FC2018: ARIMA

Samuel Orso

2018-10-02

Outline

Stationarity

Definition

If $\{y_t\}$ is a stationary time series, then for all s, the distribution of $(y_t, \ldots, y_{t+s})$ does not depend on t.

Stationarity

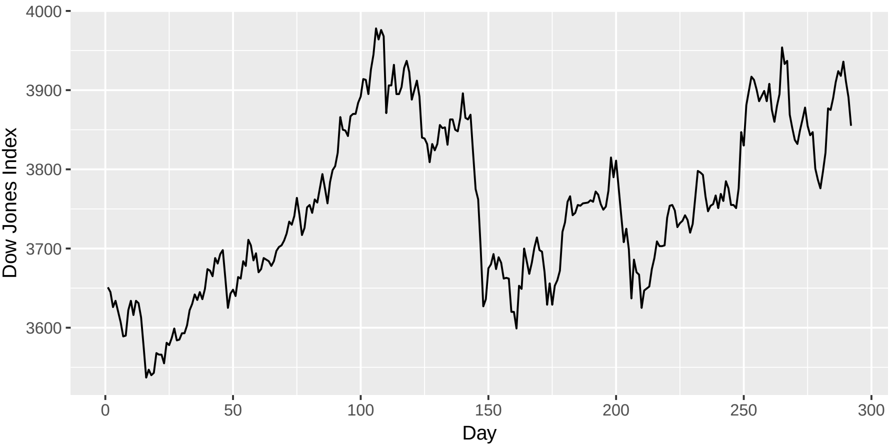
Definition

If $\{y_t\}$ is a stationary time series, then for all s, the distribution of $(y_t, \dots, y_{t+s})$ does not depend on t.

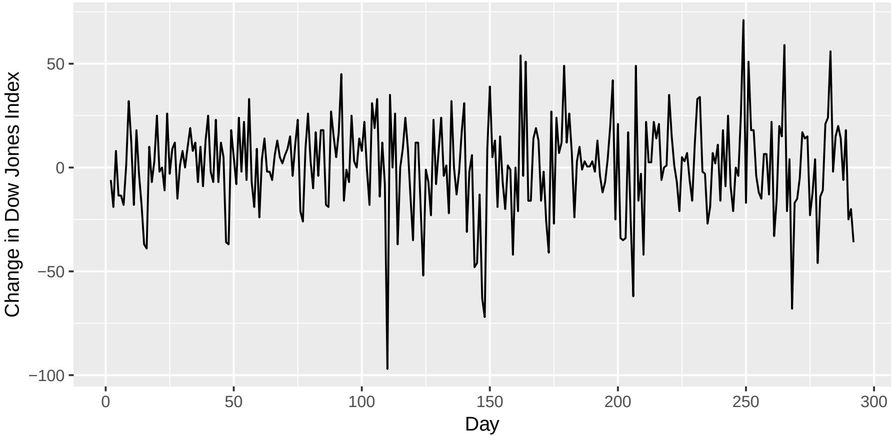
A **stationary series** is:

- roughly horizontal
- constant variance
- no patterns predictable in the long-term

Stationary?


4000
3900
3800
3700
3600
Dow Jones Index
0
50
100
150
200
250
300
Day

Stationary?

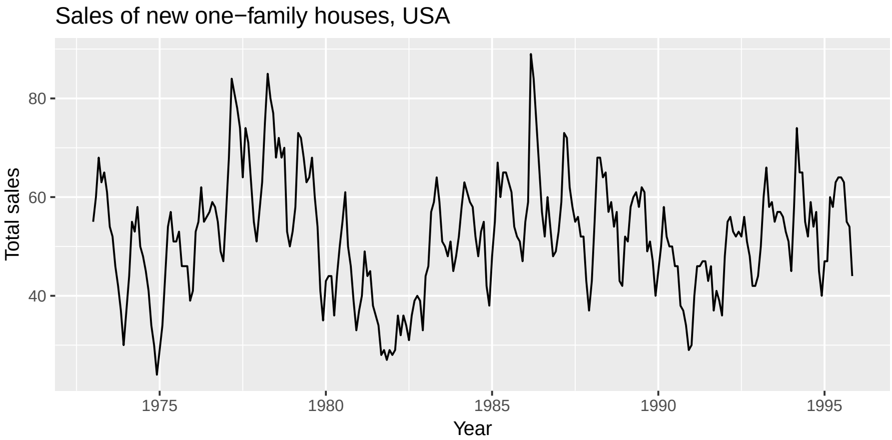


Change in Dow Jones Index
50
0
−50
−100
0
50
100
150
200
250
300
Day

Stationary?

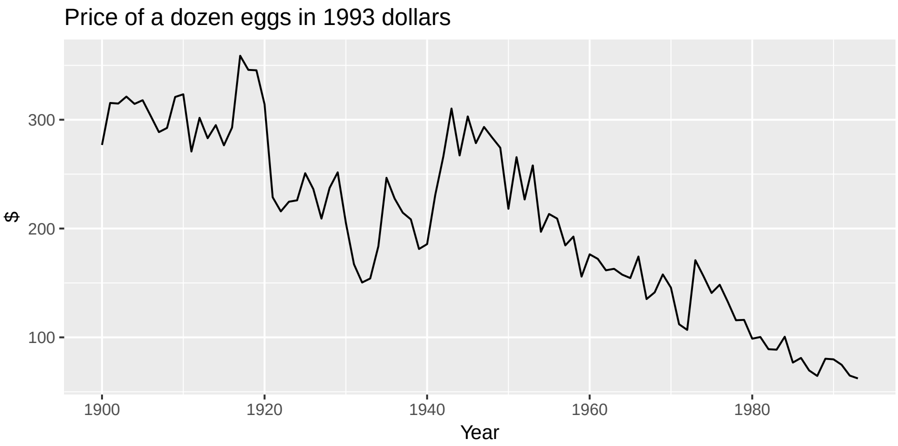

6000
5000
4000
Number of strikes
1950
1955
1960
1965
1970
1975
1980
Year

Stationary?

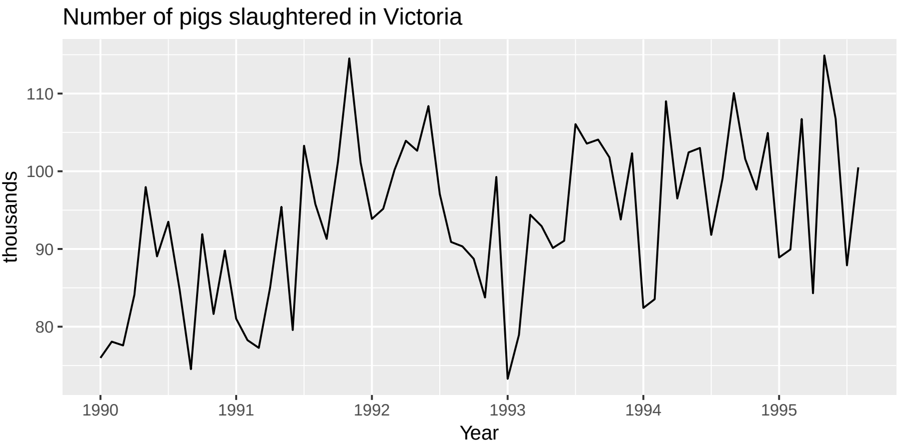


Sales of new one−family houses, USA
Total sales
80
60
40
1975
1980
1985
1990
1995
Year

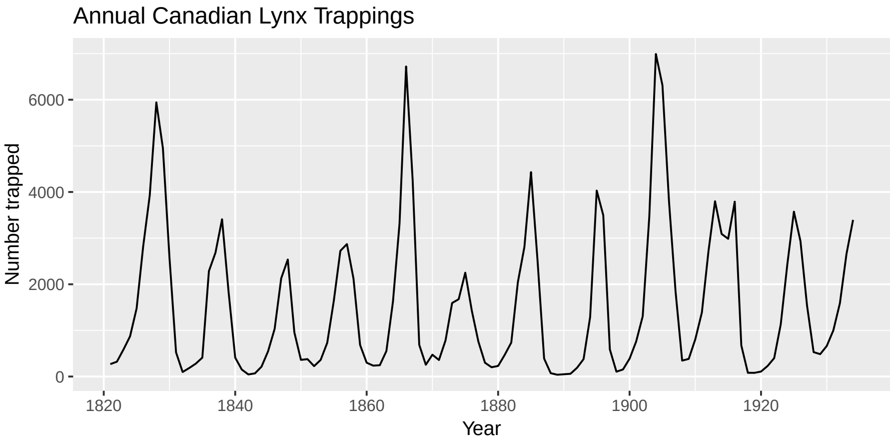
Stationary?


Price of a dozen eggs in 1993 dollars
300
200
100
$
1900
1920
1940
1960
1980
Year

Stationary?

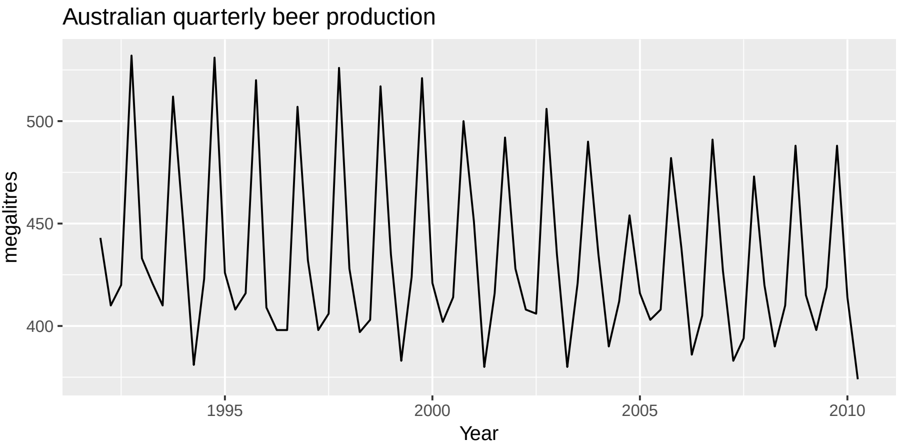

Number of pigs slaughtered in Victoria
110
100
90
80
thousands
1990
1991
1992
1993
1994
1995
Year

Stationary?



Annual Canadian Lynx Trappings
Number trapped
6000
4000
2000
0
1820
1840
1860
1880
1900
1920
Year

Stationary?


Australian quarterly beer production
megalitres
500
450
400
1995
2000
2005
2010
Year

Stationarity

Definition

If $\{y_t\}$ is a stationary time series, then for all s, the distribution of $(y_t, \dots, y_{t+s})$ does not depend on t.

Stationarity

Definition

If $\{y_t\}$ is a stationary time series, then for all s, the distribution of $(y_t, \ldots, y_{t+s})$ does not depend on t.

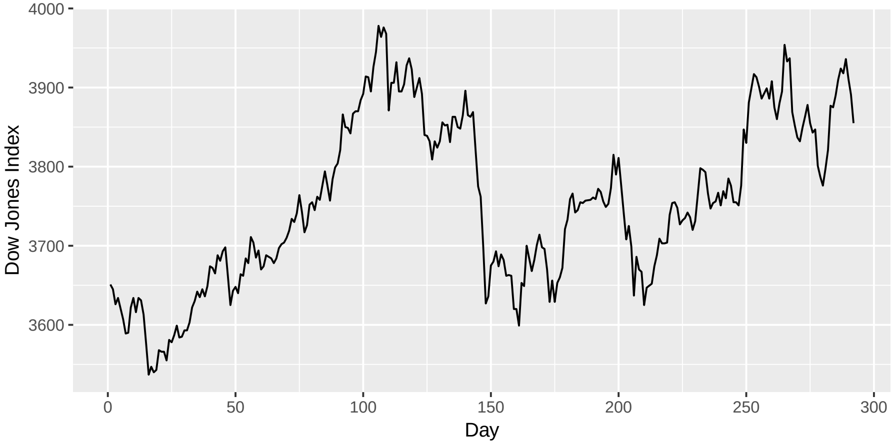
Transformations help to **stabilize the variance**.
For ARIMA modelling, we also need to **stabilize the mean**.

Non-stationarity in the mean

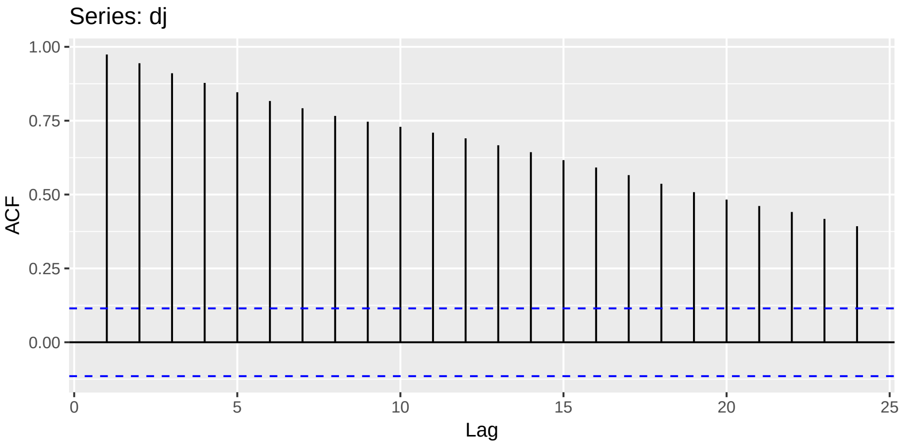
Identifying non-stationary series

- time plot.
- The ACF of stationary data drops to zero relatively quickly
- The ACF of non-stationary data decreases slowly.
- For non-stationary data, the value of r_1 is often large and positive.

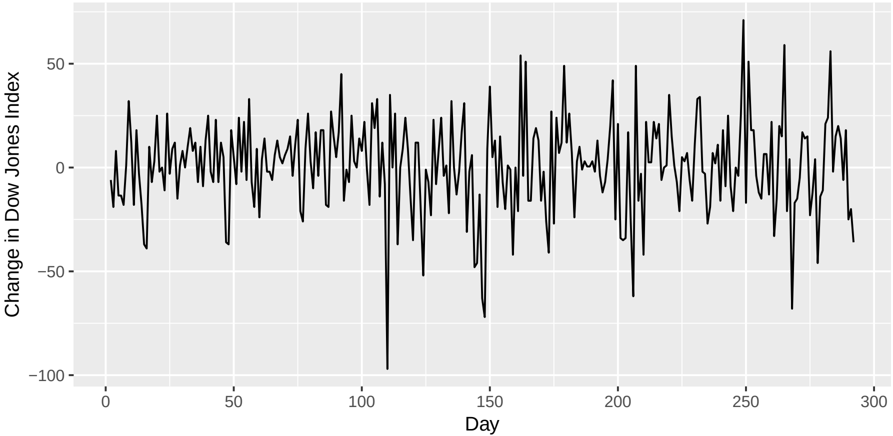
Example: Dow-Jones index



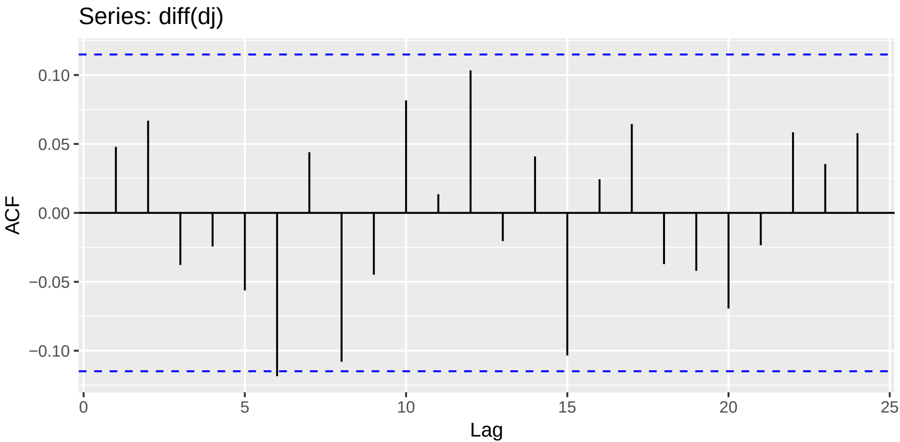
Example: Dow-Jones index



Example: Dow-Jones index



Example: Dow-Jones index



Differencing

- Differencing helps to **stabilize the mean**.
- The differenced series is the *change* between each observation in the original series: $y'_t = y_t - y_{t-1}$.
- The differenced series will have only $T - 1$ values since it is not possible to calculate a difference y'_1 for the first observation.

Seasonal differencing

A seasonal difference is the difference between an observation and the corresponding observation from the previous year.

Seasonal differencing

A seasonal difference is the difference between an observation and the corresponding observation from the previous year.

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

where $m =$ number of seasons.

Seasonal differencing

A seasonal difference is the difference between an observation and the corresponding observation from the previous year.

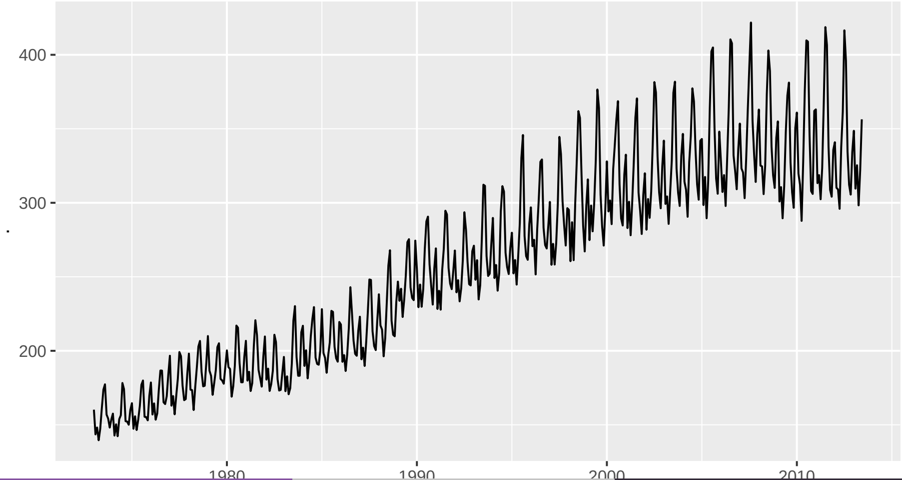
$$y'_t = y_t - y_{t-m}$$

where $m =$ number of seasons.

- For monthly data $m = 12$.
- For quarterly data $m = 4$.

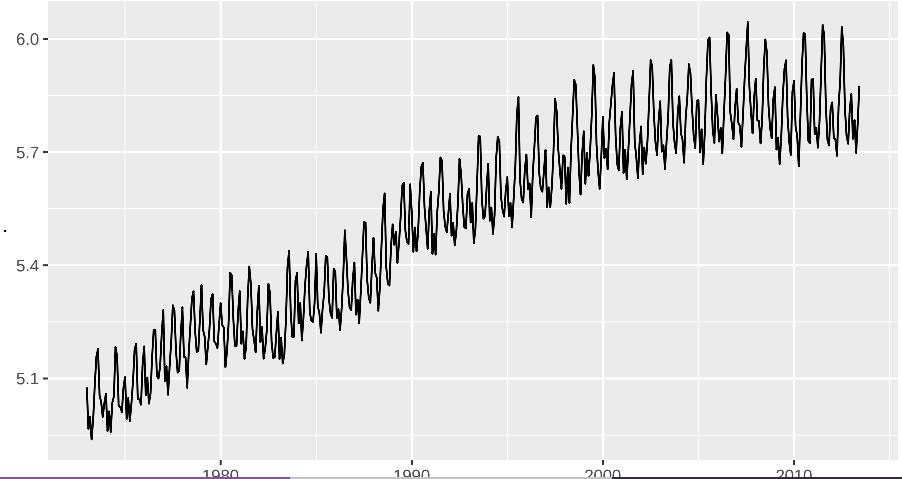
Electricity production

```
usmelec %>% autoplot()
```



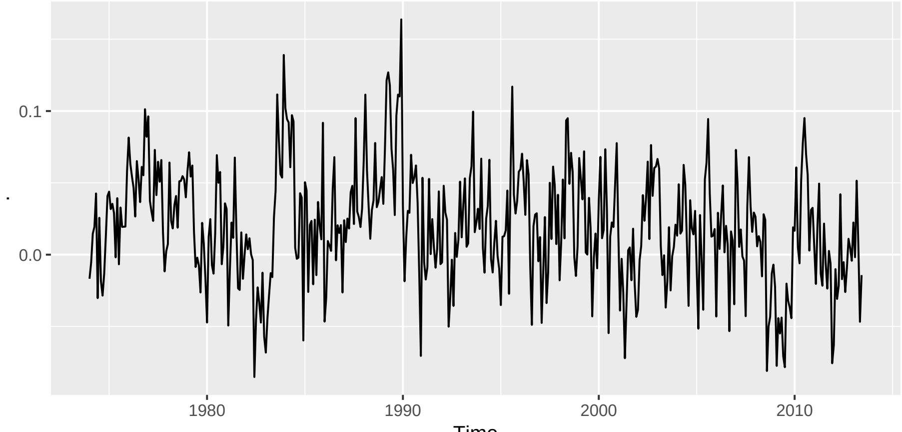
Electricity production

```
usmelec %>% log() %>% autoplot()
```



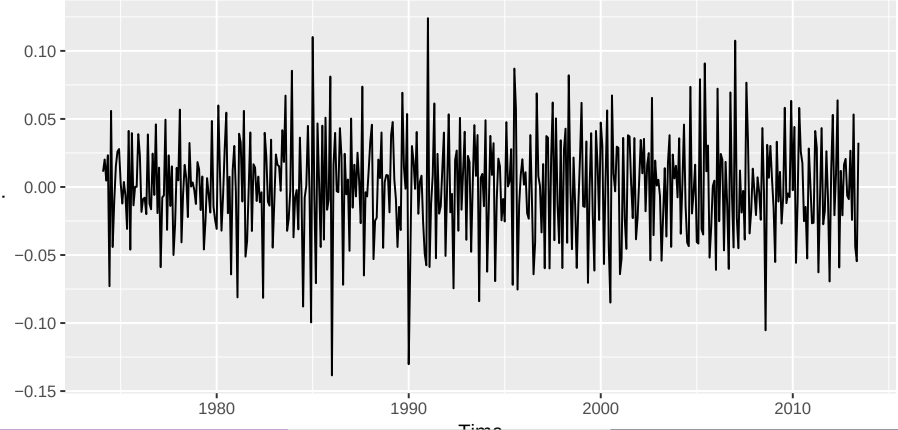
Electricity production

```
usmelec %>% log() %>% diff(lag=12) %>%
  autoplot()
```



Electricity production

```
usmelec %>% log() %>% diff(lag=12) %>%
  diff(lag=1) %>% autoplot()
```



Electricity production

- Seasonally differenced series is closer to being stationary.
- Remaining non-stationarity can be removed with further first difference.

If $y'_t = y_t - y_{t-12}$ denotes seasonally differenced series, then twice-differenced series i

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

Seasonal differencing

When both seasonal and first differences are applied…

Seasonal differencing

When both seasonal and first differences are applied. . .

- it makes no difference which is done first—the result will be the same.
- If seasonality is strong, we recommend that seasonal differencing be done first because sometimes the resulting series will be stationary and there will be no need for further first difference.

Seasonal differencing

When both seasonal and first differences are applied...

- it makes no difference which is done first—the result will be the same.
- If seasonality is strong, we recommend that seasonal differencing be done first because sometimes the resulting series will be stationary and there will be no need for further first difference.

It is important that if differencing is used, the differences are interpretable.

Interpretation of differencing

- first differences are the change between **one observation and the next**;
- seasonal differences are the change between **one year to the next**.

Interpretation of differencing

- first differences are the change between **one observation and the next**;
- seasonal differences are the change between **one year to the next**.

But taking lag 3 differences for yearly data, for example, results in a model which cannot be sensibly interpreted.

Your turn

For the `visitors` series, find an appropriate differencing (after transformation if necessary) to obtain stationary data.

Outline

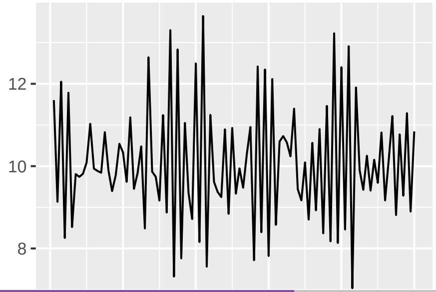
Autoregressive models

Autoregressive (AR) models:

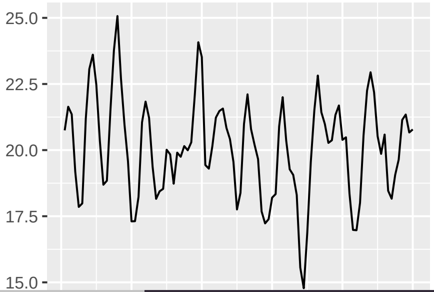
$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t,$$

where ε_t is white noise. This is a multiple regression with **lagged values** of y_t as predictors.

AR(1)



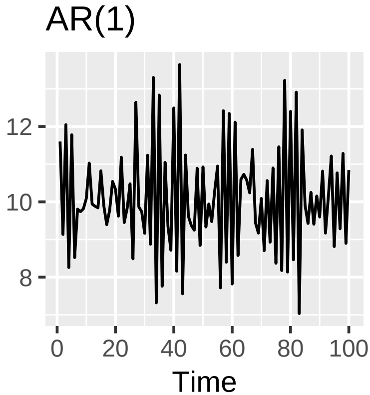
AR(2)



AR(1) model

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

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



AR(1) model

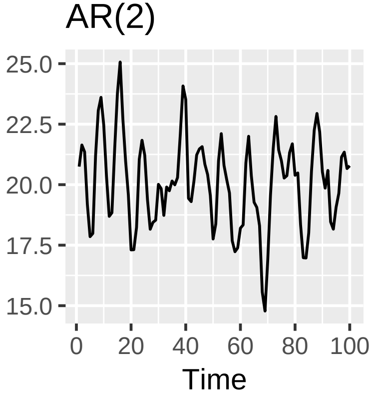
$$y_t = c + \phi_1 y_{t-1} + \varepsilon_t$$

- When $\phi_1 = 0$, y_t is **equivalent to WN**
- When $\phi_1 = 1$ and $c = 0$, y_t is **equivalent to a RW**
- When $\phi_1 = 1$ and $c \neq 0$, y_t is **equivalent to a RW with drift**
- When $\phi_1 < 0$, y_t tends to **oscillate between positive and negative values**.

AR(2) model

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

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

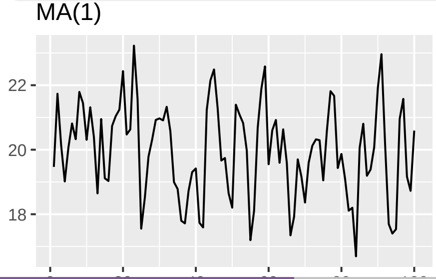


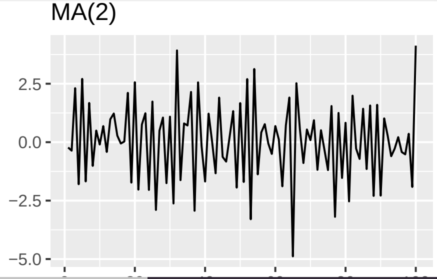
Moving Average (MA) models

Moving Average (MA) models:

$$y_t = c + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q},$$

where ε_t is white noise. This is a multiple regression with **past *errors*** as predictors. *Don't confuse this with moving average smoothing!*

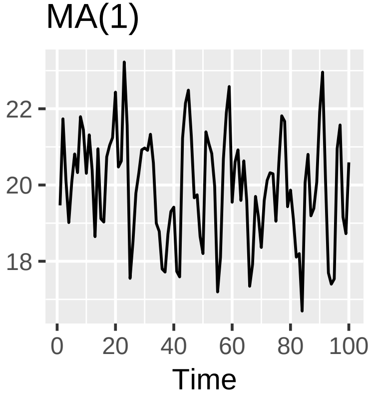




MA(1) model

$$y_t = 20 + \varepsilon_t + 0.8\varepsilon_{t-1}$$

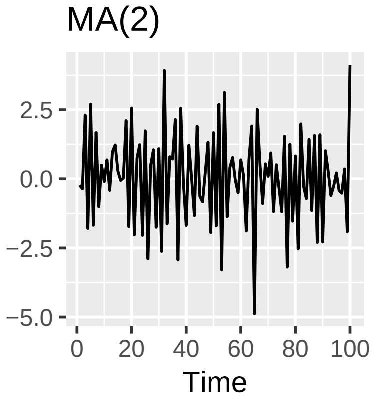
$$\varepsilon_t \sim N(0,1), \quad T = 100.$$



MA(2) model

$$y_t = \varepsilon_t - \varepsilon_{t-1} + 0.8\varepsilon_{t-2}$$

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



ARIMA models

Autoregressive Moving Average models:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t.$$

ARIMA models

Autoregressive Moving Average models:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t.$$

- Predictors include both **lagged values of y_t and lagged errors.**
- Conditions on coefficients ensure stationarity.
- Conditions on coefficients ensure invertibility.

ARIMA models

Autoregressive Moving Average models:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t.$$

- Predictors include both **lagged values of y_t and lagged errors.**
- Conditions on coefficients ensure stationarity.
- Conditions on coefficients ensure invertibility.

Autoregressive Integrated Moving Average models

- Combine ARMA model with **differencing**.

ARIMA models

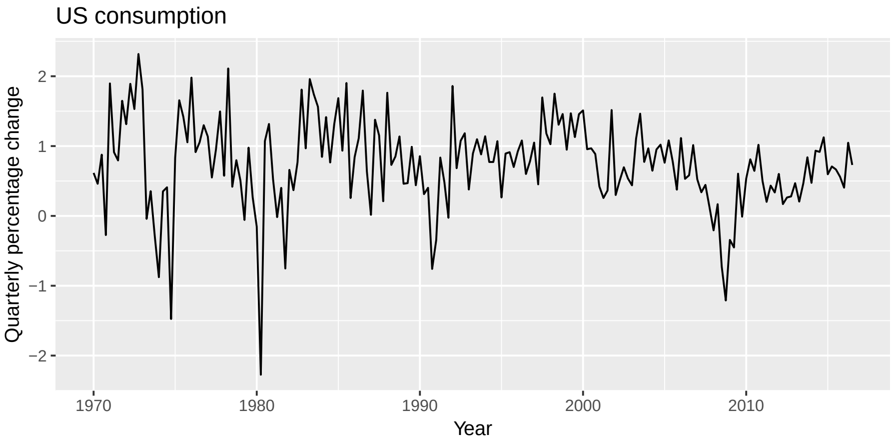
Autoregressive Integrated Moving Average models

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)

US personal consumption



US personal consumption

```r
(fit <- auto.arima(uschange[,"Consumption"]))
```

```
## Series: uschange[, "Consumption"]
## ARIMA(2,0,2) with non-zero mean
##
## Coefficients:
##          ar1      ar2      ma1     ma2    mean
##       1.3908  -0.5813  -1.1800  0.5584  0.7463
## s.e.  0.2553   0.2078   0.2381  0.1403  0.0845
##
## sigma^2 estimated as 0.3511:  log likelihood=-165.14
## AIC=342.28   AICc=342.75   BIC=361.67
```

US personal consumption

```
(fit <- auto.arima(uschange[,"Consumption"]))
```

```
## Series: uschange[, "Consumption"]
## ARIMA(2,0,2) with non-zero mean
##
## Coefficients:
##          ar1      ar2      ma1     ma2    mean
##       1.3908  -0.5813  -1.1800  0.5584  0.7463
## s.e.  0.2553   0.2078   0.2381  0.1403  0.0845
##
## sigma^2 estimated as 0.3511:  log likelihood=-165.14
## AIC=342.28   AICc=342.75   BIC=361.67
```

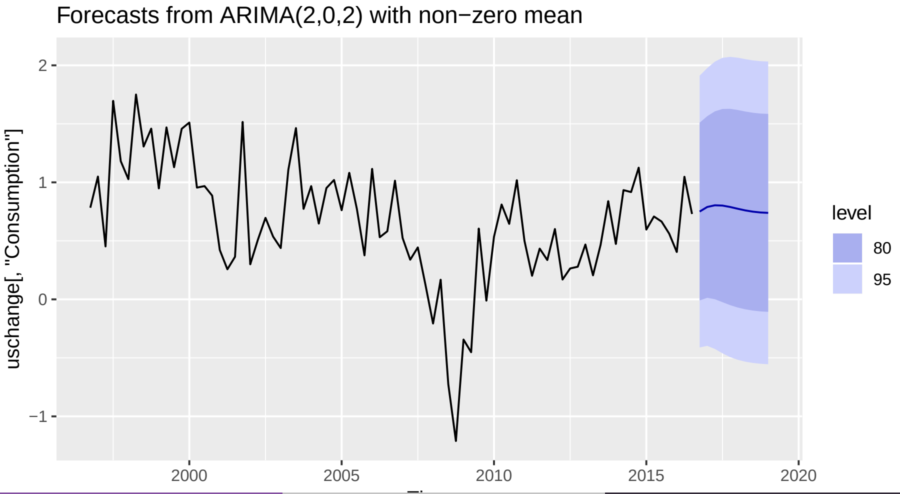
ARIMA(2,0,2) model:

$$y_t = c + 1.391y_{t-1} - 0.581y_{t-2} - 1.180\varepsilon_{t-1} + 0.558\varepsilon_{t-2} + \varepsilon_t,$$

where $c = 0.746 \times (1 - 1.391 + 0.581) = 0.142$ and ε_t is white noise with a standard deviation of $0.593 = \sqrt{0.351}$.

US personal consumption

```
fit %>% forecast(h=10) %>% autoplot(include=80)
```



Outline

Partial autocorrelations

Partial autocorrelations measure relationship between y_t and y_{t-k}, when the effects of other time lags — $1, 2, 3, \ldots, k-1$ — are removed.

Partial autocorrelations

Partial autocorrelations measure relationship between y_t and y_{t-k}, when the effects of other time lags — $1, 2, 3, \dots, k-1$ — are removed.

$$\begin{aligned}\alpha_k &= k\text{th partial autocorrelation coefficient} \\ &= \text{equal to the estimate of } b_k \text{ in regression:} \\ &\quad y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_k y_{t-k}.\end{aligned}$$

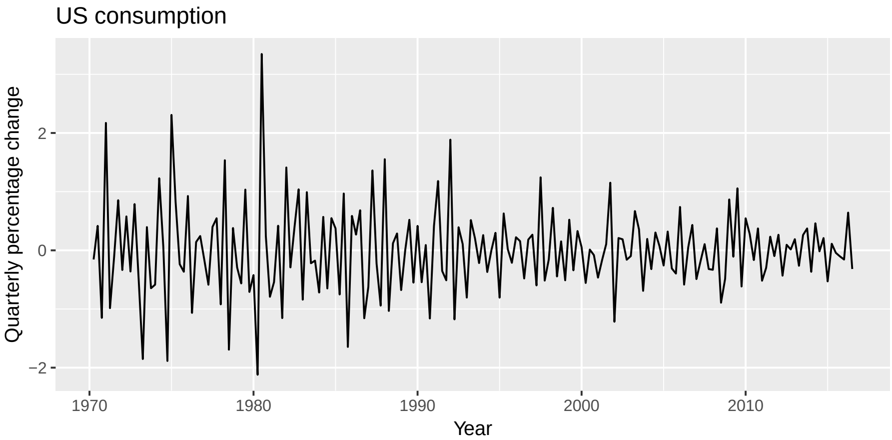
Partial autocorrelations

Partial autocorrelations measure relationship between y_t and y_{t-k}, when the effects of other time lags — $1, 2, 3, \dots, k-1$ — are removed.

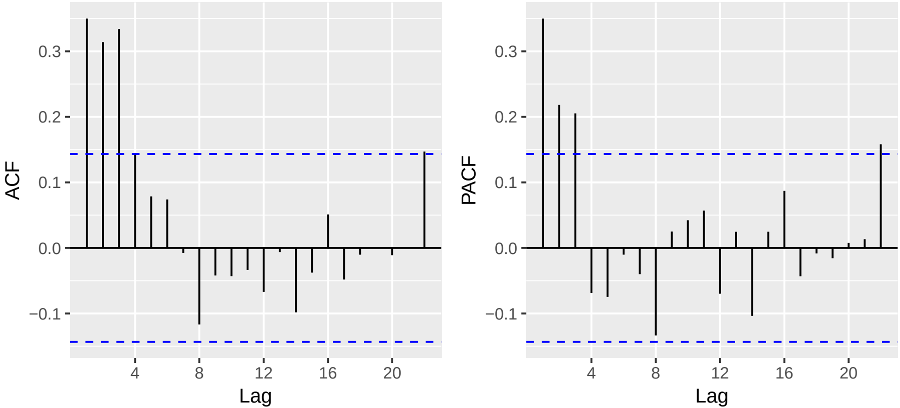
$$\begin{aligned}\alpha_k &= k\text{th partial autocorrelation coefficient}\\ &= \text{equal to the estimate of } b_k \text{ in regression:}\\ &\qquad y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_k y_{t-k}.\end{aligned}$$

- Varying number of terms on RHS gives α_k for different values of k.
- There are more efficient ways of calculating α_k.
- $\alpha_1 = \rho_1$

Example: US consumption



Example: US consumption



ACF and PACF interpretation

AR(1)

$$\rho_k = \phi_1^k \qquad \text{for } k = 1, 2, \ldots;$$
$$\alpha_1 = \phi_1 \qquad \alpha_k = 0 \qquad \text{for } k = 2, 3, \ldots.$$

So we have an AR(1) model when

- autocorrelations exponentially decay
- there is a single significant partial autocorrelation.

ACF and PACF interpretation

AR(p)

- ACF dies out in an exponential or damped sine-wave manner
- PACF has all zero spikes beyond the pth spike

So we have an AR(p) model when

- the ACF is exponentially decaying or sinusoidal
- there is a significant spike at lag p in PACF, but none beyond p

ACF and PACF interpretation

MA(1)

$$\rho_1 = \theta_1 \qquad \rho_k = 0 \qquad \text{for } k = 2, 3, \ldots;$$
$$\alpha_k = -(-\theta_1)^k$$

So we have an MA(1) model when

- the PACF is exponentially decaying and
- there is a single significant spike in ACF

ACF and PACF interpretation

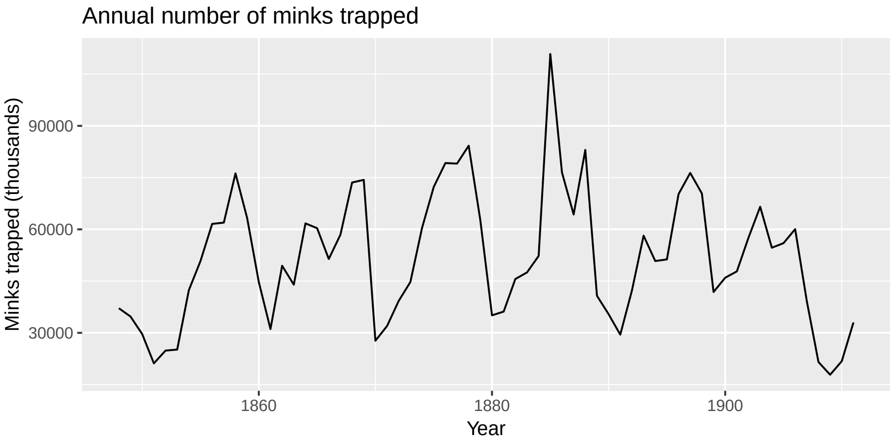
MA(q)

- PACF dies out in an exponential or damped sine-wave manner
- ACF has all zero spikes beyond the qth spike

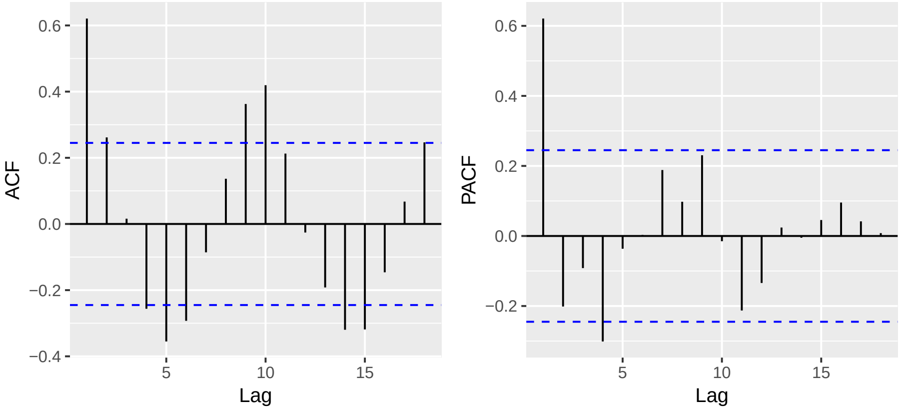
So we have an MA(q) model when

- the PACF is exponentially decaying or sinusoidal
- there is a significant spike at lag q in ACF, but none beyond q

Example: Mink trapping



Example: Mink trapping



Information criteria

Akaike's Information Criterion (AIC):

$$\mathrm{AIC} = -2\log(L) + 2(p + q + k + 1),$$

where L is the likelihood of the data,
$k = 1$ if $c \neq 0$ and $k = 0$ if $c = 0$.

Information criteria

Akaike's Information Criterion (AIC):

$$\text{AIC} = -2\log(L) + 2(p+q+k+1),$$

where L is the likelihood of the data,
$k = 1$ if $c \neq 0$ and $k = 0$ if $c = 0$.

Corrected AIC:

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

Information criteria

Akaike's Information Criterion (AIC):

$$\text{AIC} = -2\log(L) + 2(p+q+k+1),$$

where L is the likelihood of the data,
$k=1$ if $c \neq 0$ and $k=0$ if $c=0$.

Corrected AIC:

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

Bayesian Information Criterion:

$$\text{BIC} = \text{AIC} + [\log(T)-2](p+q+k-1).$$

Information criteria

Akaike's Information Criterion (AIC):

$$\text{AIC} = -2\log(L) + 2(p+q+k+1),$$

where L is the likelihood of the data,
$k = 1$ if $c \neq 0$ and $k = 0$ if $c = 0$.

Corrected AIC:

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

Bayesian Information Criterion:

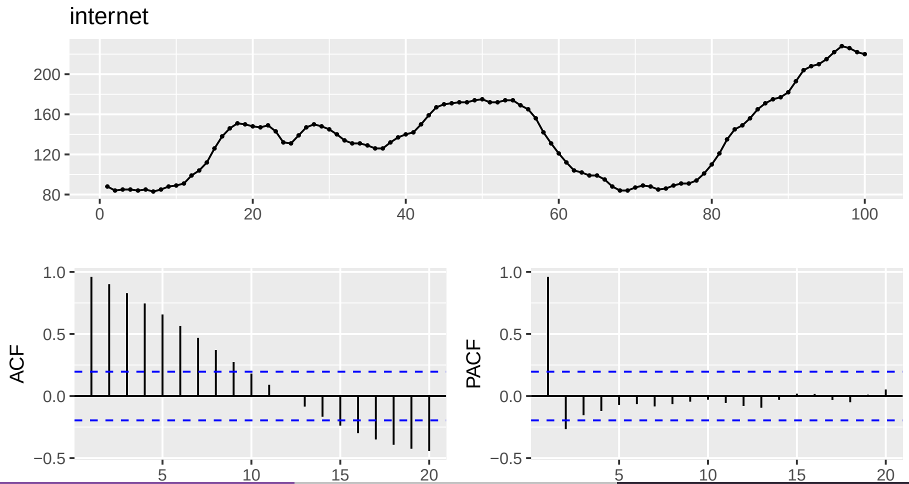
$$\text{BIC} = \text{AIC} + [\log(T) - 2](p+q+k-1).$$

Good models are obtained by minimizing either the AIC, AICc or BIC. Our preference is to use the AICc.

Outline

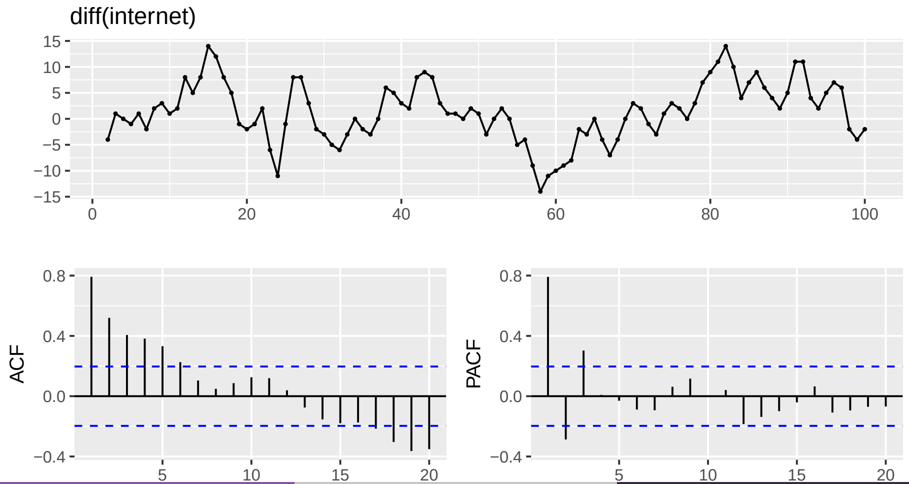
Choosing your own model

```
ggtsdisplay(internet)
```



Choosing your own model

```
ggtsdisplay(diff(internet))
```



Choosing your own model

```r
(fit <- Arima(internet,order=c(3,1,0)))
```

```
## Series: internet
## ARIMA(3,1,0)
##
## Coefficients:
##          ar1      ar2     ar3
##       1.1513  -0.6612  0.3407
## s.e.  0.0950   0.1353  0.0941
##
## sigma^2 estimated as 9.656:  log likelihood=-252
## AIC=511.99   AICc=512.42   BIC=522.37
```

Choosing your own model

```r
auto.arima(internet)
```

```
## Series: internet
## ARIMA(1,1,1)
##
## Coefficients:
##          ar1     ma1
##       0.6504  0.5256
## s.e.  0.0842  0.0896
##
## sigma^2 estimated as 9.995:  log likelihood=-254.1
## AIC=514.3   AICc=514.55   BIC=522.08
```

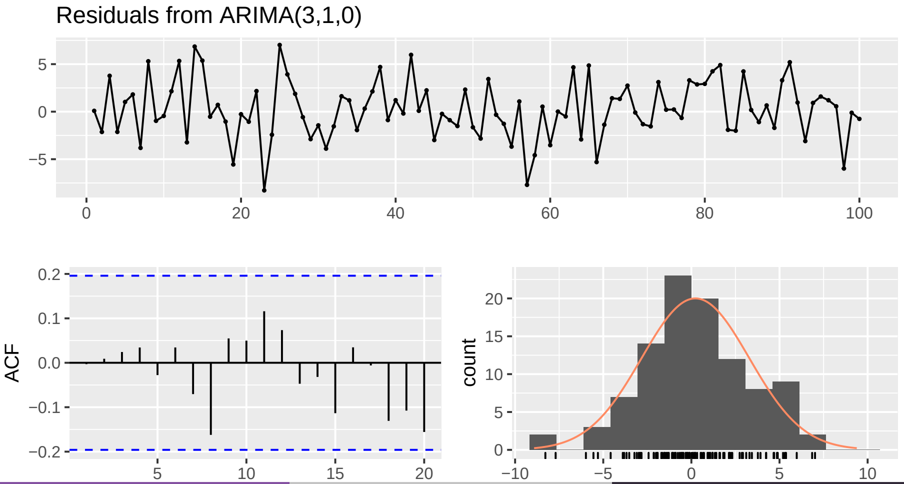
Choosing your own model

```r
auto.arima(internet, stepwise=FALSE,
  approximation=FALSE)
```

```
## Series: internet
## ARIMA(3,1,0)
##
## Coefficients:
##          ar1      ar2     ar3
##       1.1513  -0.6612  0.3407
## s.e.  0.0950   0.1353  0.0941
##
## sigma^2 estimated as 9.656:  log likelihood=-252
## AIC=511.99   AICc=512.42   BIC=522.37
```

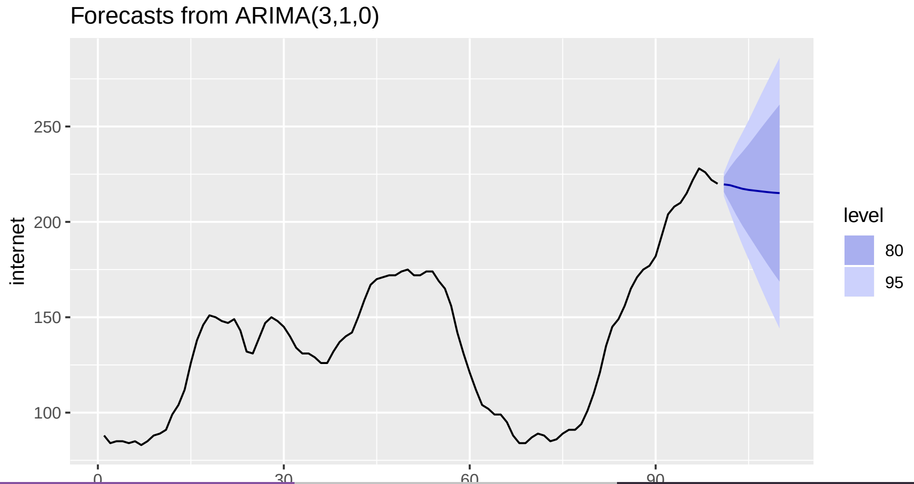
Choosing your own model

```r
checkresiduals(fit)
```



Choosing your own model

```
fit %>% forecast %>% autoplot
```



Modelling procedure with `Arima`

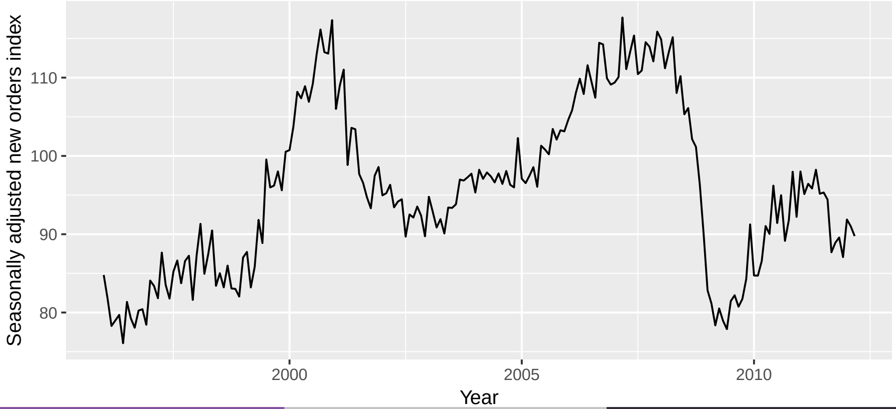
1. Plot the data. Identify any unusual observations.
2. If necessary, transform the data (using a Box-Cox transformation) to stabilize the variance.
3. If the data are non-stationary: take first differences of the data until the data are stationary.
4. Examine the ACF/PACF: Is an AR(p) or MA(q) model appropriate?
5. Try your chosen model(s), and use the AICc to search for a better model.
6. Check the residuals from your chosen model by plotting the ACF of the residuals, and doing a portmanteau test of the residuals. If they do not look like white noise, try a modified model.
7. Once the residuals look like white noise, calculate forecasts.

Modelling procedure with `auto.arima`

1. Plot the data. Identify any unusual observations.
2. If necessary, transform the data (using a Box-Cox transformation) to stabilize the variance.
3. Use `auto.arima` to select a model.
6. Check the residuals from your chosen model by plotting the ACF of the residuals, and doing a portmanteau test of the residuals. If they do not look like white noise, try a modified model.
7. Once the residuals look like white noise, calculate forecasts.

Seasonally adjusted electrical equipment

```
eeadj <- seasadj(stl(elecequip, s.window="periodic"))
autoplot(eeadj) + xlab("Year") +
  ylab("Seasonally adjusted new orders index")
```

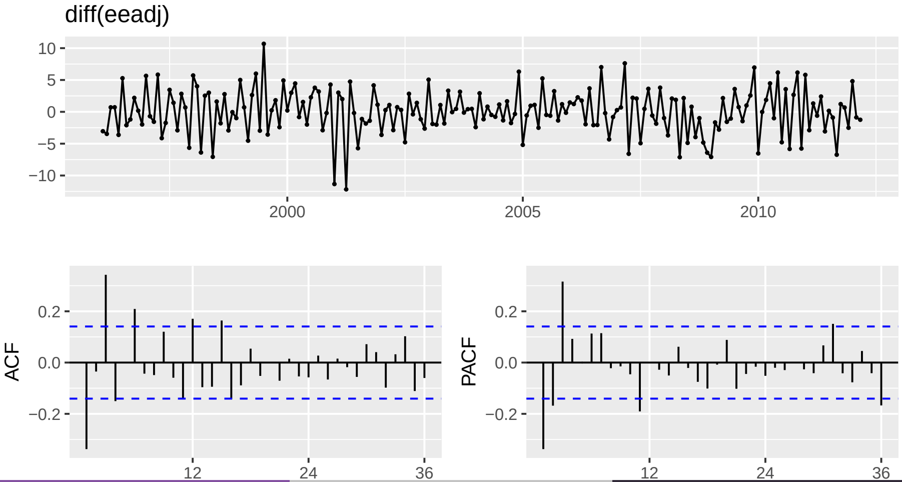


Seasonally adjusted electrical equipment

1. Time plot shows sudden changes, particularly big drop in 2008/2009 due to global economic environment. Otherwise nothing unusual and no need for data adjustments.
2. No evidence of changing variance, so no Box-Cox transformation.
3. Data are clearly non-stationary, so we take first differences.

Seasonally adjusted electrical equipment

```r
ggtsdisplay(diff(eeadj))
```



Seasonally adjusted electrical equipment

4. PACF is suggestive of AR(3). So initial candidate model is ARIMA(3,1,0). No other obvious candidates.
5. Fit ARIMA(3,1,0) model along with variations: ARIMA(4,1,0), ARIMA(2,1,0), ARIMA(3,1,1), etc. ARIMA(3,1,1) has smallest AICc value.

Seasonally adjusted electrical equipment

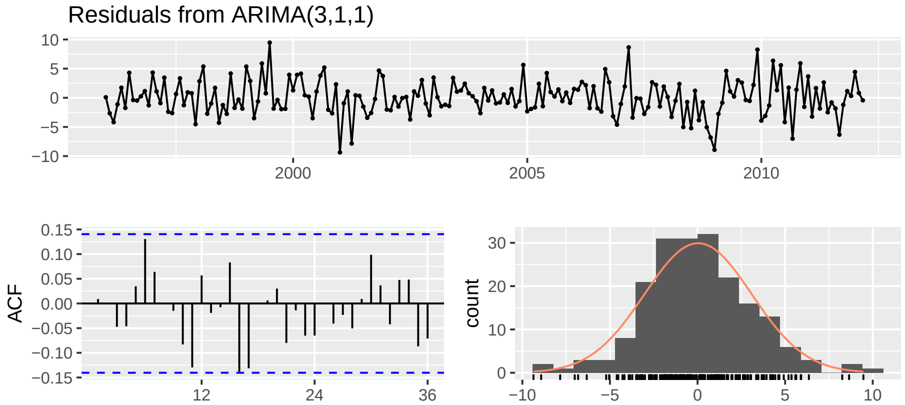
```r
(fit <- Arima(eeadj, order=c(3,1,1)))
```

```
## Series: eeadj
## ARIMA(3,1,1)
##
## Coefficients:
##          ar1     ar2     ar3      ma1
##       0.0044  0.0916  0.3698  -0.3921
## s.e.  0.2201  0.0984  0.0669   0.2426
##
## sigma^2 estimated as 9.577:  log likelihood=-492.69
## AIC=995.38   AICc=995.7   BIC=1011.72
```

Seasonally adjusted electrical equipment

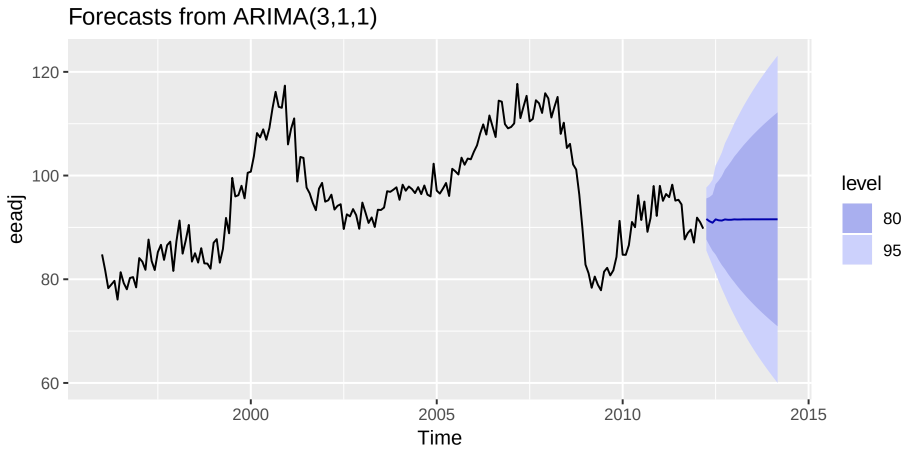
6. ACF plot of residuals from ARIMA(3,1,1) model look like white noise.

```
checkresiduals(fit)
```



Seasonally adjusted electrical equipment

```
fit %>% forecast %>% autoplot
```



Your turn

For the `usgdp` data:

- if necessary, find a suitable Box-Cox transformation for the data;
- fit a suitable ARIMA model to the transformed data using `auto.arima()`;
- check the residual diagnostics;
- produce forecasts of your fitted model. Do the forecasts look reasonable?

Outline

Seasonal ARIMA models

ARIMA	$\underbrace{(p, d, q)}$	$\underbrace{(P, D, Q)_m}$
	$\uparrow$ Non-seasonal part of the model	$\uparrow$ Seasonal part of of the model

where $m =$ number of observations per year.

Common ARIMA models

The US Census Bureau uses the following models most often:

ARIMA(0,1,1)(0,1,1)$_m$	with log transformation
ARIMA(0,1,2)(0,1,1)$_m$	with log transformation
ARIMA(2,1,0)(0,1,1)$_m$	with log transformation
ARIMA(0,2,2)(0,1,1)$_m$	with log transformation
ARIMA(2,1,2)(0,1,1)$_m$	with no transformation

Seasonal ARIMA models

The seasonal part of an AR or MA model will be seen in the seasonal lags of the PACF and ACF.

ARIMA(0,0,0)(0,0,1)$_{12}$ will show:

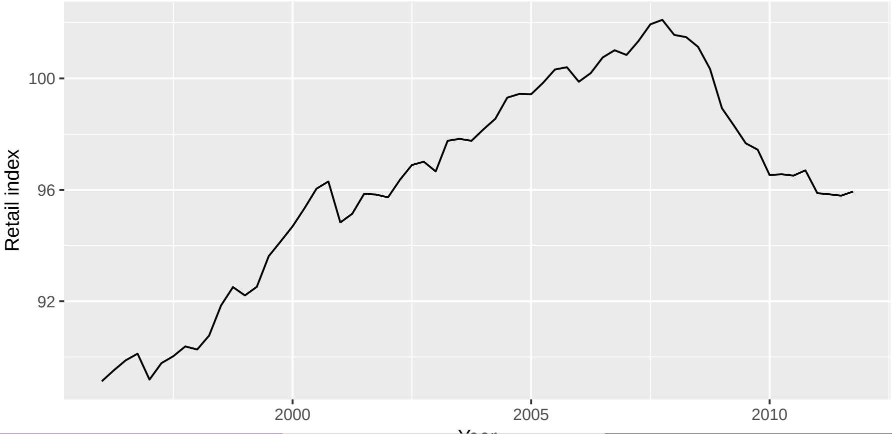
- a spike at lag 12 in the ACF but no other significant spikes.
- The PACF will show exponential decay in the seasonal lags; that is, at lags 12, 24, 36, ...

ARIMA(0,0,0)(1,0,0)$_{12}$ will show:

- exponential decay in the seasonal lags of the ACF
- a single significant spike at lag 12 in the PACF.

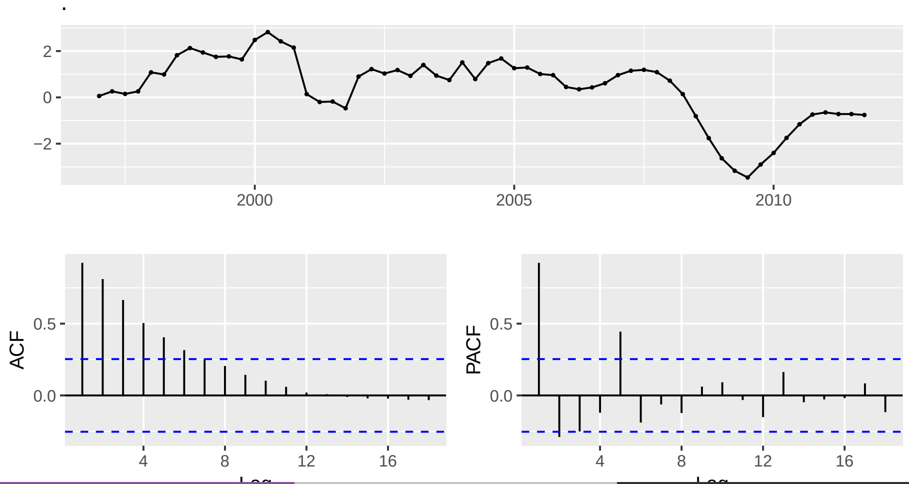
European quarterly retail trade

```r
autoplot(euretail) +
  xlab("Year") + ylab("Retail index")
```



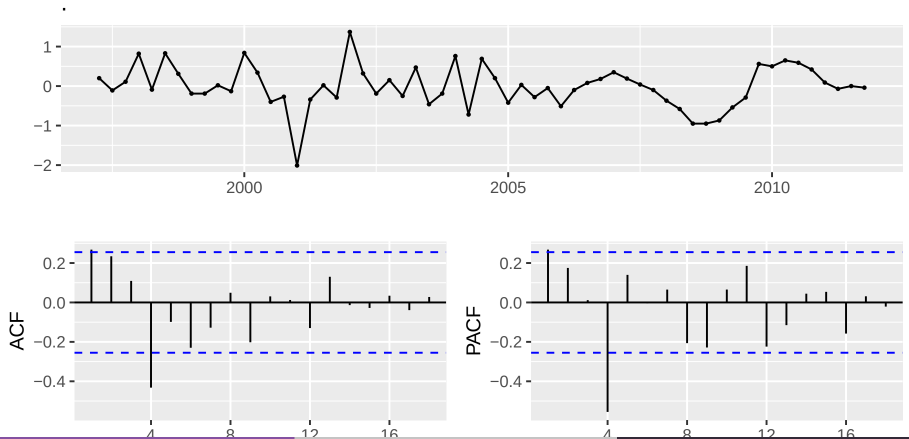
European quarterly retail trade

```
euretail %>% diff(lag=4) %>% ggtsdisplay()
```



European quarterly retail trade

```
euretail %>% diff(lag=4) %>% diff() %>%
  ggtsdisplay()
```

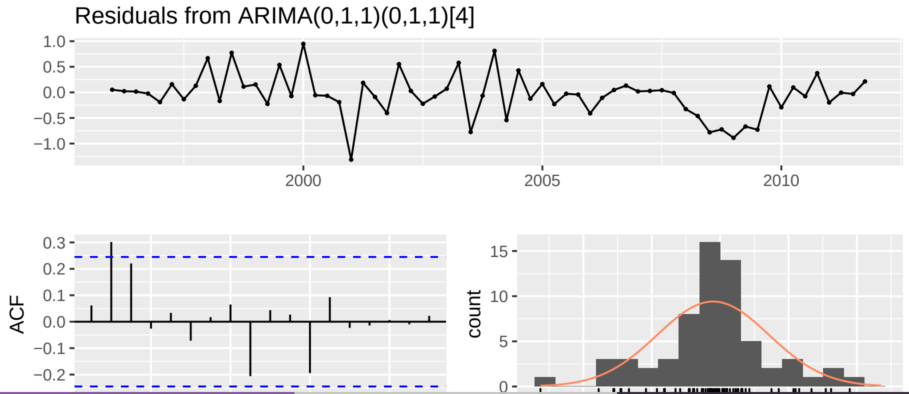


European quarterly retail trade

- $d = 1$ and $D = 1$ seems necessary.
- Significant spike at lag 1 in ACF suggests non-seasonal MA(1) component.
- Significant spike at lag 4 in ACF suggests seasonal MA(1) component.
- Initial candidate model: ARIMA(0,1,1)$(0,1,1)_4$.
- We could also have started with ARIMA(1,1,0)$(1,1,0)_4$.

European quarterly retail trade

```r
fit <- Arima(euretail, order=c(0,1,1),
  seasonal=c(0,1,1))
checkresiduals(fit)
```

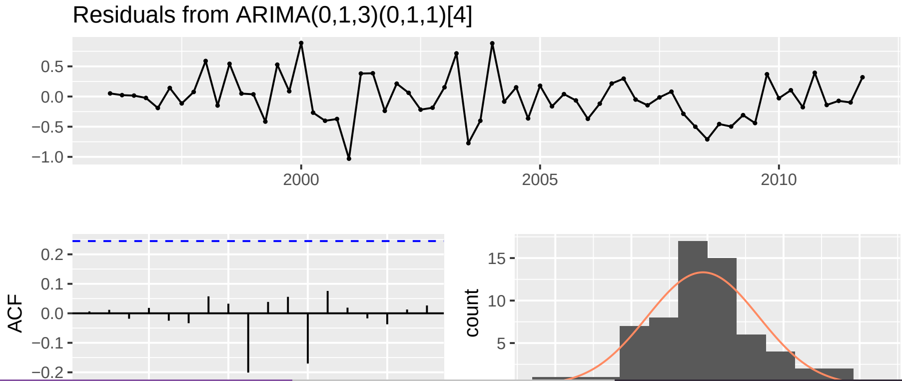


European quarterly retail trade

- ACF and PACF of residuals show significant spikes at lag 2, and maybe lag 3.
- AICc of ARIMA$(0,1,2)(0,1,1)_4$ model is 74.27.
- AICc of ARIMA$(0,1,3)(0,1,1)_4$ model is 68.39.

European quarterly retail trade

- ACF and PACF of residuals show significant spikes at lag 2, and maybe lag 3.
- AICc of ARIMA$(0,1,2)(0,1,1)_4$ model is 74.27.
- AICc of ARIMA$(0,1,3)(0,1,1)_4$ model is 68.39.

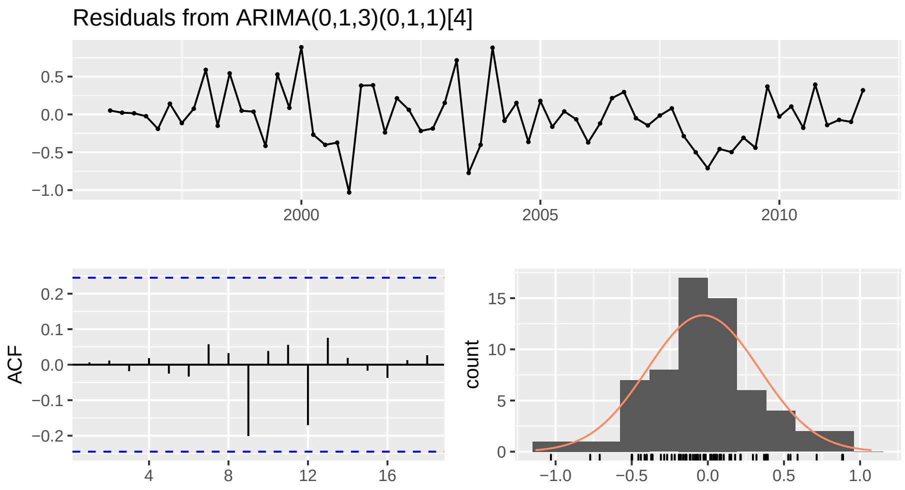


European quarterly retail trade

```
## Series: euretail
## ARIMA(0,1,3)(0,1,1)[4]
##
## Coefficients:
##          ma1     ma2     ma3     sma1
##       0.2630  0.3694  0.4200  -0.6636
## s.e.  0.1237  0.1255  0.1294   0.1545
##
## sigma^2 estimated as 0.156:  log likelihood=-28.63
## AIC=67.26   AICc=68.39   BIC=77.65
```

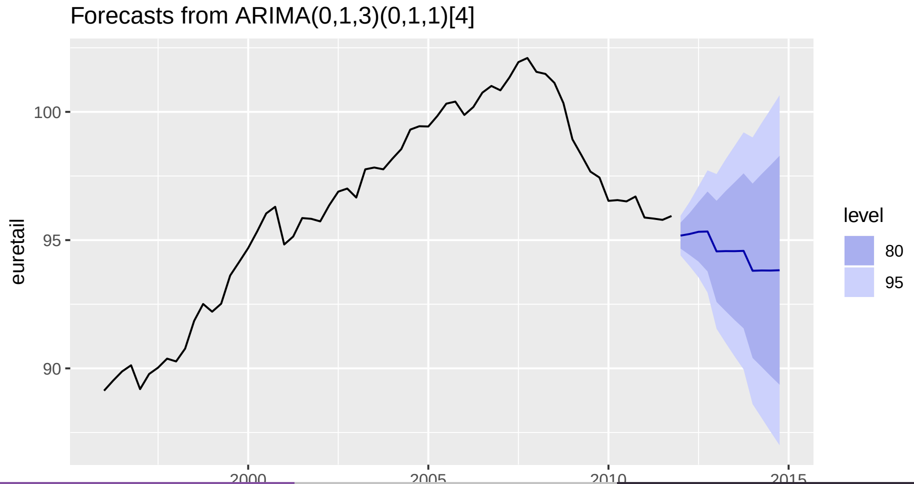
European quarterly retail trade

```r
checkresiduals(fit)
```



European quarterly retail trade

```
autoplot(forecast(fit, h=12))
```



European quarterly retail trade

```r
auto.arima(euretail)
```

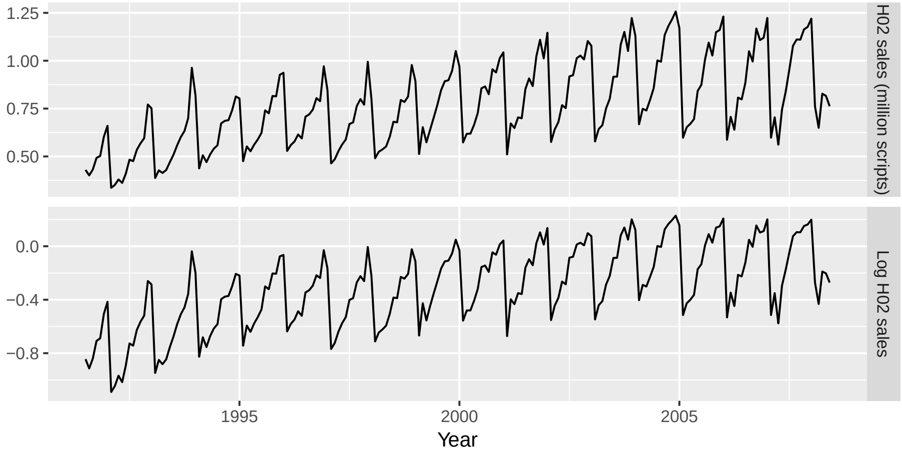
```
## Series: euretail
## ARIMA(1,1,2)(0,1,1)[4]
##
## Coefficients:
##          ar1      ma1     ma2     sma1
##       0.7362  -0.4663  0.2163  -0.8433
## s.e.  0.2243   0.1990  0.2101   0.1876
##
## sigma^2 estimated as 0.1587:  log likelihood=-29.62
## AIC=69.24   AICc=70.38   BIC=79.63
```

European quarterly retail trade

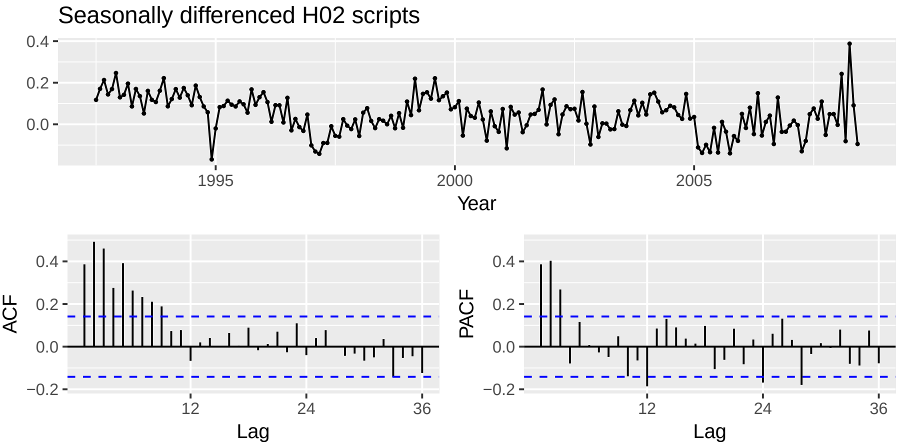
```r
auto.arima(euretail,
  stepwise=FALSE, approximation=FALSE)
```

```
## Series: euretail
## ARIMA(0,1,3)(0,1,1)[4]
##
## Coefficients:
##          ma1     ma2     ma3     sma1
##       0.2630  0.3694  0.4200  -0.6636
## s.e.  0.1237  0.1255  0.1294   0.1545
##
## sigma^2 estimated as 0.156:  log likelihood=-28.63
## AIC=67.26   AICc=68.39   BIC=77.65
```

Cortecosteroid drug sales



Cortecosteroid drug sales



Cortecosteroid drug sales

- Choose $D = 1$ and $d = 0$.
- Spikes in PACF at lags 12 and 24 suggest seasonal AR(2) term.
- Spikes in PACF suggests possible non-seasonal AR(3) term.
- Initial candidate model: ARIMA(3,0,0)(2,1,0)$_{12}$.

Cortecosteroid drug sales

Model	AICc
ARIMA(3,0,1)(0,1,2)$_{12}$	-485.48
ARIMA(3,0,1)(1,1,1)$_{12}$	-484.25
ARIMA(3,0,1)(0,1,1)$_{12}$	-483.67
ARIMA(3,0,1)(2,1,0)$_{12}$	-476.31
ARIMA(3,0,0)(2,1,0)$_{12}$	-475.12
ARIMA(3,0,2)(2,1,0)$_{12}$	-474.88
ARIMA(3,0,1)(1,1,0)$_{12}$	-463.40

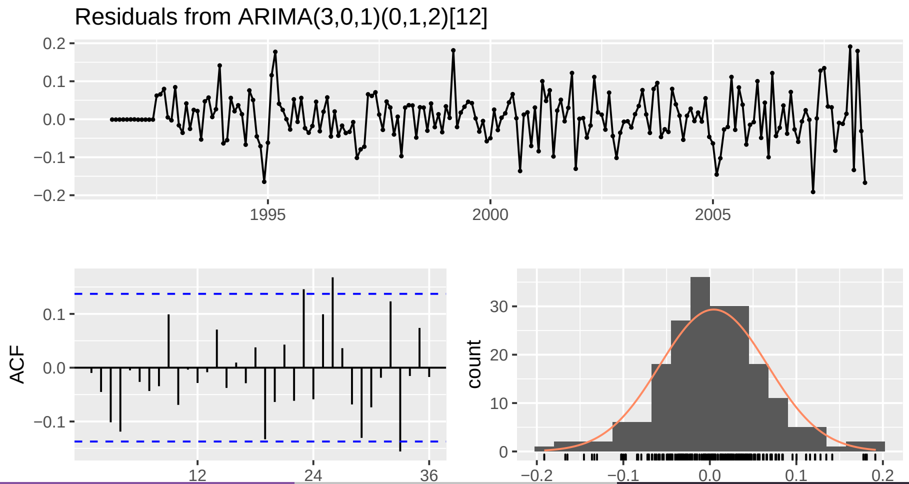
Cortecosteroid drug sales

```r
(fit <- Arima(h02, order=c(3,0,1), seasonal=c(0,1,2),
   lambda=0))
```

```
## Series: h02
## ARIMA(3,0,1)(0,1,2)[12]
## Box Cox transformation: lambda= 0
##
## Coefficients:
##           ar1     ar2     ar3     ma1     sma1     sma2
##       -0.1603  0.5481  0.5678  0.3827  -0.5222  -0.1768
## s.e.   0.1636  0.0878  0.0942  0.1895   0.0861   0.0872
##
## sigma^2 estimated as 0.004278:  log likelihood=250.04
## AIC=-486.08   AICc=-485.48   BIC=-463.28
```

Cortecosteroid drug sales

```
checkresiduals(fit, lag=36)
```



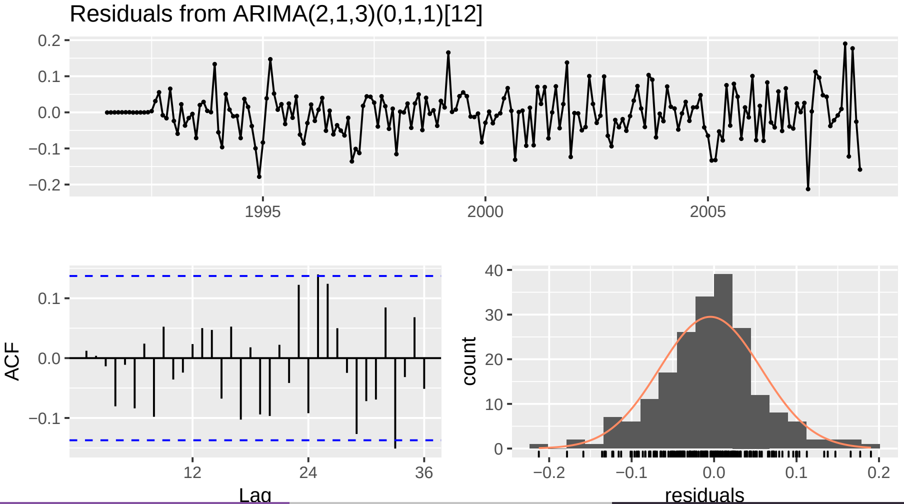
Cortecosteroid drug sales

```r
(fit <- auto.arima(h02, lambda=0))
```

```
## Series: h02
## ARIMA(2,1,3)(0,1,1)[12]
## Box Cox transformation: lambda= 0
##
## Coefficients:
##           ar1      ar2     ma1     ma2      ma3     sma1
##       -1.0194  -0.8351  0.1717  0.2578  -0.4206  -0.6528
## s.e.   0.1648   0.1203  0.2079  0.1177   0.1060   0.0657
##
## sigma^2 estimated as 0.004203:  log likelihood=250.8
## AIC=-487.6   AICc=-486.99   BIC=-464.83
```

Cortecosteroid drug sales

```
checkresiduals(fit, lag=36)
```



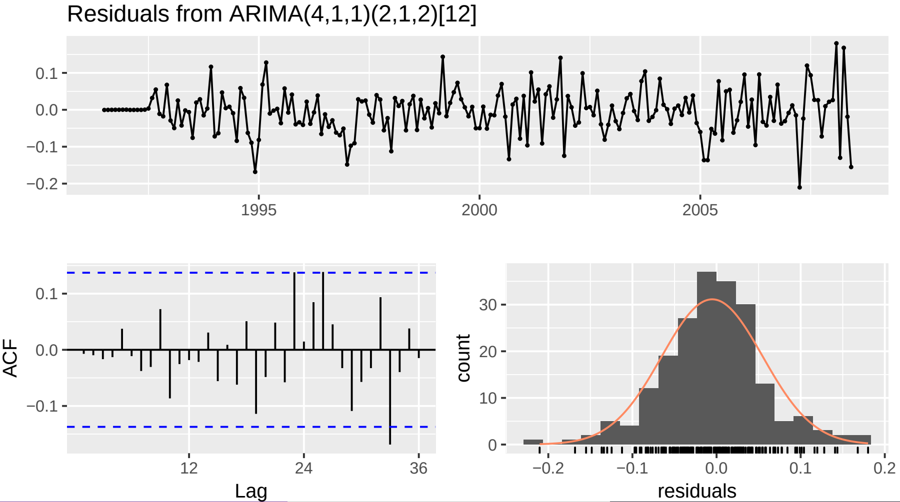
Cortecosteroid drug sales

```
(fit <- auto.arima(h02, lambda=0, max.order=9,
  stepwise=FALSE, approximation=FALSE))
```

```
## Series: h02
## ARIMA(4,1,1)(2,1,2)[12]
## Box Cox transformation: lambda= 0
##
## Coefficients:
##           ar1     ar2     ar3      ar4      ma1    sar1     sar2     sma1
##       -0.0425  0.2098  0.2017  -0.2273  -0.7424  0.6213  -0.3832  -1.2019
## s.e.   0.2167  0.1813  0.1144   0.0810   0.2074  0.2421   0.1185   0.2491
##          sma2
##        0.4959
## s.e.   0.2135
##
## sigma^2 estimated as 0.004049:  log likelihood=254.31
## AIC=-488.63   AICc=-487.4   BIC=-456.1
```

Cortecosteroid drug sales

```
checkresiduals(fit, lag=36)
```



Cortecosteroid drug sales

Training data: July 1991 to June 2006
Test data: July 2006–June 2008

```
getrmse <- function(x,h,...)
{
  train.end <- time(x)[length(x)-h]
  test.start <- time(x)[length(x)-h+1]
  train <- window(x,end=train.end)
  test <- window(x,start=test.start)
  fit <- Arima(train,...)
  fc <- forecast(fit,h=h)
  return(accuracy(fc,test)[2,"RMSE"])
}
getrmse(h02,h=24,order=c(3,0,0),seasonal=c(2,1,0),lambda=0)
getrmse(h02,h=24,order=c(3,0,1),seasonal=c(2,1,0),lambda=0)
getrmse(h02,h=24,order=c(3,0,2),seasonal=c(2,1,0),lambda=0)
getrmse(h02,h=24,order=c(3,0,1),seasonal=c(1,1,0),lambda=0)
getrmse(h02,h=24,order=c(3,0,1),seasonal=c(0,1,1),lambda=0)
```

Cortecosteroid drug sales

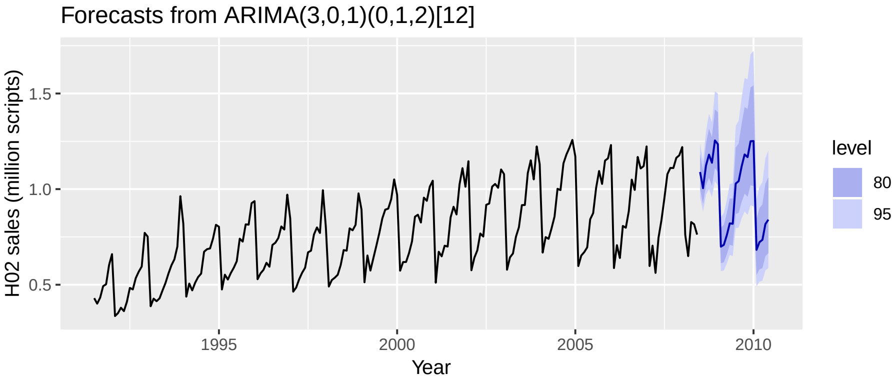
Model	RMSE
ARIMA(4,1,1)(2,1,2)[12]	0.0615
ARIMA(3,0,1)(0,1,2)[12]	0.0622
ARIMA(3,0,1)(1,1,1)[12]	0.0630
ARIMA(2,1,4)(0,1,1)[12]	0.0632
ARIMA(2,1,3)(0,1,1)[12]	0.0634
ARIMA(3,0,3)(0,1,1)[12]	0.0639
ARIMA(2,1,5)(0,1,1)[12]	0.0640
ARIMA(3,0,1)(0,1,1)[12]	0.0644
ARIMA(3,0,2)(0,1,1)[12]	0.0644
ARIMA(3,0,2)(2,1,0)[12]	0.0645
ARIMA(3,0,1)(2,1,0)[12]	0.0646
ARIMA(3,0,0)(2,1,0)[12]	0.0661
ARIMA(3,0,1)(1,1,0)[12]	0.0679

Cortecosteroid drug sales

- Models with lowest AICc values tend to give slightly better results than the other models.
- AICc comparisons must have the same orders of differencing. But RMSE test set comparisons can involve any models.
- Use the best model available, even if it does not pass all tests.

Cortecosteroid drug sales

```r
fit <- Arima(h02, order=c(3,0,1), seasonal=c(0,1,2),
  lambda=0)
autoplot(forecast(fit)) +
  ylab("H02 sales (million scripts)") + xlab("Year")
```



Outline

1. Stationarity and differencing
2. Non-seasonal ARIMA models
3. Order selection
4. ARIMA modelling in R
5. Seasonal ARIMA models
6. **ARIMA vs ETS**

ARIMA vs ETS

- Myth that ARIMA models are more general than exponential smoothing.
- Linear exponential smoothing models all special cases of ARIMA models.
- Non-linear exponential smoothing models have no equivalent ARIMA counterparts.
- Many ARIMA models have no exponential smoothing counterparts.
- ETS models all non-stationary. Models with seasonality or non-damped trend (or both) have two unit roots; all other models have one unit root.

Equivalences

ETS model	ARIMA model	Parameters
ETS(A,N,N)	ARIMA(0,1,1)	$\theta_1 = \alpha - 1$
ETS(A,A,N)	ARIMA(0,2,2)	$\theta_1 = \alpha + \beta - 2$
		$\theta_2 = 1 - \alpha$
ETS(A,A,N)	ARIMA(1,1,2)	$\phi_1 = \phi$
		$\theta_1 = \alpha + \phi\beta - 1 - \phi$
		$\theta_2 = (1 - \alpha)\phi$
ETS(A,N,A)	ARIMA(0,0,m)(0,1,0)$_m$	
ETS(A,A,A)	ARIMA(0,1,$m+1$)(0,1,0)$_m$	
ETS(A,A,A)	ARIMA(1,0,$m+1$)(0,1,0)$_m$	

Your turn

For the `condmilk` series:

- Do the data need transforming? If so, find a suitable transformation.
- Are the data stationary? If not, find an appropriate differencing which yields stationary data.
- Identify a couple of ARIMA models that might be useful in describing the time series.
- Which of your models is the best according to their AIC values?
- Estimate the parameters of your best model and do diagnostic testing on the residuals. Do the residuals resemble white noise? If not, try to find another ARIMA model which fits better.
- Forecast the next 24 months of data using your preferred model.
- Compare the forecasts obtained using `ets()`.