

MaxDiff TEST analiza (primjer)

Danijel Kopčinović, IT Market

MAXDIFF TEST

U upitniku "MAXDIFF TEST" postaviti ćemo Vam X pitanja s raznim varijantama, a Vi ćete u svakom pitanju odabrati Vama najbolju i najlošiju varijantu. Za ispunjavanje cijelog upitnika po našoj je procjeni dovoljno 15-20 minuta.

Kao znak zahvale za sudjelovanje u ovom upitniku, Y će Vam poslati mali poklon paket!

U ovom upitniku koristimo sljedeće osobne podatke:

- adresa e-pošte ispitanika
- IP adresa

Navedene osobne podatke koristit ćemo u svrhu identifikacije ispitanika i kreiranje profila ispitanika. Navedene podatke nećemo koristiti niti u koju drugu svrhu i nećemo ih dijeliti niti s kojom trećom stranom bez pismene privole ispitanika.

Više o politici zaštite osobnih podataka tvrtke Y pogledajte [ovdje](#).

Prihvaćam pravila o zaštiti osobnih podataka ☐

Uvod

U ovom dokumentu ćemo prezentirati testno istraživanje MaxDiff metodom. Istraživanje je bazirano na (pretpostavljenoj) želji lokalne pizzerije da ispita koju novu pizzu bi kupci najviše htjeli vidjeti u ponudi.

Nove pizze između kojih ispitanici biraju su:

```
## [1] miješana          povrtna          pikantna          4 vrste sira
## [5] s tunom           ribarska         4 godišnja doba calzone
## [9] slavonska        rukola/pršut     bolonjez          lovačka
## [13] losos            sa salamom
## 14 Levels: 4 godišnja doba 4 vrste sira bolonjez calzone losos ... slavonska
```

Osim glavnih proizvoda koji su nam u fokusu, zanima nas i mogući utjecaj “demografskih” faktora na rezultate:

```
## $lokacija
## $lokacija$naziv
## [1] "lokacija"
##
## $lokacija$vrjednosti
## [1] "blizu pizzerije"      "daleko od pizzerije"
```

U MaxDiff metodi ispitanici u svakom pitanju biraju najbolji i najlošiji proizvod između nekoliko alternativa. Pogledajmo neke kombinacije koje smo ispitivali, odn. postavili kao pitanja.

```
## task items best_choice worst_choice
## 1 1 4 vrste sira,slavonska,sa salamom ? ?
## 2 2 calzone,miješana,slavonska ? ?
## 3 3 4 godišnja doba,4 vrste sira,ribarska ? ?
## 4 4 bolonjez,calzone,ribarska ? ?
## 5 5 slavonska,ribarska,losos ? ?
```

Iz skupa svih 364 kombinacija proizvoda s 3 alternativa (broj alternativa smo postavili kao parametar dizajna), odabrali smo 54 reprezentativnih. S odabranim kombinacijama moguće je izračunati (prosječnu ili

individualnu) vrijednost svakog proizvoda za ispitanike.

Sve kombinacije podijelili smo na 3 upitnika, od kojih svaki ima 6 pitanja, od kojih svako ima 3 opcija/alternativa.

Molimo odaberite Vama najbolju i najlošiju opciju među navedenim opcijama

★Opcije:

NAJBOLJA		NAJLOŠIJA
<input checked="" type="radio"/>	4 vrste sira	<input type="radio"/>
<input type="radio"/>	lovačka	<input checked="" type="radio"/>
<input type="radio"/>	calzone	<input type="radio"/>

🔗 Kliknite na kružić (radio dugme) za odabir

Da bismo bolje mogli procijeniti vrijednosti različitih opcija, dodali smo i pitanje u kojem ispitanik iskazuje općenitu pozitivnu ili negativnu tendenciju prema nekoliko odabranih opcija. Opcije u tom pitanju su odabrane s obzirom na prijašnje odabire, tako da sadrže raspon pizza, od najčešće biranih kao najbolje do najčešće biranih kao najlošije. U ovom pitanju se ne uspoređuju opcije jedna s drugom nego svaka u odnosu na tzv. graničnu opciju, koja u različitim kontekstima može značiti različite stvari, ali općenito znači točku između želim/ne želim, kupio bih/ne bih kupio, zanima me/ne zanima me i sl. Ovo pitanje nam omogućuje da bolje uskladimo odgovore svih ispitanika u odnosu na graničnu opciju pa time i međusobno, čime dobivamo relevantnije procjene vrijednosti.

★Kad biste imali mogućnost kupovine ovih opcija, da li biste ih kupili ili ne?

	KUPIO BIH	NE BIH KUPIO
miješana	<input checked="" type="radio"/>	<input type="radio"/>
lovačka	<input type="radio"/>	<input checked="" type="radio"/>
4 vrste sira	<input type="radio"/>	<input checked="" type="radio"/>
povrtna	<input checked="" type="radio"/>	<input type="radio"/>
calzone	<input type="radio"/>	<input checked="" type="radio"/>

🔗 Kliknite na kružić (radio dugme) za odabir

Rezultate upitnika smo generirali računalom, s određenim parametrima slučajnosti, za 213 ispitanika.

Jednostavna analiza - prebrojavanje

Postoji više načina na koje možemo analizirati podatke MaxDiff metodom. Najjednostavniji pristup je prebrajanje koliko je puta neki proizvod odabran kao najbolji, odn. najlošiji. Pogledajmo proizvode po broju odabira najbolji/najlošiji:

alt	chosen_as_best	chosen_as_worst
4 godišnja doba	122	71
4 vrste sira	175	10
bolonjez	73	125
calzone	22	192
losos	116	57
lovačka	274	1
miješana	75	102
pikantna	165	18
povrtna	48	87
ribarska	39	83
rukola/pršut	96	65
s tunom	42	120
sa salamom	29	101
slavonska	2	246

Možemo vidjeti da je lovačka najviše puta odabrano kao najbolja opcija, a slavonska najviše puta kao najlošija opcija.

Analiza multinomnim logit modelom

Naprednija analiza uključuje modeliranje podataka (odgovora ispitanika) i zatim izračunavanje koeficijenata (“vrijednosti”) za svaki proizvod multinomnim logit modelom. Kad izvršimo analizu, dobijemo sljedeći sažetak:

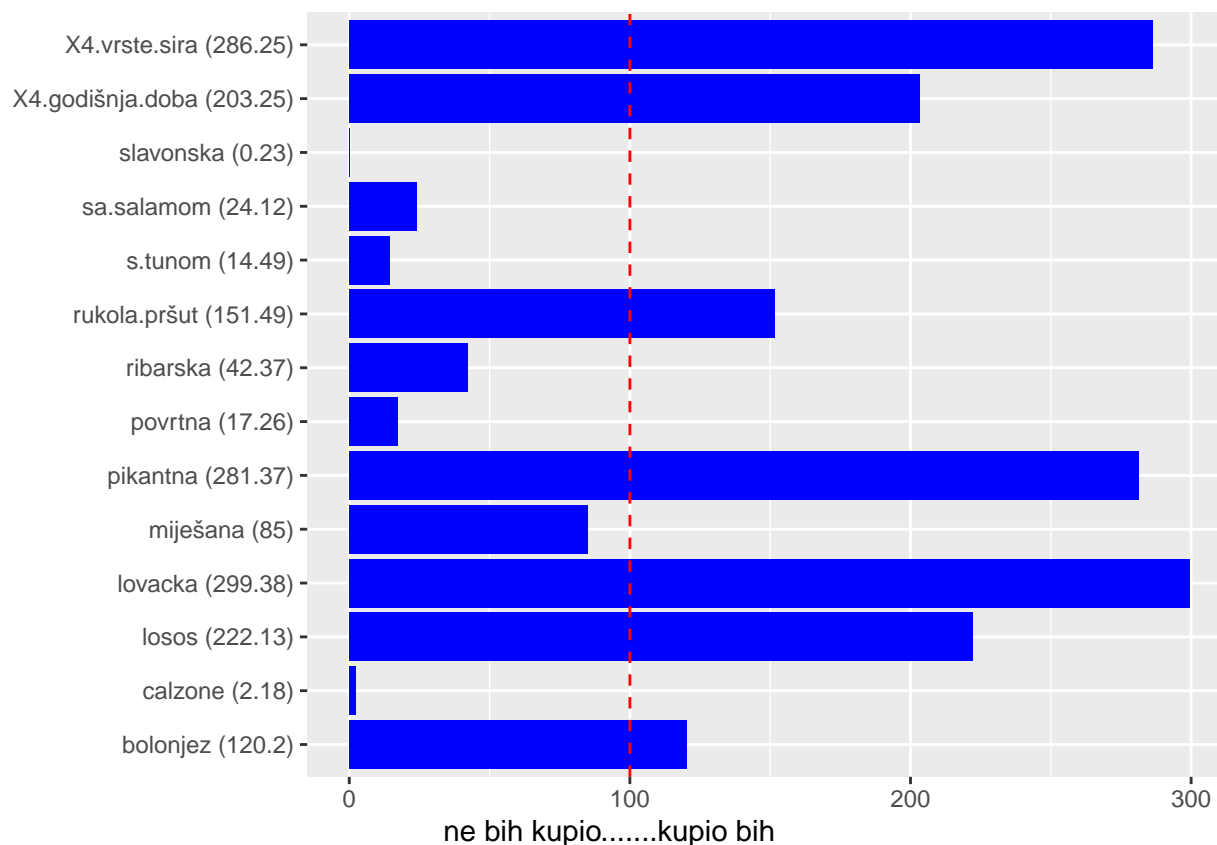
##	1st Qu.	Mean	3rd Qu.
## miješana	-3.0920891	-1.8280480	-0.5798886
## povrtna	-4.6731406	-3.6961489	-2.7276619
## pikantna	0.5397073	1.8147642	3.0274780
## X4.vrste.sira	1.0020402	2.1356078	3.2385703
## s.tunom	-5.0784475	-3.8806488	-2.6560078
## ribarska	-3.8026247	-2.7050260	-1.6177520
## X4.godišnja.doba	-1.2653653	-0.1577777	0.9236640
## calzone	-6.7619607	-5.8164645	-4.8554752
## slavonska	-9.1497087	-8.0749788	-6.9662801
## rukola.pršut	-2.0140358	-0.8801644	0.2392553
## bolonjez	-2.4283059	-1.3027477	-0.1888891
## lovačka	4.2132212	5.2759408	6.3156626
## losos	-0.9853540	0.1481920	1.2551358
## sa.salamom	-4.3412816	-3.3370907	-2.3248023

U prikazanom sažetku je dosta informacija koje nije tako lako direktno interpretirati. Ali da malo pojasnimo, evo nekoliko pojašnjenja koja slijede direktno iz sažetka:

- ‘calzone’ ima negativni (prosječni/mean) koeficijent -5.816, što znači da ga ispitanici slabije vrednuju nego osnovnu graničnu opciju koja označava točku kupio bih/ne bih kupio (usporedba koeficijenta za neku vrijednost je uvijek u odnosu na prvu/osnovnu deklariranu vrijednost, što je u ovom slučaju granična opcija) - jednostavnije rečeno ‘calzone’ ispitanici ne bi kupili, a zbog relativno velikog negativnog koeficijenta, ta je odluka prilično snažna,
- ‘lovačka’ i ‘losos’ obje imaju pozitivne (prosječne/mean) koeficijente (5.276, 0.148) što znači da ih ispitanici obje vrednuju više od osnovne granične opcije, ali ‘lovačka’ vrednuju više nego ‘losos’ - jednostavnije rečeno ispitanici bi kupili obje, ali ‘lovačka’ puno radije nego ‘losos’,

- prvi i treći kvantili označavaju raspon u kojem se, malo pojednostavljeno, nalazi svaki koeficijent sa 75% vjerojatnosti.

Ako sve koeficijente normaliziramo i skaliramo tako da granična opcija (kupio bih ili ne) bude na broju 100, lakše je interpretirati njihove vrijednosti i međusobne omjere:



Sad možemo vidjeti da je npr. najvredniji proizvod lovačka i da mu je vrijednost 1.35 puta veća od vrijednosti losos.

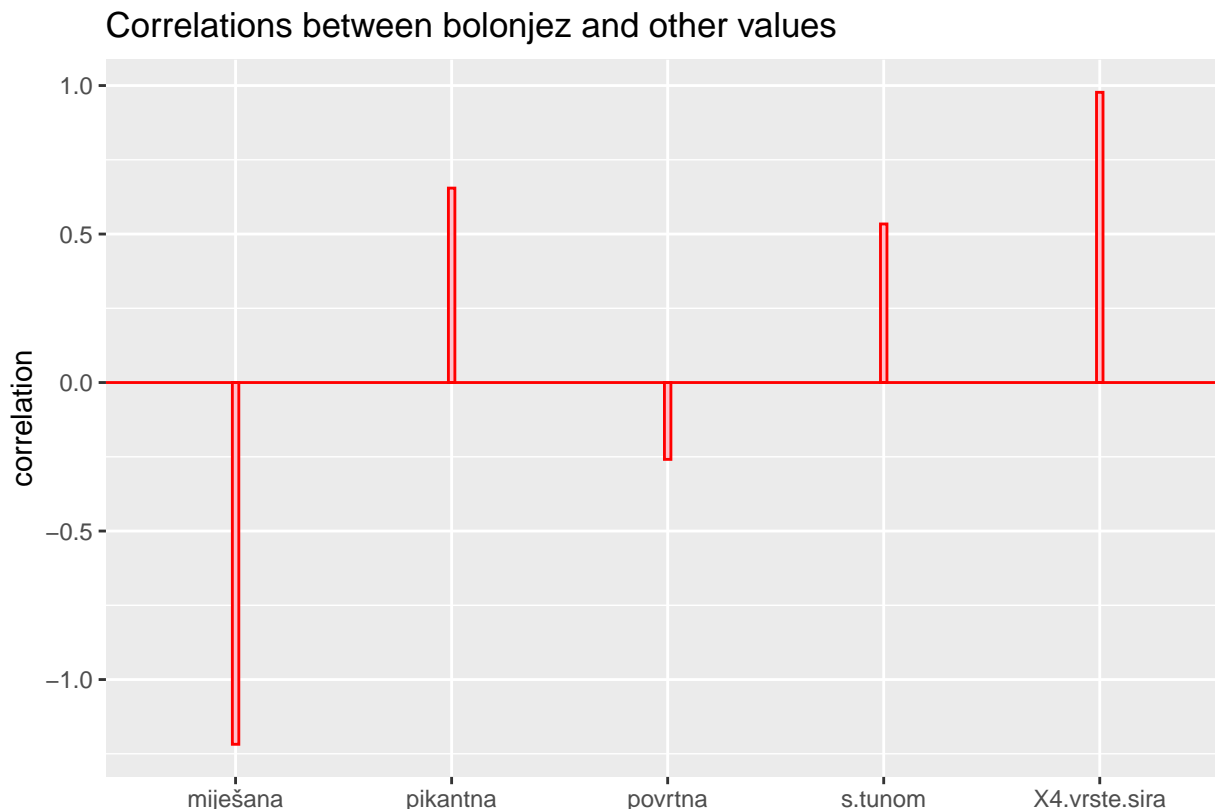
Ovi rezultati pokazuju i glavne prednosti MaxDiff analize u odnosu na jednostavno “redanje” proizvoda od najboljeg prema najlošijem koje se ponekad (prečesto :)) koristi za poslovne analize. MaxDiff nam pokazuje:

- *odnose među proizvodima* - koliko je koji vredniji ili manje vrijedan za ispitanike,
- *pozicije proizvoda u odnosu na (ključnu!) točku koja označava pozitivan ili negativan stav* - kupio bih ili ne, zanima me ili ne...

Time dobivamo kompletnu sliku svih proizvoda iz perspektive ispitanika i možemo donositi kvalitetne poslovne odluke.

Analiza koreliranosti proizvoda

Drugi element koji možemo analizirati je korelacija između vrijednosti pojedinih proizvoda.



Primjerice, u gore prikazanom grafu možemo vidjeti da je *bolonjez* pozitivno korelirana s *pikantna*, *X4.vrste.sira*, *s.tunom*, a negativno korelirana s *miješana*, *povrtna*. To znači da ako ispitanici daju veću vrijednost pizzi *bolonjez*, onda daju veću vrijednost i *pikantna*, *X4.vrste.sira*, *s.tunom*, dok za *miješana*, *povrtna* daju manju vrijednost.

Rezultate analize i procjene koeficijenata vrijednosti opcija najlakše je razumjeti preko “predviđanja udjela” (share prediction) što ćemo sada prikazati.

Predviđanje udjela (share prediction)

Jednom kada znamo vrijednosti koeficijenata svih proizvoda, možemo usporediti njihove različite kombinacije i predvidjeti udjele u prodaji. Predviđeni udjeli u prodaji su vjerojatnosti da će se određeni proizvod prodati u usporedbi s ostalim proizvodima koji su u kombinaciji koju testiramo.

Kao primjer uzet ćemo kombinaciju nekoliko slučajno odabranih proizvoda i izračunati njihove predviđene udjele u prodaji:

##	[,1]	[,2]	[,3]	[,4]	[,5]
## opcije	"pikantna"	"miješana"	"povrtna"	"s tunom"	"4 vrste sira"
## share %	"32.54"	"13.22"	"11.52"	"11.67"	"31.05"

U prethodnoj tablici možemo vidjeti kako bi ispitanici birali između 5 različitih opcija pizza. Svaka kombinacija dobiva predviđeni udio (postotak) u ukupnoj prodaji, prema procijenjenim vrijednostima pripadajućih koeficijenata.

Najčešće birana opcija (između navedenih, ne općenito) je *pikantna* s **predviđenim udjelom u prodaji 32.54%**.

Provjera modela na postojećim podacima

Korištenjem predviđanja udjela/vjerojatnosti prodaje za razne kombinacije proizvoda, možemo testirati i kako bi naš model predvidio izbor za (neke) kombinacije koje su dane u istraživanju:

```
##                      choice
## respondent/question true best predicted best true worst predicted worst
##           82/1      3      3      1      1
##           26/1      3      3      2      2
##          159/4      2      2      3      3
##          119/1      3      3      1      1
##          123/1      3      3      1      1
##          161/2      2      2      3      3
##          182/1      1      1      2      2
##          171/4      1      1      2      2
##           98/4      3      3      2      2
##          119/4      3      3      1      1
##          199/4      1      1      2      2
##           45/2      2      2      1      1
##          148/4      1      1      2      2
##          109/4      1      1      3      2
##          159/2      1      1      3      3
##          187/4      3      3      2      2
##           85/4      3      3      1      1
##          110/2      1      1      3      3
##          161/4      3      3      2      2
##           68/1      1      1      2      2

## [1] "Ukupno točno predviđanja:95.56%"
```

Ukupna točnost predviđanja daje nam sigurnost da naš model dobro funkcionira.

Segmentacija ispitanika

S obzirom na razne faktore (demografske ili sl.), ispitanike možemo podijeliti u segmente te zatim modelirati vrednovanje proizvoda za svaki segment. Broj segmenata treba zadati unaprijed. Pretpostavimo da postoje 3 različita segmenta ispitanika (s tim parametrom ćemo napraviti analizu). Modeliranjem (multinomni logit model s LCA - latent class analysis) dobivamo sljedeći sažetak:

```
##
##          segment_1 segment_2 segment_3
##   miješana    -0.421666752 -0.1928350 -0.23097801
##   povrtna     -1.023541601 -0.9933923 -0.93023531
##   pikantna     1.276050489  0.7594915  1.23078354
##   X4.vrste.sira 1.208666717  0.9930409  0.99388845
##   s.tunom      -1.070901256 -0.3983372 -0.60110445
##   ribarska     -0.834532977 -1.3050184 -1.28604449
##   X4.godišnja.doba 0.298467043  0.6136020  0.48009585
##   calzone      -2.135622549 -1.4997387 -1.59859198
##   slavonska    -3.141561217 -2.1989695 -2.12337713
##   rukola.pršut -0.002467373  0.2493077 -0.02685914
##   bolonjez     -0.259687192 -0.4068820 -0.73914492
##   lovačka      2.175849415  1.5153187  0.98609438
##   losos        0.436137449  0.7190122  0.60509715
##   sa.salamom   -0.947859577 -0.4861938 -0.71559553
```

Vidimo da se segmenti razlikuju (za više od 0.5) u vrednovanju npr. *pikantna, s.tunom, calzone, slavonska, lovačka*.

S obzirom na to da je cilj segmentacije i prepoznavanje koji korisnik (s najvećom vjerojatnošću) pripada kojem segmentu, pogledajmo i taj rezultat (navedeni su redni brojevi ispitanika koji pripadaju određenom segmentu):

```
## [1] "Segment 1"
## [1] 2 3 4 9 11 12 13 14 15 16 21 23 24 25 27 28 29 30
## [19] 34 35 36 37 38 40 41 42 44 48 49 53 55 58 61 65 67 69
## [37] 70 72 73 76 77 78 80 81 83 86 88 90 91 92 93 94 96 97
## [55] 98 99 101 104 106 108 109 110 113 115 116 117 120 121 122 123 124 125
## [73] 132 134 135 137 138 140 143 145 147 151 153 155 156 157 158 159 160 161
## [91] 162 163 164 165 166 169 172 175 178 180 181 182 183 187 194 196 197 198
## [109] 199 202 206 208 211 213
## [1] "Segment 2"
## [1] 1 5 6 7 8 10 17 18 19 20 22 26 31 32 33 39 43 45 46
## [20] 47 50 51 52 54 56 57 59 60 62 63 64 66 68 71 74 75 79 82
## [39] 84 85 87 89 95 100 102 103 105 107 111 112 114 118 119 126 127 128 129
## [58] 130 131 133 136 139 141 142 144 146 148 149 150 152 154 167 168 170 171 173
## [77] 174 176 177 179 184 185 186 188 189 190 191 192 193 195 200 201 203 204 205
## [96] 207 209 210 212
## [1] "Segment 3"
## integer(0)
```

S dobivenom segmentacijom možemo primjerice kreirati više različitih proizvoda od kojih će svaki biti najzanimljiviji određenom segmentu (s obzirom na vrednovanje).

TURF analiza

Ponudom u kojoj se nalazi više proizvoda povećat ćemo vjerojatnost da će barem neki proizvod biti kupljen od strane kupaca. Da bismo izračunali koji proizvodi bi trebali biti u ponudi da bismo postigli najveću prodaju, koristimo TURF (total unduplicated reach and frequency) analizu.

Pretpostavimo da u ponudu želimo staviti 3 proizvoda. Za svaku kombinaciju koja sadrži 3 proizvoda i za svakog ispitanika iz našeg istraživanja izračunat ćemo da li bi kupio barem jedan proizvod iz kombinacije (reach - da li je kupac 'dosegnut' barem nekim proizvodom) i koliko (ako i jedan) bi proizvoda iz kombinacije kupio (frequency - koliko je puta kupac 'dosegnut' danom ponudom proizvoda). Odluka da li bi kupac kupio neki proizvod ili ne se donosi na osnovi vjerojatnosti prodaje proizvoda u usporedbi s ostalim proizvodima iz ponude i osnovnom (graničnom) opcijom koja označava granicu kupio bih/ne bih kupio.

Pogledajmo kako izgleda nekoliko redaka iz reach & frequency tablice za naše proizvode (poredanih silazno po reach):

##	reach	frequency
## 4 vrste sira, lovačka, sa salamom	155	247
## pikantna, s tunom, lovačka	150	244
## miješana, ribarska, lovačka	147	235
## miješana, 4 godišnja doba, lovačka	147	231
## pikantna, 4 godišnja doba, lovačka	147	233
## s tunom, slavonska, lovačka	147	237

Možemo vidjeti da bi kombinaciju 4 vrste sira, lovačka, sa salamom kupilo najviše ispitanika, 155, što je 73%, s time što bi ukupno bilo 247 (potencijalnih) kupovina.

Na ovaj način možemo kreirati ponudu proizvoda kojom ćemo maksimizirati prodaju.

Zaključak

MaxDiff analiza je moćan alat za upoznavanje kupaca i njihovih vrijednosti i za kreiranje najboljih proizvoda za njih. Takav pristup uvijek vodi prema poboljšanju poslovanja, većim prihodima i većem zadovoljstvu kupaca.

Za više informacija kontaktirajte danijel.kopcinovic@itmarket.hr.