595234 53246128 59944340
## 2022 Q4       56749272 53283015 60215529
```

En este caso podemos de nuevo notar como el método de HoltWinters otorga mejores intervalos de confianza para pronósticos incluso para aquellos 'más alejados' del último valor real vs aquellos generados con el "método de suavizamiento estático"

Serie de tiempo 3 CON tendencia y CON estacionalidad (Historico mensual de Demanda de Gas Natural para el sector eléctrico en México)

Tenemos data mensual de Enero de 2005 a Agosto de 2022 sobre la demanda de gas natural para el sector eléctrico en México, son 213 datos en total. La idea es usar 201 de esos 2013 como una especie de "data de entreaniemto" para aplicar los modelos A) De suavizamiento "estático" y B) De "suavizamiento dinámico" o Holt-Winters para posteriormente predecir con cada modelo los valores de los 12 meses restantes y poder comparar contra los reales e identificar el mejor modelo entre los dos mencionados para modelar la demanda de gas natural en sector eléctrico.

Primero vamos a importar la data con los elementos del 1 al 201

```
library("readxl")
Demanda_electrico <-read_excel("Demanda_electrico_2022_full1_pp.xlsx")
```

Revisamos que la data se haya importado correctamente

```
head(Demanda_electrico)
```

```
## # A tibble: 6 x 2
##   Date                Demanded_Gas
##   <dttm>                     <dbl>
## 1 2005-01-01 00:00:00        1820.
## 2 2005-02-01 00:00:00        1895.
## 3 2005-03-01 00:00:00        1766.
## 4 2005-04-01 00:00:00        1643.
## 5 2005-05-01 00:00:00        1896.
## 6 2005-06-01 00:00:00        2052.
```

```
summary(Demanda_electrico)
```

```
##       Date                          Demanded_Gas
##  Min.   :2005-01-01 00:00:00.00   Min.   :1561
##  1st Qu.:2009-03-01 00:00:00.00   1st Qu.:2585
##  Median :2013-05-01 00:00:00.00   Median :2947
##  Mean   :2013-05-01 16:21:29.54   Mean   :2999
##  3rd Qu.:2017-07-01 00:00:00.00   3rd Qu.:3359
##  Max.   :2021-09-01 00:00:00.00   Max.   :5168
```

```
typeof(Demanda_electrico)
```

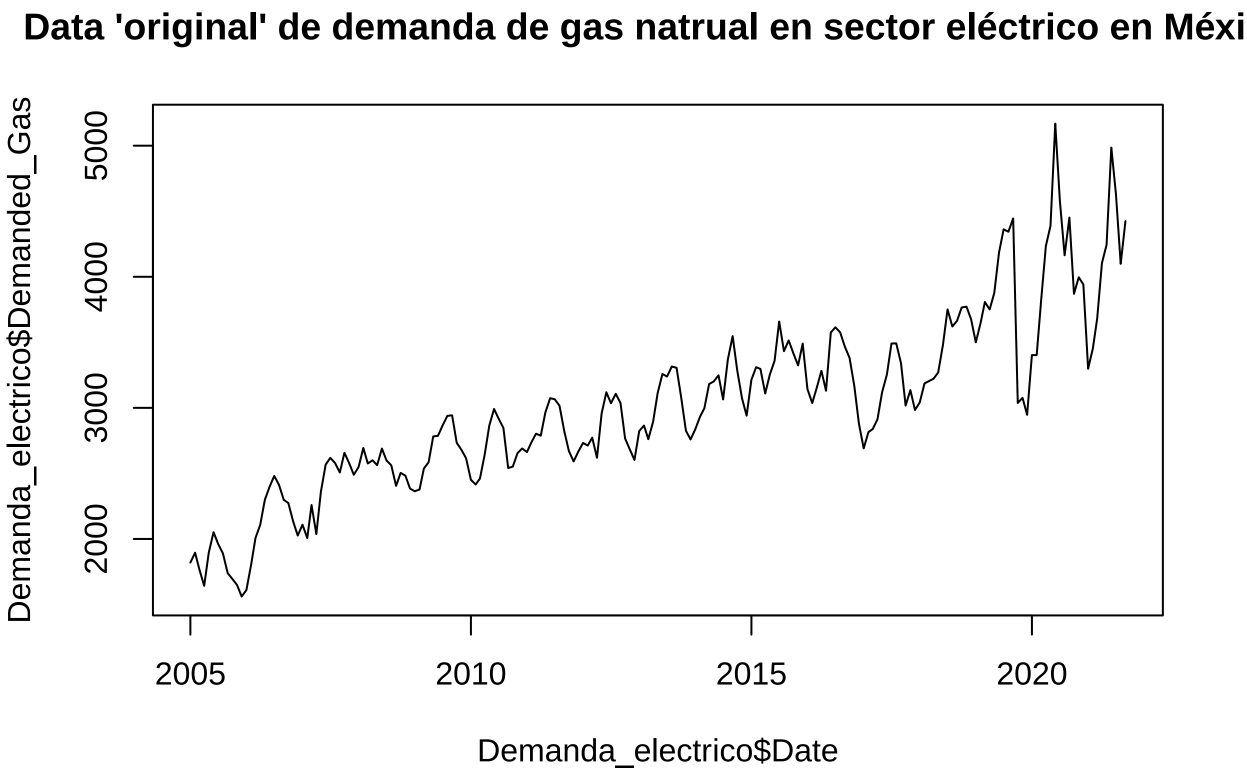
```
## [1] "list"
```

```
dim(Demanda_electrico)
```

```
## [1] 201   2
```

Graficamos la "data original" de demanda de gas natural en el sector eléctrico del año 2005 a Agoato de 2021

```
plot(Demanda_electrico$Date, Demanda_electrico$Demanded_Gas, type = "l", main="Data 'original' de demand
```



A. Método de "suavizamiento estático"

Comenzaremos tratando la serie con el método estático aplicando el siguiente flujo lógico de trabajo y considerando que estamos ene el caso de serie CON TENDECIA y CON ESTACIONALIDAD:

1. Se removerá el factor de tendencia a la serie original obteniendo una serie sin tendencia
2. A la serie sin tendencia se le removerán los Factores de Estacionalidad obtendiendo únicamente el componente de ruido blanco o componente estocástico de la serie
3. Será sobre el "ruido blanco" que se llevará a cabo el "pronóstico puntual" de los siguientes 12 meses como la media historica de dicho componente estocástico
4. Se calcularán los intervalos de pronóstico al 95% de confianza
5. Se "devolverán" los efectos estacionales y de tendencia de la serie con el fin de llevar los pronósticos a las dimensiones correctas
6. Se calculará el error tipo MAPE entre los valores pronósticados con el "método estático" y los valores reales de la serie de los últimos 12 meses

A.1 Remoción de tendencia de serie original

Primero se calculará la linea "general" de tendencia de nuestra serie de Electrico_Demanded_gas_line (que representa la tendencia de la "serie de datos originales") y se removerá dicha tendencia de la 'data original' de demanda_electrico

Calculemos la linea de tendencia

```
Electrico_Demanded_gas_line<-lm(Demanded_Gas~Date, data=Demanda_electrico)

Electrico_Demanded_gas_line
```

```
##
## Call:
## lm(formula = Demanded_Gas ~ Date, data = Demanda_electrico)
##
## Coefficients:
## (Intercept)         Date
##  -2.287e+03    3.865e-06
```

Se puede observar una "ecuación general de linea de tendencia" en este caso es: y = 2,287 + 0.0000003865 * (Date) Es muy razonable que "la ordenada al origen" sea de alrededor de 2 mil, pues todos los datos de Demanda de gas en sector eléctrico son del orden de 2 a 5 mil. También resulta lógica que el valor de la pendiente sea positivo y muy cercano a cero, pues en general para cada aumento en el valor del tiempo, es decir, para cada més de datos no existe un incremento o definición realmente signifiativo en el valor del PEA (excepto en el año 2020)

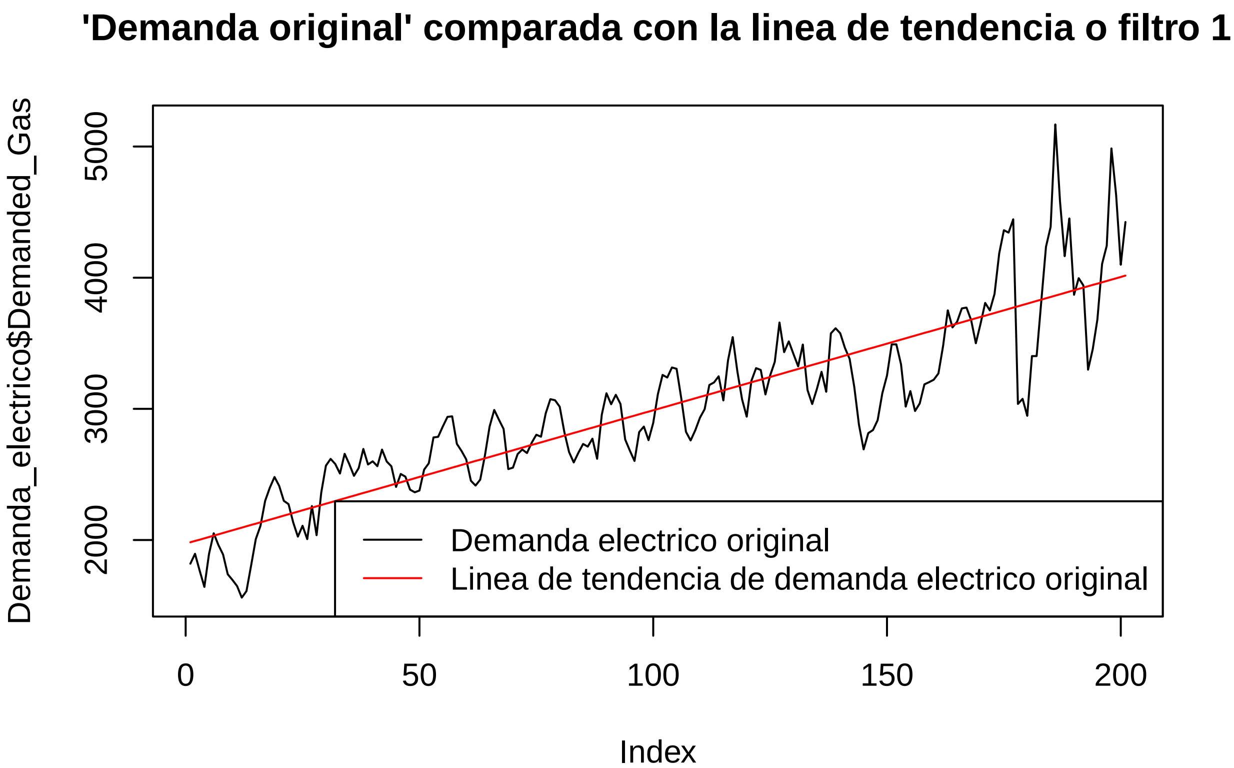
Ahora para graficar la "serie original" de demanda_electrico vs el filtro1 (Linea de tendencia) primero debemos "crear/simular" datos de Enero 2005 a Agosto 2021 a partir de nuestra linea de previamente computada de nombre: Electrico_Demanded_gas_line

```
Electrico_Demanded_gas_line_aplicada <- predict(Electrico_Demanded_gas_line, newdata = data.frame(Date =

head(Electrico_Demanded_gas_line_aplicada)
```

```
##        1        2        3        4        5        6
## 1982.917 1993.270 2002.621 2012.974 2022.993 2033.346
```

Graficamos entonces a la "data original de demanda electrico" y a la data calculada a partir de la linea tendencia

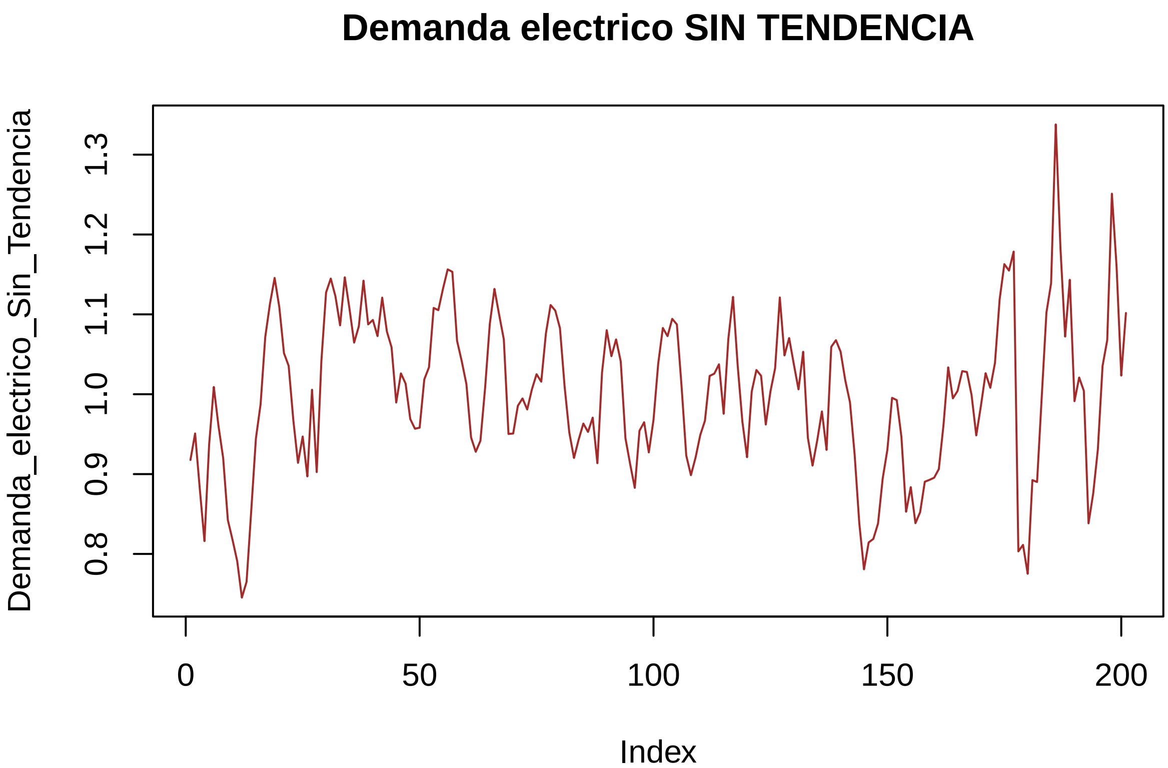
```
plot(Demanda_electrico$Demanded_Gas, type = "l", main = "'Demanda original' comparada con la linea de te

lines(Electrico_Demanded_gas_line_aplicada, col = "red")


legend("bottomright", legend = c("Demanda electrico original", "Linea de tendencia de demanda electrico
       col = c("black", "red"), lty = 1)
```



Ahora vamos a "quitar el efecto de la tendencia" de nuestra "serie original de datos" En este caso se asumirá un "efecto multiplicativo" y por lo tanto se removerá dividiendo cada elemento de la serie original / valor de la linea de tendencia en ese valor de xi

Se grafica la serie sin tendencia:

```
Demanda_electrico_Sin_Tendencia<-Demanda_electrico$Demanded_Gas/Electrico_Demanded_gas_line_aplicada

plot(Demanda_electrico_Sin_Tendencia, type="l", main="Demanda electrico SIN TENDENCIA", col='brown')
```




```
head(Demanda_electrico_Sin_Tendencia)
```

```
##         1         2         3         4         5         6
## 0.9176279 0.9508646 0.8817743 0.8160561 0.9369977 1.0090362
```

A.2 Remoción de los factores estacionales a la serie sin tendencia

Se procederá al cálculo de los Factores Estacionales de la serie.

La idea de la asignación de estaciones en la presente serie de demanda de gas natrual en el sector eléctrico es la siguiente: De manera historica el consumo de electricidad aumenta en los meses más calurosos (abril, mayo, junio) y también en los meses de extremos frio (dic, enero), contrario a los meses más 'templados' (Agosto, Sept) donde el consumo suele ser menor. Por lo tanto se propuso que el consumo de gas natural en el sector eléctrico en México puede seguir un comportamiento estacional trimestral.

De esta manera se procede a "crear las listas de los valores i-esimos de la serie" que corresponderán a los meses 1,2,3 para el triemestre 1 (T1), 4,5,6 para el triemestre 2 (T2), 7,8,9 para el triemestre 3 (T3), 10,11,12 para el triemestre 4 (T4) y de nuevo 13,14,15 para el triemestre 1 (T1), así sucesivamente hasta tener los 212 meses de datos distribuidos en los 4 trimetres.

```
# Generate indices for T1, T2, T3, and T4
# Define the indices for each subset
T1_indices <- c(seq(1, 3), seq(13, 15), seq(25, 27), seq(37, 39), seq(49, 51), seq(61, 63), seq(73, 75)
```

```r
T2_indices <- c(seq(4, 6), seq(16, 18), seq(28, 30), seq(40, 42), seq(52, 54), seq(64, 66), seq(76, 78)

T3_indices <- c(seq(7, 9), seq(19, 21), seq(31, 33), seq(43, 45), seq(55, 57), seq(67, 69), seq(79, 81)

T4_indices <- c(seq(10, 12), seq(22, 24), seq(34, 36), seq(46, 48), seq(58, 60), seq(70, 72), seq(82, 8


# Create the subsets
T1 <- Demanda_electrico_Sin_Tendencia[T1_indices]
head(T1)
```

```
##         1         2         3        13        14        15
## 0.9176279 0.9508646 0.8817743 0.7651265 0.8539887 0.9446277
```

```r
T2 <- Demanda_electrico_Sin_Tendencia[T2_indices]
head(T2)
```

```
##         4         5         6        16        17        18
## 0.8160561 0.9369977 1.0090362 0.9875578 1.0718025 1.1130431
```

```r
T3 <- Demanda_electrico_Sin_Tendencia[T3_indices]
head(T3)
```

```
##         7         8         9        19        20        21
## 0.9602052 0.9201165 0.8425581 1.1456434 1.1091477 1.0514739
```

```r
T4 <- Demanda_electrico_Sin_Tendencia[T4_indices]
head(T4)
```

```
##        10        11        12        22        23        24
## 0.8176693 0.7910745 0.7452891 1.0354835 0.9673978 0.9141021
```

Se calculan ahora los Factores Estacionarios de la Serie dividiendo la media de aquellos valores asignados a cada trimestre (T1, T2, T3 y T4) / la media de nuestra seria sin tendencia. Siendo ambos un valor escalar, serán entonces los FEi valores escalares

```r
FE1_electrico<-mean(T1)/mean(Demanda_electrico_Sin_Tendencia)
FE1_electrico
```

```
## [1] 0.9432222
```

```r
FE2_electrico<-mean(T2)/mean(Demanda_electrico_Sin_Tendencia)
FE2_electrico
```

```
## [1] 1.032675
```

```r
FE3_electrico<-mean(T3)/mean(Demanda_electrico_Sin_Tendencia)
FE3_electrico
```

```
## [1] 1.066229
```

```r
FE4_electrico<-mean(T4, na.rm = TRUE)/mean(Demanda_electrico_Sin_Tendencia)
FE4_electrico
```

```
## [1] 0.9552412
```

Interpretación de los Factores estacionales FE de la serie de demanda de gas natural en sector eléctrico en México Recordar que en general los factores estacionales nos dan una idea del PORCENTAJE DE AUMENTO o DISMINUCIÓN PROMEDIO, de la variable estudiada a lo largo de los periodos/ seasonalidades establecidas. En este caso un factor estacional 1 (que representa los primeros 3 meses del años) con un valor de

```r
FE1_electrico
```

```
## [1] 0.9432222
```

y un factor estacional 4 (que representa los ultimos 3 meses del años) con un valor de

```r
FE4_electrico
```

```
## [1] 0.9552412
```

Indica que en promedio se tiene prácticamente la misma cantidad de demanda de gas natural en el sector eléctrico en estos periodos (durante los últimos 3 meses de cada año y los primeros 3 del siguiente)

Sin embargo un valor de factor estacional 2 (que representa los meses Abril, Mayo, Junio) indica que hay un aumento aproximado del 9% en la cantidad de gas natural demanda en el sector eléctrico durante los meses de Abril, Mayo, Junio, respecto a lo observado en los de Enero, Febrero y Marzo

```r
FE2_electrico
```

```
## [1] 1.032675
```

Y un factor estacional 3 (que representa los meses Julio, Agosto, Sept) con respecto al factor 4 indica que en promedio se demanda un 10% menos de gas natural en el sector eléctrico los últomos 3 meses del año respecto a los 3 meses anteriores.

```r
FE3_electrico
```

```
## [1] 1.066229
```

```r
FE4_electrico
```

```
## [1] 0.9552412
```

A.3 Tratamiendo del ruido blanco en la serie y pronostico puntual

Una vez calculados los factores estacionales de la serie se puede obtener el unicamente el 'componente estocástico o ruido blanco de la serie' que se calcularía "quitando" el efecto de los factores estacionales previamente calculados a nuestra serie YA SIN tendencia

Esto bajo el siguiente razonamiento:

Serie total = Componente de tendencia + Componente de estacionalidad + Componente estocástico

por lo tanto al haber "extraido" los componentes de tendencia y estacionalidad tenemos únicamente el componente estocástico de la serie

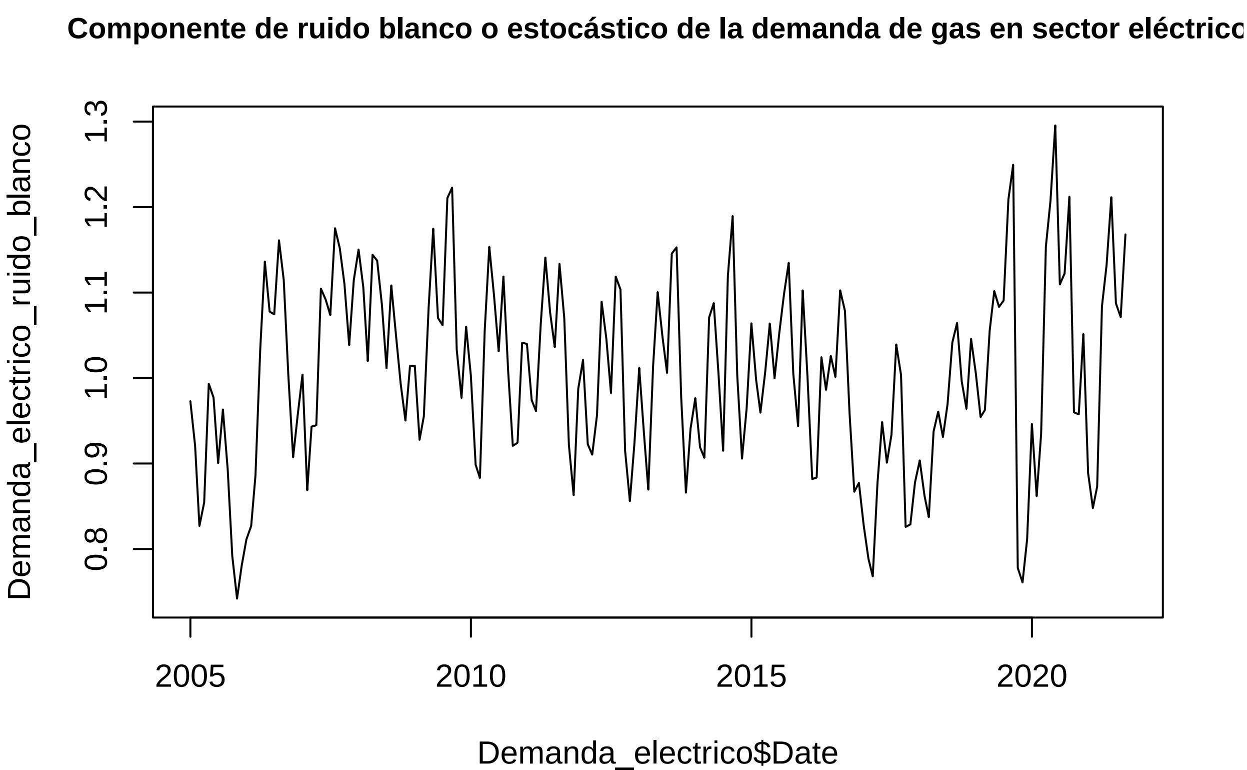
```
Demanda_electrico_ruido_blanco<-Demanda_electrico_Sin_Tendencia/c(FE1_electrico, FE2_electrico, FE3_ele
```

```
## Warning in Demanda_electrico_Sin_Tendencia/c(FE1_electrico, FE2_electrico, :
## longer object length is not a multiple of shorter object length
```

```
head(Demanda_electrico_ruido_blanco)
```

```
##         1         2         3         4         5         6
## 0.9728650 0.9207783 0.8270029 0.8542933 0.9934008 0.9771093
```

```
plot(Demanda_electrico$Date, Demanda_electrico_ruido_blanco, type="l",
     main="Componente de ruido blanco o estocástico de la demanda de gas en sector eléctrico",
     cex.main=0.9)
```



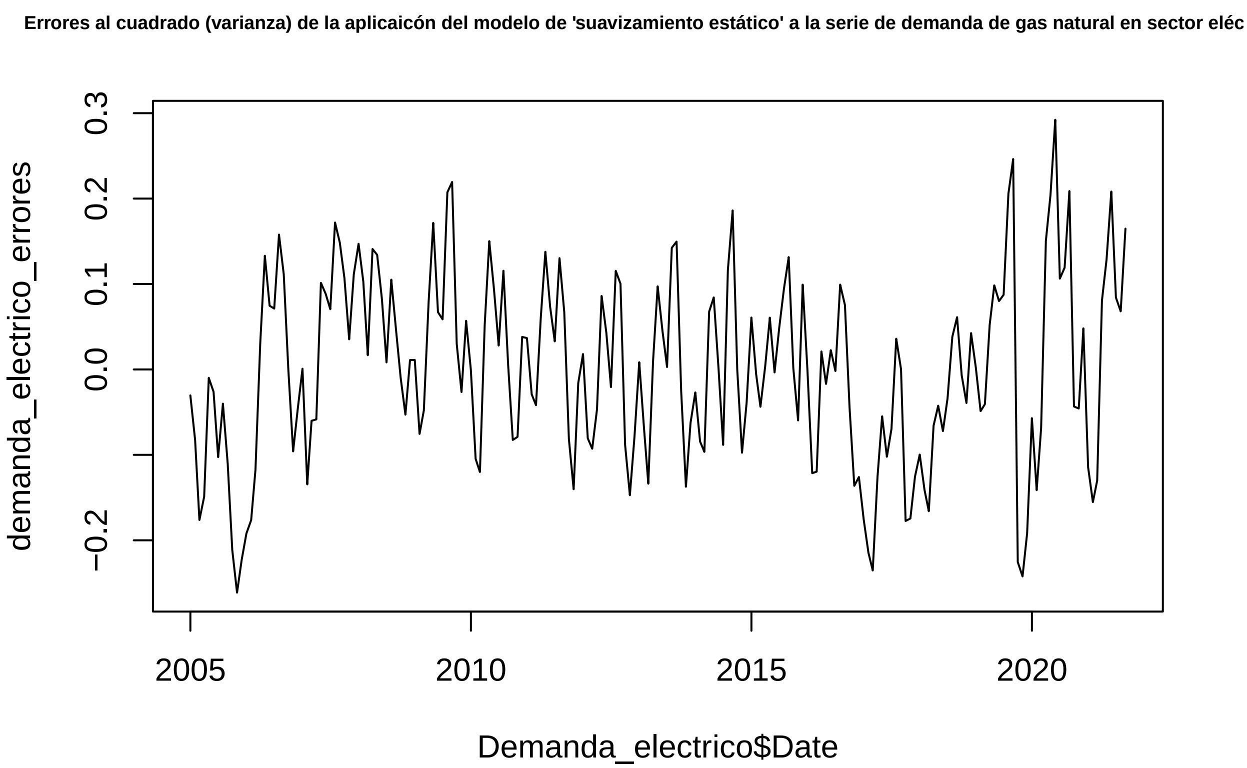
Una vez teniendo únicamente el componente estocástico de nuestra serie, es posible "extraer los errores" de la serie y conocer así el error estimado (varianza) y la desviación estandar. La idea es que extrayendo de la serie del ruido blanco su media histórica, estaríamos determinando la varianza, de la siguiente manera:

Varianza (errores al cuadrado) = Serie original - estacionalidad - tendecia - media de componete estocástico. En nuestro caso, dichos errores será nuestra serie sin tendencia ni estacionalidad - mean(Demanda_electrico_ruido_blanco). Nosotros hemos ya calculado el resultado de Serie original - estacionalidad - tendecia y lo hemos llamada justamente ruido blanco y lo lo tanto el calculo del error podría resumirse a

Errores= Componente estocástico o ruido blanco - Media del componente estocástico o ruido blanco

Determinemos entonces la varianza (errores al cuadrado) para nuestra serie de demanda_electrico

```
## Se calculan los errores como Errores= Componente estocástico o ruido blanco - Media del componente e
demanda_electrico_errores <-Demanda_electrico_ruido_blanco-mean(Demanda_electrico_ruido_blanco)

plot(Demanda_electrico$Date, demanda_electrico_errores, type="l", main="Errores al cuadrado (varianza) d
```



Se calcula la media de los errores

```
mean(demanda_electrico_errores)
```

```
## [1] 9.003955e-17
```

Se calcula la varianza y desviación estandar de los errores

```r
## Se calcula la varianza
var(demanda_electrico_errores)
```

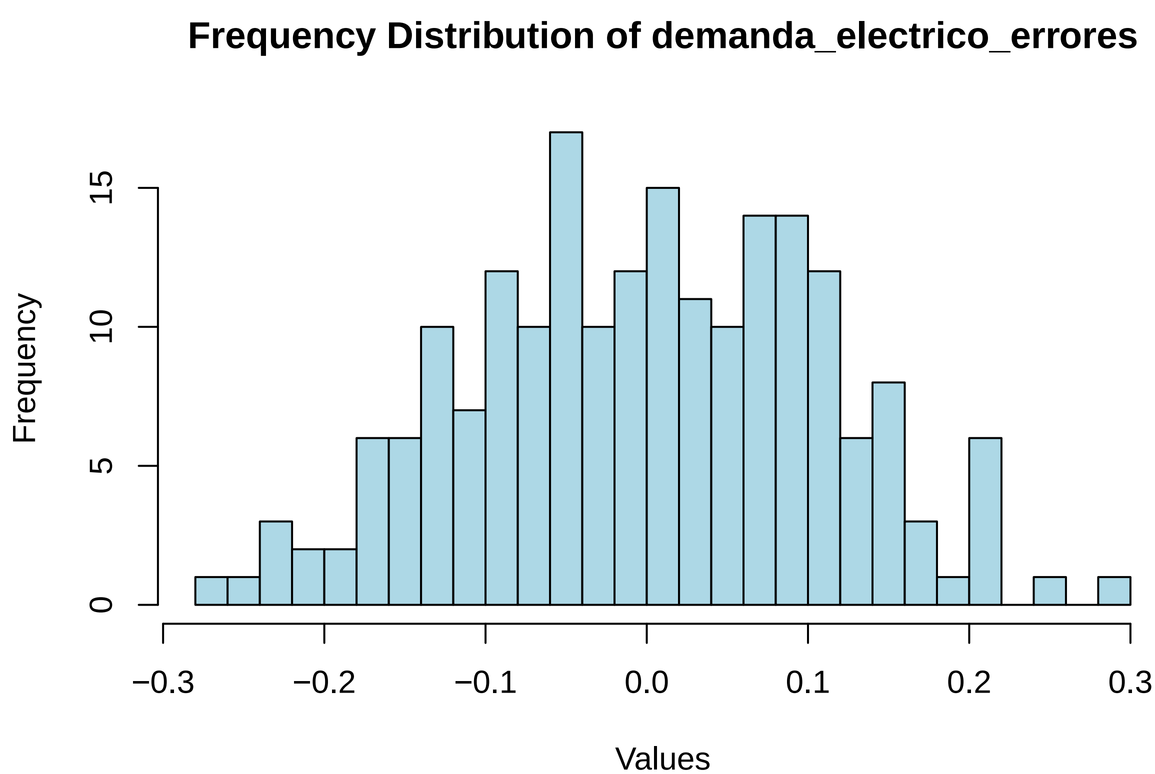
```
## [1] 0.01187763
```

```r
## Se calcula la desviación estandar
sd_est_demanda_electrico <-sd(demanda_electrico_errores)
```

Se tiene una desviación estandar de 0.01 y una media muy cercana a cero para los errores lo cual es totalmente concordante con le hecho de que tras haber removido los componentes de tendencia y estacionalidad, "el ruido blanco" o componente estocástico resultante tiende a tener una distribuión normal. Y que si a ese ruido blanco se le extrae su media tambien tiende a tener una distribución normal con media cero y en este caso una desviación estandar de 0.01

Procedamos a observar como se comporta la Distribución de dichos errores:

```r
hist(demanda_electrico_errores, breaks = 20, col = "lightblue", main = "Frequency Distribution of demand
```



Se comprueba como los errores de la serie siguen una distribución normal con media cero y desviación estandar de 0.1

Pronóstico puntual Una vez teniendo nuestra seríe SIN TENDENCIA y SIN ESTACIONALIDAD lo que como ya se mencionó, implica que, "unicamente tenemos el componente estocástico o de ruido blanco en la serie" podemos proceder a llevar a cabo el pronóstico de 12 meses equivalente 1 año de datos.

Como ha sido estudiado en la clase, nuestro mejor pronóstico puntual dadas estas condiciones en la serie es la media de los datos históricos del ruido blanco

Para ello importamos el Data Frame con la data historica completa, es decir, la que contiene los 201 meses propuestos aquí como "train data" + los 12 meses que se pronósticaran

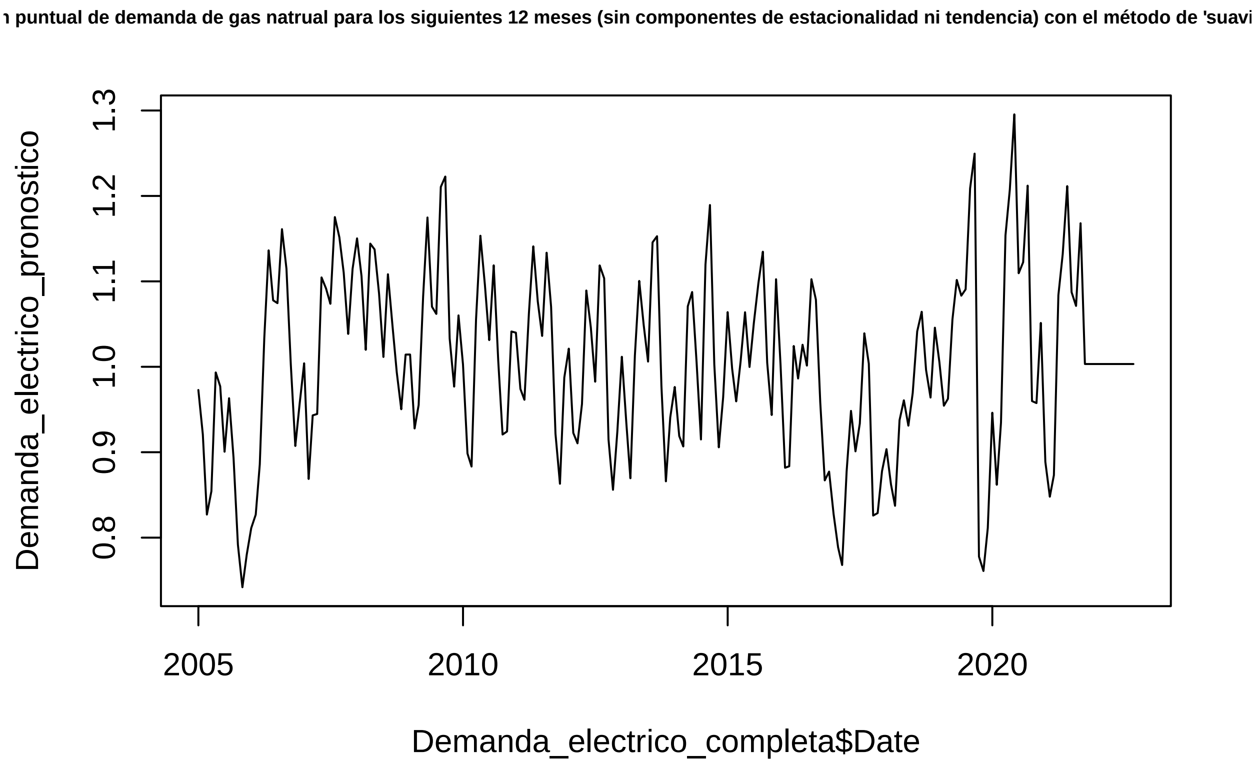
```r
Demanda_electrico_completa <-read_excel("Demanda_electrico_2022_full1.xlsx")
head(Demanda_electrico_completa)
```

```
## # A tibble: 6 x 2
##   Date                Demanded_Gas
##   <dttm>                     <dbl>
## 1 2005-01-01 00:00:00        1820.
## 2 2005-02-01 00:00:00        1895.
## 3 2005-03-01 00:00:00        1766.
## 4 2005-04-01 00:00:00        1643.
## 5 2005-05-01 00:00:00        1896.
## 6 2005-06-01 00:00:00        2052.
```

```r
tail(Demanda_electrico_completa)
```

```
## # A tibble: 6 x 2
##   Date                Demanded_Gas
##   <dttm>                     <dbl>
## 1 2022-04-01 00:00:00        3403.
## 2 2022-05-01 00:00:00        3350.
## 3 2022-06-01 00:00:00        3499.
## 4 2022-07-01 00:00:00        3351.
## 5 2022-08-01 00:00:00        3506.
## 6 2022-09-01 00:00:00        3778.
```

```r
Demanda_electrico_pronostico<-c(Demanda_electrico_ruido_blanco, rep(mean(Demanda_electrico_ruido_blanco)

plot(Demanda_electrico_completa$Date,Demanda_electrico_pronostico, type="l", main = "Estimación puntual
```



Comentario sobre el pronóstico puntual sin tendencia ni estacionalidad, que para el caso de la serie de demanda de gas natural en el sector eléctrico tiene un valor de:

```
mean(Demanda_electrico_ruido_blanco)
```

```
## [1] 1.003239
```

Este valor es bastante concordante con el hecho que. Al haber "quitado los componentes de tendencia y estacionaldad de la serie original", es decir obtenido el ruido blanco los valores de nuestra variable y: gas demandado están en este momento del algoritmo en escala de la unidad: 1. Y por lo tanto la media de dichos valores será muy cercana a 1.

En general se podría decir dados los resultados del presente trabajo, así como los resultados de los ejercicios vistos en clase, que el valor de la media estimada o pronóstico puntual para la serie sin componente de tendencia tiende al valor de 1.

A.4 Calculo de intervalos de confianza para el pronóstico

Calculamos ahora los límites de intervalos de confianza al 95% para el pronostico de los siguientes 12 meses Se considera 95% como +- 2 desviaciones estandar de la media de la serie del ruido

```
LI_demanda_electrico<-Demanda_electrico_pronostico[202:213]-2*sd_est_demanda_electrico*c(1:12)^.5
LS_demanda_electrico<-Demanda_electrico_pronostico[202:213]+2*sd_est_demanda_electrico*c(1:12)^.5

## Tambien se debe "crear la linea para nuestros limites" teniendo NA en los valores de datos historicos
```

```r
LI_demanda_electrico_lineaA<-c(rep(NA, times=201), LI_demanda_electrico)
head(LI_demanda_electrico_lineaA)
```

```
##
## NA NA NA NA NA NA
```

```r
tail(LI_demanda_electrico_lineaA)
```

```
##
## 0.4265476 0.3867299 0.3493323 0.3139608 0.2803179 0.2481725
```

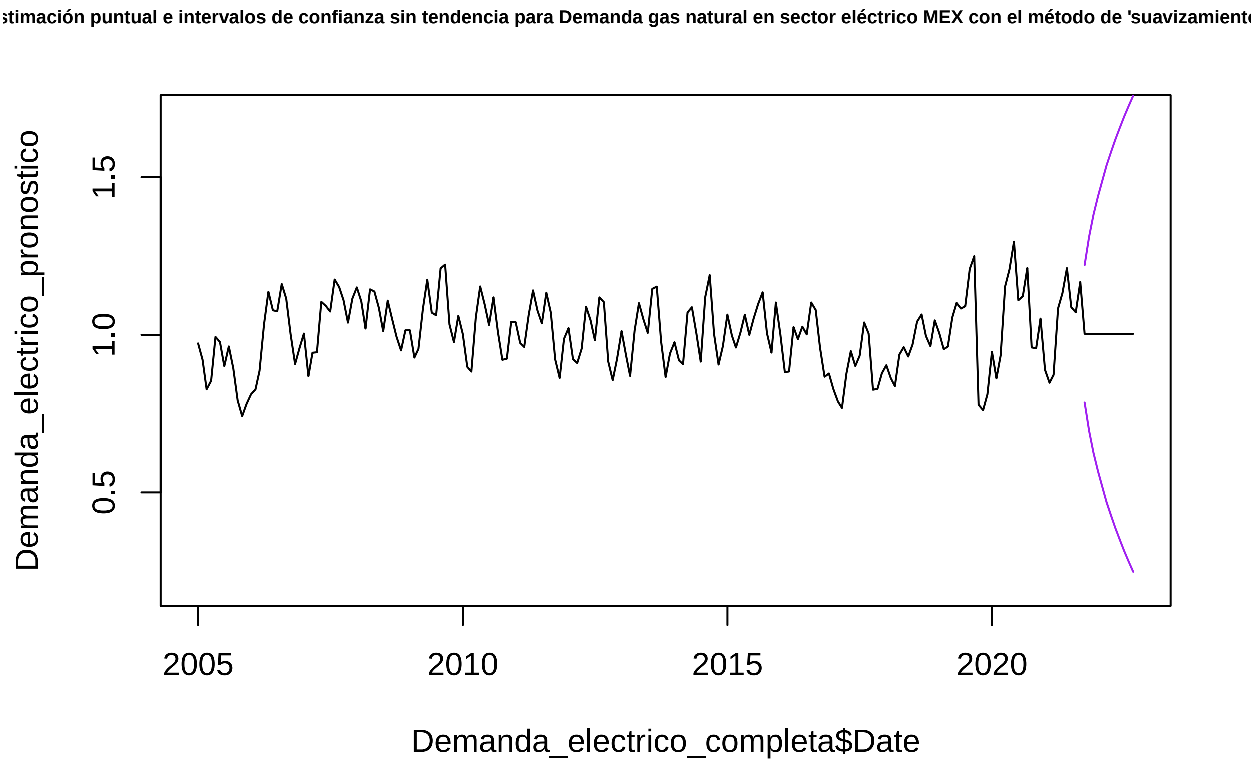
```r
LS_demanda_electrico_lineaA<-c(rep(NA, times=201), LS_demanda_electrico)
head(LS_demanda_electrico_lineaA)
```

```
##
## NA NA NA NA NA NA
```

```r
tail(LS_demanda_electrico_lineaA)
```

```
##
## 1.579931 1.619749 1.657147 1.692518 1.726161 1.758306
```

```r
plot(Demanda_electrico_completa$Date,Demanda_electrico_pronostico, type="l", main = "Estimación puntual


#lines(Demanda_electrico_pronostico, col="Red")
lines(Demanda_electrico_completa$Date, LI_demanda_electrico_lineaA, col="Purple")
lines(Demanda_electrico_completa$Date, LS_demanda_electrico_lineaA, col="Purple")
```



A.5 "Devolución de efectos de tendencia y estacionalidad a los pronósticos e intervalos calculados"

Ahora vamos a "llevar nuestro prOnóstico a las dimensiones reales" Para ello primero vamos a "devolver la esacionalidad y la tendencia a la serie" de nombre demanda_electrico_filtro_sinTendencia_NiEstacionalidad

Empezemos como "devolver estacionalidad" y para ello recordamos que los valores de FE de la seria se calcularon asumiendo efecto multiplicativo y por lo tanto para "devolver ese efecto" se debe multiplicar los valores de Demanda_electrico_pronostico (que contiene los valores historicos de la seria con unicamente el componente ruido + 12 meses pronósticados)* los valores de los Factores Estacionales de la serie

```
# Devolver estacionalidada

Demanda_electrico_pronostico_cest<-Demanda_electrico_pronostico*c(FE1_electrico, FE2_electrico,FE3_elec
```

```
## Warning in Demanda_electrico_pronostico * c(FE1_electrico, FE2_electrico, :
## longer object length is not a multiple of shorter object length
```

```
head(Demanda_electrico_pronostico_cest)
```

```
##         1         2         3         4         5         6
## 0.9176279 0.9508646 0.8817743 0.8160561 0.9369977 1.0090362
```

```r
tail(Demanda_electrico_pronostico_cest)
```

```
##
## 0.9583357 0.9462777 1.0360201 1.0696828 0.9583357 0.9462777
```

```r
LI_demanda_electrico_cest<-LI_demanda_electrico_lineaA*c(FE1_electrico, FE2_electrico,FE3_electrico, FE4
```

```
## Warning in LI_demanda_electrico_lineaA * c(FE1_electrico, FE2_electrico, :
## longer object length is not a multiple of shorter object length
```

```r
head(LI_demanda_electrico_cest)
```

```
##
## NA NA NA NA NA NA
```

```r
tail(LI_demanda_electrico_cest)
```

```
##
## 0.4074558 0.3647722 0.3607467 0.3347541 0.2677712 0.2340818
```

```r
LS_demanda_electrico_cest<-LS_demanda_electrico_lineaA*c(FE1_electrico, FE2_electrico,FE3_electrico, FE4
```

```
## Warning in LS_demanda_electrico_lineaA * c(FE1_electrico, FE2_electrico, :
## longer object length is not a multiple of shorter object length
```

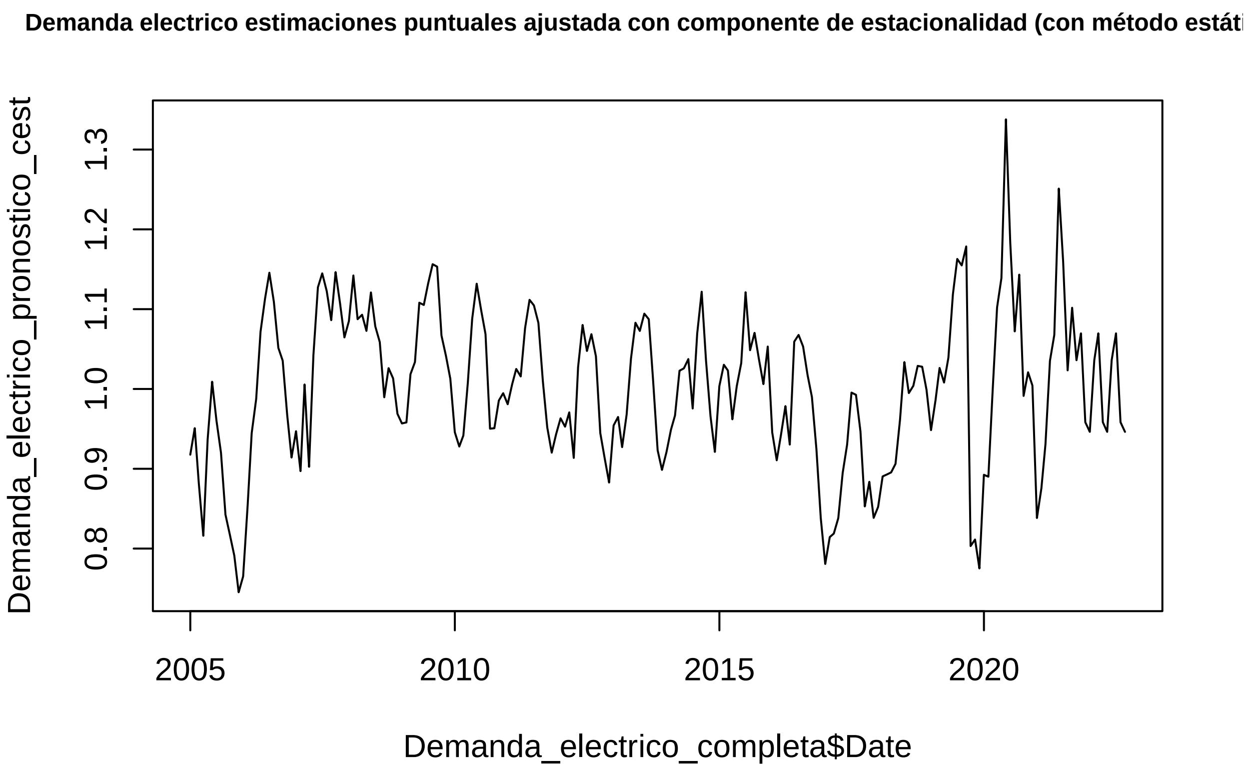
```r
head(LS_demanda_electrico_cest)
```

```
##
## NA NA NA NA NA NA
```

```r
tail(LS_demanda_electrico_cest)
```

```
##
## 1.509216 1.527783 1.711294 1.804612 1.648900 1.658474
```

```r
plot(Demanda_electrico_completa$Date, Demanda_electrico_pronostico_cest, type="l", main="Demanda electri
lines(LI_demanda_electrico_cest, col="violet")
lines(LS_demanda_electrico_cest, col="violet")
```




```r
# Repeat seasonal factors
n <- length(Demanda_electrico_pronostico)
repeated_FE <- rep(c(FE1_electrico, FE2_electrico, FE3_electrico, FE4_electrico), length.out = n)

# Compute components with seasonal factor
Demanda_electrico_pronostico_cest <- Demanda_electrico_pronostico * repeated_FE
LI_demanda_electrico_cest <- LI_demanda_electrico_lineaA * repeated_FE
tail(LI_demanda_electrico_cest)
```

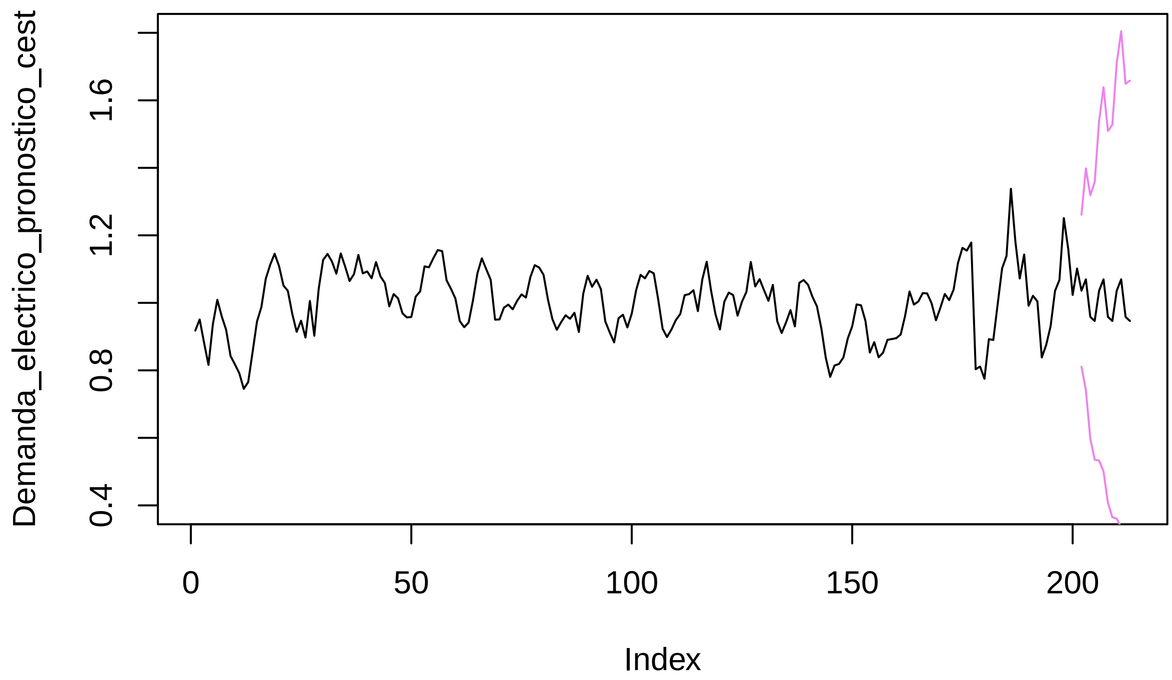
```
##
## 0.4074558 0.3647722 0.3607467 0.3347541 0.2677712 0.2340818
```

```r
LS_demanda_electrico_cest <- LS_demanda_electrico_lineaA * repeated_FE
tail(LS_demanda_electrico_cest)
```

```
##
## 1.509216 1.527783 1.711294 1.804612 1.648900 1.658474
```

```r
# Plot with confidence intervals
plot(Demanda_electrico_pronostico_cest,
     type="l", main="Demanda electrico estimación puntual e intervalos de confianza ajustada con compone
     cex.main=0.65, ylim=c(0.4,1.8))
lines(LI_demanda_electrico_cest, col="violet")
lines(LS_demanda_electrico_cest, col="violet")
```

Demanda electrico estimación puntual e intervalos de confianza ajustada con componente de estacionalidad (con método e



Y ahora "Devolvemos el efecto de la tendencia" a la serie LS_demanda_electrico_cest

Para eso primero necesitamos "crear un data frame que contenga la información histórica de demanded_gas" + 12 meses de pronóstico con un valor NA

```
# create a new data set with 12 additional months
extended_dates_electrico <- seq(as.Date("2005-01-01"), as.Date("2022-09-01"), by = "month")

# create a data frame with 225 rows
extended_demanda_electrico <- data.frame(Date = extended_dates_electrico,
                                         Demanded_Gas = rep(NA, length(extended_dates_electrico)))

# assign values to a subset of Demanded_Gas that corresponds to the length of Demanda_electrico$Demande
extended_demanda_electrico$Demanded_Gas[1:length(Demanda_electrico$Demanded_Gas)] <- Demanda_electrico$I


extended_demanda_electrico$Date <- as.POSIXct(as.character(extended_demanda_electrico$Date), format = "

head(extended_demanda_electrico)
```

```
##         Date Demanded_Gas
## 1 2005-01-01      1819.58
## 2 2005-02-01      1895.33
## 3 2005-03-01      1765.86
## 4 2005-04-01      1642.70
## 5 2005-05-01      1895.54
## 6 2005-06-01      2051.72
```

```r
tail(extended_demanda_electrico)
```

```
##           Date Demanded_Gas
## 208 2022-04-01           NA
## 209 2022-05-01           NA
## 210 2022-06-01           NA
## 211 2022-07-01           NA
## 212 2022-08-01           NA
## 213 2022-09-01           NA
```

Posteriormente calculamos los valores historicos + 12 meses pronosticados a partir de la linea de tendencia que previamente se habia calculado de nombre Electrico_Demanded_gas_line

```r
Electrico_Demanded_gas_line_aplicada2 <- predict(Electrico_Demanded_gas_line, newdata = data.frame(Date

head(Electrico_Demanded_gas_line_aplicada2)
```

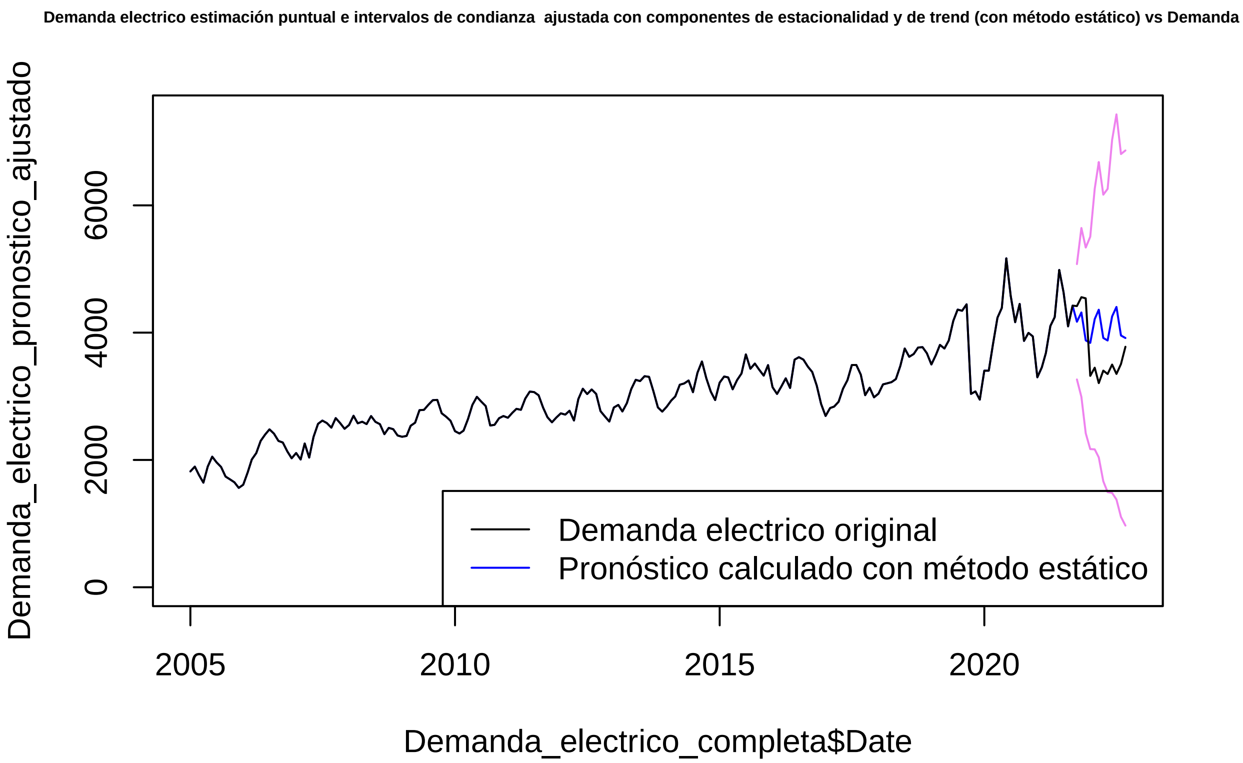
```
##        1        2        3        4        5        6
## 1983.001 1993.354 2002.705 2013.058 2023.063 2033.416
```

```r
tail(Electrico_Demanded_gas_line_aplicada2)
```

```
##      208      209      210      211      212      213
## 4086.664 4096.669 4107.022 4117.041 4127.394 4137.747
```

Entonces para finalmente "devolver el efecto de la tendnecia a los datos de LS_demanda_electrico_cest" se multiplica cada valor de la serie por el valor calculado a partir de la linea de tencia que se llama: Electrico_Demanded_gas_line_aplicada2

```r
Demanda_electrico_pronostico_ajustado<-Demanda_electrico_pronostico_cest*Electrico_Demanded_gas_line_ap

LI_demanda_electrico_ajustado<-LI_demanda_electrico_cest*Electrico_Demanded_gas_line_aplicada2
LS_demanda_electrico_ajustado<-LS_demanda_electrico_cest*Electrico_Demanded_gas_line_aplicada2


plot(Demanda_electrico_completa$Date, Demanda_electrico_pronostico_ajustado, type="l", ylim=c(0, max(LS
     cex.main=0.5, col="blue")
lines(Demanda_electrico_completa$Date,LI_demanda_electrico_ajustado, col="violet")
lines(Demanda_electrico_completa$Date, LS_demanda_electrico_ajustado, col="violet")
lines(Demanda_electrico_completa$Date, Demanda_electrico_completa$Demanded_Gas, col="black")


legend("bottomright", legend = c("Demanda electrico original", "Pronóstico calculado con método estátic
       col = c("black", "blue"), lty = 1)
```



Visualizamos el Data.Frame con los valores históricos la "data original" + los resultados del pronóstico puntual y los intervalos de confianza al 95%

```
# Create a data frame with the predicted values and intervals
df_demanda_electrico_pronostico <- data.frame(Demanded_Gas_original = c(Demanda_electrico_completa$Deman
                                              Demanda_electrico_pronostico_ajustado = Demanda_electrico_
                                              LI_demanda_electrico_ajustado = LI_demanda_electrico_ajust
                                              LS_demanda_electrico_ajustado = LS_demanda_electrico_ajust

# View the resulting data frame
View(df_demanda_electrico_pronostico)

tail(df_demanda_electrico_pronostico)
```

```
##     Demanded_Gas_original Demanda_electrico_pronostico_ajustado
## 208               3403.44                              3916.396
## 209               3350.03                              3876.587
## 210               3498.70                              4254.958
## 211               3350.97                              4403.928
## 212               3506.42                              3955.429
## 213               3778.37                              3915.458
##     LI_demanda_electrico_ajustado LS_demanda_electrico_ajustado
## 208                     1665.1350                      6167.657
## 209                     1494.3511                      6258.822
## 210                     1481.5946                      7028.321
## 211                     1378.1963                      7429.660
```

```
## 212                    1105.1974                      6805.661
## 213                     968.5715                      6862.344
```

Se observa que los pronósticos puntuales vs los reales parecen no estar muy alejados unos de otros, sin embargo los intervalos de confianza si "sufren bastante a medida que nos alejamos del último valor real" pues son cada vez más grandes hasta llegar a abarcar un rango de [1000,6800] para el 95% de intervalo de confianza en el 12vo mes pronósticado. Lo que representaría que: nuestro intervalo de confianza al 95% para nuestros valores más alejados es del orden de casi el rango total de nuestra variable y. En general, con lo observado en este ejemplo y lo visto en clase se puede mencionar que: "Los métodos de suavizamiento estático son relativamente buenos para generar pronóstcos puntuales, pero los intervalos de confianza del pronóstico se ven severamente castigados mientras más datos n 'en el tiempo t+n' quieran ser pronósticados"

A.6 Calculo del error tipo MAPE para los pronósticos hechos con los modelos estáticos de suavizamiento

Calculemos el error tipo MAPE entre los valores pronosticados con el "método de suavizamiento estático" y los valores reales de la serie

```r
##Primero formemos un vector con los valores reales
Demanda_electrico_completa_reales12 <- Demanda_electrico_completa$Demanded_Gas[202:213]
Demanda_electrico_completa_reales12
```

```
##  [1] 4417.51 4557.36 4538.89 3320.75 3449.80 3206.96 3403.44 3350.03 3498.70
## [10] 3350.97 3506.42 3778.37
```

```r
## Despues formemos un vector con únicamente los valores pronósticados
Demanda_electrico_E_forecast12 <- Demanda_electrico_pronostico_ajustado[202:213]
Demanda_electrico_E_forecast12
```

```
##
## 4170.880 4317.491 3877.669 3838.677 4213.452 4360.360 3916.396 3876.587
##
## 4254.958 4403.928 3955.429 3915.458
```

```r
##Calculemos el MAPE entre estos dos vectores

# First, calculate the absolute percentage error (APE) for each element in the vectors
APE_electrico_estatico <- abs((Demanda_electrico_completa_reales12 - Demanda_electrico_E_forecast12) / 

# Calculate the mean APE across all elements in the vectors
MAPE_electrico_estatico <- mean(APE_electrico_estatico) * 100

# Print the MAPE value
cat("El error tipo MAPE  para el forecast calculado con el método de 'suavizamiento estático'de la seri
```

```
## El error tipo MAPE  para el forecast calculado con el método de 'suavizamiento estático'de la serie 
```

B. Método de suavizamiento dinámico ó Holt-winters

Vamos a usar ahora el método de HoltWinters como "modelo de suaviazamiento y pronóstico NO ESTÁTICO" que si considera los cambios de nivel a tráves del tiempo.

Para ello debemos tener nuestra data a suavizar en el tipo de dato serie de tiempo ts en R

```
Demanda_electrico_ts <-ts(Demanda_electrico$Demanded_Gas, frequency =12, start =c(2005,1))
head(Demanda_electrico_ts)
```

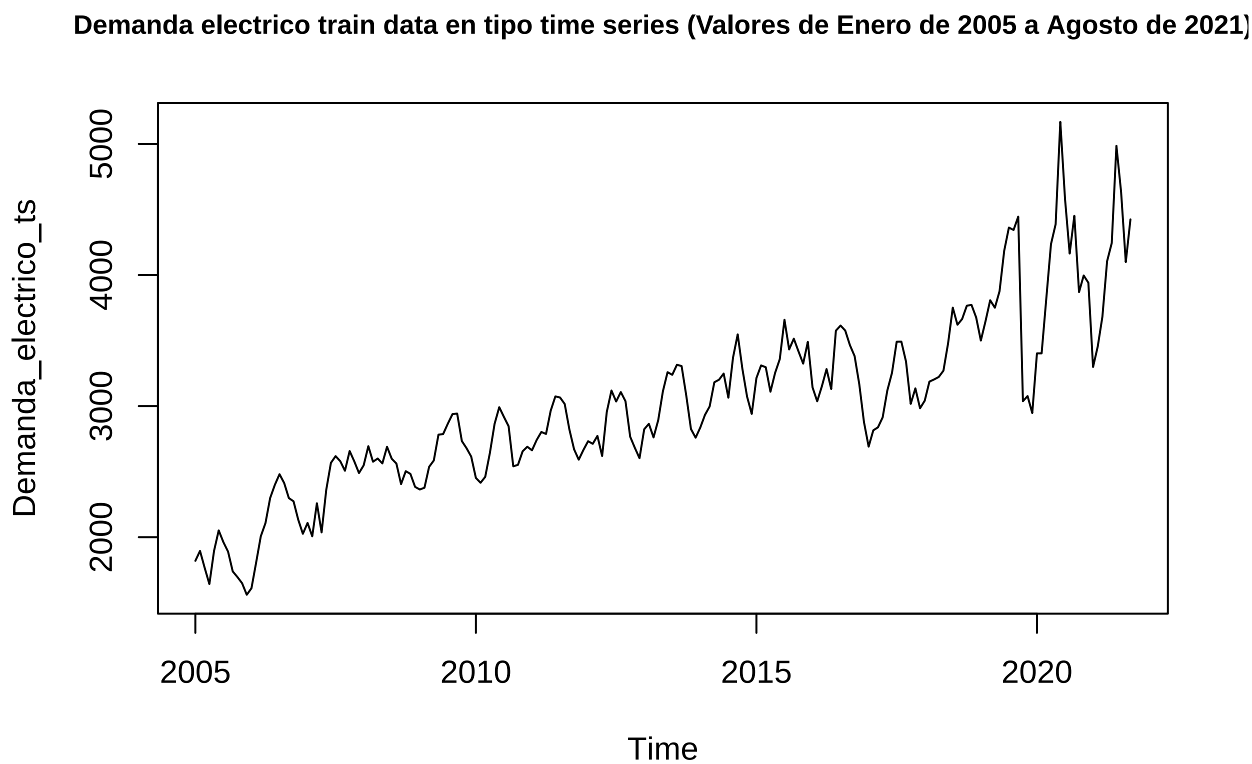
```
##          Jan     Feb     Mar     Apr     May     Jun
## 2005 1819.58 1895.33 1765.86 1642.70 1895.54 2051.72
```

```
tail(Demanda_electrico_ts)
```

```
##          Apr     May     Jun     Jul     Aug     Sep
## 2021 4104.82 4243.93 4985.53 4631.85 4098.81 4424.39
```

Graficamos la "data original de la demanda_electrico" que se usará como "train para el modelo" y que fue convertida a tipo de dato serie de tiempo

```
plot(Demanda_electrico_ts, main="Demanda electrico train data en tipo time series (Valores de Enero de 2
```



Aplicación del método de Holt Winters" a la serie CON TENDENCIA y CON ESTACIONALIDAD de demanda de gas en el sector eléctrico

Como se observará para este primer cálculo se dejará como 'grados de libertad' los valores de alpha, gamma y betha, así como los valores de arranque, de manera que el "el algoritmo de HltWinters de R" determine la "mejor combinaión posible" Recordar también que para la presente serie se está asumiendo un efecto multiplicativo de los factores estacionales

```r
Demanda_electrico_ts_hw1 <-HoltWinters(Demanda_electrico_ts, seasonal = 'multiplicative')
Demanda_electrico_ts_hw1
```

```
## Holt-Winters exponential smoothing with trend and multiplicative seasonal component.
##
## Call:
## HoltWinters(x = Demanda_electrico_ts, seasonal = "multiplicative")
##
## Smoothing parameters:
##  alpha: 0.8778578
##  beta : 0.004262188
##  gamma: 1
##
## Coefficients:
##             [,1]
## a   4083.8396737
## b     22.0244141
## s1     0.9346142
## s2     0.9328241
## s3     0.8821197
## s4     0.8525012
## s5     0.8813576
## s6     0.9252482
## s7     0.9699071
## s8     1.0226685
## s9     1.1390634
## s10    1.1151142
## s11    1.0804254
## s12    1.0833897
```

Interpretación de los valores de los coeficientes del método de Holt Winters para el caso de la serie de demanda de gas natural en el sector eléctrico en México Interpretación de los valores de coficientes α, β y γ para el caso del pronóstico usando modelo de HoltWinters SIN ESPECIFICAR DATOS DE ARRANQUE

$$0 < \alpha < 1$$

Recordemos que: Valores de α cercano a 1 reflejan caminata aleatoria Valores de α cercano a 0 reflejan una "gran influencia" de los valores pasado, es decir, que "el ruido no es un componente de gran relevancia"

En nuestro caso el valor de α es:

```r
Demanda_electrico_ts_hw1$alpha
```

```
##     alpha
## 0.8778578
```

Lo que indica que "La data historica NO tiene mayor relevancia vs el componente estocástico"

$$0 < \beta < 1$$

Recordemos que: Valores de β cercano a 1 reflejan cambios CONSTANTES de nivel a lo largo de la serie Valores de β cercano a 0 reflejan POCOS cambios de nivel a lo largo de la serie

En nuestro caso el valor de β es:

```
Demanda_electrico_ts_hw1$beta
```

```
##        beta
## 0.004262188
```

Lo que indica muy pocos cambios de nivel a lo largo de la serie o una especie de “pendiente constante”

Y entonces para el caso de valores de α cercanos a 1 y de β cercanos a cero (como es nuestra serie de demanda_electrico) indica una serie con TENDENCIA ESTABLE y con CAMBIOS DE NIVEL MUY BAJOS

Afortunadamente no obtuvimos simultaneamente valores deα cercanos a 1 y de β cercanos a 1 pues eso reflejaría una CAMINATA ALEATORIA con tendencia estocástica que sería prácticamente imposible de pronósticar.

$$0 < \gamma < 1$$

Recordemos que: Valores de γ cercano a 1 reflejan cambios CONSTANTES de los factores de estacionalidad a lo largo de la serie Valores de γ cercano a 0 reflejan POCOS cambios en los factores de estacionalidad a lo largo de la serie

En nuestro caso el valor de γ es:

```
Demanda_electrico_ts_hw1$gamma
```

```
## gamma
##     1
```

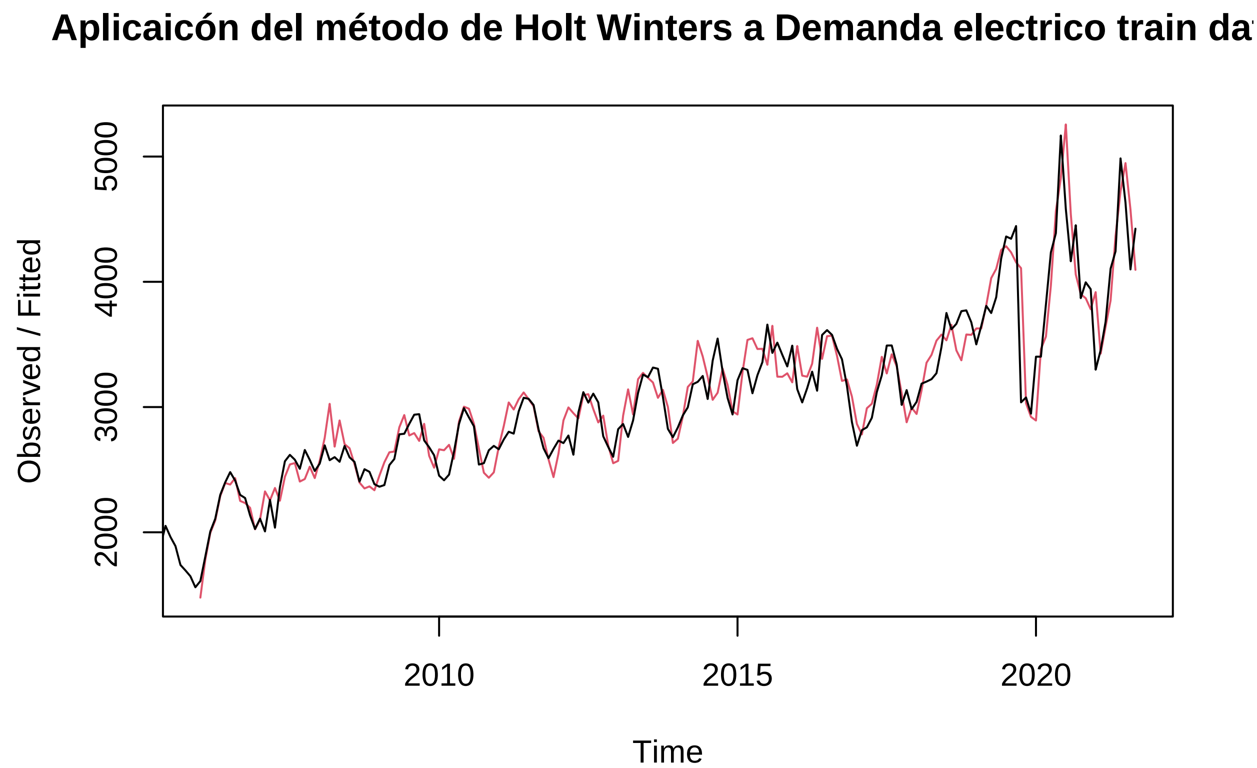
Lo que indica que cada dato nuevo en la serie cambia el valor del factor de estacionalidad

Pronóstico de la serie de demanda de gas en el sector eléctrico con el método de Holt Winters

Ahora pasemos a realizar los pronósticos de la serie pero con el método de HoltWinters que considera cambios dinámicos de nivel

Graficamos los valores originales vs suavizados con el método de Holt-Winters

```
plot(Demanda_electrico_ts_hw1, main="Aplicaicón del método de Holt Winters a Demanda electrico train da
```



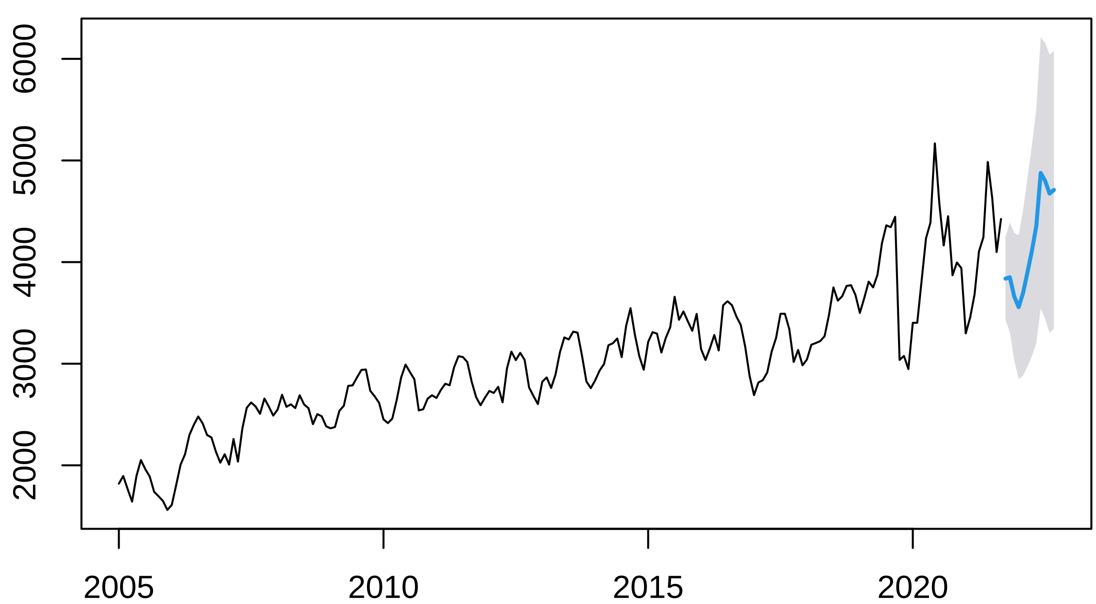

```r
# Load the forecast package
library(forecast)
```

Ahora se llevará a cabo el pronóstico con la librería de forecast y basados en el método de Holt Winters. De igual manera se usará un intervalo de confianza del 95%

```r
Demanda_electrico_ts_hw1_fc <- forecast(Demanda_electrico_ts_hw1, h=12, level=0.95)
Demanda_electrico_ts_hw1_fc
```

```
##          Point Forecast    Lo 95    Hi 95
## Oct 2021       3837.399 3431.808 4242.989
## Nov 2021       3850.594 3310.345 4390.843
## Dec 2021       3660.720 3033.263 4288.176
## Jan 2022       3556.581 2849.617 4263.546
## Feb 2022       3696.380 2882.393 4510.367
## Mar 2022       3900.834 2974.030 4827.637
## Apr 2022       4110.477 3073.869 5147.084
## May 2022       4356.604 3204.684 5508.524
## Jun 2022       4877.537 3544.382 6210.692
## Jul 2022       4799.545 3440.987 6158.102
## Aug 2022       4674.037 3304.396 6043.679
## Sep 2022       4710.722 3345.397 6076.047
```

```r
plot(Demanda_electrico_ts_hw1_fc, main="Forecast de Demanda_electrico usando el Método de Holt Winters")
```



Veamos el valor la Suma de Errores al Cuadrado del pronóstico de la demanda de gas natural en el sector eléctrico usando el Método de suaviazamiento dinámico de HoltWinters

```r
Demanda_electrico_ts_hw1$SSE
```

```
## [1] 8118743
```

Calculemos el error tipo MAPE entre los valores pronosticados con el "método de Holt Winters" y los valores reales de la serie

```r
## Formemos un vector con únicamente los valores pronósticados
Demanda_electrico_HW_forecast12 <- Demanda_electrico_ts_hw1_fc$mean
Demanda_electrico_HW_forecast12
```

```
##           Jan      Feb      Mar      Apr      May      Jun      Jul      Aug
## 2021
## 2022 3556.581 3696.380 3900.834 4110.477 4356.604 4877.537 4799.545 4674.037
##           Sep      Oct      Nov      Dec
## 2021          3837.399 3850.594 3660.720
## 2022 4710.722
```

```r
##Calculemos el MAPE entre estos dos vectores

# First, calculate the absolute percentage error (APE) for each element in the vectors
APE_electrico_HW <- abs((Demanda_electrico_completa_reales12 - Demanda_electrico_HW_forecast12) / Demand

# Calculate the mean APE across all elements in the vectors
MAPE_electrico_HW <- mean(APE_electrico_HW) * 100

# Print the MAPE value
cat("MAPE error para el forecast calculado con el método de 'HoltWinters'es:", MAPE_electrico_HW, "%")
```

```
## MAPE error para el forecast calculado con el método de 'HoltWinters'es: 22.9424 %
```

Error cuadrado calculado con el Método de HoltWinters vs el calculado con el método de suafizamiento estático

```r
Demanda_electrico_ts_hw1$SSE
```

```
## [1] 8118743
```

Entonces el error asociado al pronóstico del cálculo de pronóstico PUNTUAL para el caso Demanda_eléctirco es:

```r
sse_Demanda_electrico_hw <- ((Demanda_electrico_ts_hw1$SSE)/(length(Demanda_electrico_ts)))^0.5
sse_Demanda_electrico_hw
```

```
## [1] 200.977
```

Recordar que la SSE de la demanda_electrico para el caso del método de suavizamiento estático era:

```r
SSE_demanda_electrico_E <- sum(demanda_electrico_errores^2)
SSE_demanda_electrico_E
```

```
## [1] 2.375525
```

Lo que indica que el método de suavizamiento estático resultó bastante mejor en nuestro caso para el PRONOSTICO PUNTUAL de la demanda de gas natural en el sector eléctrico en México vs el método de HoltWinters

Comparación del MAPE del pronostico hecho con métodos estáticos vs el pronóstico hecho con el método " de suavizamiento dinámico" de HoltWinters para el pronostico puntual

En este caso y dado un cálculo con el métodO de HoltWinters sin valores de arranque especificados los errores de ambos métodos (estático vs HoltWinters) fueron los sigientes:

```r
cat("MAPE error para el forecast calculado con el método de 'suavizamiento estático'es:", MAPE_electric
```

```
## MAPE error para el forecast calculado con el método de 'suavizamiento estático'es: 16.61447 %
```

```r
cat("MAPE error para el forecast calculado con el método de 'HoltWinters'es:", MAPE_electrico_HW, "%")
```

```
## MAPE error para el forecast calculado con el método de 'HoltWinters'es: 22.9424 %
```

Lo que de nuevo indica que el método de suavizamiento estático resultó mejor en nuestro caso para el PRONOSTICO PUNTUAL de la demanda de gas natural en el sector eléctrico en México.

Comparación de los intervalos del pronostico hecho con métodos estáticos vs el pronóstico hecho con el método " de suavizamiento dinámico" de HoltWinters

Como se sabe, la parte "más sustancial" o esencial de los pronósticos es el intervalo y no el pronóstico puntual, por lo que en nuestro caso se considerará en general como "un mejor mpetodo de pronóstico" no aquel que otorge el menor error tipo MAPE en las estimaciones puntuales, sino aquel cuyos intervalos de confianza al 95% sean menores incluso para valores pronósticados relativamente lejanos del último valor real.

Intervalos calculados con el método de suavizamiento estático

```r
LI_demanda_electrico_ajustado[202:213]
```

```
##
## 3264.6930 2990.9013 2418.4465 2170.6548 2166.4737 2039.8246 1665.1350 1494.3511
##
## 1481.5946 1378.1963 1105.1974  968.5715
```

```r
LS_demanda_electrico_ajustado[202:213]
```

```
##
## 5077.067 5644.081 5336.892 5506.699 6260.431 6680.895 6167.657 6258.822
##
## 7028.321 7429.660 6805.661 6862.344
```

```r
intervalos_metodo_E<- data.frame(LI_demanda_electrico_ajustado =LI_demanda_electrico_ajustado[202:213],

intervalos_metodo_E
```

```
##    LI_demanda_electrico_ajustado LS_demanda_electrico_ajustado
## 1                      3264.6930                      5077.067
## 2                      2990.9013                      5644.081
## 3                      2418.4465                      5336.892
## 4                      2170.6548                      5506.699
## 5                      2166.4737                      6260.431
## 6                      2039.8246                      6680.895
## 7                      1665.1350                      6167.657
## 8                      1494.3511                      6258.822
## 9                      1481.5946                      7028.321
## 10                     1378.1963                      7429.660
## 11                     1105.1974                      6805.661
## 12                      968.5715                      6862.344
```

Intervalos calculados con el método de suavizamiento dinámico Holt-Witers

```
Demanda_electrico_ts_hw1_fc
```

```
##          Point Forecast    Lo 95    Hi 95
## Oct 2021       3837.399 3431.808 4242.989
## Nov 2021       3850.594 3310.345 4390.843
## Dec 2021       3660.720 3033.263 4288.176
## Jan 2022       3556.581 2849.617 4263.546
## Feb 2022       3696.380 2882.393 4510.367
## Mar 2022       3900.834 2974.030 4827.637
## Apr 2022       4110.477 3073.869 5147.084
## May 2022       4356.604 3204.684 5508.524
## Jun 2022       4877.537 3544.382 6210.692
## Jul 2022       4799.545 3440.987 6158.102
## Aug 2022       4674.037 3304.396 6043.679
## Sep 2022       4710.722 3345.397 6076.047
```

Se observa de nuevo claramente como el método de HoltWinters otorga mejores intervalos de confianza incluso para aquellos valores'más alejados' del último valor real vs aquellos generados con el "método de suavizamiento estático"

C. Conclusiones generales sobre las técnicas de suavizamiento estáticas y dinámicas (HoltWinters)

Comparación entre ambas técnicas de suavizamiento y pronóstico de series de tiempo

En general se podrían mencionar las siguientes carácterísticas sobre las "técnicas de suavizamiento estáticas":

- Al no considerar los "cambios históricos en el nivel" pueden ser relativamente buenos para generar estimaciones puntuales para valores en tiempo t+n con n pequeña, apenas más grande que 1, pero si el valor de n va creciendo, es decir, si nos alejamos del ultimo valor conocido, el pronóstico suele ir tendiendo cada vez más error.
- Suele ser "un método extremadamente sensible a valores atípicos", pues de nuevo, al tener en cuenta una especie de "nivel constante" es incapaz de caputrar datos muy lejanos a esa constante.
- Tiende producir valores de límites de pronóstico muy grandes y crecientes "de manera casi exponencial" a medida que nos alejamos del ultimo valor conocido en la serie
- Se sugiere intentar estimar solo 1 a 5 datos del tiempo t + n=5 después del último valor conocido, ya que en los siguientes datos el error de pronóstico y los intervalos son muy grandes
- Es una "ténica estática de suavizamiento" que pricipalmente se basa en la "forma histórica" de los datos e intenta reproducir esa forma en los pronósticos

En general se podrían mencionar las siguientes carácterísticas sobre las "técnicas de suavizamiento dinámicas o de HoltWinters":

- Al SI considerar los "cambios históricos en el nivel" suelen ser "menos adecuados" para calcular estimaciones puntuales en comparación con los "métodos estáticos"
- Es capaz de modelar estacionalidad y cambios en "la forma histórica de los datos" de mejor manera respecto a los métodos estáticos
- Tiende producir valores de límites de pronóstico mucho mas "constantes" incluso para valores de tiempo muy alejados del ultimo valor conocido

- Se sugiere intentar estimar de 1 a 12 datos del tiempo t + n=12 después del último valor conocido, ya que incluso para valores relativamente alejados al ultimo valor conocido produce resultados de estimación puntual e intervalos de confianza buenos.

- Es una "ténica recursiva de suavizamiento" que pricipalmente se basa en la "importancia histórica" de los datos e intenta reproducir esa importancia para estimar los siguientes valores de la serie

Ambas técnicas presentan dificicultades para modelar datos cuya carianza en el tiempo cambie significativamente y también para manekar datos atípicos "outliers"

Referencias

Bowerman, B. L., O'Connell, R. T., & Koehler, A. B. (2005). Forecasting, time series, and regression: An applied approach. Thomson/Brooks/Cole.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time series analysis: forecasting and control (5th ed.). John Wiley & Sons.