ESTATISTIKA ETA DATUEN ANALISIA

II: Aldagai bakunaren tabulazioak eta adierazpide grafikoak

Egilea: Josemari Sarasola

Gizapedia

gizapedia.hirusta.io

2.1 Taulak eta diagramak: zertarako eta nola

2.2 Aldagai kualitatiboaren kasua

2.2.1 Aldagai kualitatiboaren tabulazioa

2.2.2 Aldagai kualitatiboaren adierazpide grafikoa

2.2.2.1 Sektore-diagrama

2.2.2.2 Barra-diagrama

2.3 Aldagai kuantitatiboaren kasua

2.3.1 Aldagai diskretuaren kasua

2.3.1.1 Aldagai diskretuaren tabulazioa

2.3.1.2 Barra-diagrama

2.3.2 Aldagai jarraituaren kasua

2.3.2.1 Datu gutxiko kasua: puntu-diagrama

2.3.2.2 Datu askoko kasua: histograma

2.3.2.3 Histograma tarte-zabalera desberdinekin

2.3.2.4 Ondoko histogramak

2.3.2.5 Maiztasun-poligonoa

2.3.2.6 Maiztasun metatuen histograma eta ojiba

2.3.2.7 Adar-hostoen diagrama

2. gaia: Aldagai bakunaren tabulazioak eta adierazpide grafikoak

2.1 Taulak eta diagramak: zertarako eta nola

Azterketa estatistiko sakonago eta zorrotzagoa egin aurretik, datuak taula batean bildu eta antolatzeak eta datu-multzoa grafikoki adierazteak datuen ezaugarri nagusiak hautematen laguntzen dute. Gainera, bereziki egokiak dira datuak publikoari erakusteko. Horretarako, sinpleak eta argiak izan behar dira. Gogoratu esaldia: *irudi batek mila hitzek baino gehiago balio du.*

2.2 Aldagai kualitatiboaren kasua

18-24 urteko hainbat gazteri udako oporren helmuga galdetu zitzaien 2005 eta 2015 urteetan (Euskal Herria (EH), Europa (EU), beste kontinente bat (K). Hona hemen haien erantzunak:

2005: EU-EU-EH-EH-EH-K-EH-EH-EH-EU-EH-EH-EH-EU-EU-EU-EH-EH-K-K

2015: EU-EU-EU-EH-EU-K-EH-EH-K-EU-EU-EH-EH-EU-EU-K-EH-K-K-K-EU-EU-K-K-EU

Aurreko bi datu-multzoen tabulazioak eta aukerako adierazpide grafikoak eratu behar dira, eta interpretazio egokia egin.

2.2.1 Aldagai kualitatiboaren tabulazioa

2005eko datuen tabulazio osotua osatzeko, zenbaketa arrunta egin behar da:

Opor-tokia	Kopurua	%
Euskal Herria	11	%55
Europa	6	%30
Beste kontinentea	3	%15
Totalak	20	%100

Tabulazio (egin klik) honetan, *aldagai estatistikoa* (egin klik) opor-tokia da. Aldagaiak hartzen dituen balioetako bakoitza (EH, EU, K) *kategoria* edo *modalitate* bat da. Kategorietako gazte-kopuruak *maiztasun absolutuak* (egin klik) dira, n_i izendatuko direnak, eta portzentajeak *maiztasun erlatiboak*, f_i idatziko direnak. Galdekatutako gazte-kopuru totala *lagin-tamaina* da. Era honetako tabulazioei *maiztasun-taula* edo *maiztasun-banaketa* (egin klik) deitzen zaie orokorrean.

2005eko eta 2015eko datuak ditugunez, interesgarriena bi datu-multzoen baterako maiztasun-taula izango da. Kasu horretan, informazio adierazgarriena maiztasun erlatiboek ematen dute, bi datu-multzoak alderatze aldera:

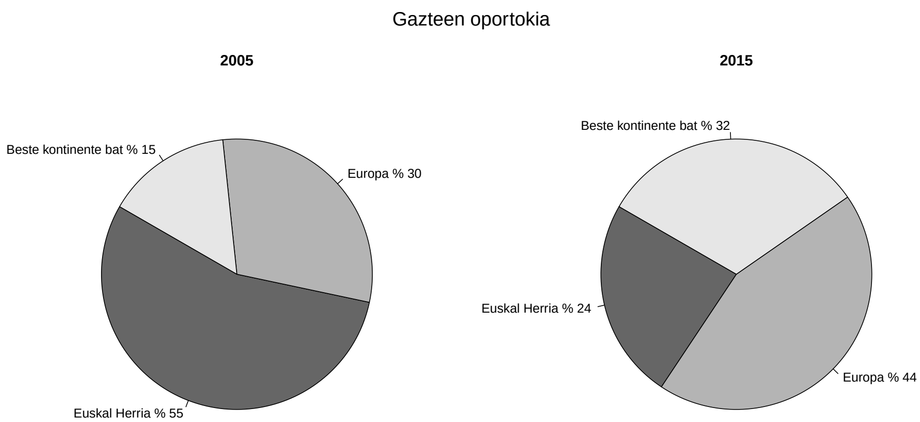
Opor-tokia	Kopuruak (2005)	% (2005)	Kopuruak (2015)	% (2015)
Euskal Herria	11	%55	6	%24
Europa	6	%30	11	%44
Beste kontinente	3	%15	8	%32
Totalak	*20*	*%100*	*25*	*%100*

Hala, hamar urte horietan zehar Euskal Herria oportokitzat hartu dutenen kopurua nabarmen murriztu da, eta Europa edo beste kontinente batera doazenak asko ugaldu, bereziki beste kontinente batera doazenak bikoiztu baino gehiago egin dira.

2.2.2 Aldagai kualitatiboaren adierazpide grafikoak

2.2.2.1 Sektore-diagrama

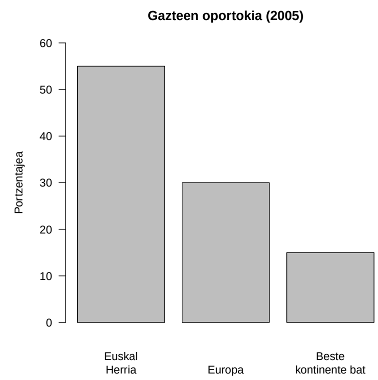
Sektore-diagraman (ingelesez, *pie chart*) kategorien maiztasun absolutuen edo erlatiboen araberako sektoreak margotzen dira zirkulu batean:



Sektore-diagrama kritikatua izan da, sektoreak edo tarta-zatiak begiz alderatzeko zailtasunagatik. Hori dela eta, *barra-diagrama* (hurrengo atalean) hobesten da.

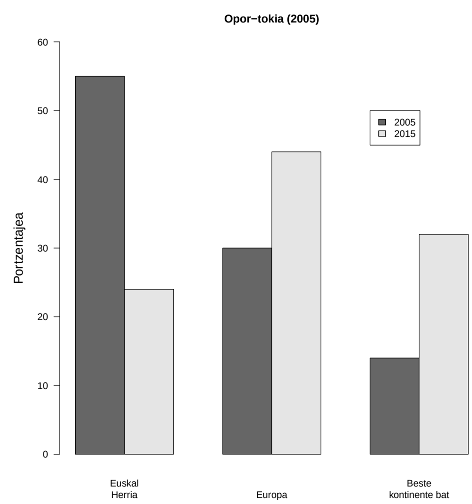
2.2.2.2 Barra-diagrama

Barra-diagraman (zutabe-diagrama ere deitua) kategorien maiztasun absolutuen edo erlatiboen araberako zutabeak eratzen dira. Arestiko datuak harturik, honelakoa litzateke barra-diagrama zutabe bertikalekin:



Barrak horizontalean (eta beraz kategoriak ardatz bertikalean) jartzea komeni da batzuetan, bereziki kategoria kopurua handia denean.

Bi datu-multzoetako kategorien maiztasun erlatiboak alderatzen dituen barra-diagrama da interesgarriena:



Batera jarritako sektore-diagrametan egiten genuenaren interpretazio berdina da, noski: Euskal Herria opor-tokitzat hartzen dutenen kopurua murriztu da. Barra-diagramarekin, ordea, askoz ere garbiago ikusten da, eta interpretazioa zehatzago egin daiteke.

2.3 Aldagai kuantitatiboaren kasua

2.3.1 Aldagai diskretuaren kasua

Aldagai diskretua deituko diogu balio desberdin gutxi hartzen dituen horri; adibidez, aldagai diskretuak dira anai-arreba kopurua familia batean (0, 1, 2, 3; eskuarki) eta ikasle batek ikasturte batean errepikatzen duen irakasgai-kopurua (0, 1, 2, ..., 6).

Hain zuzen, zenbait ikaslek errepikatu duten irakasgai-kopuruari buruzko datuak hartuko ditugu adibidetzat. Hona hemen zerrenda:

0-1-0-2-1-0-3-2-1-2-3-0-1-0-1-1-1-1-2-3-0-0-1-1-2.

2.3.1.1 Aldagai diskretuaren tabulazioa

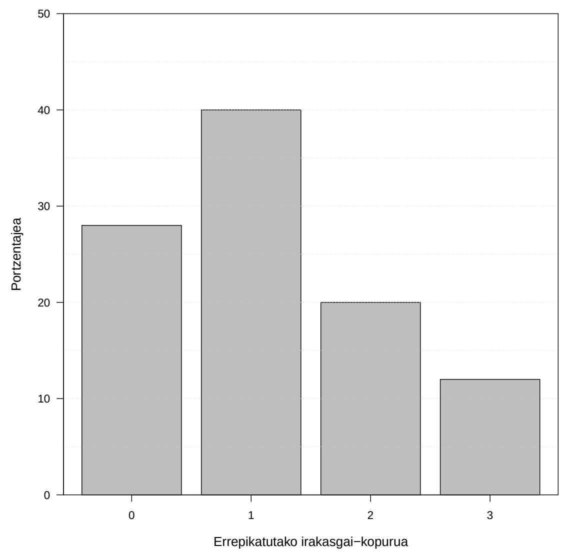
Maiztasun-taula balio bakoitzeko ikasle-kopuruaren zenbaketa eginez abiarazten da, baina kasu honetan aldagaiak (errepikatutako irakasgai-kopurua, adibidean) hartzen dituen balioak era ordenatuan ezartzen dira taulan, txikienetik handienera.

Irakasgai-kopurua (x)	Ikasleak (n)	% (f)	Ikasle metatuak (N)	% (F)
0	7	%28	7	%28
1	10	%40	17	%68
2	5	%20	22	%88
3	3	%12	25	%100
Totalak	*25*	*%100*	*25*	*%100*

Ikusten denez, aldagai kualitatiboetan ez bezala, hemen maiztasun absolutuetan zein erlatiboetan *maiztasun bakunak* eta *metatuak* ere bereizten dira. Hala, *maiztasun erlatibo metatuek*(F), esaterako, balio batetik behera edo balio hori bera duten ikasleen portzentajea jasotzen dute.

2.3.1.2 Barra-diagrama

Barra- edo zutabe-diagrama erabiltzen da, aldagai kualitatiboaren kasuan bezalaxe:



Ikusten denez, ikasle gehienek irakasgai bakarra errepikatzen dute, eta halaber balio horren inguruan biltzen dira errepikatutako irakasgai-kopurua. Beraz, irakasgai bakarra datu-banaketaren zentroa dela esan daiteke.

2.3.2 Aldagai jarraituaren kasua

Aldagai jarraitua deituko diogu balio desberdin asko hartzen dituen horri; adibidez, aldagai diskretuak dira pertsona baten altuera (160, ..., 170, ..., 180, ...), ikasle baten nota (0, 0.1, ..., 4.9, 5, 5.1, ..., 9.9, 10).

2.3.2.1 Datu gutxiko kasua: puntu-diagrama

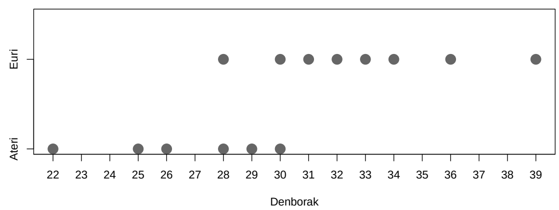
Aldagai jarraitu bati buruz datu gutxi direnean (hitzarmenez, 20-25 baino gutxiago), ez da beharrezkoa datu horiek taula batean biltzea. Grafikoki irudikatzeko, *puntu-diagramak* (ingelesez, *dot plot* edo *strip chart*) erabiltzen dira. Tabulazioa, berriz, ez da beharrezkoa, datu gutxi baitira.

Grafikoa eratzeko adibide gisa, har ditzagun ibilbide bat egiteko behar izan diren denbora batzuk, euripean zein ateri (minututan):

Ateri: 22-26-28-25-29-30

Euri: 28-32-36-39-30-31-33-34

Datu horiek irudikatu eta bide batez bi datu-multzoak alderatzeko puntu-diagramak hauek dira:



Interpretazioa: euria egiten duenean, *orokorrean* denbora luzeagoa behar da ibilbidea egiteko.

2.3.2.2 Datu askoko kasua: histograma

Datuak asko direnean (kopuru bat emateagatik, 20 baino gehiago), egokiena datuak klase-tartetan bildu eta tarte bakoitzean zenbat datu biltzen diren zenbatzea da, datuen egitura nahasiegi azal ez dadin. Tartekako banaketa horretan oinarritua histograma delakoa eratzen da, tarteetan maiztasunen araberako zutabeak altxatuz.

Ikus, gainera, Gizapediko hiztegian, klase-tarte, klase-maiztasun, klase-marka, klase-muga, tartetan bildutako datu-banaketa eta klase-zabalera.

Datuak klase-tartetan biltzerakoan, honako puntu hauek kontuan hartzea komeni da:

- tarte kopuru egokiena 5-15 bitartekoa da;
- aurreko puntuan adierazitako tartearen barruan, ikertzailearen esku geratzen da dezake datuak ikusita kopuru zehatza, baina tarte kopuru zehatza ematen duten erregelak badaude. Erabiliena Sturges-en erregela da, k tarte kopurua eta n datu kopurua izanik, eta emaitza gehiegiz borobilduz:

$$k = \frac{\ln n}{\ln 2} + 1$$

- komeni da hau ere kontuan hartzea: tarte gutxiegi eratzen dira, jatorrizko datuetako informazio asko galtzen da; tarte gehiegi eratzen badira, berriz, histogramaren interpretazioa nahasiagoa izango da;
- komeni da, *printzipioz*, klase-zabalera konstantea izatea;
- klase-mugak edo tarteak mugatzen dituzten balioak *zenbaki biribilak* izatea komeni da (10-20, 100-300, ...);
- hitzarmenez tarteak $[a, b)$ motakoak dira, azkenekoa ezik noski;
- histograma maiztasun absolutu nahiz erlatiboekin era daiteke, bi eratara histogramaren itxura berdina izango da.

Tarte kopurua erabakita, pauso hauek jarraitu daitezke datuak tartetan biltzeko:

1. datuen ibiltartea (datu handiena ken txikiena) kalkulatu;
2. datuen ibiltartea tarte kopuruarekin zatituz, tarte-zabalera ematea;
3. tarte-zabalera gehiegiz borobiltzea;
4. erabakitako tarte-zabalera bider tarte kopurua eginez, tarteek hartuko dute eremua kalkulatzea;
5. eremu hori ibiltartea baino handiagoa izango denez, soberakina hasieran eta bukaeran banatzea

.

Adibidea: Tomate barietate bateko 100 aleren pisuak jaso dira (gramutan):

$$186, 134, 165, 196, 234, 222, 226, 210, 287, 220, 243, 234, 206$$
$$208, 212, 226, 234, 238, 297, 307, 138, 145, 167, 189, 197$$
$$231, 210, 206, 214, 256, 262, 225, 227, 229, 234, 245, 216, 253$$
$$233, 216, 227, 220, 198, 187, 267, 278, 287, 302, 245, 227$$
$$227, 210, 192, 186, 145, 156, 168, 172, 174, 219, 222, 239, 226$$
$$214, 296, 314, 324, 143, 156, 206, 234, 262, 220, 221, 183$$
$$246, 192, 174, 178, 162, 174, 192, 216, 228, 234, 245, 278, 162$$
$$154, 245, 225, 230, 250, 271, 160, 252, 229, 214, 226, 195$$

Datu horietarako tabulazio egoki bat egin eta dagokion histograma marraztu behar da Sturgesen erregela erabiliz. Histograma ikusita, zein balioren inguruan biltzen da tomate baten pisua?

Eman dezagun lehenbizi eratu beharreko tarte-kopurua Sturges-en erregela baliatuz:

$$k = \frac{\ln 100}{\ln 2} + 1 = 7.64 \rightarrow 8$$

Datuen ibiltartea, datu handiena ken txikiena alegia, $324 - 134 = 190$ da. Ibiltartea 8 tartetan zatituz, tarte bakoitzak $190/8 = 23.75$ko zabalera behar duela dakigu.

Tarte-zabalera *zenbaki borobila* behar duenez izan, gehiegiz borobiltzen dugu 25era. Horrela 25 zabalerako 8 tarte izango ditugu, $25 \times 8 = 200$ gramuko ibiltartea estaliko dutenak.

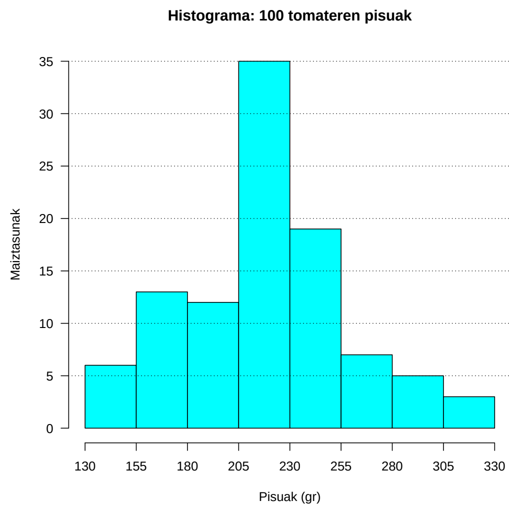
Datuen ibiltartea 190 denez, behar duguna baino gehiago estaltzen dugu 200 gramurekin. Beraz, datu txikienetik beherantz, 130 balioan hasiko gara tartean eratzen; eta 324tik gora, 330 balioan bukatuko dugu. Horrela $330 - 130 = 200$ gramuko ibiltartea izango da.

Tarteak, beraz, 130-155, 155-180, ..., 305-330 izango dira. Tarte bakoitzean zenbat datu biltzen diren zenbatuz, tarteak $[a, b)$ motakoak direla kontuan hartuz betiere, maiztasun-taula eratuko dugu, aukeran maiztasun absolutuekin eta erlatiboekin:

Pisua gramutan (x)	Maizt. abs. (n)	Maizt. erl. (f)
130-155	6	0.06
155-180	13	0.13
180-205	12	0.12
205-230	35	0.35
230-255	19	0.19
255-280	7	0.07
208-305	5	0.05
305-330	3	0.03
	100	1

2.3.2.3 Histograma tarte-zabalera desberdinekin

Batzuetan komenigarriagoa da tarte-zabalera desberdina izatea, horrekin maiztasun-taulak emango duen informazioa zehatzagoa bada. Adibidez, datu berdinekin eratutako bi maiztasun-tauletan tarte-zabalera desberdineko maiztasun-taulak informazio zehatzagoa eta aberatsagoa ematen du tarte-zabalera konstantekoak baino:



Interpretazioa: Begi bistan, tomateen pisua gutxi gorabehera 230 gr inguruan biltzen dela esan daiteke. Ezaugarri estatistiko horri *zentroa* deritzo.

1. taula: tarte-zabalera berdina.

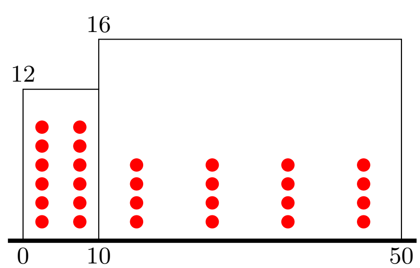
Tarteak	Maiztasunak (n)
0-20	152
20-40	28
40-60	12
60-80	6
80-100	2
	200

2. taula: tarte-zabalera desberdina.

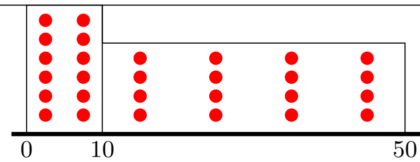
Tarteak	Maiztasunak (n)
0-5	36
5-10	64
10-20	52
20-50	34
50-100	14
	200

Tarte-zabalerak desberdinak direnean, ordea, histogramako zutabeak maiztasunen neurrian altxatzea ez da egokia. Adibidez, jarraian tarte-zabalera desberdineko maiztasun-taula, maiztasunen arabera legokiokeen histograma, oker egiten ari garela azaltzeko histogramaren barruan banakako datuak ere adierazita, eta histograma zuzena nolakoa izan beharko litzatekeen azaltzen dira:

Tarteak	Maiztasunak (n)
0-10	12
10-50	16
	28



Histograma okerra: zutabeek ez dute adierazten lehen tartean datu-dentsitatea handiagoa dela (datuak puntu gorriz adierazita), eta horixe da histograma batean azalarazi beharrekoa.



Histograma zuzena: zutabeek zehatz-mehatz adierazten dute datu-dentsitatea.

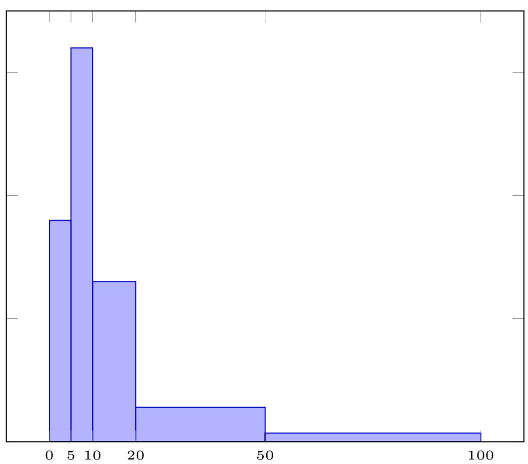
Tarte-zabalerak berdinak direnean, ez da beharrezkoa datu-dentsitateak kalkulatzea, maiztasunen baliokideak direnez. Tarte-zabalerak desberdinak direnean, ordea, honela kalkulatu behar dira dentsitateak eta, bidenabar, histogramako zutabe-altuerak, f izanik tarteko maiztasun erlatiboa, eta z tarteko zabalera:

$$d = \frac{f}{z}$$

Adibidez, aurreko adibidean:

Tarteak	Maiztasunak (n)	Maizt. erlatiboa (f)	Dentsitatea (z)
0-5	36	18	18/5=3.6
5-10	64	32	32/5=6.4
10-20	52	26	26/10=2.6
20-50	34	17	17/30=0.56
50-100	14	7	7/50=0.14
	200	100	

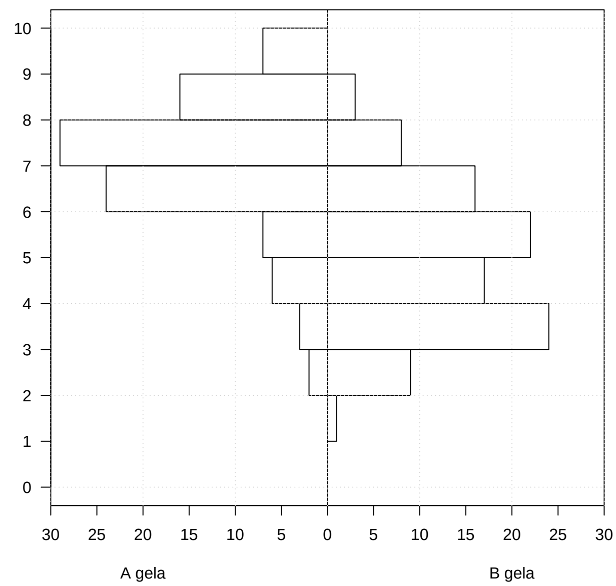
Eta dagokion histograma:



Interpretazioa: datu-dentsitate handiena 5-10 tartean suertatzen da.

2.3.2.4 Ondoko histogramak

Elkarren ondoko histogramak (ingelesez, *back-to-back histograms*) bi datu multzo alderatzeko erabiltzen dira; alderatzeko, adibidez, bi ikasgelako notak, bi sexuen pisuak, A eta B makinetako ekoizpenak eta bi enpresetako langileen adinak. Horien helburua bi datu-multzoak begi bistan aise alderatzea da.

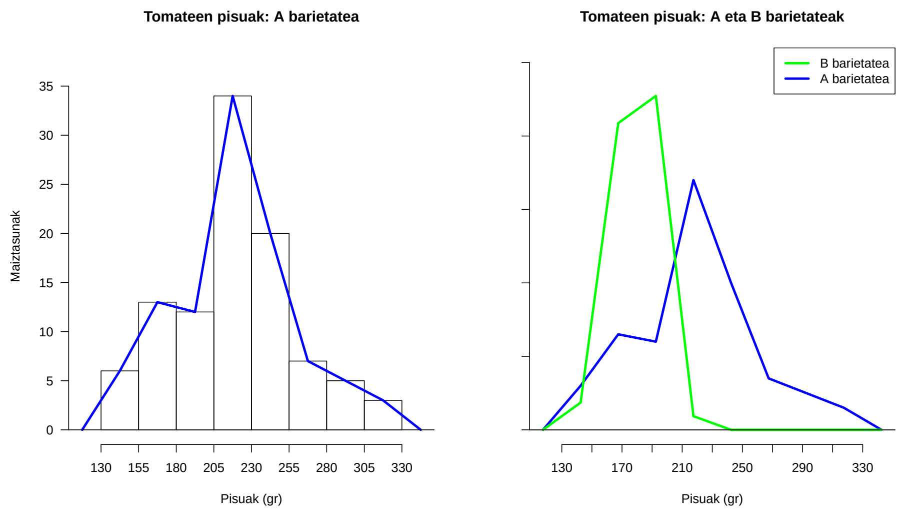


Ondoko histogramak: *Zentroari dagokionean*, A gelako notak handiagoak dira orokorrean B gelakoak baino: A gelan notak gutxi gorabehera 8 punturen inguruan biltzen diren bitartean, B gelan 5 punturen inguruan biltzen dira. *Sakabanatzeari dagokionean*, berriz, B gelako notak zertxobait sakabanatuagoak dira itxuraz: 8 punturen ibiltartean zehar biltzen dira B gelakoak, A gelakoak 7 punturen ibiltartean biltzen diren bitartean.

Komenigarria da horietan zutabeen neurria emateko *maiztasun erlatiboak* (edo dentsitateak, tarte-zabalera desberdina denean) erabiltzea, bi datu-multzoetako lagin-tamaina edo datu kopuru desberdinen eragina saihesteko. Halaber, behar bezala alderatzeko, bi datu-multzoetan erabilitako tarteak berdinak izan behar dira.

2.3.2.5 Maiztasun-poligonoa

Maiztasun-poligonoa ondorengo irudietan erakusten den bezala eraikitzen da histograman oinarrituta, histograma zutabe-goietako erdipuntuak lotuz. Histograma bezalatsu interpretatzen da zentroari, sakabanatzeari eta beste ezaugarri estatistikoei buruz, baina batzuetan ezaugarri horiek argiago agertzen da maiztasun-poligonoarekin. Maiztasun-poligonoak grafiko berean bata bestaren gainean eratuz, kolore desberdinekin, datu-multzo zenbait aldera daitezke batera.

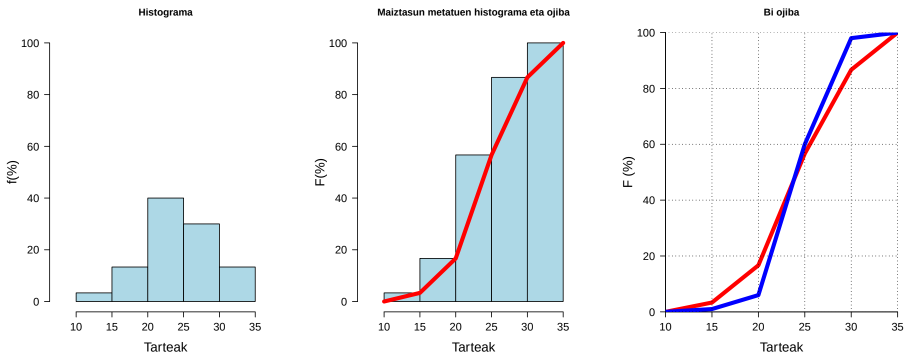


Ezkerrean, maiztasun-poligonoaren eraketa azaltzen da. **Eskuinean**, bi maiztasun-poligono batera eratu dira, horien ezaugarriak alderatzearren: B barietateko tomateen pisuak 180gr inguruan biltzen dira (zentroa), eta A bariedadekoak sakabanatuagoak dira.

2.3.2.6 Maiztasun metatuen histograma eta ojiba

Maiztasun metatuen histograma maiztasun metatuekin eratzen den histograma besterik ez da, tarte bakoitzean eta hortik behera zenbat datu dauden adierazten duena. Zutabeetako gailurrak lotuz, ojiba delakoa eskuratzen da, eta beste ojiba batzuekin batera marraztuta, datu-multzo desberdinetako maiztasun metatuen bilakaera aldera daiteke, maiztasun metatuen histogramekin baino aiseago. Adibidez, maiztasun-taula hau harturik:

Tarteak	Maiztasunak (n)	Maizt. erlatiboa (f,%)	Maiztasun metatua (F,%)
10-15	1	3.33	3.33
15-20	4	13.33	16.66
20-25	12	40.00	56.66
25-30	9	30.00	86.66
30-35	4	13.33	100.00
	30	100.00	



Ezkerrean, histograma. **Erdian**, dagokion *maiztasun metatuen histograma*, maiztasun metatuekin, kasu honetan erlatiboekin, eratua, eta dagokion *ojiba*. **Eskuinean**, bi ojiba batera: 20-30 bitartean, ojiba urdinak malda handiagoa duenez, dagokion datu-multzoak tarte horretan erlatiboki datu gehiago biltzen ditu.

2.3.2.7 Adar-hostoen diagrama

Aldagai jarraitu batean, jatorrizko datuak galdu gabe histogramaren antzeko diagrama bat eratzeko aukera da adar-hostoen diagrama (ingelesez, *stem-and-leaf plot*). Adar-hostoen diagramak badu abantaila bat histogramaren bezala: hark ez bezala, adar-hostoen diagramak jatorrizko datuen informazio osoa atxikitzen da orokorrean.

Adar-hostoen diagraman, datuek hartzen dituzten balioak bi zatitan banatzen dira: adarra balio horien lehen zifra adierazgarriek osatuko dute, eta hostoak azkeneko haiek.

Adibidea: Honako datuok harturik, adar-hostoen diagrama eratu behar da:
8.131 - 4.160 - 8.359 - 7.781 - 6.040 - 7.895- 10.107 - 6.810 - 4.758 - 7.517
4.455 - 5.358 - 5.397 - 7.436 - 3.602 - 5.769 - 10.312 - 8.315 - 9.183 - 10.465
9.058 - 8.052 - 7.527 - 9.580 - 9.279 - 11.203 - 9.824 - 4.502 - 9.097 - 6.302

```
3 | 6
4 | 2558
5 | 448
6 | 038
7 | 45589
8 | 1134
9 | 112368
10| 135
11| 2

gakoa:  3|6=3.6
adar-unitatea:  1
hosto-unitatea:  0.1
```

Ohartu behar da hosto-unitatea 0.1 denez, datuak dezimal gertuenera borobildu direla; adibidez, 4.160 → 4.2.

Aukeran, eskala zabaldu egin daiteke, bi dezimal sartuz hostoetan:

```
3 | 60
4 | 16455076
5 | 364077
6 | 043081
7 | 4452537889
```

```
8 | 05133236
9 | 061018285883
10| 113147
11| 20

gakoa:  3|60=3.60
adar-unitatea:  1
hosto-unitatea:  0.01
```

Eta eskala murriztu ere egin daiteke, adar-unitatea 2-ra aldatuz:

```
3 | 62558
5 | 448038
7 | 455891134
9 | 112368135
11| 2

gakoa:  3|6=3.6
adar-unitatea:  2
hosto-unitatea:  0.1
```

Horrela, ordea, ez dakigu datu handiena 112 edo 122 den.

Adar-hostoen diagrama-aldaera ohikoenak eman ditugu, baina beste hainbat daude aukeran.

2.4 Ariketak

1. Eskola bateko 4 urteko haurren artean, amaren ikasketa maila (baxua, b/ertaina, e/altua, a), matematika test batean lortu duten puntuazioa, eskola izaera (publikoa/pribatua) eta anai-arreba kopurua jaso dira:

b-23-pub-0	b-19-pub-1	a-20-pub-1
e-22-prib-1	e-24-pub-0	a-23-prib-1
e-25-prib-1	a-25-pub-1	b-21-pub-1
a-28-pub-2	e-23-prib-0	a-26-prib-1
a-27-prib-1	b-23-pub-1	b-27-prib-0
e-20-prib-0	e-26-pub-1	e-22-prib-1
b-20-prib-0	a-25-prib-0	b-23-prib-0
a-27-pub-2	b-18-pub-0	b-30-pub-1
a-24-prib-0	e-22-pub-0	a-29-prib-1

   (a) Amaren heziketa-mailaren arabera, eskola publikora edo pribatura joateko joera dagoen aztertu behar da, tabulazio eta diagrama egokien bitartez. Adierazi non dauden aldagai estatistikoa, kategoriak edo modalitateak, lagin tamaina, maiztasun-banaketa eta tabulazioa.

   (b) Matematika-trebetasuna aztertu behar da, amaren heziketa mailaren, eskola-motaren eta anai-arreba kopuruaren arabera.

   (c) Anai-arreba kopurua aztertu eskolaren eta amaren heziketa-mailaren arabera.

   (d) Adar-hostoen diagramaz irudikatu matematikako puntuazioa adar-unitatetzat 10, 5 eta 2 hartuz, hurrenik hurren.

2. 30 piezetan neurketa-errore hauek jaso dira bi gailu desberdin erabiliz (mm):

A gailua	10.26	3.73	0.18
	0.81	6.07	0.20
	0.13	0.38	1.98
	2.83	1.27	1.47
	0.07	3.11	0.42
	0.94	1.76	1.08
	0.17	6.88	3.11
	1.55	0.06	6.62
	0.84	3.82	4.39
	1.90	7.58	0.53
B gailua	5.64	3.69	3.89
	4.89	2.37	4.09
	1.08	0.64	3.08
	3.82	1.96	1.28
	4.57	2.34	0.37
	4.61	0.68	0.06
	3.91	3.35	0.92
	3.23	0.80	2.20
	2.35	3.69	2.28
	2.98	4.33	3.83

(a) A gailuko datuetarako tabulazio eta grafiko egokia eratu Sturges-en erregela baliatuz. Grafikorako maiztasun erlatiboak baliatuz gero desberdina aterako al litzateke?

(b) A gailuko datuetarako, 0-0.5, 0.5-1, 1-3, 3-6, 6-12 tartekako datu-tabulazioa osatu eta dagokion histograma egin.

(c) 1 zabalerako tarteekin tabulazioak eta histogramak eta maiztasun-poligonoak eratu A eta B gailuetako datuetarako, era bereizian. Beharrezkoa al da maiztasun erlatiboak kalkulatzea?

   i. Aztertu zein gailuk ematen duen errore handiena orokorrean.
   ii. Zein gailuk ematen du errore egonkorrena?
   iii. Erabaki zein den gailu egokiena neurketak egiteko.

---

3. 30 urteko 50 gizonen pisuak jaso dira (kg):

$$67-66-58-80-89-59-71-66-79-63-63-83$$

$$92-66-57-55-73-60-58-74-61-80-77-75-73$$

$$76-64-73-72-71-74-52-82-64-67-75-59$$

$$53-70-82-85-80-74-75-64-75-67-79-73-67$$

(a) Pisuen maiztasun-taula eta histograma eratu behar dira, Sturges-en erregela erabiliz, maiztasun mota esanguratsuenekin. Emaitzak interpretatu.

(b) Maiztasun metatuen histograma eta ojiba era itzazu. Ojibaren kasuan, zergatik da malda handiagoa 70 balioaren inguruan?

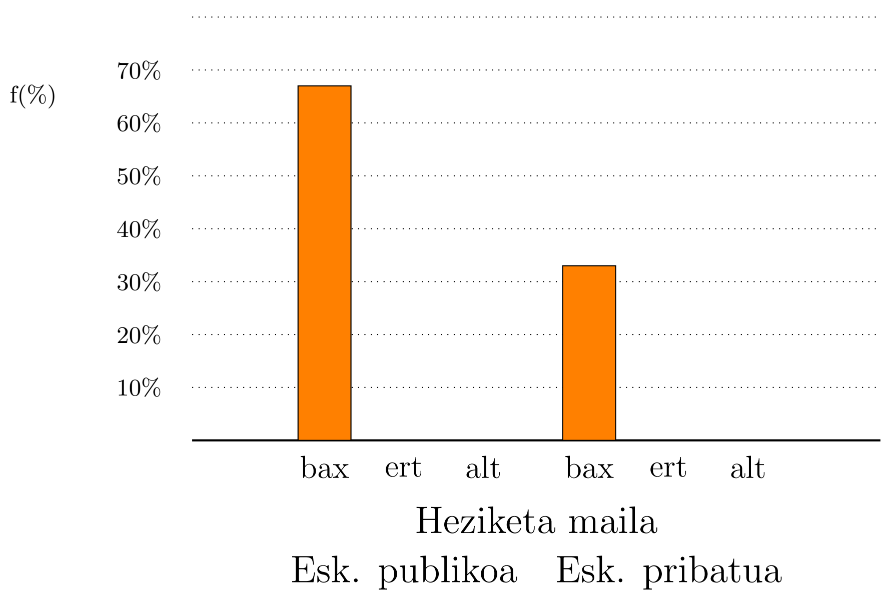
$\boxed{1a}$

AMAREN HEZIKETA MAILAREN ARABERA DATUAK BEREIZI:

Baxua	Ertaina	Altua
(1) pub	(2) prib	
	(3) prib	

Amaren heziketa maila

	Baxua		Ertaina		Altua	
Eskola-izaera	Haurrak (n)	% (f)	Haurrak (n)	%(f)	Haurrak (n)	%(f)
Publikoa	6	%67				
Pribatua	3	%33				



$\boxed{\text{1b}}$

- aldagai estatistikoa:

$\longrightarrow$
$\longrightarrow$
$\longrightarrow$

$\boxed{\text{amaren heziketa mailaren arabera}}$

Altua ....

Ertaina ....

Baxua ....

18 19 20 21 22 23 24 25 26 27 28 29 30

Puntuazioa

$\boxed{\text{eskola-motaren arabera}}$

Pribatua ....

Publikoa ....

18 19 20 21 22 23 24 25 26 27 28 29 30

Puntuazioa

$\boxed{\text{anai-arreba kopuruaren arabera}}$

2 ....

1 ....

0 ....

18 19 20 21 22 23 24 25 26 27 28 29 30

Puntuazioa

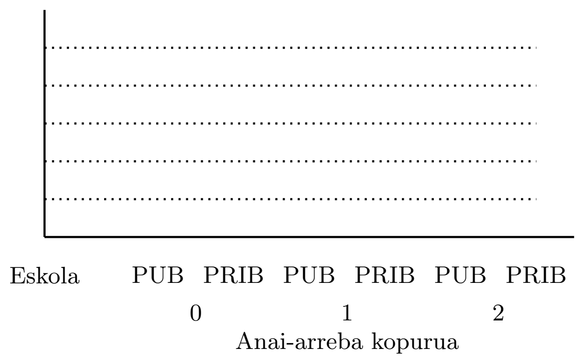
1c

- aldagai estatistikoa:

$\longrightarrow$

$\longrightarrow$

	Eskola				Amaren heziketa maila					
	Publikoa		Pribatua		Baxua		Ertaina		Altua	
Anai-arrebak	n	f(%)								
0										
1										
2										



Eskola _______________ doazen haurrek anai arreba gehiago izateko joera dute.

Heziketa maila _______________ amek haur gehiago izateko joera dute.

1d

```
1 |
2 |
3 |

gakoa:  2|3=23
adar-unitatea:  10
hosto-unitatea:  1
```

```
1 |
2 |
2 |
3 |

gakoa:  2|3=23
adar-unitatea:  5
hosto-unitatea:  1
```

```
1 |
2 |
2 |
2 |
2 |
2 |
3 |

gakoa:  2|3=23
adar-unitatea:  2
hosto-unitatea:  1
```

Tarteak (adarrak) gehitu ahala informazio finagoa eskuratzen dugu maiztasun-banaketari buruz.

$\boxed{\text{2a}}$

- aldagai estatistikoa:

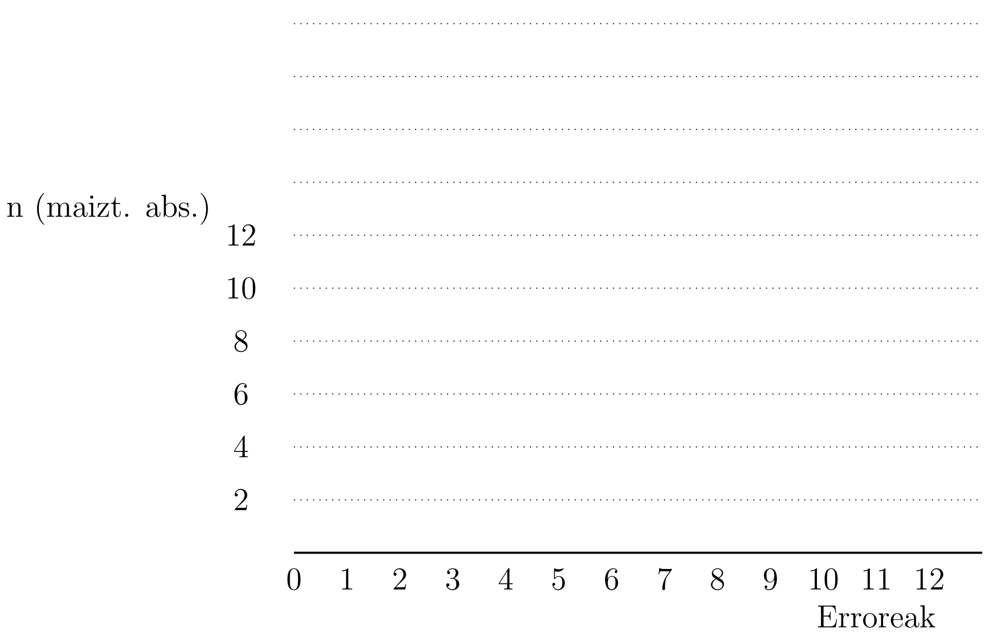
$\longrightarrow$

$\longrightarrow$

$\longrightarrow$

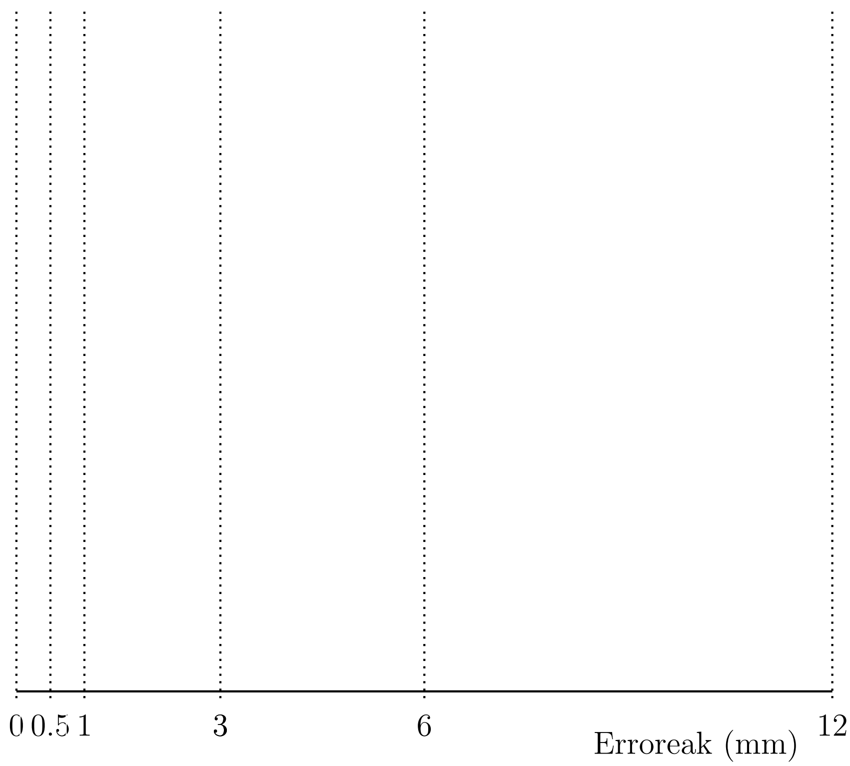
Abiaburua: tarte kopurua: Sturges-en erregela: $k = \dfrac{\ln n}{\ln 2} + 1 =$

Erroreak	Zenbaketa	Piezak n



2b

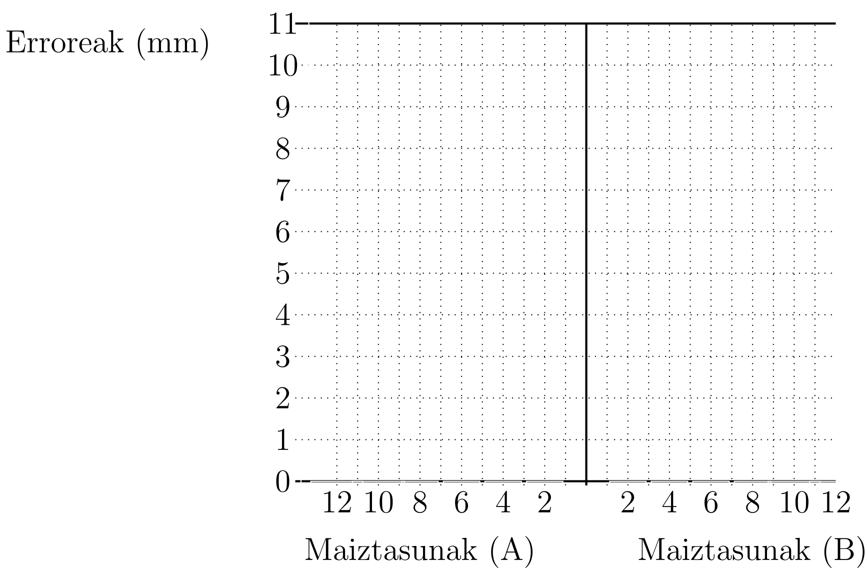
Erroreak	Zenbaketa	Piezak n	% f	Dentsitatea
0-0.5	$\vert\vert$			
0.5-1				
1-3				
3-6				
6-12	$\cancel{\vert\vert\vert\vert}$	5		
Totalak		30	100	



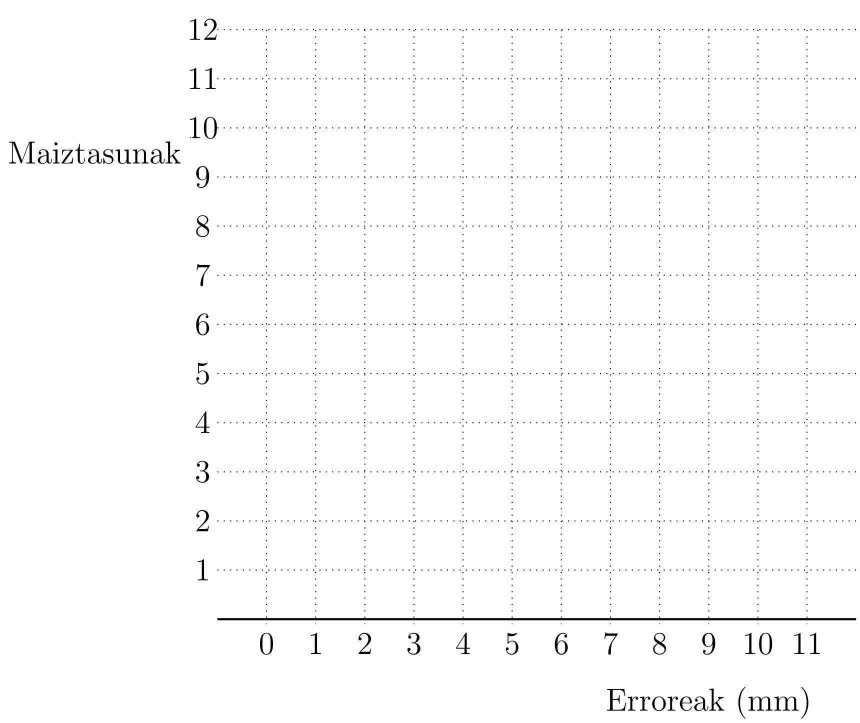
2c

Erroreak	Zenbaketa		Piezak (n)		Portzent. (f)	
	A	B	A	B	A	B
0-1	Ez da	behar	12	6	Ez da	behar
1-2			7			
2-3						
3-4						
4-5						
5-6						
6-7						
7-8						
8-9						
9-10						
10-11						

Ondoko histogramak egiteko:



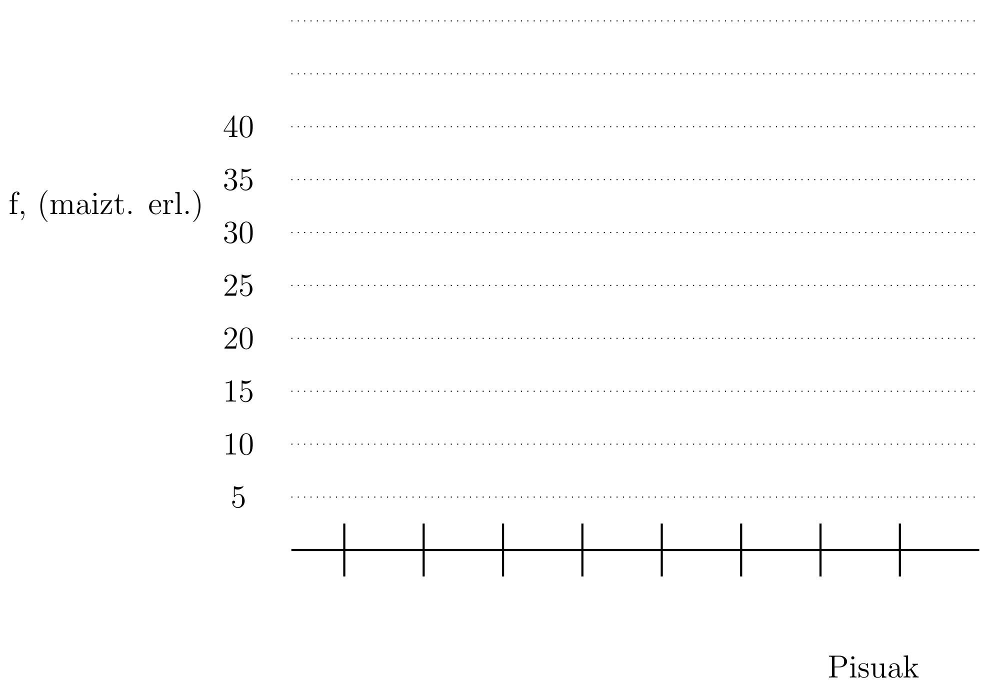
Maiztasun poligonoak marrazteko:



$\boxed{\text{3a}}$ Abiaburua: tarte kopurua: Sturges-en erregela: $k = \dfrac{\ln n}{\ln 2} + 1 =$

Pisuak (kg)	Zenbaketa	Gizonak n	f (%)	F (%)

Histograma:



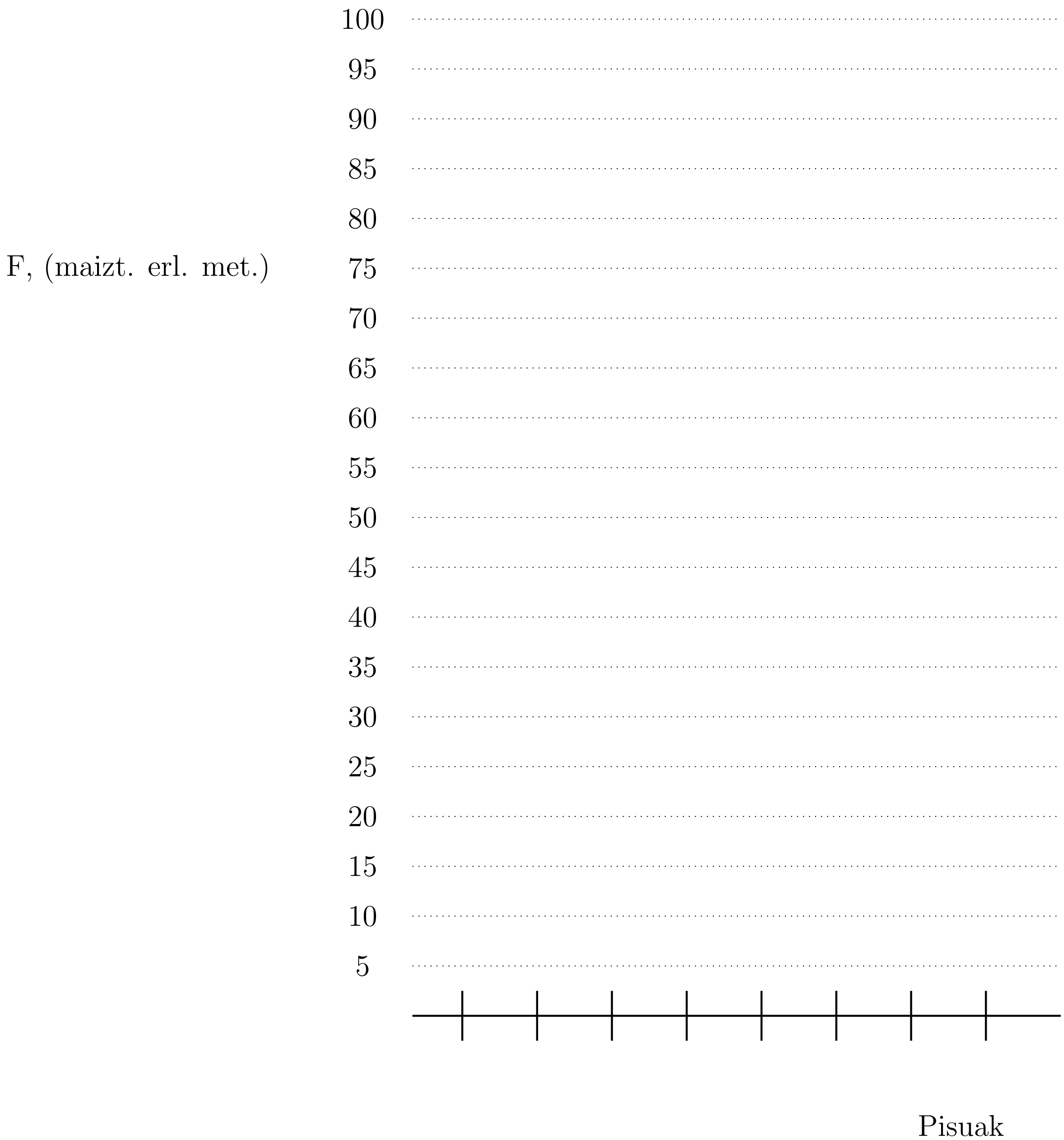
Interpretazioa:

- zentro-joerari buruz,

- sakabanatzeari buruz,

3(b)

Maiztasun metatuen histograma eta ojiba:



Zergatik da malda handiagoa 70 balioaren inguruan?

Horren inguruan datu-maiztasun handiagoa dagoelako.