Lecture 11: SQL Join

BADM/ACCY 352
Spring 2017
Instructor: Yi Yang, PhD

Class project announcement

- Previous projects:
  - Carle Hospital
  - MTD bus company
  - Uber
  - Accounting information system
  - Twitter
  - Restaurant
  - etc

Previous lectures

- Data retrieval with SELECT… FROM… command
- However, the real power of SQL is yet to come

This lecture

- We will learn SQL JOIN, the real power of relational database.

Motivation

Gallery

GAL_NUM	GAL_OWNER	GAL_PHONE
5	Alice	123-4456
6	Waters	353-2243

Painter

PTR_NUM	PTR_NAME	PTR_AREACODE	PTR_PHONE
123	Ross	901	885-4567
126	Itero	901	346-1112
127	Geoff	615	221-4456

Painting

PTNG_NUM	PTNG_TITLE	PTNG_PRICE	PTR_NUM	GAL_NUM
1338	Dawn Thunder	245.50	123	5
1339	A Faded Rose	6723.00	123	NULL
1340	The Founders	567.99	126	6
1341	Hasty Pudding Exit	145.50	123	NULL
1342	Plastic Paradise	8328.99	126	6
1343	Roamin'	785.00	127	6
1344	Wild Waters	999.00	127	5
1345	Stuff 'n Such 'n Some	9800.00	123	5

Motivation (cont.)

- List the information of painter who creates the painting 'Dawn Thunder'?

```sql
SELECT PTR_NUM FROM PAINTING WHERE PTNG_TITLE = 'Dawn Thunder';
# return 123
SELECT * FROM PAINTER WHERE PTR_NUM = 123;
```
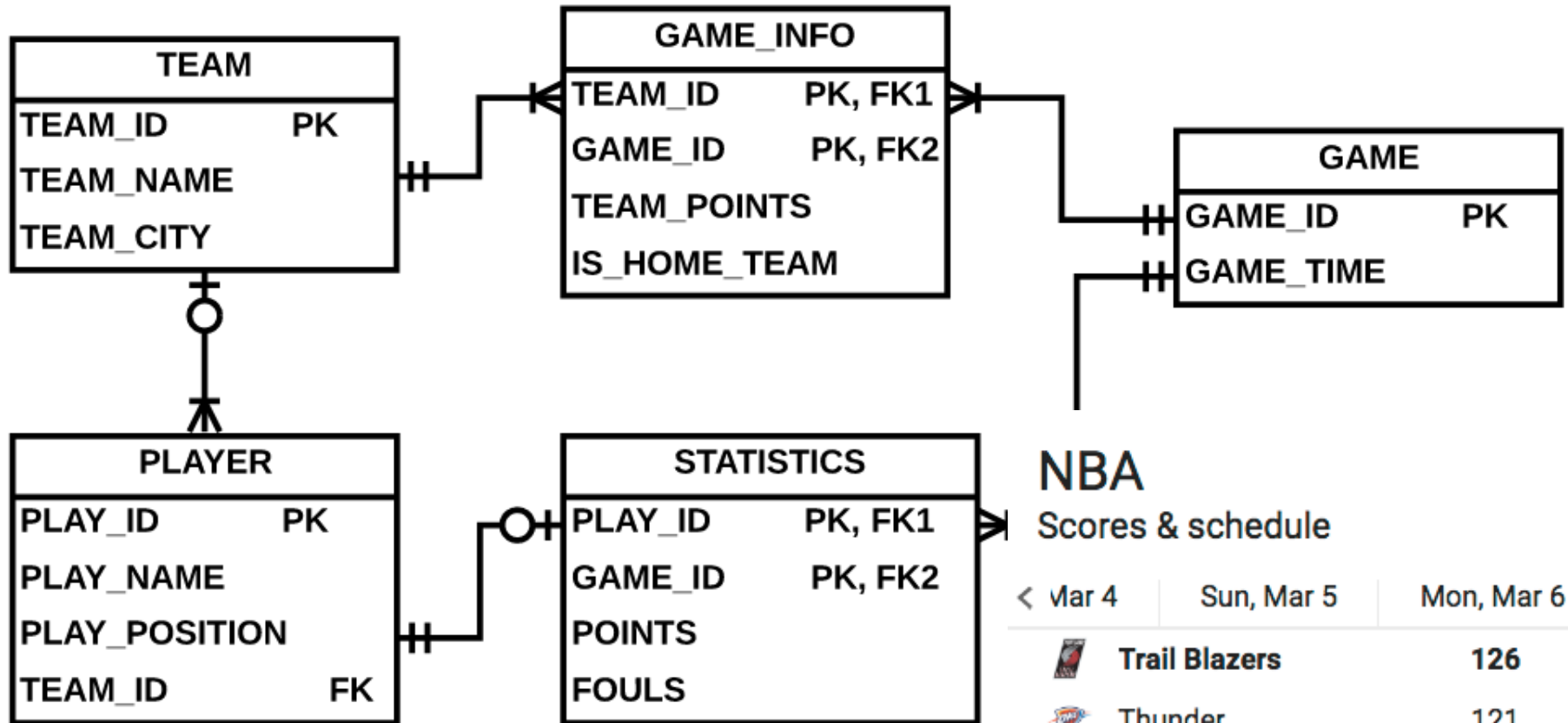
- Or use subquery

```sql
SELECT * FROM PAINTER WHERE PTR_NUM =
(SELECT PTR_NUM FROM PAINTING WHERE PTNG_TITLE = 'Dawn Thunder');
```

Motivation (cont.)

- Can you list *not only* painter information *but also* gallery information of the painting 'Dawn Thunder'.

- Things become more complicated when data are retrieved from multiple tables.
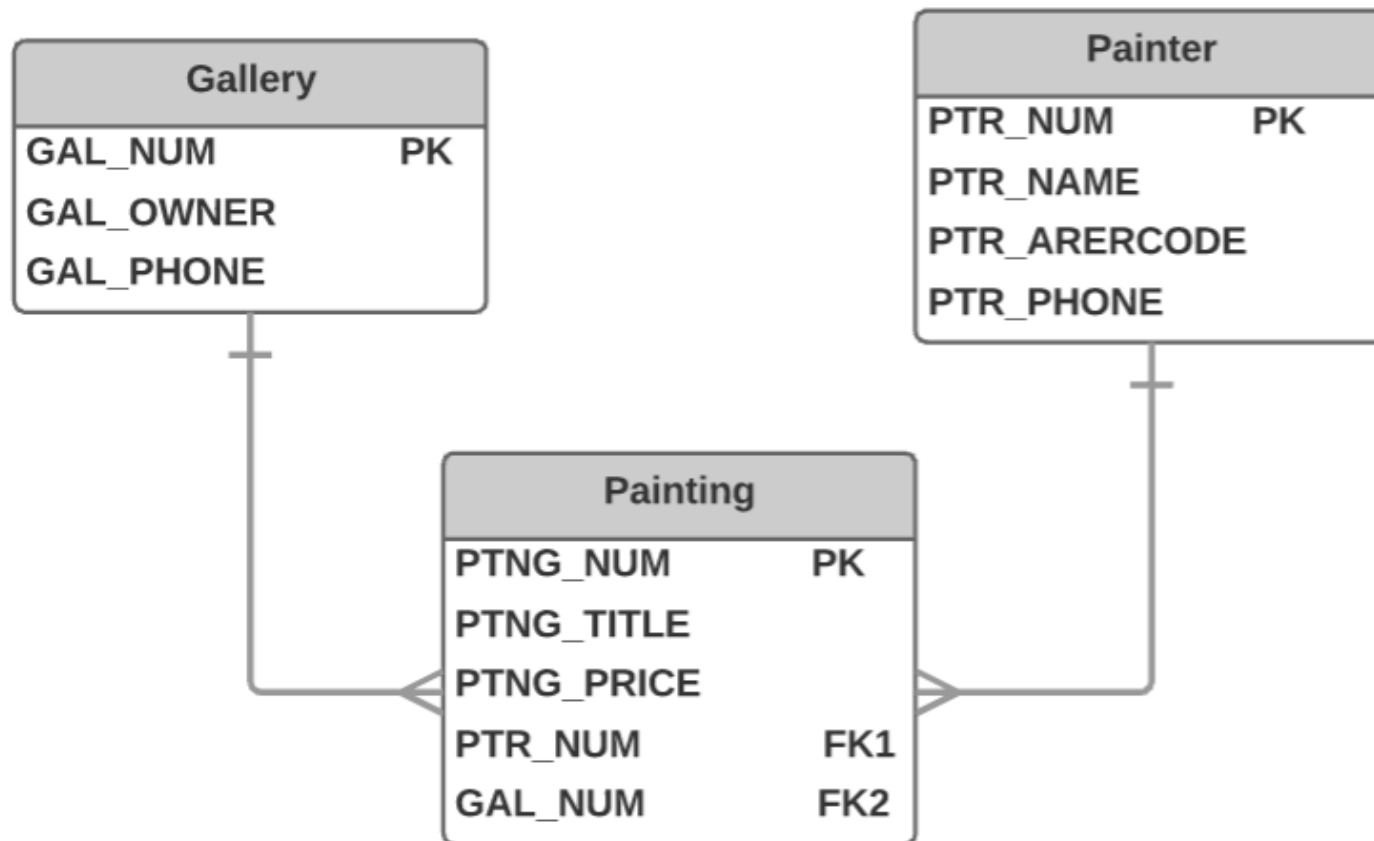
Motivation (cont.)

Query this database:
- How many games has Player A played?
- What's the average points that Player A h

Recall PK, FK

PAINTING

PTNG_NUM	PTNG_TITLE	PTNG_PRICE	PTR_NUM	GAL_NUM
1338	Dawn Thunder	245.50	123	5
1339	A Faded Rose	6723.00	123	NULL
1340	The Founders	567.99	126	6
1341	Hasty Pudding Exit	145.50	123	NULL
1342	Plastic Paradise	8328.99	126	6
1343	Roamin'	785.00	127	6
1344	Wild Waters	999.00	127	5
1345	Stuff 'n Such 'n Some	9800.00	123	5

PTNG_NUM	PTNG_TITLE	PTNG_PRICE	PTR_NUM	PTR_NAME	PTR_AREACODE	PTR_PHONE	GAL_NUM	GAL_OWNER	GAL_PHONE
1338	Dawn Thunder	245.50	123	Ross	901	885-4567	5	Alice	123-4456
1339	A Faded Rose	6723.00	123	Ross	901	885-4567	NULL	NULL	NULL
1340	The Founders	567.99	126	Itero	901	346-1112	6	Waters	353-2243
1341	Hasty Pudding Exit	145.50	123	Ross	901	885-4567	NULL	NULL	NULL
1342	Plastic Paradise	8328.99	126	Itero	901	346-1112	6	Waters	353-2243
1343	Roamin'	785.00	127	Geoff	615	221-4456	6	Waters	353-2243
1344	Wild Waters	999.00	127	Geoff	615	221-4456	5	Alice	123-4456
1345	Stuff 'n Such 'n Some	9800.00	123	Ross	901	885-4567	5	Alice	123-4456

The importance of **Entity Integrity, Referential Integrity!!**

SQL Join

- A SQL **join** clause combines records from two or more tables in a relational database.

- Natural join, Inner join, left join, right join, outer join.

Natural Join

- A natural join returns all rows (of joined tables) with <u>matching values</u> in the <u>matching columns</u> (same column name and data type) and <u>eliminates duplicate columns.</u>

- Syntax:

```
SELECT columns FROM table1 NATURAL JOIN table2;
```

PTNG_NUM	PTNG_TITLE	PTNG_PRICE	PTR_NUM	GAL_NUM
1338	Dawn Thunder	245.50	123	5
1339	A Faded Rose	6723.00	123	NULL
1340	The Founders	567.99	126	6
1341	Hasty Pudding Exit	145.50	123	NULL
1342	Plastic Paradise	8328.99	126	6
1343	Roamin'	785.00	127	6
1344	Wild Waters	999.00	127	5
1345	Stuff 'n Such 'n Some	9800.00	123	5

PTR_NUM	PTR_NAME	PTR_AREACODE	PTR_PHONE
123	Ross	901	885-4567
126	Itero	901	346-1112
127	Geoff	615	221-4456

SELECT * FROM PAINTER NATURAL JOIN PAINTING;

PTR_NUM	PTNG_NUM	PTNG_TITLE	PTNG_PRICE	GAL_NUM	PTR_NAME	PTR_AREACODE	PTR_PHONE
123	1338	Dawn Thunder	245.50	5	Ross	901	885-4567
123	1339	A Faded Rose	6723.00	NULL	Ross	901	885-4567
126	1340	The Founders	567.99	6	Itero	901	346-1112
123	1341	Hasty Pudding Exit	145.50	NULL	Ross	901	885-4567
126	1342	Plastic Paradise	8328.99	6	Itero	901	346-1112
127	1343	Roamin'	785.00	6	Geoff	615	221-4456
127	1344	Wild Waters	999.00	5	Geoff	615	221-4456
123	1345	Stuff 'n Such 'n Some	9800.00	5	Ross	901	885-4567

Only one PTR_NUM column

GAL_NUM	GAL_OWNER	GAL_PHONE
5	Alice	123-4456
6	Waters	353-2243

PTNG_NUM	PTNG_TITLE	PTNG_PRICE	PTR_NUM	GAL_NUM
1338	Dawn Thunder	245.50	123	5
1339	A Faded Rose	6723.00	123	NULL
1340	The Founders	567.99	126	6
1341	Hasty Pudding Exit	145.50	123	NULL
1342	Plastic Paradise	8328.99	126	6
1343	Roamin'	785.00	127	6
1344	Wild Waters	999.00	127	5
1345	Stuff 'n Such 'n Some	9800.00	123	5

```sql
SELECT * FROM GALLERY NATURAL JOIN PAINTING;
```

GAL_NUM	GAL_OWNER	GAL_PHONE	PTNG_NUM	PTNG_TITLE	PTNG_PRICE	PTR_NUM
5	Alice	123-4456	1338	Dawn Thunder	245.50	123
6	Waters	353-2243	1340	The Founders	567.99	126
6	Waters	353-2243	1342	Plastic Paradise	8328.99	126
6	Waters	353-2243	1343	Roamin'	785.00	127
5	Alice	123-4456	1344	Wild Waters	999.00	127
5	Alice	123-4456	1345	Stuff 'n Such 'n Some	9800.00	123

NULL gallery is not in the return result

GAL_NUM	GAL_OWNER	GAL_PHONE
5	Alice	123-4456
6	Waters	353-2243

PTR_NUM	PTR_NAME	PTR_AREACODE	PTR_PHONE
123	Ross	901	885-4567
126	Itero	901	346-1112
127	Geoff	615	221-4456

SELECT * FROM GALLERY NATURAL JOIN PAINTER;

What if there is no matching column?
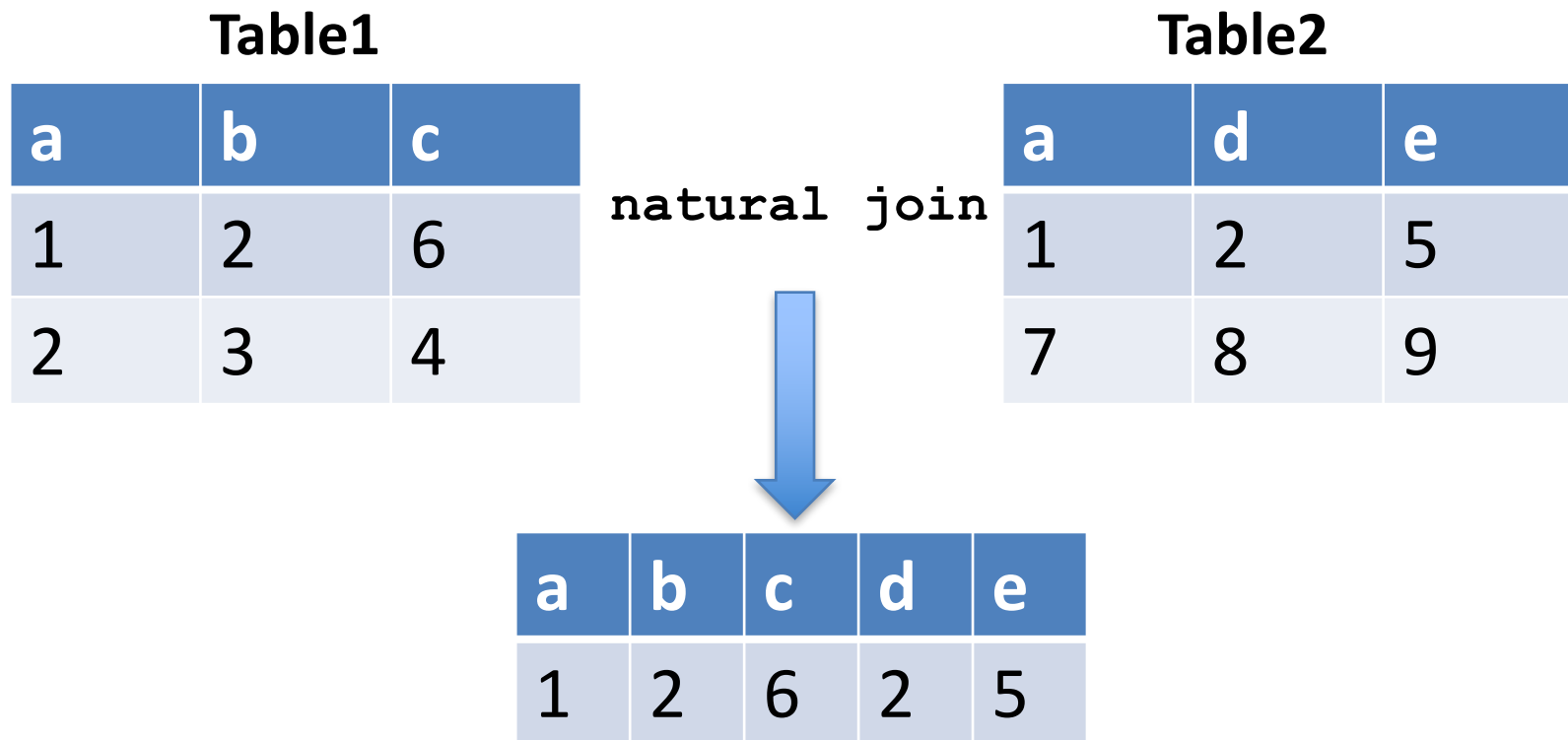Return a relational product (cartesian product) of two tables.

GAL_NUM	GAL_OWNER	GAL_PHONE	PTR_NUM	PTR_NAME	PTR_AREACODE	PTR_PHONE
5	Alice	123-4456	123	Ross	901	885-4567
6	Waters	353-2243	123	Ross	901	885-4567
5	Alice	123-4456	126	Itero	901	346-1112
6	Waters	353-2243	126	Itero	901	346-1112
5	Alice	123-4456	127	Geoff	615	221-4456
6	Waters	353-2243	127	Geoff	615	221-4456

Natural Join

- Natural join performs the following tasks
  - Determine the common column(s) by looking for columns with identical names and compatible data types;
  - Select only the rows with common values in the common columns.
  - If there are no common column, return the relational product of the tables.
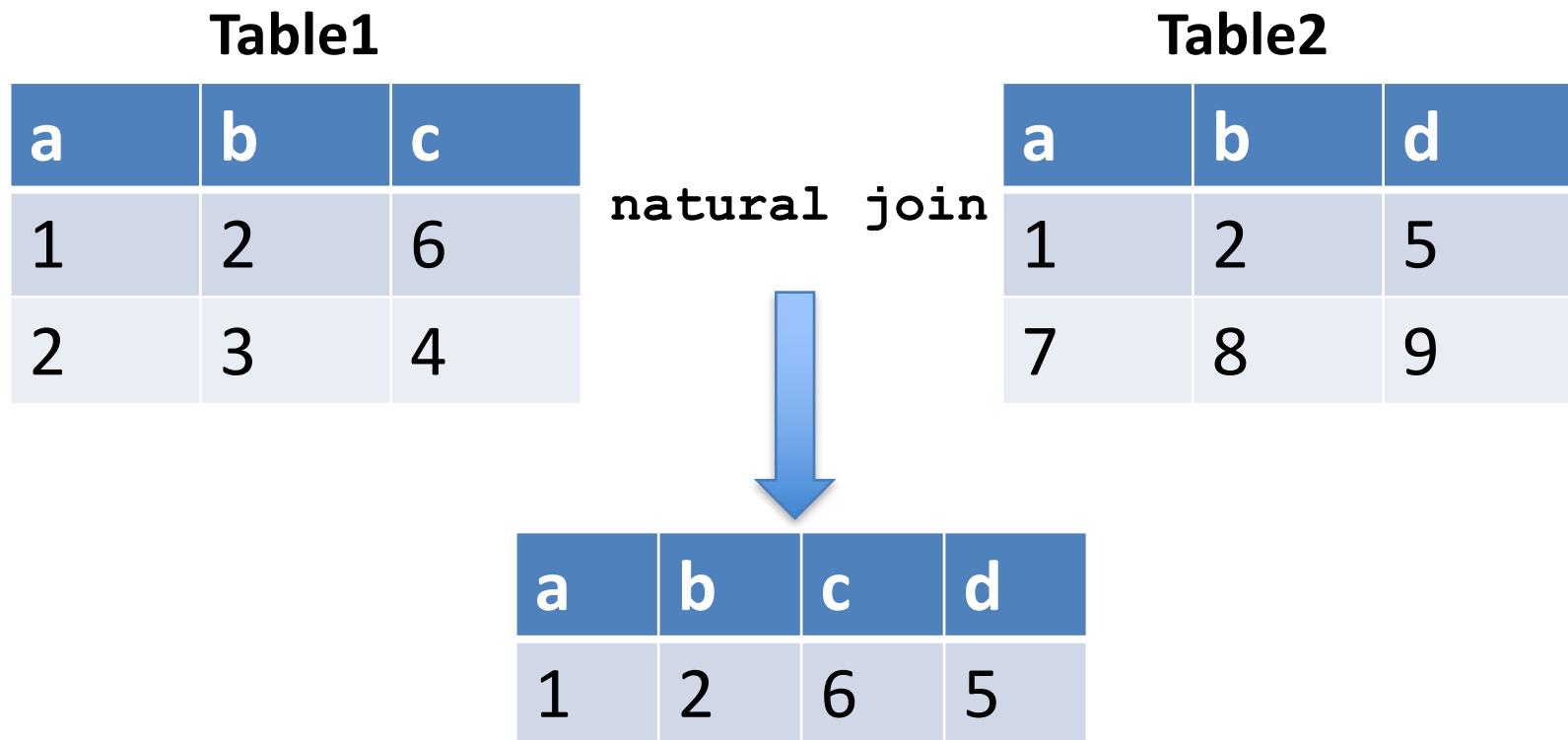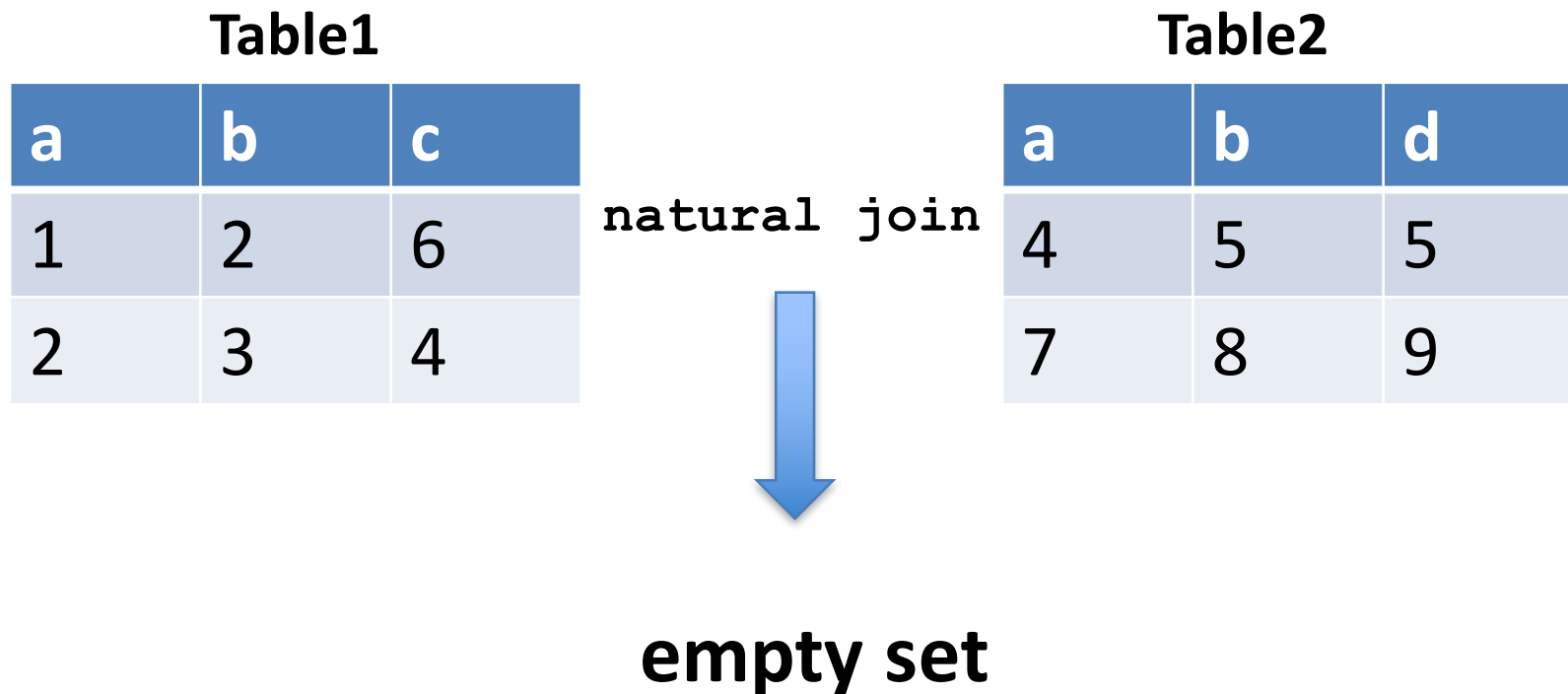
Natural Join

- One common column.

Table1

a	b	c
1	2	6
2	3	4

`natural join`

Table2

a	d	e
1	2	5
7	8	9

a	b	c	d	e
1	2	6	2	5

Natural Join (cont.)

- What if there are more than one common column?

Table1

a	b	c
1	2	6
2	3	4

natural join

Table2

a	b	d
1	2	5
7	8	9

a	b	c	d
1	2	6	5

Natural Join (cont.)

- What if there are no common values in the common columns?

Table1

a	b	c
1	2	6
2	3	4

natural join

Table2

a	b	d
4	5	5
7	8	9

empty set

Natural Join (cont.)

- What if there is no common columns?

Table1

a	b
1	2
2	3

natural join

Table2

c	d
4	5
7	8

a	b	c	d
1	2	4	5
1	2	7	8
2	3	4	5
2	3	7	8

Relational product

iClicker question 1

- How many rows will the following natural join return?

Table1

a	b
1	2
2	3

natural join

Table2

c	d
4	5
7	8
9	1

a. 4

b. 0

c. 9

d. 6

iClicker question 2

- How many rows will the following natural join return?

Table1

a	b
1	2
2	3

natural join

Table2

a	c
4	5
7	8
9	1

a. 4

b. 0

c. 9

d. 6

Structure Query Language

```
SELECT columns
FROM table
[WHERE condition]
[GROUP BY column]
[HAVING condition]
[ORDER BY column [ASC | DESC]]
[LIMIT count]
```

- Multiple tables: Natural Join Painter, Painting, Gallery.

```
SELECT * FROM PAINTING NATURAL JOIN PAINTER
NATURAL JOIN GALLERY;
```

- Where, GROUP BY, ORDER BY, etc…

```
SELECT COUNT(*), PAINTING.PTR_NUM FROM
PAINTING NATURAL JOIN PAINTER
GROUP BY PTR_NUM HAVING COUNT(*) > 2;
```

Derived table

```
SELECT columns FROM table
[WHERE ....]
[GROUP BY ....]
[ORDER BY ....]
```

a derived table

```
SELECT columns FROM table3
[WHERE ....]
[GROUP BY ....]
```

a real stored table

```
SELECT MIN(P_PRICE) FROM
(SELECT * FROM PRODUCT
WHERE P_PRICE > (SELECT AVG(P_PRICE) FROM PRODUCT)) AS T;
```

- Although join clause generates a "virtual" table, it's *not* a derived table because it's not a subquery.

INNER JOIN

- What if two columns in two tables do not have the same name, but indeed they are the same thing, or

- what if the columns you don't want to join happen to have the same name?

UserID	Name	Email
A00001	Alice	alice@
A00002	Bob	bob@
A00003	Doug	doug1@
A00004	Doug	doug2@

Facebook

UserID	Name	EmailAdd
1	Alice	alice@
2	Doug	doug2@
3	Emma	emma@
4	Fox	fox@

Instagram

INNER JOIN (cont.)

- An inner join returns all rows (of two tables) with <u>matching values</u> in the columns that are specified in the condition.

-
```
SELECT columns FROM table1
INNER JOIN table2 ON join_condition;
```

Table qualifier: use a dot to refer a table

```sql
SELECT * FROM PAINTER INNER JOIN PAINTING
ON PAINTER.PTR_NUM = PAINTING.PTR_NUM;
```

PTR_NUM	PTR_NAME	PTR_AREACODE	PTR_PHONE	PTNG_NUM	PTNG_TITLE	PTNG_PRICE	PTR_NUM	GAL_NUM
123	Ross	901	885-4567	1338	Dawn Thunder	245.50	123	5
123	Ross	901	885-4567	1339	A Faded Rose	6723.00	123	NULL
126	Itero	901	346-1112	1340	The Founders	567.99	126	6
123	Ross	901	885-4567	1341	Hasty Pudding Exit	145.50	123	NULL
126	Itero	901	346-1112	1342	Plastic Paradise	8328.99	126	6
127	Geoff	615	221-4456	1343	Roamin'	785.00	127	6
127	Geoff	615	221-4456	1344	Wild Waters	999.00	127	5
123	Ross	901	885-4567	1345	Stuff 'n Such 'n Some	9800.00	123	5

Natural Join and Inner join with the same condition will return exactly the same rows, but more columns.

table qualifier

```sql
SELECT * FROM PAINTER INNER JOIN PAINTING
ON PAINTER.PTR_NUM = PAINTING.PTR_NUM;
```
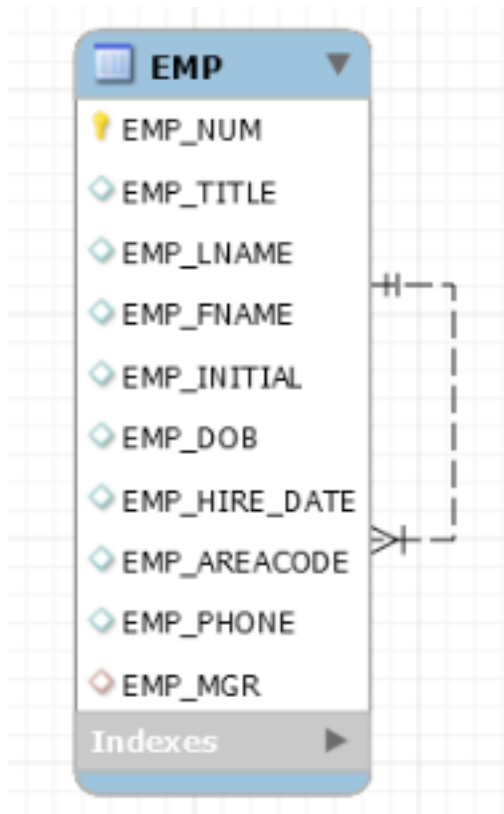
table alias

```sql
SELECT * FROM PAINTER AS P INNER JOIN PAINTING AS Q
ON P.PTR_NUM = Q.PTR_NUM;
```

Table alias

Table join itself

List all employees with their managers' names

EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_MGR
100	Kolmycz	George	NULL
102	VanDam	Rhett	100
103	Jones	Anne	100
104	Lange	John	105
105	Williams	Robert	NULL
106	Smith	Jeanine	105
107	Diante	Jorge	105
108	Wiesenbach	Paul	NULL
109	Smith	George	108
110	Genkazi	Leighla	108
111	Washington	Rupert	105
112	Johnson	Edward	100
113	Smythe	Melanie	105
114	Brandon	Marie	108
115	Saranda	Hermine	105
116	Smith	George	108
101	Lewis	Rhonda	100

EMP

- EMP_NUM
- EMP_TITLE
- EMP_LNAME
- EMP_FNAME
- EMP_INITIAL
- EMP_DOB
- EMP_HIRE_DATE
- EMP_AREACODE
- EMP_PHONE
- EMP_MGR

Indexes

Table join itself

EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_NUM	EMP_LNAME	EMP_FNAME
102	VanDam	Rhett	100	Kolmycz	George
103	Jones	Anne	100	Kolmycz	George
112	Johnson	Edward	100	Kolmycz	George
101	Lewis	Rhonda	100	Kolmycz	George
104	Lange	John	105	Williams	Robert
106	Smith	Jeanine	105	Williams	Robert
107	Diante	Jorge	105	Williams	Robert
111	Washington	Rupert	105	Williams	Robert
113	Smythe	Melanie	105	Williams	Robert
115	Saranda	Hermine	105	Williams	Robert
109	Smith	George	108	Wiesenbach	Paul
110	Genkazi	Leighla	108	Wiesenbach	Paul
114	Brandon	Marie	108	Wiesenbach	Paul
116	Smith	George	108	Wiesenbach	Paul

```
SELECT M.EMP_NUM, M.EMP_LNAME, M.EMP_FNAME,
N.EMP_NUM, N.EMP_LNAME, N.EMP_FNAME
FROM EMP M INNER JOIN EMP N ON M.EMP_MGR = N.EMP_NUM;
```

- Alias is needed when a table is joined to itself.

Difference between NATURAL JOIN and INNER JOIN

- "INNER JOIN is often used in real life."

- NATRUAL JOIN is an inner join involving all columns that happen to have the same name.

- They return different number of columns. (Natural join eliminates duplicated columns.)

Summary

- JOIN, a very powerful command in relational database.

- Natural Join

- Inner Join