CAP 5768: Introduction to Data Science

Giri NARASIMHAN

www.cis.fiu.edu/~giri/teach/5768.html

Time Series

Time Series Analysis Applications

- Economic Forecasting
- Sales Forecasting
- Budgetary Analysis
- Stock Market Analysis
- Yield Projections
- Process and Quality Control
- Inventory Studies
- Workload Projections
- Utility Studies
- Census Analysis

https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc4.htm

Average as an Estimate

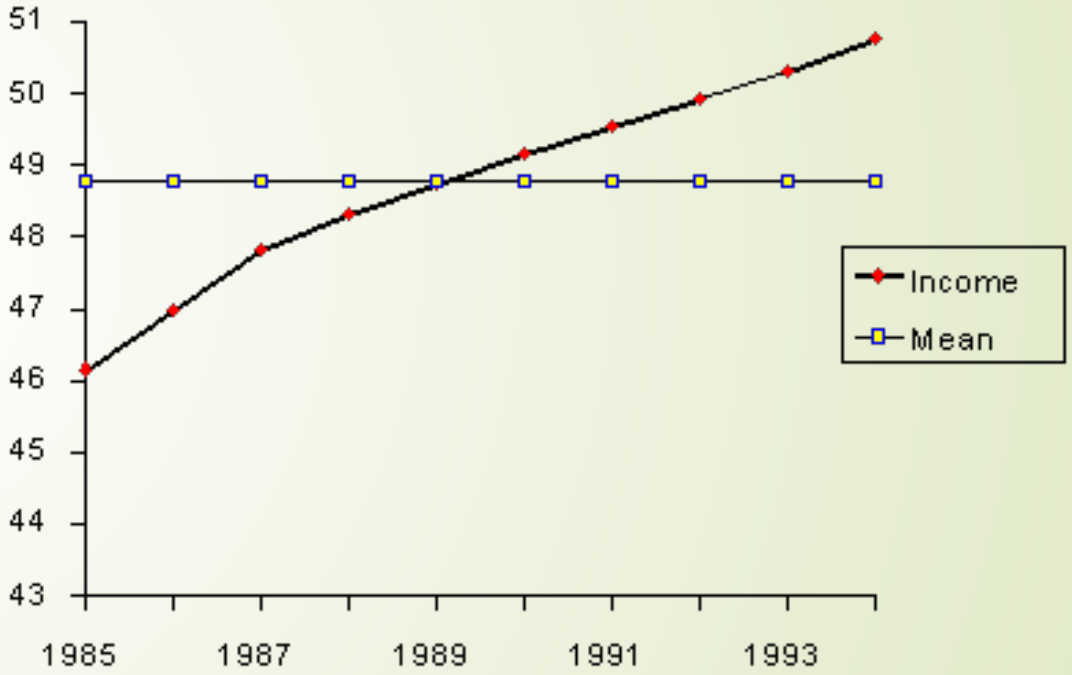
Suppl	$	Error	ES
1	9	-1	1
2	8	-2	4
3	9	-1	1
4	12	2	4
5	9	-1	1
6	12	2	4
7	11	1	1
8	7	-3	9
9	13	3	9
10	9	-1	1
11	11	1	1
12	10	0	0

Avg Est	7	9	10	12
SSE	144	48	36	84
MSE	12	4	3	7

- Is Average a good estimate?
  - Yes, it minimizes Mean Square Error (MSE)

Is Average a good predictor?

- Simple average is a bad predictor
  - Example on right: average does not show trend nor predict future



Moving Average vs Average

Suppl	$	MA	Error	ES
1	9			
2	8			
3	9	8.667	0.333	0.111
4	12	9.667	2.333	5.444
5	9	10.000	-1.000	1.000
6	12	11.000	1.000	1.000
7	11	10.667	0.333	0.111
8	7	10.000	-3.000	9.000
9	13	10.333	2.667	7.111
10	9	9.667	-0.667	0.444
11	11	11.000	0	0
12	10	10.000	0	0

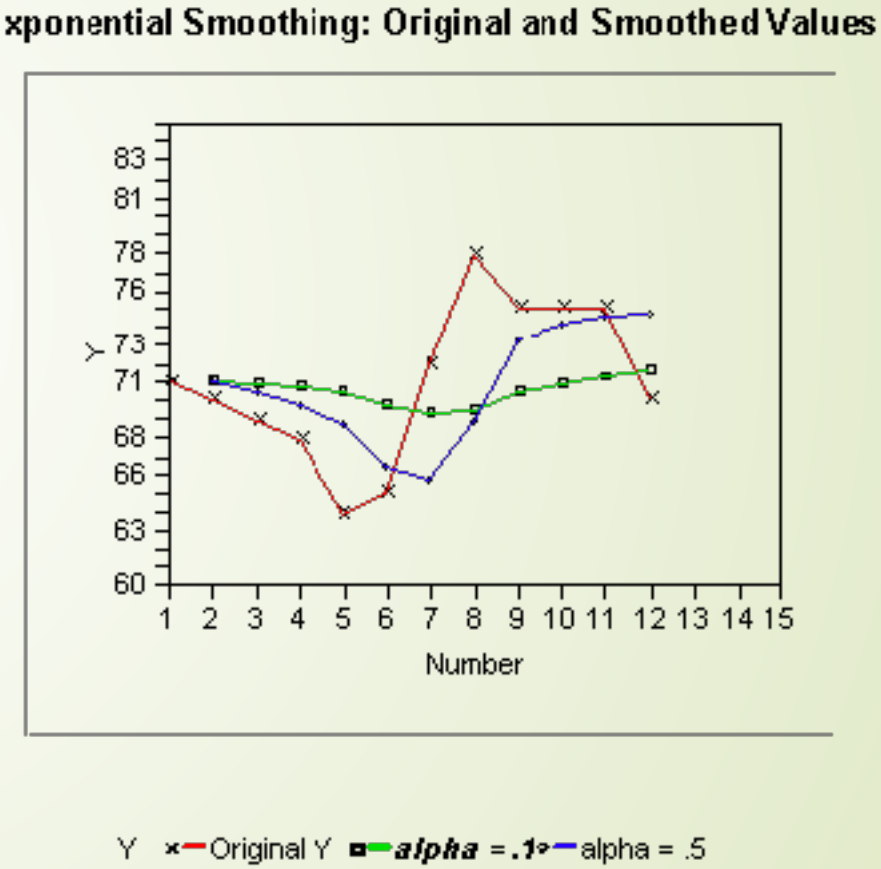
- Window size = 3
- MSE = 3.0
- Moving MSE = 2.42
- Moving Average (MA) has lower error than Average

Double Moving Averages

- Compute moving average of moving average using same window size
- Errors can be reduced further
- Do a linear regression with single and double MA to forecast

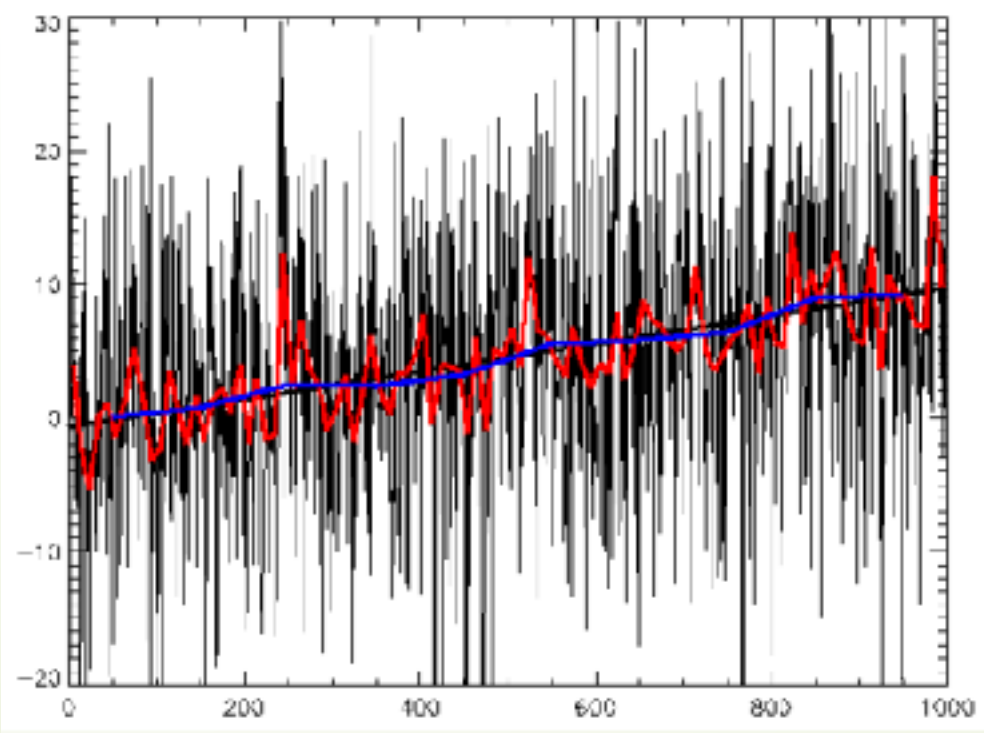
Exponential Smoothing

- **MA** gives equal weight to all items in window
- **Exp Smoothing** assigns exponentially decreasing weights for older items



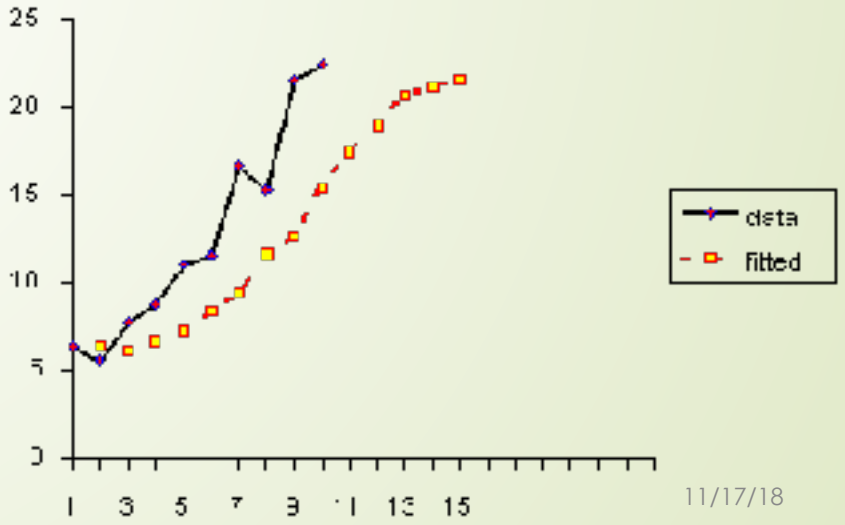
Smoothing/Averaging/Filtering

- Need to remove natural variations in data
- Shows trends unhindered by local variations



Smoothing and Forecasting

- Exponential Smoothing
  - $(y_1, y_2, \ldots, y_n)$ = sequence of observations
  - $(S_1, S_2, \ldots, S_n)$ = smoothed observations
  - $S_t = \alpha\, y_{t-1} + (1- \alpha)S_{t-1}$
- Forecasting
  - $S_{t+1} = S_t + \alpha\epsilon_t$
  - where ϵ_t is the forecast error
- More complex forecasting



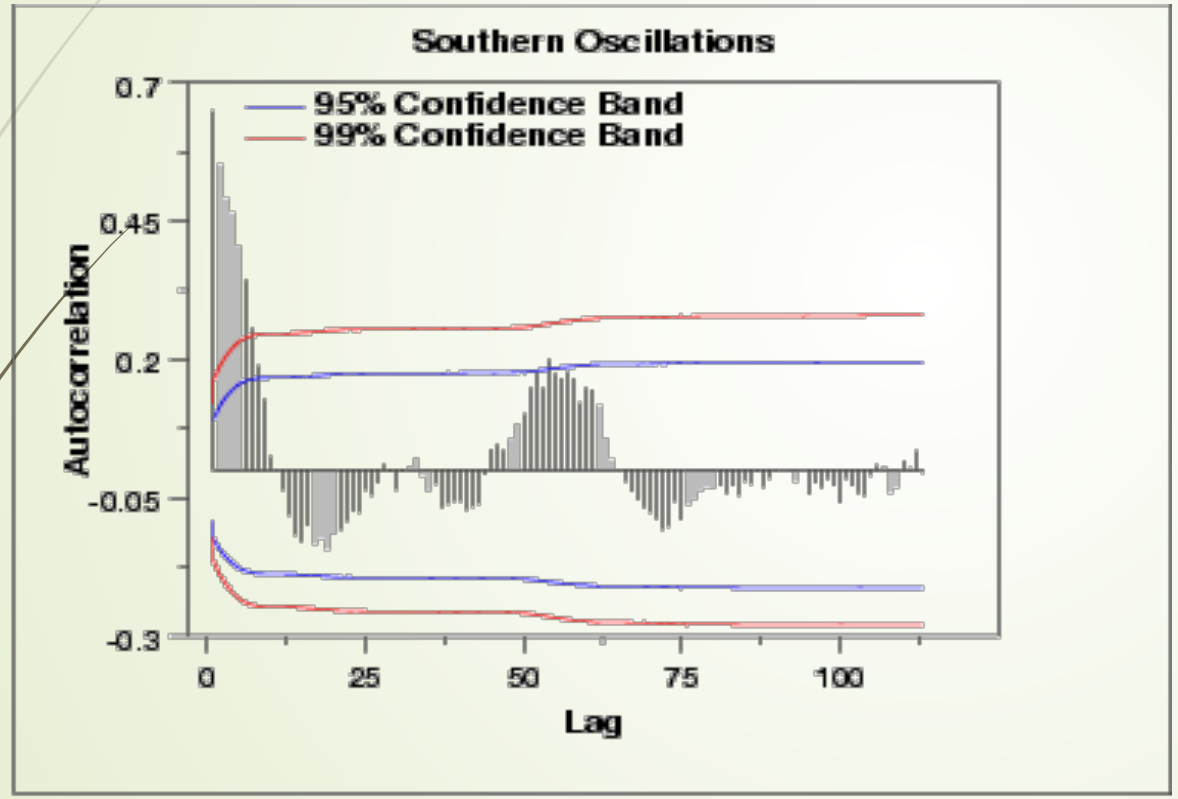
Analysis

- Check for stationarity
- Check for Trends (seasonality)
- Check for non-constant variance
  - Trends in transformed data
    - Log transform
- Check for Randomness
  - Autocorrelation plots

Autocorrelation: Box-Jenkins

- Attempt to find a regression connecting X_t with one or more prior values
  - Write down X_t as a linear combination of X_{t-1}, X_{t-2}, …, X_{t-p} with additive "white noise" and mean
- R-code available

Autocorrelation Plots



- Plot shows mixture of
  - exponentially decaying
  - damped sinusoidals
- Need autoregressive model with order > 1
- Partial autocorrelation plot needed

Spectral Plots

- Useful to analyze plots with complex cyclical structures