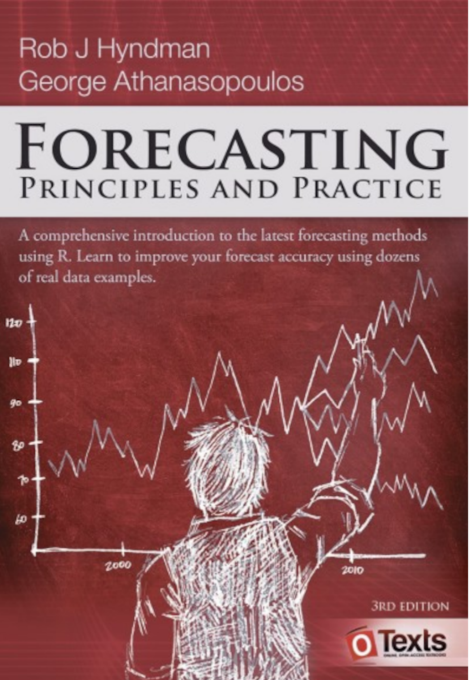


9. ARIMA models

9.1 Random walk model

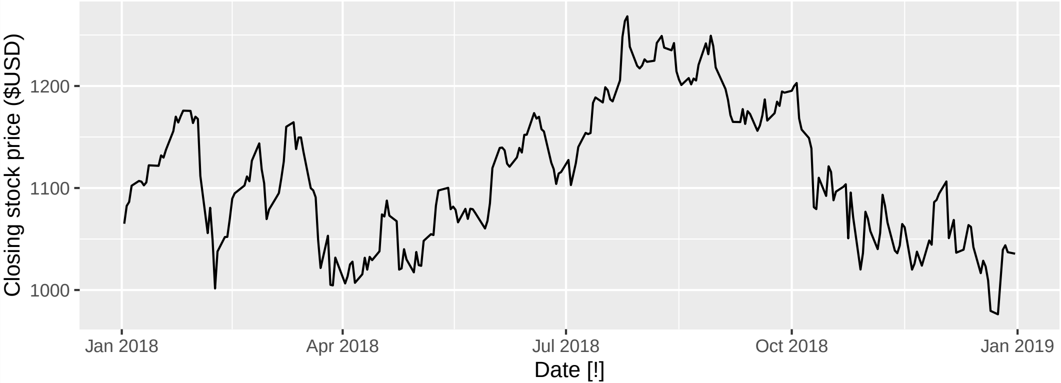
OTexts.org/fpp3/

Example: Google stock price

```r
google_2018 <- gafa_stock |>
  filter(Symbol == "GOOG", year(Date) == 2018)
```

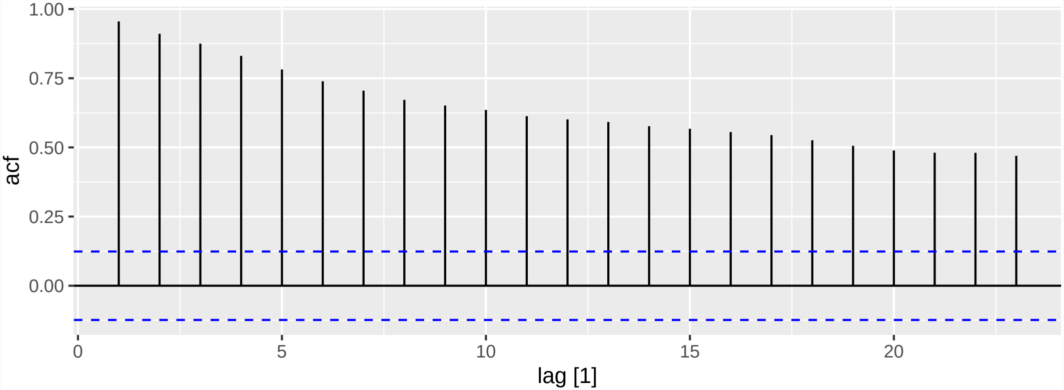
Example: Google stock price

```r
google_2018 |>
  autoplot(Close) +
  labs(y = "Closing stock price ($USD)")
```



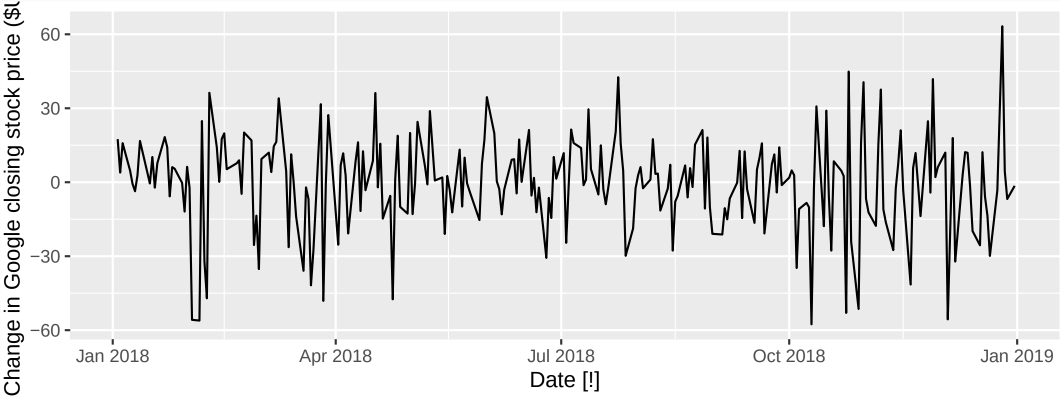
Example: Google stock price

```r
google_2018 |>
  ACF(Close) |>
  autoplot()
```



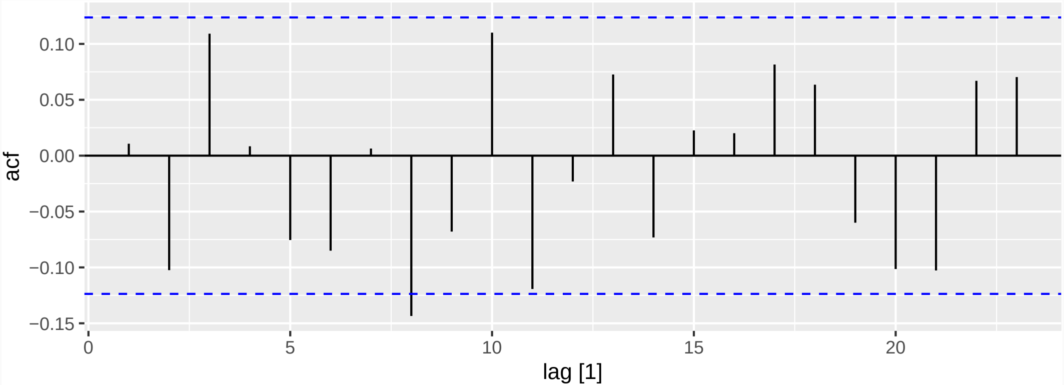
Example: Google stock price

```r
google_2018 |>
  autoplot(difference(Close)) +
  labs(y = "Change in Google closing stock price ($USD)")
```



Example: Google stock price

```r
google_2018 |>
  ACF(difference(Close)) |>
  autoplot()
```



Example: Google stock price

- The differences are the day-to-day changes.
- Now the series looks just like a white noise series:
  - No autocorrelations outside the 95% limits.
- Conclusion: The daily change in the Google stock price is essentially a random amount uncorrelated with previous days.

Random walk model

If differenced series is white noise with zero mean:

$$y_t - y_{t-1} = \varepsilon_t \qquad \text{or} \qquad y_t = y_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim NID(0, \sigma^2)$.

- Very widely used for non-stationary data.
- This is the model behind the naïve method.

Random walk model

If differenced series is white noise with zero mean:

$$y_t - y_{t-1} = \varepsilon_t \qquad \text{or} \qquad y_t = y_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim NID(0, \sigma^2)$.

- Very widely used for non-stationary data.
- This is the model behind the naïve method.
- Random walks typically have:
  - long periods of apparent trends up or down
  - Sudden/unpredictable changes in direction

Random walk model

If differenced series is white noise with zero mean:

$$y_t - y_{t-1} = \varepsilon_t \qquad \text{or} \qquad y_t = y_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim NID(0, \sigma^2)$.

- Very widely used for non-stationary data.
- This is the model behind the naïve method.
- Random walks typically have:
  - long periods of apparent trends up or down
  - Sudden/unpredictable changes in direction
- Forecast are equal to the last observation (naïve)
  - future movements up or down are equally likely.

Random walk with drift model

If differenced series is white noise with non-zero mean:

$$y_t - y_{t-1} = c + \varepsilon_t \quad \text{or} \quad y_t = c + y_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim NID(0, \sigma^2)$.

- c is the average change between consecutive observations.
- If $c > 0$, y_t will tend to drift upwards and vice versa.
- This is the model behind the drift method.

Seasonal differencing

If seasonally differenced data is white noise it implies:

$$y_t - y_{t-m} = \varepsilon_t \qquad \text{or} \qquad y_t = y_{t-m} + \varepsilon_t$$

where $\varepsilon_t \sim NID(0, \sigma^2)$.

- The model behind the seasonal naïve method.