

# 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

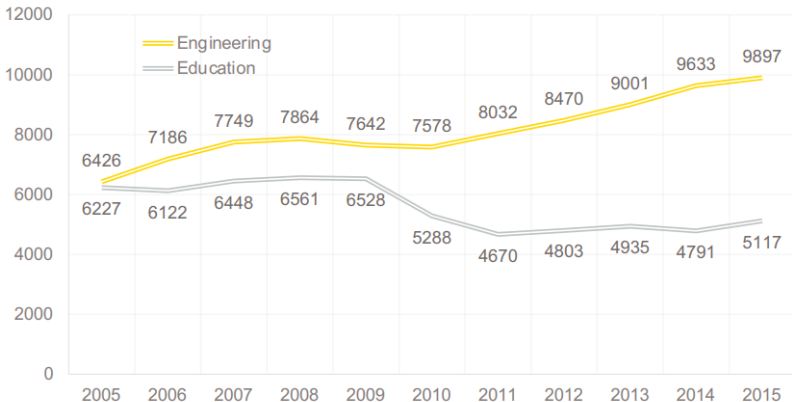
- ▶ Time Series Data: One single object (product, country, sensor, ..) observed over multiple equally-spaced time periods

# 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.

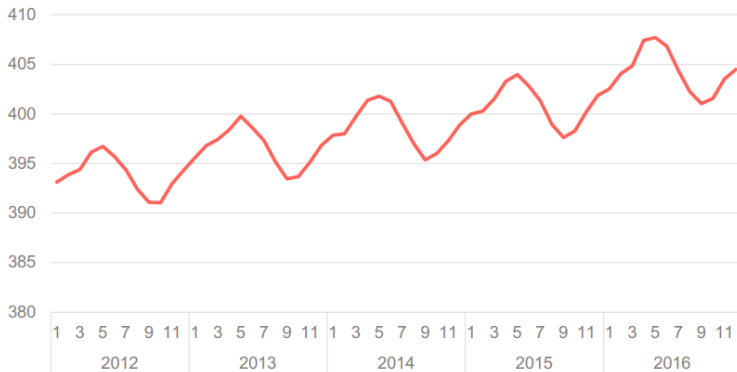
# Examples

Numbers of Doctorates Awarded in US, annual data – Engineering Vs. Education



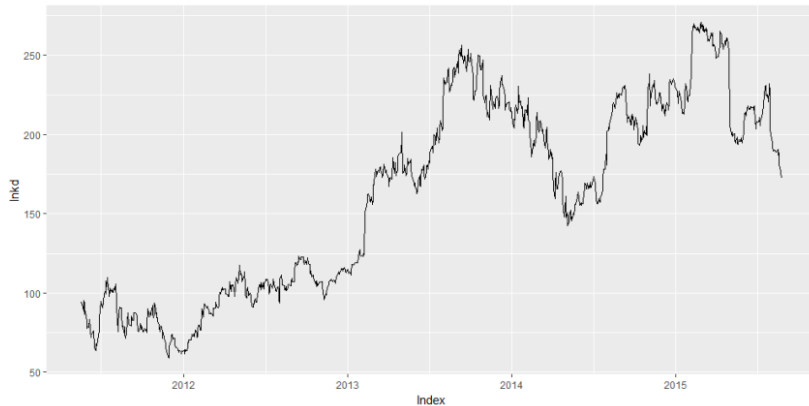
# Examples

Monthly carbon dioxide concentration (globally averaged over marine surface sites)



# Examples

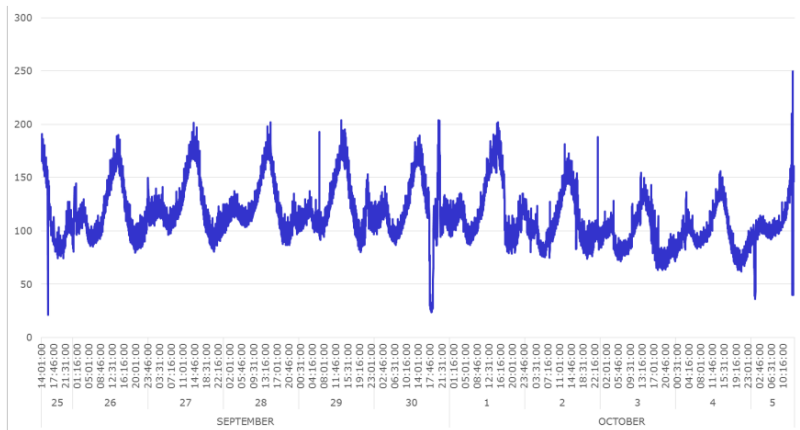
LinkedIn daily stock market closing price





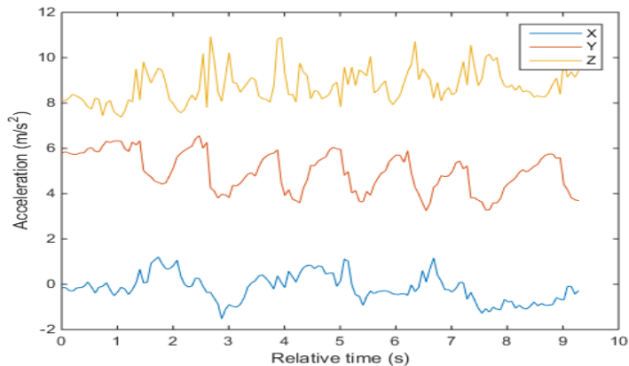
# Examples

Number of photos uploaded on the Instagram every minute (regional sub-sample)

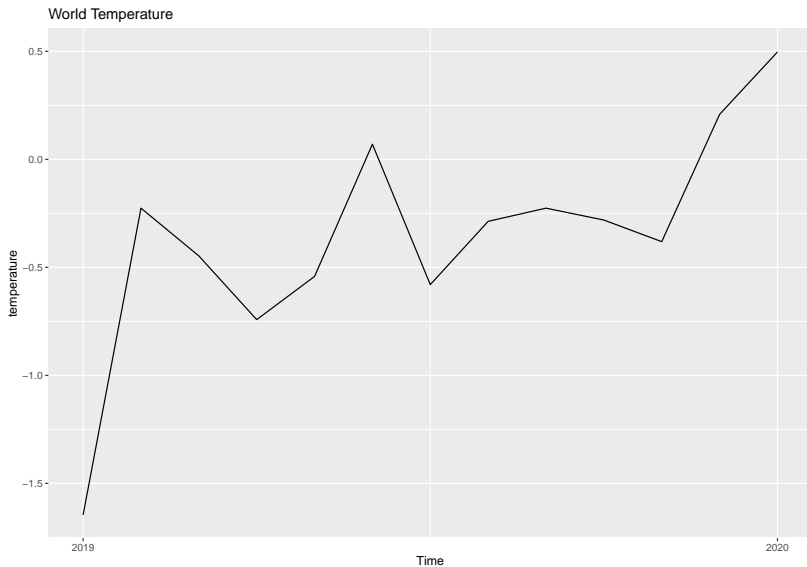


# Examples

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.

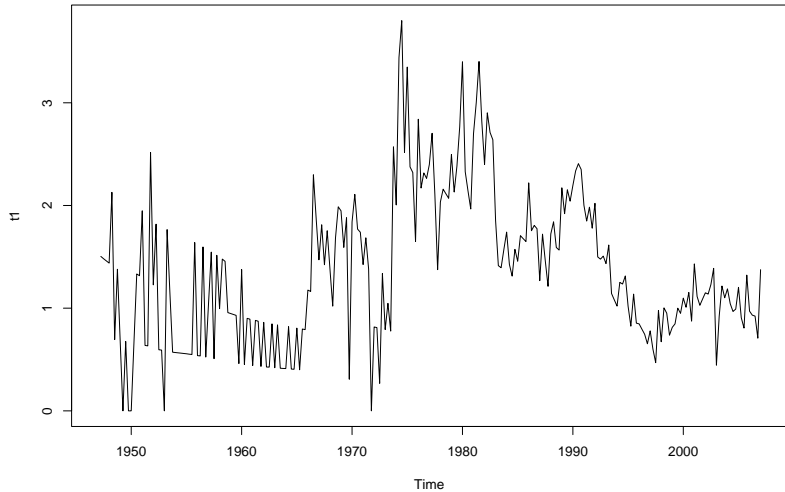
# Simple Moving Average Smoothing

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

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

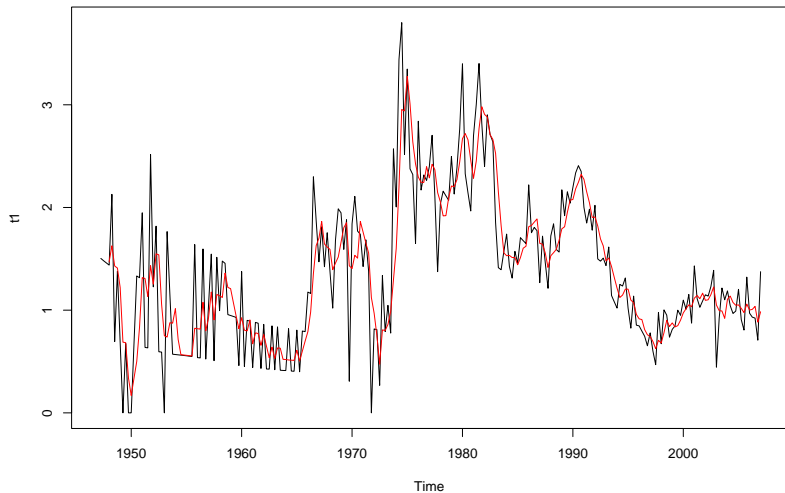
- ▶ Larger  $k$  will smooth the series more strongly

### Medical Component of the CPI

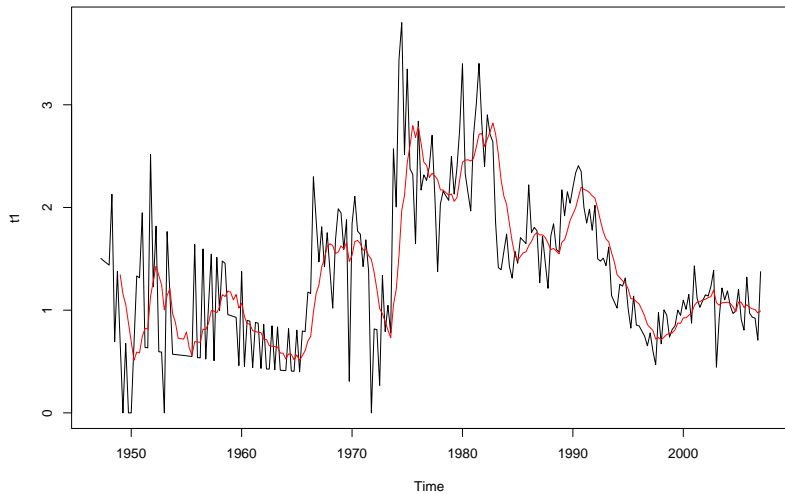




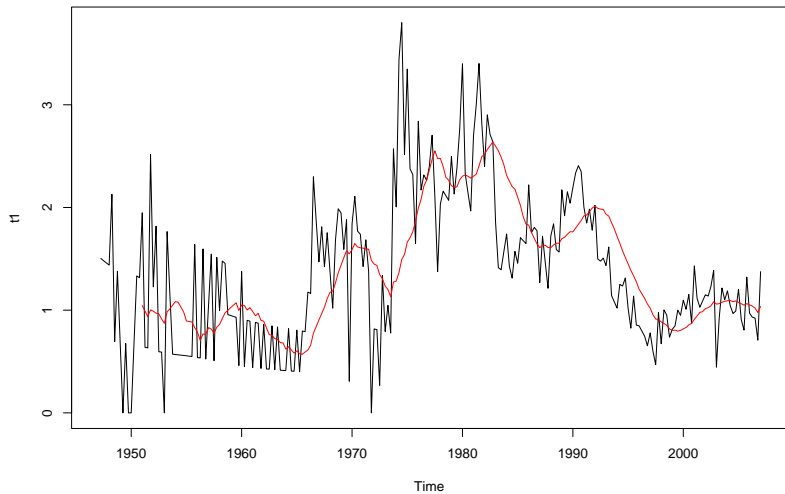
**Moving average with  $k = 4$**



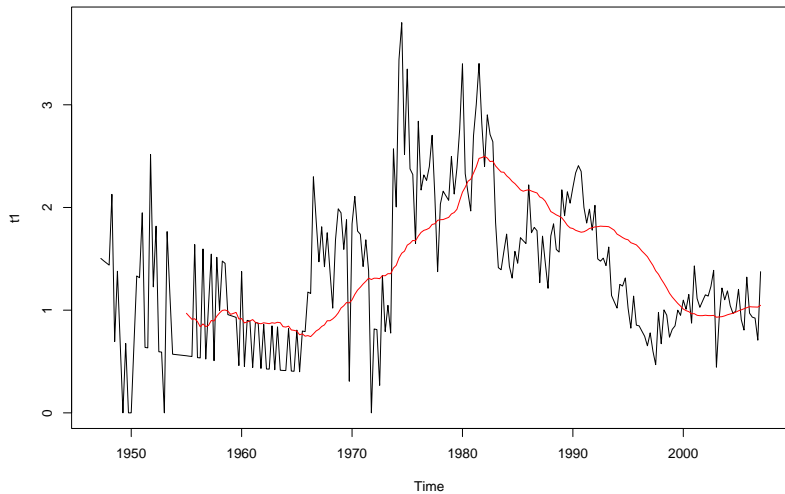
**Moving average with  $k = 8$**



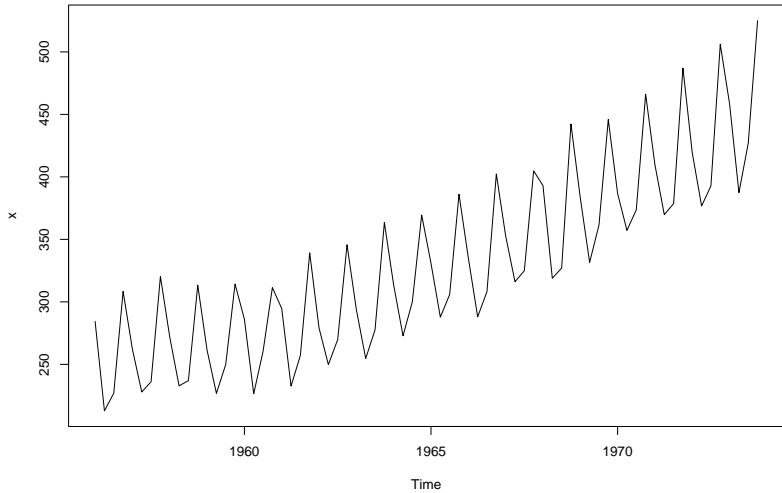
**Moving average with  $k = 16$**



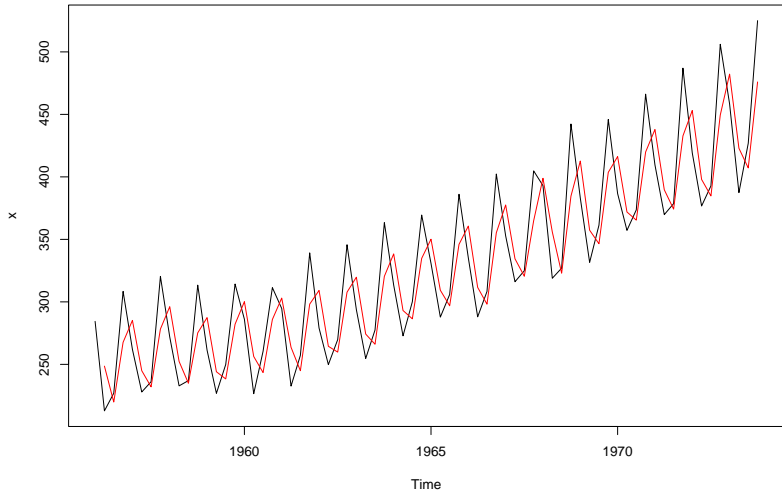
**Moving average with  $k = 32$**



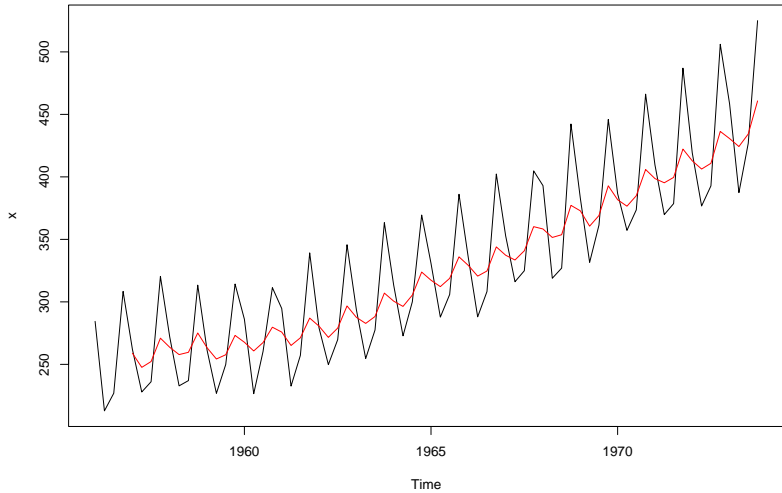
**Original Series**



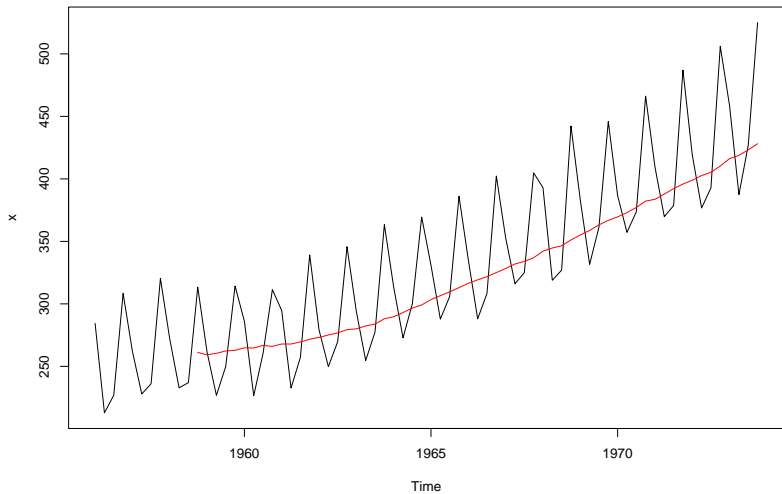
**Moving average with  $k = 2$**



**Moving average with  $k = 5$**



**Moving average with  $k = 12$**





# Forecasting

- ▶ We can use MA smoothing for forecasting
- ▶ We have

$$\begin{aligned}s_t &= \frac{y_t + y_{t-1} + \dots + y_{t-k+1}}{k} \\&= \frac{y_t + y_{t-1} + \dots + y_{t-k+1} + y_{t-k} - y_{t-k}}{k} \\&= \frac{y_t + \left( y_{t-1} + \dots + y_{t-k+1} + y_{t-k} \right) - y_{t-k}}{k} \\&= \frac{y_t + ks_{t-1} - y_{t-k}}{k} \\&= s_{t-1} + \frac{y_t - y_{t-k}}{k}\end{aligned}$$

# 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

$$\hat{y}_{T+l} = s_T$$

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

Double MA

# Double MA

- ▶ Linear trend time series:

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

- ▶ Step 1: MA Smooth the series

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

- ▶ Step 2: MA Smooth the smoothed series

$$s_t^{(2)} = \frac{s_t^{(1)} + s_{t-1}^{(1)} + \dots + s_{t-k+1}^{(1)}}{k}$$

- ▶ Step 3: Calculate the linear trend/slope

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

# Forecasting

- Forecasting  $l$  lead time into future by

$$\hat{y}_{T+l} = s_T^{(1)} + b_1 \cdot l$$

You are given the following time series

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

- ▶ Forecasting  $y_6$  using simple moving average with  $k = 2$
- ▶ Forecasting  $y_6$  using double moving average with  $k = 2$



## Example

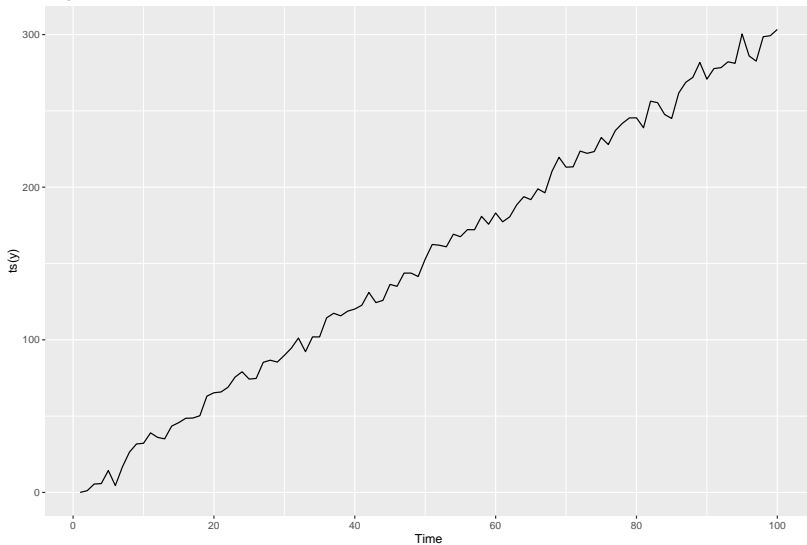
- ▶ We simulate 100 data points ( $T = 100$ ) of

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

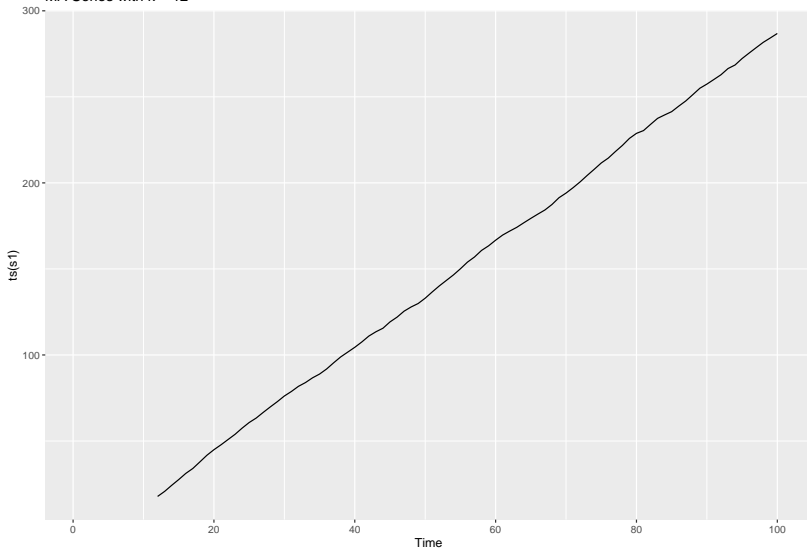
where,  $\epsilon \sim N(0, 5^2)$ .



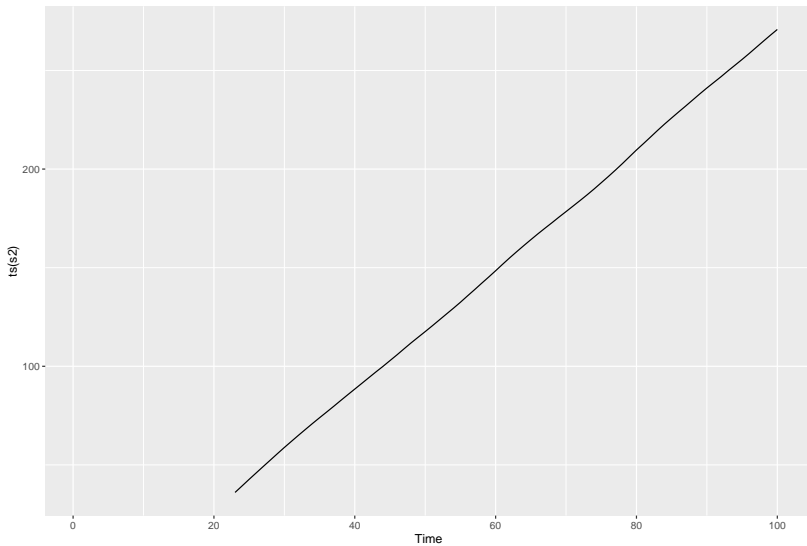
Original Series



MA Series with  $k = 12$



Double MA Series with  $k = 12$



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

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