Data Science

Deriving Knowledge from Data at Scale

Forecasting

Wee Hyong Tok

Aug 4, 2016

Forecasting

- Predict the next number in the pattern:

a) 3.7, 3.7, 3.7, 3.7, 3.7, ?

b) 2.5, 4.5, 6.5, 8.5, 10.5, ?

c) 5.0, 7.5, 6.0, 4.5, 7.0, 9.5, 8.0, 6.5, ?

Forecasting

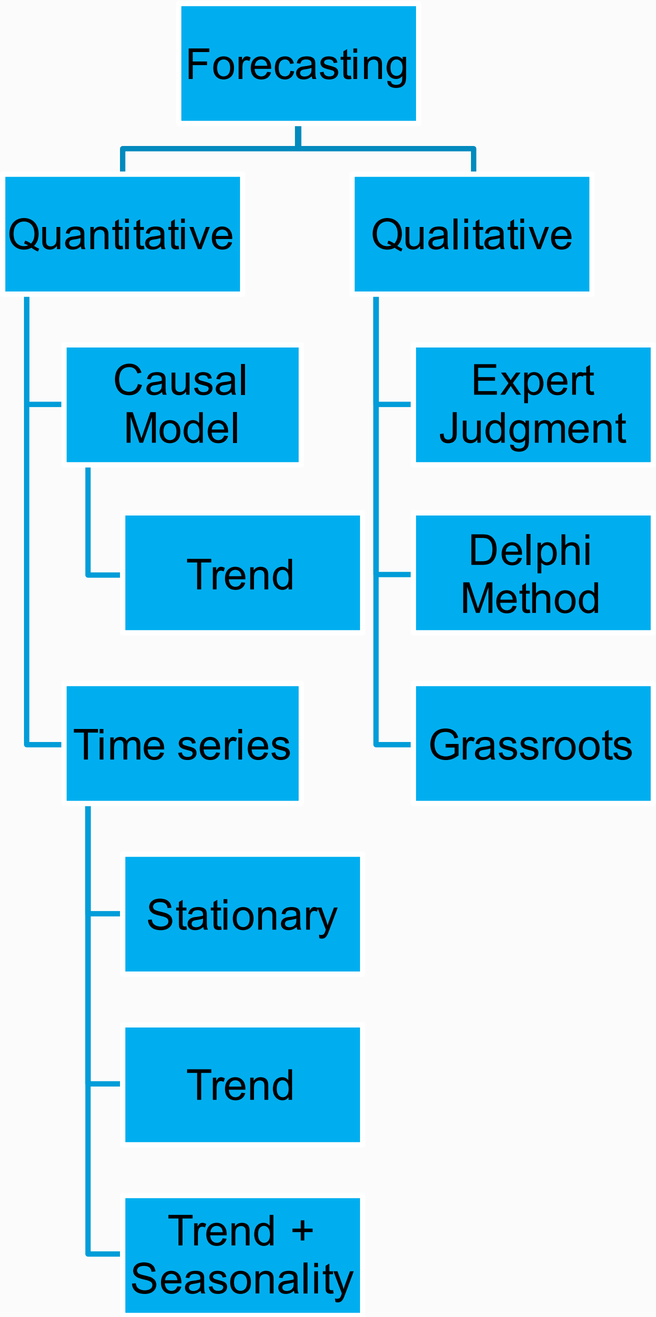
- Predict the next number in the pattern:

a) 3.7, 3.7, 3.7, 3.7, 3.7, 3.7

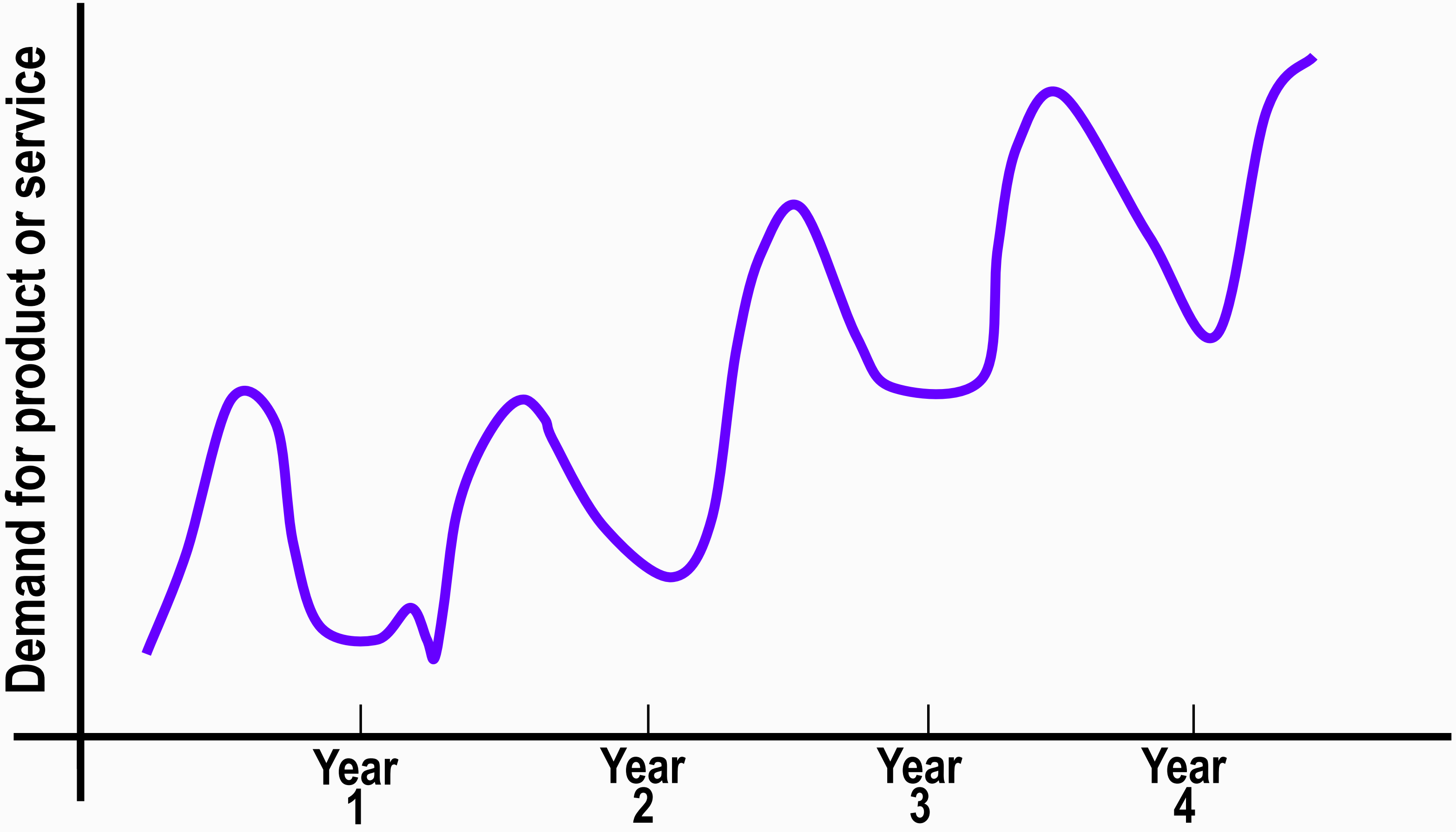
b) 2.5, 4.5, 6.5, 8.5, 10.5, 12.5

c) 5.0, 7.5, 6.0, 4.5, 7.0, 9.5, 8.0, 6.5, 9.0

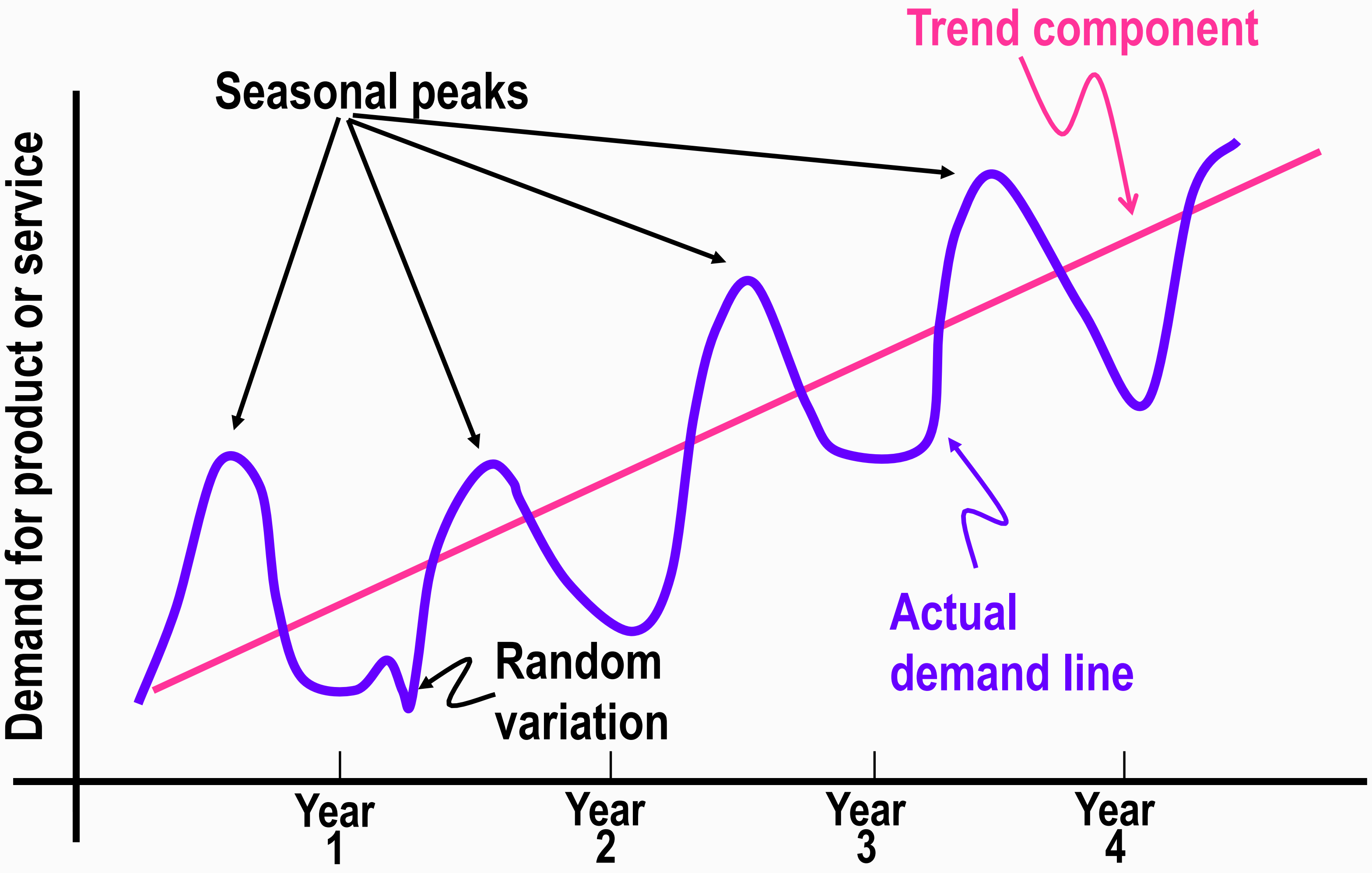
Forecasting



Product Demand over Time



Product Demand over Time



Cyclical Variation – Sample Chart

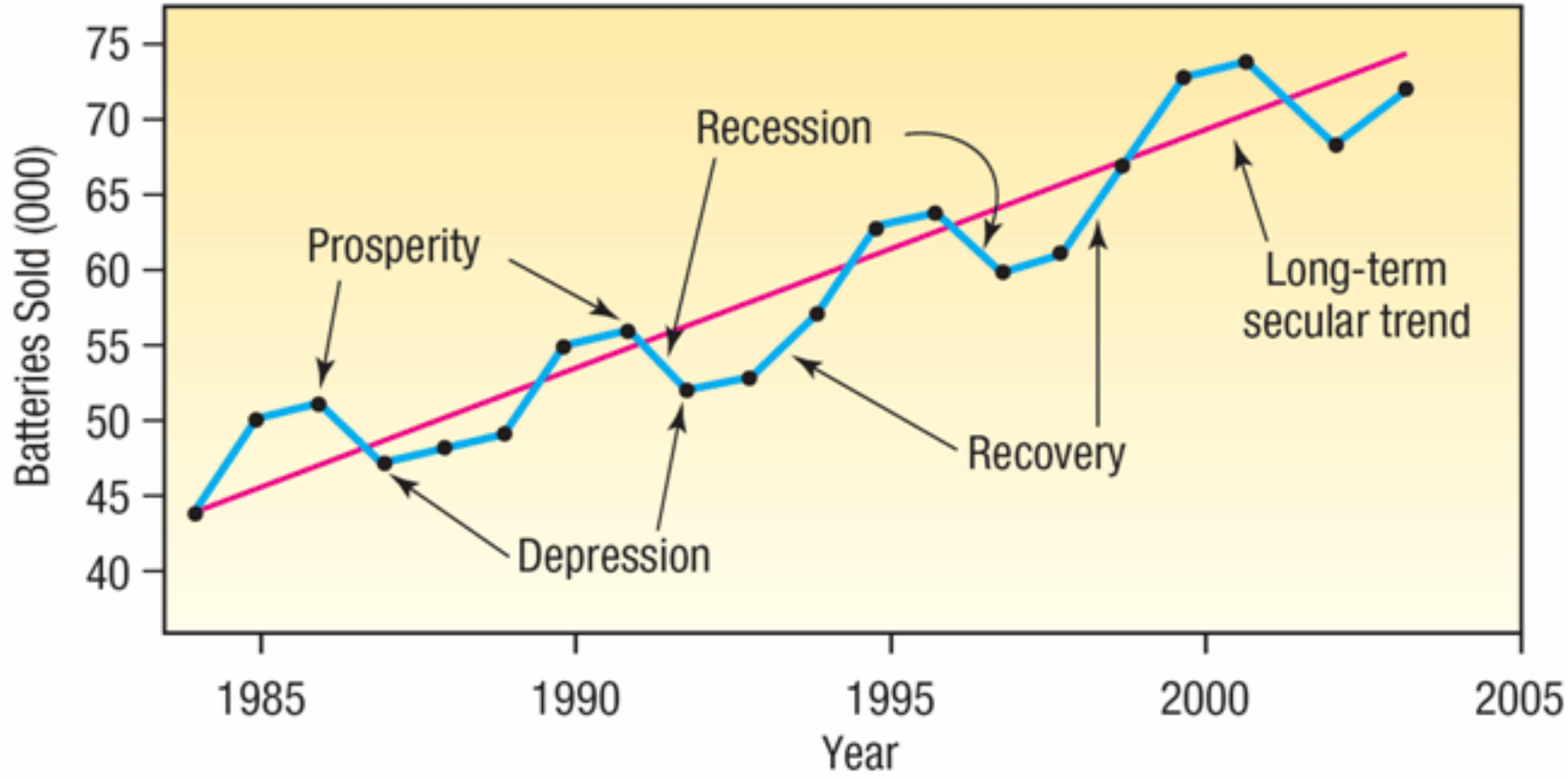


CHART 16–1 Batteries Sold by National Battery Retailers, Inc., from 1984 to 2004

Seasonal Variation – Sample Chart

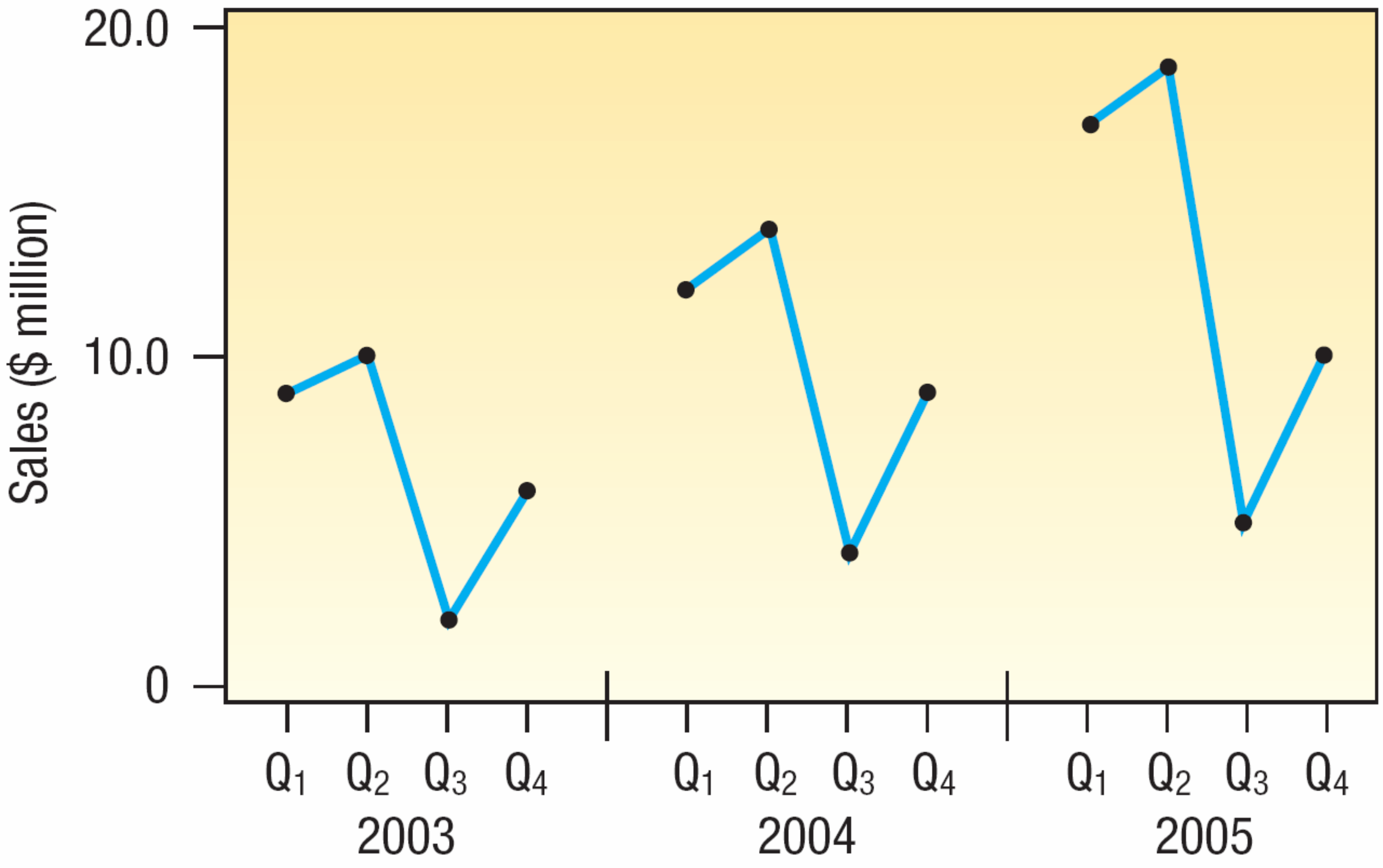


CHART 16–2 Sales of Baseball and Softball Equipment, Hercher Sporting Goods, 2003–2005 by Quarter

Outline

- What is forecasting?
- Types of forecasts
- Time-Series forecasting
  - Naïve
  - Moving Average
  - Exponential Smoothing
  - Regression
- Good forecasts

Stationary data forecasting

Naïve

- *I sold 10 units yesterday, so I think I will sell 10 units today.*

n-period moving average

- *For the past n days, I sold 12 units on average. Therefore, I think I will sell 12 units today.*

Exponential smoothing

- *I predicted to sell 10 units at the beginning of yesterday; At the end of yesterday, I found out I sold in fact 8 units. So, I will adjust the forecast of 10 (yesterday's forecast) by adding adjusted error (α * error). This will compensate over (under) forecast of yesterday.*

Naïve Model

- The simplest time series forecasting model
- Idea: “what happened last time (last year, last month, yesterday) will happen again this time”
- Naïve Model:
  - Algebraic: $F_t = Y_{t-1}$
    - Y_{t-1} : actual value in period t-1
    - F_t : forecast for period t
  - Spreadsheet: B3: = A2; Copy down

The Moving Average Method

- Useful in **smoothing time** series to see its trend
- Basic method used in measuring seasonal fluctuation
- Applicable when time series follows fairly linear trend that have definite rhythmic pattern
- Assumes average is a good estimator of future

$$F_{t+1} = \frac{A_t + A_{t-1} + A_{t-2} + \ldots + A_{t-n+1}}{n}$$

F_{t+1} = Forecast for the upcoming period, t+1
n = Number of periods to be averaged
A_t = Actual occurrence in period t

Simple Moving Average

$$F_{t+1} = \frac{A_t + A_{t-1} + A_{t-2} + ... + A_{t-n+1}}{n}$$

You're manager in Amazon's electronics department. You want to forecast ipod sales for months 4-6 using a 3-period moving average.

Month	Sales (000)	
1	4	
2	6	
3	5	
4	?	
5	?	
6	?	

Simple Moving Average

$$F_{t+1} = \frac{A_t + A_{t-1} + A_{t-2} + \ldots + A_{t-n+1}}{n}$$

You're manager in Amazon's electronics department. You want to forecast ipod sales for months 4-6 using a 3-period moving average.

Month	Sales (000)	Moving Average (n=3)
1	4	NA
2	6	NA
3	5	NA
4	?	(4+6+5)/3=5
5	?	
6	?	

Simple Moving Average

What if ipod sales were actually 3 in month 4

Month	Sales (000)	Moving Average (n=3)
1	4	NA
2	6	NA
3	5	NA
4	3	5
5	?	
6	?	

Simple Moving Average

Forecast for Month 5?

Month	Sales (000)	Moving Average (n=3)
1	4	NA
2	6	NA
3	5	NA
4	3	5
5	?	(6+5+3)/3=4.667
6	?	

Simple Moving Average

Actual Demand for Month 5 = 7

Month	Sales (000)	Moving Average (n=3)
1	4	NA
2	6	NA
3	5	NA
4	3	5
5	7	4.667
6	?	

Simple Moving Average

Forecast for Month 6?

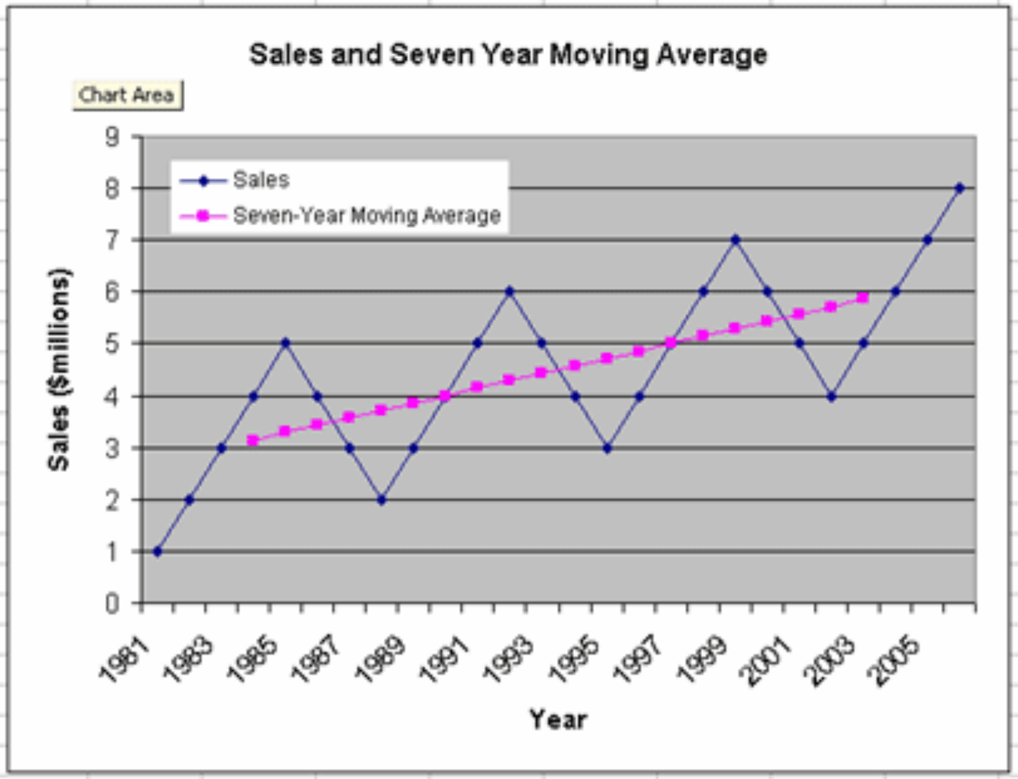
Month	Sales (000)	Moving Average (n=3)
1	4	NA
2	6	NA
3	5	NA
4	3	5
5	7	4.667
6	?	(5+3+7)/3=5

Moving Average Method - Example

Microsoft Excel - Ch16 - Forecasting

C5 =AVERAGE(B2:B8)

Year	Sales ($M)	MA7
1981	1	
1982	2	
1983	3	
1984	4	3.143
1985	5	3.286
1986	4	3.429
1987	3	3.571
1988	2	3.714
1989	3	3.857
1990	4	4.000
1991	5	4.143
1992	6	4.286
1993	5	4.429
1994	4	4.571
1995	3	4.714
1996	4	4.857
1997	5	5.000
1998	6	5.143
1999	7	5.286
2000	6	5.429
2001	5	5.571
2002	4	5.714
2003	5	5.857
2004	6	
2005	7	
2006	8	



Three-year and Five-Year Moving Averages

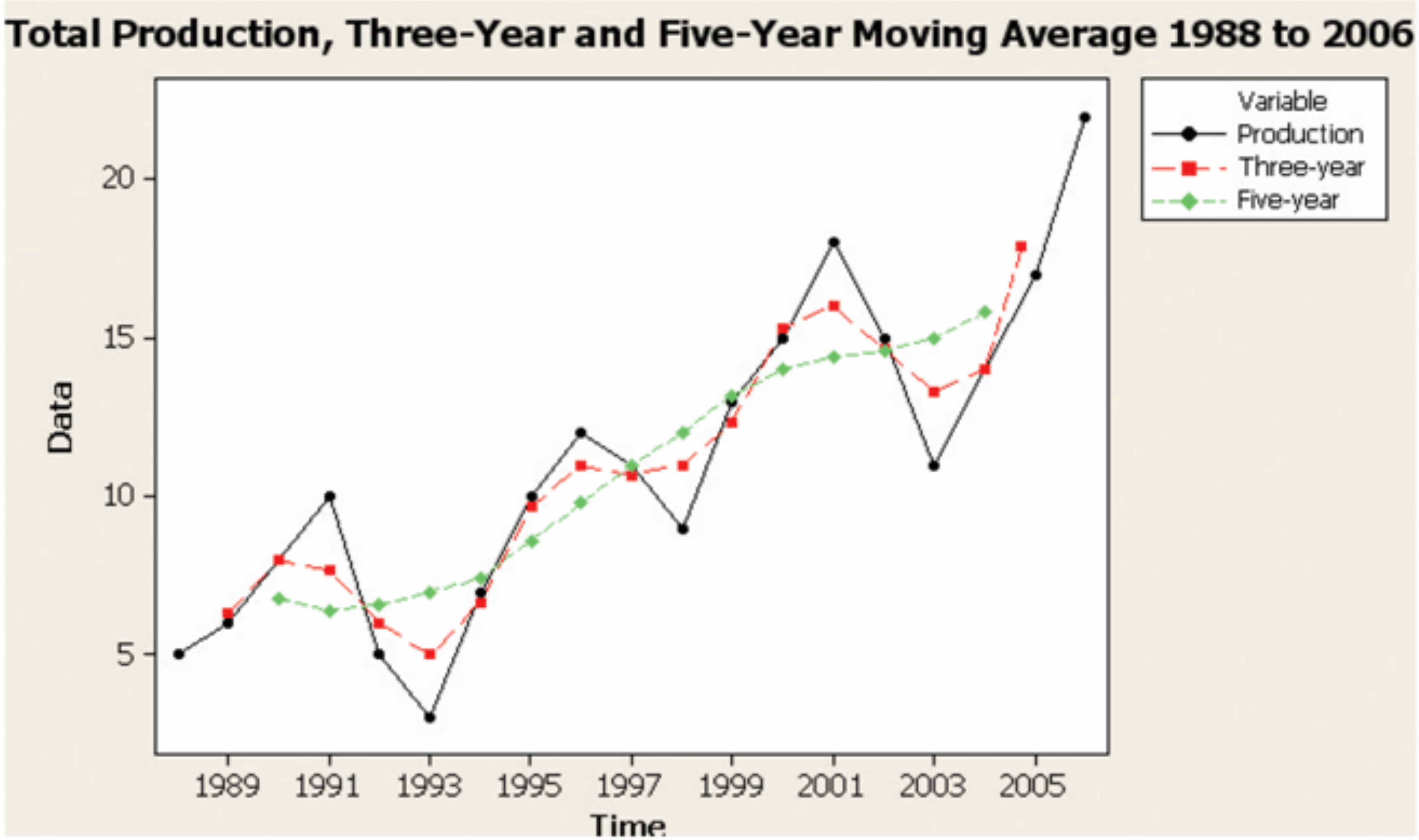


CHART 16–4 A Three-Year and Five-Year Moving Average 1988 to 2006

Weighted Moving Average

- A simple moving average assigns the same weight to each observation in averaging
- Weighted moving average assigns different weights to each observation
- Most recent observation receives the most weight, and the weight decreases for older data values
- In either case, the sum of the weights = 1

Weighted Moving Average: 3/6, 2/6, 1/6

$$F_{t+1} = w_1 A_t + w_2 A_{t-1} + w_3 A_{t-2} + \ldots + w_n A_{t-n+1}$$

Month	Sales (000)	Weighted Moving Average
1	4	NA
2	6	NA
3	5	NA
4	?	31/6 = 5.167
5	?	
6	?	

Weighted Moving Average: 3/6, 2/6, 1/6

$$F_{t+1} = w_1 A_t + w_2 A_{t-1} + w_3 A_{t-2} + \ldots + w_n A_{t-n+1}$$

Month	Sales (000)	Weighted Moving Average
1	4	NA
2	6	NA
3	5	NA
4	**3**	31/6 = 5.167
5	**7**	25/6 = 4.167
6		32/6 = 5.333

Weighted Moving Average - Example

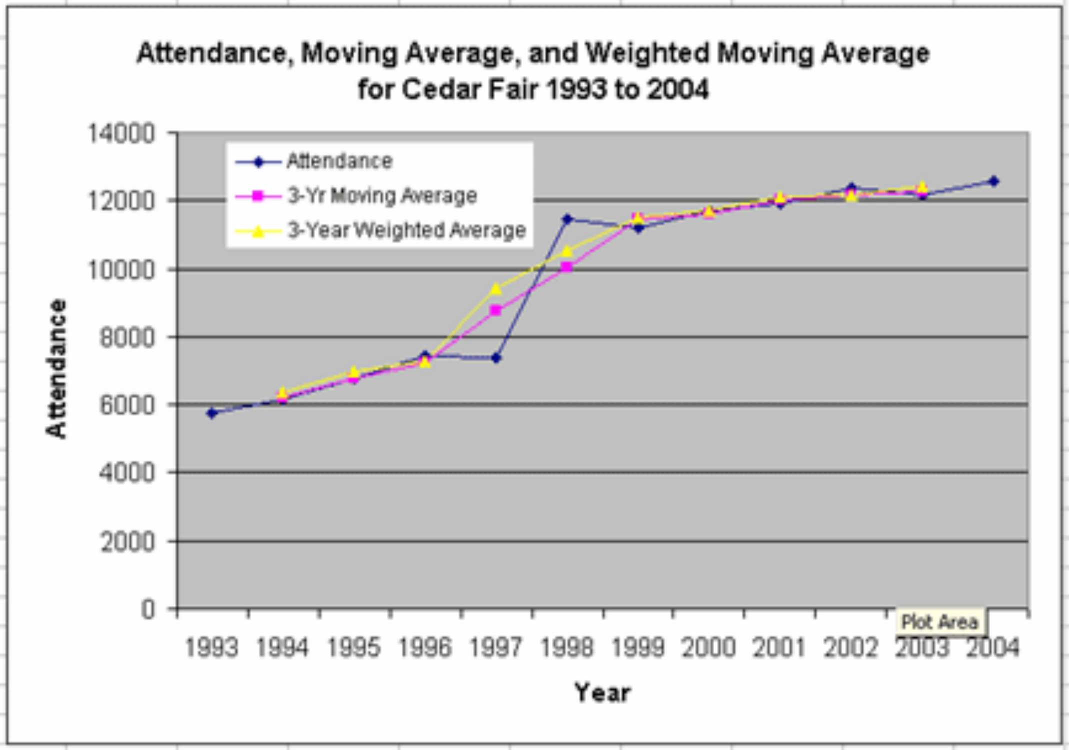
Year	Attendance (000)	Weighted Moving Average	Found by
1993	5,761		
1994	6,148	6,388	.2(5,761) + .3(6,148) + .5(6,783)
1995	6,783	6,987	.2(6,148) + .3(6,783) + .5(7,445)
1996	7,445	7,293	.2(6,783) + .3(7,445) + .5(7,405)
1997	7,405	9,436	.2(7,445) + .3(7,405) + .5(11,450)
1998	11,450	10,528	.2(7,405) + .3(11,450) + .5(11,224)
1999	11,224	11,509	.2(11,450) + .3(11,224) + .5(11,703)
2000	11,703	11,701	.2(11,224) + .3(11,703) + .5(11,890)
2001	11,890	12,098	.2(11,703) + .3(11,890) + .5(12,380)
2002	12,380	12,183	.2(11,890) + .3(12,380) + .5(12,181)
2003	12,181	12,409	.2(12,380) + .3(12,181) + .5(12,557)
2004	12,557		

Weighed Moving Average – An Example

Microsoft Excel - Ch16 - Forecasting

D3 =(0.2)*B2+(0.3)*(B3)+(0.5)*(B4)

	A	B	C	D
1	Year	Attendance	Moving Average	Weighted Average
2	1993	5761		
3	1994	6148	6230.7	6388.1
4	1995	6783	6792.0	6987
5	1996	7445	7211.0	7292.6
6	1997	7405	8766.7	9435.5
7	1998	11450	10026.3	10528
8	1999	11224	11459.0	11508.7
9	2000	11703	11605.7	11700.7
10	2001	11890	11991.0	12097.6
11	2002	12380	12150.3	12182.5
12	2003	12181	12280.5	12408.8
13	2004	12557		



Exponential Smoothing

- Concept is *simple*!
  - Make a forecast, *any* forecast
  - Compare it to the actual
  - Next forecast is
    - *Previous* forecast plus an adjustment
    - Adjustment is fraction of previous forecast error
  - Essentially
    - Not really forecast as a function of *time*
    - Instead, forecast as a function of *previous actual and forecasted value*

Exponential Smoothing

- Assumes the most recent observations have the highest predictive value
  - **gives more weight to recent time periods**

$$F_{t+1} = F_t + \alpha\underbrace{(A_t - F_t)}_{e_t}$$

F_{t+1} = Forecast value for time $t+1$

A_t = Actual value at time t

α = Smoothing constant

Need initial forecast F_t to start.

Exponential Smoothing – Example 1

$$F_{t+1} = F_t + \alpha(A_t - F_t)$$

i	Ai
Week	Demand
1	820
2	775
3	680
4	655
5	750
6	802
7	798
8	689
9	775
10	

Given the weekly demand data what are the exponential smoothing forecasts for periods 2-10 using α=0.10?

Assume $F_1=D_1$

Exponential Smoothing – Example 1

$$F_{t+1} = F_t + \alpha(A_t - F_t)$$

i	Ai	Fi
Week	Demand	α = *0.1*
1	820	820.00
2	775	
3	680	
4	655	
5	750	
6	802	
7	798	
8	689	
9	775	
10		

$F_2 = F_1 + \alpha(A_1 – F_1) = 820 + .1(820 – 820)$

$= 820$

Exponential Smoothing – Example 1

$$F_{t+1} = F_t + \alpha(A_t - F_t)$$

i	Ai	Fi
Week	Demand	α = *0.1*
1	820	820.00
2	775	820.00
3	680	
4	655	
5	750	
6	802	
7	798	
8	689	
9	775	
10		

$F_3 = F_2 + \alpha(A_2 – F_2)$ $=820+.1(775–820)$

$=815.5$

Exponential Smoothing – Example 1

$$F_{t+1} = F_t + \alpha(A_t - F_t)$$

i	Ai	Fi
Week	Demand	α = *0.1*
1	820	820.00
2	775	820.00
3	680	815.50
4	655	
5	750	
6	802	
7	798	
8	689	
9	775	
10		

This process continues through week 10

Exponential Smoothing – Example 1

$$F_{t+1} = F_t + \alpha(A_t - F_t)$$

i Week	Ai Demand	Fi α = 0.1	α = 0.6
1	820	820.00	820.00
2	775	820.00	820.00
3	680	815.50	793.00
4	655	801.95	725.20
5	750	787.26	683.08
6	802	783.53	723.23
7	798	785.38	770.49
8	689	786.64	787.00
9	775	776.88	728.20
10		776.69	756.28

What if the α constant equals 0.6

Exponential Smoothing

- How to choose **α**
  - depends on the emphasis you want to place on the most recent data

- Increasing **α** makes forecast more sensitive to recent data

Forecast Effects of Smoothing Constant α

$$F_{t+1} = F_t + \alpha\,(A_t - F_t)$$

or

$$F_{t+1} = \underbrace{\alpha}_{w_1} A_t + \underbrace{\alpha(1-\alpha)}_{w_2} A_{t-1} + \underbrace{\alpha(1-\alpha)^2}_{w_3} A_{t-2} + \ldots$$

α=	Weights		
	Prior Period	2 periods ago	3 periods ago
	α	$\alpha(1-\alpha)$	$\alpha(1-\alpha)^2$
α= 0.10	10%	9%	8.1%
α= 0.90	90%	9%	0.9%

Simple Exponential Smoothing

Properties of Simple Exponential Smoothing

- Widely used and successful model
- Requires very little data
- Larger α, more responsive forecast; Smaller α, smoother forecast

  "best" α can be found by a solver package
- Suitable for relatively stable time series

Time Series Components

Trend

- persistent upward or downward pattern in a time series

Seasonal

- Variation dependent on the time of year
- Each year shows same pattern

Cyclical

- up & down movement repeating over long time frame
- Each year does not show same pattern

Noise or random fluctuations

- follow no specific pattern
- short duration and non-repeating

Linear Trend Plot

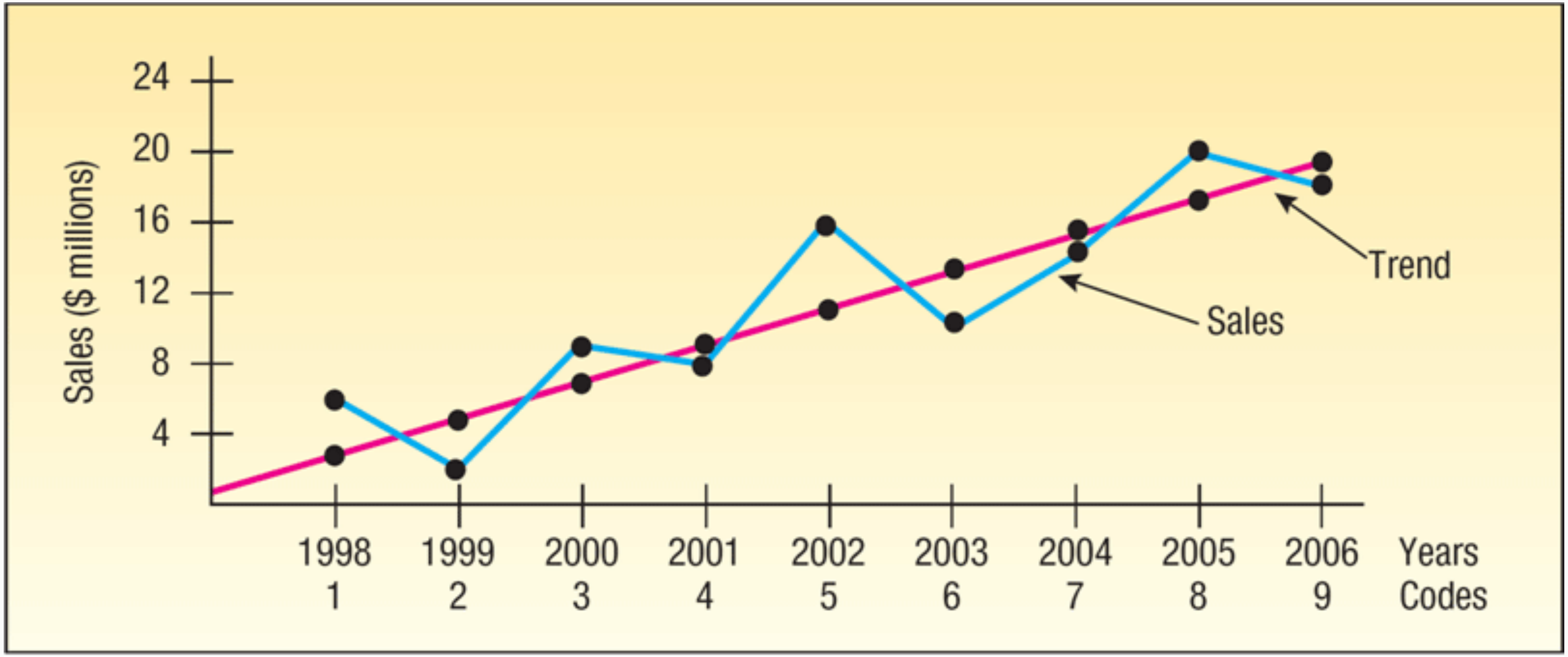
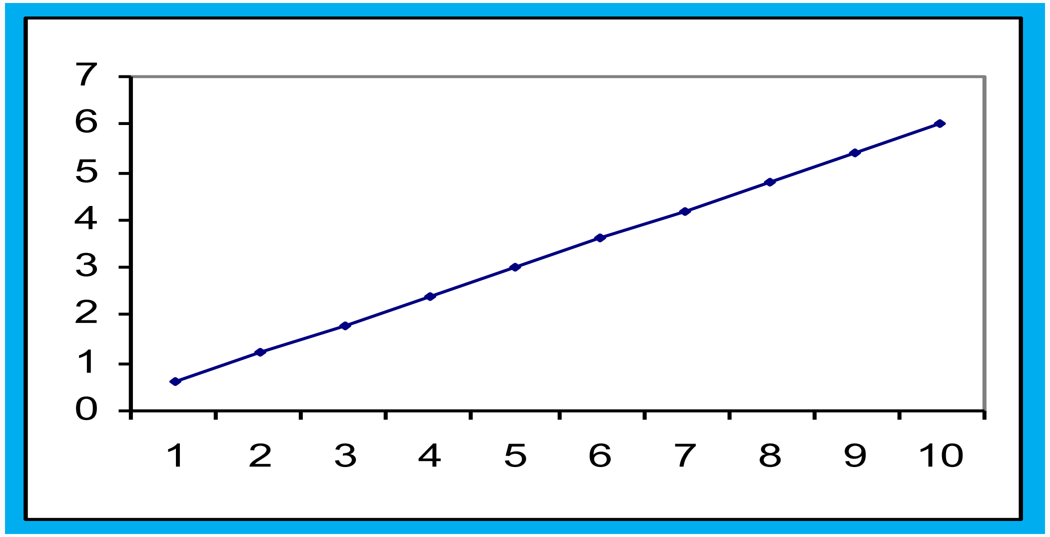


CHART 16–5 A Straight Line Fitted to Sales Data

Trend Model

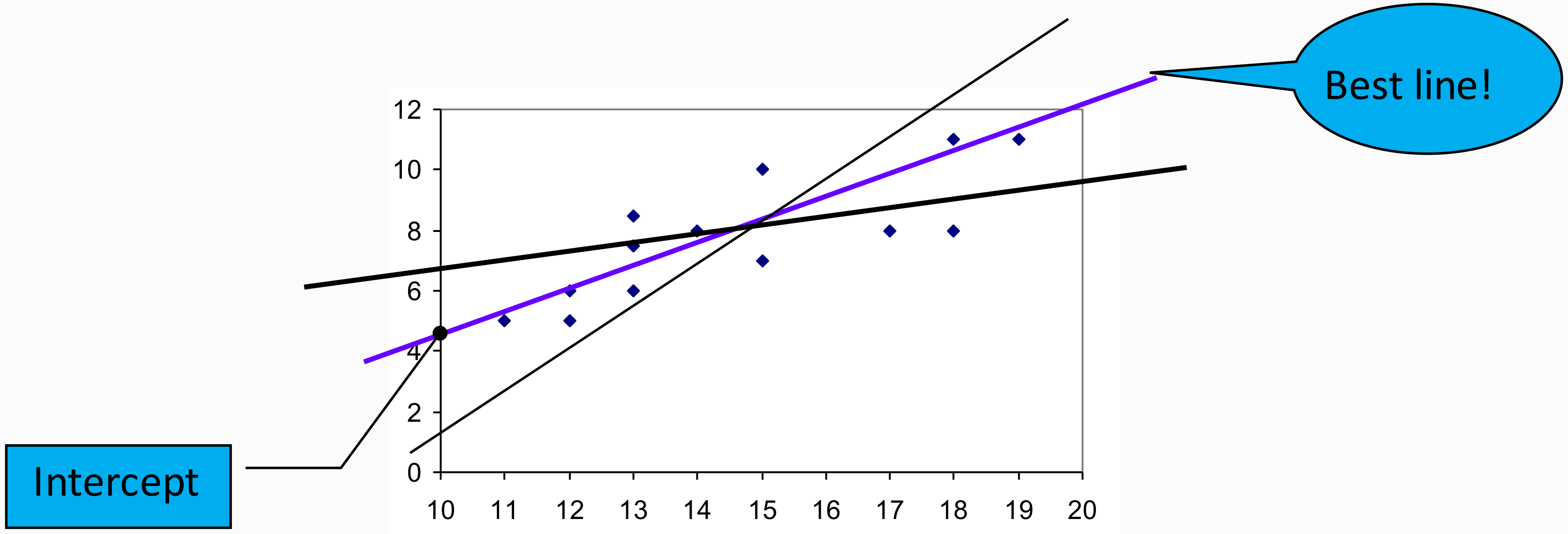
- Curve fitting method used for time series data (also called time series regression model)
- Useful when the time series has a clear trend
- Can not capture seasonal patterns
- Linear Trend Model: $Y_t = a + bt$
  - t is time index for each period, $t = 1, 2, 3, \ldots$



Pattern-based forecasting - Trend

Regression – Recall Independent Variable X, which is now time variable – e.g., days, months, quarters, years etc.

- *Find a straight line that fits the data best.*



- *y = Intercept + slope * x   (e.g. $y = b_0 + b_1 x$)*
- *Slope = change in y / change in x*

Pattern-based forecasting – Seasonal

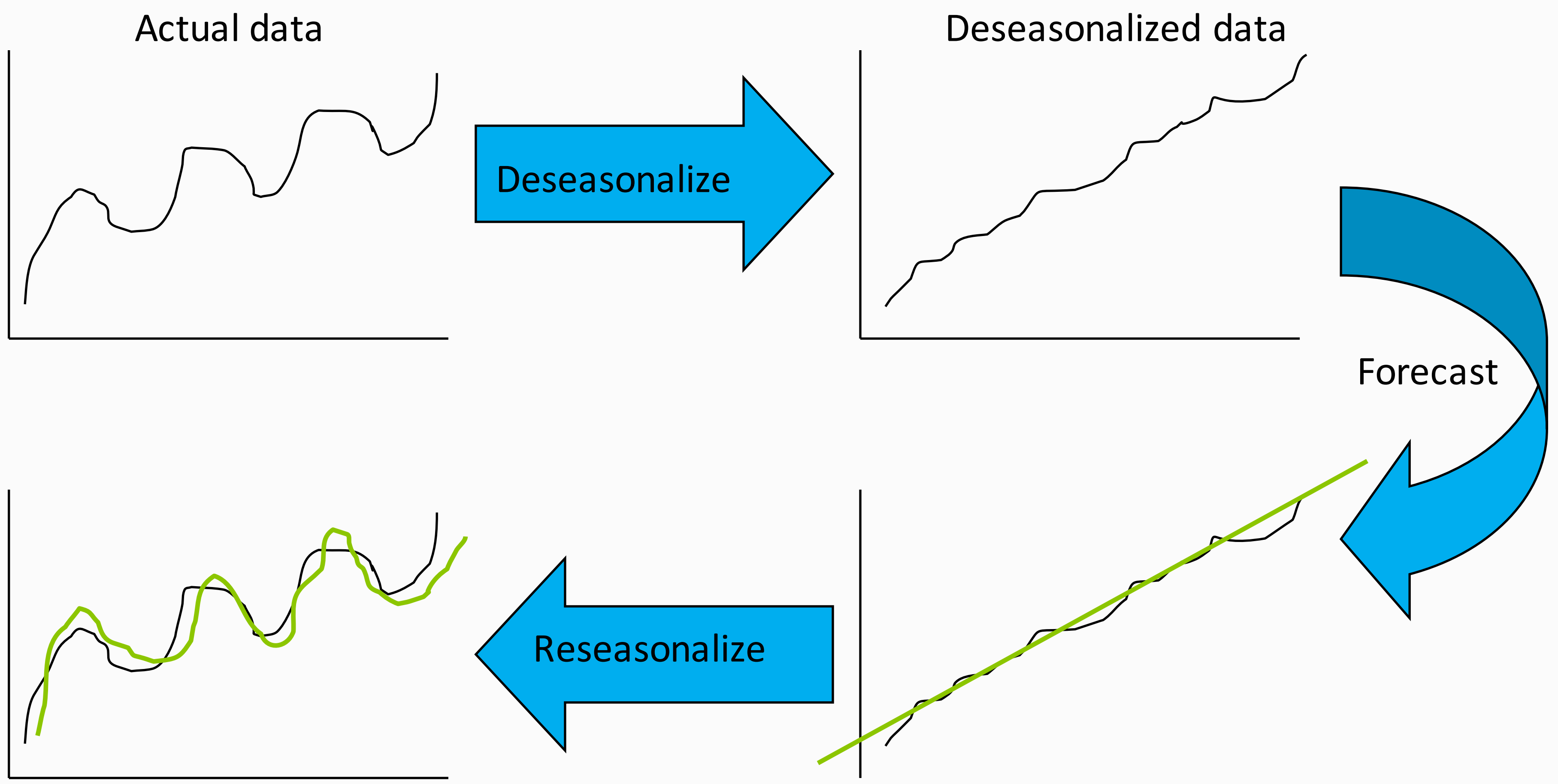
- Once data turn out to be seasonal, ***deseasonalize*** the data.
  - The methods we have learned (Heuristic methods and Regression) is not suitable for data that has pronounced fluctuations.

- Make forecast based on the deseasonalized data

- ***Reseasonalize*** the forecast
  - Good forecast should mimic reality. Therefore, it is needed to give seasonality back.

Pattern-based forecasting – Seasonal

Example (SI + Regression)



Linear Trend – Using the Least Squares Method

- Use the least squares method in Simple Linear Regression to find the best linear relationship between 2 variables

- Code time (t) and use it as the independent variable

- E.g. let t be 1 for the first year, 2 for the second, and so on (if data are annual)

Linear Trend – Using the Least Squares Method: An Example

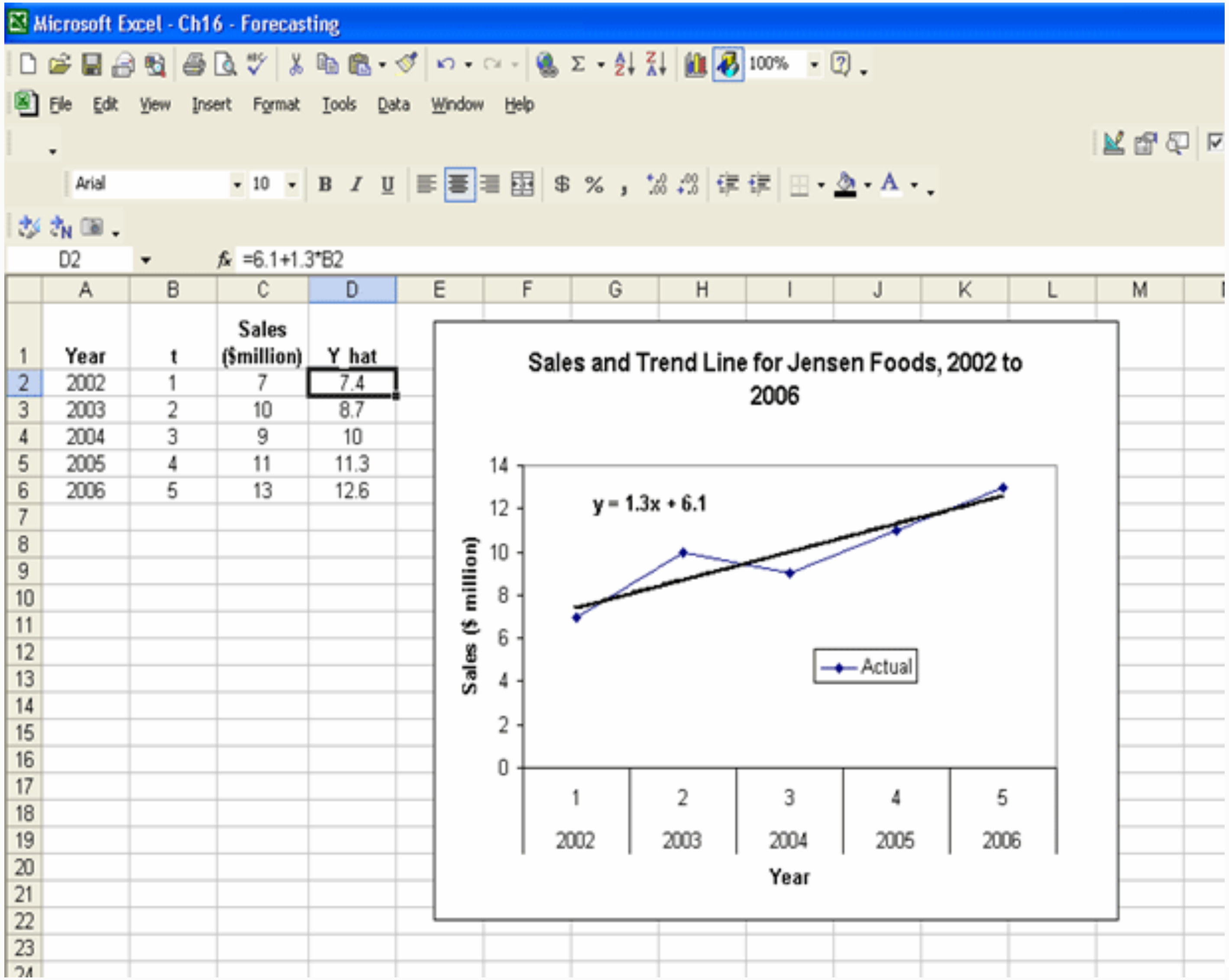
The sales of Jensen Foods, a small grocery chain located in southwest Texas, since 2002 are:

Year	Sales ($ mil.)
2002	7
2003	10
2004	9
2005	11
2006	13

Year	*t*	Sales ($ mil.)
2002	1	7
2003	2	10
2004	3	9
2005	4	11
2006	5	13

Linear Trend – Using the Least Squares Method

Example Using Excel



Year	t	Sales ($million)	Y hat
2002	1	7	7.4
2003	2	10	8.7
2004	3	9	10
2005	4	11	11.3
2006	5	13	12.6

Nonlinear Trends

- A linear trend equation is used when the data are increasing (or decreasing) by <u>equal</u> amounts
- A nonlinear trend equation is used when the data are increasing (or decreasing) by <u>increasing</u> amounts over time
- When data increase (or decrease) by equal *percents or proportions* plot will show curvilinear pattern

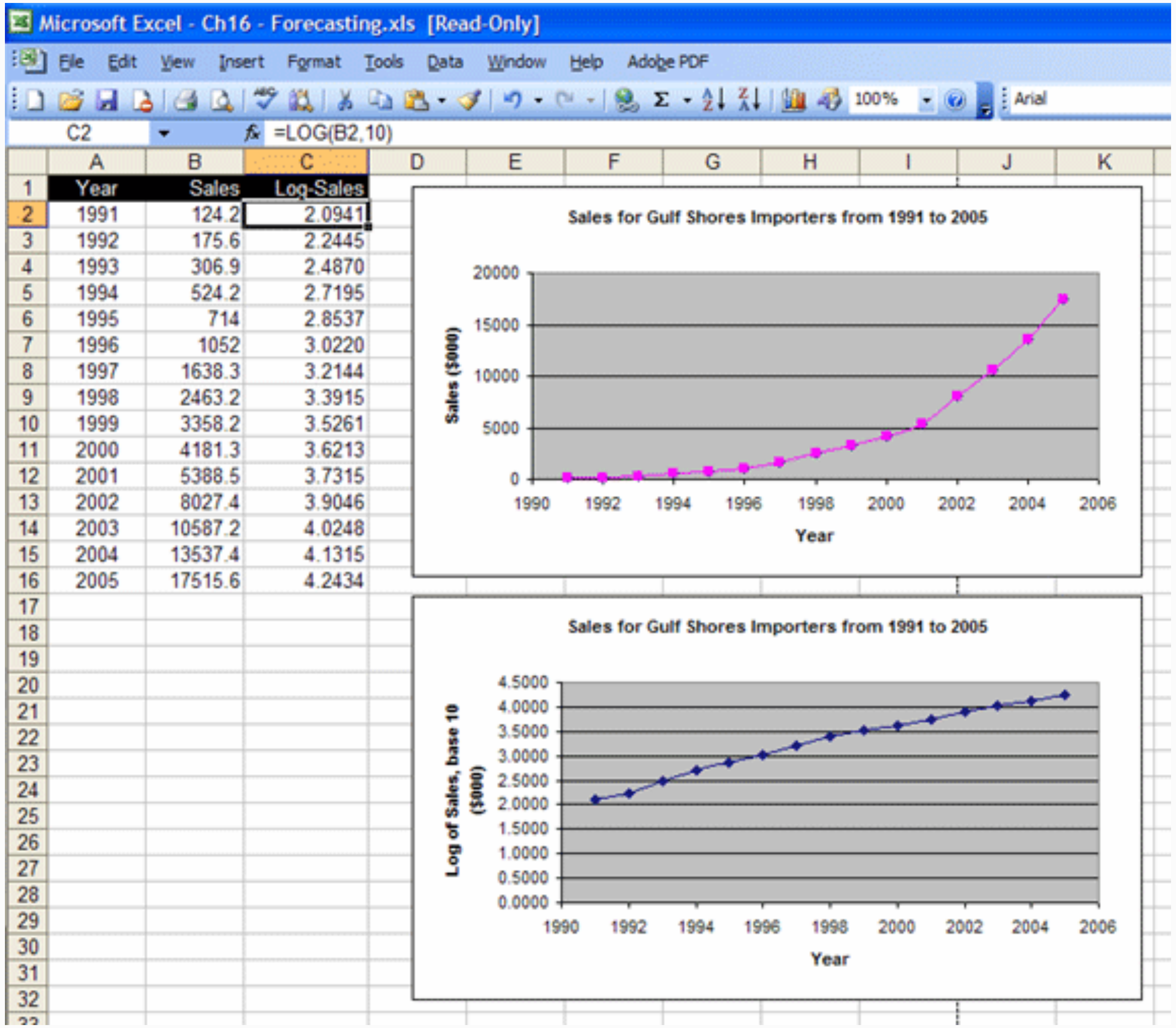
Log Trend Equation – Gulf Shores Importers Example

- Top graph is plot of the original data
- Bottom graph is the log base 10 of the original data which now is linear

  (Excel function:

   =log(x) or log(x,10)

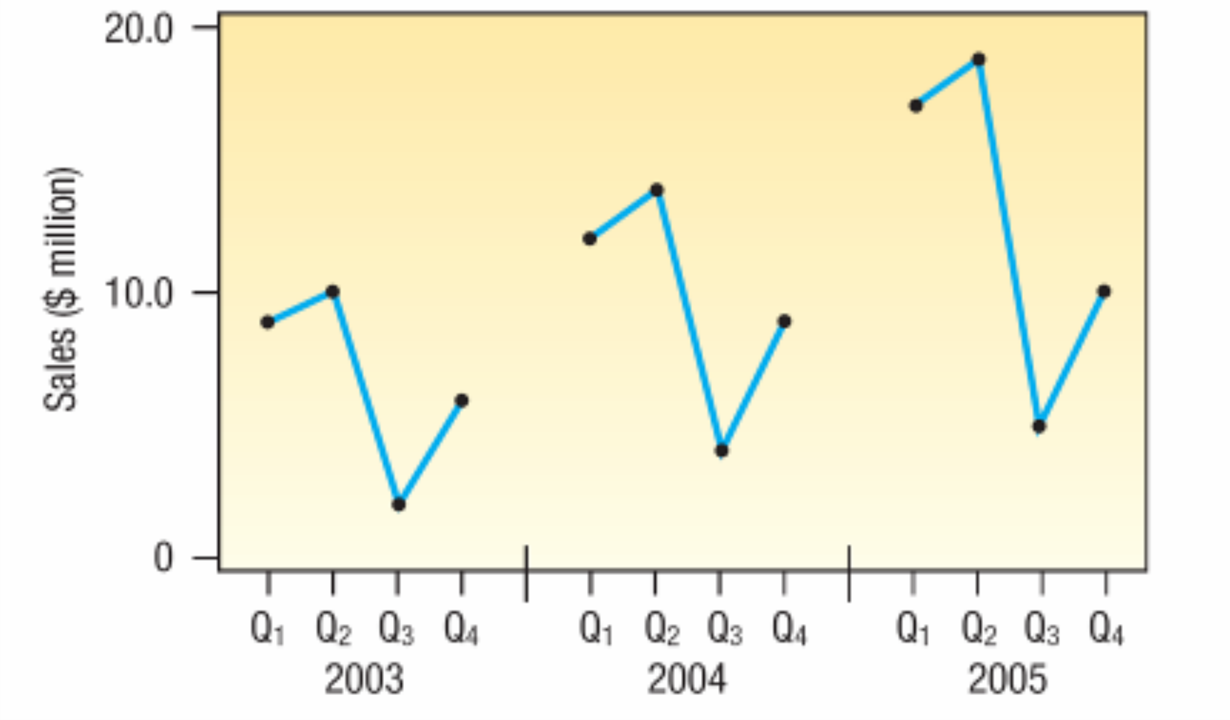
- Using Data Analysis in Excel, generate the linear equation



Year	Sales	Log-Sales
1991	124.2	2.0941
1992	175.6	2.2445
1993	306.9	2.4870
1994	524.2	2.7195
1995	714	2.8537
1996	1052	3.0220
1997	1638.3	3.2144
1998	2463.2	3.3915
1999	3358.2	3.5261
2000	4181.3	3.6213
2001	5388.5	3.7315
2002	8027.4	3.9046
2003	10587.2	4.0248
2004	13537.4	4.1315
2005	17515.6	4.2434

Seasonal Variation

- One of the components of a time series
- Seasonal variations are fluctuations that coincide with certain seasons and are repeated year after year
- Understanding seasonal fluctuations help plan for sufficient goods and materials on hand to meet varying seasonal demand
- Analysis of seasonal fluctuations over a period of years help in evaluating current sales



Pattern-based forecasting – Seasonal

- Deseasonalization
  - *Deseasonalized data = Actual / SI*
- Reseasonalization
  - *Reseasonalized forecast*

    *= deseasonalized forecast * SI*

Seasonal Index

What's an *index*?

- Ratio
- SI = ratio between actual and average demand

Suppose

- SI for quarter demand is 1.20
  - What's that mean?
  - Use it to forecast demand for next fall

So, ***where*** did the 1.20 come from?!

Calculating Seasonal Indices

- Quick and dirty method of calculating SI
  - For each ***year***, calculate average demand
  - Divide each demand by its yearly average
    - This creates a ratio and hence a ***raw index***
    - For each ***quarter***, there will be as many raw indices as there are years
  - Average the raw indices for each of the quarters
  - The result will be ***four*** values, one SI per quarter

Seasonal Index – An Example

The table below shows the quarterly sales for Toys International for the years 2001 through 2006. The sales are reported in millions of dollars. Determine a quarterly seasonal index using the ratio-to-moving-average method.

Year	Winter	Spring	Summer	Fall
2001	6.7	4.6	10.0	12.7
2002	6.5	4.6	9.8	13.6
2003	6.9	5.0	10.4	14.1
2004	7.0	5.5	10.8	15.0
2005	7.1	5.7	11.1	14.5
2006	8.0	6.2	11.4	14.9

Step (1) – Organize time series data in column form

Step (2) Compute the 4-quarter moving totals

Step (3) Compute the 4-quarter moving averages

Step (4) Compute the centered moving averages by getting the average of two 4-quarter moving averages

Step (5) Compute ratio by dividing actual sales by the centered moving averages

Year	Quarter	(1) Sales ($ millions)	(2) Four-Quarter Total	(3) Four-Quarter Moving Average	(4) Centered Moving Average	(5) Specific Seasonal
2001	Winter	6.7				
	Spring	4.6				
			34.0	8.500		
	Summer	10.0			8.475	1.180
			33.8	8.450		
	Fall	12.7			8.450	1.503
			33.8	8.450		
2002	Winter	6.5			8.425	0.772
			33.6	8.400		
	Spring	4.6			8.513	0.540
			34.5	8.625		
	Summer	9.8			8.675	1.130
			34.9	8.725		
	Fall	13.6			8.775	1.550
			35.3	8.825		
2003	Winter	6.9			8.900	0.775
			35.9	8.975		
	Spring	5.0			9.038	0.553
			36.4	9.100		
	Summer	10.4			9.113	1.141
			36.5	9.125		
	Fall	14.1			9.188	1.535
			37.0	9.250		
2004	Winter	7.0			9.300	0.753
			37.4	9.350		
	Spring	5.5			9.463	0.581
			38.3	9.575		
	Summer	10.8			9.588	1.126
			38.4	9.600		
	Fall	15.0			9.625	1.558
			38.6	9.650		
2005	Winter	7.1			9.688	0.733
			38.9	9.725		
	Spring	5.7			9.663	0.590
			38.4	9.600		
	Summer	11.1			9.713	1.143
			39.3	9.825		
	Fall	14.5			9.888	1.466
			39.8	9.950		
2006	Winter	8.0			9.888	0.801
			40.1	10.025		
	Spring	6.2			10.075	0.615
			40.5	10.125		
	Summer	11.4				
	Fall	14.9				

Seasonal Index – An Example

Year	Winter	Spring	Summer	Fall	
2001			1.180	1.503	
2002	0.772	0.540	1.130	1.550	
2003	0.775	0.553	1.141	1.535	
2004	0.753	0.581	1.126	1.558	
2005	0.733	0.590	1.143	1.466	
2006	0.801	0.615			
Total	3.834	2.879	5.720	7.612	
Mean	0.767	0.576	1.144	1.522	4.009
Adjusted	0.765	0.575	1.141	1.519	4.000
Index	76.5	57.5	114.1	151.9	

CORRECTION FACTOR FOR ADJUSTING QUARTERLY MEANS

$$\text{Correction factor} = \frac{4.00}{\text{Total of four means}}$$

$$\text{Correction factor} = \frac{4.00}{4.009} = 0.997755$$

Actual versus Deseasonalized Sales for Toys International

Deseasonalized Sales = Sales / Seasonal Index

Year	Quarter	(1) Sales	(2) Seasonal Index	(3) Deseasonalized Sales
2001	Winter	6.7	0.765	8.76
	Spring	4.6	0.575	8.00
	Summer	10.0	1.141	8.76
	Fall	12.7	1.519	8.36
2002	Winter	6.5	0.765	8.50
	Spring	4.6	0.575	8.00
	Summer	9.8	1.141	8.59
	Fall	13.6	1.519	8.95
2003	Winter	6.9	0.765	9.02
	Spring	5.0	0.575	8.70
	Summer	10.4	1.141	9.11
	Fall	14.1	1.519	9.28
2004	Winter	7.0	0.765	9.15
	Spring	5.5	0.575	9.57
	Summer	10.8	1.141	9.47
	Fall	15.0	1.519	9.87
2005	Winter	7.1	0.765	9.28
	Spring	5.7	0.575	9.91
	Summer	11.1	1.141	9.73
	Fall	14.5	1.519	9.55
2006	Winter	8.0	0.765	10.46
	Spring	6.2	0.575	10.79
	Summer	11.4	1.141	9.99
	Fall	14.9	1.519	9.81

Classical Decomposition

- Start by calculating seasonal indices
- Then, *deseasonalize* the demand
  - Divide actual demand values by their SI values

    $y' = y / SI$
  - Results in transformed data (new time series)
  - Seasonal effect removed
- Forecast
  - Regression if deseasonalized data is trendy
  - Heuristics methods if deseasonalized data is stationary
- Reseasonalize with SI

…but what is a good forecast?

A Good Forecast

Has a small error

- Error = Demand - Forecast

Measures of Forecast Error

a. MAD = Mean Absolute Deviation

$$\text{MAD} = \frac{\sum_{t=1}^{n} \overbrace{|A_t - F_t|}^{e_t}}{n}$$

b. **MSE = Mean Squared Error**

$$\text{MSE} = \frac{\sum_{t=1}^{n} (A_t - F_t)^2}{n}$$

c. RMSE = Root Mean Squared Error

$$\text{RMSE} = \sqrt{\text{MSE}}$$

- Ideal values =0 (i.e., no forecasting error)

MAD Example

$$MAD = \frac{\sum_{t=1}^{n} |A_t - F_t|}{n} = \frac{40}{4} = 10$$

What is the MAD value given the forecast values in the table below?

Month	Sales (A_t)	Forecast (F_t)	$\lvert A_t - F_t \rvert$
1	220	n/a	
2	250	255	5
3	210	205	5
4	300	320	20
5	325	315	10

$$\sum_{t=1}^{n} |A_t - F_t| = 40$$

MSE/RMSE Example

$$MSE = \frac{\sum_{t=1}^{n} (A_t - F_t)^2}{n} = \frac{550}{4} = 137.5$$

What is the MSE value?

$RMSE = \sqrt{137.5}$

$= 11.73$

Month	Sales (A_t)	Forecast (F_t)	$\lvert A_t - F_t \rvert$	$(A_t - F_t)^2$
1	220	n/a		
2	250	255	5	25
3	210	205	5	25
4	300	320	20	400
5	325	315	10	100

$$\sum_{t=1}^{n} (A_t - F_t)^2 = 550$$

Forecast Bias

- How can we tell if a forecast has a positive or negative bias?

- TS = Tracking Signal
  - Good tracking signal has low values

$$\mathrm{TS} = \frac{\mathrm{RSFE}}{\mathrm{MAD}} = \frac{{}_{t}\,(actual_t \quad forecast_t)}{\mathbf{MAD}}$$

Can you…

- describe general forecasting process?
- compare and contrast trend, seasonality and cyclicality?
- describe the forecasting method when data is stationary?
- describe the forecasting method when data shows trend?
- describe the forecasting method when data shows seasonality?

Data Science

Deriving Knowledge from Data at Scale

That’s all for tonight....