

# DBMS Languages

## Data Manipulation Language (DML)

Two classes of languages

### Procedural (Low Level )

- User specifies what data is required and how to get those data
- Example: **Relational Algebra**.
  - In RA, we specify the order in which the operations have to be performed.

### Nonprocedural (High Level)

- User specifies what data is required without specifying how to get those data
- Example: **SQL**

SQL can be

used in a standalone way (query language)

embedded in a programming language (host language)



# Relational Algebra

# Relational Algebra

Operations in RDBMS

```
graph TD; A[Operations in RDBMS] --> B[Retrieval]; A --> C[Update];
```

Retrieval

Update

**Relational Algebra** is a set of operations for specifying *retrieval requests (or queries)* in relational model

**Relational algebra expression** is a sequence of relational algebra operations

# Company Database

## 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

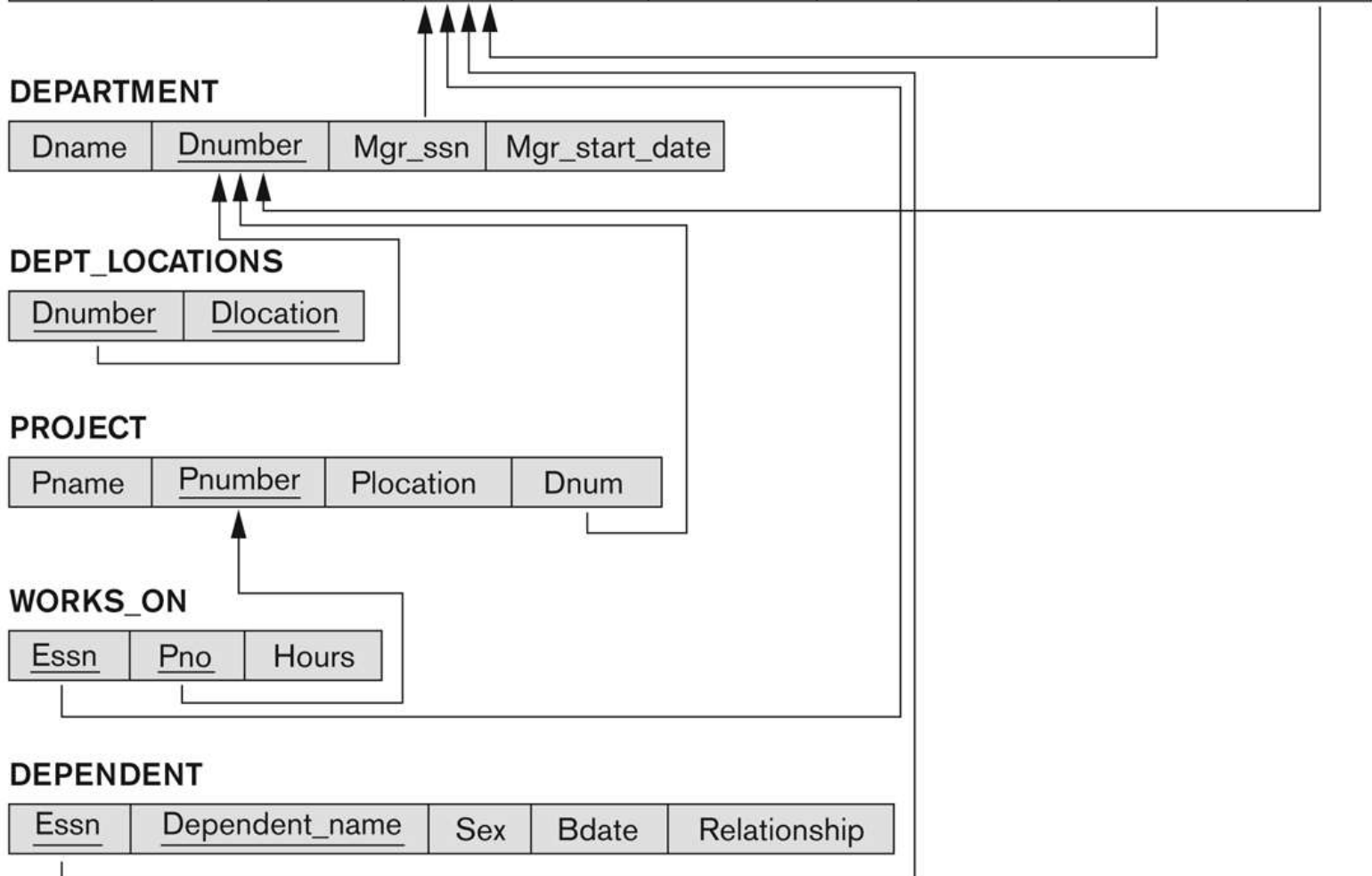
Pname	<u>Pnumber</u>	Plocation	Dnum
-------	----------------	-----------	------

## WORKS\_ON

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

## DEPENDENT

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



# Select Operation(unary operation)

This operation selects a subset of tuples from a relation that satisfy a selection condition.

Select is denoted by :  $\sigma_{\langle \text{selection condition} \rangle}(\mathbf{R})$

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	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

# Examples : Select Operation

- Select the employees whose department number is 4:

$$\sigma_{DNO = 4} (EMPLOYEE)$$

Select the employees whose salary is greater than \$35,000

EMPLOYEE									
Fname	Minit	Lname	<u>Ssn</u>	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

**Print Pnumbers of projects offered by department 5**

**Print the SSN of employees who work more than 10 hours on a project**

**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**

<u>Pname</u>	<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



# Select Operation

$\sigma$  (DNO = 4 AND Salary > 25000) OR (DNO = 5 AND Salary > 30000) (EMPLOYEE)

EMPLOYEE									
Fname	Minit	Lname	<u>Ssn</u>	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

## OUTPUT

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	Dno
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
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



# Select Operation

- **Selection condition** is a Boolean expression specified on the attributes of relation **R**
  - It can include boolean operators **AND, OR, NOT** applied on relational operators **< , > , <= , >= , != , =**

- **Select  $\sigma$  is commutative:**

$$\sigma_{\langle \text{condition1} \rangle}(\sigma_{\langle \text{condition2} \rangle}(\mathbf{R})) = \sigma_{\langle \text{condition2} \rangle}(\sigma_{\langle \text{condition1} \rangle}(\mathbf{R}))$$

- **Cascade of Select operations**

$$\sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond3} \rangle}(\mathbf{R}))) = \sigma_{\langle \text{cond1} \rangle \text{ AND } \langle \text{cond2} \rangle \text{ AND } \langle \text{cond3} \rangle}(\mathbf{R}))$$

**$\sigma$  (DNO = 4 AND Salary > 25000) OR (DNO = 5 AND Salary > 30000) (EMPLOYEE)**

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
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

# Project Operation (unary operation)

- It selects a subset of columns from the relation.
- denoted by  $\pi_{\langle \text{attribute list} \rangle} R$

## Example:

□  $\pi_{\text{LNAME, FNAME, SALARY}} (\text{EMPLOYEE})$

It removes duplicate tuples

The result of project is set of tuples

### OUTPUT

EMPLOYEE				Lname	Fname	Salary	Sex	Salary	Super_ssn	Dno
Fname	Minit	Lname	<u>Ssn</u>	Smith	John	30000				
John	B	Smith	123456789	Wong	Franklin	40000	M	30000	333445555	5
Franklin	T	Wong	333445555	Zelaya	Alicia	25000	M	40000	888665555	5
Alicia	J	Zelaya	999887777	Wallace	Jennifer	43000	F	25000	987654321	4
Jennifer	S	Wallace	987654321	Narayan	Ramesh	38000	F	43000	888665555	4
Ramesh	K	Narayan	666884444	English	Joyce	25000	M	38000	333445555	5
Joyce	A	English	453453453	Jabbar	Ahmad	25000	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	Borg	James	55000	M	25000	987654321	4
James	E	Borg	888665555				M	55000	NULL	1

# Project Operation

## Example 1

□  $\pi$  SALARY ( $\pi$  LNAME, FNAME, SALARY EMPLOYEE)

### OUTPUT

Lname	Fname	Salary
Smith	John	30000
Wong	Franklin	40000
Zelaya	Alicia	25000
Wallace	Jennifer	43000
Narayan	Ramesh	38000
English	Joyce	25000
Jabbar	Ahmad	25000
Borg	James	55000



Salary
30000
40000
25000
43000
38000
25000
25000
55000

## Example 2

□  $\pi$  LNAME, FNAME, SALARY EMPLOYEE)



Project operation is *not* commutative

Salary
30000
40000
25000
43000
38000
55000



NOW  
WHAT  
???

EMPLOYEE									
Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	555	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



# Project Operation

- Project operation is *not* **commutative**
- $\pi_{\langle \text{list1} \rangle} (\pi_{\langle \text{list2} \rangle} (R)) = \pi_{\langle \text{list1} \rangle} (R)$  as long as  $\langle \text{list2} \rangle$  contains the attributes in  $\langle \text{list1} \rangle$

No of Tuples in the result of projection  $\pi_{\langle \text{list} \rangle}(R)$  is less or equal to the number of tuples in  $R$

If the list of attributes includes a *key* of  $R$ , then the no of is *equal* to the no of tuples in  $R$



# Print the name and number of projects offered by department 5

**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**

<u>Pname</u>	<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

# Relational Algebra Expressions

*Retrieve the first name, last name, and salary of all employees who work in department number 5*

Single relational algebra expression:

□  $\pi_{\text{FNAME, LNAME, SALARY}}(\sigma_{\text{DNO}=5}(\text{EMPLOYEE}))$

Using intermediate relation:

- $\text{D5} \leftarrow \sigma_{\text{DNO}=5}(\text{EMPLOYEE})$
- $\text{RESULT} \leftarrow \pi_{\text{FNAME, LNAME, SALARY}}(\text{D5})$





# Example of applying multiple operations and RENAME

$\pi_{\text{FNAME, LNAME, SALARY}}(\sigma_{\text{DNO}=5}(\text{EMPLOYEE}))$

Fname	Lname	Salary
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000
Joyce	English	25000

$D5 \leftarrow \sigma_{\text{DNO}=5}(\text{EMPLOYEE})$

**R** (First\_name, Last\_name, Salary)  $\leftarrow \pi_{\text{Fname, Lname, Salary}} D5$

D5

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000		5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	F	40000		5
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000		5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000		5

R

First_name	Last_name	Salary
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000
Joyce	English	25000

# Union (Binary Operation)

- The result of  $R \cup S$ , is a relation that includes all tuples that are either in R or in S or in both R and S
- **Duplicate tuples are eliminated**

## Type compatible (Union compatible)

- The two relations R and S must be **Type compatible**
  - R and S must have same number of attributes
  - Each pair of corresponding attributes must have same or compatible domains

Fname	Lname	Salary
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000
Joyce	English	25000

U

Fname	Lname	Salary
John	Smith	30000
Franklin	Wong	40000



# UNION Example

To retrieve the SSN of all employees who either

- work in department 5 or
- directly supervise an employee in department 5

$D5\_EMPS \leftarrow \sigma_{DNO=5} (EMPLOYEE)$

$RESULT1 \leftarrow \pi_{SSN}(D5\_EMPS)$

$RESULT2(SSN) \leftarrow \pi_{SUPERSSN}(D5\_EMPS)$

$RESULT \leftarrow RESULT1 \cup RESULT2$

RESULT1

Ssn
123456789
333445555
666884444
453453453

RESULT2

Ssn
333445555
888665555

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Superssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000		5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000		5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000		4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000		4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000		5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000		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

RESULT

Ssn
123456789
333445555
666884444
453453453
888665555

# INTERSECTION And SET DIFFERENCE (Binary Operations)

- INTERSECTION operation: the result of  $R \cap S$ , is a relation that includes all tuples that are in both R and S
- SET DIFFERENCE operation: the result of  $R - S$ , is a relation that includes all tuples that are in R but not in S
- Two relations R and S must be “type compatible”



# RELATIONAL ALGEBRA OPERATIONS FROM SET THEORY

- Both  $\cup$  and  $\cap$  are *commutative* operations
  - $R \cup S = S \cup R$ , and  $R \cap S = S \cap R$
- Both  $\cup$  and  $\cap$  can be treated as n-ary operations
  - $R \cup (S \cup T) = (R \cup S) \cup T$
  - $(R \cap S) \cap T = R \cap (S \cap T)$
- Minus operation is not commutative
  - $R - S \neq S - R$

# Example

**STUDENT**

Fn	Ln
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

**INSTRUCTOR**

Fname	Lname
John	Smith
Ricardo	Browne
Susan	Yao
Francis	Johnson
Ramesh	Shah

**Compatible relation**

Fn	Ln
Susan	Yao
Ramesh	Shah
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert
John	Smith
Ricardo	Browne
Francis	Johnson

**Student  $\cup$  Instructor**

Fn	Ln
Susan	Yao
Ramesh	Shah

**Student  $\cap$  Instructor**

Fn	Ln
Johnny	Kohler
Barbara	Jones
Amy	Ford
Jimmy	Wang
Ernest	Gilbert

**Student – Instructor**

Fname	Lname
John	Smith
Ricardo	Browne
Francis	Johnson

**Instructor – Student**



# Example: Retrieve a list of female employee's dependents

**EMPLOYEE**

Fname	Minit	Lname	<u>Ssn</u>	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

**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

# CARTESIAN PRODUCT

- The result of Cartesian product of two relations

$$R(A_1, A_2, \dots, A_n) \times S(B_1, B_2, \dots, B_m)$$

is given as:

$$\text{Result}(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$$

- Let  $|R| = n_R$  and  $|S| = n_S$ , then  $|R \times S| = n_R * n_S$
- R and S may NOT be "type compatible"

**Cross Product is a meaningful operation only if it is followed by other operations**

## Example: Retrieve a list of female employee's dependents

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	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

$F \leftarrow \sigma_{\text{SEX}='F'}(\text{EMPLOYEE})$

# Example: Retrieve a list of female employee's dependents

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								1

$$F \leftarrow \sigma_{\text{SEX}='F'}(\text{EMPLOYEE})$$
$$\text{EmpNames} \leftarrow \pi_{\text{FNAME, LNAME, SSN}}(F)$$

F  
FEMALE\_EMPS

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
Alicia	J	Zelaya	999887777	1968-07-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5

EMPNames

Fname	Lname	Ssn
Alicia	Zelaya	999887777
Jennifer	Wallace	987654321
Joyce	English	453453453



# Problem: Retrieve a list of female employee's dependents

EMPNames

Fname	Lname	Ssn
Alicia	Zelaya	999887777
Jennifer	Wallace	987654321
Joyce	English	453453453

DEPENDENT

Essn	Dependent_name	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

EMP\_DEPENDENTS

Fname	Lname	Ssn	Essn	Dependent_name	Sex	Bdate	...
Alicia	Zelaya	999887777	333445555	Alice	F	1986-04-05	...
Alicia	Zelaya	999887777	333445555	Theodore	M	1983-10-25	...
Alicia	Zelaya	999887777	333445555	Joy	F	1958-05-03	...
Alicia	Zelaya	999887777	987654321				
Alicia	Zelaya	999887777	123456789				
Alicia	Zelaya	999887777	123456789	Alice	F	1988-12-30	...
Alicia	Zelaya	999887777	123456789	Elizabeth	F	1967-05-05	...
Jennifer	Wallace	987654321	333445555	Alice	F	1986-04-05	...
Jennifer	Wallace	987654321	333445555	Theodore	M	1983-10-25	...
Jennifer	Wallace	987654321	333445555	Joy	F	1958-05-03	...
Jennifer	Wallace	987654321	987654321	Abner	M	1942-02-28	...
Jennifer	Wallace	987654321	123456789	Michael	M	1988-01-04	...
Jennifer	Wallace	987654321	123456789	Alice	F	1988-12-30	...
Jennifer	Wallace	987654321	123456789	Elizabeth	F	1967-05-05	...
Joyce	English	453453453	333445555	Alice	F	1986-04-05	...
Joyce	English	453453453	333445555	Theodore	M	1983-10-25	...
Joyce	English	453453453	333445555	Joy	F	1958-05-03	...
Joyce	English	453453453	987654321	Abner	M	1942-02-28	...
Joyce	English	453453453	123456789	Michael	M	1988-01-04	...
Joyce	English	453453453	123456789	Alice	F	1988-12-30	...
Joyce	English	453453453	123456789	Elizabeth	F	1967-05-05	...

**Emp\_DP ← EmpNames x DEPENDENT**

# Problem: Retrieve a list of female employee's dependents

**EMPNames**

Fname	Lname	Ssn
Alicia	Zelaya	999887777
Jennifer	Wallace	987654321
Joyce	English	453453453

**DEPENDENT**

Essn	Dependent_name	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

**EMP\_DEPENDENTS**

Fname	Lname	Ssn	Essn	Dependent_name	Sex	Bdate	...
Alicia	Zelaya	999887777	333445555	Alice	F	1986-04-05	...
Alicia	Zelaya	999887777	333445555	Theodore	M	1983-10-25	...
Alicia	Zelaya	999887777	333445555	Joy	F	1958-05-03	...
Alicia	Zelaya	999887777	987654321				
Alicia	Zelaya	999887777	123456789				
Alicia	Zelaya	999887777	123456789				
Alicia	Zelaya	999887777	123456789				
Jennifer	Wallace	987654321	333445555				
Jennifer	Wallace	987654321	333445555				
Jennifer	Wallace	987654321	333445555				
Jennifer	Wallace	987654321	987654321	Abner	M	1942-02-28	...
Jennifer	Wallace	987654321	123456789	Michael	M	1988-01-04	...
Jennifer	Wallace	987654321	123456789	Alice	F	1988-12-30	...
Jennifer	Wallace	987654321	123456789	Elizabeth	F	1967-05-05	...

**Emp\_DP  $\leftarrow$  EmpNames x DEPENDENT**

**Actual\_DP  $\leftarrow \sigma_{SSN=ESSN}(\text{Emp\_DP})$**

**ACTUAL\_DEPENDENTS**

Fname	Lname	Ssn	Essn	Dependent_name	Sex	Bdate	...
Jennifer	Wallace	987654321	987654321	Abner	M	1942-02-28	...

Joyce	English	453453453					
Joyce	English	453453453					
Joyce	English	453453453					
Joyce	English	453453453					
Joyce	English	453453453	123456789	Michael	M	1988-01-04	...
Joyce	English	453453453	123456789	Alice	F	1988-12-30	...
Joyce	English	453453453	123456789	Elizabeth	F	1967-05-05	...



# Problem: Retrieve a list of each female employee's dependents

F

**FEMALE\_EMPS**

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
Alicia	J	Zelaya	999887777	1968-07-19	3321Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291Berry, Bellaire, TX	F	43000	888665555	4
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5

**EMPNames**

Fname	Lname	Ssn
Alicia	Zelaya	999887777
Jennifer	Wallace	987654321

**DEPENDENT**

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son

**EMP\_DEPENDENTS**

Fname	Lname	Ssn	Essn	Dependent_name	Sex	Bdate	Relationship
Alicia	Zelaya	999887777	333445555	Alice	F	1986-04-05	Daughter
Alicia	Zelaya	999887777	333445555	Theodore	M	1983-10-25	Son
Alicia	Zelaya	999887777	333445555	Joyce	F	1972-07-31	...
Alicia	Zelaya	999887777	987654321	Jennifer	F	1941-06-20	...
Alicia	Zelaya	999887777	123456789	Michael	M	1988-01-09	...
Alicia	Zelaya	999887777	123456789	Alice	F	1988-01-09	...
Alicia	Zelaya	999887777	123456789	Elizabeth	F	1967-05-09	...
Jennifer	Wallace	987654321	333445555	Alice	F	1986-04-05	...
Jennifer	Wallace	987654321	333445555	Theodore	M	1983-10-25	...
Jennifer	Wallace	987654321	987654321	Abner	M	1942-02-28	...
Jennifer	Wallace	987654321	123456789	Michael	M	1988-01-09	...
Jennifer	Wallace	987654321	123456789	Alice	F	1988-01-09	...
Jennifer	Wallace	987654321	123456789	Elizabeth	F	1967-05-09	...
Joyce	English	453453453	333445555	Alice	F	1986-04-05	...
Joyce	English	453453453	333445555	Theodore	M	1983-10-25	...
Joyce	English	453453453	333445555	Joy	F	1958-01-05	...
Joyce	English	453453453	987654321	Abner	M	1942-02-28	...
Joyce	English	453453453	123456789	Michael	M	1988-01-09	...
Joyce	English	453453453	123456789	Alice	F	1988-01-09	...
Joyce	English	453453453	123456789	Elizabeth	F	1967-05-09	...

$F \leftarrow \sigma_{SEX='F'}(EMPLOYEE)$

$EmpNames \leftarrow \pi_{FNAME, LNAME, SSN}(F)$

$Emp\_DP \leftarrow EmpNames \times DEPENDENT$

$Actual\_DP \leftarrow \sigma_{SSN=ESSN}(Emp\_DP)$

$Result \leftarrow \pi_{FNAME, LNAME, DEPENDENT\_NAME}(Actual\_DP)$

**ACTUAL\_DEPENDENTS**

Fname	Lname	Ssn	Essn	Dependent_name	Sex	Bdate	...
Jennifer	Wallace	987654321	987654321	Abner	M	1942-02-28	...

**RESULT**

Fname	Lname	Dependent_name
Jennifer	Wallace	Abner

# JOIN(Binary Operation)

- JOIN denoted by  $\bowtie$  *combines related tuples* from various relations
- JOIN combines CARTESIAN PRODECT and SELECT into a single operation
- General form of a join operation on two relations  $R(A_1, A_2, \dots, A_n)$  and  $S(B_1, B_2, \dots, B_m)$  is:

$$R \bowtie_{\langle \text{join condition} \rangle} S$$

# Problem: Retrieve a list of each female employee's dependents

F

FEMALE\_EMPS

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
Alicia	J	Zelaya	999887777	1968-07-19	3321Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291Berry, Bellaire, TX	F	43000	888665555	4
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5

EMPNames

Fname	Lname	Ssn
Alicia	Zelaya	999887777
Jennifer	Wallace	987654321

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
		M	1942-02-28	Spouse
		M	1988-01-04	Son

EMP\_DEPENDENTS

Fname	Lname	Ssn	Essn	Dependent_name	Sex	Bdate	...
Alicia	Zelaya	999887777					
Alicia	Zelaya	999887777					
Alicia	Zelaya	999887777					
Alicia	Zelaya	999887777					
Alicia	Zelaya	999887777					
Alicia	Zelaya	999887777					
Alicia	Zelaya	999887777					
Jennifer	Wallace	987654321					
Jennifer	Wallace	987654321					
Jennifer	Wallace	987654321					
Jennifer	Wallace	987654321					
Jennifer	Wallace	987654321					
Jennifer	Wallace	987654321					
Jennifer	Wallace	987654321					
Joyce	English	453453453	333445555	Alice	F	1986-04-05	...
Joyce	English	453453453	333445555	Theodore	M	1983-	
Joyce	English	453453453	333445555	Joy	F	1958-	
Joyce	English	453453453	987654321	Abner	M	1942-	
Joyce	English	453453453	123456789	Michael	M	1988-	
Joyce	English	453453453	123456789	Alice	F	1988-	
Joyce	English	453453453	123456789	Elizabeth	F	1967-	

$F \leftarrow \sigma_{SEX='F'}(EMPLOYEE)$

$EmpNames \leftarrow \pi_{FNAME, LNAME, SSN}(F)$

$Emp\_DP \leftarrow EmpNames \bowtie_{SSN=ESSN} DEPENDENT$

$Result \leftarrow \pi_{FNAME, LNAME, DEPENDENT\_NAME}(DP)$

RESULT

Fname	Lname	Dependent_name
Jennifer	Wallace	Abner

# Example: JOIN operation

Retrieve the name of the manager of each department

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	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						
Joyce	A	English	453453453						
Ahmad	V	Jabbar	987987987						
James	E	Borg	888665555						

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\_MGR  $\leftarrow$  DEPARTMENT  $\bowtie_{\text{MGRSSN=SSN}}$  EMPLOYEE

DEPT\_MGR

Dname	Dnumber	Mgr_ssn	...	Fname	Minit	Lname	Ssn	...
Research	5	333445555	...	Franklin	T	Wong	333445555	...
Administration	4	987654321	...	Jennifer	S	Wallace	987654321	...
Headquarters	1	888665555	...	James	E	Borg	888665555	...

# Complete Set of Relational Operations

- The set of operations including
  - SELECT  $\sigma$ ,
  - PROJECT  $\pi$ ,
  - UNION  $\cup$ ,
  - DIFFERENCE  $-$ ,
  - RENAME  $\rho$ , and
  - CARTESIAN PRODUCT  $\times$

is called a *complete set* because any relational algebra expression can be expressed using these.

- $R \cap S = (R \cup S) - ((R - S) \cup (S - R))$
- $R \bowtie_{\langle \text{join condition} \rangle} S = \sigma_{\langle \text{join condition} \rangle} (R \times S)$



# Some properties of JOIN

- Consider the following JOIN operation:

- $$R(A_1, A_2, \dots, A_n) \quad \bowtie \quad S(B_1, B_2, \dots, B_m)$$
$$R.A_i = S.B_j$$

- Result is a relation Q with degree **n + m attributes**:
  - Q(A<sub>1</sub>, A<sub>2</sub>, . . . , A<sub>n</sub>, B<sub>1</sub>, B<sub>2</sub>, . . . , B<sub>m</sub>), in that order.
- If R has n<sub>R</sub> tuples, and S has n<sub>S</sub> tuples, then no of tuples in **join result** < n<sub>R</sub> \* n<sub>S</sub> .



# Equi-Join

- EQUIJOIN is a join condition that involves only equality operator = .
- **Example:**
  - **Retrieve a list of each female employee's dependents**

$\text{FEmp} \leftarrow \sigma_{\text{SEX}='F'}(\text{EMPLOYEE})$

$\text{E\_DP} \leftarrow \text{FEmp} \bowtie_{\text{SSN}=\text{ESSN}} \text{DEPENDENT}$

$\text{Result} \leftarrow \pi_{\text{FNAME}, \text{LNAME}, \text{SSN}, \text{DEPENDENT\_NAME}}(\text{E\_DP})$

# This is EQUI-JOIN operation

Retrieve the name of the manager of each department

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	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						
Joyce	A	English	453453453						
Ahmad	V	Jabbar	987987987						
James	E	Borg	888665555						

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\_MGR  $\leftarrow$  DEPARTMENT  $\bowtie_{\text{MGRSSN=SSN}}$  EMPLOYEE

Research	5	333445555	...	Franklin	T	Wong	333445555	...
Administration	4	987654321	...	Jennifer	S	Wallace	987654321	...
Headquarters	1	888665555	...	James	E	Borg	888665555	...

# For each employee, print his project numbers

## 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

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

- For each employee, list the name of his projects

#### 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

#### 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

#### EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	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

- For each employee, retrieve the employee's name and the name of his project

**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

**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

<b>EMPLOYEE</b>										
Fname	Minit	Lname	<u>Ssn</u>	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	



# NATURAL JOIN Operation

- **Example:** Print location of each department

DEPT\_LOCS  $\leftarrow$  DEPARTMENT \* DEPT\_LOCATIONS

- Only attribute with the same name is **DNUMBER**

An implicit join condition is created based on this attribute:

DEPARTMENT.DNUMBER=DEPT\_LOCATIONS.DNUMBER

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



# NATURAL JOIN Operation

- Example: Print location of each department

DEPT\_LOCS  $\leftarrow$  DEPARTMENT \* DEPT\_LOCATIONS

- DEPT\_LOCS

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

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

# Issue with Equijoin Operation

- You have to specify the **join condition**.
  - Even if two cols in the joining tables have same name.

DEPT\_MGR

Dname	Dnumber	Mgr_ssn	...	Fname	Minit	Lname	Ssn	...
Research	5	333445555	...	Franklin	T	Wong	333445555	...
Administration	4	987654321	...	Jennifer	S	Wallace	987654321	...
Headquarters	1	888665555	...	James	E	Borg	888665555	...

- Superfluous column
- Result of EQUIJOIN always have one or more pairs of attributes that have identical values in every tuple.

# NATURAL JOIN Operation

NATURAL JOIN operation (denoted by  $*$ ) is used when

- the two join attributes, or
- each pair of corresponding join attributes

must *have the same name* in both relations

- If this is not the case, a **renaming operation** is applied first.

- NATURAL JOIN also get rid of the superfluous attribute in an EQUIJOIN condition.

# Example: Natural Join

- Consider two Relations
  - $R1(A,B,C,D)$  &  $R2(C,D,E)$
- Natural Join  $R * S$ 
  - $RES \leftarrow R1(A,B,C,D) * R2(C,D,E)$
  - The implicit join condition
    - $R1.C = R2.C$  AND  $R1.D = R2.D$

A	B	C	D

R1

C	D	E

R2

**RES(A,B,C,D,E)**

A	B	C	D	E

# Theta-join

- The general case of JOIN operation is called a Theta-join:  $R \bowtie_{\theta} S$   
*theta*
- *Theta* is a boolean expression on the attributes of R and S; for example:
  - $R.A_i < S.B_j$  AND  $(R.A_k = S.B_l$  OR  $R.A_p < S.B_q)$
- Theta can have any comparison operators  $\{=, \neq, <, \leq, >, \geq, \}$

# Theta-join Example

For each Male employee, list his colleagues who earn more than him. Retrieve only the first name and salary.

$$M(\text{Name}, \text{Sal}) \leftarrow \pi_{\text{FNAME}, \text{SALARY}} (\sigma_{\text{SEX}='M'} \text{EMPLOYEE})$$
$$\text{ECOL}(\text{CName}, \text{CSal}) \leftarrow \pi_{\text{FNAME}, \text{SALARY}} \text{EMPLOYEE}$$
$$R1 \leftarrow M \bowtie_{M.\text{Sal} < \text{ECol}.\text{CSal}} \text{ECol}$$

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	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
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1



# Theta-join Example

For each Male employee list his colleagues who earn more than him. Retrieve only the first name and salary.

$M(\text{Name}, \text{Sal}) \leftarrow \pi_{\text{FNAME}, \text{SALARY}} (\sigma_{\text{SEX}='M'} \text{EMPLOYEE})$   
 $\text{ECOL}(\text{CName}, \text{CSal}) \leftarrow \pi_{\text{FNAME}, \text{SALARY}} \text{EMPLOYEE}$   
 $R1 \leftarrow M \bowtie_{M.\text{Sal} < \text{ECol}.\text{CSal}} \text{ECol}$

Name	Sal	CName	CSal
John	30000	Franklin	40000
John	30000	Jennifer	43000
John	30000	Ramesh	38000
John	30000	James	55000
Franklin	40000	Jennifer	43000
Franklin	40000	James	55000
Ramesh	38000	Franklin	40000
Ramesh	38000	Jennifer	43000
Ramesh	38000	James	55000

# Theta-join

- For each Male employee, print the names of his peers with the same salary

$\rho_{E2}(\text{EMPLOYEE})$

$E2 \leftarrow \pi_{\text{FNAME}, \text{SALARY}} (\sigma_{\text{SEX}='M'} \text{EMPLOYEE})$

$\text{Res} \leftarrow \pi_{E1.\text{FNAME}, E2.\text{FNAME}} (E1 \bowtie_{E1.\text{SSN} \neq E2.\text{SSN} \text{ and } E1.\text{Salary}=E2.\text{Salary}} E2)$

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	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