

Opiskelijat, joiden äidinkieli ei ole suomi, saavat halutessaan käyttää sanakirjaa, jos siinä ei ole merkintöjä. Nämä henkilöt voivat halutessaan saada myös tehtäväpaperin sekä suomeksi että englanniksi tai ruotsiksi.

1. a) (8 p) Kauppaketju haluaa tehdä tietokannan sen tiliasiakkaista eli asiakkaista, jotka voivat maksaa ostoksensa kaupan luottokortilla ja lyhentää velkaansa kaupalle sovittujen ehtojen mukaisesti. Piirrä suunnittelua varten ER-kaavio seuraavien tietojen pohjalta. Käytä kurssilla opetettua notaatiota ja alleviivaa avainattribuutit.

Jokaisesta asiakkaasta tallennetaan yksilöivä asiakasnumero, nimi ja osoite. Asiakkaalla voi olla kauppaketjussa korkeintaan yksi luottotili eli tili, jolle asiakkaan ostokset ja maksut kirjataan. Samalla luottotilillä voi kuitenkin olla useampi omistaja (esimerkiksi avioparilla voi olla kaupassa yhteinen luottotili, mutta kumpikin parin jäsen kirjataan kaupassa erilliseksi asiakkaaksi). Luottotilistä on tiedossa sen omistajien lisäksi yksilöivä tunnus, tilin saldo, luoton korkoprosentti, luottoraja ja kuukausittainen minimilyhennys.

Kauppaketju koostuu useista myymälöistä. Jokaisesta myymälästä on tallennettu myymälän tunnus (yksikäsitteinen), nimi ja osoite. Jos asiakas tekee tarpeeksi usein ostoksia tietystä myymälästä, kirjataan hänet tämän myymälän kanta-asiakkaaksi. Sama asiakas voi olla samanaikaisesti usean eri myymälän kanta-asiakas (mutta hänellä on siis vain yksi luottotili, jolle hän voi tehdä ostoksia ketjun kaikissa myymälöissä).

Luottotilille tehdyt ostot tunnistetaan ostoksen numeron perusteella. Numero ei ole kuitenkaan yksikäsitteinen kaikissa ketjun myymälöissä, vaan vain yhden myymälän sisällä. Ostoksesta tiedetään myös sen tehnyt asiakas, myymälä, ostospäivä ja ostoksen loppusumma. Järjestelmässä pidetään kirjaa myös kaikista maksuista (velan lyhennyksistä), joita luottotilille on tehty. Niistä tiedetään päivämäärä, summa, yksilöivä viitenumero ja se, mille luottotilille maksu on tehty.

- b) (2 p) Muuta a-kohdan ER-kaavio relaatiomalliin. Esitä relaatiokaaviot ja alleviivaa niistä avainattribuutit.

2. Tehtävä liittyy seuraavaan tietokantakaavioon, joka kuvaa erään lentoyhtiön työntekijöitä. Relaatio Konetyyppi sisältää tiedot yhtiön käyttämistä lentokonetyypeistä. Attribuutti lentosade kertoo koneen lentosäteen kilometreinä ja attribuutti paikkalkm koneen matkustajapaikkojen lukumäärän. Relaatio Tyonteki ja sisältää sekä lentäjät että yhtiön muut työntekijät. Lentäjät erottaa siitä, että heillä on lentolupakirja vähintään yhteen konetyyppiin. Työntekijän attribuutti palkka kertoo kuukausipalkan euroina. Tiedot työntekijöiden lentolupakirjoista on esitetty relaatiossa Lentolupakirja. Relaatio Bonus sisältää tiedot yhtiön työntekijöille eri vuosina maksetuista bonuksista. Attribuutti maara kertoo bonuksen määrän euroina. Työntekijä voi saada yhtenä vuonna korkeintaan yhden bonuksen.

Konetyyppi(tunniste, tyyppinimi, lentosade, paikkalkm)

Tyontekija(tunnus, nimi, syntymavuosi, palkka)

Lentolupakirja(tyontekijatunnus, konetunniste)

Bonus(tyontekijatunnus, vuosi, maara)

Voit olettaa, että relaatioiden monikoiden attribuuteilla ei ole NULL-arvoja.

Kirjoita seuraavat kyselyt SQL-kielillä:

- a) (2 p) Niiden työntekijöiden tunnukset ja nimet, jotka ovat saaneet vuonna 2017 yli 3000 euron bonuksen.
- b) (2 p) Niiden lentäjien tunnukset ja nimet, joilla on lentolupa johonkin konetyyppiin, jossa on yli 250 paikkaa.
- c) (2 p) Niiden työntekijöiden tunnukset, nimet ja palkat, jotka eivät ole lentäjiä, mutta joiden kuukausipalkka on yli 5500 euroa.
- d) (2 p) Haetaan ne lentäjät, joiden bonusten summa (eri vuosilta yhteensä) on yli 25000 euroa. (Tässä siis tarkastellaan yhden lentäjän yhteensä saamia bonuksia.) Kyselyn pitää tuottaa näistä lentäjistä tunnus, nimi ja suurin tämän lentäjän saama bonus (ei siis summaa).

Kirjoita seuraavat kyselyt relaatioalgebran lausekkeina:

- e) (2 p) Niiden lentäjien tunnukset ja nimet, jotka ovat saaneet vuonna 2017 yli 3000 euron bonuksen.
- f) (2 p) Niiden lentäjien tunnukset ja nimet, joilla on lupakirja vähintään kahteen eri konetyyppiin, joissa kummassakin on yli 200 paikkaa.

3. Tarkastellaan relaatiota $R(A, B, C, D, E)$, jossa on voimassa funktionaaliset riippuvuudet $A \rightarrow B$, $B \rightarrow C D$ ja $E \rightarrow B$.
- (1 p) Perustele, miksi relaatio ei ole Boyce-Codd-normaali muodossa (BCNF:ssä).
 - (6 p) Osita relaatio Boyce-Codd-normaali muotoon käyttämällä kurssilla (ja oppikirjassa) esitettyä algoritmia. Perustele lyhyesti jokainen muodostamasi uusi relaatio. Jatka osittamista niin pitkälle, että jäljellä on vain BCNF:ssä olevia relaatioita. Perustele, miksi lopulliset relaatiot ovat BCNF:ssä.
4. Vastaa muutamalla lauseella seuraaviin kysymyksiin.
- (2 p) Mitä tarkoitetaan sillä, että transaktiot (tapahtumat, engl. transactions) ovat sarjallistuvia (engl. serializable)? Esitä esimerkki ongelmasta, joka voi seurata, jos transaktiot eivät ole sarjallistuvia.
 - (2 p) Mitä tarkoittaa se, että transaktioilta vaaditaan eheyttä (engl. consistency)? Selvennä vastauksesi lyhyellä esimerkillä.
 - (3 p) Mitä eroja on transaktioiden eristyvyystasoilla (engl. isolation levels) READ COMMITTED, REPEATABLE READ ja SERIALIZABLE? Miksi aina ei vaadita täyttä sarjallistuvuutta?
5. (4 p) Vastaa muutamalla lauseella kuhunkin seuraavista kysymyksistä.
- Hakemiston (engl. index) avulla voidaan yleensä nopeuttaa kyselyn suorittamista. Miksi ei kuitenkaan kannata tehdä hakemistoja tietokannan kaikkien relaatioiden kaikkien attribuuttien mukaan?
 - Kannattaako hakemisto sellaisen attribuutin mukaan, jolla tyypillisesti hyvin monella relaation monikolla (engl. tuple) on sama arvo tälle attribuutille, esimerkiksi yliopiston opiskelijatietokannassa hakemisto opiskelijan syntymävuoden mukaan (perustele vastauksesi)?
 - Oletetaan, että tehtävän 2 esimerkkietokannassa relaatio Tyontekija on tallennettu 180 levysivulle ja että keskimäärin kahdeksalla lentoyhtiön työntekijällä on sama syntymävuosi. Työntekijät ovat relaatiossa satunnaisessa järjestyksessä. Tarkastellaan seuraavaa kyselyä: haetaan kaikki työntekijät, joilla on määrätty syntymävuosi. Kuinka paljon tämän kyselyn suorittaminen nopeutuu, jos relaatiolle Tyontekija on olemassa hakemisto työntekijän syntymävuoden (attribuutin syntymavuosi) mukaan? Karkea arvio riittää, mutta perustele vastauksesi lyhyesti. Tässä ei tarvitse tarkastella mitään muita tietokantaoperaatioita kuin tätä yhtä kyselyä.