

Käsiteanalyysi

Kun alamme suunnittelemaan uutta tietokantaa ei sen rakenne ole vielä tiedossa. Suunnittelu alkaa käyttäjän tarpeista ja etenee pala palalta kohti tietorakennetta. Tätä prosessia sanotaan käsiteanalyysiksi.

Käsiteanalyysin vaiheet

1. Kirjataan tietokantaan tarvittavat ominaisuudet ja asiat sanallisesti ylös.
2. Tunnistetaan ja listataan kuvauksesta käsitteet, joista voi tulla:
 - a. tauluja (entiteetit)
 - b. taulujen sarakkeita/kenttiä (ominaisuudet)
 - c. taulujen välisiä yhteyksiä (relaatiot)
3. Normalisoidaan lista.
4. Muodostetaan taulurakenne listasta.

Käsiteanalyysiesimerkki

Asiakkaan antama kuvaus tietokannan toteuttajalle:

Olisi hyvä, jos meidän liikkeen autojen perustiedot saataisiin tallennetuksi tietokantaan. Meillä on useita eri automerkkejä sekä autojen malleja. Myyjien pitää tietää meidän autoista tietokannan avulla auton väri, hinta sekä ovien lukumäärä. Meillä on myynnissä myös bensiini- ja diesel-autojen lisäksi mm. sähköautoja.

Käsiteanalyysi

Otetaan kuvaus ja listataan käsitteitä, joita tietokantaan voisi tallentaa.

Mietitään, mitkä tiedot toistuvat useaan kertaan ja muodostetaan niistä omat kokonaisuudet, eli taulut.

Analyysin tulos

Autot:

- auton merkki (toistuva arvo → taulu)
- auton malli (toistuva arvo → taulu)
- väri (toistuva arvo → taulu)
- hinta (yksilöllinen, koska numero → sarake)
- ovien lukumäärä (yksilöllinen, koska numeroarvo TAI toistuva, jos itseasiassa korityyppi)
- bensiinautot (toistuva arvo → omaan tauluun)
- dieselautot (toistuva arvo → omaan tauluun)
- sähköautot (toistuva arvo → omaan tauluun)

Normalisointi

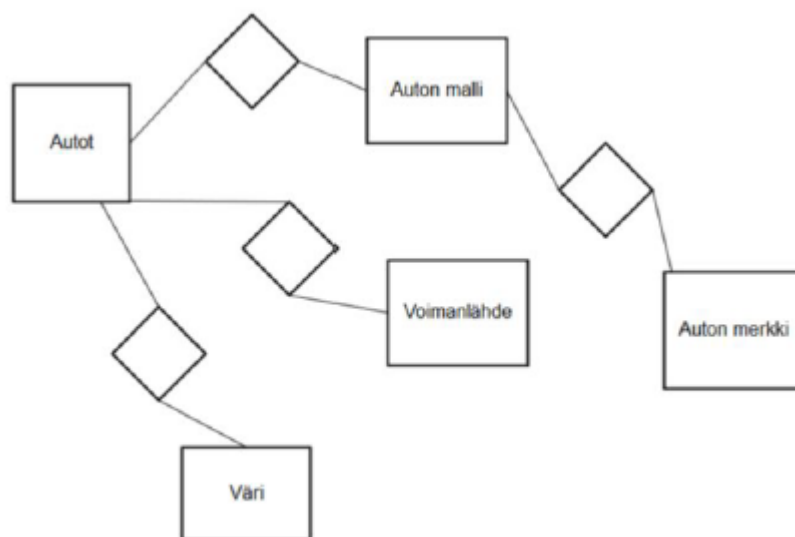
Käsiteanalyysi, normalisointi valmis (Kokonaisuudet ryhmitelty tauluiksi ja

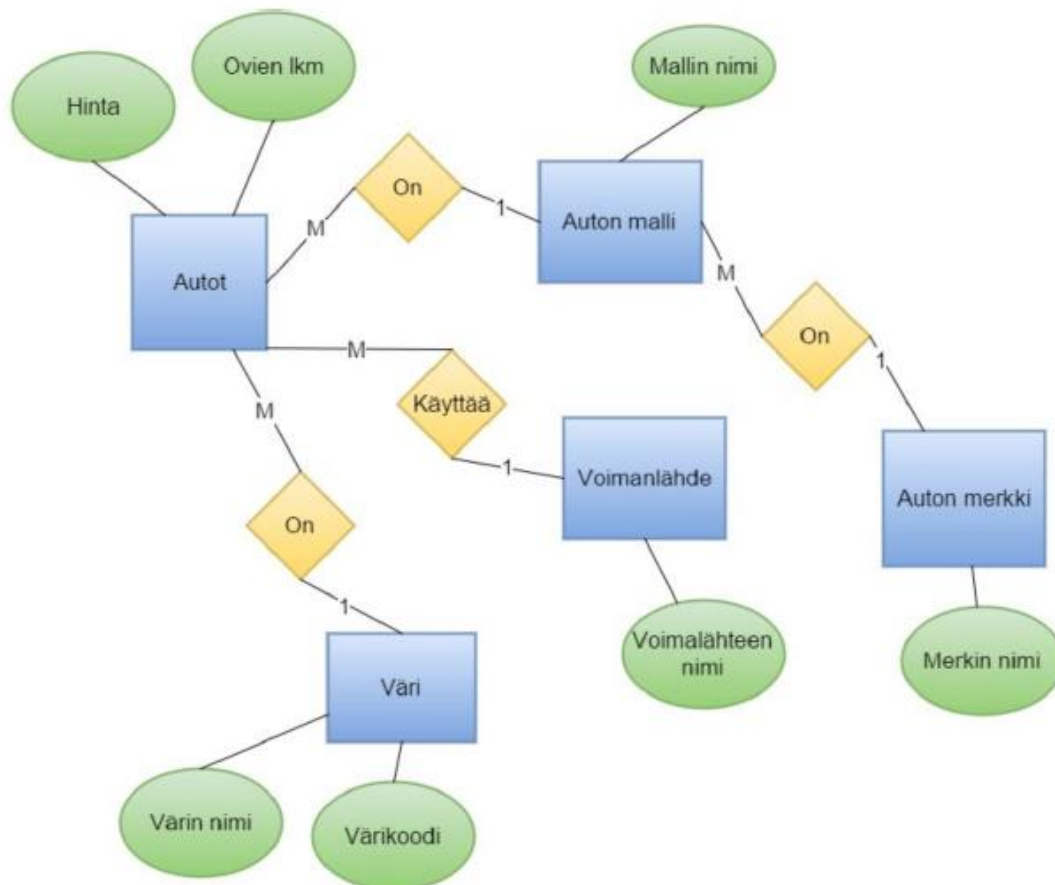
tarvittavat sarakkeet lisätty)

Autot:

- hinta
- ovien lukumäärä
- Voimanlähde:
- voimanlähteen nimi
- Auton merkki:
- auton merkin nimi
- Auton malli
- auton mallin nimi
- Väri
- värin nimi
- värikoodi(?)

Normalisoitu tietokantarakenne ER-kaavioksi





Normalisoinnin teoria

Teoria: Normalisointi

Kuten edellä olevassa esimerkissä näkyi, normalisoinnin jälkeen emme enää listanneet kolmea erilaista autoa (benssiini, diesel ja sähkö) vaan nämä olivat yhdistyneet yhdeksi entiteetiksi *Voimanlähde*.

Edelleen *Väri* taululla oli nyt ominaisuus/sarake *Värin nimi* sekä *Värikoodi*. Vastaavia muutoksia oli tapahtunut muualla tietokannan määrittelyssä.

Nämä muutokset ovat seurausta tietokantarakenteen normalisoinnista.

1. Selvitä mitä tarkoittaa tietokannan normalisointi.
2. Selvitä mitä tarkoittaa
 - a. ensimmäinen normaalimuoto,
 - b. toinen normaalimuoto,
 - c. kolmas normaalimuoto ja
 - d. neljäs normaalimuoto.
3. Selvitä mitä tarkoittaa tietokannan denormalisointi.

- a. Milloin sitä käytetään?
- b. Mitä hyötyä ja haittaa siitä on?

Linkkejä:

<https://tietokantojen-perusteet-19.mooc.fi/osa-4/1-tietokannan-normalisointi>

<https://tikape-k20.mooc.fi/luku-5/1>

https://fi.wikipedia.org/wiki/Tietokannan_normalisointi

<http://appro.mit.jyu.fi/tiedonhallinta/luennot/luento4/>