

# Difference between INNER JOIN and LEFT SEMI JOIN

Asked 7 years, 5 months ago    Active 4 days ago    Viewed 134k times

What is the difference between an `INNER JOIN` and `LEFT SEMI JOIN` ?

92 In the scenario below, why am I getting two different results?

The `INNER JOIN` result set is a lot larger. Can someone explain? I am trying to get the names within `table_1` that only appear in `table_2` .

34


```
SELECT name
FROM table_1 a
INNER JOIN table_2 b ON a.name=b.name
```

```
SELECT name
FROM table_1 a
LEFT SEMI JOIN table_2 b ON (a.name=b.name)
```


sql   hql   hive

Share   Follow

edited Feb 12 '14 at 20:24

 **Shiva**  
18.7k   13   75   104

asked Feb 12 '14 at 20:21

 **user3023355**  
947   1   7   6

2 The inner join will achieve your goal. I had never heard of a semi join until I saw this question. – Dan Bracuk Feb 12 '14 at 20:25

The `left semi join` should be returning more rows than the `inner join` . – Gordon Linoff Feb 12 '14 at 20:29

1 The `inner join` will return data only if there is a match between both tables. The `left join` will return data from the first table regardless if a matching record is found in the second table. – j03z Feb 12 '14 at 20:34

11 @GordonLinoff not necessarily, a `LEFT SEMI JOIN` will only return one row from the left, even if there are multiple matches in the right. An `INNER JOIN` will return multiple rows if there are multiple matching on the right. – D Stanley Feb 12 '14 at 20:48

1 @j03z that can't be correct. If the purpose of the left hemi-join is 1) to return only the information in the left table (as others have said) and 2) to return rows from teh left table regardless of match (as I think you say) then that is just the original left table -- no join is necessary to accomplish that. I think others must be correct that the left hemi-join 1) only returns columns from the left table, 2) only returns rows that have a match in the right table, and 3) will return a single row from the left for one or more matches. – Carl G Sep 21 '15 at 20:20

4 Answers

Active	Oldest	Votes
--------	--------	-------

138 An `INNER JOIN` can return data from the columns from both tables, and can duplicate values of records on either side have more than one match. A `LEFT SEMI JOIN` can only return columns from the left-hand table,

and yields one of each record from the left-hand table where there is one or more matches in the right-hand table (regardless of the number of matches). It's equivalent to (in standard SQL):

```
SELECT name
FROM table_1 a
WHERE EXISTS(
  SELECT * FROM table_2 b WHERE (a.name=b.name))
```

If there are *multiple* matching rows in the right-hand column, an `INNER JOIN` will return one row for each match on the right table, while a `LEFT SEMI JOIN` only returns the rows from the left table, regardless of the number of matching rows on the right side. That's why you're seeing a different number of rows in your result.

I am trying to get the names within table\_1 that only appear in table\_2.

Then a `LEFT SEMI JOIN` is the appropriate query to use.

Share Follow

edited Oct 6 '20 at 14:39

answered Feb 12 '14 at 20:28



[D Stanley](#)

140k

11

156

220

Is there really such a thing as a `LEFT SEMI JOIN` ? Isn't it just a `SEMI JOIN` ? There's no sense to a `RIGHT SEMI JOIN` , is there? – [ErikE](#) Jul 18 '14 at 23:37

In [Hive](#), yes. – [D Stanley](#) Jul 20 '14 at 2:24

- 1 great answer just what i was looking for. i would phrase the answer more accurately : "...an `INNER JOIN` will return one row for each matching **row of the right table**, while a `LEFT SEMI JOIN`... – [Barak1731475](#) Apr 20 '15 at 18:10 ✎
- 2 The opposite of this is a `LEFT ANTI JOIN` that filters out the data from the right table in the left table according to a key. Thought I'd leave this nugget here for someone who might be looking! – [shantanusinghal](#) Oct 29 '17 at 19:55

Suppose there are 2 tables TableA and TableB with only 2 columns (Id, Data) and following data:

77

**TableA:**

Id	Data
1	DataA11
1	DataA12
1	DataA13
2	DataA21
3	DataA31

**TableB:**

Id	Data
----	------

1	DataB11
2	DataB21
2	DataB22
2	DataB23
4	DataB41

**Inner Join** on column `Id` will return columns from both the tables and only the matching records:

Id	Data	Id	Data
1	DataA11	1	DataB11
1	DataA12	1	DataB11
1	DataA13	1	DataB11
2	DataA21	2	DataB21
2	DataA21	2	DataB22
2	DataA21	2	DataB23

**Left Join** (or Left Outer join) on column `Id` will return columns from both the tables and matching records with records from left table (Null values from right table):

Id	Data	Id	Data
1	DataA11	1	DataB11
1	DataA12	1	DataB11
1	DataA13	1	DataB11
2	DataA21	2	DataB21
2	DataA21	2	DataB22
2	DataA21	2	DataB23
3	DataA31		

**Right Join** (or Right Outer join) on column `Id` will return columns from both the tables and matching records with records from right table (Null values from left table):

Id	Data	Id	Data
1	DataA11	1	DataB11
1	DataA12	1	DataB11
1	DataA13	1	DataB11
2	DataA21	2	DataB21
2	DataA21	2	DataB22
2	DataA21	2	DataB23

		4	DataB41
--	--	---	---------

**Full Outer Join** on column `id` will return columns from both the tables and matching records with records from left table (Null values from right table) and records from right table (Null values from left table):

Id	Data	Id	Data
-			
1	DataA11	1	DataB11
1	DataA12	1	DataB11
1	DataA13	1	DataB11
2	DataA21	2	DataB21
2	DataA21	2	DataB22
2	DataA21	2	DataB23
3	DataA31		
		4	DataB41

**Left Semi Join** on column `id` will return columns only from left table and matching records only from left table:

Id	Data
1	DataA11
1	DataA12
1	DataA13
2	DataA21

Share Follow

edited Sep 19 '18 at 23:11



Amit Naidu

2,243 2 21 31

answered Mar 29 '17 at 17:01



Abhishek Bansal

1,119 6 8

I used to call this as "LEFT INNER Join". – [Anshul Joshi](#) Mar 31 '17 at 8:41

DISTINCT of A.\* from INNER JOIN result is equivalent to LEFT SEMI JOIN. – [Teja](#) Sep 28 '18 at 21:46

5 Distinct does not sound safe, suppose A contains two identical records. – [Dennis Jaheruddin](#) Jan 27 '19 at 12:33

Even if the result ends up being the same, using DISTINCT might have a more expensive plan compared to EXISTS – [manotheshark](#) Mar 19 '20 at 18:53

Tried in Hive and got the below output

33

table1

1,wqe,chennai,india

2,stu,salem,india

3,mia,bangalore,india

4,yepie,newyork,USA

## table2

1,wqe,chennai,india

2,stu,salem,india

3,mia,bangalore,india

5,chapie,Los angels,USA

## Inner Join

```
SELECT * FROM table1 INNER JOIN table2 ON (table1.id = table2.id);
```

1 wqe chennai india 1 wqe chennai india

2 stu salem india 2 stu salem india

3 mia bangalore india 3 mia bangalore india

## Left Join

```
SELECT * FROM table1 LEFT JOIN table2 ON (table1.id = table2.id);
```

1 wqe chennai india 1 wqe chennai india

2 stu salem india 2 stu salem india

3 mia bangalore india 3 mia bangalore india

4 yepie newyork USA NULL NULL NULL NULL

## Left Semi Join

```
SELECT * FROM table1 LEFT SEMI JOIN table2 ON (table1.id = table2.id);
```

1 wqe chennai india

2 stu salem india

3 mia bangalore india

note: Only records in left table are displayed whereas for Left Join both the table records displayed

Share Follow

edited Jun 20 '20 at 9:12

answered Apr 12 '16 at 8:14

Community

11

Kumar

770816



All above answers are correct. However in practice, it helps to associate the mental model of a `filter` when imagining LEFT SEMI JOIN.

0




The answer is a subset of rows from LEFT table, which have a match in RIGHT TABLE.



Share Follow

answered Jul 18 at 0:08

dsculptor

417



**Highly active question.** Earn 10 reputation (not counting the [association bonus](#)) in order to answer this question. The reputation requirement helps protect this question from spam and non-answer activity.