

Introduction to Databases

Dr. Eman Gouda

Course Contents

- Data Models
- Entity-Relationship Models
- Logical Database Design
- Relational Algebra
- Structured Query Language (SQL)
- Normalization
- Concurrency Control

Assessment

- Final Exam: 60
- Midterm: 20
- Project: 10
- Quizzes:10

Why Databases

- Businesses cannot operate without accurate information
- Companies must track:
 - Customers
 - Products
 - Employees
- Information is a critical organizational resource
- Databases are at the core of modern business systems.

What is a Database

- A structured system that allows fast storage, management, and retrieval of data
- Used in nearly all modern business systems
- Supports data-driven decision making

Data vs Information

Data

- Raw facts
- Unprocessed
- No immediate meaning
- **Example:**
 - University Faculty database → Raw data includes:
 - Name, Departments, Positions

Information

- Processed data
- Processing reveals meaning
- Information supports decision making
- Example:
 - What percentage of faculty are adjunct? → 20% adjunct faculty

Data vs Information

a) Data entry screen

Middle Tennessee State University

You are viewing the College console.

Baked Member

Member ID * (7 alphanumeric)
DO NOT repeat School ID (MT) to Member ID
Member ID * (7 alphanumeric) Password will be initially set to the same as Member ID

First name or initial * Middle name/initial * Last name *
 Chair/head Inactive Bypass chair for evaluation

Department * Area Email * Hire Term *

Involved? * Participating Supporting Qualification *
 Participates in the governance of the school
 Considered to be a long-term member

High Degree * Year Awarded * Rank * Assistant Professor

Back Save

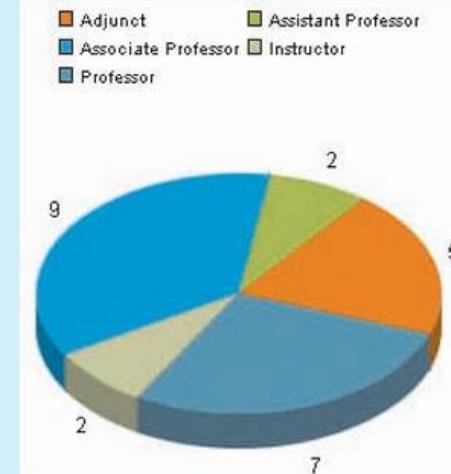
b) Raw data

ID	Last Name	Mid Name	First Name	Dept Code	Office	Email	Rank	Hire Year	Degree
1	Washington	A.	George	MGMT	N135	gwashington@mtsu.edu	Professor	2001 Ph.D	
2	Adams		John	FIN	N313	jadams@mtsu.edu	Professor	1984 Ph.D	
3	Jefferson	L.	Thomas	ECON		tjefferson@mtsu.edu	Instructor	2002 M.B.A.	
4	Madison	D.	James	FIN	N236	jmadison@mtsu.edu	Associate Professor	1994 Ph.D	
5	Murphy	N.	James	ACCT	N411	jmurphy@mtsu.edu	Assistant Professor	1995 Ph.D	
6	Adams	O.	John	ACCT	N418	jadams@mtsu.edu	Associate Professor	1993 Ph.D	
7	Jackson	C.	Andrew	ECON	N303	ejackson@mtsu.edu	Associate Professor	1999 Ph.D	
8	Van Buren	T.	Martin	FIN	N306	mvanburen@mtsu.edu	Professor	1988 Ph.D	
9	Harrison	R.	William	MKTG	N119	wjharrison@mtsu.edu	Professor	1994 Ph.D	
10	Tyler	M.	John	MGMT	N143	jtyler@mtsu.edu	Assistant Professor	2000 Ed.D	
11	Polk		Cheryl	MKTG	N145	copolk@mtsu.edu	Associate Professor	2002 Ph.D	
12	Taylor	G.	Zachery	ACCT	N415	ztaylor@mtsu.edu	Associate Professor	1998 Ph.D	
13	Fitzsimons		Meredith	JOB	N219	mfitzsimons@mtsu.edu	Professor	1992 Ph.D	
14	Pierce	A.	Franklin	MKTG	N159	fpierce@mtsu.edu	Instructor	2005 M.B.A.	
15	Buchanan	T.	James	MGMT	N146	jbuchanan@mtsu.edu	Associate Professor	1996 D.B.A.	
16	Lincoln	W.	Larry	MGMT	N155	llincoln@mtsu.edu	Associate Professor	1998 Ph.D	
17	Johnson		Andrew	ISYS	N368	ajohnson@mtsu.edu	Professor	1987 Ph.D	
18	Grant		Kate	MKTG	N128	kgrant@mtsu.edu	Assistant Professor	1989 D.B.A.	
19	Putherford		Hayes	ACCT	N403	hputherford@mtsu.edu	Professor	1992 Ph.D	
20	Grofeld	T.	Derica	ACCT		dgrofeld@mtsu.edu	Assistant Professor	2010 Ph.D	
21	Arthur		Emily	ACCT	N413	eartur@mtsu.edu	Associate Professor	2003 J.D.	
22	Cleveland	G.	Robert	ACCT	N401	rcleveland@mtsu.edu	Associate Professor	1997 Ph.D	
23	Harrison	X.	Patricia	BUSA	N406	psharrison@mtsu.edu	Associate Professor	2001 J.D.	
24	McKinley	B.	Priscilla	ISYS	N363	pmckinley@mtsu.edu	Adjunct	1994 M.S.	
25	Roosevelt	F.	Hilary	MGMT	N104	hroosevelt@mtsu.edu	Associate Professor	2002 Ph.D	
26	Wilson		Leanne	BUSE	N448	lwilson@mtsu.edu	Professor	1992 Ph.D	
27	Harding		Werner	MKTG	N114	wharding@mtsu.edu	Professor	1994 Ed.D	
28	Coolidge		Colvin	ECON	N316	ccoolidge@mtsu.edu	Professor	1975 Ph.D	
29	Hoover		Lisa	MGMT		lhoover@mtsu.edu	Adjunct	1978 M.B.A.	
30	Truman		Betty	ACCT	N416	bttruman@mtsu.edu	Professor	1971 Ed.D	
31	Johnson		Robert	BUSE	N249	rjohnson@mtsu.edu	Professor	2001 Ph.D	

c) Information in summary format

Rank	COUNT	%/INFS	TOT/COL	%/COL. TOT.	%/COL. FAC.
Adjunct	5	20.00%	23	21.74%	3.27%
Assistant Professor	2	8.00%	28	7.14%	1.31%
Associate Professor	9	36.00%	37	24.32%	5.88%
Instructor	2	8.00%	18	11.11%	1.31%
Professor	7	28.00%	47	14.89%	4.58%

d) Information in graphical format



Data Management

- Data must be:
 - Accurate
 - Well stored
 - Easy to access
 - Easy to process
- Data management includes:
 - Generation
 - Storage
 - Retrieval
 - Protection

Introducing the Database

- A database is a shared, integrated computer structure
- It stores:
 - End-user data (raw facts)
 - Metadata (data about data)
- **What is Metadata:**
 - Metadata describes the data:
 - Data element names
 - Data types (text, number, date)
 - Validation rules
 - Relationships between data

Database Management System (DBMS)

- DBMS = software that manages databases
- Controls access to stored data
- Acts as an intermediary between user and database
- Hides internal complexity
- **Role of the DBMS:**
 - Receives user requests
 - Translates them into database operations
 - Presents a single integrated view of data
 - Supports multiple users and applications

DBMS: Advantages

Data Sharing

- Data can be shared across users
- Multiple applications access the same data
- Improves collaboration

Data Security

- Protects sensitive data
- Controls user access
- Reduces security risks

Data Integration

- Combines data from different departments
- Creates a unified organizational view

DBMS: Advantages

- **Reduced Data Inconsistency**
 - Prevents duplicate or conflicting data
 - Ensures consistent records
 - Example: same employee name stored uniformly
- **Improved Data Access**
 - Supports ad hoc queries (instant questions)
 - Provides fast answers
 - Helps analyze large datasets
 - Examples:
 - Sales volume reports
 - Bonus calculations
 - Customer balances

Types of Databases: Number of users

- **Single-user database:**
 - Supports only one user at a time
 - Usually runs on a personal computer
- **Multi-user database:**
 - Supports multiple users at the same time
 - **Workgroup database:**
 - Multi-user database that supports a small group of users or a single department
 - **Enterprise database:**
 - Multi-user database that supports a large group of users or an entire organization

Types of Databases: Location of Databases

- **Centralized Database:**
 - Supports data located at a single site
- **Distributed Database:**
 - Supports data distributed across several sites
 - Connected through networks

Types of Databases: Classification by Structure

- **Structured data**
 - Organized and formatted
 - Ready for processing
 - Example: Excel
- **Unstructured data**
 - Raw, unformatted data
 - Example: videos, documents
- **Semi-structured data**
 - Partially organized
 - Example: web pages → XML

Why Database Design is Important

Database name: Ch01_Text

ID	ENum	Name	Title	HireDate	Skill1	Skill1Date	Skill2	Skill2Date	Skill3	Skill3Date
1	02345	Johnny Jones	DBA	2/14/1995	Basic Database Management	2/14/2000	Advanced Database Management	2/14/2003	Basic Web Design	8/9/2001
2	08273	Marco Bienz	Analyst	7/28/2006	Basic Web Design	3/8/2007	Advance Process Modeling	8/19/2010		
3	06234	Jasmine Patel	Programmer	8/10/2005	Basic Web Design	8/10/2005	Advanced C# programming	8/10/2005	Basic DB manipulation	1/29/2010
4	03373	Franklin Johnson, Jr.	Purchasing Agent	3/15/2002	Advanced Spreadsheets	6/20/2009				
5	13567	Almond, Robert	Analyst	9/30/2012	Basic Process Modeling	9/30/2012	Basic Database Design	5/23/2013		
6	10282	Richardson, Amanda	Clerk	4/11/2011						
7	09382	Jessica Johnson	Database Programmer	8/2/2010	Basic DB Design	8/2/2010	Basic Database Manipulation	8/2/2010	Advanced DB Manipulation	5/1/2011
8	14311	Duong, Lee	Programmer	9/1/2014	Basic Web Design	9/1/2014				
9					Master Database Programming					
10					Basic Spreadsheets					
11	09002	Ben Joiner	Clerk	5/20/2010	Advanced Spreadsheets	5/16/2011	Basic Web Design	5/16/2011		
12	13383	Raymond F. Matthews	Programmer	3/12/2012	Basic C# Programming	3/12/2012				
13	09283	Chavez, Juan	Clerk	7/4/2010						
14	04893	Patricia Richards	DBA	6/11/2004	Advanced Database Management	6/11/2004	Advanced Database Manipulation	9/20/2010		
15	13932	Lee, Megan	Programmer	9/29/2013						

Database Decomposition

Table name: EMPLOYEE

Employee_ID	Employee_FName	Employee_LName	Employee_HireDate	Employee_Title
02345	Johnny	Jones	2/14/1995	DBA
03373	Franklin	Johnson	3/15/2002	Purchasing Agent
04893	Patricia	Richards	6/11/2004	DBA
06234	Jasmine	Patel	8/10/2005	Programmer
08273	Marco	Bienz	7/28/2006	Analyst
09002	Ben	Joiner	5/20/2010	Clerk
09283	Juan	Chavez	7/4/2010	Clerk
09382	Jessica	Johnson	8/2/2010	Database Programmer
10282	Amanda	Richardson	4/11/2011	Clerk
13383	Raymond	Matthews	3/12/2012	Programmer
13567	Robert	Almond	9/30/2012	Analyst
13932	Megan	Lee	9/29/2013	Programmer
14311	Lee	Duong	9/1/2014	Programmer

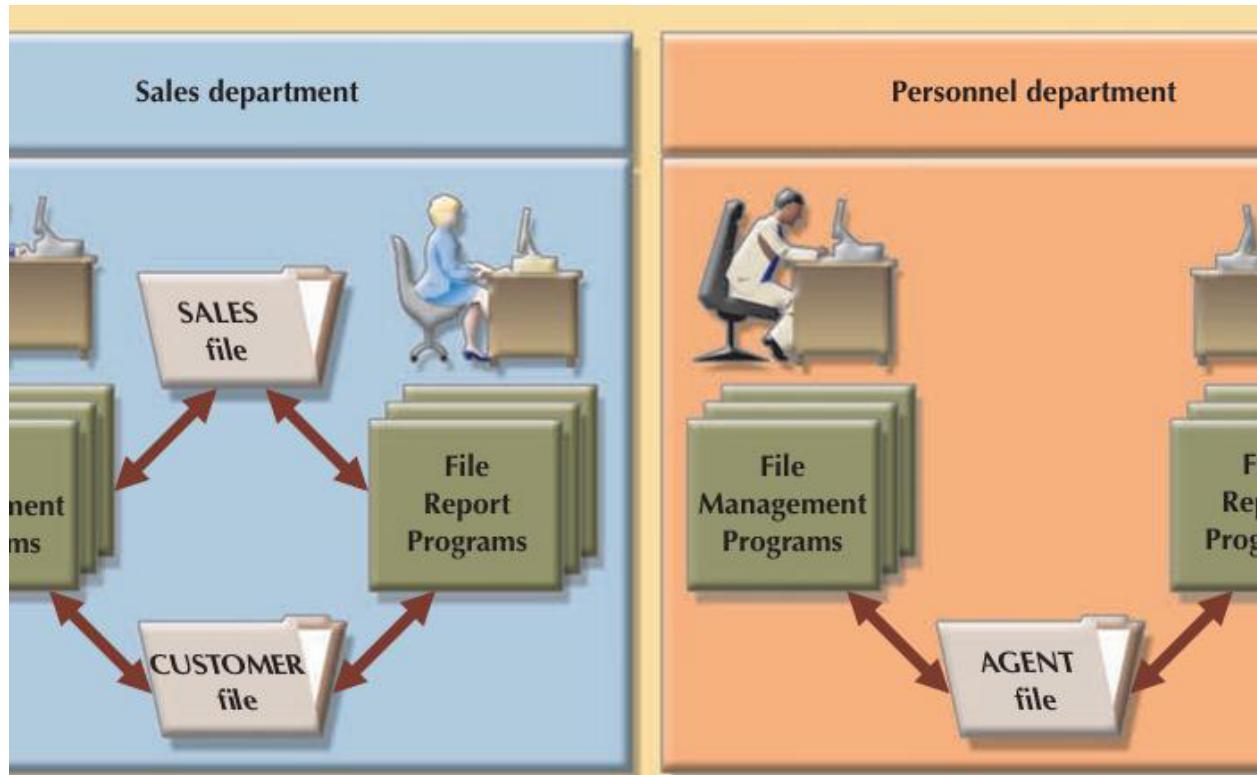
Table name: CERTIFIED

Employee_ID	Skill_ID	Certified_Date
02345	100	2/14/2000
02345	110	8/9/2001
02345	180	2/14/2003
03373	120	6/20/2009
04893	180	6/11/2004
04893	220	9/20/2010
06234	110	8/10/2005
06234	200	8/10/2005
06234	210	1/29/2010
08273	110	3/8/2007
08273	190	8/19/2010
09002	110	5/16/2011
09002	120	5/16/2011
09382	140	8/2/2010
09382	210	8/2/2010
09382	220	5/1/2011
13383	170	3/12/2012
13567	130	9/30/2012
13567	140	5/23/2013
14311	110	9/1/2014

Table name: SKILL

Skill_ID	Skill_Name	Skill_Description
100	Basic Database Management	Create and manage database user accounts.
110	Basic Web Design	Create and maintain HTML and CSS documents.
120	Advanced Spreadsheets	Use of advanced functions, user-defined functions, and macroing.
130	Basic Process Modeling	Create core business process models using standard libraries.
140	Basic Database Design	Create simple data models.
150	Master Database Programming	Create integrated trigger and procedure packages for a distributed environment.
160	Basic Spreadsheets	Create single tab worksheets with basic formulas
170	Basic C# Programming	Create single-tier data aware modules.
180	Advanced Database Management	Manage Database Server Clusters.
190	Advance Process Modeling	Evaluate and Redesign cross-functional internal and external business processes.
200	Advanced C# Programming	Create multi-tier applications using multi-threading.
210	Basic Database Manipulation	Create simple data retrieval and manipulation statements in SQL.
220	Advanced Database Manipulation	Use of advanced data manipulation methods for multi-table inserts, set operations, and correlated subqueries.

Evolution of File System Data Processing



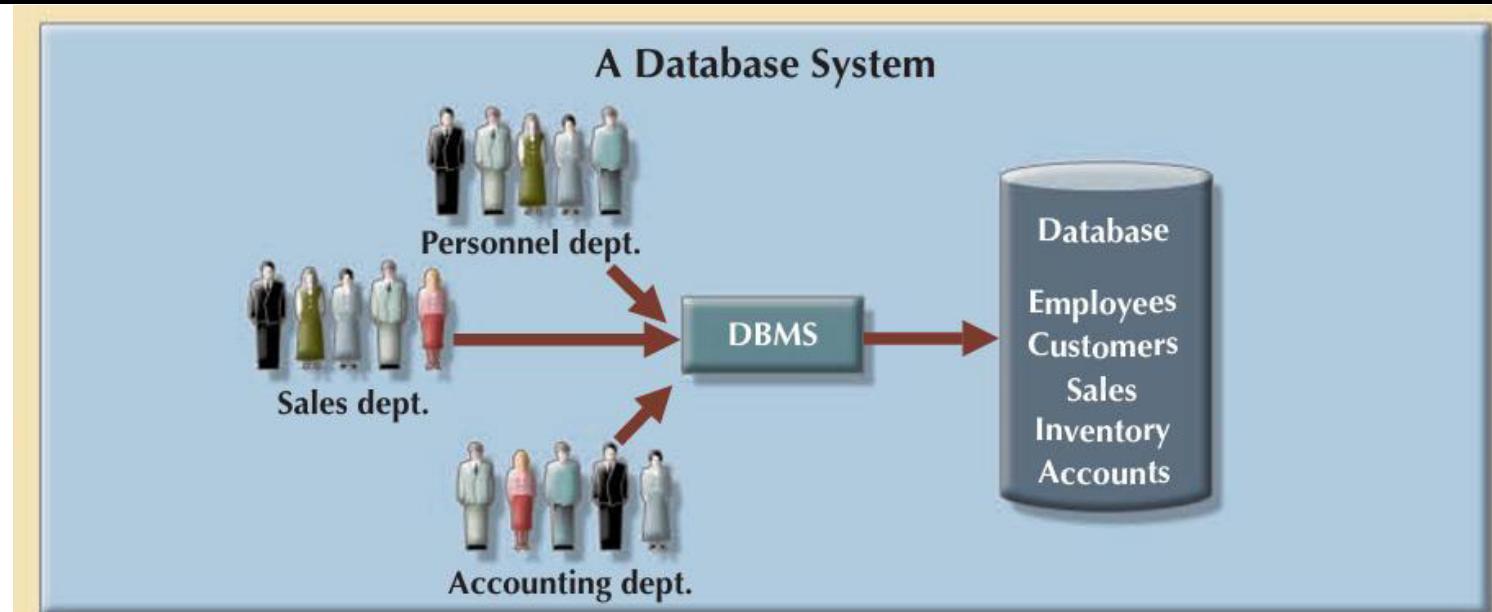
- **Manual File Systems:**

- Early systems were paper-and-pencil
- Worked well when small amounts of data
- Slow
- Hard to manage
- Prone to errors
- Time-consuming

- **Computerized File Systems**

- Computers replaced paper
- Generate reports faster
- Reduce manual effort
- Issues:
 - Data Redundancy
 - Poor Data Security
 - Data Inconsistency
 - Data Anomalies

Why Database Systems



DBMS Functions

- **Data Dictionary Management:** Stores definitions of data elements & relationships (metadata)
- **Data Storage Management:** Stores data
- **Data Transformation & Presentation:**
 - Converts data to required logical format
 - Example:
 - UK: 11/07/2015
 - US: 07/11/2015
- **Security Management:**
 - Controls user access & privacy
 - Defines what users can: Read, add, modify, delete

DBMS Functions

- **Multiuser Access Control:** Allows multiple users to access data simultaneously
- **Backup & Recovery**
 - Ensures data safety
 - Handles database failures: Disk failure, Power outages
- **Data Integrity Management**
 - Enforces integrity rules → reduces redundancy, increases consistency
- **Database Access Languages & APIs**
 - Supports query languages: SQL (Structured Query Language)

Why a Spreadsheet Is Not a Database

- Spreadsheets allow manipulation of data in a tabular format.
- Spreadsheets **do not** support basic database functionality, such as:
 - Self-documentation through metadata
 - Enforcement of data types or domains → ensures consistency within a column
 - Defined relationships among tables
 - Constraints to maintain consistency across related tables

Thank You