

Koostetietokyselyt

- Distict
- Group by
- Having
- Count, Min, Max jne

Relaatiomalli



Koostetietokyselyt

- Duplikaattien poistoon tulosjoukosta on kaksi eri tapaa
 - Distinct
 - Group by (monipuolinen käytettävyydeltään)

Koostetietokyselyt

- Tehdään kysely, joka tulostaa kaikkien ankkalinnalaisten sukunimen:

`select sukunimi from ankkalinnalainen;`

```
MariaDB [ankkalinna]> select sukunimi  
-> from ankkalinnalainen;  
+-----+  
| sukunimi |  
+-----+  
| Ankkas  
| Ankkas  
| Ankkas  
| Magia  
| Hiiri  
+-----+
```

- Huomataan että sukunimi Ankkas toistuu monta kertaa

distinct

- Tehdään kysely, joka tulostaa kaikkien ankkalinnalaisten sukunimen, mutta kukin sukunimi tulee vain kerran

`select distinct sukunimi from ankkalinnalainen;`

```
MariaDB [ankkalinna]> select distinct sukunimi  
-> from ankkalinnalainen;  
+-----+  
| sukunimi |  
+-----+  
| Anka  
| Magia  
| Hiiri  
+-----+
```

group by

- Tehdään kysely, joka tulostaa kaikkien ankkalinnalaisten sukunimen, mutta kukin sukunimi tulee vain kerran:

`select sukunimi from ankkalinnalainen group by sukunimi;`

```
MariaDB [ankkalinna]> select sukunimi
-> from ankkalinnalainen
-> group by sukunimi;
+-----+
| sukunimi |
+-----+
| Anka     |
| Hiiri    |
| Magia    |
+-----+
```

- Nyt käytetään ryhmittelyä
- Tässä tapauksessa lopputulos on sama kuin edellä distinct-avulla
- Kalvoilla esitellään myöhemmin monipuolisempia tapoja käyttää group by -komentoa.

count

- Tehdään kysely, joka tulostaa lemmikin id:n ja kuinka monta omistajaa lemmikillä on:

```
select lemmikki_id, count(*)  
from omistaa  
group by lemmikki_id;
```

```
MariaDB [ankkalinna]> select lemmikki_id, count(*)  
-> from omistaa  
-> group by lemmikki_id;
```

lemmikki_id	count(*)
1	2
2	1
3	1

- Huomataan, että tarvitaan edellisen kalvon group by
- Uutena asiana opitaan count(*)

having count(*)

- Tehdään kysely, joka listaa niiden lemmikkien id:n ja omistajien lukumäärän, joilla on vähintään kaksi omistajaa:

```
select lemmikki_id, count(*)  
from omistaa  
group by lemmikki_id  
having count(*) >= 2;
```

```
MariaDB [ankkalinna]> select lemmikki_id, count(*)  
-> from omistaa  
-> group by lemmikki_id  
-> having count(*) >= 2;  
+-----+-----+  
| lemmikki_id | count(*) |  
+-----+-----+  
|          1 |         2 |  
+-----+-----+
```

- Huomataan, että tarvitaan taas edelliseen kalvojen group by
- Uutena
 - Having count(*) (tarpeen tunnistaa usein siitä, että suomen kielinen ilmaisu on "vähintään", "enintään")

min, max, avg, sum

- Tehdään kysely, joka hakee sen lemmikin nimen, jolla on suurin id:

```
select nimi  
from lemmikki  
where id in(  
    select max(id)  
    from lemmikki);
```

```
MariaDB [ankkalinna]> select nimi  
-> from lemmikki  
-> where id in (  
-> select max(id)  
-> from lemmikki  
-> );  
+-----+  
| nimi  |  
+-----+  
| Korri |  
+-----+
```

- **Huomaa, että sisäkysely on pakollinen.** Muutoin kyllä löydetään suurin id, mutta sillä ei ole korrelaatiota tulostuvaan nimeen (tämä on tyypillinen virhe)