CHAPTER 4

Forecasting

DISCUSSION QUESTIONS

1. *Qualitative models* incorporate subjective factors into the forecasting model. Qualitative models are useful when subjective factors are important. When quantitative data are difficult to obtain, qualitative models may be appropriate.

2. Approaches are qualitative and quantitative. Qualitative is relatively subjective; quantitative uses numeric models.

3. Short-range (under 3 months), medium-range (3 to 18 months), and long-range (over 18 months).

4. The steps that should be used to develop a forecasting system are:

(a) Determine the purpose and use of the forecast

(b) Select the item or quantities that are to be forecasted

(c) Determine the time horizon of the forecast

(d) Select the type of forecasting model to be used

(e) Gather the necessary data

(f) Validate the forecasting model

(g) Make the forecast

(h) Implement the results

(i) Evaluate the results

5. Any three of: sales planning, production planning and budgeting, cash budgeting, analyzing various operating plans.

6. There is no mechanism for growth in these models; they are built exclusively from historical demand values. Such methods will always lag trends.

7. *Exponential smoothing* is a weighted moving average where all previous values are weighted with a set of weights that decline exponentially.

8. MAD, MSE, and MAPE are common measures of forecast accuracy. To find the more accurate forecasting model, forecast with each tool for several periods where the demand outcome is known, and calculate MSE, MAPE, or MAD for each. The smaller error indicates the better forecast.

9. The *Delphi technique* involves:

(a) Assembling a group of experts in such a manner as to preclude direct communication between identifiable members of the group

(b) Assembling the responses of each expert to the questions or problems of interest

(c) Summarizing these responses

(d) Providing each expert with the summary of all responses

(e) Asking each expert to study the summary of the responses and respond again to the questions or problems of interest.

(f) Repeating steps (b) through (e) several times as necessary to obtain convergence in responses. If convergence has not been obtained by the end of the fourth cycle, the responses at that time should probably be accepted and the process terminated—little additional convergence is likely if the process is continued.

10. A time series model predicts on the basis of the assumption that the future is a function of the past, whereas a causal model incorporates into the model the variables of factors that might influence the quantity being forecast.

11. A time series is a sequence of evenly spaced data points with the four components of trend, seasonality, cyclical, and random variation.

12. When the smoothing constant, α, is large (close to 1.0), more weight is given to recent data; when α is low (close to 0.0), more weight is given to past data.

13. Seasonal patterns are of fixed duration and repeat regularly. Cycles vary in length and regularity. Seasonal indexes allow "generic" forecasts to be made specific to the month, week, etc., of the application.

14. *Exponential smoothing* weighs all previous values with a set of weights that decline exponentially. It can place a full weight on the most recent period (with an alpha of 1.0). This, in effect, is the *naïve approach*, which places all its emphasis on last period's actual demand.

15. Adaptive forecasting refers to computer monitoring of tracking signals and self-adjustment if a signal passes its present limit.

16. *Tracking signals* alert the user of a forecasting tool to periods in which the forecast was in significant error.

17. The correlation coefficient measures the degree to which the independent and dependent variables move together. A negative value would mean that as X increases, Y tends to fall. The variables move together, but move in opposite directions.

18. *Independent variable* (x) is said to cause variations in the *dependent variable* (y).

19. Nearly every industry has seasonality. The seasonality must be filtered out for good medium-range planning (of production and inventory) and performance evaluation.

20. There are many examples. Demand for raw materials and component parts such as steel or tires is a function of demand for goods such as automobiles.

21. Obviously, as we go farther into the future, it becomes more difficult to make forecasts, and we must diminish our reliance on the forecasts.

ETHICAL DILEMMA

This exercise, derived from an actual situation, deals as much with ethics as with forecasting. Here are a few points to consider:

- No one likes a system they don't understand, and most college presidents would feel uncomfortable with this one. It does offer the advantage of depoliticizing the funds allocation if used wisely and fairly. But to do so means all parties must have input to the process (such as smoothing constants) and all data need to be open to everyone.
- The smoothing constants could be selected by an agreed-upon criteria (such as lowest *MAD*) or could be based on input from experts on the board as well as the college.
- Abuse of the system is tied to assigning alphas based on what results they yield, rather than what alphas make the most sense.
- Regression is open to abuse as well. Models can use *many* years of data yielding one result, or *few* years yielding a totally different forecast. Selection of associative variables can have a major impact on results as well.

ACTIVE MODEL EXERCISES

ACTIVE MODEL 4.1: Moving Averages

1. What does the graph look like when n = 1

The forecast graph mirrors the data graph but one period later.

2. What happens to the GRAPH as the number of periods in the moving average increases?

The forecast graph becomes shorter and smoother.

3. What value for n minimizes the MAD for this data?

n = 1 (a naive forecast)

ACTIVE MODEL 4.2: Exponential Smoothing

1. What happens to the graph when alpha equals zero?

The graph is a straight line. The forecast is the same in each period.

2. What happens to the graph when alpha equals one?

The forecast follows the same pattern as the demand (except for the first forecast) but is offset by one period. This is a naive forecast.

3. Generalize what happens to a forecast as alpha increases.

As alpha increases the forecast is more sensitive to changes in demand.

4. At what level of alpha is the mean absolute deviation (MAD) minimized?

Alpha = .16

ACTIVE MODEL 4.3: Exponential Smoothing with Trend Adjustment

1. Scroll through different values for alpha and beta. Which smoothing constant appears to have the greater affect on the graph?

Alpha

2. With beta set to zero, find the best alpha and observe the MAD. Now find the best beta. Observe the MAD. Does the addition of a trend improve the forecast?

Alpha = .11, MAD = 2.59; Beta above .6 changes the MAD (by a little) to 2.54.

ACTIVE MODEL 4.4: Trend Projections

1. What is the annual trend in the data?

10.54

2. Use the scrollbars for the slope and intercept to determine the values that minimize the MAD. Are these the same values that regression yields?

No they are NOT the same values. For example, an intercept of 57.81 with a slope of 9.44 yields a MAD of 7.17.

END-OF-CHAPTER PROBLEMS

4.1 (a) $\dfrac{374 + 368 + 381}{3} = 374.33$ pints

(b)

Week of	Pints Used	Weighted Moving Average
August 31	360	
September 7	389	
September 14	410	
September 21	381	
September 28	368	
October 5	374	
		Forecast 372.9

381 × .1 = 38.1
368 × .3 = 110.4
374 × .6 = 224.4
372.9 → Forecast 372.9

(c)

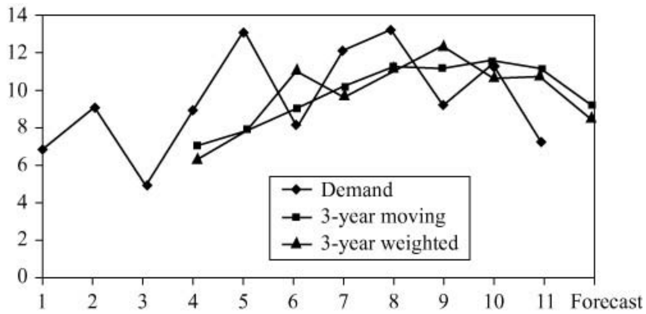
Week of	Pints	Forecast	Forecasting Error	Error × .20	Forecast
August 31	360	360	0	0	360
September 7	389	360	29	5.8	365.8
September 14	410	365.8	44.2	8.84	374.64
September 21	381	374.64	6.36	1.272	375.912
September 28	368	375.912	−7.912	−1.5824	374.3296
October 5	374	374.3296	−.3296	−.06592	374.2636

The forecast is 374.26.

4.2 (a) No, the data appear to have no consistent pattern.

	Year	1	2	3	4	5	6	7	8	9	10	11	Forecast
	Demand	7	9	5	9.0	13.0	8.0	12.0	13.0	9.0	11.0	7.0	
(b)	3-year moving				7.0	7.7	9.0	10.0	11.0	11.0	11.3	11.0	9.0
(c)	3-year weighted				6.4	7.8	11.0	9.6	10.9	12.2	10.5	10.6	8.4

(d) The three-year moving average appears to give better results.



4.5 (c) Weighted 2 year M.A. with .6 weight for most recent year.

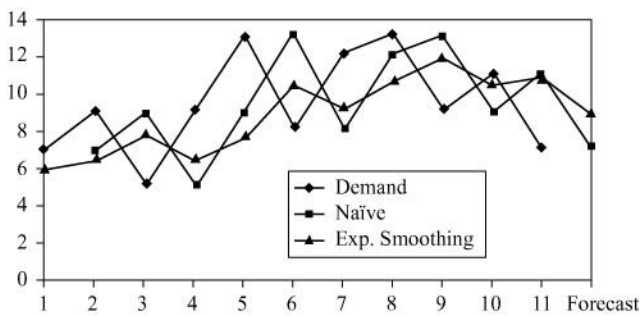
Year	Mileage	Forecast	Error	|Error|
1	3,000			
2	4,000			
3	3,400	3,600	–200	200
4	3,800	3,640	160	160
5	3,700	3,640	60	60
				420

Forecast for year 6 is 3,740 miles.

$$MAD = 140\left(= \frac{420}{3}\right)$$

4.3

Year	1	2	3	4	5	6	7	8	9	10	11	Forecast
Demand	7	9.0	5.0	9.0	13.0	8.0	12.0	13.0	9.0	11.0	7.0	
Naïve		7.0	9.0	5.0	9.0	13.0	8.0	12.0	13.0	9.0	11.0	7.0
Exp. Smoothing	6	6.4	7.4	6.5	7.5	9.7	9.0	10.2	11.3	10.4	10.6	9.2



Naïve tracks the ups and downs best, but lags the data by one period. Exponential smoothing is probably better because it smoothens the data and does not have as much variation.

TEACHING NOTE: Notice how well exponential smoothing forecasts the naïve.

4.4 (a) $F_{\text{July}} = F_{\text{June}} + 0.2(\text{Forecasting error})$

$$= 42 + 0.2(40 - 42) = 41.6$$

(b) $F_{\text{August}} = F_{\text{July}} + 0.2(\text{Forecasting error})$

$$= 41.6 + 0.2(45 - 41.6) = 42.3$$

(c) Because the banking industry has a great deal of seasonality in its processing requirements

4.5 (a) $\dfrac{3,700 + 3,800}{2} = 3,750$ ml.

(b)

Year	Mileage	Two-Year Moving Average	Error	|Error|
1	3,000			
2	4,000			
3	3,400	3,500	–100	100
4	3,800	3,700	100	100
5	3,700	3,600	100	100
		Totals	100	300

$$MAD = \frac{300}{3} = 100.$$

4.5 (d)

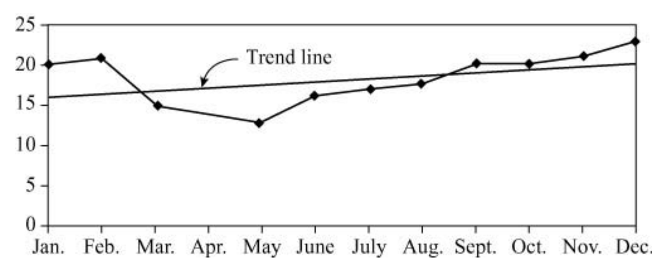
Year	Mileage	Forecast	Forecast Error	Error × α = .50	New Forecast
1	3,000	3,000	0	0	3,000
2	4,000	3,000	1,000	500	3,500
3	3,400	3,500	–100	–50	3,450
4	3,800	3,450	350	175	3,625
5	3,700	3,625	75	38	3,663
		Total	1,325		

The forecast is 3,663 miles.

4.6

	Y Sales	*X* Period	X^2	*XY*
January	20	1	1	20
February	21	2	4	42
March	15	3	9	45
April	14	4	16	56
May	13	5	25	65
June	16	6	36	96
July	17	7	49	119
August	18	8	64	144
September	20	9	81	180
October	20	10	100	200
November	21	11	121	231
December	23	12	144	276
Sum	218	78	650	1474
Average	18.2	6.5		

(a)



(b) • Naive The coming January = December = 23

• 3-month moving (20 + 21 + 23)/3 = 21.33

• 6-month weighted (0.1 × 17) + (.1 × 18) + (0.1 × 20) + (0.2 × 20) + (0.2 × 21) + (0.3 × 23) = 20.6

• Exponential smoothing with alpha = 0.3

$F_{Oct} = 18 + 0.3(20 - 18) = 18.6$

$F_{Nov} = 18.6 + 0.3(20 - 18.6) = 19.02$

$F_{Dec} = 19.02 + 0.3(21 - 19.02) = 19.6$

$F_{Jan} = 19.6 + 0.3(23 - 19.6) = 20.62 \approx 21$

• Trend $\sum x = 78,\ \bar{x} = 6.5,\ \sum y = 218,\ \bar{y} = 18.17$

$$b = \frac{1474 - (12)(6.5)(18.2)}{650 - 12(6.5)^2} = \frac{54.4}{143} = 0.38$$

$a = 18.2 - 0.38(6.5) = 15.73$

Forecast = 15.73 + .38(13) = 20.67, where next January is the 13th month.

(c) Only trend provides an equation that can extend beyond one month

4.7 Using MAD for this problem,

(1) Year	(2) Sales	(3) Marketing VP's Forecast	(4) Marketing VP's Error [(2)–(3)]	(5) Operations VP's Forecast	(6) Operations VP's Error [(2)–(5)]
1	167,325	170,000	2,675	160,000	7,325
2	175,362	170,000	5,362	165,000	10,362
3	172,536	180,000	7,464	170,000	2,536
4	156,732	180,000	23,268	175,000	18,268
5	176,325	165,000	11,325	165,000	11,325
			Totals 50,094		49,816

MAD (marketing VP) = 50,094/5 = 10,018.8.

MAD (operations VP) = 49,816/5 = 9,963.2.

Therefore, based on past data, the VP of operations has been presenting better forecasts.

4.8 (a) $\frac{(96 + 88 + 90)}{3} = 91.3$

(b) $\frac{(88 + 90)}{2} = 89$

(c)

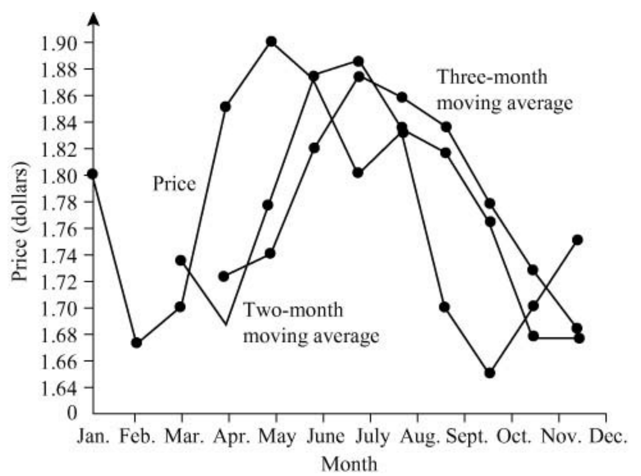
Temperature	2 day M.A.	|Error|	(Error)^2	Absolute % Error	
93	—	—	—	—	
94	—	—	—	—	
93	93.5	0.5	0.25	100(.5/93)	= 0.54%
95	93.5	1.5	2.25	100(1.5/95)	= 1.58%
96	94.0	2.0	4.00	100(2/96)	= 2.08%
88	95.5	7.5	56.25	100(7.5/88)	= 8.52%
90	92.0	2.0	4.00	100(2/90)	= 2.22%
		13.5	66.75		14.94%

MAD = 13.5/5 = 2.7

(d) MSE = 66.75/5 = 13.35

(e) MAPE = 14.94%/5 = 2.99%

4.9 (a, b) The computations for both the two- and three-month averages appear in the table; the results appear in the figure below.



(c) MAD (two-month moving average) = .750/10 = .075

MAD (three-month moving average) = .793/9 = .088

Therefore, the two-month moving average seems to have performed better.

Table for Problem 4.9 (a, b, c)

Month	Price per Chip	Forecast: Two-Month Moving Average	Forecast: Three-Month Moving Average	|Error|: Two-Month Moving Average	|Error|: Three-Month Moving Average
January	$1.80				
February	1.67				
March	1.70	1.735		.035	
April	1.85	1.685	1.723	.165	.127
May	1.90	1.775	1.740	.125	.160
June	1.87	1.875	1.817	.005	.053
July	1.80	1.885	1.873	.085	.073
August	1.83	1.835	1.857	.005	.027
September	1.70	1.815	1.833	.115	.133
October	1.65	1.765	1.777	.115	.127
November	1.70	1.675	1.727	.025	.027
December	1.75	1.675	1.683	.075	.067
			Totals	.750	.793

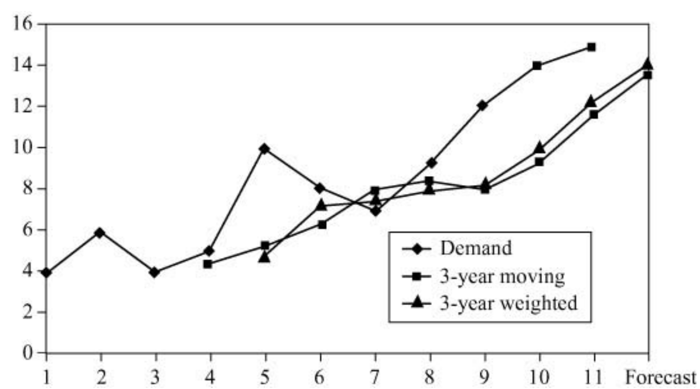
4.9 (d) Table for Problem 4.9(d).

		α = .1		α = .3		α = .5	
Month	**Price per Chip**	**Forecast**	**|Error|**	**Forecast**	**|Error|**	**Forecast**	**|Error|**
January	$1.80	$1.80	$.00	$1.80	$.00	$1.80	$.00
February	1.67	1.80	.13	1.80	.13	1.80	.13
March	1.70	1.79	.09	1.76	.06	1.74	.04
April	1.85	1.78	.07	1.74	.11	1.72	.13
May	1.90	1.79	.11	1.77	.13	1.78	.12
June	1.87	1.80	.07	1.81	.06	1.84	.03
July	1.80	1.80	.00	1.83	.03	1.86	.06
August	1.83	1.80	.03	1.82	.01	1.83	.00
September	1.70	1.81	.11	1.82	.12	1.83	.13
October	1.65	1.80	.15	1.79	.14	1.76	.11
November	1.70	1.78	.08	1.75	.05	1.71	.01
December	1.75	1.77	.02	1.73	.02	1.70	.05
	Totals		$.86		$.86		$.81
	MAD (total/12)		$.072		$.072		$.0675

α = .5 is preferable, using MAD, to α = .1 or α = .3. One could also justify excluding the January error and then dividing by n = 11 to compute the MAD. These numbers would be $.078 (for α = .1), $.078 (for α = .3), and $.074 (for α = .5)

4.10

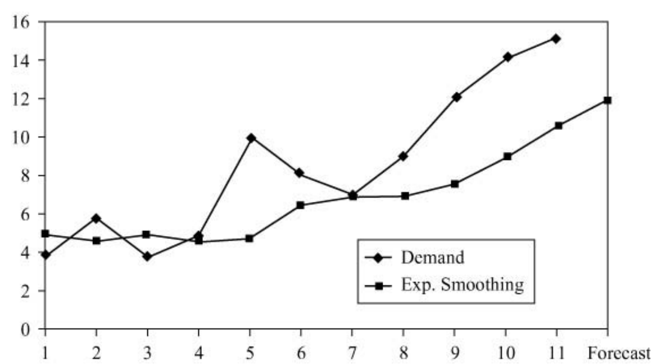
	Year	1	2	3	4	5	6	7	8	9	10	11	Forecast
	Demand	4	6	4	5.0	10.0	8.0	7.0	9.0	12.0	14.0	15.0	
(a)	3-year moving				4.7	5.0	6.3	7.7	8.3	8.0	9.3	11.7	13.7
(b)	3-year weighted				4.5	5.0	7.3	7.8	8.0	8.3	10.0	12.3	14.0



(c) The forecasts are about the same.

4.11

Year	1	2	3	4	5	6	7	8	9	10	11	Forecast
Demand	4	6.0	4.0	5.0	10.0	8.0	7.0	9.0	12.0	14.0	15.0	
Exp. Smoothing	5	4.7	5.1	4.8	4.8	6.4	6.9	6.9	7.5	8.9	10.4	11.8



4.12 |Error| = |Actual – Forecast|

Year	1	2	3	4	5	6	7	8	9	10	11	MAD
3-year moving				0.3	5.0	1.7	0.7	0.7	4.0	4.7	3.3	2.5
3-year weighted				0.5	5.0	0.8	0.8	1.0	3.8	4.0	2.8	2.3
Exp. smoothing	1	1.3	1.1	0.2	5.2	1.6	0.1	2.1	4.5	5.1	4.6	2.4

These calculations were completed in Excel. Calculations are slightly different in Excel OM and POM for Windows, due to rounding differences. The 3-year weighted average was slightly better.

4.13 (a) Exponential smoothing, $\alpha = 0.6$:

Year	Demand	Exponential Smoothing α = 0.6	Absolute Deviation
1	45	41	4.0
2	50	41.0 + 0.6(45–41) = 43.4	6.6
3	52	43.4 + 0.6(50–43.4) = 47.4	4.6
4	56	47.4 + 0.6(52–47.4) = 50.2	5.8
5	58	50.2 + 0.6(56–50.2) = 53.7	4.3
6	?	53.7 + 0.6(58–53.7) = 56.3	

$\Sigma = 25.3$
$MAD = 5.06$

Exponential smoothing, $\alpha = 0.9$:

Year	Demand	Exponential Smoothing α = 0.9	Absolute Deviation
1	45	41	4.0
2	50	41.0 + 0.9(45–41) = 44.6	5.4
3	52	44.6 + 0.9(50–44.6) = 49.5	2.5
4	56	49.5 + 0.9(52–49.5) = 51.8	4.2
5	58	51.8 + 0.9(56–51.8) = 55.6	2.4
6	?	55.6 + 0.9(58–55.6) = 57.8	

$\Sigma = 18.5$
$MAD = 3.7$

(b) 3-year moving average:

Year	Demand	Three-Year Moving Average	Absolute Deviation
1	45		
2	50		
3	52		
4	56	(45 + 50 + 52)/3 = 49	7
5	58	(50 + 52 + 56)/3 = 52.7	5.3
6	?	(52 + 56 + 58)/3 = 55.3	

$\Sigma = 12.3$
$MAD = 6.2$

(c) Trend projection:

Year	Demand	Trend Projection	Absolute Deviation
1	45	42.6 + 3.2 × 1 = 45.8	0.8
2	50	42.6 + 3.2 × 2 = 49.0	1.0
3	52	42.6 + 3.2 × 3 = 52.2	0.2
4	56	42.6 + 3.2 × 4 = 55.4	0.6
5	58	42.6 + 3.2 × 5 = 58.6	0.6
6	?	42.6 + 3.2 × 6 = 61.8	

$\Sigma = 3.2$
$MAD = 0.64$

$$Y = a + bX$$
$$b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}$$
$$a = \overline{Y} - b\overline{X}$$

X	Y	XY	X²
1	45	45	1
2	50	100	4
3	52	156	9
4	56	224	16
5	58	290	25

Then: $\Sigma X = 15$, $\Sigma Y = 261$, $\Sigma XY = 815$, $\Sigma X^2 = 55$, $\overline{X} = 3$, $\overline{Y} = 52.2$
Therefore

$$b = \frac{815 - 5 \times 3 \times 52.2}{55 - 5 \times 3 \times 3} = 3.2$$
$$a = 52.20 - 3.20 \times 3 = 42.6$$
$$Y_6 = 42.6 + 3.2 \times 6 = 61.8$$

4.14 Comparing the results of the forecasting methodologies for Problem 4.13:

Forecast Methodology	*MAD*
Exponential smoothing, α = 0.6	5.06
Exponential smoothing, α = 0.9	3.7
3-year moving average	6.2
Trend projection	0.64

Based on a mean absolute deviation criterion, the trend projection is to be preferred over the exponential smoothing with $\alpha = 0.6$, exponential smoothing with $\alpha = 0.9$, or the 3-year moving average forecast methodologies.

4.15

Year	Sales	Forecast Three Year Moving Average	Absolute Deviation
2001	450		
2002	495		
2003	518		
2004	563	(450 + 495 + 518)/3 = 487.7	75.3
2005	584	(495 + 518 + 563)/3 = 525.3	58.7
2006		(518 + 563 + 584)/3 = 555.0	

$\Sigma = 134$
$MAD = 67$

4.16

Year	Time Period X	Sales Y	X^2	XY
2001	1	450	1	450
2002	2	495	4	990
2003	3	518	9	1554
2004	4	563	16	2252
2005	5	584	25	2920
		$\Sigma = 2610$	$\Sigma = 55$	$\Sigma = 8166$

$\overline{X} = 3,\ \overline{Y} = 522$

$$Y = a + bX$$

$$b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2} = \frac{8166 - (5)(3)(522)}{55 - (5)(9)} = \frac{336}{10} = 33.6$$

$$a = \overline{Y} - bX = 22 - (33.6)(3) = 421.2$$

$$y = 421.2 + 33.6x$$

$$y = 421.2 + 33.6 \times 6 = 622.8$$

Year	Sales	Forecast Trend	Absolute Deviation
2001	450	454.8	4.8
2002	495	488.4	6.6
2003	518	522.0	4.0
2004	563	555.6	7.4
2005	584	589.2	5.2
2006		622.8	

$\Sigma = 28$
$MAD = 5.6$

4.17

Year	Sales	Forecast Exponential Smoothing $\alpha = 0.6$	Absolute Deviation
2001	450	410.0	40.0
2002	495	410 + 0.6(450 – 410) = 434.0	61.0
2003	518	434 + 0.6(495 – 434) = 470.6	47.4
2004	563	470.6 + 0.6(518 – 470.6) = 499.0	64.0
2005	584	499 + 0.6(563 – 499) = 537.4	46.6
2006		537.4 + 0.6(584 – 537.4) = 565.6	

$\Sigma = 259$
$MAD = 51.8$

Year	Sales	Forecast Exponential Smoothing $\alpha = 0.9$	Absolute Deviation
2001	450	410.0	40.0
2002	495	410 + 0.9(450 – 410) = 446.0	49.0
2003	518	446 + 0.9(495 – 446) = 490.1	27.9
2004	563	490.1 + 0.9(518 – 490.1) = 515.2	47.8
2005	584	515.2 + 0.9(563 – 515.2) = 558.2	25.8
2006		558.2 + 0.9(584 – 558.2) = 581.4	

$\Sigma = 190.5$
$MAD = 38.1$

(Refer to Solved Problem 4.1)
For $\alpha = 0.3$, absolute deviations for 2001–2005 are: 40.0, 73.0, 74.1, 96.9, 88.8, respectively. So the MAD = 372.8/5 = 74.6.

$$MAD^{\alpha=0.3} = 74.6$$

$$MAD^{\alpha=0.6} = 51.8$$

$$MAD^{\alpha=0.9} = 38.1$$

Because it gives the lowest MAD, the smoothing constant of $\alpha = 0.9$ gives the most accurate forecast.

4.18

$$MAD^{\alpha=0.3} = 74.6$$

$$MAD^{3\text{-year moving average}} = 67$$

$$MAD^{\text{trend}} = 5.6$$

One would use the trend (regression) forecast because it has the lowest MAD.

4.19 Trend adjusted exponential smoothing: $\alpha = 0.1$, $\beta = 0.2$

Month	Income	Unadjusted Forecast	Trend	Adjusted Forecast	|Error|	Error^2
February	70.0	65.0	0.0	65	5.0	25.0
March	68.5	65.5	0.1	65.6	2.9	8.4
April	64.8	65.9	0.16	66.05	1.2	1.6
May	71.7	65.92	0.13	66.06	5.6	31.9
June	71.3	66.62	0.25	66.87	4.4	19.7
July	72.8	67.31	0.33	67.64	5.2	26.6
August		68.16		68.60	24.3	113.2

$MAD = 24.3/6 = 4.05$, MSE = 113.2/6 = 18.87: note all numbers are rounded

Note: To use POM for Windows to solve this problem, a period 0, which contains the initial forecast and initial trend, must be added.

4.20 Trend adjusted exponential smoothing: α = 0.1, β = 0.8

Month	Demand (y)	Unadjusted Forecast	Trend	Adjusted Forecast	Error	|Error|	Error²
February	70.0	65.0	0	65.0	5.00	5.0	25.00
March	68.5	65.5	0.4	65.9	2.60	2.6	6.76
April	64.8	66.16	0.61	66.77	–1.97	1.97	3.87
May	71.7	66.57	0.45	67.02	4.68	4.68	21.89
June	71.3	67.49	0.82	68.31	2.99	2.99	8.91
July	72.8	68.61	1.06	69.68	3.12	3.12	9.76
Totals	419.1				16.42	20.36	76.19
Average	69.85				2.74	3.39	12.70
August Forecast				71.30	(Bias)	(MAD)	(MSE)

Note: To use POM for Windows to solve this problem, a period 0, which contains the initial forecast and initial trend, must be added.

Based upon the MSE criterion, the exponential smoothing with α = 0.1, β = 0.8 is to be preferred over the exponential smoothing with α = 0.1, β = 0.2. Its MSE of 12.70 is lower. Its MAD of 3.39 is also lower than that in Problem 4.19.

4.21
$$F_5 = \alpha A_4 + (1-\alpha)(F_4 + T_4) = (0.2)(19) + (0.8)(20.14)$$
$$= 3.8 + 16.11 = 19.91$$
$$T_5 = \beta(F_5 - F_4) + (1-\beta)T_4 = (0.4)(19.91 - 17.82)$$
$$+ (0.6)(2.32) = 0.4(2.09)$$
$$+ 1.39 = 0.84 + 1.39 = 2.23$$
$$FIT_5 = F_5 + T_5 = 19.91 + 2.23 = 22.14$$

$$F_6 = \alpha A_5 + (1-\alpha)(F_5 + T_5) = (0.2)(24) + (0.8)(22.14)$$
$$= 4.8 + 17.71 = 22.51$$
$$T_6 = \beta(F_6 - F_5) + (1-\beta)T_5 = 0.4(22.51 - 19.91) + 0.6(2.23)$$
$$= 0.4(2.6) + 1.34$$
$$= 1.04 + 1.34 = 2.38$$
$$FIT_6 = F_6 + T_6 = 22.51 + 2.38 = 24.89$$

4.22
$$F_7 = \alpha A_6 + (1-\alpha)(F_6 + T_6) = (0.2)(21) + (0.8)(24.89)$$
$$= 4.2 + 19.91 = 24.11$$
$$T_7 = \beta(F_7 - F_6) + (1-\beta)T_6 = (0.4)(24.11 - 22.51)$$
$$+ (0.6)(2.38) = 2.07$$
$$FIT_7 = F_7 + T_7 = 24.11 + 2.07 = 26.18$$

$$F_8 = \alpha A_7 + (1-\alpha)(F_7 + T_7) = (0.2)(31)$$
$$+ (0.8)(26.18) = 27.14$$
$$T_8 = \beta(F_8 - F_7) + (1-\beta)T_7 = 0.4(27.14 - 24.11)$$
$$+ 0.6(2.07) = 2.45$$
$$FIT_8 = F_8 + T_8 = 27.14 + 2.45 = 29.59$$

$$F_9 = \alpha A_8 + (1-\alpha)(F_8 + T_8) = (0.2)(28)$$
$$+ (0.8)(29.59) = 29.28$$
$$T_9 = \beta(F_9 - F_8) + (1-\beta)T_8 = (0.4)(29.28 - 27.14)$$
$$+ (0.6)(2.45) = 2.32$$
$$FIT_9 = F_9 + T_9 = 29.28 + 2.32 = 31.60$$

4.23 Students must determine the naive forecast for the four months. The naive forecast for March is the February actual of 83, etc.

(a)

	Actual	Forecast	|Error|	|% Error|
March	101	120	19	100 (19/101) = 18.81%
April	96	114	18	100 (18/96) = 18.75%
May	89	110	21	100 (21/89) = 23.60%
June	108	108	0	100 (0/108) = 0%
			58	61.16%

$$\text{E.S. MAD (for manager)} = \frac{58}{4} = 14.5$$

$$\text{MAPE (for manager)} = \frac{61.16\%}{4} = 15.29\%$$

(b)

	Actual	Naive	|Error|	|% Error|
March	101	83	18	100 (18/101) = 17.82%
April	96	101	5	100 (5/96) = 5.21%
May	89	96	7	100 (7/89) = 7.87%
June	108	89	19	100 (19/108) = 17.59%
			49	48.49%

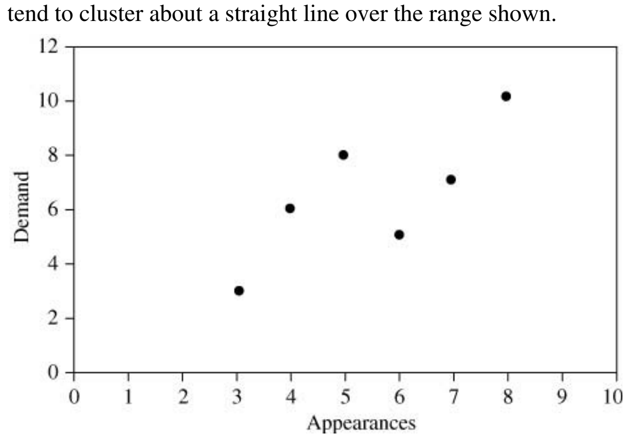
$$\text{MAD (for naive)} = \frac{49}{4} = 12.25$$

$$\text{MAPE (for naive)} = \frac{48.49\%}{4} = 12.12\%.$$

(c) MAD for the manager's technique is 14.5, while MAD for the naive forecast is only 12.25. MAPEs are 15.29% and 12.12%, respectively. So the naive method is better.

4.24 (a) Graph of Demand

The observations obviously do not form a straight line, but do tend to cluster about a straight line over the range shown.



(b) Least Squares Regression:

$$Y = a + bX$$

$$b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}$$

$$a = \overline{Y} - b\overline{X}$$

Assume

Appearances *X*	Demand *Y*	X^2	Y^2	*XY*
3	3	9	9	9
4	6	16	36	24
7	7	49	49	49
6	5	36	25	30
8	10	64	100	80
5	8	25	64	40
9	?			

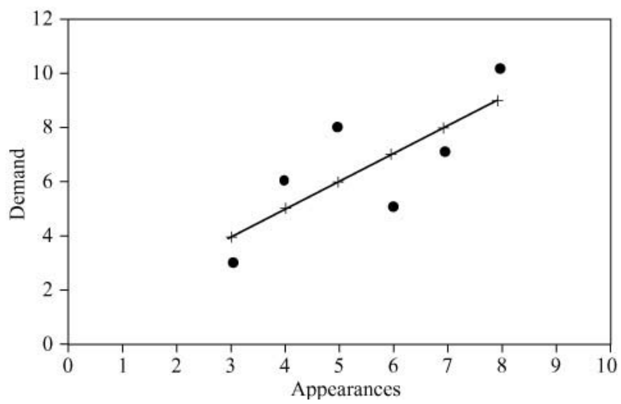
$\Sigma X = 33$, $\Sigma Y = 39$, $\Sigma XY = 232$, $\Sigma X^2 = 199$, $\overline{X} = 5.5$, $\overline{Y} = 6.5$. Therefore

$$b = \frac{232 - 6 \times 5.5 \times 6.5}{199 - 6 \times 5.5 \times 5.5} = 1$$

$$a = 6.5 - 1 \times 5.5 = 1$$

$$Y = 1 + 1X$$

The following figure shows both the data and the resulting equation:



If there are nine performances by Green Shades, the estimated sales are:

$$Y_9 = 1 + 1 \times 9 = 1 + 9 = 10 \text{ drums}$$

4.25

Month	Number of Accidents (*y*)	*x*	*xy*	x^2
January	30	1	30	1
February	40	2	80	4
March	60	3	180	9
April	90	4	360	16
Totals	220	10	650	30
Averages	$\overline{y} = 55$	$\overline{x} = 2.5$		

$$b = \frac{\sum xy - n\overline{x}\,\overline{y}}{\sum x^2 - n\overline{x}^2} = \frac{650 - 4(2.5)(55)}{30 - 4(2.5)^2} = \frac{650 - 550}{30 - 25}$$

$$= \frac{100}{5} = 20$$

$$a = \overline{y} - b\overline{x}$$

$$= 55 - (20)(2.5)$$

$$= 5$$

The regression line is $y = 5 + 20x$. The forecast for May ($x = 5$) is $y = 5 + 20(5) = 105$.

4.26

Season	Year_1 Demand	Year_2 Demand	Average Year_1–Year_2 Demand	Average Season Demand	Seasonal Index	Year_3 Demand
Fall	200	250	225.0	250	0.90	270
Winter	350	300	325.0	250	1.30	390
Spring	150	165	157.5	250	0.63	189
Summer	300	285	292.5	250	1.17	351

$$\left[\begin{array}{l}\text{Average } Yr_1 \text{ to } Yr_2 \\ \text{Demand for season}\end{array}\right] = \frac{Yr_1 \text{ Demand} + Yr_2 \text{ Demand}}{2}$$

$$\text{Average seasonal demand} = \frac{\text{Sum of Ave } Yr_1 \text{ to } Yr_2 \text{ Demand}}{4}$$

$$\text{Seasonal index} = \frac{\text{Average } Yr_1 \text{ to } Yr_2 \text{ Demand}}{\text{Average Seasonal Demand}}$$

$$Yr_3 = \frac{\text{New Annual Demand}}{4} \times \text{Seasonal Index}$$

$$= \frac{1200}{4} \times \text{Seasonal Index}$$

4.27

Day of Week	Day Average	Day Relative Index
Monday	84.75	0.903 = 84.75/93.86
Tuesday	74.25	0.791 = 74.25/93.86
Wednesday	87.00	0.927 = 87.00/93.86
Thursday	97.00	1.033 = 97.00/93.86
Friday	133.50	1.422 = 133.50/93.86
Saturday	138.75	1.478 = 138.75/93.86
Sunday	41.75	0.445 = 41.75/93.86
Average daily sales	93.86	

4.28

Quarter	2003	2004	2005	Average Demand	Average Quarterly Demand	Seasonal Index
Winter	73	65	89	75.67	106.67	0.709
Spring	104	82	146	110.67	106.67	1.037
Summer	168	124	205	165.67	106.67	1.553
Fall	74	52	98	74.67	106.67	0.700

4.29 2007 is 25 years beyond 1982. Therefore, the quarter numbers are 101 through 104.

(1) Quarter	(2) Quarter Number	(3) Forecast (77 + .43Q)	(4) Seasonal Factor	(5) Adjusted Forecast [(3) × (4)]
Winter	101	120.43	.8	96.344
Spring	102	120.86	1.1	132.946
Summer	103	121.29	1.4	169.806
Fall	104	121.72	.7	85.204

4.30 Given $Y = 36 + 4.3X$

(a) $Y = 36 + 4.3(70) = 337$

(b) $Y = 36 + 4.3(80) = 380$

(c) $Y = 36 + 4.3(90) = 423$

4.31 (a)

Year	Season	Sales (y)	(x)	(xy)	x^2
1	SS	26,825	1	26,825	1
	FW	5,722	2	11,444	4
2	SS	28,630	3	85,890	9
	FW	7,633	4	30,532	16
3	SS	30,255	5	151,275	25
	FW	8,745	6	52,470	36
Totals		107,810	21	358,436	91
		$\bar{y}$ = 17,968.33	$\bar{x}$ = 3.5		

$$b = \frac{358{,}436 - 6(3.5)(17{,}968.33)}{91 - 6(3.5)^2}$$

$$= \frac{358{,}436 - 377{,}335}{91 - 73.5} = \frac{-18{,}899}{17.5} = -1{,}080$$

$$a = 17{,}968.33 - 3.5(-1{,}080)$$

$$= 17{,}968.33 + 3{,}780 = 21{,}748.33$$

$$y = 21.748 - 1{,}080x$$

(b) The problem with this line is that it shows a decreasing trend when sales have been *rising* each year.

(c) Two separate forecast lines should be generated—one for Spring/Summer and one for Fall/Winter—or the analysis can be performed as a multiple regression.

4.32 (a)

x	y	xy	x^2
16	330	5,280	256
12	270	3,240	144
18	380	6,840	324
14	300	4,200	196
60	1,280	19,560	920

$$\bar{x} = \frac{60}{4} = 15$$

$$\bar{y} = \frac{1{,}280}{4} = 320$$

$$b = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2} = \frac{19{,}560 - 4(15)(320)}{920 - 4(15)^2} = \frac{360}{20} = 18$$

$$a = \bar{y} - b\bar{x} = 320 - 18(15) = 50$$

$$Y = 50 + 18x$$

(b) If the forecast is for 20 guests, the bar sales forecast is 50 + 18(20) = $410. Each guest accounts for an additional $18 in bar sales.

4.33 (a) See the table below.

$$b = \frac{2{,}880 - 5(3)(180)}{55 - 5(3)^2} = \frac{2{,}880 - 2{,}700}{55 - 45}$$

$$= \frac{180}{10} = 18$$

$$a = 180 - 3(18) = 180 - 54 = 126$$

$$y = 126 + 18x$$

For next year (x = 6), the number of transistors (in millions) is forecasted as y = 126 + 18(6) = 126 + 108 = 234.

Table for Problem 4.33

	Year (x)	Transistors (y)	xy	x^2	126 + 18x	Error	Error^2	|% Error|
	1	140	140	1	144	–4	16	100 (4/140) = 2.86%
	2	160	320	4	162	–2	4	100 (2/160) = 1.25%
	3	190	570	9	180	10	100	100 (10/190) = 5.26%
	4	200	800	16	198	2	4	100 (2/200) = 1.00%
	5	210	1,050	25	216	–6	36	100 (6/210) = 2.86%
Totals	15	900	2,800	55			160	13.23%
	$\bar{x}$ = 3	$\bar{y}$ = 180						

(b) $MSE = 160/5 = 32$

(c) $MAPE = 13.23\%/4 = 2.65\%$

4.34 $Y = 7.5 + 3.5X_1 + 4.5X_2 + 2.5X_3$

(a) 28

(b) 43

(c) 58

4.35 (a) $\hat{Y}$ = 13,473 + 37.65(1860) = 83,502

(b) The predicted selling price is $83,502, but this is the average price for a house of this size. There are other factors besides square footage that will impact the selling price of a house. If such a house sold for $95,000, then these other factors could be contributing to the additional value.

(c) Some other quantitative variables would be age of the house, number of bedrooms, size of the lot, and size of the garage, etc.

(d) Coefficient of determination = $(0.63)^2 = 0.397$. This means that only about 39.7% of the variability in the sales price of a house is explained by this regression model that only includes square footage as the explanatory variable.

4.36 (a) Given: $Y = 90 + 48.5X_1 + 0.4X_2$ where:

Y = expected travel cost

X_1 = number of days on the road

X_2 = distance traveled, in miles

$r = 0.68$ (coefficient of correlation)

If:

Number of days on the road $\rightarrow X_1 = 5$ and distance traveled $\rightarrow X_2 = 300$

then:

$$Y = 90 + 48.5 \times 5 + 0.4 \times 300 = 90 + 242.5 + 120 = 452.5$$

Therefore, the expected cost of the trip is $452.50.

(b) The reimbursement request is much higher than predicted by the model. This request should probably be questioned by the accountant.

(c) A number of other variables should be included, such as:

1. the type of travel (air or car)
2. conference fees, if any
3. costs of entertaining customers
4. other transportation costs—cab, limousine, special tolls, or parking

In addition, the correlation coefficient of 0.68 is not exceptionally high. It indicates that the model explains approximately 46% of the overall variation in trip cost. This correlation coefficient would suggest that the model is not a particularly good one.

4.37 (a), (b)

Period	Demand	Forecast	Error	Running sum	|error|
1	20	20	0.00	0.00	0.00
2	21	21.5	–0.50	–0.50	0.50
3	28	21.25	6.75	6.25	6.75
4	37	24.63	12.38	18.63	12.38
5	25	30.81	–5.81	12.82	5.81
6	29	27.91	1.09	13.91	1.09
7	36	28.45	7.55	21.46	7.55
8	22	32.23	–10.23	11.23	10.23
9	25	27.11	–2.11	9.12	2.11
10	28	26.06	1.94	11.06	1.94
					MAD = 4.84

RSFE = 11.06; MAD = 4.84 Tracking = 11.06/4.84 = 2.29

4.38 (a) least squares equation: $Y = -0.158 + 0.1308X$

(b) $Y = -0.158 + 0.1308(22) = 2.719$

(c) coefficient of correlation = $r = 0.966$

coefficient of determination = $r^2 = 0.934$

4.39

Year X	Patients Y	X^2	Y^2	XY
1	36	1	1296	36
2	33	4	1089	66
3	40	9	1600	120
4	41	16	1681	164
5	40	25	1600	200
6	55	36	3025	330
7	60	49	3600	420
8	54	64	2916	432
9	58	81	3364	522
10	61	100	3721	610
55	478	385	23892	2900

Given: $Y = a + bX$ where:

$$b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}$$

$$a = \overline{Y} - b\overline{X}$$

and $\Sigma X = 55$, $\Sigma Y = 478$, $\Sigma XY = 2900$, $\Sigma X^2 = 385$, $\Sigma Y^2 = 23892$, $\overline{X} = 5.5$, $\overline{Y} = 47.8$, Then:

$$b = \frac{2900 - 10 \times 5.5 \times 47.8}{385 - 10 \times 5.5^2} = \frac{2900 - 2629}{385 - 302.5} = \frac{271}{82.5} = 3.28$$

$$a = 47.8 - 3.28 \times 5.5 = 29.76$$

and $Y = 29.76 + 3.28X$. For:

$$X = 11: \; Y = 29.76 + 3.28 \times 11 = 65.8$$

$$X = 12: \; Y = 29.76 + 3.28 \times 12 = 69.1$$

Therefore:

Year 11 → 65.8 patients

Year 12 → 69.1 patients

The model "seems" to fit the data pretty well. One should, however, be more precise in judging the adequacy of the model. Two possible approaches are computation of (a) the correlation coefficient, or (b) the mean absolute deviation. The correlation coefficient:

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - \left(\sum X\right)^2\right]\left[n\sum Y^2 - \left(\sum Y\right)^2\right]}}$$

$$= \frac{10 \times 2900 - 55 \times 478}{\sqrt{\left[10 \times 385 - 55^2\right]\left[10 \times 23892 - 478^2\right]}}$$

$$= \frac{29000 - 26290}{\sqrt{\left[3850 - 3025\right]\left[238920 - 228484\right]}}$$

$$= \frac{2710}{\sqrt{825 \times 10436}} = \frac{2710}{2934.3} = 0.924$$

$$r^2 = 0.853$$

The coefficient of determination of 0.853 is quite respectable—indicating our original judgment of a "good" fit was appropriate.

Year X	Patients Y	Trend Forecast	Deviation	Absolute Deviation
1	36	29.8 + 3.28 × 1 = 33.1	2.9	2.9
2	33	29.8 + 3.28 × 2 = 36.3	–3.3	3.3
3	40	29.8 + 3.28 × 3 = 39.6	0.4	0.4
4	41	29.8 + 3.28 × 4 = 42.9	–1.9	1.9
5	40	29.8 + 3.28 × 5 = 46.2	–6.2	6.2
6	55	29.8 + 3.28 × 6 = 49.4	5.6	5.6
7	60	29.8 + 3.28 × 7 = 52.7	7.3	7.3
8	54	29.8 + 3.28 × 8 = 56.1	–2.1	2.1
9	58	29.8 + 3.28 × 9 = 59.3	–1.3	1.3
10	61	29.8 + 3.28 × 10 = 62.6	–1.6	1.6
				Σ = 32.6
				MAD = 3.26

The *MAD* is 3.26—this is approximately 7% of the average number of patients and 10% of the minimum number of patients. We also see absolute deviations, for years 5, 6, and 7 in the range 5.6–7.3. The comparison of the *MAD* with the average and minimum number of patients and the comparatively large deviations during the middle years indicate that the forecast model is not exceptionally accurate. It is more useful for predicting general trends than the actual number of patients to be seen in a specific year.

4.40

Year	Crime Rate X	Patients Y	X^2	Y^2	XY
1	58.3	36	3398.9	1296	2098.8
2	61.1	33	3733.2	1089	2016.3
3	73.4	40	5387.6	1600	2936.0
4	75.7	41	5730.5	1681	3103.7
5	81.1	40	6577.2	1600	3244.0
6	89.0	55	7921.0	3025	4895.0
7	101.1	60	10221.2	3600	6066.0
8	94.8	54	8987.0	2916	5119.2
9	103.3	58	10670.9	3364	5991.4
10	116.2	61	13502.4	3721	7088.2
Column Totals	854.0	478	76129.9	23892	42558.6

Given: $Y = a + bX$ where

$$b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}$$
$$a = \overline{Y} - b\overline{X}$$

and $\Sigma X = 854$, $\Sigma Y = 478$, $\Sigma XY = 42558.6$, $\Sigma X^2 = 76129.9$, $\Sigma Y^2 = 23892$, $\overline{X} = 85.4$, $\overline{Y} = 47.8$. Then:

$$b = \frac{42558.6 - 10 \times 85.4 \times 47.8}{76129.9 - 10 \times 85.4^2} = \frac{42558.6 - 40821.2}{76129.9 - 72931.6}$$
$$= \frac{1737.4}{3197.3} = 0.543$$
$$a = 47.8 - 0.543 \times 85.4 = 1.43$$

and $Y = 1.43 + 0.543X$

For:

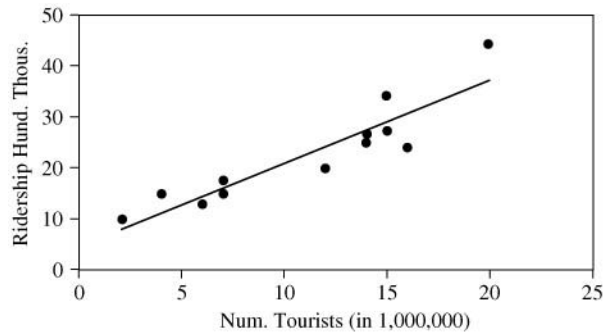
$$X = 131.2: \; Y = 1.43 + 0.543(131.2) = 72.7$$
$$X = 90.6: \; Y = 1.43 + 0.543(90.6) = 50.6$$

Therefore:

Crime rate = 131.2 → 72.7 patients
Crime rate = 90.6 → 50.6 patients

Note that rounding differences occur when solving with Excel.

4.41 (a) It appears from the following graph that the points do scatter around a straight line.



(b) Developing the regression relationship, we have:

(Summer months) Year	Tourists (Millions) (X)	Ridership (1,000,000s) (Y)	X^2	Y^2	XY
1	7	1.5	49	2.25	10.5
2	2	1.0	4	1.00	2.0
3	6	1.3	36	1.69	7.8
4	4	1.5	16	2.25	6.0
5	14	2.5	196	6.25	35.0
6	15	2.7	225	7.29	40.5
7	16	2.4	256	5.76	38.4
8	12	2.0	144	4.00	24.0
9	14	2.7	196	7.29	37.8
10	20	4.4	400	19.36	88.0
11	15	3.4	225	11.56	51.0
12	7	1.7	49	2.89	11.9

Given: $Y = a + bX$ where:

$$b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}$$
$$a = \overline{Y} - b\overline{X}$$

and $\Sigma X = 132$, $\Sigma Y = 27.1$, $\Sigma XY = 352.9$, $\Sigma X^2 = 1796$, $\Sigma Y^2 = 71.59$, $\overline{X} = 11$, $\overline{Y} = 2.26$. Then:

Then:

$$b = \frac{352.9 - 12 \times 11 \times 2.26}{1796 - 12 \times 11^2} = \frac{352.9 - 298.3}{1796 - 1452} = \frac{54.6}{344} = 0.159$$
$$a = 2.26 - 0.159 \times 11 = 0.511$$

and $Y = 0.511 + 0.159X$

(c) Given a tourist population of 10,000,000, the model predicts a ridership of:

$Y = 0.511 + 0.159X \times 10 = 2.101$ or 2,101,000 persons.

(d) If there are no tourists at all, the model predicts a ridership of 0.511, or 511,000 persons. One would not place much confidence in this forecast, however, because the number of tourists is outside the range of data used to develop the model.

(e) The standard error of the estimate is given by:

$$S_{yx} = \sqrt{\frac{\sum Y^2 - a\sum Y - b\sum XY}{n - 2}}$$
$$= \sqrt{\frac{71.59 - 0.511 \times 27.1 - 0.159 \times 352.9}{12 - 2}}$$
$$= \sqrt{\frac{71.59 - 13.85 - 56.11}{10}} = \sqrt{.163}$$

= .404 (rounded to .407 in POM for Windows software)

(f) The correlation coefficient and the coefficient of determination are given by:

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - \left(\sum X\right)^2\right]\left[n\sum Y^2 - \left(\sum Y\right)^2\right]}}$$

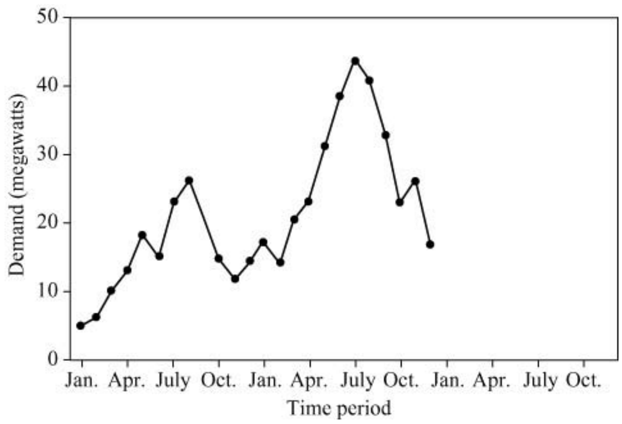
$$= \frac{12 \times 352.9 - 132 \times 27.1}{\sqrt{\left[12 \times 1796 - 132^2\right]\left[12 \times 71.59 - 27.1^2\right]}}$$

$$= \frac{4234.8 - 3577.2}{\sqrt{\left[21552 - 17424\right]\left[859.08 - 734.41\right]}}$$

$$= \frac{657.6}{\sqrt{4128 \times 124.67}} = \frac{657.6}{64.25 \times 11.166} = 0.917$$

and $r^2 = 0.840$

4.42 (a) This problem gives students a chance to tackle a realistic problem in business, i.e., not enough data to make a good forecast. As can be seen in the accompanying figure, the data contains both seasonal and trend factors.



Averaging methods are not appropriate with trend, seasonal, or other patterns in the data. Moving averages smooth out seasonality. Exponential smoothing can forecast January next year, but not further. Because seasonality is strong, a naïve model that students create on their own might be best.

One model might be: $F_{t+1} = A_{t-11}$

That is $\text{forecast}_{\text{next period}} = \text{actual}_{\text{one year earlier}}$ to account for seasonality. But this ignores the trend.

One very good approach would be to calculate the increase from each month last year to each month this year, sum all 12 increases, and divide by 12. The forecast for next year would equal the value for the same month this year plus the average increase over the 12 months of last year.

Using this model, the January forecast for next year becomes:

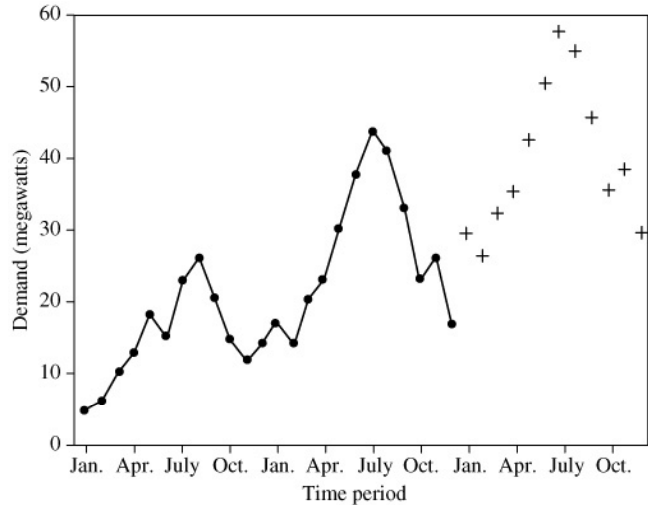
$$F_{25} = 17 + \frac{148}{12} = 17 + 12 = 29$$

where 148 = total increases from last year to this year.

The forecasts for each of the months of next year then become:

Jan	29	July	56
Feb	26	Aug	53
Mar	32	Sep	45
Apr	35	Oct	35
May	42	Nov	38
Jun	50	Dec	29

Both history and forecast for the next year are shown in the accompanying figure:



4.43 (a) and (b) See the following table.

Week t	Actual Value $A(t)$	Smoothed Value $F_t(\alpha = 0.2)$	Forecast Error	Smoothed Value $F_t(\alpha = 0.6)$	Forecast Error
1	50	+50.0	+0.0	+50.0	+0.0
2	35	+50.0	−15.0	+50.0	−15.0
3	25	+47.0	−22.0	+41.0	−16.0
4	40	+42.6	−2.6	+31.4	+8.6
5	45	+42.1	−2.9	+36.6	+8.4
6	35	+42.7	−7.7	+41.6	−6.6
7	20	+41.1	−21.1	+37.6	−17.6
8	30	+36.9	−6.9	+27.1	+2.9
9	35	+35.5	−0.5	+28.8	+6.2
10	20	+35.4	−15.4	+32.5	−12.5
11	15	+32.3	−17.3	+25.0	−10.0
12	40	+28.9	+11.1	+19.0	+21.0
13	55	+31.1	+23.9	+31.6	+23.4
14	35	+35.9	−0.9	+45.6	−10.6
15	25	+36.7	−10.7	+39.3	−14.3
16	55	+33.6	+21.4	+30.7	+24.3
17	55	+37.8	+17.2	+45.3	+9.7
18	40	+41.3	−1.3	+51.1	−11.1
19	35	+41.0	−6.0	+44.4	−9.4
20	60	+39.8	+20.2	+38.8	+21.2
21	75	+43.9	+31.1	+51.5	+23.5
22	50	+50.1	−0.1	+65.6	−15.6
23	40	+50.1	−10.1	+56.2	−16.2
24	65	+48.1	+16.9	+46.5	+18.5
25		+51.4		+57.6	
			MAD = 11.8		*MAD* = 13.45

(c) Students should note how stable the smoothed values are for $\alpha = 0.2$. When compared to actual week 25 calls of 85, the smoothing constant, $\alpha = 0.6$, appears to do a slightly better job. On the basis of the standard error of the estimate and the MAD, the 0.2 constant is better. However, other smoothing constants need to be examined.

4.44

Week t	Actual Value A_t	Smoothed Value $F_t(\alpha = 0.3)$	Trend Estimate $T_t(\beta = 0.2)$	Forecast FIT_t	Forecast Error
1	50.000	50.000	0.000	50.000	0.000
2	35.000	50.000	0.000	50.000	–15.000
3	25.000	45.500	–0.900	44.600	–19.600
4	40.000	38.720	–2.076	36.644	3.356
5	45.000	37.651	–1.875	35.776	9.224
6	35.000	38.543	–1.321	37.222	–2.222
7	20.000	36.555	–1.455	35.101	–15.101
8	30.000	30.571	–2.361	28.210	1.790
9	35.000	28.747	–2.253	26.494	8.506
10	20.000	29.046	–1.743	27.303	–7.303
11	15.000	25.112	–2.181	22.931	–7.931
12	40.000	20.552	–2.657	17.895	22.105
13	55.000	24.526	–1.331	23.196	31.804
14	35.000	32.737	0.578	33.315	1.685
15	25.000	33.820	0.679	34.499	–9.499
16	55.000	31.649	0.109	31.758	23.242
17	55.000	38.731	1.503	40.234	14.766
18	40.000	44.664	2.389	47.053	–7.053
19	35.000	44.937	1.966	46.903	–11.903
20	60.000	43.332	1.252	44.584	15.416
21	75.000	49.209	2.177	51.386	23.614
22	50.000	58.470	3.594	62.064	–12.064
23	40.000	58.445	2.870	61.315	–21.315
24	65.000	54.920	1.591	56.511	8.489
25		59.058	2.100	61.158	

Note: To use POM for Windows to solve this problem, a period 0, which contains the initial fore cast and initial trend, must be added.

To evaluate the trend adjusted exponential smoothing model, actual week 25 calls are compared to the forecasted value. The model appears to be producing a forecast approximately mid-range between that given by simple exponential smoothing using $\alpha = 0.2$ and $\alpha = 0.6$. Trend adjustment does not appear to give any significant improvement.

4.45 We begin by reordering the numbers in the table to account for the fact that enrollment lags birth by 5 years. Notice that the table in the problem contains some extraneous information.

Year	Births (x)	Enrollment 5 Years Later (y)	xy	x^2
1	131	148	19,388	17,161
2	192	188	36,098	36,864
3	158	155	24,490	24,964
4	93	110	10,230	8,649
5	107	124	13,268	11,339
Totals	681	725	103,472	99,087
	$\bar{x} = 136.2$	$\bar{y} = 145$		

$$b = \frac{103{,}472 - 5(136.2)(145)}{99{,}087 - 5(136.2)^2}$$

$$= \frac{103{,}472 - 98{,}745}{99{,}087 - 92{,}752.2} = \frac{4{,}727}{6{,}334.8} = .746$$

$$a = 145 - .746(136.2) = 145 - 101.6052 = 43.3948$$

We now can use this equation for the next 2 years.

Year	Births 5 Years Earlier	Projected Enrollment ($43.3948 + .746x$)
11	130	140.3748
12	128	138.8828

4.46

	X	Y	X^2	Y^2	XY
	421	2.90	177241	8.41	1220.9
	377	2.93	142129	8.58	1104.6
	585	3.00	342225	9.00	1755.0
	690	3.45	476100	11.90	2380.5
	608	3.66	369664	13.40	2225.3
	390	2.88	152100	8.29	1123.2
	415	2.15	172225	4.62	892.3
	481	2.53	231361	6.40	1216.9
	729	3.22	531441	10.37	2347.4
	501	1.99	251001	3.96	997.0
	613	2.75	375769	7.56	1685.8
	709	3.90	502681	15.21	2765.1
	366	1.60	133956	2.56	585.6
Column totals	6885	36.96	3857893	110.26	20299.5

Given: $Y = a + bX$ where:

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$

$$a = \bar{Y} - b\bar{X}$$

and ΣX = 6885, ΣX = 36.96, ΣXY = 20299.5, ΣX^2 = 3857893, ΣY^2 = 110.26, $\overline{X}$ = 529.6, $\overline{Y}$ = 2.843, Then:

$$b = \frac{20299.5 - 13 \times 529.6 \times 2.843}{3857893 - 13 \times 529.6^2} = \frac{20299.5 - 19573.5}{3857893 - 3646190}$$
$$= \frac{726}{211703} = 0.0034$$
$$a = 2.84 - 0.0034 \times 529.6 = 1.03$$

and $Y = 1.03 + 0.0034X$

As an indication of the usefulness of this relationship, we can calculate the correlation coefficient:

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - \left(\sum X\right)^2\right]\left[n\sum Y^2 - \left(\sum Y\right)^2\right]}}$$
$$= \frac{13 \times 20299.5 - 6885 \times 36.96}{\sqrt{\left[13 \times 3857893 - 6885^2\right]\left[13 \times 110.26 - 36.96^2\right]}}$$
$$= \frac{263893.5 - 254469.6}{\sqrt{\left[50152609 - 47403225\right]\left[1433.4 - 1366.0\right]}}$$
$$= \frac{9423.9}{\sqrt{2749384 \times 67.0}}$$
$$= \frac{9423.9}{1658.13 \times 8.21} = 0.692$$
$$r^2 = 0.479$$

A correlation coefficient of 0.692 is not particularly high. The coefficient of determination, r^2, indicates that the model explains only 47.9% of the overall variation. Therefore, while the model does provide an estimate of GPA, there is considerable variation in GPA, which is as yet unexplained. For

$$X = 350: \; Y = 1.03 + 0.0034 \times 350 = 2.22$$
$$X = 800: \; Y = 1.03 + 0.0034 \times 800 = 3.75$$

Note: When solving this problem, care must be taken to interpret significant digits.

4.47 (a) There is *not* a strong linear trend in sales over time.

(b,c) Amit wants to forecast by exponential smoothing (setting February's forecast equal to January's sales) with alpha = 0.1 Barbara wants to use a 3-period moving average

	Sales	Amit	Barbara	Amit error	Barbara error
January	400	—	—	—	—
February	380	400	—	20.0	—
March	410	398	—	12.0	—
April	375	399.2	396.67	24.2	21.67
May	405	396.8	388.33	8.22	16.67
			MAD =	16.11	19.17

(d) Note that Amit has more forecast observations, while Barbara's moving average does not start until month 4. Also note that the MAD for Amit is an average of 4 numbers, while Barbara's is only 2.

Amit's MAD for exponential smoothing (16.11) is lower than that of Barbara's moving average (19.17). So his forecast seems to be better.

4.48 (a)

Quarter	Contracts *X*	Sales *Y*	X^2	Y^2	*XY*
1	153	8	23,409	64	1,224
2	172	10	29,584	100	1,720
3	197	15	38,809	225	2,955
4	178	9	31,684	81	1,602
5	185	12	34,225	144	2,220
6	199	13	39,601	169	2,587
7	205	12	42,025	144	2,460
8	226	16	51,076	256	3,616
Totals	1,515	95	290,413	1,183	18,384
Average	189.375	11.875			

b = (18384 − 8 × 189.375 × 11.875)/(290,413 − 8 × 189.375 × 189.375) = 0.1121

a = 11.875 − 0.1121 × 189.375 = −9.3495

Sales(y) = −9.349 + 0.1121 (Contracts)

(b)

$$r = (8 \times 18384 - 1515 \times 95) / \sqrt{((8 \times 290{,}413 - 1515^2)(8 \times 1183 - 95^2))}$$
$$= 0.8963$$
$$S_{xy} = \sqrt{(1183 - (-9.3495 \times 95) - (0.112 \times 18384/6)} = 1.3408$$
$$r^2 = .8034$$

4.49 (a)

		Method → Exponential Smoothing		
		0.6 = α		
Year	Deposits (*Y*)	Forecast	|Error|	$Error^2$
1	0.25	0.25	0.00	0.00
2	0.24	0.25	0.01	0.0001
3	0.24	0.244	0.004	0.0000
4	0.26	0.241	0.018	0.0003
5	0.25	0.252	0.002	0.00
6	0.30	0.251	0.048	0.0023
7	0.31	0.280	0.029	0.0008
8	0.32	0.298	0.021	0.0004
9	0.24	0.311	0.071	0.0051
10	0.26	0.268	0.008	0.0000
11	0.25	0.263	0.013	0.0002
12	0.33	0.255	0.074	0.0055
13	0.50	0.300	0.199	0.0399
14	0.95	0.420	0.529	0.2808
15	1.70	0.738	0.961	0.925
16	2.30	1.315	0.984	0.9698
17	2.80	1.906	0.893	0.7990
18	2.80	2.442	0.357	0.1278
19	2.70	2.656	0.043	0.0018
20	3.90	2.682	1.217	1.4816
21	4.90	3.413	1.486	2.2108
22	5.30	4.305	0.994	0.9895
23	6.20	4.90	1.297	1.6845
24	4.10	5.680	1.580	2.499
25	4.50	4.732	0.232	0.0540
26	6.10	4.592	1.507	2.2712
27	7.70	5.497	2.202	4.8524
28	10.10	6.818	3.281	10.7658
29	15.20	8.787	6.412	41.1195

4.49 (a) Continued

		Method → Exponential Smoothing 0.6 = α		
Year	**Deposits (*Y*)**	**Forecast**	**|Error|**	**Error²**
30	18.10	12.6350	5.46498	29.8660
31	24.10	15.9140	8.19	67.01
32	25.60	20.8256	4.774	22.7949
33	30.30	23.69	6.60976	43.69
34	36.00	27.6561	8.34390	69.62
35	31.10	32.6624	1.56244	2.44121
36	31.70	31.72	0.024975	0.000624
37	38.50	31.71	6.79	46.1042
38	47.90	35.784	12.116	146.798
39	49.10	43.0536	6.046	36.56
40	55.80	46.6814	9.11856	83.1481
41	70.10	52.1526	17.9474	322.11
42	70.90	62.9210	7.97897	63.66
43	79.10	67.7084	11.3916	129.768
44	94.00	74.5434	19.4566	378.561
TOTALS	787.30		150.3	1513.22
AVERAGE	17.8932		3.416 (MAD)	34.39 (MSE)

Next period forecast = 86.2173 Standard error = 6.07519

		Method → Linear Regression (Trend Analysis)		
Year	**Period (*X*)**	**Deposits (*Y*)**	**Forecast**	**Error²**
1	1	0.25	–17.330	309.061
2	2	0.24	–15.692	253.823
3	3	0.24	–14.054	204.31
4	4	0.26	–12.415	160.662
5	5	0.25	–10.777	121.594
6	6	0.30	–9.1387	89.0883
7	7	0.31	–7.50	61.0019
8	8	0.32	–5.8621	38.2181
9	9	0.24	–4.2238	19.9254
10	10	0.26	–2.5855	8.09681
11	11	0.25	–0.947	1.43328
12	12	0.33	0.691098	0.130392
13	13	0.50	2.329	3.34667
14	14	0.95	3.96769	9.10642
15	15	1.70	5.60598	15.2567
16	16	2.30	7.24427	24.4458
17	17	2.80	8.88257	36.9976
18	18	2.80	10.52	59.6117
19	19	2.70	12.1592	89.4756
20	20	3.90	13.7974	97.9594
21	21	4.90	15.4357	111.0
22	22	5.30	17.0740	138.628
23	23	6.20	18.7123	156.558
24	24	4.10	20.35	264.083
25	25	4.50	21.99	305.862
26	26	6.10	23.6272	307.203
27	27	7.70	25.2655	308.547
28	28	10.10	26.9038	282.367
29	29	15.20	28.5421	178.011
30	30	18.10	30.18	145.936
31	31	24.10	31.8187	59.58
32	32	25.60	33.46	61.73
33	33	30.30	35.0953	22.9945
34	34	36.00	36.7336	0.5381
35	35	31.10	38.3718	52.8798
36	36	31.70	40.01	69.0585
37	37	38.50	41.6484	9.91266
38	38	47.90	43.2867	21.2823
39	39	49.10	44.9250	17.43
40	40	55.80	46.5633	85.3163
41	41	70.10	48.2016	479.54
42	42	70.90	49.84	443.528
43	43	79.10	51.4782	762.964
44	44	94.00	53.1165	1671.46
TOTALS	990.00	787.30		7559.95
AVERAGE	22.50	17.893		171.817 (MSE)

	Method → Least Squares–Simple Regression on GSP				
	a	b			
	–17.636	13.5936			
Coefficients:	**GPS**	**Deposits**			
Year	**(*X*)**	**(*Y*)**	**Forecast**	**|Error|**	**Error²**
1	0.40	0.25	–12.198	12.4482	154.957
2	0.40	0.24	–12.198	12.4382	154.71
3	0.50	0.24	–10.839	11.0788	122.740
4	0.70	0.26	–8.12	8.38	70.226
5	0.90	0.25	–5.4014	5.65137	31.94
6	1.00	0.30	–4.0420	4.342	18.8530
7	1.40	0.31	1.39545	1.08545	1.17820
8	1.70	0.32	5.47354	5.15354	26.56
9	1.30	0.24	0.036086	0.203914	0.041581
10	1.20	0.26	–1.3233	1.58328	2.50676
11	1.10	0.25	–2.6826	2.93264	8.60038
12	0.90	0.33	–5.4014	5.73137	32.8486
13	1.20	0.50	–1.3233	1.82328	3.32434
14	1.20	0.95	–1.3233	2.27328	5.16779
15	1.20	1.70	–1.3233	3.02328	9.14020
16	1.60	2.30	4.11418	1.81418	3.29124
17	1.50	2.80	2.75481	0.045186	0.002042
18	1.60	2.80	4.11418	1.31418	1.727
19	1.70	2.70	5.47354	2.77354	7.69253
20	1.90	3.90	8.19227	4.29227	18.4236
21	1.90	4.90	8.19227	3.29227	10.8390
22	2.30	5.30	13.6297	8.32972	69.3843
23	2.50	6.20	16.3484	10.1484	102.991
24	2.80	4.10	20.4265	16.3265	266.556
25	2.90	4.50	21.79	17.29	298.80
26	3.40	6.10	28.5827	22.4827	505.473
27	3.80	7.70	34.02	26.32	692.752
28	4.10	10.10	38.0983	27.9983	783.90
29	4.00	15.20	36.74	21.54	463.924
30	4.00	18.10	36.74	18.64	347.41
31	3.90	24.10	35.3795	11.2795	127.228
32	3.80	25.60	34.02	8.42018	70.8994
33	3.80	30.30	34.02	3.72018	13.8397
34	3.70	36.00	32.66	3.33918	11.15
35	4.10	31.10	38.0983	6.99827	48.9757
36	4.10	31.70	38.0983	6.39827	40.9378
37	4.00	38.50	36.74	1.76	3.10146
38	4.50	47.90	43.5357	4.36428	19.05
39	4.60	49.10	44.8951	4.20491	17.6813
40	4.50	55.80	43.5357	12.2643	150.412
41	4.60	70.10	44.8951	25.20	635.288
42	4.60	70.90	44.8951	26.00	676.256
43	4.70	79.10	46.2544	32.8456	1078.83
44	5.00	94.00	50.3325	43.6675	1906.85
TOTALS				451.223	9016.45
AVERAGE				10.2551 (MAD)	204.92 (MSE)

Forecasting Summary Table			
Method used:	**Exponential Smoothing**	**Linear Regression (Trend Analysis)**	**Linear Regression**
		$Y = -18.968 + 1.638 \times \text{YEAR}$	$Y = -17.636 + 13.59364 \times \text{GSP}$
MAD	3.416	10.587	10.255
MSE	34.39	171.817	204.919
Standard Error using $n - 2$ in denominator	6.075	13.416	14.651
Correlation coefficient		0.846	0.813

Given that one wishes to develop a five-year forecast, trend analysis is the appropriate choice. Measures of error and goodness-of-fit are really irrelevant. Exponential smoothing provides a forecast only of deposits for the *next* year—and thus does not address the five-year forecast problem. In order to use the regression model based upon GSP, one must first develop a model to forecast GSP, and then use the forecast of GSP in the model to forecast deposits. This requires the development of *two* models—one of which (the model for GSP) must be based solely on time as the independent variable (time is the only other variable we are given).

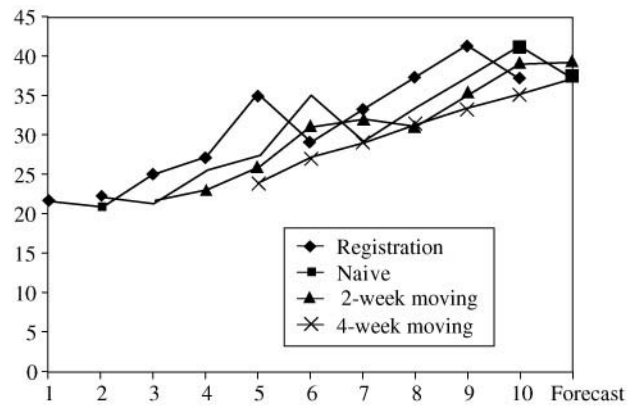
(b) One could make a case for exclusion of the older data. Were we to exclude data from roughly the first 25 years, the forecasts for the later years would likely be considerably more accurate. Our argument would be that a change that caused an increase in the rate of growth appears to have taken place at the end of that period. Exclusion of this data, however, would not change our choice of forecasting model because we still need to forecast deposits for a future five-year period.

INTERNET HOMEWORK PROBLEMS

These problems appears on our companion web site at www.prenhall.com/heizer

4.50

	Week	1	2	3	4	5	6	7	8	9	10	Forecast
	Registration	22	21	25	27	35	29	33	37	41	37	
(a)	Naïve		22	21	25	27	35	29	33	37	41	37
(b)	2-week moving			21.5	23	26	31	32	31	35	39	39
(c)	4-week moving					23.75	27	29	31	33.5	35	37



4.51

Period	Demand	Exponentially Smoothed Forecast
1	7	5
2	9	5 + 0.2 × (7 – 5) = 5.4
3	5	5.4 + 0.2 × (9 – 5.4) = 6.12
4	9	6.12 + 0.2 × (5 – 6.12) = 5.90
5	13	5.90 + 0.2 × (9 – 5.90) = 6.52
6	8	6.52 + 0.2 × (13 – 6.52) = 7.82
7	Forecast	7.82 + 0.2 × (8 – 7.82) = 7.86

4.52

Actual	Forecast	|Error|	$Error^2$
95	100	5	25
108	110	2	4
123	120	3	9
130	130	0	0
		10	38

MAD = 10/4 = 2.5, MSE = 38/4 = 9.5

4.53 (a) 3-month moving average:

Month	Sales	Three-Month Moving Average	Absolute Deviation
January	11		
February	14		
March	16		
April	10	(11 + 14 + 16)/3 = 13.67	3.67
May	15	(14 + 16 + 10)/3 = 13.33	1.67
June	17	(16 + 10 + 15)/3 = 13.67	3.33
July	11	(10 + 15 + 17)/3 = 14.00	3.00
August	14	(15 + 17 + 11)/3 = 14.33	0.33
September	17	(17 + 11 + 14)/3 = 14.00	3.00
October	12	(11 + 14 + 17)/3 = 14.00	2.00
November	14	(14 + 17 + 12)/3 = 14.33	0.33
December	16	(17 + 12 + 14)/3 = 14.33	1.67
January	11	(12 + 14 + 16)/3 = 14.00	3.00
February		(14 + 16 + 11)/3 = 13.67	

Σ = 22.00
MAD = 2.20

(b) 3-month weighted moving average

Month	Sales	Three-Month Moving Average Moving	Absolute Deviation
January	11		
February	14		
March	16		
April	10	(1 × 11 + 2 × 14 + 3 × 16)/6 = 14.50	4.50
May	15	(1 × 14 + 2 × 16 + 3 × 10)/6 = 12.67	2.33
June	17	(1 × 16 + 2 × 10 + 3 × 15)/6 = 13.50	3.50
July	11	(1 × 10 + 2 × 15 + 3 × 17)/6 = 15.17	4.17
August	14	(1 × 15 + 2 × 17 + 3 × 11)/6 = 13.67	0.33
September	17	(1 × 17 + 2 × 11 + 3 × 14)/6 = 13.50	3.50
October	12	(1 × 11 + 2 × 14 + 3 × 17)/6 = 15.00	3.00
November	14	(1 × 14 + 2 × 17 + 3 × 12)/6 = 14.00	0.00
December	16	(1 × 17 + 2 × 12 + 3 × 14)/6 = 13.83	2.17
January	11	(1 × 12 + 2 × 14 + 3 × 16)/6 = 14.67	3.67
February		(1 × 14 + 2 × 16 + 3 × 11)/6 = 13.17	

Σ = 27.17
MAD = 2.72

(c) Based on a Mean Absolute Deviation criterion, the 3-month moving average with *MAD* = 2.2 is to be preferred over the 3-month weighted moving average with *MAD* = 2.72.

(d) Other factors that might be included in a more complex model are interest rates and cycle or seasonal factors.

4.54 (a)

Week	Actual Miles	Forecast	Error	RSFE	Σ |Error|	Cum. *MAD*	Tracking Signal
1	17	17.00	0.00	–	0.00	0	
2	21	17.00	–4.00	–4.00	4.00	2	–2
3	19	17.80	–1.20	–5.20	5.20	1.73	–3
4	23	18.04	–4.96	–10.16	10.16	2.54	–4
5	18	19.03	+1.03	–9.13	11.19	2.24	–4
6	16	18.83	+2.83	–6.30	14.02	2.34	–2.7
7	20	18.26	–1.74	–8.04	15.76	2.25	–3.6
8	18	18.61	+0.61	–7.43	16.37	2.05	–3.6
9	22	18.49	–3.51	–10.94	19.88	2.21	–5
10	20	19.19	–0.81	–11.75	20.69	2.07	–5.7
11	15	19.35	+4.35	–7.40	25.04	2.28	–3.2
12	22	18.48	–3.52	–10.92	28.56	2.38	–4.6

(b) The *MAD* = 28.56/12 = 2.38

(c) The *RSFE* and tracking signals appear to be consistently negative, and at week 10, the tracking signal exceeds 5 *MAD*s.

4.55

y	*x*	x^2	*xy*
7	1	1	7
9	2	4	18
5	3	9	15
11	4	16	44
10	5	25	50
13	6	36	78
55	21	91	212

$\bar{y} = 9.17$

$\bar{x} = 3.5$

$y = 5.27 + 1.11x$

Period 7 forecast = 13.07

Period 12 forecast = 18.64, but this is far outside the range of valid data.

4.56 To compute seasonalized or adjusted sales forecast, we just multiply each seasonalized index by the appropriate trend forecast.

$$\hat{Y}_{\text{Seasonal}} = Index \times \hat{Y}_{\text{Trend forecast}}$$

Hence, for

$$\text{Quarter I: } \hat{Y}_I = 1.25 \times 120{,}000 = 150{,}000$$
$$\text{Quarter II: } \hat{Y}_{II} = 0.90 \times 140{,}000 = 126{,}000$$
$$\text{Quarter III: } \hat{Y}_{III} = 0.75 \times 160{,}000 = 120{,}000$$
$$\text{Quarter IV: } \hat{Y}_{IV} = 1.10 \times 180{,}000 = 198{,}000$$

4.57

	Mon	Tue	Wed	Thu	Fri	Sat	
Week 1	210	178	250	215	160	180	
Week 2	215	180	250	213	165	185	
Week 3	220	176	260	220	175	190	
Week 4	225	178	260	225	176	190	
Averages	217.5	178	255	218.3	169	186.3	Overall average = 204

(a) Seasonal indexes

1.066 (Mon) 0.873 (Tue) 1.25 (Wed)
1.07 (Thu) 0.828 (Fri) 0.913 (Sat)

(b) To calculate for Monday of Week 5 = 201.74 + 0.18(25) × 1.066 = 219.9 rounded to 220

Forecast 220 (Mon) 180 (Tue) 258 (Wed)
221 (Thu) 171 (Fri) 189 (Sat)

4.58 (a) 4000 + 0.20(15,000) = 7,000

(b) 4000 + 0.20(25,000) = 9,000

4.59 (a) 35 + 20(80) + 50(3.0) = 1,785

(b) 35 + 20(70) + 50(2.5) = 1,560

4.60 Given: $\Sigma X = 15$, $\Sigma Y = 20$, $\Sigma XY = 70$, $\Sigma X^2 = 55$, $\Sigma Y^2 = 130$, $\bar{X} = 3$, $\bar{Y} = 4$

(a)
$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$
$$a = \bar{Y} - b\bar{X}$$
$$b = \frac{70 - 5 \times 3 \times 4}{55 - 5 \times 3^2} = \frac{70 - 60}{55 - 45} = \frac{10}{10} = 1$$
$$a = 4 - 1 \times 3 = 4 - 3 = 1$$
$$Y = 1 + 1X$$

(b) Correlation coefficient:

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - \left(\sum X\right)^2\right]\left[n\sum Y^2 - \left(\sum Y\right)^2\right]}}$$
$$= \frac{5 \times 70 - 15 \times 20}{\sqrt{\left[5 \times 55 - 15^2\right]\left[5 \times 130 - 20^2\right]}}$$
$$= \frac{350 - 300}{\sqrt{[275 - 225][650 - 400]}} = \frac{50}{\sqrt{50 \times 250}}$$
$$= \frac{50}{111.80} = 0.45$$

The correlation coefficient indicates that there is a positive correlation between bank deposits and consumer price indices in Birmingham, Alabama—i.e., as one variable tends to increase (or decrease), the other tends to increase (or decrease).

4.60 (c) Standard error of the estimate:

$$S_{yx} = \sqrt{\frac{\sum Y^2 - a\sum Y - b\sum XY}{n - 2}} = \sqrt{\frac{130 - 1 \times 20 - 1 \times 70}{3}}$$
$$= \sqrt{\frac{130 - 20 - 70}{3}} = \sqrt{\frac{40}{3}} = \sqrt{13.3} = 3.65$$

4.61

	X	Y	X^2	Y^2	XY
	2	4	4	16	8
	1	1	1	1	1
	4	4	16	16	16
	5	6	25	36	30
	3	5	9	25	15
Column Totals	15	20	55	94	70

Given: $Y = a + bX$ where:

$$b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2}$$
$$a = \bar{Y} - b\bar{X}$$

and $\Sigma X = 15$, $\Sigma Y = 20$, $\Sigma XY = 70$, $\Sigma X^2 = 55$, $\Sigma Y^2 = 94$, $\bar{X} = 3$, $\bar{Y} = 4$. Then:

$$b = \frac{70 - 5 \times 4 \times 3}{55 - 5 \times 3^2} = \frac{70 - 60}{55 - 45} = 1.0$$
$$a = 4 - 1 \times 3 = 1.0$$

and $Y = 1.0 + 1.0X$. The correlation coefficient:

$$r = \frac{n\sum XY - \sum X \sum Y}{\sqrt{\left[n\sum X^2 - \left(\sum X\right)^2\right]\left[n\sum Y^2 - \left(\sum Y\right)^2\right]}}$$
$$= \frac{5 \times 70 - 15 \times 20}{\sqrt{\left[5 \times 55 - 15^2\right]\left[5 \times 94 - 20^2\right]}} = \frac{350 - 300}{\sqrt{[275 - 225][470 - 400]}}$$
$$= \frac{50}{\sqrt{50 \times 70}} = \frac{50}{59.16} = 0.845$$

Standard error of the estimate:

$$S_{yx} = \sqrt{\frac{\sum Y^2 - a\sum Y - b\sum XY}{n-2}} = \sqrt{\frac{94 - 1\times 20 - 1\times 70}{5-2}}$$

$$= \sqrt{\frac{94 - 20 - 70}{3}} = \sqrt{1.333} = 1.15$$

4.62 Using software, the regression equation is: Games lost = 6.41 + 0.533 × days rain.

CASE STUDIES

1 SOUTHWESTERN UNIVERSITY: B

This is the second of a series of integrated case studies that run throughout the text.

1. One way to address the case is with separate forecasting models for each game. Clearly, the homecoming game (week 2) and the fourth game (craft festival) are unique attendance situations.

		Forecasts		
Game	**Model**	**2006**	**2007**	R^2
1	y = 30,713 + 2,534x	48,453	50,988	0.92
2	y = 37,640 + 2,146x	52,660	54,806	0.90
3	y = 36,940 + 1,560x	47,860	49,420	0.91
4	y = 22,567 + 2,143x	37,567	39,710	0.88
5	y = 30,440 + 3,146x	52,460	55,606	0.93
Total		239,000	250,530	

(where y = attendance and x = time)

2. Revenue in 2006 = (239,000) ($20/ticket) = $4,780,000

Revenue in 2007 = (250,530) ($21/ticket) = $5,261,130

3. In games 2 and 5, the forecast for 2007 exceeds stadium capacity. With this appearing to be a continuing trend, the time has come for a new or expanded stadium.

1 DIGITAL CELL PHONE, INC.

Objectives:

- Selection of an appropriate time series forecasting model based upon a plot of the data.
- The importance of combining a qualitative model with a quantitative model in situations where technological change is occurring.

A plot of the data indicates a linear trend (least squares) model might be appropriate for forecasting. Using linear trend you obtain the following:

	x	y	x^2	xy	y^2
	1	480	1	480	230400
	2	436	4	872	190096
	3	482	9	1446	232324
	4	448	16	1792	200704
	5	458	25	2290	209464
	6	489	36	2934	239121
	7	498	49	3486	248004
	8	430	64	3440	184900
	9	444	81	3996	197136
	10	496	100	4960	246016
	11	487	121	5357	237169
	12	525	144	6300	275625
	13	575	169	7475	330625
	14	527	196	7378	277729
	15	540	225	8100	291600
	16	502	256	8032	252004
	17	508	289	8636	258064
	18	573	324	10314	328329
	19	508	361	9652	258064
	20	498	400	9960	248004
	21	485	441	10185	235225
	22	526	484	11572	276676
	23	552	529	12696	304704
	24	587	576	14088	344569
	25	608	625	15200	369664
	26	597	676	15522	356409
	27	612	729	16524	374544
	28	603	784	16884	363609
	29	628	841	18212	394384
	30	605	900	18150	366025
	31	627	961	19437	393129
	32	578	1024	18496	334084
	33	585	1089	19305	342225
	34	581	1156	19754	337561
	35	632	1225	22120	399424
	36	656	1296	23616	430336
Totals	666	19,366	16,206	378,661	10,558,246
Average	18.5	537.9	450.2	10,518.4	293,284.6

$$b = \frac{\Sigma xy - n\bar{x}\,\bar{y}}{\Sigma x^2 - n\bar{x}^2} = \frac{378{,}661 - 36\times 18.5\times 537.9}{16{,}206 - (36\times 18.5^2)} = \frac{20390.0}{3885.0} = 5.2$$

$$a = \bar{y} - b\bar{x} = 537.9 - 5.2\times 18.5 = 440.8$$

$$r = \frac{n\Sigma xy - \Sigma x\Sigma y}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}}$$

$$= \frac{(36)(378{,}661) - (666)(19{,}366)}{\sqrt{[(36)\times(16{,}206) - (666)^2][(36)(10{,}558{,}246) - (19{,}366)^2]}}$$

$$= \frac{13{,}631{,}796 - 12{,}897{,}756}{\sqrt{[(583{,}416) - (443{,}556)][380{,}096{,}856) - (375{,}041{,}956)]}}$$

$$= \frac{737{,}040}{\sqrt{[139{,}860][5{,}054{,}900]}} = \frac{734{,}040}{\sqrt{706{,}978{,}314{,}000}}$$

$$= \frac{734{,}040}{840{,}820} = .873$$

Y = 440.8 + 5.2 (time)
r = 0.873 indicating a reasonably good fit

The student should report the linear trend results, but deflate the forecast obtained based upon qualitative information about industry and technology trends.

VIDEO CASE STUDY

FORECASTING AT HARD ROCK CAFE

There is a short video (8 minutes) available from Prentice Hall and filmed specifically for this text that supplements this case. A 2 minute version of the video also appears on the student CD in the text.

1. Hard Rock uses forecasting for: (1) sales (guest counts) at cafes, (2) retail sales, (3) banquet sales, (4) concert sales, (5) evaluating managers, and (6) menu planning. They could also employ these techniques to forecast: retail store sales of individual (SKU) product demands; sales of each entrée; sales at each work station, etc.

2. The POS system captures all the basic sales data needed to drive individual cafe's scheduling/ordering. It also is aggregated at corporate HQ. Each entrée sold is counted as one guest at a Hard Rock Café.

3. The weighting system is subjective, but is reasonable. More weight is given to each of the past 2 years than to 3 years ago. This system actually protects managers from large sales variations outside their control. One could also justify a 50%–30%–20% model or some other variation.

4. Other predictors of café sales could include: season of year (weather); hotel occupancy; Spring Break from colleges; beef prices; promotional budget; etc.

5. $Y = a + bx$

Month	Advertising X	Guest count Y	X^2	Y^2	XY
1	14	21	196	441	294
2	17	24	289	576	408
3	25	27	625	729	675
4	25	32	625	1,024	800
5	35	29	1,225	841	1,015
6	35	37	1,225	1,369	1,295
7	45	43	2,025	1,849	1,935
8	50	43	2,500	1,849	2,150
9	60	54	3,600	2,916	3,240
10	60	66	3,600	4,356	3,960
Totals	366	376	15,910	15,950	15,772
Average	36.6	37.6			

$$b = \frac{15,772 - 10 \times 36.6 \times 37.6}{15,910 - 10 \times 36.6^2} = 0.7996$$

$$a = 37.6 - 0.7996 \times 36.6 = 8.3363$$

$$Y = 8.3363 + 0.7996X$$

At $65,000; y = 8.34 + .799 (65,000) = 8.34 + 51.97 = 60,300 guests.

For the instructor who asks other questions from this one:
$R^2 = 0.8869$
Std. Error = 5.062

INTERNET CASE STUDIES*

1 THE NORTH-SOUTH AIRLINE

Northern Airline Data

Year	Airframe Cost per Aircraft	Engine Cost per Aircraft	Average Age (hrs)
1998	51.80	43.49	6512
1999	54.92	38.58	8404
2000	69.70	51.48	11077
2001	68.90	58.72	11717
2002	63.72	45.47	13275
2003	84.73	50.26	15215
2004	78.74	79.60	18390

Southeast Airline Data

Year	Airframe Cost per Aircraft	Engine Cost per Aircraft	Average Age (hrs)
1998	13.29	18.86	5107
1999	25.15	31.55	8145
2000	32.18	40.43	7360
2001	31.78	22.10	5773
2002	25.34	19.69	7150
2003	32.78	32.58	9364
2004	35.56	38.07	8259

Utilizing the software package provided with this text, we can develop the following regression equations for the variables of interest:

Northern Airlines—Airframe Maintenance Cost:
- Cost = 36.10 + 0.0026 × Airframe age
- Coefficient of determination = 0.7695
- Coefficient of correlation = 0.8772

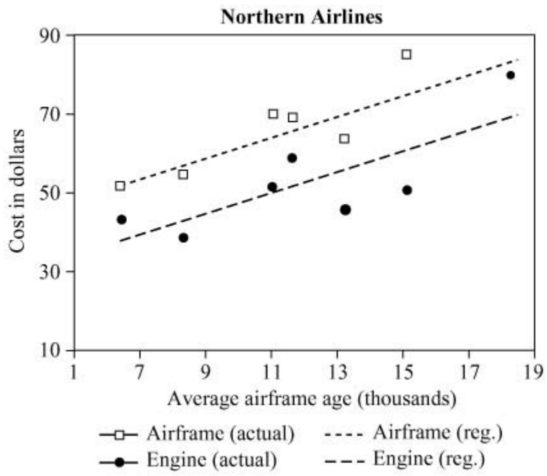
Northern Airlines—Engine Maintenance Cost:
- Cost = 20.57 + 0.0026 × Airframe age
- Coefficient of determination = 0.6124
- Coefficient of correlation = 0.7825

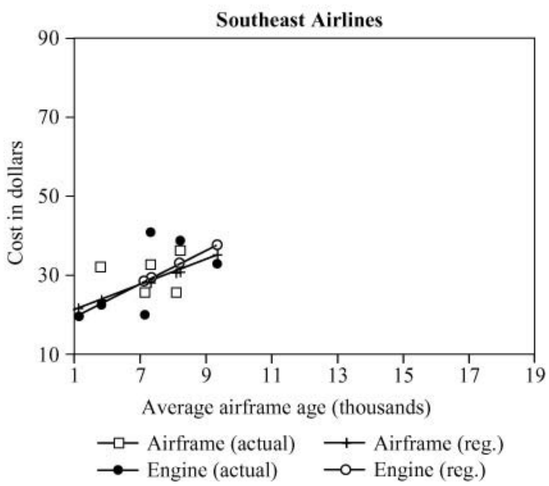
Southeast Airlines—Airframe Maintenance Cost:
- Cost = 4.60 + 0.0032 × Airframe age
- Coefficient of determination = 0.3905
- Coefficient of correlation = 0.6249

Southeast Airlines—Engine Maintenance Cost;
- Cost = –0.67 + 0.0041 × Airframe age
- Coefficient of determination = 0.4600
- Coefficient of correlation = 0.6782

The following graphs portray both the actual data and the regression lines for airframe and engine maintenance costs for both airlines.

*These case studies appear on our companion web site at www.prenhall.com/heizer.





Note that the two graphs have been drawn to the same scale to facilitate comparisons between the two airlines.

Comparison:

- *Northern Airlines*: There seem to be modest correlations between maintenance costs and airframe age for Northern Airlines. There is certainly reason to conclude, however, that airframe age is not the only important factor:
- *Southeast Airlines*: The relationships between maintenance costs and airframe age for Southeast Airlines are much less well defined. It is even more obvious that airframe age is not the only important factor—perhaps not even the most important factor.

Overall: It would seem that:

- Northern Airlines has the smallest variance in maintenance costs—indicating that its day-to-day management of maintenance is working pretty well.
- Maintenance costs seem to be more a function of airline than of airframe age.
- The airframe and engine maintenance costs for Southeast Airlines are not only lower, but more nearly similar than those for Northern Airlines. From the graphs, at least, they appear to be rising more sharply with age.
- From an overall perspective, it appears that Southeast Airlines may perform more efficiently on sporadic or emergency repairs, and Northern Airlines may place more emphasis on preventive maintenance.

Ms. Young's report should conclude that:

- There is evidence to suggest that maintenance costs *could be made to be a function of airframe age* by implementing more effective management practices.
- The difference between maintenance procedures of the two airlines should be investigated.
- The data with which she is presently working does not provide conclusive results.

Concluding Comment:

The question always arises, with this case, as to whether the data should be merged for the two airlines, resulting in two regressions instead of four. The solution provided is that of the consultant who was hired to analyze the data. The airline's own internal analysts also conducted regressions, but *did* merge the data sets. This shows how statisticians can take different views of the same data.

2 THE AKRON ZOOLOGICAL PARK

1. The instructor can use this question to have the student calculate a simple linear regression using real-world data. The idea is that attendance is a linear function of expected admission fees. Also, the instructor can broaden this question to include several other forecast techniques. For example, exponential smoothing, last-period demand, or *n*-period moving average can be assigned. It can be explained that mean absolute deviation (MAD) is but one of a few methods by which an analyst can select the more appropriate forecast technique and outcome.

First, we perform a linear regression with time as the independent variable. The model that results is:

Admissions = 44,352 + 9,197 × Year (where Year is coded as 1 = 1995, 2 = 1996, etc.)

$r = 0.88$

$MAD = 9{,}662$

$MSE = 201{,}655{,}824$

So the forecasts for 2005 and 2006 are 145,519 and 154,716, respectively. Using a weighted average of $2.875 to represent gate receipts per person, revenues for 2005 and 2006 are $418,367 and $444,808, respectively.

To complicate the situation further, students may legitimately use a regression model to forecast admission fees for each of the three categories or for the weighted average fee. This number would then replace $2.875.

Here is the result of a linear regression using weighted average admission fees as the predicting (independent) variable. Weights are obtained each year by taking 35% of adult fees, plus 50% of children's fees, plus 15% of group fees. The weighted fees each year (1989–1998) are: $0.975, $0.975, $0.975, $0.975, $1.275, $1.775, $1.775, $2.275, $2.20, and $2.875.

Gate admissions = 31,451 + (39,614 × Average fee in given year)

$r = 0.847$

$MAD = 13{,}212$

$MSE = 254{,}434{,}912$

If we assume admission fees are not raised in 2005 and 2006, expected gate admissions = 145,341 in each year and revenues = $417,856.

Comparing the earlier time-series model to this second regression, we note that the r is higher and *MAD* and *MSE* are lower in the time-series approach.

2. The student should respond that the other factors are: the variability of the weather, the special events, the competition, and the role of advertising.

3 HUMAN RESOURCES, INC.

There are three different ways to approach this case. One would be to use time-series analysis; the other two consider the use of multiple regression.

The immediate action should be to look at the data. After the time-series data has been loaded, any method can be run with POM for Windows or Excel OM. After execution, the graph can be plotted, making obvious the increasing trend. Therefore, it would be unwise to use moving averages or simple exponential smoothing. Exponential smoothing with trend is available, and students may want to use it. When running the regression, the standard error is 8.27 (and the correlation is 0.84). Of course, this is very good. But we might be able to do better.

When considering the data in a multiple regression form, one independent variable would be the period numbers (as in single regression). In addition, we could have a variable for three of the four seasons. We do not have a variable for period 4, because this would make the columns linearly dependent. The equation would be $y = 54.68 + 0.64(x_1) - 9.62(x_2) - 0.53(x_3) + 1.24(x_4)$. Students should be asked to interpret this line. The correlation coefficient is 0.89 and the standard error is 7.17.

We could make another try with the data using multiple regression and putting in the data for the previous four seasons as the independent variable. This is a common forecasting "trick." The results are:

$$y = 30.5 - 0.023\left(x_{t-1}\right) + 0.089\left(x_{t-2}\right) + 0.119\left(x_{t-3}\right) + 0.047\left(x_{t-4}\right).$$

This equation yields a standard error of 7.24 and a correlation coefficient of 0.787.

The third run was not as good as the second, but there is one more model that makes sense. Students can add to the second model a new column, in which the numbers 1, 2, 3, 4, . . . are placed in order to pick up any trend. The results are

$$y = 75.52 - 0.21\left(x_{t-1}\right) - 0.14\left(x_{t-2}\right) - 0.13\left(x_{t-3}\right) + 0.29\left(x_{t-4}\right) + 1.15\left(\text{trend}\right)$$

which has a correlation coefficient of 0.823 and a standard error of 6.67. Although our results have improved, they are not as good as the first multiple regression model.