

Laskinten tai lisämateriaalin käyttö tentissä ei ole sallittua

**TÄRKEÄÄ: ARVOSTELUN NOPEUTTAMISEKSI KIRJOITA JOKAISEN VIIDEN TEHTÄVÄN (1, 2, 3, 4 JA 5) RATKAISU OMALLE VASTAUSPAPERILLEEN JA PALAUTA SE TÄTÄ TEHTÄVÄÄ VASTAAVAAN PINOON. Älä unohda kirjoittaa nimeäsi ja opiskelijanumeroasi jokaiseen palauttamaasi paperiin.**

1. a) (8 p) Tee seuraavan kuvauksen pohjalta UML-kaavio tilauksia kuljetusroboteilla toimittavaa kauppaketjua varten. Kauppaketjuun kuuluu useita myymälöitä, joista asiakkaat voivat tilata toimitettavaksi kuljetusrobotilla kunkin myymälän valikoimaan kuuluvia tuotteita. Tuotteiden valikoima voi vaihdella eri myymälöissä, mutta sama tuote voi kuulua usean eri myymälän valikoimaan. Kuhunkin tilaukseen voi sisällyttää vain samasta myymälästä saatavia tuotteita, ja tilauksen toimittaa tämän myymälän käyttämä robotti. Robotti toimittaa aina vain yhden tilauksen kerrallaan, mutta tilaukseen voi kuulua useita tuotteita.

Jokaisesta myymälästä tiedetään sen nimi (yksikäsitteinen), osoite, puhelinnumero, mitä tuotteita tämän myymälän valikoimaan kuuluu ja mitä robotteja on tämän myymälän käytössä. Jokaisesta robotista tiedetään sen numero ja tila, joka voi olla esimerkiksi käytettävissä (robotti on kaupassa valmiina odottamassa toimitettavaa tilausta), pakattavana (robottiin ollaan pakkaamassa tilausta), toimituksessa (robotti on viemässä tilausta asiakkaalle), palaamassa (robotti on palaamassa asiakkaan luota takaisin myymälään) tai epäkunnossa. Robottien numerot ovat yksikäsitteisiä saman myymälän sisällä, mutta kahden eri myymälän roboteilla voi olla keskenään sama numero. Robotti on aina vain yhden myymälän käytössä.

Tuotteilla on yksikäsitteinen tuotenumero ja sen lisäksi nimi, hinta ja pakkauksen koko. Tietokannassa on tieto siitä, minkä myymälöiden valikoimaan tuote kuuluu ja kuinka monta kappaletta sitä on tällä hetkellä saatavissa kussakin näissä myymälöissä. Tuotteella on sama hinta kaikissa myymälöissä. Asiakkaista tiedetään asiakasnumero (yksikäsitteinen), nimi, osoite, sähköpostiosoite ja puhelinnumero. Tilauksista tiedetään tilausnumero (yksikäsitteinen), tilausaika, tilauksen tila (esim. keräilyssä, odottamassa vapaata robottia, toimituksessa tai toimitettu), mitä tuotteita tilaukseen kuuluu ja kuinka monta kappaletta kutakin tuotetta, tilauksen tehnyt asiakas ja tilauksen myymälä sekä tilausta toimittava robotti. Robotti ei ole välttämättä tiedossa siinä vaiheessa, kun tilaus tehdään, vaan tieto robotista voidaan tallentaa tietokantaan vasta myöhemmin, kun jokin robotti vapautuu tilauksen toimitusta varten. Tietokannassa säilytetään tiedot myös jo toimitetuista tilauksista esimerkiksi tilastointia ja mahdollisten valitusten käsittelyä varten.

Tehtävän yksinkertaistamiseksi UML-kaaviossa ei tarvitse mallintaa robottien reittejä tai muita sellaisia asioita, joista edellä ei ole kerrottu.

b) (2 p) Muuta a-kohdan UML-kaavio relaatiomalliin. Esitä relaatiokaaviot ja alleviivaa niistä avainattribuutit (eli PRIMARY KEY -määreellä merkittävät attribuutit).

2. Tehtävä liittyy seuraavaan tietokantakaavioon, joka kuvaa erään kauppaketjun asiakkaita, myymälöitä, asiakkaiden ostoja ja osallistumisia erilaisiin kaupan järjestämiin kampanjoihin.

**Tietokantakaavio:**

Customers(number, name, address, born)  
 Stores(name, address)  
 Purchases(custnumber, storename, time, value)  
 Campaigns(code, name, year)  
 Takenpart(custnumber, campaigncode, time)

Taulu Customers sisältää kunkin asiakkaan asiakasnumeron, nimen, osoitteen ja syntymävuoden. Taulu Stores sisältää kunkin ketjun myymälän nimen ja osoitteen. Ketjulla ei ole kahta samannimistä myymälää. Taulun Purchases yksi monikko kuvaa jonkun asiakkaan yhtä ostoskertaa jossain kauppaketjun myymälässä. Attribuutti time sisältää tiedon ostokerran päivämäärästä ja kellonajasta, ja attribuutti value tiedon ostokerralla ostettujen tuotteiden yhteishinnasta. Taulu Campaigns sisältää tiedot kauppaketjun alennusmyynti- ja muista vastaavista kampanjoista: kampanjan tunnuksen, nimen ja vuoden. Taulu Takenpart sisältää tiedon siitä, että tietty asiakas on osallistunut tiettyyn kampanjaan (esimerkiksi ostamalla tuotteen kampanjaan kuuluvasta alennusmyynnistä). Asiakas voi osallistua samaan kampanjaan useita kertoja esimerkiksi ostamalla yhden kampanjatuotteen keskiviikkona ja toisen saman viikon perjantaina.

Ostosten arvot ovat desimaalilukuja ja asiakkaiden syntymävuodet sekä kampanjoiden vuodet kokonaislukuja. Muut attribuutit ovat merkkijonoja. Voit olettaa, että taulujen monikoiden attribuuteilla ei ole NULL-arvoja.

**Kirjoita seuraavat kyselyt SQL-kielillä:**

- a) (2 p) Niiden asiakkaiden asiakasnumerot ja nimet, jotka ovat joskus tehneet yli 200 euron arvoisen ostoksen Halpis-Tapiola-nimisessä liikkeessä. (Tässä siis tarkastellaan yhtä ostoskertaa, ei useiden ostokertojen summaa.)
- b) (2 p) Niiden asiakkaiden asiakasnumerot ja nimet, joiden osoite sisältää merkkijonon 'Espoo' ja jotka ovat osallistuneet vuoden 2022 Suuri Hullutus -nimiseen kampanjaan.
- c) (2 p) Niiden asiakkaiden asiakasnumerot ja nimet, jotka ovat joskus ostaneet jotain Halpis-Tapiola-nimisestä myymälästä, mutta jotka eivät ole koskaan tehneet yhtään yli 100 euron arvoista kertaostosta ketjun mistään muusta myymälässä.
- d) (2 p) Haetaan asiakas, joiden kertaostosten arvojen keskiarvo on yli 150 euroa. Kyselyn pitää laskea näistä asiakkaista, kuinka monessa eri ketjun myymälässä he ovat tehneet ostoksia. Kyselyn pitää tuottaa asiakkaan asiakasnumero, nimi, kertaostosten arvojen keskiarvo ja em. myymälöiden lukumäärä. Keskiarvo ja myymälöiden lukumäärä lasketaan siis erikseen jokaiselle asiakkaalle.



**Kirjoita seuraavat kyselyt relaatioalgebran lausekkeina:**

e) (2 p) Kaikkien niiden kampanjoiden nimet ja vuodet, joihin Anne Asiakas -niminen asiakas (tai asiakkaat, jos tämänimisiä asiakkaita on useita) on osallistunut.

f) (2 p) Niiden asiakkaiden asiakasnumerot ja nimet, jotka ovat osallistuneet täsmälleen yhteen (ei kahteen tai useampaan) ketjun kampanjaan.

3.

a) (2 p) Oletetaan, että tehtävän 2 tietokanta on niin suuri, että se ei mahdu tietokoneen keskusmuistiin. Tietokannassa halutaan usein laskea tilastoja tehdyistä ostoksista myymälöittäin, esimerkiksi määrättyssä myymälässä tehtyjen kaikkien ostosten arvojen summa halutulla aikavälillä. Missä tilanteissa taululle Purchases kannattaa tehdä hakemisto (engl. index) attribuutin storename mukaan? Koska se ei kannata? Perustele vastauksesi muutamalla lauseella.

b) (2 p) Kerro muutamalla lauseella, mikä on laukaisin (engl. trigger) ja mikä on sen toimintaperiaate. Periaatteen kertominen riittää. Sinun ei tarvitse kirjoittaa yhtään laukaisinta tai esittää sen tarkkaa syntaksia.

4. Tarkastellaan relaatiota  $R(A, B, C, D, E)$ , jossa on voimassa funktionaaliset riippuvuudet  $A \rightarrow B$ ,  $B \rightarrow D$  ja  $E \rightarrow D$ .

a) (1 p)  
Perustele, miksi relaatio ei ole Boyce-Codd-normaali muodossa (BCNF:ssä).

b) (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.

5. (7 p) Tarkastellaan ravintolan tietokantaa, johon ravintola kirjaa tiedot tehdyistä varauksista, niihin sisältyvistä pöydistä (yhteen varaukseen voi kuulua usean pöydän varaus) ja varauksen tehneistä asiakkaista. Tietokannassa on seuraavat taulut:

```
Customers(custID, name, phone, email)
Tables(tableNo, maxSeats, window, accessible)
Reservations(resID, custID, date, startingTime, endingTime)
InReservation(resID, tableNo, seatsReserved)
```

Taulun Tables attribuutti tableNo on pöydän numero ja attribuutti maxSeats kertoo, kuinka monta henkilöä pöytään korkeintaan mahtuu. Attribuutit window ja accessible kertovat, onko kysymyksessä ikkunapöytä ja sopiiko pöytä myös pyörätuolin tai muiden apuvälineiden käyttäjille.

Taulu Customers sisältää tiedot varauksia tehneistä asiakkaista ja taulu Reservations heidän varauksistaan. Yhteen varaukseen voi kuulua useita pöytiä. Tauluun InReservation tallennetaan tieto siitä, että jokin pöytä kuuluu määrättyyn varaukseen ja kuinka monta paikkaa pöydästä on suunniteltu tässä varauksessa käytettäväksi.

Kerro lyhyesti, mitkä ovat transakzioilta (engl. transaction) vaadittavat ominaisuudet (ns. ACID-ominaisuudet) tietokannan hallintajärjestelmässä. Esitä jokaisesta ominaisuudesta jokin virhetilanne, joka voisi tapahtua annetussa esimerkkitietokannassa, jos tämä ominaisuus ei ole voimassa. C-kirjainta vastaavan ominaisuuden (suomeksi vastaavan ominaisuuden nimi alkaa E-kirjaimella) osalta esitä kolme erityyppistä virhetilannetta. Muiden kolmen ominaisuuden osalta riittää jokaisesta yhden esimerkin esittäminen.

**Muista vastata kurssin palautekyselyyn viimeistään 1.6.2023. Kyselyyn vastaamisesta saa 20 harjoitustehtävapistettä. Linkki kyselyyn on lähetetty kurssille ilmoittautuneille 12.5. sähköpostitse.**