

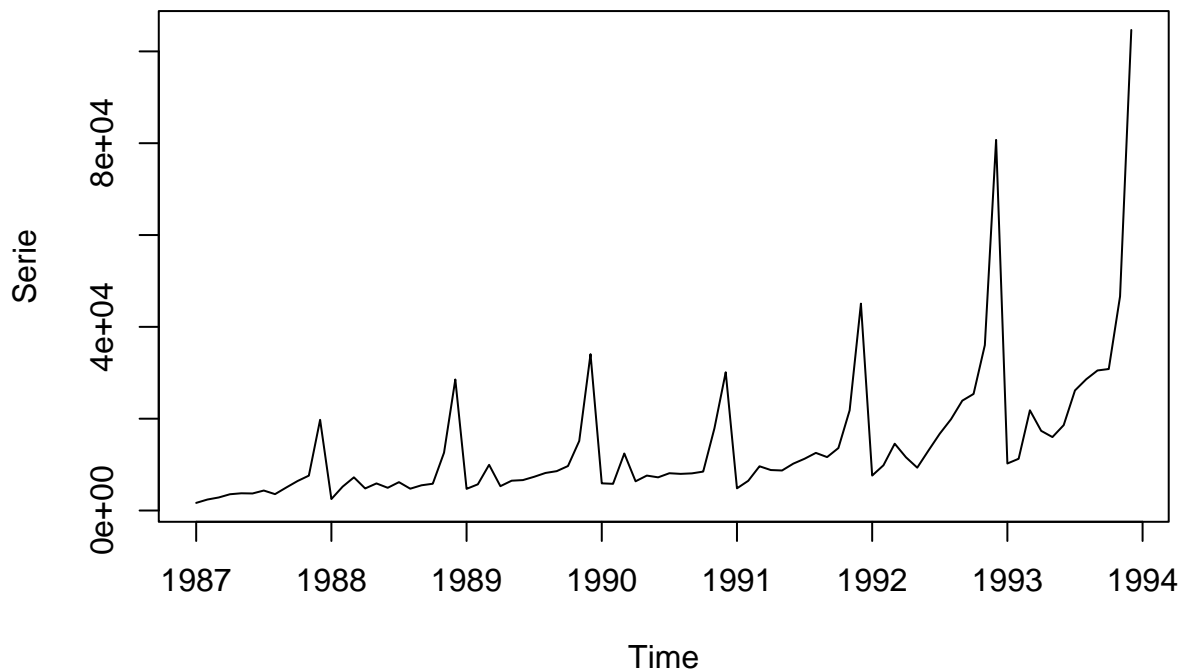
Tarea 1

Eduardo

10/8/2021

1.- Grafique los datos. Describa lo que observe de su información (varianza contante o no constante, tendencia, ciclos estacionales, periodicidad de los ciclos).

Gráfica de la serie de tiempo



Varianza

Podemos observar una varianza creciente conforme pasa el tiempo, teniendo primero un ligero aumento de 1987 a 1990 de manera lineal, después parece que se mantiene constante de 1990 a 1991 para posteriormente crecer demasiado de 1991 a 1994. Es decir, no es constante la varianza. Lo que hace que se dispare es la temporada alta vacacional (Al parecer, en noviembre y diciembre, ya que analizamos una playa en Australia, donde el verano comienza en diciembre)

Tendencia

La tendencia parece ser en general creciente, repitiendo casi el mismo patrón que la varianza: Es creciente de manera ligera (lineal, con pendiente pequeña) de 1987 a 1990, para después decrecer un poco de 1990 a 1991, sin embargo crece de manera cuadrática, al parecer, de 1991 a 1994.

Ciclos estacionales

Dado que estamos analizando una base de datos de ventas mensuales de una tienda de souvenirs en una playa en Australia, hace todo el sentido del mundo que tenga un ciclo ya que tanto en las ventas como visitas a sitios vacacionales hay una fuerte dependencia en los meses del año. Esto lo confirmamos con la gráfica, donde se observa un ciclo estacional bastante claro.

Periodicidad de los ciclos

Complementando el comentario del punto anterior, en la gráfica observamos que el ciclo es anual. De enero a febrero (o los primeros meses del año) parece crecer ligeramente, después baja un poco para crecer de manera ligera nuevamente, pero al llegar lo que parece ser noviembre y diciembre (o los últimos meses del año) crece exageradamente; posteriormente baja de diciembre a enero y se repite el ciclo.

2.- Si la base presenta datos faltantes NA. Use algún método de imputación de la paquetería `imputeTS`.

No hay ningún NA, como podemos observar

```
## [1] 0
```

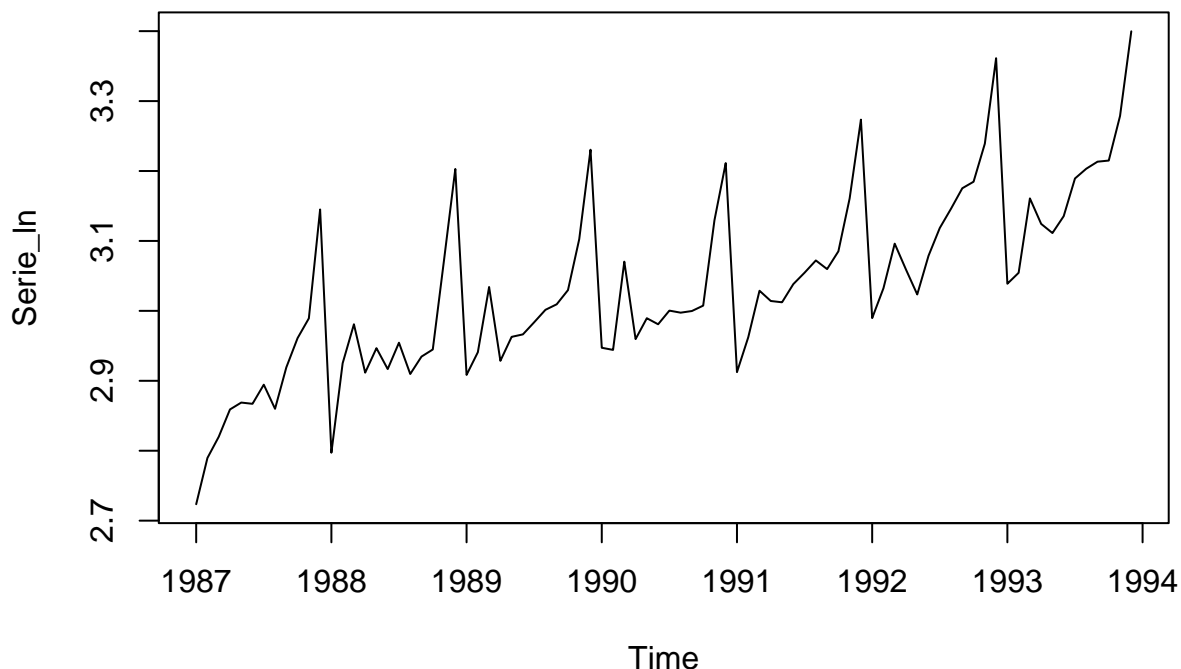
Por lo que no es necesario aplicar ningún método de imputación

3.- Use distintos métodos de descomposición de las series para obtener sus componentes (tendencia y ciclos estacionales), en específico use los siguientes:

(a) Ajuste de curvas (modelos deterministas o de regresión). Realice un pronóstico de 3 años futuros.

Primero realizaremos una transformación con el logaritmo para estabilizar la varianza

```
Serie_ln<-sqrt(log(Serie))  
ts.plot(Serie_ln)
```



Al parecer, la varianza ya se estabilizó considerablemente.

Como la tendencia parece seguir un comportamiento de un polinomio de grado 2, intentaremos ajustarle una curva con dicho polinomio

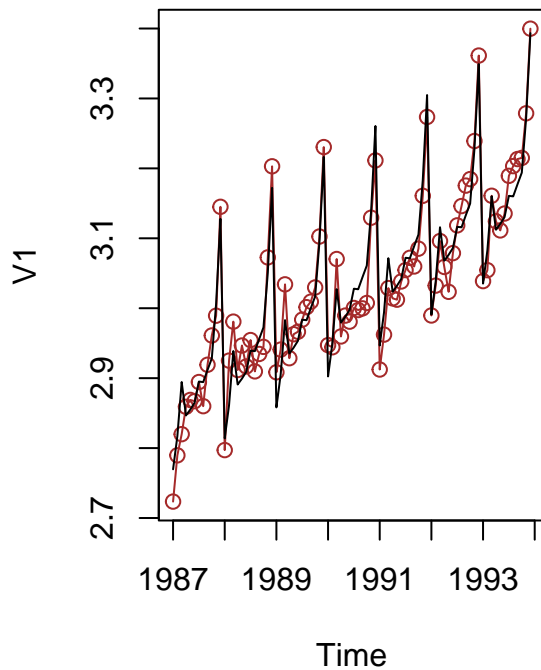
```

M <- factor(cycle(Serie_ln))
t = time(Serie_ln)-1987
regresion_1= lm(Serie_ln ~ 0+ t+M, na.action=NULL)
summary(regresion_1)

##
## Call:
## lm(formula = Serie_ln ~ 0 + t + M, na.action = NULL)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.074594 -0.019946  0.001211  0.017548  0.063774
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## t    0.044265   0.001669   26.52  <2e-16 ***
## M1  2.769652   0.012602  219.78  <2e-16 ***
## M2  2.813350   0.012658  222.26  <2e-16 ***
## M3  2.887112   0.012715  227.06  <2e-16 ***
## M4  2.835616   0.012774  221.99  <2e-16 ***
## M5  2.840288   0.012833  221.32  <2e-16 ***
## M6  2.846410   0.012894  220.75  <2e-16 ***
## M7  2.872990   0.012956  221.74  <2e-16 ***
## M8  2.868526   0.013020  220.32  <2e-16 ***
## M9  2.882184   0.013084  220.28  <2e-16 ***
## M10 2.895120   0.013150  220.16  <2e-16 ***
## M11 2.969272   0.013217  224.66  <2e-16 ***
## M12 3.087268   0.013284  232.40  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0306 on 71 degrees of freedom
## Multiple R-squared:  0.9999, Adjusted R-squared:  0.9999
## F-statistic: 6.363e+04 on 13 and 71 DF,  p-value: < 2.2e-16

par(mfrow=c(1,2))
plot(Serie_ln, type="o",col='brown')
lines(fitted(regresion_1), col='black')

```



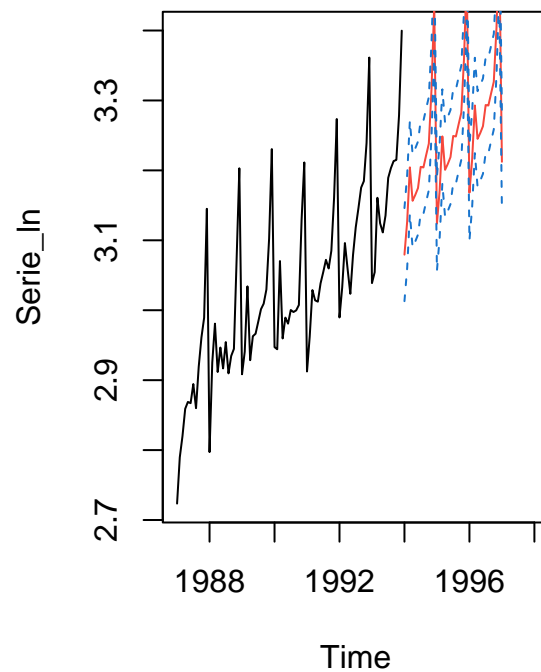
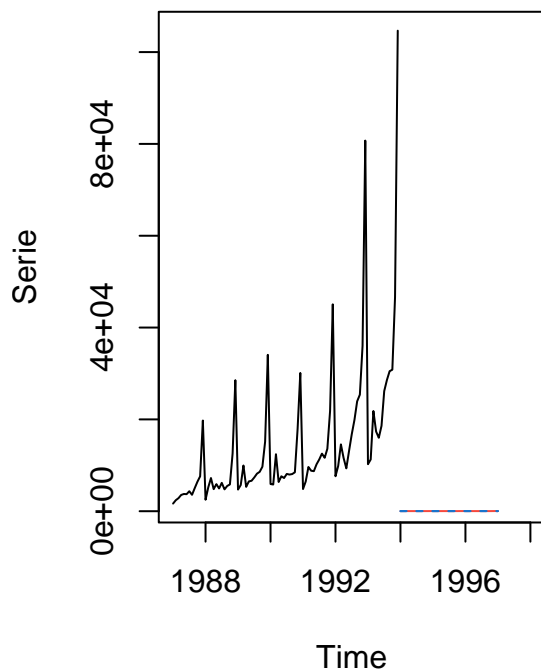
Observamos que todos los valores, tienen un p-value menor a 0.05. Entonces, quitando el efecto de ciclos estacionales, sigue una tendencia lineal.

$$\mathbb{E}[X_t] = 0.26872t + M$$

Donde M es la parte del ciclo estacional.

Ajustamos:

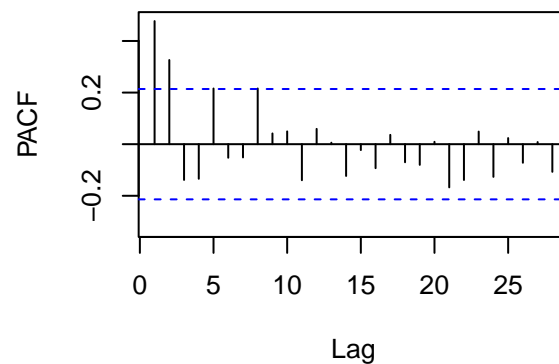
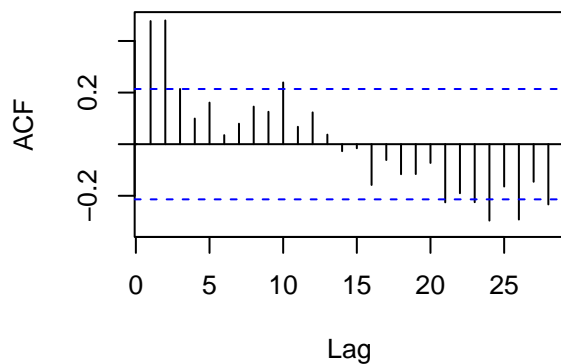
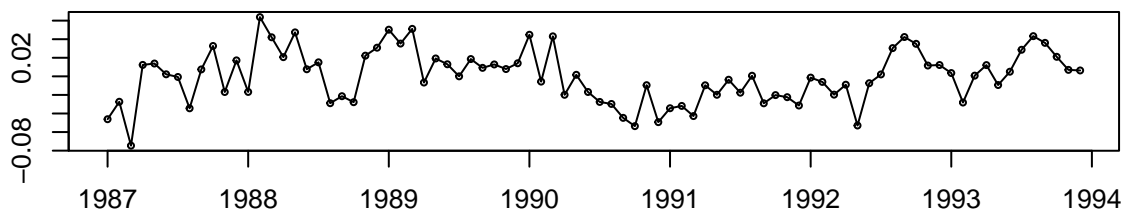
```
tnew = 7 + seq(0,3,length.out=37)
Mnew = factor(c((1:12),(1:12),(1:12),1))
pred1<- predict(regresion_1, newdata=list(t=tnew, M=Mnew), interval="prediction")
par(mfrow=c(1,2))
ts.plot(Serie, xlim=c(1987,1998))
lines(1987+tnew,exp(pred1[,1]), lty=1,col=2)
lines(1987+tnew,exp(pred1[,2]), lty=2,col=4)
lines(1987+tnew,exp(pred1[,3]), lty=2,col=4)
ts.plot(Serie_ln, xlim=c(1987,1998))
lines(1987+tnew,(pred1[,1]), lty=1,col=2)
lines(1987+tnew,(pred1[,2]), lty=2,col=4)
lines(1987+tnew,(pred1[,3]), lty=2,col=4)
```



Ahora comprobemos los supuestos de regresión

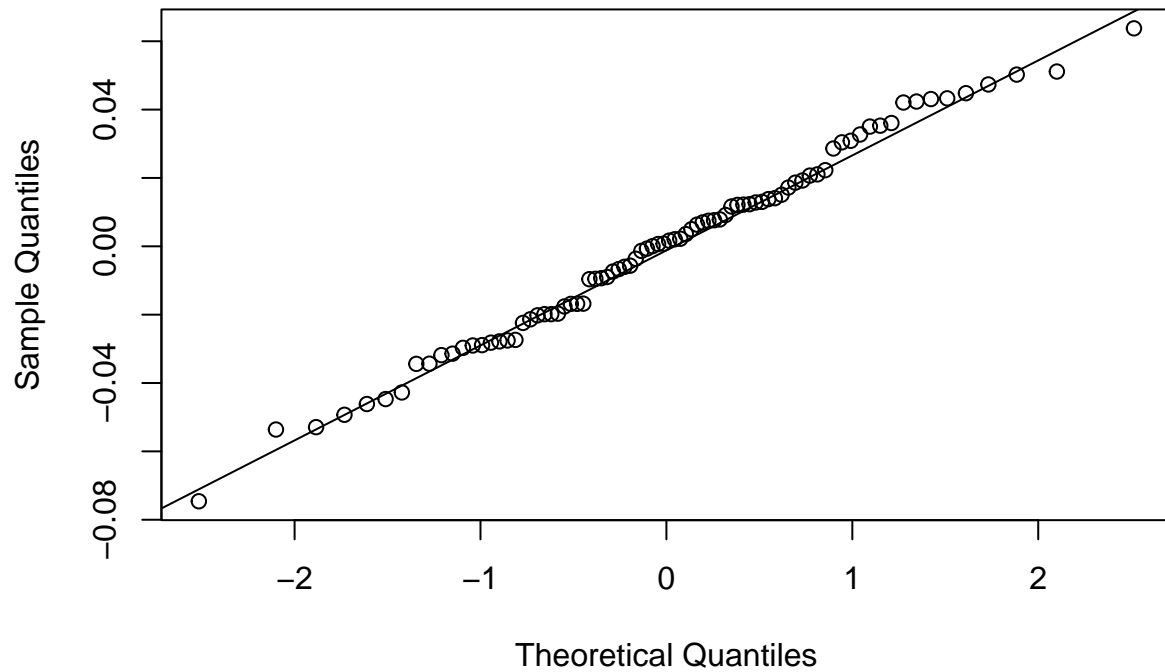
```
tsdisplay(regresion_1$res)
```

regresion_1\$res



```
qqnorm(regresion_1$res)
qqline(regresion_1$res)
```

Normal Q-Q Plot



```
ad.test(regresion_1$res)
```

```
##  
##  Anderson-Darling normality test  
##  
## data:  regresion_1$res  
## A = 0.23236, p-value = 0.7937
```

```
bptest(regresion_1)
```

```
##  
##  studentized Breusch-Pagan test  
##  
## data:  regresion_1  
## BP = 24.664, df = 12, p-value = 0.0165
```

(b) Filtros lineales o suavizamientos exponenciales. Realice un pronóstico de 3 años futuros.

(c) Diferencias.