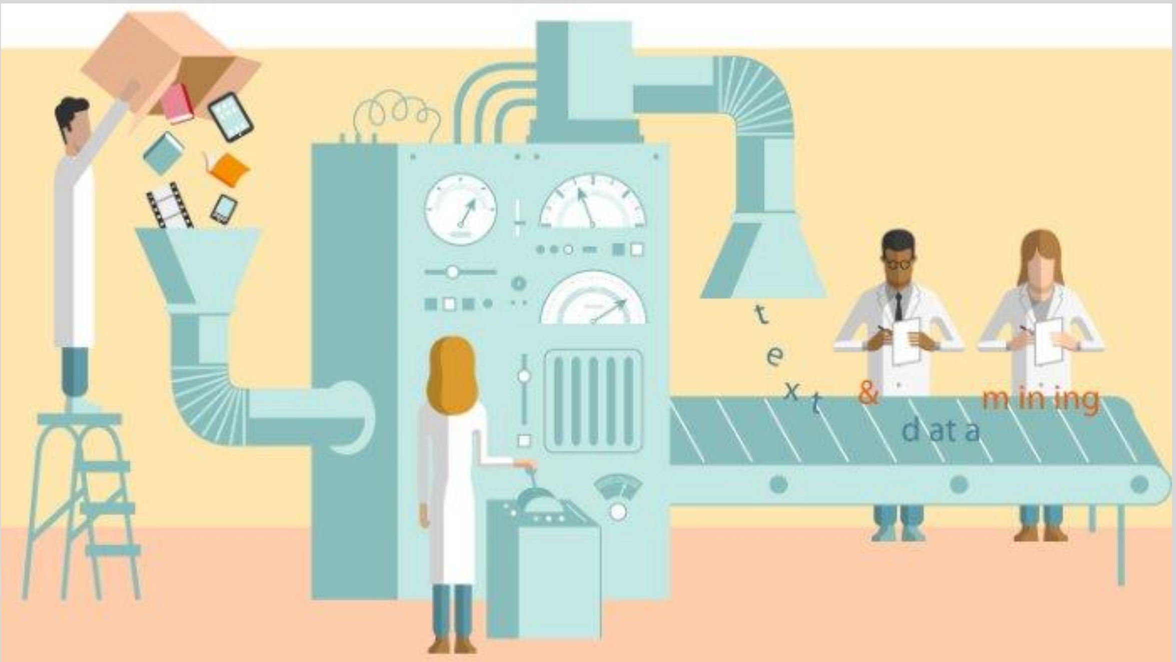


Digital House >
Coding School

DATA SCIENCE

MODULO 5

Series de Tiempo II

Transformaciones

Transformación para estabilizar la variación

- Si los datos muestran variaciones diferentes para diferentes niveles de la serie entonces una transformación puede ser útil.
- Los logaritmos, en particular, son útiles porque son más interpretables: los cambios en el log del valor son cambios relativos (porcentaje) en la escala original

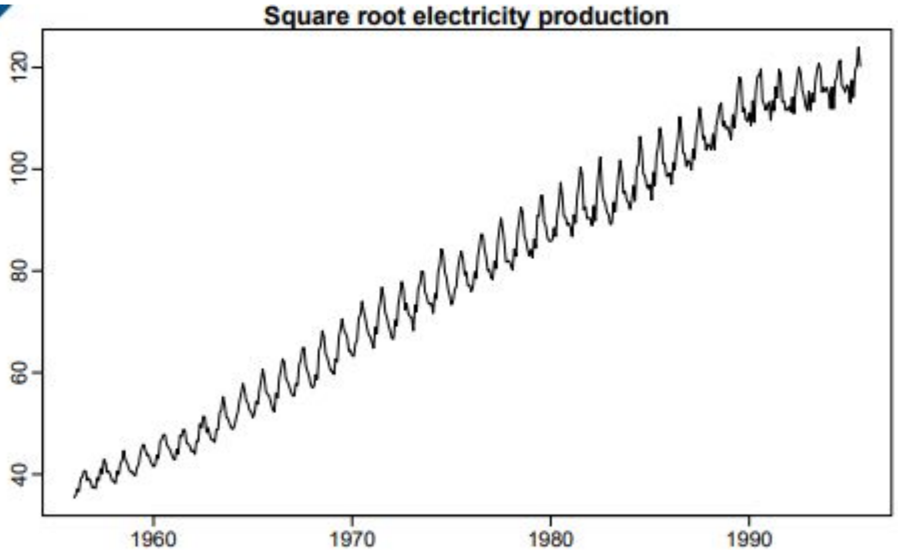
Transformaciones

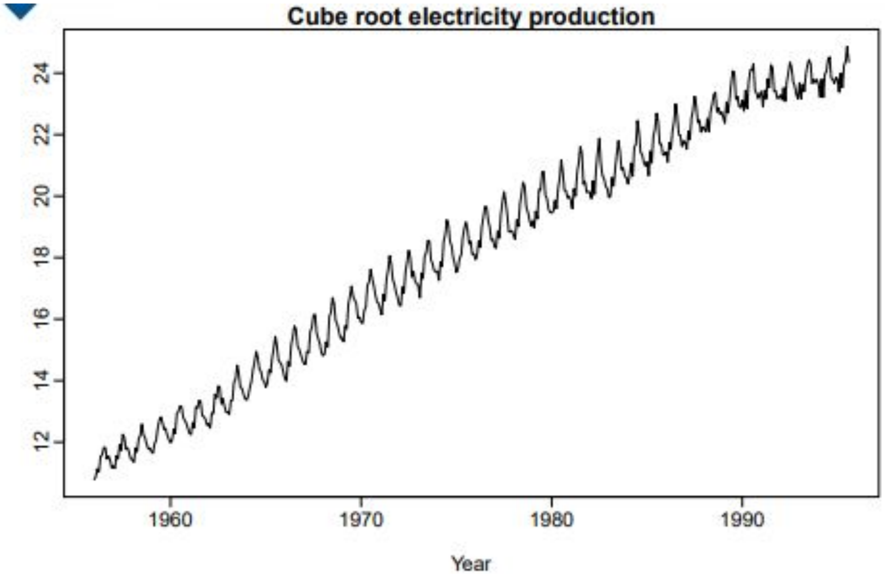
Transformación para estabilizar la variación

Square root	$w_t = \sqrt{y_t}$	$\downarrow$
Cube root	$w_t = \sqrt[3]{y_t}$	Increasing
Logarithm	$w_t = \log(y_t)$	strength

Transformaciones

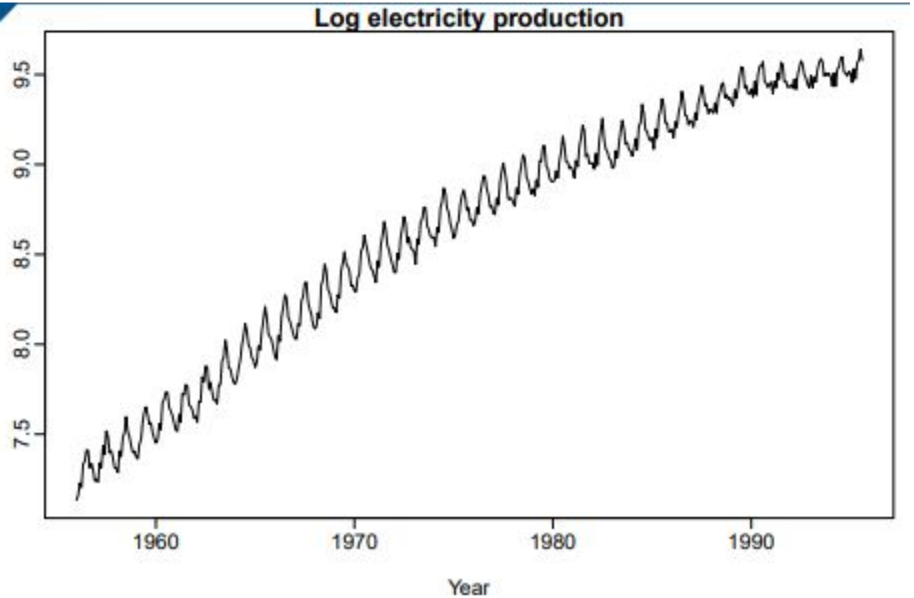
Transformación para estabilizar la variación

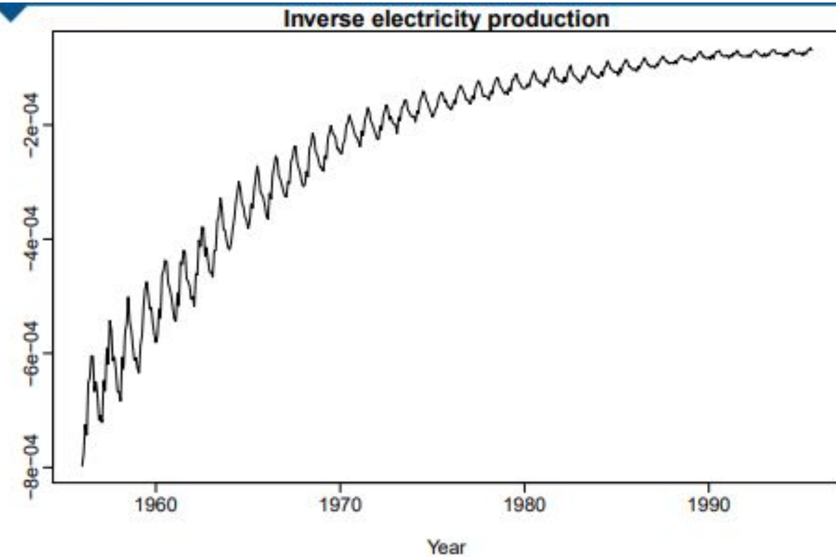




Transformaciones

Transformación para estabilizar la variación





Transformaciones

Transformación para estabilizar la variación

- Cada una de estas transformaciones es cercana a un miembro de la familia de transformaciones de Box-Cox

$$w_t = \begin{cases} \log(y_t), & \lambda = 0; \\ (y_t^\lambda - 1)/\lambda, & \lambda \neq 0. \end{cases}$$

- $\lambda = 1$: (No substantive transformation)
- $\lambda = \frac{1}{2}$: (Square root plus linear transformation)
- $\lambda = 0$: (Natural logarithm)
- $\lambda = -1$: (Inverse plus 1)

Back-transformation

- Debemos revertir la transformación (back-transform) para obtener forecasts en la escala original. Las transformaciones inversas de Box-Cox está dada por

$$y_t = \begin{cases} \exp(w_t), & \lambda = 0; \\ (\lambda W_t + 1)^{1/\lambda}, & \lambda \neq 0. \end{cases}$$

Un valor bajo de lamba puede dar intervalos de predicción extremadamente grandes

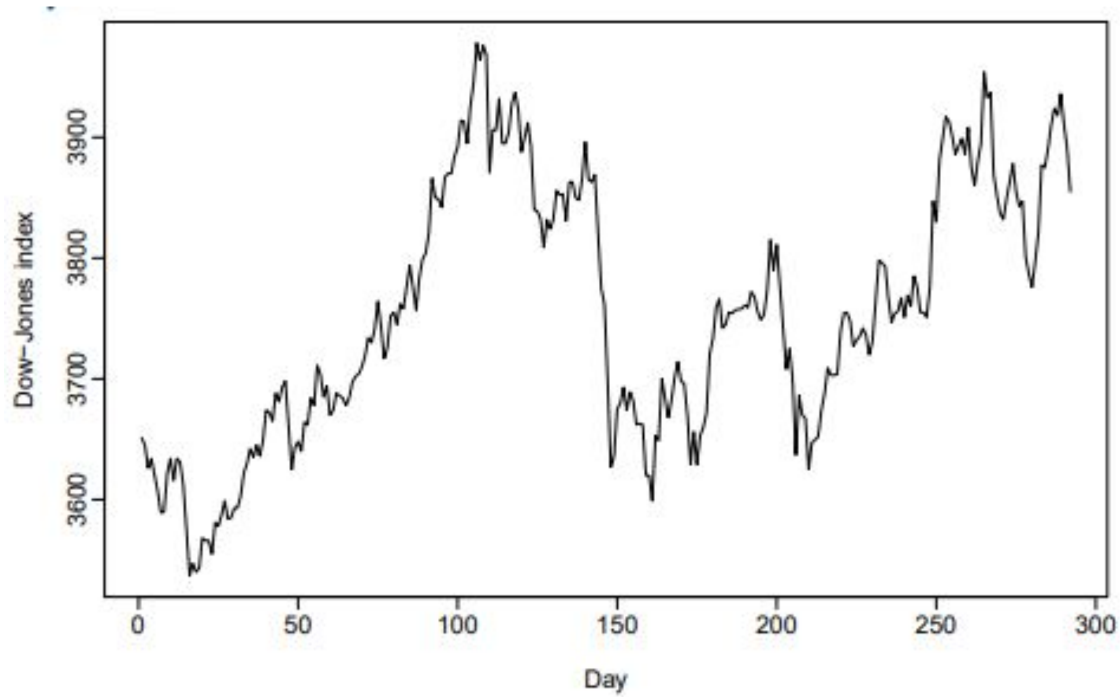
Estacionariedad

- Si $\{y_t\}$ es una serie estacionaria entonces la distribución de $(y_t, \ldots, y_{t+s})$ no depende de t.

- Una serie estacionaria es
  - Aproximadamente horizontal
  - tiene varianza constante
  - No tiene patrones predecibles en el largo plazo.

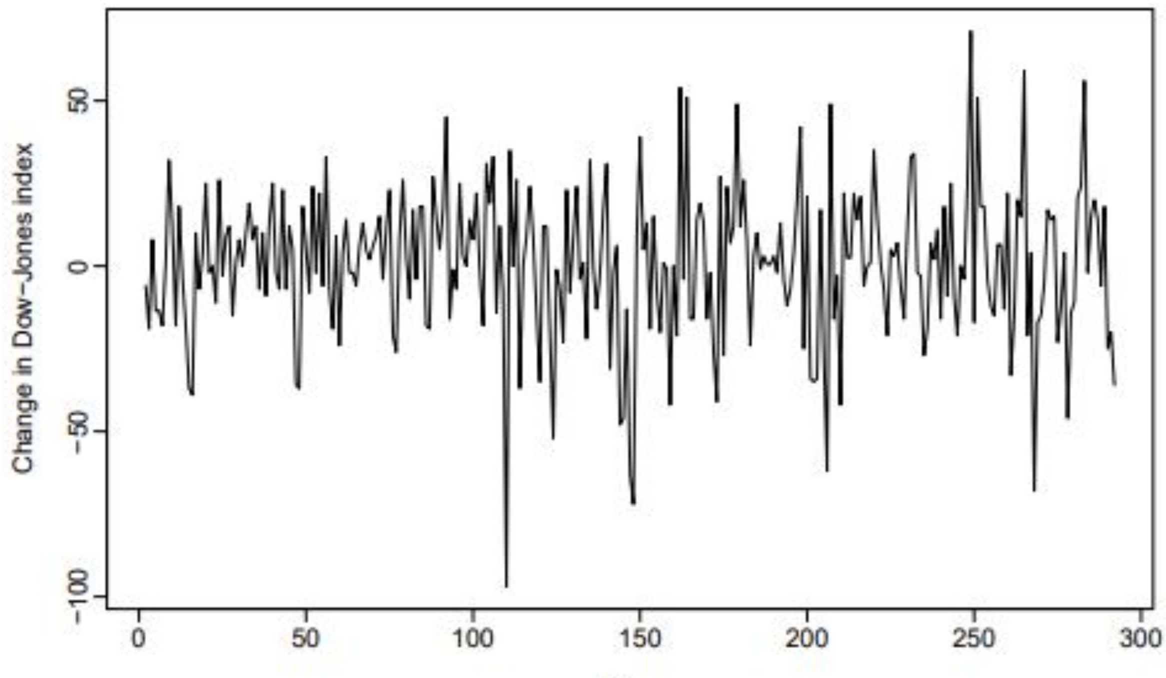
Estacionariedad

¿Estacionaria?



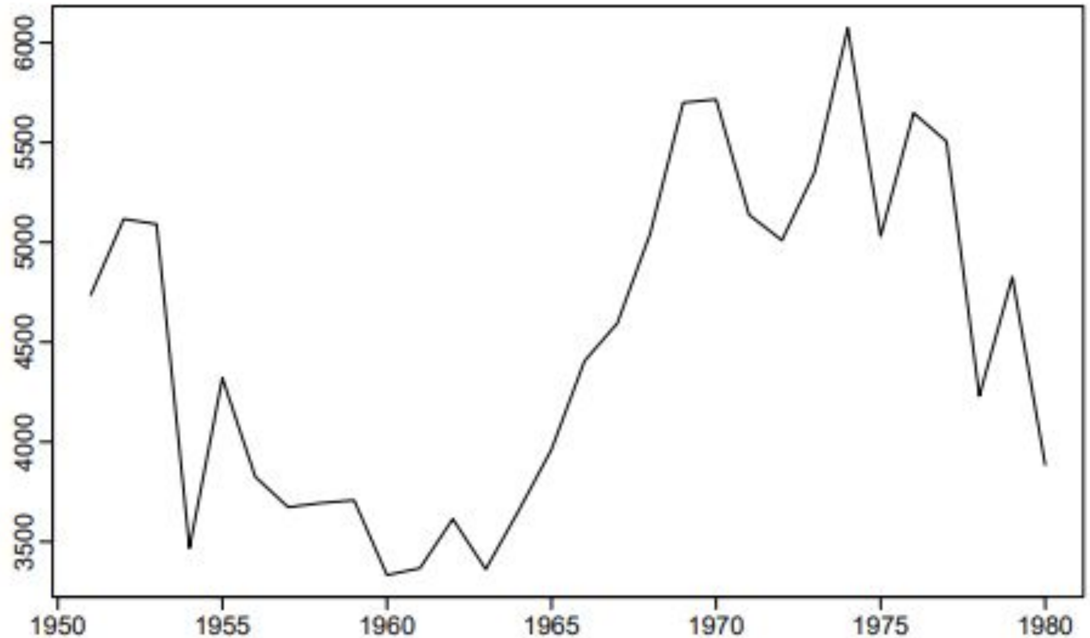
Estacionariedad

¿Estacionaria?



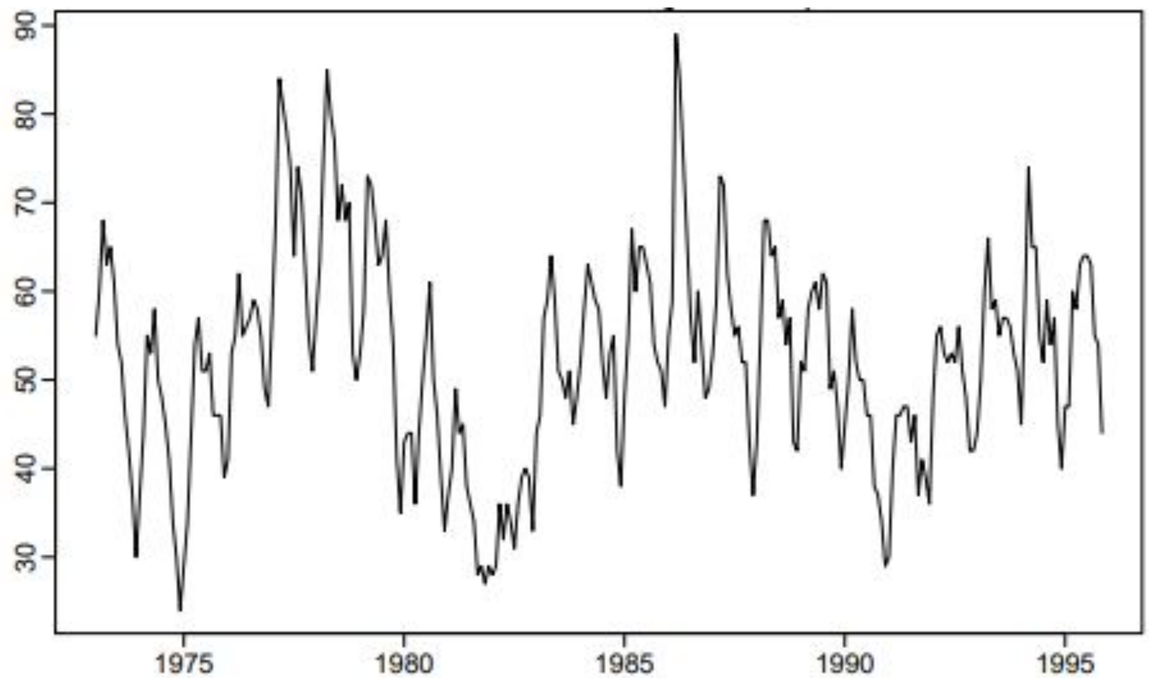
Estacionariedad

¿Estacionaria?



Estacionariedad

¿Estacionaria?

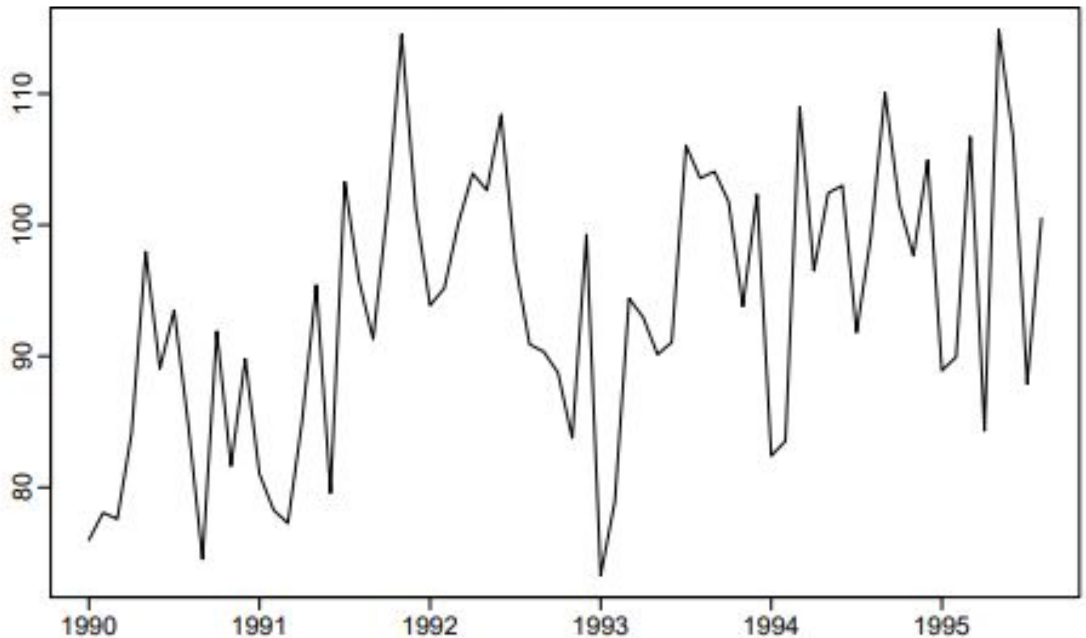


Estacionariedad

¿Estacionaria?

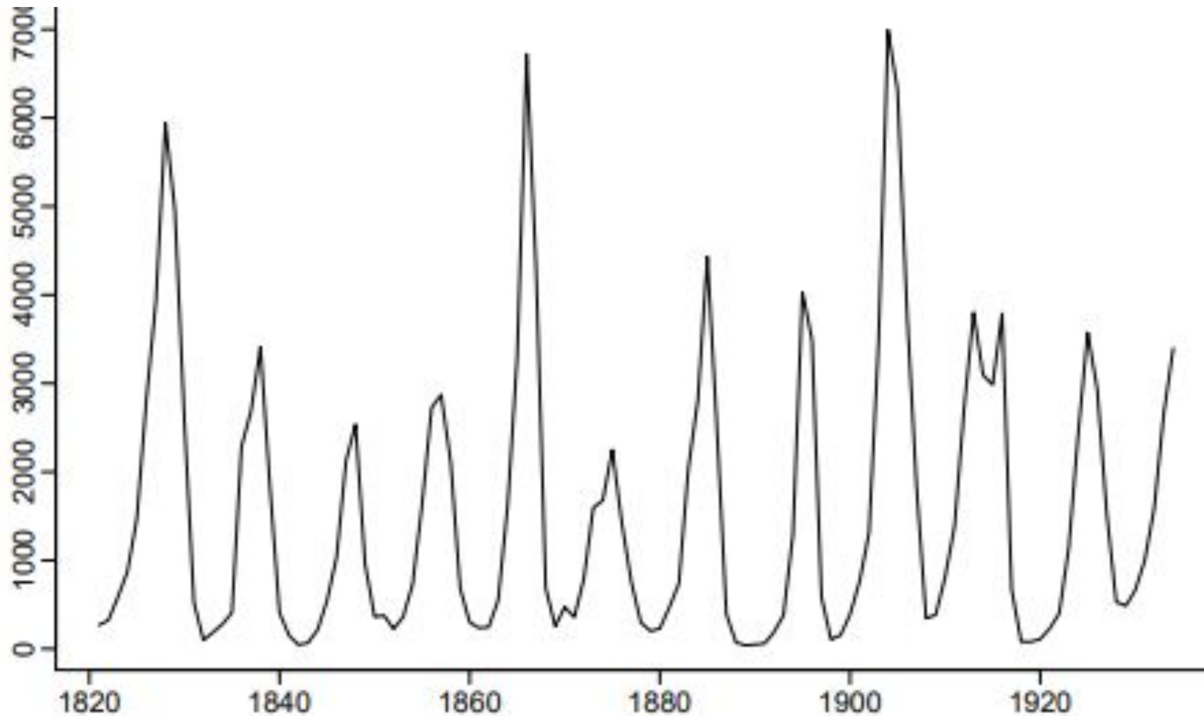
Estacionariedad

¿Estacionaria?



Estacionariedad

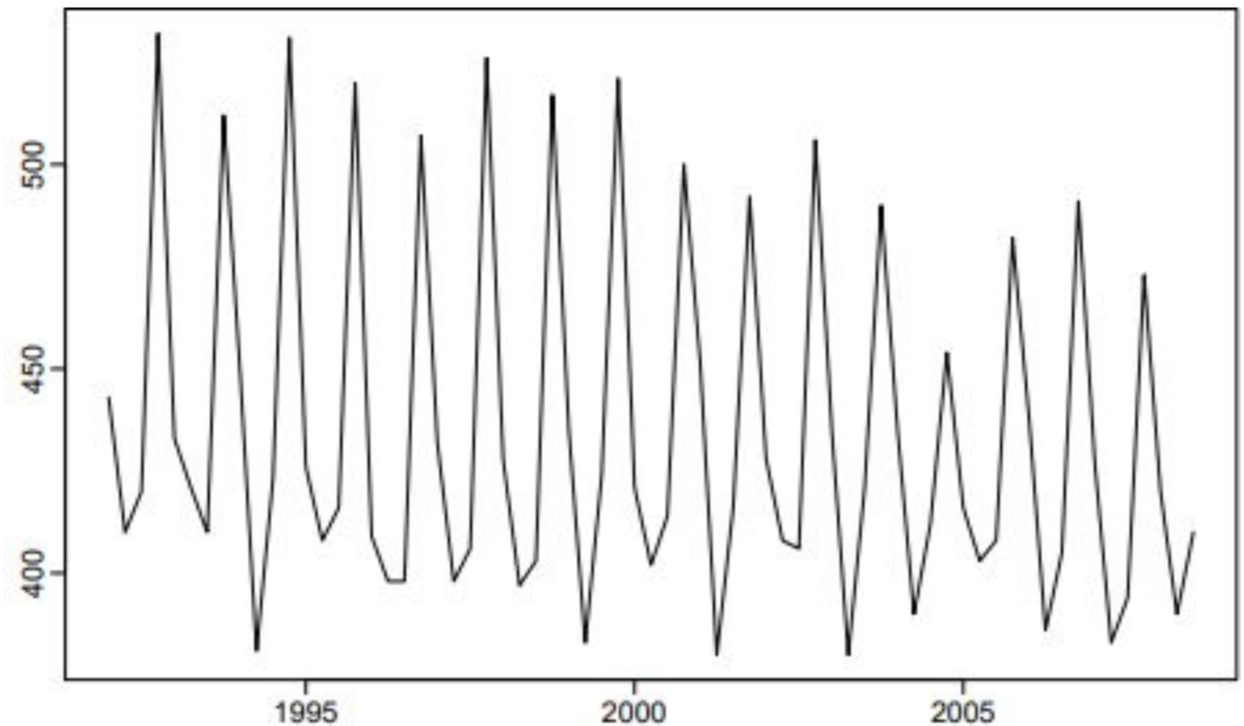
¿Estacionaria?



Estacionariedad

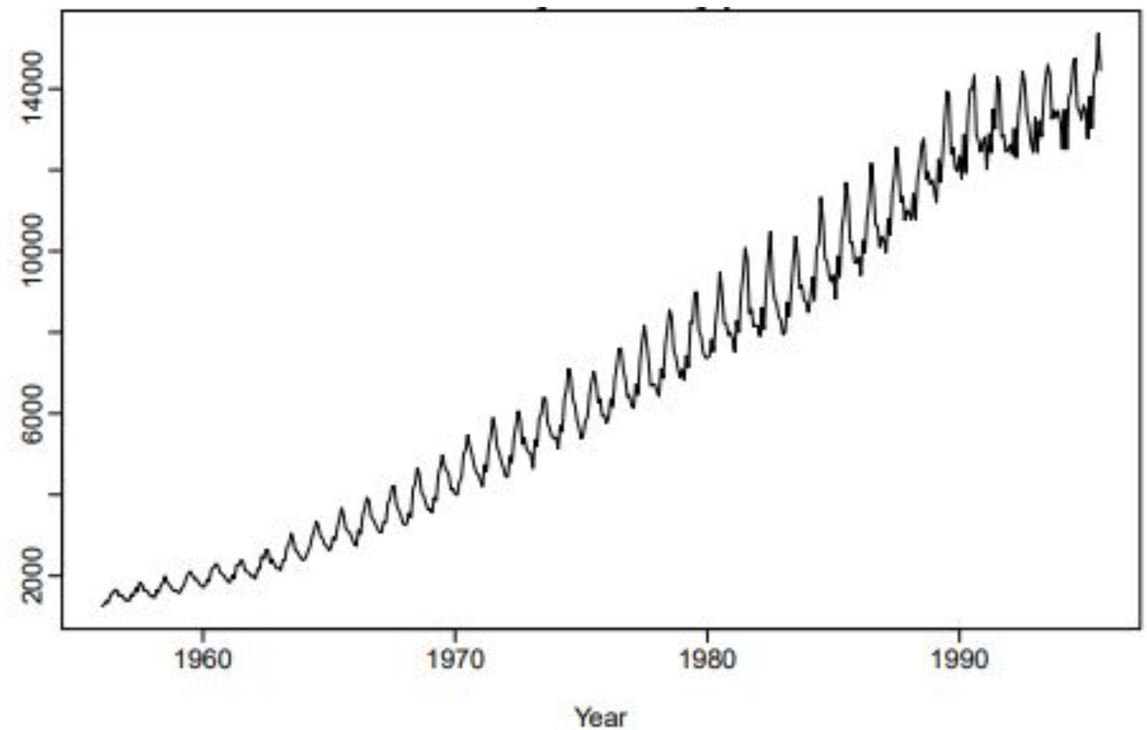
DigitalHouse>
Coding School

¿Estacionaria?



Estacionariedad

¿Estacionaria?



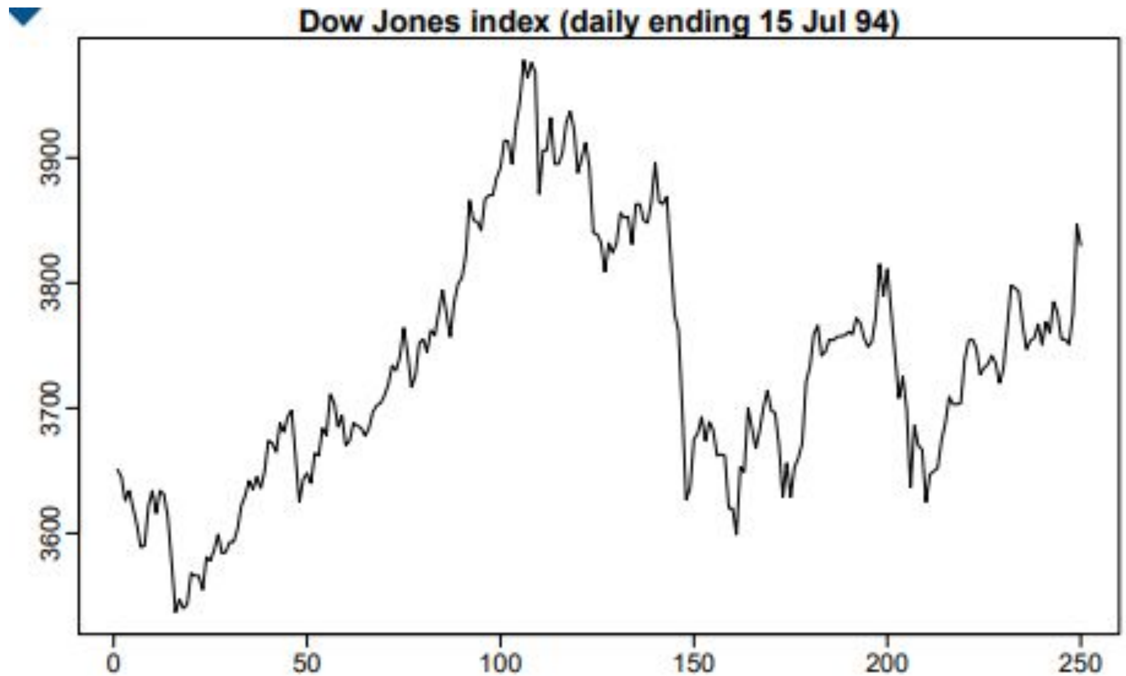
Estacionariedad

- Las transformaciones ayudan a estabilizar las variaciones
- Para el modelado ARIMA, también necesitamos estabilizar la media.

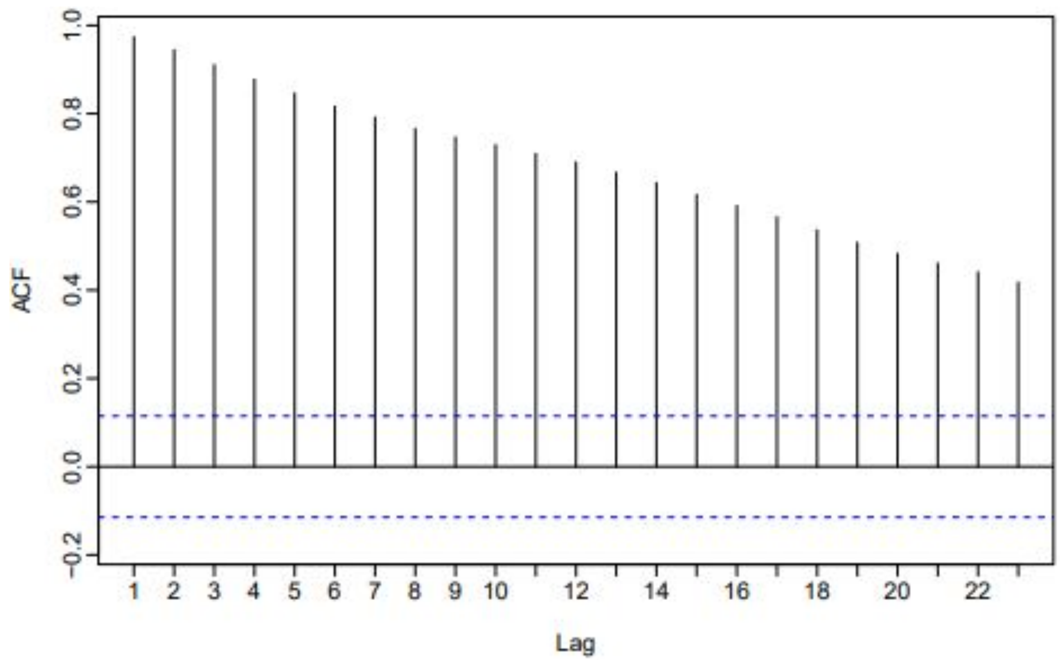
No-estacionariedad en la media

- Idea. Empecemos con un plot.
- El ACF de datos estacionarios cae a cero relativamente rápido
- El ACF de datos no estacionarios disminuye lentamente
- Para datos no estacionarios, el valor de r1 a menudo es grande y positivo

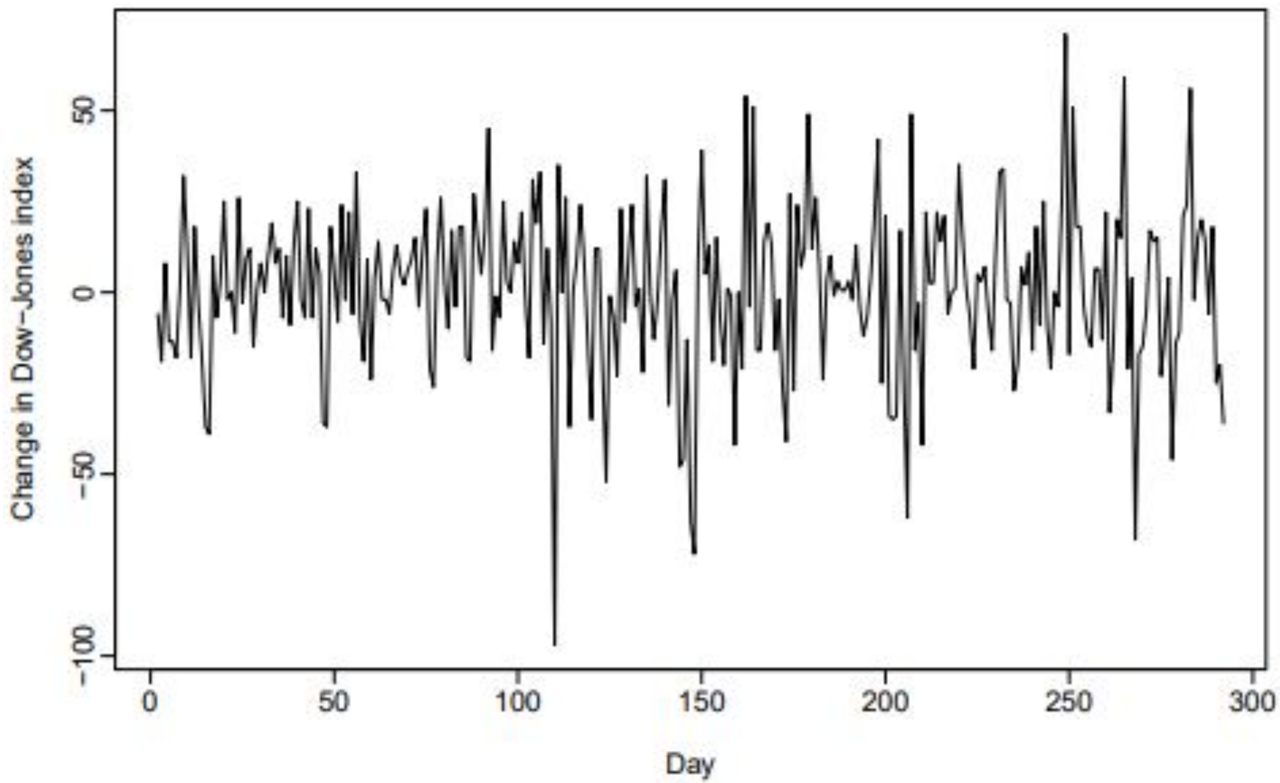
No-estacionariedad en la media



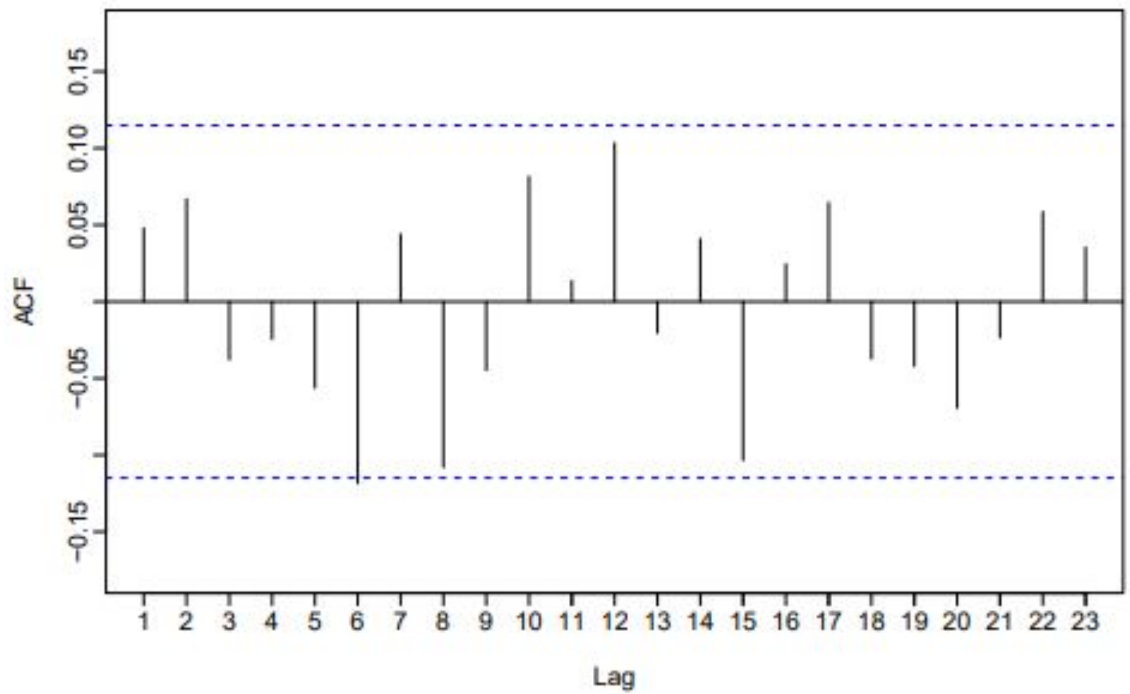
No-estacionariedad en la media



No-estacionariedad en la media



No-estacionariedad en la media



Differencing

- La diferenciación ayuda a estabilizar la media
- La diferencia de la serie es el cambio entre cada observación en la serie original $y'_t = y_t - y_{t-1}$

Differencing

- Las diferencias del índice Dow-Jones son los cambios de día a día.
- Ahora la serie se parece a una serie de ruido blanco
  - No hay autocorrelaciones fuera de los límites del 95%.
  - El estadístico de Ljung-Box tiene un valor de p 0.153 para h = 10.
- Conclusión: El cambio diario en el índice Dow-Jones es esencialmente un cantidad aleatoria no correlacionada con días previos
-

Differencing

- El gráfico de la serie en diferencias sugiere un modelo para el Índice Dow Jones

$$y_t = y_{t-1} + e_t$$

- Modelo de "paseo aleatorio" muy utilizado para datos no estacionarios.
- Los paseos aleatorios suelen tener:
    - Largos períodos de aparentes tendencias hacia arriba o hacia abajo.
    - Cambios repentinos e impredecibles en la dirección.

Differencing

DigitalHouse> Coding School

- Random Walk con drift

$$y_t = c + y_{t-1} + e_t$$

- c es el cambio promedio entre observaciones consecutivas.
- Este es el modelo detrás de drift method

Second order Differencing

- Ocasionalmente los datos diferenciados no parecen estacionarios y puede ser necesario diferenciar los datos por segunda vez:

$$\begin{aligned} y''_t &= y'_t - y'_{t-1} \\ &= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\ &= y_t - 2y_{t-1} + y_{t-2}. \end{aligned}$$

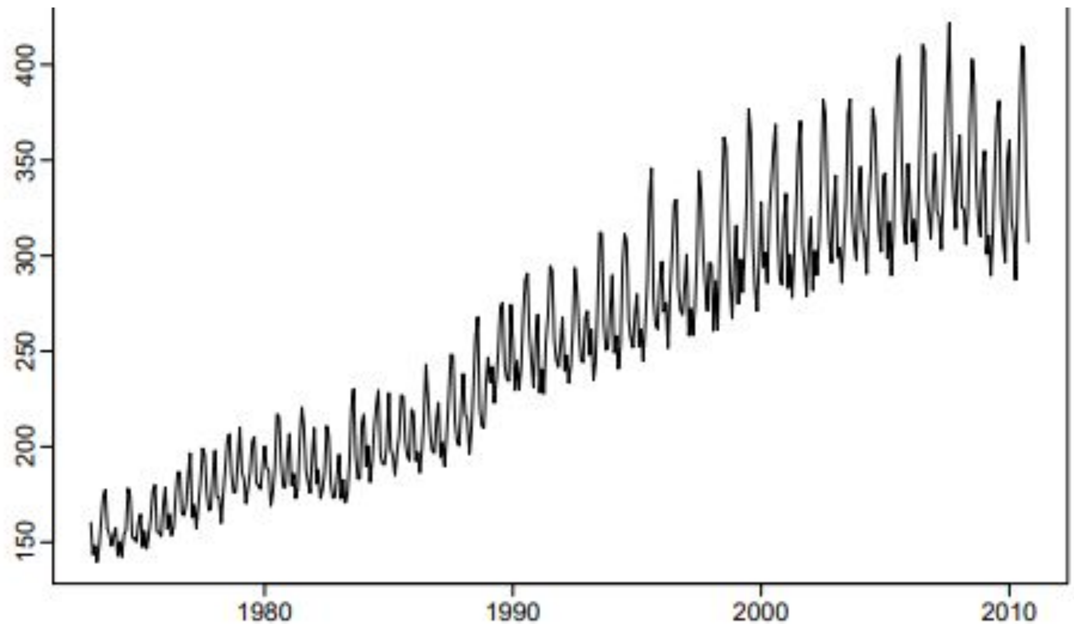
Seasonal Differencing

- Una diferencia estacional es la diferencia. entre una observación y la observación correspondiente del año anterior

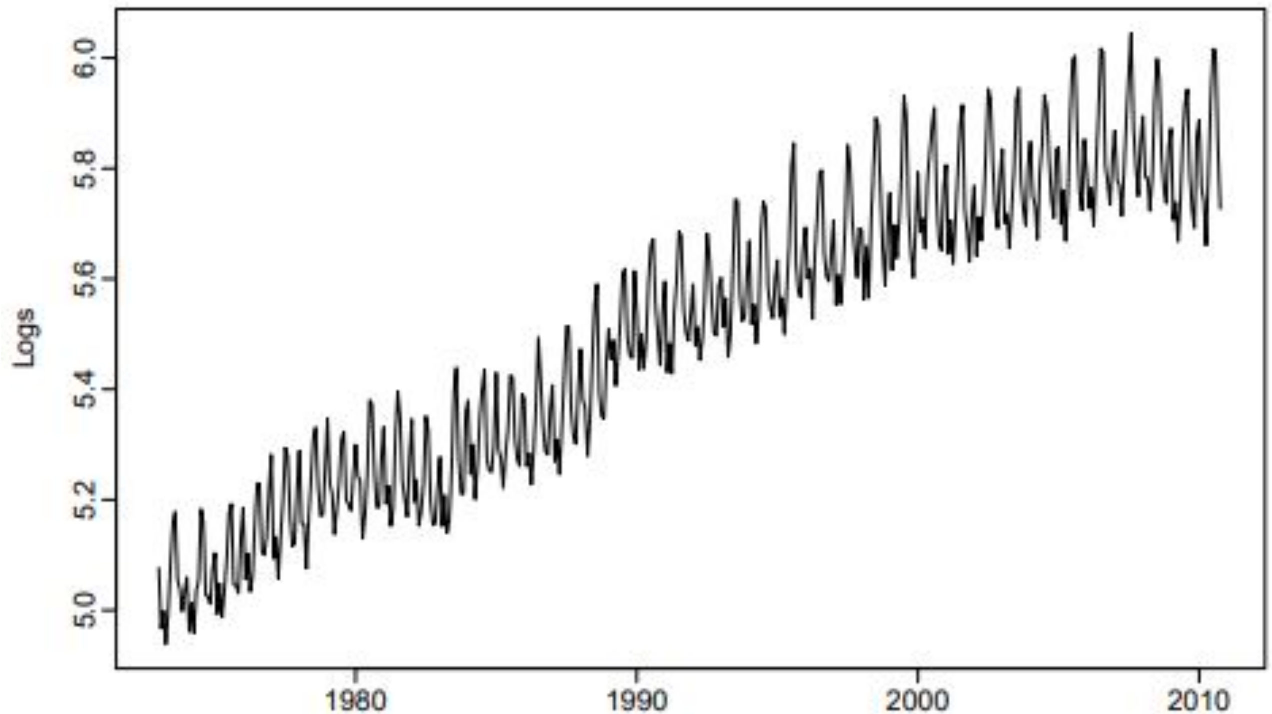
$$y'_t = y_t - y_{t-m}$$

donde m es el número de estaciones.

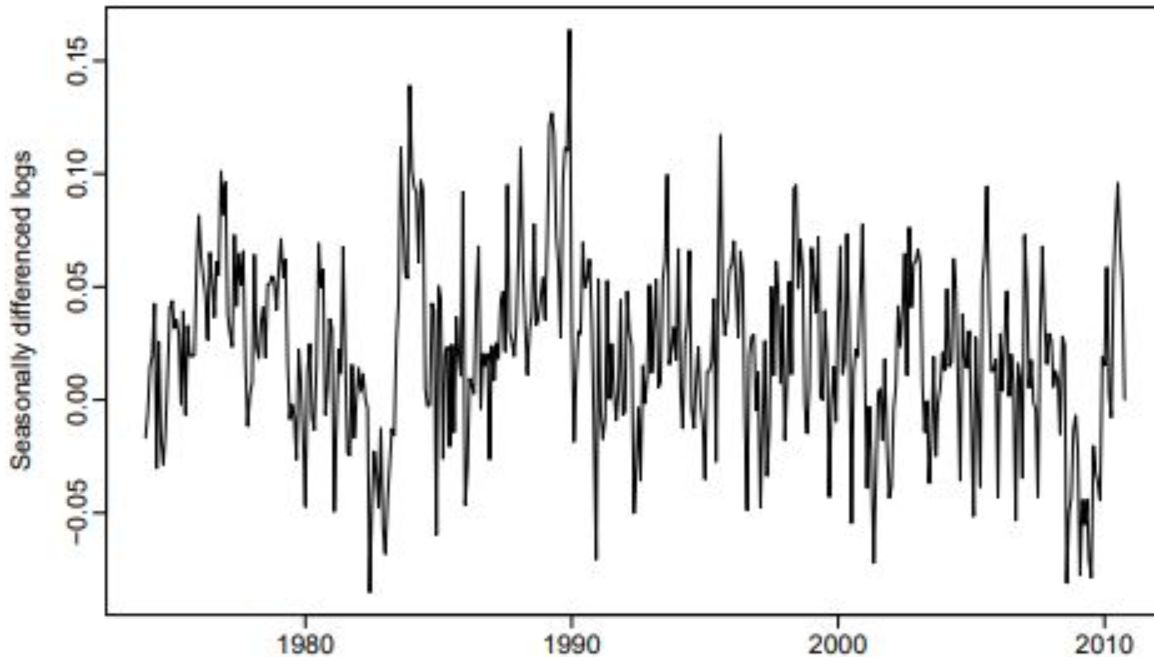
Seasonal Differencing



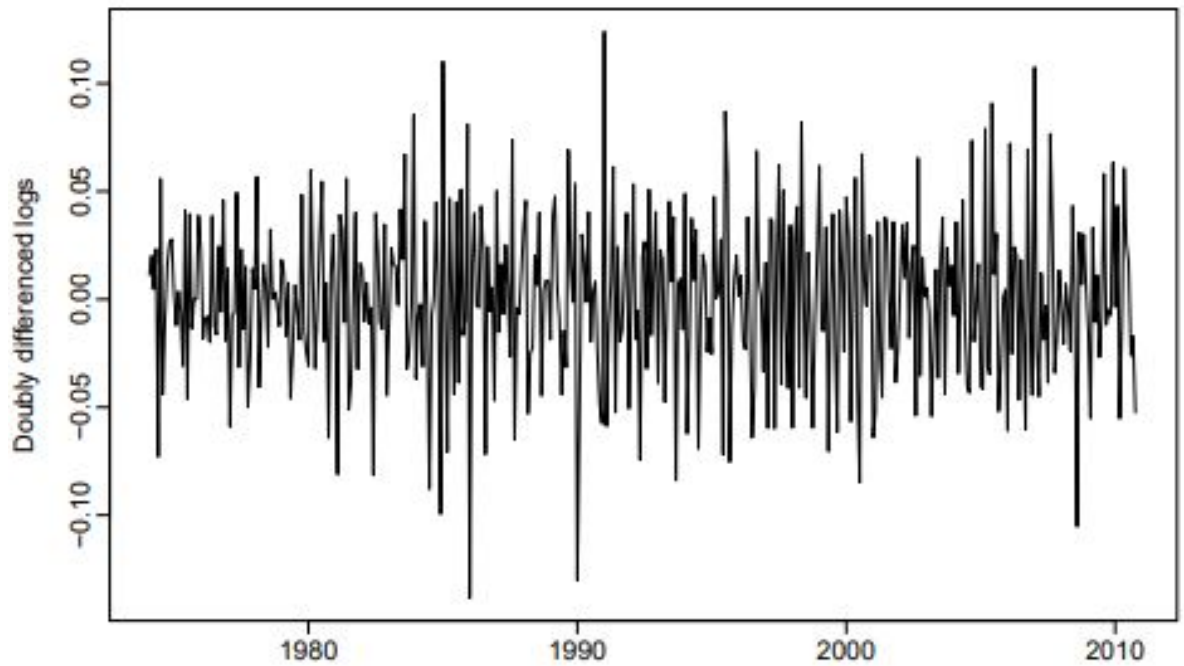
Seasonal Differencing



Seasonal Differencing



Seasonal Differencing



Seasonal Differencing

- La serie estacionalmente diferenciada está más cerca de ser estacionario.
- La no estacionariedad restante puede ser eliminado con una primera diferencia adicional

$$y'_t = y_t - y_{t-12}$$

$$\begin{aligned} y^*_t &= y'_t - y'_{t-1} \\ &= (y_t - y_{t-12}) - (y_{t-1} - y_{t-13}) \\ &= y_t - y_{t-1} - y_{t-12} + y_{t-13} \,. \end{aligned}$$

Seasonal Differencing

- Al aplicar diferencias estacionales como primera diferencia no hace ninguna diferencia el resultado en que se aplique
- Si la estacionalidad estacional es fuerte se recomienda hacer primero la diferenciación estacional porque a veces la serie resultante será estacionaria y no habrá necesidad de seguir diferenciado.

- Es importante que las diferencias sean interpretables.

Unit Root Tests

- Test estadísticas para determinar el orden diferenciación requerido
- Test de Dickey Fuller aumentada: la hipótesis nula es que los datos son no estacionario y no estacional.
- 2 Kwiatkowski-Phillips-Schmidt-Shin:hipótesis nula es que la los datos son estacionarios y no estacionales.
- Otras tests disponibles para datos estacionales

Unit Root Tests

- Estimar un modelo de regresión en las diferencias de la serie

$$y'_t = \phi y_{t-1} + b_1 y'_{t-1} + b_2 y'_{t-2} + \cdots + b_k y'_{t-k}$$

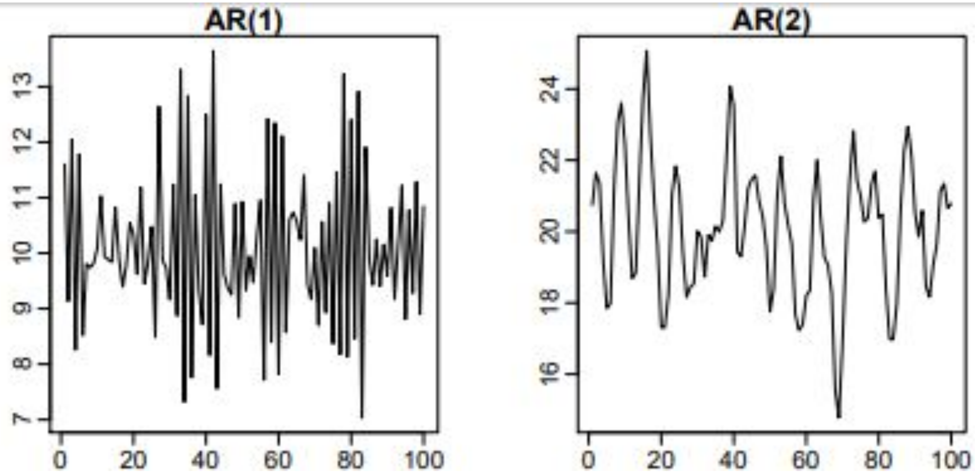
- Si la serie original necesita diferenciación entonces $\hat{\phi} \approx 0$

- Si la serie original es estacionaria entonces $\hat{\phi} < 0$

Modelos auto regresivos (AR)

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t$$

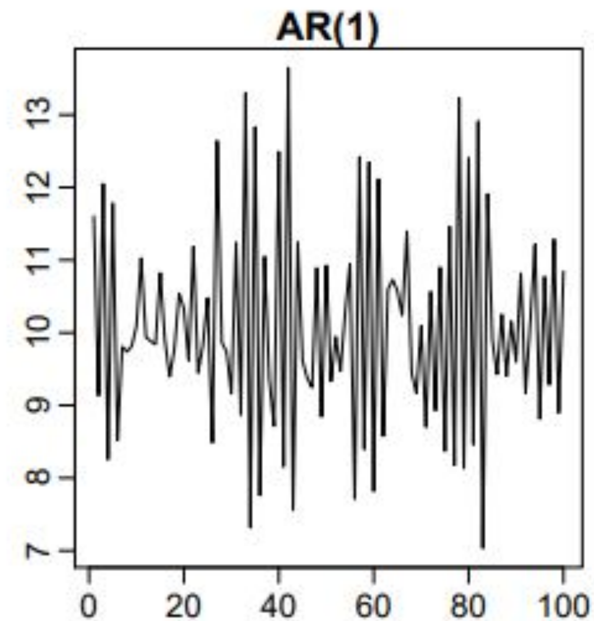
Donde et es ruido blanco



Modelos auto regresivos (AR)

$$y_t = 2 - 0.8y_{t-1} + e_t$$

$e_t \sim N(0, 1)$
$T = 100.$



Modelos auto regresivos (AR)

AR(1)

$$y_t = c + \phi_1 y_{t-1} + e_t$$

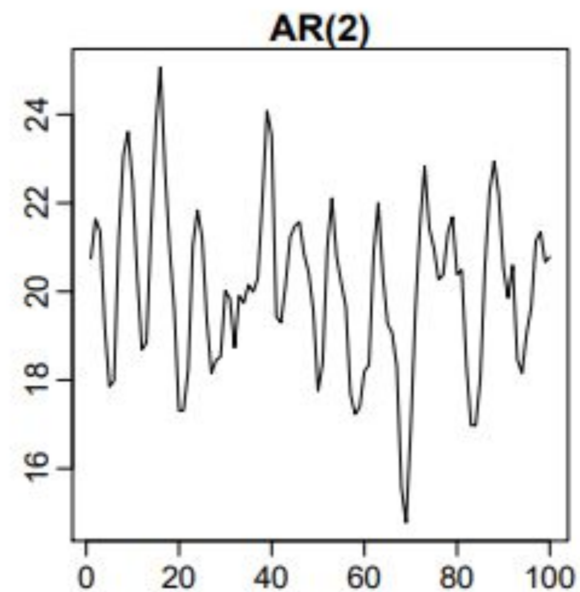
- ¿A qué modelos es equivalente para distintos valores de ϕ_1 ?

Modelos auto regresivos (AR)

AR(2)

$$y_t = 8 + 1.3y_{t-1} - 0.7y_{t-2} + e_t$$

$e_t \sim N(0,1)$
$T = 100.$



Modelos auto regresivos (AR)

DigitalHouse > Coding School

AR(1)

- Normalmente restringimos los modelos autorregresivos a datos estacionarios y luego se requieren algunas restricciones sobre los valores de los parámetros

- Condición general para estacionariedad. Las raíces complejas de $1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p$ están fuera del círculo unitario.

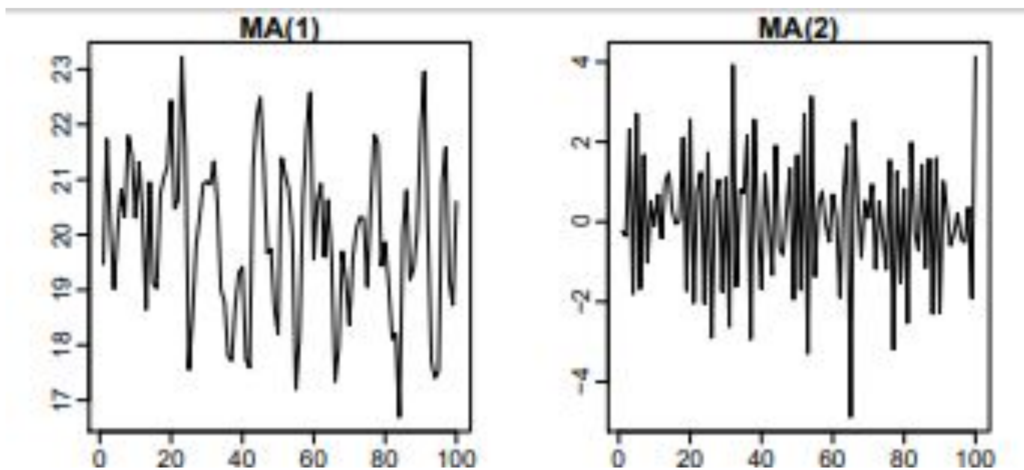
- $p = 1$: $-1 < \phi_1 < 1$.

  $p = 2$ $\quad -1 < \phi_2 < 1 \quad \phi_2 + \phi_1 < 1 \quad \phi_2 - \phi_1 < 1$

Moving Average (MA) Models

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q}$$

donde et es ruido blanco.

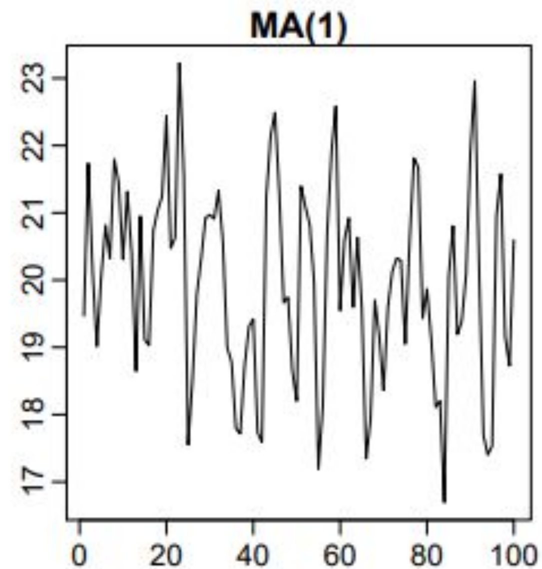


Moving Average (MA) Models

MA(1)

$$y_t = 20 + e_t + 0.8e_{t-1}$$

$e_t \sim N(0, 1)$
$T = 100.$

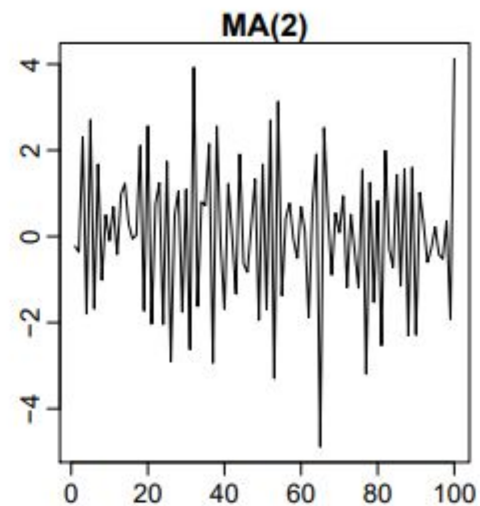


Moving Average (MA) Models

MA(2)

$$y_t = e_t - e_{t-1} + 0.8e_{t-2}$$

$e_t \sim N(0, 1)$
$T = 100.$



Invertibilidad

- Cualquier proceso MA (q) se puede escribir como un proceso AR (∞) si imponemos algunas restricciones en los parámetros de MA.
- Entonces el modelo MA se llama "invertible".
- Condición general de invertibilidad. Las raíces complejas de

$1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q$ están fuera del círculo unitario.

$$q = 1:\ -1 < \theta_1 < 1$$

$$q = 2 \quad -1 < \theta_2 < 1 \quad \theta_2 + \theta_1 > -1 \quad \theta_1 - \theta_2 < 1$$

Auto regresive moving average models (ARIMA)

ARMA

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t$$

ARIMA

Autoregressive Integrated Moving Average models

- Combinar ARMA con diferenciación
- La serie diferenciada sigue un ARMA

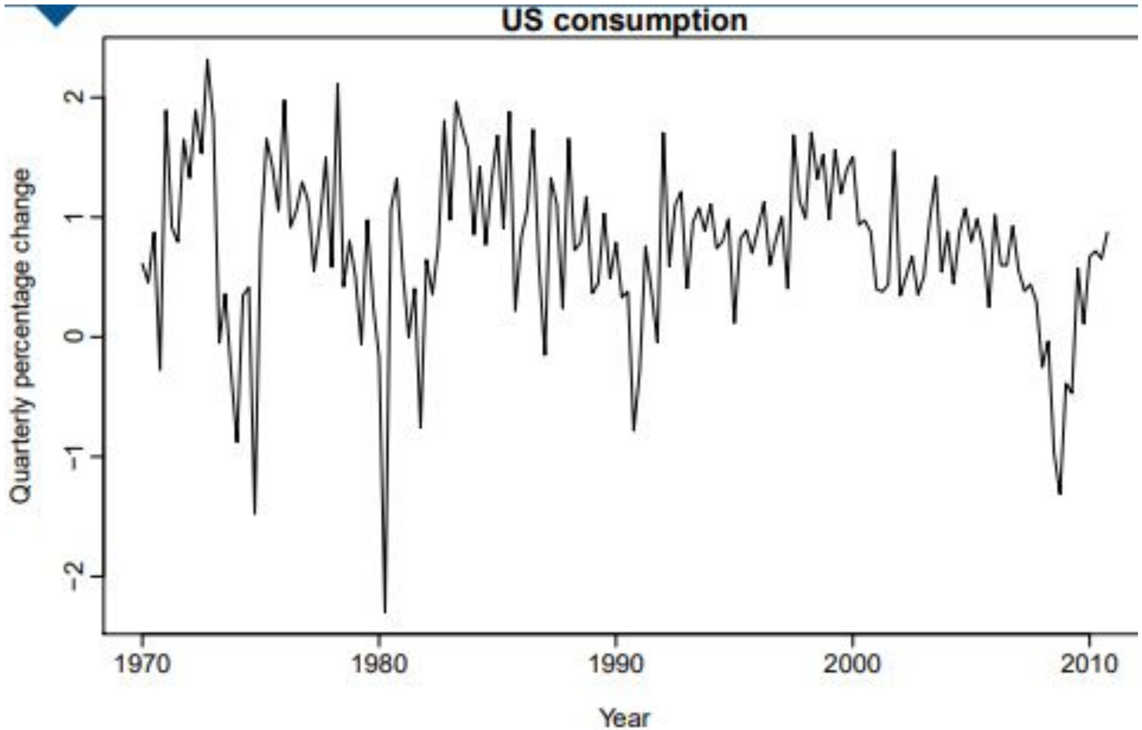
Auto regresive moving average models (ARIMA)

ARIMA(p, d, q) model

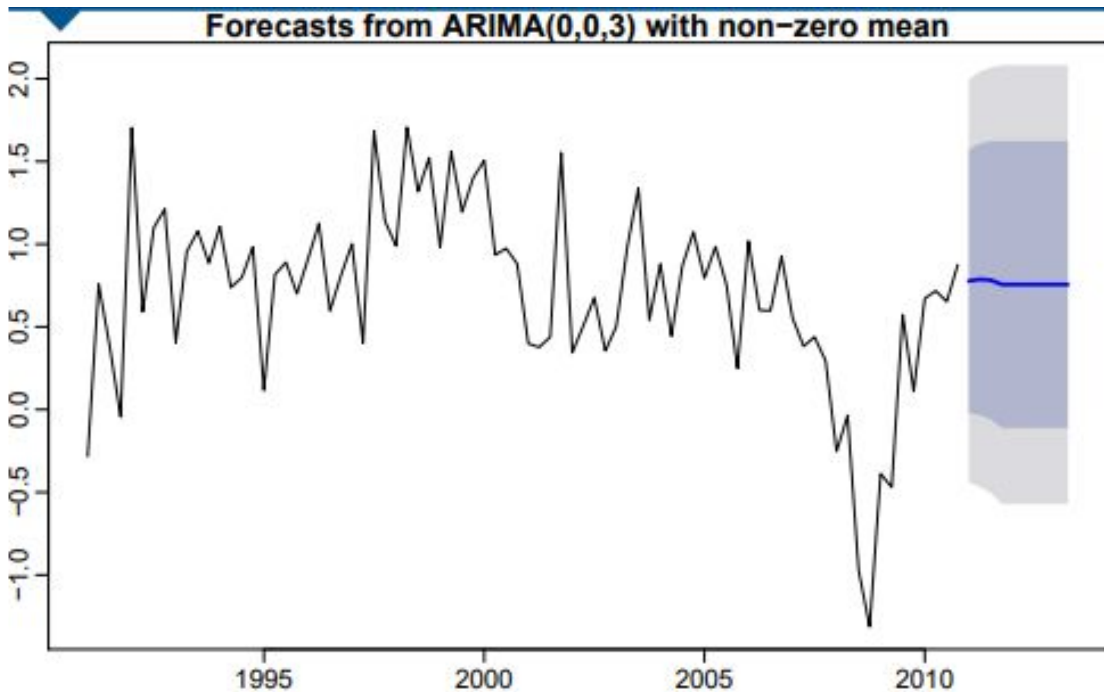
AR:	p = order of the autoregressive part
I:	d = degree of first differencing involved
MA:	q = order of the moving average part.

- White noise model: ARIMA(0,0,0)
- Random walk: ARIMA(0,1,0) with no constant
- Random walk with drift: ARIMA(0,1,0) with const.
- AR(p): ARIMA(p,0,0)
- MA(q): ARIMA(0,0,q)

Auto regresive moving average models (ARIMA)



Auto regresive moving average models (ARIMA)



Auto regresive moving average models (ARIMA)

- Si c = 0 y d = 0, los pronósticos a largo plazo irán hacia el cero
- Si c = 0 y d = 1, los pronósticos a largo plazo irán hacia una constante distinta de cero.
- Si c = 0 y d = 2, los pronósticos a largo plazo siguen una línea recta

Auto regresive moving average models (ARIMA)

- Si c distinto de 0 y d = 0, los pronósticos a largo plazo irán hacia la media de los datos.
- Si c distinto de = 0 y d = 1, los pronósticos a largo plazo serán

  una linea recta

  Si c distinto de 0 y d = 2, los pronósticos a largo plazo serán una tendencia cuadrática.

Auto regresive moving average models (ARIMA)

Forecast variance y d

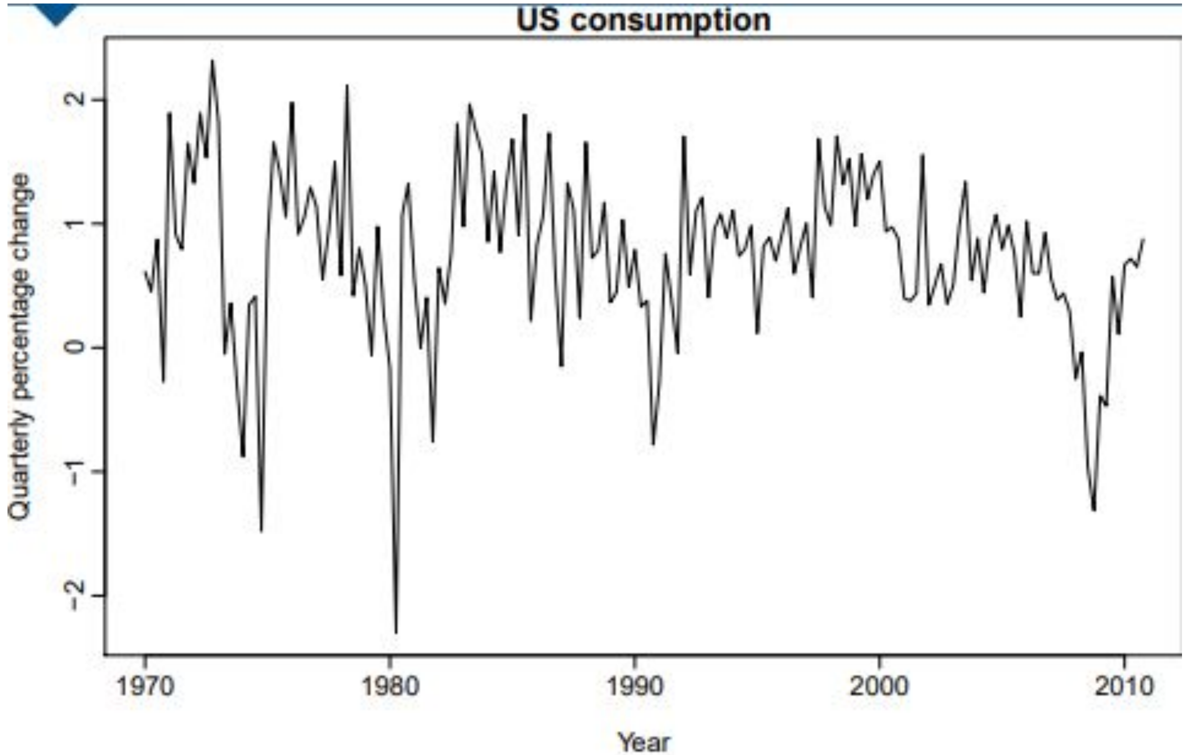
- Cuanto mayor sea el valor de d, más rápidamente los intervalos de predicción aumentan de tamaño.
- Para d = 0, el desvío estándar del pronóstico a largo plazo irá a la desviación estándar de los datos históricos.

Autocorrelación parcial

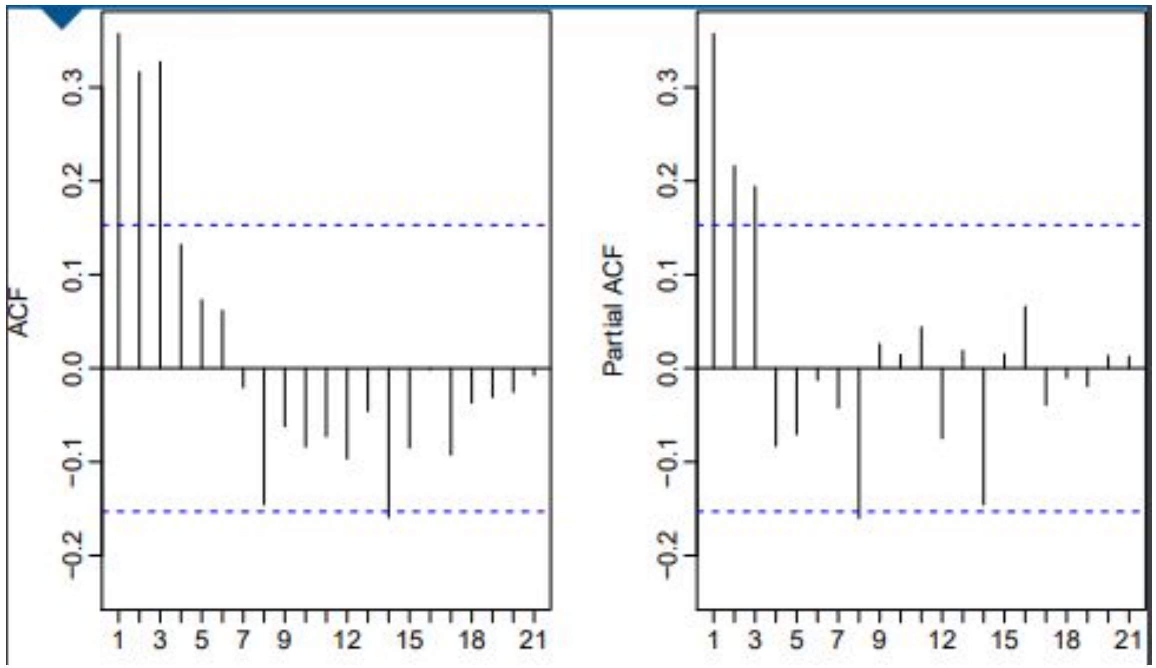
- Las autocorrelaciones parciales miden la relación entre yt e yt–k cuando los efectos de otros rezagos son removidos.
- Son equivalentes a estimar los coeficientes en la siguiente regresión

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_k y_{t-k}$$

Autocorrelación parcial



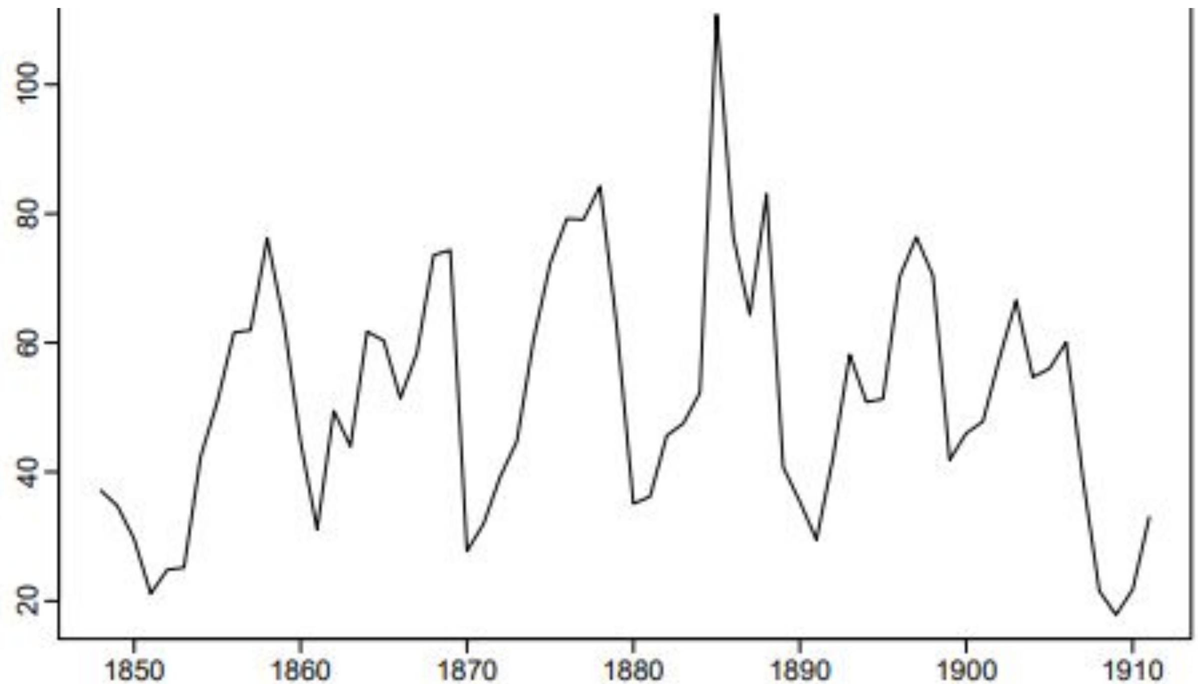
Autocorrelación parcial



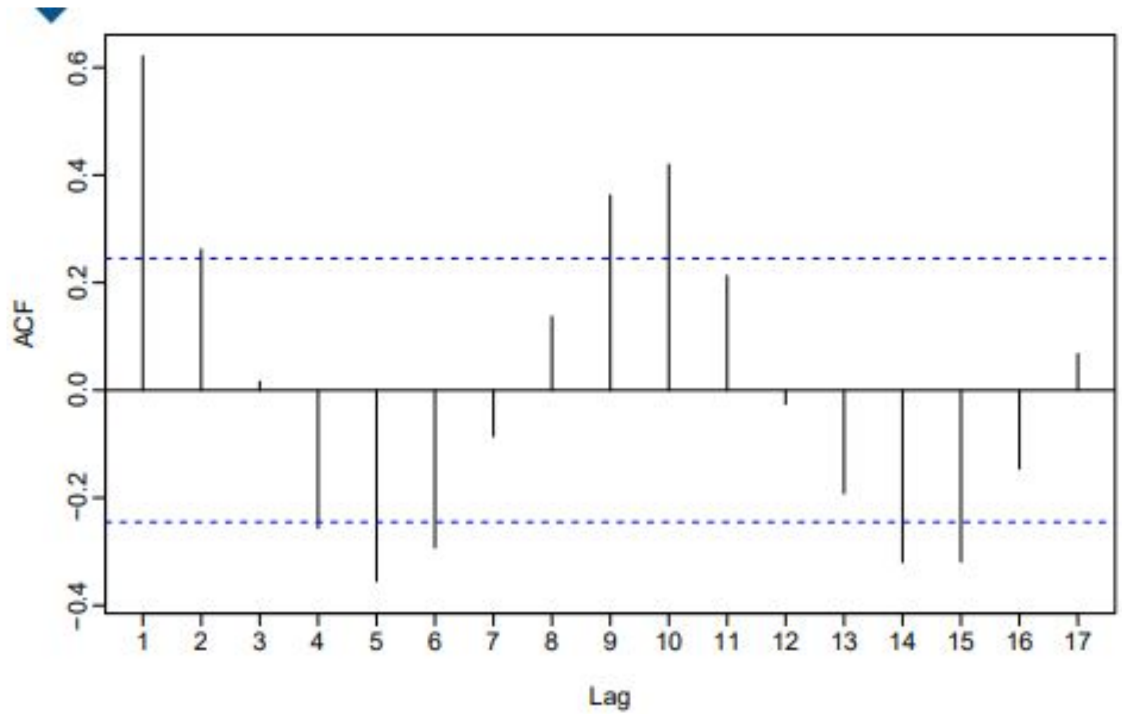
Intepretación ACF y PACF

- Modelo ARIMA (p, d, 0) si las gráficas de ACF y PACF de datos diferenciados muestran:
  - el ACF está en descomposición exponencial o sinusoidal;
  - hay un aumento significativo en el rezago p en PACF, pero ninguno más allá del rezago p.
- Modelo ARIMA (0, d, q) si los gráficos ACF y PACF de datos diferenciados muestran:
  - El PACF está decayendo exponencialmente o sinusoidal;
  - hay un aumento significativo en el rezago q en ACF, pero ninguno más allá del rezago q.

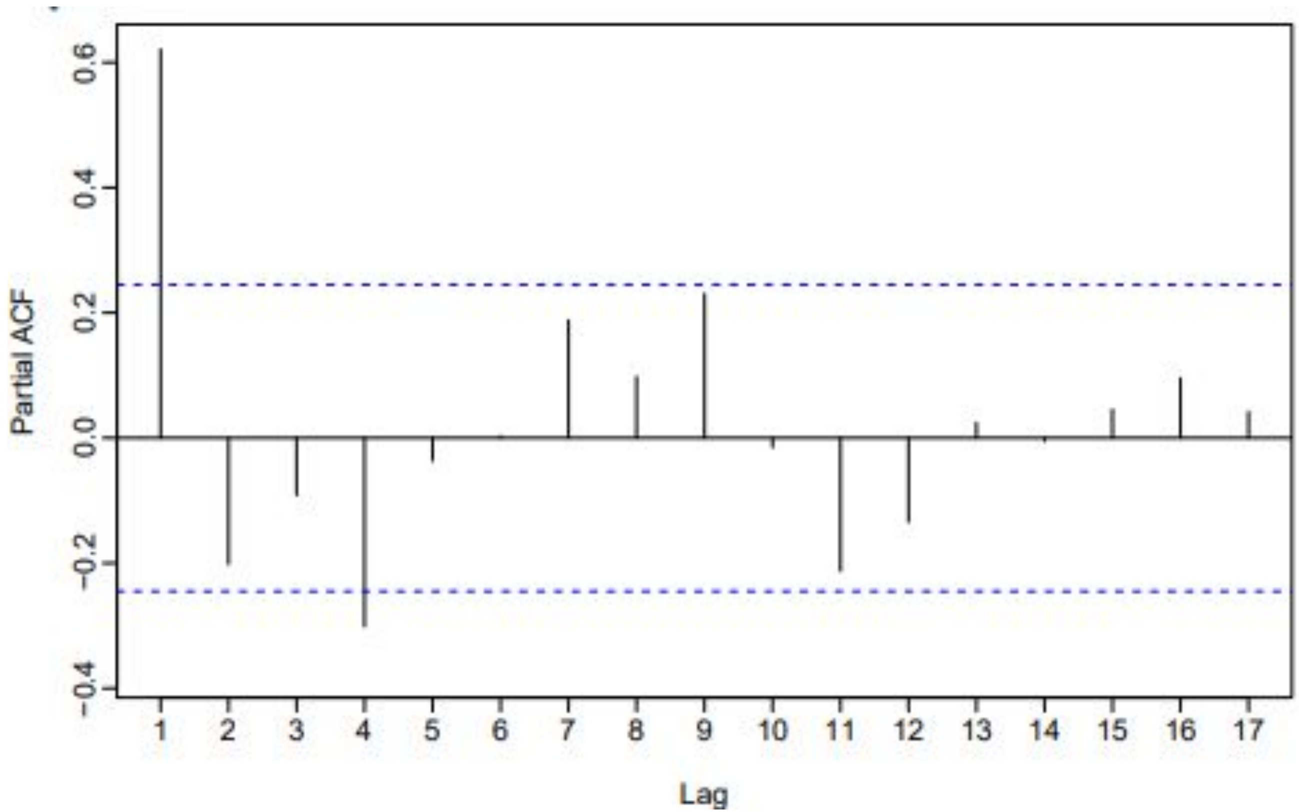
Intepretación ACF y PACF



Intepretación ACF y PACF



Intepretación ACF y PACF



Estimación

- Una vez identificado el orden del modelo necesitamos estimar los parámetros.
- Máxima verosimilitud es muy similar a mínimos cuadrados que busca minimizar

$$\sum_{t-1}^{T} e_t^2$$

Criterio de información

- Akaike

$$AIC = -2\log(L) + 2(p+q+k+1)$$

donde L es la verosimilitud de los datos. k=1 si c distinto de 0 y k=0 si c=0.

- Akaike

$$AIC_c = AIC + \frac{2(p+q+k+1)(p+q+k+2)}{T-p-q-k-2}$$

- Bayesiano

$$BIC = AIC + \log(T)(p+q+k-1)$$

Modelling procedure

- Graficar los datos. Identifique cualquier observación inusual.
- Si es necesario, transforme los datos (usando una transformación de Box-Cox) para estabilizar la varianza.
- Si los datos no son estacionarios: tomar diferencias hasta que los datos hasta los datos son estacionarios.
- Examine el ACF / PACF: ¿Es un modelo AR (p) o MA (q)apropiado?
- Pruebe los modelos elegidos y utilice el AIC para buscar un mejor modelo

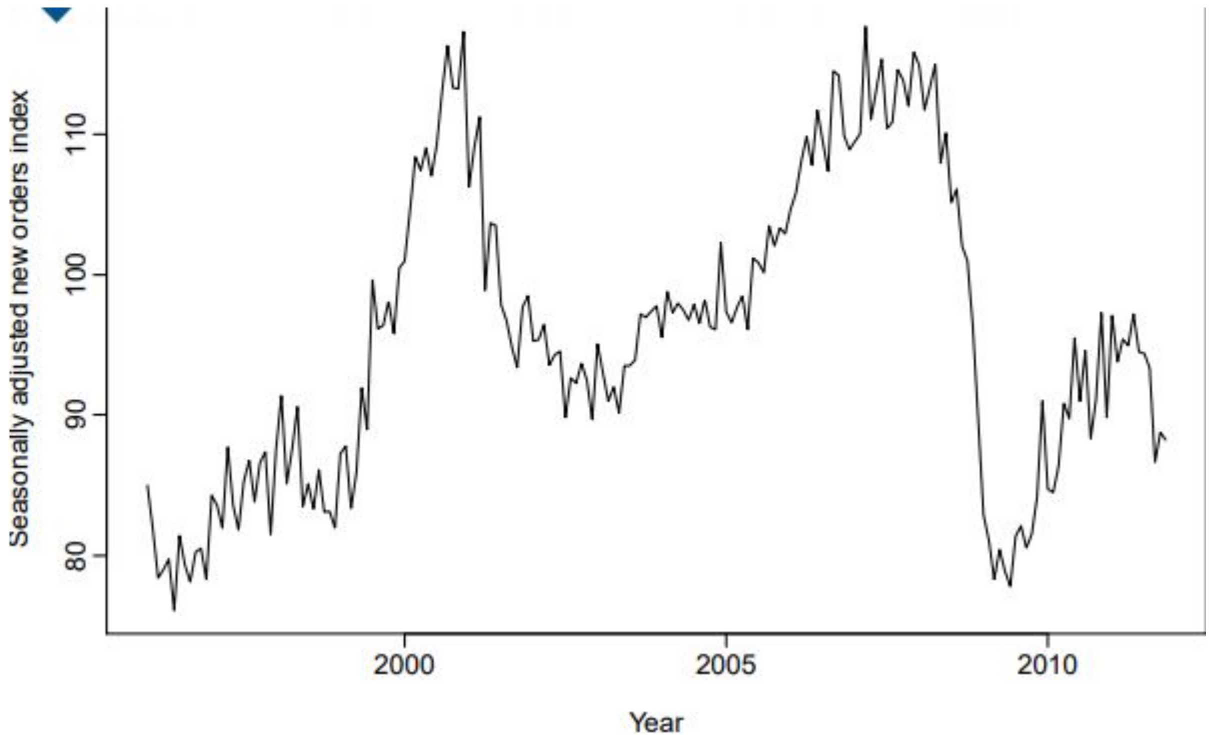
Modelling procedure

- Haga un plot del ACF de los residuos y un test de portmanteau.

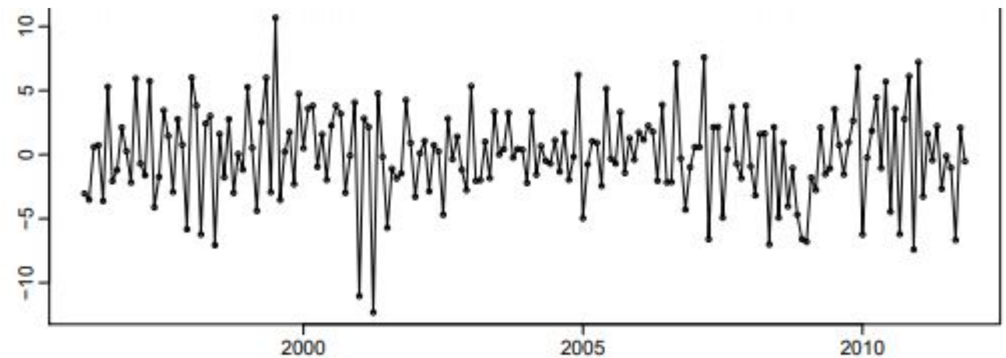
- Si no se ven como ruido blanco, pruebe con un modelo modificado.

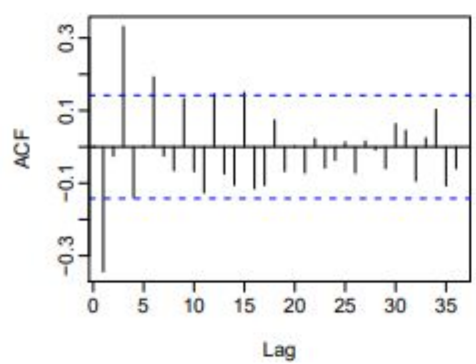
- Una vez que los residuos parezcan ruido blanco, calcule forecasts.

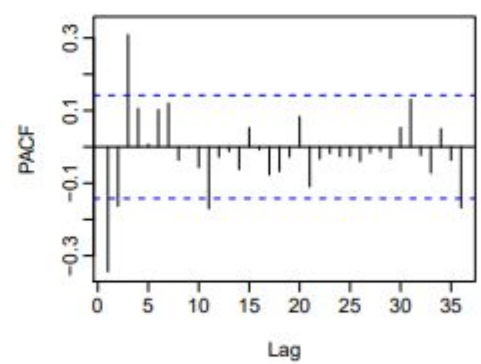
Modelling procedure



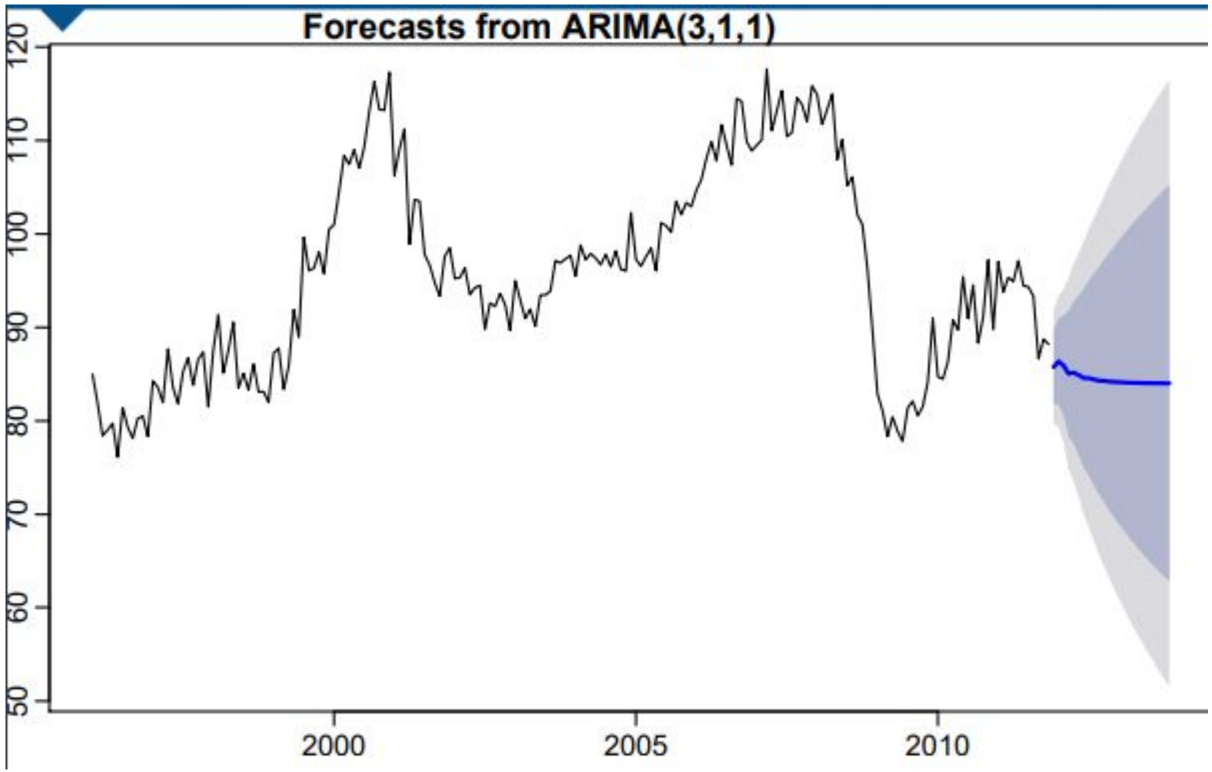
Modelling procedure







Modelling procedure



Point forecasts

DigitalHouse > Coding School

- Reorganizar la ecuación de ARIMA para que yt esté en el lado izquierdo.
- Reescribe la ecuación reemplazando t por T + h.
- En el lado derecho reemplace las observaciones futuras por
- sus forecasts, futuros errores por cero y errores pasados por la correspondiente residuos.

- Comience con h = 1. Repita para h = 2, 3

Forecasts Intervals

95% forecast interval

$$\hat{y}_{T+h|T} \pm 1.96\sqrt{v_{T+h|T}}$$

donde $v_{T+h|T}$ es la varianza estimada del forecast

Prediction Intervals

- Los intervalos de predicción aumentan de tamaño con el horizonte pronosticado
- Los intervalos de predicción pueden ser difíciles de calcular a mano
- Los cálculos suponen que los residuos son no correlacionado y normalmente distribuidos.

Prediction Intervals

- Los intervalos de predicción tienden a ser demasiado estrechos.
    - La incertidumbre en las estimaciones de los parámetros no ha sido tomada en cuenta.
    - El modelo ARIMA asume patrones históricos no cambian durante el periodo de forecast.
    - El modelo ARIMA asume errores no correlacionados en el futuro.

Práctica Guiada

Series de Tiempo