

Tehtävä 1

a)

Suurimman muuttujan arvon 5 summafrekvenssi on 100, joten tutkimukseen vastasi 100 suomalaista.

b)

Tutkimuksen perusjoukko on se joukko, josta halutaan kerätä tietoa, eli kaikki Suomessa asuvat henkilöt.

c)

Tutkimukseen on valittu vain 100 perusjoukon havaintoyksikköä, joten kyseessä on otos.

d)

Muuttujan 0 frekvenssi on 11 %, joten 11 % ei matkusta kertaakaan ulkomaille.

e)

Korkeintaan kaksi matkaa tekee 72 %, sillä se on muuttujan arvon 2 suhteellisen summafrekvenssin arvo.

Tehtävä 2

a) Määritetään pienin ja suurin arvo.

Pienin arvo	34
Suurin arvo	998

Vaihteluvälin pituus on $998 - 34 = 964$.

b)

Komentojen MOODI ja MEDIAANI avulla.

Moodi	648
Mediaani	462

$M_o = 648$ ja $M_d = 462$.

c)

Käytetään otoskeskihajontaa, jota kirjassa on käytetty.

Keskiarvo	493,020202
Keskihajonta	279,5667194

$\bar{x} \approx 493$ $s \approx 280$

d) keskihajonta kuvaa kuinka paljon keskimääräisesti muuttujien arvot poikkeavat keskiarvosta. Nyt keskihajonta on melko suurta eli arvot ovat keskittyneet vähän kauemmas keskiarvosta.

Keskiarvo kuvaa muuttujan arvon keskimääräistä suuruutta. Keskiarvo on nyt lähellä vaihteluvälin keskusta eli arvot ovat jakautuneet melko tasaisesti vaihteluvälin keskuksen molemmille puolille.

Tehtävä 3

a-c

a)

Lasketaan lasten lukumäärä.

$$1 \cdot 238258 + 2 \cdot 218750 + 3 \cdot 71378 = 889892$$

b)

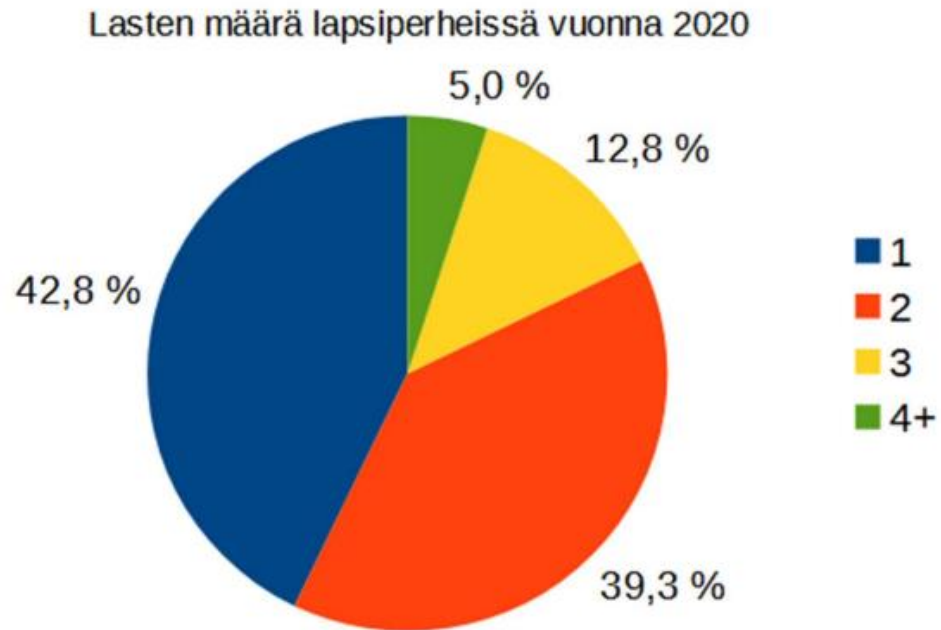
Lasten määrä	f	f%	sf	sf%
1	238258	42,8	238258	42,8
2	218750	39,3	457008	82,2
3	71378	12,8	528386	95,0
4+	27666	5,0	556052	100,0

c)

Suurin frekvenssi on muuttujan arvolla 1, joten $M_o = 1$ (lapsi)

Suhteellinen summafrekvenssi ylittää 50 % ensimmäisen kerran muuttujan arvon 2 kohdalla, joten $M_d = 2$ (lasta).

Tehtävä 3 d-f



e) Diskreetti muuttuja voi saada vain tiettyjä arvoja ja jatkuva muuttuja mitä vain arvoja joltain väliltä. Esimerkiksi metsässä tietyllä pinta-alalla kasvavien puiden lukumäärä on diskreetti ja niiden paksuus jatkuva muuttuja.

f) Aineisto kannattaa luokitella esimerkiksi silloin kun muuttuja voi saada todella montaa eri arvoa jolloin frekvenssitaulukko olisi liian iso. Luokittelu mahdollistaa tällaisen aineisto käsittelyn/tunnuslukujen laskemisen helpommin käyttämällä luokkien luokkakeskuksia.