Multivariatna analiza –

2. domača naloga - popravljena

Slavko Žitnik, 63060254

Fakulteta za računalništvo in informatiko, Ljubljana

21.4.2011

# Problem in cilj naloge

V nalogi vzamemo odvisne spremenljivke in jih razvrščamo po enotah. Cilj je, da jih razvrstimo v primerno število skupin, ki jih razložimo. Uporabili bomo hierarhično združevanje, metodo voditeljev in metodo glavnih komponent. Najdene skupine bomo smiselno poimenovali in interpretirali rezultate.

# Opis podatkov

Podatke smo pridobili na portalu ESS (*European Social Survey*). Podatki se zbirajo za akademske namene v več kot 30 državah. Cilji anketiranj so spremljanje, interpretacija javnega mnenja in raziskovanje odnosa do spreminjajočih se evropskih institucij. Želijo izboljšati metode meddržavnega anketiranja v Evropi. Navsezadnje žeijo razviti tudi seznam socialnih vedenjskih vzorcev. V prvih 4 sklopih anketiranj je sodelovalo 34 držav.

Posamezna enota anketiranja predstavlja posamezno osebo. Vsi sodelujoči so stari nad 15 let, prebivalci, ki živijo v zasebnih gospodinjstvih, ne glede na njihovo nacionalnost, jezik ali pravni status.

Uporabili bomo rezultate anketiranj iz **Avstrije**, **Nemčije**, **Slovenije** v letu **2006/2007**. V Sloveniji je bila za anketiranje odgovorna Brina Malnar iz Univerze v Ljubljani in skupno 33 ljudi je izvajalo ankete. Več si lahko o teh podatkih preberemo v [3].

## Neodvisne spremenljivke

* **EDULVLA** (*Highest level of education*)

(\* Opomba: Mislim, da je slovenska verzija vprašanja neprimerno zastavljena!)

Vprašanje: Katera je zadnja šola, ki ste jo končali, redno ali izredno?

Možni odgovori:

|  |  |
| --- | --- |
| 0 | nedokončana osnovna šola |
| 1 | dokončana osnovna šola |
| 2 | 2-3 letna poklicna šola |
| 3 | splošna gimnazija, poklicna gimnazija, štiriletna strokovna šola |
| 4 | 2-letna višja (strokovna) šola |
| 5 | visoka šola, fakulteta, akademija |
| 6 | magisterij, doktorat |
| 8 | (ne vem) |

* **PDJOBEV** (*Ever had a paid job*)

Vprašanje: Ste imeli kdaj prej plačano delo?

Možni odgovori:

|  |  |
| --- | --- |
| 1 | da |
| 2 | ne |
| 8 | ne vem |

* **MARITALA** (*Legal marital status*) (\* Enako vprašanje kot MARITAL, vendar se je uporabljal le v ESS1, ESS2)

Vprašanje: Povejte kakšen je vaš sedanji uradni zakonski stan? kateri od naslednjih opisov najbolj ustreza vaši sedanji situaciji.

Možni odgovori:

|  |  |
| --- | --- |
| 01 | poročen |
| 02 | izvenzakonska skupnost |
| 03 | živi ločeno (a legalno še poročen) |
| 04 | živi ločeno (a legalno še vedno v izvenzakonski skupnosti) |
| 05 | ločen |
| 06 | vdovec, vdova |
| 07 | nekoč v izvenzakonski skupnosti, zdaj razpadla |
| 08 | nekoč v izvenzakonski skupnosti, partner umrl |
| 09 | nikoli ni bil poročen in nikoli živel v izvenzakonski skupnosti |
| 77 | ne želi odgovoriti |
| 88 | ne vem |

* **AGE** (*Age of respondent, calculated*)

Vprašanje: (\*Avtomatsko pridobljeno iz katerega leta ste bili rojeni / je bila rojena / je bil rojen?)

Možni odgovori:

[starost]

* **GNDR** (*Gender*)

Vprašanje: Spol

Možni odgovori:

|  |  |
| --- | --- |
| 1 | Moški |
| 2 | Ženski |
| 9 | ne želi odgovoriti |

* **CTZCNTR** (*Citizen of country*)

Vprašanje: Ali ste državljan(ka) Slovenije?

Možni odgovori:

|  |  |
| --- | --- |
| 1 | da |
| 2 | ne |
| 8 | ne vem |

* **HINCTNT** (*Household's total net income, all sources*)

Vprašanje: S pomočjo te kartice povejte, katera črka najbolje označuje skupni mesečni neto dohodek vašega gospodinjstva (neto – brez davkov in prispevkov). Seštejte dohodke iz vseh virov. Če ne veste natančno, povejte približno.

Možni odgovori:

|  |  |  |
| --- | --- | --- |
| 01 | J | manj kot 650€ |
| 02 | R | od 650€ do 800€ |
| 03 | C | od 800€ do 950€ |
| 04 | M | od 950€ do 1100€ |
| 05 | F | od 1100€ do 1300€ |
| 06 | S | od 1300€ do 1600€ |
| 07 | K | od 1600€ do 1900€ |
| 08 | P | od 1900€ do 2200€ |
| 09 | D | od 2200€ do 3000€ |
| 10 | H | 3000€ in več |
| 77 |  | zavrnil odgovor |
| 88 |  | ne vem |

## Odvisne spremenljivke

1. **POLINTR** How interested in politics

Vprašanje: Kaj bi rekli, koliko se zanimate za politiko?

Možni odgovori:

|  |  |
| --- | --- |
| 1 | zelo se zanimam |
| 2 | še kar se zanimam |
| 3 | komaj kaj se zanimam |
| 4 | sploh se ne zanimam |
| 8 | ne vem |

1. **TRSTPRL** Trust in country's parliament

Vprašanje: Prosim, da s pomočjo lestvice od 0 do 10 ocenite, koliko vi osebno zaupate vsaki od naslednjih ustanov. 0 pomeni, da ustanovi sploh ne zaupate, 10 pa, da ji povsem zaupate. Najprej... (državni zbor)

Možni odgovori:

|  |  |
| --- | --- |
| 00 | sploh ne zaupam |
| 01 |  |
| 02 |  |
| 03 |  |
| 04 |  |
| 05 |  |
| 06 |  |
| 07 |  |
| 08 |  |
| 09 |  |
| 10 | povsem zaupam |
| 88 | ne vem |

1. **TRSTLGL** Trust in the legal system

Vprašanje: Prosim, da s pomočjo lestvice od 0 do 10 ocenite, koliko vi osebno zaupate vsaki od naslednjih ustanov. 0 pomeni, da ustanovi sploh ne zaupate, 10 pa, da ji povsem zaupate. Najprej... (pravni sistem)

Možni odgovori:

|  |  |
| --- | --- |
| 00 | sploh ne zaupam |
| 01 |  |
| 02 |  |
| 03 |  |
| 04 |  |
| 05 |  |
| 06 |  |
| 07 |  |
| 08 |  |
| 09 |  |
| 10 | povsem zaupam |
| 88 | ne vem |

1. **MMBPRTY** Member of political party

Vprašanje: Ali ste član kakšne politične stranke?

Možni odgovori:

|  |  |
| --- | --- |
| 1 | da |
| 2 | ne |
| 8 | ne vem |

1. **TRSTEP** Trust in the European Parliament

Vprašanje: Prosim, da s pomočjo lestvice od 0 do 10 ocenite, koliko vi osebno zaupate vsaki od naslednjih ustanov. 0 pomeni, da ustanovi sploh ne zaupate, 10 pa, da ji povsem zaupate. Najprej... (evropski parlament)

Možni odgovori:

|  |  |
| --- | --- |
| 00 | sploh ne zaupam |
| 01 |  |
| 02 |  |
| 03 |  |
| 04 |  |
| 05 |  |
| 06 |  |
| 07 |  |
| 08 |  |
| 09 |  |
| 10 | povsem zaupam |
| 88 | ne vem |

1. **STFLIFE** How satisfied with life as a whole

Vprašanje: V celoti gledno, kako zadovoljni ste sedaj s svojim življenjem? Izberite ustrezno vrednost na lestvici od 0 do 10, kjer pomeni 0 levo in 10 desno?

Možni odgovori:

|  |  |
| --- | --- |
| 00 | izjemno nezadovoljen |
| 01 |  |
| 02 |  |
| 03 |  |
| 04 |  |
| 05 |  |
| 06 |  |
| 07 |  |
| 08 |  |
| 09 |  |
| 10 | izredno zadovoljen |
| 88 | ne vem |

1. **STFHLTH** State of health services in country nowadays

Vprašanje: In kakšno je po vašem mnenju splošno stanje zdravstva danes v Sloveniji?

Možni odgovori:

|  |  |
| --- | --- |
| 00 | izjemno slabo |
| 01 |  |
| 02 |  |
| 03 |  |
| 04 |  |
| 05 |  |
| 06 |  |
| 07 |  |
| 08 |  |
| 09 |  |
| 10 | izredno dobro |
| 88 | ne vem |

# Opis izbrane metode

Pred razvrščanjem bomo podatke normalizirali (standardizirali). V nadaljevanju bomo opisali tri glavne metode, ki jih bomo uporabili.

## Hierarhično združevanje

Metoda združuje od spodaj navzgor. Na začetku vsaka enota predstavlja svojo skupino. Potem združujemo najbljižje skupine glede na določeno razdaljo, dokler ne ostane le ena, v kateri so zajete vse enote. Rezultat lahko grafično predstavimo v obliki dendrograma, ki nam prikazuje hierarhijo skupin in nam olajša odločitev, da izberemo določeno število skupin.

Slabost te metode je, da je zaradi primerjav časovno zelo zahtevna, zato jo izvajamo nad manjšimi nabori podatkov.

Različne metode hierarhičnega združevanja se med seboj razlikujejo glede na izbrano razdaljo med skupinami. V nalogi bomo uporabili minimalno, maksimalno in Wardovo razdaljo.

## Metoda voditeljev

Metoda voditeljev zahteva, da ji poleg podatkov podamo še število želenih skupin. Metoda nato sama naključno ali kako drugače izbere začetne voditelje (enega za vsako skupino). Nato vsako novo enoto dodeli k določeni skupini ter izračuna nove centroide skupin.

Ta metoda je veliko hitrejša kot hierarhično združevanje, vendar moramo sami določiti število skupin, čeprav lahko še ne poznamo podatkov. Zaradi tega bomo v nalogi izvedli metodo voditeljev tudi na način, da bomo uporabili začetne centroide iz rezultata hierarhičnega razvrščanja in nato primerjali.

## Metoda glavnih komponent

Metoda glavnih komponent nam omogoča, da zmanjšamo razsežnost podakov in pri tem izgubimo čimmanj informacij. Pri tem poiščemo nekaj prvih komponent, ki nam pojasnijo dovolj velik del variabilnosti podatkov.

## Pomembni termini iz statističnega slovarja

Pomembni termini:

* hierarhična metoda – gl. hierarhično združevanje
* hierarhična razvrstitev – razvrstitev, ki jo sestavljajo skupine iz posameznih enot ter skupina z vsemi enotami z vsemi enotami in za vsak par skupin velja, da je ena skupina podmnožica druge skupine ali da je presek med njima prazna množica; sin. drevesna razvrstitev; prim. popolna razvrstitev
* metoda glavnih komponent – v multivariatni analizi metoda v faktorski analizi (uvedel jo je K. Pearson leta 1901), ki priredi opazovanim koreliranim spremenljivkam nove spremenljivke, imenovane glavne komponente, ki so linearne kombinacije opazovanih spremenljivk, med seboj neodvisne in urejene po pomembnosti, s tem da prva komponenta pojasni največji del vsote varianc opazovanih spremenljivk, vsaka nadaljna pa kvečjemu manjši
* glavna komponenta – spremenljivka, ki je ena od linearnih kombinacij opazovanih spremenljivk, med seboj neodvisnih in urejenih po pomembnosti; prim. metoda glavnih komponent
* metoda voditeljev – nehierarhična metoda razvrščanja v izbrano število skupin, pri kateri se najprej določi množica predstavnikov posameznih skupin-voditeljev, nato se priredi vsaka enota najbližjemu voditelju in izračuna težišča tako dobljenih skupin, ki so novi voditelji, postopek se ponavlja, dokler se nova množica voditeljev ne razlikuje od množice voditeljev, dobljene korak pred njo; sin. metoda oblakov, prim. metoda razvrščanja v skupine
* metoda združevanja: metoda, pri kateri se v vsakem koraku postopka združita dve skupini ali se združi več skupin v novo skupino, pri čemer ni treba vnaprej določiti števila skupin; sin. hierarhična metoda; prim. metoda razvrščanja v skupine

# Predstavitev rezultatov

## Hierarhično združevanje

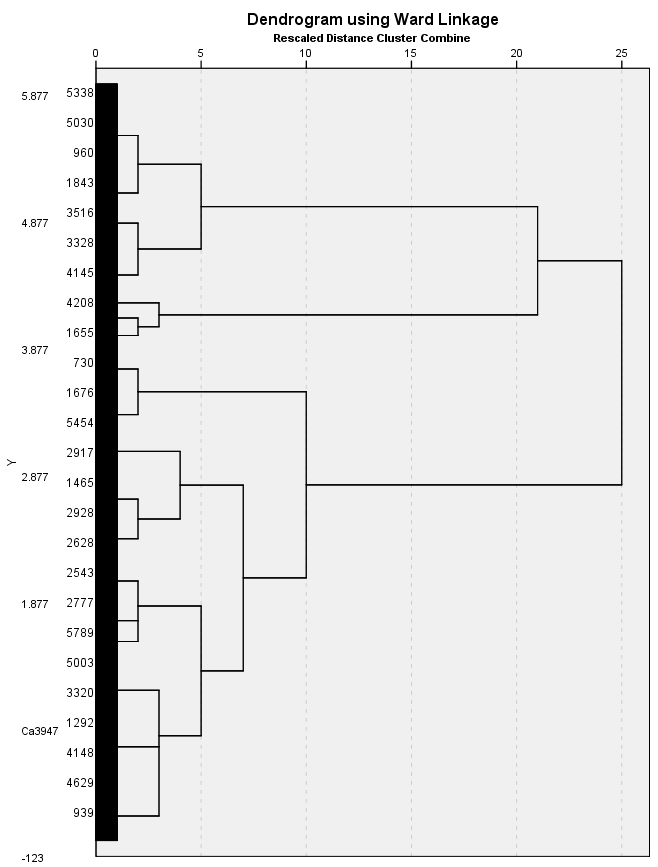
Pri združevanju bomo uporabili tri razdalje: Wardovo, minimalno in maksimalno. Neposredno pred združevanjem je metoda spremenljivke standardizirala.

Dendrogama za minimalno (Slika 2) in maksimalno razdaljo (Slika 3) se nahajata v prilogi, za Wardovo razdaljo pa na spodnji sliki (Slika 1).

Najslabša se v našem primeru izkaže minimalna metoda, kjer pride do veriženja in ne odrije nobene smiselne strukture. To je najbrž zaradi tega, ker skupine niso popolnoma disjunktne ali so enote bolj redko razporejene po prostoru.

Maksimalna metoda nam že nakaže bolj smiselno strukturo, čeprav so razlike v velikosti skupin večje.

Iz spodnjega dendrograma za Wardovo razdaljo lahko ugotovimo najbolj zanimivo strukturo v naših podatkih. Glede na velikost skupin se nam zdi smiselno, da nabor razdelimo nekje med 2 in 6 skupinami. Pri preskusu za 3, 4 in 6 skupin se glede na razlago odločimo, da je najbolj smiselno deliti podatke v tri skupine, ker jih lahko najlažje razložimo (tako je predlagala tudi prof. pri prvi oddaji naloge).



Slika 1: Dendrogram za Wardovo metodo združevanja

Iz Tabele 1 lahko vidimo, da se največ enot nahaja v skupini 2 in najmanj v skupini 3.

| **Ward Method** | | | | | |
| --- | --- | --- | --- | --- | --- |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 1 | 1659 | 24,4 | 27,8 | 27,8 |
| 2 | 3797 | 55,9 | 63,7 | 91,5 |
| 3 | 504 | 7,4 | 8,5 | 100,0 |
| Total | 5960 | 87,7 | 100,0 |  |
| Missing | System | 837 | 12,3 |  |  |
| Total | | 6797 | 100,0 |  |  |

Tabela 1: Velikost razdelitve v štiri skupine z Wardovo razdaljo

V tabeli 2 prikazujemo centroide za posamezne skupine. Na podlagi teh centroidov lahko razložimo tip enot v vsaki skupini in jih ustrezno poimenujemo.

|  | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Mean | | | | | | | |
| Skupine | How interested in politics | Trust in the legal system | Trust in country's parliament | Trust in the European Parliament | Member of political party | How satisfied with life as a whole | State of health services in country nowadays |
| 1 | 2,41 | 7,53 | 6,63 | 6,11 | 2,00 | 7,76 | 6,59 |
| 2 | 2,52 | 4,39 | 3,38 | 3,34 | 2,00 | 6,70 | 4,72 |
| 3 | 1,90 | 5,92 | 5,21 | 4,35 | 1,00 | 7,60 | 6,14 |
| Total | 2,44 | 5,39 | 4,44 | 4,20 | 1,92 | 7,07 | 5,36 |

Tabela 2: Centroidi pri 4 skupinah z Wardovo metodo

Sedaj ko imamo skupine in povprečne standardizirane vrednosti za vsak atribut, lahko določimo tip ljudi v posamezni skupini in jih poimenujemo.

V prvi skupini so ljudje, ki se bolj malo zanimajo za politiko in glede na ostale najbolj zaupajo v pravni sistem, državni in evropski parlament, niso člani političnih strank in so zadovoljni z življenjem in zdravstvenimi storitvami v državi. Ljudi v tej skupini bi imenovali ZAUPLJIVI.

V drugi skupini ljudi malenkost bolj zanima politika, vendar še vedno niso člani političnih strank. So bolj slabo zaupljivi političnim inštitucijam in slabše ocenjujejo stanje zdravstva, vendar so zadovoljni z življenjem. To skupino bi imenovali NEZAUPLJIVI.

V tretji skupini, ki je po številčnosti najmanjša so tipično ljudje bolj zainteresirani za politiko in so člani poltičnih strank. Srednje zaupajo v politične inštitucije, malo bolje ocenjujejo stanje zdravstva in so zadovoljni z življenjem. To skupino bi imenovali POVPREČNEŽI.

## Metoda voditeljev

Glede na rezultate pri hierarhičnem razvrščanju, bomo z metodo voditeljev razvrščali v 3 skupine.

Če si metoda voditeljev sama izbere začetne centroide, dobimo porazdelitev po skupinah, ki jo vidimo v tabeli 3. Če uporabimo začetne centroide iz rezultata hierarhičnega razvrščanja, dobimo skoraj enako porazdelitev po skupinah – tabela 4. Razlikuje se le v enem elementu. Torej lahko rečemo, da so naši podatki precej dobro ločeni v 3 skupine v prostoru. V prilogi smo navedli začetna in končna stanja centroidov za oba primera – v standardiziranih vrednostih. (Opomba: Pri pregledu podatkov moramo biti pozorni, da gledamo iste skupine v obeh primerih. V našem primeru glede na interpretacijo iz hierarhične metode ugotovimo, da so skupine obeh metod istoležne.)

Iz tabele 3 in 4 vidimo, da se podatki razdelijo v tri skupine tako, da jih je v prvi skupini največ in v tretji najmanj. Če to razporeditev primerjamo z razporeditvijo hierarhičnega razvrščanja, ugotovimo naslednje. Tretja skupina je ostala skoraj nespremenjena po številu. Precejšen del enot iz druge skupine se je pri metodi voditeljev preselilo v prvo skupino.

| **Number of Cases in each Cluster** | | |
| --- | --- | --- |
| Cluster | 1 | 2950,000 |
| 2 | 2508,000 |
| 3 | 502,000 |
| Valid | | 5960,000 |
| Missing | | 837,000 |

Tabela 3: Razporeditev v skupine pri naključno izbranih začetnih centroidih

| **Number of Cases in each Cluster** | | |
| --- | --- | --- |
| Cluster | 1 | 2949,000 |
| 2 | 2509,000 |
| 3 | 502,000 |
| Valid | | 5960,000 |
| Missing | | 837,000 |

Tabela 4: Razporeditev v skupine pri centroidih iz hierarhičnega razvrščanja

Zgoraj smo ugotovili, da se razporeditev enot po skupinah glede na hierarhično metodo razlikuje. Zaradi tega se bo mogoče spremenila tudi razlaga. V tabelah 5 in 6 so predstavljeni centroidi za vsako skupino. Če ju primerjamo s tabelo 2, na podlagi katere smo naredili razlago skupin iz rezultatov hierarhične metode, vidimo, da bi tudi v tem primeru sestavili enako razlago. Torej nam je metoda voditeljev malo drugače razdelila skupine, vendar so pomensko še vedno enako ločene.

|  | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Mean | | | | | | | |
| Skupina | How interested in politics | Trust in the legal system | Trust in country's parliament | Trust in the European Parliament | Member of political party | How satisfied with life as a whole | State of health services in country nowadays |
| 1 | 2,38 | 6,81 | 5,89 | 5,46 | 2,00 | 7,75 | 6,44 |
| 2 | 2,61 | 3,61 | 2,57 | 2,67 | 2,00 | 6,17 | 3,93 |
| 3 | 1,90 | 5,94 | 5,23 | 4,37 | 1,00 | 7,61 | 6,17 |
| Total | 2,44 | 5,39 | 4,44 | 4,20 | 1,92 | 7,07 | 5,36 |

Tabela 5: Končni centroidi pri uporabi začetnih centroidov iz hierarhične metode

|  | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Mean | | | | | | | |
| Skupina | How interested in politics | Trust in the legal system | Trust in country's parliament | Trust in the European Parliament | Member of political party | How satisfied with life as a whole | State of health services in country nowadays |
| 1 | 2,38 | 6,81 | 5,89 | 5,46 | 2,00 | 7,75 | 6,44 |
| 2 | 2,61 | 3,61 | 2,57 | 2,67 | 2,00 | 6,17 | 3,93 |
| 3 | 1,90 | 5,94 | 5,23 | 4,37 | 1,00 | 7,61 | 6,17 |
| Total | 2,44 | 5,39 | 4,44 | 4,20 | 1,92 | 7,07 | 5,36 |

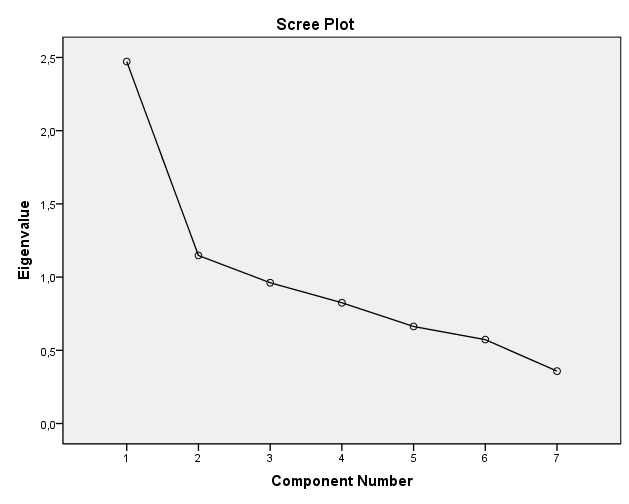
Tabela 6: Končni centroidi brez uporabe začetnih centroidov iz hierarhične metode

## Metoda glavnih komponent

Iz tabele 7 in slike 2 dve lahko opazimo, da prva komponenta očitno razloži več kot druge – t.j. 35% variance. Druga komponenta razloži še 16% variance. Pri nadaljnji analizi bomo upoštevali le ti dve komponenti.

| **Total Variance Explained** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | |
| Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 2,472 | 35,314 | 35,314 | 2,472 | 35,314 | 35,314 |
| 2 | 1,148 | 16,396 | 51,710 | 1,148 | 16,396 | 51,710 |
| 3 | ,962 | 13,736 | 65,446 |  |  |  |
| 4 | ,825 | 11,783 | 77,229 |  |  |  |
| 5 | ,663 | 9,471 | 86,700 |  |  |  |
| 6 | ,573 | 8,188 | 94,888 |  |  |  |
| 7 | ,358 | 5,112 | 100,000 |  |  |  |
|  | | | | | | |

Tabela 7: Delež varianc za posamezno komponento



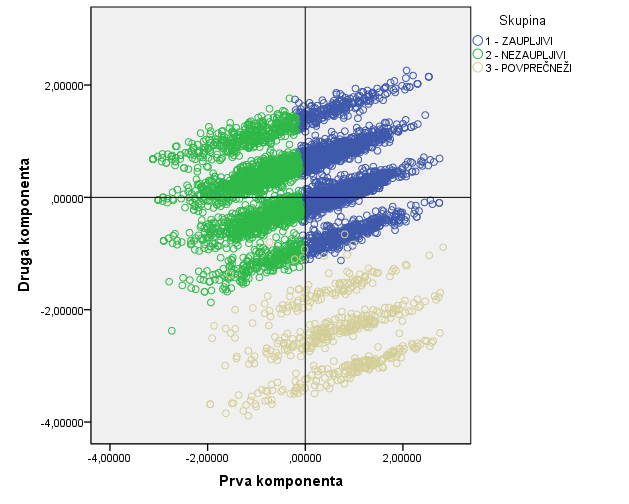
Slika 2: Scree diagram

| **Component Matrixa** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Component | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| How interested in politics | -,237 | ,**726** | ,325 | -,472 | ,279 | ,098 | ,040 |
| Trust in country's parliament | ,**836** | ,065 | -,242 | -,078 | ,043 | ,025 | ,479 |
| Trust in the legal system | ,**761** | ,108 | -,097 | -,051 | -,149 | ,567 | -,230 |
| Trust in the European Parliament | ,**700** | ,177 | -,391 | -,080 | ,323 | -,379 | -,267 |
| Member of political party | -,190 | ,**744** | -,229 | ,559 | -,210 | -,032 | ,025 |
| How satisfied with life as a whole | ,**491** | -,056 | ,595 | ,501 | ,386 | ,036 | ,005 |
| State of health services in country nowadays | ,**608** | ,133 | ,478 | -,150 | -,513 | -,309 | -,045 |
| Extraction Method: Principal Component Analysis. | | | | | | | |
|  | | | | | | | |

Tabela 8: Uteži posameznih komponent

V tabeli 8 smo odebelili pomembnejše uteži za prvi dve komponenti. Prva komponenta ima močne uteži pri zaupanju v politične inštitucije, zadovoljstvom nad življenjem in stanju zdravstvenih storitev v državi. Druga komponenta pa ima močne uteži pri pripadnosti posamezni stranki in zainteresiranosti v politiko.

Glede na razlago, ki smo jo naredili pri razvrščanju skupin sklepamo, da bo prva komponenta predvsem dobro ločila 1. in 2. skupino, t.j. ZAUPLJIVE in NEZAUPLJIVE. Druga utež pa bi morala od teh dveh skupin ločiti 3. skupino, t.j. POVPREČNEŽE, saj zanje za razliko od prvih dveh skupin velja, da so člani kakšne stranke in jih politika tudi bolj zanima.



Slika 3: Porazdelitev skupin iz hierarhičnega razvrščanja glede na prvi dve komponenti

Metoda glavnih komponent deluje tako, da išče takšne pravokotne osi, v smeri katerih je varianca največja. Na sliki 3 smo narisali prvi dve komponenti, s katerima bi lahko že dovolj dobro klasificirali primere v posamezne skupine. Slika 3 nazorno prikazuje in potrjuje našo domnevo, kako naj bi komponenti ločili skupine glede na sklepanja iz uteži.

# Interpretacija in evaluacija rezultatov

V prvem delu naloge smo nad izbranimi podatki uporabili metode za razvrščanje – hierarhično in metodo voditeljev. Ugotovili smo, da nam podatki najlepše pomensko razpadejo, če si izberemo razvrščanje v 3 skupine. Skupine smo glede na centroide skupin poimenovali ZAUPLJIVI, NEZAUPLJIVI in POVPREČNEŽI.

V drugem delu naloge smo izvedli še metodo glavnih komponent. Ugotovili smo, da prva komponenta pojasni 35% variance in druga 16%. Opisali in prikazali smo, kako dobro ti dve komponenti ločita skupine, ki smo jih dobili kot rezultat razvrščanja.

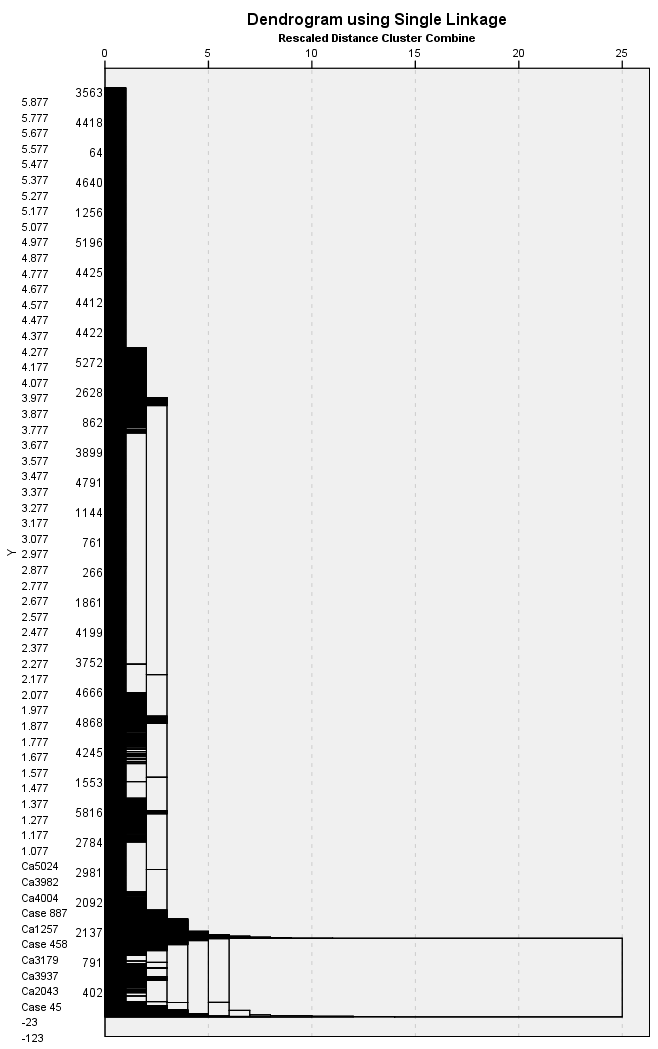
Bolj podrobna interpretacija in evaluacija rezultatov je podana že v prejšnjih razdelkih pri posameznih primerih.

# Uporabljena literatura

Prosojnice predmeta za multivariatno analizo in statistični slovar.

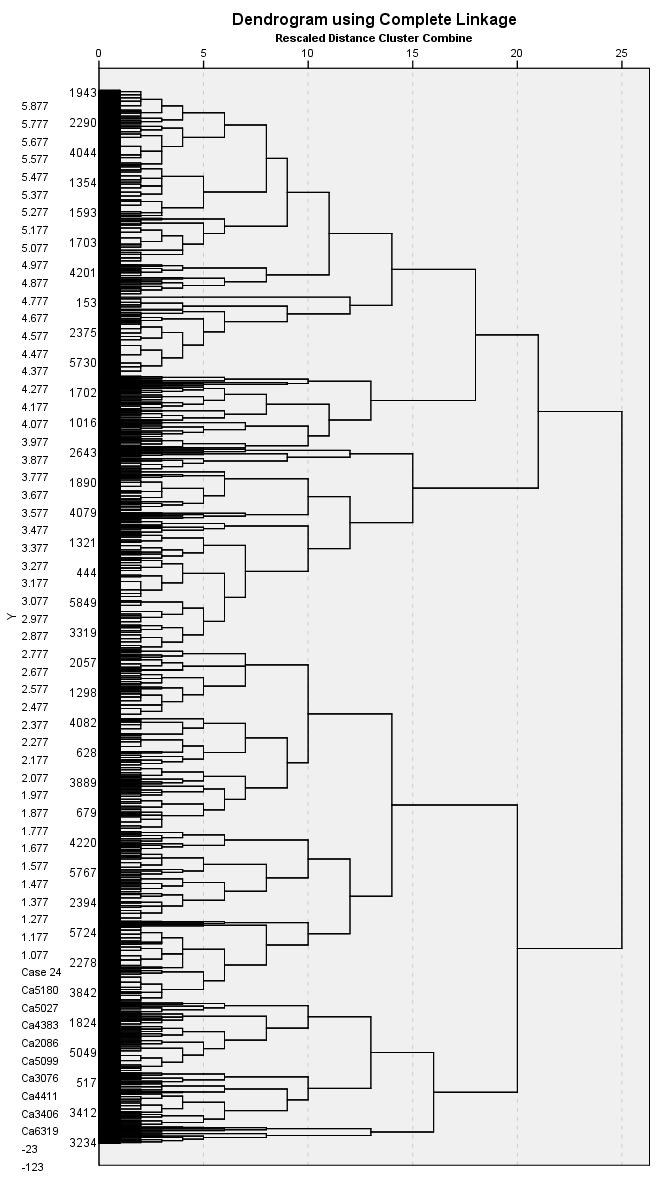
# Priloge

## Hierarhično združevanje – minimalna metoda



Slika 4:Dendrogram za minimalno metodo združevanja

## Hierarhično združevanje – maksimalna metoda



Slika 5: Dendrogram za maksimalno metodo združevanja

## Izbira centroidov pri uporabi metode voditeljev

| **Initial Cluster Centers** | | | |
| --- | --- | --- | --- |
|  | Cluster | | |
| 1 | 2 | 3 |
| Zscore: How interested in politics | ,57449 | 1,70625 | -1,68903 |
| Zscore: Trust in the legal system | 1,82039 | -2,11356 | -1,72016 |
| Zscore: Trust in country's parliament | -,99039 | -1,81609 | -,16470 |
| Zscore: Trust in the European Parliament | -1,29431 | -1,69969 | 2,35417 |
| Zscore: Member of political party | ,29404 | ,29404 | -3,40037 |
| Zscore: How satisfied with life as a whole | 1,33777 | -3,21090 | ,88291 |
| Zscore: State of health services in country nowadays | -2,11568 | 1,83366 | 1,43873 |

Tabela 9: Naključno izbrani začetni centroidi v metodi k-means

| **Final Cluster Centers** | | | |
| --- | --- | --- | --- |
|  | Cluster | | |
| 1 | 2 | 3 |
| Zscore: How interested in politics | -,12643 | ,13045 | -,67225 |
| Zscore: Trust in the legal system | ,56646 | -,69260 | ,22487 |
| Zscore: Trust in country's parliament | ,61580 | -,75599 | ,34437 |
| Zscore: Trust in the European Parliament | ,51468 | -,61608 | ,07125 |
| Zscore: Member of political party | ,29404 | ,29110 | -3,40037 |
| Zscore: How satisfied with life as a whole | ,31548 | -,40588 | ,25135 |
| Zscore: State of health services in country nowadays | ,42850 | -,56445 | ,31922 |

Tabela 10: Ciljni centroidi pri začetnih naključno izbranih

| **Initial Cluster Centers** | | | |
| --- | --- | --- | --- |
|  | Cluster | | |
| 1 | 2 | 3 |
| Zscore: How interested in politics | -,09338 | ,02813 | -,66506 |
| Zscore: Trust in the legal system | ,84841 | -,38695 | ,21559 |
| Zscore: Trust in country's parliament | ,92031 | -,42206 | ,33580 |
| Zscore: Trust in the European Parliament | ,77661 | -,34549 | ,06422 |
| Zscore: Member of political party | ,29404 | ,29404 | -3,40037 |
| Zscore: How satisfied with life as a whole | ,31974 | -,16220 | ,24663 |
| Zscore: State of health services in country nowadays | ,48579 | -,25105 | ,30956 |
|  | | | |

Tabela 11: Začetni centroidi iz hierarhične metode

| **Final Cluster Centers** | | | |
| --- | --- | --- | --- |
|  | Cluster | | |
| 1 | 2 | 3 |
| Zscore: How interested in politics | -,12629 | ,13018 | -,67225 |
| Zscore: Trust in the legal system | ,56684 | -,69254 | ,22487 |
| Zscore: Trust in country's parliament | ,61620 | -,75592 | ,34437 |
| Zscore: Trust in the European Parliament | ,51515 | -,61619 | ,07125 |
| Zscore: Member of political party | ,29404 | ,29110 | -3,40037 |
| Zscore: How satisfied with life as a whole | ,31513 | -,40519 | ,25135 |
| Zscore: State of health services in country nowadays | ,42829 | -,56381 | ,31922 |

Tabela 12: Končni centroidi pri uporabi začetnih centroidov iz hierarhične metode