Tietokannan hallinta, kevät 2002 **Harjoitus 1 (25.-27.3., 4.-5.4.)**

Läsnäolo: 3 tehtävää **Huom. Aikana 28.3. – 3.4. ei ole luentoja eikä harjoituksia.**

Tehtävien 1-3 tavoitteena on palauttaa lyhyesti mieleen SQL-kielen piirteitä. Ks. tarvittaessa Tietokantojen perusteet -kurssin materiaali ja harjoitukset tai Elmasri-Navathen (E&N) luku 8. SQL-kyselyt liittyvät luennolla hieman yksinkertaistettuna esitettyyn E&N:n esimerkkitietokantaan. (Kurssikansiossa on kopio alkuperäisen tietokannan määrittelystä (eroilla ei ole tässä merkitystä).)

1. a) Laadi kysely, joka tulostaa niiden työntekijöiden nimet, joilla on alaisina toisen osaston työntekijöitä.
   b) Muunna kyselyä niin, että tulostetaan myös alaisten nimet ja sekä esimiehen että alaisen osaston nimi.

2. a) Laadi kysely, joka tulostaa Houstonissa työskentelevien työntekijöiden nimet. (Onko taulu dept_locations yleensä tässä tietokannassa välttämätön vai selviävätkö toimipaikkakunnat aina projektien kautta?)
   b) Laadi kysely, joka tulostaa jokaisen työntekijän työtuntien kokonaismäärän, työntekijät aakkosjärjestyksessä.

3. Laadi kysely, joka tulostaa niiden työntekijöiden nimet, joilla on puoliso, mutta ei lapsia.

4. Pääkaupunkiseudulla on noin 1 000 000 ja (esimerkiksi) 2000 yhdistystä. Jokainen asukas on keskimäärin kolmen yhdistyksen jäsen. Oletetaan, että tietokannassa on tiedot asukkaista (ainakin henkilötunnus, nimi, asuinkunta ja osoite), yhdistyksistä (ainakin yhdtunnus, nimi, kotipaikka) sekä jäsenyyksistä (ainakin henkilö, yhdistys, liittymisvuosi).
   Hahmottele tietokannan rakennetta ja sisältöä sen verran, että voit karkeasti arvioida sen tilantarpeen.

5. Toteutetaan luennolla esitetyn yritystietokannan relaatio EMPLOYEE kasarakenteisena tiedostona (= järjestämättömänä peräkkäistiedostona). Työntekijöitä on 30 000 ja levyjakson koko on 6 KB.
   Määrittele tietueen (monikon) sisäinen esitysmuoto a) kiinteänmittaisia tietueita, b) Oraclen tietue-esitystä käyttäen. Anna kummassakin tapauksessa esimerkki Franklin Wongin tietueesta ('Franklin T Wong', 333445555, 08-DEC-45, '638 Voss, Houston, TX', 'M', 40000, 8886655555, 5).
   c) Arvioi tiedoston tilantarve jaksoina.

6. Sijoitetaan 200 000 jaksoa käsittävä tiedosto optimaalisesti levykölle, jossa on 20 levypintaa, levypinnalla 4000 uraa ja uralla 25 jaksoa.
   a) Paljonko aikaa vie koko tiedoston läpiluku, kun jaksot luetaan sijaintijärjestyksessä? Levyn kohdistusaika on 1-12 ms (keskimäärin 6 ms) ja jakson siirtoaika 0.2 ms. (Sektorivälejä ei huomioida.)
   b) Paljonko aikaa vie kaikkien jaksojen haku erikseen (mielivaltaisessa järjestyksessä)?
   c) Oletetaan, että on luettava X mielivaltaista tietuetta. Millä arvoilla X koko tiedoston läpiluku on erillisiä hakuja tehokkaampi vaihtoehto?

7. Kuinka seuraavat muutokset vaikuttavat levymuistin saantiajan (hakuajan) komponentteihin:
   a) levyn pyörimisnopeuden kaksinkertaistetaan,
   b) levyn fyysinen pakkaustiheys (uran merkkikapasiteetti) kaksinkertaistetaan,
   c) jokaiseen hakuvarteen sijoitetaan yhden sijasta kaksi luku/kirjoituspäätä.
   d) tiedoston jaksonkoko kaksinkertaistetaan?