بر اساس کتاب:

Business Analytics

Jeffrey D. Camm et al.

دانشکده مهندسی صنایع
دکتری مهندسی صنایع

عنوان درس:
مدلسازی داده محور

فصل هشتم:

Time Series Analysis & Forecasting

Time Series analysis & Forecasting

- **Quarterly forecasts of sales for one of company's products over the upcoming one-year period.**
  - Production schedules,
  - Raw materials purchasing,
  - Inventory policies,
  - Marketing plans,
  - Cash flows.
- **How should we go about providing the quarterly sales forecasts?**
  - Good judgment, intuition, and an awareness of the state of the economy may give us a rough idea, or feeling, of what is likely to happen in the future
  - Converting that feeling into a number that can be used as next year's sales forecast is challenging.

1

Time Series analysis & Forecasting

- Qualitative methods generally involve the use of expert judgment to develop forecasts when historical data are either unavailable or not applicable.

Quantitative forecasting methods can be used when:

(1) Past information about the variable being forecast is available,

(2) The information can be quantified

(3) It is reasonable to assume that the pattern of the past will continue into the future.

1. Causal or exploratory forecasting methods (e.g. Regression):

- The variable we are forecasting has a cause-and-effect relationship with one or more other variables.
- These methods help explain how the value of one variable impacts the value of another.
- The sales volume for many products is influenced by advertising expenditures
- Supermarket scanners allow retailers to collect point-of-sale data that can then be used to aid in planning sales, coupon targeting, and other marketing and planning efforts.
- "Which products tend to be purchased together?"

2

Time Series analysis & Forecasting

- **Time series:** A sequence of past values of the variable to be forecast at successive points in time or over successive periods of time.
  - The measurements may be taken every hour, day, week, month, year, or at any other regular interval.

2. Time series analysis: Forecasting methods that can be applied to time series.

- The objective of time series analysis is to uncover a pattern in the time series and then extrapolate the pattern to forecast the future.
- First step is to construct a time series plot, which is a graphical presentation of the relationship between time and the time series variable.

3

Time Series analysis & Forecasting

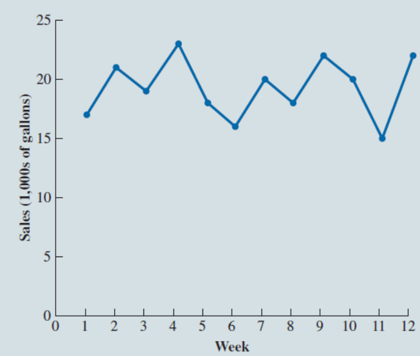
Horizontal Pattern: When the data fluctuate randomly around a constant mean over time.

The average value, or mean, for this time series is 19,250 gallons per week.

TABLE 8.1 Gasoline Sales Time Series

Week	Sales (1,000s of gallons)
1	17
2	21
3	19
4	23
5	18
6	16
7	20
8	18
9	22
10	20
11	15
12	22

FIGURE 8.1 Gasoline Sales Time Series Plot



Time Series analysis & Forecasting

Stationary time series:

- A time series whose statistical properties are independent of time.
  1. The process generating the data has a constant mean.
  2. The variability of the time series is constant over time.

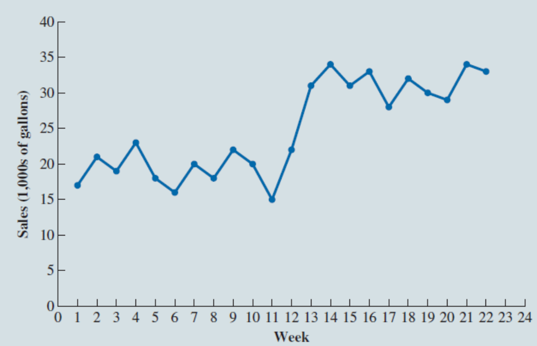
- Changes in business conditions often result in a time series with a horizontal pattern that shifts to a new level at some point in time.
- A gasoline distributor signs a contract with the Vermont State Police to provide gasoline for state police cars beginning in week 13.
- With this new contract, the distributor naturally expects to see a substantial increase in weekly sales starting in week 13.

Time Series analysis & Forecasting

TABLE 8.2 Gasoline Sales Time Series after Obtaining the Contract with the Vermont State Police

Week	Sales (1,000s of gallons)	Week	Sales (1,000s of gallons)
1	17	12	22
2	21	13	31
3	19	14	34
4	23	15	31
5	18	16	33
6	16	17	28
7	20	18	32
8	18	19	30
9	22	20	29
10	20	21	34
11	15	22	33

FIGURE 8.2 Gasoline Sales Time Series Plot after Obtaining the Contract with the Vermont State Police



6

Time Series analysis & Forecasting

Trend Pattern:

- A time series may show gradual shifts or movements to relatively higher or lower values over a longer period of time
- Trend patterns are a result of long-term factors such as population increases/decreases, shifting demographic characteristics of the population, improving technology, changes in the competitive landscape, and/or changes in consumer preferences.

7

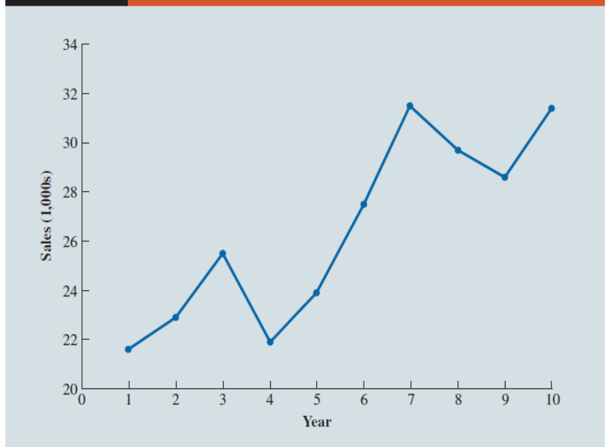
Time Series analysis & Forecasting

A linear and increasing pattern over time

TABLE 8.3 Bicycle Sales Time Series

Year	Sales (1,000s)
1	21.6
2	22.9
3	25.5
4	21.9
5	23.9
6	27.5
7	31.5
8	29.7
9	28.6
10	31.4

FIGURE 8.3 Bicycle Sales Time Series Plot



8

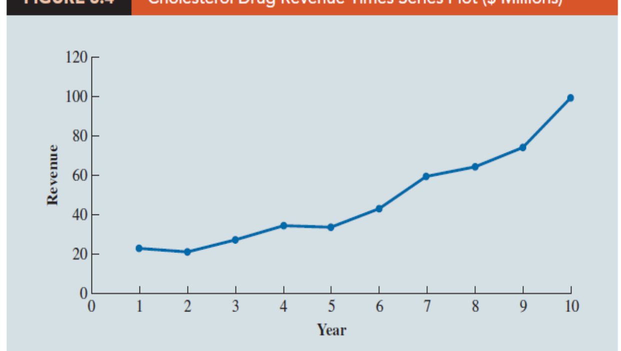
Time Series analysis & Forecasting

An exponential pattern

TABLE 8.4 Cholesterol Drug Revenue Time Series

Year	Revenue ($ millions)
1	23.1
2	21.3
3	27.4
4	34.6
5	33.8
6	43.2
7	59.5
8	64.4
9	74.2
10	99.3

FIGURE 8.4 Cholesterol Drug Revenue Times Series Plot ($ Millions)



9

Time Series analysis & Forecasting

Seasonal Pattern:

- Recurring patterns over successive periods of time that usually repeat every year.
- A retailer who sells bathing suits expects low sales activity in the fall and winter months, with peak sales in the spring and summer months to occur every year.
- Retailers who sell snow removal equipment and heavy clothing expect the opposite yearly pattern.
- Time series data can also exhibit seasonal patterns of less than one year.
- Daily traffic volume shows within-the-day "seasonal" behavior.
- Restaurant industry has sales that exhibit easily discernible seasonal patterns within a day.

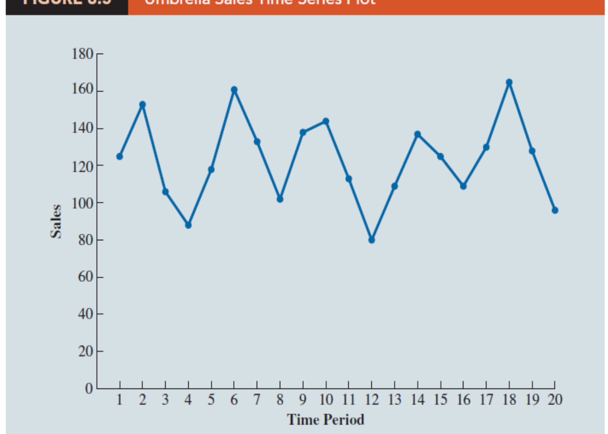
10

Time Series analysis & Forecasting

TABLE 8.5 Umbrella Sales Time Series

Year	Quarter	Sales
1	1	125
	2	153
	3	106
	4	88
2	1	118
	2	161
	3	133
	4	102
3	1	138
	2	144
	3	113
	4	80
4	1	109
	2	137
	3	125
	4	109
5	1	130
	2	165
	3	128
	4	96

FIGURE 8.5 Umbrella Sales Time Series Plot



11

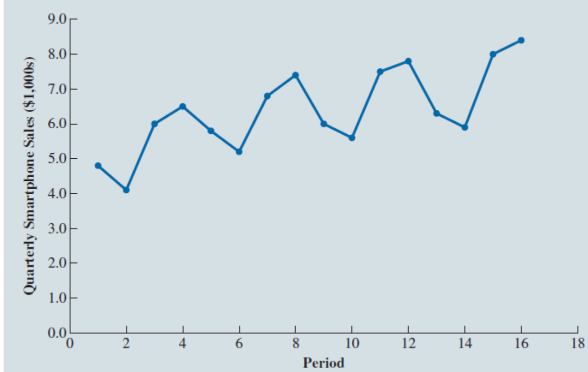
Time Series analysis & Forecasting

Both a trend and seasonal Pattern:

TABLE 8.6 Quarterly Smartphone Sales Time Series

Year	Quarter	Sales ($1,000s)
1	1	4.8
	2	4.1
	3	6.0
	4	6.5
2	1	5.8
	2	5.2
	3	6.8
	4	7.4
3	1	6.0
	2	5.6
	3	7.5
	4	7.8
4	1	6.3
	2	5.9
	3	8.0
	4	8.4

FIGURE 8.6 Quarterly Smartphone Sales Time Series Plot



12

Time Series analysis & Forecasting

Cyclical Pattern:

- An alternating sequence of points below and above the trendline that lasts for more than one year.
- Many economic time series exhibit cyclical behavior.
- Often the cyclical component of a time series is due to multiyear business cycles.
- A time series of housing costs
- Business cycles are extremely difficult or impossible to forecast. As a result, cyclical effects are often combined with long-term trend effects and referred to as trend-cycle effects.

- The underlying pattern in the time series is an important factor in selecting a forecasting method.
- A time series plot is one of the first analytic tools.

13

Time Series analysis & Forecasting

We use the most recent week's sales volume as the forecast for the next week (naïve forecasting method).

TABLE 8.7 Computing Forecasts and Measures of Forecast Accuracy Using the Most Recent Value as the Forecast for the Next Period

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	17						
2	21	17	4	4	16	19.05	19.05
3	19	21	−2	2	4	−10.53	10.53
4	23	19	4	4	16	17.39	17.39
5	18	23	−5	5	25	−27.78	27.78
6	16	18	−2	2	4	−12.50	12.50
7	20	16	4	4	16	20.00	20.00
8	18	20	−2	2	4	−11.11	11.11
9	22	18	4	4	16	18.18	18.18
10	20	22	−2	2	4	−10.00	10.00
11	15	20	−5	5	25	−33.33	33.33
12	22	15	7	7	49	31.82	31.82
		Totals	5	41	179	1.19	211.69

Time Series analysis & Forecasting

Forecast Accuracy:

- How accurate are the forecasts obtained using this naïve forecasting method?

FORECAST ERROR

$$e_t = y_t - \hat{y}_t$$

- We have n periods and k is the number of periods at the beginning for which we cannot produce a naïve forecast:

MEAN FORECAST ERROR (MFE)

$$\text{MFE} = \frac{\sum_{t=k+1}^{n} e_t}{n-k}$$

Positive mean forecast error: The method is generally under-forecasting.

Small mean forecast error: The positive and negative forecast errors tend to offset each other

Time Series analysis & Forecasting

MEAN ABSOLUTE ERROR (MAE)

$$\text{MAE} = \frac{\sum_{t=k+1}^{n} |e_t|}{n-k}$$

MEAN SQUARED ERROR (MSE)

$$\text{MSE} = \frac{\sum_{t=k+1}^{n} e_t^2}{n-k}$$

- The size of the MAE or MSE depends on the scale of the data.
- **MAPE:** To make comparisons for different time intervals or across different time series.

MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

$$\text{MAPE} = \frac{\sum_{t=k+1}^{n} \left| \left(\frac{e_t}{y_t} \right) 100 \right|}{n-k}$$

16

Time Series analysis & Forecasting

If we select a forecasting method that works well for the historical data, and we have reason to believe the historical pattern will continue into the future, our forecasts will ultimately be shown to be accurate.

TABLE 8.8 Computing Forecasts and Measures of Forecast Accuracy Using the Average of All the Historical Data as the Forecast for the Next Period

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	17						
2	21	17.00	4.00	4.00	16.00	19.05	19.05
3	19	19.00	0.00	0.00	0.00	0.00	0.00
4	23	19.00	4.00	4.00	16.00	17.39	17.39
5	18	20.00	−2.00	2.00	4.00	−11.11	11.11
6	16	19.60	−3.60	3.60	12.96	−22.50	22.50
7	20	19.00	1.00	1.00	1.00	5.00	5.00
8	18	19.14	−1.14	1.14	1.31	−6.35	6.35
9	22	19.00	3.00	3.00	9.00	13.64	13.64
10	20	19.33	0.67	0.67	0.44	3.33	3.33
11	15	19.40	−4.40	4.40	19.36	−29.33	29.33
12	22	19.00	3.00	3.00	9.00	13.64	13.64
		Totals	4.52	26.81	89.07	2.75	141.34

17

Time Series analysis & Forecasting

	Naïve Method	Average of All Past Values
MAE	3.73	2.44
MSE	16.27	8.10
MAPE	19.24%	12.85%

- Simple naïve method adjusts very rapidly to the change in level of demand in week 13 because it uses only the most recent observation as the forecast.
- When comparing different forecasting methods, we have to be careful not to rely too heavily on the measures of forecast accuracy.
- Good judgment and knowledge about business conditions that might affect the value of the variable to be forecast also have to be considered carefully when selecting a method.

18

Time Series analysis & Forecasting

Moving Averages: Appropriate for a time series with a horizontal pattern.

- Without modification, they are not appropriate when considerable trend, cyclical, or seasonal effects are present.
- These methods are easy to use and generally provide a high level of accuracy for short-range forecasts, such as a forecast for the next time period.

MOVING AVERAGE FORECAST

$$\hat{y}_{t+1} = \frac{\Sigma(\text{most recent } k \text{ data values})}{k} = \frac{\sum_{i=t-k+1}^{t} y_i}{k}$$

$$= \frac{y_{t-k+1} + \cdots + y_{t-1} + y_t}{k}$$

where

$\hat{y}_{t+1}$ = forecast of the time series for period $t + 1$

y_t = actual value of the time series in period t

k = number of periods of time series data used to generate the forecast

19

Time Series analysis & Forecasting

- In moving averages, a smaller value of k will track shifts in a time series more quickly
- Larger values of k will be more effective in smoothing out random fluctuations.
- managerial judgment based on an understanding of the behavior of a time series is helpful in choosing an appropriate value of k.

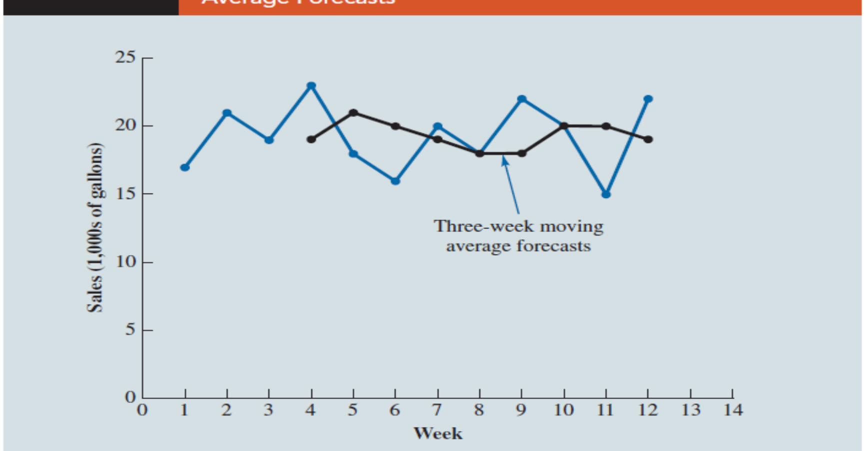
TABLE 8.9 Summary of Three-Week Moving Average Calculations

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	17						
2	21						
3	19						
4	23	19	4	4	16	17.39	17.39
5	18	21	−3	3	9	−16.67	16.67
6	16	20	−4	4	16	−25.00	25.00
7	20	19	1	1	1	5.00	5.00
8	18	18	0	0	0	0.00	0.00
9	22	18	4	4	16	18.18	18.18
10	20	20	0	0	0	0.00	0.00
11	15	20	−5	5	25	−33.33	33.33
12	22	19	3	3	9	13.64	13.64
		Totals	0	24	92	−20.79	129.21

20

Time Series analysis & Forecasting

FIGURE 8.7 Gasoline Sales Time Series Plot and Three-Week Moving Average Forecasts



- Use trial and error to determine the value of k that minimizes the MSE.
- For the gasoline sales time series, min value of MSE corresponds to k = 6 with MSE = 6.79.

21

Time Series analysis & Forecasting

Exponential Smoothing: A weighted average of past time series values

EXPONENTIAL SMOOTHING FORECAST

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t$$

where

$\hat{y}_{t+1}$ = forecast of the time series for period $t + 1$
y_t = actual value of the time series in period t
$\hat{y}_t$ = forecast of the time series for period t
α = smoothing constant ($0 \leq \alpha \leq 1$)

- The weight given to the actual value in period t is the smoothing constant α, and the weight given to the forecast in period t is 1-α.
- Exponential smoothing forecast is actually a weighted average of all the previous actual values.
- Only two pieces of information are needed to forecast for period t +1:
  - y_t: The actual value in period t
  - $\hat{y}_t$: the forecast for period t.

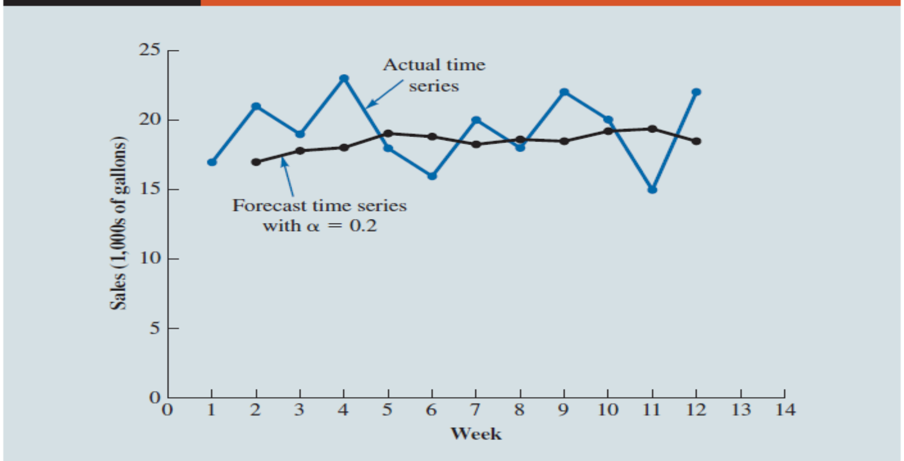
Time Series analysis & Forecasting

TABLE 8.10 Summary of the Exponential Smoothing Forecasts and Forecast Errors for the Gasoline Sales Time Series with Smoothing Constant $\alpha = 0.2$

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	17			
2	21	17.00	4.00	16.00
3	19	17.80	1.20	1.44
4	23	18.04	4.96	24.60
5	18	19.03	−1.03	1.06
6	16	18.83	−2.83	8.01
7	20	18.26	1.74	3.03
8	18	18.61	−0.61	0.37
9	22	18.49	3.51	12.32
10	20	19.19	0.81	0.66
11	15	19.35	−4.35	18.92
12	22	18.48	3.52	12.39
		Totals	10.92	98.80

Time Series analysis & Forecasting

FIGURE 8.11 Actual and Forecast Gasoline Time Series with Smoothing Constant $\alpha = 0.2$



Time Series analysis & Forecasting

Exponential Smoothing:

- If the time series contains substantial random variability, a small value of α is preferred.
- If much of the forecast error is due to random variability, we do not want to overreact and adjust the forecasts too quickly.
- With relatively little random variability, larger values of α allow the forecasts to react more quickly to changing conditions.
- we choose the value of α that minimizes the MSE.

Time Series analysis & Forecasting

Using Regression analysis for forecasting:

Linear Trend Projection:

$$\hat{y}_t = b_0 + b_1 t$$

where

$\hat{y}_t$ = forecast of sales in period t
t = time period
b_0 = the y-intercept of the linear trendline
b_1 = the slope of the linear trendline

$$\hat{y}_t = 20.4 + 1.1t$$

- We can also use more complex regression models to fit nonlinear trends.

Curvilinear trend:

$$\hat{y}_t = b_0 + b_1 t + b_2 t^2 + b_3 t^3$$

26

Time Series analysis & Forecasting

Autoregressive models:

- The independent variables are previous values of the time series

$$\hat{y}_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + b_3 y_{t-3}$$

Seasonality without trend:

- We can model a time series with a seasonal pattern by treating the season as a dummy variable.

$$\text{Qtr1}_t = \begin{cases} 1 \text{ if period } t \text{ is quarter 1} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{Qtr2}_t = \begin{cases} 1 \text{ if period } t \text{ is quarter 2} \\ 0 \text{ otherwise} \end{cases}$$

$$\text{Qtr3}_t = \begin{cases} 1 \text{ if period } t \text{ is quarter 3} \\ 0 \text{ otherwise} \end{cases}$$

- The fourth quarter will be denoted by setting all three dummy variables to 0.

$$\hat{y}_t = b_0 + b_1 \text{Qtr1}_t + b_2 \text{Qtr2}_t + b_3 \text{Qtr3}_t$$

27

Time Series analysis & Forecasting

TABLE 8.11 Umbrella Sales Time Series with Dummy Variables

Period	Year	Quarter	Qtr1	Qtr2	Qtr3	Sales
1	1	1	1	0	0	125
2		2	0	1	0	153
3		3	0	0	1	106
4		4	0	0	0	88
5	2	1	1	0	0	118
6		2	0	1	0	161
7		3	0	0	1	133
8		4	0	0	0	102
9	3	1	1	0	0	138
10		2	0	1	0	144
11		3	0	0	1	113
12		4	0	0	0	80
13	4	1	1	0	0	109
14		2	0	1	0	137
15		3	0	0	1	125
16		4	0	0	0	109
17	5	1	1	0	0	130
18		2	0	1	0	165
19		3	0	0	1	128
20		4	0	0	0	96

28

Time Series analysis & Forecasting

$$\hat{y}_t = 95.0 + 29.0\text{Qtr1}_t + 57.0\text{Qtr2}_t + 26.0\text{Qtr3}_t$$

We can use equation (8.11) to forecast sales of every quarter for the next year:

$$\text{Quarter1: Sales} = 95.0 + 29.0(1) + 57.0(0) + 26.0(0) = 124$$
$$\text{Quarter2: Sales} = 95.0 + 29.0(0) + 57.0(1) + 26.0(0) = 152$$
$$\text{Quarter3: Sales} = 95.0 + 29.0(0) + 57.0(0) + 26.0(1) = 121$$
$$\text{Quarter4: Sales} = 95.0 + 29.0(0) + 57.0(0) + 26.0(0) = 95$$

Seasonality with Trend:

$$\hat{y}_t = b_0 + b_1\text{Qtr1}_t + b_2\text{Qtr2}_t + b_3\text{Qtr3}_t + b_4 t$$

$\hat{y}_t$ = forecast of sales in period t
Qtr1_t = 1 if time period t corresponds to the first quarter of the year; 0 otherwise
Qtr2_t = 1 if time period t corresponds to the second quarter of the year; 0 otherwise
Qtr3_t = 1 if time period t corresponds to the third quarter of the year; 0 otherwise
t = time period (quarter)

The smartphone time series:

$$\hat{y}_t = 6.07 - 1.36\text{Qtr1}_t - 2.03\text{Qtr2}_t - 0.304\text{Qtr3}_t + 0.146t$$

29

Time Series analysis & Forecasting

TABLE 8.12 Smartphone Sales Time Series with Dummy Variables and Time Period

Period	Year	Quarter	Qtr1	Qtr2	Qtr3	Sales (1,000s)
1	1	1	1	0	0	4.8
2		2	0	1	0	4.1
3		3	0	0	1	6.0
4		4	0	0	0	6.5
5	2	1	1	0	0	5.8
6		2	0	1	0	5.2
7		3	0	0	1	6.8
8		4	0	0	0	7.4
9	3	1	1	0	0	6.0
10		2	0	1	0	5.6
11		3	0	0	1	7.5
12		4	0	0	0	7.8
13	4	1	1	0	0	6.3
14		2	0	1	0	5.9
15		3	0	0	1	8.0
16		4	0	0	0	8.4

For monthly data, 11 dummy variables are required.

Time Series analysis & Forecasting

Regression analysis as a Causal forecasting method:

- Advertising expenditures when sales are to be forecast.
- The mortgage rate when new housing construction is to be forecast.
- Grade point average when starting salaries for recent college graduates are to be forecast.
- The price of a product when the demand for the product is to be forecast.
- The value of the Dow Jones Industrial Average when the value of an individual stock is to be forecast.
- Daily high temperature when electricity usage is to be forecast.

Time Series analysis & Forecasting

Armand's Pizza Parlors, a chain of Italian restaurants doing business.

- The most successful locations have been near college campuses.
- It seems that quarterly sales for these restaurants (y) are related positively to the size of the student population (x).
- Management wants to forecast sales for a new restaurant that it is considering opening near a college campus.
- Because no historical data are available on sales for a new restaurant, Armand's cannot use time series data to develop the forecast.

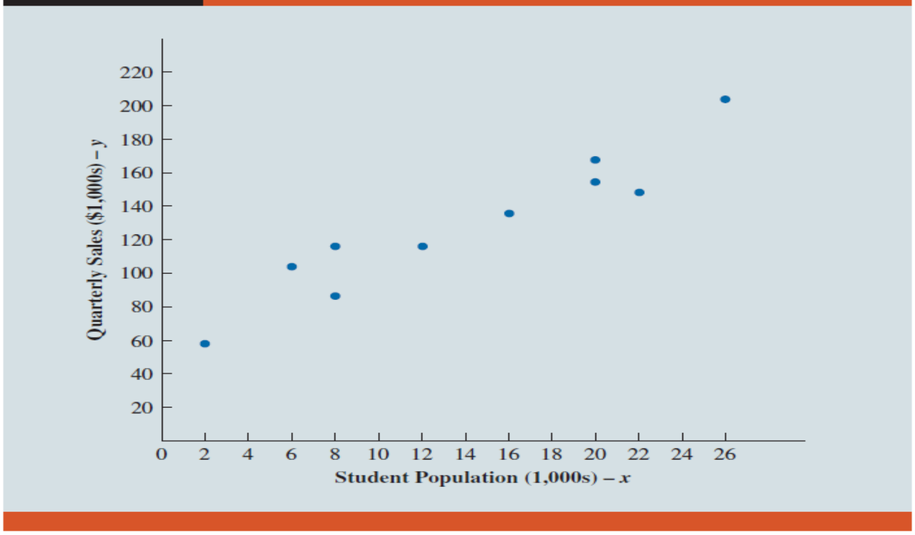
Time Series analysis & Forecasting

TABLE 8.13 Student Population and Quarterly Sales Data for 10 Armand's Pizza Parlors

Restaurant	Student Population (1,000s)	Quarterly Sales ($1,000s)
1	2	58
2	6	105
3	8	88
4	8	118
5	12	117
6	16	137
7	20	157
8	20	169
9	22	149
10	26	202

Time Series analysis & Forecasting

FIGURE 8.16 Scatter Chart of Student Population and Quarterly Sales for Armand's Pizza Parlors



34

Time Series analysis & Forecasting

- What preliminary conclusions can we draw from Figure 8.16?
- Sales appear to be higher at locations near campuses with larger student populations.
- The relationship between the two variables can be approximated by a straight line.

$$\hat{y}_i = b_0 + b_1 x_i \qquad (8.$$

where

$\hat{y}_i$ = estimated value of the dependent variable (quarterly sales) for the ith observation
b_0 = intercept of the estimated regression equation
b_1 = slope of the estimated regression equation
x_i = value of the independent variable (student population) for the i th observation

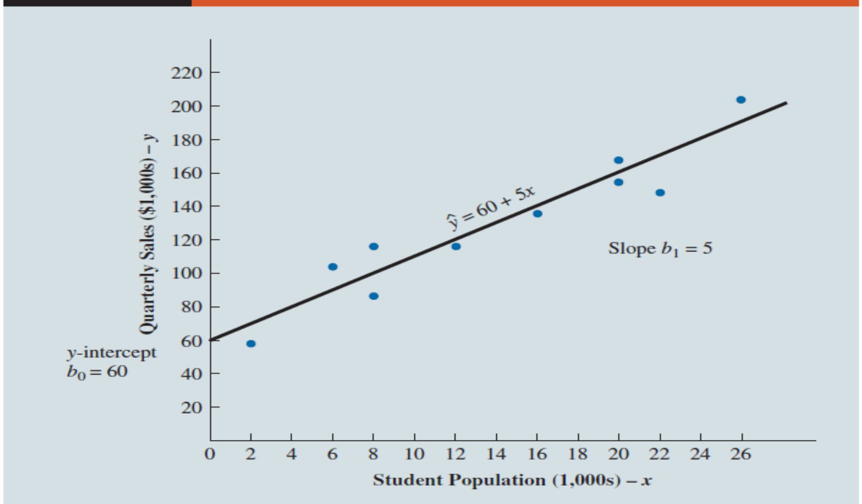
$$\hat{y}_i = 60 + 5x_i$$

Note: The values of the independent variable range from 2,000 to 26,000; thus, the y-intercept is an extrapolation of the regression line and must be interpreted with caution.

35

Time Series analysis & Forecasting

FIGURE 8.17 Graph of the Estimated Regression Equation for Armand's Pizza Parlors: $y = 60 + 5x$



36

Time Series analysis & Forecasting

FIGURE 8.18 Excel Simple Linear Regression Output for Armand's Pizza Parlors

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	*Regression Statistics*								
4	Multiple R	0.950122955							
5	R Square	0.90273363							
6	Adjusted R Square	0.890575334							
7	Standard Error	13.82931669							
8	Observations	10							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*Significance F*			
12	Regression	1	14200	14200	74.24836601	2.54887E-05			
13	Residual	8	1530	191.25					
14	Total	9	15730						
15									
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 99.0%*	*Upper 99.0%*
17	Intercept	60	9.22603481	6.503335532	0.000187444	38.72472558	81.27527442	29.04307968	90.95692032
18	Student Population (1,000s)	5	0.580265238	8.616749156	2.54887E-05	3.661905962	6.338094038	3.052985371	6.947014629

37

Time Series analysis & Forecasting

Combining causal variables with Trend and Seasonality effects:

- We had a time series of several years of quarterly sales data and advertising expenditures for a single Armand's restaurant.
- If we suspected that sales were related to advertising expenditures and that sales showed trend and seasonal effects, we could incorporate each into a single model by combining the approaches.
- If we believe that the effect of advertising is not immediate, we might also try to find a relationship between sales in period t and advertising in period $t - 1$.
- Multiple regression analysis also can be applied in these situations if additional data for other independent variables are available.
- Armand's Pizza Parlors believes that the number of competitors near the college campus is related to quarterly sales.
- Multiple regression analysis could be used to develop an equation relating quarterly sales to the size of the student population and the number of competitors.

38

Time Series analysis & Forecasting

Considerations for using Regression in forecasting:

- Although regression analysis allows for the estimation of complex forecasting models, we must be cautious about using such models and guard against the potential for overfitting to sample data.
- Simple techniques usually outperform more complex procedures for short-term forecasting.
- Using a more sophisticated and expensive procedure will not guarantee better forecasts
- Quantitative forecasting models outperform qualitative forecasts made by "experts."
- There is good reason to use quantitative forecasting methods whenever data are available.

39

Time Series analysis & Forecasting

Determining the best forecasting model:

- For a given forecasting study, how does one choose an appropriate model?
- For time series modeling, a visual inspection can indicate whether seasonality appears to be a factor and whether a linear or nonlinear trend seems to exist.
- For causal model, scatter charts can indicate whether strong linear or nonlinear relationships exist.
- When working with large data sets, it is recommended to divide your data into training and validation sets.
- With five years of monthly data, you could use the first three years as a training set for estimation.
- Based on the errors produced by the different models for the validation set, you could ultimately pick the model that minimizes MAE, MSE, or MAPE.
- **Note**: If the behavior of the time series has changed recently, a forecasting model based on the older portion of the time series will not perform well.

40

Time Series analysis & Forecasting

$$\text{SMAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|e_t|}{(|y_t| + |\hat{y}_t|)/2}$$

RMSE (Root Mean Squared Error): The square root of MSE

$$\text{MASE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|e_t|}{\frac{1}{n-1}\sum_{t=2}^{n}|y_t - y_{t-1}|}$$

41