

# Keskiluvut

- Tilastoaineistot ovat usein suuria, joten tiedon tiivistämiseksi aineistosta lasketaan usein erilaisia *tunnuslukuja*
- Tärkeimpiä tunnuslukuja ovat seuraavat *keskiluvut*:
- *Keskiarvo*  $\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$
- *Tyyppiarvo* eli *moodi* on se luku, jota havaintoaineistossa on eniten
- *Mediaani* on suuruusjärjestyksessä keskimäinen luku, tai kahden keskimäisen luvun keskiarvo, jos aineistossa on parillinen määrä lukuja.

# Tilastomuuttujat

- Tilastoitavien suureiden eli *tilastomuuttujien* lajit:
- *Diskreetti* tilastomuuttuja
  - Voi saada vain erillisiä arvoja
  - Esimerkiksi kappalemäärät, kengän koko, jne.
  - Voi olla joko *numeerinen* tai ei numeerinen
- *Jatkuva* tilastomuuttuja
  - Voi saada (ainakin periaatteessa) mitä tahansa arvoja tietyltä väliltä
  - Esimerkiksi pituus, ikä, jne.

# Asteikot

- Jos tilastomuuttujia ei voi asettaa suuruusjärjestykseen, niitä tarkastellaan vain *laatueroasteikoilla*
  - Aineistosta voidaan määrittää vain tyyppi-arvo
  - Esim. lempiruoka kyselytutkimuksessa
- Jos muuttujat voidaan asettaa suuruusjärjestykseen järkevällä tavalla, niin kyseessä on *järjestysasteikko*
  - Voidaan määrittää myös mediaani
  - Esim. yo-arvosanat, vaatekoot XS, S, M, L, XL, ...

# Asteikot

- *Välimatka-asteikossa* on määritelty suuruusjärjestyksen lisäksi myös muuttujien mahdollisten arvojen välimatka
  - Voidaan laskea myös keskiarvo
  - Esim. lämpötila celsius-asteikolla
- Jos asteikolle voidaan määrittää nollakohta mielekkäällä tavalla (usein pienin mahdollinen arvo), niin muuttujien keskinäisiä suhteita voidaan laskea järkevästi
- Tällöin on kyse *suhdeasteikosta*
  - Esim. pituus, massa, lämpötila kelvin-asteikolla mitattuna