ΣΥΣΤΗΜΑΤΑ ΔΙΑΧΕΙΡΙΣΗΣ ΚΑΙ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ

ΠΡΩΤΗ ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΗ ΕΡΓΑΣΙΑ

ΧΡΗΣΤΟΣ ΠΑΠΑΦΩΤΗΣ 3190162

ΖΗΤΗΜΑ ΠΡΩΤΟ

Εκτελούμε αρχικά το επερώτημα και καταγράφουμε τα στατιστικά στοιχεία της εκτέλεσής του.

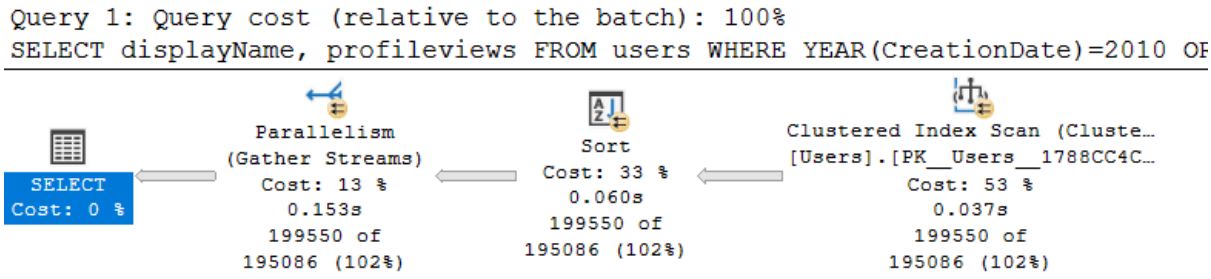
STATISTICS IO

```
(199550 rows affected)
Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Users'. Scan count 13, logical reads 6001, physical reads 3, page server reads
0, read-ahead reads 5712, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
```

```
STATISTICS TIME
```

```
SQL Server Execution Times:
   CPU time = 187 ms,  elapsed time = 1357 ms.
```

Το πλάνο εκτέλεσης



Μια λύση είναι να δημιουργήσουμε ένα ευρετήριο στο creationDate, όμως για να μπορέσουμε να το χρησιμοποιήσουμε πρέπει πρώτα να αλλάξουμε το WHERE statement σε :

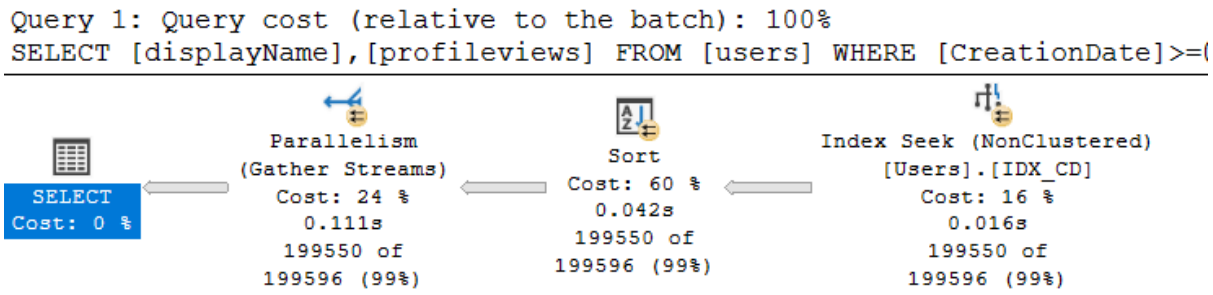
```
WHERE CreationDate >= '2010-01-01' AND CreationDate < '2011-01-01'
```

Δημιουργούμε το ευρετήριο

```
CREATE NONCLUSTERED INDEX IDX_CD
ON [dbo].[Users] ([CreationDate])
```

```
INCLUDE ([DisplayName],[ProfileViews])
```

Βλέπουμε ότι το ευρετήριο χρησιμοποιείται:




```
(199550 rows affected)
Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Users'. Scan count 13, logical reads 985, physical reads 3, page server reads
0, read-ahead reads 969, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.

(1 row affected)

STATISTICS TIME

SQL Server Execution Times:
   CPU time = 61 ms,  elapsed time = 1346 ms.
```

Άρα βελτιστοποιήσαμε το ερώτημα.

ΖΗΤΗΜΑ ΔΕΥΤΕΡΟ

Θα ήθελα να ενημερώσω σε αυτό το σημείο ότι σε αυτό το ζήτημα ενώ στην αρχή το έτρεχα κανονικά και μου εμφάνιζε αποτελέσματα, μετά τις επόμενες μέρες σταμάτησε και δεν τελείωνε ποτέ το execute. Τα αποτελέσματα από κάτω είναι από τις σημειώσεις που είχα κρατήσει όταν τα έτρεχα, όπου τότε δεν ήξερα ότι έχουν σημασία τα reads γι'αυτο και δεν υπάρχουν.

```
(70698 rows affected)
Table 'Users'. Scan count 26, logical reads 12002, physical reads 3, page server reads
0, read-ahead reads 5712, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Posts'. Scan count 26, logical reads 740104, physical reads 6, page server
reads 0, read-ahead reads 735125, page server read-ahead reads 0, lob logical reads 0,
lob physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'PostTypes'. Scan count 5, logical reads 8, physical reads 1, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
```

```
Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads
0, read-ahead reads 19068, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.


SQL Server Execution Times:
   CPU time = 3072 ms,  elapsed time = 45032 ms.
```

ΜΕ INDEX ΣΤΟ POST_TYPE_ NAME

```
CPU time = 3610 ms,  elapsed time = 13906 ms.
```

ΜΕ ΕΥΡΕΤΗΡΙΟ ΣΤΟ POST_TYPE_ID ΚΑΙ INCLUDE ΤΟ OWNER_USER_ID

```
SQL Server Execution Times:
   CPU time = 2681 ms,  elapsed time = 4061 ms.
```

ΖΗΤΗΜΑ ΤΡΙΤΟ

Εκτελούμε αρχικά το επερώτημα και καταγράφουμε τα στατιστικά στοιχεία της εκτέλεσής του.

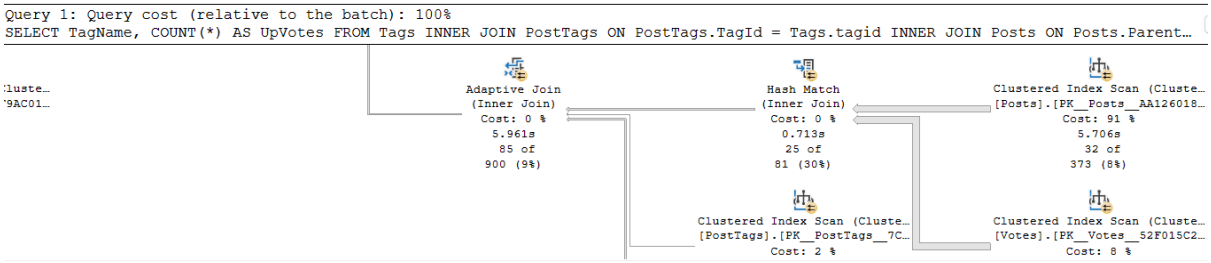
```
(39 rows affected)
Table 'Tags'. Scan count 0, logical reads 170, physical reads 37, page server reads 0,
read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob physical
reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server read-ahead
reads 0.
Table 'PostTags'. Scan count 24, logical reads 72, physical reads 25, page server
reads 0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Votes'. Scan count 13, logical reads 29495, physical reads 1, page server reads
0, read-ahead reads 29071, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Posts'. Scan count 13, logical reads 370097, physical reads 3, page server
reads 0, read-ahead reads 368412, page server read-ahead reads 0, lob logical reads 0,
lob physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'VoteTypes'. Scan count 13, logical reads 4, physical reads 1, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Users'. Scan count 13, logical reads 6001, physical reads 2, page server reads
0, read-ahead reads 5717, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
```

```
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.

(1 row affected)

SQL Server Execution Times:
   CPU time = 329 ms,  elapsed time = 6091 ms.
```

Στο πλάνο εκτέλεσης φαίνεται ότι ξοδεύεται πολύς χρόνος για τον όρο owner_id

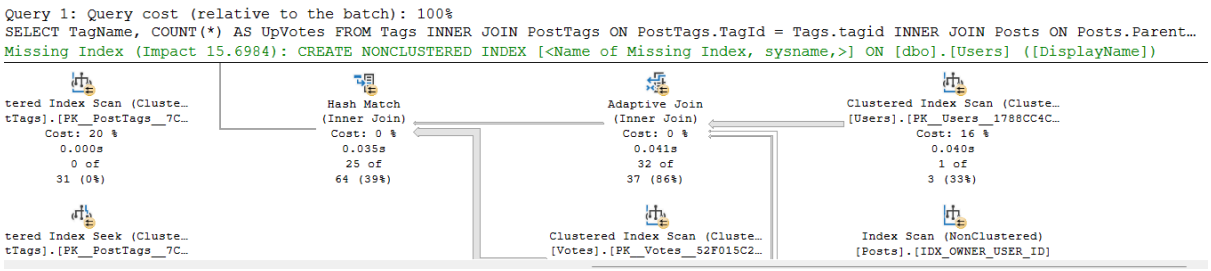


Δημιουργούμε Index στο OwnerUserId με include το parentId για να αποφύγουμε KeyLookUp

```
CREATE NONCLUSTERED INDEX IDX_OWNER_USER_ID
ON [dbo].[Posts] ([OwnerUserId])
INCLUDE (ParentId)
```

```
Table 'Posts'. Scan count 1, logical reads 3, physical reads 3, page server reads 0,
read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob physical
reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server read-ahead
reads 0.


SQL Server Execution Times:
   CPU time = 78 ms,  elapsed time = 382 ms.
```



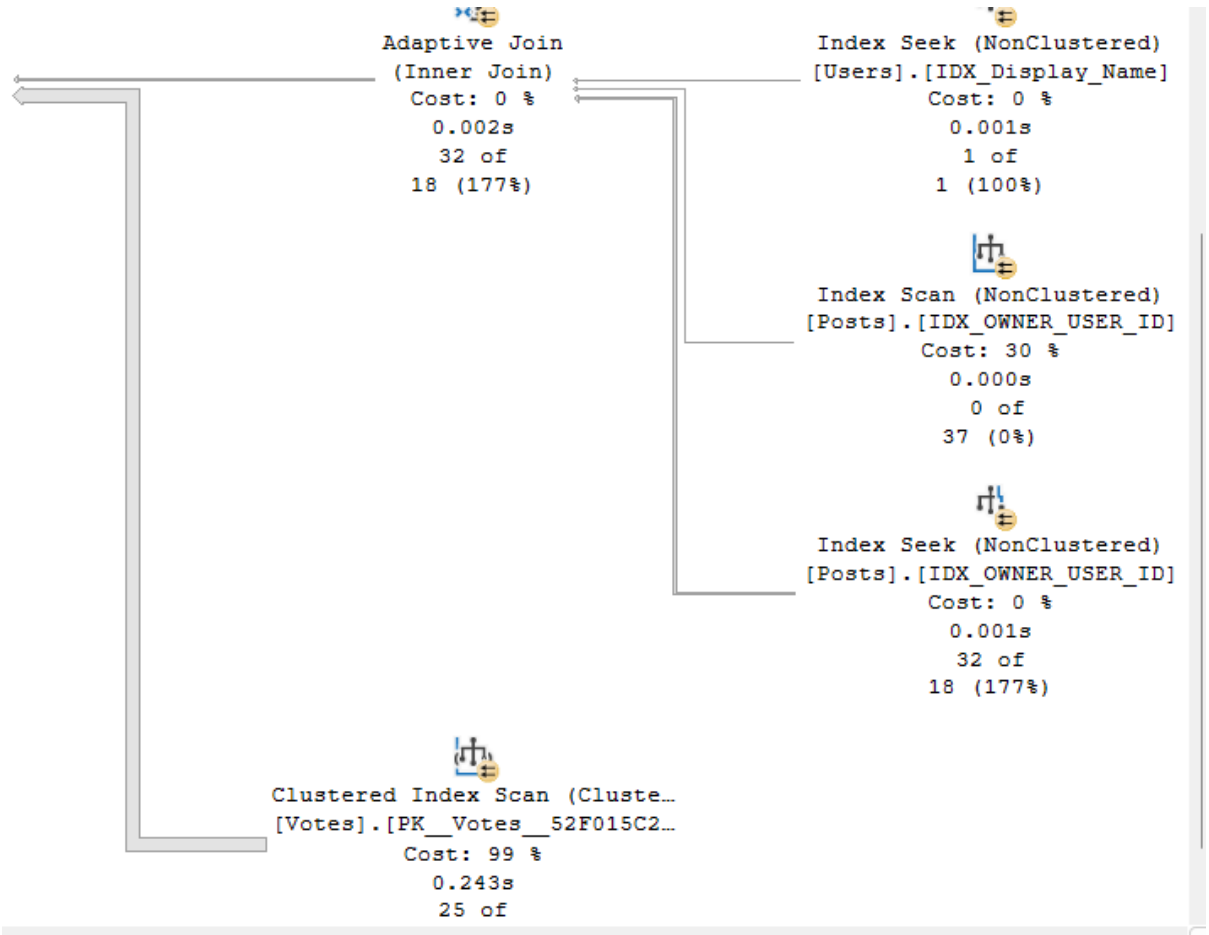
Δημιουργούμε Index στο DisplayName που φαίνεται να έχει μερικό κόστος

```
CREATE NONCLUSTERED INDEX IDX_Display_Name
ON [dbo].[Users] ([DisplayName])
```

```
Table 'Users'. Scan count 13, logical reads 6, physical reads 2, page server reads 0,
read-ahead reads 8, page server read-ahead reads 0, lob logical reads 0, lob physical
reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server read-ahead
reads 0.


SQL Server Execution Times:
   CPU time = 172 ms,  elapsed time = 324 ms.
```

Βλέπουμε στο πλάνο εκτέλεσης ότι πλέον σχεδόν όλο το κόστος έχει ο όρος postId



ME INDEX ΣΤΟ PostId

```
CREATE NONCLUSTERED INDEX Idx_PostId
ON [dbo].[Votes] ([PostId])
INCLUDE ([VoteTypeId])
```

```
Table 'Tags'. Scan count 1, logical reads 118, physical reads 1, page server reads 0,
read-ahead reads 123, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'VoteTypes'. Scan count 1, logical reads 2, physical reads 1, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
```

```
Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'PostTags'. Scan count 24, logical reads 172, physical reads 1, page server
reads 0, read-ahead reads 128, page server read-ahead reads 0, lob logical reads 0,
lob physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Votes'. Scan count 32, logical reads 96, physical reads 13, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Posts'. Scan count 1, logical reads 3, physical reads 3, page server reads 0,
read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob physical
reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server read-ahead
reads 0.
Table 'Users'. Scan count 1, logical reads 3, physical reads 3, page server reads 0,
read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob physical
reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server read-ahead
reads 0.


SQL Server Execution Times:
   CPU time = 0 ms,  elapsed time = 74 ms.
```

ΖΗΤΗΜΑ ΤΕΤΑΡΤΟ

Εκτελούμε αρχικά το επερώτημα και καταγράφουμε τα στατιστικά

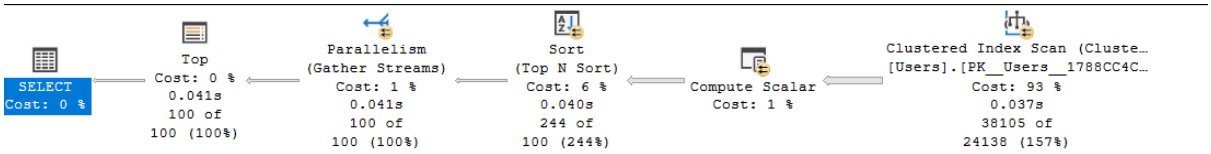
STATISTICS I/O

```
(100 rows affected)
Table 'Users'. Scan count 13, logical reads 6001, physical reads 2, page server reads
0, read-ahead reads 5717, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.

(1 row affected)

STATISTICS TIME


SQL Server Execution Times:
   CPU time = 0 ms,  elapsed time = 124 ms.
```



ΤΑ ΔΥΟ ΕΝΑΛΛΑΚΤΙΚΑ ΕΥΡΕΤΗΡΙΑ ΤΑ ΟΠΟΙΑ ΕΠΙΤΑΧΥΝΟΥΝ ΤΗΝ ΕΚΤΕΛΕΣΗ ΕΙΝΑΙ ΤΑ

```
CREATE NONCLUSTERED INDEX Idx_reputation
ON [dbo].[Users] ([Reputation])
INCLUDE ([DownVotes])
```

```
CREATE NONCLUSTERED INDEX Idx_upvotes
ON [dbo].[Users] ([UpVotes])
INCLUDE ([DownVotes])
```

Επιλέγουμε να δημιουργήσουμε το ευρετήριο με την μεγαλύτερη επιλεξιμότητα (selectivity):

```
Select CAST (count(distinct Reputation) as float) / count(*)
as LocationAverageSelectivity
FROM users
```

```
LocationAverageSelectivity

0.0543759623509922
```

```
Select CAST (count(distinct Upvotes) as float) / count(*)
as LocationAverageSelectivity
FROM users
```

```
LocationAverageSelectivity

0.0136841718520894
```

Βάσει των παραπάνω επιλέγουμε το ευρετήριο:

```
CREATE NONCLUSTERED INDEX Idx_reputation
ON [dbo].[Users] ([Reputation])
INCLUDE ([DownVotes])
```

Δημιουργούμε το ευρετήριο και εκτελούμε το επερώτημα.

```
SQL Server Execution Times:
   CPU time = 0 ms,  elapsed time = 128 ms.
```

ΖΗΤΗΜΑ ΠΕΜΠΤΟ

1.

Έστω το επερώτημα, το οποίο εμφανίζει το userId, το όνομα και το reputation των χρηστών που βρίσκονται στην Ελλάδα και έχουν το σήμα (διακριτικό) ‘Student’.

```
SELECT Users.UserId, Users.DisplayName, Users.Reputation
```

```
FROM Users
JOIN Badges ON Users.UserId = Badges.UserId
WHERE Users.UserLocation LIKE '%Greece' AND Badges.Bname = 'Student';
```

2. Εκτελούμε αρχικά το επερώτημα και καταγραφουμε τα στατιστικά

```
(152 rows affected)
Table 'Badges'. Scan count 13, logical reads 5942, physical reads 1, page server reads
0, read-ahead reads 5868, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Users'. Scan count 13, logical reads 6001, physical reads 2, page server reads
0, read-ahead reads 5717, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.

(1 row affected)

 SQL Server Execution Times:
   CPU time = 109 ms,  elapsed time = 234 ms.
```

Δημιουργούμε ευρετήριο στο όρισμα Badges και κάνουμε include το UserId.

```
CREATE NONCLUSTERED INDEX IDX_Bname
ON [dbo].[Badges] ([Bname])
INCLUDE ([UserId])
```

```
(152 rows affected)
Table 'Badges'. Scan count 13, logical reads 567, physical reads 3, page server reads
0, read-ahead reads 539, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Users'. Scan count 13, logical reads 6001, physical reads 2, page server reads
0, read-ahead reads 5717, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads
0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.

(1 row affected)

 SQL Server Execution Times:
   CPU time = 47 ms,  elapsed time = 128 ms.
```

Ευρετήριο στο UserLocation

```
CREATE NONCLUSTERED INDEX IDX_UserLocation
ON [dbo].[Users] ([UserLocation])
```

```
(152 rows affected)
Table 'Badges'. Scan count 1, logical reads 538, physical reads 3, page server reads
0, read-ahead reads 539, page server read-ahead reads 0, lob logical reads 0, lob
```

```
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.
Table 'Users'. Scan count 1, logical reads 2401, physical reads 4, page server reads
0, read-ahead reads 2959, page server read-ahead reads 0, lob logical reads 0, lob
physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server
read-ahead reads 0.


(1 row affected)

 SQL Server Execution Times:
   CPU time = 125 ms,  elapsed time = 249 ms.
```