

# Database System

## Relational Data Model

2<sup>nd</sup> Week

### Goal

1. Understanding the relational model

# Relational Model

## History

Introduced by Dr. Edgar Frank “Ted” Codd in 1970

A declarative method for specifying data and queries

- Set theory

- 1<sup>st</sup> order logic (predicate logic)

*“Database is a collection of relations”*

# Relational Model

## Definition

Data model that adopted the *relation* concepts

Data Structure

Relation

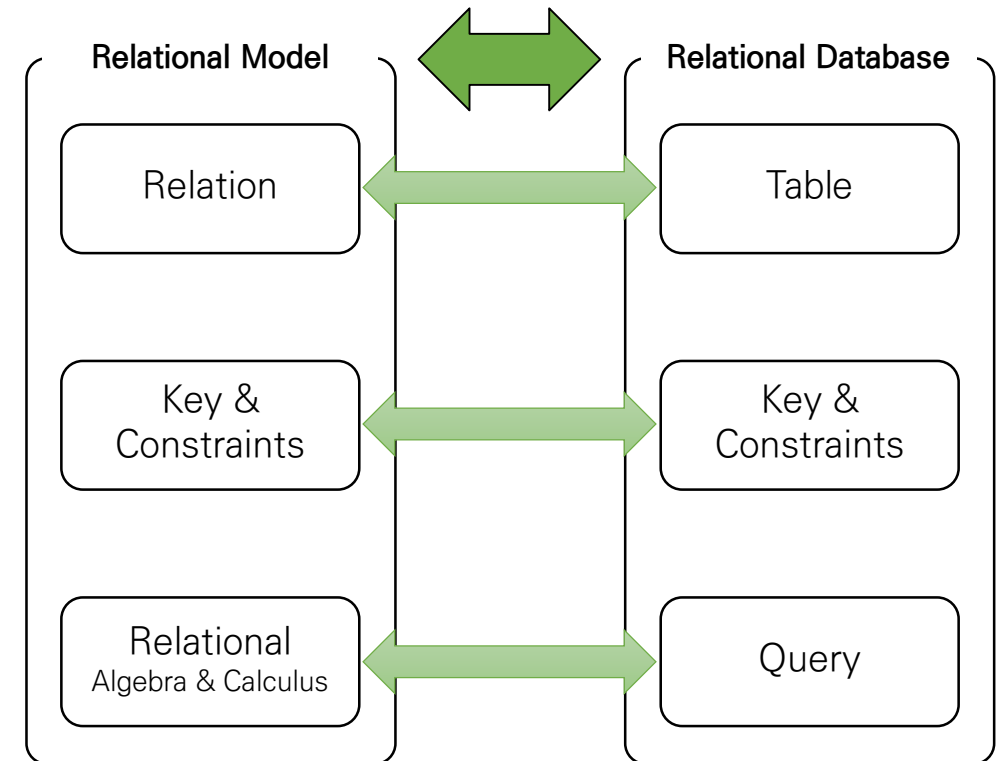
Data Integrity

Key & Constraints

Data Manipulation

Relational Algebra

Relational Calculus



# Relational Model

## Data Structure

<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

Relation  
Schema

Relation  
Intension

Head  
(Fixed)

Body  
(Time-varying)

Relation

Relation  
Instance

Relation  
Extension

⟨Instructor⟩



Relation Name

# Relational Model

## Data Structure

	Attribute				Degree
	<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>	
Tuple	10101	Srinivasan	Comp. Sci.	65000	Cardinality
	12121	Wu	Finance	90000	
	15151	Mozart	Music	40000	
	22222	Einstein	Physics	95000	
	32343	El Said	History	60000	
	33456	Gold	Physics	87000	
	45565	Katz	Comp. Sci.	75000	
	58583	Califieri	History	62000	
	76543	Singh	Finance	80000	
	76766	Crick	Biology	72000	
	83821	Brandt	Comp. Sci.	92000	
	98345	Kim	Elec. Eng.	80000	

# Relational Model

## Data Structure

Primary Key			
<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	<del>80000</del>
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000
Domain			

# Relational Model

## Data Structure

### Domain

A set of values that has a **name**, **data type** and **format**

Every domain has the special value *null*

### Relation Schema (a.k.a. Relational Intension)

$R(A_1, A_2, \dots, A_d)$ : A relation name  $R$  and a set of attributes  $\{A_d\}$

### Relation (a.k.a. Relation Extension or Relation Instance)

$r(R)$  of relation schema  $R(A_1, A_2, \dots, A_d)$ : An set of tuples  $\{t_c\}$

A tuple is an ordered list of values from domains of corresponding attributes

$t_1 = \{20211234, \text{"Anderson"}, \text{"DB Design"}\}$

# Relational Model

## Data Structure

Relation (cont.)

$r(R)$  of relation schema  $R(A_1, A_2, \dots, A_d)$ : An set of tuples  $\{t_c\}$

A tuple is a set of (attribute, value) pairs

then  $t_c: R \rightarrow D$ ,  $D = (\text{domain}(A_1) \times \text{domain}(A_2) \times \dots \times \text{domain}(A_d))$

$t_1 = \{20211234, \text{"Anderson"}, \text{"DB Design"}\}$

$t_1 = \{(\text{"ID"}, 20211234), (\text{"Name"}, \text{"Anderson"}), (\text{"Class"}, \text{"DB Design"})\}$

$$\underline{r(R) \subseteq (\text{domain}(A_1) \times \text{domain}(A_2) \times \dots \times \text{domain}(A_d))}$$



# Relational Model

## Data Structure

### Properties

1. Uniqueness of tuples
2. Uniqueness of attributes
3. No ordering of tuples
4. No ordering of attributes
5. Homogeneous data type in an attribute
  - Domain Integrity
6. Atomicity of attribute values
  - Normalization (1<sup>st</sup> Normal Form Relation)

<i>ID</i>	<i>name</i>	<i>dept_name</i>	<i>salary</i>
22222	Einstein	Physics	95000
12121	Wu	Finance	90000
32343	El Said	History	60000
45565	Katz	Comp. Sci.	75000
98345	Kim	Elec. Eng.	80000
76766	Crick	Biology	72000
10101	Srinivasan	Comp. Sci.	65000
58583	Califieri	History	62000
83821	Brandt	Comp. Sci.	92000
15151	Mozart	Music	40000
33456	Gold	Physics	87000
76543	Singh	Finance	80000

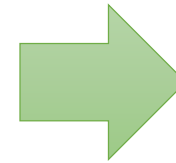
# Relational Model

## Data Structure

Name	Order
Anderson	Coffee, Cake
Andrew	Coffee
Mary	Coffee, Cookie
Kim	Cake, Cookie, Bread
Dona	Cake

Repeating 1-D Group

Normalize



Name	Order
Anderson	Coffee
Anderson	Cake
Andrew	Coffee
Mary	Coffee
Mary	Cookie
Kim	Cake
Kim	Cookie
Kim	Bread
Dona	Cake

1NF Relation

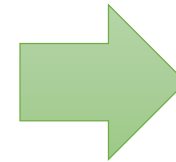
# Relational Model

## Data Structure

Name	Spendings
Anderson	{('Meal', \$6), ('Movie', \$11)}
Andrew	{('Meal', \$5)}
Mary	{('Meal', \$7), ('Book', \$15)}
Kim	{('Game', \$40)}
Dona	{('Meal', \$4), ('Game', \$11)}

Repeating 2-D Group

Normalize



Name	Category	Price
Anderson	Meal	\$6
Anderson	Movie	\$11
Andrew	Meal	\$5
Mary	Meal	\$7
Mary	Book	\$15
Kim	Game	\$40
Dona	Meal	\$4
Dona	Game	\$11

1NF Relation

# Relational Model

## Data Structure

### Key (a.k.a. Candidate Key)

Attribute (or attributes combination) that can uniquely identify every single tuples in a relation schema

Key is a minimal **superkey** for given attributes combination **Minimum**

Superkey (SK) is a subset of attributes of relation schema that satisfies  $t_i[SK] \neq t_j[SK], \forall t_i, t_j, i \neq j$  **Uniqueness**

Every relation has at least one default superkey

### Primary Key (PK)

Chosen key from candidate keys

### Alternate Key

The remaining candidate keys

### Foreign Key (FK)

An attribute that is referencing the PK of other relation schema.

# Relational Model

## Data Structure

