

Esimene registripõhine prooviloendus Eestis

Registripõhise rahva ja eluruumide loenduse (REGREL) prooviloenduse käigus katsetati loenduse tunnuste moodustamisel kasutatavate registrite^a kvaliteeti, meetodika toimimist ja loendust toetava tarkvara valmisolekut registripõhiseks loenduseks. Nagu loenduse puhul, nii määratakse ka prooviloenduse jaoks loendusmoment ja moodustatakse selle hetke seisuga loendusandmestik. Tekkivad andmetabelid kontrollitakse ja selgitatakse välja, missugune tase andmetel kehtestatud kriteeriumite ja statistika kvaliteedinõuete kohaselt on. Prooviloenduse käigus avastatud puuduste kõrvaldamiseks võetakse vajalikud meetmed, mille tulemust kontrollitakse teisel prooviloendusel 2019. aastal.

Esimese prooviloenduse tulemus kinnitab Statistikaameti valmisolekut registripõhiseks loenduseks. Loenduse vajadusest on lähtunud registrite olukorra analüüsis ja antud hinnang loendustunnuste kvaliteedile (tabel 4).

Prooviloendus hõlmas Euroopa Liidu liikmesriikidele kohustuslikke loendustunnuseid, välja arvatud tunnuseid „amet“ ja „töökoha asukoht“. Lisaks analüüsiti tunnuseid „rahvus“ ja „emakeel“.

Prooviloenduse õiguslik alus:

- riikliku statistika seadus
- Vabariigi Valitsuse 17. septembri 2015. aasta korraldus nr 391 „Statistikaameti 2015.–2019. aastal tehtavate statistikatööde loetelu“
- Euroopa Parlamendi ja nõukogu 9. juuli 2008. aasta määrus (EÜ) nr 763/2008 rahva ja eluruumide loenduste kohta ja selle rakendamise määrus (EÜ) nr 1201/2009, 30. november 2009 ja määrus (EÜ) nr 519/2010, 16. juuni 2010

Prooviloenduse moment: 31.12.2015

Prooviloenduse tegevuste toimumise aeg: 2.1–8.12.2016

Prooviloenduse üldkogum:

- Eesti alaline rahvastik
- Eestis asuvad eluruumid asustatusest olenemata ja asustatud mitteiluruumid

Prooviloenduse eesmärk

Katsetada rahva ja eluruumide loenduse tegemist registripõhiselt, järgides Eurostati eeskirja ja loenduse kvaliteedinõudeid. Andmeid tuleb hõivata 24 riiklikust andmekogust ning teha andmetöötlus ja -analüüs.

Metoodika

REGREL on registripõhine loendus, see tähendab, et REGRELi andmestik moodustatakse riiklike andmekogude põhjal. Enne loendusandmetena kasutuselevõttu kontrollitakse andmete kvaliteeti põhiliste statistikakriteeriumite alusel. Kui andmed sisalduvad eri registrites, kasutatakse neid ühelt poolt andmete kvaliteedi kontrollimiseks ja teiselt poolt kõige usaldusväärsemate väärtuste valimiseks vastavalt meetodikas välja töötatud eeskirjadele.

^a Dokumendis on registrite all mõeldud riiklikke andmekogusid.

Loendustunnused ei ole üldjuhul üks-üheselt registritest saadavad, sest registrid on ette nähtud muude, mitte statistiliste ülesannete täitmiseks ja registrites kasutatakse enamasti teistsuguseid määratlusi kui statistikas. See tähendab, et mõne loendustunnuse moodustamiseks tuleb kasutada mitme, isegi üle 10 registri andmeid (nt hõiveseisund) ja mõne tunnuse kohta on olemas dubleeriv info mitmes registris.

REGRELi meetodikud on lahendanud loendustunnuste moodustamisel järgmised põhilised ülesanded:

- analüüsinud loendusel kasutatava määratluse ja väärtuste skaala vahekorda registrites kasutatavate määratluste ja väärtuste skaalaga;
- kontrollinud registrite kvaliteeti ja teinud pingutusi selle nimel, et registripidajad puudujäägid kõrvaldaksid;
- määranud iga loendustunnuse puhul kindlaks selle moodustamiseks ja vajaduse korral kontrollimiseks tarviliku registrite hulga või kui mõne tunnuse jaoks vajalikku infot registrites ei ole, siis astunud samme vastava registri või registriosa loomiseks;
- määranud optimaalse eeskirja iga loendustunnuse moodustamiseks registriandmete põhjal ja loonud selleks vajaliku tarkvara, kasutades optimeerimise alusena väljundtunnuse kvaliteeti;
- kui andmelüngad ei võimalda optimaalset eeskirja kasutada, näinud ette strateegia dubleerivate eeskirjade kasutamiseks, samuti mõnel juhul väärtuste imputeerimiseks statistiliste või loogiliste eeskirjade alusel.

Isikute üldkogum

Isikute üldkogumi all mõistetakse riigi elanike arvu ehk rahvaarvu. Selle moodustamisel ei saa tugineda üksnes rahvastikuregistri andmetele, sest eelmine rahva ja eluruumide loendus tõestas, et rahvastikuregister on ülekaetud. Probleemi lahendamiseks töötati välja residentsuse indeks, mis mõõdab inimese aktiivsust registrites, ja selle koondinfo põhjal otsustatakse, kas inimene on Eesti elanik või mitte. Mida aktiivsem on inimene registrite põhjal, s.t mida rohkem on tal nn elumärke registrites, seda suurema tõenäosusega on ta Eesti elanik. Lähemalt saab residentsuse indeksi meetodika kohta lugeda Statistikaameti veebilehelt (<http://www.stat.ee/400790?highlight=residentsus>) ja kvartalikirjast (http://www.stat.ee/valjaanne-2016_eesti-statistika-kvartalikirj-3-16).

Loenduse üldkogumisse kuulus 1 315 944 isikut.

Leibkondade üldkogum

REGRELi esimesel prooviloendusel jagati tavaleibkondadesse 1 254 705 inimest ja asutusleibkondadesse 12 320 inimest (tabel 1). Arvestades tavaleibkonna keskmist suurust, asutusleibkonda kuuluvate inimeste osatähtsust ja nende inimeste arvu, kes jäid leibkondadesse jagamata puuduliku elukohainfo tõttu, on tavaleibkondade kogum alakaetud umbes 20 000 leibkonna võrra.

Tabel 1. Inimesed ja leibkonnad

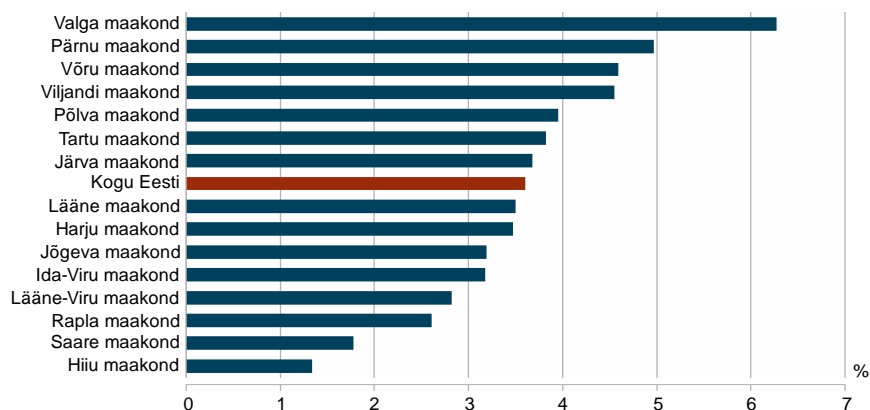
	Inimeste arv	Leibkondade arv
Tavaleibkond	1 254 705	527 443
Asutusleibkond ^a	12 320	–
Elukoht omavalitsus-üksuse tasemel	47 358	19 500 ^b
Elukoht teadmata	1 561	500 ^b
Kokku	1 315 944	

^a Asutusleibkondade liikmeid leibkondadesse ei jagata.

^b Hinnanguline arv.

Inimese alaline elukoht on rahvastikuregistris registreeritud elukoht. Kõigil inimestel aga ei ole elukoht rahvastikuregistris registreeritud eluruumi täpsusega. Maakonniti on omavalitsusüksuse täpsusega elukohaga inimeste osatähtsus erinev (joonis 1). Suurim on nende osatähtsus Valga ja väikseim Hiiu maakonnas. Mida suurem on eluruumi tasemel registreeritud elukohaga inimeste osatähtsus, seda täpsem ja parem on leibkonna ja perekonna tunnuste kvaliteet.

Joonis1. Omavalitsusüksuse täpsusega registreeritud elukohaga elanikud maakonna järgi



Eluruumide üldkogum

Tabelis 2 on eluruumide üldkogumi jaotus eluruumi liigi järgi.

Tabel 2. Eluruumide üldkogum eluruumi liigi ja asustatuse järgi

Eluruumi liik	Eluruumide arv	Asustamata eluruumide arv	Asustatud eluruumide arv	Asustatud eluruumid, %
Tavaeluruumid	725 539	199 074	526 465	72,6
Muud elamuüksused	978	0	978	100,0
Ühiseluruumid	4 359	0	4 359	100,0
Kokku	730 876	199 074	531 802	72,8

Ühiseluruumid ja muud elamuüksused peavad definitsiooni järgi olema asustatud. Tavaeluruumid võivad olla asustatud või asustamata. Et eluruumidesse on jaotatud 96,3% rahvastikust, on asustatud eluruumide puhul üldkogumis tegemist alakaetusega ja asustamata eluruumide puhul ülekaetusega. Hinnanguliselt oleks, juhul kui kõik inimesed oleksid jagatud eluruumidesse, asustatud tavaeluruumide 546 800 ehk 75,4% tavaeluruumide koguarvust.

Tabel 3. Eluruumide infoallikad

Infoallikas	Asustamata tavaeluruumid	Asustatud tavaeluruumid	Muud elamuüksused	Ühiseluruumid	Kokku	Allika osatähtsus, %
ADS	8 625	19 082	479	1 726	29 912	4,1
ADS ja EHR	190 449	507 383	499	2 633	700 964	95,9
Kokku	199 074	526 465	978	4 359	730 876	100,0

4,1% eluruumidest on ainult ADSis, tehnilisi tunnuseid on nende jaoks võimalik otsida ainult REL 2011 andmestikust.

Loendustunnuste kvaliteet

Loendustunnuste kvaliteedi hindamisel tuleb jälgida, kas tunnuse definitsioon vastab rahvusvaheliselt määratud, loenduse väljundtunnuse definitsioonile.

Loenduse puhul on üks põhilisi kvaliteedikriteeriume kaetus. Seda hinnatakse üksiktunnuste puuduvate väärtuste osatähtsuse põhjal. Kui puuduvad väärtusi on soovitud rohkem, tuleb need asendada – imputeerida –, kasutades selleks matemaatilisi või statistilisi eeskirju. Kõige põhjalikumalt on andmete kvaliteeti võimalik kontrollida loendustunnuste jaotuste võrdlemisel mõne teise allika samade tunnuste jaotusega. Selle meetodi puhul on oluline arvestada, et mitmesugustel põhjustel ei tarvitse mõlemas allikas olla ühesugused jaotused: valikuuringutega

võrdlemisel võib erinevus tulla valimiveast, eelmise loendusega võrdlemisel aja jooksul toimunud muutustest, mõne muu (kõikse) allika puhul teistsugustest määratlustest.

Käesolevas REGRELi prooviloenduse kvaliteedi analüüsis hinnati puuduvate väärtuste osatähtsust. Samuti hinnati määratluste vastavust määruse määratlustele, enamikul juhtudest oli vastavus täielik. Kooskõla võrdlusallikaga hinnati sobivalt tuletatud näitaja abil.

Üksiktunnuste keskmine hinne viiepallisüsteemis on 3,4, s.t rahuldav, kusjuures isikuandmete keskmine hinne on 3,9 ning eluruumi- ja leibkonnaandmete keskmine hinne 3,1. Seitsme tunnuse kvaliteeti hinnati kasinaks, need on elukohaga seotud tunnused (tabel 4).

Peale üksiktunnuste kvaliteedi, mille hindamise aluseks olid objektiivsetele kriteeriumitele tuginevad eksperdi-hinnangud, hinnati ka väljundkuupide^a kvaliteeti. Kuupide keskmine kvaliteedihinne skaalal 0–1 on 0,73. Hinnat ei arvestatud kuuel kuubil, mis sisaldasid kaht loendustunnust, mida prooviloendusel ei moodustatud. 13 kuubil 41 väljundkuubist on hinne väga hea ja 17-l rahuldav. Viie kuubi kvaliteedihinne on kasin – need on väikese tunnuste arvuga kuubid, milles on info elukoha asukoha ja leibkonna/perekonna tüübi kohta. Kõiki kriteeriume arvestades on REGRELi tunnuste kvaliteet rahuldav, seejuures isikutunnuste kvaliteet hea. Et tegemist on alles esimese prooviloendusega, on tulemus igati vastuvõetav. Kõige suuremat tööd tuleb loenduseni jäänud aastate jooksul teha elukohaga seotud tunnustega.

Kokkuvõte

REGRELi esimene prooviloendus on edukalt tehtud. Prooviloendus näitas, et registripõhine rahva ja eluruumide loendus on teostatav ja loenduse ettevalmistus on olnud sihipärane.

Esimese prooviloenduse kõik üldkogumid ja prooviloenduse eesmärgiks seatud kohustuslikud tunnused on moodustatud. Suurim töö on tehtud üldkogumite metoodika väljatöötamisega. Inimeste üldkogumi moodustamiseks on loodud residentsusindeks. Eluruumide üldkogumi moodustamiseks kasutatakse ADSi ja EHRi kombinatsiooni, sest EHR on alakaetud. Asustatud eluruumide ja tavaleibkondade arv on alakaetud, sest 3,7% inimestel ei ole elukoht teada eluruumi täpsusega.

Pärast REL 2011 on suurim samm REGRELi ettevalmistuses olnud ADSi standardi kasutuselevõtt. See on teinud elukoha ja eluruumide sidumise oluliselt lihtsamaks.

REGRELi esimesel prooviloendusel kasutati 24 registri andmeid, millest 3 registri (RR, ADS, KOPIS) omad hõivati automaatselt ja 21 registrit andmebaasi väljavõttena. Enim andmekogusid on vaja inimeste üldkogumi ehk kogurahvastiku moodustamiseks (tabel 5).

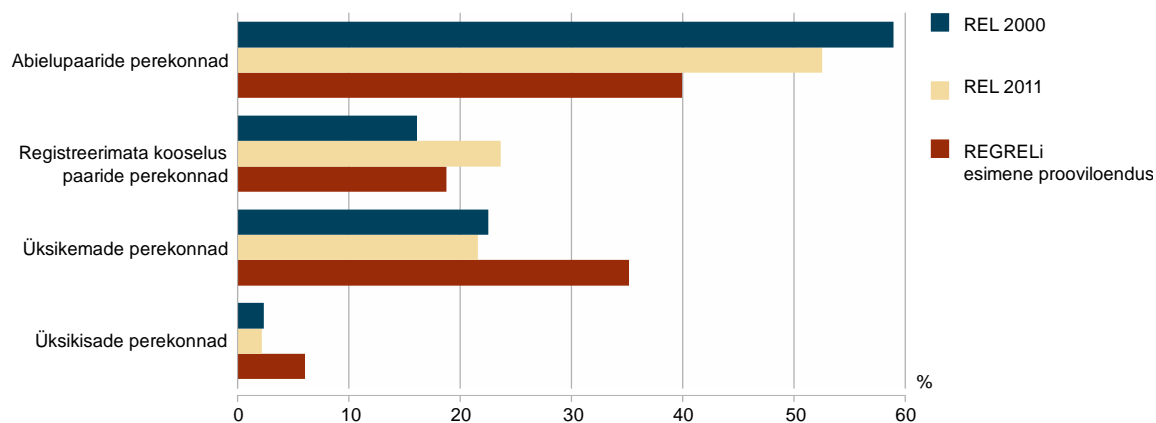
Kontrollitud on loendustunnuste kvaliteedi vastavust kvaliteedikriteeriumitele (Euroopa Parlamendi ja nõukogu määruse (EÜ) nr 763/2008 järgi). Moodustatud 38 loendustunnusest on parima kvaliteediga järgmised: sugu, vanus, seaduslik perekonnaseis, sünniriik, kodakondsus, kogurahvastik, rahvus, emakeel, eluruumi asukoht, eluruumid hoonetüübi järgi (tabel 4).

Hinnatud on REGRELi kohustusliku väljundi (hüperkuupide) kvaliteeti prooviloenduse kvaliteedikriteeriumite kohaselt. 13 kuubil 41 väljundkuubist on hinne väga hea ja 17-l rahuldav.

REGRELi suurim probleem on registreeritud ja tegeliku elukoha erinevus. See mõjutab alalise elukoha madalaima taseme (omavalitsusüksus) jaotust ning kõiki leibkonna ja perekonnaga seotud tunnuseid (joonis 2). Võimalus olukorda parandada, s.t tegelikke ja registreeritud andmeid kokku viia, on motiveerida elanikke oma tegelikke aadresse rahvastikuregistris registreerima. Sellele võib kaasa aidata kohalike omavalitsuste ja meedia teavitustöö.

^a Loenduse väljund esitatakse kuupidena, mis on 6–8-mõõtmelised tabelid. Kuubi kvaliteedihinne arvutati kuupi kuuluvate üksiktunnuste hinnete keskmisena.

Joonis 2. Perekonnad tuumperekonna tüübi järgi



Tabel 4. Tunnuste kvaliteedi koondhinne

Hinne	Tunnus
5 (väga hea)	sugu, vanus, seaduslik perekonnaseis, sünniriik, kodakondsus, kogurahvastik, rahvus, emakeel, eluruumi asukoht, eluruumid hoonetüübi järgi
4	töölane staatus, haridustase, eluruumi tüüp, elanike arv eluruumis, kasulik põrandapind, eluaseme tubade arv, veevarustussüsteem, tualettruum, küttesüsteem, eluaseme asustustiheduse standard
3 (rahuldav)	alaline elukoht aasta enne loendust, hõiveseisund, majandusharu, riiki saabumise aasta, seisund leibkonnas, tavaleibkonna suurus, tuumperekonna suurus, elukorraldus, tavaeluruumide valdusõigus, vannituba
2	eluruumid ehitusaja järgi
1 (kasin)	alaline elukoht, leibkonna valduse staatus, paikkond, seisund perekonnas, tuumperekonna tüüp, tavaleibkonna tüüp, omandi liik

Tabel 5. Andmekogud^a ja loendustunnused

Tunnus	RR	EHR	ADS	PKR	KIR	TOR	EHIS	EMSR	MKR	ARIREG	EMPIS	ETR	KIRST	KR	RKOARR	SAP	STAR	E-toimik	KMAIS	KOPIS	KVKR	Liiklusregister	RETS	SPR	REL 2011
Alaline elukoht																									
Sugu																									
Vanus																									
Seaduslik perekonnaseis																									
Sünniriik ja -koht																									
Kodakondsus																									
Elukoht aasta enne loendust																									
Leibkonnaliikmete seosed																									
Töökoha asukoht ^b																									
Hetke hõiveseisund																									
Amet ^b																									
Majandusharu																									
Tööalane staatus																									
Haridustase																									
Riiki saabumise aasta																									
Leibkondade valduse staatus																									
Kogurahvastik																									
Paikkond																									
Seisund leibkonnas																									
Seisund perekonnas																									
Tuumperekonna tüüp																									
Tuumperekonna suurus																									
Tavaleibkonna tüüp																									
Tavaleibkonna suurus																									
Rahvus																									
Emakeel																									
Elukorraldus																									
Elamispinna tüüp																									
Elamispinna asukoht																									
Tavaeluruumide valdusõigus																									
Omandi liik																									
Elanike arv																									
Põrandapind / tubade arv																									
Veevarustussüsteem																									
Tualettruum																									
Vannituba																									
Küttetüüp																									
Eluruumid hoone tüübi järgi																									
Eluruumid ehitusaja järgi																									
Asustustiheduse standard																									

^a Andmekogude lühendite selgitused on lisas 1.

^b Töökoha asukohta ja ameti kohta saab tulevikus andmed töötamise registrist (TÖR).

Lisa 1. Lühendid

ADS	Aadressiandmete süsteem
ARIREG	Äriregister
EHIS	Eesti Hariduse Infosüsteem
EHR	Riiklik ehitisregister
EMPIS	Töötuna ja töotsijana arvel olevate isikute ning tööturuteenuste osutamise register
EMSR	Meditsiiniline sünniregister
ETR	Riiklik elamis- ja töölubade register
KIR	Vangide ja kriminaalhooldusaluste register
KIRST	Ravikindlustuse andmekogu
KMAIS	Isikut tõendavate dokumentide andmekogu
KOPIS	Kohustusliku kogumispensioni register
KR	Kinnistusraamat
KVKR	Kaitseväekohustuslaste register
MKR	Maksudkohustuslaste register
PKR	Riiklik pensionikindlustuse register (PKR-nimelist registrit RIHAs enam pole, on SKAIS – Sotsiaalkindlustusameti infosüsteem)
REL 2000	2000. aasta rahva ja eluruumide loendus
REL 2011	2011. aasta rahva ja eluruumide loendus
RETS	Retseptikeskus
RKOARR	Riigi- ja kohaliku omavalitsuse asutuste riiklik register
RR	Rahvastikuregister
SAP	Riigi personali- ja palgaarvestuse andmekogu
SPR	Surma põhjuste register
STAR	Sotsiaalteenuste ja -toetuste andmeregister
TÖR	Töötamise register