

5. VJEŽBA

Korisne informacije:

- Sljedeće tvrdnje nisu vidljive u dokumentaciji, ali čini se da za trenutačno korištenu verziju sustava SQL Server vrijedi sljedeće:
 - ako je automatsko kreiranje i ažuriranje statističkih podataka za bazu podataka isključeno, te niti jedan objekt sa statističkim podacima nije kreiran, SQL Server ipak, uvijek na raspolaganju ima statistički podatak $N(r)$ za svaku relaciju r
 - ako niti jedan drugi statistički podatak osim $N(r)$ nije raspoloživ, broj n -torki u rezultatu operacije $\sigma_{A=c}(r)$ procjenjuje se na $(N(r))^{1/2}$. Dakle, drugačije nego na predavanju, gdje se koristio izraz $N(r)/10$.
- U grafičkom prikazu plana izvršavanja u alatu SSMS, operacija selekcije postaje vidljiva tek kad se nad sliku operatora *Table scan* postavi kazaljka miša. Predikat selekcije tada je vidljiv pod naslovom *Predicate*.

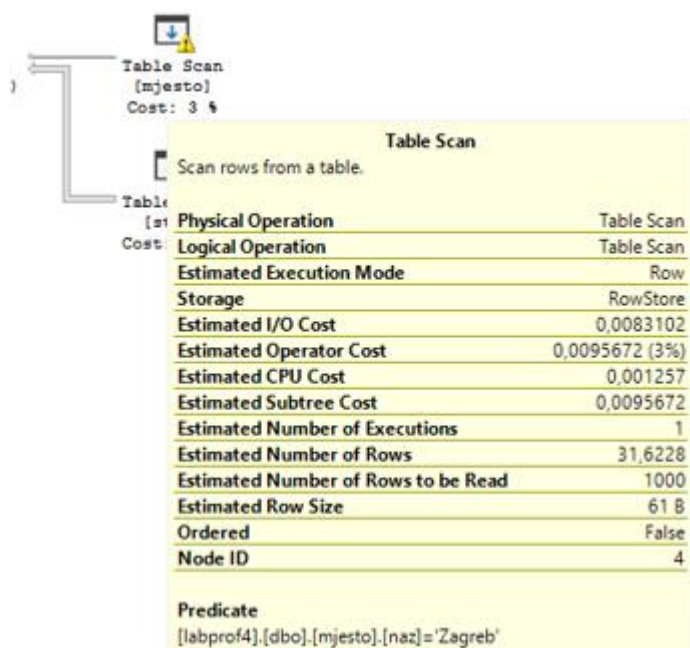


Table Scan (mjesto) Cost: 3 %	Table Scan Scan rows from a table.
Physical Operation	Table Scan
Logical Operation	Table Scan
Estimated Execution Mode	Row
Storage	RowStore
Estimated I/O Cost	0,0083102
Estimated Operator Cost	0,0095672 (3%)
Estimated CPU Cost	0,001257
Estimated Subtree Cost	0,0095672
Estimated Number of Executions	1
Estimated Number of Rows	31,6228
Estimated Number of Rows to be Read	1000
Estimated Row Size	61 B
Ordered	False
Node ID	4
Predicate	[labprof4].[dbo].[mjesto].[naz]='Zagreb'

1. zadatak: Kreirati testnu bazu podataka labprof5

Procjena potrebnog vremena: 5 minuta
U rješenjima je potrebno dostaviti: ništa

Preuzeti datoteke *.csv i izvršiti skripta labprof5.sql.

Uočiti da skripta sadrže naredbe koje isključuju automatsko kreiranja i ažuriranje statističkih podataka.

2. zadatak: Heuristička optimizacija

Procjena potrebnog vremena: 120 minuta

U rješenjima je potrebno dostaviti: definirano je uz svaki pojedinačni zadatak.

- 2.1. Nacrtati inicijalni plan izvršavanja i pripadni izraz relacijske algebre (to načiniti prema načelima s predavanja) za sljedeći upit:

```
SELECT DISTINCT mbr, prez
FROM ispit, stud, mjesto
WHERE mbr = mbri
      AND pbr = pbrs
      AND ocj = 1
      AND naz = 'Zagreb';
```

U rješenjima je potrebno dostaviti inicijalni izraz relacijske algebre i sliku inicijalnog plana izvršavanja, bez ucrtanih fizičkih operatora i bez metoda prosljeđivanja međurezultata. Slike se mogu crtati bilo kojim alatom (može i u Word-u) ili ih se može nacrtati na papiru i skenirati ili fotografirati, ali moraju biti uredne i čitljive.

- 2.2. Nad inicijalnim planom iz 2.1. "na papiru" provesti heurističku optimizaciju (prema načelima s predavanja). Na raspolaganju su sljedeći statistički podaci (tijekom optimizacije se ne može koristiti niti jedan drugi statistički podatak):

- $N(\text{mjesto}) = 1\ 000$
- $N(\text{stud}) = 5\ 000$
- $N(\text{ispit}) = 50\ 000$
- $V(\text{pbr}, \text{mjesto}) = V(\text{pbrs}, \text{stud}) = N(\text{mjesto})$
- $V(\text{mbr}, \text{stud}) = V(\text{mbri}, \text{ispit}) = N(\text{stud})$

Radi pojednostavljenja, veličinu međurezultata izražavati u n-torkama (a ne blokovima), tj. procjenu troška temeljiti na broju n-torki. Pri optimizaciji u obzir uzeti i mogućnost poboljšanja plana promjenom redoslijeda spajanja relacija ili međurezultata.

U rješenjima je potrebno dostaviti sliku plana izvršavanja nakon provedene heurističke optimizacije. Nije potrebno crtati međukorake optimizacije. U sliku upisati procijenjene veličine međurezultata, izražene u broju n-torki. Ucrtati i pretpostavljene fizičke operatore. Metode prosljeđivanja međurezultata nije potrebno ucrtavati.

- 2.3. Proučiti plan izvršavanja kojeg za upit iz 2.2 predviđa sustav SQL Server. Usporediti ga s planom iz 2.2.

U rješenjima je potrebno dostaviti sliku plana iz sustava SQL Server (*copy-paste* slike iz SSMS alata). U dopunskom tekstu navesti gdje se i s kojim predikatima obavljaju operacije selekcije (jer se na samoj slici plana iz alata SSMS taj podatak ne vidi).

Ako se plan dobiven u sustavu SQL Server razlikuje od plana iz 2.2, objasniti zašto.

- 2.4. Kreirati objekt sa statističkim podacima koji će sadržavati histogram (ali ne kreirati nikakav indeks!) za attribute mjesto.naz i ispit.ocj. Usporediti novi plan s planom iz 2.3.

U rješenjima je potrebno dostaviti naredbe s kojima su kreirani objekti sa statističkim podacima i sliku plana izvršavanja iz sustava SQL Server (*copy-paste* slike iz SSMS alata). U dopunskom tekstu navesti gdje se i s kojim predikatima obavljaju operacije selekcije (jer se na samoj slici plana iz alata SSMS taj podatak ne vidi).

Ako se plan dobiven u sustavu SQL Server razlikuje od plana iz 2.3, objasniti zašto.

2.5. Uništiti objekte sa statističkim podacima kreirane u zadatku 2.4. Proučiti plan kojeg je sustav predvidio za sljedeći upit:

```
SELECT *  
  FROM ispit, stud  
 WHERE mbr = mbri  
    AND ocj = 1;
```

U rješenjima je potrebno dostaviti sliku plana izvršavanja iz sustava SQL Server (*copy-paste* slike iz SSMS alata). U dopunskom tekstu navesti gdje se i s kojim predikatima obavljaju operacije selekcije (jer se na samoj slici plana iz alata SSMS taj podatak ne vidi).

Napisati kratko objašnjenje plana i razlog zašto je sustav za obavljanje operacije spajanja odabrao dotični fizički operator.

2.6. Kreirati indeks za atribut stud.mbr i objekt sa statističkim podacima za atribut ispit.ocj. Proučiti plan kojeg je sustav sada predvidio za upit iz zadatka 2.5.

U rješenjima je potrebno dostaviti naredbe za kreiranje indeksa i objekta sa statističkim podacima, te sliku plana izvršavanja iz sustava SQL Server (*copy-paste* slike iz SSMS alata). U dopunskom tekstu navesti gdje se i s kojim predikatima obavljaju operacije selekcije (jer se na samoj slici plana iz alata SSMS taj podatak ne vidi).

Napisati kratko objašnjenje plana i razlog zašto je sustav za obavljanje operacije spajanja odabrao dotični fizički operator.

2.7. Uništiti objekt sa statističkim podacima za atribut ispit.ocj. Proučiti plan kojeg je sustav sada predvidio za upit iz zadatka 2.5.

U rješenjima je potrebno dostaviti naredbu za uništavanje objekta sa statističkim podacima, te sliku plana izvršavanja iz sustava SQL Server (*copy-paste* slike iz SSMS alata). U dopunskom tekstu navesti gdje se i s kojim predikatima obavljaju operacije selekcije (jer se na samoj slici plana iz alata SSMS taj podatak ne vidi).

Napisati kratko objašnjenje plana i razlog zašto, iako indeks za atribut stud.mbr i dalje postoji, sustav opet koristi plan kao u 2.5, umjesto plan kao u 2.6.

Tijekom nadziranih provjera/vježbi očekuju se sljedeća znanja i vještine:

- u nekoj zadanoj bazi podataka, za neki zadani upit, napisati inicijalni izraz relacijske algebre, nacrtati plan izvršavanja, provesti heurističku optimizaciju "na papiru", usporediti s planom kojeg producira SQL Server, objasniti razloge zašto je optimizator odabrao određeni plan.