# 1. GAIA

## 1.1 KOMUNIKAZIOARI BURUZKO IKERKETA EGITEA

Ikerketa akademikoa→ Jakintza sortzea helburu:

Gaiak:

* Gai zehatzen trataera komunikabideetan
* Agenda mediatikoa
* Saio, pelikula edo serieen publikoaren soslaia/profila
* Ikus-entzuneko produktuen harrera

Marketinerako ikerketa:

* Publizitate ereduak
* Salmenta estrategien baliagarritasuna eta arrakasta maila

Zehaztasuneko kazetaritza:

* Kazetariek, beren lanaren baitan egiten dituzten ikerketak
* Normalean luzeak izaten dira, denbora eta dedikazioa eskatzen dute
* Erreportaia sorta moduan publikatu ohi dira, edo aldizkarietako atal gisa, baita dokumental gisara ere.

## 1.2 BESTEEK EGINDAKO IKERKETA KOMUNIKATZEKO MODUAK

* Enpresek sortutako datuak
* Gobernu edo instituzioek sortutako datiak, kazetariak landuak
* Instituzioek edo ikerlariek sortutako eta landutako txostenak

# 2. GAIA

## 2.1 IKERKETA

* Enpirikoa: aztertzen den errealitatean oinarritu, ez ikertzen duenaren iritzi hutsetan
* Objektiboa: arauak eta prozedurak ditu.
* Egiaztagarria da eta errepikatzeko posiblea: beste ikerlari batek egindakoa berriz egin dezakegu.
* Ez perfektua: ikerketa orok bere mugak eta akatsak dauzka.
* Metakorra: aurretik egindakoetan oinarritzen da eta etorkizunean baliagarria izan
* Publikoa: bere emaitzak konpartitzeko aukera egon behar da, eta baita ikerketa nola egin den adierazi, gardentasunez.

## 2.2 IKERKETAREN BIDEA

Ezagutza teorikoa→ Galdera, helburu eta hipotesiak→ Datuen analisia→ Ondorioak

↑ ↓

← ← ← ← ← ← ← ← ← ← ← ← ← ← ← ← ← ← ← ← ← ← ← ←

Zer genekien? → Zer jakin nahi dugu? → Nora “begiratuko” dugu hori jakiteko→ Zer ikasi dugu?

## 2.3 FASEAK

1. **Gaia identifikatu:** Zer ikertuko dugu eta zertarako

Gaia ikertzeak zerbait berria al dakar? → Jakintza praktikoa, ekarpen teorikoa... Erabilgarritasun sozialik du? Norentzat? Askorentzat ala gutxi batzuentzat?

1. **Ikerketa galdera egin:** gure gaia galdera bezala adierazi. Adibidez: zeintzuk dira jendearen motibazioak gure fakultatean ikasteko?

Galdera zehatza izan behar da eta egingarria da: ikerketaren bidez erantzun dezakegu?

Gure neurrikoa: ba al dugu galdera erantzuteko denbora, jendea, dirua, ezagutza...?

1. **Berrikuspen bibliografikoa:** orain arte gaiari buruz idatzi/ikertu dena irakurri behar dugu!

Nun bilatu? Artikuluak, txostenak edo liburuak bilatuko ditugu liburutegietan

Ez dugu egin nahi jada egina dagoen ikerketarik!

Irakurri eta informatu ostean hipotesiak eta helburuak zehaztea ere errazagoa izango zaiguz

1. **Helburuak zehaztu:** Zer ikertuko dugun argi izateko, ikerketa bideratzeko

Helburu nagusia eta azpi helburuak izan ditzakegu. Zehatzak izan behar dira. Esaldi bakoitzean helburu bat. Gure ikerketa gidatuko dute.

1. **Hipotesiak sortu:** gure ikerketa galderari buruzko suposizio edo susmoak dira. Beti ez dira egiten.

Ez dira besterik gabe usteak, ezagutzen dugunean oinarritzen dira

Ikerketa amaitzean gai izan behar gara zuzen edo oker zeuden jakiteko.

Zehatzak izan behar dira.

1. **Ikertuko dugun gaia operazionalizatu:** darabiltzagun kontzeptu eta ideiak “neurgarri” bihurtu Ideia abstraktuetatik (adibidez, ongizatea) zehatzagoetara goaz (adibidez, maila ekonomikoa, osasuna...)

Nola neurtuko ditugun erabaki (eskala batean? Zenbakitan?), alegia zer galdetuko dugu?

1. **Erabiliko ditugun datuei buruz erabaki:**

Lehen mailako datuak: guk geuk sortzen ditugu

* Abantailak:
  + Gure ikerketa galdera eta helburuei ondo egokitu ditzakegu
  + Prozesuaren gaineko kontrola dugu
* Zailtasunak:
  + Denbora eta diru kostu handia
  + Datuak beste datu batzuekin konparatzeko zailtasunak

Bigarren mailakoak: jada existitzen diren datuak erabiliko ditugu

* Abantailak:
  + Denbora eta diru baliabide gutxiago behar ditugu
  + Lagin handiak erabili ditzakegu
  + Tendentziak aztertu eta konparazioak egin ditzakegu
* Zailtasunak:
  + Nahi ditugun datuak ez daude beti eskuragarri
  + Datuok ez dute gure helburuekin bat egiten
  + Datuak fidagarriak al dira?

1. **Teknika aukeratu:** nola lortuko ditugu datuak?

Behaketa edo kontaketa sistematikoa

Data-mining

Inkesta

* Asko erabiltzen da
* Galdetegi estandarizatua erabili behar da
* Informazioa biltzen laguntzen digu bai informazio “faktikoen” inguruan (adina, soldata, bizilekua...), baita iritzien inguruan ere
  + Abantailak:
    - Kostu baxuarekin info asko
    - Emaitza adierazgarria da, ondo eginez gero
  + Zailtasunak:
    - Lortzen dugun info ez da sakona
    - Benetan adierazgarriak diren datuak lortzeko oso koste alatua du
    - Akatsak ez dira erraz atzematen

1. **Behaketa unitateak mugatu eta aukeratu:** erabiliko dugun teknika nori aplikatuko diogu? Galdera klabea: populazio osoaren datuak biltzea posible al da?

* Bai→ galdetegia populazio osoari pasako diogu.
* Ez→ Populazioa “ordezkatu” dezakeen lagin bat sortu behar dugu

1. **Datuak bildu, txukundu eta aztertu**
2. **Ondorioak atera**
3. **Txostena sortu eta gure ikerketaren berri eman**

### 2.4. IKERKETAREN URRATSAK

1. Iker-arazoa aukeratzea→ zer ikertu nahi duen idenfikatu behar dugu

Informazio-iturriak:

* Zientzia alor horretarako aldizkari bereziak
* Gizarteko komunikabide arruntak
* Norberaren interesak eta eguneroko egoerak
* Adituen iritziak
* Datu-baseak eta beste ikerketak

1. Ikerketaren Corpus Teorikoa eraiki

* Gure ikerketaren gaiarekin lotzen diren liburu, artikulu era aldizkariak irakurriz
* Gure ikerketaren kontzeptuak argitzen ditugu

1. Ikerketaren helburuak zehaztu

* Helburu argiak izan behar ditugu
* Jakin nahi duguna adierazi behar dute: “ezagutu, aztertu, deskeibatu … “
* Helburu nagusia eta azpi helburu zehatzagoak
* Gure ikerketa gero eta konkretuago egiten laguntzen digute
* Ikergaia hautatzerakoan galdera hauek egin behar:
* Zabalegia al da ikertu nahi den arazoa?
* Ikergai hori errealki iker al daiteke?
* Ikergai hori interesgarri eta baliagarria ote da?
* Zein da ikerketaren tamaina?
* Hautatutako metodologiak egokiak ote dira?
* Emaitzak unibertsora orokor ote litezke?
* Nahiko baliabide ba ote dugu ikerketa egin ahal izateko?
* Ikerketaren lan-plangintza ongi prestatu ote da?

1. Hipotesiak sortu

* Ikerketa egin aurretik uste duguna adierazten dute
* Egia ala gezurra izan daitezke, hori da hain zuzen ere ikerketarekin jakin nahi duguna
* Hipotesi mota ezberdinak daude: harreman eza, norabide gabeko harremana, norabideduna edo kausaduna adierazten dutenak, adibidez
* Lanerako tresna dira, jaso behar ditugun datuak zeintzuk diren esaten digutelako.

# 3. DATU KUANTITATIBOEKIN LAN EGITEKO OINARRIAK

## 3.1. OINARRIZKO KONTZEPTUAK

Ikerketa gehienetan ez da posible populazio osoa aztertzea→ Horren ordez, populazioa “ordezkatzen” duen lagina sortzen dugu→ Lagin horren parte izango diren banakoak aukeratzeko era ugari daude

**Populazioa:** : ikerketaren xedea den elementuen multzo osoa

**Banakoa:** gure lagin barruan dago. Aldagai guztiek banako bakoitzean hartzen duten balioa neurtzen dugu.

**Lagina:** populazioa ordezkatzea helburu duen populazioaren zatia.

* **Ausazkoa / probabilistikoa:**Populazioko kide guztiek hautatuak izateko probabilitatea dute
  + **Ausazko sinplea:** populazio osoa biltzen duen zerrendatik ausaz aukeratzen dira subjektuak. Adibidez: EHUko ikasle eta langileen zerrendatik.
  + **Ausazko sistematikoa:** ausazkotasuna bermatzeko sistema bat ezarriz. Adibidez: EHUko ikasle eta langileen zerrendatik bat aukeratu eta hurrengo 50ak laginean sartu.
  + **Estratifikatu sinplea:** populazioan dauden ezaugarriak betetzen dituen ausazko lagina aukeratuz. Adibidez: EHUko ikasle eta langileen zerrendatik 50 ikasle, 50 irakasle, 50 garbitzaile eta 50 administrazioko langile.
  + **Estratifikatu proportzionala:** populazioan dauden ezaugarriak betetzen dituen ausazko lagina, baina era proportzionalean. Adibidez: EHUko ikasle eta langileen zerrendatik 50 ikasle, 10 irakasle, 5 garbitzaile eta 7 administrazioko langile.
  + **Konglomeratuetan oinarritua:** aztertu nahi dugun populazioaren ezaugarri eta proportzioak betetzen dituen “berezko” multzo bat.
* **Ez ausazkoa / ez-probabilistikoa:** Populazioko kide guztiek ez dute hautatuak izateko probabilitatea.
  + **Kuoten bidezkoa:** aldez aurretik kopuruak ezartzen dira baina ez dira populazio osoa biltzen duen zerrendatik hartzen. Adibidez: 50 irakasle, 50 ikasle, 50 garbitzaile, 50 administrazioko langile, gertuen ditugunak edo ezagutzen ditugunak.
  + **Kuoten bidezko proportzionala:** aldez aurretik kopuruak ezartzen dira proportzionalki baina ez dira populazio osoa biltzen duen zerrendatik hartzen. Adibidez: 50 irakasle, 10 ikasle, 5 garbitzaile, 7 administrazioko langile, gertuen ditugunak edo ezagutzen ditugunak.
  + **Intentzionalak**: beste ikerketa batzuetan emaitza adierazgarriak eman dituzten guneak edo bereziki interesatzen zaizkigunak. Adibidez: Arte Ederretako fakultatea.
  + **Elur** **bola**: lagineko banako bakoitzak beste batzuen kontaktua errazten digu. Adibidez: ikasle pisuetan bizi diren ikasleek pisuetan bizi diren beste ikasle batzuen kontaktua ematen digute, lagina osatu arte. Lagin mota honetan kuotak ere ezarri daitezke. Adibidez: guztira 20 bizkaitar, 20 arabar eta 20 nafar behar ditugu.

ALDAGAI MOTAK:

* **Aldagai** **kategorikoak** (**Kualitatiboak**): haien baloreak modalitateak edo kategoriak dira eta ez dira zenbaki bidez adierazten.
  + **Nominalak**: hauen baloreak ezin dira hierarkikoki ordenatu (adibidez: generoa, nazionalitatea, erlijioa, probintzia...)
  + **Ordinalak**: baloreak ordenatu daitezke (adibidez: ikasketa maila, adostasun maila, euskara maila)
* **Aldagai** **numerikoak** (**Kuantitatiboak**): baloreak zenbakiak dira
  + **Diskretuak**: zenbaki osoak bakarrik izan ditzakete (anai-arreba kopurua, irakurritako liburu kopurua, zenbatetan bozkatu duzun)
  + **Jarraiak**: zenbaki osoen arteko bitarteko baloreak har ditzakete (pisua, altuera, adina)

DATU MATRIZEA:

Datu matrizeak lortu ditugun datu guztiak biltzen ditu

* Lerroetan gure lagineko banakoak daude
* Zutabeetan aldagaiak daude
* Laginaren tamainaren lerro kopurua eta aldagai beste zutabe izango ditu.

## 3.2. GRAFIKOAK

Adierazpen grafikoek:

* Informazioa era azkar eta argian eskaintzen dute
* Irakurterrazak izan behar dute eta egiazko adierazpenak eskaini.
* Aldagaiei buruzko informazio fidagarria eskaini behar dute.
* Ezin dute emaitzarik desitxuratu (eskalak aldatuz adibidez)

Grafiko guztiek:

* Izenburua dute beti
* Iturria adierazten dute beti
* Ardatzetan aldagaien inguruko informazioa eraman ohi dute.

Barra-diagrama bertikalak

* Aldagai kualitatibo/kategoriko guztiekin zein aldagai kuantitatibo/numeriko diskretuekin.
* Ardatz horizontalean aldagaiak eta ardatz bertikalean, balioak

Barra-diagrama horizontalak

* Aldagai kualitatibo/kategoriko guztiekin zein aldagai kuantitatibo/numeriko diskretuekin.
* Ardatz bertikalean aldagaiak eta ardatz horizontalean, balioak.
* Aldagaien kategorien izenak oso luzeak direnean erabiltzen da.

Barra diagrama pilatu edo metatuak

* Aldagai kualitatibo/kategoriko guztiekin zein aldagai kuantitatibo/numeriko diskretuekin.
* Bertikalak zein horizontalak izan daitezke
* Zutabe bakarrean aldagai baten totala eta haren kategoriak erakusten dituzte

Sektore diagramak

* %100 batzen duten aldagaiekin bakarrik erabiltzen da.
* Aldagai kualitatibo/kategoriko guztiekin edo kuantitatibo diskretuekin.
* Aldagaiak kategoria asko baditu, ez da oso argia izaten.

Lerro-diagramak (diagrama linealak)

* Denboraren aldaera adierazteko erabilgarriak
* Kategorien artean jarraikortasuna egon behar da
* Lerro diagramek lerro bat baino gehiago izan dezakete, informazio gehiago erakutsi ahal izateko

Kartogramak:

* Mapetan oinarritzen dira
* Lekuen arteko ezberdintasunak adierazteko balio dute

# 4. GAIA: ALDAGAI BAKARREKO ESTATISTIKA DESKRIBATZAILEA

## 4.1. MAIZTASUN TAULAK

Maiztasun taulak, grafikoak bezala:

* Datuak modu ordenatuan eta laburbilduan jasotzeko erabiltzen ditugu
* Emaitzak aurkezteko erabilgarriak dira baina baita hasierako analisiak egiteko.
* Adi! Irakurterrazak eta argiak izan behar dira

**ni**: maiztasun absolutua **n**: laginaren tamaina

**Ni**: maiztasun metatua **N**: populazioaren tamaina

**fi:** maiztasun erlatiboa

**Fi**: maiztasun eraltibo metatua

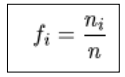
**pi**: maiztasun erlatiboa ehunekotan (portzentaia)

**Pi**: maiztasun erlatibo metatua ehunekoetan (portzentaia)

**Maiztasun absolutua (ni):**

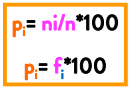
Aztergai dugun taldean, balio (xi) edo kategoria bat zenbat aldiz errepikatzen den adierazi

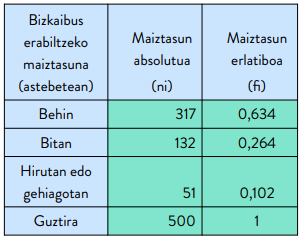
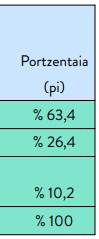
**Maiztasun erlatiboa (fi):**

Aztergai dugun taldean, balio (xi) edo kategoria bat **batetik** zenbat aldiz errepikatzen den adierazten digu.

**Maiztasun absolutua (ni):**

aztergai dugun taldean, balio (xi) edo kategoria bat zenbat aldiz errepikatzen den adierazten digu.

**Portzentaia (pi):** aztergai dugun taldean, balio (xi) edo kategoria bat ehunetik zenbat aldiz errepikatzen den adierazten digu.



Behin fi: 317/500= 0,634

Bitan fi: 132/500= 0,264

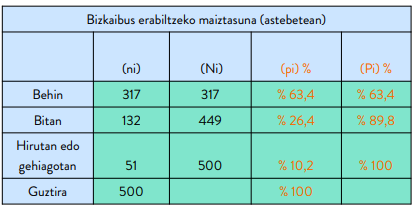
Hirutan edo gehiagotan fi: 51/500=0,102

Behin pi: 312/500 \* 100 = % 63,4

Bitan pi: 132/500 \* 100 = % 26,4

**Maiztasun METATUAK**

Maiztasun metatua kalkulatzeko ezinbesteko baldintza da balioak ordenatu ahal izatea hierarkikoki (txikitik handira) maiztasun → metatua aukeratzen dugun balioa eta txikiagoak diren balioen batura da.



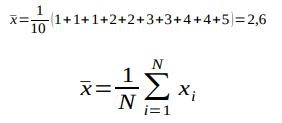
## 4.2. ZENTRALIZAZIO NEURRIAK

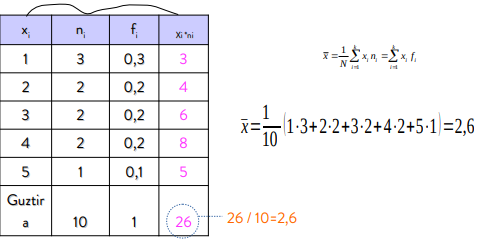
***BATEZBESTEKOA***

Aldagai numeriko diskretu eta jarraietan datuen banaketa ordezkatzen duen balio zentrala da. Aldagaiaren balio guztiak batu eta balioen kopuruaz zatitzean lortuko dugu. Eguneroko bizitzan sarritan erabiltzen dugun ideia da.

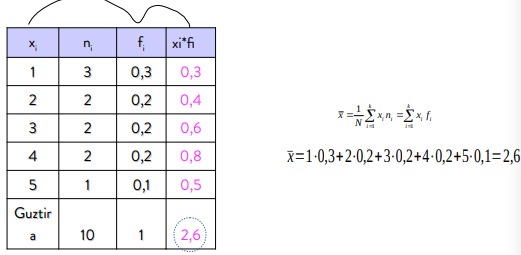
**1. Taldekatu gabeko datuekin:**

Estatistikako kontrol batean 10 pertsonen notak: 1, 1, 1, 2, 2, 3, 3, 4, 4, 5.



**2. Maiztasun taulekin:**

(ni erabiliz)



(fi erabiliz)

**Soilik aldagai KUANTITATIBOEKIN erabil dezakegu**

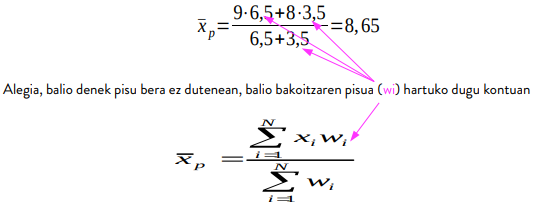
* Datuen gutxienezko eta gehiegizko balioen artean kokatu behar da, baina ez du zertan balio enpiriko bat izan (ez da gure zenbaki zerrendan egon behar).
  + Adinak: 20, 23, 27, 30 Batezbestekoa = 25 (inork 25 urte izan ez arren)
* Aldagai diskretuak ditugunean: batezbestekoaren balioa borobildu daiteke (gora edo behera)
  + Seme alaba kopurua 2,8 bada, batezbestekoa 3 dela esan dezakegu
* Aldagaia jarraiekin: hamarrekoak/dezimalak erabili (normalean bi)
  + Azterketako batezbesteko nota 6,75 izan da
* Batezbestekoak datuak ondo ordezkatuko ditu banaketaren sakabanaketaren arabera:
  + Banaketa homogeneoa batezbestekoa neurgailu ona da → 1, 1, 1, 2, 2, 3, 3, 4, 4, 5 batezbestekoa = 2,6
  + Banaketa heterogeneoa, edo sakabanaketa handiarekin batezbestekoak ez da hain neurgailu ona→1, 1, 1, 2, 2, 3, 3, 4, 4, 527 Batezbestekoa = 50,27

**3. Lagin ezberdina duten taldeen batezbestekoak kalkulatzea:**

50 ikasle daude eta azterketan 6 da atera duten batezbesteko nota. Euskaraz ikasten duen taldean aldiz, 20 ikasle daude eta 5 da azterketan atera duten batezbesteko nota. Orokorrean? 

**4. Batezbesteko ponderatuak kalkulatzen:**

Azterketa teorikoak 6,5 puntu balio ditu eta praktikek 3,5, azterketan 9 eta praktiketan 8 atera badut, zein da lortu dudan batezbesteko emaitza?



***MEDIANA***

* Datuak txikienetik handienera ordenatzean, mediana erdian dagoen balioa da
* Kasu honetan posizioa da garrantzitsua, alegia balioa non dagoen kokatuta.
* Datuen multzoa tamaina berdineko bi zatitan banatzen duen balioa da.
  + Balioen %50-a medianaren berdina edo txikiagoa da eta beste %50-a berdina edo handiagoa

**Nola kalkulatu?**

* **Datu kopurua txikia / bakoitia**→ datuak ordenatu eta erdiko balioa hartu →

Gure koadrilakoen adinak 20, 22, 22, 24, 27 badira, mediana erdian dagoen banakoaren adina da 22 urte. →Koadrilaren % 50ak 22 urte edo gutxiago ditu eta beste %50ak 22 urte edo gehiago.

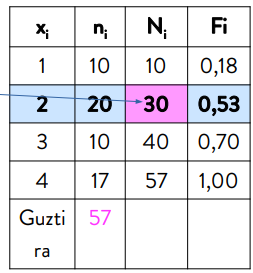
* **Datu kopurua bikoitia**→ erdian bi balio izaten dira, eta mediana bi balio horien batezbestekoa da→

Alboko klaseko ikasleen adina 32, 33, 25, 46, 23, 25, 39, 42. Mediana kalkulatzeko datuak ordenatzen dira: 23, 25, 25, 32, 33, 39, 42, 46. Erdian 32 eta 33 gelditzen dira, hortaz mediana: 32,5 da. Beraz, klasearen % 50ak 32,5 urte edo gutxiago ditu eta beste %50ak 32,5 urte edo gehiago.

* **Maiztasun taulekin, bakoitiak direnean→**

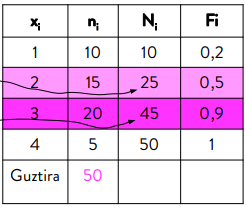
57 kasurekin, medianaren balioa, 29. posizioan dagoena izango da 57/2=28,5 baita.

Alegia, mediana 29. “postuan” kokatzen da, eta bere balioa 2 da. Datu hori maiztasun metatuen zutabean ikusiko dugu.



* **Maiztasun taulekin, bikoitiak direnean→**

50 kasurekin medianaren balioa 25 eta 26. posizioan dagoena izango da, 50/2 = 25 delako. Kasu honetan, 25. posizioan “2” balorea dugu, 26. posizioan aldiz “3” balorea dugu. Bien arteko erdiko puntua izango da gure mediana (2+3)/2, alegia, 2,5.



* Aldagai kuantitatiboekin soilik
* Medianaren balioak ez du zertan zenbaki osoa izan behar
* Medianan eragin gutxiago dute muturreko balioek batezbestekoan baino
* Datu sakabanatuak edo polarizatuak dituzten banaketetan, egokiagoa da mediana erabiltzea batezbestekoa baino batezbestekoak banaketaren deskribapen erratua eman diezaguke →
  + 1, 1, 1, 2, 2, 3, 3, 4, 4, 5 Batezbestekoa = 2,6 eta Mediana = 2,5
  + 1, 1, 1, 2, 2, 3, 3, 4, 4, 527 Batezbestekoa = 50,27 eta Mediana = 2,5

***MODA***

Datu multzo batean maiztasun handiena duen balioa da

* Klase honetako ikasleen adinak 20, 22, 22, 24, 27 badira, maiztasun handiena duen balioa 22 da.

Maiztasun tauletan, moda kalkulatzeko nahikoa da maiztasun handiena zein baliori dagokion ikustea

Modarik ez duten datu taldeak aurki ditzakegu ez dago errepikatzen den daturik →

1, 2, 4, 7, 10

Moda bakarra duten datu taldeak aurki ditzakegu datu batek besteak baino maiztasun handiagoa du

1, 1, 2, 3, 3, 3, 4, 7, 10 Mo =3 →banaketa unimodala

Moda bat baino gehiago duten datu taldeak ere aurki ditzakegu → banaketa bimodala, trimodala edo multimodala

1, 1, 2, 2, 3, 3, 4, 4, 7, 10 Mo = 1, 2, 3 eta 4 → banaketa multimodala

Aldagaiak hartzen duen balio bat da beti

Gogoratu: batez bestekoak eta medianak ez du zertan aldagaiak hartutako balio bat izan behar

Aldagai kualitatibo zein kuantitatiboetan kalkula daiteke aldagai nominalekin kalkula dezakegun zentroko neurri bakarra da

Batezbesteko edo medianaren kalkuluan datu guztiak hartzen dira kontutan, modaren kasuan aldiz, ez → gehien errepikatzen dena soilik.

Moda balio zentral gisa definitzen bada ere, ez du zertan banaketaren zetroan egon behar, muturretatik hurbil egon daiteke → egokia da banaketa asimetrikoetarako

1, 2, 3, 5, 7, 7, 10, 10, 10,10

## 4.3. POSIZIO NEURRIAK

Posizio neurriek, datuak txikienetik handienera ordenatzean, datu konkretu batek hartzen duen posizioa adierazten digute.

Medianaz gain (erdiko posizioa), estatistikan baditugu posizioa gehiago zehazteko beste neurri batzuk ere: kuartilak, dezilak eta pertzentilak.

Askotan balore absolutuek ez digute behar dugun informazioa ematen aztertzen ari garen banakoaz.

Lagin bateko banako baten balioa datuen multzoan kokatzeko posizio neurriak erabiltzen dira. Nola?

1. Balio guztiak ordenean jarri, txikitik handira
2. Interesatzen zaigun banakoa seinalatu talde horretan
3. Laginaren tamaina adierazi gure banakoaren kokapena ondo ulertzeko

***KUARTILAK***→ lau zati berdinetan zatitzen du→ Q da bere sinboloa

Q1→ lehenengo kuartilla, datuen laurdena ezker aldean uzten du

Q2→ bigarren kuartilla, medianarekin bat dator

Q3→ hirugarren kuartilla, datuen ¾ ak ezker aldean uzten ditu

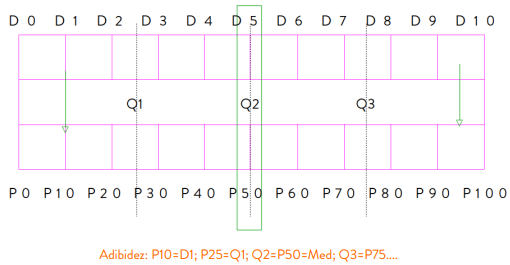
Adibidez: Q1=5 →Klaseko %25ak 5 edo gutxiago atera du azterketan, beraz gutxienez klaseko %75ak gainditu du

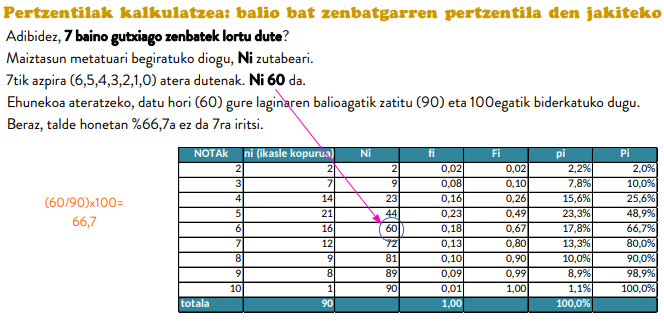
***DEZILAK→*** Banaketa hamar zati berdinetan zatitzen du D1, D2, D3 …. D9

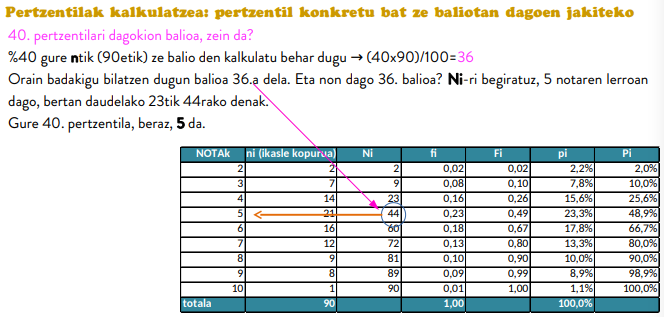
Adibidez: D1=0→ Gutxienez laginaren %10ak ez du neba-arrebarik

***PERTZENTILAK→*** Banaketa ehun zati berdinetan zatitzen du P1, P2, P3….. P99

Adibidez: P85=80 → Etxeen %85ak 80 litro ur edo gehiago kontsumitzen du egunean.







## 4.4. SAKABANATUTAKO NEURRIAK

* Datuen arteko hurbiltasuna (edo urruntasuna) adierazten dute, alegia haien arteko berdintasun edo desberdintasuna…
* Datu sortako balioek balio zentralarekiko (batezbestekoarekiko) duten hurbiltasun edo urruntasuna neurtzeko dira sakabanaketa neurriak.
* Zentralizazio edo posizio neurrien fidagarritasun maila kalkulatzen lagundu al digute: adibidez, ze punturaino dira esanguratsuak kalkulatu ditugun media edo mediana?

***IBILTARTEA***

Ibiltarteak banaketako balio handiena eta balio txikienaren arteko aldeari egiten dio erreferentzia **Ibiltartea= (xM - xm)**

Zenbat eta ibiltartea handiagoa izan, banaketako datuek berain artean duten sakabanaketa handiago izango dela pentsa genezake.

Bertuteak:

* Azkar eta erraz kalkulatzen da
* Emaitza aldagaiaren unitate berean eman dezakegu: Euro, metro, litro…-tan

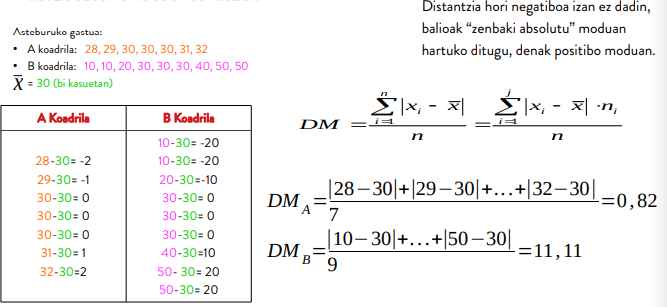
Mugak

* Muturreko datuak baino ez ditu erabiltzen (balio maximoa eta minimoa) eta ez banaketako balio guztiak
* Muturreko datuek era nabarmenean eragiten dute ibiltartea kalkulatzerakoan
  + A) 2, 20, 21, 22, 50 → ibiltartea = 48
  + B) 19, 20, 21, 22, 23 → ibiltartea = 4

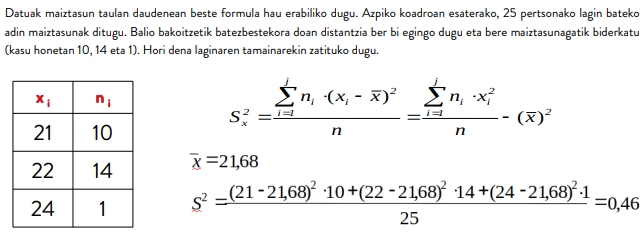
***BATEZBESTEKO DESBIDERATZEA***

Lehen eta azken balioa hartu ordez, multzoan dauden datu guztiak kontuan hartuta kalkulatzen dira hurrengo sakabanaketa neurriak.

Horretarako, datuak balio zentraletik (batezbestekotik) zein distantziara dauden kalkulatzen dugu.



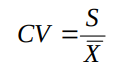
***BARIANTZA***



***DESBIDERATZE TIPIKOA (EDO DESBIDERATZE ESTANDARRA)***

******

***ALDAKORTASUN KOEFIZIENTEA***

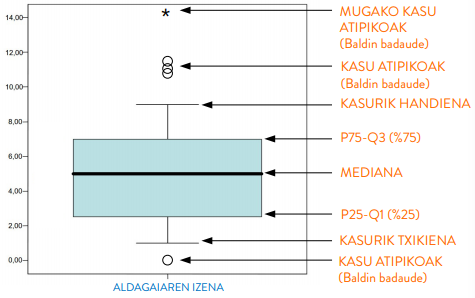
Aldakortasun koefizienteak desbideratze erlatiboa neurtzen du, desbideratze tipikoa batez bestekoaren artean zatituz

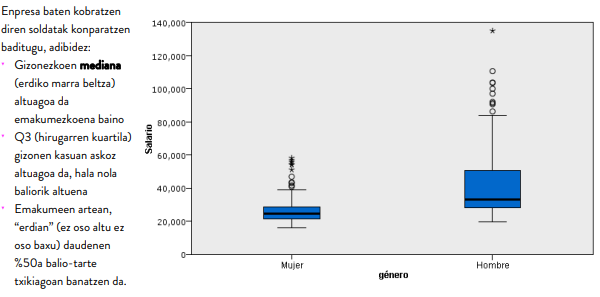
* Pisuaren aldakortasun koefizientea %29koa da, beraz oso talde heterogeneoa gara pisuari dagokionez. Aldakortasun handia duenean dfierentzia handiago izan bere barnean.
* Altueraren aldakortasun koefizientea %17 da, beraz nahiko talde homogeneoa gara altuerari dagokionez. Aldakortasun txikiagoa denean ez dago hain diferentzia handia.

## 4.5. KUTXA DIAGRAMAK ETA BANAKETAREN ITXURA NEURRIAK

***KUTXA DIAGRAMAK***

* Datuak nola banatzen diren erakusten digute
* Datu talde multzo bakarra deskribatzeko erabili daitezke baina normalean talde ezberdinen arteko ezberdintasunak eta aldakortasuna aztertzeko erabiltzen dira

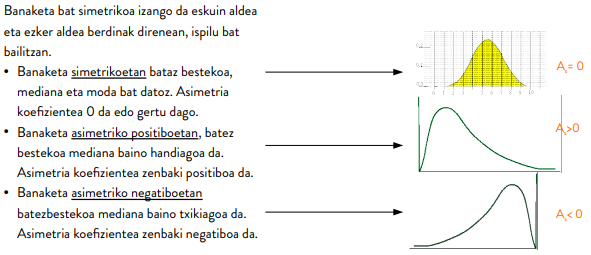


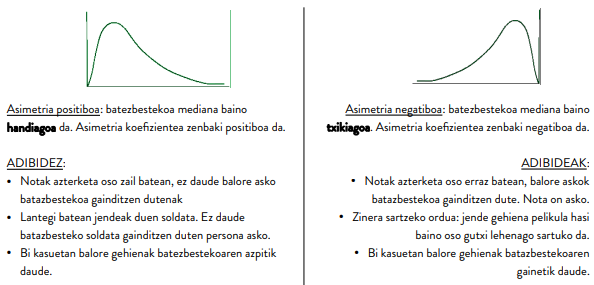


***BANAKETAREN ITXURAREN ITXURA-NEURRIAK***

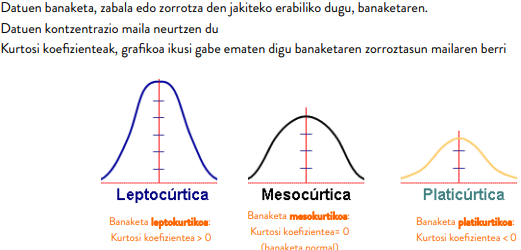
* Datuak era orekatuan sakabanaturik dauden ala ez ikusteko aukera ematen digute.
* Erreferentzia puntua batezbestekoa izango da, eta datuak bere ze aldetan pilatzen diren ikusiko dugu.
* Bi neurri ezagutuko ditugu:
  + Datuak era simetrikoan banatzen ote diren jakiteko → Asimetria koefizienteak
  + Banaketaren zorroztasun maila ezagutzeko → Kurtosi koefizienteak

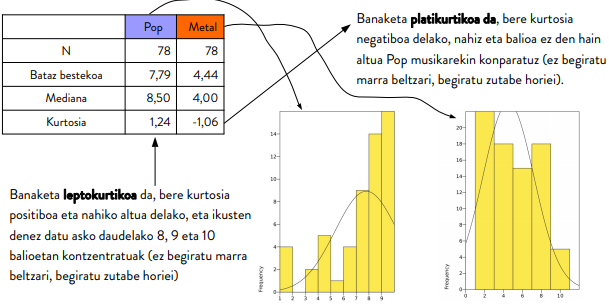
*ASIMETRIA KOEFIZIENTEA*





*KURTOSIA*

**



# ALDAGAI BIKO ESTATISTIKA DESKRIBATZAILEA

## 4.1. ALDAGAI NUMERIKOEN ARTEKO HARREMANA

* Aldagai askea: menpekoan eragin dezake (ala ez, hori da hain zuzen ere jakin nahi duguna)
* Menpeko aldagaia: aldagai askearen eragina jaso dezake (ala ez, hori da hain zuzen ere jakin nahi duguna)

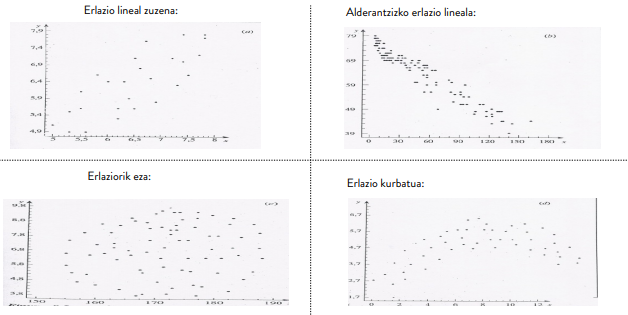
***KORRELAZIOA***

Aldagai kuantitatiboekin, dispertsio diagramak erabiltzen dira haien arteko harremana aztertzeko.

* Dispertsio diagraman, puntu bakoitzak banako bat adierazten du.
* Dispertsio diagramak bi aldagaien artean loturarik ote dagoen adieraziko digu.

Erlazio lineala bi eratakoa izan daiteke:

* Erlazio zuzena: Aldagai baten balioak handitzen direnean bestearenak ere handitzen badira, edo batenak txikitzean bestearenak ere txikitzen direnean.
* Alderantzizko erlazioa: Aldagai baten balioak handitzen direnean bestearenak txikitu egiten badira, edo batenak txikitzean bestearenak handitu egiten direnean.



***PEARSON-EN R***

Bi aldagaien artean korrelazio zuzenik dagoen adierazten digu.

Aldagaia neurtzeko erabilitako magnitudeak ez du emaitzan eragiten (bere balioa beti da -1 eta 1 bitartekoa)

* Emaitza -1 bada → alderantzizko erlazioa
* Emaitza 0 bada →erlaziorik eza
* Emaitza 1 bada →erlazio lineal zuzena

***DETERMINAZIO KOEFIZIENTEA***

Determinazio koefizienteak aldagai bati (x) esker azal daitekeen beste aldagaiaren (y) aldakortasun maila adierazten du

Adibidez: Gure klase kideen pisua eta altueraren artean harremanik dagoen jakin nahi dugu, eta ia zein puntu arte altuerak pisua baldintzatzen duen.

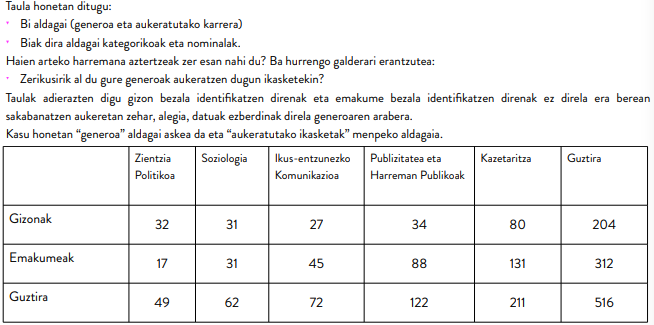
X= altuera eta Y=pisua

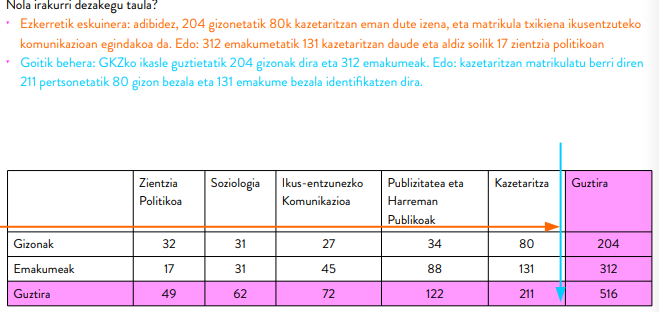
Pearson-en koefizientea, Rxy = 0,95 da: pisua eta altueraren artean erlazio lineal zuzen sendoa dago, altuerak gora egiten duten heinean, pisuak ere gora egiten du.

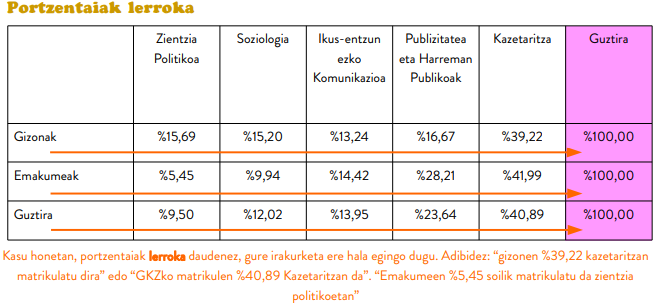
Determinazio koefizientea, R²= 0,95² = 0,90 da: pisuaren aldakortasunaren %90-a altueraren aldakortasunak azaltzen du.

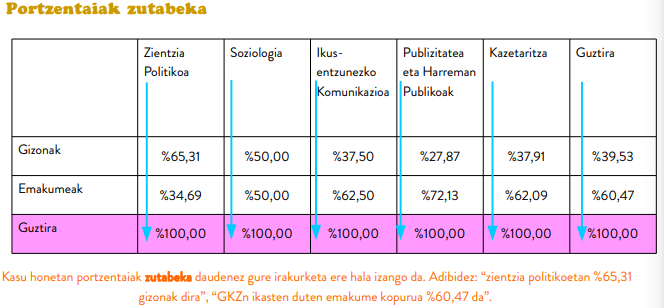
## 4.2. ALDAGAI KATEGORIKOEN ARTEKO HARREMANA

***KONTINGENTZIA TAULAK***





******

******

# ESTATISTIKA INFERENTZIALA

Ezin ditugunez populazio osoaren datuak bildu, laginak erabiltzen ditugu, baina gure asmoa lagin horrekin lortzen ditugun ondorioak populazio osoari buruzko baieztapenak egiteko baliagarriak izatea da.

Hori ezin da edonola egin, arau batzuk daude! Inferentzia estatistikoak probabilitatearen teorian eta bere banaketetan du oinarria. Laginean ikusitakoa populazio osora estrapolatzen laguntzen digu (ezaguna den) akats tarte determinatuarekin.

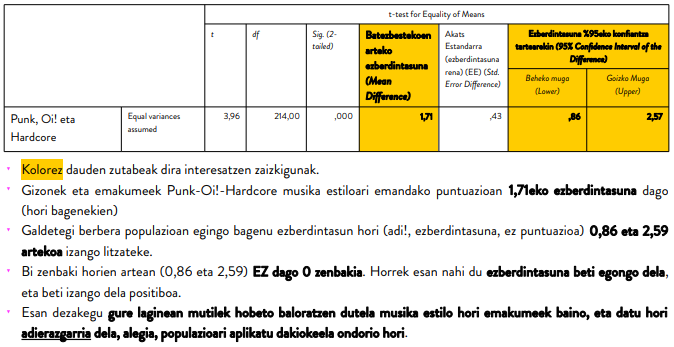
## 5.1. KONFIANTZA TATEZKO ESTIMAZIOAK

**Parametro estatistikotik estimaziora:** “Balio erreala” populazio osoaren inguruko datuekin soilik jakin dezakegu. Laginarekin balio horretara hurbil gaitezke: ESTIMAZIOA Autobus eta trenen ordutegietan →bezala, heltzeko denbora estimatua dakigu.

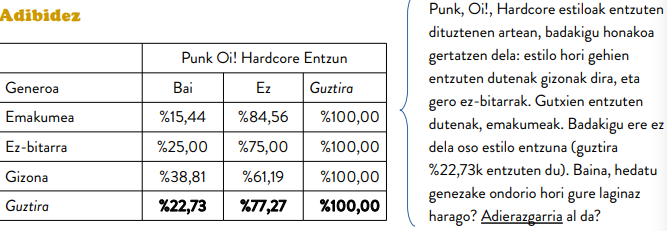
**Konfiantza-tartezko** **estimazioa**: tarte bat kalkulatzen dugu (bi balio) zeinek balio erreala barnebilduko duen probabilitate determinatu batekin. Horretarako konfiantza tartea definitu behar dugu.

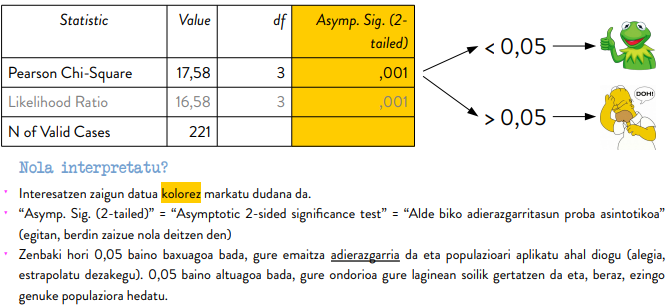
## 5.2. ADIERAZGARRITASUN FROGAK

Bi batezbesteko haien artean ezberdinak dira, eta jakin nahi dugu hori gure laginean soilik gertatzen den ala adierazgarria den, alegia, populaziora estrapolatu edo hedatu dezakegun ondorioa ote den.



## 5.3. JI KARRATUAN OINARRITZEN DEN ADIERAZGARRITASUN TEST-A





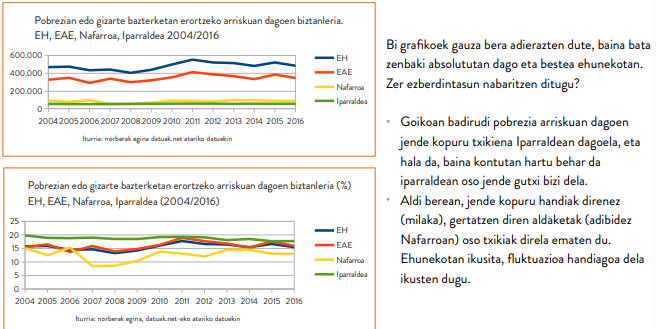
# BIGARREN MAILAKO DATUEKIN LANEAN

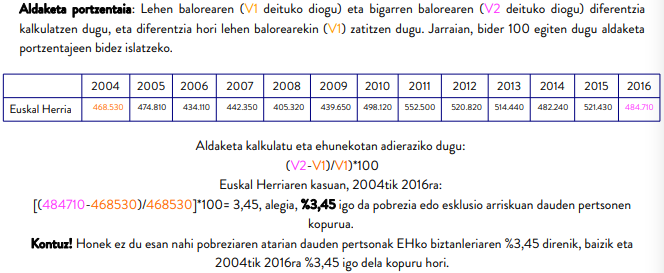
* Lehen mailako datuak: guk sortutakoak (adibidez, zuek egin duzuen inkesta)
* Bigarren mailako datuak: beste batzuek sortutako datuak, guk erabiltzen ditugunak. Gizarteko prozesu ezberdinen inguruan hausnartzeko eta informatzeko aukera ematen digu. Datu horiek gure kabuz biltzea oso zaila litzateke (denbora, gastuak, eskuragarritasuna...).

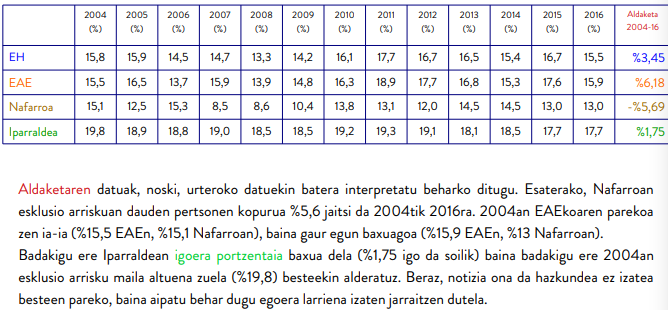
Errealitatearen inguruko informazioa sarritan portzentaietan dago. Portzentaiak zati bat osoarekiko harremanetan jartzen du. Bi elementu ditu:

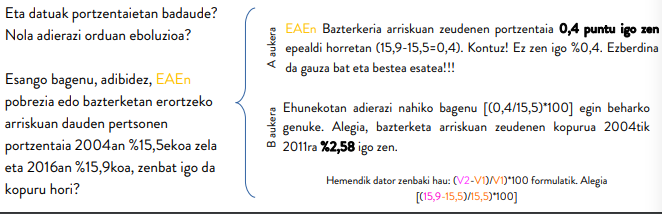
* Zenbakitzailea: harremanetan jarri nahi dugun zatia (emakumeak, langabetuak)
* Izendatzailea: zeinekiko jarri nahi dugun harremanetan (populazio osoa, populazio aktiboa)

## 6.1. NOLA AZTERTU DENBORA TARTE BATEAN DATUETAN GERTATZEN DIREN ALDAKETAK?









## 6.2. NOLA AZALDU BI DATUREN ARTEAN DAGOEN ALDEA?

EAEn batezbesteko errenta gizonezkoen kasuan 25.258 euro dira urtean eta emakumezkoen kasuan 14.534.

Bien arteko ezberdintasuna 10.724 eurokoa da. Ezberdintasun hori portzentaietan eman dezakegu emakumezkoen soldata oinarri hartuta [(10724/14.534)\*100], gizonen errenta emakumeena baino %73,8 handiagoa dela esanez.

Hartu genezake ere gizonena oinarri bezala (askotan hau egiten da), orduan [(10724/25.258)\*100]: emakumeen errenta gizonena baino %42,4 txikiagoa da