Database Processing: Fundamentals, Design, and Implementation

Chapter One

Introduction

semicolon

Learning Objectives (1 of 2)

1.1 To understand the importance of databases in Internet Web applications and smart-phone apps

1.2 To understand the nature and characteristics of databases

1.3 To survey some important and interesting database applications

1.4 To gain a general understanding of tables and relationships

1.5 To describe the components of a Microsoft Access database system and explain the functions they perform

1.6 To describe the components of an enterprise-class database system and explain the functions they perform

1.7 To define the term *database management system* (DBMS) and describe the functions of a DBMS

1.8 To define the term *database* and describe what is contained within the database

Learning Objectives (2 of 2)

1.9 To define the term *metadata* and provide examples of metadata

1.10 To define and understand database design from existing data

1.11 To define and understand database design as new systems development

1.12 To define and understand database redesign of an existing database

1.13 To understand the history and development of database processing

How Did We Get Here?
The Internet World I

- Personal Computers
  - 1977: Apple II
  - 1981: IBM PC
- Local Area Networks
  - Ethernet networking technology
    - Early 1970s: Xerox Palo Alto Research Center
    - 1983: U.S. National Standard

How Did We Get Here?
The Internet World II

- The Internet
  - 1969: ARPANET
- World Wide Web (WWW)
  - 1993: First Web browser (Netscape) available
  - Mid 1990s: Online retail sites
    - 1995: Amazon
    - Followed by Best Buy
- Early 2000s: Web 2.0

How Did We Get Here?
The Smartphone World

Mid 1970s: Mobile Phone (Cell Phone)

Smartphone

- 2007: Apple iPhone
- 2008: Google Android Operating System

Tablets

- 2010: Apple iPad

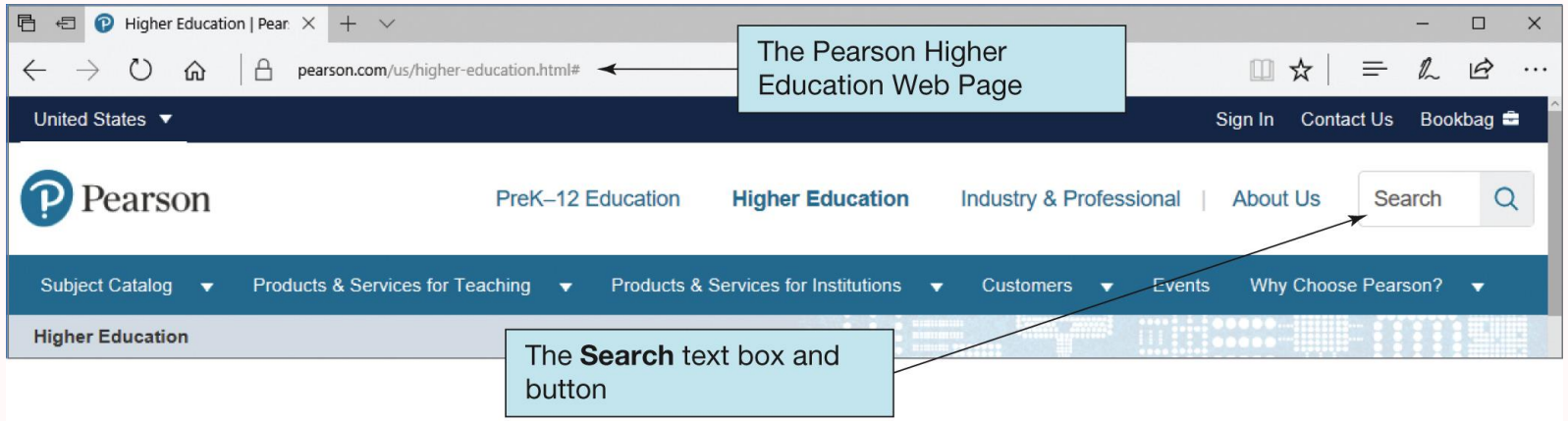
Apps

All of these examples depend on databases

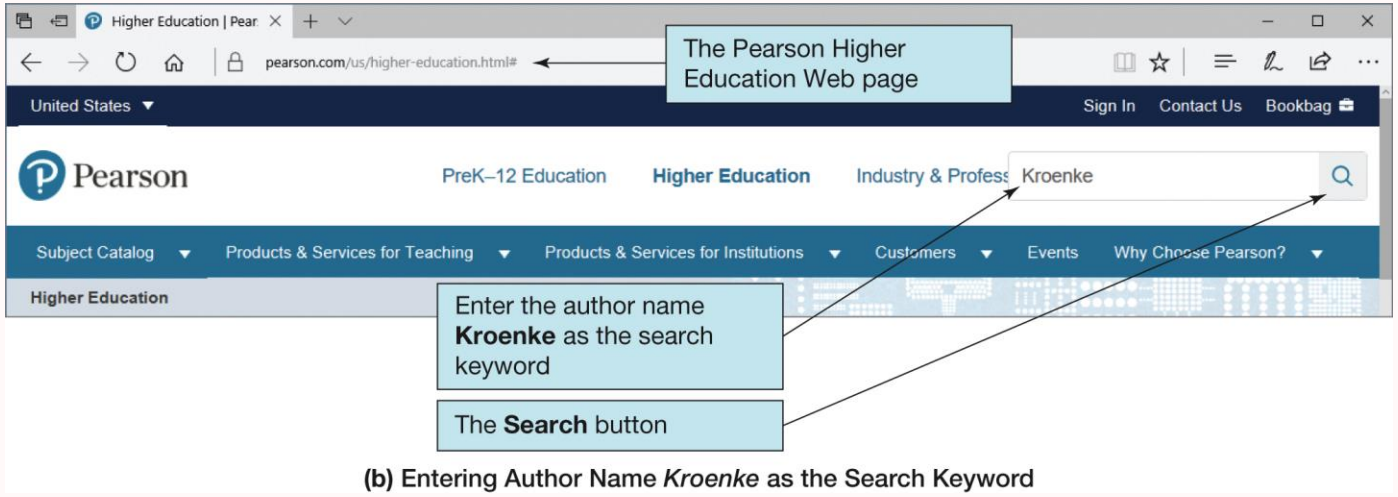
Databases in the Internet and Mobile Device World

- Databases are important because they are everywhere and are used daily:
  - Facebook
    - Posts
    - Likes
  - Twitter
    - Tweets
  - Online shopping
    - Amazon.com – [Do an actual search]

Figure 1-1
Searching a Database in a Web Browser (1 of 2)

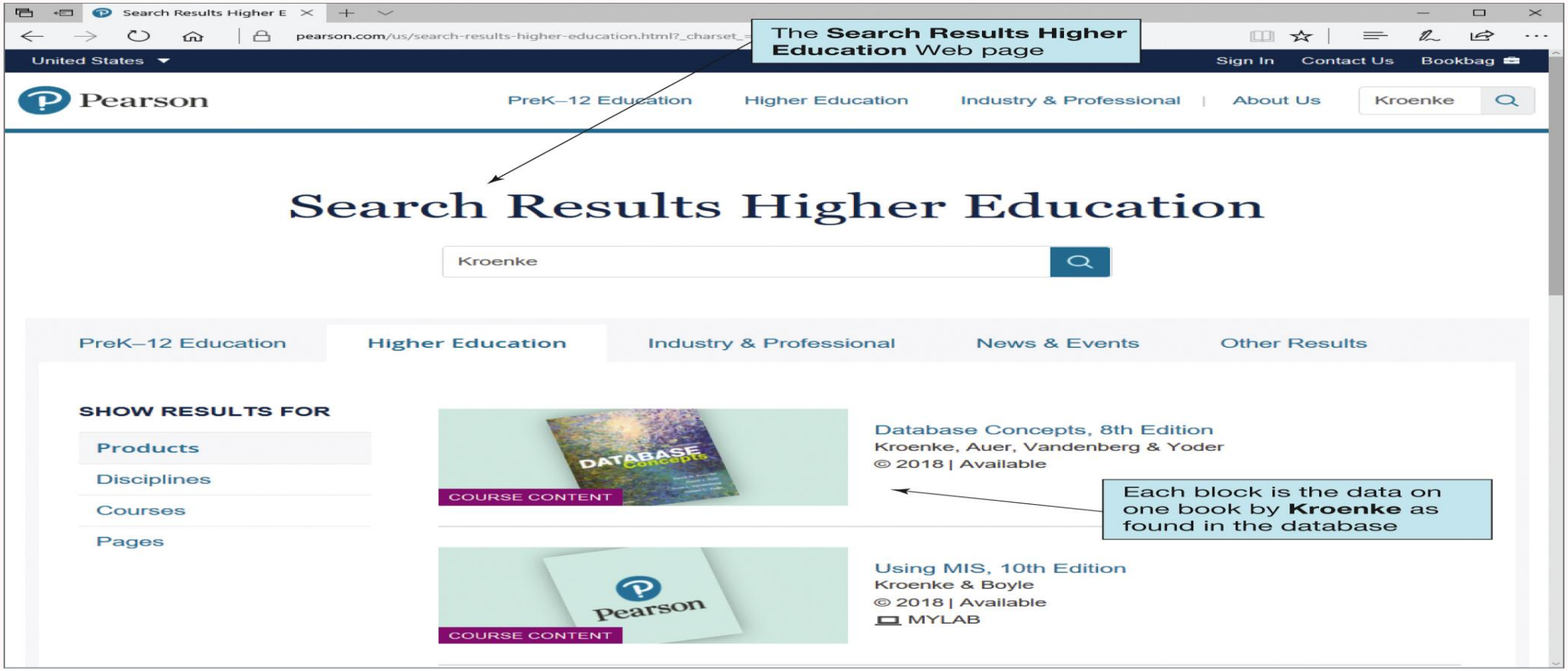


(a) The Pearson Higher Education Web Page



(b) Entering Author Name *Kroenke* as the Search Keyword

Figure 1-1
Searching a Database in a Web Browser (2 of 2)

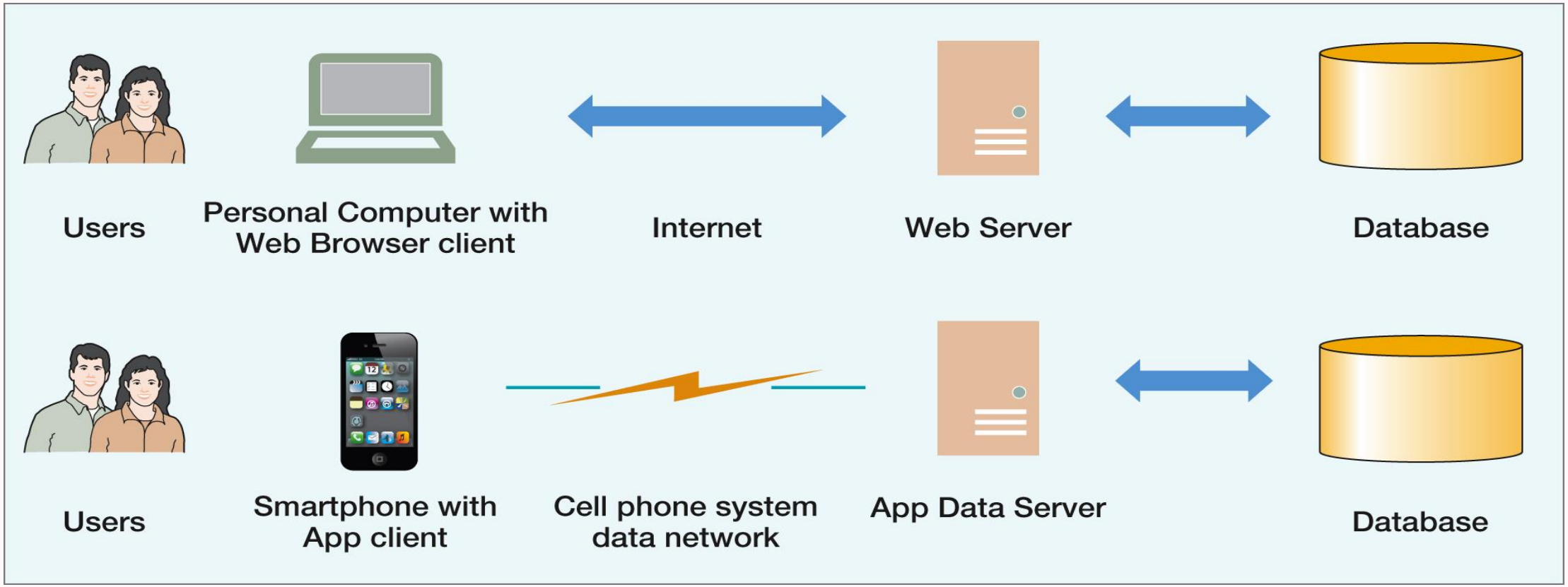


(c) Books by Author *Kroenke* Found in the Database

The Characteristics of Databases

- The purpose of a **database** is to help people track things of interest to them.
- Data is stored in **tables**, which have rows and columns like a spreadsheet.
- A database may have multiple tables, where each table stores data about a different thing.
- Each row in a table stores data about an occurrence or **instance** of the thing of interest.
- A database stores **data** and **relationships**.

Figure 1-2
The Internet and Mobile Device World



Naming Conventions in this Slide

- **Table names** are written with all capital letters:
  - STUDENT, CLASS, GRADE, COURSE_INFO
- **Column names** are written with an initial capital letter, and compound names are written with a capital letter on each word:
  - Term, Section, ClassNumber, StudentName

Figure 1-3
The STUDENT and CLASS Tables

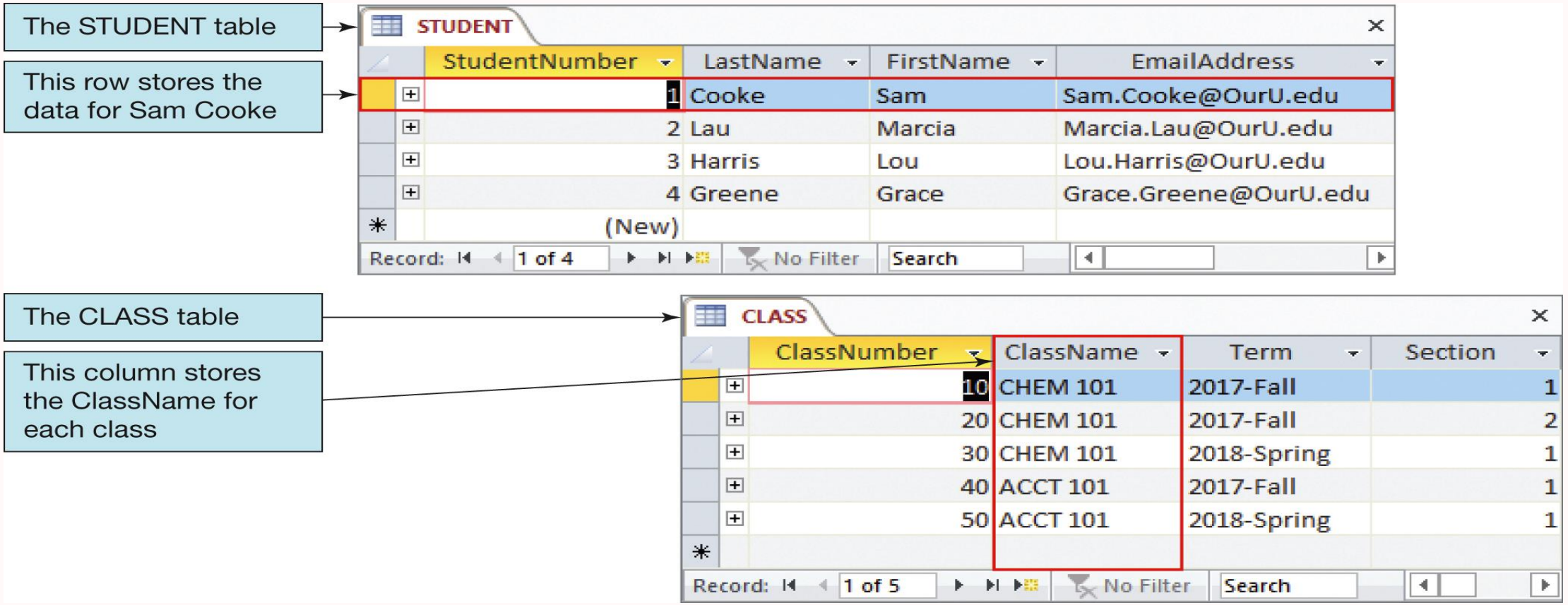


Figure 1-4
The STUDENT, CLASS, and GRADE Tables

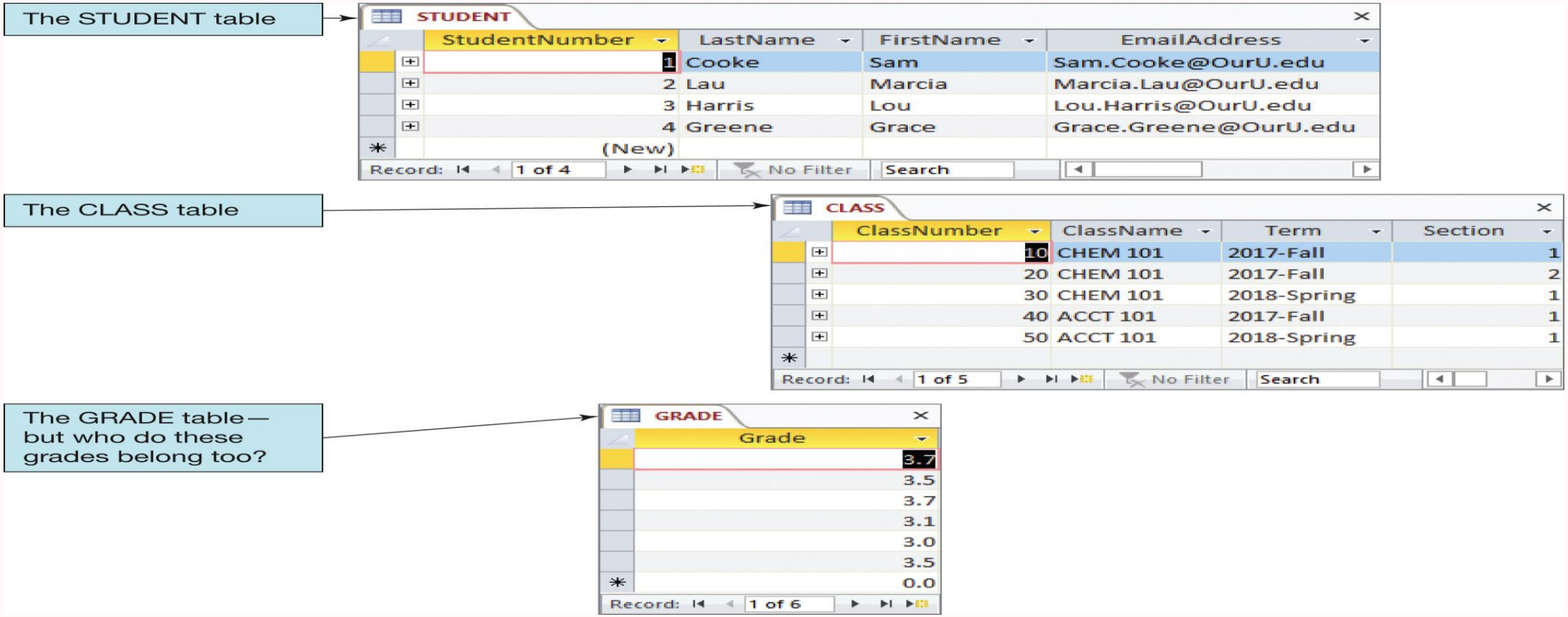


Figure 1-5
The Key Database Characteristics: Related Tables

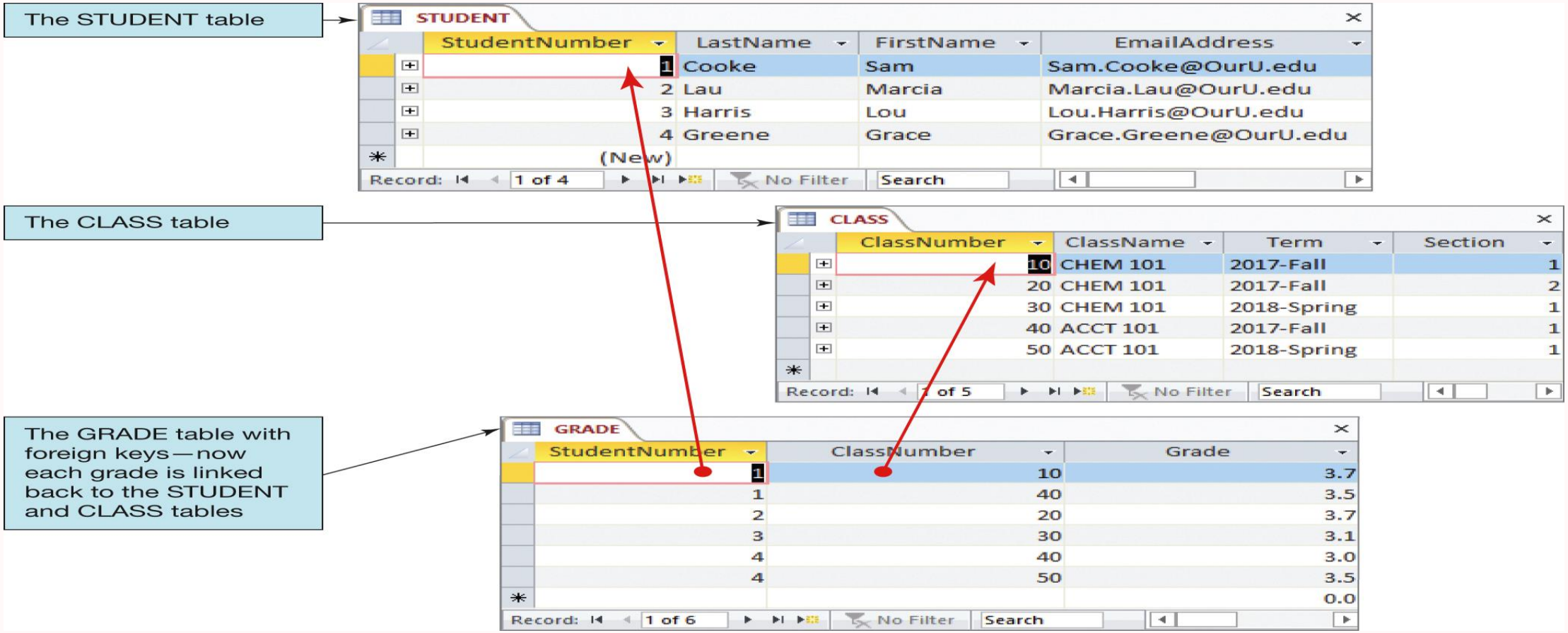
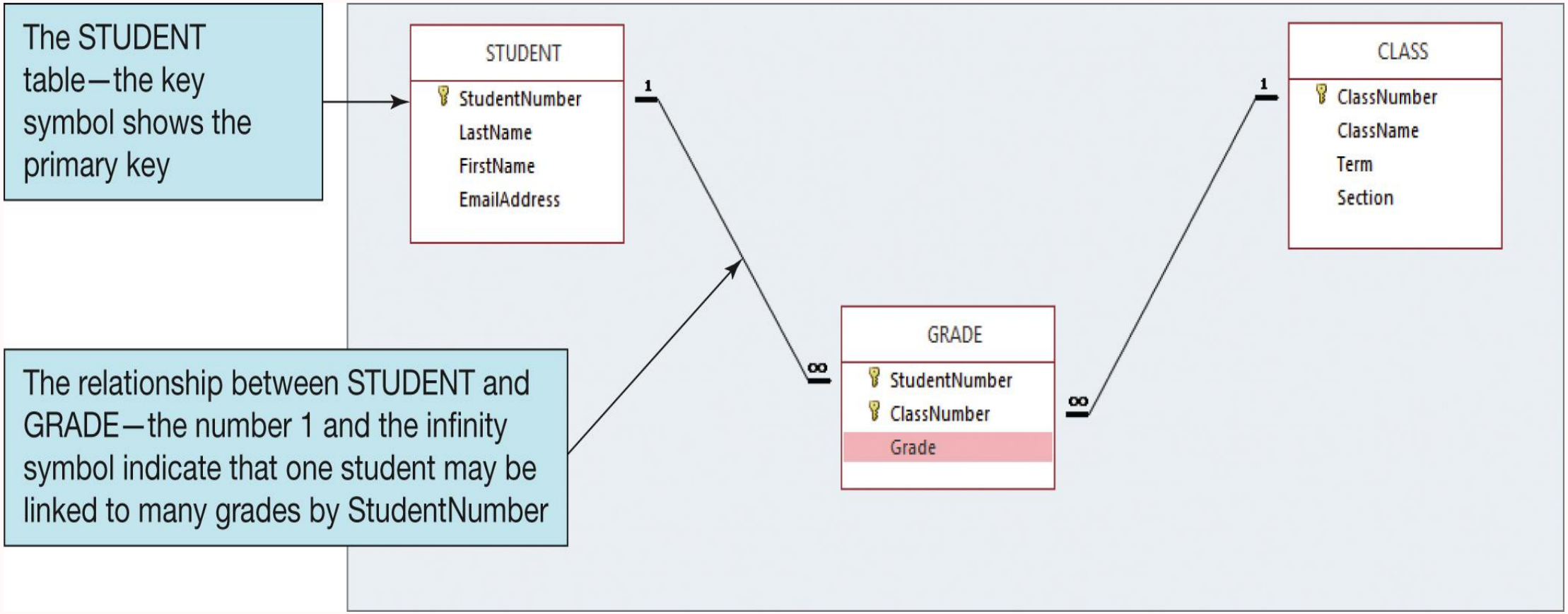


Figure 1-6
Microsoft Access 2016 View of Tables and Relationships



Databases Create Information

- **Data** = recorded facts and figures
- **Information** = knowledge derived from data
- Databases record data, but they do so in such a way that we can produce information from the data
  - The data on STUDENTs, CLASSes, and GRADEs could produce information about each student's GPA.

Figure 1-7
Example Database Applications

Application	Example Users	Number of Users	Typical Size	Remarks
Sales Contact Manager	Salesperson	1	2,000 rows	Products such as GioldMine and Act! Are database centric
Patient appointment (doctor, dentist)	Medical office	15 to 50	100,000 rows	Vertical market software vendors incorporate databases into their software products
Customer relationship management (CRM)	Sales, marketing, or customer service departments	500	10 million rows	Major vendors such as Microsoft and Oracle PeopleSoft Enterprise build applications around the database
Enterprise resource planning (ERP)	An entire organization	500	10 million+ rows	SAP uses a database as a central repository for ERP data.
E-commerce site	Internet users	Possibly millions	1 billion+ rows	Drugstore.com has a database that grows at the rate of 20 million rows per day!
Digital dashboard	Senior managers	500	100,000 rows	Extractions, summaries, and consolidations of operational databases.
Data mining	Business analysts	25	100,000 to millions+	Data are extracted, reformatted, cleaned, and filtered for use by statistical mining tools.

Figure 1-8
The Components of a Database System

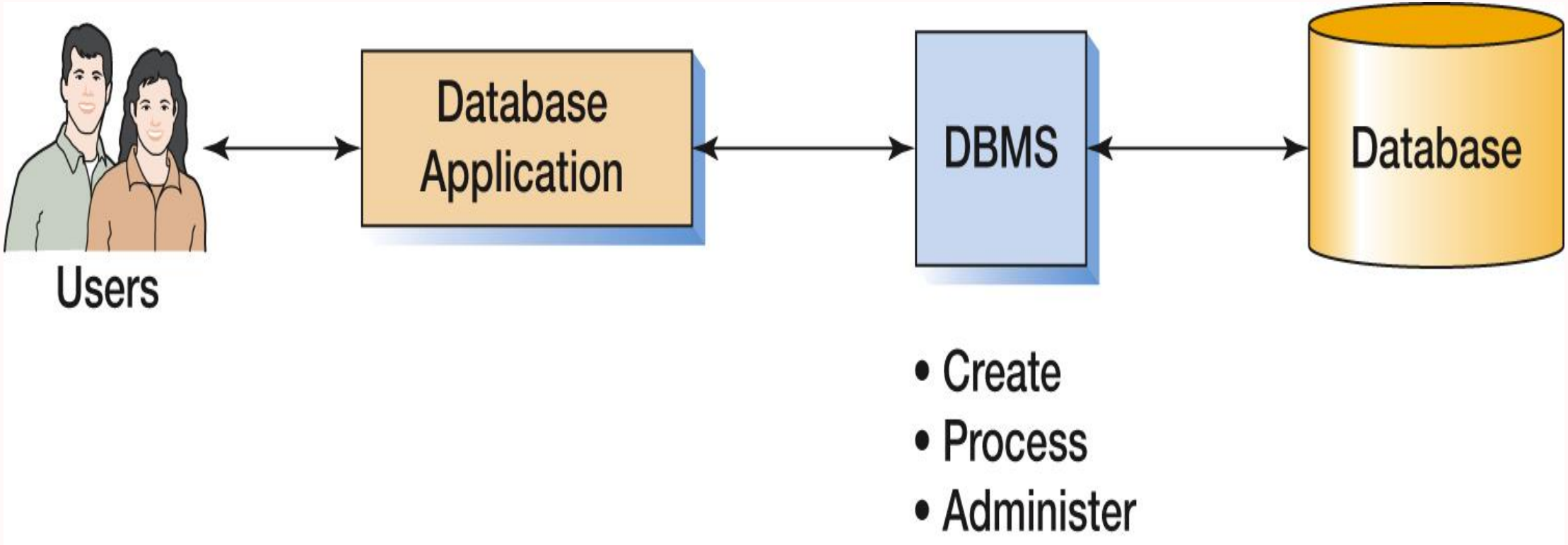
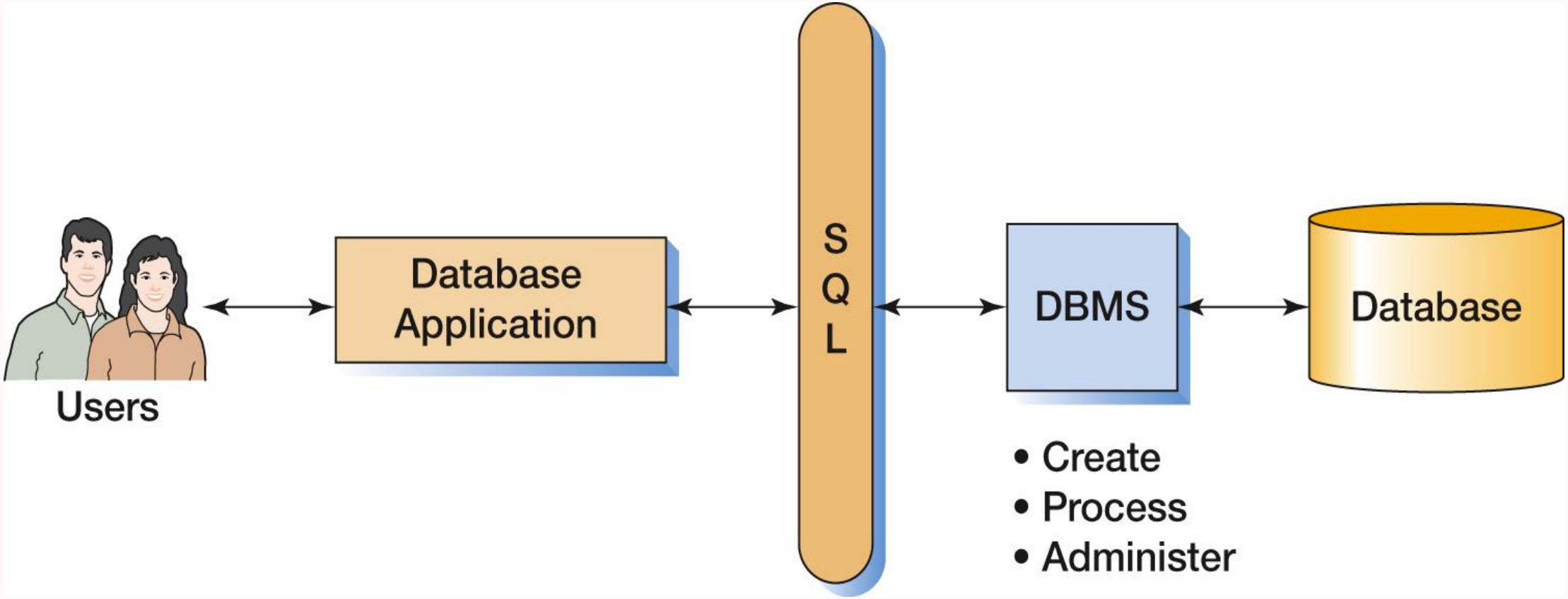


Figure 1-9
The Components of a Database System with SQL



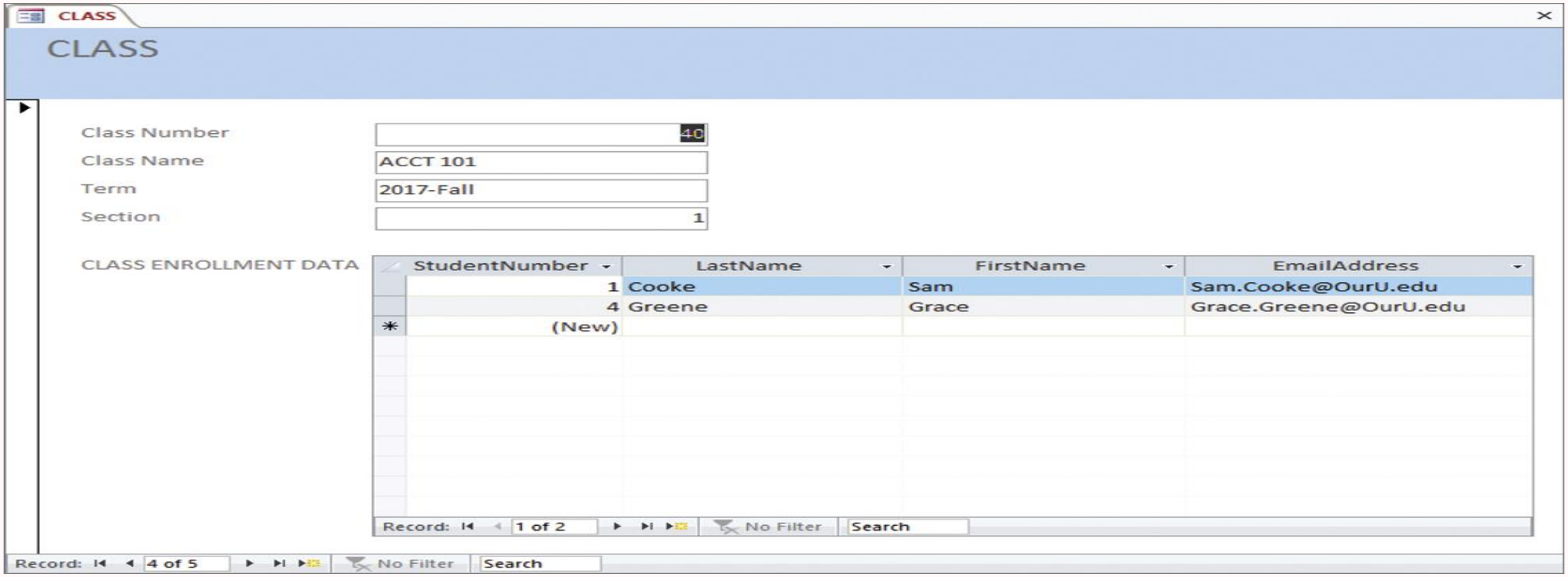
Applications, the DBMS, and SQL

- **Applications** are the computer programs that users work with.
- The **Database Management System (DBMS)** creates, processes, and administers databases.
- **Structured Query Language (SQL)** is an internationally recognized standard database language that is used by all commercial DBMSs.

Figure 1-10
Basic Functions of Application Programs

Basic Functions of Application Programs
Create and process forms
Process user queries
Create and process reports
Execute application logic
Control the application itself

Figure 1-11
An Example Data Entry Form



semicolon

Figure 1-12
Example SQL Query Results

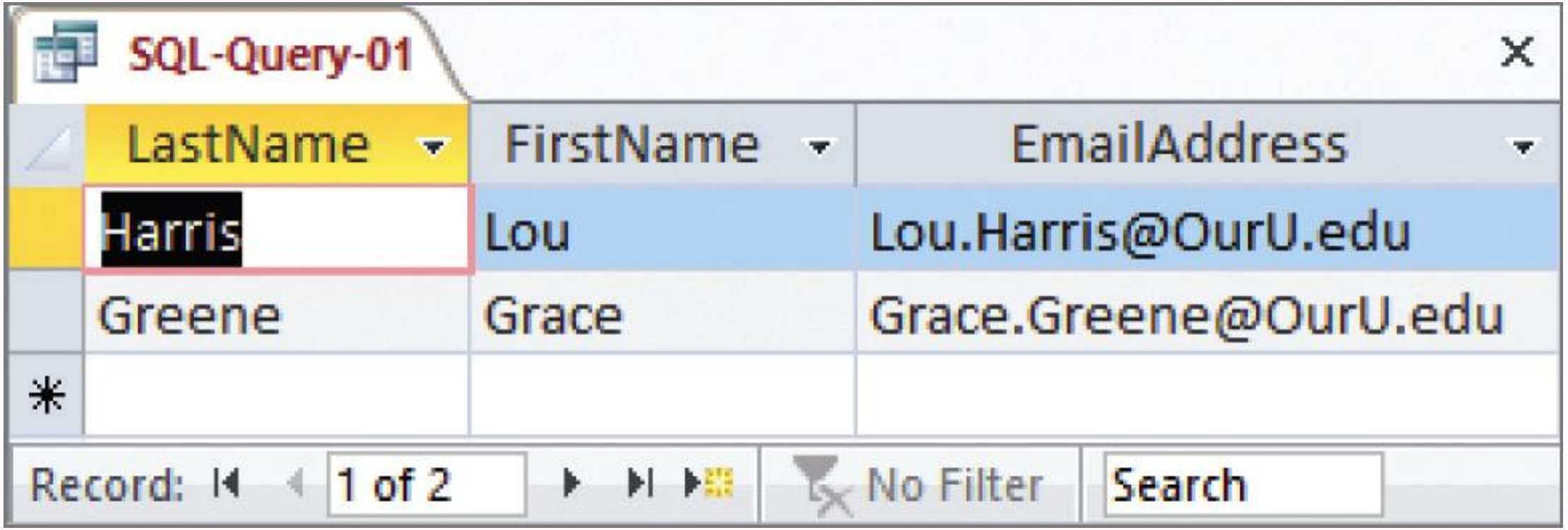


Figure 1-13
Example Report

Class Grade Report

ClassNumber	ClassName	Term	Section	LastName	FirstName	Grade
10	CHEM 101	2017-Fall	1			
				Cooke	Sam	3.7
20	CHEM 101	2017-Fall	2			
				Lau	Marcia	3.7
30	CHEM 101	2018-Spring	1			
				Harris	Lou	3.1
40	ACCT 101	2017-Fall	1			
				Cooke	Sam	3.5
				Greene	Grace	3.0
50	ACCT 101	2018-Spring	1			
				Greene	Grace	3.5

Figure 1-14
Functions of a DBMS

Functions of a DBMS
Create database
Create tables
Create supporting structures (e.g., Indexes)
Modify (insert, update, or delete) database data
Read database data
Maintain database structures
Enforce rules
Control concurrency
Perform backup and recovery

The Database

A **database** is a self-describing collection of integrated tables.

The tables are called **integrated** because they store data about the relationships between rows of data.

A database is called **self-describing** because it stores a description of itself.

The self-describing data is called **metadata**, which is data about data.

Figure 1-15
Typical Metadata Tables (1 of 2)

TableName	NumberColumns	PrimaryKey
STUDENT	4	StudentNumber
CLASS	4	ClassNumber
GRADE	3	(StudentNumber, ClassNumber)

Figure 1-15
Typical Metadata Tables (2 of 2)

ColumnName	TableName	DataType	Length (bytes)
StudentNumber	STUDENT	Integer	4
LastName	STUDENT	Text	25
FirstName	STUDENT	Text	25
EmailAddress	STUDENT	Text	100
ClassNumber	CLASS	Integer	4
Name	CLASS	Text	25
Term	CLASS	Text	12
Section	CLASS	Integer	4
StudentNumber	GRADE	Integer	4
ClassNumber	GRADE	Integer	4
Grade	GRADE	Decimal	(2,1)

Microsoft Access

- **Microsoft Access** is a low-end product intended for individual users and small workgroups.
- Microsoft Access tries to hide much of the underlying database technology from the user.
- This is a good strategy for beginners, but not for database professionals.

What is Microsoft Access?

- Microsoft Access is a DBMS plus an application generator:
  - The DBMS creates, processes, and administers Microsoft Access databases.
  - The application generator includes query, form, and report components.
- The Microsoft Access DBMS engine is called the Access Data Engine (ADE).
- Microsoft Access 2000 thru 2010 can be used as an application generator for the Microsoft SQL Server DBMS.

Figure 1-17
Components of a Microsoft Access Database System

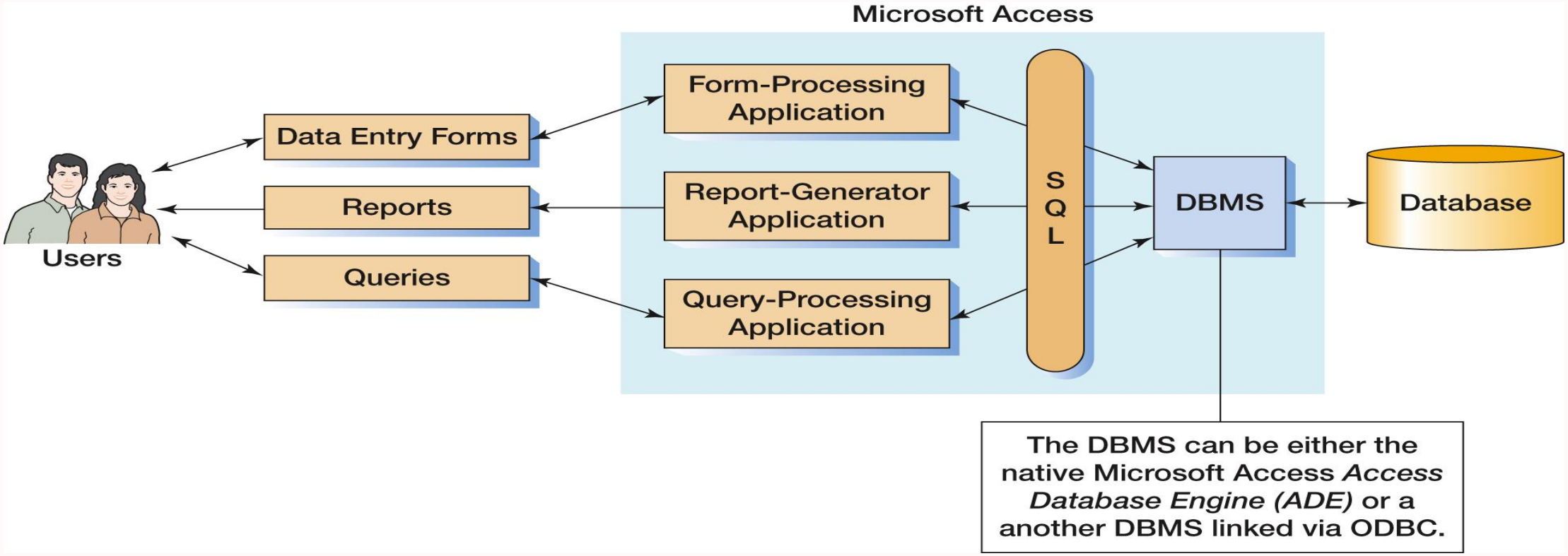


Figure 1-18
Components of an Enterprise-Class Database System

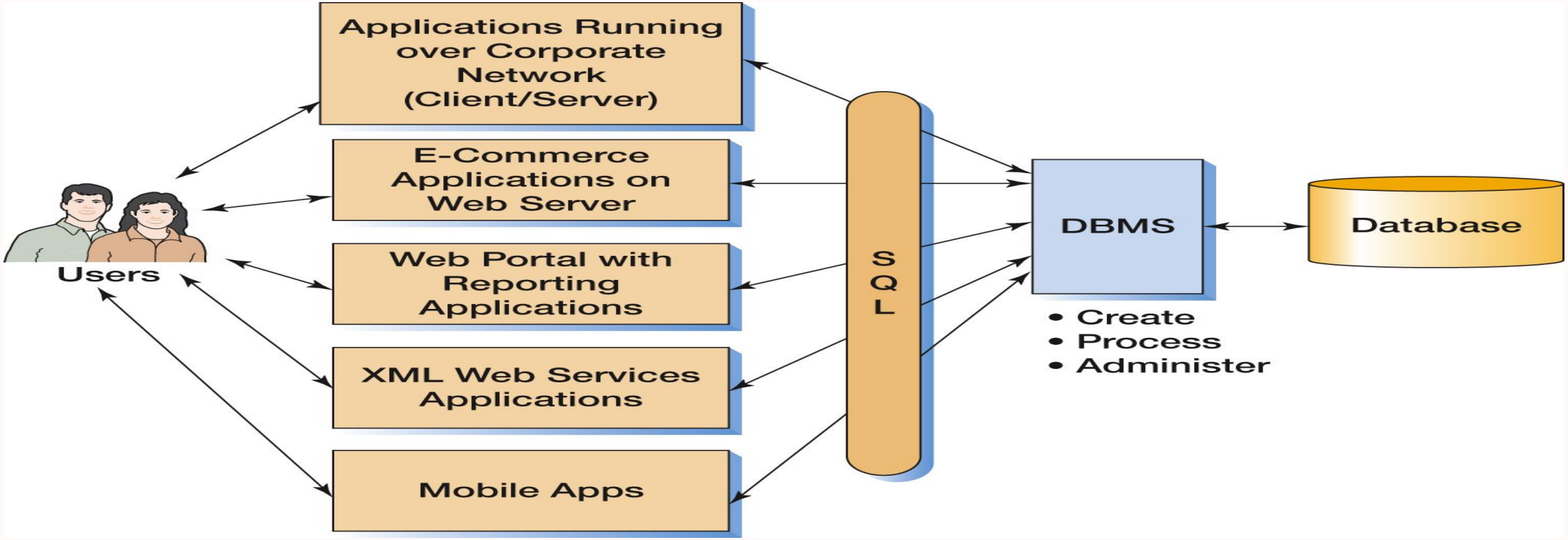


Figure 1-20
Three Types of Database Design

Types of Database Design Process
• From existing data (Chapters 3 and 4)
Analyze spreadsheets and other data tables
Extract data from other databases
Design using normalization principles
• New systems development (Chapters 5 and 6)
Crete data model from application requirements
Transform data model into database design
• Database redesign (Chapter 8)
Migrate databases to newer databases
Integrate two or more databases
Reverse-engineer and design new database using normalization principles and data model transformation

Figure 1-21
Databases Originating from Existing Data

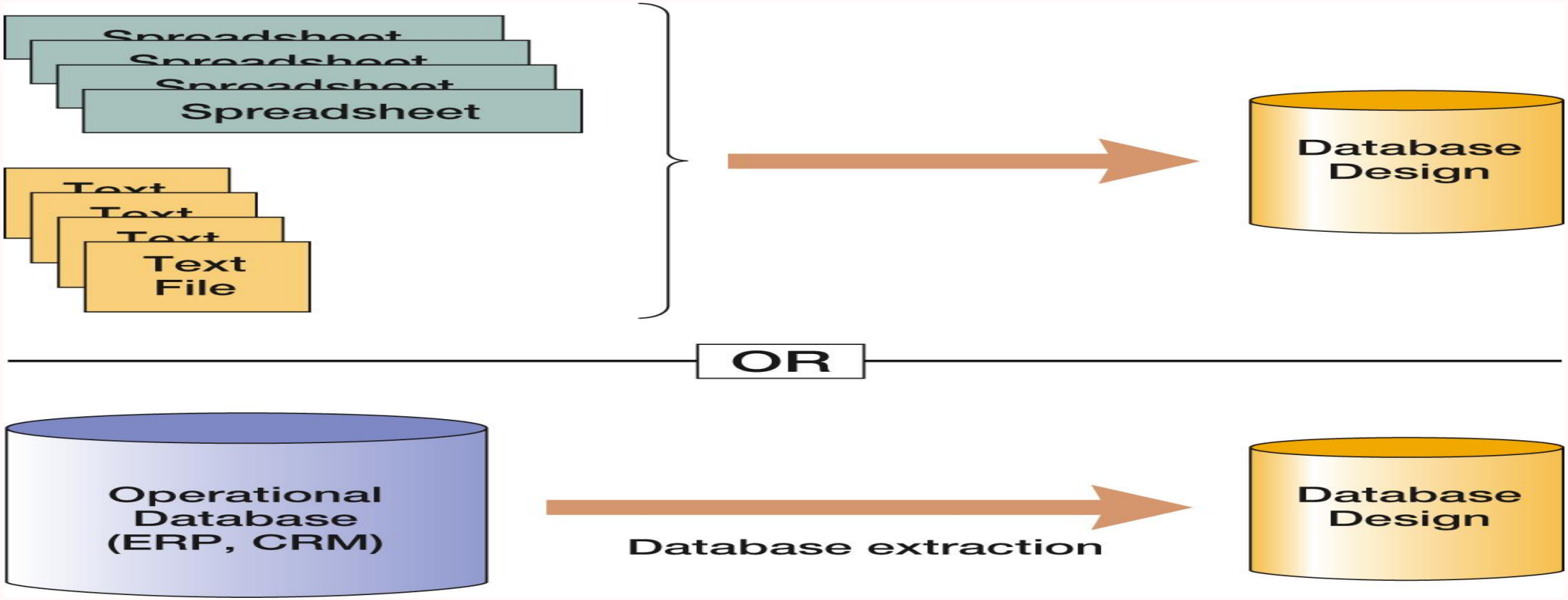


Figure 1-22
Data Import: One or Two Tables?

EmpNum	EmpName	DeptNum	DeptName
100	Jones	10	Accounting
150	Lau	20	Marketing
200	McCauley	10	Accounting
300	Griffin	10	Accounting

(a) One-Table Design

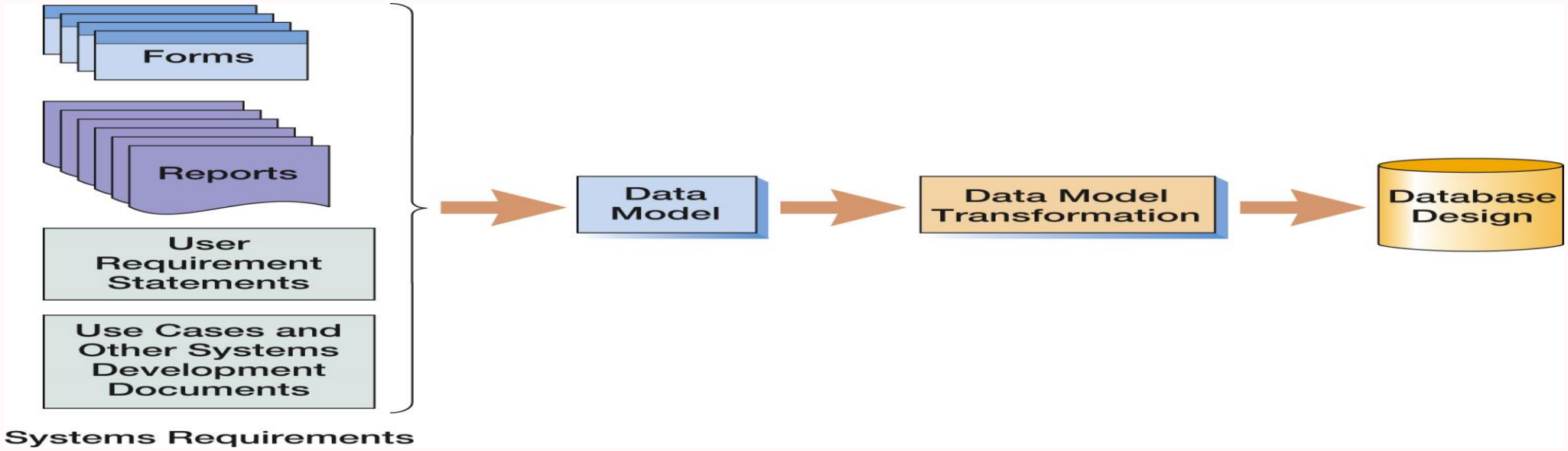
OR?

DeptNum	DeptName
10	Accounting
20	Marketing

EmpNum	EmpName	DeptNum
100	Jones	10
150	Lau	20
200	McCauley	10
300	Griffin	10

(b) Two-Table Design

Figure 1-23
Database Originating from New Systems Development



- **Entity-Relationship data modeling** is covered in Chapter 5, and **data model transformations to database designs** are covered in Chapter 6.

Figure 1-24
Databases Originating from Database Redesign

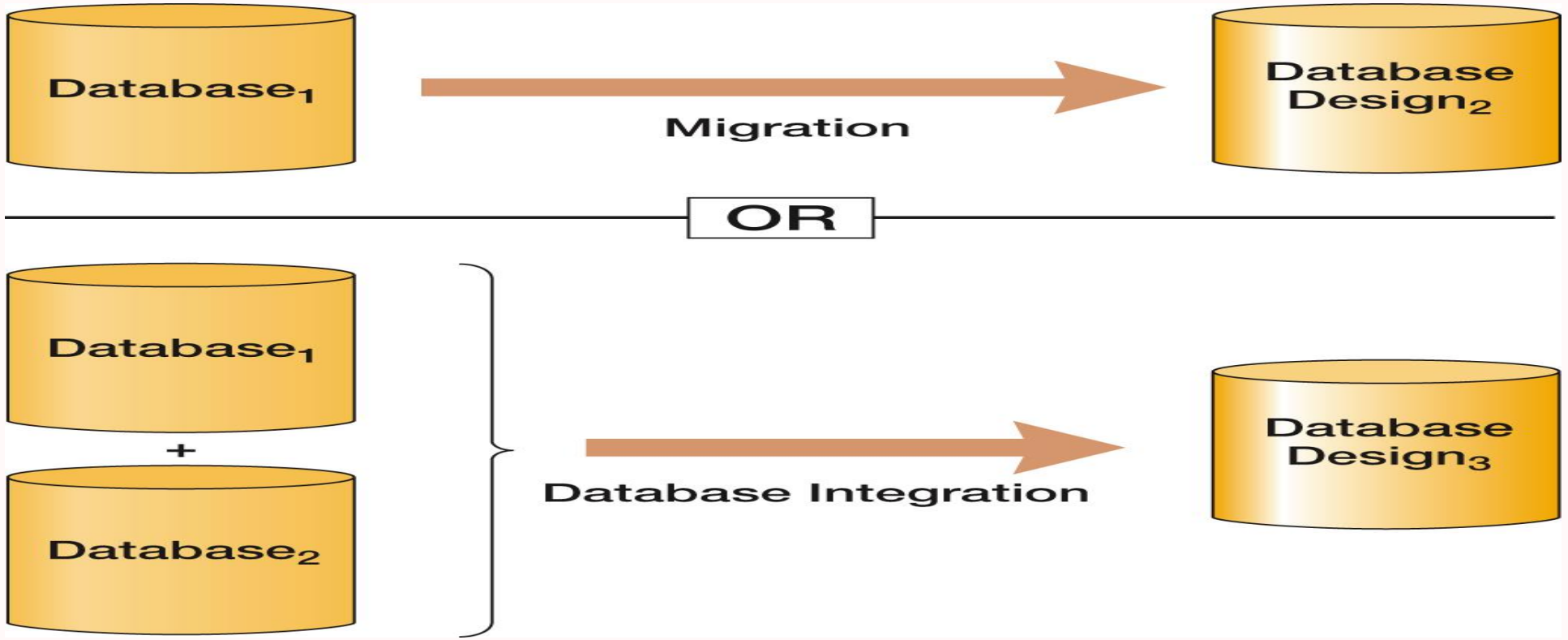


Figure 1-25
Working Domains of Knowledge Workers, Programmers, and Database Administrators

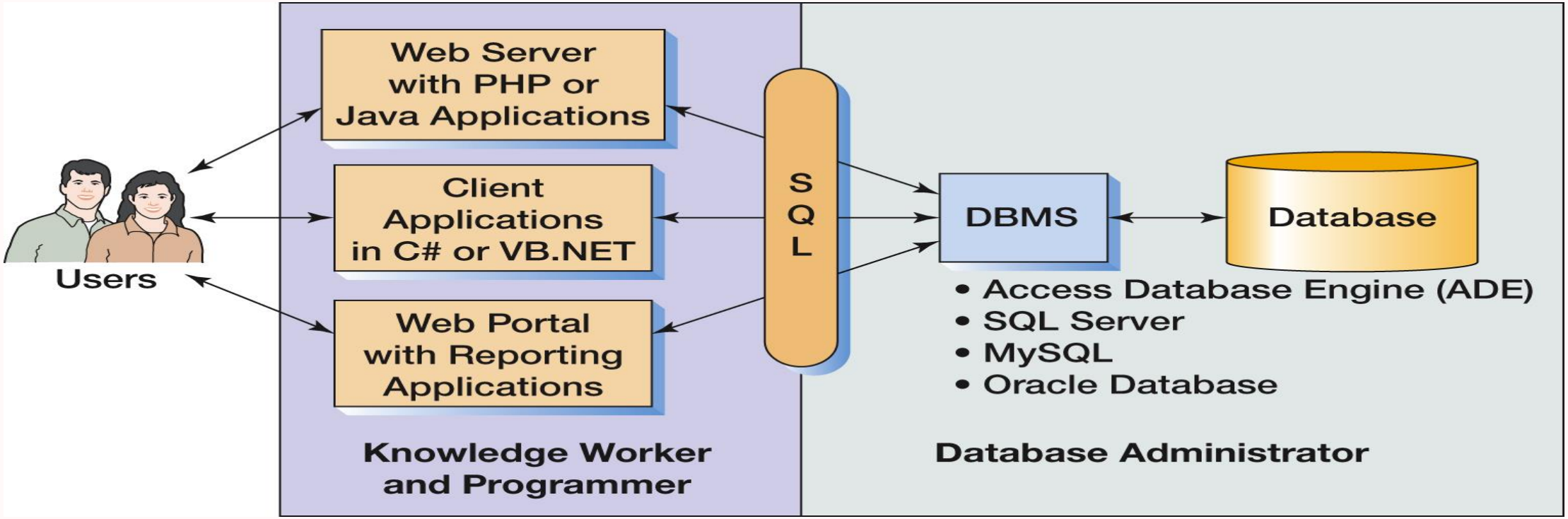


Figure 1-27
Database History (1 of 2)

Era	Years	Important Products	Remarks
Predatabase	Before 1970	File Managers	All data were stored in separate files. Data integration was very difficult. File storage space was expensive and limited.
Early Database	1970-1980	ADABAS, System2000, Total, IDMS, IMS	First products to provide related tables. CODASYL DBTG and hierarchical data models (DL/I) were prevalent.
Emergence of relational model	1978-1985	DB2, Oracle Database, Ingres	Early relational DBMS products had substantial inertia to overcome. In time, the advantages weighed out.
Microcomputer DBMS products	1982-1992+	dBase-II, R:base, Paradox, Microsoft Access	Amazing! A database on a micro. All micro DBMS products were eliminated by Microsoft Access in the early 1990s.
Object-oriented DBMS	1985-2000	Oracle ODBMS, Gemstone, O2, Versant	Never caught on. Required relational database to be converted. Too much work for perceived benefit.

Figure 1-27
Database History (2 of 2)

Era	Years	Important Products	Remarks
Web Databases	1995-Present	IIS, Apache, PHP, ASP.NET, and Java	Stateless characteristic of HTTP was a problem at first. Early applications were simple one-stage transactions. Later, more complex logic developed.
Open source DBMS products	1995-Present	MySQL, PostgresQL, and other products	Open source DBMS products provide much of the functionality and features of commercial DBMS products at reduced cost.
XML, and Web services	1998-Present	XML, SOAP, WSDL, UDDI, and other standards	XML provides tremendous benefits to Web-based database applications. Very important today. May replace relational databases during your career. See Chapter 11 and Appendix I.
Big Data and the NoSQL movement	2009-present	Hadoop, Cassandra, Hbase, CouchDB, Arango DB, MongoDB, JSON and other products	Web applications such as Facebook and Twitter use Big Data technologies. The NoSQL movement is geared toward processing large data sets using NoSQL data models which replace relational databases with nonrelational data structures such as XML and JSON, and which may supplant relational databases during your career. See Chapter 12 and Appendices K and L.

The Relational Database Model

- The dominant database model is the **relational database model**—all. current major DBMS products are based on it.
- It was created by IBM engineer **E. F. Codd** in 1970.
- It was based on mathematics called **relational algebra**.

The NoSQL Movement and Big Data

- Recent developments in Internet and mobile computing have resulted in the development of non-relational DBMSs.
  - NoSQL movement
  - Big Data
- These do not replace the relational model, but rather complement it.

End of Presentation:

Chapter One

semicolon