Συστήματα Διαχείρησης και Ανάλυσης Δεδομένων 1° Project

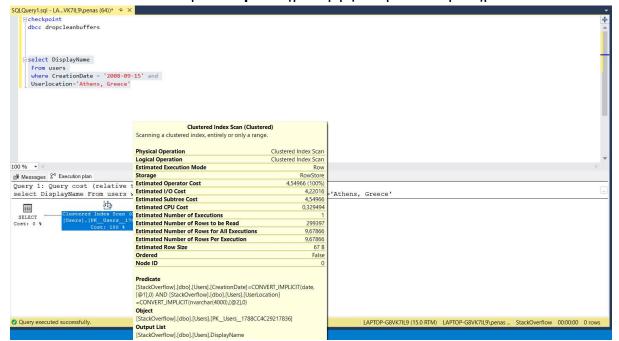
Χρυσικός Παναγιώτης – 3180295

Ζήτημα Πρώτο (20 μονάδες)

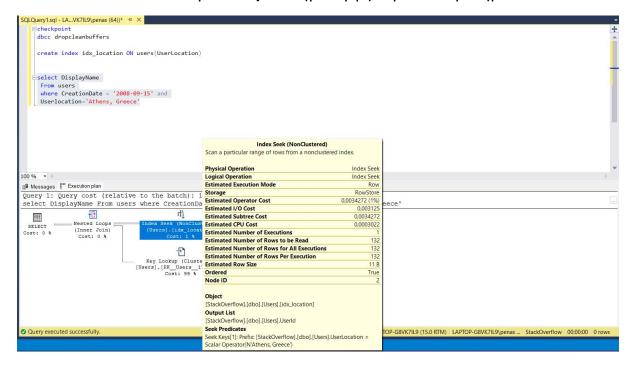
1) Έχουμε να κάνουμε με μία βάση δεδομένων παγκοσμίου βεληνηκές. Επομένως θα ήταν πιο λογικό αλλά και προς το συμφέρον μας να βάλουμε index στο UserLocation γιατί αφού πρόκειται για βάση παγκόσμιας εμβέλειας είναι πολύ πιο πιθανό να υπάρχουν περισσότερα αποτελέσματα για μια συγκεκριμένη ημερομηνία από μια συγκεκριμένη πόλη.

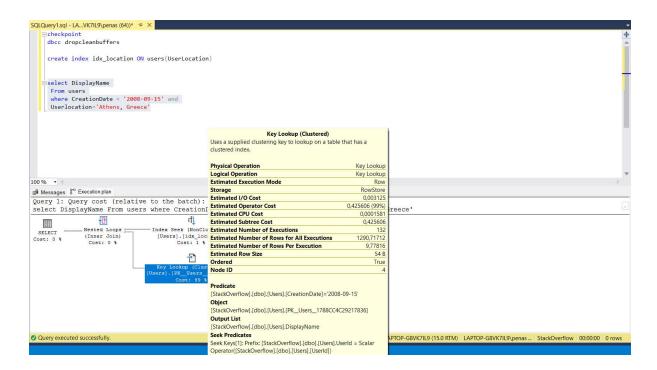
2)

Αυτά είναι τα αποτελέσματα πριν δημιουργήσουμε το ευρετήριο

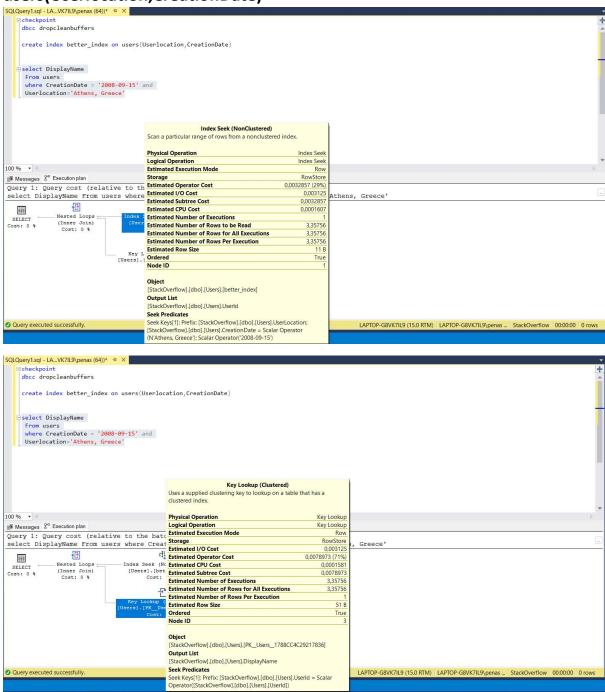


Αυτά είναι τα αποτελέσματα αφού δημιουργήσαμε το ευρετήριο



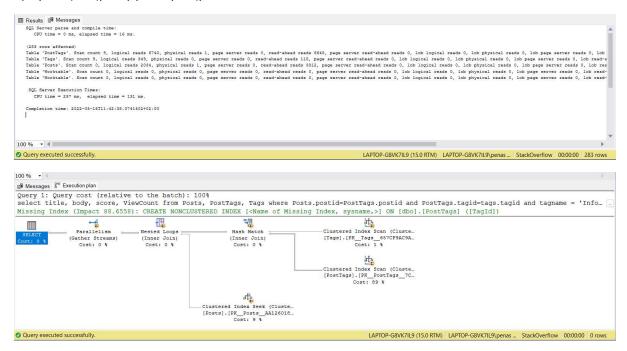


3) Ένα καλύτερο ευρετήριο είναι το 'create index better_index on users(Userlocation,CreationDate)'

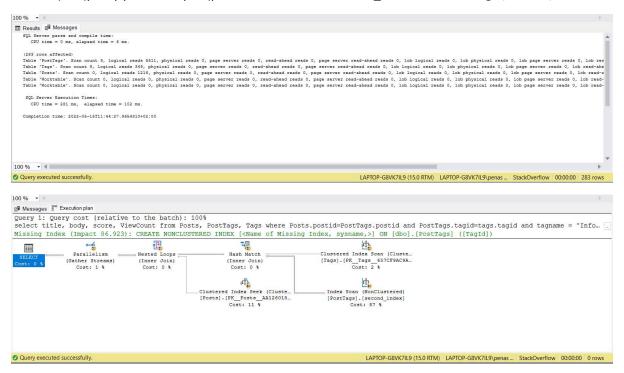


Ζήτημα Δεύτερο (20 μονάδες)

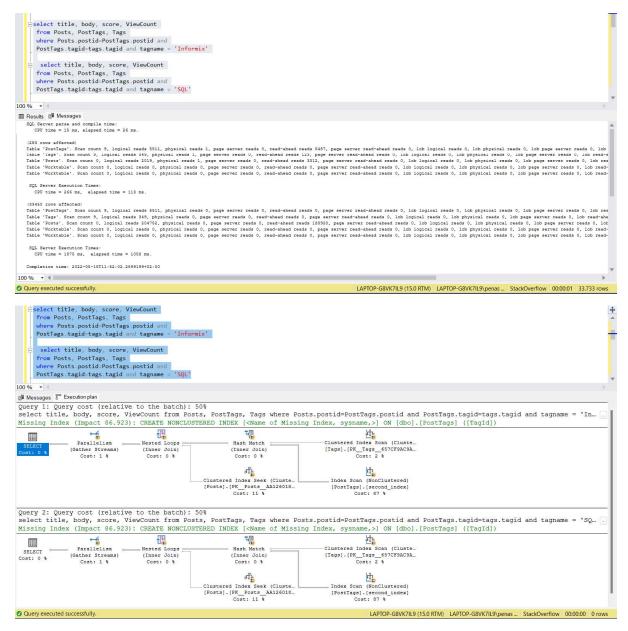
1) Πριν την δημιουργία ευρετηρίων



Μετα την δημιουργία του ευρετηρίου create index second_index on PostTags(PostId)

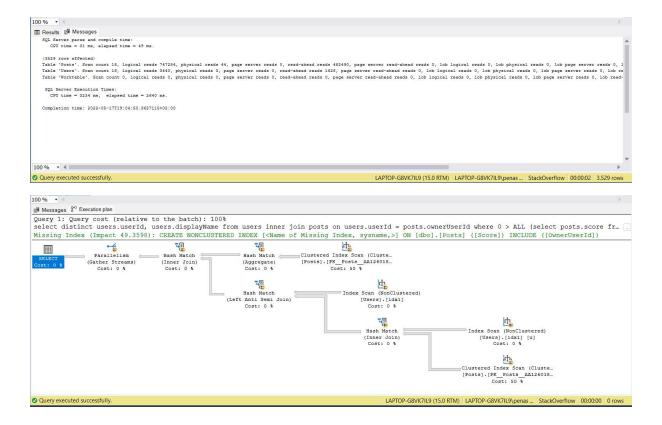


2) Παρατηρούμε πως το δεύτερο ερώτημα είναι πολύ πιο αργό αφού υπάρχουν πολύ περισσότερες εγγραφές με tagname = 'SQL'(33.450) σε σχέση με tagname = 'Informix'(283)

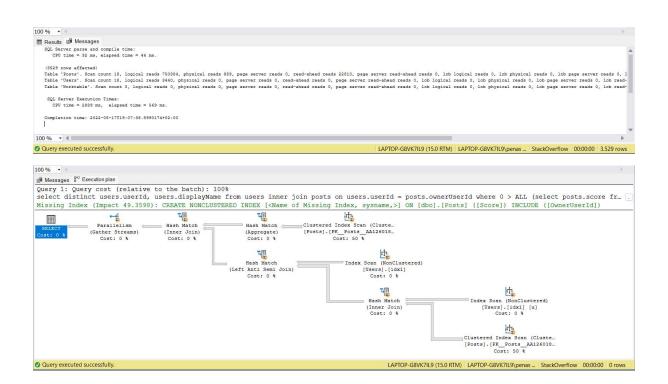


Ζήτημα 3 (20 μονάδες)

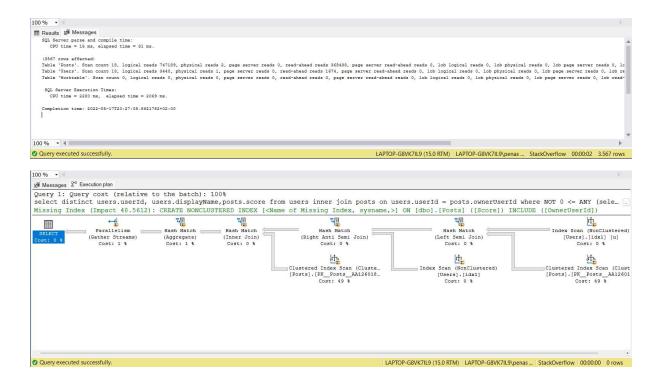
```
i)select distinct users.userId, users.displayName
from users
inner join posts
on users.userId = posts.ownerUserId
where 0 > ALL
(select posts.score
from posts
inner join users u
on posts.ownerUserId = u.userId and users.userId = u.userId)
```



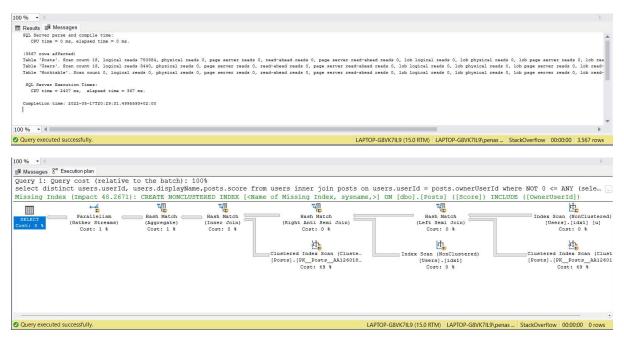
Μετά την δημιουργία του create index third_index on posts(score)



```
ii) select distinct users.userId, users.displayName,posts.score
from users
inner join posts
on users.userId = posts.ownerUserId
where NOT 0 <= ANY
(select posts.score
from posts
inner join users u
on posts.ownerUserId = u.userId and users.userId = u.userId)</pre>
```



$Μετά την δημιουργία του create index third_index1 on posts(score)$

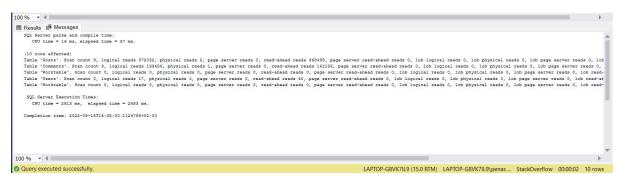


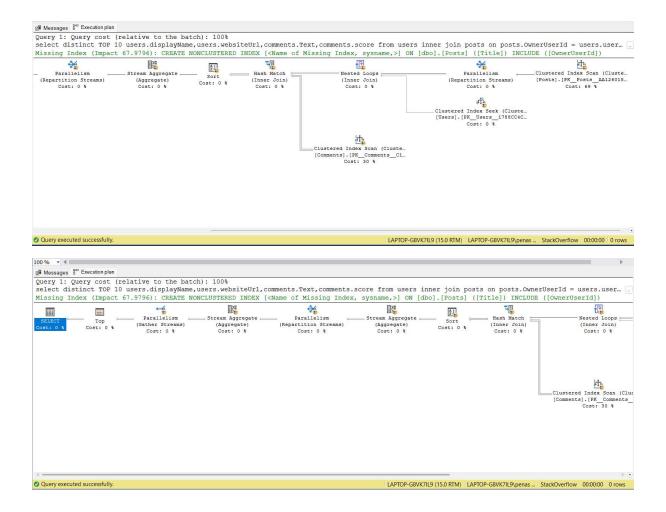
Ζήτημα Τέταρτο (40 μονάδες)

i) Βρίσκει τα 10 κορυφαία σχόλια για το θέμα 'Encrypting Passwords' και τα εμφανίζει μαζί με το όνομα του χρήστη το URL της ιστοσελίδας του και το 'σκορ' του σχολίου του.Τα αποτελέσματα εμφανίζονται σε φθίνουσα σειρά βάση το 'σκορ' σχολίου

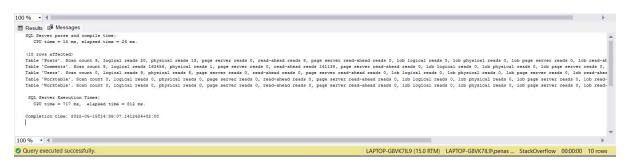
```
select distinct TOP 10 users.displayName,users.websiteUrl,comments.Text,comments.score
from users
inner join posts
on posts.OwnerUserId = users.userId
inner join comments
on comments.userId = users.userId
where posts.Title like 'Encrypting Passwords'
order by comments.score DESC
```

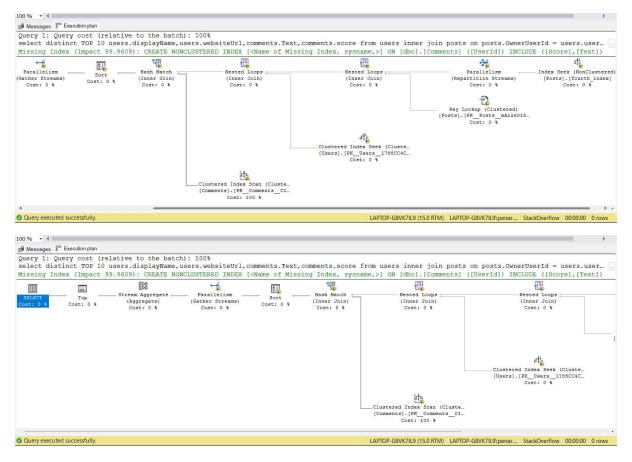
Πριν την δημιουργία ευρετηρίου





Μετά την δημιουργία του ευρετηρίου 'create index fourth_index on posts(title)'



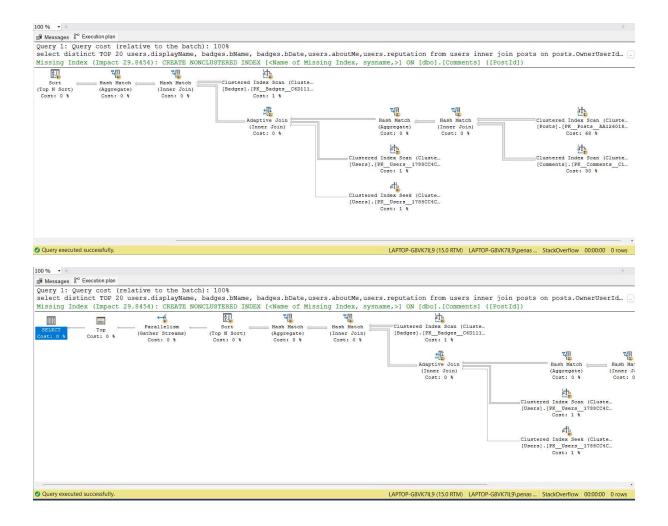


ii) Εμφανίζει τους κορυφαίους 20 χρήστες με το έμβλημα 'Scholar' μαζί με την ημερομηνία απόκτησης του, την περιγραφή για τον εαυτό τους και την φήμη τους.Τα αποτελέσματα εμφανίζονται βάση την φήμη σε φθίνουσα σειρά

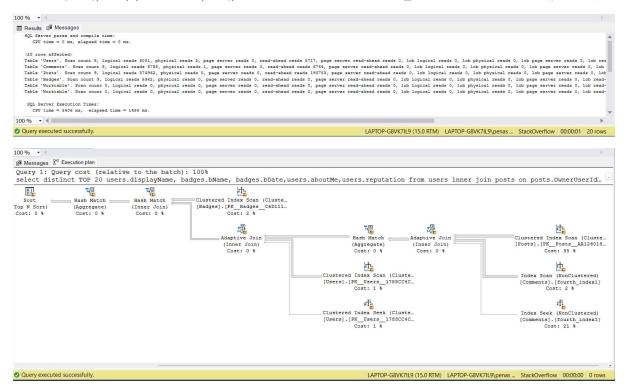
```
select distinct TOP 20 users.displayName, badges.bName,
badges.bDate,users.aboutMe,users.reputation
from users
inner join posts
on posts.OwnerUserId = users.userId
inner join badges
on badges.userId = users.UserId
inner join comments
on comments.PostId = posts.postId
WHERE badges.bname = 'Scholar'
order by reputation DESC
```

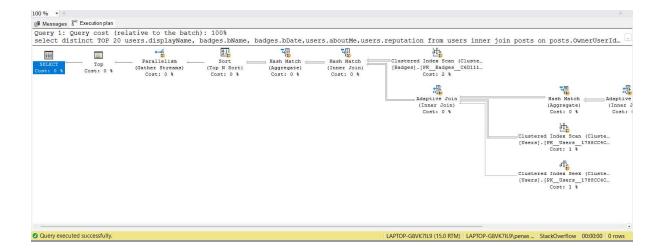
Πριν την δημιουργία ευρετηρίου





Μετά την δημιουργία του ευρετηρίου create index fourth_index1 on comments(postid)

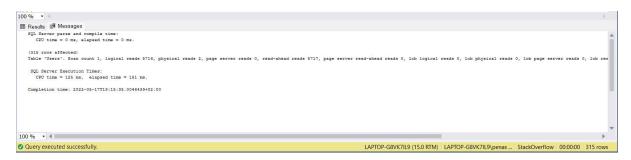


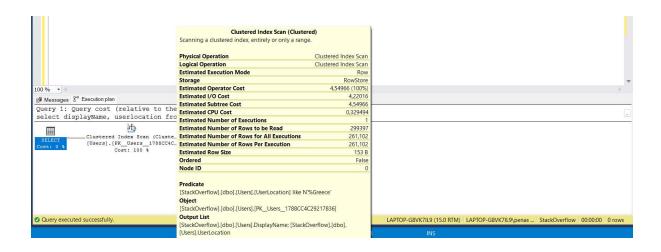


Ζήτημα Πέμπτο(10 μονάδες)

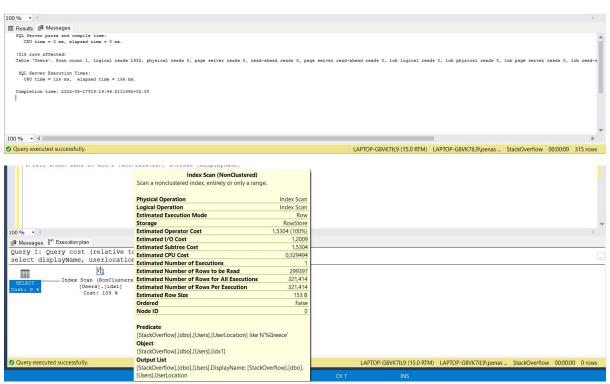
1)) Το ευρετήριο **idx1** επιταχύνει την εκτέλεση.

Πριν την δημιουργία ευρετηρίου





Μετά την δημιουργία του ευρετηρίου



2) Στην ουσία αφού το Greece βρίσκεται στο τέλος θα το διαβάζει πάντα στο τέλος.Επομένως μια λύση θα ήταν να διαγράψουμε το Greece από το τέλος και να το προσθέσουμε στην αρχή έτσι ώστε να το διαβάζει πρώτο