CS245: Databases

SQL

Vijaya saradhi

Department of Computer Science and Engineering
Indian Institute of Technology Guwahati

Intersection & Difference - limitations using IN

- Duplicate rows cause problems when used IN predicate
- We will provide two to three alternate ways of implementing Intersect, Difference
- The implementation will be discussed after we go through (1) join operator (2) nested queries and (3) querying on the go method

Reading Data From Tables

- Set/Bag operations
  - Union of two tables
  - Intersection between two tables
  - Difference of two tables
  - Cross product of two tables
  - Joining two tables
- Join
  - Natural join
  - Inner join (theta join)
  - Left outer join
  - Right outer join
  - Full outer join

Cartesian (Cross) product

Definition

$$\mathrm{A} \times \mathrm{B} = \{ (\mathrm{a}, \mathrm{b}) \mid a \in \mathrm{A} \textbf{ AND } b \in \mathrm{B}\}$$

Example

Table R_1

A	B
1	2
3	4

Table S_1

B	C	D
2	5	6
4	7	8
9	10	11

Table $R_1 \times S_1$

R_1.A	R_1.B	S_1.B	S_1.C	S_1.D
1	2	2	5	6
1	2	4	7	8
1	2	9	10	11
3	4	2	5	6
3	4	4	7	8
3	4	9	10	11

Example Database

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

Boats		
bid	bname	color
101	Interlake	blue
102	Interlanke	red
103	Clipper	green
104	Marine	red

Cross Join Execution - 01/10

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

×

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

SELECT * FROM Sailors CROSS JOIN Reserves						
sid	sname	rating	age	sid	bid	day
22	Dustin	7	45.0	22	101	10-Oct-2019
22	Dustin	7	45.0	22	102	10-Oct-2019
22	Dustin	7	45.0	22	103	08-Oct-2019
22	Dustin	7	45.0	22	104	07-Oct-2019
22	Dustin	7	45.0	31	102	10-Nov-2019
22	Dustin	7	45.0	31	103	06-Nov-2019
22	Dustin	7	45.0	31	104	12-Nov-2019
22	Dustin	7	45.0	64	101	05-Sep-2019
22	Dustin	7	45.0	64	102	08-Sep-2019
22	Dustin	7	45.0	74	103	08-Sep-2019

Cross Join Execution - 02/10

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

×

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

SELECT * FROM Sailors CROSS JOIN Reserves						
sid	sname	rating	age	sid	bid	day
29	Brutus	1	33.0	22	101	10-Oct-2019
29	Brutus	1	33.0	22	102	10-Oct-2019
29	Brutus	1	33.0	22	103	08-Oct-2019
29	Brutus	1	33.0	22	104	07-Oct-2019
29	Brutus	1	33.0	31	102	10-Nov-2019
29	Brutus	1	33.0	31	103	06-Nov-2019
29	Brutus	1	33.0	31	104	12-Nov-2019
29	Brutus	1	33.0	64	101	05-Sep-2019
29	Brutus	1	33.0	64	102	08-Sep-2019
29	Brutus	1	33.0	74	103	08-Sep-2019

Cross Join Execution - 03/10

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

$\times$

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

SELECT * FROM Sailors CROSS JOIN Reserves						
sid	sname	rating	age	sid	bid	day
31	Lubber	8	55.5	22	101	10-Oct-2019
31	Lubber	8	55.5	22	102	10-Oct-2019
31	Lubber	8	55.5	22	103	08-Oct-2019
31	Lubber	8	55.5	22	104	07-Oct-2019
31	Lubber	8	55.5	31	102	10-Nov-2019
31	Lubber	8	55.5	31	103	06-Nov-2019
31	Lubber	8	55.5	31	104	12-Nov-2019
31	Lubber	8	55.5	64	101	05-Sep-2019
31	Lubber	8	55.5	64	102	08-Sep-2019
31	Lubber	8	55.5	74	103	08-Sep-2019

Cross Join Execution - 10/10

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

×

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

SELECT * FROM Sailors CROSS JOIN Reserves						
sid	sname	rating	age	sid	bid	day
95	Bob	3	63.5	22	101	10-Oct-2019
95	Bob	3	63.5	22	102	10-Oct-2019
95	Bob	3	63.5	22	103	08-Oct-2019
95	Bob	3	63.5	22	104	07-Oct-2019
95	Bob	3	63.5	31	102	10-Nov-2019
95	Bob	3	63.5	31	103	06-Nov-2019
95	Bob	3	63.5	31	104	12-Nov-2019
95	Bob	3	63.5	64	101	05-Sep-2019
95	Bob	3	63.5	64	102	08-Sep-2019
95	Bob	3	63.5	74	103	08-Sep-2019

Natural Join Execution - 01/10

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

NATURAL JOIN

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

Join condition: Sailors.sid = Reserves.sid (join identical column names from both tables)

SELECT * FROM Sailors NATURAL JOIN Reserves						
sid	sname	rating	age	sid	bid	day
22	Dustin	7	45.0	22	101	10-Oct-2019
22	Dustin	7	45.0	22	102	10-Oct-2019
22	Dustin	7	45.0	22	103	08-Oct-2019
22	Dustin	7	45.0	22	104	07-Oct-2019
31	Lubber	8	55.5	31	102	10-Nov-2019
31	Lubber	8	55.5	31	103	06-Nov-2019
31	Lubber	8	55.5	31	104	12-Nov-2019
64	Horatio	7	35.0	64	101	05-Sep-2019
64	Horatio	7	35.0	64	102	08-Sep-2019
74	Horatio	9	35.0	74	103	08-Sep-2019

Theta Join

Natural Join

- Join condition is implicit
- Join condition is on attributes whose name is identical
- Attributes are always joined by equality condition

Theta Join

- Specify the condition on which on row from each table to be joined by using ON clause
- Specify filtering condition by using WHERE clause

Theta Join Execution - 01/23

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

JOIN ON Sailors.sid < Reserves.sid

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

Join condition: Sailors.sid < Reserves.sid (join on specified condition)

SELECT * FROM Sailors JOIN Reserves ON Sailors.sid < Reserves						
sid	sname	rating	age	sid	bid	day
22	Dustin	7	45.0	31	102	10-Nov-2019
22	Dustin	7	45.0	31	103	06-Nov-2019
22	Dustin	7	45.0	31	104	12-Nov-2019
22	Dustin	7	45.0	64	101	05-Sep-2019
22	Dustin	7	45.0	64	102	08-Sep-2019
22	Dustin	7	45.0	74	103	08-Sep-2019
29	Brutus	1	33.0	31	102	10-Nov-2019
29	Brutus	1	33.0	31	103	06-Nov-2019
29	Brutus	1	33.0	31	104	12-Nov-2019
29	Brutus	1	33.0	64	101	05-Sep-2019
29	Brutus	1	33.0	64	102	09-Sep-2019
29	Brutus	1	33.0	74	103	08-Sep-2019
31	Lubber	8	55.5	64	101	05-Sep-2019
31	Lubber	8	55.5	64	102	08-Sep-2019
31	Lubber	8	55.5	74	103	08-Sep-2019
32	Andy	8	25.5	64	101	05-Sep-2019
32	Andy	8	25.5	64	102	08-Sep-2019
32	Andy	8	25.5	74	103	08-Sep-2019
58	Rusty	10	35.0	64	101	05-Sep-2019
58	Rusty	10	35.0	64	102	08-Sep-2019
58	Rusty	10	35.0	74	103	08-Sep-2019
64	Horatio	7	35.0	74	103	08-Sep-2019
71	Zobra	10	16.0	74	103	08-Sep-2019

Theta Join Execution - 02/04

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

JOIN ON Sailors.sid $<$ Reserves.sid where Sailros.rating ≥ 9

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

Join condition: Sailors.sid $<$ Reserves.sid (join on specified condition) where clause is Sailors.rating ≥ 9

SELECT * FROM Sailors JOIN Reserves ON Sailors.sid $<$ Reserves where Sailors.rating ≥ 9

sid	sname	rating	age	sid	bid	day
58	Rusty	10	35.0	64	101	05-Sep-2019
58	Rusty	10	35.0	64	103	08-Sep-2019
58	Rusty	10	35.0	74	103	08-Sep-2019
71	Zobra	10	16.0	74	103	08-Sep-2019

(Left/Right) Outer Join

Outer join

Definition: In addition to the Natural join, add any dangling tuples from R or S.
Dangling tuples: The tuples that did not meet the Natural join criteria

Left Outer Join Execution - 01 (Left-outer-join) /15

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

LEFT OUTER JOIN

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

Join condition: Sailors.sid = Reserves.sid (join identical column names from both tables)

SELECT * FROM Sailors LEFT OUTER JOIN Reserves ON Sailors.sid = Reserves.sid

sid	sname	rating	age	sid	bid	day
22	Dustin	7	45.0	22	101	10-Oct-2019
22	Dustin	7	45.0	22	102	10-Oct-2019
22	Dustin	7	45.0	22	103	08-Oct-2019
22	Dustin	7	45.0	22	104	07-Oct-2019
29	Brutus	1	33.0	⊥	⊥	⊥
31	Lubber	8	55.5	31	102	10-Nov-2019
31	Lubber	8	55.5	31	103	06-Nov-2019
31	Lubber	8	55.5	31	104	12-Nov-2019
32	Andy	8	25.5	⊥	⊥	⊥
58	Rusty	10	35.0	⊥	⊥	⊥
64	Horatio	7	35.0	64	101	05-Sep-2019
64	Horatio	7	35.0	64	102	08-Sep-2019
71	Zorba	10	16.0	⊥	⊥	⊥
74	Horatio	9	35.0	74	103	08-Sep-2019
85	Art	3	25.5	⊥	⊥	⊥
95	Bob	3	63.5	⊥	⊥	⊥

Left Outer Join - Example

U		
A	B	C
1	2	3
4	5	6
7	8	9

LEFT OUTER JOIN

V		
B	C	D
2	3	10
2	3	11
6	7	12

SELECT * FROM U LEFT OUTER JOIN V					
A	B	C	B	C	D
1	2	3	2	3	10
1	2	3	2	3	11
4	5	6	$\perp$	$\perp$	$\perp$
7	8	9	$\perp$	$\perp$	$\perp$

Right Outer Join - Example

U		
A	B	C
1	2	3
4	5	6
7	8	9

RIGHT OUTER JOIN

V		
B	C	D
2	3	10
2	3	11
6	7	12

SELECT * FROM U RIGHT OUTER JOIN V

A	B	C	B	C	D
1	2	3	2	3	10
1	2	3	2	3	11
$\perp$	$\perp$	$\perp$	6	7	12

Full Outer Join

Full outer join

TableA left outer join TableB

$\cup$

TableA right outer join TableB

FULL Outer Join

FULL Outer Join - Example (Not in MySQL)

U		
A	B	C
1	2	3
4	5	6
7	8	9

FULL OUTER JOIN

V		
B	C	D
2	3	10
2	3	11
6	7	12

SELECT * FROM U LEFT OUTER JOIN V					
A	B	C	B	C	D
1	2	3	2	3	10
1	2	3	2	3	11
4	5	6	$\perp$	$\perp$	$\perp$
7	8	9	$\perp$	$\perp$	$\perp$
$\perp$	$\perp$	$\perp$	6	7	12