

1. atala. Estatistikaren zergatiak eta nolakoak.. 3	
■ Estatistikak ezaguera sortzen laguntzen digu.....	2
■ Estatistika, errealitatearen ezaugarriak ezagutzeko tresna.....	12
Estatistikaren aztergaiak: errealitatearen egoera aldakorrak.....	12
Espainiako etxebizitzaren prezioa	20
■ Estatistika eta ikerketa estatistikoen ahalmena	24
Europar Batasuneko nazioek hezkuntzara bideratzen dituzten baliabideak.....	24
Estatistika eta administrazioaren erabakiak. Mamografiak hautatuentzat soilik.....	28
2. atala. Estatistikaren erabilera aldagaien aldakortasuna aztertzeko..... 31	
■ Datuen lehenengo miaketa	32
Datuak, aldagaiak, populazioak	32
Zeri deitzen diogu datuen lehenengo miaketa? Zertan datza?	33
Aldagaiaren aldakortasuna ezagutzeko lehen pausoa: ibiltartearen kalkulua	34
Elementu arraroak eta egon daitezkeen hutsegiteak	35
■ Balioen barne-aldakortasunaren analisisa.....	40
Aldagaiaren barne-banaketaren kontzeptua.....	40
Balio-banaketaren analisisa.....	45
Maiztasun-banaketaren IRUDIkapen grafikoa: Histograma.....	51
■ Nola azaldu eta deskribatu aldagaien aldakortasuna.....	56
Tresna deskribatzaileak aukeratzeko irizpideak.....	59
■ Kanpai-itxura duten banaketen analisisa	62
Kanpai-itxurako banaketen erdigunea. batezbestekoa eta mediana	63
Notazio estatistikoaren oinarriak	64
Batezbestekoaren erabilera.....	67
Banaketen erdigunea finkatzeko beste neurria: mediana.....	70

Elementu arraroren bat duten banaketen batezbestekoa	74
Banaketaren maiztasun gehieneko tartea: desbidERATZE estandarra	80
Maiztasun gehieneko tartea.....	80
Desbideratze estandarraren kalkulua	82
Desbideratze estandarraren erabilera	89
Aldakuntza-koefizientea.....	95
Simetriaren analisisa	99
Simetriaren analisi grafikoa.....	100
Asimetria-analisiaren ekarpena.....	107
Asimetria-mailaren kuantifikazioa.....	109
Balioen taldekatze-mailaren kuantifikazioa: kurtosiA	111
Kanpai-itxurako banaketak deskribatZeko neurrien erabilera elkartua	119
Hezetasun desberdineko behatokiak.....	123
Aldakortasun handiko prezipitazio kopururak.....	124
Eranskinak	128
1. taula. Pertsonako errenta erabilgarria. 2001 eta 2003.....	129
2. taula. Espainiako hainbat hiriren etxebizitzaren prezioa. Etxebizitza berriei buruzko batez besteko datuak. 2007	136
3. taula. Zaragoza probintziako hirietako etxebizitzaren prezioa. 2007	141
4. taula. Toledoko eta San Fernandoko (Cádiz) urteko prezipitazioak. 1910-1995.....	143
5. taula. Pertsonako batez besteko gastu osoa. 2006	146
6. taula. Pertsonako batez besteko gastu osoa	147
7. taula. Pertsonako batez besteko gastu osoa	148
8. taula. Nazioarteko analfabetismoari buruzko datuak. 2006.....	149



*1. atala. Estatistikaren
zergatiak eta nolakoak*

ESTATISTIKAK EZAGUERA SORTZEN LAGUNTZEN DIGU

Estatistika zer den eta zertan lagundu diezagukeen jakiteko hoberena izango da haren laguntzaren beharra izango genukeen egoera batean murgiltzea. Horretarako, lehenik esan beharra dago Estatistikak ezaguera lortzeko baliabideak eskaintzen dizkigula, munduari eta errealitateari buruz ezagutzen ez diren gauzak jakiteko. Mendeetan gizateriak eskuratu dituen ezaguera moten kopurua ikusgarria izan arren, are luzeagoa da ezagutzen ez diren gauzen zerrenda.

Egia esanda, gero eta ohikoagoa dugu aurkikuntzei buruzko berriak zabalkunde handiko komunikabideetan agertzea. Garrantzi handiko gaiak dira batzuetan, eta huskeriak edo bitxikeriak, besterik ez, besteetan. Ikertzeari eta ezaguerari ematen zaien garrantzia da gure mahai gainean dagoena. Horren adibide esanguratsuak aipa daitezke.

1996ko apirilaren 24an *El Mundo* egunkariak albiste bat zabaldu zuen, urdaiazpikoa jatearen abantailei buruzkoa. Egunkariak zioenez, estatistika-metodoetan oinarrituta zegoen ikerketa baten emaitzak ziren.

Kolesterolaren kontra ezkur-txerria. Ezkur-urdaiazpikoak osasunarentzat onuragarria dela erakutsi digu ikerketa batek.

Ehun eta hemeretzi urdaiazpiko, zortzi ikertzaile eta Badajozeko herri bateko biztanle talde baten jateko gogo behar izan dira gustuak diosenari arrazoia emateko. Osasunarentzat ezkur-urdaiazpikoa, iberiarra, ona da.

Aukeratutako lekua Badajozeko Valdebotoa herria izan da. Han, hemeretzi emakumek, 66 urte ingurukoek, onartu zuten Extremadurako Unibertsitateak, Badajozeko Ospitaleko Barne Medikuntzako Zerbitzuak eta Extremadurako Juntako Nekazaritza Sailak antolatutako ikerketa batean parte hartzea.

***El Mundo* egunkaria, 1996ko apirilaren 24a.**

2004ko urrian, gizadiarentzat askozaz garrantzitsuagoa zen ikerketa baten emaitza ezagutu genuen:

Malariaren kontrako txerto berriak hiltzeko arriskuan dauden jaioberrien % 65 infekzioetatik babesten ditu. Mozambiken lan egiten duen Pedro Alonso espainiarrak egindako saiakuntza klinikoak Afrikako eskala handiko immunizaziorako bidea prestatzen du.

***ABC* egunkaria. 2007ko urriaren 18a.**

Hiru urte geroago, 2007an, zortzi urteko ikerketa sakonak behar izan zituzten Estatu Batuetako ikertzaileek, gizonezkoek emakumezkoek beste berba egiten dutela jakiteko.

Arizonan dagoen Tucson-eko Unibertsitateko ikertzaileek metodo zientifikoa erabili dute, emakumezkoek gizonezkoek baino gehiago hitz egiten dutela dioen estereotipoa aztertzeko. [...]

Hainbat egunetan ehunka pertsonak izaniko elkarrizketa naturalak grabatu ziren ikerketan. Horren emaitzen arabera, emakumezkoen eta gizonezkoen arteko aldea, estatistikaren ikuspuntutik, ez da esanguratsua: emakumezkoek 16.215 hitz inguru erabiltzen dituzte egunero, eta gizonezkoek, 15.669 inguru.

***El País* egunkaria. 2007ko uztailaren 7a.**

Egunero gauza berriak argitzen dira, eta mota askotako aurkikuntzak ere egiten dira. Guztien garrantzia eta oihartzuna berdina ez bada ere, guztiek dute antzekotasun bat: ikerketa-prozesuen bidez (funtsezko ezaugarri bertsuak dituzten ikerketen bidez, alegia) lortu izana, hain zuzen ere.

Erantzun nahi diren galderetatik abiatzen dira ikerketak; galdera horiek zerbait ezagutzeko grinak bultzatzen ditu. Horrela hasten dira gizakiok egiten ditugun gauza asko ere, jakin-minak bultzatuta edo batzuetan gure lanaren beharrak behartuta.

Honaino iritsita, jarduera zientifikoek eta profesionalek egin ohi dituzten jarduera intelektualen arteko paralelismo bat dagoela esan dezakegu, bietan zeregin antzekoak izaten dira eta. Hezkuntza eta teknologia digitalari buruzko gaiak jorratzen dituen artikulu batean, egileak, Ferran Ruiz Tarragók, aipaturiko paralelismoaren funtsezko ezaugarriak laburbiltzen ditu :

Pertsona guztiek antolatu eta kontrolatu behar dute beren lana; pertsona guztiek jasotzen, gordetzen eta bilatzen dute informazioa, hainbat modutara; entzunez, irakurriz eta behatuz jasotzen dugu denok informazioa, baita beste zentzumenen bidez ere; jarduera profesional guztietan ekoizten da informazioa, ahozkoa, idatzia, edo beste mota batekoa; pertsona guztiek elkarri eragiten diote, eta beste pertsonekin eta beren inguruarekin komunikatu egiten dira.

Informazioarekin egiten den buru-lana da jarduera horien guztien oinarri komuna, informazioaren kodifikazioa, formatua edo euskarriak albo batera utzita.¹

Ikusten dugunez, Ruiz Tarragók parekatzen ditu jarduera zientifikoaren ohiko zereginak eta gizakiarentzat ere ohikoak diren buru-lanak. Azken batean, ikertzaile denak, baita ez denak ere, informazioarekin lan egiten du; informazioa bilatzen, prozesatzen, ordenatzen du... Informazioarekin egiten den lanaren norainokoa eta, batez ere, informazioa kudeatzeko erabiltzen den metodoa dira jarduera zientifikoaren bereizgarriak. Baina, edonola ere, lanaren testuingurua edozein dela ere, ezaguera sortzeko prozesuak metodoa eskatzen du: zerbait ezagutzen ez dugunean eta hura ezagutzeko bideari ekin nahi badiogu, metodoa erabili beharko dugu –prozesuan zehar gidatuko gaituen prozedura-segida, alegia–.


Ezagueraren bilaketa hasten da, beraz, ikertzaileek, profesionalek edo


¹ Ferran Ruiz Tarragó (1999): «Trabajo intelectual, información y tecnología digital». *Educar* 25, 32. or.


partikularrek arazo bat daukatenean, ezagutzen ez duten zerbaitekin topo egiten dutenean, eta arazoa konpontzeko behar duten informazioa berek sortzea erabakitzen dutenean. Baina batzuetan ikertzaileak ere ez daki kezka sortzen dion arazoari buruzko informaziorik ote dagoen –norbaitek lehenago horren inguruko ikerketa egin ote duen, alegia—. Beraz, logikoa denez, miaketa egiten da –hain zuzen, miaketa bibliografikoa—, edozein ikerketatan hasi aurretik arazoa konpontzeko behar den ezaguera dagoen ala ez ziurtatzeko.

Gaiari buruzko ezaguera bilatu ondoren informazio asko aurkitu ez bada ere, gehienetan gaiari buruzko hausnarketa sakonagoa egiten laguntzen du. Gutxiago ezagutzen diren alderdiak edo ezagutzeko erabil daitezkeen datu-iturriak zehazteko ere balio izaten du. Horren ondorioz, lan bibliografikoa egin ondoren, ikertzaileak behar duena hobeto definitzen du, hala nola horretarako erabil zitezkeen datu motak eta horiekin egin beharreko analisisa. Hortik aurrera lanaren plangintza zehazteko une egokiena hasten da.

Aipatu dugun prozesua da, labur azalduta, zientzialariek *ikerketa esploratzaile* deritzotena. Hain zuzen ere, gauza ezezagunarekin egin ohi diren miaketa-lanetatik datorkio izena ikerketa mota horri. *Ikerketa esploratzailearen* aurkezpena erabilgarri suertatzen zaigu Estatistikaren erabilera ikerketa-prozesuaren barnean kokatzeko; estatistika zertan gerta daitekeen lagungarria hobeto aurkezteko edo ikerketa zientifikoaren ohikoa den jarduera-segida aurkezteko:


 Badago arazo, behar edo jakin-min bat², zerbait ezagutzeko premia dakarrena.


 Arazoa konpontzeko baliabideei buruzko hausnarketa-prozesua egiten da, behar diren erantzunak lortzeko, jarraitu beharreko bidea aurkitzeko, hain zuzen ere.


 Gaiaren inguruan egon daitekeen ezaguera bilatzen da. Ikerketaren parte horretan, gaiari buruzko teoriak aurki daitezke,


² Jakin-mina, zerbait ezagutzeko grina, ase nahi denean, irtenbideen bilaketa den neurrian, arazotzat hartzen da.

gaiaren hainbat alderdi azaltzen dituztenak eta ezagutzen ez diren beste alderdi batzuk ere agerian uzten dituztenak.


 Lan-egitasmoa gauzatu egiten da. Zehazten dira ikerketaren helburu zehatzak, (hau da, ezagutu nahi diren alderdiak), eta horretarako bilatu eta jaso behar diren datuak.


 Aukeratutako datuak bilatu eta jaso egiten dira.


 Gero, jasotako datuen analisia dator.

 Azken pausoa emaitzak kaleratzea da.

Gero, aipatutako eskemari egokitzen zaion ikerketa-kasua azalduko dugu. Edozein jardura intelektualek eta lan zientifikoak dituzten parekotasunak erakutsi nahirik, norbaiten bizitza praktikoan gerta zitekeen egoera bat asmatu dugu:

 Gure familiako batek merkataritza-lokal bat dauka, eta bertan denda bat jartzea erabaki du, bizimodua ateratzeko. Arazoa da denda mota hautatzea. Aukera batzuk baditu buruan, baina egokiena edo errentagarriena zein izan litekeen erabakitzeko oinarririk ez dauka. Bestalde, ondo dakienez negozio baten arrakasta faktore askoren arabera dela, hartu beharreko erabakiak hobeto oinarritzeko informazioa bilatzea pentsatu du.

 Esperientziaz badaki hiriaren inguruko auzoetan denda ugari egoten direla, egunero produktuak saltzen dituztenak, auzotarrek hirigunera edo merkataritzagune handietara egunero joan beharrik eduki ez dezaten. Ideia hori abiapuntutzat hartuta, argi ikusten du horrelako dendek ezaugarriak aztertze beharra. Zer-nolako produktu edo zerbitzuak izaten dira jendeak egunero behar dituen horiek? Nola jakin daiteke auzo batean zer denda falta diren eta beharrezkoak edo egokiak izan litezkeen? Horrelako burutazioak zituela, hiriko beste auzoetan dauden dendak ikustera joan zen, ideien bila hain zuzen ere. Erabakia hartzeko nahikoa ez bazen ere, auzoen arteko konparazioa egitea komeni zela sumatu zuen; beste auzoetan zeuden dendak, bere auzoan falta zirenak, ezagutzea erabaki zuen.

 Gora eta behera luzaro ibili ondoren, jakin zuen bazegoela berak behar zuen

informazioa zeukan erakunde bat, Merkataritza Ganbera izeneko. Bertara jo zuen, eta ohartu zen hirian dauden dendei buruz informazio ugari daukela eta berea salgai dagoela. Horretaz gain, jakin zuen erakundearen bulegoetan hirien merkataritzari buruzko liburuak eta aldizkariak ere badituztela, eta kontsulta daitezkeela. Hortik aurrera datuak eta bestelako informazioak lortzea erraza izan zen.

■ Auzoetako dendei buruzko datuak eskuan zituela, denden kokapena hiri-plano batean egokitzea erabaki zuen, auzoko denda-banaketa argiago ikusteko. Informazio osagarriaren bila auzoetara berriro joatea ere deliberatu zuen. Hilabeteren buruan egoera honako hau zen:

■ Jasotako datu guztiak mahai gainean zituen aztertzeko prest.

■ Auzoetako denden banaketako kartografia ere eginda zeukan.

■ Datu-analisia (estatistika-analisia) egiteaz arduratuko zen lagun baten hitza zeukan.

■ Datu-analisiar arduratuko den lagunari eskaera edo galdera zehatzak egin dizkio:

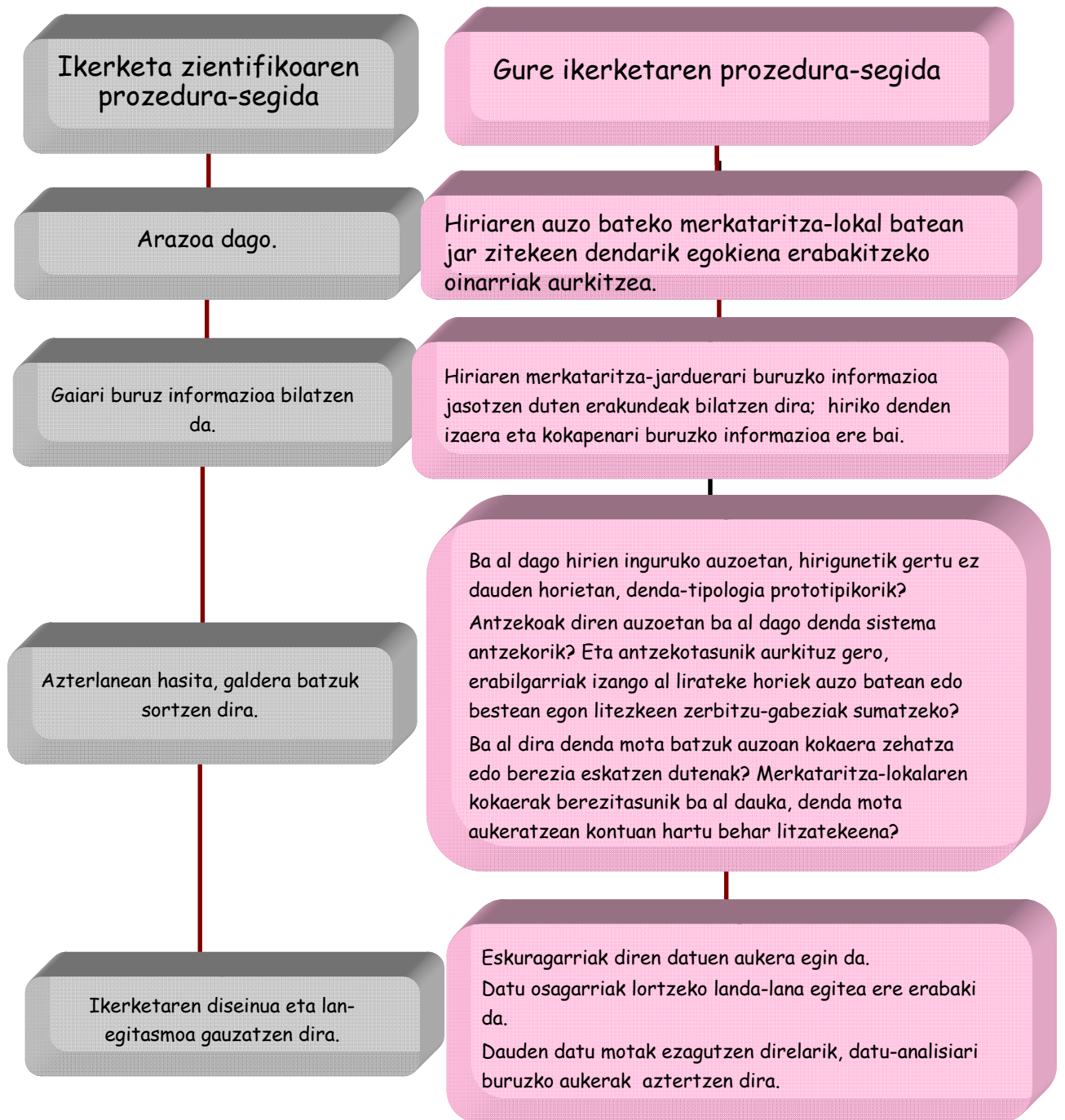
■ Tamaina antzekoa duten auzoetan antzekoa al da batez besteko denda kopurua?

■ Ba al dago denda motaren bat bestek baino ugariagoa dena?

■ Ba al dago denda mota bat beste auzoetan maiz agertzen dena eta bere auzoan, berriz, gutxitan ikus daitekeena?

■ Auzoen eta hirigunearen arteko distantziak diferenteak direla kontuan hartuz, ba al dago nolabaiteko harremanik denda mota eta kopuruaren eta auzoetatik hirigunera dagoen distantziaren artean?

Osatu dugun *kontakizunak* ikerketa zientifikoaren ohiko osagaiak ditu; osagai horiek ez daude modu formalean adierazita baina erraz suma daitezke:



Gure adibide hipotetikoaren edozein ikerketa zientifikotan egiten den jarduerarekin egokitzen zaiola ikusten dugu, eta edozein ikertzailek erabiliko luke, bere eremuko aztergairaren bat landu nahi izanez gero. Bestaldetik, pertsona zahurrak eta ausartak har lezakeen bidea ere izan liteke proposaturiko egoera, hain zuzen ere arazo pertsonalerako irtenbidea bilatzeko³. Dena den, gure kasua ez da hiri-merkataritzari buruzko azterketa orokor eta sakona egin nahi duen ikertzaile batena; gure pertsonaiak ez dauka ez gaiaren inguruko ezaguerarik, ezta heziketa espezializaturik ere. Ondorioz, ez ditu kontuan hartu metodo zientifikoaren betebeharrak eta ñabardura guztiak; ez ditu kontuan hartu, esate baterako, datuak jasotzean eta aztertzean analisiaren emaitzak zientifikotzat hartuak izateko bete behar diren baldintzak. Horren ondorioei berehala antzematen zaie:

- Gure ahaideak lortuko dituen emaitzak lagungarriak izango dira, seguruenean, dendaren inguruko erabakiak askoz hobeto oinarritzeko eta, oro har, bere negozioa hobeto kudeatzeko.
- Baina emaitza horiek ez lukete inolaz ere balioko hiriaren merkataritza alorrari buruzko ezaguera zientifikoa eskuratzeko. Ezinezkoa litzateke, beraz, hirien merkataritzari buruzko balio orokorreko ezaugarriak ondorioztatzea.

Edonola ere, gure helburua ez da lan zientifikoaren ezaugarriak sakonean jorratzea, metodo zientifikoaren eta Estatistikaren arteko harremanak erakustea baizik. Horregatik, arreta zeri egin ahalik eta argien jakiteko, berriro laburbilduko ditugu zein diren horrelako adibideen helburuak:

- Gai bati buruzko ezagutza sortzeko egiten diren buru-jardueren oinarritzeko

³ Guk geuk, hori frogatu nahirik, Internetera jo genuen; Google bilatzailea erabiliz, guk proposaturiko egoeraren antzeko adibide ugari aurkitu genituen. Bilaketa egiteko *merkatu-azterketa* hitza erabili genuen. Atera ziren ehunka emaitzaren artean, pertsona askoren galderak zeuden, denda bat jartzeko asmoa zeukaten pertsonenak. Galderak ere antzekoak izaten ziren: nola egin, beren kontura, azterketa bat beren hirien edo auzoen inguruko denda eta zerbitzuen egoera eta beharrak ezagutzeko.

pauso-segida erakutsi eta, adibideen bidez, adierazi.

- Zerbait ezagutzeko erabil daitekeen edozein metodoren eta metodo zientifikoaren arteko aldeak erakutsi eta adibideen bidez adierazi.
- Ikerketa zientifikoak eta Estatistikak nolako harremana duten erakutsi edo, beste era batera esanda, zein den Estatistikaren parte-hartzea prozedura zientifikoaren bidez ezaguera lortzeko.
- Estatistika metodo osagarria dela erakutsi, ezaguera sortzeko erabiltzen diren metodoei zeharo lotuta baitago.

Jarritako adibidean, ikertzaile batek, gure ahaideak ez bezala, hiri-merkataritzari buruzko oinarri teorikoak ezagutzen ditu, hala nola erabil litezkeen datu-iturriak (Merkataritza Ganbera eta beste). Eta batez ere, datuak jaso baino lehenago, ikerketaren hasieratik, estatistikari buruzko ezaguera erabiltzen du ikertzaileak, hainbat jarduera egiteko. Areago, analisirako tresna estatistikoak daudenez, hainbat analisi mota egin daitezkeela jakingo du ikertzaileak. Jakingo du, esaterako, hiriko denda-banaketaren azterketa egin daitezkeela eta zein motatako emaitzak espero daitezkeen; horregatik, analisi mota hori egiteko datuak jasotzea erabakitzen du. Hitz batez esanda, gure ikertzaileak lagungarria izango du estatistikaren ezaguera ikerketa-prozesu osoan zehar:

- Ikerketaren helburuak zehazteko unean, hiri-merkataritzaren ezaugarriak, azter daitezkeenak, aukeratzen eta definitzen laguntzen dio Estatistikak.
- Gaia aztertzeko erabil daitezkeen datu motak aukeratzeko eta jasotzeko unean ere zeharo lagungarria da Estatistika. Inkestak beharrezkoak izanez gero, galderak osatzeko eta egituratzeko ezinbestekoak dira.
- Jasotzen diren datuen analisisa egiteko ordua iristen denean, Estatistikak eskaintzen ditu ez bakarrik analisisa egiteko tresna zehatzak, baita hainbat tresnaren artean komenigarriena aukeratzeko oinarri teorikoak ere.

Ikerketaren prozesu osoan Estatistika lagungarria dugula konturatzen garenean, Estatistika formula multzo bat baino askozaz gehiago dela argi ulertzen dugu. Estatistikak eskaintzen duen ezagueraren barnean arrazoitzeko

tresnak ditugu, ezagutza adierazteko kontzeptuak, errealitateari buruz hitz egiteko kontzeptuak... Haren baliabideak ez dira, beraz, eragiketa matematikoak egiteko unean soilik erabiltzen.

Mario de Miguel-ek eta Roberto Behar-ek, estatistikari ezagunek, Estatistikaren definizio interesgarria eskaintzen digute, guk hemen aurkeztu nahi dugunaren funtsa laburbiltzen duena:

Estatistikak] Multzoei buruzko ezaguera lortzeko printzipio eta tresnak eskaintzen ditu, multzo horiek ezagutzeko jaso diren datuetan oinarrituz. Haren printzipioek eta tresnek metodologia osatzen dute, nolako datuak bilatu behar diren eta nola lor daitezkeen jakiteko. Datuak lortu ondoren antolatzeke eta ustiatzeko metodo eta prozedurak eskaintzen dizkigu, datu horietatik gure helburu eta interesen arabera atera daitekeen ahalik eta emaitza gehien lortzeko. Datu-analisitik, aztertzen den multzoari buruzko iritzi eztabaidaekin lortzen dira.⁴

Era berean, uste badugu Estatistika metodo edo metodo multzo bat dela, ezaguera sortzeko prozesuan laguntzen duena, analisi estatistikoa metodo zientifikoaren parte dela argi ulertzen dugu. Analisi estatistikoa da, beraz, ikerketa zientifikoan erabil daitekeen metodo bat; beste batzuen artekoa.

⁴ Mario Miguel Ojeda, M.; Behar, R. (2006): *Estadística, Productividad y Calidad*. Secretaría de Educación de Veracruz, Mexiko, 53. or.

ESTADISTIKA, ERREALITATEAREN EZAUGARRIAK EZAGUTZEKO TRESNA

Behin eta berriro aipatzen dugunez, Estatistika erabiltzen da zerbaiti buruzko ezagutza bilatzen den egoeretan, errealitateari edo munduari buruz zerbait jakin nahi dugunean alegia. Baina askotan errealitatea guztiz bestelakoa da, eta edozein testuingurutan erabiltzera ohituta gaudela aitortu behar dugu. Askotan, Estatistika ikasteen eragiketa matematikoak egin behar izaten ditugu, formulatxoan erabilera barneratu behar delakoan. Beraz, zeharo arrunta izaten da datu-zerrenda aurrean dugula kalkuluak egin behar izatea, horren bidez bilatzen den ezaguera mota jakin gabe. Ez da harritzekoa, beraz, horrelako bizipenen ondorioz Estatistika formula multzo bat besterik ez dela ondorioztatzea; formula multzoa hainbatetan ulertezina eta balio edo erabilgarritasun gutxiko dela pentsatzea.

Aipatu dugun egoera, esperientzia txarrarena, ezaguna da estatistikari eta Estatistika irakasleen artean. Bestaldetik, ezagunak ditugu, orobat, egoera tamalgarri hori aldatzeko egiten ari diren ahaleginak. Estatistika gero eta beharrezkoagoa da, zenbakien eta kontabilitateen erabilera gero eta handiagoa den gizarte batean. Biltzarrak egiten dira arazoa aztertzeko; lantaldeak sortzen dira, Estatistika irakasteko prozedura berriak sortu nahirik... Saio horietan guztietan leloa hau entzuten da etengabe: Estatistikaren erabilera ezagutzaren bilaketarekin batera gertatzen bada, Estatistikaren irakaskuntza antzeko egoeretan garatu behar da.

Gure kasuan, urteetako esperientzia eduki ondoren, aipaturiko planteamenduarekin bat gatoz. Hori dela eta, material honen helburua horixe dugu: Estatistikaren balioa eta erabilera errealitatea edo mundua ezagutzeko egiten diren azterketen barnean azaltzea.

Estatistikaren aztergaiak: errealitatearen egoera aldakorrak

Estatistika erabiltzeko, argi dugu arazoak edo aztergaiak behar direla. Hortik abiatuz, aipatuko ditugu, adibide gisa, interesgarriak izan daitezkeen aztergai batzuk:

 Denok ezagutzen ditugu azken urteotan mundu- mailan gertatzen ari diren

populazio-mugimenduak. Jakin badakigu, halaber, Europako nazio garatuenak direla etorkin askoren helburua. Arrunta da gaur egun gero eta biztanle atzerritar gehiago ikustea gure hirietan. Egoera horren garrantzia handia denez, gaian interesaturik daudenen artean galderak sortzen dira: Zein dira Espainian etorkin gehien jaso dituzten hiriak? Ba al dira, beraz, hiri batzuk beste batzuk baino erakargarriagoak, eta, hala bada, zergatik? Nola banatzen dira etorkinak hirien barnean: biltzen al dira etorkin-auzoetan edo hirian zehar banatzen dira?

- Etorkinen gaiari lotuta egon zitekeen beste gai bat udalen gizarte-laguntza da. Udalek eskaintzen dituzten gizarte-laguntzen ezaugarriak eta horretarako erabiltzen dituzten diru kopuruak desberdinak izaten dira. Hori dela eta, ez litzateke arraroa izango desberdintasun horiek etorkinen erabakietan eragina edukitzea. Hortik sor zitekeen galderari berehala antzematen zaio: udalek dituzten gizarte-politiken arteko aldeek eduki al dute eraginik migrazio-mugimenduetan. Kontuan hartzen al dute etorkinek udalen laguntzen politika beren bizitokia aukeratzeko orduan?
- Kezka handia sortzen duen beste gai bat euroa dugu. Denok dakigun bezala, 2004an hasi zen Europako Diru Batasuna; Espainiako kasuan, pezetatik eurora pasatu zen. Harrez geroztik, bizitza inoiz baino gehiago garestitu dela sumatu dugu denok. Horri buruzko berriak agertzen dira maiz komunikabideetan. Adostasun handiko gaia da, beraz. Baina hainbatetan ikusgarria izan den garestitze hori neurri berdinekoa izan al da produktu eta zerbitzu guztietan? Hau da, beharbada, galdera interesgarria: zein izan dira eragin handiena jaso duten produktuak eta zerbitzuak?
- Klima-aldaketa eta ingurumenaren hondamena direla eta, erakunde publikoek kontzientziazio-kanpaina ugari egin dituzte. Horien gastuak handiak izan direla kontutan hartuz, komenigarria litzateke horrelako kanpainen lortu duten arrakasta ezagutzea. Eduki al dute kanpainen zabalduetako mezuek eraginik biztanleriaren usadio eta ohituretan? Handitu egin al da biztanleek inguruarekiko daukaten arreta? Eta hala bada, zertan nabari da? Azken urteotako biztanleriaren ohiturak aztertuz gero, aurki al daiteke joera positiborik?

Egun inork ez du zalantzan jartzen proposatu ditugun galderari erantzuteko

ikerketa estatistikoak erabiltzen direla. Adibide gisa egin ditugun galdera horiek gaurkotasan handiko gaiei lotuta daude, eta, ondorioz, erraza da horrelako gaiak jorratu dituzten ikerketen emaitzak egunkarietan aurkitzea. Bestaldetik, arrunta izaten da gure eguneroko bizitzan halako ikerketetatik sorturiko informazioak erabiltzea.

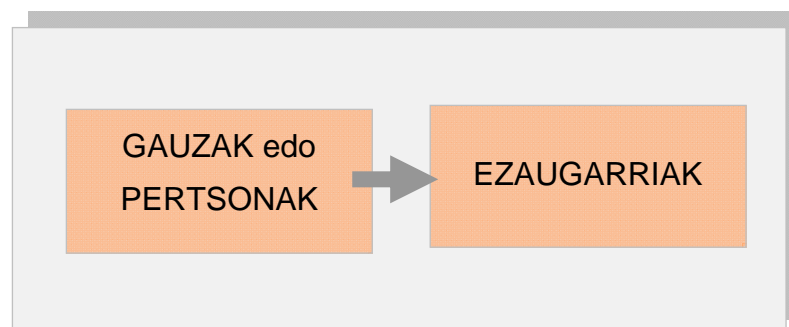
Esan bezala, aipaturiko gai guztiei buruzko ikerketak egin dira, eta ikerketa gehienetan ere metodo estatistikoak erabili izan dira (askotan oso antzekoak). Horrela bada, nola ulertzen da orduan hain desberdinak diren gaien azterketa metodo berdintsuen bidez egitea? Erantzuna aurkitzeko, egindako galderei arretaz erreparatu behar diegu, eta orduan berehala ikusiko dugu ezaugarri komunak dituztela:

📌 Kasu guztietan pertsonen edo gauzen ezaugarriak aipatzen dituzte; materialak edo materiagabeak izan daitezkeen gauzak:

📌 Migrazio-mugimenduen adibidean, bizitokia izango litzateke etorkinen ezaugarri bat. Europako nazio bateko hiriek jaso dituzten etorkinen kopurua hiri horien ezaugarri bat izango litzateke. Lehenengo kasuan, pertsonen buruzko ezaugarri batetaz ari gara, bigarrean, berriz, hiriei buruzko ezaugarri batetaz ari gara.

📌 Euroaren inguruko planteamenduan, prezioa izango litzateke produktu edo zerbitzuen ezaugarria.

📌 Ingurumenari buruzko adibidean, azterkizunak biztanleak izango lirateke, eta horien ezaugarria ingurumenean eragina eduki lezakeen ohituraren bat.

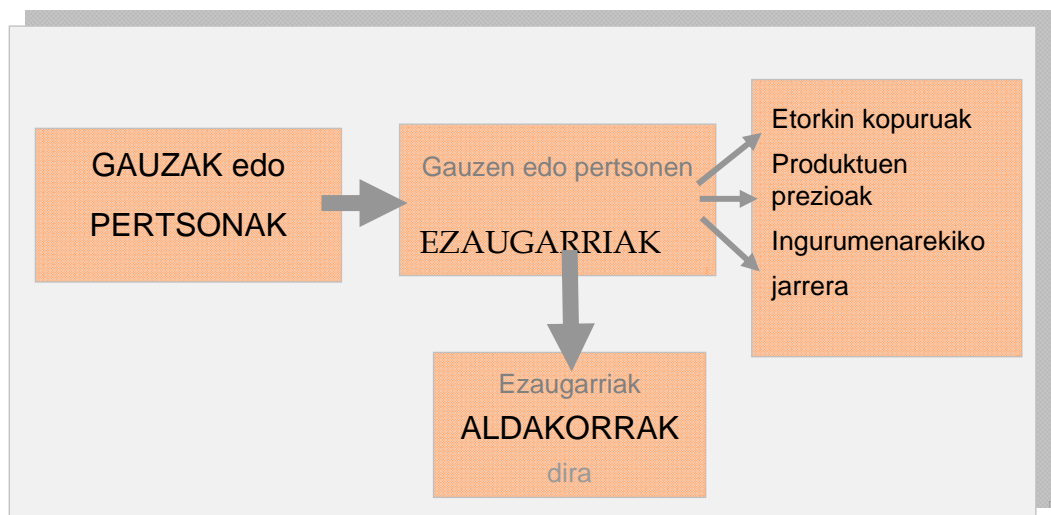


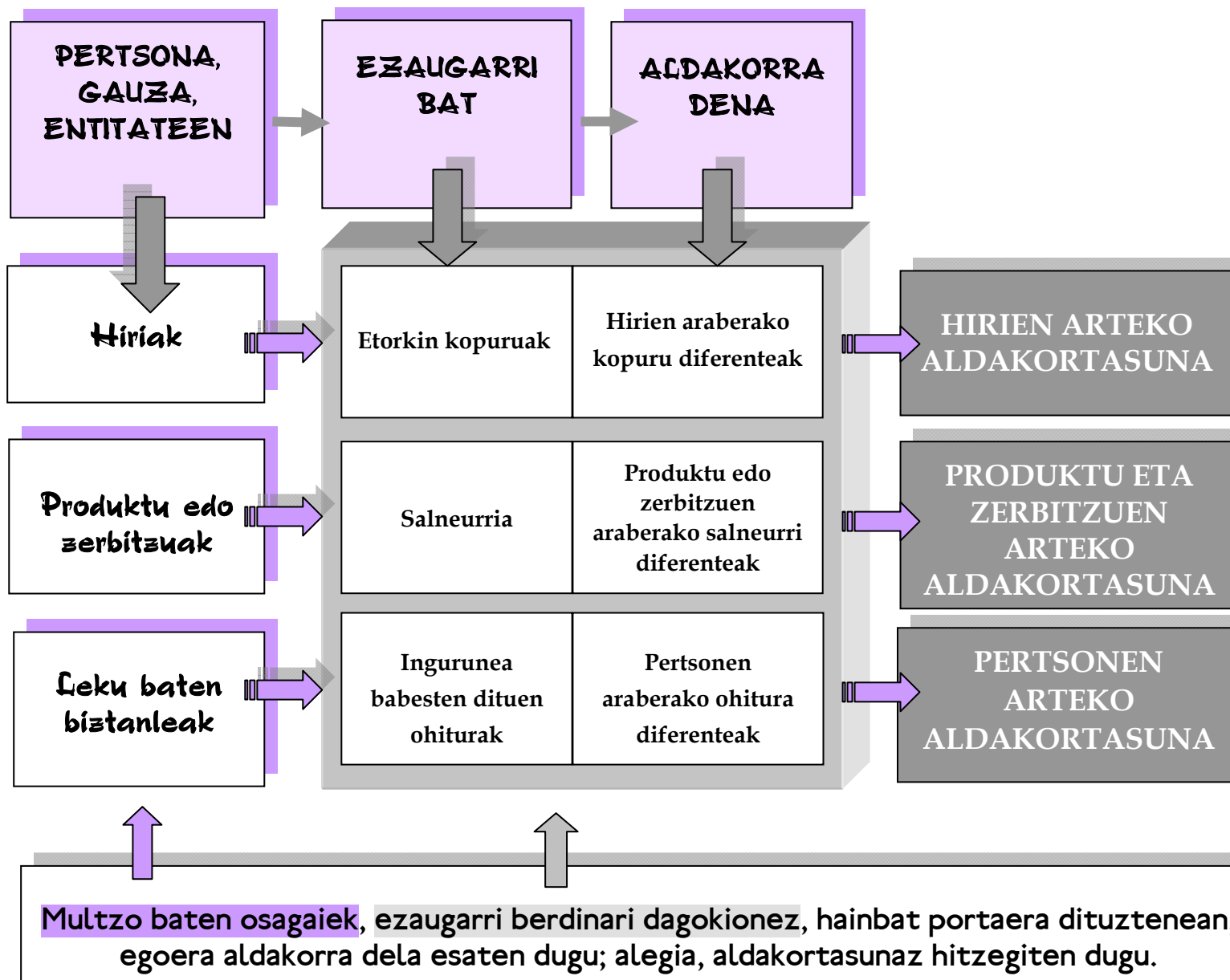
■ Egindako galderen erantzunak ezagutzen ez baditugu ere, argi sumatzen da galdera guztiek erantzun aldakorrak izango dituztela.

■ Etorkinei buruzko adibidean, erantzunak zeharo alda daitezke, esaterako, aztertzen dugun unearen edo aldiaren arabera: hiri bakoitzak jasotzen duen etorkin kopurua aldatu egin daiteke urte batetik bestera, eta, ondorioz, etorkin gehien edo gutxien jasotzen dituzten hiriak ere beste batzuk izan daitezke.

■ Euroaren sarreraren ondoren gertatu diren prezio-igoerak ez dira alor guztietan berdinak izan; igoerak aldatzen dira lekuaren, produktuen eta zerbitzuen arabera.

■ Hiri bateko biztanleek ingurumenaren aurrean duten jarrera eta babestu nahirik hartu dituzten ohiturak zeharo desberdinak izango dira pertsonen izaeren arabera. Kontzientziazeko egin diren kanpainen eragina ere diferentea izango da biztanle bakoitzarengan. Kezka eta arreta handia daukatenen eta zeharo axolagabeak direnen artean kontzientzia-gradu ugari egon daitezke.





Jarri ditugun adibideak lagin txiki bat besterik ez dira, errealitatean, munduan, ageri diren ezaugarri aldakorrek konparatuta. *Aldakortasun* hitza multzo bateko osagaien arteko desberdintasunak aipatzeko erabiltzen dela konturatuz gero, berehala ohartuko gara bukaezinak direla aurki daitezkeen multzoak, osagaien arteko ezaugarri desberdinak dituztenak. Edozein multzo hartuta, erraz sumatuko dugu ezaugarri sorta bat aldakorra dela bere osagaien artean:

■ Mende baten urte kopurua multzoa litzateke; multzo horren ezaugarriak, beste askoren artean aipa litezkeenak, honako hauek dira:

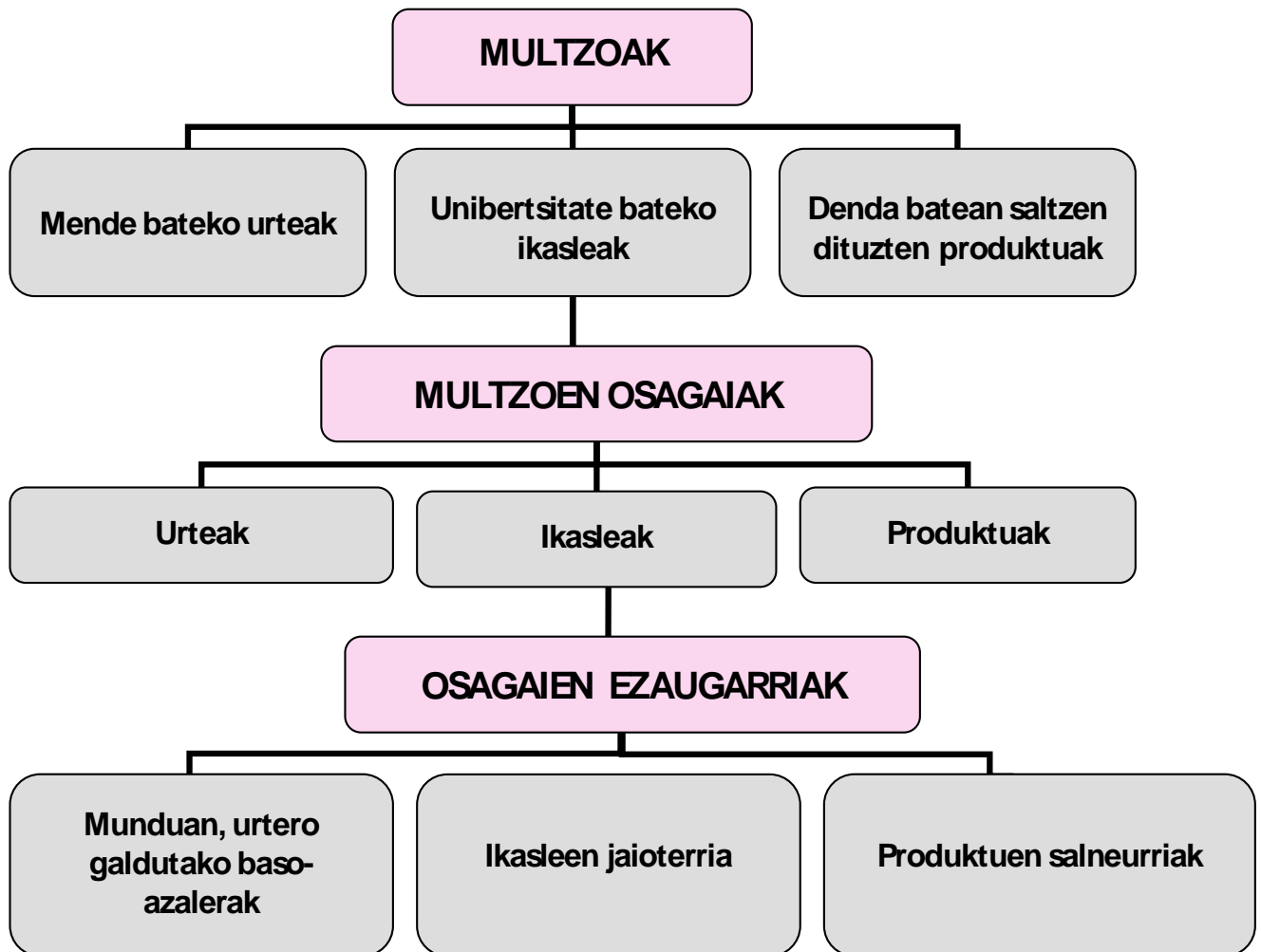
■ Gerren ondorioz hil diren pertsonen kopuruak aldatzen dira osagai batetik bestera (alegia, urte batetik bestera).

■ Leku batean jaso diren prezipitazioen kopuruak ere aldakorrak dira urteen arabera.

■ Munduan desagertu diren basoen azalerak ere aldakorrak dira urteen arabera.

■ Unibertsitateko ikasleek multzo bat osatuko lukete, baita nazio bateko herrialdeek ere; urte eta hilabete berdinetan jaio direnek; ibai-arro bereko ibaiek... Zerrenda bukaerarik gabea da.

Egia esanda, aldakortasunez inguraturik bizi gara, eta, konturatu ala ez, ohituta gaude bizitzaren arlo guztietan aldakortasunarekin moldatzen:



Medikuarengana joaten garenean gehienetan ez dakigu alde zuretik zenbat denbora emango dugun, itxaronaldia gehienetan ere aldakorra izaten baita. Gure bizitza osoan medikuei egindako bisita guztiek osa lezakete aipaturiko multzo horietako bat. Ezaugarria, kasu honetan, itxaroten eman dugun denbora da; oraingo honetan ere, aldakorra izango da bisita batetik bestera.

Multzoa : epealdi batean medikuei egindako bisitaldiak

Multzoaren osagaiak

Ezaugarria : itxaroten eman dugun denbora

1998ko otsailaren 1a

45 minutu

1998ko ekainaren 19a

10 minutu

1999ko urriaren 10a

3 minutu

2000eko apirilaren 6a

20 minutu

2000eko abenduaren 8a

75 minutu

2001eko martxoaren 21a

15 minutu

2002ko uztailaren 13a

10 minutu

Zerbait erostea erabakitzen dugunean, argi edukitzen dugu askotan, hainbat arrazoiengatik, gauza horren prezioa aldakorra izan daitekeela; zer esanik ez, etxebizitzetara buruz hitz egiten badugu. Hiriko edo herriko inguru berean egonda ere, alde handiak egon daitezke bi etxebizitzaren prezioen artean.

Mediku eta erosketen adibideak, guztiok eduki ditugun esperientziak direnez, erabilgarriak suertatzen zaizkigu erakusteko eguneroko bizitzan aldakortasunaren aurrean jokatzeko estrategiak erabiltzen ditugula:

- Mediku-kontsulten kasuan, jendeak erabiltzen dituen estrategien artean dugu denbora pasatzeko zereginen bat bilatzea. Modu horretan, itxaronaldia motza edo luzea izan, irakurriz edo eskulana eginez eramangarriagoa izaten da.
- Gauzak erosterakoan hainbat pertsonaren estrategia da denda batera baino gehiagotara joatea; modu horretan, nahi duten produktuaren prezio guztiak jakinda, merkeena erosteko aukera izaten dute.

Espainiako etxebizitzaren prezioa

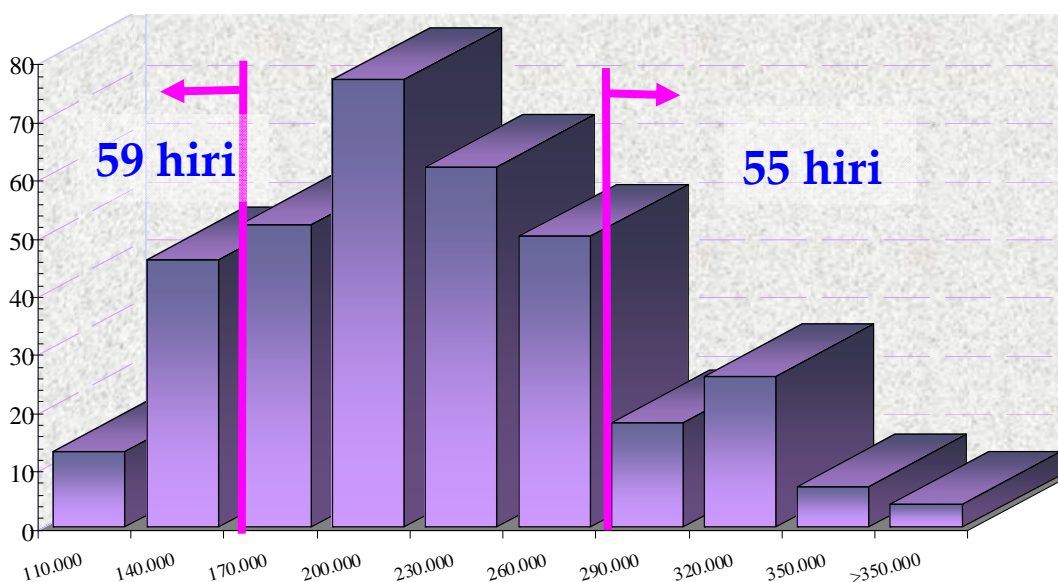
Denok dakigu etxebizitzaren prezioa asko alda daitekeela urte batetik bestera, eta gehienok badakigu azken urteotan prezioen igoera kezagarria gertatu dela Espainiako Estatuan. Espainiako probintzien edo hirien artean dauden prezio-aldeak ere ezagunak izaten dira, komunikabideetan aipamen ugari egiten baitira. Aldeak zein mailatakoak izan daitezkeen izango da, beharbada, gutxien ezagutzen duguna.

Demagun Espainiako probintzien edo hirien artean dauden etxebizitzaren prezioen arteko aldeak zehatzago ezagutu nahi ditugula, eta horretarako datu-zerrenda bat lortu dugula; hor, hainbat hiritako etxebizitzaren prezioak agertzen dira.

Gure jakin-mina asetzeko informazio nahikoa dugu dagoeneko, hain zuzen ere, [355 datu](#), hiri bakoitzaren batez besteko prezioei dagozkienak. Datuak banan-banan irakurriz gero, hirien artean alde nabarmenak daudela baieztatu besterik ezin dugu egin. Prezio-aldakortasunaren ikuspegi orokorra eduki nahi izanez gero, datu multzoaren analisia egin beharko da.

Espainiako etxebizitzaren prezioen igoerak edo eskualdeen arteko diferentziak direla eta, ohikoa izaten da zabalkunde handiko komunikabideek gaiaren inguruko informazioak kaleratzea. Haiek informazioa emateko duten moduaren eraginez pentsa daiteke nahikoa dela prezio altuena eta baxuena aipatzea egoeraren ikuspegi orokorra emateko. Errealitatea guztiz bestelakoa dela ohartzeko, nahikoa izango dugu grafiko honi begiratzea:

Espainiako 355 hiritako etxebizitzaren batez besteko prezioa. 2007⁵



Grafikoa ikusirik, etxebizitzaren prezioei dagokienez, hirien artean dagoen aldakortasun ikusgarria da, seguruena, ezaugarri nabarmenena:

- Hirietatik 59tan prezioa 140.000 eurotik beherakoa da.
- Beste 55 hiritan, berriz, 260.000 eurotik gorakoa da.

⁵ Ikusi datu multzoa eranskinean. 2. Taula.

Etxebizitza merke eta garestien arteko aldea ez da txikia, hain zuzen ere, 130.000 eurokoa (hau da, lanbide arteko gutxieneko soldata jasotzen duen langileak 16,3 urtean lanean lortuko lukeen kantitate bera).⁶

Aipatutako datuen aurrean, zalantzarikoa da batezbesteko prezioen datua ezertarako baliagarria den;

Ez dugu uste, beraz, hiri guztietako etxebizitzaren batez besteko prezioaren datuak (201.711,43 euro) egoera deskribatzeko balio duenik.



Era berean, hiririk garestienaren batez besteko prezioa 445.400 eurotik gorakoa dela edo merkeena 86.200 eurotik beherakoa dela jakitea beharrezkoa dela onartzen dugu, baina ahaztu gabe datu horiek (prezio-eskalaren muturretako egoerei dagozkienek) ez digutela inolaz ere ematen egoerari buruzko ikuspegi orokorra.

Horren ohikoak diren deskribatzaile horiek (batezbestekoa eta muturreko balioak) ez badira inondik ere nahikoak egoeraren irudi adierazgarri bat osatzeko, argi dago bestelako tresna batzuk bilatu beharko ditugula. Hiri gehienaren prezioak eta prezio-aldeak, baita hiri gehienaren artekoak ere, aipatu beharko genituzke, egoeraren irudi esanguratsu bat osatzeko. Argiago ikusteko, adibide batez baliatuko gara.

⁶ 2007rako EUSTATEk eskaintzen dituen datuen araberrako kalkulua (570,60 euro).

Espainiako etxebizitzaren prezioak:

Alde handia hirien artean:

Hiri garestiena, Bartzelona: batez beste 445.400 €

Hiri merkeena, Villanueva de la S. (Badajoz): batez beste 86.200 €

Hiri gehieneko egoera: 141.157,71 - 262.265,95 €

Hirietatik % 70ean 290.920 eurotik beherako prezioak daude

Laburpen moduko oharretan, batezbestekoaren datua aipatu ere ez dugu egin. Horren aldakorra den egoera batean informatu beharrean desinformatuko luke. Esan genezake, kasu horretan, batezbestekoaren datuak ondorio okerrak ateraraz diezazkiokeela.

Etxebizitzaren prezioen aldakortasuna erakustez gain, adibide horrek balio digu Estatistikak eskaintzen diguna berriro aipatzeko. Argi eta garbi ikus daiteke zeregin nagusia ez dela izan *formulatxo*en kalkulua. Kasu horretan, datuek erakusten duten egoera deskribatzeko adierazle esanguratsuak aukeratzea da datu-analisiaren zereginik nabarmena.

ESTATISTIKA ETA IKERKETA ESTATISTIKOEN AHALMENA

Ondorengo bi atalen helburua da Estatistikak gaur eguneko gizartean daukan garrantzia edo, zergatik ez esan, daukan boterea erakustea. Gure iritziz, zerbait ikasten denean, lehenik jorratu eta barneratu behar dena zera da: gaiaren garrantzia eta kokaera ezaguera-maparen barnean. Horrek ematen digun ikuspuntua gaia jorratzeko gida lagungarria izaten da.

Estatistikaren garrantzia erakusteko, askotariko adibide ugari ditugu, eta, horren ondorioz, zaila da aukeratzea. Guk aukeratu dugun lehenengo adibideak erakusten du nolako ekarpena egiten duen Estatistikak errealitateari buruzko pentsamoldeak eta irudiak osatzeko. Bigarren adibidean, jendearen bizitzan daukan eraginagatik, administrazio publikoan egin ohi den erabilera mota bat aipatuko dugu.

Europar Batasuneko nazioek hezkuntzara bideratzen dituzten baliabideak

Europar Batasuneko hainbat nazioek osatuko dute lehenengo kasuaren multzoa. Multzo horren osagaiak (nazioak, alegia) desberdinak dira; ezaugarri desberdinak dituzte. Desberdinak diren ezaugarrien artean dugu, besteak beste, nazio bakoitzak hezkuntzara urtero bideratzen duen diru kopurua. Alde horiek ohiko aztergaiak izaten dira, nazioen beste egoera batzuen adierazle egokitzat hartzen baitira.

Nazio baten politikarako biztanleen hezkuntzak daukan garrantzia neurtzen da, neurri batean, hezkuntzara bideratzen diren giza baliabide eta baliabide materialetan oinarrituz. Hezkuntzara bideratzen den BPGaren⁷ ehunekoia izaten da aipaturiko baliabideen adierazle orokorra; hor sartzen dira gastu publikoa eta familiek egiten dutena. Indize horren interpretazio egokia egiteko

⁷ Barne-produktu gordina.

kontuan hartu behar dira beste magnitude batzuk: biztanleko BPGa, eskola-adineko biztanleriaren proportzioa, eskolatze-tasak eta hezkuntza-sistema, besteak beste.⁸

Adostasun orokorra dago, beraz, hezkuntzaren balioari dagokionez; nazioen garapenaren zutabea den neurrian, arlo horretara bideratzen diren baliabideak nazioen inbertsio handiena eta errentagarriena dira. Nazioen hezkuntza maila, horren kalitatea eta kantitatea ezinbestekoak dira garapenerako, hazkunde ekonomikorako eta gizartearen ongizaterako.

Hezkuntzari buruzko informazioari ematen zaion garrantzia kontuan harturik, ez da harrizkoa nazioen egoerak konparatzen dituzten ikerketak aldika egitea. Nazioen hezkuntza-sistemaren bikaintasuna balioestea eginkizun konplexua da; erabaki behar da zein diren azter daitezkeen edo aztertu behar diren alderdiak, hezkuntzaren egoera modu egokienean islatuko dutenak. Hezkuntza neurtzeko adierazle erabiliena nazioen urteko gastua da, kalitatea baino gehiago kantitatea neurtzen duena; urteko gastua, nazioen aurrekontuen proportzio gisa adierazita, nazioek hezkuntzari ematen dioten garrantziaren islatzat hartzen da.

Hezkuntzara bideratzen diren baliabideei buruzko ikerketak erabiliz, inbertsio mailaren arabera nazio zerrendak osatzen dira.

Bistakoa da ranking mota horiek, albiste izan ohi diren heinean, gaitasun handia daukatela herrialdeen irudia osatzeko; barneko irudia (bertako biztanleena, alegia), baita herrialdeek kanpora agertzen dutena ere. Ildo horretatik, zerrendak Espainiako hezkuntzaren egoerari buruzko irudi ilun bat sorrarazten du: zerrendaren muturrean dago, azkenaurreko lekuan, Grezia soilik atzetik duela.

Aipatzen ditugun azterketen ondorioak ez dira hor amaitzen. Biztanleen hezkuntza estatuaren ardura denez gero, arlo horretan gertatzen diren

⁸ «Sistema estatal de indicadores de la educación, 2004», INECSE (Instituto Nacional de Evaluación y Calidad del Sistema Educativo).

arrakastak eta bereziki porrotak estatuaren lana ebaluatzeko tresnak izaten dira. Halaber, askotan ebaluazioa barik azpijoko politikoa izaten da; oso arrunta da, hauteskunde garaian bereziki, horrelako informazioak erabiltzea alderdi politikoek elkarren aurkako salaketak egiteko.

El Periódico de Catalunya izeneko egunkarian, 2003ko irailaren 17ko zenbakian, aipatu berri dugunaren adibide esanguratsu bat aurki daiteke. Egungutxi batzuk lehenago, Parisen kaleraturiko ikerketa baten emaitzek –hain zuzen ere, hezkuntzarako inbertsioei buruzko rankingak– Espainiako hainbat politikariren erreakzioa ekarri zuten. Aipatutako egunkariak sortu zen zalapartaren berria jaso zuen, eta, noski, haren oihartzuna ere bai:

Espainiak galdu du urrezko aukera bat hezkuntzarloan dituen gabeziak gutxitzeko. Hazkunde ekonomikoa gertatu arren, hezkuntzan inbertitzen duenari dagokionez, OCDE (Ekonomia Lankidetzeta eta Garapenerako Erakundea) osatzen duten 30 herrialdeen arteko zerrendan, hogeigarren lekuan dugu. [...]

PSOEk eskatu zuen hezkuntza-ministroa, Pilar del Castillo, Kongresuan agertzeko. Datuek, lotsagorritzeko modukoek, Gobernuaren hezkuntza-politika irrigarri uzten dutela salatu zuen Amparo Valcarce-k, sozialisten bozeramaileak.

Urte berean, 1983an alegia, Espainiako hezkuntza-sistemaren egoerari buruzko txostenean, Estatuko Eskola Kontseiluak ohartarazi zuen Gobernuak Espainiako hezkuntza-inbertsioa inguruko herrialdeena baino apalagoa zela. Egoera konpondu behar zelakoan, hezkuntzarako gastu publikoa handitzeko eskatu zion.

Iskanbila eta, batez ere, bilatutako oihartzuna lortzen direla ikusi dugu. Hortik aurrera ezer gutxi balio du horrelako mailaketen jatorria edo ikerketa egiteko erabili diren datuen zilegitasuna zalantzan jartzeak. Izan ere, eginez gero, seguruenean, emaitzak estatistikako adituen arloan geratuko lirarteke, eta, batez ere, egunkariak zabalduko ideiek ez lukete ezabatuko. Granadako Unibertsitateko bi ikertzailek, mailaketak direla eta, horien balioa erabat zalantzan jartzen dute:

Ukaezina da mailaketek erabakiak hartzeko baino kazetaritzarako balio handiagoa dutela; ukaezina da, baita ere, haien fidagarritasun metodologikoa (Berry, 1999) eta egiteko erabiltzen diren oinarrien baliagarritasuna (Morrison, 1995) eztabaidagarriak direla. Aipaturiko erabilera dela eta, edozein datu mota argitaratu orduko egin ohi diren mailaketak saihestea ezinezkoa da. Ikerketek azalera ditzaketen ondorioen beldurra ezin da aitzakia izan, erakundeek beren erabiltzaileei ematen dieten informazioa ez hobetzeko.⁹

Metodoen fidagarritasunaz gain, ikerketa eta mailaketak kutsatuta daudela salatu dute askok. 1980ko hamarkadan, Frantziako bi soziologo ospetsuk argi utzi zuten horrelako eztabaidek, hala nola mailaren kontzeptuak berak, duten izaera ideologikoa (Baudelot eta Establet, 1989). Hezkuntzaren garapena modu askozaz zehatzago eta eraginkorrean balioetsi behar dela alde batera utzita, haiek diote hezkuntza-mailaren inguruko eztabaidek zama ideologikoa dutela. Horiek eztabaida eta konparatze eraikitzailea galarazten dute¹⁰.

Zer-nolako ondorioak atera daitezke egoera horretatik?

Ondoriorik nabarmena izan liteke, beharbada, ikerketa estatistikoaren botereari lotuta dagoena. Iritzi publikoa bideratzeko dela, administrazioaren erabakiak babesteko dela, estatistika-emaitzek indarrez hornitzen dute edozein diskurtso.

⁹ Javier Vidal: «¿Cómo se utiliza la información estadística en las Instituciones de Educación Superior en España?». Encuentro Internacional de responsables de Información Estadística de las Instituciones de Educación Superior. UNAM, Mexiko, 2001eko azaroak 7-9.

¹⁰ Alejandro Tiana Ferrer, A. (2001): «Logros y desafíos de la educación al inicio del siglo XXI». Lázaro Llorente, L.M. (ed.): «Problemas y desafíos para la educación en el siglo XXI en Europa y América». Valentzia, Valentziako Unibertsitatea.

Matematikaren erabilerak, alde batetik, eta metodo zientifikoarenak, bestetik, fidagarritasuna ematen diete datu estatistikoei, eta horiek duten ahalmena azaltzen dute.

Horrelako ikerketen oihartzun publikoari dagokionez, argi dago ezer gutxi egin daitekeela haien erabilera interesatua edo garrantzi mediatikoa saihesteko. Eragin txarretatik salbu egoteko irtenbide bakarra norberaren esku dago. Irtenbidea norberaren ahalmen kritikoa jorratzean datza. Ikuspegi kritikoz hornituta, edozein ondorio onartu baino lehenago, emaitza estatistikoen egokitasuna, datuen jatorria edo datuak lortzeko prozeduren balioa aztertuko ditugu. Aipatutako ikuspegi kritikoa lortzeko ezinbestekoa da estatistikari buruzko oinarriko ezaguerak edukitzea. Ezaguera horiek eta tresna estatistikoen balizko ahalmenari buruzko ezaguerak lagunduko digute hainbat informazioen eta egunkarietako titulu batzuen egitasunaz eztabaidatzen.

Estatistika eta administrazioaren erabakiak. Mamografiak hautatuentzat soilik

Azkenik, aipatu nahi dugu administrazio publikoaren kudeaketarako ikerketa estatistikoek daukaten zeregina edota eragina, ikerketa horiek erabakiak hartzeko oinarria izaten diren heinean. Aipatzekoak dira, bereziki, gizartean, oro har, edo horren sektore batean eragin nabarmenak edukitzen dituzten erabakien kasuak.

Egoera horien adibide esanguratsu bat agertu zen *El País* egunkarian 2007ko urriaren 25ean. Berriaren izenburua adierazgarria izan zen: «Mamografiak berezientzat soilik». «Bereziak» Eusko Legebiltzarreko parlamentariak dira. Aseguru mediko berezia dute. Aseguruaren ezaugarriengatik, 40 urtetik aurrera bularreko minbiziaren diagnostiko goiztiarraren azterketa medikoak egiteko eskubidea daukate. Osasun-sistema publikoak, ordea, berrogeita hamar urtetik aurrera ordaintzen ditu aipaturiko azterketak. Bi egoeren arteko aldea agerian jarri zuen berriak, gaiaren inguruko eztabaida piztu zuen. Mamografiak zer urtetatik aurrera egin behar diren erabakitzeke oinarri zientifikoei buruzko eztabaidak ere berpiztu egin zituen albisteak.

Administrazio publikoa da, Osakidetza kasu honetan, azterketa medikoa egiteko gutxieneko adina ezartzen duena. Hori egiteko oinarriak, berriro ere,

ikerketa estatistikoak dira, gaixotasunak adinaren arabera daukan eragina aztertzen dutenak. Aipatzen ditugun hainbat erabakiren arrazoiak ekonomikoak dira: azterketa medikoak egiteko adina aurreratzeak gastuen igoera ikusgarria ekarriko lioke administrazioari. Mamografiei buruzko berriak zioenez:

Aurten, Osakidetzak 64 urtetik 69 urtera bitarteko emakumezkoak ere sartu zituen mamografien programaren barnean. Horrek ekarri zuen 1,1 milioiko gastua, teknologia eta giza baliabideak ordaintzeko; horrenbestez, beste 50.000 emakumeri zerbitzua eman zitzaien.

Argi dago gure helburua ez dela mamografiei buruzko eztabaidan sartzea, baizik eta adibide gisa erabiltzea, ikerketa estatistikoek izan dezaketen boterea erakusteko, bereziki emaitzek giza taldeetan eragin nabarmena daukaten kasuetan. Osasun publikoaren alorrean edo diru-baliabide gutxiagoko biztanleriari eskaintzen zaizkion gizarte-laguntzen alorrean gertatzen dira, besteak beste, aipaturiko egoerak. Ohikoa izaten da, esaterako, horietan nork jasoko dituzten diru-laguntzak edo nork edukiko dituzten doako zerbitzuak erabaki behar izatea; eta ohikoa izaten da, halaber, aipatutako erabakiak hartzeko metodo estatistikoez baliatzea.

Argi dago, orobat, hemen eskaintzen den materiala azterketa estatistikoak hasi dituzten ikasleei zuzenduta dagoela, eta ikasle horiek ez dutela parte hartuko laster –edo, beharbada, inoiz ez– garrantzi handiko ikerketa estatistikoetan. Baina gauza bat da ikerketak egiteko gaitasuna, eta beste bat, oso bestelakoa, ikerketen emaitzen esanahia, garrantzia eta aplikazio-eremuak ezagutzeko eta ulertzeko gaitasuna. Desberdintasun horretatik abiatuz, argi azal daiteke orain gure helburua: estatistikaren erabilera hain zabaldua dagoen gizarte batean moldatzeko hasi-masiak eskaintzea. Hasi-masi horietatik garrantzitsuenetariko bat, zalantzarik gabe, Estatistikaren zeregina dugu.

Ez gara originalak kultura estatistikoaren garrantzia aipatzean; adostasun zabaleko iritzi bat da. Richard Blanco Peck estatistikari ezagunaren hitzak izango dira, beharbada, aipaturiko ideiaen isla egokiena:

Prestakuntza egokia garapenaren motorra da, eta ez bakarrik estatistikak egiten dituzten teknikarien prestakuntza. Funtsezkoa da, halaber, interpretatu eta erabakiak hartzeko erabili behar dituzten profesionalen eta biztanleen prestakuntza, hala nola behar diren datuak jasotzeko laguntza ematen dutenena.¹¹

Haren hitzetatik bereziki azpimarratuko genituzke biztanleen prestakuntzari dagozkionak.

Peck ez da bakarra, beste aditu askok ere gauza bera aldarrikatzen dute eta. Aditu gehienek iritziz, biztanleek ezagutu behar dute noraino iristen den Estatistikaren erabilera gure gizartean eta erabilera hori, askotan, ez dela batere neutrala. Ikerketa estatistikoak antolatzen dira batzuetan, gehiegitan beharbada, egileek nahi dituzten emaitzak lortzeko moduan. Guztion onurarako izango delakoan, eta zientziaren aterpeak babestuta, ezkututzen dira batzuen probetxurako izango diren gauzak. Halakoak gertatzen direla jakin beharko genuke guztiok. Jakin beharko genuke, baita ere, Administrazio publikoaren erabakiek, batzuetan, batezbestekoetatik gertu ez dauden pertsonak kanpoan utziko dituztela.

¹¹ Blanco Peck, R. (2006): «Los Enfoques Metodológicos y la Administración Pública Moderna». *Revista Cinta de M. Revista de Epistemología de Ciencias Sociales*, Txileko Unibertsitatea, 2006ko abendua.

*2. atala. Estatistikaren
erabilera aldagaien
aldakortasuna aztertzeko*

DATUEN LEHENENGO MIAKETA

DATUAK, ALDAGAIAK, POPULAZIOAK

Batzuetan, kostatzen da prozedura estatistikoak ulertzea, txarto ulertzen direlako, edo ez direlako ulertzen, oinarrizko kontzeptuak, ondorengo azalpen guztietan etengabeko erreferentzia izango direnak. Horiek kontuan hartuz, oinarrizko kontzeptuei buruzko hainbat argibide azalduko ditugu.

Osagai multzo bat = POPULAZIO BAT	Osagaien ezaugarriak = ALDAGAIAK
Euskadiko ibaiak	Luzera
	Udako batez besteko ur-emia
	Zeharkatzen duten lurralde historikoa
	Bokaleko lekua
	Ur-kalitatea

Goiko zerrendan jaso ditugu ibaien deskribapena egiteko aipatu ohi diren ezaugarriak. Hainbat motatako ezaugarriak dira. Luzera metrotan edo kilometrotan adierazten da. Ibaiek zeharkatzen duten lurralde historikoa beste ezaugarri bat da, leku-izenen bidez adierazten dena. Ur-kalitatea aipatzeko zenbait kategoria edo maila bereizten dituzten sailkapenak erabiltzen dira.

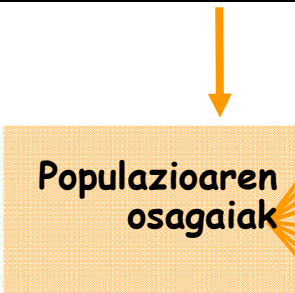

Aipatu ditugun ezaugarrien aldakortasuna aztertuz gero, zenbait prozedura erabili beharko genituzke. Ibaien luzerari buruzko datuak, esaterako, zenbakizko datuak izango lirateke. Ur-kalitateari buruzko datuak, berriz, kualitatezko datuak izango lirateke, hots, kalitate-maila deskribatzen duten adierazpideak. Bi kasuetan datu multzo bat edukiko genuke, baina, izaera desberdinekoak. Zer dira, orduan, datuak? Datuak dira ikusi den zerbaiten emaitza, informazio unitate edo zatiak. Datuak dira, beraz, zenbakien bidez edo hitzen bidez aipatzen diren informazio-unitateak.

Estatistika-metodoen bidez azter daitezke hainbat motatako datuak. Badira, beraz, ezaugarrien edo aldagaien aldakortasuna aztertzeko prozedura bat baino gehiago, datuen tankeraren arabekoak. Dena den, datozen orrietan soilik

azalduko ditugu zenbakien bidez adierazten diren datuei dagozkien prozedurak.

Aldakortasun-analisia aldakorrak diren ezaugarrien (aldagaien) analisia da. Aldakortasuna aztertzen dugunean bi datu multzo ditugu:

- ▣ Alde batetik aztertuko den osagaien multzoa, eta,
- ▣ bestetik, multzoaren osagaiei dagozkien balioak, beraz, ezaugarriaren (aldagaiaren) balioen multzoa:

Populazioa Euskadiko ibaiak		Ezaugarria = ALDAGAIA ibaien LUZERA (km)
 <p>Populazioaren osagaiak</p>	Nerbioi	43,8
	Cadagua	48,5
	Ibaizabal	43,5
	Deba	57
	Urola	51
	Oria	66
	Bidasoa	70
	Zadorra	79
		 <p>Aldagaiaren balioak</p>

▣ Ikasten ari garen ezaugarria/aldagaia ibaien luzera denez, ibaien luzeraei buruzko datuak aldagaiaren balioak dira.

ZERI DEITZEN DIOGU DATUEN LEHENENGO MIAKETA? ZERTAN DATZA?

Datuak ordenatzeari, berrikusteari eta grafikoetan ordezkatzeko lanari datuen lehenengo miaketa deritzogu. Horren helburua da egon daitezkeen hutsegiteak edo anomaliak aurkitzea hala nola datuetatik ahalik eta informazio gehien ateratzea. Lehenengo fase honetan ez dugu formula matematikorik erabiltzen, datuen miaketa hutsa edo ordezkazpen grafikoa baino.

Datuak ordenatu ondoren egiten dugun lehenengoa da balio handiena eta txikiena bilatzea. Lan erraza da, baita probetxuzkoa ere, aztertzen ari garen

aldagaiaren aldakortasunari buruz lehenengo informazioa ematen baitigu. Datu multzoaren balio handienaren eta txikienaren arteko diferentziari *ibiltartea* deritzogu.

Ondoren, datu multzoan akatsik egon litekeen begiratzen dugu. Lortuko ditugun emaitza guztiak datuetan oinarrituko direla ohartuz gero, erraz barneratuko dugu datuen baliagarritasuna ondo begiratzeko beharra.

ALDAGAIAREN ALDAKORTASUNA EZAGUTZEKO LEHEN PAUSOA: IBILTARTEAREN KALKULUA

Lehenago esan dugunez, aldagai baten aldakortasuna ezagutzen hasteko datu-multzoaren balio handienaren eta txikienaren arteko diferentzia kalkulatu dugu, aldagaiaren ibiltartea alegia. Beraz, guk ere bai, aldakortasunari buruzko azterketarekin hasteko, aldagai baten ibiltartea kalkulatu dugu. Horretarako, erabiliko ditugu *Euskadiko udalerrietako batez besteko errenta pertsonal erabilgarriari* buruzko datuak, 2001koak.¹²

Laburrean esanda, pertsonako errenta erabilgarria da gastatzeko edo aurrezteko norberak daukan diru kopurua. *Erabilgarria* den diru kopuru hori sortzen da pertsona fisikoek jasotzen dituzten mota guztietako diru-sarrerei zuzeneko zergatan ordaintzen dutena kenduz. (pertsona fisikoen errentaren gaineko zerga eta Gizarte Segurantzari ordaindutako kotizazioak)

Landuko dugun aldagaiari dagokionez, interesatzen zaigu ezagutzea zein neurritan diren aldakorrak errenta-balioak udalerrien artean. Esaterako, udalerrri guztietako edo gehienetako errenta nahiko antzekoa den edo ba ote diren, berriz, besteak baino errenta askoz altuagoa duten udalerrriak.

¹² Ikusi datu-zerrenda osoa eranskinean. 1. taula.

Lanestosa	7.192
Laukiz	20.627


Ikusten dugunez, Lanestosa eta Laukiz dira batez besteko errenta handiena eta txikiena duten udalerrriak. Biek duten errenta-balioen kenketa eginda 13.435 euroren aldea lortzen dugu. Hau dugu, beraz, 2001eko *Euskadiko udalerrietako batez besteko errenta pertsonal erabilgarria* aldagaiaren ibiltartea.


Dagoeneko lehen ideia dugu aztertzen ari garen aldagaiari buruz. Kasu honetan, gainera, esan dezakegu ikusgarria dela Euskadiko udalerrien arteko errenta-aldea, edo, bestela esanda, udalerrien arteko errentaren aldakortasuna handia dela: Laukizeko errenta, autonomia-erkidegoko handiena, Lanestosako errenta baina ia hiru aldiz handiagoa da.

ELEMENTU ARRAROAK ETA EGON DAITEZKEEN HUTSEGITEAK

Aldagaiaren balio handienaren eta txikienaren arteko aldea handia denean, Euskadiko errenta-balioen kasuan gertatzen den bezala, datu-zerrenda arretaz aztertu behar dugu. Helburua izango da aztertzea ikusitako balio-aldea egiazkoa **den** ala datuen hutsegite baten ondorioa **ote den**.

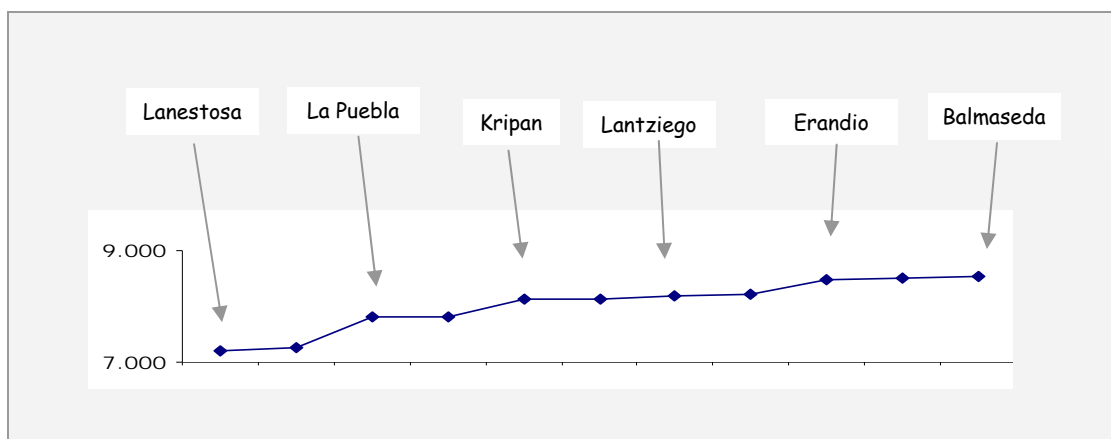
Estatistikan hasi berria den pertsona baten zalantza izan daiteke, beharbada, jakitea muturreko bi balioen arteko aldea susmagarria ote den. Horretarako, muturreko balioak beste balioekin konparatu behar ditugu:

 Bi balioetako edozein hutsegitea balitz, datuak jasotzean edo informatizatzean gertatutako akats baten ondoriozkoa, balio okerra beste balioen oso desberdina izango litzateke (salbuespen bat, alegia).

 Lanestosako eta Laukizeko balioak salbuespenak ote diren jakiteko datu-zerrenda aztertzen dugu bi udalerrietakoak bezalako baliorik ote dagoen.

Errenta-balio baxueneko udalerrriak	
Lanestosa	7.192
Karrantza Harana	7.275
Lapuebla de Labarca	7.800
Turtzioz	7.813
Kripan	8.131
Ekora	8.138
Lantziego	8.185
Harana	8.218
Erandio	8.492
Elantxobe	8.507
Balmaseda	8.547

Taulan ikusten dugunez, Lanestosako datuaren jarraian dauden balioek mailakako segida osatzen dute: balio batetik bestera igoera txikia da. Badira, beraz, Lanestosako errenta antzekoa –zertxobait handiagoa– duten beste udalerrri batzuk. Badira, esaterako, 8.000 euro baino gutxiago duten udalerrriak. Bi ardatzeko grafiko baten bidez argiago ikusiko dugu esandakoa:

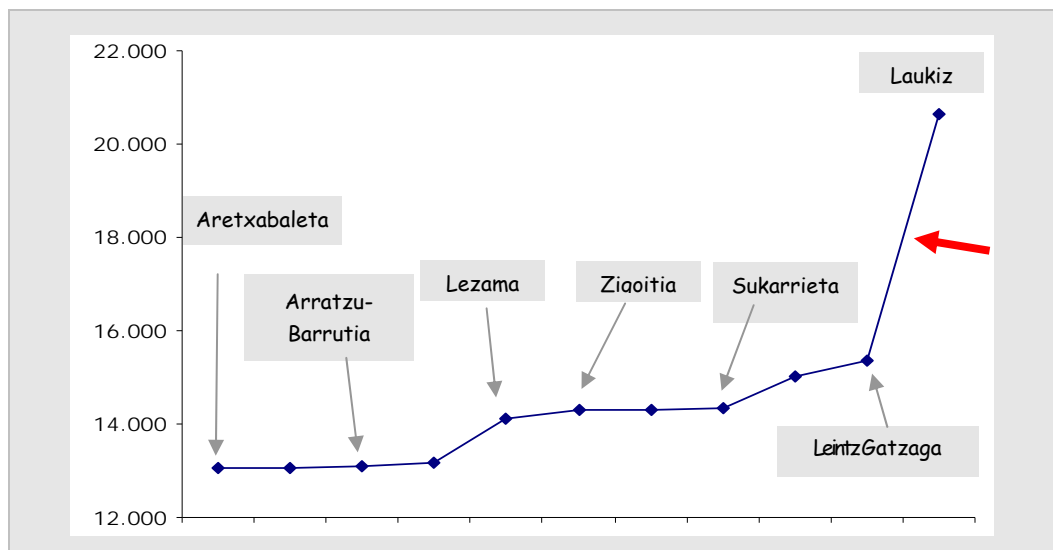


Grafikako lerroaren aldapa leunak islatzen du udalerrri batetik bestera dagoen alde txikia.

Ondoren, gauza bera egingo dugu errenta-balio altuenak dituzten udalerriekin:

Errenta-balio altueneko udalerriak	
Aretxabaleta	13.044
Legutio	13.069
Arratzu-Ubarrundia	13.100
Lezama	13.157
Burgelu	14.105
Getxo	14.291
Zigoitia	14.294
Gordexola	14.335
Sukarrieta	15.016
Leintz Gatzaga	15.346
Laukiz	20.627

Oraingo honetan ere, mailakakoa da udalerrri batetik bestera dagoen errenta-balioaren igoera. Azken bi udalerrien arteko aldea, berriz, askoz handiagoa da: Leintz Gatzagako eta Laukizeko balioen arteko aldea 5.281 eurorena da. Dirualde hori oso handia dela hobeto ulertzeko urte bateko pertsonen irabazia dela gogoratu behar dugu.



Oraingo honetan, grafikoak argi erakusten du, aldapa piko baten bidez, Leintz Gatzagako eta Laukizeko errenta-balioen arteko aldea.


Dagoeneko, prest gaude baieztatzeko Laukizeko errenta-balioa salbuespena dela, balio hori bezalakorik ez baitago. Lanestosako errenta-balioa, berriz, ez da salbuespena, berea bezalako beste balio batzuk daude eta.


Datu multzoko balioen artetik besteak baino nabarmen handiagoak edo txikiagoak direnei *elementu arraro* deritzegu.

Datu multzoetan agertzen diren elementu arraroak, berez, susmagarriak dira, baina horrek ez du esan nahi beti hutsegiak izan behar dutenik. Hori dela eta, Laukizeko balioa *elementu arraroa* dela ohartu ondoren, egiazkoa edo hutsegitea den egiaztatu behar dugu. Egia al da udalerrri horretako pertsonen errenta, 2001ean, beste udalerritakoena baino 5.281 euro handiagoa zela? Zalantza hori argitzeko geratzen zaigun bide bakarra da datu-bilketako prozesua berrikustea, jatorrizko datu-iturrira iritsi arte.

Gure kasuan datuen jatorria da Euskal Estatistika Erakundea (EUSTAT). Erakunde horrek bere webgunean eskaintzen dituen datu-zerrendak berrikusi ditugu: haien datua gure zerrendan daukagun datu bera da.

Honaino iritsita, datu-bilketan inolako hutsegiterik ez dela egon ikusirik, aukera batzuk ditugu:

 Hutsegitea dela susmatzeko bestelako arrazoirik edukiz gero eta konpontzeko biderik ezean, elementu arraroa datu-zerrendatik kentzea.

 Ondoren egingo ditugun analisi estatistiko guztiak bi eratara egitea: elementu arraroa hartuta eta kenduta.

Hartzen dugun aukera hartzen dugula, analisiaren txostenean gertatutako guztia azaldu beharra dago.

Euskadiko errentaren kasuan badugu Laukizeko datua egiazkoa dela pentsatzeko beste argudiorik, prentsan agertutako albiste bat, hain zuzen ere. 2007ko maiatzaren 12an, *El Correo* egunkarian, Laukizeko egoera berezia, errentari dagokionez, azaltzen zen.

Herri aberatsa, herri pobrea

Behi lodiak, behi argalak

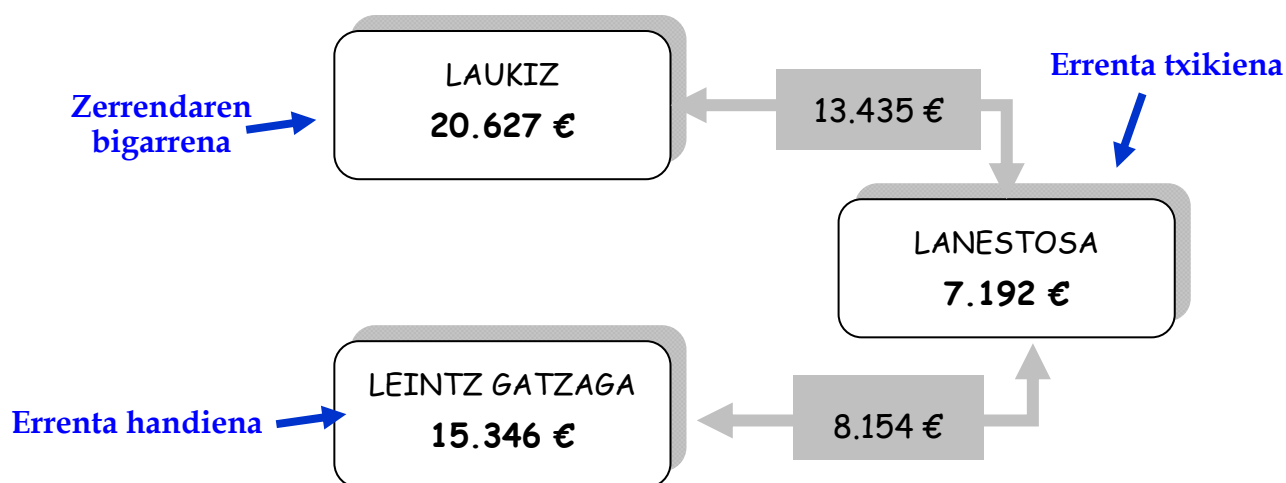
Laukizeko eta Karrantza Haraneko per capita errenten artean 15.000 euroren aldea dago. Udalerri hauek dira, nahi gabe, Bizkaiko aberatsena eta pobreena. «Ez horrenbeste, ezta horren gutxi ere, diote bertakoek».¹³

Egunkariko berriaren data (2007) eta gure datuen data (2001) desberdinak izan arren, Laukizeko egoera berezia dela pentsatzeko behar besteko argudiorik badago.

¹³ http://www.elcorreodigital.com/vizcaya/prensa/20070512/politica/vacas-gordas-vacas-flacas_20070512.html. 2008/06/20

BALIOEN BARNE-ALDAKORTASUNAREN ANALISIA

Dagoeneko argi dugu Euskadiko udalerrien arteko aldeak, errenta-balioei dagokienez, ez direla hasieran ematen zuten bezain handiak. Udalerri bakarra dago errenta bereziki handia daukana, eta horixe da ibiltartearen zabaleraren eragilea. Lanestosako errenta eta Leintz Gatzagakoa (datu-zerrendaren bigarrena) konparatuz gero, aldea 8.154 eurora jaisten da.



Arestian esan dugun bezala, ibiltarte da aldagaiaren aldakortasunarekiko lehen hurbilketa, baina ez digu ematen ibiltarte barneko aldakortasunari buruzko informaziorik. Euskadiko udalerrietako errentari buruz badakigu udalerri guztietako errentak, Lanestosakoa izan ezik, 7.192 eta 15.346 euroren artean daudela, baina ez dakigu nola banatzen diren errenta-balioak tarte horren barnean. Esaterako, nagusi al dira Lanestosakoa bezalako errenta baxuak ala errenta-balio handiagoetatik gertu daudenak (hots, Leintz Gatzagakoa bezalako altuak)? Edo, beharbada, ez bata ez bestea, balio gehienak kokatzen dira ibiltartearen erdialdean? Segidan ikusiko ditugu horien antzeko galderei erantzuteko baliabideak.

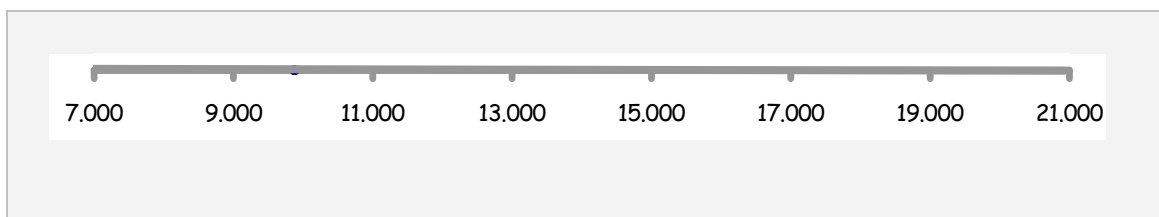
ALDAGAIAREN BARNE-BANAKETAREN KONTZEPTUA

Datu multzoaren balio gehienak antzekoak edo desberdinak ote diren ikusteko ordezkapen grafiko bat erabiliko dugu. Horren bitartez oso erraza izango da lehen egindako galderei erantzuna ematea.

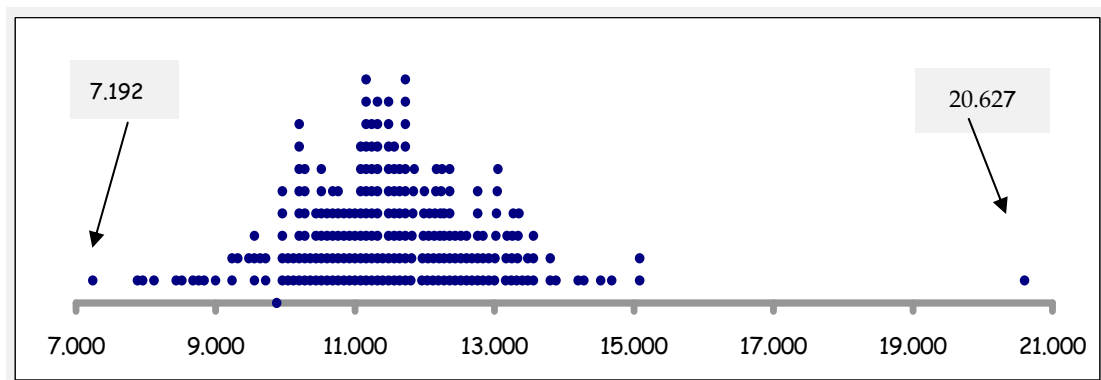
Irudikapen grafikoa egiten hasteko zuzen-segmentu bat marraztuko dugu. Segmentu horren bidez irudikatuko dugu aldagaiaren balio handienaren eta txikienaren arteko aldea (*ibiltartea*, alegia). Segmentuko muturretan egokituko ditugu aldagaiaren muturreko balioak.



Ondoren, segmentua zati berdinetan banatuko dugu; zati horiek errenta-balioen igoera berdinak irudikatzen dituzte. Eraitza eskala **graduatu**a izango da.



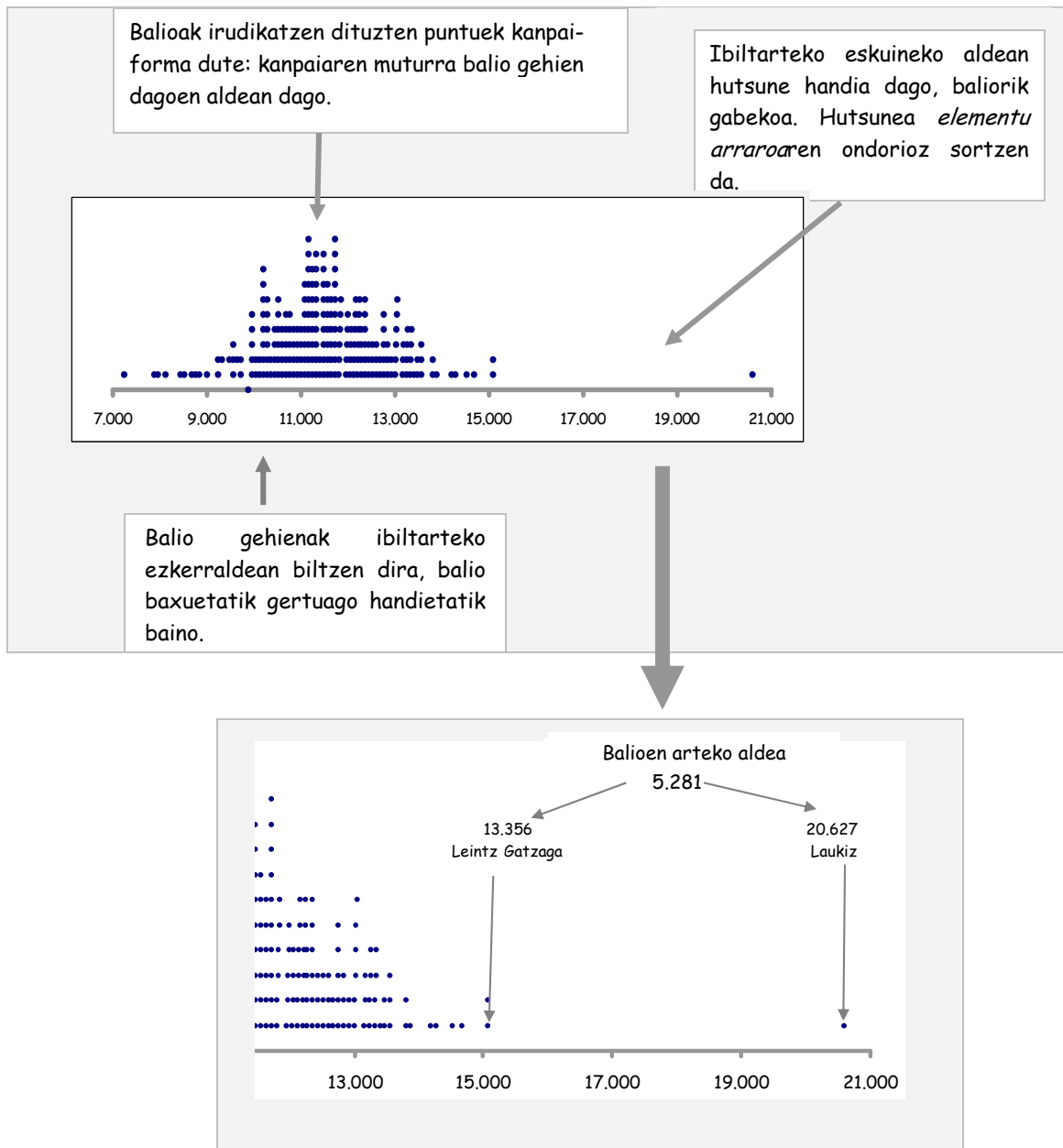
Egindako eskalaren gainean kokatuko ditugu errenta-datuei dagozkien ikurrak, bakoitza, bere balioaren arabera, eskalan dagoen lekuan. Datu bati dagokion lekua beteta egonez gero, ikur berria aurrekoaren gainean jarriko dugu. Eraitza izango da irudi honetan ikusten dugunaren antzekoa:



Lortu dugun irudiak erakusten du datu multzoaren banaketa (aldagaien balio-banaketa, alegia) ibiltartearen barnean. Hori dela eta, grafikoak erakusten duen informazioa aipatzeko, eskuarki erabiltzen da *aldagaiaren balio-banaketa* adierazpena.

Irudia sinplifikazio bat da. Beraz, ez dugu ahaztu behar grafikoari begira bakarrik lortuko dugula aldagaiaren balioei buruzko gutxi gorabeherako informazioa. Haren abantaila da, datu-zerrendarekin konparatuta, aldagaiaren balio-banaketari eta balioen arteko distantziari begirada batez antzemateko aukera ematen digula. Hitz batean, grafikoaren bidez balio-aldakortasuna ikusten dugu zuzenaren barneko distantzia gisa. Balioak elkarren artean gertuago edo urrunago egotea haien aldakortasunaren adierazlea da.

Irudiak balioen barneko antolaketa erakusten digunez, tresna egokia da planteatzeko zer-nolako informazioa atera daitekeen, hain zuzen ere, datuen banaketatik. Galdera modura esanda, zer-nolako informazioa atera daiteke *errenta-balioaren banaketari buruz* grafikoari begiratuta?

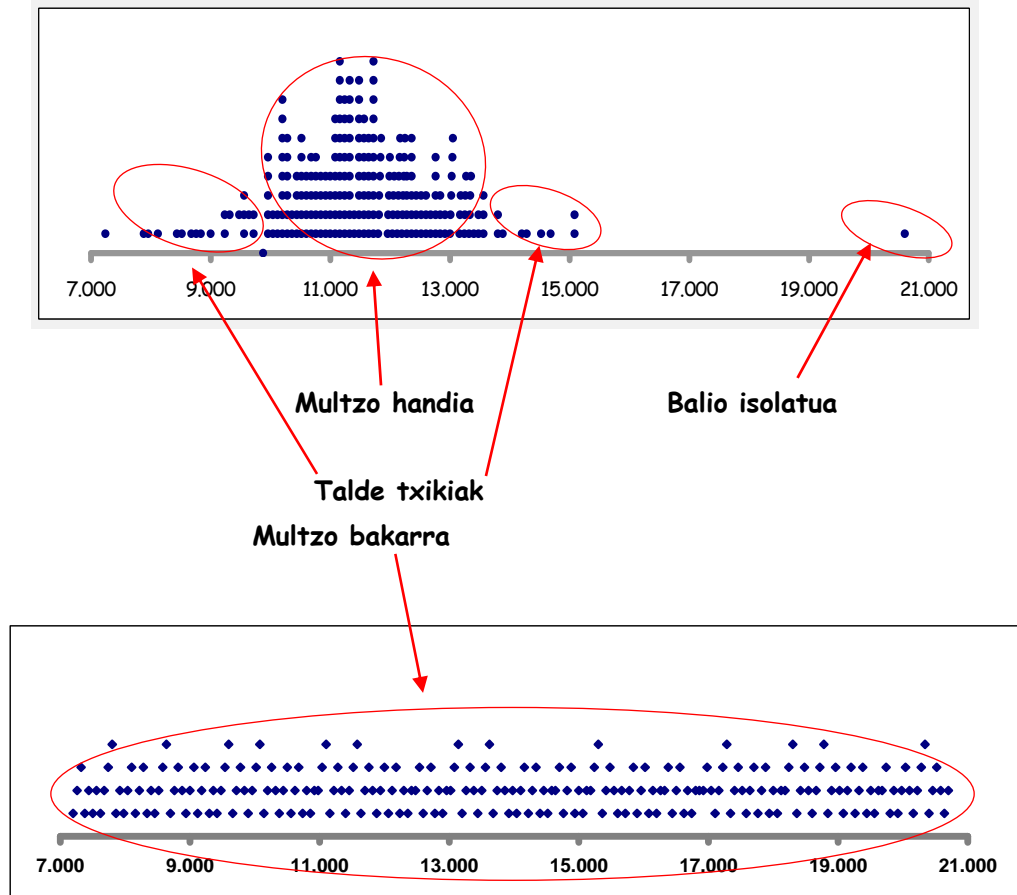


Datu multzo baten analisia egiten dugunean aztertzen dugun ezaugarri garrantzitsuenetariko bat balioen taldekatzea da, aldakortasunaren adierazle zuzena da eta. Zenbat eta handiagoa izan balioen taldekatzea, orduan eta txikiagoa da haien aldakortasuna, eta alderantziz.

Udalerrietako errentaren adibidean ikusi dugunez, taldekatze nabarmena dago. Errenta-balio batzuk *isolaturik* edo talde txikitan bilduta badaude ere, balioen

gehiengoa taldekatuta dago, multzo handi batean, ibiltarteko sektore batean.

Ondoren ikusiko dugu udalerrietako errentaren grafikoa egoera zeharo desberdina irudikatzen duen beste grafiko batekin konparatuta:




Oraingo irudi honetan ikusten dugun balioen banaketan ez dago inolako taldekatzerik. Balioak, berriz, *sakabanatzen* dira ibiltarte osoan zehar; horrek erakusten du ez dagoela balioen artean antzekotasunik.

BALIO-BANAKETAREN ANALISIA

Aurreko atalean erabili dugun grafikoa zeharo baliagarria da ibiltarte barneko balio-banaketari buruzko kontzeptuak barneratzeko. Grafiko hori tresna didaktikoa da, baina ez da analisi estatistikoetan datuen aldakortasuna aztertzeko eta deskribatzeko erabili ohi den grafiko mota bat:

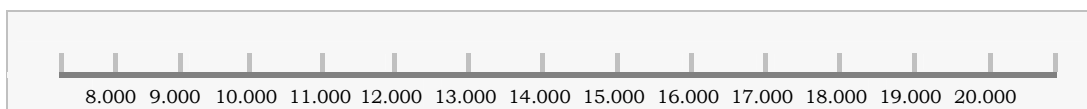
 Ez da eroso, ezta praktikoa ere, balioak, banan-banan, puntutxo bidez irudikatzea.

 Balio-banaketaren irudi orokorra ematen digu baina informazioa irakurtzeko zehastasuna falta zaio. Esaterako, ezin da zehaztu ibiltarteko zein puntutan – zer baliotan – hasten den balio-bilketa.

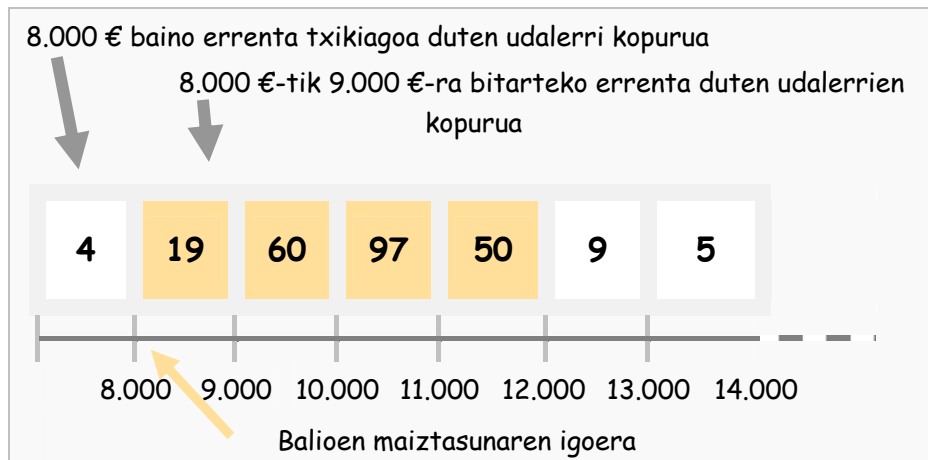
Hori dela eta, balio-banaketa aztertzeko erabil daitezkeen prozedurak eta hura irudikatze ohikoak diren grafikoak aurkeztuko ditugu segidan.

Aldagai baten balioen barneko banaketa aztertzen dugunean gure helburua izaten da datuen taldeko portaera ezagutzea. Jakin nahi dugu datuak elkartzen ote diren eta, elkartuz gero, ibiltarteko zer sektoretan gertatzen den. Ezaugarri horiek aztertzeko bide zuzena eta erraza da ibiltartea sektoreka miaztea, sektore bakoitzean zenbat balio biltzen diren ikusteko. Hobeto ulertzeko *errenta erabilgarriaren* datu-banaketa, Euskadiko udalerrietakoa, aztertuko dugu:

Errenta erabilgarria aldagaiaren ibiltartea milako sektore edo zatitan banatuko dugu:

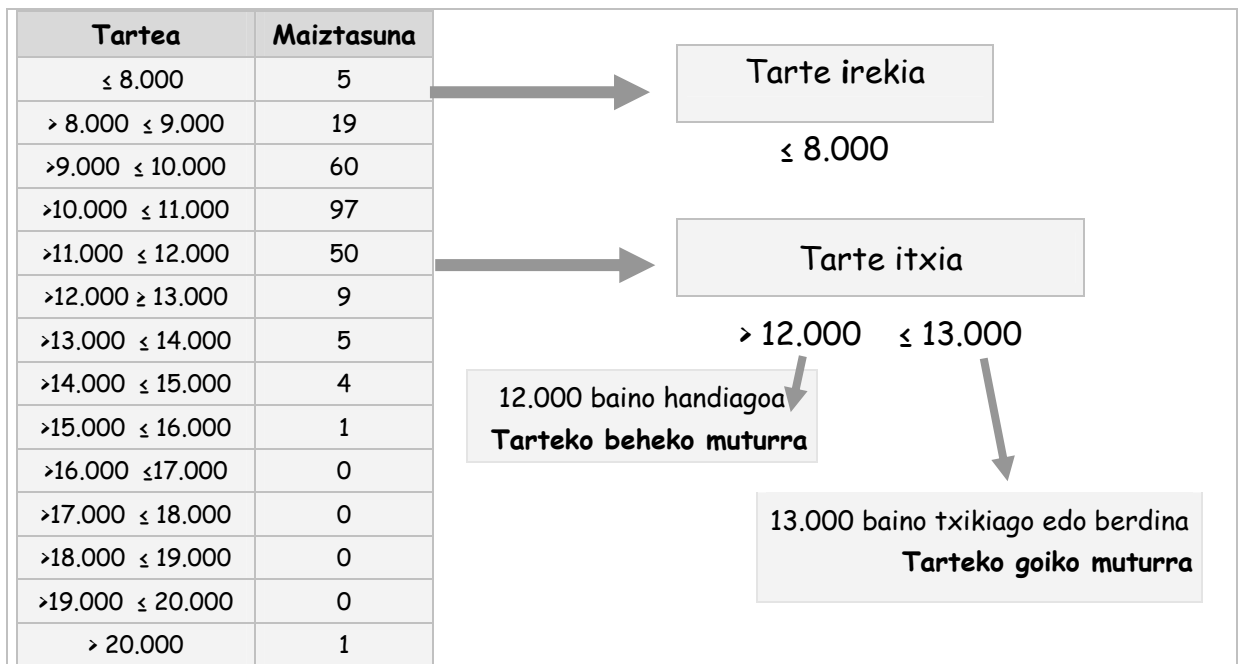


Ondoren, sektore bakoitzean biltzen den balio kopurua zenbatzen dugu:

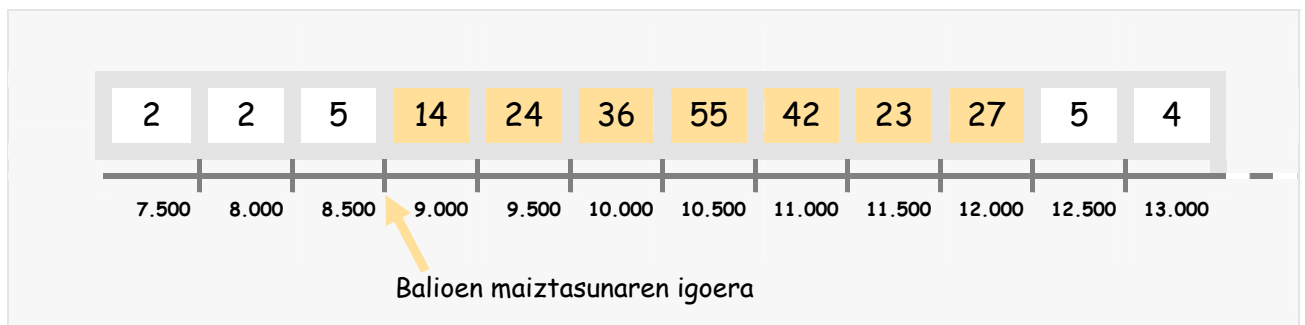


Aldagaiaren balioak ibiltarteko sektoreetan biltzen ditugunean balioen barneko banaketa *agertzen* da. Hori dela eta, erraz zehaztu daiteke ibiltarteko zer puntutan hasten den datu-maiztasunaren igoera: 8.000 €-ri dagokion puntuan datu-maiztasunaren aldaketa nabaria ikusten da; 4 udalerrri dira, soilik, 8.000 € baino errenta txikiagoa dutenak. Hortik aurrera maiztasun kopuruek modu nabarmenean gora egiten dute: ibiltarteko hurrengo sektoreak, 8.000 eta 9.000 euroko errenta duten udalerriei dagokiena, 19ko maiztasuna du.

Aldagaiaren balioak tartetan elkartzen ditugunean lortzen dugu *aldagaiaren maiztasun-banaketen taula* deritzona. Bi zutabeko taula da; ezkerreko zutabean ditugu aldagaiaren sektoreak edo tartearak, eta eskuinekoan, tarte horien barneko balioa duten datuen kopurua (maiztasuna, alegia).



Aurrekoan ibiltartea banatu dugu 1.000ko tartetan. Baina beste aukera batzuk bazeuden, esaterako, 500eko zatiak egitea, irudi honetan erakusten den bezala:




500eko tartetan sailkatzerakoan ibiltarte barneko balioen banaketa zertxobait aldatzen da. Oraingo honetan balioen maiztasunaren aldaketa ez da 8.000n gertatzen, 8.500ean baizik.

Tartea	Maiztasuna
> 7.500	2
$\geq 7.500 \leq 8.000$	2
$\geq 8.000 \leq 8.500$	5
$\geq 8.500 \leq 9.000$	14
$\geq 9.000 \leq 9.500$	24
$\geq 9.500 \leq 10.000$	36
$\geq 10.000 \leq 10.500$	55
$\geq 10.500 \leq 11.000$	42
$\geq 11.000 \leq 11.500$	23
$\geq 11.500 \leq 12.000$	27
$\geq 12.000 \leq 12.500$	5
$\geq 12.500 \leq 13.000$	4
$\geq 13.000 \leq 13.500$	5
$\geq 13.500 \leq 14.000$	0
$\geq 14.000 \leq 14.500$	4
$\geq 14.500 \leq 15.000$	0
$\geq 15.000 \leq 15.500$	2
$\geq 15.500 \leq 16.000$	0
$\geq 16.000 \leq 16.500$	0
$\geq 16.500 \leq 17.000$	0
$\geq 17.000 \leq 17.500$	0
$\geq 17.500 \leq 18.000$	0
> 18.000	1

Balio handien aldetik berriz maiztasunaren aldaketa lehengo puntu berean gertatzen da, 12.000 euroan hain zuzen ere.

Aldagaiaren maiztasun-banaketen taulak egiteko prozedura ondo ulertzeko eta erabiltzeko hainbat oinarri argi eduki behar ditugu:

 Jatorrizko datu-zerrendaren (informazioaren) sinplifikazioa da, hasierako datuen multzoa murrizten da eta. Erabilitako prozeduraren bidez, errentaren adibidean ikusi dugun bezala, oso luzea den datu multzo baten ordez askoz txikiagoa den datu-taula bat lortzen dugu. Begi-ukaldi batez identifikatzen ditugu nagusitzen diren balioak edo oso urriak direnak. Irabazten dugun aukera horren ordainez galera bat ere badago: maiztasun-taulan banakako datuen informazioa ez dugu ikusten, ezta datu bakoitza populazioaren zein elementuri dagokion ere.

Jatorrizko datu-zerrenda		Maiztasun-banaketaren taula	
Populazioaren elementuak	Balioa	Tartea	Maiztasuna
Lanestosa	7.192	> 7.500	2
Karrantza Harana	7.275		
Lapuebla de Labarca	7.800	$\geq 7.500 \leq 8.000$	2
Turtzioz	7.813		
Kripan	8.131	$\geq 8.000 \leq 8.500$	5
Ekora	8.138		
Lantziego	8.185		
Harana	8.218		
Erandio	8.492		

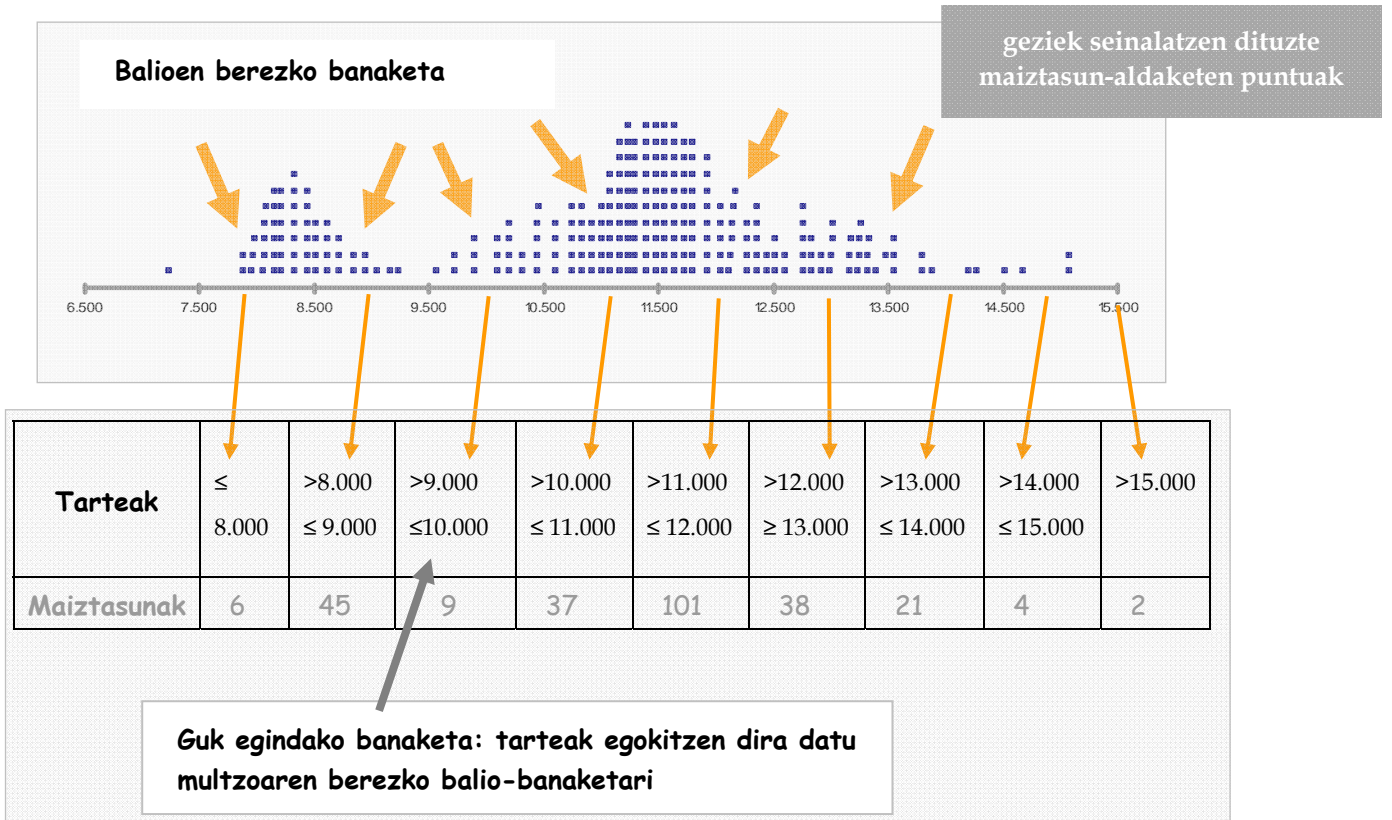
Maiztasun-banaketaren taula bat egiten dugunean, tarteen zabalera eta, ondorioz, tarte kopurua guk aukeratzen ditugu. Tartek txikiagoak edo handiagoak izan daitezke, aurreko adibidean ikusi dugun bezala. Eta hala bada, nola jakin zenbat multzo egin behar dugun? Lehenengo eta behin esan behar da ez dagoela tartek egiteko inolako araurik. Tarte kopurua aukeratzeko hoberena da zenbait saiakera egitea, balioen barne-banaketa modu egokian erakusten duen sailkapen bat aurkitu arte.

Balioen sailkapen egoki bat izango da erakusten duena ibiltarteko zer puntutan gertatzen diren balio-maiztasunen aldaketak. Aipatutako puntu horiek banaketaren berezko etendurak dira; egiten ditugun tartek, ahal den neurrian, etendurak erakutsi beharko dituzte.

Errentaren maiztasun-taula egiteko bi saiakera egin ditugu, bakoitzean zabalera bateko tartek aukeratuz. Oro har, zabalera txikiagoko tartek zehatzago erakusten dituzte banaketaren berezko etendurak.

Ondoko grafikoa asmatutako balio-banaketa baten irudikapena da. Bertan, geziek seinalatzen dituzte maiztasun-aldaketen puntuak, hots, **banaketaren berezko etendurak**. Horiek dira tarteen mugak ezartzeko, ahal den neurrian, erabili behar diren ibiltarteko puntuak. Azken finean egin behar duguna zera da: datu multzoaren berezko balio-banaketa eta maiztasun-taularen banaketa

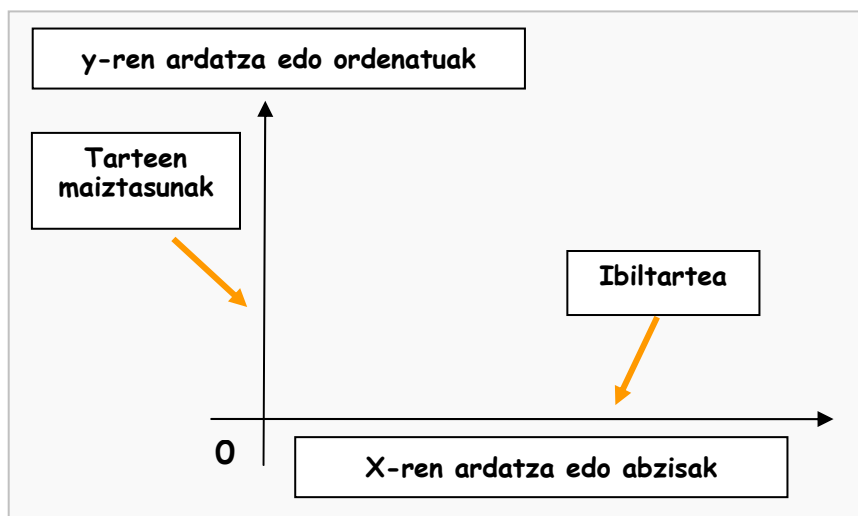
bat egin edo, bestela esanda, bigarrena lehenengoari egokitu.



MAIZTASUN-BANAKETAREN IRUDIKAPEN GRAFIKOA: HISTOGRAMA

Histograma grafiko mota bat da, aldagai baten ibiltarte barneko balioen maiztasun-banaketa erakusteko erabiltzen dena. Tresna grafikoa da, banaketaren ezaugarri nabarmenak ahalik eta adierazgarrien irudikatzeko. Histogramaren bidez ikusizko informazioaren irakurketa arina bilatzen da.

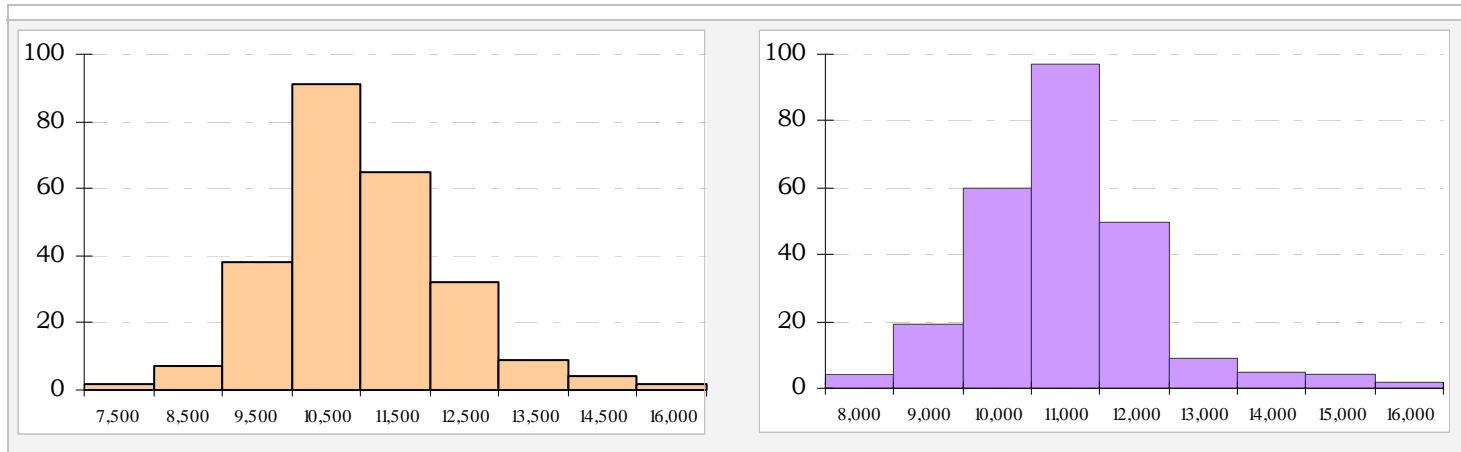
Histogramak egiteko aurreko atalean ikusi ditugun balio-maiztasunaren taulak erabiltzen dira. Grafikoaren oinarria ardatz kartesiarra da:



Irudian ikusten den bezala, ardatz kartesiarrek bi lerro dira 0 balioari dagokion puntuan gurutzatzen direnak. **Ardatz horizontalean** aldagaiaren ibiltartea irudikatzen da, tartetan banatuta; **ardatz bertikalean** tarteen maiztasunak.

Esan daiteke histograma egitea erraza dela, baldin eta, aldezturik, maiztasun-banaketaren taula egitean behar diren erabakiak hartu baditugu. Lehenago ikusi dugun bezala, datu-multzoaren ezaugarriak ondo erakusteko, garrantzitsuena da tarte kopurua eta etendura egokiak aukeratzea. Berrituz esango dugu egileari dagokiola erabaki horiek hartzea eta egiten dituen aukeren arabera histogramak eskaintzen duen irudia zertxobait alda daitezkeela. Errentaren maiztasun-banaketaren taulak egiterakoan zabalera bateko eta besteko tarteak egiteak nolako emaitzak dakartzan ikusi dugu. Ondoren ikusiko ditugu Euskadiko udalerrietako errentaren datuekin egin daitezkeen histogramak:

Euskadiko udalerrietako batez besteko errenta erabilgarria. 2001

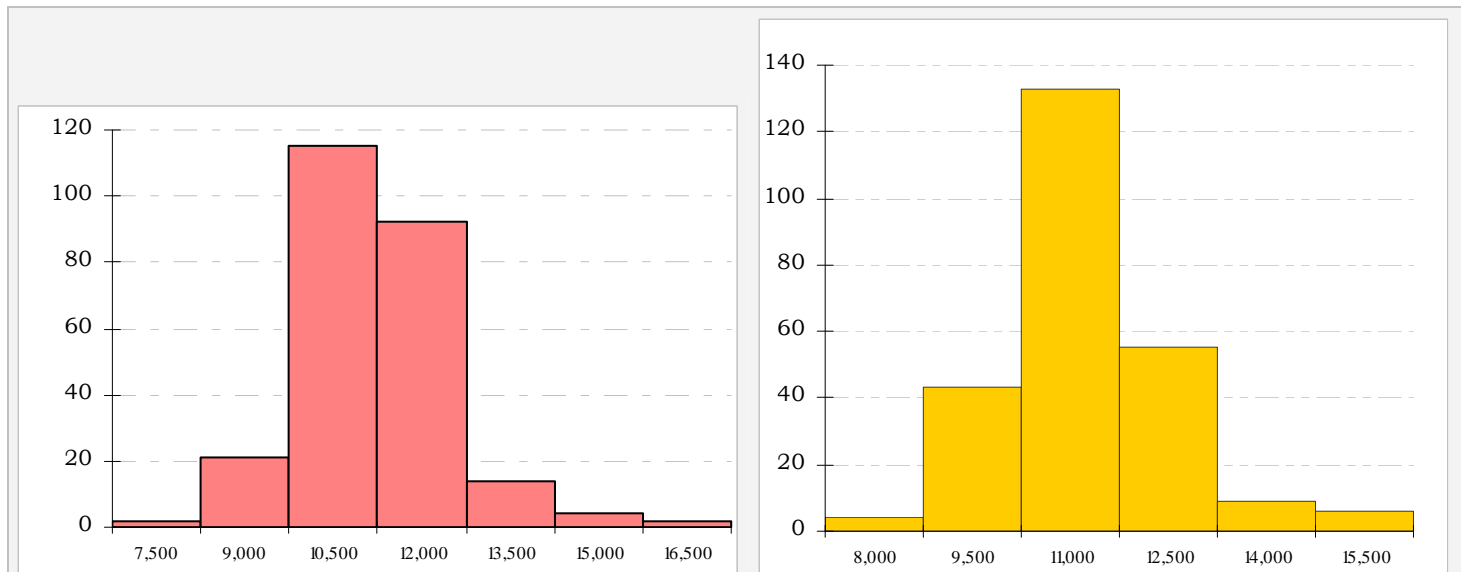


1.000ko zabalerako tartekak

Hasierako tartea: 6.500 -7.500 €

Hasierako tartea: 7.000- 8.000 €

Bi histogramen hasierako puntua desberdina izanda, bakoitzak eskeintzen duen irudia zertxobait aldatzena da: ezkerreko histogramari begiratuta, ematen du balio gehiago biltzen direla grafikoaren eskuineko partean, balio handien aldean. Eskuineko histograman kontrakoa gertatzen da.

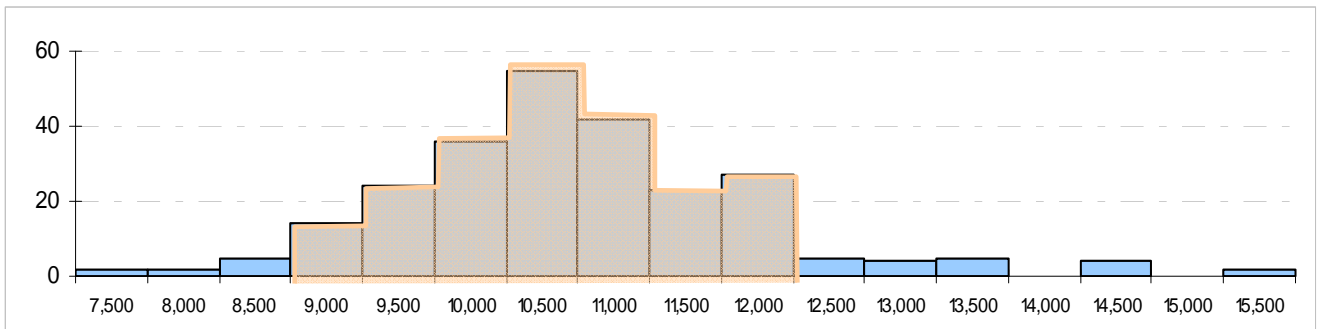


1.500eko zabalerako tartekak

Hasierako tartea: 6.000- 7.500 €

Hasierako tartea: 6.500- 8.000 €


Histograma hauetan, goikoekin konparatuta, tarteen zabalera handiagoa denez, datu-bilketa handiagoa dela ematen du. Taldekatze handiagoaren itxura areagotu egiten da eskuineko histograman.




500eko zabalera-ko tartekak

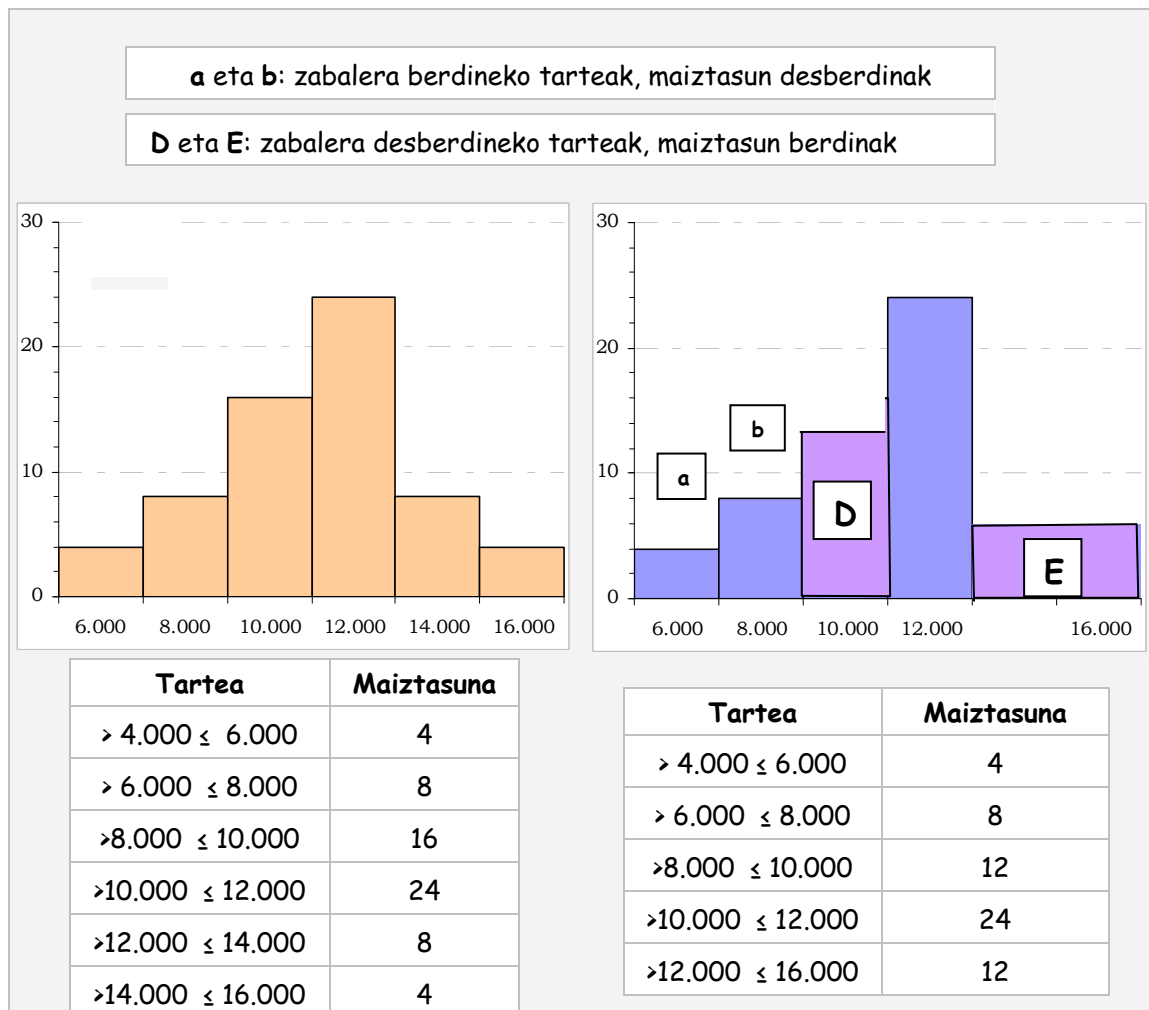
Egin ditugun histogrametan askotariko tarte kopuruak eta tarte-zabalera erabili ditugu; lehenengo tarteen hasierak ere desberdinak izan dira: batzuk 6.000 €-an hasita, besteak 6.500 €-an. Begi-bistan dago irudiak desberdinak direla. Horien aurrean galdera batzuk sortzen dira: Zein izan daiteke histogramarik egokiena? Tarte gutxiago ala gehiago egin?

Ez dago egoera guztietarako erantzun bakarra, aztertzen ari den datuen ezaugarrien arabera aukera egokiena desberdina izan daiteke eta. Aztertzen ari garen adibideari dagokionez zera esan daiteke:

 500 €-ko zabalera duen histogramak zeharo ondo erakusten du ibiltarteko zer puntutan gertatzen diren balio-maiztasunaren aldaketak. Udalerri gehienetako errenta erabilgarriak 8.500 eurotik 12.000 €-ra bitartean daudela bistakoa da.

 1.000 €-ko zabalera-koak, 8.000 €-an hasten denak, askoz sinpleagoa izanda ere, maiztasun-banaketa-aren ezaugarri nagusiak erakusten ditu. Alabaina, 8.500 euroan dagoen etendurari buruzko informazio galdu egiten da.

Proposatu ditugun histograma guztiek badute ezaugarri komun bat: histograma bakoitzeko tarteen zabalera berdina da, zabalera berdineko zutabeak ere egin ahal izateko. Hori ulertzeko ezagutu behar da zutabeen tamainaren esanahia. Zutabeen balioa irakurtzeko ez zaio begiratu behar zutabeen altuerarik azalerari baizik. Irudi baten bidez argiago ulertuko da:



Ezkerreko histograman zabalera berdineko tartekak ditugu, denak 2.000koak. Hori dela eta, tarteen arteko konparazioa egiteko, tarteen altuerari begiratu behar diogu bakarrik: zutabeen altuerak maiztasunekiko proportzionalak dira. Esaterako, bigarren zutabearen maiztasuna, lehenengoaren bikoitza denez, altuera ere bikoitza da. Hirugarren zutabearena ere, aurrekoaren bikoitza...

Eskuineko histograman, berriz, zabalera desberdina daukan tarte bat dugu: 12.000 €-tik 16.000 €-ra bitartekoa (E), beste guztien zabalera bikoitza duena. Bestalde, (E) tartearen maiztasuna 12koa da, (D) tartearena bezala. Bi tarte horien maiztasuna berdina izan arren, zutabeen altuerak desberdinak dira. Bi tarteen azalera, berriz, berdina da. Ondorioz, esan dezakegu maiztasun berdina duten tarteei azalera berdineko zutabeak dagozkiela.

Ez dugu gaia bukatu nahi histogramei buruzko ideia nagusiak berriro gogoratu gabe:

- ❖ Histogramak egiteko hartzen diren erabakien arabera, datuen banaketari buruzko irudia oso desberdina izan daiteke. Desberdinak izan daitezke tarte kopuruak eta, ondorioz, haien zabalera, lehenengo tartearen hasierako puntua, zabalera berdineko edo zabalera desberdineko tarteak...
- ❖ Ez dago araurik tarte-zabalerari eta tarte kopuruari dagokienez. Hori dela eta, egilea bera da erabaki behar duena, datuen ezaugarriak aztertu ondorenean, ibiltarteko zer etendura irudikatu.
- ❖ Histogramak egiten direnean, komenigarriena da saiakera batzuk egitea, datu-banaketa egokien irudikatzen duen histograma bat aurkitu arte.

NOLA AZALDU ETA DESKRIBATU ALDAGAIEN ALDAKORTASUNA

Datu multzo baten lehenengo miaketa eta ibiltarte barneko balioen banaketa aztertu ondoren, datu-analisiaren egileak nahiko ondo ezagutzen ditu datu horien banaketaren ezaugarri nagusiak. Euskadiko udalerrietako errenta-datuen hasierako miaketa egin ondoren, datu horien ezaugarri nagusiak deskribatzeko prest gaude. Argi eta garbi nabarmentzen dira ezaugarri hauek:

- Euskadiko udalerrietako batez besteko errentaren aldakortasuna, ibiltarte osoari begiratuz gero, nabarmen handia da.
- Aldagaiaren ibiltarte horren handia izatearen arrazoia: *elementu arraroa*, bereziki handia, dagoela.
- Udalergi guztien arteko aldea (aldakortasuna, alegia) handia den arren, badago udalergi kopuru handi bat errenta antzekoa duena. Azken finean, gutxi dira errenta-balio oso handia edo oso txikia dutenak.
- Udalergi askotako errenta 8.500 €-tik 12.000 €-ra bitarteko tartean dago.
- Badaude udalergi gehiago ibiltarteko balio baxuen aldean altuen aldean baino.

Kontua da orain nola osatu egoeraren irudi bat; nola egin aldagaiaren deskribapen sintetikoa eta, azken finean, nola komunikatu aipatutako ezaugarriak gai horretaz jakin nahi dutenei. Goian egin dugun deskribapena ez da inolaz ere egokia, aipamen gehienak zehaztu gabeko baieztapenak dira eta. Lortu behar dugun deskribapenak hainbat baldintza bete behar ditu:

- Deskribapenaren baieztapenek tresna edo prozedura estatistiko ezagunetan eta onartuetan oinarrituak izan behar dute. Erabilitako prozeduren emaitzak ere aipatu behar dira. Esaterako, aldagaiaren ibiltarte handia dela esateak ez du balio, baldin eta ibiltartearen balioa edo balio handienaren eta txikienaren balioak ez badira aipatzen. Ibiltarte tresna estatistikoa da eta gaia ezagutzen duen edonork badaki ibiltarte zer den eta nola kalkulatu den.

- Datuen ezaugarriak aztertzeko eta deskribatzeko, erabiltzen diren prozedurek datuen izaerarako egokiak izan behar dute. Prozedura estatistiko guztiak ez dira mota guztietako datuetarako egokiak.
- Deskribapenak datu multzoaren eta banaketaren ezaugarri adierazgarri guztiak kontuan hartu behar ditu.

Estatistikak baditu datu multzoen aldakortasuna aztertzeko eta deskribatzeko baliabide ugari, mota askotako datuak eta aldagaiak miatzeko egokiak direnak. Baina Estatistikaren tresna deskribatzaile guztiak ezin dira erabili egoera guztietan, eta, horregatik, datu multzoen ezaugarrien arabera tresna egokiak aukeratu behar dira. Kontua da, beraz, baliabide egokiak aukeratzen jakitea, egoera bakoitzean dugun datu multzoaren arabera jokatzeko gaitasuna edukitzea.

Aukera egokien garrantzia barneratzeko egiten ditugun deskribapenen eragina argi eduki behar dugu. Hasteko, argi eta garbi eduki behar dugu deskribapenak irudikatzen duela jatorrizko datu multzoa. Gehienetan, datu multzo baten deskribapen estatistikoak entzuten edo irakurtzen dituzten pertsonak ez daukate jatorrizko datuak ikusteko aukerarik. Ondorioz, pertsona horiek egoerari buruz aterako dituzten emaitzak deskribapenean oinarrituak izango dira. Datuen irudikapena (deskribapena, alegia) ona edo txarra, egokia edo desegokia izanda, egoerari buruz zabaltzen den irudiaren arduraduna da.

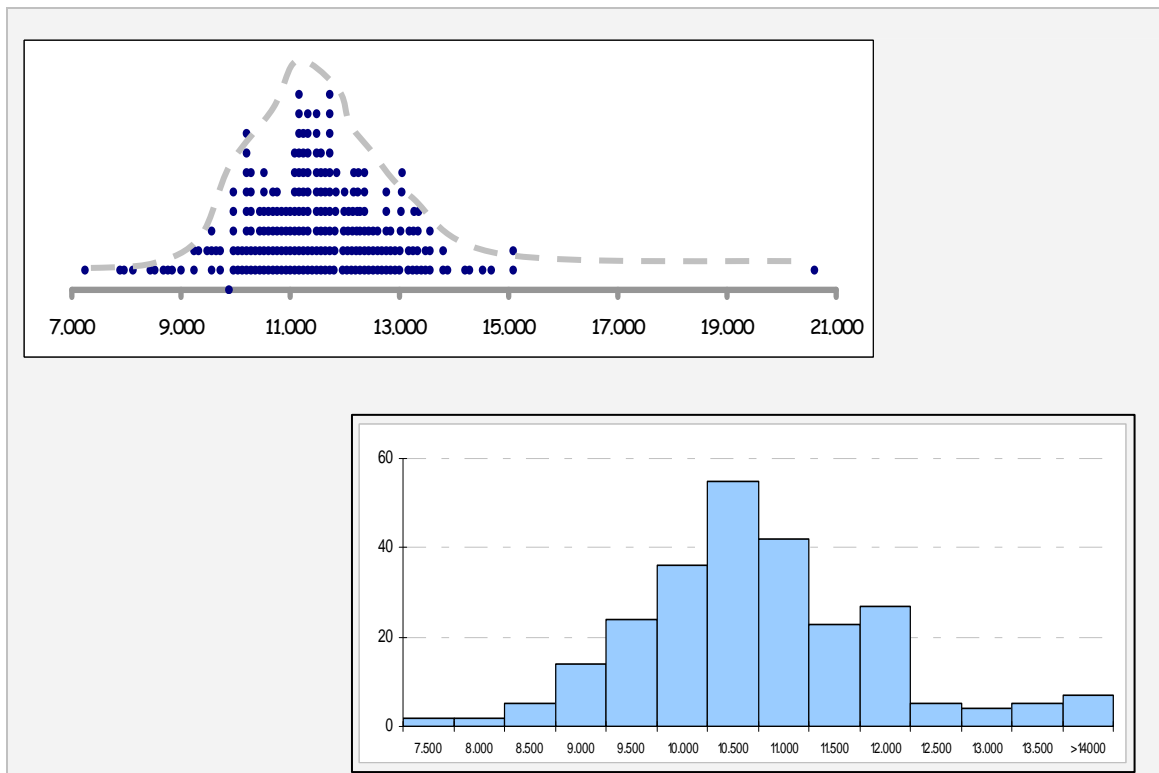
Ezin dugu ahaztu adituen analisiek jendearen iritzia bideratzeko gaitasuna izaten dutela. Zinema-kritikarien analisiek filmei buruz zabaltzen diren irudietan eragina duten bezala, errealitateko egoerei buruzko analisi estatistikoek jendearen iritzietan eragina izateko gaitasuna izaten dute. Herri-administrazioak, esaterako, ekonomiari, gizarteari, hezkuntzari eta beste arlo askori buruzko txostenak zabaltzen ditu etengabe. Txosten horiek izaten dira aipatutako gai eta egoerei buruzko deskribapenak, datu multzoen analisi estatistikoetan oinarrituak. Deskribapen horien eragina ukazina da, askotan errealitateko hainbat gairi buruz daukagun informazio bakarra izaten da eta. Ildo horretatik, esan genezake estatistika-analisiak izaten direla errealitatearen eta gizakien arteko bitartekariak.



TRESNA DESKRIBATZAILEAK AUKERATZEKO IRIZPIDEAK

Datu multzo baten aldakortasuna aztertzeko erabil daitezkeen tresna eta prozedura estatistikoak aldatu egiten dira datuen ezaugarrien arabera. Aztertzen ari garen datuetarako egokiak diren prozedurak aukeratzeko irtenbide erraza eta zuzena da datuen banaketari begiratzea.

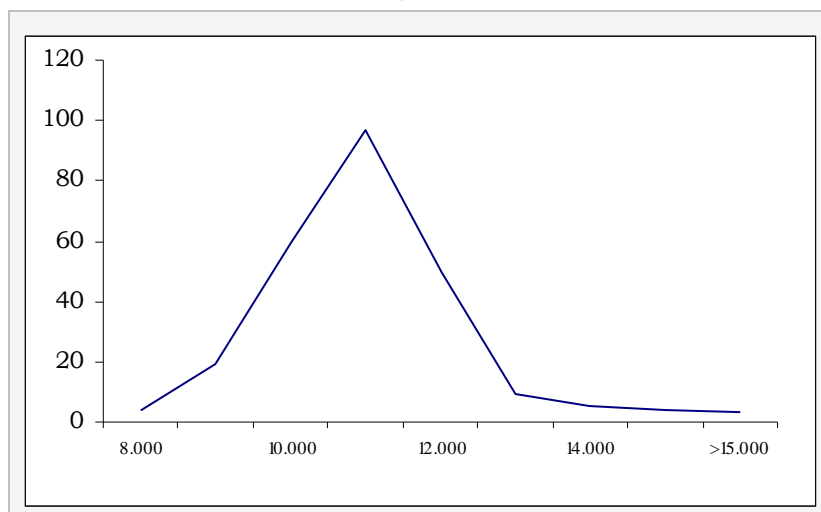
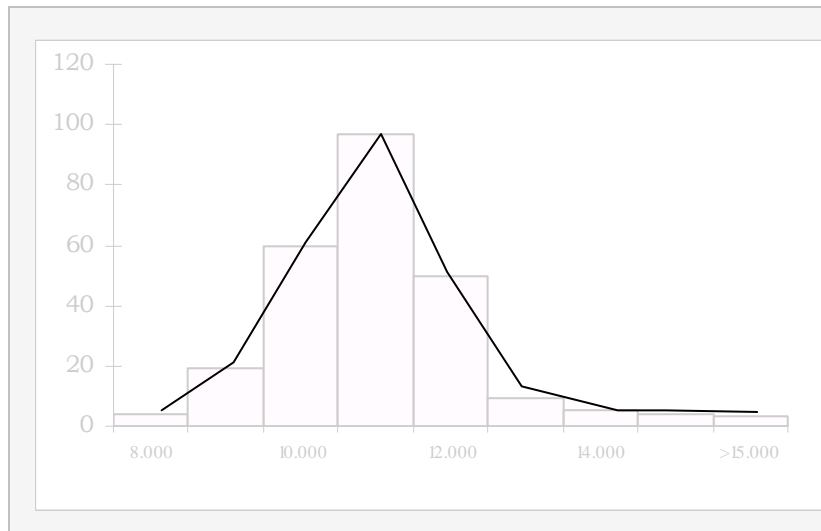
Datu-banaketaren kontzeptua azaldu dugunean, adibide gisa Euskadiko udalerrien errenta balioak erabili ditugu. Ondoren, histogramari buruzko azalpenak datu berdinekin egin ditugu. Bi kasuetan eduki dugu aukera ikusteko datu horien banaketak hartzen duen forma, *kanpai*-itxurakoa.



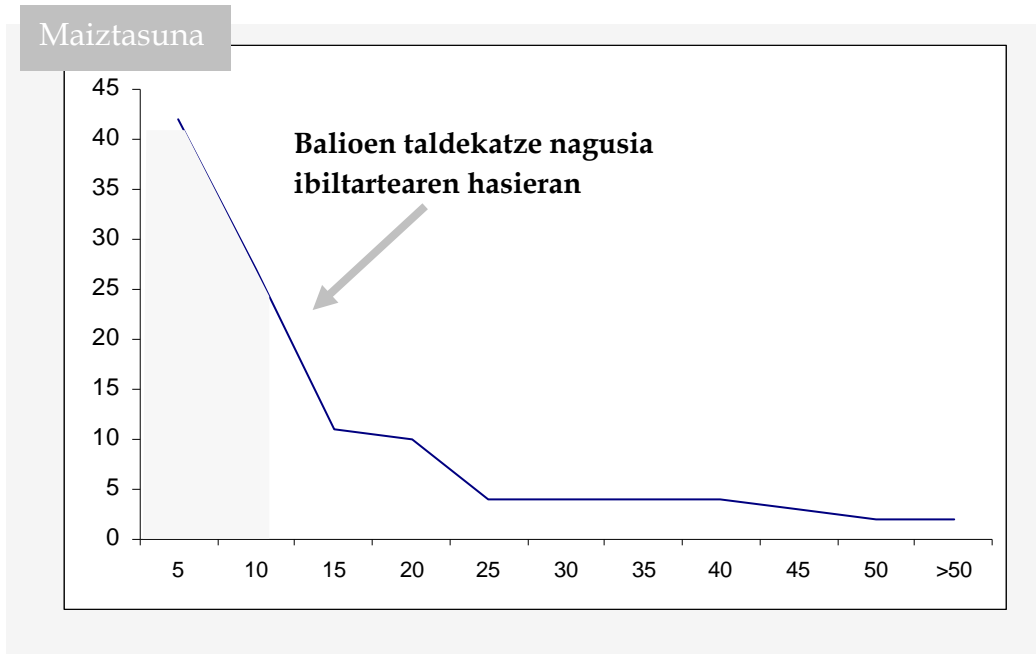
Bi grafikoaren inguradako marra baten bidez inguratuko bagenu, argiago ikusiko genuke kanpaiaren itxura.

Datu-banaketaren forma histogramarekin baino argiago ikusteko, histograma bera erabiliko dugu beste grafiko bat sortzeko, *maiztasun-poligonoa* izeneko grafikoa. Maiztasun-poligonoaren grafikoa egiteko prozedura simplea da.

Histogramaren zutabe guztien goiko muturren erditik pasatzen den lerro bat egitea nahikoa da *maiztasun-poligonoa* lortzeko.



Aldagai askoren balioak kanpai-itxura duen banaketara egokitzen diren arren, badira aldagai asko ere bestelako formak dituztenak. Azken finean, banaketen forma datu multzoen ezaugarrien isla da, edo, bestela esanda, datuen taldekatzearen ondorioa.

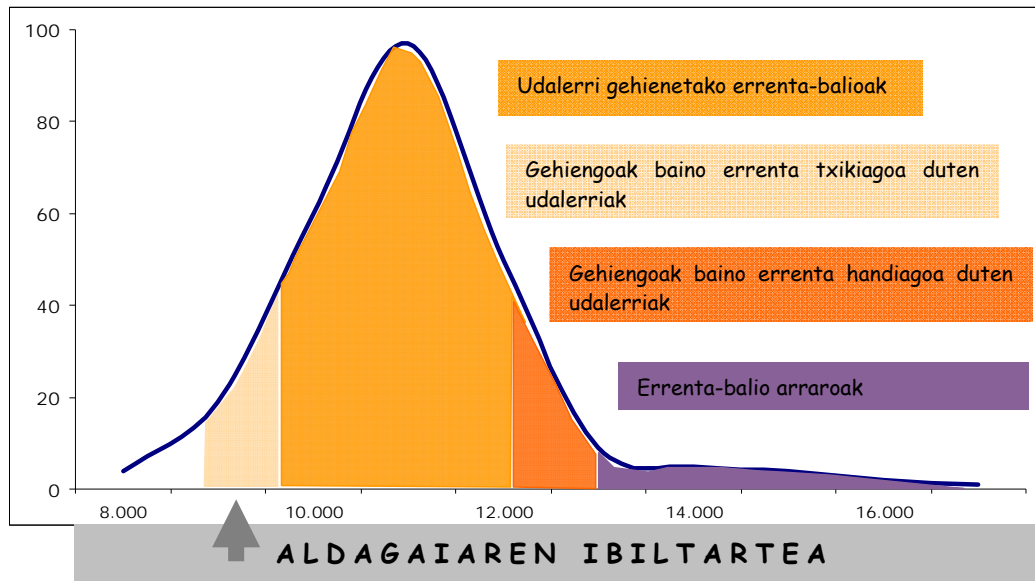


Goian ikusten dugun grafikoak (maiztasun poligonoak) ez dauka kanpai-itxurako banaketaren antzik batere, balioen taldekatzea ere desberdina delako. Banaketa forma horri *alderantzizko J* deritzo. Banaketa-formaren eredu horrekin bat datozen balioen taldekatze nagusia ibiltartearen hasieran gertatzen da, aldagaiaren balio baxuenen aldean.

Aldagaien aldakortasun-azterketari dagokionez, balioen banaketen formak bi multzotan banatuko ditugu: kanpai-itxura dutenak eta bestelako formak dituztenak. Biak aztertzeko prozedurek desberdinak izan behar dute. Datozen orrietan, kanpai-itxurako balioen banaketaz arduratuko gara.

KANPAI-ITXURA DUTEN BANAKETEN ANALISIA

Begira diezaiogun, berriro, Euskadiko udalerrietako batez besteko errenta erabilgarriaren maiztasun-banaketari (2001):

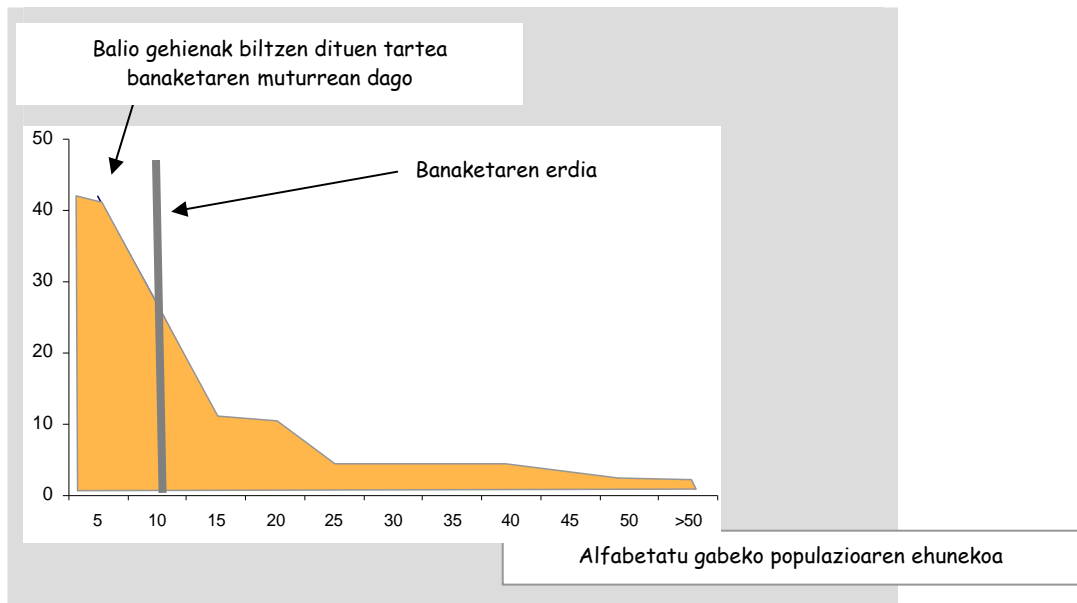
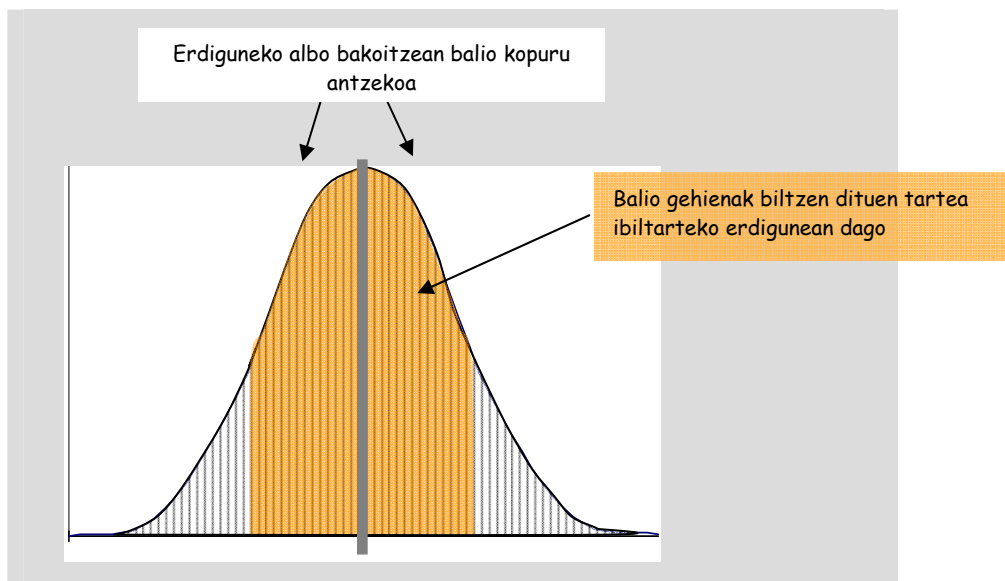


Koloreen bidez bereizi ditugu balio-azpimultzoak, tresna estatistikoaren bidez deskribatu nahi ditugunak:

- Datu multzoaren balio gehienak jasotzen dituen azpimultzoa.
- Udalerrri gehienetako errentatik gora eta behera dauden udalerrrien azpimultzoa.
- Errenta-balio handienak dituztenen azpimultzoa.

KANPAI-ITXURAKO BANAKETEN ERDIGUNEA. BATEZBESTEKOA ETA MEDIANA

Balio gehienaren egoera deskribatzeko lehen pausoa da banaketaren erdiguneari dagokion balioa kalkulatzeko. Kanpai-itxurako banaketetan bat datoz balio gehienak biltzen dituen tartearen erdigunea eta banaketaren erdigunea.



Badira hainbat prozedura datu-banaketaren erdigunea kalkulatzeko. Dauden

prozeduretatik batezbestekoa da, inolako zalantzarik gabe, gehien erabiltzen dena. Batezbestekoa kalkulatzeko erraztasuna dela eta, komenigarria ez denean ere erabiltzen da. Arazoa da, beharbada, kalkulu errazak zailagoa den beste zerbait ezkututzen duela. Batezbestekoaren kalkulua baino zailagoa da jakitea noiz erabil daitekeen batezbestekoa datu-banaketa erdigunea kalkulatzeko.

Abiapuntu gisa, baita printzipio orokor gisa ere, esan daiteke batezbestekoa erabil daitekeela banaketen erdigunea finkatzeko **soilik** kanpai-itxurako banaketen kasuetan. Datozen orrietan emango dira azalpenak, egindako baieztapena ulertzeko.

Orain arte eman diren azalpenetan ez ditugu ikur estatistikoak erabili, ideiak ulertzeko oztopo izan ez zitezen. Hemendik aurrera ezinbestekoa izango da ikur estatistikoak *ikustea*, erabiltzen diren prozedurak azaltzeko. Horretarako, baita ikur estatistikoak irakurtzea erraza dela erakusteko ere, gaiaren oinarriak azalduko ditugu ondoren.

Notazio estatistikoaren oinarriak

■ Estatistikan, aldagai bati buruz hitz egiteko, oro har, **X** letra erabiltzen da, larriz idatzita. Ikur edo letra orokorra erabiltzen da aldagaiaren kontzeptua aipatzeko, edozein aldagairi dagokion informazioa azaldu nahi denean. Demagun, adibide gisa, honako esaldia:

Aldagaien balioak batu behar dira...

Aldagaiaren kontzeptua aipatzeko dagokion ikurra erabiliko dugu orain:

X-ren balioak batu behar dira...

Modu horretan eginda, ulertu egiten da aldagaia edozein izanda ere haren balioak batu behar direla.

■ Aldagaiaren balioak aipatzeko, edozein aldagaiarenak, **x** letra erabiltzen dugu, kasu horretan xehez idatzita. Aldagaiaren edozein balio aipatu nahi denean, azpiindizea gehitzen zaio: **x_i**

Aldagaiaren edozein balio hartuz gero...

X-ren edozein balio hartuz gero

X-ren x_i hartuz gero

Aldagaiaren balio batzuen aipamena egin nahi denengan x letrari gehitzen zaio zenbaki bat, azpiindize moduan: x_1 edo x_2 edo x_3 .

Arrazoibide horri jarraituz, aldagaiaren balioak batu behar direla adierazi nahi denean:

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \dots$$

Demagun esan nahi dugula aldagaiaren balio guztiak batu behar direla. Luzeegia litzateke, baita zentzurik gabekoa ere, ikur-zerrenda luze bat egitea. Datu guztiez ari garela adierazteko irtenbidea:

$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \dots + x_n$
Lehenengo baliotik azken balioraino:
x_n : datu multzoaren azken balioa, datu kopurua edozein dela ere

📌 Datu sorta bat batu behar dela adierazteko Σ ikurra erabiltzen da. Ikur horri *batukari* deritzo. Batu behar diren datuen aurrean jartzen da. Batukaria erabiliz, goiko adierazpenak honako forma hau hartuko du:

$$\Sigma \quad x_1, x_2, x_3, x_4, x_5, x_6, \dots, x_n$$

📌 Badago oraindik beste baliabide bat aurreko adierazpen matematikoa laburtzeko: aldagaiaren balioen aipamena banan-banan egin ordez, honako adierazpena erabiltzen da:

$\sum_{i=1}^n x_i$	<p>Aldagaiaren balio guztiak batu behar direla adierazteko x-ren i azpiindizearen esanahia zehazten da: aldagaiaren balio guztiak, lehenengo baliotik ($i=1$) azken balioraino (n)</p>
<p>Batu behar dira (Σ) aldagaiaren balioak (x_i) lehenengo baliotik ($i=1$) azken balioraino (n)</p>	

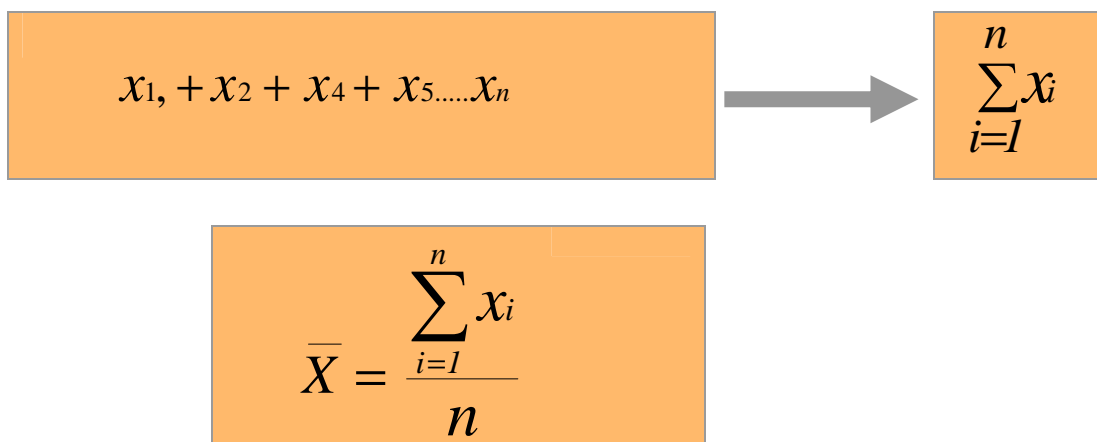
■ Honaino iritsita erraza izango da, batezbestekoaren definiziotik abiatuz, haren adierazpen matematikoa osatzea.

Datu multzo baten batura datu kopuruaz zatitzean lortzen den emaitzari batezbesteko aritmetiko deritzo:

$$\bar{X} = \frac{x_1 + x_2 + x_4 + x_5 \dots x_n}{n}$$

Non \bar{X} batezbestekoa adierazteko ikurra baita eta n -k datu kopurua adierazten baitu.

Lehenago azaldutako notazio estatistikoaren oinarriari jarraituz:

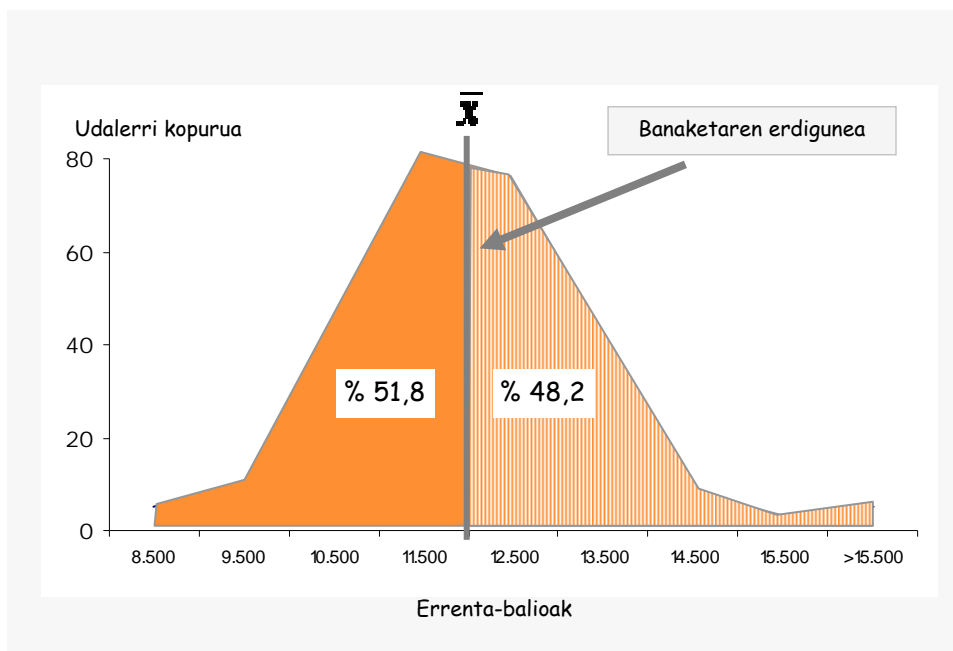


Batezbestekoaren erabilera

Aldagai baten balioen batezbestekoa erabiltzen da kanpai-formako banaketaren erdigunea *bilatzen*. Lehenago esan dugun bezala, kanpai-forma edukiz gero, ondo dakigu batezbestekoaren inguruan (hots, bi alboetan) biltzen direla datu gehienak. Baina, kontuz, hala gertatzen da soilik, kanpai-itxura duten banaketetan, batezbestekoa banaketaren erdigunean dagoenean.

Batezbestekoaren erabilera eta esanahia hobeto ulertzeko, kanpai-itxurako banaketa baten adibidea eta bestelako forma duen banaketa baten adibidea ikusiko ditugu.

Lehenbizi ikusiko dugu kanpai-itxurako banaketa bat. Bertan irudikatzen dira *2003ko Euskadiko udalerrietako batez besteko errenta pertsonal erabilgarria* aldagaiaren balioak.



Irudiak erakusten digu atal honen hasieran esaten genuena:

- Kanpai-itxurako banaketetan balio gehienak biltzen dituen tartearen erdigunea eta banaketaren erdigunea bat datozela, eta, ondorioz,

■ batezbestekoaren balioa ere banaketaren erdigunean kokatzen da, bere bi alboetan balio kopuru antzekoa izanik.

Errentari dagokionez, aipatutako ezaugarrien esanahia da, esaterako, batezbestekoaren balioa udalerrri gehienetako errenta-balioen erdian dagoela. Antzekoak dira, beraz, batezbestekoa baino errenta handiagoa duten udalerrrien kopurua eta batezbestekoa baino errenta txikiagoa dutenena.

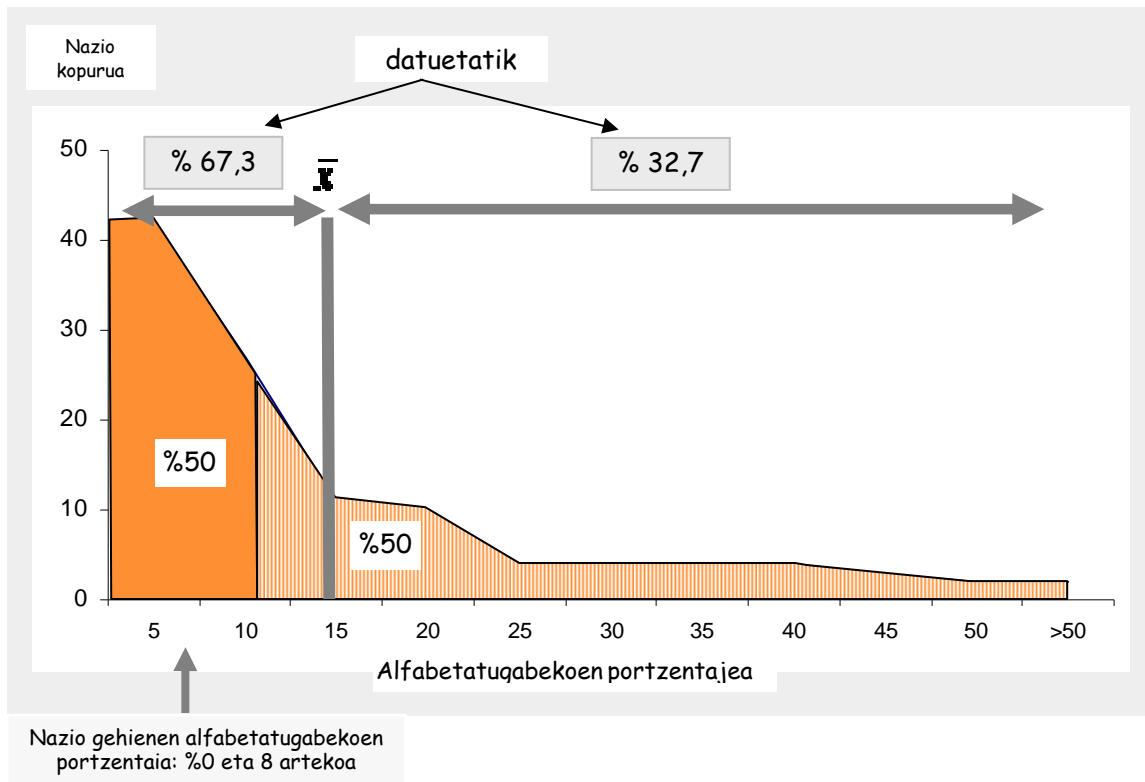
Aipatutako ezaugarrien ondorioz argiena honako hau da:

Kanpai-formako banaketara egokitzen diren aldagaien batezbestekoa **datu multzoaren adierazpen esanguratsua da.**

Udalerrien batezbesteko errenta, 11.506 eurokoa, aztertzen ari garen errentaren balio adierazgarria da, balio horren antzeko asko baitaude. Taula honetan ikusten dugunez, badira 193 udalerrri 10.000 eurotik 13.000 eurora bitarteko errentak dituztenak. 193 udalerrri horiek udalerrien % 77 dira.

Tarteak	Balio kopurua	Ehunekoa	Ehuneko metatua
>10.000 ≤10.500	29	12	12
> 10.500 ≤ 11.000	27	11	23
> 11.000 ≤ 11.500	47	19	41
> 11.500 ≤ 12.000	37	15	56
> 12.000 ≤ 12.500	32	13	69
> 12.500 ≤ 13.000	21	8	77
	193	77	

Ondoren, munduko hainbat nazioen alfabetatugabeen ehunekoari buruzko adibide bat ikusiko dugu. Oraingo honetan mota guztietako garapen-mailak dituzten 113 naziok osatutako populazioa dugu. Aldagaia da nazioetako alfabetatugabeen ehunekoa. Aurrekoan bezala, grafikoak erakusten digu ibiltarte barneko balioen banaketa. Koloreen bidez seinatu ditugu banaketaren bi alboak; albo bakoitzak balioen % 50 jasotzen du.



Argi dago kasu honetan ez gaudela kanpai-itxurako banaketaren aurrean. Zertan datza desberdintasuna? Zein dira bi banaketen arteko desberdintasunak?

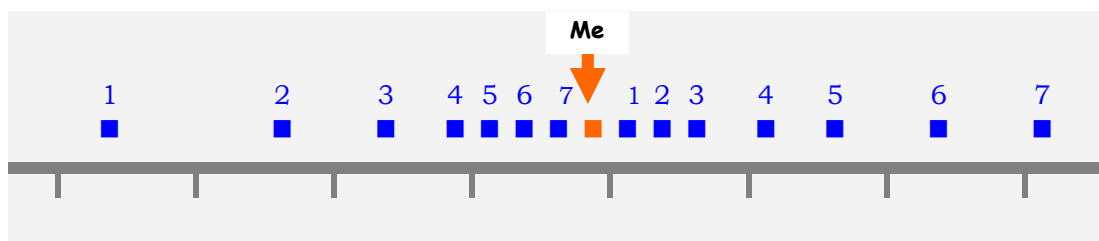
- Bigarren banaketa honetan, balio-maiztasun handienak biltzen dituzten tarteak ez daude banaketaren erdian mutur batean baino, alfabetatugabeen portzentajea % 8 baino txikiagoa daukatenean artean.
- Batezbestekoaren balioa ez dago maiztasun handieneko tartean barnean, balio handien artean baizik. Batezbestekoaren emaitzaren arabera nazioen alfabetatugabeen batez besteko portzentajea % 12,2koa da. Grafikoan ikusten da, berriz, nazio gehieneko alfabetatugabeen portzentajeak batezbestekoaren balioa baino txikiagoak direla. Hain zuzen, nazioetatik % 67,3n alfabetatugabeen portzentajea batezbestekoaren azpian dago.

Alfabetatugabeen adibidetik atera daitekeen ondorio argiena honako hau da:

Kanpai-formako banaketara egokitzen ez diren aldagaien batezbestekoa ez da datu multzoaren adierazgarri esanguratsua.

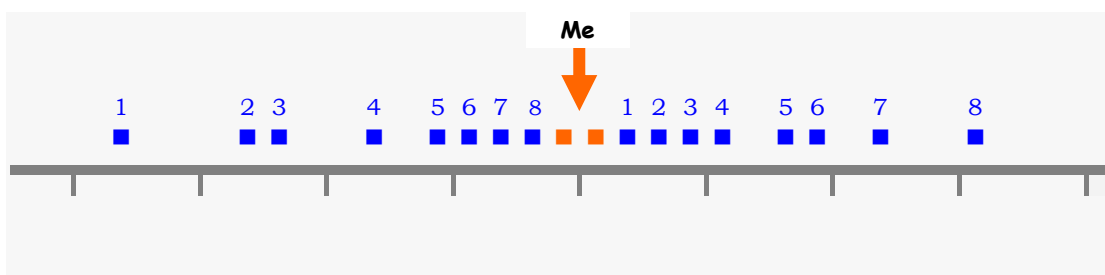
Banaketen erdigunea finkatzeko beste neurria: mediana

Aurreko ataletan behin eta berriko ikusi dugu batezbestekoaren balioa banaketaren erdialdean kokatzen dela, baldin eta banaketa kanpai-itxurakoa bada. Alfabetatugabeen buruzko adibidean, ordea, batezbestekoa banaketaren erdigunetik alboratuta agertzen da. Halako kasuetan, batezbestekoaren balioa adierazgarria ez dela-eta, bestelako neurri batzuk erabili behar dira banaketaren erdigunea finkatzeko.



Batezbestekoaren balioa adierazgarria ez den kasuetan, erdigunea finkatzeko erabil daitekeen beste neurria mediana izeneko da. Aldagai baten balioak hartu eta ordenatu ondoren, datu multzoa bitan banatzen duen balioari mediana deritzo. Datu multzoak bakoitiak direnean, mediana erdiko balioa da.

Datu multzoa bikoitia izanez gero, erdiko balio bakararra ez dagoenez, erdiko bi balioen batazbestekoa da mediana.



Taula honetan ikusten dira Espainiako autonomia-erkidegoetan, 2004an, gaikako zabor-bilketaren bidez jaso ziren paper eta kartoia kopuruak, hain zuzen, urteko eta pertsonako kilo kopurua. Erkidegoen kopurua bikoitia denez, erdiko bi balioen batezbestekoa kalkulatu dugu mediana lortzeko:

Autonomia-erkidegoak	Papera eta kartoia (kg)
Extremadura	8,3
Ceuta eta Melilla	9,3
Gaztela eta León	10
Andaluzia	11
Gaztela-Mantxa	12,6
Valentziako erkidegoa	13
Murtzia	13,6
Galizia	14
Kanariak	14,5
Kantabria	14,8
Aragoi	15,8
Madril	20,3
Katalunia	21,2
Errioxa	21,2
Asturias	22,8
Balear Uharteak	25,7
Nafarroa	26
Euskadi	38,9

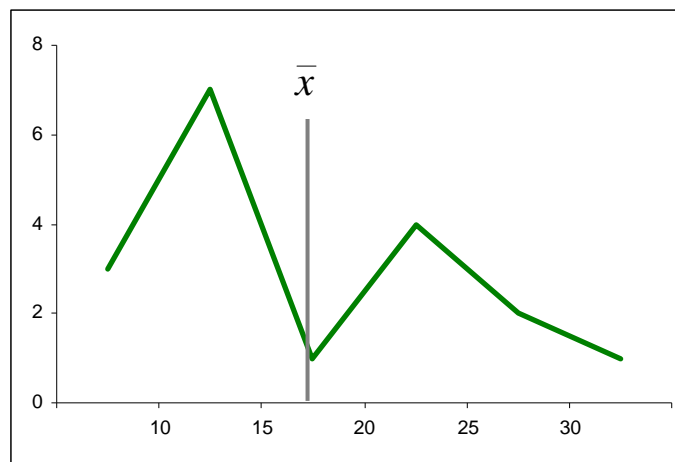
$$Me = \frac{14,5 + 14,8}{2} = 14,65$$

Medianaz gain, batezbestekoa ere kalkulatu dugu:

$$\frac{\sum_{i=1}^n x_i - \bar{x}}{n} = \frac{313}{18} = 17,39$$

	Batezbestekoa	Mediana
	17,39	14,65
Diferentzia	2,74	

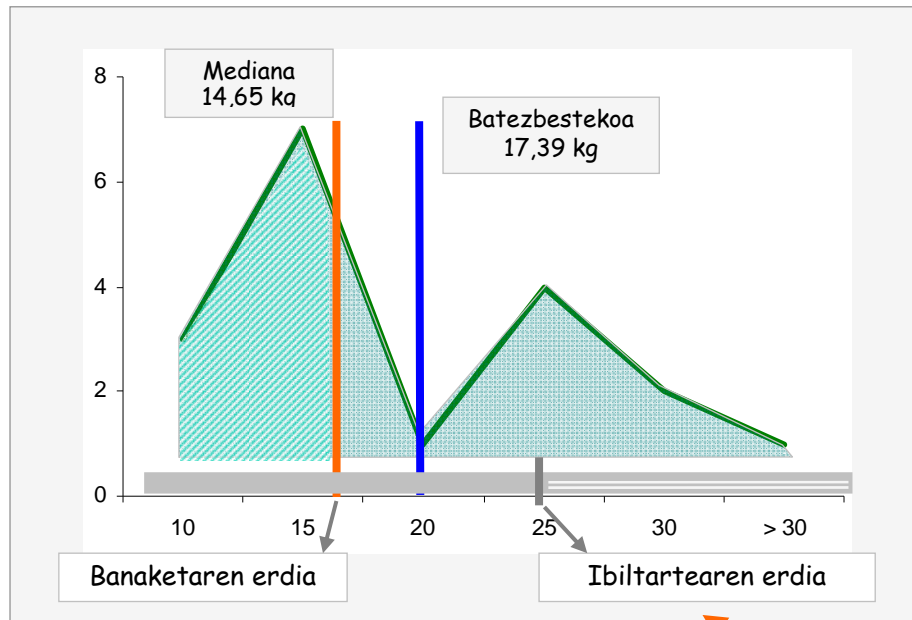
Bi neurrien emaitzak konparatuta, esan daiteke batezbestekoaren eta medianaren arteko diferentzia nabarmena dela, pertsonako jasotzen diren kopuruei dagokienez 2,74 kiloko diferentzia handia da eta. Balioen maiztasun-poligonoaren grafikoa ikusita, berehala ohartuko gara diferentzia horren zergatiaz:



Argi ikusten dugu orain erkidegoen portaera, paperaren eta kartoiaren bilketari dagokionez, zeharo desberdina izan zela. Hamaika erkidegotan pertsonako jasotako kopurua 16 kiloz azpikoa izan bazen ere, beste erkidegoetan jasotakoak nabarmen handiagoak izan ziren. Hori dela eta, balio multzo nagusiaren alboan badago bigarren azpimultzo bat balio handiagok osatutakoa. Azken horien eraginez batezbestekoaren emaitza igo da, eta bi multzoen arteko balioari dagokion lekua hartu.

Batezbestekoa banaketaren erdigunetik urrun geratzen dela hobeto ikusteko,

erakutsiko dugu orain batezbestekoa eta medianaren kokaera balio-banaketaren barnean:

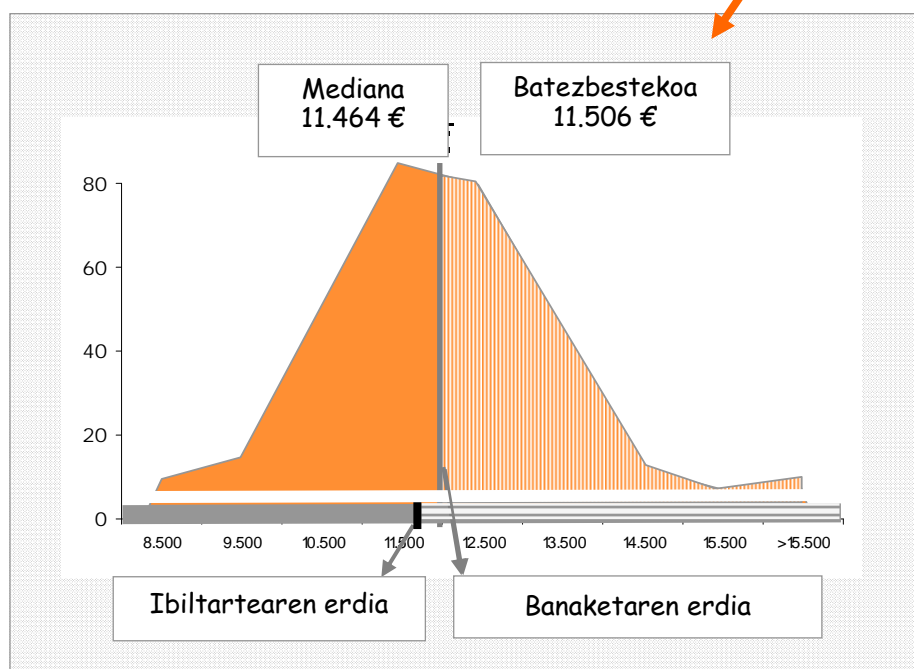


Kanpai-itxura ez daukan balio-banaketa

Banaketaren erdia, batezbestekoa eta ibiltartearen erdia: **balio oso desberdinak**

Banaketaren erdia, batezbestekoa eta ibiltartearen erdia: **balio oso antzekoak**

Kanpai-itxura daukan balio-banaketa



Estatistikaren arloan, batezbestekoa eta mediana, beste batzuekin batera, neurri mota berdinen barnean sailkatzen dira, *joera zentraleko neurrien* barruan, hain zuzen. Dagoeneko bistakoa da izenaren zergatia, argi ikusi dugulako banaketaren erdigunea finkatzeko erabiltzen diren neurriak direla.

Elementu arraroren bat duten banaketen batezbestekoa

Datu multzoetan egoten diren *elementu arraroek* eragin handia izan dezakete batezbestekoaren emaitzan, hain zuzen ere, batezbestekoa kalkulatzeko prozeduraren ondorioz. Beste neurriekin gertatzen ez den bezala, batezbestekoa kalkulatzeko aldagaiaren datu guztiak erabiltzen dira, eta, ondorioz, balio guztiak dira emaitzaren parte. Logikoa den hori arazo bihurtzen da balio baten eragina handiegia denean, alegia, batezbestekoaren emaitza nabarmenki handitzen edo jaisten duenean.

Datu multzoetan elementu arraroak egotea ohikoa da. Kanpai-itxurako banaketetan ere arrunta izaten da besteak baino askoz handiagoa edo txikiagoa den balioen bat egotea. Hori da, esaterako, Aragoiko hirietako etxebizitza berrien prezioen datuekin gertatzen dena. Taula honetan¹⁴ (2007ko datuak jasotzen ditu), ikusten dugu zein handia den prezio-aldea zerrendako hiri garestienaren (Zaragozaren) eta hurrengo garestienaren (Jacaren) artean. Zaragozako etxeen prezioa, Jacarena baino ia 104.000 euro handiagoa, batezbestekoaren emaitza nabarmenki igotzen duen *elementu arraroa* da.

¹⁴ Etxebizitza berriaren batez besteko prezio nominala, eurotan, 2007ko ekainean. Sociedad de Tasación. SA. <http://web.st-tasacion.es/html/index.php>. Balio nominala: balio honetan KPI (kontsumo-mailan ohikoa den diru-unitatea) barnean dago. Balio erreala, 1985etik aurrera KPIren igoera 0 izan dela jota.

	Hiriak	Prezioa (€)	Diferentziak (€)
1	Caspe	103.900	
2	Fraga	127.000	23.100
3	Barbastro	128.500	1.500
4	Alcañiz	129.700	1.200
5	Ejea de Los Caballeros	155.400	25.700
6	Calatayud	158.600	3.200
7	Teruel	170.600	12.000
8	Huesca	194.900	24.300
9	Jaca	197.600	2.700
10	Zaragoza	301.500	103.900

Zaragozako datuaren eragina ikusteko nahikoa da batez bestekoa bi aldiz kalkulatzea: batean Zaragozako datua barnean eta, bestean, datua kenduta. Emaitzak adierazgarriak dira, bi batez bestekoen arteko diferentzia ia 15.000 eurokoa baita.

Batezbestekoa	Zaragozako datua barne	Zaragozako daturik gabe
	166.770	151.800

Ondoren ikusiko ditugu beste datuak, etxebizitzaren batez besteko prezioei buruzkoak, berriro 2007koak eta Aragoikoak, baina kasu honetan eskualde-mailakoak eta bakarrik Zaragozako probintziakoak¹⁵. Aurreko datuetan ez bezala, hauetan jasotzen dira etxebizitza berrien prezioak eta bigarren eskuko etxebizitzarenak. Helburua ez da egitea inolako konparaziorik bi datu multzoen artean, baizik eta erakustea elementu arraroek ez dutela beti eragin nabarmena batezbestekoaren emaitzan.

¹⁵ Zaragozako eskualdeetako etxebizitzaren batez besteko prezioa (€). Caja de Ahorros de la Inmaculada de Aragón. *Mercado Inmobiliario de Aragón 2007*.

	Eskualdea	Prezioa €	Diferentzia €
1	Bajo Martín	52.647	
2	Bajo Aragón-Caspe	78.960	26.313
3	Campo de Daroca	90.090	11.130
4	Sierra de Albarracín	92.584	2.494
5	Campo de Belchite	95.184	2.600
6	Campo de Borja	97.468	2.284
7	La Ribagorza	100.152	2.684
8	Matarraña	100.386	234
9	Sobrarbe	104.958	4.572
10	Andorra-Sierra de Arcos	105.840	882
11	Jiloca	107.334	1.494
12	Ribera Alta del Ebro	108.864	1.530
13	Bajo Aragón	109.896	1.032
14	Valdejalón	112.800	2.904
15	Cinca Medio	121.869	9.069
16	Cinco Villas	122.400	531
17	Comunidad de Calatayud	123.172	772
18	Bajo Cinca	127.846	4.674
19	Gúdar-Javalambre	142.389	14.543
20	Somontano de Barbastro	145.376	2.987
21	Tarazona eta El Moncayo	153.510	8.134
22	La Jacetania	167.014	13.504
23	Ribera Baja del Ebro	176.157	9.143
24	Los Monegros	178.830	2.673
25	Hoya de Huesca	181.764	2.934
26	Comunidad de Teruel	185.472	3.708
27	La Litera	186.400	928
28	Campo de Cariñena	186.944	544
29	Alto Gállego	192.975	6.031
30	D.C. Zaragoza	236.672	43.697

Kasu horretan ere Zaragozako datuak ez dauka zerikusirik besteekin, askoz altuagoa baita. Goiko taulan, prezioen zutabearen eskuinean beste zutabe bat egokitu dugu. Bertan jaso dugu eskualde baten prezioetik hurrengora dagoen diferentzia. Azken lerroan ageri dira, azpimarratuta, garestienak diren bi eskualdeen arteko prezio-aldeak: 43.697 eurokoa da *Alto Gállego* eta *D.C. Zaragoza* izeneko eskualdeen artekoa; ehunekotan emanez gero, esan daiteke Zaragozako prezioa *Alto Gállego*koa baino % 22,6 handiagoa dela.

Azter dezagun orain zein den Zaragozako datuaren eragina batezbestekoan.

	Zaragozako datua barne	Zaragozako daturik gabe
Batezbestekoa	132.865,1	129.285,55
Diferentzia	3.579,55	

Oraingo honetan ere, Zaragozako datuak batezbestekoaren emaitza 3.579,55 euro igotzen du. Kantitatea ez da txikia, baina aurreko datuekin lortutako 15.000 eurotik urrun dago. Argi dago, beraz, eskualdeko prezioen elementu arraroa, Zaragoza eskualdeko datua, ez dela aurreko adibidearen bezain apartekoa. Dena den, badago beste faktore bat azaltzen duena zergatik duen oraingo elementu arraroak batezbestekoaren emaitzan eragin txikiagoa. Ikusitako bi datu multzoen tamaina –hots, populazio kopurua– desberdina da. Hirietako prezioen kasuan populazioaren tamaina 10 osagaikoa da; eskualdeko osagai kopurua, berriz, 30koa da. Hori dela eta, Zaragozako eskualdeko balio altua 30 osagaien artean banatzen da, eta, hala, batezbestekoan daukan eragina gutxitzen da.

Demagun, eskualdeko prezioen kasuan, Zaragozako datua jatorrizkoa baino altuagoa dela; 236.672 eurokoa izan ordez, 290.000 eurokoa dela. Ikus dezagun orain zer eragin daukan eskualdeko prezioen batezbestekoan asmatutako datuak:

Batezbestekoa	Zaragozako datua, asmatutakoa	Zaragozako datua barne	Zaragozako daturik gabe
	134.643	132.865,1	129.285,5

Zaragozako datua aldatuta, batezbestekoaren emaitza igo da; haren eragina, beraz, nabaritzen da. Alabaina, diferentzia 5.357,5 eurokoa da, hirien prezioen kasuan *elementu arraroak* sortzen duena (ia 15.000 eurokoa) baina askoz txikiagoa. Agerian geratzen da, ondorioz, *elementu arraroen* eragina, neurri handian, datu multzoaren tamainaren menpean dagoela.

Datu multzoaren tamainaren garrantzia are argiago ikusten da Euskadiko udalerrietako errentaren adibidean. Horretan, 251 udalerrietatik batek badu errenta erabilgarriaren balioa besteena baino askoz handiagoa; elementu arrarotzat hartu behar da. Hala ere, osagaien kopurua handia denez, balio horrek ez du ia inolako eraginik eduki batezbestekoaren emaitzan:

Euskadiko udalerrietako batezbesteko pertsonako errenta erabilgarria, 2001en (€)		
Balio txikiena	Lanestosa	7.192
Balio handiena	Laukiz	20.627
Bigarren balio handiena	Leintz Gatzaga	15.346
	Elementu arraroa barne	Elementu arrarorik gabe
Batezbestekoa	10.514,47	10.474
40,45 euroren aldea		

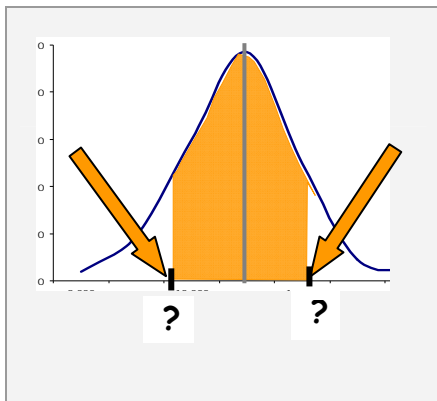
Zer esan daiteke, azken finean, elementu arraroren bat daukaten datu multzoen analisiari buruz?

- ❖ Eskuarki aldagaiaren balio kopurua txikia ez denean, elementu arraroek ez dute eragin nabarmenik batezbestekoaren emaitzan.
- ❖ Ez dago araurik balio kopuru bat handia edo txikia den esateko. Hori dela eta, *elementu arraroak* agertzen diren kasu guztietan haien eragina aztertu eta balioetsi behar da.
- ❖ *Elementu arraroen* eragina handia dela pentsatuz gero, batezbestekoa kalkulatzeko, eragin handiegia duen datua ken daiteke. Analisiaren egileak hartu beharko du mantentzeko edo kentzeko erabakia.
- ❖ Edozein kasutan, *elementu arraroak*, hutsegiteak ez badira, errealitatean existitzen duten neurrian ezin dira ahaztu edo baztertu. Hori dela eta, nahiz eta batezbestekoaren kalkulutik kanpoan utzi, elementu arraroek dagokien lekua eduki behar dute analisisian.

BANAKETAREN MAIZTASUN GEHIENEO TARTEA: DESBIDERATZE ESTANDARRA

Maiztasun gehieneko tartea

Kanpai-itxurako banaketei buruz hitz egin dugunean, esan dugu batezbestekoa ibiltartearen erdialdean eta baita maiztasun gehieneko tartearen erdian ere kokatzen dela.



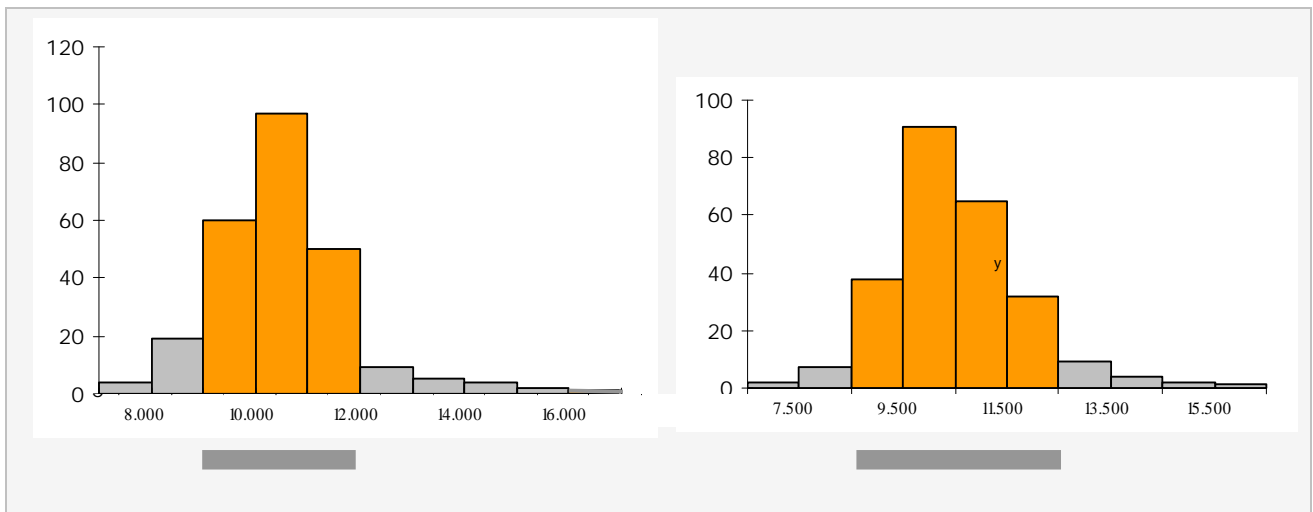
Banaketaren erdigunea balio zehatzari dagokio, batezbestekoaren bidez kalkulaten dena.

Kontua da orain nola zehaztu edo kalkulatu maiztasun gehieneko tartea mugatzen duten bi balioak.

Maiztasun gehieneko tartearen mugak bilatzearen helburua da banaketaren ezaugarri adierazgarriak deskribatzea. Horretarako ezinbestekoa da esatea ibiltarteko zein balioen artean biltzen diren datu multzoaren balio gehienak, informazioa hori banaketen ezaugarri adierazgarria izaten da eta.




Histogramak eta maiztasun-poligonoak azaldu ditugunean, argi ikusi dugu ezin direla erabili grafiko horiek maiztasun gehieneko tartearen mugak zehazteko, aukeratzen diren tarteen arabera histogramen eta maiztasun-poligonoen itxura aldakorra baita. Ez dugu begien bistatik galdu behar maiztasun gehieneko tartearen zabalera, handiagoa edo txikiagoa izatea, aldagaiaren aldakortasun-mailaren adierazlea dela.

Alboko irudiak erakusten ditu bi histograma, datu eta zabalera berdineko tartea dituztenak. Bietan, berriz, desberdina da lehen tartearen hasiera. Datuak dira lehenago erabili ditugunak: Euskadiko udalerrietako batezbesteko errenta erabilgarria (2001). Bi histogramek aldagaiaren aldakortasunari buruzko ideia bat adierazten dute.



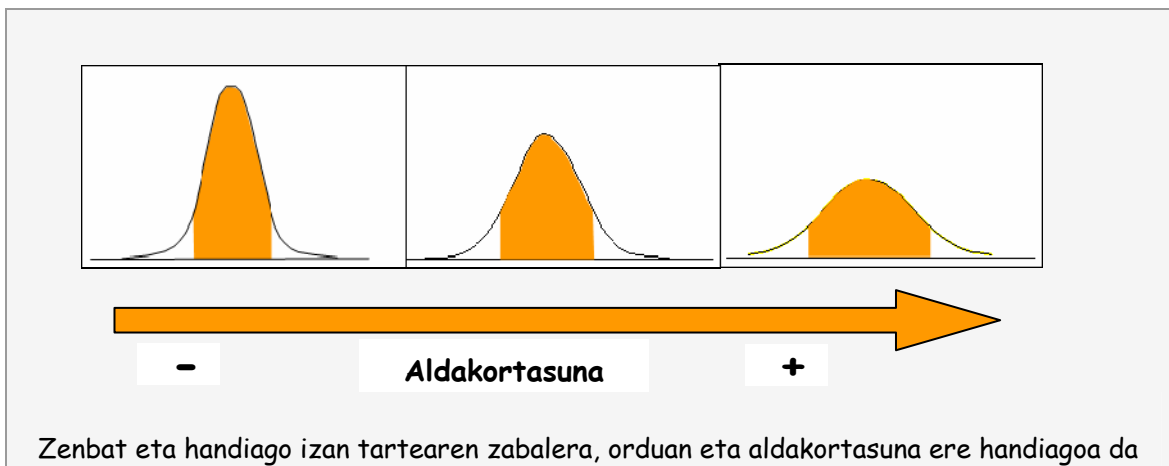
Eskuineko histogramak, ezkerreko histogramarekin konparatuta, aldagaiaren aldakortasuna handiagoa dela ematen du. Ezkerrekoan, datu gehienak biltzen dituen tartea 9.000 eurotik 12.000 eurora bitartekoa da. Eskuinekoan, datuen taldekatzea –tarterik zabalena– 8.500 eurotik 12.500 eurora bitartean agertzen da.

Maiztasun gehieneko tartea zehazteko histograma erabiliz gero, hainbat emaitza lor daitezke, egileak hartutako erabakien arabera. Ondorioz, ibiltarteko datu-taldekatzea zehazteko, argi dago prozedura estandarra erabili behar dugula:

-  Egileak hartzen dituen erabakien menpean ez dagoen prozedura.
-  Edonork erabiliz gero, emaitza berdinak ematen dituen prozedura.
-  Adituen artean ezaguna, ohikoa eta onartua den prozedura.

Aipatutako baldintzak betetzen dituen prozedura erabiliz gero, hainbat ikerketaren arteko emaitzak konparatzeko aukera egongo da; emaitzen arteko aldeak aldagaien ezaugarrien ondorioak izango dira, eta ez egileek erabili dituzten irizpideen ondorioak.

Maiztasun gehieneko tartea zehazteko erabiltzen dugun prozedura une batez alde batera utziz, gogoratu behar dugu banaketen datu gehienak biltzen dituen tartearen-zabalera desberdina izan daitekeela eta izaten dela. Datuak egon daitezke batezbestekotik gertu, taldekatuak, edo tarte zabalago baten barnean banatuta. Modu batean edo bestean egotea aldagaiaren aldakortasun-mailaren ondorioa da.

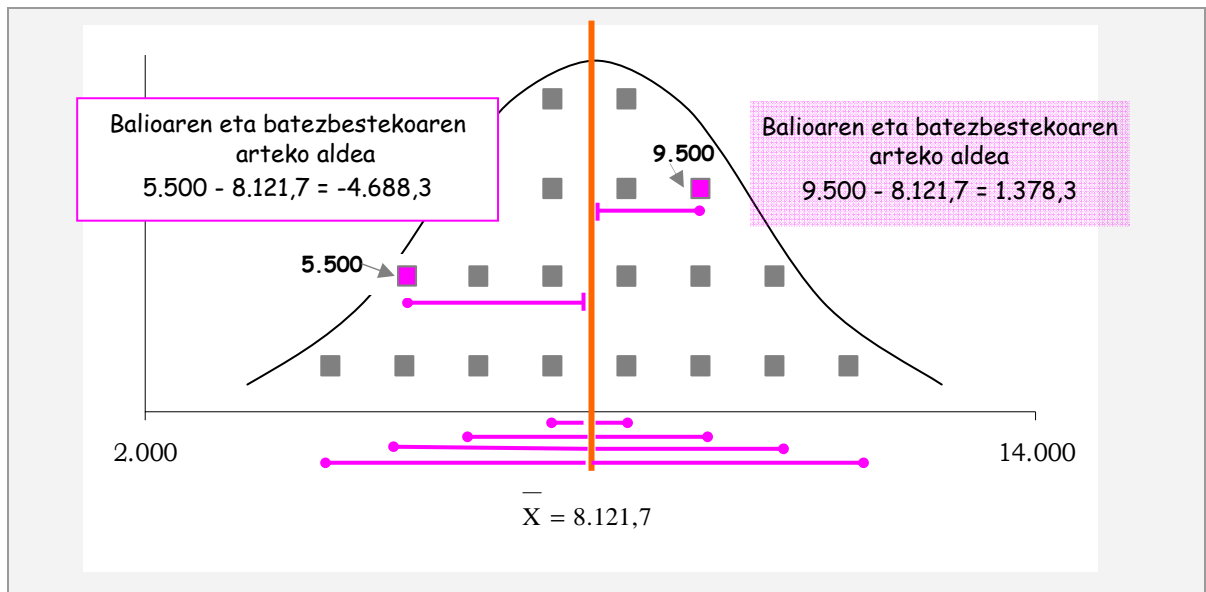


Desbideratze estandarraren kalkulua

Desbideratze estandarra da estatistikak eskaintzen digun prozedura bat aldagaiaren balioen taldekatzea zehazteko, hots, aldagaiaren datu gehienak ibiltarteko zein punturen artean biltzen diren finkatzeko.

Lehenengo eta behin ezagutu eta barneratu behar dugu aldagaiaren balio gehienak biltzen dituen tartea finkatzeko prozeduraren oinarria. Aipatzen dugun oinarria batezbestekoaren eta balio guztien arteko diferentzia neurtzean datza; alegia, balioek, banaka hartuta, batezbestekotik daukaten aldakortasunaren analisisan.

Maiztasun gehieneko tarteko muturrak ibiltarte barneko bi balio dira, batezbestekotik gertuago edo urrunago daudenak. Hori dela eta, bi puntu edo balio horiek finkatzeko, banakako balioen eta batezbestekoaren arteko tartea kalkulatu behar da, edo, bestela esanda, datuen eta batezbestekoaren arteko balio-aldea.



Goiko irudian erakusten den bezala, aldagaiaren balio guztiak zerrendatuta, banakako kenketak egiten dira, balio guztien eta batezbestekoaren arteko kenketak, alegia. Eragiketa horren ondorioz bigarren datu-zerrenda lortzen dugu, oraingo honetan kenketen emaitzez osatutako zerrenda. Prozeduraren pausoak argiago ikusteko adibide bat erabiliko dugu.

Beheko taulan ikusten ditugun datuak Euskadiko hirietako etxebizitzaren prezioak dira. Datuak 2007koak dira. Balioak bi zutabetan banandu ditugu, batezbestekoa baino txikiagoak direnen taldea eta handiagoena:

Aldagaiaren balioak: etxebizitza berrien prezioak, €/m ²				
Batezbestekoa baino txikiagoak diren balioak: diferentzia negatiboak			Batezbestekoa baino handiagoak diren balioak: diferentzia positiboak	
Basauri	2.120	Batezbestekoa 2.881,31	2.916	Barakaldo
Hernani	2.255		2.988	Gasteiz
Erandio	2.571		3.097	Irun
Arrasate	2.577		3.268	Bilbo
Portugalete	2.753		3.272	Getxo
Santurtzi	2.781		4.061	Donostia
Leioa	2.798			

¹⁶Ikus ditzagun orain balioen eta batezbestekoaren arteko aldeak:

Diferentzia negatiboak		Diferentzia positiboak	
Hiriak	$x_i - \bar{x}$	$x_i - \bar{x}$	Hiriak
Basauri	-761,31	34,69	Barakaldo
Hernani	-626,31	106,69	Gasteiz
Erandio	-310,31	215,69	Irun
Arrasate	-304,31	386,69	Bilbo
Portugalete	-128,31	390,69	Getxo
Santurtzi	-100,31	1.179,69	Donostia
Leioa	-83,31		
$\sum x_i - \bar{x}$	-2.314,1	2.314,1	

Lehen esaten genuen bezala, diferentziak kalkulatuta beste datu-zerrenda bat lortzen dugu. Alabaina, guri interesatzen zaigu balioen eta batezbestekoaren arteko aldeari buruzko balio bat, adierazgarria den balioa. Diferentzia guztien adierazgarri egokia izan liteke haien batezbestekoa. Izan litekeela esaten dugu, diferentzien batezbestekoa egiterik ez baitago:

Diferentzia negatiboak	+	Diferentzia positiboak	=	0
-2.314,1	+	2.314,1	=	0

Batezbestekoaren eta balioen arteko diferentzien batuketak, kasu honetan eta

¹⁶ Etxebizitza berriaren batez besteko prezio nominala, eurotan, 2007ko ekainean. Sociedad de Tasación, SA. <http://web.st-tasacion.es/html/index.php>.

guztietan, emaitza berdina ematen du: 0. Horrela da, hain zuzen ere, batezbestekoaren propietate bat delako.

Balioen eta batezbestekoaren arteko diferentzien batezbestekoa egiterik ez dagoela barneratuz gero, erraz ulertuko da diferentziei buruzko balio bat sortzeko prozedura.

Estatistikak eskaintzen ditu irtenbide batzuk aldeei buruzko balio adierazgarri bat sortzeko. Balio horren bidez maiztasun gehieneko tartearen mugak finkatu egiten dira.

Bilatzen dugun balioa sortzeko prozedura bat *batez besteko desbideratzea* da. Prozedura horren bidez, balioen eta batezbestekoaren arteko aldean ikurra kentzen da, eta, ondoren, haien batezbestekoa kalkulatzen da. Balioen eta batezbestekoaren arteko diferentzia guztiak positibotzat hartzeari *balio absolutuak* erabiltzea deritzo. Batez besteko desbideratzearen kalkulua eta erabilera bereziki erraza den arren, errealitatean oso gutxitan erabiltzen da, hain zuzen ere balio absolutuak hartu behar direlako.

$$D\bar{x} = \frac{\sum_{i=1}^n x_i - \bar{x}}{n}$$

Euskadiko hirien etxebizitzetarako buruzko balioen batez besteko desbideratzearen kalkulua:

$$\frac{\sum_{i=1}^n x_i - \bar{x}}{n} = \frac{4.628,31}{13} = 356,02$$

Aldagaiaren balioen eta batezbestekoaren arteko diferentzien balioa finkatzeko, errealitatean gehien erabiltzen den prozedura **desbideratze estandarra** da. Kasu horretan ere, diferentzien batezbestekoa kalkulatu da, baina balioen batuketaren arazoari beste modu batez egiten zaio aurre. Desbideratze estandarraren kalkulua egiteko –alegia, diferentzien batezbestekoa lortzeko–, diferentzien karratuak erabiltzen dira. Balioen diferentzien karratuak egitean balio negatiboak desagertzen dira.

Euskadiko hirietako etxebizitzaren prezioak aztertzen jarraituko dugu, desbideratze estandarraren aplikazioa erakusteko:

Euskadiko hiriak	Aldagaiaren balioak	Balioen eta batezbestekoaren arteko diferentziak	Diferentzien karratuak
	x	$(x - \bar{x})$	$(x - \bar{x})^2$
Basauri	2.120	-761,31	579.592,92
Hernani	2.255	-626,31	392.264,22
Erandio	2.571	-310,31	96.292,30
Arrasate	2.577	-304,31	92.604,58
Portugalete	2.753	-128,31	16.463,46
Santurtzi	2.781	-100,31	10.062,10
Leioa	2.798	-83,31	6.940,56
Barakaldo	2.916	34,69	1.203,40
Gasteiz	2.988	106,69	11.382,76
Irun	3.097	215,69	46.522,18

Bilbo	3.268	386,69	149.529,16
Getxo	3.272	390,69	152.638,68
Donostia	4.061	1.179,69	1.391.668,50
$\sum_{i=1}^n (x_i - \bar{x})^2$			2.947.164,77

Goiko taulan ikusten dugunez, diferentzien karratuak eginda horien batuketa egin daiteke, eta, bukatzeko, diferentzien batezbestekoa lortu:

$$\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{2.974.164,77}{13} = 226.704,98$$

Diferentzien batezbestekoaren emaitza ikusita konturatzen gara horren balioak ez duela zerikusirik aldagaiaren balioekin. Diferentzien batezbestekoaren emaitza balio izugarri handia da, aldagaiaren balioak baina askoz handiagoa. Izan ere, batezbestekoa kalkulatzeko diferentzien karratuak erabili ditugu. Hori dela eta, diferentzien batezbestekoaren emaitzaren igoera orekatzeko, horren erro karratua egiten dugu, eta, horrenbestez, desbideratze estandarraren formula honela geratzen da:

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

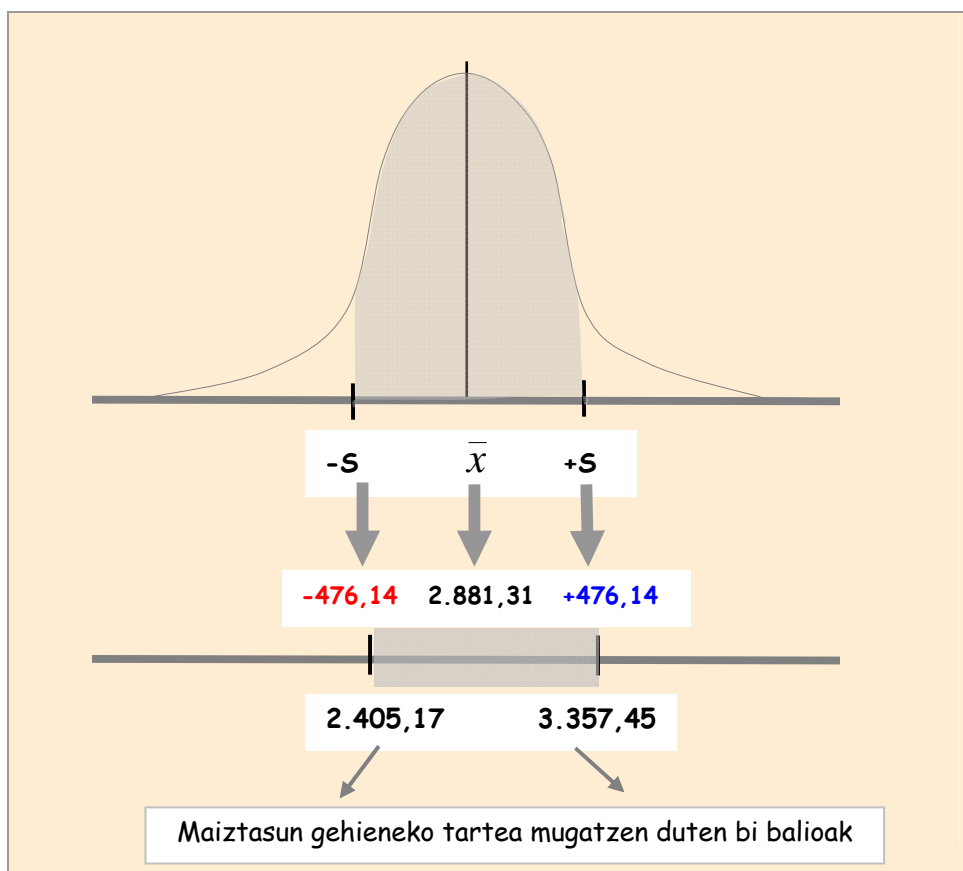
Hauxe litzateke erabili dugun adibidearen emaitza –alegia, Euskadiko hirietako etxebizitzaren prezioen desbideratze estandarra–:

$$S = \sqrt{\frac{2.947.164,77}{13}} = 476,14$$


Desbideratze estandarraren erabilera


Atalaren hasieran esan dugu gure helburua dela balioen aldakortasunari buruzko datua aurkitzea. Bilatzen dugu balio bat, banaketaren maiztasun gehieneko tarte finkatzeko erabil daitekeena. Esan dugu baita ere maiztasun gehieneko tartearen muturrak ibiltarte barneko bi balio direla, desbideratze estandarraren bidez finka daitezkeenak. Orain, desbideratze estandarraren kalkulua azaldu ondoren, ikusiko dugu nola erabil daitekeen desbideratzearen neurria balio kopuru gehien biltzen duen ibiltarteko sektorea zehazteko.

Onartzen badugu desbideratze estandarrak neurtzen duela zer neurritan aldentzen diren aldagaiaren balioak batezbestekotik, argi ikusiko dugu nahikoa dela batezbestekoari neurteta hori batzea eta kentzea bilatzen dugun tarte finkatzeko. Aurrekoetan bezala, adibide bat erabiltzea izango da modu egokiena prozeduraren zehaztasunak hobeto ulertzeko. Euskadiko hirietako etxebizitzaren prezioei buruzko datuak aztertzen jarraituko dugu:



Goiko irudiak erakusten digu zein den prozedura etxebizitzaren datu-banaketa maiztasun gehieneko tarteen mugak lortzeko:


 Balioen batezbestekoari (2.881,31 €) desbideratze estandarraren balioa (476,14 €) kendu diogu. Emaitzak (2.405,17) maiztasun gehieneko tartearen behe-muga finkatzen du.


 Balioen batezbestekoari (2.881,31 €) desbideratze estandarraren balioa (476,14 €) batu diogu. Emaitzak (3.357,45) maiztasun gehieneko tartearen goi-muga finkatzen du.


Behin kalkuluak bukatuta, prest gaude baieztatzeko hiri gehienetan etxebizitzaren metro karratuko prezioa 2.405,17 eurotik 3.357,45 eurora bitartean dagoela. Beraz, **hiri guztietako prezioak** 2.120 eurotik 4.061 eurora bitartean badaude ere, **hiri gehienetako prezioak** 2.405,17 eurotik 3.357,45 eurora bitartean daude.

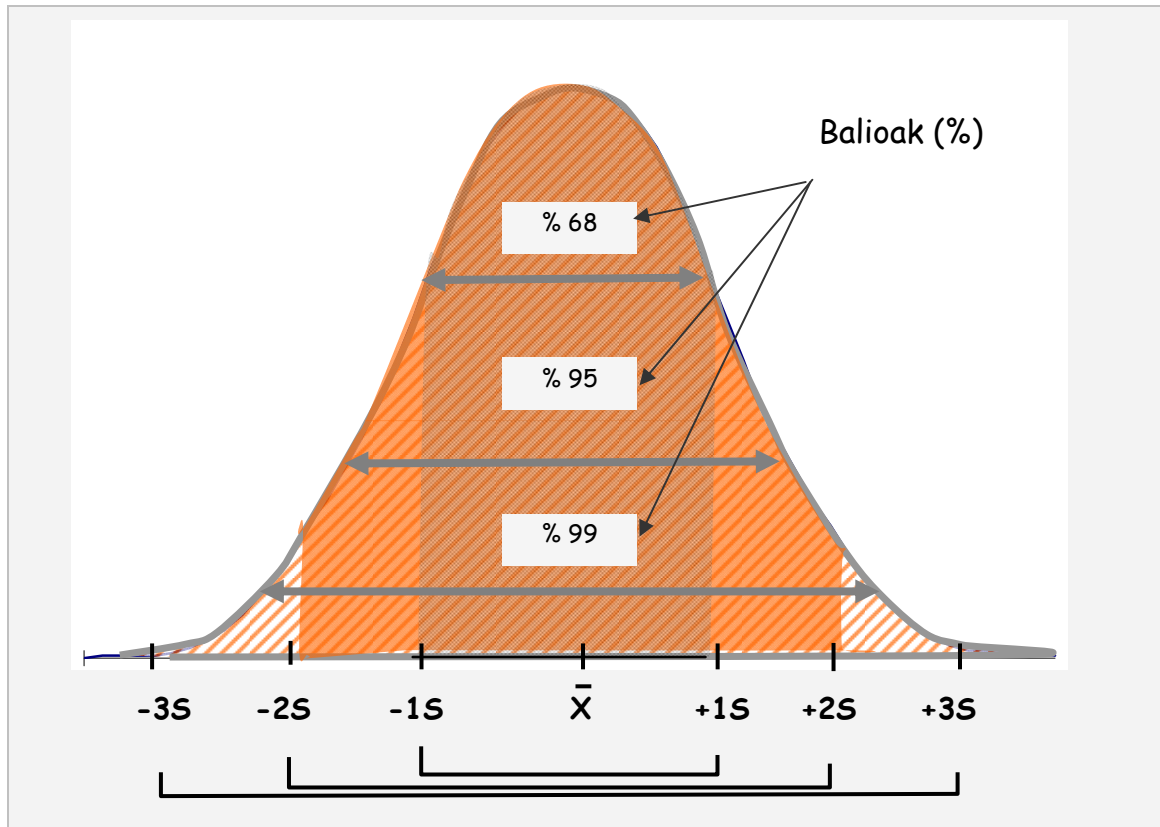
Badago beste modu bat desbideratze estandarraren emaitza aipatzeko, hots, balioen aldakortasuna adierazteko: maiztasun gehieneko tartearen mugak aipatu ordez, desbideratze estandarraren emaitza zuzenean ematea. Etxebizitzaren prezioen adibideari aplikatuz zera esan genezake: **Euskadiko hirietako etxebizitzaren prezio-aldakortasuna batezbestekoarekiko 476,14 eurokoa da.**

Aldagaiaren balio gehienei buruzko baieztapena egin daiteke, hain zuzen ere, aldagai horren balioak kanpai-itxurako banaketara egokitzen direlako. Arrazoi beragatik (kanpai-itxurakoa izateagatik) beste baieztapen hauek ere egin daitezke:


 Balioetatik % 68 inguru tarte honetan daude: batezbestekoa +/- desbideratze estandarra.


 Balioetatik %95 inguru tarte honetan daude: batezbestekoa +/- bi desbideratze estandarrak.

 Balioetatik %99 inguru tarte honetan daude: batezbestekoa +/- bi desbideratze estandarrak.



Desbideratze estandarra askotan erabiltzen da antzemateko aldagaiaren balioei, batezbestekotik aldenduta samar daudenak. Helburua da hainbat osagaien portaera berezia nabarmentzea. Horretarako, hainbat ikerketatan, $\bar{X} \pm 2S$ tarte edo $\bar{X} \pm 3S$ tarte erabiltzen dira. Tarte horietatik kanpo geratzen diren balioek aldagaiaren egoera berezia osatu ohi dute. Horrelako tarteen aipamena egiteko esaten ohi da batezbestekoari horrenbeste *desbideratze unitate* gehitzen edo kentzen zaizkiola:

 $\bar{X} \pm 2S$: batezbestekoa plus-minus bi *desbideratze unitate*


 $\bar{X} \pm 3S$ batez bestekoa plus-minus hiru *desbideratze unitate*


Euskadiko hirietako etxebizitzaren prezioen kasuan sortuko liratekeen tartearik ikusiko ditugu:

Batezbestekoa	2.281,31
Desbideratze estandarra	476,14

$\bar{X} \pm 2S$	2.881,31	-	476,14	-	476,14	=	1.929,04		
	2.881,31	+	476,14	+	476,14	=	3.833,58		
$\bar{X} \pm 3S$	2.881,31	-	476,14	-	476,14	-	476,14	=	1.452,90
	2.881,31	+	476,14	+	476,14	+	476,14	=	4.309,71

Etxebizitzen prezioen inguruko egoera bereziei antzemateko $\bar{X} \pm S$ tartek erabiliz lortuko genituzkeen ondorioak:

 Ez dago hiririk $\bar{X} \pm 3S$ tarteko muturretik kanpoko preziorik daukanik.

 Hiri bat badago $\bar{X} \pm 2S$ tarteko muturretik kanpo prezioa daukana: Donostiak 4.061€/m²-ko prezioa dauka.

Hasieratik, hirietako etxebizitzen prezioak ikusi ditugunetik bertatik, Donostiako balioa besteena baino askoz handiagoa dela ohartu gara. Orain proposatzen duguna da $\bar{X} \pm 2S$ tartea erabiltzea Donostiako balioa zein neurritan den handia nabarmentzeko.

Mexikoko Zerbitzu Meteorologikoak 2008an gertatutako ekaitz gogorren kopuru handia nabarmentzeko honela zioen:

2008ko abuztuan, 126 ekaitz gogor gertatu ziren estatu osoan. Ekaitz kopuru hori hilabeteko batezbestekoa (87,5 ekaitz) baino handiagoa da. 2008ko abuztuko errekorrak batezbestekoa plus desbideratzearen mugak gainditzen ditu.¹⁷

Madriko Erkidegoko industriako atlasean erabiltzen dira batezbestekoa plus-minus desbideratzeko tartek, udalerrien espezializazio ekonomikoa zehazteko:

ESPEZIALIZAZIO-BALIOEN KALKULUA

Nelson-en indizea erabili da industria-jardueren espezializazio-maila kalkulatzeko. Indizea oinarritzen da desbideratze tipikoak banaketa baten datuen sakabanaketa neurtzeko daukan gaitasunean, hain zuzen ere zehaztutako mailaren gainetik dauden balioak bereizteko. Hiriko barruti bakoitzerako hamaika industria-jardueraren parte-hartzeak, ehunekotan, kalkulatu dira. Ondoren, jarduera bakoitzaren batezbestekoa eta desbideratze tipikoa kalkulatu dira. Hiriko barrutia espezializatuta dagoela esateko muga finkatzen da, batezbestekoari desbideratze tipikoaren balio bat, bi edo hiru gehituz. Gure kasuan muga jarri dugu: batezbestekoa + bi desbideratzeen batuketa. Hainbat barrutitan gerta daiteke jarritako mugak ez gainditzea, eta, ondorioz, barruti horiek espezializazio kategoriarik gabe geratzea. Gerta daiteke, baita ere, barruti batean bi edo espezializazio gehiago agertzea. Azken kasu horretan, barrutiei egokitu zaie maila handiagoko espezializazioa.¹⁸

¹⁷ CONAGUA. Comisión Nacional del Agua. (2008): «Informe de tormentas intensas registradas durante el mes de agosto de 2008 en México».

<http://smn.cna.gob.mx/noticias/tormentas-jul08.pdf>

¹⁸ Comunidad de Madrid, Consejería de Economía e Innovación Tecnológica (2007): «Atlas

Andaluziako urteko prezipitazioen bilakaerari buruzko ikerketa batean, egileak erabiltzen ditu batezbestekoa plus-minus desbideratzeko tarteak¹⁹. Ikerketaren egilea tarte horietaz baliatu da hezetasun-mailen sailkapen bat sortzeko, urteko prezipitazio kopuruaren arabera. Egindako sailkapena erabiliz edozein urtetako prezipitazio kopurua jakinda (Andaluziakoa), urte horren hezetasun maila balioetsi egin daiteke.


Sailkapena egiteko egileak erabili duen prozedura aurreko atalean azaldu duguna da. Aztertu dituen urteko prezipitazio kopuruaren batezbestekoari desbideratze estandarraren balioa gehitu eta kendu dio. Prozedura horren bidez hiru tarte sortu ditu, batezbestekoari desbideratze unitate bat, bi edo hiru gehituz eta kenduz.


Urteko hezetasun-maila	Sailkatzeko irizpidea
Urte oso lehorra	$R \text{ urtea} < \bar{X} - S$
Urte lehorra	$R \text{ urtea} < \bar{X} - 1/4 S \quad \text{eta} \quad > \bar{X} - S$
Urte normala/lehorra	$R \text{ urtea} < \bar{X} \quad \text{eta} \quad > \bar{X} - S$
Urte normala/hezea	$R \text{ urtea} < \bar{X} \quad \text{eta} \quad > \bar{X} + S$
Urte hezea	$R \text{ urtea} < \bar{X} \quad \text{eta} \quad > \bar{X} + S$
Urte oso hezea	$R \text{ urtea} < \bar{X} + 1/4 S \quad \text{eta} \quad > \bar{X} + S$
Urte bereziki hezea	$R \text{ urtea} < \bar{X} + S \quad \text{eta} \quad > \bar{X} + 2S$
Urte izugarri hezea	$R \text{ urtea} < \bar{X} + 3S$

R: eskualdeko batez besteko prezipitazioa

de la industria en la comunidad de Madrid». 167. or.

¹⁹ Castillo Requena, J.M. (2000): «Evolución de la precipitación anual en las regiones pluviométricas de Andalucía». Cuadernos Geográficos, 30, 123-141. or.

 Urtea oso lehorra dela esango da, baldin eta urteko prezipitazio kopurua, dagokion eskualdeko batezbestekoaren eta desbideratze estandarraren arteko kendura baino txikiagoa denean.

 Urtea lehorra dela esango da, baldin eta urteko prezipitazio kopurua, dagokion eskualdeko batezbestekoaren eta desbideratze estandarraren arteko kendura baino handiagoa izanda, batezbestekoaren eta desbideratze estandarraren arteko kenduraren laurdena baino txikiagoa denean.

Desbideratze estandarrari buruzko azalpenak bukatzeko, neurri honen ezaugarri nagusiak berrikusiko ditugu:

- ❖ Desbideratze estandarra edo tipikoa da datu multzoen aldakortasuna zehazteko tresna estatistikoa. Estatistikako tresnen barnean datuen sakabaneta neurtzeko prozeduren kategorian sartzen da.
- ❖ Desbideratzearen balioa erabil daiteke bakarrik batezbestekoaren osagarri modura edo, batezbestekoari gehituta eta kenduta, maiztasun gehieneko tartearen zabalera finkatzeko. Edozein kasutan, batezbestekoaren derrigorrezko osagaia da.
- ❖ Desbideratze estandarrak aldagaiaren unitate berdina erabiltzen du: ikusitako adibidean metro karratuko euroak.

Aldakuntza-koefizientea

Lehenago ikusi dugun bezala, desbideratze estandarrak erabiltzen ditu aldagaien unitate berdinak. Etxebizitzen prezioei buruzko adibidean ikusi dugu aldagaiaren unitatea euroa zela (metro karratukoa); beraz, desbideratze estandarraren unitatea ere euroa zen. Desbideratze estandarraren ezaugarri hori zeharo erabilgarria suertatzen da, aldagaien aldakortasuna aztertzen dugunean; batezbestekoaren eta desbideratzearen emaitzak lortu ondoren, ez da beste eragiketarik egin behar, eta, ondorioz, nahikoa da bi emaitzak elkarren alboan jartzea aldagaien datu gehienen balioak sumatzeko.

Desbideratzeari buruz aipatutako ezaugarria ez da egoera guztietara egokitzen, zeren bi aldagai edo gehiagoren arteko aldakortasuna konparatzeko beti ez

baita egokia. Horren inguruan sor daitezkeen, eta sortzen diren, arazoak adibide baten bidez azalduko ditugu.

Etxebizitzen prezioak aztertzen jarraituko dugu, baina kasu honetan bi autonomia-erkidegotako hirietako prezioak, 2007koak, ikusiko ditugu.

Euskadiko hiriak	Etxebizitzen prezioak €/m ²	Galiziako hiriak	Etxebizitzen prezioak €/m ²
Basauri	2.120	Monforte de Lemos	898
Hernani	2.255	O Barco de Valdeorras	1.080
Erandio	2.571	Ames	1.206
Arrasate	2.577	Ferrol	1.391
Portugalete	2.753	Pontevedra	1.477
Santurtzi	2.781	Viveiro	1.526
Leioa	2.798	Lugo	1.540
Barakaldo	2.916	Ourense	1.742
Gasteiz	2.988	Santiago Compostelakoa	1.786
Irun	3.097	Coruña	1.999
Bilbo	3.268		
Getxo	3.272		
Donostia	4.061		

Taulako datuak ordenatuta daudenez (balio txikienetatik handienetara), erraza da ikustea hirietako prezioen aldakortasuna handia dela. Horretaz jabetzeko nahikoa da ikustea bi erkidegoetan hiri garestieneko prezioa hiri merkeeneko prezioaren bikoitza edo ia bikoitza dela. Baina horretaz gain, esan daiteke antzekoa dela bi erkidegotako aldakortasuna? Galderari erantzun ahal izateko, bi datu multzoen batezbestekoa eta desbideratze estandarra kalkulatuko ditugu.

	Euskadi	Galizia
Batezbestekoa	2.881,31	1.513
Desbideratze estandarra	476,14	340,51

Erkidegoen arteko aldeak handiak dira, batez ere batezbestekoei dagokienez. Batezbesteko eta desbideratze-balio handienak Euskadikoak dira.

Desbideratze estandarraren emaitzei begira pentsa genezake Euskadiko aldakortasuna Galiziakoa baino handiagoa dela. Alabaina, hori akatsa litzateke. Gogoratu behar dugu desbideratze estandarra lortzen dela batezbestekoaren eta balioen arteko diferentziak kalkulatu. Hori dela eta, desbideratzearen emaitzaren magnitudea eta balioen magnitudea lotuta daude. Hortaz, zenbat eta handiagoak izan joera izango da aldagaiaren balioak batezbestekoaren eta desbideratzearen emaitzak ere gero eta handiagoak izatea. Azken finean, aldakortasun txikiagoko aldagai baten desbideratzea altua izan daiteke, eta alderantziz. Hori da, hain zuzen ere, jarritako adibidean gertatzen dena: Euskadiko hirietako prezioen aldakortasuna Galiziakoena baino txikiagoa izanda ere, erkidego horren desbideratzea altuagoa da, hango prezio altuek eragina baitute.

Datu multzo baten baino gehiagoren aldakortasunak konparatzeko tresna badago, datu multzoen balioen magnitudearen menpean ez dagoena. Aipatzen dugun tresna aldakuntza-koefizientea da. Aldakuntza-koefizientearen bidez kalkulatu egiten da desbideratze estandarraren tamaina batezbestekoaren tamainaren arabera:

$$ak = \frac{S}{\bar{X}} \cdot 100$$

Ikusten dugunez, eragiketa erraza da, desbideratze estandarren balioa batezbestekoaren balioaz zatitzean datza eta. Horren ondoren, ehunekoa lortzeko biderketa egiten da.

Euskadi	$ak = \frac{S}{\bar{x}} = \frac{476,14}{2.881,31} = 0,1652$	→	%16,52
Galizia	$ak = \frac{S}{\bar{x}} = \frac{340,51}{1.513} = 0,2251$	→	%22,51

Segidan kalkulatuko ditugu bi autonomia-erkidegoetako datuen aldakuntza koefizienteak. Horrek emango digu lehen egindako baieztapena frogatzeko aukera, alegia, Euskadi ez dela aldakortasun handieneko erkidegoa.

Aldakuntza-koefizienteen emaitzen aurrean argi ikusten da Euskadiko erkidegoaren prezioen aldakortasuna bi erkidegoetatik txikiena dela, edo, bestela esanda, askoz antzekoagoak direla Euskadiko hirietako etxebizitzaren prezioak Galiziako hirietakoak baino.

Lehen aipatu dugun Andaluziako prezipitazioei buruzko ikerketan²⁰ aldakuntza-koefizientea erabili dute, erkidego horretako prezipitazio-aldakortasun handia agerian jartzeko. Horretarako, Andaluziako eskualdeetako aldakuntza-koefizienteak eta Espainia penintsularreko koefizientea konparatzen dira:

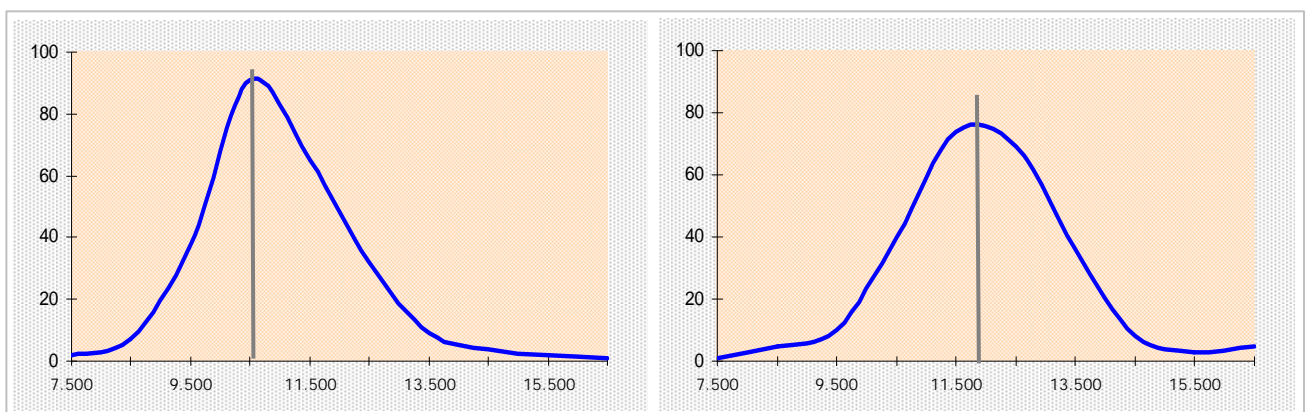
²⁰ Castillo Requena, J.M. (2000)

Desbideratze handiagoa eduki arren, aldakuntza-koefizientea txikiagoa da

	Batezbestekoa	Desbideratze estandarra	Aldakuntza-koefizientea
Guadalquivirreko harana	600,47	180,45	30,05
Hegoaldeko harana	537,82	165,72	30,81
Erdialdeko unitatea	605,26	127,55	21,07
Mediterraneo aldea	527,69	116,60	22,10
Espainia penintsularra	667,09	119,25	17,88

Kasu honetan, Hegoaldeko eta Guadalquivirreko haranetako neurriei begiratu behar diegu. Hegoaldeko haranak Guadalquivirrekoa baino desbideratze txikiagoa eduki arren, aldakuntza-koefiziente handiagoa dauka. Erdialdeko unitatea eta Mediterraneo aldearen kasuetan antzekoa gertatzen da.

SIMETRIAREN ANALISIA



Irudian ikusten ditugun bi maiztasun-poligonoak Euskadiko udalerrietako errenta pertsonal erabilgarriaren datuei dagozkie. Ezkerreko grafikoa 2001ekoa da; eskuinekoa, 2003koa.

Banaketa-formari dagokionez, badago bi maiztasun-poligonoen artean oso nabarmena den desberdintasun bat. 2001eko grafikoan, banaketako erdiguneak bereizten dituen bi alboen forma desberdina da. Hori dela eta, banaketa asimetrikoa dela esaten dugu. 2003ko banaketan, berriz, aipatutako bi alboen forma oso antzekoa da. Kasu horretan, banaketa simetrikoa dela esaten dugu.

Simetria kontzeptuaren esanahia argitzeko hiztegi joko dugu:

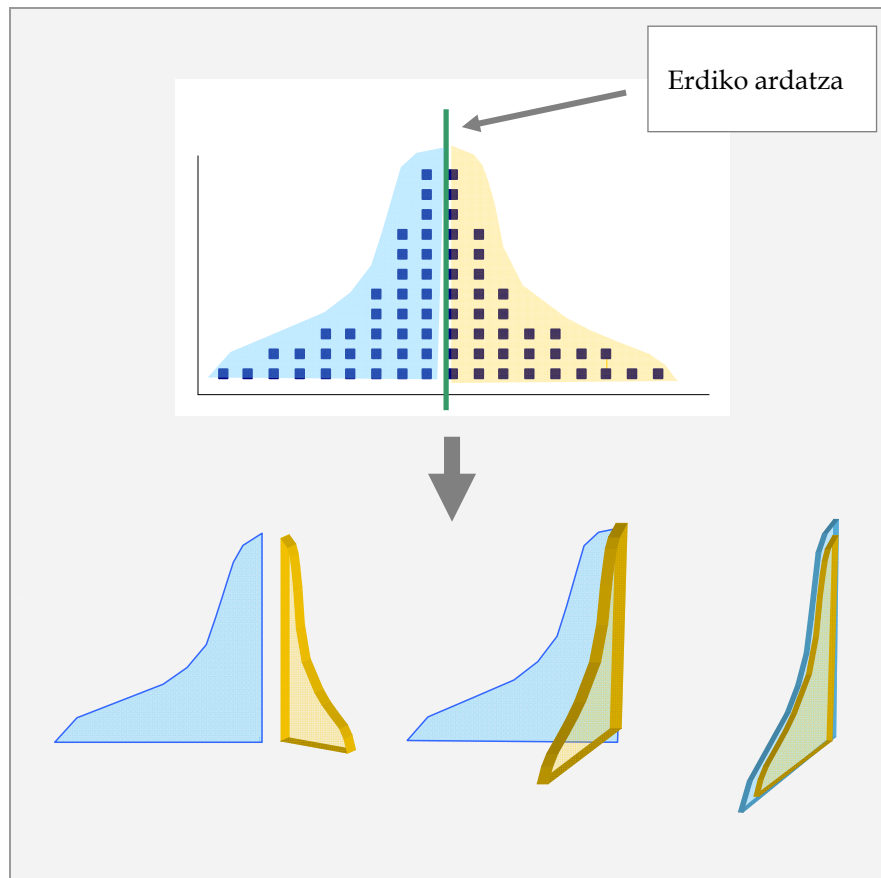
Multzo bateko elementuek, edo multzo bi edo gehiagok, puntu, lerro edo plano batekiko duten posizio-, neurri- edo forma-korrespondentzia. Delako puntuari simetria-zentro, lerroari simetria-ardatz eta planoari simetria-plano esaten zaie. *Zientzia eta Teknologiaren Hiztegi Entziklopedikoa*

Geom. Irudi edo gauzaki baten zatien egokitasuna, baten eta bestearen zati bakoitza ardatz, puntu edo planoko bati buruz (simetria ardatz, puntu edo planoari buruz) distantzia eta posizio berean egotea duena ezaugarri. *Entziklopedia laburra*

Ikusten dugunez, geometria-arloko kontzeptua da, datu-multzoen analisisian erabiltzen dena, datuen banaketari buruzko ezaugarri adierazgarriak zehazteko baliotsua baita.

Simetriaren analisi grafikoa

Aldagai baten datu-banaketaren simetria aztertzeko maiztasun-banaketako grafikoa erabil daiteke. Goian jaso ditugun definizioak aplikatuz, esango dugu banaketa bat simetrikoa dela, baldin eta, simetria-ardatzetik bananduta, sortzen diren bi alboen formak berdinak badira.

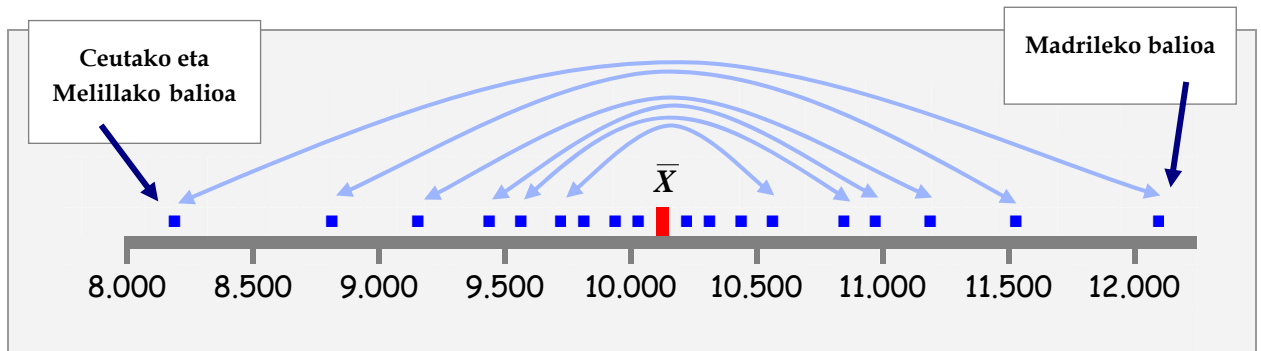


Banaketa simetriko baten grafikoa bitan tolestuz gero erdiko ardatzetik alboko bi zatiak bat datoz.

Banaketen bi alboen arteko antzekotasunak uzten digu finkatzen banaketaren izaera simetrikoa edo asimetrikoa. Horretaz gain, guri interesatzen zaiguna da ondo ezagutzea eta barneratzea simetriaren esanahia, datu multzo bati dagokionez; hortik aurrera funtsezkoa izango da, baita ere, jakitea nola finka daitekeen simetria- edo asimetria-neurria.

Banaketa baten bi alboak berdinak izateak esan nahi du bi alboetan balio kopurua antzekoa dela eta, beraz, balioen sakabanaketak ere –hots, batezbestekarekiko distantziak– antzekoa izan behar duela.

Argiago ikusteko, ibiltartea ordezkatzeko duen grafiko batean datu multzo hipotetiko baten balioak jasoko ditugu. Asmatutako datu multzoa da estatuko autonomia-erkidego bakoitzeko pertsona talde batek 2006an gastatu zuten batez besteko diru kopurua.



Lerroaren erdiko ikur gorria batezbestekoari dagokio (10.150,98 €).

Ikur urdinak erkidego bakoitzeko batez besteko gastuari dagozkio.

Lerro urdinez lotzen dituzte albo banatako balioak, batezbestekotik distantzia berera daudenak.

Ibiltarte barneko banaketaren irudiari eta datu-taulari begiratu gero²¹, konturatuko gara batezbestekoa baino gutxiago gastatu zuten taldeek eta gehiago gastatu zutenek neurri berdinean egin zutela. Esaterako, Madrileko taldeak gastatu zuten batezbestekoa baino 1.960,85 € gehiago; Ceutakoak eta Melillakoak gastatu zuten batezbestekotik aldentzen den kantitate berdina (600,86 €), baina, kasu horretan, batezbestekoa baino txikiagoa den kantitatea. Kataluniako eta Zaragozako balioak konparatu gero, emaitza berdina lortuko dugu: Kataluniakoek gastatu zuten batezbestekoa baino 1.359,99 € gehiago eta Zaragozakoek baino 1.359,99 € gutxiago.




²¹ Sortutako datu multzoa eranskinean ikusi. 6. taula.

Autonomia-erkidegoak	Gastua €	Batezbestekoarekiko aldeak
Ceuta eta Melilla	8.190,13	-1.960,85
Extremadura	8.565,99	-1.359,99
Gaztela-Mantxa	8.945,31	-980,67
Kanariak	9.257,72	-740,26
Errioxa	9.320,97	-563,22
Murtzia	9.391,17	-438,06
Andaluzia	9.554,63	-312,90
Kantabria	9.668,21	-187,74
Gaztela eta León	10.000,27	-62,58
Galizia	10.083,93	62,58
Asturias	10.662,51	187,74
Aragoi	10.684,19	312,90
Valentziako Erkidegoa	10.720,13	438,06
Balear Uharteak	11.208,57	563,22
Nafarroa	11.542,70	740,26
Euskadi	11.855,74	980,67
Katalunia	11.994,94	1.359,99
Madril	12.111,83	1.960,85


Taulako balio guztiekin eragiketa berdina eginez gero, 9 erkidego bikote lortuko ditugu. Bikote bakoitzeko balioak batezbestekotik distantzia berdiner daude, bata batezbestekoaren ezkerraldean eta bestea eskuinaldean.

Erkidego bikote distantziakideak		Batezbestearekiko distantzia
Ceuta eta Melilla	Madril	± 1.960,85
Extremadura	Katalunia	± 1.359,99
Gaztela-Mantxa	Euskadi	± 980,67
Kanariak	Nafarroa	± 740,26
Errioxa	Balearr Uharteak	± 563,22
Murtzia	Valentziako Erkidegoa	± 438,06
Andaluzia	Aragoi	± 312,90
Kantabria	Asturias	± 187,74
Gaztela eta León	Galizia	± 62,58

Azkenik, laburpen modura, asmatutako balioen banaketari buruz atera daitezkeen emaitzak honako hauek dira:

-  Batezbestekoaren albo banatan balio kopuru berdina dago.
-  Albo banatako balioen sakabanaketa ere berdina da.
-  Banaketa, beraz, zeharo simetrikoa da.

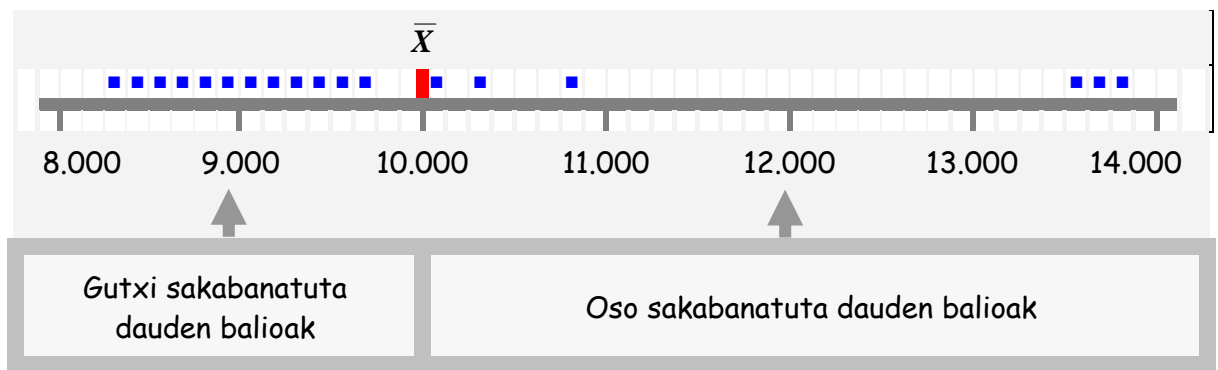
Alabaina, zein da simetriaren esanahia aztertu dugun datu multzoan? Nola interpreta daiteke banaketa simetrikoa izatea?

-  Ez dagoela erkidego multzorik, edo azpimultzorik, beste portaera bat, gastatu duten diru kopuruari dagokionez, eduki duena. Talde gehienek baino gutxiago edo gehiago gastatu dutenek batez besteko gastuarekiko diferentzia antzekoak eduki dituzte.




Banaketa simetroari buruzko adibidea ikusi ondoren, hurrengo pausoa izango da aurkeztea nola izan daitezkeen banaketa asimetrikoa. Oraingo honetan, aurreko aldagai berdina eta datu hipotetikoak erabiliko ditugu, baina erkidego

bakoitzaren balioak aldatu²².

Batezbestekoa 10.051,61 eurokoa da.




Irudian ikusten dugunez, oraingo datu-banaketak ez dauka zerikusirik aurrekoarekin:


-  18 balioetatik 12 daude batezbestekotik gertu, haren ezkerrean.
-  Batezbestekoaren eskuinean balioetatik heren bat besterik ez dago. Horietatik hiru daude ibiltarteko eskuineko muturrean, beste balioetatik oso urrun.
-  Balioen batezbestekoa, 10.051,61 eurokoa, ibiltartearen ezkeralderantz alboratuta dago.

Bistakoa da datu-banaketa erabat asimetricoa dela, batezbestekoaren albo banatako sakabanaketa guztiz desberdina baita. Ezkerreko balioak elkarrengandik gertu daude; eskuinekoak, berriz, elkarren artean urrun.

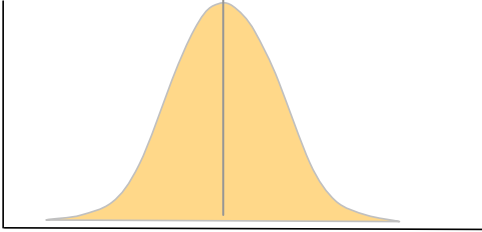
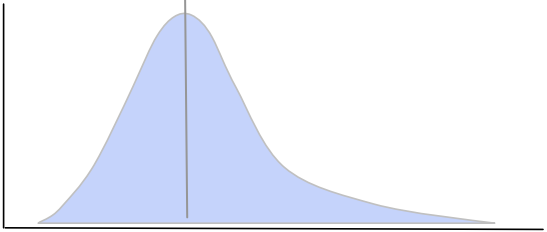
Zein da asimetriaren esanahia aztertu dugun datu multzoan? Nola interpreta daiteke banaketa asimetricoa izatea?

²² Sortutako datu multzoa eranskinean ikusi. 7. taula.

 Aurreko adibidean ez bezala, oraingo honetan taldeen portaera ez da homogenea izan, talde gehienen gastuak ez baitira antzekoak izan.

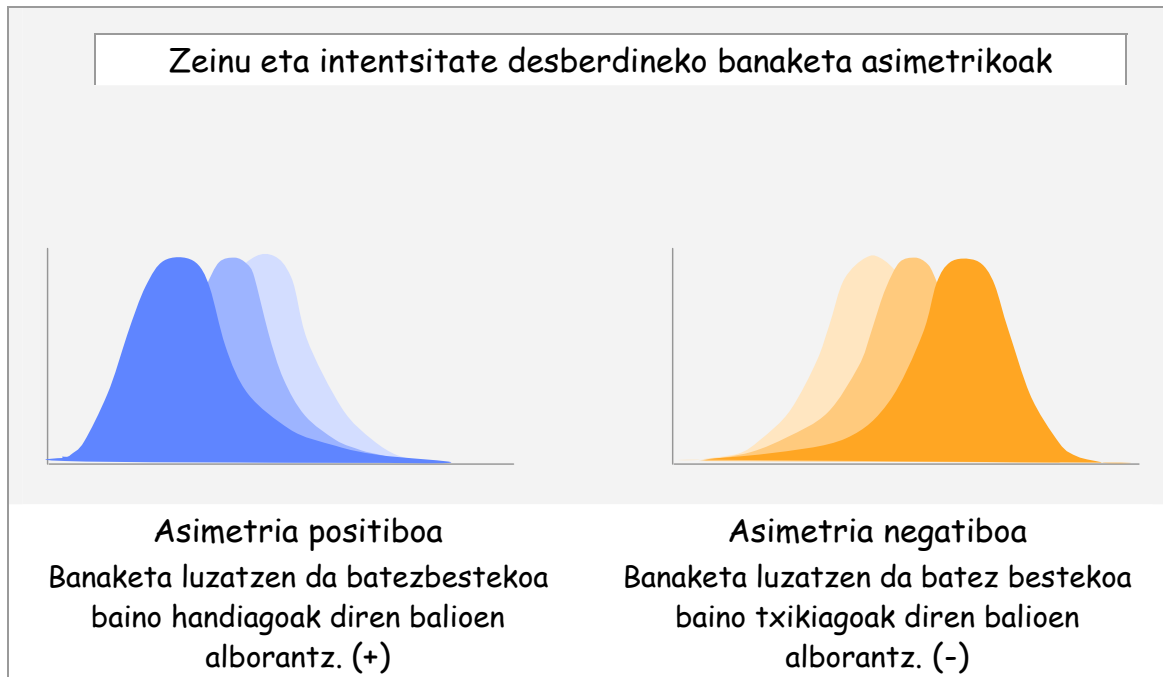
 Badago multzo bat, 12 taldek osatutakoa, gastu txikiak eduki dituena. Beste sei taldeek nabarmen gastu handiagoak egin dituzte. Sei horietatik hiruk portaera zeharo berezia izan dute, gastu bereziki handiak egin dituzte eta.

Adibide gisa erabili ditugun bi banaketaren maiztasun-poligonoak egingo bagenitu, irudi hauetan ikusten diren grafikoaren oso antzekoak izango lirateke.

Banaketa simetriko baten grafikoa	Banaketa asimetriko baten grafikoa
	
<p>Banaketaren bi alboak berdina: banaketa simetrikoa</p>	<p>Batezbestekoa baino nabarmen handiagoak diren balioak egoteagatik, banaketaren eskuineko alboa luzatzen da, eta, horrekin batera, banaketa ere bai.</p>

Banaketa asimetrikoa erakusteko erabili dugun adibidean luzatu egiten da banaketaren eskuinaldea, sakabanaketa handiko balioak datu multzoko altuenak direlako. Horrela gertatzen denean, banaketak asimetria positiboa duela esaten da. Luzamendu bat gerta daiteke, baita ere, banaketaren ezker aldean. Kasu horretan, sakabanaketa handieneko balioak datu multzoko baxuenak dira, alegia, banaketaren ezkerreko alloan kokatzen diren balioak. Horrela gertatuz gero, esaten da banaketak asimetria negatiboa daukala.

Banaketen eskuinaldean edo ezker aldean gertatzen diren luzamenduak izan daitezke hainbat luzeratakoak, handiagoak edo txikiagoak, balioen sakabanaketaren mailaren arabera.



Dagoeneko ikusi dugu banaketak asimetrikoak direla batezbestekoaren albo bateko edo besteko sakabanaketa-mailak desberdinak direnean. Ikusi dugu, baita ere, banaketen asimetria aipatzen dugunean datu multzo baten ezaugarri bati buruz ari garela, alegia, aztertu eta deskribatu behar den ezaugarriari buruz. Segidan ikusiko dugu zein den asimetria-analisiaren ekarpena datu multzoa deskribatzeko eta, azkenik, nola lortu asimetriaren neurri bat.

Asimetria-analisiaren ekarpena

Banaketen asimetria-maila neurtzeko beharraz eta balioaz ohartzeko, ikusiko ditugu batezbestekoaren eta desbideratze estandarren bidez deskribatzen diren ezaugarriak. Horrela lortuko dugu, baita ere, neurri horiek azaltzen ez dutena argitzea, hain zuzen ere asimetria-analisiaren bidez ezagutu daitekeena. Horretarako, lehengo datu berak erabiliko ditugu, autonomia-erkidegoetako pertsonako batez besteko gastuenak (alegiak).

Autonomia-erkidegoetako pertsonako batez besteko gastua				
	Banaketa simetrikoa A		Banaketa asimetrikoa B	
Batezbestekoa	10.150,98		10.051,61	
Desbideratze estandarra	933,95		1.747,47	
Batezbestekoa \pm desbideratzea	9.217,03	11.084,93	8.304,14	11.799,09
$\bar{X} \pm S$ tarteen zabalera	1.868		3.495	
Balio handiena eta txikiena	8.190,13	12.111,83	8.400,13	13.836,83
Ibiltartea	3.921,7		5.436,7	

Taulan ikusten diren batez bestekoen emaitzen arabera bi datu multzoak nahiko antzekoak dira, bien artean 100 euroren aldea besterik ez dago eta. Ondorio zeharo desberdinak lortzen dira, bi aldagaien desbideratze estandarren emaitzei begiratzen diegunean, bi aldagaien emaitzak zeharo desberdinak dira eta. **B** banaketaren desbideratzea askoz handiagoa da **A** banaketarena baino, ia bikoitza dela esan daiteke.

Bi aldagaien aldakortasunaren diferentziak agerian geratzen dira. **B** banaketaren ibiltartea eta batezbestekoa \pm desbideratze estandarreko tartea dira diferentziaren adierazle esanguratsuak.

Taulan jasotzen diren neurrietan aldakortasun-aldeak nabarmenak badira ere, ez da inon ikusten diferentzia horien kausa edo zergatia. Ez da ikusten, esaterako, badirela **B** datu multzoan 3 balio, 3 autonomia-erkidego, aparteko portaera bat izan dutenak, pertsonako batez besteko gastuei dagokienez. Asimetriari buruzko neurria erabili beharko dugu aparteko gastuak izan dituzten erkidegoen egoera deskribatzeko.


Asimetria-mailaren kuantifikazioa


Banaketa baten maiztasun poligonoari begiratzea nahikoa da banaketa simetrikoa edo asimetrikoa den jakiteko. Alabaina, datu multzo baten aldakortasuna deskribatu nahi dugunean, erosoagoa da asimetriaren neurri bat edukitzea.

Estatistikak baditu tresna batzuk datu-banaketen asimetriari buruzko neurri bat lortzeko. Horietatik bat Fisher-en asimetria-koefizientea da. Koefiziente hori oso erabilia izateaz gain, kalkulatzeko erraza da.

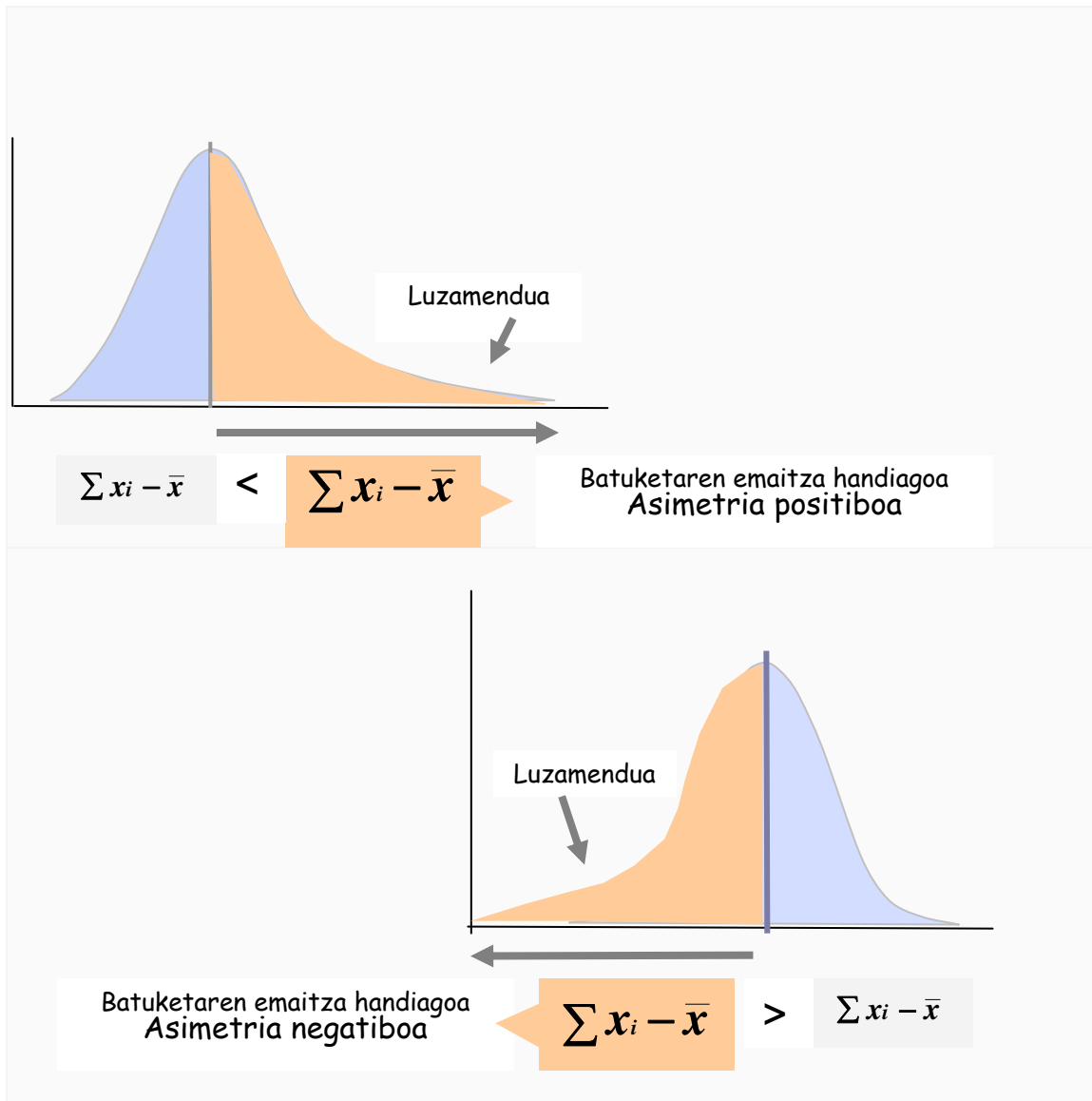
$$G_1 = \frac{\sum (x_i - \bar{x})^3}{N \cdot S^3}$$

Oraingo honetan ere formularen oinarria da banakako balioen eta batezbestekoaren arteko diferentzien kalkulua. Alabaina, desbideratze estandarrean ez bezala, asimetria-koefizientearen batezbestekoaren eta balioen arteko diferentzien zeinua mantentzen da, diferentzien kuboaren batuketara egiten baita. Diferentzien zeinuak mantentzen direnez, asimetria-koefizientearen emaitza izan daiteke positiboa edo negatiboa:

 Batezbestekoaren ezker aldeko balioen diferentziak eskuinaldekoak baino handiagoak direnean, koefizientearen emaitza negatiboa da. Hori gertatzen da, banaketaren asimetria negatiboa denean, alegia, balio txikiagoen alborantz luzatzen denean.

 Batezbestekoaren eskuinaldeko balioen diferentziak ezker aldekoak baino handiagoak direnean, koefizientearen emaitza positiboa da. Hori gertatzen da banaketaren asimetria positiboa denean, alegia, balio handiagoen alborantz luzatzen denean.

Batezbestekoaren albo bateko eta besteko balioen diferentziak antzekoak direnean, asimetria-koefizientearen emaitza 0tik gertu dago.



BALIOEN TALDEKATZE-MAILAREN KUANTIFIKAZIOA: KURTOSIA

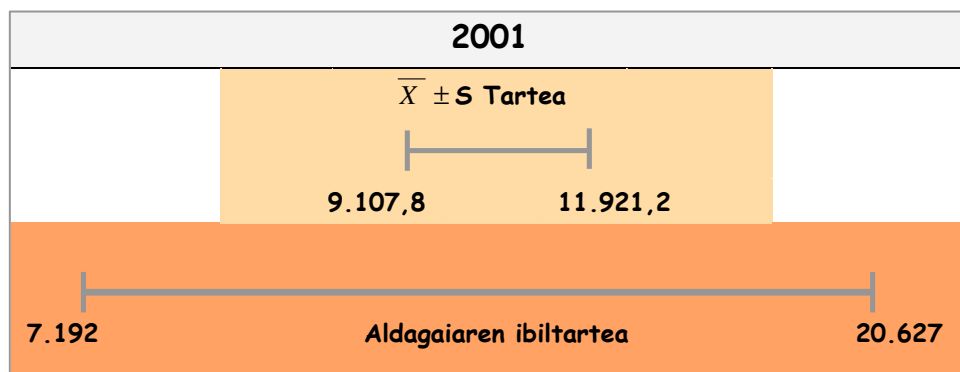
Aurreko ataletan ikusi dugu batezbestekoa +/- desbideratze estandarreko tartearen bidez lortzen dugula datu multzo baten balio gehienen aldakortasunari buruzko deskribapen adierazgarria. Aipatutako tartearen zabalerak erakusten digu zein neurritan diren antzekoak datu multzoaren balio gehienak.

Gogora ditzagun berriro *Euskadiko udalerrietako pertsonako batez besteko errenta erabilgarriaren* kasuan lortutako tartearen balioak:



Maiztasun gehieneko tarteari begira badakigu udalerrietatik % 68 inguruan biztanleen errenta erabilgarria izango dela 9.107,76 eurotik 11.921,18 eurora bitartekoa. Tartearen zabalera, 2.813,43 eurokoa dena, balio gehienen arteko aldakortasunari dagokio. Aldakortasun tarte hori oso txikia da balio guztien aldakortasunarekin –hots, ibiltartearekin– konparatuta.

Bi aldakortasunen maila ikusita, helburua izango da lortzea tresna bat haien maila desberdinak deskribatzeko erabil litekeena. Bilatu beharko dugu, beraz, neurri mota bat deskribatzeko maiztasun gehieneko tartearen aldakortasuna balio guztien aldakortasunarekiko. Ibiltartearen kontzeptua irudikatzeko grafikoa berriro erabiliko dugu:



Irudian ikusten ditugun bi zuzen zatiek $\bar{X} \pm S$ tartearen zabalera eta ibiltartearen zabalera irudikatzen dituzte. Lehen zuzen zatia oso txikia da bigarrenarekin konparatuta. Luzera-desberdintasun horren esanahia, dagoeneko, argi dago: nahiz eta udalerrri guztietako errentaren aldakortasuna handia izan, gehienetakoa nahiko antzekoa da, 9.107,8 euroren 11.921,2 euroren bitartean dago eta. Hortaz, $\bar{X} \pm S$ tartearen barneko balioak taldekatzen dira ibiltarteko zati txikian.

Ondoren, aurreko irudi berdina ikusiko dugu, baina, oraingo honetan, 2003ko errenta balioei dagokiena:




Bi irudien arteko desberdintasunak, eta ondorioz bi banaketen desberdintasunak, begi-bistan daude: 2003ko ibiltartea askoz txikiagoa da 2001ekoa baino. Bestalde, $\bar{X} \pm S$ tarteko balioen taldekatzea zertxobait txikiagoa da.


Bi banaketen analisi konparatiboa egin nahi izanez gero, ikusitako

desberdintasunak, balio guztien eta gehienen arteko harremani dagokienez, aztertu eta deskribatu beharko genituzke. Badago horretarako tresna bat, kurtosi izenekoa, $\bar{X} \pm S$ tartearen zabalera eta ibiltartearen zabalera erlazionatzen dituena.

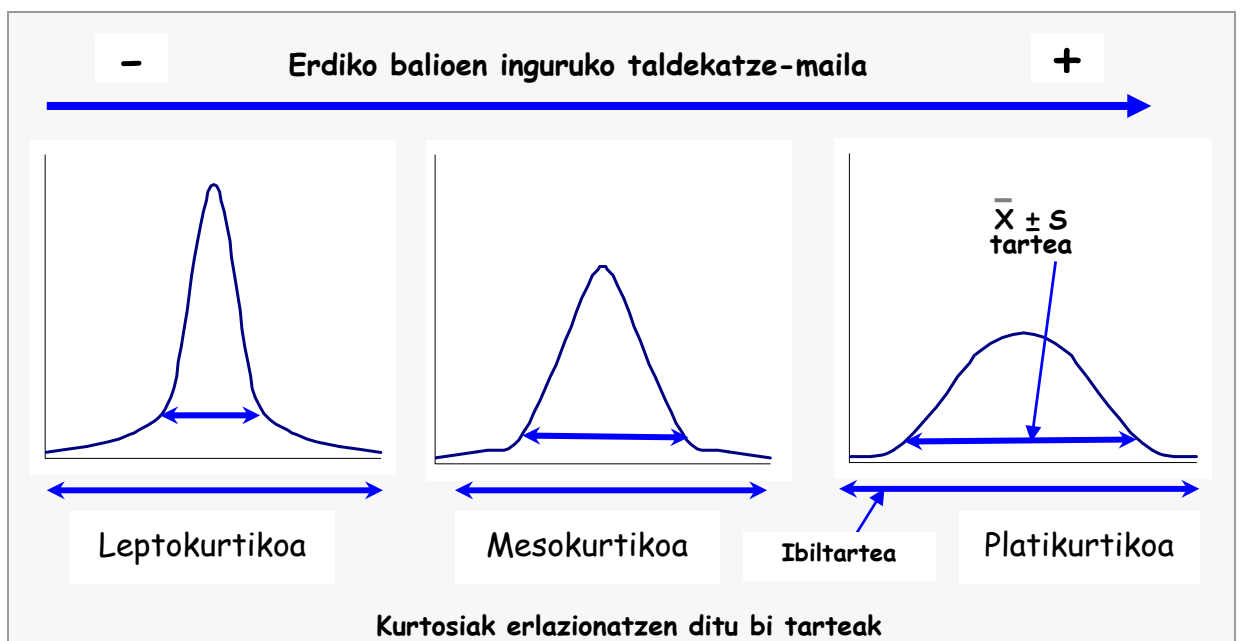
Asimetria lantzean ikusi dugun bezala, kurtosia bi modutan azter daiteke: grafiko baten bidez eta koefiziente baten bidez.


Ibiltartearen eta $\bar{X} \pm S$ tarteen artean gertatzen diren harreman desberdinen ondorioz, itxura desberdineko banaketak egoten dira:


 $\bar{X} \pm S$ tartearen barneko balioen taldekatzea oso handia denean, mutur luzea duten kanpai-itxurako banaketak gertatzen dira.


 Zenbat eta handiagoa edo zabalagoa izan $\bar{X} \pm S$ tartea, barneko balioen sakabanaketa handiagoren ondorioz, orduan eta itxura zapalagoa duten kanpai-formak agertzen dira.

Eskuarki banaketak hiru motatan sailkatzen dira, kanpaiaren luzamenduaren arabera:



 Aldagaiaren balio gehienak banaketaren erdiko balioen inguruan pilatzen direnean, banaketa leptokurtikoa dela esaten da.


 Aldagaiaren balio gehien taldekatzea, banaketaren erdiko balioen ingurukoa, txikia denean, banaketa platikurtikoa dela esaten da.


 Aldagaiaren balio gehien taldekatzea, banaketaren erdiko balioen ingurukoa, maila ertainekoa denean, banaketa mesokurtikoa dela esaten da.


Lehen esaten genuen bezala, badira tresna estatistikoak banaketen muturraren luzamendua neurtzeko. Horietako bat erabiliz banaketen luzamenduaren maila kuantifika daiteke. Asimetriaren kasuan bezala, Fisher-en neurria erabiliko dugu kurtosia neurtzeko:

$$G_2 = \frac{\sum (x_i - \bar{x})^4}{N \cdot S^4} - 3$$

Fisher-en kurtosiaren emaitza interpretatzeko irizpideak:

 $G_2 > 0$. Emaitza positiboa izanez gero –alegia, zero baino handiagoa –, banaketaren forma leptokurtikoa da.

 $G_2 < 0$. Emaitza negatiboa izanez gero –alegia, zero baino txikiagoa –, banaketaren forma platikurtikoa da.

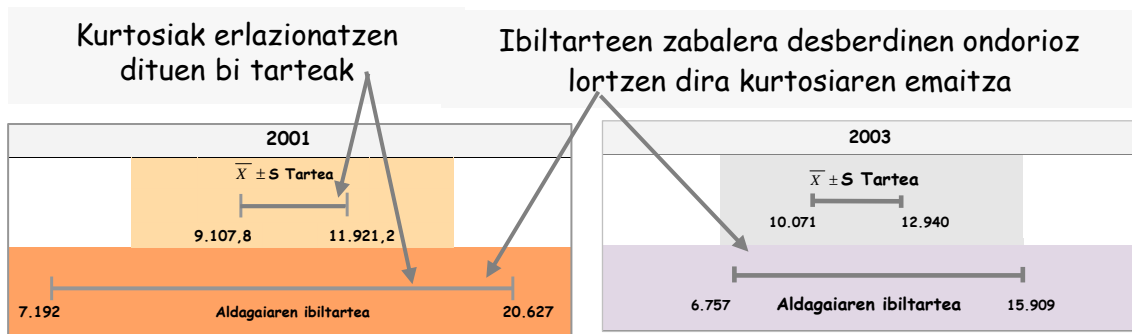
 $G_2 \approx 0$. Emaitza zero edo zero ingurukoa izanez gero, banaketaren forma mesokurtikoa da.

Segidan, *Euskadiko udalerrietako pertsonako batez besteko errenta erabilgarriaren* kurtosiaren balioak, 2001ekoak eta 2003koak, kalkulatu ditugu:

	2001	2003
Diferentzien 4. berreduren batuketa $\sum (x_i - x)^4$	13.315.132 · 10 ⁹	4.352.128 · 10 ⁹
Desbideratze estandarra	1.403,91	1.403,91
Desbideratze estandarraren 4. berredura	3.884.685 · 10 ⁶	4.200.998 · 10 ⁶

2001
$G_2 = \frac{\sum (x_i - x)^4}{N \cdot S^4} - 3 = \frac{13.315.132 \cdot 10^9}{251 \cdot 3.884.685 \cdot 10^6} - 3 = \frac{13.315.132 \cdot 10^9}{975.056 \cdot 10^9} - 3 = 10,65$
2003
$G_2 = \frac{\sum (x_i - x)^4}{N \cdot S^4} - 3 = \frac{4.352.128 \cdot 10^9}{251 \cdot 4.200.998 \cdot 10^6} - 3 = \frac{4.352.128 \cdot 10^9}{1.054.450 \cdot 10^9} - 3 = 1,27$

Lortutako emaitzek bi banaketen arteko desberdintasuna agerian jartzen dute: 2001eko banaketaren ibiltartea askoz zabalagoa izanda, erdiko balioen taldekatzea handiagoa suertatzen da, eta, ondorioz, kurtosiaren emaitzak ere handiagoa izan behar du. Dena den, balio oso desberdinak eduki arren, bi banaketak eredu leptokurtikoaren barnean sartu behar dira, eta, ondorioz, balio gehien aldakortasun txikia duten banaketen artean.

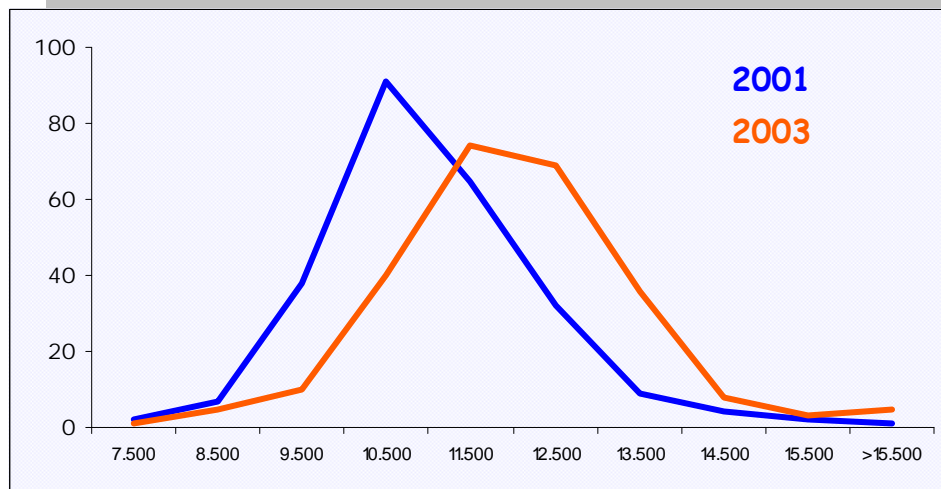


Kurtosiak bi tarteen arteko harremana neurtzen duenez, bi banaketen erdiko balioen taldekatzea antzekoa izan arren, 2001eko banaketaren emaitza 2003koa baino askoz handiagoa da.

Bestalde, badakigu 2001eko banaketaren ibiltarteen zabalera elementu arraro baten ondorioa dela. Errenta handieneko balioak (Laukiz, 20.627€) luzatzen du ibiltartea 15.346 baliotik 20.627 balioraino. Hori dela eta, kurtosiaren kalkulu guztiak errepikatu ditugu, baina, oraingo honetan, elementu arraroaren balioa kenduta:

$$G_2 = \frac{\sum (x_i - \bar{x})^4}{N \cdot S^4} - 3 = \frac{2.857.357.374 \cdot 10^6}{250 \cdot 2.459.110 \cdot 10^6} - 3 = \frac{2.857.357.374 \cdot 10^6}{614.776.103 \cdot 10^6} - 3 = 1,65$$

Bigarren honetan, 2001eko banaketarako lortutako emaitza askoz txikiagoa da eta askoz gertuago dago 2003ko banaketaren emaitzatik. Dena den, elementu arraroa kenduta ere, zertxobait handiagoa da 2001eko kurtosia, balio gehienek aldakortasuna ere zertxobait txikiagoa delako. Bi banaketek maiztasun-poligonoak ikusterakoan hobeto ohartuko gara aipatutako desberdintasunaz:



Irudi honetan ez dira ikusten ibiltarteen arteko desberdintasunak, abzisen ardatza bakarrik 15.500era iristen delako. Ondorioz, irudi honen bidez konpara daitezke bi banaketen erdiko sektoreak. Lehenago esaten genuen bezala, maiztasun gehieneko tarteei dagokienez desberdintasun txikia dago bi banaketen artean: 2003ko banaketaren aldakortasuna zertxobait handiagoa denez, *kanpaia* ere zertxobait zapalagoa da.


Azkenik, aldagaien aldakortasuna deskribatzeko erabil daitezkeen neurri guztien emaitzak ikusiko ditugu:


	2001		2003
	Balio guztiak	Elementu arrarorik gabe	
Batezbestekoa	10.514,47	10.474	11.506
Mediana	10.316	10.314	11.464
Desbideratze estandarra	1.403,91	1.252,26	1.431,65
Aldakuntza-koefizientea	13,35	11,96	12,44
Kurtosia	10,9	1,65	1,27
Asimetria-koefizientea	1,92	0,62	0,25
Ibiltartea	13.435	8.154	9.152
Balio txikiena	7.192	7.192	6.757
Balio handiena	20.627	15.346	15.909

Aldakortasuna pixka bat handiagoa da

Balio txiki gehiago: asimetria handiagoa da

Bi aldagaien ezaugarriak oso antzekoak izan arren, badira haien artean desberdintasun txikiak, hainbat emaitzaren bidez agerian jartzen direnak:





 2003ko balioen aldakortasuna pixka bat handiagoa da 2001koa baino. Aldakuntza koefizienteak eta kurtosisak dira ezaugarri honen adierazleak.

 2001ean baziren udalerrri gehiago, errenta-balioa batezbestekoa baino txikiagoa zutenak. Banaketa horren asimetria pixka bat handiagoa da 2003koa baino.

Kanpai-itxurako banaketak deskribatzeko neurrien erabilera elkartua

Kanpai-itxurako banaketen deskribapenari buruzko azalpena bukatzeko, aplikatuko ditugu ikusitako neurriak bi datu multzoren analisi konparatiboa²³ egiteko. Aztertuko ditugu bi meteorologia-behatokiren datuak, prezipitazioei buruzkoak hain zuzen ere. Toledoko eta San Fernandoko (Cádiz) behatokitako datuak dira, 1910etik 1995era bitartean bildutakoak.

Analisiaren helburua izango da bi behatokitako prezipitazioen aldakortasuna ikastea. Saiatuko gara honelako galderei erantzuten:

-  1910etik 1995era urteko prezipitazio kopuruak homogeneoak izan al dira (beren artean antzekoak)?
-  Ba al dago urteko prezipitazio kopurua edo tartea urtetan zehar ohikoa izan dena?
-  Bi behatokitako eurien portaera antzekoa izan al da? Edo, osterantzean, zertan datza desberdintasuna?
-  Egon al da aldi bereziki lehorrik edo bereziki hezerik?

Aztertzen ari garen aldagaien aldakortasunari buruz hasierako ideia bat izateko bien ibiltarteak kalkulatu ditugu:


²³ Ikusi, baita ere, eranskinetako 3. taularekin batera Zaragoza probintziako hirietako etxebizitzaren prezioen analisia.


	Toledo		San Fernando	
	Urtea	Prezipitazio kopurua	Urtea	Prezipitazio kopurua
Balio handiena	1955	573,3	1955	999
Balio txikiena	1994	188,1	1994	282,8
Ibiltartea	385,2		716,2	

Aldagai baten ibiltartea bestearena baino askoz handiagoa dela ikusten dugu. San Fernandoko behatokiko urte lehorrenaren eta hezeenaren arteko diferentzia Toledoko behatokikoaren ia bikoitza da. Horren handia den diferentzia gertatzen da, batez ere bi behatokietako urte hezeeneko balioei dagokionez. 1995ean, bi behatokietako urte euritsuenean, San Fernandon jaso ziren 425,7 mm gehiago Toledon baino.

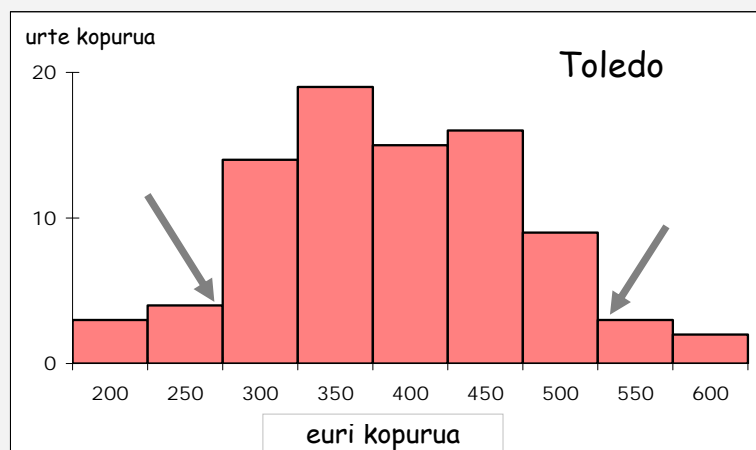
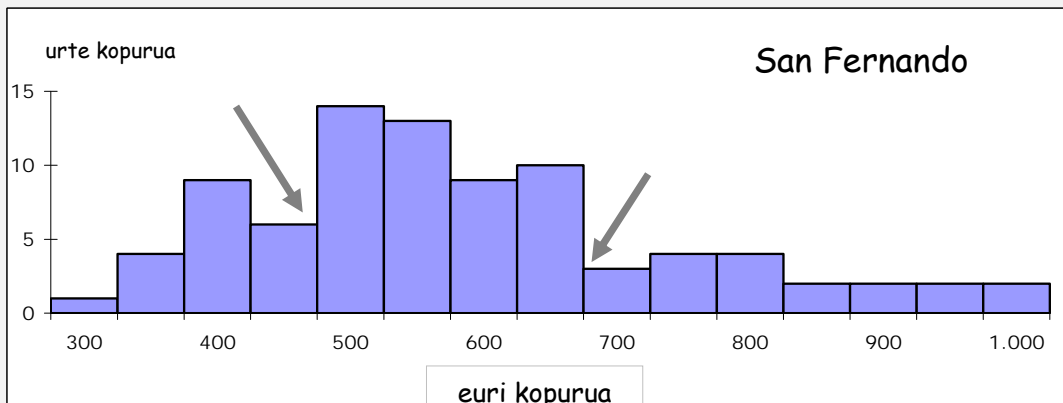
Ibiltarteen balioei atxikirik, San Fernandoko behatokiko aldakortasuna askoz handiagoa da Toledokoa baino; hori ukaezina da. Dena den, oraingoz ezin dugu jakin San Fernandoko aldakortasuna aldi osoko ezaugarria den ala urte euritsu bakan batzuen ondorioa. Histogramek lagunduko digute horren erantzuna bilatzen.

Dena den, histogramak egin baino lehenago grafiko mota horien bi ezaugarri garrantzitsuak berriro aipatuko ditugu:


 Histogramak ez dira erabilgarriak bakarrik azterketa estatistiko baten emaitzak aurkezteko. Histogramak tresnak zeharo erabilgarriak dira datuen miaketa egiteko eta, oro har, aldagaiaren banaketan ezaugarriak ezagutzeko. Histogramen helburuen arabera, irizpideak ere desberdinak izaten dira.


 Histogramek komunikatzen duten ideia, deskribatzen duten aldagaiari buruzkoa, aldakorra izaten da aukeratu diren tarte kopuruaren eta zabaleraren arabera.

Gure kasuan, histogramekin hasiko gara datuen miaketa egiteko asmoz. Hori dela eta, tarte askotako histogramak, eta, ondorioz, zabalera gutxikoak, egingo ditugu. Saiatuko gara ikusten maiztasunen bat-bateko beherakadarik edo gorakadarik dauden eta, horrela bada, ibiltarteko zein puntuetan ematen diren.



Egindako histogramek erakusten dute bi kasuetan maiztasunen bat-bateko aldaketak gertatzen direla:

 San Fernandoko histograman, bat-bateko maiztasunen aldaketak 450-500 tartearen azpitik eta 600-650 tartearen gainetik gertatzen dira.

 Toledoko histograman badago bat-bateko aldaketa nabarmena 200-250 eta 250-300 milimetroko tarteen artean. Banaketako beste alboan ere gertatzen da antzeko aldaketarik, 450-500 eta 500-550 tarteen artean.

Maiztasunen aldaketak erakusteaz gain, histogramek jartzen dute agerian bi behatokitako aldakortasuna handia dela: ez dago balio multzo bat banaketaren erdialdean taldekatzen dena; balio gehienak, berriz, ibiltartean zehar sakabanatzen dira. Ezaugarri hori bereziki nabarmena da Toledoko histograman, non sakabanaketa ia ibiltarte osoan hedatzen baita. Nekez esan liteke, ikasten ari garen aldiari buruz, ohiko euri kopururik egon denik.

San Fernandoko behatokiko histogramak nahiko argi erantzuten dio lehen egiten genuen beste galdera bati, balio handien ugaritasunari buruzkoari: 800 mm baino handiagoak diren balioak gutxitan gertatu dira. Izan ere, histogramaren azken lau tarteek, 800 eta 1.000 milimetro artekoek, 8 balio besterik ez dute biltzen.

San Fernandoko banaketako asimetria ere nabaria da histograman. Urte euritsuenei dagozkien balioek banaketa eskuinaldera luzatzen dute. Banaketaren ezker aldean, balio txikien aldeko muturrean, ez dago luzamendurik.

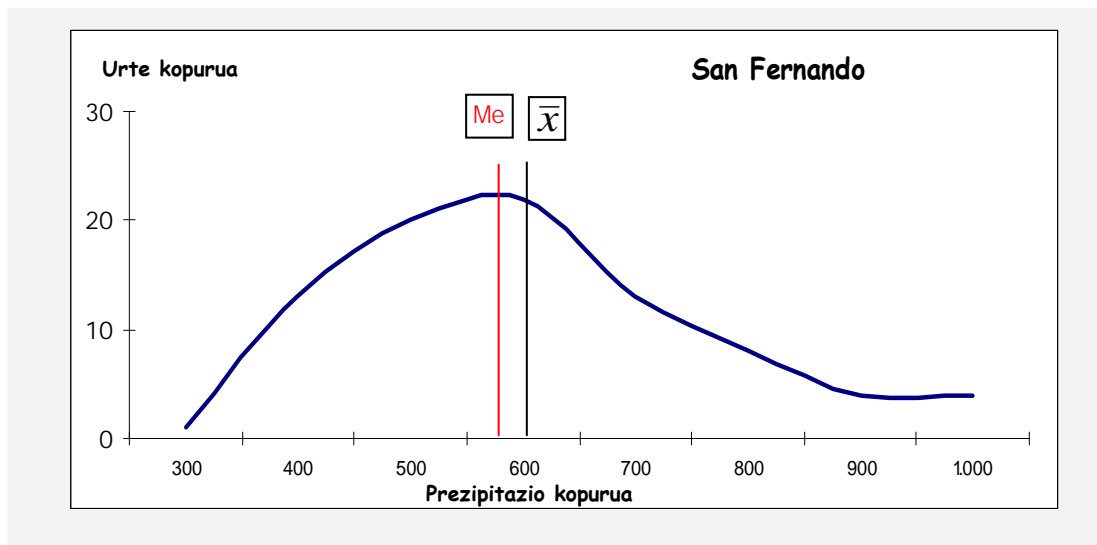
Dagoeneko esan daiteke histogramek informazio nahikoa eman digutela ikasten ari garen aldagaiei buruz, alegia, Toledoko eta San Fernandoko behatokitatan 1910etik 1995era jasotako prezipitazioen banaketari buruz. Ondoren, ikasi ditugun neurri estatistikoak erabiliko ditugu aldagaien aldakortasunari buruzko deskribapena egiteko.

	Toledo		San Fernando	
Batezbestekoa	367,1		558,0	
Mediana	359,8		522,1	
Desbideratze estandarra	83,2		160,0	
Batezbestekoa ± Desbideratzea tartea	283,9	450,3	397,93	717,98
Asimetria	0,12		0,81	
Aldakuntza-koefizientea	22,67		28,68	
Kurtosia	-0,27		0,19	
Balio handiena	573,3		999	
Balio txikiena	188,1		282,8	
Ibiltartea	385,2		716,2	

Dagoeneko bi behatokietako prezipitazioen aldakortasuna, aztertu dugun aldian, argia da. Orain kontua da nola egin deskribapen bat modu egokian oinarrituta, hau da, baieztapenak frogatzen dituzten neurrietan finkatuta.

Hezetasun desberdineko behatokiak

Toledoko eta San Fernandoko behatokietan, ikasitako aldian behintzat, jaso diren euri kopuruak nahiko desberdinak izan dira. San Fernandoko behatokia euritsuagoa izan da Toledokoa baino. Batezbestekoen emaitzen arabera 189,6 milimetroren aldea dago haien artean. Dena den, **medianen emaitzei** begiratzen badiegu, bi behatokiaren arteko diferentzia txikiagoa dela ikusten dugu, 162,3 milimetrokoa da eta. San Fernandoko batezbestekoaren emaitza altuagoa da, aldiko urte euritsuaren eraginez. Datu multzoaren balio handiek banaketa luzatzen dutenek; **bultzatu** dute batezbestekoa, eta haren emaitza igo.



Arretaz begiratu gero ohartuko gara batezbestekoaren desplazamenduz; batezbestekoa ez dago banaketaren erdian, eskuinalderantz mugituta baizik.

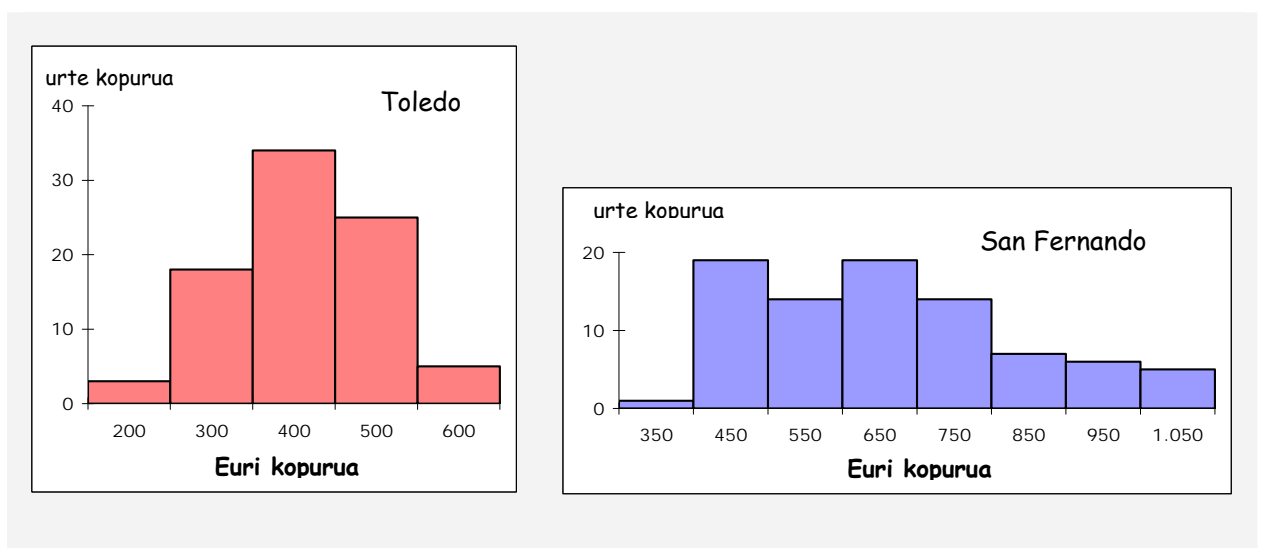
Aldakortasun handiko prezipitazio kopururak

Bi behatokietan urte batetik bestera gertatzen diren prezipitazio kopuruen aldaketak ikusgarriak dira. Ikasitako aldi osoari begiratzen badiogu prezipitazioen aldakortasuna batezbestekoaren balioa baino handiagoa da. Berez, oso deigarria suertatzen da San Fernandoko behatokiko aldakortasuna, 700 milimetro baino handiagoa dela eta. Toledokoa, berriz, 400 milimetrora ez da iristen.

Bi behatokiaren arteko desberdintasuna nabarmena da urte bereziki euritsuak dagokienez. San Fernandoko behatokian egon dira zenbait urte, prezipitazio-balio handiak eduki dituztenak, 1.000 milimetro ingurukoak. Toledoko behatokian, berriz, ez da jaso bere batezbestekotik horren urrun dagoen prezipitazio kopururik. Bion arteko desberdintasun hori deskribatzeko, agerian jartzeko, neurri estatistiko egokiena **asimetria-koefizientea** da: San Fernandoko koefizientearen emaitza (0,81) eskuinerako luzamenduaren seinalea da; Toledoko emaitza (0,12) banaketa nahiko simetrikoaren lekukoa da.

Ohiko aldakortasunari dagokienez –alegia, urte gehienetan gertatu den prezipitazioen arteko diferentziei dagokienez–, bi behatokietako portaera antzekoagoa izan da. Bi kasuetan esan daiteke ohikoa den prezipitazio

kopururik ez dagoela. Urtetik urtera prezipitazio kopuruak oso desberdinak izatea ez da noizbehinkako gauza, urterokoa baizik. Bi banaketen **kurtosi-koefizientearen emaitzek** argi asko erakusten dute esandakoa, alegia, aldiko urte guztien prezipitazio-balioak ia ibiltarte osoan sakabanatzen direla. Toledoko balioa (0,27) eta San Fernandokoa (0,19) banaketa platikurtikoen lekuko adierazgarriak dira. Histogramek ere mutur edo luzamendurik gabeko irudiak erakusten dituzte:

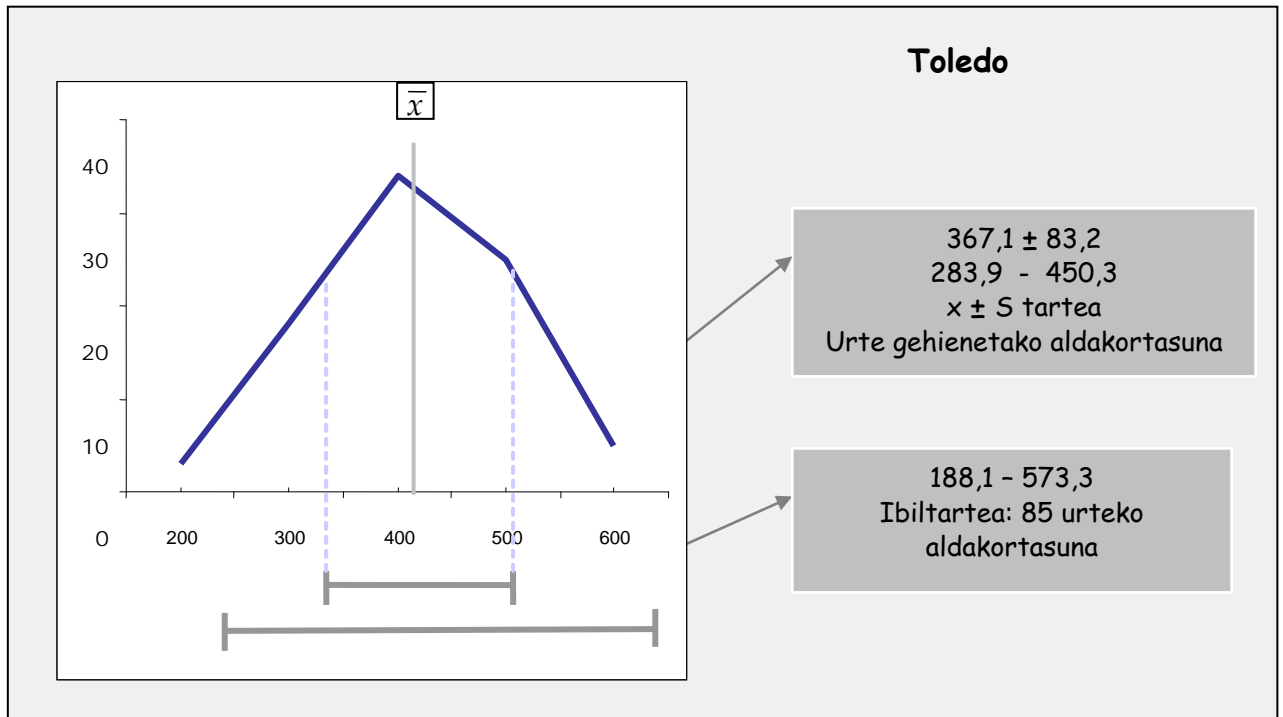


Histogramek nabarmentzen dute ohiko prezipitazio-tarteen gabezia, edo bestela esanda, prezipitazio-tarte gehienek maiztasuna oso antzekoa dela. San Fernandoko kasuan, esaterako, 350-450 milimetroko prezipitazioak eta 650-750 milimetrokoak urte kopuru antzekoetan gertatu dira.

Aztertutako bi aldagaien aldakortasunari buruzko deskribapen zehatzagoa egiteko aipatuko ditugu **batez bestekoa +/- desbideratze estandarreko tarteak**. Balio horiek lagunduko digute zehazten ohiko aldakortasunaren maila, hau da, zer prezipitazio-balioaren artean egon diren urte gehienetan kopuruak.

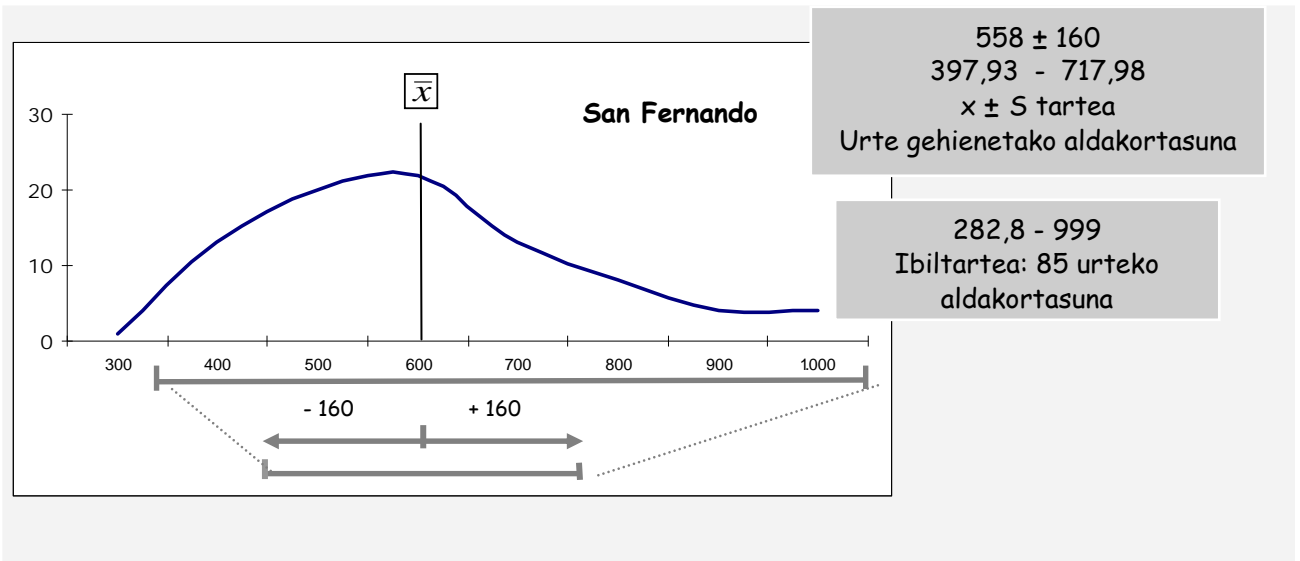
Toledoko datu multzoen desbideratze estandarren emaitza 83,2 milimetrokoa da; batezbestekoaren balioari gehituta eta kenduta, 283,9 eta 450,3 milimetroko tarteak sortzen dira. Aldakortasun-tarteak gordinik hartuta txikia dela ematen du.

Alabaina, behatoki honetan jasotzen diren prezipitazio kopuru txikiak kontuan hartuz, handia dela esan daiteke. Bestela, begira diezaiogun aldakuntza-koefizienteari (22,67).



San Fernandoko behatokiaren kasuan, maiztasun gehieneko tartea –hots, urte gehienetan jasotako euri kopuruak – 397,93-717,98 milimetrokoak izan dira.

2. atala. *Estatistikaren erabilera aldagaien aldakortasuna aztertzeko*



Eranskinak

1. taula. Pertsonako errenta erabilgarria. 2001 eta 2003

UDALERRIAK	2001	2003
Alegría-Dulantzi	11.775	13.225
Amurrio	10.377	11.067
Aramaio	11.876	12.653
Artziniega	11.013	10.453
Armiñón	10.574	10.947
Arrazua-Ubarrundia	13.100	15.695
Asparrena	10.052	10.442
Aiara	10.271	12.005
Mañueta	9.040	8.935
Barrundia	11.277	13.755
Berantevilla	9.202	10.330
Bernedo	10.268	11.310
Kanpezu	9.127	10.149
Zigoitia	14.294	15.001
Kripan	8.131	8.397
Kuartango	10.602	11.679
Burgelu	14.105	15.689
Elciego	9.522	11.344
Bilar	9.259	9.198
Iruraiz-Gauna	10.705	11.822
Bastida	9.060	10.140
Lagrán	9.292	10.182
Laguardia	9.607	10.604
Lantziego	8.185	8.771
Lapuebla de Labarca	7.800	7.845
Leza	9.145	10.803
Llodio	9.953	10.686
Arraia-Maeztu	9.807	10.883
Moreda de Álava	9.094	8.465
Navaridas	10.081	11.026
Okondo	9.268	10.227
Oion	9.573	9.533
Peñacerrada-Urizaharra	10.007	10.718

Ribera Alta	11.619	12.727
Erribera Beitia	10.986	12.702
Añana	9.198	10.134
Agurain	10.085	10.466
Samaniego	10.022	11.281
Donemiliaga	10.268	11.098
Urkabustaiz	10.502	12.154
Gaubea	10.127	11.083
Harana	8.218	8.713
Eskuernaga	8.916	9.636
Legutiano	13.069	12.710
Vitoria-Gasteiz	11.829	12.343
Yécora	8.138	8.633
Zalduondo	10.199	11.611
Zambrana	10.023	9.923
Zuia	13.043	15.029
Iruña Oka	10.702	12.144
Lantarón	8.828	10.224
Abaltzisketa	11.296	12.390
Aduna	9.281	11.798
Aizarnazabal	10.218	11.700
Albiztur	10.990	11.685
Alegia	10.013	11.034
Alkiza	9.982	10.968
Altzo	11.315	12.997
Amezketeta	11.225	12.208
Andoain	9.832	10.828
Anoeta	10.251	11.145
Antzuola	11.936	13.300
Arama	10.776	11.903
Aretxabaleta	13.044	13.510
Asteasu	10.603	12.547
Ataun	10.461	12.178
Aia	8.883	10.063
Azkoitia	10.905	11.802
Azpeitia	11.243	12.192

Beasain	11.659	12.539
Beizama	10.586	12.099
Belauntza	11.253	10.546
Berastegi	9.974	11.578
Berrobi	9.738	11.166
Bidegoian	9.719	10.441
Zegama	10.301	11.100
Zerain	12.553	13.179
Zestoa	10.039	11.069
Zizurkil	10.457	11.098
Deba	10.713	11.491
Eibar	11.023	11.639
Elduain	11.284	12.291
Elgoibar	11.058	11.677
Elgeta	10.682	11.508
Eskoriatza	12.095	12.493
Ezkie-Itsaso	11.009	12.016
Hondarribia	11.231	12.461
Gaintza	12.505	12.827
Gabiria	11.920	11.530
Getaria	10.217	10.881
Hernani	10.151	10.915
Hernalde	9.800	11.488
Ibarra	10.170	11.223
Idiazabal	11.810	12.814
Ikaztegieta	11.502	13.870
Irun	9.517	10.361
Irura	11.210	13.094
Itsasondo	10.451	11.061
Larraul	8.795	9.643
Lazkao	10.883	12.235
Leaburu	11.516	11.614
Legazpi	10.645	11.550
Legorreta	10.816	11.464
Lezo	9.656	10.278
Lizartza	10.393	11.433

Arrasate	11.758	12.108
Mutriku	9.490	10.528
Mutiloa	9.517	12.306
Olaberría	11.872	12.245
Oñati	12.755	13.321
Orexa	11.319	13.509
Orio	10.466	11.779
Ormaiztegi	11.564	13.106
Oiartzun	10.569	11.602
Pasaia	8.720	9.535
Soraluze	10.137	10.699
Errezil	9.034	9.531
Errenteria	9.163	9.916
Leintz Gatzaga	15.346	13.116
Donostia	12.399	12.955
Segura	11.281	11.780
Tolosa	11.667	12.434
Urnieta	10.058	10.858
Usurbil	10.723	11.690
Bergara	12.270	12.872
Villabona	10.161	11.261
Ordizia	11.547	12.055
Urretxu	10.799	11.688
Zaldibia	10.745	11.238
Zarautz	11.684	12.355
Zumarraga	10.086	11.112
Zumaia	10.885	11.703
Mendaro	10.902	11.894
Lasarte-Oria	10.215	10.998
Astigarraga	10.300	11.161
Baliarrain	8.930	10.136
Orendain	9.524	11.143
Alzaga	9.964	12.372
Gaztelu	9.530	11.609
Abadiño	10.503	12.287
Abanto-Zierbena	9.413	10.523
Amorebieta-Etxano	10.620	11.947

Amoroto	9.853	11.179
Arakaldo	10.894	11.893
Arantzazu	11.441	13.248
Munitibar-Arbatzegi-Gerrikaitz	9.976	10.986
Artzentales	8.996	10.058
Arrankudiaga	10.316	12.695
Arrieta	9.756	11.473
Arrigorriaga	10.011	11.509
Bakio	11.501	14.184
Barakaldo	9.199	10.014
Barrika	11.634	12.730
Basauri	9.443	10.023
Berango	10.791	13.171
Bermeo	9.035	9.882
Berriatua	9.779	11.053
Berriz	10.308	11.436
Bilbao	10.837	11.798
Busturia	9.975	10.505
Karrantza Harana	7.275	7.915
Artea	11.668	12.037
Zeanuri	11.656	13.258
Zeberio	10.709	11.925
Dima	11.287	12.625
Durango	11.505	12.106
Ea	10.133	12.171
Etxebarri	9.192	10.564
Etxebarria	10.501	11.877
Elantxobe	8.507	9.174
Elorrio	10.908	11.757
Ereño	9.854	12.275
Ermua	9.242	9.918
Fruiz	10.332	12.185
Galdakao	10.850	11.812
Galdames	9.731	11.408
Gamiz-Fika	11.346	13.454
Garai	11.554	13.183

Gatika	11.391	12.788
Gautegiz-Arteaga	11.465	13.409
Gordexola	14.335	11.645
Gorliz	12.189	13.486
Getxo	14.291	15.787
Güeñes	8.963	10.297
Gernika-Lumo	10.186	11.230
Gizaburuaga	10.244	12.080
Ibarrangelu	10.689	12.573
Ispaster	9.139	10.382
Izurtza	10.892	12.920
Lanestosa	7.192	6.757
Larrabetzu	11.513	12.987
Laukiz	20.627	15.860
Leioa	11.148	12.016
Lemoa	10.489	11.534
Lemoiz	10.768	11.873
Lekeitio	9.784	10.681
Mallabia	10.652	12.476
Mañaria	11.432	13.262
Markina-Xemein	10.095	11.120
Maruri-Jatabe	12.239	13.796
Mendata	10.157	11.198
Mendexa	10.176	11.701
Meñaka	10.907	12.062
Ugao-Miraballes	9.847	10.665
Morga	10.081	11.100
Muxika	10.324	11.347
Mundaka	10.178	11.254
Mungia	11.531	12.997
Aulesti	9.848	10.408
Muskiz	9.590	10.403
Otxandio	10.149	11.401
Ondarroa	8.893	9.467
Urduña-Orduña	9.563	10.269
Orozko	10.072	11.180

Sukarrieta	15.016	15.909
Plentzia	12.970	14.636
Portugalete	9.455	10.189
Errigoiti	10.300	11.811
Trapagaran	9.397	10.169
Lezama	13.157	14.483
Santurtzi	9.132	9.915
Ortuella	8.735	9.605
Sestao	8.710	9.286
Sopelana	11.560	12.951
Sopuerta	9.646	10.268
Turtzioz	7.813	8.050
Ubide	10.584	11.289
Urdúliz	10.054	11.246
Balmaseda	8.547	9.436
Atxondo	10.533	11.438
Bedia	10.294	10.838
Areatza	9.958	10.632
Igorre	10.470	11.274
Zaldibar	10.353	11.121
Zalla	9.681	10.470
Zaratamo	10.154	11.003
Derio	9.853	10.775
Erandio	8.492	9.307
Loiu	11.972	14.236
Sondika	10.856	12.314
Zamudio	10.504	11.466
Forua	9.563	10.221
Kortezubi	9.974	11.248
Murueta	11.857	12.976
Nabarniz	10.312	10.663
Iurreta	9.617	10.745
Ajangiz	10.762	11.260
Alonsotegi	8.694	10.172
Zierbena	10.362	10.664
Arratzu	11.246	11.403

Datu-iturria: EUSTAT

2. taula. Espainiako hainbat hiriren etxebizitzaren prezioa.
Etxebizitza berriei buruzko batez besteko datuak. 2007

UDALERRIA	Batez besteko prezioa, eurotan	UDALERRIA	Batez besteko prezioa, eurotan
Villanueva De La Serena	86.200	Fraga	127.000
Pozoblanco	88.000	Melilla	128.000
Monforte de Lemos	88.800	Barbastro	128.500
Don Benito	99.000	Langreo	128.700
Benavente	100.900	Lora del Rio	128.800
Molina de Aragon	101.100	Alcazar de San Juan	128.900
Caspe	103.900	Alcañiz	129.700
Baeza	104.200	Xativa	130.500
Loja	104.700	Valdepeñas	132.100
Barco de Valdeorras (O)	105.900	Mieres	133.600
Villena	106.200	Requena	133.700
Petrer	108.200	Vall d' Uixo (La)	134.400
Cocentaina	109.400	Mora d' Ebre	134.900
Cuellar	111.800	Santa Fe	135.200
Daimiel	113.600	Medina de Pomar	135.800
Mora	113.900	Huerca-Overa	135.900
Elda	114.400	Balaguer	135.900
Ubeda	114.800	Segorbe	136.100
Ubrique	115.500	Ecija	136.800
Jumilla	116.200	Ferrol	137.100
Alcoi	116.500	Solsona	138.500
Medina del Campo	116.600	Montilla	138.600
Ames	117.900	Torrijos	138.800
Astorga	118.200	Santa Cruz de La Palma	139.700
Puente Genil	119.400	Moron de La Frontera	140.200
Andujar	119.800	Yecla	140.600
Tomelloso	119.800	Reinosa	141.400
Cantalejo	120.200	Sueca	141.900
Najera	122.300	Linares	143.600
Ponferrada	122.400	Falset	144.300
Merida	123.600	Plasencia	144.900
Gandesa	125.100	Pontevedra	146.200

Valverde Del Camino	125.400	Tortosa	146.200
Aguilar De Campoo	125.700	Deltebre	146.600
Puertollano	126.800	Utrera	146.900
Borges Blanques (Les)	147.400	Cuenca	172.000
Ceuta	147.700	Sagunto/Sagunt	172.100
Badajoz	149.200	Antequera	172.200
Mazarron	149.600	Campello (El)	172.200
Lugo	150.300	Tarrega	172.500
Aranda De Duero	150.500	Torrent	172.700
Lucena	150.800	Sant Carles De La Rapita	173.100
Amposta	152.400	Linea De La Concepcion (La)	173.200
Alcanar	152.900	Roquetas De Mar	174.200
Jerez De La Frontera	153.500	Orihuela	174.500
Villarreal/Vila-Real	154.000	Teguise	175.100
Montblanc	154.200	Guardamar Del Segura	175.100
Santa Marta De Tormes	154.900	Velez-Malaga	175.800
Ejea De Los Caballeros	155.400	Aviles	175.900
Cervera	156.800	Santiago De Compostela	175.900
Lorca	157.200	Ayamonte	176.000
Mollerussa	158.200	Dos Hermanas	176.600
Calatayud	158.600	Chiclana De La Frontera	177.000
Pont De Suert (El)	158.700	Caceres	177.500
Sant Joan´ Dalacant	159.000	Burjassot	178.000
Oliva	159.000	Miranda De Ebro	178.200
Puerto Real	159.000	San Pedro Del Pinatar	179.200
Zamora	159.300	Alcala De Guadaira	179.400
Algeciras	161.400	Sort	179.500
Manacor	162.600	Cartagena	179.700
Viveiro	163.100	Valls	180.100
Catarroja	163.200	Quart De Poblet	180.200
Elche/Elx	163.600	Leon	180.200
Puerto Del Rosario	163.600	Santa Cruz De Tenerife	180.300
Ronda	163.600	San Javier	180.400
Ejido (El)	163.900	Bisbal D´ Emporda (La)	180.500
Cullera	164.000	Vila-Seca	181.600
Orotava (La)	164.400	Vila Joiosa (La)	182.300
Murtzia	165.600	Motril	182.600

Mairena del Aljarafe	165.900	Torre vieja	182.600
Arbucies	166.200	Inca	183.200
Vinaros	166.600	Santa Coloma De Farners	183.300
Palentzia	168.200	Ciudad Real	183.300
Sanlucar De Barrameda	169.200	Ripoll	183.800
San Fernando	169.300	Talavera De La Reina	184.600
Teruel	170.600	Tutera	185.100
Telde	170.800	Paterna	185.300
Ourense	170.800	Avila	185.600
San Sebastian de La Gomera	170.900	Arona	186.600
San Cristobal de La Laguna	171.000	Benicarlo	186.800
Jaen	171.200	Tias	186.900
Gandia	172.000	Huelva	187.100
Manlleu	187.500	Olot	211.300
Mogan	188.200	Alcudia	211.400
Puerto De Santa Maria (El)	189.600	Torrox	211.600
San Bartolome	189.800	Llança	213.100
Torrelavega	190.000	Salou	213.500
Palmas de Gran Canaria (Las)	190.600	San Lorenzo de El Escorial	214.800
Manises	190.600	Aranjuez	215.100
Lleida	191.100	Cambrils	215.500
Puerto de La Cruz	191.300	Estepona	216.000
Castello de La Plana	191.500	Sant Celoni	216.000
Ribadesella	191.600	Creixell	216.100
Oviedo	192.000	Segovia	216.500
Soria	193.600	Vic	217.300
Toledo	194.200	Punta Umbria	218.200
Almuñecar	194.300	Benicassim	218.600
Alacant	194.700	Calonge	218.700
Huesca	194.900	Begur	220.200
Vandellos i L´ Hospitalet de L´ infant	195.600	Granada	220.300
Orpesa	196.100	Gijon	220.600
Torroella de Montgri	196.700	Torremolinos	221.600
Vigo	197.000	Nerja	221.700
Jaca	197.600	Escala (L´)	221.800
Coruña (A)	199.400	Manresa	222.600
Seu D´ Urgell (La)	200.500	Figueres	223.500

Mislata	201.800	Parla	224.000
Logroño	205.100	Malaga	224.200
Albacete	205.300	Hernani	224.500
Banyoles	205.500	Escorial (EI)	224.500
Reus	205.500	Torredembarra	225.800
Kordoba	205.500	Malgrat de Mar	226.700
Rincon De La Victoria	205.800	Calvia	227.300
Pals	206.800	Palamos	227.800
Castello d'Empuries	207.300	Roses	228.500
Fuengirola	207.500	Valladolid	228.500
Vendrell (EI)	207.500	Torrejon de Ardoz	229.000
Peñíscola	208.100	Sant Feliu de Guixols	229.100
Almeria	208.400	Roda de Bara	229.500
Igualada	208.500	Eivissa	229.900
Ciudadella de Menorca	209.300	Tossa De Mar	230.200
Sant Antoni de Portmany	209.400	Castro-Urdiales	230.400
Salt	209.700	Pineda de Mar	231.000
Palafrugell	210.000	Collado Villalba	232.700
Basauri	210.000	Cunit	233.700
Santa Pola	210.100	Vilafranca del Penedes	234.200
Benlmadena	210.300	Castellar del Valles	234.800
Mont-Roig del Camp	210.400	Tarragona	235.300
Mahon	210.700	Valentzia	235.300
Cadiz	235.700	Cerdanyola del Valles	271.500
Cubelles	237.000	Girona	272.300
Lloret de Mar	238.000	Portugalete	273.300
Cardedeu	238.200	San Fernando de Henares	273.300
Laredo	239.200	Santurtzi	275.000
Barbera del Valles	240.200	Leioa	275.300
Canet de Mar	240.700	Sant Adria de Besos	277.900
Sant Pere de Ribes	240.800	Leganes	280.500
Altafulla	241.400	Boadilla del Monte	284.400
Algete	241.500	Vilassar de Mar	284.500
Burgos	241.600	Sant Feliu de Llobregat	284.700
Mostoles	241.700	Barakaldo	290.200
Sant Pol de Mar	243.400	Prat de Llobregat (EI)	291.400
Guadalajara	244.100	Tres Cantos	291.400
Palma de Mallorca	244.200	Sant Boi de Llobregat	293.600
Benidorm	244.700	Rozas de Madrid (Las)	294.000
Calafell	245.100	Mataro	295.200
Pinto	245.700	Puigcerda	295.300
Martorell	245.800	Arenys de Mar	296.200
Ripollet	246.100	Gasteiz	297.600
Vilanova I La Geltru	246.700	Terrassa	298.900

Galapagar	246.900	Viladecans	301.300
Coslada	247.300	Zaragoza	301.500
Salamanca	248.000	Sabadell	301.800
Caldes de Montbui	248.100	Gava	302.300
Alcala de Henares	248.600	Irun	303.500
Montcada I Reixac	248.800	Montgat	303.600
Santander	248.900	Sant Joan Despi	305.600
Sevilla	250.600	Cornella de Llobregat	306.200
Calella	251.500	Santa Coloma de Gramenet	306.500
Marbella	251.500	San Sebastian de Los Reyes	306.800
Iruñea	251.700	Masnou (El)	307.100
Erandio	253.300	Majadahonda	311.200
Fuenlabrada	254.100	Badalona	311.500
Sant Andreu de Llavaneres	254.600	Premia de Mar	314.400
Arrasate	255.300	Castelldefels	316.500
Blanes	255.900	Alcobendas	317.300
Colmenar Viejo	258.400	Bilbo	320.500
Villanueva de La Cañada	258.800	Getxo	322.000
Torreloanes	258.800	Pozuelo de Alarcon	327.900
Cadaques	259.400	Esplugues de Llobregat	330.900
Rubi	260.300	Sitges	336.700
Getafe	260.500	Hospitalet de Llobregat (L´)	337.800
Granollers	263.100	Sant Cugat del Valles	350.000
Castell-Platja d´ Aro	265.300	Sant Just Desvern	387.700
Alcorcon	265.400	Madril	397.800
Molins de Rei	266.300	Donostia	399.900
Mollet del Valles	267.100	Bartzelona	445.400

Etxebizitza berriaren batez besteko prezio nominala, eurotan, 2007ko ekaina.
Sociedad de Tasación, SA. <http://web.st-tasacion.es/html/index.php>

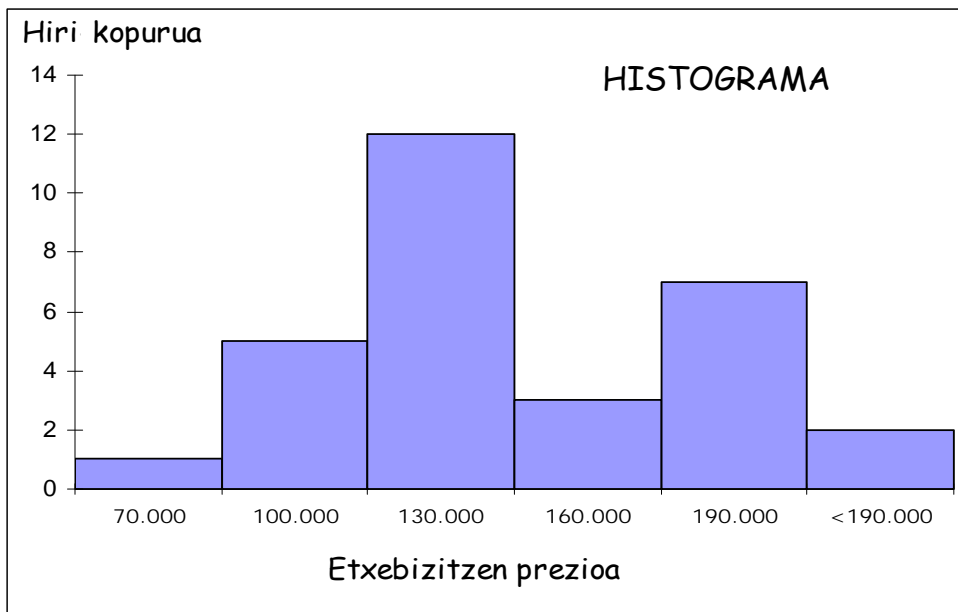
3. taula. Zaragoza probintziako hirietako etxebizitzaren prezioa. 2007

Batezbestekoa	132.865,1
Mediana	122.135
Desbideratze estandarra	37.885,91
$\bar{X} \pm S$	94.979,19 170.751,01
Asimetria	0,65
Kurtosia	0,77
Balio handiena	236.672
Balio txikiena	52.647
Ibiltartea	184.025

Balio handienek batezbestekoaren emaitza igorazi dute. Mediana eta batezbestekoa desberdinak.

Aldakortasuna handia da, hiri guztietakoa eta gehienetakoa: kurtosiaren emaitza baxua; batezbestekoa \pm desbideratzeko tartea

Prezio handienek luzatzen dute banaketa eskuinaldera: asimetria positiboa



	GUZTIRA	Etxebizitza berriak	Bigarren eskuko etxebizitzak	
		m ² /euro		m ²
La Jacetania	2.330	2.243	2.531	80
Alto Gállego	2.573	2.584	2.537	75
Sobrarbe	1.652	1.963	1.610	88
La Ribagorza	1.987	1.469	2.061	90
Cinco Villas	1.020	1.109	981	120
Hoya de Huesca	1.782	1.815	1.750	102
Somontano de Barbastro	1.290	1.458	997	119
Cinca Medio	1.231	1.531	947	99
La Litera	856	1.118	838	117
Los Monegros	702	1.337	576	143
Bajo Cinca	1.318	1.399	1.288	97
Tarazona y El Moncayo	1.175	1.198	1.172	96
Campo de Borja	826	1.022	566	118
Ribera Alta del Ebro	1.587	1.665	1.382	111
Valdejalón	1.478	1.728	1.068	113
Ribera Baja del Ebro	652	1.143	584	142
Bajo Aragón-Caspe	658	1.169	614	120
Comunidad de Calatayud	1.162	1.522	1.127	106
Campo de Cariñena	1.472	2.015	1.109	127
Campo de Belchite	661	1.085	564	144
Bajo Martín	483	444	485	109
Campo de Daroca	693	1.481	553	130
Jiloca	801	876	797	134
Andorra-Sierra de Arcos	784	1.033	705	135
Bajo Aragón	964	1.334	738	114
Comunidad de Teruel	1.472	1.511	1.460	126
Sierra de Albarracín	1.071	1.272	787	98
Gúdar-Javalambre	1.217	1.329	979	117
Matarraña	756	904	646	144
D.C. Zaragoza	2.752	2.696	2.811	86

Caja de Ahorros de la Inmaculada de Aragón: Mercado Inmobiliario de Aragón 2007.

4. taula. Toledoko eta San Fernandoko (Cádiz) urteko prezipitazioak. 1910-1995

Urtea	Prezipitazio kopurua (mm)	
	San Fernando	Toledo
1910	460,3	458
1911	568,7	343
1912	456,7	292,9
1913	568,3	359,8
1914	522,4	438,3
1915	712,3	327
1916	415	349,9
1917	505,1	246,5
1918	412,6	284,9
1919	764	357
1920	562,2	393,5
1921	373,3	320,3
1922	464,8	318,7
1923	500,9	277
1924	617,8	311,3
1925	474,4	393,7
1926	439,2	347,1
1927	562,1	401,1
1928	538,1	461,9
1929	370,6	385,3
1930	711,8	482,3
1931	483,9	297,9
1932	608,9	382,4
1933	786,5	334
1934	375,3	285,5
1935	313	411,6
1936	816,1	464,9
1937	608,1	557,6
1938	410,1	256,7

1939	697,6	438,1
1940	625,4	387,8
1941	717,3	494,1
1942	738,9	427,8
1943	387,3	408,7
1944	435,9	349
1945	474,5	276,6
1946	639,5	387,4
1947	667,4	522,4
1948	579,6	299,4
1949	524	318,1
1950	385,7	191
1951	580,5	347,1
1952	509,9	342,5
1953	493,9	293,8
1954	325,7	225,9
1955	999	573,3
1956	515,9	468,5
1957	520,5	385,5
1958	655,2	354,6
1959	481,4	416,1
1960	928,8	424,8
1961	893,3	422,7
1962	837,5	399,1
1963	969,3	471,4
1964	524,9	299,6
1965	622,5	320,1
1966	455,3	386,5
1967	474,9	341,6
1968	496,7	345,8
1969	933,5	537,6
1970	637,6	196,1
1971	556,2	418,1

1972	777,4	488,1
1973	414,6	302,9
1974	348,6	346
1975	604,5	444,7
1976	767,7	428,9
1977	643,3	504,2
1978	473,9	421,9
1979	516	406,7
1980	454,9	332
1981	344,2	343,1
1982	517,9	375,2
1983	522,1	284,4
1984	556	461,4
1985	512,4	283,1
1986	374,8	251,4
1987	605,8	365,1
1988	398,9	403,1
1989	857,7	381,3
1990	360,2	233,3
1991	553,1	243,5
1992	374,3	294,7
1993	477,3	406,3
1994	282,8	188,1

Instituto Nacional de Meteorología (1996): *Homogeneidad y variabilidad de los registros históricos de precipitación de España.*

5. taula. Pertsonako batez besteko gastu osoa. 2006

Autonomia-erkidegoak	Gastua (€)
Andaluzia	9.554,63
Aragoi	10.684,19
Asturias	10.662,51
Kanariak	9.257,72
Kantabria	9.668,21
Gaztela-Mantxa	8.945,31
Gaztela eta León	10.000,27
Katalunia	11.994,94
Ceuta eta Melilla	8.190,13
Valentziako Erkidegoa	10.720,13
Extremadura	8.565,99
Galizia	10.083,93
Balear Uharteak	11.208,57
Errioxa	9.320,97
Madril	12.111,83
Murtzia	9.391,17
Nafarroa	11.542,70
Euskadi	11.855,74

Datu-iturria: Familiako aurrekontuen inkesta. INE. 2006

Espainiako Estatistika Erakundeak bildutako datuak dira, 24.000 egoitza ingurutan egindako inkesten bidez sortutakoak. Jasotzen den gastuei buruzko informazioa 12 atal hauei buruzkoa da:

- Elikagaiak eta alkoholik gabeko edariak
- Edari alkoholikoak eta tabakoa
- Arropa eta zapatak
- Etxebizitza, ura, elektrizitatea eta erregaiak
- Altzariak, ekipamendua eta etxebizitzaren beste gastu batzuk
- Osasuna
- Garraioak
- Komunikazioak
- Aisialdia, ikuskizunak eta kultura

- Irakaskuntza
- Hotelak, kafetegiak eta jatetxeak
- Bestelako ondasunak eta zerbitzuak

6. taula. Pertsonako batez besteko gastu osoa

Banaketa simetrikoaren datu fikziozkoak

Autonomia-erkidegoak	Gastua (€)
Ceuta eta Melilla	9.838,09
Aragoi	10.463,89
Asturias	10.338,73
Balear Uharteak	10.714,21
Ceuta eta Melilla	8.190,13
Errioxa	9.587,77
Euskadi	11.131,65
Extremadura	8.790,99
Galizia	10.213,57
Gaztela eta León	10.088,41
Gaztela-Mantxa	9.170,31
Kanariak	9.410,72
Kantabria	9.963,25
Katalunia	11.510,97
Madril	12.111,83
Murtzia	9.712,93
Nafarroa	10.891,24
Valentziako Erkidegoa	10.589,05

7. taula. Pertsonako batez besteko gastu osoa
Banaketa asimetrikoaren datu fikziozkoak

Autonomia-erkidegoak	Gastua (€)
Andaluzia	9.115,13
Aragoi	9.740,13
Asturias	9.615,13
Balear Uharteak	10.290,13
Ceuta eta Melilla	8.400,13
Errioxa	8.865,13
Euskadi	13.590,13
Extremadura	8.490,13
Galizia	9.490,13
Gaztela eta León	9.365,13
Gaztela-Mantxa	8.615,13
Kanariak	8.740,13
Kantabria	9.240,13
Katalunia	13.715,13
Madril	13.836,83
Murtzia	8.990,13
Nafarroa	10.790,13
Valentziako Erkidegoa	10.040,13

8. taula. Nazioarteko analfabetismoari buruzko datuak. 2006

ESTATUAK	Biztanle kopurua	Alfabetatugabeen kopurua	Alfabetatugabeen ehunekoa
Aljeria	33.351.136	6.622.796,0	19,86
Bahrain	738.912	64.544,0	8,74
Egipto	74.166.499	16.451.048,8	22,18
Jordania	5.728.966	276.801,3	4,83
Kuwait	2.778.650	151.864,1	5,47
Libia	6.038.644	597.874,1	9,90
Mauritania	3.043.638	1.024.021,3	33,64
Maroko	30.852.969	11.529.038,5	37,37
Oman	2.546.323	287.801,9	11,30
Palestina	3.889.267	168.019,7	4,32
Qatar	821.311	68.990,9	8,40
Saudi Arabia	24.174.941	2.655.485,9	10,98
Siria	19.407.555	2.466.731,4	12,71
Tunisia	10.215.220	1.863.140,5	18,24
Arabiar emirerri batuak	4.248.477	364.316,9	8,58
Yemen	21.732.245	6.078.952,3	27,97
Albania	3.172.156	27.204,4	0,86
Belarus	9.742.123	28.198,7	0,29
Bulgaria	7.692.544	138.975,9	1,81
Kroazia	4.556.018	54.218,4	1,19
Estonia	1.339.974	2.822,1	0,21
Hungaria	10.058.461	111.677,0	1,11
Letonia	2.289.095	5.308,7	0,23
Lituania	3.408.147	10.427,8	0,31
Polonia	38.140.108	258.253,8	0,68
Moldavia	3.581.110	28.311,1	0,79
Errumania	21.531.731	523.652,3	2,43
Errusia	143.221.293	675.760,4	0,47
Eslovenia	2.000.833	6.005,1	0,30
Mazedonia	2.036.372	55.987,2	2,75
Turkia	73.921.766	6.792.377,4	9,19
Ukraina	46.557.424	144.500,9	0,31

Armenia	3.009.548	14.203,2	0,47
Azerbaijan	8.406.027	43.141,9	0,51
Kazakhstan	15.314.345	50.800,5	0,33
Kirgizistan	5.258.626	31.398,9	0,60
Mongolia	2.604.629	74.400,8	2,86
Tajikistan	6.639.839	17.744,5	0,27
Turkmenistan	4.899.455	19.821,8	0,40
Brunei Darussalam	381.951	15.054,0	3,94
Cambodia	14.196.612	2.683.738,6	18,90
Txina	1.320.864.228	74.934.670,2	5,67
Indonesia	228.864.479	15.278.471,5	6,68
Laos	5.759.404	1.195.572,6	20,76
Malaysia	26.113.733	1.616.855,1	6,19
Papua Guinea Berria	6.201.690	2.012.714,7	32,45
Filipinak	86.263.714	4.660.291,1	5,40
Samoa	185.363	1.691,0	0,91
Thailandia	63.443.952	3.209.491,4	5,06
Tonga	99.812	596,0	0,60
Vanuatu	220.775	34.576,0	15,66
Argentina	39.134.296	761.347,3	1,95
Aruba	104.936	1.671,4	1,59
Bolivia	9.353.848	627.060,1	6,70
Brasil	189.322.987	15.095.603,8	7,97
Txile	16.465.420	475.027,2	2,88
Kolonbia	45.558.453	2.642.747,8	5,80
Costa Rica	4.398.770	150.094,2	3,41
Kuba	11.266.706	19.500,8	0,17
Dominikar errepublika	9.614.667	794.514,7	8,26
Ekvador	13.201.994	767.910,0	5,82
El Salvador	6.762.317	793.033,2	11,73
Guatemala	13.028.572	2.443.013,8	18,75
Haiti	9.445.945	2.702.751,6	28,61
Honduras	6.968.686	879.406,1	12,62

Jamaika	2.698.606	299.120,7	11,08
Mexico	105.342.120	6.438.338,5	6,11
Holandarren Antillak	188.838	5.958,1	3,16
Nikaragua	5.532.366	832.492,6	15,05
Panama	3.287.538	177.134,2	5,39
Paraguai	6.015.701	297.511,3	4,95
Peru	27.588.575	2.247.197,6	8,15
Surinam	455.272	36.033,0	7,91
Trinidad eta Tobago	1.328.433	15.746,7	1,19
Uruguai	3.331.197	62.925,4	1,89
Zipre	778.684	16.825,0	2,16
Grezia	11.122.510	298.879,1	2,69
Italia	58.778.774	603.397,3	1,03
Malta	404.748	30.541,0	7,55
Portugal	10.578.658	490.398,8	4,64
Espainia	43.886.817	1.004.337,5	2,29
Bangladesh	155.990.775	57.566.712,8	36,90
Bhutan	648.766	238.047,1	36,69
India	1.151.751.462	311.702.537,1	27,06
Iran	70.270.179	8.556.000,4	12,18
Maldivak	300.293	7.545,0	2,51
Nepal	27.641.363	8.847.932,1	32,01
Pakistan	160.943.054	58.210.994,5	36,17
Sri Lanka	19.207.440	1.429.053,1	7,44
Benin	8.759.653	3.829.450,8	43,72
Botswana	1.858.162	241.327,7	12,99
Burkina Faso	14.358.500	7.652.201,5	53,29
Cape Verde	518.566	57.452,1	11,08
Comoros	818.434	142.192,1	17,37
Kongo	3.689.299	313.494,1	8,50
Gabon	1.310.818	132.465,3	10,11
Ghana	23.008.442	6.163.778,5	26,79
Guinea-Bissau	1.645.528	354.504,9	21,54

Liberia	3.578.922	1.076.884,0	30,09
Malawi	13.570.714	2.577.537,4	18,99
Mali	11.968.377	6.563.654,3	54,84
Maurizio	1.251.528	131.391,0	10,50
Mozambique	20.971.446	8.543.095,5	40,74
Namibia	2.046.553	190.442,4	9,31
Niger	13.736.722	6.507.134,0	47,37
Nigeria	144.719.951	27.621.319,1	19,09
Sao Tome and Principe	155.125	13.038,0	8,40
Senegal	12.072.475	5.291.093,1	43,83
Sierra Leone	5.742.694	2.587.945,4	45,07
Hegoafrika	48.282.459	4.579.512,2	9,48
Uganda	29.898.598	5.046.680,0	16,88
Tanzania	39.458.708	7.952.591,0	20,15
Zimbabwe	13.228.195	815.986,2	6,17

Datu-iturria: UNESCOren Estatistika Institutua: *Compendio mundial de la educación 2006. Comparación de las estadísticas de educación en el mundo. Kanada*