

## COMPSCI 105: Lecture #21 Joins

©2015-2019 Dr. William T. Verts

## JOINS

- Something spreadsheets can't do.
- Synthesize new tables from two or more source tables.
- Source tables must have a field in common (doesn't have to have the same name, but should have compatible data types).
- Result may contain more records than either source table, or may be empty.

©2015 Dr. William T. Verts

## Types of Joins

- Inner Join (intersection of tables)
- Outer Join
  - Left Outer Join
  - Right Outer Join
  - Full Outer Join (union of tables)

©2015 Dr. William T. Verts

## Source Tables (ID in common)

PEOPLE			PAYMENTS		
	ID	NAME		ID	SALARY
1	12345	Fred	1	60233	\$30,000
2	72401	Joe	2	54321	\$20,000
3	22222	Mary	3	11111	\$40,000
4	54321	Sam	4	97330	\$50,000
5	20202	Martha	5	72401	\$35,000
6	11111	Bob	6	13333	\$45,000
7	47904	Tom			

©2015-2019 Dr. William T. Verts

## Inner Join on ID Field (Intersection)

### Inner Join

	ID	NAME	SALARY
1	11111	Bob	\$40,000
2	54321	Sam	\$20,000
3	72401	Joe	\$35,000

©2015 Dr. William T. Verts

## Inner Join

- Contains only records where join-field in one table matches join-field in the other table.
- In this case, there are only three matches. Those three results will be filled in with data from both tables.

©2015 Dr. William T. Verts

## Left Outer Join on ID Field

*Left Outer Join*

	ID	NAME	SALARY
1	11111	Bob	\$40,000
2	12345	Fred	
3	20202	Martha	
4	22222	Mary	
5	47904	Tom	
6	54321	Sam	\$20,000
7	72401	Joe	\$35,000

©2015 Dr. William T. Verts

## Left Outer Join

- Contains all data from the left table.
- The matches with the right table will have their records completely filled in.
- Answer records from the left with no match in the right will have "holes".

©2015 Dr. William T. Verts

## Right Outer Join on ID Field

*Right Outer Join*

	ID	NAME	SALARY
1	11111	Bob	\$40,000
2	13333		\$45,000
3	54321	Sam	\$20,000
4	60233		\$30,000
5	72401	Joe	\$35,000
6	97330		\$50,000

©2015 Dr. William T. Verts

## Right Outer Join

- Contains all data from the right table.
- The matches with the left table will have their records completely filled in.
- Answer records from the right with no match in the left will have "holes".

©2015 Dr. William T. Verts

## Full Outer Join on ID Field (Union)

*Full Outer Join*

	ID	NAME	SALARY
1	11111	Bob	\$40,000
2	12345	Fred	
3	13333		\$45,000
4	20202	Martha	
5	22222	Mary	
6	47904	Tom	
7	54321	Sam	\$20,000
8	60233		\$30,000
9	72401	Joe	\$35,000
10	97330		\$50,000

©2015 Dr. William T. Verts

## Full Outer Join

- All records from both source tables will be in the answer.
- Those that match will be completely filled in (and appear only once).
- All other records will contain "holes".

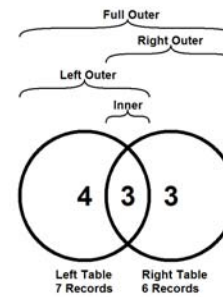
©2015 Dr. William T. Verts

### For our example

- There are 7 records in the left table.
- There are 6 records in the right table.
- There are 3 matches (the inner join).
- The left outer join will contain 7 records, 3 filled in and 4 with holes.
- The right outer join will contain 6 records, 3 filled in and 3 with holes.
- The full outer join contains 4 (from left) + 3 (inner) + 3 (from right) = 10 records.

©2015 Dr. William T. Verts

### Records in Result of Joins



©2015 Dr. William T. Verts

### Final Concepts

- Simple Joins return either the *intersection* or the *union* (in several ways) of two tables.
- However, we have assumed all along in this lecture that for any record in one table there is at most one match in the other table.
- What if that isn't true? What if there could be multiple matches, in either direction, between two tables? How might indexing affect join speed? Next lecture!

©2019 Dr. William T. Verts

©2019 Dr. William T. Verts