CS699

Introduction to Time Series Analysis

References

- This note is based on and uses the material from the following sources:
  - *NIST/SEMATECH e-Handbook of Statistical Methods*, http://www.itl.nist.gov/div898/handbook/, NIST, U.S. Department of Commerce, April 2012
  - G. Shmueli, N.R. Patel, and P.C. Bruce, Data Mining for Business Intelligence, 2nd Ed., Wiley, 2010
  - J. Han, M. Kamber, and J. Pei, Data Mining Concepts and Techniques, 3rd Ed., Morgan Kaufmann, 2012

Introduction

- Time series: *An ordered sequence of values of a variable at equally spaced time intervals*.

- Usage: The usage of time series models is twofold
  - Descriptive: Obtain an understanding of the underlying forces and structure that produced the observed data
  - Predictive: Fit a model and use it to forecast future values

Introduction

- Applications:
  - Economic Forecasting
  - Sales Forecasting
  - Budgetary Analysis
  - Stock Market Analysis
  - Inventory Studies
  - Workload Projections
  - Utility Studies
  - Census Analysis
  - and many more …

Introduction

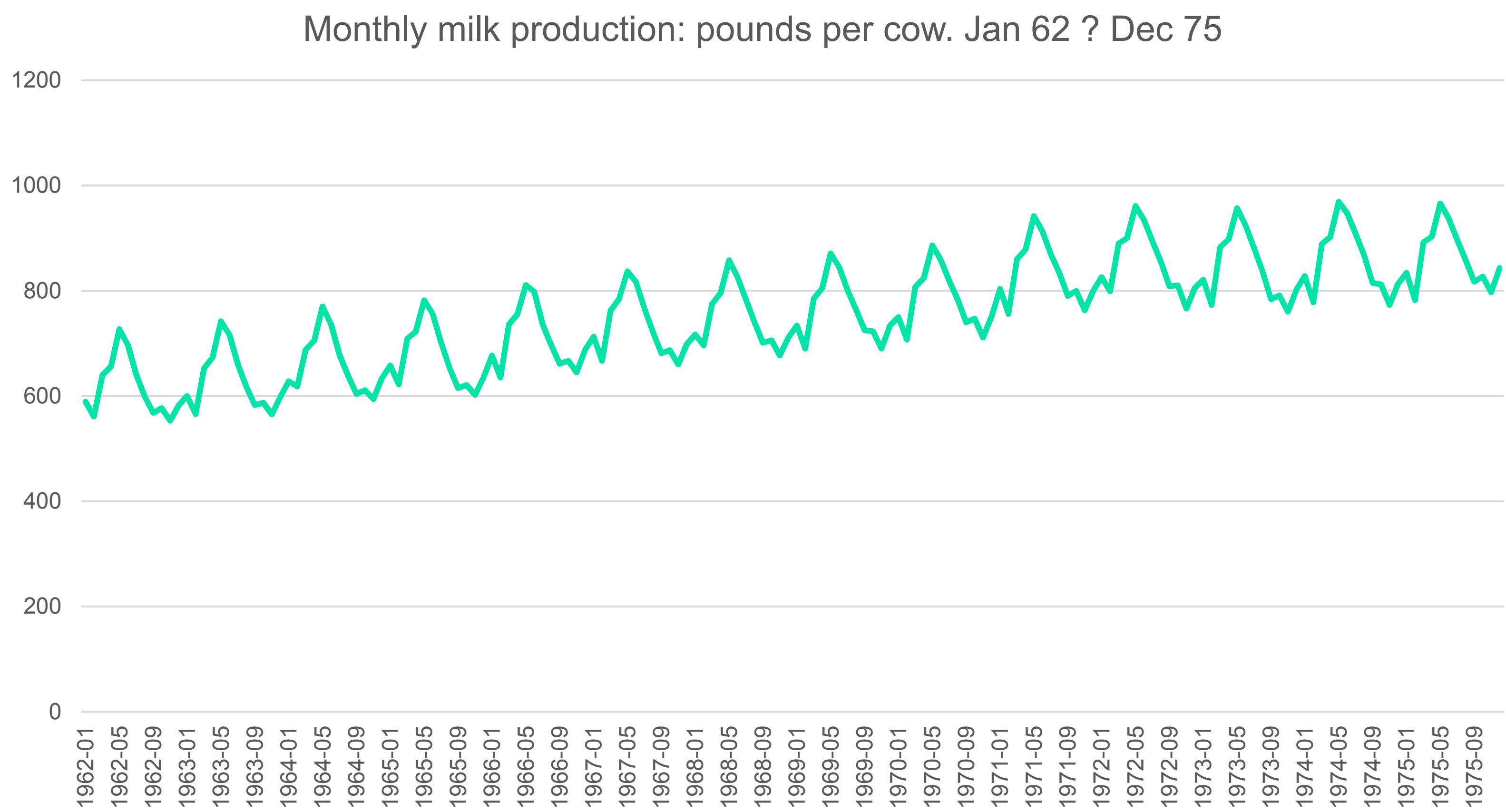
- Methods:
  - Regression
  - Smoothing
  - Hybrid
- Will discuss regression and smoothing with univariate time series

Introduction

- Time Series Components
  - Level: average value of the time series
  - Trend: general direction of the change of values over time
  - Seasonality: identical patterns are observed during corresponding seasons of successive years
  - Noise: random variations or sporadic changes due to chance events

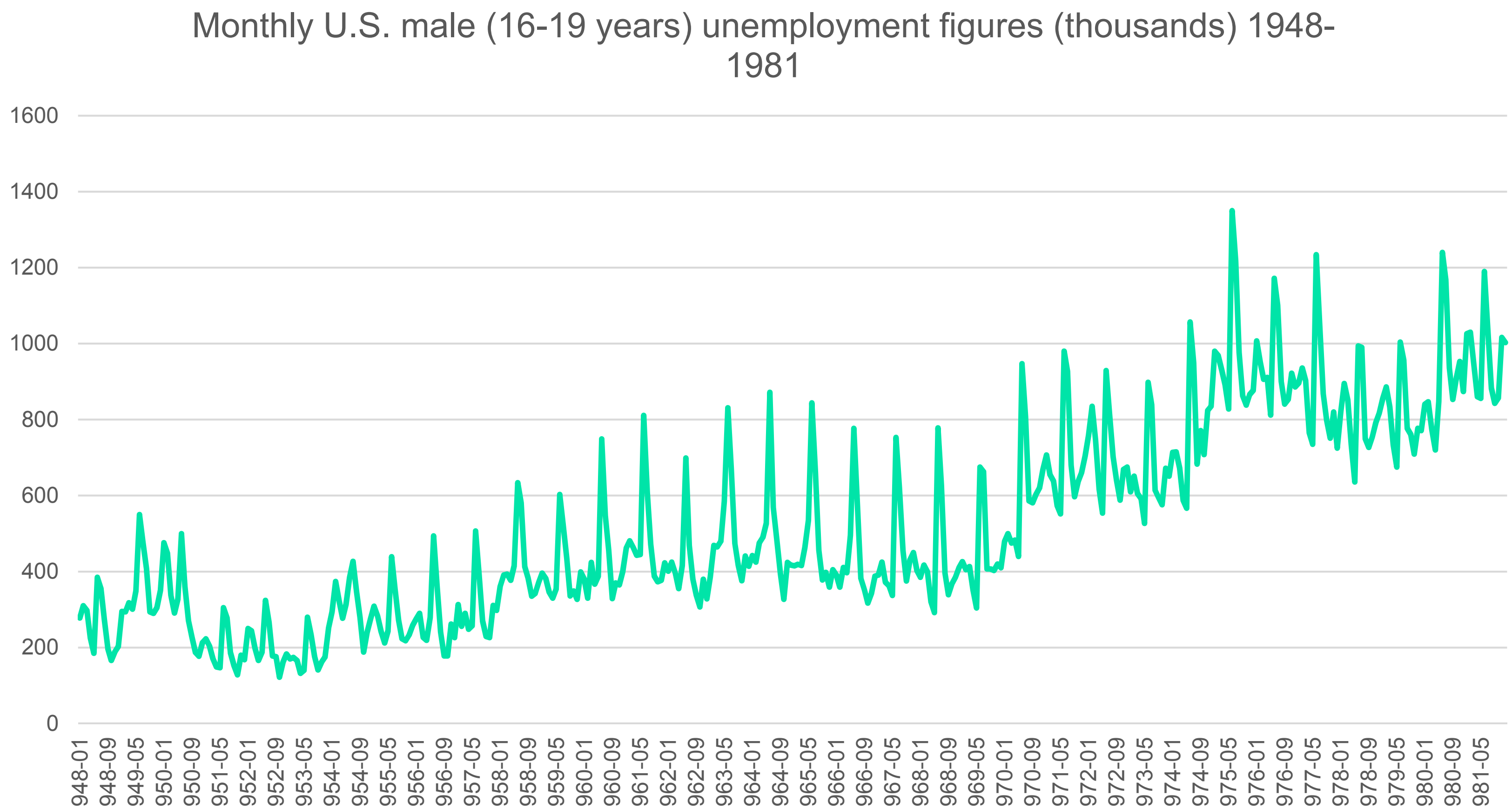
Introduction

- Example:



Introduction

- Example:



Forecasting by Regression

- A regression model is generated from the data and used for future prediction.
- Can capture
  - Linear trend
  - Exponential trend
  - Polynomial trend

Forecasting by Regression

- Linear trend
  - Output variable Y is expressed as a linear function of input variable (or predictor variable) X.

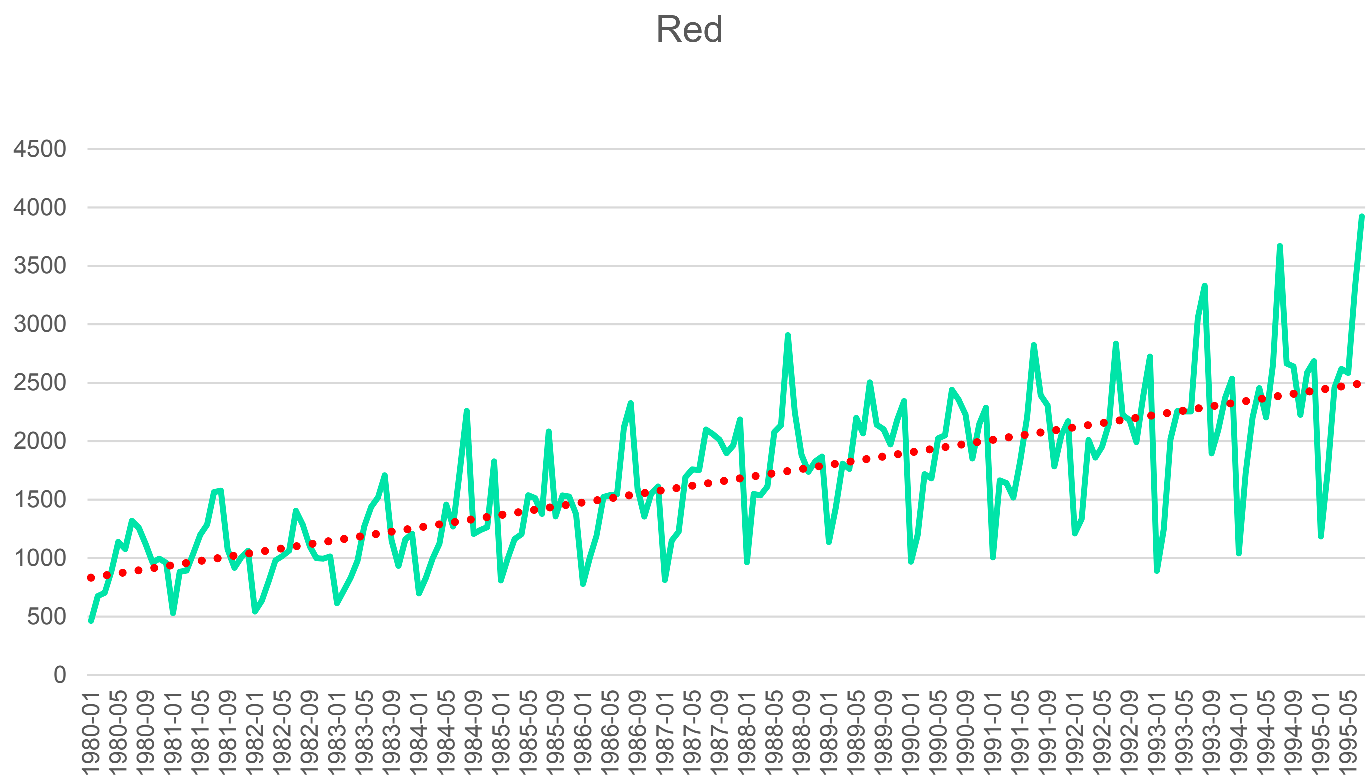
$$Y_t = b_0 + b_1 X_t + e, \text{ where}$$

b_0 and b_1 are regression coefficients

e is a noise (also called white noise)

Forecasting by Regression

- Linear trend



Forecasting by Regression

- Exponential trend
  - Captures multiplicative increase/decrease of time series values over time.

$$\log(Y_t) = b_0 + b_1 X_t + e$$

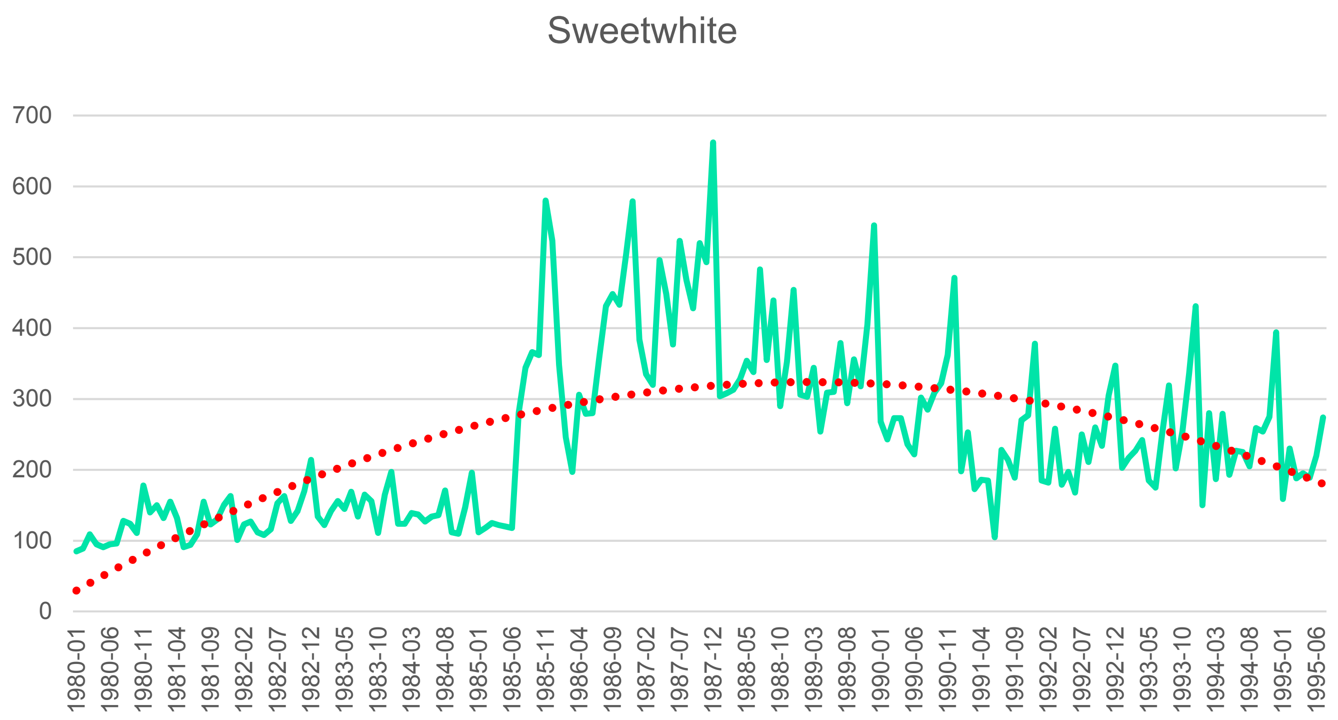
Forecasting by Regression

- Polynomial trend
  - Trend follows a polynomial shape.

$$Y_t = b_0 + b_1 X_t + b_2 X_t^2 + e$$

Forecasting by Regression

- Polynomial trend



Forecasting by Smoothing

- Data driven (i.e., not based on a theoretical model such as a regression model)
- Data taken over time contains some form of random variations
- Smoothing reduces or cancels the effect due to random variations
- Smoothing can reveal more clearly the underlying trend, seasonal and cyclic components.
- Will discuss moving average and single exponential smoothing

Moving Average

- Forecasting with simple average

Monthly Supply

Supplier	$	Error	Error Squared
1	9	-1	1
2	8	-2	4
3	9	-1	1
4	12	2	4
5	9	-1	1
6	12	2	4
7	11	1	1
8	7	-3	9
9	13	3	9
10	9	-1	1
11	11	1	1
12	10	0	0

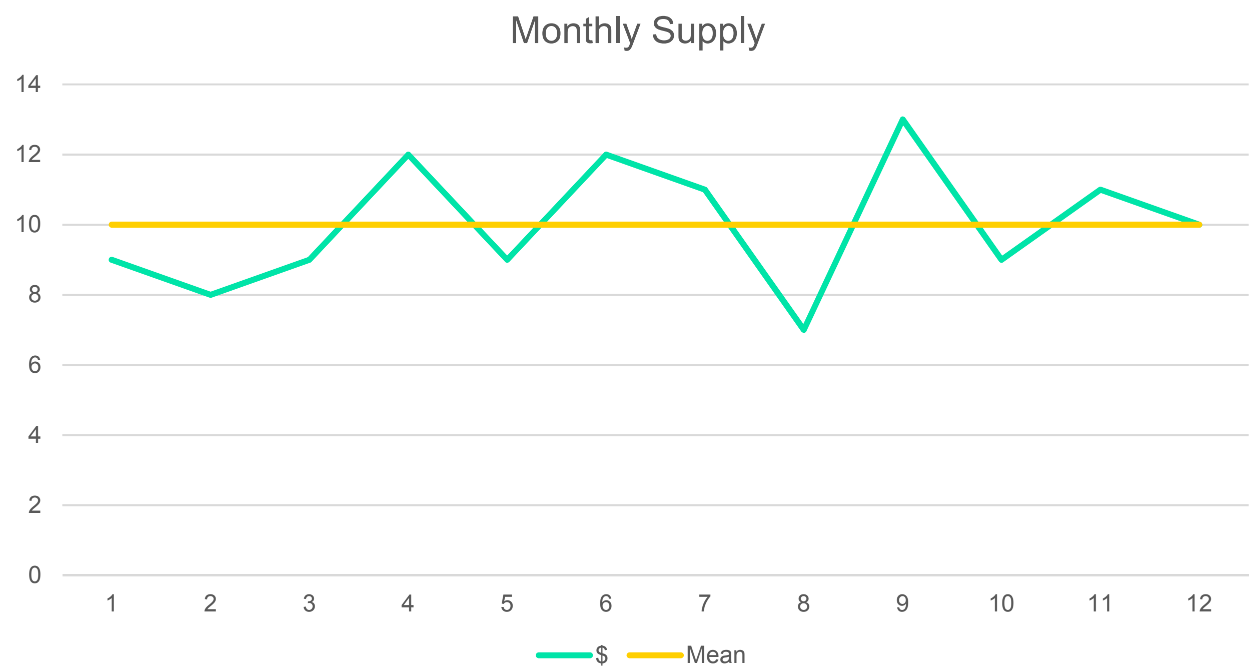
Estimator (mean) = 10

SSE = 36

MSE = 36/12 = 3

Moving Average

- Forecasting with simple average



Moving Average

- Forecasting with simple average

Income before tax of a company

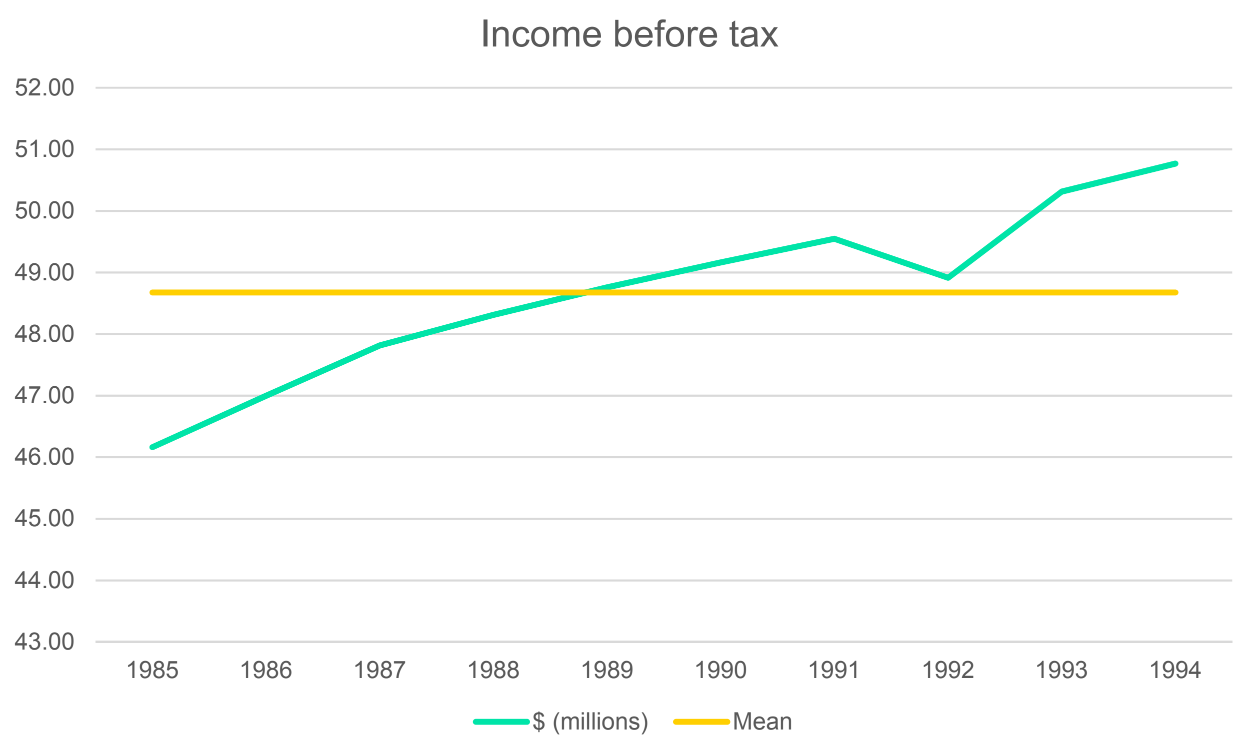
Year	$ (millions)	Error	Squared Error
1985	46.163	-2.513	6.313
1986	46.998	-1.678	2.814
1987	47.816	-0.860	0.739
1988	48.311	-0.365	0.133
1989	48.758	0.082	0.007
1990	49.164	0.488	0.239
1991	49.548	0.872	0.761
1992	48.915	0.239	0.057
1993	50.315	1.639	2.688
1994	50.768	2.092	4.378

Estimator (Mean) = 48.676

MSE = 1.8129

Moving Average

- Forecasting with simple average



Moving Average

- Forecasting with moving average

  Moving average with window = 3

  $M_t = (M_t + M_{t-1} + M_{t-2})/3$

Year	$ (millions)	Mean	MA3
1985	46.16	48.68	
1986	47.00	48.68	
1987	47.82	48.68	46.99233
1988	48.31	48.68	47.70833
1989	48.76	48.68	48.295
1990	49.16	48.68	48.74433
1991	49.55	48.68	49.15667
1992	48.92	48.68	49.209
1993	50.32	48.68	49.59267
1994	50.77	48.68	49.99933

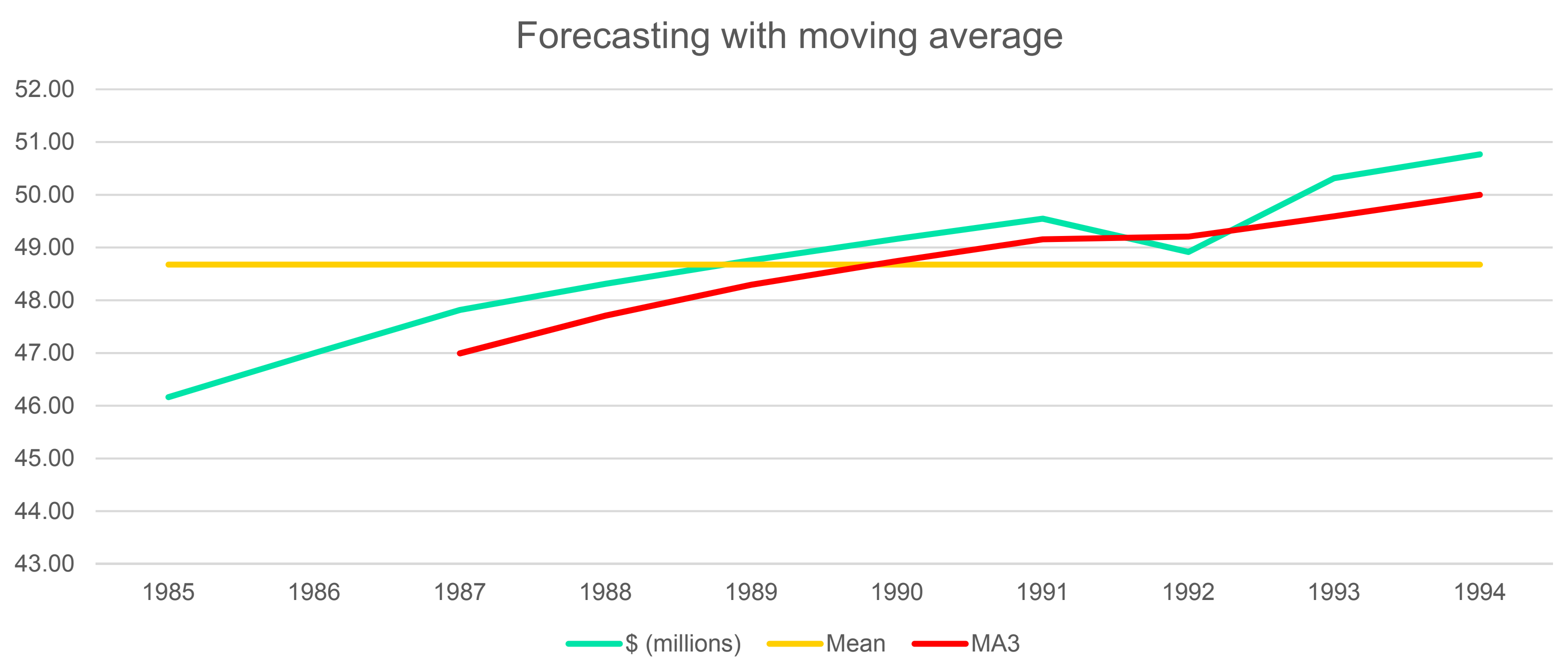
MSE = 0.313

Moving Average

- Forecasting with moving average

  Moving average with window = 3

  $M_t = (M_t + M_{t-1} + M_{t-2})/3$



Moving Average

- Centered moving average

  Centered moving average with window = 3

  $M_t = (M_{t+1} + M_t + M_{t-1})/3$

Year	$ (millions)	Mean	MA3	CMA3
1985	46.16	48.68		
1986	47.00	48.68		46.99233
1987	47.82	48.68	46.99233	47.70833
1988	48.31	48.68	47.70833	48.295
1989	48.76	48.68	48.295	48.74433
1990	49.16	48.68	48.74433	49.15667
1991	49.55	48.68	49.15667	49.209
1992	48.92	48.68	49.209	49.59267
1993	50.32	48.68	49.59267	49.99933
1994	50.77	48.68	49.99933	

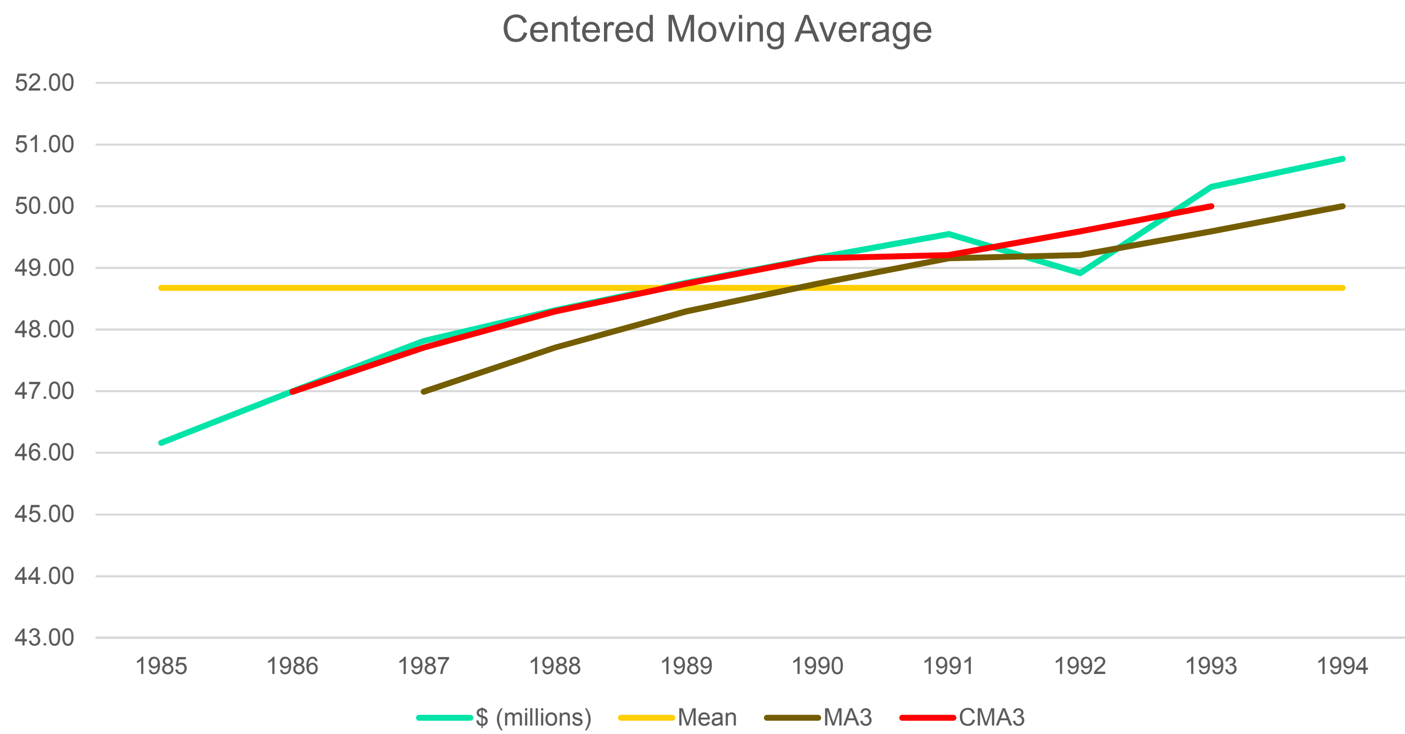
MSE = 0.086

Moving Average

- Centered moving average

  Moving average with window = 3

  $M_t = (M_{t+1} + M_t + M_{t-1})/3$



Moving Average

- Centered moving average is good for visualizing trends because it suppresses seasonality and noise.
- Not for forecasting because future values are not known.

Moving Average

- Window size
  - Large window: reveals more global trends
  - Small window: reveals local trends
  - Better to empirically decide by experimenting with different window sizes

Exponential Smoothing

- In moving average, observations in the past used in calculating the average are weighted equally.
- Exponential Smoothing assigns *exponentially decreasing weights* as the observations get older.
- In other words, *recent observations are given relatively more weight in forecasting than older observations*.

Exponential Smoothing

- Single exponential smoothing

$$S_2 = X_1$$

$$S_t = \alpha X_{t-1} + \alpha\,(1 - \alpha) X_{t-2} + \alpha\,(1 - \alpha)^2 X_{t-3} + \ldots,$$

$$0 < \alpha \leq 1$$

$$S_t = \alpha X_{t-1} + (1 - \alpha) S_{t-1}$$

Exponential Smoothing

- Single exponential smoothing

Time	Xt	S(α=0.1)	S(α=0.5)
1	71		
2	70	71	71
3	69	70.9	70.5
4	68	70.71	69.75
5	64	70.44	68.88
6	65	69.8	66.44
7	72	69.32	65.72
8	78	69.58	68.86
9	75	70.43	73.43
10	75	70.88	74.21
11	75	71.29	74.61
12	70	71.67	74.80

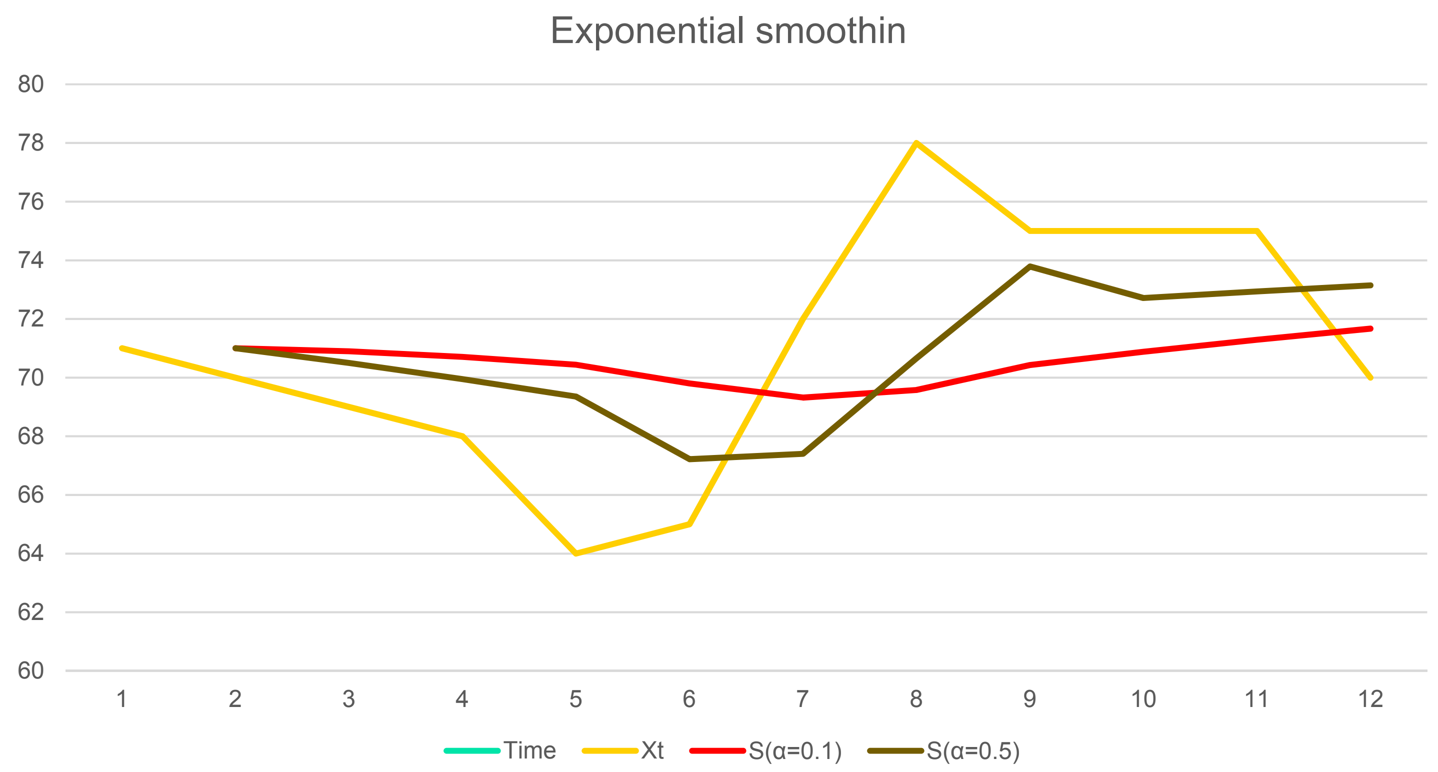
$\alpha=0.1$
MSE = 19.0

$\alpha=0.5$
MSE = 16.29

$\alpha=0.5$ is better

Exponential Smoothing

- Single exponential smoothing



Exponential Smoothing

- Single exponential smoothing does not do well when the time series data has a trend
- Double exponential smoothing is better in handling trends.
- If the data has both trend and seasonality, double exponential smoothing does not work.
- Triple exponential smoothing (Holt-Winters method) works well with time series with trend and seasonality.

Seasonality

- Detecting seasonality:
  - Time plot
  - Seasonal subseries plot
  - Multiple boxplots
  - Autocorrelation plot

Seasonality

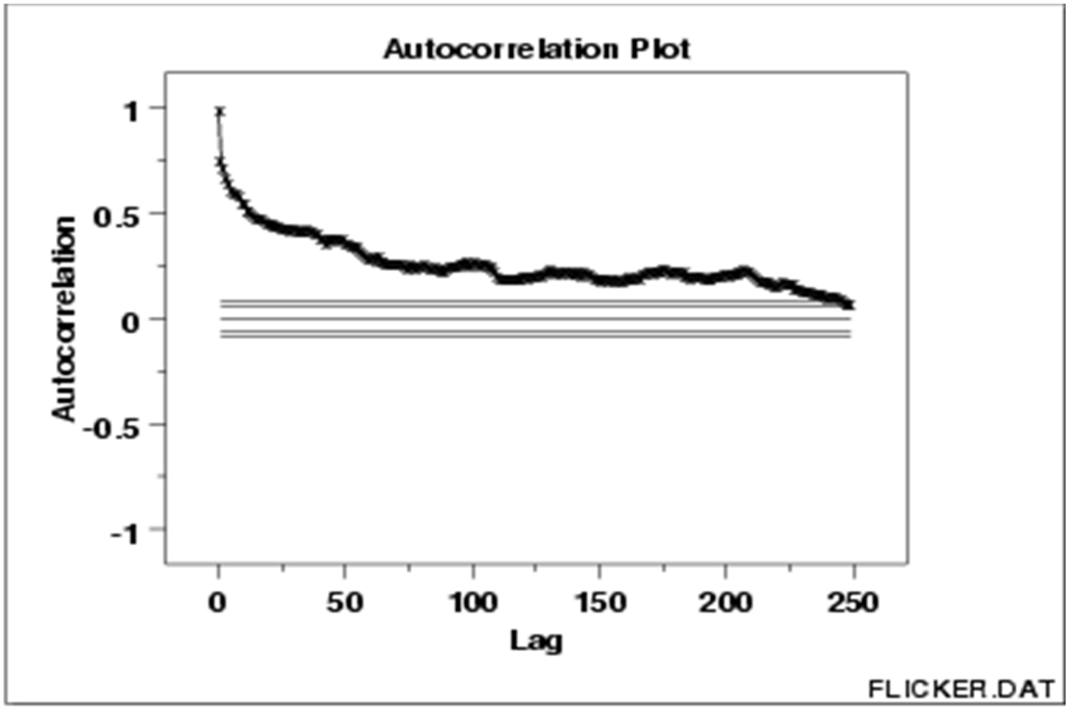
- Seasonal subseries plots are formed by:
  - Vertical axis: output variable
  - Horizontal axis: Time ordered by season. For example, with monthly data, all the January values are plotted (in chronological order), then all the February values, and so on.
- This plot is only useful if the period of the seasonality is already known. In many cases, this will in fact be known. For example, monthly data typically has a period of 12.

Seasonality

- Autocorrelation
  - Autocorrelation plots are a commonly-used tool for checking randomness in a data set.
  - This randomness is ascertained by computing autocorrelations for data values at varying time lags.
  - If random, such autocorrelations should be near zero for any and all time-lag separations.
  - If non-random, then one or more of the autocorrelations will be significantly non-zero.

Seasonality

- Autocorrelation



This sample autocorrelation plot shows that the time series is not random, but rather has a high degree of autocorrelation between adjacent and near-adjacent observations.

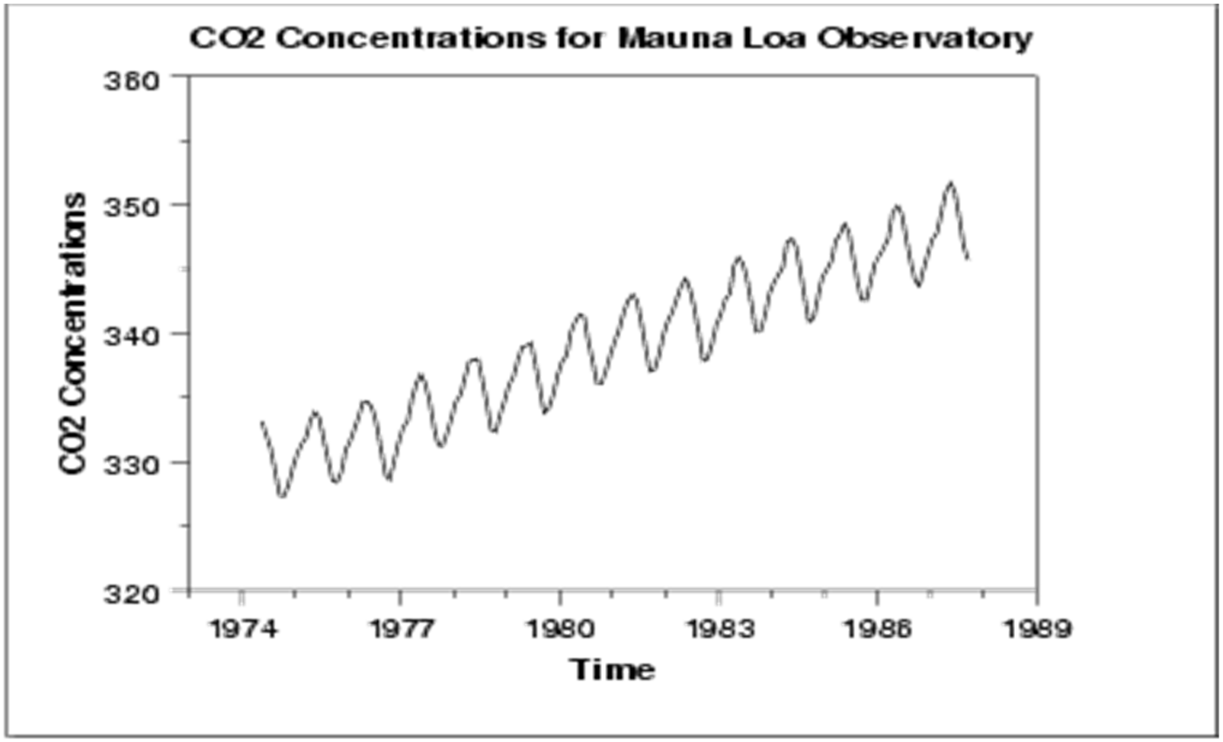
Example: CO2 Concentration

- This data set contains selected monthly mean CO_2 concentrations at the Mauna Loa Observatory from 1974 to 1987.
- Dataset (part)

CO2	Year&Month	Year	Month
333.13	1974.38	1974	5
332.09	1974.46	1974	6
331.1	1974.54	1974	7
329.14	1974.63	1974	8
327.36	1974.71	1974	9
327.29	1974.79	1974	10
328.23	1974.88	1974	11
329.55	1974.96	1974	12

Example: CO2 Concentration

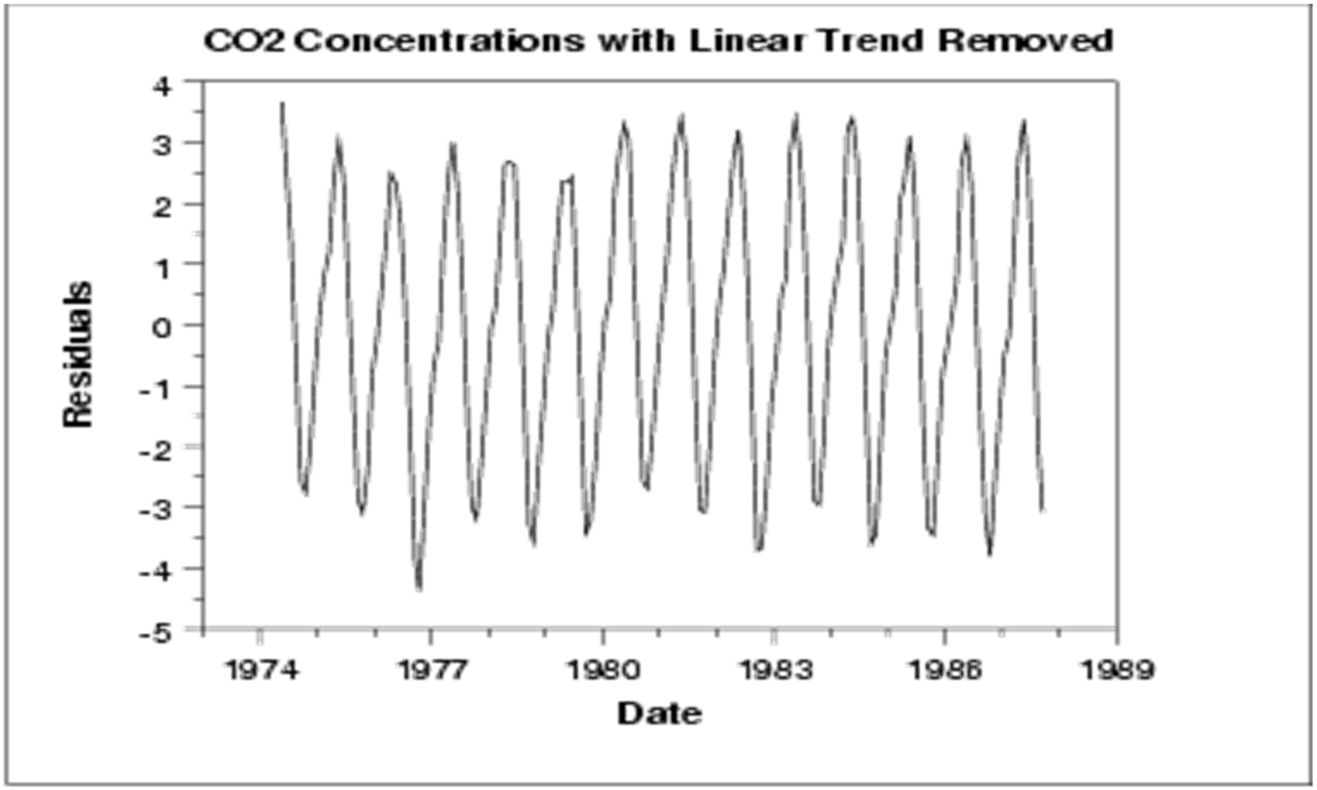
- Time plot



- Can observe a liner increasing trend.

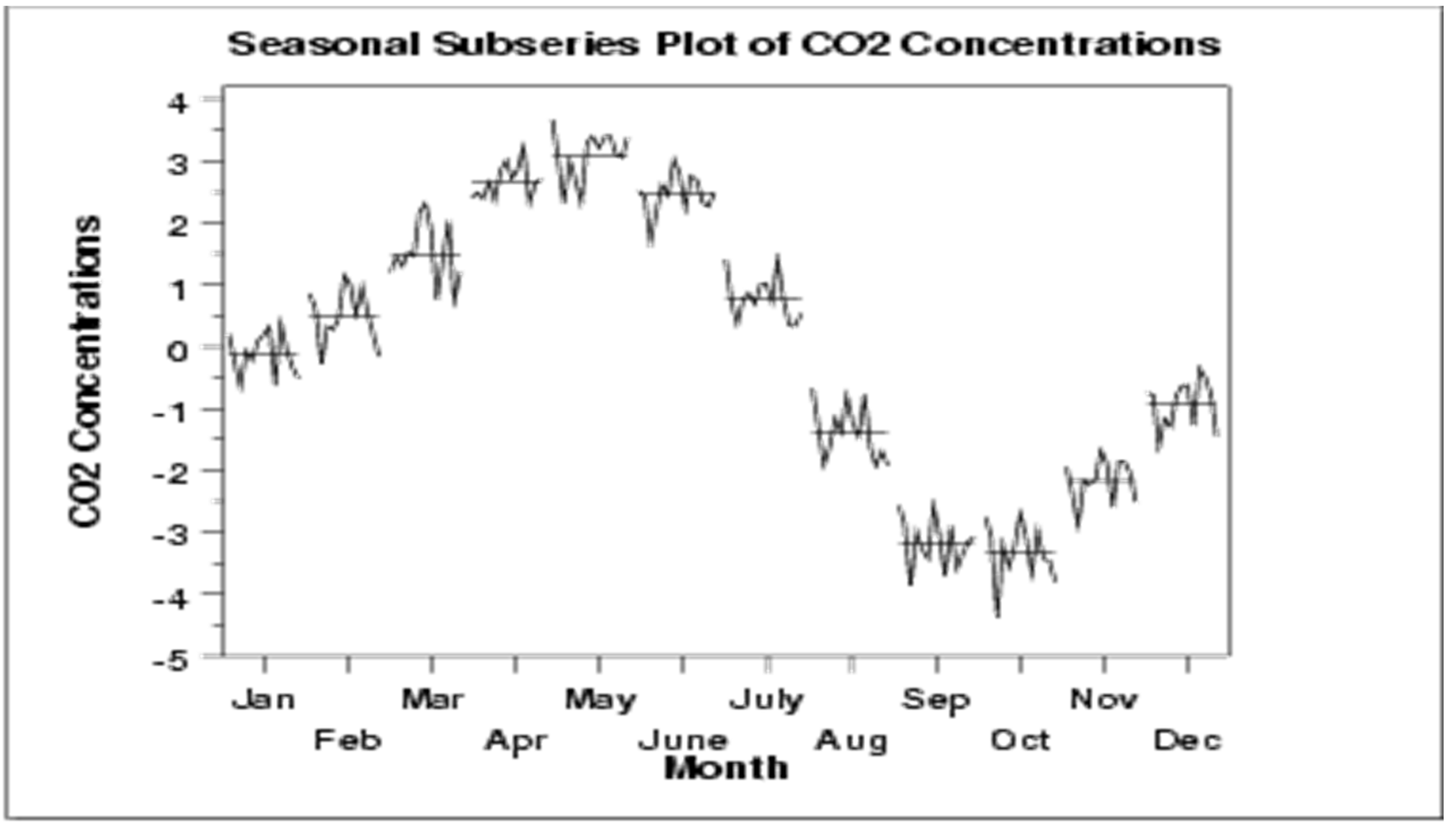
Example: CO2 Concentration

- After removing linear trend



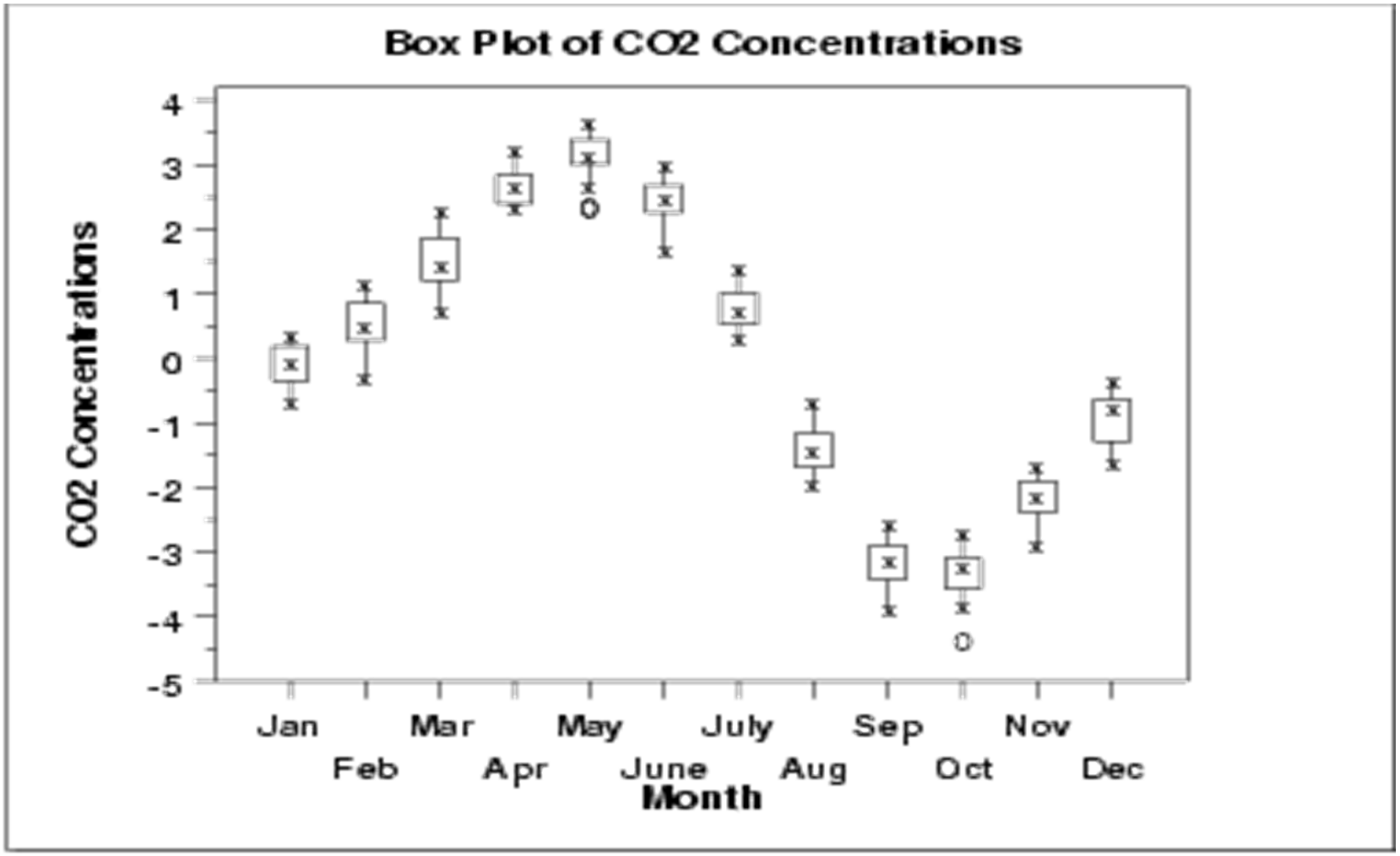
Example: CO2 Concentration

- Seasonal subseries plot



Example: CO2 Concentration

- Boxplots



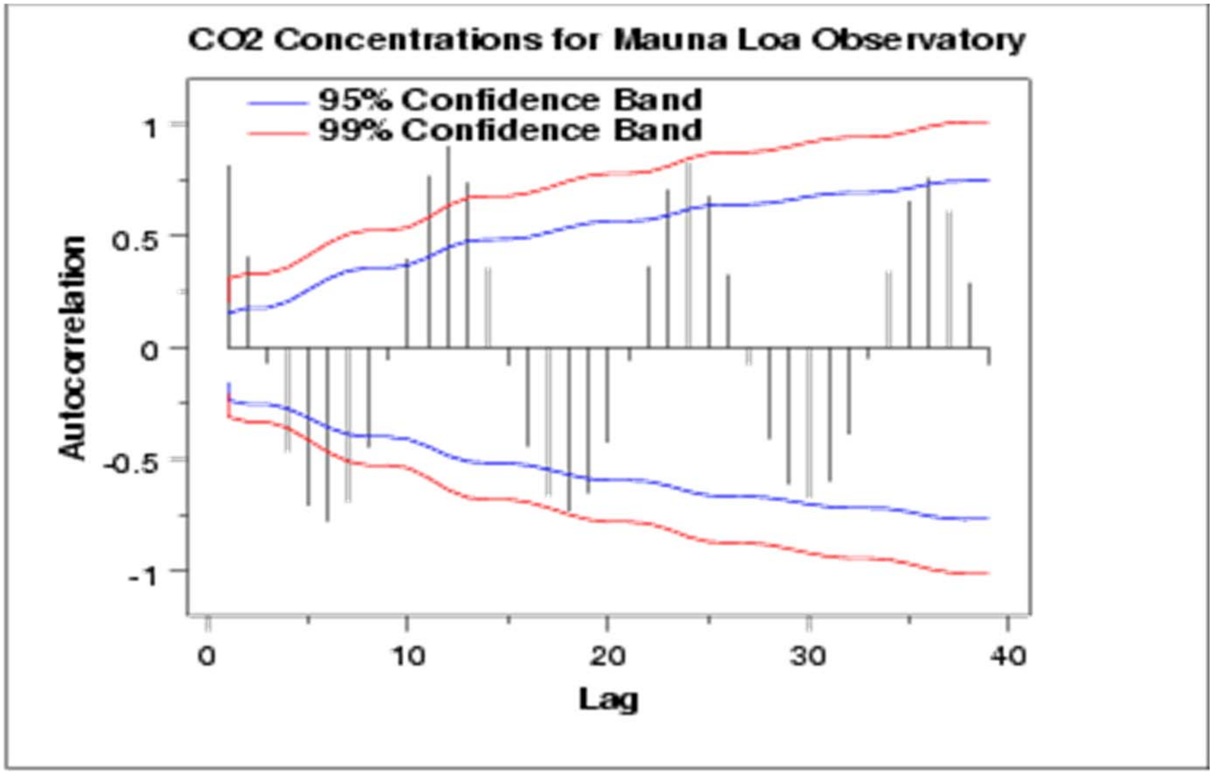
Example: CO2 Concentration

- Autocorrelation (ACF)

CO2	Lag1	Lag2	Lag3	Lag4
333.13				
332.09	333.13			
331.1	332.09	333.13		
329.14	331.1	332.09	333.13	
327.36	329.14	331.1	332.09	333.13
327.29	327.36	329.14	331.1	332.09
328.23	327.29	327.36	329.14	331.1
329.55	328.23	327.29	327.36	329.14
330.62	329.55	328.23	327.29	327.36
331.4	330.62	329.55	328.23	327.29
331.87	331.4	330.62	329.55	328.23

Example: CO2 Concentration

- Autocorrelation (ACF)



Shows strong correlation between neighbors and repeating pattern every 12 lags (12 months)

Example: Southern Oscillation

- The southern oscillation is defined as the barometric pressure difference between Tahiti and the Darwin Islands at sea level.
- The southern oscillation is a predictor of El Nino which in turn is thought to be a driver of world-wide weather.
- Specifically, repeated southern oscillation values less than -1 typically defines an El Nino.

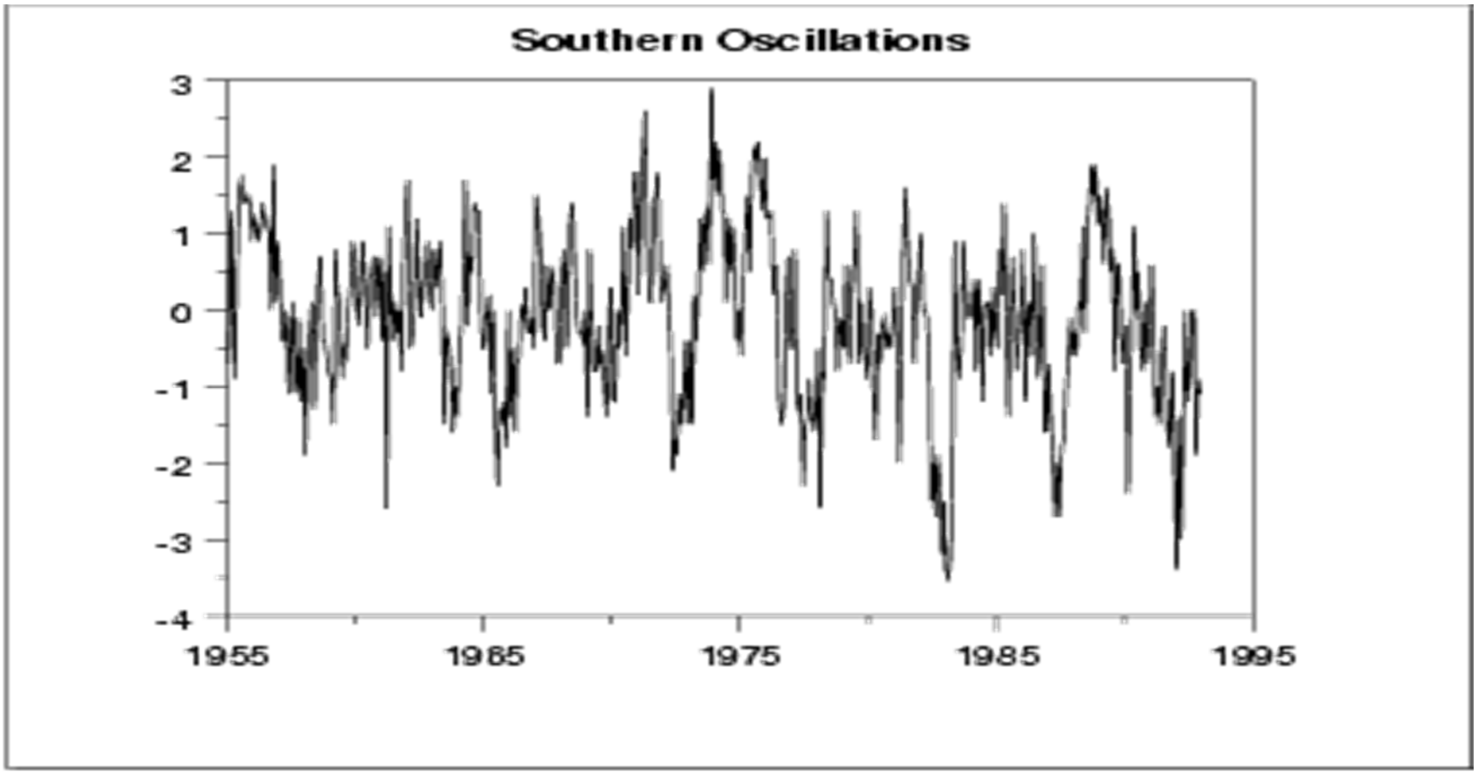
Example: Southern Oscillation

- Data (part)

Oscillation	Year_fraction	Year	Month
-0.7	1955.04	1955	1
1.3	1955.13	1955	2
0.1	1955.21	1955	3
-0.9	1955.29	1955	4
0.8	1955.38	1955	5
1.6	1955.46	1955	6
1.7	1955.54	1955	7
1.4	1955.63	1955	8
1.4	1955.71	1955	9
1.5	1955.79	1955	10
1.4	1955.88	1955	11

Example: Southern Oscillation

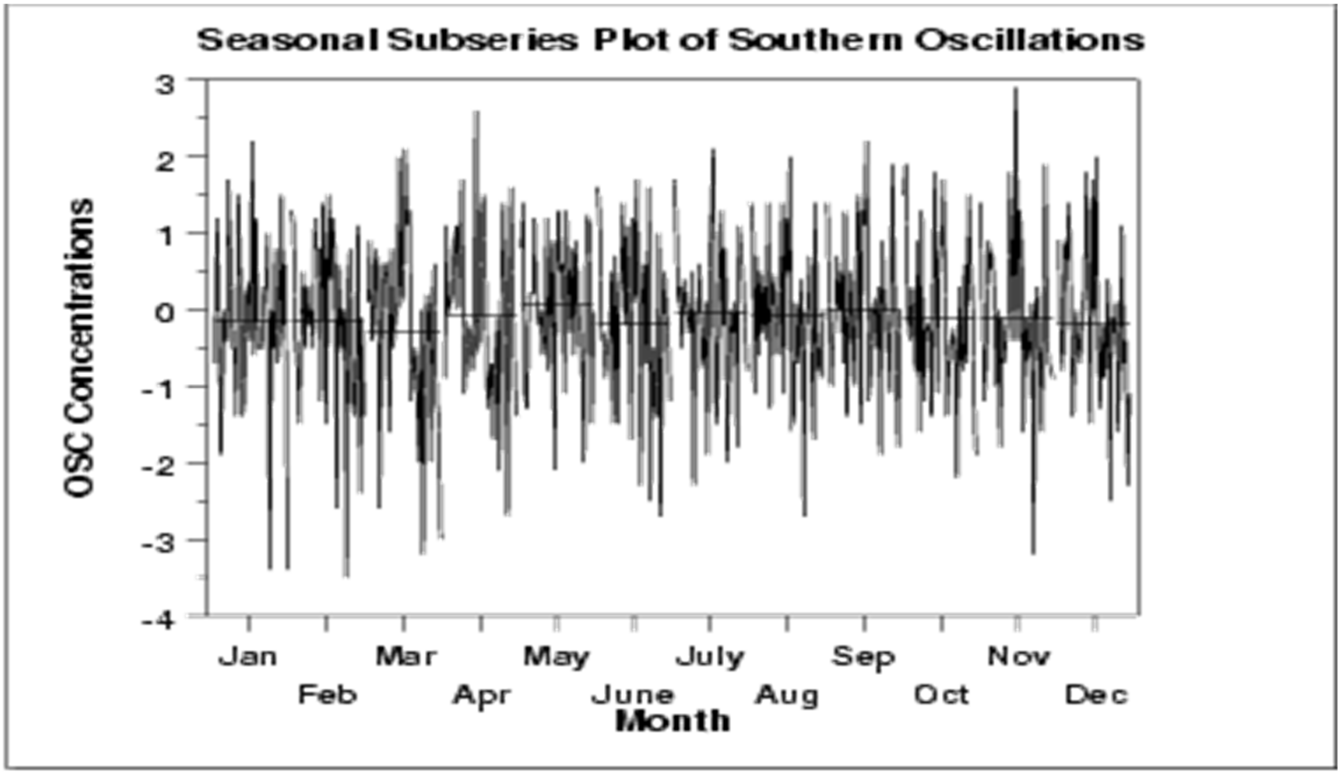
- Time plot



- No obvious periodic patterns

Example: Southern Oscillation

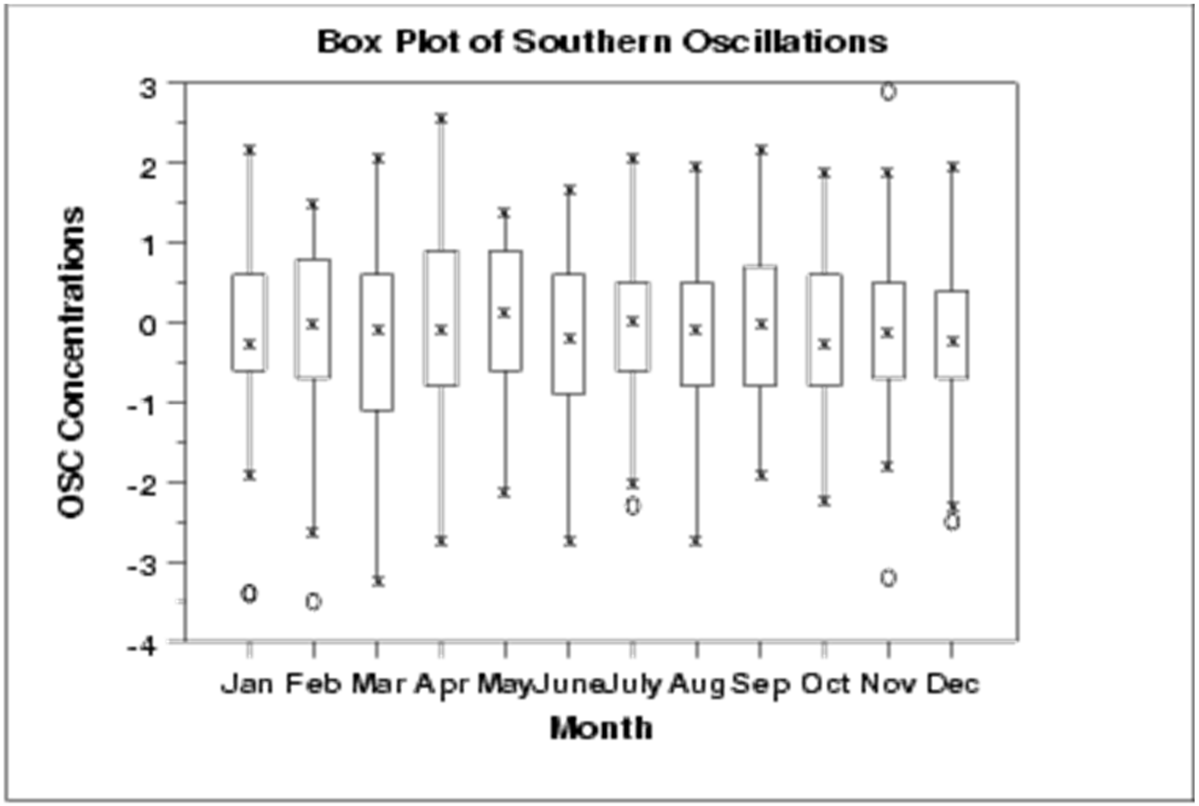
- Seasonal subseries plot



- No obvious seasonal pattern

Example: Southern Oscillation

- Boxplots



- No obvious seasonal pattern

Models

- Box-Jenkins ARMA model
  - Combines AR (autoregressive model) and MA (moving average model)
  - AR is similar to regression model, except that past predictor values are used.
  - For example, an autoregression model of order 2 (AR(2)) is:

    $Y_t = b_0 + b_1X_{t-1} + b_1X_{t-2} + e$

Models

- ARIMA (autoregressive integrated moving average) model

Performance Evaluation and Prediction

- Test of randomness
- Predictive performance