Introduction to Information Systems

Data Science Education Program

Chapter #7

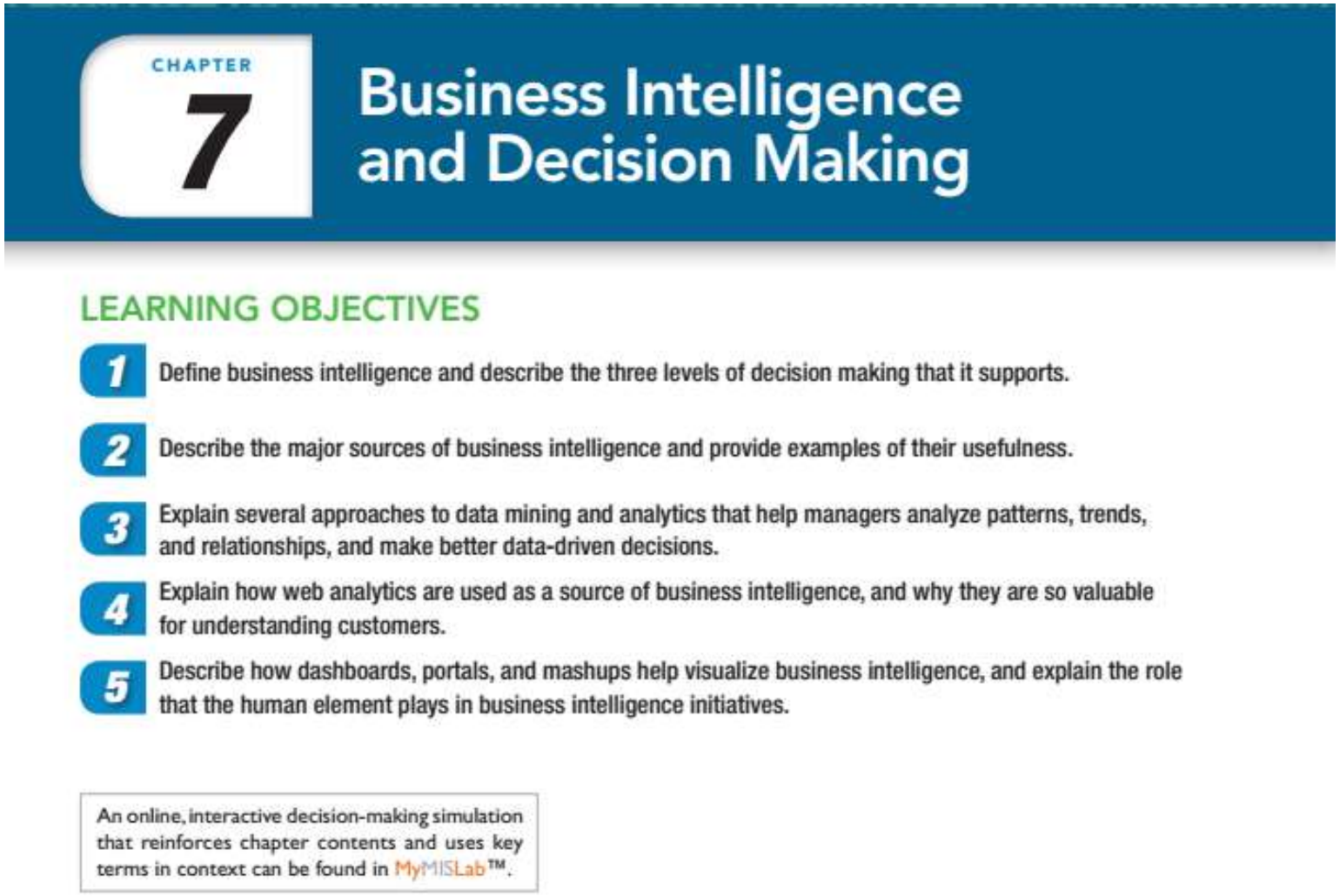
Business intelligence and decision making

Key Terms and Concepts

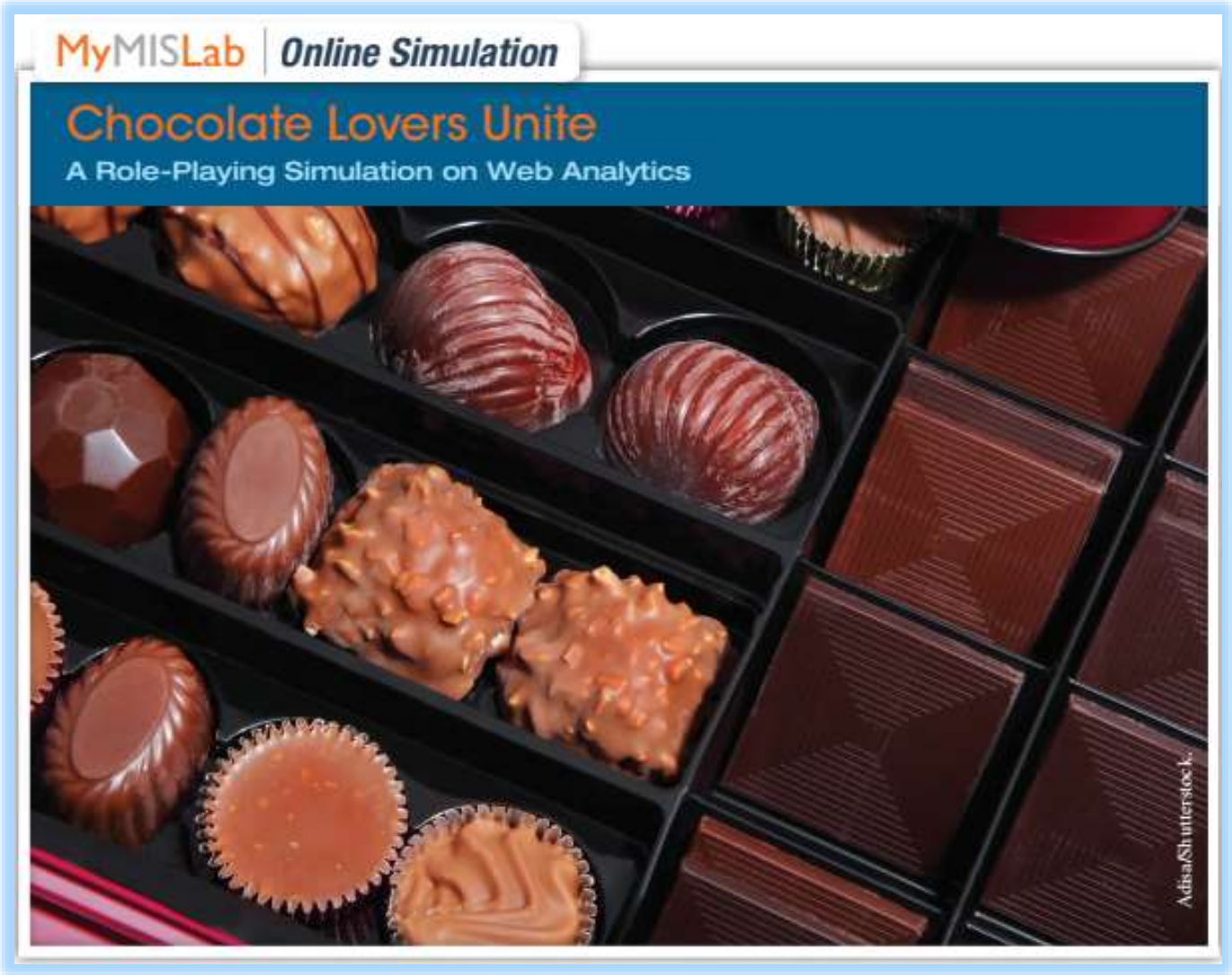
KEY TERMS AND CONCEPTS

intelligent agents
predictive analytics
online analytical processing (OLAP)
market basket analysis
text mining
what-if analysis
goal seeking
optimization
forecasting
artificial intelligence (AI)
CAPTCHA
expert system
neural network
clickstream data
stickiness
dashboard
key performance indicators (KPIs)
portals
mashup
web feed

Chapter #7 Overview



On-line Simulation



Introduction

Chapter #7 Topics

- Levels of Decision-making
  - Operational level
  - Technical level
  - Strategic level
- Sources of business intelligence
  - Transactional databases
  - Data warehouses
  - Internal data sources
- Data mining and analytics
  - Analysing patterns, trends, and relationships
  - Simulating, optimising, and forecasting
  - Artificial intelligence
- Web analytics
  - Web metrics
  - Analysing traffic and achieving success
- Putting it all together: dashboards, portals, and mashups
  - Dashboards
  - Portals
  - Mashups
  - Business intelligence: the human element

Levels of decision making

Levels of Decision-Making

- We have discussed decision-making with the levels and the related information used:
  - This section repeats many of the points previously introduced
- In summary:
  - There are three levels of decisions and information usage:
    - *Strategic*
    - *Tactical*
    - *Operational*
  - In practice there is generally no clear demarcation between the levels
  - Additionally: information may be used interchangeably

Levels of Decision-Making

- The figure repeats the previous figure presented in this course

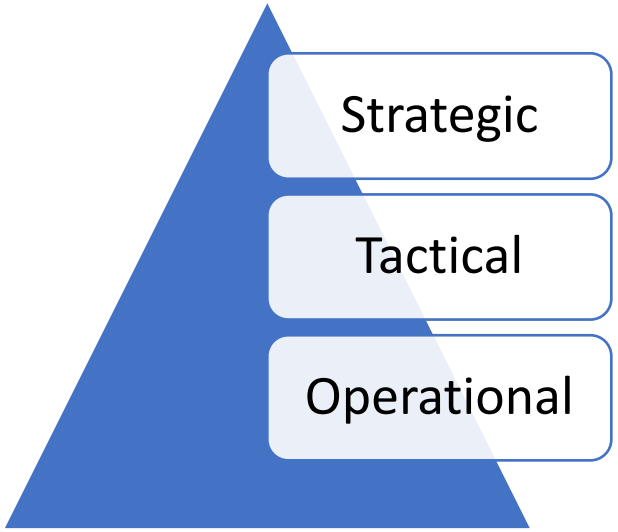


FIGURE 7-1

Different levels of decision making in an organization rely on different mixes of structured and unstructured information.

Operational Decisions

- Operational decisions are:
  - Decisions made in interactions with customers
  - Routine transactions
  - The application of an organization's policies
- Operational decisions are made by:
  - Sales staff dealing directly with customers
  - Staff dealing with everyday operational matters
- The effectiveness of operational staff is often measured against pre-defined targets

Operational Decision Metrics

Metric	Minutes and Seconds
Current call length	**13:42**
Your average call length today	12:41
Today's average call length, for all agents	10:25
This week's average call length, all agents	13:32

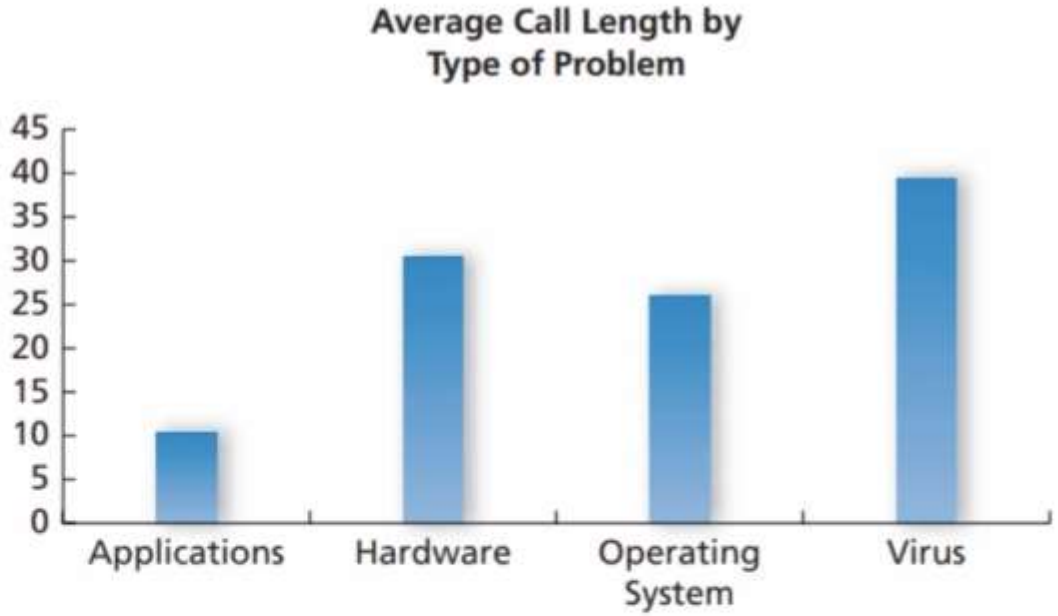


FIGURE 7-2
Call center agents can see timely data on their own and their coworkers' activities, especially to support operational decision making and improved performance.

Tactical and Strategic Decisions

- Tactical decisions are made by mid-level managers and include:
  - Sales and marketing plans
  - Business development
  - Budgeting
  - Etc
- Strategic decisions are made by top-level management (such as a CEO, executives, or directors) and include:
  - Capital expenditure
  - Research and development projects
  - The organisation strategic direction

The Time Scale

- In considering operational, tactical, and strategic decisions *time* forms an important factor in both the nature of the decision and the nature of the information used to reach decisions:
  - *Operational* decisions are generally immediate (*real-time*) decisions and are frequently automated using real-time data and information from a limited range of sources
  - *Tactical* decisions are reflected in relatively *short-to-medium term* time scales and use a broader range of data and information from an increased range of sources
  - *Strategic* decisions are long term objectives for an organisation and rely on a very broad range of data and information

Sources of business intelligence

Sources of Business Intelligence

- Typical intelligence sources include both *internal* and *external* sources of data and information
- Internal sources include:
  - Transactional databases, data warehouses, other internal data sources
- External data sources include:
  - Intelligent agents (see Figure 7.4.)
  - Business intelligence (BI) and big data
  - Cloud-based systems including the Internet-of-things (IoT)
  - External databases
  - Web-sites, social networks, and text messaging etc

Useful Data

U.S. Census Bureau

International Data Base (IDB)

Search This Site

IDB Main | Data Access | Country Rankings | World Population | Methodology | Glossary

Demographic Indicators | Tables | Population Pyramids | Source Information

Morocco

Source: U.S. Census Bureau, International Data Base

DEMOGRAPHIC INDICATORS	2006	2007	2008	2009	2010	1995	2005	2015	2025
Population									
Midyear population (in thousands)	30,248	30,594	30,940	31,285	31,627	26,148	29,901	33,323	36,484
Growth rate (percent)	1.1	1.1	1.1	1.1	1.1	1.6	1.2	1.0	0.8
Fertility									
Total fertility rate (births per woman)	2.4	2.3	2.3	2.3	2.2	3.3	2.4	2.1	2.0
Crude birth rate (per 1,000 population)	21	20	20	20	19	26	21	18	16
Births (in thousands)	622	620	619	617	614	677	622	606	575
Mortality									
Life expectancy at birth (years)	75	75	75	75	76	71	75	77	79
Infant mortality rate (per 1,000 births)	33	32	31	30	29	51	35	24	16
Under 5 mortality rate (per	39	38	36	35	33	62	41	27	19

Source: U.S. Census Bureau.

FIGURE 7-3

Data useful for business intelligence can be downloaded from the U.S. Census Bureau's website.

Useful Data

FIGURE 7-4

To build a bot, the designer carries out the steps a human being would perform to capture data on public websites, and the software creates the agent that will carry out the tasks on its own.

PRODUCTIVITY TIP

You can create an agent to carry out online searches for recent articles relevant to your upcoming term paper. Yahoo, Google, and many online library databases offer this service, usually called "alerts." Schedule it to run daily and email you whatever it finds.

PRODUCTIVITY TIP

How well are you managing your own big data? You might not have petabytes, but you do have great variety and growing volume. Cloud-based services such as Evernote offer tools to organize your web links, tweets, online articles, notes to self, and other resources, and also make them searchable so you can find them again.

intelligent agents

Software programs or "bots" that are sent out to conduct a mission and collect data from web pages on behalf of a user.

Ethical Considerations

THE ETHICAL FACTOR The Ethics of Tagging Faces in Photos

Tagging online photos of faces with people's names is wildly popular and very helpful as a means to add structure to information. Those old group photos come alive when you don't have to struggle to recall long-forgotten names. However, these tags are raising serious privacy concerns and ethical dilemmas.

Although you may have no reluctance to tag yourself in your own photos, will your friends and family want you to tag them? They may think the photo is unflattering. Or they may be concerned that their employers will stumble upon tagged images you thought were amusing, but that employers think show poor judgment. Parents may also object to tagged images of their children appearing online.

Though photo sharing sites offer assurances and many choices about privacy, the potential for harm is not trivial, especially because you can't control what others are doing. Uploaded photos might include metadata that you may not even know is attached, such as GPS coordinates that can "geotag" your photo to indicate location, and the date and time the photo was taken.

The services offer facial recognition software to ease the tedium of individually tagging each photo. After you tag some faces, the software can find those people in other photos and tag them on its own. Though very handy, this tool greatly amplifies the chance that people will be tagged without their knowledge or consent, and the information will be used in unpredictable ways by marketers, employers, relatives, or law enforcement. Facebook implemented this kind of service, but regulators in Europe insisted that Facebook delete all facial recognition data for European users because of privacy concerns.[5] With cell phone cameras widespread and photo uploads so simple to do, the "anonymous face in the crowd" may become rare indeed, despite mounting privacy regulations.

Data mining and analytics

Data Mining and Analytics

- Analysing patterns, trends, and relationships
  - On-line analytical processing (OLAP)
  - Statistics and modelling techniques
  - Text mining
- Simulating, optimising, and forecasting
  - What-if analysis
  - Goal seeking
  - Optimizing
  - Forecasting
- Artificial intelligence
  - Expert systems
  - Artificial neural networks

Data Visualization

- The figure shows a graphical representation of population density in the USA

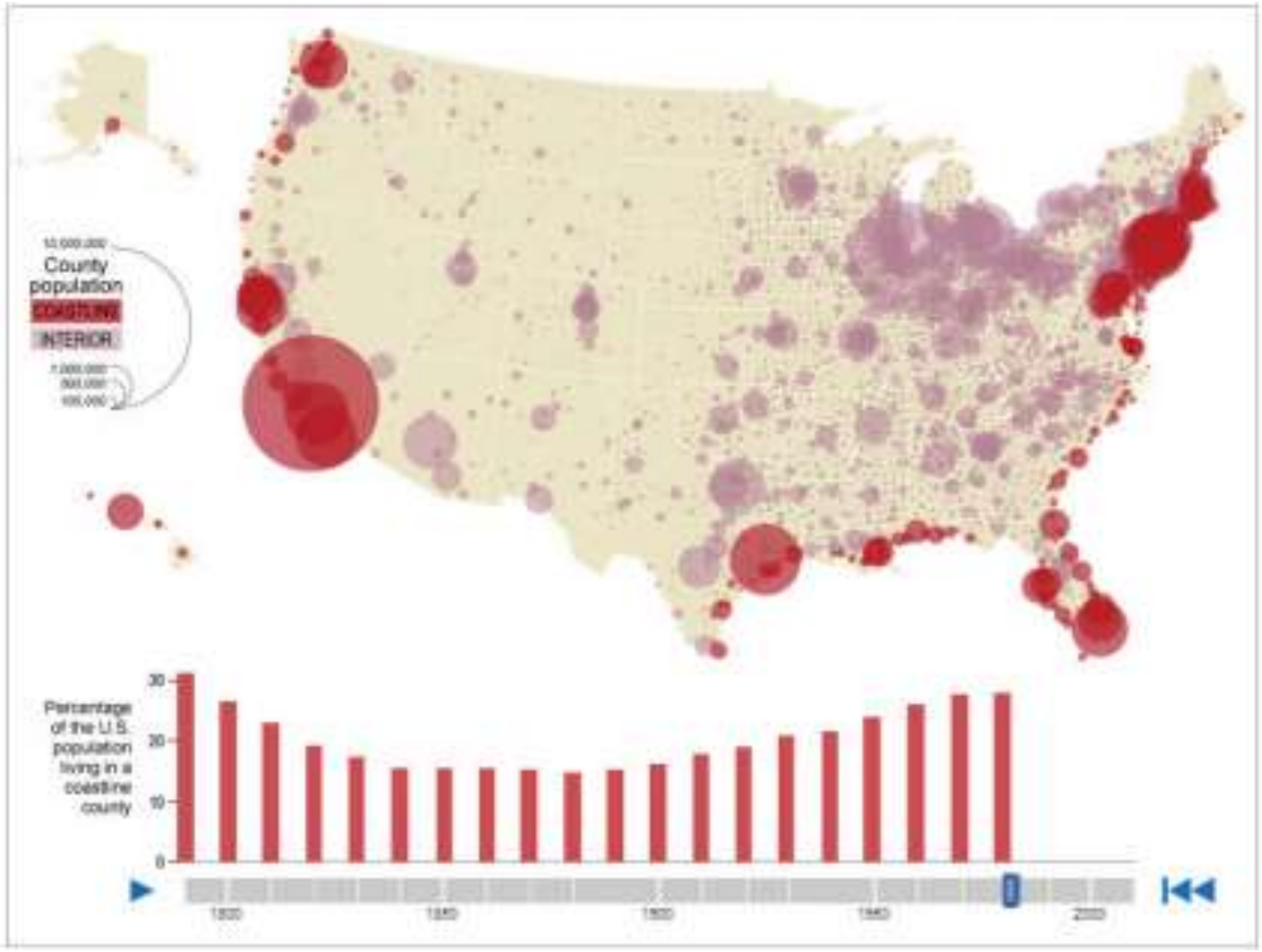


FIGURE 7-5
Data visualization using an interactive map that shows changes in population in coastal counties over time.

OLAP

PRODUCTIVITY TIP

Excel is a powerful OLAP tool itself, thanks to its pivot tables and charting capabilities. It uses in-memory computing so your charts appear almost instantly, even with very large worksheets.

predictive analytics
Data mining approaches and statistical techniques used to predict future behavior, especially to unlock the value of business intelligence for strategy.

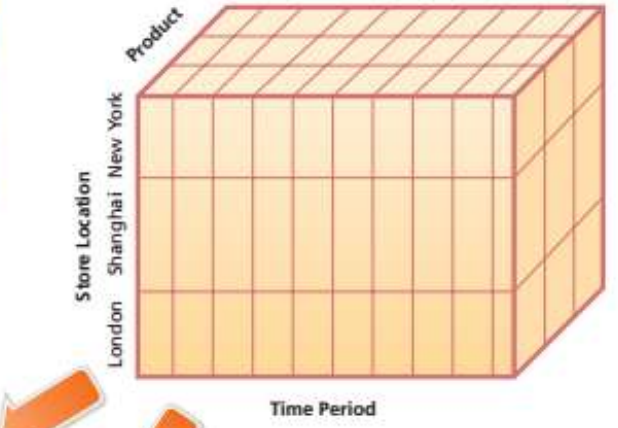
online analytical processing (OLAP)
Software that allows users to "slice and dice" or drill down into massive amounts of data stored in data warehouses to reveal significant patterns and trends.

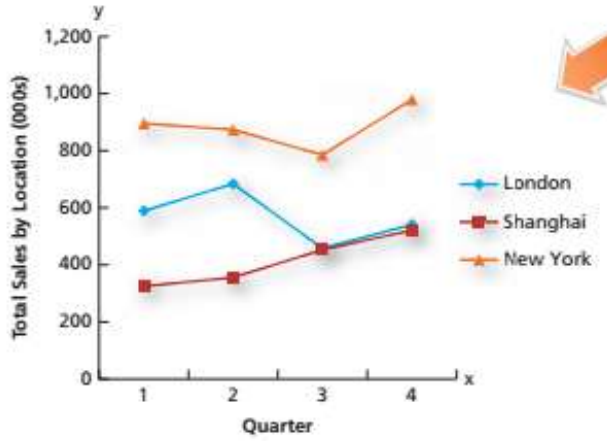
Analysing Patterns

- The figure shows examples of:
  - Online analytical processing
  - Statistics and modeling techniques
  - Text mining

FIGURE 7-6
Multidimensional cube created from tables in a data warehouse and used for OLAP.

Date	Store Location	Product	Quantity
5/14/2012	LONDON	XB395	15
5/14/2012	LONDON	RS585	26
5/15/2012	SHANGHAI	XB395	25
5/15/2012	LONDON	RS585	2
5/15/2012	NEWYORK	XB395	10





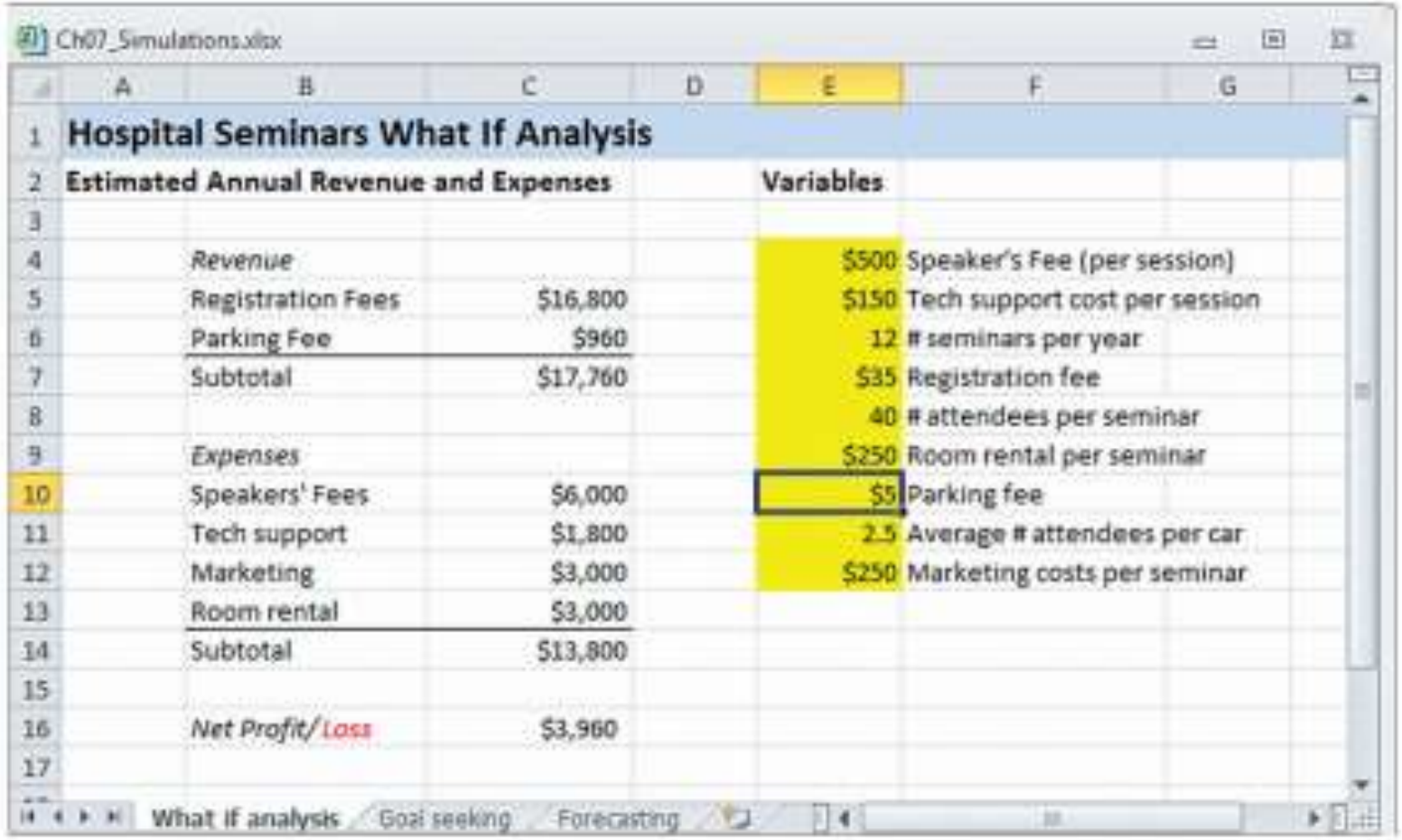
Total Sales by Location and Time Period

	Qtr 1	Qtr 2	Qtr 3	Qtr 4
London	589	685	458	541
Shanghai	325	356	452	521
New York	896	874	785	980

Estimating Revenue and Expenses

FIGURE 7-8
What-if spreadsheet to estimate revenue and expenses for a hospital seminar series. The user can change the estimates for any variable in yellow, and the spreadsheet recomputes Net Profit/Loss.

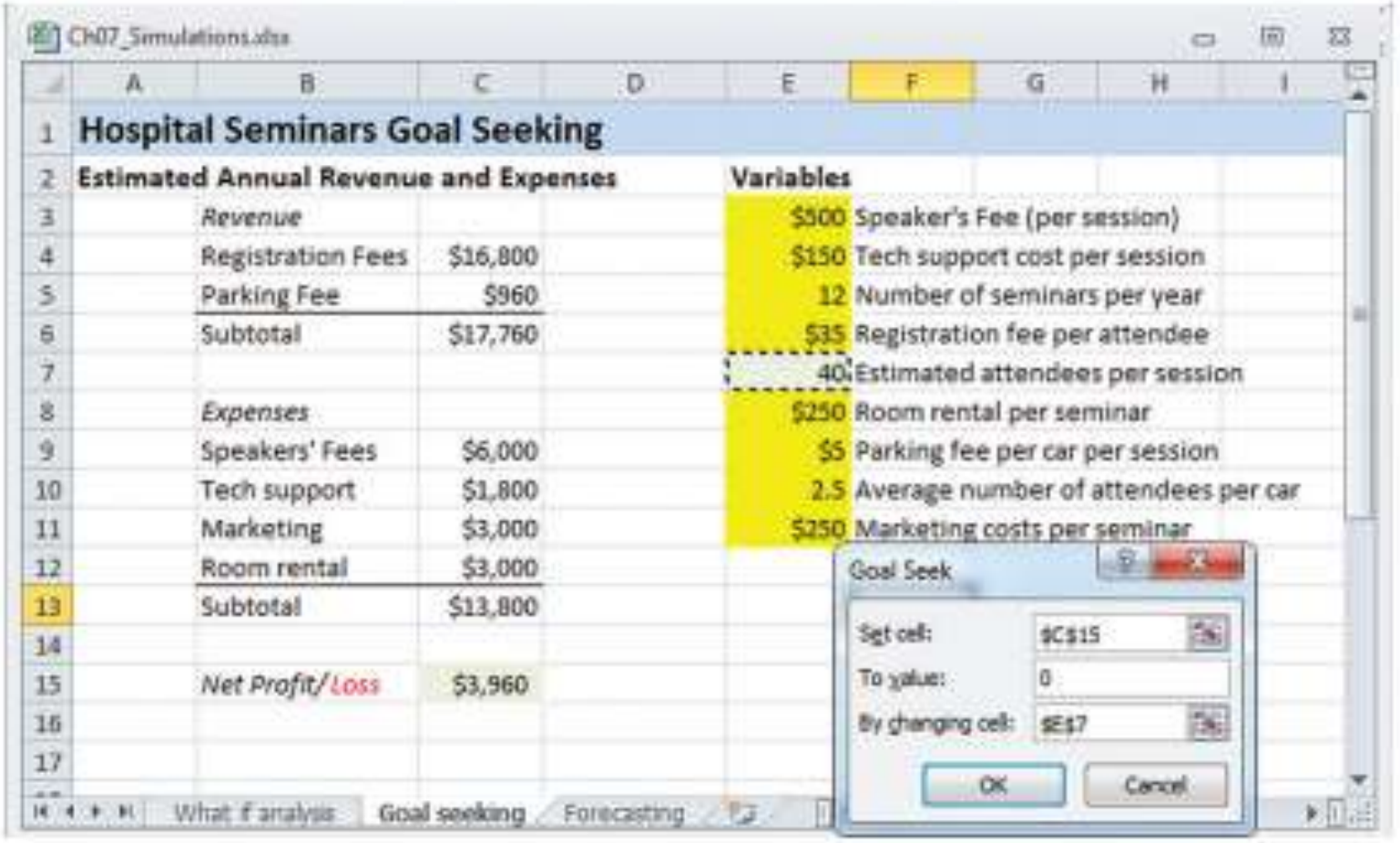
- *What-if* analysis
  - Goal seeking
  - Optimization
  - Forecasting



Goal setting

FIGURE 7-9

Goal seeking. The user sets a target value for one cell, such as Net Profit/Loss, and then enters the cell that Excel can change to reach the target.



Forecasting

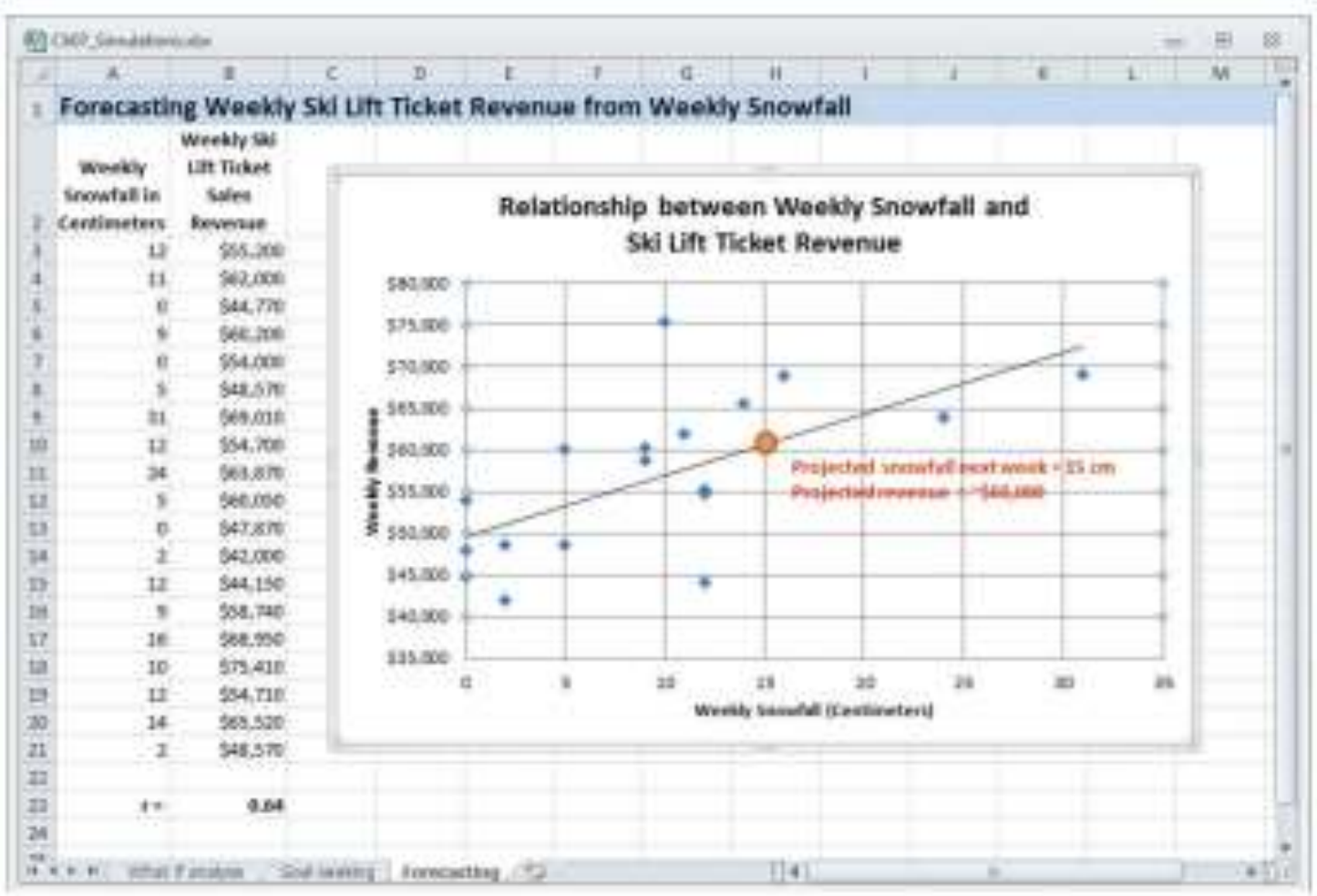


FIGURE 7-10
Forecasting ski lift ticket sales revenue from weekly snowfall based on historical patterns.

Analysis Techniques

what-if analysis
A simulation model, often constructed using Excel, that calculates the relationships between many variables; users can change some variables to see how others are affected.

goal seeking
A decision support tool, often based on an Excel model, in which the user sets a target value for a particular variable, such as profit/loss, and tells the program which variable to change to try to reach the goal.

optimization
An extension of goal seeking in which the user can change many variables to reach some maximum or minimum target, as long as the changes stay within the constraints the user identifies.

forecasting
A statistical decision support tool used to analyze historical trends and other business intelligence to estimate some variable of interest, such as customer demand.

artificial intelligence (AI)
The capability of some machines to mimic aspects of human intelligence, such as learning, reasoning, judging, and drawing conclusions from incomplete information.

Artificial Intelligence Methods

- Machine intelligence
  - Pattern matching
  - Neural networks
  - Expert systems
- Applications:
  - Robotics
  - Machine *vs* human players in games
- Can be used for positive reasons
- Challenges and negative uses

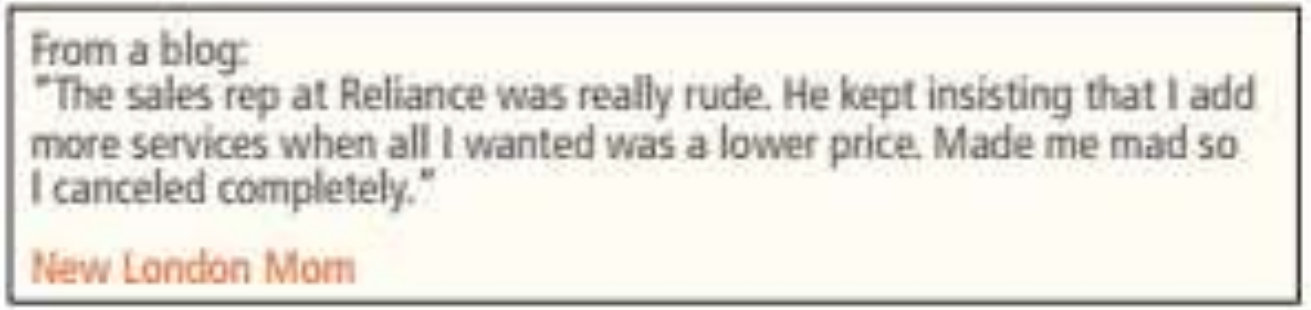
PRODUCTIVITY TIP

An add-in for Excel that you might want to install is called "Analysis ToolPak." Tools include descriptive statistics with mean, mode, median, standard deviation, correlation, and several modeling and hypothesis-testing tools. Some knowledge of statistics is needed.

Text mining on historical documents turned up a fascinating tidbit about how views of the United States changed. Authors wrote "the United States are . . ." well into the 1800s, when they began using "the United States is . . . ." Apparently, Americans didn't think of the United States as a single nation rather than a collection of states until long after the country was born.[12]

FIGURE 7-7

Text mining software can extract useful business intelligence from blogs.



Artificial Intelligence Methods

A usability expert calls the CAPTCHA one of the most hated user interactions on the web. Each one can take a user from 10 to 15 seconds to figure out, assuming they don't give up, which collectively adds up to 500,000 hours every day.[15] But the hours aren't totally wasted. Your human eyes help digitize old books with fonts too strange or blurry for optical character readers, so the text can be searched. CAPTCHA is an acronym for "Completely Automated Public Turing test to tell Computers and Humans Apart."

FIGURE 7-11
A CAPTCHA designed to ensure visitors are actually human beings and not bots.

market basket analysis
A statistical technique that reveals customer behavior patterns as they purchase multiple items.

text mining
A technique used to analyze unstructured text that examines keywords, semantic structures, linguistic relationships, emotion-laden words, and other characteristics to extract meaningful business intelligence.

PRODUCTIVITY TIP
Before you travel outside your home country, it's wise to let your credit card company know where you are headed. A neural net may automatically block your card if unexpected foreign charges appear and the card company is unable to reach you.

Artificial Intelligence Application

FIGURE 7-12
The neural-net called "20Q" plays the game 20 Questions with visitors, very often guessing correctly. The training data includes the millions of games users play at the site (www.20Q.net, from 20Q.net Inc.).

CAPTCHA
A test created by software developers that the visitor must pass before continuing to register or enter the site; designed to thwart software bots.

expert system
Software that mimics the reasoning and decision making of a human expert, drawing from a base of knowledge about a particular subject area developed with the expert's assistance.

neural network
An information system that attempts to mimic the way the human brain works; often used to spot suspicious activity in financial transactions

Web analytics

Web Analytics

- Web metrics
  - Web site metrics
  - Social media metrics
  - E-commerce metrics
- Analysing Internet traffic achieving success
  - Analytics software
  - Reaching goals
  - Analysing the effectiveness of advertisements

Web Visitor Related Metrics

Web Visitor-Related Metrics	Description
Visitors	Number of visitors to the website. (Returning visitors will be counted again if they return within the time period.)
Unique visitors	Number of unique visitors. (Returning visitors are not counted again.)
Average time on site	Average amount of time visitors spent on the site.
New visitors	Number of new visitors to the site.
Depth of visit	The number of page views per visit, which shows how extensively visitors interact with and navigate around your site.
Languages	The number (or percentage) of visitors based on the language they configured to use on their computer.
Traffic sources	The sources from which visitors arrive at your site, such as a keyword search in a search engine, an ad, or from a link on related sites. Direct traffic is a visit from someone who used a bookmark or typed the URL in the browser.
Service providers	The number of visits coming from people using different Internet service providers.

FIGURE 7-13
Website metrics.

Web Content Related Metrics

Web Content-Related Metrics	Description
Page views	The number of visits per page on the site, showing analysts the most popular content.
Bounce rate	Percentage of visits in which the user left the site from the same page he or she entered it. This can mean that the page the user landed on was not very relevant.
Top landing pages	The number of entrances to your site for each page.
Top exit pages	The number of exits from the site for each page.

clickstream data
Business intelligence data that includes every click by every visitor on a website along with associated data such as time spent on the page and the visitor's IP address.

FIGURE 7-13
Website metrics.

Twitter Activity Metrics

Metric	Platform	Description
Number of fans	Facebook	Number of users who "like" a page
Fan reach	Facebook	Number of users who view a particular post
Number of impressions	Twitter	Number of times users saw the tweet on Twitter
Number of retweets	Twitter	Number of times users retweet a company's post
Engagement	Instagram	Number of likes and comments divided by number of followers
Follower industry demographics	LinkedIn	Profile of followers by the industry in which they work

FIGURE 7-14
Sample metrics for social media activity.

E-Commerce Metrics

E-Commerce Metric	Description
Conversion rate	The ratio of visitors who complete some activity (such as buying a product) divided by the total number of visitors.
Clickthrough rate (CTR)	The ratio of clicks on an ad divided by the number of times the ad was delivered.
Cost per clickthrough (CPC)	The amount an advertiser pays each time a visitor clicks on the ad to navigate to the advertiser's site.
Cost per impression (CPM is cost per thousand impressions)	For banner and display ads, the cost the advertiser pays each time the ad loads onto a user's screen from any site on which it appears.
Position on page	The position in which a sponsored link appears on a page in keyword advertising on search engines.

FIGURE 7-15
E-commerce metrics.

E-Commerce Metrics

Metric	Description
Number of downloads	Number of users who download the app
Geographic distribution	Describes users by location
Average session length	Average amount of time users have the app open
User paths	Path that users take through the app

FIGURE 7-16
Sample metrics for mobile apps.

Figure 7.17.

FIGURE 7-17

In the online simulation for this chapter called "Chocolate Lovers Unite," you compare the metrics for two versions of the company's website.

Analyzing Internet Traffic

FIGURE 7-19

Sample output from web analytics software, describing the number of website visitors by day and the most keywords used to reach the site.

Most Popular Keywords	Visits
Peru	987
Latin America	745
South America	158

- Analyzing software
- Reaching goals
- Ad effectiveness

FIGURE 7-19

Online ads can include variables to customize the text according to the phrase the user entered as the search term.

Search term: phones for children	Search term: kids' phones
Phones for Children The Fun & Easy Way to Find **Phones** for **Children** at Low Prices	**Kids' Phone** 100s of Kids Phone Top **Brands** at **Low** Prices!

Putting it all together: dashboards, portals, and mashups

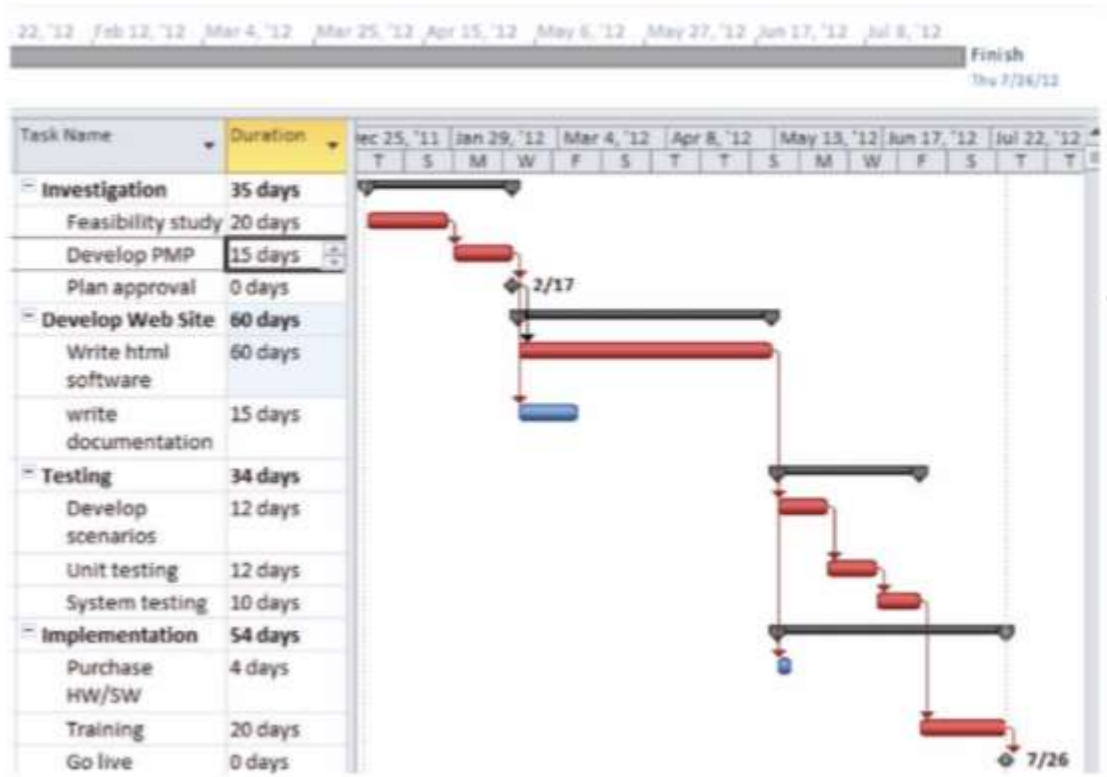
Dashboard Design

Dashboard Feature	Design Tip
Key performance indicators (KPIs)	Choosing KPIs that are most important to the organization and the person using the dashboard is the most important success factor.
Data quality	Dashboards make data look good, but the charts and graphs are only as useful as the quality of the underlying data. Users should ask for warnings when data is stale or not altogether reliable.
Timeliness	Update the dashboard as often as needed for the user's situation—daily, hourly, or minute by minute, if necessary.
Density	Use seven or fewer graphs, charts, maps, or tables on one dashboard, to avoid information overload.
Chart formats	Keep tables small and charts simple, using familiar types. Avoid 3D and unnecessary animations. Be cautious about pie charts, which can be more difficult to interpret.
Maps and visual displays	When relevant, populate actual maps, seating plans, campus layouts, or other visual displays that combine real images with data.

FIGURE 7-22
Dashboard design tips.

Example Dashboard

- Graphical user interface
- Key performance indicators



Portal Design

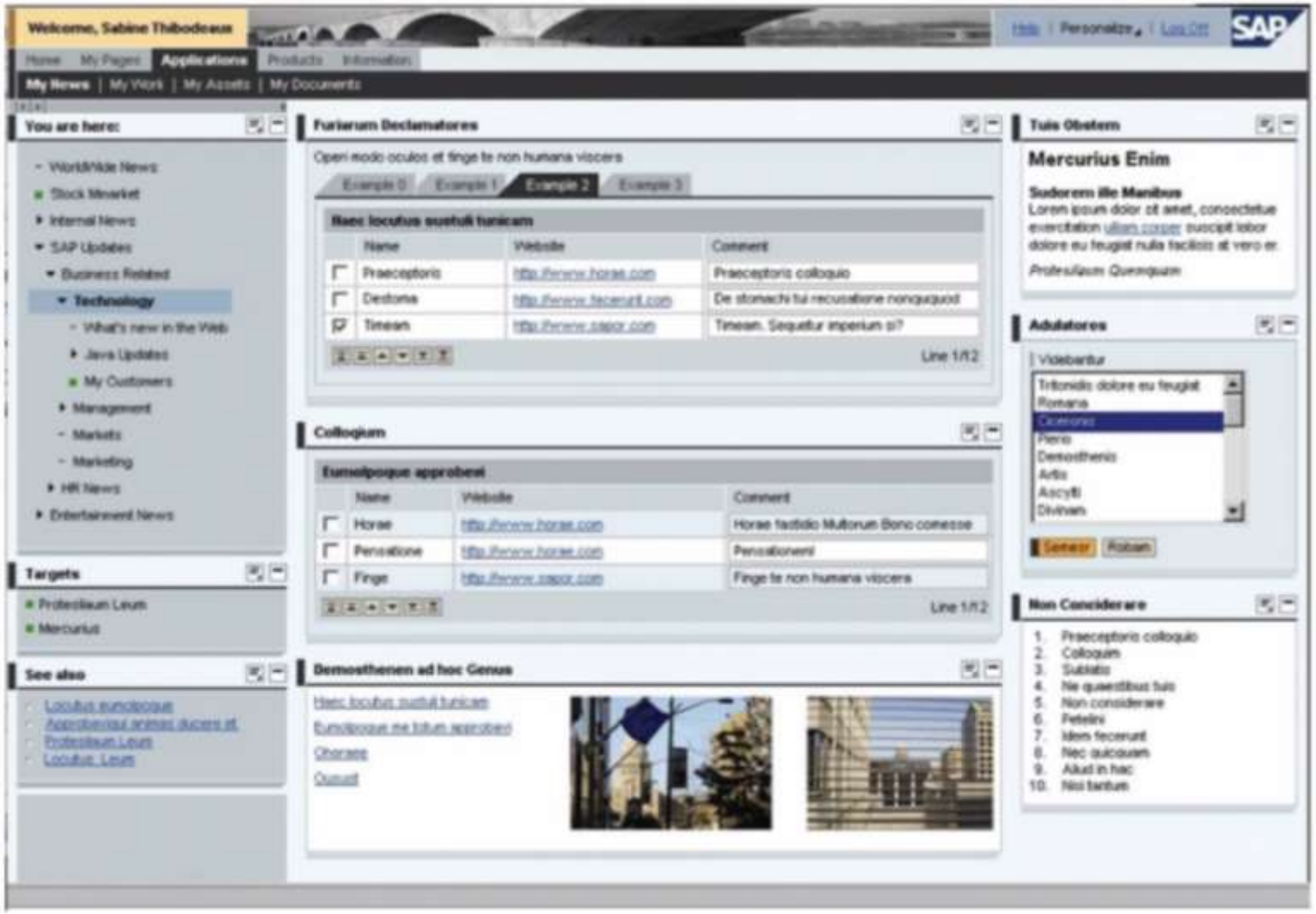


FIGURE 7-23

An enterprise portal.

- Enterprise portal
- Variety of information on one screen

Mashup

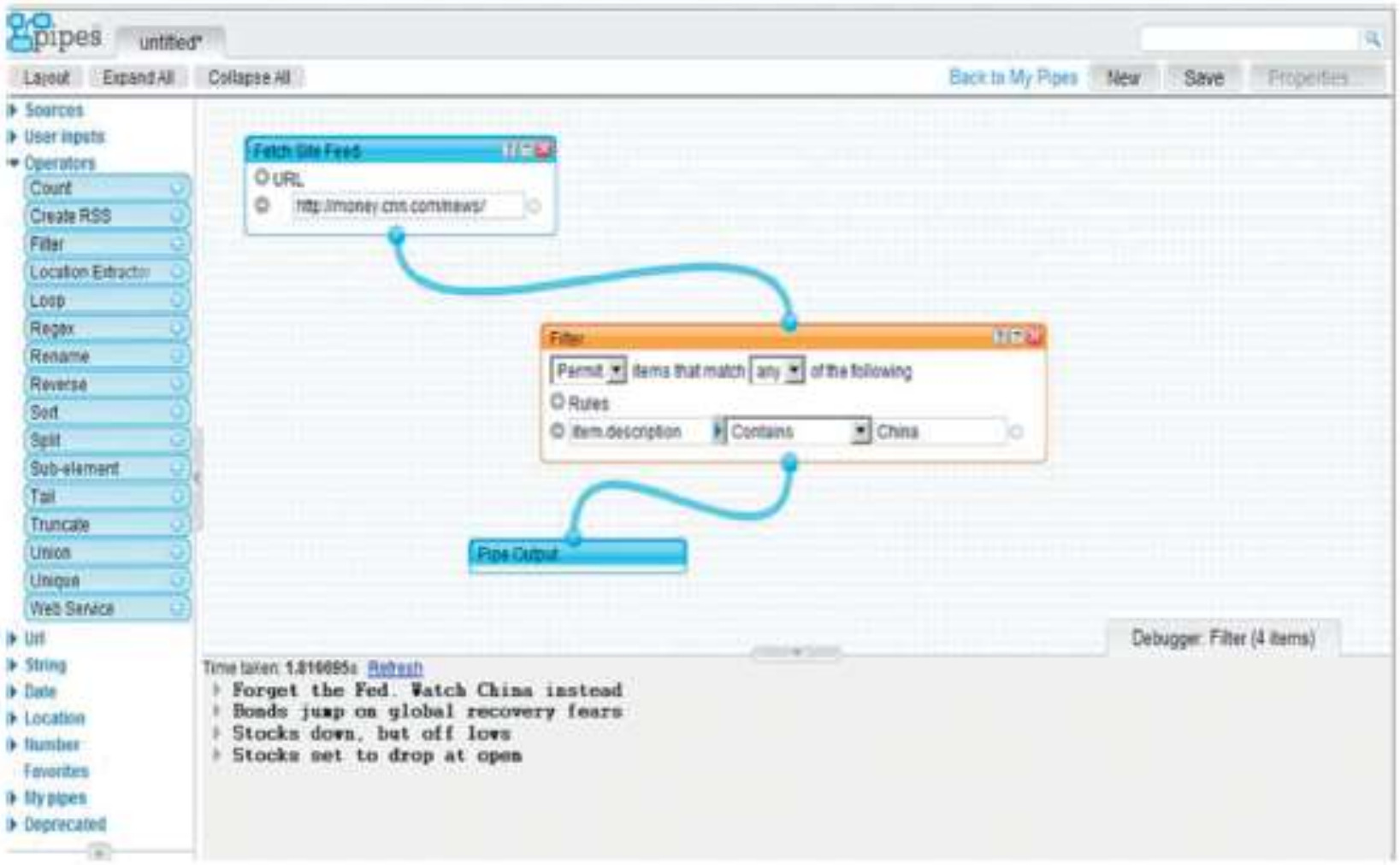


FIGURE 7-24
Mashup modules can be created using software tools such as Yahoo! Pipes. In this example, the module will display updated business news from CNN with the word *China* in the description.

- Aggregate data from multiple sources
- Customizable web pages

Portals and Mashups

PRODUCTIVITY TIP

If your university or college offers a portal, you can experiment to see what portlets are available, what functionality you can access, and how much you can customize it. Set it as your startup page in your browser if you find it useful enough.

portal

A gateway that provides access to a variety of relevant information from many different sources on one screen; for an enterprise, the portal provides a secure gateway to resources needed by employees, customers, and suppliers.

mashup

An approach to aggregating content from multiple internal and external sources on customizable web pages that relies on Web 2.0 technologies.

RSS Web Feeds

- RSS web feeds can:
  - Provide regular and frequent news feeds for selected topics at regular intervals
  - The feeds can be user configured

FIGURE 7-23
This symbol indicates that the website offers a web feed.

PRODUCTIVITY TIP

When you visit a website that displays the orange symbol (Figure 7-23), you can subscribe to that publisher's web feed for your own consumer portal or website. Adding too many, though, could cause your portal to load more slowly.

Business intelligence: the human element

The Human Element

- People form a critical component in organisations and their culture
- Business intelligence must consider the human element in all its aspects:
  - Humans choose intelligence, tools, and interpretation
  - However: humans have weaknesses
- Information systems design and implementation
  - Must consider human motives in relation to information system
  - The attitude of stakeholders in information systems will determine the effectiveness of such systems

Chapter #7 summary, case studies, and coursework

Review

- In this tutorial session we have introduced:
  - Decision making
  - Business intelligence
  - Data mining
  - Web analytics
  - Putting it all together: dashboards and portals
  - The ethical factor
- In the following slides we set out

Chapter # Reading and Coursework

- Read and understand subjects and concepts introduced in Chapter #7
- Learn the meanings of the *key terms and concepts* introduced
- Read and work through:
  - The *Chocolate Lovers Unite* on-line simulation
  - The (two) *case studies* (located on page 253 and 254 at the end of Chapter #7)
  - The (two) *e-projects* (located on page 255 and 256 at the end of Chapter #7)
- At the end of the chapter you will find:
  - Chapter review questions
  - Projects and discussion questions
  - Application exercises
  - Work through these to understand the subjects and concepts in Chapter #7

On-Line Simulation

MyMISLab | *Online Simulation*

Chocolate Lovers Unite

A Role-Playing Simulation on Web Analytics

"Can you believe this?" whispered the VP to you, as a loud argument broke out among the marketing people, sales people, web designers, and just about everyone else in the room. They all thought they knew what design would be most effective for Chocolate Lovers Unite (CLU)—an online chocolate retailer with some of the most sumptuous products on the market. The VP asked for quiet in the room, and when everyone settled down, explained that you would be working with them to evaluate different website approaches.

As a web analyst, you were asked to help CLU resolve these arguments using data-driven decision making and business intelligence. Which marketing pitch will work best? Your job is to find out with real data, not guesswork. Time is short with the gift-giving holiday season fast approaching, so you better get started. Check your email and voice messages after you log in. . . .

CHAPTER 7

Chapter Summary

LEARNING OBJECTIVES

1. Business intelligence encompasses a vast array of information sources that can contribute to better decision making, along with the software applications, technologies, and practices used to analyze it. Levels of decision making that draw on different types of information sources include operational, tactical, and strategic.

2. A primary source of business intelligence is the transactional database, or data warehouse, used by the organization itself for operations, or by its suppliers and customers. Data available online can also be sources, including websites, blogs, email, downloadable tables, wikis, and business reports. Along with data from sensors, scanners, and the Internet, these sources form immense big data collections. Both internal and external sources can include structured, semi-structured, or unstructured data.

3. Data mining and decision support tools used to analyze patterns, trends, and relationships rely on data warehouses and newer platforms such as Hadoop that can handle less structured big data. Tools include online analytical processing (OLAP), statistics and modeling techniques, and text mining software. These software systems can analyze immense quantities of data to identify patterns, spot relationships, test hypotheses, and assess sentiments in online comments. Several software approaches are useful for simulating business events, forecasting the future, or determining optimal solutions to business problems given a set of constraints. The what-if analysis, for example, involves building a model based on relationships among variables that the user can change. Other tools in this category include goal seeking, optimization, and statistical forecasting. Artificial intelligence research contributes many important decision support tools, especially in the fields of robotics, expert systems, and neural nets. These all mimic some aspects of human intelligence, such as learning or expert decision making.

4. The organization's website is a key source of business intelligence with its own metrics. Total visits, number of unique visitors, traffic sources, page views, bounce rates, and other measures reveal how well the site is meeting its goals. For e-commerce and advertising, web analysts rely on display ads and search engine ads, with their own metrics and payment schemes. Web analytic software tracks and summarizes all the clickstream data.

5. Dashboards provide graphic displays that summarize key performance indicators (KPIs), and their content can be customized to meet the needs of individual users. These help reduce information overload and focus attention on the most important metrics. Portals are gateways that aggregate content on the screen and provide access to the individual's resources from a personalized website. Enterprise portals control access to the organization's resources, and the login determines access rights. Mashups also aggregate content and are similar to portals in concept. However, they use Web 2.0 technologies and standards that provide more flexibility to incorporate external resources of all kinds. The human element plays a critical role in decision making, and only people can decide which intelligence to draw on, which tools to use, and how to interpret the results.

KEY TERMS AND CONCEPTS

intelligent agents
predictive analytics
online analytical processing (OLAP)
market basket analysis
text mining
what-if analysis
goal seeking
optimization
forecasting
artificial intelligence (AI)
CAPTCHA
expert system
neural network
clickstream data
stickiness
dashboard
key performance indicators (KPIs)
portals
mashup
web feed

CHAPTER REVIEW QUESTIONS

7-1. How do you define business intelligence?

7-2. Can business intelligence support decisions based on both unstructured and structured information?

7-3. What are the most important sources of business intelligence inside the organization? What makes them useful?

7-4. What is the role of intelligent agents and big data in extracting business intelligence for firms?

7-5. How can managers use data mining techniques to analyze patterns, trends, and relationships? How does this lead to better data-driven decision making?

7-6. What is text mining?

7-7. What are examples of statistical techniques that managers can use to simulate business situations, optimize variables, and forecast sales or other figures?

7-8. What are examples of applications that draw on artificial intelligence for decision support?

7-9. Discuss how analytics can be used to improve the effectiveness of advertisements.

7-10. How do dashboards, portals, and mashups support decision making?

7-11. How does the human element affect decision making?

PROJECTS AND DISCUSSION QUESTIONS

7-12. Why do organizations use external data as a source of business intelligence? What are examples of sources of external data? How might retail giant Walmart use external data to make tactical-level decisions? How might its decision makers use external data to make strategic-level decisions?

7-13. How can an intelligent agent assist with a term paper? Visit your university library's home page to locate the "Search Databases" feature. If your library offers the "ABI/INFORM Complete" database, choose that and enter several keywords (for example, "social media in organizations") into the Basic Search dialog box. (If your library does not offer ABI/INFORM, try doing this exercise on a different database.) Review the results, then select "Refine Search" to select additional databases and/or specify additional search criteria. When you have the results you want, select the "Set Up Alert" option to schedule an alert. Prepare a brief report that describes the alert options that are available for your search. How frequently can you receive updates? How long can you receive updates? Are there options other than frequency and duration? Would you recommend using this intelligent agent to other students working on term papers?

7-14. First Class Salons maintains a company website to promote its chain of 12 health salons. The website includes links to information about its locations, special offers, and FAQs about its services, as well as "About Us" and "Contact Us" links. How can First Class Salons use information from its website to gain business intelligence? Consider the various visitor-related and content-related web metrics and suggest at least six specific metrics that First Class Salons would want to analyze. Prepare a brief report of your suggestions.

7-15. The Springfield Family Community Center has an outdoor pool that operates May through October. The director is interested to learn whether the center can afford the $57,000 cost of installing a pool-covering dome so patrons can swim year-round. It will also cost about $200 a month for power to keep the dome inflated for 6 months each year. How can the director use forecasting to evaluate the likelihood of selling sufficient tickets to pay for this improvement? Prepare a brief report to the director that explains forecasting. Be sure to include suggestions on both internal and external data that would be useful for this analysis.

7-16. Digital dashboards began to appear in the 1990s as organizations looked for ways to consolidate and display data to make it accessible and useful for busy executives. Visit www.dashboardinsight.com or www.dashboardsbyexample.com or search the Internet to learn more about digital dashboards. What is the relationship between digital dashboards and key performance indicators? Work in a small group with classmates to consider how a digital dashboard can be used by a Radio Shack or other electronics store manager. What specific daily performance indicators would he or she want to see on a digital dashboard? What design tips would you offer to the dashboard developer? As a group, create a hand-drawn sketch of a dashboard design for the Radio Shack manager.

APPLICATION EXERCISES

7-17. EXCEL APPLICATION: Analyzing Revenue and Expenses for City Hospital Seminars

Figure 7-25 shows the Excel spreadsheet that Bora uses to evaluate the variables relating to the hospital seminar series. She has asked you to use Excel to create a similar spreadsheet to conduct additional what-if and goal seeking analyses. You will need to use the following formulas:

Revenue

Registration Fees = Attendees per seminar × Registration fee × Seminars per year

Parking Fees = (Attendees per seminar / Average number attendees per car) × Seminars per year × Parking fee

Expenses

Speakers' Fees = Speaker's fee per session × Seminars per year

Tech Support = Tech support cost per session × Seminars per year

Marketing = Marketing cost per seminar × Seminars per year

Room Rental = Room rental per seminar × Seminars per year

What-If Questions

After answering each question, be sure to return the variables to their original values shown in Figure 7-25 before testing the impact of changing another one.

1. What is the impact on net profit if the average attendance per seminar increases to 45?
2. What is the impact on net profit if the average attendance drops to 35?
3. What is the impact on net profit if the parking fee is reduced to $3?
4. What is the impact on net profit if the speaker's fee increases to $550 per seminar?
5. What is the impact on net profit of increasing the marketing expense per seminar to $350, which increases average attendance per seminar to 50?
6. What is the impact on net profit of an increase in room rental per seminar to $300?
7. If Bora can negotiate a room rental fee of $160 per seminar, how much will net profit increase?
8. If technical support is included in the room rental per seminar, what is net profit?

FIGURE 7-25
The hospital seminar series data.

	A	B
1	**Hospital Seminars Revenues and Expenses**	
2		
3	**Revenue**	
4	Registration Fees	$16,800
5	Parking Fees	$ 960
6	Subtotal	$17,760
7		
8	**Expenses**	
9	Speakers' Fees	$ 6,000
10	Tech support	$ 1,800
11	Marketing	$ 3,000
12	Room rental	$ 3,000
13	Subtotal	$13,800
14		
15	**Net Profit/Loss**	$ 3,960
16		
17	**Variables**	
18	Speaker's Fee (per session)	$ 500
19	Tech Support cost per session	$ 150
20	Seminars per year	12
21	Registration fee	$ 35
22	Attendees per seminar	40
23	Room rental per seminar	$ 250
24	Parking fee	$ 5
25	Average # attendees per car	3
26	Marketing cost per seminar	$ 250

Goal Seeking Questions

1. Given the expenses and variables presented in the figure, how many attendees per seminar are required to generate a net profit of $5,500?
2. What parking fee results in a net profit of $4,150?
3. What registration fee per attendee results in a net profit of $5,750?

7-18. ACCESS APPLICATION: Marketing City Hospital Seminars

Colin is the assistant director of marketing at a hospital that conducts seminars on topics such as sports injuries, arthritis, hip and knee pain, knee replacement, and joint replacement. He is working on a marketing campaign for a new seminar on minimally invasive knee surgery that the hospital is planning to offer. Colin has asked you to help identify potential patients who may be interested in this seminar.

Download the City Hospital database, Ch07Ex02. Write a query that sorts registrants by the type of seminar they have attended. Include the session date as well as attendee information. Modify the query to identify registrants who attended a Knee Replacement seminar. Use the report wizard to create a report that lists the session dates and the names and phone numbers of those who have attended Knee Replacement seminars. This report serves as a "patient contact sheet" that hospital staff will use to call previous attendees to invite them to attend the new seminar. How many patients are listed on the report? Review the attendees table. Is there additional patient information the hospital could collect that may be useful for future marketing campaigns?

CASE STUDY #1

Cracking Fraud with Government's Big Data

Driving around Massachusetts, you might notice an RV frequently parked in your neighborhood with Montana license plates. Perhaps it is someone visiting from out west, but it might also be a Massachusetts resident who bought the RV online and registered it in Montana, which has no sales tax. The RV owner saves thousands of dollars, but Massachusetts loses the tax revenue. This kind of fraud is illegal, but difficult to catch. A Massachusetts state agency, for example, would not have access to Montana's vehicle registrations, and would not be able to match them up against Massachusetts' tax forms or employment records. The data are there, but they are not integrated into big data that can be analyzed.

The health care system is especially plagued by fraud, at both federal and state levels. Analysts estimate that fraud and abuse cost $125 to $175 billion each year, but problems often goes unnoticed because they are difficult to identify. Only 3% to 5% of fraudulent cases are discovered, and the detection often happens so late that the funds cannot be recovered.

Predictive analytics are the tools that can spot suspicious activity and unusual patterns. The potential to reduce government waste and fraud in general is enormous; combined with big data, these tools can arm investigators with ways to track fraudulent billing patterns buried in millions of legitimate claims, picking out unusual trends that no human being working alone could ever see.

Health Care Service Corp. (HCSC), for example, implemented a fraud detection system, and it paid off almost immediately. An allergist in Illinois was submitting fraudulent bills, but the individual amounts were never high enough to trigger any suspicion. Something was amiss, however, and the analysts for the insurance company were able to compare what other allergists were charging for the same procedures. The results helped uncover an $800,000 scam.

With access to big data from multiple sources, fraud detection systems can spot a large variety of suspicious activities that need investigation, particularly if data can be drawn from state and federal databases. In health care, for example, such systems first start with rules that flag unusual behavior in near real time, such as when a provider bills for many services in a short time window, or when a person enrolls in Medicaid in more than one state.

Predictive analytics can also learn from the data to build more sophisticated models, especially as the tools have access to more and more data. Fraud rings commit a large number of the scams in health care and government services, and the software can examine relationships. Analyzing the linkages can reap huge benefits and uncover large criminal gangs, even if each individual transaction on its own does not trigger any flags. One ring caught in 2012, for instance, included 107 providers who billed Medicare for over $400 million in services that were not performed.

Near real-time analysis is especially important because of the need to spot fraud before any claim is paid. It is much easier to deny payment than to recover funds that have already been paid out. The time window is short, however, given pressure on payers to reimburse quickly. With big data analytical tools, fraud detection systems can operate quickly enough to catch fishy claims before they are paid.

For the 50 states with different information systems, and all the counties that maintain their own records, the challenge of reducing fraud is daunting. The goal is to create a big data view of citizens, one that would, for instance, inform a state agency if someone who is receiving benefits purchases a luxury vehicle in a state with no sales tax. With states and counties struggling with budget woes, the drive to catch fraud is strong.

Will "big data" become "big brother?" Privacy advocates voice concerns over the growing access to big data across government agencies, particularly as ways are found to integrate the data to paint a meaningful and comprehensive picture of citizen's financial transactions and government benefits. Disclosures about the extent of the government's electronic data gathering for national security has intensified those concerns. Balancing privacy and the need to reduce fraud will be particularly important.

Discussion Questions

7-19. What are some ways that data mining could be used to detect fraud in health insurance claims?

7-20. How could private insurance companies and public government agencies collaborate to combat insurance fraud?

7-21. Identify some privacy concerns that can arise from the government's electronic data gathering of citizen's financial transactions.

7-22. What business processes are necessary to complement the IS component of data mining?

Sources: Efficient use of big data could reduce instances of healthcare fraud. (2013). GovPlace, http://www.govplace.com/2012/05/efficient-use-of-big-data-could-reduce-instances-of-healthca00000000re-fraud/, accessed April 2, 2013.
Hanson, W. (March 19, 2013). Can big data crack fraud? Government Technology, http://www.govtech.com/e-government/Can-Big-Data-Crack-Fraud.html, accessed April 2, 2013.
Horowitz, B. T. (December 28, 2012). Big data can fight fraud in health insurance exchanges: Operat. eWeek, http://www.eweek.com/enterprise-apps/big-data-can-fight-fraud-in-health-insurance-exchanges-operal, accessed April 2, 2013.
Schrieber, R. (2013). Examine the broader context to identify healthcare fraud. *Managed Healthcare Executive, 23*(2), 44–45.
Thorpe, N., Deslich, S., Sikula, S., & Coustasse, A. (2012). Combating Medicare Fraud: A Struggling Work in Progress. *Franklin Business & Law Journal, 2012*(4), 95–107.

CASE STUDY #2

TV and Twitter: How Nielsen Rates Programs with "Social TV"

To understand the audience for TV shows, TV ratings giant Nielsen relies on electronic "People Meters" placed into a representative sample of homes throughout the United States to track viewing patterns. In half the homes, Nielsen installs "cross-platform" people meters to detect TV viewing on computers or mobiles, and also track web traffic. The company also asks viewers to fill out paper-and-pencil diaries about their TV viewing habits.

Founded in 1936 when there were very few televisions in the country, Nielsen grew into the ratings giant that can make or break any new program, or any TV producer's career. The ratings affect not only the show's survival, but the cost of the ads that appear during the show. Super Bowl ads, for instance, are most expensive of all because the Super Bowl has the largest audience of any show on TV.

In the age of Twitter, however, TV viewing is becoming a social experience that involves many more people than those in a single home. Viewers, especially those in the coveted 18 to 34 age group, share their thoughts in real time as they watch a program, trashing the actors, praising the costumes, or mocking the script. Nielsen teamed up with SocialGuide, a social media metrics company that measures the engagement of the TV audience by tracking tweets and posting results on real-time dashboards for the company's clients.

SocialGuide's software picks up tweets related to a particular TV show by using relevant keywords—actors' names, characters, incidents, and other tracking tools. It relies also on the hashtags that identify the topic in many tweets, such as #bigbangtheory, #projectrunway, or #gameofthrones. The dashboard shows relevant statistics by time period, such as the number of related tweets with positive or negative spin, and it can scroll the tweets as they flash by. Because the dashboard is showing real-time data, clients can not only see gross statistics such as overall number of viewers. They can also see how viewers are reacting to particular scenes or characters as they appear. The software tracks over 30,000 programs, so it can generate comprehensive comparison ratings for "social TV" viewers, their demographics, and their preferences.

The use of Twitter feeds to analyze social TV patterns adds a great deal to Nielsen's capabilities. For example, Nielsen's set-top boxes do not easily capture who in the family is actually watching—even if the set is just turned on and no one at all is watching. Many people leave the TV on so they can record shows they like, though they may not have time to actually view them later. Even if they do, they may fast-forward through the ads. The set-top boxes are also unable to assess viewer attitudes during the show or the ads.

Twitter feeds also have disadvantages as an audience rating tool, however. The tweets are not generated from a representative sample, for instance, so their content is biased toward a certain population of viewers. Those who don't use Twitter are not in the sample, and the feeds may be overwhelmed by a small group of frequent tweeters who are loud and vocal.

Despite the drawbacks, research confirms a relationship between Twitter activity and TV ratings measured by Nielsen's other tools. For example, premiere episodes that generated an 8.5% increase in Twitter volume showed a 1% increase in TV ratings for viewers in the 18–34 age group. The relationship was weaker for other age groups, probably because fewer people outside that group use Twitter.

The relationship between Twitter volume and TV ratings becomes stronger as the season continues, and is highest for the season finale. This suggests that Twitter metrics are not just reflecting a show's appeal. The chatter may be creating TV buzz that draws more viewers into the social TV experience.

Social TV may also be drawing people back to viewing shows live, rather than recording them. If Twitter volume is high during the show, it means that people have points to make in real time. For instance, they may prefer to weigh in on American Idol singers right after their song, rather than wait until the next day. This trend may mean more live viewers for the ads, as well.

While Nielsen will not drop its careful sampling techniques and set-top boxes, the company is leading the way toward new ways to learn about social TV. The networks may also be looking for ways to make viewers want to share thoughts in real time so they will see some ads.

Discussion Questions

7-23. What potential value does Nielsen intend to add to their ratings by data mining Twitter to analyze social TV patterns?

7-24. What are the drawbacks of using Twitter as a rating tool? Do these disadvantages compromise the value of the Nielsen ratings?

7-25. Should Nielsen stop using its earlier techniques for TV ratings? Why or why not?

7-26. If Nielsen extended their data mining of social media to include Facebook as well as Twitter, what differences might they expect in the audience being analyzed? Would this analysis have any value to the networks? Why or why not?

Sources: Humphrey, M. (2012). Nielsen acquires social TV metrics company SocialGuide. *Forbes*, http://www.forbes.com/sites/michaelhumphrey/2012/11/12/nielsen-acquires-social-tv-metrics-company-socialguide/, accessed April 10, 2013.
McLellan, M. (2013). Nielsen Media Research, Inc. Hoover's Online Database, http://subscriber.hoovers.com.proxy3.library.jhu.edu/H/company360/overview.html?companyId=57187000000000, accessed April 13, 2013.
New study confirms correlation between Twitter and TV ratings. (2013, Mar 20). *PR Newswire*, http://search.proquest.com/docview/1317970375?accountid=11752, accessed June 20, 2013.
Nielsen Website, www.nielsen.com, accessed April 10, 2013.
Pomerantz, D. (2013). Can Twitter Save Live TV? Forbes.Com, 9. SocialGuide Website, www.socialguide.com, accessed April 10, 2013.

E-PROJECT 1 Detecting Suspicious Activity in Insurance Claims

Detecting unusual patterns in drug prescriptions is the focus of this e-project. To begin, download the Excel file called Ch07_MedicalCharges.

The worksheet contains columns showing a sample of hypothetical prescription drug claims over a period of years.

7-27. Create a pivot table and chart to show the total amounts paid by year for this pharmacy, by dragging Year to the Axis Fields box and Amount to the Values box. Be sure you are looking at the sum of Amounts in your chart. Which year had the highest sales for prescription drugs?

7-28. Change the pivot table to show total sales by month by removing Year from the Axis Fields and dragging Month to that box. During which month of the year does this pharmacy tend to sell the most prescription drugs?

7-29. Remove Month and put Prescriber ID in the Axis Fields box. Which prescriber generates the most income for this pharmacy?

7-30. Remove PrescriberID and put PatientID in the Axis Fields box. Which patient generates the most income for the pharmacy?

7-31. Let's take a closer look at this patient by filtering the records. Click on PatientID in the PivotTable Field List and uncheck all boxes except for this patient. Drag Year under PatientID in the Axis Fields box so you can see how this person's spending patterns have changed. Which year shows the most spending?

7-32. Let's see who is prescribing for this patient. Remove Year from the Axis Fields box and drag PrescriberID to the box. Which prescriber has the highest spending total?

7-33. Now, let's see what is being prescribed. Drag DrugName to the Axis Fields box under PrescriberID. What might you conclude from this chart?

E-PROJECT 2 Analyzing Nielsen TV Ratings with Excel

In this e-project, you will explore TV ratings and analyze them with Excel. Download the Excel file called Ch07_NielsenRatings. This file contains ratings for popular network programs for two separate weeks in 2013 (http://www.nielsen.com/us/en/top10s.html). The rating represents the percent of U.S. households that were watching that channel at the time (of those whose TV was turned on).

7-34. Calculate three new columns:
- a. Percent change (up or down) in number of viewers from the March 25 data to the April 1 data.
- b. Percent change (up or down) in rating.
- c. Absolute change in the number of viewers.

7-35. Answer the following questions:
- a. Which show gains the largest number of viewers from March 25 to April 1?
- b. Which show is the biggest loser from March 25 to April 1, in terms of change in ratings?
- c. Compute the total viewers for these shows for March 25 and for April 1. How many total viewers watched one of the TV shows in this list during the week of March 25.
- d. What is the percent change in total viewers for the shows in this list from March 25 to April 1?