Time Series Data Analysis

Assoc. Prof. Nguyen Binh Minh Ph.D.

1

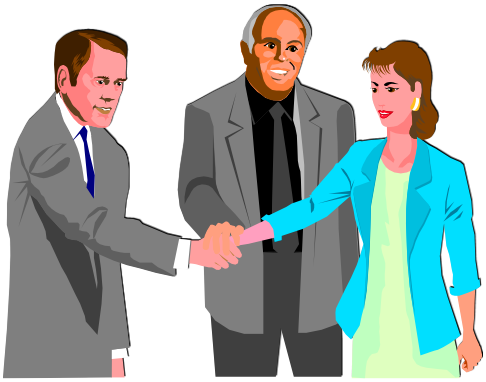
Learning Objectives

- Describe what forecasting is
- Explain time series & its components
- Smooth a data series
  - Moving average
  - Exponential smoothing
- Forecast using trend models
  Linear Regression Simple Auto-regressive

2

What Is Forecasting?

- Process of predicting a future event
- Underlying basis of all business decisions
  - Production
  - Inventory
  - Personnel
  - Facilities



3

Forecasting Approaches

Qualitative Methods

- Used when situation is vague & little data exist
  - New products
  - New technology
- Involve intuition, experience
- e.g., forecasting sales on Internet

Quantitative Methods

4

Forecasting Approaches

Qualitative Methods

- Used when situation is vague & little data exist
  - New products
  - New technology
- Involve intuition, experience
- e.g., forecasting sales on Internet

Quantitative Methods

- Used when situation is 'stable' & historical data exist
  - Existing products
  - Current technology
- Involve mathematical techniques
- e.g., forecasting sales of color televisions

5

Quantitative Forecasting

- Select several forecasting methods
- 'Forecast' the past
- Evaluate forecasts
- Select best method
- Forecast the future
- Monitor continuously forecast accuracy

6

Quantitative Forecasting Methods

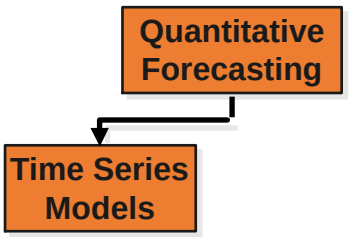
7

Quantitative Forecasting Methods

Quantitative Forecasting

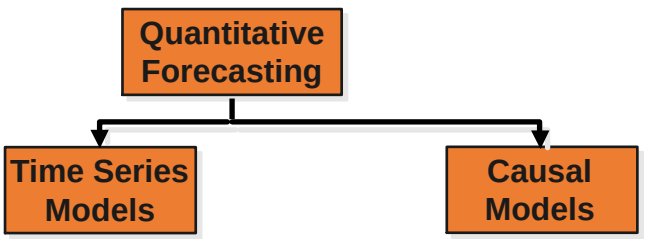
8

Quantitative Forecasting Methods



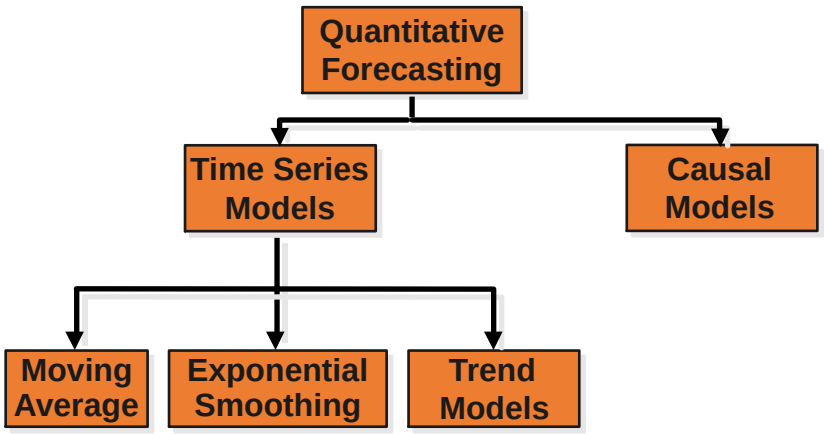
9

Quantitative Forecasting Methods



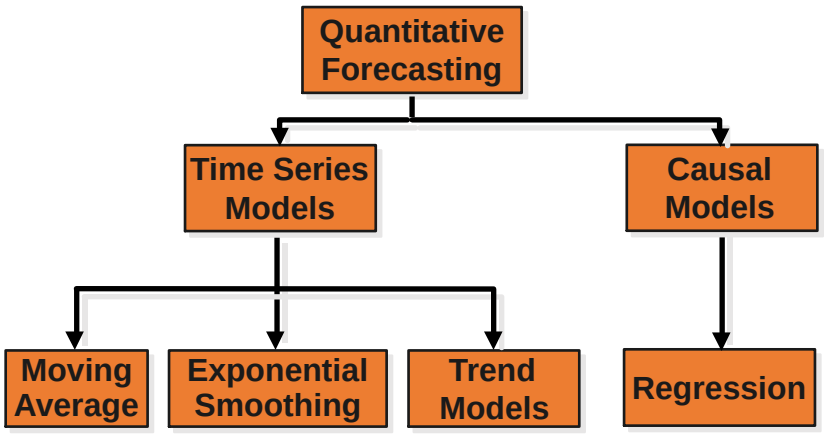
10

Quantitative Forecasting Methods



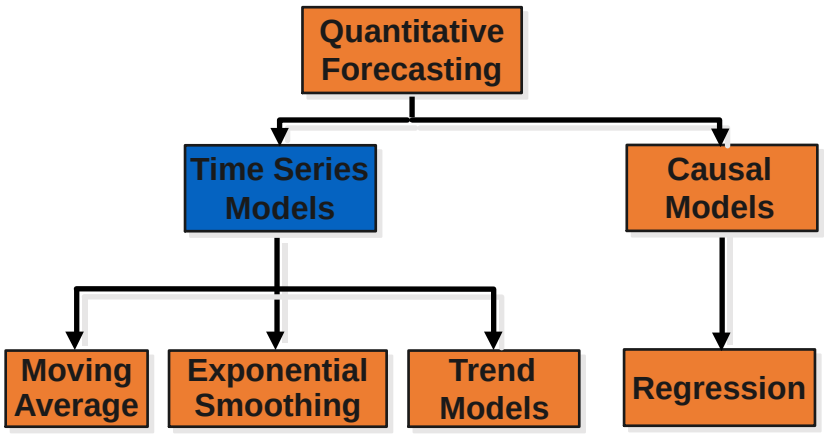
11

Quantitative Forecasting Methods



12

Quantitative Forecasting Methods



13

What is a Time Series?

- Set of evenly spaced numerical data
  - Obtained by observing response variable at regular time periods
- Forecast based only on past values
  - Assumes that factors influencing past, present, & future will continue
- Example
  - Year: 1995 1996 1997 1998 1999
  - Sales: 78.7 63.5 89.7 93.2 92.1

14

Time Series vs. Cross Sectional Data

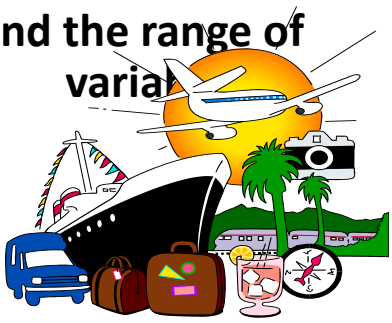
Time series data is a sequence of observations

- **collected from a process**
- **with equally spaced periods of time.**

15

Time Series vs. Cross Sectional Data

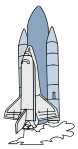
Contrary to restrictions placed on cross-sectional data, the major purpose of forecasting with time series is to extrapolate beyond the range of the explanatory varia



16

Time Series vs. Cross Sectional Data

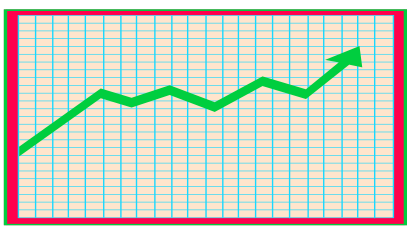
Time series is dynamic, it does change over time.



17

Time Series vs. Cross Sectional Data

When working with time series data, it is paramount that the data is plotted so the researcher can view the data.

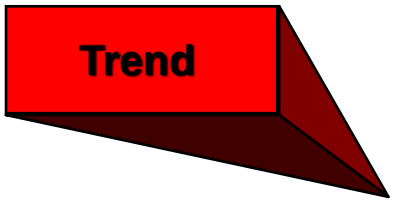


18

Time Series Components

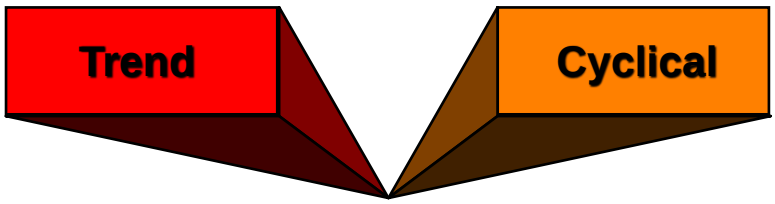
19

Time Series Components



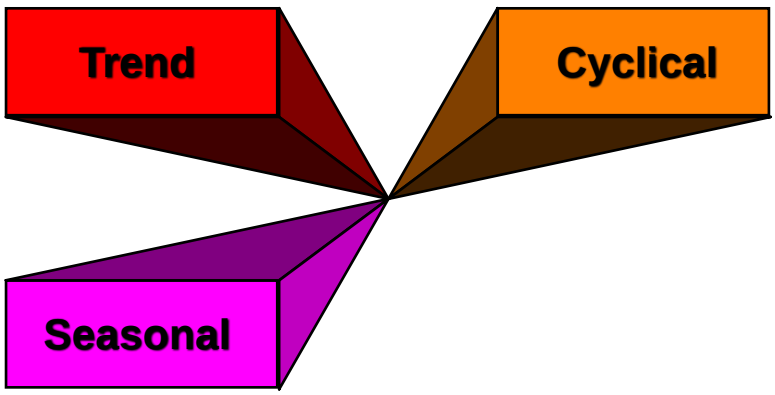
20

Time Series Components



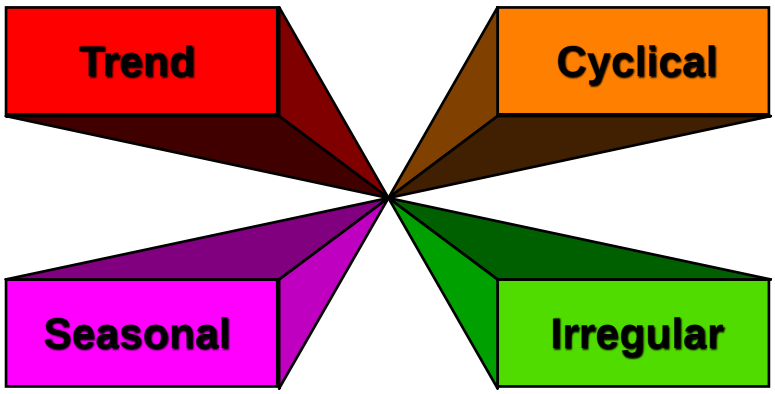
21

Time Series Components



22

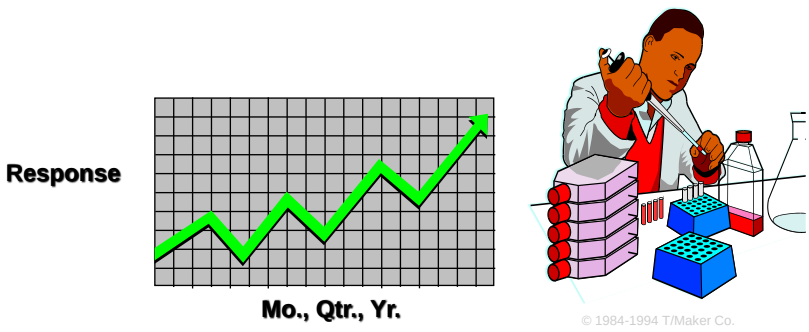
Time Series Components



23

Trend Component

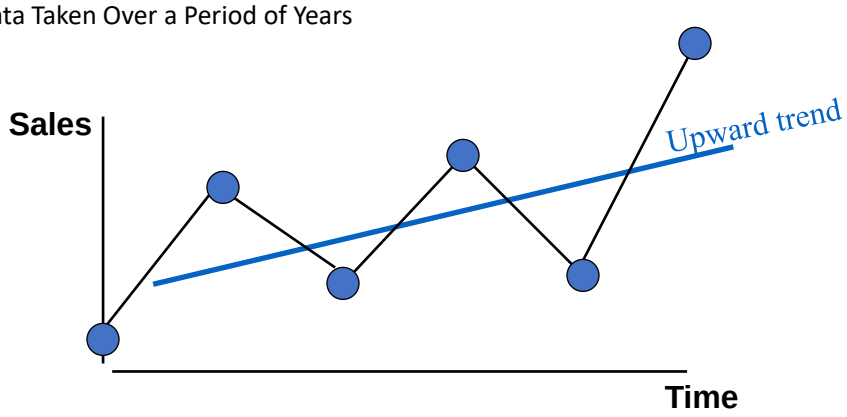
- Persistent, overall upward or downward pattern
- Due to population, technology etc.
- Several years duration



24

Trend Component

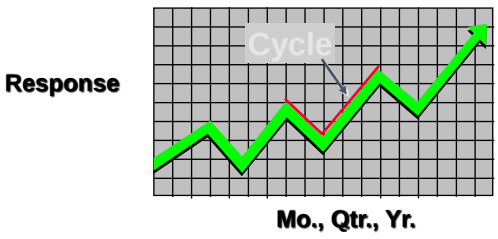
- Overall Upward or Downward Movement
- Data Taken Over a Period of Years



25

Cyclical Component

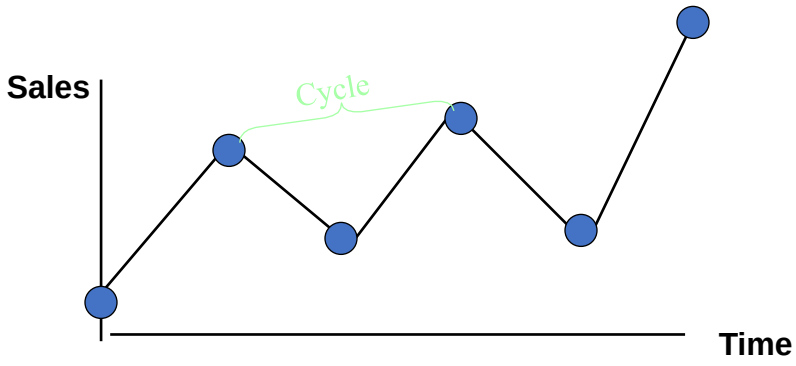
- Repeating up & down movements
- Due to interactions of factors influencing economy
- Usually 2-10 years duration



26

Cyclical Component

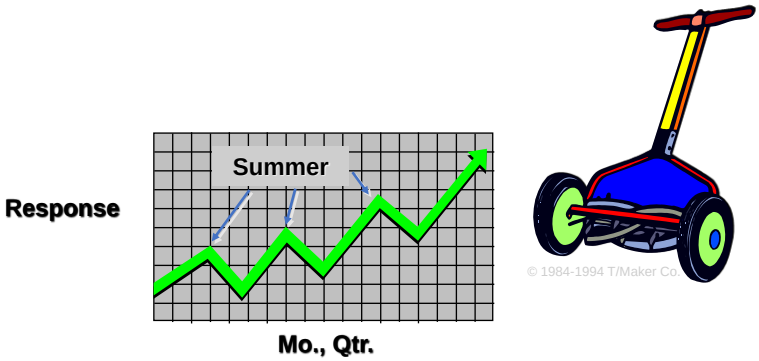
- Upward or Downward Swings
- May Vary in Length
- Usually Lasts 2 - 10 Years



27

Seasonal Component

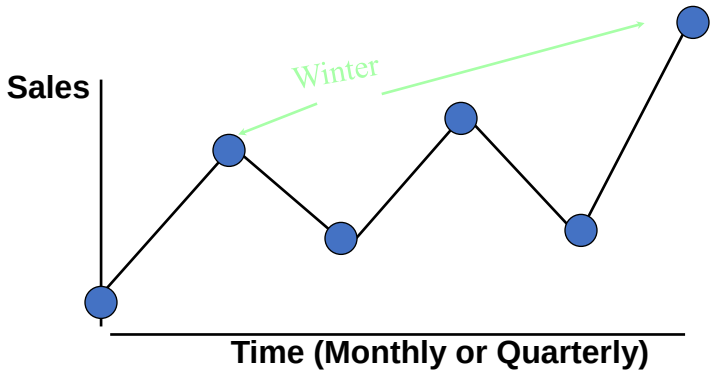
- Regular pattern of up & down fluctuations
- Due to weather, customs etc.
- Occurs within one year



28

Seasonal Component

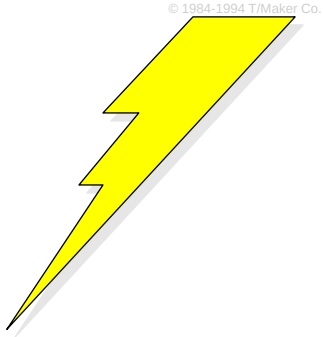
- Upward or Downward Swings
- Regular Patterns
- Observed Within One Year



29

Irregular Component

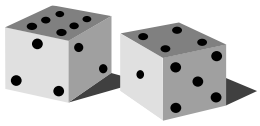
- Erratic, unsystematic, 'residual' fluctuations
- Due to random variation or unforeseen events
  - Union strike
  - War
- Short duration & nonrepeating



30

Random or Irregular Component

- Erratic, Nonsystematic, Random, 'Residual' Fluctuations
- Due to Random Variations of
  - Nature
  - Accidents
- Short Duration and Non-repeating



31

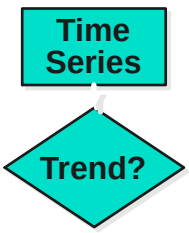
Time Series Forecasting

32

Time Series Forecasting

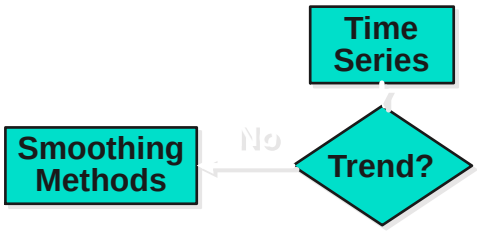
33

Time Series Forecasting



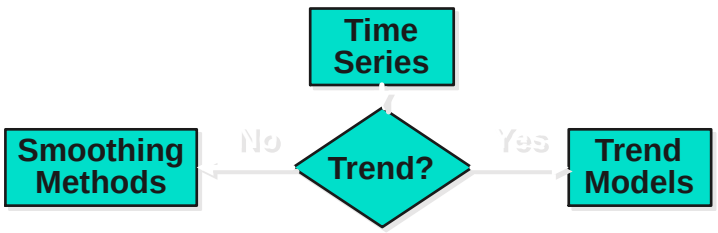
34

Time Series Forecasting



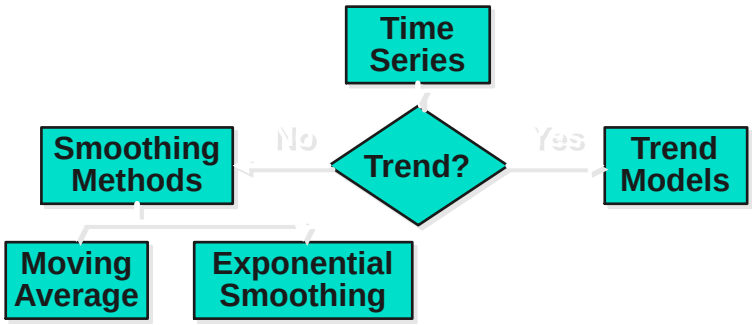
35

Time Series Forecasting



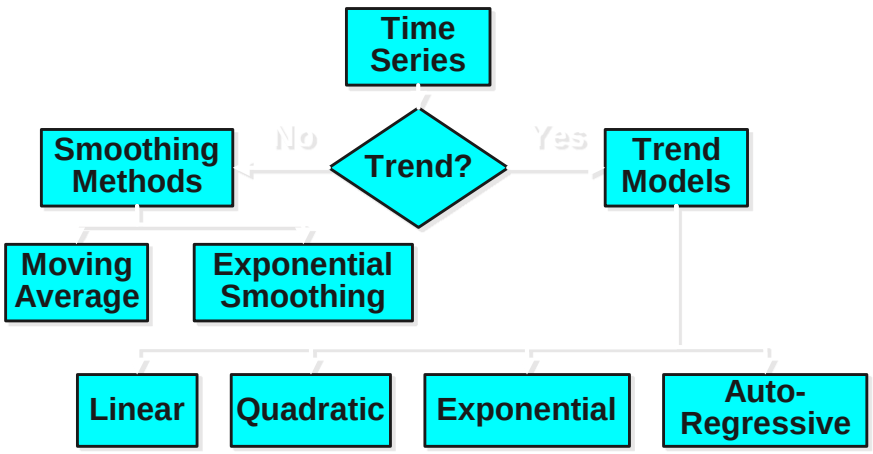
36

Time Series Forecasting



37

Time Series Forecasting

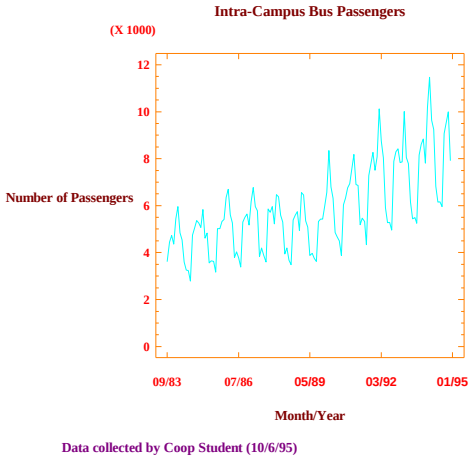


38

Time Series Analysis

39

Plotting Time Series Data

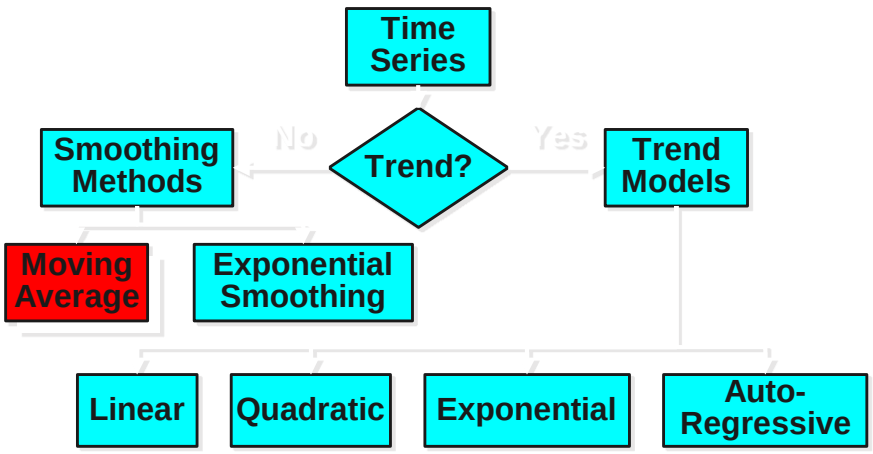


40

Moving Average Method

41

Time Series Forecasting



42

Moving Average Method

- Series of arithmetic means
- Used only for smoothing
  - Provides overall impression of data over time

43

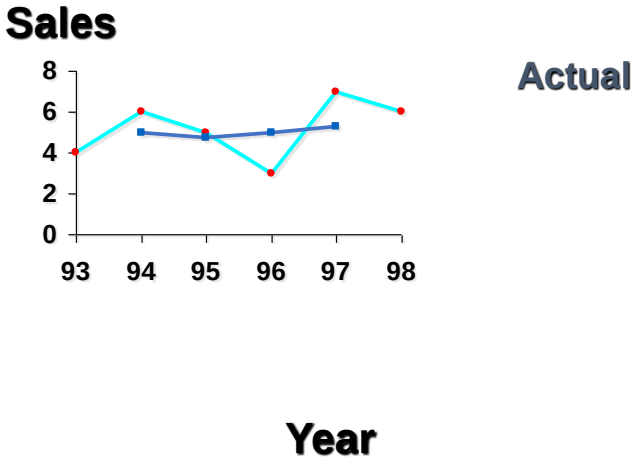
Moving Average Method

- Series of arithmetic means
- Used only for smoothing
  - Provides overall impression of data over time

Used for elementary forecasting

44

Moving Average Graph



45

Moving Average

[An Example]

You work for Firestone Tire. You want to smooth random fluctuations using a 3-period moving average.

1995	20,000
1996	24,000
1997	22,000
1998	26,000
1999	25,000

46

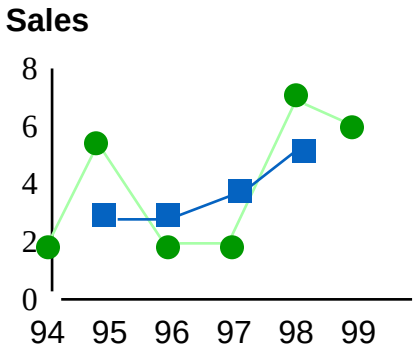
Moving Average
[Solution]

Year	Sales	MA(3) in 1,000
1995	20,000	NA
1996	24,000	(20+24+22)/3 = 22
1997	22,000	(24+22+26)/3 = 24
1998	26,000	(22+26+25)/3 = 24
1999	25,000	NA

47

Moving Average

Year	Response ●	Moving Ave ■
1994	2	NA
1995	5	3
1996	2	3
1997	2	3.67
1998	7	5
1999	6	NA

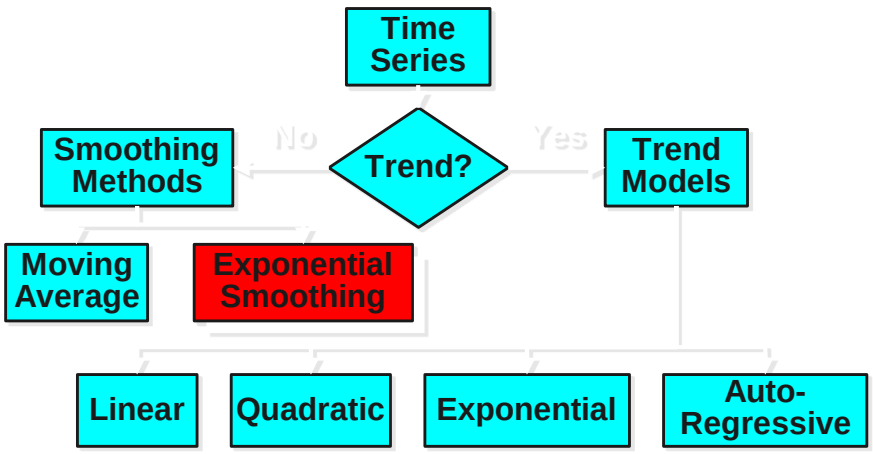


48

Exponential Smoothing Method

49

Time Series Forecasting



50

Exponential Smoothing Method

- Form of weighted moving average
  - Weights decline exponentially
  - Most recent data weighted most
- Requires smoothing constant (*W*)
  - Ranges from 0 to 1
  - Subjectively chosen
- Involves little record keeping of past data

51

Exponential Smoothing
[An Example]

You're organizing a Kwanza meeting. You want to forecast attendance for 1998 using exponential smoothing (α = .20). Past attendance (00) is:

1995	4
1996	6
1997	5
1998	3
1999	7

52

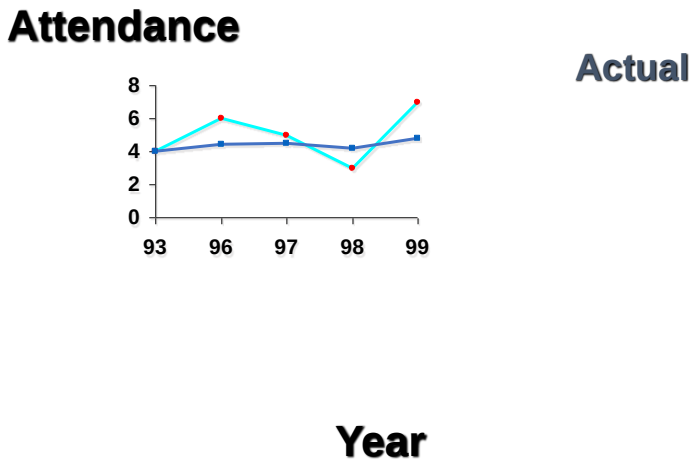
Exponential Smoothing

$$E_i = W \cdot Y_i + (1 - W) \cdot E_{i-1}$$

Time	Y_i	Smoothed Value, E_i (W = .2)	Forecast $\hat{Y}_{i+1}$
1995	4	4.0	NA
1996	6	(.2)(6) + (1-.2)(4.0) = 4.4	4.0
1997	5	(.2)(5) + (1-.2)(4.4) = 4.5	4.4
1998	3	(.2)(3) + (1-.2)(4.5) = 4.2	4.5
1999	7	(.2)(7) + (1-.2)(4.2) = 4.8	4.2
2000	NA	NA	4.8

53

Exponential Smoothing [Graph]



54

Forecast Effect of Smoothing Coefficient (*W*)

$$\hat{Y}_{i+1} = W \cdot Y_i + W \cdot (1-W) \cdot Y_{i-1} + W \cdot (1-W)^2 \cdot Y_{i-2} + \ldots$$

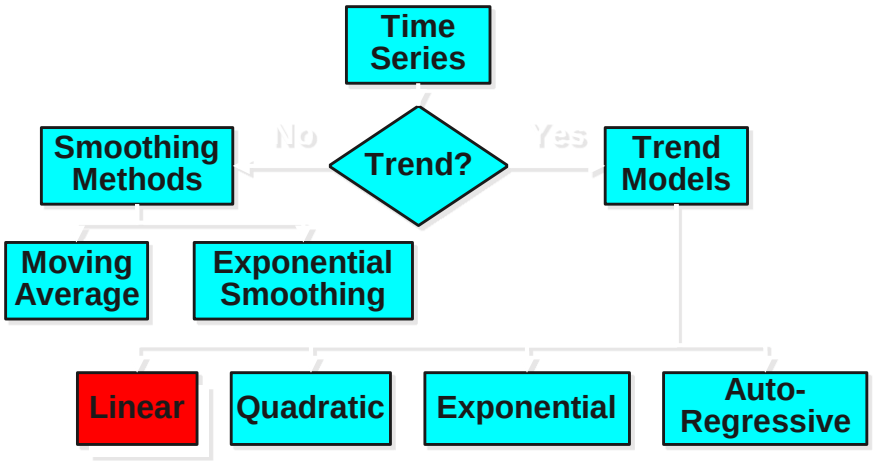
W is...	Weight: Prior Period	Weight: 2 Periods Ago	Weight: 3 Periods Ago
	W	*W*(1-*W*)	*W*(1-*W*)^2
0.10	10%	9%	8.1%
0.90	90%	9%	0.9%

55

Linear Time-Series Forecasting Model

56

Time Series Forecasting



57

Linear Time-Series Forecasting Model

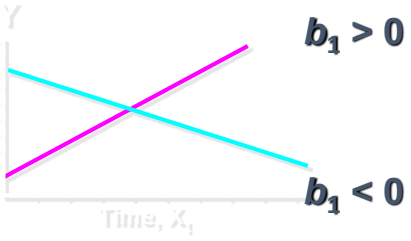
- Used for forecasting trend
- Relationship between response variable *Y* & time *X* is a linear function
- Coded *X* values used often

• Year *X*:	1995	1996	1997	1998	1999
• Coded year:	0	1	2	3	4
• Sales *Y*:	78.7	63.5	89.7	93.2	92.1

58

Linear Time-Series Model

$$\hat{Y}_i = b_0 + b_1 X_{1i}$$



59

Linear Time-Series Model [An Example]

You're a marketing analyst for Hasbro Toys. Using coded years, you find $Y_i = .6 + .7X_i$.^

1995	1
1996	1
1997	2
1998	2
1999	4

Forecast 2000 sales.

60

Linear Time-Series [Example]

Year	Coded Year	Sales (Units)
1995	0	1
1996	1	1
1997	2	2
1998	3	2
1999	4	4
2000	5	?

2000 forecast sales: $Y_i = .6 + .7 \cdot (5) = 4.1$

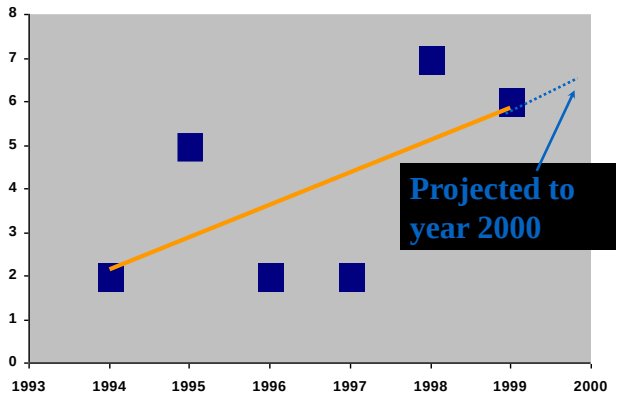
The equation would be different if 'Year' used.

61

The Linear Trend Model

Year	Coded	Sales
94	0	2
95	1	5
96	2	2
97	3	2
98	4	7
99	5	6

$$\hat{Y}_i = b_0 + b_1 X_i = 2.143 + .743 X_i$$

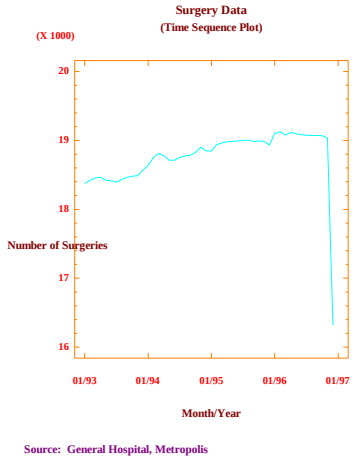


Excel Output

	Coefficients
Intercept	2.14285714
X Variable	0.74285714

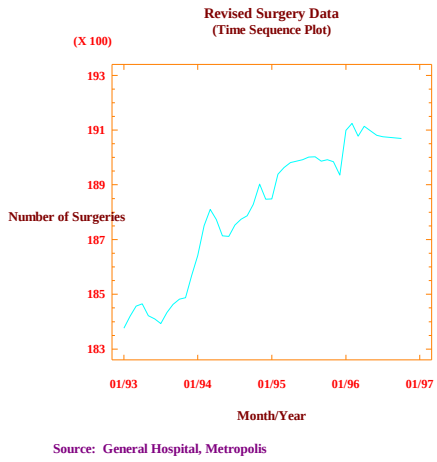
62

Time Series Plot



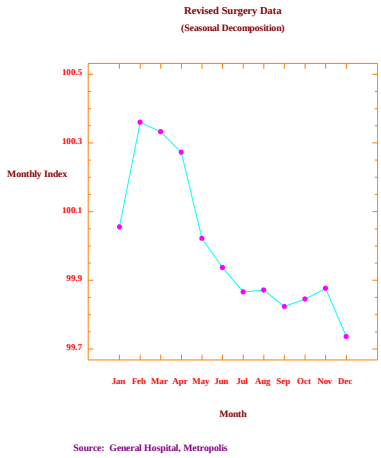
63

Time Series Plot [Revised]



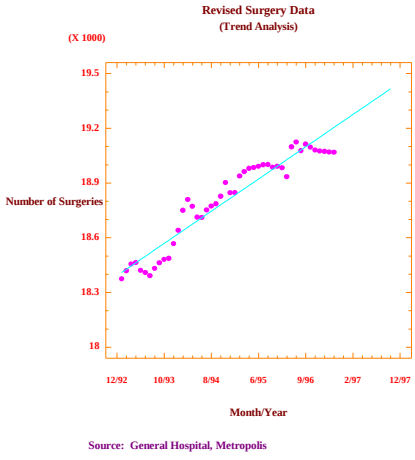
64

Seasonality Plot



65

Trend Analysis

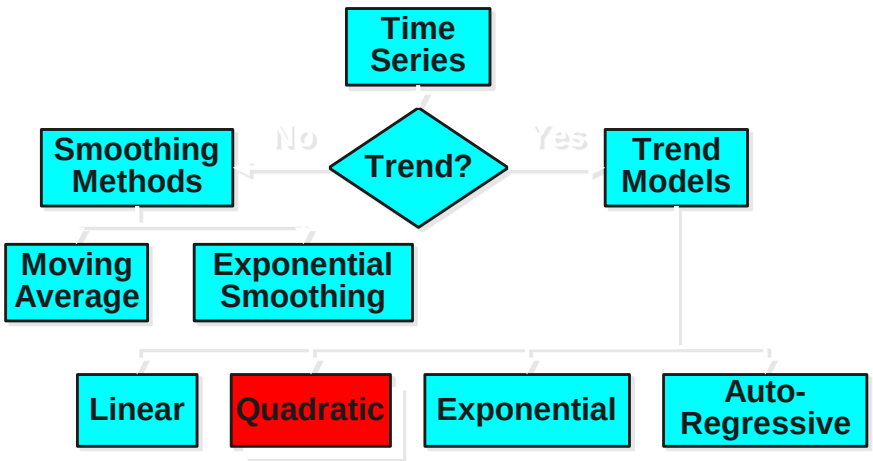


66

Quadratic Time-Series Forecasting Model

67

Time Series Forecasting



68

Quadratic Time-Series Forecasting Model

- Used for forecasting trend
- Relationship between response variable *Y* & time *X* is a quadratic function
- Coded years used

69

Quadratic Time-Series Forecasting Model

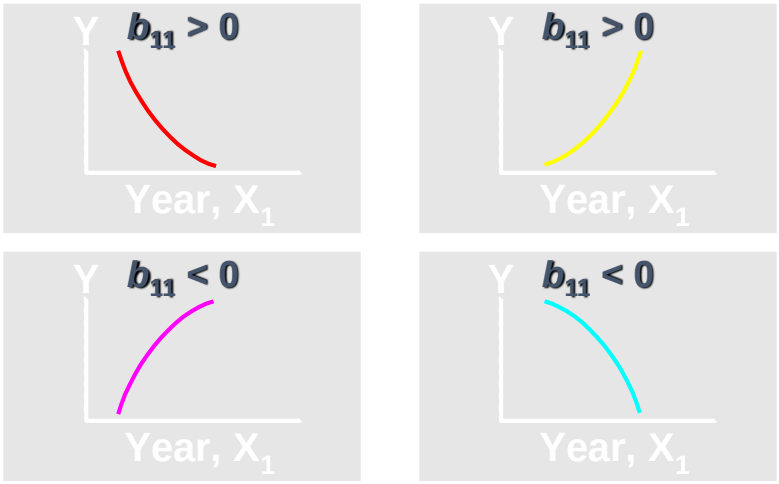
- Used for forecasting trend
- Relationship between response variable *Y* & time *X* is a quadratic function
- Coded years used
- Quadratic model

$$\hat{Y}_i = b_0 + b_1 X_{1i} + b_{11} X_{1i}^2$$

70

Quadratic Time-Series Model

Relationships



71

Quadratic Trend Model

Year	Coded	Sales
94	0	2
95	1	5
96	2	2
97	3	2
98	4	7
99	5	6

$$\hat{Y}_i = b_0 + b_1 X_i + b_2 X_i^2$$

	Coefficients
Intercept	2.85714286
X Variable 1	-0.3285714
X Variable 2	0.21428571

Excel Output

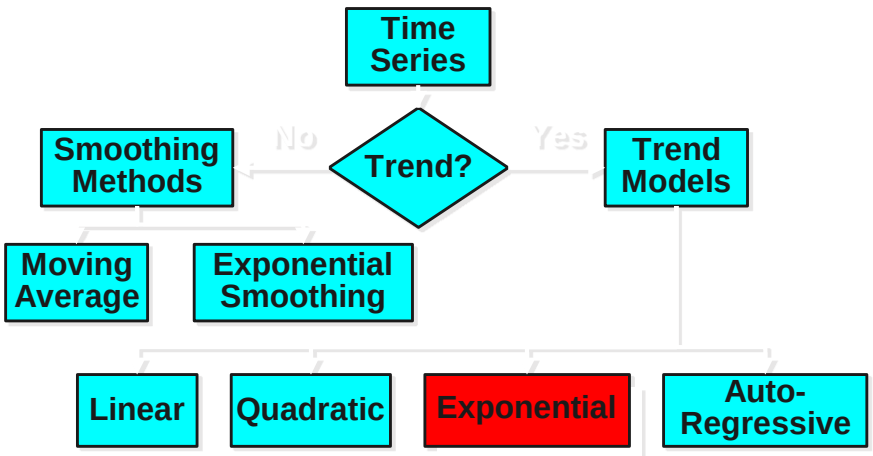
$$\hat{Y}_i = 2.857 - 0.33 X_i + .214 X_i^2$$

72

Exponential Time-Series Model

73

Time Series Forecasting



74

Exponential Time-Series Forecasting Model

- Used for forecasting trend
- Relationship is an exponential function
- Series increases (decreases) at increasing (decreasing) rate

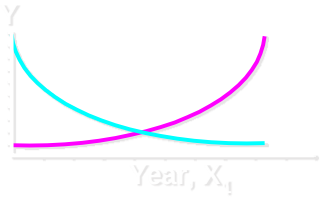
75

Exponential Time-Series Forecasting Model

- Used for forecasting trend
- Relationship is an exponential function
- Series increases (decreases) at increasing (decreasing) rate

76

Exponential Time-Series Model Relationships

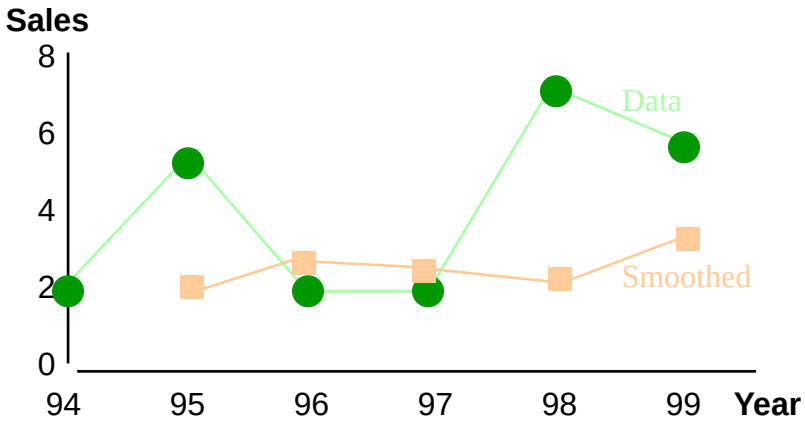


$b_1 > 1$

$0 < b_1 < 1$

77

Exponential Weight [Example Graph]



78

Exponential Trend Model

$$\hat{Y}_i = b_0 b_1^{X_i} \quad \text{or} \quad \log \hat{Y}_i = \log b_0 + X_1 \log b_1$$

Year	Coded	Sales
94	0	2
95	1	5
96	2	2
97	3	2
98	4	7
99	5	6

	Coefficients
Intercept	0.33583795
X Variable	0.08068544

Excel Output of Values in logs

antilog(.33583795) =	2.17
antilog(.08068544) =	1.2

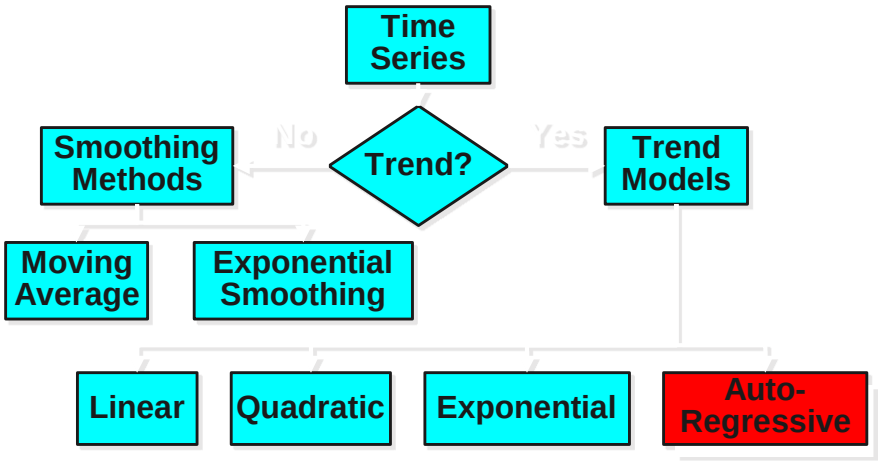
$$\hat{Y}_i = (2.17)(1.2)^{X_i}$$

79

Autoregressive Modeling

80

Time Series Forecasting



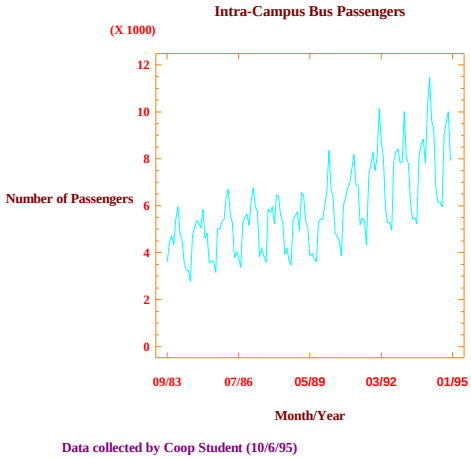
81

Autoregressive Modeling

- Used for forecasting trend
- Like regression model
  - Independent variables are lagged response variables Y_{i-1}, Y_{i-2}, Y_{i-3} etc.
- Assumes data are correlated with past data values
  - 1st Order: Correlated with prior period
- Estimate with ordinary least squares

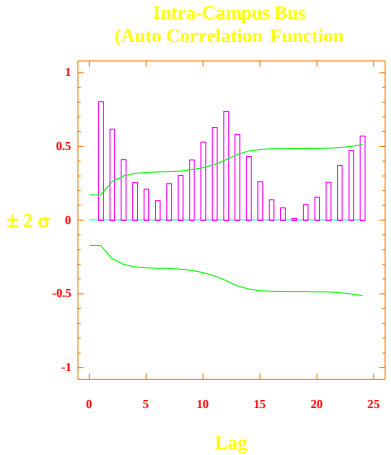
82

Time Series Data Plot



83

Auto-correlation Plot

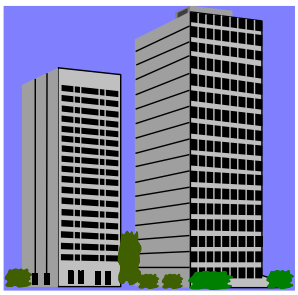


84

Autoregressive Model [An Example]

The Office Concept Corp. has acquired a number of office units (in thousands of square feet) over the last 8 years. Develop the 2nd order Autoregressive models.

Year	Units
92	4
93	3
94	2
95	3
96	2
97	2
98	4
99	6



Autoregressive Model [Example Solution]

- Develop the 2nd order table
- Use Excel to run a regression model

Excel Output

	Coefficients
Intercept	3.5
X Variable 1	0.8125
X Variable 2	-0.9375

Year	Y_i	Y_{i-1}	Y_{i-2}
92	4	---	---
93	3	4	---
94	2	3	4
95	3	2	3
96	2	3	2
97	2	2	3
98	4	2	2
99	6	4	2

$$Y_i = 3.5 + .8125Y_{i-1} - .9375Y_{i-2}$$

Evaluating Forecasts

87

Quantitative Forecasting Steps

- Select several forecasting methods
- 'Forecast' the past
- **Evaluate forecasts**
- Select best method
- Forecast the future
- Monitor continuously forecast accuracy

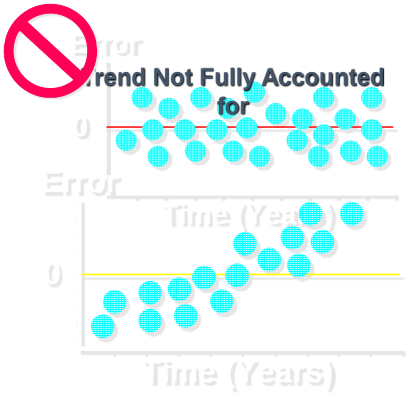
88

Forecasting Guidelines

- No pattern or direction in forecast error
  - e_i = (Actual Y_i - Forecast Y_i)
  - Seen in plots of errors over time
- Smallest forecast error
  - Measured by mean absolute deviation
- Simplest model
  - Called principle of parsimony

89

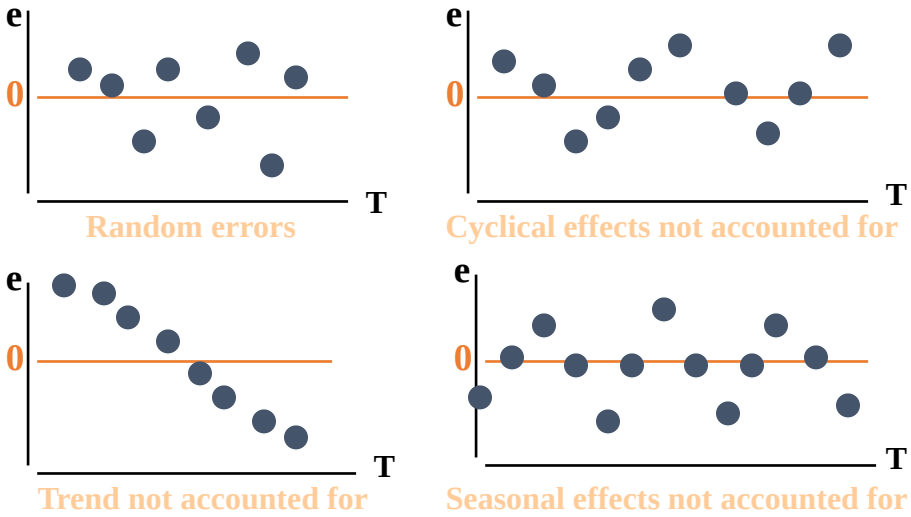
Pattern of Forecast Error



Desired Pattern

90

Residual Analysis



91

Principal of Parsimony

- Suppose two or more models provide good fit for data
- Select the Simplest Model
  - Simplest model types:
    - least-squares linear
    - least-square quadratic
    - 1st order autoregressive
  - More complex types:
    - 2nd and 3rd order autoregressive
    - least-squares exponential

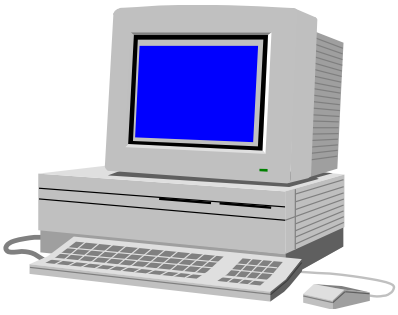
92

Summary

- Described what forecasting is
- Explained time series & its components
- Smoothed a data series
  - Moving average
  - Exponential smoothing
- Forecasted using trend models

93

You and StatGraphics



- **Specification**

  [Know assumptions underlying various models.]
- **Estimation**

  [Know mechanics of StatGraphics Plus Win].
- **Diagnostic checking**

94

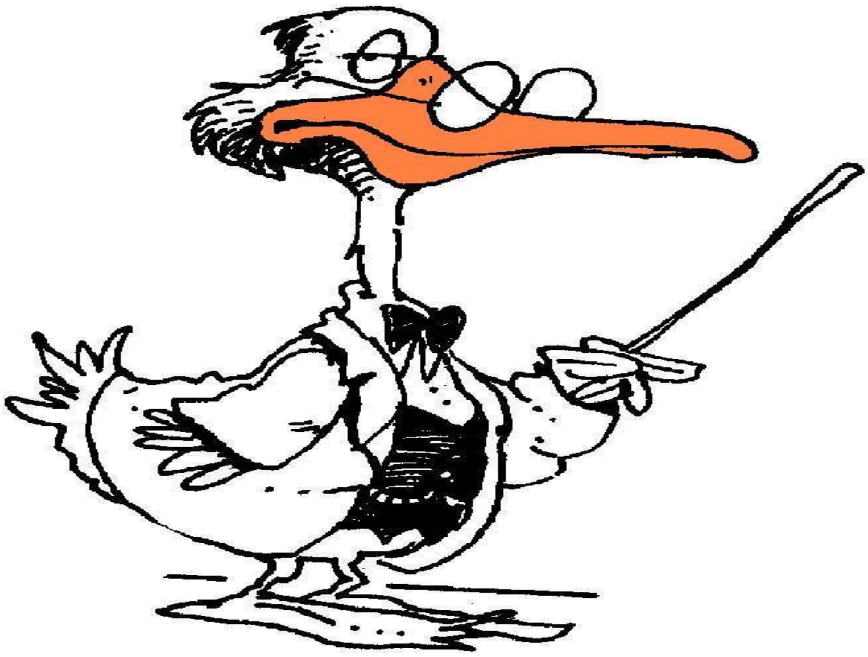
Questions?

95

Source of Elaborate Slides

Prentice Hall, Inc
Levine, et. all, First Edition

96



97

End of Chapter

98