

4 JOHDATUS TILASTOTIETEeseen

4.1 Tilastojen esittäminen

ALOITA PERUSTEISTA

- 401. a)** Vuonna 2018 kirjastoja oli yhteensä $282 + 438 + 17 = 737$ ja vuonna 2020 niitä oli $280 + 439 + 12 = 731$.

Vastaus: 737 vuonna 2018 ja 731 vuonna 2020

- b)** Kirjastojen määrä väheni a-kohdan mukaan $737 - 731 = 6$ kirjastolla.

Vastaus: 6:lla kirjastolla

- c)** Lasketaan vuosien 2020 ja 2018 lainausmäärien suhde

$$\frac{70\,366}{84\,517} = 0,83256\dots = 83,256\dots\%$$

Näin ollen lainausten määrä väheni

$$100\% - 83,256\dots\% = 16,74\dots\% \approx 16,7\%$$

Vastaus: 16,7 %

- d)** Tilastomuuttujina ovat sarakkeiden otsikot: Pääkirjastojen, sivukirjastojen, laitoskirjastojen ja lainausten määrät.

Vastaus: Pääkirjastojen, sivukirjastojen, laitoskirjastojen ja lainausten määrät

- e)** Tilastoyksikköinä ovat vuodet 2018 ja 2020.

Vastaus: vuodet 2018 ja 2020

- 402.** Luku 674 € on myynnin arvo eli havaintoarvo, joten A ja II kuuluvat yhteen.

Luku 9 on frekvenssi, joten B ja I kuuluvat yhteen.

Lukuväli 1500–1999 € on luokka, joten C ja IV kuuluvat yhteen.

Taulukon tilastomuuttuja on myynti, joten D ja III kuuluvat yhteen.

Vastaus: A: II, B: I, C: IV ja D: III

- 403.** a) Kuvion mukaan 20–29-vuotiaita naisia oli noin 320 000.

Vastaus: 320 000

- b) Kuvion mukaan 30–39-vuotiaita miehiä on eniten, joten miesten suurin ikäluokka oli 30–39-vuotiaat.

Vastaus: 30–39-vuotiaat

- c) Luokkavälin pituus on peräkkäisten alarajojen erotus eli esimerkiksi 20 vuotta – 10 vuotta = 10 vuotta.

Vastaus: 10 vuotta

404. Täydennetään taulukkoon frekvenssit pylväiden korkeuksien avulla.

Arvosana	Frekvenssi	Suhteellinen frekvenssi
4	2	
5	4	
6	3	
7	9	
8	8	
9	6	
10	1	
Yhteensä		

Arvosanoja on yhteensä $2 + 4 + 3 + 9 + 8 + 6 + 1 = 33$.

Täydennetään taulukkoon suhteelliset frekvenssit jakamalla frekvenssit arvosanojen yhteismäärällä.

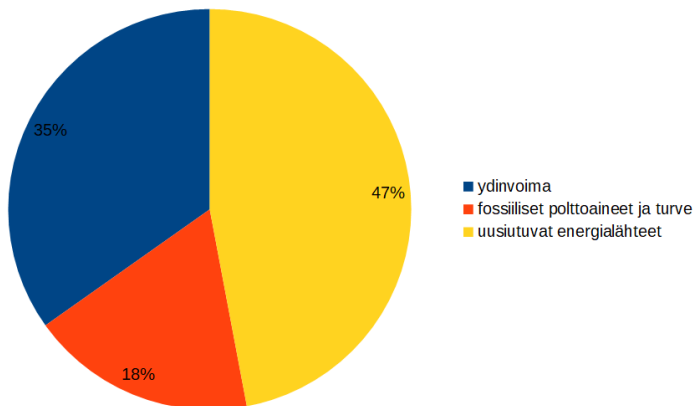
Arvosana	Frekvenssi	Suhteellinen frekvenssi
4	2	$\frac{2}{33} \approx 6,1\%$
5	4	$\frac{4}{33} \approx 12,1\%$
6	3	$\frac{3}{33} \approx 9,1\%$
7	9	$\frac{9}{33} \approx 27,3\%$
8	8	$\frac{8}{33} \approx 24,2\%$
9	6	$\frac{6}{33} \approx 18,2\%$
10	1	$\frac{1}{33} \approx 3,0\%$
Yhteensä	33	100,0 %

Vastaus:

Arvosana	Frekvenssi	Suhteellinen frekvenssi
4	2	6,1 %
5	4	12,1 %
6	3	9,1 %
7	9	27,3 %
8	8	24,2 %
9	6	18,2 %
10	1	3,0 %
Yhteensä	33	100,0 %

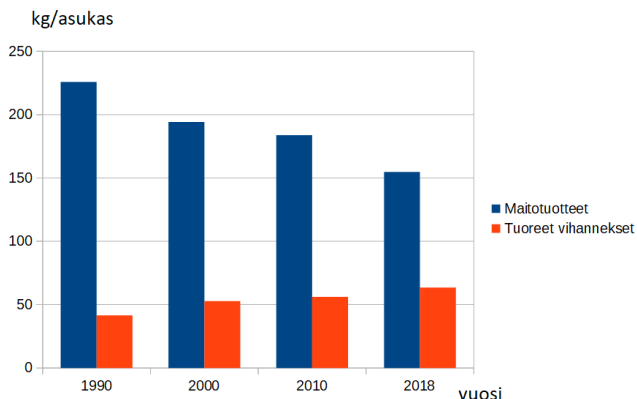
405. Piirretään ympyräkuvio taulukkolaskentaohjelmalla.

Vastaus:



406. a) Laaditaan pylväskuvio taulukkolaskentaohjelmalla siten, että kunkin vuoden kulutustietopylväät ovat rinnakkain.

Vastaus:



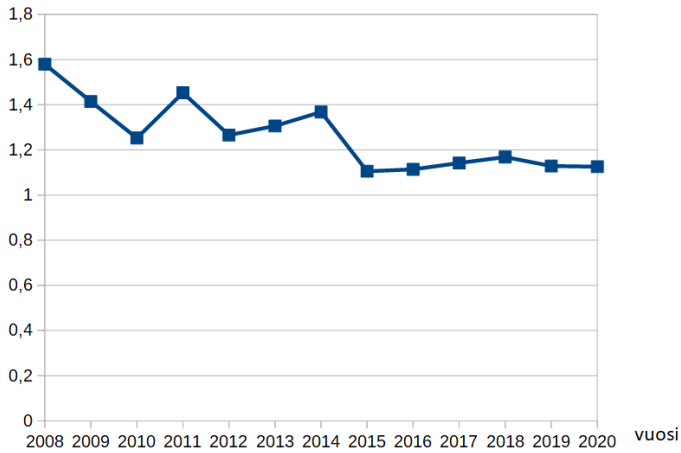
- b)** Taulukon ja kuvion avulla voidaan havaita, että maitotuotteiden asukaskohtainen kulutus vuosien myötä vähenee ja tuoreiden vihannesten kulutus kasvaa.

Vastaus: maitotuotteiden asukaskohtainen kulutus vuosien myötä vähenee ja tuoreiden vihannesten kulutus kasvaa

407. Piirretään viivakuvio taulukkolaskentaohjelmalla.

Vastaus:

euron arvo (USD)



Euron arvo on pienentynyt dollariin verrattuna.

408. a) Luokitellaan tulokset taulukon luokkiin.

Tulos (m)	1300–1790	1800–2290	2300–2790	2800–3290
Frekvenssi	2	7	10	5

Suurin frekvenssi on luokalla 2300–2790 m.

Vastaus:

Tulos (m)	1300–1790	1800–2290	2300–2790	2800–3290
Frekvenssi	2	7	10	5

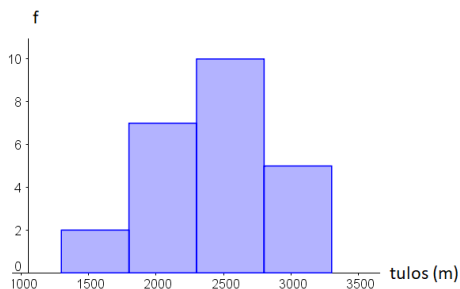
2300–2790 m

b) Luokkavälin pituus on peräkkäisten alarajojen erotus.
Ensimmäisten alarajojen erotus on $1800 \text{ m} - 1300 \text{ m} = 500 \text{ m}$.

Vastaus: 500 m

c) Laaditaan luokitellusta aineistosta histogrammi GeoGebralla.

Vastaus:



409. a) Täydennetään taulukkoon summafrekvenssit, suhteelliset frekvenssit ja suhteelliset summafrekvenssit. Ensimmäisessä kuvassa näkyvät komennot, joilla toisen kuvan luvut on saatu taulukkolaskentaohjelmalla.

1	Hinta (€)	f	Summafrekvenssi	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
2	160 000–179 999	1	=B2	=B2/\$B\$9*100 =D2	
3	180 000–199 999	1	=C2+B3	=B3/\$B\$9*100 =E2+D3	
4	200 000–219 999	2	=C3+B4	=B4/\$B\$9*100 =E3+D4	
5	220 000–239 999	4	=C4+B5	=B5/\$B\$9*100 =E4+D5	
6	240 000–259 999	3	=C5+B6	=B6/\$B\$9*100 =E5+D6	
7	260 000–279 999	2	=C6+B7	=B7/\$B\$9*100 =E6+D7	
8	280 000–299 999	2	=C7+B8	=B8/\$B\$9*100 =E7+D8	
9		=SUMMA(B2:B8)			
10	Lähde: Oikotie Asumat Oy (2020)				

1	Hinta (€)	f	Summafrekvenssi	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
2	160 000–179 999	1	1	6,7	6,7
3	180 000–199 999	1	2	6,7	13,3
4	200 000–219 999	2	4	13,3	26,7
5	220 000–239 999	4	8	26,7	53,3
6	240 000–259 999	3	11	20,0	73,3
7	260 000–279 999	2	13	13,3	86,7
8	280 000–299 999	2	15	13,3	100,0
9		15			

Tapa 2: Ilman taulukkolaskentaohjelman kaavojen hyödyntämistä:

Hinta (€)	Frekvenssi	Summa- frekvenssi	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
160 000–179 999	1	1	$\frac{1}{15} = 0,0666... \approx 6,7 \%$	$\frac{1}{15} = 0,0666... \approx 6,7 \%$
180 000–199 999	1	$1 + 1 = 2$	$\frac{1}{15} = 0,0666... \approx 6,7 \%$	$\frac{2}{15} = 0,1333... \approx 13,3 \%$
200 000–219 999	2	$2 + 2 = 4$	$\frac{2}{15} = 0,1333... \approx 13,3 \%$	$\frac{4}{15} = 0,2666... \approx 26,7 \%$
220 000–239 999	4	$4 + 4 = 8$	$\frac{4}{15} = 0,2666... \approx 26,7 \%$	$\frac{4}{15} = 0,5333... \approx 53,3 \%$
240 000–259 999	3	$8 + 3 = 11$	$\frac{3}{15} = 0,2 = 20,0 \%$	$\frac{11}{15} = 0,7333... \approx 73,3 \%$
260 000–279 999	2	$11 + 2 = 13$	$\frac{2}{15} = 0,133... \approx 13,3 \%$	$\frac{13}{15} = 0,8666... \approx 86,7 \%$
280 000–299 999	2	$13 + 2 = 15$	$\frac{2}{15} = 0,133... \approx 13,3 \%$	100 %

Vastaus:

Hinta (€)	Frekvenssi	Summa- frekvenssi	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
160 000–179 999	1	1	6,7 %	6,7 %
180 000–199 999	1	$1 + 1 = 2$	6,7 %	13,3 %
200 000–219 999	2	$2 + 2 = 4$	13,3 %	26,7 %
220 000–239 999	4	$4 + 4 = 8$	26,7%	53,3 %
240 000–259 999	3	$8 + 3 = 11$	20,0 %	73,3 %
260 000–279 999	2	$11 + 2 = 13$	13,3 %	86,7 %
280 000–299 999	2	$13 + 2 = 15$	13,3 %	100 %

- b)** a-kohdan taulukon perusteella hinta oli alle 220 000 euroa
26,6 prosentissa yksioista.

Vastaus: 26,6 %

- c)** a-kohdan taulukon perusteella hinta oli vähintään 240 000 mutta alle
280 000 euroa $20,0 + 13,3 = 33,3$ prosentissa yksioista.

Vastaus: 33,3 %

- 410.** **a)** Kuvion A pystyakselin arvot alkavat nolasta ja yksi ruutu vastaa
500 000:a liittymää. Kuvion B pystyakselin arvot alkavat 2 400 000:sta
ja yksi ruutu vastaa 50 000:a liittymää.
- b)** Jotta liittymien määrän nousu näyttäisi suhteellisesti suuremmalta kuin
se todellisuudessa on tai jotta voidaan nähdä tarkemmin, kuinka monta
tuhatta liittymää vuosittain on tullut lisää.

VAHVISTA OSAAMISTA

- 411.** Aikasarjan havainnollistamiseen sopii viivakuviio, joten A ja II kuuluvat yhteen. Aikasarjaa voidaan havainnollistaa myös pylväskuviolla, joten myös kuvio I sopii aineiston A havainnollistamiseen.

Suhteellisten osuuksien havainnollistamiseen sopii ympyräkuviio, joten B ja III kuuluvat yhteen. Suhteellisia osuuksia voidaan havainnollistaa myös pylväskuviolla, joten myös kuvio I sopii aineiston B havainnollistamiseen.

Lukumäärien havainnollistamiseen sopii pylväskuvio, joten C ja I kuuluvat yhteen.

Vastaus: A: I ja II, B: I ja III, C: I

- 412.** a) Frekvenssi on tietyn tilastomuuttujan arvon lukumäärä. Jos esimerkiksi 6 opiskelijaa sai kurssista arvosanan 8, arvosanan 8 frekvenssi on 6.
- b) Summafrekvenssi on niiden tilastomuuttujan arvojen lukumäärä, jotka ovat korkeintaan tietyn arvon suuruisia. Esimerkiksi, jos 13 opiskelijaa on korkeintaan 180 cm pitkiä, niin pituuden 180 cm summafrekvenssi on 13.
- c) Jakauma on kaikkien muuttujien arvojen frekvenssien tai suhteellisten frekvenssien kokoelma. Esimerkiksi opiskelijoiden massojen jakauma kilogrammoina voi olla seuraava:

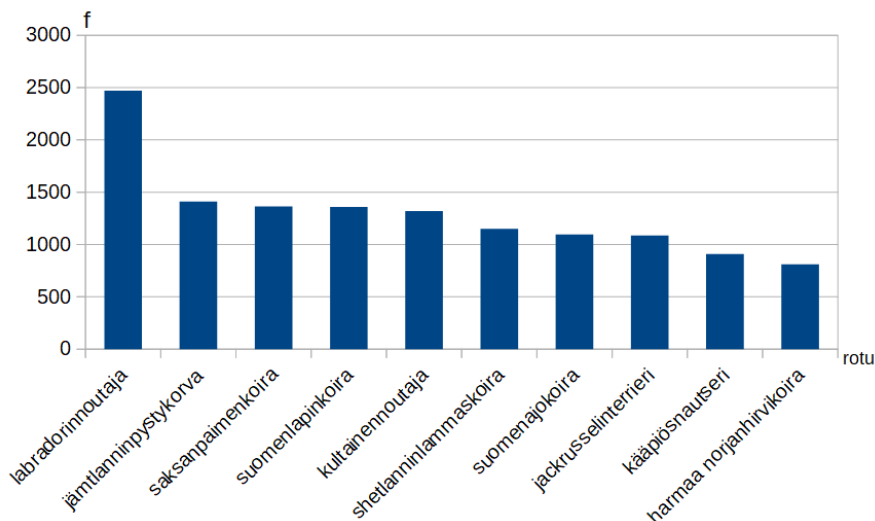
Massa (kg)	Frekvenssi
54	2
60	1
62	1
71	3
89	1

- d) Luokka on väli, jolla osa tilastomuuttujan arvoista on. Esimerkiksi luokkaan 170–174 cm kuuluvat ne opiskelijat, joiden pituus on 170 cm, 171 cm, 172 cm, 173 cm tai 174 cm.

- e) Luokkavälin pituus on peräkkäisten luokkien alarajojen erotus. Esimerkiksi jos opiskelijoiden sähköpostilaatikossa olevien viestien lukumäärät on luokiteltu luokkiin 0–49, 50–99, 100–149 ja niin edelleen, niin luokkavälin pituus on $50 - 0 = 50$.
- f) Kertymä on niiden havaintoarvojen lukumäärä tai suhteellinen osuus, jotka ovat korkeintaan tietyn arvon suuruisia. Kertymällä tarkoitetaan siis joko summafrekvenssien tai suhteellisten summafrekvenssien kokoelmaa. Esimerkiksi, jos 17 opiskelijalta kysytään, kuinka monta kertaa he olivat ulkomailla viime kalenterivuoden aikana, kertymä voi olla seuraava:

Ulkomaanmatkojen määrä viime kalenterivuoden aikana	Summafrekvenssi
0–1	8
2–3	14
4–5	16
6–7	17

413. a) Piirretään pylväskuvio taulukkolaskentaohjelmalla.



- b) Jos jakaumaa havainnollistettaisiin ympyräkuviolla, voisi saada käsityksen, että kuviossa on kaikki Suomen koirarodut, eli muita rotuja näiden 10 rodun lisäksi ei ole.

Vastaus: Voi syntyä kuva, että kuviossa on kaikki Suomen koirarodut.

- 414. a)** Viljelijän eli alkutuotannon osuus saadaan vähentämällä muut osuudet 100 %:sta.

$$100 \% - 4 \% - 12 \% - 32 \% - 48 \% = 4 \%$$

Vastaus: 4 %

- b)** Merkitään leivän hintaa kirjaimella x .

Viljelijän osuus on 10 snt = 0,1 €.

a-kohdan ja perusteella saadaan yhtälö $0,04x = 0,1$.

Ratkaistaan yhtälö.

$$\begin{aligned} 0,04x &= 0,1 & || : 0,04 \\ x &= 2,5 \end{aligned}$$

Leivän hinta on 2,5 €.

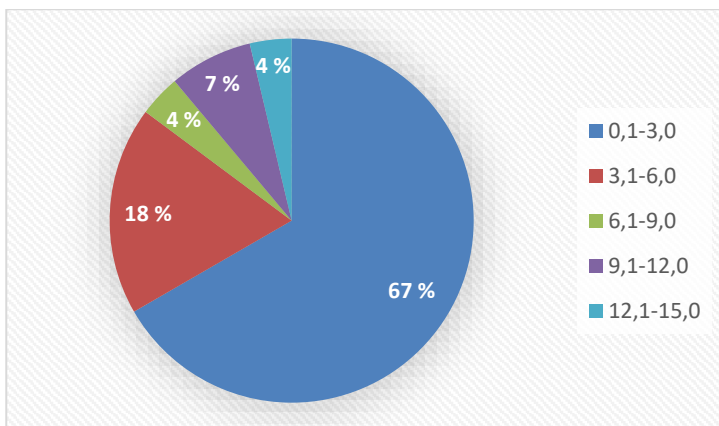
Vastaus: 2,5 €

415. Luokitellaan taulukkolaskentaohjelmalla koulumatkojen pituudet taulukkoon. Ensimmäisessä kuvassa näkyy komennot, joilla toisen kuvan arvot on saatu.

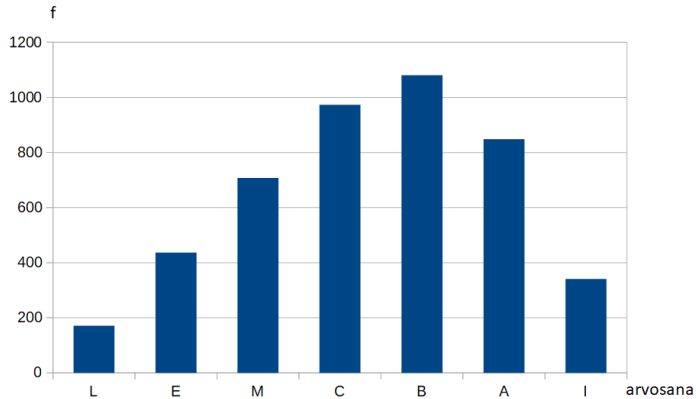
	A	B	C	D	E
1	Koulumatkan pituus (km)		Koulumatkan pituus	Yläraja	Frekvenssi
2	0,7		0,1–3,0	3	{=TAAJUUS(A2:A28;D2:D6)}
3	1,2		3,1–6,0	6	{=TAAJUUS(A2:A28;D2:D6)}
4	0,9		6,1–9,0	9	{=TAAJUUS(A2:A28;D2:D6)}
5	2,1		9,1–12,0	12	{=TAAJUUS(A2:A28;D2:D6)}
6	0,2		12,1–15,0	15	{=TAAJUUS(A2:A28;D2:D6)}
7	2,1				{=TAAJUUS(A2:A28;D2:D6)}
8	5,1				
9	0,6				

	A	B	C	D	E
1	Koulumatkan pituus (km)		Koulumatkan pituus	Yläraja	Frekvenssi
2	0,7		0,1–3,0	3	18
3	1,2		3,1–6,0	6	5
4	0,9		6,1–9,0	9	1
5	2,1		9,1–12,0	12	2
6	0,2		12,1–15,0	15	1
7	2,1				0
8	5,1				

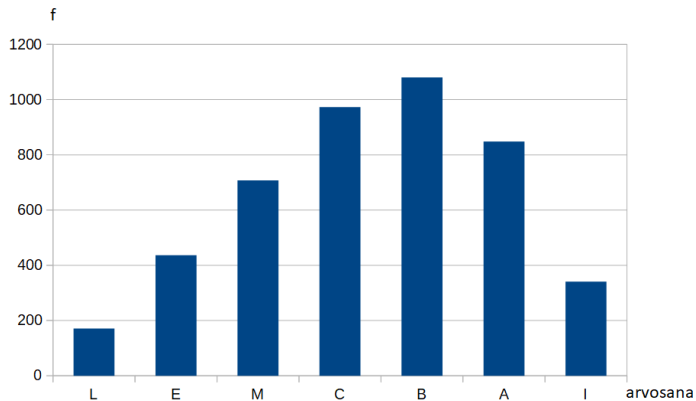
Havainnollistetaan jakaumaa ympyräkuviolla.



416. a) Piirretään pylväskuvio taulukkolaskentaohjelmalla. Kuviosta huomataan, että jakauma on epäsymmetrinen, koska kolmea alinta arvosanaa on enemmän kuin kolmea ylintä eikä huippu ole keskimmäisen arvosanan kohdalla.



Vastaus:



Jakauma on epäsymmetrinen, koska kolmea alinta arvosanaa on enemmän kuin kolmea ylintä eikä huippu ole keskimmäisen arvosanan kohdalla.

- b)** Terveystiedon arvosanoja saaneita oli yhteensä
 $171 + 436 + 707 + 973 + 1080 + 848 + 341 = 4556$.

Heistä arvosanan E tai L sai $436 + 171 = 607$.

Tämä on $\frac{607}{4556} = 0,1332... \approx 13,3\%$.

Vastaus: 13,3 %

- c)** Hyväksyttyjä suorituksia oli yhteensä $4556 - 341 = 4215$.

Hyväksytyistä suorituksista korkeintaan arvosanan C arvoisia oli
 $973 + 1080 + 848 = 2901$.

Tämä on $\frac{2901}{4215} = 0,6882... \approx 68,8\%$ terveystiedon hyväksytysti kirjoittaneista.

Vastaus: 68,8 %

417. Työntekijöiden määrät ja sektorien keskuskulmat ovat suoraan verrannollisia. Taulukoidaan tunnetut tiedot ja merkitään tuntemattomia lukumääriä kirjaimilla.

Työntekijöiden paikka	Työntekijöiden määrä	Keskuskulma (°)
töissä	x	260
koulutuksessa	8	40
sairaana	y	60

Muodostetaan verranto ja ratkaistaan siitä töissä olevien työntekijöiden määrä.

$$\begin{aligned}\frac{x}{8} &= \frac{260}{40} \quad || \cdot 8 \\ x &= \frac{260}{40} \cdot 8 \\ x &= 52\end{aligned}$$

Töissä oli siis 52 työntekijää.

Muodostetaan verranto ja ratkaistaan siitä sairaana olevien työntekijöiden määrä.

$$\begin{aligned}\frac{y}{8} &= \frac{60}{40} \quad || \cdot 8 \\ y &= \frac{60}{40} \cdot 8 \\ y &= 12\end{aligned}$$

Sairaana oli siis 12 työntekijää.

Vastaus: töissä 52, sairaana 12

418. a) Lasketaan viennin kokonaisarvo kumpanakin vuonna.

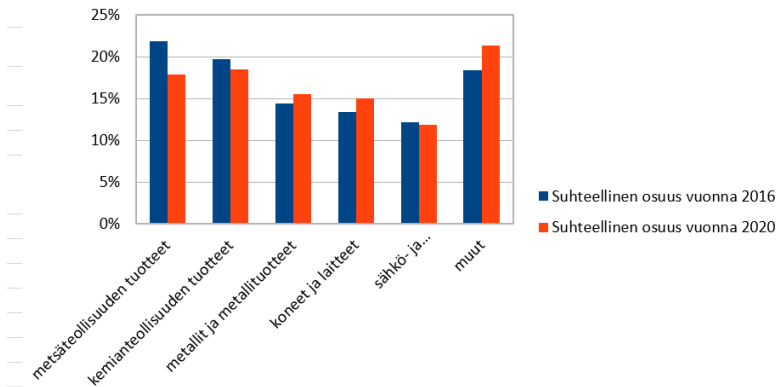
	A	B	C	
1	Vientiala	Viennin arvo vuonna 2016 (miljoonaa euroa)	Viennin arvo vuonna 2020 (miljoonaa euroa)	
2	metsäteollisuud en tuotteet	11 356	10 270	
3	kemianteollisuu den tuotteet	10 229	10 603	
4	metallit ja metallituotteet	7 473	8 870	
5	koneet ja laitteet	6 940	8 603	
6	sähkö- ja elektroniikkateo llisuuden tuotteet	6 327	6 773	
7	muut	9 541	12 213	
8	yhteensä	51866	57332	

Lasketaan ventialojen suhteelliset osuudet kumpanakin vuonna.
 Kahdessa ensimmäisessä kuvassa näkyvät komennot, joilla kahden seuraavan kuvan luvut on saatu.

	A	B	C
1	Vientiala	Viennin arvo vuonna 2016 (miljoonaa euroa)	Viennin arvo vuonna 2020 (miljoonaa euroa)
2	metsäteollisuuden tuotteet	11 356	10 270
3	kemianteollisuuden tuotteet	10 229	10 603
4	metallit ja metallituotteet	7 473	8 870
5	koneet ja laitteet	6 940	8 603
6	sähkö- ja elektroniikkateollisuuden tuotteet	6 327	6 773
7	muut	9 541	12 213
8	Yhteensä	=SUMMA(B2:B7)	=SUMMA(C2:C7)
9			
10	Vientiala	Suhteellinen osuus vuonna 2016	Suhteellinen osuus vuonna 2020
11	metsäteollisuuden tuotteet	=B2/\$B\$8*100	=C2/\$C\$8*100
12	kemianteollisuuden tuotteet	=B3/\$B\$8*100	=C3/\$C\$8*100
13	metallit ja metallituotteet	=B4/\$B\$8*100	=C4/\$C\$8*100
14	koneet ja laitteet	=B5/\$B\$8*100	=C5/\$C\$8*100
15	sähkö- ja elektroniikkateollisuuden tuotteet	=B6/\$B\$8*100	=C6/\$C\$8*100
16	muut	=B7/\$B\$8*100	=C7/\$C\$8*100

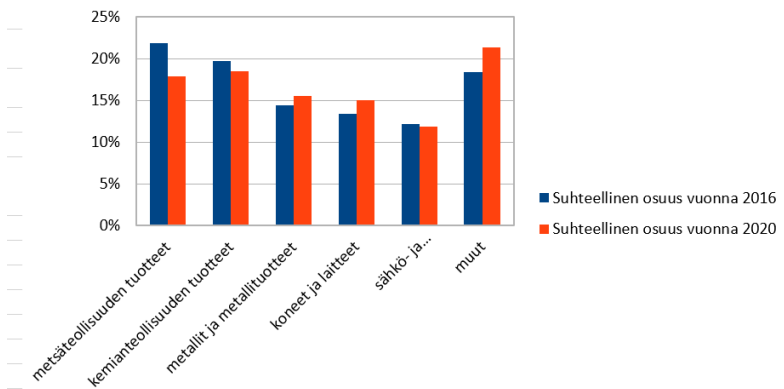
	A	B	C
1	Vientiala	Viennin arvo vuonna 2016 (miljoonaa euroa)	Viennin arvo vuonna 2020 (miljoonaa euroa)
2	metsäteollisuuden tuotteet	11 356	10 270
3	kemianteollisuuden tuotteet	10 229	10 603
4	metallit ja metallituotteet	7 473	8 870
5	koneet ja laitteet	6 940	8 603
6	sähkö- ja elektroniikkateollisuuden tuotteet	6 327	6 773
7	muut	9 541	12 213
8	Yhteensä	51866	57332
9			
10	Vientiala	Suhteellinen osuus vuonna 2016 (%)	Suhteellinen osuus vuonna 2020 (%)
11	metsäteollisuuden tuotteet	22	18
12	kemianteollisuuden tuotteet	20	18
13	metallit ja metallituotteet	14	15
14	koneet ja laitteet	13	15
15	sähkö- ja elektroniikkateollisuuden tuotteet	12	12
16	muut	18	21

Piirretään pylväskuvio.



Vastaus:

Vientiala	Suhteellinen osuus vuonna 2016	Suhteellinen osuus vuonna 2020
metsäteollisuuden tuotteet	22 %	18 %
kemianteollisuuden tuotteet	20 %	18 %
metallit ja metallituotteet	14 %	15 %
koneet ja laitteet	13 %	15 %
sähkö- ja elektroniikkateollisuuden tuotteet	12 %	12 %
muut	18 %	21 %



b) Kokonaisvientien suhde on

$$\frac{57332}{51866} = 1,10538... = 110,538...\%$$

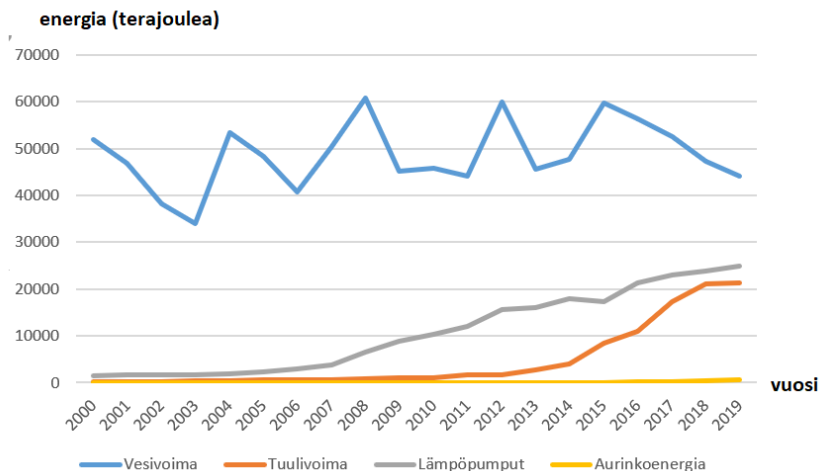
joten kokonaisvienti kasvoi

$$110,538...\% - 100\% = 10,538...\% \approx 10,5\%$$

vuodesta 2016 vuoteen 2020.

Vastaus: 10,5 %

419. a) Piirretään viivakuvio taulukkolaskentaohjelmalla.



b) Lasketaan kunkin energianlähteen suhteellinen muutos vuodesta 2000 vuoteen 2019. Ensimmäisessä kuvassa näkyvät komennot, joilla toisen kuvan luvut on saatu.

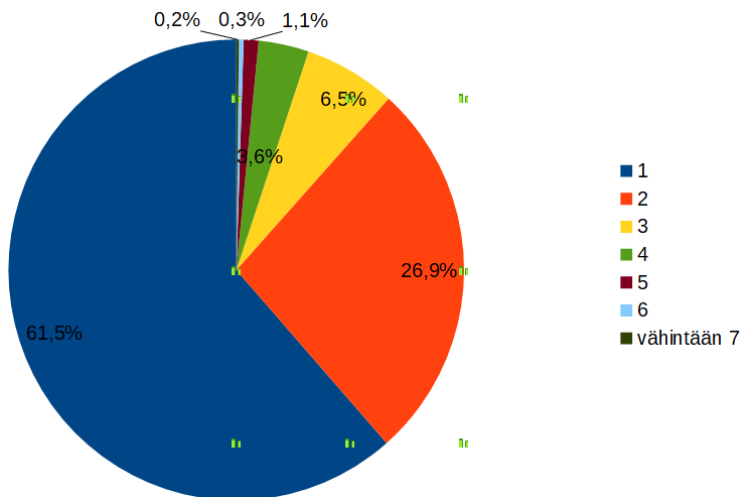
	A	B	C	D	E
37	Energialähteen käyttö terajouleina				
38	Vuosi	Vesivoima	Tuulivoima	Lämpöpumput	Aurinkoenergia
39	2000	52030	276	1450	22
40	2019	44231	21420	24881	733
41	Suhde	=B40/B39	=C40/C39	=D40/D39	=E40/E39
42	Prosentteina	=B41*100	=C41*100	=D41*100	=E41*100
43	Muutosprosentti	=(B40-B39)/B39	=(C40-C39)/C39	=(D40-D39)/D39	=(E40-E39)/E39

	A	B	C	D	E
37	Energialähteen käyttö terajouleina				
38	Vuosi	Vesivoima	Tuulivoima	Lämpöpumput	Aurinkoenergia
39	2000	52030	276	1450	22
40	2019	44231	21420	24881	733
41	Suhde	0,8501	77,6087	17,1593	33,3182
42	Prosentteina	85,0106	7760,8696	1715,9310	3331,8182
43	Muutosprosentti	-0,1499	76,6087	16,1593	32,3182

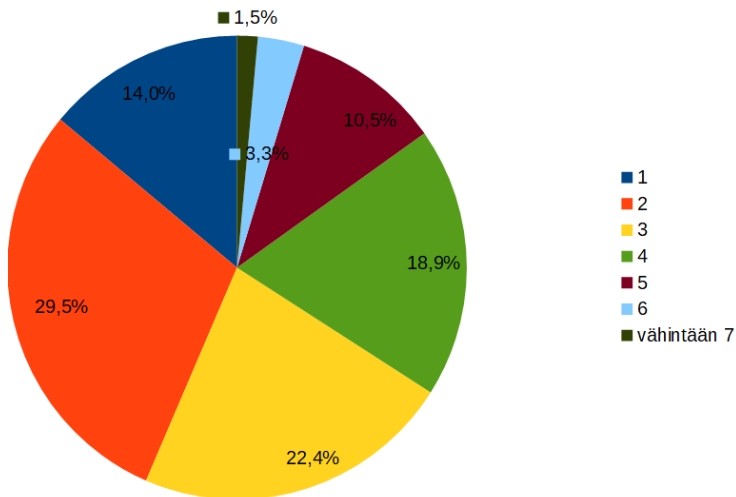
Tuulivoiman käyttö on muuttunut eniten. Se on kasvanut 7660,87 %.
Suuren muutoksen syynä on pyrkimys fossiilittomaan energiantuotantoon ja se että tuulivoimaa käytettiin vuonna 2000 olemattoman vähän.

Vastaus: Tuulivoimalla tuotetun energian määrä on muuttunut eniten.

420. a) Piirretään taulukkolaskentaohjelmalla ympyräkuvio, josta näkyy henkilömäärältään erikokoisten asuntokuntien suhteelliset osuudet vuonna 2019.



- b) Piirretään taulukkolaskentaohjelmalla ympyräkuvio, josta näkyy huoneluvultaan erikokoisten asuntojen suhteelliset osuudet vuonna 2019



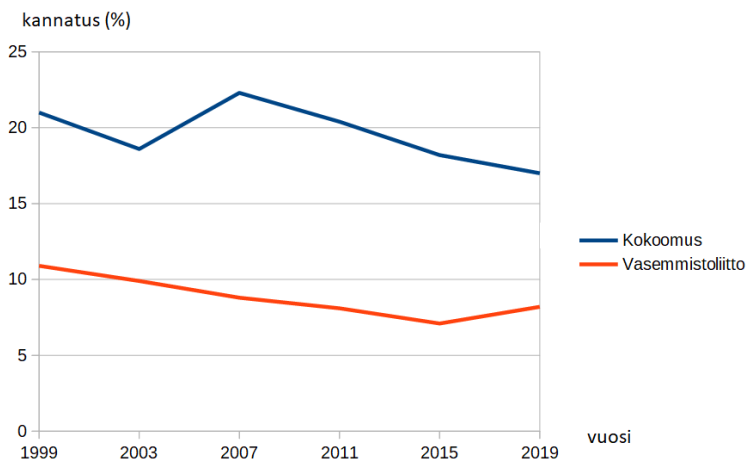
- c) Huomataan, että kerrostalojen asuntokunnista suurin osa on 1–2 hengen asuntokuntia, kun taas asuntojen huonelukumäärät ovat jakautuneet melko tasaisesti erikokoisiin asuntoihin. Jakaumien ero johtunee siitä, että asuntokuntien koot on tilastoitu vain kerrostaloista ja huonelukutilastossa on kaiken tyyppiset asunnot. Omakoti- ja rivitaloasuntojen huonelukumäärä on suurempi kuin kerrostaloasunnoissa, vaikka asuntokunnan koko olisi 1–2 henkeä.

Vastaus: Kerrostalojen asuntokunnista suurin osa on 1–2 hengen asuntokuntia, kun taas asuntojen huonelukumäärät ovat jakautuneet melko tasaisesti erikokoisiin asuntoihin.

- 421. a)** Kuvion perusteella syntyy vaikutelma, että Vasemmistoliiton kannatus olisi vuonna 2019 ollut noin viidesosa Kansallisen Kokoomuksen kannatuksesta, vaikka se on lähes puolet. Tämä johtuu siitä, että pystyakseli ei ala nollasta.

Vastaus: Pystyakseli ei ala nollasta.

- b)** Piirretään uusi kuvio katkaisematta pystyakselia arvioimalla kuvaajasta puolueiden kannatukset kunakin vuonna.

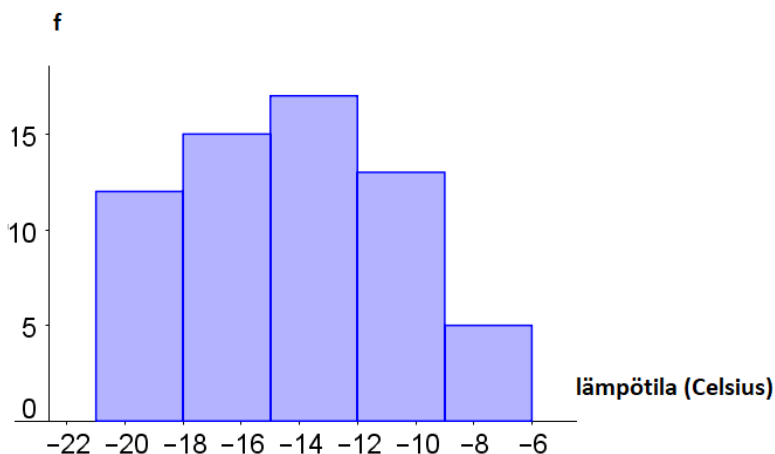


422. Täydennetään taulukkoon summafrekvenssit, suhteelliset frekvenssit ja suhteelliset summafrekvenssit taulukkolaskentaohjelmalla. Ensimmäisessä kuvassa näkyy komennot, joilla toisen kuvan luvut on saatu.

	A	B	C	D	E
1	Lämpötila (°C)	Frekvenssi		Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
2	-21,0...-18,1	12	=B2	=B2/\$B\$7*100	=C2/\$B\$7*100
3	-18,0...-15,1	15	=C2+B3	=B3/\$B\$7*100	=C3/\$B\$7*100
4	-15,0...-12,1	17	=C3+B4	=B4/\$B\$7*100	=C4/\$B\$7*100
5	-12,0...-9,1	13	=C4+B5	=B5/\$B\$7*100	=C5/\$B\$7*100
6	-9,0...-6,1	5	=C5+B6	=B6/\$B\$7*100	=C6/\$B\$7*100
7	Yhteensä	=SUMMA(B2:B6)			

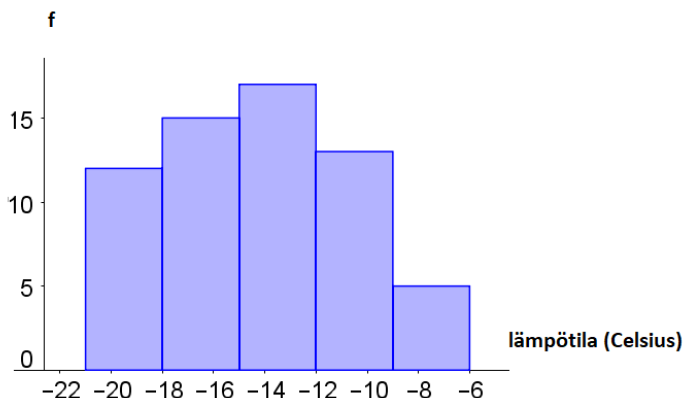
	A	B	C	D	E
1	Lämpötila (°C)	Frekvenssi		Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
2	-21,0...-18,1	12	12	19,35	19,35
3	-18,0...-15,1	15	27	24,19	43,55
4	-15,0...-12,1	17	44	27,42	70,97
5	-12,0...-9,1	13	57	20,97	91,94
6	-9,0...-6,1	5	62	8,06	100,00
7	Yhteensä	62			

Havainnollistetaan aineistoa histogrammilla.

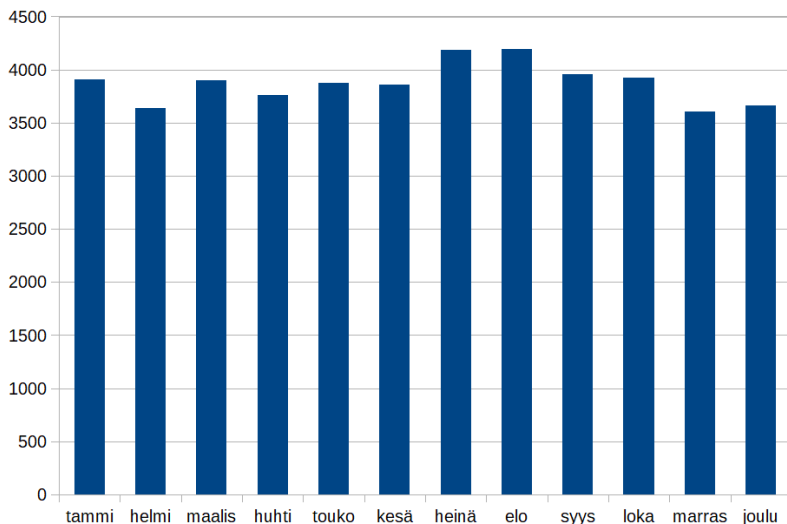


Vastaus:

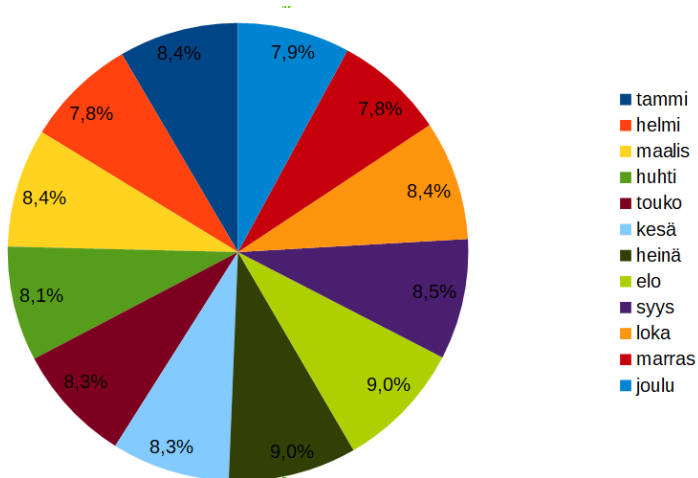
Lämpötila (°C)	Frekvenssi	Summafrekvenssi	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
-21,0...-18,1	12	12	19,4 %	19,4 %
-18,0...-15,1	15	27	24,2 %	43,5 %
-15,0...-12,1	17	44	27,4 %	71,0 %
-12,0...-9,1	13	57	21,0 %	91,9 %
-9,0...-6,1	5	62	8,1 %	100 %



423. a) Piirretään pylväskuvio taulukkolaskentaohjelmalla.



Piirretään samasta aineistosta ympyräkuvio taulukkolaskentaohjelmalla.



- b) Pylväskuvion avulla on helpompi asettaa kuukaudet suuruusjärjestykseen kuin ympyräkuvion avulla, koska pylväiden korkeuksia on helpompi verrata kuin sektoreiden keskuskulmia.

Vastaus: pylväskuvio

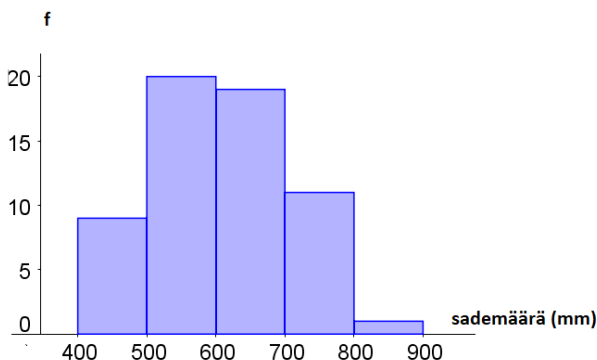
- c) Syntymien määrät vaihtelevat kuukausittain, joten kaikki kuukaudet eivät ole yhtä todennäköisiä.

Vastaus: eivät ole

- 424.** Pienin vuosittainen sademäärä on 402,8 mm ja suurin 863,4 mm. Jos luokista tehdään tasavälisiä, on luokkavälin oltava millimetreinä vähintään $\frac{863,4 - 402,8}{5} = 92,12$. Yksinkertaisuuden vuoksi valitaan luokkaväliksi 100 mm. Luokitellaan taulukon arvot taulukkolaskentaohjelmalla ja piirretään histogrammi. Ensimmäisessä kuvassa näkyy komennot, joilla toisen kuvan luvut on saatu.

	A	B	C	D	E	F
1	Vuosi	Sademäärä (mm)		Sademäärä	Yläraja	
2	1961	520,8		400,0–499,9	499,9	{=TAAJUUS(B2:B61;E2:E6)}
3	1962	760,8		500,0–599,9	599,9	{=TAAJUUS(B2:B61;E2:E6)}
4	1963	440,3		600,0–699,9	699,9	{=TAAJUUS(B2:B61;E2:E6)}
5	1964	444,3		700,0–799,9	799,9	{=TAAJUUS(B2:B61;E2:E6)}
6	1965	543		800,0–899,9	899,9	{=TAAJUUS(B2:B61;E2:E6)}
7	1966	626,4				{=TAAJUUS(B2:B61;E2:E6)}
8	1967	581				

	A	B	C	D	E	F
1	Vuosi	Sademäärä (mm)		Sademäärä	Yläraja	
2	1961	520,8		400,0–499,9	499,9	9
3	1962	760,8		500,0–599,9	599,9	20
4	1963	440,3		600,0–699,9	699,9	19
5	1964	444,3		700,0–799,9	799,9	11
6	1965	543		800,0–899,9	899,9	1
7	1966	626,4				0
8	1967	581				



425. a) Määritetään frekvenssit, summafrekvenssit, suhteelliset frekvenssit ja suhteelliset summafrekvenssit taulukkolaskentaohjelmalla. Ensimmäisessä kuvassa näkyy kaavat, joilla toisen kuvan luvut on saatu.

	A	B	C	D	E
1	Ikä (vuotta)	Frekvenssi	Summafrekvenssi	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
2	18–24	5	=B2	=B2/\$B\$7*100	=D2
3	25–34	18	=C2+B3	=B3/\$B\$7*100	=E2+D3
4	35–44	14	=C3+B4	=B4/\$B\$7*100	=E3+D4
5	45–54	15	=C4+B5	=B5/\$B\$7*100	=E4+D5
6	55–74	9	=C5+B6	=B6/\$B\$7*100	=E5+D6
7		=SUMMA(B2:B6)			

	A	B	C	D	E
1	Ikä (vuotta)	Frekvenssi	Summafrekvenssi	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
2	18–24	5	5	8,20	8,20
3	25–34	18	23	29,51	37,70
4	35–44	14	37	22,95	60,66
5	45–54	15	52	24,59	85,25
6	55–74	9	61	14,75	100,00
7		61			

Vastaus:

Ikä (vuotta)	Suhteellinen frekvenssi (%)	Summafrekvenssi	Suhteellinen summafrekvenssi (%)
18–24	8,2	5	8,2
25–34	29,5	23	37,7
35–44	23,0	37	60,7
45–54	24,6	52	85,2
55–74	14,8	61	100

- b) Vastaus saadaan laskemalla yhteen 25–34- ja 35–44- sekä 45–54- vuotiaiden määrät ja jakamalla kokonaismäärällä 61.

$$\frac{18+14+15}{61} = 0,7704... \approx 77,0 \%$$

Työntekijöistä 77,0 % on 25–54-vuotiaita.

Vastaus: 77,0 %

- c) Vertaamalla yrityksen ikäjakaumaa valtakunnalliseen työssä käyvien ikäjakaumaan huomataan, että yrityksessä on suhteellisesti enemmän 25–34-vuotiaita kuin valtakunnallisesti, joten ei voida sanoa, että yrityksen ikäjakauma noudattaisi valtakunnallista ikäjakaumaa.

Vastaus: ei voida

SYVENNÄ YMMÄRRYSTÄ

426. a) Taulukoidaan kuvaajan avulla suhteelliset summafrekvenssit.

Massa (g)	Suhteellinen summafrekvenssi
486–490	2
491–495	12
496–500	35
501–505	59
506–510	88
511–515	98
516–520	100

Lisätään taulukkoon sarake suhteelliselle frekvenssille ja lasketaan suhteelliset frekvenssit yksi kerrallaan.

Massa (g)	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
486–490	2	2
491–495	$12 - 2 = 10$	12
496–500	$35 - 12 = 23$	35
501–505	$59 - 35 = 24$	59
506–510	$88 - 59 = 29$	88
511–515	$98 - 88 = 10$	98
516–520	$100 - 98 = 2$	100

Vastaus:

Massa (g)	Suhteellinen frekvenssi
486–490	2
491–495	10
496–500	23
501–505	24
506–510	29
511–515	10
516–520	2

b) Korkeintaan 490 grammaa painavia juustoja on 2 % kaikista juustoista. 2000 kappaleen juustoerästä 2 % on $2000 \cdot 0,02 = 40$, joten B-luokan juustoja on odotettavasti 40.

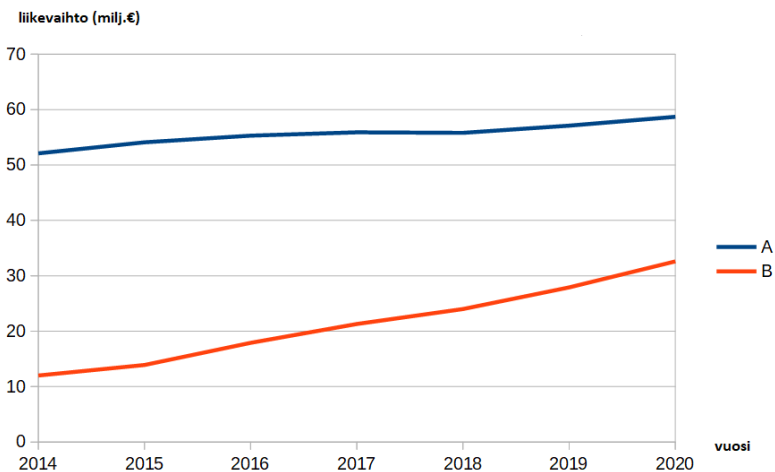
Vastaus: 40 juustoa

427. Lasketaan kummankin yrityksen kasvuprosentit taulukkolaskentaohjelmalla. Ensimmäisessä kuvassa näkyy kaavat, joilla toisen kuvan luvut on saatu.

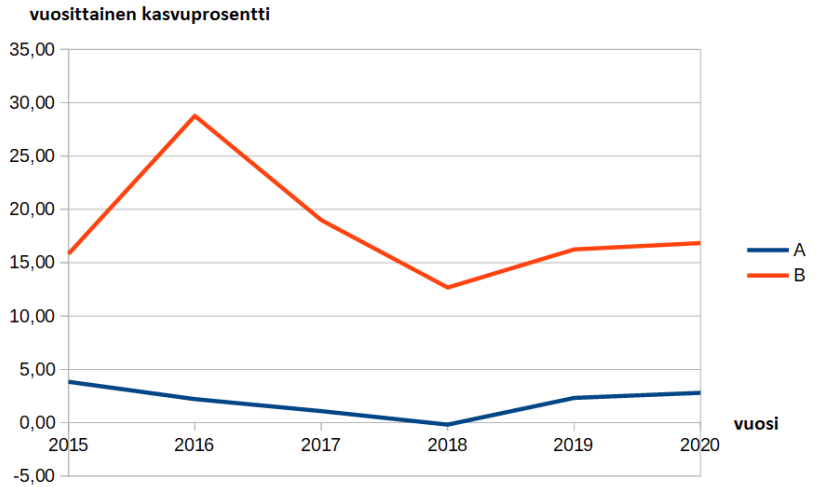
	A	B	C	D	E	F	G	H
1	Vuosi	2014	2015	2016	2017	2018	2019	2020
2	Yrityksen A liikevaihto (miljoonaa €)	52,1	54,1	55,3	55,9	55,8	57,1	58,7
3	Yrityksen A kasvuprosentti		$=(C2-B2)/B2*100$	$=(D2-C2)/C2*100$	$=(E2-D2)/D2*100$	$=(F2-E2)/E2*100$	$=(G2-F2)/F2*100$	$=(H2-G2)/G2*100$
4	Yrityksen B liikevaihto (miljoonaa €)	12	13,9	17,9	21,3	24	27,9	32,6
5	Yrityksen B kasvuprosentti		$=(C4-B4)/B4*100$	$=(D4-C4)/C4*100$	$=(E4-D4)/D4*100$	$=(F4-E4)/E4*100$	$=(G4-F4)/F4*100$	$=(H4-G4)/G4*100$

	A	B	C	D	E	F	G	H
1	Vuosi	2014	2015	2016	2017	2018	2019	2020
2	Yrityksen A liikevaihto (miljoonaa €)	52,1	54,1	55,3	55,9	55,8	57,1	58,7
3	Yrityksen A kasvuprosentti		3,84	2,22	1,08	-0,18	2,33	2,80
4	Yrityksen B liikevaihto (miljoonaa €)	12	13,9	17,9	21,3	24	27,9	32,6
5	Yrityksen B kasvuprosentti		15,83	28,78	18,99	12,68	16,25	16,85

- a) Yrityksen A liikevaihto on koko ajan ollut suurempaa kuin yrityksen B, joten yrityksen A kannalta edullista on vertailla liikevaihtoja. Piirretään viivakuvio liikevaihdoista taulukkolaskentaohjelmalla.



- b) Yrityksen B liikevaihdon kasvuprosentti on koko ajan ollut suurempi kuin yrityksen A. Yrityksen B kannalta edullista on vertailla kasvuprosentteja. Piirretään niistä viivakuviotaulukkolaskentaohjelmalla.



428. Taulukon suurin summafrekvenssi on 50, joten havaintoaineiston koko on 50. Tällöin yksi henkilö on $\frac{1}{50} = 0,02$ eli 2 % koko aineistosta.

Täydennetään tällä perusteella frekvenssin ja suhteellisen frekvenssin sarakkeita.

Kengän-numero	Frekvenssi	Suhteellinen frekvenssi	Summa-frekvenssi	Suhteellinen summafrekvenssi
36	3	$2\% \cdot 3 = 6\%$		
37	9	$2\% \cdot 9 = 18\%$		
38	8	$2\% \cdot 8 = 16\%$		
39				
40	$14\% : 2\% = 7$	14 %		
41	$6\% : 2\% = 3$	6 %		
42	$8\% : 2\% = 4$	8 %	45	
43			47	
44			50	

Täydennetään frekvenssisaraketta summafrekvenssien perusteella.

Kengän-numero	Frekvenssi	Suhteellinen frekvenssi	Summa-frekvenssi	Suhteellinen summafrekvenssi
36	3	6		
37	9	18		
38	8	16		
39				
40	7	14		
41	3	6		
42	4	8	45	
43	$47 - 45 = 2$		47	
44	$50 - 47 = 3$		50	

Kaikkien frekvenssien summan on oltava 50. Muita kokoja kuin kokoa 39 on yhteensä $3 + 9 + 8 + 7 + 3 + 4 + 2 + 3 = 39$ kpl, joten koon 39 frekvenssi on $50 - 39 = 11$. Täydennetään tämä frekvenssi ja loput suhteelliset frekvenssit taulukkoon.

Kengän-numero	Frekvenssi	Suhteellinen frekvenssi	Summa-frekvenssi	Suhteellinen summafrekvenssi
36	3	6 %		
37	9	18 %		
38	8	16 %		
39	11	$2 \% \cdot 11 = 22 \%$		
40	7	14 %		
41	3	6 %		
42	4	8 %	45	
43	2	$2 \% \cdot 2 = 4 \%$	47	
44	3	$2 \% \cdot 3 = 6 \%$	50	

Täydennetään taulukkoon summafrekvenssit ja suhteelliset summafrekvenssit.

Kengän-numero	Frekvenssi	Suhteellinen frekvenssi	Summafrekvenssi	Suhteellinen summafrekvenssi
36	3	6 %	3	6 %
37	9	18 %	$3 + 9 = 12$	$12 \cdot 2 \% = 24 \%$
38	8	16 %	$12 + 8 = 20$	$20 \cdot 2 \% = 40 \%$
39	11	22 %	$20 + 11 = 31$	$31 \cdot 2 \% = 62 \%$
40	7	14 %	$31 + 7 = 38$	$38 \cdot 2 \% = 76 \%$
41	3	6 %	$38 + 3 = 41$	$41 \cdot 2 \% = 82 \%$
42	4	8 %	45	$45 \cdot 2 \% = 90 \%$
43	2	4 %	47	$47 \cdot 2 \% = 94 \%$
44	3	6 %	50	$50 \cdot 2 \% = 100 \%$

Vastaus:

Kengännumero	Frekvenssi	Suhteellinen frekvenssi	Summafrekvenssi	Suhteellinen summafrekvenssi
36	3	6	3	6
37	9	18	12	24
38	8	16	20	40
39	11	22	31	62
40	7	14	38	76
41	3	6	41	82
42	4	8	45	90
43	2	4	47	94
44	3	6	50	100

429. a) Koska yhdenmuotoisten kappaleiden tilavuuksien suhde on pituuksien suhteen kolmas potenssi, kuutioiden tilavuuksien suhde on

$$\frac{10,4^3}{5,5^3} = 6,761... \approx 6,8. \text{ Ruotsin väkiluku vaikuttaa kuutioiden}$$

perusteella noin seitsemänkertaiselta Suomen väkilukuun verrattuna,

vaikka se on todellisuudessa $\frac{10,4}{5,5} = 1,890... \approx 1,9$ -kertainen.

Vastaus: Ruotsin väkiluku vaikuttaa kuutioiden perusteella noin seitsemänkertaiselta Suomen väkilukuun verrattuna, vaikka se on todellisuudessa noin 1,9-kertainen.

- b) Tilavuuksien suhde on mittakaavan kuutio. Muodostetaan yhtälö ja ratkaistaan sen avulla Ruotsin väkilukua esittävän kuution särmän pituus x .

$$\left(\frac{5,5}{x}\right)^3 = \frac{5,5}{10,4}$$

$$\frac{5,5^3}{x^3} = \frac{5,5}{10,4}$$

$$5,5x^3 = 10,4 \cdot 5,5^3 \quad \| : 5,5$$

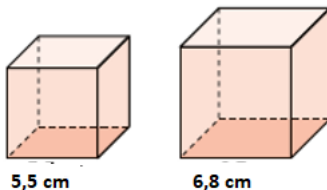
$$x^3 = \frac{10,4 \cdot 5,5^3}{5,5}$$

$$x = \sqrt[3]{\frac{10,4 \cdot 5,5^3}{5,5}}$$

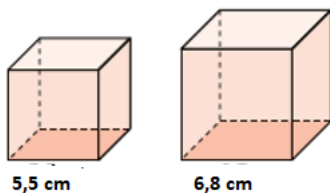
$$x = 6,801...$$

Ruotsin väkilukua esittävän kuution särmän pituuden pitäisi olla noin 6,8 cm.

Piirretään kuutiot.

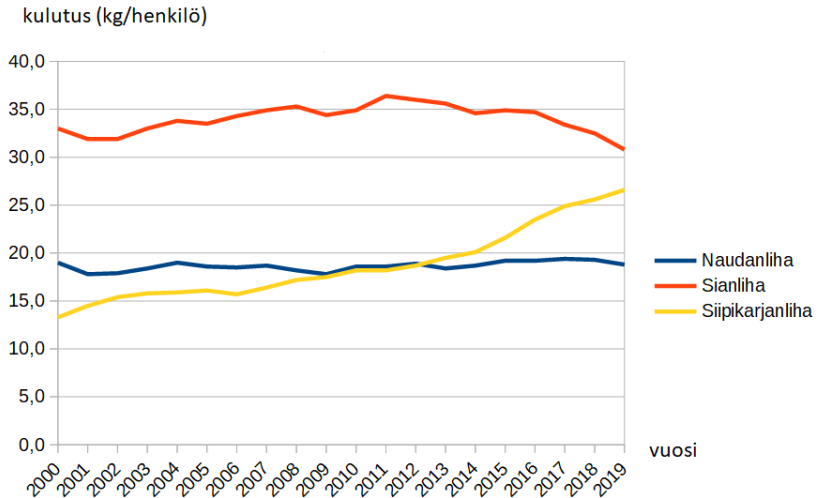


Vastaus: 6,8 cm,



- 430.** Taulukossa on luonnonvarakeskuksen tilastoista vuonna 2020 löytynyt aineisto lihan vuosikulutuksesta (kg/henkilö).

	Naudanliha	Sianliha	Siipikarjanliha
2000	19,0	33,0	13,3
2001	17,8	31,9	14,5
2002	17,9	31,9	15,4
2003	18,4	33,0	15,8
2004	19,0	33,8	15,9
2005	18,6	33,5	16,1
2006	18,5	34,3	15,7
2007	18,7	34,9	16,4
2008	18,2	35,3	17,2
2009	17,8	34,4	17,5
2010	18,6	34,9	18,2
2011	18,6	36,4	18,2
2012	18,9	36,0	18,7
2013	18,4	35,6	19,5
2014	18,7	34,6	20,1
2015	19,2	34,9	21,6
2016	19,2	34,7	23,5
2017	19,4	33,4	24,9
2018	19,3	32,5	25,6
2019	18,8	30,8	26,6



Aineiston ja sen perusteella laaditun viivakuvion avulla voidaan sanoa, että siipikarjan lihan kulutus (kg/henkilö) on kasvanut 2000-luvun alusta lähtien. Sianlihan kulutus on kääntynyt laskuun vuonna 2016, kun taas naudanlihan kulutus on säilynyt jokseenkin ennallaan. Vuoden 2012 aikoihin siipikarjanlihan kulutus on ylittänyt naudanlihan kulutuksen.

4.2 Sijaintilukuja

ALOITA PERUSTEISTA

431. Moodi merkitään Mo , mediaani merkitään Md ja keskiarvo merkitään $\bar{x}$.

Vastaus: A: II, B: III ja C: I

432. Moodi Mo on yleisin arvo ja mediaani Md suuruusjärjestykseen asetetuista arvoista keskimmäinen tai kahden keskimmäisen keskiarvo.

a) Lukujen 1, 1, 2, 4 ja 7 moodi on $Mo = 1$ ja mediaani $Md = 2$.

$$\text{Keskiarvo on } \bar{x} = \frac{1+1+2+4+7}{5} = \frac{15}{5} = 3.$$

Vastaus: $Mo = 1$, $Md = 2$, $\bar{x} = 3$

b) Kaikki neljä lukua ovat yhtä yleisiä, joten moodeja ovat 3, 6, 7 ja 8.

$$\text{Keskimmäisiä ovat luvut 6 ja 7, joten mediaani on } Md = \frac{6+7}{2} = 6,5.$$

$$\text{Keskiarvo on } \bar{x} = \frac{3+6+7+8}{4} = \frac{24}{4} = 6.$$

Vastaus: $Mo = 3$, $Mo = 6$, $Mo = 7$ ja $Mo = 8$, $Md = 6,5$, $\bar{x} = 6$

c) Järjestetään ensin luvut suuruusjärjestykseen: 10, 30, 40, 40.
Moodi on $Mo = 40$.

Keskimmäisiä ovat luvut 30 ja 40, joten mediaani on

$$Md = \frac{30+40}{2} = 35.$$

$$\text{Keskiarvo on } \bar{x} = \frac{10+30+40+40}{4} = \frac{120}{4} = 30.$$

Vastaus: $Mo = 40$, $Md = 35$, $\bar{x} = 30$

- 433.** Keskilukujen määrittämistä varten järjestetään vastaukset suuruusjärjestykseen: 1, 1, 2, 2, 2, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5. Vastauksia on yhteensä 16.

a) Vastausten moodi on 4, koska se on yleisin vastaus.

Vastaus: $M_o = 4$

b) Vastauksia on 16, joten 8. ja 9. vastaus ovat keskimmäiset vastaukset. Sekä 8. että 9. vastaus on 4, joten mediaani on 4.

Vastaus: $M_d = 4$

- 434.** Mediaani on se opintopisteiden määrä, jonka kohdalla suhteellinen summafrekvenssi on 50 %. Kuvaajan perusteella mediaani on noin 78.

Vastaus: 78 opintopistettä

435. Määritetään keskiarvo, moodi ja mediaani taulukkolaskentaohjelman avulla. Ensimmäisessä kuvassa näkyvät komennot, joilla toisen kuvan luvut on saatu.

	A	B
1	Askelten määrä	
2	6790	
3	7712	
4	7543	
5	8924	
6	9972	
7	5214	
8	4505	
9	9972	
10	10076	
11	9521	
12	10374	
13	10642	
14	11021	
15	10989	
16	8528	
17	11153	
18	keskiarvo	=KESKIARVO(A2:A17)
19	moodi	=MOODI(A2:A17)
20	mediaani	=MEDIAANI(A2:A17)

	A	B
1	Askelten määrä	
2	6790	
3	7712	
4	7543	
5	8924	
6	9972	
7	5214	
8	4505	
9	9972	
10	10076	
11	9521	
12	10374	
13	10642	
14	11021	
15	10989	
16	8528	
17	11153	
18	keskiarvo	8933,5
19	moodi	9972
20	mediaani	9746,5

Keskiarvo on 8933,5, moodi 9972 ja mediaani 9746,5.

Vastaus: $\bar{x} = 8933,5$, $Mo = 9972$, $Md = 9746,5$

436. Määritetään keskiarvo ja mediaani.

Tapa 1: GeoGebralla

n	209
Keskiarvo	501735.5742
σ	2343564.9056
s	2349191.7202
Σx	104862735
Σx^2	1200503326071971
Min	62
Q1	7446.5
Mediaani	33887
Q3	249234.5
Max	27720709

Tapa 2: Taulukkolaskentaohjelmalla

Ensimmäisessä kuvassa näkyy kaavat, joilla toisen kuvan luvut on saatu.

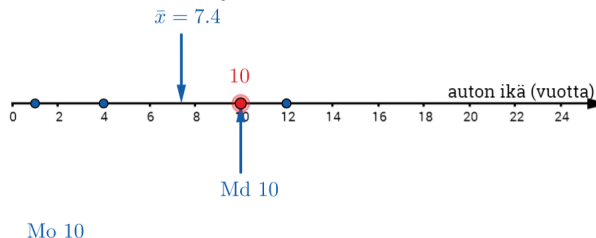
	A	B	C	D	E	F
1	Lyhenne	Sijoitus	Maa / alue	Bruttokansantuote dollareina v. 2023		
2	USA	1	Yhdysvallat	27720709		
3	CHN	2	Kiina	17794783	keskiarvo	=KESKJARVO(D2:D218)
4	DEU	3	Saksa	4525704	mediaani	=MEDIAANI(D2:D218)
5	JPN	4	Japani	4204495		
6	IND	5	Intia	3567552		

	A	B	C	D	E	F
1	Lyhenne	Sijoitus	Maa / alue	Bruttokansantuote dollareina v. 2023		
2	USA	1	Yhdysvallat	27720709		
3	CHN	2	Kiina	17794783	keskiarvo	501735,57
4	DEU	3	Saksa	4525704	mediaani	33887
5	JPN	4	Japani	4204495		
6	IND	5	Intia	3567552		

Keskiarvoksi saadaan 501 735,57 miljoonaa dollaria ja mediaaniksi 33 887 miljoonaa dollaria.

Vastaus: keskiarvo 501 735,57 miljoonaa dollaria ja mediaani 33 887 miljoonaa dollaria

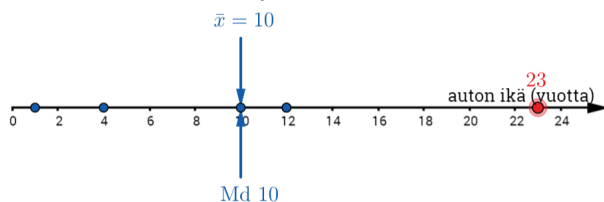
437. a) Tutkitaan appletin avulla, mikä viidennen auton iän pitäisi olla, jotta ikien moodi olisi 10 vuotta.



Viidennen auton iän on oltava 10 vuotta.

Vastaus: 10 vuotta

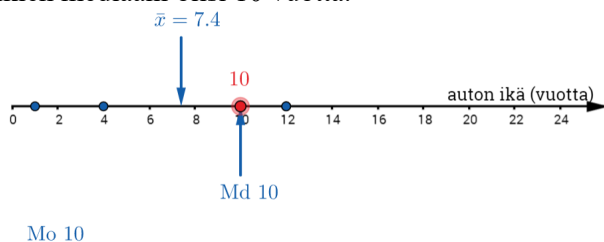
- b) Tutkitaan appletin avulla, mikä viidennen auton iän pitäisi olla, jotta ikien keskiarvo olisi 10 vuotta.



Viidennen auton iän on oltava 23 vuotta.

Vastaus: 23 vuotta

- c) Tutkitaan appletin avulla, mikä viidennen auton iän pitäisi olla, jotta ikien mediaani olisi 10 vuotta.



Ikien mediaani on 10 vuotta, kun viidennen auton ikä on vähintään 10 vuotta.

Vastaus: vähintään 10 vuotta

438. a) Täydennetään taulukkoon summafrekvenssit ja suhteelliset summafrekvenssit.

Poissaoloja (oppituntia)	Frekvenssi	Summafrekvenssi	Suhteellinen summafrekvenssi
0–1	11	11	$\frac{11}{29} = 0,379... \approx 38\%$
2–3	10	$11 + 10 = 21$	$\frac{21}{29} = 0,724... \approx 72\%$
4–5	5	$21 + 5 = 26$	$\frac{26}{29} = 0,896... \approx 90\%$
6–7	3	$26 + 3 = 29$	100 %

Vastaus:

Poissaoloja (oppituntia)	Summafrekvenssi	Suhteellinen summafrekvenssi
0–1	11	38 %
2–3	21	72 %
4–5	26	90 %
6–7	29	100 %

- b) Mediaaniluokka on luokka, jonka kohdalla summafrekvenssi ylittää 50 %. Mediaaniluokka on siis 2–3 oppituntia.

Vastaus: 2–3 oppituntia

- 439. a)** Yöunen pituuden moodiluokka on luokka, jonka frekvenssi on suurin. Moodiluokka on siis 8–9 h.

Vastaus: 8–9 h

- b)** Täydennetään taulukkoon todelliset rajat ja luokkakeskukset.

Yöunen pituus (h)	Todellinen alaraja (h)	Todellinen yläraja (h)	Luokkakeskus (h)	Frekvenssi
4–5	3,5	5,5	$\frac{3,5 + 5,5}{2} = 4,5$	3
6–7	5,5	7,5	6,5	11
8–9	7,5	9,5	8,5	14
10–11	9,5	11,5	10,5	2

Vastaus:

Yöunen pituus (h)	Todellinen alaraja (h)	Todellinen yläraja (h)	Luokkakeskus (h)	Frekvenssi
4–5	3,5	5,5	4,5	3
6–7	5,5	7,5	6,5	11
8–9	7,5	9,5	8,5	14
10–11	9,5	11,5	10,5	2

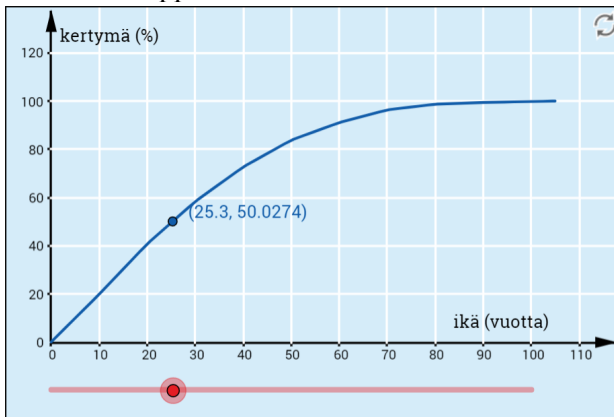
- c)** Lasketaan keskimääräinen yöunen pituus luokkakeskusten avulla.

$$\bar{x} = \frac{3 \cdot 4,5 + 11 \cdot 6,5 + 14 \cdot 8,5 + 2 \cdot 10,5}{3 + 11 + 14 + 2} = \frac{225}{30} = 7,5$$

Lukiolaisen keskimääräinen yöunen pituus on 7,5 h.

Vastaus: 7,5 h

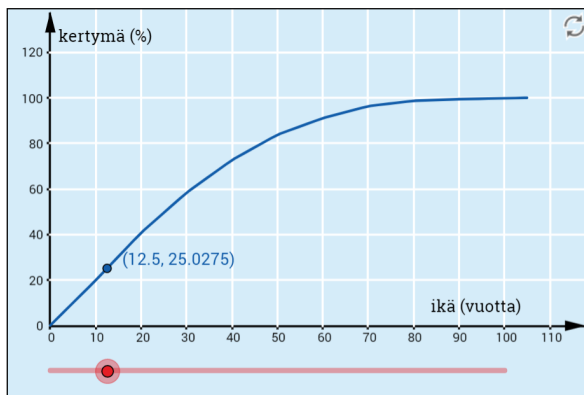
440. a) Määritetään appletin avulla mediaani-ikä.



Mediaani-ikä on appletin perusteella noin 25 vuotta.

Vastaus: 25 vuotta

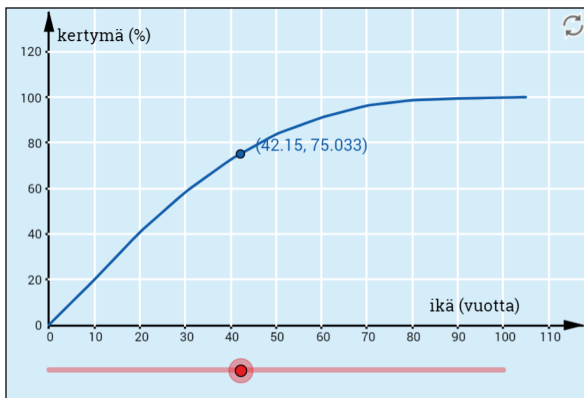
b) Määritetään appletin avulla ikä, jota nuorempia on 25 % intialaisista.



Appletin perusteella noin 25 % intialaisista on alle 13-vuotiaita.

Vastaus: 13 vuotta

- c) Määritetään appletin avulla ikä, jota vanhempia on 25 % intialaisista.



Appletin perusteella 25 % intialaisista on yli 42-vuotiaita.

Vastaus: 42 vuotta

- d) Intian syntyvyys ja toisaalta myös lapsikuolleisuus ovat suhteellisesti suurempia kuin Suomessa. Tästä johtuen lapsia on suhteellisesti paljon, mutta he eivät elä niin vanhoiksi kuin Suomessa.

441. Lasketaan erikokoisten lasien frekvenssit suhteellisten frekvenssien avulla, kun tiedetään, että lasia on yhteensä 300.

Tilavuus (dl)	Suhteellinen frekvenssi	Frekvenssi
2,0	48 %	$0,48 \cdot 300 = 144$
3,4	35 %	$0,35 \cdot 300 = 105$
5,6	17 %	$0,17 \cdot 300 = 51$

Lasketaan tilavuuksien keskiarvo frekvenssien avulla.

$$\bar{x} = \frac{144 \cdot 2,0 + 105 \cdot 3,4 + 51 \cdot 5,6}{300} = 3,102 \approx 3,1$$

Juomalasien tilavuuksien keskiarvo on noin 3,1 dl.

Vastaus: 3,1 dl

442. Lasketaan kunkin moduulin laajuuden ja arvosanojen tulot yhteen
 $1 \cdot 8 + 3 \cdot 6 + 2 \cdot 7 + 2 \cdot 6 + 2 \cdot 8 = 68$.

B-oppimäärän laajuus on yhteensä $1 + 3 + 2 + 2 + 2 = 10$.

Painotettu keskiarvo on

$$\bar{x} = \frac{68}{10} = 6,8 \approx 7.$$

Villen vieraan kielen B-oppimäärän arvosana on 7.

Vastaus: 7

VAHVISTA OSAAMISTA

443. a) Silmien värien moodi on sininen, koska se on yleisin havaintoarvo.

Vaatekokojen moodi on 42, koska se on yleisin havaintoarvo.

Kokoja on yhteensä $5 + 8 + 16 + 18 + 15 = 62$.

Puolet siitä on $\frac{62}{2} = 31$, joten 31. ja 32. koko ovat keskimmäisiä.

Kokoja 36, 38 ja 40 on yhteensä $5 + 8 + 16 = 29$ ja niitä seuraavat 18 ovat kokoa 42, joten keskimäinen koko on 42.

Lasketaan aineiston B keskiarvo

$$\bar{x} = \frac{5 \cdot 36 + 8 \cdot 38 + 16 \cdot 40 + 18 \cdot 42 + 15 \cdot 44}{62} = \frac{2540}{62} = 40,967... \approx 41.$$

Vastaus: A: Mo = sininen, B: Mo = 42, Md = 42, $\bar{x} \approx 41$

- b) Mediaania ei ole mahdollista määrittää, koska havaintoarvoja ei voida asettaa suuruusjärjestykseen. Keskiarvoa ei voi määrittää, koska havaintoarvot eivät ole lukuarvoja.

444. a) Moodiluokka on luokka, jonka frekvenssi on suurin. Moodiluokka on siis 160–164 cm.

Vastaus: 160–164 cm

- b) Lasketaan summafrekvenssit mediaanin määrittämistä varten.

Pituus (cm)	Frekvenssi	Summafrekvenssi
150–154	3	3
155–159	6	$3 + 6 = 9$
160–164	17	$9 + 17 = 26$
165–169	15	$26 + 15 = 41$
170–174	12	$41 + 12 = 53$
175–179	11	$53 + 11 = 64$
180–184	14	$64 + 14 = 78$
185–189	6	$78 + 6 = 84$
190–194	2	$84 + 2 = 86$

Mediaaniluokka on luokka, jossa keskimäinen pituus on.

Pituuksia on yhteensä 86 ja $\frac{86}{2} = 43$.

Pituudet 43. ja 44. ovat keskimmäisiä.

43. ja 44. pituus ovat luokassa 170–174 cm, joten se on mediaaniluokka.

Vastaus: 170–174 cm

- c) Täydennetään keskiarvon laskemista varten taulukkoon luokkakeskukset.

Pituus (cm)	Frekvenssi	Luokkakeskus (cm)
150–154	3	$\frac{149,5 + 154,5}{2} = 152$
155–159	6	157
160–164	17	162
165–169	15	167
170–174	12	172
175–179	11	177
180–184	14	182
185–189	6	187
190–194	2	192

Keskiarvo on

$$\bar{x} = \frac{3 \cdot 152 + 6 \cdot 157 + 17 \cdot 162 + \dots + 2 \cdot 192}{86} = 171,186\dots$$

Keskiarvo on noin 172 cm.

Vastaus: 172 cm

445. Kopiaidaan arvot taulukkolaskentaohjelmaan.
Määritetään keskiarvo, moodi ja mediaani ohjelman avulla.

	A	B	C
1	Voitto		
2	25	Keskiarvo	=KESKIARVO(A2:A21)
3	-3	Moodi	=MOODI(A2:A21)
4	-2	Mediaani	=MEDIAANI(A2:A21)
5	-9		
6	17		
7	-6		
8	-8		
9	-2		
10	0		
11	-4		
12	-11		
13	-5		
14	-3		
15	-2		
16	-9		
17	-12		
18	10		
19	-9		
20	-2		
21	-1		

	A	B	C
1	Voitto		
2	25	Keskiarvo	-1,8
3	-3	Moodi	-2
4	-2	Mediaani	-3
5	-9		
6	17		
7	-6		
8	-8		
9	-2		
10	0		
11	-4		
12	-11		
13	-5		
14	-3		
15	-2		
16	-9		
17	-12		
18	10		
19	-9		
20	-2		
21	-1		

Voiton keskiarvo on -1,8 euroa, moodi -2 euroa ja mediaani -3 euroa.
Negatiiviset voitot tarkoittavat sitä, että pelaajat ovat jääneet tappiolle.

Vastaus: $\bar{x} = -1,8$ €, $M_o = -2$ €, $M_d = -3$ €. Pelaajat jäivät keskimäärin tappiolle.

- 446.** Keskilukujen määrittämistä varten järjestetään arvot suuruusjärjestykseen: 1, 5, 6, 6, 7, 9, 9, 9, 12, 12. Arvoja on yhteensä 10 kpl.

Aineiston moodi on $Mo = 9$, koska se on yleisin arvo.

Aineiston mediaani on viidennen ja kuudennen arvon keskiarvo, eli

$$Md = \frac{7+9}{2} = 8.$$

$$\text{Aineiston keskiarvo on } \bar{x} = \frac{1+5+2 \cdot 6+7+3 \cdot 9+2 \cdot 12}{10} = \frac{76}{10} = 7,6.$$

- a)** Moodi on 9 ja mediaani 8, joten moodi on suurempi kuin mediaani.
Väite on siis epätosi.

Vastaus: epätosi, suurempi

- b)** Mediaani on 8, joten väite on tosi.

Vastaus: tosi

- c)** Mediaani on 8 ja keskiarvo 7,6, joten väite on tosi.

Vastaus: tosi

- d)** Keskiarvo on 7,6, joten väite on epätosi.

Vastaus: epätosi; 7,6

- 447.** Mediaani on 7 kuviossa II, koska havaintoarvot ovat 4, 5, 6, 6, 7, 7, 7, 7, 8 ja 9. Havaintoarvoja on yhteensä 10, joten viides ja kuudes havaintoarvo ovat keskimmäisiä. Viides ja kuudes havaintoarvo ovat molemmat 7, joten mediaani on 7. Tunnusluku A kuuluu kuvioon II.

Moodi on 6 kuviossa III, koska lukua 6 kuvaava sektori on suurin.

Tunnusluku B kuuluu kuvioon III.

Keskiarvo on 7 kuviossa I, koska luvut ovat jakautuneet symmetrisesti luvun 7 ympärille. Tunnusluku C kuuluu kuvioon I.

Vastaus: A: II, B: III ja C: I

448. Appletin avulla saadaan lukujen 1, 2, 3, 4, 7, 7, 8 ja 9 moodiksi 7, mediaaniksi 5,5 ja keskiarvoksi 5,13.

- a) Kun pienin luvuista kasvaa seitsemällä, moodina ovat luvut 7 ja 8, mediaani on 7 ja keskiarvo on 6. Moodiksi tulevat 7 ja 8, mediaani kasvaa 1,5:llä ja keskiarvo kasvaa 0,87:llä.

Vastaus: Moodiksi tulevat 7 ja 8, mediaani kasvaa 1,5:llä ja keskiarvo kasvaa 0,87:llä.

- b) Kun suurin luvuista kaksinkertaistuu, on moodi 7, mediaani 5,5 keskiarvo 6,25. Moodi ja mediaani eivät muutu, keskiarvo kasvaa 1,12:lla.

Vastaus: Moodi ja mediaani eivät muutu, keskiarvo kasvaa 1,12:lla.

- c) Kun jokainen luku kasvaa seitsemällä, moodi 14, mediaani 12,5 ja keskiarvo 12,13. Kaikki kolme tunnuslukua kasvavat 7:llä.

Vastaus: Kaikki kolme tunnuslukua kasvavat 7:llä.

- d) Kun jokainen luku kaksinkertaistuu, on moodi 14, mediaani 11 ja keskiarvo 10,26. Jokainen tunnusluku siis kaksinkertaistuu.

Vastaus: Kaikki kolme tunnuslukua kaksinkertaistuvat.

449. Valitaan luvut esimerkiksi niin, että luvuista kaksi on samoja ja kaikki muut keskenään eri lukuja, ja yksi luku poikkeaa muista suuresti. Tällaiset luvut ovat esimerkiksi 0, 0, 1, 2 ja 10. Lukujen moodi on 0 ja mediaani 1.

Keskiarvo on $\frac{0+0+1+2+10}{5} = 2,6$.

Vastaus: esim. 0, 0, 1, 2 ja 10

- 450.** Koska moduulien opintopistemäärä ovat yhtä suuret voidaan tehtävä ratkaista aritmeettisen keskiarvon avulla. Merkitään viidennen moduulin vähimmäisarvosanaa kirjaimella x ja ratkaistaan se yhtälön avulla.

$$\begin{aligned}\frac{8+7+6+8+x}{5} &= 7,5 \\ \frac{29+x}{5} &= 7,5 \quad || \cdot 5 \\ 29+x &= 37,5 \\ x &= 37,5 - 29 \\ x &= 8,5\end{aligned}$$

Karoliinan pitäisi saada viidennestä moduulista vähintään arvosana yli 8,5 ja koska moduulien arvosanat ovat kokonaislukuja, on hänen saatava arvosana 9.

Vastaus: 9

- 451. a)** Kokojen moodiluokka on 40, koska se on yleisin koko.

Määritetään suhteelliset summafrekvenssit taulukkolaskentaohjelmalla. Ensimmäisessä kuvassa näkyy kaavat, joilla toisen kuvan luvut on saatu.

	A	B	C
1	Kokonro	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
2	32	0,20	=B2
3	34	3,20	=C2+B3
4	36	8,40	=C3+B4
5	38	17,80	=C4+B5
6	40	20,00	=C5+B6
7	42	13,70	=C6+B7
8	44	11,10	=C7+B8
9	46	10,60	=C8+B9
10	48	7,60	=C9+B10
11	50	4,10	=C10+B11
12	52	1,50	=C11+B12
13	54	1,10	=C12+B13
14	56	0,70	=C13+B14

	A	B	C
1	Kokonro	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
2	32	0,20	0,20
3	34	3,20	3,4
4	36	8,40	11,8
5	38	17,80	29,6
6	40	20,00	49,6
7	42	13,70	63,3
8	44	11,10	74,4
9	46	10,60	85
10	48	7,60	92,6
11	50	4,10	96,7
12	52	1,50	98,2
13	54	1,10	99,3
14	56	0,70	100

Suhteellinen summafrekvenssi ylittää 25 % koon 38 kohdalla, joten alakvartiili on 38.

Suhteellinen summafrekvenssi ylittää 50 % koon 42 kohdalla, joten mediaani on 42.

Suhteellinen summafrekvenssi ylittää 75 % koon 46 kohdalla, joten yläkvartiili on 46.

Koska vaatteiden kokonaismäärää ei ole kerrottu, merkitään sitä kirjaimella n . Tällöin kokoa 32 olevien vaatteiden määrä on $0,002n$ ja kokoa 34 olevien vaatteiden määrä on $0,032n$ jne.

Lasketaan vaatekoon keskiarvo.

$$\begin{aligned}
 \bar{x} &= \frac{0,002n \cdot 32 + 0,032n \cdot 34 + \dots + 0,007n \cdot 56}{n} \\
 &= \frac{41,918n}{n} \\
 &= \frac{41,918n^{(n)}}{n} \\
 &= 41,918 \\
 &\approx 42
 \end{aligned}$$

Kokojen keskiarvo on noin 42.

Vastaus: $Mo = 40$, $Q_1 = 38$, $Q_2 = 42$, $Q_3 = 46$, $\bar{x} \approx 42$

- b) Koska kokonaistilausmäärä on noin 100, on kutakin kokoa tilattava koon suhteellisen frekvenssin verran.

Pyöristetään saadut desimaaliluvut ylöspäin seuraavaan kokonaislukuun, jotta jokaista kokoa on jakauman mukaan riittävästi.

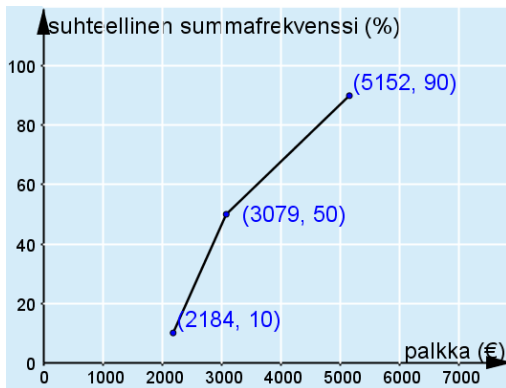
	A	B	C	D
1	Kokonro	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi	Pyöristys seuraavaan kokonaislukuun
2	32	0,20	=B2	=PYÖRISTÄ.DES.YLÖS(B2)
3	34	3,20	=C2+B3	=PYÖRISTÄ.DES.YLÖS(B3)
4	36	8,40	=C3+B4	=PYÖRISTÄ.DES.YLÖS(B4)
5	38	17,80	=C4+B5	=PYÖRISTÄ.DES.YLÖS(B5)
6	40	20,00	=C5+B6	=PYÖRISTÄ.DES.YLÖS(B6)
7	42	13,70	=C6+B7	=PYÖRISTÄ.DES.YLÖS(B7)
8	44	11,10	=C7+B8	=PYÖRISTÄ.DES.YLÖS(B8)
9	46	10,60	=C8+B9	=PYÖRISTÄ.DES.YLÖS(B9)
10	48	7,60	=C9+B10	=PYÖRISTÄ.DES.YLÖS(B10)
11	50	4,10	=C10+B11	=PYÖRISTÄ.DES.YLÖS(B11)
12	52	1,50	=C11+B12	=PYÖRISTÄ.DES.YLÖS(B12)
13	54	1,10	=C12+B13	=PYÖRISTÄ.DES.YLÖS(B13)
14	56	0,70	=C13+B14	=PYÖRISTÄ.DES.YLÖS(B14)

	A	B	C	D
1	Kokonro	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi	Pyöristys seuraavaan kokonaislukuun
2	32	0,20	0,20	1
3	34	3,20	3,4	4
4	36	8,40	11,8	9
5	38	17,80	29,6	18
6	40	20,00	49,6	20
7	42	13,70	63,3	14
8	44	11,10	74,4	12
9	46	10,60	85	11
10	48	7,60	92,6	8
11	50	4,10	96,7	5
12	52	1,50	98,2	2
13	54	1,10	99,3	2
14	56	0,70	100	1

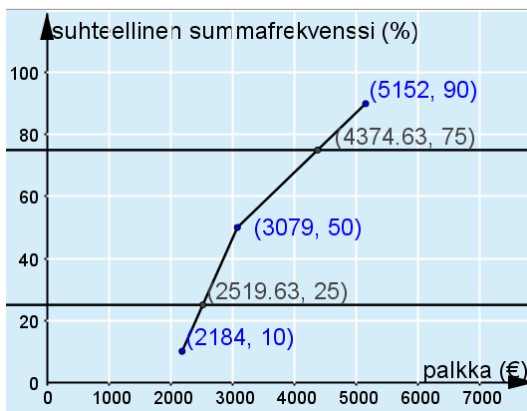
Vastaus: esim.

Koko	32	34	36	38	40	42	44	46	48	50	52	54	56
Tilattava määrä	1	4	9	18	20	14	12	11	8	5	2	2	1

452. Suhteellisen summafrekvenssin kuvaaja kulkee pisteiden (2184, 10), (3079, 50) ja (5152, 90) kautta. Piirretään niiden kautta kertymän kuvaaja.



Määritetään ala- ja yläkvartiili piirtämällä suorat $y = 25$ ja $y = 75$ ja määrittämällä niiden ja kertymän kuvaajan leikkauspisteet.



Leikkauspisteet ovat (2519,63; 25) ja (4374,63; 75), joten alakvartiili on $2519,63 \approx 2520$ euroa ja yläkvartiili $4374,63 \approx 4375$ euroa.

Vastaus: $Q_1 = 2520$ €, $Q_3 = 4375$ €

- 453. a)** Jos aineiston jokaisesta havaintoarvosta vähennetään yksi, havaintoarvojen summa pienenee havaintojen kokonaislukumäärän verran, joten keskiarvokin pienenee yhdellä. Väite on siis tosi.

Vastaus: tosi

- b)** Jos vain suurin havaintoarvo kasvaa muiden pysyessä samoina, havaintoarvojen suuruusjärjestys ei muutu. Tällöin mediaanikaan ei muutu. Väite on epätosi.

Vastaus: epätosi, pysyy samana

- c)** Alakvartiilin alapuolella on 25 % havainnoista, joten sen yläpuolella on 75 % havainnoista, joten väite on tosi.

Vastaus: tosi

- 454.** Kaikkien viiden luvun summa on $2 \cdot 42 + 3 \cdot 54 = 246$.

Kaikkien viiden luvun keskiarvo on $\frac{246}{5} = 49,2$.

Vastaus: 49,2

- 455. a)** Kaikkien yhdentoista henkilön pituuksien summa on $10 \cdot 172 \text{ cm} + 182 \text{ cm} = 1902 \text{ cm}$. Keskiarvo on $\frac{1902 \text{ cm}}{11} = 172,909... \text{ cm} \approx 173 \text{ cm}$.

Vastaus: 173 cm

- b)** Kaikkien 21 henkilön pituuksien summa on $20 \cdot 172 \text{ cm} + 182 \text{ cm} = 3622 \text{ cm}$. Keskiarvo on $\frac{3622 \text{ cm}}{21} = 172,476... \text{ cm} \approx 172 \text{ cm}$.

Vastaus: 172 cm

- c) Merkitään henkilön pituutta senttimetreinä kirjaimella x . Muodostetaan yhtälö ja ratkaistaan se.

$$\frac{20 \cdot 172 + x}{21} = 173 \quad || \cdot 21$$

$$3440 + x = 3633$$

$$x = 3633 - 3440$$

$$x = 193$$

Henkilön pitäisi olla 193 cm pitkä.

Vastaus: 193 cm

456. a) Luokkakeskus on luokan todellisen ala- ja todellisen ylärajan keskiarvo. Täydennetään taulukkoon todelliset rajat ja luokkakeskukset taulukkolaskentaohjelmalla. Ensimmäisessä kuvassa näkyy kaavat, joilla toisen kuvan luvut on saatu.

	A	B	C	D	E
1	Odotusaika (min)	Frekvenssi	Todellinen alaraja (min)	Todellinen yläraja (min)	Luokkakeskus (min)
2	0–4	3	0	4,5	$=(C2+D2)/2$
3	5–9	12	4,5	9,5	$=(C3+D3)/2$
4	10–14	40	9,5	14,5	$=(C4+D4)/2$
5	15–19	31	14,5	19,5	$=(C5+D5)/2$
6	20–24	10	19,5	24,5	$=(C6+D6)/2$
7	25–29	2	24,5	29,5	$=(C7+D7)/2$

	A	B	C	D	E
1	Odotusaika (min)	Frekvenssi	Todellinen alaraja (min)	Todellinen yläraja (min)	Luokkakeskus (min)
2	0–4	3	0	4,5	2,25
3	5–9	12	4,5	9,5	7
4	10–14	40	9,5	14,5	12
5	15–19	31	14,5	19,5	17
6	20–24	10	19,5	24,5	22
7	25–29	2	24,5	29,5	27

Keskiarvo on

$$\bar{x} = \frac{3 \cdot 2,25 + 12 \cdot 7,0 + 40 \cdot 12,0 + 31 \cdot 17,0 + 10 \cdot 22,0 + 2 \cdot 27,0}{3 + 12 + 40 + 31 + 10 + 2}$$

$$= \frac{1371,75}{98} = 13,997... \approx 14,0.$$

Vastaus:

Odotusaika (min)	Luokkakeskus (min)
0–4	2,25
5–9	7,0
10–14	12
15–19	17,0
20–24	22
25–29	27

$$\bar{x} \approx 14,0 \text{ min}$$

b) Luokitellaan odotusajat 10 minuutin pituisiin luokkiin.

Aika (min)	Todellinen alaraja (min)	Todellinen yläraja (min)	Luokkakeskus (min)	Frekvenssi
0–9	0	9,5	$\frac{0+9,5}{2} = 4,75$	$3 + 12 = 15$
10–19	9,5	19,5	$\frac{9,5+19,5}{2} = 14,5$	$40 + 31 = 71$
20–29	19,5	29,5	$\frac{19,5+29,5}{2} = 24,5$	$10 + 2 = 12$

Keskiarvo on

$$\bar{x} = \frac{15 \cdot 4,75 + 71 \cdot 14,5 + 12 \cdot 24,5}{98} = 14,232... \approx 14,2.$$

a- ja b-kohdan tulokset ovat likimain yhtä suuria.

Vastaus:

Odotusaika (min)	Frekvenssi
0–9	15
10–19	71
20–29	12

$\bar{x} \approx 14,2$ min. Tulokset ovat likimain yhtä suuria.

457. Merkitään ryhmän B opiskelijoiden keskiarvoa kirjaimella x . Opiskelijoita oli yhteensä $27 + 34 = 61$. Muodostetaan yhtälö ja ratkaistaan siitä x .

$$\frac{27 \cdot 7,25 + 34x}{61} = 7,14$$

$$\frac{195,75 + 34x}{61} = 7,14 \quad \parallel \cdot 61$$

$$195,75 + 34x = 435,54$$

$$34x = 435,54 - 195,75$$

$$34x = 239,79 \quad \parallel : 34$$

$$x = 7,052...$$

Ryhmän B opiskelijoiden keskiarvo oli noin 7,05.

Vastaus: 7,05

458.

Veronalaiset tulot (€)	Tulonsaajat (tuhatta)	Valtionveronalaiset tulot (Milj.€)
– 9999	871	4446
10 000– 19 999	1046	15592
20 000–29 999	929	23207
30 000–39 999	782	27093
40 000–49 999	455	20249
50 000–59 999	247	13483
60 000–69 999	140	9046
70 000–79 999	81	6019
80 000–89 999	50	4210
90 000–99 999	31	2980
100 000–149 999	61	7233
150 000–	31	10101
Yhteensä	4724	143659

Huomataan, että moodiluokka on 10 000–19 999 €.

Tulojen keskiarvo saadaan jakamalla viimeisen sarakkeen summa 143 659 000 000 € tulonsaajien kokonaismäärällä 4 724 000.

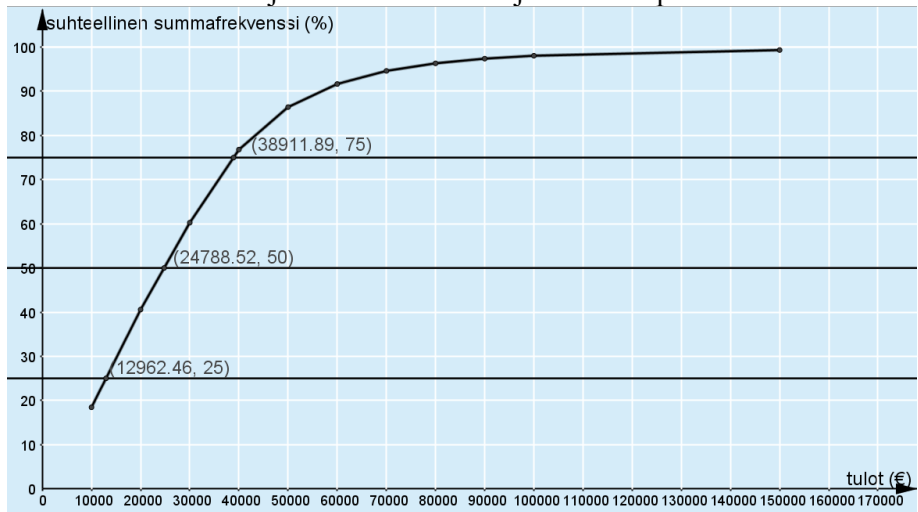
Keskiarvo on $\bar{x} = \frac{143\,659\,000\,000\text{ €}}{4\,724\,000} = 30\,410,457\dots\text{€} \approx 30\,400\text{ €}$.

Kvartiilien määrittämiseksi lasketaan jakauman todelliset ylärajat ja suhteelliset summafrekvenssit. Ylimmällä tuloluokalla ei ole ylärajaa, joten kyseistä riviä ei voida täyttää. Tällä ei kuitenkaan ole merkitystä kvartiilien kannalta.

	A	B	C	D	E
1	Veronalaiset tulot (€)	Todellinen yläraja (€)	Frekvenssi	Suhteellinen frekvenssi	Suhteellinen summafrekvenssi
2	– 9999	9999,5	871	18,44 %	18,44 %
3	10 000– 19 999	19999,5	1046	22,14 %	40,58 %
4	20 000–29 999	29999,5	929	19,67 %	60,25 %
5	30 000–39 999	39999,5	782	16,55 %	76,80 %
6	40 000–49 999	49999,5	455	9,63 %	86,43 %
7	50 000–59 999	59999,5	247	5,23 %	91,66 %
8	60 000–69 999	69999,5	140	2,96 %	94,62 %
9	70 000–79 999	79999,5	81	1,71 %	96,34 %
10	80 000–89 999	89999,5	50	1,06 %	97,40 %
11	90 000–99 999	99999,5	31	0,66 %	98,05 %
12	100 000–149 999	149999,5	61	1,29 %	99,34 %
13	150 000–		31	0,66 %	100,00 %
14	Yhteensä		4724	100,00 %	
15					

Kvartiilien määrittämiseksi piirretään kertymää kuvaava viivakuvio. Viivakuvion kuvaajan pisteiden x -koordinaatti on luokan todellinen yläraja ja y -koordinaatti luokan suhteellinen frekvenssi.

Selvitetään kvartiilit piirtämällä suorat $y = 25$, $y = 50$ ja $y = 75$ ja määrittämällä suorien ja viivakuvion kuvaajan leikkauspisteet.



Alakvartiili on noin 13 000 euroa, mediaani noin 24 800 euroa ja yläkvartiili noin 38 900 euroa.

Vastaus: $M_0 = 10\,000 - 19\,999\text{ €}$, $\bar{x} = 30\,400\text{ €}$, $Q_1 = 13\,000\text{ €}$,
 $Q_2 = 24\,800\text{ €}$ ja $Q_3 = 38\,900\text{ €}$

- b) Keskiarvo on noin $30\,400\text{ €} - 24\,800\text{ €} = 5\,600\text{ €}$ suurempi kuin mediaani. Keskiarvo- ja mediaanitulon välinen ero johtuu väestön tuloeroista. Suuret poikkeamat vaikuttavat keskiarvoon enemmän kuin mediaaniin. Jos pienellä joukolla ihmisiä on todella suuret tulot, heidän tuloensa nostavat keskiarvoa selvästi, mutta eivät juuri vaikuta mediaaniin, koska hyvin suurituloisia ihmisiä on vain vähän.

Vastaus: Keskiarvo on $5\,600\text{ €}$ suurempi kuin mediaani.

459. Keskiarvo on satunnaisista kovista pakkasista huolimatta noin -5 °C . 25 asteen pakkaset poikkeavat keskiarvosta noin 20 asteen verran alaspäin, mutta yhtä suuria poikkeamia keskiarvosta ylöspäin ei ole. Koska suuria poikkeamia keskiarvosta ylöspäin ei ole, täytyy sellaisia päiviä, joina lämpötilan keskiarvo ylittyy, olla enemmän kuin päiviä, joina keskiarvo alittuu. Lämpötilan mediaanin on siten oltava korkeampi kuin keskiarvo.

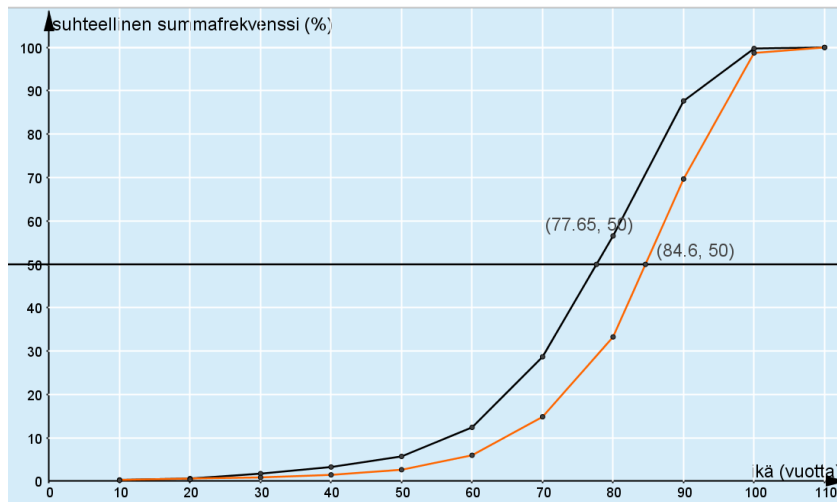
Vastaus: mediaani

Miesten kuoliniän keskiarvo on noin 75,34 vuotta ja naisten noin 82,2 vuotta.

Mediaanin määrittämiseksi tehdään sarakkeet suhteellisille summafrekvensseille (Suhteellinen summafrekvenssimiehet ja Suhteellinen summafrekvenssinaiset).

Piirretään naisten ja miesten suhteellisista summafrekvensseistä kertymän kuvaaja. Mediaani määritetään kertymän kuvaajasta. Kuvaajan pisteiden x -koordinaatit ovat luokkien ylärajoja ja y -koordinaatit suhteellisia summafrekvenssejä.

Mediaani määritetään piirtämällä suora $y = 50$.



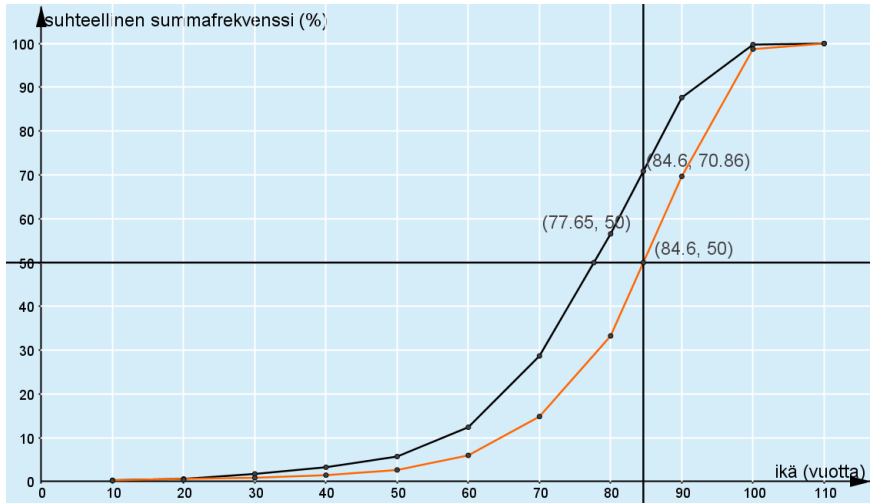
Naisten kuolinikien mediaani on noin 84,6 vuotta.

Miesten kuolinikien mediaani on noin 77,7 vuotta.

Vastaus: miehet: $\bar{x} = 75,34$ vuotta ja $Md = 77,7$ vuotta,

naiset: $\bar{x} = 82,3$ vuotta ja $Md = 84,6$ vuotta

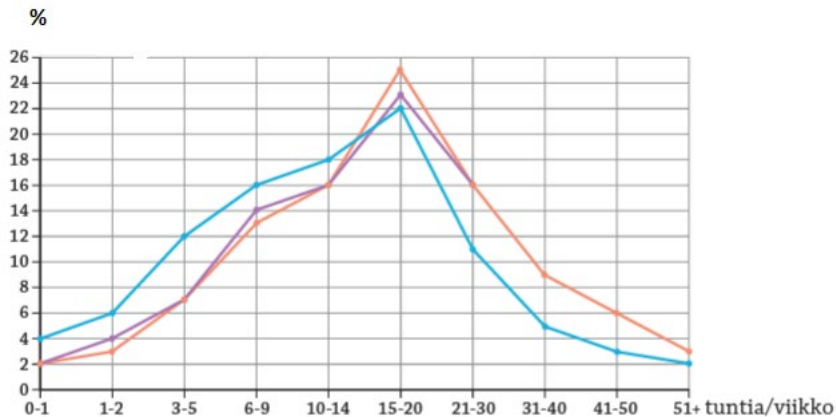
- b) Naisten kuoliniän mediaani oli 84,6 vuotta. On selvítettävä, kuinka monta prosenttia miehistä kuoli yli 84,6-vuotiaana. Piirretään x -akselille normaali pisteen $(84,6; 0,5)$ kautta ja määritetään normaalin ja miesten kuoliniän kertymän kuvaajan leikkauspiste.



Miehistä noin $70,86 \% \approx 71 \%$ kuoli naisten kuoliniän mediaania nuorempina, joten $100 \% - 71 \% = 29 \%$ kuoli kyseistä ikää vanhempina.

Vastaus: 29 %

462. Muodostetaan kuvan perusteella frekvenssitaulukko 13–17-vuotiaiden sosiaalisen median palveluissa viettämästä ajasta.



Lähde: Ebrand Group Oy & Oulun kaupunki (2019)

Aika (h)	Suhteellinen frekvenssi
0–1	2
1–2	4
3–5	7
6–9	14
10–14	16
15–20	25
21–30	16
31–40	9
41–50	6
51–70	3

Keskiarvon laskemista varten täydennetään taulukkoon todelliset rajat ja luokkakeskukset.

Aika (h)	Todellinen alaraja (h)	Todellinen yläraja (h)	Luokkakeskus (h)	Suhteellinen frekvenssi
0–1	0	1,5	$\frac{0+1,5}{2} = 0,75$	2
1–2	1,5	2,5	1,5	4
3–5	2,5	5,5	4	7
6–9	5,5	9,5	7,5	14
10–14	9,5	14,5	12	16
15–20	14,5	19,5	17,5	25
21–30	20,5	30,5	25,5	16
31–40	30,5	40,5	35,5	9
41–50	40,5	50,5	45,5	6
51–70	50,5	70,5	60,5	3

Koska nuorten kokonaismäärää ei ole kerrottu, merkitään sitä kirjaimella n . Tällöin luokan 0–1 h frekvenssi on $0,01n$ jne. Lasketaan käytetyn ajan keskiarvo.

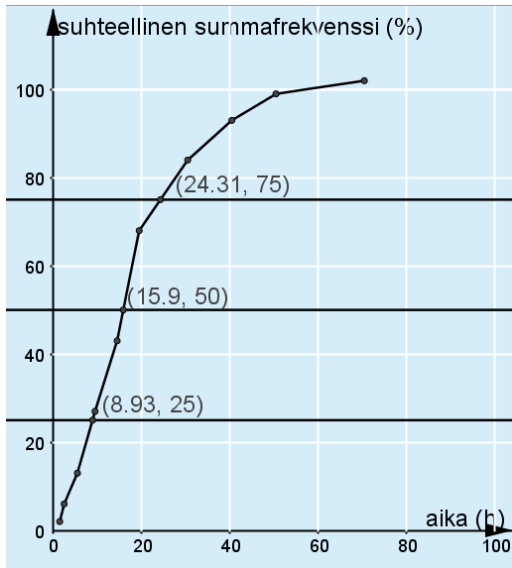
$$\begin{aligned}\bar{x} &= \frac{0,02n \cdot 0,75 + 0,04n \cdot 1,5 + \dots + 0,03n \cdot 60,5}{n} \\ &= \frac{19,52n}{n} \\ &= 19,52\end{aligned}$$

Keskiarvoksi saadaan $19,52 \text{ h} = 19 \text{ h} + 0,52 \cdot 60 \text{ min} \approx 19 \text{ h } 31 \text{ min}$.

Kvartiilien määrittämiseksi lasketaan jakauman suhteelliset summafrekvenssit.

Aika (h)	Todellinen yläraja (h)	Suhteellinen summafrekvenssi
0–1	1,5	2
1–2	2,5	6
3–5	5,5	13
6–9	9,5	27
10–14	14,5	43
15–20	19,5	68
21–30	30,5	84
31–40	40,5	93
41–50	50,5	99
51–70	70,5	102

Kvartiilien määrittämiseksi piirretään kertymää kuvaava viivakuvio. Viivakuvion kuvaajan pisteiden x -koordinaatti on luokan todellinen yläraja ja y -koordinaatti luokan suhteellinen frekvenssi. Määritetään kvartiilit piirtämällä suorat $y = 0,25$, $y = 0,5$ ja $y = 0,75$ ja määrittämällä niiden ja viivakuvion kuvaajan leikkauspisteet.



Alakvartiili on 8,93 h, keskikvartiili eli mediaani 15,9 h ja yläkvartiili 24,3 h. Tunneiksi ja minuuteiksi muutettuina nämä ovat noin 8 h 56 min, 15 h 54 min ja 24 h 18 min.

Vastaus: $\bar{x} = 19$ h 31 min, $Q_1 = 8$ h 56 min, $Q_2 = 15$ h 54 min, $Q_3 = 24$ h 18 min

- 463.** Merkitään tuntematonta frekvenssiä x :llä ja muodostetaan keskiarvon avulla yhtälö. Ratkaistaan yhtälöstä x .

$$\begin{aligned}\frac{4 \cdot 0 + 9 \cdot 1 + 12 \cdot 2 + 11 \cdot 3 + x \cdot 4 + 5 \cdot 5 + 3 \cdot 6 + 1 \cdot 7}{4 + 9 + 12 + 11 + x + 5 + 3 + 1} &= 2,72 \\ \frac{4x + 116}{45 + x} &= 2,72 && \| \cdot (45 + x) \\ 4x + 116 &= 2,72 \cdot 45 + 2,72x \\ 4x + 116 &= 122,4 + 2,72x \\ 4x - 2,72x &= 122,4 - 116 \\ 1,28x &= 6,4 && \| : 1,28 \\ x &= 5\end{aligned}$$

Tuntematon frekvenssi on 5.

Vastaus: 5

- 464.** Merkitään oikeaa keskiarvoa kirjaimella x ja oppilaiden määrää kirjaimella n . Arvosanojen oikea summa on tällöin nx .

16 % oppilaista on $0,16n$. Koska $0,16n$ oppilasta sai yhden numeron liian pienen arvosanan, jäi summa $0,16n$ liian pieneksi.

6 % oppilaista on $0,06n$. Koska $0,06n$ oppilasta sai kaksi numeroa liian pienen arvosanan, jäi summa $0,06n \cdot 2 = 0,12n$ liian pieneksi.

Merkittyjen arvosanojen summa on $nx - 0,16n - 0,12n = nx - 0,28n$.

Keskiarvoksi tuli $\frac{nx - 0,28n}{n} = x - 0,28$. Keskiarvo jäi siis $0,28$ numeroa liian pieneksi.

Vastaus: $0,28$ numeroa alaspäin

465. a) Lasketaan kunkin moduulin laajuuden ja arvosanojen tulot yhteen
 $2 \cdot 6 + 2 \cdot 8 + 2 \cdot 8 + 2 \cdot 6 + 2 \cdot 7 + 1 \cdot 7 + 1 \cdot 8 = 85$.

Oppimäärän laajuus on yhteensä $2 + 2 + 2 + 2 + 2 + 1 + 1 = 12$.

Painotettu keskiarvo on

$$\bar{x} = \frac{85}{12} = 7,083... \approx 7.$$

Tuukan matematiikan oppimäärän arvosana on 7.

Vastaus: 7

- b) Jotta päättöarvosana nousisi yhdellä numerolla eli numeroksi 8, on painotetun keskiarvon oltava vähintään 7,5. Tutkitaan taulukkolaskentaohjelmalla, mikä arvosanojen painotetuksi keskiarvoksi tulee kaikilla moduulien MAB8 ja MAB9 arvosanojen yhdistelmillä.

Täytetään taulukkoon jo olemassa olevat arvosanat ja asetetaan moduulien MAB8 ja MAB9 arvosanoiksi 7–10 ja lasketaan arvosanojen painotettu keskiarvo.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Moduuli	MAY1	MAB2	MAB3	MAB4	MAB5	MAB6	MAB7	MAB8	MAB9	Painotettu keskiarvo			
2	Opintopisteitä	2	2	2	2	2	2	1	1	2	2			
3	Arvosana	6	8	8	6	7	7	8	7	7	7,0625	$= (2 \cdot B3 + 2 \cdot C3 + 2 \cdot D3 + 2 \cdot E3 + 2 \cdot F3 + 1 \cdot G3 + 1 \cdot H3 + 2 \cdot I3 + 2 \cdot J3) / 16$		
4		6	8	8	6	7	7	8	7	8	7,1875			
5		6	8	8	6	7	7	8	7	9	7,3125			
6		6	8	8	6	7	7	8	7	10	7,4375			
7		6	8	8	6	7	7	8	8	8	7,3125			
8		6	8	8	6	7	7	8	8	9	7,4375			
9		6	8	8	6	7	7	8	8	10	7,5625			
10		6	8	8	6	7	7	8	9	9	7,5625			
11		6	8	8	6	7	7	8	9	10	7,6875			
12		6	8	8	6	7	7	8	10	8	7,5625			
13		6	8	8	6	7	7	8	10	9	7,6875			
14		6	8	8	6	7	7	8	10	10	7,8125			

Muodostetaan taulukko kaikista arvosanayhdistelmistä, joilla Tuukan päättöarvosana nousee yhdellä numerolla.

MAB8 arvosana	8	10	9	10	10
MAB9 arvosana	10	8	10	9	10

Vastaus:

MAB8 arvosana	8	10	9	10	10
MAB9 arvosana	10	8	10	9	10