

# Binomi 5 – Luku 4 – Tehtävien malliratkaisut

## 4.1

Tutkitaan histogrammeja:

I

Histogrammin keskellä on korkeat palkit ja reunoilla matalat.  
Tässä aineistossa pieniä arvoja sekä suuria arvoja on suhteellisesti vähän.  
Oikea kuvaus on siis B.

II

Histogrammin pylväät ovat keskenään melko saman korkuisia.  
Tässä aineistossa arvot ovat melko tasaisesti jakautuneita.  
Oikea kuvaus on siis C.

III

Histogrammissa on korkeat pylväät pienten ja suurten arvojen kohdalla.  
Keskellä on matalammat pylväät.  
Tässä aineistossa pieniä ja suuria arvoja on suhteellisesti paljon.  
Oikea kuvaus on siis A.

IV

Histogrammissa on matalat pylväät pienten arvojen kohdalla ja korkeat pylväät suurten arvojen kohdalla. Aineistossa on siis vähän pieniä arvoja ja paljon suuria arvoja.  
Oikea kuvaus on siis D.

**Vastaus:**

B – I

C – II

A – III

C – IV

## 4.2

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.  
Katso tarvittaessa ohjeet Esimerkin 1 videolta.

a)

Lajitellaan aineisto nousevaan järjestykseen eli pienimmästä suurimpaan.

Lajittelun jälkeen nähdään, että pienin arvo on 5 ja suurin 19.

Kooltaan pienimmässä laumassa on siis 5 leijonaa ja suurimmassa 19 leijonaa.

b)

Taulukon luokittelu on tasavälinen eli jokaisen luokan leveys on sama.

Lasketaan ensimmäisen luokan 5–8 pituus eli luokan ylä- ja alarajan erotus.

$$8 - 5 = 3$$

Luokkavälin pituus on 3 (leijonaa).

c)

Määritetään kuhunkin luokkaan kuuluvien havaintojen lukumäärät.

### Tapa 1.

Lasketaan kuhunkin luokkaan kuuluvien havaintojen lukumäärä suuruusjärjestyksessä olevasta aineistosta.

	A	B	C	D
1	Leijonalauman koko			
2	7		5	
3	11		6	
4	12		6	Luokka 5 – 8
5	7	Lajittelu	7	11 havaintoa
6	15		7	
7	8		7	
8	19		7	
9	8		7	
10	5		8	
11	10		8	
12	14		8	
13	12		9	
14	16		9	
15	7		9	Luokka 9 – 12
16	6		10	13 havaintoa
17	17		10	
18	9		11	
19	14		11	
20	13		11	
21	9		11	
22	12		12	
23	18		12	
24	8		12	
25	11		12	
26	7		13	
27	19		14	Luokka 13 – 16
28	10		14	5 havaintoa
29	11		15	
30	7		16	
31	12		17	
32	6		18	Luokka 17 – 20
33	11		19	4 havaintoa
34	9		19	

## Tapa 2.

Käytetään luokkien ylärajoja ja Taajuus-toimintoa.

Luokka	Yläraja
'5-8	8
'9-12	12
'13-16	16
'17-20	20

→  
Taajuus-  
toiminto

Luokka	Yläraja	Frekvenssi
'5-8	8	11
'9-12	12	13
'13-16	16	5
'17-20	20	4
		0

Kootaan tiedot taulukkoon.

Lauman koko	$f$
5-8	11
9-12	13
13-16	5
17-20	4

Luokkaan kuuluvien havaintojen lukumäärät, eli luokkien frekvenssit, voidaan määrittää esimerkiksi luokkien ylärajojen perusteella ohjelman Taajuus-toiminnolla.

Katso yksityiskohdat esimerkin 1 videolta.

d)

Luokan 9-12 frekvenssi on suurin eli tässä luokassa on eniten havaintoja.

Lasketaan luokan 9-12 luokkakeskus eli luokan keskimäinen arvo.

$$\frac{9 + 12}{2} = \frac{21}{2} = 10,5$$

Luokkakeskus on 10,5 (leijonaa)

**Vastaus:**

a) pienin 5, suurin 19

b) 3 (leijonaa)

c)

Lauman koko	$f$
5-8	11
9-12	13
13-16	5
17-20	4

d) Luokassa 9-12 (leijonaa), luokkakeskus on 10,5 (leijonaa)

### 4.3

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.  
Katso tarvittaessa ohjeet Esimerkin 2 videolta.

a)

Lajitellaan aineisto nousevaan järjestykseen eli pienimmästä suurimpaan.

Lajittelun jälkeen nähdään, että aineiston pienin arvo on 62 ja suurin arvo on 360.  
Lisäksi aineistossa on tieto yhteensä 30 rivillä eli havaintoja on yhteensä 30.

Pienin arvo kuuluu ensimmäiseen luokkaan **61**–120.

- Ensimmäisen luokan yläraja on **120**.
- Seuraavan luokan alaraja on siis **121**.
- Alarajojen erotus on  $121 - 61 = 60$ .
- Luokkavälin pituus on **60**.

Lasketaan luokkien ylärajat lisäämällä edelliseen ylärajaan luku 60.  
Alarajat voidaan tämän jälkeen päätellä.

luokan alaraja	luokan yläraja
<b>61</b>	<b>120</b>
<b>121</b>	$120 + 60 = 180$
181	$180 + 60 = 240$
241	$240 + 60 = 300$
301	$300 + 60 = 360$

Suurin arvo 360 kuuluu viimeiseen luokkaan 301–360.

Määritetään kuhunkin luokkaan kuuluvien havaintojen lukumäärät esimerkiksi Taajuustoiminnolla.

Luokka	Yläraja
61-120	120
121-180	180
181-240	240
241-300	300
301-360	360

→  
Taajuus-  
toiminto

Luokka	Yläraja	
61-120	120	4
121-180	180	6
181-240	240	12
241-300	300	5
301-360	360	3
		0

Kootaan tiedot taulukkoon.

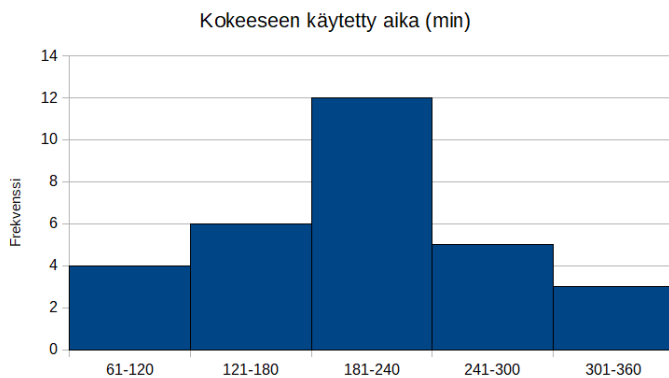
Aika (min)	$f$
61–120	4
121–180	6
181–240	12
241–300	5
301–360	3

Luokkaan kuuluvien havaintojen lukumäärät, eli luokkien frekvenssit, voidaan määrittää esimerkiksi luokkien ylärajojen perusteella ohjelman Taajuustoiminnolla.

Katso yksityiskohdat esimerkin 2 videolta.

b)

Piirretään luokitellun aineiston perusteella histogrammi taulukkolaskentaohjelmalla.



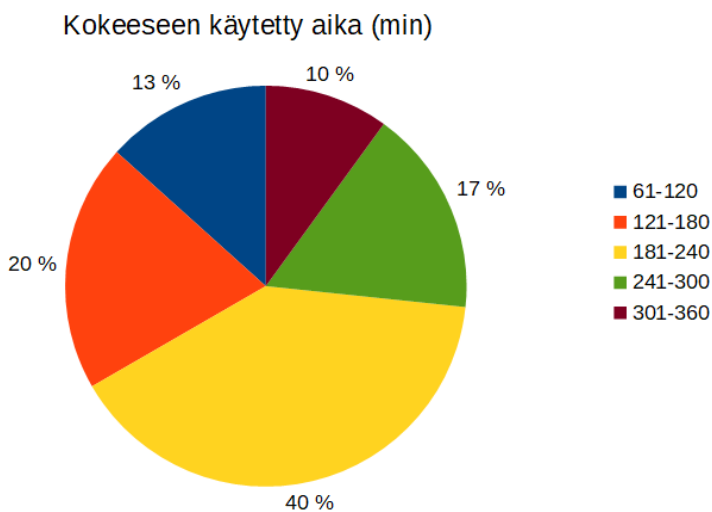
Histogrammin piirtämisen yksityiskohdat käyvät ilmi Esimerkin 2 videolta.

Histogrammi piirretään pylväskaaviona, jonka asetuksista

- objektiväliksi säädetään 0 %
- reunan leveydeksi säädetään esimerkiksi 0,02 cm.

c)

Piirretään luokitellun aineiston perusteella ympyräkaavio taulukkolaskentaohjelmalla.



Ympyräkaavion piirtämisen yksityiskohdat käyvät ilmi Esimerkin 2 videolta.

Ympyräkaavion asetuksista

- lisätään arvopisteiden otsikot
- muotoillaan arvopisteiden otsikot prosenttiluvuiksi.

Huomaa: ympyräkaaviolla havainnollistetaan prosentiosuuksia.

d)

Preliminäärikokeeseen käytetty aika jakautui välille 62–360 minuuttia eli noin 1–6 tuntia. Opiskelijat käyttivät preliminäärikokeeseen tyypillisesti aikaa 181–240 minuuttia eli noin 3–4 tuntia. Pieni osa opiskelijoista (4 opiskelijaa eli 13 %) käytti kokeeseen alle kaksi tuntia ja vielä pienempi osa (3 opiskelijaa eli 10 %) yli viisi tuntia.

#### 4.4

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

a)

Lajitellaan aineisto nousevaan järjestykseen eli pienimmästä suurimpaan. Lajittelun jälkeen nähdään, että pienin arvo on 1 ja suurin 12.

Ensimmäisen luokan alaraja on aineiston pienin arvo 1.

Muuttujana on lukumäärä, joka saa vain kokonaislukuarvoja.

Luokkavälin pituus on 2, joten ensimmäisen luokan yläraja on  $1 + 2 = 3$ .

Ensimmäinen luokka on 1–3 (maata).

Toinen luokka on tällöin 4–6, kolmas 7–9 ja neljäs 10–12.

Suurin arvo 12 kuuluu viimeiseen luokkaan.

Määritetään kuhunkin luokkaan kuuluvien havaintojen lukumäärä.

Maiden lukumäärä	$f$
1–3	8
4–6	12
7–9	6
10–12	1

Frekvenssit voidaan laskea suuruusjärjestykseen järjestetystä aineistosta (kuva oikealla).

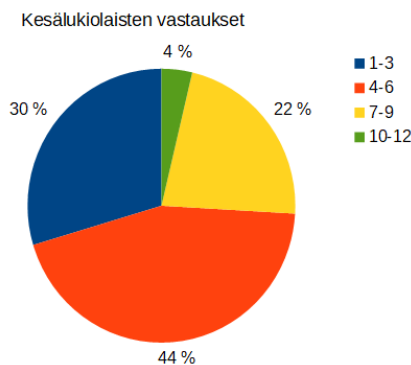
Frekvenssit, voidaan määrittää myös luokkien ylärajojen ja ohjelman Taajuus-toiminnon avulla.

3	1	
4	1	Luokka 1-3
5	2	8 havaintoa
6	2	
7	2	
8	3	
9	3	
10	3	
11	4	
12	4	Luokka 4 – 6
13	4	12 havaintoa
14	4	
15	5	
16	5	
17	5	
18	5	
19	5	
20	5	
21	6	
22	6	
23	7	

b)

Piirretään luokitellun aineiston perusteella ympyräkaavio taulukkolaskentaohjelmalla.

Kuinka monessa maassa olet kotimaan lisäksi käynyt?



Ympyräkaavion piirtämisen yksityiskohdat käyvät ilmi Esimerkin 2 videolta.

Huomaa: ympyräkaaviolla havainnollistetaan prosenttiosuuksia.

c)

Luokittelun etuina ovat selkeys ja kokonaiskuvan hahmottaminen.

Haittana, että tieto yksittäisistä vastauksista menetetään.

## 4.5

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.  
Katso tarvittaessa ohjeet Esimerkin 3 videolta.

a)

Lajitellaan jonotusajat nousevaan järjestykseen eli pienimmästä suurimpaan.  
Lajittelun jälkeen nähdään, että pienin arvo on 0 ja suurin 28.

Jonotusajan vaihteluväli on siis 0–28 minuuttia.

b)

Ensimmäisen luokan alaraja on 0.

Muuttujana on jonotusaika, joka voi periaatteessa saada mitä vain epänegatiivisia arvoja.

Luokkavälin pituus on 6, joten toisen luokan alaraja on  $0 + 6 = 6$ .

Ensimmäinen luokka on silloin 0–5 (min).

Päätellään muut luokat:

luokan alaraja	luokan yläraja
0	5
6	11
$6 + 6 = 12$	17
$12 + 6 = 18$	23
$18 + 6 = 24$	29

Seuraavan luokan alaraja saadaan lisäämällä edelliseen alarajaan luokkavälin pituus 6.

Suurin arvo 28 kuuluu viimeiseen luokkaan 24–29.

Määritetään kuhunkin luokkaan kuuluvien havaintojen lukumäärä.

Jonotusaika (min)	Yläraja	$f$
0–5	5	
6–11	11	
12–17	17	
18–23	23	
24–29	29	

→  
Taajuus-  
toiminto

Jonotusaika (min)	$f$
0–5	21
6–11	14
12–17	13
18–23	7
24–29	5

c)

Mintun jonotusaika 22 minuuttia kuuluu luokkaan 18–23.

Tämän luokan frekvenssi on 7 eli tässä luokassa on yhteensä 7 havaintoa.

Mintun kanssa samassa luokassa on siis (Minttu mukaan lukien) 7 asiakasta.

**d)**

Mikon ei tarvinnut jonottaa, joten jonotusaika on 0 minuuttia. Aika kuuluu luokkaan 0–5. Tämän luokan frekvenssi on 21 eli tässä luokassa on yhteensä 21 havaintoa.

Havaintoja on yhteensä  $21 + 14 + 13 + 7 + 5 = 60$ .

Lasketaan, kuinka monta prosenttia 21 on kokonaismäärästä 60.

$$\frac{21}{60} = 0,35 = 35 \%$$

Mikon kanssa samassa luokassa on siis (Mikko mukaan lukien) 35 % asiakkaista.

**e)**

Mintun jonotusaika on luokassa, johon kuuluu yhteensä 7 asiakasta.

Mikon jonotusaika on luokassa, johon kuuluu yhteensä 21 asiakasta.

Mikon kanssa samassa luokassa on suurempi määrä asiakkaita ( $21 > 7$ ), joten Mikon jonotusaika oli tyypillisempi.

**Vastaus:**

**a)** 0–28 (min)

**b)**

Jonotusaika (min)	$f$
0–5	21
6–11	14
12–17	13
18–23	7
24–29	5

**c)** Minttu mukaan lukien 7 asiakasta.

**d)** 35 %

**e)** Mikon jonotusaika oli tyypillisempi.



## 4.6

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

a)

Kirjoitetaan luokkien ylärajat taulukkoon.

Määritetään frekvenssit aineiston ja ylärajojen perusteella ohjelman Taajuus-toiminnolla.

Yläraja	Frekvenssi
14,9	
24,9	
34,9	
54,9	
100	

→ Taajuus-toiminto

Yläraja	Frekvenssi
14,9	9
24,9	17
34,9	12
54,9	4
100	0
	0

Taajuus-toiminnon ohjeet löytyvät esimerkin 3 videolta.

Kootaan tiedot taulukkoon.

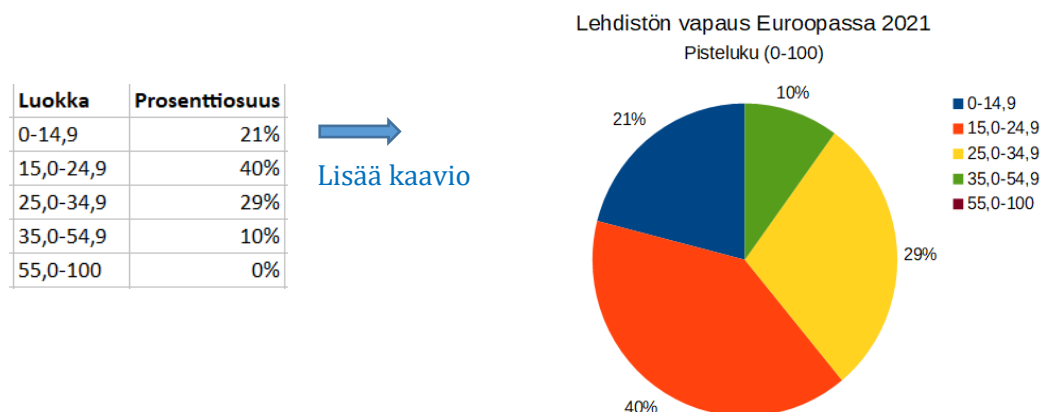
Pisteluku	$f$
0–14,9	9
15,0–24,9	17
25,0–34,9	12
35,0–54,9	4
55,0–100	0

b)

Havaintoja on yhteensä  $9 + 17 + 12 + 4 = 42$ . Lasketaan luokkien suhteelliset frekvenssit.

Pisteluku	$f$		$f\%$
0–14,9	9	$9/42 = 0,214... \approx 21\%$	21
15,0–24,9	17	$17/42 = 0,404... \approx 40\%$	40
25,0–34,9	12	$12/42 = 0,285... \approx 29\%$	29
35,0–54,9	4	$4/42 = 0,095... \approx 10\%$	10
55,0–100	0		0

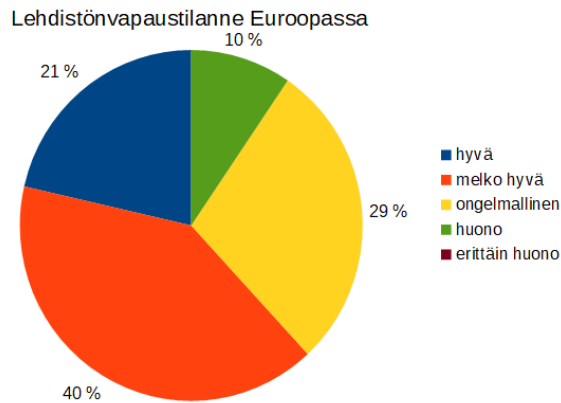
Piirretään luokkien ja suhteellisten frekvenssien perusteella ympyräkaavio.



TAI:

Luokka	Prosentiosuus
hyvä	21%
melko hyvä	40%
ongelmallinen	29%
huono	10%
erittäin huono	0%

→  
Lisää kaavio



c)

a-kohdan taulukko:

Pisteluku	Tilanne	<i>f</i>
0–14,9	hyvä	9
15,0–24,9	melko hyvä	17
25,0–34,9	ongelmallinen	12
35,0–54,9	huono	4
55,0–100	erittäin huono	0

b-kohdan taulukkoon lasketut prosentiosuudet ovat pyöristettyjä arvoja.

Pyöristetyillä arvoilla laskeminen saattaa johtaa pyöristysvirheeseen vastauksessa.

Kannattaa siis laskea lukumäärillä, jotka ovat tarkkoja arvoja.

Lasketaan taulukon perusteella, kuinka monessa maassa lehdistönvapaustilanne on ”hyvä” tai ”melko hyvä”.

$$9 + 17 = 26$$

Prosentteina tämä on

$$\frac{26}{42} = 0,619 \dots \approx 62 \%$$

62 prosentissa maista tilanne on hyvä tai melko hyvä.

d)

Selvitetään aineiston suurin arvo esimerkiksi käyttämällä ohjelman lajittele-toimintoa ja järjestämällä havaintoarvot laskevaan järjestykseen eli suurimmasta pienempään.

Aineiston suurin arvo on 50,8.

”Erittäin huono” tilanteen pisteluku on vähintään 55,0. Tällaisia arvoja ei ole aineistossa, joten aineiston mukaan Euroopassa ei ole maata, jossa tilanne olisi erittäin huono.

## 4.7

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

a)

Selvitetään aineistojen pienimmät ja suurimmat arvot esimerkiksi lajittelemalla aineistot nousevaan järjestykseen eli pienimmästä suurimpaan.

- Aineisto A: pienin arvo on 13, suurin arvo on 118
- Aineisto B: pienin arvo on 15, suurin arvo on 112

Ensimmäisen luokan alaraja on 10. Luokkavälin pituus on 20.

- Toisen luokan alaraja on  $10 + 20 = 30$ .  
Muiden luokkien alarajat saadaan samaan tapaan, lisäämällä edelliseen alarajaan luku 20.
- Luokkien ylärajat voidaan päätellä alarajoista.

Luokittelu on taulukon mukainen:

luokan alaraja	luokan yläraja
10	29
$10 + 20 = 30$	49
$30 + 20 = 50$	69
$50 + 20 = 70$	89
$70 + 20 = 90$	109
$90 + 20 = 110$	129

Pienimmät arvot 13 ja 15 kuuluvat ensimmäiseen luokkaan 10-29.

Suurimmat arvot 118 ja 112 kuuluvat viimeiseen luokkaan 110-129.

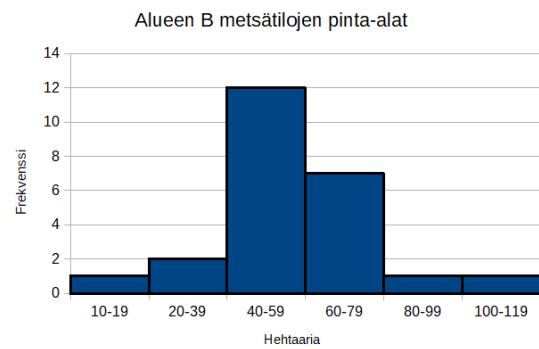
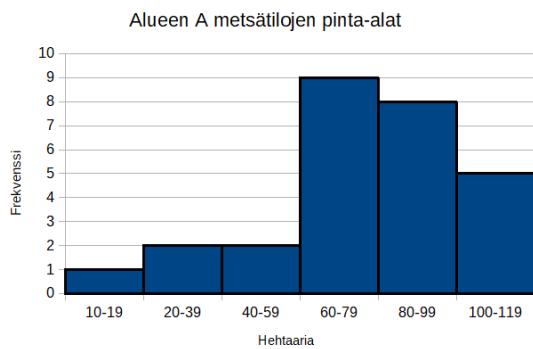
Määritetään kuhunkin luokkaan kuuluvien havaintojen lukumäärät.

Pinta-ala (hehtaaria)	Yläraja	Frekvenssi	
		A	B
10-29	29		
30-49	49		
50-69	69		
70-89	89		
90-109	109		
110-129	129		

Esim. Taajuus-toiminto

Pinta-ala (hehtaaria)	Frekvenssi	
	A	B
10-29	1	1
30-49	2	2
50-69	2	12
70-89	9	7
90-109	8	1
110-129	5	1

Piirretään luokiteltujen aineistojen perusteella histogrammit taulukkolaskentaohjelmalla.



Histogrammin piirtämisen yksityiskohdat käyvät ilmi Esimerkin 2 videolta.

Histogrammi piirretään pylväskaaviona, jonka asetuksista

- objektiväliksi säädetään 0 %
- reunan leveydeksi säädetään esimerkiksi 0,02 cm.

Alueella A on yhteensä  $1 + 2 + 2 + 9 + 8 + 5 = 27$  metsätilaa. Tiloista suurin osa, eli yhteensä  $9 + 8 + 5 = 22$  tilaa, on vähintään 60 hehtaaria olevia suuria tiloja. Pieniä tiloja on vähän.

Alueella B on yhteensä  $1 + 2 + 12 + 7 + 1 + 1 = 24$  metsätilaa. Tilat ovat keskikokoisia, tyypillisimmillään 40–59 hehtaaria. Tällä alueella pieniä ja suuria tiloja on vähän.

Yhteenvedona voidaan todeta, että alueen A tilat ovat pääosin melko suuria kun taas alueella B tilat ovat pääosin keskikokoisia.

## 4.8

Ratkaistaan tehtävä taulukkolaskentaohjelmassa.

a)

Käytetään ohjelman lajittelu-toimintoa ja järjestetään lipputulot nousevaan järjestykseen eli pienimmästä suurimpaan. Lasketaan, kuinka moni havainto kuuluu mihinkin luokkaan.

Lipputulot (miljoonaa USD)			
541,3		345,5	
707,3	→	389,0	
1 105,2	Lajittelu	461,7	Luokka 300,0-599,9
1 229,6		516,4	9 havaintoa
916,5		541,3	
612,9		543,0	
785,6		580,2	
999,6		589,3	
543,0		595,5	
839,2		612,9	
794,6		641,5	
589,3		658,5	Luokka 600,0-899,9
516,4		707,3	9 havaintoa
389,0		753,8	
461,7		785,6	
345,5		794,6	
641,5		808,7	
580,2		839,2	
595,5		916,5	
658,5		999,6	Luokka 900,0-1199,9
808,7		1 017,9	4 havaintoa
753,8		1 105,2	
1 326,7		1 229,6	Luokka 1200-1499,9
1 017,9		1 326,7	2 havaintoa

Frekvenssit voidaan määrittää myös luokkien ylärajojen avulla, ohjelman Taajuus-toiminnolla.

Kootaan tiedot taulukkoon.

Lipputulot (miljoonaa USD)	$f$
300,0-599,9	9
600,0-899,9	9
900,0-1199,9	4
1200-1499,9	2

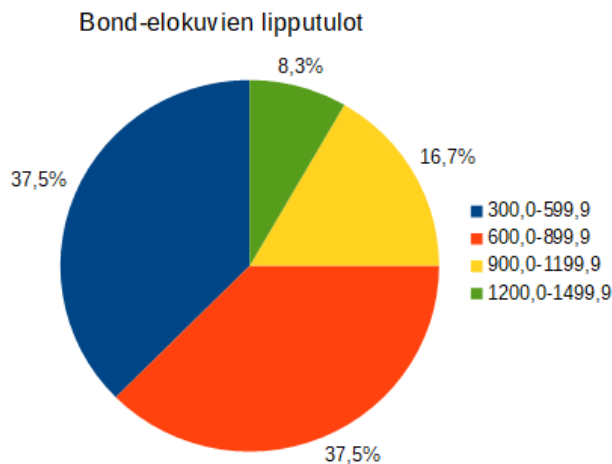
b)

Ympyräkaavio havainnollistaa suhteellisia frekvenssejä eli prosenttiosuuksia.

Lasketaan suhteelliset frekvenssit taulukkoon. Havaintoja on yhteensä  $9 + 9 + 4 + 2 = 24$ .

Lipputulot (miljoonaa USD)	$f$		$f\%$
300,0-599,9	9	$9/24 = 0,375 = 37,5 \%$	37,5
600,0-899,9	9	$9/24 = 0,375 = 37,5 \%$	37,5
900,0-1199,9	4	$4/24 = 0,1666... \approx 16,7 \%$	16,7
1200-1499,9	2	$2/24 = 0,0833... \approx 8,3 \%$	8,3

Piirretään luokitellun aineiston ja prosentiosuuksien perusteella ympyräkaavio.

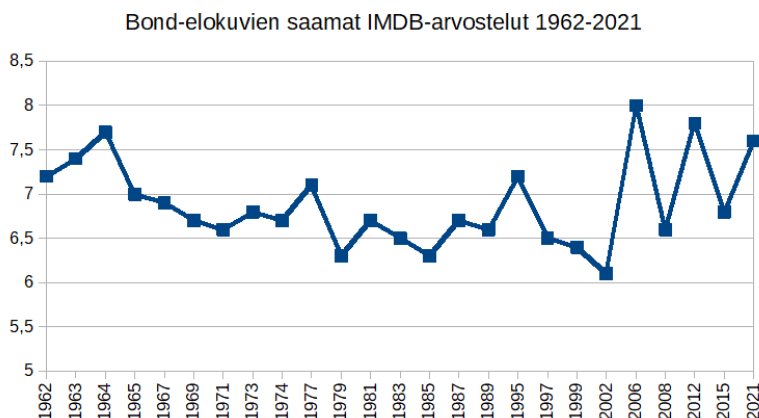


Ympyräkaavion piirtämisen yksityiskohdat käyvät ilmi Esimerkin 2 videolta.

Huomaa: ympyräkaaviolla havainnollistetaan prosentiosuuksia.

c)

Piirretään vuosilukujen ja IMDB-pisteiden perusteella viivakaavio.



Kertaa tarvittaessa viivakaavion piirtäminen luvun 3 esimerkistä 2.

d)

Viivakaaviosta nähdään, että IMDB-pistemäärä on noussut eniten vuosien 2002 ja 2006 välillä.

Vuoden 2002 elokuva on "Kuolema saa odottaa". Pistemäärä: 6,1

Vuoden 2006 elokuva on "Casino Royal". Pistemäärä: 8,0

Lasketaan, kuinka monta prosenttia nousu on.

$$\frac{8,0 - 6,1}{6,1} = \frac{1,9}{6,1} = 0,311 \dots \approx 31 \%$$

**Vastaus:**

**d)** Pistemäärä on noussut eniten elokuvien Kuolema saa odottaa (2002) ja Casino Royal (2006) välillä. Nousu on 31 %.

#### 4.9

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.  
Katso tarvittaessa ohjeet Esimerkin 3 videolta.

a)

Lajitellaan kappaleiden pituudet nousevaan järjestykseen eli pienimmästä suurimpaan.  
Lajittelun jälkeen nähdään, että pienin arvo on **155** ja suurin **314**.

Tunnuskappaleiden pituuden vaihteluväli on siis 155–314 sekuntia.

b)

Ensimmäisen luokan alaraja on aineiston pienin arvo **155**.

Muuttujana on kappaleen pituus, joka voi periaatteessa saada mitä vain epänegatiivisia arvoja.

Luokkavälin pituus on 50, joten toisen luokan alaraja on  $155 + 50 = 205$ .

Ensimmäinen luokka on silloin 155–204 (sekuntia).

Päätellään muut luokat:

luokan alaraja	luokan yläraja
<b>155</b>	204
<b>205</b>	254
$205 + 50 = 255$	304
$255 + 50 = 305$	354

Seuraavan luokan alaraja saadaan lisäämällä edelliseen alarajaan luokkavälin pituus 50.

Suurin arvo **314** kuuluu viimeiseen luokkaan 305–354.

Määritetään kuhunkin luokkaan kuuluvien havaintojen lukumäärä.

Pituus (s)	Yläraja	$f$
155–204	204	
205–254	254	
255–304	304	
305–354	354	

→  
Taajuus-  
toiminto

Pituus (s)	$f$
155–204	12
205–254	7
255–304	6
305–354	1

c)

Skyfallin pituus 286 sekuntia kuuluu luokkaan 255–304.

Tämän luokan frekvenssi on 6 eli tässä luokassa on yhteensä 6 havaintoa.

Skyfallin kanssa samassa luokassa on siis (Skyfall mukaan lukien) 6 tunnuskappaletta.

d)

Thunderballin pituus 183 sekuntia kuuluu luokkaan 155–204.

Tämän luokan frekvenssi on 12 eli tässä luokassa on yhteensä 12 havaintoa.

Havaintoja on yhteensä  $12 + 7 + 6 + 1 = 26$ .

Lasketaan, kuinka monta prosenttia 12 on kokonaismäärästä 26.

$$\frac{12}{26} = 0,461 \dots = 46 \%$$

Thunderballin kanssa samassa luokassa on siis (Thunderball mukaan lukien) noin 46 % tunnuskappaleista.

**Vastaus:**

**a)** 155–314 (sekuntia)

**b)**

Pituus (s)	<i>f</i>
155–204	12
205–254	7
255–304	6
305–354	1

**c)** 6 tunnuskappaletta (Skyfall mukaan lukien).

**d)** 46 %



#### 4.10

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

a)

Aineistossa on tieto yhteensä 54 rivillä eli tieto pesien lukumäärästä yhteensä 54 yhdyskunnassa.

Aineiston mukaan merimetsoyhdyskuntia oli yhteensä 54.

b)

Lajitellaan aineisto pesämäärän mukaan nousevaan järjestykseen eli pienimmästä suurimpaan. Lajittelun jälkeen nähdään, että

- pienin arvo on **2**, Lumijoki
- suurin arvo on **3600**, Rauma.

Pienin yhdyskunta (2 pesää) oli Lumijoella. Suurin yhdyskunta (3600 pesää) oli Raumalla.

c)

Ensimmäinen luokka on 0–399.

Muuttujana on pesien lukumäärä, joka saa vain kokonaislukuarvoja.

Luokkavälin pituus on tällöin  $399 - 0 = 399$ .

Toisen luokan alaraja on 400 ja yläraja  $400 + 399 = 799$ , eli toinen luokka on 400–799.

Päätellään muut luokat:

Pesien lukumäärä	
0–399	
400–799	
800–1199	
1200–1599	
1600–1999	
2000–2399	
2400–2799	
2800–3199	
3200–3599	
3600–3999	

Suurin arvo **3600** kuuluu viimeiseen luokkaan 3600–3999.

Määritetään kuhunkin luokkaan kuuluvien havaintojen lukumäärä.

Pesien lukumäärä	Yläraja	$f$
0-399	399	
400-799	799	
800-1199	1199	
1200-1599	1599	
1600-1999	1999	
2000-2399	2399	
2400-2799	2799	
2800-3199	3199	
3200-3599	3599	
3600-3999	3999	

→  
Esim.  
Taajuus-  
toiminto

Pesien lukumäärä	$f$
0-399	35
400-799	9
800-1199	3
1200-1599	6
1600-1999	0
2000-2399	0
2400-2799	0
2800-3199	0
3200-3599	0
3600-3999	1

Aineiston mukaan Espoon yhdyskunnassa on 443 pesää. Tämä määrä kuuluu välille 400-799. Espoon merimetsöyhdyskunta kuuluu siis pesämäärältään luokkaan 400-799 (pesää).

**d)**

Aineistossa on yhteensä  $35 + 9 + 3 + 6 + 1 = 54$  havaintoa.

Lasketaan luokkien suhteelliset frekvenssit.

Pesien lukumäärä	$f$		$f\%$
0-399	35	$35/54 = 0,6481\dots$ $\approx 64,8 \%$	64,8
400-799	9	$9/54 = 0,1666\dots$ $\approx 16,7 \%$	16,7
800-1199	3	$3/54 = 0,0555\dots$ $\approx 5,6 \%$	5,6
1200-1599	6	$6/54 = 0,1111\dots$ $\approx 11,1 \%$	11,1
1600-1999	0		0
2000-2399	0		0
2400-2799	0		0
2800-3199	0		0
3200-3599	0		0
3600-3999	1	$1/54 = 0,0185\dots$ $\approx 1,9 \%$	1,9

Luokitellun aineiston perusteella pieniä yhdyskuntia on paljon, keskikokoisia on jonkin verran ja suuria vähän. Joukossa on yksi poikkeuksellisen suuri yhdyskunta (Rauma, 3600 pesää).

#### 4.11

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

a)

Järjestetään aineisto ohjelman lajittelu-toiminnolla nousevaan järjestykseen eli pienimmästä suurimpaan. Nähdään, että aineiston pienin arvo on 200,9 ja suurin arvo on 407,6.

Juoksuajan vaihteluväli on siis 200,9–407,6 (minuuttia).

b)

Määritetään kuhunkin luokkaan kuuluvien havaintojen lukumäärä.

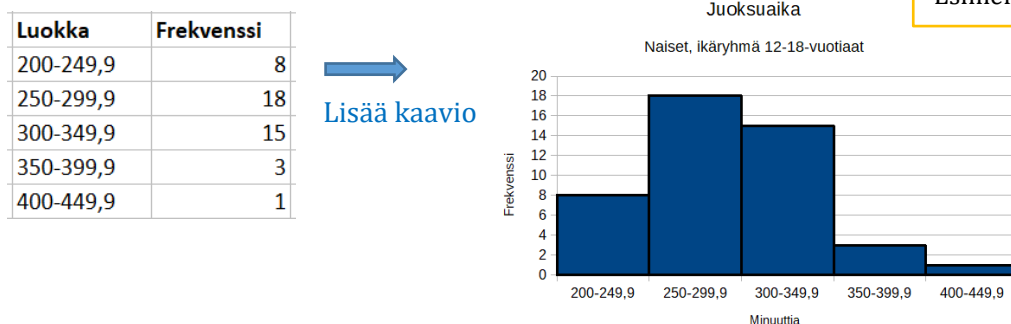
Juoksuaika (min)	Yläraja	$f$
200,0–249,9	249,9	
250,0–299,9	299,9	
300,0–349,9	349,9	
350,0–399,9	399,9	
400,0–449,9	449,9	

→  
Esim.  
Taajuus-  
toiminto

Juoksuaika (min)	$f$
200,0–249,9	8
250,0–299,9	18
300,0–349,9	15
350,0–399,9	3
400,0–449,9	1

Piirretään luokittelun aineiston perusteella histogrammi.

Histogrammin piirtämisen yksityiskohdat käyvät ilmi Esimerkin 2 videolta.



c)

Kirjoitetaan nuorimman juoksijan aika 5 h 56 min 42 sekuntia minuuteiksi.

$$5 \text{ h} = 5 \cdot 60 \text{ min} = 300 \text{ min}$$

$$42 \text{ s} = 42/60 \text{ min} = 0,7 \text{ min}$$

$$300 \text{ min} + 56 \text{ min} + 0,7 \text{ min} = 356,7 \text{ min}$$

Nuorimman juoksijan aika (356,7 min) kuuluu b-kohdan luokittelussa välille 350–399,9 minuuttia. Tähän luokkaan kuuluu yhteensä 3 juoksijaa eli prosentteina

$$\frac{3}{45} = 0,0666 \dots \approx 6,7 \%$$

Ikäryhmän juoksijoista vain 6,7 % oli nuorimman kanssa samassa tulosluokassa. Tämän perusteella ei voida sanoa, että nuorimman juoksijan aika olisi tyypillinen tässä ikäryhmässä. (Sen sijaan tyypillinen aika tässä ikäryhmässä on b-kohdan luokittelussa välillä 250–299,9 minuuttia, sillä tämän luokan frekvenssi on suurin.)

#### 4.12

Ratkaistaan tehtävä taulukkolaskentaohjelmalla.

a)

Järjestetään aineisto ohjelman lajittelu-toiminnolla nousevaan järjestykseen eli pienimmästä suurimpaan. Nähdään, että aineiston pienin arvo on 1559 ja suurin arvo on 2284.

Cooperin testissä tulos on sitä parempi, mitä suurempi se on.

Ryhmän huonoin tulos oli 1559 metriä ja paras 2284 metriä.

b)

Ensimmäisen luokan alaraja on 1500 ja hyvän tuloksen raja 2000 metriä on yhden luokan alarajana.

Jos valitaan luokkavälin pituudeksi 250, niin

- toisen luokan alaraja on  $1500+250 = 1750$
- kolmannen luokan alaraja on  $1750+250 = 2000$
- neljännen luokan alaraja on  $2000+250 = 2250$ .

Tällöin luokittelu olisi taulukon mukainen.

Tulos (m)	<i>f</i>
1500–1749	
1750–1999	
2000–2249	
2250–2499	

Suurin arvo 2284 kuuluu nyt viimeiseen luokkaan 2250–2499, joten tämä luokittelu kelpaa tehtävänantoon.

Lasketaan kuhunkin luokkaan kuuluvien havaintojen lukumäärä suuruusjärjestyksessä olevasta aineistosta.

Tulokset (metriä)	
1559	
1658	Luokka 1500-1749
1680	7 havaintoa
1700	
1712	
1725	
1745	
1806	
1822	
1832	
1834	Luokka 1750-1999
1843	12 havaintoa
1865	
1867	
1905	
1916	
1935	
1955	
1994	
2010	
2034	Luokka 2000-2249
2085	7 havaintoa
2120	
2140	
2144	
2154	
2284	Luokka 2250-2499
	1 havainto

← Frekvenssit voidaan määrittää myös luokkien ylärajojen avulla, ohjelman Taajuus-toiminnolla.



Kootaan tiedot taulukkoon.

<b>Tulos (m)</b>	<b><i>f</i></b>
1500-1749	7
1750-1999	12
2000-2249	7
2250-2499	1

c)

Luokittelun etuina ovat selkeys ja ryhmän kokonaiskuvan hahmottaminen. Haittana on se, että tieto yksittäisistä juoksutuloksista menetetään.

#### 4.13

a)

Aineistossa on tieto 198 rivillä, joten tieto metsäpinta-alasta on 198 maan osalta.

b)

Tehdään luokittelu taulukkolaskentaohjelmalla.

Kirjoitetaan luokkien ylärajat taulukkoon. Määritetään luokkien frekvenssit ohjelman Taajuus-toiminnolla.

Yläraja	Frekvenssi
19,9	
39,9	
59,9	
79,9	
99,9	

→  
Taajuus-  
toiminto

Yläraja	Frekvenssi
19,9	74
39,9	52
59,9	44
79,9	21
99,9	7
	0

Taajuus-toiminnon ohjeet löytyvät esimerkin 3 videolta.

Kootaan tiedot taulukkoon.

Metsää (%)	$f$
0-19,9	74
20-39,9	52
40-59,9	44
60-79,9	21
80-99,9	7

c)

Lasketaan luokkien suhteelliset frekvenssit. Aineistossa on yhteensä 198 havaintoa.

Metsää (%)	$f$		$f\%$
0-19,9	74	$74/198 = 0,3737... \approx 37,4\%$	37,4
20-39,9	52	$52/198 = 0,2626... \approx 26,3\%$	26,3
40-59,9	44	$44/198 = 0,2222... \approx 22,2\%$	22,2
60-79,9	21	$21/198 = 0,1060... \approx 10,6\%$	10,6
80-99,9	7	$7/198 = 0,03535... \approx 3,5\%$	3,5

Laskut voidaan tehdä taulukkolaskenta-ohjelmassa.

Tehtävänannon perusteella Suomen pinta-alasta metsää on 73,7 %.

Suomi kuuluu tämän perusteella metsäisyydessä luokkaan 60-79,9 (%).

Suomen kanssa samaan tai korkeampaan luokkaan kuuluu yhteensä  $21 + 7 = 28$  maata.

Prosentteina tämä on

$$\frac{28}{198} = 0,1414... \approx 14,1\%$$

**Tapa 2.** Suomen kanssa samassa tai korkeammassa luokassa olevia maita on yhteensä  $10,6\% + 3,5\% = 14,1\%$ .

**d)**

Suomen pinta-alasta on metsää 73,7 %.

Venäjän pinta-alasta on metsää 49,8 %.

$$\frac{73,7 - 49,8}{49,8} = \frac{23,9}{49,8} = 0,479 \dots \approx 48 \%$$

Metsän osuus Suomen pinta-alasta on 48 % suurempi Venäjän metsäosuuteen verrattuna.

**Vastaus:**

**a)** 198 maata

**c)** 14,1 %

**d)** 48 %



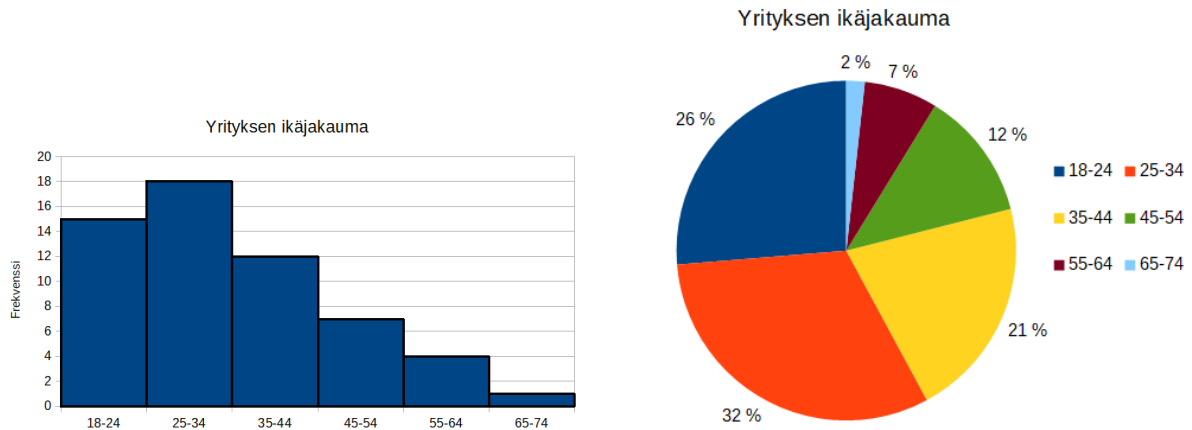
#### 4.14

a)

Yrityksen työntekijöiden ikäjakaumaa voidaan havainnollistaa histogrammilla tai ympyräkaaviolla.

Piirretään kaavio taulukkolaskentaohjelmalla.

Katso tarvittaessa ohjeet kaavion piirtämiseen Esimerkin 2 videolta.



Huomaa: ympyräkaaviolla havainnollistetaan prosenttiosuuksia.

b)

Laaditaan frekvenssijakaumat taulukkolaskentaohjelmalla.

Kertaa tarvittaessa jakaumien muodostaminen luvun 2 Esimerkistä 3.

Lasketaan frekvenssien summa. Käytetään ohjelman summa-toimintoa.

Frekvenssien summa, eli havaintojen kokonaismäärä, on 57.

Lasketaan suhteelliset frekvenssit. Ilmaistetaan suhdeluvun prosentteina.

	A	B	C
1	YRITYKSEN IKÄJAKAUMA		
2		Frekvenssi	f %
3	18-24	15	=B3/57
4	25-34	18	
5	35-44	12	
6	45-54	7	
7	55-64	4	
8	65-74	1	
9		57	

→

Kopioi kaava.  
Muotoile prosentiksi.

	A	B	C
1	YRITYKSEN IKÄJAKAUMA		
2		Frekvenssi	f %
3	18-24	15	26,32 %
4	25-34	18	31,58 %
5	35-44	12	21,05 %
6	45-54	7	12,28 %
7	55-64	4	7,02 %
8	65-74	1	1,75 %
9		57	

Lasketaan summafrekvenssit.

	A	B	C	D
1	YRITYKSEN IKÄJAKAUMA			
2		Frekvenssi	f %	sf
3	18-24	15	26,32 %	15
4	25-34	18	31,58 %	=D3+B4
5	35-44	12	21,05 %	
6	45-54	7	12,28 %	
7	55-64	4	7,02 %	
8	65-74	1	1,75 %	
9		57		

	A	B	C	D
1	YRITYKSEN IKÄJAKAUMA			
2		Frekvenssi	f %	sf
3	18-24	15	26,32 %	15
4	25-34	18	31,58 %	33
5	35-44	12	21,05 %	45
6	45-54	7	12,28 %	52
7	55-64	4	7,02 %	56
8	65-74	1	1,75 %	57
9		57		

Lasketaan lopuksi suhteelliset summafrekvenssit.

Esimerkiksi luokka 18–24:  $15/57 = 0,26315... \approx 26,32 \%$

	A	B	C	D	E
1	<b>YRITYKSEN IKÄJAKAUMA</b>				
2		<b>Frekvenssi</b>	<b>f %</b>	<b>sf</b>	<b>sf %</b>
3	18-24	15	26,32 %	15	26,32 %
4	25-34	18	31,58 %	33	57,89 %
5	35-44	12	21,05 %	45	78,95 %
6	45-54	7	12,28 %	52	91,23 %
7	55-64	4	7,02 %	56	98,25 %
8	65-74	1	1,75 %	57	100,00 %

c)

Tulkitaan frekvenssijakaumaa:

Yli puolet (57,89 %) työntekijöistä on alle 35-vuotiaita ja yli 90 % (91,23 %) työntekijöistä on alle 55-vuotiaita. Yrityksen työntekijät ovat siis verraten nuoria.

d)

Verrataan yrityksen ikäjakaumaa valtakunnalliseen ikäjakaumaan.

Ikä (vuotta)		18–24	25–34	35–44	45–54	55–64	65–74
f %	valtakunnallinen	9,2	22,2	24,3	23,1	19,8	1,4
	yritys	26,3	31,6	21,1	12,3	7,0	1,8

Valtakunnallisesti ikäluokat ovat nuorinta ja vanhinta luokkaa lukuun ottamatta suurin piirtein yhtä suuria (noin 20 % työssä käyvistä).

Yrityksen ikäjakaumassa ikäluokat eivät ole yhtä suuria. Nuorimmat ikäluokat 18–24 ja 25–34 ovat suurempia kuin valtakunnallisesti. Vastaavasti vanhemmat ikäluokat 45–54 ja 55–64 ovat pienempiä kuin valtakunnallisesti.

Tällä perusteella ei voida sanoa, että yrityksen ikäjakauma noudattaisi valtakunnallista jakaumaa.