

5.1

Ratkaistaan tehtävä LibreOfficen Calc-ohjelmalla.

Kirjoitetaan sarakkeeseen A muuttujan arvot suuruusjärjestyksessä ja sarakkeeseen B niiden suhteelliset frekvenssit. Lasketaan sarakkeeseen C suhteelliset summafrekvenssit.

	A	B	C
1	Yöunen pituus (h)	$f\%$	$sf\%$
2	1-2	2	'=B2
3	3-4	8	'=C2+B3
4	5-6	19	'=C3+B4
5	7-8	48	'=C4+B5
6	9-10	19	'=C5+B6
7	11-12	4	'=C6+B7

Kirjoita kaava soluun C3 ja kopioi sitä alaspäin.

	A	B	C
1	Yöunen pituus (h)	$f\%$	$sf\%$
2	1-2	2	2
3	3-4	8	10
4	5-6	19	29
5	7-8	48	77
6	9-10	19	96
7	11-12	4	100

Suhteellinen summafrekvenssi ylittää 50 % rajan luokassa 7–8, joten yöunen mediaaniluokka on 7–8.

Vastaus

mediaaniluokka 7–8

5.2

Ratkaistaan tehtävä LibreOfficen Calc-ohjelmalla.

Kirjoitetaan sarakkeeseen A muuttujan arvot suuruusjärjestyksessä ja sarakkeeseen B niiden suhteelliset frekvenssit. Lasketaan sarakkeeseen C suhteelliset summafrekvenssit.

	A	B	C
1	Aika (h)	<i>f</i> %	<i>sf</i> %
2	0-2	1	'=B2
3	3-5	2	'=C2+B3
4	6-9	3	'=C3+B4
5	10-14	7	'=C4+B5
6	15-20	15	'=C5+B6
7	21-30	21	'=C6+B7
8	31-40	21	'=C7+B8
9	41-50	20	'=C8+B9
10	51-	10	'=C9+B10

Kirjoita kaava soluun C3 ja kopioi sitä alaspäin.

	A	B	C
1	Aika (h)	<i>f</i> %	<i>sf</i> %
2	0-2	1	1
3	3-5	2	3
4	6-9	3	6
5	10-14	7	13
6	15-20	15	28
7	21-30	21	49
8	31-40	21	70
9	41-50	20	90
10	51-	10	100

Suhteellinen summafrekvenssi ylittää 50 % rajan luokassa 31–40, joten Internetissä vietetyn ajan mediaaniluokka on 31–40.

Vastaus

mediaaniluokka 31–40

5.3

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelmalla.

1. Syötä sarakkeeseen A muuttujan luokat, sarakkeeseen B suhteelliset frekvenssit ja sarakkeeseen C todelliset luokkarajat. Laske sarakkeeseen D suhteelliset summafrekvenssit.

	A	B	C	D
1	Pituus (cm)	$f\%$	Todellinen yläraja	$sf\%$
2	-40	0	40,5	0
3	41-43	2	43,5	'=D2+B3
4	44-46	14	46,5	'=D3+B4
5	47-49	33	49,5	'=D4+B5
6	50-52	34	52,5	'=D5+B6
7	53-55	14	55,5	'=D6+B7
8	56-58	3	58,5	'=D7+B8

Suhteellisen summafrekvenssin ensimmäisen arvon tulee olla 0. Lisätään sen takia ensimmäiseksi luokaksi alkuperäisen taulukon arvojen alle jäävät pituudet.

Kirjoita kaava soluun D3 ja kopioi sitä alaspäin.

	A	B	C	D
1	Pituus (cm)	$f\%$	Todellinen yläraja	$sf\%$
2	-40	0	40,5	0
3	41-43	2	43,5	2
4	44-46	14	46,5	16
5	47-49	33	49,5	49
6	50-52	34	52,5	83
7	53-55	14	55,5	97
8	56-58	3	58,5	100

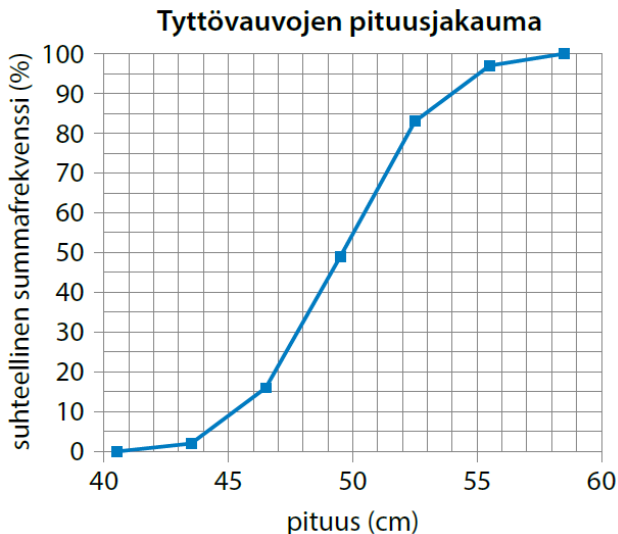
2. Piirrä kertymäkuvaaja **Ohjattu kaavion luonti** -toimintoa käyttäen.

- Maalaa solut C1–D8.
- Valitse **Lisää**-valikosta **Kaavio**.
- Vaiheessa **1. Kaaviotyyppi** valitse kaaviotyyppiä **XY (hajonta)** ja **Pisteet ja viivat**.
- Vaiheessa **2. Tietoaalue** valitse **Arvosarjat sarakkeissa**, **Ensimmäinen rivi sisältää otsikoita** ja **Ensimmäinen sarake sisältää otsikoita**.
- Vaiheessa **3. Arvosarja** älä tee muutoksia.
- Vaiheessa **4. Kaavioelementit** syötä otsikoksi ”Tyttövauvojen pituusjakauma”, X-akselin otsikoksi ”pituus (cm)”, Y-akselin otsikoksi ”suhteellinen summafrekvenssi (%)” ja poista valinta kohdasta Näytä selite. Valitse Näytä ruudukot -kohdasta X-akseli ja Y-akseli.
- Valitse Valmis.

Ohjelma piirtää kertymäkuvaajan, mutta sitä täytyy vielä muokata. Mitä tiheämpi asteikko on, sitä tarkemmin arvoja voidaan lukea.

Kertymäkuvaajassa pystyakselin asteikon tulee olla välillä 0–100.

- Aktivoi piirtoalue kaksoisnapauttamalla tilastokuviota.
- Napauta hiiren oikealla painikkeella muokattavaa akselia ja valitse **Muotoile akseli**. Vaihda Y-akselin asteikko välille 0–100. Voit tihentää X-akselin asteikkoa.



- b)** Arvioidaan tyttövauvojen syntymäpituuksien mediaani kertymäkuvaajasta. Kuvaaja ylittää 50 % rajan noin pituuden 49,6 cm kohdalla. Tyttövauvojen mediaanipituus on 49,6 cm.
- c)** Arvioidaan pituus, jota pidempiä on 20 % tyttövauvoista. Kuvaaja ylittää 80 % rajan noin 52,2 cm kohdalla. Pisimmät 20 % tyttövauvoista ovat vähintään 52,2 cm pitkiä.
- d)** Arvioidaan pituus, jota lyhyempiä on 25 % tyttövauvoista. Kuvaaja ylittää 25 % rajan noin 47,3 cm kohdalla. Lyhyimmät 25 % tyttövauvoista ovat korkeintaan 47,3 cm pitkiä.

Vastaus

- b)** 49,6 cm
c) 52,2 cm
d) 47,3 cm

5.4

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelmalla.

1. Syötä sarakkeeseen A muuttujan luokat, sarakkeeseen B suhteelliset frekvenssit ja sarakkeeseen C todelliset luokkarajat. Laske sarakkeeseen D suhteelliset summafrekvenssit.

	A	B	C	D
1	Aika (h)	$f\%$	Todellinen yläraja	$sf\%$
2			0	0
3	0-4	2	4,5	'=D2+B3
4	5-10	4	10,5	'=D3+B4
5	11-14	9	14,5	'=D4+B5
6	15-20	16	20,5	'=D5+B6
7	21-30	21	30,5	'=D6+B7
8	31-40	22	40,5	'=D7+B8
9	41-50	18	50,5	'=D8+B9
10	51-60	8	60,5	'=D9+B10

Suhteellisen summafrekvenssin ensimmäisen arvon tulee olla 0.

Kirjoita kaava soluun D3 ja kopioi sitä alaspäin.

	A	B	C	D
1	Aika (h)	$f\%$	Todellinen yläraja	$sf\%$
2			0	0
3	0-4	2	4,5	2
4	5-10	4	10,5	6
5	11-14	9	14,5	15
6	15-20	16	20,5	31
7	21-30	21	30,5	52
8	31-40	22	40,5	74
9	41-50	18	50,5	92
10	51-60	8	60,5	100

2. Piirrä kertymäkuvaaja **Ohjattu kaavion luonti** -toimintoa käyttäen.

- Maalaa solut C1–D10.

- Valitse **Lisää**-valikosta **Kaavio**.

- Vaiheessa **1. Kaaviotyyppi** valitse kaaviotyyppiä **XY (hajonta)** ja **Pisteet ja viivat**.

- Vaiheessa **2. Tietoa** valitse **Arvosarjat** sarakkeissa,

Ensimmäinen rivi sisältää otsikoita ja Ensimmäinen sarake sisältää otsikoita.

- Vaiheessa **3. Arvosarja** älä tee muutoksia.

- Vaiheessa **4. Kaavioelementit** syötä otsikoksi ”Opiskelijoiden Internetissä viettämä aika”, X-akselin otsikoksi ”aika (h)”, Y-akselin otsikoksi ”suhteellinen summafrekvenssi (%)” ja poista valinta kohdasta Näytä selite. Valitse Näytä ruudukot -kohdasta X-akseli ja Y-akseli.

- Valitse Valmis.

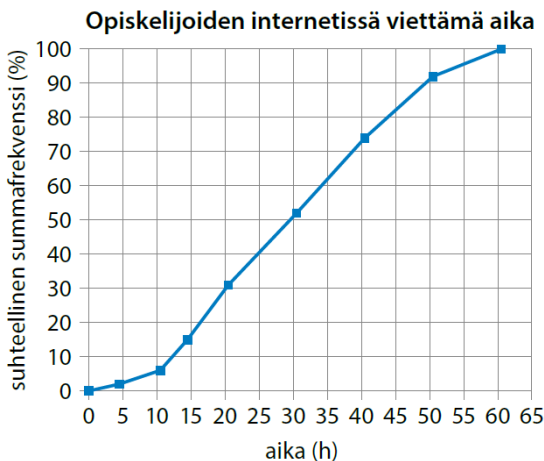
Ohjelma piirtää kertymäkuvaajan, mutta sitä täytyy vielä muokata. Mitä tiheämpi asteikko on, sitä tarkemmin arvoja voidaan lukea.

Kertymäkuvaajassa pystyakselin asteikon tulee olla välillä 0–100.

- Aktivoi piirtoalue kaksoisnapauttamalla tilastokuviota.

- Napauta hiiren oikealla painikkeella muokattavaa akselia ja valitse

Muotoile akseli. Vaihda Y-akselin asteikko välille 0–100. Voit tihentää X-akselin asteikkoa.



- b)** Arvioidaan Internetissä vietetyn ajan mediaani kertymäkuvaajasta. Kuvaaja ylittää 50 % rajan noin ajan 30 h kohdalla. Internetissä vietetyn ajan mediaani on 30 h.
- c)** Arvioidaan aika, jota enemmän 25 % opiskelijoista vietti Internetissä. Kuvaaja ylittää 75 % rajan ajan 41 h kohdalla. Opiskelijoista 25 % vietti Internetissä yli 41 h.

Vastaus

- b)** 30 h
- c)** 41 h

5.5

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelmalla.

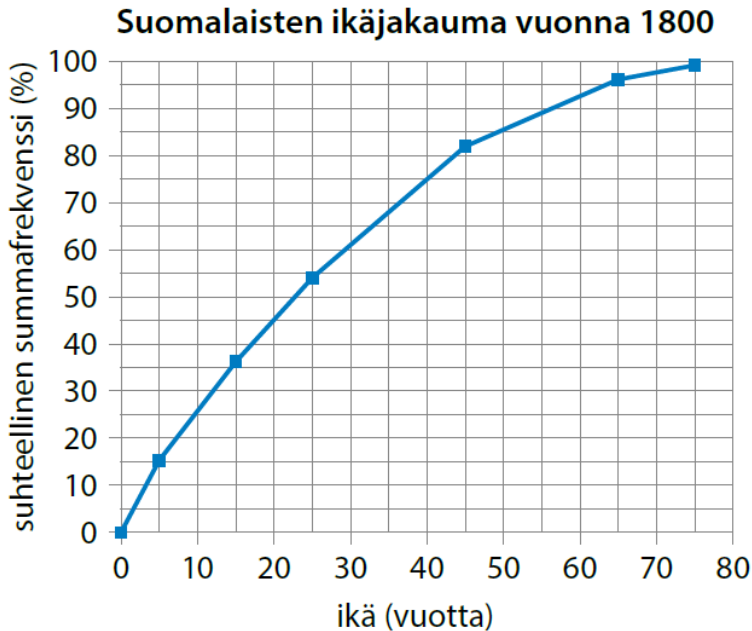
Syötetään tarvittavat tiedot sarakkeisiin A–C ja lasketaan suhteelliset summafrekvenssit sarakkeeseen D.

	A	B	C	D
1	Ikä (vuotta)	$f\%$	Todellinen yläraja	$sf\%$
2			0	0
3	0-4	15,3	5	'=D2+B3
4	5-14	20,9	15	'=D3+B4
5	15-24	18,1	25	'=D4+B5
6	25-44	27,6	45	'=D5+B6
7	45-64	14,2	65	'=D6+B7
8	65-74	3,0	75	'=D7+B8
9	75-	0,9		'=D8+B9

Viimeiselle luokalle ei ole annettu ylärajaa, joten se jää avoimeksi.

	A	B	C	D
1	Ikä (vuotta)	$f\%$	Todellinen yläraja	$sf\%$
2			0	0
3	0-4	15,3	5	15,3
4	5-14	20,9	15	36,2
5	15-24	18,1	25	54,3
6	25-44	27,6	45	81,9
7	45-64	14,2	65	96,1
8	65-74	3,0	75	99,1
9	75-	0,9		100

Piirretään kertymäkuvaaja Ohjattu kaavion luonti -toimintoa käyttäen.



- b) Arvioidaan suomalaisten mediaani-ikä kertymäkuvaajasta. Kuvaaja ylittää 50 % rajan noin 23 vuoden kohdalla.
- c) Arvioidaan ikä, jota nuorempia oli 25 % suomalaisista. Kuvaaja ylittää 25 % rajan noin 10 vuoden kohdalla. Nuorin neljännes on siis alle 10 vuotta.
- d) Arvioidaan ikä, jota vanhempia oli 25 % suomalaisista. Kuvaaja ylittää 75 % rajan noin 40 vuoden kohdalla. Vanhin neljännes on siis yli 40 vuotta.

Vastaus

- b) 23 vuotta
- c) alle 10 vuotta
- d) yli 40 vuotta

5.6

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelmalla.

Syötetään annetut tiedot sarakkeisiin A–B ja lasketaan summafrekvenssit sarakkeeseen C.

	A	B	C
1	Ikä (vuotta)	<i>f</i>	<i>sf</i>
2	-19	0	0
3	20-29	2	'=C2+B3
4	30-39	6	'=C3+B4
5	40-49	15	'=C4+B5
6	50-59	10	'=C5+B6
7	60-69	2	'=C6+B7

	A	B	C
1	Ikä (vuotta)	<i>f</i>	<i>sf</i>
2	-19	0	0
3	20-29	2	2
4	30-39	6	8
5	40-49	15	23
6	50-59	10	33
7	60-69	2	35

Lasketaan seuraavaksi suhteelliset summafrekvenssit sarakkeeseen E.

	A	B	C	D	E
1	Ikä (vuotta)	<i>f</i>	<i>sf</i>	Todellinen yläraja	<i>sf</i> %
2	-19	0	0	20	0
3	20-29	2	2	30	'=C3/\$C\$7*100
4	30-39	6	8	40	'=C4/\$C\$7*100
5	40-49	15	23	50	'=C5/\$C\$7*100
6	50-59	10	33	60	'=C6/\$C\$7*100
7	60-69	2	35	70	'=C7/\$C\$7*100

Suhteellisen summafrekvenssin saa laskettua jakamalla summafrekvenssin laulajien kokonaismäärällä (solu C7).

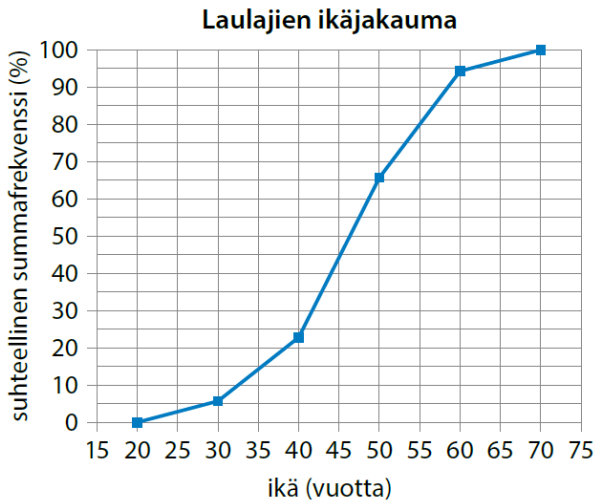
Todellinen yläraja on merkitty taulukkoon kertymäkuvaajan piirtämistä varten.

	A	B	C	D	E
1	Ikä (vuotta)	<i>f</i>	<i>sf</i>	Todellinen yläraja	<i>sf</i> %
2	-19	0	0	20	0
3	20-29	2	2	30	5,7
4	30-39	6	8	40	22,9
5	40-49	15	23	50	65,7
6	50-59	10	33	60	94,3
7	60-69	2	35	70	100,0

Taulukko summafrekvensseistä ja suhteellisista summafrekvensseistä:

Ikä (vuotta)	<i>sf</i>	<i>sf</i> %
20-29	2	5,7
30-39	8	22,9
40-49	23	65,7
50-59	33	94,3
60-69	35	100

b) Piirretään kertymäkuvaaja Ohjattu kaavion luonti -toimintoa käyttäen.



5.7

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelman avulla.

Syötetään annetut tiedot sarakkeisiin A, B ja D ja lasketaan summafrekvenssit sarakkeisiin C ja E.

	A	B	C	D	E
1	Alkoholinkäyttö	2008		2016	
2		<i>f</i> %	<i>sf</i> %	<i>f</i> %	<i>sf</i> %
3	Ei koskaan	5	'=B3	10	'=D3
4	Harvemmin kuin kerran kuukaudessa	8	'=C3+B4	14	'=E3+D4
5	Kerran pari kuukaudessa	29	'=C4+B5	33	'=E4+D5
6	Kerran viikossa	27	'=C5+B6	25	'=E5+D6
7	Pari kertaa viikossa	26	'=C6+B7	14	'=E6+D7
8	Vähintään 4 kertaa viikossa	5	'=C7+B8	4	'=E7+D8

	A	B	C	D	E
1	Alkoholinkäyttö	2008		2016	
2		<i>f</i> %	<i>sf</i> %	<i>f</i> %	<i>sf</i> %
3	Ei koskaan	5	5	10	10
4	Harvemmin kuin kerran kuukaudessa	8	13	14	24
5	Kerran pari kuukaudessa	29	42	33	57
6	Kerran viikossa	27	69	25	82
7	Pari kertaa viikossa	26	95	14	96
8	Vähintään 4 kertaa viikossa	5	100	4	100

Summafrekvenssit voi laskea myös käsin.

Vuonna 2008 ylittyi 50 % raja luokassa kerran viikossa. Vuonna 2008 alkoholin käytön mediaani oli kerran viikossa.

Vuonna 2016 ylittyi 50 % raja luokassa kerran pari kuukaudessa. Vuonna 2016 alkoholin käytön mediaani oli kerran pari kuukaudessa.

- b)** 75 % raja ylittyi vuonna 2008 luokassa pari kertaa viikossa. Alkoholia useimmin juova neljännes käytti alkoholia pari kertaa viikossa tai useammin.

Vastaavasti vuonna 2016 sama raja ylittyi luokassa kerran viikossa. Alkoholia useimmin juova neljännes käytti alkoholia kerran viikossa tai useammin.

Vastaus

- a)** 2008: kerran viikossa, 2016: kerran pari kuukaudessa
b) 2008: pari kertaa viikossa tai useammin, 2016: kerran viikossa tai useammin

5.8

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelman avulla.

Syötetään annetut tiedot sarakkeisiin A ja B ja lasketaan summafrekvenssit sarakkeeseen C.

	A	B	C
1	Arvosana	<i>f</i> %	<i>sf</i> %
2	L	5,9	'=B2
3	E	15,7	'=C2+B3
4	M	21,3	'=C3+B4
5	C	25,1	'=C4+B5
6	B	19,1	'=C5+B6
7	A	9,3	'=C6+B7
8	I	3,6	'=C7+B8

	A	B	C
1	Arvosana	<i>f</i> %	<i>sf</i> %
2	L	5,9	5,9
3	E	15,7	21,6
4	M	21,3	42,9
5	C	25,1	68,0
6	B	19,1	87,1
7	A	9,3	96,4
8	I	3,6	100,0

Summafrekvenssit voi laskea myös käsin.

50 % raja ylittyy luokassa C. Mediaaniarvosana oli siis C.

- b) Kokelaat, jotka saivat vähintään arvosanan E, saivat siis arvosanan L tai arvosanan E. Summafrekvenssi arvosanan E kohdalla on 21,6 % eli vähintään arvosanan E sai 21,6 % kokelaista.
- c) Vähintään arvosanan A sai 96,4 % kokelaista. Vähennetään tästä ne kokelaat, jotka saivat paremman arvosanan kuin M eli kokelaat, jotka saivat vähintään arvosanan E.

$$96,4\% - 21,6\% = 74,8\%$$

Kokelaista 74,8 % sai arvosanan, joka oli korkeintaan M, mutta vähintään A.

Vastaus

- a) C
b) 21,6 %
c) 74,8 %

5.9

- a) Kertymäkuvaajassa 50 % raja ylittyy noin iän 21 vuotta kohdalla.
Kotoamuuttoiaan mediaani on siis 21 vuotta.
- b) Kertymäkuvaajasta nähdään, että 20-vuotiaana kotoa on muuttanut noin 47 % suomalaisista.
- c) Kertymäkuvaajasta nähdään, että 25-vuotiaana kotoa on muuttanut noin 85 % suomalaisista.
- d) 20–24-vuotiaana kotoa muuttaneiden osuus saadaan vähentämällä alle 25-vuotiaana kotoa muuttaneiden prosenttiosuudesta alle 20-vuotiaana kotoa muuttaneiden osuus.

$$85\% - 47\% = 38\%$$

20–24-vuotiaana kotoa on muuttanut 38 % suomalaisista.

- e) Kertymäkuvaajassa 90 % raja ylittyy noin iän 27 vuotta kohdalla.
Kotoa on muuttanut 90 % suomalaisista 27 vuoden ikään mennessä.

Vastaus

- a) 21 vuotta
- b) 47 %
- c) 85 %
- d) 38 %
- e) 27 vuotta

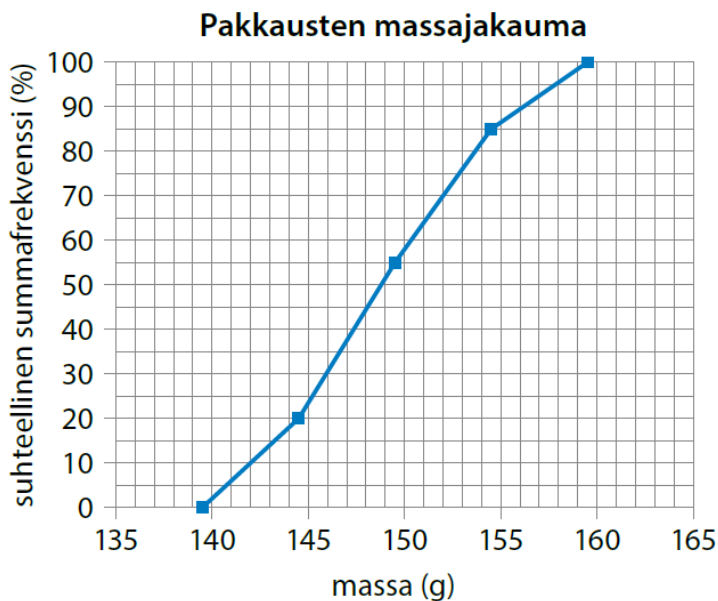
5.10

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelmalla.

Syötetään histogrammista saadut tiedot tiedot sarakkeisiin A–B ja lasketaan suhteelliset summafrekvenssit sarakkeeseen D.

	A	B	C	D
1	Massa (g)	$f\%$	Todellinen yläraja	$sf\%$
2	-139		0	0
3	140-144	20	144,5	20
4	145-149	35	149,5	55
5	150-154	30	154,5	85
6	155-159	15	159,5	100

Piirretään kertymäkuvaaja Ohjattu kaavion luonti -toimintoa käyttäen.



- b) Arvioidaan massa, jota kevyempi on 75 % pakkauksista. Kuvaaja ylittää 75 % rajan noin 153 gramman kohdalla. Pakkauksista 75 % massa on alle 153 g.
- c) Arvioidaan, kuinka monta prosenttia pakkauksista painaa yli 152 g. Kertymäkuvaajasta nähdään, että kertymä 152 g kohdalla on noin 70 % eli 70 % pakkauksista painaa alle 152 g.

$$100\% - 70\% = 30\%$$

Pakkauksista 30 % painaa yli 152 g.

Vastaus

- b) 153 g
c) 30 %

5.11

Ratkaistaan tehtävä LibreOfficen Calc-ohjelmalla.

Kirjoitetaan sarakkeeseen A muuttujan arvot suuruusjärjestyksessä ja sarakkeeseen B niiden suhteelliset frekvenssit. Lasketaan sarakkeeseen C suhteelliset summafrekvenssit.

	A	B	C
1	Pinta-ala (m²)	<i>f</i> %	<i>sf</i> %
2	35-49	13,8	'=B2
3	50-64	20,7	'=C2+B3
4	65-79	27,6	'=C3+B4
5	80-94	24,1	'=C4+B5
6	95-109	13,8	'=C5+B6

Kirjoita kaava soluun C3 ja kopioi sitä alaspäin.

	A	B	C
1	Pinta-ala (m²)	<i>f</i> %	<i>sf</i> %
2	35-49	13,8	13,8
3	50-64	20,7	34,5
4	65-79	27,6	62,1
5	80-94	24,1	86,2
6	95-109	13,8	100

Suhteellinen summafrekvenssi ylittää 50 % rajan luokassa 65–79, joten pinta-alojen mediaaniluokka on 65–79.

Vastaus

mediaaniluokka 65–79

5.12

Ratkaistaan tehtävä LibreOfficen Calc-ohjelmalla.

Kirjoitetaan sarakkeeseen A muuttujan arvot suuruusjärjestyksessä ja sarakkeeseen B niiden suhteelliset frekvenssit. Lasketaan sarakkeeseen C suhteelliset summafrekvenssit.

	A	B	C
1	Aika (h)	$f\%$	$sf\%$
2	0-2	6	'=B2
3	3-5	7	'=C2+B3
4	6-9	14	'=C3+B4
5	10-14	16	'=C4+B5
6	15-20	23	'=C5+B6
7	21-30	16	'=C6+B7
8	31-40	9	'=C7+B8
9	41-50	6	'=C8+B9
10	51-	3	'=C9+B10

Kirjoita kaava soluun C3 ja kopioi sitä alaspäin.

	A	B	C
1	Aika (h)	$f\%$	$sf\%$
2	0-2	6	6
3	3-5	7	13
4	6-9	14	27
5	10-14	16	43
6	15-20	23	66
7	21-30	16	82
8	31-40	9	91
9	41-50	6	97
10	51-	3	100

Suhteellinen summafrekvenssi ylittää 50 % rajan luokassa 15–20, joten sosiaalisessa median palveluissa vietetyn ajan mediaaniluokka on 15–20.

Vastaus

mediaaniluokka 15–20

5.13

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelmalla.

Syötetään sarakkeisiin A–C tarvittavat tiedot ja lasketaan summafrekvenssi sarakkeeseen C.

	A	B	C	D
1	Aika (h)	$f\%$	Todellinen yläraja	$sf\%$
2			0	0
3	0-4	10	4,5	'=D2+B3
4	5-10	14	10,5	'=D3+B4
5	11-14	16	14,5	'=D4+B5
6	15-20	23	20,5	'=D5+B6
7	21-30	18	30,5	'=D6+B7
8	31-40	11	40,5	'=D7+B8
9	41-50	6	50,5	'=D8+B9
10	51-60	2	60,5	'=D9+B10

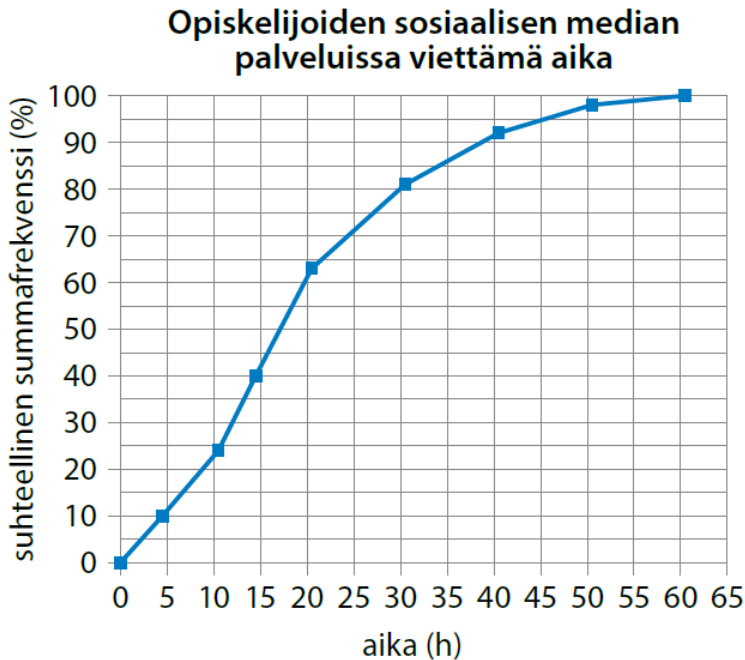
Suhteellisen summafrekvenssin ensimmäisen arvon tulee olla 0.

Lisätään sen takia soluihin C ja D arvo 0.

Kirjoita kaava soluun D3 ja kopioi sitä alaspäin.

	A	B	C	D
1	Aika (h)	$f\%$	Todellinen yläraja	$sf\%$
2			0	0
3	0-4	10	4,5	10
4	5-10	14	10,5	24
5	11-14	16	14,5	40
6	15-20	23	20,5	63
7	21-30	18	30,5	81
8	31-40	11	40,5	92
9	41-50	6	50,5	98
10	51-60	2	60,5	100

Piirretään kertymäkuvaaja Ohjattu kaavion luonti -toimintoa käyttäen.



- b) Arvioidaan sosiaalisen median palveluissa vietetyn ajan mediaani kertymäkuvaajasta. Kuvaaja ylittää 50 % rajan noin ajan 17 tuntia kohdalla. Sosiaalisen median palveluissa vietetyn ajan mediaani on 17 tuntia.
- c) Arvioidaan aika, jota enemmän sosiaalisen median palveluissa vietti 25 % opiskelijoista. Kuvaaja ylittää 75 % rajan noin tunnin 27 kohdalla. Opiskelijoista 25 % vietti sosiaalisessa mediassa yli 27 tuntia.

Vastaus

- b) 17 tuntia
c) 27 tuntia

5.14

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelmalla.

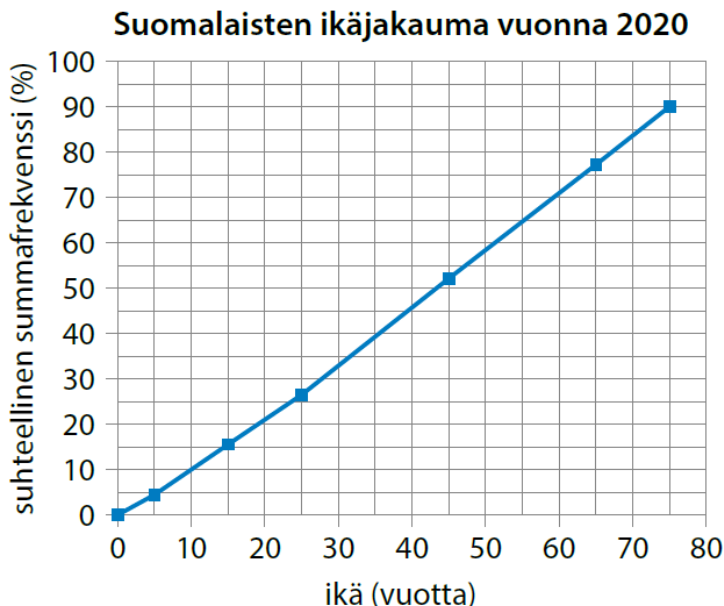
Syötetään tarvittavat tiedot sarakkeisiin A–C ja lasketaan suhteelliset summafrekvenssit sarakkeeseen D.

	A	B	C	D
1	Ikä (vuotta)	<i>f</i> %	Todellinen yläraja	<i>sf</i> %
2			0	0
3	0-4	4,46	5	'=D2+B3
4	5-14	11,09	15	'=D3+B4
5	15-24	10,96	25	'=D4+B5
6	25-44	25,67	45	'=D5+B6
7	45-64	25,12	65	'=D6+B7
8	65-74	12,80	75	'=D7+B8
9	75-	9,90		'=D8+B9

Viimeiselle luokalle ei ole annettu ylärajaa, joten se jää avoimeksi.

	A	B	C	D
1	Ikä (vuotta)	<i>f</i> %	Todellinen yläraja	<i>sf</i> %
2			0	0
3	0-4	4,46	5	4,46
4	5-14	11,09	15	15,55
5	15-24	10,96	25	26,51
6	25-44	25,67	45	52,18
7	45-64	25,12	65	77,30
8	65-74	12,80	75	90,10
9	75-	9,90		100

Piirretään kertymäkuvaaja Ohjattu kaavion luonti -toimintoa käyttäen.



Arvioidaan suomalaisten mediaani-ikä kertymäkuvaajasta. Kuvaaja ylittää 50 % rajan noin 43 vuoden kohdalla.

- b) Arvioidaan ikä, jota nuorempia oli 25 % suomalaisista. Kuvaaja ylittää 25 % rajan noin 24 vuoden kohdalla. Nuorin neljännes on siis alle 24 vuotta.
- c) Arvioidaan ikä, jota vanhempia oli 25 % suomalaisista. Kuvaaja ylittää 75 % rajan noin 63 vuoden kohdalla. Vanhin neljännes on siis yli 63 vuotta.

Vastaus

- b) 43 vuotta
- c) alle 24 vuotta
- d) yli 63 vuotta

5.15

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelmalla.

Syötetään annetut tiedot sarakkeisiin A–B ja lasketaan summafrekvenssit sarakkeeseen C.

	A	B	C
1	Ikä (vuotta)	<i>f</i>	<i>sf</i>
2		0	0
3	0-14	20	'=C2+B3
4	15-29	11	'=C3+B4
5	30-44	11	'=C4+B5
6	45-59	24	'=C5+B6
7	60-74	4	'=C6+B7

	A	B	C
1	Ikä (vuotta)	<i>f</i>	<i>sf</i>
2		0	0
3	0-14	20	20
4	15-29	11	31
5	30-44	11	42
6	45-59	24	66
7	60-74	4	70

Lasketaan seuraavaksi suhteelliset summafrekvenssit sarakkeeseen E.

	A	B	C	D	E
1	Ikä (vuotta)	<i>f</i>	<i>sf</i>	Todellinen yläraja	<i>sf</i> %
2		0	0	0	0
3	0-14	20	20	15	'=C3/\$C\$7*100
4	15-29	11	31	30	'=C4/\$C\$7*100
5	30-44	11	42	45	'=C3/\$C\$7*100
6	45-59	24	66	60	'=C6/\$C\$7*100
7	60-74	4	70	75	'=C7/\$C\$7*100

Suhteellisen summafrekvenssin saa laskettua jakamalla summafrekvenssin laulajien kokonaismäärällä (solu C7).

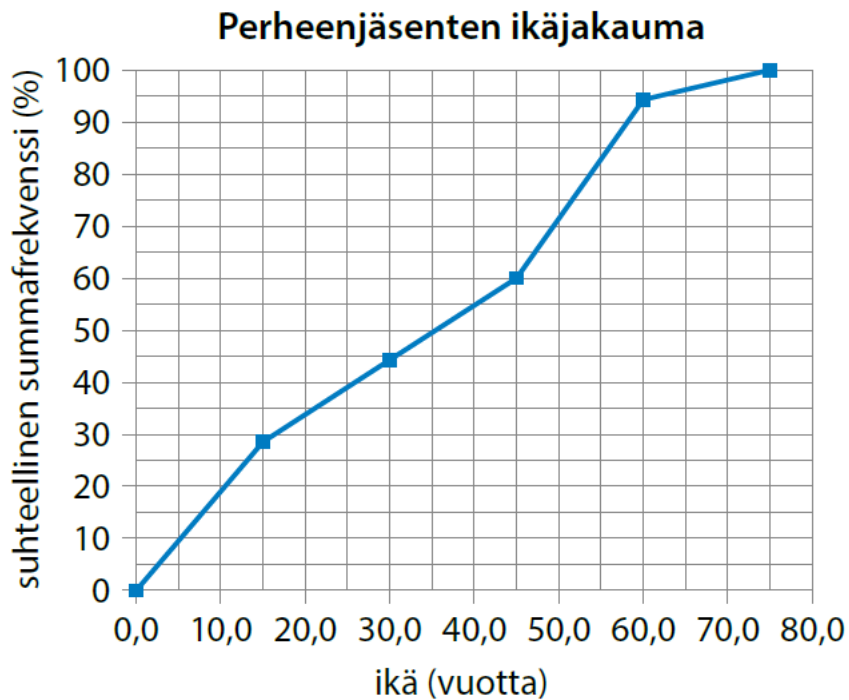
Todellinen yläraja on merkitty taulukkoon kertymäkuvaajan piirtämistä varten. Kertymäkuvaajaa varten ensimmäiselle riville on merkitty myös nollarivi.

	A	B	C	D	E
1	Ikä (vuotta)	<i>f</i>	<i>sf</i>	Todellinen yläraja	<i>sf</i> %
2		0	0	0	0
3	0-14	20	20	15	28,6
4	15-29	11	31	30	44,3
5	30-44	11	42	45	60,0
6	45-59	24	66	60	94,3
7	60-74	4	70	75	100,0

Taulukko summafrekvensseistä ja suhteellisista summafrekvensseistä:

Ikä (vuotta)	<i>sf</i>	<i>sf</i> %
0–14	20	28,6
15–29	31	44,3
30–44	42	60,0
45–59	66	94,3
60–74	70	100

b) Piirretään kertymäkuvaaja Ohjattu kaavion luonti -toimintoa käyttäen.



5.16

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelman avulla.

Syötetään annetut tiedot sarakkeisiin A, B ja D ja lasketaan summafrekvenssit sarakkeisiin C ja E.

	A	B	C	D	E
1	Alkoholinkäyttö	2000		2016	
2		<i>f</i> %	<i>sf</i> %	<i>f</i> %	<i>sf</i> %
3	Ei koskaan	11	'=B3	24	'=D3
4	Harvemmin kuin kerran kuukaudessa	28	'=C3+B4	27	'=E3+D4
5	Kerran pari kuukaudessa	35	'=C4+B5	32	'=E4+D5
6	Kerran viikossa	20	'=C5+B6	13	'=E5+D6
7	Pari kertaa viikossa	6	'=C6+B7	3	'=E6+D7
8	Vähintään 4 kertaa viikossa	1	'=C7+B8	1	'=E7+D8

	A	B	C	D	E
1	Alkoholinkäyttö	2000		2016	
2		<i>f</i> %	<i>sf</i> %	<i>f</i> %	<i>sf</i> %
3	Ei koskaan	11	11	24	24
4	Harvemmin kuin kerran kuukaudessa	28	39	27	51
5	Kerran pari kuukaudessa	35	74	32	83
6	Kerran viikossa	20	94	13	96
7	Pari kertaa viikossa	6	100	3	99
8	Vähintään 4 kertaa viikossa	1	101	1	100

Summafrekvenssit voi laskea myös käsin.

Kun lasketaan pyöristetyillä arvoilla, saattaa suhteellisten frekvenssien summa poiketa hieman 100 prosentista. Pyöristysten takia suhteellisten frekvenssien summa voi olla esimerkiksi 99 % tai 101 %, kuten vuoden 2000 aineistossa.

Vuonna 2000 ylittyi 50 % raja luokassa kerran pari kuukaudessa.
Vuonna 2000 alkoholin käytön mediaani oli kerran pari kuukaudessa.

Vuonna 2016 ylittyi 50 % raja luokassa harvemmin kuin kerran kuukaudessa. Vuonna 2016 alkoholin käytön mediaani oli harvemmin kuin kerran kuukaudessa.

- b) 80 % raja ylittyi vuonna 2000 luokassa kerran viikossa. Alkoholia useimmin juova 20 % käytti alkoholia kerran viikossa tai useammin.

Vastaavasti vuonna 2016 sama raja ylittyi luokassa kerran pari kuukaudessa. Alkoholia useimmin juova 20 % käytti alkoholia kerran pari kuukaudessa tai useammin.

Vastaus

- a) 2000: kerran pari kuukaudessa,
2016: harvemmin kuin kerran kuukaudessa
- b) 2000: kerran viikossa tai useammin,
2016: kerran pari kuukaudessa tai useammin

5.17

Ensimmäisessä luokassa summafrekvenssin pitää olla sama kuin luokan frekvenssin. Viimeisessä luokassa suhteellisen summafrekvenssin pitää olla 100. Täydennetään nämä tiedot taulukkoon.

	A	B	C	D	E
1	Ikä (vuotta)	f	$f\%$	sf	$sf\%$
2	0-4	3		3	
3	5-14		15		
4	15-24				55
5	25-44	15			
6	45-64			60	100

Nyt taulukosta nähdään, että absoluuttista summafrekvenssiä 60 vastaa suhteellinen summafrekvenssi 100 %.

Rivillä 4 suhteellista summafrekvenssiä 55 % vastaa siten absoluuttinen summafrekvenssi 33.

Rivillä 2 absoluuttista summafrekvenssiä 3 vastaa suhteellinen summafrekvenssi 5 %. Koska kyseessä on ensimmäinen luokka, pitää myös suhteellisen frekvenssin olla 5 %.

	A	B	C	D	E
1	Ikä (vuotta)	f	$f\%$	sf	$sf\%$
2	0-4	3	5	3	5
3	5-14		15		
4	15-24			33	55
5	25-44	15			
6	45-64			60	100

Rivillä 3 suhteellinen frekvenssi on 15 %, joten suhteellisen summafrekvenssin pitää olla $5\% + 15\% = 20\%$. Tätä arvoa vastaava absoluuttinen summafrekvenssi on 12.

	A	B	C	D	E
1	Ikä (vuotta)	<i>f</i>	<i>f</i> %	<i>sf</i>	<i>sf</i> %
2	0-4	3	5	3	5
3	5-14		15	12	20
4	15-24			33	55
5	25-44	15			
6	45-64			60	100

Rivillä 3 absoluuttinen summafrekvenssi on 12. Kun vähennetään siitä edellisen luokan summafrekvenssi, saadaan selville rivin 3 absoluuttinen frekvenssi $12 - 3 = 9$.

Vastaavasti rivin 4 absoluuttinen frekvenssi saadaan erotuksen avulla $33 - 12 = 21$. Myös suhteellinen frekvenssi saadaan erotuksen avulla $55\% - 20\% = 35\%$.

	A	B	C	D	E
1	Ikä (vuotta)	<i>f</i>	<i>f</i> %	<i>sf</i>	<i>sf</i> %
2	0-4	3	5	3	5
3	5-14	9	15	12	20
4	15-24	12	35	33	55
5	25-44	15			
6	45-64			60	100

Rivin 5 summafrekvenssi saadaan lisäämällä edellisen luokan absoluuttiseen summafrekvenssiin rivin 5 absoluuttinen frekvenssi eli $33 + 15 = 48$. Tätä vastaa suhteellinen summafrekvenssi 80 %.

Rivin 5 suhteellinen frekvenssi saadaan vähentämällä rivin 5 suhteellisesta summafrekvenssistä edellisen luokan summafrekvenssi $80\% - 55\% = 25\%$.

	A	B	C	D	E
1	Ikä (vuotta)	<i>f</i>	<i>f</i> %	<i>sf</i>	<i>sf</i> %
2	0-4	3	5	3	5
3	5-14	9	15	12	20
4	15-24	12	35	33	55
5	25-44	15	25	48	80
6	45-64			60	100

Viimeisen rivin eli rivin 6 puuttuvat tiedot saadaan vähentämällä rivin summafrekvensseistä edellisen rivin summafrekvenssit.

$$60 - 48 = 12$$

$$100\% - 80\% = 20\%$$

Täydennetty taulukko:

Ikä (vuotta)	<i>f</i>	<i>f</i> %	<i>sf</i>	<i>sf</i> %
0-4	3	5	3	5
5-14	9	15	12	20
15-24	21	35	33	55
25-44	15	25	48	80
45-64	12	20	60	100

Taulukkoa voi täydentää useassa eri järjestyksessä ja päätellä puuttuvat luvut myös muilla tavoin kuin tässä ratkaisussa on esitetty.

5.18

a) Ratkaistaan tehtävä LibreOffice Calc -ohjelman avulla.

Syötetään annetut tiedot sarakkeisiin A ja B ja lasketaan summafrekvenssit sarakkeeseen C.

	A	B	C
1	Arvosana	<i>f</i> %	<i>sf</i> %
2	L	8,9	'=B2
3	E	22,5	'=C2+B3
4	M	21,8	'=C3+B4
5	C	22,1	'=C4+B5
6	B	16,6	'=C5+B6
7	A	6,7	'=C6+B7
8	I	1,4	'=C7+B8

	A	B	C
1	Arvosana	<i>f</i> %	<i>sf</i> %
2	L	8,9	8,9
3	E	22,5	31,4
4	M	21,8	53,2
5	C	22,1	75,3
6	B	16,6	91,9
7	A	6,7	98,6
8	I	1,4	100,0

Summafrekvenssit voi laskea myös käsin.

50 % raja ylittyy luokassa M. Mediaaniarvosana oli siis M.

- b) Kokelaat, jotka saivat vähintään arvosanan M, saivat siis arvosanan L, arvosanan E tai arvosanan M. Summafrekvenssi arvosanan M kohdalla on 53,2 % eli vähintään arvosanan M sai 53,2 % kokelaista.
- c) Vähintään arvosanan A sai 98,6 % kokelaista. Vähennetään tästä ne kokelaat, jotka saivat paremman arvosanan kuin C eli kokelaat, jotka saivat vähintään arvosanan M.

$$98,6\% - 53,2\% = 45,4\%$$

Kokelaista 45,4 % sai arvosanan, joka oli korkeintaan C, mutta vähintään A.

Vastaus

- a) M
b) 53,2 %
c) 45,4 %

5.19

- a) Kertymäkuvaajassa 50 % raja ylittyy noin iän 44 vuotta kohdalla.
Kokoaikaisten palkansaajien mediaani-ikä on siis 44 vuotta.
- b) Kertymäkuvaajassa 20 % raja ylittyy noin 32 vuoden kohdalla.
Kokoaikaisten palkansaajien nuorin 20 % on siis alle 32-vuotiaita.
- c) Kertymäkuvaajassa 75 % raja ylittyy noin 54 vuoden kohdalla.
Kokoaikaisten palkansaajien vanhin neljännes on siis yli 54-vuotiaita.
- d) 30–50-vuotiaiden palkansaajien osuus saadaan vähentämällä alle 50-vuotiaiden prosenttiosuudesta alle 30-vuotiaiden osuus.

$$64\% - 14\% = 50\%$$

Kokoaikaisista palkansaajista 50 % on 30–50 -vuotiaita.

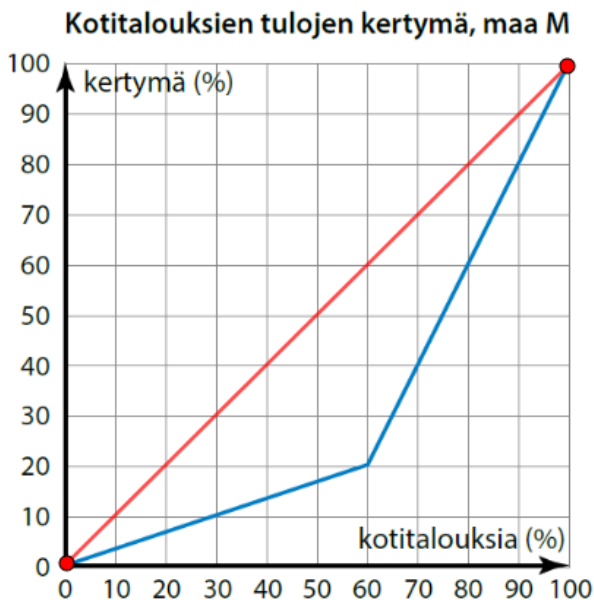
Vastaus

- a) 44 vuotta
- b) alle 32 vuotta
- c) yli 54 vuotta
- d) 50 %

5.20

- a) Tutkitaan maan N kertymäkuvaajaa. Luetaan tulojen kertymä kohdasta, jossa kotitalouksien kertymä on 10 %. Tulojen kertymä on tällä kohdalla 5 %. Vähiten ansaitseva 10 % ansaitsee siis 5 % kotitalouksien tuloista.
- Luetaan sitten tulojen kertymä kohdasta, jossa kotitalouksien kertymä on 90 %. Tulojen kertymä on tällä kohdalla 70 %. Eniten ansaitseva 10 % ansaitsee siis loput kotitalouksien tuloista.
- $$100\% - 70\% = 30\%$$
- Eniten ansaitseva 10 % ansaitsee siis 30 % kotitalouksien tuloista.

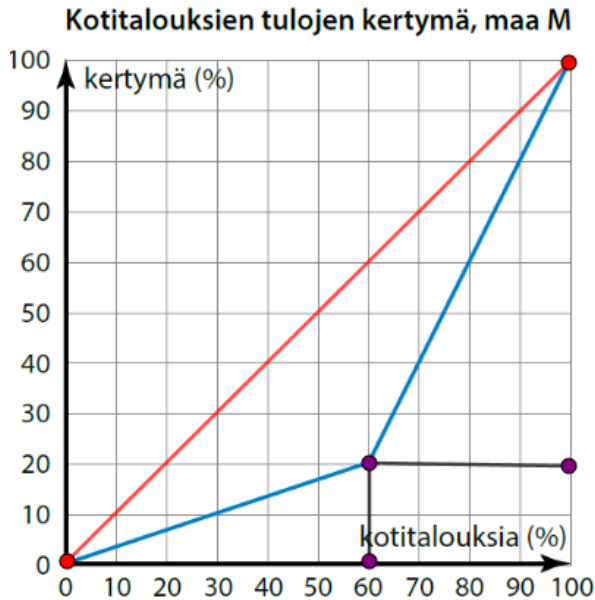
- b) Hahmotellaan kolmio maan M kertymäkuvaajaan.



Lasketaan pisteiden (0, 0), (100, 0) ja (100, 100) väliin jäävän kolmion pinta-ala. Käytetään yksikkönä koordinaatiston ruutuja.

$$\frac{10 \cdot 10}{2} = 50$$

Lasketaan seuraavaksi kertymäkuvaajan ja punaisen janan väliin jäävä pinta-ala. Vähennetään edellä lasketusta pinta-alasta kertymäkuvaajan ja vaaka-akselin väliin jäävä pinta-ala.



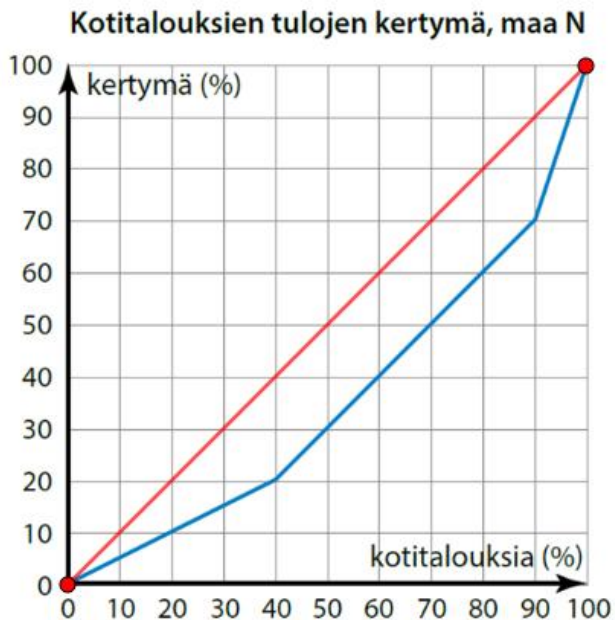
Vähennettävä pinta-ala koostuu kahdesta suorakulmaisesta kolmiosta ja suorakulmiosta (apukuva yllä).

$$50 - \frac{4 \cdot 8}{2} + \frac{6 \cdot 2}{2} + 4 \cdot 2 = 20$$

Gini-kerroin saadaan vertaamalla jälkimmäistä pinta-alaa ensiksi laskettuun pinta-alaan.

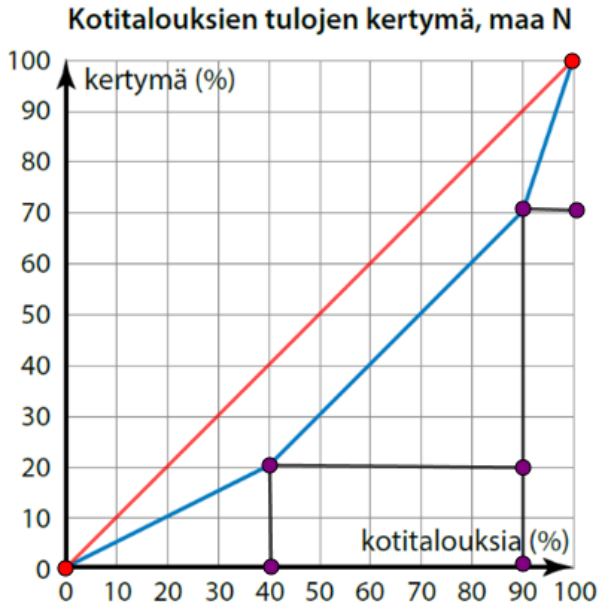
$$\frac{20}{50} = 0,4 = 40\%$$

Hahmotellaan kolmio maan N kertymäkuvaajaan.



Pisteiden $(0, 0)$, $(100, 0)$ ja $(100, 100)$ väliin jäävän kolmion pinta-ala on 50. Käytetään yksikkönä koordinaatiston ruutuja.

Lasketaan seuraavaksi kertymäkuvaajan ja punaisen janan väliin jäävä pinta-ala. Vähennetään edellä lasketusta pinta-alasta kertymäkuvaajan ja vaaka-akselin väliin jäävä pinta-ala.



Vähennettävä pinta-ala koostuu kolmesta suorakulmaisesta kolmiosta ja kahdesta suorakulmiosta (apukuva yllä).

$$50 - \frac{4 \cdot 2}{2} + \frac{5 \cdot 5}{2} + \frac{1 \cdot 3}{2} + 5 \cdot 2 + 1 \cdot 7 = 15$$

Punaisen janan ja kertymäkuvaajan väliin jäävän pinta-alan voi selvittää myös muilla tavoin. Kertymäkuvaajan ja vaaka-akselin väliin jäävän pinta-alan voi selvittää myös esimerkiksi suunnikkaan pinta-aloja käyttämällä.

Gini-kerroin saadaan vertaamalla jälkimmäistä pinta-alaa ensiksi laskettuun pinta-alaan.

$$\frac{15}{50} = 0,3 = 30\%$$

Vastaus

- Vähiten ansaitseva 10 % ansaitsee 5 % tuloista.
Eniten ansaitseva 10 % ansaitsee 30 % tuloista.
- Maa M 40 ja maa N 30