The TOP(n) Clause

.NET Cohort

Coding Bootcamp

SOFTWARE GUILD

Lesson Goals

- Learn to use the TOP function to limit result sets
- Learn about the utility of TOP with programming concepts like WHILE loops in large data processing tasks

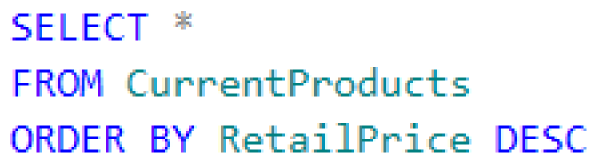
TOP(n) Queries

Oftentimes in reporting applications, leader boards, etc., we will be asked to find a limited set of items matching certain criteria.

For example: “What are the 3 most expensive products we sell?”

Answering the Question

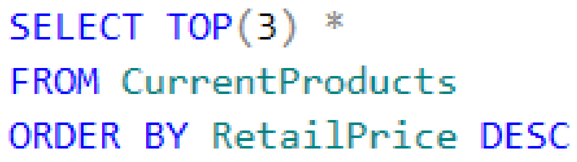
We can easily write a query at this point to order the products in SWCCorp by RetailPrice descending and visually intuit the top 3, like so:



```
SELECT *
FROM CurrentProducts
ORDER BY RetailPrice DESC
```

Limiting the Result Set to 3 rows

However, the request clearly states that we only want the top 3 results, not all of them. The *top* keyword allows us to tell the database to limit the results of a query. Try this:



```
SELECT TOP(3) *
FROM CurrentProducts
ORDER BY RetailPrice DESC
```

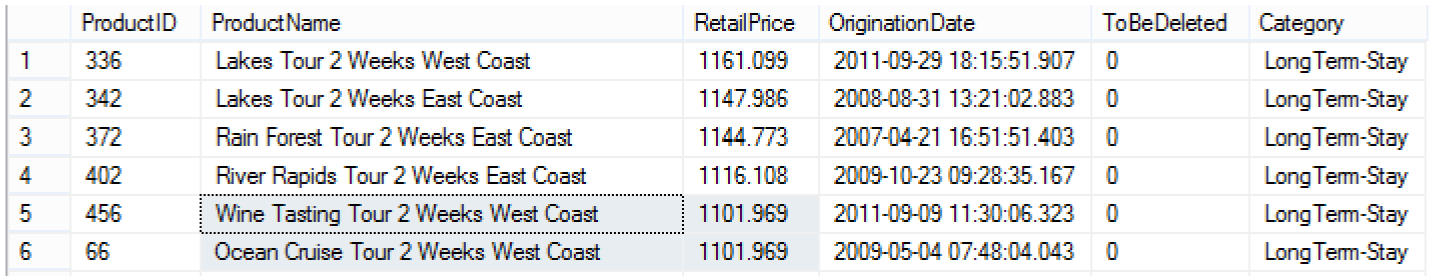
FAQ: Is there a Bottom() clause?

No. If you change the order by statement, you can easily get records from the other end of the query, so it is unnecessary.

What about ties?

Run the original query and look at the result set, taking note of the 5th and 6th rows:

```
SELECT *
FROM CurrentProducts
ORDER BY RetailPrice DESC
```



	ProductID	ProductName	RetailPrice	OriginationDate	ToBeDeleted	Category
1	336	Lakes Tour 2 Weeks West Coast	1161.099	2011-09-29 18:15:51.907	0	LongTerm-Stay
2	342	Lakes Tour 2 Weeks East Coast	1147.986	2008-08-31 13:21:02.883	0	LongTerm-Stay
3	372	Rain Forest Tour 2 Weeks East Coast	1144.773	2007-04-21 16:51:51.403	0	LongTerm-Stay
4	402	River Rapids Tour 2 Weeks East Coast	1116.108	2009-10-23 09:28:35.167	0	LongTerm-Stay
5	456	Wine Tasting Tour 2 Weeks West Coast	1101.969	2011-09-09 11:30:06.323	0	LongTerm-Stay
6	66	Ocean Cruise Tour 2 Weeks West Coast	1101.969	2009-05-04 07:48:04.043	0	LongTerm-Stay

How many rows will be returned if we do a Top(5) query?

TOP(n) WITH TIES

If we want to include rows that have ties in a top query, we can simply append the command “WITH TIES” to the statement and it will expand the result set to include the tied records.

```
SELECT TOP(5) WITH TIES *
FROM CurrentProducts
ORDER BY RetailPrice DESC
```

SOFTWARE GUILD

Other uses for TOP(n)

Oftentimes, doing large deletes or mass updates will cause a lock on the table until all of the records are processed.

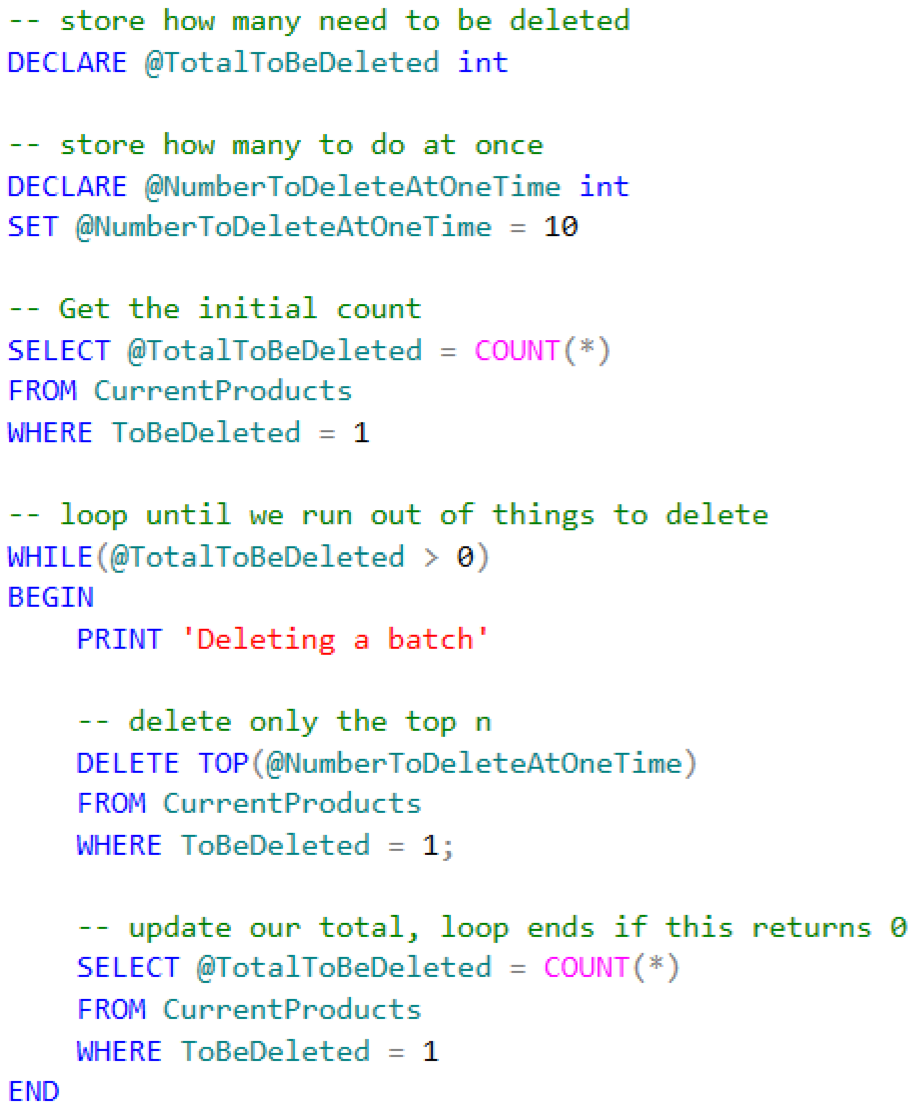
If we are editing millions of rows during productions hours, this can be unacceptable. In that case, we may want to delete the top 100 at a time until all the records are gone.

Deleting Products in Batches

The CurrentProducts table has a column called “ToBeDeleted.”

When set to true (1), it signifies that the product is discontinued and we can remove it from the database.

If you write a query to count the rows where ToBeDeleted = 1, you will find there are 80. Let’s write a SQL Script to delete 10 rows at a time.



```sql
-- store how many need to be deleted
DECLARE @TotalToBeDeleted int

-- store how many to do at once
DECLARE @NumberToDeleteAtOneTime int
SET @NumberToDeleteAtOneTime = 10

-- Get the initial count
SELECT @TotalToBeDeleted = COUNT(*)
FROM CurrentProducts
WHERE ToBeDeleted = 1

-- loop until we run out of things to delete
WHILE(@TotalToBeDeleted > 0)
BEGIN
    PRINT 'Deleting a batch'

    -- delete only the top n
    DELETE TOP(@NumberToDeleteAtOneTime)
    FROM CurrentProducts
    WHERE ToBeDeleted = 1;

    -- update our total, loop ends if this returns 0
    SELECT @TotalToBeDeleted = COUNT(*)
    FROM CurrentProducts
    WHERE ToBeDeleted = 1
END
```

New Concepts

1. You can set a variable using the SET keyword (statement 3) or by including it as an expression in a SELECT statement (statement 4) .
2. SQL is a programming language and has a WHILE loop similar to C#. The main difference is that you use BEGIN/END instead of curly braces to denote a code block in SQL.
    1. SQL also has an IF statement which is often used with the EXISTS command.
    2. SQL <u>does not</u> have a FOR loop

SQL has some powerful language features. You are not limited to single queries.

```sql
-- store how many need to be deleted
DECLARE @TotalToBeDeleted int

-- store how many to do at once
DECLARE @NumberToDeleteAtOneTime int
SET @NumberToDeleteAtOneTime = 10

-- Get the initial count
SELECT @TotalToBeDeleted = COUNT(*)
FROM CurrentProducts
WHERE ToBeDeleted = 1

-- loop until we run out of things to delete
WHILE(@TotalToBeDeleted > 0)
BEGIN
    PRINT 'Deleting a batch'

    -- delete only the top n
    DELETE TOP(@NumberToDeleteAtOneTime)
    FROM CurrentProducts
    WHERE ToBeDeleted = 1;

    -- update our total, loop ends if this returns 0
    SELECT @TotalToBeDeleted = COUNT(*)
    FROM CurrentProducts
    WHERE ToBeDeleted = 1
END
```

Lab Exercises (SWCCorp)

1. Find the oldest two employees in the employee table.
2. Find the 6 largest grants in the grant table, including ties.
3. Display the 10 most expensive single-day trips found in the CurrentProducts table (Category = No-Stay)