



# **COSC 3380**

# **Design of Database Systems**

## **Data Modeling with Entity-Relationship (ER) Model**

February 21, 2024

# Relational Model Constraints

- **Flat relational model**
  - Composite and multivalued attributes not allowed
  - Multivalued attributes
    - Must be represented by separate relations
- **Inherent model-based constraints or implicit constraints**
- **Schema-based constraints or explicit constraints**
- **Application-based or semantic constraints or business rules**



# Schema-based: Domain Constraints

- Typically include:
  - Numeric data types for integers and real numbers
  - Characters
  - Boolean
  - Fixed-length strings
  - Variable-length strings
  - Date, time, timestamp
  - Money
  - Other special data types

# Key Constraints

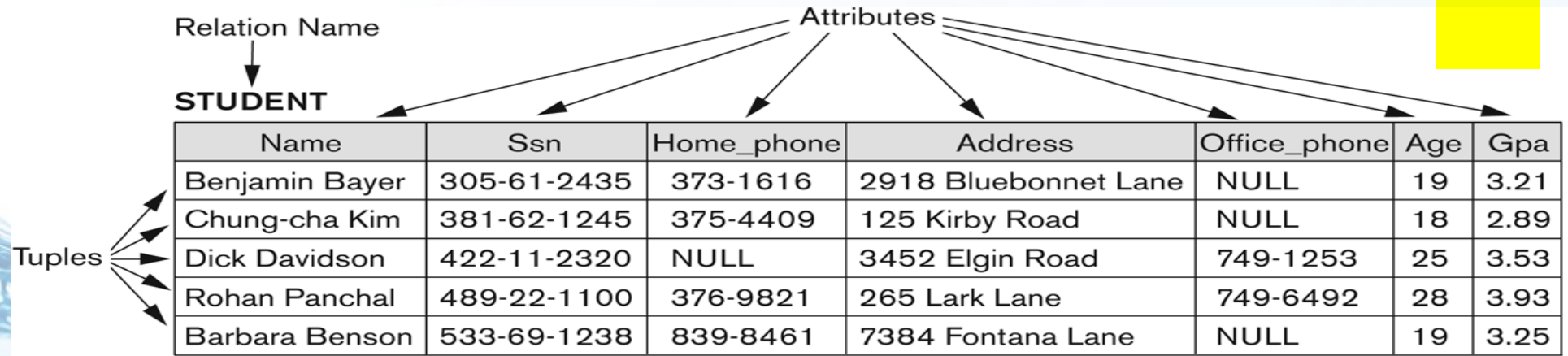
- No two tuples can have the same combination of values for all their attributes.
- **Superkey**
  - No two distinct tuples in any state  $r$  of  $R$  can have same values for a subset of attributes, SK
  - Such a subset SK is the Superkey of  $R$
  - Specifies a ***uniqueness constraint***
  - Every relation has at least one default superkey
  - Superkey can have redundant attributes

# Key Constraints

- **Key**
  - Superkey of  $R$
  - Property: Removing any attribute  $A$  from key,  $K$  leaves a set of attributes  $K'$  that is not a superkey of  $R$  any more
  - No redundancy
  - Determined from the meaning of attributes
  - The property is ***time-invariant***




# Key Constraints



The diagram illustrates a relation named **STUDENT**. Above the table, the label "Relation Name" has an arrow pointing to **STUDENT**. The label "Attributes" has arrows pointing to each of the seven columns: Name, Ssn, Home\_phone, Address, Office\_phone, Age, and Gpa. To the left of the table, the label "Tuples" has arrows pointing to each of the five rows of data. A yellow square is located in the top right corner of the diagram area.

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

- SSN is a key of STUDENT
  - Any set of attributes including SSN is a Superkey
    - {SSN, Name, Age}
  - That superkey is not a key of STUDENT. Why?
  - Any superkey formed from a single attribute is also a key.
  - A key with multiple attributes must have all its attributes together for uniqueness
- 

# Key Constraints

- **Key**
  - Superkey of R
  - No redundancy
  - Determined from the meaning of attributes
  - The property is ***time-invariant***

# Key Constraints

- Key satisfies two properties:
  - Two distinct tuples in any state of relation cannot have identical values for (all) attributes in a key
  - Minimal superkey
    - Cannot remove any attributes and still have uniqueness constraint hold
    - Minimality property is required for a key, though optional for a superkey
- A key is superkey, but may not be the other way around, depending on minimality!



# Key Constraints & Constraints on NULL Values

- **Candidate keys**
  - Relation schema may have more than one key
- **Primary key** of the relation
  - Chosen from the candidate keys
  - A single attribute or a small number of attributes
  - Underline attribute(s)
- Other candidate keys are designated as **unique keys**
- **NOT NULL constraint**

# Relational Databases and Schemas

- **Relational database schema  $S$** 
  - Set of relation schemas  $S = \{R_1, R_2, \dots, R_m\}$
  - Set of integrity constraints  $IC$
- **Relational database state**
  - Set of relation states  $DB = \{r_1, r_2, \dots, r_m\}$
  - Each  $r_i$  is a state of  $R_i$
  - The  $r_i$  relation states satisfy integrity constraints specified in  $IC$

# Relational Databases and Schemas

- **Invalid state**
  - Does not obey all the integrity constraints
- **Valid state**
  - Satisfies all the constraints in the defined set of integrity constraints IC



# Relational Databases and Schemas

## EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
-------	-------	-------	------------	-------	---------	-----	--------	-----------	-----

## DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
-------	----------------	---------	----------------

## DEPT\_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
----------------	------------------

## PROJECT

Pname	<u>Pnumber</u>	Plocation	Dnum
-------	----------------	-----------	------

## WORKS\_ON

<u>Essn</u>	<u>Pno</u>	Hours
-------------	------------	-------

## DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
-------------	-----------------------	-----	-------	--------------

# Relational Databases and Schemas

## EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
-------	-------	-------	------------	-------	---------	-----	--------	-----------	-----

## DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
-------	----------------	---------	----------------

## DEPT\_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
----------------	------------------

## PROJECT

<u>Pname</u>	<u>Pnumber</u>	Plocation	<u>Dnum</u>
--------------	----------------	-----------	-------------

## WORKS\_ON

<u>Essn</u>	<u>Pno</u>	Hours
-------------	------------	-------

## DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
-------------	-----------------------	-----	-------	--------------

# Integrity, Referential Integrity, Foreign Keys

- **Entity integrity constraint**
  - No primary key value can be NULL
  - Used for identification of individual tuples in a relation
- **Ensures ability to distinguish tuples when referenced from other relations**



# Integrity, Referential Integrity, Foreign Keys

- **Referential integrity constraint**
  - Specified between two relations
  - **Maintains consistency among tuples in two relations**
  - A tuple in one relation that refers to another relation must refer to an existing tuple in that relation

# Integrity, Referential Integrity, Foreign Keys

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	<u>Dno</u>
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT\_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS\_ON

<u>Essn</u>	<u>Pno</u>	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

PROJECT

Pname	<u>Pnumber</u>	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

# Integrity, Referential Integrity, Foreign Keys

- **Foreign key rules:**

- A set of attributes, FK in  $R_1$  have the same domain(s) as the primary key attributes PK in  $R_2$
- Value of FK in a tuple  $t_1$  of the current state  $r_1(R_1)$  either occurs as a value of PK for some tuple  $t_2$  in the current state  $r_2(R_2)$  or is NULL
  - Tuple  $t_1$  **references** or **refers to** tuple  $t_2$
  - **R1** – *referencing relation*
  - **R2** – *referenced relation*
- *Referential integrity constraint from R1 to R2 holds*



# Integrity, Referential Integrity, Foreign Keys

- Diagrammatically display referential integrity constraints
  - Directed arc from each foreign key to the relation it references
- All integrity constraints should be specified in the relational database schema

# Integrity, Referential Integrity, Foreign Keys

## EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
-------	-------	-------	------------	-------	---------	-----	--------	-----------	-----

## DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
-------	----------------	---------	----------------

## DEPT\_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
----------------	------------------

## PROJECT

Pname	<u>Pnumber</u>	Plocation	Dnum
-------	----------------	-----------	------

## WORKS\_ON

<u>Essn</u>	<u>Pno</u>	Hours
-------------	------------	-------

## DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
-------------	-----------------------	-----	-------	--------------



# Integrity, Referential Integrity, Foreign Keys

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	<u>Super_ssn</u>	<u>Dno</u>
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	880 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT\_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS\_ON

<u>Essn</u>	<u>Pno</u>	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

PROJECT

Pname	<u>Pnumber</u>	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse



# Other Types of Constraints

- **Semantic integrity constraints**
  - May have to be specified and enforced on a relational database
    - 40 - Maximum number of hours an employee can work per week
    - Salary of an employee should not be more than his/her supervisor's salary
  - Use **triggers** and **assertions**
  - More common to check for these types of constraints in the application programs

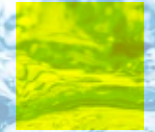
# Other Types of Constraints

- **State constraints**

- Constraints discussed so far... 
- Define the constraints that a valid state of the database must satisfy

# Other Types of Constraints

- **Transition constraints**
  - Defined to deal with state changes in the database
    - Salary of an employee cannot decrease!
  - Enforced by application programs





# Other Types of Constraints

- **Functional dependency constraint**
  - Establishes a functional relationship among two sets of attributes  $X$  and  $Y$
  - Value of  $X$  determines a unique value of  $Y$
  - Enforced using ***Validation*** checks

# Operations on Relations

- INSERT a tuple
- MODIFY a tuple
- DELETE a tuple
- Integrity constraints **should not be violated** by these update operations.
- Several update operations may have to be grouped together.
- Updates may *propagate* to cause other updates automatically - to maintain integrity constraints.

# Possible violations on INSERT

- INSERT may violate any of the constraints:
  - **Domain constraint**
    - if one of the attribute values provided for the new tuple is not of the specified attribute domain
  - **Key constraint**
    - if the value of a key attribute in the new tuple already exists in another tuple in the relation
  - **Referential integrity**
    - if a foreign key value in the new tuple references a primary key value that does not exist in the referenced relation
  - **Entity integrity**
    - if the primary key value is null in the new tuple



# Possible violations on INSERT

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

- Insert <'Cecilia', 'F', 'Kolonsky', NULL, '1960-04-05', '6357 Windy Lane, Katy, TX', F, 28000, NULL, 4> into EMPLOYEE
- Insert <'Alicia', 'J', 'Zelaya', '999887777', '1960-04-05', '6357 Windy Lane, Katy, TX', F, 28000, '987654321', 4> into EMPLOYEE
- Insert <'Cecilia', 'F', 'Kolonsky', '677678989', '1960-04-05', '6357 Windswept, Katy, TX', F, 28000, '987654321', 7> into EMPLOYEE