

MTTTP1, luento 13.9.2018

KERTAUSTA

- Havaintomatriisi

Tilastoyksiköt: $a_1, a_2, a_3, \dots, a_n$

Muuttujat: $x_1, x_2, x_3, \dots, x_p$

Havaintomatriisissa n riviä ja p saraketta

	x_1	$x_2 \dots$	$x_j \dots$	x_p
a_1	x_{11}	$x_{12} \dots$	$x_{1j} \dots$	x_{1p}
a_2	x_{21}	$x_{22} \dots$	$x_{2j} \dots$	x_{2p}
..				
a_i	x_{i1}	$x_{i2} \dots$	$x_{ij} \dots$	x_{ip}
.				
a_n	x_{n1}	$x_{n2} \dots$	$x_{nj} \dots$	x_{np}

- Mitta-asteikot

Nominaaliasteikko

Järjestysasteikko

Intervalliasteikko

Suhdeasteikko

Absoluuttinen asteikko

http://www.sis.uta.fi/tilasto/mtt1/syksy2018/mitta_asteikot_kuva.pdf

Esim. CTESTI-aineiston muuttujien mitta-asteikot

IKÄ	Oppilaan ikä 1.11.1974 täysinä vuosina
PITUUS	Oppilaan pituus 1 cm tarkkuudella
PAINO	Oppilaan paino 100 g tarkkuudella.
HENGITIL	Hengitystilavuus 100 cm ³ tarkkuudella
COOPER	Cooperin testin tulos
LEUANVET	Leuanvetojen lukumäärä
KOULU	Oppilaan koulu MYP=0, TKYL = 1
LKASTE	8-luokkaisen koulun I, III, V, VII, Vastaavuus peruskoulun 5, 7, 9 ja lukion 2.
ÄIDINK1, ÄIDINK2, MATEM, KIELI1, KIELI2, KIELI3, LIIKUNTA	Ko. aineitten koulunumerot
SOSRYHMÄ	Isän ammatin koodaus 1-9, Rauhala: 1960-luvun suomalaisten yhteiskunnan sosiaalinen kerrostuneisuus ammatin arvostuksen valossa
KIPPI	Koodaus pääsee kipin= 1, ei pääse =0
PUOLIV	Koodaus pääsee = 1, ei pääse = 0
HAVOPP	Oppilaan numero aineistossa

5 EMPIIRISET JAKAUMAT

5.1 Yksiulotteinen jakauma

Havaintomatriisin sarakkeilla on muuttujien $x_1, x_2, x_3, \dots, x_p$ jakaumat.

5.1.1 Frekvenssijakauma

Esim. Opintojaksolle ilmoittautuneet, $n=442$, 10.9.

Opiskelijat tutkinto-ohjelmittain (%)

Matematiikka & Tilastotiede	10
Tietojenkäsittelytieteet	31
Kauppatieteet	24
Hallintotieteet	20
Muut	15

Tutkinto-ohjelmaopiskelijoista ensimmäisen vuoden opiskelijoita 17 %

Esim. 5.1.2. Yritykset toimialoittain

<u>Toimiala</u>	<u>Frekvenssi</u>
Elintarviketeollisuus	7
Ilmailu	18
IT	22
Lääketeollisuus	12
Paperinjalostus	11
Öljynjalostus	27
	<hr/>
	97

Esim. 5.1.1. Lepopulssin jakauma, PULSSI-aineisto, liite 4

Pulssi -muuttujan luokat	Tilastoyksiköiden määrä luokassa
42-52	4
53-63	7
64-74	31
75-85	25
86-96	12
97-107	1

Esim. 5.1.5 ja 5.1.6 Lepopulssin mittaustarkkuus 1, pyöristetyt luokkarajat, todelliset luokkarajat, luokkakeskukset, luokan pituus 11, frekvenssit, summafrequenssit

<u>pyöristetyt luokkarajat</u>	<u>todelliset luokkarajat</u>	<u>Luokkak.</u>	<u>Frekv.</u>	<u>Summafrev.</u>
42-52	41,5-52,5	47	4	4
53-63	52,5-63,5	58	7	11
64-74	63,5-74,5	69	31	42
75-85	74,5-85,5	80	25	67
86-96	85,5-96,5	91	12	79
97-107	96,5-107,5	102	1	80

-
Frekvenssit ja summafrequenssit voidaan esittää myös prosentteina.

Esim. 5.1.11. Miesopiskelijoiden pituus.

<u>pyöristetyt luokkarajat</u>	<u>frekv.</u>	<u>summa- frekv.</u>	<u>luokka- keskus</u>	<u>todelliset luokkarajat</u>
154-160	5	5	157	153,5-160,5
161-167	20	25	164	160,5-167,5
168-174	39	64	171	167,5-174,5
175-181	28	92	178	174,5-181,5
182-188	8	100	185	181,5-188,5

Mittaustarkkuus 1, luokan pituus 7

Ehdolliset frekvenssijakaumat, tarkastellaan frekvenssijakaumaa toisen muuttujan mukaan ryhmiteltynä.

Esim. 5.1.7. Lepopulssin jakauma miehillä ja naisilla

	Mies	Nainen
42-52	4 9,1 %	0 0,0 %
53-63	6 13,6 %	1 2,8 %
64-74	18 40,9 %	13 36,1 %
75-85	13 29,5 %	12 33,3 %
86-96	3 6,8 %	9 25,0 %
97-107	0 0 %	1 2,8 %
	44	36

Esim 5.1.4. Tampereen yliopistosta maisterin tutkinnosta 2016 valmistuneiden työelämään sijoittuminen, työtilanne koulutusaloittain vuosi valmistumisen jälkeen

<https://www.uta.fi/opiskelunopas/tyoelama/valmistuneet-tyoelamassa>

<https://intra.uta.fi/portal/documents/159280/44060654/sijoittumis seuranta+2016.pdf/71ca38b5-90a6-4378-bbbf-1e697853e49a> (s. 14)

Esim. Huoneiden lukumäärä alueittain, aineisto

Tre_myydyt_asunnot_2012.sav sivulla

<https://coursepages.uta.fi/mhttp1/esimerkkiaineistoja/>

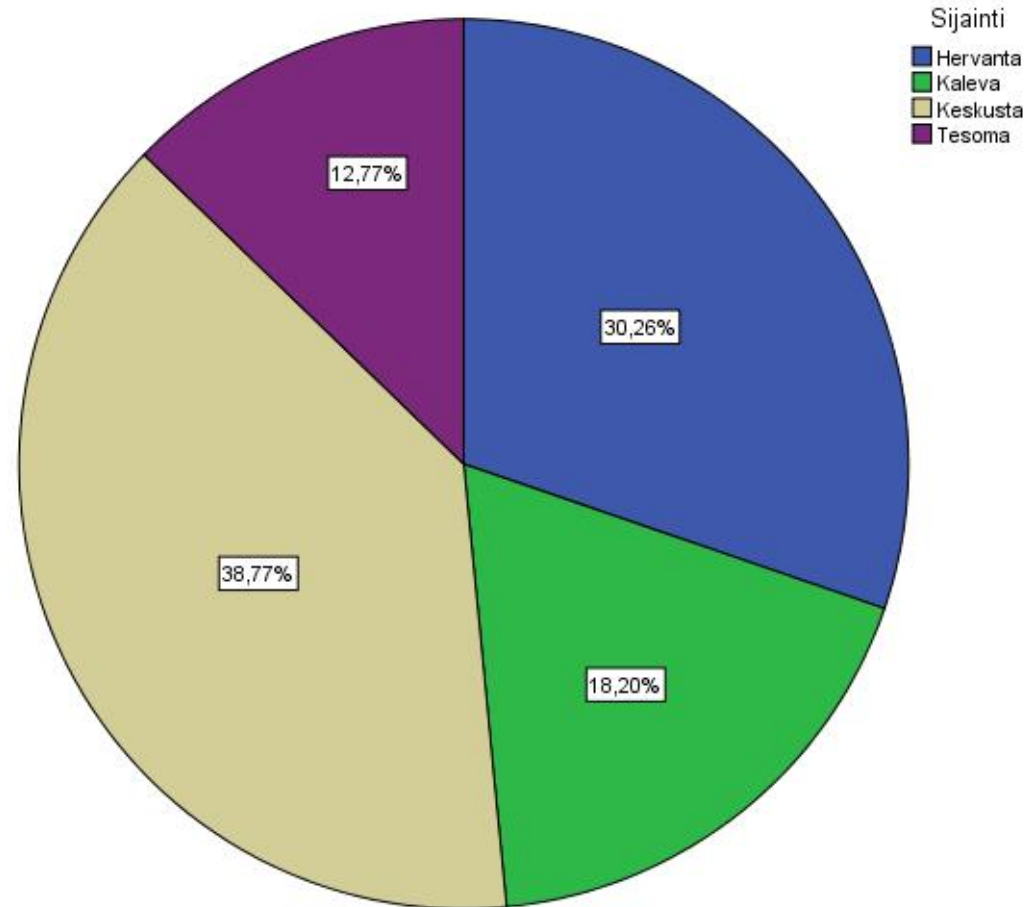
		Sijainti					
		Hervanta	Kaleva	Keskusta	Tesoma	Total	
Huoneiden lukumäärä	1	Count	21	15	45	5	86
		% within Sijainti	16,4%	19,5%	27,4%	9,3%	20,3%
	2	Count	53	43	66	14	176
		% within Sijainti	41,4%	55,8%	40,2%	25,9%	41,6%
	3	Count	42	15	41	19	117
		% within Sijainti	32,8%	19,5%	25,0%	35,2%	27,7%
	4	Count	12	4	12	16	44
		% within Sijainti	9,4%	5,2%	7,3%	29,6%	10,4%
Total		Count	128	77	164	54	423
		% within Sijainti	100,0%	100,0%	100,0%	100,0%	100,0%

5.1.2 Frekvenssijakaumien graafiset esitykset

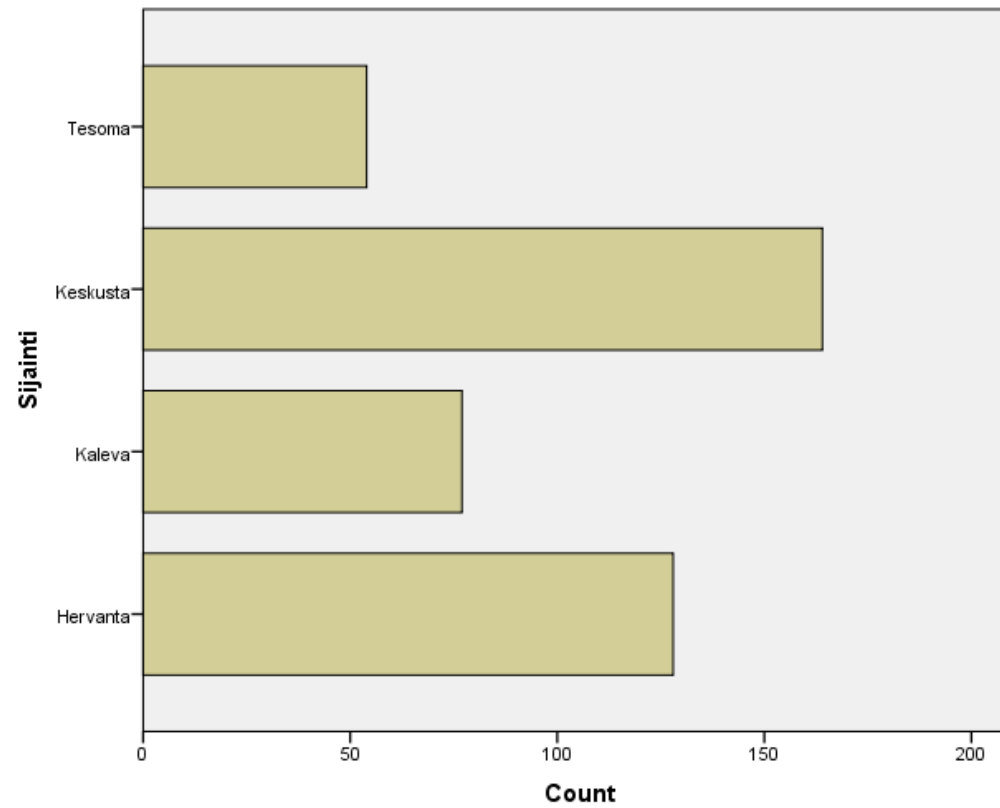
- Piirakkakuvio
- Pylväs-, vaakapylväs-, janadiagrammi
- Frekvenssihistogrammi

Esim. Aineisto Tre_myydyt_asunnot_2012.sav, muuttujien sijainti, huoneiden lukumäärä, neliöhinta graafiset esitykset ovat piirakkakuvio tai vaakapylväsdiagrammi, pylväsdiagrammi tai janadiagrammi, frekvenssihistogrammi. Neliöhintaa voidaan tarkastella myös sijainnin mukaan ehdollistettuna samoin huoneiden lukumäärää.

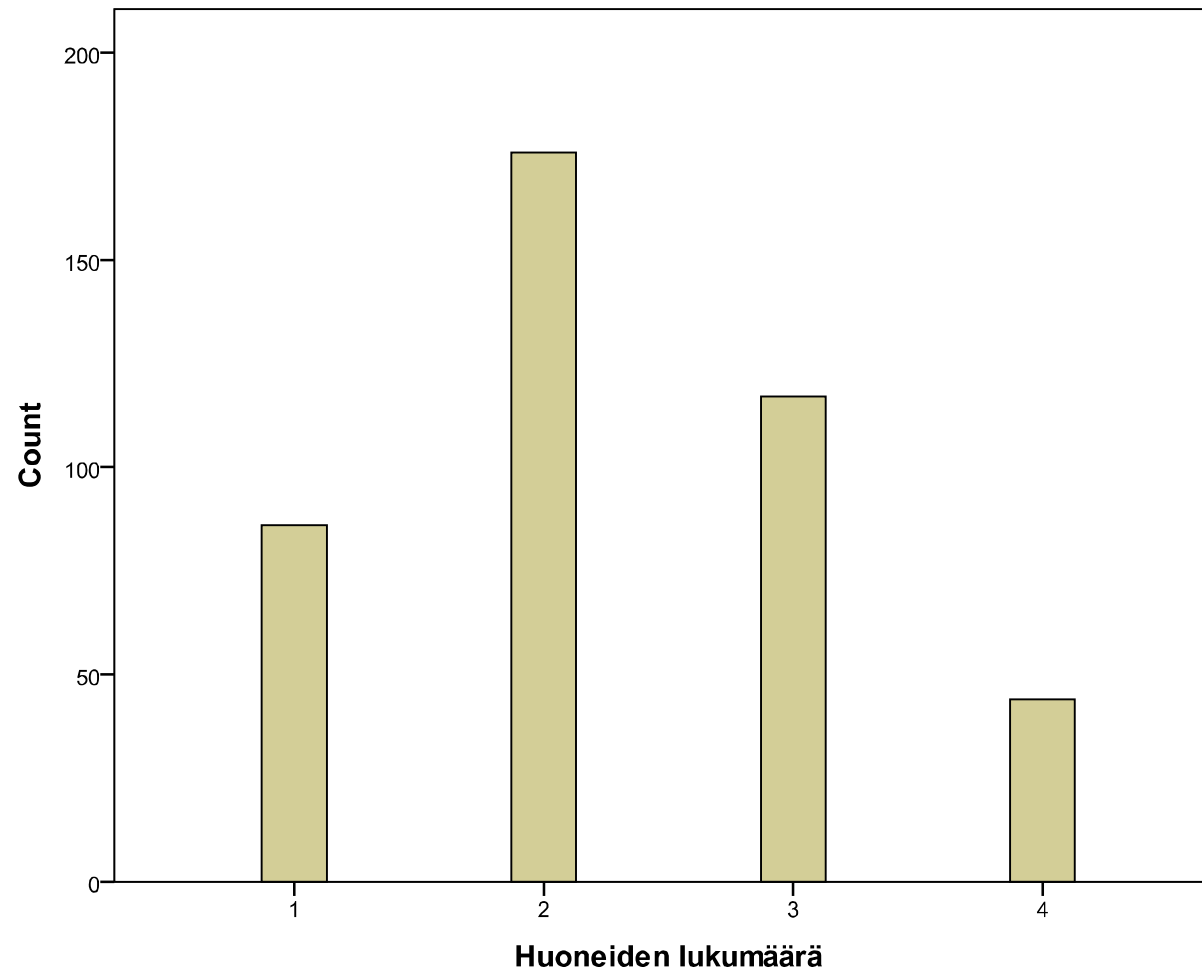
- Sijainnin jakauma, piirakkakuvi



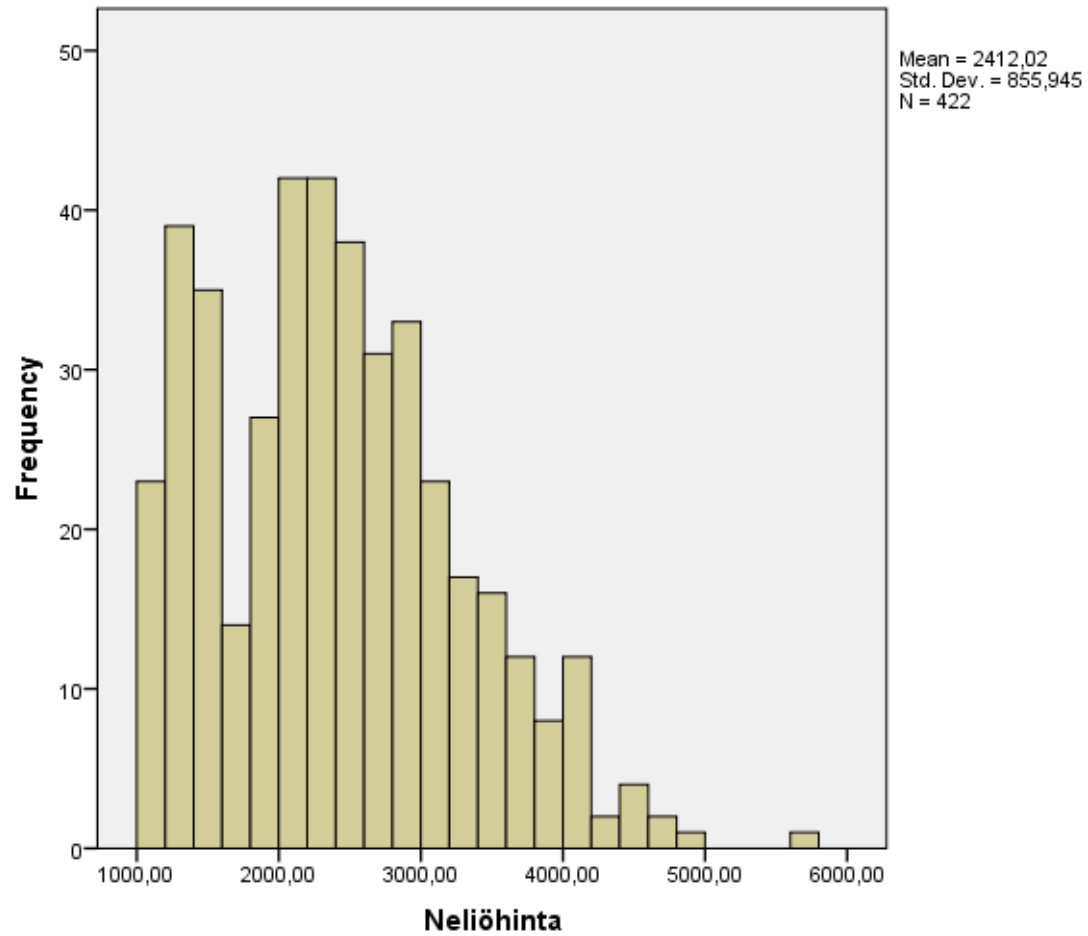
- Sijainnin jakauma, vaakapylväsdiagrammi



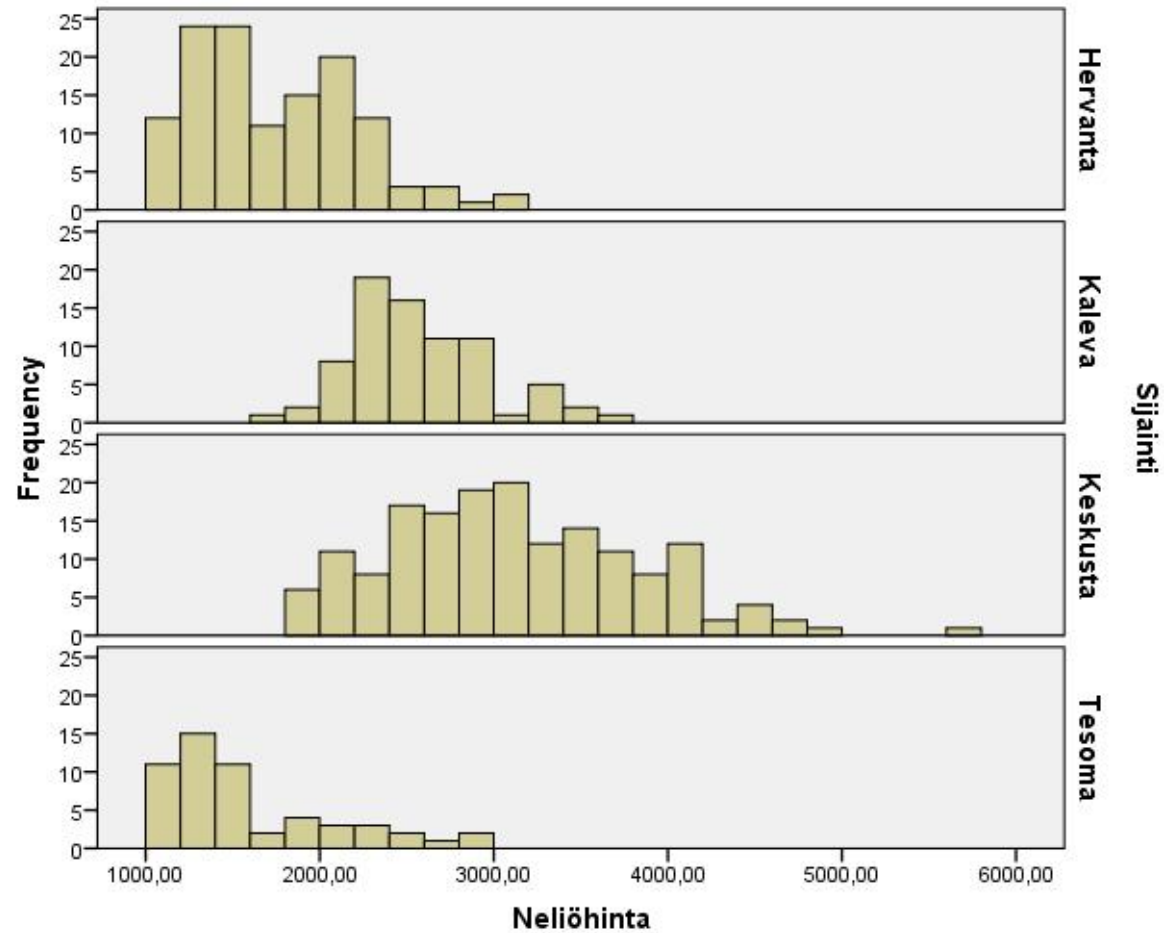
- Huoneiden lukumäärän jakauma, pylväsdiagrammi



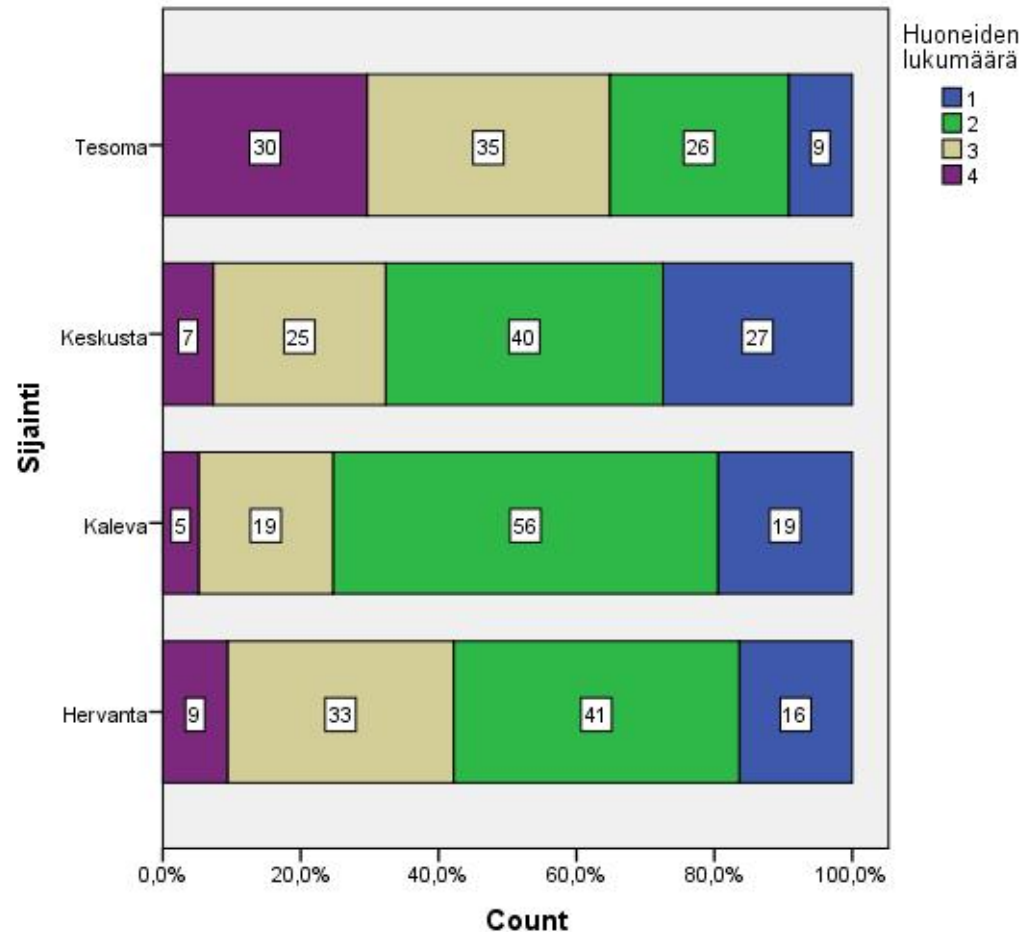
- Neliöhinnan jakauma, frekvenssihistogrammi



- Neliöhinnan ehdolliset histogrammit



- Huoneiden lukumäärä sijainnin mukaan, summapylväsdiagrammi



Grafiikan valinnasta esimerkkejä ks.

http://www.sis.uta.fi/tilasto/mttp1/syksy2013/grafiikan_valinnasta.pdf