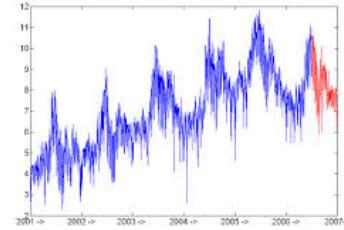


# Series temporales y minería de flujos de datos



# Contenido

- Predicción
- Herramientas básicas
- Regresión
- Descomposición de series temporales
- Alisado exponencial
- Modelos ARIMA
- Modelos de predicción avanzados

# Bibliografía

- C. Chatfield, «The analysis of time series: An Introduction», CRC Press, 2003
- J.D. Hamilton, «Time Series Analysis», Princeton University Press, 1994
- R. Hyndman, G. Athanasopoulos, «Forecasting: principles and practice» 2013,  
[www.otexts.org/fpp](http://www.otexts.org/fpp)
- *Id.* 2nd version, 2018 [www.otexts.org/fpp2](http://www.otexts.org/fpp2)
- P.J. Brockwell, R.A. Davis, «Introduction to Time Series and Forecasting», 2nd ed., Springer, 2002

# Bibliografía (2)

- S.G. Makridakis, S.C. Wheelwright, R.J. Hyndman, «Forecasting», 3rd Ed., Wiley & Sons, 1998
- A.K.Palit, D. Popovic, «Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications», Springer, 2005
- R.H. Shumway, D.S. Stoffer, «Time Series Analysis and Its Applications», Springer, 2nd Ed., 2006

# Software

- Plataforma R
- CRAN Task View: [“Time Series Analysis”](#)
- Paquetes de referencia:
  - forecast
  - ts, xts, zoo
  - fpp

# Predicción

# Contenido

- Predicción
- Herramientas básicas
- Regresión
- Descomposición de series temporales
- Alisado exponencial
- Modelos ARIMA
- Modelos de predicción avanzados

# Predicción

- La predicción ha fascinado a los seres humanos durante miles de años
- Algunas referencias históricas:
  - Los profetas de la Biblia
  - Sacerdotes de Babilonia
  - Los griegos consultando el oráculo de Delfos
  - Adivinos ¿Rappel?



# ¿Qué se puede predecir?

- La predecibilidad de un evento o magnitud depende de varios factores:
  - Cómo de bien entendemos los factores que influyen
  - Cuántos datos están disponibles
  - Si la predicción puede afectar lo que se predice

# Ejemplos

- Predicción de la demanda de energía eléctrica: muy precisa
- Predicción de la tasa de cambio: más o menos
- Combinación de euromillones: no



# Factores que afectan la predicción

- Horizonte temporal
- Tipo de patrones en los datos

# Elección del método

- Depende de la disponibilidad de datos y de la prediccibilidad de la cantidad a predecir
- Cada método tiene sus propias características: precisiones, restricciones, costes

# Definición

- **Predicción:** Averiguar el futuro tan precisamente cómo sea posible, a partir de la información disponible incluyendo datos históricos y conocimiento de cualquier evento futuro que pueda influir en la predicción
- Habitualmente, es parte de un proceso de toma de decisiones

# Clasificación basada en el horizonte

- A corto plazo
- A medio plazo
- A largo plazo

# Determinar qué predecir

- ¿Qué predecir?
  - Una cantidad
  - Un resumen
- Horizonte de predicción
- Frecuencia de predicción

# Predicción cuantitativa

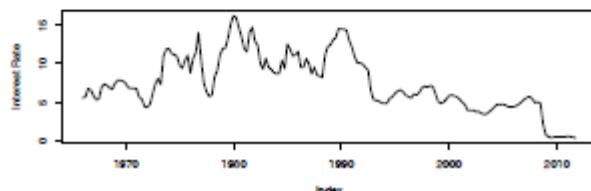
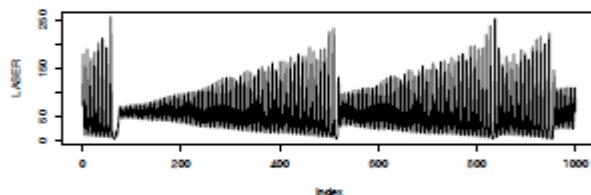
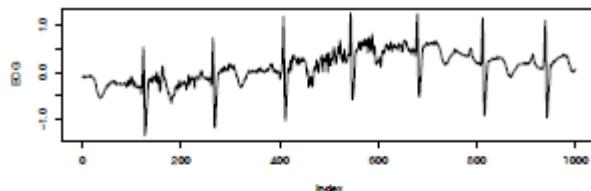
- Se puede aplicar cuando:
  - Hay datos numéricos del pasado disponibles
  - Es razonable asumir que algunos aspectos de patrones que ocurren en el pasado seguirán en el futuro

# Datos *cross-sectional*

- Datos recogidos en un mismo instante
- Predicción de datos *cross-sectional* pretende predecir el valor de algo no observado en base a valores observados (en el mismo periodo temporal):
  - El precio de todas las casas vendidas en 2020 en un área concreta
  - Consumo de combustible para los vehículos de 2030

# Series temporales

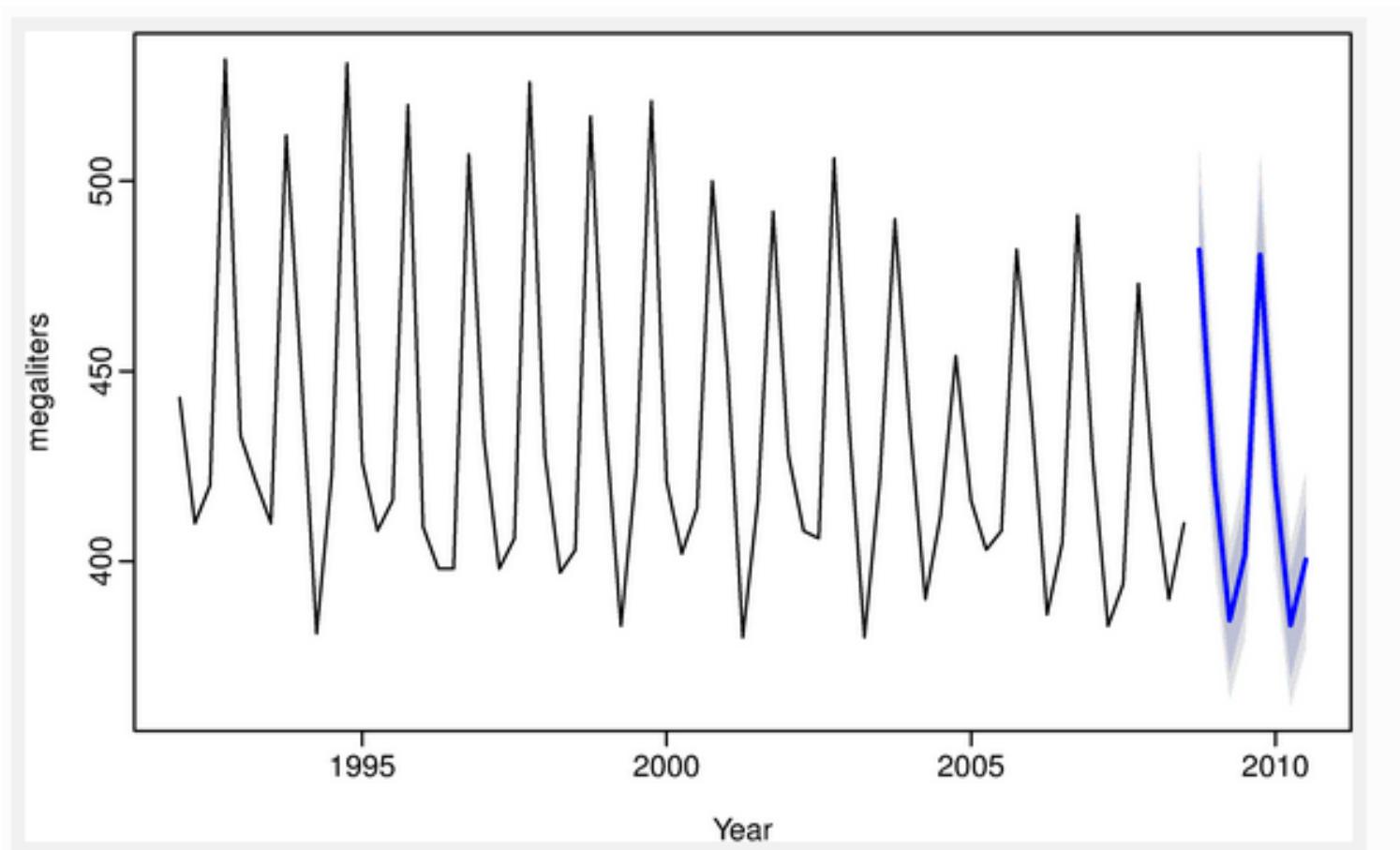
- Cualquier magnitud observada a lo largo del tiempo es una **serie temporal**
- Series temporales observadas en intervalos regulares (cada minuto, hora, día, semana, ...)



# Predicción de series temporales

- Los datos de las series temporales son útiles cuando se está prediciendo algo que cambia con el tiempo (p.ej. precios de acciones, ventas, beneficios, ...)
- La predicción de series temporales pretende calcular cómo continúa la serie temporal en el futuro

# Predicción de producción de cerveza



# Variables predictores

- Predecir la demanda eléctrica (DE)
- $DE = f(\text{temperatura actual, nivel de la economía, población, hora del día, día de la semana, error})$

# Pasos básicos en predicción

1. Definición del problema
2. Reunir datos e información
3. Análisis preliminar (exploratorio)
4. Elección del modelo de ajuste
5. Usar y evaluar los modelos

# Perspectiva estadística de la predicción

- La cantidad a predecir se puede considerar una variable aleatoria
- Cuanto más alejado sea el horizonte de predicción, más incertidumbre hay
- Predicción: aproximar la media del rango de los valores posibles, o en general, la esperanza de la variable aleatoria

# Contenido

- Predicción
- **Herramientas básicas**
- Regresión
- Descomposición de series temporales
- Alisado exponencial
- Modelos ARIMA
- Modelos de predicción avanzados

# Las herramientas del analista

- Gráficos
- Resúmenes numéricos
- Transformaciones
- Evaluación de la precisión del modelo
- Diagnóstico de residuos
- Intervalos de predicción

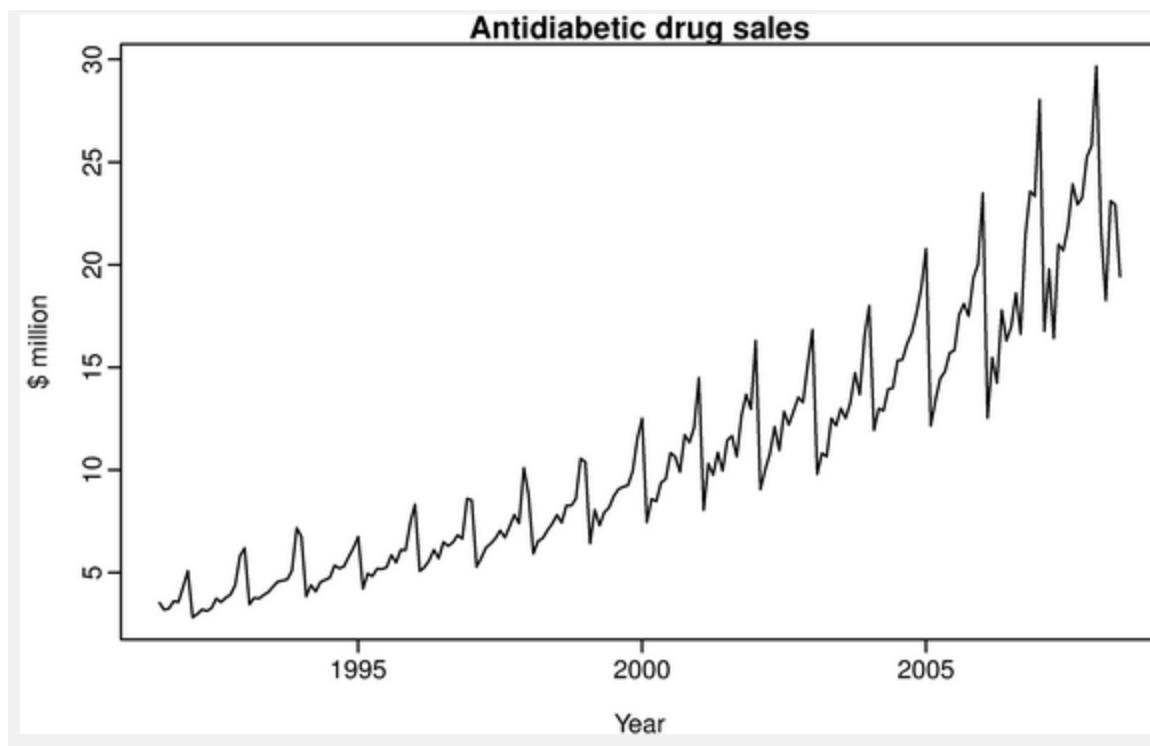
# Gráficos

- Gráfico temporal
- Gráfico de puntos

# Gráfico temporal

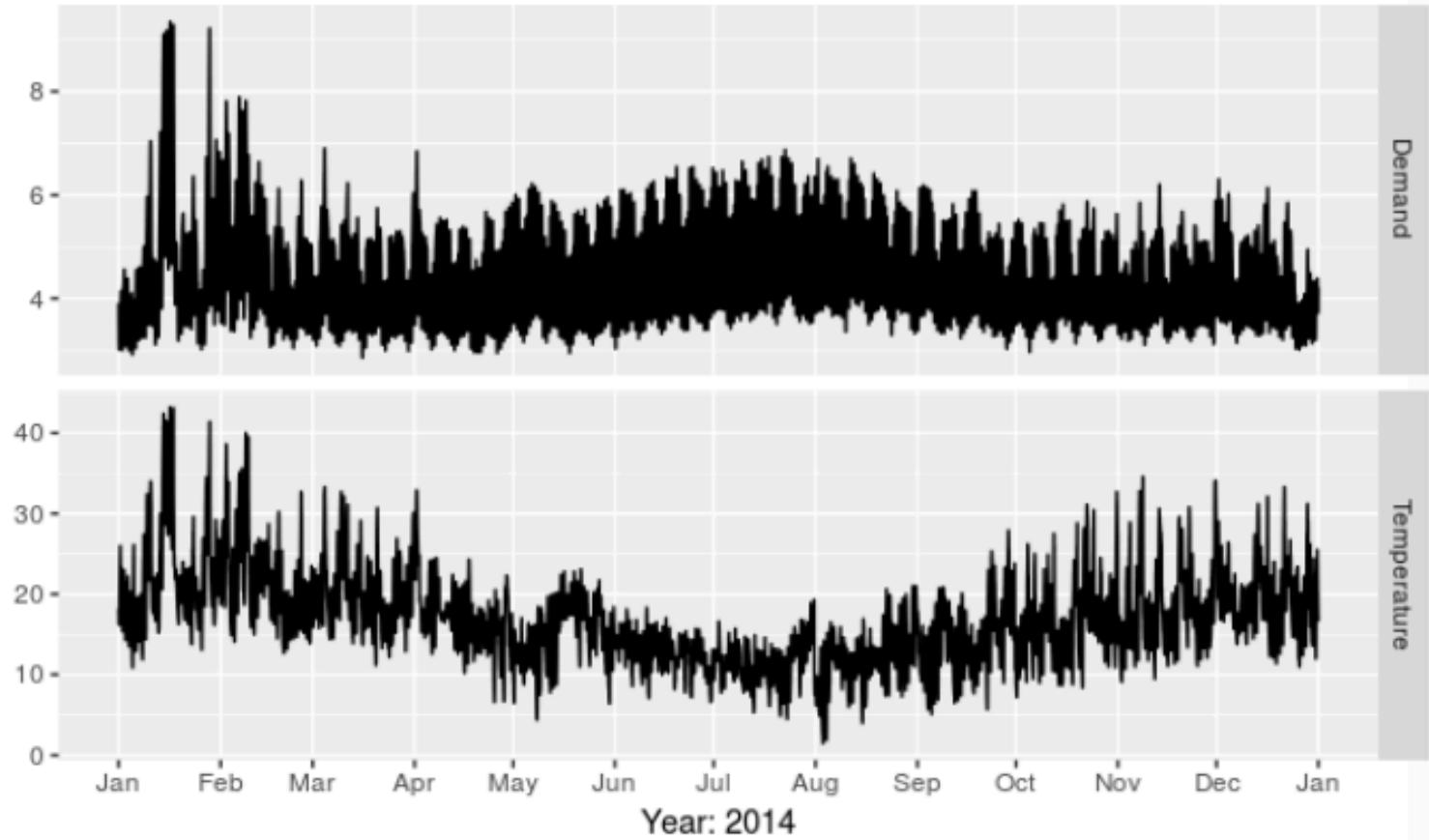


# Gráfico temporal

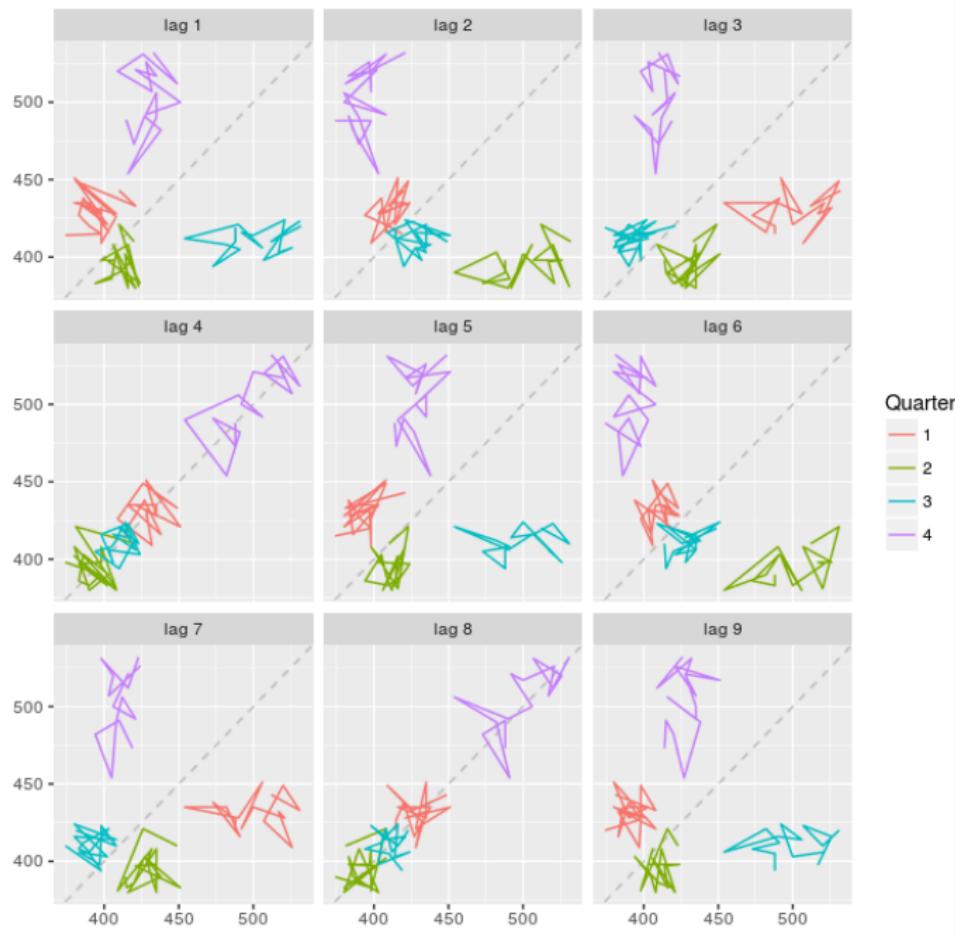


# ScatterPlots

Half-hourly electricity demand: Victoria, Australia



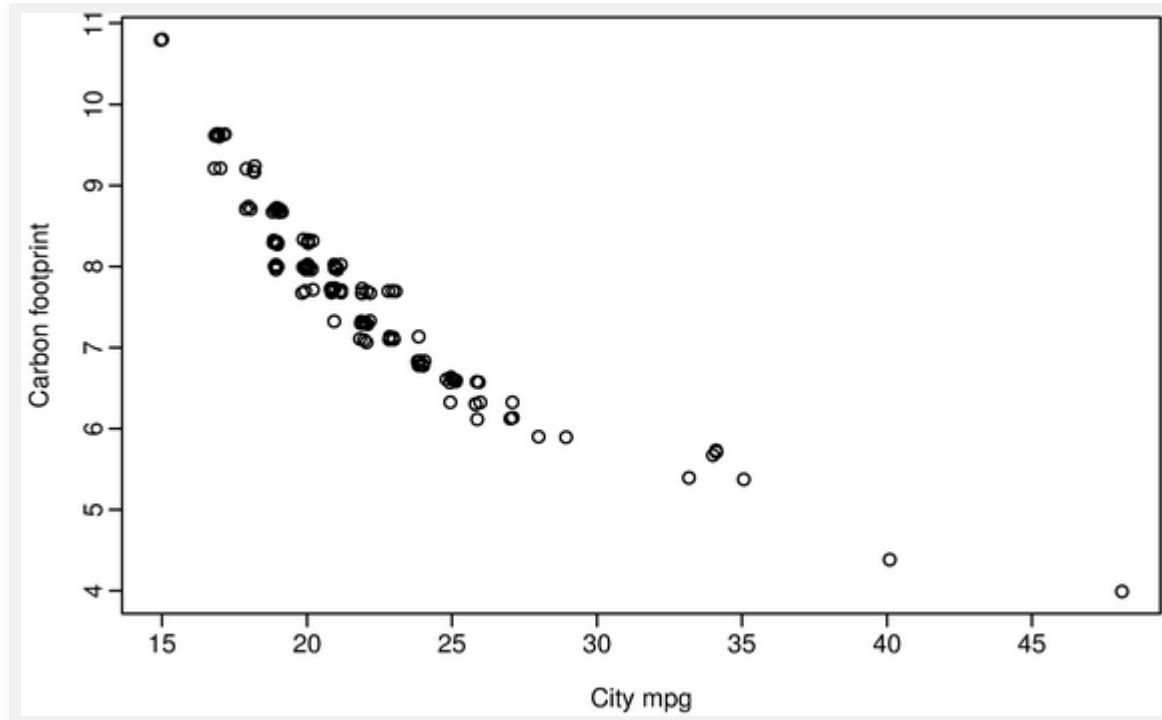
# Gráficos de retardos



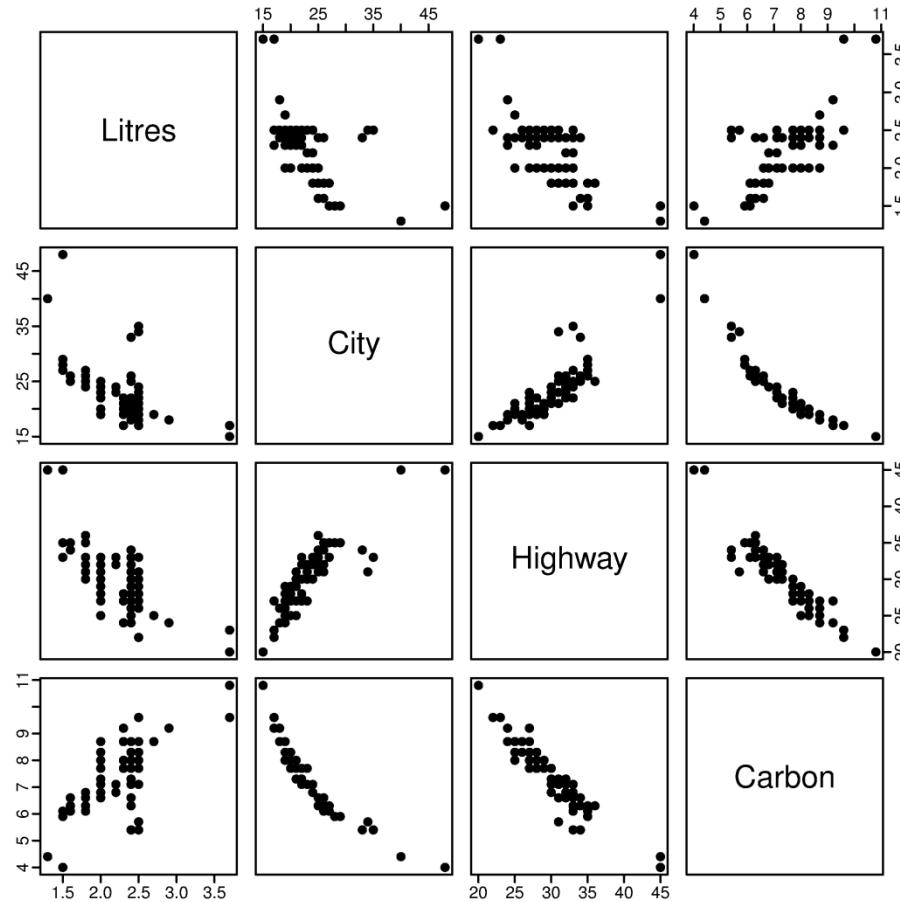
# Patrones en las series temporales

- *Tendencia*: incrementos o decrementos a largo plazo en los datos
- *Patrón estacional*: datos afectados por un patrón estacional tales como el día del año o el día de la semana
- *Ciclo*: los datos muestran subidas o bajadas pero no asociados a un periodo fijo; la longitud suele ser variable y desconocida

# Gráfico de puntos



# Matriz de gráficos de puntos



# Resúmenes numéricos

- Estadísticas univariadas

- Media
- Mediana
- Percentiles
- Rango Intercuartílico (IQR)
- Desviación estándar

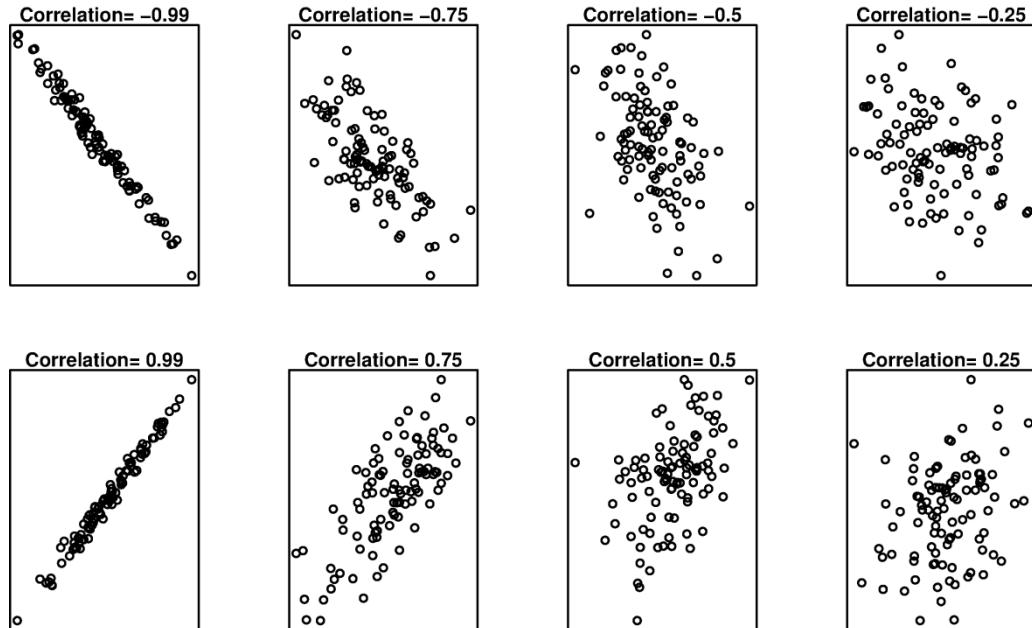
$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = (x_1 + x_2 + x_3 + \cdots + x_N)/N.$$

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}.$$

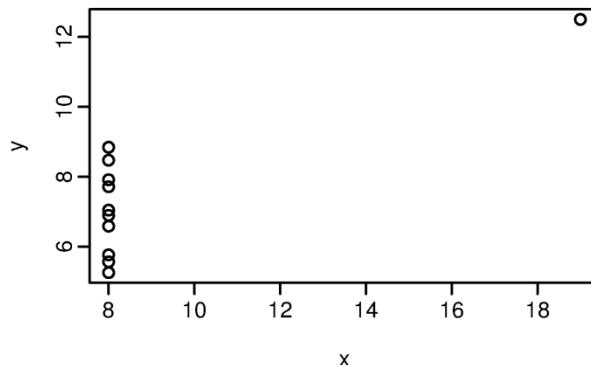
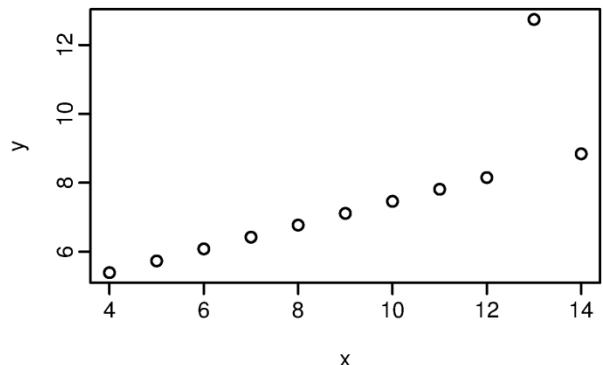
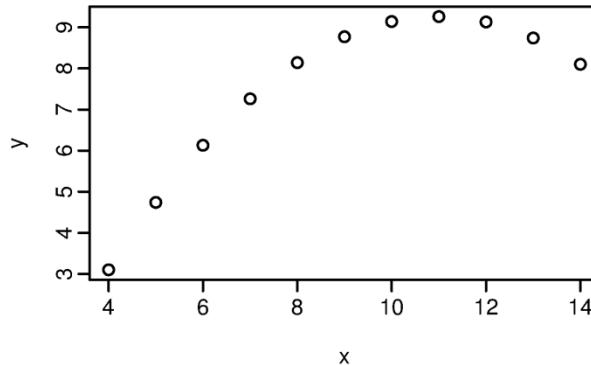
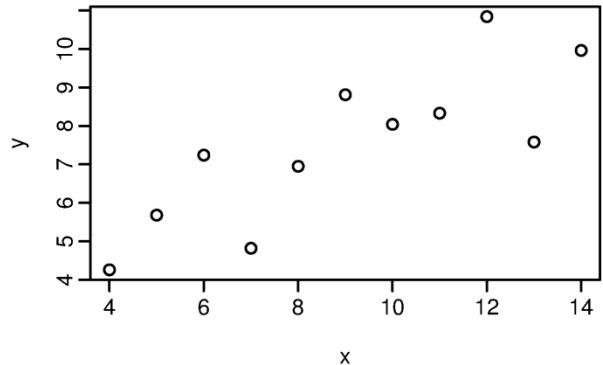
# Coeficiente de correlación

Fuerza de la  
relación lineal

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2} \sqrt{\sum(y_i - \bar{y})^2}},$$



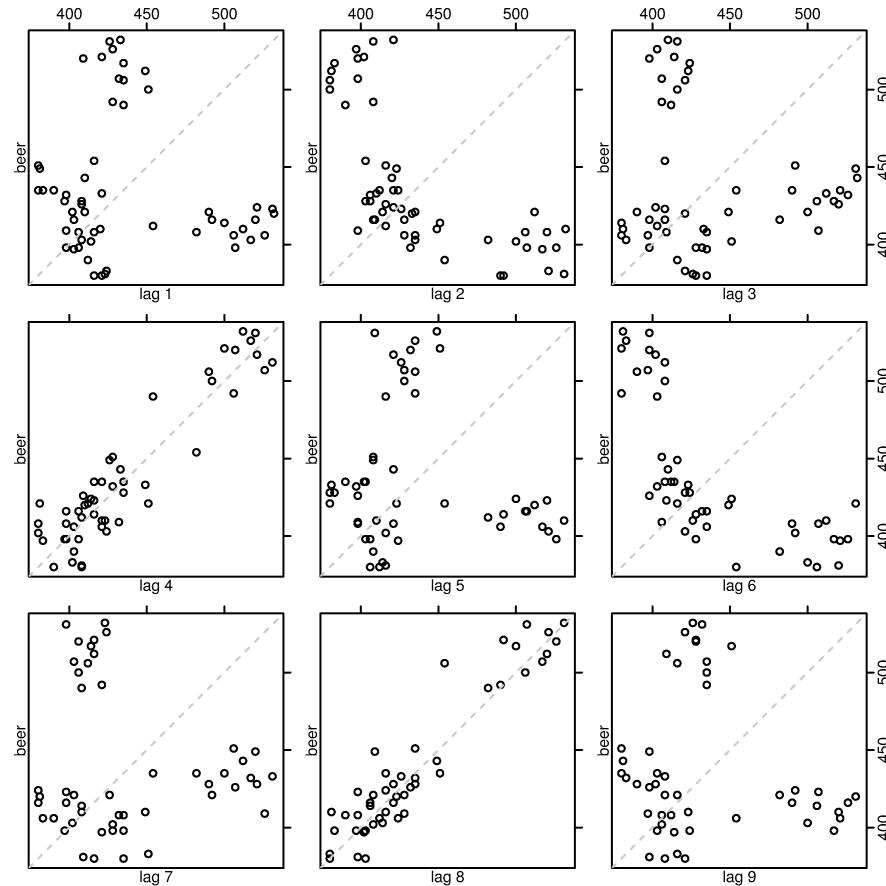
# Coeficiente de correlación 0,82



# Autocorrelación

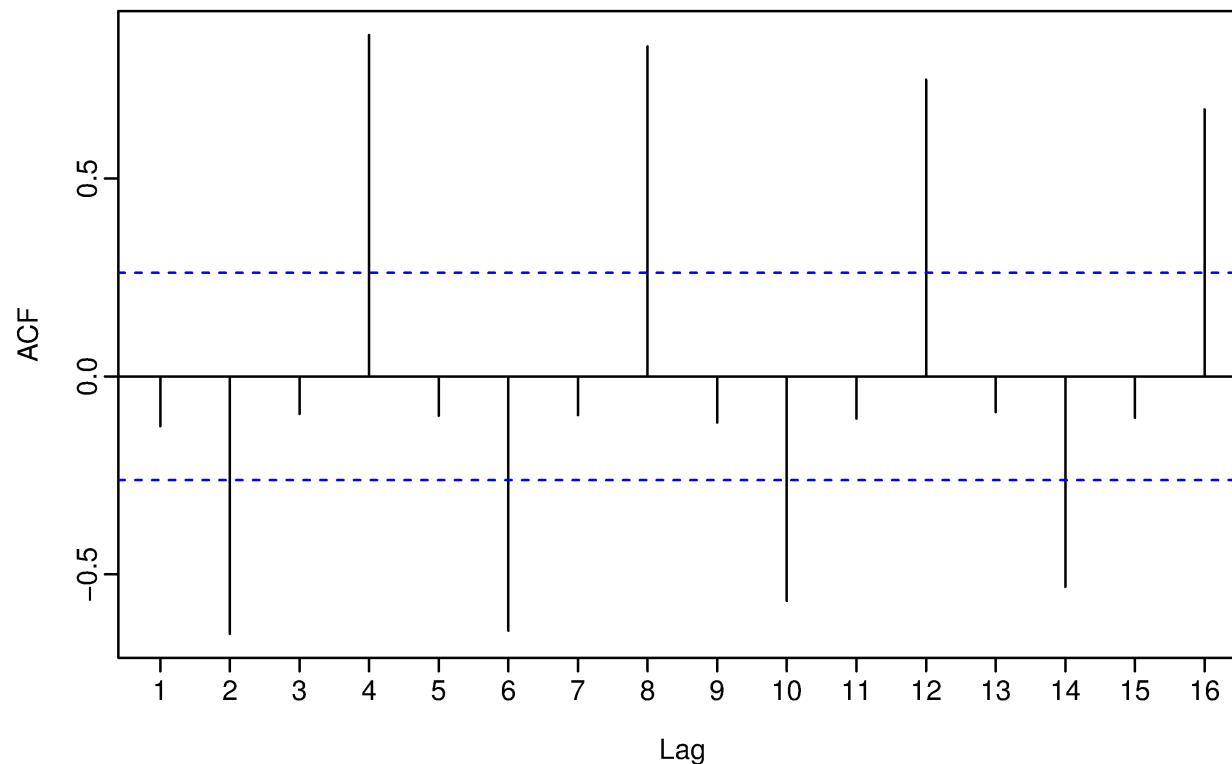
- Relación entre valores retrasados de una serie temporal

$$r_k = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2},$$



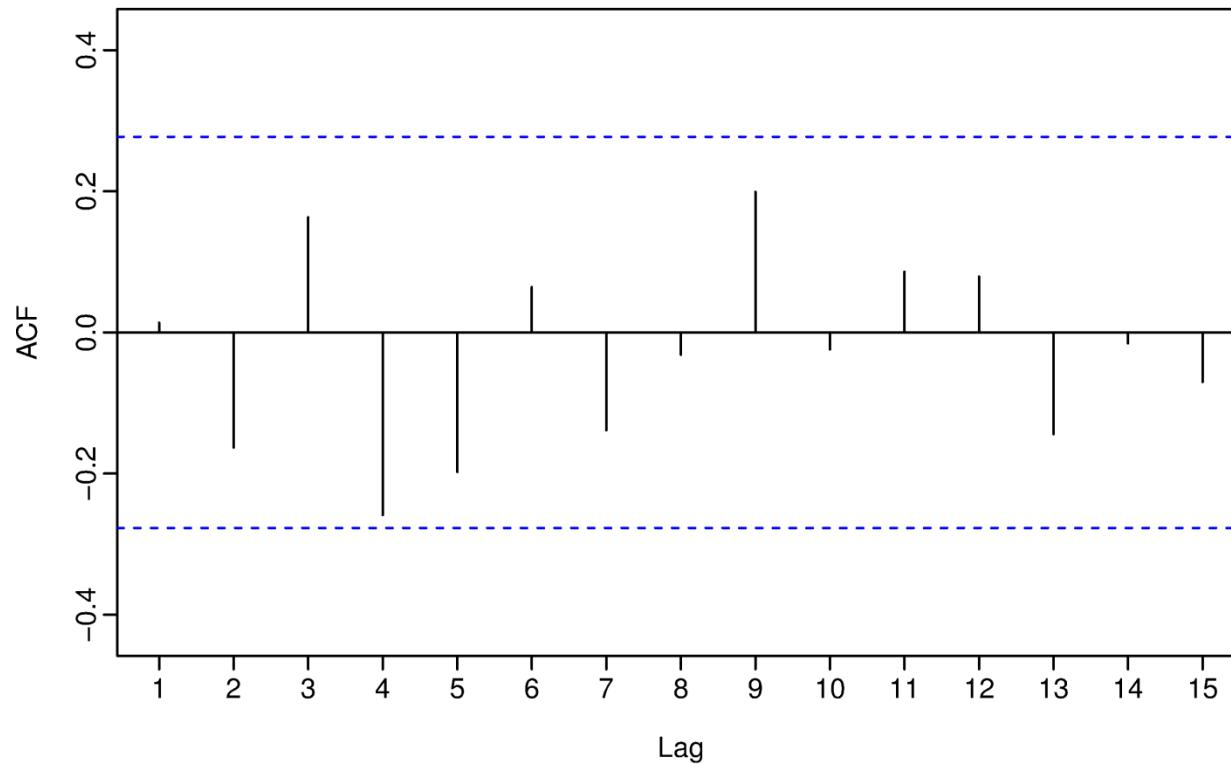
# Función de autocorrelación (ACF)

$r_1$	$r_2$	$r_3$	$r_4$	$r_5$	$r_6$	$r_7$	$r_8$	$r_9$
-0.126	-0.650	-0.094	0.863	-0.099	-0.642	-0.098	0.834	-0.116



# Ruido blanco

- Series que no muestran autocorrelación



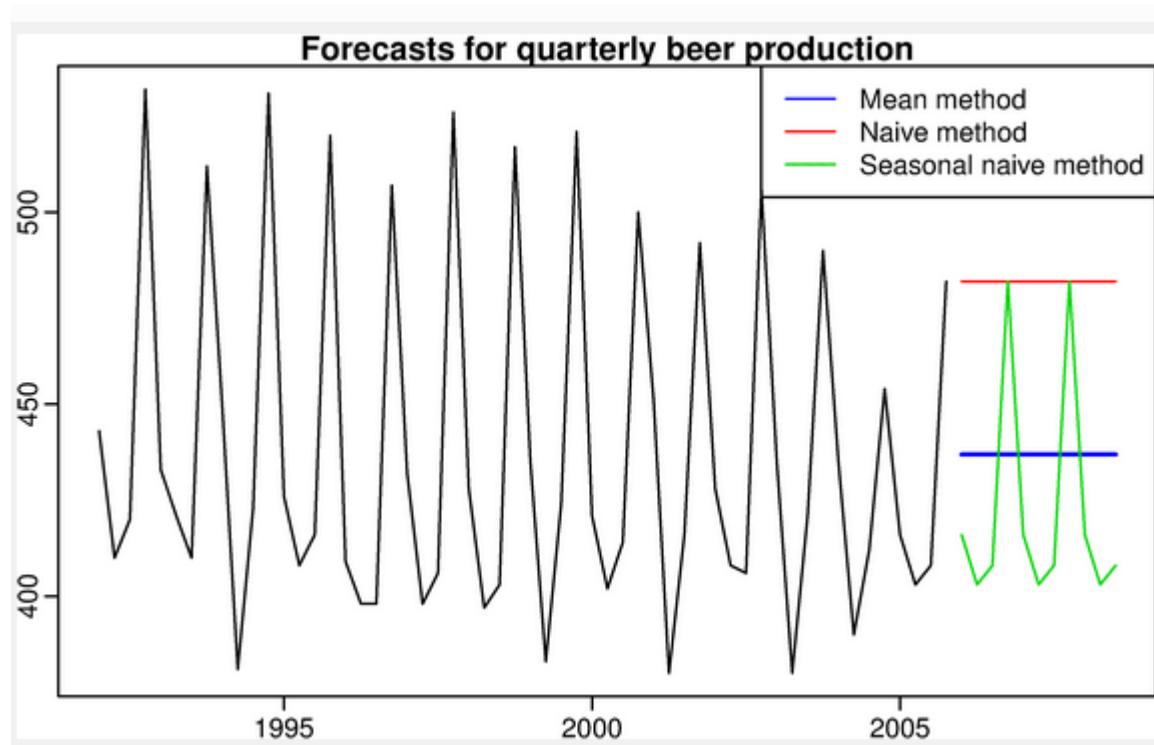
# Métodos de predicción sencillos

- Media

$$\hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T.$$

- Predictor ingenuo
  - Predicción: último valor
- Predictor ingenuo estacional
- Método de deriva
  - La predicción crece o decrece a lo largo del tiempo, en base a la media histórica

# Métodos de predicción sencillos (2)



# Transformaciones

- Ajuste de datos históricos puede llevar a un modelo de predicción más sencillo
- Transformación matemática
- Ajuste de calendario
- Ajuste de población
- Ajuste de inflación

# Evaluando la calidad de la predicción

- Error:

$$e_i = y_i - \hat{y}_i$$

- Error dependiente de escala

$$\text{MAE} = \text{mean}(|e_i|),$$

$$\text{RMSE} = \sqrt{\text{mean}(e_i^2)}.$$

- Error en %:

$$p_i = 100e_i/y_i$$

- Errores escalados

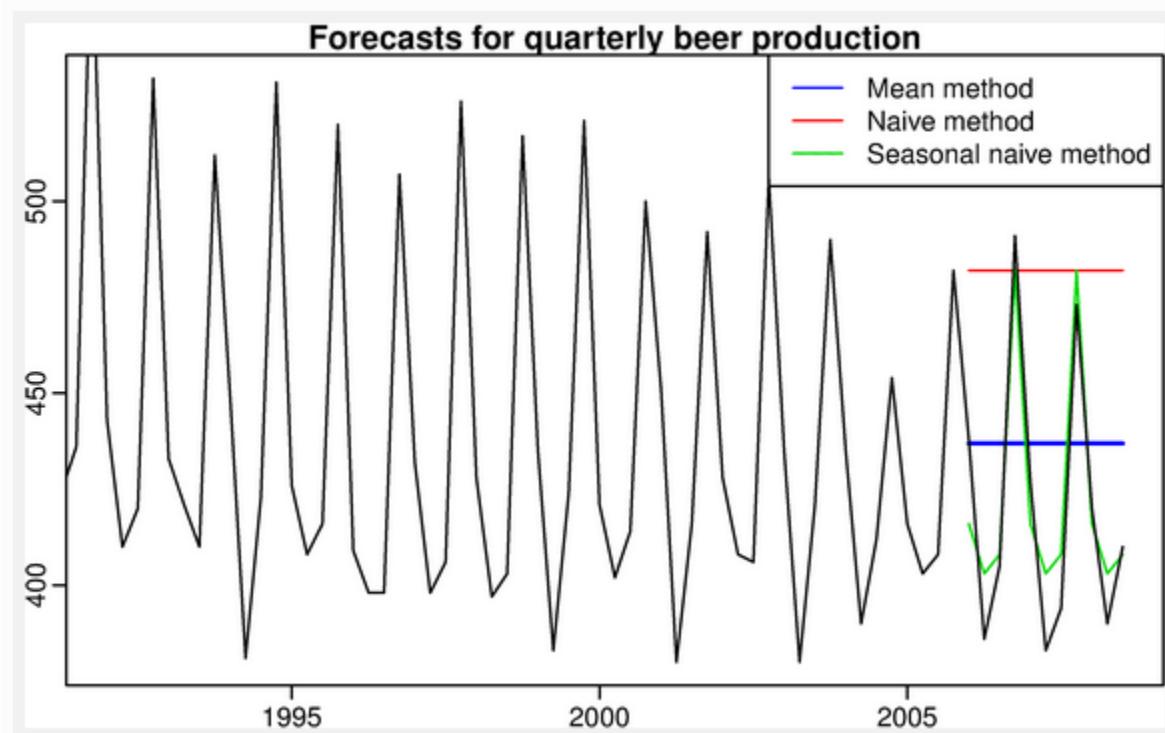
$$\text{MAPE} = \text{mean}(|p_i|).$$

$$\text{sMAPE} = \text{mean} (200|y_i - \hat{y}_i|/(y_i + \hat{y}_i))$$

$$q_j = \frac{e_j}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|}.$$

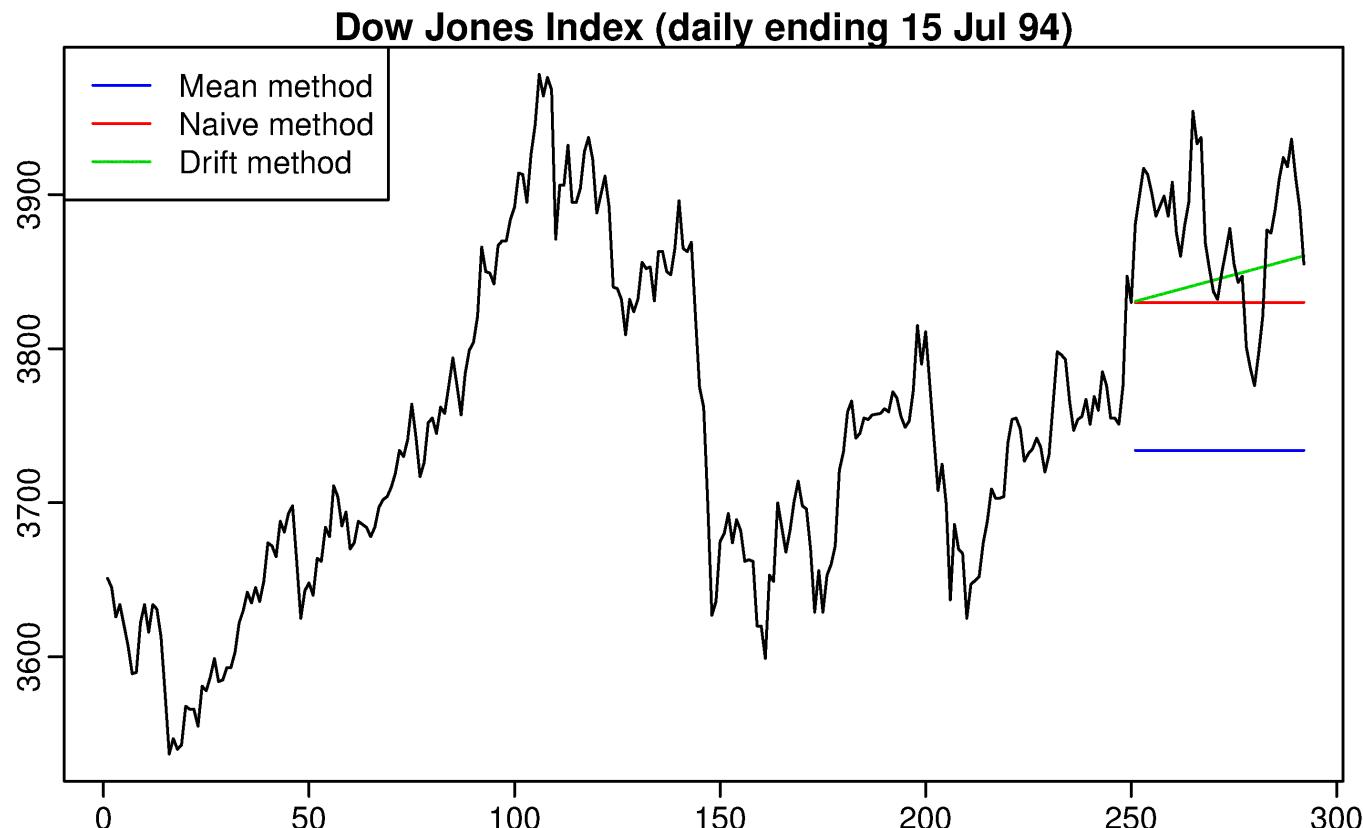
$$\text{MASE} = \text{mean}(|q_j|).$$

# Series temporales estacionales



Method	RMSE	MAE	MAPE	MASE
Mean method	38.01	33.78	8.17	0.61
Naïve method	70.91	63.91	15.88	1.15
Seasonal naïve method	12.97	11.27	2.73	0.20

# Series temporales no estacionales



Method	RMSE	MAE	MAPE	MASE
Mean method	148.24	142.42	3.66	8.70
Naïve method	62.03	54.44	1.40	3.32
Drift method	53.70	45.73	1.18	2.79

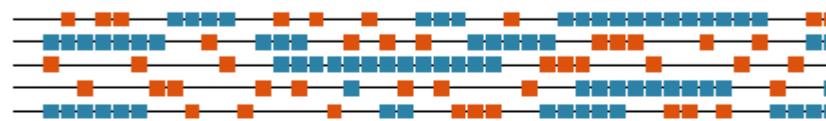
# Metodología

- Como en cualquier otra tarea de modelado es esencial realizar una evaluación correcta
- Dividir datos en: entrenamiento y prueba
- Validación cruzada mejorada
- Blocked Cross-Validation

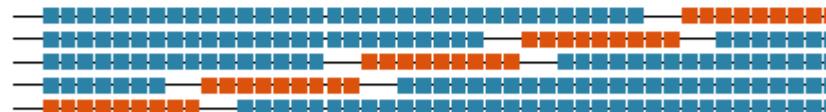
# Procedimientos de selección de modelos



cross-validation



non-dep. cross-validation



blocked cross-validation



last block

# Diagnóstico de residuos

- Residuo: diferencia entre el valor observado y la predicción
- Un buen método de predicción debe producir residuos que cumplan:
  - No están correlados
  - La media es cero
- Cualquier método que no verifique las propiedades previas puede ser mejorado
- Además, es interesante si:
  - Los residuos tienen varianza constante
  - Los residuos se distribuyen según una normal

# Intervalos de predicción

- Un intervalo de predicción es un intervalo dentro del cual está el valor esperado con una probabilidad especificada
- Cuando la predicción es de un paso, la desviación estándar del predictor es casi la misma que la desviación estándar de los residuos

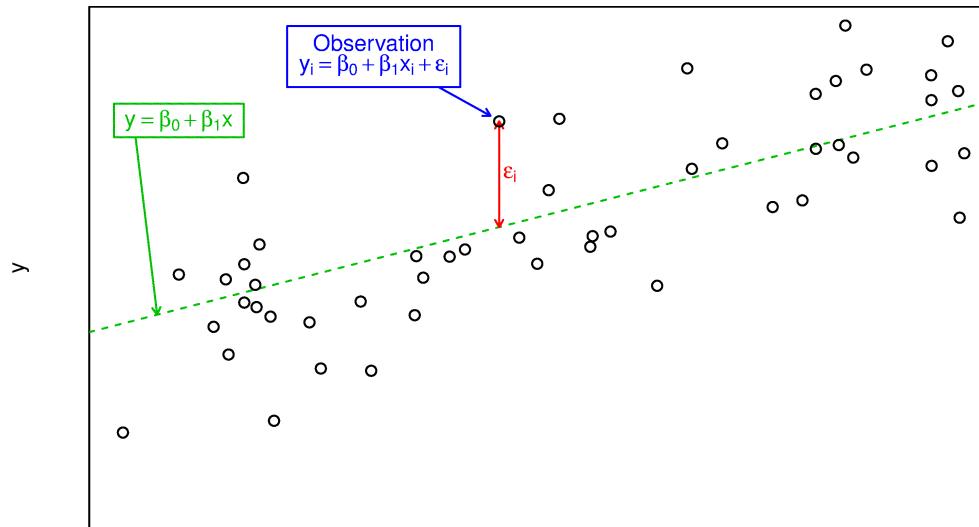
# Contenido

- Predicción
- Herramientas básicas
- **Regresión**
- Descomposición de series temporales
- Alisado exponencial
- Modelos ARIMA
- Modelos de predicción avanzados

# Modelo lineal sencillo

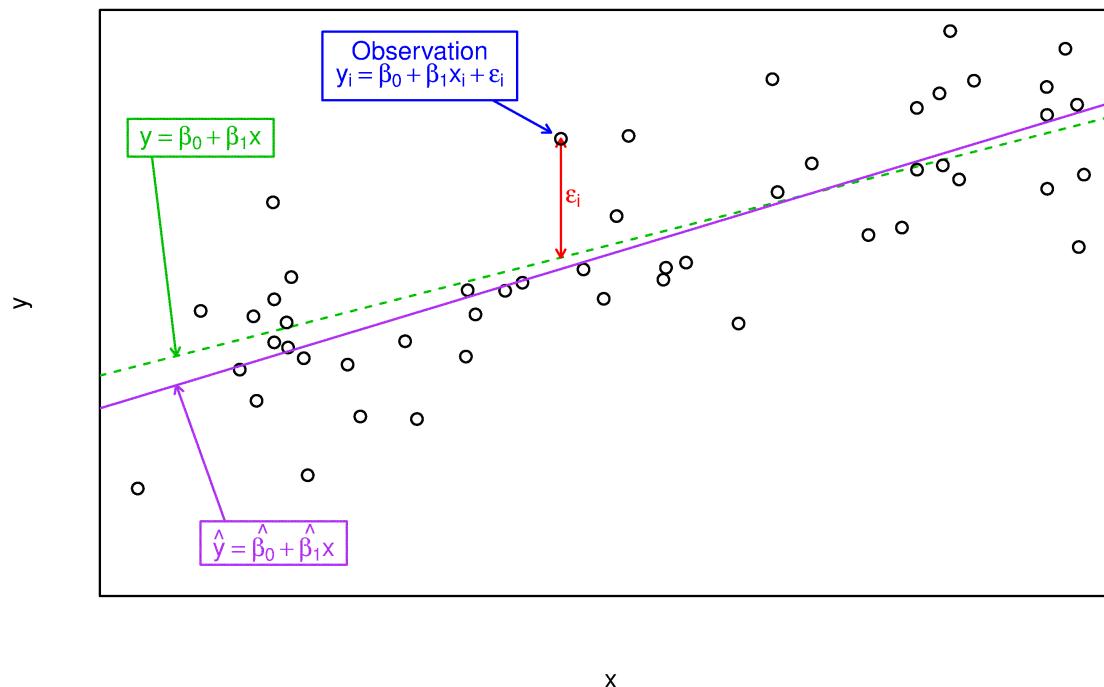
- Asumamos una relación lineal entre el predictor y la variable de salida

$$y = \beta_0 + \beta_1 x + \varepsilon$$



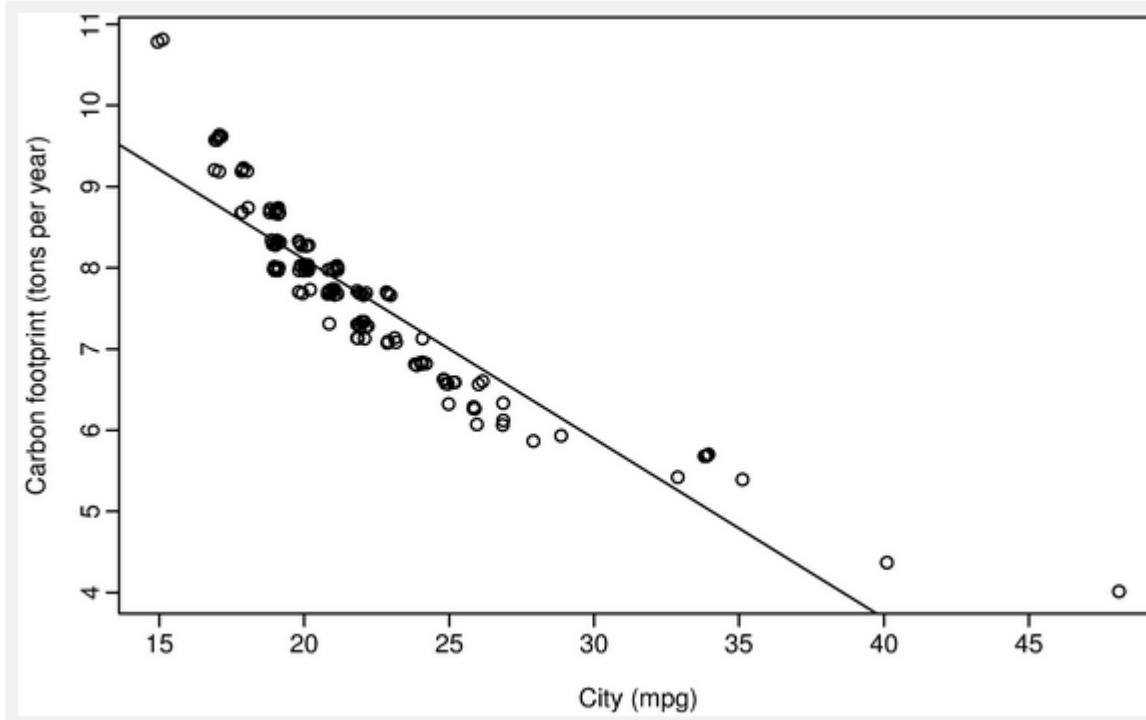
# Aproximación por mínimos cuadrados

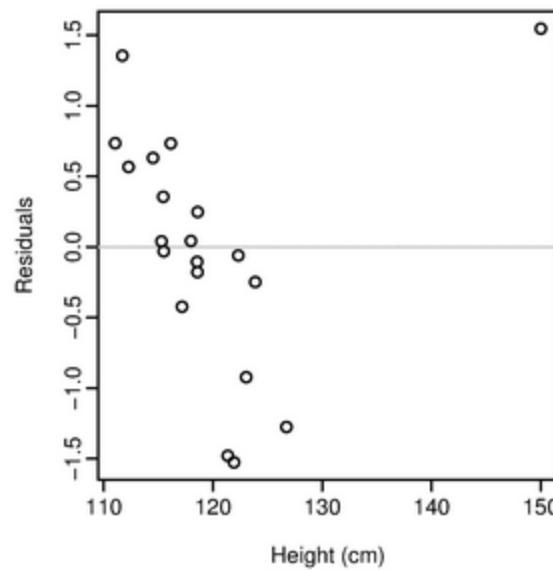
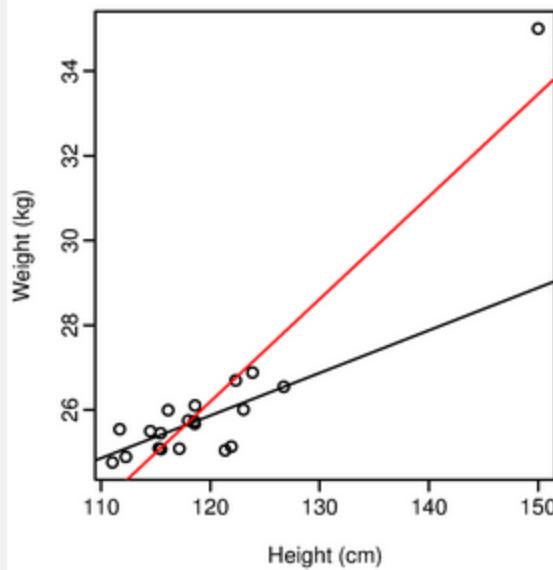
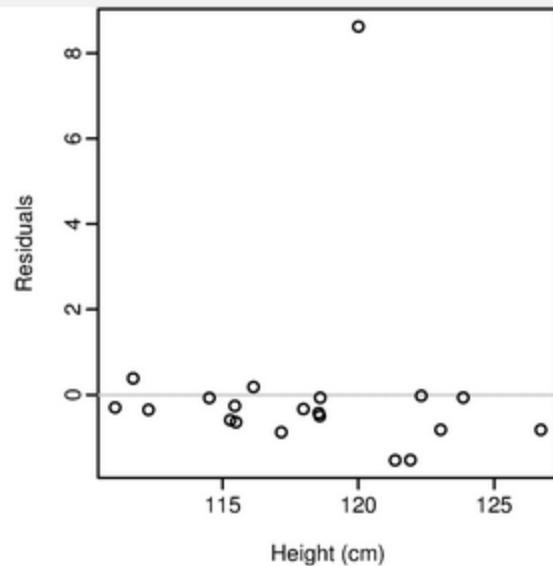
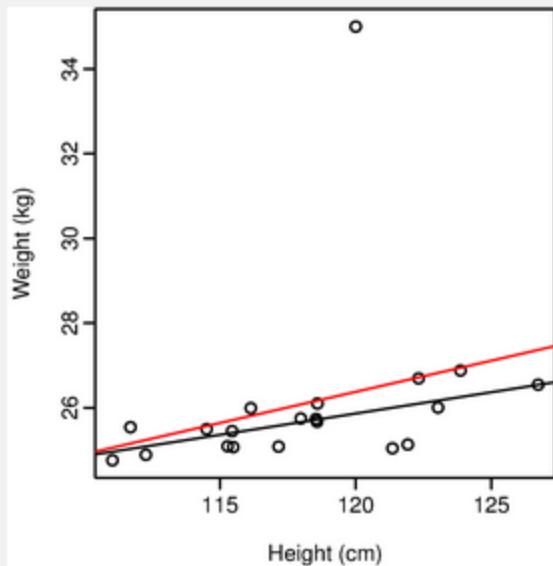
$$\sum_{i=1}^N \epsilon_i^2 = \sum_{i=1}^N (y_i - \beta_0 - \beta_1 x_i)^2$$



# Correlación y regresión

$$\hat{\beta}_1 = r \frac{s_y}{s_x}$$





# Calidad del ajuste

- Coeficiente de ajuste,  $R^2$ :

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2},$$

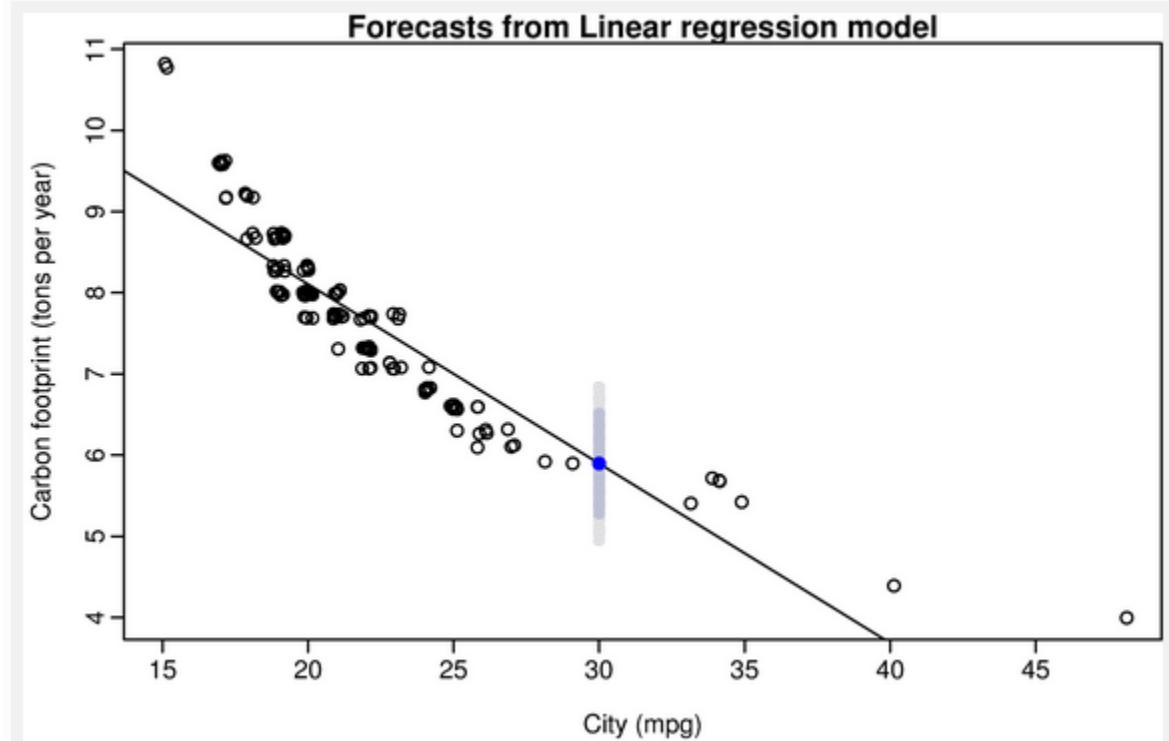
- Error estándar de la regresión

$$s_e = \sqrt{\frac{1}{N-2} \sum_{i=1}^N e_i^2}.$$

# Predicción con regresión

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

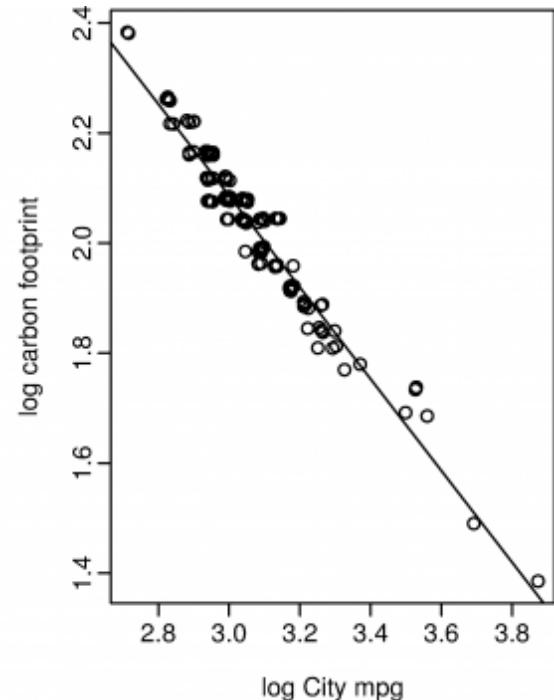
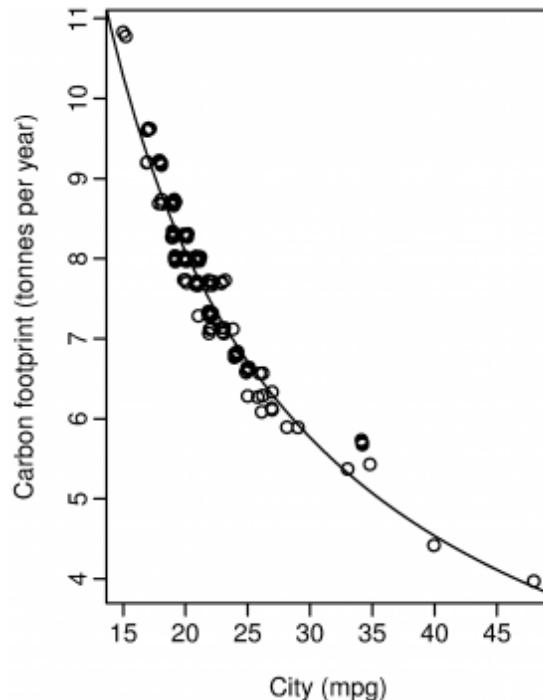
$$\hat{y} \pm 1.96 s_e \sqrt{1 + \frac{1}{N} + \frac{(x - \bar{x})^2}{(N - 1)s_x^2}}$$



# Regresión no lineal

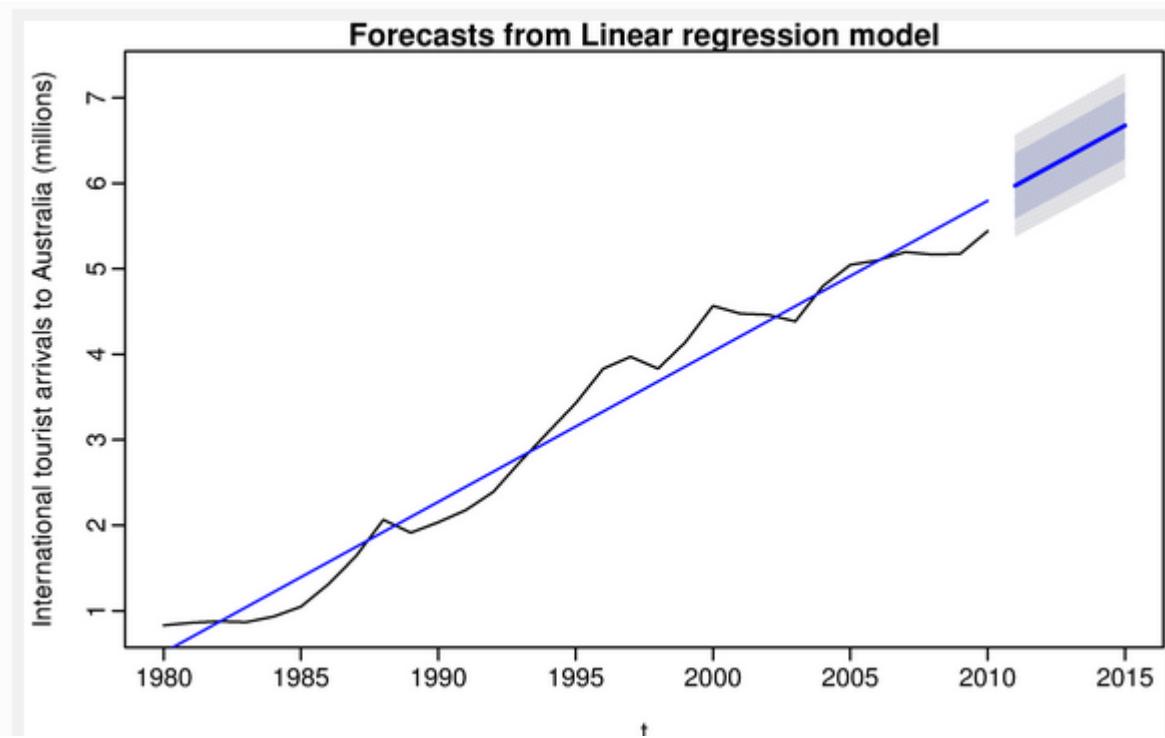
- Existen problemas para los que una función lineal puede ser más adecuada que una lineal
- Se puede obtener mediante la transformación de **x** o **y**

$$\log y_i = \beta_0 + \beta_1 \log x_i + \varepsilon_i.$$

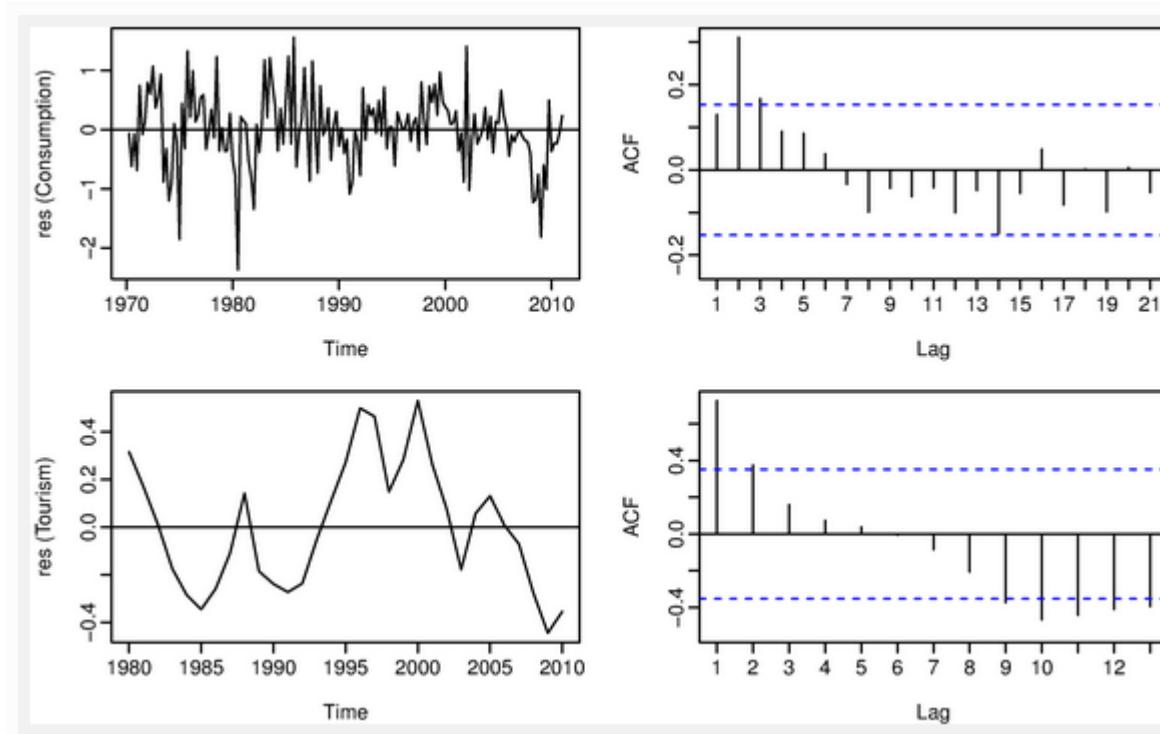


# Regresión con datos de series temporales

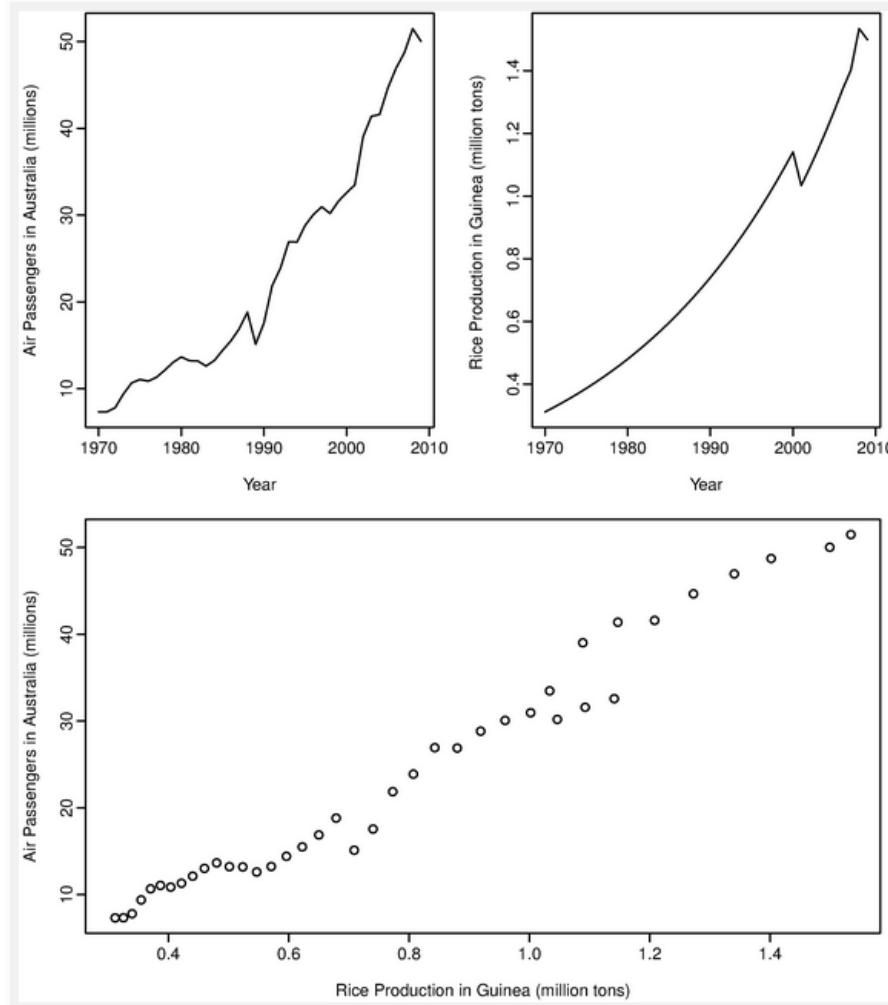
$$y_t = \beta_0 + \beta_1 t + \varepsilon_t$$



# Autocorrelación de los residuos



# Regresión espúrea



# Regresión multivariable

# Regresión múltiple

Una variable para predecir y varias independientes

$$y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \cdots + \beta_k x_{k,i} + e_i.$$

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1,i} - \cdots - \beta_k x_{k,i})^2.$$

# Selección de modelos

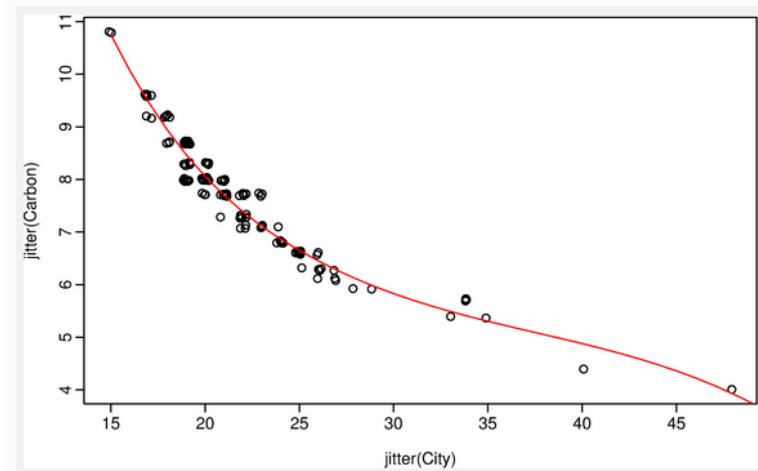
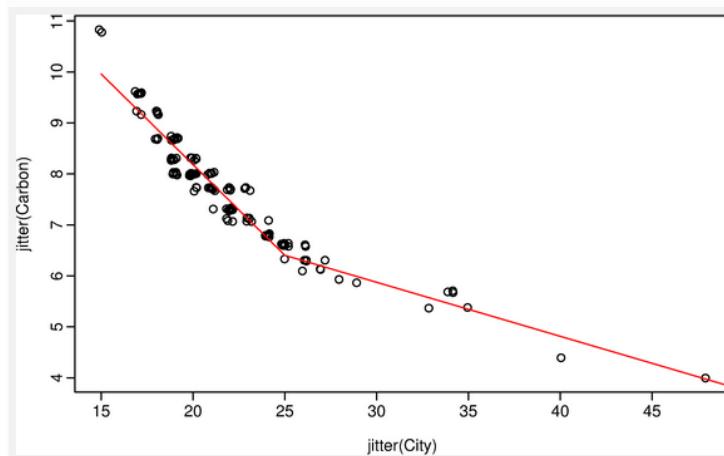
- $R^2$  ajustado
- Validación cruzada
- Criterio de información de Akaike (AIC)
- Criterio de información de Akaike Corregido (AIC)
- Mejor subconjunto de regresión
- Regresión por pasos

# Diagnóstico de residuos: gráficos

- Gráfico de puntos de residuos frente a variables independientes
- Gráfico de puntos de residuos frente a valores ajustados
- Autocorrelación de los residuos
- Histograma de los residuos

# Regresión no lineal

$$y = f(x) + e$$



# Correlación no es causa

- Una variable **x** puede ser útil para predecir la variable **y**, pero eso no significa que **x** causa **y**
- Las correlaciones son útiles para la predicción, aunque no haya relaciones causales entre dos variables

# Contenido

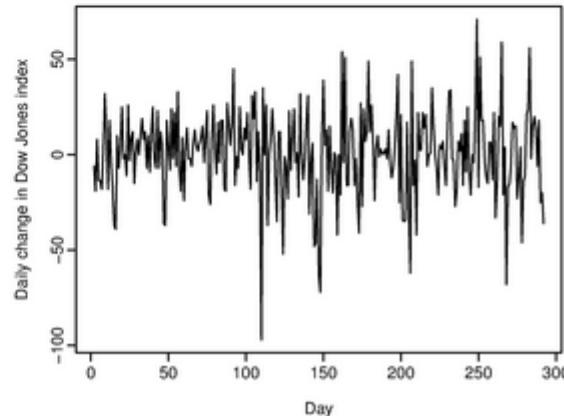
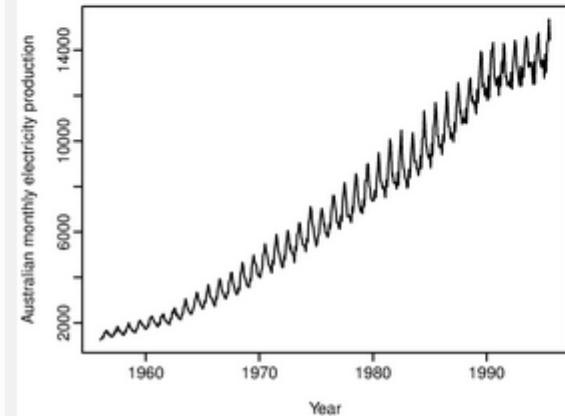
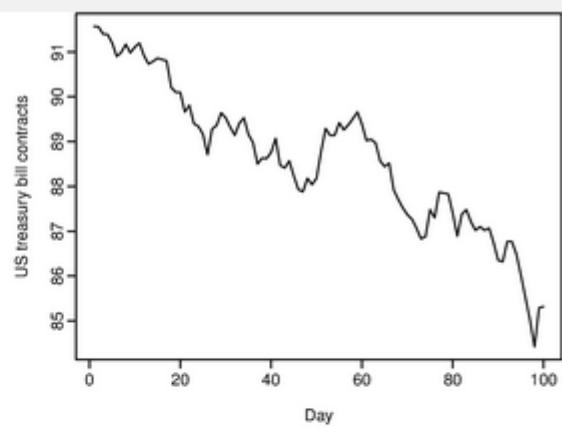
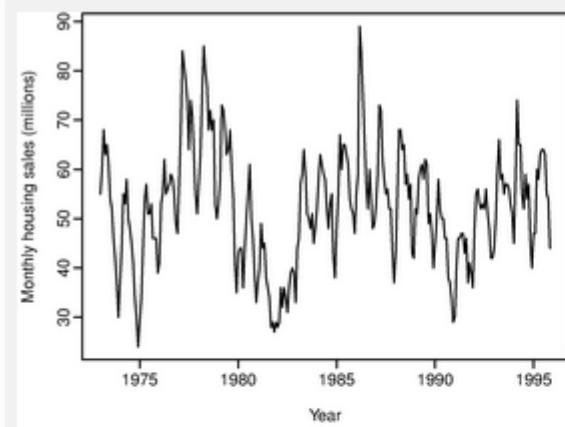
- Predicción
- Herramientas básicas
- Regresión
- **Descomposición de series temporales**
- Alisado exponencial
- Modelos ARIMA
- Modelos de predicción avanzados

# Descomposición de series temporales

- Las series temporales pueden mostrar una gran variedad de patrones y es útil categorizarlos y los comportamientos que se pueden visualizar
- A veces es útil descomponer una serie en componentes diferenciados, cada uno de los cuales representa un aspecto

# Componentes de las series

Tendencia, estacionalidad, ciclo



# Descomposición

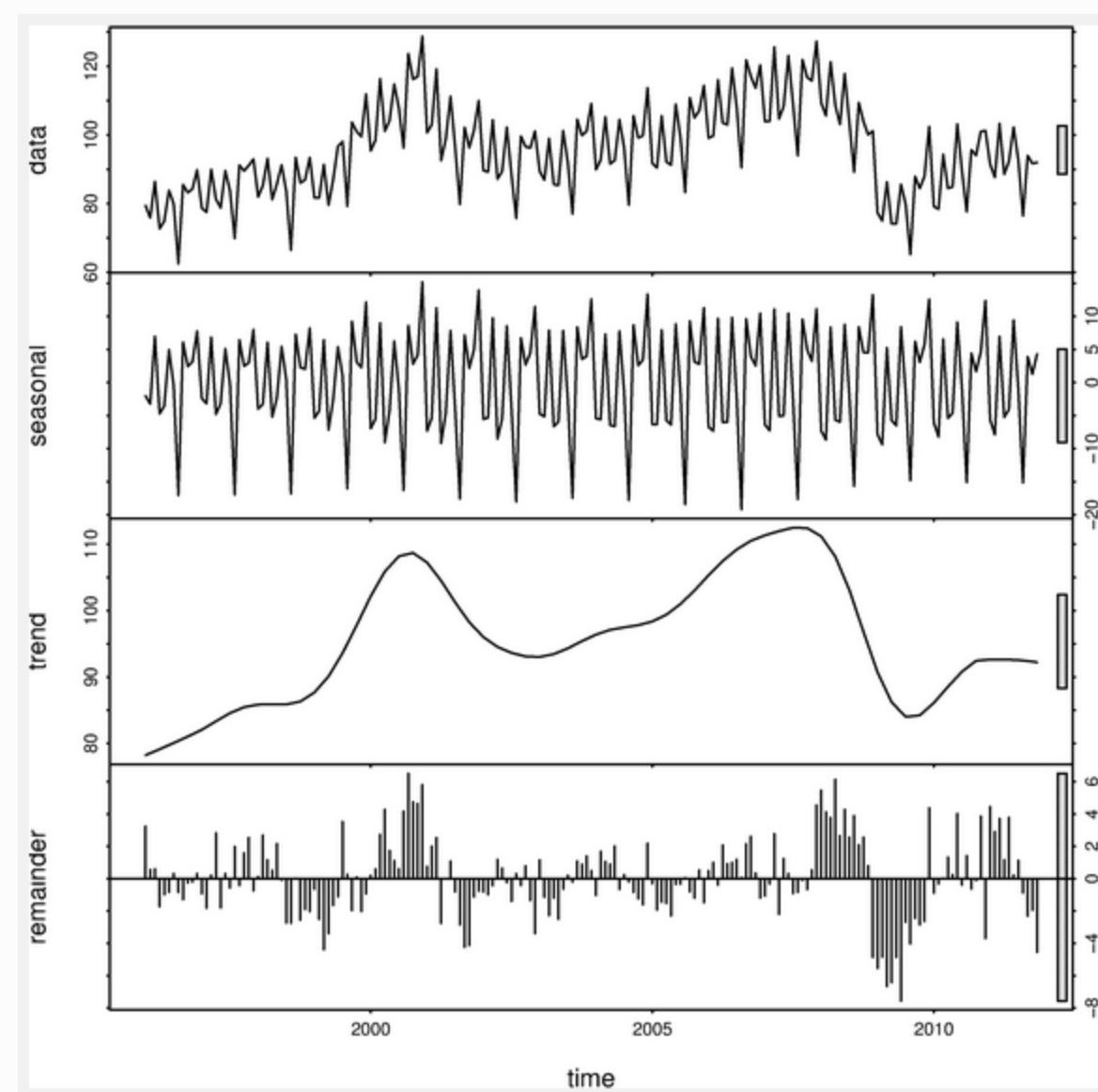
## Descomposición aditiva

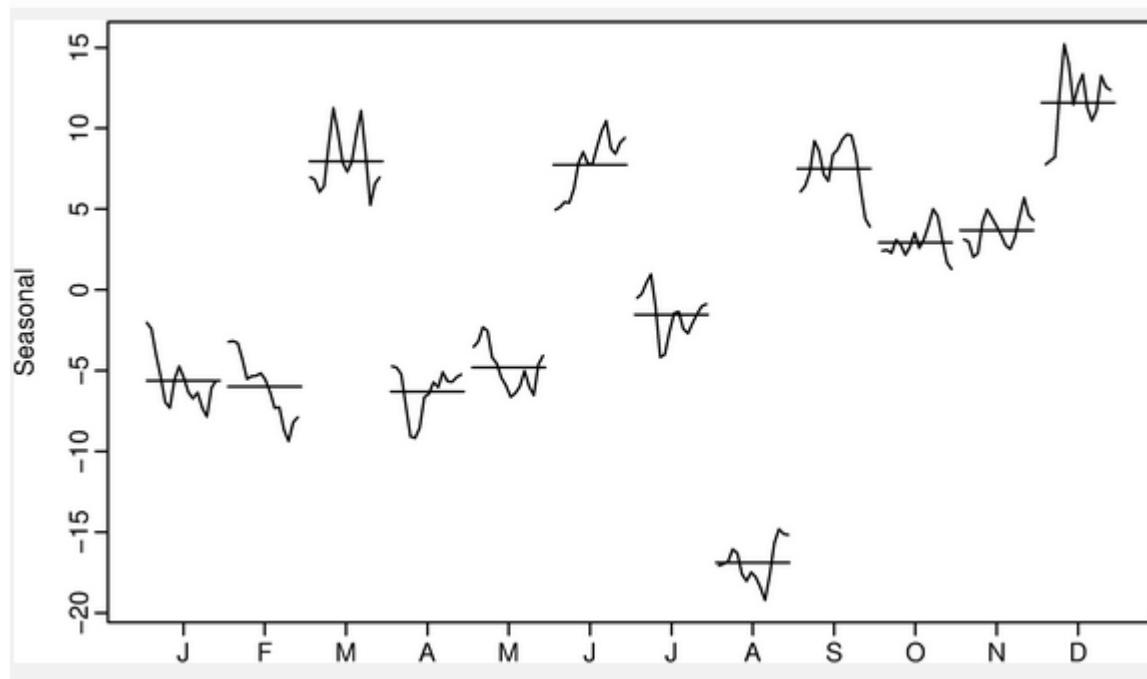
$$y_t = S_t + T_t + E_t$$

Indicada cuando la magnitud de las fluctuaciones estacionales por las variaciones entorno a la tendencia no varían con el valor de la serie temporal

## Descomposición multiplicativa

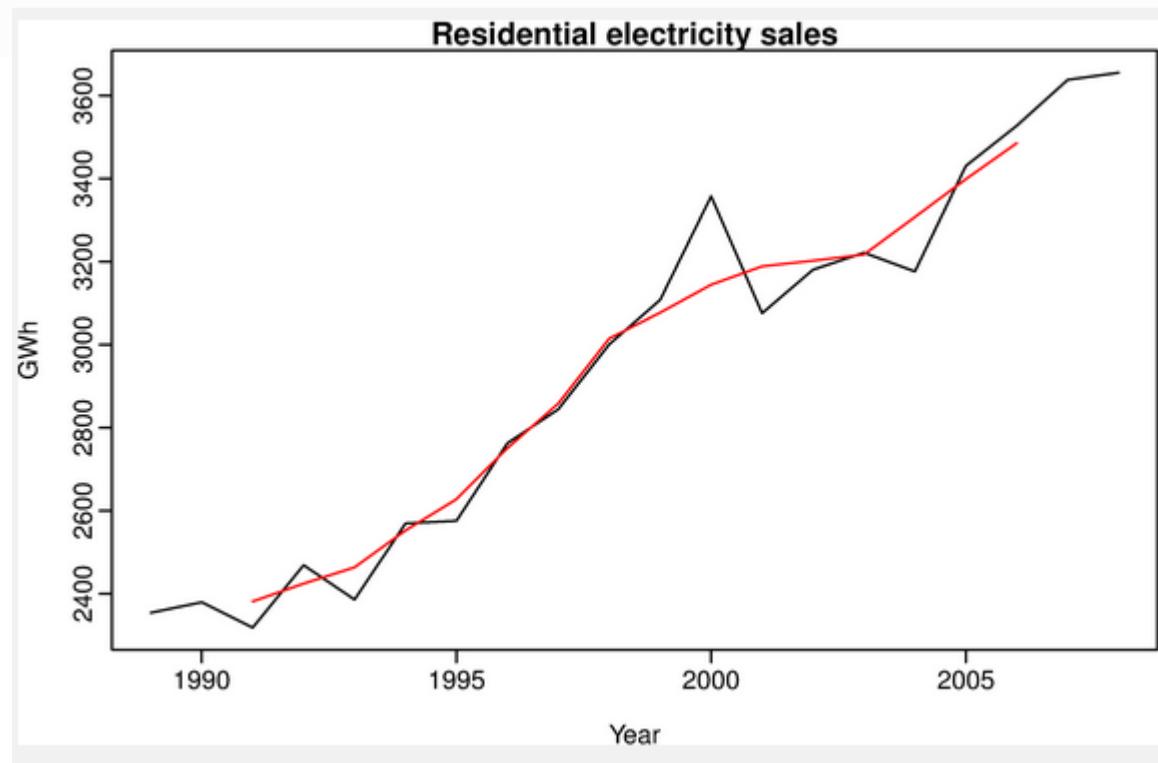
$$y_t = S_t \times T_t \times E_t$$

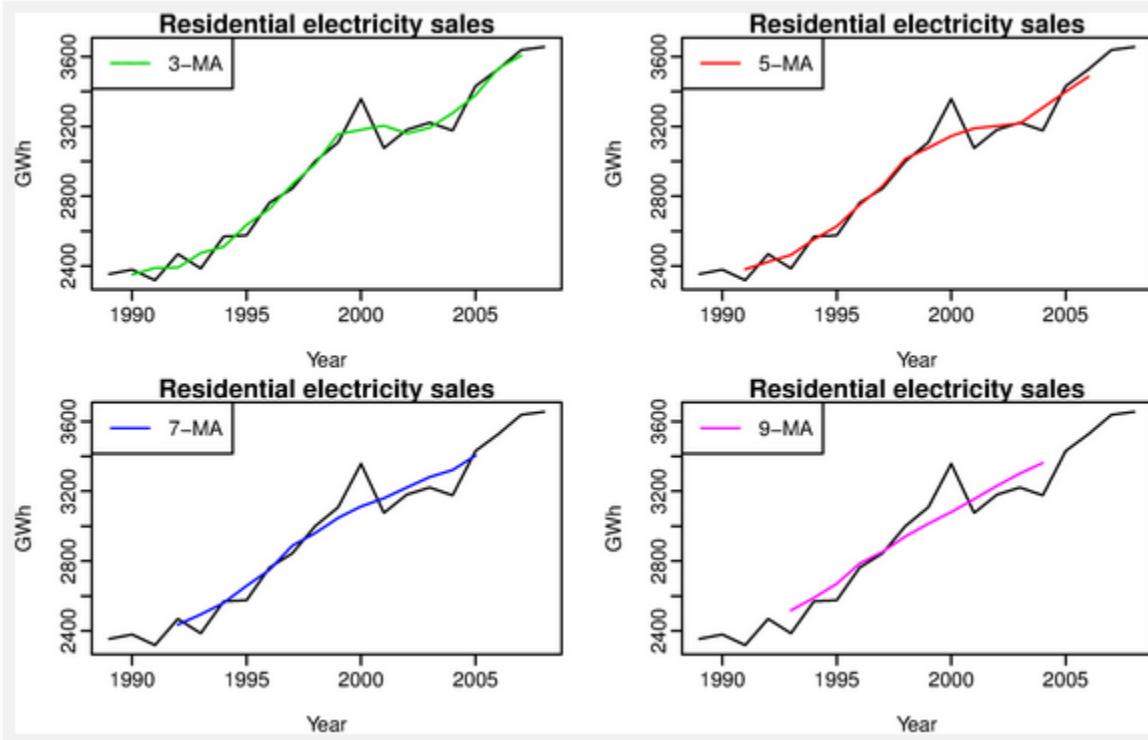




# Medias móviles

$$\hat{T}_t = \frac{1}{m} \sum_{j=-k}^k y_{t+j}$$





Utilizado frecuentemente para detectar tendencia y ciclo en datos estacionales

# Medias móviles ponderadas

$$\hat{T}_t = \sum_{j=-k}^k a_j y_{t+j}$$

Su principal ventaja es que proporcionan una aproximación más “suave” de la tendencia-ciclo

# Descomposición aditiva clásica

## Step 1

If  $m$  is an even number, compute the trend-cycle component using a  $2 \times m$ -MA to obtain  $\hat{T}_t$ . If  $m$  is an odd number, compute the trend-cycle component using an  $m$ -MA to obtain  $\hat{T}_t$ .

## Step 2

Calculate the detrended series:  $y_t - \hat{T}_t$ .

## Step 3

To estimate the seasonal component for each month, simply average the detrended values for that month. For example, the seasonal index for March is the average of all the detrended March values in the data. These seasonal indexes are then adjusted to ensure that they add to zero. The seasonal component is obtained by stringing together all the seasonal indices for each year of data. This gives  $\hat{S}_t$ .

## Step 4

The remainder component is calculated by subtracting the estimated seasonal and trend-cycle components:  $\hat{E}_t = y_t - \hat{T}_t - \hat{S}_t$ .

# Descomposición multiplicativa clásica

## Step 1

If  $m$  is an even number, compute the trend-cycle component using a  $2 \times m$ -MA to obtain  $\hat{T}_t$ . If  $m$  is an odd number, compute the trend-cycle component using an  $m$ -MA to obtain  $\hat{T}_t$ .

## Step 2

Calculate the detrended series:  $y_t/\hat{T}_t$ .

## Step 3

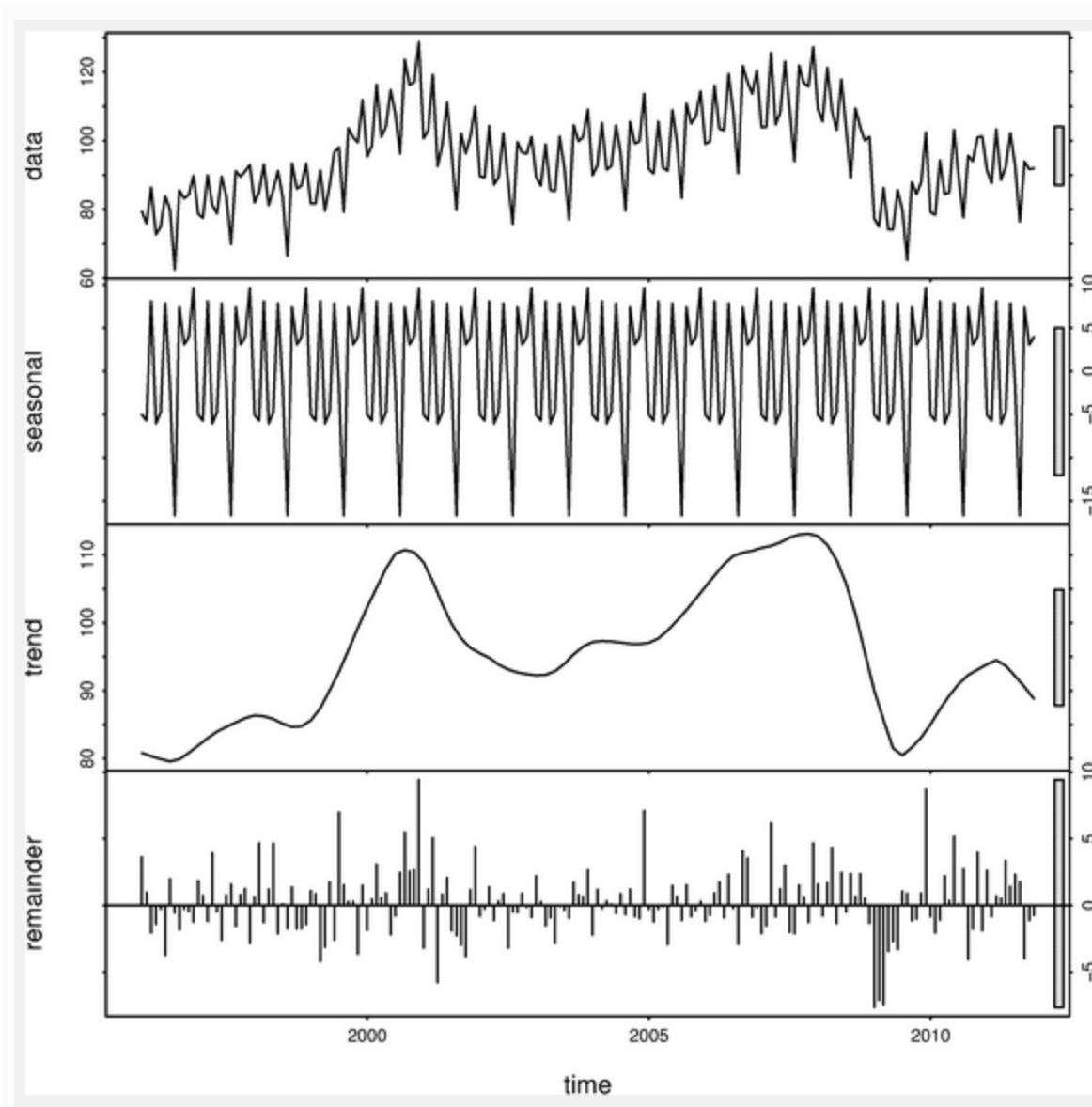
To estimate the seasonal component for each month, simply average the detrended values for that month. For example, the seasonal index for March is the average of all the detrended March values in the data. These seasonal indexes are then adjusted to ensure that they add to  $m$ . The seasonal component is obtained by stringing together all the seasonal indices for each year of data. This gives  $\hat{S}_t$ .

## Step 4

The remainder component is calculated by subtracting the estimated seasonal and trend-cycle components:  $\hat{E}_t = y_t / (\hat{T}_t \hat{S}_t)$ .

# Descomposición STL

- STL: método robusto y versátil de descomposición
  - It can handle any type of seasonality
  - The seasonal component is allowed to change over time, within a range controllable by the user
  - The smoothness of the trend-cycle can also be controlled by the user
  - It is robust to outliers



# Predicción con descomposición

- Simple: predecir los componentes individuales y agregarlos

$$y_t = \hat{S}_t + \hat{A}_t$$

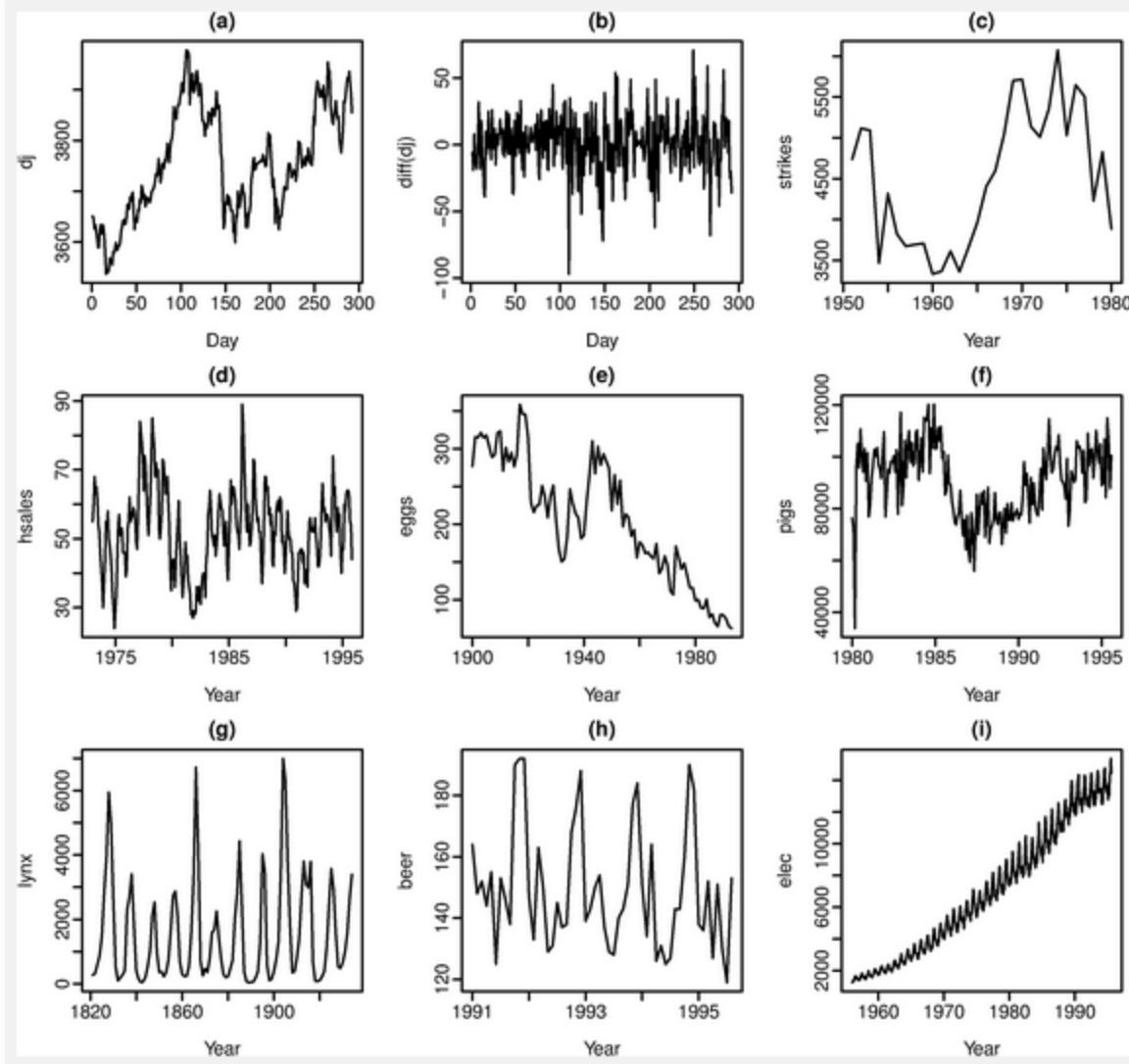
$$y_t = \hat{S}_t \hat{A}_t,$$

# Contenido

- Predicción
- Herramientas básicas
- Regresión
- Descomposición de series temporales
- **Modelos ARIMA**
- Modelos de predicción avanzados

# Estacionariedad

- Una serie estacionaria es aquélla cuyas propiedades no dependen del momento en que se observa la serie
- El gráfico ACF es útil para identificar series no estacionarias. Para series estacionarias se aproxima a cero rápidamente, mientras que para las series no estacionarias lo hace más lentamente



# Calcular diferencias

- Calcular las diferencias entre observaciones sucesivas
- Las transformaciones como el logaritmo pueden ayudar a estabilizar la varianza de las series temporales
- Las diferencias pueden ayudar a estabilizar la media de la serie temporal eliminando cambios en el valor de la serie, y por tanto, eliminando la tendencia y la estacionalidad

# Modelo de paseo aleatorio

- Es una serie construida añadiendo el término de error a cada nuevo término de la serie:

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

donde la media de  $e_t$  es cero y su desviación típica es constante

- Los paseos aleatorios típicamente tienen:
  - Periodos largos de tendencias aparentes alcistas o bajistas
  - Cambios de dirección repentinos e impredecibles

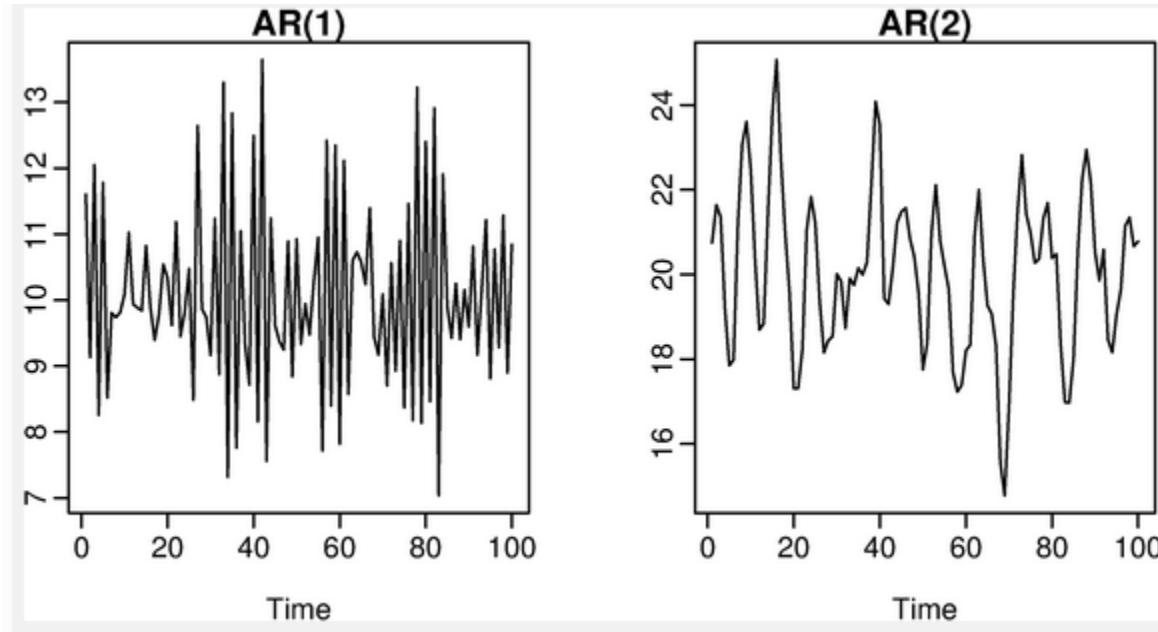
# Pruebas de raíz unidad

- Test de hipótesis estadístico de estacionariedad diseñados para determinar si es necesario hacer diferencias
- Test Dickey-Fuller ampliado

$$y'_t = \phi y_{t-1} + \beta_1 y'_{t-1} + \beta_2 y'_{t-2} + \cdots + \beta_k y'_{t-k},$$

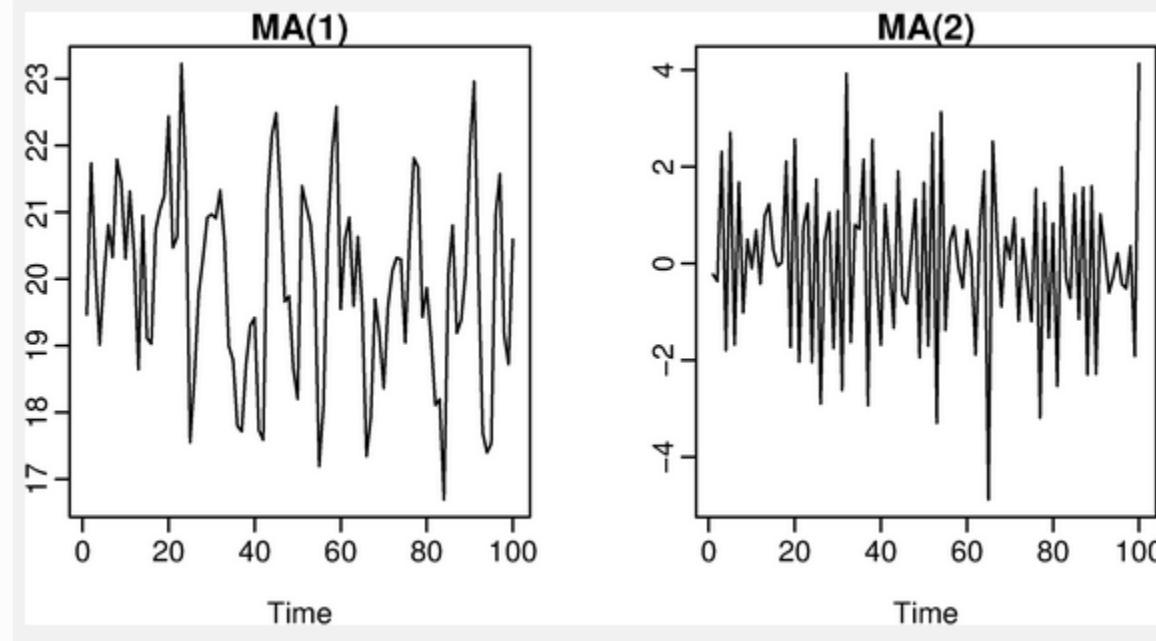
# Modelos autorregresivos

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



# Modelos de medias móviles

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



# Modelos ARIMA no estacionales

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t$$

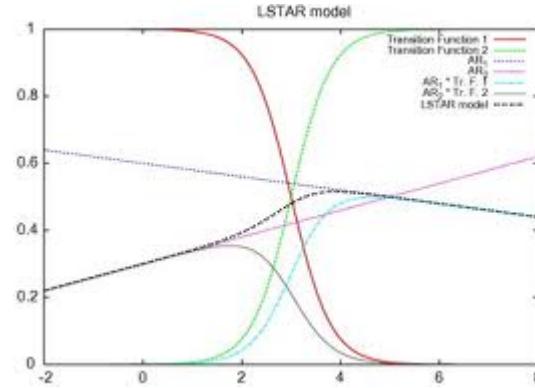
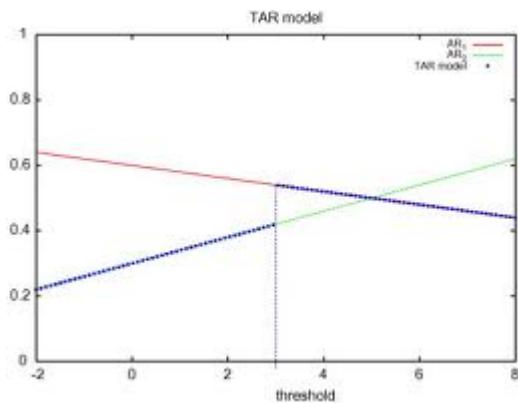
- ARIMA(p,d,q)
  - p: orden de la parte autorregresiva
  - d: grado de las diferencias
  - q: orden de la parte de medias móviles

White noise	ARIMA(0,0,0)
Random walk	ARIMA(0,1,0) with no constant
Random walk with drift	ARIMA(0,1,0) with a constant
Autoregression	ARIMA(p,0,0)
Moving average	ARIMA(0,0,q)

# Métodos de predicción avanzados

# Threshold Autoregressive Models

- Múltiples métodos autorregresivos, definidos por funciones indicador



# Familia de modelos TAR

- TAR: Threshold AR
- STAR: Smooth TAR
- LSTAR: Logistic TAR
- NCSTAR: Neural Coefficient STAR

# Conexiones entre las familias de modelos

- Muchas familias son realmente distintas caras de la misma idea básica
- Existen relaciones cercanas entre esas familias, lo que permite intercambiar propiedades y procedimientos
- ANN = FRBS
- TAR = FRBS
- SVR = FRBS

# Incrustamiento de series temporales

lag4	lag3	lag2	lag1	target
	...	...		
14	10	26	11	-13
10	26	11	-13	-15
<b>26</b>	<b>11</b>	<b>-13</b>	<b>-15</b>	<b>-8</b>
11	-13	-15	-8	35
-13	-15	-8	35	40
<b>-15</b>	<b>-8</b>	<b>35</b>	<b>40</b>	<b>-8</b>
-8	35	40	-8	-16
35	<b>40</b>	<b>-8</b>	<b>-16</b>	7
40	-8	-16	7	17
	...	...		

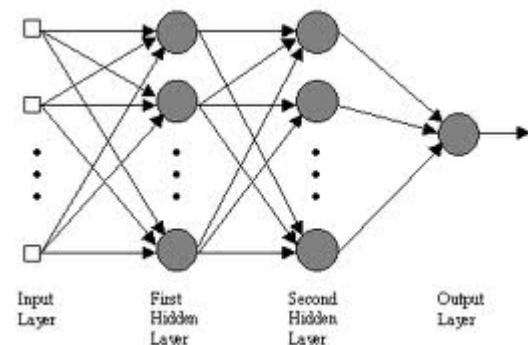
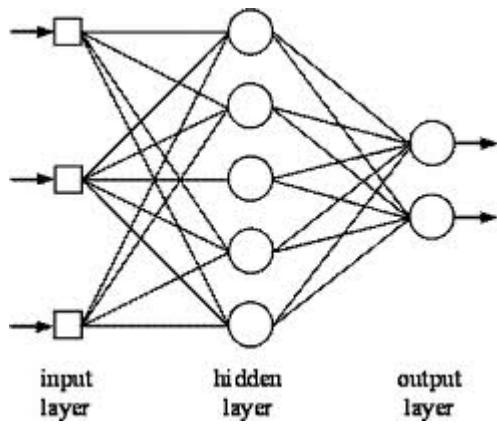
# Redes Neurales Artificiales

- Perceptrones multicapa
- RBF
- Redes Neuronales Recurrentes
  - Elman
- ...

# MLP

- Los MLPs son los modelos más conocidos y usados
- Debido a su eficacia en problemas de regresión se aplican frecuentemente en predicción de series temporales
- Las mismas consideraciones que se hacen para problemas de regresión regular deben hacerse para la predicción de series temporales

# Capas ocultas

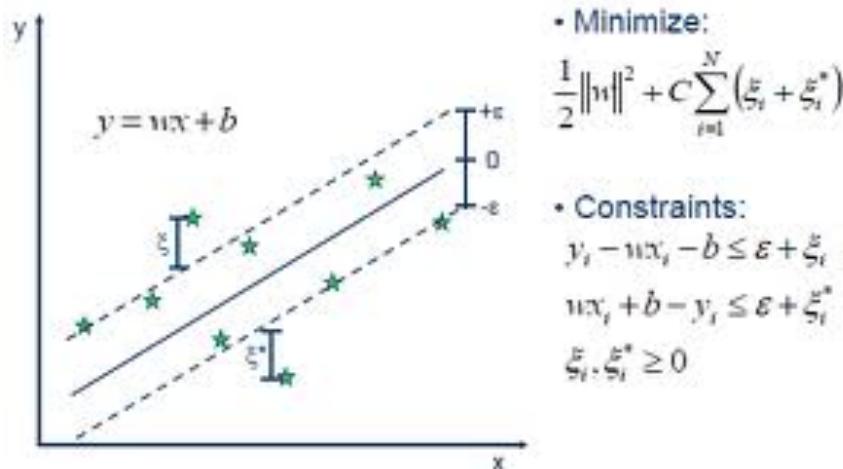


# Pasos para la aplicación de MLP

- Definir el problema: entradas y salidas
- Preprocesamiento: transformaciones sobre los datos
- Definir la arquitectura de la red
  - Número de capas; número de unidades en cada capa
  - Función de activación
- Definir algoritmos de aprendizaje y parámetros
- Ajustar el modelo
- Validar el modelo
- Aplicarlo

# Support Vector Regression

- Esta es la versión para regresión de las SVM

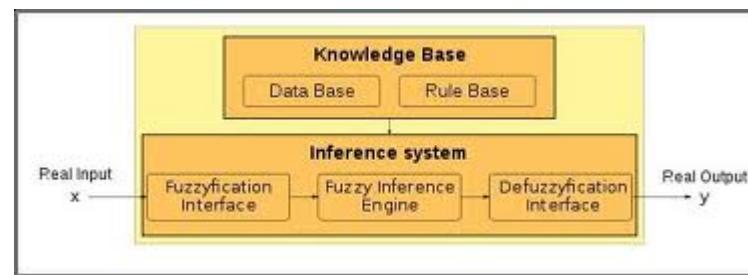


# Pasos para la aplicación de SVR

- Definir el problema: entradas y salidas
- Preprocesamiento: transformaciones sobre los datos
- Definir los parámetros de la SVM
  - Tipo de la función kernel
  - Parámetros
- Ajustar el modelo
- Validar el modelo
- Aplicarlo

# Sistemas basados en reglas difusas

- En particular, los TSK, más adecuados para regresión
- Pueden aportar una interpretación de los modelos

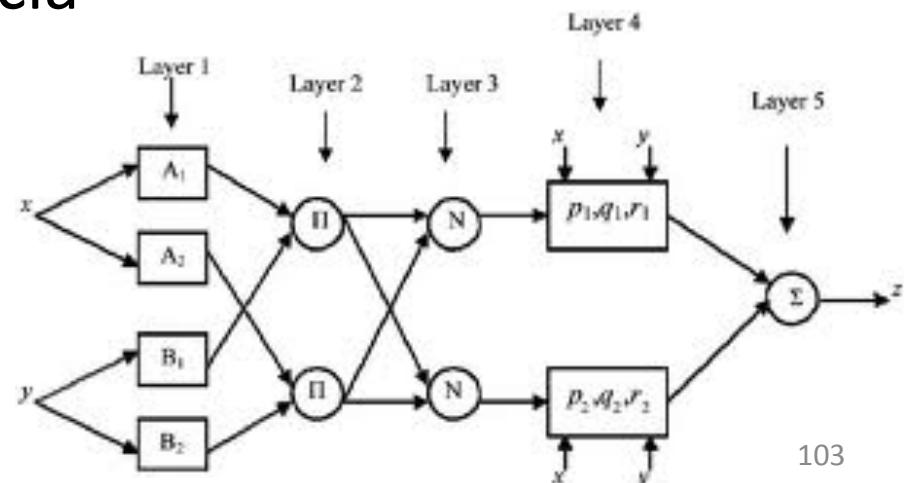


# Pasos para la aplicación de FRBS

- Define the problem: inputs and outputs
- Apply possible transformations to data
- Define linguistic label partition:
  - Granularity: number of labels
  - Semantic:
    - Kind of membership functions
    - Params for each label
  - Structure of consequents
- Create initial rule base:
  - From experts after knowledge acquisition
  - Derived from data
- Define the learning/tuning algorithms and parameters
  - Gradient descent
  - Evolutionary
  - Hybrid model
- Fit the model
- Validate the model
- Deploy it

# ANFIS

- Modelo neuro-difuso con algoritmo de entrenamiento híbrido:
  - Optimización basada en mínimos cuadrados para los coeficientes de salida
  - Descenso en gradiente para los parámetros de las funciones de pertenencia



# Otras técnicas

- Wavelets
- Spectral Analysis
- State-space models
- Hidden Markov models
- MARS
- Functional analysis
- ...

# Referencias

- C. Chatfield, «The analysis of time series: An Introduction», CRC Press, 2003
- J.D. Hamilton, «Time Series Analysis», Princeton University Press, 1994
- R. Hyndman, G. Athanasopoulos, «Forecasting and time series» 2013
- P.J. Brockwell, R.A. Davis, «Time Series: Theory and Methods», 2nd Ed., Springer, 1991
- J.S. Armstrong (ed), «Principles of Forecasting: A Handbook for Researchers and Practitioners», Springer, 2001

## Referencias (2)

- S.G. Makridakis, S.C. Wheelwright, R.J. Hyndman, «Forecasting», 3rd Ed., Wiley & Sons, 1998
- P.J. Brockwell, R.A. Davis, «Introduction to Time Series and Forecasting», 2nd ed., Springer, 2002
- A.K. Palit, D. Popovic, «Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications», Springer, 2005
- R.H. Shumway, D.S. Stoffer, «Time Series Analysis and Its Applications», Springer, 2nd Ed., 2006

