

# Database Management System – 14 (Relation - Formal and Informal Definitions)

Ajay James  
Asst. Prof in CSE  
Government Engineering College Thrissur

## Outline

- Relations
- Informal Definitions
- Formal Definitions
- Characteristics of Relations

## Relational Data Model

- Based on the concept of a ***Relation***
- Mathematical concept based on the ideas of sets
- Strength of the relational approach
  - Formal foundation provided by the theory of relations
- Model was first proposed by Dr. E.F. Codd of IBM in 1970
  - "A Relational Model for Large Shared Data Banks,"  
Communications of the ACM, June 1970
- *ACM Turing Award*

## Relations

### STUDENT

Name	Student_number	Class	Major
Smith	17	1	CS
Brown	8	2	CS

### 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

### SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	07	King
92	CS1310	Fall	07	Anderson
102	CS3320	Spring	08	Knuth
112	MATH2410	Fall	08	Chang
119	CS1310	Fall	08	Anderson
135	CS3380	Fall	08	Stone

## INFORMAL DEFINITIONS

- **RELATION:** A table of values
  - a **set of rows**
  - a **set of columns**
  - Each row represents an **entity** or **relationship**
  - Each row has a value of an item or set of items that **uniquely** identifies that row in the table
  - Sometimes **row-ids** or sequential numbers are assigned to identify the rows in the table
  - Each column typically is called by its column name or column header or **attribute name**

## FORMAL DEFINITIONS

- **Schema** of a Relation:  $R(A_1, A_2, \dots, A_n)$
- Relation schema  $R$  is defined over **attributes**  $A_1, A_2, \dots, A_n$
- $D$  is called the **domain of  $A_i$**  and is denoted by  **$\text{dom}(A_i)$**
- $R$  is called the **name** of this relation
- **Degree (or arity)** of a relation is the number of attributes  $n$  of its relation schema
- Example -  
 $\text{STUDENT}(\text{Name}, \text{Ssn}, \text{Home\_phone}, \text{Address}, \text{Office\_phone}, \text{Age}, \text{Gpa})$
- $\text{STUDENT}(\text{Name: string}, \text{Ssn: string}, \text{Home\_phone: string}, \text{Address: string}, \text{Office\_phone: string}, \text{Age: integer}, \text{Gpa: real})$

## FORMAL DEFINITIONS

- **Tuple** is an ordered set of values
- CUSTOMER (Cust-id, Cust-name, Address, Phone#)
- **<632895, "John Smith", "101 Main St. Atlanta, GA 30332", "(404) 894-2000">**
- A relation may be regarded as a **set of tuples** (rows)
- Columns in a table are also called attributes of the relation

## Example

Diagram illustrating the structure of a relation (table) with annotations:

- Relation Name:** STUDENT
- Attributes:** Name, Ssn, Home\_phone, Address, Office\_phone, Age, Gpa
- Tuples:** The rows of data in the table.

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

## FORMAL DEFINITIONS

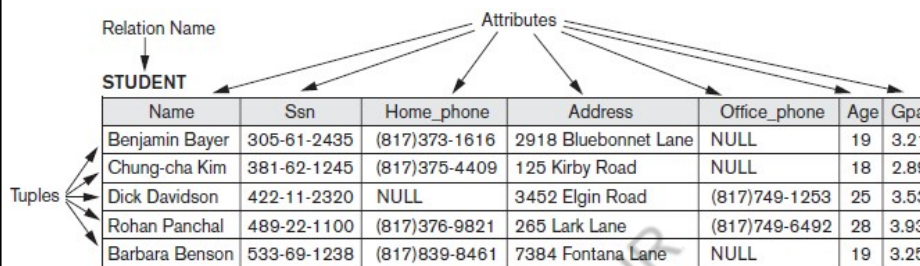
- **Domain**  
“USA\_phone\_numbers” are the set of 10 digit phone numbers valid in the U.S
- A domain may have a data-type or a format defined for it

## FORMAL DEFINITIONS

- **Relation (or relation state)  $r$**  of the relation schema  $R(A_1, A_2, \dots, A_n)$ , also denoted by  $r(R)$ , is a set of  $n$ -tuples  $r = \{t_1, t_2, \dots, t_m\}$
- Each tuple  $t$  is an ordered list of  $n$  values  $t = \langle v_1, v_2, \dots, v_n \rangle$ , where each value  $v_i$ ,  $1 \leq i \leq n$ , is an element of  $\text{dom}(A_i)$  or is a special NULL value

## FORMAL DEFINITIONS

- A relation (or relation state)  $r(R)$  is a mathematical relation of degree  $n$
- $r(R) \subseteq (\text{dom}(A_1) \times \text{dom}(A_2) \times \dots \times \text{dom}(A_n))$



## Definition

- $R$ : schema of the relation
- $r(R)$  -  $r$  of  $R$ : a specific "value" or population of  $R$
- $R$  is also called the **intension** of a relation
- $r$  is also called the **extension** of a relation

## DEFINITION SUMMARY

<u>Informal Terms</u>		<u>Formal Terms</u>
Table		Relation
Column		Attribute/Domain
Row		Tuple
Values in a column		Domain
Table Definition		Schema of a Relation
Populated Table		Extension

## CHARACTERISTICS OF RELATIONS

- **Ordering of tuples in a relation  $r(R)$ :**
  - Tuples are *not* considered to be ordered
- **Ordering of attributes in a relation schema  $R$** 
  - We will consider the attributes in  $R(A_1, A_2, \dots, A_n)$  and the values in  $t = \langle v_1, v_2, \dots, v_n \rangle$  to be *ordered*
- **Values in a tuple:**
  - All values are considered *atomic* (indivisible)
  - A special **null** value is used to represent values that are unknown or inapplicable to certain tuples

## Ordering – alternative definition

**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

$t = \langle (\text{Name, Dick Davidson}), (\text{Ssn, 422-11-2320}), (\text{Home\_phone, NULL}), (\text{Address, 3452 Elgin Road}), (\text{Office\_phone, (817)749-1253}), (\text{Age, 25}), (\text{Gpa, 3.53}) \rangle$

$t = \langle (\text{Address, 3452 Elgin Road}), (\text{Name, Dick Davidson}), (\text{Ssn, 422-11-2320}), (\text{Age, 25}), (\text{Office\_phone, (817)749-1253}), (\text{Gpa, 3.53}), (\text{Home\_phone, NULL}) \rangle$

## Relational Model Notation

- A relation schema R of degree n is denoted by R(A1, A2, ... , An).
- Uppercase letters Q, R, S denote relation names
- Lowercase letters q, r, s denote relation states
- Letters t, u, v denote tuples
- $t = \langle \text{'Barbara Benson'}, \text{'533-69-1238'}, \text{'(817)839-8461'}, \text{'7384 Fontana Lane'}, \text{NULL}, 19, 3.25 \rangle$
- $t[\text{Name}] = \langle \text{'Barbara Benson'} \rangle$ , and  $t[\text{Ssn, Gpa, Age}] = \langle \text{'533-69-1238'}, 3.25, 19 \rangle$ .



## Reference

- Elmasri R. and S. Navathe, Database Systems: Models, Languages, Design and Application Programming, Pearson Education 6<sup>th</sup> edition and 7<sup>th</sup> edition

Thank you