

Kvantitatiivne andmeanalüüs II

KODUNE TÖÖ 3

Tähtaeg: 08.01.2024.

Pikkus: 2-3 lehekülge (sh. tabelid/joonised ja tekst) + skript (+ lisa(d))

Kodutöö eesmärgiks on esiteks tõlgendada **faktoranalüüsi** (ülesanne 1), kusjuures analüüsi ei tule ise teha, vaid tõlgendada olemasolevaid tulemusi. Teiseks aga rakendada **logistilist regressioonanalüüsi** (ülesanne 2), mille käigus tuleb teostada iseseisvalt analüüs ning esitada korrektselt vormistatuna analüüsi tulemused. Analüüsi teostamine sisaldab endas nii analüüsiks sobilike tunnuste valikut, nende teisendamist kui ka analüüsis teostamist. Tulemuste esitamine sisaldab endas tulemuste esitamiseks parima viisi (tabel, joonis, tekst jne) valikut ja teostust ning tulemuste statistilist ja sisulist tõlgendust.

ÜLESANNE

- 2017.a aasta Euroopa Väärtuste Uuringus (European Value Survey) oli järgmine küsimuste plokk:

Q44 Please tell me for each of the following whether you think it can always be justified, never be justified, or something in between, using this card.

		never									always	DK	NA
v149	Claiming state benefits which you are not entitled to	1	2	3	4	5	6	7	8	9	10	88	99

v150	Cheating on tax if you have the chance	1	2	3	4	5	6	7	8	9	10	88	99
v151	Taking the drugs marijuana or hashish	1	2	3	4	5	6	7	8	9	10	88	99
v152	Someone accepting a bribe in the course of their duties	1	2	3	4	5	6	7	8	9	10	88	99
v153	Homosexuality	1	2	3	4	5	6	7	8	9	10	88	99
v154	Abortion	1	2	3	4	5	6	7	8	9	10	88	99
v155	Divorce	1	2	3	4	5	6	7	8	9	10	88	99
v156	Euthanasia (terminating the life of the incurably sick)	1	2	3	4	5	6	7	8	9	10	88	99
v157	Suicide	1	2	3	4	5	6	7	8	9	10	88	99
v158	Having casual sex	1	2	3	4	5	6	7	8	9	10	88	99
v159	Avoiding a fare on public transport	1	2	3	4	5	6	7	8	9	10	88	99
v160	Prostitution	1	2	3	4	5	6	7	8	9	10	88	99
v161	Artificial insemination or in-vitro fertilization	1	2	3	4	5	6	7	8	9	10	88	99
v162	Political violence	1	2	3	4	5	6	7	8	9	10	88	99
v163	Death penalty	1	2	3	4	5	6	7	8	9	10	88	99

Kvantitatiivse andeanalüüsi kursuse tudeng otsustas analüüsida neid andmeid Venemaa kohta, et saada teada:

- a) Milline käitumisviis on Vene ühiskonnas enim aktsepteeritud ja milline kõige vähem?
- b) Kas ja millistele üldisematele dimensioonidele oleks võimalik neid mõõdetud väärtusi taandada?

Selleks tegi ta faktoranalüüsi ja sai järgmised tulemused:

Kirjeldav analüüs

Tabel 1. Kirjeldatavad näitajad (min, max, keskmine, st.hälve)

	Min	Max	Keskmine	St.hälve
Põhjuseta riiklike toetuste/hüvitiste saamine	1	10	2,12	2,007
Maksupettus	1	10	2,08	2,019
Kanepi ja hašiši kasutamine	1	10	2,45	2,389
Altkäemaksu võtmine	1	10	1,57	1,472
Homoseksuaalsus	1	10	4,97	3,797
Abort	1	10	5,22	3,303
Lahutus	1	10	6,26	3,152
Eutanaasia	1	10	5,09	3,386
Enesetapp	1	10	2,87	2,671
Juhuseks	1	10	4,03	3,155
Piletita sõit ühistranspordis	1	10	2,62	2,382
Prostitutsioon	1	10	2,97	2,616
Kunstlik viljastamine	1	10	6,75	3,242
Poliitiline vägivald	1	10	1,63	1,551
Surmanuhtlus	1	10	3,31	2,976

Faktoranalüüs (4-faktoriline lahendus)

Peakomponentide meetod varimax pööramisega.

KMO = 0.864,

Tabel 2. Faktorite kirjeldusvõime ja faktoraaldungid pärast pööramist

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4,850	32,335	32,335	2,896	19,305	19,305
2	1,860	12,399	44,734	2,779	18,524	37,829
3	1,244	8,296	53,030	2,022	13,478	51,307
4	1,062	7,082	60,112	1,321	8,805	60,112

	Faktorlaaduningud			
	Faktorid			
	1	2	3	4
Lahutus	,803			
Abort	,764			
Eutanaasia	,585			
Kuntslik viljastamine	,582			
Juhuseks	,569	,401		
Kanepi ja hašiši kasutamine		,762		
Altkäemaksu võtmine		,691		
Homoseksuaalsus		,683		
Enesetapp		,599		
Poliitiline vägivald		,522		,499
Prostitutsioon	,446	,490		
Põhjusteta riiklike toetuste/hüvitiste saamine			,826	
Maksupettus			,809	
Piletita sõit ühistranspordis	,445		,587	
Surmanuhtlus				,860

Teie ülesanne on neid tulemusi tõlgendades vastata eelpool toodud kahele uurimisküsimusele ehk siis 1) millist/milliseid käitumisviise aktsepteeritakse enim, mida vähem. (selleks piisab kirjeldava analüüsi tulemuste tõlgendamisest) ning 2) kas ja milliseid üldisemaid väärtusdimensioone/faktoreid võib nende näitajate põhjal mõtte eristada. (selleks peaksite tõlgendama faktoranalüüsi tulemusi). Lisaks tulemuste statistilisele ja sisulisele tõlgendamisele andke omapoolne üldisem hinnang ka faktoranalüüsi mudeli/tulemuste headusele ning kus seda ei ole võimalik teha, kirjutage miks seda teha ei saa ning millist infot/näitajaid oleks veel hea vaadata (selle aluseks võtke praktikumis läbi arutatud näitajad).

2. Teie ülesandeks on analüüsida hoiakuid nn kodanikupalga (*basic income*) süsteemi osas – kes või mis seostub selle pooldamisega või mitte-pooldamisega.

Analüüsi läbiviimiseks kasutage juba tuttavat Euroopa Sotsiaaluuringu (European Social Survey) 8. vooru (2016. aasta) andmestikku. See on vabakasutusega rahvusvaheline võrdlev

uuring, mille leiate selle projekti kodulehelt <http://www.europeansocialsurvey.org/>. Andmefaili ja codebook'i leiate Moodle's kodutöö ülesande kaustast. NB! Andmestik on kõige lihtsam formaadis, st kogu info on vaid numbrilisel kujul (ESS8_num.csv). Seega pidage meeles enne analüüsi juurde asumist analüüsiks vajalikud tunnused sobivasse formaati panna. Andmestikus võib olla tunnuseid, mida ankeedis sellisel kujul otseselt ei ole (nt ümberkodeeritud haridustunnused) – kui soovite mõnda neist tunnustest kasutada, siis uurige lisainfot projekti kodulehelt või siis küsige abi õppejõult.

Igale üliõpilasele on analüüsiks antud üks Euroopa maa (vt maade jaotust allpool). Tegemist on kõiki maid koondava andmestikuga, seega soovitan alustuseks sealt välja filtreerida vaid Teile analüüsiks antud maa ning selekteerida vaid analüüsiks vajalikud tunnused.

Kodanikupalga suhtes valitseva hoiaku mõõtmiseks kasutage tunnust *basinc*:

E36 CARD 54 Some countries are currently talking about introducing a basic income scheme¹¹². In a moment I will ask you to tell me whether you are against or in favour of this scheme. First, I will give you some more details. The highlighted box at the top of this card shows the main features of the scheme. A basic income scheme includes all of the following:
...**READ OUT**...

- The government pays everyone a monthly income to cover essential living costs.
- It replaces many other social benefits.
- The purpose is to guarantee everyone a minimum standard of living.
- Everyone receives the same amount regardless of whether or not they are working.
- People also keep¹¹³ the money they earn from work or other sources.
- This scheme is paid for by taxes.

INTERVIEWER: PAUSE TO GIVE THE RESPONDENT TIME TO READ CARD.

Overall, would you be against or in favour of having this scheme in [country]?
Please choose your answer from the options at the bottom of the card.

Strongly against	1
Against	2
In favour	3
Strongly in favour	4

Kuna sellel tunnusel on neli kategooriat, siis soovitan selle alustuseks ümber kodeerida binaarseks: toetab (poolt + tugevalt poolt) ning ei toeta (vastu + tugevalt vastu). Analüüsi tegemisel otsustage, millist shanssi (toetada versus mitte toetada) soovite hinnata/esitada ning vastavalt sellele kodeerige tunnus või järjestage tunnuse väärtused ümber (kasutades näiteks releval funktsiooni).

Analüüsi tegemisel kasutage binaarset logistilist regressiooni.

Töös tuleb esitada/kõrvutada/võrrelda kolme mudelit:

Mudel 1: seletavad tunnused on indiviidi tasandi sotsiaal-demograafilised karakteristikud nagu sugu (gndr), vanus (aega), haridus (edulevlb, eisced vms), rahvus (lnghom1, blgetmg vms). Sellesse mudelisse tuleb kaasata vähemalt sõltumatut ehk seletavat kolm tunnust.

Mudel 2: mudel 1 tunnused + lisaks vähemalt üks respondendi või leibkonna sotsiaalmajandusliku staatuse tunnus (nt tööturustaatus (mnactic), sissetulekudetsiil (hinctnta), hinnang leibkonna sissetulekule (hincfel) vms)

Mudel 3: mudel 1 ja 2 tunnused + vähemalt üks väärtusi, hoiakuid, tõekspidamisi, maailmavaadet vms iseloomustav tunnus (näiteks usklikkuse tase (rlgdgr), maailmavaade vasak-parem skaalal (lrscale), jms))

NB! Seletavate tunnuste hulgas peab olema nii kvantitatiivseid kui kategoriaalseid tunnuseid.

Analüüsi teostamiseks:

a) Sõnastage üks kuni kaks uurimisküsimust selle kohta, milliseid seoseid ootate. Püüdke põhjendada, miks sellist seost uurite ning ootate. Siin ei eeldata, et lisaksite viiteid vms (kui seda teete, siis on see boonuseks), piisab kui kasutate oma senistel teadmistel ning üldisel loogikal põhinevaid argumente/seletusi.

b) Kirjeldage andmeid ning tunnuseid, mille analüüsimiseks olete välja valinud. Püüdke oma tunnuste valikut ka veidi põhjendada. Analüüsi kaasatud muutujate kirjeldava statistika esitamiseks võite kasutada ka tabeleid või jooniseid. Hoidke see osa siiski lühikese ja kompaktsena – oluline on, et saaks aru, milliseid tunnuseid kasutate, mis on nende skaalad jms.

d) Tehke regressioonanalüüs, st kolm regressioonimudelit Teile antud maa kohta. Esitage oma tulemused (mudelid) ühe tabelina võrreldes kolme erinevat mudelit ning esitades sellest olulise info tunnuste kohta (nt kategooriate nimed, referentgrupid jms), mõjude (vabalt valitud moel, st log-koefitsientidena, suhtekorrajatena või average marginal effects'idena) ja mõjude olulisuse (standardviga, Sig vms) kohta. Tabeli juures peaks olema ka info analüüsis olevate indiviidide (arv, selektsioonide kirjeldus, kui neid on jms) ning mudeli sobivuse/headuse kohta. Tõlgendage oma tulemusi nii statistiliselt kui sisuliselt. Mh võrrelge mudeleid, sh mis juhtub mudelis siis kui lülitate mudelisse uusi tunnuseid

NB! Analüüsi tegemise juures jälgige hoolega, mis tüüpi on need tunnused, mida plaanite kasutada ning sellele vastavalt, et kas ja kuidas neid Teie poolt valitud analüüsis korrektselt kasutada.

e) Sõnastage lähtuvalt oma uurimisküsimustest/hüpoteesidest kokkuvõtte, paigutades oma tulemused laiemasse maa konteksti.

Nõuded kodutööde vormistusele

- o Igal kodutööl peab olema autori nimi ning uurimuse kohta sõnastatud vastav peakiri;
- o Kodutöö peab olema korrektselt vormistatud, seda nii keeleliselt, sisuliselt kui vormiliselt;
- o Töös esitatud tabelid/joonised ja tekst peavad olema eesti keeles;
- o Pöörake tähelepanu tabelite ja jooniste vormistusele
- o Tabelid/joonised peavad olema alati peal/allkirjastatud;
- o Tabeli/joonise all peab olema viide andmeallikale;
- o Tunnuste/kategooriate nimed peavad olema sisukad ja lugejale mõistetavad, mitte lihtsalt andmefailist/süntaksist ülevõetud;
- o Täpsustage, millega on tabeli/joonise näol tegu (protsendid, keskmised jne);
- o Iga tabel/joonis peab moodustama terviku, st olema mõistetav ka tekstist eraldi; tabelid peavad olema korrektselt vormistatud
- o Kodutööle peab olema lisatud skript, mis sisaldab mh kodutöö tegemiseks kasutatud andmeteisendusi. NB! Skript peab olema 'puhas', st selles on esitatud õiges järjekorras vaid need käsud, mida töös esitatud tulemuste saamiseks realselt vaja läks.

Maade jaotus:

Üliõpilane		Maa
Helena	Tihkan	AT – Austria
Eleonor	Kalamets	CZ- Tšehhi
Claudia Isabel	Lopez Ortiz	FI – Soome
Lisbeth	Ilves	DE – Saksamaa
Emilia	Heero	RU – Venemaa
Artjom	Gnezdilov	NL- Holland
Daniil	Stõk	PL – Poola
Karolina Helene	Kaukver	NO – Norra
Äli	Bergmann	ES – Hispaania
Kelly	Kohal	FR – Prantsusmaa
Hosanna Sofia	Mäekallas	HU – Ungari
Kertu	Saks	GB – Suur-Britannia
Anna Marie	Vasar	BE – Belgia
Kaarel	Leedo	EE – Eesti
Deborah	Šapovalov	IT – Itaalia
Ekaterina	Filippova	SE – Rootsi
Hele	Simson	CH – Šveits
Marta	Bogatõr	PT – Portugal
Kaleria	Frolovskaja	IE – Iirimaa
		IL – Iisrael
		IS – Island
		LT – Leedu

		SI – Slovenia
--	--	---------------