

SQL Survival Guide

Joins

Office of Quality and Patient Safety

May 18, 2015

Outline

- 1 What Is A Join Clause?
- 2 Types of SQL Joins
- 3 Examples
- 4 Questions?

What Is A Join Clause?

What Is A Join Clause?

- Combines records from two or more tables.
- Returns a set of records which can be saved as a table or used as is.
- Joins are probably the hardest core SQL skill to master.

Why Do We Have To Do This?

- Normalization:
 - There aren't enough Database Administrators to go around.
 - Eliminates duplication of data.
 - Makes it easier to maintain data integrity, at the cost of speed.
- DOH DataMart is NOT consistently normalized:
 - Normalization deserves its own presentation.
 - Combines OLTP with Transactional Structures.
 - Sometimes easier. Sometimes Harder
 - Denormalized data == Mistakes.
 - Use normalized data structure unless data is COMPLETELY static.

Types of Joins

There are 5 ANSI joins, plus one special case that deserves additional discussion:

- ① Cross Join
- ② Inner Join
- ③ Left Outer Join
- ④ Right Outer Join
- ⑤ Full Outer Join
- ⑥ Self Join

Data

- We need example data.
- Two tables of data!
 - Yes, you can join more than two tables at a time.
 - No, we won't do that today.
- Keep track of the next two slides.
- Code provided:
 - sql: All SQL code needed to build tables and run queries. (Tested in Vertica.)
 - data: Prepared sqlite db files. Can be used by any JDBC or ODBC compliant front-end.

Table: Departments

DEPT_ID	DEPT_NAME	DEPT_FLOOR
31	Sales	1
33	Engineering	3
34	Clerical	2
35	Marketing	3

Note: Aliased as 'empl'.

Table: Employees

EID	DEPT_ID	LAST_NAME	FIRST_NAME	GENDER
1	31	Rafferty	Gerry	M
3	33	Jones	Jon	M
5	33	Heisenberg	Werner	M
7	34	Robinson	Elizabeth	F
9	34	Smith	Jefferson	M
11	NULL	Williams	Serena	F

Note: Aliased as 'empl'.

SQL Clauses

The following protected SQL commands START a clause:

- SELECT
- FROM
- WHERE
- HAVING

SQL Clauses

The following protected SQL commands START a clause:

- SELECT
- FROM
- WHERE
- HAVING

Style:

- Align SQL Clauses
- Predicates should be indented consistently.
- Feel free to ask me why I structure SQL queries the way I do.

Cross Join: Discussion

- Returns the Cartesian product of the tables in the FROM clause.
- This is bad.
- These tables are tiny, so it doesn't really matter.
- But if you do this on a table like RECIP_PROFILE, you are going to get a nasty call from Samir.

Explicit Cross Join: Example

Question: How many rows will the following query return?

```
1  SELECT *  
2  FROM EMPLOYEES empl cross join DEPARTMENTS dept  
3  ;
```

Source: <https://github.com/Choens/sql-survival-guide/blob/master/sql/04-joins/cross-joins.sql>

Explicit Cross Join: Example

Question: How many rows will the following query return?

```
1  SELECT *  
2  FROM EMPLOYEES empl cross join DEPARTMENTS dept  
3  ;
```

Source: <https://github.com/Choens/sql-survival-guide/blob/master/sql/04-joins/cross-joins.sql>

Answer: 24 rows!

Cross Join: Results

EID	DEPT_ID	LAST_NAME	FIRST_NAME	GENDER	DEPT_ID	DEPT_NAME	DEPT_FLOOR
1	31	Rafferty	Gerry	M	31	Sales	1
1	31	Rafferty	Gerry	M	33	Engineering	3
1	31	Rafferty	Gerry	M	34	Clerical	2
1	31	Rafferty	Gerry	M	35	Marketing	3
3	33	Jones	Jon	M	31	Sales	1
3	33	Jones	Jon	M	33	Engineering	3
3	33	Jones	Jon	M	34	Clerical	2
3	33	Jones	Jon	M	35	Marketing	3
5	33	Heisenberg	Werner	M	31	Sales	1
5	33	Heisenberg	Werner	M	33	Engineering	3
5	33	Heisenberg	Werner	M	34	Clerical	2
5	33	Heisenberg	Werner	M	35	Marketing	3
7	34	Robinson	Elizabeth	F	31	Sales	1
7	34	Robinson	Elizabeth	F	33	Engineering	3
7	34	Robinson	Elizabeth	F	34	Clerical	2
7	34	Robinson	Elizabeth	F	35	Marketing	3
9	34	Smith	Jefferson	M	31	Sales	1
9	34	Smith	Jefferson	M	33	Engineering	3
9	34	Smith	Jefferson	M	34	Clerical	2
9	34	Smith	Jefferson	M	35	Marketing	3
11	[NULL]	Williams	Serena	F	31	Sales	1
11	[NULL]	Williams	Serena	F	33	Engineering	3
11	[NULL]	Williams	Serena	F	34	Clerical	2
11	[NULL]	Williams	Serena	F	35	Marketing	3

Implicit Cross Join: Example

```
1  SELECT *  
2  FROM EMPLOYEES empl, DEPARTMENTS dept  
3  ;
```

Source: <https://github.com/Choens/sql-survival-guide/blob/master/sql/04-joins/cross-joins.sql>

- Also returns 24 rows.
- You can do the same thing implicitly. (Don't!)
- What does this look like?

Cross Join: Results

- Cross Joins Are Dangerous!
- RECIP_PROFILE == Phone Call From Samir.
- IMPLICIT CROSS JOIN Looks Like IMPLICIT INNER JOIN.

Inner Join: Discussion

```
1  select * from EMPLOYEES;
```

Inner Join: Example

Left Outer Join: Discussion

TODO

Left Outer Join: Example

TODO

Right Outer Join: Discussion

TODO

Right Outer Join: Example

TODO

Full Outer Join: Discussion

TODO

Full Outer Join: Example

TODO

Self Join: Discussion

TODO

Self Join: Example

TODO

Additional Information

- [https://en.wikipedia.org/wiki/Join_\(SQL\)](https://en.wikipedia.org/wiki/Join_(SQL))

Questions?