Welcome to Forecasting Using R

FORECASTING IN R

Rob J. Hyndman

Professor of Statistics at Monash University

What you will learn

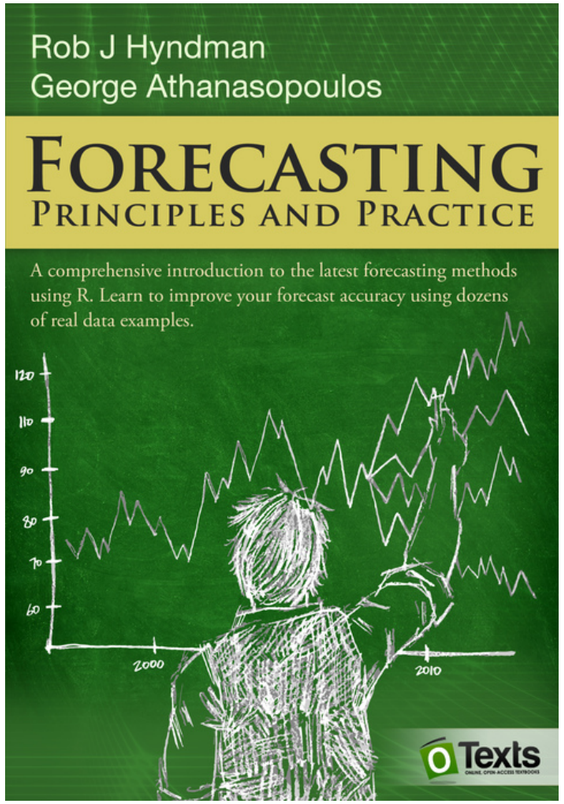
- Exploring and visualizing time series
- Simple benchmark methods for forecasting
- Exponential smoothing and ARIMA models
- Advanced forecasting methods
- Measuring forecast accuracy
- Choosing the best method

Course textbook

Hyndman, R. J. &
Athanasopoulos, G. (2017)

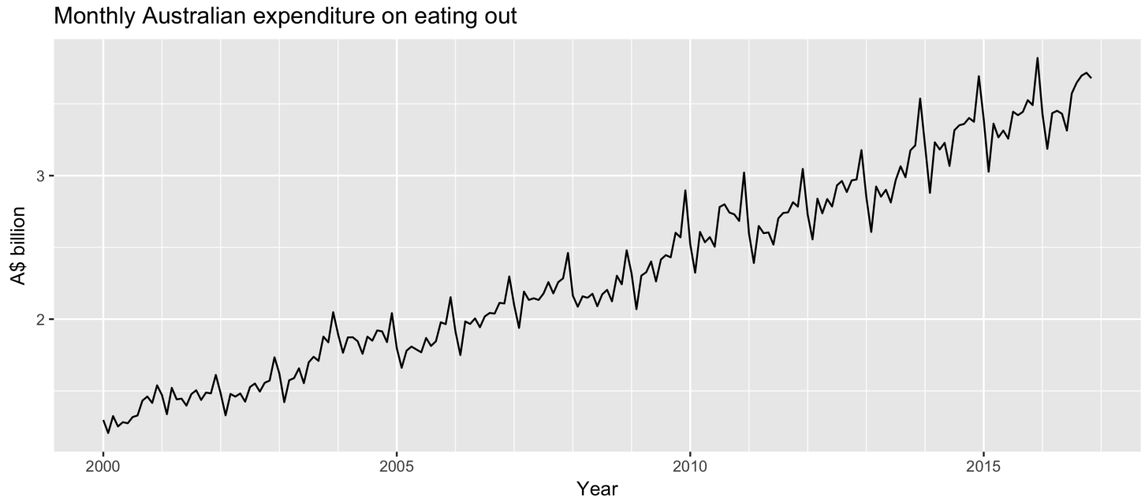
Forecasting: principles and practice, 2nd edition

- Free and online at OTexts.org/fpp2/
- Data sets in associated R package fpp2
- R code for all examples



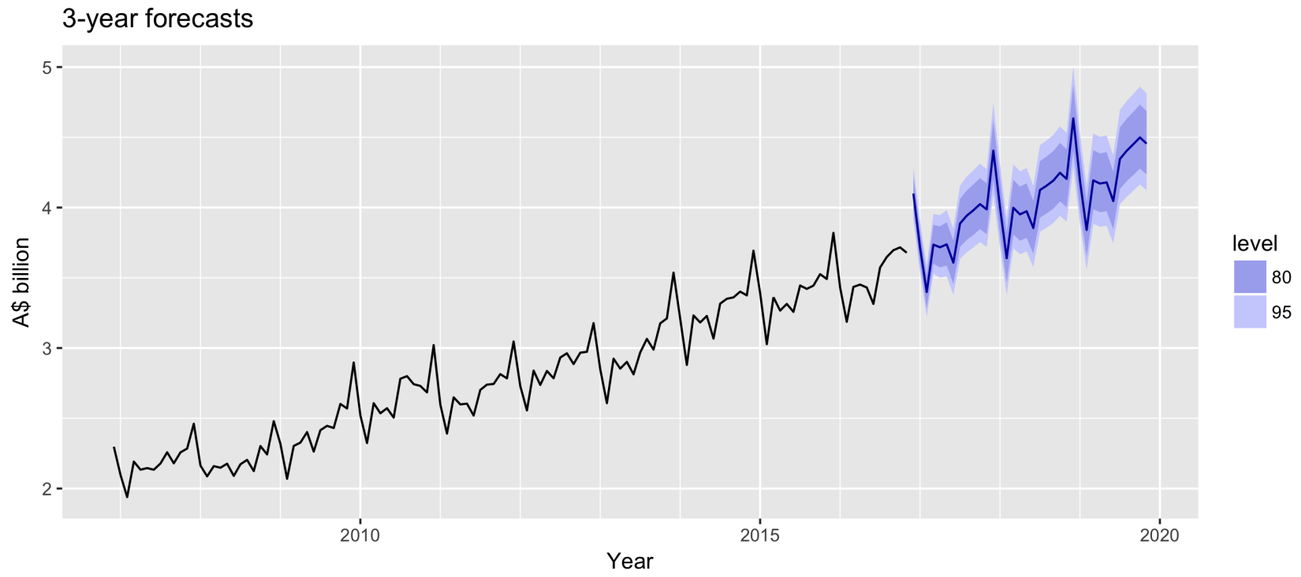
Time series data

- Series of data observed over time
- Eg.: Daily IBM stock prices, monthly rainfall in London,...



Forecasting is estimating how the sequence of observations will continue into the future.

Forecasts of monthly Australian expenditure on eating out



- What forecasting methods are available that take account of trend, seasonality and other features of the data?
- How to measure the accuracy of your forecasts?
- How to choose a good forecasting model?

Let's practice!

FORECASTING IN R

Trends, seasonality, and cyclicity

FORECASTING IN R

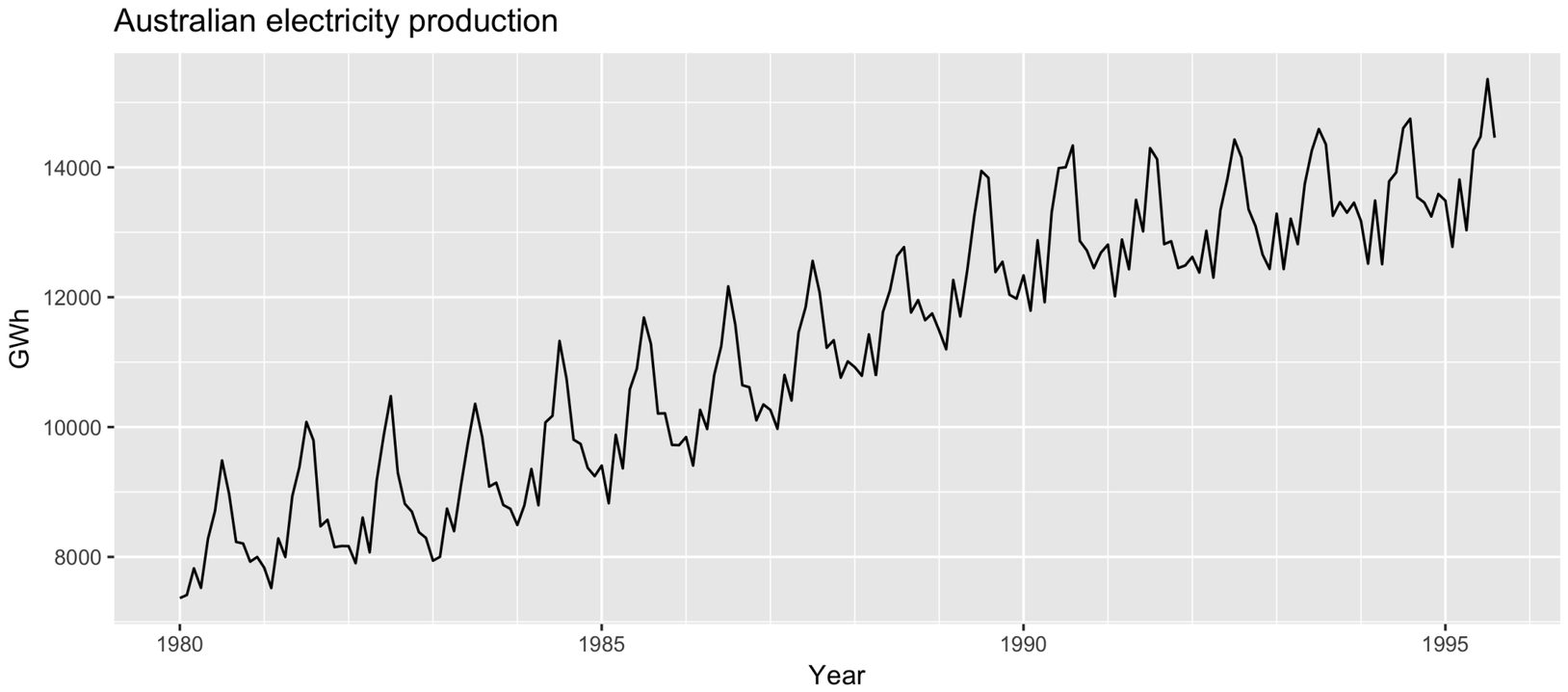
Rob J. Hyndman

Professor of Statistics at Monash University

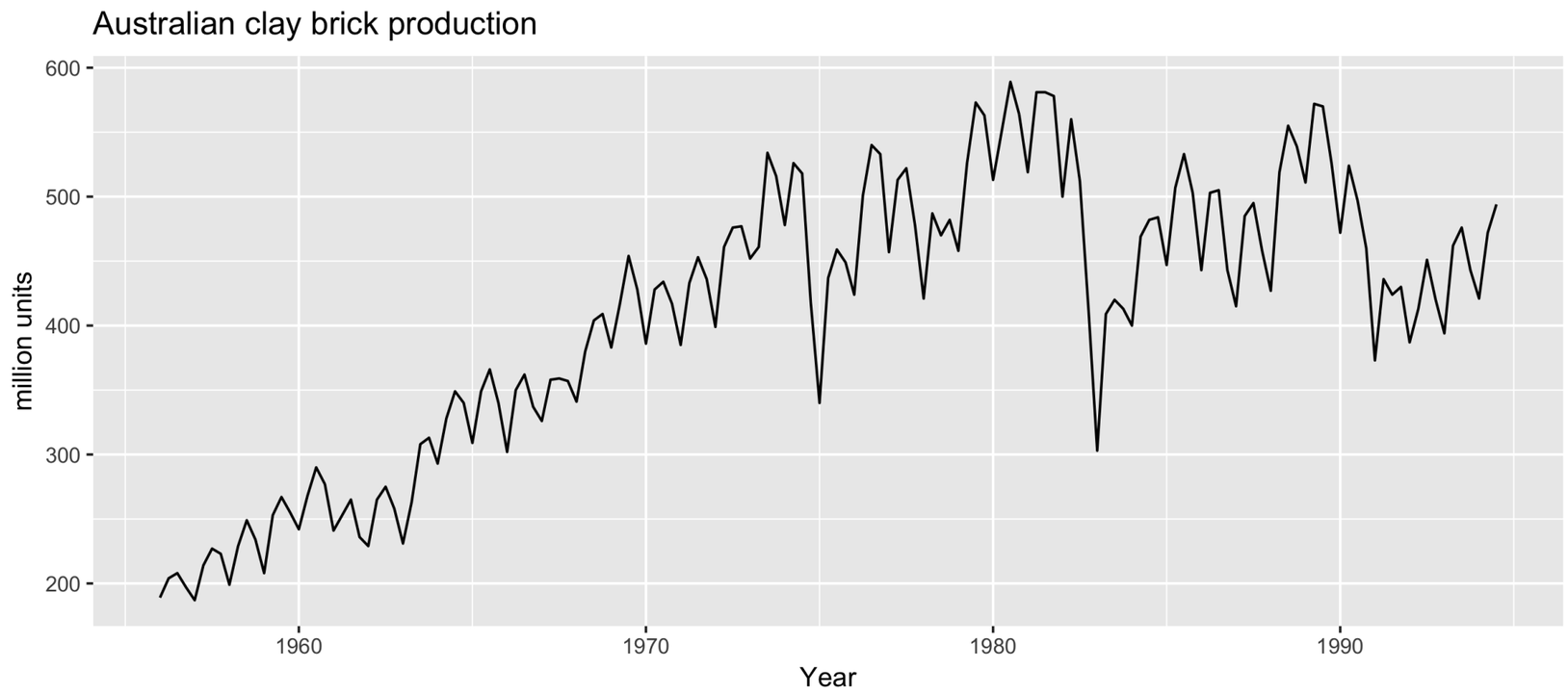
Time series patterns

Pattern	Description
Trend	A pattern exists involving a long-term increase OR decrease in the data
Seasonal	A periodic pattern exists due to the calendar (e.g., the quarter, month, or day of the week)
Cyclic	A pattern exists where the data exhibits rises and falls that are not of fixed period (duration usually of at least 2 years)

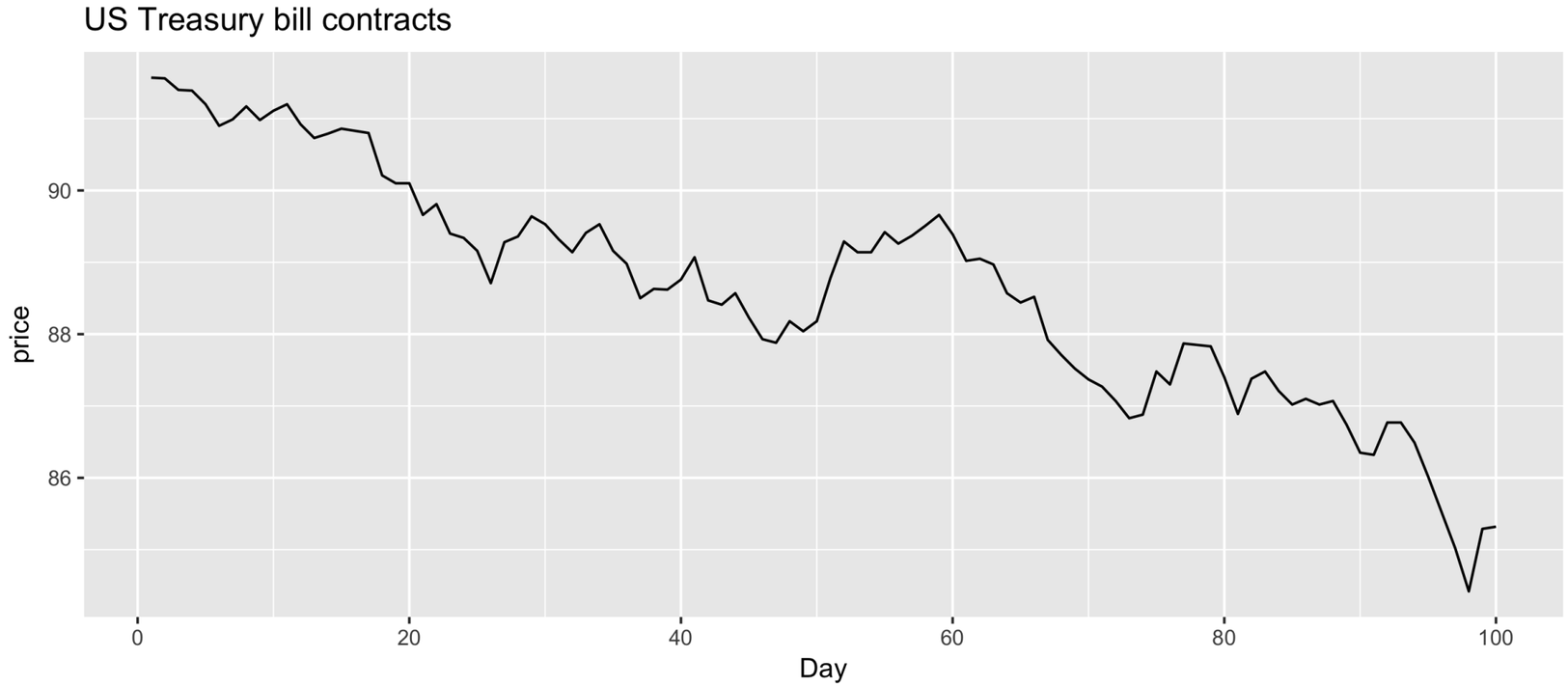
Examples of time series patterns



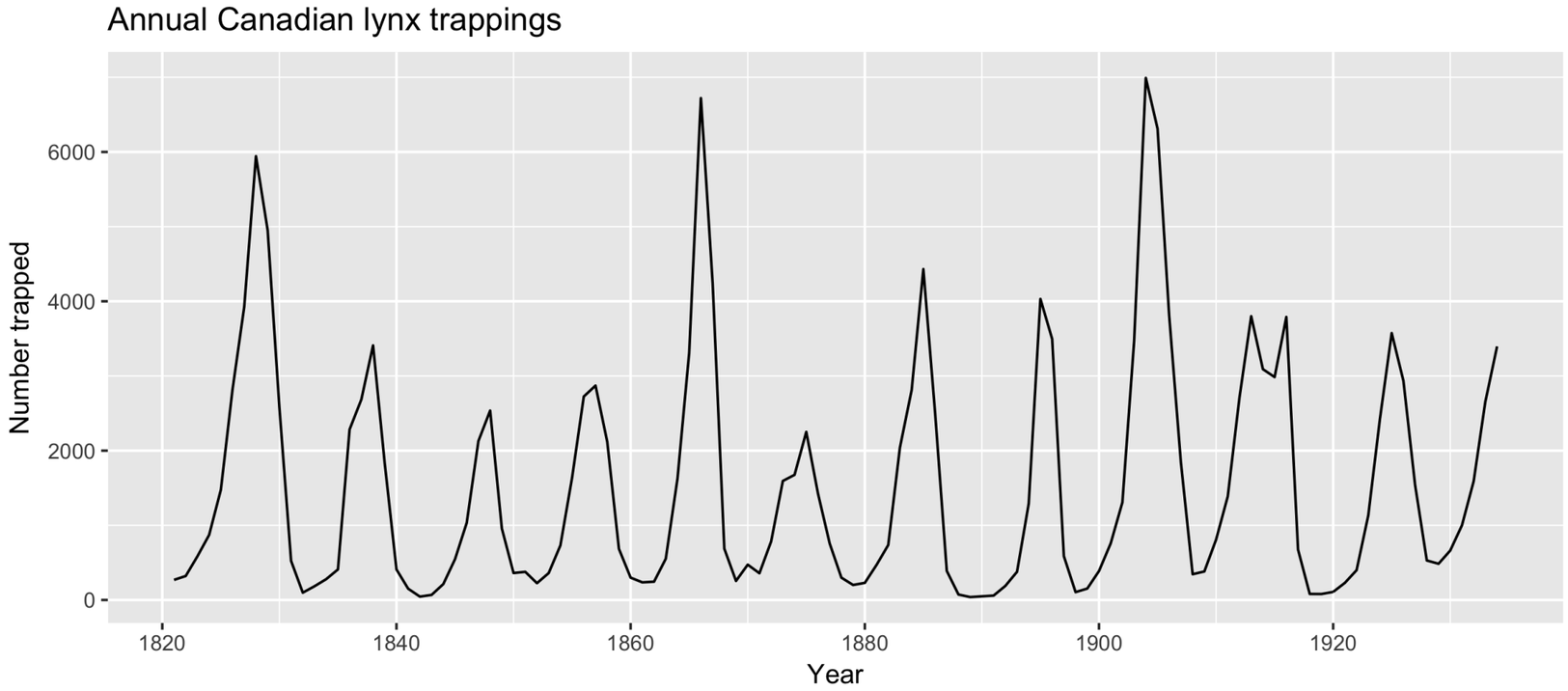
Examples of time series patterns



Examples of time series patterns



Examples of time series patterns



Seasonal or cyclic?

Differences between seasonal and cyclic patterns:

- Seasonal pattern constant length vs. cyclic pattern variable length
- Average length of cycle longer than length of seasonal pattern
- Magnitude of cycle more variable than magnitude of seasonal pattern

The timing of peaks and troughs is predictable with seasonal data, but unpredictable in the long term with cyclic data.

Let's practice!

FORECASTING IN R

White noise

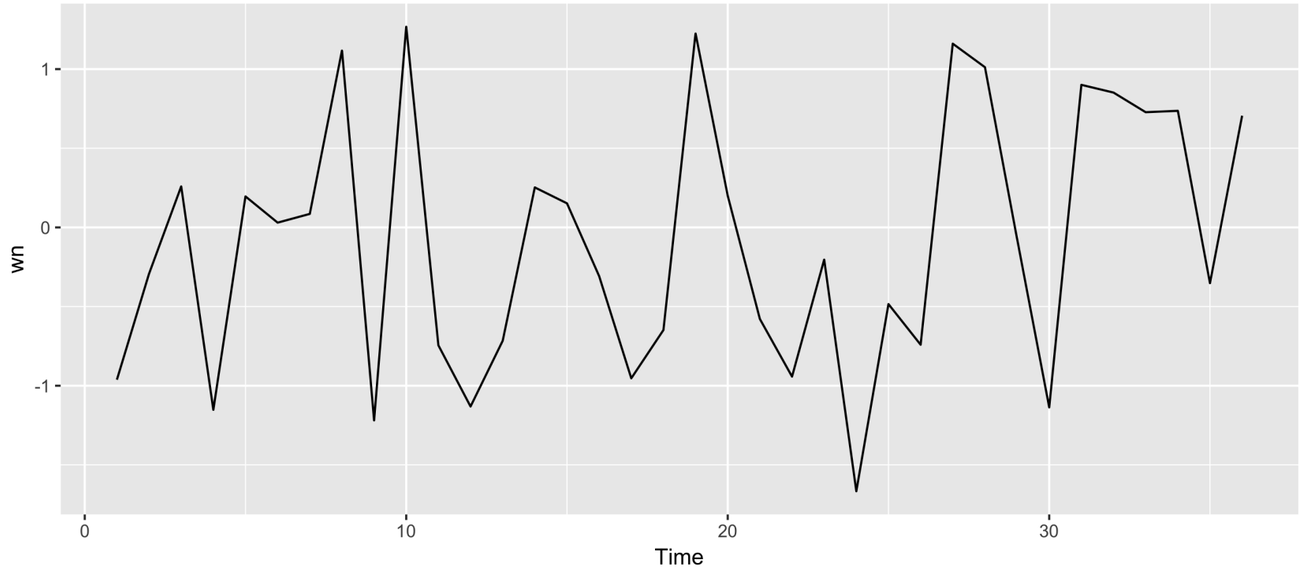
FORECASTING IN R

Rob J. Hyndman

Professor of Statistics at Monash University

White noise

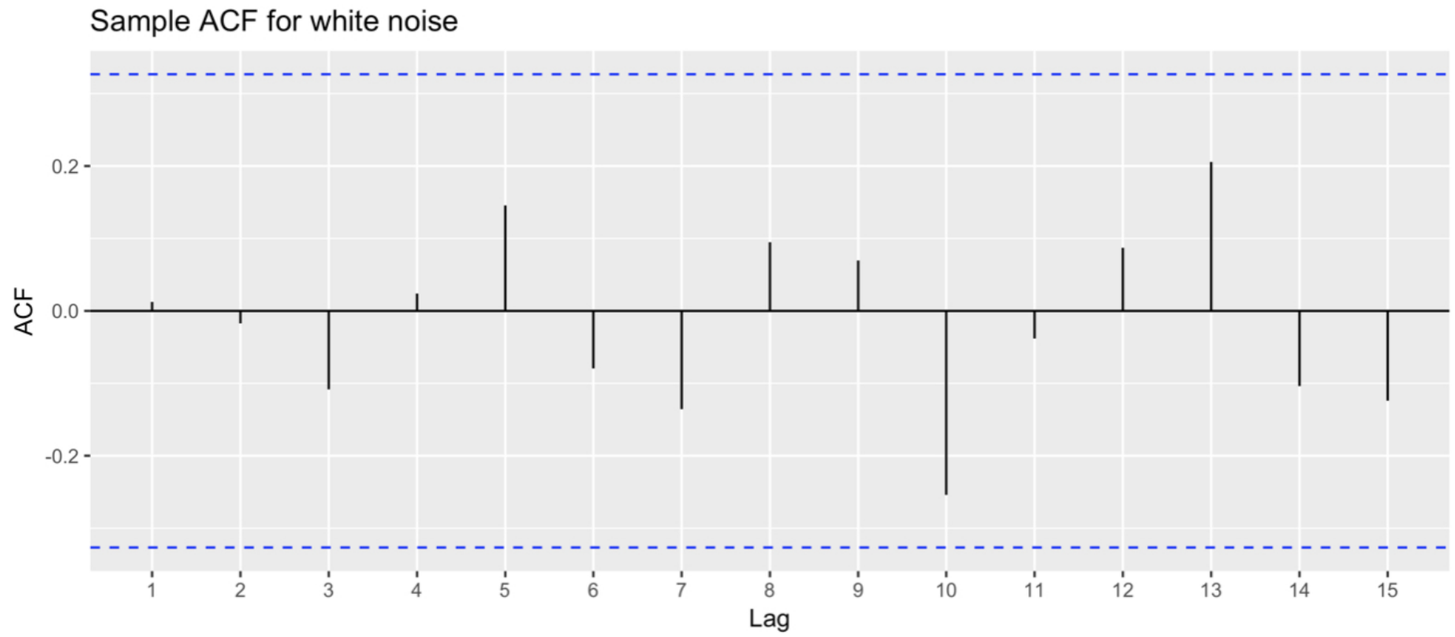
```
set.seed(3)           # Reproducibility
wn <- ts(rnorm(36))   # White noise
autoplot(wn)          # Plot!
```



"White noise" is just a time series of iid data

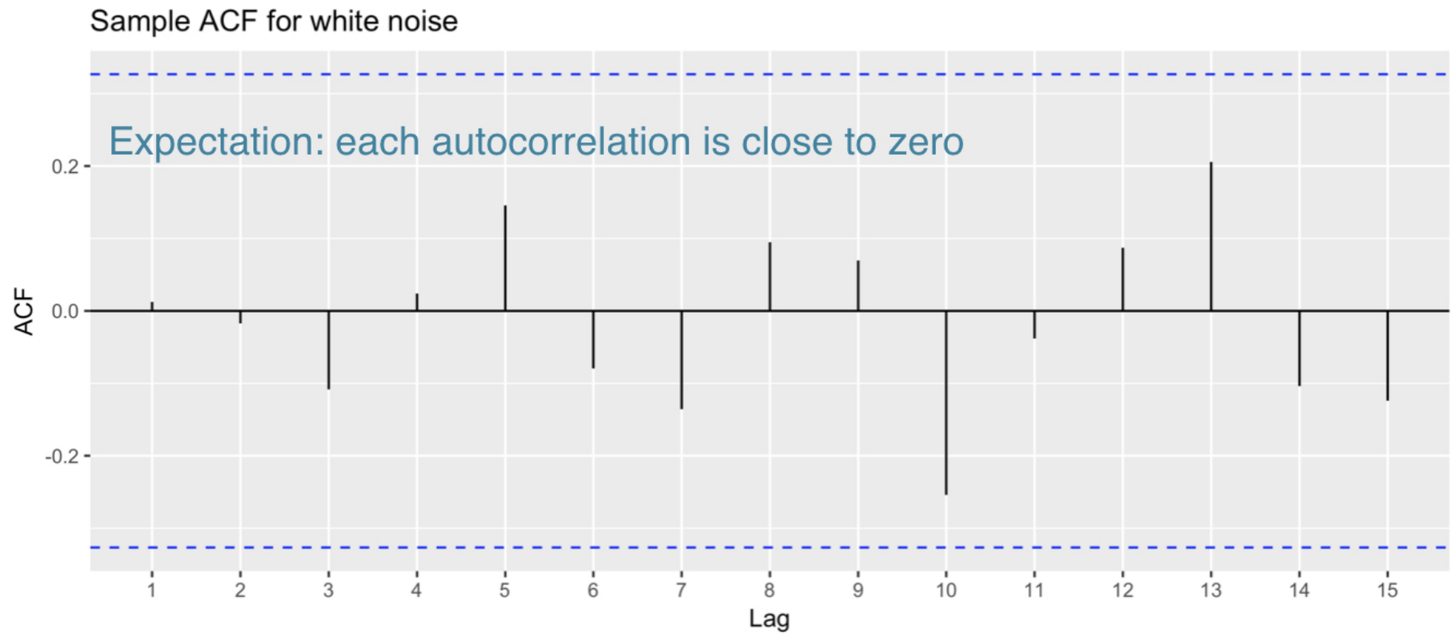
White noise ACF

```
ggAcf(wn) +
    ggtitle("Sample ACF for white noise")
```



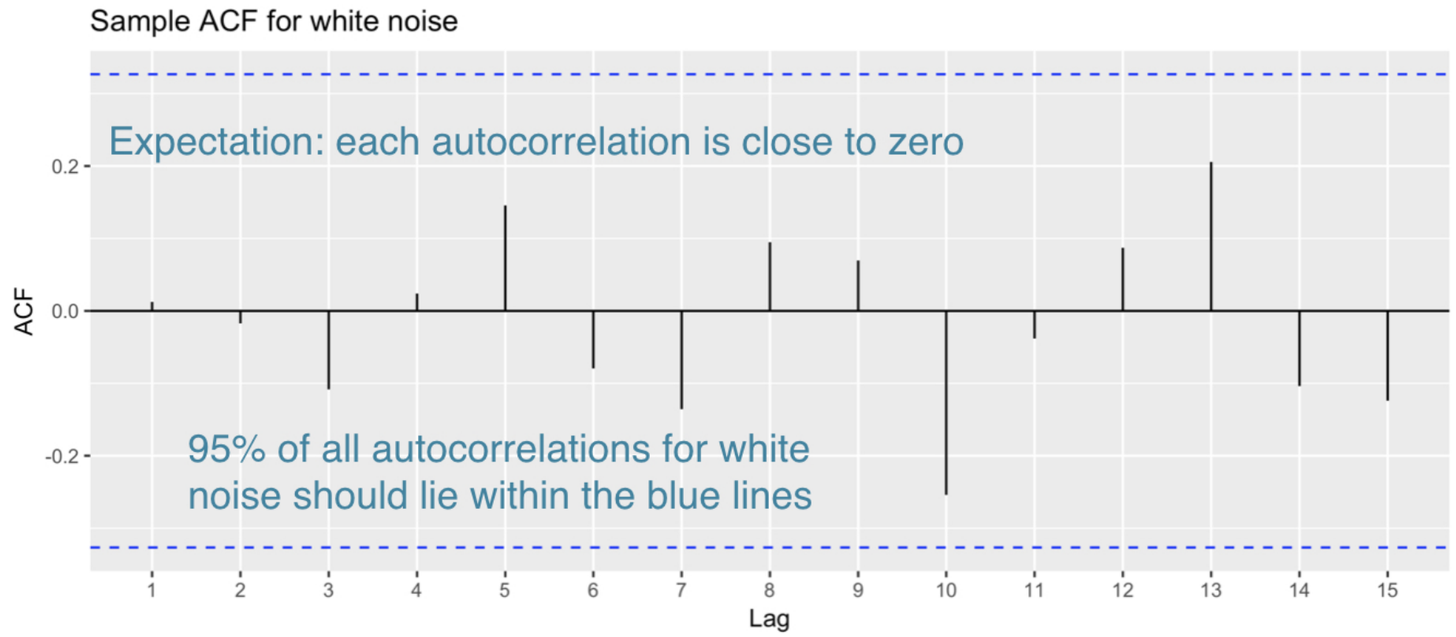
White noise ACF

```
ggAcf(wn) +
    ggtitle("Sample ACF for white noise")
```



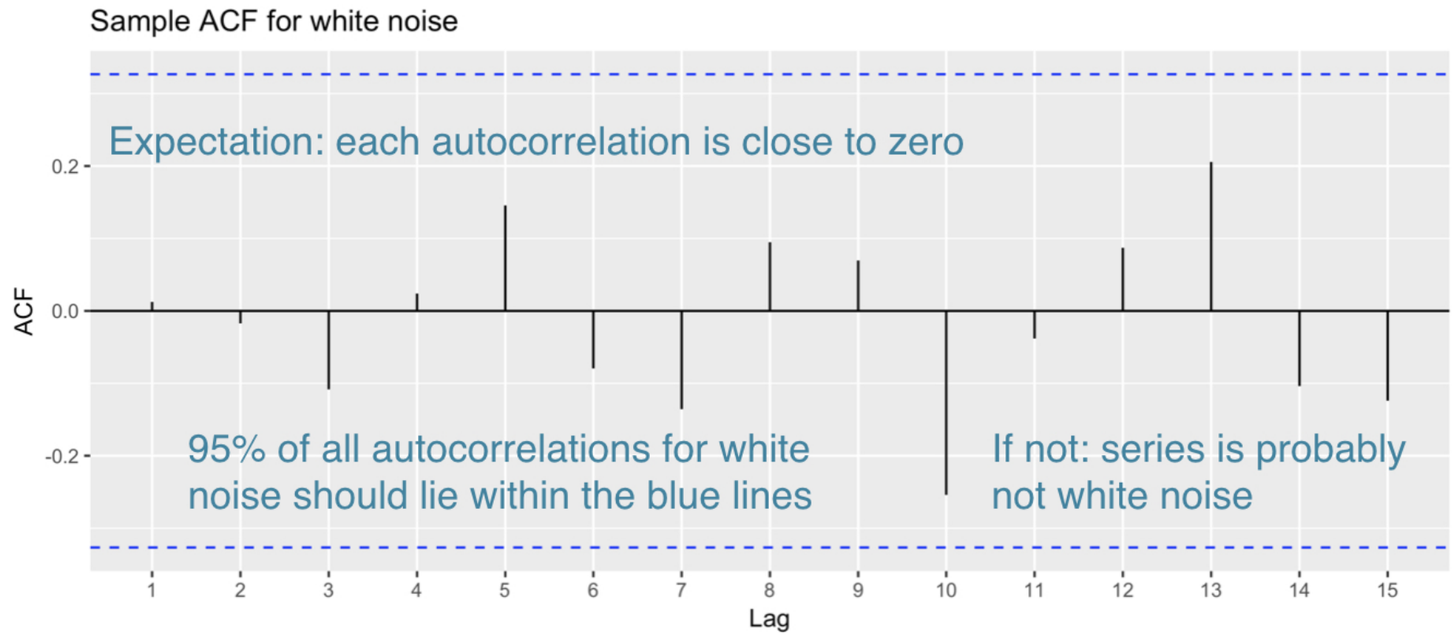
White noise ACF

```
ggAcf(wn) +
    ggtitle("Sample ACF for white noise")
```



White noise ACF

```
ggAcf(wn) +
    ggtitle("Sample ACF for white noise")
```

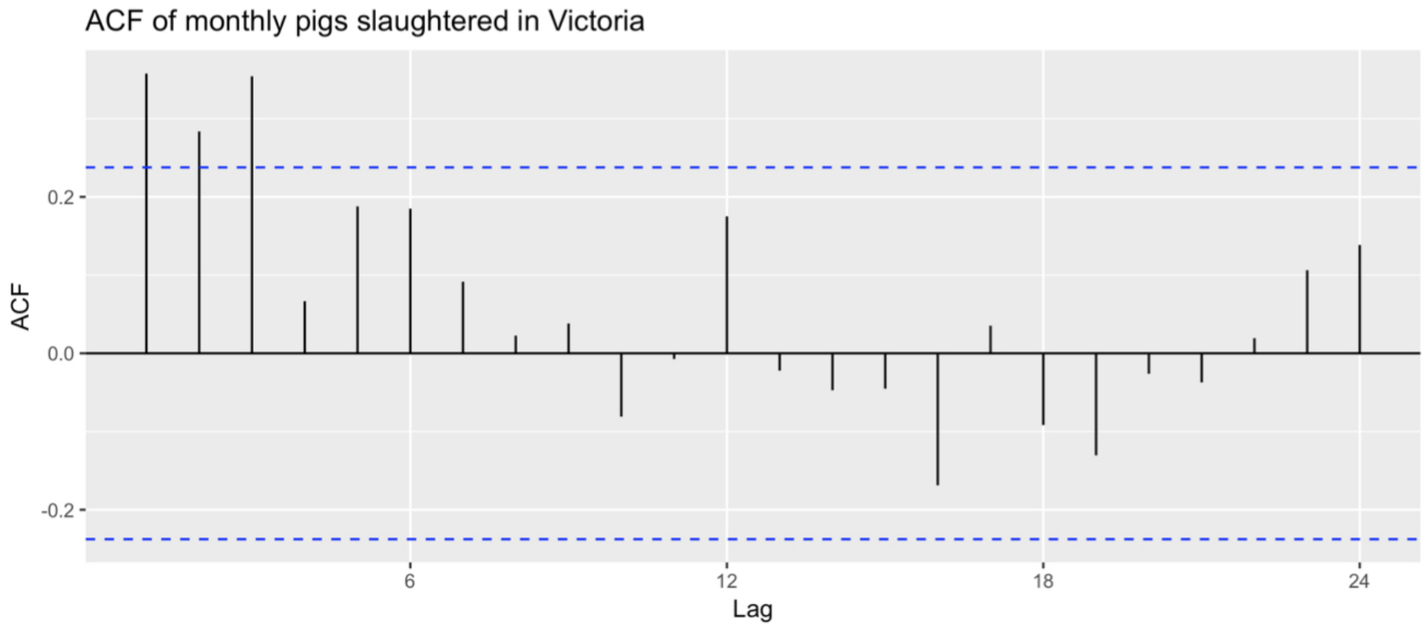


Example: Pigs slaughtered

```
pigs <- window(pigs, start=1990)
autoplot(pigs/1000) +
  xlab("Year") +
  ylab("thousands") +
  ggtitle("Monthly number of pigs slaughtered in Victoria")
```

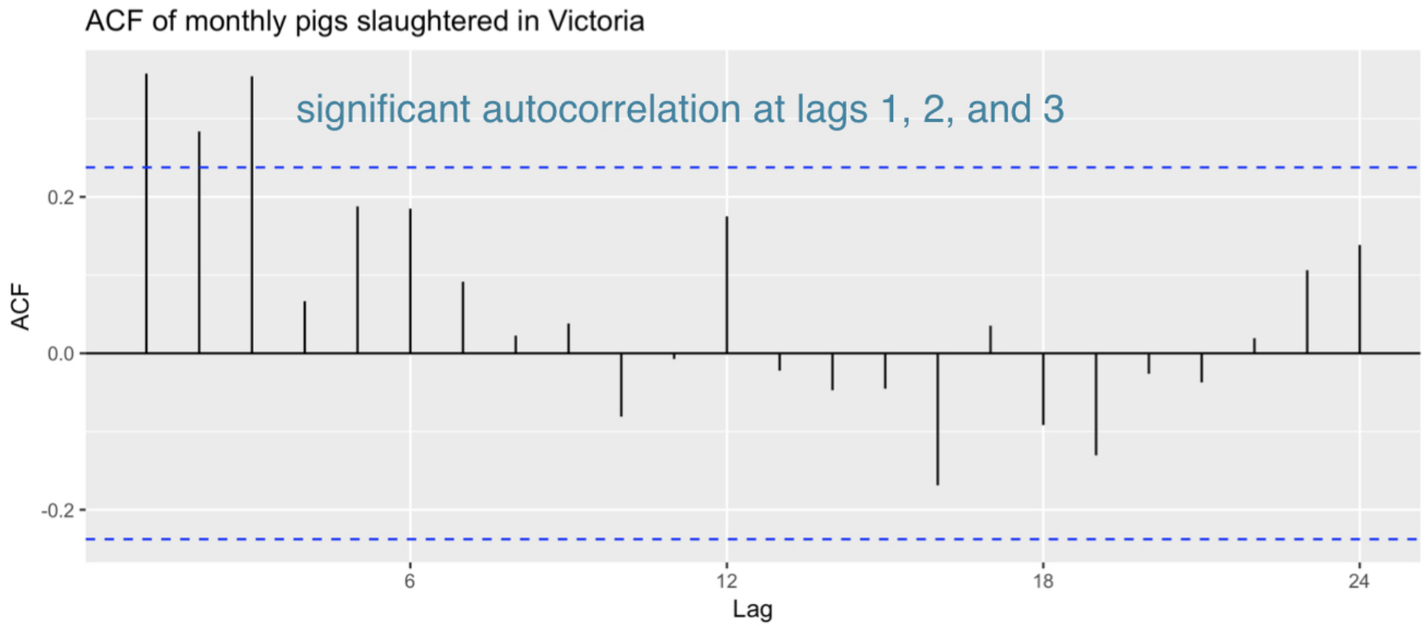
Example: Pigs slaughtered

```
ggAcf(pigs) +
  ggtitle("ACF of monthly pigs slaughtered
           in Victoria")
```



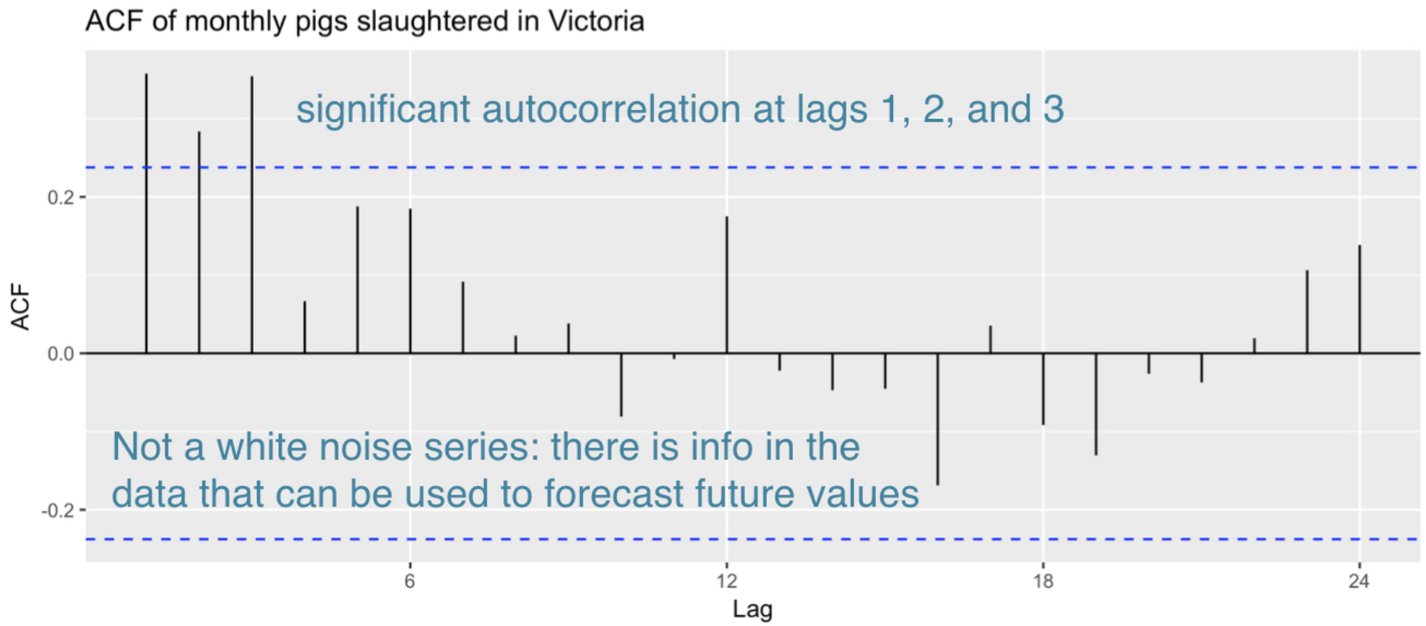
Example: Pigs slaughtered

```
ggAcf(pigs) +
  ggtitle("ACF of monthly pigs slaughtered
           in Victoria")
```



Example: Pigs slaughtered

```
ggAcf(pigs) +
  ggtitle("ACF of monthly pigs slaughtered
           in Victoria")
```



Ljung-Box test

The Ljung-Box test considers the first h autocorrelation values **together**.

A significant test (small p-value) indicates the data are probably not white noise.

```
Box.test(pigs, lag = 24, fitdf = 0, type = "Lj")
```

```
Box-Ljung test
data:  pigs
X-squared = 634.15, df = 24, p-value < 2.2e-16
```

White noise summary

- White noise is a time series that is purely random
- We can test for white noise by looking at an ACF plot or by doing a Ljung-Box test

Let's practice!

FORECASTING IN R