Chapter 1: Introduction

Why Study Database as a Topic?

- A company's database is its biggest asset
- Database processing is the heart of all applications today

- It's important to understand what is stored in a database and what can be retrieved from it

- So, if we are going to create a database, what is going into the database?

  – Data

Data vs. Information

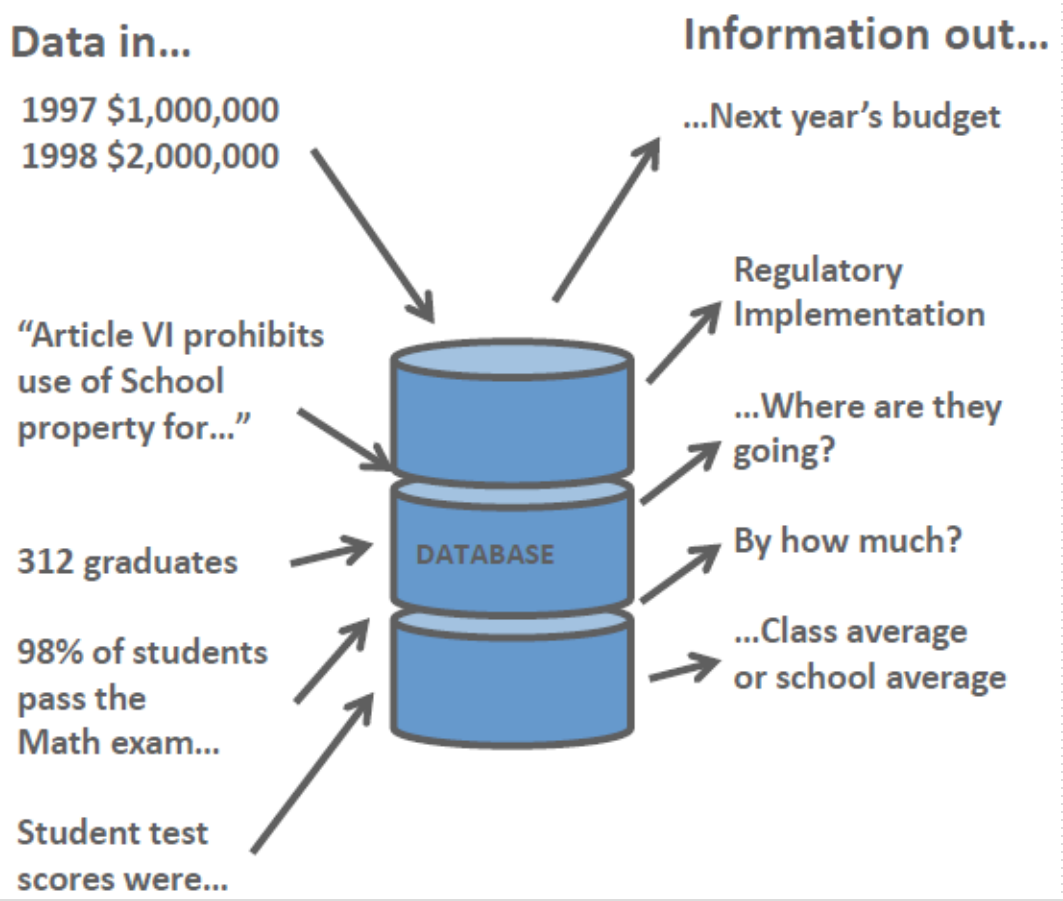
- "data" and "information" are often used to mean the same thing
- However, they have different meanings in database terminology:
- What are data?
  - Unorganized or unprocessed raw facts
- What is Information?
  - Data with special meaning
  - The result of sorting, combining, comparing, analyzing or performing calculations on data (raw facts)

Data versus Information

Data vs. Information

Data: Each student's exam score is one piece of data

Information: The class average for the exam is individual pieces or data transformed into information



Data vs. Information Examples

- Data
  - (1) Students test scores
  - (2) Weekly sales for each salesperson
  - (3) Inventory count for each product at each warehouse
- Information
  - (1) Class average
  - (2) Bonus as a percent of sales for the week
  - (3) The total company inventory for each product

Data versus Information

- Is your student number data or information?
  - When is it data?
  - When is it information?

- Is your grade in this course data or information?
  - When is it data?
  - When is it information?

Database Management System (DBMS)

- What is a Database Management System (DBMS)?

- A collection of software programs that control the storage, organization, and retrieval of data in a database

- Gives us the ability to access and manipulate data without knowledge of the structure of the database

Relational Database

- What is a Relational Database?

- An organized collection of related tables containing data that can be manipulated and tables joined to provide information to users
- Provides facilities for:
  - Retrieving, adding, modifying, and deleting data
  - Transforming retrieved data into meaningful information

NULL Value

- Some data stored in a database can contain a NULL value
- What is a NULL value?

NULL Value

- NULL means missing, unknown, or unassigned
- It is not zero (numeric) or space (alphanumeric) – it is NULL
- Consider a Customer table:
  - Missing – Perhaps a customer, such as Sally, does not divulge her age to the customer service representative
  - Unknown – An employee's termination date is usually some event in the unforeseen future
  - Unassigned (doesn't apply) – If the customer is a business, then Gender does not apply and thus is unassigned
- Data that can be NULL, is optional data

Database Development Life Cycle (DDLC)

1. Database Planning
2. Requirements Analysis
3. Database Design
4. Database Build
5. Database Testing
6. Database Deployment
7. Database Maintenance

Step 1: Database Planning

- Starts when a customer (user) submits a request for the development of a database
- Four major activities are performed:
  - Review and approve the database project request
  - Prioritize the database project request
  - Allocate resources such as money, people, and tools
  - Assign development team to develop the database project

Step 2: Requirements Analysis

- Also known as the systems analysis phase
- Includes investigation and analysis of the request
- Results in a set of requirements that the database must support
- Includes:
  - What data is to be stored
  - What are the relationships between the data
  - What processes are involved
  - Business rules

Step 3: Database Design

- Process of creating a detailed data model of the proposed database
- Three common phases in database modeling:
  - Entity Relationship Diagram (ERD)
  - Normalization
  - Relational Data Model

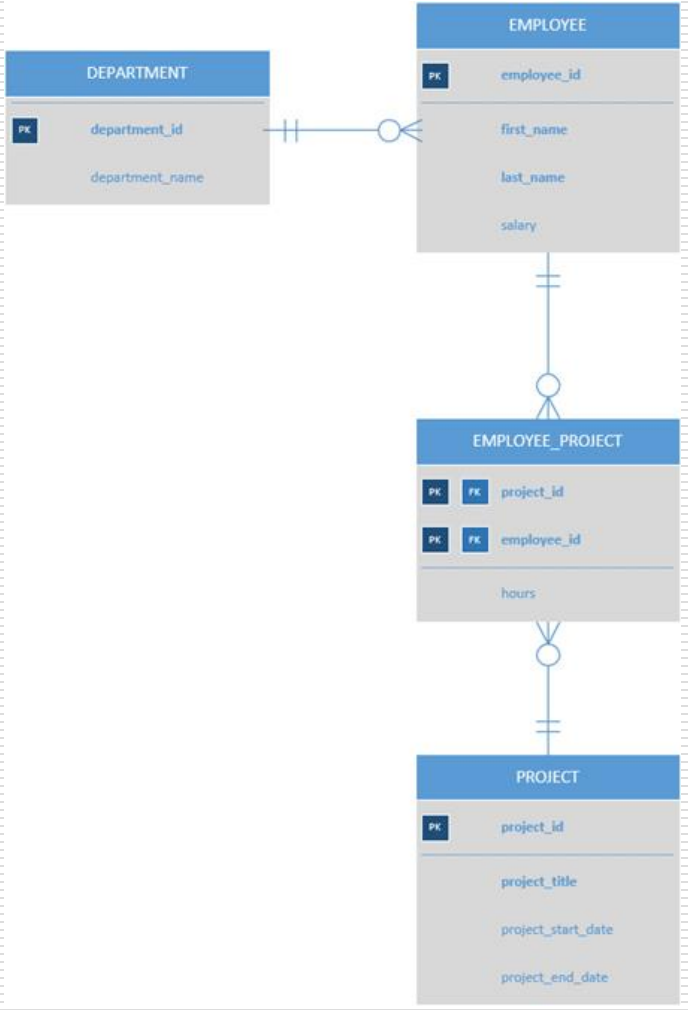
Entity Relationship Diagram (ERD)

- Graphical representation of the proposed database
- Stays the same regardless of what type of DBMS the system is eventually built with

Entity Relationship Diagram (ERD)

- Answers the questions:
  - What entities (person, place, or thing) are being represented?
  - What attributes (data) are stored about each entity?
  - What are the relationships between the entities?
- Basically, what data do we want to capture and what are the business rules surrounding that data
- Crow's foot notation
  - Information Engineering (IE) notation
  - Barker notation

Example ERD


DEPARTMENT
PK
department_id
department_name
EMPLOYEE
PK
employee_id
first_name
last_name
salary
EMPLOYEE_PROJECT
PK
FK
project_id
PK
FK
employee_id
hours
PROJECT
PK
project_id
project_title
project_start_date
project_end_date

Normalization

- A technique used during database design to identify redundancy within the database

- Objective is to correct un-normalized relations (tables that contain repeating groups) using a process of normal forms
- Chapter 5

Relational Data Modeling

- A data model expressed in terms of a relational database structure (DB2, Oracle, MS SQL Server, MySQL, and others)
- The ER data model is transformed into a relational data model based on the DBMS being used for implementation
- The transformation of the ER diagram representing entities, attributes, and relationships into a relational data model representing tables, columns, an relationships

Step 4: Database Build

- Create the database
  - SQL Data Definition Language (DDL) statements
    - For the selected DBMS according to the requirements specified in the physical data model
    - DBMS specific – DB2 is used in this course
  - Implement integrity constraints from business rules

Step 5: Database Testing

- Test all constraints
- Verify that all requirements have been met

Step 6: Database Deployment

- Allocate storage requirements
- Place the database into production

Step 7: Database Maintenance

- Database Maintenance
  - Maintain the database on an on-going basis according to user requirements

Introduction to SQL Statements

- SQL pronounced "S-Q-L" stands for Structured Query Language
- The industry-standard language of relational database management systems (RDBMS) for manipulating and querying data in a relational database

SQL Statements

- SQL commands are categorized into four categories:
  - DDL (Data Definition Language)
  - DML (Data Manipulation Language)
  - DCL (Data Control Language)
  - TCL (Transaction Control Language)

DDL (Data Definition Language)

- SQL commands used to create and modify the structure of database objects:
  - CREATE – create the database or its objects (table, index, function, views, store procedure and triggers)
  - DROP – delete objects from the database
  - ALTER – alter the structure of the database
  - TRUNCATE – remove all rows (records) from a table
  - COMMENT – add comments to the data dictionary
  - RENAME – Rename an object existing in the database

DML (Data Manipulation Language)

- SQL commands that deal with the manipulation of data in a database:
  - SELECT – retrieve data from the a database
  - INSERT – insert data into a table
  - UPDATE – update existing data within a table
  - DELETE – delete rows (records) from a database table

DCL (Data Control Language)

- SQL commands that deal with the rights, permissions, and other controls of the database system:
  - GRANT – gives user's access privileges to database
  - REVOKE – withdraw user's access privileges given by using the GRANT command

TCL (Transaction Control Language)

- SQL commands that deal with the transactions within the database:
  - `COMMIT`– commits a transaction
  - `ROLLBACK` – rollbacks or reverses a transaction when errors occur
  - `SAVEPOINT` – sets a SAVEPOINT within a transaction
  - `SET TRANSACTION` – specifies characteristics for a transaction

Creating a Schema

- A schema must be created before the creation of tables and other SQL objects
- A schema is an object that serves as a container for database objects, such as tables, views, indexes, stored procedures, and other object types
- A schema is created by entering the `CREATE SCHEMA` command, followed by the name of the schema, followed by a semicolon
- `CREATE SCHEMA myschema;`


Q & A
Questions & Answers