



## Review Questions

- 4.1 Summarize the premise of this chapter.
- 4.2 When you receive a set of tables, what steps should you take to assess their structure and content?
- 4.3 Show SQL statements to count the number of rows and to list the top 15 rows of the RETAIL\_ORDER table.
- 4.4 Suppose you receive the following two tables:

**DEPARTMENT** (DepartmentName, BudgetCode)

**EMPLOYEE** (EmployeeNumber, EmployeeLastName, EmployeeFirstName, Email, DepartmentName)

and you conclude that EMPLOYEE.DepartmentName is a foreign key to DEPARTMENT.DepartmentName. Show SQL for determining whether the following referential integrity constraint has been enforced:

**DepartmentName in EMPLOYEE must exist in DepartmentName in DEPARTMENT**

- 4.5 Summarize how database design principles differ with regards to the design of updatable databases and the design of read-only databases.
- 4.6 Describe two advantages of normalized tables.
- 4.7 Why do we say that data duplication is only reduced? Why is it not eliminated?
- 4.8 If data duplication is only reduced, how can we say that the possibility of data inconsistencies has been eliminated?
- 4.9 Describe two disadvantages of normalized tables.
- 4.10 Suppose you are given the table:

**EMPLOYEE\_DEPARTMENT** (EmployeeNumber, EmployeeLastName, EmployeeFirstName, Email, DepartmentName, BudgetCode)

and you wish to transform this table into the two tables:


**DEPARTMENT** (DepartmentName, BudgetCode)

**EMPLOYEE** (EmployeeNumber, EmployeeLastName, EmployeeFirstName, Email, DepartmentName)

Write the SQL statements needed for filling the EMPLOYEE and DEPARTMENT tables with data from EMPLOYEE\_DEPARTMENT.

- 4.11 Summarize the reasons explained in this chapter for not placing ZIP code values into BCNF.
- 4.12 Describe a situation, other than the one for ZIP codes, in which one would choose not to place tables into BCNF. Justify your decision not to use BCNF.

- 4.13 According to this text, under what situations should you choose not to remove multi-valued dependencies from a relation?
- 4.14 Compare the difficulty of writing subqueries and joins with the difficulty of dealing with anomalies caused by multivalued dependencies.
- 4.15 Describe three uses for a read-only database.
- 4.16 How does the fact that a read-only database is never updated influence the reasons for normalization?
- 4.17 For read-only databases, how persuasive is the argument that normalization reduces file space?
- 4.18 What is denormalization?
- 4.19 Suppose you are given the DEPARTMENT and EMPLOYEE tables in Review Question 4.10 and asked to denormalize them into the EMPLOYEE\_DEPARTMENT relation. Show the design of the EMPLOYEE\_DEPARTMENT relation. Write an SQL statement to fill this table with data.
- 4.20 Summarize the reasons for creating customized duplicated tables.
- 4.21 Why are customized duplicated tables not used for updatable databases?
- 4.22 List four common design problems when creating databases from existing data.
- 4.23 Give an example of a multivalued, multicolumn table other than one discussed in this chapter.
- 4.24 Explain the problems caused by the multivalued, multicolumn table in your example in Review Question 4.23.
- 4.25 Show how to represent the relation in your answer to Review Question 4.23 with two tables.
- 4.26 Show how the tables in your answer to Review Question 4.25 solve the problems you identified in Review Question 4.24.
- 4.27 Explain the following statement: "The multivalued, multicolumn problem is just another form of multivalued dependency." Show how this is so.
- 4.28 Explain ways in which inconsistent values arise.
- 4.29 Why are inconsistent values in foreign keys particularly troublesome?
- 4.30 Describe two ways to identify inconsistent values. Are these techniques certain to find all inconsistent values? What other step can be taken?
- 4.31 What is a null value?
- 4.32 How does a null value differ from a blank value?
- 4.33 What are three interpretations of null values? Use an example in your answer that is different from the one in this book.
- 4.34 Show SQL for determining the number of null values in the column EmployeeFirstName of the table EMPLOYEE.
- 4.35 Describe the general-purpose remarks column problem.
- 4.36 Give an example in which the general-purpose remarks column makes it difficult to obtain values for a foreign key.
- 4.37 Give an example in which the general-purpose remarks column causes difficulties when multiple values are stored in the same column. How is this problem solved?
- 4.38 Why should one be wary of general-purpose remarks columns?



## Project Questions

The Elliot Bay Sports Club owns and operates three sports club facilities in Houston, Texas. Each facility has a large selection of modern exercise equipment, weight rooms, and rooms for yoga and other exercise classes. Elliot Bay offers 3-month and 1-year memberships. Members can use the facilities at any of the three club locations.


Elliot Bay maintains a roster of personal trainers who operate as independent consultants. Approved trainers can schedule appointments with clients at Elliot Bay facilities, as long as their client is a member of the club. Trainers also teach yoga, Pilates, and other classes. Answer the following questions, assuming you have been provided the following three tables of data (PT stands for personal trainer):

**PT\_SESSION** (Trainer, Phone, Email, Fee, ClientLastName, ClientFirstName, ClientPhone, ClientEmail, Date, Time)

**CLUB\_MEMBERSHIP** (ClientNumber, ClientLastName, ClientFirstName, ClientPhone, ClientEmail, MembershipType, EndingDate, Street, City, State, Zip)

**CLASS** (ClassName, Trainer, StartDate, EndDate, Time, DayOfWeek, Cost)

- 4.39 Identify possible multivalued dependencies in these tables.
- 4.40 Identify possible functional dependencies in these tables.
- 4.41 Determine whether each table is either in BCNF or in 4NF. State your assumptions.
- 4.42 Modify each of these tables so that every table is in BCNF and 4NF. Use the assumptions you made in your answer to question 4.41.
- 4.43 Using these tables and your assumptions, recommend a design for an updatable database.
- 4.44 Add a table to your answer to question 4.43 that would allow Elliot Bay to assign members to particular classes. Include an AmountPaid column in your new table.
- 4.45 Recommend a design for a read-only database that would support the following needs:
  - A. Enable trainers to ensure that their clients are members of the club.
  - B. Enable the club to assess the popularity of various trainers.
  - C. Enable the trainers to determine if they are assisting the same client.
  - D. Enable class instructors to determine if the attendees to their classes have paid.



## Case Questions

### Marcia's Dry Cleaning Case Questions

Marcia Wilson, the owner of Marcia's Dry Cleaning, is in the process of creating databases to support the operation and management of her business. For the past year, she and her staff have been using a cash register system that collects the following data:

**SALE** (InvoiceNumber, DateIn, DateOut, Total, Phone, FirstName, LastName)



Unfortunately, during rush times, not all of the data are entered, and there are many null values in Phone, FirstName, and LastName. In some cases, all three are null; in other cases, one or two are null. InvoiceNumber, DateIn, and Total are never null. DateOut has a few null values. Also, occasionally during a rush, phone number and name data have been entered incorrectly. To help create her database, Marcia purchased a mailing list from a local business bureau. The mailing list includes the following data:

**HOUSEHOLD (Phone, FirstName, LastName, Street, City, State, Zip, Apartment)**

In some cases, a phone number has multiple names. The primary key is thus the composite (Phone, FirstName, LastName). There are no null values in Phone, FirstName, and LastName, but there are some null values in the address data.

There are many names in SALE that are not in HOUSEHOLD, and there are many names in HOUSEHOLD that are not in SALE.

- A. Design an updatable database for storing customer and sales data. Explain how to deal with the problems of missing data. Explain how to deal with the problems of incorrect phone and name data.
- B. Design a read-only database for storing customer and sales data. Explain how to deal with the problems of missing data. Explain how to deal with the problems of incorrect phone and name data.

The Queen  
Anne  
Curiosity  
Shop



The Queen Anne Curiosity Shop questions in Chapter 3 asked you to create a set of relations to organize and link The Queen Anne Curiosity Shop typical sales data shown in Figure 3-33 and the typical purchase data shown in Figure 3-34. The set of relations may look like the following:

**CUSTOMER (CustomerID, LastName, FirstName, Phone, Email)**

**SALE (SaleID, CustomerID, InvoiceDate, PreTaxTotal, Tax, Total)**

**SALE\_ITEM (SaleID, SaleItemID, PurchaseID, SalePrice)**

**PURCHASE (PurchaseID, PurchaseItem, PurchasePrice, PurchaseDate, VendorID)**

**VENDOR (VendorID, Vendor, Phone)**

Use these relations and the data in Figures 3-33 and 3-34 to answer the following questions.

- A. Follow the procedure shown in Figure 4-1 to assess these data.
  1. List all functional dependencies.
  2. List any multivalued dependencies.
  3. List all candidate keys.
  4. List all primary keys.
  5. List all foreign keys.
  6. State any assumptions you make as you list these components.
- B. List questions you would ask the owners of The Queen Anne Curiosity Shop to verify your assumptions.
- C. If there are any multivalued dependencies, create the tables needed to eliminate these dependencies.
- D. Do these data have the multivalued, multicolumn problem? If so, how will you deal with it?
- E. Do these data have the inconsistent data problem? If so, how will you deal with it?
- F. Do these data have a null value data problem? If so, how will you deal with it?
- G. Do these data have the general-purpose remarks problem? If so, how will you deal with it?

