

Tietokantojen perusteet, K 2004

### SQL kysely

Kyselyn yleisrakenne:

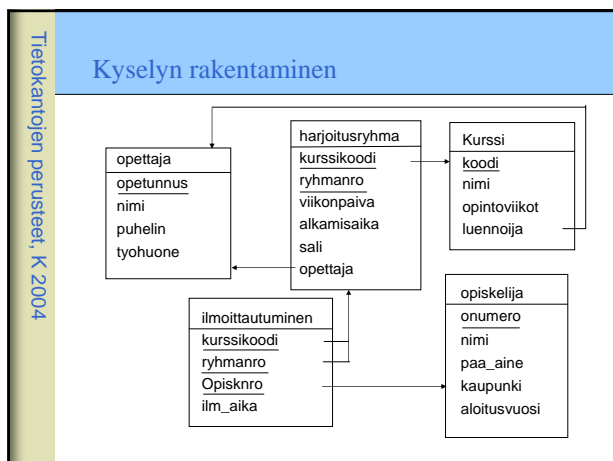
```
select tulostietomäärittely
from taulukkeet
[where valintaehdot]
[group by ryhmitystekijät]
[having ryhmärajoitteet]
[order by järjestysperusta]
```

Kysely tuottaa nimettömän tulostaulun.

Tietokantojen perusteet, K 2004

### SQL-kysely

- Tulostietomäärittelyn elementeille lasketaan normaalitapauksessa arvo jokaista valintaehdot täyttävää riviyhdistelmää kohden



Tietokantojen perusteet, K 2004

### Alikyselyt

- Alikyselyllä tarkoitetaan kyselyyn upotettua toista kyselyä. Upotettua kyselyä voidaan käyttää kyselyn from osassa mutta myös **where osassa valintaehto** operandina.
- Alikyselykin tuottaa tuloksenaan **taulun**
- Alikyselyiden käyttöön valintaehdoissa on omia predikaatteja ja lisätarkenteita, jotka määrittelevät, miten ehdon operandia sovelletaan alikyselyn tulokseen
- IN; NOT IN; tarkenteet any, some, all; EXISTS, NOT EXISTS

Tietokantojen perusteet, K 2004

### Alikyselyt

- Vuoden 92 standardissa from-osaan sallittiin normaalien taulujen myös alikyselyiden tulostaulut
- from (alikusely) [[as] alias [( sarakeluettelo)]]**
  - alikusely on normaali kysely
  - sarakeluettelo uudelleennimeää alikyselyn tulossarakkeet
  - tästä rakenteesta on hyötyä, jos halutaan yhdistää yksityiskohtaista tietoa ja yhteenvetotietoa tai eri perustein laskettuja yhteenvetotietoja samalle riville. Yksityiskohtien kyselyssä rakennetta ei tarvita.
  - yleensä rakenteen käyttö vain sotkee asioita - VÄLTÄ

Tietokantojen perusteet, K 2004

### Alikyselyt

- Opettajat, jotka luennoivat jotain kurssia

```
select nimi from opettaja
where opetunnus in
(select luennoija from kurssi)
order by nimi;
```

- Opettajat, jotka eivät luennoi mitään kurssia

```
select nimi from opettaja
where opetunnus not in
(select luennoija from kurssi)
order by nimi;
```

alikusely voidaan suorittaa erillään pääkyselystä

Tietokantojen perusteet, K 2004

### Alikyselyt

- Luennoivat opettajat, kytketyllä alikyselyllä

```
select nimi from opettaja
where
exists (select luennoija from kurssi
        where luennoija= opettaja.opetunnus)
order by nimi;
```

alikusely on evaluoitava erikseen jokaista opettajariviä kohti

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt

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

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt

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

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt

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

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt

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

Tietokantojen perusteet, K 2004

### 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?

Tietokantojen perusteet, K 2004

### 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;
```

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt -ryhmät

- Jos kyselyyn liitetään ryhmittelymääre (Group by) muuttuu tuulosrivien 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-ehtoa rivien karsintaan.

Tietokantojen perusteet, K 2004

### 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

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt -ryhmät

- Group by -lausetta 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, ryhmanro, count(*)
from kurssi, ilmoittautuminen
where ilmoittautuminen.kurssikoodi=
kurssi.koodi
group by koodi, nimi, ryhmanro;
```

**Ei anna tyhjen ryhmien ilmoittautujamäärää**

koska nimi on tulostietoluettelossa sen pitää olla myös ryhmittelytekijänä vaikka ilman sitä saataisiin samat ryhmät

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt -ryhmät

kurssi

1132
1133
1135

Pariton putoaa

ilmoittautuminen

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

Ryhmät where-ehdon täyttävien perusteella

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt -ryhmät

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

vakioarvo

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt -ryhmät

- Edellisen voisi hoitaa myös ulkoliitoksella

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

- huom: tämä syntaksi ei käy Oraclessa eikä siten myöskään Trainer-harjoituksissa

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt -ryhmät

- Oraclessa edellisen voisi hoitaa myös ulkoliitoksella seuraavasti

```
select nimi, harjoitusryhma.ryhmanro,
count ilmoittautuminen.ryhmanro)
from kurssi,
harjoitusryhma, ilmoittautuminen
where harjoitusryhma.kurssikoodi =
ilmoittautuminen.kurssikoodi (+) and
harjoitusryhma.ryhmanro =
ilmoittautuminen.ryhmanro (+) and
harjoitusryhma.kurssikoodi=kurssi.koodi
group by nimi, harjoitusryhma.ryhmanro
```

osoittavat sarakkeen johon voi tulla tyhjäarvo

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt -ryhmät

- Ryhmäkohtaisen rivin ottamista mukaan tulokseen voidaan rajoittaa **having** määreellä.
- Having-ehto toimii kuten where-ehto, mutta se perustuu ryhmäkohtaisesti lasketun yhteenvetofunktion arvoon
- Ryhmät, joihin on ilmoittautunut yli 20 opiskelijaa

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

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt

- Yhteenvetofunktiolla voi laskea, mutta niitä ei voi ketjuttaa (eli toinen on toisen argumenttina)
- Millä kurseilla on suurin keskimääräinen ryhmäko, ei onnistu seuraavasti

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

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

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt -ryhmät

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

Tietokantojen perusteet, K 2004

### Yhteenvetokyselyt -ryhmät

- XXXXX=**

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

tyhjä kurssit eivät tule mukaan, mutta niiden keskimääräinen ryhmäko ei voi olla suurin