

196. Delete Duplicate Emails

July 10, 2017 | 81.5K views

Previous

Next

★★★★★

Average Rating: 3.91 (34 votes)

Write a SQL query to **delete** all duplicate email entries in a table named **Person**, keeping only unique emails based on its *smallest Id*.

```
+-----+
| Id | Email |
+-----+
| 1  | john@example.com |
| 2  | bob@example.com  |
| 3  | john@example.com |
+-----+
Id is the primary key column for this table.
```

For example, after running your query, the above **Person** table should have the following rows:

```
+-----+
| Id | Email |
+-----+
| 1  | john@example.com |
| 2  | bob@example.com  |
+-----+
```

Note:

Your output is the whole **Person** table after executing your sql. Use **delete** statement.

Solution

Approach: Using **DELETE** and **WHERE** clause [Accepted]

Algorithm

By joining this table with itself on the *Email* column, we can get the following code.

```
SELECT p1.*
FROM Person p1,
      Person p2
WHERE
  p1.Email = p2.Email
;
```

Then we need to find the bigger id having same email address with other records. So we can add a new condition to the **WHERE** clause like this.

```
SELECT p1.*
FROM Person p1,
      Person p2
WHERE
  p1.Email = p2.Email AND p1.Id > p2.Id
;
```

As we already get the records to be deleted, we can alter this statement to **DELETE** in the end.

MySQL

```
DELETE p1 FROM Person p1,
      Person p2
WHERE
  p1.Email = p2.Email AND p1.Id > p2.Id
```

Rate this article: ★★★★★

Previous

Next

Comments: 87

Sort By

- Type comment here... (Markdown is supported)

Preview

Post
- legend1

★ 80

January 29, 2019 9:09 PM

THIS WORKED on MySQL!

```
DELETE FROM Person WHERE Id NOT IN
(SELECT * FROM(
SELECT MTN(Id) FROM Person GROUP BY Email) as n);
```

80

Share

Reply

SHOW 14 REPLIES
- jalonzou

★ 108

June 22, 2018 2:31 PM

Why group by Email doesn't work here?

106

Share

Reply

SHOW 8 REPLIES
- Bayesic

★ 26

April 11, 2018 10:46 AM

Is there something wrong with this problem? Every time I submit code, even if I just say select ID from Person, it returns the original table. This was my solution; it should work no?

```
SELECT
MIN(Id) as Id, Email
```

26

Share

Reply

SHOW 4 REPLIES
- alexapnw

★ 27

November 4, 2018 1:52 AM

DELETE FROM Person WHERE
Id NOT IN (SELECT MIN(p.Id) FROM (SELECT * FROM Person) p
Group by p.Email);

25

Share

Reply

SHOW 2 REPLIES
- jl4343

★ 29

July 23, 2018 1:14 AM

select Id, Email from
(select *,
row_number() over (partition by Email order by Id desc) as row_rank
from Person) base0
where row_rank = 1

24

Share

Reply

Read More
- GJIER

★ 40

September 20, 2018 5:49 AM

WHY THIS IS NOT ACCEPTABLE?

```
select min(id) as Id, Email
from person
group by email;
```

40

Share

Reply

SHOW 3 REPLIES
- tys9227

★ 31

October 23, 2018 9:57 PM

I don't understand why we could delete p1 from the join of itself, actually it is my first time seeing this format. How does this work exactly? And the condition **p1.Id > p2.Id** can only rule out the dual duplicates, it wouldn't apply to the case where there were more than 2 duplicates, right?

21

Share

Reply

SHOW 2 REPLIES
- wei_jung

★ 36

March 18, 2018 9:20 AM

SELECT MIN(Id) as Id, Email
FROM Person
GROPY BY Email
;

36

Share

Reply
- luoanfast000

★ 38

March 31, 2019 2:36 AM

DELETE FROM Person
WHERE ID NOT IN (
SELECT MIN(id)
FROM Person
Group by Email)

14

Share

Reply

Read More

SHOW 2 REPLIES
- a.shoaibsbx

★ 11

July 15, 2018 4:09 AM

delete from Person
where id NOT IN
(select z.y from (select min(x.id) as y
from person x
group by x.email) z);

9

Share

Reply

Read More