Time Series

Cross Sectional vs. Time Series Data

- Cross Sectional Data: Multiple objects observed at a particular point of time

Cross Sectional vs. Time Series Data

- Cross Sectional Data: Multiple objects observed at a particular point of time
- Examples: customers' behavioral data at today's update,companies' account balances at the end of the last year,patients' medical records at the end of the current month.

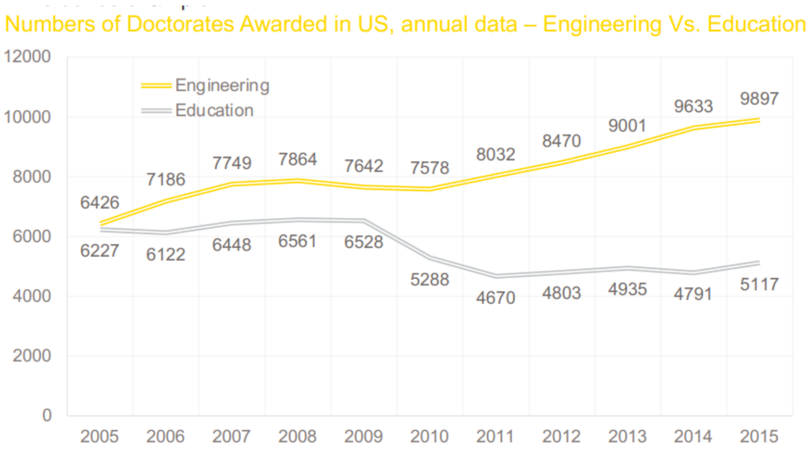
Cross Sectional vs. Time Series Data

- Time Series Data: One single object (product, country, sensor, ..) observed over multiple equally-spaced time periods

Cross Sectional vs. Time Series Data

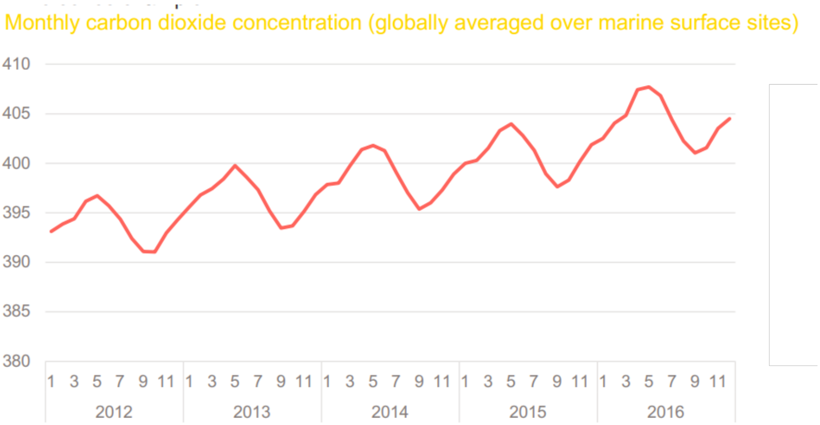
- Time Series Data: One single object (product, country, sensor, ..) observed over multiple equally-spaced time periods
- Examples: quarterly Italian GDP of the last 10 years, weekly supermarket sales of the previous year, yesterday's hourly temperature measurements.

Examples



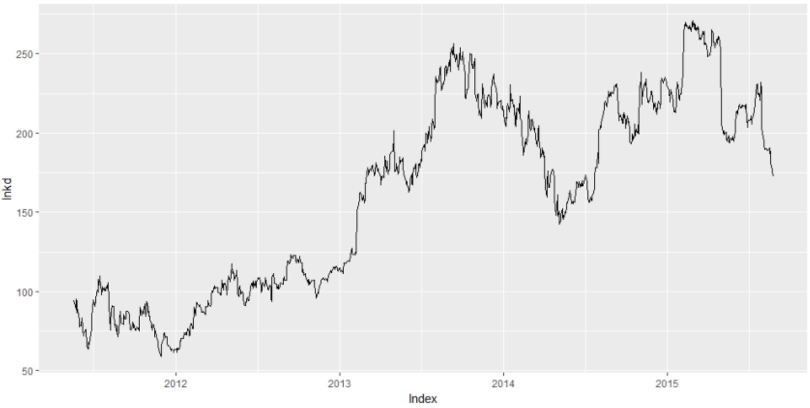
Numbers of Doctorates Awarded in US, annual data – Engineering Vs. Education
Engineering
Education
12000
10000
8000
6000
4000
2000
0
6426
7186
7749
7864
7642
7578
8032
8470
9001
9633
9897
6227
6122
6448
6561
6528
5288
4670
4803
4935
4791
5117
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015

Examples

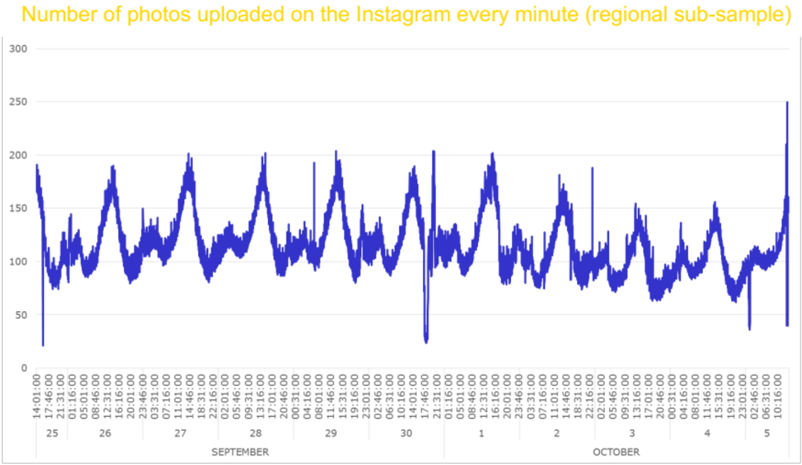

Monthly carbon dioxide concentration (globally averaged over marine surface sites)
410
405
400
395
390
385
380
1 3 5 7 9 11 1 3 5 7 9 11 1 3 5 7 9 11 1 3 5 7 9 11 1 3 5 7 9 11
2012
2013
2014
2015
2016

Examples

LinkedIn daily stock market closing price



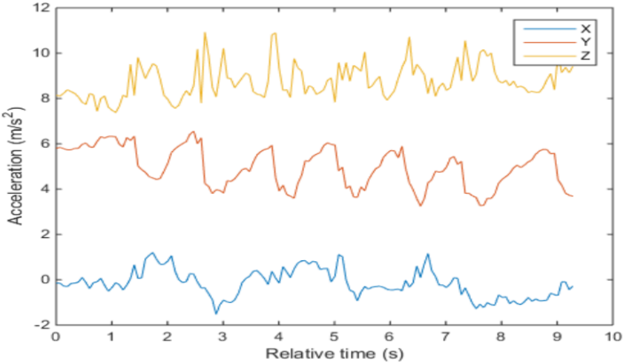
Examples



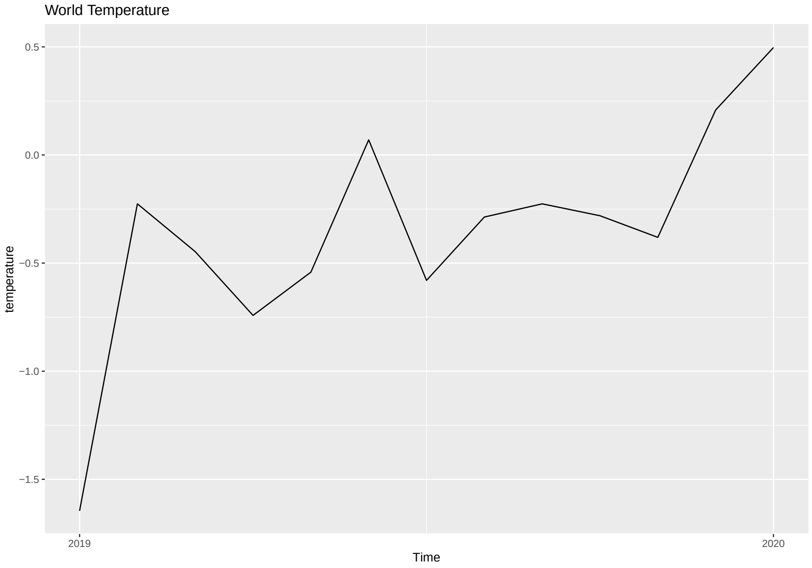
Number of photos uploaded on the Instagram every minute (regional sub-sample)
300
250
200
150
100
50
0
25
26
27
28
29
30
1
2
3
4
5
SEPTEMBER
OCTOBER

Examples

Acceleration detected by a smartphone sensors during a workout session (10 seconds)



Examples



What to do with time series?

- Understanding of specific series features or pattern
- Forecasting

Smoothing

Smoothing

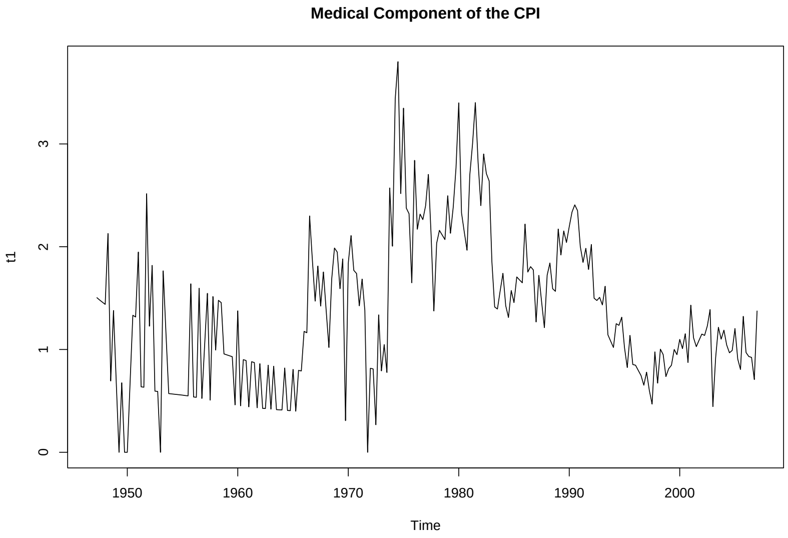
- Smoothing is usually done to reveal the series patterns and trends.

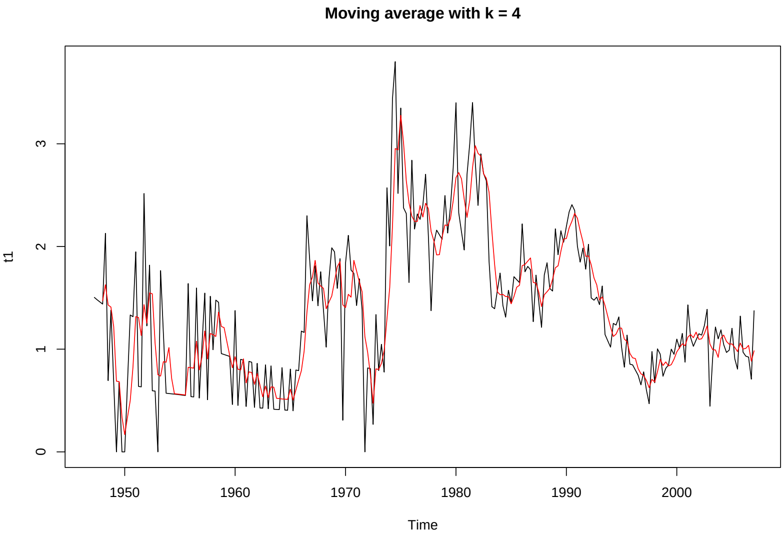
Simple Moving Average Smoothing

- Moving Average (MA) creates a new series by averaging the most recent observations from the original series.
- MA(k) creates s_t as follows.

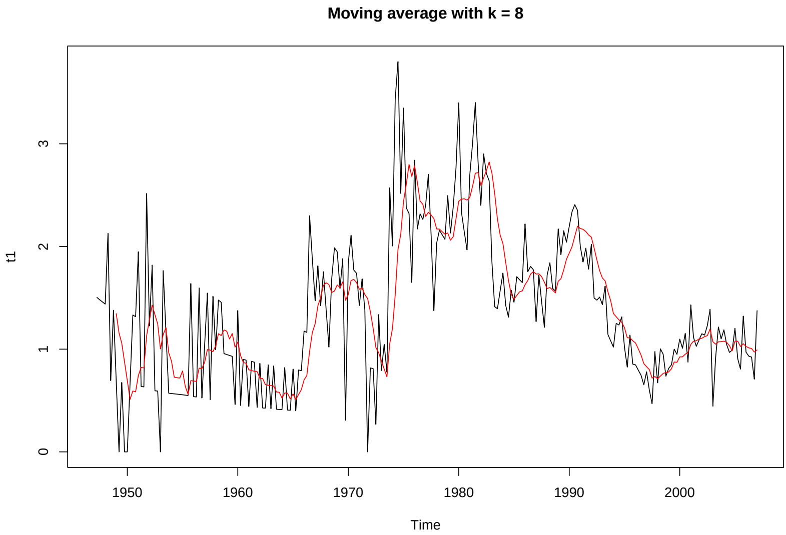
$$s_t = \frac{y_t + y_{t-1} + \dots + y_{t-k+1}}{k}$$

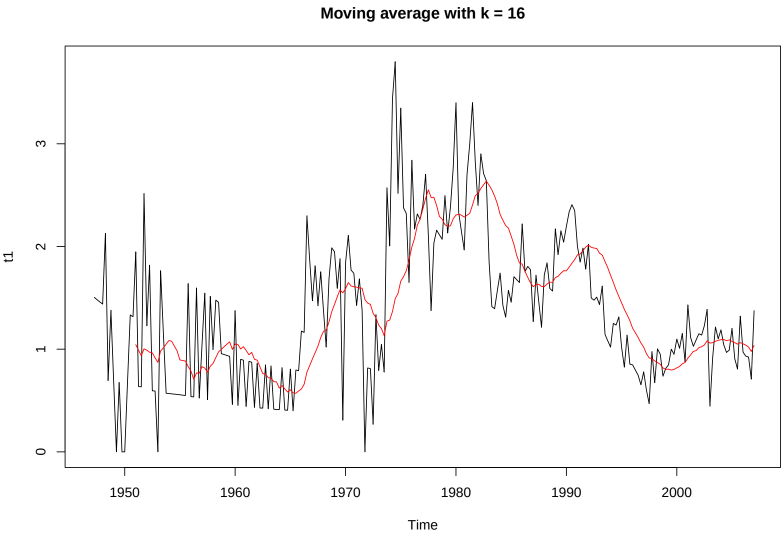
- Larger k will smooth the series more strongly


Medical Component of the CPI
t1
0
1
2
3
1950
1960
1970
1980
1990
2000
Time

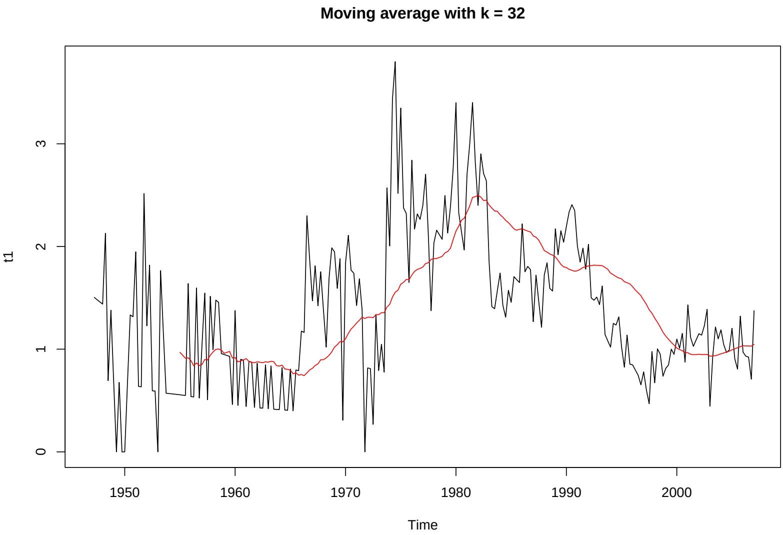


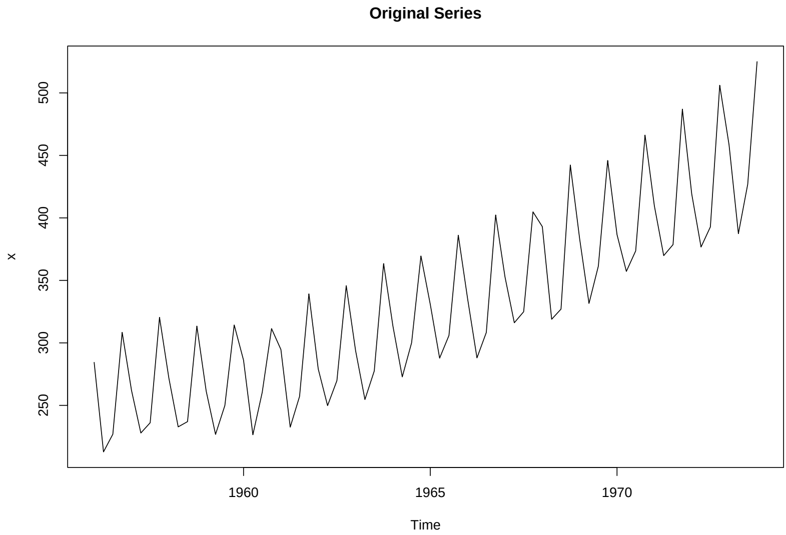
Moving average with k = 4
t1
0
1
2
3
1950
1960
1970
1980
1990
2000
Time


Moving average with k = 8
t1
Time
0
1
2
3
1950
1960
1970
1980
1990
2000

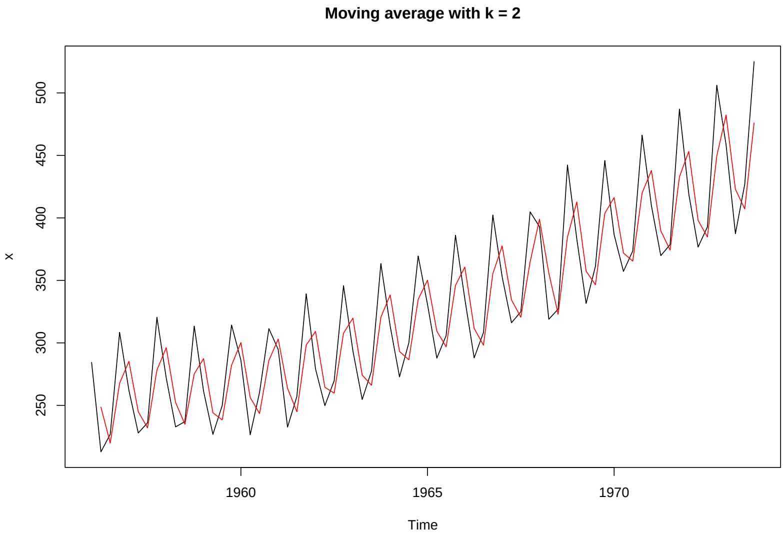


Moving average with k = 16
t1
0
1
2
3
Time
1950
1960
1970
1980
1990
2000

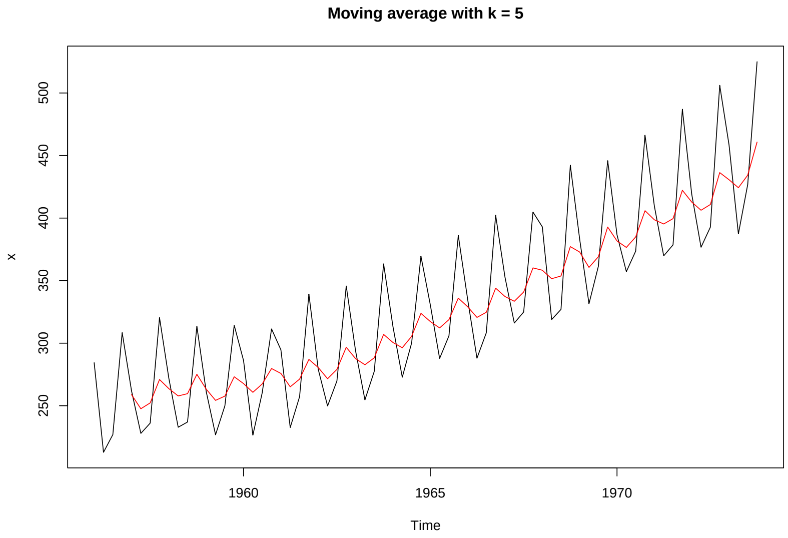

Moving average with k = 32
t1
Time
0
1
2
3
1950
1960
1970
1980
1990
2000

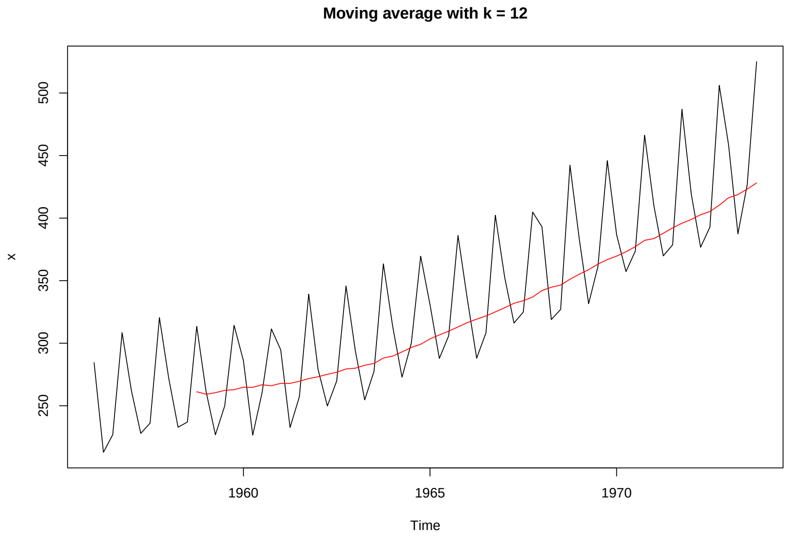


Original Series
x
500
450
400
350
300
250
1960
1965
1970
Time



Moving average with k = 2
x
250
300
350
400
450
500
Time
1960
1965
1970


Moving average with k = 5
x
500
450
400
350
300
250
1960
1965
1970
Time


Moving average with k = 12
x
500
450
400
350
300
250
1960
1965
1970
Time

Forecasting

- We can use smoothing for forecasting
- We have

$$
\begin{aligned}
\hat{s}_t &= \frac{y_t + y_{t-1} + \ldots + y_{t-k+1}}{k} \\
&= \frac{y_t + y_{t-1} + \ldots + y_{t-k+1} + y_{t-k} - y_{t-k}}{k} \\
&= \frac{y_t + \left(y_{t-1} + \ldots + y_{t-k+1} + y_{t-k} \right) - y_{t-k}}{k} \\
&= \frac{y_t + k\hat{s}_{t-1} - y_{t-k}}{k} \\
&= \hat{s}_{t-1} + \frac{y_t - y_{t-k}}{k}
\end{aligned}
$$

Forecasting

- If there is no trend in y_t the second term $(y_t - y_{t-k})/k$ can be ignored
- Forecasting l lead time into future by

$$\hat{y}_{T+l} = \hat{s}_T$$

- If there is a linear trend in a series, we can use the double moving average to estimate the trend (slope)

Double MA

7. Double MA

- Linear trend time series:

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

- Step 1: Smooth the series

$$\hat{s}_t^{(1)} = \frac{y_t + y_{t-1} + ... + y_{t-k+1}}{k}$$

- Step 2: Smooth the smoothed series

$$\hat{s}_t^{(2)} = \frac{\hat{s}_t^{(1)} + \hat{s}_{t-1}^{(1)} + ... + \hat{s}_{t-k+1}^{(1)}}{k}$$

- Step 3: Calculate the trend

$$b_1 = \hat{\beta}_1 = \frac{2}{k-1}\left(\hat{s}_T^{(1)} - \hat{s}_T^{(2)}\right)$$

Forecasting

- Forecasting l lead time into future by

$$\hat{y}_{T+l} = \hat{s}_T + b_1 \cdot l$$

You are given the following time series

t	1	2	3	4	5
y_t	1	3	5	8	13

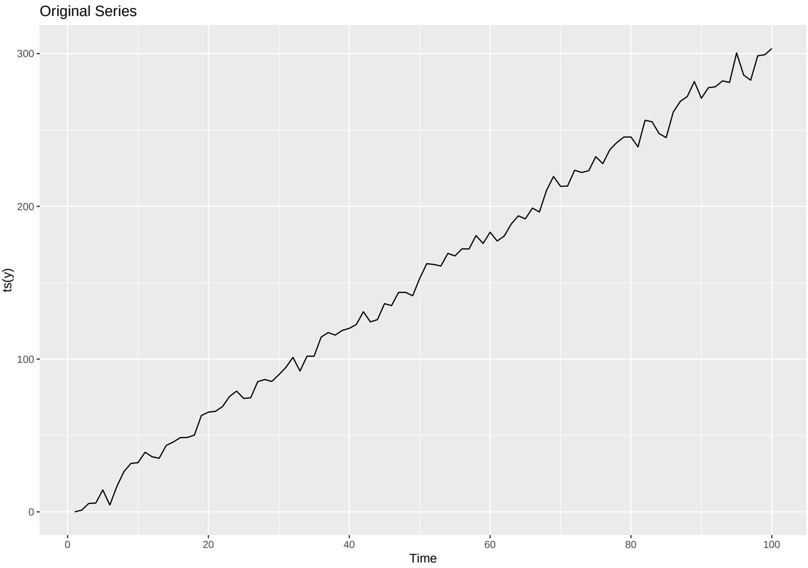
- Forecasting y_7 using simple moving average with $k = 2$
- Forecasting y_7 using double moving average with $k = 2$

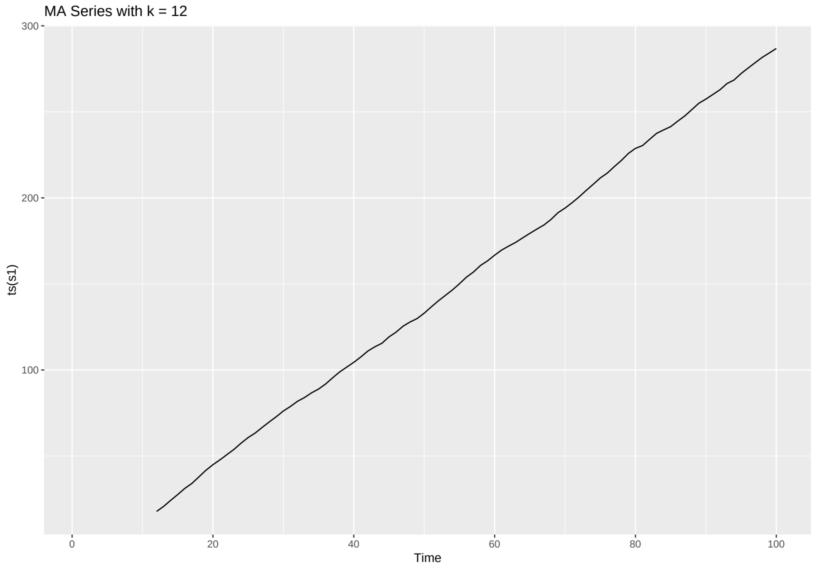
Example

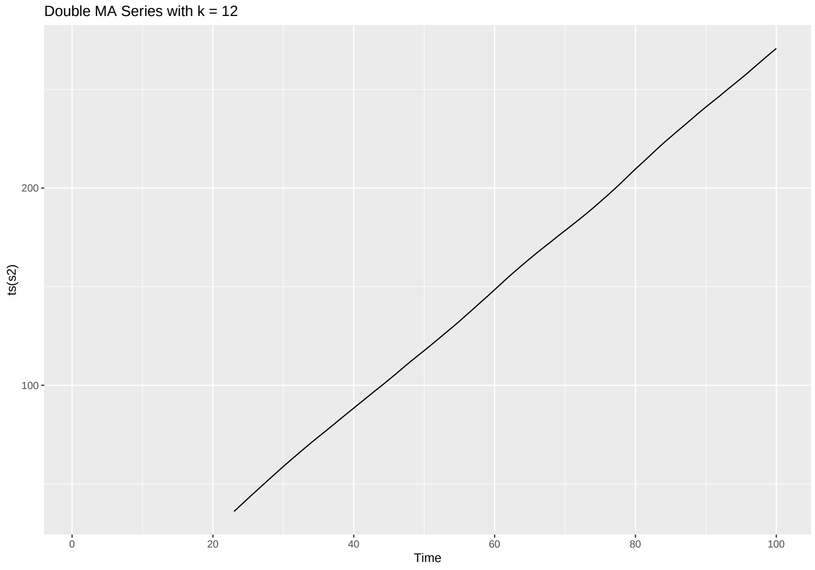
- We simulate 100 data points ($T = 100$) of

$$y_t = 1 + 3t + \epsilon,$$

where, $\epsilon \sim N(0, 5^2)$.


Original Series
ts(y)
300
200
100
0
0
20
40
60
80
100
Time


MA Series with k = 12
300
200
100
ts(s1)
0
20
40
60
80
100
Time


Double MA Series with k = 12
ts(s2)
200
100
0
20
40
60
80
100
Time

- Using the above steps, the estimated trend is $b_1 = 2.92$
- The forecast for the next points from y_{100} is

$$\hat{y}_{100+l} = \hat{s}_{100} + b_1 \cdot l = \hat{s}_{100} + 2.92 \cdot l$$

Exponential Smoothing

Exponential Smoothing

- MA distributes the weight equally to the recent observations
- Exponential Smoothing controls the weights of the recent observations by w

$$\hat{s}_t = \frac{y_t + w y_{t-1} + w^2 y_{t-2} + ... + w^t y_0}{1/(1-w)}$$

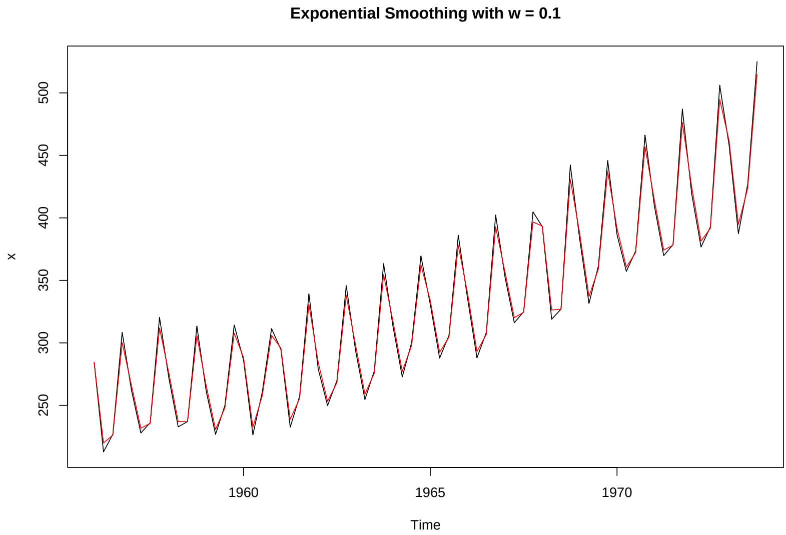
- Smaller w ($w \to 0$) gives higher weights to the more recent observations
- Smaller w smooths the series more lightly.

Exponential Smoothing

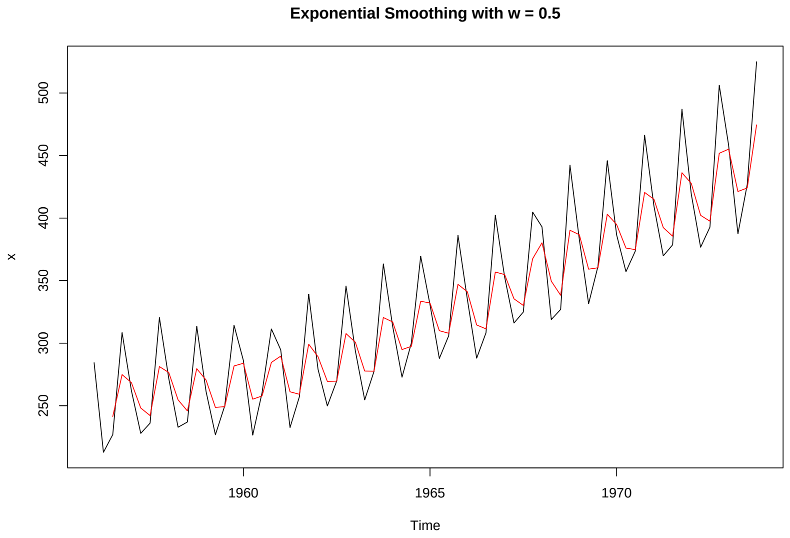
- We have

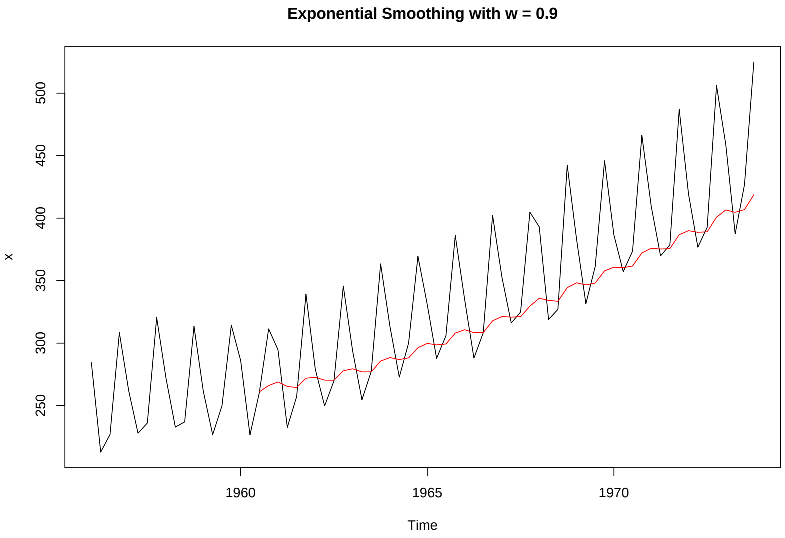
$$\begin{aligned}\hat{s}_t &= \hat{s}_{t-1} + (1-w)(y_t - \hat{s}_{t-1}) \\ &= (1-w)y_t + w\hat{s}_{t-1}\end{aligned}$$

- When $w \to 0$, $\hat{s}_t \to y_t$, or little smoothing has taken



Exponential Smoothing with w = 0.1
x
Time
250
300
350
400
450
500
1960
1965
1970


Exponential Smoothing with w = 0.5
x
250
300
350
400
450
500
1960
1965
1970
Time



Exponential Smoothing with w = 0.9
500
450
400
350
300
250
x
1960
1965
1970
Time

Double Exponential Smoothing

Double Exponential Smoothing

We can use double smoothing to identify the trend and forecast linear trend time series as follows.

- Step 1: Create a smoothed series: $\hat{s}_t^{(1)} = (1-w)y_t + w\hat{s}_{t-1}^{(1)}$
- Step 2: Create a double smoothed series: $\hat{s}_t^{(2)} = (1-w)\hat{s}_t^{(1)} + w\hat{s}_{t-1}^{(2)}$
- Step 3: Estimate the trend:

$$b_1 = \frac{1-w}{w}(\hat{s}_T^{(1)} - \hat{s}_T^{(2)})$$

- Step 4: Forecast

$$\hat{y}_{T+l} = 2\hat{s}_T^{(1)} - \hat{s}_T^{(2)} + b_1 \cdot l$$

Example

You are given the following time series

t	1	2	3	4	5
y_t	1	3	5	8	13

- Forecasting y_7 using exponential smoothing with $w = .8$
- Forecasting y_7 using double exponential smoothing with $w = .8$

Citation

- Photos are taken from KNIME Hub