

Datuen bistaratzea estatistika ofizialetan

Enara Galbete Ahechu, Yosu Yurramendi Mendizabal¹

¹ Konputazio Zientziak eta Adimen Artifiziala Saila,
Euskal Herriko Unibertsitatea (UPV/EHU)

enara.galbete@gmail.com; yosu.yurramendi@ehu.es

Jasota: 16/05/2013

Onartua: 27/06/2013

Laburpena: Ordenagailuek zein informatikak azken urteetan jasandako aurrerape-
nek eragin zuzena izan dute datuen bistaratzearen esparruan. Datuen bistaratzea de-
lako analisi prozedurak funtsean ahalbidetzen du jasotako datuak era adierazgarrian
sintetizatu edo laburtzea; datuen analisi klasikoaren osagarria da, zenbakizko kalku-
luko prozeduretan oinarritua. Aipaturiko berrikuntzek, ahalbidetu dute bistaratze es-
tatiko edo klasikoez gainera, interaktibo deritzen beste bistaratze sofistikuago ba-
tzuk sortzen hastea.

Estatistika ofizialen erakundeak, eta bereziki Eustat Euskal Estatistika Erakundea,
hasiak dira informazioaren teknologia berri hauek eskaintzen dizkieten aukera eta be-
rrikuntzez baliatzen egoera berrira moldatu ahal izateko. Hori dela eta, artikulu hone-
tan azalduko dira zein izan diren Eustat Euskal Estatistika erakundeak 2 urtean zehar
datuen bistaratzearen gaiaren inguruan egindako ikerketa lanean lorturiko emaitza eta
ondorioak; horretarako, bereziki erreparatuko zaie erakundearen datu mota ezberdine-
tara egokituriko grafiko estatiko eta batez ere dinamikoei.

Abstract: The advances made in recent years in computers and computing have had a
direct impact on the data visualization area. Data visualization, essentially, is an analysis
procedure that allows synthesizing or summarizing significantly the collected data,
which is complementary to classical data analysis and which is based on numerical
calculus procedures. These innovations have permitted the beginning of the production
of more sophisticated visualizations, in addition to the static and classic ones.

Official statistical institutions, and especially Eustat, the Basque Statistics Institute,
have started to apply these new technologies, taking advantage of the opportunities and
innovations that they provide, so as to adapt to the new scene. In this sense, in the next
article the results and conclusions obtained in a two years research work in Eustat in
the field of data visualization are explained, giving extra importance to the static and
especially dynamic graphs adjusted to the different data-types of the organization.

SARRERA

Ordenagailuen teknologien zein interneten aurrerapenek eragin dute informazio estatistikoaren hedapen edo zabaltze metodoak erabat aldatzea. Aurrerapen hauei esker denbora laburrean aldatu gara era askotariko emaitza estatistikoak paperean zabaltzetik elektronikoki zabaltzera eta nabarmen aldatu eta erraztu dira informazio estatistikoaren taula zein grafiko bidezko aurkezpen moduak. Gainera estatistiken erabiltzaileen eskaera kopurua igo eta estatistikak modernizatzeak izugarri eragin diote ikerkuntza eta garapenaren arloari.

Teknologia eta teknika-estatistika berriek ahalbidetzen dute informazio estatistikoa ahalik eta gehien sintetizatzea eta erabiltzaileei era ulergarriago batean eskaintzea. Honela, informatika eta ordenagailuen garapenari esker, komunitate zientifikoak aurrerapen handiak egin ditu datuen sintesi zein bistaratze metodoen esparruan. Duela gutxi arte ezinezkoak ziruditen aurrerapen hauek erabiltzaileei aukera eman diete estatistika ofizialaren emaitzak era erraz eta intuitiboago batean ulertu eta erabiltzeko.

Datuen bistaratzea da ordenagailuen zein informatikaren aurrerapenek ahalbidetu duten datuen difusio-metodo berrietako bat. Datu estatistikoaren grafiko edo bistaratze delako analisi prozedurak funtsean, ahalbidetzen du jasotako datuak era adierazgarrian sintetizatu edo laburtzea. Metodoa, datuen analisi klasikoaren osagarria da eta zenbakizko kalkuluko prozeduretan oinarritzen da. Dagoeneko XVIII. mendean datu kuantitatiboen irudikapen grafikoak aurki bazitezkeen ere, era honetako irudikapenak ohikoak eta askotan beharrezkoak bihurtu dira, bistaratze-teknikak sortzea ahalbidetu baitute.

Ordenagailuei esker, bistaratze dinamikoak (animatuak eta interaktiboak) sor daitezke eta horrela, datu multzo handiak bistara daitezke.

Hori dela eta, estatistika ofizialen erakundeak informazioaren teknologia berriak aplikatu eta etengabe panorama berriari egokitu beharrean gertatzen dira, hauek eskaintzen dizkieten aukera eta berrikuntzez baliatuz.

Estatistika ofizialetako erakunde askok interneterako bidea hartu dute, berau informazioa kanpora komunikatzeko kanal edo hedabide nagusi delarik. Horrek guztiak, erakundeak behartu ditu metodologia zein estatistika-ekoizpenerako estrategiak berriaz planteatzera eta horietako askoren helburuetako bat da estatistikako informazioaren zabalkunde argi eta ulergarria egitea. Horretarako, ahalegin handiak egiten ari dira estatistikako datuen bistaratzearen ikerkuntzan, erabiltzaileek era dinamiko eta interaktiboan azter eta erabil ditzaten datuak.

Artikulu hau bi urtean zehar Eustat Euskal Estatistika Erakundearen datuen bistaratzearen inguruan egindako ikerketa lanean oinarritzen da eta lorturiko emaitza eta ondorioak estatistika ofizialetako erakunde askotan aplikatu daitezke.

DATUEN BISTARATZEARI EGIN AURRETIK KONTUAN HARTU BEHARREKOAK

Zenbait irizpide hartu behar dira kontuan datuak bistaratzeko unean, sorturiko bistaratzeak «onak» izatea eta berauen interpretaziotik ahalik eta informazio gehien ateratzea lortu ahal izateko. Hurrengo hauek dira W.S. Cleavelandek 1985eko *The Elements of Graphing Data* liburuan aipatutako irizpide eta gomendioak:

Ikuspen edo ikusker garbia. Komenigarria da beharrezkoa ez dena baztertzea, eta horretarako datuak erakusteko elementu grafiko ikusgarri eta nabariak erabiliko dira. Aldagai bakoitzeko eskala-zuzen pare bat erabiliko da eta datu eremua eskala-zuzen horiek osatzen duten angelu zuzenaren barrualdea izango da, hain zuzen. Eskaletako markak datu eremutik kanpo ezarriko dira, marka gehiegi ez erabiltzeko.

Garrantzitsua da datu eremua ez nahaspilatzeke oharrak, giltzak, seinaleak eta datuen etiketak ekiditea; horretarako, oharrak legendan edo testuan jarriko dira eta gainontzekoak datu eremutik kanpo.

Bestalde, gomendatzen da grafikoa zeharkatu behar duen balio bat dagoen aldiro datuekin nahastuko ez den erreferentzia-zuzen bat erabiltzea. Gainera, marraztutako ikurrak zein datu multzoak gainjartzen diren kasuan, hauek garbi bereizi behar dira elkarrengandik. Azkenik, grafikoak ez du bereizmenik galdu behar laburtu edo erreproduzitzen denean.

Ulertze garbia. Ondorio nagusiak era grafikoan jartzea eta informazio-legenda ulergarriak egitea gomendatzen da. Komenigarria da era berean datuen aldagarritasuna (desbideratze estandarra, errore estandarra, konfiantza tarte) adierazten duten errore barren esanahia ondo azaltzea. Kontuan hartu behar da aldagai baten logaritmoak marrazten direnean bat etorri behar dutela eskalen etiketek eta marken etiketek. Azkenik, ezinbestekoa da grafikoetako akatsak zuzendu eta garbitasuna zaintzea.

Eskalak. Komenigarria da marken tarte datu guztien edo ia guztien tarte estaltzeko moduan eta eskalak datu multzo osoa hartzeko moduan aukeratzea. Batzuetan lagungarri gertatzen da aldagai baterako bi eskala-zuzen erabiltzea, bi eskala ezberdin erakusteko. Komenigarria da grafikoak elkarren artean alderatu behar direnean eskala egokiak aukeratzea. Ez da beharrezkoa magnitude bat adierazten duen eskalan beti zeroa azaltzea.

Kontuan har bedi elkarrengandik bereizteko lagungarria izan daitekeela datuak eskala logaritmikoan erakustea.

Hain zuzen, eskala logaritmikoa erabiltzea gomendatzen da faktore biderkatzailearen portzentaje-aldakuntza ulertzea garrantzitsua denean edo balio handien alborapena zuzendu nahi denean. Azkenik, eskala bat behar-beharrezkoa denean bakarrik eten behar da; etendura saihestu ezin denean,

erabateko eskala-etendura erabiliko da eta etenduraren bi aldeetako zenbakizko balioak ez dira lotuko.

Estrategia orokorra. Irudikapen grafikoek ahalbidetzen dute eremu txiki batean informazio kuantitatibo asko sartzea. Grafikoak marraztea iterazio edo esperimentu prozesu bat izan daitekeenez, komenigarria da beharrezana dagoenean datuen grafikoa behin baino gehiagotan egitea. Kontuan izan bedi grafiko erabilgarri askok azterketa xehe eta arduratsua behar dutela.

DATU ESTATISTIKOEN SAILKAPENA

Datu estatistikoak izaera eta jatorri ezberdinetakoak izan daitezke, eta haien araberakoak dira aztertu eta aurkezteko erak, hala nola, grafikoak. Hain zuzen, zailtasun eta mota ezberdinetako grafikoak erabiliz datuak begi-bistaz analizatu, sintetizatu eta irudikatzea da datuen bistaratzearen helburu nagusia. Honela, datu mota ezberdinei bistaratze ezberdinak egokitzen zaizkie, eta horrela informazioa era grafiko eta nahiko intuitiboan aurkeztu eta aztertzeo aukera ematen zaie (datuen zein grafikoen zailtasunaren arabera). Grafikoa ona da sinplea, erabilerraza eta interpretatzeko erraza denean. Grafiko onek estatistika ofizialen interpretazio eta esplorazioan laguntzen dute, batez ere erabakiak hartzeko erabiltzen direnean.

Beraz, datu multzo bati bistaratze bat egokitu aurretik sailkatu egin behar da datu multzoa.

Behin Eustateko datu mota ezberdinak aztertuta, datuok bi irizpideren arabera sailkatu dira:

- *datuen izaeraren* arabera eta
- *jatorriaren* arabera.

Izaeraren araberako sailkapenean 5 talde bereiz daitezke:

- *Aldagai bakarreko datuak.*
- *Aldagai biko datuak.*
- *Aldagai anitzeko datuak:* hiru aldagai edo gehiago dituzten datuak dira.
- *Denborazko datuak:* denboraren une jakinetan neurtu, kronologikoki ordenatu eta elkarren artean uniformeki banatzen diren datu, behaketa edo balio sekuentziak dira.
- *Espaziozko datuak:* espazioko kokapen jakin batzuei lotutako aldagaiak dira.

Jatorriaren araberakoak, bi mota daude:

- *Maiztasun taulak:* datu bakoitzari dagokion maiztasuna esleitzen dioten taulak.

- *Mikrodatuak*: indibiduo bakoitzari aldagai zerrenda bat egokitzen dioten datuak.

Eustateko datuak izaera eta jatorriaren arabera sailkatu ondoren, datu mota bakoitzari grafiko mota bat egokitu zaio.

Grafiko hauek R programazio-lengoaia erabiliz sortu dira eta bi talde handitan sailka daitezke: grafiko estatikoak, pdf formatuan, eta grafiko dinamiko eta interaktiboak, interneteko web-orrialde batean txertatuz bistara daitezkeenak.

GRAFIKO ESTATIKOAK

Bistaratze estatikoak dira datuak aurkezteko erabilienak oraindik, eta mugimendua zein elkarrekintza duten beste bistaratze sofistikatuago bartzuen oinarria dira. Mota honetako grafiko asko sortu direnez eta erabiltzaile gehienentzat ezagunagoak suerta daitezkeenez, grafiko hauen artean interesgarrienak eta ezohikoenak azalduko dira.

Aldagai bakarreko datuei egokitutako grafikoen artean baditugu *barra diagramak*, *puntu diagramak* edo *dotplotak* (barra diagramen aldaera, aldagaiaren modalitateen aldakuntzak bistartzeko), *sektore diagramak*, *histogramak* eta *kutxa diagramak*.

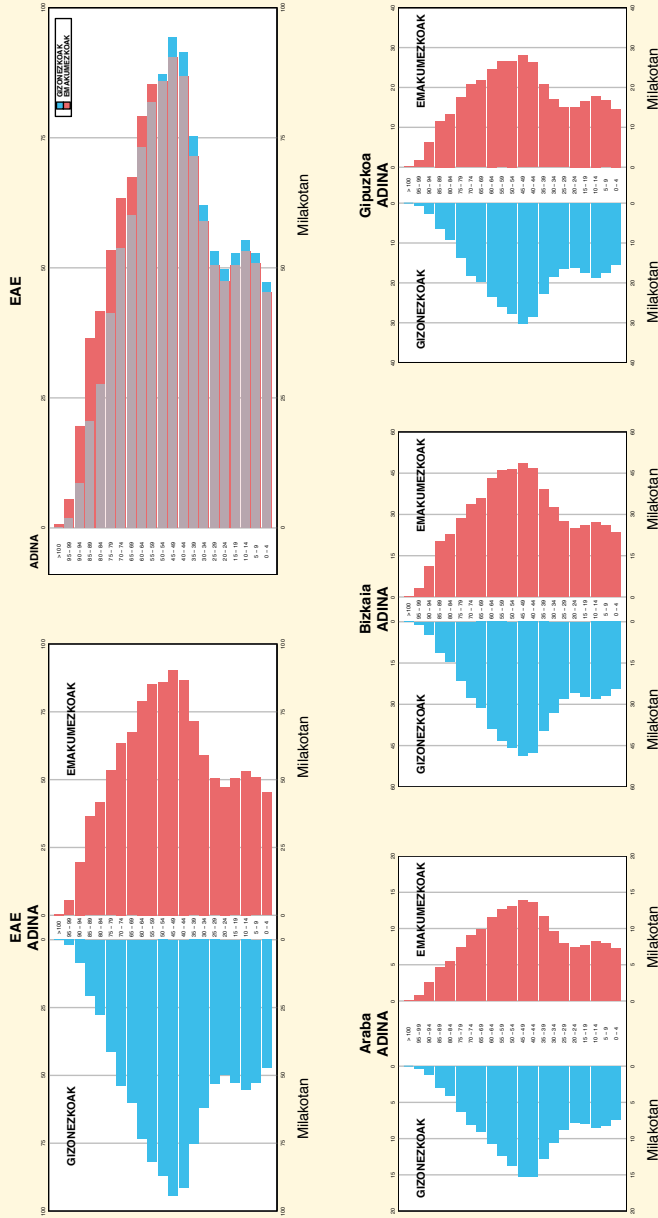
Aldagai biko datuei egokitutako grafikoen artean, badaude *ehunekoen* edo *maiztasun erlatiboen barra diagrama* eta *barra diagrama kategorizatua* edo *maiztasun absolutuen barra diagrama*. Grafiko hauek egoikiak dira aldagai kualitatibo bat beste baten arabera bistartzeko, baldin eta lehen aldagaiak 4 modalitate baditu eta horietako bat gainontzekoen bildura bada. Honen adibide da «jatorria» aldagaia, hurrengo modalitateak dituen: Araba, Bizkaia, Gipuzkoa eta Euskal AE (gainontzekoen bildura dena). 1. irudian erakusten da bi grafiko hauek eta barra diagrama bakunak azaltzen dituen irudikapen grafikoa.

Piramideak ere sailkatu dira aldagai biko datuei egokituriko grafikoen artean. Batetik baditugu piramide bakunak (betiko piramideak), eta bestetik piramide alternatiboak. Azken hauek, piramide bakunaren antzeko irudikapena dira, baina bata bestearen gainean irudikatzen dira ohiko piramidean erdiko ardatz bertikalaren bi alboetara doazen barrak (ikus bedi 2. irudia).

Mota honetako datuei egokitu zaien azken grafikoa zatikako erregresio lineal bat izan da, eta beraren bidez azter daiteke bi aldagai kuantitatiboren arteko erlazioa zatika lineala ote den.

Aldagai anitzeko datuei egokitutako grafikoen artean aipatzekoak dira *bi-piramidea*, eta *piramide kategorizatua*. Lehenengoan, bata bestearen

EAEko Biztanleria, 2020



2. irudia. Bistarate honen goi-ezker partean 2020 urteko Euskal AE-ko biztanleriaren proiektzioaren *piramide bakuna* erakusten da eta goi-eskuin partean, *piramide alternatiboa*. Behean, aldez, hiru lurralde historikoetako *piramide bakunak* irudikatzen dira.

gainean irudikatzen dira bi urte ezberdinei dagozkien piramide bakunak. Bigarrena, piramide bakun bat aldagai anitzeko kasura egokitzean lortzen den irudikapena da.

Aldagai anitzeko datuetarako beste zenbait grafiko sortu dira. Bi aldagai kuantitatiboren arteko erlazioa aztertzeke *mosaiko diagrama bika-tegorikoa* erabiltzen da. Kontingentzia-taula baten aldagai kategorikoen arteko erlazioa aztertzeke egokia da *korrespondentzia-analisiaren irudikapen grafikoa*, hain zuzen korrespondentzia-analisisa grafikoki irudikatzean datzana.

Bestalde, *Clevelanden dotplot grafikoa* puntu diagramaren aldagai anitzeko kasurako hedapena da (ikus 3. irudia).

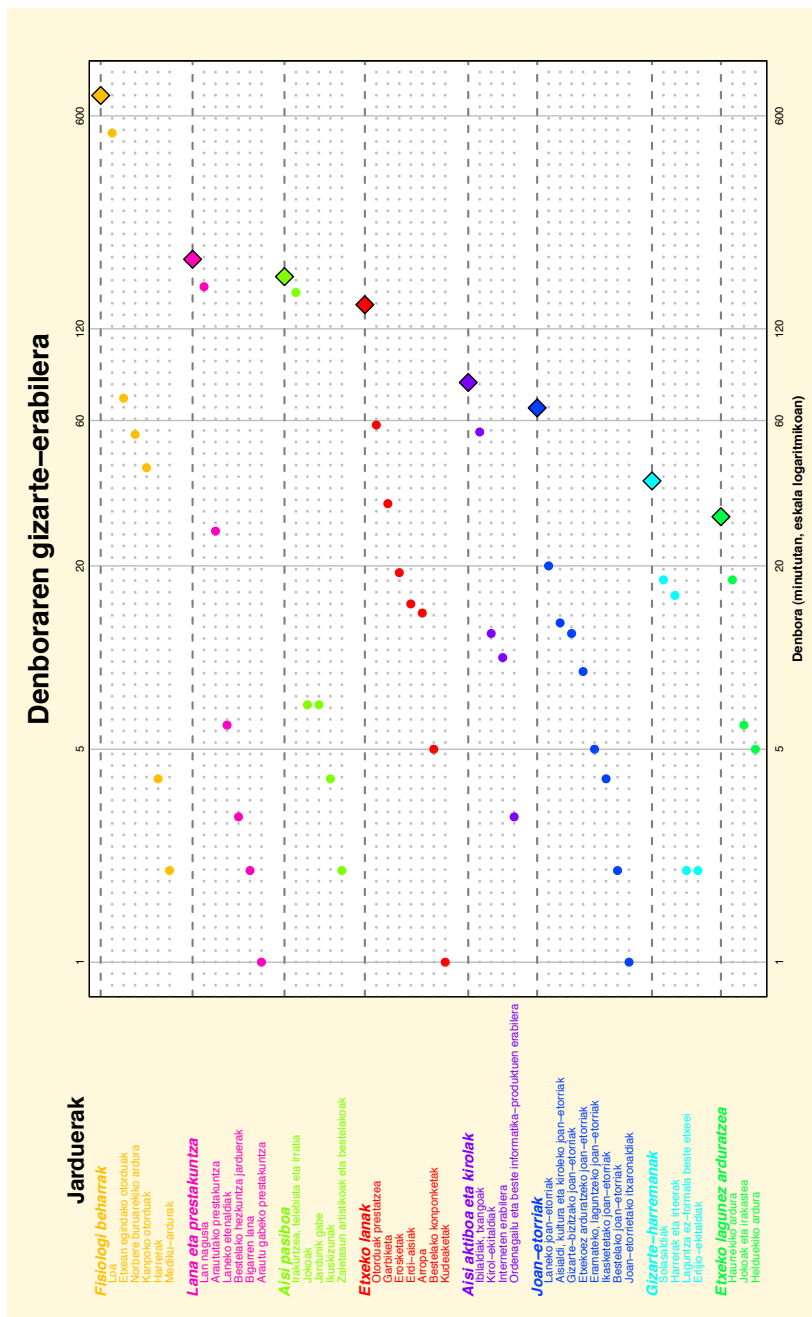
Zatikako erregresio lineala, sektoreka, egokia da bi aldagai kuantitatiboren arteko erlazioa beste aldagai kualitatibo baten modalitateekiko zatika lineala den aztertzeke (ikus 4. grafikoko gelaxka ez-diagonaletako grafikoa). Sorturiko beste grafiko bat *dendrograma* da; zuhaitz itxurako irudikapen grafiko honetan, zatitzen doazen azpikategoriatan antolatzen dira datuak, nahi den xehetasun mailara iritsi arte. Bestalde, *Scatter plot matrix grafikoa*, aldagai biko diagramaz osaturiko $n \times n$ ($3 \leq n \leq 5$ izanik) dimentsioko matrize batean datza (ikus 4. irudia). Azkenik, *koordinatu paraleloen grafikoak* beste guztiekiko paraleloa den eskala moduan irudikatzen du aldagai bakoitza eta behaketak aldiz, dagokien ardatzetan; horretarako, modalitate bakoitzeko kolore bat erabiltzen da eta ondoren kolore bereko puntu guztiak elkartzen dira.

Denborazko datuei honakoak egokitu zaizkie: *indizeen grafikoak*, indizeak eta indize desestazionalizatuak (datuei egutegi-efektua deritzona kentzean sortuak) irudikatu eta alderatzen dituztenak; eta denborazko serie asko aldi berean eta espazio mugatu batean bistartzeko era eskaintzen duten *horizon grafikoak* (ikus 5. irudia).

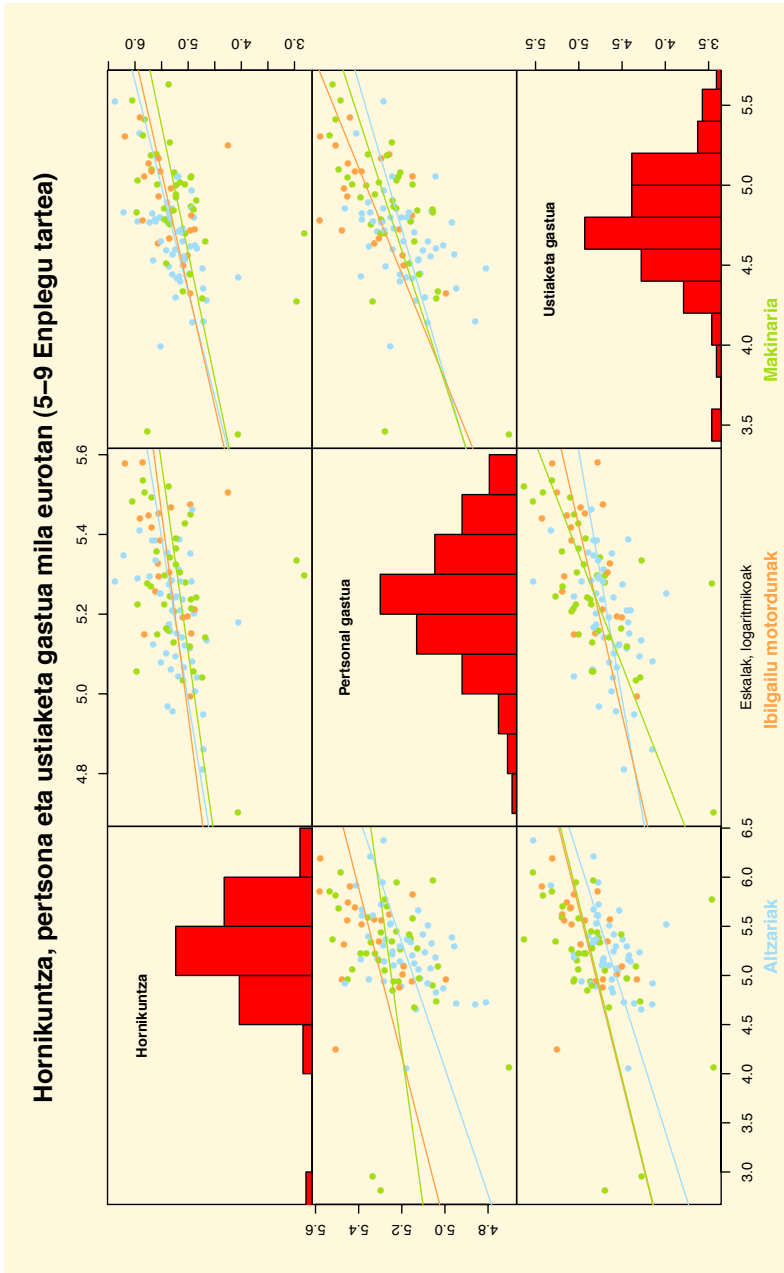
GRAFIKO INTERAKTIBOAK

Grafiko edo bistaratze bat interaktiboa dela esaten da erabiltzaileak aldatetaren bat sor dezakeenean bertan. Alegia, hala da erabiltzaileak eta bistaratzeak elkarri eragiten diotenean, grafikoan isla daitezkeen datuen inguruko galderak eginez.

Grafikoek interakzio gradu bat dutenean, batzuetan erabiltzaileak denbora tarte edota aztertu nahi dituen aldagaiak aukera ditzake; besteetan «zoom efektua» erabil dezake, eta beste batzuetan «data brushing» delakoa erabil dezake. «Data brushing» deritzo arratoia puntu baten gainean kokatuz informazio jakin bat agertzeari edo arratoia erabiliz irekitzen den leiho baten bidez datu-multzo bat hautatu edo tamaina edo kolorea aldatzeari.

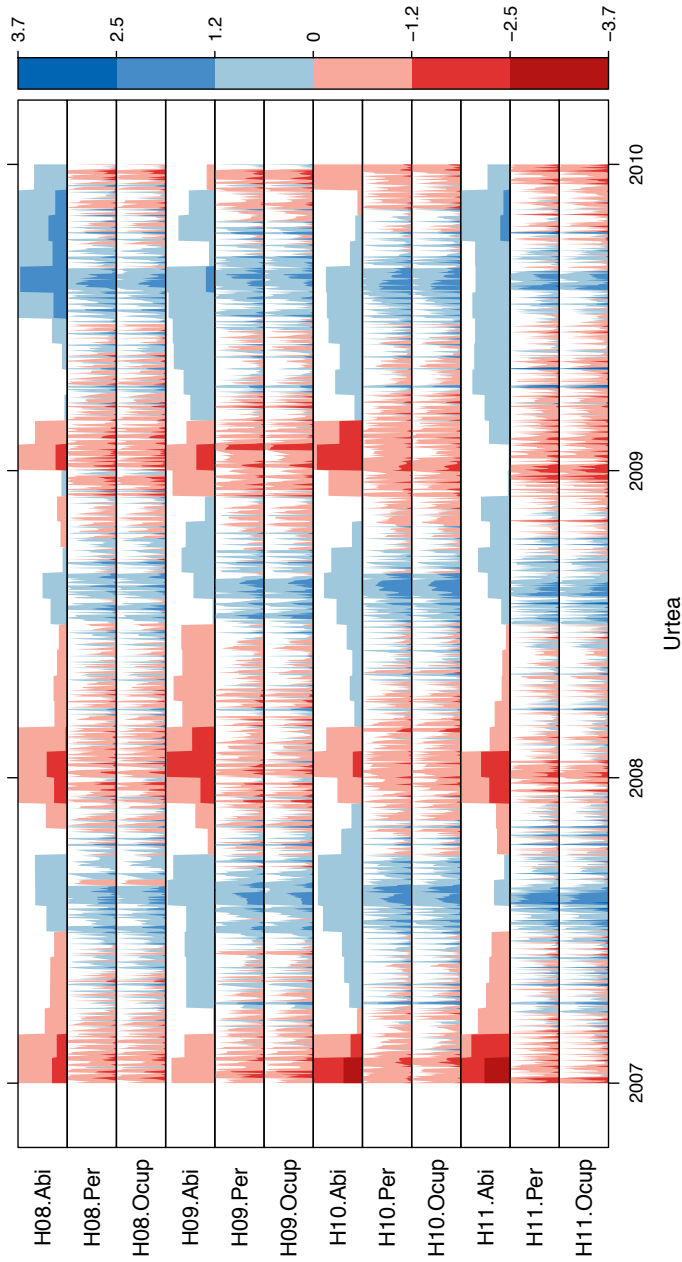


3. irudia. Bistaratzeko honetan Euskal AE-ko biztanleen denboraren gizarte-erabilerean banaketa aztertzen da, jarduerak eta azpigarriaren arabera.



4. irudia. Bistaratze honetan Euskal AE-ko hornikuntza, pertsona eta ustiaketa gastuak aztertzen dira, 5-9 enplegu tarteen arabera. Diagonaleko gelaxketan *histogramak* irudikatzen dira eta gainontzekoetan *zatikako erregresio linealak*, *sektoreka*.

Gipuzkoako hotelak



5. irudia. Bistaratzeta honek 2007. urtetik 2010era Gipuzkoako hotel mota ezberdinen plaza ireki kopurua (Abi), ostatu-gau totalen kopurua (Per) eta okupazioa (Ocup) irudikatzen ditu. Hotel motak hauek izan ziren: Donostiako goi-mailako hotelak (H08), Donostiako behe-mailako hotelak (H09), Gipuzkoar-kostaldeko hotelak (H10), Gipuzkoa barnealdeko hotelak (H11).

Zehazki ikerketa honetan sortu diren grafiko interaktiboak denborazko datuekin eta espaziozko datuekin egin dira.

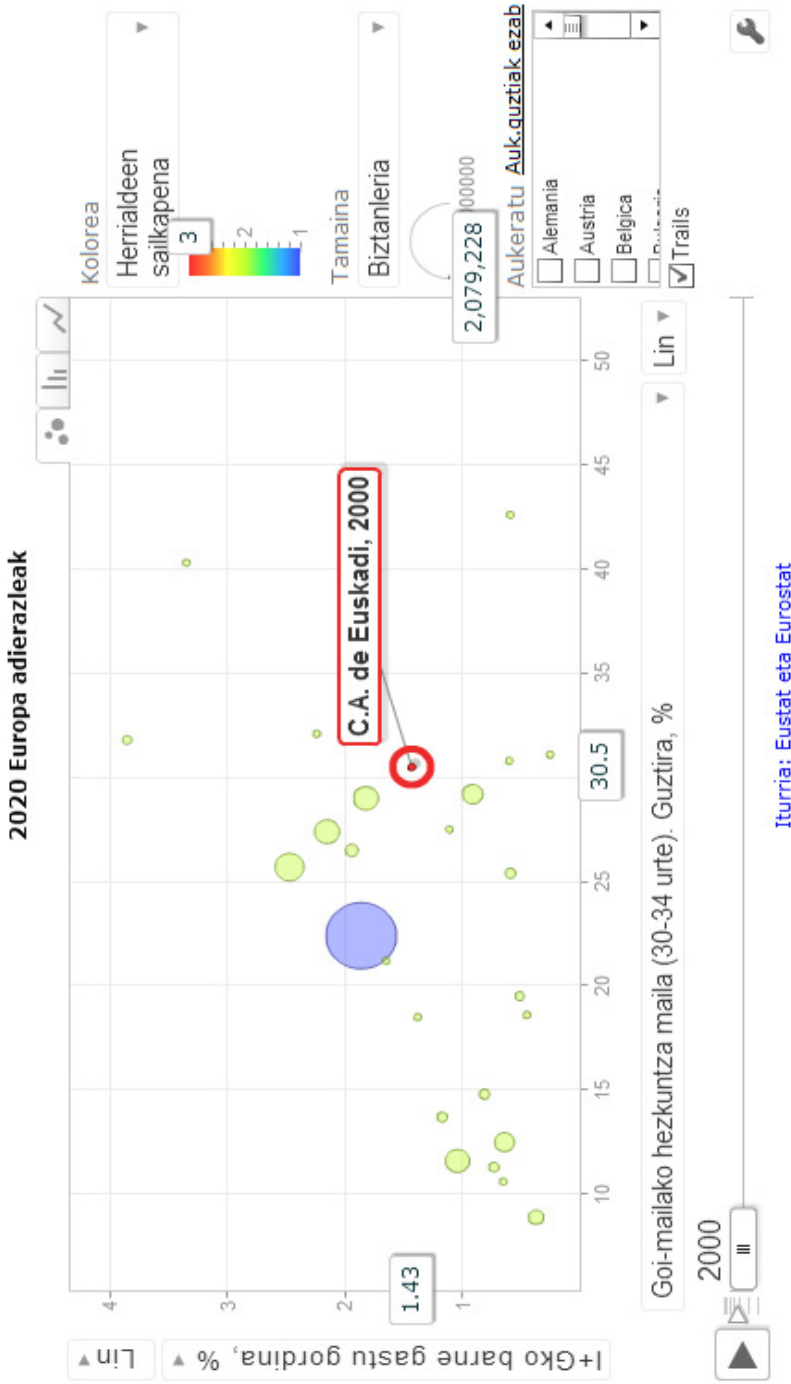
Denborazko datuekin bi motako grafiko interaktiboak sortu dira: *Motion Chart* deritzenak eta *Time Line* deiturikoak. Biak sortzeko «Google Visualization API» delakoa erabiltzen da. APIa edo Aplikazioak Programatzeko Interfazea (ingelesez, Application Programming Interface) funtzio eta prozedura (edo, objektuei zuzenduriko programazioan, metodo) multzoa da, beste software batek erabiltzeko liburutegi bat eskaintzen duena.

Google Visualization API delakoa bistaratze teknologia aurreratuaren erabilpena errazteko diseinatua dagoen JavaScript APIa da eta funtsezko elementu bitan oinarritzen da: erraztasunean eta nonahikotasunean. Erabiltzaileak APIko galeriako grafiko bat aukeratu eta berau sortzeko kodea programatu behar du, tresnaren baldintzetara egokituz. Kode horrek APIari deia egiten dion agindu bat izango du, eta berari esker kodea online exekutatatu eta aukeraturiko grafikoa bistaratuko da.

Motion Chart grafikoa aldagai anitzeko kasurako dispersio- edo sakabanatze-diagrama baten hedapena dela esan daiteke. Grafiko dinamiko da, hots, mugimendua duen bistaratze motakoa (ikus 6. irudia). Grafiko dinamiko gehienetan bezala, kasu honetan ere mugimendua zehazten duen faktorea denboran zehar sortzen den aldaketa da, baina ez da hau faktore bakarra. Kasu honetan, erabiltzaileak aukera ditzake x eta y ardatzetako aldagaiak eta burbuilen tamaina eta kolorea zehaztuko dituzten aldagaiak eta ikus dezake burbuilek denboran zehar duten bilakaera. Grafiko honek beste grafiko tradizional batzuetan ezin ikus daitezkeen joera eta ereduak identifikatzeko aukera ematen du, datuak denboran zehar aztertzea ahalbidetzen baitu. Horrez gain, interpretazioa errazten duten tresnak eta grafikoarekin berarekin jolasteko aukera eskaintzen ditu. Hain zuzen, hauek dira eskaintzen dituen tresnak: eskala lineal edo logaritmikoaren artean hautatzeko gaitasuna, burbuilek denboran zehar duten ibilbidea bistaratzeko aukera, erreproduktzio-abiadura kontrolatzeko gaitasuna, zona jakin batean zentratzeko aukera, zooma eta hautatu gabeko burbuilen gardentasuna doitzeko aukera.

Gainera, arratoia burbuila baten gainean jartzean, 4 koadrotxo agertzen dira; bertan, burbuilaren identifikatzailea (kasu honetan, herrialdearen izena), x eta y ardatzetako aldagaien balioak, eta burbuilaren tamaina zein kolorea finkatzen dituzten aldagaien balioak erakusten dira. Hauxe da «data brushing» delakoa.

Azalduriko aukera guztiez gain, posible da era berean datuak *Motion Chart* formatuan aurkeztu beharrean *barra diagrama dinamiko* edo *lerro diagrama estatiko* moduan aurkeztea.



6. irudia. Bistaratzeta interaktibo honek «2020 Europa adierazleak» egiturazko adierazleen *Motion Chart* grafikoa irudikatzen du. Eustateko web orrian argitaratua.

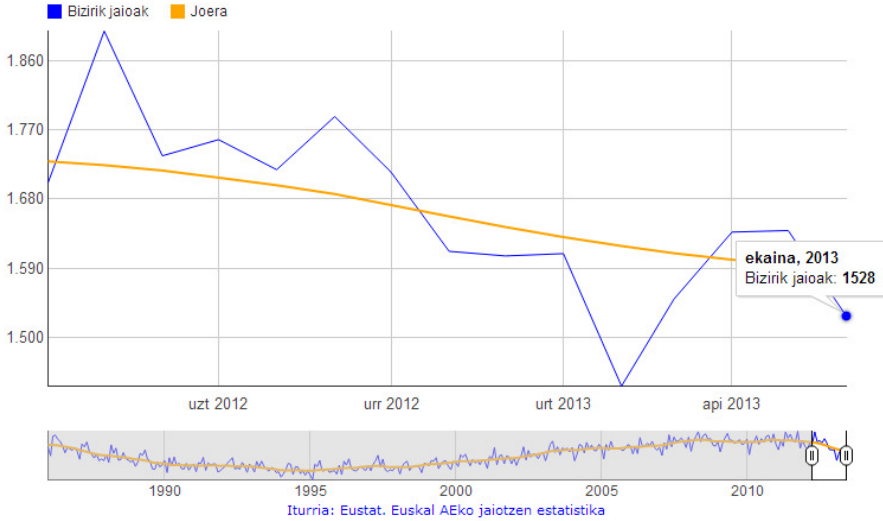
Time Line delako grafikoa, diagrama interaktibo bat da, aldi berean denborazko serie bat baino gehiago bistartzeko egokia. Lerro diagrama batez eta «kontrol» deritzon egitura batez osatuta dago. Lerro diagramaren ordez posible da halaber azalera diagrama edo sakabanatze-diagrama edo «Combo Chart» deritzona erabiltzea. Azken hau serie ezberdinak irudikatzeke lerroak, azalera, barrak, kandelak edo mailakatutako azalera konbina ditzakeen grafikoa da. Kontrol-egiturak erabiltzaileari aukera ematen dio nahi duen denbora tartean zentratzeko, grafikoen y ardatzeko eskala automatikoki aukeratutako tartera egokituz (ikus 7.1 eta 7.2 irudiak). Arra-toia grafikoen denborazko serieetako bat irudikatzen duen lerroa (edo azalerara, edo barrara edo kandelara edo mailakatutako azalerara) hurbiltzean, elkarriketa koadro bat azaltzen da, eta bertan erakusten dira serieak puntu horretan duen balioa eta dagokion data, data «brushing» efektua, alegia. Gainera, grafikoen gaineko legendako serieetako bat aukeratzean dagokion lerroa loditu egiten da.

Espaziozko datuak bistartzeko mapei elkartutako grafikoak erabiltzen dira. Estatistika ofizial asko leku geografikoekin lotuta daudenez, askotan informazioa erakusteaz gain mapa batean irudikatzen da. Egun informazio geografikoa maneiatzeko badira *GIS* deritzen sistema informatikoak, eta estatistikaren barruan bada *espaziozko estatistika* deituriko diziplina bat, informazio geografikoari lotutako informazio estatistikoa aztertzen duena. Duela gutxi arte, estatistika-erakundeetan informazio geografikoa eta datu estatistikoak bakoitza bere aldetik gordetzen bazen ere, espaziozko datuak bistartzeko (*datuen geo-bistaratzeta*) aplikazioen garapena hastapenetan dago. Datuak eta eremu geografikoak lotu ahal izateak datu estatistiko ofizialen erlazio eta egiturak miatu eta ulertzeko aukera garrantzitsua eskaintzen du eta analisi estatistikoan eragin handia du.

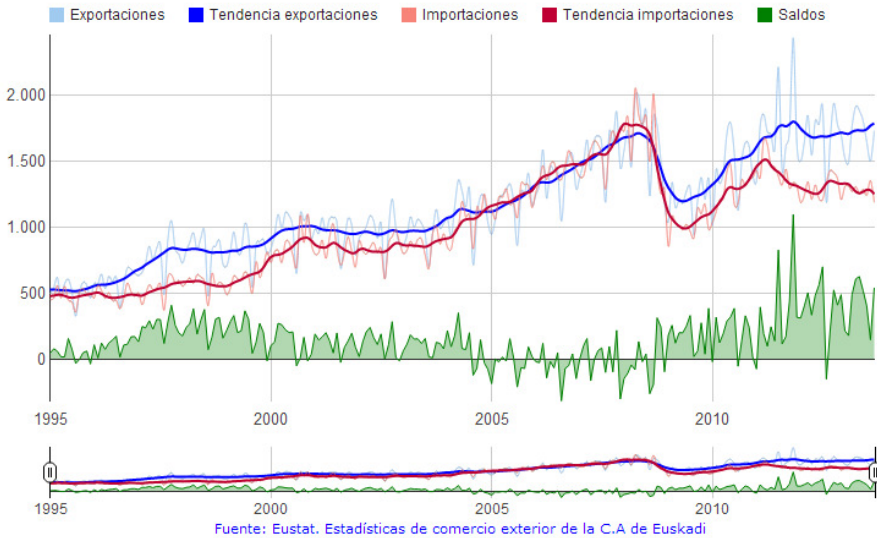
Egin den datu geografikoen bistaratzeta Euskal AE-ko zentsu sekzioen tipoei buruzkoa izan da. Horretarako, lehenik Euskal AE-ko zentsu sekzioen tipoen analisi faktorial anizkoitz bat eta sailkapen bat egin dira, zentsu sekzioak ezaugarri berdinen arabera biltzeko helburuarekin. Analisiaren ondorioz, Euskal AE-ko zentsu sekzioak 12 tipotan sailkatu dira, bakoitza bere ezaugarri nagusien arabera deskribatuz. Behin tipoak lortuta, emaitzak Google Earthen bidez bistaratu ahal izateko kmz motako fitxategi bat sortu da (ikus 8. irudia). Google Earth Informazio GIS baten antzekoa den programa informatikoa da; programa honek, ahalbidetzen du planetako irudiak bistaratzeta, sateliteek azken hiru urteetan ateratako argazkiak, mapak eta datu-base oso on bat konbinatuta.

Sortutako kmz fitxategia irekitzean Euskal AE-ko mapa koloreztatua agertzen da, zentsu sekzio bakoitza dagokion tipoaren arabera koloreztatuta egonik. Gainera, edozein zentsu sekzioen gainean sakatzean hari buruzko informazioa duen elkarriketa koadroa agertzen da; horri, «data brushing» efektua deritzo.

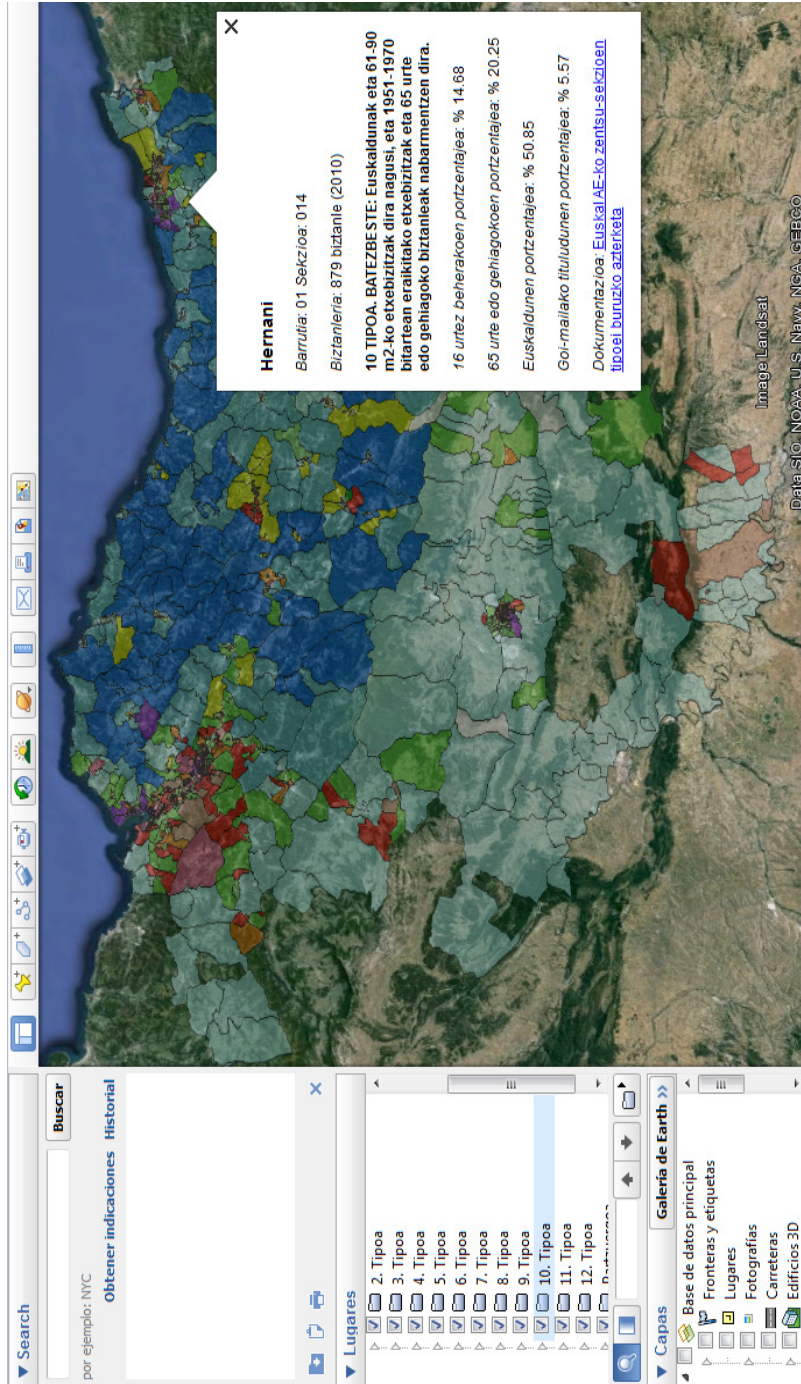
Euskal AEn erregistratuta dauden eta Euskal AEn bizi den ama batengandik bizirik jaiokoak. Hileroako datuak



Evolución del comercio exterior de la C.A. de Euskadi. (Millones de euros)



7.1 eta 7.2 irudiak. Bi bistaratze interaktibo hauetan «Euskal AEn erregistratuta dauden eta Euskal AEn bizi den ama batengandik bizirik jaiotakoak» eta «Euskal Aeko kanpo-merkataritzaren bilakaera milioi eurotan» ikus daitezke, hurrenez hurren. Eustateko web orrialdean argitaratuak.



8. irudia. Google Earth bidez sorturiko bistaratze interaktibo honetan «Euskal AE-ko zentsu sekzioen tipoen analisiaren» ondoren lorturiko emaitzak erakusten dira. Eustateko web orriari argitaratua.

Ezkerreko goitibehera-koadroari esker erabiltzaileak interesatzen zaizkion tipoak soilik aukera ditzake, eta bakarrik aukeraturikoak koloretatuko dira mapan. Horrez gain, tresnak eskaintzen dituen aukerez balia daiteke erabiltzailea informazio gehiago lortzeko, hala nola, zoomaz, bilatzaileaz, ...

ONDORIOAK

Artikuluari zehar esan den bezala datuen bistaratzeak baditu hainbat abantaila, hala nola taula eta fitxategietako datuak behatuz soilik ikus ezin daitezkeen ereduak, datuen egiturak eta joerak ikustea, datuak jasotzean sor daitezkeen ezuste eta gaizkiulertuak ekiditea edota datuei doitutako eredu konplexuak antzeman eta beren baliagarritasuna ebaluatzea. Hain zuzen, datuak bistaratu eta sintetizatzeak haien erabilpena indartu egiten du, datu estatistikoak ezagutza bihurtuz. Hala ere, bistaratzeak eskaintzen dituen abantailaz baliatzeko, beharrezkoa da metodologia jakin bat eta testu honetan aipaturiko oinarritzko printzipioak kontuan hartzea.

Hasieran esan den bezala, teknologiaren eta ordenagailuen aurrerape- nek errazago bilakatu dute orain dela gutxi arte ezinezko zena. Hala ere, estatistika ofizialak teknika hauek datuen zabalkunde zein analisi arloetan eskaintzen dituzten aukerez baliatzen hasi besterik ez dira egin eta lan handia dute egiteke.

BIBLIOGRAFIA

- [1] GALBETE E. 2013. *Datuen bistaratzea estatistika ofizialetan*. Euskal Estatistika Erakundea, Vitoria-Gasteiz.
- [2] CLEVELAND W.S. 1985. *The Elements of Graphing Data*. Monterey, CA: Wadsworth.
- [3] YOUNG F.W., VALERO-MORA P.M. eta FRIENDLY M. 2006. *Visual Statistics. Seeing Data with Dynamic Interactive Graphics*. Wiley, New Jersey.
- [4] GALBETE E.; ADIN A.; ARAMENDI J.; IZTUETA A. eta YURRAMENDI Y. 2012. *Análisis de tipologías de las secciones censales de la CAE. Visualización en Google Earth*. IX Congreso Vasco de Sociología y Ciencia Política.
- [5] ESCOFIER B. eta PAGÈS J. 1992. *Análisis factoriales simples y múltiples. Objetivos, métodos e interpretación*. Euskal Herriko Unibertsitatea, Argitalpen zerbitzua.
- [6] LÊ S., JOSSE J. eta HUSSON F. 2008. «*FactoMineR: An R Package for Multivariate Analysis*». *Journal of Statistical Software*, **Volume 25**, Issue 1.
- [7] EUSTAT .1996. *Análisis de tipologías de las secciones censales de la C.A. de Euskadi*. Eustat Instituto Vasco de Estadística, Vitoria-Gasteiz.

- [8] EUSTAT .2001. *Análisis de tipologías de las secciones censales de la C.A. de Euskadi*. Eustat Instituto Vasco de Estadística, Vitoria-Gasteiz.
- [9] **R-Project**: <http://www.r-project.org/>. Eguneratze-data: 2013ko maiatzaren 16a.
- [10] **Google Visualization API**: <https://developers.google.com/chart/>. Eguneratze-data: 2012 apirilaren 3a.
- [11] **API. Wikipedia**: http://es.wikipedia.org/wiki/Interfaz_de_programaci%C3%B3n_de_aplicaciones. Eguneratze-data: 2013ko martxoaren 23a.