

## Tilastotieteen peruskäsitteitä

-tilastotietoa kerätään jäsentääksemme tutkittavaa asiaa/ilmiötä, ja tilastotietoja voidaan havainnollistaa tilastollisilla tunnusluvuilla (moodi, mediaani, keskiarvo, keskihajonta, vaihteluväli) ja kuvaajilla/taulukkoilla.

-tilastoyksikkö on se yksikkö, jota tilastointi koskee. Esimerkiksi väestöä tutkittaessa yksi ihminen.

-tutkimuksen perusjoukko on se populaatio, joka sisältää kaikki tilastoyksiköt. Esim. Väestössä kaikki ihmiset.

-otanta: Jos tilastotiedon kerääminen kaikilta tilastoyksiköiltä ei onnistu tai on kohtuuttoman työlästä, tehdään tällöin otanta perusjoukosta, jonka tulisi olla luotettava ja edustaa koko perusjoukon ominaisuuksia. Otanta tulee tehdä siten, että siihen valitut perusyksiköt eivät ole valittu minkään ominaisuuden (alue, sukupuoli, oppilaitos, yms.) perusteella vaan otanta on valittu sattumanvaraisesti

## Tilastollisia tunnuslukuja

Moodi: -muuttujan arvojen (=havaintoarvot) tyypillisin arvo

Mediaani: -muuttujan järjestetyn arvojoukon keskimmäisin arvo tai kahden keskimmäisimmän keskiarvo

Keskiarvo: -muuttujan saamien arvojen summa jaetaan niiden lukumäärällä

Vaihteluväli: -muuttujan saaman suurimman ja pienimmän arvon erotus

**Esimerkki**: Ohessa on taulukoitu matikan kurssiarvosanat frekvenssitaulukkona. Määritä keskiarvo,  $M_o$ ,  $M_d$  ja vaihteluväli. Piirrä tilastosta pylväsdiagrammi.

Teht. Alk. S. 20 t.1-2, 11, 15, 18.