WEEK 4

THE RELATIONAL ALGEBRA BINARY OPERATION OF INNER JOIN (NATURAL JOIN AND EQUI JOIN)

STUDENT OBJECTIVES

- Upon completion of this video, you should be able to:
  - Identify the symbols for natural join and equi join.
  - Determine if the join will be possible based on the given tables and join information.
  - Write a relational algebra expression that uses JOINS given two tables based on a given query.
  - Given 2 tables and a JOIN relational algebra expression, show the new table that would be returned once the expression is performed.
  - Given a JOIN relational algebra expression and two tables, explain in simple English what query is answered by the expression.

JOIN

- A join is just a Cartesian Product **X** with a Selection **σ** to find matches. The selection will remove some of the rows/tuples from the returned Cartesian Product.
- The attributes that will be in the new table depends on if you are doing an equi join or a natural join

- Symbol → ⋈

- Example Expression:

Table1 ⋈ Table2 (natural join)

Table1 ⋈ columnnametable1=columnnametable2 Table2 (equi join)

ANSWER FOR Table1 ⋈ Table2

First Table Name

ANSWER FOR Table1 ⋈ Age=ID Table2

Symbol for Join and maybe a subscript saying which column to join on.

Second Table Name

JOIN

- A join is just a Cartesian Product **X** with a Selection **σ** to find matches. The selection will remove some of the rows/tuples from the returned Cartesian Product.
- T… pends on if you are doing an equi join or a natural join
- S…
- Example Expression:

Table1

ID	FirstName	LastName	Age
12	Homer	Smith	24
24	Gene	Simpson	33
45	Walter	Reid	33

Table2

ID	FirstName	LastName	Age
33	Marg	Jones	28
24	Gene	Simpson	33

Table1 ⋈ Table2 (natural join)

Table1 ⋈ columnnametable1=columnnametable2 **Table2 (equi join)**

join fields that 1) have same name 2) have same value

ANSWER FOR Table1 ⋈ Table2

ANSWER FOR Table1 ⋈ $_{Age=ID}$ Table2

Symbol for Join and maybe a subscript

Second Table Name

ID	FirstName	LastName	Age	Table2.ID	Table2.FirstName	Table2.LastName	Table2.Age
12	Homer	Smith	24	24	Gene	Simpson	13
24	Gene	Simpson	33	33	Marg	Jones	28
45	Walter	Reid	33	33	Marg	Jones	28

MORE ON JOINS

- $R \bowtie_{\text{Boolean formula}} S$: Produces a relation that contains tuples satisfying a condition from the Cartesian Product of R X S where the formula can contain comparisons using one of <,<=,>,>=,=,<> connected by and, or and not
- Example:

$$Department \bowtie_{ManagerSSN = SuperSSN} Employee$$

$$Staff \bowtie_{(Salary > Salary \text{ and } JobLevel=JobLevel)} Manager$$

NATURAL JOIN MORE INFO:

- If the two tables being joined have the same attribute name(s), it is a natural join and the attribute will only show up ONCE in the resulting table
- The natural join includes EACH pair of attributes with the same name, "AND" ed together, for example:
  - $Q \leftarrow R(A,B,C,D) \bowtie S(C,D,E)$
  - Result would only keep one copy of each pair
    - R.C=S.C AND R.D = S.D
  - and would give: Q(A,B,C,D,E)

- **Equi Join:** when all of the comparisons are =, then it is called an equi join and pairs of the attributes are returned that are equal (i.e. attribute from both sides of the equals are returned).
- **Natural Join:** when all of the comparisons are = and it matches any attribute in Table1 that has the same name as the attribute in Table2. The attribute is only shown once.

Table1

A	B	C	D
7	Cow	Pink	22
8	Dog	Pink	33
9	Cow	Red	44

Table2

A	E	C	F	D
6	Cow	Blue	Hat	33
6	Cow	Blue	Sock	44
8	Cow	Pink	Shoe	44
8	Cat	Pink	Hat	33

A
7
8
9

A	B	C	D
7	Cow	Pink	22
8	Dog	Pink	33
9	Cow	Red	44

				F	D	le2.A	E	Table2.C	F	Table2.D	
				Shoe	44		Cow	Pink	Shoe	44	
8	7	8	Cat	Pink	Hat	33		Cat	Pink	Hat	33
	8		Dog	Pink	33	8		Cow	Pink	Shoe	44

A	B	C	D	Table1.A	Table1.B	Table1.C	Table1.D
7	Cow	Pink	22	8	Dog	Pink	33
7	Cow	Pink	22	9	Cow	Red	44
8	Dog	Pink	33	9	Cow	Red	44

MORE REALISTIC EXAMPLES OF JOINS:

Example 1:

Equi Join on Department and Project:

Project ⋈ $_{\text{Project.DeptNumber = Department.DeptNumber}}$ Department

ProjectName	ProjectNumb	ProjectLocation	project.DeptNum	Department.DeptNun	DeptName	ManagerS
Accounting Upd	A1	Toronto	S7G	S7G	Safety Department	
Inventory2	I2	London	S7G	S7G	Safety Department	
Acct6	A6	Toronto	S7G	S7G	Safety Department	
Payroll	P1	Springfield	G8H	G8H	Head Office	
Acc3	A3	Springfield	G8H	G8H	Head Office	
Payroll2	P2	London	G8H	G8H	Head Office	
Payroll3	P3	London	G8H	G8H	Head Office	
Inventory	I1	Toronto	G8H	G8H	Head Office	

Example 2:

Natural Join on Department and Project:

Project ⋈ Department

ProjectName	ProjectNumb	ProjectLocation	Department.DeptNun	DeptName	ManagerSSN	ManagerS
Accounting Upd	A1	Toronto	S7G	Safety Department	3	
Inventory2	I2	London	S7G	Safety Department	3	
Acct6	A6	Toronto	S7G	Safety Department	3	
Payroll	P1	Springfield	G8H	Head Office	4	
Acc3	A3	Springfield	G8H	Head Office	4	
Payroll2	P2	London	G8H	Head Office	4	
Payroll3	P3	London	G8H	Head Office	4	
Inventory	I1	Toronto	G8H	Head Office	4	

QUESTION: What will be the resulting table of the following relational algebra expression?

$\pi_{\text{LastName, DeptName}}$ (Employee ⋈ Department)

remember - to remove duplicate.

Department

DeptNun	DeptName	ManagerEmpID	ManagerStartdate	Cli
G8H	Head Office	4	12/12/1999	
S7G	Safety Department	3	11/11/1998	
Y5J	Research Department	6	12/24/1998	
		0		

Record: 1 of 3 No Filter Search

4	I1
4	I2
4	P1
4	P2
4	P3
6	I2
12	A3
0	

Record: 1 of 16

Employee

EmpID	LastName	FirstName	DeptNumber	Sex	BDate	SuperSSN	Salary
1	Simpson	Bart	G8H	M	2/2/1995	2	$1,000.00
2	Smithers	Waylan	S7G	M	1/1/1960	4	$2,000.00
3	Beauvieau	Patty	Y5J	F	3/3/1959	6	$4,000.00
4	Burns	Monty	S7G	M	7/7/2020		$5,000.00
6	Simpson	Lisa	S7G	F	6/6/1990	2	$1,000.00
12	Simpson	Homer	G8H	M	8/8/1961	2	$2,000.00
0						0	$0.00

LastName	DeptName
Simpson	Head Office
Simpson	Head Office
Smithers	Safety Department
Burns	Safety Department
Simpson	Safety Department
Beauvieau	Research Department

QUESTION: What does the above result represent in English?

ANSWER: Show me all the Employee's last names and the name of the Department that they have been assigned to.

QUESTION: What will be the difference that will result in the answers in the following two relation algebra expressions?

Expression 1: Employee ⋈ Department

Expression 2: Employee $\bowtie_{\text{DeptNumber = DeptNumber}}$ Department

DeptNumber	DeptName	ManagerEmpID	ManagerStartdate	EmpID	LastName	FirstName	Sex	BDate	SuperSSN	Salary
G8H	Head Office	4	12/12/1999	1	Simpson	Bart	M	2/2/1995	2	$1,000.00
G8H	Head Office	4	12/12/1999	12	Simpson	Homer	M	8/8/1961	2	$2,000.00
S7G	Safety Department	3	11/11/1998	2	Smithers	Waylan	M	1/1/1960	4	$2,000.00
S7G	Safety Department	3	11/11/1998	4	Burns	Monty	M	7/7/2020		$5,000.00
S7G	Safety Department	3	11/11/1998	6	Simpson	Lisa	F	6/6/1990	2	$1,000.00
Y5J	Research Department	6	12/24/1998	3	Beauvieau	Patty	F	3/3/1959	6	$4,000.00

Department.DeptNumber	DeptName	ManagerEmpID	ManagerStartdate	EmpID	LastName	FirstName	Employee.DeptNumber	Sex	BDate	SuperSSN	Salary
G8H	Head Office	4	12/12/1999	1	Simpson	Bart	G8H	M	2/2/1995	2	$1,000.00
G8H	Head Office	4	12/12/1999	12	Simpson	Homer	G8H	M	8/8/1961	2	$2,000.00
S7G	Safety Department	3	11/11/1998	2	Smithers	Waylan	S7G	M	1/1/1960	4	$2,000.00
S7G	Safety Department	3	11/11/1998	4	Burns	Monty	S7G	M	7/7/2020		$5,000.00
S7G	Safety Department	3	11/11/1998	6	Simpson	Lisa	S7G	F	6/6/1990	2	$1,000.00
Y5J	Research Department	6	12/24/1998	3	Beauvieau	Patty	Y5J	F	3/3/1959	6	$4,000.00

QUESTION: What will the following expression return?

Temp(LN, FN, SSN) ← $\pi_{\text{LastName,FirstName,EmpID}}$ (Employee)

Result ← Temp $\bowtie_{\text{SSN=SuperSSN}}$ ($\pi_{\text{LastName,FirstName,SuperSSN,EmpID}}$(Employee))

LN	FN	SSN
Simpson	Bart	1
Smithers	Waylan	2
Beauvieau	Patty	3
Burns	Monty	4
Simpson	Lisa	6
Simpson	Homer	12

Employee

EmpID	LastName	FirstName	DeptNumber	Sex	BDate	SuperSSN	Salary
1	Simpson	Bart	G8H	M	2/2/1995	2	$1,000.00
2	Smithers	Waylan	S7G	M	1/1/1960	4	$2,000.00
3	Beauvieau	Patty	Y5J	F	3/3/1959	6	$4,000.00
4	Burns	Monty	S7G	M	7/7/2020		$5,000.00
6	Simpson	Lisa	S7G	F	6/6/1990	2	$1,000.00
12	Simpson	Homer	G8H	M	8/8/1961	2	$2,000.00

LN	FN	SSN	LastName	FirstName	SuperSSN	EmpID
Smithers	Waylan	2	Simpson	Bart	2	1
Smithers	Waylan	2	Simpson	Lisa	2	6
Smithers	Waylan	2	Simpson	Homer	2	12
Burns	Monty	4	Smithers	Waylan	4	2
Simpson	Lisa	6	Beauvieau	Patty	6	3

QUESTION: In English, what does the above expression represent?

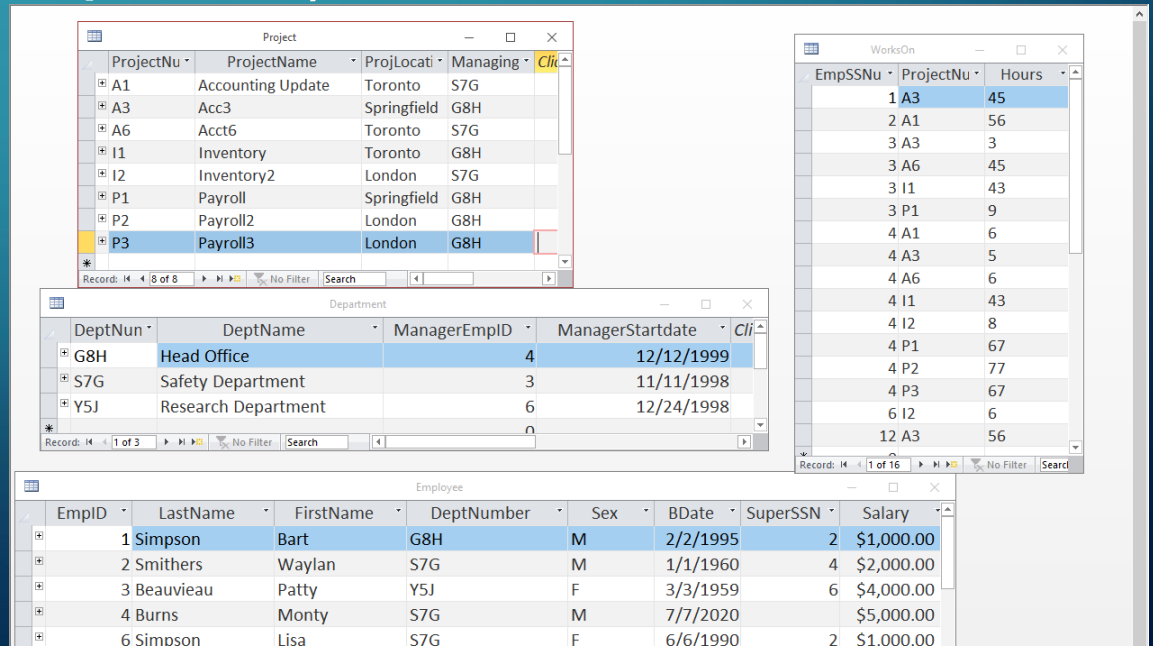
ANSWER: Show me the id and first and last names of the supervisors and the ids and first and last names of the employees that they supervise.

QUESTION: Write the Relational Algebra expression to print the department name, and the manager's first and last name:

$$\pi_{\text{DeptName, LastName, FirstName}} (\text{Employee} \bowtie_{\text{EmpID = ManagerEmpID}} \text{Department})$$

QUESTION: Write the Relational Algebra expression to display department names of departments who have people making more than 4,000 dollars.

$$\text{TEMP} \leftarrow \pi_{\text{DeptNumber}} (\sigma_{\text{Salary > 4000}} (\text{Employee}))$$
$$\text{ANSWER} \leftarrow \pi_{\text{DeptName}} (\text{TEMP} \bowtie \text{Department})$$



Project

ProjectNu	ProjectName	ProjLocati	Managing
A1	Accounting Update	Toronto	S7G
A3	Acc3	Springfield	G8H
A6	Acct6	Toronto	S7G
I1	Inventory	Toronto	G8H
I2	Inventory2	London	S7G
P1	Payroll	Springfield	G8H
P2	Payroll2	London	G8H
P3	Payroll3	London	G8H

Record: 8 of 8

Department

DeptNun	DeptName	ManagerEmpID	ManagerStartdate
G8H	Head Office	4	12/12/1999
S7G	Safety Department	3	11/11/1998
Y5J	Research Department	6	12/24/1998

Record: 1 of 3

WorksOn

EmpSSNu	ProjectNu	Hours
1	A3	45
2	A1	56
3	A3	3
3	A6	45
3	I1	43
3	P1	9
4	A1	6
4	A3	5
4	A6	6
4	I1	43
4	I2	8
4	P1	67
4	P2	77
4	P3	67
6	I2	6
12	A3	56

Record: 1 of 16

Employee

EmpID	LastName	FirstName	DeptNumber	Sex	BDate	SuperSSN	Salary
1	Simpson	Bart	G8H	M	2/2/1995	2	$1,000.00
2	Smithers	Waylan	S7G	M	1/1/1960	4	$2,000.00
3	Beauvieau	Patty	Y5J	F	3/3/1959	6	$4,000.00
4	Burns	Monty	S7G	M	7/7/2020		$5,000.00
6	Simpson	Lisa	S7G	F	6/6/1990	2	$1,000.00

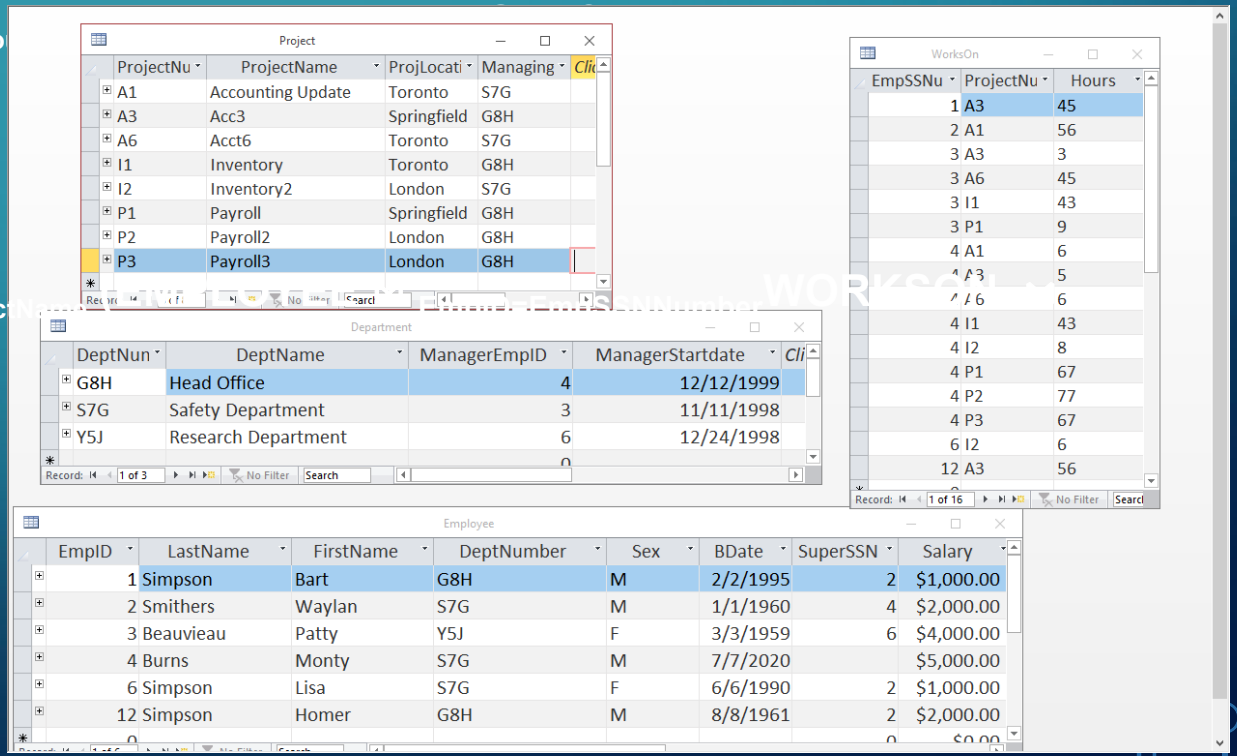
QUESTION: Write the Relational Algebra expression to print the first and last name of all employees, the name of the projects they work on and the number of hours they work on each project. ***(This is a join for a many to many relationship)***

TEMP ← π $_{\text{FirstName, LastName, ProjectNumber,Hours}}$ **(EMPLOYEE ⋈** $_{\text{EmpID=EmpSSNNumber}}$**WORKSON)**

ANSWER← π $_{\text{FirstName, LastName, Ho}}$

OR

ANSWER← π $_{\text{FirstName, LastName, Hours,Projec}}$ **PROJECT)**



Project

ProjectNu	ProjectName	ProjLocati	Managing
A1	Accounting Update	Toronto	S7G
A3	Acc3	Springfield	G8H
A6	Acct6	Toronto	S7G
I1	Inventory	Toronto	G8H
I2	Inventory2	London	S7G
P1	Payroll	Springfield	G8H
P2	Payroll2	London	G8H
P3	Payroll3	London	G8H

WorksOn

EmpSSNu	ProjectNu	Hours
1	A3	45
2	A1	56
3	A3	3
3	A6	45
3	I1	43
3	P1	9
4	A1	6
4	A3	5
4	A6	6
4	I1	43
4	I2	8
4	P1	67
4	P2	77
4	P3	67
6	I2	6
12	A3	56

Department

DeptNun	DeptName	ManagerEmpID	ManagerStartdate
G8H	Head Office	4	12/12/1999
S7G	Safety Department	3	11/11/1998
Y5J	Research Department	6	12/24/1998

Employee

EmpID	LastName	FirstName	DeptNumber	Sex	BDate	SuperSSN	Salary
1	Simpson	Bart	G8H	M	2/2/1995	2	$1,000.00
2	Smithers	Waylan	S7G	M	1/1/1960	4	$2,000.00
3	Beauvieau	Patty	Y5J	F	3/3/1959	6	$4,000.00
4	Burns	Monty	S7G	M	7/7/2020		$5,000.00
6	Simpson	Lisa	S7G	F	6/6/1990	2	$1,000.00
12	Simpson	Homer	G8H	M	8/8/1961	2	$2,000.00