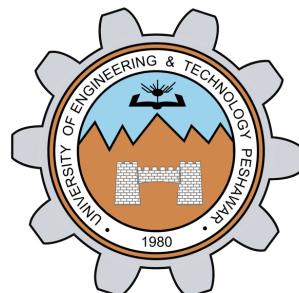


Database Management System

Lecture 1: Introduction

Prof. Dr. Sadeeq Jan

Department of Computer Systems Engineering
University of Engineering and Technology Peshawar



Instructor

Prof. Dr. Sadeeq Jan

- Director, National Centre for Cyber Security
- Professor, DCSE

Qualification:

- *BSc Engg (Computer Information Systems Engineering) - UET Peshawar*
- *IZAZ-E-SABQAT (Presidential Award), Gold Medalist, CGPA 3.98/4.00*
- *MS (Info/Comm System Security) with Excellent grade - KTH Sweden*
- *PhD (Security Testing) with Excellent grade – University of Luxembourg*

Contact:

sadeeqjan@uetpeshawar.edu.pk

sadeeqjan@yahoo.com

Grading Criteria

Exams

- Midterm: 20%
- Final Term: 50%

Sessional

- Assignments/Quizzes: 5%
- Semester Project 25%

Books

- Modern Database Management, Jeffrey A. Hoffer, Mary B. Prescott, Fred R. McFadden
- R. Ramakrishnan and J. Gehrke. Database Management Systems. latest Edition. WCB/McGraw-Hill 1998
- Database System Principles, Design, and Implementation by CATHERINE M. RICARDO
- Database System Concepts by Abraham Silberschatz, Henry F. Korth S. Sudarshan, latest Edition
- P. O'Neil and E. O'Neil. Database: Principles, Programming, and Performance. latest Edition. Morgan Kaufmann 2000.

Objectives

- Understand the fundamental concepts of database systems
- Understand Relational Database Management System in particular to analyse, design, and implement the Relational Database Application.
- Determine the user requirements for data of given problem and develop Conceptual Data Model
- Transform Conceptual Data Model into Logical and Physical Specifications to guide database implementation.
- Efficiently use Structure Query Language (SQL).

Course Learning Outcomes (CLOs)		Level of Learning (Bloom's Taxonomy)	Program Learning Outcomes (PLOs)
CLO #			
CLO- 1	Describe the fundamental concepts of database systems in general and Relational Database Management System in particular.	Cog-2 (Comprehension)	PLO1 (Engineering Knowledge)
CLO- 2	Develop data models to represent user requirements for data of given problem using Conceptual Data Modelling.	Cog-5(Synthesis)	PLO3 (Design/Development of Solutions)
CLO- 3	Develop Logical and Physical Specifications to guide database implementation.	Cog-5 (Synthesis)	PLO3 (Design/Development of Solutions)
CLO- 4	Use Structure Query Language (SQL) efficiently to implement Physical Specification in Database Management Systems such as MySQL/Oracle.	Cog-3 (Application)	PLO5 (Modern Tool Usage)

Mapping of CLOs with Course Assessment Tools

Course	CLO1	CLO2	CLO3	CLO4
Assessment Tools				
Assignments		✓		
Quizzes			✓	
Mid Term	✓	✓		
Final Term			✓	
Semester Project				✓

Weekwise Content

Week	Topics
01	Introduction to the Database and Database Management System Database definitions, Databases & Conventional File Processing Systems, Advantages of Database Approach, Functions of DBMS, and Components of Database Environment
02	Data Models Logical, Physical, and Conceptual view of Data; Types of Data Models, Types of Database Design; System Development Life Cycle, The Database Life Cycle; Front-end and Back-end Databases
03	Entity-Relationship Modeling ER Model Constructs, Classification of Entity Types, Attributes, and Relationships
04	Entity-Relationship Modeling (Cont.) Super & Subtype Entities, Relationships and Cardinalities between Entities, Conceptual & Logical Database Design
05	Relational Data Model Converting E-R to Relations, Relational Algebra, and Relational Calculus
06	Functional Dependency and Normalization Functional Dependency, Second Normal Form, Third Normal Form, Boyce-Codd Normal Form, Higher Normal Form, and Normalization Summary
07	Structured Query Language Basic Commands and Functions of SQL, Data Definition Language, and Data Manipulation Language
08	Structured Query Language (Cont.) SQL Queries, Characteristics/Types of Views, Relational Operators, and SQL Join

Cont.....

Week	Topics
09	Structured Query Language (Cont.) Integrity Constraints, Updating Multiple Tables, SQL Sub-Queries, and SQL Built-in Functions
10	Advanced Structured Query Language Stored Programs, Stored Procedures, Stored Functions, and Triggers
11	Advanced Structured Query Language (Cont.) Prepared Statement, General-Purpose Stored Procedures, Cursors, Transactions, Locks, and Security
12	Transaction Processing Concepts The Concept of Transaction, Transaction & Schedules, Concurrent Execution of Transactions, Lock-based Concurrency Control, Recovery Management, and Security
13	File Organization and Indexing Data Storage Principles, Alternative File Organization, Indexes, Index Classification, Tree-Structured Indexes, Range Searches, ISAM, B+-Trees, Bulk Loading
14	Query Execution Query Processor: Architecture, Query Parser, External Sorting, Hash Join, Query Optimization
15	Database Architectures Client/Server, Distributed Databases NOSQL Databases The Database Warehouse Concept of Data Warehouse, Its Architecture & Characteristics

Introduction to the Database

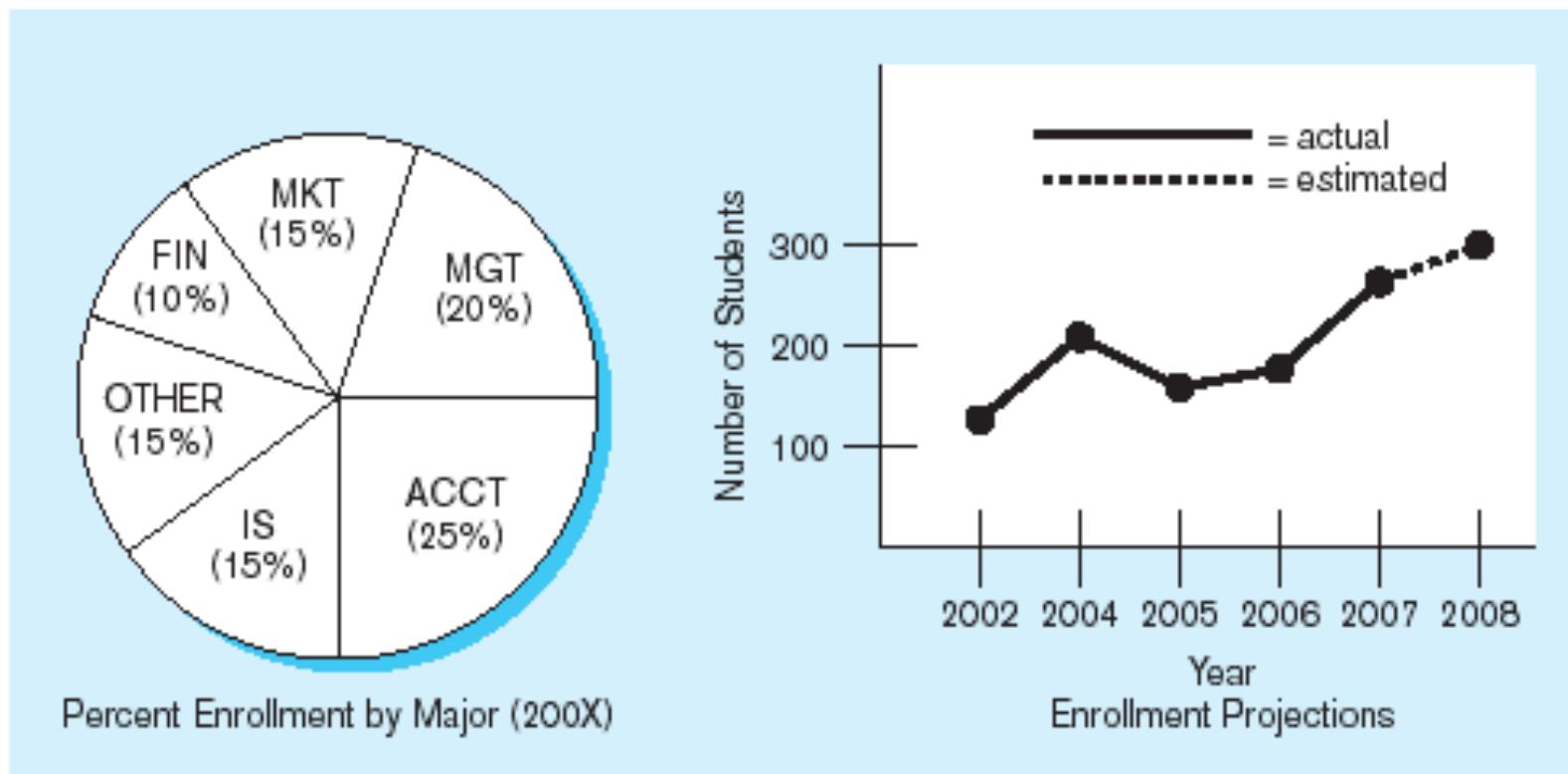
- Database Definitions
- Databases & Traditional File Processing Systems
- Advantages of Database Approach
- Costs & Risks of the Database Approach
- Functions of Database Management Systems (DBMS)
- Range of Database Applications
- Components of the Database Environment
- Evolution of Database Systems

Data: facts, figures,

Class Roster			
Course:	MGT 500 Business Policy	Semester:	Spring 200X
Section:	2		
Name	ID	Major	GPA
Baker, Kenneth D.	324917628	MGT	2.9
Doyle, Joan E.	476193248	MKT	3.4
Finkle, Clive R.	548429344	PRM	2.8
Lewis, John C.	551742186	MGT	3.7
McFerran, Debra R.	409723145	IS	2.9
Sisneros, Michael	392416582	ACCT	3.3

Context helps users understand data

Figure 1-1b Summarized data



Graphical displays turn data into useful information that managers can use for decision making and interpretation

Definitions

- **Data:** stored representations of meaningful objects and events
 - Structured: numbers, text, dates
 - Unstructured: images, video, documents
- **Information:** data processed to increase knowledge in the person using the data
- **Metadata:** data that describes the properties and context of user data
- **Database:** organized collection of logically related data

File-Processing Systems

- Information system that stores groups of records in separate files
- Separate applications for individual units of an organization
- Each application has its own set of files

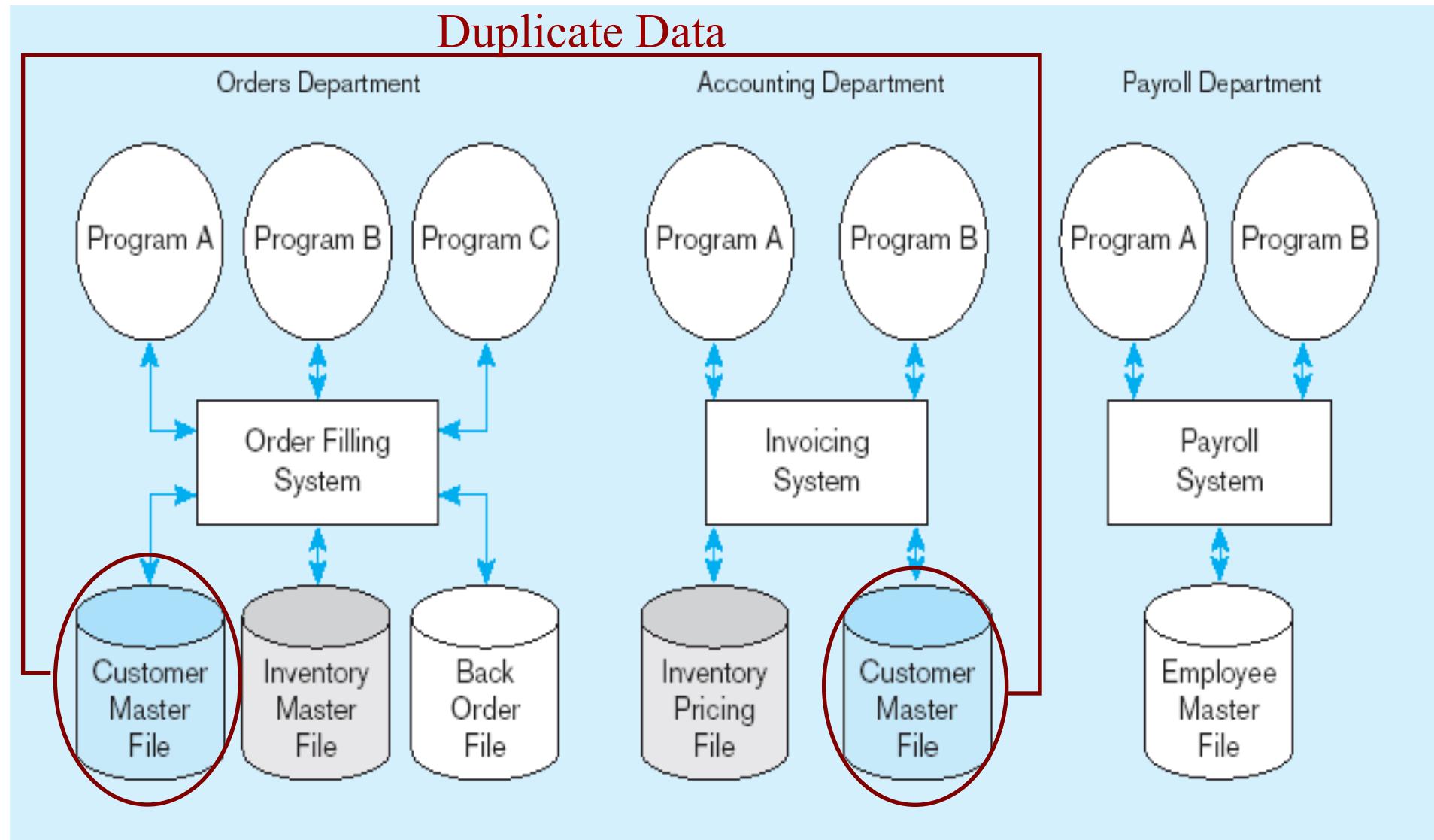
Disadvantages of File Processing

- Program-Data Dependence
- Duplication of Data
- Limited Data Sharing
- Lengthy Development Times
- Excessive Program Maintenance

Problems with Data Dependency

- Each application programmer must maintain his/her own data
 - changing the customer address field length in the customer file will require changes to all programs that uses customer file
- Each application program needs to include code for the metadata of each file
- Each application program must have its own processing routines for reading, inserting, updating, and deleting data
- Lack of coordination and central control
- Non-standard file formats

Figure 1-3 Old file processing systems at Pine Valley Furniture Company



Problems with Data Redundancy (Duplication)

- Different systems/programs have separate copies of the same data
- Waste of space to have duplicate data
- Causes more maintenance headaches
- The biggest problem:
 - **Data changes in one file could cause inconsistencies**
 - Compromises in *data integrity*
- Synonym (same data item having different names)
- Homonym (different data items having the same name)

Limited Data Sharing

- No centralized control of data
- Each application has its own private files
- Generating a report which requires data from several incompatible files in separate systems, is difficult

Lengthy Development Times

- Each new application has to be started from scratch by designing new file formats and descriptions

Excessive Program Maintenance

- Data definitions are hard coded in the program logic
- 80% of information systems budget

SOLUTION: The DATABASE Approach

- Central repository of shared data
- Data is managed by a controlling agent
- Stored in a standardized, convenient form

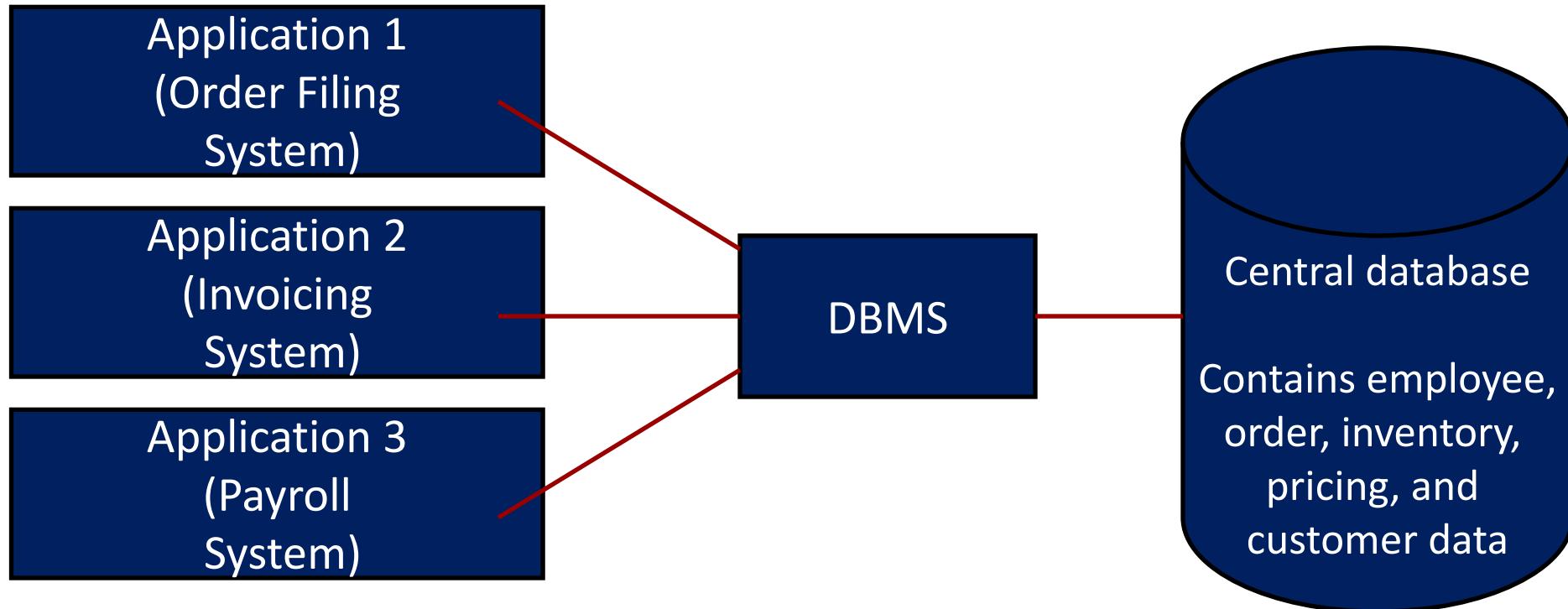
Requires a Database Management System (DBMS)

Definitions (Cont.)

- Database System: system consisting of a database, a DBMS, hardware and people
- Database Design: The design of the database structure that will be used to store and manage data (not the design of the DBMS)
- Data Dictionary (DD) or Metadata: describes the characteristics of data stored in a database and the inter-relationships among data

Database Management System

A software system that is used to create, maintain, and provide controlled access to user databases



DBMS manages data resources like an operating system manages hardware resources

Table 1-1 Example Metadata for Class Roster

<i>Data Item</i>		<i>Value</i>				
Name	Type	Length	Min	Max	Description	Source
Course	Alphanumeric	30			Course ID and name	Academic Unit
Section	Integer	1	1	9	Section number	Registrar
Semester	Alphanumeric	10			Semester and year	Registrar
Name	Alphanumeric	30			Student name	Student IS
ID	Integer	9			Student ID (SSN)	Student IS
Major	Alphanumeric	4			Student major	Student IS
GPA	Decimal	3	0.0	4.0	Student grade point average	Academic Unit

Descriptions of the properties or characteristics of the data, including data types, field sizes, allowable values, and data context

Figure 1-2: Three File Processing

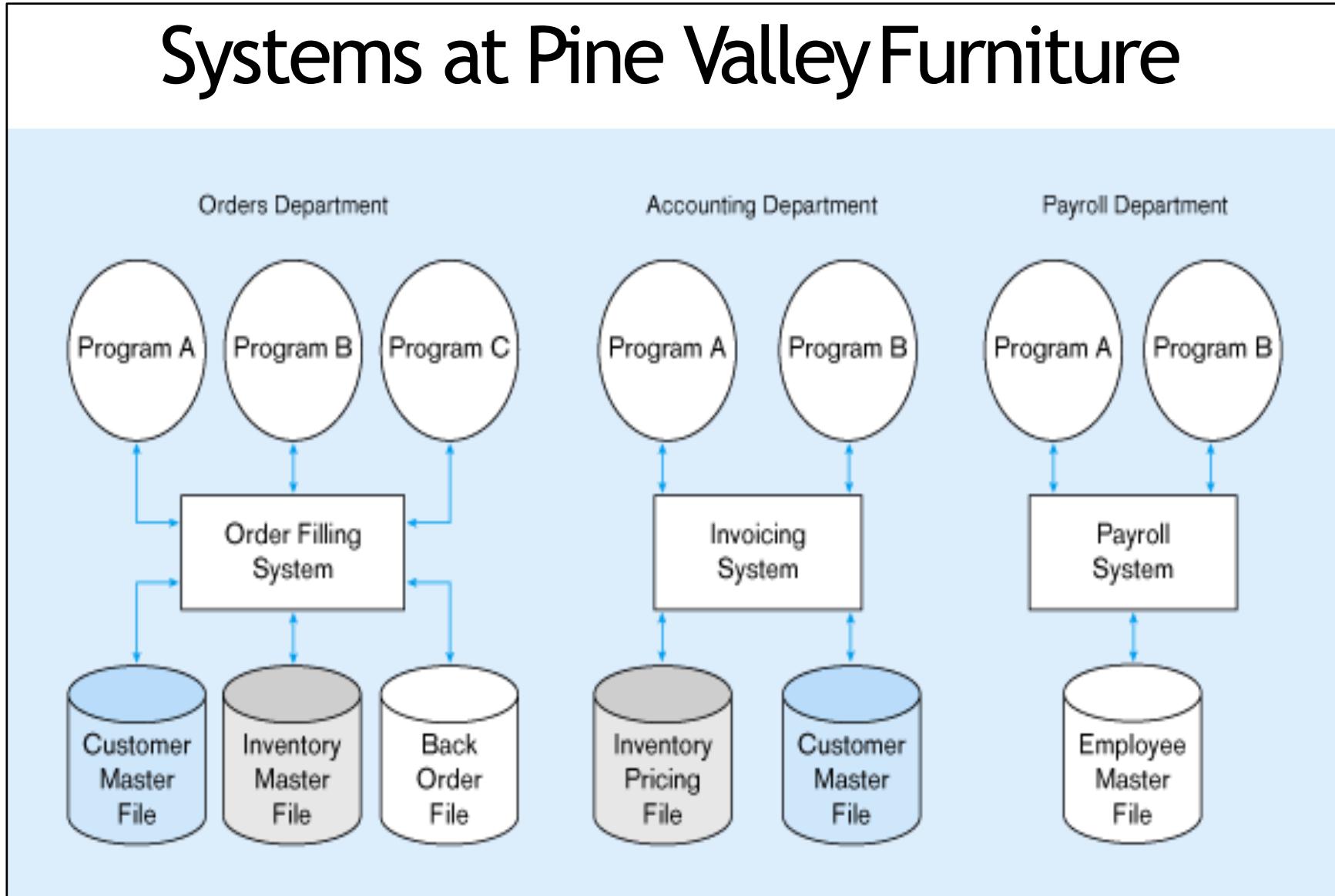
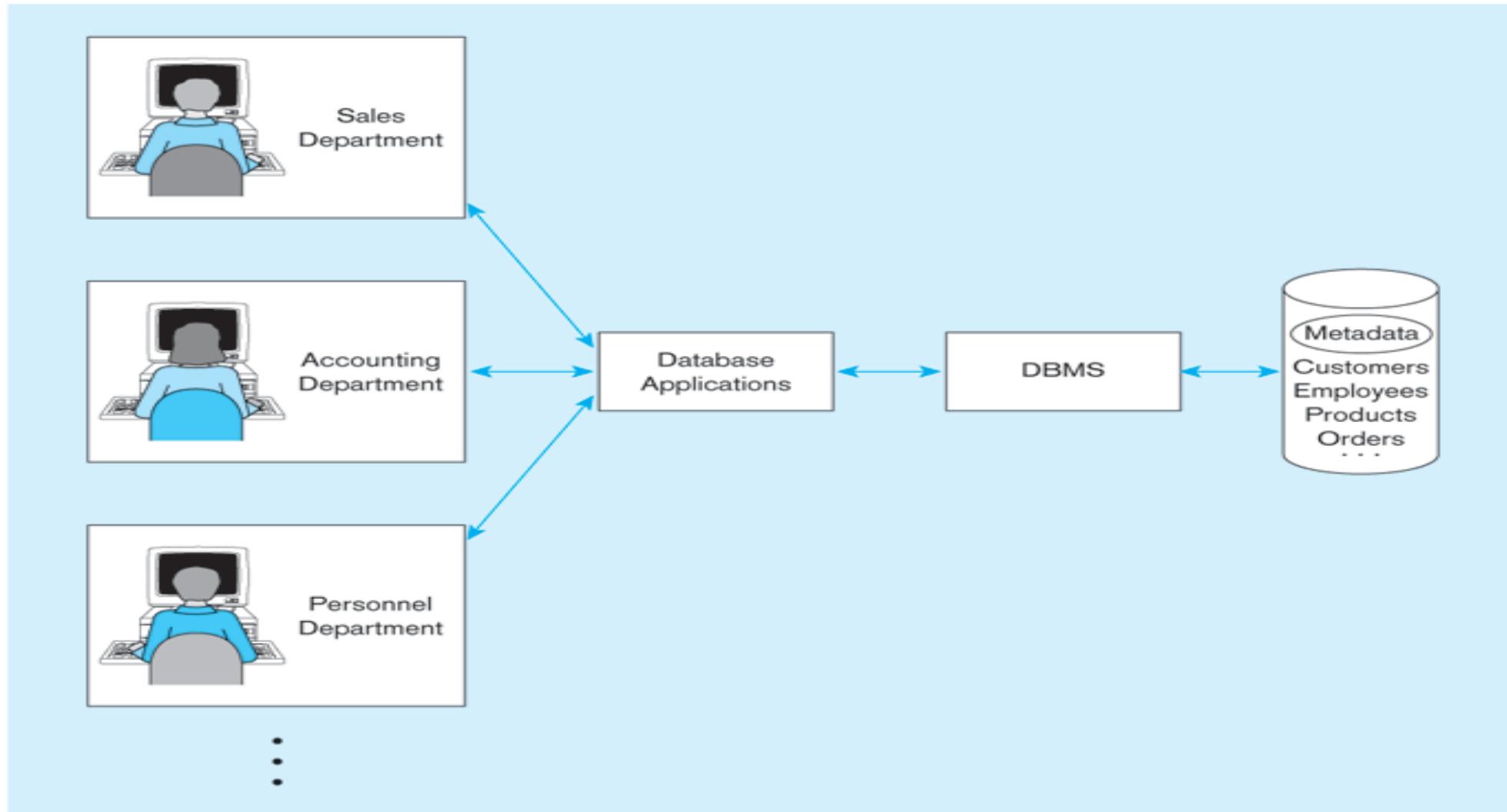


Figure 1-3 Database approach at Pine Valley Furniture Company



Advantages of the Database Approach

- Program-data independence
- Minimal data redundancy
- Improved data consistency
- Improved data sharing
- Increased application development productivity
- Enforcement of standards
- Improved data quality
- Improved data accessibility and responsiveness
- Reduced program maintenance
- Improved decision support
- Security, Backup/Recovery, Concurrency

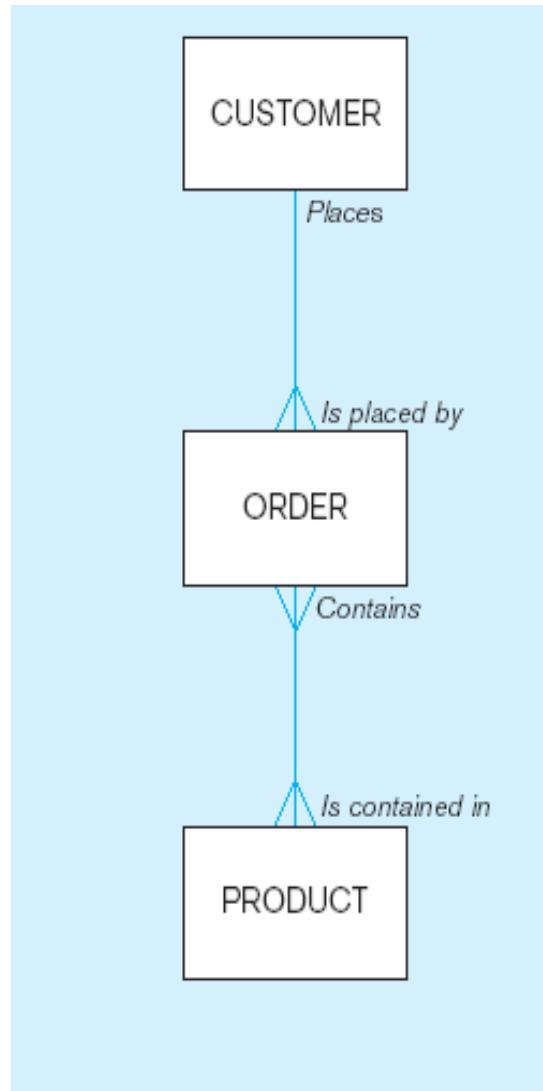
Costs and Risks of the Database Approach

- Up-front Costs
 - Installation Management Cost and Complexity
 - Conversion costs
- Ongoing Costs
 - New, specialized personnel
 - Need for explicit backup and recovery
- Organizational conflict

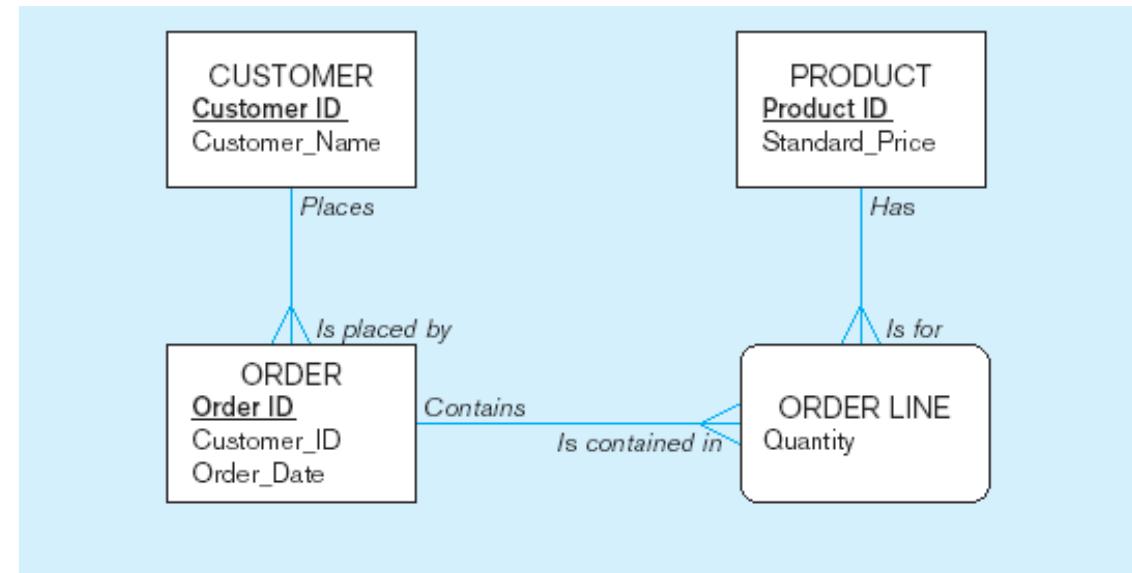
Elements of the Database Approach

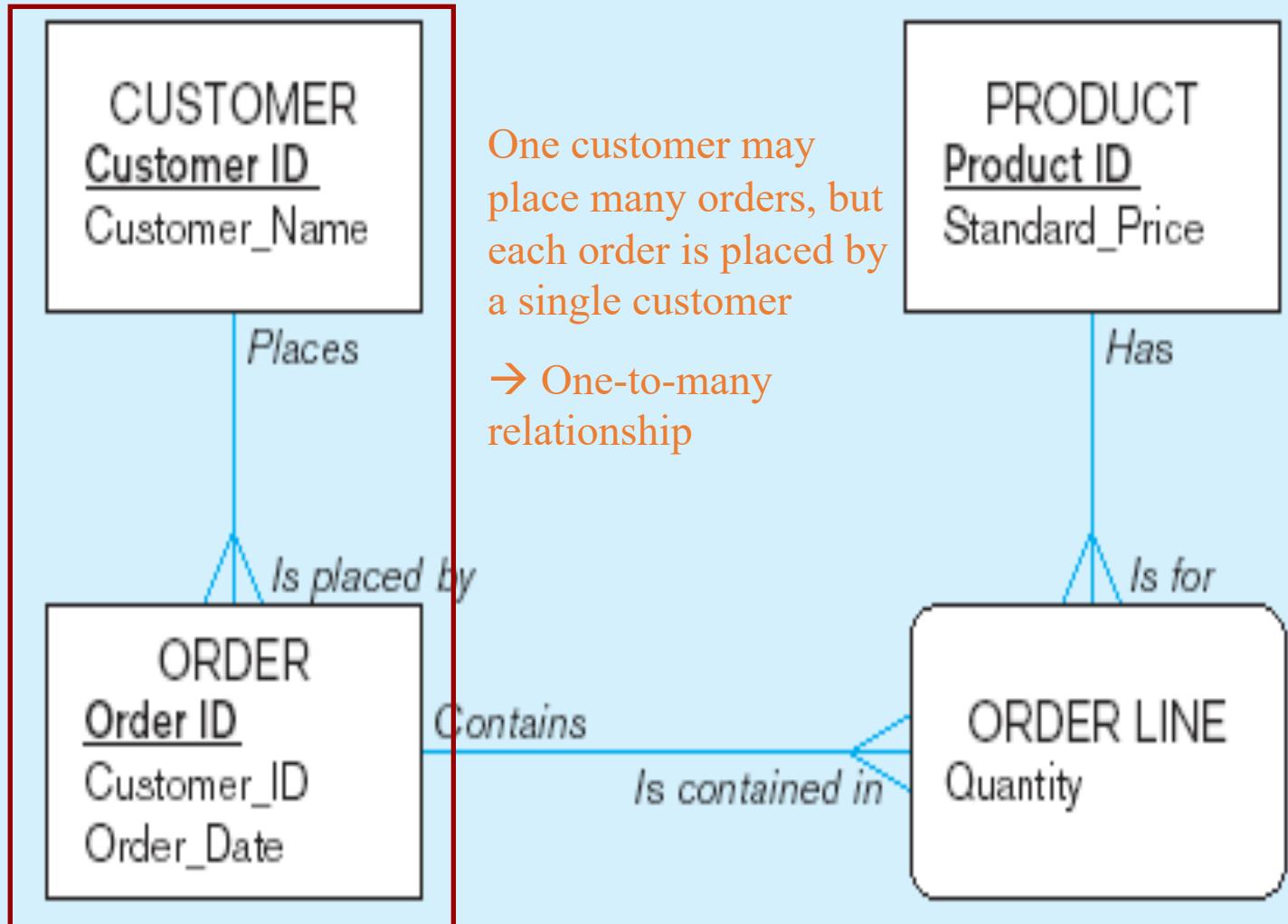
- Data models
 - Graphical system capturing nature and relationship of data
 - Enterprise Data Model—high-level entities and relationships for the organization
 - Project Data Model—more detailed view, matching data structure in database or data warehouse
- Relational Databases
 - Database technology involving tables (relations) representing entities and primary/foreign keys representing relationships
- Use of Internet Technology
 - Networks and telecommunications, distributed databases, client-server, and 3-tier architectures
- Database Applications
 - Application programs used to perform database activities (create, read, update, and delete) for database users

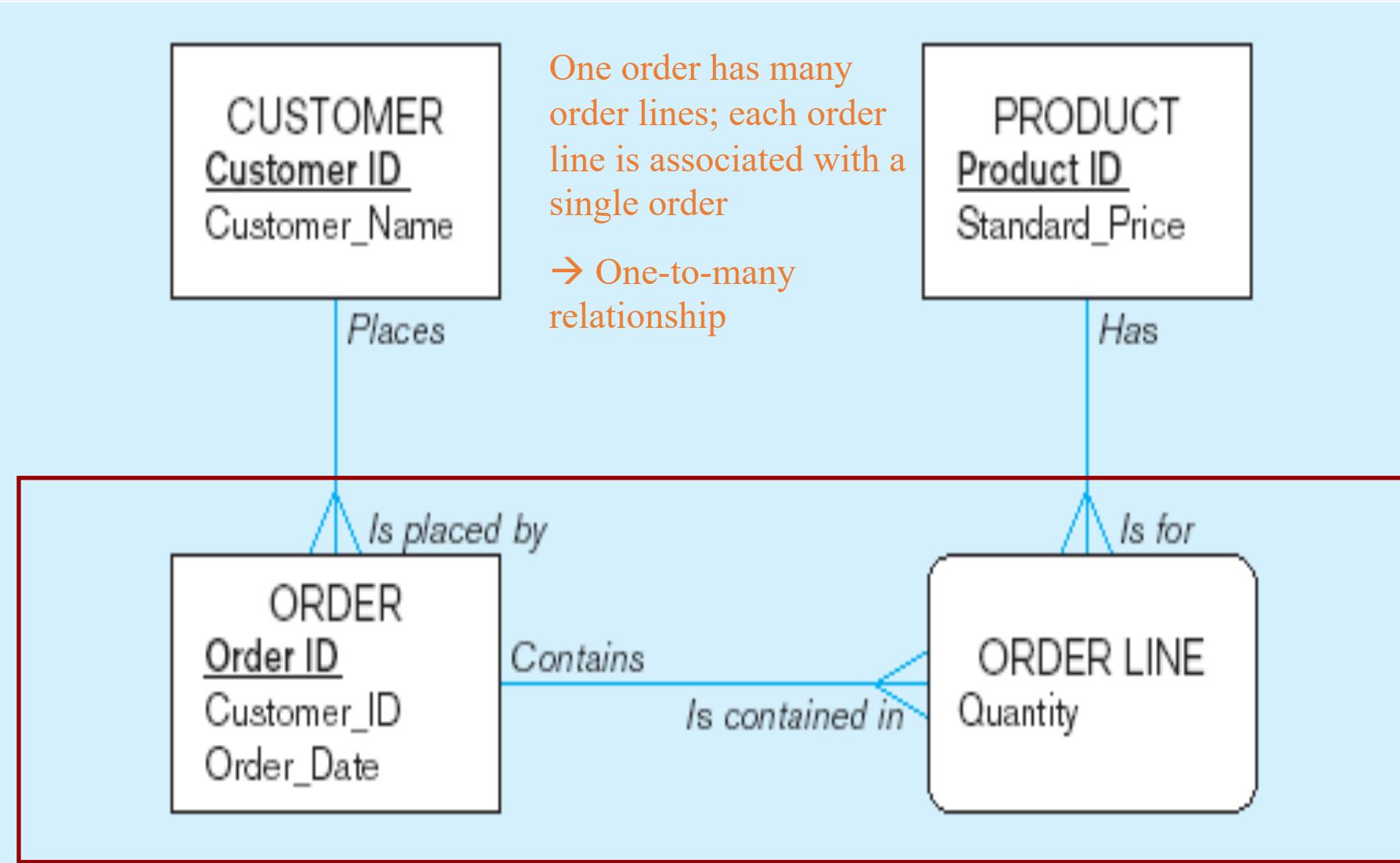
Segment of an Enterprise Data Model



Segment of a Project-Level Data Model

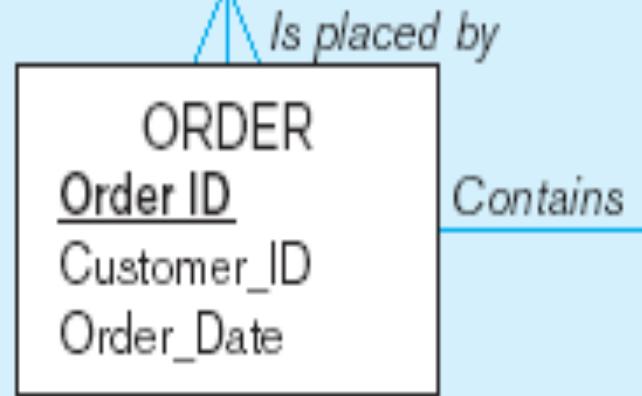
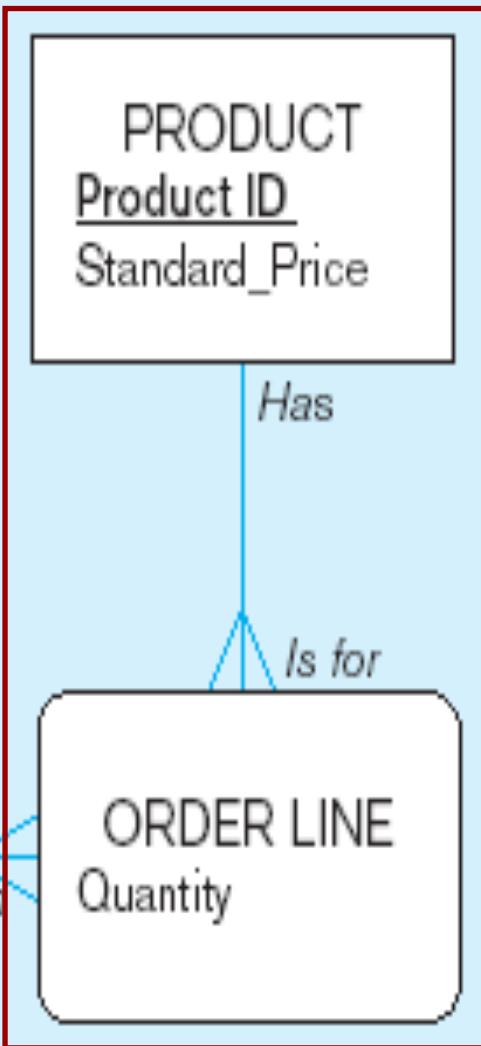






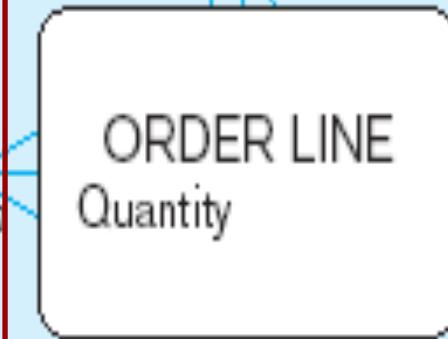


One product can be in many order lines, each order line refers to a single product
→ One-to-many relationship



Contains

Is contained in



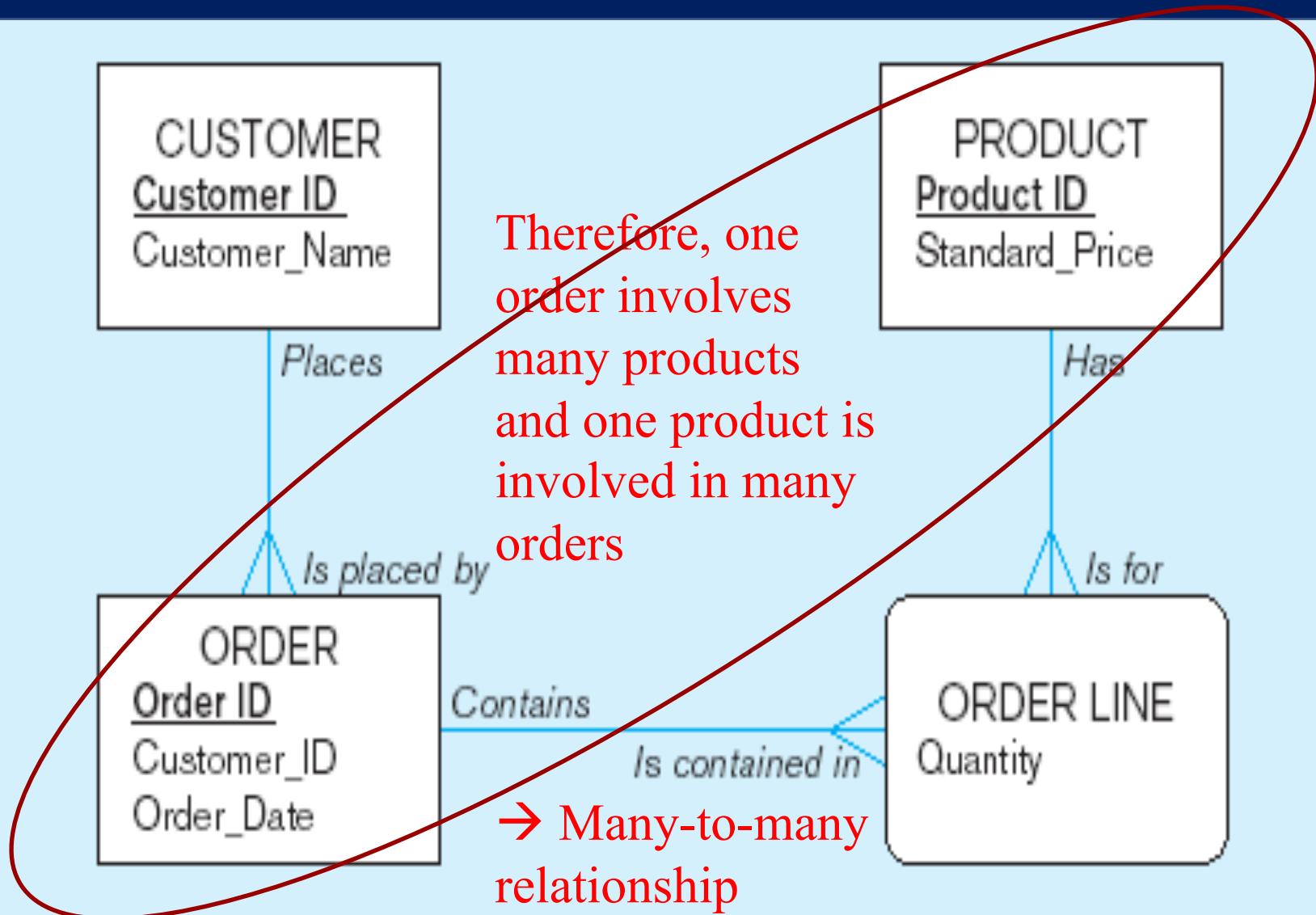


Figure 1-5a Four relations (Pine Valley Furniture Company) - ORDER and Order_line tables

The screenshot shows the Microsoft Access application interface with two open tables:

- ORDER_t : Table**: This table contains 10 records. The columns are Order_ID, Order_Date, and Customer_ID. The Order_ID values range from 1001 to 1010. The Customer_ID values are 1, 8, 15, 5, 3, 2, 11, 12, 4, and 1 respectively.
- Order_line_t : Table**: This table contains 18 records. The columns are Order_ID, Product_ID, and Ordered_Quantity. The Order_ID values are 1001 through 1010, and 1006 through 1009. The Product_ID values are 1, 2, 4, 3, 3, 6, 8, 4, 4, 5, 7, 1, 2, 3, 8, 4, 7, 8, and 0. The Ordered_Quantity values are 2, 2, 1, 5, 3, 2, 2, 4, 1, 2, 2, 3, 2, 3, 2, 3, 3, 10, and 0.

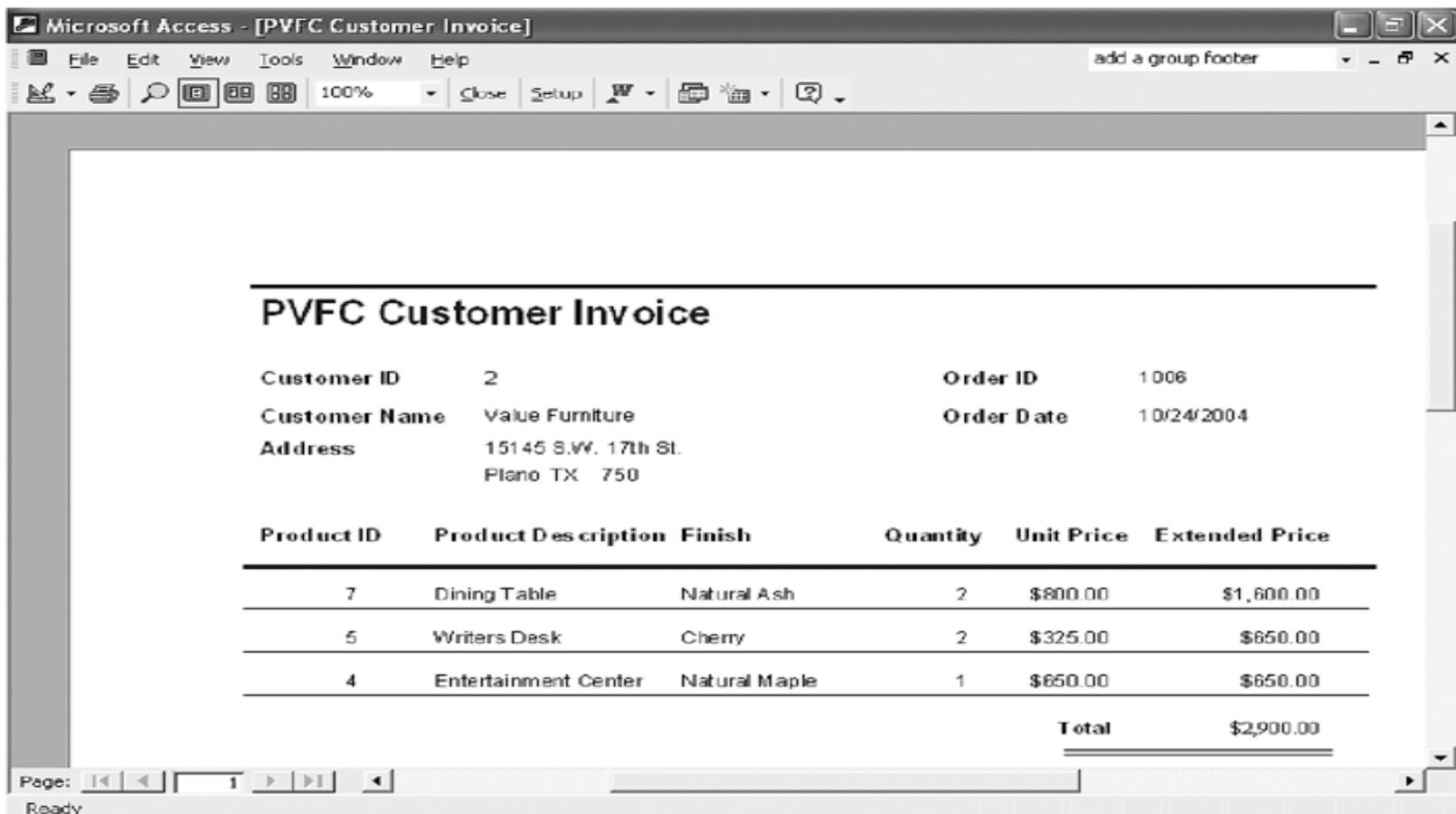
Red boxes highlight the Order_ID column in both tables, and red arrows point from the highlighted cells in the ORDER table to the corresponding cells in the Order_line table, illustrating a relationship.

	Order_ID	Order Date	Customer_ID
1	1001	10/21/2004	1
2	1002	10/21/2004	8
3	1003	10/22/2004	15
4	1004	10/22/2004	5
5	1005	10/24/2004	3
6	1006	10/24/2004	2
7	1007	10/27/2004	11
8	1008	10/30/2004	12
9	1009	11/5/2004	4
*	1010	11/5/2004	1
*	0		0

	Order_ID	Product_ID	Ordered_Quantity
1	1001	1	2
2	1001	2	2
3	1001	4	1
4	1002	3	5
5	1003	3	3
6	1004	6	2
7	1004	8	2
8	1005	4	4
9	1006	4	1
10	1006	5	2
11	1006	7	2
12	1007	1	3
13	1007	2	2
14	1008	3	3
15	1008	8	3
16	1009	4	2
17	1009	7	3
18	1010	8	10
*	0	0	0

Relationships established in special columns that provide links between tables

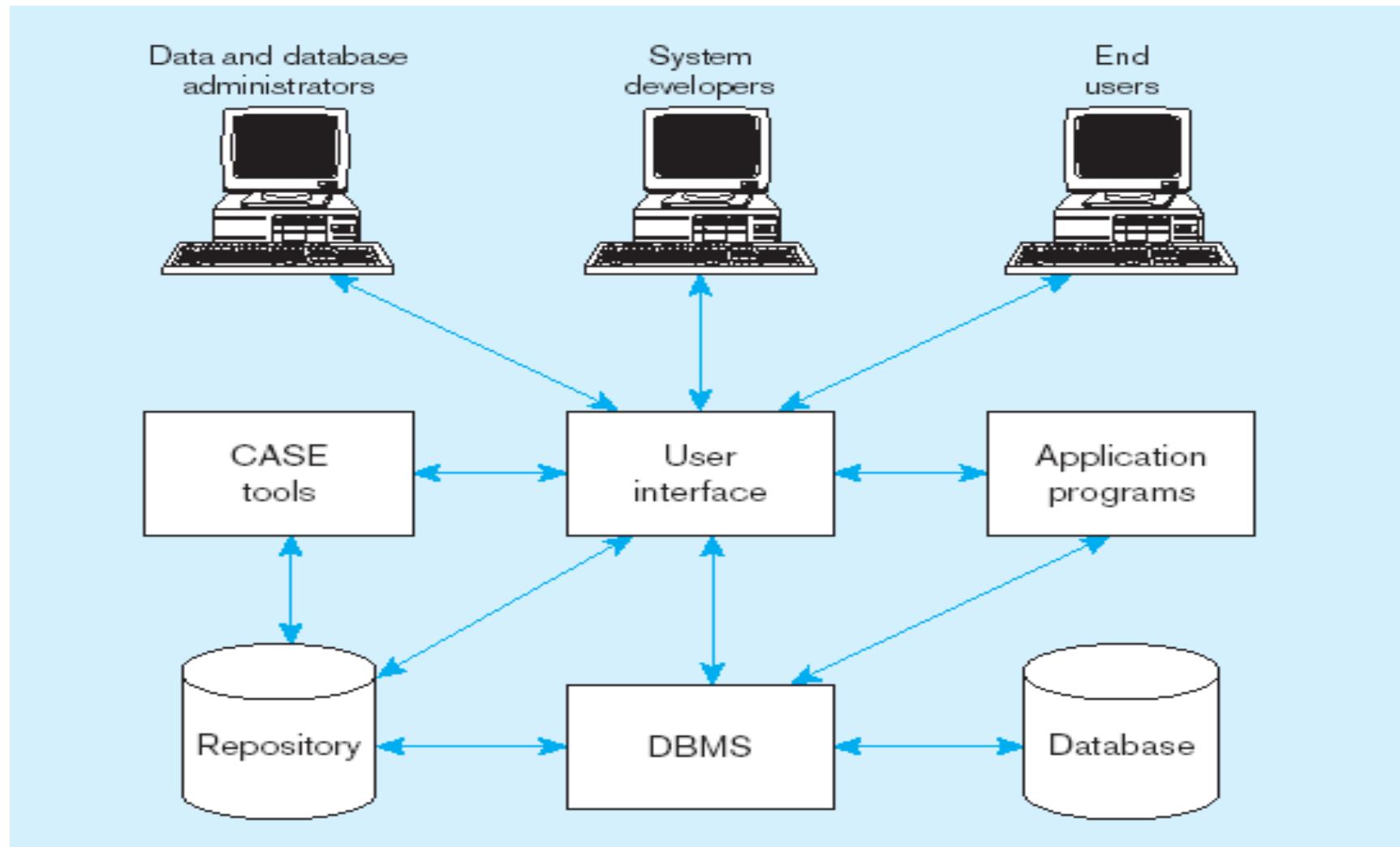
Figure 1-7 Customer invoice (Pine Valley Furniture Company)



Application program functions:

inserting new data, *updating* existing data,
deleting existing data, *reading* data for display

Figure 1-5 Components of the Database



The Range of Database Applications

- **Personal Computer Databases**
 - Support one user with a standalone personal computer
- **Workgroup Database**
 - Workgroup is a relatively small team of people (5-25) who collaborate on the same project/application or on a group of similar projects/applications
- **Department Databases**
 - Department databases are designed to support the various functions and activities of a department
- **Enterprise Databases**
 - An enterprise database is one whose scope is the entire organization or enterprise
 - Data Warehouse
 - An integrated decision support database whose content is derived from the various operational databases

Components of the Database Environment

- **CASE Tools**—computer-aided software engineering
- **Repository**—centralized storehouse of metadata
- **Database Management System (DBMS)** —software for managing the database
- **Database**—storehouse of the data
- **Application Programs**—software using the data
- **User Interface**—text and graphical displays to users
- **Data/Database Administrators**—personnel responsible for maintaining the database
- **System Developers**—personnel responsible for designing databases and software
- **End Users**—people who use the applications and databases

Customer

Customer Name:
Multi Media, Inc.

Address:
1000 River Road

City:
San Antonio

State:
TX

Zip:
76235

Phone:
(219) 864-2000

Next Contact Date: **10/17/2006**

Time: **10:30 AM**

Contact History for Customer

Date	Time	Contact	Comments
08/04/2006	10:00 AM	Roberts	Review proposal
08/19/2006	08:00 AM	Roberts	Revise schedule
09/10/2006	09:00 AM	Pearson	Sign contract
09/21/2006	02:00 PM	Roberts	Follow up

Figure 1-6
Typical data from a personal database

Figure 1-7 Workgroup database with wireless local area network

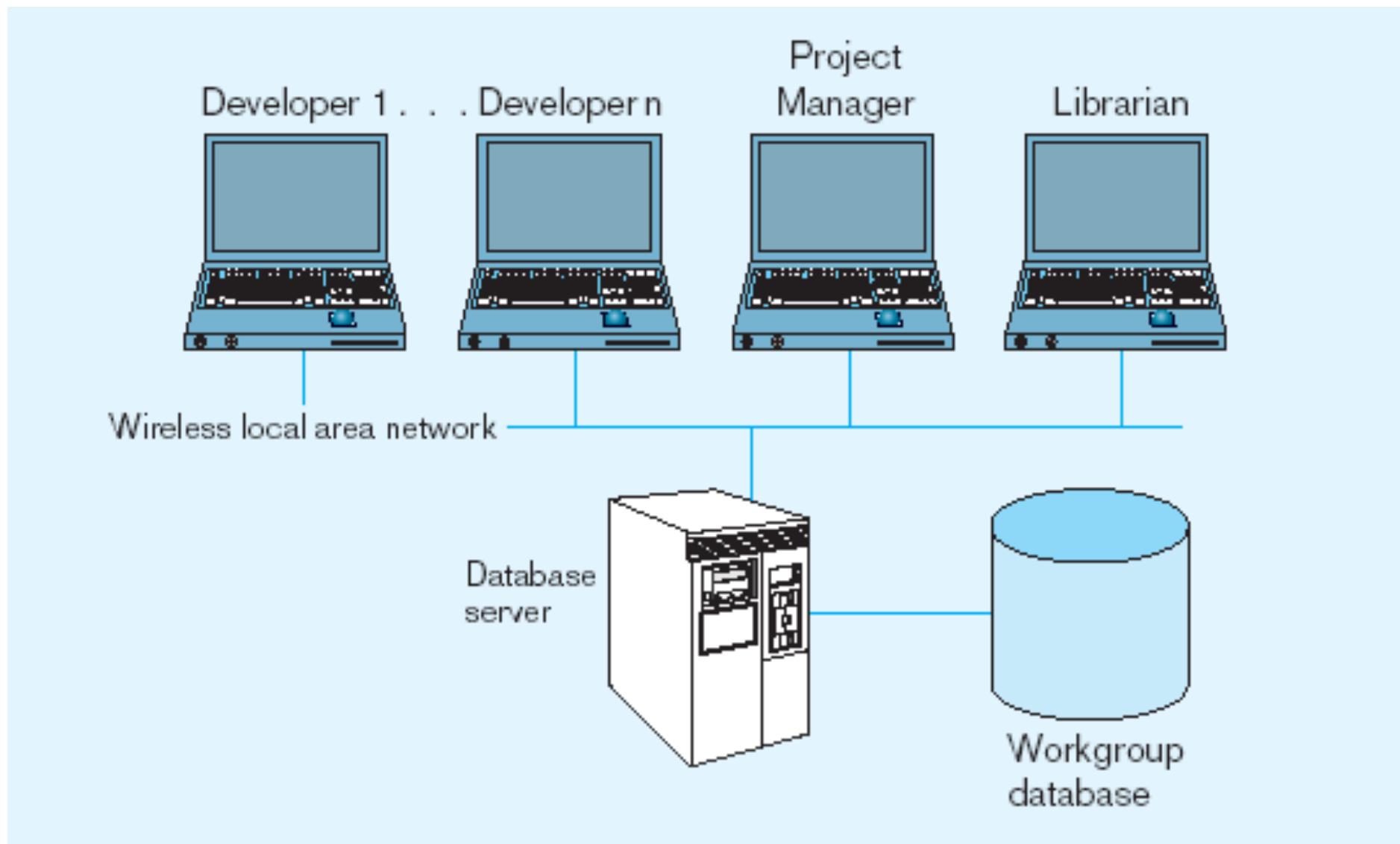


Figure 1-7: Three-Tier Database with Local Area Network

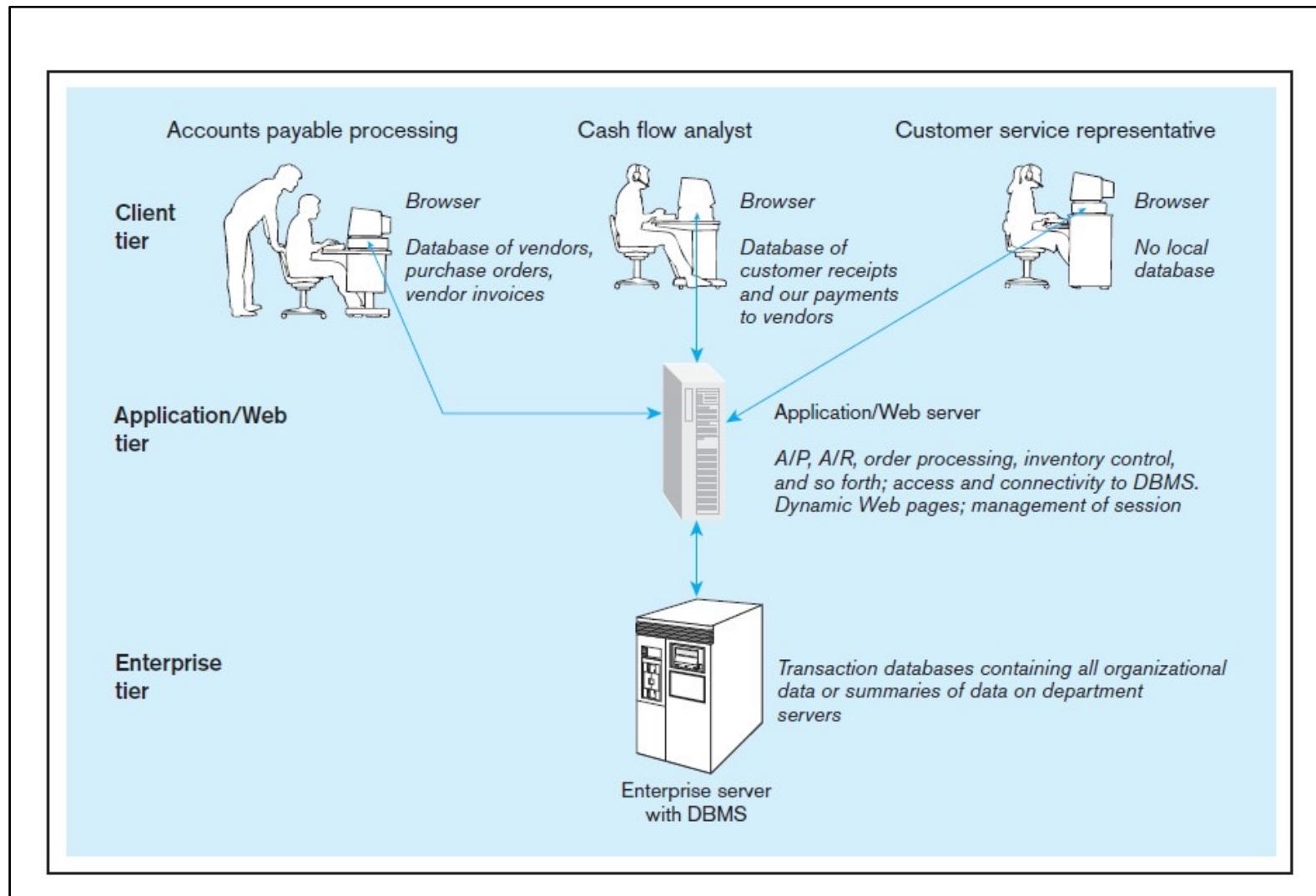


TABLE 1-5 Summary of Database Applications

Type of Database / Application	Typical Number of Users	Typical Size of Database
Personal	1	Megabytes
Two-tier	5–100	Megabytes–gigabytes
Three-tier	100–1000	Gigabytes
Enterprise resource planning	>100	Gigabytes–terabytes
Data warehousing	>100	Terabytes–petabytes

Enterprise Database Applications

- Enterprise Resource Planning (ERP)
 - Integrate all enterprise functions (manufacturing, finance, sales, marketing, inventory, accounting, human resources)
- Data Warehouse
 - Integrated decision support system derived from various operational databases

Figure 1-8 An enterprise data warehouse

