## Join Operations

Theta, Equi, Natural, Outer and Semi

### Relational Algebra operations

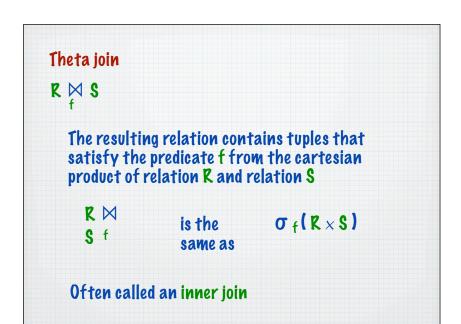
Selection	σ
Projection	π
Cartesian Product	X
Union	U
Set Difference	-
Join	M
Intersection	n
Division	÷

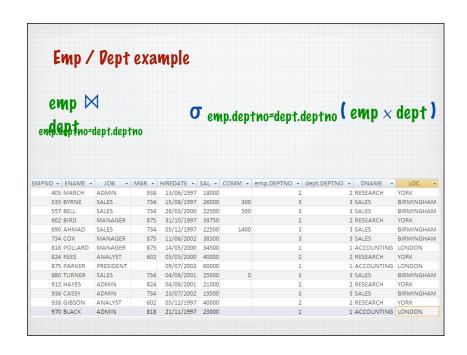
We have seen Inner Join performed using CP and selection

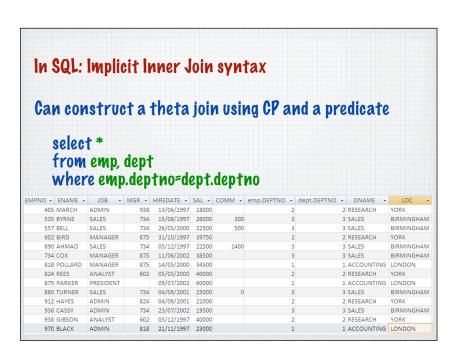
### Relational Algebra operations

In fact, there are 6 different types of Join

el., 1, 1, 1,			
Theta join	M		
Equi join	M		
Natural join	M		
Outer join	×		
Semi join	×		
Anti join	⊲		







# SQL 92: Explicit Inner Join syntax Can construct a theta join using specific inner join syntax select \* or expression from relation inner join relation on field operator field the predicate goes here The on clause must be included

### SQL 92: Explicit Inner Join syntax Can construct a theta join using specific inner join syntax select \* from emp inner join dept on emp.deptno=dept.deptno EMPNO - ENAME - JOB - MGR - HIREDATE - SAL - COMM - emp.DEPTNO - dept.DEPTNO - DNAME - LOC 405 MARCH ADMIN 938 13/06/1997 18UUU 535 BYRNE SALES 734 15/08/1997 26000 734 26/03/2000 22500 3 SALES BIRMINGHAM 26/03/2000 BIRMINGHAM 602 BIRD MANAGER 875 31/10/1997 39750 2 RESEARCH YORK SALES 734 05/12/1997 22500 MANAGER 875 11/06/2002 38500 690 AHMAD 3 SALES 3 SALES BIRMINGHAM 734 COX BIRMINGHAM 818 POLLARD MANAGER 875 14/05/2000 34500 824 REES ANALYST 602 05/03/2000 40000 1 ACCOUNTING LONDON 824 REES 2 RESEARCH 875 PARKER PRESIDENT 09/07/2002 1 ACCOUNTING LONDON 912 HAYES ADMIN 04/06/2001 21000 2 RESEARCH YORK 912 HAYES ADMIN 824 04/06/2001 21000 936 CASSY ADMIN 734 23/07/2002 19500 3 SALES BIRMINGHAM 938 GIBSON ANALYST 2 RESEARCH 938 GIBSON ANALYSI 602 05/12/1997 40000 970 BLACK ADMIN 818 21/11/1997 23000 1 ACCOUNTING LONDON

### Theta join

R M S The predicate f may use any of the comparison operators =, <, >, <=. >=, ≠

Example: Theta join with >= and <=

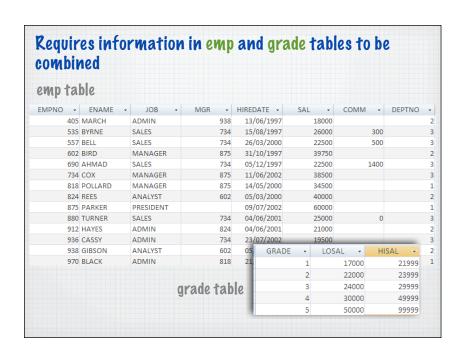
A high earner, is defined as someone who is earning salary in the top 10% of the grade they are in.

Show the ename, grade, sal and hisal for all high earners

### Example: Theta join with >= and <=

A high earner, is defined as someone who is earning salary in the top 10% of the grade they are in.

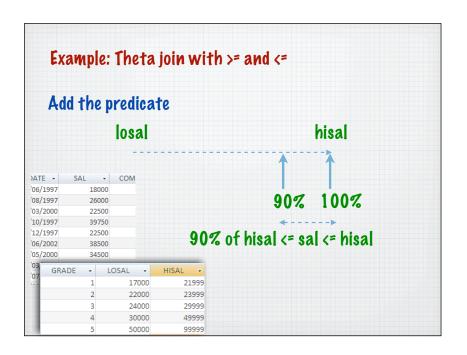
Show the ename, grade, sal and hisal for all high earners

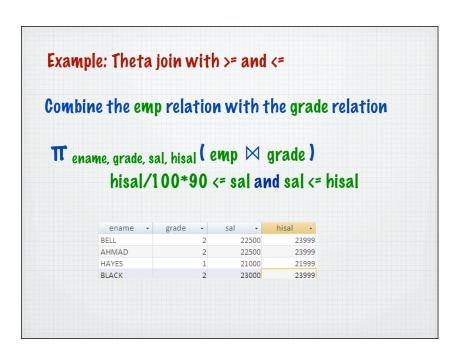


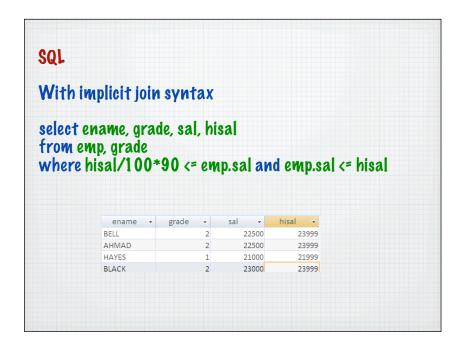
Example: Theta join with >= and <=

Combine the emp relation with the grade relation

emp ⋈ grade







### SQL

With explicit join syntax

select ename, grade, sal, hisal from emp inner join grade on hisal/100\*90 <= emp.sal and emp.sal <= hisal

Fails in many DBs- can't do explicit join syntax without common columns

### Equi join

R M S

A particular type of Theta join that only uses equality in the predicate/condition

emp M dept

emp.deptno=dept.deptno

### Natural join

RMS

The equi-join of relation R and S over all common attributes. One occurrence of each common attribute is eliminated from the result.

No condition required as the join occurs across ALL common attributes

### SQL

### With explicit natural join syntax (rare support)

# select \* from emp natural join dept

# No Access support

EMPNO	*	ENAME	JOB	MGR	-	HIREDATE →	SAL ▼	COMM →	emp.DEPTNO •	DNAME +	LOC
4	105	MARCH	ADMIN		938	13/06/1997	18000		2	RESEARCH	YORK
5	35	BYRNE	SALES		734	15/08/1997	26000	300	3	SALES	BIRMINGHAI
5	57	BELL	SALES		734	26/03/2000	22500	500	3	SALES	BIRMINGHAI
6	02	BIRD	MANAGER		875	31/10/1997	39750		2	RESEARCH	YORK
6	90	AHMAD	SALES		734	05/12/1997	22500	1400	3	SALES	BIRMINGHA
7	34	COX	MANAGER		875	11/06/2002	38500		3	SALES	BIRMINGHA
8	18	POLLARD	MANAGER		875	14/05/2000	34500		1	ACCOUNTING	LONDON
8	24	REES	ANALYST		602	05/03/2000	40000		2	RESEARCH	YORK
8	75	PARKER	PRESIDENT			09/07/2002	60000		1	ACCOUNTING	LONDON
8	80	TURNER	SALES		734	04/06/2001	25000	0	3	SALES	BIRMINGHA
9	12	HAYES	ADMIN		824	04/06/2001	21000		2	RESEARCH	YORK
9	36	CASSY	ADMIN		734	23/07/2002	19500		3	SALES	BIRMINGHA
9	38	GIBSON	ANALYST		602	05/12/1997	40000		2	RESEARCH	YORK
9	70	BLACK	ADMIN		818	21/11/1997	23000		1	ACCOUNTING	LONDON

### Outer join

Often when joining relations there are no matching values in the join columns. The condition used will exclude these values. To include values that don't match, use an Outer Join

Right Outer Join

Left Outer Join

### Best explained by example

Show a list of all departments and the employees who work in them

First attempt using implicit join syntax

select dept.deptno, dname, loc, ename from emp, dept where emp.deptno=dept.deptno

### Best explained by example Show a list of all departments and the employees who work in them deptno - dname -1 ACCOUNTING LONDON BLACK 1 ACCOUNTING LONDON PARKER 1 ACCOUNTING LONDON POLLARD 2 RESEARCH YORK GIBSON 2 RESEARCH HAYES Fails to show 2 RESEARCH YORK REES 2 RESEARCH YORK BIRD 2 RESEARCH YORK MARCH deptno 4 BIRMINGHAM CASSY 3 SALES BIRMINGHAM TURNER 3 SALES 3 SALES BIRMINGH DEPTNO • DNAME • LOC 3 SALES BIRMINGH 1 ACCOUNTING LONDON 3 SALES BIRMINGH 2 RESEARCH YORK 3 SALES BIRMINGH 3 SALES BIRMINGHAM

### Best explained by example

Show a list of all departments and the employees who work in them

select dept.deptno, dname, loc, ename from emp, dept where emp.deptno=dept.deptno

> We need to include unmatched rows from the dept relation

A Right Outer Join

### Right Outer join

Show a list of all departments and the employees who work in them

Emp × Pept
emp.deptno=dept.deptno

Includes unmatched tuples from the right

### **Explicit Outer Join syntax**

Can construct an Outer join using explicit right or left join syntax

select \* or expression from relation left join relation on field operator field

 $R \underset{f}{\times} S$ 

select \* or expression from relation right join relation on field operator field

R × S

### SQL

With explicit right outer join syntax

select dept.deptno, dname, loc, ename from emp right join dept on emp.deptno = dept.deptno

deptno	*	dname 🔻	loc ⋅	ename 🔻
	1	ACCOUNTING	LONDON	POLLARD
	1	ACCOUNTING	LONDON	PARKER
	1	ACCOUNTING	LONDON	BLACK
	2	RESEARCH	YORK	MARCH
	2	RESEARCH	YORK	BIRD
	2	RESEARCH	YORK	REES
	2	RESEARCH	YORK	HAYES
	2	RESEARCH	YORK	GIBSON
	3	SALES	BIRMINGHAM	BYRNE
	3	SALES	BIRMINGHAM	BELL
	3	SALES	BIRMINGHAM	AHMAD
	3	SALES	BIRMINGHAM	COX
	3	SALES	BIRMINGHAM	TURNER
	3	SALES	BIRMINGHAM	CASSY
	4	OPERATIONS	LEEDS	

### Full Outer join

Includes tuples from both sides where there are null values

R × S

Full Outer Join

### Full Outer join Change to emp table to demonstrate this - the president doesn't have a department EMPNO + ENAME + JOB + MGR + HIREDATE + SAL - COMM - DEPTNO 405 MARCH 13/06/1997 535 BYRNE SALES 734 15/08/1997 26000 557 BELL SALES 26/03/2000 22500 MANAGER 602 BIRD 875 31/10/1997 39750 690 AHMAD SALES 05/12/1997 22500 1400 MANAGER 875 11/06/2002 734 COX 38500 818 POLLARD MANAGER 14/05/2000 34500 602 05/03/2000 824 REES ANALYST 40000 PRESIDENT 875 PARKER 09/07/2002 60000 734 04/06/2001 880 TURNER SALES 25000 04/06/2001 912 HAYES ADMIN 824 21000 936 CASSY ADMIN 734 23/07/2002 19500 938 GIBSON ANALYST 602 05/12/1997 23000

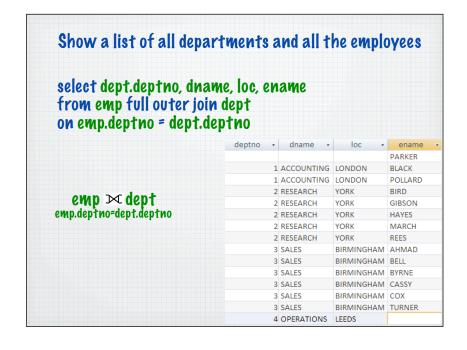
### Full Outer join Show a list of all departments and all the employees select dept.deptno, dname, loc, ename from emp, dept where emp.deptno=dept.deptno deptno - dname - loc 1 ACCOUNTING LONDON 1 ACCOUNTING LONDON PARKER 1 ACCOUNTING LONDON 2 RESEARCH YORK GIBSON 2 RESEARCH HAYES 2 RESEARCH YORK REES We lose both 2 RESEARCH YORK BIRD department 4 and 3 SALES BIRMINGHAM CASSY 3 SALES BIRMINGHAM TURNER the President BIRMINGHAM COX

3 SALES

SALES

BIRMINGHAM AHMAD BIRMINGHAM BELL

BIRMINGHAM BYRNE



### Full Outer join

However - not many databases support the full outer join syntax

### Emulate it with a Union

$$\mathbf{R} \underset{f}{\bowtie} \mathbf{S} \quad \cup \quad \mathbf{R} \underset{f}{\bowtie} \mathbf{S}$$

### Show a list of all departments and all the employees

select dept.deptno, dname, loc, ename from emp left join dept on emp.deptno = dept.deptno UNION

select dept.deptno, dname, loc, ename

from emp right join dept on emp.deptno = dept.deptno

deptno	۳	dname →	loc +	ename 🕶
				PARKER
	1	ACCOUNTING	LONDON	BLACK
	1	ACCOUNTING	LONDON	POLLARD
	2	RESEARCH	YORK	BIRD
	2	RESEARCH	YORK	GIBSON
	2	RESEARCH	YORK	HAYES
	2	RESEARCH	YORK	MARCH
	2	RESEARCH	YORK	REES
	3	SALES	BIRMINGHAM	AHMAD
	3	SALES	BIRMINGHAM	BELL
	3	SALES	BIRMINGHAM	BYRNE
	3	SALES	BIRMINGHAM	CASSY
	3	SALES	BIRMINGHAM	COX
	3	SALES	BIRMINGHAM	TURNER
		ODERATIONS	LEEDS	

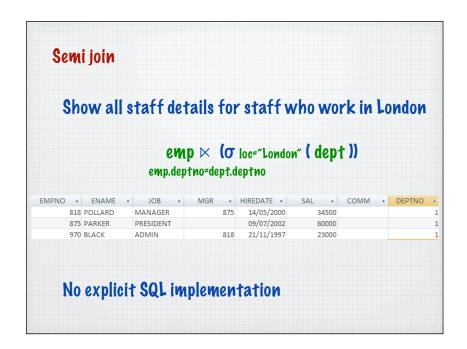
### Semi join

Theta join, but only includes columns from one side

R × S

Only includes columns from R

Same symbol as an outer join



# Anti join Left for you to find out .... Hint: Sometimes described as the opposite of a semi join