Introduction to Data Science

# Time Series

Gordon Anderson

# Time Series

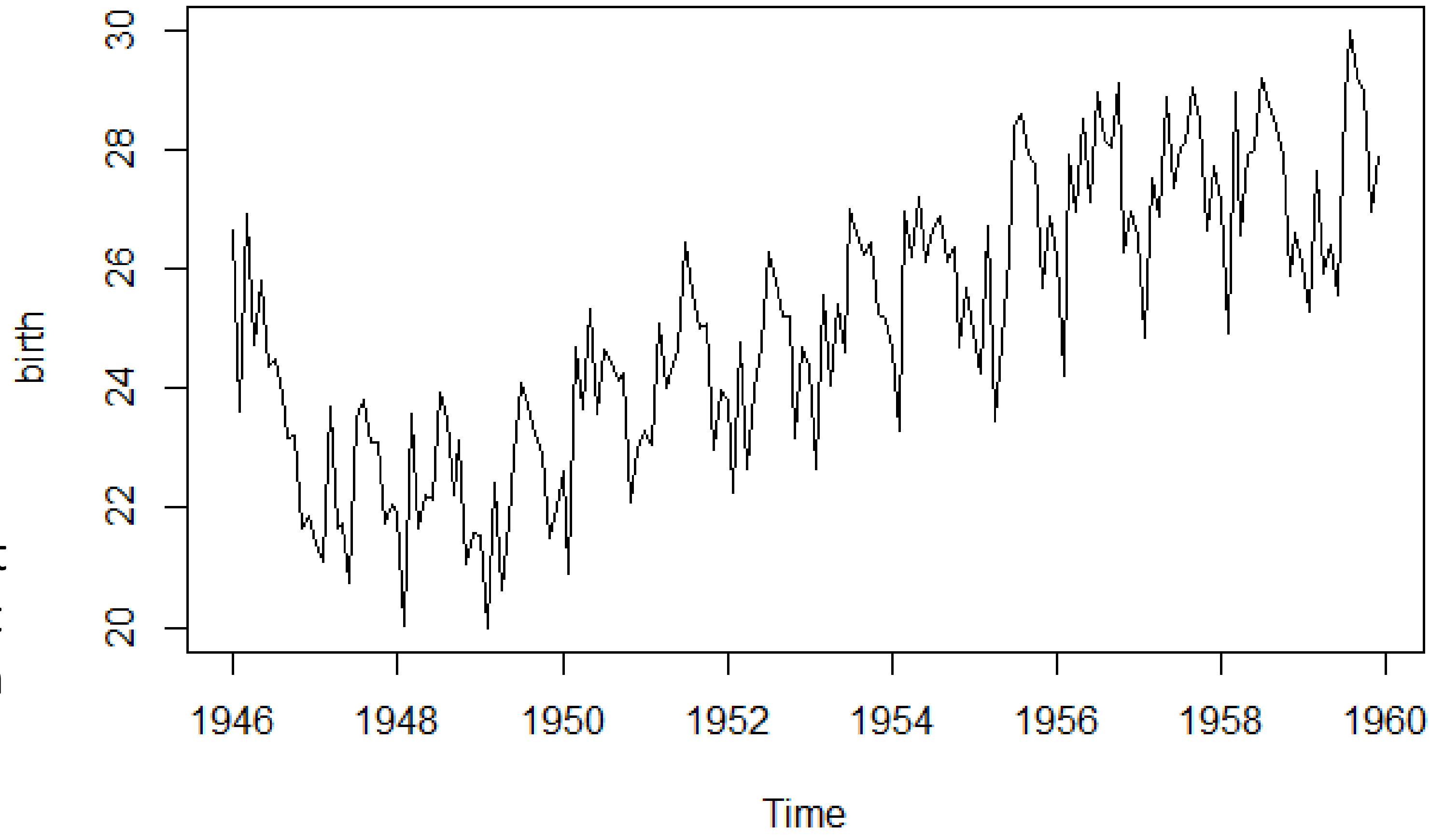
- A time series is a number of observations recorded over time.
- Time series data is consistently measured at equally spaced intervals.
- Data recorded irregularly is not a time series.
- A time series has a time unit and a frequency.
- For example, consider the number of births recorded from 1946 to 1958 on a monthly basis. The time unit is a year, with a frequency of 12 recordings per year.

# Time Series

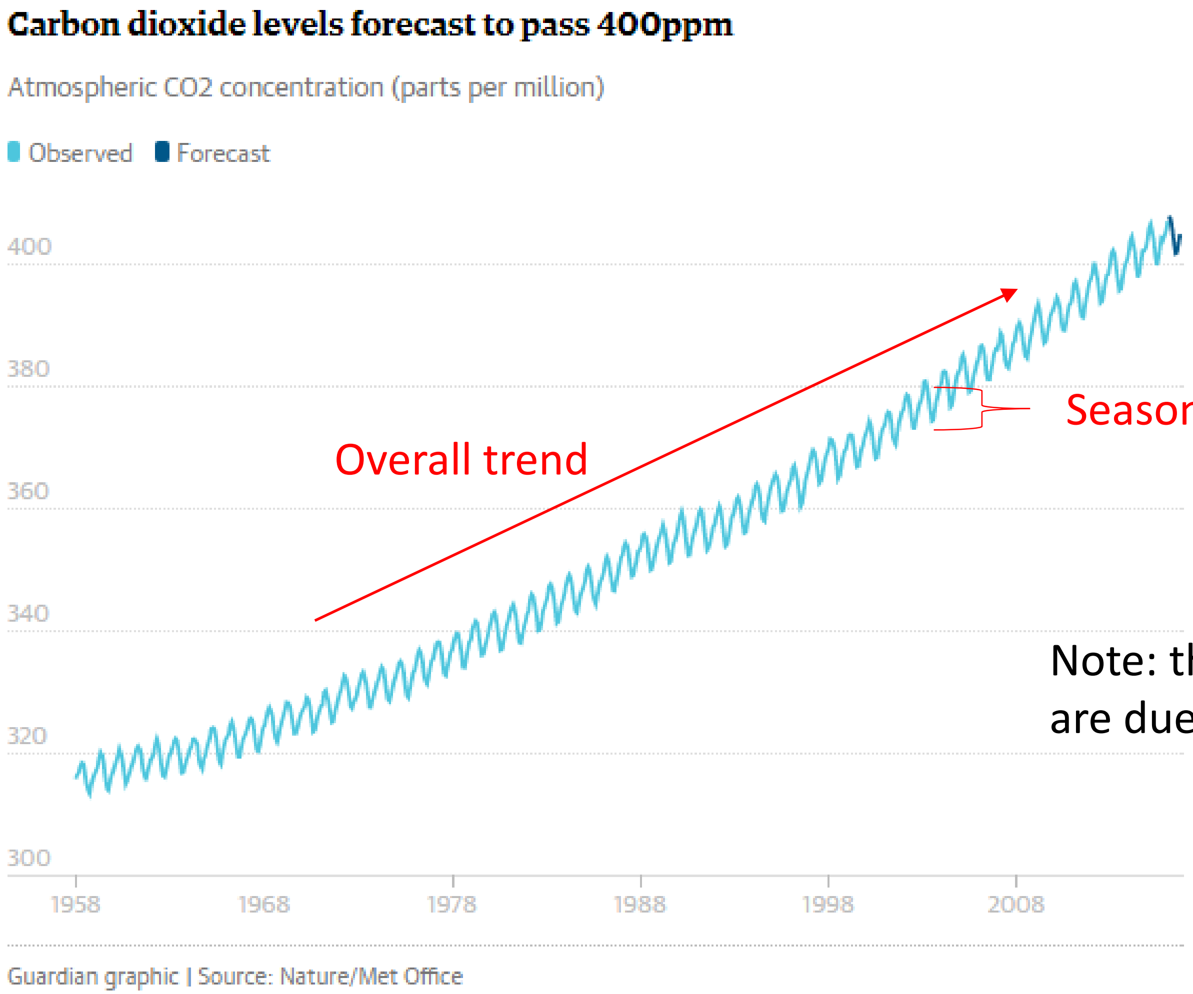
A time series has 4 components:

1. Trend- the level over time.
2. Trend cycles.
3. Seasonal pattern.
4. Noise-related fluctuations

In the plot of births, the trend is Increasing. There are no cycles, but We may not see it if our data is not Long enough. The seasonal pattern is a peak of births in summer. The fluctuations are pretty regular Overall.



# Example Time Series:




Predicted values, or “forecast”

Seasonal cycles

Note: the seasonal cycles in the data are due to the changes in foliage.

# Smoothing

Use a *smoothing* function to observe  just the trend component.

Simple Moving Average (SMA):
The average of *n* values in the series calculated on a moving “window”. All windows equally “weighted”.

Observed data= 26.66, 23.560, 26.93, 24.74, 25.80, 24.36

25.73

Observed data= 26.66, 23.560, 26.93, 24.74, 25.80, 24.36

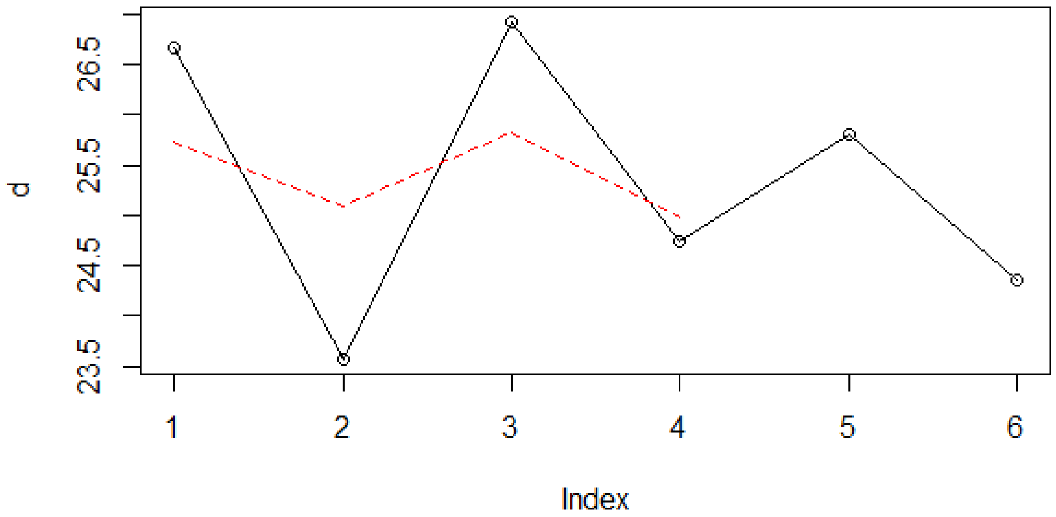
25.09

Observed data= 26.66, 23.560, 26.93, 24.74, 25.80, 24.36

25.82

Observed data= 26.66, 23.560, 26.93, 24.74, 25.80, 24.36

24.97

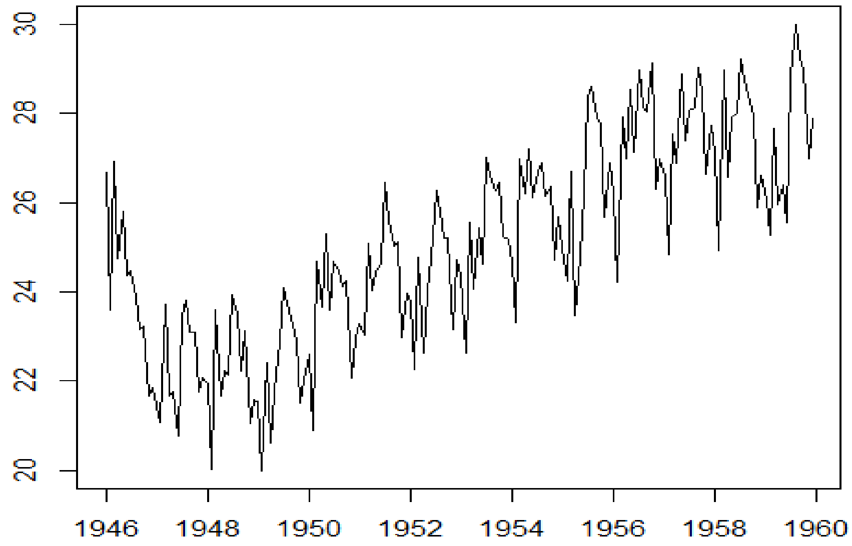


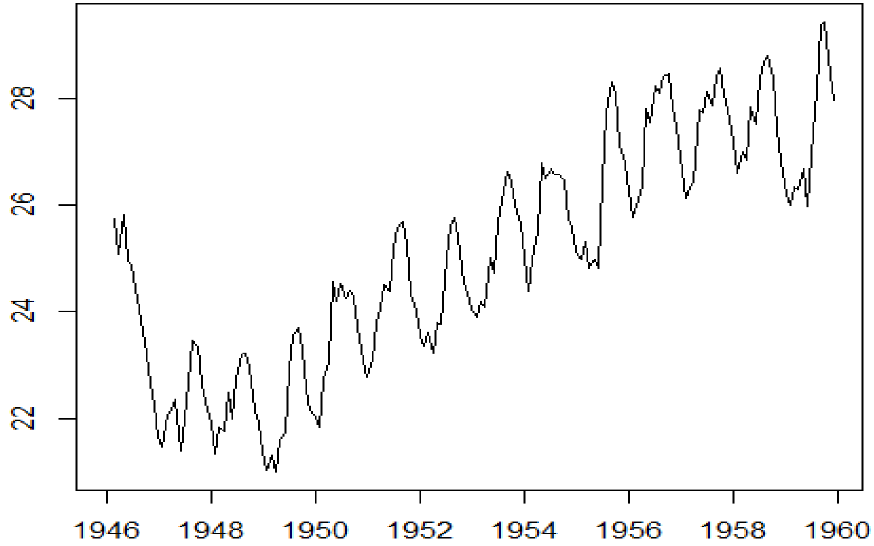

Black: data
Red: SMA

# Smoothing

The entire “births” series smoothed with SMA.

Top: original plot
Bottom: smoothed with *n*=3



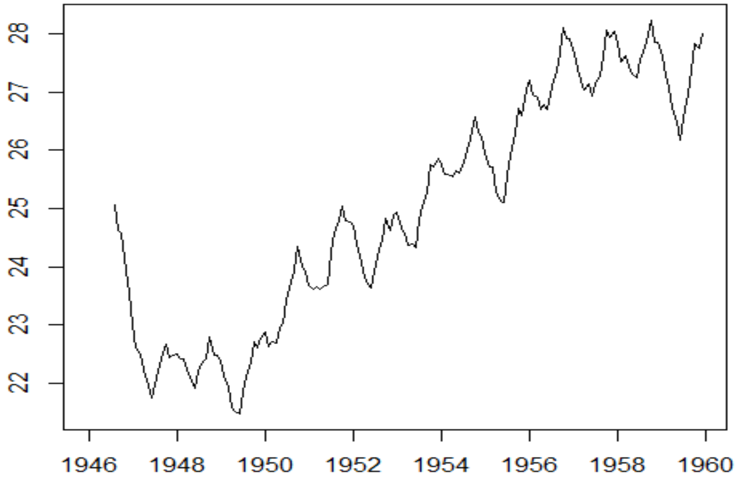



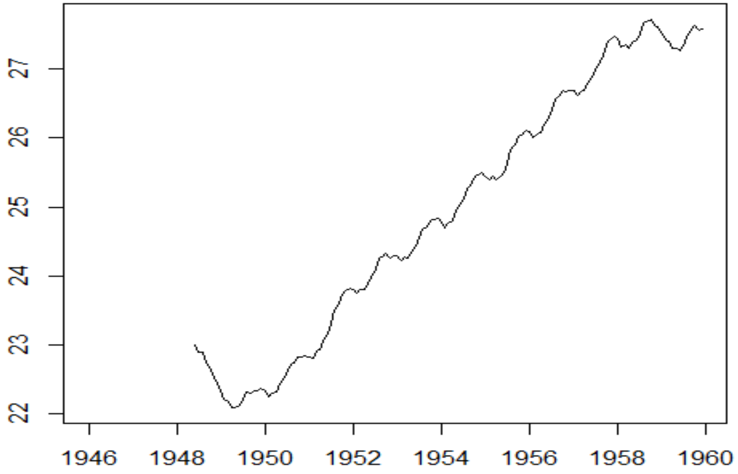
# Smoothing

Smoothing methods are a “low-pass” filter- they filter out the high frequency noise.

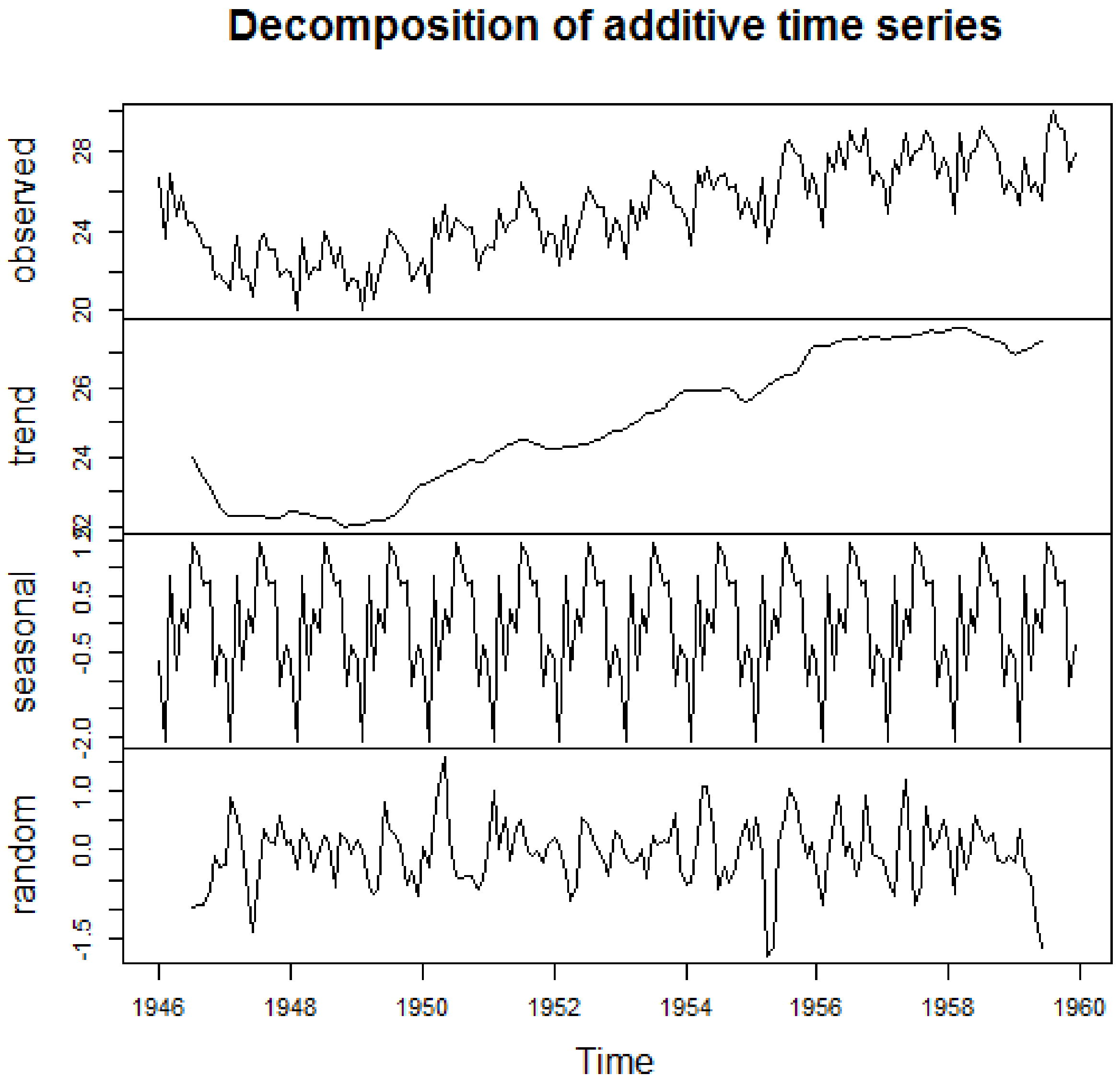
Top: smoothed with *n*=8
Bottom: smoothed with *n*=30

As *n* increases, the trend is more apparent.

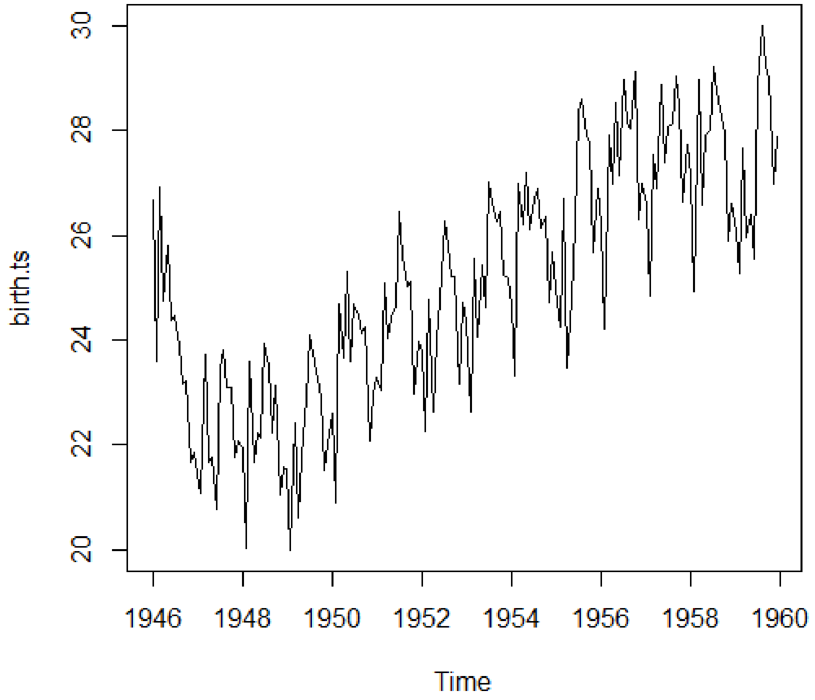


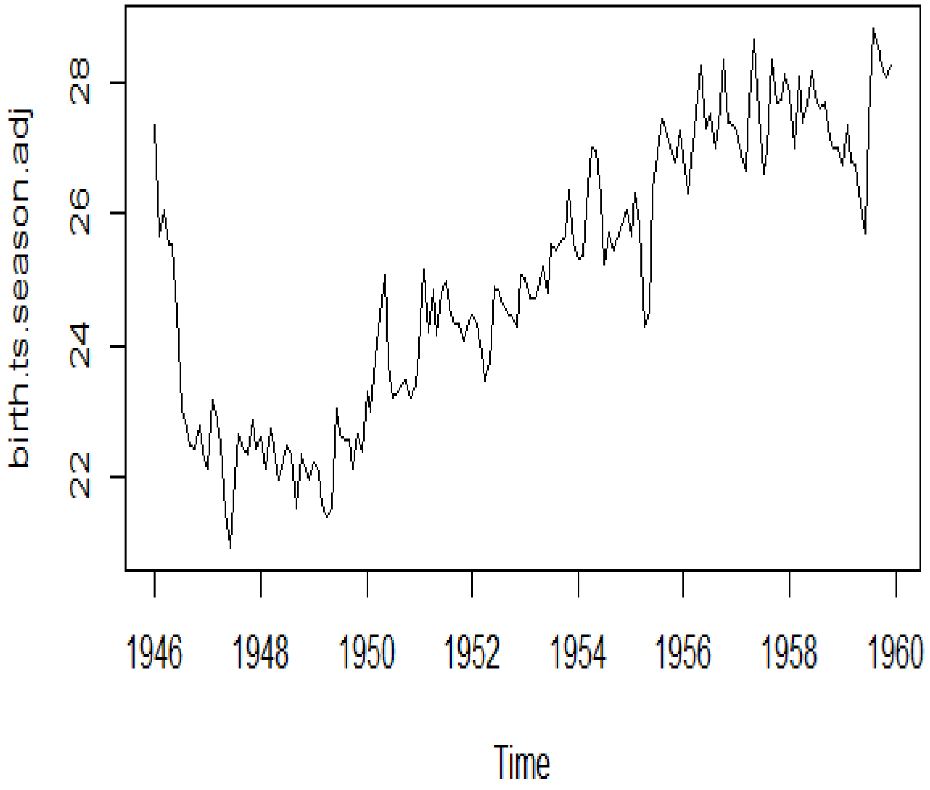
# TS Components



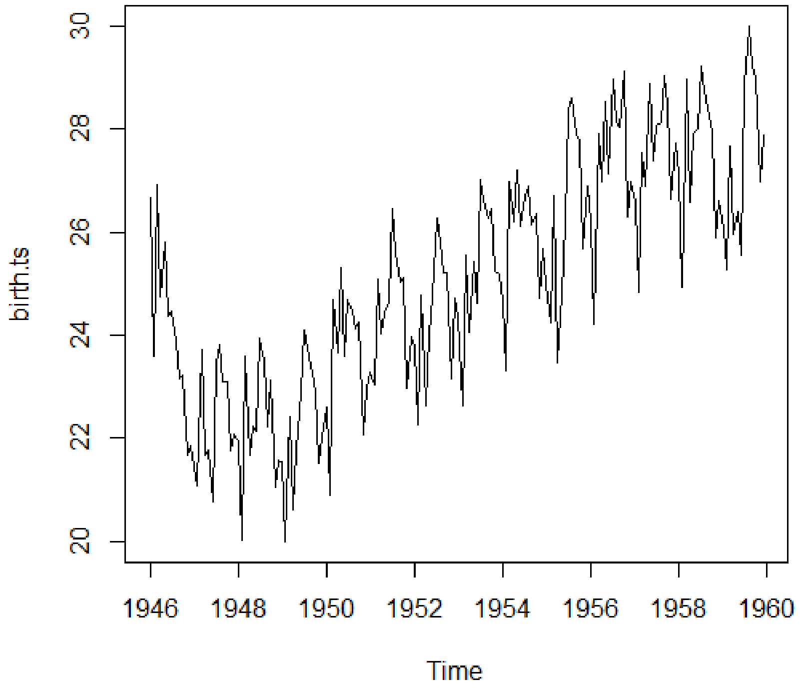
# Seasonal Adjustment

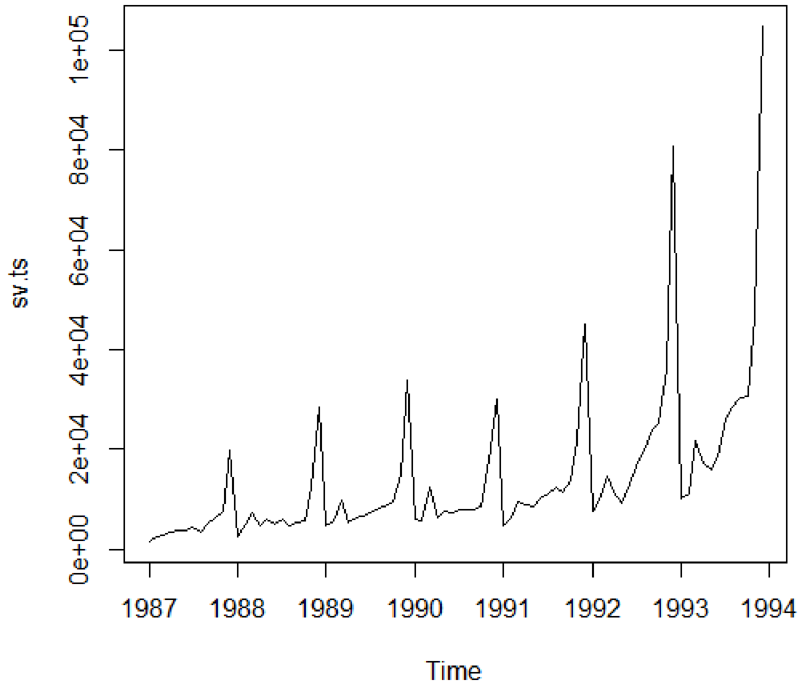



The birth data with the seasonal Component removed.



# Additive vs. Multiplicative Time Series



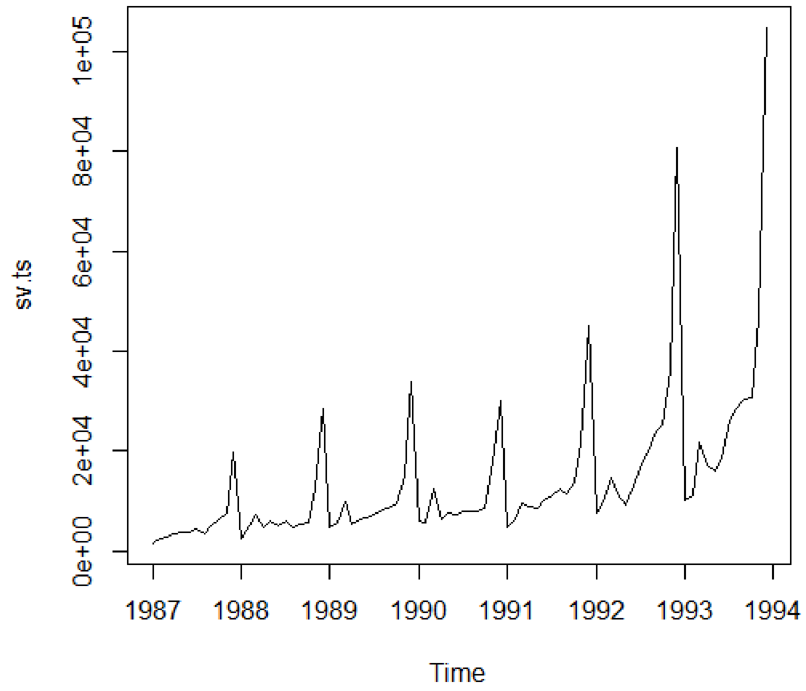

Additive: amplitude constant

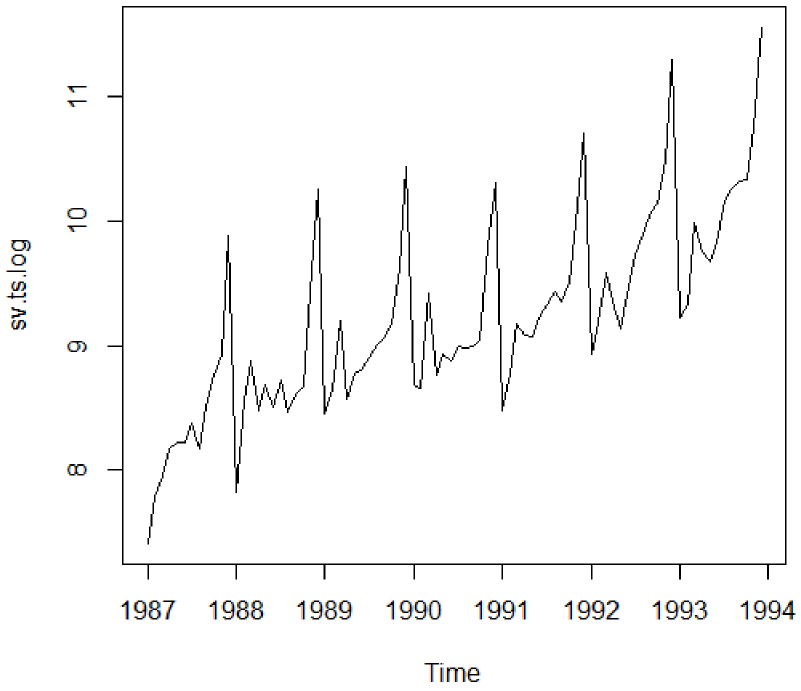
AddTS = trend + cycle + noise

Multiplicative: amplitude changes

MultTS = trend * cycle * noise

# Log transform




Multiplicative: amplitude changes

MultTS = trend * cycle * noise

Additive: amplitude constant

log(MultTS) = log(trend) + log(cycle) + log(noise)

# Exponential Smoothing

*Exponential smoothing*. Differs from SMA in that the window "weights" exponentially decrease.

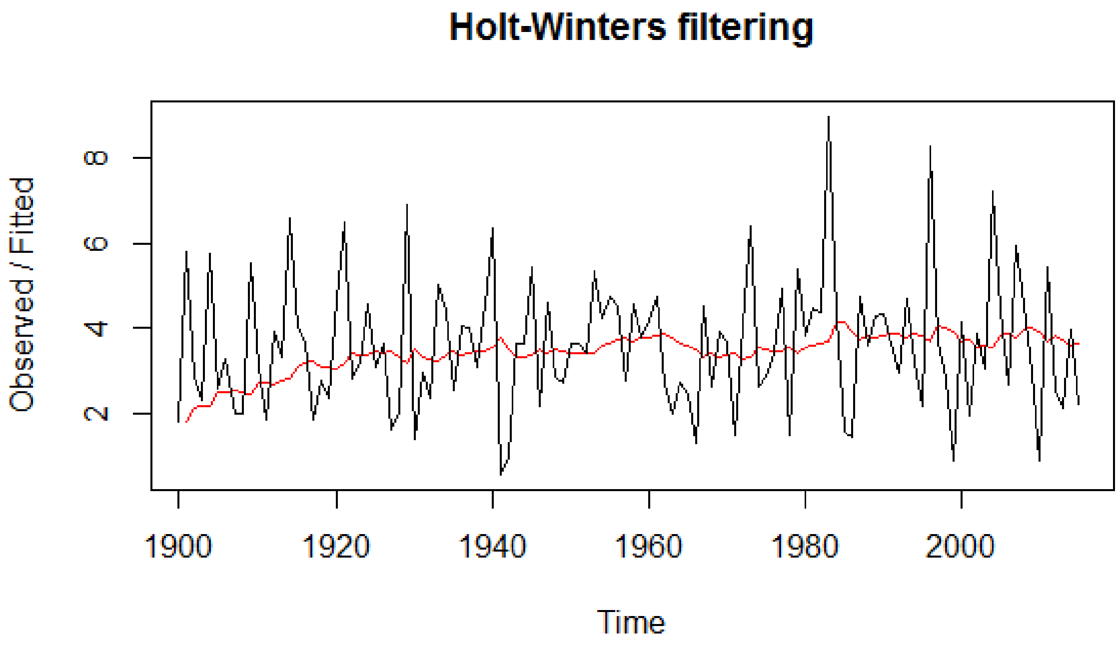
Smoothing parameter alpha used to create a smoothed sequence:

$$s_t = \alpha x_t + (1 - \alpha) s_{t-1}$$

and

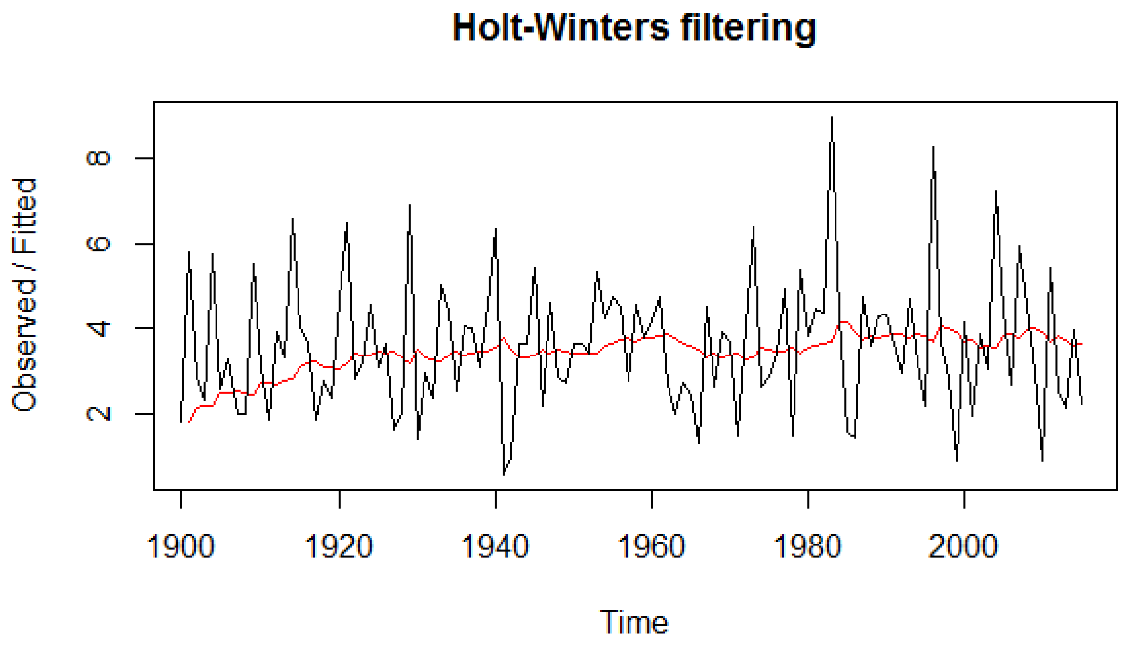
$$0 < \alpha < 1$$

The smaller the value of alpha, the more weight given to the previous values.



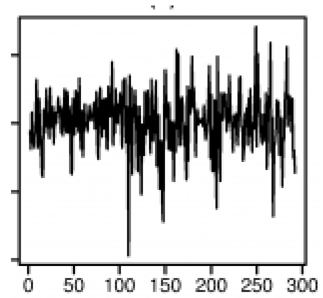
# Exponential Smoothing

The Holt-Winters method uses three smoothing equations-
one for the level, trend, and seasonal component,
with corresponding smoothing parameters alpha, beta, gamma.

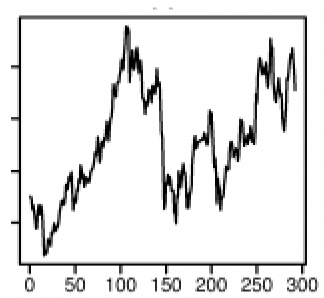


# TS Stationary vs Non-stationary

Stationary:
No change in trend or seasonal variations.
It “looks the same” everywhere.
Mean and variance pretty much constant.




Non-stationary:
Has a cyclical or changing trend or seasonal component.



# TS Stationary vs Non-stationary

***Differencing***: compute the differences between consecutive observations.

$$y'_t = y_t - y_{t-1}$$

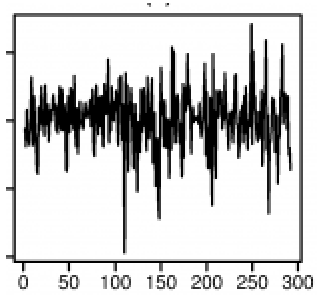
Makes a time series stationary by removing differences in the mean.

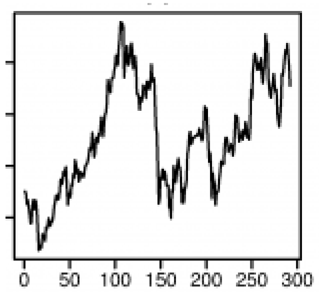
Several ways to difference.
Second order:

$$y''_t = y'_t - y'_{t-1}$$

Seasonal (where m is the number of seasons):

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



# TS Modeling- ARIMA

ARIMA:  AutoRegressive Moving Average.
Exponential smoothing is based on describing trend and seasonality components.
ARIMA describes the autocorrelations in the data. Works with ***Stationary*** TS.

$$ARIMA(p, d, q)$$

Parameters:
**p** is the order of the Autoregressive model
**d** is the degree of differencing
**q** is the order of the Moving-average model

# Lagged Variables

A linear model where the variables are time-lagged.

$$\mathbf{Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t\text{-}1} + \ldots + \beta_q X_{t\text{-}q} + e_t}$$

So, in the above model, the parameter $\beta_1$ describes the effect of the data point at time t-1, etc.

Also, $X_{t-1}$ is "lagged" $X_t$, and $X_2$ is "lagged" $X_3$, and so on.

The lag operator: $L(X_t) + X_{t-1}$, so, for example: $L(X_5) = X_4$

Of course, we can extend this to a lag period greater than 1.

$L_k(X_t) = X_{t-k}$, and then: $L_{20}(X_{25}) = X_5$

# TS Prediction

Prediction with 85% and 95% error bands