

عنوان: روش‌هایی برای حذف رکوردهای تکراری

نویسنده: وحید نصیری

تاریخ: ۱۳۸۸/۰۵/۲۷ ۲۲:۲۵:۰۰

آدرس: www.dotnettips.info

برچسب‌ها: SQL Server

هر برنامه نویسی در طول عمر کاری خود حداقل یکبار با این مساله مواجه خواهد شد: "چگونه یک سری رکورد تکراری موجود را باید حذف کرد؟"

ابتدا جدول موقتی زیر را که در آن یک سری رکورد تکراری ثبت شده است در نظر بگیرید:

```
CREATE TABLE #Employee
(
    ID          INT,
    FIRST_NAME  NVARCHAR(100),
    LAST_NAME   NVARCHAR(300)
)

INSERT INTO #Employee VALUES ( 1, 'Vahid', 'Nasiri' );
INSERT INTO #Employee VALUES ( 2, 'name1', 'lname1' );
INSERT INTO #Employee VALUES ( 3, 'name2', 'lname2' );
INSERT INTO #Employee VALUES ( 2, 'name1', 'lname1' );
INSERT INTO #Employee VALUES ( 3, 'name2', 'lname2' );
INSERT INTO #Employee VALUES ( 4, 'name3', 'lname3' );
```

روش‌های حذف رکوردهای تکراری جدول موقتی Employee

الف) استفاده از یک جدول موقتی دیگر بر اساس تمامی فیلدهای موجود

```
SELECT DISTINCT *
FROM #Employee

SELECT * INTO #DuplicateEmployee
FROM #Employee

INSERT #DuplicateEmployee
SELECT DISTINCT *
FROM #Employee

BEGIN TRAN
DELETE #Employee
INSERT #Employee
SELECT *
FROM #DuplicateEmployee

COMMIT TRAN

DROP TABLE #DuplicateEmployee

SELECT DISTINCT *
FROM #Employee
```

ب) استفاده از یک جدول موقتی دیگر بر اساس یک مجموعه از فیلدهای موجود

```
SELECT DISTINCT * FROM #Employee

SELECT * INTO #DuplicateEmployee FROM #Employee

INSERT #DuplicateEmployee
SELECT ID,
       FIRST_NAME,
       LAST_NAME
```

```

FROM #Employee
GROUP BY
    ID,
    FIRST_NAME,
    LAST_NAME
HAVING COUNT(*) > 1

BEGIN TRAN
DELETE #Employee
FROM #DuplicateEmployee
WHERE #Employee.ID = #DuplicateEmployee.ID
    AND #Employee.FIRST_NAME = #DuplicateEmployee.FIRST_NAME
    AND #Employee.LAST_NAME = #DuplicateEmployee.LAST_NAME

INSERT #Employee
SELECT *
FROM #DuplicateEmployee

COMMIT TRAN
DROP TABLE #DuplicateEmployee

SELECT DISTINCT * FROM #Employee

```

ج) استفاده از rowcount بر اساس یک مجموعه از فیلدهای موجود

```

SELECT DISTINCT *
FROM #Employee

SET ROWCOUNT 1
SELECT 1
WHILE @@rowcount > 0
    DELETE #Employee
    WHERE 1 < (
        SELECT COUNT(*)
        FROM #Employee a2
        WHERE #Employee.ID = a2.ID
            AND #Employee.FIRST_NAME = a2.FIRST_NAME
            AND #Employee.LAST_NAME = a2.LAST_NAME
    )

SET ROWCOUNT 0

SELECT DISTINCT *
FROM #Employee

```

د) استفاده از Analytical Functions بر اساس یک مجموعه از فیلدهای موجود

```

SELECT DISTINCT *
FROM #Employee;

WITH #DeleteEmployee AS (
    SELECT ROW_NUMBER()
        OVER(PARTITION BY ID, First_Name, Last_Name ORDER BY ID) AS
        RNUM
    FROM #Employee
)

DELETE
FROM #DeleteEmployee
WHERE RNUM > 1

SELECT DISTINCT *
FROM #Employee

```

ه) استفاده از یک فیلد identity جدید بر اساس یک مجموعه از فیلدهای موجود

```

SELECT DISTINCT *

```

```
FROM #Employee;

ALTER TABLE #Employee ADD UNIQ_ID INT IDENTITY(1, 1)

DELETE
FROM #Employee
WHERE UNIQ_ID < (
    SELECT MAX(UNIQ_ID)
    FROM #Employee a2
    WHERE #Employee.ID = a2.ID
        AND #Employee.FIRST_NAME = a2.FIRST_NAME
        AND #Employee.LAST_NAME = a2.LAST_NAME
)

ALTER TABLE #Employee DROP COLUMN UNIQ_ID

SELECT DISTINCT *
FROM #Employee
```

و در آخر

```
DROP TABLE #Employee
```

نظرات خوانندگان

نویسنده: salarblog
تاریخ: ۱۵:۵۶:۵۹ ۱۳۸۸/۰۵/۲۸

بسیار عالی، تشکر.

نویسنده: مهران زند
تاریخ: ۰۹:۱۱:۴۶ ۱۳۸۸/۰۵/۳۱

جناب نصیری من دقیقا کد های شما را نخواندمو اما من از تیکه کد زیر برای حذف رکورد های تکراری استفاده میکنم :

```
DELETE
FROM MyTable
WHERE ID NOT IN
(
(SELECT MAX(ID)
FROM MyTable
GROUP BY COLUMN1, COLUMN2
```

که در اینجا ستون ها ستون هایی هستند که امکان یونیک بودن داده ها در آن وجود دارد.

با سپاس