



链滴

# LeetCode196 Delete Duplicate Emails

作者: [houjie](#)

原文链接: <https://ld246.com/article/1512980247260>

来源网站: [链滴](#)

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

# LeetCode196. Delete Duplicate Emails

## 题目

Write a SQL query to delete all duplicate email entries in a table named **Person**, keeping only nique 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  |
+-----+-----+
```

题目大意: 删除重复的邮箱,只保留id最小的邮箱

## 解题思路

是否删除这条记录有两个条件

1. 之前存在这个邮箱
2. 之前存在的这个邮箱的id比和这条记录的id小

由此 使用 EXISTS 即可求解

```
-- 一开始写成了这样,一直报错 DELETE 能否使用相关子查询?
DELETE FROM Person p1 WHERE Id > (
SELECT min(Id) FROM Person p2 WHERE p1.Email = p2.Email)
```

## DELETE a FROM table a,table b 形式

```
-- 双表做笛卡尔积 使用相同思路取结果
DELETE p1
FROM Person p1, Person p2
WHERE p1.Email=p2.Email AND p1.Id>p2.Id
```

## NOT IN

先找出不用删除元组的id 然后删除剩下的 (速度较快)

```
DELETE FROM Person WHERE Id NOT IN (  
  SELECT NOT_DELETE_Id FROM (  
    SELECT min(Id) AS NOT_DELETE_Id FROM Person GROUP BY Email  
  ) temp  
);
```