

Bookdown - Grupo B

Andrés Peralta Alean, David Barrera Barrera, Luis Gonzalo Guerra J

2023-05-10

Contents

1	Propuesta	5
2	Analisis Exploratorio	7
2.1	Creacion del objeto de analisis temporal indice.ts	7
2.2	Analisis Descriptivo	10
3	Estacionalidad y Descomposicion	13
3.1	Estacionalidad	13
3.2	Ahora, hagamos una descomposicion del objeto y la graficamos para su analisis	14
4	Methods	21
4.1	math example	21
5	Applications	23
5.1	Example one	23
5.2	Example two	23
6	Final Words	25

Chapter 1

Propuesta

Análisis de riesgo crediticio y su importancia La evaluación del riesgo crediticio es una tarea crítica para cualquier empresa que ofrezca préstamos o créditos. El incumplimiento de los términos del préstamo o la falta de pago pueden generar grandes pérdidas financieras y afectar la estabilidad de la empresa. Es por ello que es importante contar con herramientas y técnicas que permitan evaluar el riesgo crediticio de manera eficiente.

El análisis de riesgo crediticio utiliza históricos para evaluar el comportamiento de los clientes a lo largo de varios periodos. De esta forma, se pueden obtener patrones y características que permiten segregar grupos de clientes y determinar un posible perfil de riesgo. Los resultados obtenidos a partir de este análisis pueden ayudar a la empresa a tomar decisiones más informadas y a reducir el riesgo crediticio.

Además, el análisis de riesgo crediticio permite identificar posibles oportunidades para la empresa. Por ejemplo, puede ayudar a la empresa a identificar clientes que presenten un bajo riesgo crediticio y, por lo tanto, puedan recibir préstamos con tasas de interés más bajas. De esta manera, la empresa puede aumentar su base de clientes y mejorar su rentabilidad.

En cuanto a las fuentes de información, se pueden utilizar diversas fuentes para recopilar los datos necesarios para el análisis de riesgo crediticio. Entre ellas se encuentran bases de datos públicas y privadas, encuestas, registros gubernamentales, entre otras. Es importante verificar la calidad de la información obtenida para asegurar la precisión y confiabilidad de los resultados.

En el caso específico de la empresa ABC, se cuenta con los permisos necesarios por parte del área de gestión de cobranzas y se eliminaron los datos personales de los clientes con los cuales se extrajeron las características.

En resumen, el análisis de riesgo crediticio es una técnica muy útil para evaluar el riesgo crediticio de los clientes y para identificar posibles oportunidades para

la empresa. Se pueden utilizar diversas fuentes de información para recopilar los datos necesarios para este análisis, pero es importante asegurarse de contar con los permisos necesarios y verificar la calidad de la información obtenida.

Chapter 2

Analisis Exploratorio

2.1 Creacion del objeto de analisis temporal indice.ts

2.1.1 Carga de librerias y datasource

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

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##   as.Date, as.Date.numeric

## Successfully loaded changepoint package version 2.2.4
## See NEWS for details of changes.

## # A tibble: 643,570 x 8
##   Periodo Sub_Tipo N_Clientes DIAS_DE_MORA      Saldo Genero grupo_actividad_eco
##   <chr>    <chr>          <dbl>      <dbl>      <dbl> <chr>  <chr>
## 1 2018-01 CDC              4          0 15824105. Femen~ Dependiente privado
## 2 2018-01 CDC              1          0  6810373. Femen~ Dependiente privado
## 3 2018-01 CDC              6          6 28819502. Femen~ Dependiente privado
```

```
## 4 2018-01 CDC 12 63 81343674. Femen~ Dependiente privado
## 5 2018-01 CDC 1 21 7524344. Femen~ Dependiente privado
## 6 2018-01 CDC 4 0 12974213. Femen~ Dependiente privado
## 7 2018-01 CDC 3 1 21348609. Femen~ Dependiente privado
## 8 2018-01 CDC 2 0 11475858. Femen~ Dependiente privado
## 9 2018-01 CDC 10 38 60012355. Femen~ Dependiente privado
## 10 2018-01 CDC 1 0 9034715. Femen~ Dependiente privado
## # i 643,560 more rows
## # i 1 more variable: Ciudad_res <chr>
```

2.1.2 Modificamos el df para que tenga el formato adecuado y lo mostramos

```
# Cambio el tipo de dato de la columna temporal(Periodo)
datos$Periodo <- as.Date(paste0(datos$Periodo, "-01"))

# Consolido el df en funcion de la variable de interes (Saldo)
datos <- aggregate(Saldo ~ Periodo, data = datos, sum)

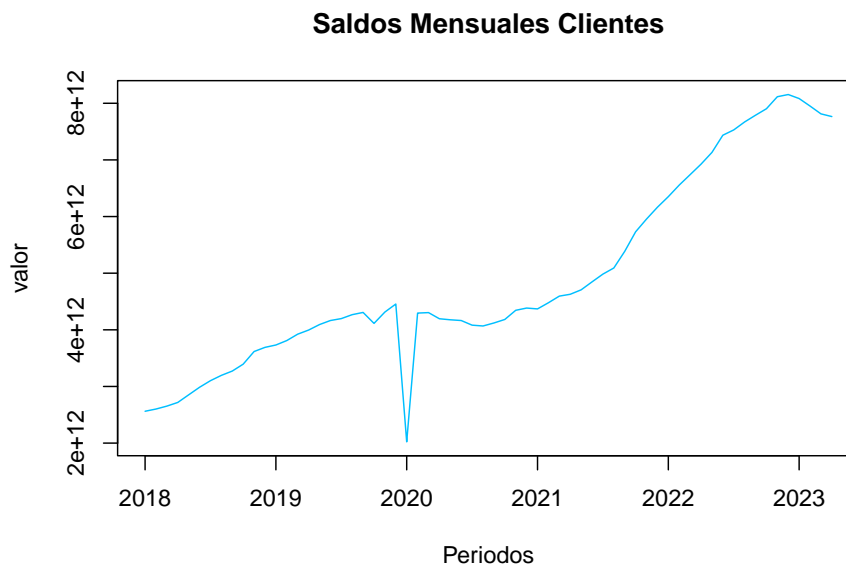
# Genero mi objeto ts para el analisis
indice.ts <- ts(datos$Saldo, start = c(2018,1), frequency = 12)
indice.ts
```

```
##           Jan           Feb           Mar           Apr           May
## 2018 2.562223e+12 2.601532e+12 2.653315e+12 2.717915e+12 2.852608e+12
## 2019 3.732137e+12 3.810740e+12 3.923380e+12 3.995810e+12 4.092819e+12
## 2020 2.022405e+12 4.295886e+12 4.304185e+12 4.195404e+12 4.177219e+12
## 2021 4.369300e+12 4.478002e+12 4.595047e+12 4.627874e+12 4.705783e+12
## 2022 6.351501e+12 6.555819e+12 6.739823e+12 6.925001e+12 7.132542e+12
## 2023 8.083445e+12 7.951094e+12 7.812947e+12 7.765132e+12
##           Jun           Jul           Aug           Sep           Oct
## 2018 2.986446e+12 3.102493e+12 3.196138e+12 3.272388e+12 3.394038e+12
## 2019 4.164386e+12 4.198177e+12 4.267921e+12 4.307340e+12 4.113506e+12
## 2020 4.164434e+12 4.082507e+12 4.067948e+12 4.120926e+12 4.182877e+12
## 2021 4.846473e+12 4.983882e+12 5.091594e+12 5.385784e+12 5.730839e+12
## 2022 7.435792e+12 7.529328e+12 7.670212e+12 7.789046e+12 7.903472e+12
## 2023
##           Nov           Dec
## 2018 3.617129e+12 3.690229e+12
## 2019 4.314034e+12 4.455499e+12
## 2020 4.344709e+12 4.384188e+12
## 2021 5.955740e+12 6.164705e+12
## 2022 8.115444e+12 8.154174e+12
## 2023
```


2.1. CREACION DEL OBJETO DE ANALISIS TEMPORAL INDICE.TS 9

2.1.3 Graficamos la serie

```
plot(indice.ts, main = "", ylab="valor", col="deepskyblue", xlab="Periodos")
title(main = "Saldos Mensuales Clientes")
```



2.1.4 Chequeos basicos para confirmar la estructura del contenedor ts

```
## [1] "El tipo de datos del df indice.ts es: "
```

```
## [1] "ts"
```

```
## [1] "La serie de tiempo indice.ts empieza en: "
```

```
## [1] 2018    1
```

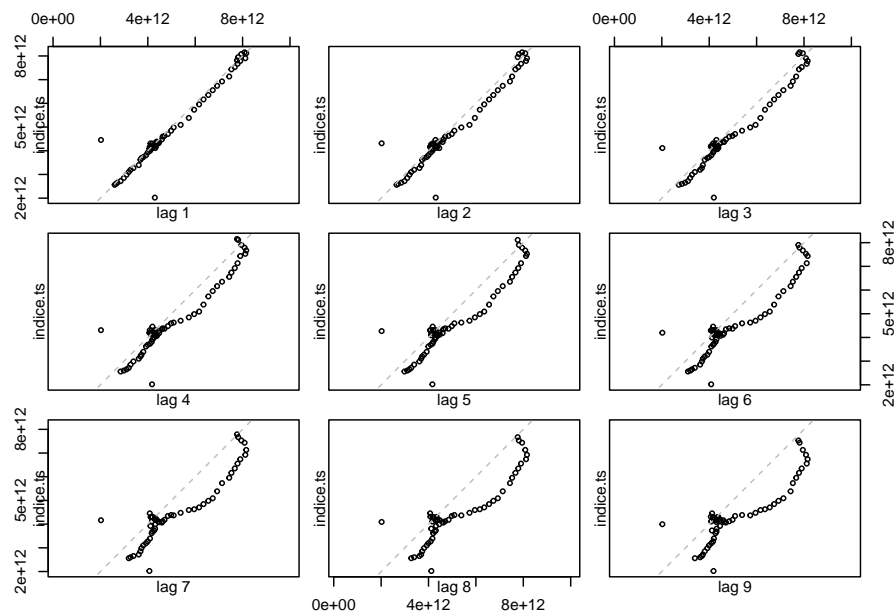
```
## [1] "La serie de tiempo indice.ts termina en: "
```

```
## [1] 2023    4
```

2.2 Analisis Descriptivo

2.2.1 Grafica de Rezagos

```
lag.plot(indice.ts, 9, do.lines = FALSE)
```



Conclusion: Se observa con claridad que existe una tendencia XXXXXX

2.2.2 Media Movil

Crearemos a continuacion 3 medias moviles para el objeto ts. Estas tendran 3, 5 y 7 periodos para su calculo.

```
mm3 <- rollmean(indice.ts,k=3)
cat("Media Movil con 3 meses: ", mm3,"\n\n")
```

```
## Media Movil con 3 meses:  2.60569e+12 2.657587e+12 2.741279e+12 2.852323e+12 2.9805...
```

```
mm5 <- rollmean(indice.ts,k=5)
cat("Media Movil con 5 meses: ", mm5,"\n\n")
```

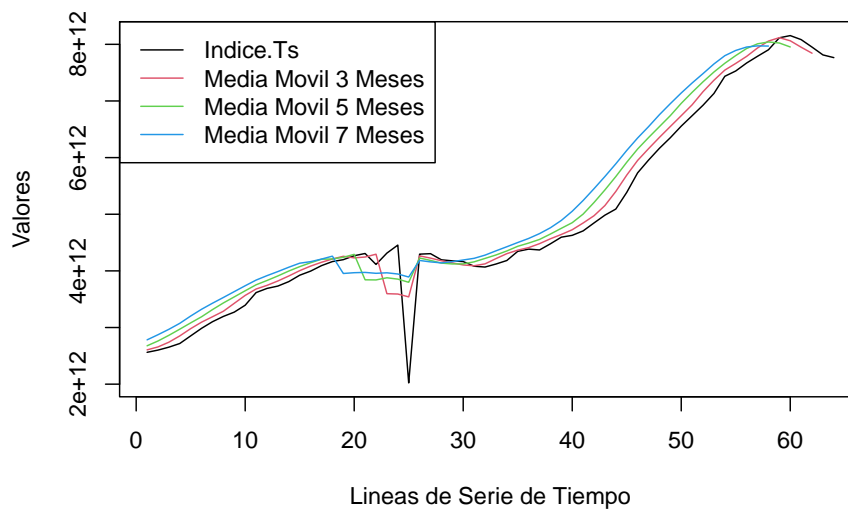
```
## Media Movil con 5 meses:  2.677519e+12 2.762363e+12 2.862556e+12 2.97112e+12 3.0820...
```

```
mm7 <- rollmean(indice.ts,k=7)
cat("Media Movil con 7 meses: ", mm7)
```

```
## Media Movil con 7 meses:  2.782362e+12 2.872921e+12 2.968758e+12 3.074575e+12 3.203034e+12 3.3
```

Veamos como es el comportamiento de las mismas en comparacion con lod datos originales de la serie de tiempo

```
plot(1:length(indice.ts), indice.ts, type = "l",
     ylim = c(min(indice.ts), max(indice.ts)),
     xlab = "Lineas de Serie de Tiempo", ylab = "Valores")
lines(1:length(mm3),mm3,type = "l", col=2)
lines(1:length(mm5),mm5,type = "l", col=3)
lines(1:length(mm7),mm7,type = "l", col=4)
legend("topleft",
      c("Indice.Ts", "Media Movil 3 Meses", "Media Movil 5 Meses", "Media Movil 7 Meses"),
      lty = 1, col = 1:4)
```



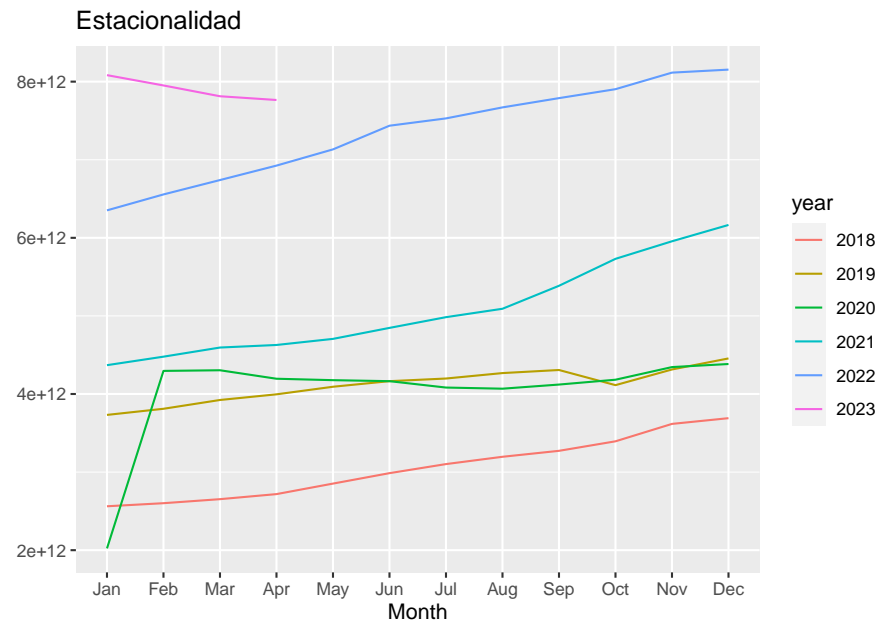
Chapter 3

Estacionalidad y Descomposicion

3.1 Estacionalidad

```
library(forecast)
library(timsac)
library(ggplot2)
library(changepoint)
library(readxl)

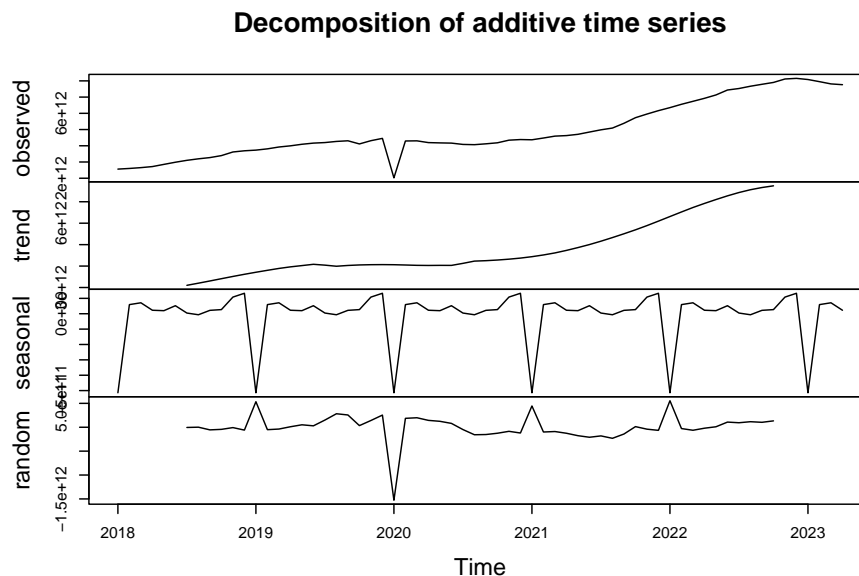
ggseasonplot(x = indice.ts,
             main= "Estacionalidad")
```



3.2 Ahora, hagamos una decomposicion del objeto y la graficamos para su analisis

```
descomposicion <- decompose(x=indice.ts)
plot(descomposicion)
```

3.2. AHORA, HAGAMOS UNA DECOMPOSICION DEL OBJETO Y LA GRAFICAMOS PARA SU ANALISIS



A pesar de presentar un patron recurrente en el componente de estacionalidad, se puede observar un trend en la serie de datos. De igual manera el error no se ve aleatorio sino que por el contrario, presenta un patron constante. Dicho esto, vamos a comprobar por medio del Augmented Dickery-Fuller test (adf) la estacionalidad del conjunto de datos.

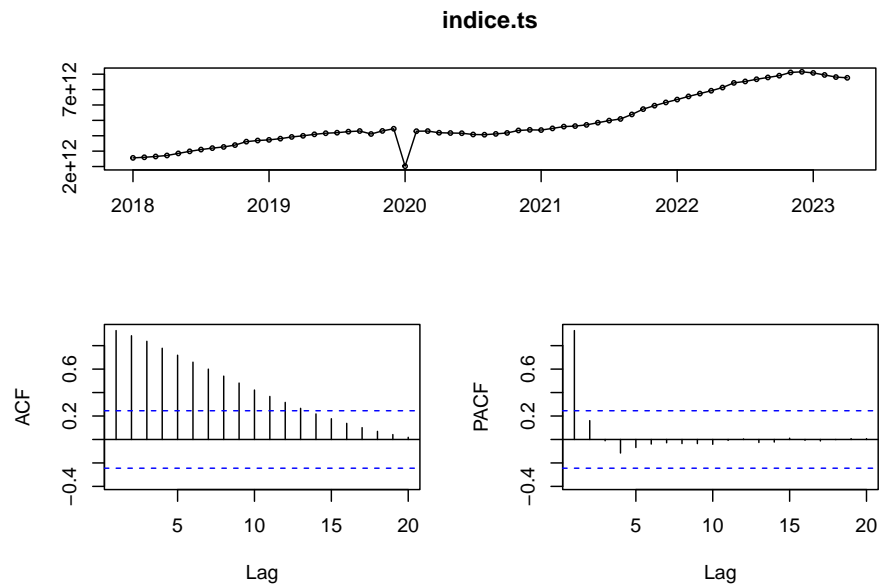
```
library(tseries)
adf.test(indice.ts)
```

```
##
## Augmented Dickey-Fuller Test
##
## data: indice.ts
## Dickey-Fuller = -1.2017, Lag order = 3, p-value = 0.8984
## alternative hypothesis: stationary
```

De acuerdo al resultado ($p\text{-value} > 0.05$), debemos aceptar la H_0 la cual nos confirma la no-estacionalidad del conjunto. Esto quiere decir que el objeto `indice.ts` requerira una transformacion para su posterior procesamiento en el modelo.

Ahora veamos si existe autocorrelacion total o parcial.(acf y pacf tests)

```
tsdisplay(indice.ts, lag.max = 20)
```



Ahora separamos la Estacionalidad y la graficamos

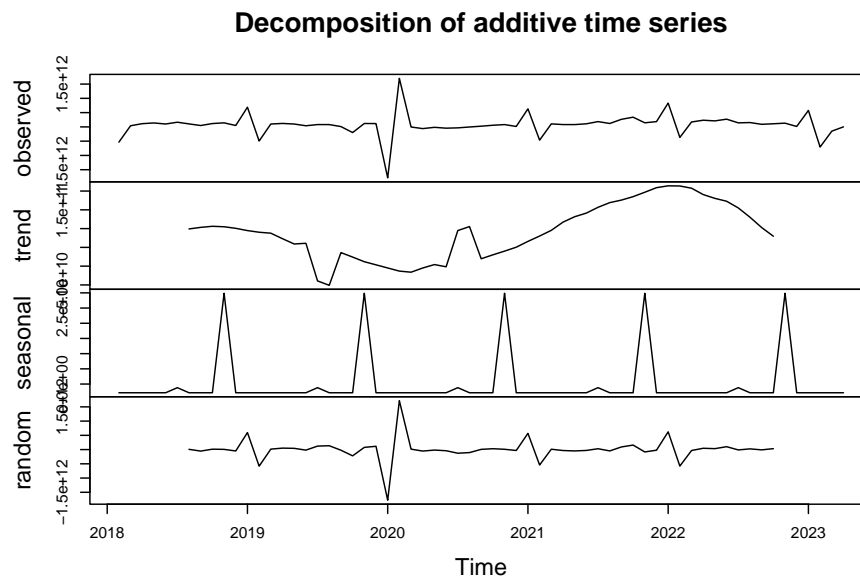
```
library(forecast)

# Ajustamos la estacionalidad de la serie de tiempo por medio del comando seasadj
indice_adj <- seasadj(descomposicion)

#Luego removemos la tendencia (trend) con el comando diff
modelo_ts <- diff(indice_adj)

# Grafiquemos la nueva serie de tiempo
nueva_descomp <- decompose(modelo_ts)
plot(nueva_descomp)
```


3.2. AHORA, HAGAMOS UNA DECOMPOSICION DEL OBJETO Y LA GRAFICAMOS PARA SU ANALISIS

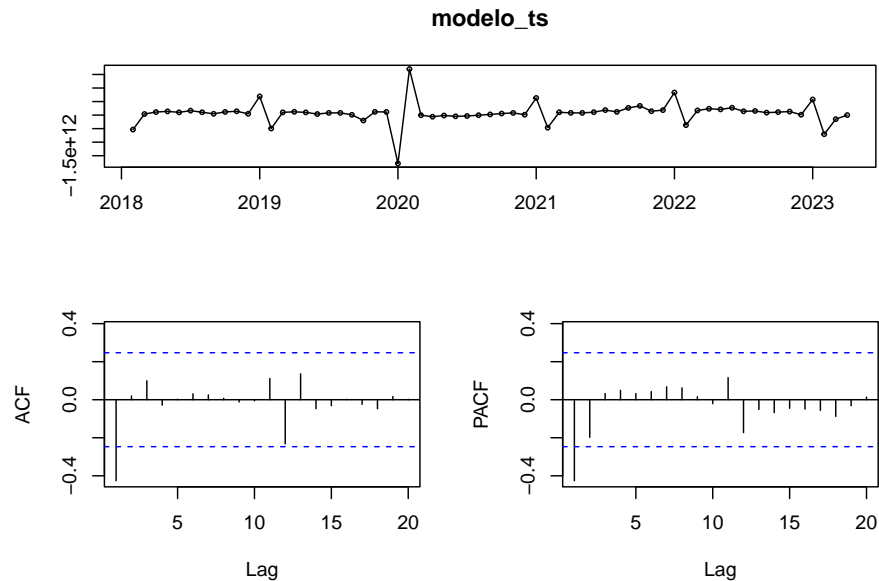


Por ultimo, apliquemos los mismos test que corrimos para nuestra serie de tiempo original.

```
# Test de estacionalidad  
adf.test(modelo_ts)
```

```
##  
## Augmented Dickey-Fuller Test  
##  
## data: modelo_ts  
## Dickey-Fuller = -3.7052, Lag order = 3, p-value = 0.0316  
## alternative hypothesis: stationary
```

```
# Test de autocorrelacion total y parcial  
tsdisplay(modelo_ts, lag.max = 20)
```



Podemos encluir que nuestra nueva serie de tiempo, etc

Ahora separamos la Estacionalidad y la graficamos

```
##### descomposicion

# Descomponer la serie de tiempo en sus componentes
#descomposicion <- decompose(indice.ts)

# Obtener la componente de estacionalidad
estacionalidad <- descomposicion$seasonal

# Convertir la componente de estacionalidad a un data frame
df_estacionalidad <- data.frame(periodo = time(estacionalidad), valor = estacionalidad)

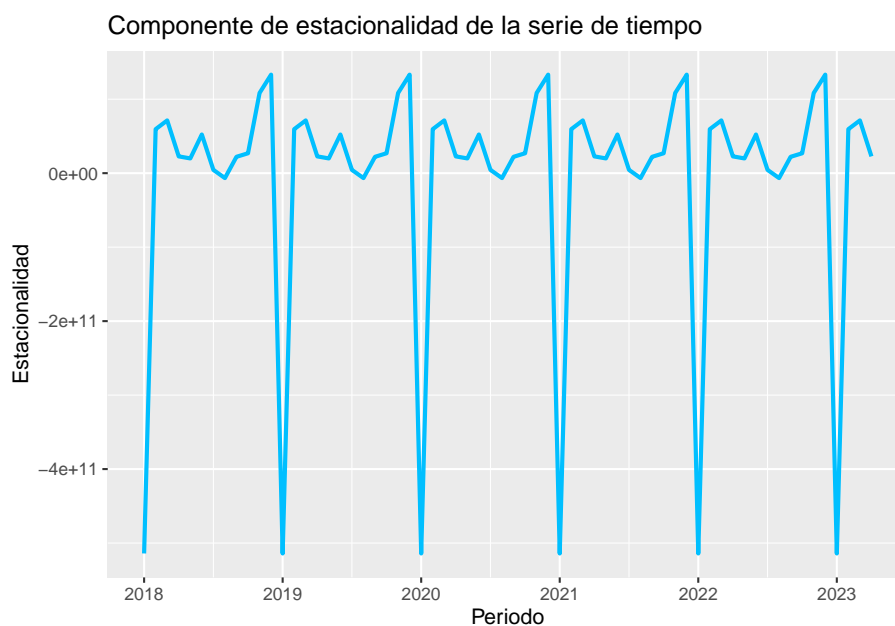
# Graficar la componente de estacionalidad
ggplot(df_estacionalidad, aes(x = periodo, y = valor)) +
  geom_line(color = "deepskyblue", size = 1) +
  xlab("Periodo") + ylab("Estacionalidad") +
  ggtitle("Componente de estacionalidad de la serie de tiempo")

## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
```

3.2. AHORA, HAGAMOS UNA DECOMPOSICION DEL OBJETO Y LA GRAFICAMOS PARA SU ANALISIS19

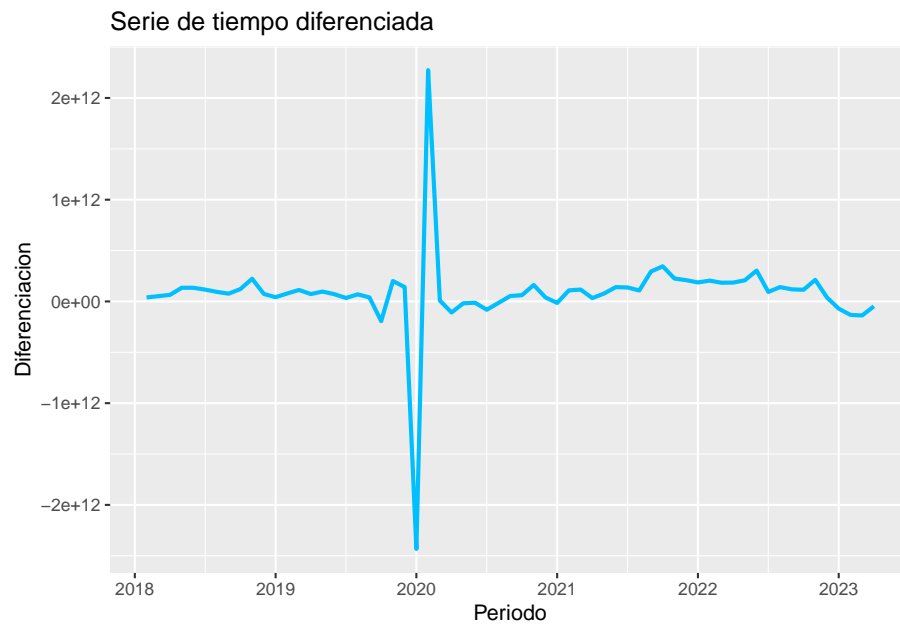
```
## This warning is displayed once every 8 hours.  
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was  
## generated.
```

```
## Don't know how to automatically pick scale for object of type <ts>. Defaulting  
## to continuous.  
## Don't know how to automatically pick scale for object of type <ts>. Defaulting  
## to continuous.
```



```
##### diferenciación  
  
# Calcular la diferenciación  
diferenciacion <- diff(indice.ts)  
  
# Convertir la serie de tiempo diferenciada a un data frame  
df_diferenciacion <- data.frame(periodo = time(diferenciacion), valor = diferenciacion)  
  
# Graficar la serie de tiempo diferenciada utilizando "ggplot2"  
ggplot(df_diferenciacion, aes(x = periodo, y = valor)) +  
  geom_line(color = "deepskyblue", size = 1) +  
  xlab("Periodo") + ylab("Diferenciacion") +  
  ggtitle("Serie de tiempo diferenciada")
```

```
## Don't know how to automatically pick scale for object of type <ts>. Defaulting  
## to continuous.  
## Don't know how to automatically pick scale for object of type <ts>. Defaulting  
## to continuous.
```



Chapter 4

Methods

We describe our methods in this chapter.

Math can be added in body using usual syntax like this

4.1 math example

p is unknown but expected to be around $1/3$. Standard error will be approximated

$$SE = \sqrt{\left(\frac{p(1-p)}{n}\right)} \approx \sqrt{\frac{1/3(1-1/3)}{300}} = 0.027$$

You can also use math in footnotes like this¹.

We will approximate standard error to 0.027^2

¹where we mention $p = \frac{a}{b}$

² p is unknown but expected to be around $1/3$. Standard error will be approximated

$$SE = \sqrt{\left(\frac{p(1-p)}{n}\right)} \approx \sqrt{\frac{1/3(1-1/3)}{300}} = 0.027$$

Chapter 5

Applications

Some *significant* applications are demonstrated in this chapter.

5.1 Example one

5.2 Example two

Chapter 6

Final Words

We have finished a nice book.