

Eesti Statistika Kvartalikirj

Quarterly Bulletin of Statistics Estonia

Intervjuu Statistikaameti
peadirektoriga

Mõtteid arvamusfestivalidelt

Suurandmed statistikas

Kaupade ja teenuste
väliskaubandus

Maa hind

Balti loenduse seminarid

4/2017



Eesti Statistika Kvartalikirj

Quarterly Bulletin of Statistics Estonia

4/2017

EESTI STATISTIKA
STATISTICS ESTONIA

TALLINN 2017

MÄRKIDE SELETUS

EXPLANATION OF SYMBOLS

...	andmeid ei ole saadud või need on avaldamiseks ebakindlad <i>data not available or too uncertain for publication</i>
..	mõiste pole rakendatav <i>category not applicable</i>
x	andmete avaldamist ei võimalda andmekaitse põhimõte / <i>data are confidential</i>
M/M	mehed/ <i>males</i>
N/F	naised/ <i>females</i>

Väljaandes on kasutatud Statistikaameti andmeid, kui ei ole viidatud teisiti.

The publication is based on Statistics Estonia's data, unless specified otherwise.

Toimetuskolleegium / *Editorial Council*: Birgit Hansson, Riina Kerner, Toomas Kirt, Marika Kivilaid, Märt Leesment, Robert Mürsepp, Evelin Puura, Taimi Rosenberg, Mihkel Servinski, Eve Telp, Ene-Margit Tiit

Toimetanud Liis Haugas, Taimi Rosenberg, Kairit Saar
Inglise keelde tõlkinud Kati Coleman, Pille Peensoo, Kai Takkis, Triangular OÜ
Kaane kujundanud Uku Nurges
Küljendanud Nele Lumiste

Edited by Liis Haugas, Taimi Rosenberg, Kairit Saar
Translation into English by Kati Coleman, Pille Peensoo, Kai Takkis, Triangular OÜ
Cover by Uku Nurges
Layout by Nele Lumiste

Kirjastanud Statistikaamet,
Tatari 51, 10134 Tallinn
Trükkinud Ofset OÜ,
Paldiski mnt 25, 10612 Tallinn

Detsember 2017

Published by Statistics Estonia,
Tatari 51, 10134 Tallinn
Printed by Ofset OÜ,
Paldiski mnt 25, 10612 Tallinn

Detsember 2017

ISSN-L 1736-7921
ISSN 1736-7921 (trükis / *hard copy*)
ISSN 2346-6049 (PDF)

Autoriõigus/*Copyright*: Statistikaamet, 2017

Väljaande andmete kasutamisel või tsiteerimisel palume viidata allikale.

When using or quoting the data included in this issue, please indicate the source.

SISUKORD

Uudisnopeid statistika vallast	4
I Mart Mägi: Statistikaameti pakutav informatsioon on usaldusväärne	8
Intervjueerinud Mihkel Servinski	
II Mõtteid arvamusefestivalidelt	26
Märt Leesment, Mihkel Servinski	
III Suurandmed statistikas	33
Toomas Kirt	
IV Kaupade ja teenuste väliskaubandus ettevõtlusnäitajate järgi	59
Allan Aron	
V Maa hind: kui palju on Eestimaa väärt?	74
Reelika Parve, Robert Mürsepp	
VI Balti loenduse seminar tähistas 2017. aastal 20. aastapäeva	89
Diana Beltadze	
Noppeid Eesti statistika ajaloost	100
Põhinäitajad	104
Eesti statistika kvartalikirjas avaldatud artiklid 2009–2017	112

CONTENTS

News picks from the field of statistics	6
I Mart Mägi: the information offered by Statistics Estonia is reliable	17
Mihkel Servinski	
II Thoughts from the Opinion Festivals	29
Märt Leesment, Mihkel Servinski	
III Big data in statistics	48
Toomas Kirt	
IV Foreign trade of goods and services by enterprise characteristics	69
Allan Aron	
V Land value: how much is Estonia worth?	84
Reelika Parve, Robert Mürsepp	
VI The Baltic Census Seminar celebrated its 20th anniversary in 2017	95
Diana Beltadze	
Picks from the history of Estonian Statistics	100
Main indicators	104
Articles published in the Quarterly Bulletin of Statistics Estonia, 2009–2017	112

UUDISNOPPEID STATISTIKA VALLAST

Robert Mürsepp

Nopete allikas on uuemad Eurostati andmed^a.

Veerand EL-i kodumajapidamiste kulutustest läheb eluasemele

2016. aastal kulutasid Euroopa Liidu kodumajapidamised ligikaudu 2 triljonit eurot üürile, veele, elektrile, gaasile ja muudele kütustele. See moodustab 13,3% EL-i SKP-st ja peaaegu veerandi (24,5%) kodumajapidamiste kogukulutustest. See on ka kodumajapidamiste suurim väljaminekuallikas. Samuti tasub märkida, et võrreldes 2006. aastaga on selle kulurühma osatähtsus kõigis väljaminekutes EL-is keskmiselt suurenenud 2,3 protsendipunkti võrra.

Suurim on eluasemekulude osatähtsus kogukuludes Taanis (29,1%). Järgnevad Soome (28,2%), Suurbritannia (27,1%), Prantsusmaa (26,2%) ja Rootsi (26%). Erinevalt meie põhjanaabritest on Eestis eluasemekulude osatähtsus üks EL-i väiksemaid – 17,6%. Veel pisem on see näitaja vaid Maltal (10,3%), Leedus (15,6%) ja Küprosel (15,9%). Kui eluasemekulude osatähtsuse kasv iseloomustab enamikku EL-i maid, siis ka siin on Eesti üks vähestest, kus on aset leidnud vastupidine. Viimasel kümnendil on eluasemekulude osatähtsus Eestis vähenenud 0,5 protsendipunkti. Suurim langus on aga olnud Slovakkias – 2,3 protsendipunkti. Sarnaselt Eestile on kogunud kuni 1 protsendipunkti vähenemist Poola, Malta, Saksamaa ja Rootsi. Enim on eluasemekulude osatähtsus samal perioodil suurenenud Portugalis – 4,5 protsendipunkti, järgnevad Hispaania (4,3), Soome (4,1) ja Iirimaa (3,8).

Kodumajapidamistes on suuruselt teine kuluallikas transpordikulud, mis 2016. aastal moodustasid 12,9% kogukuludest. Suurim on transpordikulude osatähtsus Sloveenias (15,9%), järgnevad Luksemburg (15,4%) ja Leedu (15,1%). Väikseima transpordikulude osatähtsusega riigid on Slovakkia (7,5%), Tšehhi (9,8%), Belgia (11%) ja Hispaania (11%). Eestis on transpordikulude osatähtsus 11,3% ja sellega kuulume taas väiksema osatähtsusega riikide hulka.

Kolmas olulisem kuluallikas on kodumajapidamistel toit ja mittealkohoolsed joogid, mis EL-is keskmiselt moodustasid 2016. aastal 12,2% kõigist tarbimiskulutustest. Selles valdkonnas on Eesti koos teiste Baltimaadega kõrgeimate kulutuste osatähtsusega riikide seas. Kõige suurem osatähtsus toidu ja mittealkohoolsete jookide kulutustel on Leedus (22,2%), järgnevad Eesti (20,3%), Bulgaaria (19,5%) ja Läti (18,2%). Väikseim kulutuste osatähtsus on aga Suurbritannias (8,1%), Iirimaal (9,1%), Luksemburgis (9,4%) ja Austrias (9,7%).

Pool EL-i sisesest kaubavahetusest tuleb väikestest ja keskmise suurusega ettevõtetest

Väikesteks ja keskmise suurusega ettevõtete (VKE) loetakse ettevõtteid, kus on alla 250 töötajat. Kuigi majandusuudistest võib sageli lugeda eelkõige suurettevõtete kohta, siis nii töökohtade kui ka majandusliku väärtuse loomise poolest ei ole VKE-d vähemtähtsad. 2015. aastal moodustasid nad 98% kõigist EL-is kaubavahetusega tegelevatest ettevõtetest. Sellest umbes kaks kolmandikku on koguni mikroettevõtted (0–9 töötajat). VKE-de EL-i sisesest kaubavahetuse väärtus moodustab omakorda poole kogu kaubeldavate kaupade väärtusest. Ka on igas liikmesriigis VKE-de osatähtsus EL-i siseses kaubavahetuses üle 90%. VKE-de vahendatavate kaupade väärtus kogu kaubanduses varieerub riikide võrdluses aga kõvasti.

EL-i sisesest ekspordist moodustavad VKE-d 45% (mikroettevõtted 13,2%). Suurim on VKE-de osatähtsus kogu eksporditud kauba väärtuses Küprosel (88%) ja Lätis (81%). Ka järgnevad on

^a <http://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20171121-1>.
<http://ec.europa.eu/eurostat/documents/2995521/8467294/6-21112017-AP-EN.pdf/ab123dd2-0901-4c64-95b0-f071a5f235e3>.

väikeriigid – Belgia (70%), Eesti (68%) ja Holland (67%). Väikseim VKE-de osatähtsus ekspordi koguväärtuses on Prantsusmaal (21%), Saksamaal (26%), Slovakkias (30%) ja Iirimaa (32%). Mikroettevõtete puhul pole erinevused küll nii suured, kuid siiski märkimisväärsed. Kui VKE-de osa on suurim Belgias (46%) ja Maltal (37%), siis vähim panustavad need Prantsusmaal, Saksamaal ja Soomes – kõigis riikides ligikaudu 4%. Eestis panustavad mikroettevõtted 17% kogu eksporditud kaupade väärtusest.

VKE-de osatähtsus EL-i siseses impordis sarnaneb ekspordiga. Suurim on VKE-de osatähtsus kogu impordi väärtuses Lätis (85%), Küprosel (82%), Eestis (79%) ja Leedus (78%). Alla poole jääb VKE-de panustatud kaubavahetuse väärtus vaid viies riigis – Prantsusmaal (31%), Saksamaal (34%), Tšehhis (47%), Slovakkias (48%) ja Poolas (49%). Ka impordis on mikroettevõtete panus suurim Maltal (45%) ja Belgias (38%). Väikseimaks jääb see Prantsusmaal, Saksamaal ja Tšehhis – kõigis riikides alla 7%. Eestis on mikroettevõtete panus kogu EL-i siseses impordis 26% kogu kaupade väärtusest.

NEWS PICKS FROM THE FIELD OF STATISTICS

Robert Mürsepp

The picks are based on recent Eurostat data^a.

A quarter of household expenditure of EU is allocated to housing

In 2016, households of the European Union (EU) spent nearly 2 trillion euros on housing, water, electricity, gas and other fuels. This is equivalent to 13.3% of the EU GDP and forms almost a quarter (24.5%) of the total expenditure of households. This is by far the most significant expenditure of EU households. It is also worth noting that compared to 2006, the share of housing expenditure in total household expenditure in the EU has increased on average by 2.3 percentage points.

The households of Denmark devoted the largest share of their total expenditure to housing (29.1%). Denmark is followed by Finland (28.2%), the United Kingdom (27.1%), France (26.2%) and Sweden (26%). Unlike our northern neighbours in Finland, the share of housing expenditure in Estonia is one of the lowest in the EU – 17.6%. This indicator is lower only in Malta (10.3%), Lithuania (15.6%) and Cyprus (15.9%). If the vast majority of Member States can be characterised by the growth of the share of housing expenditure, Estonia is one of the few countries where the opposite has occurred. In the past decade, the share of housing expenditure has dropped in Estonia by 0.5 percentage points. The largest drop – 2.3 percentage points – has occurred in Slovakia. Similarly to Estonia, the share of housing expenditure in total household expenditure also dropped by up to 1 percentage point in Poland, Malta, Germany and Sweden. In the same period, the share of housing expenditure increased the most in Portugal – 4.5 percentage points, followed by Spain (4.3 pp), Finland (4.1 pp) and Ireland (3.8 pp).

The second largest source of costs is expenditure on transport, forming 12.9% of the total expenditure. The share of transport expenditure is highest in Slovenia (15.9%), followed by Luxembourg (15.4%) and Lithuania (15.1%). Countries with the lowest share of transport expenditure are Slovakia (7.5%), the Czech Republic (9.8%), Belgium (11%) and Spain (11%). In Estonia, the share of transport expenditure is 11.3%, leaving us once again among the countries with a lower share of expenditures.

The third most significant expenditure of households is food and non-alcoholic beverages, forming on average 12.2% of the total consumption expenditure in the EU in 2016. In this sector, Estonia and the other Baltic States are among the countries with the highest share of expenditure. The highest share of expenditure on food and non-alcoholic beverages is in Lithuania (22.2%), followed by Estonia (20.3%), Bulgaria (19.5%) and Latvia (18.2%). The lowest share of expenditure was in the United Kingdom (8.1%), Ireland (9.1%), Luxembourg (9.4%) and Austria (9.7%).

SMEs in the European Union generate half of the intra-EU trade in goods

Small and medium sized enterprises (SMEs) are companies with up to 249 persons employed. Although economic news often cover the activities of large corporations, SMEs are as important in creating jobs and economic value. In 2015, SMEs formed 98% of companies trading goods within the EU. Around two-thirds of them are micro-enterprises (0–9 employees). The value of intra-EU trade of SMEs forms a half of the value of the total traded goods. Also, in all Member States, the share of SMEs in intra-EU trade is over 90%. The value of goods mediated by SMEs in total trade, however, varies greatly among different countries.

^a <http://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20171121-1>.
<http://ec.europa.eu/eurostat/documents/2995521/8467294/6-21112017-AP-EN.pdf/ab123dd2-0901-4c64-95b0-f071a5f235e3>.

SMEs form 45% of the intra-EU exports (micro-enterprises 13.2%). In the intra-EU exports of goods, the share of SMEs in the value of exported goods is largest in Cyprus (88%) and Latvia (81%). They are followed by another set of small countries – Belgium (70%), Estonia (68%) and the Netherlands (67%). The share of SMEs in the total value of exports is lowest in France (21%), Germany (26%), Slovakia (30%) and Ireland (32%). Differences in the shares of micro-enterprises are not that large but still noteworthy. If the share of SMEs is largest in Belgium (46%) and Malta (37%), then in France, Germany and Finland they contribute the least – approximately 4% in each listed country. In Estonia, micro-enterprises account for 17% of the total value of exported goods.

The share of SMEs in intra-EU imports is similar to that of intra-EU exports. Considering the total value of imports, the largest contributors are the SMEs of Latvia (85%), Cyprus (82%), Estonia (79%) and Lithuania (78%). The contribution of SMEs to the value of trade remains under 50% only in five countries – France (31%), Germany (34%), the Czech Republic (47%), Slovakia (48%) and Poland (49%). In intra-EU imports, the micro-enterprises of Malta (45%) and Belgium (38%) contribute the most. The smallest contributors are France, Germany and the Czech Republic (all under 7%). In Estonia, the contribution of micro-enterprises to intra-EU imports is 26% of the total value of goods.

MART MÄGI: STATISTIKAAMETI PAKUTAV INFORMATSIOON ON USALDUSVÄÄRNE

Statistikaameti peadirektor Mart Mägi on ametis olnud veidi üle poole aasta. Statistikaamet on talle esimene töökoht avalikus sektoris. Peadirektorile esitab küsimusi Statistikaameti peaanalüütik Mihkel Servinski, enamik küsimusi on pärit Statistikaameti töötajatelt.

Milline on sinu arvates Eesti statistikaameti omapära võrreldes teiste Euroopa riikide statistikaametitega? Kas me üldse peaksime teistest statistikaametitest eristuma, ja kui, siis mille poolest?

Eesti statistikaameti eristumine teistest statistikaametitest ei ole otseselt vajalik. Oluline on, et Statistikaamet Eesti riigi andmete agentuurina annaks Eesti ühiskonnale ja otsuste vastuvõtjatele vajalikku teavet kiiresti ja et teave oleks usaldusväärne. Meie omapära on, et Eesti statistikaamet on juba praegu üks Euroopa kõige efektiivsemaid statistikaameteid. See, et oleme efektiivsed, et Eesti riik on väike, annab meile ainukordse võimaluse katsetada uusi andmemajanduse mudeleid, kui oleme selleks valmis ja julged seda tegema.

Statistikaameti visioonis on mõte, et oleme kõige efektiivsem, innovaatilisem ja tarbijasõbralikum usaldusväärse statistika tegija. Kas mõnes valdkonnas võiks Statistikaamet olla Euroopa parim? Kas meil on selline ambitsioon?

Jah, ambitsioon on kõigis nendes neljas valdkonnas. Meie visioon on väga selge: tahame olla kõige efektiivsem statistikaamet Euroopas; tahame olla Euroopa kõige innovaatilisem.

Efektiivsusest oli eespool juba juttu ja selles oleme eesmärgile üsna lähedal. Innovatsioonis, kui vaatame, milliseid andmeallikaid saaksime veel hõivata ja kuidas saaksime infot tarbijatele paremini edastada, on meil veel üsna pikk tee minna.

Usaldusväärseusega on nii, et ebausaldusväärset statistikat ei tohiks üldse olemas olla. Praegu võime öelda, et oleme teistega võrdselt usaldusväärsed. Usaldusväärseusega tuleb loomulikult edasi töötada, aga eesmärk pole olla kõige usaldusväärsem, sest kõik statistikaametid peavad tegema usaldusväärset statistikat. Kõik Euroopa statistikaametid lähtuvad Euroopa statistika kvaliteedi juhendist.

Tarbijasõbralikkuse puhul saame taas rääkida parimaks olemisest. Hinnangu sellele, kui tarbijasõbralikud me oleme, annavad statistika tarbijad ja meie partnerid statistika tegemisel.

Statistikaameti tarbijasõbralikkuse kõige tähtsamaks hindajaks pean Eesti avalikkust. Pean siin silmas kolme sihtrühma. Esimene on avaliku huvi esindajad, ministeeriumid ja kohalikud omavalitsused, kellele peame tagama otsustamiseks piisavalt aja- ja asjakohast usaldusväärset infot. Teine rühm on teadlased ja analüütikud, kellele peame suutma pakkuda aja- ja asjakohaseid usaldusväärseid andmeid kuni mikroandmete tasandini, järgides muidugi andmekaitse-nõudeid. Kolmas sihtrühm on laiem avalikkus, kellele peame samuti pakkuma infot avalikkust huvitavatel teemadel.

Oleme ka senini kõigile sihtrühmadele teavet pakkunud ja Statistikaametil pole põhjust häbeneda hinnanguid meie tarbijasõbralikkusele, aga et senist üsna kõrget hinnangut hoida ja ka parandada, peame toime tulema uute ülesannetega. Küsimus on selles, kas iga inimene saab personaalselt seda infot, mida ta tahab võimalikult lihtsalt ja võimalikult kiiresti saada. Andmete koguhulk kasvab meeletu kiirusega. Järjest keerulisem on sellest hulgast just mulle, just praegu vajalikku teavet leida. Olen veendunud, et Statistikaamet peab lähiaastatel selle teemaga senisest märksa rohkem tegelema.

Toon näitena Amazoni: kui oled korra Amazonis käinud, siis uuel sisenemisel tunneb keskkond sind ära ja suhtleb sinuga kui vana tuttavaga, pakkudes infot sinu huvidest lähtuvalt. Selles peab Statistikaamet arenguhüppe tegema.

Statistika mainega ühiskonnas on nii ja naa. Väita, et see väga hea oleks, on liialdus. Väita, et ühiskond statistikat ei hinda, oleks samuti liialdus. Milline on statistika maine Eestis sinu hinnangul? Kas saaksid olukorda Eestis võrrelda olukorraga teistes riikides?

Tegelikult peame siin vaatama kahte asja: statistika mainet ja Statistikaameti mainet. Need on kaks omavahel seotud, aga samas täiesti erinevat teemat.

Eestis on riikliku statistika maine, mida mõõdetakse sageli usaldusväarsusega, paljudele ehk üllatavalt väga hea. Tarbijauuringute tulemused kinnitavad, et meie pakutavad näitajad on tegelikult väga hinnatud: ajakirjanduses, avalikkuses ja otsustajate seas.

Statistikaameti maine on halvem, kuid miks? Statistikaamet ongi ju peamine riikliku statistika tegija Eestis ja kui meie pakutavat kõrgelt hinnatakse, oleks ootuspärane, et ka pakkuja on kõrgelt hinnatud. Küllap see nii ka oleks, kui Statistikaameti maine oleneks vaid meie pakutavast – riiklikust statistikast. Tegelikult kujundavad ühiskonna hinnangut Statistikaametile suures osas andmeesitajad ehk Statistikaameti olulised partnerid. Andmete esitamine Statistikaametile on andmeesitajatele töö, mis pole otseselt seotud nende põhitööga, vaid on üks riigi kehtestatud kohustus. Kui Statistikaameti tekitatud halduskoormus tundub andmeesitajale ülemäära suur ja andmete esitamisest tagasi saadu väärtus liiga väike, pole põhjust loota, et andmeesitaja Statistikaametile kõrge positiivse hinnangu annaks.

Selge on, et praegu pole Statistikaametil võimalik andmeesitajatele tekitatud halduskoormust päris ära kaotada, aga selle nimel tuleb tegutseda. Kui andmeesitajad tunnetavad Statistikaameti jõupingutusi, siis küllap paraneb ka nende hinnang meile. See tähendab Statistikaametile vajadust teha pidevalt tööd selleks, et muuta andmekorjet efektiivsemaks, andmete esitamist mugavamaks ja andmeesitajatele tagasiantu väärtust suuremaks. Viimast ei tohi mingil juhul unustada: halduskoormuse vähendamisest ei piisa, kui meie tehtav statistika pole usaldusväärne ja vajalik. Vajalikkuse juures on oluline rääkida riikliku statistika tegemise kiirusest ja kiiruse suurendamise tähtsusest: kui praegu kogutud andmete põhjal suudame infot anda alles kahe aasta pärast, ei ole see tänapäeval ootustele vastav andmete töötlemise kiirus, sest kahe aasta pärast on praegused andmed juhtimisotsuste tegemiseks lootusetult aegunud. Kui meie pakutav info on vananenud, on loomulik, et Statistikaametit nähakse vananeva ja vanamoodsa organisatsioonina.

Tegelikult tegeleb Statistikaamet täiesti tänapäevaste, võib öelda, et kõige moodsamate andmehõive meetoditega. Selle näiteks on mobiilpositsioneerimise andmed, elektritarbimise andmed reaajas jne. Peeglisse vaadates tuleb kahjuks tunnistada, et nendest meie tegemistest teatakse Eestis suhteliselt vähe.

Kas see tähendab, et Statistikaamet peaks olema avatum, oma tegemistest ja plaanidest rohkem rääkima?

Jah, peaksime avatumad olema! Peaksime tegema rohkem koostööd ministeeriumidega, kohalike omavalitsustega ja eriti teadlastega. Tänapäeva andmemailmas ei saa keegi, sh Statistikaamet hakkama, kui kavatseb kõike ise ja iseene tarkusest teha. Statistikaamet peab senisest enam jagama teadmisi valdkondades, kus ta on kompetentne, ja kuulama teisi valdkondades, kus omal kompetentsi napib. Teiste kuulamine ja teistega senisest rohkem arvestamine on kindlasti vajalik.

Statistikaamet töötab Eesti riigi ja rahva heaks. Peamine ei ole, et Statistikaametil läheks hästi, vaid see, et Eesti riigil oleks olemas vajalik informatsioon. Statistikaamet ei pea mõtlema üksnes sellele, et meie tehtav statistika oleks kvaliteetne, vaid peab kaasa aitama sellele, et ministeeriumides, kohalikes omavalitsustes, riiklikes registrites jm pakutav statistiline info oleks usaldusväärne. See tähendab koostööd. See tähendab, et Statistikaamet ei võta kogu statistika tegemist oma kõhu alla, vaid võimaldab ka teistel teha häid analüüse.

Selge on, et Statistikaametil on iseenda tarbeks infot vaja vaid sellel määral, kui asutuse juhtimiseks vajalik on. Kogu riiklik statistika tehakse Eesti ühiskonna, mitte Statistikaameti tarbeks. Riikliku statistika tegemise kava kinnitab Vabariigi Valitsus. Samas on vale öelda, et Statistikaamet ei osale riikliku statistika programmi kokkupanemisel. Kuidas

hindad, kas Statistikaameti roll riikliku statistika programmi kokkupanekul on piisav? Kas ministeeriumid, kes enamasti on statistikavaldkondades peamised avaliku huvi esindajad, pole meid riikliku statistika programmi koostamisest kõrvale tõrjunud? Kas ministeerium, paljuski poliitilise huvi esindajana, on piisavalt motiveeritud lähtuma riikliku statistika programmi kokkupanekul demokraatliku ühiskonna huvidest?

Viimased kümme aastat kindlasti ja tõenäoliselt kauemgi on Statistikaamet töötanud riikliku statistika programmi järgi. See on ajalooliselt kujunenud programm, mis muutub vaid siis, kui mõnel avaliku huvi esindajal on väga suur, lausa kriitiline huvi seda muuta. Oleme jõudnud olukorda, kus pakume teiste hulgas ka näitajaid, mis on selgelt vananenud, ega paku kõiki näitajaid, mida muutunud maailmas vaja oleks, sest riikliku statistika programmi muutmise pole piisavalt paindlik. Kuidas seda olukorda muuta? Oleme jõudnud väga selge arusaamani, et kahte liiki näitajaid peab Statistikaamet kindlasti pakkuma: rahvusvahelisi näitajaid, mis on kokku lepitud organisatsioonides, kuhu Eesti riik vabatahtlikult kuulub, ning näitajaid, mis on kokku lepitud avaliku sektori arengukavades püstitatud eesmärkide seireks. Kui avaliku huvi alusel luuakse arengukava ja selles seatakse eesmärgid, on ülimalt oluline, et nende eesmärkide täitmist ka objektiivselt ja usaldusväärsetl mõõdetaks. Võib öelda, et praegu on avaliku sektori arengukavades püstitatud eesmärkide täitmise jälgimises vajakajäämisi. Peeglisse peab vaatama Statistikaamet ja ehk veel mõni teinegi asutus.

Avaliku sektori arengukavade eesmärkide täitmise jälgimine on loomulikult tähtis, aga ehk peaks Statistikaamet senisest märksa aktiivsemalt osalema ka püstitatud eesmärkide mõõdikute väljatöötamisel.

Jah, kindlasti! Tegelikult viib ka see teema selleni, et Statistikaameti roll ühiskonnas peab muutuma ja selle muutumise eelduseks on Statistikaameti senisest oluliselt suurem koostöö ja parem partnerlus avaliku huvi esindajatega riikliku statistika tööde tellimisel. Olen veendunud, et Statistikaamet peab osalema avaliku sektori arengukavade väljatöötamisel, pakkuma lahendusi, kuidas mõõta püstitatud eesmärgid ja kuidas nende eesmärkide täitmist jälgida. Veelgi olulisem aga on, et annaksime arengukavade koostajatele andmeid ja infot arengukavade eesmärkide seadmiseks, et Statistikaametis tehtavad analüüsid pööraksid julgelt tähelepanu ühiskonnas toimuvale, et suudaksime statistika abil sõnastada ühiskonna tugevusi ja kitsaskohti. Statistikaameti käes on väga suur hulk andmeid ja meil on võimalus märgata ühiskonnas toimuvaid muutusi nende algaasis. Oluline on see võimalus realiseerida ja esitada tulemused nii usaldusväärsetena, et need saaksid ka tegelikult otsuste aluseks.

Intervjuu alguses ütlesin, et Statistikaamet peab saama Eesti riigi andmeagentuuriks. See tähendab, et Statistikaamet peab suutma otsa vaadata riigi andmearhitektuurile, andmestandarditele ja nõustama andmekogusid nende standardite rakendamisel. See on osa Statistikaameti muutuvast rollist Eesti ühiskonnas: Statistikaamet ei ole ainult kokkulepitud näitajate pakkuja, vaid peab olema ka avaliku sektori andmete kogumise, töötlemise ja sidumise koordineerija.

Statistikaametil on suur potentsiaal sotsiaaluuringute tegemiseks. Seda potentsiaali aga ei kasutata täiel määral ja nii teevad paljusid riigis vajalikke sotsiaaluuringuid ettevõtted, kus puudub vajalik meetodiline ja analüütiline võimekus ning kus uuringu kvaliteeti suhtutakse väga pinnapealselt. Statistikaameti probleem on, et me ei ole piisavalt paindlikud kiirete tellimuste täitmiseks ega piisavalt paindlikud tellija soovide rahuldamiseks. Me ei oska turul piisavalt hästi oma võimekusest teada anda ja nii jäetakse Statistikaameti poole pöördumata ka nendel kordadel, kus pöördumine oleks igati põhjendatud. Kas sellises kirjelduses on tõde?

Jah. On olnud juhtumeid, kus oleme teinud pakkumise nelja kuuga olukorras, kus tegelikult oleme võimelised seda tegema nelja päevaga. Statistikaameti üks probleeme on, et me ei ole proaktiivsed. Statistikaamet ei ole proaktiivne ministeeriumides, teistes avaliku sektori asutustes ega ka eraettevõtluses. Reageerime, kui ministeerium meie poole pöördub ja küsib, kas teeksime pakkumise. Tegelikult peaksime olema olukorraga kursis nii hästi, et suudaksime ise pakkuda ministeeriumile (ka erialaliitudele) teemasid, mida on vaja uurida, peame suutma välja tuua

küsimused, millele ei saa olemasoleva riikliku statistika abil vastata, aga millele vastamata jätta samuti ei saa. Ja olgem ausad, me räägime, et meil on võimekus, et meie töö on kvaliteetne, et see on meie eelis võrreldes paljude teistega, kes analoogset teenust pakuvad, aga ... Ma ei sea kahtluse alla nende väidete õigsust, ent kui me pole suutnud Statistikaameti eeliseid selgelt esile tuua ja võimalikele tellijatele selgeks teha, siis maksavad meie tugevused suhteliselt vähe. Meil võivad olla selged kvaliteedikriteeriumid, võime kvaliteedinõudeid suurepäraselt täita, aga kui see ei lähe tellijale korda, oleme pildilt väljas. Peame oma rolli tagasi saama. See on üks meie eesmärke.

Toon näite Euroopa statistikapäeva konverentsilt, kus Eesti Panga esindaja ütles selgelt, et usaldab Maksu- ja Tolliameti andmeid rohkem kui Statistikaameti omi. Kuidas siis nii? Miks on tekkinud olukord, kus üks riigi institutsioon ütleb, et usaldab ühe institutsiooni andmeid rohkem kui teise omi? Põhjus on levinud ettekujutus, et kui andmete esitamine registrisse või avalikule institutsioonile on tugevalt survestatud kohustus, siis ollakse andmete esitamisel täpsed. Statistikaametile andmete esitamine on samuti seadusest tulenev kohustus, aga see kohustus on väga nõrgalt survestatud ja seepärast arvatakse, et Statistikaametile võib esitada andmeid nii, nagu parasjagu jõuab kirja panna. Kas see nii on, seda oskavad ettevõtjad paremini hinnata, aga küsigem, kas Maksu- ja Tolliameti ning Statistikaameti tehtavat statistikat võrreldes võtavad hinnangu andjad arvesse, et mitte kõik Eesti ettevõtted ei pea Maksu- ja Tolliametile maksudeklaratsioone esitama. Kui käive on alla 30 000 euro, ei pea esitama käibemaksudeklaratsiooni; kui ettevõttes pole töötajaid, ei pea esitama töötamise deklaratsiooni jne. Aus vastus on, et sellised asjad unustatakse ära. Tegelikult Maksu- ja Tolliamet ei arvuta välja, mis on selle n-õ pika saba ehk puuduolevate andmete mõju ühiskonnas. Need arvutused teeb Statistikaamet. Aga miks Eesti Pank ütleb, et Maksu- ja Tolliameti andmed on usaldusväärsemad? On üks väga lihtne põhjus – Eesti Pank saab endale vajalikud andmed Maksu- ja Tolliametist kiiremini ja kompaktsemalt. Hinnangut ei anta kahjuks mitte selle alusel, kui usaldusväärne mingi näitaja tegelikult on, vaid selle põhjal, kui kiiresti ja mugavalt andmed kätte saadakse. See on ohtlik tendents, aga Statistikaamet peab siin peeglisse vaatama, sest me ise pole suutnud olukorda selgitada ja ka andmete töötlemise kiiruses on meil selge reserv olemas.

Jah, aga kindlasti on siin ka teisi tahke. Kui Statistikaamet küsib asju, mille kohta ettevõtte arvestust ei pea, saame statistilise aruande, mida andmeesitajate seas nimetatakse rahvaluuleks. Näiteks ei pea ettevõtte arvestust selle üle, kas autokütus osteti Soomest, Rootsist, Eestist või Lätist, sest see ettevõtjat ei huvita ja Statistikaameti pärast arvepidamises muudatust ei tehta. Nii et kütuse kogukulu on õige, aga selle jaotus riigiti väljamõeldis.

Aga miks me seda küsime? Statistikaamet ei küsi andmeid statistikute isikliku huvi rahuldamiseks. Statistikaamet küsib andmeid (peamise) avaliku huvi esindaja tellimusel, mille on kinnitanud Vabariigi Valitsus. Üsna sageli juhtub, et avaliku huvi esindaja on oma soovi andmete saamiseks esitanud, aga tegelikult ta Statistikaameti kogutud andmete põhjal saadud infot ei vaja. Näpuga näidatakse muidugi Statistikaameti peale. Ei ütle, et selleks põhjust poleks, sest andmeid kogub ikkagi Statistikaamet, aga probleemi lahendus ei olene ainult Statistikaametist.

Toon näite piirkondliku statistika edendamise töörühma tööst. Üks ministerium soovis infot isadusmäära ja naiste sünnitusvanuse kohta omavalitsusüksuste kaupa. Põnevad näitajad mõlemad, aga kas neid on ikka vaja nii detailselt? Arutelu käigus selgus, et ei ole. Tegelikult on see positiivne näide, sest koostöös selgus tegelik vajadus.

Laiemas tähenduses on eelmine küsimus üks näide sellest, milles Statistikaamet peaks olema proaktiivsem. Peame senisest märksa paremini aru saama, milleks avaliku huvi esindaja näitajat vajab, ja kui meie analüüs näitab, et tegelikult näitajat ei vajata, peame julgema teha ettepaneku selle avaldamine lõpetada. Võib olla ka vastupidi: kui ühiskonnas on tekkinud põhjendatud huvi mõne näitaja vastu, peab Statistikaamet peamise avaliku huvi esindajaga koostöös hakkama seda näitajat pakkuma.

Kuigi Statistikaamet ei kogu andmeid iseendale, peame suutma andmeesitajatele selgitada, miks me neid kogume, ja seda hoopis täpsemalt kui viitega Vabariigi Valitsusele, kes riikliku statistika programmi kinnitab.

Sinu jutust tuleb selgelt välja, et hoolimata sellest, et oleme efektiivselt toimiv asutus, on muudatuste tegemine töökorralduses vajalik. Kas ülesanded on püstitatud?

Ülesanded on püstitatud. Hakkame peale sellest, et täpsustame, kes ja kuidas Statistikaameti avaliku huvi esindajatega suhtlevad. Kui Statistikaamet avaliku huvi esindajatega proaktiivselt ei suhtle, siis kuidas saame tellimustöid või kuidas saame olla veendunud, et meie tehtav riiklik statistika on see, mida ühiskond tegelikult vajab? Praegu on meil väljaspool riiklikku tellimust vähe töid, kolm kuni neli tööd kuus. Kui arvestada võimalike tellimuste hulka, peaks neid olema viis kuni kümme korda rohkem. Võib-olla veelgi rohkem, aga alustame sellestki.

Suhtlemisoskus on omaette keeruline kunst. Hea statistik ei pruugi olla hea suhtleja. Ainult heast suhtlemisoskusest ei piisa avaliku huvi esindajatega suhtlemisel, vaja on ka erialaseid oskusi. Kas Statistikaameti on mõeldud sellele, kuidas spetsialistide suhtlusoskust parandada?

Oleme juhtkonnas arutanud, kes võiks olla ministeeriumis meie suhtluspartner ja kes võiks olla Statistikaameti esindaja. Kipume minema seda teed, et meie esindaja on tippspetsialist ja on loogiline, et ta ka suhtleb ministeeriumi tippspetsialistiga. Kindlasti on vaja, et Statistikaameti esindaja oleks hea suhtleja ja veelgi enam – hea kommunikatsiooniekspert. Hea spetsialist saab teha konkreetset statistikatööd, aga ta ei pruugi osata saavutada kokkulepet, et see töö üldse kavva võetaks. Kindlasti on vaja mõelda, kuidas selle teemaga edasi minna, ja koolitusel on edasiminekus oma kindel koht.

Teen sinu senisest jutust järelduse, et väga suurt põhjust nuriseda selle üle, kuidas Eesti ühiskond Statistikaametisse suhtub, ei ole, kuigi arenguruumi on. Räägime nüüd Statistikaameti kõige olulisemast partnerist andmeesitajana ehk ettevõtjast. Kas ettevõtja suhtumine Statistikaametisse väärriks eraldi tähelepanu?

Väärrib ja väga mitmest aspektist. Käisin hiljuti ühe andmeesitaja juures. Tegemist on tuntud Eesti kontserniga, kes esitab aastas Statistikaametile 440 aruannet. See tähendab, et kontsern esitab rohkem kui ühe aruande päevas. Nende tegemiseks kulub poolteist kuud ühe inimese tööd. Kui lisame siia vajaduse esitada andmeid ka Maksu- ja Tolliametile, Eesti Pangale jt-le, kulub riigile andmete esitamiseks veelgi rohkem aega. Riigile andmete esitamise kohustus tõenäoliselt veel niipea ära ei kao, aga andmekogumismeetodid peavad kindlasti tänapäevaseks ja efektiivseks muutuma ning andmeesitaja peab ka ise neid muutusi tajuma.

Teine aspekt on andmekorjel veel. Ettevõtja on nõus andmeid esitama ja tegelikult saab enamik ettevõtjaid selle kohustuse vajalikkusest aru, aga ta tahab ka midagi vastu saada. Loomulikult saab ta vastu riikliku statistika, aga ettevõtja soovib ka midagi personaalset.

Statistikaamet peab mõlema aspektiga tõsiselt tegelema. Andmekogumisega tegeleme praegu väga aktiivselt, näiteks halduskoormuse vähendamise projektis „Aruandlus 3.0“. Püüame saada kogu andmekorje, v.a küsitlused, masinatevahelise liidestuse peale.

Tagasiside puhul oleme vähem aktiivselt tegutsenud. Statistikaameti eelmine juhtkond on selles minu arvates head tööd teinud ja loonud head eeldused edasiminekuks, aga see, kuidas luua personaliseeritud infovaateid andmeesitajatele ja teistele riikliku statistika tarbijatele, on üks meie suurimaid ülesandeid praegusel strateegiaperioodil.

Rõhutan veel kord, et keegi ei ole kurtnud, et nad peavad andmeid esitama. Kuna teeme usaldusväärset statistikat, siis meid tunnustatakse selle eest. Ma ei saa öelda, et meie maine ettevõtjate või üldsuse silmis oleks negatiivne. Statistikaameti mainet kujundab peamiselt see, kuidas suudame korraldada andmete kogumise, ja märksa vähem see, mida me avaldame. Meid tunnustatakse selle eest, et oleme seadnud sihiks andmekogumist lihtsustada ja tarbija-sõbralikumaks teha.

Statistikaamet on hakanud rakendama sunniraha ettevõtetele, kes pole andmeid esitanud. Oleme seda teinud peaaegu pool aastat ja seetõttu olen tõesti väga paljudele andmeesitajatele helistanud. Mitmega olen ka kohtunud. Ükski neist pole kurnud, et peab andmeid esitama, aga küsimusi on esitatud küll. Väga hea küsimuse esitas näiteks Büroomaailma omanik ja juht Jüri Ross. Ta ütles, et saab aru kõigest, välja arvatud sellest, miks me küsime harilike pliiaatsite ja täidetava sisuga harilike pliiaatsite kohta eraldi, mitu tükki ta neid toonud on. Harilik pliiaats on harilik pliiaats ja inimesed ei osta sinna juurde täiendusi. Kui inimesel on harilik pliiaats, siis ta kulutab sisu lõpuni ja viskab pliiaatsi ära.

Kas sinu juhtimisfilosoofias on midagi sellist, mida sinu meelest teistel juhtidel ei ole? Kui jah, siis mis see on? Kui ei, siis milline on sinu juhtimisfilosoofia kõige olulisem punkt?

Ma ei usu, et minu juhtimisfilosoofias oleks midagi sellist, mida ühelgi teisel juhil pole. Kõige olulisem punkt on see, mida näen oma rollina juhtimises.

Minu ülesanne organisatsioonis on suurendada organisatsiooni väärtust. Mida ma selle all mõtlen? Erasektoris koosneb asutuse väärtus viiest, avalikus sektoris neljast osast. Ühtegi osa ei saa teise arvelt muuta. Esimene on väärtus tarbijatele. Tegelikult loome tarbijate jaoks väärtust kogu aeg, minu ülesanne on seda väärtust suurendada. Teine on väärtus meie partneritele (erasektoris hankijad). Partneritele, kes annavad sisendi (nt andmed), tuleb toota lisandväärtust. Kolmas on väärtus töötajatele. See võib olla mitmel kujul, mitte ainult palgatõus, vaid ka uhkus, maine ja tunnus. Neljas osa on väärtus riigile ja ühiskonnale ning viies on väärtus omanikule. See viimane väärtus ühtib avalikus sektoris neljanda väärtusega.

Need väärtused siin on esitatud juhuslikus järjekorras. Kui tahta neid prioriteetide järgi reastada, siis see oleneb organisatsioonist ja sellest, kus organisatsioon konkreetsel hetkel asub.

Mulle on tähtis suurendada organisatsiooni koguväärtust, väärtust kõigile osalistele, võtmata kelleltki midagi ära. Kui mõtled, et see juhtimisfilosoofia on väga lihtne, siis on see eksitav. Kiusatus suurendada ühe osa väärtust teise arvelt on suur. Kui ma 1990ndate alguses ärisse tulín, vahetasin kord töökohta, sest selle omanikud soovisid ühte väärtust suurendada teise arvelt, aga see ei lähe kokku minu väärtushinnangutega.

Küllap on sinulgi olnud juhuseid, et kohtud head muljet jätva inimesega, kes tegelikult on tühi kõmisev kest, ja inimesega, kes end eriti väljendada ei oska, aga kelle ideed on sisutihedad. Võib-olla on ka Statistikaametiga nii, et oleme sisemiselt palju tugevamad, kui ühiskonnale ennast näidata oskame.

Ma olen sinuga nõus: valik võib olla väga keeruline. Mul on su küsimusele keeruline vastata ehk ka seepärast, et olen aastaid olnud olukorras, kus olen pidanud panema kokku juhtide meeskonda. Niisuguse meeskonna kokkupanek erineb oluliselt konkreetset tööd tegeva meeskonna kokkupanekust. Juhtide puhul on mul aga väga selge arusaam sellest, mida nende juures hindan.

Esimene asi, mis juhil peab olema, on visioon. Kui juhil oma valdkonnas oma visiooni ei ole, siis tegelikult ei muutu midagi, kõik jääb nii, nagu on. Juht võib küll kuuletuda, aga see, et ta vaid kuuletub, tähendab, et ta on koormav organisatsiooni ja kõigi teiste jaoks, et tegelikult on ta tuulelipp. Juhil peab olema visioon.

Juhil peab olema ka julgust otsuseid teha. Ei saa olla nii, et ta otsustab alles siis, kui kõik on üksmeelselt nõus. Juht peab võtma vastutuse ja tegema otsuseid.

Kolmandaks peab juhil olema võime asju ära teha ja teha seda tulemuslikult.

Mulle meeldib inimestele öelda, mis on meie eesmärk, ja jätta tegevusviis – kokkulepitud eesmärkide raamides muidugi – vabaks tegija enda jaoks. Inimestel on erinevad käekirjad, üks teeb ühtpidi ja teine teistpidi, aga oluline on saavutada tulemus, mis muidugi ei tohi väärtushinnangutega vastuollu minna. Tähtis on, et saavutaksime oodatud tulemuse.

Mulle meeldib, kui minu meeskond on võimalikult eripalgeline, sest niisuguses meeskonnas tehakse palju vähem vigu kui kindla, ühesuguse mõtteviisiga liikmetest koosnevas meeskonnas. Aga meeskond peab suutma koos töötada ja meeskond suudab koos töötada, kui

meeskonnaliikmetel on ühesugused väärtushinnangud. Oleme Statistikaametis väärtushinnangute teemat seni veel vähe puudutanud, aga hakkame nendest üha rohkem rääkima ja inimeste värbamisel sellele üha suuremat tähelepanu pöörama.

Milline on sinu hinnang kaadrivoolavusele Statistikaametis ja milline on Statistikaameti senine suhtumine voolavusse? Kui mineja ütleb, et läheb kõrgema palga peale, siis mõistame teda ja selle puu taga metsa ei näe. Madal palk pole ju ainus kaadrivoolavuse põhjus.

Ma olen sinuga täiesti nõus, see on väga tähtis küsimus. Vahemärkusena olgu öeldud, et mina olen selle põlvkonna esindaja, kes ütleb, et ühes organisatsioonis töötamise paras kestus – eriti tegevjuhil – on kolm kuni viis aastat, sest selle ajaga ei jõua veel mugavustsooni sattuda. Aga räägime nüüd muudest lahkumise põhjustest.

Lisaks palgale on töökoha valimisel sageli olulised kolm omavahel tihedasti seotud asja, mis tihti on ka lahkumise põhjused: kas töö on huvitav, kas organisatsiooni meeskond sobib mulle ja milline on organisatsiooni maine.

Huvitav töö saab otsa, tekib rutiin. Tahaks teha midagi põnevamat, aga ei saa, sest peab tegema oma igapäevaseid rutiinseid tegevusi. See on kindlasti probleem ka Statistikaametis ja sellega peab pidevalt tegelema. Meil on tööprotsesse, mida uuendame kümne aasta tagant. Kui võtame praegu inimese tööle ja paneme ta näiteks FoxProga töötama, võib olla üsna kindel, et ta lahkub meilt kiiresti.

Meeskond. Kas meie meeskonna liikmed toetavad üksteist? Kas meie juhid toetavad töötajaid? Juhtide toetus meeskonnale on väga oluline ja algab sellest, kas juhil on aega oma meeskonda toetada. Viimase struktuurimuudatusega andsime tiimijuhtidele õigused, vastutuse ja voli oma meeskondi juhtida. Ent ei ole võimalik juhtida meeskonda, kus on 140 inimest. Juht ei suuda sel juhul süveneda töötaja probleemidesse ja talle tuge pakkuda. Keeruline on mõelda, millised võiksid olla töötaja uued ülesanded ja milles teda koolitada. Seepärast reorganiseerisime tiiminduse Statistikaametis.

Maine. Loomulikult tahavad organisatsiooni liikmed tunda, et töötavad edukas organisatsioonis. Kui see ei ole teadlik soov, siis alateadlikult soovib seda enamik inimesi. Mis teeb ühest organisatsioonist eduka organisatsiooni? Üks asi on organisatsiooni väline maine, teine aga meie oskus märgata ja tähistada väikesi võite. Kui oskame tähistada oma võite, kui teeme seda koos oma partneritega, ka siis loome positiivset mainet. Seda suunda peame kindlasti rohkem arendama.

Oled Statistikaametis töötanud juba üle poole aasta. Mulle tundub, et üks oluline paradigma muutus on Statistikaametis juba toimunud või siis kohe-kohe toimumas. Siiani oleme otsinud vastust küsimusele, miks üht või teist asja teha ei saa, aga nüüd otsime vastust küsimusele, kuidas üht või teist vajalikku oleks võimalik ära teha. Kas olen olukorrast õigesti aru saanud?

Jah, täiesti õigesti. Mina olen avalikus sektoris tööl esimest korda. Alguses oli minu jaoks kõige šokeerivam erinevus avaliku ja erasektori vahel see, et siin kõik rääkisid, et meil ei ole raha, et me ei saa midagi teha. Erasektoris ei ole samamoodi raha, aga keegi ei räägi sellest, vaid üritatakse Eesti talupoja kombel leida mingi viis, kuidas raha saada, kuidas raha teha. Kuidas teha, kui seda raha pole või pole piisavalt? Kas annab mõne tegevuse ära jätta, et rohkem vajalikke asju teha? Jah, oled olukorrast õigesti aru saanud.

Tuletan meelde Statistikaameti kvartalikirja selle aasta kolmandat numbrit, kus Eesti mõttes suure ettevõtte juht Emõke Sogenbits ütleb: „Kui jätab inimesed unarusse, on asjade käik ettearvatu.“ Kas Statistikaametis on inimesed sinu hinnangul unarusse jäetud?

Ma arvan, et osas valdkondades on, aga usun, et liigume kiiresti selle poole, et kõik meie asutuse teemandid ja kalliskivid üles leida, et ükski meie inimene poleks unarusse jäetud.

Möte, et robotid võtavad kord maailma üle, ei tundu enam täiesti ulmeline olevat. Milline on sinu prognoos, millal võtavad robotid üle riikliku statistika tegemise? Milline on viimane töö, mille robotid Statistikaametis üle võtavad, ja kui kaua on Statistikaameti arsenalis silmast silma tehtavad küsitlused?

Viimasest küsimusest alustades ütlen, et Statistikaamet ei tee praegu ühtegi kvalitatiivset uuringut. Tee, mille valinud oleme, on see, et hakkame peale kvantitatiivsete uuringute tegema ka kvalitatiivseid, mis tähendab, et silmast silma tehtavad küsitlused kaduma ei hakka. Ma ei ütle, et niisuguste küsitluste osatähtsus riikliku statistika tegemisel kasvaks, aga kindlasti on kvantitatiivsete uuringute asendumine kvalitatiivsetega üks suund. Silmast silma küsitlemine annab infot, mida muud moodi pole võimalik saada, aga mida me praegu veel ei kasuta. Peame õppima seda kasutama.

Küsimuse esimese osa juurde tulles ütlen, et oleme Statistikaametis alustanud töid, mille käigus paljud tegevused muutuvad automaatsemaks. Näiteks kopeerime praegu sageli andmeid ühest Exceli failist teise, mis on üsna mõttetu, aga samas aega nõudev tegevus. Selliste tegevuste käsitsi tegemine peab kaduma, need tuleb asendada tehniliste lahendustega.

Mis siis Statistikaametisse jääb? Kõigi robotiteemaliste arenduste juures pole praeguseks veel suudetud välja mõelda lahendust inimeste asendamiseks algoritmide väljamõttelisel.

Viimasel ajal on Statistikaametis toimunud reform reformi otsa. Sageli pole reform veel lõpule jõudnud, kui algab reformi muutmise reform. Üksikult võttes on iga muudatus olnud ehk ka põhjendatud, kuid olukord, kus uus reform algab enne, kui vana on saanud võimaluse tulemusi näidata, tekitab segadust. Kuidas sina reformidesse suhtud?

Evolutsioon on olemuselt väga pikk protsess. Revolutsioonid kiirendavad protsessi oluliselt. On etappe, kus organisatsioon vajab revolutsioonilist arengut. Ühest küljest püüame praegu Statistikaametis teha võimalikult vähe muudatusi, aga teisest küljest, arvestades ühiskonna tugevat survet riikliku statistika tegemisele, on mõnes statistika tegemise lõigus, näiteks statistika tegemise tehnoloogias, vaja teha väga kiireid muudatusi. Seda muudatust ei saa aastate peale venitada. Vabandust, aga kolm või viis aastat andmelao tegemist ei ole normaalsus.

Muutused ei tähenda seda, et kõik tuleb pea peale pöörata. Ma lootan, et mina ja ka minu järglased suudavad Statistikaametis säilitada kõik selle, mis on siin väärtuslikku, et säilivad positiivsed traditsioonid. Uuel juhil on kindlasti mõtet jätkata eelmise juhi pooleli jäänud vajalikke muudatusi. Minu eelkäijad Priit Potisepp ja Andres Oopkaup on teinud väga head tööd. Olen seda korduvalt öelnud ja olen selles sügavalt veendunud. Mina olen siin selleks, et säilitada kõik Statistikaameti väärtuslikud traditsioonid ja seda väärtuslikku säilitades teha muudatusi, et Statistikaameti kui organisatsiooni koguväärtust võimalikult kiiresti kasvatada.

Praegused muutused Statistikaametis võivad mõne inimese jaoks tunduda väga kiiretena, mõnele väga aeglastena. Statistikaametis ei tee peadirektor otsuseid üksinda. Otsuseid teeb juhtkond, kuhu kuuluvad kõik ameti oluliste valdkondade eestvedajad. Me vaidleme, arutame mõnd otsust mitu korda. Mõned küsivad, miks see protsess peab nii pikk olema. Teised jällegi imestavad, et oi-oi, miks nii kiiresti, eriti siis, kui otsus inimest isiklikult puudutab. Tasakaalu otsimine on lõpmatu teema ja arutelude käigus ei tohi kordagi unustada eesmärki saavutada oodatud tulemus ja et organisatsiooni liikmetele läheks asi paremaks.

Kas olen õigesti aru saanud, kui väidan, et sinu arvates pole erilist vahet riigiasutuse ja erasektori ettevõtte juhtimises?

Juhtimisprintsipiidid on erasektoris ja avalikus sektoris samad. Selles mõttes pole vahet. Asutuse või ettevõtte kõige suurem vara on inimesed ja meeskond. Kõige olulisem on, kuidas me asju ära teeme ja tulemuse saavutame.

Jäänud on üks küsimus, aga enne seda minu traditsiooniline küsimus, mida olen senini esitanud kõikidele, kellega mul on olnud au intervjuud teha. Kas mõni oluline küsimus jäi esitamata? Kui jah, siis mis see on ja kuidas kõlab vastus sellele?

Otseselt ei jäänud ükski küsimus esitamata. Arvestades, kui keeruliseks on võimalik ühte teemat statistika kontekstis mõelda, siis arusaamise võti peitub sageli hoopis selles, et asju on vaja lihtsustada. Kui suudame asju lihtsamaks teha, on meil kergem oma sõnumit edastada.

Oli kunagi üks Rooma riigimees (senaator Cato Vanem), kes kõik oma kõned lõpetas sõnumiga, et Kartaago tuleb hävitada. Lõpuks Kartaago hävitatigi. Milline on Statistikaameti sõnum Eesti ühiskonnale? Mis peaks pidevalt kõlama?

Ma ei vasta lühidalt. Soovitan lugeda Yuval Noah Harari raamatut „Homo Deus: A Brief History of Tomorrow“. Raamatut pole kahjuks veel eesti keelde tõlgitud, aga tegemist on kultusraamatuks saanud teosega. Ma ei hakka raamatut ümber jutustama, aga raamatu huvitav mõte on, et inimkonna valitsev usk pole mitte kristlus, islam jne, vaid inimese valitsevaks usuks on kujunenud dataism ehk usk andmetesse. Usume masinate langetatud otsustesse rohkem kui kunagi varem. Kui tahame näiteks sõita Järva-Jaani, paneme Google'i kaarti sihtkoha kirja ja sõidame selle järgi, uskudes, et see tee on parem, kui meie enda valitud. Kui usaldame masinaid ja andmeid, tekib küsimus, kas olukorras, kus maailm kubiseb valeuudistest ja valeandmetest, on masinate usaldamine kõige targem tegu.

Statistikaameti sõnum võiks olla, et meie majas andmete töötlemiseks kasutatavad algoritmid ja meie pakutav informatsioon on usaldusväärsed.

Intervjuu on tehtud 15. novembril 2017.

MART MÄGI: THE INFORMATION OFFERED BY STATISTICS ESTONIA IS RELIABLE

Mart Mägi, the director general of Statistics Estonia, has been in office for a little over six months. Statistics Estonia is his first work place in the public sector. The director general is interviewed by Mihkel Servinski, a principal analyst of Statistics Estonia. The majority of questions originate from the employees of Statistics Estonia.

In your opinion, what is different about Statistics Estonia compared to other national statistical offices in Europe? Should we even distinguish ourselves from other statistical offices, and if so, how?

The distinction between Statistics Estonia and other statistical offices is not that necessary. It is important that Statistics Estonia as the data agency of the Estonian state would quickly provide the Estonian society and decision-makers with necessary and reliable information. A difference of Statistics Estonia is that we are already one of the most efficient statistical offices in Europe. Our efficiency and the small size of Estonia gives us a unique opportunity to experiment with new data management models, as long as we are willing and brave enough to do so.

The vision of Statistics Estonia includes the idea that we are the most efficient, innovative and user-friendly producer of reliable statistics. Is it possible for Statistics Estonia to be the best in Europe in any of these areas? Do we have such an ambition?

Yes, there is ambition in all four areas. We have a very clear vision: we want to be the most efficient statistical office in Europe, we want to be the most innovative in Europe.

Efficiency was already discussed above, and in this regard, we are pretty close to the objective. As for innovation, we still have a rather long way to go in terms of considering what other data sources we could employ and how to improve the dissemination of information to users.

As for reliability, unreliable statistics should not even exist. Today, we can say that we are as reliable as any other statistics provider. Naturally, reliability requires continuous work, but the goal here is not about being the most reliable, because all statistical offices must produce reliable statistics. All European national statistical institutes follow the quality guidelines for European statistics.

Speaking of user-friendliness, we can once again talk about being the best. Our level of user-friendliness is rated by users of statistics and our partners in producing statistics.

In my opinion, the Estonian public is most important in rating user-friendliness of Statistics Estonia. Here, I am referring to three target groups. The first one consists of representatives of public interest, ministries and local governments whom we must provide sufficient, up-to-date and relevant reliable information, so that they could make decisions. The second group is researchers and analysts whom we must provide up-to-date and relevant reliable data to the micro data level, naturally in compliance with data protection requirements. The third target group is the general public, who also expects us to offer information on the topics of public interest.

So far we have managed to offer information to all target groups, and Statistics Estonia has no reason to be ashamed of the ratings of our user-friendliness, but in order to maintain and perhaps even improve the high ratings, we must be able to cope with new tasks. The question is whether each person can personally receive the requested information as easily and quickly as possible. The total amount of data increases at a frantic pace. It is more and more difficult to retrieve information that is relevant for the particular person at the particular moment. I am convinced that in the coming years this issue requires considerably more attention from Statistics Estonia.

Let's use Amazon as an example: once you have visited Amazon, the website will recognize you when you visit it again and it interacts with you as if you were an old acquaintance, offering you information based on your interests. Statistics Estonia should pursue this development direction.

Statistics has a so-so reputation in the society. It would be an overstatement to say that it has a great reputation, but at the same time, we cannot say that the society does not appreciate statistics at all. What do you think is the reputation of statistics in Estonia? Can you compare the situation in Estonia and in other countries?

Actually, we need to look at two things here: the reputation of statistics and the reputation of Statistics Estonia. These two are interrelated, but they are still completely different topics.

The reputation of official statistics, which is often measured by reliability, is very high in Estonia, which might come as a surprise to many. User surveys confirm that the indicators we produce are actually very much appreciated by the press, public and decision-makers.

The reputation of Statistics Estonia is worse, but why? After all, Statistics Estonia is the primary producer of official statistics in Estonia, and considering the high ratings to what we offer, it would only make sense that the producer of such statistics is highly appreciated as well. Perhaps it would be so, if the reputation of Statistics Estonia depended only on what we offer – official statistics. In reality, the rating of Statistics Estonia by the society is largely shaped by data providers, i.e. Statistics Estonia's important partners. For respondents, submitting data to Statistics Estonia is a task that is not directly related to their main job but an obligation imposed by the state. If the administrative burden caused by Statistics Estonia seems excessive to data providers and the gain from submitting data too small, there is no reason to expect that respondents would give a very positive rating to Statistics Estonia.

It is clear that currently it is impossible for Statistics Estonia to reduce the administrative burden of respondents to zero, but we have to work in that direction. If respondents perceive the efforts made by Statistics Estonia, probably their rating to us would improve. For Statistics Estonia, this means continuous work to increase the efficiency of data collection, convenience of data submission and value returned to data providers. The latter must never be forgotten: reducing administrative burden is not enough if the statistics produced are unreliable or unnecessary. As for necessity, it is important to talk about the speed of producing official statistics and the importance of increasing it: if it takes two years to give information based on the data collected today, we do not meet today's expectations for data processing speed, as in two years these data will be hopelessly outdated for making any governance decisions. If the information we produce is outdated, it is natural that Statistics Estonia seems like an aging and old-fashioned organisation.

Actually, Statistics Estonia uses contemporary or even the most up-to-date methods of data collection. Examples include mobile positioning data, real-time data on electricity consumption, etc. Unfortunately, we have to admit that this is not well-known in Estonia.

Does this mean that Statistics Estonia should be more open and share its activities and plans more?

Yes, we should be more open! We should cooperate more with ministries, local governments and especially with researchers. In the modern data world, no one – including Statistics Estonia – can do everything on one's own and by using only one's own knowledge. Statistics Estonia must share its knowledge more than before in the areas where it has competence and listen to others in the matters where its own competence is lacking. Listening to and considering others more is definitely necessary.

Statistics Estonia serves the Estonian state and people. The primary goal does not consist in the wellbeing of Statistics Estonia, it consists in providing the Estonian state with the information it needs. Statistics Estonia not only has to think about producing quality information, but also has to contribute to reliable statistical information offered by ministries, local governments, state registers, etc. This requires cooperation. It means that Statistics Estonia does not keep the production of all statistical information to itself, but allows others to produce good analyses as well.

It is clear that Statistics Estonia needs information for its own purposes only to the extent that is necessary for the purposes of running the organisation. All official statistics are produced for the Estonian society and not for Statistics Estonia. The official statistical

programme is approved by the Government of the Republic. However, it is inaccurate to say that Statistics Estonia does not take part in the preparation of the official statistical programme. In your opinion, is the role of Statistics Estonia in the preparation of the official statistical programme sufficient? Have we not been cast aside by ministries who mainly act as main representatives of public interest for statistical domains? Considering that ministries also represent political interest, are they motivated enough to consider the interests of a democratic society when preparing the official statistical programme?

For the last decade or perhaps even longer, Statistics Estonia has worked according to the official statistical programme. This is a historically established programme, which changes only if a representative of public interest shows great or even critical interest in changing it. We have reached a situation where some of the indicators we produce are clearly outdated and we are not offering all the indicators that are required in the changing world, because we lack flexibility in making changes to the official statistical programme. How can we change this situation? We have come to a clear understanding that Statistics Estonia definitely has to produce two types of indicators: international indicators established by organisations where the Estonian state is a voluntary member, and indicators agreed upon for monitoring the objectives set in public sector development plans. When preparing a development plan and establishing the relevant objectives based on public interest, it is crucial to ensure impartial and reliable measurement of these objectives. It can be said that today there are certain shortcomings in monitoring the achievement of objectives set in the development plans of the public sector. This is an issue that Statistics Estonia and perhaps a few other institutions have to deal with.

Monitoring the achievement of objectives established in public sector development plans is of course important, but perhaps Statistics Estonia should also be more actively involved in developing measures for the established objectives.

Yes, most certainly! Actually, this relates to the topic that the role of Statistics Estonia must change in the society. A prerequisite for such a change consists in significantly more cooperation and better partnership with the representatives of public interest ordering statistical activities. I believe that Statistics Estonia must participate in the preparation of public sector development plans, offer solutions for measuring established objectives and for monitoring the attainment of these objectives. Even more importantly, we should provide the people preparing the development plans with data and information for setting development plan goals; the analyses made by Statistics Estonia should boldly draw attention to current trends in the society; we should be able to use statistics to formulate the strengths and weaknesses in the society. Statistics Estonia holds a very large amount of data and we have a chance to notice changes taking place in the society in their early stages. It is crucial to seize that opportunity and present results with the reliability that allows using them in the decision-making process.

At the beginning of the interview, I said that Statistics Estonia must become the Estonian information agency. This means that Statistics Estonia must be able to face the state's data architecture, data standards and consult databases in applying these standards. This is part of the changing role of Statistics Estonia in the Estonian society: Statistics Estonia does not just offer agreed indicators, but it must also coordinate the public sector data collection, processing and linking.

Statistics Estonia has a great potential for conducting social surveys. This potential is not fully utilised, and thus many social surveys required in the country are carried out by enterprises that lack methodological and analytical capacity and where the attitude towards quality is very superficial. The problem with Statistics Estonia is that we are not flexible enough to respond to urgent orders and satisfy customers' wishes. We are not good enough at marketing our capabilities, and thus potential customers do not place their requests with Statistics Estonia even if they have every reason to do so. Is there any truth in this description?

Yes. There have been cases, where we have made an offer in four months, when we could have made it in four days. One of the problems of Statistics Estonia is that we are not proactive. Statistics Estonia is not proactive in ministries, other public sector entities or the private sector.

We act after the ministry comes to us and asks if we could make an offer. In fact, we should have so much knowledge of the current situation that we could be the ones to suggest to the ministry (or professional associations) the areas that need further research. We should be able to point out the issues that currently cannot be answered by means of available official statistics but which cannot remain unanswered either. Let's be honest, we may talk about having the capability, quality of work, advantage compared to many others who offer similar service, but... I will not question the accuracy of these arguments, but our strengths do not matter much if we are not able to clearly highlight the advantages of Statistics Estonia and explain them to our potential customers. We might have clear quality criteria. We could be excellent at meeting such quality criteria, but if customers do not know that, we are out of the game. We have to get back in the game. This is one of our goals.

Let me give an example from the conference of the European Statistics Day, where a representative of Eesti Pank expressly said that he trusts the data of the Tax and Customs Board more than that of Statistics Estonia. How so? Why have we come to a situation where a state institution expresses greater trust in the data of one institution compared to the data of another? The reason for that is the common assumption that there is greater accuracy if data submission to a registry or public institution is a strictly enforced obligation. While submission of data to Statistics Estonia is also required by the law, such obligation is less enforced, and thus it is thought that you can submit as much data as you can write down at the moment. Perhaps enterprises can tell whether this is the case or not. However, we can ask whether when comparing the statistics produced by the Tax and Customs Board and Statistics Estonia the evaluators take into account that not all Estonian enterprises have to submit tax returns to the Tax and Customs Board. An enterprise with turnover under 30,000 euros is not required to submit a VAT declaration; an enterprise with no employees is not required to submit the employment-related declarations, etc. An honest answer would be that such aspects tend to be forgotten. The Tax and Customs Board does not actually calculate the impact of this "long tail", i.e. missing data, in the society. Such calculations are made by Statistics Estonia. Then why does Eesti Pank refer to better reliability of the data of the Tax and Customs Board? There is a very simple reason for that – Eesti Pank receives necessary data from the Tax and Customs Board more quickly and in a more compact form. Unfortunately, the assessment is not based on the actual reliability of the indicator, but on the speed and convenience of getting hold of information. This is a dangerous trend, but Statistics Estonia has to look in the mirror – we have failed to clarify the situation and we have obvious reserves in the speed of data processing.

Yes, but there are certainly other aspects as well. If Statistics Estonia asks an enterprise to provide data about things that are not recorded by the enterprise, we get a statistical report that the respondents refer to as "fiction". For instance, an enterprise does not keep records of whether motor fuel was bought from Finland, Sweden, Estonia or Latvia, because this is not of interest to the enterprise and they are not going to change their way of keeping records just for the sake of reporting to Statistics Estonia. Therefore, the total fuel consumption is correct, but its distribution by countries is "fiction".

Why do we ask such things? Statistics Estonia does not ask questions to satisfy personal curiosity of its statisticians. Statistics Estonia asks information according to the order of the (main) representative of public interest, approved by the Government of the Republic. It happens rather frequently that the representative of public interest has requested data when it actually does not need the information based on the data collected by Statistics Estonia. However, Statistics Estonia is blamed. I can understand why they do that because Statistics Estonia is still the one to collect the data, but the solution to this problem does not depend on Statistics Estonia alone.

Here is an example of the work of the task force promoting regional statistics. A certain ministry requested information on paternity rate and mother's age at childbirth broken down by municipalities. These are both interesting indicators, but do we really need them at such level of detail? The discussion revealed that this was not necessary. This is in fact a positive example, because the actual need was clarified by cooperating.

In the broader sense, the previous question is an example of where Statistics Estonia should be more proactive. We must understand much better why the representative of public interest needs a particular indicator, and if our analysis shows that such an indicator is actually not needed, we have to be brave enough to suggest that its publishing be discontinued. It could also be the other way around – if the society has developed reasonable interest in a certain indicator, then Statistics Estonia must start offering it in cooperation with the main representative of public interest.

Although Statistics Estonia does not collect data for itself, we must be able to explain to data providers why we collect these data and do it better than just by referring to the Government of the Republic, which approves the official statistical programme.

From what you are saying, it clearly appears that despite being an efficient institution, it is necessary to make changes in the organisation of work. Have the relevant goals been set?

The goals have been set. We start by specifying who and how communicates at Statistics Estonia with the representatives of public interest. If Statistics Estonia fails to be proactive in interaction with the representatives of public interest, how will we get orders or how can we be sure that the official statistics we produce are what the society actually needs. Today, we have few orders, 3–4 activities a month, in addition to what is ordered by the state. When considering the potential number of orders, they should increase five to ten times. Perhaps even more, but it is a start.

Communication is an intricate art. A good statistician is not necessarily a good communicator. Good communication skills are not enough when interacting with the representatives of public interest, it also requires specific professional skills. Has it been considered at Statistics Estonia how to improve the communication skills of specialists?

Our top management has discussed potential communication partners at the ministries and who would represent Statistics Estonia. We tend to take the road that our representative is a senior specialist and it is logical that he or she also communicates with the relevant senior specialist of the ministry. The representative of Statistics Estonia should definitely be a good communicator and not just that – a great communications expert. A good specialist can perform a particular statistical activity, but might not be able to come to an agreement that a particular action is included in the programme in the first place. Further consideration is clearly needed about how to proceed with this topic and training is important for making progress.

Based on what you have said so far, I conclude that there aren't very significant reasons to complain about the attitude of the Estonian society towards Statistics Estonia, although there is still room for improvement. Let's continue our discussion about the most important partner of Statistics Estonia in terms of data submission, i.e. enterprises. Is the attitude of enterprises towards Statistics Estonia worth special attention?

Yes, and in several aspects. I lately visited one respondent, a well-known Estonian enterprise group, which annually submits 440 reports to Statistics Estonia. This means that they submit more than one report a day. It takes one and a half month's work for one person to complete these reports. If we add to this the obligation to submit data to the Tax and Customs Board, Eesti Pank, etc., it takes even more time to submit data to the state. The obligation to submit data to the state will probably not disappear any time soon, but the methods of data collection should definitely become more modern and efficient, and such changes should also be felt by the respondent.

There is another aspect to data collection. Enterprises agree to submit data, in fact most of them understand the necessity of this obligation, but they also want something in return. Sure, they get official statistics, but they want to receive something more personal in return.

Statistics Estonia should work seriously on both aspects. We are currently working very actively on data collection, e.g. as part of the project "Reporting 3.0", which is concerned with reducing administrative burden. We are trying to make data collection, except surveys, entirely based on machine-to-machine interfaces.

We have been less active with regard to feedback. The previous management of Statistics Estonia has, in my opinion, done a good job and created prerequisites for further development, but one of our greatest challenges during this strategic period is how to create personalised information views for respondents and other users of official statistics.

I emphasise once more, nobody has complained about the obligation to submit data. We produce reliable statistics and we are acknowledged for that. I cannot say that we have negative reputation among enterprises or the public. The reputation of Statistics Estonia depends mostly on our ability to organise data collection and to a much lesser extent on what we publish. We are recognised for setting the target of simpler and more user-friendly data collection.

Statistics Estonia has started to impose penalty payments on enterprises that fail to submit reports. We have done this for almost six months and for that reason I have called very many data providers. I have also met several of them. None of them has complained about the obligation to submit data, but they had various questions. For instance, a good question was asked by Jüri Ross, the owner and manager of Büroomaailm. He said that he understands everything, except why we ask separately how many pencils and how many re-fillable pencils he has bought. A pencil is a pencil and people do not buy the refills. If you have a pencil, you use it until the lead is used up and then you throw it away.

Does your management philosophy include anything that you don't think other managers have? If yes, what would it be? If no, what is the most significant concept in your management philosophy?

I do not believe that my management philosophy would include something that no other manager has. The most important is how I perceive my role as a leader.

My role in an organisation is to grow its value. What do I mean by that? There are five parts to the value of an institution in the private sector and four parts in the public sector. None of the parts can be changed at the expense of another. The first is value for users. We actually create value for users all the time, and my task is to increase that value. The second is value for our partners (buyers in the private sector). We have to create value added for the partners who provide input (e.g. data). The third is value for employees. This can occur in several forms, not just as a pay rise, but also as pride, reputation and being known. The fourth part is value for the state and society, and the fifth is value for owner. The latter value is combined with the fourth value in the public sector.

The above values are listed in random order. Listing them by priority may vary depending on the organisation and its current situation.

Perhaps you have also had instances where you have met a person who despite making a good impression turns out to be void of ideas, or a person who is not good at expressing oneself but who has excellent ideas. Perhaps it is the same with Statistics Estonia – we are much stronger on the inside than we can show to the society.

I agree with you, this can be a very complicated choice. Perhaps the reason why it is difficult to answer this question is that for years I have been in a situation where I have had to put together a team of managers. Forming such a team is rather different from putting together a team for performing a particular task. As for managers, I have a clear understanding of the qualities I value in them.

Vision is the first thing that a manager must have. If a manager has no vision in one's area of activity, nothing will actually happen and everything will remain as it is. The manager might obey, but if he or she does nothing else, this is a burden for the organisation and for everyone else, and the manager serves as a weather vane. A manager has to have vision.

A manager must also have courage to make decisions. He or she cannot wait until consensus has been reached. A manager must take responsibility and make decisions.

Thirdly, a manager must be able to make things happen and get results.

I like to tell people what our purpose is and leave it up to them how they do it – naturally, within the agreed framework. People have various ways of working, they do things differently, but it is important to achieve results, provided that it does not conflict with values. What matters is that we achieve the expected results.

I like when my team is as versatile as possible, because such a team does not make as many mistakes as a team consisting of members with one solid way of thinking. However, a team must be able to work together, and it can work together if its members share the same values.

So far we have not discussed values much in Statistics Estonia, but we will talk about them more and pay more attention to that when recruiting people.

What is your opinion on staff turnover at Statistics Estonia and the organisation's position on it so far? When the person leaving says that they leave for better pay, we understand them and do not see a forest behind that tree. Low pay is not the only reason for high staff turnover, though.

I completely agree with you. This is a very important question. I should mention that I represent the generation who believes that the right duration of working for a single organisation – especially for an executive manager – ranges from three to five years, as this time is not long enough to find a comfort zone. But let us look at other reasons for leaving.

Besides pay, three interrelated factors are important when choosing a job, and they are often also the reason for leaving: is the job interesting, does the organisation's team suit you and what is the reputation of the organisation.

Interest in the work eventually ends and routine takes over. You would like to do something more exciting, but it is impossible due to having to perform routine daily activities. This problem certainly exists in Statistics Estonia as well, and it needs constant attention. We have work processes that are renewed every ten years. If we hire a person today and make them use FoxPro, we can be quite sure that the person will leave soon.

Team. Do the members of our team support each other? Do our managers support our employees? Managers' support for their team is crucial, starting with whether the manager has time to support the team. With the most recent structural change we granted team leaders rights, responsibility and authority to lead their team. But it is impossible to lead a team consisting of 140 people. The leader cannot focus on the problems of an employee and support him or her. It is difficult to think what could be the new tasks for each employee and what training they might need. Therefore, we have reorganised the teams in Statistics Estonia.

Reputation. It is natural that the members of an organisation want to feel that they work for a successful organisation. Even if this is not a wish you are aware of, most people still have it. What makes an organisation successful? Its external reputation is one thing, but it is also about the ability to notice and celebrate small victories. If we can celebrate our victories and do it with our partners, we also create positive reputation. We must definitely improve in this direction.

You have worked at Statistics Estonia for more than half a year already. It seems to me that Statistics Estonia has already undergone or is about to undergo a significant paradigm shift. Previously, we have focused on why it is impossible to do one or another thing, but now we are looking for an answer to how we can get necessary things done. Have I understood that correctly?

Yes, you are completely right. This is my first time working in the public sector. Initially, the most shocking difference between the public and the private sector was that everybody talked about having no money and it being impossible to do anything. Private sector lacks money as well, but nobody talks about it. Instead, they use their common sense to attempt to find some way how to get money, how to make money. How can we make things happen if there is little or no money? Is it possible to stop doing something so that the necessary activities can be done more? Yes, you understood it correctly.

It reminds me of what Emõke Sogenbits, head of a large enterprise in Estonian terms, said in this year's third issue of the Quarterly Bulletin of Statistics Estonia: "If you neglect the people, the course of events is unpredictable". In your opinion, have people in Statistics Estonia been neglected?

I think that in certain areas they have been, but I believe that we are quickly moving towards finding all the diamonds and gems in our organisation, so that no one would be neglected in Statistics Estonia.

The thought of robots taking over the world one day does not seem like science fiction anymore. When do you think will robots take over the production of official statistics? What would be the last job in Statistics Estonia that robots will take over and how long will Statistics Estonia carry out face-to-face interviews?

I will answer in reverse order. Statistics Estonia does not currently conduct any qualitative studies. The road we have chosen is that we will also perform qualitative studies in addition to quantitative studies, meaning that face-to-face interviews are not about to disappear. I'm not going to say that the share of such interviews is growing in the production of official statistics, but swapping quantitative studies with qualitative studies is definitely one direction. Face-to-face interviews provide information that cannot be obtained otherwise, but which we currently do not use yet. We have to learn how to use it.

As for the first part of the question, Statistics Estonia has launched activities which involve automation of several activities. For example, today we often copy data from one Excel file to another, which is rather pointless, yet time-consuming work. Manual performance of such tasks must be replaced by technical solutions.

What will be left for Statistics Estonia? Considering all the robotics-related developments, today there is still no known solution for how to replace people in developing algorithms.

Recently, Statistics Estonia has had one reform after another. Frequently, one reform cannot end before a new one is started to amend it. Perhaps there is a reasonable cause for each individual change, but it creates confusion when a new reform begins before the previous one has had a chance to show results. What is your position on reforms?

Evolution is essentially a very long process. Revolutions speed it up significantly. There are phases where an organisation requires revolutionary developments. On the one hand, we are trying to make as few changes at Statistics Estonia as possible, but on the other hand, considering strong social pressure on producing official statistics, there are parts of statistics production, e.g. in the statistics production technology, that require urgent changes. This change cannot be years-long. I'm sorry, but three or five years for making a data warehouse is not a normal timeframe.

Change does not mean that everything should be turned upside down. I hope that my successors and I can maintain everything that is valuable in Statistics Estonia, and that positive traditions will survive. It makes sense for the new leader to continue with necessary changes started by the previous one. My predecessors Priit Potisepp and Andres Oopkaup have done a great job. I have said it before and I am convinced in it. I am here to preserve all the valuable traditions of Statistics Estonia, and while doing so, to introduce changes to increase the total value of Statistics Estonia as an organisation as quickly as possible.

The current changes in Statistics Estonia may seem very fast for some people and very slow for others. The director general does not make decisions alone in Statistics Estonia. Decisions are made by the management consisting of the leaders of all the significant areas in the organisation. We argue and discuss some decisions many times. Some people ask why does it have to be such a lengthy process. Others are surprised that it happens so quickly, especially when the decision personally affects them. Finding balance is an endless issue and in the course of discussions, it is important to keep in mind the objective to achieve the expected results and to improve the situation for the members of the organisation.

Have I understood correctly that in your opinion there is no major difference in leading a public institution or a private sector enterprise?

Management principles are the same in both the private and the public sector. In that sense, there is no difference. The staff and team are major assets of an institution or enterprise. The most important is how we make things happen and get results.

There is one more question left, but before that I want to ask the traditional question which I have asked from each person whom I have had the honour to interview so far. Is there an important question that I did not ask? If yes, what is it and how would you answer it?

I do not think there was such a question. Considering how complex a topic can be in the context of statistics, the key to understanding often lies in simplifying things. If we can simplify, it is easier to communicate our message.

There was once a Roman statesman (Senator Cato the Elder), who ended all his speeches with a message that Carthage should be destroyed. In the end, Carthage was destroyed. What is the message of Statistics Estonia to the Estonian society? What should be heard consistently?

I will not give a brief response. I recommend reading "Homo Deus: A Brief History of Tomorrow" by Yuval Noah Harari. Unfortunately, this book has not been translated into Estonian yet, but it has become a bestseller. I will not retell the book, but the interesting idea of the book is that the ruling religion of humankind is not Christianity, Islam, etc., but instead Dataism, i.e. belief in data, has become the ruling religion. More than ever before, we believe in the decisions made by machines. For example, if we want to drive to Järva-Jaani, we enter the destination in Google Maps and follow the suggested route to get there, believing that this route is better than the one we would have chosen ourselves. We place our trust in machines and data. However, it is questionable whether trusting machines is the smartest thing to do at a time when the world is full of false news and data.

The message from Statistics Estonia could be that the algorithms we use for processing data and the information we offer are reliable.

The interview took place on 15 November 2017.

MÕTTEID ARVAMUSFESTIVALIDELT

Märt Leesment, Mihkel Servinski

Kahel viimasel korral on Statistikaamet võtnud osa Paide arvamusfestivalist. Kui 2016. aastal arutleti teemal „Mis kasu on meile statistikast?“, siis 2017. aastal oli arutelu teema „Paremad otsused = vähem privaatsust?“. Teemade fookus oli küll erinev, kuid siiski olid aruteludes kõlanud mõtted piisavalt lähestikku, et neid sai koos ning teineteist täiendavana vaadelda. Järgnevalt on kokku võetud mõlema aasta arutelud ja heidetud põgus pilk kõlama jäänud mõtetele, lõpus on toodud ka autorite järeldused.

Statistika vajalikkusest

Pole saladus, et ettevõtjad näevad Statistikaametit väga sageli tüütuna, sest nad ei saa aru, miks peab kogu aeg metsikus koguses andmeid edastama või mis kasu sellest saab. Samasugused küsimused võivad tekkida ka eraisikul. Lühidalt võib vastata, et parem elu eeldab paremaid otsuseid, nende otsuste tegemiseks on aga tarvis sisendit – paremaid ja rohkem andmeid. Nii ettevõtted/organisatsioonid kui ka inividid/leibkonnad on kriitilise tähtsusega andmeallikad. Andmete baasil tehakse kindlaks ühiskonnas toimuvad protsessid, aga koostatakse ka prognoosid, näiteks ettevõtjat huvitab üldine majanduskliima, makstav palk, tööjõu paiknemine ning oskused jne. Hüvedega käib kaasas ka vastutus ja see ei tähenda mitte üksnes andmete esitamist, vaid õigete andmete esitamist. Teisisõnu: kuna avalikkuse huvi on saada võimalikult korrektne peegeldus tegelikkusest, peaksid andmeesitajad edastama korrektseid andmeid. Küsimus on, kuidas saaks neid motiveerida seda tegema. Üks olulisi probleeme peitub statistika vajalikkust puudutavas teavitustöös (eelkõige selle vähesuses). Olukorda ilmestab väliskogemustega ettevõtjate kommentaar, et võrreldes mitme teise riigiga on Eestis aruandluskeskkond hea ja nõutavate andmete maht ei ole liiga metsik. Tegemist pole küll meetilgaga tõrvapotis, kuid vastupidine arvamus kipub sellegipoolest domineerima.

Harimisest

Mõlemal aastal toodi üksmeelselt esile, et inimesi peab harima. Välja koorus kolm tasandit:

- parandada statistilist kirjaoskust, sh tõlgendamist;
- teavitada andmekaitseõigustest;
- arendada kriitilist mõtlemist, tuua ausse tõendus põhisis.

Statistika mõistmiseks on tarvis statistilist kirjaoskust. Andmetest on kasu üksnes siis, kui nendest saadakse aru ja osatakse tõlgendada. Statistilist teavet saadakse tihtilugu teiste allikate kaudu. Probleem on aga selles, et andmeid vahendades tõlgendatakse neid erinevalt. See omakorda võib suurendada usaldamatust statistika vastu. Seega võiks Statistikaamet aidata tagada, et inimesed teaksid otsida ja leiaksid üles esmased andmed Statistikaameti andmebaasist. Selle teeb keeruliseks asjaolu, et peale äriorganisatsioonide pakuvad oma statistikat ka teised riigiasutused (näiteks palgastatistika nii Maksu- ja Tolliametilt kui ka Statistikaametilt võib tavatarbijas tekitada küsimuse, keda uskuma peaks).

Kunagi varem pole inimkonnal kasutada olnud nii palju kogutud andmeid kui praegu, kuid sageli jäädakse jänni nende tõlgendamisega. Küsimus on selles, kuidas peaks tõlgendama ja kas tõlgendamise ning arvamuse kujundamisega peaks igaüks ise tegelema või peaks meedia abiks olema. Või hoopis Statistikaamet. Mida teha siis, kui andmeid saab tõlgendada erinevalt?

Konkreetse teadmise kohta võib olla mitu arvamust ja üsna sageli on arvamus hakanud asendama teadmist. Kui aga arvamus ei tuginegi teadmisele, siis kuidas peaks info vastuvõtja aru saama, kas tegemist on arvamuse või teadmiselega?

Riikliku statistikana avaldatud andmed on neutraalsed, tõlgendus ei pruugi aga olla. Andmete väänamise ja kallutatud tõlgendamise puhul ei kahjustu mitte üksnes tõlgendaja, vaid ka andmete maine. Niisuguse loogika kohaselt võib andmete väär kasutamine kahandada ka usaldust Statistikaameti vastu. Statistilise kirjaoskuse õpetamisega peaks kaasama ka kriitilise mõtlemise arendamine ja tõenduspõhisuse väärtustamine. Niinimetatud tõejärgses olukorras (kus emotsioonid ja subjektiivne arvamus muutub olulisemaks objektiivsest faktist) on olulisel kohal oskus orienteeruda mitmesuguste tõdede rägastikus ja suutlikkus kahelda.

Kuigi igaüks vastutab oma rumaluse või tarkuse eest ise, on andmekaitse aspektidest teavitamine suuresti ikkagi riigi ülesanne. Esimene koht oleks kõikidel juhtudel (üldharidus)kool. Keerulisem lugu on juba koolitee minevikku jätnud täiskasvanutega, kelle küber(turvalisuse) teadmised kipuvad noorte omadele alla jääma. Peab arvestama, et vanemate ellu on tehnoloogia hiljem tulnud, aga noorem generatsioon on tunduvalt suurema osa oma elust arvutimaailmaga kokku puutunud.

Kuna elu liigub üha enam küberruumi, tuleks aina rohkem teadvustada ka sealseid riske, olulisemad neist on seotud privaatsusega. Riigi kui andmete koguja seisukohalt tähendab rohkem andmeid suuremaid võimalusi teha paremaid otsuseid ja selle kaudu luua paremat elukeskkonda. Teisalt peavad elaniku selle hüvanguks loovutama üha enam oma privaatsusest. Ühiskondlikud kokkulepped määravad, kust jookseb piir saadava kasu ja privaatsusest loobumise vahel. Sääraste kokkulepete saavutamine pole aga lihtne, eriti veel olukorras, kus andmete turvalisusega seotud aspektid pole päevapoliitika jaoks piisavalt seksikaks osutunud.

Kõige taustal tuleks aga meeles pidada, et riik ei ole ainus andmete koguja. Mitmesuguste rakenduste alla laadimisel, Google'i või muu pakkuja teenuste kasutamisel küsitakse nõusolekut inimest puudutavate andmete jagamise ja kasutamise kohta. Teisisõnu tuleb teenuse kasutamiseks loobuda teatud osast privaatsusest. Teenuseosutaja kasutab neid andmeid (väidetavalt) teenuse parandamiseks. Kuigi säärane olukord ei ole tarbijate jaoks just ülearu meeldiv, ei ole ka tulist vastuseisu ja seda väljendab ka üha suurenev tarbijate hulk. Kui aga inimesi jälgib riik, siis nähakse seda sageli hoopis halvemas valguses. Võib-olla on olukord tingitud sellest, et erinevalt erafirma teenuse kasutamisest on kodanikel keeruline riigi teenusepakkumisest keelduda ja oma andmeid mitte jagada. Sellest hoolimata ei süüvi tarbijad liiga sageli sellesse, millele nad erafirma teenusega liitumisel nõusoleku annavad. Vaikimisi ollakse kõigi nõudmistega päri. Et selline olukord on tekkinud, siis tuleb näpuga ka teenuseosutaja peale näidata, sest tema pole suutnud asjaolusid tarbijatele arusaadavaks teha. Nii või teisiti on vaja paremat selgitustööd nii teenuseosutajalt kui ka riigilt. Isegi sel juhul, kui inimestel polegi selle vastu midagi, et keegi nende kohta andmeid tahab.

Tulevikust

Ühiskond soovib kiiremini, rohkem ja kvaliteetsemad andmeid, aga väiksemat koormust andmeesitajale. Kuna infot tekib üha enam, siis üks võimalikke lahendusi on suurandmete laialdasem kasutamine. Sääraselt kogutud andmeid saadakse operatiivsemalt ja andmeesitajaid säästetakse väga palju. Probleem on aga andmete omanikega/haldajatega kokkuleppele jõudmine. (Pikemalt saab suurandmete kohta lugeda artiklist „Suurandmed statistikas“).

Suurandmete kättesaadavuse paranemisega võib laieneda andmete hulk, mis omakorda parandab võimalusi paremate otsuste tegemiseks. Samas, mida rohkem teavet andmesubjekti kohta kogutakse, seda suuremad on ka (privaatsus)riskid. Oluline on tähele panna, et kogumine oleks mõtestatud, andmed peaksid olema kogutud ja töödeldud eesmärgipäraselt ja tulemusest peaks olema ühiskonnale kasu.

Eksimisvõimalused ja riskid on aga suuresti tingitud inimfaktorist. Siiski on inimtegevuse roll andmete kogumisel ja töötlemisel vähenenud ning masinate tähtsus suureneb märgatavalt ka edaspidi. Kas võib juhtuda, et masinad hakkavad tegema otsuseid? Kui paljusid standardseid otsuseid juba teevad masinad (kiiruskaamerad, enamakstud tulumaksu tagasiarvestamine maksuametist), siis missugused otsused võiksid kindlasti inimese hooleks jääda? Kui masin on raudkindlalt ratsionaalne, siis inimese otsuseid mõjutavad tihti emotsioonid. Küsimus on, kas

masinad teevad inimesest paremaid otsuseid. Või mida teha siis, kui inimene ei mõista enam masinat, kui masina töötamise loogika muutub hoomamatuks. Ei ole kindel, kas laseksime sel juhul masinatel enda asemel otsustada. Kuna otsustusõigusega kaasneb ka kohustus vastutada, võiks olla selge, kuidas vastutavad masinad. Oluline on mõelda ka sellele, kas kõige paremad lahendused peavad alati ja üksnes ratsionaalsetele otsustele tuginema või tuleb mõnel juhul kasuks ka inimlikkus (sh eksimisevõimalus).

Kokkuvõte

Kui jälgida Eesti ühiskonna mõtteavaldusi riikliku statistika teemal, siis võib jääda mulje, et tegemist on nähtusega, mille pealt on võimalik palju raha kokku hoida. On isegi arvatud, et Statistikaameti likvideerimine leevendaks olukorda Eesti tööturul, kus paljud töötajate nappuse üle kurdavad.

2016. aasta arvamusfestivalile minnes soovis Statistikaamet selliste arvamuste üle arutleda. Õnneks ei saanud diskussioonist asja. Vestlusringis osalenud ettevõtjad, meedia ja poliitikute esindajad (Statistikaameti esindajast rääkimata) teatasid avaringis otsustavalt, et küsimus „Kas statistikat on vaja?“ on nii rumal, et selle esitamisest rumalam oleks ainult niisuguse diskussiooni alustamine. Leiti aga, et mõtet on arutleda teemadel, millist statistikat toota, kuidas vajaliku statistika tootmist mõistuspäraselt korraldada ja kuidas toodetud statistikat paremate juhtimisotsuste tegemiseks kasutada.

Kokkuvõtteks võib öelda, et diskussioonis osalenud ja mõnekümnepealine publik jõudsid arusaamisele, et kindlasti on vaja tänapäevastada riikliku statistika kogumise meetodeid, sh on eriti tähtis suurandmete kasutuselevõtt ja Statistikaameti e-võimekuse suurendamine. Oluline on, et Statistikaamet teeks tunduvalt rohkem koostööd andmeesitajate ja kasutajatega, et kogutaks vaid neid andmeid, mille alusel vajaliku infot toodetakse. Tähtis on, et see info oleks esitatud tarbijatele kasutamiseks mugaval kujul, arvestades, et tarbijatel on erinev ettekujutus info mugavast kasutamisest. Leiti ka, et statistika vajalikkuse küsimus on ühiskonnas siiski olemas ja et ainus meetod selle küsimuse kaotamiseks on parandada statistilist kirjaoskust.

Statistikaameti vaatevinklist oli ehk kõige ootamatum vestlusringis osalenud ettevõtjate hinnang Eesti aruandluskeskkonnale: see on võrreldes paljude Euroopa Liidu riikidega hea ja esitada tulevate andmete maht ei ole tegelikult metsik. Et arutlusringis esinenud ettevõtjatel on olemas märkimisväärne töökogemus väljaspool Eestit, siis võib sellist hinnangut usaldusväärseks pidada. Lisades muidugi, et positiivne hinnang ei tähenda seda, et aruandluskeskkonda ei tuleks arendada.

Statistika taga on inimene. Andmeid kogutakse inimtegevuse kohta, neid töötlevad ja analüüsivad inimesed. Neid tõlgendavad ja nende põhjal teevad otsuseid samuti inimesed. Isegi kui mõni protsess usaldatakse masinate hooleks, siis tehtud otsused puudutavad ikkagi inimest. Seetõttu on vaja leida tasakaal, et ühelt poolt oleks andmekogumine võimalikult inimlik (ehk riivaks inimese huve, õigusi, privaatsust ning kaasaks kohustusi võimalikult vähe) ja teisalt võimaldaksid kogutud andmed langetada elu võimalikult heaks tegevaid otsuseid. On tähtis, et statistika oleks arusaadav ja kõnetaks inimesi, oleks inimlik. Ehk läheb siis statistika ka rohkematele korda.

Statistikaamet tänab aruteludel osalenuid (2016. aastal: Mihkel Servinski (Statistikaameti peaanalüütik), Siim Krusell (Statistikaameti peaanalüütik), Riin Aljas (Postimehe teadustoimetuse juhataja), Heilir-Valdor Seeder (Riigikogu aseesimees), Peter Wüthrich (ettevõtja ja OÜ Eesti Valgus omanik), Sami Seppänen (Elisa juht); 2017. aastal: Heidy Roosimägi (Statistikaameti juhtivstatistik-metoodik), Baldur Kubo (Cybernetica AS, Sharemindi projektijuht), Gregor Sibold (tehnoloogiaudiste portaali geenius.ee toimetaja), Liisa-Ly Pakosta (soolise võrdõiguslikkuse ja võrdse kohtlemise volinik), Viljar Peep (Andmekaitseinspektsiooni peadirektor), Jaak Madison (Riigikogu liige)), kuulajaid, arutelude toimumisele kaasa aidanud vabatahtlikke ja arvamusfestivali korraldajaid. Statistikaamet ei ühine pärast 2017. aasta arvamusfestivali avaldatud arvamusel, et üritus on vaid kasutu jututuba. Jah, arvamusfestival on tõesti jututuba, aga mitte kasutu.

THOUGHTS FROM THE OPINION FESTIVALS

Märt Leesment, Mihkel Servinski

Statistics Estonia took part in the last two Opinion Festivals in Paide. In 2016, the discussion topic was “How do we benefit from statistics?”, and in 2017, it was “Better decisions = less privacy?” The central theme of these topics was different, but the ideas heard during the discussions were close enough to be considered together. In the following, the discussions of both years are outlined and the prominent ideas are touched upon. At the end, the authors’ conclusions are also presented.

About the need for statistics

It is no secret that enterprises often view Statistics Estonia as an annoyance, as it is not understood why enormous amounts of data continuously need to be submitted and how this can be beneficial. A private person might also have such questions. The short answer is that better life depends on better decisions and to make these decisions input is needed – more data with better quality. Enterprises/organisations as well as individuals/households are critical data sources. On the basis of data, social processes are identified, but also projections are made. For example, an entrepreneur is interested in the general economic climate, salaries paid, location and skills of the work force, etc. With benefits comes responsibility and not just in the sense of submitting data, but in the sense of submitting correct data. In other words, if the society is interested in getting as accurate a reflection of reality as possible, data providers should submit correct data. The question is how to motivate them to do that. To a great extent the problem lies in informing (mostly lack of informing) about the need for statistics. The situation is characterized by comments of entrepreneurs with experience outside Estonia who have said that compared to several other countries the reporting environment in Estonia is good and the amount of required data is not colossal. Although the situation is not entirely bad, the opposite opinion still tends to dominate.

About educating

In both years, there was consensus that people need to be educated. Three levels of educating emerged:

- *improving statistical literacy, incl. interpreting data;*
- *informing about data protection rights;*
- *developing critical thinking, raising the value of evidence-based information.*

To understand statistics, statistical literacy is required. Data are useful only when people understand them and know how to interpret them. Statistical information is often received from secondary sources. The problem lies in the fact that when data are mediated, they are twisted and interpreted differently. This might in turn increase mistrust of statistics. Therefore, Statistics Estonia could help to ensure that people would know to look for and find the primary data in the database of Statistics Estonia. This is made complicated by the fact that besides business organisations other state authorities publish statistics as well (for example, wages and salaries statistics are offered by both the Tax and Customs Board and Statistics Estonia, and this might make an average user wonder who to believe).

People have never had available as much collected data as we do now, but there is often trouble with interpreting these data. The question is about how to interpret and whether everyone should interpret data and form their own opinions or should the media help. Or Statistics Estonia? What to do when data can be interpreted differently?

There can be many opinions about particular knowledge, and quite often opinions have started to replace knowledge. But if opinions are not even based on knowledge, how should the information recipient know whether the issue at hand is opinion or knowledge?

Data published as official statistics are neutral, but the interpretation might not be. Twisted data and biased interpretation might hurt not only the reputation of the interpreter but also that of data. According to this logic, wrong use of data might diminish trust in Statistics Estonia as well. Teaching statistical literacy should go together with developing critical thinking and valuing evidence-based information. In the so-called post-truth environment (where emotions and subjective opinion become more important than objective facts), the skill to find one's way in a maze of truths and the ability to doubt have an important place.

Although everyone is responsible for their own foolishness or cleverness, informing of data protection aspects is still greatly the task of the state. In any case, the first place for this would be (general education) schools. The situation is more complicated when it comes adults who left school a long time ago and whose knowledge in cyber (security) matters tends to lag behind that of young people. It should be taken into account that technology entered the lives of older people later and the younger generation has been involved with the computer world for a significantly longer part of their lives.

As our daily life increasingly takes place in the cyber space, the risks that are found there should also be more acknowledged. The more important risks are privacy related. From the viewpoint of the state as the data collector, more data means greater opportunities to make better decisions, thereby creating a better living environment. On the other hand, citizens have to give up more of their privacy for this. Social agreement determines where the border between benefits and giving up privacy is. Reaching such agreement is not easy, though, especially considering that the aspects related to data security have not been intriguing enough for the political agenda.

At the same time, it should be kept in mind that the state is not the only data collector. When downloading various applications, using the services of Google or other providers, people are asked to agree to personal data sharing and use. In other words, in order to use the service, some privacy has to be relinquished. The service provider uses these data (allegedly) to improve the service. Although this situation is not especially liked by consumers, there is no strong opposition, and this is evident in the ever higher numbers of consumers. But when it is the state keeping an eye on people, this is often viewed more negatively. Perhaps the reason for this situation is that in contrast to using the service of a private company, citizens have a hard time refusing a service offer by the state and not sharing their data. Nevertheless, consumers often do not pay enough attention to what they agree to when starting to use a private company's service. All requirements are agreed to by default. Some of the blame for this situation lies with the service provider, because it has not been able to make the conditions clear to consumers. In any case, better information work is required from both the service provider as well as the state, even if people actually have nothing against anybody wanting data about them.

About the future

The society expects to get data faster, with more quantity and quality, but a smaller burden for data providers. As there is increasingly more information generated, one of the possible solutions is more wide-spread use of big data. Big data is collected efficiently and data providers are spared much more. The problem here has to do with coming to agreement with data owners/managers. (Read more about big data in the article "Big data in statistics", p. 48).

With better access to big data, the amount of data might increase, which improves the possibilities of making better decisions. However, the more information about a data subject is collected, the greater the (privacy) risks as well. It is important to keep in mind that data collection would be justified, that data are collected and processed with a purpose and the society would benefit from the results.

The chances and risks of being wrong are mainly caused by the human factor. However, the role of people in collecting and processing data has decreased and the importance of machines continues to increase considerably in the future. Is it possible that machines will start making decisions? If many standard decisions are already being made by machines (speed cameras, calculating tax refunds at the Tax and Customs Board), then which ones should definitely be made by people? While a machine is completely rational, a person's decision is often influenced by emotions. Do machines make better decisions than people? What to do when the machine's work logic becomes unfathomable for people? The question is whether we would in this case let machines make decisions for us. As with the right to make decisions also comes responsibility, a question arises about how machines can be responsible. It is important to consider whether the best solutions always and only depend on rational decisions or is humanness (incl. the chance of errors) sometimes useful.

Conclusion

When regarding the opinions expressed in the Estonian society on the topic of official statistics, it might seem that it is an area where a lot of money could be saved. It has even been expressed that closing Statistics Estonia would alleviate the situation on the Estonian labour market where many complain about the lack of workers.

Going to the 2016 Opinion Festival, Statistics Estonia was prepared to discuss such opinions. Fortunately, this discussion did not come about. The entrepreneurs, media and politicians' representatives (and of course the representative of Statistics Estonia) participating in the panel discussion declared in the opening remarks decisively that the question of whether statistics was needed was silly and starting such a discussion was even sillier. It was found, however, that it made sense to discuss which kind of statistics to produce, how to organise the production of necessary statistics in a rational way and how to use the produced statistics for making better governance decisions.

To sum it up, it can be said that the discussion participants and an audience of a couple of dozen reached an understanding that the methods for collecting data for official statistics definitely need to be updated and the adoption of big data and increasing the e-capabilities of Statistics Estonia are especially relevant. It is crucial that Statistics Estonia would cooperate significantly more with data providers and users, so that only these data are collected which are used to produce necessary information. It is important that this information would be presented in a format convenient for consumers to use, taking into account that consumers have a different understanding of convenient use. It was also concluded that the issue of whether statistics are needed indeed exists in the society and the only way to solve this issue is to improve statistical literacy.

The discussion participants' evaluation of the Estonian reporting environment surprised Statistics Estonia the most: compared to many other European countries, the reporting environment is good and the amount of data to be submitted is actually not enormous. As the entrepreneurs who took part in the discussion have considerable working experience outside Estonia, this evaluation can be considered trustworthy. It should be added, of course, that a positive evaluation does not mean that the reporting environment should not be improved.

There are people behind statistics. Data are collected about the activity of people and they are processed and analysed by people. They are also interpreted by people and the decisions based on them are made by people. Even if a process is trusted with machines, the decisions made still concern people. Therefore, it is necessary to find balance, so that on the one hand, data collection would spare people as much as possible (i.e. it would interfere with people's interests, rights, privacy and bring along duties as little as possible), and on the other hand, the collected data would enable making decisions which improve life as much as possible. It is essential that statistics would be comprehensible, would address people and be humane. Perhaps then statistics would matter to more.

Statistics Estonia expresses gratitude to discussion participants (in 2016, Mihkel Servinski (principal analyst of Statistics Estonia), Siim Krusell (principal analyst of Statistics Estonia), Riin Aljas (head of science news of Postimees), Helir-Valdor Seeder (Vice-President of the Riigikogu), Peter Wüthrich (entrepreneur and owner of OÜ Eesti Valgus), Sami Seppänen (head of Elisa), and in 2017, Heidy Roosimägi (leading statistician-methodologist of Statistics Estonia), Baldur Kubo (project manager of Cybernetica AS, Sharemind), Gregor Sibold (editor of the technology news portal geenius.ee), Liisa-Ly Pakosta (Gender Equality and Equal Treatment Commissioner), Viljar Peep (Director General of the Estonian Data Protection Inspectorate), Jaak Madison (Member of the Riigikogu)), as well as the audience, volunteers who helped with discussion organisation and the organisers of the Opinion Festival. Statistics Estonia does not agree with the opinion expressed after the 2017 Opinion Festival that the event is just a useless chat room. Yes, the Opinion Festival is a chat room, but it is not useless.

SUURANDMED STATISTIKAS

Toomas Kirt

Artikkel annab ülevaate, kuidas on defineeritud suurandmed ja milliste vahenditega neid töödeldakse. Seejärel tutvustatakse suurandmete potentsiaalset kasutamist riiklikus statistikas: kuidas Statistikaamet osaleb suurandmete katseprojekti, mille raames uuritakse elektri kaugmõõtjatega salvestatud andmete potentsiaalset kasutamist statistika tootmisel.

Sissejuhatus

Järgnev püüab anda selguse, mis on peidus termini „suurandmed“ taga ja kuidas need on seotud statistikaga. Suurandmed ei tähenda, et numbrid oleksid kirjapildilt suured või arvud oleksid väärtuse poolest suured, vaid pigem nähtust, et neid arve on hästi palju. Suurandmeteks saamiseks ei pea andmed olema esitatud numbrina, vaid need võivad olla ka mingite muude sümbolitega, näiteks kirjatekstid. Kuid suurandmeteks saamiseks peab neid olema väga palju, st tavaliste meetoditega ei saaks neid enam töödelda.

Suurandmed tekivad valdavalt siis, kui mingit pidevat või regulaarset protsessi mõõdetakse ühe või mitme parameetri järgi või salvestatakse muud protsessi tulemit. Mida rohkem on parameetreid ja mida sagedamini mõõdetakse, seda enam andmeid saadakse. Rahvaloendus toimub korra kümne aasta tagant ja siis loendatakse ära kõik Eesti elanikud (1,3 miljonit), selle käigus saadakse vastused sajakonnale küsimusele. Nii tekib suur hulk andmeid, kuid need pigem ei ole suurandmed. Sama hulk inimesi, kes kasutab iga päev mobiiltelefoni, tekitab juba päevas palju rohkem andmeid kui võimalik saada rahvaloendusega. Nende andmete salvestamiseks ja töötlemiseks on juba vaja erinevaid meetodeid ja need loetakse suurandmeteks.

Suurandmed

Enamikule andmetega kokku puutuvatele inimestele on jäänud kõrvu termin „suurandmed“. See on defineeritud EKI terminibaasis (2014) kui suured, mahukad, struktureeritud ja struktureerimata andmed, mille haldamine tavapäraste relatsiooniliste andmebaasi- ja andmetöötlemise vahenditega on raskendatud, kui mitte võimatu (maht, formaat, kiirus). Põhiliselt iseloomustab suurandmeid vajadus kasutada traditsioonilisest erinevaid töötlusvahendeid. Kuna andmetöötlusvõimekus suureneb ajas pidevalt, siis ei pruugi praegused suurandmed seda enam olla tulevikus. Teisalt iseloomustab suurandmeid ka see, kuidas on võimalik nendest ülevaade saada. Paari tuhandet kirjet andmebaasis on inimene võimeline rida haaval üle vaatama, kuid teatud mahust alates muutub selle tegevuse jaoks kuluv aeg ebaproportsionaalselt pikaks ja andmetest ülevaade saamiseks tuleb kasutada visualiseerimist või muid andmete struktuuri avavaid lahendusi. Kuid andmed ei pea olema ainult mahult suured, vaid võivad seda olla ka tähtsusest, kui nad kajastavad unikaalse nähtuse vaatlust (Lynch 2008), mida ei saa taastada ning seetõttu on oluline tagada sellega seotud andmete säilimine.

Joonis 1. Suurandmed^a*Figure 1. Big data^a*

12

2658756354657538

345453,4545354,5645645640,9
24154564545,4564,4564564,45
364536,4564564564,78645454,
45456345,4534534,4534534,45
34536453,454545454564619,3

^a Suurandmed ei tähenda suures formaadis numbrite kirjutamist (ülemine) või suuri arve (keskmine), vaid suurt hulka numbrilisel või mittenumbrilisel esitatud andmeid (alumine).

^a *Big data do not mean writing large-format numbers (top) or large numbers (middle), but a large number of data presented as numbers or other symbols (bottom).*

Suurandmete mõiste pärineb 1990. aastate keskpaigast (Diebold 2012). Suurte andmehulkade defineerimise raamistikuks võib lugeda Doug Laney pakutud kolme V mudelit (2001), mis hindab plahvatuslikult suurenenud andmete haldamist mahu (*volume*), kiiruse (*velocity*) ja struktuursuse (*variety*) kaudu. Kolme V mudel määratleb, et e-teenuste pakkumisel suureneb plahvatuslikult andmete hulk, kasvab märgatavalt andmete liikumise ja töötlemise kiirus ning nende formaat muutub mitmekesisemaks, sisaldades nii struktureeritud (andmed, mida saab paigutada tavalisse relatsioonilisse andmebaasi) kui ka struktureerimata andmeid (ilma kindla struktuurita andmed, näiteks tekstid ja videosalvestised). Suurandmete määratlemisel on välja toodud, et kui tavalised andmed esindavad pigem valimit mingist nähtusest, siis suurandmed pigem tervet populatsiooni ning need on alati seotud andmete tekkimise ja töötlemise kiirusega (Kitchin ja McArdle 2016). Kolme V mudel on saanud täiendusi ka teiste v-tähega sõnadega (näiteks andmete väärtus (*value*), mis kajastab nii hoidmise kulu kui ka saadavat ärilist tulu). Välja on pakutud ka teistel tähtedel põhinevaid tunnuseid, kuid kolme V määratlus on valdavalt piisav, et kajastada suurandmete töötlemisel eesiseivaid ülesandeid.

Seoses salvestatavate andmemahtude suurenemisega on statistikas toimunud paradigma muutus (Japac jt 2015) ja vaadatakse kaugemale ka töö aluseks olevast juhuvalikul ning küsitlusel põhinevast metoodikast. Statistikud hakkavad üha enam kasutama administratiivseid andmeallikaid / registreid. Uus paradigma tähendab ka, et on võimalik digitaalselt koguda, seostada, agregeerida ja korreleerida andmeid. Saadakse palju mitmekesisem andmete esitus kui pelgalt küsitlusega. Suurandmete kasutuselevõtt võimaldab aga siduda veelgi detailirohkemaid andmeallikaid ja tuua välja uusi seoseid.

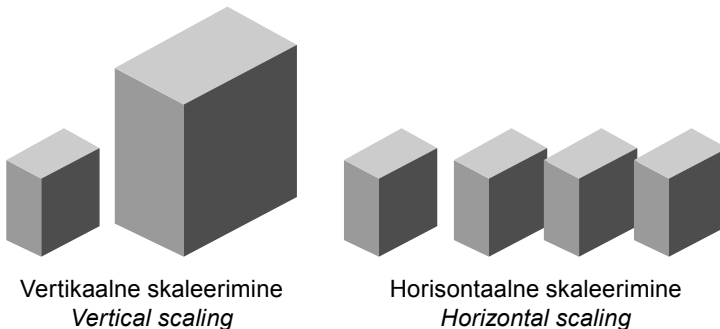
Suurandmete analüüsi teeb keeruliseks andmeelementide vaheliste seoste arvu eksponentsiaalne kasv. Näiteks kui soovitakse leida kõikide objektide vahelisi seoseid, siis tuleb leida seoste ruutmatriks, mille elementide arv suureneb objektide arvu ruuduna. Seetõttu on suurandmete analüüsiks vaja teistsuguseid lähenemisi, eriti kui soovitakse teha kiireid järeldusi andmete voo põhjal. Suurandmete kasutamisega kaasnevad ka riskid (Japac jt 2015) ja andmete töötlemiseks on vaja statistikutel uusi teadmisi ning oskuseid. Suurandmete kasutuselevõtul tuleb näiteks arvestada, et andmed võivad olla kallutatud (*biased*). Kui toetuda näiteks Twitteri andmetele, siis saab teha kallutatud järeldusi – kasutajad on pigem noored ja aktiivsed. Suurandmed võivad olla vigased, selektiivsed ja mittetäielikud ning need on sageli väljund protsessidest, mille eesmärk ei ole toota andmeid analüüsi jaoks.

Töötlusvahendid

Suurandmete defineerimisel viidatakse, et nende töötlusvahendid erinevad tavalistest. Järgnev püüab selgitada üht suurandmete töötlemise süsteemi – Hadoop-i. See on kasutusel ka Statistikaametis. Põgusalt kirjeldatakse ka selle peamisi tööpõhimõtteid. Suurandmete töötlemiseks kasutatakse kergesti laiendatavaid süsteeme. Laiendada saab vertikaalselt, mis tähendab palju suurema jõudlusega serveri hankimist. Teine võimalus on seda teha horisontaalselt, mis on töötluste jagamine mitme serveri vahel, jõudluse suurendamiseks lisatakse süsteemi veel üks server (Singh, Reddy 2014) (joonis 2). Kui kasutusel on tavaline andmebaasisüsteem ja andmeid tuleb juurde ning töötleva masina võimsus muutub ebapiisavaks, siis lisatakse andmebaasi-serverile andmete töötlemiseks vajalikku mälu, protsessorivõimsust või kõvakettamahtu ning töö saab jätkuda, st kasutatakse vertikaalset laiendamist. Kuid mingist hetkest muutub andmete mahu suurenedes ühel serveril põhineva süsteemi laiendamine äärmiselt kulukaks, sest suure võimsusega serverid on tavaliselt üsna kallid. Et tulla toime suureneva andmetöötlusmahuga, on välja töötatud horisontaalselt laiendatav suurandmete töötlemise süsteem – Apache Hadoop. Hajus arhitektuur pakub eelist, et süsteemi laiendamiseks ei pea hankima juurde kõige kallimat riistvara. Hajusus tagab ka suurema töökindluse, sest hajusa arhitektuuri tõttu ei mõjuta ühe sõlme töötamast lakkamine terviksüsteemi tööd. Kahtlemata on ka teisi lähenemisi (Singh, Reddy 2014), kuid Hadoopi platvorm on vabavara ja see on laialdaselt kasutusel suurtes tehnoloogiaettevõtetes.

Joonis 2. Serverarvutite jõudluse suurendamine vertikaalselt ja horisontaalselt

Figure 2. Increasing the performance of server computers vertically and horizontally



Hadoopi tekkimise aluseks sai Google'i failisüsteemi tööpõhimõtteid kirjeldava artikli avaldamine 2003. aastal (Ghemawat jt 2003) ja sellest järgmisel aastal ilmunud MapReduce'i algoritmi tööpõhimõtet tutvustav artikkel (Dean jt 2004). Nendest inspireerituna hakati hajusal arhitektuuril põhinevat süsteemi arendama Apache'i projekti „Nutch“ raames ja ning 2006. aastal tegi selle projekti vedaja Doug Cutting eraldi alamprojekti Hadoop, mis sai nime tema poja kasukaru järgi. Hadoopis on kaks peamist komponenti: juhtsõlm (vt detailsemalt White 2015) (*NameNode*) ja andmesõlm(ed) (*DataNode*). Juhtsõlmes hoitakse andmete metaandmeid ja juhitakse ning koordineeritakse andmesõlmede tööd. Andmesõlmedes hoitakse andmeid (tööpõhimõtte kirjeldus allpool) ja tegeletakse saadud päringute töötlemisega MapReduce'i tehnoloogiaga (kirjeldus allpool) kasutades.

Hadoop ei ole ainult failisüsteem, vaid sellega on seotud terve hulk töövahendeid, mida kokku kutsutakse Hadoopi ökosüsteemiks ja see on liidestatud ka mitme andmeanalüüsiplatvormiga. Hadoopi ökosüsteemist võiks näitena välja tuua Hive'i, mis võimaldab eri formaatides salvestatud andmetest formeerida struktureeritud andmetabeleid ja teha tavalisi SQL-päringuid. Näiteks võib tuua ka Sparki, andmetöötlussüsteemi, mis päringute kiirendamiseks hoiab andmeid võimalikult palju serveri mälus; Yarni, mis tegeleb päringute ja ressursside koordineerimisega ning ajastamisega; Sqoopi, mis võimaldab liidestada Hadoopi teiste andmebaasisüsteemidega jpt

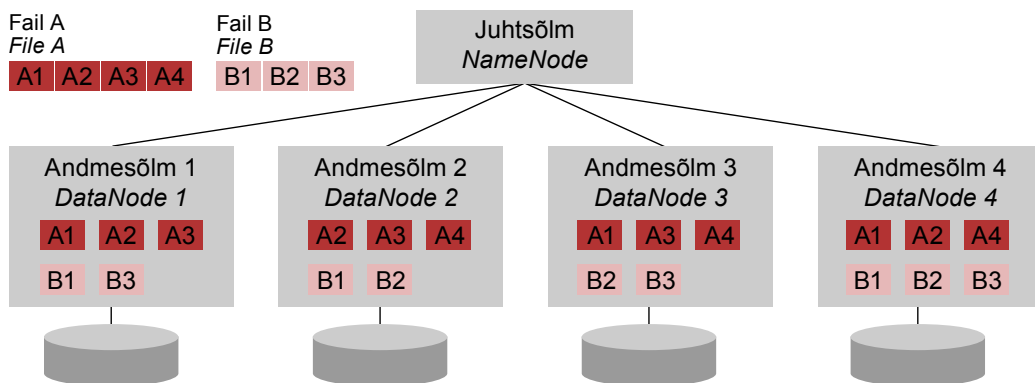
vahendid. Peale selle on administreerimiseks ja liidestamiseks arendatud mitu süsteemi, näiteks Ambari, Hue ja Zeppelin, samuti on võimalik kasutada andmeanalüütikute hulgas sageli tarvitusel olevat Pythoni Jupyteri Notebooki ja RStuudio. Hadoop pakub ühest küljest võimekust tulla toime suurte andmemahitudega, aga teisalt on see paindlik ning pakub võimalusi analüüsi tulemuste sobivaks esitamiseks.

Hadoopi tööpõhimõte

Hadoopi süsteemi alus on hajus failisüsteem – HDFS (*Hadoop Distributed File System*). Sinna võib salvestada igas vormingus ja suuruses faile. Suured failid tükeldatakse (vastavalt seadetele näiteks 64 MB tükkideks) ja igast tükist tehakse kolm koopiat, mis salvestatakse eri andmesõlmedesse (joonis 3). HDFS-is paistavad failid tavaliste failidena, kuigi füüsiliselt võivad selle tükid paikneda maailma eri otstes. Failist päringut tehes jagab juhtsõlm päringu tükkideks ja saadab teatud failiosa kohta päringu andmesõlmele nii, et süsteemi ressursid oleksid optimaalselt kasutatud ja päringule saaks kiiresti vastuse. Juhtsõlm jälgib, et kui teatud andmesõlmest ei tule vastust, siis esitatakse sama päring mõnele teisele, kus hoitakse koopiat samast tükist. See tagab süsteemi suure veakindluse, sest ühe andmesõlme kadu ei sega süsteemi edasi toimimast.

Joonis 3. Failid Hadoopi failisüsteemis

Figure 3. Files in the Hadoop file system



MapReduce

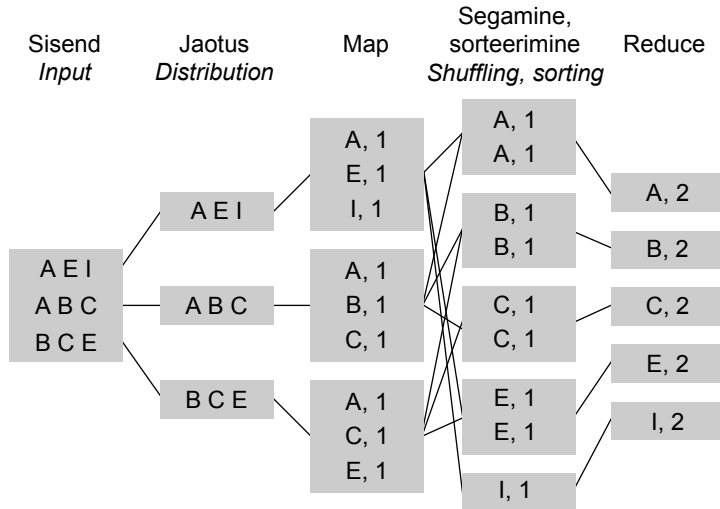
MapReduce on tehnoloogia, mille abil saab päringuid teha hajusalt^a. Süsteem koosneb kahest faasist: Map, mis kogub algandmed kokku võtme ja väärtuse paariks, ning Reduce, mis agregeerib võtme ja väärtuse paarid soovitud vastuseks. Joonisel 4 on toodud näide, kus on vaja teada saada failis asuvate erinevate tähtede arv. Kui tegemist on hajusa failisüsteemiga, siis on algfail jagatud kolmeks ning seetõttu käivitatakse kolm paralleelset Mapi tööd, mille tulemusena tekib praegusel juhul loetelu tähe ja väärtuse paaridest. Mapi protsess lõppeb segamis- ja sorteerimisfaasiga, kus ühesugused võtme ja väärtuse paarid kogutakse kokku. Seejärel algab Reduce'i protsess, et summerida võtme järgi väärtused ning iga tähe kohta tekib üks võtme ja väärtuse paar ehk siis vastus, mitu korda see täht tekstis esines. Protsessi vahetulemused salvestatakse tööfailidesse ja samuti toimub aktiivne andmevahetus sõlmede vahel, mistõttu nõuab protsess palju sisend-/väljundressurssi (I/O). Hajusas failisüsteemis pakub see veakindlat

^a Vt näiteks Tutorialspoint 2017 (https://www.tutorialspoint.com/map_reduce/pdf/map_reduce_introduction.pdf).

ja tõhusat päringute tegemise viisi. Lõppkasutaja jaoks on MapReduce'i kasutamine tehtud automaatseks ja näiteks Apache Hive'is saab kirjutada tavalise SQL-süntaksis (struktureeritud päringukeel andmebaasides) päringu ning see teisendatakse süsteemi tasemel automaatselt MapReduce'i päringuteks.

Joonis 4. MapReduce'i tööpõhimõte

Figure 4. Operating principle of MapReduce



Suurandmed riiklikus statistikas

Kasvanud võimekus töödelda efektiivselt suuri andmehulki on köitnud ka statistikuid, kes loodavad leida alternatiivseid viise statistika tootmiseks. Senine statistikasüsteem on valdavalt valikuuringupõhine (Särndal, Wretman, Swensson 1992), mida täiendatakse administratiiv-andmetega riiklikest registritest. Kuid suurandmete kasutuselevõtt toob statistikutele uued võimalused teha statistikat ka üldkogumi põhjal. Samas on märgatud, et suurandmete kasutamisel on ka piirid ja kõiki uuringute või loenduste käigus kogutavaid andmeid ei saa asendada suurandmetega, sest need ei pruugi katta kogu populatsiooni ega mõõda täpselt seda, mis kajastub loendusstatistikas (Tiit 2016). Suurandmete kasutuselevõtt eeldab ka teiste tehnoloogiate ja meetodite kasutamist, mis praegu ei pruugi statistikaametites kasutusel olla. Ka võtab aega suurandmete integreerimine standarditud statistika tööprotsessi (UNECE 2013). Suurandmete statistikas kasutuselevõttuga seotud probleeme lahendatakse aktiivselt ja sama kehtib uute andmeallikatega tutvumise puhul. Loetletud on paarkümmend projekti, mille raames püütakse kasutada suurandmeid statistikas (UNECE ... 2017).

Statistika püüab kajastada mingit üldkogumit kirjeldavaid arvilisi väärtusi, näiteks võib üldkogum olla kogu rahvastik või siis teatud uuringu puhul puidutööstusettevõtted. Kõige täpsem statistika saadakse, kui õnnestub koguda andmeid kõigi üldkogumi elementide kohta, kuid see on tavaliselt väga kulukas ja nõuab palju ressursi. Seepärast kasutatakse statistikas tavaliselt juhuvalimi põhiseid uuringuid, mille tulemusi üldistatakse üldkogumile. Meetod on üldjuhul väga efektiivne ja võimaldab saada piisavalt head statistikat. Kahtlemata on meetodil ka omad nõrkused, st kuigi valimi moodustamine on juhuslik, võib väikeste valimite korral olla üldistatud tulemus erinev tegelikkusest. Samuti mõjutab tulemust see, kui täpselt on andmeid esitatud. Seetõttu on oluline töötada välja meetodika, mis kõiki neid nähtusi arvesse võtab. Selle kaudu saab parandada tulemuse kvaliteeti.

Üldkogumist parema ülevaate saamiseks on statistikas hakatud laialdaselt kasutama registreid ja teisi administratiivseid andmeallikaid, sest registrid peaksid valdavalt kajastama tervikut. Suurandmete kasutuselevõtuga on võimalik liikuda veelgi täpsema tulemuse juurde, sest need kajastavad mingit nähtust väga detailselt ja katavad üldkogumi eri aspekte.

Statistikas kasutatakse sagedasti ka mudeleid, mis võimaldavad teatud andmelünki täita. Näiteks kui on teada, et üks perekond püüab oma naabritest alati kaks korda parem olla, siis saab naabrite tulemust teades tuletada lihtsa valemiga teise pere väärtused. Kui naabritel on üks auto, siis teisel perel on kaks autot, kui naabritel on üks maja, siis teisel kaks, kui naabrite peres on üks mees, siis ... Statistikas kasutatakse valdavalt lihtsaid lineaarseid mudeleid, mille muutujate vahelist seost saab selgelt kontrollida. Suurandmete korral kasutatakse pigem masinõppe põhised lähenemist, kus mudeli parameetreid optimeeritakse õppimisalgoritmidega ja leitakse tunnustevahelisi seoseid. Samas mudeli enda sisu kohta selget ülevaadet pole ja see on n-õ must kast. See on kriitiline asjaolu statistikas, kus üritatakse olla maksimaalselt läbipaistev. Seetõttu on nimetatud teemal ka märkimisväärne diskussioon (Braaksma, Zeelenberg 2015; Breiman 2001), kas ja millal võib mudeleid statistikas kasutada.

Suurandmete kasutuselevõtt võimaldaks märgatavalt muuta statistika tegemise protsessi. Florescu (2014) ja tema kolleegid on välja toonud viis võimalust suurandmete kaasamiseks statistikasse:

- asendada senine andmeallikas (valikuuring) suurandmetega ja säilitada senise statistika väljund;
- asendada andmeallikas (valikuuring) osaliselt suurandmetega ja säilitada senise statistika väljund;
- pakkuda statistilist lisainfot teisest vaatenurgast – lisaväljundid statistikale;
- täiendada ja täpsustada senise andmeallika põhjal tehtud hinnanguid – täiendatud statistikaväljund;
- pakkuda täiesti uut statistikat, mis põhineb uutel andmeallikatel – uus ja alternatiivne statistikaväljund.

Suurandete kasutamine statistikas pakub võimalusi ja proovikive, kuid selles on ka riske. Need on kokkuvõtlikult toodud järgmises tabelis.

Tabel 1. Võimalused, proovikivid ja riskid suurandmete kasutamisel statistika tootmiseks

Võimalused	Proovikivid	Riskid
Täiendada, asendada, parandada olemasolevaid andmeallikaid	Koostöö suurandmete omanikega, et leida mõlemale poolele kasulikke lahendusi	Statistikaametite missiooni muutus – keskendutakse liigselt saadaolevatele andmetele
Toota ajakohasemat väljundit	Saada juurdepääs andmetele	Aegriidade järjepidevus, kui andmeallikad muutuvad
Pakkuda alternatiivi küsimustikel põhinevale statistikale ja lihtsustada andmeesitajate tööd	Saada juurdepääs metaandmetele	Reputatsiooni ja usalduse kadu, kui kvaliteet muutub ebaühtlaseks
Täiendada valikuuringu tulemusi mikroandmetega ja kasutada reaalselt mõõdetud tulemusi, mitte hinnangulisi	Saada aru, millist nähtust andmed tegelikult kirjeldavad	Privaatsuspiirangud ja andmete turvalisus
Parandada olemasoleva statistika kvaliteeti	Õiguslikud ja regulatiivsed küsimused, sh intellektuaalomand ja isikuandmete kaitse	Andmetele juurdepääsu mittejärgipidevus, kui andmeid salvestavat protsessi muudetakse
Võimalus siduda olemasolevaid andmeid uute andmeallikatega ja leida uusi seoseid	Andmete sobivus statistika eesmärgiga	Andmeomanike otsus andmeid mitte enam jagada
Võimalus kasutada uusi andmeanalüüsi meetodeid ja saada parem ülevaade andmetest	Hinnata andmete kvaliteeti	Nähtuste kirjeldamise definitsioonid muutuvad riikideüleselt olenevalt jurisdiktsioonist
Vähendada kulusid, sest kolmas osapool on juba andmeid kogunud	Tehniline võimekus andmeid töödelda	Ressursside äravõtmine enne suurandmete kasutuse juurutamist
Suurendada automatiseerimise kaudu statistika tootmise efektiivsust	Metoodiline võimekus toota andmetest ametlikku ja usaldusväärset statistikat	Andmekonkurents, sest eraettevõtted kasutavad andmeid kasumi teenimiseks
Töötajad saavad tegeleda keerukamate ülesannetega	Suurandmete andmeallikad on väga eripalgelised ja tuleb leida viis, kuidas rakendada eri andmeallikate puhul ühtseid standardeid	
Parem koostöö eri valdkondade vahel andmete vaheliste seoste kaudu		
Statistika parem nähtavus		

Allikas: Kitchen 2015.

Enne suurandmete kasutuselevõttu statistikas tuleb ületada mitu uuest andmestikust tulenevat proovikivi, aga tuleb tähele panna, et näivad anomaaliad võivadki olla andmestiku normaalne käitumine. Järgnevalt on refereeritud hollandlaste kogemust suurandmete kasutamisel statistikas (Daas jt 2015).

Andmete põhjalik tundmaõppimine. Tavaliselt ei ole suurandmete andmestikud loodud selleks, et teha nende põhjal analüüsi ja statistikat. Pigem on need mingi muu protsessi tulem ja seetõttu tuleb väga põhjalikult süüvida andmetesse, leida lisaks kirjeldavaid metaandmeid ning uurida põhjalikult struktuuri. Tuleb uurida andmete kvaliteeti, puuduvaid andmeid, anomaaliaid jm. Üks võimalus leida anomaaliaid on visualiseerimine.

Puuduvad andmed. Mõni protsess, mis tekitab andmeid, ei ole pidev. Seetõttu tekivad andmetes augud. Nii tulebki leida viise, kuidas selles andmestikus olevad tühimikud kõige paremini ära täita. Need meetodid ei pruugi olla samad, mis on statistikas praegu kasutusel.

Volatiilsus. Mõned suurandmete allikad paistavad silma suure volatiilsusega ja näiteks mõõtmise ebaühtlane intervall tekitab äkilisi muutusi andmetes. Vajadusel saab need siluda kas või libiseva keskmise meetodil.

Selektiivsus. Suurandmed võivad esindada vaid teatud osa kogupopulatsioonist. Näiteks Twitteri andmestikku kasutades tuleb tähele panna, et seal on aktiivsed pigem nooremad ja üldse tegusama meelelaadiga inimesed. Samuti ei pruugi olla andmestikuga seotud need üksused, kelle kohta statistikat tehakse.

Juriidilised ja privaatsuse probleemid. Suurandmete kasutuselevõtuga tekivad kohe juriidilised probleemid ja see on seotud ka inimeste privaatsusega. Teatud andmeallikatele on seatud juriidilised piirangud ja nende kasutuselevõtt statistika tegemiseks on väga raske. Andmeallikaid kombineerides võib saada ühe indiviidi kohta ilmnevat teadmist.

Andmete haldamine. Statistika eeldab tavaliselt pikemate aegridade uurimist, kuid suurandmete puhul on oma mahu tõttu töötlemiseks vaja märkimisväärset arvutusvõimsust. Samuti muutuvad andmeallikad kiiresti ja algselt ühte nähtust mõõtnud protsess transformeerub milleski teiseks. Seetõttu tuleb leida paindlikke andmete haldamise viise ning säilitada maksimaalselt metaandmestikku, mis aitab tuvastada vanemate andmete olemust.

Suur arvutusjõudlus. Pikkade ja suurandmetel põhinevate aegridade analüüsiks on vaja ka suurt arvutusvõimsust. Seetõttu peab aegsasti planeerima vajalikku riistvara ja pigem võiks kaaluda horisontaalselt laiendatavat arhitektuuri.

Uued oskused. Suurandmete töötlemiseks on vaja oskusi, mida traditsioonilisel statistikul ei ole kas siis täielikult või osaliselt. Peamiselt on need seotud arvuti kasutamise oskuse ja programmeerimisega, kuid samuti on suurandmetega vaja teisi meetodikaid ja tehnikaid võrreldes tavaliste andmeridadega. Suurandmete töötlemine eeldab ka julgemat andmete uurimist ja uute seoste leidmist.

Siinkohal mõned näited, kus on suurandmeid statistika tootmisel kasutatud.

Liiklusstatistika. Hollandi statistikaamet on kasutanud liikluse monitoringufot liiklusstatistika tegemisel (Daas jt 2015). Ühe päeva liiklusandmeid visualiseerides tuvastasid nad selged tippnunid hommikul kella 8 ja õhtul kella 5 ajal. Andmestikust oli võimalik tuvastada sõiduki suurus, ka olid suurte sõidukite tippnunid hilisemad kui väikeste omad. Töö näitas, et liikluse monitoringu põhjal saab peale liiklusstatistika hinnata ka majanduslikku aktiivsust ja majanduskasvu.

Tarbijate kindlustunne. Hollandi statistikaamet analüüsis ka sotsiaalmeedia andmeid ja püüdis nende sisu ning tunnetuse alusel hinnata tarbijate kindlustunnet (Daas 2015). Uurimuses kombineeriti Facebooki ja Twitteri andmestikku ja selekteeriti kasutajate tundeid väljendavad sõnumid, nii saavutati kokkulangevus tarbija kindlustunde indeksiga 88% ulatuses.

Turismistatistika ja rahvastiku dünaamika. Eestis on kasutatud mobiilandmeid, et hinnata inimeste liikumist olenevalt kellaajast ja nädalapäevast (Järv jt 2012). Töös leiti, et reedeõhtune tippnund on tunduvalt erinev teiste nädalapäevade omast, sest liiklusvoog on intensiivsem ja seotud pikema maa läbimisega ehk siis seotud siseturismi ehk nädalalõpu veetmisega väljaspool elukohta.

ESSneti suurandmete projekt

2002. aasta Palermo kohtumisel leidsid Euroopa statistikaametite peadirektorid, et oleks vaja paremini arendada ja jagada kogemusi Euroopa Statistikasüsteemis (ESS). Eesmärgi saavutamiseks otsustati luua kompetentsivõrgustike ja keskuste projekte (mida nüüd kutsutakse ESSnetiks^a), mille kaudu oleks võimalik panna kokku kõikide osalevate statistikaametite eksperdid ja nende kogemused ning arendada ühiselt kõigile süsteemi osalistele kasulikke teadmisi. ESSneti projekt on ESS-i organisatsioonide võrgustik, mis pakub kõigile osalistele kasulikke lahendusi. Üks selliseid on ka ESSneti suurandmete projekt.

Euroopa statistikaametid (National Statistics Institute – NSI) on märganud suurandmete potentsiaali ja Scheveningeni memorandumiga (European ... 2013) seati eesmärk leida võimalusi suurandmete integreerimiseks Euroopa statistikasüsteemi. Memorandumiga pandi siht uurida võimalusi suurandmete kasutamiseks statistika tootmisel ja sh hinnata, millised on kasutuselevõttuga seotud metodoloogilised, kvaliteedi ja infotehnoloogiaga seotud probleemid. Täpsema tegevuskavaga (European ... 2014) pandi alus projektile prooviuringute tegemiseks. Dokumendis on välja toodud ka potentsiaalseid prooviuringu valdkondi.

Tabel 2. Potentsiaalsed andmeallikad ja prooviuringud

Andmeallikas	Andmetüüp	Statistikavaldkond
Mobiilne side	Mobiiliandmed	Turismi- ja rahvastikustatistika
Veeb	Veebiotsingud	Töajõu- ja migratsioonistatistika
	E-kaubanduse veebilehed	Hinnastatistika
	Ettevõtete veebilehed	Innovaatilisus, toimla aadressid
	Töökuulutuste veebilehed	Töajõustatistika
	Kinnisvara veebilehed	Hinnastatistika
Sensorid	Liikluse mõõtajad	Liiklus- ja transpordistatistika
	Kaugmõõtajad	Energiastatistika
	Satelliidipildid	Maakasutus-, põllumajandus- ja keskkonnastatistika
Tehingud	Veesõidukite automaatne identifitseerimissüsteem (AIS)	Transpordi- ja keskkonnastatistika
	Kantavad seadmed	Ajakasutusstatistika
	Lendude registreerimine	Transpordi- ja keskkonnastatistika
Ühisloome (<i>crowdsourcing</i>)	Kaupluste lojaalsuskaardid ja seotud ostud	Hinna ja leibkonna kulude statistika
	Finantstehingud	Leibkonna kulude statistika
	Vabatahtlikud geograafilised veebilehed (OpenStreetMap, GeoWiki)	Maakasutusstatistika

Allikas: European ... 2014.

Euroopa statistikasüsteemi (ESS) riigid on algatanud projekti ESSnet Big Data, et leida võimalusi suurandmete kasutamiseks statistika tootmisel. Selles on 22 partnerit 20 riigist, sh ka Eesti. Tulemusi kajastatakse projekti veebilehel (ESSnet ... 2017). Projektiga alustati 2016. aasta veebruaris ja see kestab 2018. aasta maini. Kokku on selles konkreetse andmeallika või meetodikaga seotud 8 alamprojekti ning peale selle veel juhtimise ja kommuniqueerimise alamprojektid. Esimesed viis alamprojekti tegelevad reaalsete suurandmete töötlemise katseprojektidega.

^a https://ec.europa.eu/eurostat/cros/page/essnet_en.

Alamprojektide eesmärgid:

- vabade töökohtade veebikammimine (*webscraping*) – eesmärk on otsida veebist ja koguda kokku tööjuportaalides olevad töökuulutused ning teha selle põhjal vabade töökohtade statistikat;
- ettevõtete andmete veebikammimine – eesmärk on leida veebist ettevõtetega seotud andmeid ja neid kasutades täiendada ettevõtteregestrit näiteks ettevõtte peamise tegutsemiskoha ja toimlate aadressidega;
- kaugmõõtjad – kasutada kaugloetavate elektriarvestite andmeid, et asendada olemasolevat vaikuuringul põhinevat elektrienergia statistikat ja arendada uusi statistikatooteid;
- AIS-i andmed – kasutada veesõidukite automaatse identifitseerimise süsteemi asukohaandmeid (Automatic Identification System) olemasoleva statistika parandamiseks või uue loomiseks;
- mobiiliandmed – uurida võimalusi mobiiliandmetele juurdepääsu saamiseks ja leida viise, kuidas mobiiliandmeid statistikas kasutada;
- kiirhinnangud – hinnata võimalust andmeallikate kasutamiseks statistika kiirhinnangute tootmisel;
- eri valdkonnad – leida uusi andmeallikaid ja kombineerida neid võimalike statistikatoodete jaoks;
- meetodika – arendada meetodikat suurandmete kasutuselevõtuks statistikas.

Elektriandmete katseprojekt

Administratiivandmete kasutamine statistika tegemisel võimaldaks vähendada kulu ühiskonnale, mis tuleneb aruannete esitamisest ja sellega seotud halduskoormusest. Paljude suurandmete näitel võimaldaks andmeallikas kajastada ka tervet nähtust korraga. Näiteks elektriandmete korral oleks võimalik saada ülevaade kogu riigi elektritarbimisest. Administratiivandmete kasutamine võimaldaks ka märgatavalt lihtsustada statistika tegemise protsesse ja neid automatiseerida. Teisalt ei pruugi suurandmetes olla mõõtmistulemusi nende üksuste kohta, mille kohta avaldatakse statistikat. Suurandmete kasutamise kogemuste saamiseks osaleb Eesti koos Taani, Rootsi, Austria, Portugali ja Itaaliaga elektri kaugmõõtjate katseprojektis. Järgnevalt tutvustatakse katseprojekti esimesi tulemusi.

Katseprojektile seati kolm eesmärki:

- kas elektri mõõteandmete põhjal saaks teha seni küsitluse teel saadud elektrienergia statistikat ettevõtete kohta;
- kas elektriandmetest saaks toota kodumajapidamiste elektritarbimise statistikat ja
- kas elektriandmetest oleks võimalik tuvastada eluruume, milles elatakse, mis on tühjad või mida kasutatakse periooditi.

Katseprojektis kasutati Eleringi^a andmelao 2013.–2015. aasta andmeid, mis sisaldavad mõõteandmeid ligi 730 000 mõõtepunkti (elektriarvesti) kohta. Andmestikus on infot sõlmitud lepingute, lepingu omaniku (äriettevõtte, eraisik), lepingu kehtivuse, lepinguga seotud mõõtepunkti ja selle aadressi kohta, samuti mõõteandmed. Statistikatöodes kasutatavad andmed anonümiseeritakse Statistikaametis. Andmestikus on mõõteandmeid kokku üle 18 miljardi kirje. Kuna andmestikku kirjeldavaid metaandmeid ei olnud, siis tuvastati ise lõpptarbimises mitteosalevate mõõtepunktide võimalikud tunnused. Pärast lõpptarbimises mitteosalevate mõõtepunktide väljaarvamist arvatati 2015. aasta lõpptarbimiseks 8,2 TWh, mis erineb Eleringi veebilehel avaldatud lõpptarbimisest

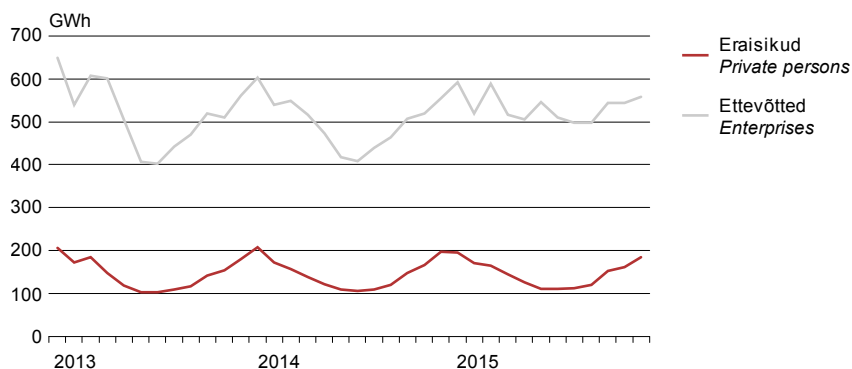
^a <https://elering.ee>.

7,4 TWh, st lõpptarbimise arvutusreegleid tuleb veel korrigeerida. Ettevõtetega seotud mõõtepunkte oli 2015. aastal 12% ja eraisikutega 88%, samas moodustasid ettevõtted 78% ning eraisikud 22% tarbimisest.

Suurandmete kasutuselevõtuks tuleb leida meetodika, kuidas andmetes olevad üksused siduda statistiliste üksustega ja kuidas hinnata selle sidumise kvaliteeti. Sellest tulenevalt on peaprobleem, kuidas siduda mõõdetav üksus (kaugmõõtjaga mõõtepunkt ehk koht, kus mõõdetakse elektritarbimist) statistilise üksusega, nt ettevõtte / institutsioon või leibkond. Kui esmapilgul tundub see lihtne ülesanne, siis detailsemal uurimisel tuleb välja keerukus. Probleem on see, et mõõtepunkti seotud elektritarbimise lepingu omanik ei pruugi olla see, kes on energia lõpptarbija. Seda ettevõtetes, kus näiteks suuri ärihooneid haldavad kinnisvaraettevõtted, mis üürivad neid välja eri tegevusvaldkondade ettevõtetele, kuid ka eraisikute puhul, kes samuti elavad üüripindadel. Üks paljude tarbijate kooslus on ka korteriühistu, kus võivad olla koos peale majapidamiste ka ettevõtted. Probleemi võiks saada lahendada, kui siduda lõpptarbijad nende reaalse tarbimisega aadressi kasutades, kuid ka siin tekivad probleemid. Esiteks võib ühel aadressil olla mitu mõõtepunkti ja mitu tarbijat korraga ning siis ei ole selget reeglit, mille alusel lõpptarbijaid tarbimisega kokku panna. Teiseks muudab ülesande tunduvalt keerulisemaks see, et nii äriregistris kui ka elektriandmete andmelaos on aadress vabatekstina ja see suurendab märkimisväärselt ebatäpsusi aadressides.

Joonis 5. Eraisikute ja ettevõtete summeeritud elektritarbimine kuus^a, 2013–2015

Figure 5. Monthly aggregated electricity consumption of private persons and enterprises^a, 2013–2015



^a Esialgsed tulemused.

^a Preliminary results.

Aadressid tuli korrastada, et siduda mõõteandmeid teiste registritega, samuti tuli leida linkimiseks vajalik unikaalse aadressi identifikaator. Maa-ametis on välja arendatud aadressiandmete süsteemi infosüsteem^a, mis seob iga unikaalse aadressiobjektiga aadressi. Nii on igal objektil, mis nimetatud aadressil asub, ka oma unikaalne aadressiobjekti identifikaator. Näiteks kui ühel aadressil on mitu hoonet, siis igal hoonel on unikaalne identifikaator. Statistikaamet kasutas Maa-ameti veebileheküljel pakutavat teenust, et viia vabatekstilised aadressid normaliseeritud kujule ja leida aadressile vastav aadressi ning aadressil asuva objekti identifikaator. Kui aadressi tekst sisaldas sõnu, mis polnud aadressiga seotud, siis aadressi ei normaliseeritud või leiti sellele lähim aadress. See suurendas märgatavalt varjatud vigasid. Kõigile mõõtepunktide aadressidele leiti 93% juhtudel vastava aadressi identifikaator, sh ettevõtetel 81% ja eraisikutel 95% juhtudest. Ettevõttereegistrist linkimiseks selekteeritud ettevõtetest õnnestus aadressi identifikaator leida 96% juhtudest. See ei tähenda, et mõlema andmekogu aadressid oleksid kohe seotavad, sest süsteem võis leida samale aadressile asustusüksuse tasemel aadressi identifikaatori või siis hoone osa identifikaatori, mistõttu neid siduda ei õnnestu.

^a <http://xgis.maaamet.ee/adsavalik/ads>.

Ettevõtete elektritarbimise hindamiseks kasutati võrdlusandmestikuna valikuuringu küsitlusega toodetud statistikat. Elektriandmete kliendibaas ei sisalda infot ettevõtte tegevusvaldkonna kohta ja peale selle võib ettevõtte peamine tegevusvaldkond muutuda vastavalt käibestruktuurile. Seetõttu tuli ettevõtte klienditabelis siduda valitud ettevõtetega äriregistris (150 000 ettevõtet), mille kohta talletab Statistikaamet igal aastal tegevuskoha ja peamise tegevusala. Sidumiseks / linkimiseks kasutati kolme strateegiat. Esimene oli mõõtepunkti lepingu omaniku sidumine äriregistriga registrikoodi kasutades, teiseks kasutati nii mõõtepunkti kui ka ettevõtte aadressi identifikaatorit (selekteerides välja ainult ettevõtetega seotud mõõtepunktid) ning kolmandaks kasutati nii mõõtepunkti kui ka ettevõtte aadressi identifikaatorit (kasutusel kõik mõõtepunktid). Esimesel juhul õnnestus siduda 22 000, teisel puhul 25 000 ning kolmandal 63 000 ettevõtet. Linkimise tegid keeruliseks mitu-mitmele-seosed, kus ühel aadressil oli mitu mõõtepunkti ja ettevõtet, mistõttu oli raskendatud reaalse lõpptarbija sidumine mõõdetud näiduga. Võrreldes 2015. aasta küsitluse tulemusega kaeti esimese strateegiaga 88% ettevõtete elektritarbimisest, teisel juhul 92% ja kolmandal 95%. Samas on selged erisused tegevusvaldkonniti ja näiteks kinnisvarasektori elektritarbimine moodustab 160% uuringu tulemusest ning hulgi- ja jae-kaubanduse tarbimine vaid 50%. Kui vaadata absoluutarve, siis elektriandmete põhjal arvatud kinnisvarasektori ülejääk katab ära hulgi- ja jae-kaubanduse ning majutusteenuste puudujäägi võrreldes uuringuandmetega.

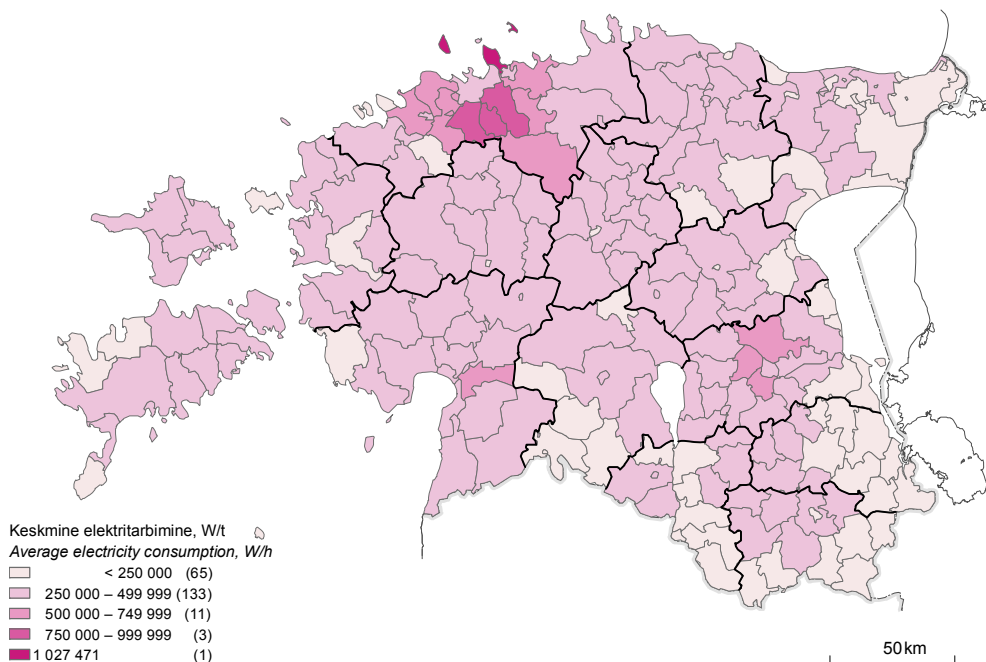
Teine ülesanne oli leida võimalusi leibkondadega seotud statistika tootmiseks. Leibkondade puhul elektrikasutuse uuringut regulaarselt ei tehta, mistõttu võrdlusandmeid ei olnud. Linkimiseks kasutati Statistikaameti leibkondade ja eluruumide tabelit, mis sisaldas leibkonnaliikmete anonümiseeritud infot ja aadressi identifikaatorit. Elektriandmetega sidumiseks leiti leibkonnad, kus aadressi identifikaator ühtis mõõtepunkti omaga ja ainult ühel leibkonnaliikmel oli elektrileping. Nii õnnestus siduda 48% leibkondadest elektritarbimise infoga. Samuti oli võimalik siduda elektritarbimisega eluruumi info, millega saab hinnata näiteks elektritarbimise seotust hoone ehitusaasta ja eluruumi suurusega ning see pakuks võimalust hinnata eluruumi energiatõhusust. Saadud tulemus võimaldaks toota statistikat, mida Statistikaametil ei ole siiani olnud võimalik toota andmete puudumise tõttu.

Kolmandaks hinnati, kas elektriandmete põhjal saaks kinnitada eluruumide kasutamist püsiva elukohana. Leiti tuhandeid aadresse, kus rahvastikuregistri järgi peaksid elama inimesed, kuid elektritarbimine oli 0 või siis vastupidi: oli suur elektritarbimine, kuid registri järgi seal kedagi ei elanud. Seetõttu on andmestikul oluline potentsiaal rahvastiku paiknemise analüüsil.

On hinnatud, et ettevõtete uuringupõhise elektritarbimise statistika asendamiseks mõõteandmete põhjal arvatud statistikaga on vaja lisaanalüüsi selle kohta, kuidas paremini leida vastavust lõpptarbija ja mõõtepunkti vahel ning kuidas lahendada mitu-mitmele-seosed, mis tekivad aadressiga linkides.

Eratarbijate suurt osatähtsust arvestades võimaldab juba olemasolev linkimise tase hinnata majapidamiste elektritarbimist. Et varem sellises mahus andmeid statistika tegemiseks kogutud ei ole, siis võiks sellest saada uus statistikatoode. Samuti saaks elektriandmeid kasutades parandada rahvastiku paiknemise statistikat.

Oluline järeldus on ka, et registrite kvaliteeti on vaja parandada, et minna üle suuranndmete põhisele statistikale ja automatiseerida protsesse. Kõik infosüsteemide arendajad võiksid kasutada standardseid aadresse koos aadressi identifikaatoritega, sest peale täpsuse muudab see aadresside uuendamise palju lihtsamaks. Aadresse ei pea uuesti infosüsteemi sisestama, vaid saab automaatsete reeglitega teisendada uuele kujule. Mida rohkem standardiseeritud lahendusi kasutatakse, seda efektiivsemalt on võimalik toota ka statistikat.

Kaart 1. Eraisikute elektritarbimine asustusüksuse tasemel^a, jaanuar 2014*Map 1. Electricity consumption of private persons on settlement unit level^a, January 2014*^a Esialgsed tulemused.^a Preliminary results.**Kokkuvõte**

Suurandmete kasutuselevõtt tähendab, et kuskil on kogunenud suur hulk infot ühe objekti / subjekti kohta. Konkreetselt Jüri või Mari tegemised ei paku statistikutele huvi, vaid pigem nende käitumine koos sarnase kogumiga – see loob andmete reaalse dünaamika. Üksikute andme-komponentide koostoisest ilmneb näiteks riigis loodud lisandväärtuse juurdekasv. Nagu igast sipelgate kokku toodud okkast saab nähtus, mida nimetatakse pesaks. Ilma konkreetsete okasteta ei oleks see just selline nagu on, kuid pesa kui tervikut vaadates ei ole vaja täpselt teada, milline sipelgas konkreetse okka kohale tõi. Seetõttu ei ole vaja karta ka oma andme-tükikeste jagamist, sest vaid nii saab luua reaalse tervikpildi. Suurandmed pakuvad just seda võimalust, et saada parim tervikpilt võimalikult palju detailseid algandmeid kasutades.

Euroopa statistikaametid on suurandmete kasutusvõimaluste hindamiseks statistika tootmisel loonud ESSneti suurandmete projekti. Projekti eesmärk on konkreetsete näidete varal töötada välja meetodika ja paremad viisid suurandmete kasutuselevõtuks statistikas. Esialgsed tulemused viitavad suurandmete kasutamise suurele potentsiaalile ja võimalusele vähendada halduskoormust ühiskonnale, kuid edukaks rakendamiseks on vaja veel teha tööd, et luua selge ja läbipaistev meetodika suurandmetest statistika tootmiseks.

Suurandmete kasutamisega võivad kaasneda ka teatud riskid ja andmed võivad olla kallutatud, mittetäielikud või siis kajastada nähtust, mille kohta ei tehta statistikat. Niisugune oli kogemus ka elektriandmete katseprojekti, kus andmestik on lepingu omaniku põhine, kuid statistikatulem peaks olema lõpptarbijapõhine. Katseprojekti tulemustest võib järeldada, et alati ei pruugi saada ka kogu nähtust kirjeldava andmestiku olemasolul kohe toota traditsioonilist statistikat. Samas annavad sellised andmestikud alternatiivseid võimalusi info pakkumiseks ühiskonnale. Uue statistika väljatöötamine on ka suurandmete jätkuprojekti teema.

Suurandmete kasutamine võimaldab teha ka täiesti uut statistikat nagu näitab Eurostati korraldatud Hackathon'i töö (Õpik jt 2017), kus visualiseeritakse ameteid ja nendega seotud oskusi tööportaalide suurandmeid kasutades. Sellised tooted annavad ühiskonnale parema tagasiside ja pakuvad vaadet tulevikku. Tänapäeva andmepõhises maailmas võiks kaaluda ka lahendust, kus andmeid saaks vaadata avaliku hüvena (Mitchell, Brynjolfsson 2017) ning ettevõtte, kelle tegevuse tulemusena koguneb suurem hulk andmeid, võiksid ühiskonnale peegeldada seda, mida nad ühiskonna käest on kokku kogunud. Suurandmete kasutuselevõtt võimaldab avaldada täiesti uut infot, see vähendab märgatavalt halduskoormust, samuti saab jälgida nähtusi soovi korral reaalsajas ja ühiskond saab otsuste tegemiseks vajaliku info tunduvalt kiiremini. Kui kasutada suurandmeid metodoloogiliselt läbimõeldult, siis on võimalik toota statistikat odavamalt, kiiremini ja palju täpsemalt, aga selleks peab olema ka andmeomanike valmisolek jagada andmeid statistika tootmiseks.

Allikad Sources

Braaksma, B., Zeelenberg, K. (2015). Re-make/Re-model: Should big data change the modelling paradigm in official statistics? – *Statistical Journal of the IAOS*, 31(2).

Breiman, L. (2001). *Statistical Modeling: The Two Cultures* *Statistical Science*, 16(3).

Daas, P. J. H., Puts, M. J. H., Buelens, B. and van den Hurk, P. A. M. (2015): Big data as a source for official statistics. – *Journal of Official Statistics*, 31, 249–269.

Dean, J., Ghemawat, S. (2004). MapReduce: simplified data processing on large clusters, *Proceedings of the Sixth Symposium on Operating System Design and Implementation*, p.10.

Diebold, F. X. (2012) Personal Perspective on the Origin(s) and Development of 'Big Data': The Phenomenon, the Term, and the Discipline, Second Version (November 26, 2012). – PIER Working Paper No. 13-003. [doi] <http://dx.doi.org/10.2139/ssrn.2202843> (16.10.2017).

Eesti Keele Instituut. (2014). Suurandmed. Eesti Keele Instituudi mitmekeelne terminibaas. [www] <http://termin.eki.ee/esterm/concept.php?id=79384&term=suurandmed> (10.10.2017).

ESSnet Big Data. (2017). Big data: from exploration to exploitation. [www] https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/index.php/Main_Page (11.10.2017).

European Statistical System Committee (2013): Scheveningen Memorandum on Big Data and Official Statistics.

European Statistical System Committee (2014): Big Data Action Plan and Roadmap.

Florescu, D., Karlberg, M., Reis, F., Reis, P. R. Del Castillo, Skaliotis, M., Wirthmann, A. (2014). Will 'big data' transform official statistics? [www] http://www.q2014.at/fileadmin/user_upload/ESTAT-Q2014-BigDataOS-v1a.pdf (11.10.2017).

Ghemawat, S., Gobiuff, H., Leung, S.-T. (2003). The Google file system, *ACM SIGOPS Operating Systems Review*, 37(5).

Japac, L., Kreuter, F., Berg, M., Biemer, P., Decker, P., Lampe, C., Lane, J., O'Neil, C., Usher, A. (2015). Big Data in Survey Research: AAPOR Task Force Report. *Public Opinion Quarterly*, 79(4).

Järv, O., Ahas, R., Saluveer, E., Derudder, B., Witlox, F. (2012). Mobile Phones in a Traffic Flow: A Geographical Perspective to Evening Rush Hour Traffic Analysis Using Call Detail Records. *PLoS ONE* 7(11): e49171. [doi] doi:10.1371/journal.pone.0049171.

- Kitchin, R., McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets – *Big Data & Society*, January-June 1–10.
- Kitchin, R. (2015). The opportunities, challenges and risks of big data for official statistics. *Statistical Journal of the IAOS*, 31, 471–481.
- Laney, D. (2001). 3-d data management: controlling data volume, velocity and variety. META Group Research Note, 6 February.
- Lynch, C. (2008). Big Data: How Do Your Data Grow? – *Nature*, 455.
- Mitchell, T, E. Brynjolfsson. (2017). Track how technology is transforming work. – *Nature*, 544(290). [doi] doi:10.1038/544290a.
- Õpik, R., Kirt, T., Liiv, I. (2017). Megatrend and Intervention Impact Analyser for Jobs: A European Big Data Hackathon Entry. arXiv:1708.08262.
- Särndal, C.-E., Wretman, J., Swensson, B. (1992). *Model Assisted Survey Sampling*. Springer, New York.
- Singh, D, Reddy, C.K. (2014). A survey on platforms for big data analytics. – *Journal of Big Data*,1(8).
- Tiit, E.-M. (2016). Suurandmed statistikas [www] <https://blog.stat.ee/2016/08/25/suurandmed-statistikas/> (9.10.2017).
- Tutorialspoint (2017). MapReduce – Introduction. [www] https://www.tutorialspoint.com/map_reduce/pdf/map_reduce_introduction.pdf (15.11.2017).
- UNECE Wiki. (2017). UNECE Big Data Inventory Home. UNECE Statistics Wikis. [www] <https://statswiki.unece.org/display/BDI/UNECE+Big+Data+Inventory+Home> (11.10.2017).
- UNECE. (2013). *Generic Statistical Business Process Model GSBPM*. United Nations Economic Commission for Europe.
- White, T. (2015). *Hadoop: The Definitive Guide*, 4th Edition. O'Reilly Media.

BIG DATA IN STATISTICS

Toomas Kirt

In this article, an overview is given of how big data have been defined and what kind of tools are used to process such data. After that, the potential use of big data in official statistics is examined: how is Statistics Estonia involved in the big data pilot project which looks into the potential use of the data recorded by electricity smart meters in the production of statistics.

Introduction

The following is an attempt to elaborate on what lies behind the term 'big data' and how big data are related to statistics. Big data do not mean that numbers are large by size or value but rather a phenomenon of a large number of such data. In order to be called big data, data do not necessarily have to be presented as numbers but can just as well take the form of other symbols, e.g. written text. But in order to become big data, there must be a large amount of these symbols, so that one would not be able to process them using conventional methods.

Big data are generated mainly when a constant or regular process is measured on the basis of one or many parameters, or other results of the process are recorded. The higher the number of parameters and the greater the frequency of measuring, the more data are received. Population censuses are conducted every ten years and all the Estonian citizens are counted (1.3 million); as a result of the census-taking, answers to approximately 100 questions are received. Consequently, a large amount of data are captured, but these are, rather, not considered to be big data. The same amount of people who use their mobile phones daily generate more data in a single day than is possible to collect with a population census. The recording and processing of such data already require special methods, and these are considered to be big data.

Big data

Most people who work with data have heard the term 'big data'. The term has been defined in the terminological database of the Institute of the Estonian Language (2014) as large, extensive, structured and unstructured data which are difficult, if not impossible, to manipulate with standard relational database and data processing tools (volume, format, speed). Big data are mainly characterised by the need to use other than traditional processing tools. As the data processing capacity is steadily increasing in time, what are big data today may not be big data in the future. Big data are also characterised by how to get an overview of them. It is possible to physically check row by row a couple of thousand records in a database, but from certain quantities, a disproportionate amount of time is required for this activity, and to get an overview of the data, visualisation or other solutions that open up the structure of the data should be used. But data may not only be big in volume – they can also be big in terms of importance, if they reflect an observation of a unique phenomenon (Lynch 2008), which cannot be restored, and therefore it is important to ensure the preservation of the associated data. (Figure 1, p. 34)

The term 'big data' originates from the mid-1990s (Diebold 2012). The model of 3Vs (2001), introduced by Doug Laney, which assesses the management of the explosively increased amount of data through volume, velocity and variety, could be considered the framework for defining large data volumes. The model of 3Vs states that with the offering of e-services, the amount of data and the speed of data movement and processing increase considerably and the data format becomes more varied, including both structured data (data which can be entered in a typical relational database) as well as unstructured data (data with no specific structure, e.g. texts and video recordings). When big data have been defined, it has been pointed out that if conventional data rather represent a sample of some phenomenon, big data rather represent the

whole population and are always related to the speed of data generation and processing (Kitchin and McArdle 2016). The model of 3Vs has been supplemented with other v-words (such as data value, which indicates both the costs of storage as well as commercial revenue). Also other characteristics based on other letters have been proposed, but the definition of 3Vs is usually enough to describe the challenges faced during the processing of big data.

In connection to the increasing volume of recorded data, a paradigm shift has taken place in statistics (Japac et al. 2015), and statisticians are also looking beyond the methodology based on random sampling and surveys that has been the basis of their work. They are increasingly using administrative data sources / registers. The new paradigm also means that it is possible to digitally collect, link, aggregate and correlate various data. This way, a more varied data representation is achieved than solely with surveys. The introduction of big data enables to combine even more detailed data sources and identify new relations.

The analysis of big data is complicated by the exponential growth of the number of relations between data elements. For example, if one wished to find the relations between all objects, the square matrix of these links must be found, the number of the elements of which increases by the square of the number of objects. Therefore, the analysis of big data requires different approaches, especially if one wishes to draw quick conclusions from the data flow. The use of big data also involves risks (Japac et al. 2015), and to process these data, statisticians need new knowledge and skills. When introducing big data, it should also be taken into account that the data may be biased. Based on, for example, Twitter data, we may draw biased conclusions, as Twitter users are mostly young and active. Big data may be flawed, selective and incomplete, and are often the result of processes, the aim of which is not to produce data for analysis.

Processing tools

Definitions of big data refer to the fact that the tools used to process such data differ from the conventional tools. The following attempts to describe a system used for the processing of big data – the Hadoop. This system is used also in Statistics Estonia. A short description has been provided also of its main operating principles. Big data are processed using easily scalable systems. Scaling can be vertical, i.e. acquiring a server with much higher performance, or horizontal, i.e. distributing processing between many servers and adding another server to the system to increase performance (Singh, Reddy 2014) (Figure 2. p. 35). If a conventional database system is used, data are continuously added and the capacity of the machine processing the data proves to be inadequate, additional memory, processor capacity or hard drive capacity needed for the processing of data are added to the database server, and work may continue – i.e. vertical scaling is used. However, at one point, given the increasing volume of data, a system based on one server becomes extremely expensive, as servers with large capacity usually cost quite a lot. In order to manage the increasing volume of data processing, a horizontally scalable big data processing system has been developed – the Apache Hadoop. A distributed architecture provides the advantage that to expand the system, one does not need to purchase the most expensive hardware. Distribution also ensures better reliability, as in the case of a distributed architecture, the failure of one component does not affect the work of the whole system. There are, no doubt, other approaches (Singh, Reddy 2014), but the Hadoop platform is freeware and is widely used by large technology companies.

The creation of Hadoop was based on the publication of an article in 2003 describing the operating principles of the Google file system (Ghemawat et al. 2003) and the article published in the following year introducing the operating principles of the MapReduce algorithm (Dean et al. 2004). Inspired by these ideas, a system based on a similar distributed architecture was started to be developed within the framework of the Apache Nutch project, and in 2006, the project leader Doug Cutting initiated a separate subproject Hadoop, which was named after his son's teddy bear. In Hadoop there are two main components: the NameNode (See in more detail in White 2015) and DataNode(s). In the NameNode, metadata of data are stored and the work of DataNode(s) is managed and coordinated. In the DataNodes, data are stored (See description of the operating principle below) and the queries received are processed using the MapReduce technology (See description below).

Hadoop is not just a file system, but a number of tools are associated with it, which are collectively called the Hadoop ecosystem, and it is also interfaced with various data analysis platforms. Tools in the Hadoop ecosystem include, e.g., Hive, which enables to form structured data tables from the data in various formats and make regular SQL queries. Other examples are Spark, a data processing engine, which keeps data as much as possible in the server memory to speed up queries, Yarn, which manages resources and schedules jobs, Scoop, which enables to interface Hadoop with other database systems, and many other tools. In addition, many managing and interfacing systems have been developed, such as Ambari, Hue and Zeppelin, and it is also possible to use Python's Jupyter Notebook and R Studio, which are often used by data analysts. Hadoop offers the capacity to manage large quantities of data, and is also flexible and provides opportunities for presenting the results of the analysis in an appropriate form.

Operating principle of Hadoop

The Hadoop system is based on a distributed file system – HDFS (Hadoop Distributed File System). Files in all formats and sizes can be stored in Hadoop. Large files are split into parts (depending on settings, e.g., into parts 64 MB in size) and three copies are made of each, which are stored in different DataNodes (Figure 3, p. 36). In HDFS, the files look like ordinary files, although physically, parts of these files may be located at different ends of the world. When a query is made from the file, the NameNode divides it into parts and sends a query about a certain part of the file to the DataNode in a way that the system's resources are optimally used and the query would get a fast reply. The NameNode monitors that if there is no reply from a particular DataNode, the same information is requested from another DataNode where a copy of the same part is stored. This makes the system largely fault tolerant, as the loss of one DataNode does not prevent the system from working.

MapReduce

MapReduce is a technology which enables to make distributed queries^a. The system consists of two phases: Map, which groups the source data into a key and value pair, and Reduce, which aggregates the key and value pairs into the sought for reply. In Figure 4 (p. 37), an example has been provided, where the number of different letters in the file has to be counted. In the case of a distributed file system, the source file has been split into three and, therefore, three parallel Map operations are initiated, as a result of which, in this case, a list of letter and value pairs is drawn up. Next, the shuffling and sorting phase of the Map process starts, as a result of which, similar key and value pairs are grouped together. After this, the Reduce process begins to sum up the values by key, and one key and value pair is generated for each letter, i.e. the reply to how many times the letter appeared in the text. The interim results of the process are stored in work files, and active data exchange between the nodes also takes place, which means that the process demands a lot of input/output (I/O) resources. This provides a fault tolerant and efficient way for making queries in a distributed file system. For the end user, the use of MapReduce has been made automatic, and, e.g., in Apache Hive, the query can be written in regular SQL syntax (structured query language in databases), and this is automatically converted on system level into MapReduce queries.

Big data in official statistics

The increased capacity to effectively process large quantities of data has attracted also statisticians, who are hoping to find alternative ways to produce statistics. The current statistical system is mostly based on sample surveys (Särndal, Wretman, Swensson 1992), which is supplemented with administrative data from state registers. The introduction of big data, however, offers statisticians new opportunities to produce statistics also on the basis of the parent population. At the same time, it has been observed that the use of big data has its limits, and that all the data collected with surveys or censuses cannot be replaced with big data, as the latter do not necessarily cover the whole population, nor do they measure exactly what is reflected with

^a See, for example, Tutorialspoint 2017 (https://www.tutorialspoint.com/map_reduce/pdf/map_reduce_introduction.pdf).

census statistics (Tiit 2016). The introduction of big data also requires the use of other technologies and methods, which may not be currently available to statistical offices. Also the integration of big data into the standardised statistical workflow (UNECE 2013) takes time. Problems caused by the introduction of big data into statistics are being solved and new data sources are being explored actively. There are some 20 projects in the course of which big data are tried to be used in statistics (UNECE ... 2017).

With statistics, the numerical values describing some parent population are tried to be communicated. The parent population may, for example, include the whole population or, in the case of a certain survey, wood processing enterprises. The most accurate statistics are produced if it is possible to collect data about all the elements of the parent population, although this is usually very expensive and requires vast resources. This is why in statistics mostly surveys based on random samples are used, the results of which are extrapolated to the parent population. The method is in most cases very effective and enables to produce good enough statistics. Without a doubt, the method also has disadvantages, i.e. even though sampling is random, the extrapolated result may differ from reality in the case of small samples. The result is also affected by how precise information has been submitted. Therefore it is important to develop a methodology which takes into account all these phenomena, and thereby it is possible to improve the quality of the result.

For a better overview of the parent population, the use of registers and other administrative data sources has become widespread, as registers should reflect the whole to a large extent. The introduction of big data should produce an even more accurate result, as big data reflect some phenomenon in great detail and cover various aspects of the parent population.

Often in statistics, models are used which enable to fill in certain gaps in data. For example, if one has the information that a certain family always tries to be twice as good as its neighbours, we can calculate the values of this family if we have information about the neighbours. If the neighbours have one car, the family has two; if the neighbours have one house, the family has two; if in the neighbouring family there is one man... . In statistics, mostly simple linear models are used, the relationship between the variables of which can clearly be verified. In the case of big data, most likely an approach based on machine learning is used, where the model's parameters are optimised with learning algorithms, and relationships between the variables are found. However, there is no clear overview of the content of the model itself, and it is, so to say, a black box. This is risky in statistics where transparency is very important. Therefore, there is a lot of discussion (Braaksma, Zeelenberg 2015; Breiman 2001) about whether and when models can be used in statistics.

The introduction of big data would enable to make considerable changes to the statistics production process. Florescu (2014) and his colleagues have outlined five ways of using big data in statistics:

- to replace the current data source (sample survey) with big data, keeping the current statistical output;
- to partly replace the current data source (sample survey) with big data, keeping the current statistical output;
- to provide additional statistical information from another perspective – supplementary statistical output;
- to supplement and specify the estimations made based on the current data source – supplemented statistical output;
- to provide completely new statistics based on new data sources – new and alternative statistical output.

The use of big data in statistics provides opportunities and poses challenges but also involves risks. These have been summed up in the following table.

Table 1. Using big data in the production of statistics: opportunities, challenges and risks

Opportunities	Challenges	Risks
Supplement, replace, improve existing data sources	Cooperation with owners of big data to find beneficial solutions for both	Changing of the mission of statistical offices – focusing too much on available data
Produce more up-to-date output	Get access to data	Continuity of time series if data sources change
Provide an alternative to statistics based on questionnaires and facilitate the work of respondents	Get access to metadata	Loss of reputation and reliability if quality becomes uneven
Supplement sample survey results with microdata and use results actually measured, not estimates	To understand what phenomenon the data actually describe	Privacy restrictions and data security
Improve the quality of existing statistics	Legal and regulative issues, incl. intellectual property and protection of personal data	Discontinuity of access to data if the data recording process is changed
Opportunity to combine existing data with new data sources and find new relations	Usability of data for statistical purposes	Decision of data owners to no longer share data
Opportunity to use new data analysis methods and get a better overview of data	Assess the quality of data	Definitions describing phenomena change on a transnational level depending on jurisdiction
Reduce costs, as data have already been collected by third party	Technical capacity for data processing	Withdrawal of resources before the use of big data has been implemented
Increase the efficiency of statistics production through automation	Methodological capacity to produce official and reliable statistics from data	Competing for data because private enterprises are using data to earn profit
Employees can work on more complex tasks	The sources of big data are very different and a way must be found for applying common standards to various data sources	
Better cooperation between various domains through links between data		
Better visibility of statistics		

Source: Kitchin 2015.

Before the incorporation of big data into statistics, many challenges posed by the new dataset have to be met, but one should note that what may appear as anomalies may in fact be normal behaviour of the dataset. Next, an overview of the experience of the Dutch in using big data in statistics has been presented (Daas et al. 2015).

In-depth examination of data. Usually the datasets of big data are not intended for the purpose of conducting analysis and producing statistics. They are rather the result of some other process, and therefore a thorough examination of the data is required, additional descriptive metadata need to be found and the structure is to be thoroughly examined. One needs to investigate the quality of the data, missing data, anomalies, etc. One possible way to find anomalies is visualisation.

Missing data. Some processes which generate data are not constant, and therefore gaps appear in data. This means that ways need to be found for filling in the gaps in the dataset as well as possible. These methods are not necessarily the same as those used in statistics at the moment.

Volatility. Some big data sources show great volatility and, e.g., an irregular interval of measuring causes sudden changes in data. If necessary, these can be overcome with, e.g., the method of moving averages.

Selectivity. Big data can represent only a part of the whole population. When using the Twitter dataset, for example, one should keep in mind that younger people mostly and people with more active mindsets in general prevail there. Also the units that statistics are produced for may not be related to the dataset.

Legal and privacy problems. With the introduction of big data, legal issues arise immediately, and it also has implications on the privacy of people. Legal restrictions have been set to certain data sources, and the incorporation of these into the production of statistics is very difficult. When different data sources are combined, information about individuals may emerge.

Data management. Statistics usually require examination of longer time series, but, due to their large volume, big data require considerable computing capacity. Data sources are also changing quickly and a process that initially measured one phenomenon transforms into something else. Therefore, flexible ways for data management need to be found and the metadata set, which helps to identify the nature of older data, be maintained as much as possible.

High computing capacity. The analysis of long time series based on big data also requires high computing capacity. Therefore, the necessary hardware needs to be planned early enough and horizontally scalable architecture should be preferred.

New skills. The processing of big data requires skills that a traditional statistician does not have at all or partly. These are mainly related to computer use and programming, but big data also require other methodologies and techniques compared to regular time series. Processing of big data also requires a braver examination of data and finding of new relations.

Some examples of domains where big data have been used in the production of statistics:

Traffic statistics. Statistics Netherlands has used traffic monitoring information to produce traffic statistics (Daas et al. 2015). When they visualised the traffic data for one day, they identified clear rush hours in the morning at 8 a.m. and in the evening at 5 p.m. The dataset enabled to identify the size of vehicles, and the rush hours of large vehicles occurred later than those of small vehicles. The work showed that based on the traffic monitoring information, besides traffic statistics, it is also possible to assess economic activity and economic growth.

Confidence of consumers. Statistics Netherlands also analysed social media data and, based on the content and sentiment of the data, tried to assess the confidence of consumers (Daas et al. 2015). After having combined the datasets of Facebook and Twitter and having selected the posts expressing the consumer's feelings, an 88% coincidence with the user confidence index was achieved.

Tourism statistics and population dynamics. In Estonia, mobile data have been used to assess the movement of people depending on the time and day (Järv et al. 2012). As a result, it was found that the rush hour on Fridays is significantly different from that on other days of the week, as on Fridays, the traffic flow is more intense and related to covering longer distances, i.e. related to the domestic tourism, or spending the weekend outside one's place of residence.

ESSnet Big Data project

At the 2002 Palermo meeting, Directors General of European national statistics institutes decided that expertise in the European Statistical System (ESS) should be developed and shared better. In order to achieve this aim, they proposed to create Centres and Networks of Excellence (now called ESSnet^a) projects to bring together the experts of all the participating statistical offices and their experience, and develop knowledge beneficial for the whole system. ESSnet Big Data project is one such project.

The European national statistics institutes (NSI) have noticed the potential of big data, and with the Scheveningen Memorandum (European ... 2013), a goal was set to find opportunities for integrating big data into the European Statistical System. With the memorandum, a target was set to examine the possibilities of using big data in the production of statistics, and also assess what kind of methodological, quality and information technology issues are associated with the introduction of big data. A more precise action plan was compiled (European ... 2014), which lay the ground for the conducting of pilot studies of the project. In the document, also potential domains of pilot studies have been outlined.

Table 2. Potential data sources and pilot studies

Data source	Type of data	Statistical subject matter domain
Mobile communications	Mobile data	Tourism and population statistics
Web	Web searches	Labour force and migration statistics
	E-commerce websites	Price statistics
	Websites of enterprises	Innovation, addresses of local units
	Websites of vacancy notices	Labour force statistics
Sensors	Real estate websites	Price statistics
	Traffic meters	Traffic and transport statistics
	Smart meters	Energy statistics
	Satellite images	Land use, agricultural and environmental statistics
Transactions	Automatic identification system of vessels (AIS)	Transport and environmental statistics
	Wearable devices	Time use statistics
	Flight registration	Transport and environmental statistics
Crowdsourcing	Loyalty cards of shops and scanner data	Price and household expenditure statistics
	Financial transactions	Household expenditure statistics
	Voluntary geographical websites (OpenStreetMap, GeoWiki)	Land use statistics

Source: European ... 2014.

The countries participating in the European Statistical System (ESS) have initiated the ESSnet Big Data project to find opportunities for using big data in the production of statistics. The participants of the project include 22 partners from 20 countries, incl. Statistics Estonia. The results are published on the website of the project (ESSnet ... 2017). The project was launched in February 2016 and it ends in May 2018. The project includes eight subprojects related to a particular data source or methodology, and, in addition, management and communication subprojects. The first five subprojects are pilot projects that deal with the processing of real big data.

^a https://ec.europa.eu/eurostat/cros/page/essnet_en.

Aims of subprojects:

- *webscraping of vacancies – to search the Web for and collect vacancy notices from job portals and use this information to produce job vacancy statistics;*
- *webscraping of enterprise data – to find on the Web the data related to enterprises, and with the help of this information to supplement the commercial register with the addresses of the enterprise’s main place of business and local units;*
- *smart meters – to use the data of smart electricity meters to replace the current electricity statistics produced on the basis of a sample survey and develop new statistical products;*
- *AIS data – to use the location data of the automatic identification system of vessels to improve the existing statistics or produce new statistics;*
- *mobile data – to investigate the opportunities for getting access to mobile data and find ways for using mobile data in statistics;*
- *flash estimates – to assess the prospect of using data sources in the production of flash estimates of statistics;*
- *various domains – to find new data sources and combine them for potential statistical products;*
- *methodology – to develop a methodology for the introduction of big data into statistics.*

Electricity data pilot project

Using administrative data in the production of statistics would help reduce the cost to the society arising from submitting reports and the associated administrative burden. Many sources of big data suggest that data sources could also reflect the whole phenomenon at a time. In the case of electricity data, for example, an overview of the electricity consumption in the whole country could be given. The use of administrative data would also allow considerable simplification and automation of statistics production processes. On the other hand, big data may not include measured values for units for which statistics are produced. In order to gain experience in using big data, Estonia, along with Denmark, Sweden, Austria, Portugal and Italy, participates in the pilot project of electricity smart meters. Next, the first results of the pilot project are presented.

Three goals were set for the pilot project:

- *whether it would be possible to produce the currently survey-based electricity statistics for enterprises on the basis of the metering data of electricity meters;*
- *whether it would be possible to produce electricity consumption statistics of households from electricity data and*
- *whether it would be possible to identify occupied, vacant and seasonally vacant dwellings from electricity data.*

In the pilot project, data of 2013–2015 from Elering^a data hub were used, which include metering data of approximately 730,000 metering points (electricity meters). The dataset includes information about concluded agreements, the agreement owner (commercial enterprise, private person), duration of agreement, metering point in the agreement and its address and also metering data. Data used in statistical activities are anonymised in Statistics Estonia. There are more than 18 billion metering data records in the dataset. As there were no available metadata describing the dataset, possible characteristics of metering points not included in the final consumption had to be detected. After excluding the metering points not included in the final consumption, the calculated final consumption for 2015 was 8.2 TWh, which is different from the final consumption figure 7.4 TWh published on the Elering website, which means that the rules

^a <https://elering.ee>.

for calculating final consumption need to be further improved. 12% of metering points were related to enterprises and 88% to private persons in 2015, while 78% of the electricity was consumed by enterprises and 22% by private persons.

For the introduction of big data, methodology must be developed for how to link units in data to statistical units and how to assess the quality of the linking. Consequently, the main problem lies in how to link the measured unit (metering point with a smart meter, i.e. the place where electricity consumption is measured) to a statistical unit, e.g. an enterprise / institution or a household. While it seems like an easy task at first glance, detailed investigation reveals its complexity. The problem is that the owner of an electricity consumption agreement associated with a metering point may not be the end consumer of electricity. This applies to enterprises, where, e.g., real estate enterprises manage large commercial buildings, which they rent out to enterprises of various economic activities, as well as to private persons, who also live on rented premises. There are many consumers also in apartment associations, which may include, besides households, also enterprises. The problem could be solved by linking the end consumers to their real consumption by address, but problems arise also here. First of all, there may be more metering points and consumers than just one at one address and, in this case, there is no clear rule about how to link end consumers to consumption. Secondly, the task is made a lot more complicated by the fact that the addresses in both the commercial register as well as in the electricity data hub are recorded as free text, which significantly increases the inaccuracy of addresses. (Figure 5, p. 43)

Addresses had to be sorted out in order to link the metering data to other registers, and unique address identifiers required for linking had to be found. The Land Board has developed the Address Data System (ADS)^a, in which each address is given a unique address identifier. In addition, each object situated at the given address has its own unique address object identifier. For example, if there are many buildings at an address, each of these buildings has a unique address object identifier. Statistics Estonia used the service provided on the website of the Land Board to convert the addresses recorded as free text into normalised format, and find the corresponding address and address object identifiers. If the address text included words not connected to the address, the address was not normalised, or the closest address was found, which considerably increased the number of non-observable errors. Corresponding address identifiers were found for 93% of the metering point addresses, incl. for 81% of enterprises and 95% of private persons. For enterprises selected for linking from the commercial register, address identifiers were found in 96% of cases. This does not mean that the addresses of the two databases are immediately linkable, as the system could have found for the same address an address identifier on settlement unit level or an identifier for a part of building, whereby they cannot be linked.

Statistics produced with a sample survey questionnaire were used as a comparative dataset to assess the electricity consumption of enterprises. The customer base of electricity data does not include information about the economic activity of enterprises, and the main economic activity of an enterprise may also change based on the turnover structure. Therefore, enterprises in the customer table had to be linked with selected enterprises in the commercial register (150,000 enterprises), for which, in Statistics Estonia, the place of business and main economic activity are registered yearly. Three strategies were used for linking. The first was linking the metering point agreement owner to the commercial register using the registry code; the second strategy involved using both the metering point as well as the address identifier of the enterprise (selecting only the metering points associated with enterprises), and in the case of the third strategy, both the metering point as well as the address identifier of the enterprise were used (using all the metering points). In the first case, it was possible to link 22,000, in the second, 25,000 and in the third, 63,000 enterprises. Linking was difficult due to many-to-many relationships, in which case there were many metering points and many enterprises at one address, whereby it was difficult to link

^a <http://xgis.maaamet.ee/adsavalik/ads>.

the real end consumer with the measured value. Compared to the 2015 survey results, with the first strategy, 88%, with the second strategy, 92% and with the third strategy, 95% of the electricity consumption of enterprises was covered. However, there are clear differences by economic activity, and, for example, the electricity consumption of the real estate sector accounts for 160% of the survey results and the consumption of wholesale and retail trade for only 50% of the survey results. Looking at absolute figures, the surplus of the real estate sector calculated based on the electricity data covers the deficit of the wholesale and retail trade and accommodation services compared to survey data.

Another task was to find ways for producing statistics on households. Households are not regularly surveyed on electricity consumption and, therefore, there were no reference data available. For linking, Statistics Estonia's table on households and dwellings was used, which included anonymised information and address identifiers of household members. For linking with electricity data, the households were used whose address identifier coincided with the address of the metering point and only one household member had an electricity agreement. Thereby we were able to link 48% of households with electricity consumption information. It was also possible to link dwelling information with electricity consumption, whereby, e.g., the correlation between the electricity consumption and year of construction and dwelling size can be assessed, which would later on provide the opportunity to assess the energy efficiency of the dwelling. The obtained result would enable to produce statistics that Statistics Estonia has thus far not been able to do due to lack of data.

Third, it was assessed whether it would be possible to confirm the use of a dwelling as a permanent place of residence based on electricity data. Thousands of addresses were found which should have been occupied according to the population register but where the electricity consumption was 0, or, on the contrary, addresses with high electricity consumption but no occupants according to the population register. This is why the dataset is of great potential when analysing the distribution of the population.

It has been estimated that in order to replace the survey-based electricity consumption statistics of enterprises with statistics calculated on the basis of metering data, further analysis is needed about how to better match the end consumer and the metering point and how to solve the many-to-many relationships which occur when linking with the address.

Considering the large share of private consumers, already the current level of linking enables to assess the electricity consumption of households. As such amount of data have not previously been collected for the production of statistics, this could become a new statistical product. Electricity data could also be used to improve the statistics on the distribution of the population.

An important conclusion is also that the quality of registers requires improvement so that it would be possible to switch to statistics produced on the basis of big data and to automate processes. All the developers of information systems could use standard addresses along with address identifiers, because in addition to being precise, this makes the updating of addresses much easier. Addresses do not need to be entered into the system all over again but can be converted to the new format with the help of automatic rules. The more the standard solutions are used, the more efficiently also statistics can be produced. (Map 1, p. 45)

Summary

The introduction of big data means that somewhere a large amount of information about an object / subject has accumulated. Statisticians are not interested in the doings of John or Mary individually but rather in their behaviour in combination with a similar aggregate – this creates the real dynamics of data. From the interaction between individual data components, e.g., a growth in the value added generated in this country is revealed. It is like every needle brought in by ants becomes a phenomenon called an anthill. Without these particular needles, the anthill would not be what it is, but when we are looking at the anthill as a whole, we do not need to know exactly

which are brought in which needle. That is why we should not be afraid of sharing our pieces of data, because only then can a true and complete picture be created. Big data provide this opportunity exactly – to get the best complete picture by using as much detailed source data as possible.

In order to assess the potential uses of big data in the production of statistics, European national statistics institutes have set up the ESSnet Big Data project. The goal of the project is to develop, on the basis of specific examples, a methodology and best practices for the introduction of big data into statistics production. Preliminary results refer to the great potential of the use of big data, and to the possibility of reducing the administrative burden of the society, but more work needs to be done for the success of the implementation, so that a clear and transparent methodology could be created for the production of statistics from big data.

The use of big data may also involve certain risks and the data may be biased, incomplete or reflect a phenomenon that statistics are not produced for. This was the experience gained from the pilot project involving electricity data, where the dataset is based on the agreement owner but the statistical outcome should be based on the end consumer. It can be concluded from the results of the pilot project that even if we have a dataset describing the whole phenomenon, we may not be able to immediately produce traditional statistics. At the same time, such datasets provide alternative opportunities for providing information to the society. The development of new statistics is also the theme of the follow-up project on big data.

The use of big data also enables to produce completely new statistics, as was proved by the prototype created at the hackathon organised by Eurostat (Õpik et al. 2017), which visualises jobs and related skills using the big data of job portals. Such tools provide the society with better feedback and offer the opportunity to look to the future. In today's data-based world, it could be considered whether data could be seen as a public good (Mitchell, Brynjolfsson 2017), and enterprises, as a result of whose activities a large amount of data accumulates, could mirror to the society what they have collected from it. The introduction of big data enables to publish completely new information; it enables to considerably reduce the administrative burden, phenomena can be monitored real-time and the society receives a lot faster the information needed for decision-making. If big data are used with methodological prudence, the production of statistics can be cheaper, quicker and more accurate, but for this, the respondents need to be ready to share their data for the production of statistics.

KAUPADE JA TEENUSTE VÄLISKAUBANDUS ETTEVÕTLUSNÄITAJATE JÄRGI

Allan Aron

Artikkel vaatleb Eesti väliskaubandust valitud ettevõtlusnäitajate (hõivatute arv, tegevusala, omandivorm) järgi. Selline lähenemine võimaldab saada lisainfot rahvusvahelisel turul tegutsevate ettevõtete kohta. Analüüs hõlmab nii kaupade kui ka teenuste väliskaubandust ning aitab välja tuua nii ühisosa kui ka erinevusi kaupade ja teenuste eksportijate vahel.

Ettevõtlusnäitajate tähtsus väliskaubandusstatistikas

Väliskaubandusstatistika vaatleb traditsiooniliselt kaubavoogusid eri riikide vahel. See võimaldab jälgida, milliseid kaupu eksporditakse ja imporditakse ning millised on tähtsamad partnerriigid. Seejuures on põhitähelepanu kaupadel, mida on võimalik tuvastada väga detailsel tasemel ja millel on peale kauba väärtuse olemas ka lisainfo eksporditud ja imporditud kauba koguse ja kaalu kohta. Samal ajal pakub väliskaubandusstatistika üsna vähe infot nende ettevõtete kohta, kes on väliskaubandustehingute taga. Ettevõtted ei ole identifitseeritavad andmekaitse ja konfidentsiaalsusnõuete tõttu. Ei ole teada, kas mingit kaupa ekspordivad või importivad ettevõtted on suured või väikesed, milline on nende tegevusala või kas nad on kohalikud ettevõtted või välisomanduses. Väliskaubandusstatistika andmete sidumine ettevõtlusnäitajatega aitab leida nendele küsimustele vastuseid. See võimaldab välja tuua uusi statistilisi näitajaid ning annab väliskaubandusstatistikale lisamõõtme. Euroopa Liidu statistikaamet Eurostat on seetõttu algatanud projektid „Kaupade väliskaubandus ettevõtlusnäitajate järgi“ (*Trade of Goods by Enterprise Characteristics* ehk TEC) ja „Teenuste väliskaubandus ettevõtlusnäitajate järgi“ (*Services Trade by Enterprise Characteristics* ehk STEC). Nende projektidega kogutakse liikmesriikidelt lisaandmeid ja Eurostat avaldab tulemused. Väga oluline on siinjuures põhimõte, et lisanäitajate arvutamisel kasutatakse olemasolevaid andmeid ega suurendata andmeesitajate halduskoormust.

Kaupade ja teenuste ühendamine väliskaubandusstatistikas

Traditsiooniliselt vaadeldakse väliskaubandusstatistikas kaupade ja teenuste väliskaubandust eraldi. Andmeid kaupade ja teenuste kohta kogutakse eri vormis ja eri küsimustikega, sageli on isegi andmeid koguvad ja avaldavad asutused erinevad. Samuti põhineb kaupade ja teenuste väliskaubandusstatistika erinevatel klassifikaatoritel, meetodikal ja õigusaktidel. Kõik see toob kaasa olukorra, kus kaupade ja teenuste statistikat on raske võrrelda ning statistika kasutajal on keeruline saada olukorrast head ülevaadet. Sageli on vaja isegi täpsustada, kas ekspordist rääkides mõeldakse ainult kaupade ekspordi või kaupu ja teenuseid koos. Andmete kvaliteedi ja võrreldavuse täielikuks mõistmiseks tuleb väga põhjalikult tunda mõlema valdkonna meetodikat. Ühise klassifikaatori puudumine tähendab, et analüüsiks vajalikud algandmed tuleb leida eri allikatest.

Samal ajal on rahvusvaheline kaubandus kiiresti globaliseeruv ja nii kaupade kui ka teenuste olemus on muutumas. Kaubad on kõik füüsilised tooted, mida eksporditakse või imporditakse üle Eesti riigipiiri (v.a transiit ja tolliladudevaheline kaubandus). Teenus on immateriaalne kaup ning teenuste eksport ja import põhineb teenuste ostu-müügitehingutel Eesti residentide ja mitte-residentide vahel. Kaupu eksporditakse sageli koos kaasnevate teenustega ja mõnikord on isegi keeruline eristada, kas tegemist on kauba või teenusega. Näiteks kui tarkvara müüakse interneti

kaudu, käsitatakse seda teenusena, kuid kui täpselt sama tarkvara müüakse füüsilisel andmekandjal (DVD vms), on tegemist kaubaga. Sellist eristamist võib olla teinekord raske mõista. Internetikaubanduse areng on olnud väga kiire ja järjest enam on kaubad kaotamas oma füüsilist vormi ja muutumas virtuaalseks (peale tarkvara ka filmid, muusika, raamatud jne). Statistika-süsteem peab tehnoloogia arenguga kaasas käima ja suutma pakkuda kaupade ja teenuste väliskaubandusest ühtset ülevaadet. Järgnevalt on püütudki seda saavutada, ühendades omavahel nii kaupade kui ka teenuste eksportijate ja importijate ettevõtlusnäitajad.

Mõisted ja meetodika

Analüüsis kasutatud algandmed pärinevad TEC-i ja STEC-i projekti jaoks Eurostatile koostatud andmetest (Compilers Guide 2016, 2017). Uuringu periood hõlmab aastaid 2013–2014, sest STEC-i aegrida algab 2013. aastaga. Et aegrida piirdub praegu kahe aastaga, ei ole kahjuks võimalik välja tuua pikemaid trende ja põhitähelepanu on aastal 2014.

Kaupade ja teenuste eksportijad on omavahel ühendatud ettevõtte registrikoodi alusel. Algandmetes on olemas ettevõtete kaupade ja teenuste aastakäive. Ettevõtlusnäitajate info pärineb Statistikaametis kasutusel olevast statistilisest ettevõtete registrist, mille aluseks on äriregister ja Statistikaameti kogutud lisaandmed.

Andmed kaupade väliskaubanduse kohta põhinevad tollideklaratsioonidel, Intrastati (EL-i riikide omavahelise kaubavahetuse andmete kogumise süsteem) küsimustikuga kogutud andmetel ja käibemaksudeklaratsioonidel. Käibemaksudeklaratsioonidel näidatud käivet kasutati juhul, kui ettevõtte kohta Intrastati andmeid ei olnud. Teenuste ekspordi ja impordi andmed on kogunud Eesti Pank.

Nii kaupade kui ka teenuste kaubavahetuse andmed sisaldavad kaupade töötlemise infot. Kaupade väliskaubanduses sisaldab töötlemise summa kauba algset ja lisatud materjali väärtust ning samuti osutatud töötlemise teenuse maksumust. Teenuste kaubavahetuse statistika hõlmab vaid teenuse maksumust. Töötlemise arvestamine nii kaupade kui ka teenuste juures tekitab olukorra, kus töötlemise väärtust arvestatakse topelt. Seetõttu on siinses analüüsis eemaldatud kaupade käibest kaupade töötlemise väärtus.

Ettevõtlusnäitajatena kasutatakse analüüsis suurusklassi, tegevusala ja omandivormi. Suurusklass põhineb ettevõtte hõivatute arvul ja selle järgi jagunevad ettevõtted mikro- (kuni 9 hõivatut), väike- (10–49 hõivatut), keskmisteks (50–249 hõivatut) ja suurettevõteteks (vähemalt 250 hõivatut). Tegevusala põhineb Eesti majanduse tegevusalade klassifikaatoril (EMTAK) ja näitab ettevõtte põhilist tegutsemisvaldkonda. Kuigi ettevõttel võib olla mitu tegevusala, lisatakse kogu tema väliskaubanduse käive põhitegevusala alla. Analüüsis kasutatakse tegevusala tähtkoodi. Osa sarnaseid tegevusalasid on grupeeritud ja väikese käibega tegevusalad on koondatud muude tegevusalade alla. Omandivormi järgi jagunevad Eesti ettevõtted kodumaisteks ja väliskapitali kontrolli all olevateks. Kui ettevõtte kapitaliosalusest üle 50% on välismaise ettevõtte omanduses, loetakse ettevõtte väliskapitali kontrolli all olevaks.

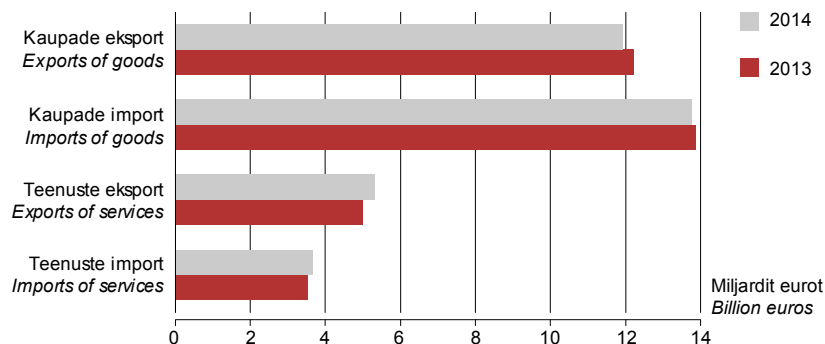
Nii kaupade kui ka teenuste väliskaubanduse andmed sisaldavad eksportijaid ja importijaid, kelle kohta ei ole ettevõtlusnäitajaid. Nende hulka kuuluvad näiteks eraisikud ja mitteresidentid äriühingud. Mitteresident on välismaa äriühing, kes ei ole registreeritud Eesti äriregistris, kuid on teinud Eestis väliskaubandustehinguid. Siinkohal on oluline, et mitteresidenti ei aetaks segamini välisomanduses oleva Eesti ettevõttega. Mitteresidentide ja eraisikute väliskaubanduse käive on hõlmatud kaupade ja teenuste kaubavahetuse üldises analüüsis, kuid detailses analüüsis ettevõtlusnäitajate järgi on nende andmed eemaldatud, sest nende ettevõtlusnäitajad (tegevusala, suurusklass, omandivorm) ei ole teada.

Kaupade ja teenuste väliskaubandus aastatel 2013–2014

Aastal 2014 vähenes kaupade eksport 2013. aastaga võrreldes 2% ja kaupade import 1% (joonis 1). Väliskaubandust mõjutasid 2014. aastal Ukraina sündmuste tõttu kehtestatud sanktsioonid Venemaa vastu ja Venemaa kehtestatud vastusanktsioonid, mis mõjutasid kaupade eksporti Eestist Venemaale (Puura, Silla 2016). Teenuste väliskaubandusele sanktsioonid olulist mõju ei avaldanud ning teenuste eksport kasvas 2014. aastal 7% ja import 4%.

Joonis 1. Kaupade ja teenuse eksport ja import, 2013–2014

Figure 1. Exports and imports of goods and services, 2013–2014



Väliskaubandus tegevusala järgi

Tegevusala järgi (tabel 1) oli kaupade ekspordis 2014. aastal suurima osatähtsusega töötleva tööstuse (60% kaupade ekspordi väärtusest) ning hulgi- ja jaekaubanduse tegevusala (25%). Kaupade impordis oli seevastu esikohal hulgi- ja jaekaubandus (52%), millele järgnes töötlev tööstus (32%). Kaubavahetuse ülejääk (ekspordi väärtus ületab impordi väärtust) oli suurim töötlevas tööstuses (2,7 miljardit eurot) ning puudujääk oli suurim hulgi- ja jaekaubanduses (–4,3 miljardit eurot).

Teenuste ekspordis ja impordis ei olnud ühegi tegevusala domineerimine nii suur. Suurima ekspordikäibega oli veonduse ja laonduse tegevusala, mis hõlmas 31% teenuste ekspordist. Info ja side tegevusala ettevõtete eksport oli teenuste ekspordist 10%. Teenuste impordis oli samuti suurima käibega veondus ja laondus (18%) ning info ja side (8%). Teenuste käibes on suur osatähtsus ka sellistel teenustel, mida ei ole võimalik ettevõtlusnäitajatega siduda. Valdavalt on siin tegemist turismi- ja finantsteenustega.

Mõne tegevusala ettevõtted ekspordivad peamiselt kaupu, samas kui teistel tegevusaladel on tähtsam teenuste eksport (joonis 2). Kaupade ekspordi osatähtsus on suurem töötleva tööstuse (C), põllumajanduse, metsanduse ja mäetööstuse (A–B), hulgi- ja jaekaubanduse (G) ning elektri-, gaasi- ja veevarustuse (D–E) tegevusalal. Teenuste eksport domineerib info ja side (J), finants- ja kindlustustegevuse (K), haldus- ja abitegevuste (N), kutse-, teadus- ja tehnikaalase tegevuse (M), ehituse (F) ning veonduse ja laonduse (H) tegevusalal.

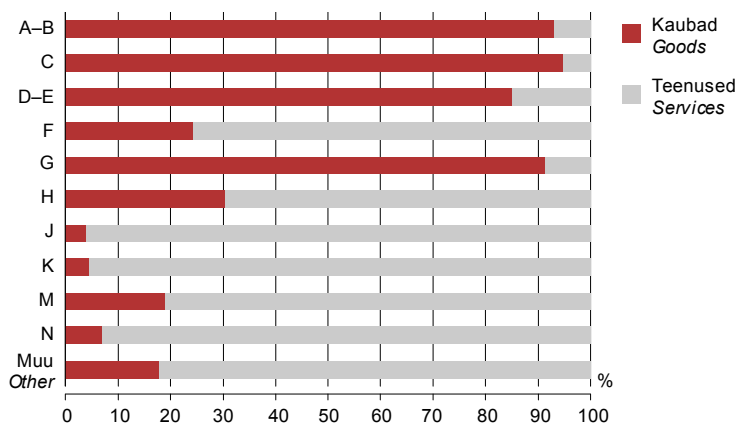
Tabel 1. Ekspordi- ja impordikäive tegevusala järgi, 2014

Table 1. Value of exports and imports by economic activity, 2014
(miljonit eurot – million euros)

Tegevusala Economic activity	Kaubad Goods		Teenused Services	
	eksport exports	import imports	eksport exports	import imports
A–B Põllumajandus, metsandus ja mäetööstus A–B Agriculture, forestry and mining	127	48	10	14
C Töötlev tööstus C Manufacturing	7 152	4 467	397	261
D–E Elektri-, gaasi- ja veevarustus D–E Electricity, gas and water supply	327	333	57	128
F Ehitus F Construction	68	152	212	96
G Hulgi- ja jaekaubandus G Wholesale and retail trade	2 944	7 195	280	243
H Veondus ja laondus H Transportation and storage services	710	660	1 636	656
J Info ja side J Information and communication	21	118	509	292
K Finants- ja kindlustustegevus K Financial and insurance activities	6	62	119	94
M Kutse-, teadus- ja tehnikaalane tegevus M Professional and scientific activities	67	97	288	127
N Haldus- ja abitegevused N Administrative services	25	82	342	258
Muud tegevusalad Other economic activities	36	153	169	131
Tegevusala teadmata Economic activity unknown	453	406	1 304	1 374
Kokku Total	11 936	13 774	5 323	3 673

Joonis 2. Kaupade ja teenuste eksport tegevusala järgi, 2014

Figure 2. Exports of goods and services by economic activity, 2014

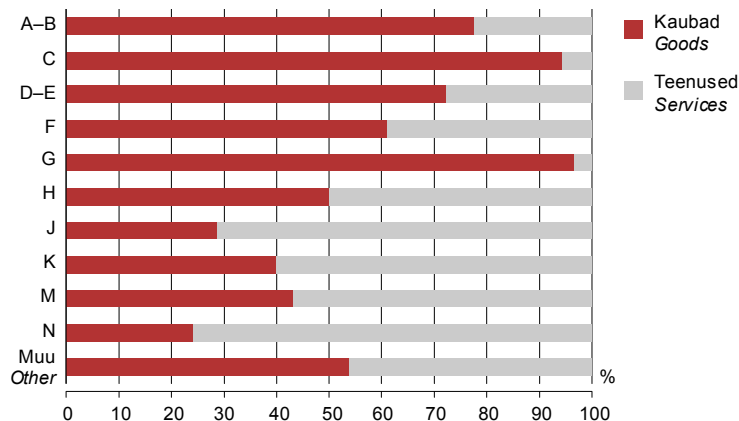


Kaupade ja teenuste impordis on kaupade osatähtsus samuti suurem hulgi- ja jaekaubanduse (G) ning töötleva tööstuse (C) tegevusalal (joonis 3). Kõige suurem erinevus võrreldes ekspordiga ilmneb ehituses (F). Kui ekspordis oli suurem osatähtsus teenustel, siis impordis on

suurem osatähtsus kaupadel. Eesti ehitusettevõtted tegutsevad sageli ka välisurgudel ja ekspordivad ehitusteenuseid. Samal ajal imporditakse ehitusmaterjale ja seadmeid siseturu vajadusest lähtuvalt. Kaupade impordi osatähtsus on suurem ka mitmel teisel tegevusalal (finants- ja kindlustusteenused (K), info ja side (J)).

Joonis 3. Kaupade ja teenuste import tegevusala järgi, 2014

Figure 3. Imports of goods and services by economic activity, 2014

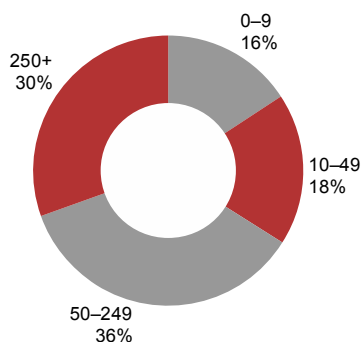


Väliskaubandus ettevõtte suurusklassi järgi

Ettevõtte suurusklass võimaldab uurida, kas väliskaubanduses domineerivad suurettevõtted või väikese ja keskmise suurusega ettevõtted. Suurriikides on suurettevõtete osatähtsus rahvusvahelises kaubanduses tavaliselt suurem ja väikeriikides nagu Eesti on see mõnevõrra tagasihoidlikum (Statistical Insights 2016). 2014. aastal oli Eestis suurettevõtete osatähtsus kogu kaupade ekspordi väärtuses 30% (joonis 4). Euroopa Liidus keskmiselt oli suurettevõtete osatähtsus 2010. aastal 53% EL-i siseses ekspordis (International trade 2012). Eestis olid kaupade ekspordis 2014. aastal suurima osatähtsusega keskmised ettevõtted, kes andsid ekspordikäibest 36%.

Joonis 4. Kaupade eksport ettevõtte suurusklassi järgi, 2014

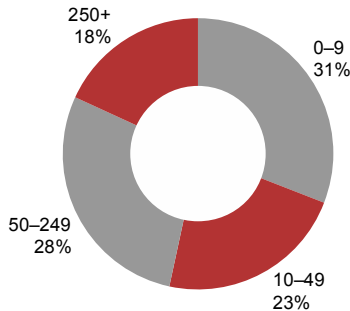
Figure 4. Exports of goods by enterprise size class, 2014



Teenuste ekspordis ettevõtte suurusklassi järgi avaneb mõneti teistsugune pilt. Suurettevõtete osatähtsus oli siin veelgi väiksem, vaid 18% teenuste ekspordist (joonis 5). Suurima osatähtsusega olid teenuste ekspordis hoopis mikroettevõtted (0–9 hõivatut), kes andsid 31% teenuste ekspordist. Siit võib teha järelduse, et kaupade eksport nõuab märksa enam tööjõudu kui teenuste eksport ning ka väga väikese hõivatute arvuga teenuste eksportijad võivad olla rahvusvahelisel turul edukad.

Joonis 5. Teenuste eksport ettevõtte suurusklassi järgi, 2014

Figure 5. Exports of services by enterprise size class, 2014

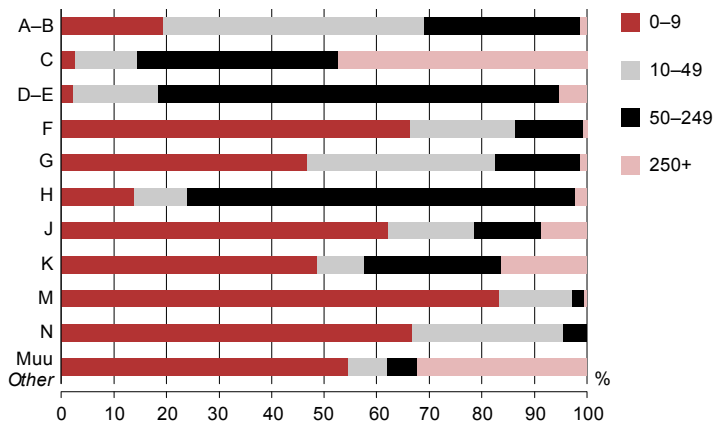


Erinevused kaupade ja teenuste importijate vahel on väiksemad. Suurettevõtete osatähtsus oli 24% nii kaupade kui ka teenuste impordi käibes. Mikro- ja väikeettevõtete osatähtsus oli 45% kaupade impordis ja 50% teenuste impordis. Mikro- ja väikeettevõtte suurem osatähtsus impordis näitab, et hõivatute arv ei ole impordi kaubavoos nii oluline näitaja. Paljud importijad tegutsevad ka hulgi- ja jaekaubanduse tegevusalal ning on ennekõike kaupade maaletoojad ja vahendajad. Need tegevused ei ole väga tööjõumahukad.

Kaupade ekspordis ettevõtte suurusklassi ja tegevusala järgi oli suurettevõtete osatähtsus kõige suurem (47%) töötlevas tööstuses (C). Kõigil teistel tegevusaladel oli suurettevõtete osatähtsus tagasihoidlik (joonis 6). Keskmise suurusega ettevõtted domineerisid elektri-, gaasi- ja veevarustuse (D–E) ning veonduse ja laonduse (H) tegevusalal.

Joonis 6. Kaupade eksport ettevõtte suurusklassi ja tegevusala järgi, 2014

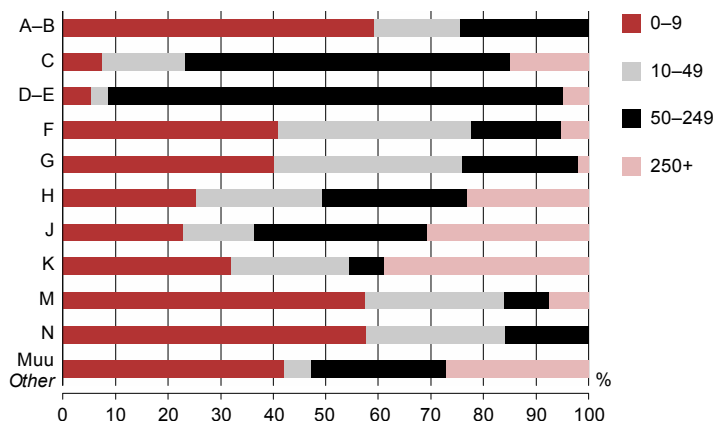
Figure 6. Exports of goods by enterprise size class and economic activity, 2014



Teenuste ekspordis ettevõtte suurusklassi ja tegevusala järgi oli suurettevõtete osatähtsus suurem finants- ja kindlustustegevuse (K) ning info ja side (J) tegevusalal (joonis 7). Keskmise suurusega ettevõtetele oli suurem osatähtsus elektri-, gaasi- ja veevarustuse (D–E) ning töötleva tööstuse (C) tegevusalal.

Joonis 7. Teenuste eksport ettevõtte suurusklassi ja tegevusala järgi, 2014

Figure 7. Exports of services by enterprise size class and economic activity, 2014



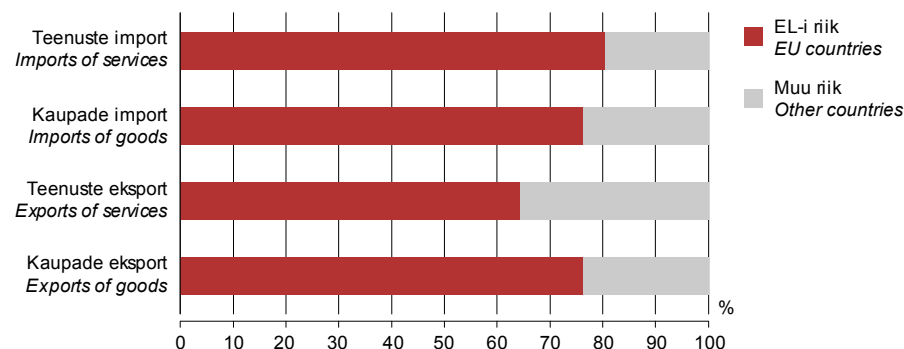
Euroopa Liidu osatähtsus väliskaubanduses

Eesti väliskaubandus toimub peamiselt Euroopa Liidu liikmesriikidega. 2014. aastal hõlmas eksport EL-i riikidesse 76% kaupade ja 64% teenuste ekspordist (joonis 8). Kaupade ekspordi peamised sihtriigid on sageli lähemad naaberriigid, sest kaugematesse riikidesse eksportides kasvab nii transpordikulu kui ka transpordile kuluv aeg. Teenuste ekspordi puhul on nende piirangute mõju väiksem ja seetõttu on teenuseid lihtsam pakkuda ka globaalsel turul. Sellega on seletatav ka EL-i väliste riikide suurem osatähtsus teenuste ekspordis.

Suurettevõtted on rohkem keskendunud EL-i turule. 2014. aastal ekspordisid nad 84% kaupadest ja 74% teenustest EL-i riikidesse. Väiksemate ettevõtete puhul oli ekspordil EL-i riikidesse väiksem tähtsus. Näiteks mikroettevõtete puhul oli EL-i riikide osatähtsus 75% kaupade ja 64% teenuste ekspordis.

Joonis 8. Kaupade ja teenuste eksport ja import partnerriigi järgi, 2014

Figure 8. Exports and imports of goods and services by partner countries, 2014

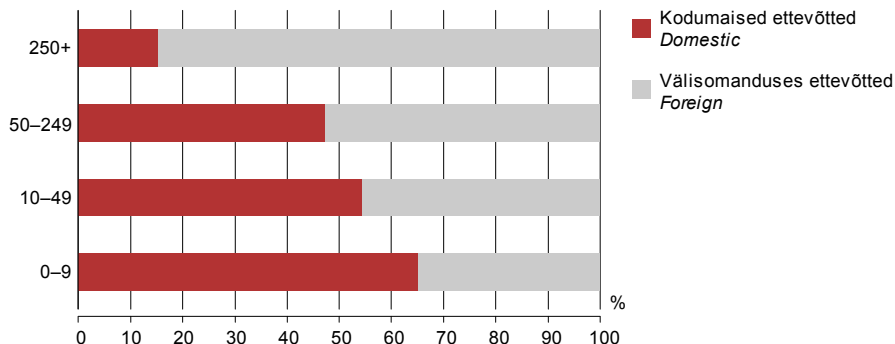


Väliskaubandus ettevõtte omandivormi järgi

Omandivorm näitab, kas ettevõtte on kodumaises või välisomanduses. 2014. aastal oli kodumaises omanduses 93% ning välisomanduses 7% ettevõtetest.^a Rahvusvahelisel turul tegutsevate ettevõtete seas on välisomanduses ettevõtete osatähtsus kindlasti suurem. Paljudel välisfirmadel on Eestis tütarettevõtted või osalus mõnes Eesti ettevõttes. Nii hõlmas 2014. aastal kodumaises omanduses firmade kaupade eksport vaid 42% kogueksporti väärtusest. Kaupade ekspordis on märgatav ka selge seos ettevõtte suurusklassi ja omandivormi vahel (joonis 9). Väiksemate ettevõtete seas annavad suurema osa ekspordikäibest kodumaises omanduses ettevõtted, samas kui suurettevõtete seas annavad valdava osa ekspordist välisomanduses ettevõtted.

Joonis 9. Kaupade eksport ettevõtte omandivormi ja suurusklassi järgi, 2014

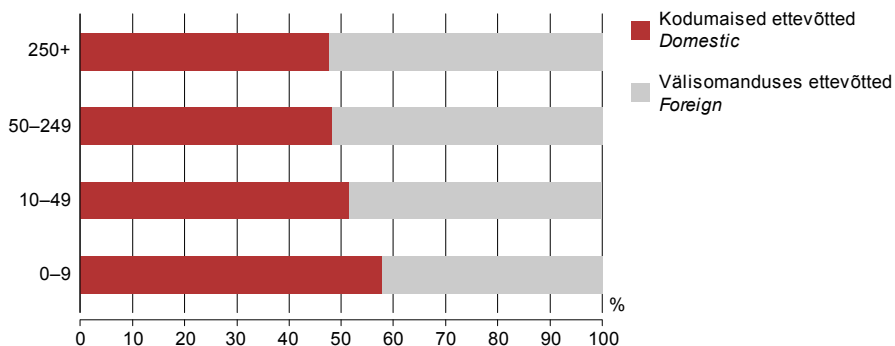
Figure 9. Exports of goods by type of ownership and enterprise size class, 2014



Teenuste ekspordis on välisomanduses ettevõtete osatähtsus mõnevõrra väiksem. 52% teenuste ekspordist oli kodumaises omanduses ettevõtete käes. Välismaises omanduses ettevõtete osatähtsus teenuste ekspordis oli kõigis suurusklassides väiksem kui kaupade ekspordis (joonis 10). Sarnaselt kaupade ekspordiga on ka teenuste ekspordis märgatav trend, et suurettevõtete seas on välismaises omanduses ettevõtete ekspordi osatähtsus suurem kui väiksemate ettevõtete seas. Suurettevõtete teenuste ekspordis oli välisomanduses ettevõtete osatähtsus 52%, samas kui kaupade ekspordis oli välisomanduses suurettevõtte osatähtsus koguni 85%.

Joonis 10. Teenuste eksport ettevõtte omandivormi ja suurusklassi järgi, 2014

Figure 10. Exports of services by type of ownership and enterprise size class, 2014



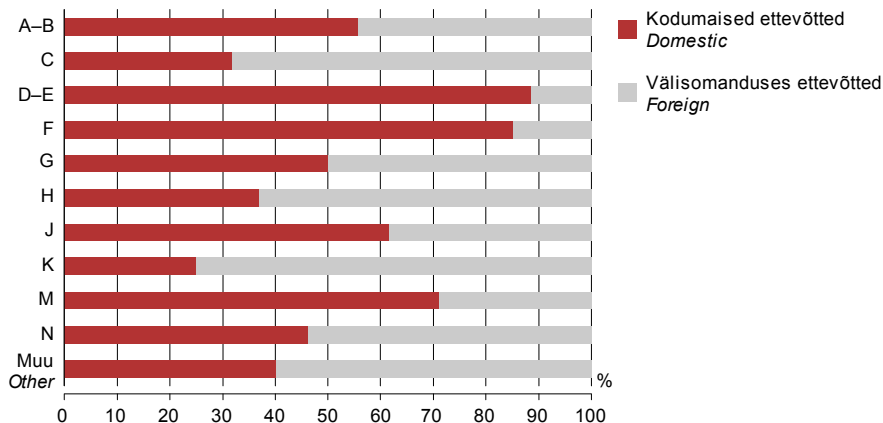
^a Statistika andmebaas, tabel ER026.

Kaupade ja teenuste import ettevõtte omandivormi ja suurusklassi järgi oli sarnane ekspordiga.

Kaupade ekspordis oli välisomanduses ettevõtete ekspordi osatähtsus suurim finants- ja kindlustustegevuse (K), töötleva tööstuse (C) ning veonduse ja laonduse (H) tegevusalal (joonis 11). Kodumaises omanduses ettevõtete ekspordi osatähtsus oli seevastu suurem elektri-, gaasi- ja veevarustuse (D–E) ning ehituse (F) tegevusalal.

Joonis 11. Kaupade eksport ettevõtte omandivormi ja tegevusala järgi, 2014

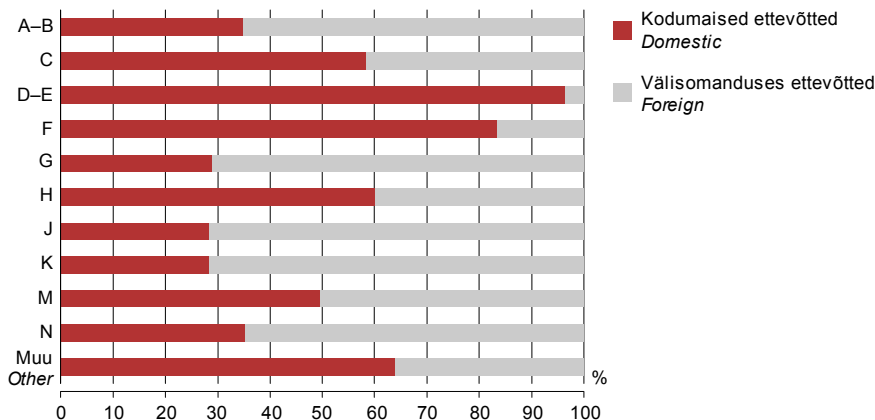
Figure 11. Exports of goods by type of ownership and economic activity, 2014



Teenuste ekspordis oli välisomanduses ettevõtete ekspordi osatähtsus suurim finants- ja kindlustustegevuse (K), info ja side (J) ning hulgi ja jaekaubanduse (G) tegevusalal (joonis 12). Kodumaises omanduses ettevõtete teenuste ekspordi osatähtsus oli sarnaselt kaupade ekspordiga suurim elektri-, gaasi- ja veevarustuse (D–E) ja ehituse (F) tegevusalal.

Joonis 12. Teenuste eksport ettevõtte omandivormi ja tegevusala järgi, 2014

Figure 12. Exports of services by type of ownership and economic activity, 2014



Kokkuvõte

Kaupade ja teenuste väliskaubanduse ühendamine ja analüüsimine ettevõtlusnäitajate alusel võimaldab välja tuua uusi statistilisi näitajaid ning annab lisainfot eksportijate ja importijate olemuse kohta. Kaupade eksportijad ja importijad tegutsevad peamiselt töötleva tööstuse ning hulgi- ja jaekaubanduse valdkonnas. Teenuste eksport on olulisem veonduse ja laonduse ning info ja side tegevusalal. Kaupade ekspordis on suurema osatähtsusega suured ja keskmise suurusega ettevõtted, seevastu teenuste ekspordis on tähtsam roll mikro- ja väikeettevõtetel. Välisomanduses ettevõtetel on suurem roll kaupade ekspordis ja kodumaises omanduses ettevõtetel teenuste ekspordis. Nii kaupade kui ka teenuste ekspordis on välisomanduses ettevõtete osatähtsus suurem suurettevõtete seas.

Allikad Sources

Compilers guide on European statistics on international trade in goods by enterprise characteristics (TEC). Eurostat 2016 (23.11.17).

Compilers Guide for statistics on Services Trade by Enterprise Characteristics (STEC). Eurostat-OECD. 2017 (23.11.17).

International trade by enterprise characteristics. Eurostat. 2012 (1.11.17).

Puura, E., Silla, M. (2016). Kaupade väljavedu Venemaale. Eesti Statistika Kvartalikirj. 4/16. Quarterly Bulletin of Statistics Estonia, lk 25–43. Tallinn: Statistikaamet. [www] https://www.stat.ee/valjaanne-2016_eesti-statistika-kvartalikirj-4-16.

Statistical Insights: Who's Who in International Trade: A Spotlight on OECD Trade by Enterprise Characteristics data. 25.04.2016. (1.11.17).

FOREIGN TRADE OF GOODS AND SERVICES BY ENTERPRISE CHARACTERISTICS

Allan Aron

This article examines Estonian foreign trade by chosen enterprise characteristics (number of employees, economic activity, type of ownership). This approach enables obtaining additional information about traders acting in the international market. The analysis includes both the foreign trade of goods and services and allows identifying the differences and common traits of exporters of goods and services.

The importance of enterprise characteristics in foreign trade statistics

Traditionally, foreign trade statistics observes trade flows between countries. It allows analysing goods exported and imported and identifying the main partner countries. The main focus here is on the goods – it is possible to identify them on a very detailed level, and in addition to the value of goods, there is also information about the net mass and quantity of goods exported and imported. However, foreign trade statistics provide quite little information about the companies involved in foreign trade transactions. Identifying the enterprises is not possible due to data protection and confidentiality reasons. It is not known whether the enterprises exporting or importing certain goods are small or large, what their economic activity is or whether they are domestically or foreign owned. Linking trade data with enterprise characteristics helps to address these questions. It provides new statistical indicators and an additional dimension to foreign trade statistics. For this reason, the European Union statistical office Eurostat has launched initiatives called Trade of goods by Enterprise Characteristics (TEC) and Services Trade by Enterprise Characteristics (STEC). With these projects, additional data are collected from the Member States and disseminated by Eurostat. It is very important that the new indicators are calculated on the basis of existing data and that the administrative burden of data providers does not increase.

Linking trade of goods and trade of services in foreign trade statistics

The traditional approach in foreign trade statistics is that the trade of goods and trade of services are studied separately. The data for goods and services are collected by different surveys and in different formats, often even the national institutions responsible for data collection and dissemination are different. The statistical domains have different classifications, methodology and legal background as well. This means that the trade of goods and trade of services statistics are difficult to compare, and the situation is confusing for the users of statistics who are trying to get a complete overview of the situation. Often, it is even necessary to clarify whether total exports cover both goods and services or only goods. Full understanding of data quality and comparability requires in-depth knowledge of the methodology of both statistical domains. The lack of a common classification for goods and services means that separate data sources must be used for analysis.

At the same time, international trade is globalizing and the very nature of trade and services is changing. Goods are all physical products, which are exported or imported across the Estonian border (excl. goods in transit and trade between customs warehouses). Services are intangible goods, and the exports and imports of services are based on purchase and sale transactions between Estonian residents and non-residents. Goods are often exported with accompanying services, and sometimes it is even difficult to draw a line between goods and services. For example, if software is sold via the internet, the transaction is considered as trade of services, but if the same software is sold in a physical format (DVD, etc.), it is considered as trade of

goods. Such distinction could sometimes be difficult to understand. The development of internet commerce has been rapid and goods are losing their physical existence and becoming virtual (in addition to software, the same applies to movies, music, books, etc.). Technological progress means that the statistical system must also develop and respond by providing a combined overview of the foreign trade of goods and services. The current analysis attempts to address these issues by linking importers and exporters of goods and services by their enterprise characteristics.

Concepts and methodology

The source data for the analysis are based on the datasets of TEC and STEC projects compiled for Eurostat (Compilers Guide 2016, 2017). The survey period covers years 2013–2014, as the STEC data series starts with the year 2013. As the time series covers only two years at the moment, it is not possible to identify long-term trends and the main focus is on year 2014.

The traders of goods and services are linked by their business registration number. The source data includes the annual value of enterprises' goods and services. The enterprise characteristics are extracted from the statistical business register, which is based on the national business register and additional data collected by Statistics Estonia.

The trade of goods data are based on customs declarations, the Intrastat survey (data collection system for intra-EU trade) and value added tax (VAT) declarations. The VAT data were used when the Intrastat data were missing. The trade of services data were collected by Eesti Pank.

Both the trade of goods and trade of services statistics include data of processing. In trade of goods statistics, the value of processing includes the cost of original and added materials as well as the cost of processing service. The trade of services data only records the cost of processing service. If processing was included both for goods and services, it would create a situation where processing value is counted twice. Therefore, the value of processing of goods was excluded from the trade of goods dataset.

Enterprise characteristics used in the analysis include size class, economic activity and type of ownership. Size class is based on the number of employees and the traders were divided into the following categories: micro (0–9 employees), small (10–49), medium-sized (50–249) and large (250+) enterprises. The economic activity is determined according to the Estonian Classification of Economic Activities (EMTAK, based on NACE) and it shows the main area of operation of the enterprise. Although an enterprise can be engaged in many economic activities, all of its foreign trade turnover is added under its main economic activity. The EMTAK letter-code is used in the analysis. Some similar activities are grouped together and activities with low value are grouped under other economic activities. By type of ownership, Estonian enterprises are divided between domestic and foreign-controlled enterprises. An enterprise is considered foreign-controlled when more than 50% of its fixed capital is owned by a foreign company.

The foreign trade data of both goods and services include exporters and importers for which enterprise characteristics are missing. These include, for example, individuals and non-resident companies. A non-resident enterprise is a foreign company which is not registered in the Estonian business register but which has foreign trade transactions in Estonia. Here, it is important to make sure that a non-resident is not confused with a foreign-owned Estonian enterprise. The foreign trade value of non-residents and individuals is included in the general analysis of trade in goods and services, but in the detailed analysis by enterprise characteristics, their data are excluded, as their enterprise characteristics (economic activity, size class, type of ownership) are not known.

Trade of goods and services in 2013–2014

In 2014, the exports of goods decreased by 2% and imports by 1% compared to 2013 (Figure 1, p. 61). In 2014, foreign trade was influenced by the sanctions imposed on Russia in relation to events in Ukraine and counter-sanctions imposed by Russia, which had an effect on the exports of goods from Estonia to Russia (Puura, Silla 2016). The sanctions did not have a significant impact on the foreign trade of services. In 2014, the exports of services increased by 7% and imports by 4%.

Foreign trade by economic activity

In exports of goods by economic activity (Table 1, p. 62), manufacturing (60% of the total value of exports of goods) and wholesale and retail trade (25%) had the largest shares in 2014. However, in imports of goods, wholesale and retail trade had the largest share (52%), followed by manufacturing (32%). The highest trade surplus (value of exports exceeds the value of imports) occurred in manufacturing (2.7 billion euros) and the highest deficit occurred in wholesale and retail trade (–4.3 billion euros).

In the exports and imports of services, no single economic activity held such a big share. Transportation and storage services had the highest export value, accounting for 31% of the total value of exports of services. Information and communication held the second place with 10% share in the exports of services. In the imports of services, also transportation and storage services (18%) and information and communication (8%) had the biggest shares. In the total trade value of services, services which cannot be connected to enterprise characteristics have a significant share as well. These are mainly tourism and financial services.

Some sectors are mainly exporting goods while others are mainly involved in exports of services (Figure 2, p. 62). The share of exports of goods is larger in the economic activities of manufacturing (C), agriculture, forestry and mining (A–B), wholesale and retail trade (G) and electricity, gas and water supply (D–E). Exports of services dominate in information and communication (J), financial and insurance activities (K), administrative services (N), professional and scientific activities (M), construction (F) and transportation and storage services (H).

In the imports of goods and services, the share of goods was higher also in wholesale and retail trade (G) and manufacturing (C) (Figure 3, p. 63). The biggest difference compared to exports was in the construction sector (F). While the exports of services dominated in exports, goods had a larger share in imports. Estonian construction enterprises are often operating on foreign markets and exporting construction services. At the same time, raw materials and construction equipment are imported based on domestic demand. The share of imports of goods is larger also in several other economic activities (financial and insurance activities (K), information and communication (J)).

Foreign trade by enterprise size class

Enterprise size class helps to identify whether foreign trade is dominated by large or small and medium-sized enterprises. In big countries, the share of large enterprises in international trade is usually larger and in small countries such as Estonia it is somewhat more modest (Statistical Insights 2016). In 2014, the share of large enterprises in Estonia was only 30% in the total value of exports of goods (Figure 4, p. 63). In comparison, large companies contributed 53% of the total intra-EU exports of the European Union as a whole in 2010 (International trade 2012). In the Estonian exports of goods, medium-sized enterprises had the largest share in 2014, accounting for 36% of the total export value.

The situation is somewhat different in the exports of services by size class. Here, the share of large enterprises was even lower – only 18% of the total export value (Figure 5, p. 64). Micro enterprises (0–9 employees) had the largest share in the exports of services, contributing 31% of

export value. Micro and small enterprises together contributed over half of the total value of exports of services. It can be concluded that exporting goods requires more labour force compared to exporting services, and the exporters of services with a very small number of employees can also be successful in the international market.

The differences between importers of goods and services are smaller. Large companies are responsible for 24% of the imports value of both goods and services. Micro enterprises and small companies contribute 45% of the imports of goods and 50% of the imports of services. The larger share of micro enterprises and small enterprises in imports shows that the number of employees is not such a significant factor in imports trade flow. Many importers operate also in wholesale and retail trade and are primarily mediators of goods. These activities are not very labour-intensive.

In the exports of goods according to size class and economic activity, the share of large enterprises was greatest (47%) in manufacturing (C). In all other economic activities, their share was limited (Figure 6, p. 64). Medium-sized enterprises dominated in electricity, gas and water supply (D–E) and transportation and storage services (H).

In the exports of services by size class and economic activity, the share of large enterprises was higher in financial and insurance activities (K) and information and communication activities (J) (Figure 7, p. 65). Medium-sized enterprises held the largest share in electricity, gas and water supply (D–E) and manufacturing (C) sectors.

The share of the European Union in foreign trade

Estonian foreign trade is mostly with the European Union Member States. In 2014, intra-EU trade accounted for 76% of the exports of goods and 64% of the exports of services (Figure 8, p. 65). The main destination countries for the exports of goods are often the neighbouring countries, because transportation costs and logistical reasons make it more time-consuming and expensive to export to more distant countries. In the exports of services, the effect of such constraints is smaller, and therefore, it is easier to export services to the global market. This explains the bigger share of extra-EU trade in the exports of services.

Large companies are more focused on the EU market. In 2014, they exported 84% of their goods and 74% of their services to the EU countries. In the case of smaller enterprises, exports to the EU countries were less important. For example, micro companies exported 75% of their goods and 64% of their services to the EU countries.

Foreign trade by type of ownership

Type of ownership provides information on whether the enterprise is domestically or foreign-owned. In 2014, 93% of enterprises were domestically owned and 7% were foreign-owned.^a Among enterprises engaged in international trade, the share of foreign-controlled enterprises is definitely larger. Many foreign companies have affiliates in Estonia or a share in an Estonian enterprise. In 2014, domestically owned enterprises contributed only 42% of the total export value. In the exports of goods, there is also clear correlation between the type of ownership and enterprise size class (Figure 9, p. 66). Among smaller enterprises, the major share of export value is contributed by domestically controlled enterprises, while among large enterprises, foreign controlled enterprises account for the largest share.

In the exports of services, the share of foreign-controlled enterprises is somewhat lower. 52% of the total value of services is exported by domestic enterprises. In the exports of services, the share of foreign-controlled enterprises was smaller in all size classes compared to the exports

^a Statistical database, table RE026.

of goods (Figure 10, p. 66). But similarly to the exports of goods, there is a trend also in the exports of services that in the case of large enterprises the share of exports of foreign-controlled enterprises is higher than in the case of smaller enterprises. In the exports of services by large enterprises, 52% of the value was exported by foreign-controlled large enterprises. In the exports of goods, the share of foreign-controlled large enterprises was as much as 85%.

The type of ownership according to the size class in imports of goods and services was similar to the respective breakdowns in exports trade flow.

In the exports of goods, the share of foreign-controlled enterprises was highest in finance and insurance activities (K), manufacturing (C) and transportation and storage services (H) (Figure 11, p. 67). On the other hand, the share of domestically controlled enterprises was highest in electricity, gas and water supply (D–E) and construction (F).

In the exports of services, the share of foreign-controlled enterprises was highest in finance and insurance activities (K), information and communication (J) and wholesale and retail trade (G) (Figure 12, p. 67). Similarly to the exports of goods, in the exports of services, the share of domestically controlled enterprises was largest in electricity, gas and water supply (D–E) and construction (F) sectors.

Conclusion

The linking and analysis of trade of goods and services by enterprise characteristics provides new statistical indicators and additional information about the nature of traders of goods and services. The exporters and importers of goods are mainly operating in manufacturing, and wholesale and retail trade sectors, but the exporters of services are concentrated in transportation, and information and communication sector. The exports of goods are dominated by medium-sized and large enterprises, but in the exports of services, micro and small enterprises are playing a more important part. The exporters of goods are mainly foreign-controlled companies, but in the exports of services, domestically controlled enterprises are playing a more important part. Both in the exports of goods as well as services, the share of foreign-controlled enterprises is highest among large enterprises.

MAA HIND: KUI PALJU ON EESTIMAA VÄÄRT?

Reelika Parve, Robert Mürsepp

Iga majandusliku tegevuse jaoks on tarvis maad: kas rohkem või vähem, kas otseselt tootmisprotsessis kasutamiseks või tootmisüksuse asukohaks oleva maatükina. Artikkel annab ülevaate, kui palju on maa Eestis väärt ja kui palju tuleb eri otstarbega maa soetamisel selle eest maksta.

Sissejuhatus

Eelmainitu taustal on oluline teada maa väärtust. Maatüki väärtuse kindlaks tegemisel (olenemata maa liigist ja kasutaja vajadustest) tuleb silmas pidada mitut tegurit. Eelkõige võib maa hinda mõjutavatest teguritest esile tuua asukohta ehk läheduse linnakeskustele või vastupidi kauguse linnapiirkondadest ja vaigse ümbruse; maalapi läheduses asuva praeguse või tulevase taristu, mullaviiljakuse jne. Kõigi maa hinda mõjutavate tegurite loetelu tuleks siin mainimiseks liiga pikk. Maa väärtus on küll seotud mitmesuguste näitajatega, kuid konkreetse maalapi väärtust mõjutavate tunnuste loetelu sõltub eelkõige selle otstarbest. Maa hinna teadasaamiseks võib paluda eksperthinnangut või üritada muid allikaid kasutades ise hinnangut anda, näiteks võrreldes tunnuste poolest sarnaste maatükkide väärtust. Loomulikult ei sobi selline kõiki eripärasid arvesse võttev hindamine kogu riigi maa väärtusele hinnangu andmiseks.

Siinses lühiartiklis esitatakse selle Eurostati projekti^a tulemusi, mille eesmärk oli hinnata maa väärtust Eestis viimase 20 aasta jooksul. Maa väärtuse meetodika oli vaja välja töötada Euroopa rahvamajanduse ja regionaalse arvepidamise süsteemi ESA 2010 (European System of Accounts 2010) raamistiku põhjal. Teoreetilist tuge pakkus projektile Eurostati ja OECD koostöös valminud arvestusjuhend. (Eurostat-OECD ... 2015)

Metoodika

Riigi bilansis on maa kirjas kui vara. Maa hindamise meetod (otsene või kaudne) sõltub sellest, millist tüüpi andmed on saadaval. Otsese meetodi puhul hinnatakse maa väärtust võimalikult detailselt maaliigi kaupa. Selleks tuleb:

- hinnata maa pindala liigi kaupa ühel või mitmel aastal;
- hinnata aastaseid muutuseid liikides, et luua aegrida;
- hinnata umbkaudsed hinnad eri liikide kohta ühel või mitmel aastal;
- modelleerida eri liikide hinnamuutused, et saada ühiku hinna aegridu;
- viia kokku pindala ja hinna andmed, et luua maa väärtuse aegrida;
- täpsustada aastased mahu- ja hinnamuutused.

Kaudsete meetoditega saab maa väärtust hinnata muude agregeeritumate näitajate kaudu või koos mõne teise näitajaga. Näiteks:

- hinnata maa väärtust koos sellel asuvate hoonete ja rajatistega;
- leida maa väärtus jääkliikmena mõnest laiema haardega arvestusest;
- hinnata maa väärtust maa ja rajatiste suhtarvude kaudu;
- hinnata maa väärtust regressioonimudeliga.

^a Action to support specific improvements in National Accounts. Objective 3: Improvements in availability, timeliness and quality of annual and quarterly national accounts, supply and use tables and regional accounts.

OECD (2015) soovitus järgi tuleks maa väärtust hinnata võimalikult detailselt ja võtta arvesse nii piirkondlikke kui ka maaliigist tulenevaid erinevusi. Eestis on selliste hinnangute andmise aluseks kõige parem allikas Maa-amet, kust võib saada andmeid nii maa suuruse kui ka hinna kohta. Osa lisateavet on pärit ka Keskkonnaagentuurist, kust saab detailsed aastaandmed metsamaa omanike, kasutamise, tunnuste ja turuhinna kohta.

Eesti riigi maa väärtuse hindamisel kasutati otsesest meetodit, sest on olemas vajaliku detailsusega andmestik. Otsese meetodi korral korrutatakse iga maatüki suurus selle liigi hinnaga. Seejärel summeeritakse kõigi maatükkide väärtused, et saada riigi territooriumil oleva maa koguväärtus. Isegi kui maa väärtuse otsene hindamine tundub teoreetiliselt lihtsam, on see praktikas siiski väga keeruline. Ideaalis peaks olema info iga maatüki hinna ja pindala kohta olemas, kuid sellise detailsusega andmete kättesaamine võib olla üsna raske.

Maa väärtust aastal t arvutatakse järgmise valemi järgi:

$$LV_t = \sum_{i=1}^n x_{i,t} * p_{i,t}$$

kus:

i = maaliik (põllumajandusmaa, metsamaa jne);

t = vaatlusaasta (1995–2015);

$x_{i,t}$ = pindala maa liigi i järgi aastal t ;

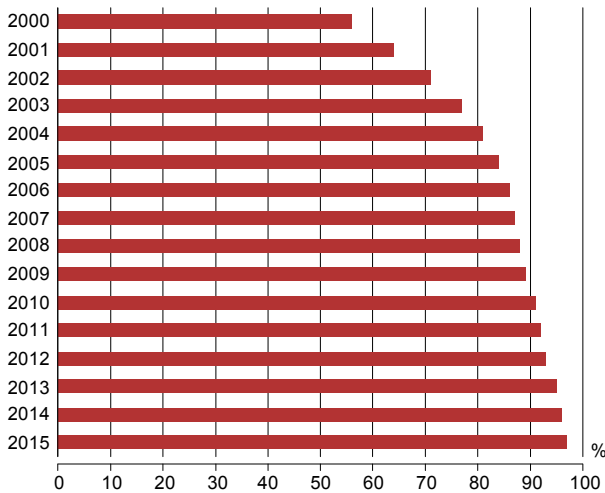
$p_{i,t}$ = maa liigi i hind aastal t .

Maaliikide hinnaerinevuste arvestamiseks on maa väärtus leitud eraldi iga liigi kohta ja selleks on kasutatud maakondlikku jaotust. Andmed katastriüksuste, tehingute ja hindade kohta on kättesaadavad isegi detailsemalt, aga konfidentsiaalsusnõuete tõttu läheb see osa infost kaduma.

1991. aasta oktoobris võeti vastu maareformi seadus, mille eesmärk oli tagastada riigistatud maaomand eraisikutele, anda varem riigile kuulunud maa eraõigusliku isiku, avalik-õigusliku juriidilise isiku või kohaliku omavalitsusüksuse omandisse ning määrata kindlaks riigi omandisse jäetav maa. (Eesti ...1991) 1992. aasta detsembris registreeriti maakatastris esimesed maatükid, enamik registreerimisi tehti aastatel 1995–2003. Aastal 2015 oli Maa-ameti andmetel reformitud 97% varem riigile kuulunud maast.

Joonis 1 näitab seda, kui suur osa kogu territooriumist on 16 aasta jooksul maakondades keskmiselt registreeritud. Maakatastris oli maa registreerimise algusaastatel Maa-ameti registreeritud maatükkide osatähtsuse erinevus maakonniti märkimisväärne^a, aga aja jooksul on need piirkondlikud vahed vähenenud. Kui 2000. aastal oli registreeritud vaid 56% maakondade territooriumist, siis 2004. aastal ületati juba 80% tase ning 2015. aastaks oli maakatastris registreeritud 95–99% maakondade territooriumist.

^a 2000. aastal oli maakatastris registreeritud vaid 32% Harju ja Saare maakonnas asuvast maast, samal ajal oli Viljandi ja Põlva maakonna maast registreeritud enam kui 70%.

Joonis 1. Maa-ametis registreeritud maa osatähtsus kogu territooriumis, 2000–2015*Figure 1. Share of land registered with the Land Board in the whole territory, 2000–2015*

Maa väärtuse hindamise täpsuse suurendamiseks peaks maa liikide klassifikaator hõlmama katastris kõige tähtsamaid maa alamtüüpe. Kuna maa hind sõltub olulisel määral selle otstarbest, muudaks maa liikide detailse jaotuse lisamine arvestuse tuntavalt täpsemaks. Rahvusvahelise võrdluse tagamiseks soovitas OECD (2015) hinnata maa väärtust järgmiste liikide alusel: hoonete ja rajatiste alune maa, kultiveeritav maa, puhkealad koos pinnaveega ja muu maa. Need neli kategooriat jagunevad omakorda alamkategooriateks, et teha maa väärtuse hindamine võimalikult täpseks, näiteks hoonete ja rajatiste alune maa jaguneb eluruumide aluseks maaks ning muude hoonete ja rajatiste aluseks maaks (äriksel, tööstuslikul jne eesmärgil kasutamiseks).

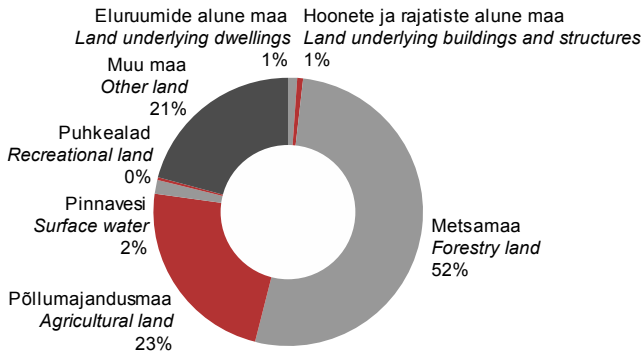
Et saada iga maa liigi puhul suurim vastavus OECD käsiraamatuga (2015), jaotati Maa-ameti andmebaasist saadud andmed kõlvikute ja maa sihtotstarbe järgi. Kõlvik on ühetaolise majandusliku sihtotstarbe ja/või loodusliku seisundiga katastriüksuse osa, mida ei piiritleta piirimärkidega. Sihtotstarve viitab maa tegelikule ja määratud kasutamisele. (Maakatastriseadus 1994) Näiteks võib ühe põllumajandusettevõtte maa koosneda mitmest kõlvikust (põllu-, heina-, karjamaa jne). Maa-ametist saadi 2000.–2015. aasta 15 maakonna katastriandmed järgmiste liikide kohta:

- eluruumide alune maa;
- hoonete ja rajatiste alune maa;
- põllumajandusmaa;
- metsamaa;
- pinnavesi;
- puhkealad;
- muu maa.

Joonisel 2 on näha maa jaotus 2015. aastal. Maa-ameti andmetel on üle poole (52%) Eestis registreeritud maast metsamaa. Osatähtsuse poolest järgnevad põllumajandusmaa (23%) ja muu kasutusega maa (21%). See tähendab, et mitmesuguse hoonestusega, puhkealadega ja pinnaveega on kaetud vaid 4% kogupinnast. Aastate jooksul pole see jaotus eriti muutunud.

Joonis 2. Maa jaotus liigi järgi, 2015

Figure 2. Land structure by type, 2015



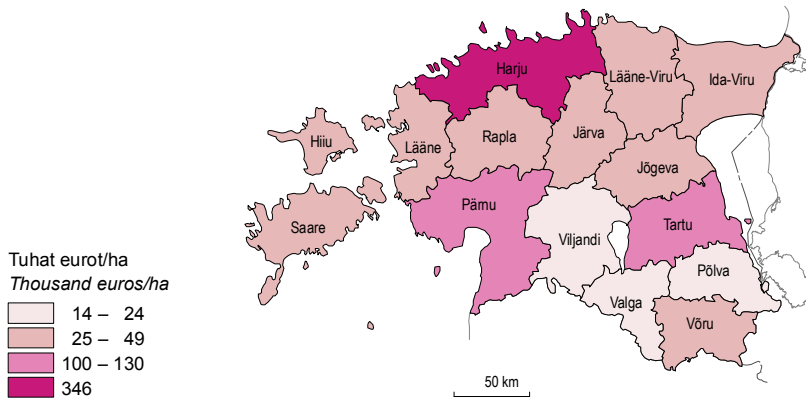
Allikas: Maa-amet.
Source: the Land Board.

Otsesel meetodil maa väärtuse hindamise eeltingimus on turuhindade kättesaadavus iga maa liigi kohta, sest väärtuse otseseks hindamiseks on vaja nii pindala andmeid kui ka vastavat hinnainfot. Hind peaks peegeldama tegelikku turutehingu hinda või sellele lähedast hinda. Viimase puhul kasutatakse õiglase hinna kontseptsiooni, mille järgi leitakse turutehingu puudumisel hind sellele võimalikult sarnase tehingu järgi.

Maa hind on mitme teguri suhtes väga tundlik ja seda mitte ainult maakasutuse alusel, vaid ka piirkondlikul tasandil. Järgmised kaardid näitavad suuri maakondlikke erinevusi maa hinnas: pole sugugi üllatav, et tootmis- ja ärimaa väärtus hektari kohta on suurem kui elamumaa oma ning et võrreldes metsa- või põllumajandusmaa hinnaga on erinevus veelgi suurem.

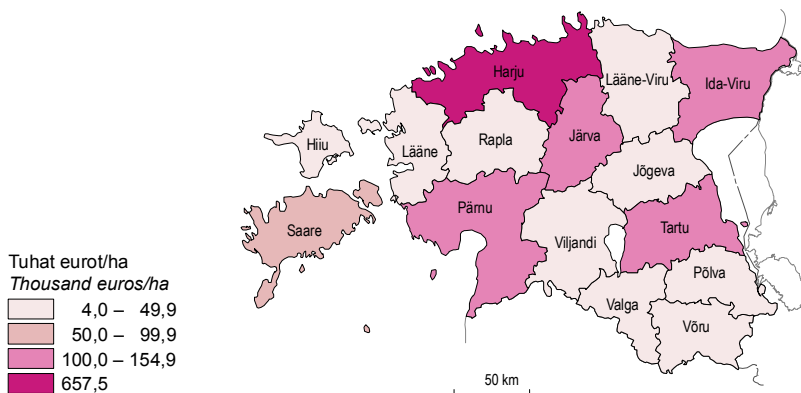
Eestis hinnati viimati korralliselt maad (maa korralisi hindamisi tehakse maksustamise eesmärgil) 2001. aastal. Et katastriväärtused ei pruugi peegeldada turuhindu, siis kasutati arvestustes hoopis Maa-ameti andmeid tehingute kohta. Maa väärtuse hindamiseks on olemas kaks andmestikku: üks koos hoonetega ja teine ilma hooneteta. Tehingute andmeid on kogutud alates 1996. aastast ja neid avaldatakse Maa-ameti veebilehel vaid siis, kui on olnud vähemalt viis tehingut.

Maa väärtus leitakse maa turuhinda kasutades, sellest on maha arvatud maa omandiõiguse üleandmisega kaasnevad kulud.

Kaart 1. Eluruumide aluse maa väärtus, 2015*Map 1. The value of land underlying dwellings, 2015*

Allikas: Maa-amet.
Source: the Land Board.

Kaardilt 1 on näha, et suurim on elumumaa väärtus Harjumaal – 346 000 eurot/ha. Võrreldes 2005. aastaga, mil sama näitaja oli 273 000 eurot/ha, on selle maaliigi väärtus suurenenud ligikaudu 27%. Samal ajal on väärtuselt järgmiste maakondade Tartu (123 000 (2005. a) ja 130 000 (2015. a)) ja Pärnu (107 000 ja 104 000) väärtus püsinud peaaegu muutumatuna. Kiireimini on elumumaa väärtus suurenenud Hiiumaal (+236%) ja Põlvamaal (+138%). Suurim väärtuse langus on võrreldaval perioodil olnud Valgamaal (–78%), Viljandimaal (–56%) ja Läänemaal (–52%).

Kaart 2. Tootmis- ja ärimaa ilma hooneteta, 2015*Map 2. Production and commercial land without buildings, 2015*

Allikas: Maa-amet.
Source: the Land Board.

Tootmis- ja ärimaa hinnaerinevused on Harjumaa ja ülejäänud Eesti vahel võrreldes eluruumide aluse maaga veegi suuremad. Harjumaa (658 000 eurot/ha) ja maa hinna poolest teisel kohal oleva Pärnumaa (153 000) erinevus on enam kui neljakordne. Odavaim tootmis- ja ärimaa on Jõgevamaal (5000), Põlvamaal (8000) ja Hiiumaal (8000). Ka eluruumide aluse maa on neis maakondades Eesti odavaim. Kui eluruumide aluse maa väärtus on mitmel pool vahemikus 2005–2015 tõusnud, siis tootmis- ja ärimaa puhul on protsess vastupidine. Kuna 2005. aastal oli Eestis

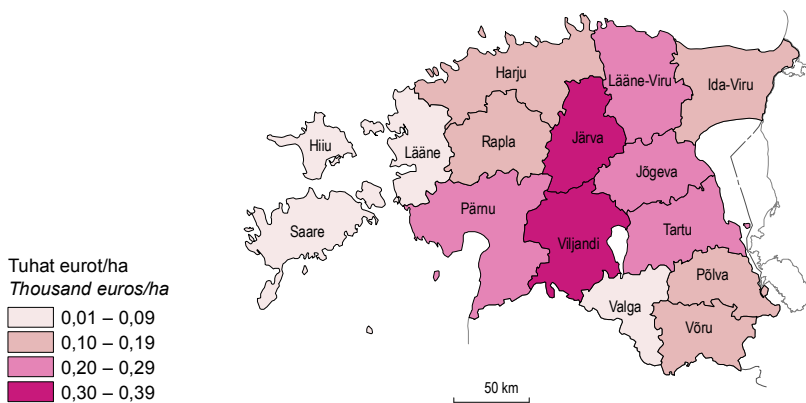
juba kiirenev majanduskasv ja oli näha märke majandusbuumist, siis suurenesid toona ka tootmis- ja ärimaa hinnad üsna kiirelt. Kuigi hinnad saavutasid lae aastatel 2006 ja 2007, olid juba 2005. aastal mitmes maakonnas hinnad kõrged. Nii tuli Harjumaal hektari eest välja käia veidi üle miljoni euro, Pärnumaal 870 000 ja Tartumaal 492 000.

Elamumaa ning tootmis- ja ärimaa hinnad erinevad piirkonniti suurel määral eelkõige piirkondlike erinevuste tõttu majandustegevuses ja sellest tulenevalt ka rahvastiku paiknemises. Majanduslikud põhjused taoliseks jaotuseks võivad olla nii logistilised kui ka ajaloolised. Äri asupaiga valimisel on oluline nii ligipääs transpordihendusele kui ka muule taristule. Samuti ei pea alati paika asjaolu, et inimesed liiguvad sinna, kuhu rajatakse ettevõtteid. Sageli rajatakse uusi suuri ettevõtteid hoopis sinna, kus on sobiva kvalifikatsiooniga tööjõudu lihtsam leida. Neil põhjustel eelistataksegi nii elu- kui ka ettevõtluspiirkonnana valdavalt ajaloolisi suuremaid keskuseid.

Primaarsektoris (tegeleb põllumajanduse, metsanduse ja kalandusega) ei ole aga eespool mainitud majanduslikud ja tööjõu vajadused niivõrd spetsiifilised. Kuigi joonisel 2 on näha, et maad on selles sektoris palju, on võimalik kasutada keskustest kaugemal asuvat maad. Seega on selles sektoris ka maa väärtus madalam ja piirkondadevahelised erinevused väiksemad.

Kaart 3. Põllumajadusmaa, 2015

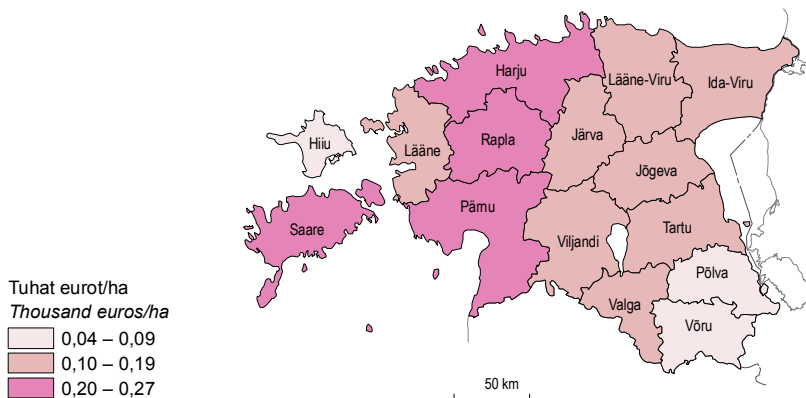
Map 3. Agricultural land, 2015



Kõige kallim põllumajadusmaa Eestis oli 2015. aastal Järvamaal ja see maksis 330 eurot hektari kohta. Järgnesid Viljandi (320 eurot/ha) ja Lääne-Viru maakond (290 eurot/ha). Odavaim põllumajadusmaa oli Hiiu (10 eurot/ha), Valga (70 eurot/ha) ja Lääne (90 eurot/ha) maakonnas. Muidu kallima kommerts- ja elamumaaga keskused jäävad siin keskmike hulka – Harju maakonnas oli põllumajadusmaa hind 140 eurot/ha, Tartu maakonnas 210 eurot/ha ja Pärnu maakonnas 220 eurot/ha. Et põllumajadusmaa hinnad on mitme teguri tõttu piirkonniti ühtlasemad, siis on seda olnud ka põllumajadusmaa väärtuse muutus. Võrreldes 2005. aastaga on põllumajadusmaa väärtus kasvanud kõigis maakondades mitu korda.

Kaart 4. Metsamaa, 2015

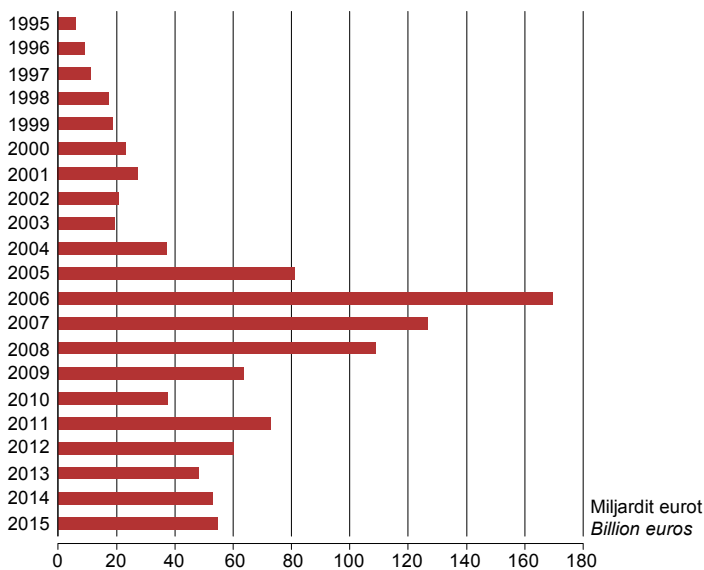
Map 4. Forestry land, 2015



Metsamaa väärtust iseloomustavad põllumajandusmaa väärtusega sarnased suurusjärgud, kuid veel väiksemate piirkondlike erinevustega. Kalleim metsamaa oli 2015. aastal Pärnu maakonnas (270 eurot/ha). Sellele järgnesid Saare (260 eurot/ha), Harju (230 eurot/ha) ja Rapla maakond (230 eurot/ha). Väikseima väärtusega metsamaa oli Hiiu (40 eurot/ha), Võru (80 eurot/ha) ja Põlva (80 eurot/ha) maakonnas. Kuigi ka metsamaa puhul oli vaatlusaluse kümnendi puhul toimunud märgatav väärtuse tõus, siis jäi see enamikul juhtudel 50–150% vahele. Erand on vaid Pärnu maakond, kus metsamaa väärtuse kasv on olnud 207%, ja Hiiu maakond, kus metsamaa väärtus suurenes 10 aasta jooksul vaid 14%.

Joonis 3. Maa väärtus Eestis, 1995–2015

Figure 3. Land value in Estonia, 1995–2015



Kui 1995. aastal oli kogu Eesti maa väärt ligikaudu 6 miljardit eurot, siis aja jooksul on väärtus tublisti suurenenud. Jooniselt 3 selgub, et kuni 2003. aastani ületas see vaid korraks 20 miljardi piiri. Koos 2004. aasta kiireneva majanduskasvuga hakkas ka maa väärtus suurenema ja kasvas majandusbuumi ajal hüppeliselt. 2004. aastal oli maa väärt 37 miljardit eurot, 2005. aastal

81 miljardit eurot ja 2006. aastal juba 170 miljardit eurot. Majanduskriisi saabudes hakkas ka maa väärtus kiirelt langema. 2010. aastaks oli varasemast kõrgest väärtusest järele jäänud 38 miljardit eurot ehk peaaegu sama palju kui majandusbuumi alguses 2004. aastal.

Ka pärast majanduskriisi on maa väärtus liikunud sarnaselt üldisele majanduse käekäigule. Kiirele kriisist väljumisele järgnes mõningane stagnatsioon. Nii hüppas maa väärtus 2011. aastal veidi optimistlikult kohe 73 miljardi euroni, kuid seejärel langes taas veidi ja stabiliseerus 2015. aastal 54 miljardi euro juures.

Kokkuvõte

Maa on kõige vanem mittetoodetud kapitali vorm ja seda võib pidada nii keskkondlikuks kui ka majanduslikuks varaks. Eurostati grandiprojekti eesmärk oli hinnata maa väärtust rahvamajanduse arvepidamises. Projekti käigus töötati välja olemasolevatel allikatel põhinev arvutusmetoodika rahvusvaheliselt heaks kiidetud arvestuspõhimõtetest lähtudes. Maa väärtuse arvutamise otsest meetodit rakendades saadi Eestis kasutusele võetud maa väärtuse aegrida alates 1995. aastast.

Eesti on väike ja avatud majandusega riik, mistõttu on ka sinne majandusareng volatiilsem kui suuremate majandustega riikides. Kinnisvara on üks majanduse tugisammas (vara). See tähendab, et ka kinnisvara väärtus muutub koos majanduse tsüklitega. Sellest tulenevalt on kinnisvaratehingute hinnad nii hoonestatud kui ka hoonestamata maa puhul olnud viimase 20 aasta jooksul väga muutlikud.

Viimase majandusbuumi ja kriisi ajal (2005–2009) kõikus maa väärtus rohkem kui majanduskasv. Eelkõige oli see tingitud kõikumistest elamu- ning tootmis- ja ärimaa hindades. Kui perioodil 2005–2015 muutus nende maaliikide väärtus kokkuvõttes maakonniti erinevalt, siis majandusbuumi ajal aastatel 2006–2007 läksid hinnad üles igal pool. Põllumajandus- ja metsamaa, mis on majandustsüklite mõju suhtes vähem tundlikud, on kogenud stabiilses tempos hinnakasvu.

Perioodil 1995–2015 on Eestis maa väärtus suurenenud pea 6 korda – 6 miljardilt eurost 1995. aastal 54 miljardi euroni 2015. aastal. Eespool viidatud suurte hinnakõikumiste tõttu jõudis maa väärtus vahepeal aga veelgi kõrgemale. 2006. aastal oli maa Eestis hinnanguliselt väärt koguni 170 miljardit eurot. Sarnaselt viimaste aastate majanduse üldisele arengule viitab ka maa väärtuse trend edasisele stabiilsele kasvule.

Tabel 1. Eluruumide alune maa maakonna järgi, 2005, 2015

Table 1. Land underlying dwellings by county, 2005, 2015
(tuhat eurot – thousand euros)

	2005	2015	Muutus %, Change, %
Harju	273,0	345,9	27
Hiiu	11,9	40,0	236
Ida-Viru	51,2	42,6	-17
Järva	40,9	30,9	-24
Jõgeva	24,6	36,5	48
Lääne	79,2	37,8	-52
Lääne-Viru	44,4	37,6	-15
Pärnu	106,7	103,9	-3
Põlva	8,6	20,5	138
Rapla	34,8	48,0	38
Saare	25,2	29,8	18
Tartu	123,8	130,1	5
Valga	61,7	13,7	-78
Viljandi	44,1	19,2	-56
Võru	33,3	26,6	-20

Tabel 2. Tootmis- ja ärimaa maakonna järgi, 2005, 2015*Table 2. Production and commercial land by county, 2005, 2015*

(tuhat eurot – thousand euros)

	2005	2015	Muutus, % Change, %
Harju	1 021,0	657,5	-36
Hiiu	..	8,4	..
Ida-Viru	27,6	113,4	312
Järva	30,9	123,3	299
Jõgeva	23,0	4,7	-80
Lääne	241,1	32,2	-87
Lääne-Viru	106,0	17,0	-84
Pärnu	869,9	152,6	-82
Põlva	69,2	7,5	-89
Rapla	41,7	14,2	-66
Saare	123,6	91,2	-26
Tartu	492,5	101,1	-79
Valga	95,8	27,8	-71
Viljandi	15,2	22,9	51
Võru	..	26,2	..

Tabel 3. Metsamaa maakonna järgi, 2005, 2015*Table 3. Forestry land by county, 2005, 2015*

(tuhat eurot – thousand euros)

	2005	2015	Muutus, % Change, %
Harju	0,1	0,2	81
Hiiu	0,0	0,0	14
Ida-Viru	0,1	0,2	124
Järva	0,1	0,2	70
Jõgeva	0,1	0,1	107
Lääne	0,1	0,2	139
Lääne-Viru	0,1	0,2	46
Pärnu	0,1	0,3	207
Põlva	0,0	0,1	125
Rapla	0,1	0,2	72
Saare	0,1	0,3	82
Tartu	0,0	0,1	258
Valga	0,1	0,1	115
Viljandi	0,1	0,1	70
Võru	0,1	0,1	56

Tabel 4. Põllumajandusmaa maakonna järgi, 2005, 2015

Table 4. Agricultural land by county, 2005, 2015
(tuhat eurot – thousand euros)

	2005	2015	Muutus, % Change, %
Harju	0,0	0,1	382
Hiiu	0,0	0,0	891
Ida-Viru	0,0	0,1	320
Järva	0,1	0,3	411
Jõgeva	0,1	0,2	325
Lääne	0,0	0,1	708
Lääne-Viru	0,1	0,3	354
Pärnu	0,1	0,2	176
Põlva	0,0	0,1	1 007
Rapla	0,0	0,2	447
Saare	0,0	0,1	334
Tartu	0,0	0,2	353
Valga	0,0	0,1	126
Viljandi	0,0	0,3	539
Võru	0,0	0,1	1 052

Allikad Sources

Eesti Vabariigi maareformi seadus. (1991). [www] <https://www.riigiteataja.ee/akt/30543> (4.01.2017).

Eurostat-OECD compilation guide on land estimation. (2015). [www] <http://ec.europa.eu/eurostat/documents/3859598/6893405/KS-GQ-14-012-EN-N.pdf/> (6.01.2017).

Land Cadastre Act. (1994). [www] <https://www.riigiteataja.ee/en/eli/ee/Riigikogu/act/522062016005/consolide> (5.01.2017).

Land Reform Act. (1991). [www] <https://www.riigiteataja.ee/en/eli/ee/Riigikogu/act/504092017013/consolide> (4.01.2017).

Maakatastriseadus. (1994). [www] <https://www.riigiteataja.ee/akt/28705> (5.01.2017).

LAND VALUE: HOW MUCH IS ESTONIA WORTH?

Reelika Parve, Robert Mürsepp

Each economic activity requires land: some do more and some less, and land is required either to be used directly in the production process or indirectly as the plot of land where the production unit is situated. This article gives an overview of how much land is worth in Estonia and how much it costs to purchase different types of land.

Introduction

Considering all of the above, it is important to know land value. When determining the value of a piece of land (irrespective of the land type and user's needs), multiple factors should be considered. One could highlight its good location and proximity to urban centres or, on the contrary, its remoteness and the quiet of its surroundings, nearby current or future infrastructure, its soil productivity, etc. There are too many factors affecting land value to list them here. Indeed, the value of land correlates with various indicators, but the list of the relevant variables depends primarily on the intended use. To determine the price for land one can ask for an expert evaluation or try to make an estimation by using other available sources of information, for example, by comparing the value of plots of land with similar characteristics. Naturally, this kind of evaluation which takes into account each and every characteristic is not suitable for providing an estimate of the land value of the whole country.

This short article highlights the findings of the Eurostat project^a which had the aim to estimate Estonian land value over the last 20 years. It was necessary to develop a methodology for estimating land value based on the framework of the European System of National and Regional Accounts 2010 (ESA 2010). Theoretical support to the project was found in the manual compiled by Eurostat and the OECD. (Eurostat-OECD ... 2015)

Methodology

Land represents an asset on the balance sheet of a country. The method for estimating land value (direct or indirect) depends on which type of data is available. In case of the direct method, land value is estimated in as much detail as possibly by land type. In order to do that, it is necessary to:

- estimate land area by type in one or more years;
- estimate annual changes in land types to create a time series;
- estimate approximate prices for different land types in one or more years;
- model price changes of different land types to obtain time series for unit prices;
- combine area and price data to create a time series of land value;
- specify annual volume and price changes.

With indirect methods, it is possible to estimate land value with other more aggregated indicators or together with some other indicator. For example:

- estimate the value of land with the buildings and structures situated on it;
- find land value as an extra indicator from an estimation with a broader scope;
- estimate the value of land through the ratios of land and structures;
- estimate land value with a regression model.

^a Action to support specific improvements in National Accounts. Objective 3: Improvements in availability, timeliness and quality of annual and quarterly national accounts, supply and use tables and regional accounts.

According to recommendations of the OECD (2015), the evaluation of land should be made at the most detailed level possible, incorporating both regional as well as land type differences. In Estonia, the best data source for this kind of estimates is the Land Board, which has both land area and price data available. Some additional information, such as detailed annual data about the ownership, use, characteristics and market prices of forestry land, was obtained from the Estonian Environment Agency.

The estimation of Estonian land was carried out using the direct method as data with the required level of detail is available. In the case of the direct method, each piece of land is multiplied by an appropriate price for this land type. Then the values of all parcels are added to obtain the total value of land in the territory of a country. Even if the direct land value calculation appear easy in theory, it is still very demanding in practice. Ideally, the price and area information of every parcel of land should be accessible, but obtaining data with this level of detail might be rather difficult.

The land value in year t is estimated by the following formula:

$$LV_t = \sum_{i=1}^n x_{i,t} * p_{i,t}$$

where:

i = land type (agricultural, forestry land, etc.);

t = reference year (1995–2015);

$x_{i,t}$ = land area by type i in year t ;

$p_{i,t}$ = price of land type i in year t .

To capture price differences of land types, land value has been calculated separately for each land type using county level estimation. Data about cadastral units, transactions and prices are available at even a lower level of detail, but due to confidentiality requirements this part of the information is lost.

In October 1991, the Land Reform Act was passed. The aim of the act was to transfer nationalized land ownership to private persons, transfer the land previously owned by the state into the ownership of persons in private law, legal persons in public law or local government and determine the land to be kept in state ownership. (Land ... 1991) In December 1992, the first plots of land were registered in the land cadastre, the majority of registrations took place between 1995 and 2003. According to the Land Board, 97% of the land previously owned by the state had been reformed in 2015.

Figure 1 (p. 76) presents the average share of land registered in counties over the last 16 years. In the first years of land registrations, in the land cadastre, the differences in the share of plots registered with the Land Board were considerable between counties^a, but over time this regional discrepancy has decreased. Whereas only 56% of the territory of counties had been registered in 2000, 80% share was surpassed in 2004, and 95–99% of the territory of counties had been registered in the land cadastre by 2015.

To increase the accuracy of land estimates, the classification of land types in the cadastre should contain the most important subtypes of land. As the price of land depends to a great extent on its intended use, adding a detailed division of land types would make the estimates significantly more precise. For the purpose of international comparison, the OECD (2015) recommended to estimate land according to the following types: land underlying buildings and structures, land under cultivation, recreational land and associated surface water and other land. These four categories are divided into additional sub-categories to make the estimation as accurate as possible: for example, land underlying buildings and structures is divided between land underlying dwellings and land underlying other buildings and structures (for commercial, industrial use, etc.).

^a In 2000, only 32% of the land in Harju and Saare counties was registered in the land cadastre, while more than 70% of the land in Viljandi and Põlva counties had been registered.

In order to find the most appropriate match for each particular land type according to the OECD manual (2015), the data extracted from the Land Board database was divided between land use type and intended use. Land use type is a part of a cadastral unit which has the same intended economic use and/or natural status and which is not delimited by boundary markers. Intended use refers to the land's actual and designated utilization. (Land ... 1994) For example, land belonging to one agricultural enterprise can comprise many land use types, e.g. agricultural land, grassland, pasture land, etc. Cadastral records of 15 counties for years 2000–2015 were obtained from the Land Board for the following land types:

- land underlying dwellings;
- land underlying other buildings and structures;
- agricultural land;
- forestry land;
- surface water;
- recreational land;
- other land.

Figure 2 (p. 77) shows land structure in 2015. According to the Land Board, more than half of the land (52%) registered in Estonia is forestry land. Next in terms of share are agricultural land (23%) and land used for other purposes (21%). This means that 4% of the total area is under various buildings, recreational areas and surface water. This structure has not changed much over the years.

The precondition for estimating land by the direct approach is the availability of market prices for each land type, as for the direct approach both data on surface areas as well as appropriate price information is required. The price should reflect the actual market transaction price or its equivalent. In the latter case, fair price concept is used, which means that if there is no market transaction, the price is determined based on a similar transaction.

In Estonia, the last regular assessment of land (regular assessments are conducted for taxation purposes) took place in 2001. As cadastral values might not reflect market prices, the Land Board's data of transactions were used instead. Two datasets are available for land estimates: one with buildings and the other without buildings. The data on transactions have been collected since 1996, and they are published on the website of the Land Board only if there have been at least five transactions.

Land value is estimated using the market price, excluding the costs related to the transfer of ownership.

Map 1 (p. 78) shows that Harju county has the highest residential land value – 346,000 euros/ha. Compared to 2005 when the indicator was 273,000 euros/ha, the value of this land type has increased approximately 27%. At the same time, in the counties following in terms of value it has remained almost unchanged – in Tartu county the value has increased from 123,000 (2005) to 130,000 (2015) and in Pärnu county from 107,000 to 104,000. The value of residential land has increased faster in Hiiu (236%) and Põlva (138%) counties. The greatest decrease in value during the observed period was in Valga (–78%), Viljandi (–56%) and Lääne (–52%) counties.

The price differences of production and commercial land are even greater between Harju county and the rest of Estonia compared to land underlying dwellings (Map 2, p. 78). The difference between Harju county (658,000 euros/ha) and Pärnu county (153,000), which comes second in land value, is more than four-fold. Production and commercial land is cheapest in Jõgeva county (5,000), Põlva county (8,000) and Hiiu county (8,000). These counties also have the cheapest residential land in Estonia. Whereas the value of residential land has increased in many areas in 2005–2015, in the case of production and commercial land the process has been the opposite. As in 2005 economic growth was becoming faster in Estonia and there already appeared boom signs,

the prices of production and commercial land rose quite fast. Although prices reached their maximum in 2006 and 2007, they were high in several counties already in 2005. For example, a hectare cost a little over a million euros in Harju county, 870,000 in Pärnu county and 492,000 in Tartu county.

The prices of land underlying dwellings and production and commercial land differ by regions to a great extent primarily due to regional differences in economic activity and correspondingly in population distribution. Economic reasons for this distribution may be both logistical as well as historical. When choosing a business location, access to transportation connections as well as to other infrastructure are both important. Also, the fact that people move where businesses are established does not always ring true. Oftentimes, new large enterprises are established where it is easier to find labour force with suitable qualifications. For these reasons, mainly larger historical centres are preferred as residential and business regions.

In the primary sector (agriculture, forestry and fishing), however, the aforementioned economic and labour force needs are not so specific. Although Figure 2 (p. 77) shows that there is a lot of land in this sector, it is possible to use land further away from centres. Therefore, land value in this sector is also lower and regional differences smaller.

The most expensive agricultural land in Estonia in 2015 was in Järva county amounting to 330 euros per hectare. This was followed by Viljandi (320 euros/ha) and Lääne-Viru county (290 euros/ha). Agricultural land was least expensive in Hiiu (10 euros/ha), Valga (70 euros/ha) and Lääne (90 euros/ha) counties. The centres with otherwise more expensive commercial and residential land are average here – the price of agricultural land was 140 euros/ha in Harju county, 210 euros/ha in Tartu county and 220 euros/ha in Pärnu county. As the prices of agricultural land are more even by regions due to several factors, the same applies to change in the value of agricultural land. Compared to 2005, the value of agricultural land has increased many-fold in all counties.

The value of forestry land is characterized by similar ranges as the value of agricultural land, but the regional differences are even smaller. In 2015, the most expensive forestry land was in Pärnu county (270 euros/ha), followed by Saare (260 euros/ha), Harju (230 euros/ha) and Rapla (230 euros/ha) counties. Forestry land was least valuable in Hiiu (40 euros/ha), Võru (80 euros/ha) and Põlva (80 euros/ha) counties. Although there was also a remarkable rise in value of forestry land in the observed decade, in most cases it fell between 50–150%. The exceptions are Pärnu county, where the increase in forestry land value was 207%, and Hiiu county, where the value of forestry land increased only 14% in 10 years.

While in 1995 the total land value in Estonia was approximately six billion euros, the value has grown considerably over time. It is evident from Figure 3 (p. 80) that before 2003 it exceeded the 20 billion threshold only once. In 2004, in line with faster economic growth, the value of land started to rise, increasing markedly during the economic boom. In 2004, the land value was 37 billion euros, in 2005, it was 81 billion euros and, in 2006, already 170 billion euros. By 2010, the previously high value had decreased to 38 billion euros, i.e. almost as much as during the economic boom in 2004.

After the economic crisis, the value of land has fluctuated similarly to the overall condition of the economy. A swift move out of the crisis was followed by some stagnation. Immediately land value soared somewhat optimistically to 73 billion euros, but then experienced some decline again and stabilised at 54 billion euros in 2015.

Conclusion

Land is the oldest form of non-produced capital, and it can be considered as an environmental as well an economic asset. The Eurostat grant project was aimed at estimation of land value in the national accounts. This led to the development of the calculation methodology according to the available data sources proceeding from internationally approved estimation methods. In applying the direct approach to calculating the value of land utilised in Estonia, a time series was produced starting from 1995.

Estonia is a small country with an open economy, which is the reason why economic development here is more volatile than in countries with larger economies. Real estate is one pillar (asset) of the economy. This means that real estate values change with economic cycles as well. Due to this, the prices of real estate transactions in both land with or without buildings have been very sensitive in the last 20 years.

During the last economic boom and crisis (2005–2009), land value fluctuated more than economic growth. This was caused primarily by fluctuations in the prices of residential and production and commercial land. Whereas overall in 2005–2015, the value of these land types changed differently in counties, during the economic boom in 2006–2007, prices went up everywhere. Agricultural and forestry land that are more stable in terms of economic activity have experienced a stable growth trend.

In 1995–2015, land value has increased almost six times in Estonia – from six billion euros in 1995 to 54 billion euros in 2015. Due to the aforementioned significant price fluctuations, land value reached even higher levels at times. In 2006, the estimated land value in Estonia was 170 billion euros. Similarly to overall economic development, land value is indicating continuing stable growth.

BALTI LOENDUSE SEMINAR TÄHISTAS 2017. AASTAL 20. AASTAPÄEVA

Diana Beltadze

Balti riikide rahva ja eluruumide loenduse seminaride ajalugu on eelkõige Eesti, Läti ja Leedu statistikute, loendusjuhtide ja ekspertide koostöö ajalugu. Artiklis tutvustatakse Baltimaade statistikaametite koostöö olulisemaid rõhuasetusi loenduse teemal.

Rahva ja eluruumide loendust tehakse tavaliselt iga kümne aasta järel ja see on rahvastikuarengu kirjeldamisel kõige autoriteetsem andmeallikas. Selle eesmärk on koguda andmeid riigi rahvaarvu, rahvusliku koostise ja paiknemise ning eluruumide kohta. Taasiseseisvunud Eesti, Läti ja Leedu on teinud kaks loendust, mille tulemused on rahvusvaheliselt võrreldavad, sest on järgitud ÜRO soovitusi loendusvoorudeks (Principles ... 2008; Main ... 2009). Loendusmetoodika valik on riigiti vaba, kuid see peab järgima Euroopa Liidu kvaliteedinõudeid. Seetõttu on rahvusvaheline koostöö oluline: aitab mõista võimalusi loenduste korraldamisel.

Järgmises loendusvoorus plaanivad Eesti ja Läti esimest korda regiooni ajaloos teha loenduse registre andmetele tuginedes, st tulemused pannakse kokku riiklikes andmekogudes juba olemasoleva info põhjal. Leedu on teatanud, et mitmesuguseid asjaolusid arvestades tuleb kavandada järgmine loendus veel kombineeritud metoodikat kasutades, aga selle täpsem kuju ei ole veel teada. Seega on praegu vara analüüsida Balti riikide valmisolekut järgmiseks loenduseks, sest ettevalmistustööd alles käivad, kuid tagantjärele on võimalik anda ülevaade Balti riikide proovikividest loendustel pärast taasiseseisvumist, kasutades selleks Balti loenduse seminaride materjale. 2017. aastal toimus Balti loenduse seminar 20. korda.

Balti loenduse seminar kui koostöövorm

Esimene Baltimaade rahvaloenduse seminar toimus 1995. aastal Leedus. Sellest ajast peale on seminare korraldatud Eestis ja Lätis seitsmel korral ja Leedus kuuel korral. Seminaride teemad on alati olnud pühendatud loendustöö kõige olulisematele ülesannetele (nt loendusmeetodi valik, andmekvaliteedi hindamine, tehnoloogia kasutamine) või on seal selgitatud võimalusi registre kasutamiseks loendusstatistikas. Teemad on kooskõlastatud riikide eri osapoolte vahel, niisamuti on määratud prioriteedid teemade käsitlemisel. Seminarist osavõtjate piirarvu ja seminari kestuse määramisel on alates 2007. aastast lähtutud statistikaametite majanduslikest võimalustest. Varem oli seminaridest osavõtt arvukam ja need toimusid pikemalt.

Koostööetappe võib eristada mitmeti. Neid võib käsitleda vastavalt koostööle eri osapooltega, nt Põhjamaade abiprogrammi ajajärk 2000. aasta loendusvooru ettevalmistamiseks Baltikumis; võib eristada teemade järgi 2000. ja 2011. aasta loenduse ettevalmistamise etappe ning 2011. aasta loenduse järgset aega. Seminare võib liigendada ka osavõtjate arvu järgi (Conference ... 2017).

Kuidas koostöö alguse sai ja kes oli algataja, ei ole enam täpselt võimalik teada saada, aga koostöö vajadust loenduse ettevalmistamisel Balti riikides tutvustavad mitu Eesti Statistikaameti arhiividokumenti 1995. aastast. Koostööettepanekuid loenduse teemadel võib leida eelkõige Statistikaameti juhtivtöötajate Mati Sundja ja Lembit Tepi (Tepp) ametkondlikust kirjavahetusest (1993–1995) Läti ja Leedu loendusjuhtidega.

Esimene loendusseminar toimus Leedus. Algaastail aitasid seminare korraldada Põhjamaad, kelle koordineeriva käe all toimusid ettevalmistustööd 2000. aasta loendusvooruks. Hiljem, kui Eestist, Lätist ja Leedust said Euroopa Liidu riigid, koostöövõrgustik laienes ja Balti riikide loendusjuhid said seminaride korraldamisega ise hakkama. Praeguseks jagavad Balti riigid teistele riikidele loendusteadmust eelmisel loendusel tulemuslikult kasutatud kombineeritud loendusmeetodi kohta.

Võib nentida, et kahekümne aasta jooksul on arendatud koostööoskusi ja välja on kujunenud koostööaldis hoiak. Aastakümnete jooksul tehtud koostööd iseloomustab kolme riigi erinevate loendusmeeskondade sisuliste, tehniliste ja ka emotsionaalsete oskuste arendamine, osapoolte ära kuulamine ning vigadest õppimise väärtustamine jne.

On tavaline, et igaks loenduseks tuleb suuremal või vähemal määral võtta tööle ja koolitada uus arvukas meeskond. Enda loendusteemaga kurssi viimine on aeganõudev protsess igale loendusorganisatsioonile. Tihti on uutel, loendusteemade puhul kogenematutel meeskonnaliikmetel raske eristada loendusstatistika töös olulist infot ebaolulisest. Aegkriitilises loendustöös tundub kõik ülitähtis. Sellistel rasketel hetkedel on abi saanud infovahetusest Läti ja Leedu kolleegidega. Seetõttu on Balti seminar olnud loenduste tegemise supervisiooni loomise koht mitte ainult loendusjuhtidele, vaid ka loendusvõrgustiku spetsialistidele. Üheskoos teemade läbi arutamine ja mõtlemine on aidanud kaasa parema mõistmise tekkimisele loendustööde korraldamisel.

Seminaridel käsitletud teemad

Balti loenduse seminaride ajalugu on teemade kaupa jagatud kolme perioodi (tabel 1). Teemasid saab eristada loendusmeetodi valiku järgi.

- Esimene loendus pärast taasiseseisvumist (2000. aasta) toimus kõigis kolmes riigis traditsioonilisel viisil;
- Eelmine loendus (2011. aasta) tehti Baltimaades kombineeritud loendusmetoodikat rakendades.

Balti seminaridel on olnud üksmeelel selles, et statistikaametitel on võimalus pakkuda erinevaid meetodeid loendusandmete kogumiseks, kuid samas peab valitud meetodil tehtud statistika kajastama võimalikult tõetruult tegelikkust olukorda, et tarbijad ei kaotaks usaldust loendustulemuste vastu.

Loenduse tegemine eeldab peale loenduse organisatsioonide loomise veel seda, et statistikaorganisatsioonides on loodud keskkond, mis võimaldaks suuremat tootlust, mitmekesiste dūnaamiliste infoallikate kasutamist, juurdepääsu täpsele ja asjakohasele infole.

Uue tehnoloogia kasutusele võtmisel on esmane ootus olnud tõestada kuluefektiivsust ning tagada loendusstatistika tootmine lühikese aja jooksul. Selleks on tulnud eelkõige parandada statistikatootmise protsessi jõudlust ja luua ajajärgule omaseid uusi teenuseid, pidades silmas tarbija vajadusi.

Läti, Leedu ja Eesti statistikaamet on loendust ettevalmistavatelt meeskondadelt lisaks nõudnud, et loenduseks loodud lahendused oleksid jätkusuutlikud (ülevõetavad) ning toetaksid jooksvat statistika tootmist, optimeerides töömeetodeid ja tehnoloogiat.

Balti seminaride arhiveeritud andmestik on väga mitmekülgne (info on kättesaadav arhiivis). Materjalid on koostanud eri riikide statistikategijad, aga ka muude valdkondade (ehitised, aadressid, registrid jne) eksperdid (info on olemas loendusteemade ettekannetes). Nimetatud materjal pakub loenduse ettevalmistamiseks juhiseid õigusruumi loomise kohta, samuti loenduse korralduse, andmekvaliteedi parandamise jpt teemadel.

Nagu juba mainitud, on iga kord pärast ÜRO väljakuulutatud loendusvooru tulnud alustada loenduse ettevalmistustöödega. Esimene teemakäsitus on alati olnud seotud loenduskulude planeerimisega: kulude arvutamise meetodika on olnud erinev mitte ainult maailmas, vaid ka Balti regioonis, seetõttu ei saa öelda, missuguses Balti riigis on loendusi korraldatud kõige efektiivsemalt, kuid loenduskulude võrdluses on kõige kallim töö andmete kogumine (Measuring ... 2008).

Peale selle on teemade valik iga loendusvooru alguse seminaridel lähtunud Balti loendusjuhtide ühisest eesmärgist – selgitada välja parim meetod järjekordseks loenduseks (põhineb riikide otsustel), analüüsida loendust takistavaid tegureid, õppida vigadest ja teiste riikide kogemustest ning leida strateegiad, mis aitaksid loenduste korraldamisel tekkivaid probleeme lahendada meetoodiliselt ja lihtsalt.

Tabel 1. Balti loenduse seminaride ajalugu

Seminar	Toimumiskoht	Toimumisaeg	Kestus päevades
1.	Palanga, Leedu	12.–16. juuni 1995	5
2.	Jurmala, Läti	19.–23. august 1996	5
3.	Tallinn, Eesti	8.–12. september 1997	5
4.	Trakai, Leedu	7.–11. september 1998	5
5.	Jurmala, Läti	23.–27. august 1999	5
6.	Tallinn, Eesti	10.–12. oktoober 2000	3
7.	Kaunas, Leedu	17.–18. september 2001	2
8.	Riia, Läti	19.–20. september 2002	2
9.	Tallinn, Eesti	25.–26. september 2003	2
10.	Vilnius, Leedu	18.–19. mai 2006	2
11.	Riia, Läti	27.–28. september 2007	2
12.	Tallinn, Eesti	12.–13. november 2008	2
13.	Vilnius, Leedu	2.–3. detsember 2009	2
14.	Riia, Läti	9.–10. detsember 2010	2
15.	Tallinn, Eesti	10.–11. november 2011	2
16.	Riia, Läti	25.–26. november 2012	2
17.	Tallinn, Eesti	13.–14. november 2013	2
18.	Vilnius, Leedu	17.–18. detsember 2014	2
19.	Riia, Läti	30.–31. märts 2016	2
20.	Tallinn, Eesti	26.–27. aprill 2017	2

Allikas: Conference ... 2017.

Suurem osa seminaride materjalidest on seotud loenduse korraldust puudutava ettevalmistustööga eri loendusmetoodikate rakendamiseks. Metoodikad on kompleksed ja seda materjali tuleks kajastada eraldi. Seminari arhiivmaterjaliga tutvudes on võimalik jälgida riikide arengut ja valmidust kasutada nüüdisaegset tehnoloogiat loendustöös, samuti võimekust kasutada teisi andmeallikaid, nt registrite andmeid loendusstatistika tegemisel.

Aastatel 1995–2017 olid seminaridel käsitletud probleemikohad (1995–2017 päevakavade põhjal) üsna sarnased.

Need olid:

- loendusandmete kvaliteet;
- puuduvad andmed;
- registrite ühilduvuse infotehnoloogiline tagamine (identifikaatorid, klassifikaatorid);
- aadressiandmed;
- õiguslik raamistik;
- ankeet;
- andmekogumise tehnoloogiad, sh GIS;
- loendusprogramm;
- eluruumide andmed.

Järgmise ehk 2021. aasta loendusvooru tarbeks EL-is tehakse ettevalmistusi Eestis ja Lätis registripõhiseks loenduseks, kuid väärib märkimist, et huvi registrite teema vastu on olnud juba esimesel seminaril 1995. aastal. Seega on registrite kasutamise teema loendusel olnud aktuaalne pikka aega.

Kuid nii viimast kui ka eelviimast loendust ette valmistades tuli riikidel tõdeda, et Baltimaades ei ole loendusstatistika tegemiseks usaldusväärseid registreid. See järelendus tuleneb seminarideks koostatud andmeallikate (sh registrite) kvaliteedi analüüsist ja ülevaadetest kõigis Balti riikides.

Analüüsivõimekuse suurenedes on tekkinud lisasoovid. Selle kohta mõned näited: 2013. aastal avaldasid loendusjuhid soovi ja valmidust teha rohkem koostööd eri asutuste ja ametkondadega, et vahetada infot riikide loenduskogemuste kohta; 2014. aastal tutvustas Eesti rahvastikuregistri pidaja Riias andmekvaliteedi parandamiseks loodud andmekorrasustööde projekti ja Maa-amet andis ülevaate ADS-süsteemist ning aadresside korrastamisest; 2016. aastal tutvustas Läti Statistikaamet Riia linna aadresside korrastamise projekti; 2017. aasta seminaril andis Eesti Riigi Infosüsteemi Amet Balti seminarist osavõtjatele ülevaate riigi andmevahetuse taristust.

Seminare läbiv teema on olnud ülevaated andmekogumis- ja andmetöötlussüsteemidest perioodil 2000–2012. Materjalidest selgub, et 2000. aasta loenduse andmetöötluste strateegia välja-töötamist alustati 1995. aastal. Andmetöötluste loomisel osalesid paljud firmad, kes pidid tagama loendajate täidetud loenduslehtede skannimise ning seejärel loenduslehtede andmete tõlgenda-mise, kontrollimise ja kinnitamise. Baltimaades otsiti eraldi tarkvara loendusandmete kodeerimiseks, loogilisuse kontrollide ja dubleerivate andmete eemaldamise küsimuste lahenda-miseks. Loendustulemused plaaniti avaldada peale paberväljaande elektrooniliselt (CD-ROM, internet), samuti GIS-i vahenditega ehk kaartidena. Nende riikide ülevaadete puhul kerkib esile asjaolu, et infotehnoloogia areng on pärast 2000. aasta loendust olnud kiire. Platvormid (nt FoxPro 2.5DOS või Oracle 8 Developer 2000) ja süsteemid (nt Microsoft NT4.0 Workstation või Map info 5.5), mis võeti Balti riikides kasutusele perioodil 1998–2003, ei olnud efektiivselt kasutatavad viimast loendust ette valmistades. Uue ehk 2011. aasta loenduse jaoks tuli taristute loomist alustada nullist kõigis kolmes riigis.

Peale selle väärub märkimist, et kogu loendusprotsessis pöörati erilist tähelepanu andmete turvalisusele ja konfidentsiaalsuse tagamisele. Suurimaks riskiks peeti tollel ajal loendus-materjalide käideldavuse häireid, seejuures hävimist, kardeti ka andmebaaside sattumist võõrastesse kättesse. Riskide maandamiseks ja turvameetmete rakendamiseks sõlmiti koostöö-leppeid ettevõtetega erasektorist. 2000. aasta loenduse tegemise ja lõpetamise kohta ei ole seminari materjalides suuremaid andmemetoodiliste tööde ülevaateid ega ka loenduse analüüse. Nähtavasti olid loendusmeeskonnad siis hõivatud loenduse töö korraldamisega andmetöötlemise ja avaldamise teemal.

Pärast loendusandmete avaldamist 2004. aastal oli Balti seminari töös paariaastane paus. Anti välja Balti riikide loendustulemusi kajastav trükis „2000 Round of Population and Housing Censuses in Estonia, Latvia and Lithuania“ (2004). Loendusseminare hakati taas korraldama 2006. aastal, sellest ajast võib leida visioonidokumente koostöövajaduse kohta loendusmeetodi valikul regioonis. Valituks osutus kombineeritud loendusmeetodika. Arhiivmaterjalide põhjal saab esile tõsta süstemaatilise tegevuse alguse registrite kvaliteedi ja loenduseks sobivuse hindamisel kõigis kolme riigis alates aastast 2007.

2008. aastal olid statistikaametid analüüsinud rahva ja eluruumide loenduse näitajate olemasolu riiklikes andmekogudes. Uurimustest järeldati, et registriandmete kvaliteet oli ebapiisav loendusel kasutamiseks, sest andmekogude ja registrite täitumus ei olnud loenduse jaoks vajalike näitajate puhul küllaldane. Kuid registriandmeid oli võimalik kasutada loenduse ettevalmistamiseks ja eeltäitmiseks nii Lätis, Leedus kui ka Eestis.

2008.–2011. aastal toimunud seminaride päevakavad olid pühendatud loenduse õiguslike regulatsioonide teemale, kombineeritud loendusmeetodi kasutamisele, e-loenduse eripärale, loendusandmete mitmele aspektile, pikemale loendusperioodile, loenduse kohustuslikule väljundile jne (Rahvaloenduse ... ; Rahvaloendus ...).

Registripõhise loenduse teema kerkis esile seoses 2020. aasta loendusvooru meetodi valiku kinnitamisega Eestis 2010. aastal. (Vabariigi ... 2010) Eesti registripõhise loenduse ettevalmistustöödest on seminaridel antud ülevaateid alates 2013. aastast. 2017. aasta seminaril anti juba ülevaate esimestest sammudest registripõhise meetodi kasutamisel prooviloendusel Eestis. Balti regioonis tõrkavad silma registriandmete kasutamise probleemid: tegelikud elukohad ei lange kokku registreeritud elukohtadega. See asjaolu on kõige suurem proovikivi registripõhisele loendusele üleminekul regioonis.

Kokkuvõte

Balti loendusseminaride seeria on taasiseseisvunud riikide jaoks olnud regiooni üks olulisemaid statistikakoostöö vorme, mis on aidanud statistikaametitel olla kursis loendusstatistika suundumustega maailmas ja saavutada loendusteks tase, mis vastab ajastuomasele tehnoloogia arengule.

Balti seminaride ajalugu räägib sellest, kuidas eri rollides loenduse tegijad saavad üksteist toetada ja ergutada kogemustest õppima. Üheskoos areneda, enda riigi kogemust teiste riikidega võrrelda, võimalikele probleemidele või keerukatele situatsioonidele lahendusi otsida ja õnnestumiste üle rõõmustada, mis ongi olnud kohtumiste põhiidee.

Üks tähtsamaid teemasid loendusstatistika tegemise kontekstis on olnud registrite temaatika. Balti riigid taotleavad registripõhiselt loenduselt kuluefektiivsust, head andmekvaliteeti ning väljatöötatud lahenduste mitmekülgsemat rakendust tulevikus registripõhise statistika tegemisel.

Vähem tähtsad teemad (nt loogilised kontrollid, mida selles artiklis ei ole käsitletud) on siiski loendusspetsialistide jaoks kogemuste ja info vahetamisel olnud üliolulised, et leida keerulistele küsimustele vastuseid.

Balti riikide loendusseminaride lugu on väärtuslik, sest see räägib tulevastele statistika tegijatele, kuidas taasiseseisvunud Balti riikides valmistati ette ja tehti rahva ja eluruumide loenduseid eri meetodikaid ja tehnoloogiat kasutades. Praegu on selge vajadus teha loenduse teemadel sisutihedat koostööd, Balti seminaril osalejad soovivad järgmise loenduse ettevalmistamisel teha rohkem rahvusvahelist koostööd, milles nähakse võtmerolli hea loendustulemuse saavutamiseks.

Allikad Sources

2000 round of Population and Housing Censuses in Estonia, Latvia and Lithuania. (2004). Vilnius.

Ametkondlik kirjavahetus 1993–1995. Official correspondence 1993–1995.

Balti loendusseminaride päevakavad perioodil 1995–2017. Baltic Census Seminar Programmes 1995–2017.

Conference of European Statisticians. Group of Experts on Population and Housing Censuses. (2017). Twentieth Meeting. Geneva, 3–5 October 2017. Session H: Cooperation models among countries. Baltic Population and Housing Census Seminars as cooperation model to study experiences and to develop future strategies for Censuses in Estonia, Latvia, and Lithuania. Note by Statistics Estonia, Central Statistical Bureau of Latvia and Statistics Lithuania. Prepared by Diana Beltadze, Peteris Vegis, Dalia Ambrozaitiene.

Loendusteemalised ettekanded: Soome, Norra, Rootsi, Taani, Iiri, Hollandi, Luksemburgi, Hispaania, Poola, Sloveenia, Iisrael jt riigid. Presentation on census topics: Finland, Norway, Sweden, Denmark, Ireland, the Netherlands, Luxembourg, Spain, Poland, Slovenia, Israel and other countries.

Main Results of the UNECE-UNSD Survey on the 2010 Round of Population and Housing Censuses. (2009). Conference of European statisticians. Twelfth Meeting. Economic Commission For Europe: Geneva.

Measuring Population and Housing – Practices of UNECE countries in the 2000 round of censuses. (2008). New York and Geneva: United Nations.

Principles and Recommendations for Population and Housing Censuses. (2008). Revision 2. Statistical Papers Series M, No. 67/Rev.2, ST/ESA/STAT/SER.M/67/Rev.2. United Nations Department of Economic and Social Affairs (UNDESA), Statistics Division, New York.

Rahvaloendus 2000. Balti seminari materjalid. Tallinn, 2000. Balti seminari materjalid, Tallinn 1997. Teiste riikide loenduste materjal. [arhiividokumendid] Nr 180.

Rahvaloenduse Balti seminaride materjalid. I seminar. Loendus Palangas 12.–16. juuni 1995. VI seminar Eestis Tallinnas, 10.–12. oktoober 2000. II seminar Lätis Jūrmalas, 19. –23. august 1996. [arhiividokumendid] Nr 181.

Statistikaameti ametkondlik kirjavahetus ministeeriumitega registrite teemal loenduse kontekstis perioodil 2000–2009. Official correspondence between Statistics Estonia and ministries on register topics in the census context in 2000–2009.

Statistikaametite koostatud analüüsid andmeallikate, sh registrite kvaliteedi kohta. Statistics Estonia's analyses on the quality of data sources, incl. registers.

Vabariigi Valitsuse rahvaloenduskomisjoni protokollid 2010. aastast. The minutes of the Government of the Republic Census committee from 2010.

THE BALTIC CENSUS SEMINAR CELEBRATED ITS 20TH ANNIVERSARY IN 2017

Diana Beltadze

The history of the Baltic States Population and Housing Census Seminars is first and foremost a history of cooperation between Estonian, Latvian and Lithuanian statisticians, census leaders and experts. This article presents the more important census-related themes in international cooperation between the statistical offices of the Baltic States.

A Population and Housing Census is usually conducted every ten years and it is the most authoritative source of data in describing demographic change. Its aim is to collect data on population, ethnic composition, distribution and dwellings. After the restoration of independence, Estonia, Latvia and Lithuania have conducted two censuses, the results of which are internationally comparable, because the UN recommendations for census rounds were considered (Principles ... 2008; Main ... 2009). The choice of a census method is free for countries, but the European Union quality requirements have to be adhered to. Therefore, international cooperation is important – it helps to understand options available in organising censuses.

In the next census round, Estonia and Latvia are planning for the first time in the history of the region to conduct a register-based census, i.e. results are obtained by combining information already available in state registers. Lithuania has declared that, considering various circumstances, the next census still has to be organised using a combination of methods, but its exact format is not yet known. It is still early to analyse the readiness of the Baltic States for the next census as preparations are ongoing, but it is possible to give an overview of census-related challenges that the Baltic countries have faced after regaining independence by using the historical materials of the Baltic Census Seminars. In 2017, the Baltic Census Seminar took place for the 20th time.

Baltic Census Seminar as a form of cooperation

The first census seminar of the Baltic countries took place in 1995 in Lithuania. Since then, the seminars have been organised seven times in Estonia and Latvia, and six times in Lithuania. The seminar topics have always been dedicated to the most important tasks in enumeration work (e.g. choice of the census method, assessment of data quality, use of technology), or the use of registers in census statistics has been explained. The topics are agreed upon between different parties, and priorities have been identified in treating the topics. Since 2007, the maximum number of seminar participants and seminar duration have depended on financial capabilities of the statistical offices. Prior to that, the participation numbers were higher and the duration was longer.

The stages of cooperation can be grouped in many ways. They can be identified according to cooperation with different parties, e.g. the era of Nordic support when preparing for the 2000 round of censuses in the Baltic countries. The stages can be divided by themes as the periods preparing for the 2000 and 2011 censuses and the period following the 2011 census. The seminars can also be grouped by the number of participants (Conference ... 2017).

Exactly how the cooperation started and who initiated it cannot be found out anymore, but the need for cooperation in preparing for censuses in the Baltic States is described in several Statistics Estonia's archive documents from 1995. Cooperation proposals on census topics can

be found primarily in the official correspondence (1993–1995) between Statistics Estonia's leading staff Mati Sundja and Lembit Tepp and the census leaders of Latvia and Lithuania.

The first census seminar took place in Lithuania. In the first years, the Nordic countries helped to organise the seminars, and the preparatory work during the 2000 census round took place with their coordination. Later when Estonia, Latvia and Lithuania became European Union Member States, the cooperation network expanded and the census leaders of the Baltic countries managed the organisation of the seminars themselves. Nowadays the Baltic countries are sharing with other countries knowledge about the combined census method, which was successfully used during the last census.

It can be acknowledged that in two decades, cooperation skills and readiness to cooperate have developed. The cooperation over the decades is characterised by improvement of competences, technical and emotional skills of the census teams of the three countries, listening to various parties and valuing the chance to learn from mistakes, etc.

It is customary that for every census a new team of many people has to be hired and trained. Achieving an understanding of census topics is a time-consuming process for every census organisation. Often the less experienced team members have difficulty in separating important information from unimportant information in census-related statistics. In time-critical census work, everything seems of utmost importance. In such difficult moments, information exchange with Latvian and Lithuanian colleagues has been helpful. Thus, the Baltic Census Seminars have become a place for providing census-related supervision not only to organisers but also to the specialists of the enumeration network. Discussing topics and thinking together have contributed to better understanding in organising census work.

Seminar topics

The history of the Baltic Census Seminars is divided into three periods by seminar topics (Table 1). The topics can be differentiated based on the choice of the enumeration method.

- The first census (2000) after the restoration of independence was conducted in all three countries in the traditional way;
- the last census (2011) was conducted in the Baltic countries by using a combined census methodology.

There has been agreement in the Baltic seminars that the statistical offices can offer different methods for collecting census data, but at the same time, the statistics produced by the chosen method have to depict the actual situation as precisely as possible, so that consumers would not lose trust in census results.

Carrying out a census requires, in addition to establishing census organisations, that the statistical offices have created an environment which would enable better capacity to produce results, use of varied dynamic information sources, access to accurate and relevant information.

When adopting new technologies, the first expectation has always been the need to prove cost-efficiency and ensure the production of census statistics in a short time. For that it has been necessary to, first and foremost, improve the efficiency of the process of producing statistics and create new up-to-date services while considering the needs of the consumer.

The statistical offices of Latvia, Lithuania and Estonia have also asked the census preparation teams that the solutions created for censuses would be sustainable (transferable) and would support day-to-day production of statistics by optimizing work methods and technology.

The archived materials on the Baltic Census Seminars are very diverse (the information is available in the archive). The material has been produced by the statisticians of different countries as well as by experts of other areas (housing, addresses, registers, etc.) (the information is available in the presentations on census topics). The aforementioned materials

provide direction for census preparation in terms of legal environment, as well as census organisation, improvement of data quality, etc. The greatest value of the material lies in a detailed overview of information technology developments in census statistics.

As mentioned earlier, each time after the UN has announced a new census round, it has been necessary to start with census preparation work. The first discussion topic has always been related to planning census costs: the methodology for calculating census costs has been different all over the world as well as in the Baltic States, and therefore, it cannot be said in which Baltic country has census organisation been the most cost-effective. However, in comparing census costs, data collection has been most expensive (Measuring ... 2008).

In addition, the choice of topics for the first seminars of each census round have been based on the common goal of the Baltic census leaders – to identify the best method for the upcoming census (based on the decisions of the countries), analyse the hindering factors for the census, learn from mistakes and other countries' experiences and find strategies that would help to solve the problems arising during census organisation in a methodical and simple way.

Table 1. History of the Baltic Census Seminars

Seminar	Location	Date	Duration in days
1	Palanga, Lithuania	12–16 June 1995	5
2	Jurmala, Latvia	19–23 August 1996	5
3	Tallinn, Estonia	8–12 September 1997	5
4	Trakai, Lithuania	7–11 September 1998	5
5	Jurmala, Latvia	23–27 August 1999	5
6	Tallinn, Estonia	10–12 October 2000	3
7	Kaunas, Lithuania	17–18 September 2001	2
8	Riga, Latvia	19–20 September 2002	2
9	Tallinn, Estonia	25–26 September 2003	2
10	Vilnius, Lithuania	18–19 May 2006	2
11	Riga, Latvia	27–28 September 2007	2
12	Tallinn, Estonia	12–13 November 2008	2
13	Vilnius, Lithuania	2–3 December 2009	2
14	Riga, Latvia	9–10 December 2010	2
15	Tallinn, Estonia	10–11 November 2011	2
16	Riga, Latvia	25–26 November 2012	2
17	Tallinn, Estonia	13–14 November 2013	2
18	Vilnius, Lithuania	17–18 December 2014	2
19	Riga, Latvia	30–31 March 2016	2
20	Tallinn, Estonia	26–27 April 2017	2

Source: Conference ... 2017.

The majority of the seminar materials concern census organisation related preparations for applying different census methods. The methodologies are complex, and this material would require a separate discussion. When examining the archive materials on seminars, it is possible to observe the developments in the countries and readiness to use up-to-date technologies in census work, as well as the capacity to use other data sources, e.g. register data in producing census statistics.

The problem areas discussed during the seminars in 1995–2017 have been quite similar (based on the seminar programmes 1995–2017).

These include:

- quality of census data;
- missing data;
- ensuring IT compatibility of registers (identifiers, classifiers);
- address data;

- *legal framework;*
- *questionnaire;*
- *data collection technologies, incl. GIS;*
- *census programme;*
- *data on dwellings.*

For the next, i.e. 2021 census round in the EU, preparations are made in Estonia and Latvia for a register-based census, but it is worth mentioning that there was interest in the topic of registers already during the first seminar in 1995. Therefore, using registers in censuses has been topical for a long time.

However, when preparing for both the last census and the one before the last, it had to be admitted that there are no reliable registers for producing census statistics. This conclusion is based on the analysis and overviews of prepared data sources, incl. registers in all Baltic countries.

As the analysis capabilities have increased, additional wishes have arisen. Some examples about this: in 2013, census leaders expressed the desire and readiness to cooperate more with different institutions and authorities to exchange information on countries' census experiences; in 2014, the Estonian population register introduced in Riga a data organisation project for improving data quality, and the Land Board gave an overview of the Address Data System (ADS) and organising addresses; in 2016, the Central Statistical Bureau of Latvia presented the data organisation project of Riga city; in the 2017 seminar, the Estonian Information System Authority gave the seminar participants an overview of the country's data exchange infrastructure.

One of the topics throughout the seminars has been overviews of data collection and processing systems in 2000–2012. It is evident in the materials that the development of the data processing strategy for the 2000 census was started in 1995. Many companies participated in the development of data processing, and they had to ensure the scanning of census questionnaires filled in by the enumerators, and consequently the interpretation, checking and approving of the data of the questionnaires. In the Baltic countries, there was an effort to find a separate software for coding census data, and for solving the issues of logic checks and removal of double data. It was planned to disseminate the census results in addition to a paper publication also electronically (CD-ROM, internet), as well as by GIS solutions, i.e. maps. It becomes apparent from these country overviews that information technology development has been rapid since the 2000 census. Platforms (e.g. FoxPro 2.5DOS or Oracle 8 Developer 2000) and systems (e.g. Microsoft NT4.0 Workstation or Map info 5.5) introduced in the Baltic countries in 1998–2003 were no longer efficiently usable during the preparation for the most recent census. For the new 2011 census, it was necessary to create an infrastructure from scratch in all three countries.

In addition, it is worth mentioning that during the whole census process special attention was paid to data security and ensuring confidentiality. The greatest risk at the time was considered to be issues in processing census materials, including their loss; it was also feared that wrong people might get hold of the databases. In order to manage the risks and apply security measures, cooperation agreements were concluded with private sector enterprises. Regarding carrying out and finishing the 2000 census, the seminar materials do not include major overviews of data methodology work or census analyses. Probably, at the time, the census teams were engaged in organising data processing and dissemination for the census.

After the publishing of census data in 2004, the Baltic Census Seminars were put on hold for a couple of years. A publication titled "2000 round of Population and Housing Censuses in Estonia, Latvia and Lithuania" (2003) was published on census results of the Baltic States. Census seminars were organised again starting from 2006; vision documents can be found from this time on cooperation needs for the selection of the census method in the region. The combined census method was chosen. Based on archive materials, the beginning of systematic activities towards quality and suitability assessment of registers could be observed in all three Baltic countries since 2007.

In 2008, the statistical offices had analysed the existence of the Population and Housing Census indicators in state databases. The studies revealed that the quality of register data was inadequate for census use, as the databases and registers did not have enough data on the indicators required for the census. However, the data of registers could be used for preparing for census and for pre-filling in Latvia, Lithuania as well as in Estonia.

The programmes of the seminars in 2008–2011 were concerned with census-related legal regulations, the use of the combined census method, the special character of e-census, multiple aspects of the census questionnaire, longer enumeration period, compulsory output of the census, etc. (Rahvaloenduse ... ; Rahvaloendus ...).

The issue of register-based census arose in Estonia in 2010 in connection with approving the census method for the 2020 census round (Vabariigi ... 2010). Overviews of the preparations of Estonia's register-based census have been presented at the seminars since 2013. During the 2017 seminar, an overview was already given on the first steps towards using a register-based census method during the first trial census in Estonia. In the Baltic region, there are noticeable problems for the use of register data: the actual residences do not coincide with the registered ones. This aspect is the greatest challenge for transferring to a register-based census in the region.

Conclusion

After the restoration of independence, the Baltic Census Seminars have been one of the most important forms of cooperation in statistics for the Baltic countries, which have helped the national statistical offices to be aware of the developments in census statistics around the world and achieve the level for censuses which corresponds to current technological developments.

The history of the Baltic Census Seminars shows how census organisers with different roles can support each other and encourage to learn from experience. The core idea of these meetings has been to develop together, compare the experiences of involved countries, find solutions to potential problems or complex situations and be happy for successes.

One of the more important topics in the context of producing census-based statistics has been registers. The Baltic countries are expecting from the register-based census cost-efficiency, good data quality and more diverse use of developed solutions in producing register-based statistics in the future. Less important topics (e.g. logic checks which have not been discussed in this article) have nevertheless been essential for enumeration specialists while exchanging experiences and information, in order to find solutions to complicated questions.

The story of the Baltic Census Seminars is valuable because it tells the future statistics producers how population and housing censuses were prepared and conducted in re-independent Baltic States by using different methods and technology. Currently, there is an obvious need for substantial cooperation on census topics; the participants of the Baltic Census Seminars hope to cooperate more while preparing for the next census, and this is seen as a key to achieving great census results.

ÜLEVAADE RIIGI STATISTIKA KESKBÜROO VÄLJAANNETEST OVERVIEW OF PUBLICATIONS OF THE STATE STATISTICAL CENTRAL BUREAU

Mihkel Servinski, Sille Tiitsmaa

Sel korral on ajaloolistes nopetes nimekiri Riigi Statistika Keskbüroo väljaannetest (1918–1940). Kui kellelgi peaks tekkima vajadus varasemate aastate statistika järele, siis võib järgnev nimekiri kiirendada andmete leidmist. Praegu on statistilised andmed andmebaasides, pilvedes ja mujal e-avarustes. Eesti esimese iseseisvusperioodi statistika on enamjaolt kättesaadav raamatutest. Elektrooniliselt on niisugune statistika olemas vaid juhul, kui see on digitaliseeritud. Seda on tehtud, kuid mitte kõikide Riigi Statistika Keskbüroo väljaannete puhul. Arvestama peab, et kogu esimesel iseseisvusperioodil tehtud statistikat ei avaldatud Riigi Statistika Keskbüroo väljaannetes, aga selliste väljaannete nimekirja pole autorite teada keegi kokku pannud. Näiteks olgu raamatud „Tartu I. Linnavalitsuse tegevuse ülevaade 1919–1930. Statistika 1919–1930. Korteriolud 1927“ ja „Eesti majandus. Eesti panga aastaraamat 1935 a“.

Eesti Statistika Keskbüroo väljaanded võib jagada kaheks: kuukirjad^a ja muud statistikaväljaanded. Perioodil 1922–1940 ilmus 223 kuukirja. Muid väljaandeid oli kokku 89. Loetelu ei pruugi olla täielik.

Kõige süsteemsemalt avaldati Eesti esimesel iseseisvusperioodil Riigi Statistika Keskbüroos raamatuid väliskaubanduse ja põllumajanduse kohta. Küllap peegeldab see seda, milliste teemade statistika vastu tollane Eesti ühiskond enim huvi tundis. Vaadeldava perioodi väljaannetest tuleb kindlasti eraldi esile tuua 1922. aasta üleriigilise rahvaloenduse andmete avaldamine neljateistkümnes raamatus. Rahvaloenduse andmeid on sellises mahus raamatutes avaldatud hilisematel aastatel vaid 2000. aasta rahvaloenduse puhul ja tõenäoliselt hõlmatavas tuleviks seda enam ei tehta.

These picks from the history of Estonian statistics include a list of publications (1918–1940) of the State Statistical Central Bureau. If anyone needs to access statistics of prior years, this list might facilitate locating the data. Currently, statistical data are in databases, clouds and other electronic environments. The statistics of the first independence period of Estonia are mainly available in books. These statistics exist electronically only if they have been digitalized. This has been done to some extent, but not for all the publications of the State Statistical Central Bureau. It should be taken into account that not all statistics produced during the first independence period were published in the publications of the State Statistical Central Bureau, but the authors are unaware whether such a list of publications has ever been compiled. An example of this are the books “Tartu I. Linnavalitsuse tegevuse ülevaade 1919–1930. Statistika 1919–1930. Korteriolud 1927” on activities of Tartu City Government and apartment conditions in 1927 and “Eesti majandus. Eesti panga aastaraamat 1935 a”, the 1935 yearbook of Eesti Pank.

The publications of the State Statistical Central Bureau of Estonia can be divided into two: monthly bulletins^b and other statistical publications. During the period 1922–1940, 223 monthly bulletins were published. Other statistical publications numbered 22. The list might not be complete.

During Estonia’s first independence period, books on foreign trade and agriculture were published most systematically by the State Statistical Central Bureau. It likely reflects on which statistical domains were of most interest to the Estonian society at the time. Of the publications of this period, it is definitely worth highlighting the publishing of the data of the 1922 Population Census in fourteen books. In later years, census data has been published in print in such volume only in the case of the 2000 Population and Housing Census, and probably in the foreseeable future this will not happen again.

^a Kõik esimesel iseseisvusperioodil ilmunud kuukirjad on digiteerituna olemas aadressil <http://www.digar.ee/arhiiv/et/periodika?id=2628>.

^b All monthly bulletins published during the first independence period are available digitally at <http://www.digar.ee/arhiiv/et/periodika?id=2628> (in Estonian, French).

Eesti Statistika Keskbüroo väljaanded, 1918–1940

Publications of the State Statistical Central Bureau, 1918–1940

Nr No	Pealkiri Title	Aasta Year	Sari Series	Digitaalselt olemas Available digitally
VÄLISKAUBANDUS FOREIGN TRADE				
1	Väliskaubandus 1921 ja 1922 a.	1923		
2	Väliskaubandus 1923 a.	1924	Eesti majandus. Vihik I	
3	Väliskaubandus 1924 a.	1925	Eesti majandus. Vihik IV	
4	Väliskaubandus 1925 a.	1926	Eesti majandus. Vihik VII	
5	Väliskaubandus 1926 a.	1927	Eesti majandus. Vihik VIII	
6	Väliskaubandus 1927 a.	1928	Eesti majandus. Vihik X	
7	Eesti väliskaubandus 1928 a.	1929	Eesti majandus. Vihik XI	
8	Väliskaubandus 1929 a.	1930	Eesti majandus. Vihik XII	
9	Väliskaubandus 1930 a.	1931	Eesti majandus. Vihik XIII	
10	Väliskaubandus 1931 a.	1932	Eesti majandus. Vihik XIV	
11	Väliskaubandus 1932 a.	1933	Eesti majandus. Vihik XV	
12	Väliskaubandus 1933 a.	1934	Eesti majandus. Vihik XVI	
13	Väliskaubandus 1934 a.	1935	Eesti majandus. Vihik XVII	
14	Väliskaubandus 1935 a.	1936	Eesti majandus. Vihik XVIII	
15	Väliskaubandus 1936 a.	1937	Eesti majandus. Vihik XIX	
16	Väliskaubandus 1937 a.	1938	Eesti majandus. Vihik XX	
17	Väliskaubandus 1938 a.	1939	Eesti majandus. Vihik XXI	
RAHVASTIK POPULATION				
1922. aasta üldrahvalugemine 1922 census				
18	Eesti Vabariigi jaotusüksused 1922 a. rahvaloendusel	1923		
19	Üldrahvalugemise eelkõikuvõtted	1923		
20	1922 a. üldrahvalugemise andmete läbitöötamise plaan ja tabelite sisu	1923		
21	Rahva demograafiline koosseis ja korteriolud Eestis	1924	1922 a. üldrahvalugemise andmed. Vihik I	x
22	Üleriikline kokkuvõte : tabelid	1924	1922 a. üldrahvalugemise andmed. Vihik II	x
23	Rahva tööala ja ühiskondline kihitus	1925	1922 a. üldrahvalugemise andmed. Vihik III	x
24	Viru maakond	1924	1922 a. üldrahvalugemise andmed. Vihik IV	x
25	Järva maakond	1923	1922 a. üldrahvalugemise andmed. Vihik V	x
26	Harju maakond ja Tallinna linn	1924	1922 a. üldrahvalugemise andmed. Vihik VI	x
27	Lääne maakond	1924	1922 a. üldrahvalugemise andmed. Vihik VII-a	x
28	Saare maakond	1924	1922 a. üldrahvalugemise andmed. Vihik VII-b	x
29	Pärnu maakond	1923	1922 a. üldrahvalugemise andmed. Vihik VIII	x
30	Viljandi maakond	1924	1922 a. üldrahvalugemise andmed. Vihik IX.	x
31	Tartu ja Valga maakonnad	1924	1922 a. üldrahvalugemise andmed. Vihik X	x
32	Võru ja Petseri maakonnad	1927	1922 a. üldrahvalugemise andmed. Vihik XI	x
1934. aasta üldrahvalugemine 1934 Census				
33	Valdade rahvastik	1934	1. III 1934 rahvaloenduse andmed. Vihik I	x

Eesti Statistika Keskbüroo väljaanded, 1918–1940

Publications of the State Statistical Central Bureau, 1918–1940

Järg – Cont.

Nr No	Pealkiri Title	Aasta Year	Sari Series	Digitaalselt olemas Digitally available
34	Rahvastiku koostis ja korteriolud	1935	1. III 1934 rahvaloenduse andmed. Vihk II	x
35	Tööhärad ja leibkonnad	1935	1. III 1934 rahvaloenduse andmed. Vihk III	x
36	Rahvastikuprobleeme Eestis	1937	II rahvaloenduse tulemusi. Vihk IV	x
37	Eesti arvudes : 1934. a. rahvaloenduse mälestuseks PÖLLUMAJANDUS AGRICULTURE Aastaraamatud Yearbooks	1934		
38	Eesti põllumajandus: statistiline aastaraamat I	1923	Lühendatud väljaanne riigi statistika keskbüroo põllumajandusliste korrespondentide jaoks.	
39	Eesti põllumajandus. Vihk II	1924	Eesti põllumajandus. Vihk II	
40	Loomapidamine Eestis 1924	1925	Eesti põllumajandus. Vihk III	
41	Eesti põllumajandus 1925 : statistiline aastaraamat IV	1926	Eesti põllumajandus. Vihk IV	
42	Eesti põllumajandus 1926 : statistiline aastaraamat V	1927	Eesti põllumajandus. Vihk V	
43	Eesti põllumajandus 1927 : statistiline aastaraamat VI	1928	Eesti põllumajandus. Vihk VI	
44	Eesti põllumajandus 1928 : statistiline aastaraamat VII	1929	Eesti põllumajandus. Vihk VII	
45	Eesti põllumajandus 1929 : statistiline aastaraamat VIII	1930	Eesti põllumajandus. Vihk VIII	
46	Eesti põllumajandus 1930 : statistiline aastaraamat IX	1931	Eesti põllumajandus. Vihk IX	
47	Eesti põllumajandus 1931 : statistiline aastaraamat X	1932	Eesti põllumajandus. Vihk X	
48	Eesti põllumajandus 1932 : statistiline aastaraamat XI	1933	Eesti põllumajandus. Vihk XI	
49	Eesti põllumajandus 1933 : statistiline aastaraamat XII	1934	Eesti põllumajandus. Vihk XII	
50	Eesti põllumajandus 1934 : statistiline aastaraamat XIII	1935	Eesti põllumajandus. Vihk XIII	
51	Eesti põllumajandus 1935 : statistiline aastaraamat XIV	1936	Eesti põllumajandus. Vihk XIV	
52	Eesti põllumajandus 1936 : statistiline aastaraamat XV	1937	Eesti põllumajandus. Vihk XV	
53	Eesti põllumajandus 1937 : statistiline aastaraamat XVI	1938	Eesti põllumajandus. Vihk XVI	
54	Eesti põllumajandus 1938 : statistiline aastaraamat XVII	1939	Eesti põllumajandus. Vihk XVIII	
55	Eesti põllumajandus 1939 : statistiline aastaraamat XVIII Põllumajandusloendused ja muu põllumajandusstatistika <i>Agricultural censuses and other agricultural statistics</i>	1940	Eesti põllumajandus. Vihk XVIII	
56	Asumaa majapidamised Eestis 1919–1923	1924		
57	1925 a. põllumajandusliku üleskirjutuse andmed	1926		
58	1929 a. põllumajandusliku üleskirjutuse andmed. Vihk I	1930		
59	1930 a. põllumajandusliku üleskirjutuse andmed. Vihk II	1932		
60	Põllumajandusloenduse juhised	1939		

Eesti Statistika Keskbüroo väljaanded, 1918–1940
Publications of the State Statistical Central Bureau, 1918–1940

Järg – *Cont.*

Nr <i>No</i>	Pealkiri <i>Title</i>	Aasta <i>Year</i>	Sari <i>Series</i>	Digitaalselt olemas <i>Digitally available</i>
61	Talundite arv, rahvastik, maapidamine, maakasutus, põlluviljade kasvupindalad, viljapuud, marjapõõsad, kõõgiviljad ja loomad	1940	III põllumajandusloendus. Vihik I	
SOTSIAALSTATISTIKA <i>SOCIAL STATISTICS</i>				
63	Sugemed vigaste laste kohta Eestis	1923		
64	II Riigikogu valimised	1923		
65	III Riigikogu valimised	1927		
66	IV Riigikogu valimised	1929		
67	V Riigikogu valimised	1933		
68	Haridus Eestis : lühendatud väljaanne (tekst ja diagrammid)	1924	Eesti demograafia. Vihik I	
69	Haridus Eestis	1924	Eesti demograafia. Vihik I	x
70	Sündivus, surevus, abielluvus ja rahvalikumine 1921–1923	1925	Eesti demograafia. Vihik II	
71	Tevishoid Eestis	1925	Eesti demograafia. Vihik III	
72	Rahvastik ja tervishoid Eestis	1930	Eesti demograafia. Vihik IV	
73	Surmapõhjuste registreerimiskord ja rahvusvaheline nomenklatuur	1931		
74	Sotsiaalkindlustus Eestis	1937	Eesti demograafia. Vihik V	
MAJANDUS <i>ECONOMY</i>				
75	Eesti sadamad, veeteed, kaubalaevastik ja laevasõit 1923 ja 1924 a.	1925	Eesti majandus. Vihik II	
76	Raudtee, post, telegraaf ja telefon Eestis	1925	Eesti majandus. Vihik III	
77	Tööstus Eestis 1924	1926	Eesti majandus. Vihik V	
78	Riigiraudteede tegevus 1924 a.	1926	Eesti majandus. Vihik VI	
79	Aksiaseltside, osühistute, ja ühisteg.asutuste. Äriseisund 1. I 1926	1927	Eesti majandus. Vihik IX	
80	Tööstus	1939	1937. a. majandusloenduse andmed. Vihik I	
81	Käsitööstus	1939	1937. a. majandusloenduse andmed. Vihik II	
82	Kaubandus ja transport	1939	1937. a. majandusloenduse andmed. Vihik III	
83	Büdzetiuurimus	1940	1937. a. majandusloenduse andmed. Vihik IV	
MUUD ÜLDVÄLJAANDED <i>OTHER GENERAL PUBLICATIONS</i>				
84	Eesti statistiline album. Vihik I. Maa ja rahvas	1925		x
85	Eesti statistiline album. Vihik II. Majandus	1926		x
86	Eesti statistiline album. Vihik III. Põllumajandus	1928		x
87	Eesti arvudes 1920–1930. Arvuline ülevaade (ilukõites)	1931		x
88	Eesti arvudes 1920–1935	1937		x
89	Eesti statistika: kuukirjad 1922–1940. Nr. 1–223	1940		x

PÕHINÄITAJAD, 2012–2017 MAIN INDICATORS, 2012–2017

Tabel 1. Põhinäitajad aastate ja kvartalite kaupa, 2012–2017
Table 1. Main indicators by years and quarters, 2012–2017

Period	Keskmine brutokuupalk, eurot ^a	Keskmise brutokuupalga muutus eelmise aasta sama perioodiga võrreldes, % ^a	Keskmine vanaduspension kuus, eurot ^b	Hõivatud ^c tuhat	Töötud ^c
<i>Period</i>	<i>Average monthly gross wages and salaries, euros^a</i>	<i>Change of average monthly gross wages and salaries over corresponding period of previous year, %^a</i>	<i>Average monthly old-age pension, euros^b</i>	<i>Employed^c</i> thousands	<i>Unemployed^c</i>
2012	887	5,7	312,9	614,9	68,5
2013	949	7,0	327,4	621,3	58,7
2014	1 005	5,9	345,1	624,8	49,6
2015	1 065	6,0	365,6	640,9	42,3
2016	1 146	7,6	386,0	644,6	46,7
2012					
I kvartal	847	6,9	303,4	604,5	77,4
II kvartal	900	5,0	316,2	614,0	68,8
III kvartal	855	5,7	316,1	625,8	65,9
IV kvartal	916	5,9	315,9	615,4	61,9
2013					
I kvartal	900	6,3	315,9	610,1	67,5
II kvartal	976	8,5	331,3	632,1	55,0
III kvartal	930	8,8	331,4	627,1	53,3
IV kvartal	986	7,6	331,0	616,1	58,9
2014					
I kvartal	966	7,3	330,9	605,8	56,6
II kvartal	1 023	4,8	349,9	629,5	47,7
III kvartal	977	5,0	350,0	633,7	51,3
IV kvartal	1 039	5,3	349,6	630,3	42,7
2015					
I kvartal	1 010	4,5	349,5	623,1	44,2
II kvartal	1 082	5,8	371,3	640,1	44,4
III kvartal	1 045	6,9	370,9	661,0	36,5
IV kvartal	1 105	6,4	370,7	639,4	43,9
2016					
I kvartal	1 091	8,1	370,6	630,0	43,6
II kvartal	1 163	7,6	391,4	657,0	45,3
III kvartal	1 119	7,1	390,2	653,3	52,9
IV kvartal	1 182	7,0	390,3	638,2	45,1
2017					
I kvartal	1 153	5,7	390,7	646,8	38,4
II kvartal	1 242	6,8	409,9	653,5	49,0
III kvartal	1 201	7,4	409,1	666,6	36,5

^a 1999. aastast ei hõlma keskmine brutokuupalk ravikindlustushüvitist.

^b Sotsiaalkindlustusameti andmed.

^c 15–74-aastased.

^a Since 1999, the average monthly gross wages and salaries do not include health insurance benefits.

^b Data of the Social Insurance Board.

^c Population aged 15–74.

Tööjõus osalemise määr ^a	Tööhõive määr ^a	Töötuse määr ^a	Tarbijahinna indeks	Tööstustoodangu tootjahinnaindeks	Period
	%		muutus eelmise aasta sama perioodiga võrreldes, %		
<i>Labour force participation rate^a</i>	<i>Employment rate^a</i>	<i>Unemployment rate^a</i>	<i>Consumer price index</i>	<i>Producer price index of industrial output</i>	<i>Period</i>
	%		change over corresponding period of previous year, %		
67,6	60,8	10,0	3,9	2,3	2012
68,0	62,1	8,6	2,8	4,1	2013
68,0	63,0	7,4	-0,1	-1,6	2014
69,4	65,2	6,2	-0,5	-2,0	2015
70,4	65,6	6,8	0,1	-0,7	2016
					2012
67,5	59,8	11,3	4,4	3,3	1st quarter
67,5	60,7	10,1	3,9	2,0	2nd quarter
68,4	61,9	9,5	3,7	1,9	3rd quarter
67,0	60,9	9,1	3,7	2,1	4th quarter
					2013
67,7	61,0	10,0	3,5	4,6	1st quarter
68,7	63,2	8,0	3,4	4,7	2nd quarter
68,0	62,7	7,8	2,8	3,9	3rd quarter
67,5	61,6	8,7	1,5	3,3	4th quarter
					2014
66,8	61,1	8,5	0,6	-1,2	1st quarter
68,3	63,5	7,0	0,0	-2,0	2nd quarter
69,1	63,9	7,5	-0,6	-1,1	3rd quarter
67,9	63,6	6,3	-0,5	-2,0	4th quarter
					2015
67,8	63,3	6,6	-0,9	-1,6	1st quarter
69,6	65,1	6,5	0,0	-1,7	2nd quarter
70,9	67,2	5,2	-0,5	-2,7	3rd quarter
69,5	65,0	6,4	-0,5	-2,1	4th quarter
					2016
68,6	64,1	6,5	-0,4	-1,4	1st quarter
71,5	66,9	6,5	-0,7	-1,6	2nd quarter
71,9	66,5	7,5	0,4	-1,1	3rd quarter
69,6	65,0	6,6	1,3	1,5	4th quarter
					2017
70,2	66,3	5,6	3,0	2,8	1st quarter
72,0	66,9	7,0	3,1	3,7	2nd quarter
72,0	68,3	5,2	3,7	4,4	3rd quarter

^a 15–74-aastased.

^a Population aged 15–74.

Tabel 1. Põhinäitajad aastate ja kvartalite kaupa, 2012–2017
Table 1. Main indicators by years and quarters, 2012–2017

Period	Tööstus- toodangu mahuindeks ^a	Elektrienergia toodangu mahuindeks ^a	Eksporti- hinnaindeks	Impordi- hinnaindeks	Ehitushinna- indeks	Ehitusmahu- indeks ^b
muutus eelmise aasta sama perioodiga võrreldes, %						
Period	Volume index of industrial production ^a	Volume index of electricity production ^a	Export price index	Import price index	Construction price index	Construction volume index ^b
change over corresponding period of previous year, %						
2012	1,1	-7,0	1,8	4,0	4,6	16,6
2013	4,1	10,9	-1,1	-1,6	5,2	-0,1
2014	3,9	-6,3	-2,6	-2,2	0,5	-2,1
2015	0,3	-16,8	-3,9	-3,8	0,5	-4,5
2016	2,4	16,8	-0,5	-2,3	-0,8	2,6
2012						
I kvartal	2,2	-17,3	5,0	5,8	5,0	25,8
II kvartal	1,2	-8,0	1,4	4,4	4,7	27,9
III kvartal	-1,3	-3,1	0,5	3,0	5,0	12,7
IV kvartal	2,3	1,9	0,4	2,9	3,7	6,8
2013						
I kvartal	3,8	21,7	-0,8	-0,1	5,6	0,8
II kvartal	5,4	16,0	-0,9	-2,6	5,2	-0,4
III kvartal	5,1	14,7	-1,2	-2,1	5,3	3,6
IV kvartal	2,1	-4,7	-1,7	-1,5	4,7	-4,7
2014						
I kvartal	1,6	-19,2	-2,3	-2,4	2,3	-2,9
II kvartal	2,6	-2,4	-2,2	-1,7	0,8	-3,5
III kvartal	4,8	-7,0	-2,2	-1,1	-0,2	-7,4
IV kvartal	6,7	2,7	-3,7	-3,6	-0,7	6,5
2015						
I kvartal	3,5	-0,1	-4,3	-4,7	0,1	-1,2
II kvartal	1,3	-24,0	-3,3	-1,9	0,7	-4,2
III kvartal	-1,2	-22,2	-4,5	-4,3	0,6	-2,7
IV kvartal	-2,2	-20,5	-3,6	-4,2	0,7	-5,0
2016						
I kvartal	-2,6	-5,6	-3,0	-4,0	-0,7	3,4
II kvartal	0,5	4,1	-2,4	-4,5	-1,3	4,2
III kvartal	4,0	41,8	-0,1	-2,2	-0,7	0,5
IV kvartal	7,6	32,9	3,6	1,7	-0,5	2,9
2017						
I kvartal	13,0	31,0	6,7	6,6	0,7	20,5
II kvartal	10,3	42,0	5,7	4,3	1,5	17,6
III kvartal	3,6	-13,1	5,3	3,7	1,7	19,8

^a Andmed põhinevad lühiajastatistikal. 2016.–2017. aasta andmeid võidakse korrigeerida.

^b Ehitustööd Eestis ja välisriikides, 2016.–2017. aasta andmeid võidakse korrigeerida.

Tööstustoodangu mahuindeksi ja ehitusmahuindeksi puhul statistika Eesti majanduse tegevusalade klassifikaatori EMTAK 2008 järgi.

^a Short-term statistics. The data for 2016–2017 may be revised.

^b Construction activities in Estonia and in foreign countries. The data for 2016–2017 may be revised.

In the case of volume index of industrial production and construction volume index, statistics according to the Estonian Classification of Economic Activities (EMTAK 2008, based on NACE Rev. 2).

Järg – Cont.

Põllumajandus- saaduste tootjahinna- indeks	Põllumajandus- saaduste toot- mise vahendite ostuhinnaindeks	Sisemajanduse koguprodukt (SKP) aheldamise meetodil ^a	Jooksevkonto osatähtsus SKP-s, % ^b	Ettevõtete müügitulu, miljonit eurot, jooksev- hindades ^c	Periood
<i>Agricultural output price index</i>	<i>Agricultural input price index</i>	<i>Gross domestic product (GDP) by chain-linking method^a</i>	<i>Balance of current account as percentage of GDP, %^b</i>	<i>Net sales of enterprises, million euros, current prices^c</i>	<i>Period</i>
change over corresponding period of previous year, %					
1,4	4,0	4,3	-1,9	46 262,7	2012
6,7	3,0	1,9	0,5	50 357,2	2013
-5,7	-2,3	2,9	0,3	50 328,6	2014
-13,0	-0,8	1,7	2,0	49 065,8	2015
-2,5	-1,9	2,1	1,9	50 194,5	2016
2012					
4,1	3,2	5,1	-1,0	10 624,9	1st quarter
-5,8	2,7	5,0	-3,8	11 684,7	2nd quarter
-2,9	4,4	3,5	-3,1	11 821,2	3rd quarter
7,4	5,7	3,7	0,1	12 131,9	4th quarter
2013					
12,9	5,5	3,1	-0,6	12 054,1	1st quarter
27,4	4,8	1,0	0,9	12 733,1	2nd quarter
14,5	2,2	1,6	0,1	12 808,7	3rd quarter
-12,4	-0,4	2,1	1,6	12 761,3	4th quarter
2014					
4,0	-2,7	1,8	-3,8	11 798,0	1st quarter
-4,5	-2,8	3,0	0,8	12 869,6	2nd quarter
-10,0	-2,1	2,5	0,7	12 666,7	3rd quarter
-9,8	-1,4	4,1	2,9	12 994,3	4th quarter
2015					
-23,4	-1,1	1,5	-1,5	11 531,1	1st quarter
-18,6	-0,4	2,3	3,9	12 475,7	2nd quarter
-8,9	1,0	2,1	3,2	12 359,5	3rd quarter
-4,1	-0,7	0,9	2,0	12 699,5	4th quarter
2016					
-3,3	-1,0	2,2	-2,2	11 726,0	1st quarter
-7,7	-2,4	0,9	2,5	12 651,7	2nd quarter
-5,3	-2,3	2,0	5,5	12 619,2	3rd quarter
3,0	-1,8	3,1	1,4	13 197,6	4th quarter
2017					
21,6	0,1	4,6	1,4	12 686,9	1st quarter
28,5	1,3	5,7	2,9	13 969,9	2nd quarter
28,8	2,2	4,2	4,2	13 898,9	3rd quarter

^a Referentsaasta 2010 järgi. 2013. aasta I kvartali – 2017. aasta I kvartali andmed on korrigeeritud.^b Eesti Panga andmed. 2013. aasta I kvartali – 2017. aasta I kvartali andmed on korrigeeritud.^c Andmed põhinevad lühiajastatistikal. Statistika Eesti majanduse tegevusalade klassifikaatori EMTAK 2008 järgi.^a Reference year 2010. The data for 1st quarter 2013 – 1st quarter 2017 have been revised.^b Data of Eesti Pank. The data for 1st quarter 2013 – 1st quarter 2017 have been revised.^c Short-term statistics. Statistics according to the Estonian Classification of Economic Activities (EMTAK 2008, based on NACE Rev. 2).

Tabel 1. Põhinäitajad aastate ja kvartalite kaupa, 2012–2017

Table 1. Main indicators by years and quarters, 2012–2017

Period	Riigieelarve tulud ^a	Riigieelarve kulud ^a	Riigieelarve tulude ülekaal kuludest ^a	Eksport ^b	Import ^b	Kaubavahtuse bilanss ^b
	miljonit eurot, jooksevhindades					
Period	Revenue of state budget ^a	Expenditure of state budget ^a	Surplus of state budget ^a	Exports ^b	Imports ^b	Balance of trade ^b
	million euros, current prices					
2012	6 427,2	6 567,2	-140,0	12 521,1	14 096,5	-1 575,4
2013	6 556,2	6 853,0	-296,9	12 288,2	13 902,5	-1 614,4
2014	6 677,5	6 488,4	189,1	12 006,0	13 788,1	-1 782,0
2015	6 792,7	7 157,3	-364,6	11 570,5	13 104,7	-1 534,2
2016	7 318,8	7 326,8	-8,0	11 897,2	13 492,9	-1 595,8
2012						
I kvartal	1 519,9	1 472,7	47,2	2 996,9	3 340,9	-344,0
II kvartal	1 602,4	1 500,1	102,3	3 083,8	3 520,0	-436,2
III kvartal	1 484,8	1 767,5	-282,7	3 295,5	3 618,9	-323,4
IV kvartal	1 820,1	1 826,9	-6,8	3 144,9	3 616,6	-471,7
2013						
I kvartal	1 395,0	1 490,3	-95,3	3 098,1	3 405,8	-307,7
II kvartal	1 862,9	1 593,7	269,2	3 173,3	3 611,9	-438,6
III kvartal	1 697,3	1 763,3	-66,1	2 977,4	3 431,1	-453,7
IV kvartal	1 601,0	2 005,7	-404,7	3 039,4	3 453,7	-414,3
2014						
I kvartal	1 565,0	1 506,8	58,2	2 837,8	3 276,0	-438,2
II kvartal	1 730,4	1 537,0	193,4	3 005,3	3 492,8	-487,5
III kvartal	1 591,6	1 546,6	45,0	3 042,7	3 470,4	-427,7
IV kvartal	1 790,5	1 898,0	-107,5	3 120,3	3 549,0	-428,7
2015						
I kvartal	1 601,1	1 810,7	-209,6	2 832,7	3 187,3	-354,6
II kvartal	1 739,1	1 692,6	46,5	2 990,6	3 339,9	-349,4
III kvartal	1 676,3	1 709,8	-33,5	2 831,6	3 261,8	-430,3
IV kvartal	1 776,2	1 944,2	-168,0	2 920,6	3 307,6	-387,2
2016						
I kvartal	1 850,6	1 874,8	-24,2	2 778,8	3 229,1	-450,5
II kvartal	1 844,1	1 643,7	200,4	3 025,9	3 491,0	-465,0
III kvartal	1 742,3	1 775,1	-32,8	3 017,8	3 319,5	-301,7
IV kvartal	1 881,8	2 033,1	-151,3	3 069,5	3 482,2	-412,7
2017						
I kvartal	1 722,8	1 766,4	-43,6	3 096,6	3 738,1	-641,5
II kvartal	1 963,0	1 926,8	36,3	3 269,5	3 708,5	-439,0
III kvartal	1 876,6	1 835,5	41,1	3 144,9	3 535,1	-390,1

^a Rahandusministeeriumi andmed. Alates 2017. aastast on riigieelarve täitmise arvestus tekkepõhine.^b Jooksva aasta andmeid täpsustatakse iga kuu, eelmiste aastate andmeid kaks korda aastas.^a Data of the Ministry of Finance. Since 2017, the accounting of state budget execution is accrual-based.^b Data for the current year are revised monthly; data for the previous years are revised twice a year.

Järg – Cont.

Jaemüügi mahuindeksi muutus eelmise aasta sama perioodiga võrreldes, % ^a	Sõitjatevedu, tuhat sõitjat ^b	Kaubavedu, tuhat tonni ^b	Lihatoodang (eluskaalus) ^c	Piima- toodang ^c	Muna- toodang ^c	Period
muutus eelmise aasta sama perioodiga võrreldes, %						
<i>Change of retail sales volume index over corresponding period of pre- vious year, %^a</i>	<i>Carriage of passengers, thousands^b</i>	<i>Carriage of goods, thousand Tonnes^b</i>	<i>Production of meat (live weight)^c</i>	<i>Production of milk^c</i>	<i>Production of eggs^c</i>	<i>Period</i>
<i>change over corresponding period of previous year, %</i>						
8	200 746,5	78 142	-2,4	4,1	-2,3	2012
6	216 040,5	78 726	1,4	7,0	5,8	2013
7	211 015,1	75 141	1,2	4,3	5,0	2014
8	213 990,2	66 219	3,1	-2,7	2,5	2015
6	207 531,7	65 354	-4,3	0,0	-2,6	2016
2012						
12	50 840,5	19 577	-0,4	8,1	-1,1	1st quarter
8	50 919,1	19 396	-3,1	1,2	-2,9	2nd quarter
6	50 166,2	18 630	-3,4	3,2	-6,1	3rd quarter
5	48 820,8	20 538	-2,8	4,2	0,9	4th quarter
2013						
5	55 234,3	21 040	3,3	2,8	-0,9	1st quarter
6	53 601,1	19 463	0,0	6,9	-2,7	2nd quarter
5	53 297,5	18 749	1,7	8,7	18,1	3rd quarter
6	53 907,6	19 474	0,6	9,7	9,9	4th quarter
2014						
6	54 844,4	19 220	5,3	10,1	18,1	1st quarter
6	52 806,9	17 376	0,0	4,7	2,6	2nd quarter
7	51 113,9	18 559	0,0	4,2	-6,7	3rd quarter
7	52 249,9	19 986	-0,3	-1,4	7,4	4th quarter
2015						
9	57 669,1	18 063	2,7	-4,6	-8,6	1st quarter
7	54 095,2	15 958	4,9	-4,2	0,8	2nd quarter
8	50 425,1	15 954	-0,3	-2,9	6,5	3rd quarter
8	51 800,7	16 245	5,1	0,9	11,9	4th quarter
2016						
7	52 968,6	16 177	-7,8	4,0	15,1	1st quarter
7	53 418,5	15 352	-0,7	2,9	5,0	2nd quarter
4	49 779,6	16 763	0,0	-2,3	-10,7	3rd quarter
5	51 365,0	17 062	-8,3	-4,4	-17,1	4th quarter
2017						
5	53 406,4	13 830	-8,8	-1,7	-9,9	1st quarter
4	52 999,1	15 298	-9,2	-0,7	-1,2	2nd quarter
3	-12,4	1,7	10,8	3rd quarter

^a Andmed põhinevad lühiajastatistikal. 2017. aasta andmeid võidakse korrigeerida. Statistika Eesti majanduse tegevusalade klassifikaatori EMTAK 2008 järgi.

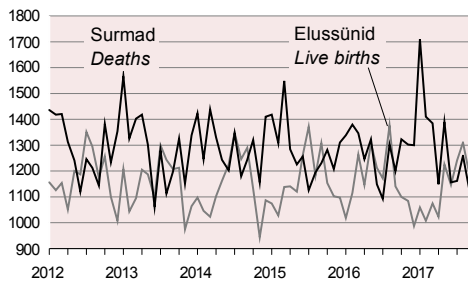
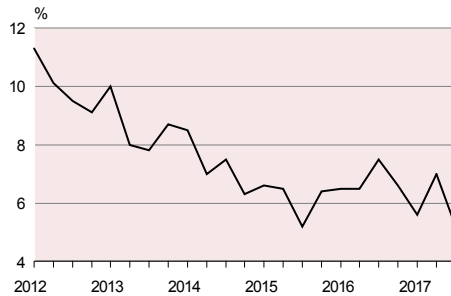
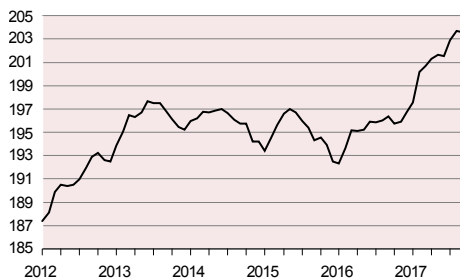
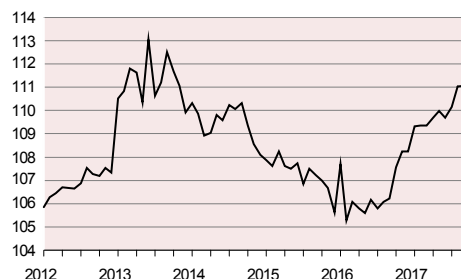
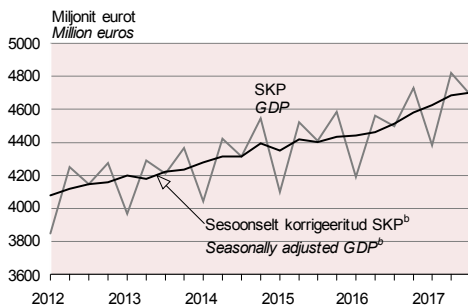
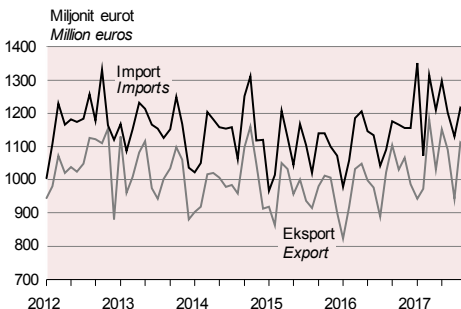
^b Eesti veandusettevõtete veoandmed. Andmeid võidakse korrigeerida järgmise kvartali andmete lisamisel.

^c 2017. aasta andmed on esialgsed.

^a Short-term statistics. The data for 2017 may be revised. Statistics according to the Estonian Classification of Economic Activities (EMTAK 2008, based on NACE Rev. 2).

^b Carriage data of Estonian transport enterprises. The data may be revised when the data for the next quarter are added.

^c Preliminary data for 2017.

Loomulik rahvastikumuu
Natural change of population

15–74-aastaste töötuse määr
Unemployment rate of population aged 15–74

Tarbijahinnaindeks, 1997 = 100
Consumer price index, 1997 = 100

Tööstustoodangu tootjahinnaindeks, 2010 = 100
Producer price index of industrial output, 2010 = 100
Producer price index of industrial output, 2010 = 100

Sisemajanduse koguprodukt aheldatud väärtustes (referentsaasta 2010 järgi)^a
Gross domestic product at chain-linked volume (reference year 2010)^a

Väliskaubandus
Foreign trade


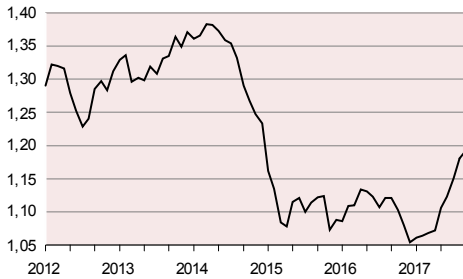
^a Referentsaasta järgi ahelindeksiga arvatud väärtused (referentsaasta väärtused korrutatakse arvestusperioodi ahelindeksiga). Referentsaasta on püsivhindades näitajate esitamiseks kasutatav tinglik aasta, indeksite seeria alguspunkt. Ahelindeks on järjestikuste perioodide aheldamiseks loodud kumulatiivne indeks, mis näitab komponendi kasvu võrreldes referentsaastaga. 2013. aasta I kvartali – 2017. aasta I kvartali andmed on korrigeeritud.

^b Aegriidade sesoonse korrigeerimine tähendab kindlaks teha ja kõrvaldada regulaarsed aastasisesed mõjud, et esile tuua majandusprotsesside pika- ja lühiajaliste trendide dünaamikat. SKP on sesoonselt ja tööpäevade arvuga korrigeeritud.

^a Values calculated by chain-linked index of reference year (values at reference year are multiplied by chain-linked index of the calculated period). Reference year is a conditional year for calculating chain-linked data and starting point of the series of chain-linked indices. Chain-linked index is a cumulative index for chain-linking sequential periods and it expresses the growth rate of a component compared to the reference year. The data for 1st quarter 2013 – 1st quarter 2017 have been revised.

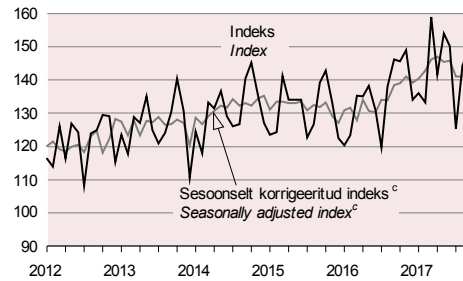
^b Seasonal adjustment of time series means identifying and eliminating regular within-a-year influences to highlight the underlying trends and short-run movements of economic processes. GDP is seasonally and working-day adjusted.

USA dollari kuukeskmine kurss euro suhtes
Average monthly exchange rate of the US dollar against the euro



Allikas: Euroopa Keskpang
Source: European Central Bank

Tööstustoodangu mahuindeks, 2010 = 100^a
Volume index of industrial production, 2010 = 100^a



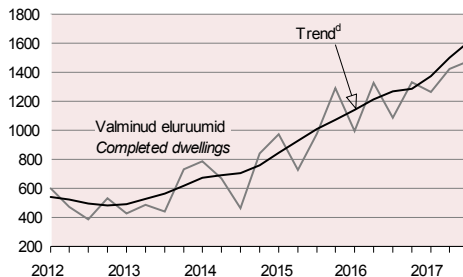
^a Statistika Eesti majanduse tegevusalade klassifikaatori EMTAK 2008 järgi.

^c Aegriidade sesoone korrigeerimine tähendab kindlaks teha ja kõrvaldada regulaarsed aastasisesed mõjud, et esile tuua majandusprotsesside pika- ja lühiajaliste trendide dünaamikat.

^a Statistics according to the Estonian Classification of Economic Activities (EMTAK 2008, based on NACE Rev. 2).

^c Seasonal adjustment of time series means identifying and eliminating regular within-a-year influences to highlight the underlying trends and short-run movements of economic processes.

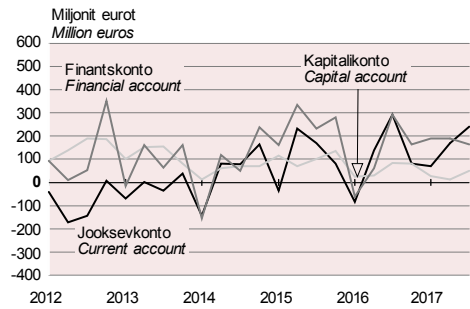
Valminud eluruumid
Completed dwellings



^d Trend – aegrea pikaajaline arengusuund.

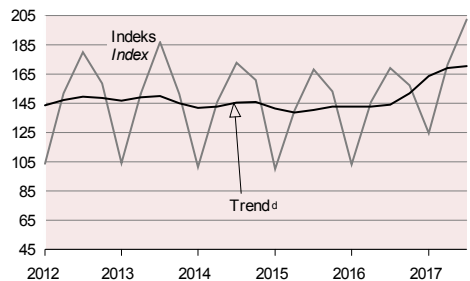
^d Trend – the long-term general development of the time series.

Maksebilanss
Balance of payments



Allikas/Source: Eesti Pank

Ehitismahuindeks, 2010 = 100^b
Construction volume index, 2010 = 100^b



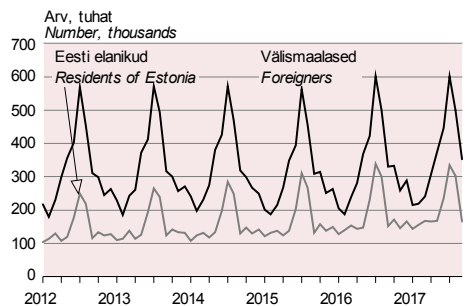
^b Ehitustööd Eestis ja välisriikides. Statistika Eesti majanduse tegevusalade klassifikaatori EMTAK 2008 järgi.

^d Trend – aegrea pikaajaline arengusuund.

^b Construction activities in Estonia and in foreign countries. Statistics according to the Estonian Classification of Economic Activities (EMTAK 2008, based on NACE Rev. 2).

^d Trend – the long-term general development of the time series.

Majutatute ööbimised
Nights spent by accommodated persons



EESTI STATISTIKA KVARTALIKIRJAS AVALDATUD ARTIKLID ARTICLES PUBLISHED IN THE QUARTERLY BULLETIN OF STATISTICS ESTONIA

Nr 1, 2009

Eakate taandumine tööturul. *Retirement of the elderly from the labour market.* Mai Luuk

Eesti edemused ja vajakajäämised innovatsiooni tulemuskaardil. *Innovation scoreboard: Estonia's advantages and shortcomings.* Aavo Heinlo

Mittetulundusühendused ja kodanikuühiskonna areng. *Non-profit organisations and development of civil society.* Helmut Hallemaa, Mihkel Servinski

Põllumajanduslikud majapidamised Balti- ja Põhjamaades. *Agricultural holdings in the Baltic and Nordic countries.* Eve Valdvee, Andres Klaus

Nr 2, 2009

Kriisist kriisini ehk Eesti praegu ja 10 aastat tagasi. *From crisis to crisis or Estonia now and 10 years ago.* Toomas Rei

Eesti transport aastal 2008. *Estonia's transport in 2008.* Piret Pukk

Eesti residentide töötamine välisriigis. *Employment of Estonian residents abroad.* Siim Krusell

Täiskasvanute tasemehariduses õppimine. *Adults continuing studies in the formal education system.* Tiiu-Liisa Rummo

Nr 3, 2009

Keskkonnamaksud – keskkonnakaitse majanduslikud meetmed. *Environmental taxes – economic instruments for environmental protection.* Eda Grüner, Kersti Salu, Kaia Oras, Tea Nõmmann

Kõrgtehnoloogiliste kaupade kaubavahetus Eestis aastail 2004-2008. *Trade in high-tech goods in Estonia in the years 2004–2008.* Riina Kerner, Allan Aron

Turism 2008. *Tourism 2008.* Anneken Metsoja, Helga Laurmaa

Arengutrendid kriisiaja tööturul. *Trends on labour market during the crisis.* Mai Luuk

Nr 4, 2009

Elamuehitus- ja kinnisvaraturu areng viimasel kümnendil. *Development of the dwelling construction and real estate market during the last decade.* Olga Smirnova, Merike Sinisaar

Koostootmine kui energiasäästuvõimalus. *Combined heat and power generation as an energy saving opportunity.* Helle Truuts, Rita Raudjärv

Eesti kodumajapidamiste käitumine kiire majanduskasvu aastatel. *Behaviour of households in Estonia in the years of fast economic growth.* Liisi Läänemets, Tõnu Mertsina

Kulutused – kas mõistlik tarbimine või kulutamine? *Expenditure – reasonable consumption or spending.* Piret Tikva

Nr 1, 2010

Eesti kaubavahetus majanduse põhikategooriate järgi. *Foreign trade of Estonia by Broad Economic Categories.* Riina Kerner, Allan Aron

Majanduslanguse põhjustatud muutused tööturul. *Changes on the labour market caused by economic recession.* Siim Krusell

Põllumajandustootmise tasuvus. *Profitability of agricultural production.* Ivar Himma, Elve Ristsoo, Andres Tekkel

Uus strateegia rahva ja eluruumide loenduse korraldamisel 2011. aastal. *A new strategy for organisation of the Population and Housing Census in 2011*. Diana Beltadze, Ene-Margit Tiit

Välisosalusega äriühingud Valga maakonnas 2007. aastal. *Companies with foreign shareholding in Valga county in 2007*. Mihkel Servinski

Nr 2, 2010

Valitsemissektori võlg ja defitsiit. *Government debt and deficit*. Agnes Naarits

2010. aasta põllumajandusloendus. *2010 Agricultural Census*. Eve Valdvee, Andres Klaus

Äriühingute majandustegevus. *Business activity of non-financial corporations*. Tiina Pärson

Kuritegude ohvriks langemine Eestis. *Crime victimisation in Estonia*. Kutt Kommel

Euroopa Liidu rahvastik aastal 2061. *The population of the European Union in 2061*.

Mihkel Servinski

Nr 3, 2010

Prooviloenduse õppetunnid. *Lessons of the pilot census*. Ene-Margit Tiit

Esimese ja teise põlvkonna immigrantrahvastik tööturul. *First and second generation immigration population on labour market*. Siim Krusell

Kaubavahetuse puudujääk 2009. aastal. *Foreign trade deficit in 2009*. Allan Aron, Riina Kerner

Uuenduslikkus luubi all. *Innovativeness under magnifier*. Aavo Heinlo

Vägivald paarisuhtes – müüdid ja tegelikkus. *Intimate partner abuse – myths and the reality*.

Merle Paats

Nr 4, 2010

Tööelu kvaliteedi subjektiivne mõõde. *Subjective measure of quality of work life*.

Karolin Kõrreveski

Töötlev tööstus taastub kriisist. *Manufacturing is recovering from the crisis*. Rita Raudjärv

Eesti majandust läbiva materjalivoo arvestus. *Economy-wide material flow account of Estonia*.

Eda Grüner

Kaubanduse areng viimasel viiel aastal. *Development of trade during the last five years*. Jaanika Tiigiste

Eesti rändestatistika ja piirkondlik rahvaarv Statistikaameti andmetel. *Estonian migration statistics and regional population according to the data of Statistics Estonia*. Helerin Rannala, Alis Tammur

Nr 1, 2011

Eesti riiklik statistika 90. *Official statistics of Estonia 90*. Priit Potisepp

Toidukaupade hinnatrendidest Eestis ja mujal maailmas viimastel aastatel. *Price developments of food products in Estonia and in the world during last years*.

Viktoria Trasanov

Sesoonselt korrigeeritud aegridade kvaliteet. *Quality of seasonally adjusted time series*. Mihkel Täht

Setomaa om hüa elamise, olõmisõ ja tulõmisõ kotus. *Setomaa is a good place for living, staying and coming*. Mihkel Servinski, Merli Reidolf, Garri Raagmaa

Nr 2, 2011

Statistika ajaloost. *About the history of statistics*. Ene-Margit Tiit

Eesti kinnisvaraturg muutuste keerises 2006-2010. *Estonian real estate market in a whirl of changes, 2006–2010*. Olga Smirnova

Tööturg majanduskriisi teises pooles. *Labour market in the second half of the economic crisis*.
Yngve Rosenblad

Transiitkaupade veost möödunud dekaadil. *Goods in transit over the last decade*. Piret Pukk

Pakkumise ja kasutamise tabelid eelmise aasta hindades. *Supply and use tables at the previous year's prices*. Iljen Dedegkajeva

Nr 3, 2011

Aasta lõpus algab rahva ja eluruumide loendus. *At the end of the year the Population and Housing Census will begin*. Diana Beltadze

Põllumajandustootmise olulisemad tulemused 2010. aastal. *Main results of agricultural production in 2010*. Helina Uku

Eesti ettevõtete kaubavahetus tehnoloogilise vaatepunktist. *Trade of Estonia's enterprises from point of view of technology*. Riina Kerner

Leibkonna kulutused kultuurile. *Household expenditure on culture*. Kutt Kommel

Nr 4, 2011

Rahvaloenduse ankeetidest. *About census questionnaires*. Ene-Margit Tiit

Eesti elanike sündimus ja suremus aastatel 2001–2010. *Fertility and mortality of the Estonian population in the years 2001–2010*. Alis Tammur, Jaana Rahno

Heaolu ja elukvaliteedi mõõtmisest OECD näitajate põhjal. *Measuring well-being and quality of life using OECD indicators*. Karolin Kõrreveski

Kas energia on Eestis odav või kallis? *Is energy in Estonia cheap or expensive?*
Rita Raudjärv, Ljudmilla Kuskova

20 aastat taasiseseisvumisest ja 90 aastat Eesti riikliku statistika rajamisest. *20 years from the restoration of independence and the 90th anniversary of official statistics in Estonia*.

Priit Potisepp, Kaja Sõstra, Allan Randlepp

Nr 1, 2012

Ülemaailmse finantskriisi mõju statistikale. *Impact of the global financial crisis on statistics*. Agnes Naarits

Majandusaasta aruannete kasutamine riikliku statistika tegemisel. *Use of annual reports in the production of official statistics*. Ene Saareoja, Merike Põldsaar

Põllumajanduse trendid ja hetkeseis 2010. aasta loenduse andmetel. *The trends and current state of agriculture based on the 2010 census*. Eve Valdvee, Andres Klaus

Tervena elada jäänud aastad ning nende arvutamine. *The concept and calculation of disability-free life expectancy*. Julia Aru, Jaana Rahno, Helerin Rannala

Viljandi maakonna ja Viljandi linna rahvastiku areng. *Population trends in Viljandi county and Viljandi city*. Mihkel Servinski, Helerin Rannala.

Nr 2, 2012

Eesti kaubavahetus üleilmsel taustal. *Estonia's trade in the world of globalisation*.
Riina Kerner

Raudteekaubaveod Eestis ja mujal Euroopas. *Rail freight transport in Estonia and elsewhere in Europe*. Liivi Adamson, Piret Pukk

Piima kokkuost ja piimatoodete tootmine aastail 2007–2011. *Purchase of milk and production of milk products*. Helina Uku

Valitsemissektori kulud funktsiooni järgi. *General government expenditure by COFOG*.
Maria Vassiljeva

Mida näitas Eesti esimene e-loendus? *E-census of the 2011 Population and Housing Census*. Diana Beltadze, Ene-Margit Tiit

Rahvaloendajate tegevus küsitluse järel. *Enumerators' activity after the Census*. Ene-Margit Tiit, Mare Vähi

Nr 3, 2012

Uuenduslikkus tasub end ära. *Innovativeness pays off*. Aavo Heinlo

Eesti eksportiva ettevõtte portree. *A portrait of the Estonian exporter*. Riina Kerner

Säästva arengu näitajad arenevad suurema lõimitavuse suunas. *Sustainable development indicators are moving towards greater integration*. Kaia Oras

Rahvastikuarengust Eesti meedia vahendusel. *Views on the future of the population in the Estonian media (summary)*. Mihkel Servinski

Rahvaloenduse üldkogumi hindamine. *Assessment of the target population of the census*. Ene-Margit Tiit, Mare Vähi, Koit Meres

Eesti rahvastikusündmustest üheksa kümnendi jooksul. *Vital events in Estonia through nine decades*. Helena Anijalg

Nr 4, 2012

Eesti rahvaarv, rahvastiku koosseis ja paiknemine 2011. aasta rahvaloenduse tulemuste põhjal. *The size, structure and distribution of the population of Estonia based on the 2011 census*. Diana Beltadze

Elatustase ja tööturuarengud Eestis – kriisiaastate võrdlus teiste Euroopa Liidu riikidega. *Standard of living and labour market trends in Estonia – a comparison with other European Union countries during the recession*. Siim Krusell

Tööjõu ühikukulu majanduse konkurentsivõime näitajana. *Unit labour cost as an indicator of the competitiveness of the economy*. Tõnu Mertsina, Tamara Jänes

Ettevõtlussektori investeeringud. *Business sector's investments*. Merike Sinisaar, Rita Raudjärv, Lilian Bõkova

Kinnisvara. *Real estate*. Kadi Leppik

Ettevõtluse demograafia. *Business demography*. Maret Helm-Rosin

Teenuste tootjahinnaindeksid. *Services producer price indices*. Eveli Šokman

2011. aasta rahva ja eluruumide loenduse alakaetuse hinnang. *Assessment of under-coverage in the 2011 Population and Housing Census*. Ene-Margit Tiit

Nr 1, 2013

Energiatarbimine kodumajapidamistes. *Energy consumption in households*. Rita Raudjärv, Ljudmilla Kuskova

Kes, kus ja miks internetti kasutab? *Who, where and why uses the Internet?* Mari Soiela

Palgatöötajatest ja töötasust töötasu struktuuri uuringu põhjal. *Employees and their earnings based on the Structure of Earnings Survey*. Mare Kusma

Vaesuse individuaalne ja sotsiaalne olemus. *The individual and social nature of poverty*. Tiiu-Liisa Laes

Välispäritolu ja põlisrahvastik tööturul. *The native and immigrant population in the labour market*. Siim Krusell

Nr 2, 2013

Kuidas ohjeldada juhuslikkust? *How to control randomness?* Kalev Pärna

Euroopa Liidus maksustatakse jätkuvalt tarbimist. *Taxation of consumption is still prevalent in the European Union*. Anu Lill

Liiklusõnnetustest taasiseseisvunud Eestis. *Traffic accidents in re-independent Estonia.*

Piret Pukk

Ligi kolmandik Eesti leibkondadest on aiapidajad. *Nearly a third of households in Estonia have kitchen gardens.* Eve Valdvee, Andres Klaus

Maksustamine ning aktsiisid Eestis. *Taxation and excise duties in Estonia.* Monika Sadilov, Kandela Õun

Nr 3, 2013

Rahvastikust – positiivselt. *A positive view of demographic trends.* Andres Oopkaup, Mihkel Servinski

Sisseränne Eestisse 2000–2011. *Immigration to Estonia in 2000–2011.* Alis Tammur, Koit Meres

Eesti-sisene töötamise ja õppimisega seotud pendelränne rahvaloenduste andmetel. *Commuting inside Estonia for work and study purposes based on the data of population censuses.* Ene-Margit Tiit

Kaupade ja teenuste eksport ning neis sisalduv kodumaine väärtus. *Exports of goods and services and the domestic value added embodied in exports.* Riina Kerner

Üleilmsed väärtusahelad. *Global value chains.* Tiina Pärson, Aavo Heinlo

Tööpuudusest laiemalt: vaeghõive ning potentsiaalne lisatööjõud. *Unemployment on a broader scale: underemployment and potential additional labour force.* Meelis Naaber

Nr 4, 2013

Ehitus buumist kriisini. *Construction from boom to crisis.* Merike Sinisaar

Keskonnamaksude arvepidamine võimaldab nende makromajanduslikku analüüsi. *Environmental taxes account enables analysing the taxes macroeconomically.* Kaia Oras, Kersti Salu

Tootemaksud ja -subsiidiumid aastail 2008–2010. *Taxes and subsidies on products in 2008–2010.* Anastassia Medkova

Teekaart registripõhise loenduseni. *Roadmap to a register-based census.* Doris Matteus

Elukvaliteet Tallinnas ja teistes Euroopa Liidu riikide pealinnades. *Quality of life in Tallinn and in the capitals of other European Union Member States.* Marika Kivilaid, Mihkel Servinski

Nr 1, 2014

Tallinn ja selle linnaosad. *Tallinn and its districts.* Marika Kivilaid, Mihkel Servinski

Eesti sisekaubandus. *Internal trade in Estonia.* Jaanika Tiigiste

Mis iseloomustab keskmist töövõimetuspensionäri? *What characterises an average incapacity pensioner?* Marianne Leppik

Välispäritolu rahvastiku teine põlvkond Eesti tööturul. *Second generation of immigrant population in the Estonian labour market.* Siim Krusell, Ellu Saar

Vanemaealiste tööhõive Eestis aastail 2007–2012. *Employment of the elderly in Estonia in 2007–2012.* Annely Saik, Kandela Õun

Nr 2, 2014

Metoodilised muudatused rahvamajanduse arvepidamise näitajate arvestuses. *Methodological changes in national accounts.* Annika Laarmaa

Mobiilpositsioneerimisandmete kasutamise võimalikkusest turismistatistika tegemisel. *Feasibility study on the use of mobile positioning data for tourism statistics.* Maiki Ilves, Epp Karus

Eesti suuremad linnad võrdluses naaberriikide suuremate linnadega. *The biggest Estonian cities in comparison with bigger cities in the neighbouring countries*. Mihkel Servinski, Marika Kivilaid

Hariduslike valikute küsimus – kas kombineerida erinevaid õppesuundi? *Making educational choices – to combine different fields of study or not?* Koit Meres, Kaia Kabanen

Rahvastikustatistika metoodikast. *On the methodology of demographic statistics*. Ene-Margit Tiit

Nr 3, 2014

Toimepiirkonnad Eestis. *Local activity spaces in Estonia*. Anu Tõnurist

Pärast taasisesivsustumist Eestisse elama asunute positsioon tööturul ja seal toimetulek. *Labour market position of those who immigrated after Estonia regained independence*.

Siim Krusell

Kuidas on kutseharidus seotud rändega? *How is vocational education linked with migration?* Koit Meres, Kaia Kabanen

Andmete võrreldavuse probleemidest: tööhõive ja töötuse näitajad tööjõu-uuringu ja rahvaloenduse järgi. *On the issues of data comparability: employment and unemployment indicators according to the Labour Force Survey and the Census*. Yngve Rosenblad

Nr 4, 2014

Eesti põllumajandussaaduste ja toidukaupade eksportijate käekäik aastail 1993–2013. *Activity of exporters of agricultural products and food preparations in 1993–2013*. Mirgit Silla, Evelin Puura

Tööjõukulu ja selle muutus aastail 2008–2012. *Labour costs and their change in 2008–2012*. Mare Kusma

Olulisemad muutused Eesti põllumajanduses pärast 2010. aasta loendust. *Main changes in Estonian agriculture after Agricultural Census 2010*. Eve Valdvee, Andres Klaus

Nr 1, 2015

Viljandi vald stardijoonel. *Viljandi rural municipality at the start line*. Mihkel Servinski, Anu Tõnurist, Koit Meres

Tartu eeslinnastumine. *The suburbanisation of Tartu*. Antti Roose, Rivo Noorkõiv, Martin Gauk

Sotsiaalne ettevõtetus Eestis. *Social entrepreneurship in Estonia*. Riinu Lepa, Agnes Naarits

Leibkonnad ja perekonnad registripõhises loenduses. *Households and families in the register-based census*. Kairiin Kütt

Nr 2, 2015

Väliskülastajad Eestis. *Foreign visitors in Estonia*. Anu Tõnurist, Kaja Sõstra

Noorte hõive ja töötus ning majanduskriisi mõju. *Youth employment and unemployment and the impact of the economic crisis*. Siim Krusell

Eesti säästva arengu näitajate valguses. *Estonia in the light of sustainable development indicators*. Kaia Oras, Evelin Enno

Statistika tegemisel on abiks majandusaasta aruanded ja statistilised mudelid. *Use of annual reports and statistical models in the production of statistics*. Ebu Tamm, Merike Põldsaar, Reet Nestor

Nr 3, 2015

Eesti regionaalne areng piirkondliku lisandväärtuse järgi. *Regional development in Estonia based on regional value added*. Robert Mürsepp

Eestlased ja mitte-eestlased tööturul – rahvuse ja eesti keele oskuse mõju tööturupositsioonile. *Estonians and non-Estonians in labour market – impact of ethnic nationality and command of Estonian on labour market positions.* Siim Krusell

Registripõhise rahva ja eluruumide loenduse meetodika ja selle arengusuundumused. *Register-based population and housing census: methodology and developments thereof.* Ene-Margit Tiit

Üksikandmete statistiline paljastumine ja selle vältimine. *Statistical disclosure of individual data and prevention thereof.* Ebu Tamm

Interneti kasutamine 65–74-aastaste seas. *Internet use among people aged 65–74.* Jaanika Ait

Nr 4, 2015

Keskonnakaupade ja -teenuste sektor kui arenev majandusvaldkond. *Environmental goods and services sector as a developing economic sector.* Kaia Oras, Kaia Aher

Piimatootmise võimalikkusest Eestis. *On possibility of milk production in Estonia.* Eve Valdvee, Andres Klaus

Millega veetakse kaupa, millega reisijaid? *Modes of transport: freight versus passengers.* Piret Pukk

Eesti ja välismaise päritoluga kaupade osatähtsus Eesti ekspordis. *Share of Estonian and foreign goods in Estonian exports.* Allan Aron

Teenidussektor Eestis ja mujal Euroopa Liidus. *Service sector in Estonia and the rest of the European Union.* Jaanika Tiido

Nr 1, 2016

Pärnu maakonna üldine ülevaade. *General overview of Pärnu county.* Marika Kivilaid, Greta Tischler

Pärnu maakonna rahvastik. *Population of Pärnu county.* Mihkel Servinski

Pärnu maakonna majandusareng. *Economic development of Pärnu county.* Märt Leesment

Pärnu maakond vaadatuna läbi tööturuprisma. *Pärnu county through labour market prism.* Siim Krusell

Kohalike omavalitsuste eelarved. *Local government budgets.* Marika Kivilaid

Leibkonnad ja nende elamistingimused. *Households and their living conditions.* Anu Tõnurist

Prognoosimisest ja seirest Pärnu maakonna näitel. *Projections and monitoring illustrated by Pärnu county.* Urmas Kase

Külavanemal põhinev kogukondlik mudel. *Village elders-based community model.* Rivo Noorkõiv

Nr 2, 2016

Eesti statistikasüsteemi 95. aastapäeva konverentsile pühendatud eriväljaanne.

Special edition dedicated to the conference of the 95th anniversary of the Estonian Statistical System

Nr 3, 2016

Kuidas koguda andmeid võimalikult vähese vaevaga ja pakkuda ettevõtjale vajalikku infot. *How to collect data with minimal work and provide entrepreneurs with data they need.*

Andres Oopkaup

E-MTA loob uued e-teenused ja vähendab ajakulu maksukohustuse täitmisel. *E-ETCB creates new e-services and saves time on meeting tax liabilities.* Marek Helm

Briti statistikasüsteem väikeriigi esindaja pilgu läbi. *British statistical system through the eyes of a representative of a small country*. Allan Randlepp

Statistiliste andmete kvaliteet ja nende kuvand avalikkuses. *Quality and public image of statistical data*. Peeter Annus

Residentsuse indeks ning selle rakendamine loendusel ja rahvastikustatistikas. *Residency index and its applications in censuses and population statistics*. Ene-Margit Tiit, Ethel Maasing

Eesti majanduse käekäik Euroopa mittefinantskontode taustal. *Progress of Estonian economy in light of European non-financial accounts*. Robert Mürsepp

Noorte tööelu kvaliteet kriisijärgses Euroopas. *Quality of work life of the young in post-crisis Europe*. Heidy Roosimägi

Turismist ja majutusest 2015. aastal statistikauuringute põhjal. *Tourism and accommodation in 2015 based on statistical research*. Helga Laurmaa

Liikumiskeskonna areng Tallinnas: väljakutse ühistranspordile. *Development of mobility environment in Tallinn: a challenge for public transport*. Rivo Noorkõiv, Dago Antov

Nr 4, 2016

Mida tõi endaga kaasa elektrituru avanemine? *The impact of opening the electricity market*. Rita Raudjärv

Kaupade väljavedu Venemaale. *Exportation of goods to Russia*. Evelin Puura, Mirgit Silla

Keskonnakaitseteenused Eesti turul. *Environmental protection services on the Estonian market*. Grete Luukas

Statistika olulisus lastega perede heaolu tagamisel. *The importance of statistics in ensuring the well-being of families with children*. Hede Sinisaar, Kristiina Kukk

Veel kord palgalõhest *Once more on the pay gap*. Maret Seppel

Registripõhise rahva ja eluruumide loenduse tarbijaküsitlus *User survey of the register-based population and housing census*. Ene-Margit Tiit

Nr 1, 2017

Andres Oopkaup: „Mul on põhimõte: alati tuleb tõtt rääkida!“. *Andres Oopkaup: "I have a principle – always tell the truth!"*. Mihkel Servinski

Põlisrahvastik ja välispäritolu rahvastik Eestis. *Native and foreign-origin population in Estonia*. Alis Tammur

Rahvaarvu arvutamine: residentsuse indeks vs. rahvastikuregister. *Calculation of population size: residency index vs population register*. Koit Meres

Registripõhise ja tegeliku elukoha kattuvus. *Coincidence of actual place of residence with population register records*. Helerin Äär

Põllumajanduslike majapidamiste arv väheneb endiselt. *The number of agricultural holdings continues to decrease*. Eve Valdvee, Andres Klaus

Kulutused suhkrurikastele toodetele eri tervisehinnanguga leibkondades. *Expenditure on high-sugar products in households with different self-reported health statuses*. Märt Leesment

Nr 2, 2017

Tuulikki Sillajõe: Euroopa Liidu Nõukogu eesistumine on kindlasti suur võimalus Statistikaametile. *Tuulikki Sillajõe: The Presidency of the Council of the European Union is certainly a great opportunity for Statistics Estonia*. Mihkel Servinski

- Majandusülevaade. *Economic overview*. Robert Mürsepp
- Sotsiaalne kaitse. *Social protection*. Eve Telpt
- Keskkond. *Environment*. Kaia Oras
- Rahvastik. *Population*. Alis Tammur
- Haridus. *Education*. Marti Lillemägi
- Kultuur. *Culture*. Kutt Kommel
- Sotsiaalne tõrjutus ja vaesus. *Social exclusion and poverty*. Tiitu-Liisa Rummo
- Tööturg. *Labour market*. Heidy Roosimägi
- Brutokuupalk ja vabad ametikohad. *Wages and salaries*. Kai Maasoo
- Hinnad. *Prices*. Eveli Šokman
- Rahvamajanduse arvepidamine. *National accounts*. Annika Laarmaa
- Valitsemissektori rahandus. *Government finance*. Agnes Naarits
- Ettevõtete majandusnäitajad. *Financial statistics of enterprises*. Hille Vares, Reet Nestor
- Majandusüksused. *Economic units*. Katrin Aasmäe, Svetlana Šutova
- Väliskaubandus. *Foreign trade*. Evelin Puura, Mirgit Silla
- Põllumajandus. *Agriculture*. Aili Maanso, Ivika Aasa
- Kalandus. *Fishing*. Evelin Enno-Sakwan
- Metsamajandus. *Forestry*. Mati Valgepea
- Energeetika. *Energy*. Helle Truuts
- Transport. *Transport*. Piret Pukk, Rita Raudjärv
- Infotehnoloogia. *Information technology*. Tiina Pärson, Jaanika Ait
- Turism ja majutus. *Tourism and accommodation*. Helga Laurmaa
- Teadus- ja arendustegevus. *Research and development*. Rita Raudjärv, Tiina Pärson

Nr 3, 2017

- Emöke Sogenbits: Kui jätab inimesed unarusse, on asjade käik ettearvamatult. *Emöke Sogenbits: If you neglect the people, the course of events is unpredictable*. Mihkel Servinski, Asso Uibo
- Lõuna-Eesti ja ühine statistiliste territoriaalüksuste liigitus (NUTS). *Southern Estonia and Nomenclature of Territorial Units for Statistics (NUTS)*. Marika Kivilaid, Greta Tischler
- Lõuna-Eesti elanikud. *Inhabitants of Southern Estonia*. Ene-Margit Tiit
- Lõuna-Eesti majandus: (majanduslikult) eduka piirkonna olemus. *Economy of Southern Estonia: the nature of an (economically) successful region*. Märt Leesment
- Tööturg Lõuna-Eestis. *Labour market in Southern Estonia*. Mihkel Servinski
- Kutseharidusega inimesed Lõuna-Eestis. *People with vocational education in Southern Estonia*. Koit Meres
- Elukvaliteet Lõuna-Eestis. *Quality of life in Southern Estonia*. Kadri Raid
- Maakondlike arenduskeskuste ühendatud kompetentsid aitavad luua edukat Eestit. *The joined competences of county development centres help to make Estonia successful*. Silja Talviste, Kaarel Lehtsalu, Asso Uibo
- Haldusreform: ülevaade toiminust ja esimesed järeldused. *Administrative reform: an overview and first conclusions*. Rivo Noorkõiv, Mikk Lõhmus