RELATIONAL DATA MODEL

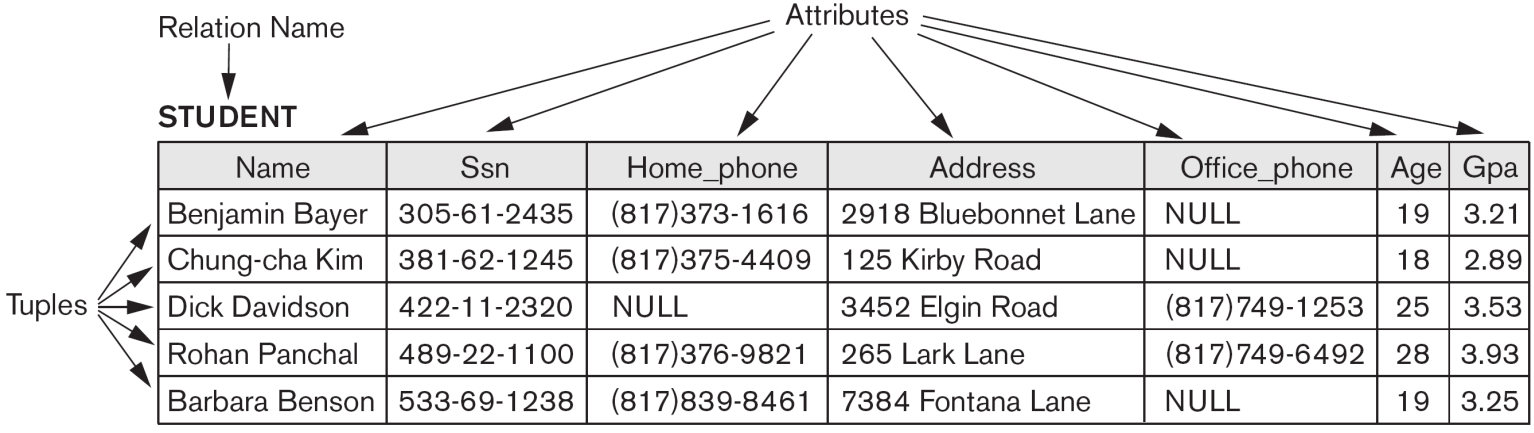
Corresponding Reading: Chapter 5.1

Relational Model

- First introduced by Ted Codd (IBM Research – 1970)
  - Model uses the concept of a mathematical **relation** as its basic building block
    - Looks similar to a table of values
    - Theoretical basis is set theory and predicate logic
- Simple model with high level query language.
- First commercial implementation in 1980's
- Today: Oracle DBMS, MySQL, PostgreSQL

Basic Construct

- **Relation** (also referred to as **table**).
- Each relation has a name.
- Each relation has a set of named **attributes** (or columns).
- Each **tuple** (or **row**) has a value for each attribute.
- Each attribute has a **type** (or **domain**).



Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	(817)373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	(817)375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	(817)749-1253	25	3.53
Rohan Panchal	489-22-1100	(817)376-9821	265 Lark Lane	(817)749-6492	28	3.93
Barbara Benson	533-69-1238	(817)839-8461	7384 Fontana Lane	NULL	19	3.25

University Database

- Database is a set of named relations (or tables).

STUDENT

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	(817)749-1253	25	3.53
Barbara Benson	533-69-1238	(817)839-8461	7384 Fontana Lane	NULL	19	3.25
Rohan Panchal	489-22-1100	(817)376-9821	265 Lark Lane	(817)749-6492	28	3.93
Chung-cha Kim	381-62-1245	(817)375-4409	125 Kirby Road	NULL	18	2.89
Benjamin Bayer	305-61-2435	(817)373-1616	2918 Bluebonnet Lane	NULL	19	3.21

- COURSE

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

Relation Schema

- **Schema** is a structural description of relations in the database: name, attributes, types of these attributes.
- **Instance** actual contents of a table at any given time.
- Schema: constant; Instance: varies.

Domains

- Each attribute has a domain (attribute type).
- Name the domain.
  - Examples:
    - USA_phone_numbers: The set of 10 digit phone numbers valid in USA.
    - Social_security_numbers: The set of valid 9 digit SSN
    - Names: The set of character strings that represent names of persons
    - Academic_dept_codes: Department codes, such as 'CS', 'ECON', 'ENG'
- A domain is given a name, data type, and format.
  - Additional information can be given to help interpret values
    - Example: Weight in pounds or kilograms, etc.
    - age: integer x; $0 < x < 100$.

Ordering of Tuples

- Tuples in a relation do not have a particular order.
- The relation is not sensitive to the ordering of tuples.
- The order of attributes/values in a tuple is important.

STUDENT

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	(817)749-1253	25	3.53
Barbara Benson	533-69-1238	(817)839-8461	7384 Fontana Lane	NULL	19	3.25
Rohan Panchal	489-22-1100	(817)376-9821	265 Lark Lane	(817)749-6492	28	3.93
Chung-cha Kim	381-62-1245	(817)375-4409	125 Kirby Road	NULL	18	2.89
Benjamin Bayer	305-61-2435	(817)373-1616	2918 Bluebonnet Lane	NULL	19	3.21

Formal definition of a Tuple

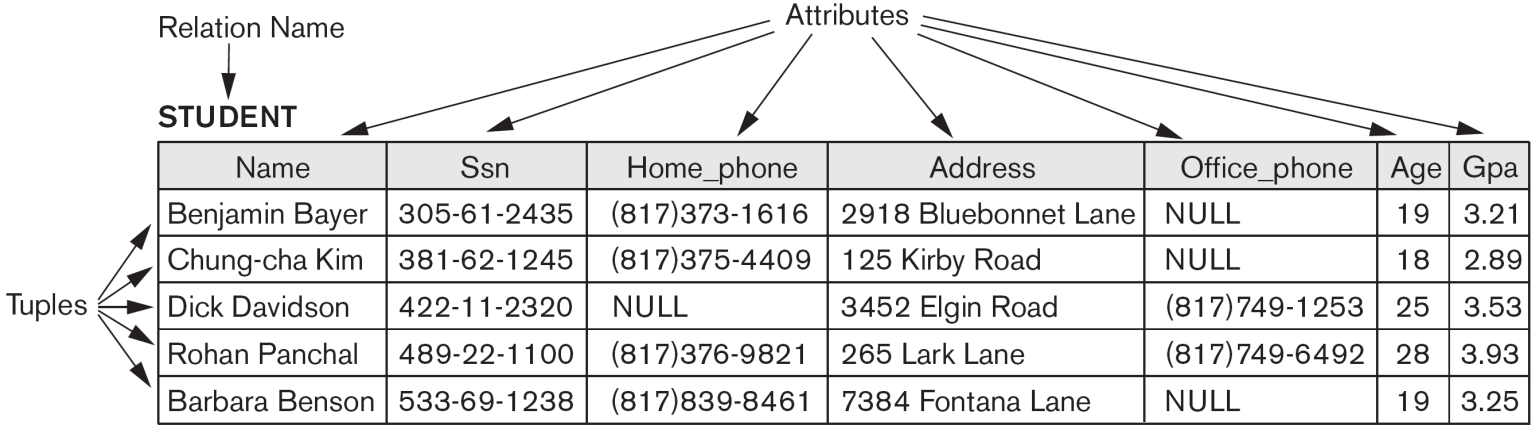
- A tuple can be considered as a set of (<attribute>,<value>) pairs, where each pair gives the value of the mapping from an attribute A_i to a value v_i from $dom(A_i)$.
- The ordering of attributes is not important since the attribute name appears with its value.
  - Example: These two tuples are identical.

$t = <$ (Name, Dick Davidson),(Ssn, 422-11-2320),(Home_phone, NULL),(Address, 3452 Elgin Road), (Office_phone, (817)749-1253),(Age, 25),(Gpa, 3.53)>

$t = <$ (Address, 3452 Elgin Road),(Name, Dick Davidson),(Ssn, 422-11-2320),(Age, 25), (Office_phone, (817)749-1253),(Gpa, 3.53),(Home_phone, NULL)>

Null Value

- NULL: unknown or undefined
  - query: students with phone area code (817)
  - query: students whose phone area code not (817)



Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	(817)373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	(817)375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	(817)749-1253	25	3.53
Rohan Panchal	489-22-1100	(817)376-9821	265 Lark Lane	(817)749-6492	28	3.93
Barbara Benson	533-69-1238	(817)839-8461	7384 Fontana Lane	NULL	19	3.25

Key

- All tuples in a relation must be distinct (*set definition*).
- relation **key:** attribute whose value is unique in each tuple OR set of attributes that together can uniquely identify each tuple in a relation.
  - Example: SSN is a key of the STUDENT relation
  - Example: Course #, Dept #

Key

- A relation may have more than one key. Each of the keys is called a **candidate key**.
- Designate one of the candidate keys as the **primary key** of the relation. (Often arbitrary choice)
  - Attributes that form the primary key are underlined, however the other candidate keys are not.

Example

- CAR relation
  - Candidate Keys are: License_number and Engine_serial_number

CAR

License_number	Engine_serial_number	Make	Model	Year
Texas ABC-739	A69352	Ford	Mustang	02
Florida TVP-347	B43696	Oldsmobile	Cutlass	05
New York MPO-22	X83554	Oldsmobile	Delta	01
California 432-TFY	C43742	Mercedes	190-D	99
California RSK-629	Y82935	Toyota	Camry	04
Texas RSK-629	U028365	Jaguar	XJS	04

Relational Database Schema

- Example: COMPANY relational DB schema

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date

DEPT_LOCATIONS

Dnumber	Dlocation

PROJECT

Pname	Pnumber	Plocation	Dnum

WORKS_ON

Essn	Pno	Hours

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship

Summary

- Examined the relational model of data
- Relational data model is used in most DBMS systems today including MySQL, which we will use
- These concepts will directly relate to real operations that can be performed with SQL (Structured Query Language) on real databases.