**Chapter 14**

**Electronic Commerce, Databases, and Personal Privacy**

**A Guide to this Instructor’s Manual:**

We have designed this Instructor’s Manual to supplement and enhance your teaching experience through classroom activities and a cohesive chapter summary.

This document is organized chronologically, using the same headings that you see in the textbook. Under the headings you will find: Lecture Notes that summarize the section, Teaching Tips, Class Discussion Topics, and Additional Projects and Resources. Pay special attention to teaching tips and activities geared toward quizzing your students and enhancing their critical thinking skills.

In addition to this Instructor’s Manual, our Instructor’s Resources also contain PowerPoint Presentations, Test Banks, and other supplements to aid in your teaching experience.

|  |
| --- |
| **At a Glance** |

**Instructor’s Manual Table of Contents**

* Overview
* Learning Objectives
* Teaching Tips
* Class Discussion Topics
* Additional Projects
* Additional Resources
* Key Terms

|  |
| --- |
| Lecture Notes |

**Overview**

Chapter 14 examines electronic commerce, a field that has grown in importance tremendously over the past 15 years. The chapter introduces different purposes an organization might have for an ecommerce site. It discusses the positives and negatives of creating an ecommerce website. The chapter explains the design process for deciding what an ecommerce site will contain, and how online transactions must be managed. The chapter then discusses a key system required by most or all ecommerce sites: databases. It describes the purpose of a database management system, and how SQL is used to retrieve data from tables within a database. The chapter includes a discussion of database integrity, and how that is maintained. It ends with a section on personal privacy in the era of data mining and massive databases.

# **Learning Objectives**

* Explain the issues, pros, and cons of whether to create an ecommerce website
* List the components of online transactions, and explain what is involved in each step
* List one other ecommerce model besides the online retailer selling to the general public and explain how it works
* List one other electronic payment system besides traditional credit cards and explain how it works
* Describe the relational database model
* Frame simple queries in SQL to retrieve information from one or more tables in a relational database
* Define data mining
* List three sources data brokers used to collect data on individuals
* Give an example of data science used "for the greater good"

# **Teaching Tips**

**14.1 Introduction**

1. In today’s society, if you own just about any type of business, you need a web presence, such as advertising, providing follow-up service to an offline transaction, online transactions that are not retail sales, business-to-business sales, and business-to-consumer sales. Use examples.
2. Inform students that this chapter focuses mostly on the idea of business to consumer sales, and how it fits into the world of computer science, specifically *data science*.

**14.2 Ecommerce**

1. Introduce the idea of ecommerce and emphasize that modern ecommerce often integrates advertising, sales, credit-checking, shipping, billing, and inventory.
2. Business owners need to consider many issues when deciding how to create an ecommerce site. Students should know that this, like any business decision, must be attached to one or more goals. Does the business want to broaden the customer base? Better serve the existing customer base? Better integrate their internal departments?

|  |  |
| --- | --- |
| ***Teaching Tip*** | Refer students to these websites for more information and examples of ecommerce sites, <http://www.ecommerce-blog.org/>, <http://kaleazy.com/top-25-best-e-commerce-sites-of-2011/> and encourage them to look for more lists of good and bad sites and practices, as well. |

1. Go over some of the risks associated with developing an ecommerce site. Will customers simply move from in-store to online? Can you offer something unique? Discuss with the students in detail.
2. Decisions around development and maintenance of the site are critical. Discuss decisions such as personnel, hardware, and software. Will the site be developed and maintained in-house or by outsourcing? Introduce the term **ASP (Application Service Provider)**.
3. Discuss the steps in an online transaction, and the fundamental need for convenience. A site must attract customers, keep them on the site, and set up a convenient and secure way to complete a purchase.
4. To attract customers, an ecommerce site must advertise and have a useful **domain name**. Introduce the term **ICANN (Internet Corporation for Assigned Names and Numbers)**, the organization that gives out domain names, seek to be a sponsored link for a web search engine, or place information on a **portal** site related to your business.
5. Discuss search engines and how one may be able to help put their results closer to the top of search results. Talk about sponsored results.
6. Personalize the web shopping experience, either by having user’s register and log in, or by leaving information on the user’s computer. Introduce the term **cookie** and describe what it can and cannot do.
7. Sites must provide authentication that the site is what it claims to be, and must provide encryption so that sensitive user data are transmitted securely.

|  |  |
| --- | --- |
| ***Teaching Tip*** | Plan a lab activity while talking about ecommerce site design. Ask each student to pick an ecommerce site. While different features are discussed, have the students look for those features on their site. Discuss as a class how site organization differs, what is effective, what is not. |

1. A website should be designed to reflect a business’ personality and how its customers shop. Introduce the terms **site map** and **navigation bar**, and talk about how electronic “shopping carts” and checkout forms work.
2. Go over the different kinds of information an ecommerce website needs to have on it. Talk about the issues of working with multiple browsers and plug-ins.
3. Discuss legacy code and why it can prove to be a problem at times.
4. Introduce the term **middleware**, which is software that allows separate, existing programs to communicate and work together seamlessly.
5. Stress the importance of planning for things to go wrong, from machine failures to natural disasters to hackers: having a **disaster recovery strategy**.
6. Note that business owners can be online in other ways, such as through social networking sites or through other sites such as eBay, Craigslist, and Groupon. Introduce the term peer-to-peer (collaborative consumption) economy.
7. New forms of payment are arriving quickly as well. Talk briefly about Apple Pay, PayPal, and Bitcoin.

**Quick Quiz 1**

1. (True or False) An ecommerce site is always designed to facilitate retail sales.

Answer: False

1. (True or False) A website designed to work with a particular browser will automatically work well with other browsers.

Answer: False

1. \_\_\_\_\_\_\_\_\_\_\_\_\_\_ is a kind of software that helps connect existing programs with a website.

Answer: Middleware

1. List at least three ways an ecommerce site can attract customers.

Answer: Any three of the following: advertising, direct communication with existing customers, easy-to-guess domain names, being a sponsored link, or a popular link, on a web search engine, getting a link from a portal.

1. Name one alternative payment method discussed here.

Answer: Any one of the following: Apple Pay, PayPal, Bitcoin

**14.3 Databases**

1. Computers must organize data into sensible pieces. Databases use certain terms to describe various data. Introduce the terms **field**, **record**, **data file**, and **database**. Emphasize that records are often information about a specific real-world person or object, whereas fields are categories of information we might have about a particular kind of person or object.
2. Introduce the terms **database management system (DBMS)** and the **relational database model**. Use an example like the Rugs-For-You one from the book to illustrate these ideas. Introduce relational database terminology: **tuples**, **attributes**, and **primary keys** and explain how they work together.
3. Talk about the query model for accessing data in a DBMS, and introduce the term **SQL (Structured Query Language)**, along with examples of its use. Given an example table, ask students for queries to make, and then turn those into SQL commands. Talk about the power and complexity of working with multiple tables of data.
4. Introduce the term **foreign key**.
5. Introduce the three main database operations: **project**, **restrict**, and **join**. If possible, work with an actual database system, and show the queries in action.
6. When database information can be modified, care must be taken to keep the data consistent. Use inventory as an easy example: if an item is sold, it must be removed from the database properly so that another customer does not try to purchase it. Introduce the terms **entity integrity**, **data integrity**, and **referential integrity**.
7. Other issues include fast performance, security, and privacy. Introduce the term **distributed database**, which allows the physical data to reside oat separate and independent locations that are networked.

**14.4 Data Science**

1. Talk about how much data exist in the world today (in 2017, it was 25 zettabytes – or 25 trillion gigabytes).
2. Talk about *Big Data* and *data analysis*.
3. Introduce the term **data science**.
4. Talk about a *data warehouse*.
5. Discuss **data mining** and define data warehouse, big data, business intelligence, and business analytics. Discuss how data mining seeks to uncover previously hidden patterns in a massive database.
6. Discuss the use of a decision tree for data mining.
7. Introduce the term **data brokers**, companies that collect data on a large population and then sell that data to clients.

**Quick Quiz 2**

1. List the three main relational database operations.

Answer: projection, restriction, and join

1. (True or False) A record is made up of multiple data files.

Answer: False

1. (True or False) SQL allows users to retrieve and compile information from a database.

Answer: True

1. In an actual SQL command, how is the restriction operation typically indicated?

Answer: By the word “WHERE”

1. The \_\_\_\_\_\_\_ rule says that no primary key value, or no component of a composite primary key value, can be missing (“null”) in a tuple.

Answer: entity integrity

# **Class Discussion Topics**

1. When might a business decide **not** to have a web presence? What conditions might cause a business to avoid a web presence, and what changes might cause the business to reconsider?
2. What components necessary for an ecommerce business-to-consumer site would a “bricks and mortar” business already have, and what would be entirely new?
3. Given the kinds of information databases can contain, and how SQL queries may be used to retrieve the data, what are some restrictions a business might put on which employees can access what information in a business’ database?
4. Discuss the role of Big Data in day to day life. Are Data Brokers ethical? Should our personal information, purchasing history, and various proclivities be for sale?

# **Additional Projects**

1. Suppose you are hired by a family practice clinic to create a website for their use. List the purpose or purposes you would expect such a website to have. What information would be critical to have on the website? What decisions would you recommend about designing and maintaining the website? What kind of transactions would you expect to have? What concerns might you expect a “customer” of such a website to have?
2. Suppose you were going to design a database to hold information about student organizations on your campus. Decide what tables you might want to have, and what fields each table would have. What kind of queries could your database allow?
3. Create a dataset and demonstrate how it would fit into a decision tree.

# **Additional Resources**

1. A blog about ecommerce design about the “Best Worst eCommerce Site Ever”: <http://blog.brandlabs.us/just-for-fun/humor/best-worst-ecommerce-site-ever>
2. A website for a company that provides software for ecommerce sites: <http://www.bigcommerce.com/>
3. An SQL tutorial that includes online demos/examples for students to try live: <http://www.w3schools.com/sql/default.asp>

**Key Terms**

* **ASP (Application Service Provider)**:An organization that, for a fee, provides software services over the Web, such as hosting a website.
* **Attribute**: The name of one field in a database tuple.
* **Cookie**: A small text file that gets stored on the hard drive of a visitor to a website allowing a more customized page to be presented on the user's next visit.
* **Data brokers**: Collect data on virtually everyone and then in turn sell that data to clients, who use it to target customers with “personalized” advertising.
* **Data file**: A collection of related records.
* **Data integrity**: A database constraint that says that any attribute in a tuple must contain a data value appropriate for that attribute.
* **Dara mining**: Used to discover previously hidden patterns that a big dataset might contain.
* **Data science**: Not only incorporates many of the tasks of data analysis but also involves knowledge of the enterprise in order to formulate useful queries.
* **Database**: A collection of related data files.
* **Database management system (DBMS)**: Software that manages the files in a database.
* **Disaster recovery strategy**: A plan for recovering data and getting computer systems back up after some emergency event.
* **Distributed database**: Allows the physical data to reside at separate and independent locations that are electronically networked together.
* **Domain name**: The URL for the homepage of a business or organization.
* **Entity**: A fundamental distinguishable component in a dataset.
* **Entity integrity**: A database constraint that says that no tuple can be missing any of its primary key values.
* **Field**: A group of bytes that represents one piece of information, such as a name.
* **Foreign key**: An attribute in a relational database table that is a primary key in another table, used to establish a relationship between the tables.
* **ICANN (Internet Corporation for Assigned Names and Numbers)**: A nonprofit corporation that handles domain name management.
* **Join**: A relational database operation that matches tuples from two different relational tables using a common attribute.
* **Middleware**: Software that allows separate, existing programs to communicate and work together seamlessly.
* **Navigation bar**: A graphical set of links to the major parts of a website so visitors can easily move about the site.
* **Peer-to-peer (collaborative consumption) economy**: An economy based on selling directly from one individual to another individual, oftentimes with the platform taking a cut (e.g., eBay and Groupon).
* **Portal**: An entry point webpage with links to other webpages on some topic.
* **Primary key**: One or more attributes that uniquely identifies a tuple within a relational database table.
* **Project**: A relational database operation that picks out certain attributes from a set of tuples.
* **Record**: A collection of related fields, for example, all information about a single employee.
* **Referential integrity**: A database constraint that says any foreign key attribute value in a given table must match a value in the corresponding primary keys of the related table.
* **Relational database model**: A conceptual representation of a data file as a two-dimensional table.
* **Restrict**: A relational database operation that picks out tuples that meet a certain condition.
* **Site map**: An overview of the organization of a website so visitors can tell where they are and easily move about the site.
* **SQL (Structured Query Language)**: An international standard database query language.
* **Tuple**: One row in a relational database table.

**Solutions to End-of-Chapter Exercises**

**1**. Answers will vary. Possible sites that can be used include Amazon.com, Ebay.com, and Dell.com.

**2**. Answers will vary. Poor retail Web sites have cluttered or “noisy” presentation, unclear or inconsistent navigability, lack of features, lack of flexibility, do not let the customer know where they are in the sales process or what is coming next, etc.

**3**. Answers will vary depending on sites the student has recently visited. To aid students, consider providing instructions on how to access the cookies folder/file.

**4**.

| **ID** | **LastName** | **FirstName** | **Birthdate** | **PayRate** | **HoursWorked** |
| --- | --- | --- | --- | --- | --- |
| 116 | Kay | Janet | 3/29/1980 | $16.60 | 94 |
| 165 | Honou | Morris | 6/9/1997 | $6.70 | 53 |

**5**. SELECT FirstName, LastName, PayRate  
 FROM Employees  
 ORDER BY PayRate;

**6**.

| **ID** | **PlanType** |
| --- | --- |
| 149 | B2 |
| 149 | A1 |
| 149 | C2 |

**7**. SELECT FirstName, LastName, HoursWorked, PlanType  
 FROM Employees, InsurancePolicies  
 WHERE HoursWorked < 100  
 AND ID = EmployeeID;

(The result will be the “null set” - no employee fits this description)

|  |  |  |
| --- | --- | --- |
| A1 | Basic medical/single | 83.54 |
| A4 | Major medical/family | 256.80 |
| B2 | Health management/single | 68.90 |

**8.** Examples of InsurancePlans tuples:

**9**. SELECT LastName, FirstName, InsurancePolicies.PlanType, MonthlyCost  
 FROM Employees, InsurancePlans, InsurancePolicies

WHERE LastName = "Kay" And FirstName = "John"

AND ID = EmployeeID

AND InsurancePolicies.PlanType = InsurancePlans.PlanType;

**10.** SELECT \*

FROM Doctor;

SELECT \*

FROM Patient;

SELECT \*

FROM ClinicAppointment;

**11.** SELECT Doctor.FirstName, Doctor.LastName, Patient.FirstName, Patient.LastName,   
 HomePhone, CellPhone, AppTime

FROM Doctor, Patient, ClinicAppointment

WHERE Doctor.DoctorID = ClinicAppointment.DoctorID

AND Patient.PatientID = ClinicAppointment.PatientID

AND ClinicAppointment.AppDate = #2/28/2018#;

| **Doctor.FirstName** | **Doctor.LastName** | **Patient.FirstName** | **Patient.LastName** | **HomePhone** | **CellPhone** | **AppTime** |
| --- | --- | --- | --- | --- | --- | --- |
| Anne | Davis | Gail | Perez | 333-777-1212 | 333-410-7777 | 8:30:00 AM |
| Estelle | Villanueva | Gordon | Zhang | 332-555-9999 | 332-217-4321 | 10:30:00 AM |

**12.** SELECT FirstName, LastName, AppTime

FROM Doctor, ClinicAppointment

WHERE Doctor.DoctorID = ClinicAppointment.DoctorID

AND ClinicAppointment.AppDate=#2/27/2018#

AND Doctor.DoctorID ="DO1"

ORDER BY AppTime;

| **FirstName** | **LastName** | **AppTime** |
| --- | --- | --- |
| Vladimir | Yevgeny | 11:30:00 AM |
| Vladimir | Yevgeny | 1:00:00 PM |

**13.** SELECT Patient.FirstName, Patient.LastName, AppTime, AppDate,   
 Doctor.FirstName, Doctor.LastName

FROM Patient, Doctor, ClinicAppointment

WHERE Doctor.DoctorID = ClinicAppointment.DoctorID

AND Patient.PatientID=ClinicAppointment.PatientID

AND Patient.PatientID="PA3"

AND AppDate=#2/27/2018#;

| **Patients.FirstName** | **Patients.LastName** | **AppTime** | **AppDate** | **Doctors.FirstName** | **Doctors.LastName** |
| --- | --- | --- | --- | --- | --- |
| DuWayne | Martin | 1:00:00 PM | 2/27/2018 | Vladimir | Yevgeny |

**14.** High blood pressure seems to be a small risk factor because there are only small differences in the percentages between low and high blood pressure individuals who otherwise classify the same on age and obesity. Age plays a slightly higher part because the percentage difference between the lows in either age classification is small and so is the percentage difference between the highs in either age classification. Obesity is the greatest risk factor, causing a split between lows and highs regardless of age or blood pressure.

**15.**  a. The functions to use are: MIN, MEDIAN, AVERAGE, MAX

b. After the chart is created, right-click in the chart area and choose Select Data. Then click Switch Row/Column. Then Edit the Legend Series to change the names to Site 1, Site 2, etc. The result should now agree with Figure 10.9. c. Student should be able to just follow the instructions in the problem statement. Results should agree with the R results (with perhaps fewer decimal places) and, because the matrix is symmetric, only the lower triangular portion is shown.

**Challenge Work**

The appointment two days later would be March 1, 2018, because 2018 is not a leap year so there is no February 29. If you try to enter an appointment of the form

DO1 PA03 2/29/2018 9:00 AM Revisit chest pain

you will get an error message:

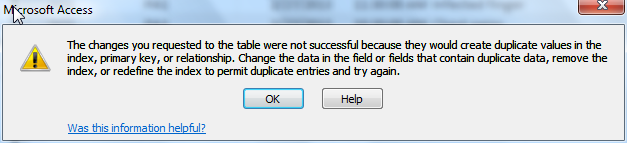
‘The value you entered does not match the Date/Time data type in this column’.

The system is enforcing the data integrity rule. This is the advantage of using the Date/Time data type rather than just Text. But this is not the major problem

If you try to insert

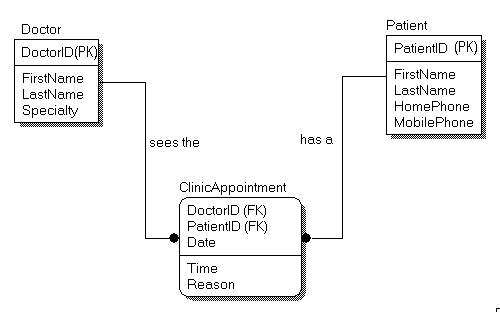
DO1 PA03 3/01/2018 9:00 AM Revisit chest pain

you get the message shown below. Why?



With the current database design, the composite key DoctorID/PatientID is the primary key of the ClinicAppointment table, which means it should uniquely identify the tuple. Therefore there cannot be two different tuples with the same DoctorID/PatientID pair.

You need to modify the design of the database to look as follows.



Then you can add the new record. The SQL query is

SELECT \*

FROM ClinicAppointment

ORDER BY AppDate, AppTime;

And the result is

**ClinicAppointment**

| **DoctorID** | **PatientID** | **AppDate** | **AppTime** | **Reason** |
| --- | --- | --- | --- | --- |
| DO1 | PA1 | 2/27/2018 | 11:30:00 AM | Infected finger |
| DO1 | PA3 | 2/27/2018 | 1:00:00 PM | Chest pains |
| DO4 | PA2 | 2/28/2018 | 8:30:00 AM | Prenatal checkup |
| DO3 | PA4 | 2/28/2018 | 10:30:00 AM | Poison ivy |
| DO1 | PA3 | 3/1/2018 | 9:00:00 AM | Revisit chest pain |