

SQL - додатне могућности

Ненад Митић

Математички факултет
`nenad.mitic@matf.bg.ac.rs`

Увод

- OLAP (енг. *OnLine Analytical Processing*) се може дефинисати као "интерактивни процес формирања, управљања, анализирања и приказа података"
- Обично се подаци са којима се ради посматрају и са њима се управља као да се чувају у вишедимензионалном низу

Агрегација података

- 1 Процес анализе захтева одређену агрегацију података, обично на различите начине и према различитим груписањима
- 2 Пример: посматрајмо следеће упите над базом
 - 1 Наћи просечне оцене на положеним испитима
 - 2 Наћи просечне оцене на положеним испитима по предметима
 - 3 Наћи просечне оцене на положеним испитима по испитном року (без обзира на годину)
 - 4 Наћи просечне оцене на положеним испитима по предметима и року (без обзира на годину рока)

Агрегација података

Одговарајући упити који су решење су

- 1 olap.primer1.sql
- 2 olap.primer2.sql
- 3 olap.primer3.sql
- 4 olap.primer4.sql

Агрегација података

Проблеми:

- прављење више сличних али незнатно различитих упита је досадно и заморно за корисника
- Извршавање свих упита може да буде јако скупа операција

Са више нивоа агрегације у једном упиту

- олакшава се посао кориснику
- нуди се могућност да се све агрегације израчунају много ефикасније (у једном пролазу)

Опција GROUPING SETS

olap.primer5.sql - комбинација упита из
olap.primer2.sql и olap.primer3.sql

```
select naziv, oznaka_roka,  
       dec(avg(ocena*1.0),4,2) as prosek  
from   ispit a, predmet b  
where  ocena >5  
and    a.id_predmeta=b.id_predmeta  
group  by grouping sets ((naziv), (oznaka_roka))
```

Резултат се приказује у облику једне табеле (која се врло тешко може назвати релацијом)

ROLLUP

GROUP BY ROLLUP се користи када је потребно да се анализира скуп података преко једне димензије, али на више нивоа детаља.

```
select naziv, oznaka_roka,dec(avg(ocena*1.0),4,2) as prosek
from   ispit a, predmet b
where  ocena >5
and    a.id_predmeta=b.id_predmeta
group  by rollup ((naziv), (oznaka_roka))
```

rollup =

grouping sets ((naziv,oznaka__roka),(naziv),())

Пример: olap.primer6.sql

ROLLUP

```
group by rollup (A,B,....,Z)=  
  grupisanje preko(A,B,....,Z)  
    (A, B, ...)  
    .....  
    (A, B)  
    (A)  
    ()
```


CUBE

GROUP BY CUBE се користи када је потребно да се анализира скуп података преко више димензија.

```
select naziv, oznaka_roka,  
       dec(avg(ocena*1.0),4,2) as prosek  
from   ispit a, predmet b  
where  ocena >5  
and    a.id_predmeta=b.id_predmeta  
group by cube ((naziv), (oznaka_roka))
```

Primer: olap.primer7.sql

CUBE

CUBE - у OLAP-у се подаци посматрају као да се налазе у ћелијама вишедимензионалног низа односно хиперкоцке

GROUP BY CUBE(A,B, ...,Z) значи груписање по сваком могућем подскупу скупа A,B,...,Z (овим упитом се добијају резултати сва 4 почетна упита)

GROUPING

Агрегатна функција GROUPING се користи заједно са GROUPING SETS, ROLLUP или CUBE и враћа вредност која је индикатор да ли слог враћен са GROUP BY клаузулом садржи или не садржи атрибут који је наведен као аргумент функције.

Пример: `olap.primer8.sql`, `olap.primer9.sql`

Агрегатна функција GROUPING

Разликовање NULL-ова у претходној табели се постиже упитом

```
select case grouping (naziv)
        when 1 then '??'
        else naziv
    end as naziv,
    case grouping (oznaka_roka)
        when 1 then '!!'
        else oznaka_roka
    end as oznaka_roka,
    dec(avg(ocena*1.0),4,2) as prosek
from    ispit a, predmet b
where   ocena >5
and     a.id_predmeta=b.id_predmeta
group  by grouping sets ((naziv), (oznaka_roka))
```

Аналитичке функције

RANK() - olap.primer10.sql

DENSE_RANK() - olap.primer11.sql

ROW_NUMBER() OVER - olap.primer12.sql,

olap.primer13.sql

FIRST_VALUE, LAST_VALUE - olap.primer14.sql

LAG, LEAD - olap.primer15.sql

OLAP AVG() - olap.primer16.sql

SQL reference V11.5 део о изразима од стране 159

*Поглавље 3 у приручнику SQL Fundamentals for IBM DB2
Scholars*

Укрштене табеле

- OLAP софтверски пакети често приказују резултате не у облику SQL табела већ у облику укрштених табела
- Укрштена табела је вишедимензиона табела која садржи вредности зависних атрибута и која је индексирана са кључним атрибутима SQL табела

Укрштене табеле

	apr	feb	jan	jun
Algebarska geometrija	9		10	
Algebarska topologija	10		10	9,7
Algebarska topologija 2			10	
Algebra 1	7,47		8,3	8,31
Algebra 1A	6,64	6,33	9	6,81
Algebra 1B	7,32	6,66	7,12	8,12
Algebra 2			8,77	7,88
Algebra 2A	8,33	10	9,5	9,6
Algebra 2B		8		9,33
Algebra 3	9,33		10	
Algebra 4	10		10	
Analiza 1	6,05		6,8	7,16

Рекурзивни SQL, LOB, MQT

Primeri:

- Рекурзивни SQL 10.primeri.sql\
{reksql.primer1, reksql.primer2, reksql.primer3}
- LOB 10.primeri.sql\LOB
- MQT 10.primeri.sql\MQT