Prefacio

La econometría financiera es una disciplina que aúna la utilización de modelos estadísticos y matemáticos con la programación computacional yel estudio riguroso de instrumentos financiertos. Como subdisciplina de las finanzas cuantitativas, emplea un trilingüismo como enfoque natural, alternando lenguaje común con código y matemáticas. Este es precisamente el enfoque empleado en este manual, cuyo objetivo es el de dar un tratamiento riguroso a la par que pragmático de la disciplina. Para ello, se alterna el español con la formulación matemática y el código en el lenguaje de programación en estadística computacional R. Para el correcto abordaje de este libro, se precisan conocimientos básicos de inferencia estadística, matemática aplicada, y finanzas. Conocimientos elementales de programación son un añadido importante pero no un requisito indispensable.

El *leit motif* de este libro es claramente el estudio del análisis temporales. La causa de esto es clara: gran parte de los instrumentos financieros operan diaramiente en los mercados financieros, por lo que el estudio riguroso de las series temporales es por tanto una de las habilidades más demandadas en los trabajos de índole financiera en la industria y academia.

De esta forma, el primer capítulo del libro comienza ofreciendo una introducción a las series temporales. El segundo capítulo introduce aspectos fundamentales para la modelización econométrica de series temporales financieras, tales como las tendencias o los supuestos de estacionareidad. Una vez cubiertas las necesidades básicas indispensables en los dos capítulos anteriores, el tercer capítulo expone los aspectos fundamentales del análisis de regresión de series temporales.

Los tres primeros capítulos pueden pensarse como una introducción básica a la econometría financiera, mientras que los tres últimos son de corte más avanzado. En particular, el capítulo 4 muestra como emplear procesos estocásticos para modelizar series temporales, con a menudo óptimos resultados en ámbitos como la predicción del valor de instrumentos financieros. El quinto capítulo, por otra parte, introduce al lector al análisis cuantitativo del riesgo financiero mediante modelos para la volatilidad. Finalmente, el sexto capítulo muestra como modelizar shocks de diversos tipos, además de conceptos avanzados de comovimiento de series financieras tales como la cointegración.

A lo largo del libro, el contenido teórico se intercala con ejemplos prácticos para el refuerzo de la comprensión de los contenidos. Además, muchos capítulos cuentan con ejercicios a resolver por el lector, de forma que se asegure el correcto aprendizaje de los conocimientos.

1 Introducción a las series temporales

Los datos de series temporales permiten responder preguntas de inferencia y predicción en relación a una variable cuyo valor no varía entre sujetos, sino que varía conforme pasa el tiempo. Esto implica trabajar con variables del tipo

$$X_t : X \in \mathbb{R}, t = 1, \ldots, n$$

Es decir, la variable X toma un valor de los números reales para cada periodo t , que varía desde el primero (1) hasta el último (n) .

Para trabajar con R en series temporales, se usa la función `ts()`

```
x = data.frame(date=c(2000,2001,2002,2003,2004), z=c(4, 7, 21, 54, 64))
y = ts(x)
class(x)
```

```
[1] "data.frame"
```

```
class(y)
```

```
[1] "mts"    "ts"     "matrix"
```

```
head(y)
```

```
     date  z
[1,] 2000  4
[2,] 2001  7
[3,] 2002 21
[4,] 2003 54
[5,] 2004 64
```

Si los datos son anuales, R solo necesita la fecha de inicio.

```
y=ts(c(4, 7, 21, 54, 64), start = 2000)
head(y)
```

```
[1]  4  7 21 54 64
```

En caso de que las observaciones de la variable se produzcan en una frecuencia distinta a la anual, basta con añadir el comando `frequency`

```
x=rnorm(1200, mean=50, sd=20)
y=ts(x, start=2000, frequency=12)
summary(y)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -8.978  35.662  50.334  49.638  62.902 116.484
```

De forma que frecuencias menores a la anual se especifican teniendo en cuenta el número de veces (k) que hay que dividir el año, o *k-ésimos* de año. Si la frecuencia es mensual, el año se divide en 12 meses ($k = 12$), si es trimestral, el año se divide en 4 trimestres ($k = 4$), y así sucesivamente.

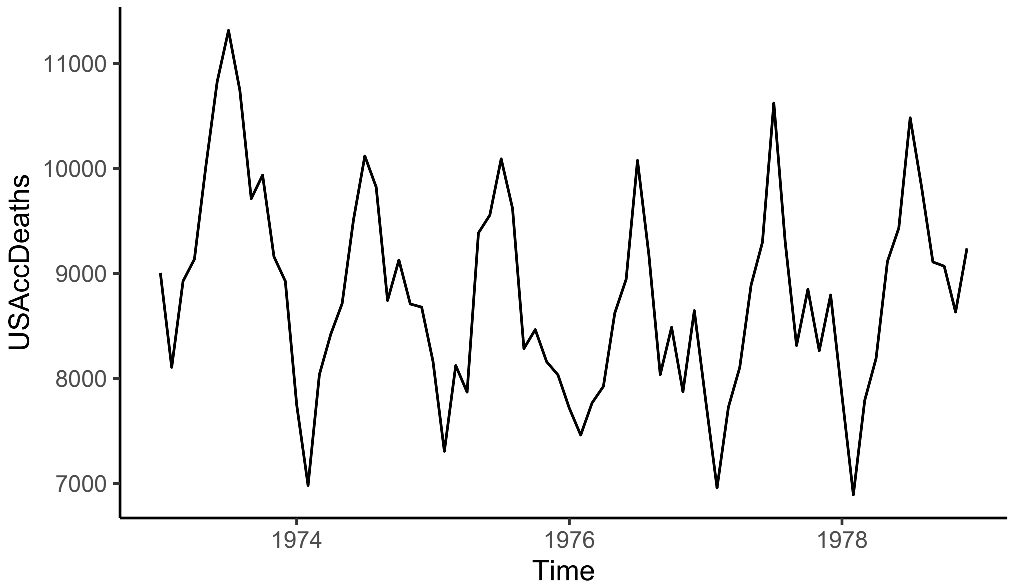
Frecuencia	k
Anual	1
Semestral	2
Trimestral	3
Cuatrimestral	4
Mensual	12
Semanal	52
Diaria	365

1.1 Gráfica de series temporales

El dicho popular *"una imagen vale más que mil palabras"* es perfectamente aplicable al caso de las variables de tipo serie temporal, donde el visualizar tablas de datos es poco informativo. La opción más útil en estos casos es empezar el análisis con un gráfico de serie temporal, que tenga el valor de la variable en el eje y, y el paso del tiempo en el eje x.

La librería `forecast` es especialmente útil para este tipo de gráficos. Si la variable es de series temporales, el comando `autoplot` producirá un gráfico de series temporales.

```
library(forecast, quietly=TRUE)
library(ggplot2, quietly=TRUE)
data("USAccDeaths")
autoplot(USAccDeaths) + theme_classic()
```



El gráfico de series temporales arroja mucha más información que sus estadísticos descriptivos. Por ejemplo, es fácilmente interpretable que el pico de muertos por accidentes se produce en el primer año de la serie, que sigue a una caída drástica en el número de muertes por accidentes.

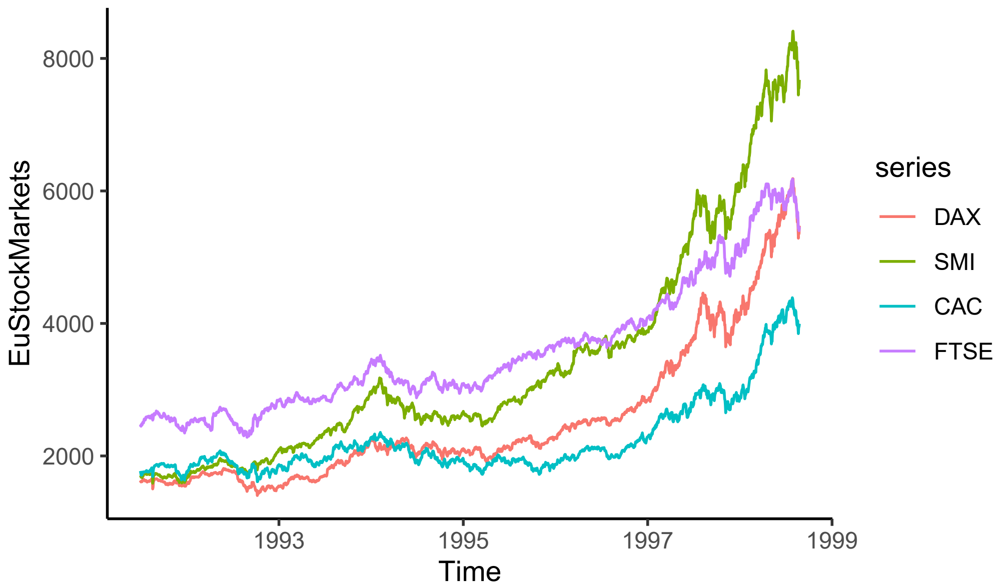
Además, es facilmente percibible el hecho de que, a mitad de año, el número de muertes por accidentes aumenta sustancialmente, para después caer a finales del año y principios del siguiente. A mediados del año la estación del año es verano, fecha en la que los trabajadores suelen tomar vacaciones y aumentan actividades como el nado en playas y piscinas que pueden dar lugar a ahogamientos. También aumentan viajes por carretera que pueden dar lugar a accidentes. A finales del año y principios del siguiente, es invierno, fecha en la que se reducen considerablemente tanto el baño como el desplazamiento. Esto es lo que se denomina como *patrón estacional* o simplemente *estacionalidad,* haciendo evidente que el comportamiento de la serie depende de la estación del año en que se encuentre. Otros ejemplos claros de patrones estacionales tienen que ver con las temperaturas o la lluvia.

En el paquete `Forecast`, además de en otros paquetes como `Ecdat`, se pueden encontrar muchos conjuntos de datos disponibles para el análisis. Sin embargo, el primero se caracteriza por contener variables de series temporales, como por ejemplo *EuStockMarkets.*

```
data("EuStockMarkets")
head(EuStockMarkets)
```

```
Time Series:
Start = c(1991, 130)
End = c(1991, 135)
Frequency = 260
             DAX    SMI    CAC   FTSE
1991.496 1628.75 1678.1 1772.8 2443.6
1991.500 1613.63 1688.5 1750.5 2460.2
1991.504 1606.51 1678.6 1718.0 2448.2
1991.508 1621.04 1684.1 1708.1 2470.4
1991.512 1618.16 1686.6 1723.1 2484.7
1991.515 1610.61 1671.6 1714.3 2466.8
```

```
autoplot(EuStockMarkets) + theme_classic()
```



Como se puede comprobar, si se tienen varias series temporales apiladas en un mismo conjunto de datos, autoplot representa cada una por separado.

Este tipo de gráficos es muy útil pues permite obtener mucha información de manera compacta. Por ejemplo:

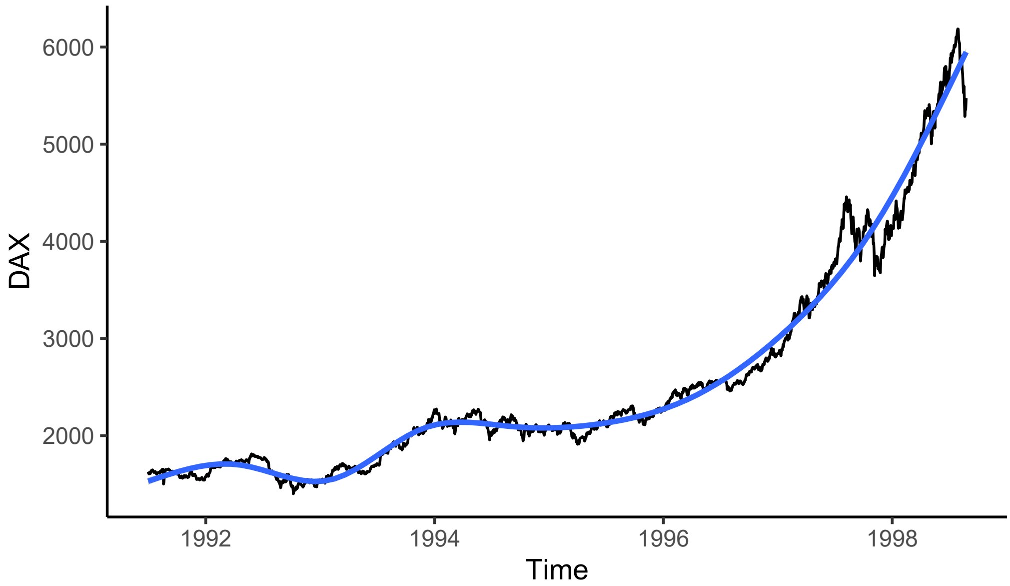
- El índice bursátil de acciones alemán (DAX) superó al francés (CAC) en 1994.
- Suiza (SMI) pasó de tener el penúltimo índice bursátil a tener el primero en valor alrededor de 1997.
- El crecimiento del índice bursátil inglés (FTSE) presenta una tendencia de tipo lineal y cierto declive, mientras que los índices suizo, alemán y francés presentan tendencias exponenciales de crecimiento. Esto podría implicar que los inversores interpretan que los últimos países experimentarán un gran crecimiento económico y, por tanto, prefieren depositar sus fondos en estos índices, mientras que esperan que el crecimiento económico de Inglaterra se estanque.
- Suiza (SMI) y Alemania (DAX) presentan patrones muy similares de crecimiento desde finales de 1997. Podría decirse que *co-mueven.*

1.2 Aspectos avanzados de las series temporales

Además del componente estacional, aquel componente repetitivo que afecta a la serie temporal de forma fija y conocida, existen otro tipo de componentes fundamentales de las series temporales.

La *tendencia* de una serie es su comportamiento a largo plazo. Se caracteriza por un crecimiento o decrecimiento lineal o no lineal. Se denota con T . En el contexto del modelo de regresión lineal, es posible capturar esta tendencia incorporando el parámetro T, el cual es igual a 1 en la primera observación, a 2, en la segunda, etc., con la respectiva forma funcional que corresponda (lineal, polinómica, exponencial, etc.).

```
library(ggplot2, quietly=TRUE)
forecast::autoplot(EuStockMarkets[,1]) + geom_smooth() + ylab("DAX") +
  theme_classic()
```



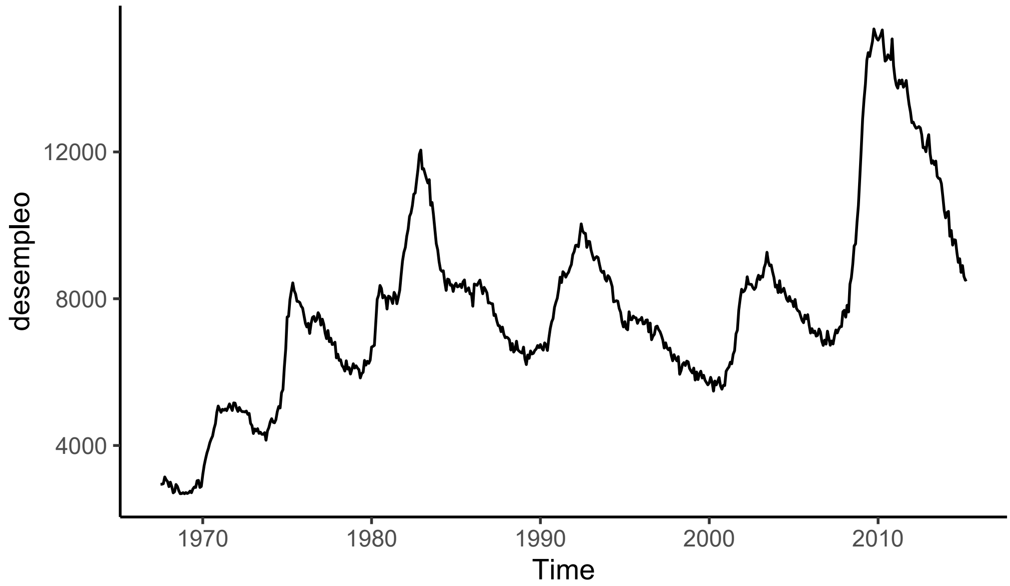
Se aprecia como la serie presenta una tendencia creciente al comienzo, una decreciente desde 1992 a 1993, una creciente desde 1993 a 1994, una relativamente estable entre 1994 y 1996, y un crecimiento exponencial desde 1996 hacia adelante.

El *ciclo* de una serie ocurre cuando esta presenta oscilaciones crecientes y decrecientes que no tienen una frecuencia fija, es decir, que no son de carácter estacional. A menudo suceden en series temporales económicas y están relacionadas con lo que se conoce como “ciclo económico”. Existe una popular descomposición, denominada la descomposición de Hodrick-Prescott, por la cual una serie se puede representar como

$$Y = T + C + \epsilon$$

siendo Y la serie, T la tendencia, C el ciclo y ϵ el término error.

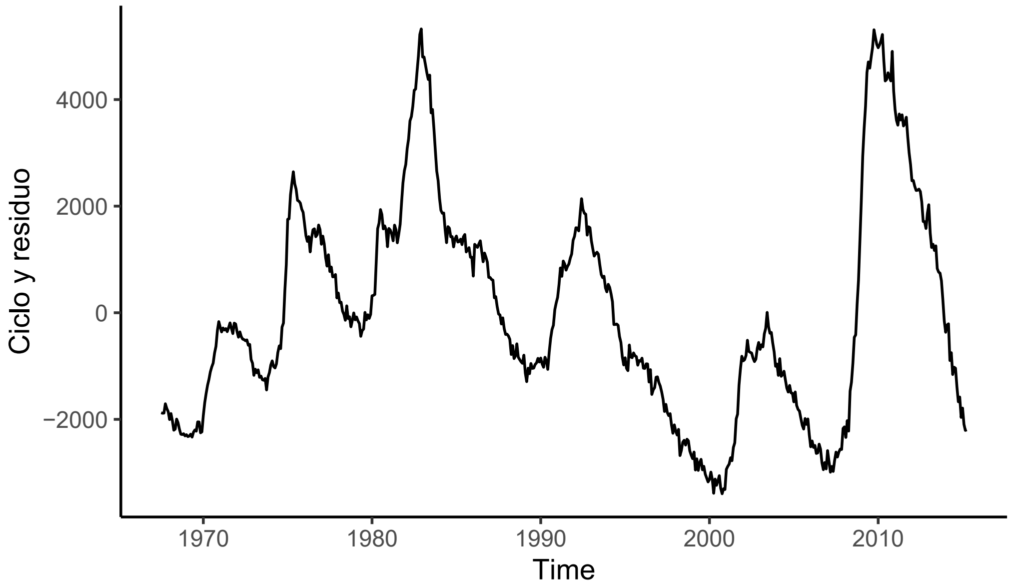
```
data(economics)
desempleo <- ts(economics$unemploy, economics$date, start=c(1967,7),
                end=c(2015,4), frequency=12)
autoplot(desempleo) + theme_classic()
```



Parece que la serie incorpora un componente cíclico bastante claro. Está claro que a los periodos de aumento de desempleados le siguen periodos de disminución en consonancia con las circunstancias económicas EEUU. No obstante, estos intervalos no siguen un patrón regular como en el caso del componente estacional. Por lo tanto, hablamos de ciclos, y no de estaciones del año.

Por otra parte, parece que el número de desempleados sigue una tendencia creciente a largo plazo. Esto es lógico si se tiene en cuenta el crecimiento poblacional (a más población, más desempleados y empleados). Vamos a capturar la tendencia a largo plazo con un modelo de regresión, y a representar el residuo, que corresponderá a los términos C y ϵ de la descomposición HP.

```
modelo <- lm(unemploy ~ date, data = economics)
residuos <- residuals(modelo)
serie <- ts(residuos, economics$date, start=c(1967,7), end=c(2015,4),
            frequency=12)
forecast::autoplot(serie) + ylab("Ciclo y residuo") + theme_classic()
```



Se aprecia claramente el componente cíclico de la serie de desempleados. El número de desempleados aumenta sustancialmente en:

- 1970-1980: Crisis del petróleo.
- 1980-1985: Segunda crisis del petróleo e inflación.
- 1990-1993: Recesión global de los primeros años de la década de 1990.
- 2001-2003: Burbuja financiera dot.com.
- 2007-2010: Crisis financiera o gran recesión.

lo que se corresponde con los periodos de recesión, crisis o contracción del ciclo económico. Por contra, el número de desempleados desciende con fuerza en los periodos anteriores y posteriores a los periodos de crisis, en lo que se conoce como periodos de expansión.

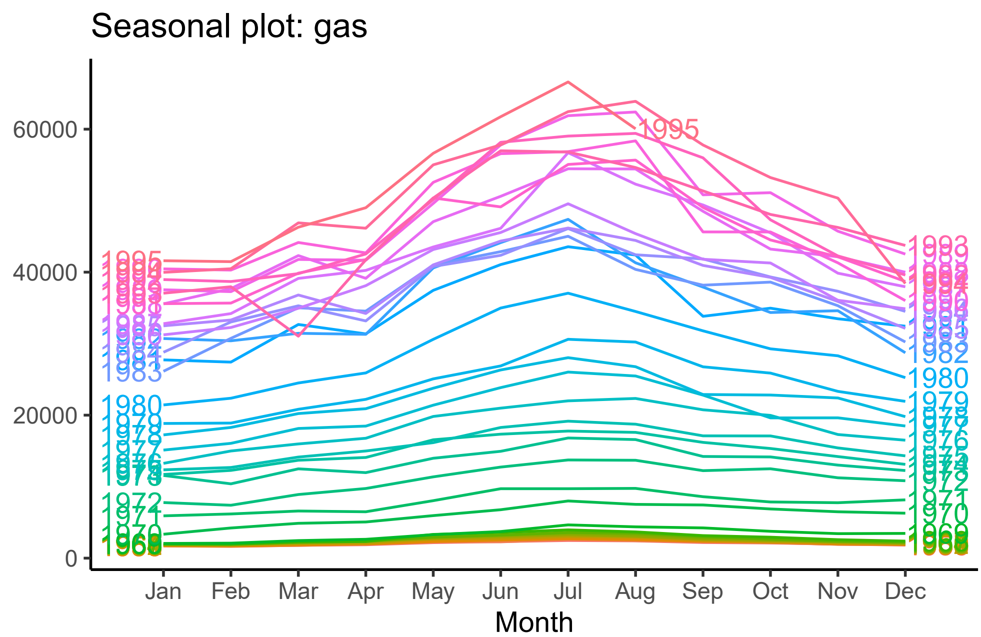
1.3 Tipos de gráficos para analizar series temporales

Existen algunos gráficos útiles para analizar más en profundidad las propiedades de las series temporales.

1.3.1 Gráfico estacional

Un gráfico estacional permite visualizar claramente el patrón estacional de la serie:

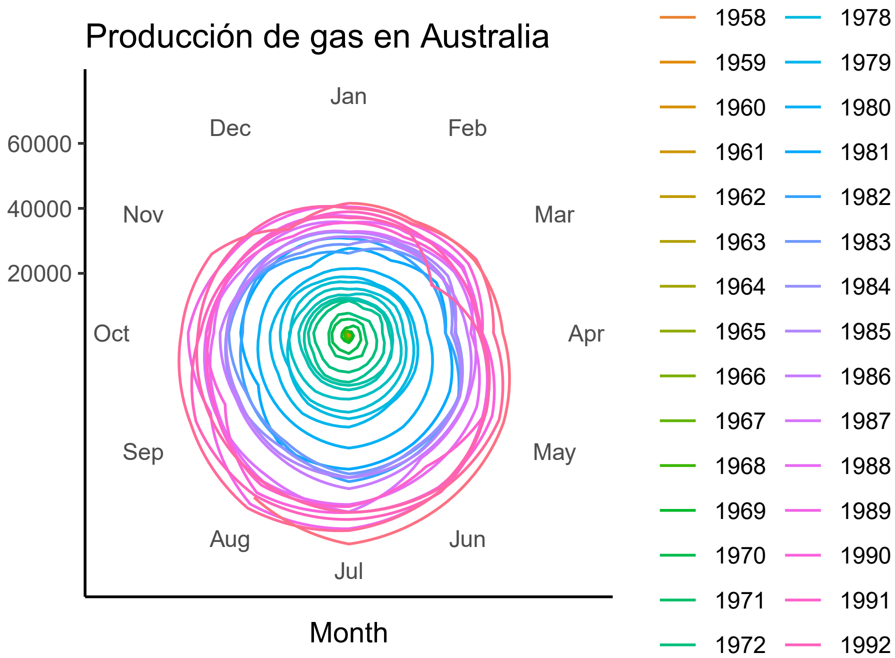
```
library(forecast, quietly=TRUE)
data(gas)
forecast::ggseasonplot(gas, year.labels=TRUE, year.labels.left=TRUE) +
  theme_classic()
```



Se observa que la producción de gas en Australia ha aumentado considerablemente con los años (tendencia creciente). En cuanto al componente estacional, se observa que la producción crece considerablemente en los meses de calor, con máximos en verano, y disminuye en los meses de frío, con mínimos en invierno.

Si se quiere obtener mayor comparativa entre meses, el comando ‘polar=TRUE‘ es bastante útil.

```
library(ggplot2, quietly=TRUE)
forecast::ggseasonplot(gas, polar=TRUE) +
  ggtitle("Producción de gas en Australia") + theme_classic()
```

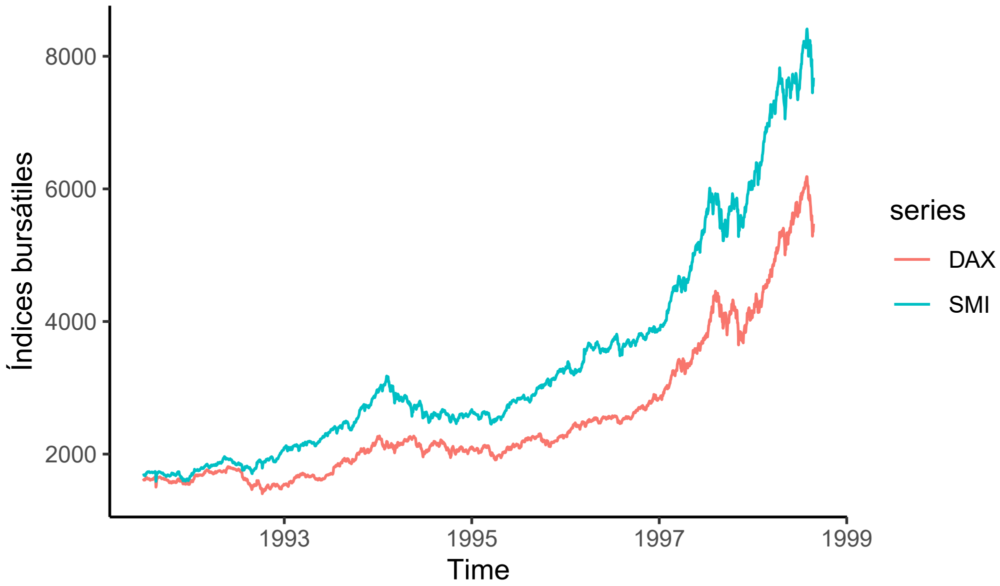


Claramente, los meses líderes en producción son julio y agosto, y los de menor producción, diciembre y enero.

1.3.2 Correlación entre series

Además de analizar los aspectos individuales de una serie temporal, se pueden calcular correlaciones entre series para medir el grado de relación entre dos o más series.

```
data("EuStockMarkets")
autoplot(EuStockMarkets[,c(1:2)]) + ylab("Índices bursátiles") +
  theme_classic()
```



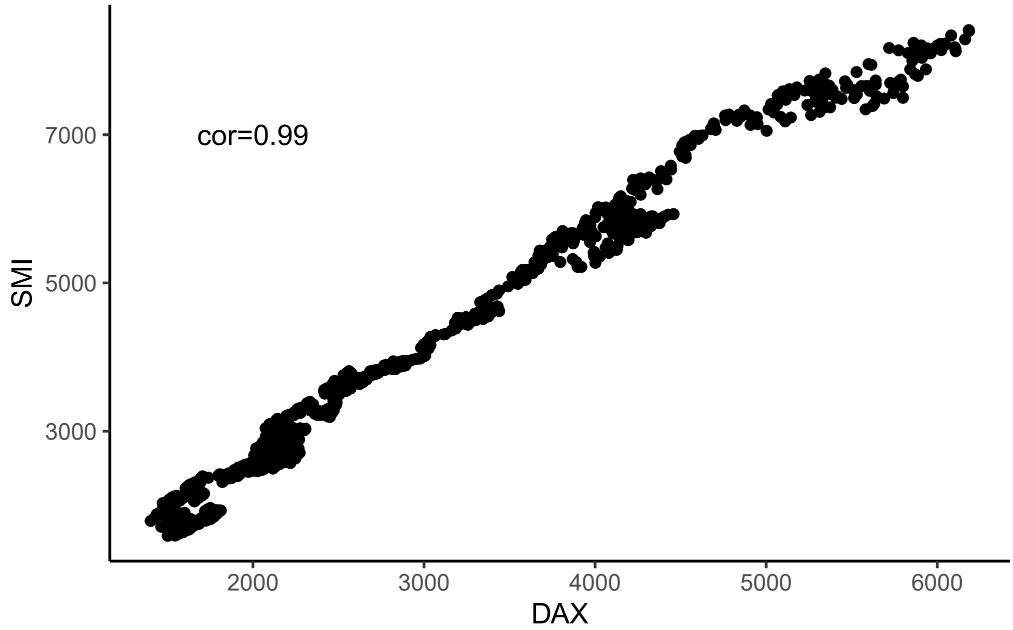
A priori, parece que ambos índices bursátiles guardarán una fuerte correlación. Para obtener una métrica formal de dicha correlación, es útil calcular el coeficiente de correlación entre las dos series:

$$\rho(X_t, Y_t) = \frac{\sigma(X_t Y_t)}{\sigma(X_t)\sigma(Y_t)} = \frac{Cov(X_t, Y_t)}{\sqrt{Var(X_t)Var(Y_t)}}$$

```
EUdf <- as.data.frame(EuStockMarkets)
cor(EUdf$DAX, EUdf$SMI)
```

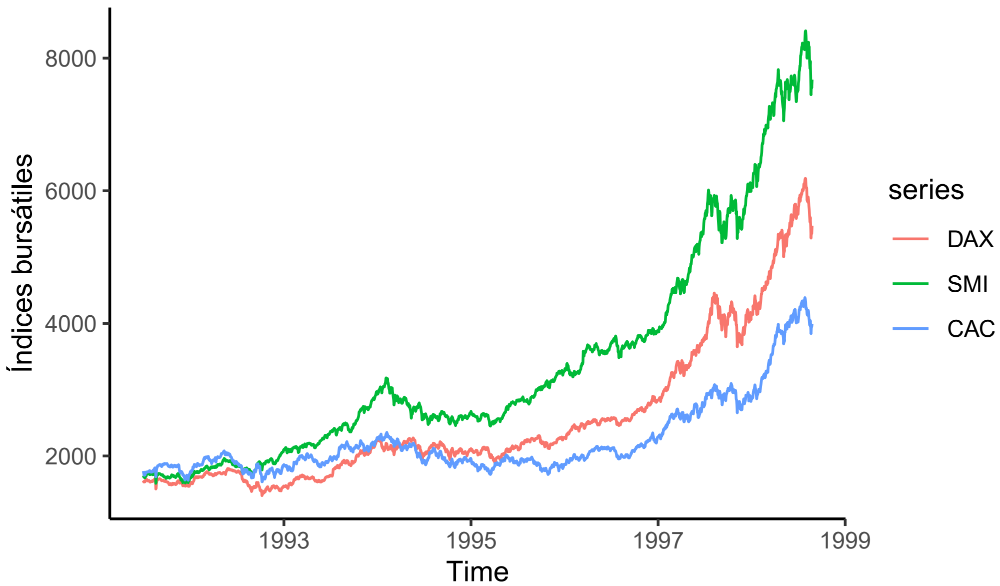
```
[1] 0.9911539
```

```
ggplot(data=EUdf, aes(x=DAX, y=SMI)) +
  geom_point() +
  annotate("text", label="cor=0.99", x=2000, y=7000, size=4, color="black")+
  theme_classic()
```



La correlación es casi perfecta. Veamos entre tres índices bursátiles a la vez

```
autoplot(EuStockMarkets[,c(1:3)]) + ylab("Índices bursátiles") +
  theme_classic()
```

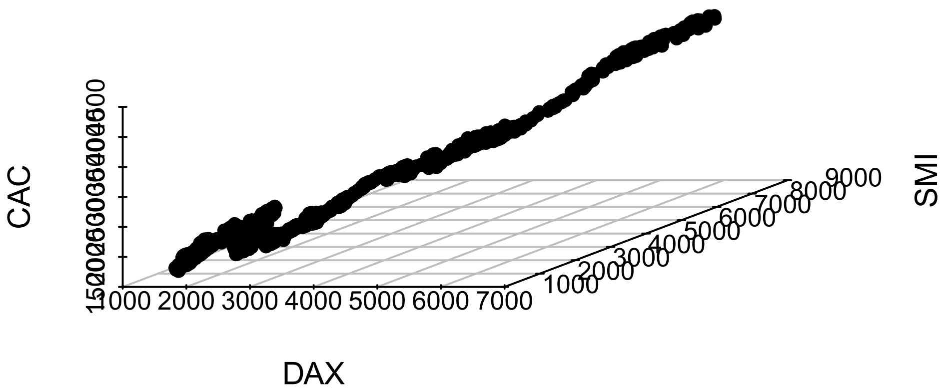


A priori debería de haber una correlación muy alta entre los tres.

```
attach(EUdf)
library(scatterplot3d, quietly=TRUE)
```

```
Warning: package 'scatterplot3d' was built under R version 4.2.3
```

```
scatterplot3d(DAX, SMI, CAC, grid=TRUE, box=FALSE, color = "black",
              pch = 16)
```



Efectivamente, hay una alta correlación, lo cual se interpreta de la variación lineal creciente de la nube de puntos.

Ejercicio: Cargue y analize la estacionalidad de la serie AirPassengers.