

# **Sailkapen-analisisa aldagai hertsiaik eta irekiak batera tratatuz.**

## **“EHUdenda” inkesta euskaraz erantzun zutenen kasua**

**Karmele Fernandez Aguirre**  
**EHUko irakaslea eta UEUko kidea**

Artikulu honen helburua estatistika esploratzaile anizkoitzaren aplikazio aurreratu bat aurkeztea da. Alde batetik, faktore-analisiak eta sailkapen-analisiak osagarri gisa erabiltzen dira baina, beste alde batetik, taula anizkoitzen kasuan kokaturik, izaera desberdina duten taulak batera aztertzen ditugu aldagai hertsiaik eta irekiak edo testualak batera tratatuz. Aplikazioa merkatuen ikerketaren arloan kokatzen da eta *On-Line* inkesta elebidun baten azpimultzo baterako garatzen da.

The object of this work is to show an advanced application in multivariate exploratory statistics. On one side, factor analysis and cluster analysis are applied in a complementary way; on the other side, in a framework of multiple tables of different kinds of variables, we jointly analyze tables of closed and open-ended (text) questions by means of a conjoint treatment of both types of variables. The applied work is placed in the area of market research and it is performed on part of an bilingual on-line survey.

## 1. Sarrera

Duela gutxi, 2006ko abenduaren 20an, hain zuzen, Euskal Herriko Unibertsitatearen *Uniberts-EHU Denda* zabaldu zen. Inauguraziorako gonbidapenak zioena hauxe izan zen: «Denda-kontzeptu berria sortu da unibertsitatearen munduari begira. Elkarretaratzeko eremua, oparirako objektuetan... unibertsitatea gizartera hurbildu nahi duena. Esperientzia berritzaile eta irekia, izanen dituen bisitarien heinera handitu, egokitu eta moldatuko dena».

Dendaren bidegarritasunaren proiektua gauzatzeko, bi inkesta antolatu genituen. Bata, hiru campusetako hiriburuetan EHUrekin zerikusia ez zutenentzat, kuota-laginketa aplikatuz. Bestea, EHUko ikasle, irakasle eta administrazioko langileen unibertsoa hartuz, web gunean egindako *On-line* inkesta, probabilitatezko laginketa aplikatuz.

*On-line* inkesta oso aberatsa izan zen eta erantzunetatik informazio esanguratsu franko atera ditugu. Lan honetan bi galdera hertsia eta horiekin zerikusia duen galdera testuala, hau da irekia, batera aztertuko ditugu. Lan honen bigarren atalean, aplikatuko ditugun analisi estatistikoaren ideia orokorra emango dugu; hirugarren atalean, lanaren helburua eta datuak aurkeztuko ditugu; laugarren atalean, aplikazioa aztertuko dugu; eta, amaitzeko, ohar batzuk egingo ditugu konklusio gisa.

## 2. Analisi estatistiko anizkoitzak

Aldagai-multzo bi edo gehiago batera, estatistikaren bidez, tratatzeko erabiliko dugun teknika Faktore Analisi Anizkoitza (FAA) da. Aldagaien izaeraren arabera, aldagai-multzo bakoitzari faktore-analisi mota bat dagokio: hots, aldagai kuantitatiboak direnean Osagai Nagusien Analisia (ONA); aldagai kualitatiboak edo kategorikoak direnean Korrespondentzi Analisi Anizkoitza (KAA), eta maiztasuntaula baten aurrean gaudenean Korrespondentzi Analisi Faktoriala (KAF), hain zuzen.

Estatistika esploratzaile anizkoitzaren barnean FAA estatistika-metodoak, aldagai-multzo (edo taula) bakoitzari dagokion analisia egin ondoren, taulak batera hartuz eta aldagai-multzo horien garrantziaren berdinketarako konpromisorik hoberena lortuz faktore orokorrak ateratzen ditu. Ponderazio egoki baten bidez taulen eraginaren oreka lortzen da (ikus Escofier eta Pagès, 1992). Metodoak, beste analisi faktorialak bezala, interpretaziorako laguntza ugari ditu: koefizienteak eta irudiak. Faktore-planoak garrantzi handiko irudiak dira, non indibiduoak, aldagaiak eta aldagai-multzoak proiektaturik aztertzen baitira. Elementu horien aldibereko aurkezpen grafikoa lagungarria izango da, halaber, faktoreen lorpenean parte hartu ez duten beste aldagai batzuk edota beste indibiduo batzuk, elementu gehigarriak deitutakoak, hain zuzen.

Sailkapen-analisia egin aurretik analisi faktorialak egingo ditugu, horien bidez, datuetan agertzen diren estrukturak aztertuko ditugu, azken finean, datuetan dagoen informazio nagusia lehen faktoreak interpretatzerakoan lortuko baitugu.

Ondoren, faktore-analisan lortutako lehen faktoreak hartuz, eta, bide batez, gorabehera aleatorioak jasotzen dituzten azken faktoreak baztertuz, indibiduen sailkapena egingo da (Lebart, 1994). Sailkapen baten bidez talde edo *cluster* disjuntutan indibiduen tipologia lortzen dugu. Sailkapen baten helburua talde homogeneoak barnealdean eta heterogeneoak elkarren artean lortzea da. Tipologiak dituen talde desberdinak neurtutako aldagaien arabera karakterizatuko dira.

### 3. Datuak eta Helburua

“EHUdenda” deitutako *On-line* inkestaren bidez 2005eko otsailaren 7tik 2005eko martxoaren 7ra informazioa bildu genuen. Hurrengo taulan, fitxa teknikoa ikus dezakegu. Laginaren tamaina finkatu ondoren, 2.289 ikasle, 768 administrazio eta zerbitzuko pertsonala eta 1.499 irakasle pertsonalki *On-line* inkestaren proiektuan parte hartzeko gonbidatuak izan ziren.

<b>IKASLEAK</b>
<b>Populazioa:</b> 48.995
<b>Laginaren tamaina:</b> 547
<b>Laginketa mota:</b> Geruzatua sexu, campus eta zikloaren arabera proportziozko esleipenarekin.
<b>Laginketa-errorea:</b> 0.042
<b>Konfiantza-maila:</b> 0.95
<b>ADMINISTRAZIO ETA ZERBITZUKO PERTSONALA</b>
<b>Populazioa:</b> 1128
<b>Laginaren tamaina:</b> 444
<b>Laginketa mota:</b> Geruzatua sexu, campus eta adinaren arabera proportziozko esleipenarekin.
<b>Laginketa-errorea:</b> 0.036
<b>Konfiantza-maila:</b> 0.95
<b>IRAKASLEAK</b>
<b>Populazioa:</b> 3982
<b>Laginaren tamaina:</b> 754
<b>Laginketa mota:</b> Geruzatua sexu, campus eta adinaren arabera proportziozko esleipenarekin.
<b>Laginketa-errorea:</b> 0.032
<b>Konfiantza-maila:</b> 0.95

1. taula.

Hasieran, estamentu bakoitzean aukeratutako laginaren tamaina laginketa-errore maximoa % 2koa izateko izan zen. Lortu ziren erantzunak % 23,9; % 57,81 eta % 50,30 izan ziren, hurrenez hurren. Ikasleen kasuan ehuneko txikia da, hala eta guztiz, mota hauetako inkestetan erantzunak ez dira izaten % 20koa baino gehiagokoak eta ez zen emaitza txarra izan.

Lortutako informazioarekin lan estatistiko asko egin ditugu (Fernandez Aguirre *et al.*, 2005 eta 2007). Aldagai kuantitatiboko zein kualitatiboko taulak aztertu ditugu, ikuspuntu eta helburu desberdinak finkatuz. Oraingo honetan, dugun helburua honako hau da: *Uniberts-EHU Dendan* salgai egongo ziren logotipodun produktuen aurrean zein den erosteko edo ez erosteko joera eta erakundean lan edo ikasten dihardugunen artean zein tipologia aurkezten den. Helburua lortzeko aldagai hertsia eta irekia (testualak) batera aztertuko ditugu, oso gutxitan egin dena, hain zuzen, estatistika anizkoitza erabiliz garatuko dugu.

Bi aldagai hertsiaik honako hauek dira:

- *Oro har, pozik nago UPV/EHUko kidea izateaz. Analisisetan, Satis2 laburduraz deitutako aldagaiak hiru kategoria ditu: Satis2=1, Satis2=2 eta Satis2=3. Hasieran, esaldi horren bost erantzun posibleak 1etik (ez nago batere ados), 5era (guztiz ados nago), hiru erantzunetan honela birkodifikatu ditugu: Satis2=1 (hasierako 1, 2 eta 3), Satis2=2 (hasierako 4) eta Satis2=3 (hasierako 5)<sup>1</sup>.*
- *Erosiko zenuke UPV/EHUren logotipoa daraman produkturen bat, zeuk erabiltzeko edo beste inori oparitzeko? (Adibidez, iduneko zapiak, giltzak-eta uzteko platertxoa, kamisetak, erlojuak, katiluak...).* Galderak, BuyLogo laburduraz analisisetan, bi kategoria ditu: Buylo=1 eta Buylo=2, Bai eta Ez, hurrenez hurren.

Aldagai irekia honako hau da:

- *Azaldu hemen zergatik, mesedez.*

*Erosiko zenuke...* galderaren ondotik inkestatuak nahi beste idatziz eman zezakeen zergatia eta inkesta elebiduna izan zenez, gaztelaniaz zein euskeraz. Azalpen testualak 1.553 erantzun izan zituen guztira eta horietatik 304 soilik euskaraz. Lan honetan dugun erronka aldagai hertsiaik eta irekia batera erabiliz erantzuleen sailkapen bat lortzea da.

## 4. Analisi estatistikoaren emaitzak

### 4.1. Analisi testualaren zenbait berezitasun

Faktore-analisiak aplikatu aurretik analisi testualak (Etxeberria *et al.*, 1995; Lebart *et al.*, 1998; Lebart *et al.*, 2000) dituen prozedura bereziak aplikatu behar ditugu. Lehen pausoa testu-corpora finkatzea da, lexiko-unitateak<sup>2</sup> bere maiztasunekin programa egoki baten bidez finkatuz, adibidez (SPAD, 2006).

Lexiko-unitateak finkatzerakoan lanaren helburua garrantzi handikoa da (Fernandez Aguirre, 2003), batzuetan, forma grafiko esanguratsuak hartzen dira; hau da, izenak, aditzak, izenordeak eta aditzondoak, beste guztiak baztertuz. Hau da, lematizazioa egingo da lemaren inguruan forma grafikoak bilduz. Lematizazioaren prozedura ezin da guztiz automatikoki egin, badaude programa lagungarriak baina edozein kasutan ikertzaileak milaka forma grafikoak ikusi eta aztertu beharko ditu.

Testu-corpora finkatu ondoren maiztasun-taula lexikala behar da. Taulak inkestatu ditugun indibiduo beste errenkada eta lexiko-unitate beste zutabe ditu, *i.* errenkada eta *j.* zutabea gurutzatzen diren tokian *j.* lexiko-unitateak duen maiztasuna *i.* inkestatuaren kasurako dugu<sup>3</sup>.

1. Hasierako 1 eta 2 kategorien aukera hartu zuten inkestatuen urritasuna izan da birkodifikazioaren arrazoia.

2. Lexiko-unitatea: bi hutsune edo banatzaile artean idatzitako *forma grafikoa* da eta *segmentu erreplikatu* bi edo lexiko-unitate gehiagoz osatzen dena eta corpusean erreplikatu dagoena da.

3. Halaber, maiztasun-taula agregatuak antola daitezke, lexiko-unitateak eta inkestatuen kategoriak: adin-taldeak, gizonezkoak eta emakumezkoak... guruzatuz, hain zuzen.

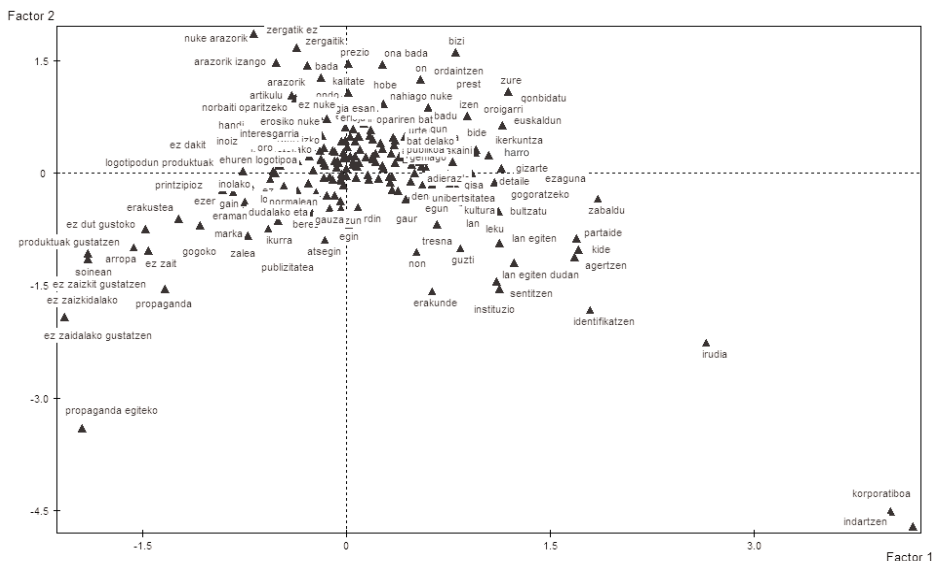
## 4.2. Korrespondentzi Analsiaren emaitzak

Korrespondentzi Analsi Faktoriala egin ahal izateko, analsi testualaren kasuan komeni da corpusak minimo batzuk betetzea eta gure kasuan ozta-ozta bete dira. Erantzunak 304 izan ziren, 4.756 hitz guztira eta 1.637 hitz desberdin. Horrela, lematizazio-prozedurak baino maiztasun handiagoak ematen dizkigun tematizazio-prozedura jorratu dugu.

Mintzagaiak kontutan hartuz, tematizazioa egin ondoren 4 bider edo gehiago errepikatzen diren 223 forma grafiko eta segmenturekin geratu gara, horrela (304 x 223) maiztasun edo kontingentzi taula dugu.

Lehen analsi estatistikoa lortutako kontingentzi taularen KAF izango da, aldagai hertsia gehigarriko aldagai gisa erabiliz. Hau da, Satis2 eta BuyLogo aldagaien kategoriak ardatz eta plano faktorialetan proiektatuko dira, ardatz faktorialearen eraketan parte hartu gabe. KAFren propietate garrantzitsu bat, honako hau da: plano faktorialetan proiektatutako elementuak, aktiboak zein gehigarrikoak, aukeratu dituzten indibiduo inkestatuen grabitate-zentroan kokatzen dira.

Analsiaren bi lehen ardatz faktorialek —lehen horizontala eta bigarrena bertikala— osatutako plano nagusia ikus dezakegu 1. grafikoan.



1. grafikoa.

Dakusagun irudia, KAF baten irudi tipikoa da, parabola baten antzekoa, hain zuzen. Logotipodun produktuak erostearen alde agertzen direnen zergatiak islatzen dira planoaren laugarren koadrantean, eta hortxe daude: *korporatiboa*, *indartzen*, *irudia*, *identifikatzen*, *sentitzen*, *lan egiten dudun*, *instituzio*, *erakunde*, *agertzen*, *partaide*, *zabaldu*... Lehen koadrantean, joera berdina ikusten dugu: *harro*, *bide*, *ikerkuntza*, *euskaldun*, baina gero ta gorago joatean eta parabolaren erdian kokaturik, lehen koadrantean zein bigarren koadrantean ikusten dugun hiztegiak

erosketa baldintzatu edota zalantzazko batena dirudi, horrela: *gonbidatu, oroigarri, opariren bat, nahiago nuke, ona bada, kalitate, prezio, zergatik ez, arazorik izango, norbaiti oparitzeko, interesgarria, ez dakit*. Azkenik, hirugarren koadrantean erostearen kontra daudenen hiztegia islatzen da, ikus dezakegu: *inolako, ezer, ez dut gustuko, ez zaizkit gustatzen, soinean, arropa, propaganda, ez zaidalako gustatzen, propaganda egiteko...*

Ondoren, aldagai hertsien kategorien proiektzioen *test-values*, 2. taulan, ikus ditzakegu lehen eta bigarren ardatz faktorialetan. Aldagai horiek ez dute parte hartu aldagai testualaren KAFn baina lortutako faktoreetan proiektatzen dira, horrela batzuetan kokapen pribilegiatua edukita, asko lagundu ahal dute emaitzen azalpenetan.

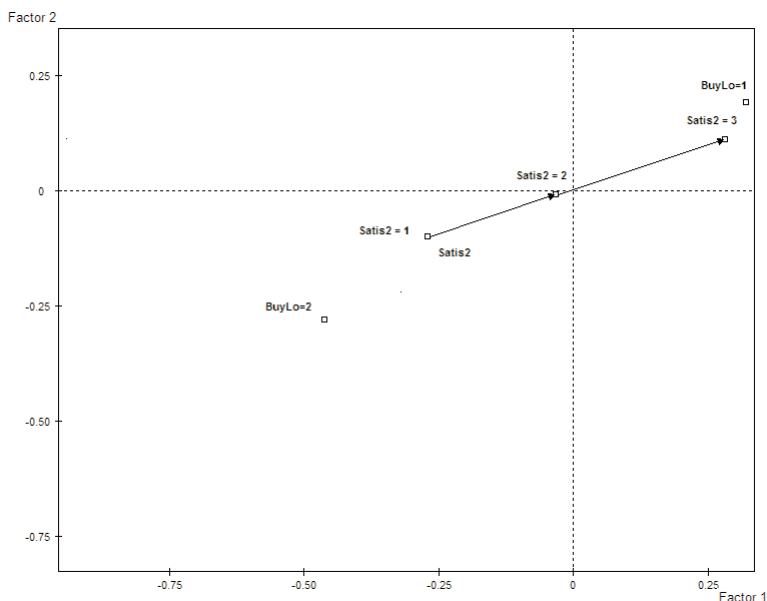
**Test-values of supplementary categories**

Label	Count	Absolute weight	Distance to origin	Axis 1	Axis 2
<b>BuyLogo</b>					
BuyLo=1	175	2306,00	1,86071	6,51	3,87
BuyLo=2	128	1575,00	2,62490	-6,86	-4,16
missing category	1	14,00	277,21400	-0,94	0,11
<b>Satis2</b>					
Satis2 = 1	76	914,00	5,84224	-2,72	-1,02
Satis2 = 2	135	1742,00	2,45687	-0,50	-0,16
Satis2 = 3	86	1163,00	4,66058	3,08	1,20
missing category	7	76,00	57,97770	-0,85	-0,59

**2. taula.**

Test-balio horien balio absolutua 2 edo handiagoa bada kategoriak posizio esanguratsua du —% 5 edo gehiagoko esangura-mailaz— ardatzaren gainean. BuyLogo aldagaiaren kasuan, erostearen alde daudenen kategoriak (BuyLo=1) ardatzaren alde positiboan oso posizio esanguratsua du eta berdin gertatzen da erostearen kontra daudenen kategoriarekin (BuyLo=2), alde negatiboan izanik kasu honetan. Bi kategoriek badute ere posizio esanguratsua bigarren ardatzean, baina ez hain garrantzizkoa. Satis2 aldagaiaren bi muturretan dauden kategoriek posizio esanguratsua dute lehen ardatzean. Horrela, pozik ez daudenak (Satis2=1) alde negatiboan daude eta oso pozik daudenak (Satis2=3), berriz, alde positiboan.

Hurrengo irudian, bi lehen ardatz faktorialek osatutako plano nagusia ikus dezakegu, non aipatutako gehigarritzko kategoriak proiektatuta baitaude. Hiztegiarekin aztertu dugun plano berbera da, baina, puntuak ikusi ahal izateko bikoiztu dugu plano eta elementu gehigarritzkoak soilik proiektatu ditugu 2. grafikoan.



2. grafikoa.

Satis2 aldagaiaren bigarren kategoriak posizio guztiz neutrala du, planoaren jatorriaren ondoan, hau da, inkestaren erantzuleen grabitate-zentroaren ondoan, hain zuzen. Planoaren eskuin aldean eta gutxi gorabehera erostearen alde hiztegi nabaria dutenen tokian EHUko partaideak izateaz oso pozik daudenak eta produktu logotipodunak erostearen alde azaldu direnak. Ezker aldean, berriz, eta erostearen kontrako hiztegia dutenen tokian erakundearekin pozik ez daudenak eta produktu logotipodunak erostearen kontra azaldu direnak.

### 4.3. Faktore Anizkoitza eta sailkapenaren emaitzak

FAA estatistika-metodoaren bidez, 3.2. atalean aztertu ditugun aldagai-multzo biak aktibotzat hartuz, faktore orokor batzuk aterako ditugu. Analsiari aldagai-multzo (edo taula) bakoitzari dagokion analisisia egin ondoren, taulak batera hartuta ateratzen ditu faktore orokorrak. Kasu horretan, aldagai hertsia —lehen multzokoak— kategorikoak direnez eta irekia —bigarren multzokoak— kontingentzi edo maiztasun-taulakoak, ez da FAA arrunta egiten dena, maiztasun-taulan ditugun indibiduen maiztasun erlatiboaren arabera kontuan hartu behar baititugu (ikus, Bécue- Bertaut eta Pagès, 2007). Aztertzen ari garen kasuan, pisu horiek inkestatuek idatzi dutenaren arabera dira.

Analisi partzialeko lehen autobalioak, honako hauek dira: 0,6576 lehen multzoaren kasuan eta 0,37 bigarrenaren kasuan. Hau da, FAAk bi multzoen garrantzia orekatzen du eta analisi orokorraren lehen ardatzak jasotzen duen inertzian —sakabanatze-neurriaren— balioari multzo bakoitzak egindako ekarpena % 54,69 da aldagai hertsien kasuan eta % 45,31 testualaren kasuan, beraz, esan dezakegu dispersio- edo sakabanatze-norabide bakarra dagoela bi multzoetarako eta erosketa-joeraren “maila” gisa interpreta dezakegu. Bestalde, analisi orokorraren lehen autobalioa 1,513 da, 2ren —balio handienaren— nahiko hurbila eta horrek

esan nahi du dispertsio-norabide hori bi multzoek bakoitzak bere aldetik dutenaren baliokidea dela.

FAAren lehen bost faktoreak aldagai berritzat hartuta sailkapen hierarkikoa aplikatu dugu eta zuhaitz hierarkikoaren mozketan egokia hartuz 4 sail lortu ditugu. Sailetan dauden inkestaturen ehunekoak eta horien ezaugarriak aldagai hertsien arabera, 3. taulan ikus ditzakegu. Aldagai hertsien kategoria bakoitzaren test-balioa “*Test-value*” bi ehunekoen parekatzeaz lortzen da, kategoriak sailean duen ehunekoa eta kategoriak multzo osoan duena, hain zuzen. Adibidez, Satis2 aldagaiaren 2. kategoria (Satis2=2) lehen sailean dauden 89 inkestaturen artean ehuneko altu batean aukeratua izan da, multzo osoan ditugun 304 inkestaturekin parekatuz.

**Characterisation by categories of groups of**

**CUT "a" OF THE TREE INTO 4 CLUSTERS**

**Group: Cluster 1 / 4 (Count: 29 - Percentage: 29.26)**

Variable label	Characteristic categories	Test-value
Satis2	Satis2 = 2	6,81
BuyLogo	BuyLo=1	5,74

**Group: Cluster 2 / 4 (Count: 22 - Percentage: 22.34)**

Variable label	Characteristic categories	Test-value
Satis2	Satis2 = 3	7,97
BuyLogo	BuyLo=1	4,71

**Group: Cluster 3 / 4 (Count: 25 - Percentage: 24.64)**

Variable label	Characteristic categories	Test-value
BuyLogo	BuyLo=2	6,64
Satis2	Satis2 = 2	2,44

**Group: Cluster 4 / 4 (Count: 24 - Percentage: 23.77)**

Variable label	Characteristic categories	Test-value
Satis2	Satis2 = 1	9,51
BuyLogo	BuyLo=2	2,80

**3. taula.**

Dakusagunez, inkestaturen ehunekoak lau sailetan berdintsuak dira. Bi lehen sailetan erostearen aldeko inkestatuak dira nagusi, lehengoan, bide batez, pozik daudenak dira nagusi eta bigarrenean guztiz pozik daudenak. Hirugarren eta laugarren sailetan erostearen kontrakoak dira nagusi, halaber, hirugarrenean pozik daudenak eta laugarrenean pozik ez daudenak.

Azkenik, alde batetik, aldagai testualaren aldetik sailen hitz eta segmentu karakteristikoak, hau da, gain-gertatuak eta horien test-balioak, eta bestetik, esaldi modalak, hau da, sailean hitz karakteristikoak dituztenak ikusiko ditugu.



**Cases Group : Cluster 1/4**

Characteristic words or segments	Internal percentage	Global percentage	Test-Value
lan	1,14	0,50	3,048
artikulu	0,57	0,20	2,709
aukera	0,76	0,31	2,678
zergaitik	1,33	0,70	2,589
on	1,24	0,64	2,518
indartzen	0,38	0,11	2,434
korporatiboa	0,38	0,11	2,434
bat delako	0,38	0,11	2,430
irudia	0,76	0,34	2,397
nola	0,48	0,17	2,329
ezaguna	0,95	0,48	2,297

**Cases Group : Cluster 1/4**

Classification criterion	Number	Answer label
2,028	1	polita delako
1,610	2	unibertsitatearen irudia indartu eta suspertzeko
1,600	3	konnotazioak orokorrean onak direlako
1,597	4	lankide batentzat opari egokia delako
1,499	5	irudi korporatiboa indartzen duelako. Opari egokia izan daitekeelako eremu profesionalean
1,483	6	laguntzeko modu on bat delako
1,374	7	bere irudi instituzionala erakargarria delako, pozik sentitzen naizelako kolektibo honetan lan egiteaz
1,340	8	publizitatea baita eta hori da empresa baten helburua, ezaguna izatea.
1,233	9	korporazioarekin identifikatzen naizelako
1,097	10	egokiak izan daitezke ikerketetan (batez ere soziologian, psikologian, etab.) elkarrizketatzen diren pertsonen oparitzeko edo gonbidatzen diren kanpoko irakasleei oparitzeko

Lehen sailean, logotipodun produktuak erostearen alde eta erakundearekin pozik daudenen artean, hiztegia eta esaldiak begiratzean, unibertsitatearen indartzeaz eta bere irudiaz arduratzen direla argi dago, zentzu horretan lagungarria dirudi dendaren proiektuak erakundearen identitatea suspertzeko. Gainera, erakundearekin identifikatzen dira.

**Cases Group : Cluster 2/4**

Characteristic words or segments	Internal percentage	Global percentage	Test-Value
ikasi	2,38	1,04	3,720
harro	0,63	0,14	3,261
euskal	0,75	0,20	3,195
publikoa	0,63	0,17	2,781
egon	2,01	1,04	2,689
bat	3,88	2,61	2,358
bertako	0,88	0,36	2,225
unibertsitatea	3,01	1,96	2,189
herriko	0,63	0,22	2,124
eskaini	0,63	0,22	2,124

**Cases Group : Cluster 2/4**

Classification criterion	Number	Answer label
2,088	1	unibertsitate honen partaide izateak ikastegi batean ikastea baino harago egon behar lukeelako
1,934	2	ikasten dudan lekuaz harro nagoelako
1,595	3	unibertsitatea gizarteari zabaltzeko
1,522	4	laguntza bezala
1,405	5	unibertsitatea beharrezkoa dela aldarrikatzeko
1,280	6	Euskal Herrian kokatutako erakunde kultural baten produktua litzatekeelako
1,266	7	zure unibertsitateko garaiko oroigarri bat izango litzatekeelako
1,242	8	unibertsitate publikoari propaganda egiteko
1,240	9	ni oso harro nago EHU ikasten ari naizelako; izan ere, betiko nire nahia bertan ikastea baitzen. Behin EHU ikasten egonik, oso harro egongo nitzateke holako produktua nire eskutan edukitzeaz eta jendeak ikustez ni EHUKo ikasle bat, eta beraz bere parte bat naizela
1,111	10	azkenean hemengo produktu bat erosiko nuke eta horrekin unibertsitateari lagunduz euskal gizarte osoari laguntzeko ahalegin txiki bat izango litzateke

Bigarren sailean, erostearen alde eta oso pozik direnen artean, hiztegia eta esaldietan ikusten dena, honako hau da: euskal unibertsitate publikoa, bertakoa, erakunde kultural nagusia... dela azaltzen dute eta horregatik harro izateaz nabarmentzen dira. EHU gure unibertsitatea izanik, zabaltzeko eta aldarrikatzeko ongi ikusten dute dendaren proiektua eta gainera gizarte osorako lagungarria izango dela uste dute.

**Cases Group : Cluster 3/4**

Characteristic words or segments	Internal percentage	Global percentage	Test-Value
ez	11,23	7,62	4,346
gogoko	4,05	2,24	3,766
ez dut uste	1,27	0,42	3,751
ez zait	2,43	1,15	3,555
produktuak gustatzen	0,58	0,14	3,140
horrela	1,50	0,64	3,133
ez zaizkit gustatzen	0,81	0,25	3,058
uste	1,97	0,98	2,983
arropa	0,46	0,11	2,705
eraman	2,20	1,26	2,534

**Cases Group : Cluster 3/4**

Classification criterion	Number	Answer label
3,343	1	logotipoak ez zaizkit gustatzen
3,122	2	ez zaizkidalako horrelako produktuak gustatzen
2,841	3	logotipoa gustatzen zait
2,527	4	logotipoa daramaten produktuak gustukoak ez ditudalako
2,526	5	ez zaizkit logotipoak dauzkaten produktuak gustatzen
2,367	6	logotipoak ez eramateko joera daukat
2,306	7	ez dudalako gustuko horrelako propaganda eramatea gainean
2,270	8	horrelako gauzak erostea ez dut gustuko
2,148	9	halakoak erabiltzea ez dudalako gustuko
2,129	10	ez zait gustatzen edozein publizitate eramatea

Hirugarren sailaren kasuan, erostearen alde ez direnen baina pozik direnen artean, ez dituzte gogokoak eta ez zaizkie gustatzen logotipodun produktuak, ez zaie gustatzen eramatea ezta propagandarik egitea. Merkatuen ikerketaren aldetik ikusirik, talde horretan egon daitezkeen % 25 erostearen alde dauden talderako berreskura daitekeela dirudi.

**Cases Group : Cluster 4/4**

Characteristic words or segments	Internal percentage	Global percentage	Test-Value
zalea	0,82	0,25	3,086
ez	9,81	7,62	2,637
asko	0,93	0,36	2,633
propaganda	1,52	0,78	2,426
logo	3,97	2,80	2,194
ikusi	1,29	0,73	1,885
batere	0,58	0,25	1,736
atsegin	0,35	0,11	1,693
ez zait	1,75	1,15	1,646
diseinua	0,47	0,20	1,543
interesgarria	0,93	0,53	1,536

**Cases Group : Cluster 4/4**

Classification criterion	Number	Answer label
1,921	1	ez zaidalako inoren ez ezeren propagandarik egitea gustatzen. Ez nioke begiratuko logoa duen edo ez
1,908	2	ez naiz 'marka' zalea
1,807	3	ez zaidalako propaganda egitea gustatzen
1,807	4	ez zaidalako gustatzen propaganda egitea
1,727	5	propaganda egiteko
1,583	6	ez naiz logotipo instituzionalen zalea
1,550	7	ez zaidalako gustatzen inolako eratako propagandarik ibiltzea
1,530	8	ez naiz horrelako propaganda egin zalea
1,514	9	ez dudalako interesgarririk ikusten
1,491	10	ez zaizkit produktu logotipodunak gustatzen

Laugarren sailean, argi eta garbi ezezko joera dutenak dira nagusi, erostearen kontra agertzen dira eta ez daude pozik erakundearekin. Ez dira marka zaleak, ez zaie inoren ez ezeren propagandarik egitea gustatzen, eta berdin zaie instituzionala izatea edo ez. Erostearen alde daudenen talderako berreskurazekin dirudite.

## 5. Konklusioak

- Analisi testualak edozein inkestaren analisia asko aberasten du. Zentzu horretan, KAFren plano nagusiak oso informatiboak izaten dira, aztertu dugun kasuan gertatzen den bezala. Gainera, funtzio azaltzailea betetzen duten aldagai gehigarrikoak proiektatuz, interpretaziorako lagungarriak dira.
- FAA birmoldatu baten bidez, mota desberdinetako aldagaiak batera “*mixed data*” azter ditzakegu eta garrantzi berdintsua multzo bakoitzak edukiz, globalki dagoen dimentsio nagusia atera dezakegu, gure kasuan erosteko joeraren ardatza eta horren “maila” neurtzeko balio duena.

- Azkenik, aplikatutako FAA bereziak distantzia orokor bat indultzatzen du, non aldagai-multzoen eragina orekatzen baita. Horrela, faktore nagusiak hartuz, aldagai guztietan oinarritutako inkestatuen sailkapena lortu dugu. Sailak aldagai hertsien bidez eta hiztegi-erantzunen bidez argiro karakterizatuta ditugu eta, aztertu dugun kasuan kokaturik, merkatuaren ikerketarako guztiz lagungarriak.

## Bibliografia

- Bécue-Bertaut, M. eta Pagès, J. (2007): "Clustering units from frequency and nominal variables: definition of a global distance. Application to survey data with closed and open-ended questions", *Journal of Computational Statistics and Data Analysis*, (argitaratzeko zain).
- Etxeberria, J.; Garcia, E.; Gil, E. eta Rodriguez, G. (1995): *Análisis de Datos y Textos*, RA-MA, Madrid.
- Escofier, B. eta Pagès, J. (1992): *Análisis Factoriales Simples y Múltiples*, EHU-Argitalpen Zerbitzua, Bilbao.
- Fernandez Aguirre, K. (2003): "Análisis textual: generación y aplicaciones", *Metodología de Encuestas*, 5. lib., 1. zk.
- Fernandez Aguirre, K.; Martín, A. M.; Modroño, J. I. eta Zorrilla, P. (2005): *Análisis de la viabilidad de la tienda universitaria de la UPV/EHU "EHUdenda" de acuerdo con la imagen percibida de la institución*, txosten teknikoa, UPV/EHU, Bilbao.
- Fernandez Aguirre, K.; Landaluce Calvo, M. I.; Martín Arroyuelos, A. M. eta Modroño Herran, J. I. (2007): "Data Mining of an On-Line Survey. A Market Research Application", *Data Analysis, Machine Learning and Applications*, The 31<sup>st</sup> Annual Conference of the German Classification Society (GfKl), Freiburg.
- Lebart, L. (1994): "Complementary use of correspondence analysis and cluster analysis" in M. J. Greenacre, eta Blasius (arg.), *Correspondence Analysis in the Social Sciences*, Academic Press, London, 162-78.
- Lebart, L.; Salem, A. eta Berri, E. (1998): *Exploring Textual Data*, Dordrecht. Kluwer. Academic Publisher.
- Lebart, L.; Salem, A. eta Bécue, M. (2000): *Análisis estadístico de textos*, Milenio
- SPAD 6.0, (2006): *Système Portable d'Analyse des Données Numeriques*, <<http://www.decisia.fr>>, Paris.

