Binomi 5 – Luku 6 – Tehtävien malliratkaisut

6.1

a)

Vaaka-akselin asteikolta nähdään, että pienin arvo on 17 (%).

b)

Vaaka-akselin asteikolta nähdään, että suurin arvo on 27 (%).

c)

Vaihteluväli ilmaisee pienimmän ja suurimman arvon. Vaihteluväli on siis 17–27 (%).
Vaihteluväli voidaan merkitä myös hakasulkumerkinnällä: [17 %, 27 %].

d)

Suurimman ja pienimmän arvon erotus on 27 – 17 = 10.

Vaihteluvälin pituus on 10 (prosenttiyksikköä).

Vastaus:

a) Min = 17 %

b) Max = 27 %

c) [17 %, 27 %]

d) Δ = 10 (prosenttiyksikköä)

6.2

Keskihajonta kuvaa sitä, kuinka tiiviisti arvot ovat sijoittuneet keskiarvon läheisyyteen.

Keskihajonta on

- pieni, kun arvot ovat sijoittuneet pääosin lähelle keskiarvo.
- suuri, kun arvot ovat sijoittuneet pääosin kauas keskiarvosta.

Jakauma A:

- Jakauma on melko symmetrinen, joten keskiarvo on suurin piirtein vaihteluvälin [4, 10] puolivälissä eli $\bar{x} \approx 7$.
- Pienten ja suurten arvojen kohdalla on korkeat pylväät. Keskiarvosta kaukana olevia arvoja on paljon.
- Jakauman keskellä on matalat pylväät. Keskiarvon lähellä olevia arvoja on vähän.
- Keskihajonta on siis melko suuri.

Jakauma B:

- Jakauma on melko symmetrinen, joten keskiarvo on suurin piirtein vaihteluvälin [4, 10] puolivälissä eli $\bar{x} \approx 7$.
- Jakauman keskellä on korkeat pylväät. Keskiarvon lähellä olevia arvoja on paljon.
- Pienten ja suurten arvojen kohdalla on matalat pylväät. Keskiarvosta kaukana olevia arvoja on vähän.
- Keskihajonta on siis melko pieni.

Jakauma C:

- Jakauma on melko symmetrinen, joten keskiarvo on suurin piirtein vaihteluvälin [4, 10] puolivälissä eli $\bar{x} \approx 7$.
- Pylväät ovat melko tasakorkuiset eli kaikkia arvoja on suunnilleen yhtä paljon. Lähellä keskiarvoa olevia arvoja on suurin piirtein yhtä paljon kuin kaukana keskiarvosta olevia arvoja.

Havaintojen perusteella keskihajonta on suurinta jakaumassa A ja pienintä jakaumassa B.

Keskihajonnan mukainen suuruusjärjestys pienimmästä suurimpaan on B < C < A.

Vastaus:

B, C, A

6.3

Tunnusluvut voidaan määrittää GeoGebran taulukko-sovelluksella tai taulukkolaskentaohjelmalla. Katso tarvittaessa ohjeet Esimerkin 1 videolta.

Tapa 1. GeoGebra

Valitaan elokuvien pituudet. Valitaan Yhden muuttujan analyysi –työväline.

Tilastot

n	40
Keskiarvo	103.025
σ	19.6894
s	19.9403
Σx	4121
Σx²	440073
Min	62
Q1	89
Mediaani	101.5
Q3	119.5
Max	150

GeoGebran tulosteessa on kaksi keskihajontalukua: σ ja *s*.

Näissä ratkaisuissa käytetään sitä hajontalukua, joka on merkitty tulosteeseen kirjaimella *s*.

Luetaan kysytyt tunnusluvut ohjelman tulosteesta:

a)

Pienin arvo on Min= 62 (minuuttia) ja suurin arvo Max = 150 (minuuttia).
Vaihteluväli on [62 min, 150 min].
Vaihteluvälin pituus on 150 – 62 = 88 (min).

b)

Keskiarvo on $\bar{x} = 103{,}02 \ldots \approx 103$ (min).
Keskihajonta on $s = 19{,}94\ldots \approx 20$ (min).

Tapa 2. Taulukkolaskenta

Valitaan elokuvien pituudet. Valitaan Data – Tilastotiedot – Tunnusluvut.

	Sarake 1
Keskiarvo	103,025
Keskivirhe	3,152835008
Moodi	99
Mediaani	101,5
First Quartile	89,5
Kolmas neljännes	119,25
Varianssi	397,6147436
Keskihajonta	19,94027943
Kurtoosi	-0,142653628
Vinous	0,297624373
Alue	88
Minimi	62
Maksimi	150
Summa	4121
Lukumäärä	40

Luetaan kysytyt tunnusluvut ohjelman tulosteesta:

a)

Pienin arvo on 62 minuuttia ja suurin arvo 150 minuuttia.
Vaihteluväli on [62 min, 150 min].
Vaihteluvälin pituus on 150 – 62 = 88 (min).

b)

Keskiarvo on $\bar{x} = 103{,}02 \ldots \approx 103$ (min).
Keskihajonta on $s = 19{,}94\ldots \approx 20$ (min)

Vastaus:

a) Vaihteluväli on [62 min, 150 min] ja $\Delta = 88$ min.

b) $\bar{x} \approx 103$ (min) ja $s \approx 20$ (min)

6.4

Ratkaistaan tehtävä GeoGebran taulukko-sovelluksella.
Katso tarvittaessa ohjeet Esimerkin 2 videolta.

GeoGebra 6

Ohje: Kirjoita arvot ja niiden frekvenssit taulukkoon.
Valitse aineisto ja Yhden muuttujan analyysi -työväline.

	A	B
1	0	85
2	1	71
3	2	65
4	3	35
5	4	9
6	5	7

Yhden muuttujan analyysi

Tilastot

n	272
Keskiarvo	1.386
σ	1.2754
s	1.2778
Σx	377
Σx²	965
Min	0
Q1	0
Mediaani	1
Q3	2
Max	5

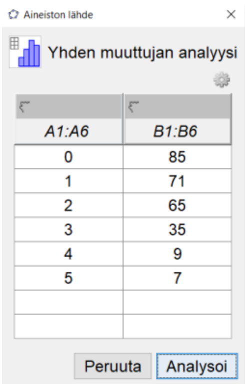
GeoGebra 5

Ohje: Kirjoita arvot ja niiden frekvenssit taulukkoon.

- Valitse arvot ja Yhden muuttujan analyysi -työväline.
- Valitse avautuvan ikkunan asetuksista vaihtoehto ”Data ja frekvenssit”.
- Valitse frekvenssit ja paina ”Lisää valinta”.
- Paina ”Analysoi”.

	A	B
1	0	85
2	1	71
3	2	65
4	3	35
5	4	9
6	5	7

Yhden muuttujan analyysi



Tilastot	
n	272
Keskiarvo	1.386
σ	1.2754
s	1.2778
Σx	377
Σx²	965
Min	0
Q1	0
Mediaani	1
Q3	2
Max	5

Luetaan kysytyt tunnusluvut ohjelman tulosteesta:

a)

Tapa 1. Aineiston koko on $n = 272$. Yhteensä 272 henkilöautoa oli mukana tutkimuksessa.

Tapa 2. Lasketaan frekvenssien summa: $85 + 71 + 65 + 35 + 9 + 7 = 272$.

b)

Mediaani on Md = 1 (matkustaja).
Keskiarvo on $\bar{x} = 1{,}38\ldots \approx 1{,}4$ (matkustajaa).
Keskihajonta on $s = 1{,}27\ldots \approx 1{,}3$ (matkustajaa).

Vastaus:

a) 272 henkilöautoa

b) Md = 1 matkustaja, $\bar{x} \approx 1{,}4$ matkustajaa ja $s \approx 1{,}3$ matkustajaa

6.5

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

Kirjoitetaan arvot taulukkoon ja määritetään tunnusluvut tilasto-toiminnoilla.

	A	B	C	D
1	54			**Sarake 1**
2	63		Keskiarvo	59,83333333
3	75		Keskivirhe	2,953768352
4	58		Moodi	56
5	41		Mediaani	57
6	51		First Quartile	54,75
7	55		Kolmas neljännes	64,5
8	56		Varianssi	104,6969697
9	69		Keskihajonta	10,23215372
10	77		Kurtoosi	0,007066082
11	56		Vinous	0,161088066
12	63		Alue	36
13			Minimi	41
14			Maksimi	77
15			Summa	718
16			Lukumäärä	12

Luetaan kysytyt tunnusluvut ohjelman tulosteesta:

a)

Pienin arvo on 41 vuotta. Suurin arvo on 77 vuotta.

Iän vaihteluväli on [41 vuotta, 77 vuotta].

Keskiarvo on $\bar{x} = 59{,}83\ldots \approx 59{,}8$ (vuotta).

Keskihajonta on $s = 10{,}23\ldots \approx 10{,}2$ (vuotta).

b)

Mediaani on 57 vuotta.

Tämä tarkoittaa, että puolet (50 %) Suomen presidenteistä oli virkaan astuessaan korkeitaan 57-vuotiaita ja puolet (50 %) vähintään 57-vuotiaita.

Vastaus:

a) [41 vuotta, 77 vuotta], $\bar{x} \approx 59{,}8$ vuotta ja $s \approx 10{,}2$ vuotta

b) Md = 57 vuotta

6.6

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.
Määritetään autojen rekisteröintimäärän tunnusluvut tilasto-toiminnoilla.

Vuosi	Lukumäärä		Sarake 1
2000	134603	Keskiarvo	120827,0952
2001	109429	Keskivirhe	3677,807803
2002	117034	Moodi	#ARVO!
2003	147405	Mediaani	118587
2004	142642	First Quartile	109429
2005	148161	Kolmas neljännes	134603
2006	145700	Varianssi	284051675
2007	125608	Keskihajonta	16853,83265
2008	139669	Kurtoosi	-0,81859877
2009	90574	Vinous	0,216842212
2010	111968	Alue	57587
2011	126123	Minimi	90574
2012	111251	Maksimi	148161
2013	103450	Summa	2537369
2014	106236	Lukumäärä	21
2015	108812		
2016	119000		
2017	118587		
2018	120499		
2019	114203		
2020	96415		

Luetaan kysytyt tunnusluvut ohjelman tulosteesta:

a)

Pienin arvo on 90 574. Suurin arvo on 148 161.

Rekisteröintimäärän vaihteluväli on [90 574 autoa, 148 161 autoa].

b)

Suurin arvo 148 161 on vuodelta 2005, joten vuonna 2005 rekisteröitiin eniten autoja.

c)

Keskiarvo on $\bar{x} = 120\,827{,}0\ldots \approx 120\,827$ (autoa/vuosi).

Keskihajonta on $s = 16\,853{,}8\ldots \approx 16\,854$ (autoa).

d)

Tapa 1. Ohjelmiston tulostaman taulukon perusteella arvojen summa on 2 537 369. Aikavälillä 2000–2020 rekisteröitiin siis yhteensä 2 537 369 autoa.

Tapa 2. Lasketaan vuosittaisten rekisteröintimäärien summa. Käytetään ohjelman summa-toimintoa. Summaksi saadaan 2 537 369, joten aikavälillä 2000–2020 rekisteröitiin yhteensä 2 537 369 autoa.

Vastaus:

a) [90 574 autoa, 148 161 autoa]

b) Vuonna 2005.

c) $\bar{x} \approx 120\,827$ autoa ja $s \approx 16\,854$

d) 2 537 369 autoa

6.7

Aineisto on luokiteltu, joten tunnusluvut määritetään luokkakeskusten avulla.

Ratkaistaan tehtävä GeoGebran taulukko-sovelluksessa.

	A	B	C	D
1	Tod.alaraja	Tod.yläraja	Luokkakeskus	f
2	1	5	3	84
3	6	10	8	234
4	11	15	13	43
5	16	20	18	17
6	21	25	23	5
7	26	30	28	2

Ohjeet:

1. Kirjoita todelliset luokkarajat taulukkoon.
2. Laske luokkakeskukset.
3. Kirjoita frekvenssit taulukkoon.
4. Määritä tunnusluvut luokkakeskusten ja frekvenssien perusteella Yhden muuttujan analyysi -työvälineellä.

Katso tarvittaessa ohjeet Esimerkin 3 videolta.

Yhden muuttujan analyysi

Tilastot

n	385
Keskiarvo	8.2078
σ	4.1975
s	4.203
Σx	3160
Σx²	32720
Min	3
Q1	8
Mediaani	8
Q3	8
Max	28

Luetaan kysytyt tunnusluvut ohjelman tulosteesta:

a)

Keskiarvo on $\bar{x} = 8{,}20\ldots \approx 8{,}2$ (kirjainta).

Keskihajonta on $s = 4{,}20\ldots \approx 4{,}2$ (kirjainta).

b)

Lasketaan, kuinka kaukana keskiarvosta $\bar{x}$ suurin arvo 29 on.

$29 - 8{,}2077\ldots = 20{,}792\ldots$ (kirjainta)

Lasketaan, kuinka monta keskihajontaa ero on.

Huomaa!
Älä käytä liian pyöristettyjä välituloksia.

$$\frac{20{,}792\ldots}{4{,}202\ldots} = 4{,}94\ldots \approx 4{,}9$$

Suurin arvo on 4,9 keskihajonnan päässä keskiarvosta.

Vastaus:

a) $\bar{x} \approx 8{,}2$ kirjainta ja $s \approx 4{,}2$ kirjainta

b) 4,9 keskihajonnan päässä

6.8

Ratkaistaan tehtävä taulukkolaskentaohjelmassa.
Katso tarvittaessa ohjeet Esimerkin 4 videolta.

a)

Määritetään tunnusluvut tilasto-toiminnoilla.

	A	B	C	D
1	Valkohain pituuksia (m)			
2				
3	5,70			**Sarake 1**
4	5,00		Keskiarvo	4,751363636
5	4,02		Keskivirhe	0,117154688
6	5,82		Moodi	4,02
7	3,75		Mediaani	4,805
8	5,09		First Quartile	4,14
9	4,82		Kolmas neljännes	5,18
10	4,94		Varianssi	0,603909725
11	5,67		Keskihajonta	0,777116288
12	5,43		Kurtoosi	0,456372048
13	4,36		Vinous	0,220798251
14	6,95		Alue	4,08
15	5,00		Minimi	2,87
16	4,94		Maksimi	6,95
17	5,06		Summa	209,06
18	5,12		Lukumäärä	44
19	4,79			

Keskiarvo on $\bar{x} = 4{,}751\ldots \approx 4{,}75$ (m). Keskihajonta on $s = 0{,}777\ldots \approx 0{,}78$ (m).

b)

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa pienempi.

$\bar{x} - 2s = 4{,}751\ldots - 2 \cdot 0{,}777\ldots = 3{,}197\ldots \approx 3{,}20$ (m)

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa suurempi.

$\bar{x} + 2s = 4{,}751\ldots + 2 \cdot 0{,}777\ldots = 6{,}305\ldots \approx 6{,}31$ (m)

Tee laskut tarkoilla arvoilla taulukossa soluviittausten avulla.

Katso tarvittaessa ohjeet Esimerkin 4 videolta.

c)

Verrataan aineiston pienintä ja suurinta arvoa b-kohdassa laskettuihin arvoihin:

Aineiston pienin arvo on 2,87 < 3,20. Pienin arvo on yli kaksi keskihajontaa keskiarvoa pienempi, joten se poikkeaa merkitsevästi keskiarvosta.

Aineiston suurin arvo on 6,95 > 6,31. Suurin arvo on yli kaksi keskihajontaa keskiarvoa suurempi, joten se poikkeaa merkitsevästi keskiarvosta.

Vastaus:

a) $\bar{x} \approx 4{,}75$ m ja $s \approx 0{,}78$ m

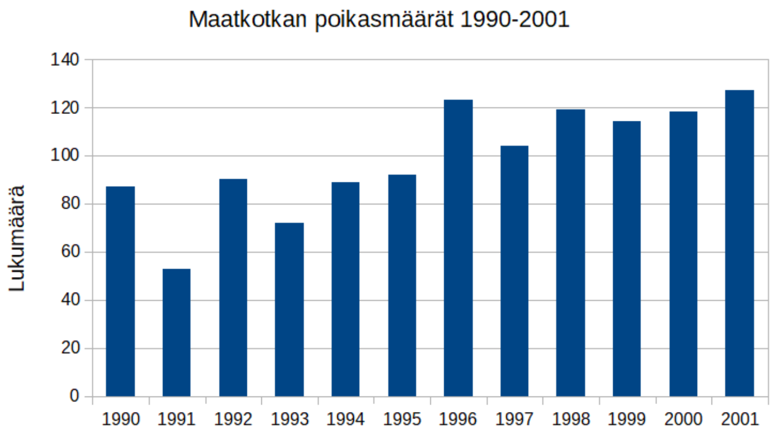
b) 3,20 m ja 6,31 m

c) Molemmat poikkeavat merkitsevästi keskiarvosta.

6.9

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

Piirretään pylväsdiagrammi.



Ohje:

1. Valitse poikasmäärät.
2. Valitse Lisää – Kaavio.
3. Valitse ”Pylväs”.
4. Valitse vaihtoehto ”Ensimmäinen sarake sisältää otsikoita”.
5. Lisää kaaviolle otsikko ja y-akselille otsikko.

Määritetään tunnusluvut tilasto-toiminnoilla.

	A	B	C	D	E
1	Maakotkan poikasmäärät				
2					
3	**Vuosi**	**Poikasmäärä**			**Sarake 1**
4	1990	87		Keskiarvo	99
5	1991	53		Keskivirhe	6,507571348
6	1992	90		Moodi	#ARVO!
7	1993	72		Mediaani	98
8	1994	89		First Quartile	88,5
9	1995	92		Kolmas neljännes	118,25
10	1996	123		Varianssi	508,1818182
11	1997	104		Keskihajonta	22,54288842
12	1998	119		Kurtoosi	-0,1669286
13	1999	114		Vinous	-0,6350114
14	2000	118		Alue	74
15	2001	127		Minimi	53
16				Maksimi	127
17				Summa	1188
18				Lukumäärä	12

Poikasmäärän keskiarvo on $\bar{x} = 99$ (poikasta) ja keskihajonta on $s = 22{,}54\ldots \approx 23$ (poikasta).

Tutkitaan vuosien 1991 ja 2001 poikasmääriä.

Tapa 1.

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa pienempi.

$\bar{x} - 2s = 99 - 2 \cdot 22{,}54 \ldots = 53{,}91 \approx 53{,}9$ (poikasta)

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa suurempi.

$\bar{x} + 2s = 99 + 2 \cdot 22{,}54 \ldots = 144{,}08 \ldots \approx 144$ (poikasta)

Tee laskut tarkoilla arvoilla taulukossa soluviittausten avulla.

Katso tarvittaessa ohjeet Esimerkin 4 videolta.

Verrataan vuosien 1991 ja 2001 arvoja laskettuihin arvoihin:

Vuoden 1991 poikasmäärä on 53 < 53,9. Tämä poikasmäärä on yli kaksi keskihajontaa keskiarvoa pienempi, joten se poikkeaa merkitsevästi keskiarvosta.

Vuoden 2001 poikasmäärä on 127. Tämä poikasmäärä on välillä $53{,}9 < 127 < 144$ eli alle kahden keskihajonnan päässä keskiarvosta. Vuoden 2001 poikasmäärä ei poikkea merkitsevästi keskiarvosta.

Tapa 2.

Lasketaan, kuinka monen keskihajonnan päässä keskiarvosta vuosien 1991 ja 2001 poikasmäärät ovat.

1991: 53 poikasta

$$\frac{99-53}{22{,}54\ldots} = 2{,}040\ldots > 2$$

Laske, kuinka paljon arvo poikkeaa keskiarvosta. Jaa erotus keskihajonnalla.

Vuoden 1991 poikasmäärä on yli kahden keskihajonnan päässä keskiarvosta, joten se poikkeaa merkitsevästi keskiarvosta.

2001: 127 poikasta

$$\frac{127-99}{22{,}54\ldots} = 1{,}242\ldots < 2$$

Vuoden 2001 poikasmäärä on alle kahden keskihajonnan päässä keskiarvosta, joten se ei poikkea merkitsevästi keskiarvosta.

Vastaus:

$\bar{x} = 99$ poikasta ja $s \approx 23$ poikasta

Vuoden 1991 poikasmäärä poikkeaa merkitsevästi keskiarvosta, vuoden 2001 poikasmäärä ei.

6.10

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

a)

Kukin havaintoarvo ilmaisee pesien lukumäärää. Arvojen summa ilmaisee tällöin pesien kokonaismäärän.

Lasketaan arvojen summa. Käytetään ohjelman summa-toimintoa.

Summaksi saadaan 25 761, joten vuonna 2020 oli yhteensä 25 761 merimetson pesää.

b)

Määritetään pesien lukumäärän tunnusluvut tilasto-toiminnoilla.

	Sarake 1
Keskiarvo	477,05556
Keskivirhe	84,178923
Moodi	80
Mediaani	308
First Quartile	87,5
Kolmas neljännes	482
Varianssi	382648,92
Keskihajonta	618,58623
Kurtoosi	11,459325
Vinous	2,8829252
Alue	3598
Minimi	2
Maksimi	3600
Summa	25761
Lukumäärä	54

Huomautus a-kohtaan: Tunnuslukujen yhteenveto-taulukosta nähdään myös summa eli pesien kokonaismäärä.

Luetaan kysytyt tunnusluvut yhteenvetotaulukosta:

Pesien lukumäärän keskiarvo on $\bar{x} = 477{,}05\ldots \approx 477$ (pesää).
Pesien lukumäärän keskihajonta on $s = 618{,}58\ldots \approx 619$ (pesää).

c)

Lasketaan, kuinka monen keskihajonnan päässä keskiarvosta Rauman merimetsoyhdyskunnan pesämäärä 3600 on.

$$\frac{3600 - 477{,}05\ldots}{618{,}58\ldots} = 5{,}048\ldots \approx 5{,}0 > 2$$

Laske, kuinka paljon arvo 3600 poikkeaa keskiarvosta. Jaa erotus keskihajonnalla.

Tee lasku tarkoilla arvoilla taulukossa. Katso tarvittaessa ohjeet Esimerkin 4 videolta.

Rauman yhdyskunnan pesämäärä on noin viisi keskihajontaa keskiarvoa suurempi. Ero on yli kaksi keskihajontaa, joten arvo poikkeaa merkitsevästi keskiarvosta. Tämän perusteella voidaan sanoa, että Rauman yhdyskunta oli poikkeuksellisen suuri muihin yhdyskuntiin verrattuna.

Vastaus:

a) 25 761 pesää

b) $\bar{x} \approx 477$ pesää ja $s \approx 619$ pesää

c) Noin viiden keskihajonnan päässä. Rauman yhdyskunta oli poikkeuksellisen suuri.

6.11

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa pienempi.

$\bar{x} - 2s = 7{,}1 - 2 \cdot 2{,}3 = 2{,}5$

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa suurempi.

$\bar{x} + 2s = 7{,}1 + 2 \cdot 2{,}3 = 11{,}7$

Arvot, jotka ovat

- pienempiä kuin 2,5 tai
- suurempia kuin 11,7

poikkeavat merkitsevästi keskiarvosta. Tällaisia arvoja ovat 1 ja 2 sekä 12.

Arvoja 1 ja 2 esiintyy pylväskaavion perusteella molempia yksi kappale.
Arvoa 12 esiintyy kaksi kappaletta.

Keskiarvoa merkitsevästi pienempiä arvoja on siis kaksi (1 ja 2), ja keskiarvoa merkitsevästi suurempia arvoja on kaksi (12 ja 12).

Vastaus:

Arvot 1 ja 2 sekä 12 poikkeavat merkitsevästi keskiarvosta.

6.12

Moodi on se muuttujan arvo, jota esiintyy eniten, eli jonka kohdalla on korkein pylväs. Mo = 2.

Selvitetään muiden keskilukujen (Md ja keskiarvo) suuruudet.

Tapa 1.

Mediaani on suuruusjärjestyksessä olevien arvojen keskimmäinen arvo. Jakauman puoliväli ei pylväiden korkeuksien perusteella ole vielä arvon 2 kohdalla. Mediaani on siis suurempi kuin moodi.

Jakaumassa on suurien arvojen kohdalla matalat pylväät. Tämä tarkoittaa, että joukossa on muutamia suuria arvoja. Suuret arvot kasvattavat keskiarvoa, joten keskiarvo on mediaania suurempi.

Keskilukujen suuruusjärjestys on $\bar{x}$ > Md > Mo.

Tapa 2.

Määritetään mediaanin ja keskiarvon tarkat arvot GeoGebran taulukko-sovelluksella.

Luetaan pylväskaaviosta arvot ja niiden frekvenssit. Kirjataan tiedot taulukkoon ja tehdään yhden muuttujan analyysi.

	A	B
1	Arvo	f
2	1	7
3	2	21
4	3	19
5	4	17
6	5	13
7	6	8
8	7	5
9	8	3

Yhden muuttujan analyysi

Tilastot

n	93
Keskiarvo	3.7204
σ	1.7862
s	1.7959
Σx	346
Σx²	1584
Min	1
Q1	2
Mediaani	3
Q3	5
Max	8

Yhteenvetotaulukon perusteella muuttujan arvon mediaani on Md = 3 ja keskiarvo on $\bar{x} \approx 3{,}7$.

Keskilukujen suuruusjärjestys on 3,7 > 3 > 2 eli $\bar{x}$ > Md > Mo.

Vastaus:

$\bar{x}$ > Md > Mo

6.13

Aineisto on luokiteltu, joten tunnusluvut määritetään luokkakeskusten avulla.

Ratkaistaan tehtävä GeoGebran taulukko-sovelluksessa.

	A	B	C	D
1	Tod.alaraja	Tod.yläraja	Luokkakeskus	f
2	9.5	14.5	12	10
3	14.5	19.5	17	6
4	19.5	24.5	22	18
5	24.5	29.5	27	22
6	29.5	34.5	32	6
7	34.5	39.5	37	4

Ohje:

1. Kirjoita todelliset luokkarajat taulukkoon.
2. Laske luokkakeskukset.
3. Kirjoita frekvenssit taulukkoon.
4. Määritä tunnusluvut luokkakeskusten ja frekvenssien perusteella Yhden muuttujan analyysi -työvälineellä.

Katso tarvittaessa ohjeet Esimerkin 3 videolta.

Yhden muuttujan analyysi

Tilastot

n	66
Keskiarvo	23.51515
σ	6.79626
s	6.84834
Σx	1552
Σx²	39544
Min	12
Q1	22
Mediaani	22
Q3	27
Max	37

Luetaan kysytyt tunnusluvut ohjelman tulosteesta:

a)

Keskiarvo on $\bar{x} = 23{,}51\ldots \approx 23{,}5$ (cm).

Keskihajonta on $s = 6{,}84\ldots \approx 6{,}8$ (cm).

b)

Lasketaan, kuinka kaukana keskiarvosta $\bar{x}$ pienin arvo 11 on.

$23{,}515\ldots - 11 = 12{,}515\ldots$ (cm)

Lasketaan, kuinka monta keskihajontaa ero on.

Huomaa!
Älä pyöristä välituloksia liikaa.

$$\frac{12{,}515\ldots}{6{,}848\ldots} = 1{,}82\ldots \approx 1{,}8$$

Pienin arvo on 1,8 keskihajonnan päässä keskiarvosta.

c)

Tapa 1. Aineiston koko on n = 66. Ryhmä mittasi yhteensä 66 heinän pituuden.

Tapa 2. Lasketaan frekvenssien summa: 10 + 6 + 18 + 22 + 6 + 4 = 66.

Vastaus:

a) $\bar{x} \approx$ 23,5 cm ja $s \approx$ 6,8 cm

b) 1,8 keskihajontaa

c) 66 heinää

6.14

Ratkaistaan tehtävä taulukkolaskentaohjelmassa.

a)

Määritetään tunnusluvut tilasto-toiminnoilla.

	A	B	C	D	E
1	**Yhdysvaltojen presidenttien virkaanastumisiät**				
2					
3	**Presidentti**	**Ikä**			**Sarake 1**
4	George Washington	57		Keskiarvo	55,47826087
5	John Adams	61		Keskivirhe	1,081124451
6	Thomas Jefferson	57		Moodi	51
7	James Madison	57		Mediaani	55
8	James Monroe	58		First Quartile	51
9	J. Q. Adams	57		Kolmas neljä	59,5
10	Andrew Jackson	61		Varianssi	53,76618357
11	Martin Van Buren	54		Keskihajonta	7,332542777
12	W. H. Harrison	68		Kurtoosi	0,973317887
13	John Tyler	51		Vinous	0,756031684
14	James K. Polk	49		Alue	36
15	Zachary Taylor	64		Minimi	42
16	Millard Fillmore	50		Maksimi	78
17	Franklin Pierce	48		Summa	2552
18	James Buchanan	65		Lukumäärä	46
19	Abraham Lincoln	52			

Keskiarvo on $\bar{x} = 55{,}47\ldots \approx 55{,}5$ (vuotta).

Keskihajonta on $s = 7{,}33\ldots \approx 7{,}3$ (vuotta).

b)

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa pienempi.

$\bar{x} - 2s = 55{,}47\ldots - 2 \cdot 7{,}33\ldots = 40{,}81\ldots \approx 40{,}8$ (vuotta)

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa suurempi.

$\bar{x} + 2s = 55{,}47\ldots + 2 \cdot 7{,}33\ldots = 70{,}14\ldots \approx 70{,}1$ (vuotta)

Tee laskut tarkoilla arvoilla taulukossa soluviittausten avulla.

Katso tarvittaessa ohjeet Esimerkin 4 videolta.

c)

Verrataan Bidenin ikää (78 vuotta) b-kohdassa laskettuun arvoon:

Bidenin ikä on 78 > 70,1. Ikä on yli kaksi keskihajontaa keskiarvoa suurempi, joten ikä poikkeaa merkitsevästi keskiarvosta.

Tällä perusteella Bidenia voidaan pitää poikkeuksellisen vanhana muihin presidentteihin verrattuna.

Vastaus:

a) $\bar{x} \approx 55{,}5$ vuotta ja $s \approx 7{,}3$ vuotta

b) 40,8 vuotta ja 70,1 vuotta

c) Bidenia voidaan pitää poikkeuksellisen vanhana presidenttinä.

6.15

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.
Määritetään tunnusluvut tilasto-toiminnoilla.

	A	B	C	D
1	Simon Newcomb (1882)			
2	Mittaukset valon nopeuden määrittämiseksi			
3				
4	28			**Sarake 1**
5	22		Keskiarvo	26,21212121
6	36		Keskivirhe	1,322658048
7	26		Moodi	28
8	28		Mediaani	27
9	28		First Quartile	24
10	26		Kolmas neljä	30,75
11	24		Varianssi	115,4620047
12	32		Keskihajonta	10,74532478
13	30		Kurtoosi	28,61493577
14	27		Vinous	-4,598484611
15	24		Alue	84
16	33		Minimi	-44
17	21		Maksimi	40
18	36		Summa	1730
19	32		Lukumäärä	66
20	31			

a)

Mittaustulosten keskiarvo on $\bar{x} = 26{,}21\ldots \approx 26{,}2$ ja keskihajonta on $s = 10{,}74\ldots \approx 10{,}7$.

b)

Käytetään ohjelman lajittelu-toimintoa ja järjestetään mittaustulokset nousevaan järjestykseen eli pienimmästä suurimpaan.

4	-44
5	-2
6	16
7	16
8	19
9	20
10	21
11	21
12	22

Lajitellusta aineistosta nähdään, että joukossa on kaksi poikkeuksellisen pieneltä vaikuttavaa arvoa: -44 ja -2, kun muut arvot ovat välillä 16–40.

Lasketaan, kuinka monen keskihajonnan päässä keskiarvosta pieneltä vaikuttavat arvot ovat.

Arvo -44:

$$\frac{26{,}21\ldots - (-44)}{10{,}74\ldots} = 6{,}53\ldots\ldots \approx 6{,}5 > 2$$

Arvo -2:

$$\frac{26{,}21\ldots - (-2)}{10{,}74\ldots} = 2{,}62\ldots\ldots \approx 2{,}6 > 2$$

Laske, kuinka paljon arvo poikkeaa keskiarvosta. Jaa erotus keskihajonnalla.

Tee lasku taulukossa soluviittauksilla:

	A	B	C	D	E	F
1	Simon Newcomb (1882)					
2	Mittaukset valon nopeuden määrittämiseksi					
3						
4	-44			**Sarake 1**		
5	-2		Keskiarvo	26,21212121		=(D5-A4)/D12
6	16		Keskivirhe	1,322658048		
7	16		Moodi	28		
8	19		Mediaani	27		
9	20		First Quartile	24		
10	21		Kolmas neljä	30,75		
11	21		Varianssi	115,4620047		
12	22		Keskihajonta	10,74532478		
13	22		Kurtoosi	28,61493577		

Arvot −44 ja −2 ovat yli kahden keskihajonnan päässä keskiarvosta, joten ne poikkeavat merkitsevästi keskiarvosta.

c)

Määritetään tunnusluvut uudestaan, ilman arvoja −44 ja −2.

	Sarake 1
Keskiarvo	27,75
Keskivirhe	0,635428864
Moodi	28
Mediaani	27,5
First Quartile	24,75
Kolmas neljän	31
Varianssi	25,84126984
Keskihajonta	5,083430912
Kurtoosi	0,150160314
Vinous	0,154168423
Alue	24
Minimi	16
Maksimi	40
Summa	1776
Lukumäärä	64

Mittaustulosten keskiarvo on nyt $\bar{x} = 27{,}75 \approx 27{,}8$ ja keskihajonta on $s = 5{,}08... \approx 5{,}1$.

d)

c-kohdan aineistossa on yhteensä 64 arvoa.

Lasketaan arvo, joka on yksi keskihajontaa keskiarvoa pienempi.

$\bar{x} - s = 27{,}75 - 5{,}08\ldots = 22{,}66\ldots \approx 22{,}7 < 23$

Tee laskut taulukossa soluviittausten avulla.

Lasketaan arvo, joka on yksi keskihajontaa keskiarvoa suurempi.

$\bar{x} + s = 27{,}75 + 5{,}08\ldots = 32{,}83\ldots \approx 32{,}8 > 32$

Alle yhden keskihajonnan päässä keskiarvosta olevat arvot ovat välillä 23–32.

Lasketaan suuruusjärjestyksessä olevasta aineistosta, kuinka monta arvoa on tällä välillä: arvoja on yhteensä 46. Prosentteina tämä on

$$\frac{46}{64} = 0{,}71875 \approx 72\ \%$$

Arvoista 72 % on alle yhden keskihajonnan päässä keskiarvosta.

Vastaus:

a) $\bar{x} \approx 26{,}2$ ja $s \approx 10{,}7$

b) Arvot −44 ja −2.

c) $\bar{x} \approx 27{,}8$ ja $s \approx 5{,}1$

d) 72 %

6.16

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.
Määritetään tunnusluvut tilasto-toiminnoilla.

Alue A:

	Sarake 1
Keskiarvo	51,57
Keskivirhe	5,7486570514
Moodi	25,6
Mediaani	58,5
First Quartile	27,775
Kolmas neljännes	67,975
Varianssi	660,94115789
Keskihajonta	25,708775893
Kurtoosi	-0,9919135358
Vinous	0,0548482914
Alue	88,8
Minimi	12,5
Maksimi	101,3
Summa	1031,4
Lukumäärä	20

Alue B:

	Sarake 1
Keskiarvo	56,883333333
Keskivirhe	4,0407555375
Moodi	#ARVO!
Mediaani	57,2
First Quartile	48,05
Kolmas neljännes	61,875
Varianssi	391,86492754
Keskihajonta	19,795578485
Kurtoosi	1,8501191853
Vinous	0,5123458375
Alue	96,2
Minimi	15,3
Maksimi	111,5
Summa	1365,2
Lukumäärä	24

Poimitaan kysytyt tunnusluvut ohjelman yhteenvetotaulukoista:

Alueella A metsätilojen

- vaihteluväli on [12,5 ha; 101,3 ha]
- Md = 58,5 (hehtaaria)
- $\bar{x} = 51,57 \approx 51,6$ (hehtaaria)
- $s = 25,70 \ldots \approx 25,7$ (hehtaaria).

Alueella B metsätilojen

- vaihteluväli on [15,3 ha; 111,5 ha]
- Md = 57,2 (hehtaaria)
- $\bar{x} = 56,88 \ldots \approx 56,9$ (hehtaaria)
- $s = 19,79 \ldots \approx 19,8$ (hehtaaria).

b)

Alueen B metsätilojen koon mediaani on suurempi (57,2 > 51,6) ja keskiarvo on suurempi (56,9 > 51,6). Alueen B tilat ovat siis keskimäärin suurempia.

Alueen A metsätilojen koon keskihajonta on suurempi (25,7 > 19,8), joten alueen A tiloissa esiintyy suurempaa hajontaa.

c)

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa pienempi.

Alue A: $\bar{x} - 2s = 51,57 - 2 \cdot 25,70 \ldots = 0,152 \ldots \approx 0,2$ (ha)
Alue B: $\bar{x} - 2s = 56,88 \ldots - 2 \cdot 19,79 \ldots = 17,29 \ldots \ldots \approx 17,3$ (ha)

Lasketaan arvo, joka on kaksi keskihajontaa keskiarvoa suurempi.

Alue A: $\bar{x} + 2s = 51,57 + 2 \cdot 25,70 \ldots = 102,98 \ldots \approx 103,0$ (ha)
Alue b: $\bar{x} + 2s = 56,88 \ldots + 2 \cdot 19,79 \ldots = 96,47 \ldots \approx 96,5$ (ha)

Verrataan alueiden pienimpiä ja suurimpia tiloja laskettuihin arvoihin:

Alueen A pienin tila on 12,5 ha > 0,2 ha ja suurin 101, 3 ha < 103,0 ha. Kumpikaan näistä ei poikkea merkitsevästi keskiarvosta, joten alueella A ei ole merkitsevästi keskiarvosta poikkeavia arvoja.

Alueen B pienin tila on 15,3 ha < 17,3 ha ja suurin 111,5 ha > 96,5 ha. Sekä pienin että suurin tila ovat yli kahden keskihajonnan päässä keskiarvosta eli poikkeavat merkitsevästi keskiarvosta.

Tutkitaan vielä, onko alueella B muita tiloja, jotka poikkeavat merkitsevästi keskiarvosta. Lajitellaan tilat pienimmästä suurimpaan.

- Toiseksi pienin tila on 27,4 ha > 17,3 ha.
- Toiseksi suurin tila on 83,6 ha < 96,5 ha.

Kumpikaan näistä ei poikkea merkitsevästi keskiarvosta, joten alueella B ei ole muita merkitsevästi keskiarvosta poikkeavia arvoja.

Vastaus:

a)

Alueella A:

- vaihteluväli on [12,5 ha; 101,3 ha]
- Md = 58,5 ha
- $\bar{x} \approx 51{,}6$ ha
- $s \approx 25{,}7$ ha

Alueella B:

- vaihteluväli on [15,3 ha; 111,5 ha]
- Md = 57,2 ha
- $\bar{x} \approx 56{,}9$ ha
- $s \approx 19{,}8$ ha

b) Alueen B tilat ovat keskimäärin suurempia. Alueella A tilojen koossa esiintyy suurempaa hajontaa.

c) Alueella A ei ole merkitsevästi keskiarvosta poikkeavia arvoja. Alueen B tiloista sekä pienin että suurin poikkeavat merkitsevästi keskiarvosta.

6.17

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.
Määritetään tunnusluvut tilasto-toiminnoilla.

	Sarake 1
Keskiarvo	157,2925926
Keskivirhe	32,89643029
Moodi	65
Mediaani	84
First Quartile	44
Kolmas neljännes	270
Varianssi	29218,7284
Keskihajonta	170,934866
Kurtoosi	1,119706216
Vinous	1,400199062
Alue	632,7
Minimi	0,3
Maksimi	633
Summa	4246,9
Lukumäärä	27

a)

Pinta-alan keskiarvo on $\bar{x} = 157{,}29 \ldots \approx 157$ eli $\bar{x} \approx 157\,000$ km^2.
Pinta-alan mediaani Md = 84 eli Md = 84 000 km^2.

Keskiarvo on huomattavasti mediaania suurempi.

b)

Pinta-alan keskihajonta on $s = 170{,}93\ldots \approx 171$ (tuhatta km^2) eli $s \approx 171\,000$ km^2.

Lajitellaan aineisto pinta-alan mukaan laskevaan järjestykseen eli suurimmasta pienimpään.

	A	B	C
1	**EU-jäsenmaiden pinta-alat**		
2	*Lähde: EU*		
3	**Maa**	**Pinta-ala (/1000 km²)**	
4	Ranska	633	
5	Espanja	506	
6	Ruotsi	439	
7	Saksa	357	

Kaksi suurinta arvoa ovat 633 (Ranska) ja 506 (Espanja).

Lasketaan, kuinka monen keskihajonnan päässä keskiarvosta nämä arvot ovat:

Ranska: $\frac{633-157{,}29\ldots}{170{,}93\ldots} = 2{,}782\ldots \approx 2{,}78 > 2$

Suurin arvo poikkeaa noin 2,8 keskihajontaa keskiarvosta.

Espanja: $\frac{506-157{,}29\ldots}{170{,}93\ldots} = 2{,}040\ldots \approx 2{,}04 > 2$

Laske, kuinka paljon arvo poikkeaa keskiarvosta. Jaa erotus keskihajonnalla.

Tee laskut tarkoilla arvoilla taulukossa. Katso tarvittaessa ohjeet Esimerkin 4 videolta.

Toiseksi suurin arvo poikkeaa noin 2,0 keskihajontaa keskiarvosta.

Molemmat arvot ovat yli kahden keskihajonnan päässä keskiarvosta, eli arvot poikkeavat merkitsevästi keskiarvosta.

c)

Aineistossa on kaksi poikkeuksellisen suurta arvoa (Ranska 633, Espanja 506). Nämä poikkeuksellisen suuret arvot suurentavat keskiarvoa, jolloin keskiarvo on mediaania suurempi.

Vastaus:

a) $\bar{x} \approx 157\ 000\ \text{km}^2$ ja Md = $84\ 000\ \text{km}^2$

b) $s \approx 171\ 000\ \text{km}^2$

c) Suurin arvo 2,8 keskihajontaa, toiseksi suurin arvo 2,0 keskihajontaa.

d) Aineistossa on kaksi poikkeuksellisen suurta arvoa, jotka suurentavat keskiarvoa.

6.18

a)

Lasketaan lipputulojen summa. Käytetään ohjelman summa-toimintoa.

Summaksi saadaan 17 759,5.

Bond-elokuvat ovat tuottaneet yhteensä 17 759,5 miljoonaa dollaria.

Huomaa: Uusimman elokuvan (No Time to Die) tuotto ei ole tiedossa.

b)

Määritetään lipputulojen keskiarvo tilasto-toiminnoilla.

Keskiarvoksi saadaan $\bar{x} = 739{,}97 \ldots \approx 740{,}0$.

Bond-elokuvat ovat tuottaneet keskimäärin 740,0 miljoonaa dollaria elokuvaa kohden.

c)

Määritetään elokuvien keston keskiarvo tilasto-toiminnoilla.

Keskiarvoksi saadaan $\bar{x} = 128{,}24 \approx 128$.

Bond-elokuvien keskimääräinen kesto on 128 minuuttia.

128 min = 60 min + 60 min + 8 min = 2 h 8 min

d)

Lasketaan elokuvien kestojen summa. Uusimman elokuvan (No Time to Die) tuotto ei ole tiedossa, joten ei huomioida tämän elokuvan kestoa summassa.

Käytetään ohjelman summa-toimintoa.

Summaksi saadaan 3043 (min).

- Elokuvasarja on tuottanut yhteensä 17 759,5 miljoonaa dollaria.
- Elokuvasarjan kokonaiskesto on 3043 minuuttia.

Lasketaan elokuvasarjan tuotto minuuttia kohden.

$$\frac{17\ 759{,}5}{3043} = 5{,}83 \ldots \approx 5{,}8$$

Elokuvasarja on tuottanut 5,8 miljoonaa dollaria jokaista minuuttia kohden.

Vastaus:

a) 17 759,5 miljoonaa USD

b) 740,0 miljoonaa USD

c) 2 h 8 min

d) 5,8 miljoonaa USD

6.19

Ratkaistaan tehtävä GeoGebran taulukko-sovelluksessa.

a)

Kirjoitetaan lasten lukumäärät ja perheiden lukumäärät taulukkoon.

	A	B
1	Lasten lukumäärä	Perheitä
2	1	241709
3	2	220116
4	3	75326
5	4	18409
6	5	5493
7	6	2289
8	7	1235
9	8	751
10	9	476
11	10	262
12	11	117
13	12	41
14	13	12
15	14	3
16	15	0
17	16	3

Taulukon perusteella eniten on niitä perheitä, joissa on 1 lapsi.
Lasten lukumäärän moodi on siis 1 lapsi.

Määritetään lasten lukumäärän mediaani ja keskiarvo tilastotoiminnoilla tekemällä yhden muuttujan analyysi.

Tilastot

n	566242
Keskiarvo	1.8479
σ	1.01
s	1.01
Σx	1046336
Σx²	2511160
Min	1
Q1	1
Mediaani	2
Q3	2
Max	16

Lasten lukumäärän keskiarvo on $\bar{x} = 1{,}847\ldots \approx 1{,}85$ (lasta) ja mediaani on Md = 2 (lasta).

b)

Tapa 1.

Jos perheessä on esimerkiksi 3 lasta, on jokaisella lapsella kaksi sisarusta.

- Tällaisessa perheessä on 3 lasta, joilla on kaksi sisarusta.
- Perheitä, joissa jokaisella lapsella on kaksi sisarusta, on yhteensä 75 326 kappaletta.
- Niitä lapsia, joilla on kaksi sisarusta, on yhteensä $3 \cdot 75\,326 = 225\,978$.

Samoin saadaan laskettua lasten lukumäärä, joilla on 0–15 sisarusta: kerrotaan lasten lukumäärä perheiden lukumäärällä.

Laaditaan taulukko, josta käy ilmi sisarusten lukumäärä sekä lapsien lukumäärä.

	A	B	C	D
1	Lasten lukumäärä	Perheitä	Sisarusten lukumäärä	Lapsia
2	1	241709	0	241709
3	2	220116	1	440232
4	3	75326	2	225978
5	4	18409	3	73636
6	5	5493	4	27465
7	6	2289	5	13734
8	7	1235	6	8645
9	8	751	7	6008
10	9	476	8	4284
11	10	262	9	2620
12	11	117	10	1287
13	12	41	11	492
14	13	12	12	156
15	14	3	13	42
16	15	0	14	0
17	16	3	15	48

Ohje:

1. Kirjoita soluun D2 kaava: =A2*B2

2. Kopioi kaava sarakkeessa alaspäin.

Määritetään sisarusten lukumäärän keskiarvo tilastotoiminnoilla tekemällä sisarusten lukumäärälle yhden muuttujan analyysi.

	A	B	C	D
1	Lasten lukumäärä	Perheitä	Sisarusten lukumäärä	Lapsia
2	1	241709	0	241709
3	2	220116	1	440232
4	3	75326	2	225978
5	4	18409	3	73636
6	5	5493	4	27465
7	6	2289	5	13734
8	7	1235	6	8645
9	8	751	7	6008
10	9	476	8	4284
11	10	262	9	2620
12	11	117	10	1287
13	12	41	11	492
14	13	12	12	156
15	14	3	13	42
16	15	0	14	0
17	16	3	15	48

Yhden muuttujan analyysi

Tilastot

n	1046336
Keskiarvo	1.39996
σ	1.40298
s	1.40299
Σx	1464824
Σx²	4110260
Min	0
Q1	1
Mediaani	1
Q3	2
Max	15

Sisarusten lukumäärän keskiarvo on $\bar{x} = 1{,}399\ldots \approx 1{,}40$ (sisarusta).

Tapa 2.

a-kohdan tulosteesta nähdään lasten kokonaismäärä lapsiperheissä.

Tilastot

n	566242
Keskiarvo	1.8479
σ	1.01
s	1.01
Σx	1046336
Σx²	2511160
Min	1
Q1	1
Mediaani	2
Q3	2
Max	16

Lapsia on yhteensä 1 046 336.

Lasketaan sisarusten kokonaismäärä. Jos perheessä on esimerkiksi 3 lasta, niin jokaisella heistä on kaksi sisarusta.

- Tällaisessa perheessä on 3 lasta, joilla on kaksi sisarusta.
- Perheitä, joissa jokaisella lapsella on kaksi sisarusta, on yhteensä 75 326 kappaletta.
- Niitä lapsia, joilla on kaksi sisarusta, on yhteensä $3 \cdot 75\,326$.

Sisarusten kokonaismäärä on tällöin

$2 \cdot 3 \cdot 75\,326 = 451\,956$.

Lasketaan samalla tavoin sisarusten määrä kussakin perhekoossa, ja lasketaan lopuksi sisarusten määrät yhteen.

	A	B	C	D
1	Lasten lukumäärä	Sisaruksia per lapsi	Perheitä	Sisaruksia
2	1	0	241709	0
3	2	1	220116	440232
4	3	2	75326	451956
5	4	3	18409	220908
6	5	4	5493	109860
7	6	5	2289	68670
8	7	6	1235	51870
9	8	7	751	42056
10	9	8	476	34272
11	10	9	262	23580
12	11	10	117	12870
13	12	11	41	5412
14	13	12	12	1872
15	14	13	3	546
16	15	14	0	0
17	16	15	3	720
18				1464824

Ohje:

1. Kirjoita soluun D2 kaava:
=A2*B2*C2

2. Kopioi kaava sarakkeessa alaspäin.

3. Laske yllä olevissa soluissa olevien lukujen summa.

Esimerkiksi:
=SUMMA(D2:D17)

Sisaruksia on yhteensä 1 464 824 kappaletta.

Lasketaan, kuinka monta sisarusta yhdellä lapsella on keskimäärin, kun lapsia on yhteensä 1 046 336 kappaletta.

$$\frac{1\,464\,824}{1\,046\,336} = 1{,}399\ldots \approx 1{,}40$$

Jokaisella lapsella on keskimäärin 1,40 sisarusta.

Vastaus:

a) Mo = 1 lapsi, Md = 2 lasta, $\bar{x} \approx 1{,}85$ (lasta)

b) $\bar{x} \approx 1{,}40$ (sisarusta)

2.20

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

a)

Kukin havaintoarvo ilmaisee pesien lukumäärää. Arvojen summa ilmaisee tällöin pesien kokonaismäärän.

Lasketaan arvojen summa. Käytetään ohjelman summa-toimintoa.

Summaksi saadaan 25 685, joten vuonna 2019 oli yhteensä 25 685 merimetson pesää.

b)

Lajitellaan aineisto pesien lukumäärän mukaan nousevaan järjestykseen eli pienimmästä suurimpaan. Lajittelun jälkeen nähdään, että pienin yhdyskunta on Helsingissä (7 pesää) ja suurin Raumalla (2080 pesää).

Määritetään aineistosta tunnusluvut tilasto-toiminnoilla.

	Sarake 1
Keskiarvo	546,4894
Keskivirhe	84,08384
Moodi	30
Mediaani	300
First Quartile	109
Kolmas neljännes	872,5
Varianssi	332294
Keskihajonta	576,4498
Kurtoosi	0,40959
Vinous	1,231508
Alue	2073
Minimi	7
Maksimi	2080
Summa	25685
Lukumäärä	47

Huomautus a-kohtaan:
Tunnuslukujen yhteenveto-taulukosta nähdään myös summa eli pesien kokonaismäärä.

Luetaan tunnusluvut yhteenvetotaulukosta:

Pesien lukumäärän keskiarvo on $\bar{x} = 546{,}48\ldots \approx 546$ (pesää).
Pesien lukumäärän keskihajonta on $s = 576{,}44\ldots \approx 576$ (pesää).

c)

Keskiarvo (546) on pienempi kuin keskihajonta (576).

- Keskiarvo on yhden luokan luokkarajana.
- Valitaan: keskiarvo 546 on ensimmäisen luokan yläraja. Valitaan alarajaksi 0.
- Ensimmäinen luokka on siis 0–546.
- Seuraavan luokan alaraja on tällöin 547.

Ensimmäisen luokan alarajaksi voitaisiin valita myös pienin arvo 7.

Lasketaan suuruusjärjestykseen lajitellun aineiston perusteella, kuinka monta arvoa kuuluu ensimmäiseen luokkaan, eli välille 0–546.

	A	B	C
1	**Yhdyskunta**	**Pesiä**	
2	Helsinki	7	
3	Kristiinankaupunki 2	30	
4	Parainen	30	
5	Espoo	36	
6	Ii	40	
7	Sauvo 2	40	
8	Vöyri 2	45	
9	Hamina 4	73	
10	Raasepori 5	75	
11	Hamina 3	85	
12	Naantali	88	
13	Hamina 2	91	
14	Sauvo 1	127	
15	Porvoo 4	132	
16	Raasepori 4	136	
17	Porvoo 3	158	
18	Siikajoki	163	
19	Maalahti	192	
20	Uusikaarlepyy 3	231	
21	Raasepori 3	239	
22	Hamina 1	261	
23	Porvoo 2	280	
24	Loviisa 3	294	
25	Vöyri 1	300	
26	Loviisa 2	345	
27	Raasepori 2	373	
28	Sipoo	387	
29	Raasepori 1	395	
30	Uusikaarlepyy 2	462	MIN → keskiarvo (546)
31	Hanko	536	Yhteensä 30 yhdyskuntaa.
32	Pyhtää	562	
33	Eurajoki 2	563	
34	Loviisa 1	563	

Ensimmäiseen luokaan kuuluu yhteensä 30 yhdyskuntaa.

Muut luokat ovat keskihajonnan (576) pituisia:

- Toisen luokan alaraja on 547. Yläraja on 547 + 576 = 1123.
- Kolmannen luokan alaraja on tällöin 1124 ja yläraja 1124 + 576 = 1700.
- Neljännen luokan alaraja on tällöin 1701 ja yläraja 1701 + 576 = 2277.
- Suurin arvo (2080) kuuluu neljänteen luokkaan 1701–2277.

Lasketaan pesämäärältään toiseen, kolmanteen ja neljänteen luokkaan kuuluvien yhdyskuntien lukumäärä suuruusjärjestykseen lajitellusta aineistosta.

30	Uusikaarlepyy 2	462	MIN → keskiarvo (546)
31	Hanko	536	Yhteensä 30 yhdyskuntaa.
32	Pyhtää	562	
33	Eurajoki 2	563	
34	Loviisa 1	563	
35	Kokkola	600	
36	Kristiinankaupunki 1	870	
37	Kirkkonummi 2	875	
38	Porvoo 1	990	
39	Uusikaarlepyy 1	1051	547 – 1123
40	Kustavi	1055	Yhteensä 9 yhdyskuntaa.
41	Virolahti	1148	
42	Pori	1246	
43	Eurajoki 1	1561	1124-1700
44	Uusikaupunki	1670	Yhteensä 4 yhdyskuntaa.
45	Mustasaari	1710	
46	Turku	1766	
47	Kirkkonummi 1	1793	1701-2277
48	Rauma	2080	Yhteensä 4 yhdyskuntaa.

Pesiä	Yhdyskuntia
0–546	**30**
547–1123	**9**
1124–1700	**4**
1701–2277	**4**

Luokiteltu aineisto on:

Pesiä	Yhdyskuntia
0–546	30
547–1123	9
1124–1700	4
1701–2277	4

Tässä on yksi esimerkki luokittelusta. Muitakin mahdollisia, tehtävänannon mukaisia luokitteluja voidaan muodostaa.

Tehtävänannon ehdot ovat:
keskiarvo (546) on yhtenä luokkarajana
luokat ovat (ensimmäistä lukuun ottamatta) keskihajonnan (576) pituisia

Kokkolan yhdyskunnassa on pesiä 600, joten Kokkola yhdyskunta kuuluu toiseen luokkaan 547–1123 (pesää).

d)

Aineistossa on yhteensä 47 havaintoa (30 + 9 + 4 + 4 = 47).
Lasketaan luokkien suhteelliset frekvenssit eli prosenttiosuudet.

Pesiä	f		$f\,\%$
0–546	30	30 / 47 = 0,6382… ≈ 63,8 %	63,8
547–1123	9	9 / 47 = 0,1914… ≈ 19,1 %	19,1
1124–1700	4	4 / 47 = 0,0851… ≈ 8,5 %	8,5
1701–2277	4	4 / 47 = 0,0851… ≈ 8,5 %	8,5

Suhteellisen jakauman ($f\,\%$) perusteella voidaan todeta, että pieniä yhdyskuntia on paljon, keskikokoisia jonkin verran ja suuria todella vähän.

Vastaus:

a) 25 685 pesää

b) Pienin Helsingissä (7 pesää) ja suurin Raumalla (2080 pesää)
$\bar{x} \approx 546$ (pesää) ja $s \approx 576$ (pesää)

c) Kokkolan yhdyskunta kuuluu toiseen luokkaan (esimerkiksi [547, 1123]).

d)

Pesiä	$f\,\%$
0–546	63,8
547–1123	19,1
1124–1700	8,5
1701–2277	8,5

Pieniä yhdyskuntia on paljon, keskikokoisia jonkin verran ja suuria todella vähän.