

Tietokantojen perusteet, viikko 4

Kotona tehtävät harjoitukset tuodaan esimerkiksi paperilla laskuharjoitustilaisuuteen, jossa ne käydään läpi ryhmittäin. Harjoitukset käsitellään harjoitustilaisuudessa 8.2 alkavalla viikolla.

Tehtävä 1.

Yksi relaatiotietokantojen suunnittelun perusideoista on toistuvan tiedon toiston välttäminen (esim. nimen uudelleen ja uudelleen tallentaminen). Tähän on syynä mm. se, että toistuva tieto vie "turhaa" tilaa, jonka lisäksi toistuvan tiedon ylläpitäminen on työlästä, sillä jokaista kopiota on muokattava muutoksen yhteydessä.

Seuraavassa on esitetty Pankkitapahtuma-niminen tietokantataulu. Pohdi miltä tietokantataulun sisältö näyttäisi jos sinne lisätään useampi rivi; missä sarakkeissa on erityisesti toistuvaa tietoa?

Jaa tietokantataulu useampaan osaan siten, että suurin osa toisteisesta tiedosta vältetään.

Pankkitapahtuma(id, pankkitilin_numero, asiakkaan_etunimi, asiakkaan_sukunimi, asiakkaan_henkilötunnus, otto_vai_pano, summa, tapahtumapäivä, pankkivirkailija_etunimi, pankkivirkailija_sukunimi, tapahtuman_lisakuvaus)

Tehtävä 2.

Haluat tehdä tietokantakyselyn tehtävässä 1 esittämäsi toisteista dataa vähentävään tietokantaan. Kysely on "Anna kaikki asiakkaaseen 'Aimo Porkkana' liittyvät tapahtumat, joissa pankkivirkailijan etunimenä on ollut 'Väiski' ja summana on ollut yli 500 euroa. Aimo porkkanan henkilötunnus on '070777-177C'".

Kuinka montaa taulua joudut käymään läpi tehtävän 1 ratkaisuehdotuksessasi? Perustele edellä mainittua kyselyä esimerkkinä käyttäen (a) miksi olisi parempi pitää kaikki tieto yhdessä tietokantataulussa, ja (b) miksi olisi parempi jakaa tieto erillisiin tietokantatauluihin.

Tehtävä 3.

Tee osoitteessa <http://sqlbolt.com> olevasta interaktiivisesta SQL-oppaasta oppitunnit 8-13 ja 16.

Tehtävä 4.

Jatketaan paikan päällä tehtävistä laskuharjoituksista tutun [Chinook-tietokannan](#) parissa. Chinook kuvaa digitaalisen musiikin myyntipalvelua, ja sen tietokantakaavio löytyy osoitteesta [http://chinookdatabase.codeplex.com/wikipage?title=Chinook Schema&referringTitle=Documen](http://chinookdatabase.codeplex.com/wikipage?title=Chinook+Schema&referringTitle=Documen)

[tation](#) -- jos käytössäsi ei ole sqliteä, on sen lataussivu osoitteessa <https://www.sqlite.org/download.html> apua asennukseen löytynee netistä, esim. http://www.tutorialspoint.com/sqlite/sqlite_installation.htm voi olla hyödyksi.

Kurssia varten tietokannasta on ladattu SQLite-versio osoitteeseen

http://tietokantojen-perusteet.github.io/dbs/Chinook_Sqlite.sqlite

Ladatkaa tietokanta, ja avatkaa se SQLitessä. Kun tietokanta on avattu, antaa komento `.tables` seuraavanlaisen tulostuksen.

```
sqlite> .tables
Album          Employee      InvoiceLine   PlaylistTrack
Artist         Genre         MediaType    Track
Customer       Invoice       Playlist
```

Ennen tehtäviä kannattaa antaa komento **.headers on** jonka jälkeen sqlite näyttää tulostaulujen sarakkeiden nimet.

Toteuta SQL-kyselyt, joilla saat vastaukset seuraaviin kysymyksiin:

1. Kuinka monta kappaletta kuhunkin genreen liittyy?
2. Kuinka monta kappaletta kustakin genrestä on ostettu? Voitte olettaa, että jos lasku on olemassa, kappale on ostettu.
3. Ketkä artistit esiintyvät useimmilla levyillä? Järjestäkää artistit levymäärän mukaan ja tulostakaa top 10. Top 10 tulostukseen auttaneen avainsana "LIMIT".
4. Ketkä ovat viisi eniten kappaleita ostanutta asiakasta?
5. Mitkä 10 kappaletta ovat tuottaneet eniten rahaa? Voit olettaa, että tuotto lasketaan suoraan yksikköhinnan kautta.

Tehtävä 5.

Sormilinna-nimisen sarjakuvan piirtäjä Pertsu Jaarli haluaa mustaa huumoria sisältävälleen sarjakuvallensa web-julkaisualustan. Hän listaa järjestelmälle seuraavia vaatimuksia:

- Käyttäjien on kirjauduttava sisään nähdäkseen sarjakuvia.
- Superkäyttäjät voivat lisätä ja poistaa sarjakuvia.
- Kun sarjakuvia lisätään, niihin voidaan liittää myös teemasanoja.
- Teemasanojen perusteella pitää pystyä hakemaan sarjakuvia.
- Jokaisen sarjakuvan vieressä tulee olla "tykkään"-nappi, jota voi klikata.
- Kun käyttäjä klikkaa "tykkää"-nappia, tulee järjestelmän rekisteröidä sarjakuva josta käyttäjä tykkää. (Tätä toimintoa on tarkoitus käyttää kohderyhmien löytämiseen myöhemmin.)

Tee ylläolevasta kuvauksesta käsiteanalyysi, ja piirrä niiden perusteella käsitekaavio.

Tehtävä 6.

Tee tehtävän 5 pohjalta tietokantakaavio.

Tehtävä 7.

Luo tehtävän 6 pohjalta SQL-kyselyt, jotka luovat tietokantakaaviossa kuvatut tietokantataulut ja luo niistä sqlite3 -tietokanta.

Tehtävä 8.

Minkälaisilla kyselyillä saat selville tehtävässä 7 toteuttamastasi tietokannasta seuraavat Pertsan haluamat tiedot:

- Kuinka monta kertaa kutakin sarjakuvaa on tykätty?
- Mihin sarjakuviin liittyy avainsanat "paljastus" ja "kummitus"? Kerro kaksi erilaista tapaa tämän tiedon hakemiseen!
- Mihin sarjakuviin liittyy avainsana "paljastus", mutta ei avainsanaa "kummitus"? Kerro kaksi erilaista tapaa tämän tiedon hakemiseen!
- Ketkä 5 käyttäjää tykkäävät eniten sarjakuvista (laske tämä tykkäysten perusteella)?

Jos tietokantasi ei mahdollista em. kyselyitä, muokkaa tietokantaasi sopivasti.