

# Data Information Systems Management Database Systems

## LECTURE 1 - DATABASES

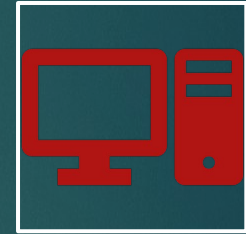
# CHAPTER OUTLINE



Managing Data



The Database  
Approach

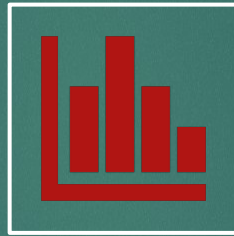


Database  
Management  
Systems

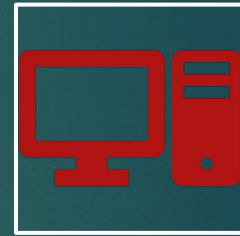
# What's the big deal about databases?



Who has used a database today?



What data are they recording



What data is stored in these databases



# Annual Flood of Data from.....

**Credit/Debit card swipes**

**E-mails**

**Digital video**

**CCTV**

**RFID tags**

**Social Media**

**Telemetry**

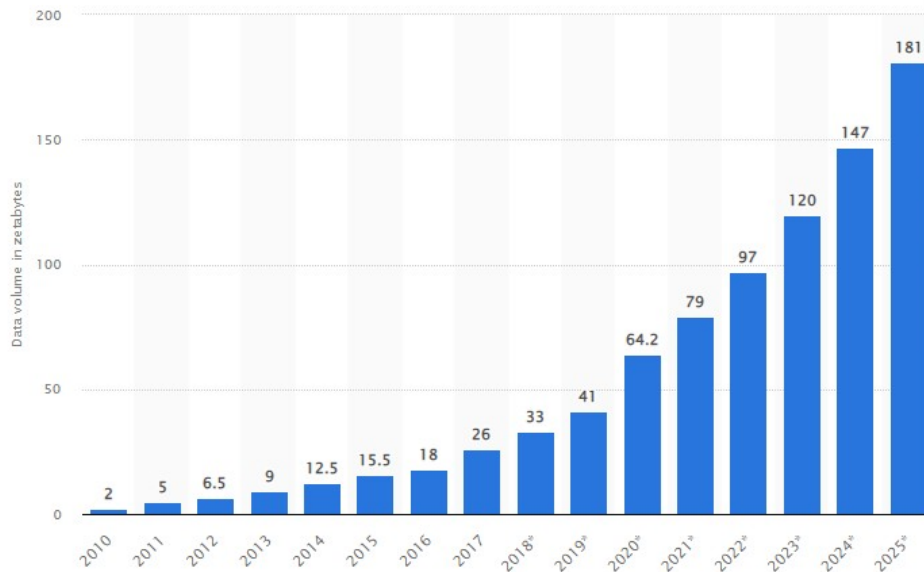
**Radiology scans**



Source: Media Bakery

# Annual Flood of New Data!

(in zettabytes)



© Statista 2022

Show source

**In the zettabyte range**

**A zettabyte is 1000 exabytes**



# The Difficulties of Managing Data

- ▶ Amount of data increasing exponentially
- ▶ Data are scattered throughout organizations and collected by many individuals using various methods and devices.
- ▶ Data come from many sources.
- ▶ Data security, quality, and integrity are critical.

There is a need for Data Governance



# Data Governance

- **Data Governance:** is an approach to managing information across an entire organization
- **Master Data Management :** is a process that spans all of an organization's business processes
- **Master Data:** are a set of core data that span all of an enterprise's information systems.

See [video](#) The benefits of a Master Data Management system



# Master Data Management

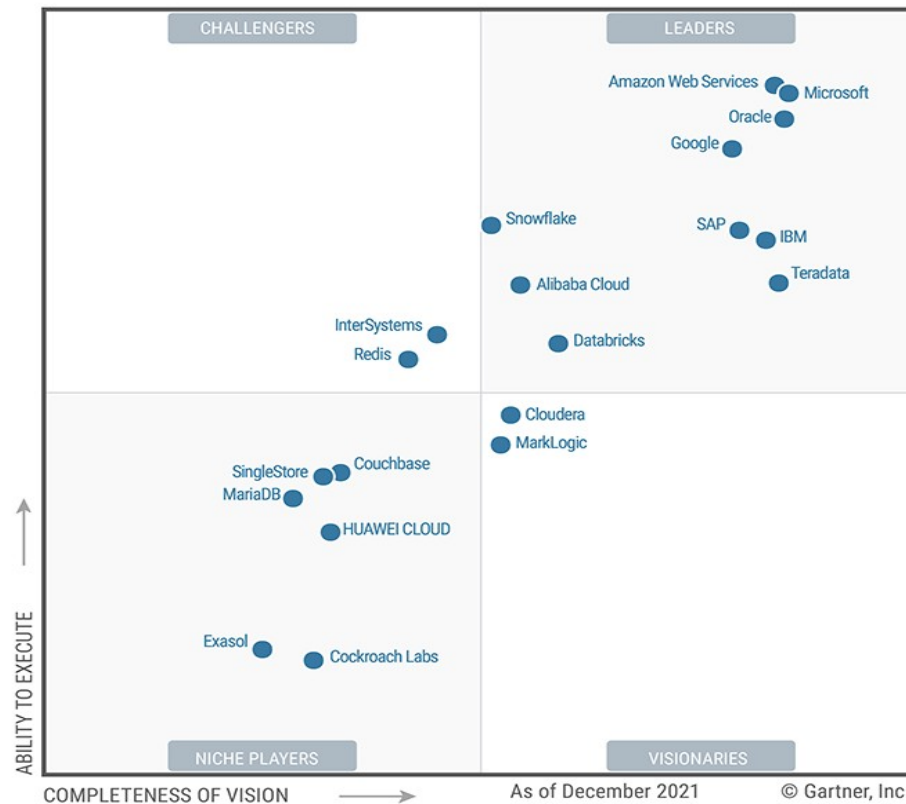
- ▶ Top 10 Enterprise Database Systems to consider in 2023

\* Note of caution depending on who is producing these stats we must just accept these are indicative in nature



# Data Management systems 2021

Figure 1: Magic Quadrant for Cloud Database Management Systems



# Data Management systems 2022 (December)

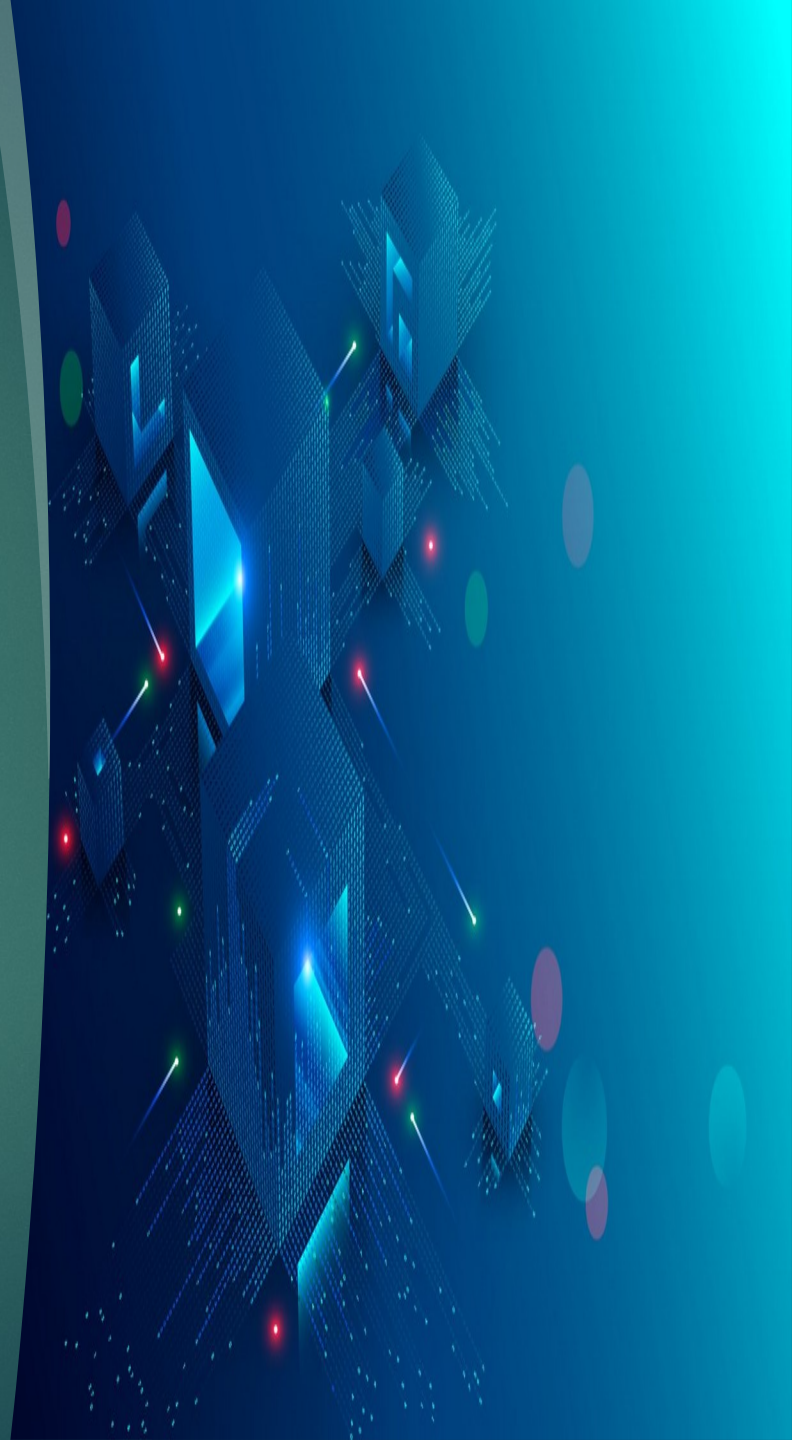
Figure 1: Magic Quadrant for Cloud Database Management Systems



Source: Gartner (December 2022)

# The Database Approach

- ▶ Oracle Database concepts
  - ▶ Oracle database concepts
  - ▶ Top differences between Oracle and Microsoft SQL server
  - ▶ MySQL versus Oracle Features/Functionality





# The Database Approach

Database management system (DBMS) minimize the following problems:

**Data redundancy:** The same data are stored in many places.

**Data isolation:** Applications cannot access data associated with other applications.

**Data inconsistency:** Various copies of the data do not agree.





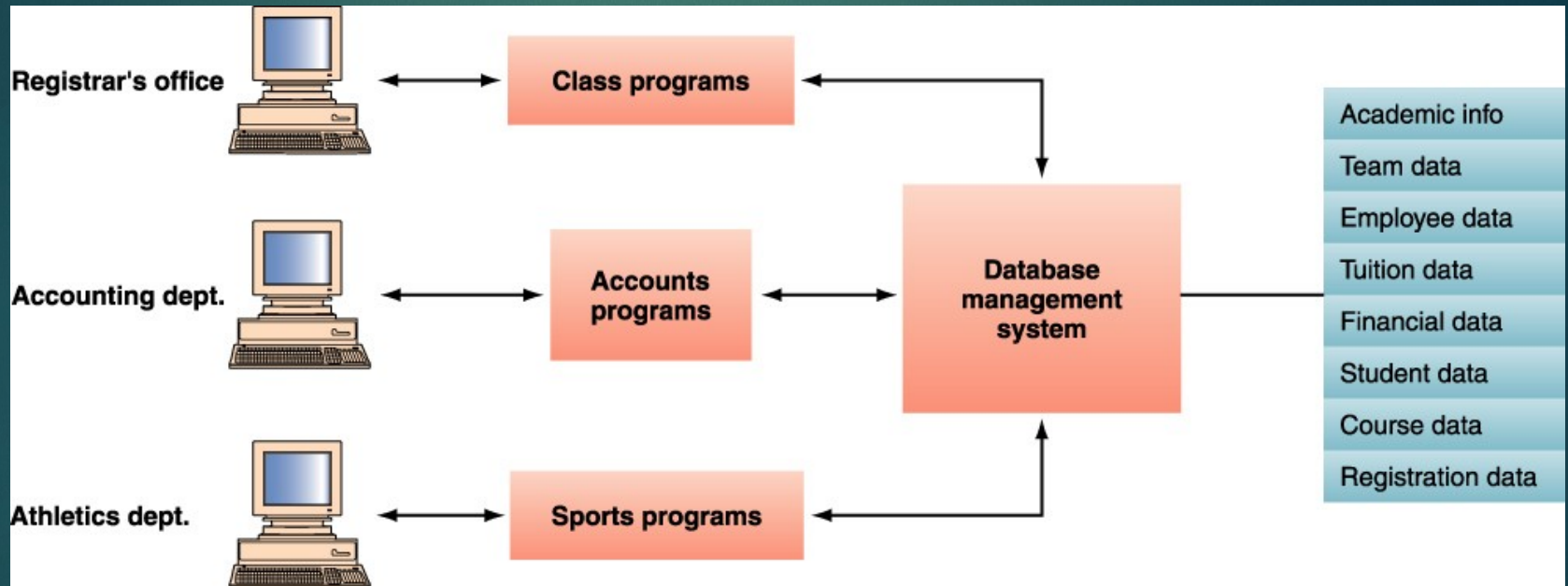
# Database Approach (continued)

DBMSs maximize the following issues:

- ▶ **Data security:** Keeping the organization's data safe from theft, modification, and/or destruction.
- ▶ **Data integrity:** Data must meet constraints (e.g., student grade point averages cannot be negative).
- ▶ **Data independence:** Applications and data are independent of one another. Applications and data are not linked to each other, meaning that applications are able to access the same data.



# Database Management Systems





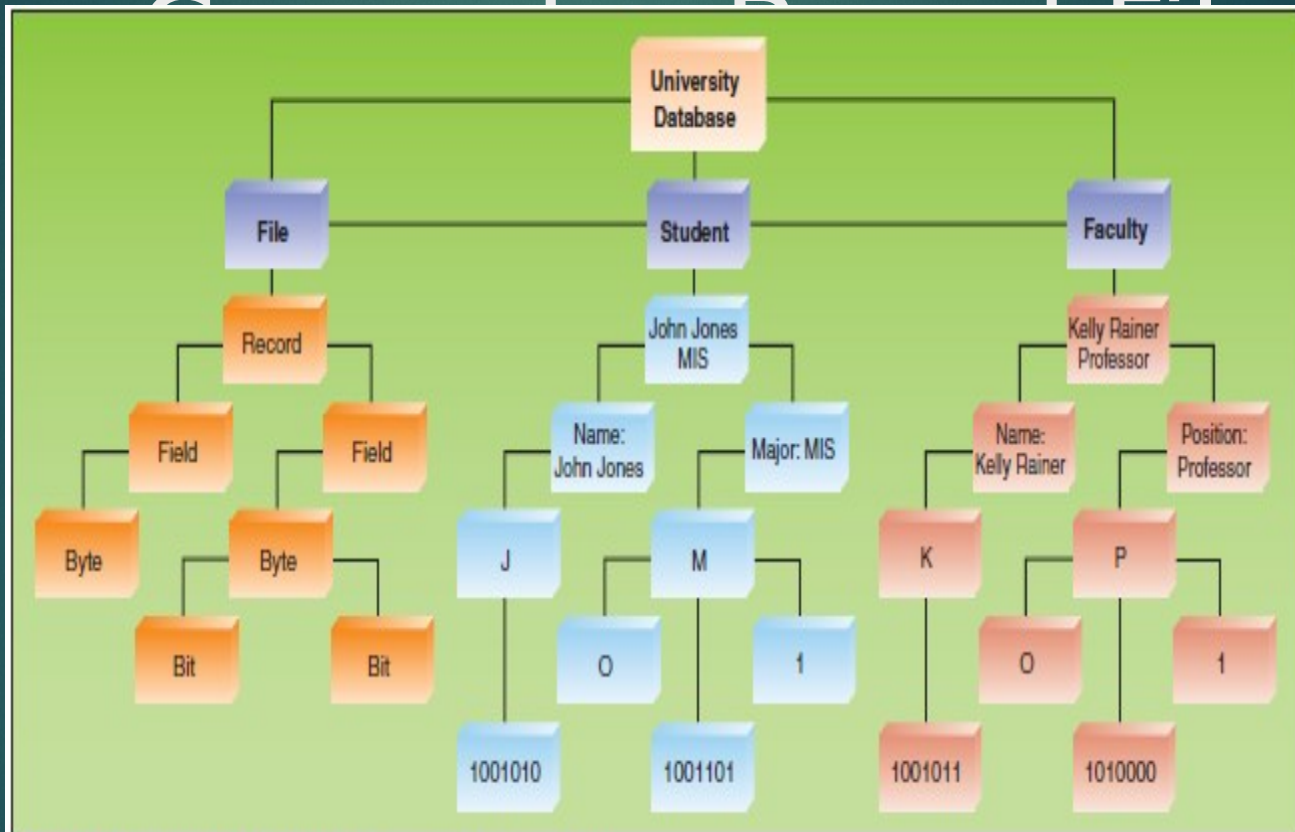
# Data Hierarchy

- A **bit** is a binary digit, or a “0” or a “1”.
- A **byte** is eight bits and represents a single character (e.g., a letter, number or symbol).
- A **field** is a group of logically related characters (e.g., a word, small group of words, or identification number).
- A **record** is a group of logically related fields (e.g., student in a university database).
- A **file** is a group of logically related records.
- A **database** is a group of logically related files.



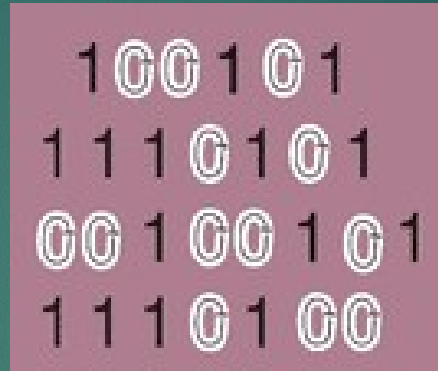
# Hierarchy of Data for

a



# Data Hierarchy (continued)

Bit (binary digit)



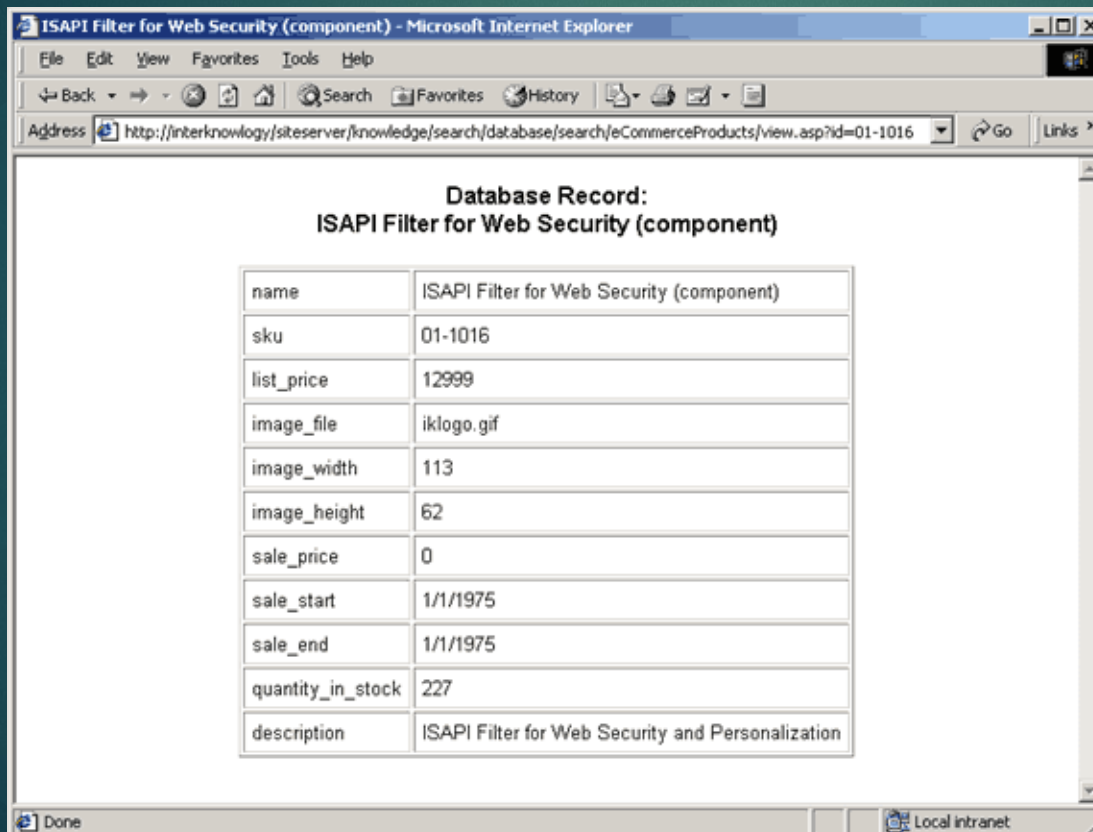
Byte (eight bits)

**01101010**



# Data Hierarchy (continued)

## Example of Field and Record



The screenshot shows a Microsoft Internet Explorer browser window with the title 'ISAPI Filter for Web Security (component) - Microsoft Internet Explorer'. The address bar displays the URL 'http://interknowlogy/siteserver/knowledge/search/database/search/eCommerceProducts/view.asp?id=01-1016'. The main content area is titled 'Database Record: ISAPI Filter for Web Security (component)' and contains a table with the following data:

Field	Value
name	ISAPI Filter for Web Security (component)
sku	01-1016
list_price	12999
image_file	iklogo.gif
image_width	113
image_height	62
sale_price	0
sale_start	1/1/1975
sale_end	1/1/1975
quantity_in_stock	227
description	ISAPI Filter for Web Security and Personalization

The browser window also shows standard menu bars (File, Edit, View, Favorites, Tools, Help) and a status bar at the bottom with 'Done' and 'Local intranet' indicators.

# Data Hierarchy (continued)

## Example of Field and Record

The screenshot shows a software window titled "IvtterrainDB - [epasites.lvt]". The menu bar includes File, Edit, View, Insert, Format, Records, Scripts, Window, and Help. On the left, there is a sidebar with a map icon, "Records: 1499", "Found: 2", and "Unsorted". The main area has a header with radio buttons for "GNIS map features", "EPA Sites" (selected), and "Locations", along with a "View Facility List" button. The facility details are as follows:

Facility Name:	RITZ TOWER HOTEL				
Address:	465 PARK AVENUE				
City:	NEW YORK	State:	NY	Zip:	10022
EPA ID:	110001609536	Latitude:	+40.761670	Longitude:	-73.970940
State-County Code:	36061				
SIC Code:	6514	OPER OF DWELL OTHER THAN APART			

Below the details, a section titled "----- EPA programs this facility is regulated under -----" contains the following fields:

Air Emissions:	<input checked="" type="checkbox"/>
Superfund Site:	<input type="checkbox"/>
Wastewater Discharges:	<input type="checkbox"/>
Hazardous Waste:	<input type="checkbox"/>
Toxic Release Inventory:	<input type="checkbox"/>

At the bottom, there are two buttons: "Show on Map" and "Show on Envirofacts Web Site".

# Designing the Database

## Data model

- ▶ The **data model** is a diagram that represents the entities in the database and their relationships.
- ▶ An **entity** is a person, place, thing, or event about which information is maintained. A record generally describes an entity.
- ▶ An **attribute** is a particular characteristic or quality of a particular entity.
- ▶ The **primary key** is a field that uniquely identifies a record.
- ▶ **Secondary keys** are other field that have some identifying information but typically do not identify the file with complete accuracy.



# Entity-Relationship Modeling

Database designers plan the database design in a process called **entity-relationship (ER) modeling**.

**ER diagrams** consists of entities, attributes and relationships.

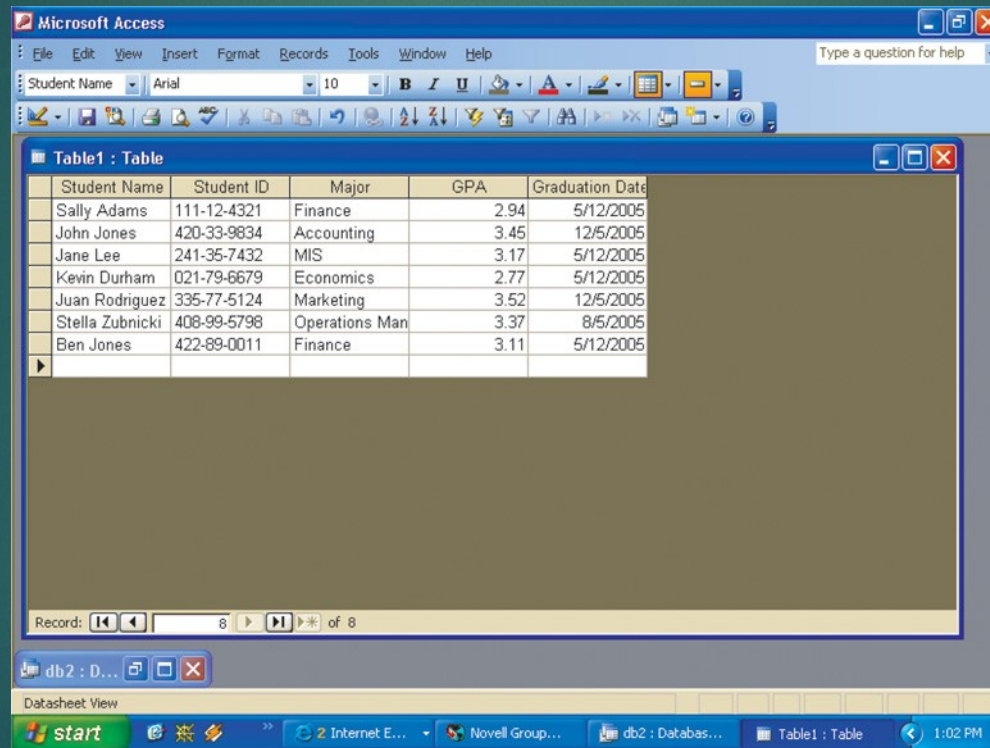
- ▶ **Entity classes** are groups of entities of a certain type.
- ▶ An **instance** of an entity class is the representation of a particular entity.
- ▶ Entity instances have **identifiers**, which are attributes that are unique to that entity instance.

# Database Management Systems



- ▶ A **database management system** is a set of programs that provide users with tools to add, delete, access, and analyze data stored in one location.
- ▶ The **relational database model** is based on the concept of two-dimensional tables.
- ▶ **Structured query language** allows users to perform complicated searches by using relatively simple statements or keywords.
- ▶ **Query by example** allows users to fill out a grid or template to construct a sample or description of the data he or she wants.

# Student Database Example



The screenshot shows the Microsoft Access application window. The title bar reads "Microsoft Access". The menu bar includes File, Edit, View, Insert, Format, Records, Tools, Window, and Help. The toolbar shows various icons for editing and formatting. The main window displays a table named "Table1 : Table" with the following data:

Student Name	Student ID	Major	GPA	Graduation Date
Sally Adams	111-12-4321	Finance	2.94	5/12/2005
John Jones	420-33-9834	Accounting	3.45	12/5/2005
Jane Lee	241-35-7432	MIS	3.17	5/12/2005
Kevin Durham	021-79-6679	Economics	2.77	5/12/2005
Juan Rodriguez	335-77-5124	Marketing	3.52	12/5/2005
Stella Zubnicki	408-99-5798	Operations Man	3.37	8/5/2005
Ben Jones	422-89-0011	Finance	3.11	5/12/2005

At the bottom of the table, there is a status bar that reads "Record: 8 of 8". The taskbar at the bottom shows the Start button, several open applications (Internet Explorer, Novell Group..., db2 : Databas..., Table1 : Table), and the system clock showing 1:02 PM.



# Normalization

- ▶ **Minimum redundancy**
- ▶ **Maximum data integrity**
- ▶ **Best processing performance**

Normalized data occurs when attributes in the table depend only on the primary key.

**Normalization** is a method for analyzing and reducing a relational database to its most streamlined form for minimum redundancy, maximum data integrity, and best processing performance.



# Non-Normalized Relation

Microsoft Access - [Non-Normalized Relation : Table]

File Edit View Insert Format Records Tools Window Help

Type a question for help

Order Number Arial 10 B I U

	Ord Nu	Part N	Part Description	Unit	Sup	Supplier Name	Supplier Address	Order Da	Delivery	Order	Cust	Customer	Customer Add
▶ 11	1	1374	Left side panel	150	15	AAA Automotive	123 Main St.	1/15/06	2/1/06	950	121	J. Smith	14 Wind St
11	1	1759	Hood	225	15	AAA Automotive	123 Main St.	1/15/06	2/1/06	950	121	J. Smith	14 Wind St
11	2	2273	Head Light	75	17	NAPA	178 Green Ave.	1/15/06	2/1/06	950	121	J. Smith	14 Wind St
11	1	3451	Grill	425	15	AAA Automotive	123 Main St.	1/15/06	2/1/06	950	121	J. Smith	14 Wind St
12	1	2394	Windshield	550	19	All About Glass	145 Highway 1	1/31/06	2/15/06	1775	122	S. Page	110 Fifth Av
12	2	1125	Windshield Wip	25	17	NAPA	178 Green Ave	1/31/06	2/15/06	1775	122	S. Page	110 Fifth Av
12	1	1759	Hood	225	15	AAA Automotive	123 Main St	1/31/06	2/15/06	1775	122	S. Page	110 Fifth Av
12	1	1888	Roof panel	650	15	AAA Automotive	123 Main St	1/31/06	2/15/06	1775	122	S. Page	110 Fifth Av
12	1	1374	Left side panel	150	15	AAA Automotive	123 Main St	1/31/06	2/15/06	1775	122	S. Page	110 Fifth Av
12	1	1375	Right side pane	150	15	AAA Automotive	123 Main St	1/31/06	2/15/06	1775	122	S. Page	110 Fifth Av
13	4	1655	Radial tires	175	21	Tire World	153 Highway 1	2/2/06	2/8/06	1650	129	R. Cox	19 Mall Dr.
13	4	1699	Chrome Hubcap	225	29	Chrome Center	197 Beulah Ave	2/2/06	2/8/06	1650	129	R. Cox	19 Mall Dr.
13	1	1991	Gas cap	50	17	NAPA	178 Green Ave	2/2/06	2/8/06	1650	129	R. Cox	19 Mall Dr.
14	1	1766	Trunk lid	450	15	AAA Automotive	123 Main St	2/9/06	2/26/06	1625	133	T. Carr	92 Star Ct.
14	1	2395	Rear windshield	550	19	All About Glass	145 Highway 1	2/9/06	2/26/06	1625	133	T. Carr	92 Star Ct.
14	2	2274	Tail Light	65	17	NAPA	178 Green Ave	2/9/06	2/26/06	1625	133	T. Carr	92 Star Ct.
14	1	2497	Rear bumper	495	15	AAA Automotive	123 Main St	2/9/06	2/26/06	1625	133	T. Carr	92 Star Ct.
*													

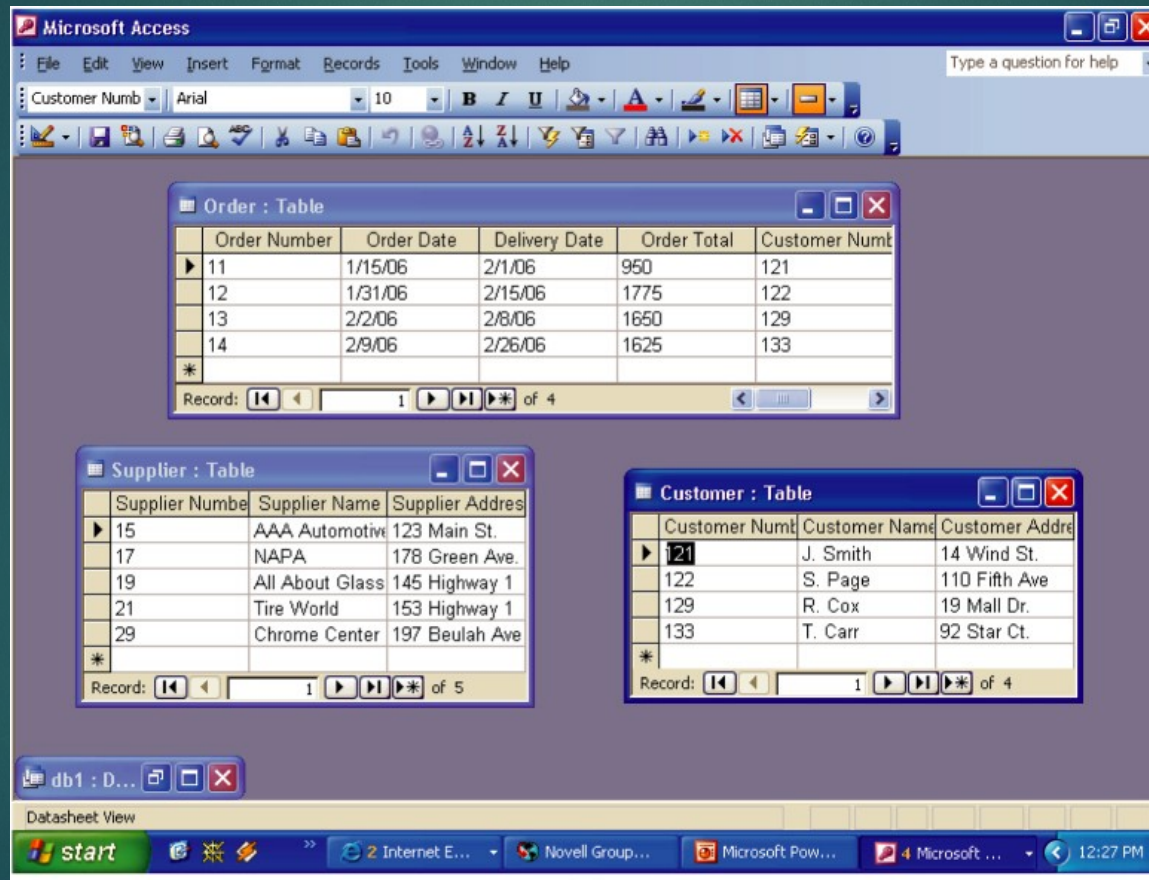
Record: 14 of 18

Datasheet View

start Internet E... Novell Group... Microsoft Pow... 4 Microsoft ... 12:34 PM



# Normalizing the Database (part A)





# Normalizing the Database (part B)

The screenshot shows the Microsoft Access interface with two tables displayed in Datasheet View. The 'Ordered Parts : Table' is on the left, and the 'Part : Table' is on the right. The 'Part : Table' has a record with 'Part Number' 1699 highlighted.

**Ordered Parts : Table**

Order Number	Part Number	Number of Parts
12	1125	2
11	1374	1
12	1374	1
12	1375	1
13	1655	4
13	1699	4
11	1759	1
12	1759	1
14	1766	1
12	1888	1
13	1991	1
11	2273	2
14	2274	2
12	2394	1
14	2395	1
14	2497	1
11	3451	1

Record: 1 of 17

**Part : Table**

Part Number	Part Description	Unit Price	Supplier Number
1699	Chrome Hubcap	225	29
2394	Front Windshield	550	19
1991	Gas cap	50	17
3451	Grill	425	15
2273	Head Light	75	17
1759	Hood	225	15
1374	Left side panel	150	15
1655	Radial tire	175	21
2497	Rear bumper	495	15
2395	Rear windshield	550	19
1375	Right side panel	150	15
1888	Roof panel	650	15
2274	Tail Light	65	17
1766	Trunk lid	450	15
1125	Windshield Wip	25	17

Record: 1 of 15

Taskbar: start, 2 Internet E..., Novell Group..., Microsoft Pow..., 3 Microsoft ..., 12:30 PM