



Latvijas Statistiķu asociācijas,
Centrālās statistikas pārvaldes,
Latvijas Universitātes Doktorantūras skolas “Baltijas jūras reģiona valstu integrācija ES nozīmīgākās sadarbības dimensijās”
lasījums

KLASTERANALĪZES LIETOJUMS REĢIONĀLĀS DEMOGRĀFISKĀS ATTĪSTĪBAS NOVĒRTĒJUMĀ

Dr. demogr. Aleksandrs Dahs
Latvijas Universitāte

2024. gada 10. aprīlī, Rīgā, Latvijā

Pētījumu atbalsta Nacionālā pētniecības programmas
“**Letonika latviskas un eiropēiskas sabiedrības attīstībai**”
projekts Nr.VPP Letonika - 2021/4 - 0002

“**Jauni risinājumi demogrāfijas un migrācijas procesu izpētē latviskas un eiropēiskas
zināšanu sabiedrības attīstībai**”

(DemoMig Pro)

Literatūras apskats

Latvijas teritoriālo vienību klasteranalīze pirms 2015. gada:

- Melluma, A., Ramute, L., and Rivza, P. (1996). Cluster analysis for the classification of rural communities. *Latvijas Lauksaimniecības Universitātes Raksti – Lauku attīstība*, no. 7, 99-102.
- Brauksa, I. (2013). Use of cluster analysis in exploring economic indicator differences among regions: The case of Latvia. *Journal of Economics, Business and Management*, 1(1), 42-45.

Klasteranalīzes pielietojums demogrāfisko procesu izpēti citur Eiropā:

- Hasek, O. (2020). Regionální diferenciace plodnosti podle typologie venkova. (*The Regional Differentiation of Fertility by Rural Typology in Czechia*). *Demografie*. Vol. 62, 3-13.
- Kuznetsova, T.Y., (2015). Geo-demographic typology of municipalities of the Kaliningrad region. *Mediterranean Journal of Social Sciences*, Issue 6, Vol. 6. pp. 345-350.
- Recaño, J., (2017). La sostenibilidad demográfica de la España vacía. (*The demographic sustainability of empty Spain*). *Perspectives démographiques*. Issue 7, 1-4.
- Stonawska, K., Vaishar, A. (2018). Differentiation and Typology of the Moravian Countryside. *European Countryside*. Vol. 10, 127-140.

Metodoloģija:

- Peters, W. S. (1958). Cluster analysis in urban demography. *Social Forces*, Vol. 37, 38-48.
- Kaufman, L. and Rousseeuw, P.J., (2009). *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons.

DemoMig projekta pētnieku grupas publikācijas:

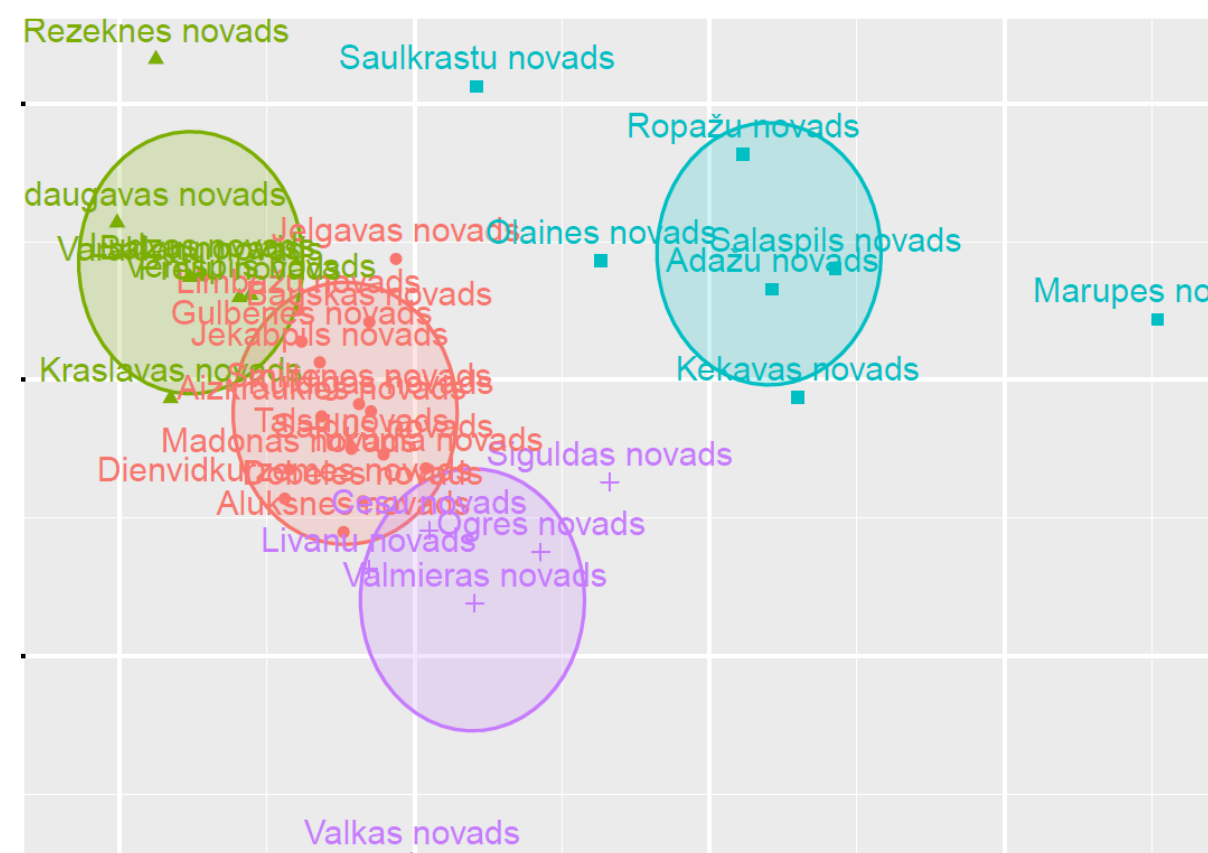
- Krisjane, Z., Krumins, Ju., Dahs, A., Lece, K., Krumins, Ja. (2021). Teritoriālo un individuālo datu lietojuma inovācijas Latvijas demogrāfiskās situācijas analīzē par 2011.–2021. gadu. *Latvijas Zinātņu Akadēmijas Vēstis, A daļa* (2021), No. 4., 76-93.
- Krumins, J., Berzins, A., Dahs, A. Regional Demographic Trends in Accordance with the New Model of Territorial Division in Latvia and Future Dynamics in Statistical Regions of the Baltic States. *Economic Science for Rural Development Conference Proceedings*, 2020, Issue 54, 233-240.

Klasteranalīzes pieejas izvēle

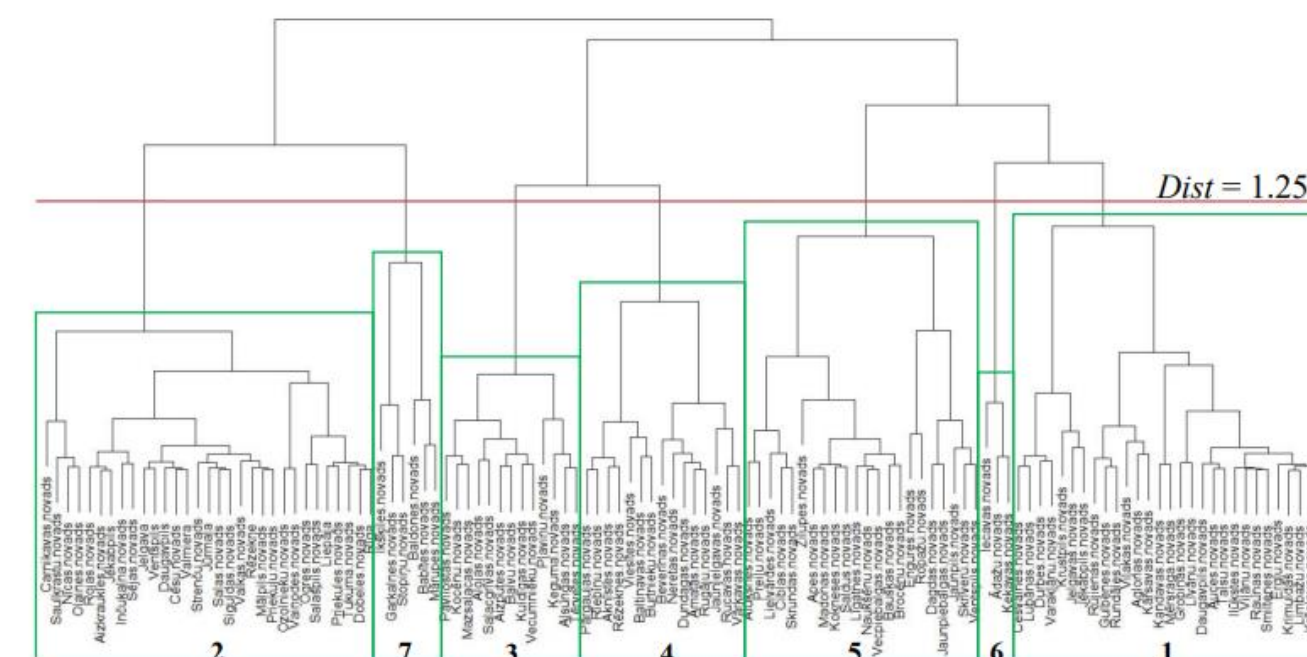
Ilgspējīgas attīstības kontekstā ir svarīgi izvērtēt reģionālo demogrāfisko situāciju un tās attīstības perspektīvas.

Lai spētu pielāgoties dinamiskām situācijas izmaiņām un palielināt izpratni par reģionālo demogrāfisko situāciju, ir nepieciešams efektīvs un salīdzinoši vienkāršs teritoriālo vienību kategorizācijas rīks.

Reģionālās demogrāfijas pētījumos tiek pielietoti gan hierarhiskās, gan nehierarhiskās klasteranalīzes metodes.



Latvijas lauku pašvaldību (2021) nehierarhiskās (PAM) klasteranalīzes procesa shematiskais attēlojums.



Koka diagramma, kas attēlo Latvijas pašvaldību teritoriju (2017. gadā) hierarhiskās klasteranalīzes rezultātus pēc to iedzīvotāju skaita trendu līdzības (korelācijas) 1959.-2011. gadu periodā

- **Hierarhiskā klasteranalīze** tiek pielietota novērojumu grupēšanai pēc neliela indikatoru skaita un galvenokārt balstās uz atšķirības distances novērtējuma (piem. korelācijas) starp novērojumiem.
- **Nehierarhiskā klasteranalīze** tiek pielietota plašākām indikatoru kopām un sniedz iespēju veikt periodisku situācijas novērtējumu.

Avots: autora izstrāde, pamatojoties uz Oficiālās statistikas portāla datiem .

Nehierarhiskās klasteranalīzes priekšrocības



Neatkarība no pētnieku pieņēmumiem

Klasteru noteikšana notiek izmantojot mašīnmācīšanās algoritmu.
Rezultātu neietekmē iepriekš noteiktie novērojumu atšķirības sliekšņi vai mērvienības



Lielāka datu izvēles brīvība

Praktiski neierobežots indikatoru un novērojumu skaits.
Rezultātus neietekmē indikatoru multikolinearitāte.



Iespēja mainīt klasteru skaitu un eksperimentēt

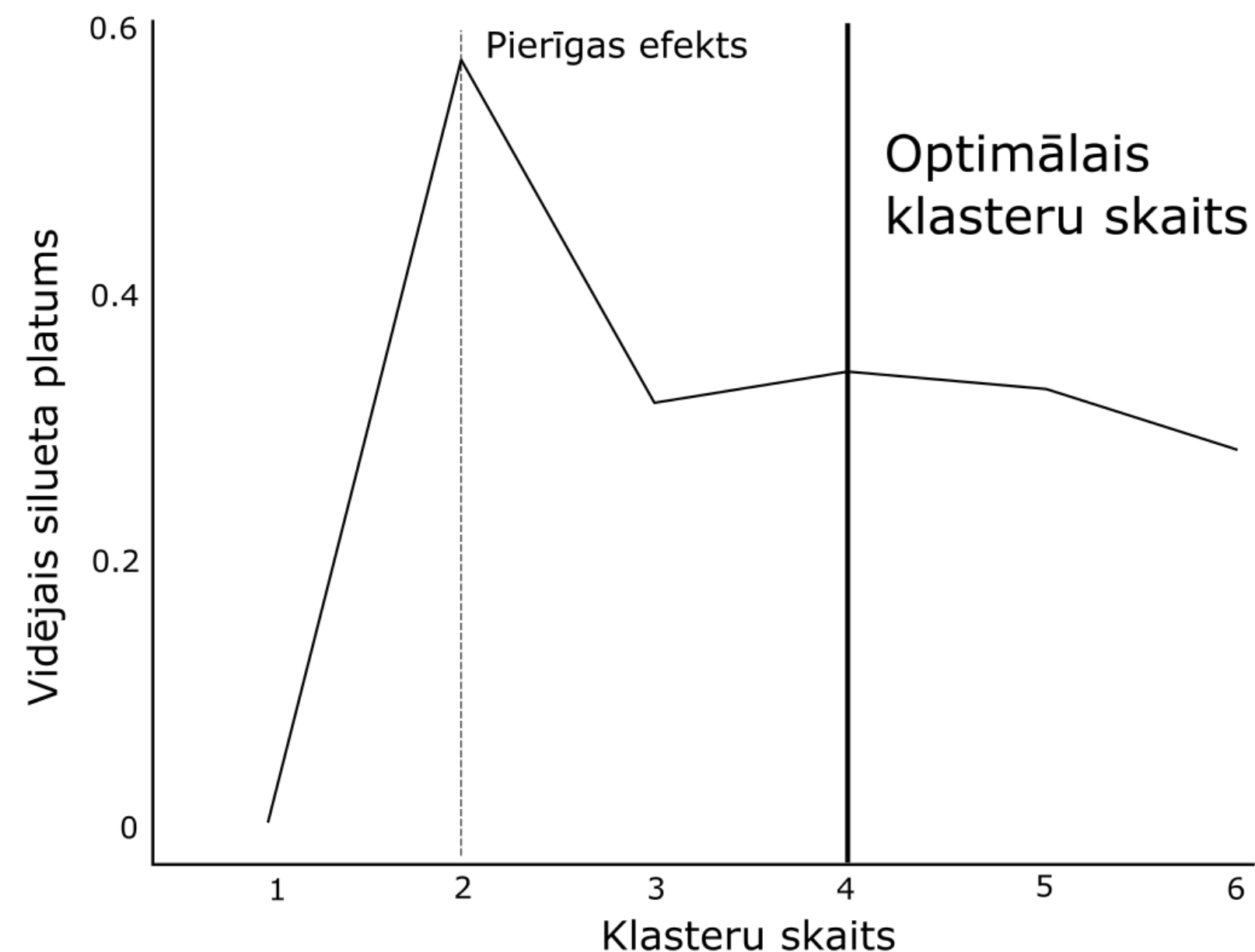
Iespēja veidot jebkādu klasteru skaitu no «1» līdz «n».
Pieejamas vairākas metodes optimāla klasteru skaita noteikšanai.
Labāko rezultātu var iegūt arī secīgi palielinot klasteru skaitu un eksperimentējot.

Metodoloģija

Tā kā teritoriālajos datos pastāv anomālu novērojumu risks, visdrošāk ir izmantot pārbaudītu **k-medoidu PAM (*Partitioning around Medoids*)** algoritmu.

Lai noteiktu optimālo klasteru skaitu, var izmantot **Vidējo siluetu metodi**, kas paredz PAM algoritma izmantošanu atšķirīgam klasteru skaitam ar atbilstošu vidējo klasteru siluetu noteikšanu katram no rezultātiem.

Sākotnēji PAM algoritma lietojuma aprēķinu rezultāti liecina, ka optimālais klasteru skaits Latvijas novadu grupēšanai ir divi. Aplūkojot klasteru sadalījumu, var secināt, ka šādu rezultātu ietekmē Prierīgas novadu efekts. Tāpēc par optimālu klasteru skaitu ir atzīstams nākamais labākais rezultāts – **četri klasteri**.

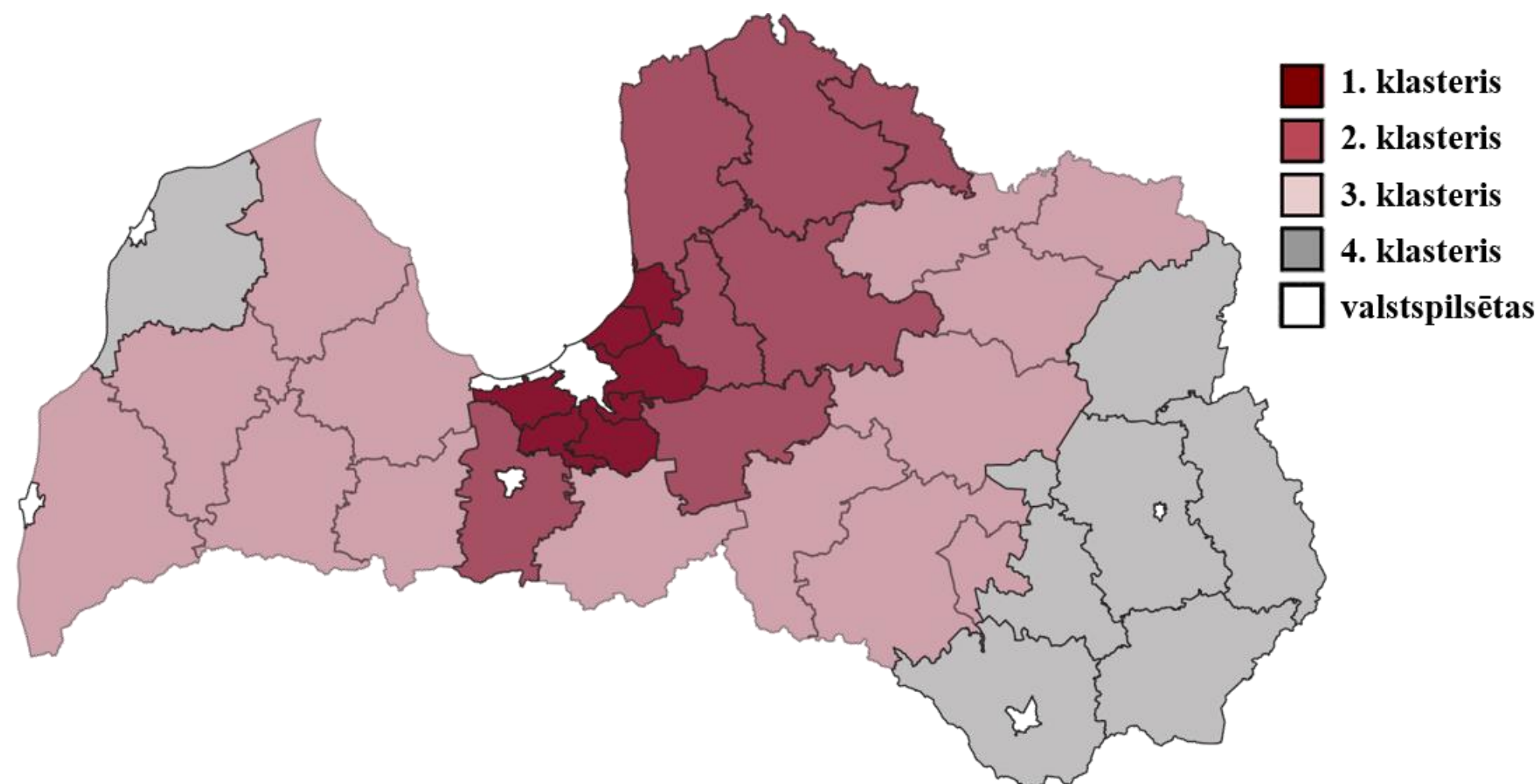


Avots: autora izstrāde, pamatojoties uz Oficiālās statistikas portāla datiem .

Rezultāti Latvijas kontekstā (1)

Lai nodrošinātu pētījuma aktualitāti, tiek izmantota Latvijas pašvaldību struktūra, kas izveidota pēc 2021.gada administratīvi teritoriālās reformas. Pētījumā tiek apskatīti tikai novadi, kuru sastāvā ir arī lauku teritorijas.

Klasteris	Pašvaldība	Iedzīvotāju blīvums (iedz. uz km2, 2021)	Dabiskā kustība (%) (2018-2021)	Migrācijas saldo (%) (2018-2021)	Vīriešu īpatsvars 15-63 vecuma grupā (%) (2021)	Bērnu skaits (0-14) uz 100 senioriem (64+) (2021)	Demogrāfiskā slodze (2021)
1	Ādažu novads	93.94	0.53	10.36	50.05	130.37	60.67
2	Ogres novads	32.61	-1.55	1.17	49.54	86.21	63.65
3	Talsu novads	13.35	-2.75	-2.81	51.32	65.66	61.23
4	Balvu novads	8.13	-5.39	-3.70	52.27	56.13	57.15



Lai novērtētu iegūtos klasterus, ir lietderīgi aplūkot atsevišķas pašvaldības, kas pārstāv katru grupu.

Klasteru analīzes rezultāti liecina par ievērojamām atšķirībām iedzīvotāju dinamikā starp pašvaldībām, kas iedalītas četros klasteros:

1. demogrāfiskās izaugsmes apgabali,
2. līdzsvarotu iedzīvotāju skaita izmaiņu apgabali,
3. mērenās depopulācijas apgabali,
4. paātrinātās depopulācijas apgabali.

Avots: autoru izstrāde, pamatojoties uz Oficiālās statistikas portāla datiem

Rezultāti Latvijas kontekstā (2)

- **Demogrāfiskās izaugsmes apgabali** – pašvaldības, kas atrodas Rīgas pilsētas metropoles teritorijā ar augstu iedzīvotāju blīvumu, labvēlīgu vecuma struktūru, pozitīvu iedzīvotāju skaita pieaugumu un vidējo demogrāfisko slodzi. Šajās pašvaldībās iedzīvotāju sastāvā ir stabils darbības vecuma sieviešu īpatsvars.
- **Līdzsvarotu iedzīvotāju skaita izmaiņu apgabali** – pašvaldības ar vidējo iedzīvotāju blīvumu, nedaudz negatīvām dabiskajām iedzīvotāju skaita izmaiņām un stabilu vai pozitīvu migrācijas saldo. Šajās pašvaldībās parasti ir vislielākie demogrāfiskie slodzes rādītāji, nedaudz nestabila iedzīvotāju vecuma struktūra ar gandrīz vienādu nepilngadīgo un senioru skaitu. Šie apgabali uzrāda dzimumu sadalījumu ar lielāku sieviešu īpatsvaru darbības vecuma grupā.
- **Mērenās depopulācijas apgabali** – pašvaldības ar zemu vidējo iedzīvotāju blīvumu, salīdzinoši augstu demogrāfisko slodzi un augstiem emigrācijas rādītājiem. Šajos apgabalos vērojamas arī nestabilas iedzīvotāju struktūras, stagnējošas vai negatīvas dabiskās populācijas izmaiņas un darbības vecuma vīriešu pārsvars virs valsts vidējā līmeņa.
- **Paātrinātās depopulācijas apgabali** – pašvaldības ar īpaši zemu iedzīvotāju blīvumu, ļoti zemu gados jaunu iedzīvotāju īpatsvaru un ārkārtīgi augstu depopulācijas līmeni, ko ietekmē negatīva dabiskā kustība un negatīvs migrācijas saldo. Šajās pašvaldībās ir arī vislielākais iedzīvotāju īpatsvars darbības vecuma grupā un lielākās vīriešu un sieviešu dzimumu atšķirības darbības vecuma iedzīvotāju vidū.



Secinājumi Latvijas kontekstā



Rīgas metropoles areāla ietekme

Rīgas metropoles sociālās un ekonomiskās sekundārā ietekme ir acīmredzamas otrajā klasterī iekļauto pašvaldību vidū, kur stabilu iedzīvotāju skaitu galvenokārt atbalsta pozitīva iekšzemes migrācija, ko stimulē Rīgas metropoles teritorijas paplašinātie suburbanizācijas procesi.



Samazinās citu valstspilsētu ietekme

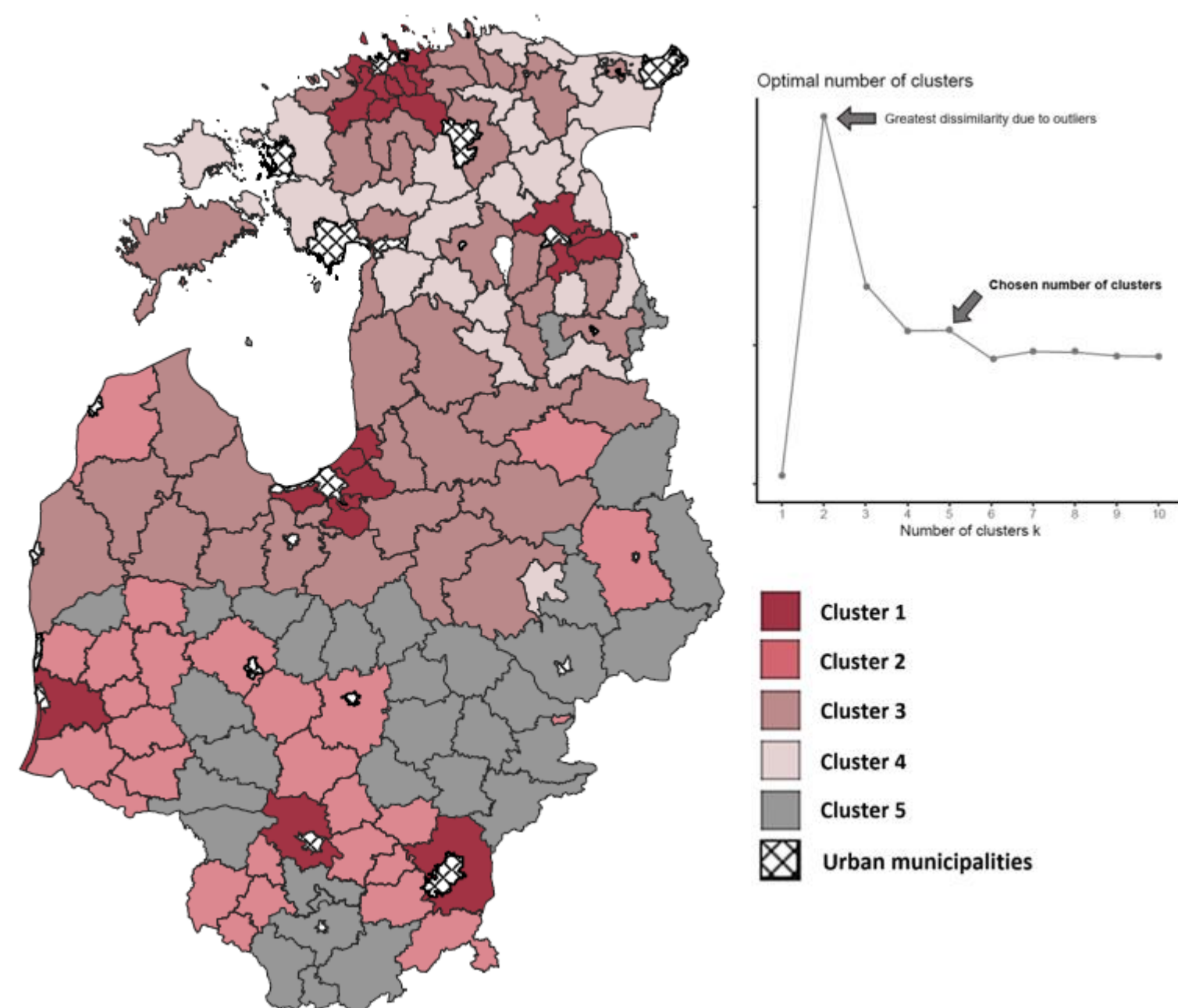
Valstspilsētām un reģionālajiem attīstības centriem, kas ģeogrāfiski atrodas trešajā un ceturtajā klasterī, praktiski nav pozitīvas suburbanizācijas ietekmes apkārtesošajās teritorijās. Ievērojot minēto, ārpus Rīgas metropoles zonas esošo apdzīvoto vietu suburbanizācijas procesu atsākšana un veicināšana ir ļoti svarīga lauku teritoriju revitalizācijai.



Iedzīvotāju blīvums un sociālie pakalpojumi

Atšķirība starp otro un trešo klasteri galvenokārt ir saistīta ar iedzīvotāju blīvumu un novecošanas līmeni. Analīzes rezultāti liecina, ka ģimenes ar bērniem izvēlas blīvāk apdzīvotās vietas, kur ir vairāk pieejami nepieciešamie sociālie pakalpojumi, savukārt seniori turpina dzīvot rajonos, kas pamazām iztukšojas.

Rezultāti Baltijas kontekstā



Cluster	Typical municipality representing the cluster	Population density (2022) (pers/km sq)	Share of working age population (15-64) (2022) (%)	Share of population below working age (0-14) (2022) (%)	Share of population above working age (65+) (2022) (%)	Total change of population number from 2018 to 2023 (%)	Natural change of population number from 2018 to 2023 (%)	Change of population number due to migration from 2018 to 2023 (%)
1	LV_Ropažu novads	63,52	64,97%	20,33%	14,70%	17,37%	0,85%	16,52%
2	LT_Šilutės rajono savivaldybė	22,48	66,00%	14,20%	19,80%	-1,89%	-3,97%	-1,89%
3	LV_Dobeles novads	16,94	62,32%	15,35%	22,33%	-6,39%	-3,92%	-2,46%
4	EE_Tõrva rural municipality	9,10	60,06%	14,35%	25,59%	-5,22%	-5,76%	0,53%
5	LT_Kupiškio rajono savivaldybė	14,93	64,80%	11,40%	23,80%	-2,97%	-6,27%	-2,97%

- Izaugsmes zonas** – pašvaldības, kas atrodas lielo pilsētu tuvumā un pilsētu metropoles teritorijās, kurās ir augsts iedzīvotāju blīvums, stabila vecuma struktūra un pozitīvs iedzīvotāju skaita pieaugums.
- Demogrāfiski stagnējošas teritorijas** – pašvaldības ar vidēju iedzīvotāju blīvumu, augstu darbības vecuma iedzīvotāju īpatsvaru un nenozīmīgu iedzīvotāju skaita samazināšanos. Raksturīgi Lietuvai.
- Emigrācijas teritorijas** – pašvaldības ar vidēju iedzīvotāju blīvumu, salīdzinoši augstu demogrāfisko slogu un jūtamu negatīvu iedzīvotāju skaita samazināšanos gan negatīvā dabiskā pieauguma, gan migrācijas saldo ietekmē. Raksturīgi Latvijai.
- Negatīvā dabiskā pieauguma teritorijas** – pašvaldības ar zemu iedzīvotāju blīvumu, augstu senioru īpatsvaru un augstu iedzīvotāju skaita samazināšanās tempu negatīvā dabiskā pieauguma dēļ. Raksturīgi Igaunijas lauku rajonos.
- Demogrāfiskās krīzes zonas** – pašvaldības ar zemu iedzīvotāju blīvumu, ļoti zemu gados jaunu iedzīvotāju īpatsvaru un ārkārtīgi augstu depopulācijas līmeni gan dabiskās kustības, gan negatīvas migrācijas saldo dēļ. Pārsvarā sastopamas Latvijas un Lietuvas perifērijas teritorijās.

Diskusijai



Metodoloģija

- Kādus vēl klasteranalīzes vai mašīnmācīšanās rīkus varētu pielietot demogrāfijas pētījumos?

Dati un rezultāti

- Datu izvēle ietekmē klasteranalīzes rezultātus. Kādas vēl datu kopas (oficiālās vai eksperimentālās) varētu pielietot reģionu demogrāfiskās attīstības klasteranalīzei?

Citas pētījumu jomas

- Kurās sociālo pētījumu nozarēs būtu lietderīgi pielietot šādu klasteranalīzes pieeju?

Publikācijas

Saistīto pētījumu rezultāti tiks atspoguļoti 2024. gada publikācijās:

- **Vanguard and laggard rural population clusters in Latvia**
autori Aleksandrs Dahs ; Juris Krūmiņš; Atis Bērziņš un Kristīne Lece
- **Machine-learning technique for drafting regional population policy: cluster-based demographic typology of rural municipalities in the three Baltic countries**
autori Aleksandrs Dahs; Juris Krūmiņš

Paldies par uzmanību!

aleksandrs.dahs@lu.lv

DEMOMIG@lu.lv

<http://www.statistikuasociacija.lv>