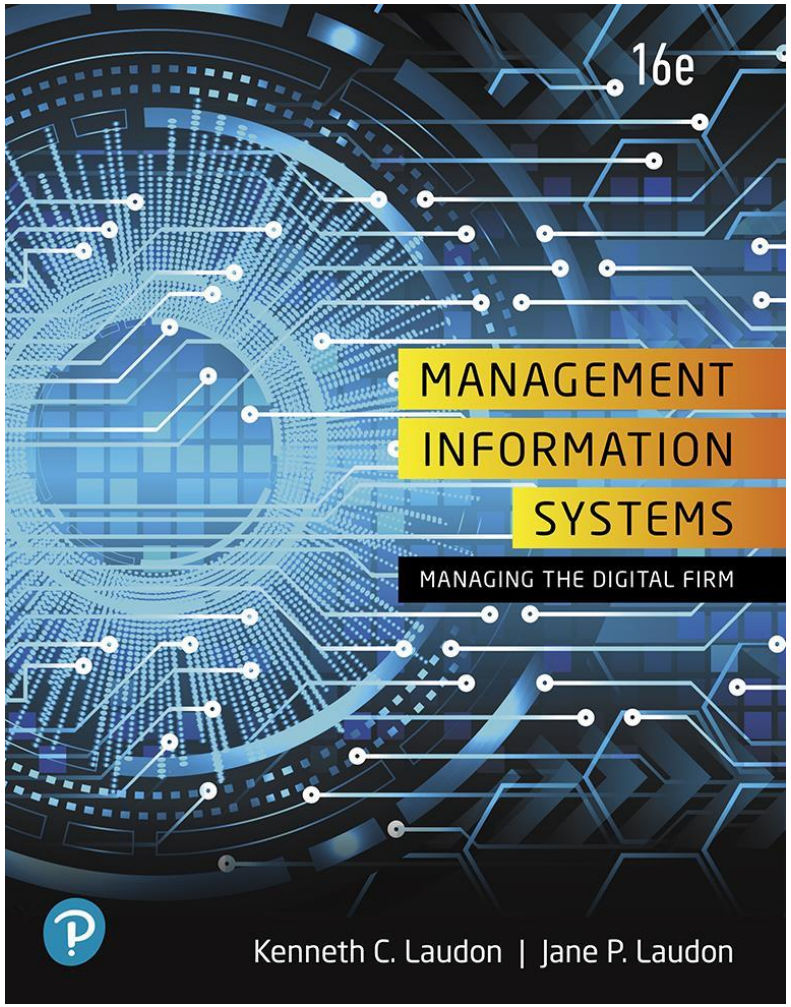
Management Information Systems: Managing the Digital Firm

Sixteenth Edition



Chapter 6

Foundations of Business Intelligence: Databases and Information Management

Pearson

Learning Objectives

6.1 What are the problems of managing data resources in a traditional file environment?

6.2 What are the major capabilities of database management systems (DBMS), and why is a relational DBMS so powerful?

6.3 What are the principal tools and technologies for accessing information from databases to improve business performance and decision making?

6.4 Why are information policy, data administration, and data quality assurance essential for managing the firm's data resources?

6.5 How will MIS help my career?

Video Cases

- Case 1: Dubuque Uses Cloud Computing and Sensors to Build a Smarter City
- Case 2: Brooks Brothers Closes In on Omnichannel Retail
- Case 3: Maruti Suzuki Business Intelligence and Enterprise Databases

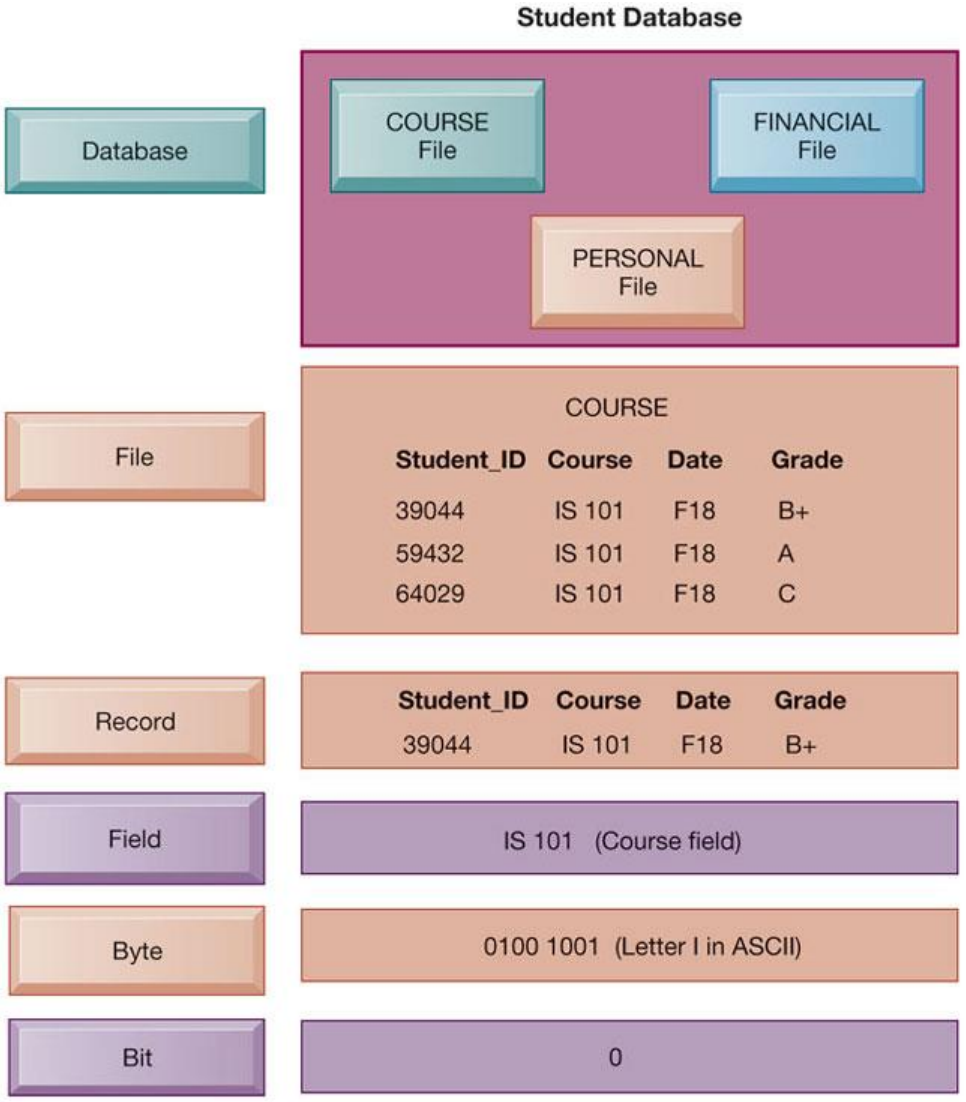
Data Management Helps the Charlotte Hornets Learn More About Their Fans

- Problem
  - Large volumes of data in isolated databases
  - Outdated data management technology
- Solutions
  - SAP HANA
  - Data warehouse
  - FanTracker
- Illustrates the importance of data management for better decision making and customer analysis

File Organization Terms and Concepts

- Database: Group of related files
- File: Group of records of same type
- Record: Group of related fields
- Field: Group of characters as word(s) or number(s)
- Entity: Person, place, thing on which we store information
- Attribute: Each characteristic, or quality, describing entity

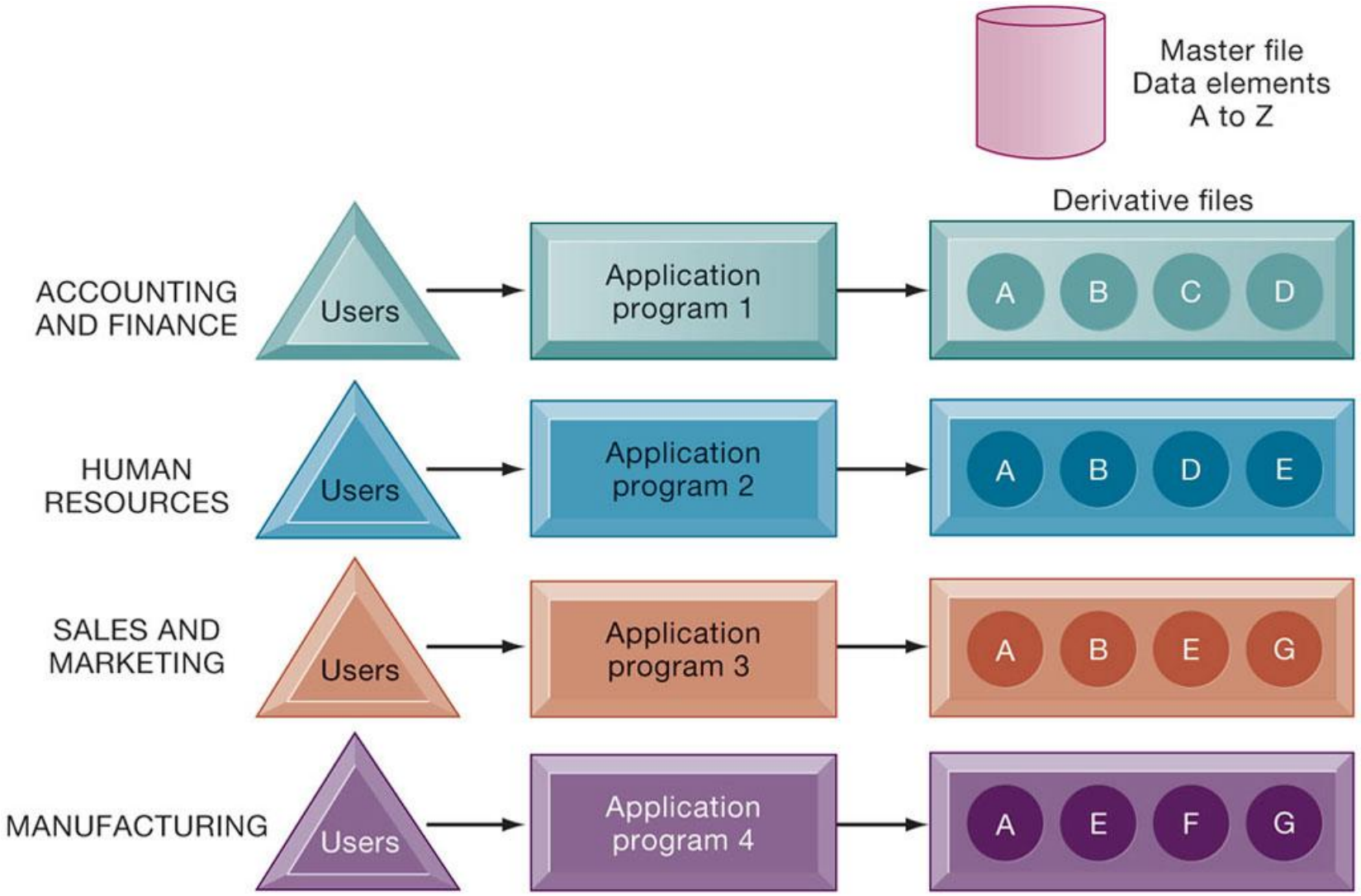
Figure 6.1 The Data Hierarchy



Problems with the Traditional File Environment

- Files maintained separately by different departments
- Data redundancy
- Data inconsistency
- Program-data dependence
- Lack of flexibility
- Poor security
- Lack of data sharing and availability

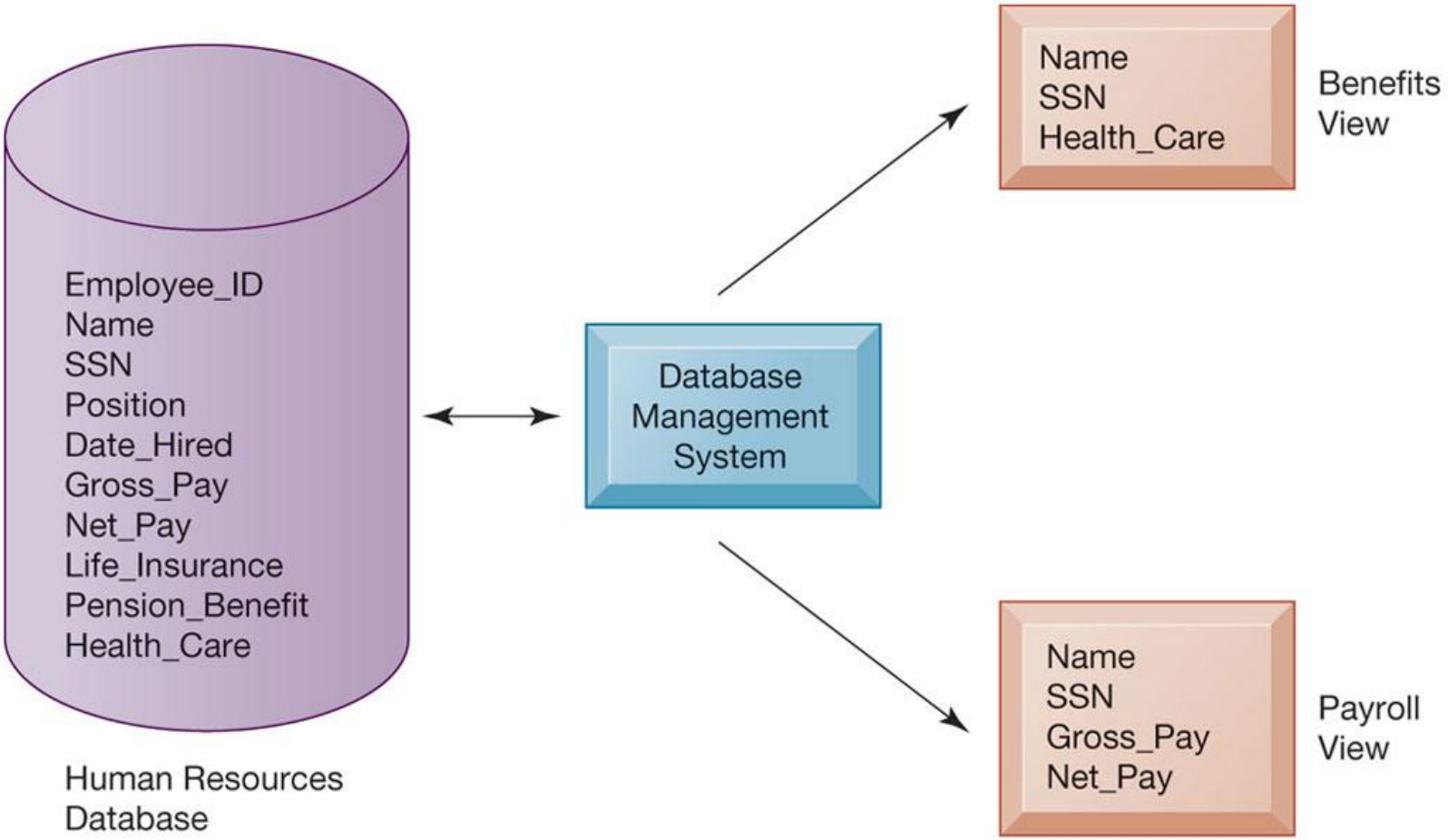
Figure 6.2 Traditional File Processing



Database Management Systems

- Database
  - Serves many applications by centralizing data and controlling redundant data
- Database management system (DBMS)
  - Interfaces between applications and physical data files
  - Separates logical and physical views of data
  - Solves problems of traditional file environment
    - Controls redundancy
    - Eliminates inconsistency
    - Uncouples programs and data
    - Enables organization to centrally manage data and data security

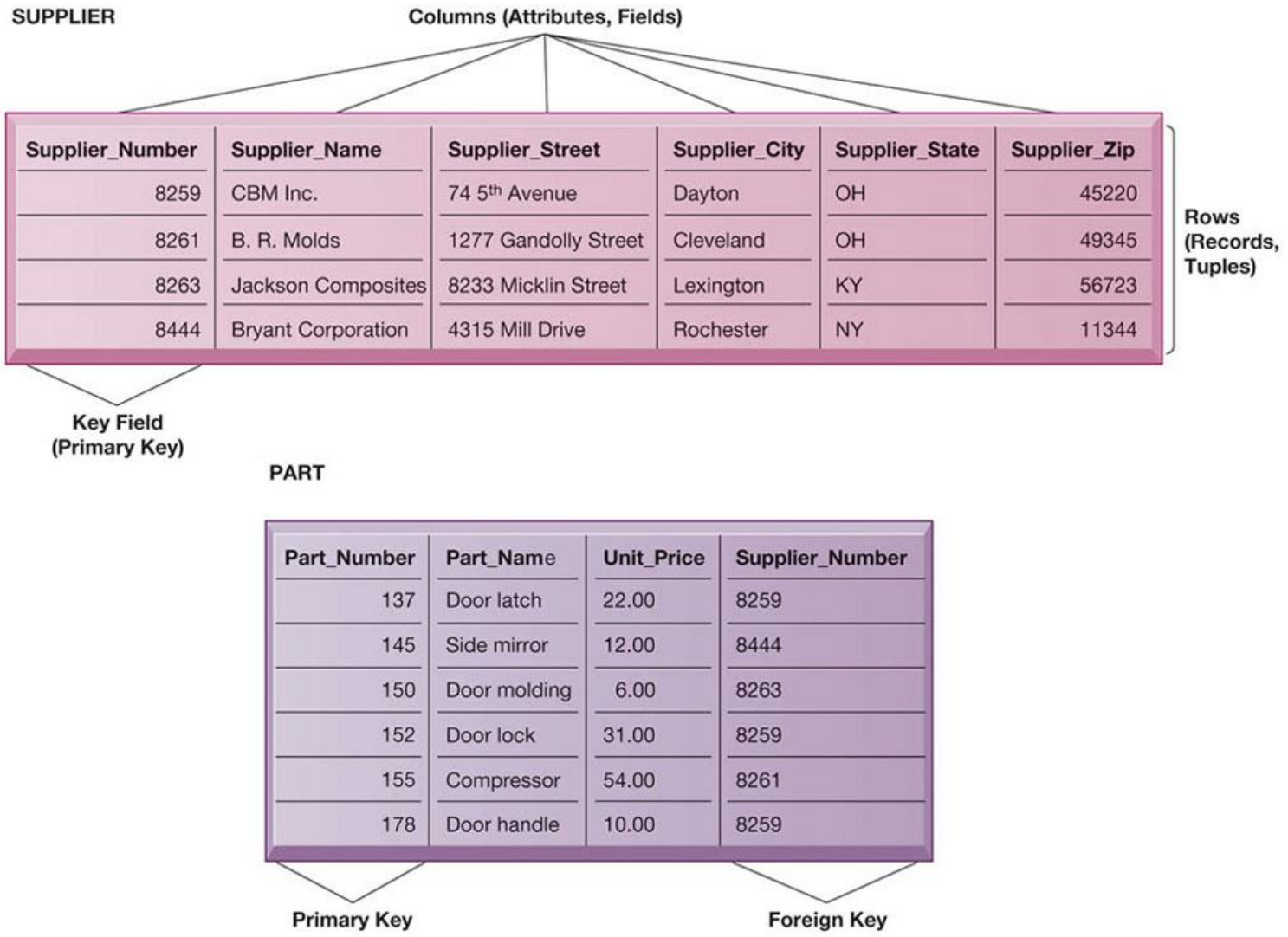
Figure 6.3 Human Resources Database with Multiple Views



Relational DBMS

- Represent data as two-dimensional tables
- Each table contains data on entity and attributes
- Table: grid of columns and rows
  - Rows (tuples): Records for different entities
  - Fields (columns): Represents attribute for entity
  - Key field: Field used to uniquely identify each record
  - Primary key: Field in table used for key fields
  - Foreign key: Primary key used in second table as look-up field to identify records from original table

Figure 6.4 Relational Database Tables



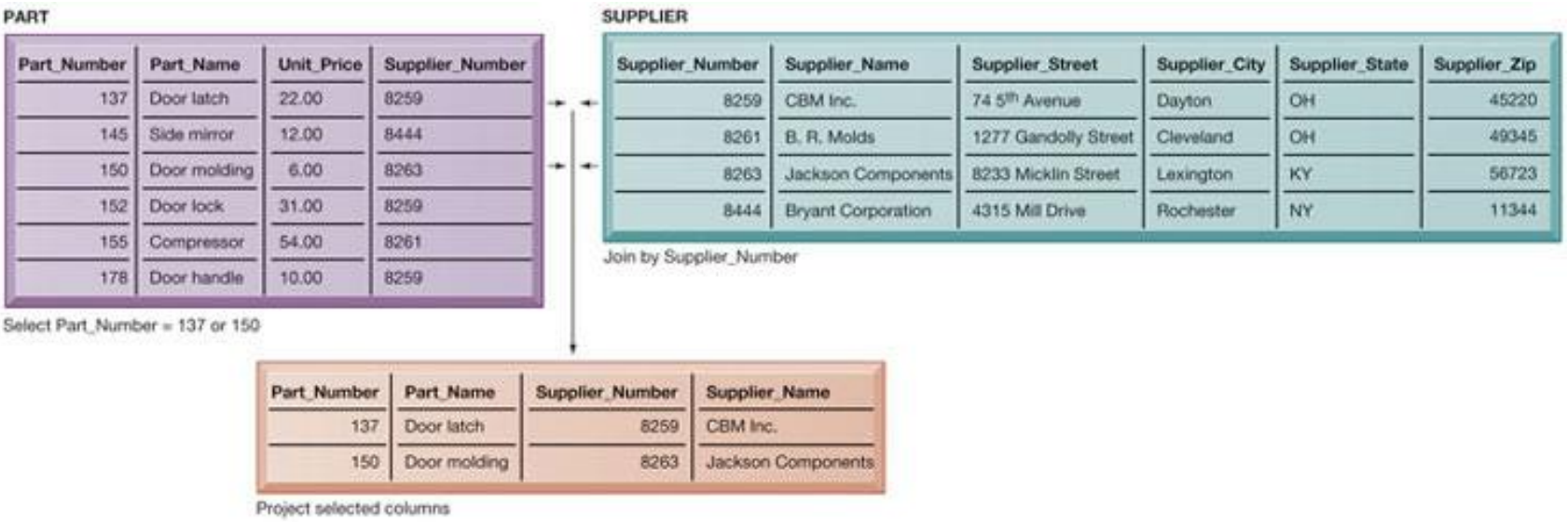
Supplier_Number	Supplier_Name	Supplier_Street	Supplier_City	Supplier_State	Supplier_Zip
8259	CBM Inc.	74 5th Avenue	Dayton	OH	45220
8261	B. R. Molds	1277 Gandolly Street	Cleveland	OH	49345
8263	Jackson Composites	8233 Micklin Street	Lexington	KY	56723
8444	Bryant Corporation	4315 Mill Drive	Rochester	NY	11344

Part_Number	Part_Name	Unit_Price	Supplier_Number
137	Door latch	22.00	8259
145	Side mirror	12.00	8444
150	Door molding	6.00	8263
152	Door lock	31.00	8259
155	Compressor	54.00	8261
178	Door handle	10.00	8259

Operations of a Relational DBMS

- Three basic operations used to develop useful sets of data
  - SELECT
    - Creates subset of data of all records that meet stated criteria
  - JOIN
    - Combines relational tables to provide user with more information than available in individual tables
  - PROJECT
    - Creates subset of columns in table, creating tables with only the information specified

Figure 6.5 The Three Basic Operations of a Relational DBMS



Capabilities of Database Management Systems

- Data definition capability
- Data dictionary
- Querying and reporting
  - Data manipulation language
    - Structured Query Language (SQL)
- Many DBMS have report generation capabilities for creating polished reports (Microsoft Access)

Figure 6.6 Access Data Dictionary Features

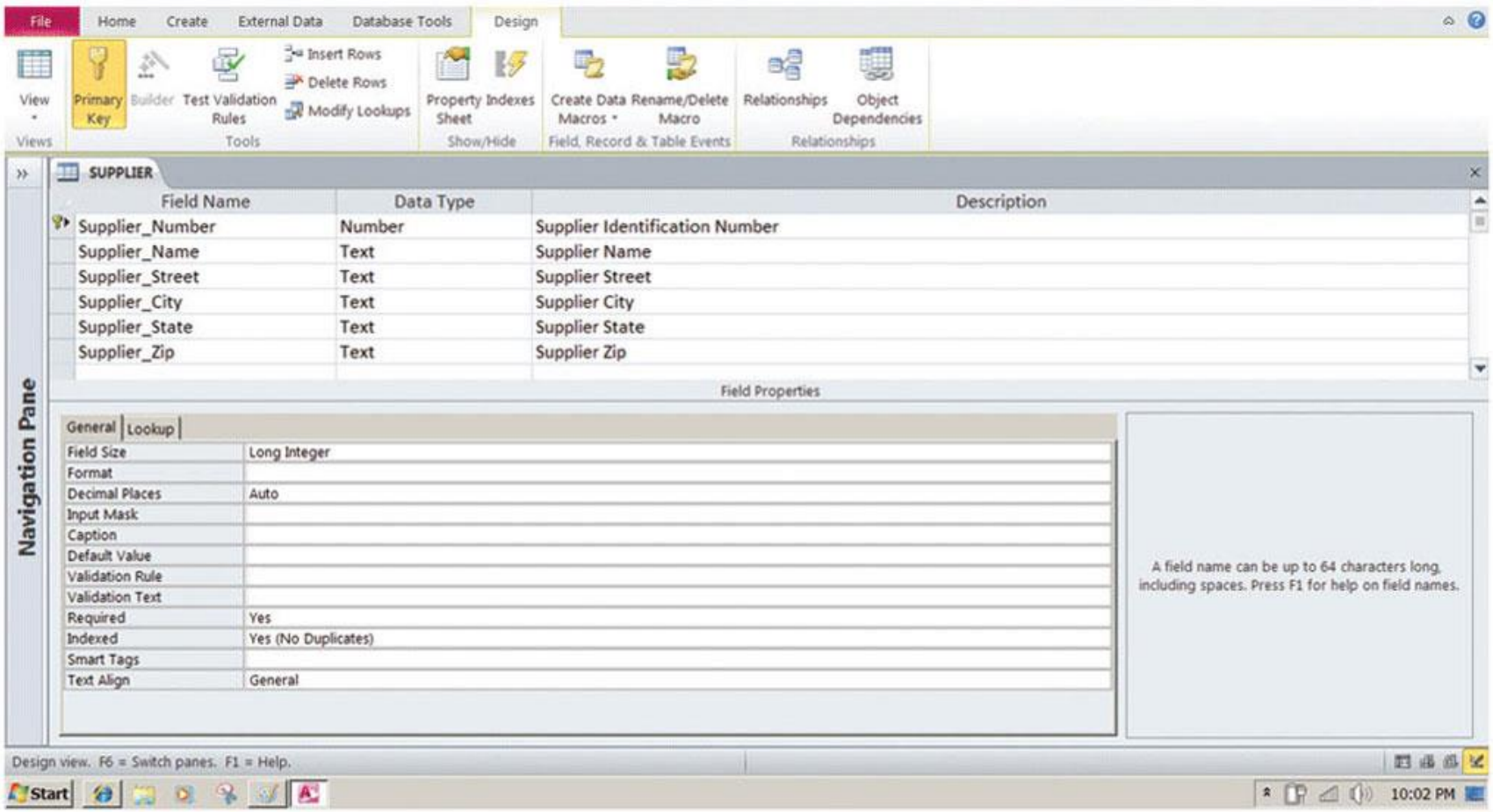
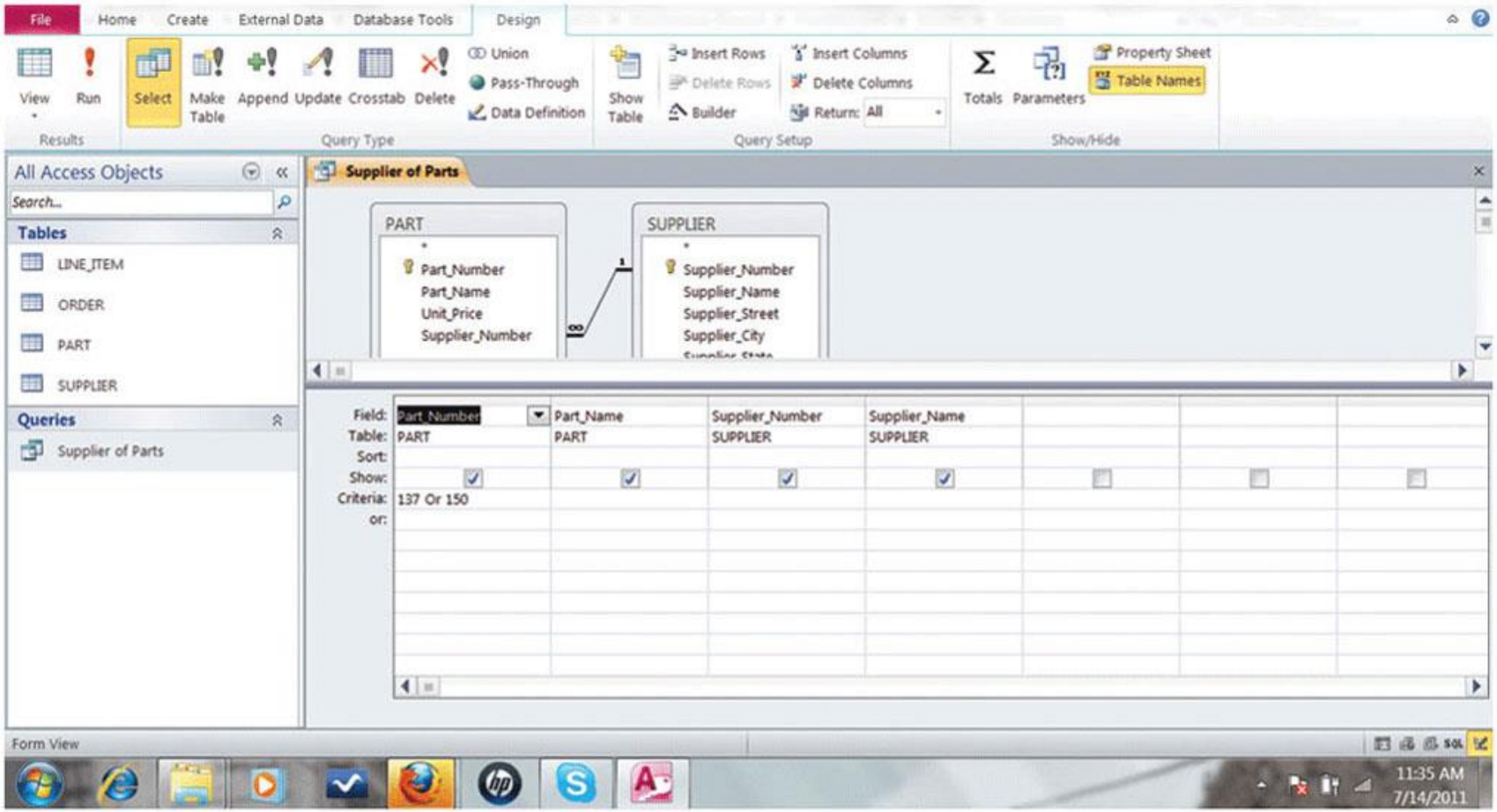


Figure 6.7 Example of an SQL Query

```
SELECT PART.Part_Number, PART.Part_Name, SUPPLIER.Supplier_Number,
SUPPLIER.Supplier_Name
FROM PART, SUPPLIER
WHERE PART.Supplier_Number = SUPPLIER.Supplier_Number AND
Part_Number = 137 OR Part_Number = 150;
```

Figure 6.8 An Access Query



Designing Databases

- Conceptual design vs. physical design
- Normalization
  - Streamlining complex groupings of data to minimize redundant data elements and awkward many-to-many relationships
- Referential integrity
  - Rules used by RDBMS to ensure relationships between tables remain consistent
- Entity-relationship diagram
- A correct data model is essential for a system serving the business well

Figure 6.9 An Unnormalized Relation for Order

Figure 6.10 Normalized Tables Created from Order

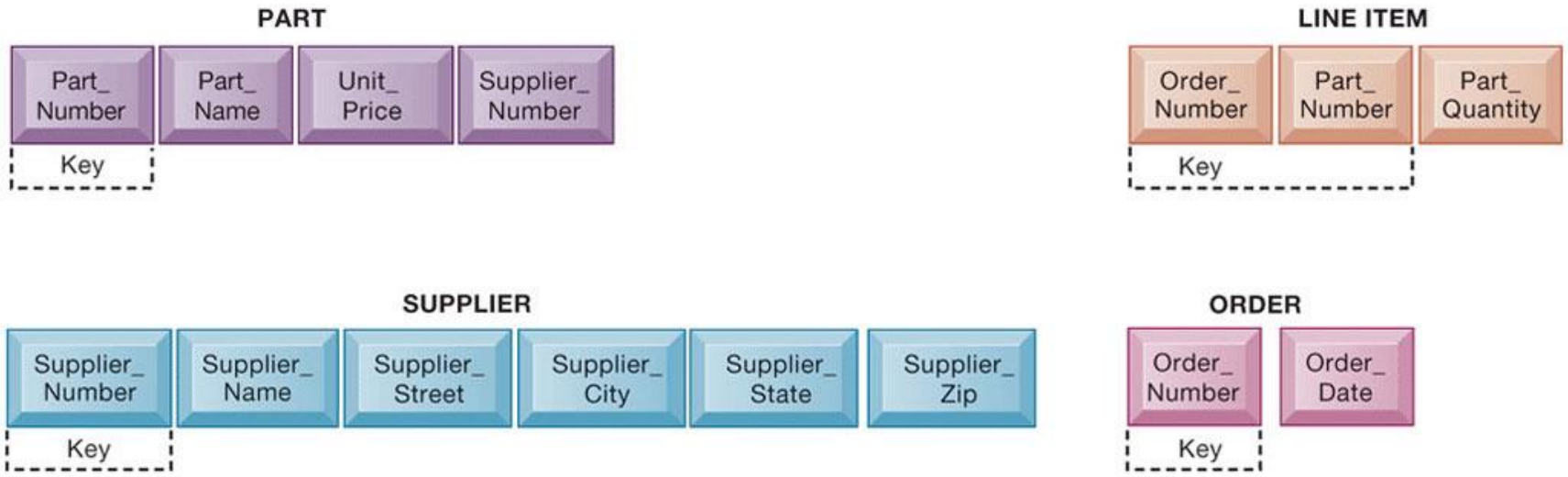
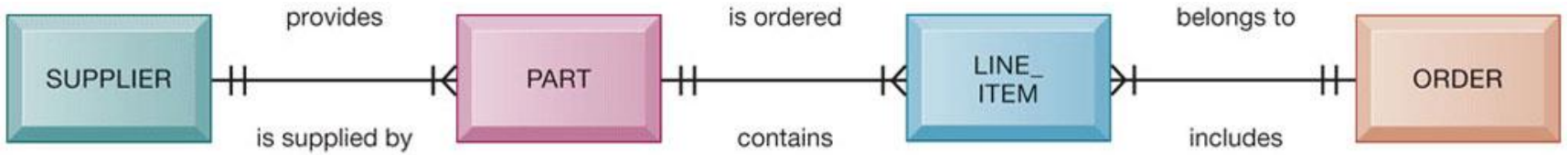


Figure 6.11 An Entity-Relationship Diagram



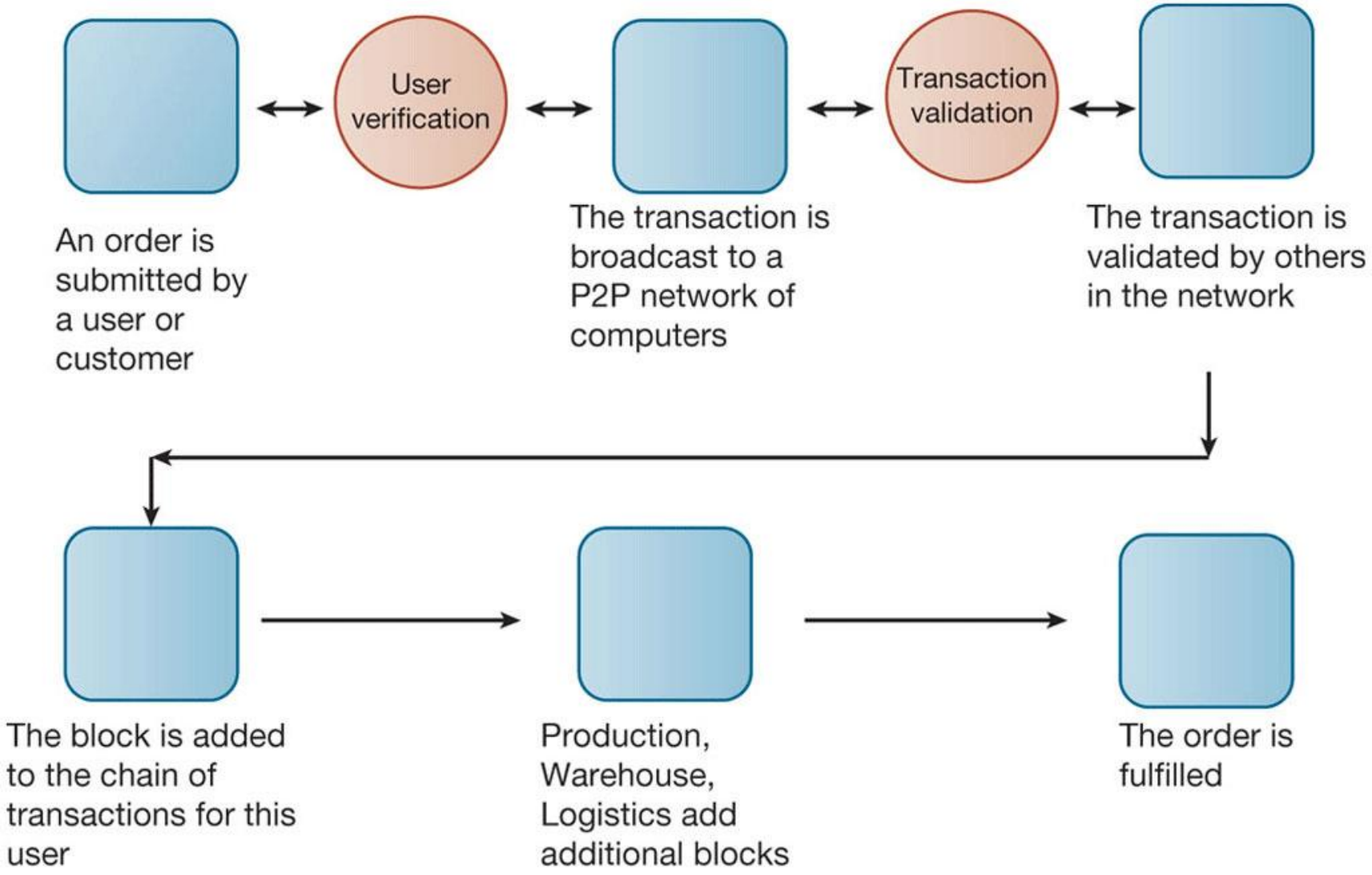
Non-Relational Databases and Databases in the Cloud

- Non-relational databases: “No SQL”
  - More flexible data model
  - Data sets stored across distributed machines
  - Easier to scale
  - Handle large volumes of unstructured and structured data
- Databases in the cloud
  - Appeal to start-ups, smaller businesses
  - Amazon Relational Database Service, Microsoft SQL Azure
  - Private clouds

Blockchain

- Distributed ledgers in a peer-to-peer distributed database
- Maintains a growing list of records and transactions shared by all
- Encryption used to identify participants and transactions
- Used for financial transactions, supply chain, and medical records
- Foundation of Bitcoin, and other crypto currencies

Figure 6.12 How Blockchain Works



The Challenge of Big Data

- Big data
  - Massive sets of unstructured/semi-structured data from web traffic, social media, sensors, and so on
- Volumes too great for typical DBMS
  - Petabytes, exabytes of data
- Can reveal more patterns, relationships and anomalies
- Requires new tools and technologies to manage and analyze

Business Intelligence Infrastructure (1 of 3)

- Array of tools for obtaining information from separate systems and from big data
- Data warehouse
  - Stores current and historical data from many core operational transaction systems
  - Consolidates and standardizes information for use across enterprise, but data cannot be altered
  - Provides analysis and reporting tools

Business Intelligence Infrastructure (2 of 3)

- Data marts
  - Subset of data warehouse
  - Typically focus on single subject or line of business
- Hadoop
  - Enables distributed parallel processing of big data across inexpensive computers
  - Key services
    - Hadoop Distributed File System (HDFS): data storage
    - MapReduce: breaks data into clusters for work
    - Hbase: No SQL database
  - Used Yahoo, NextBio

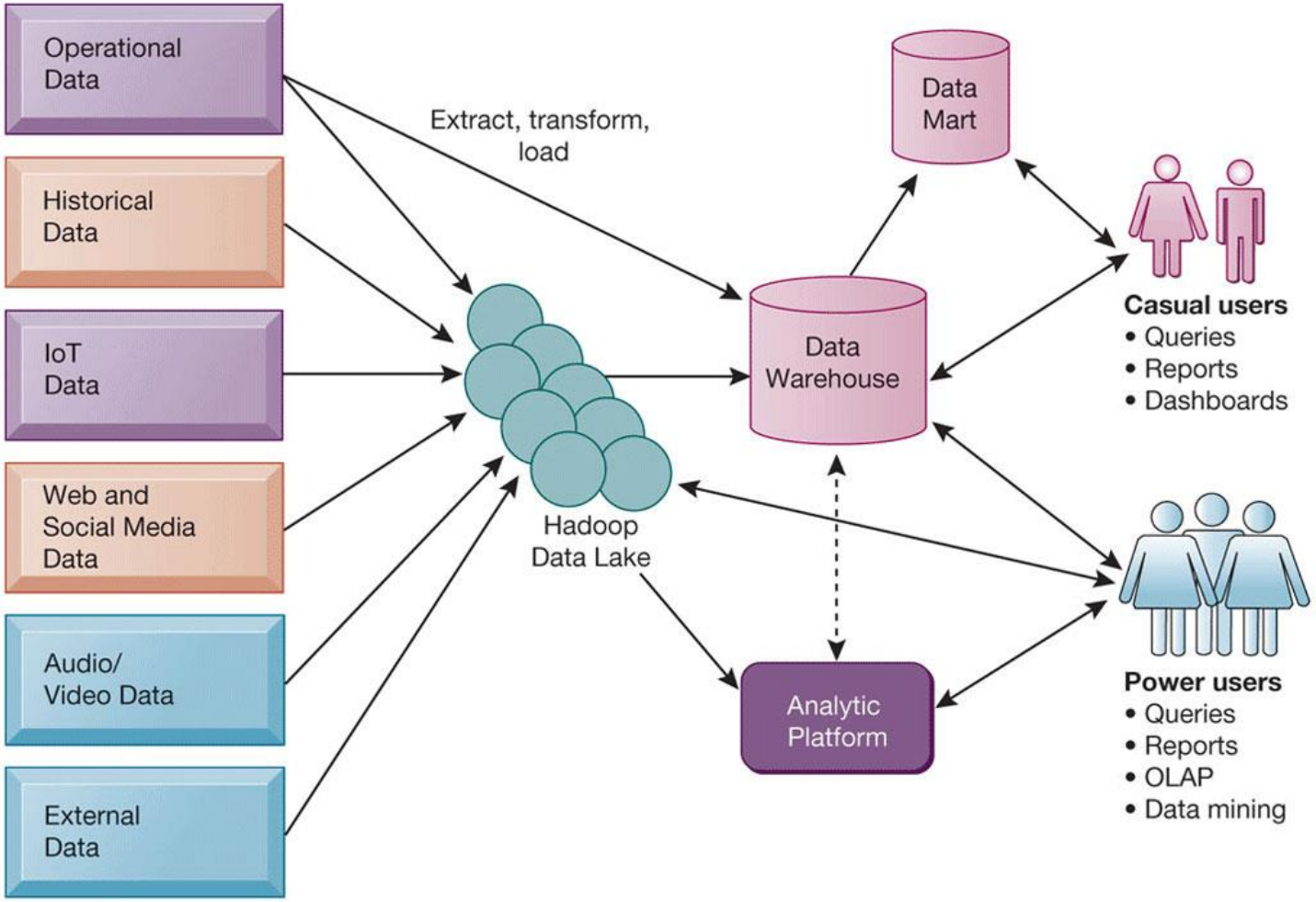
Business Intelligence Infrastructure (3 of 3)

- In-memory computing
  - Used in big data analysis
  - Uses computers main memory (RAM) for data storage to avoid delays in retrieving data from disk storage
  - Can reduce hours/days of processing to seconds
  - Requires optimized hardware
- Analytic platforms
  - High-speed platforms using both relational and non-relational tools optimized for large datasets

Interactive Session: Technology: Kraft Heinz Finds a New Recipe for Analyzing Its Data

- Class discussion
  - Identify the problem in this case study. To what extent was it a technology problem? Were any management and organizational factors involved?
  - How was information technology affecting business performance at Kraft Heinz?
  - How did new technology provide a solution to the problem? How effective was the solution?
  - Identify the management, organizational, and technology factors that had to be addressed in selecting and implementing Kraft-Heinz's new data warehouse solution.

Figure 6.13 Contemporary Business Intelligence Infrastructure



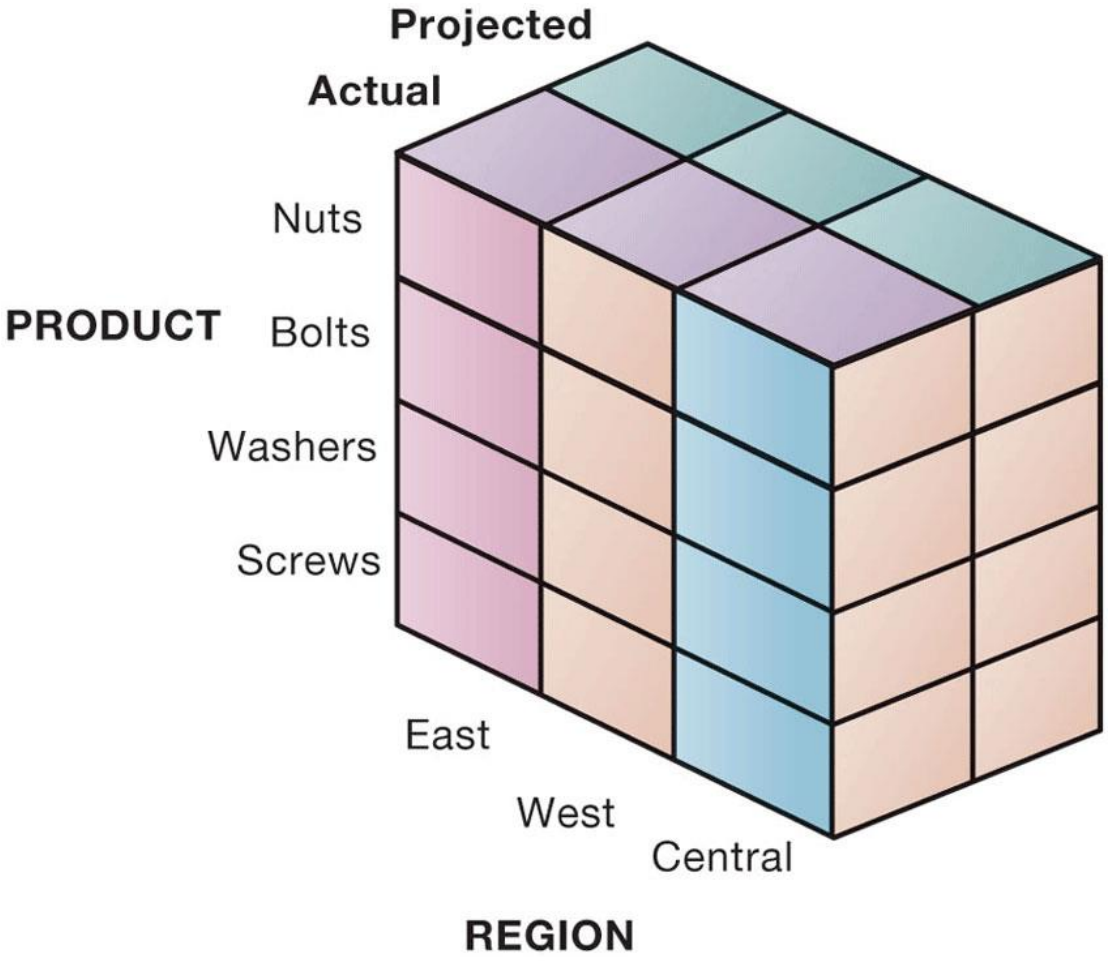
Analytical Tools: Relationships, Patterns, Trends

- Tools for consolidating, analyzing, and providing access to vast amounts of data to help users make better business decisions
  - Multidimensional data analysis (OLAP)
  - Data mining
  - Text mining
  - Web mining

Online Analytical Processing (OLAP)

- Supports multidimensional data analysis
  - Viewing data using multiple dimensions
  - Each aspect of information (product, pricing, cost, region, time period) is different dimension
  - Example: How many washers sold in the East in June compared with other regions?
- OLAP enables rapid, online answers to ad hoc queries

Figure 6.14 Multidimensional Data Model



Data Mining

- Finds hidden patterns, relationships in datasets
  - Example: customer buying patterns
- Infers rules to predict future behavior
- Types of information obtainable from data mining:
  - Associations
  - Sequences
  - Classification
  - Clustering
  - Forecasting

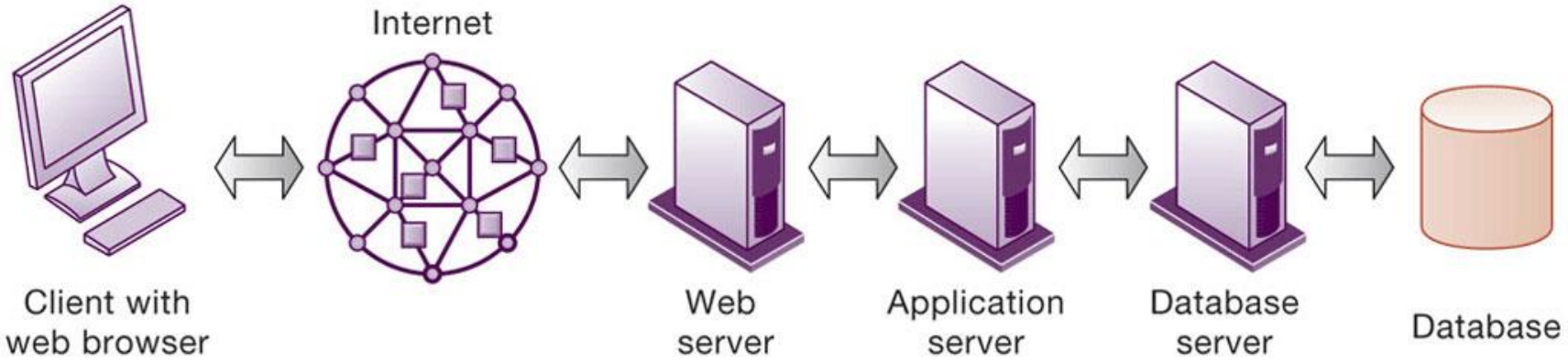
Text Mining and Web Mining

- Text mining
  - Extracts key elements from large unstructured data sets
  - Sentiment analysis software
- Web mining
  - Discovery and analysis of useful patterns and information from web
  - Web content mining
  - Web structure mining
  - Web usage mining

Databases and the Web

- Many companies use the web to make some internal databases available to customers or partners
- Typical configuration includes:
  - Web server
  - Application server/middleware/CGI scripts
  - Database server (hosting DBMS)
- Advantages of using the web for database access:
  - Ease of use of browser software
  - Web interface requires few or no changes to database
  - Inexpensive to add web interface to system

Figure 6.15 Linking Internal Databases to the Web



Establishing an Information Policy

- Firm's rules, procedures, roles for sharing, managing, standardizing data
- Data administration
  - Establishes policies and procedures to manage data
- Data governance
  - Deals with policies and processes for managing availability, usability, integrity, and security of data, especially regarding government regulations
- Database administration
  - Creating and maintaining database

Ensuring Data Quality

- More than 25 percent of critical data in Fortune 1000 company databases are inaccurate or incomplete
- Before new database is in place, a firm must:
  - Identify and correct faulty data
  - Establish better routines for editing data once database in operation
- Data quality audit
- Data cleansing

Interactive Session: Organizations: Databases Where the Data Aren't There

- Class discussion
  - Define the problem described in this case. How serious a problem is it?
  - What management, organization, and technology factors contributed to this problem?
  - What is the political and social impact of incomplete recordkeeping in the FBI NCIC and NICS databases?

How Will MIS Help My Career?

- The Company: Mega Midwest Power
- Position Description: Entry-level data analyst
- Job Requirements
- Interview Questions
- Author Tips

Copyright