Database Systems -

Introduction to Databases and Data Warehouses

CHAPTER 1 - Introduction

KEY TOPICS

- Basic Terms
  - Data, Metadata, Database, Database Metadata
  - DBMS, Database System
  - Front-end application, End user, Direct/Indirect interaction
  - 2-Tier and 3-Tier Database System Architecture
  - Back-end processing
  - SQL interface
  - Relation, Column, Row, Relational Database
  - Relational and Non-Relational Table
  - Primary Key
- People Involved with Database Systems
- Database Scope

INITIAL TERMINOLOGY

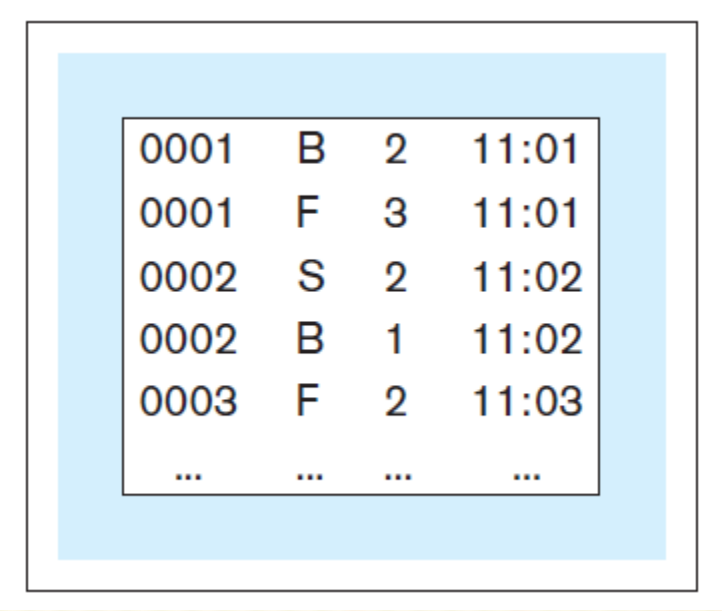
- **Data** - facts that are recorded and can be accessed
  - Data formats – text, numbers, figures, graphics, images, audio/video recordings and more
  - Data is recorded and kept because it is considered to be of use to an intended user
- **Information** - data that is accessed by a user for some particular purpose
  - Typically, getting the needed information from a collection of data requires performing an activity, such as searching through, processing, or manipulating the data in some form or fashion
  - Data that we need
- Data, Information – synonyms in the textbook
  - Differentiated in other books

INITIAL TERMINOLOGY

- **Metadata**
  - Data that describes the structure and the properties of the data
  - Metadata is essential for the proper understanding and use of the data
  - Data dictionary

INITIAL TERMINOLOGY

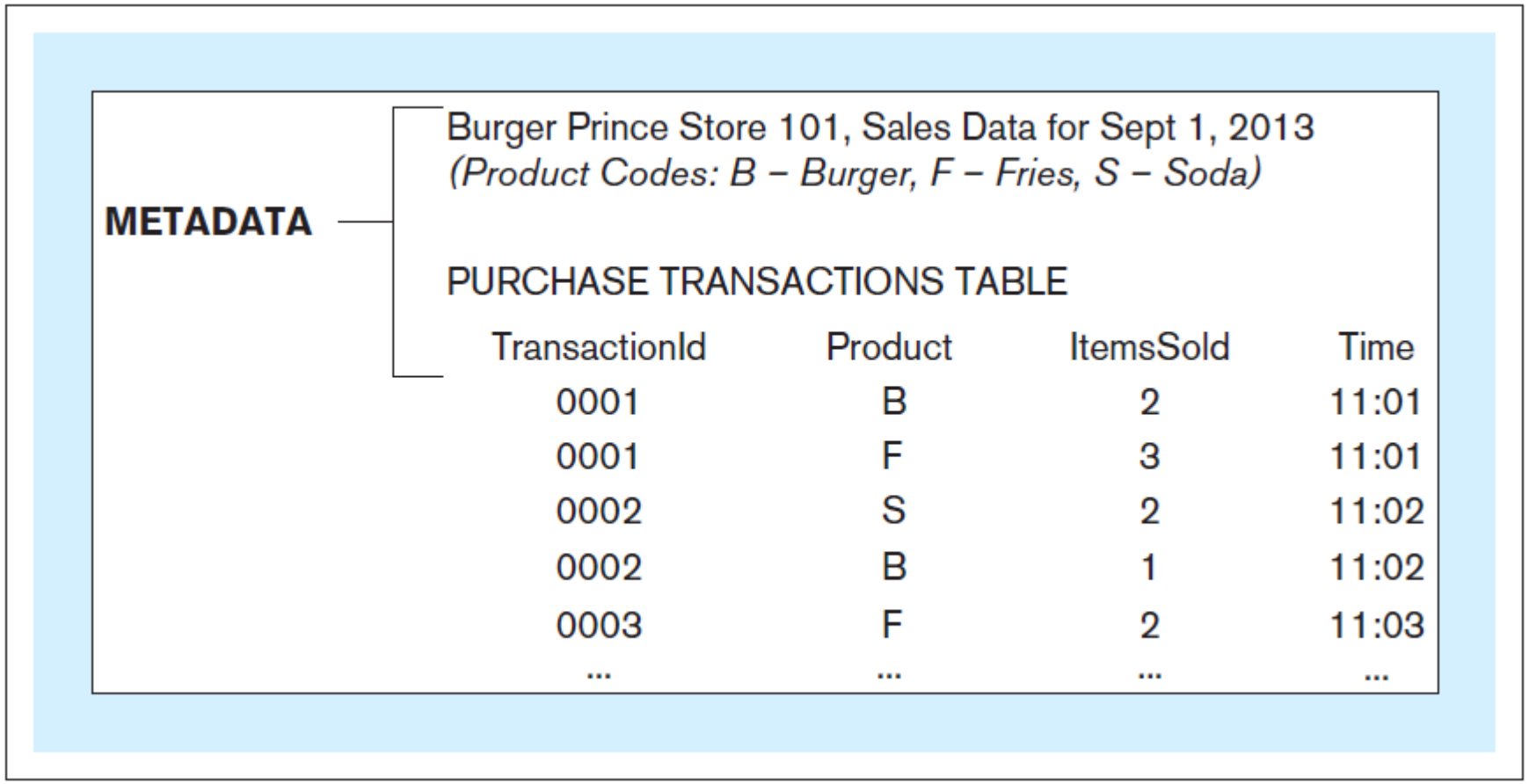
Data without metadata - example



0001	B	2	11:01
0001	F	3	11:01
0002	S	2	11:02
0002	B	1	11:02
0003	F	2	11:03
...	...	...	...

INITIAL TERMINOLOGY

Data with metadata - example



INITIAL TERMINOLOGY

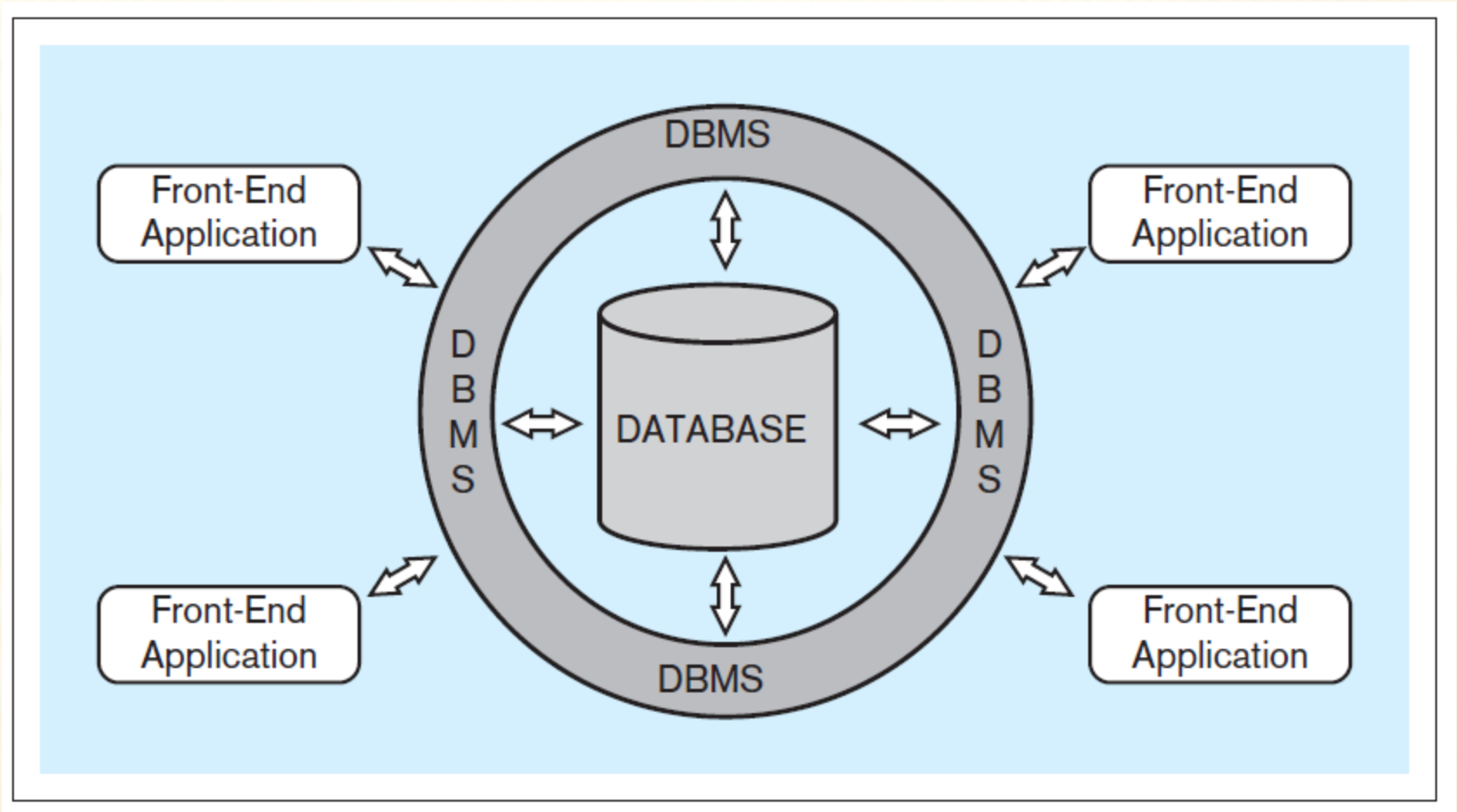
- **Database** - structured collection of related data stored on a computer medium
  - Organizes the data in a way that facilitates efficient access to the information captured in the data
- **Database metadata** – represents the structure of the database
  - Database content that is not the data itself (data about the data)
  - Contains:
    - Names of data structures (e.g. table name)
    - Data types (e.g. NUMBER as table column type)
    - Data descriptions (e.g. table column name)
    - Other information describing the characteristics of the data

INITIAL TERMINOLOGY

- **Database management system (DBMS)** - software used for:
  - Creation of databases
  - Insertion, storage, retrieval, update, and deletion of the data in the database
  - Maintenance of databases
- **Database and DBMS like this PPT file and MS PowerPoint**
- **Database system**
  - Computer-based system
  - Purpose: enable an efficient interaction between the users and the information captured in a database

INITIAL TERMINOLOGY

Typical database system architecture

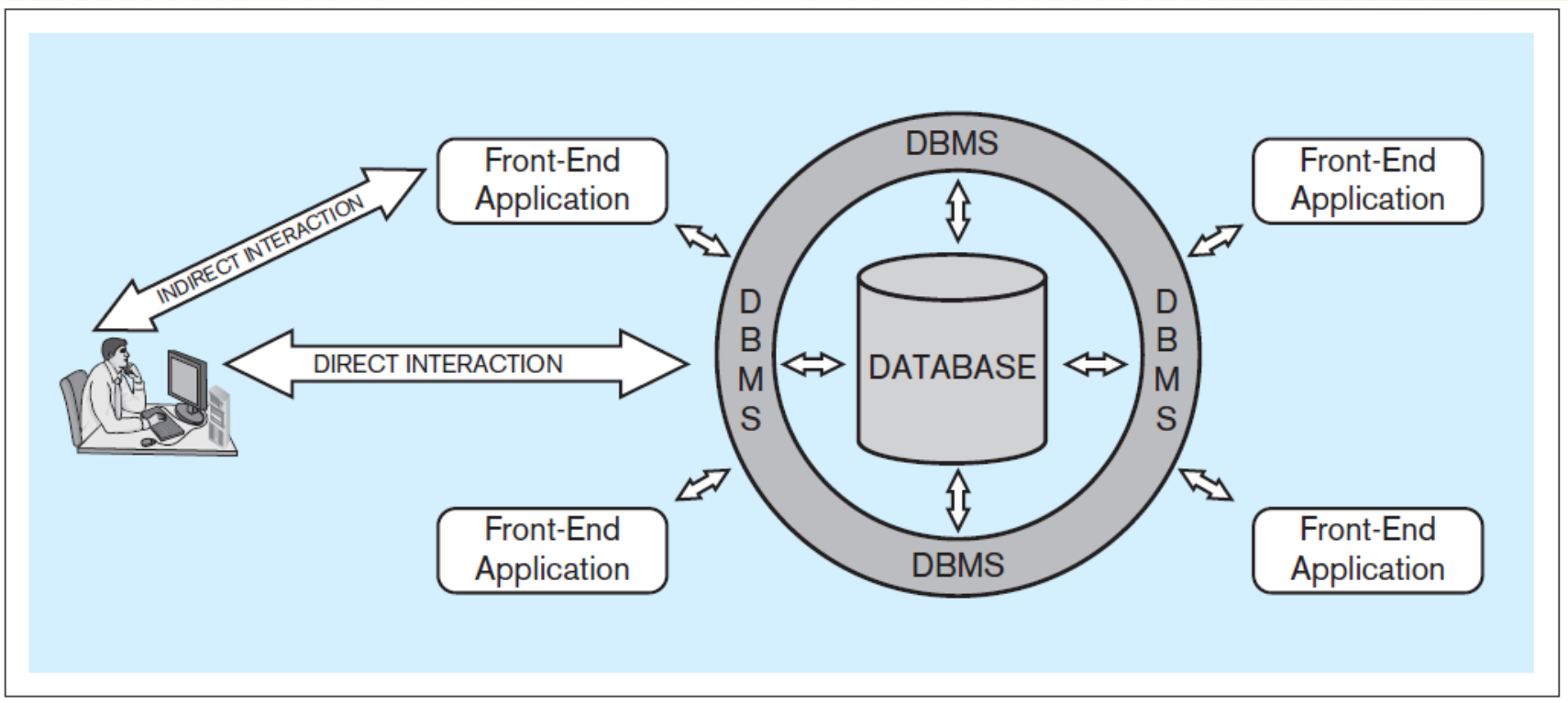


INITIAL TERMINOLOGY

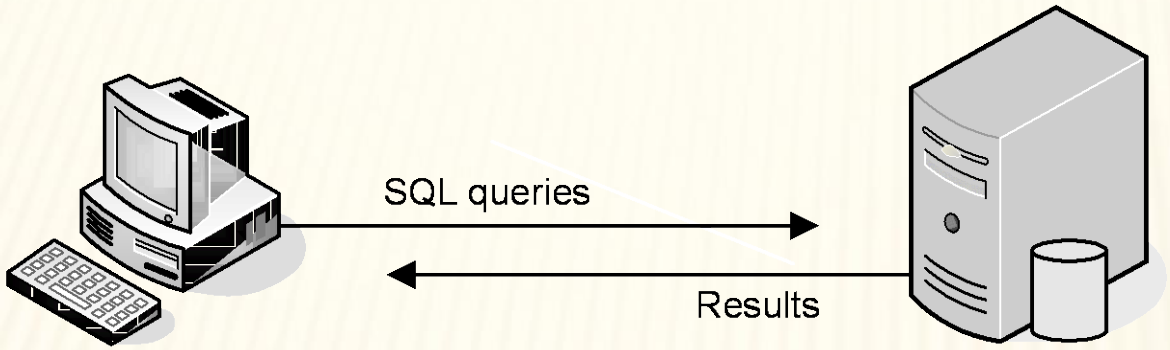
- **Front-end applications** - provide a mechanism for easy interaction between the users and the DBMS
  - Example: ATM withdrawal
- **End-users** (**business-users**) - users using a database system to support their work- or life-related tasks and processes
  - Different from technical personnel that test-use or maintain the database systems
- **Indirect interaction** - end-user communicating with the database through front-end applications
  - Requires little knowledge of DBMS
  - Most common
- **Direct interaction** - end-user communicating with the database directly through DBMS
  - Requires technical knowledge of DBMS

INITIAL TERMINOLOGY

Typical database system architecture



Client software, server software, and the SQL interface



Client
Application software
Data access API

Database server
Database management system
Database

Server software

- Database management system (DBMS)
- The DBMS does the *back-end processing*

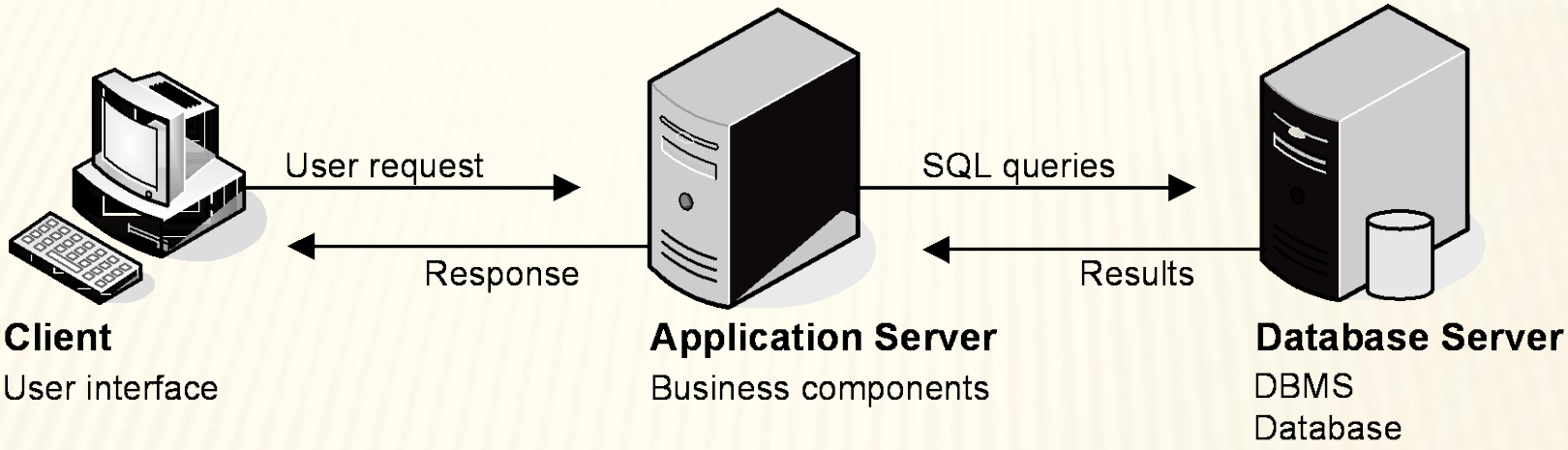
Client software

- Application software
- Data access API (application programming interface)
- The client software does the *front-end processing*

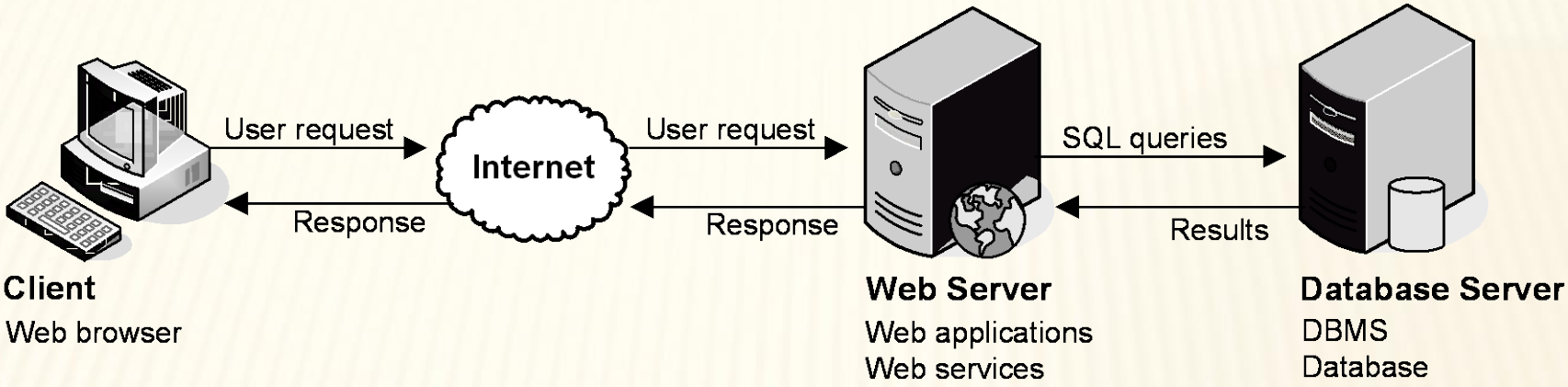
The SQL interface

- SQL queries
- *SQL* stands for *Structured Query Language*

An application that uses an application server



A simple web-based system



Other client/server components

- Application servers store business components
- Web servers store web applications and web services

How web applications work

- Web browser on a client sends a request to a web server.
- Web server processes the request.
- Web server passes any data requests to the database server.
- Database server returns results to web server.
- Web server returns a response to the browser.

INTRODUCTION

Terminology

TABLE 3.1 Synonyms Used in the Relational Database Model

Relation	=	Relational Table	=	Table
Column	=	Attribute	=	Field
Row	=	Tuple	=	Record

INTRODUCTION

- **Relational database**
  - A collection of related relations
    - Each relation must have a unique name within one collection
- **Relation**
  - A table in a relational database
  - A table containing rows and columns
  - The main construct in the relational database model
  - Every relation is a table, not every table is a relation

INTRODUCTION

- **Relation** - table in a relational database
  - Conditions for a table to be a relation:
    - Must have a name for each column
      - Unique column name within each table
    - Unique row within each table
    - Single-valued entry
      - Within each row, each value in each column must be single valued
    - Must have same (predefined) domain for all values in each column
    - Irrelevant ordering of columns and rows

INTRODUCTION

- **Relation** - table in a relational database
  - Conditions for a table to be a relation:
    - Example: Employee information
      - Domains of each column
        - Employee ID – 4 digits
        - Employee Name – 0 to 20 chars
        - Employee Gender – ‘M’ or ‘F’
        - Employee Phone – “xddd” //d: digit
        - Employee Bdate – date (day, month, year)

INTRODUCTION

Example of relational and non-relational tables

Relational Table (Relation)

EmpID	EmpName	EmpGender	EmpPhone	EmpBdate
0001	Joe	M	x234	1/11/1985
0002	Sue	F	x345	2/7/1983
0003	Amy	F	x456	4/4/1990
0004	Pat	F	x567	3/8/1971
0005	Mike	M	x678	5/5/1965

Not a Relational Table

EmpID	EmpInfo	EmpInfo	EmpPhone	EmpBdate
0001	Joe	M	x234	1/11/1985
0002	Sue	F	x345	2/7/1983
0001	Joe	M	x234	1/11/1985
0004	Pat	F	x567, x789	3/8/1971
0005	Mike	M	x678	a long time ago

INTRODUCTION

Different ordering of rows and columns in a relation

- But same information about Employee => same relation

A Relation

EmpID	EmpName	EmpGender	EmpPhone	EmpBdate
0001	Joe	M	x234	1/11/1985
0002	Sue	F	x345	2/7/1983
0003	Amy	F	x456	4/4/1990
0004	Pat	F	x567	3/8/1971
0005	Mike	M	x678	5/5/1965

Exact Same Relation (order of rows and columns is irrelevant)

EmpName	EmpID	EmpGender	EmpBdate	EmpPhone
Joe	0001	M	1/11/1985	x234
Amy	0003	F	4/4/1990	x456
Sue	0002	F	2/7/1983	x345
Pat	0004	F	3/8/1971	x567
Mike	0005	M	5/5/1965	x678

PRIMARY KEY

- **Primary key**
  - A single column (or a set of columns) that uniquely identify each row
  - Must have one primary key for each relation
  - Notation:
    - Underlined column(s)

PRIMARY KEY

Relation with the primary key underlined

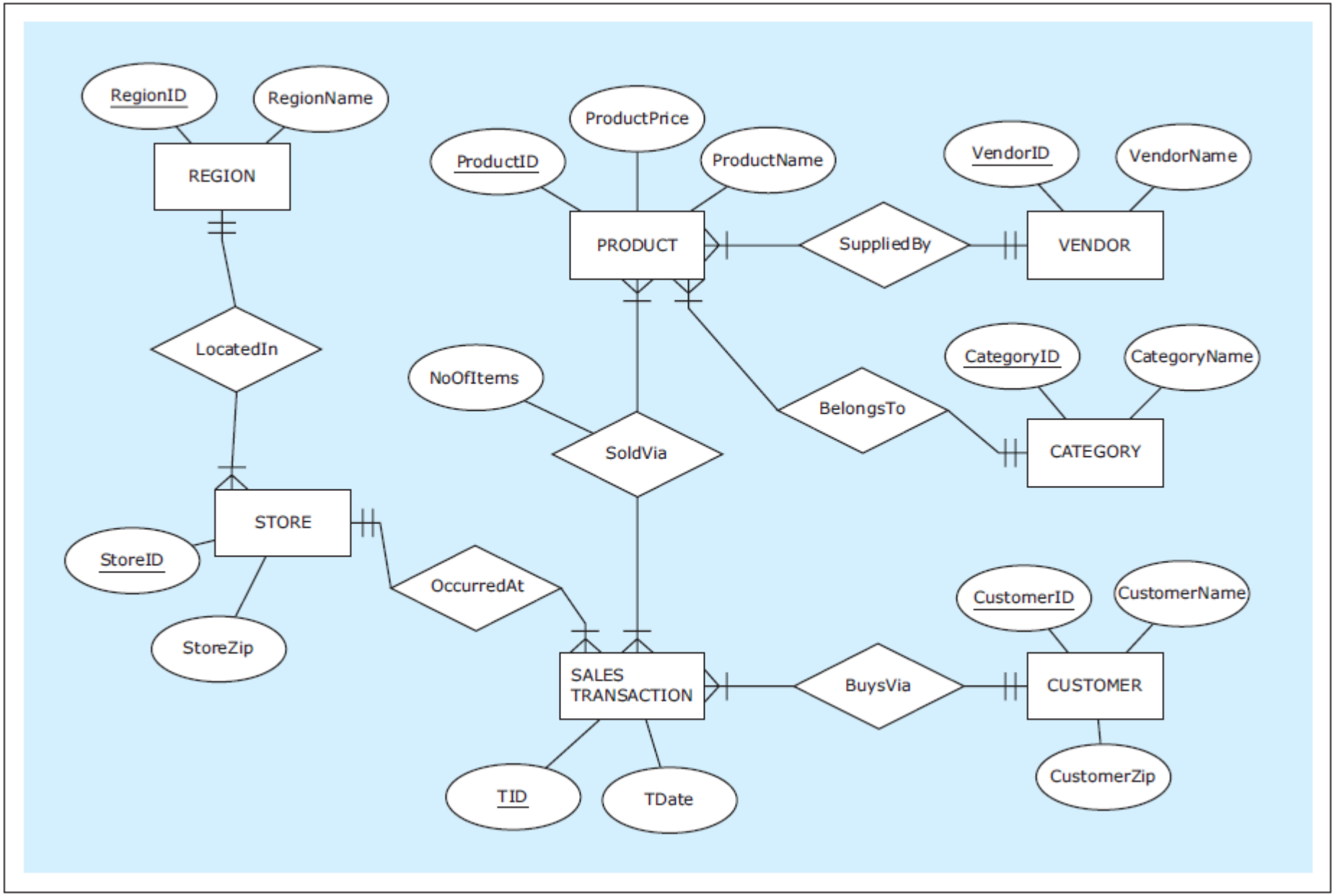
EMPLOYEE

EmpID	EmpName	EmpGender	EmpPhone	EmpBdate
0001	Joe	M	x234	1/11/1985
0002	Sue	F	x345	2/7/1983
0003	Amy	F	x456	8/4/1990
0004	Pat	F	x567	3/8/1971
0005	Mike	M	x678	5/5/1965
0010	Mike	M	x666	8/1/1974
0007	Barbara	F	x777	4/5/1980
0011	Ivan	M	x777	3/4/1981
0009	Amy	F	x777	1/11/1985

PEOPLE INVOLVED WITH DATABASE SYSTEMS

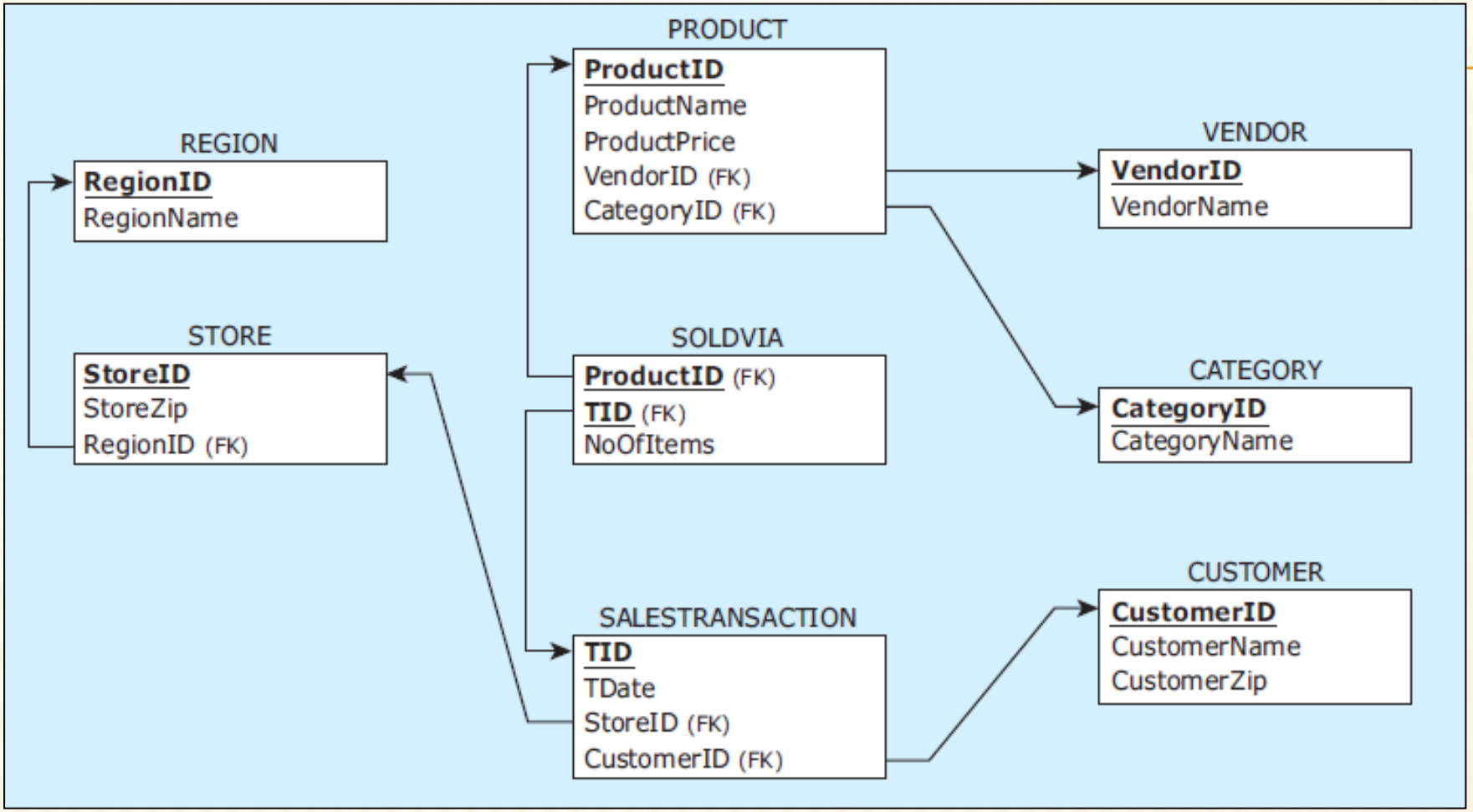
- **Database analysts, designers, and developers**
  - **Database analysts** - involved in the requirements collection, definition, and visualization stage
  - **Database designers** (a.k.a. **database modelers** or **architects**) - involved in the database modeling stage
  - **Database developers** – in charge of implementing the database model as a functioning database using the DBMS software
  - One person multiple roles possible in small companies

ER diagram example: ZAGI Retail Company Sales Department Database



(See notes page for details)

Example mapped relational schema: ZAGI Retail Company Sales Department Database



PEOPLE INVOLVED WITH DATABASE SYSTEMS

- **Front-end applications analysts and developers**
  - **Front-end application analysts** - in charge of collecting and defining requirements for front-end applications
    - User interfaces (how many, what kinds etc.)
  - **Front-end applications developers** - in charge of creating the front-end applications

PEOPLE INVOLVED WITH DATABASE SYSTEMS

- **Database administrators (DBAs)**
  - Perform the tasks related to the maintenance and administration of a database system
    - Security
    - Storage space
    - Backup and recovery
    - Etc.

PEOPLE INVOLVED WITH DATABASE SYSTEMS

- **Database end users**
  - Use a database system to support their work- or life-related tasks and processes
  - Arguably the most important category of people involved with database systems
    - Reason for database existence
    - Database quality = ease of use
  - Users differ in:
    - Level of technical sophistication
    - Amount of data that they need
    - Frequency with which they access the database system

DATABASE SCOPE

- Databases can vary in their scope
  - Small single-user (personal) databases
  - Large enterprise databases used by thousands of end-users