Self Inner Join

This lesson discusses how to join a table with itself.

Self Inner Join

The simplest join one can use is the inner join. Rows from two tables are joined together using a common column between them.

Syntax

SELECT *

FROM **table1**

INNER JOIN **table1**

ON <**join condition**>;

Connect to the terminal below by clicking in the widget. Once connected, the command line prompt will show up. Enter or copy and paste the command **./DataJek/Lessons/25lesson.sh** and wait for the MySQL prompt to start-up.

```
-- The lesson queries are reproduced below for convenient copy/paste into the terminal.

-- Query 1
SELECT * FROM Actors a INNER JOIN Actors b;

-- Query 2
SELECT * FROM Actors a INNER JOIN Actors b USING(FirstName);

-- Query 3
SELECT * FROM Actors a INNER JOIN Actors b USING(NetWorthInMillions);
```

```
SELECT * FROM Actors a INNER JOIN Actors b USING(NetWorthInMillions);
```

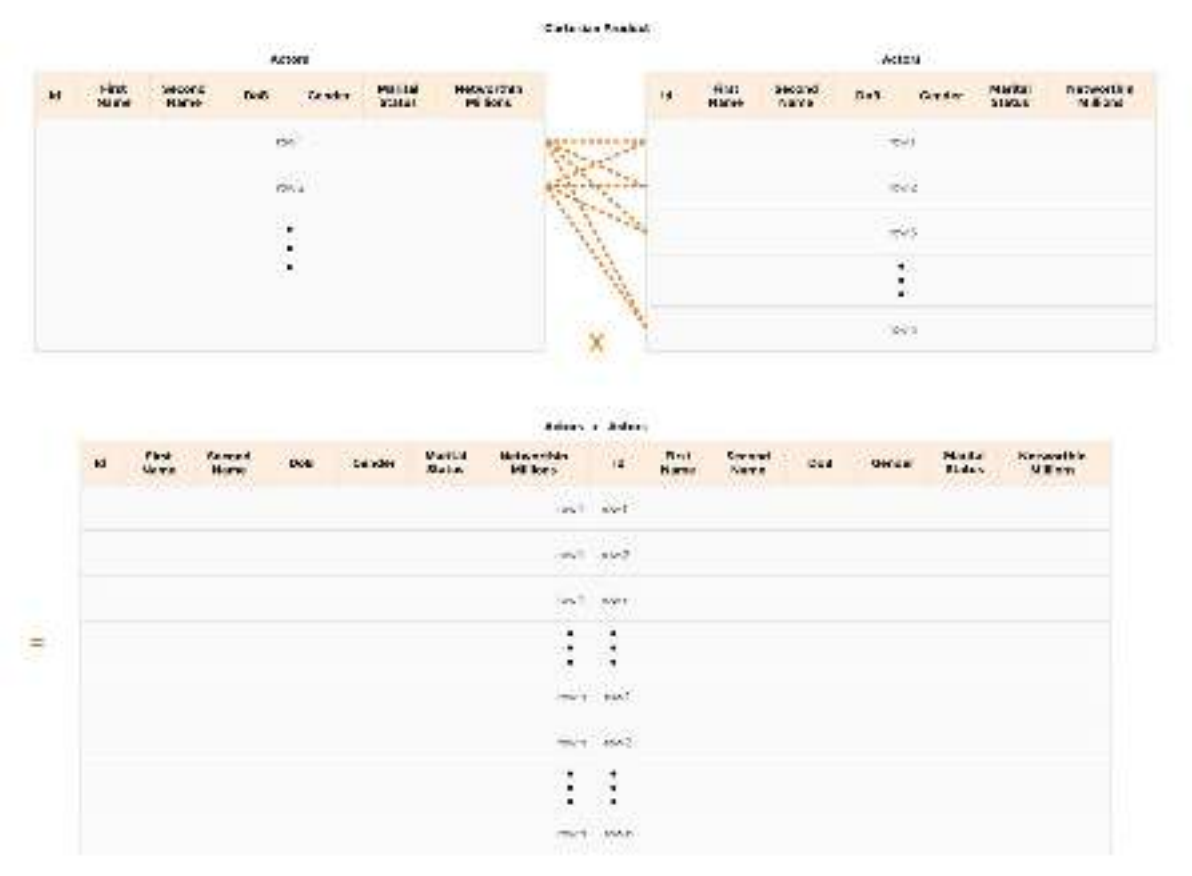
Terminal

1. It may come as a surprise, but we can also join a table with itself. However, we need to use aliases as the **INNER JOIN** clause requires the two tables to be unique.

   ```
   SELECT * FROM Actors a INNER JOIN Actors b;
   ```

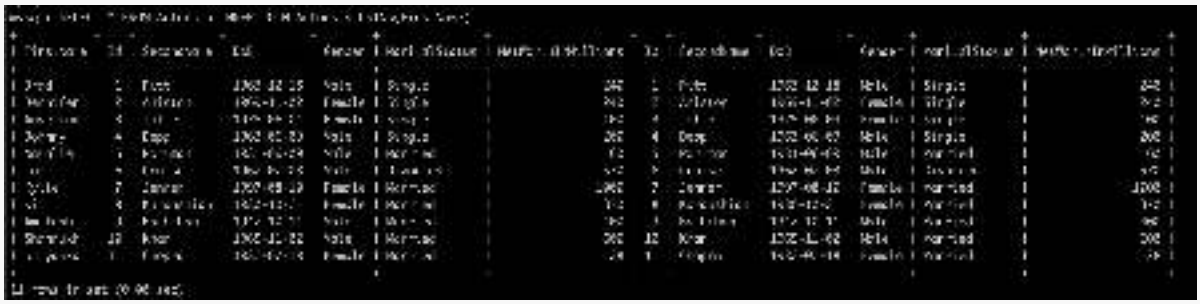
   If you run the above query, the result will be a cartesian product, i.e., each row will join with every other row of the second table. The total number of rows in the resulting table will be 121 because the table has 11 rows. The count for the inner join is shown below:

   ```
   mysql> SELECT COUNT(*) FROM Actors a INNER JOIN Actors b;
   +----------+
   | COUNT(*) |
   +----------+
   |      121 |
   +----------+
   1 row in set (0.00 sec)
   ```



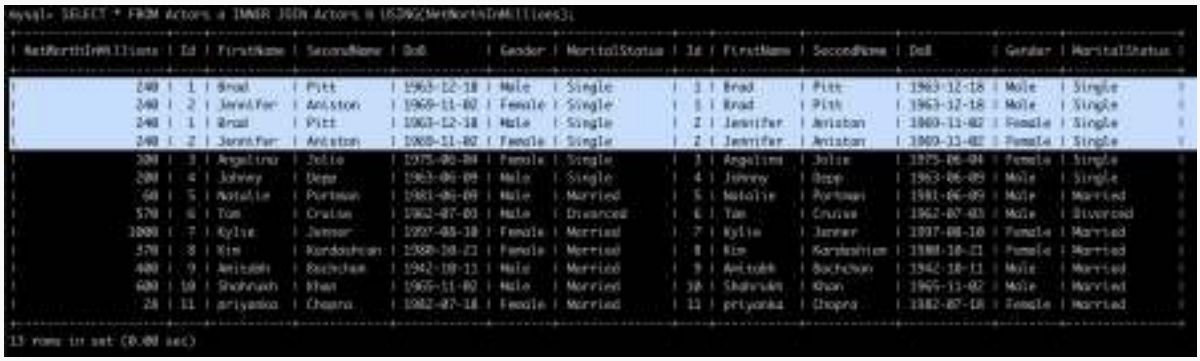
2. We can use the **USING** clause to specify the column to join the two tables on. For example:

```
SELECT * FROM Actors a INNER JOIN Actors b USING(FirstName);
```



Note we have exactly 11 rows now because each row in the first table matches exactly one row in the second table. However, if we change the query and specify the NetWorthInMillions column in the **USING** clause, we'll get 13 rows in the result because the two rows with value 240 for the NetWorthInMillions column match twice for a total of four rows.

```
SELECT * FROM Actors a INNER JOIN Actors b USING(NetWorthInMillio
ns);
```



3. Remember, the **USING** clause defines one or more columns that are in both tables or results and used to join or match rows. Furthermore, if any rows from the two tables don't match, they aren't included in the output. This obviously, will not happen in the case of a self-join.