

SQL kysely

Kyselyn yleisrakenne:

```
select tulostietomäärittely
from taulut
[where valintaehdot]
[group by ryhmitystekijät]
[having ryhmärajoitteet]
[order by järjestysperusta]
```

1

SQL-kysely

- n Tulostietomäärittelyn elementeille lasketaan normaalitapauksessa arvo jokaista valintaehdot täyttävää riviyhdistelmää kohden

2

Yhteenvetokyselyt

- n SQL:ssä on joukko yhteenvetofunktioita (aggregate function, koostefunktio)
 - n AVG keskiarvo
 - n MIN pienin arvo (minimi)
 - n MAX suurin arvo (maksimi)
 - n SUM summa
 - n COUNT lukumäärä
- n Yhteenvetofunktioita käytettäessä tulosrivia ei muodostetakaan jokaisesta valintaehdon täyttävästä riviyhdistelmästä vaan, ellei ryhmittelyä ole määritetty, muodostetaan yksi tulosrivi koko aineistosta

3

Yhteenvetokyselyt

- n Opiskelijoiden lukumäärä:
 - n `select count(*) from opiskelija;`
- n Count:n argumenttina voisi käyttää myös mitä tahansa vakiota, tulos olisi sama eli rivien lukumäärä
 - n `select count(1) from opiskelija;`
- n Jos parametrina annetaan sarake saadaan siinä olevien **ei-tyhjien** arvojen määrä
 - n `select count(Numero) from opiskelija;`

4

Yhteenvetokyselyt

- n Milloin pisimpään opiskelut helsinkiläinen opiskelija on aloittanut opintonsa?
`select min(aloitusvuosi) from opiskelija where kaupunki='Helsinki';`
- n Keskiarvoa, summaa, minimiä ja maksimia laskettaessa tyhjäarvot jätetään huomioimatta
- n Kurssien keskimääräinen opintoviikkomäärä
 - n `select avg(opintoviikot) from kurssi;`

5

Yhteenvetokyselyt

- n Yhteenvedon laskenta voidaan rajata vain keskenään erilaisiin arvoihin (projektiioon). Tällainen rajaus on yleensä järkevä vain lukumääriä laskettaessa
- n Monellako eri paikkakunnalla opiskelijat asuvat?
 - n `select count(distinct kaupunki) from opiskelija;`

6

Yhteenvetokyselyt

- Yhteenvetofunktion sisältävään kyselyyn **ei voi ottaa mukaan** dataa niiltä yksittäisiltä riveiltä, joilta funktio lasketaan
- Minkä kurssien opintoviikkomäärä on suurin?
- Ei siis ole mahdollista:**

```
select nimi, max(opintoviikot)
from kurssi;
```

yksittäisen rivin dataa
miltä riviltä nimi poimitaisiin?

7

Yhteenvetokyselyt

- Minkä kurssin opintoviikkomäärä on suurin? Toimivia vaihtoehtoja:

```
select nimi, opintoviikot from kurssi
where opintoviikot >=
ALL (select opintoviikot from kurssi);
```
- ```
Select nimi, opintoviikot from kurssi
where opintoviikot =
(select max(opintoviikot) from kurssi);
```
- ```
select nimi,maksi
from kurssi,
(select max(opintoviikot) maksi from kurssi) as m
where opintoviikot =m.maksi;
```

8

Yhteenvetokyselyt -ryhmät

- Jos kyselyyn liitetään ryhmittelymääre (Group by) muuttuu tulosrivien muodostusperiaate jälleen:
 - muodostetaan **yksi tulosrivi kutakin ryhmää kohti**
 - group by** -määreessä luetellaan sarakkeet, joiden arvojen perusteella ryhmittely tehdään
 - kaikki ne rivit, joilla on sama arvo luetelluissa sarakkeissa muodostavat ryhmän**
 - ryhmät muodostetaan sen jälkeen kun on **ensin sovellettu where-ehdot** rivien karsintaan.

9

Yhteenvetokyselyt -ryhmät

Group by A

Taulu X

A	B	C	D
1	4	6	7
1	1	4	2
1	5	5	2
2	4	8	7
2	3	5	1
3	1	5	2
3	2	4	6

Select A, sum(B) from X group by A;

A	B
1	10
2	7
3	3

10

Yhteenvetokyselyt -ryhmät

- Group by -lauseetta käytettäessä tulostietoluettelossa voi olla yhteenvetofunktioiden lisäksi **vain niitä sarakkeita**, jotka esiintyvät group by -lauseessa.
- Kaikkien ryhmittelyyn käytettyjen sarakkeiden ei tarvitse olla mukana, mutta yleensä ne ovat

```
select koodi, nimi, ryhmänro, count(*)
from kurssi, ilmoittautuminen
where ilmoittautuminen.kurssikoodi=
kurssi.koodi
group by koodi, nimi, ryhmänro;
```

koska **nimi** on tulostietolistalla sen pitää olla myös ryhmittelytekijänä vaikka ilman sitä saataisiin samat ryhmät.

11

Yhteenvetokyselyt -ryhmät

kurssi
1132
1133
1135

ilmoittautuminen
1132 1 A
1132 1 B
1132 2 C
1135 1 D
1135 1 E
1135 1 F

Pariton putoaa

Ryhmät where-ehdon täyttävien perusteella

12

Yhteenvetokyselyt -ryhmät

```
select nimi, ryhmänro, count(*)
from kurssi, ilmoittautuminen
where ilmoittautuminen.kurssikoodi=kurssi.koodi
group by nimi, ryhmänro
union
(select nimi, ryhmänro, 0
from kurssi, harjoitusryhma H,
where koodi=H.kurssikoodi and
(koodi, H.ryhmänro) not in
(select kurssikoodi, ryhmänro
from ilmoittautuminen));
```

vakioarvo

13

Yhteenvetokyselyt -ryhmät

n Edellisen voisi hoitaa myös ulkoliitoksella

```
select nimi, harjoitusryhma.ryhmänro,
count(ilmoittautuminen.ryhmänro)
from kurssi,
harjoitusryhma left outer join ilmoittautuminen on
harjoitusryhma.kurssikoodi=
ilmoittautuminen.kurssikoodi and
harjoitusryhma.ryhmänro=
ilmoittautuminen.ryhmänro
where harjoitusryhma.kurssikoodi=kurssi.koodi
group by nimi, harjoitusryhma.ryhmänro
```

Ei voi olla count(*)
Koska tyhjiäkin ryhmiä
vastaa 1 rivi

14

Yhteenvetokyselyt -ryhmät

n Ryhmäkohtaisen rivin ottamista mukaan tulokseen voidaan rajoittaa **having** määreellä.

n Having-ehto toimii kuten where-ehto, mutta se perustuu ryhmäkohtaisesti lasketun yhteenvetofunktion arvoon

n Ryhmät, joihin on ilmoittautunut yli 20 opiskelijaa

```
select nimi, ryhmänro, count(*)
from kurssi, ilmoittautuminen
where ilmoittautuminen.kurssikoodi=kurssi.koodi
group by nimi, ryhmänro
having count(*) >20;
```

15

Yhteenvetokyselyt

n Yhteenvetofunktiolla voi laskea, mutta niitä ei voi ketjuttaa (eli toinen toisen argumenttina)

n Millä kurseilla on suurin keskimääräinen ryhmäkoko, ei onnistu seuraavasti:

```
select nimi, H.ryhmänro, max(avg(count(*)))
from kurssi, harjoitusryhma H, ilmoittautuminen I
where koodi=H.kurssikoodi and
H.kurssikoodi=I.kurssikoodi and
H.ryhmänro=I.ryhmänro
group by nimi, H.ryhmänro
```

Ajettuna Oraclessa: VIRHE rivillä 1:
ORA-00937: tämä ei ole yhden ryhmän koostefunktio

16

Yhteenvetokyselyt -ryhmät

```
select koodi, nimi, opiskelijoita/ryhmiä
from kurssi,
(select kurssikoodi, count(*) opiskelijoita
from ilmoittautuminen
group by kurssikoodi) as ilm,
(select kurssikoodi, count(*) ryhmiä
from harjoitusryhma
group by kurssikoodi) as ryhm
where kurssi.koodi= ilm.kurssikoodi and
kurssi.koodi= ryhm.kurssikoodi and
opiskelijoita/ryhmiä = (XXXXX seuraavalla sivulla)
```

17

Yhteenvetokyselyt -ryhmät

n XXXXX=

```
select max(opiskelijoita/ryhmiä)
from
(select kurssikoodi, count(*) opiskelijoita
from ilmoittautuminen
group by kurssikoodi) as i,
(select kurssikoodi, count(*) ryhmiä
from harjoitusryhma
group by kurssikoodi) as ry
where i.kurssikoodi = ry.kurssikoodi
```

tyhjet kurssit eivät tule mukaan, mutta niiden keskimääräinen ryhmäkoko ei voi olla suurin

18