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)

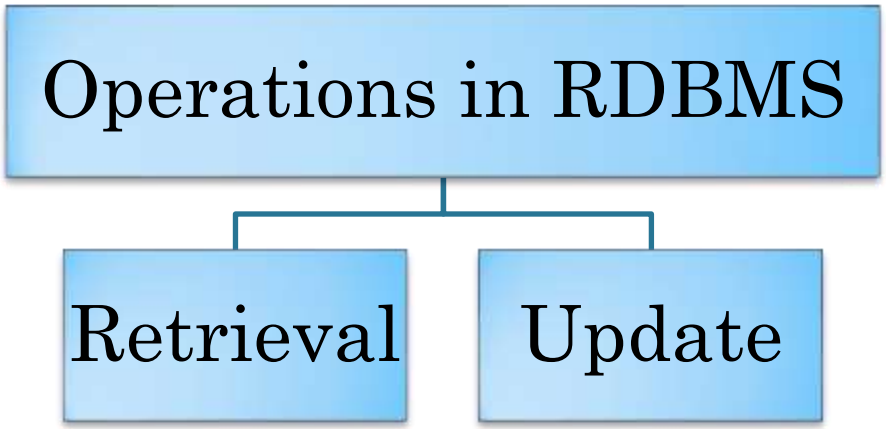
- User specifies what data is required without specifying how to get those data
- Example: **SQL**

SQL can be

- used in a standalone way (query language)
- embedded in a programming language (host language)

Relational Algebra

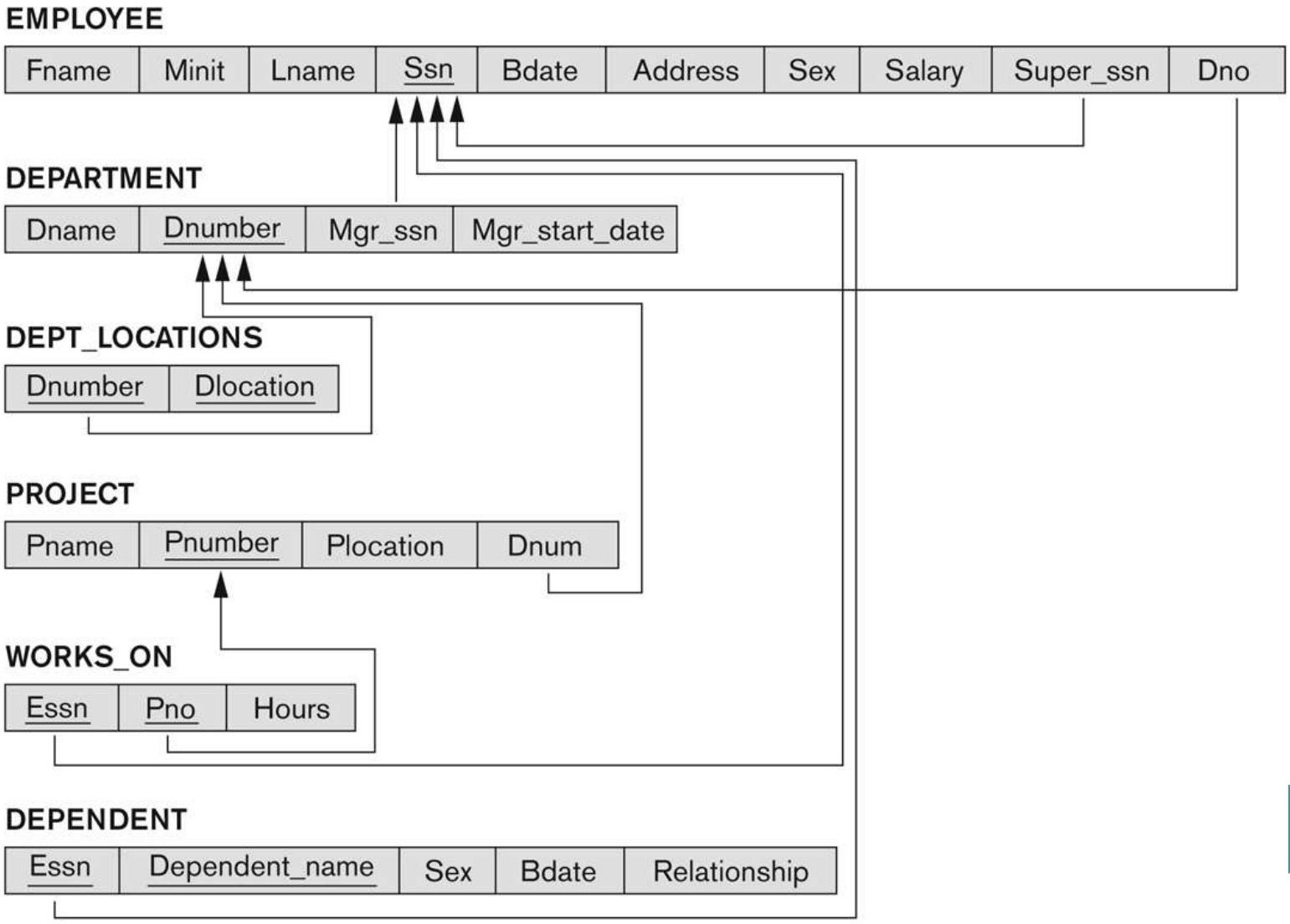
Relational Algebra



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



Select Operation(unary operation)

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

Select is denoted by : $\sigma_{\text{<selection condition>}}(R)$

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

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	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

List projects offered by department 5

List employees who work more than 10 hours on a project

WORKS_ON

Essn	Pno	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	Pnumber	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

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

Select Operation

$\sigma_{(DNO = 4\ AND\ Salary > 25000)\ OR\ (DNO = 5\ AND\ Salary > 30000)}(EMPLOYEE)$

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

OUTPUT

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

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 σ is commutative:**

  $\sigma_{<condition1>}(\sigma_{<condition2>}(R)) = \sigma_{<condition2>}(\sigma_{<condition1>}(R))$

- **Cascade of Select operations**

  $\sigma_{<cond1>}(\sigma_{<cond2>}(\sigma_{<cond3>}(R))) = \sigma_{<cond1> \text{ AND } <cond2> \text{ AND } <cond3>}(R)))$

$\sigma_{(DNO = 4 \text{ AND } Salary > 25000) \text{ OR } (DNO = 5 \text{ AND } Salary > 30000)}$ (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 $\pi_{<\text{attribute list}>}R$

Example:

- π LNAME, FNAME, SALARY (EMPLOYEE)

It removes **duplicate tuples**

The result of project is set of tuples

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

EMPLOYEE

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

Project Operation

Example 1

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

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
55000

Example 2

- $\pi_{\text{LNAME, FNAME, SAL}\ldots\text{ARY}}\text{EMPLOYEE})$

Salary
30000
40000
25000
43000
38000
55000

Project operation is *not* commutative

NOW WHAT ???

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Sa…	…ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30…	…555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40…	…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**

- **$\pi_{\text{<list1>}}(\pi_{\text{<list2>}}(R)) = \pi_{\text{<list1>}}(R)$ as long as <list2> contains the attributes in <list1>**

No of Tuples in the result of projection $\pi_{\text{<list>}}(R)$ 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

Essn	Pno	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	Pnumber	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

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

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:

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

Example of applying multiple operations and **RENAME**

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

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

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

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

D5

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

R

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

Print the details of the manager of each department

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					321	4
James	E	Borg	888665555						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

$D\_M \leftarrow DEPARTMENT \times EMPLOYEE$

$DEPT\_MGR \leftarrow \sigma_{MGRSSN=SSN}\ (D\_M)$

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	...

CARTESIAN PRODUCT

- The result of Cartesian product of two relations

  R(A1, A2, . . ., An) x S(B1, B2, . . ., Bm)

  is given as:

  Result(A1, A2, . . ., An, B1, B2, . . ., Bm)

- 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: JOIN operation

Retrieve the name of the manager of each department

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						
Joyce	A	English	453453453						
Ahmad	V	Jabbar	987987987						
James	E	Borg	888665555						

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

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

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	...

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	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

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

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	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

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

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

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

F

FEMALE_EMPS

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
Alicia	J	Zelaya	999887777	1968-07-19	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
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						
Joyce	English	4					
Joyce	English	4					
Joyce	English	4					
Joyce	English	4					
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

Actual_DP ← $\sigma_{SSN=ESSN}$(**Emp_DP**)

ACTUAL_DEPENDENTS

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

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	98765432

DEPENDENT

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

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

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

$Emp\_DP \leftarrow EmpNames \times DEPENDENT$

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

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

EMP_DEPENDENTS

Fname	Lname	Ssn	Essn				
Alicia	Zelaya	999887777	333445555				
Alicia	Zelaya	999887777	333445555				
Alicia	Zelaya	999887777	333445555				
Alicia	Zelaya	999887777	987654321				
Alicia	Zelaya	999887777	123456789				
Alicia	Zelaya	999887777	123456789				
Alicia	Zelaya	999887777	123456789				
Jennifer	Wallace	987654321	333445555				
Jennifer	Wallace	987654321	333445555	Theodore	M	1983-10-25	...
Jennifer	Wallace	9876543					
Jennifer	Wallace	9876543					
Jennifer	Wallace	9876543					
Jennifer	Wallace	9876543					
Jennifer	Wallace	9876543					
Joyce	English	4534534					
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-	

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 ⋈ *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(A1, A2, . . ., An) and S(B1, B2, . . ., Bm) is:

$$R \bowtie_{\text{<join condition>}} S$$

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-	

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

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

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

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

RESULT

Fname	Lname	Dependent_name
Jennifer	Wallace	Abner

Some properties of JOIN

Consider the following JOIN operation:

- **R(A1, A2, . . ., An)** $\bowtie_{R.Ai=S.Bj}$ **S(B1, B2, . . ., Bm)**

- Result is a **relation Q** with degree n + m attributes:
  Q(A1, A2, . . ., An, B1, B2, . . ., Bm), in this order.

- If R has n_R tuples, and S has n_S tuples, then no of tuples in join result $< n_R * n_S$.

Equi-Join

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

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

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

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

This is EQUI -JOIN operation

Retrieve the name of the manager of each department

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						
Joyce	A	English	453453453						
Ahmad	V	Jabbar	987987987						
James	E	Borg	888665555						

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

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

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

Essn	Pno	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		
999887777		
999887777		
987987987		
987987987		
987654321		
987654321		
888665555		

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

- **For each employee, list the name of his projects**

WORKS_ON

Essn	Pno	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		
999887777		
999887777		
987987987		
987987987		
987654321		
987654321		
888665555		

PROJECT

Pname	Pnumber	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	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

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

WORKS_ON

Essn	Pno	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		
999887777		
999887777		
987987987		
987987987		
987654321		
987654321		
888665555		

PROJECT

Pname	Pnumber	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	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

NATURAL JOIN Operation

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

  DEPT_LOCS ← 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	Dnumber	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

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

NATURAL JOIN Operation

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

  DEPT_LOCS ← DEPARTMENT * DEPT_LOCATIONS

- A

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	Dnumber	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

Dnumber	Dlocation
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

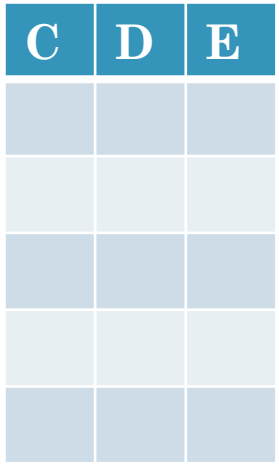
- If this is not the case, a renaming operation is applied first.

- 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)**

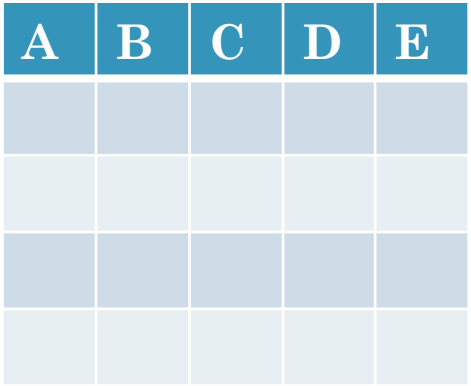
R1



R2

- Natural Join R*S
  - **RES ← R1(A,B,C,D) * R2(C,D,E)**
  - The implicit join condition
    - **R1.C=R2.C AND R1.D=R2.D**

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



Theta-join

- The general case of JOIN operation is called a **Theta-join:** $R \bowtie_{theta} S$
- ***Theta*** is a boolean expression on the attributes of R and S; for example:
  - R.Ai<S.Bj AND (R.Ak=S.Bl OR R.Ap<S.Bq)
- Theta-join 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.

$M(Name, Sal) \leftarrow \pi_{\mathbf{FNAME,\ SALARY}} (\sigma_{SEX='M'} EMPLOYEE)$

$ECOL(CName, CSal) \leftarrow \pi_{\mathbf{FNAME,\ SALARY}} EMPLOYEE$

$R1 \leftarrow M \bowtie_{M.Sal < ECol.CSal} ECol$

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
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.

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

$ECOL(CName, CSal) \leftarrow \pi_{FNAME,\ SALARY}\ EMPLOYEE$

$R1 \leftarrow M \bowtie_{M.Sal < ECol.CSal} ECol$

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

ρ_{E2}(EMPLOYEE)

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

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

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

Union (Binary Operation)

- The result of R ∪ 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

∪

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	Ssn	Bdate	Address	Sex	Sala		Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	300		5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	400		5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	250		4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	430		4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	380		5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	250		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 ∪ Instructor

Fn	Ln
Susan	Yao
Ramesh	Shah

Student ∩ 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

Examples of Queries in RA

Retrieve the name and address of all employees who work for the 'Research' department.

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

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

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
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

$$RESEARCH\_DEPT \leftarrow \sigma_{DNAME='Research'}(DEPARTMENT)$$

$$RESEARCH\_EMPS \leftarrow (RESEARCH\_DEPT \bowtie_{DNUMBER=DNO} EMPLOYEE)$$

$$RESULT \leftarrow \pi_{FNAME, LNAME, ADDRESS}(RESEARCH\_EMPS)$$

EXAMPLE: Retrieve the names of employees who have no dependents.

ALL_EMPS ← π SSN(EMPLOYEE)

EMPS_WITH_DEPS(SSN) ← π ESSN(DEPENDENT)

EMPS_WITHOUT_DEPS ← (ALL_EMPS - EMPS_WITH_DEPS)

RESULT ← π LNAME, FNAME (EMPS_WITHOUT_DEPS * EMPLOYEE)

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

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

RENAME OPEARATION

Rename operator is denoted by ρ (rho)

$D5 \leftarrow \pi_{Fname, Lname, Salary} (\sigma_{DNO=5}(EMPLOYEE))$

$\rho_{R(First\_Name, Last\_Name, Salary)}(D5)$

R

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

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

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

RENAME OPEARATION

$\rho_S(R)$ rename the *relation R* to S

R

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

S

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

$\rho_{(B1, B2, \ldots, Bn)}(R)$ rename the *attributes* to B1, B2, …..Bn

R

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

R

F_	L	Salary
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000

$\rho_{S\ (B1, B2, \ldots, Bn)}(R)$ rename R to S & attributes to B1, …..Bn

R

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

S

F_	L	Salary
John	Smith	30000
Franklin	Wong	40000
Ramesh	Narayan	38000

Aggregate Functions

- Mathematical Aggregate Functions applied to collections of numeric values include
  - SUM, AVERAGE, MAXIMUM, and MINIMUM.
  - COUNT function is used for counting tuples or values.

Examples:

Retrieve the average or total salary of all employees

Retrieve total number of employee tuples

Aggregate Functions $\mathcal{F}$

- $\mathcal{F}_{\text{MAX Salary}}$ (EMPLOYEE)
- $\mathcal{F}_{\text{MIN Salary}}$ (EMPLOYEE)
- $\mathcal{F}_{\text{SUM Salary, AVERAGE Salary}}$ (EMPLOYEE)
- $\mathcal{F}_{\text{COUNT SSN}}$ (EMPLOYEE)

COUNT (*) returns the no. of rows in the result of the query ***(it counts without removing duplicates)***

NULL values are **discarded** when aggregate functions are applied to a particular column (attribute).

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

Using Grouping with Aggregation

- Grouping can be combined with Aggregate Functions
- **Example:**
  - For each department, retrieve the DNO, COUNT of employees and AVERAGE SALARY
  - **DNO** $\mathcal{F}$**COUNT SSN, AVERAGE Salary** EMPLOYEE

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

Grouping with Aggregation

DNO $\mathcal{F}$ **COUNT SSN, AVERAGE Salary** EMPLOYEE

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

Dno	Count_ssn	Average_salary
5	4	33250
4	3	31000
1	1	55000

Grouping with Aggregation

DNO $\mathcal{F}$ **COUNT SSN, AVERAGE Salary** EMPLOYEE

Dno	Count_ssn	Average_salary
5	4	33250
4	3	31000
1	1	55000

$\mathcal{F}$ **COUNT SSN, AVERAGE Salary** EMPLOYEE

Count_ssn	Average_salary
8	35125

ρ**R**(Dno, No_of_employees, Average_sal) (**DNO** $\mathcal{F}$ **COUNT SSN, AVERAGE Salary** EMPLOYEE)

R

Dno	No_of_employees	Average_sal
5	4	33250
4	3	31000
1	1	55000

EXAMPLE: RETRIEVE THE NAMES OF ALL EMPLOYEES WITH TWO OR MORE DEPENDENTS.

$$T1(Ssn, No\_of\_dependents) \leftarrow Essn\ \mathcal{F}_{COUNT\ Dependent\_name}(DEPENDENT)$$

$$T2 \leftarrow \sigma_{No\_of\_dependents > 1}(T1)$$

$$RESULT \leftarrow \pi_{LNAME, FNAME}(T2 * EMPLOYEE)$$

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

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

Example

- List the employees name and the department name that they manage.
- Temp $\leftarrow$ (Employee $\bowtie_{Ssn=Mgr\_Ssn}$ Department)
- Result $\leftarrow \pi_{Fname, Minit, Lname, Dname}$(Temp)

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

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

Left Outer Join

- List the employees name and the department name that they manage. **If they don't manage one, then indicate this with a null value.**
- Temp ← (Employee $\bowtie_{Ssn=Mgr\_Ssn}$ Department)
- Result ← $\pi_{Fname, Minit, Lname, Dname}$(Temp)

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

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

Left Outer Join

- List the employees name and the department name that they manage. **If they don't manage one, then indicate this with a null value.**
- Temp ← (Employee $\overset{\text{Ssn=Mgr\_Ssn}}{\bowtie}$ Department)
- Result ← $\pi_{\text{Fname, Minit, Lname, Dname}}$(Temp)

RESULT

Fname	Minit	Lname	Dname
John	B	Smith	NULL
Franklin	T	Wong	Research
Alicia	J	Zelaya	NULL
Jennifer	S	Wallace	Administration
Ramesh	K	Narayan	NULL
Joyce	A	English	NULL
Ahmad	V	Jabbar	NULL
James	E	Borg	Headquarters

Right Outer Join

- List the employees name and the department name that they manage. If they don't manage one, then indicate this with a null value.
- Temp ← (**Department** $\overset{\text{⟕}}{}_{\text{Mgr\_Ssn= Ssn}}$ **Employee**)
- Result ← $\pi_{\text{Fname, Minit, Lname, Dname}}$(Temp)

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

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

Full Outer Join

List the employees name and the department name that they manage. If they don't manage one **or the department have no manager,** then indicate this with a null value.

$$Temp \leftarrow Employee \overset{}{\bowtie}_{Ssn = Mgr\_Ssn} Department$$

$$Result \leftarrow \pi_{Fname,\ Lname,\ Dname}(Temp)$$

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
CS	6		

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

Full Outer Join vs Cartesian Product

What is the difference ?
OR ...
are they same ... ?

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
CS	6		

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

Outer Join Operation

- In INNER JOIN, tuples without a *matching* are eliminated from the join result
  - Tuples with null are also eliminated
  - This amounts to loss of information.

OUTER joins operations are used when we want to keep

- **all the tuples in R** in the join result , or
- **all the tuples in S** in the join result, or
- **all tuples in both relations R and S** in the join result

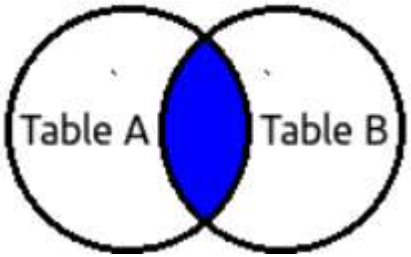
Outer Join Operation

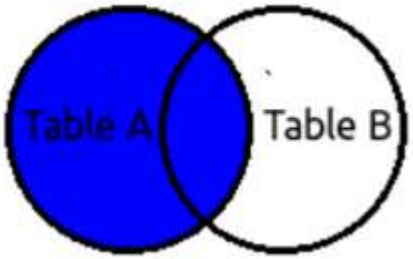
- **Left outer join**: keeps every tuple in R, denoted as R ⟕ S
  - if no matching tuple is found in S, then the attributes of S in the join result are filled with null values.

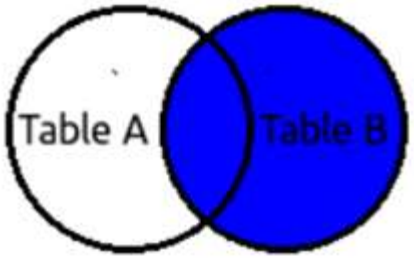
- **Right outer join:** keeps every tuple in S in the result of R ⟖ S.

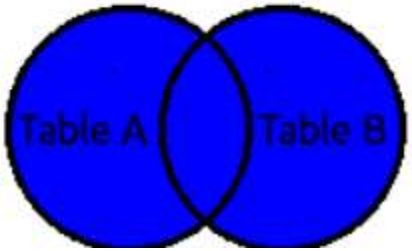
- **Full outer join:** keeps all tuples in both the left and the right relations. It is denoted by ⟗

Inner and Outer Joins









Another Example Outer Join

List the employees name and the Project name that they work on. If they don't work on any project or a project have no employee working on it, then indicate this with a null value.

PROJECT

Pname	Pnumber	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

WORKS_ON

Essn	Pno	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0

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

Yet another Example

Find SSN of employees who work on all the projects of Dnum= 4

PROJECT

Pname	Pnumber	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

WORKS_ON

Essn	Pno	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

- PD4(Pno) ← π_{Pnumber} ($\sigma_{\text{DNUM=4}}$ Project)
- Ssn_Pnos ← $\pi_{\text{Essn,Pno}}$ (Works_on)
- SSNS(ssn)← Ssn_Pnos ??? PD4

DIVISION

Yet an other Example

Find SSN of employees who work on all the projects of Dnum= 4

PROJECT

Pname	Pnumber	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

PD4

Pno
10
30

SSN_PNOS

Essn	Pno
123456789	1
123456789	2
666884444	3
453453453	1
453453453	2
333445555	2
333445555	3
333445555	10
333445555	20
999887777	30
999887777	10
987987987	10
987987987	30
987654321	30
987654321	20
888665555	20

- PD4(Pno) ← π_{Pnumber} ($\sigma_{\text{DNUM=4}}$ Project)
- Ssn_Pnos ← $\pi_{\text{Essn,Pno}}$ (Works_on)
- **SSNS(ssn)← Ssn_Pnos ÷ PD4**

DIVISION

DIVISION (Binary Operation)

Division operation is applied to two relations R1 and R2

R1(Attributes_R1) ÷ R2(Attributes_R2)

- where Attributes_R2 ⊂ Attributes_R1.

Let **Result = R1 ÷ R2**

Attr_Res = Attributes_R1 - Attributes_R2

- Attr_Res is a set of attributes of R1 that are not the attributes of R2.

R2

A
a1
a2
a3

Result

B
b1
b4

R1

A	B
a1	b1
a2	b1
a3	b1
a4	b1
a1	b2
a3	b2
a2	b3
a3	b3
a4	b3
a1	b4
a2	b4
a3	b4

For a **tuple t** to appear in the result of the DIVISION, the values in t must appear in R1 in combination with *every* tuple in R2.

Example of DIVISION

Find **SSN** of employees who work on all the projects that *John Smith* works on

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO

PROJECT

PNAME	PNUMBER	PLOCATION	DNUM

WORKS_ON

ESSN	PNO	HOURS

- Smith ← $\sigma_{\text{fname='John' and lname='Smith'}}$ (Employee)
- Smith_Pnos ← π_{Pno} (Works_on $\bowtie_{\text{essn=ssn}}$ Smith)
- Ssn_Pnos ← $\pi_{\text{Essn,Pno}}$ (Works_on)
- SSNS(ssn)← Ssn_Pnos ÷ Smith_Pnos

SMITH_PNOS

Pno
1
2

SSNS

Ssn
123456789
453453453

SSN_PNOS

Essn	Pno
123456789	1
123456789	2
666884444	3
453453453	1
453453453	2
333445555	2
333445555	3
333445555	10
333445555	20
999887777	30
999887777	10
987987987	10
987987987	30
987654321	30
987654321	20
888665555	20

Examples of Queries in RA

Find the names of employees who work on *all* the projects controlled by department number 5.

$$T1(Pno) \leftarrow \pi_{Pnumber}(\sigma_{Dnum=5}(Project))$$

$$T2 \leftarrow \pi_{Essn, Pno}(Work\_On)$$

$$T3 \leftarrow (T2 \div T1)$$

$$R \leftarrow \pi_{LNAME, FNAME}(T3 * Employee)$$

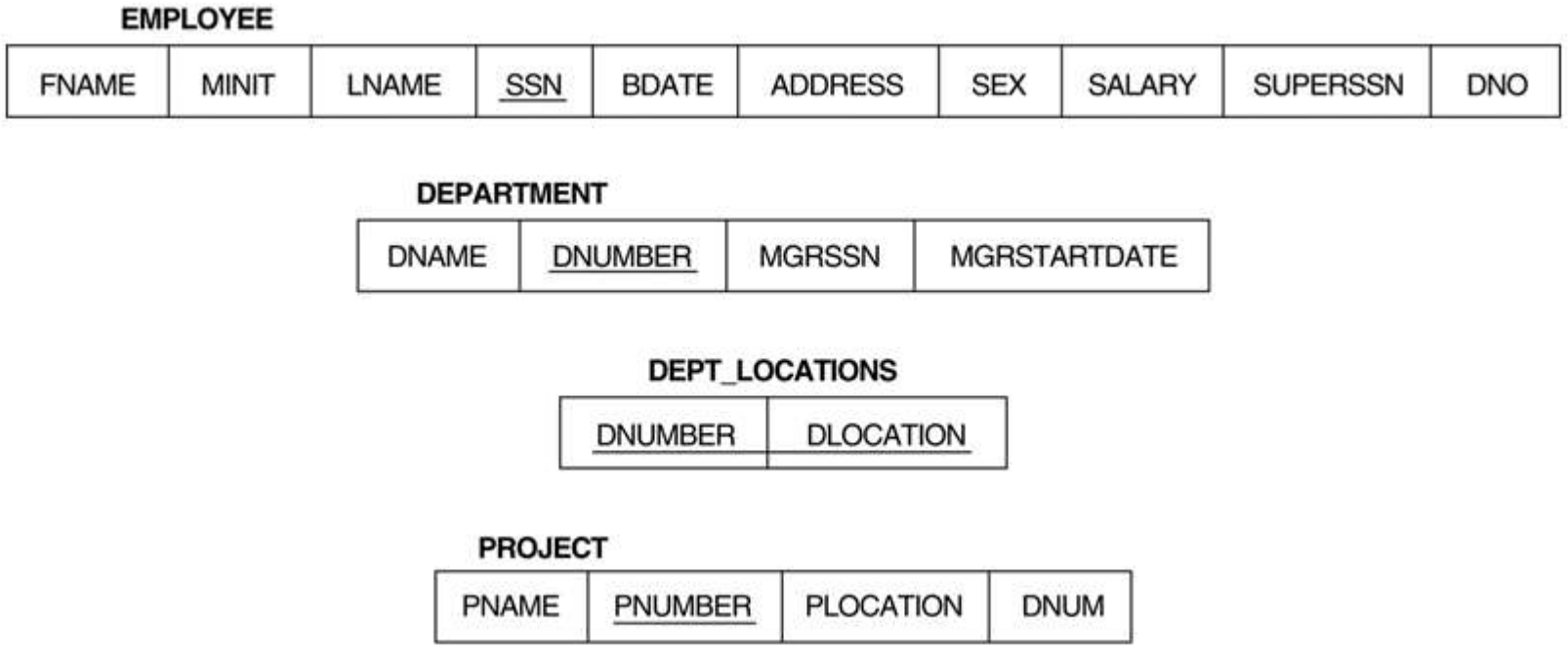
PROJECT

Pname	Pnumber	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

Essn	Pno
123456789	1
123456789	2
666884444	3
453453453	1
453453453	2
333445555	2
333445555	3
333445555	10
333445555	20
999887777	30
999887777	10
987987987	10
987987987	30
987654321	30
987654321	20
888665555	20

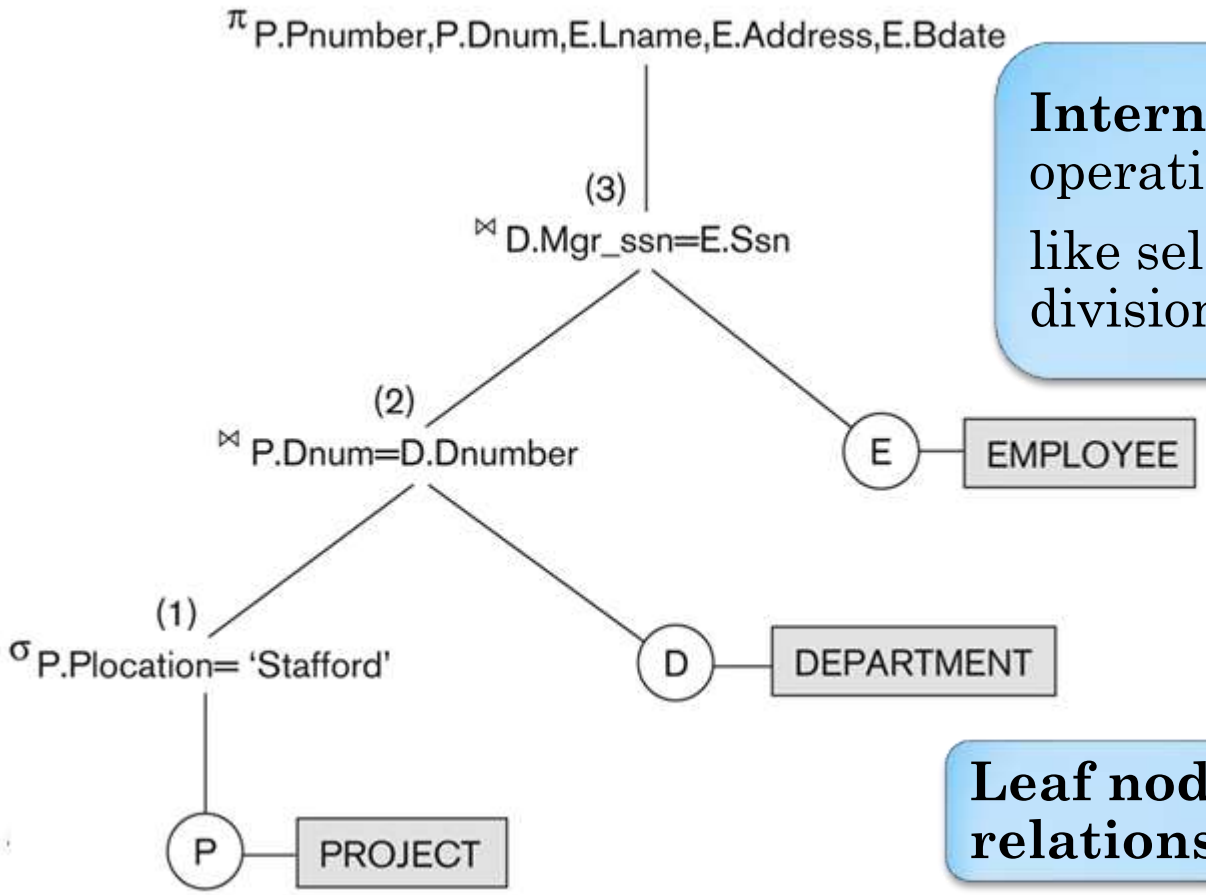
Example

For every project located in **'Stafford'**, list the **project no**, the **controlling department no**, and the **department manager's last name, address, and birth date**.



Example of Query Tree

For every project located in **'Stafford'**, list the **project no**, the **controlling department no**, and the **department manager's last name, address, and birth date**.



Internal Nodes stand for operations

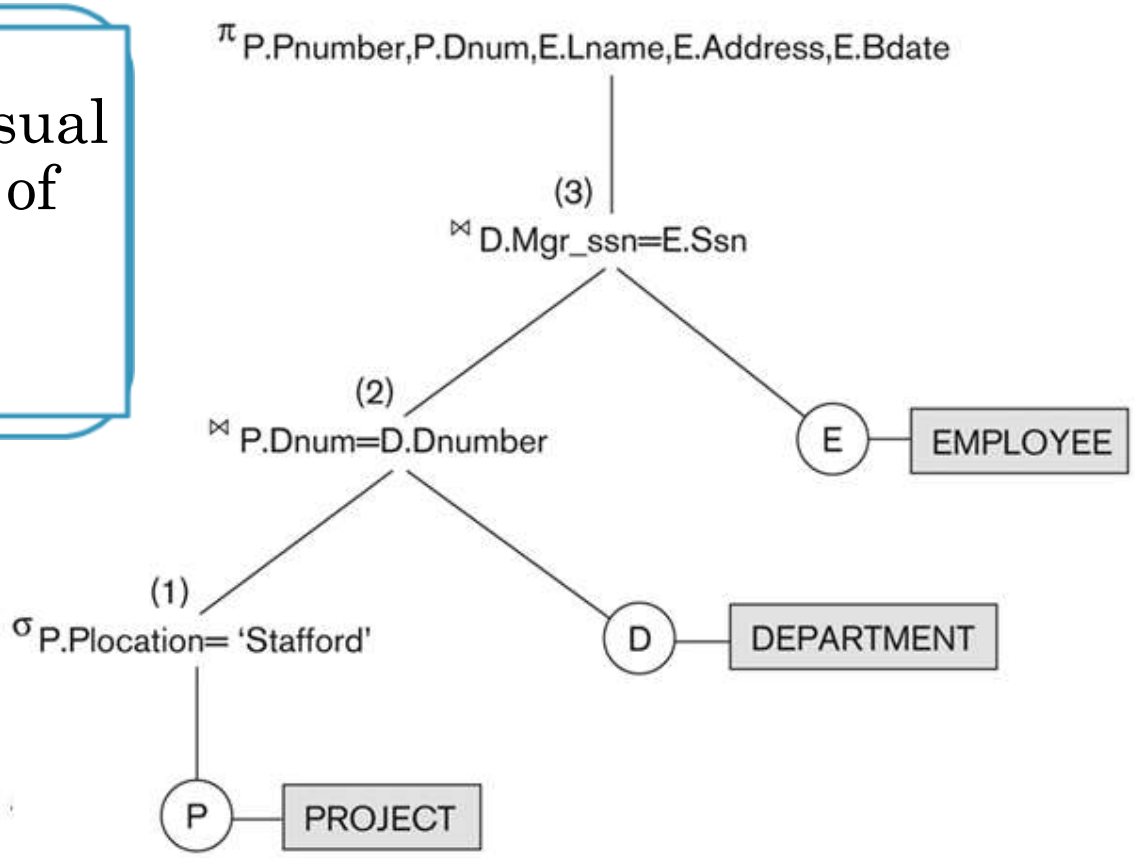
like selection, projection, join, division, ….

Leaf nodes represent **base relations**

Query Tree is an internal data structure to represent a query

Standard technique to **estimate the work done** in executing the query, and the ***optimization of execution***

A tree gives a good visual feel of the complexity of the query and the operations involved



Recursive Closure Operation

- This can't be specified in general using **Relational Algebra**
- **Example:** Retrieve all SUPERVISEES of an EMPLOYEE **e** at all levels — that is,
  - all employees **e`** directly supervised by **e**;
  - all employees **e``** directly supervised by each employee **e`**;
  - all employees **e```**directly supervised by each employee **e``**;
  - and so on.

We can retrieve employees at each level and then take their union, however, we cannot specify a query such as

"retrieve the supervisees of 'James Borg' at all levels" without utilizing a looping mechanism.

The SQL3 standard includes syntax for recursive closure.

Recursive Closure Operation

(Borg's SSN is 888665555)

	(SSN)	(SUPERSSN)
SUPERVISION	SSN1	SSN2
	123456789	333445555
	333445555	888665555
	999887777	987654321
	987654321	888665555
	666884444	333445555
	453453453	333445555
	987987987	987654321

RESULT 1	SSN
	333445555
	987654321

(Supervised by Borg)

RESULT 2	SSN
	123456789
	999887777
	666884444
	453453453
	987987987

(Supervised by Borg's subordinates)

RESULT	SSN
	123456789
	999887777
	666884444
	453453453
	987987987
	333445555
	987654321

(RESULT1 ∪ RESULT2)

PRACTICE QUESTION

- **Do example queries and the questions at the end of Relational Algebra Chapter in**

  - *Fundamentals of Database Systems* (6th Edition), Ramez Elmasri
  - *Database Systems: The Complete Book*, Hector Garcia-Molina, Jeffrey Ullman, Jennifer Widom
  - *Database Management Systems*, Raghu Ramakrishnan

Relational Algebra Operators

- Relational Algebra consists of several groups of operations
  - **Unary Relational Operations**
    - SELECT (symbol: **σ** (sigma))
    - PROJECT (symbol: **π** (pi))
    - RENAME (symbol: **ρ** (rho))
  - **Relational Algebra Operations From Set Theory**
    - UNION (∪), INTERSECTION (∩), DIFFERENCE (–)
    - CARTESIAN PRODUCT (**x**)
  - **Binary Relational Operations**
    - JOIN (several variations of JOIN exist)
    - **DIVISION**
  - **Additional Relational Operations**
    - OUTER JOINS, OUTER UNION
    - AGGREGATE FUNCTIONS (These compute summary of information: for example, SUM, COUNT, AVG, MIN, MAX)

SQL