

Relational Databases

Lecture 2

Robb T. Koether

Hampden-Sydney College

Wed, Jan 17, 2018

- 1 Types of Databases
- 2 Relational Databases
- 3 Queries
- 4 Schemas
- 5 Primary Keys

Types of Databases

- Originally, “databases” were simply collections of data, arranged in whatever way the programmer saw fit.
- The “DBMS” was simply a program that could process the data.

Types of Databases

- In the early 1970s, the concept of a **relational database (RDB)** was introduced.
- Its data were organized into tables of rows and columns.
- Queries were organized into select, insert, delete, and update.
- Given the standardized forms, queries could be highly optimized by the DBMS.
- Today, SQL (Structured Query Language) is the standard for relational databases.

Types of Databases

Item	Price	Stock
123	3.99	1000
456	6.99	1500
:	:	:

Catalog of Items

Cust	Item	Qty
Alice	123	2
Bob	123	4
Bob	456	1
:	:	:

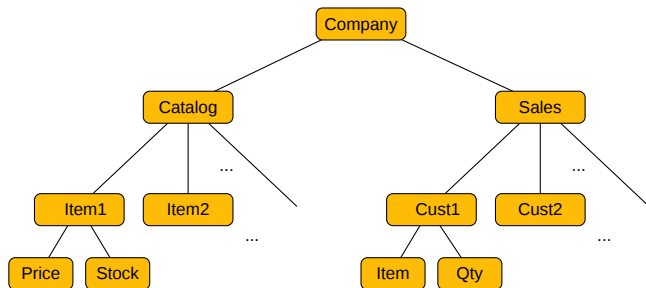
Sales

A Relational Database (SQL)

Types of Databases

- More recently, other forms have appeared, most notably **Extensible Markup Language (XML)**.
- An XML database is structured hierarchically.
- The underlying theory is less mature.
- The supporting software is less extensive.
- The queries are less optimized.
- On the other hand, the structure is much more flexible.

Types of Databases



A Hierarchical Database (XML)

Relational Databases

- A **relational database** is a database in which the data are stored in tabular form, arranged in rows and columns.
- Traditionally, rows are called **records** and columns are called **fields**.
- In modern terminology, rows are called **tuples** and columns are called **attributes**.

Relational Databases

- Consider a database of company employees.

fname	lname	ssn	bdate	sex	salary	dept
Alice	Smith	123-45-6789	1968-05-22	F	35000.00	2
Barbara	Brown	135-79-2468	1985-10-12	F	40000.00	3
James	Green	246-81-3579	1974-02-15	F	100000.00	1
Jennifer	Wallace	321-54-9876	1985-12-02	F	50000.00	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Employees table

Relational Databases

- In the example, the **tuples** are
 - (Alice, Smith, 123-45-6789, 1968-05-22, ...)
 - (Barbara, Brown, 135-79-2468, 1985-10-12, ...)
 - etc.
- The **attributes** are
 - fname
 - lname
 - ssn
 - bdate
 - etc.

Relational Databases

- The table expresses a **relation**, namely
“The employee *fname lname* has SS number *ssn*, was born on *bdate*, is *sex*, earns a salary of *salary*, and works in department *dept*.”
- For example, the first row represents the statement
“The employee *Alice Smith* has SS number *123456789*, was born on *1968-05-22*, is *female*, earns a salary of *\$35000.00*, and works in department *#2*.”
- Thus, each tuple of the database represents a true statement, an **instance** of the relation.

Relational Databases

- A **query** is a request that is sent to the DBMS.
- The DBMS parses the query, optimizes it, and sends it to the **storage engine**, which executes the query on the database.
- There are four basic types of query.
 - Select
 - Insert
 - Delete
 - Update

Relational Databases

- A **selection** query might request all employees in Dept #2.
- The result would be

fname	lname	ssn	bdate	sex	salary	dept
Alice	Smith	123-45-6789	1968-05-22	F	35000.00	2
Jennifer	Wallace	321-54-9876	1985-12-02	F	50000.00	2
Ernest	Roth	642-18-9753	1986-06-12	M	60000.00	2
John	Kohler	789-01-2345	1966-11-24	M	40000.00	2
Raymond	Jones	963-41-8527	1974-08-30	M	80000.00	2

Employees table

Relational Databases

- An **insertion** query might add the tuple
(Susan, Lane, 654-87-2109, 1959-03-31, F, 70000.00, 2).
- The result would be

fname	lname	ssn	bdate	sex	salary	dept
Alice	Smith	123-45-6789	1968-05-22	F	35000.00	2
Jennifer	Wallace	321-54-9876	1985-12-02	F	50000.00	2
Ernest	Roth	642-18-9753	1986-06-12	M	60000.00	2
John	Kohler	789-01-2345	1966-11-24	M	40000.00	2
Raymond	Jones	963-41-8527	1974-08-30	M	80000.00	2
Susan	Lane	654-87-2109	1959-03-31	F	70000.00	2

Employees table

Relational Databases

- A **deletion** query might delete the tuple with SSN 654872109.
- The result would be

fname	lname	ssn	bdate	sex	salary	dept
Alice	Smith	123-45-6789	1968-05-22	F	35000.00	2
Jennifer	Wallace	321-54-9876	1985-12-02	F	50000.00	2
Ernest	Roth	642-18-9753	1986-06-12	M	60000.00	2
John	Kohler	789-01-2345	1966-11-24	M	40000.00	2
Raymond	Jones	963-41-8527	1974-08-30	M	80000.00	2

Employees table

Relational Databases

- An **update** query might change the department number from 2 to 5.
- The result would be

fname	lname	ssn	bdate	sex	salary	dept
Alice	Smith	123-45-6789	1968-05-22	F	35000.00	5
Jennifer	Wallace	321-54-9876	1985-12-02	F	50000.00	5
Ernest	Roth	642-18-9753	1986-06-12	M	60000.00	5
John	Kohler	789-01-2345	1966-11-24	M	40000.00	5
Raymond	Jones	963-41-8527	1974-08-30	M	80000.00	5

Employees table

Schemas

Definition (Schema)

A **relation schema** consists of the relation's name, a list of its attributes, and their **domains**, i.e., the set of legitimate values of each attribute. A **database schema** is a set of relation schemas.

Schemas

- The schema of the Employees relation is

Employees

```
(  
    fname string,  
    lname string,  
    ssn string,  
    bdate date,  
    sex string,  
    salary float,  
    dept integer  
)
```

- Furthermore, $1 \leq \text{dept} \leq 5$ and `sex` is 'M' or 'F'.

Primary Keys

- Every table must have a **primary key**.
- The primary key is an attribute, or set of attributes, whose value, or values, is unique for each tuple.
- In a relation schema, we underline the attributes that comprise the primary key.
- What would serve as the primary key in the `Employees` table?