

10. Suppose you are selling the data warehouse idea to your users. How would you define multidimensional data analysis for them? How would you explain its advantages to them?
11. Before making a commitment, the data warehousing project group has invited you to provide an OLAP overview. The group's members are particularly concerned about the OLAP client/server architecture requirements and how OLAP will fit the existing environment. Your job is to explain to them the main OLAP client/server components and architectures.
12. One of your vendors recommends using an MDBMS. How would you explain this recommendation to your project leader?
13. The project group is ready to make a final decision, choosing between ROLAP and MOLAP. What should be the basis for this decision? Why?
14. The data warehouse project is in the design phase. Explain to your fellow designers how you would use a star schema in the design.
15. Briefly discuss the decision support architectural styles and their evolution. What major technologies influenced this evolution?
16. What is OLAP, and what are its main characteristics?
17. Explain ROLAP and give the reasons you would recommend its use in the relational database environment.
18. Explain the use of facts, dimensions, and attributes in the star schema.
19. Explain multidimensional cubes and describe how the slice-and-dice technique fits into this model.
20. In the star schema context, what are attribute hierarchies and aggregation levels, and what is their purpose?
21. Discuss the most common performance improvement techniques used in star schemas.
22. Explain some of the most important issues in data warehouse implementation.
23. What is data mining, and how does it differ from traditional decision support tools?
24. How does data mining work? Discuss the different phases in the data mining process.

## P R O B L E M S



### O N L I N E   C O N T E N T

The databases used for this problem set are found in the Student Online Companion for this book. These databases are stored in Microsoft Access 2000 format. The databases, named **Ch13\_P1.mdb**, **Ch13\_P3.mdb**, and **Ch13\_P4.mdb**, contain the data for Problems 1, 3, and 4, respectively. The data for Problem 2 are stored in Microsoft Excel format in the Student Online Companion for this book. The spreadsheet filename is **Ch13\_P2.xls**.

1. The university computer lab's director keeps track of lab usage, measured by the number of students using the lab. This particular function is important for budgeting purposes. The computer lab director assigns you the task of developing a data warehouse in which to keep track of the lab usage statistics. The main requirements for this database are to:
  - Show the total number of users by different time periods.
  - Show usage numbers by time period, by major, and by student classification.
  - Compare usage for different majors and different semesters.

Use the **Ch13\_P1.mdb** database, which includes the following tables:

- USELOG contains the student lab access data.
- STUDENT is a dimension table containing student data.

Given the three bulleted requirements and using the **Ch13\_P1.mdb** data, complete Problems 1a–1g.

- Define the main facts to be analyzed. (*Hint:* These facts become the source for the design of the fact table.)
  - Define and describe the appropriate dimensions. (*Hint:* These dimensions become the source for the design of the dimension tables.)
  - Draw the lab usage star schema, using the fact and dimension structures you defined in Problems 1a and 1b.
  - Define the attributes for each of the dimensions in Problem 1b.
  - Recommend the appropriate attribute hierarchies.
  - Implement your data warehouse design, using the star schema you created in Problem 1c and the attributes you defined in Problem 1d.
  - Create the reports that will meet the requirements listed in this problem's introduction.
2. Ms. Victoria Ephanor manages a small product distribution company. Because the business is growing fast, Ms. Ephanor recognizes that it is time to manage the vast information pool to help guide the accelerating growth. Ms. Ephanor, who is familiar with spreadsheet software, currently employs a small sales force of four people. She asks you to develop a data warehouse application prototype that will enable her to study sales figures by year, region, salesperson, and product. (This prototype is to be used as the basis for a future data warehouse database.)

Using the data supplied in the **Ch13\_P2.xls** file, complete the following seven problems:

- Identify the appropriate fact table components.
- Identify the appropriate dimension tables.
- Draw a star schema diagram for this data warehouse.
- Identify the attributes for the dimension tables that will be required to solve this problem.
- Using a Microsoft Excel spreadsheet (or any other spreadsheet capable of producing pivot tables), generate a pivot table to show the sales by product and by region. The end user must be able to specify the display of sales for any given year. (The sample output is shown in the first pivot table in Figure P13.2E.)
- Using Problem 2e as your base, add a second pivot table (see Figure P13.2E) to show the sales by salesperson and by region. The end user must be able to specify sales for a given year or for all years and for a given product or for all products.

**FIGURE P13.2E** Using a pivot table

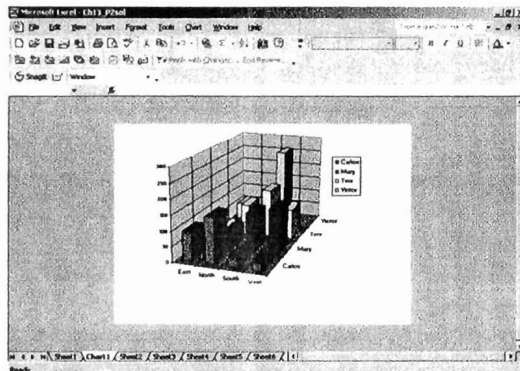
The screenshot shows two pivot tables in Microsoft Excel. The first pivot table is titled 'PivotTable1' and shows sales by product and region for the year 2005. The second pivot table is titled 'PivotTable2' and shows sales by salesperson and region for the year 2005.

Year	2005				
Sum of Values	Region				
Product	East	North	South	West	Total
1. Balls	55	70	100	225	
2. Banners	30	90	100	190	
3. Pencils	45	80	75	190	
4. Postcards	100	100	100	300	
5. Total	330	340	375	1045	

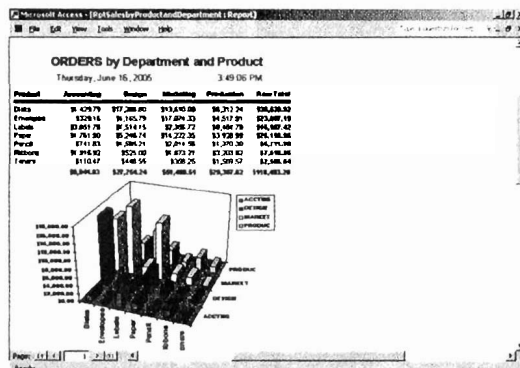
  

Year	2005				
Sum of Values	Region				
Salesperson	East	North	South	West	Total
1. Carlos	95	150	30	25	300
2. Henry	80	125	145	350	
3. Terry	12	100	180	100	372
4. Victor	55	30	250	335	
5. Total	242	305	475	270	1392

**FIGURE P13.2G** 3-D bar graph showing the relationships among agent, product, and region



**FIGURE P13.3** Crosstab report: orders by product and department



g. Create a 3-D bar graph to show sales by salesperson, by product, and by region. (See the sample output in Figure P13.2G.)

- Mr. David Suker, the inventory manager for a marketing research company, is interested in studying the use of supplies within the different company departments. Mr. Suker has heard that his friend, Ms. Ephanor, has developed a small spreadsheet-based data warehouse model (see Problem 2) that she uses to analyze sales data. Mr. Suker is interested in developing a small data warehouse model like Ms. Ephanor's so he can analyze orders by department and by product. He will use Microsoft Access as the data warehouse DBMS and Microsoft Excel as the analysis tool.
  - Develop the order star schema.
  - Identify the appropriate dimensions attributes.
  - Identify the attribute hierarchies required to support the model.
  - Develop a crosstab report (in Microsoft Access), using a 3-D bar graph to show orders by product and by department. (The sample output is shown in Figure P13.3.)

- Develop the order star schema.
- Identify the appropriate dimensions attributes.
- Identify the attribute hierarchies required to support the model.
- Develop a crosstab report (in Microsoft Access), using a 3-D bar graph to show orders by product and by department. (The sample output is shown in Figure P13.3.)

- ROBCOR, whose sample data are contained in the database named **Ch13\_P4.mdb**, provides "on-demand" aviation charters, using a mix of different aircraft and aircraft types. Because ROBCOR has grown rapidly, its owner has hired you to be its first database manager. (The company's database, developed by an outside consulting team, already has a charter database in place to help manage all of its operations.) Your first critical assignment is to develop a decision support system to analyze the

charter data. (Review Problems 30–34 in Chapter 3, The Relational Database Model, in which the operations have been described.) The charter operations manager wants to be able to analyze charter data such as cost, hours flown, fuel used, and revenue. She would also like to be able to drill down by pilot, type of airplane, and time periods.

Given those requirements, complete the following:

- Create a star schema for the charter data.
- Define the dimensions and attributes for the charter operation's star schema
- Define the necessary attribute hierarchies.
- Implement the data warehouse design, using the design components you developed in Problems 4a–4c.
- Generate the reports that will illustrate that your data warehouse meets the specified information requirements.

Using the data provided in the SaleCo snowflake schema in Figure 13.25, solve the following problems.