**TOMŠIČ GAŠPER**

**21150317**

**REDNI ŠTUDIJ**

**DI-UNI-1**

**2P**

Univerza

*v Ljubljani*

Fakulteta

*za družbene vede*



1. **STATISTIČNA NALOGA**

**MULTIVARIATNA ANALIZA**

**NEMČIJA**

**Študijsko leto 2016/2017**

**Nosilec: izr. prof. dr. Aleš Žiberna**

**Tomšič Gašper, 21150317, gasper.sola@gmail.com**

**Družboslovna Informatika**

**Redni študij, Družboslovna Informatika UNI, 2P**

**KONČNA OCENA 2. NALOGE: \_\_\_\_\_\_\_\_\_\_**



**IZJAVA O AVTORSTVU**

statistične naloge

Spodaj podpisani/-a **Gašper Tomšič**, z vpisno številko 21150317, sem avtor/-ica oddane statistične naloge pri predmetu Statistika v študijskem letu 2015/16.

S svojim podpisom zagotavljam, da:

* je predložena statistična naloga izključno rezultat mojega lastnega dela;
* sem poskrbel/-a, da so dela in mnenja drugih avtorjev oz. avtoric, ki jih uporabljam v predloženem besedilu, navedena oz. citirana v skladu fakultetnimi navodili ter navedena v seznamu virov, ki je sestavni element tega besedila;
* se zavedam, da je plagiatorstvo – predstavljanje tujih del, v katerem so tuji izračuni, interpretacije oz. ideje predstavljane kot moje lastne – kaznivo po zakonu (Zakon o avtorstvu in sorodnih pravicah, Uradni list RS št. 21/95), prekršek pa podleže tudi ukrepom Fakultete za družbene vede v skladu z njenimi pravili;
* se zavedam posledic, ki jih dokazano plagiatorstvo lahko predstavlja za predloženo delo in za moj status na Fakulteti za družbene vede.

V Ljubljani, dne \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Podpis avtorja/-ice \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Osnovne informacije o podatkovni datoteki in izbranih spremenljivkah**

**Podatkovna datoteka:**

World values survey – Nemčija

**Izbrane spremenljivke:**

Justifiable v198-v210 (brez v201 in v207A)

Spol v240

Starost v242

Izobrazba v248

Zaposlen v229

**Osnovne informacije o podatkovni datoteki in izbranih spremenljivkah**

**Podatkovna datoteka:**

World values survey – Nemčija

**Izbrane spremenljivke:**

Justifiable v198-v210 (brez v201 in v207A)

Spol v240

Starost v242

Izobrazba v248

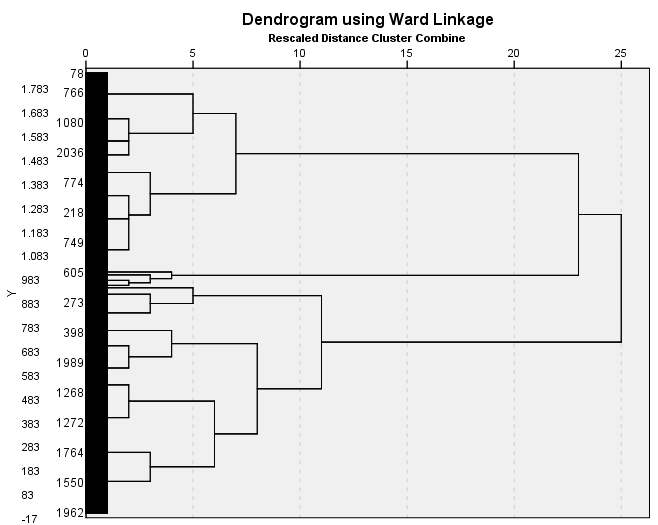
Zaposlen v229

**HIEARHIČNO RAZVRŠČANJE:**

**Hierarhična razvrstitev po Wardovi metodi z kvadrirano evklidsko razdaljo:**

Za razvrščanje smo uporabili Wardovo metodo z kvadrirano evklidsko razdaljo. Na sliki 1 se lepo vidi, da bo najbolj smiselno uporabiti 3 skupine, saj je tukaj največji preskok(razdalja) med skupinami. Videti je, da so nastale 2 nekoliko večje skupine in ena malce manjša. Iz tabele 1 je razvidno, da smo razvrščali na 13ih spremenljivkah, Wardova kriterijska funkcija za razvrščanje po Wardovi metodi z kvadrirano evklidsko razdaljo je 18042,371.

**Slika 1: Hierarhična razvrstitev po Wardovi metodi z kvadrirano evklidsko razdaljo**

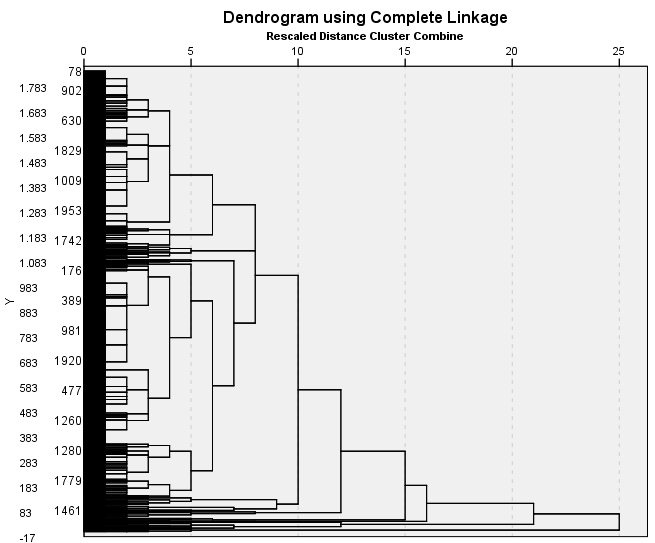


|  |  |  |
| --- | --- | --- |
| **Tabela 1: Wardova kriterijska funkcija pri Wardovi metodi z evklidskim izračunom razdalje**  **Descriptive Statistics** | | |
|  | N | Sum |
| SumofSquares Sum of Squares | 13 | 18042,371 |
| Valid N (listwise) | 13 |  |

**Hiearhična razvrstitev po metodi oddaljenega soseda z kvadirano evklidsko razdaljo**

Za naslednje razvrščanje smo uporabili metodo oddaljenega soseda z kvadrirano evklidsko razdaljo. Iz slike 2 sicer ni najbolje razvidno, vendar bi bilo najbolj smiselno uporabiti 3, saj je tukaj največji preskok. Smiselno bi bilo lahko bilo tudi vzeti dve skupini. Sicer pa je precej slabo razvrščanje, saj je nastala 1 skupina, ki zajema skoraj vse enote in dve skupine z zelo malo enotami. Iz tabele 2 je razvidno, da smo razvrščali na 13ih spremenljivkah, Wardova kriterijska funkcija za razvrščanje po metodi oddaljenega soseda z kvadrirano evklidsko razdaljo je 21281,593.

**Slika 2: Hiearhična razvrstitev po metodi oddaljenega soseda z kvadirano evklidsko razdaljo**



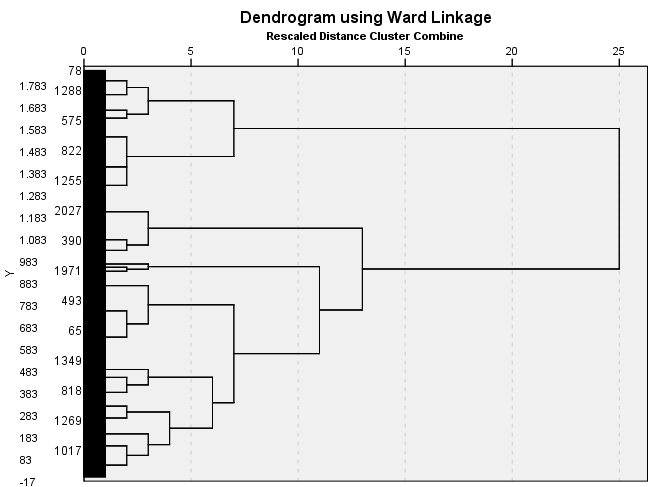
**Tabela 2: Wardova kriterijska funkcija pri maksimalni metodi z evklidskim izračunom razdalje**

|  |  |  |
| --- | --- | --- |
| **Descriptive Statistics** | | |
|  | N | Sum |
| SumofSquares Sum of Squares | 13 | 21281,593 |
| Valid N (listwise) | 13 |  |

**Hierarhična razvrstitev po Wardovi metodi z Manhattan izračuni razdalje**

Za naslednje razvrščanje smo uporabili Wardovo metodo z Manhattan izračunom razdalje. Iz slike 3 je razvidno, da je najbolj smiselno uporabiti dve skupini saj je tukaj največji preskok med skupinami. Pri razvrščanju je nastala ena večja skupina in ena manjša skupina. Iz tabele 3 je razvidno, da smo razvrščali na 13ih spremenljivkah, Wardova kriterijska funkcija za razvrščanje po Wardovi metodi z Manhattan izračunom razdalje je 20819,783.

**Slika 3: Hierarhična razvrstitev po Wardovi metodi z Manhattan izračuni razdalje**



**Tabela 3: Wardova kriterijska funkcija pri Wardovi metodi z manhattan izračunom razdalje**

|  |  |  |
| --- | --- | --- |
| **Descriptive Statistics** | | |
|  | N | Sum |
| SumofSquares Sum of Squares | 13 | 20819,783 |
| Valid N (listwise) | 13 |  |

**Izbor najboljšega razvrščanja**: Za nadaljnjo analizo bomo uporabili razvrščanje po Wardovi metodi z kvadrirano evklidsko razdaljo, ker ima najmanjšo Wardovo kriterijsko funkcijo, poimenovana bo CLU3wardSEk3.

**METODA VODITELJEV**

Za razvrščanje po metodi voditeljev (k-means) smo uporabili makro »qcluMacroRnd« in sicer smo izbrali 3 skupine in 100 ponovitev. V tabeli 8 je 8 najbolj optimalnih razvrstitev po metodi voditeljev (od 100ih poskusov). Vidimo, da smo razvrstitev delali na 1847 enotah. Najboljša razvrstitev je kar privzeta različica SPSS-a, KMDISTSQopt z Wardovo kriterjsko funkcijo 16735,35. Za nadaljnjo analizo bomo uporabili KMCLUopt.

**Tabela 4: Razvrstitev skupin po vrednosti Wardove kriterjske funkcije**

|  |  |  |
| --- | --- | --- |
| **Descriptive Statistics** | | |
|  | N | Sum |
| KMDISTSQopt | 1847 | 16735,35 |
| KMDISTSQ79 | 1847 | 16735,35 |
| KMDISTSQ50 | 1847 | 16735,45 |
| KMDISTSQ32 | 1847 | 16735,45 |
| KMDISTSQ60 | 1847 | 16735,56 |
| KMDISTSQ31 | 1847 | 16735,56 |
| KMDISTSQ78 | 1847 | 16735,57 |
| KMDISTSQ4 | 1847 | 16735,57 |

**PRIMERJAVA RAZBITIJ**

**Izbor najboljšega razbitja:** Naredili smo razvrščanja po štirih različnih metodah, ki so bile opisane v prejšnjih dveh poglavjih. Najboljšo razbitje smo dobili po metodi voditeljev (k-means) z Wardovo kriterijsko funkcijo 16735,35. Drugo najboljšo razbitje smo dobili po Wardovi metodi z kvadriranim evklidskim izračunom razdalj z Wardovo kriterjsko funkcijo 18042,371.

**Primerjava dveh najboljših razbitij**

V tabeli 5 lahko vidimo pripadnost enot dvema najboljšima razvrstitvama. Če bi bile razvrstitve popolnoma enake pri obeh metodah, bi bila razmerja skrajnih vrednosti enaka (vendar ne nujno v istih skupinah, ker so to le naključne označbe). Pri našem primeru vidimo, da sta si razvrstitvi precej podobni. Vidimo, da so pri obeh metodah nastale dve zelo veliki skupini in ena zelo majhna skupina. Vidimo, da je prva skupina po Wardu (n = 831) zelo podobna tretji skupini po metodi voditeljev (n = 821), ujema se 681 enot. Druga skupina po Wardu (n = 952) je zelo podobna prvi skupini po metodi voditeljev (n = 920), ujema se 779 enot. Tudi najmanjši skupini se precej ujemajo, tretja skupina po Wardu je zelo podobna drugi skupini po metodi voditeljev, ujema se 60 enot od 64.

**Tabela 5: Pripadnost skupinam narejenih po Wardu in metodi voditeljev**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Ward, kvadratna evklidska, k = 3 \* Metoda voditeljev Crosstabulation** | | | | | |
| Count | | | | | |
|  | | Metoda voditeljev | | | Total |
| 1 | 2 | 3 |
| Ward, kvadratna evklidska, k = 3 | 1 | 141 | 9 | 681 | 831 |
| 2 | 779 | 37 | 136 | 952 |
| 3 | 0 | 60 | 4 | 64 |
| Total | | 920 | 106 | 821 | 1847 |

Zgornjo interpretacijo potrjujejo tudi izpis Randovega indeksa (0,735) in popravljenega Randovega indeksa (0,466). Iz teh vrednosti lahko rečemo, da sta razvrstitvi srednje močno povezane.

**Izpis 1: Randov indeks za primerjavo dveh skupin CLU3wardSEk3 (wardova metoda) KMCLUopt (metoda voditeljev)**

Run MATRIX procedure:

Rand Index

,7347025806

adjusted Rand Index

,4660084942

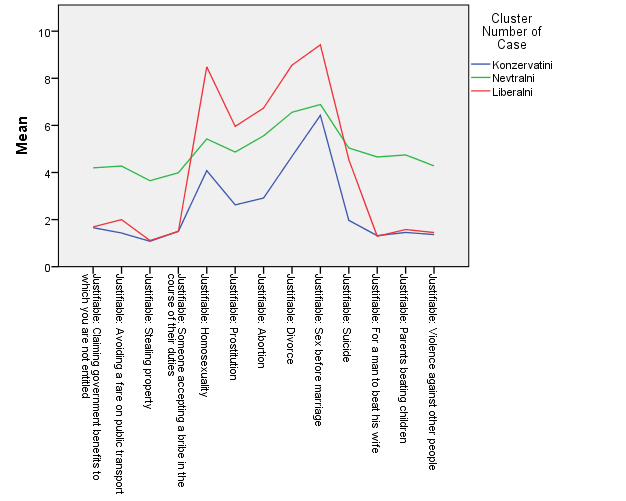
------ END MATRIX -----

**PREDSTAVITEV REZULTATOV, POVEZANOST, INTERPRETACIJA**

Najboljše razbitje je bilo po metodi voditeljev zato bomo za te skupine nadaljnje interpretirali rezultate.

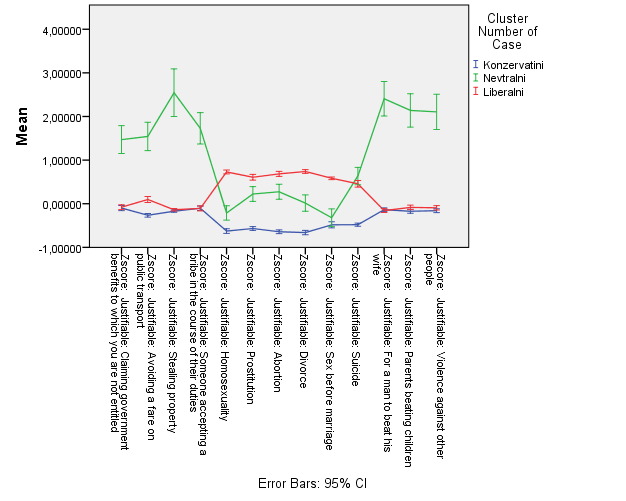
Na sliki 4 je prikaz povprečij skupin po posameznih tematskih spremenljivkah. Možno je videti razlike med skupinami. Pri sklopih materialno kaznivih dejanj in odnosu so glavne dve skupini zelo podobni, izjema je izjemno majhna skupina. Prvo skupino smo poimenovali konzervativni, njihovo povprečje je pri materialno kaznivih dejanjih in pri odnosu do nasilja zelo nizek, zelo opazno pa je da so neopravičljivi do liberalnih dejanj (z izjemo spolnega odnosa pred poroko), zaradi katerega smo jim dodelili tako ime. Drugo večjo skupino smo poimenovali liberalni (prej skupina 3), kot pri konzervativni skupini je povprečje pri materialno kaznivih dejanjih in odnosu do nasilja zelo nizek. Opazna pa je zelo velika opravičljivost liberalnih dejanj. Najmanjšo skupino smo poimenovali nevtralni (prej skupina 2), ki je zelo neopredeljena pri vseh treh sklopih, sicer se pa vidi, da je podobna ostalim dvema skupinam – nekoliko bolje opravičujejo liberalna dejanja, manj pa materialno kazniva dejanja in odnos do nasilja. Na splošno je razvidno, da vse skupine najbolj opravičujejo spolnost pred poroko ter homoseksualnost, najmanj krajo lastnine in da mož tepe ženo. Precej zanimivi sta tudi spremenljivki opravičljivost prostitucije in splava, ki je pri vseh skupinah v liberalnem sklopu ocenjen najslabše.

**Slika 4: Linijski grafikon za temaskte spremenljivke razvrščene v skupine po metodi voditeljev**



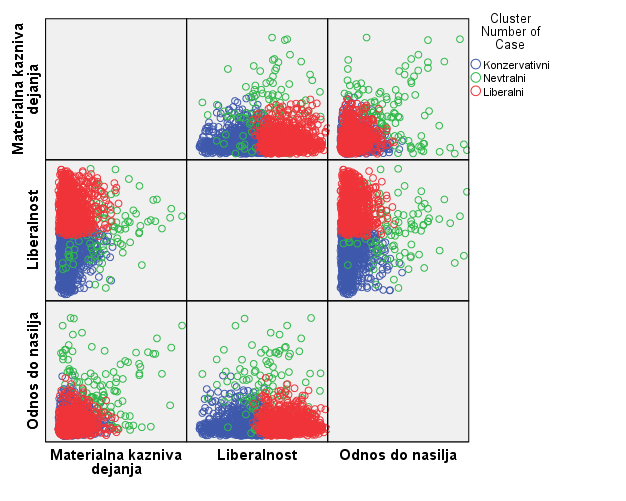
Na sliki 5 so standardizirane vrednosti skupin po posameznih tematskih spremenljivkah. Videti je, da se liberalna in konzervativno skupina po materialno kaznivih dejanjih ne razlikujeta, izjema je upravičljivost izogibanja plačevanja javnega prevoza. Kot smo že videli na sliki 5 je razlika med skupinama pri sklopu liberalnih dejanj. Intervali zaupanja so pri obeh skupinah zelo majhni, kar pomeni, da so vse enote znotraj skupin odgovarjale precej podobno. Skupina nevtralni se v vseh sklopih precej razlikuje od ostalih dveh skupin. Intervali zaupanja so precej bolj široki pri skupini nevtralnih, enote znotraj skupin so odgovarjale precej različno (razprešeno).

**Slika 5: Linijski grafikon za standardizirane tematske spremenljivke razvrščene v skupine po metodi voditeljev**



Podobno vidimo tudi na sliki 6, kjer vidimo, da skorajda ni razlik med konzervativno in liberalno skupino pri sklopu materialno kazniva dejanja in odnosu do nasilja. Edine razlike med skupinama so razvidne po osi, kjer je sklop liberalnosti. Iz slike 6 je tudi razvidno zelo velika razpršenost nevtralne skupine, kar se prikazuje na sliki 5 v obliki zelo široki intervalov zaupanja.

**Slika 6: Razsevni grafikoni za likartove spremenljivke razdeljene po skupinah**



V tabeli 6 je vidna porazdeljenost skupin po spolu. Pri liberalni in konzervativni skupini so razlike v porazdeljenosti spolov zelo majhne. Moških je v obeh skupinah nekoliko manj (48,2% - konzervativna, 47,9% - liberalna). Je pa nekoliko večja razlika pri nevtralni skupini, kjer več enot moškega spola (66,0%) kot enot ženskega spola (34,0%).

**Interpretacija hi-kvadrat testa:**

Vrednost hi-kvadrata je 12,996 pri dveh prostorskih stopnjah, natančna stopnja značilnosti p = 0,002. Ker je natančna stopnja značilnosti p manjša od 0,05, ničelno domnevo pri 5% stopnji značilnost zavrnemo. Z 5% gotovostjo lahko trdimo, da na populaciji obstaja statistično značilna povezanost med spolom in porazdeljenostjo v skupinah. Iz tabele 6 smo že videli, da je žensk nekoliko več v konzervativni in liberalni skupini, moških pa je več v nevtralni skupini.

**Interpretacija Cramerjevega koeficienta**:

Kramerjev koeficient (0,084) kaže na zelo šibko moč povezanosti med spremenljivkama.

**Tabela 6: Porazdeljenost skupin po spolu**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **V240 Sex \* KMCLUopt Cluster Number of Case Crosstabulation** | | | | | | |
|  | | | KMCLUopt Cluster Number of Case | | | Total |
| Konzervativni | Nevtralni | Liberalni |
| V240 Sex | 1 Male | Count | 443 | 70 | 393 | 906 |
| % within KMCLUopt Cluster Number of Case | 48,2% | 66,0% | 47,9% | 49,1% |
| 2 Female | Count | 477 | 36 | 428 | 941 |
| % within KMCLUopt Cluster Number of Case | 51,8% | 34,0% | 52,1% | 50,9% |
| Total | | Count | 920 | 106 | 821 | 1847 |
| % within KMCLUopt Cluster Number of Case | 100,0% | 100,0% | 100,0% | 100,0% |

|  |  |  |  |
| --- | --- | --- | --- |
| **Tabela 7: Pearsonov hi-kvadrat**  **Chi-Square Tests** | | | |
|  | Value | df | Asymp. Sig. (2-sided) |
| Pearson Chi-Square | 12,996a | 2 | ,002 |
| Likelihood Ratio | 13,187 | 2 | ,001 |
| Linear-by-Linear Association | ,005 | 1 | ,945 |
| N of Valid Cases | 1847 |  |  |
| a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 52,00. | | | |

**Tabela 8: Kramerjev koeficient**

|  |  |  |  |
| --- | --- | --- | --- |
| **Symmetric Measures** | | | |
|  | | Value | Approx. Sig. |
| Nominal by Nominal | Phi | ,084 | ,002 |
| Cramer's V | ,084 | ,002 |
| N of Valid Cases | | 1847 |  |

**POVZETEK**

V tej analizi smo razvrščali enote po štirih različnih metodah. 3 metode so bile hierarhične in sicer metoda po Wardu z kvadratnim evklidskim izračunom razdalj, metoda oddaljenega soseda z kvadratnim evklidskim izračunom razdalj in metoda po Wardu z manhattan izračunom razdalj. Zadnja metoda, ki smo jo uporabili je bila metoda voditeljev, ki smo jo ponovili 100x vsakič z novimi začetnimi enotami. Na koncu se je izkazalo, da je metoda voditeljev imela najmanjšo Wardovo kriterjsko funckijo (16735,35), drugo najmanjšo pa Wardova metoda kvadratnim evklidskim izračunom razdalj (18042,371). Za najboljši dve metodi smo preverili povezanost porazdeljenosti med skupinami s pomočjo kontingenčne tabele in randovega indeksa. Izkazalo se je, da je so bile obe metodi dokaj podobno razdeljene. Popravljeni randov indeks je bil 0,466, kar nakazuje na srednje močno povezanost.

Skupine dobljene iz najboljše metode smo poimenovali in sicer dve največji skupini sta dobili ime konzervativna in liberalna skupina. Najmanjša skupina pa je dobila ime nevtralni. Skupine smo primerjali po treh tematskih sklopih. Med liberalno in konzervativno skupino ni bilo nobene razlike pri opravičevanju materialno kaznivih dejanj in nasilju, medtem ko je bila zelo velika razlika pri opravičevanju liberalnih dejanj, kjer je liberalna skupina precej bolj upravičevala liberalna dejanja kot konzervativna. Med liberalnim sklopom je bila pri vsek skupinah precej manj upravičevanja prostitucije in splava. Pri povezanosti med skupinami in spolom smo ugotovili, da je statistično značilna povezanost in sicer je več žensk v konzervativni in liberalni skupini, medtem ko je več moških v skupini nevtralnih.

**SINTAKSA:**

**Branje makrota:**

INCLUDE FILE='C:\Users\predavalnica24\Desktop\MULTIVAR\makriRazvrscanje.sps'.

ScaleListwise V198 TO V210.

**Hierarhično razvrščanje po wardovi metodi z kvadrirano evklidsko razadaljjo**

CLUSTER ZV198 ZV199 ZV200 ZV202 ZV203 ZV203A ZV204 ZV205 ZV206 ZV207 ZV208 ZV209 ZV210

/METHOD WARD

/MEASURE=SEUCLID

/PLOT DENDROGRAM

/PRINT NONE

/SAVE CLUSTER(3).

**Hiearhično razvrščanje po metodi najdaljšega soseda z kvadrirano evklidsko razdaljo**

\*Furthest neighbour.

CLUSTER ZV198 ZV199 ZV200 ZV202 ZV203 ZV203A ZV204 ZV205 ZV206 ZV207 ZV208 ZV209 ZV210

/METHOD COMPLETE

/MEASURE=SEUCLID

/PLOT DENDROGRAM

/PRINT NONE.

/SAVE CLUSTER(3).

**Hiearhično razvrščanje po Wardovi metodi z Manhattan izračunom razdalje**

CLUSTER ZV198 ZV199 ZV200 ZV202 ZV203 ZV203A ZV204 ZV205 ZV206 ZV207 ZV208 ZV209 ZV210

/METHOD WARD

/MEASURE=BLOCK

/PLOT DENDROGRAM

/PRINT NONE.

/SAVE CLUSTER(2).

**Wardove kriterisjke funckije za vse tri možne razvrstitve**

WardCF group=CLU3wardSEk3 varlist= ZV198 ZV199 ZV200 ZV202 ZV203 ZV203A ZV204 ZV205 ZV206 ZV207 ZV208 ZV209 ZV210.

WardCF group= CLU3maxSEk3 varlist= ZV198 ZV199 ZV200 ZV202 ZV203 ZV203A ZV204 ZV205 ZV206 ZV207 ZV208 ZV209 ZV210.

WardCF group= CLU2wardBLk2 varlist= ZV198 ZV199 ZV200 ZV202 ZV203 ZV203A ZV204 ZV205 ZV206 ZV207 ZV208 ZV209 ZV210.

**Generiranje razvrstitev 100x in izpis tabele kriterijskih funkcij**

qcluMacroRnd from = 1 till = 100 k=3 varlist= ZV198 ZV199 ZV200 ZV202 ZV203 ZV203A ZV204 ZV205 ZV206 ZV207 ZV208 ZV209 ZV210/.

**Linijski grafikon za tematatske spremenljivke razvrščene po metodi voditeljev**

\* Chart Builder.

GGRAPH

/GRAPHDATASET NAME="graphdataset" VARIABLES=MEAN(V198) MEAN(V199) MEAN(V200) MEAN(V202) MEAN(V203) MEAN(V203A) MEAN(V204) MEAN(V205) MEAN(V206) MEAN(V207) MEAN(V208) MEAN(V209) MEAN(V210) KMCLUopt MISSING=LISTWISE REPORTMISSING=NO

TRANSFORM=VARSTOCASES(SUMMARY="#SUMMARY" INDEX="#INDEX")

/GRAPHSPEC SOURCE=INLINE.

BEGIN GPL

SOURCE: s=userSource(id("graphdataset"))

DATA: SUMMARY=col(source(s), name("#SUMMARY"))

DATA: INDEX=col(source(s), name("#INDEX"), unit.category())

DATA: KMCLUopt=col(source(s), name("KMCLUopt"), unit.category())

GUIDE: axis(dim(2), label("Mean"))

GUIDE: legend(aesthetic(aesthetic.color.interior), label("Cluster Number of Case"))

SCALE: cat(dim(1), include("0", "1", "2", "3", "4", "5", "6", "7", "8", "9", "10", "11"

, "12"))

SCALE: linear(dim(2), include(0))

ELEMENT: line(position(INDEX\*SUMMARY), color.interior(KMCLUopt), missing.wings())

END GPL.

**Linijski grafikon za standardizirane tematske spremenljivke razvrščene po metodi voditeljev**

\* Chart Builder.

GGRAPH

/GRAPHDATASET NAME="graphdataset" VARIABLES=MEANCI(ZV198, 95) MEANCI(ZV199, 95) MEANCI(ZV200, 95) MEANCI(ZV202, 95) MEANCI(ZV203, 95) MEANCI(ZV203A, 95) MEANCI(ZV204, 95) MEANCI(ZV205, 95) MEANCI(ZV206, 95) MEANCI(ZV207, 95) MEANCI(ZV208, 95) MEANCI(ZV209, 95) MEANCI(ZV210, 95) KMCLUopt MISSING=LISTWISE REPORTMISSING=NO

TRANSFORM=VARSTOCASES(SUMMARY="#SUMMARY" INDEX="#INDEX" LOW="#LOW" HIGH="#HIGH")

/GRAPHSPEC SOURCE=INLINE.

BEGIN GPL

SOURCE: s=userSource(id("graphdataset"))

DATA: SUMMARY=col(source(s), name("#SUMMARY"))

DATA: INDEX=col(source(s), name("#INDEX"), unit.category())

DATA: KMCLUopt=col(source(s), name("KMCLUopt"), unit.category())

DATA: LOW=col(source(s), name("#LOW"))

DATA: HIGH=col(source(s), name("#HIGH"))

GUIDE: axis(dim(2), label("Mean"))

GUIDE: legend(aesthetic(aesthetic.color.interior), label("Cluster Number of Case"))

GUIDE: text.footnote(label("Error Bars: 95% CI"))

SCALE: cat(dim(1), include("0", "1", "2", "3", "4", "5", "6", "7", "8", "9", "10", "11"

, "12"))

SCALE: linear(dim(2), include(0))

ELEMENT: line(position(INDEX\*SUMMARY), color.interior(KMCLUopt), missing.wings())

ELEMENT: interval(position(region.spread.range(INDEX\*(LOW+HIGH))), shape.interior(shape.ibeam), color.interior(KMCLUopt))

END GPL.

**Kontingenčna tabela pripadnosti enot skupinam narejeni po Wardu in po metodi voditeljev**

CROSSTABS

/TABLES=CLU3wardSEk3 BY cluKmeansSPPSk3

/FORMAT=AVALUE TABLES

/CELLS=COUNT

/COUNT ROUND CELL.

**Randov indeks za skupini CLU3wardSEk3 (Wardova metoda) in KMCLUopt (metoda voditeljev) – uporaba makrota.**

RandIndex CLU3wardSEk3 cluKmeansSPPSk3.

**Razsevni grafikoni likartovih lestvic po skupinah**

GGRAPH

/GRAPHDATASET NAME="graphdataset" VARIABLES=KMCLUopt matKD liberalnost nasilje MISSING=LISTWISE

REPORTMISSING=NO

/GRAPHSPEC SOURCE=INLINE.

BEGIN GPL

SOURCE: s=userSource(id("graphdataset"))

DATA: KMCLUopt=col(source(s), name("KMCLUopt"), unit.category())

DATA: matKD=col(source(s), name("matKD"))

DATA: liberalnost=col(source(s), name("liberalnost"))

DATA: nasilje=col(source(s), name("nasilje"))

GUIDE: axis(dim(1.1), ticks(null()))

GUIDE: axis(dim(2.1), ticks(null()))

GUIDE: axis(dim(1), gap(0px))

GUIDE: axis(dim(2), gap(0px))

GUIDE: legend(aesthetic(aesthetic.color.exterior), label("Cluster Number of Case"))

TRANS: matKD\_label = eval("Materialna kazniva dejanja")

TRANS: liberalnost\_label = eval("Liberalnost")

TRANS: nasilje\_label = eval("Odnos do nasilja")

ELEMENT: point.jitter(position((matKD/matKD\_label+liberalnost/liberalnost\_label+nasilje/nasilje\_label)\*

(matKD/matKD\_label+liberalnost/liberalnost\_label+nasilje/nasilje\_label)),color.exterior(KMCLUopt))

END GPL.

**Kontingenčna tabela za porazdeljenost skupin (razvrščena po metodi voditeljev) po spolu, ter pearsonov hi kvadrat in kramerjev koeficient**

CROSSTABS

/TABLES=V240 BY KMCLUopt

/FORMAT=AVALUE TABLES

/STATISTICS=CHISQ PHI

/CELLS=COUNT COLUMN

/COUNT ROUND CELL.