

WRITE A QUERY TO DELETE DUPLICATE RECORDS IN SQL

Duplicate records in a SQL Server table can be a very serious issue. . Now by using these queries, let's delete duplicate values from the table.

I've used it on a table where a PHP bug of mine created duplicate rows. ID AND b. In this tip let's examine some ways to solve data de-duplication needs. In the code below there are six solutions to find that duplicate values which should be deleted leaving only one value : Finding duplicate values in a table with a unique index --Solution 1 SELECT a. Despite the fact that it is practically impossible to find official documentation about this feature, it can be used as an analog to ROWID pseudo column in Oracle. All three will generate a plan with a self join though that can be avoided. Check out the examples below to get real world examples on how to delete duplicate records from a table. For example, we can store distinct rows in a temporary table, then delete all data from our table and after that insert distinct rows from temporary table to our permanent table. Looking at the execution plan we can see that latest - the most 'compact' solution 'Solution 6' has a highest cost in our example there is a primary key on the 'ID' column, so 'NULL' values are not possible for that column, therefore 'NOT IN' will work without any problem , and the second has the lowest cost: Deleting Duplicate Rows in a SQL Server Table Now by using these queries, let's delete duplicate values from the table. Let's investigate and compare some common ways. While using this approach, it is very important to realize that this is an undocumented feature of SQL Server and, therefore, developers should be very careful. A big table was created and populated as described here: [sqlinthewild](#). When the loading process completes, table should be cleaned or clean data should be loaded to a permanent table, so after that duplicates are no longer needed. In this tip we explained some techniques which hopefully will be helpful to solve these types of problems. Then I proceeded with the execution plans, and guess what? By Jagan Mohan Create a table EmpDtl1 with some duplicate rows as shown below to understand different methods of delete duplicate rows. However, sometimes we need to work with databases where these rules are not followed or exceptions are possible when these rules are bypassed knowingly. One place where the GROUP BY solution might be preferred is situations where a hash aggregate would be chosen in preference to a stream aggregate. For example, when a staging table is used and data is loaded from different sources where duplicate rows are possible. The second case is that table does not have a primary key or any unique indexes and contains duplicate rows which should be removed.