

TIETO & TRENDIT

Tilastokeskus 

Talous- ja hyvinvointikatsaus

2•2015



ISO DATA

Vienti tökkii –
miksi?

Töihin pyörällä
tai julkisilla

Kotitalouksien varat
ja velat kasvussa

14

Taloudessa pulaa enemmän luottamuksesta kuin resursseista

Painopaperiin panostaminen, työn kustannukset, Nokian vaikeudet, tuotannon siirtyminen ulkomaille, Venäjä, markkinointiosaaminen... ekonomisti Pasi Kuoppamäki näkee Suomen vientivaikeuksiin useita syitä.

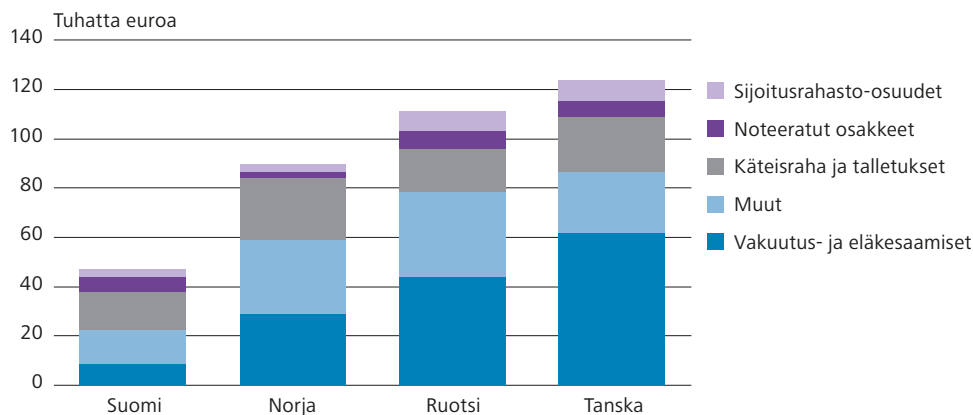
Lehtikuva Oy



20

Suomalaisilla pienet rahoitusvarat verrattuna muihin Pohjoismaihin

Kotitalouksien varallisuus on kasvanut, mutta se on edelleen selvästi pienempi kuin muissa Pohjoismaissa. Syynä on etenkin suomalaisten rahoitusvarallisuuden vähäisyys.



PÄÄKIRJOITUS

- 5** Hype laantuu, iso data arkipäivästy
Marjo Bruun
- 6** TRENDIT
- 10** T&k-investointien osuus on pudonnut huolestuttavasti
Aku Alanen
- 14** Viennin kasvu vahvistaisi luottamusta koko taloudessa
Pasi Kuoppamäki
- 17** Vienti ei vedä vielääkään
Christina Telasuo
- 20** Kotitalouksien varallisuus kasvaa Suomessa – muista Pohjoismaista ollaan silti jäljessä
Martti Pykäri

[TOIMITUKSELTA]

Big data, iso data, massadata

Iso data eli suuret ja järjestelemättömät tietomassat on noussut laajemman kiinnostuksen kohteeksi viimeisen parin vuoden aikana.

Teema-osiossa pohditaan mitä ilmiö merkitsee tilastotieteessä: Mitä uusia tietomassoja tilastoinnissa voitaisiin käyttää? Mitä hyviä ja huonoja puolia on eri aineistoissa? Ratkaiseeko iso data tiedonkeruun ongelmat vai tuoko se tusinan lisää?

Iso data on sen verran uusi ilmiö, ettei käsitteelle ole vakiintunutta suomenosta. Englanninkielinen *big data* kääntyy suoraan isoksi dataksi, mutta tarjottu on myös massadataa.

Jääkö iso data käsitteenä elämään? Tietokoneiden laskentakapasiteetin huima kasvu on tehnyt eilisen isosta datasta tämän päivän pikkudataa – tietoa, joka mahtuu yhdelle koneelle. Ehkä jatkossa puhumme vain datasta. ■

23 Iso data tilastotieteessä

Pasi Piela

28 Ison datan alkulähteillä

Hannes Heikinheimo ja Antti Ukkonen

34 Soveltuvatko kaupan kanta-asiakasrekisterit kulutuksen tilastointiin?

Tarja Hatakka

39 Oman datan hallinta on perusoikeus

40 Miksi isoon dataan hukutaan?

Reijo Sund

46 Tilastovirastot vahvoilla erilaisen tiedon yhdistämisessä

Jari Nieminen

49 Reittiopas taipuu myös työmatkalaskelmiin

Topias Pyykkönen

56 Johtaminen edellyttää taitoa ja tietoa

Juha Saukkonen

[TEEMA]

23

Isoa dataa on tuettava avoimilla tilastotiedoilla

Iso data haastaa perinteistä tilastotietoa nopeutensa vuoksi. Tilastotiedon vahvuuksia ovat sen sijaan tiedon ajallinen vertailtavuus ja edustavuus. Yhteistyöllä on saavutettavissa molempia tiedontuotannon muotoja hyödyttäviä etuja.



[TEEMA]

49

Pääkaupunkiseudulla työmatka taittuu julkisilla puolella tunnissa

HSL:n Reittioppaan ja Tilastokeskuksen työssäkäyntitilaston avulla on mahdollista arvioida, miten kauan työmatkaan menisi, jos jokainen käyttäisi julkisia liikennevälineitä.





57

Henkilötiedonkeruu verkkoon

Tilastokeskus antaa kansalaisille mahdollisuuden vastata kyselyihinsä verkossa vuodesta 2016 alkaen, lupaa Laura Hulkko.



- 57** Henkilötiedonkeruuta tehdään pian myös verkossa
Laura Hulkko
- 60** Negatiivinen inflaatio, hintavakaus vai deflaatio?
Mari Ylä-Jarkko
- 62** Indeksilaskentaa Myanmarin markkinoilla
Mari Ylä-Jarkko
- 64** Maailman ensimmäinen väestötilasto
Mauri Nieminen
- 70** Kansantalouden tilinpitoa on kehitetty etenkin kriisiaikoina
Eeva Hamunen

KIRJOJA

- 76** Arvaamattomat arvot
Kirsti Ahlqvist
- 78** Verkostojen voima
Kaisa Saarenmaa
- 81** Vaalipulssi.fi taltioi vaalien somekohinan
- 86** Entä jos kaikki polkisivat töihin?
Pasi Piela

64

Maailman ensimmäinen väestötilasto

Vuonna 1750 Ruotsi-Suomessa aloitettiin vuosittaisen väestötilaston laatiminen. Ruotsissa ja Suomessa on tällä hetkellä maailman pisimmät aikasarjat väkiluvusta ja väestömuutoksista.

KESKUSTELE BLOGISSA
TIETOTRENDITBLOGI.STAT.FI
Nuoret naiset menestyivät "miehisissä" vaaleissa
777 metriä vaaliurnille
Nollatuntisopimukset – epävarmuutta vai toivottua joustoa?

Tilastokeskus 

JULKAISIJA
Tilastokeskus
www.tilastokeskus.fi/tietotrendit
Työpajankatu 13, 00022
Tilastokeskus
tietotrendit@stat.fi
etu.sukunimi@stat.fi
ISSN 1795–9799

• VASTAAVA PÄÄTOIMITTAJA Heli Mikkeliä puh. 029 551 3200 ARTIKKELIPÄÄTOIMITTAJA Riitta Harala TOIMITUS: Kirsti Ahlqvist, Jouni Kotkavuori, Päivi Linna-Koski, Maija Metsä-Pauri, Kaisa Saarenmaa, Informaatikko: Mikko T. Mäkinen GRAAFINEN ILME JA TAITTO Riikka Turunen PAINATUS Edita Prima Oy TILAUSEHDOT: Tieto&trendit – talous- ja hyvinvointikatsaus ilmestyy viisi kertaa vuodessa. Kestotilauksen hinta on 98 euroa vuodessa. Tilaus alkaa heti seuraavasta numerosta ja jatkuu kestotilauksena ilman eri ilmoitusta. Irtonumeron hinta on 22 euroa. TILAUKSET: Edita Publishing Oy, puh. 020 450 05, asiakaspalvelu.publishing@edita.fi tai tietotrendit@stat.fi TOIMITUSNEUVOSTO: Jaana Huhta, Marjaana Järvensivu, Johanna Laiho-Kauranne (Luke), Aura Pasila, Pekka Räsänen (Turun yliopisto), Olli Savela, Eero Savolainen (Suomen Pankki), Tuukka Tammi (THL).
• Julkaisun artikkelit eivät ole Tilastokeskuksen kannanottoja vaan niistä vastaavat tekijät. Julkaisun tietoja saa lainata, lähde mainiten.



MARJO BRUUN
PÄÄJOHTAJA

Hype laantuu, iso data arkipäiväistyy

Pahin hypetyspöly big datan, eli suomalaisittain ison datan, ympärillä on hiljalleen laskeutumassa ja tilalle on nousemassa tilastoinninkin kannalta mahdollisuuksia. Iso data tarjoaa vakavasti otettavan lisän tilastotuotannon modernisoimiseksi.

Räjähdysmäisesti kasvava tiedon määrä sähköisissä järjestelmissä on jo vuosia ruokkinut visioita datavallankumouksesta, joka johdattaa yhteiskunnan uudelle tekniselle ja tiedolliselle tasolle. Reaaliaikaisesti tarjolla oleva lähes rajaton tietomäärä mahdollistaa edullisen perustan uusien palveluiden tuottamiseksi ja vanhojen uudistamiseksi.

Uskoa vahvistavat konkreettiset esimerkit uudesta tulevaisuudesta ovat toistaiseksi olleet vähissä. Ponnistelut alkavat kuitenkin viimein tuottaa hedelmää. EU-maiden tilastovirastot ovat yhteisessä strategiassaan sitoutuneet ison datan hyödyntämiseen tilastokäytössä (ns. Scheveningen Memorandum), ja uusien ideoiden testauksesta ollaan siirtymässä tuotannollisiin ratkaisuihin.

Pisimmällä ollaan varmaankin Hollannissa. Maanteiden massiivisen sensoriverkoston raakadataa on analysoitu nykyaikaisin menetelmin, jotta tieliikennetilastoja voitaisiin rikastaa. Hollannissa on myös tuloksellisesti testattu väestön taloudellisia odotuksia kuvaavien tilastojen täydentämistä Twitterin ja Facebookin tarjoamalla datalla. Sosiaalisen median mahdollinen käyttö tilastotoimessa on kaikkiaan kansainvälisen tilastoyhteisön mielenkiinnon kohteena.

Myös muualla on edistytty ison datan hyödyntämisessä. Italiassa kuluttajahintaindeksin lentomatkatietoja ja kodinelektronikkatietoja täydennetään internetistä kerättävillä tiedoilla. Myös Tilastokeskuksessa on meneillään kehittämishanke, jolla pyritään selvittämään kahden uuden datalähteen eli verkkokauppojen ja kassapääteaineistojen hintatietojen hyödyntämistä.

Useat maat ovat jo testanneet kännyköiden käytöstä kertyvää dataa matkailu-, työssäkäynti- ja muuttoliiketilastojen tekemisessä. Suuria toiveita asetetaan satelliittikuvannuksen hyödyntämiseen maataloustilastoissa.

Tyypillisimmät datalähteet tilastojen täydentämiseksi ovat ainakin toistaiseksi olleet sosiaalinen media, mobiililaitteista kertyvä data ja internetin tietojen hyödyntäminen. Lyhytkin kokemus on osoittanut niiden tarjoavan usein nykyistä nopeamman ja edullisemman tavan tuot-

YRITYSTEN HALLUSSA OLEVAN
DATAN KÄYTTÖ TILASTOJEN POHJANA
VAATII KUMPPANUUKSIEN LUOMISTA.

taa tietoa. Samanaikaisesti EU:n tietosuojadirektiiviä ollaan uudistamassa. Sen vaikutus esimerkiksi mobiilidatan käyttöön on vielä epäselvää.

Ison datan hyödyntämistä on tilastovirastoissa arvioitu myös menetelmällisistä lähtökohdista. Uusi ja vanha tuotantomalli eivät sulje toisiaan pois, vaan niiden välillä on myös yhtäläisyyksiä. Jo nyt kansantalouden tilinpito käyttää erilaisia lähteitä, joiden luotettavuutta arvioidaan kokonaistaloudellisessa kehikossa. Hallinnollisten aineistojen tiedot koostetaan tilastovirastojen välittömän ohjauksen ulkopuolella, mutta aineistojen laatua pystytään silti arvioimaan riittävän luotettavasti.

Edistysaskeleista huolimatta haasteita toki riittää. Hallinnollisista aineistoista poiketen suurin osa isosta datasta on yritysten hallussa. Datan käyttö tilastojen pohjana edellyttää uusien, molempia osapuolia hyödyttävien kumppanuuksien luomista. Monissa tapauksissa myös kansallista lainsäädäntöä tulisi kehittää yhteiskunnan muuttuvien tarpeiden myötä. Samalla on huolehdittava siitä, että tilastotuotannon kulmakivi, yksilön tietosuoja, on varmistettu asianmukaisella tavalla.

Tehokkaasti organisoitu tilastotoimi edellyttää uusien tietovarantojen innovatiivista hyödyntämistä. Tarve korostuu tiukassa taloudellisessa tilanteessa. On selvää, ettei myöskään Tilastokeskus voi jättää huomioimatta uusien tietolähteiden tuomia mahdollisuuksia.

	Bkt:n vuosi- muutos, %	Teollisuus- tuot. vuosi- muutos, %	Yhdenmuk. inflaatio v. muutos %	Työttö- myys- aste, %
	4. nelj. 2014	4. nelj. 2014	maalisk. 2015	helmik. 2015
Suomi	-0,1	-0,9	0,0	9,1
Alankomaat	1,4e	-0,9	-0,3	7,1
Belgia	1,0	2,0	-0,1	8,5
Espanja	2,1e	1,3	-0,8	23,2
Irlanti	4,1	29,5	-0,3	9,9
Italia	-0,6	-0,9	0,0	12,7
Itävalta	-0,2	0,6	0,9e	5,3
Kreikka	1,2e	5,0	-1,9	25,7*
Kypros	-2,0	-2,3	-1,4	16,3
Latvia	2,1	-1,7	0,5	10,2*
Liettua**	2,4	3,2	-1,1	10,0
Luxemburg	4,0*	8,7	0,1	5,8
Malta	4,0	0,5	0,5	5,9
Portugali	0,5	-1,8	0,4	14,1
Ranska	0,4	-0,6	0,0	10,6
Saksa	1,6	0,7	0,1	4,8
Slovakia	2,4	0,4	-0,4	12,3
Slovenia	2,4	5,1	-0,4	9,4
Viro	3,0	6,3	0,0	6,2*
Euroalue	0,9	1,0	-0,1e	11,3
Britannia	2,7	2,2	0,0	5,5*
Bulgaria	1,6e	3,7	-1,1	10,2
Kroatia	0,3	4,2	0,0	18,5
Puola	2,9	3,6	-1,2	7,8
Romania	2,7e	4,8	0,8	6,4
Ruotsi	2,7	-2,1	0,7	7,9
Tanska	1,5	1,5	0,3	6,2
Tšekki	1,2	5,6	0,1	5,5
Unkari	3,4	4,8	-0,5	7,5*
EU28	1,2	1,3	-0,1e	9,8
Norja	3,2	4,3	1,7	3,9*
Sveitsi	1,9	2,7	-0,5	3,5
Venäjä	0,4	2,1	16,9	5,8
Kiina	7,3	6,7	1,4	4,1*
Intia	7,5	3,8*	5,2	4,9*
Japani	-0,8	0,1*	2,2*	3,5
Yhdysvallat	2,4	4,1	-0,1	5,5

Lähteet: Eurostat, Venäjän ja Latvian tilastovirastot, Tradingeconomics.com

(e) ennakkoluku

* POIKKEUKSET:

Bkt: Luxemburg 3. nelj. 2014

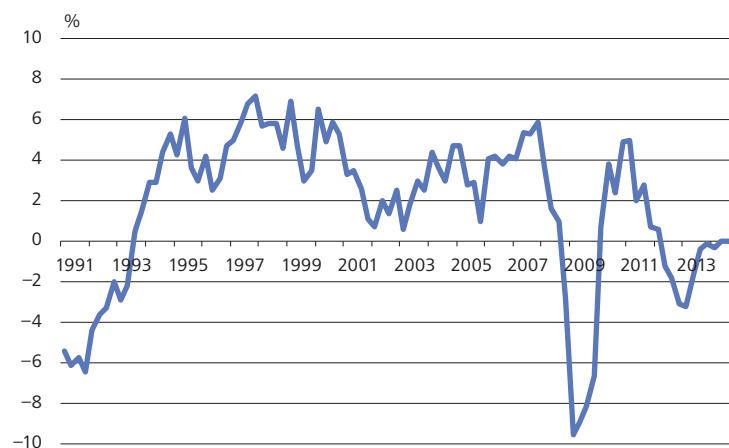
Teollisuustuotanto, työpäiväkorjattu: Intia, Japani jouluk. 2014

Yhdenmukaistettu inflaatio: Japani, helmik. 2015

Työttömyysaste: Kreikka, Viro, Britannia, Unkari, Norja tammik. 2015; Latvia, Kiina 4. nelj. 2014; Intia 2013

** Liettua liittyi euroalueeseen 1.1.2015

Bruttokansantuotteen volyymin muutos neljänneksittäin 1991–2014



Lähde: Tilastokeskus. Neljännesvuositilinpito

Suomen kasvuennuste Japanin tasoa

Vuosi 2014 oli jo kolmas peräkkäinen taantumavuosi Suomen taloudessa. Kansantalouden neljännesvuositilinpidon mukaan bkt:n volyymi pieni 0,1 prosenttia viime vuonna. Viimeisellä neljänneksellä bkt väheni 0,2 prosenttia edellisestä neljänneksestä ja saman verran myös vuoden 2013 loka–joulukuuhun verrattuna.

Kuluvalle vuodelle eri laitosten talousennusteet lupaavat 0,5–1 prosentin ja seuraavalle 1,4–1,7 prosentin kasvua. Suomen talouskasvu on jäämässä Japanin tasolle ja vaisummaksi kuin euroalueella.

Eurostatin mukaan euroalueen bkt kasvoi 0,3 prosenttia heinä–syyskuulta loka–joulukuulle. Koko vuodelle 2014 kasvua kertyi 0,9 prosenttia. USA:ssa bkt kasvoi 2,4 prosenttia viime vuonna. ■

Bkt:n kasvuennusteet Suomelle vuosimuutos-%

	2015	2016
Danske Bank	0,5	1,5
Nordea	0	1,5
OP Ryhmä	1,0	1,7
VM	0,5	1,4
Euroopan komissio	0,8	1,4

Bkt-kasvu ja ennusteet vuosimuutos-%

	2014	2015	2016
Euroalue	0,9	1,4	2,0
USA	2,4	3,1	3,0
Japani	0	1,0	1,4
Kiina	7,4	7,0	6,9
Intia	7,3	7,7	8,0

Lähde: OECD, Interim Global Economic Assessment, March 2015

Konjunkturi-institutit:

Vienti vetää Ruotsin taloutta

Ruotsin tutkimuslaitos Konjunkturitutet arvioi maaliskuun suhdannekatsauksessaan Ruotsin viennin kasvavan 5 prosentin vuosivauhtia seuraavat kaksi vuotta. Viennin vauhdittumista ajavat kruunun suhteellisen heikko kurssi ja maailmantalouden odotettu elpyminen. Brutto-

kansantuote kasvaa viennin vetämänä arviolta 3 prosenttia sekä tänä että ensi vuonna. ■

Lähteet: Konjunkturitutet, Konjunkturläget, Mars 2015

Suomen kasvihuonekaasupäästöt olivat 63,2 miljoonaa hiilidioksiditonnia Kioton pöytäkirjan toisen velvoitekauden ensimmäisenä vuotena 2013.

Kokonaispäästöt kasvoivat edellisestä vuodesta noin prosenttia; energiasektorin päästöt nousivat 2 prosenttia, teollisuusprosessien ja tuotteiden käytön sekä maatalouden päästöt pysyivät edellisen vuoden tasolla, jättesektorin päästöt vähenivät 5 prosenttia.

Energiasektorilla kivihien käyttö lisääntyi yli kolmanneksella, kun taas turpeen kulutus pieneni 12 prosenttia ja maakaasun 7 prosenttia.

Maankäyttö, maankäytön muutos ja metsätalous -sektorin (kuviassa LULUCF) poistumat pienenevät edelliseen vuoteen verrattuna noin 27 prosenttia ollen 20,4 miljoonaa hiilidioksiditonnia. Tätä sektoria ei sisällytetä kokonaispäästöihin.



Energiasektorilta leijonanosa päästöistä

Koko kasvihuonekaasupäästöjen aikasarja vuosille 1990-2013 on päivitetty vastaamaan muuttuneita raportointivaatimuksia. Tiedot perustuvat Tilastokeskuksen julkaisemiin tarkennettuihin päästötietoihin.

”Suomen kasvihuonepäästöt vaihtelevat paljon vuosittain sen mukaan miten Pohjoismaissa sataa”, toteaa kasvihuonekaasujen tilastoinnista vastaava kehittämisspäällikkö **Riitta Pipatti**. ”Pohjoismaisilla sähkömarkkinoilla Suomi ostaa sateisina vuosina paljon vesivoimalla tuotettua sähköä Ruotsista ja Norjasta, jolloin Suomen päästöt vähenevät. Vähäsaateisina vuosina Suomi myy näille maille turpeella ja kivihielellä tuotettua sähköä, jolloin Suomen päästöt kohoavat.”

”Koko totuutta vaihtelun syistä tämä ei kerro. Tekijöitä on monia muitakin, kuten yleinen taloustilanne, sähkön tuotimäärät Venäjältä ja vuotuinen lämmitystarve.”

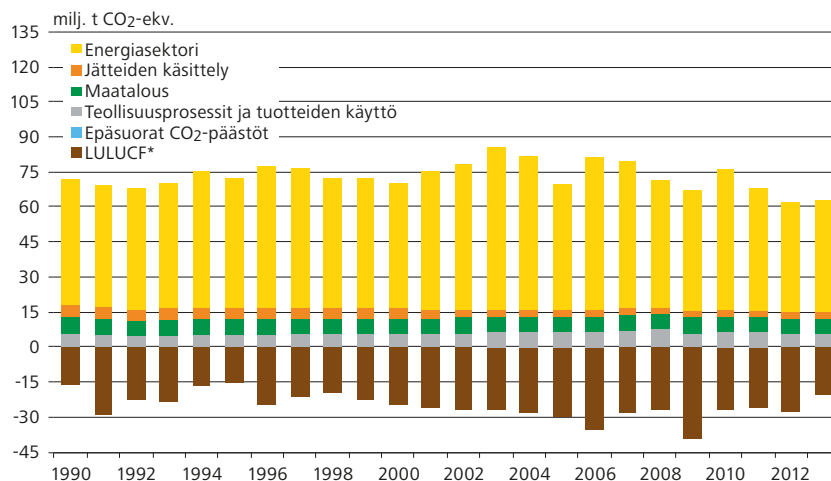
Vuosittaisista vaihteluista huolimatta kasvihuonekaasupäästöjen trendi on ollut hitaasti laskeva. Vuonna 2013 päästöt olivat 12 prosenttia alhaisemmat kuin vuonna 1990.

”Päästöjen vuosittainen vaihtelu on merkittävää ja noudattaa pääsääntöisesti energiasektorilla tapahtuvia muutoksia”, Pipatti sanoo. ”Päästötrendi oli kasvava

[YMPÄRISTÖ]

Suomen kasvihuonekaasupäästöt vaihtelevat sateiden mukaan

Suomen kasvihuonekaasupäästöt ja -poistumat sektoreittain



* LULUCF tarkoittaa maankäyttöä, maankäytön muutosta ja metsätalous -sektoria.

Lähde: Kasvihuonekaasujen inventaario. Tilastokeskus

2000-luvun alkuvuosiin asti, sen jälkeen päästötrendi on ollut laskeva. Trendiin vaikuttavat monet tekijät kuten toteutetut päästövähennystoimet mutta myös Suomen taloudellinen tilanne." ■

Lähde: Kasvihuonekaasujen inventaario. Tilastokeskus

[VÄESTÖ]

Imeväiskuolleisuus ennätysellisen pientä

Ensimmäisenä ikävuonnaan kuoli 124 lasta vuonna 2014. Määrä on toiseksi pienin itsenäisyyden aikana. Vuotta aiemmin määrä oli toistaiseksi pienin eli 102 lasta.

Imeväiskuolleiden määrä tuhatta elävänä syntyneitä lasta kohti oli viime vuonna 2,2. Vuosia 2014 ja 2013 ennen pienin imeväiskuolleisuusluku oli 2,3 promillea vuodelta 2010.

Vielä 1990-luvun alussa alle vuoden ikäisenä kuolleita oli vuosittain yli 350, vuoden 2000 jälkeen luku ei ole enää yhtenä vuonna ylittänyt kahtasataa.

Suomen keskimääräinen imeväiskuolleisuus vuosina 2011–13 oli toiseksi pienin Pohjoismaissa. Suomea pienempi imeväiskuolleisuus oli Islannissa, jossa se oli 1,3 promillea. Vuonna 2014 Ruotsin imeväiskuolleisuus oli yhtä suuri kuin Suomen.

[VÄESTÖ]

Syntyneiden määrä väheni yhä

Vuonna 2014 lapsia syntyi 57 232, mikä oli 902 vähemmän kuin edellisellä vuonna. Viimeksi syntyneiden määrä oli pienempi vuonna 2003, jolloin syntyi 56 630 lasta. Syntyneiden määrä on nyt pienentynyt neljänä vuonna peräkkäin.

Vuoden 2014 syntyvyyden mukaan nainen synnyttää keskimäärin 1,71 lasta, edellisen vuoden syntyvyyden mukaan 1,75 lasta. Vielä vuoden 2010 syntyvyyden mukaan nainen olisi synnyttänyt keskimäärin 1,87 lasta. Syntyvyys on ollut vuodesta 1969 lähtien alle väestön uusiutumistason, joka on 2,1 lasta naisesta kohden. ■

Lähde: Väestö- ja oikeustilastot. Tilastokeskus

Elinajanodote nousee edelleen

Kuolleiden määrä vuonna 2014 oli 52 190, mikä on 710 enemmän kuin vuonna 2013. Kuolleiden naisten määrä kasvoi 600 ja miesten 120 henkeä.

Kuolleiden määrä on ollut suurempi vain sotavuosina ja 1920-luvulla. Naisia on kuollut viimeksi 70 vuotta sitten enemmän kuin viime vuonna. Miehiä on kuollut viimeksi enemmän vuonna 1929, jos ei oteta huomioon sotavuosia.

Vastasyntyneen miehen elinajanodo-

te kasvoi 0,3 ja naisen 0,1 vuotta vuonna 2014. Vastasyntyneen elinajanodote kertoo vuosien määrän, jonka verran 0-vuotias eläisi kuolleisuuden pysyessä ennallaan. Vastasyntyneiden elinajanodotteen kasvu osoittaa, että huolimatta iäkkäiden henkilöiden määrän kasvusta ja väestön vanhenemisesta kuolleisuus on alentunut. ■

Lähde: Väestö- ja kuolemansytilastot. Tilastokeskus



Väestönmuutokset

	3. nelj. 2013	4. nelj. 2013	1. nelj. 2014	2. nelj. 2014	3. nelj. 2014*	4. nelj. 2014*
Elävänä syntyneet	15 406	13 738	13 944	14 552	15 343	13 240
Kuolleet	11 973	12 934	13 138	13 021	12 843	12 593
Syntyneiden enemmyys	3 433	804	806	1 531	2 500	647
Maahan muuttaneet	10 175	7 342	6 646	7 056	9 984	8 261
Maasta muuttaneet	4 827	3 025	3 502	2 927	4 988	2 995
Nettomuuttamuutos	5 348	4 317	3 144	4 129	4 996	5 266
Väestönlisäys ⁽¹⁾	8 781	5 007	3 950	5 660	7 496	5 913
Väkiluku⁽²⁾	5 445 699*	5 450 614*	5 455 220*	5 460 880*	5 468 376*	5 474 289*

* Ennakkotieto

⁽¹⁾ Sisältää rekisterikorjaukset lopullisten tietojen 4. neljänneksellä

⁽²⁾ Ajankohdan lopussa

[VAALIT]

Nuoremmissa kansanedustajissa jo naisenemmistö

Uusi eduskunta on edeltäjänsä hie- man nuorempi ja uuden eduskun- nan nuorekkuudesta vastaavat nai- set, kirjoittaa Tieto&trendit-blogissa **Miina Keski-Petäjä** Tilastokeskuksen tasa- arvotilastoista.

Valittujen naisten keski-ikä oli 44,4, miesten 49,3 vuotta. Alle 40-vuotiaista edustajista yli puolet oli naisia. 40–49-vuotiaissa puntit ovat lähes tasan. Miehiä on selkeä enemmistö 50–59-vuotiaista sekä erityisesti 60 vuotta täyttäneistä.

Vuosia kestänyt taantuma ei siis aiheuttanut samankaltaista notkahdusta sukupuolten välisissä osuuksissa kuin mitä 90-luvun laman jälkeisissä vaaleissa nähtiin, Keski-Petäjä kirjoittaa. Naisten osuus eduskuntaan valituista, annetuista äänistä ja ehdokkaista on pysytellyt samalla tasolla miltei koko 2000-luvun.

Vielä 1990-luvun laman aikaan nais-

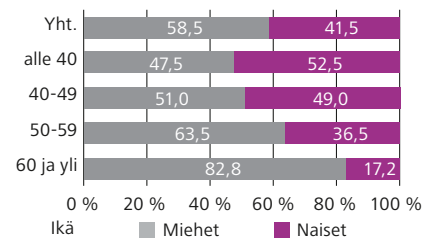
ehdokkaiden osuus, naisten saama ääniosuus ja eduskuntaan valittujen naisten osuus tipahtivat aiemmista vaaleista, vaikka naisten äänestysaktiivisuus laski miesten aktiivisuutta vähemmän.

Eniten naisia vasemmistolla ja kristillisillä

Uudessa eduskunnassa on suhteellisesti eniten naisia SDP:n, Kristillisdemokraattien ja Vasemmistoliiton valituissa, yli puolet kunkin ryhmän edustajista. Ehdokkaita näilläkin puolueilla oli naisia alle puolet.

Vain vihreillä oli ehdokkaissa naisenemmistö, valituissa on miehiä enemmän heilläkin. Naisten osuus valituista

Eduskuntaan valitut



on pienin Keskustalla, Perussuomalaisilla ja RKP:llä.

Naisia valittiin miehiä enemmän Lapin ja Varsinais-Suomen vaalipiireissä. Naisten osuus valituista on puolestaan pienin Keski-Suomessa (30 %), Vaasan vaalipiirissä (31 %) ja yllättäen Uudella- maalla (34 %). ■

Enemmän vaaleista myös Tilastokeskuksen karttapalvelussa: http://pxweb2.stat.fi/explorer/eduskuntavaalit_2015/index.shtml

Kommentoi blogissa:

tietotrenditblogi.stat.fi

[MATKAILU]

Lisää matkailijoita kotimaasta, vähemmän ulkomailta

Kotimaisten matkailijoiden yöpymisiä kirjattiin Suomen majoitusliikkeissä lähes 1,1 miljoonaa vuorokautta helmikuussa, mikä oli 4 prosenttia enemmän kuin viime vuonna. Sen sijaan ulkomaisten matkailijoiden yöpymiset vähenivät 1,8 prosenttia ja heille tilastoitiin majoitusliikkeissä vajaat 400 000 vuorokautta.

Yöpymisten kokonaismäärä majoitusliikkeissä oli runsaat 1,4 miljoonaa yöpymisvuorokautta, mikä oli 2,4 prosenttia enemmän kuin helmikuussa 2014.

Venäjältä helmikuussa jo vähemmän kuin Ranskasta

Venäläisyöpymisten jyrkkä lasku jatkui helmikuussa 2015. Heille kirjattiin Suomen majoitusliikkeissä 33 000 yöpymisvuorokautta, mikä oli peräti 54,5 prosenttia vähemmän kuin helmikuussa 2014.

Saksalaiset, britit ja ranskalaiset nou-

sivat venäläisten ohi suurimmiksi ulkomaisten matkailijoiden ryhmiksi helmikuussa. Saksalaisten ja ranskalaisten yöpymisvuorokaudet kasvoivat kummatkin kymmenisen prosenttia. Vuodenvaihteen matkailu pitää kuitenkin venäläiset vielä koko alkuvuoden tilaston kärjessä. ■

Lähde: Majoitustilasto. Tilastokeskus

Yöpymiset¹ kaikissa majoitusliikkeissä, tammi–helmik. 2015*

Yöpyneiden asuinmaa	Yöpymisvuorokaudet majoitusliikkeissä	
	Kaikki majoitusliikkeet	Yöpymisten vuosimuutos, %
Kaikki yhteensä	2 736 253	-2,6
Suomi	1 852 037	3,6
Ulkomaat	884 216	-13,4
Venäjä	212 014	-44,7
Britannia	86 366	1,6
Saksa	74 961	10,3
Ranska	66 760	9,3
Ruotsi	49 281	8,0
Alankomaat	39 468	-1,9
Japani	30 449	-14,9
Sveitsi	29 021	7,4
Viro	28 938	2,6
Yhdysvallat	21 032	7,4
Norja	20 421	23,2

* Ennakkotieto

Majoitustoiminta*¹

	Helmikuu 2015	Vuosimuutos %	Tammi-helmikuu 2015	Vuosimuutos %
Yöpymisiä kaikissa majoitusliikkeissä, 1 000	1 459,0	2,4	2 736,3	-2,6
ulkomaalaisten yöpymisiä, 1 000	396	-1,8	884,2	-13,4
Yöpymisiä hotelleissa, 1 000	1 244,4	3,5	2 351,3	-1,7
ulkomaalaisten yöpymisiä, 1 000	352	-0,6	761,4	-13,3
Hotellihuoneiden käyttöaste, %	51,4	2,0	47,6	0,1

* Ennakkotieto

¹ Hotellit ym. majoitusliikkeet, joissa on vähintään 20 vuodepaikkaa, tai sähköpistokkeella varustettua matkailuvaunu-/autopaikkaa sekä retkeilymajat

AKU ALANEN

T&k-investointien osuus on pudonnut huolestuttavasti

Investoinneista on pulaa koko EU:n alueella. T&k-menojen lukeminen investoinneiksi kansantalouden tilinpidon uudistuksen myötä nosti Suomen investointiastetta suhteessa useimpiin EU-maihin. Toisaalta vaatimattomatkin investoinnit ovat riittäneet pitämään asteen kohtuullisena, kun sen nimittäjäpuoli eli bkt on supistunut Suomessa viime vuosina.

Suomen investointiaste eli investointien suhde bruttokansantuotteen on ollut koko 2000-luvun EU:n keskitasoa korkeampi erityisesti kansantalouden tilinpidon uuden EKT2010-laskutavan mukaan. Tämä vaikuttaa hieman yllättävältä, kun ajatellaan viimeaikaista suomalaistakin keskustelua investointien vähäisyydestä. Kyse onkin siis koko EU:n tasoisesta investointilamasta eikä pelkästään meidän.

Tutkimus- ja kehittämistoiminnan ymmärtäminen investoinniksi on eniten investointiastetta nostanut tekijä uusissa bkt-laskelmissa. Toki viime vuonna käyttöön otettu EKT2010 toi myös muita metodisia muutoksia, mutta niiden mer-

kitys ei yllä lähellekään t&k:n vaikutusta.

Uudella tavalla laskien Suomen ero EU:n vanhojen jäsenmaiden keskitasoon kasvoi lähes kahteen prosenttiyksikköön (Taulukko 1). Suomi säilyttää asemansa vertailussa, vaikka EU:n keskitasona käytettäisiin vanhojen EU-maiden (EU15) sijasta EU28-maita tai euroaluetta. Maiden väliset erot ja vuosittaiset vaihtelut ovat suuria (Kuvio 1).

On muistettava, että investointiasteen uudessa laskentatavassa tutkimus- ja kehittämistoiminta on otettu mukaan sekä osoittajaan eli investointeihin että nimittäjään eli bruttokansantuotteen. T&k:n suhteellinen rooli on suurempi investoinneissa kuin bkt:ssa. Tä-

mä näkyy myös Suomen investointiasteen laskussa vuonna 2013, kun t&k-menot alkoivat Nokian mukana meillä vaihua (Kuvio 1). Huoli Suomen investointien suunnasta on siis aiheellinen.

Bkt:n arvoa t&k:n mukaanotto laskelmiin nosti suhteellisesti eniten EU-maista Ruotsissa, Suomessa ja Irlannissa. Irlannissa kyse oli pääosin ulkomailta globaalien yritysten omiin tytäryhtiöihin tuomasta t&k-panoksesta. Vastaavasti myös t&k-investoinnit nousivat näissä maissa suhteellisesti eniten.

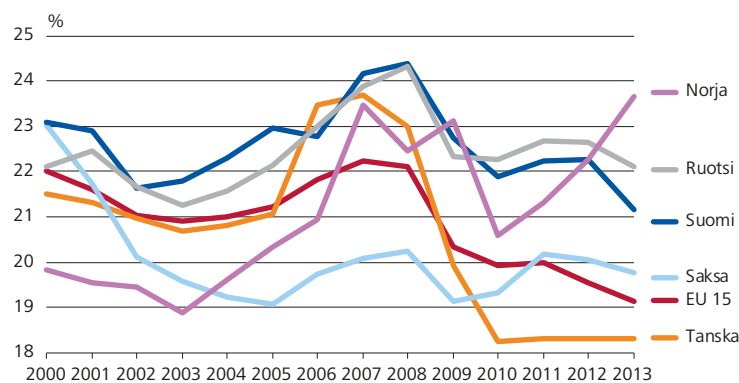
Suomen ja Ruotsin investointiasteet ovat kulkeneet suurin piirtein samaa tahtia viime vuosina. 2000-luvun alkuvuosina Suomen taso oli hivenen korkeam-

TAULUKKO 1. Investointiaste keskimäärin 2000–2013

	%
EU15	20,9
Tanska	20,7
Saksa	20,1
Suomi	22,6
Ruotsi	22,5
Norja	21,1

Lähde: Kansantalouden tilinpito. GDP and main components, ESA2010, Eurostat

Investointiaste 2000–2013, %



Lähde: Kansantalouden tilinpito. GDP and main components, ESA2010, Eurostat

TAULUKKO 2. Maakuntien investointiaste 2000-luvulla

	2000–12 %	2007–09 %	2010–12 %
Ahvenanmaa	26,5	29,1	35,7
Pohjois-Pohjanmaa	28,8	29,6	29,2
Kainuu	26,0	34,1	28,7
Pirkanmaa	24,2	24,2	24,8
Etelä-Pohjanmaa	22,9	23,4	23,5
Lappi	24,8	26,7	23,4
Keski-Suomi	23,5	23,8	22,8
KOKO MAA	22,7	23,8	22,1
Uusimaa	22,5	24,1	21,4
Etelä-Savo	20,6	18,8	21,4
Keski-Pohjanmaa	22,0	27,2	21,4
Varsinais-Suomi	22,0	22,4	21,2
Etelä-Karjala	20,5	20,8	20,6
Pohjois-Savo	21,5	23,7	20,4
Pohjois-Karjala	20,8	21,3	20,2
Kymenlaakso	22,6	23,9	20,1
Päijät-Häme	20,0	20,2	19,6
Pohjanmaa	20,2	19,2	19,5
Kanta-Häme	21,6	21,8	19,3
Satakunta	18,8	20,2	18,4

Lähde: Aluetilinpito, Tilastokeskus

pi, kun taas taluskriisin jälkeen Ruotsi on ollut hieman edellä. Muista pienistä maista myös Itävallan investointiaste on viime vuosina noussut Suomen edelle, vaikkei Itävallassa t&k:n osuus olekaan meidän tasoamme.

Suomen osalta myös heikko yleinen talouskehitys selittää investointiastetta. Kun nimittäjä eli bkt laskee, niin osoittajan ei tarvitse olla kummoinkaan suhdeluvun silti pysyessä kohtuullisena.

Vain asuntoinvestoinnit ovat kasvaneet kriisin aikana

Viime vuosien lama on ollut Suomen investoinneille jopa kohtalokkaampi kuin tuotannolle. Ainoa merkittävä investointityyppi, jonka volyyymi on kasvanut vuodesta 2008, on ollut asuntoinvestoinnit. Ne muodostavat vajaan kolmasosan kaikista investoinneista.

Itse asiassa asuntoinvestointien määrä on ollut t&k:takin tärkeämpi tekijä investointiasteen pysymisessä suhteellisen

korkealla tasolla. Ilman asuntoinvestointeja romahdus olisi ollut vielä rajumpi.

Kokonaisinvestointien volyyymi oli vuonna 2013 peräti 15 prosenttia pienempi kuin vuonna 2008. Käyvin hinnoin laskien pudotus on hieman pienempi, 10 prosenttia.

Vähiten, 5 prosenttia, on laskenut kuljetusvälineiden investointien volyyymi. Pienemmistä eristä myös tietokoneiden hankintainvestoinnit ovat hivenen kasvaneet, mutta niiden osuus on vain prosentin luokkaa kaikista investoinneista.

Kaikkein eniten investoinnit ovat vähentyneet t&k-toiminnassa sekä muussa talonrakennuksessa. T&k-investoinnit ovat supistuneet neljänneksen, liike- ym. rakentaminen kolmanneksen verrattuna taluskriisiä edeltäneeseen tilanteeseen. T&k-investointien osuus on ollut viime vuosina noin 15 prosentin luokkaa kaikista eli samalla tasolla muiden kone- ja laiteinvestointien kanssa.

T&k-investoinnit ovat vähentyneet vuoden 2008 jälkeen myös käyvin hin-

Aineettoman pääoman rooli kasvaa investointilaskelmissa

Pääoman käsitteen konkretisoiminen on yksi kaikkein tärkeimpiä taloustieteen ja myös taloustilastojen kysymyksiä. Taloustieteessä pääoman määrittelystä on kiistelty kautta aikojen ja käsitykset ovat muuttuneet. Myös taloustilastoissa on tapahtunut muutoksia siinä, mitä asioita nimitetään ja lasketaan pääomaksi. Jos joku menoerä katsotaan pääoman osaksi, joudutaan sen osalta tekemään myös tilastomittaukset investointina, poistoina jne.

Nykykaikaisessa kansantalouden tilinpidossa pääomaksi ja sitä myöten investoinneiksi katsottiin aluksi pääosin vain aineellisia asioita kuten rakennuksia, teitä ja koneita. Pikkuhiljaa myös uusien aineettomien asioiden kuten tietokoneohjelmien katsottiin kuuluvan pääoman piiriin.

Vuonna 2014 myös t&k-laskelmien katsottiin olevan kypsiä saamaan investointien nimityksen. Tä-

täkin siirtoa valmisteltiin hartaasti ja pitkään. On kiinnostavaa seurata, mikä aineettoman pääoman elementti voisi olla seuraavana vuorossa pääsemään osaksi pääomalakentaa. Tarjokkaita riittää sekä inhimillisen pääoman, sosiaalisen pääoman että luontopääoman alueilla.

Periaatteessahan mitä tahansa talouden menoerää voi pitää pääoman lisäyksenä, investointina, jos se on tarkoitettu lisäämään tulevaa kulutusta nykyisen sijaan. Voi toki olla, että uusien elementtien hyväksyminen investoinneiksi lykkääntyy pitkälle tulevaisuuteen, koska ei päästä sopuun niiden laskentamenetelmistä. Lisäksi kovin mielellään ei tehdä isoja metodisia muutoksia vertailukelpoisuuden säilyttämisen takia. Toistaiseksi niistä tehdäänkin laskelmia ja tutkimuksia irrallaan kansantalouden tilinpidon virallisista laskelmista.

noin noin 13 prosenttia. Itse asiassa käypähintaiset t&k-investoinnit olivat vielä vuonna 2013 alemmalla tasolla kuin vuonna 2004. Johtopäätös on selvä, vaikka Suomen t&k-panostus on edelleen maavertailussa kohtuullinen: nykytrendin valossa t&k-investointien tilanne on vakava pitemmän päälle.

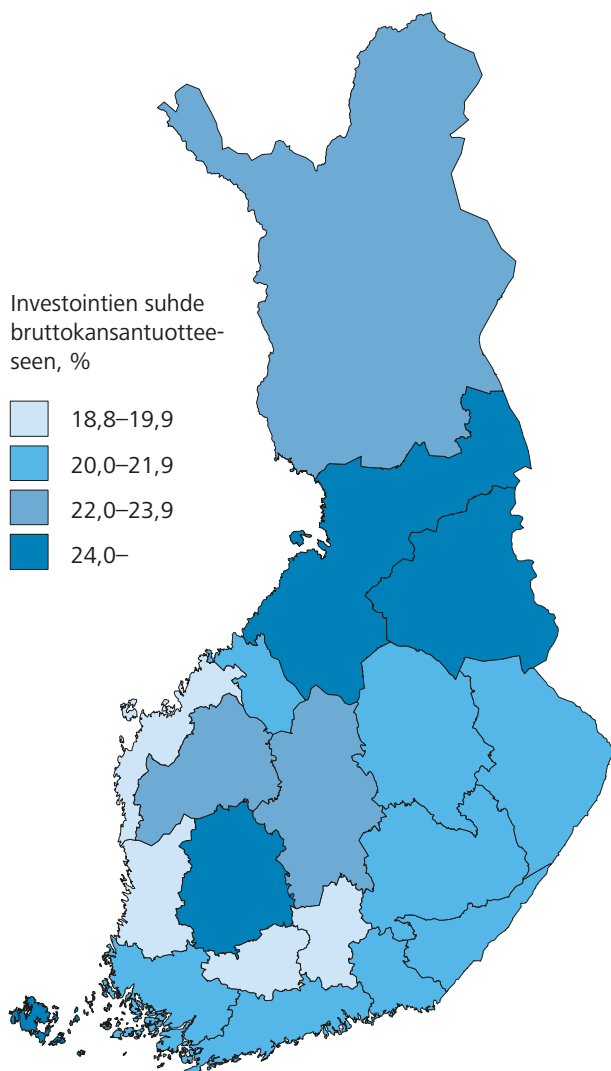
Investointiaste nousi t&k:n pääomittamisen myötä kaikilla alueilla

Suomen maakuntien investointiaste (kiinteän pääoman bruttomuodostuksen suhde bkt:hen) on noussut asuntoinvestointien ja t&k:n pääomittamisen johdosta keskimäärin muutaman prosentin tällä vuosituohannella. Voimakkainta kasvua on siellä, missä t&k:n suhteellinen merkitys on suurin eli Pohjois-Pohjanmaalla, Varsinais-Suomessa ja Pirkanmaalla. (Taulukko 2)

Taluskriisi on johtanut vuosittaisten investointiasteiden suureen vaihteluun, ▶

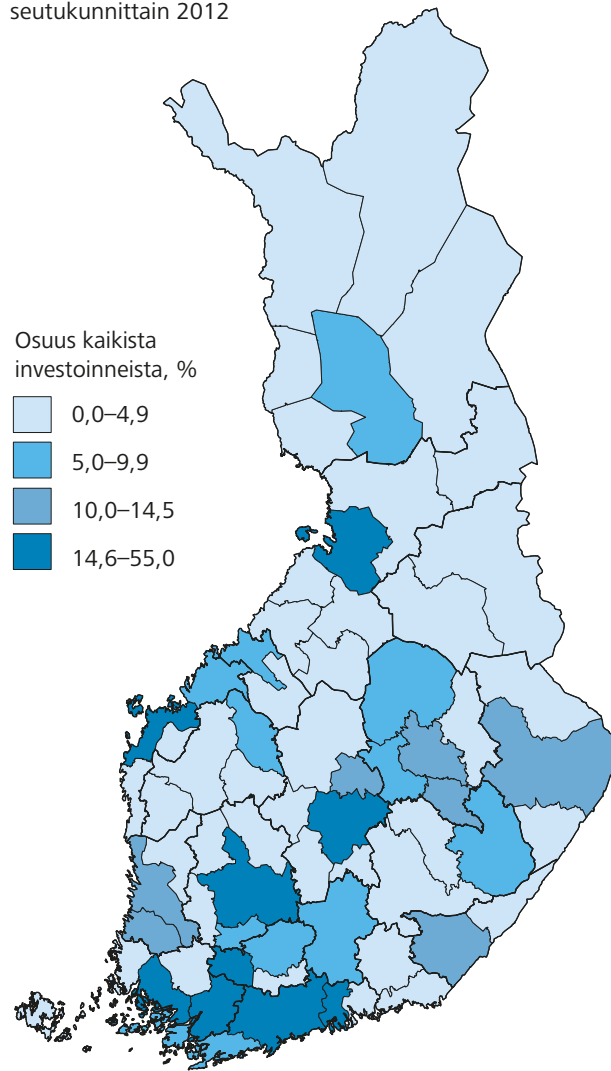
[INVESTOINNIT]

Investointiaste 2010–12



Kartta 1

Tutkimus- ja kehittämistoimen investoinnit seutukunnittain 2012



Kartta 2

Lähde: Aluetilinpito, Tilastokeskus

koska sekä nimittäjä että osoittaja ovat vaihdelleet voimakkaasti.

Investointiasteen alueellinen taso selittyy monilla tekijöillä. Ahvenanmaan muita korkeampi investointiaste johtuu meriliikenteen kuljetusvälineiden hankinnan suuresta arvosta muuten pienellä alueella. Muilla alueilla kuljetusvälineiden osuus investoinneista liikkuu viiden prosentin tienoilla. Kainuun korkea sijoitus taas johtuu kaivosteollisuuden buumista. Sen sijaan Pohjois-Pohjanmaalla ja Pirkanmaalla juuri t&k selittää niiden sijoitusta, vaikka muutkin investoinnit ovat mittavia. (Kartta 1)

Itä-Suomen alueilla ja Päijät-Hämeessä, Pohjanmaalla ja Keski-Suomessa asuininvestoinnit ovat taas suhteellisesti suuremmat kuin muualla. Näillä alueilla ne vastaavat reilusti yli kolmannelle kaikista investoinneista.

Kone- ja laiteinvestoinnit edustavat kaikkiaan vain viidennestä kaikista investoinneista, ja niiden osuus pysyttelee useimmilla alueilla vuosittain suurin piirtein samalla tasolla. Silti joillakin alueilla tapahtui isojakin koneinvestointeja vuonna 2012. Kone- ja laitehankintoja tehtiin suhteellisesti eniten Lapissa vuonna 2012. Siellä teollisuuden

ja kaivostoiminnan koneiden ja laitteiden investoinnit kolminkertaistuivat aiempaan verrattuna ja muodostivat noin 40 prosenttia kaikista investoinneista. Myös Kymenlaaksossa, Kanta-Hämeessä ja Päijät-Hämeessä kasvoivat teollisuuden kone- ja laiteinvestointien määrät voimakkaasti.

Pohjois-Pohjanmaa ollut kärjessä t&k:n osuudessa

Yliopistojen sekä yritysten t&k-keskusten sijainti selittää alueiden suhteellista asemaa, kun verrataan t&k-investointien

TAULUKKO 3. T&k-investointien osuus alueen kaikista investoinneista, %

	2010	2011	2012
Pohjois-Pohjanmaa	29,5	25,6	23,1
Varsinais-Suomi	23,6	22,7	22,6
Pirkanmaa	30,5	26,2	19,9
Uusimaa	18,7	18,6	17,2
Pohjanmaa	16,0	13,4	15,7
KOKO MAA	17,3	15,7	14,6
Keski-Suomi	11,1	10,3	12,4
Satakunta	8,3	6,6	10,1
Etelä-Karjala	9,9	9,4	9,6
Pohjois-Savo	10,7	9,3	9,5
Pohjois-Karjala	9,4	7,8	8,8
Kanta-Häme	8,0	7,7	6,9
Päijät-Häme	5,5	4,9	5,7
Keski-Pohjanmaa	3,9	3,2	4,7
Lappi	5,2	5,0	4,5
Etelä-Savo	3,8	3,7	4,4
Kainuu	3,7	3,2	3,8
Etelä-Pohjanmaa	3,0	3,5	3,3
Kymenlaakso	2,6	2,5	1,9
Ahvenanmaa	0,7	0,5	0,7

TAULUKKO 4. T&k-investointien osuus aineettomista investoinneista maakunnittain, %

	2010	2011	2012
Uusimaa	66,7	65,2	60,2
Varsinais-Suomi	80,1	79,6	82,5
Kanta-Häme	63,7	63,0	78,0
Päijät-Häme	50,0	55,2	54,1
Kymenlaakso	33,9	35,9	68,8
Etelä-Karjala	65,0	67,7	73,6
Satakunta	63,2	60,7	66,4
Pirkanmaa	82,6	81,6	80,4
Keski-Suomi	71,3	73,8	84,8
Etelä-Pohjanmaa	43,9	54,9	53,2
Pohjanmaa	73,4	75,7	76,5
Etelä-Savo	56,3	57,5	62,2
Pohjois-Savo	72,4	70,8	76,4
Pohjois-Karjala	72,8	60,2	69,8
Keski-Pohjanmaa	50,6	55,3	59,7
Pohjois-Pohjanmaa	86,2	83,8	80,1
Kainuu	37,8	36,8	45,6
Lappi	50,1	48,3	53,3
Ahvenanmaa	10,4	18,3	22,8
KOKO MAA	71,9	70,6	68,5

Taulukoiden lähde: Aluetilinpito, Tilastokeskus

osuutta kaikista investoinneista. Suomen keskitason ylittävät vain ne maakunnat, joissa on yksi tai useampi yliopisto (Taulukko 3). Niillä on monesti myös yhteyksiä isojen yritysten toimintaan. Keskitason alle jää vain Lappi, jonka yliopistossa ei ole teknologista tiedekuntaa.

Seutukuntatasolla it-talouden ja korkeakoulujen rooli näkyy myös selvästi kartassa 2. Korkeimmillaan t&k-investointien osuus on niissä seutukunnissa, joissa on sekä korkeakouluja että it-taloutta. Esimerkiksi Oulun seutukunta on seutukuntien kakkostilalla, mutta Pohjois-Pohjanmaan muissa seutukunnissa t&k:n rooli on hyvin vaatimaton.

Valtaosa ns. aineettomista investoinneista kohdistuu tutkimukseen ja kehittämiseen. T&k-investoinnit ovat olleet muita aineettomia investointeja suurempia useissa maakunnissa viime vuosina (Taulukko 4).

Uudellamaalla ohjelmointi-investoinnit ovat suhteellisesti suurimmat, kymmenesosa kaikista investoinneista. Seu-

raavina tulevilla Pohjois-Pohjanmaalla ja Pirkanmaalla ne olivat suhteessa puolta pienempiä, viisi prosenttia kaikista investoinneista. Vain Kymenlaaksossa ja Kainuussa ohjelmointi-investointien osuus alittaa yhden prosentin.

Kaikilla alueilla t&k-investointien taso ylittää huomattavasti ohjelmointi-investointien tason. Näin on myös Uudellamaalla, vaikka siellä tehdään yli 60 prosenttia kaikista ohjelmointi-investoinneista. Tämä johtuu siitä, että lähes puolet kaikista t&k-investoinneista tehdään Uudellamaalla.

Mineraalien etsintää pidetään myös aineettomana investointina. Etsintä on viime vuosina kohdentunut pohjoiseen, ennen kaikkea Kainuuseen, Pohjois-Pohjanmaalle ja Lappiin.

Taiteen ja viihteen alkuperäisteoksiin tehdyt investoinnit kohdentuvat lähes täysin (3/4) Uudellemaalle. Asehan-kintoihin kohdistuneita investointeja on absoluuttisesti eniten suuntautunut Kymenlaaksoon.

Pientä valoa tunnelissa pienemmissä maakunnissa

Miljarditason t&k-investointialueilla, Uudellamaalla, Pirkanmaalla ja Pohjois-Pohjanmaalla, investoinnit ovat vähentyneet vuoden 2009 laman pohjan jälkeen kussakin maakunnassa 200-300 miljoonaa euroa eli yhteensä lähes 800 miljoonaa euroa.

Pieni positiivinen oire paremmasta on se, että muilla alueilla sinänsä vaatimattomat, yleensä muutamien miljoonien tai kymmenien miljoonien t&k-investoinnit ovat kasvaneet vuoden 2009 jälkeen. ■

Kirjoittaja työskentelee Kansantalouden tilinpidon Aluetilinpidoissa

Enemmän bkt-laskennan uudistuksesta ks. Tieto&trendit 3-4/2014.

Viennin kasvu vahvistaisi luottamusta koko taloudessa

Lehtikuva Oy

PASI KUOPPAMÄKI

Avainroolia viennin kasvuun saamisessa näyttelee länsimaiden talouksien voimistuminen, painottaa Danske Bankin pääekonomisti Pasi Kuoppamäki.

Suomen bruttokansantuote supistui kolmatta vuotta putkeen vuonna 2014. Talouden kasvunäkymät pysyvät heikkoina myös 2015–2016. Uutistoimisto Bloombergin keräämien lukujen mukaan suhdanneennustajat odottavat bruttokansantuot-

teen kasvavan keskimäärin 0,5 prosenttia vuonna 2015 ja 1,3 prosenttia vuonna 2016. Näyttää melko todennäköiseltä, että Suomen taloudessa koetaan menetetty vuosikymmen, jolloin bruttokansantuote ei saavuta finanssikriisiä edeltänyttä vuoden 2008 huipputasoa

ennen tämän vuosikymmenen loppua.

Sekä Suomen Pankin että valtiovarainministeriön arviot Suomen talouden keskipitkän aikavälin kasvunäkymistä jäävät lähelle 1 prosentin vuotuista kasvua, ellei talouden rakenteellisia heikkouksia saada korjattua. Heikot kasvunäky-

mät vaikuttavat haitallisesti niin hyvinvointivaltion rahoituspohjaan kuin yritysten investointihalukkuuteen, mikä lisää negatiivisen kierteen riskiä.

Heikon taloustilanteen takaa löytyy iso joukko tekijöitä: esimerkiksi painopaperin kysynnän supistuminen, Nokian kännykkäliiketoiminnan vaikeudet, viennin supistuminen Venäjälle, kaupakumppaneita nopeammin nousseet yksikkötyökustannukset, väestön ikääntyminen ja julkisen sektorin nopean vakaantumisen aiheuttamat rajoitteet finanssipolitiikalle. Osa näistä asioista jää hallituksen vaikutusmahdollisuuksien ulkopuolelle, mutta esimerkiksi teollisuuden kilpailukykyyn tai verotuksen kannustinvaikutuksiin voidaan vaikuttaa talouspoliittisin päätöksin.

Suomesta löytyy myös paljon vahvuuksia, kuten hyviä oppimistuloksia tuottanut peruskoulutus, erittäin alhainen korkotaso niin kotitalouksille kuin yrityksille sekä vähäinen korruptio. Vahvuuksia ajatellen onkin toisaalta paradoksaalista, että Suomen talouden näkymät ovat EU:n heikoimpia.

Suomen talous on avoin talous, vaikka vienti suhteessa bruttokansantuotteeseen on laskenut huippuvuoden 2008 yli 44 prosentista 36 prosenttiin vuonna 2014. Suomen tavaroiden ja palveluiden viennin volyymi supistui 0,4 prosenttia vuonna 2014. Tavaravienti kasvoi 1,1 prosenttia, mutta palveluiden vienti supistui 4 prosenttia (Viennin kehityksestä tarkemmin alkaen s. 19).

Talushistoriassa Suomen suhdannevaihtelut ovat usein aiheutuneet viennissä koetuista vaikeuksista tai menestyksistä. Tälläkin kertaa monet poliitikot toi-

VAHVUUKSIA AJATELLEN
ON PARADOKSAALISTA,
ETTÄ SUOMEN TALOUDEN
NÄKYMÄT OVAT
EU:N HEIKOIMPIA.

vovat ja ennustajat ennustavat viennin toimivan talouden elpymisen moottorina. Viennin merkitystä tosin usein liioitellaan, sillä Suomi ei vie 36 prosenttia kokonaistuotannosta, vaikka vienti suhteessa bruttokansantuotteeseen onkin 36 prosenttia. Ulkomailta hankittujen välituotteiden huomioimisen jälkeen kotimaisen arvonlisän osuus, josta maksetaan mm. palkat Suomessa, on alle 30 prosenttia bruttokansantuotteesta. Läpikulkumaana toimivassa Irlannissa bruttovienti on yli 100 prosenttia suhteessa bruttokansantuotteeseen.

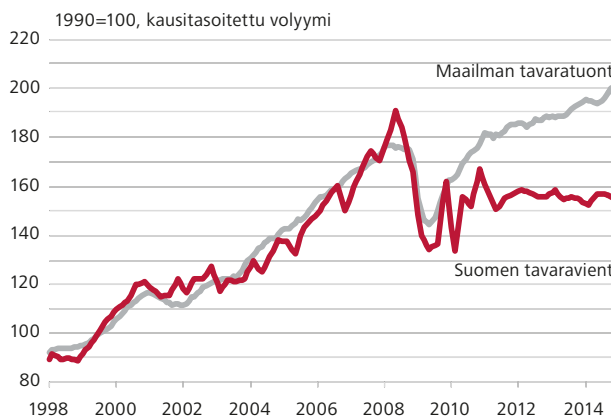
Useat muut euroalueen maat, esimerkiksi Saksa ja Espanja, ovat jo pystyneet kasvattamaan vientiä, mikä vähitellen ruokkii vientiteollisuuden investointien ja työllisyyden kautta myös kotimaista kysyntää. Suomen läntisten vienti-

markkinoiden kasvun odotetaan kiihtyvän tänä vuonna. Heikentynyt euro tukee viennin hintakilpailukykyä. Suomen vienti Yhdysvaltoihin ja Saksaan onkin kasvanut selvästi. Saksa ohitti Venäjän Suomen suurimpana kaupakumppanina 2014 ja Yhdysvallat ohitti Venäjän kolmanneksi suurimpana vientikohteena joulutammikuussa.

Valitettavasti Suomen vienti Venäjälle supistuu edelleen. Lisäksi vienti polkee paikoillaan, vaikka maailmankauppa kasvaa (Kuvio 1). Suomen markkinaosuus supistui Ruotsin ja Ison-Britannian tuonnista 2014, vaikka niiden taloudet olivat kasvussa. Venäjän ruplan ja Ruotsin kruunun heikkous vähentävät Suomen saamaa hyötyä dollarin vahvistumisesta. Tullin ulkomaankauppatilastojen mukaan Suomen koko viennin ar-

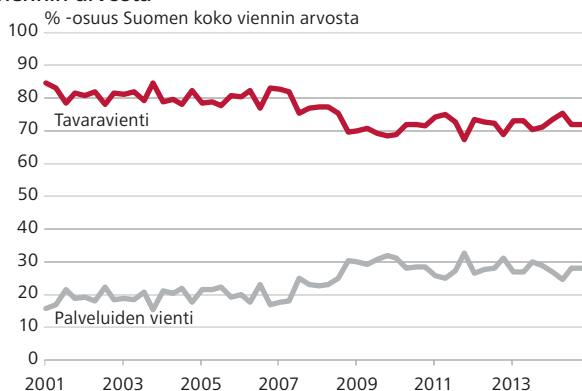
VENÄJÄN RUPLAN JA RUOTSIN KRUUNUN HEIKKOUS VÄHENTÄVÄT SUOMEN SAAMAAN HYÖTYÄ DOLLARIN VAHVISTUMISESTA.

KUVIO 1. Maailmankauppa ja Suomen vienti

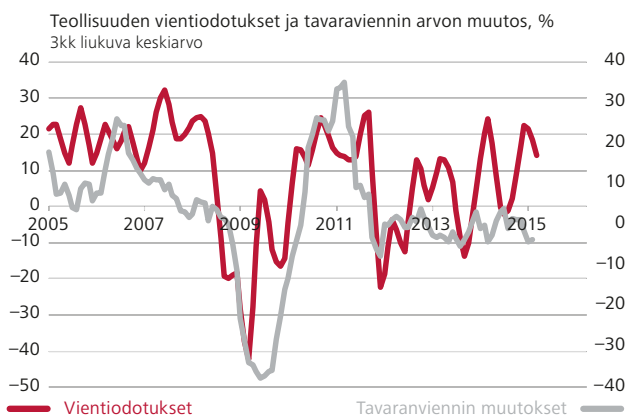


Lähde: Macrobond

KUVIO 2. Tavaroiden ja palveluiden osuus Suomen koko viennin arvosta



KUVIO 3. Tavaraviennin kasvu ja teollisuuden vientitodotukset



Kuvioiden lähde: Macrobond

vo putosi 7 prosenttia ja Tilastokeskuksen mukaan teollisuuden uudet tilaukset supistuivat 3,3 prosenttia tammi-helmikuussa 2015.

Suomen heikkoa vientimenestystä saattavat selittää painottuminen painopaperin kaltaisiin taantuneisiin hyödyke-ryhmiin, heikentynyt kustannuskilpailukyky, suuryritysten päätös siirtää tuotantoa kolmansiin maihin tai kilpailijoita vaatimattomampi markkinointiosaaminen.

Hallitus ja työmarkkinajärjestöt ovat pyrkineet parantamaan toimintaympäristöä laskemalla yhteisöverotusta alle EU:n keskiarvon ja sopimalla maltillisen palkkaratkaisun, mutta nämä vaikuttavat hitaasti eivätkä poista kaikkia ongelmia.

Aivan kaikkea tuotantoa ei kannata Suomessa kynsin ja hampain pitääkään, esimerkiksi Nokian kännykkätuotanto muuttui vähitellen arvoa tuhoavaksi, mutta toki vahva teollinen pohja luo edellytyksiä myös muulle taloudelliselle aktiiviteetille. Palveluista on toivottu uutta kasvua vientiin ja niiden osuus kokonaisviennistä onkin noussut jo noin 30 prosenttiin (Kuvio 2).

Valitettavasti palveluiden vienti supistui 2013–2014. Palveluiden viennistä kertovat tilastot ovat hankalasti tulkittavia, mutta ainakin Nokian kännykkäliiketoiminnan tyrehtyminen näkyy myös palveluiden viennin heikkenemisenä. Toisaalta peliteollisuuden liikevaihdon vahva nousu kertoo siitä, miten nopeasti palveluiden saralla voidaan saavuttaa menestystä.

PELITEOLLISUUDEN
LIIKEVAIHDON VAHVA
NOUSU KERTOO SIITÄ,
MITEN NOPEASTI
PALVELUIDEN SARALLA
VOIDAAN SAAVUTTA
MENESTYSTÄ.

Sähkö- ja elektroniikkateollisuuden alamäki ja kemian teollisuuden kasvu ovat tasoittaneet Suomen viennin toimialarakennetta viime vuosina. Vuonna 2014 osuuttaan viennissä kasvatti kone- ja kulkuneuvoteollisuus. Mielenkiintoinen yksityiskohta Tullin ulkomaankauppatilastoissa on sähkö- ja elektroniikkateollisuuden syöksyn taantumisen nousuun 2014, mikä kertoo Nokian romahduksen vaikutusten vihdoin olevan takana päin tavaraviennissä.

Venäjän talouden taantumisen seu-

rauksena avainroolia viennin kasvuun saamisessa näyttelee länsimaiden talouksien voimistuminen. Suhteellisten yksikkötyökustannusten nousun tuoma kilpailuhaitta vähenee maltillisten palkankorotusten ansiosta lähivuosina. Heikosta alkuvuoden tavaraviennistä huolimatta on mahdollista, että vienti kasvaa hieman vuonna 2015. Investointien vilkastuminen Euroopassa voisi tuoda tilauksia myös Suomen teollisuudelle, missä investointihyödykkeillä on perinteisesti ollut suuri paino. Teollisuuden vientitodotukset ovatkin nousseet normaalimmalle tasolle (Kuvio 3).

Kotimaisen kysynnän kannalta näkyvät pysyvät vaatimattomina. Inflaation painuminen nollan tuntumaan ja matalat korot näyttäisivät piristäneen kuluttajien luottamusta, mutta ostovoima pysyy vaatimattomana. Työllisyydessä ei odoteta merkittävää nousua ja palkankorotukset pysyvät matalina. Valtion ja kuntien velkaantuminen sekä väestön ikääntymisen tuomat kustannuspaineet aiheuttanevat mittavia julkisten menojen leikkauksia lähivuosina. Rakennusalan ja teollisuuden investoinnit näyttävät nekin vaatimattomilta. Kyseessä näyttäisi olevan enemmänkin luottamus-pula kysynnän kasvuun kuin resurs-sipula. Esimerkiksi pörssiyritykset pystyvät maksamaan suuret osingot ja työvoimaa-kin olisi tarjolla. Kasvava vienti saattai-si piristää luottamus-pulasta kärsiviä yri-tyksiä ja kotitalouksia, jolloin myös koti-mainen kysyntä voimistuisi. ■

CHRISTINA TELASUO

Vienti ei vedä vielääkään

Suomen tavaraviennin arvon kehityksessä on ollut historiallisen pitkä stagnaation aika. Viimeisen kolmen vuoden ajan viennin arvo ei ole kasvanut lainkaan tai on jopa hieman vähentynyt. Venäjän viennin kutistuminen synkentää myös kuluvan vuoden näkymiä.

Suomen talouden nousua on odotettu jo kolmen vuoden ajan. Pienelle avoimelle taloudelle viennin elpyminen on usein talouskasvun kannalta avainasemassa. Finanssikriisin aiheuttamasta vuoden 2009 viennin romahduksesta on Suomessa toivuttu vain osittain. Viennin arvo on yhä alhaisempi kuin vuonna 2006. Viime vuonna viennin arvo laski alle puoli prosenttia. Vuonna 2013 se laski yhden prosentin ja vuonna 2012 viennin arvo pysyi edellisvuoden tasolla.

Kauppataaseissa oli viestintäteknologian viennin ansiosta poikkeuksellisen suuria ylijäämiä vuodesta 1991 aina vuoteen 2010 asti. Matkapuhelinten viennin arvo syöksyi vuoden 2007 lähes seitsemästä miljardista eurosta neljässä vuodessa 1,4 miljardiin euroon. Vuonna 2011 kauppatase romahti 3,7 miljardia alijäämäiseksi. Sen jälkeen alijäämät ovat pienentyneet ja viime vuonna vajetta oli 1,8 miljardia euroa. Kauppataseen ylijäämät olivat parhaimpina vuosina 12–15 prosenttia kokonaiskauppavaihdon arvosta. Vuonna 2011 kauppataseen alijäämä oli suurimmillaan ja noin kolme prosenttia kokonaiskauppavaihdosta, mutta viime vuonna enää 1,6 prosenttia.

Pelkästään viennin arvon tarkasteleminen antaa kuitenkin yksipuolisen kuvan viennin kehittymisestä, varsinkin nyt kun yleinen hintataso on laskussa. EU-maiden inflaatio on ollut alhainen jo pitkään, ja kääntyi deflaation puolelle joulukuussa. Suomenkin inflaatio painui miinukselle alkuvuodesta.

Vientihinnoissa laskua jo vuodesta 2012

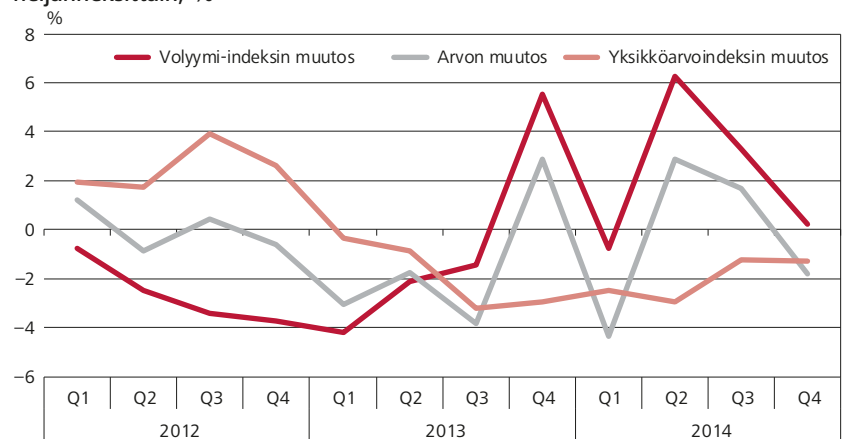
Vientihintojen lasku alkoi jo vuonna 2012. Vuonna 2013 sitä kertyi 1,9 prosenttia viennin yksikköarvoindeksillä mitattuna. Viime vuonna vientihinnat alenivat kaksi prosenttia. Tuontihinnat laskivat myös, jopa enemmän kuin vientihinnat. Kun viennin arvo on laskenut vähemmän kuin vientihinnat, on viennin volyyymi noussut.

Viime vuonna viennin volyyymi kasvoi 2,3 prosenttia volyyymi-indeksillä mitattuna. Volyymin kasvu alkoi jo vuoden 2013 viimeisellä neljänneksellä ja jatkui ensimmäisen neljänneksen pientä notkaidusta lukuun ottamatta koko vuoden 2014.

Öljyn hinnanlasku vaikutti vienti- ja tuontihintojen laskuun eniten viime kesästä lähtien. Suomen kokonaisviennistä yli kymmenen prosenttia oli öljytuotteita viime vuonna. Öljytuotteiden viennin arvo aleni 13 prosenttia, kun taas raakaöljyn tuonnin arvo supistui kymmenen prosenttia. Raakaöljyn tuonnin volyyymi pieneni vain 2,6 prosenttia, mutta öljytuotteiden viennin volyyymi väheni yli kuusi prosenttia viime vuonna.

Öljyn hinnanlasku on jatkunut kuluvaan vuoteen puolella ja tulee vaikuttamaan yhä sekä viennin että tuonnin arvoihin alentavasti. Kuluvaan vuoden tammiukuussa öljytuotteiden viennin arvo kutistui yli 40 prosenttia.

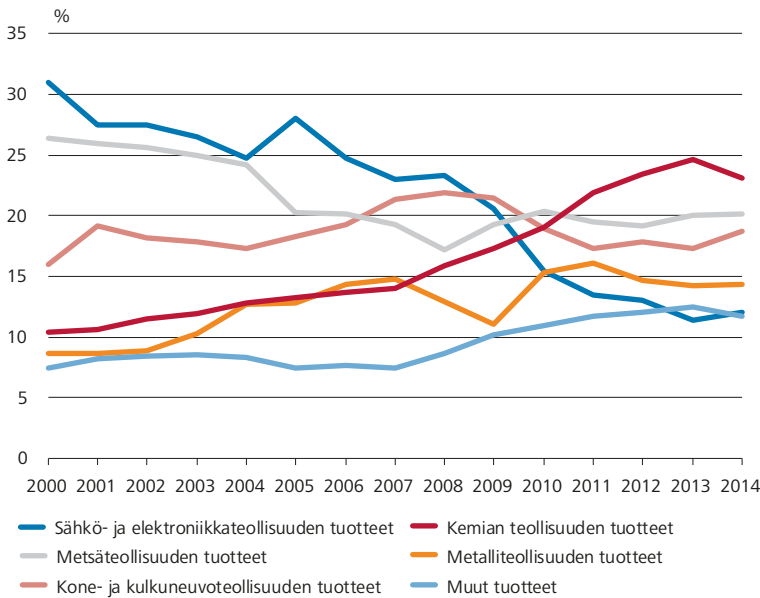
KUVIO 1. Suomen viennin volyyymi-indeksin, yksikköarvoindeksin ja arvon muutos neljänneksittäin, %



Lähde: Tulli

Huom! Tullin volyyymi-indeksi perustuu yksikköarvoindeksiin, Tilastokeskuksen volyyymi-indeksi vientihintaindeksiin.

KUVIO 2. Eri toimialojen tuotteiden osuudet koko viennistä 2000–2014



Lähde: Tulli

Kauppataaseen kehitykseen öljyn hinnanlasku vaikuttanee sen sijaan positiivisesti tänä vuonna. Suomen tuonnista 17 prosenttia on raakaöljyä ja öljytuotteita ja viennistä vain noin 10 prosenttia on öljytuotteita. Raakaöljyn tuontihinnat ovat laskeneet enemmän kuin öljytuotteiden vientihinnat. Vaikutus tuskin on kuitenkaan kovin järjestyttävä, koska öljytuotteil-

la on merkittävä osuus kokonaisviennistä. Toinen merkittävä tekijä, jolla oli suuri vaikutus Suomen viennin kehitykseen viime vuonna, oli vienti Venäjälle. Maalla on ollut viimeiset vuodet lähes kymmenen prosentin osuus Suomen kokonaisviennistä. Vuonna 2014 osuus pieneni kahdeksaan prosenttiin. Suomen vienti Venäjälle supistui viime vuonna

14 prosenttia edellisvuoteen verrattuna ja tuonti 18 prosenttia. Energiatuotteiden osuus koko tuonnista Venäjältä oli lähes 80 prosenttia viime vuonna.

Viennin rakenne monipuolistunut

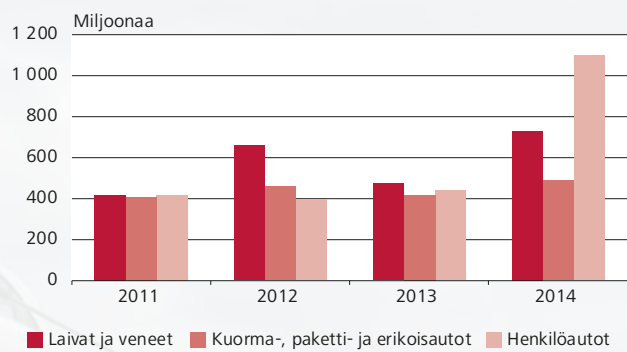
Suomen viennin rakenteessa ei ole tapahtunut suuria muutoksia viimeisten vuosien aikana. Viennin rakenne on kuitenkin huomattavasti tasaisempi verrattuna vuosituhannen alkuun, jolloin sähkö- ja elektroniikkateollisuuden viennillä oli yli 30 prosentin osuus ja metsäteollisuuden tuotteiden viennillä yli neljänneksen osuus kokonaisviennistä.

Viime vuonna kemianteollisuus oli suurin yksittäinen teollisuudenala öljytuotteiden viennin suuren osuuden vuoksi. Metsäteollisuustuotteet olivat yhä toiseksi suurin vientisektori, vaikka kone- ja kulkuneuvoteollisuuden tuotteiden osuus oli lähes yhtä suuri. Sähkö- ja elektroniikkateollisuuden viennin osuus kääntyi viime vuonna kasvuun usean heikomman vuoden jälkeen.

Viennin perinteisen kivijalan, puu- ja paperiteollisuuden vienti kokonaisuudessaan kehittyi vielä viime vuoden alussa myönteisesti, mutta kääntyi sittemmin loivaan laskuun. Tämä johtui ennen kaikkea paperin viennin vähemisestä. Sekä paperimassan että mekaa-

Uudenkaupungin tehtaan valmistamille Mercedes-Benzin A-sarjan autoille riitti kysyntää Saksassa viime vuonna.

KUVIO 3. Kulkuneuvojen vienti 2011–2014



Lähde: Tulli

Lehtikuva Oy

KUVIO 4. Viennin muutos Venäjälle ja muihin maihin neljänneksittäin



Lähde: Tulli

nisen metsäteollisuuden vienti lisääntyi, puutavaran vienti näistä eniten.

Koneiden ja laitteiden viennin supistuminen alkoi jo vuonna 2012, viime vuonna niiden vienti ei enää vähentynyt yhtä paljon kuin edellisellä vuonna. Kemianteollisuuden viennin kasvun esteenä oli öljytuotteiden hinnanlasku ja lääketieteellisuuden viennin pienentyminen.

Sähkö- ja elektroniikkateollisuuden tuotteiden sekä metallien vienti piristyi viime vuonna. Erityisesti mobiiliteknologian laitteiden vienti kääntyi nousuun usean vaimumman vuoden jälkeen. Metallien ja metallituotteiden viennin volyyymi kasvoi huomattavasti, vaikka niiden hinnat olivat laskussa.

Kulkuneuvojen vienti kasvoi viime vuonna eniten, lähes 50 prosenttia. Nopeaa kasvua selittää etenkin henkilöautojen vienti Saksaan. Myös laivojen ja tavarankuljetusautojen vienti kasvoi. Terveysteknologian kojeet ja laitteet olivat yksi viennin kasvualoista, vaikka osuus kokonaisviennistä onkin vielä sängen vaatimaton. Myös erilaisten tieteellisten mittaus- ja analyysikojeiden, elektroniikan komponenttien ja puolustusteollisuuden tuotteiden vienti on lisääntynyt vauhdilla viime vuosina.

Mistä lisää kysyntää?

Öljyn hinnanlaskusta ja Venäjän viennin vaikeuksista huolimatta viennin kokonaisarvo oli viime vuonna lähes en-

nallaan. Yhteenlaskettu vienti kaikkiin muihin maihin paitsi Venäjälle olikin lisääntymään päin viime vuonna. Vienti EU-maihin kasvoi kolme prosenttia ja euroalueelle kuusi prosenttia. EU-maiden osuus Suomen viennistä kasvoi viime vuonna 57 prosenttiin. Myös tuonnissa EU-maiden osuus oli nousussa ollessaan viime vuonna 58 prosenttia.

Vienti ei lähde elpymään ilman ulkomaista kysyntää. Venäjältä sitä tuskin tulee ainakaan lähiaikoina. Länsi-Euroopan suurin kansantalous Saksa on EU-maista melkein yksin kasvattanut tuontiaan Suomesta merkittävästi. Saksa nousi suurimmaksi kauppakumppaniksi sekä viennissä että tuonnissa. Espanjan ja Latvian tuonti Suomesta koheni myös viime vuonna.

Ruotsi oli pitkään tärkein vientimaamme, mutta viime vuonna Ruotsin talouskasvu oli vaatimatonta. Ruotsi ostikin neljä prosenttia vähemmän tavaroita Suomesta viime vuonna. Tosin maan talous kasvoi odotettua enemmän viime vuoden viimeisellä neljänneksellä, joka voi heijastua Ruotsin tuontikysyntään myöhemmin tänä vuonna. Toistaiseksi Ruotsin kokonaistuonti on kuitenkin laskenut ja alkuvuonna tuonti Suomesta väheni erityisen paljon.

Kasvavia ulkomaanmarkkinoita on jonkin verran löytynyt kaukaisemmista maista. Suomen vienti Yhdysvaltoihin lisääntyi seitsemän prosenttia viime vuonna. Se on suunnilleen yhtä paljon kuin EU-maiden viennin kasvu maa-

han. Yhdysvallat oli viime vuonna neljänneksi suurin vientimaamme. Joulukuussa 2014 viennin arvo Yhdysvaltoihin oli korkeampi kuin Venäjän viennin arvo. Maa nousi Suomen kolmanneksi suurimmaksi vientimaaksi.

Kiina on yksi maailman suurimmista talouksista ja olisi potentiaalinen kasvumarkkina-alue tulevaisuudessa, vaikka maan talousnäkökulmat ovatkin heikentyneet. Suomen vienti Kiinaan väheni kuitenkin viime vuonna seitsemän prosenttia. EU28-maiden yhteenlaskettu vienti Kiinaan nousi samaan aikaan yli kymmeneksen.

Vientimme Japaniin ja Etelä-Koreaan kasvoi hienoisesti viime vuonna. Suurimmat kasvuluvut löytyvät ei niin perinteisistä vientimaista, kuten esimerkiksi Meksikosta, Indonesiasta, Singaporesta ja Thaimaasta. Kaikilla näillä mailla on kuitenkin vain alle prosentin osuus Suomen kokonaisviennistä.

Ulkomaankaupan tilastoista ei tätä kirjoittaessa löydy merkkejä viennin kääntymisestä kasvuun. Viime vuoden viimeisellä neljänneksellä viennin orastava kasvu laantui. Tammikuussa 2015 viennin arvo laski 11 prosenttia, helmikuussa 4 prosenttia. Tammikuussa vienti Venäjälle vajosi jopa 43 prosenttia ja lasku jatkui helmikuussa. Venäjän viennin osuus tammikuussa oli enää alle viisi prosenttia kokonaisviennistämme.

Öljyn hinnanlasku ja Venäjän viennin raju pudotus selittävät vain noin puolet koko viennin vähenemisestä tammikuussa. Toisaalta molemmat tekijät vaikuttivat jo viime vuonna vientiin, ja vienti pysyi kuitenkin lähes edellisvuoden tasolla. Viennin kasvu erityisesti Saksaan ja myös Yhdysvaltoihin on korvannut Venäjän viennin kutistumista.

Yli 70 prosenttia Suomen viennistä menee Euroopan maihin. Nopean talouskasvun maat sijaitsevat tällä hetkellä enimmäkseen Euroopan ulkopuolella. Nähtäväksi jää, vetävätkö länsimaat vientimme kasvuun tai löytyykö kasvu uusilta markkina-alueilta. ■

Christina Telasuo työskentelee Tullin tilastointiyksikössä.

MARTTI PYKÄRI

Kotitalouksien varallisuus kasvaa Suomessa – muista Pohjoismaista ollaan silti jäljessä

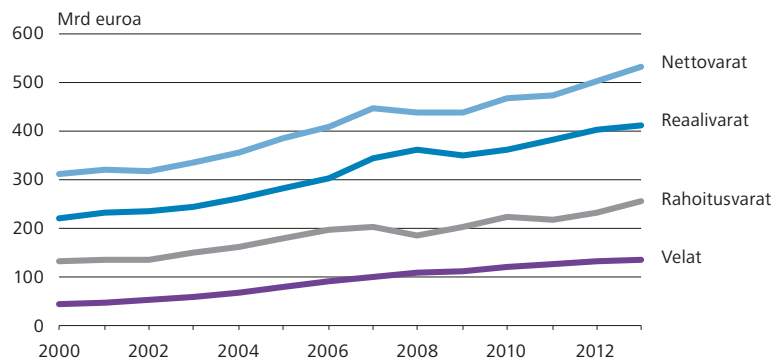
Suomalaisten kotitalouksien varallisuus perustuu omistusasumiseen. Asuntojen hintojen nousu on taannut nettovarallisuuden kasvun viime vuosikymmeninä. Varallisuuden rakenteessa ei ole tapahtunut merkittävää muutosta 2000-luvun aikana, vaikka sijoitusrahasto-osuuksien suosio onkin kasvanut.

Keskustelussa suomalaisten kotitalouksien varallisuudesta on tapana mainita ainakin kaksi asiaa: Suomalaiset kotitaloudet eivät ole erityisen varakkaita, ja varallisuus on sitoutunut asuntoihin. Lisäksi suomalaisten on sanottu makuuttavan rahojaan pankkitileillä. Muun muassa näitä väitteitä voidaan arvioida Tilastokeskuksen varallisuustilinpidoon pohjalta.

Kokonaisvarallisuuden kansainvälinen vertailu on vaikeaa. Reaalivaroista on olemassa hyvin vähän tilastotietoa ja laskentamenetelmät eroavat maittain. Kansantalouden tilinpidon kehikko on laajentumassa näiltä osin, mutta EU-maita velvoittava kattava raportointi tekee vasta tuluaan. Tilastotuotannon käynnistäminen ja menetelmien harmonisointi ottaa aikansa.

Rahoitusvarojen osalta kattavat tiedot ovat sen sijaan olleet saatavilla jo pitkään. Tältä osin kansainvälistä vertailua ei tosin voi tehdä kovin suoraviivaisesti erityisesti maiden eläkejärjestelmien eroista johtuen. Suomessa työeläkelaitosten varoja ei lasketa kotitalouksien vaan julkisyhteisöjen säästökseen. Monessa muussa maassa eläkesäästöt on luettu kotitalouksien rahoitusvaroihin.

KUVIO 1. Kotitalouksien varallisuus 2000–2013



Lähde: Tilastokeskus, varallisuustilinpito, rahoitustilinpito

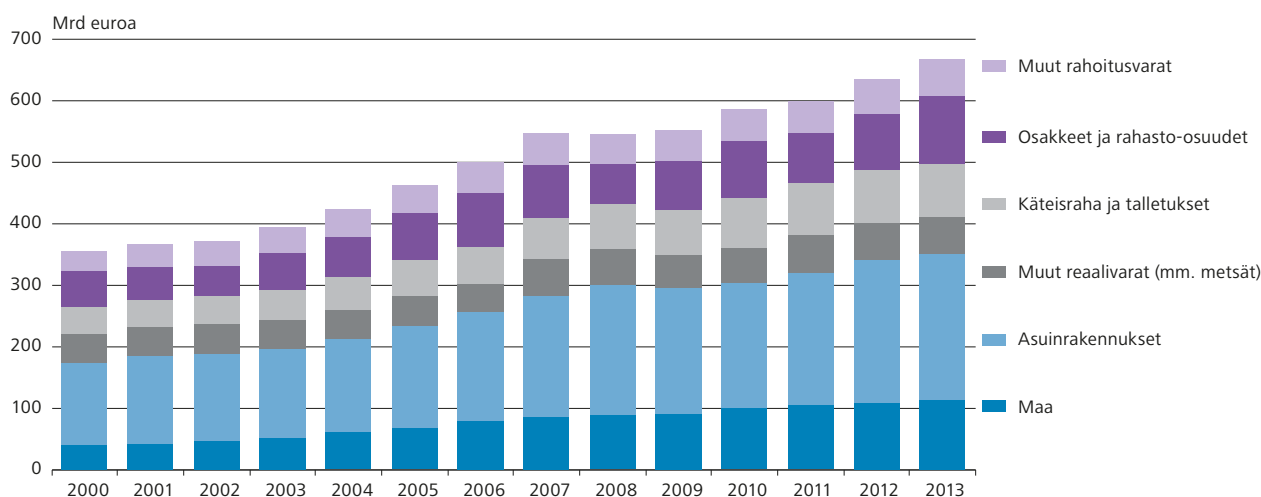
Tästä järjestelmäerosta johtuen suomalaisten kotitalouksien rahoitusvarallisuudesta ja sen rakenteesta saa helposti vääristyneen kuvan esimerkiksi EU:n tilastoviraston Eurostatin (Eurostat) tai OECD:n laatimista maavertailuista. Jos työeläkelaitosten nettoraahoitusvarat laskettaisiin suoraan kotitalouksien säästökseen, nousisi suomalaisten kotitalouksien rahoitusvarallisuus merkittävästi, mutta tämäkään ei tekisi suomalaisista varakkaita verrattuna muihin saman tulotason maihin.

Nettovarallisuus on kasvanut

Nykyisestä taloustaantumasta huolimatta suomalaisten kotitalouksien nettovarallisuus, eli varojen ja velkojen erotus, ovat kasvaneet ainakin vuoden 2013 loppuun saakka. Vain finanssikriisin alkuvuosina nettovarallisuus laski, kun sekä rahoitus- että asuntomarkkinat notkahtivat (kuvio 1).

Vuonna 2013 kotitalouksien nettovarallisuus olivat 532 miljardia euroa. Suurin yksittäinen varallisuuden erä olivat asunnot, 237 miljardia euroa (kuvio 2).

KUVIO 2. Kotitalouksien varallisuuden rakenne 2000–2013



Lähde: Tilastokeskus, varallisuustilinpito, rahoitustilinpito

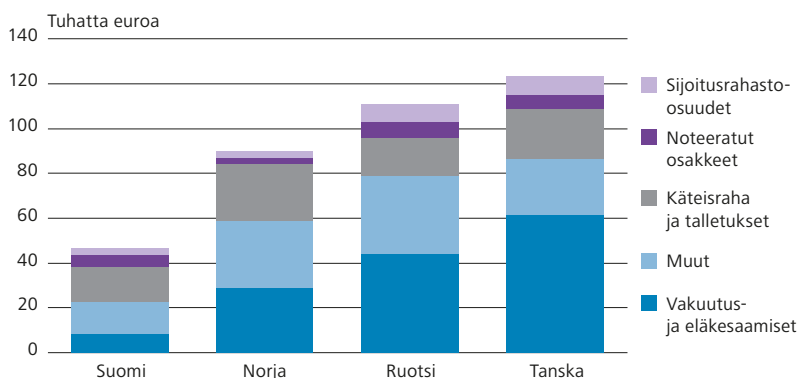
Kotitalouksien varallisuuden rakenne on säilynyt ennallaan koko 2000-luvun. Kokonaisvarallisuudesta vajaan 40 prosenttia on rahoitusvarallisuutta ja reilut 60 prosenttia reaaliavarallisuutta. Asuntojen osuus varallisuudesta on pysynyt runsaassa 35 prosentissa. Toiseksi suurimman reaaliavarallisuuden erän, maan, osuus oli 17 prosenttia vuonna 2013.

Rahoitusvarallisuuden rakenne on sen sijaan muuttunut hieman, koska kotitaloudet ovat alkaneet sijoittaa kertynyttä varallisuuttaan sijoitusrahasto-osuuksiin.

Talletusten ja käteisrahan osuus kotitalouksien rahoitusvaroista on pysynyt koko 2000-luvun reilussa 30 prosentissa. Vielä 1990-luvun lopulla yli 50 prosenttia rahoitusvarallisuudesta oli talletuksia. Kokonaisvarallisuudesta talletusten osuus oli 13 prosenttia vuoden 2013 lopussa. Määräaikaistalletusten houkuttelevuus on heikentynyt jatkuvasti. 2000-luvun aikana kotitaloudet ovat sijoittaneet nettomääräises-

**KOTITALOUDET OVAT
ALKANEET SIOITTA
SIOITUSRAHASTO-
OSUUKSIIN.**

KUVIO 3. Rahoitusvarat asukasta kohti eräissä Pohjoismaissa 2013



Muut rahoitusvarat sisältävät mm. noteeraamattomat osakkeet.

Lähteet: Tilastokeskus, Statistiska centralbyrån, Danmarks Statistik, Statistisk sentralbyrå

ti yli 11 miljardia euroa sijoitusrahasto-osuuksiin, ja niiden osuus kotitalouksien rahoitusvaroista on noussut hiljalleen vajaan neljästä prosentista seitsemään.

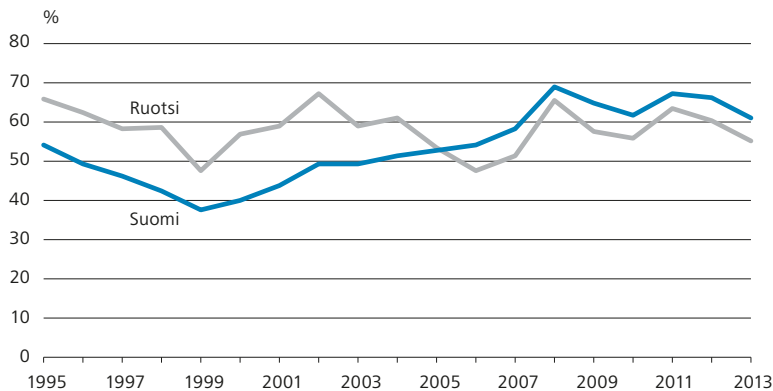
Suomessa pienet rahoitusvarat verrattuna muihin Pohjoismaihin

Suomessa talletusten suhteellinen osuus rahoitusvaroista on pohjoismaisittain korkea, vaikka jättäisi järjestelmäeroita johtuen eläkesäästöt kokonaan huomioon ottamatta. Talletusten korkea suhteellista osuutta selittää kuitenkin enemmän muun rahoitusvarallisuuden

vähyyks kuin talletusvarojen suuruus. Esimerkiksi vuoden 2013 lopussa suomalaisilla oli asukasta kohti vajaan 16 000 euroa talletusvaroja, ja 9 000 euroa noteerattuja osakkeita ja rahasto-osuuksia. Vastaavat luvut Ruotsista ovat 17 000 ja 15 000 euroa. (Kuvio 3.)

Suomalaisten kotitalouksien riskit painottuvat selvästi asuntomarkkinoille. Asuntokannan arvon osalta vertailua esimerkiksi Ruotsiin hankaloittavat laskentamenetelmien eroavaisuudet. Sen uskaltaa kuitenkin sanoa, että naapurimaassa asunnot muodostavat selvästi pienemmän osan kotitalouksien kokonaisvarallisuude-

KUVIO 4. Kotitalouksien lainavelkojen suhde rahoitusvaroihin¹ Suomessa ja Ruotsissa 1995–2013



¹ Järjestelmien eroavaisuuksista johtuen rahoitusvaroissa ei ole tässä otettu huomioon vakuutus- tai eläkesäästöjä (rahoitustilinpidon varaluokka F6).
Lähteet: Tilastokeskus, Statistiska centralbyrån

ta. Osa erosta selittyy yksinkertaisesti korkeammalla kokonaisvarallisuuden tasolla. Lisäksi omistusasuminen on Ruotsissa hieman vähemmän suosittua kuin Suomessa.

Suomen kotitalouksien velkaantumistaso EU-maiden keskitasoa

Kotitalouksien velat ja velkaantumistaso ovat läheisesti kytköksissä asuntomarkkinoihin. Pitkään jatkunut matala korkotaso on osaltaan mahdollistanut kotitalouksien velkaantumistasoen nousun, joka on tosin viime vuosina hidastunut merkittävästi. Kotitalouksien lainavelkakanta kasvoi vuonna 2013 vain reilut 3 prosenttia, kun vielä 2000-luvun puolivälissä vuotuiset kasvuprosentit olivat kaksinumeroisia.

Kotitalouksien velkaantumistaso, eli lainavelkojen määrä suhteessa käytettävissä olevaan tuloon, oli vuoden 2013 lopussa 119 prosenttia, mikä on melko lähellä EU-maiden keskiarvoa ja Pohjoismaiden matalin. Esimerkiksi Ruotsissa vastaava luku oli yli 150 ja Tanskassa peräti yli 260 prosenttia. Velkojen suhteuttaminen käytettävissä olevaan tuloon on perusteltua, koska velat maksetaan yleensä tuloilla. Asiaa voi kuitenkin tarkastella myös suhteuttamalla velat tulojen sijaan rahoitusvaroihin.

Vuodesta 2006 lähtien suomalaisilla kotitalouksilla on ollut rahoitusvaroihin nähden enemmän velkaa kuin ruotsalaisilla (kuvio 4).

Suomalaisten kotitalouksien velkaantumistasoen kehitys jää nähtäväk-

SUOMALAISTEN KOTITALOUKSIEN RISKIT PAINOTTUVAT ASUNTOMARKKINOILLE

si. Velkaantumisesta puhuttaessa on syytä muistaa, että kotitaloudet ovat keskenään hyvin erilaisessa asemassa. Varallisuuden jakautumisen, velkaantuneisuuden ja sosioekonomisten ryhmien tarkastelun mahdollistaa parhaiten Tilastokeskuksen huhtikuun alussa julkaistava varallisuustutkimus, joka perustuu mikrotason otosaineistoon (Tilastokeskus 2015).

Mikrotason otosaineiston ja kansantalouden tilinpidon kehittöön kuuluvan makroaineiston lähtökohtaisista eroavaisuuksista on kuitenkin hyvä olla tietoinen. Mikroaineiston tiedot koostetaan Suomessa eri lähteistä. Esimerkiksi pörssirosakkeet ja suorat rahastosijoitukset perustuvat rekisteritietoon, mutta talletukset haastattelutietoon. Tyypillisesti kotitalouksien rahoitusvarallisuus on mikrotason varallisuustutkimuksissa selvästi pienempi kuin rahoitustilinpidoissa. Myös käsite- ja arvotuserot vaikeuttavat varallisuustutkimuksen tietojen käyttöä makrotason tarkastelussa.

[Kommentoi verkossa tietotrendit.stat.fi](#)

Varallisuustaseiden tilastointia kehitetään edelleen

Tilastokeskus julkaisi sektorikohtaiset varallisuustaseet ensimmäisen kerran syyskuussa 2014. Reaalivarojen osalta varallisuustaseiden tilastoinnissa on edelleen puutteita ja kehitettävää. Kiinteän pääoman tiedot perustuvat pääomakan- tamallin mukaiseen nimellishintaiseen nettokantaan, joka arvotetaan vuoden keskihinnoin. Periaatteessa kaikki varat tulisi arvottaa taseajankohdan markkinahintaan, mikä ei kuitenkaan kiinteälle pääomalle ole kovin yksinkertaista.

Nettokanta koostuu investointien kumuloituneesta arvosta vähennettynä kumuloituneella kiinteän pääoman kulumisella. Maan arvo on laskettu maankäyttölajeittain alueellisesti kiinteistöjen kauppahintarekisterin tietojen pohjalta. Menetelmä on laadittu yhteistyössä Maanmittauslaitoksen kanssa. Tarkastelun ulkopuolella on edelleen tärkeitä varallisuuseriä kuten mineraali- ja malminvarat, joilla ei kuitenkaan kotitalouksien osalta ole juuri merkitystä. ■

Kirjoittaja on yliaktuaari Tilastokeskuksen talous- ja ympäristötilastot -yksikössä.

Lähteet:

Danmarks Statistik. Financial accounts. <http://www.dst.dk/en/Statistik/emner/offentlige-finanser/finansielle-offentlige-finanser.aspx>.

Eurostat. Household financial assets and liabilities. http://ec.europa.eu/eurostat/statistics-explained/index.php/Household_financial_assets_and_liabilities.

Statistisk centralbyrå. Financial accounts. <http://www.ssb.no/en/nasjonalregnskap-og-konjunkturer/statistikk/finsek/kvartal>.

Statistiska centralbyrån. Financial accounts. <http://www.scb.se/FM0103-EN/>.

Tilastokeskus. Varallisuus- ja rahoitustilinpito. <http://tilastokeskus.fi/til/rtp/index.html>.

Tilastokeskus 2015. Kotitalouksien varallisuus 2013. Tulot ja kulutus 2015. http://tilastokeskus.fi/til/vtutk/2013/vtutk_2013_2015-04-01_fi.pdf.

PASI PIELA

Iso data tilastotoimessa

Iso data haastaa monessakin mielessä perinteistä tilastotietoa. Ylivoimaisin se on nopeutensa ansiosta. Perinteisen tilastotiedon vahvuuksia ovat sen sijaan tiedon ajallinen vertailtavuus ja edustavuus. Yhteistyöllä on saavutettavissa molempia tiedontuotannon muotoja hyödyttäviä etuja.

Nykysin keskustellaan paljon tiedon vallankumouksesta ja massadatasta, jotka teknologinen kehitys ja tiedon digitalisoituminen ovat mahdollistaneet. Myös big data -aiheisten kirjoitusten määrä on lisääntynyt eksponentiaalisesti siinä missä Google-hakujenkin. Kyseessä on big data -hypetyt.

Myös julkinen tilastotoimi on merkittävästi reagoimassa ison datan haasteeseen sekä kansainvälisellä (YK), eurooppalaisella (Euroopan komissio) että kansallisella tasolla. Julkishallinnon tehtävänä on toimia ison datan mahdollistajana, hyödyntäjänä ja voimavarojen kohdentajana.

Käsittelen artikkelissani isoa dataa ja sen haasteita tilastotoimen kannalta. Keskityn ensin tilastoinnin laatunäkökohtiin ja esittelen sitten joitakin esimerkkejä massadatan tilastollisesta hyödyntämisestä.

Isoa dataa ei ole mielekästä käsitellä kokonaisuutena, sillä määritelmällisesti eheä se ei ole. Iso data on jatkuvaa datavirtaa jollakin elämän osa-

alueella vaikkapa sosiaalisen median päivityksinä tai luottokorttitransaktioina. Voitaisiinko transaktiodatasta saada tietojen julkaisemista nopeutavaa täydennystä talustilastoihin? Voitaisiinko kauppojen kanta-asiakaskorttiaineistoilla saada lisäpotkua kotitalouksien kulutuksen tilastointiin?

Toisaalta luonnollinen suuntaus etsiä uusia tietolähteitä nykyisten tilastojen rikastamiseksi ja mahdollisesti uudentyypisten tietojen tuottamiseksi on ollut olemassa jo kauan (ks. Jari Niemisen artikkeli tässä numerossa). Tilastotoimessa ollaan silti vielä esivaiheessa varsinaisen ison datan suhteen, mutta siellä tunnistetaan hyvin se, että yksityisellä puolella iso data voi olla jo arkipäivää tietojen tuottamisessa.

Muuttuuko tilaston käsite?

Tilastotoimi on julkista, ja sitä säätelevät tiukat laatu- ja luotettavuuskriteerit, jotka on useimmiten standardoitu kansainvälisesti ja erityisesti osana Euroopan tilasto- ➤



U. Östlund

järjestelmää. Tilastotiedot perustuvat vertailtavuuteen pitkin aikasarjoihin – ei väliaikaisuuteen.

Iso data haastaa monessakin suhteessa perinteisen virallisen tilastotiedon. Voiko meillä olla tulevaisuudessa kilpailevia yksityisen sektorin tuottamia tilastoihin rinnastettavia tietoja? Joka puolella generoidaan dataa ja myös osataan analysoida sitä. Lisäksi menestyvät tahot voivat rekrytoida parhaimmat osaajat; näistä tunnetuimpia esimerkkejä on luonnollisesti Google.

taa perinteisen tilastotiedon. Iso data elää tässä hetkessä jatkuvana tapahtumien virtana. Ison datan ongelma on siinä, että se pitää pilkkoa pieneksi, jotta sitä voi käsitellä. Tällaiset pilkkomistyökälyt on mahdollistanut viime vuosien teknologinen kehitys datan tehokkaamman varastoinnin ohella. Tunnetuin avoimen lähdekoodin työkalu on Hadoop. Hadoop on suurten datamassojen tallentamiseen ja tehokkaaseen rinnakkaiskäsitelyyn optimoitu järjestelmä (ks. Hannes Heikinheimon ja Antti Ukkosen artikkeli tässä numerossa).

Tilastokeskuksen hallinnollisia aineistoja voi visualisoida ja esittää ilman pilkkomistakin – olkoonkin, että se voi olla vaativaa. Näin tilastoaineistot eivät ole määritelmällisesti isoa dataa. Entisestäään kasvavien valtavien hallinnollisten aineistomassojen käsittelyssä ja yhdistelyssä modernit tietovarastoteknologiat tulevat kuitenkin ilman big dataakin ajankohtaisiksi (Piela 2013a; 2013b).

Nopeuden lisäksi toinen keskeinen syy ison datan ja uusien tietolähteiden etsimiselle ovat niiden mahdollistamat aiempaa tehokkaammat tiedonkeruumuodot. Voitaisiko esimerkiksi hintatiedonkeruuta korvata internetistä löytyvillä hinnoilla? Tilastokeskus selvittelee tänä vuonna vastausta kysymykseen erillisrahoituksen turvin.

Internet-hinnat ja transaktiot ovat hyviä esimerkkejä keskeisestä ison datan haasteesta – edustavuudesta. Otoksen edustavuuden klassinen tilastotieteellinen määritelmä on: suhteellisten jakaumien otoksessa pitää olla samat kuin kohdeperusjoukon jakaumien (Bethlehem 2009).

Ison datan piirissä olevat eivät edusta välttämättä tilastoinnin tai tutkimuksen kohdeperusjoukkoa. Kaikilla ei ole esimerkiksi luottokortteja käytössään, eikä voida ajatella luottokorttien käyttäjien edustavan myös niitä, joilla sitä ei ole.

Sosiaalinen media on yksi kiinnostuksen kohde kansainvälisessä tilastotoimen maailmassa. Ajankohtainen kysymys tähän liittyen on, kertovatko Twitter- tai Facebook-päivitykset jotakin onnellisuuden ja hyvinvoinnin kokemuksesta? Jos kertovat, niin miten laajaan väestöryhmään tulokset voidaan yleistää?

Nopeasta isosta datasta etua ennustamiseen

Nopeus on massadatan keskeisiä etuja, ja siksi se voi parantaa ennusteiden laatua. Tästä hyvänä esimerkkinä on ETLAn julkaisema tutkija Joonas Tuhkurin selvitys (Tuhkuri 2014), jossa internet-hakuaineistolla hyödynnettiin nykyhetken ja lähitulevaisuuden työttömyyden ennustamista. Työttömyystiedot tulivat Tilastokeskuksesta ja hakutiedot Google Trends -palvelusta.

Euroopan komissio rahoittaa tutkimusta, jonka avulla on tarkoitus edistää ison datan hyödyntämistä tilastotietojen tuotannossa.

Lukijalle ovat jo nyt varmasti tuttuja erilaiset internetissä julkaistut taulukot, joilla ei ole mitään tekemistä virallisen tilaston kanssa. Jos etsii asuntoa tietyltä alueelta, saa helposti kyseisen alueen tuoreen hintakehityksen, ja ehkäpä jopa tietystä kiinteistöstä. Tällaisia tilastotoimi ei tarjoa.

Keskustelua herättää ajoittain massadatan eräs erityinen ominaisuus tilastotoimen näkökulmasta: sitä hallinnoi tyypillisesti yksityinen sektori. Entä jos datalle tuleekin hintalappu? Miten käy tiedon tuotannon jatkuvuuden?

Yritykset luovuttavat yleisesti tilastoviranomaisille tietoja toiminnastaan, mutta eivät hallinnoimistaan tietovarannoista. Muutoinkin avoimen datan yhteydessä kuulee erittäin harvoin *open business data*sta. Avoin tieto kuuluu julkisen sektorin tuotettavaksi.

Entä sitten big datan laatu? On hyvä huomata, että tilastotieteellisesti kaikkein tärkein laatu-dimensio, tarkkuus ja luotettavuus, on vain yksi virallisen tilaston laadun ulottuvuuksista. Muut ovat relevanssi, puolueettomuus ja läpinäkyvyys, laadun hallinta, tietosuojat, tehokkuus, yhtenäisyys ja vertailukelpoisuus, saatavuus ja selkeys sekä oikea-aikaisuus ja ajantasaisuus (Tilastokeskus 2010).

Laatukriteereistä tarkkuus ja ajantasaisuus ovat usein ristiriidassa keskenään. Tilastoinnissa ristiriitaa voidaan ratkaista tuottamalla ennakkotietoja, jotka tarkentuvat aineiston karttuessa. Vielä nopeampia tilastotietoja toivotaan niin kutsuttuina *flash*-estimaatteina, pikaennakkoina, ennen virallisia ennakkotietoja.

Nopeus on erityisesti yksi ison datan keskeisistä perusominaisuuksista ja siinä mielessä se haas-

TAULUKKO. Ison datan tietolähteitä ja mahdollisia tilastointikohteita

NO.	AINEISTOLÄHDE	AINEISTON TYYPI	TILASTOLLINEN SOVELTAMISALUE
1	TELELIKENNE	Mobiilidata	Matkailutilastot Väestötilastot
2	INTERNET	Web-haut	Työvoimatilastot Muuttoliike
		Verkkokaupat	Hintatilastot
		Yritysten verkkosivut	Tietoyhteiskuntatilastot Yritysrekisteri
		Avoimet työpaikat verkossa	Työllisyystilastot
		Kiinteistöjen myynti-ilmoitukset verkossa	Hintatilastot (asuntojen hinnat)
		Sosiaalinen media	Kuluttajien luottamus Komission Beyond GDP -ohjelmaan liittyvä hyvinvoinnin ja onnellisuuden mittaaminen Tietoyhteiskuntatilastot
3	SENSORIT	Liikenteen automaattiset mittausasemat	Liikennetilastot Kuljetustilastot, tavaraliikennetilastot
		Etäluettavat sähkömittarit	Energiatilastot
		Satelliittikuvat	Maankäyttötalastot Maataloustalastot Ympäristötalastot
		Laivojen automaattinen tunnistusjärjestelmä (AIS)	Liikennetilastot Päästötalastot
4	PROSESSIN GENEROIMA DATA	Lentoreitit	Liikennetilastot Päästötalastot
		Kauppojen hintaskanneri- ja myyntidata	Hintatilastot Kulutustutkimus, kotitalouksien kulutus
		Taloudellinen transaktiodata	Kulutustutkimus
5	JOUKKOUTTAMINEN (CROWDSOURCING)	Osallistava paikkatieto, VGI (OpenStreetMap, Wikimapia, Geowiki)	Maankäyttö
		Yhteisölliset valokuvakokoelmat, CPC (Flickr, Instagram, Panoramio)	–

Lähde: Eurostat 2014 (ESS Big Data Action Plan and Roadmap 1.0)

Työttömyyteen liittyvien Google-hakujen yleisyyttä kuvaavalla muuttujalla oli tutkijan mukaan onnistuttu parantamaan aikasarjamallin tarkkuutta siten, että se ennustaa nykyhetken ja lähitulevaisuuden työttömyyttä paremmin kuin sama malli ilman Google-muuttujaa. Tällaisen internet-hakuja hyödyntävän mallin etu korostuu erityisesti käännekohtien ennustamisessa, mikä on luonnollisesti ennustuslaitosten suurimpia haasteita.

Myös Eurostat on julkaissut tutkimuksia Google Trends -palvelun käytöstä virallisen tilaston tutkimus- ja menetelmäportaalissaan muiden juuri nyt relevanttien iso data -tutkimusten ohella (Eurostat).

Kansainvälistä kehittämistä

Euroopan komissio rahoittaa tutkimusta, jonka avulla on tarkoitus edistää ison datan hyödyntämistä tilastotietojen tuotannossa. Euroopan unionin tilastovirastolla Eurostatilla on oma työryhmänsä (ESS Task-force on Big Data and Official Statistics), jossa Suomen Tilastokeskuksella on edustajansa. Työryhmässä pohditaan ja kehitetään tulevia iso data -pilotteja eurooppalaisen tilastotoimen tarpeisiin. Taulukkoon on koottu työryhmän johtopäätöksiä iso data -aineistotyypeistä ja mahdollisista tilastoista, joissa aineistoja voitaisiin hyödyntää. ➤

Mobiilipaikannusdataa väestö- ja matkailutilastoihin

Tilastojen rikastamiseksi mobiilipaikannusdata auttaisi matkustustilastojen laadinnassa. Esimerkiksi Hollannin tilastovirasto käyttää yö- ja päiväväestön paikantamiseen kolmannen osapuolen aggregoimia tietoja, jotka perustuvat teleoperaattori Vodafone:n mobiilipäätteiden paikannusdataan. Televiestintälaki estää tilastoviranomaiselta mikrotason tietojen saamisen välillisenkin tunnistamisen riskin nojalla. Aineistolla on hintansa ja sopimus yhtiön ja tilastoviraston välille on tehty vuodeksi kerrallaan.

Hollannin tilastoviraston suunnitelmissa on saada huomattavasti nykyistä enemmänkin tietoa ihmisten liikkuvuudesta. Suomessa on työssäkäynnin tarkastelemiseen hyvät hallinnolliset aineistot asuin- ja työpaikkakoordinaattitietoineen ilman mobiilidataakin.

Visa Europen kulutusbarometri: yhteistyöllä flash-estimaatteihin?

Luottokorttiyhtiö Visa Europe tuottaa indeksitilastoa ”EU Consumer Spending Barometer”, misä hyödynnetään tosiaikaisia luottokorttitransaktioita (Visa Europe 2015). Vuosineljänneksittäin tuotetaan aikasarjat kaksi kuukautta referenssiperiodin jälkeen. Samoin tuotetaan Britannian ”UK Expenditure Index” ja Ruotsin ”Sweden Expenditure Index”.

Visa pystyy tuottamaan informaatiota jaoteltuna useiden taustamuuttujien mukaan (ikä, sukupuoli, tulot, aviosääty jne.), jotka talletetaan haettaessa korttia. Lisäksi kokonaiskulutus voidaan todennäköisesti jakaa varsin tarkasti eri kulustavararyhmiin ostopaikan perusteella (vaatekauppa, lääkäri, apteekki, ravintola jne.).

Euroopassa on enemmän kuin 419 miljoonaa Visa-korttia, ja niiden käytöstä johdetulla barometrilla on Eurostatin tutkimuksen (ks. Eurostat 2014) mukaan merkittävä yhteys viralliseen tilastotietoon kotitalouksien kulutuksesta. Euroopassa virallisia kulutustutkimuksia varten kerätään tietoja suoraan kotitalouksilta noin joka viides vuosi ja aika tiedonkeruusta julkistukseen on yhdestä neljään vuotta. Kulutustietojen harmonisointi Euroopan tasolla on vaativaa, koska tiedonkeruulle ei ole olemassa lainsäädännöllistä perustaa. Suomessa Tilastokeskuksen kulutustutkimus toteutetaan epäsäännöllisin väliajoin.

Eurostat pitää mahdollisena pääsyä aggregoituihin Visa-tietoihin, kun kaikki mahdollinen yksikkökohtainen tieto on poistettu. Visaa voisi siis käyttää flash-estimaattien tuottamiseen rakenteellisesta kuluttamisesta ja sen määrästä. Eurostat kuitenkin korostaa, ettei Visan barometri voi



Tilastotoimen on tuettava ison datan tuotantoa entistä paremmin linkitettävillä avoimilla tilastotiedoilla.

korvata virallista tilastoa ja sen sisältämiä useita indikaattoreita. Sitä voidaan käyttää ylimääräisenä rikastuttajana viralliselle tilastolle. (Suomen osalta ks. Tarja Hatakan artikkeli tässä numerossa.)

Sensoritiedolla matka-aikamallinnukseen

Sensordata on tyyppiesimerkki isosta datasta. Esimerkki pienehköstä sensordatavirrasta julkishallinnossa on Liikenneviraston yli 400 tieliikenteen automaattisen mittausaseman data.

Tilastokeskuksessa on laskettu erilaisia saavutettavuuteen liittyviä tietoja kuten työmatkat kaikille suomalaisille (Piela 2014). Saavutettavuuteen sisältyy yleensä myös matkustusaika tai pelkkä ajoaika. Tyypillisesti tähän käytetään tien toiminnallisuuteen liittyviä malleja. Päästäkseen parempiin matka-aikamalleihin erityisesti ruuhka-aikoina Tilastokeskus on soveltanut Liikenneviraston automaattisten mittausasemien tietoja keskinopeuksista eri tieosuuksilla sopivilla aikaikkunoilla. Näiden tuottamaa tietoa pidetään suhteellisen edustavana Suomen pääteille ja kantateille.

Tiedon edustavuus on silti suuri haaste tässäkin hankkeessa, mutta mittausasemien data on tarpeen yleistettynäkin, koska parempaakaan tietoa ei ole helposti saatavilla nopeutusrajotusten ja tien eri ominaisuusluokkien lisäksi. Tässä siis rikastetaan aiempia matka-aikamalleja uudella tiedolla liikennesensoreista osana saavutettavuustutkimusta.

Yksityisen sektorin hallinnoiman ison datan rikkaus ja moninaisuus nousee esiin matka-aikamallinnuksissakin: jokainen voi käydä laskemassa työmatkansa haluttuna kellonaikana esimerkiksi Googlen karttaa käyttäen ottaen huomioon tiehäiriöt ja ruuhkaisuudet yms. Näin tarkkaan estimointiin ei Tilastokeskuksen matkalaskelmissa päästä.

Lisäksi navigointi- ja karttajärjestelmien kehittäjät kuten Tomtom myyvät tie-elementtikohaitaisia (Tomtomilla TeleAtlas) keskinopeustietoja haluttuna kellonaikana perustuen todellisiin tietoihin. Hinnoittelu voi olla kilometripohjainenkin. Tomtom tarjoaa myös evaluointimahdollisuuksia näppärään käyttöliittymäänsä.

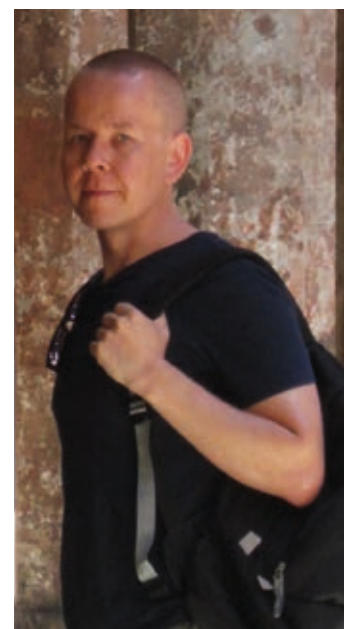
Kaupalliset aineistot eivät ole maksuttomia käyttäjilleen, mutta Liikennevirasto on avannut sekä digitalisoidun liikenneverkon ”Digiroad” että liikenneasemiensa sensoridatan Digitraffic.fi-palvelun kautta laajalle käyttäjäryhmälle. (Esimerkkinä työmatkojen pituuden laskemisesta Digiroad-aineiston pohjalta ks. Huhta & Pyykkönen 2013; ks. myös Topias Pyykkösen artikkeli tässä numerossa.)

Yhteistyössä entistä parempaan tiedontuotantoon

Virallisen tilastotoimen rooli ison datan maailmassa on moninainen. Yhtäältä isoa dataa on ehdottomasti tuettava entistä paremmin linkitettävillä, avoimilla tilastotiedoilla. Toisaalta tilastotoimi tarvitsee isoa dataa myös rikastaakseen ja tehostaakseen omaa tiedon tuotantoaan.

Keskeisiä kysymyksiä ja selvitettäviä asioita ison datan ja tilastotuotannon yhteistyölle ovat mm. tiedon nopeuteen ja edustavuuteen liittyvät haasteet, yksityisen sektorin hallinnoimat parhaimmat massadatarannot, aineistojen hinta ja kuinka julkisen sektorin tarjoamat avoimet aineistot saatai-

Kirjoittaja on kehittäispäällikkö Tilastokeskuksen väestö- ja elin-olotilastot -yksikössä.



siin tukemaan myös iso data -kehitysympäristöä.

Se, mitä tilastotoimi ylipäätään voi tarjota moninaisille massadatan kanssa toimiville kansainvälisille yrityksille ikään kuin vastineeksi tiedosta, on tiedon käsitteiden, määritelmien ja luokitus-ten kansainvälinen harmonisointi ja pitkä kokemus laatustandardeista puhumattakaan mm. tilasto-, yhteiskunta-, ympäristö- ja taloustieteellisesti osaamisesta. ■

Lähteet:

Bethlehem, Jelke 2009. *Applied Survey Methods: A Statistical Perspective*. Hoboken, New Jersey: John Wiley & Sons.

Eurostat 2014. *Big data. Papers on Big Data in the context of Official Statistics*. <http://www.cros-portal.eu/content/papers-big-data-context-official-statistics>.

Huhta, Jaana & Pyykkönen, Topias 2013. Linnuntieltä oikeille teille. Hyvinvointikatsaus 4/2013. http://www.tilastokeskus.fi/artikkelit/2013/art_2013-12-09_003.html.

Piela, Pasi 2013a. *Avataan Big Data! Tieto&trendit-blogi* 15.5.2013. <http://tietotrenditblogi.stat.fi/avataan-big-data/>.

Piela, Pasi 2013b. *Teollisuuden big data tulee olemaan sensoridataa. Tieto&trendit-blogi* 31.3.2013. <http://tietotrenditblogi.stat.fi/teollisuuden-big-data-tulee-olemaan-sensoridataa/>.

Piela, Pasi 2014. *Commuting time for every employed: combining traffic sensors and many other data sources for population statistics. European Forum for Geography and Statistics (EFGS) Krakow Conference 2014. Puolan tilastovirasto*. <http://geo.stat.gov.pl/efgs/programme>.

Tilastokeskus 2010. *Suomen virallisen tilaston laatukriteerit*. <http://www.tilastokeskus.fi/meta/svt/svt-laatukriteerit.html>.

Tuhkuri, Joonas 2014. *Big Data: Google-haut ennustavat työttömyyttä Suomessa. ETLA*. <http://www.etla.fi/julkaisut/33195/>.

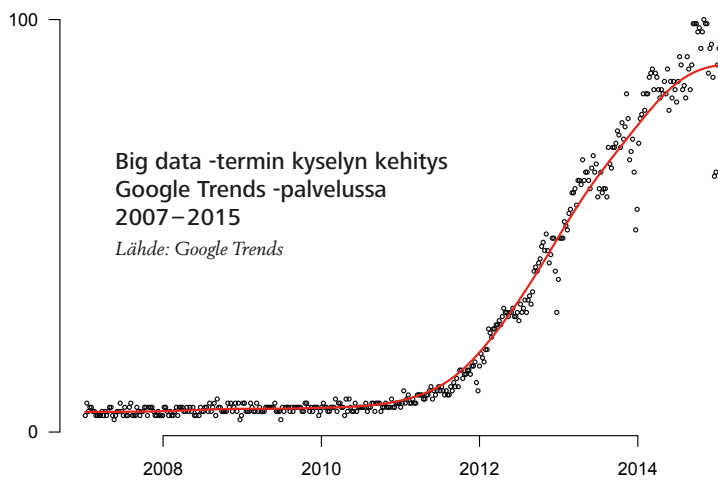
Visa Europe 2014. *Visa Europe Expenditure Index. Visa Europe*. <http://www.visaeurope.com/newsroom/all-reports>.

[Kommentoi verkossa tietotrendit.stat.fi](http://www.tietotrendit.stat.fi)

HANNES HEIKINHEIMO JA ANTTI UKKONEN

Ison datan alkulähteillä

Iso data -käsitteellä voidaan viitata isoihin tietoaisteihin, aineiston käsittelyyn soveltuviin työkaluihin tai itse ilmiöön, jossa datan keräämisestä ja tarkastelusta on tullut verraten halpaa ja nopeaa. Useimmat iso data -teknologiat on kehitetty ratkaisemaan globaalien internet-yritysten melko rajattuja tietojenkäsittelyongelmia. Teknologioiden soveltuvuus muuhun käyttöön on aina punnittava erikseen.



Tässä artikkelissa avaamme iso data -ilmiötä sen historiallisessa viitekehyksessä. Iso data voi oikein ymmärrettynä ja käytettynä olla tutkimukselle ja liiketoiminnalle erittäin hyödyllinen voimavara. Ei kuitenkaan pidä kuvitella sen ratkaisevan kaikkia organisaation ongelmia tai johtavan automaattisesti uusiin tieteellisiin läpimurtoihin.

Dataa eri muodoissa on ollut olemassa iät ja ajat. Myös “tiedon louhinnasta” on puhuttu jo 1990-luvulta asti. Iso data on käsitteenä kuitenkin kohtuullisen uusi.

Kuviossa on esitetty normalisoitu indeksi, joka kuvaa hakusanan *big data* yleisyyttä hakukone-yhtiö Googlen hakukoneessa vuodesta 2007 vuoden 2015 tammikuuhun (Google Trends, <http://trends.google.com>). Kiinnostus asiaa kohtaan on kasvanut huikeasti reilussa parissa vuodessa.

Kuvio on myös havainnollistava esimerkki siitä, mitä iso data voi käytännössä olla, ja minkälaisia asioita sillä voidaan tehdä.

Tässä tapauksessa Googlen tallentamista hakusanoista käy ilmi, kuinka maailmanlaajuinen kiinnostus jotakin ilmiötä kohtaan on kehittynyt. Samalla työkalulla on myös pyritty mittaamaan kausi-influenssan leviämistä (*Google Flu Trends*, Ginsberg ym. 2009; ks. myös Lazer ym. 2014).

Suomessa Elinkeinoelämän tutkimuslaitos (Etila) julkaisi hiljattain hakukonedataan perustuvan työttömyysindikaattorin (www.etla.fi/etlanow, ks. myös Tuhkuri 2014).

Hakukoneiden lisäksi uudenlaista dataa tuottavat erilaiset sosiaalisen median ja vähittäiskaupan palvelut sekä teollisuuden järjestelmät. Googlen rooli tässä kehityksessä on kuitenkin ollut erittäin merkittävä.

Siinä missä ohjelmistoyritys Microsoft toi henkilökohtaiset tietokoneet ensin toimistoihin ja myöhemmin koteihin, monet keskeisistä ison datan teknologioista on alun perin kehitetty Googlella.

Muita tärkeitä toimijoita menetelmien ja työkalujen kehittämisessä ovat olleet internet-yritykset Yahoo, Facebook ja Twitter. Erityisesti pilvipalveluiden kehittäjänä verkkokauppa Amazon on ollut keskeisessä roolissa.

Perinteisille tiedonhallintaan erikoistuneille yrityksille kuten IBM ja Oracle on usein jäänyt perässähihtäjän rooli.

Käytännössä kaikki tärkeät ison datan teknologiat perustuvat Linux-käyttöjärjestelmän tapaan avoimeen lähdekoodiin ja maailmanlaajuiseen kehittäjäyhteisöön.

Iso data ilmiönä

Iso data voi tarkoittaa jotakin tietynlaista dataa tai sen käsittelyyn kehitettyä teknologiaa. Iso data -ilmiöön liittyy eri tahojen kasvava kiinnostus ja valmius kerätä ja hyödyntää dataa omassa toiminnassaan ja jakaa keräämäänsä dataa muiden tahojen käyttöön.

Iso data -ilmiön voi myös kuvata ajattelu- tai toimintatavaksi, joka on seurausta siitä, että datan kerääminen ja käyttäminen on muuttunut halvaksi. Tähän kehitykseen ovat myötävaikuttaneet etupäässä digitalisaatio, tallennuskapasiteetin hinnan romahtaminen sekä pilvipalveluiden yleistäminen.

Digitalisaation seurauksena yhä suurempi osa yhteiskunnan toiminnasta on tallennettavissa tietokoneilla helposti käsiteltävään muotoon. Dataa tuottavat niin ihmiset kuin koneetkin. Sitä syntyy yrityksissä, julkisella sektorilla sekä tieteessä ja tutkimuksessa.

Digitaalisessa muodossa oleva data voidaan siirtää tehokkaasti paikasta toiseen ja siitä voidaan tehdä loputtomasti kopioita. Datan käyttö on siis yleistynyt, koska sitä on saatavilla jatkuvasti enemmän eri yhteyksissä. Koska datan käyttöä on ollut monille tahoille huomattavaa hyötyä, sitä vastaavasti myös kerätään entistä enemmän.

Tallennuskapasiteetin hinnan lasku merkitsee, että yhä useammalla toimijalla ei ole taloudellisia esteitä kerätä ja tallentaa kaikkea tuottamaansa dataa. Suhteutettuna organisaatioiden muihin kustannuksiin datan varastointi onkin varsin halpaa.

Pilvipalvelut ovat mahdollistaneet kaikille toimijoille kustannustehokkaan tavan kerätä ja hallita tietoa. Pieni yhden työntekijän aloitteleva yritys voi käyttää samaa teknologiaa kuin kansainvälinen suuryritys ja maksaa siitä vain käyttönsä mukaan.

Yleisimmät ison datan työkalut ja teknologiat

Kenties tunnetuin ja yleisesti käytetyin iso data -työkalu tällä hetkellä on **Hadoop** (<http://hadoop.apache.org>). Se on alun perin pääosin Yahoolla kehitetty avoimen lähdekoodin toteutus Googlen tiedonhallintainfrastruktuurista. Hadoop tuo mukanaan samat edut ja haitat kuin Googlen alkuperäiset järjestelmät. Erittäin suurten tietomassojen käsittely on sen avulla yleensä verraten edullista ja tehokasta, ja infrastruktuuri on helposti laajennettavissa, kun liiketoiminta kasvaa. Mutta sen hyödyt tulevat esiin vasta, kun dataa on todella paljon, mielellään vähintään kymmeniä teratavuja. Prosessoinnin on myös oltava luonteeltaan Hadoopille sopivaa.

Käytännössä Hadoop on kokoelma erilaisia avoimen lähdekoodin ohjelmistoja. Sen käyttöönotto vaatii varsin korkeaa tietoteknistä asiantuntemusta ja useasta eri tietokoneesta koostuvan laskentaympäristön. Hadoop ei siis ole esimerkiksi SPSS:n tai SAS:n kaltainen loppukäyttäjälle valmis työkalu. Pikemminkin Hadoop on alusta, jonka päälle voidaan rakentaa erilaisia työkaluja.

Toimivan Hadoop-järjestelmän käytön monimutkaisuus riippuu paljon sillä ratkottavista ongelmista. Yksinkertaisimmillaan se voi tarjota samanlaisia toiminnallisuksia kuin perinteiset tietovarastot, mutta käytännössä Hadoop voi suorittaa mielivaltaista datan käsittelyä. Tämä kuitenkin edellyttää räätälöityjen ohjelmistojen laadintaa.

Työkalut myös kehittyvät nopeasti. Alkuperäinen Hadoop soveltuu huonosti monimutkaisempien algoritmien toteutukseen. Erityisesti Hadoopille aiheuttavat ongelmia sellaiset tehtävät, joita ei ole mahdollista ratkaista lukemalla data kertaalleen läpi alusta loppuun.

Tällaisia ongelmia esiintyy usein esimerkiksi koneoppimisessa. Parin viime vuoden aikana kehitetyt uudet teknologiat sopivat paremmin myös tällaisten algoritmien toteuttamiseen. Tällä hetkellä näistä suosituin on **Spark** (<http://spark.apache.org>).

Googlella on ollut merkittävä rooli ison datan työkalujen kehittäjänä.

Suuriin investointeihin ei ole tarvetta. Järjestelmät mukautuvat joustavasti käytön edellyttämälle tasolle.

Data kiinnostaa erityisesti, jos sillä katsotaan olevan rahallista arvoa. Esimerkiksi valtaosa erilaisista verkossa toimivista palveluista on mainosrahoitteisia.

Mainosten esittämisen lisäksi on yhä tärkeämpää, että niihin reagoidaan. Äärimmäisessä tapauksessa palveluntarjoaja saa tuloja ainoastaan, jos mainos johtaa konkreettiseen ostotapahtumaan.

Palveluntarjoajat ovatkin ryhtyneet kohdentamaan mainoksia siten, että kullekin käyttäjälle näytettävät mainokset ovat todennäköisesti tätä kiinnostavia. Mitä enemmän dataa käyttäjistä on tarjolla, sitä paremmin kohdentaminen voidaan tehdä ja sitä todennäköisemmin mainokset johdavat tuloihin palveluntarjoajalle.

Kuinka tähän on tultu? Iso data -ilmiötä voi olla vaikea ymmärtää syvemmin ilman, että tuntee hieman sen historiallista taustaa. Myös iso osa aiheeseen liittyvistä väärinkäsityksistä saattaa juontaa juurensa innosta hyödyntää uusimpia työkaluja ymmärtämättä täysin niille alun perin suunniteltua käyttötarkoitusta.

Tässä artikkelissa kuvaamme minkälaisista lähtökohdista ensimmäiset ison datan teknologiat syntyivät, ja kuinka ne ovat kehittyneet kulueneen 15 vuoden aikana. Iso data onkin ilmiönä huomattavasti vanhempi kuin iso data -käsite.

Google on toiminut pioneerina

Modernit ison datan hallintajärjestelmät saivat alkunsa niistä teknologisista ratkaisuista, joita hakukoneyhtiö Google teki 2000-luvun alussa. Sekä Googlen ongelmat että Googlen niihin kehittä-

tautuneen tiedonhakumenetelmän ja perustivat hakukoneyhtiön (Brin & Page 1998). Tästä hakukoneesta tuli nopeasti erittäin suosittu. Internetin suosio kasvoi voimakkaasti, ja Googlella tehtyjen hakujen määrä alkoi lisääntyä.

Yrityksen oli vastattava niihin teknisiin haasteisiin, joita miljoonien verkkosivujen indeksointi ja miljoonien yhtäaikaisten verkkohakujen käsittely sille aiheuttivat. Käytännössä tämä tarkoitti laskennan suorittamisen rinnakkain hajautetussa ympäristössä usealla eri tietokoneella.

Aikaisemmin laskennan hajautusta vaativat sovellukset olivat löytyneet lähinnä tieteellisen laskennan puolelta, esimerkiksi sään ennustamisesta, ilmastotutkimuksesta, molekyylibiologiasta, aerodynamiikasta, kosmologiasta, hiukkasfysiikasta sekä kryptografiasta. Näitä ongelmia ratkoivat valtiolisten laitteiden kalliit supertietokoneet, jotka tyypillisesti perustuivat satoihin tai tuhansiin huippunopeilla yhteyksillä toisiinsa kytkettyihin prosessoreihin ja jaettuun muistiin. Tällaiset tietokoneet olivat monimutkaisia ja kalliita erikoislaitteita.

Supertietokoneiden käytön sijaan Googlella haluttiin toisenlainen ratkaisu. Yritys päätyi rakentamaan kokonaan oman laskentainfrastruktuurin, jonka keskeisiä osia olivat hajautetun laskennan kannalta epäluotettavat, pääasiassa kuluttajille suunnatut PC-tietokoneet ja hitaat verkkoyhteysteknologiat. Syitä tähän ratkaisuun oli monia.

Yksi oli hinta. Googlen laskentavaatimukset kehittyivät jatkuvasti vaikeasti ennustettavalla tavalla. Verkkosivujen määrä kasvoi eksponentiaalisesti, samoin Googlen käyttäjäkunta. Laskentakapasiteettia tulitaisiin tarvitsemaan jatkuvasti yhä enemmän.

Jos Google olisi sijoittanut kalliisiin supertietokoneisiin, lasku olisi ollut hirvittävä. Yksinkertainen PC- ja verkkoteknologia oli sen sijaan halpaa ja helposti saatavilla. Kun laskenta- tai tallennuskapasiteettia piti kasvattaa, ei tarvinnut kuin ostaa iso kasa uusia PC-tietokoneita ja liittää ne osaksi olemassa olevaa järjestelmää.

Toinen syy olivat Googlen tieteellisestä laskennasta poikkeavat tiedonkäsittelyn ja -hallinnan haasteet. Tieteellinen laskenta on tyypillisesti nimensä mukaisesti laskentaintensiivistä: tarvitaan paljon nopeita prosessoreita suorittamaan liukuoperaatioita.

Googlen ongelmat liittyvät numeronmurskauksen sijaan enimmäkseen erittäin suurten tietomasojen käsittelyyn. Esimerkiksi kaikkien www-sivujen indeksointi edellyttää ainoastaan, että jokaiselta www-sivulta luetaan kaikki ko. sivulla esiintyvät sanat. Sivuja on huikea määrä, mutta ne voidaan kaikki käsitellä rinnakkain, toisistaan riippumatta. Monimutkaiseen numeeriseen laskentaan verrattuna tämä on hyvin erilainen tehtävä.

Työkalut ja menetelmät ovat huomattavasti vanhempia kuin käsitteet.

mät työkalut olivat kuitenkin melko tarkkaan rajattuja. Näin ollen kaikki Googlen ratkaisut eivät sellaisenaan välttämättä vastaa useimpien organisaatioiden tietovarastointia ja analytiikkaa koskeviin haasteisiin. Mitä Googlella siis oikeastaan tehtiin?

Google syntyi, kun 1990-luvun lopulla kaksi Stanfordin yliopiston jatko-opiskelijaa, Sergey Brin ja Larry Page, kehittivät ylivoimaiseksi osoit-

Kolmas syy oli se, että Google on ohjelmistotalo. Hajautetun laskentainfrastruktuurin ylläpitoon ja toimintahäiriöihin liittyvät ongelmat se päätti ratkaista itse, softan avulla.

Vastoin yhtä keskeistä isoa dataa koskevaa harhakäsitystä pyrkimyksenä ei niinkään ollut mahdollisimman tehokas suoritusnopeus vaan laskennan luotettavuus yllättäviä virhetilanteita kohtaan sekä järjestelmän helppo laajennettavuus ("horizontal scaling").

Yrityksen jatkuvan kasvun vuoksi oli myös tärkeää, että käytettävät ohjelmointiympäristöt olivat mahdollisimman helposti omaksuttavissa.

Tyypillisesti rinnakkaisohjelmointi vaatii erityisosaamista, jota aloittelevan hakukoneyrityksen olisi ollut vaikea rekrytoida. Pyrkimys oli siis luoda ohjelmistoarkkitehtuuri, joka tarjoaisi helpon mutta tehokkaan tavan hallita satojen teravujen kokoisia aineistoja.

Näistä lähtökohdista Google kehitti kolme teknologiaa, jotka osoittautuivat keskeisiksi iso data -ajattelun kannalta. Nämä ovat useammalle tietokoneelle hajautettu tiedostojärjestelmä GFS (Google file system, Shvachko ym. 2010), yksinkertainen hajautetun laskennan ohjelmointiparadigma *MapReduce* (Dean & Ghemawat 2008), sekä *Bigtable* (Chang ym. 2008), tehokas tietorakenne erilaisten online-palveluiden taustajärjestelmiä varten.

Näiden teknologioiden toteutus on myös keskeinen osa avoimen lähdekoodin Hadoop-järjestelmää (ks. tietolaatikko), joka on tällä hetkellä tunnetuin työkalu ison datan käsittelyä varten.

On huomattava, että Googlen alkuperäiset teknologiat eivät olleet avoimia. Hadoop on syntynyt kansainvälisen kehittäjäyhteisön toimesta ja muilla internet-yrityksillä on ollut tässä tärkeä rooli. Käytännössä Google on kuitenkin ollut merkittävä suunnannäyttävä, ja siellä tehdyt innovaatiot heijastuvat selvästi koko alan toimintaan.

Koneoppiminen ja iso data

Koneoppiminen ja ennustava mallinnus ovat niinkään iso data -ilmiön myötä suuren suosion saavuttaneita tapoja käyttää dataa. Tämä ei ole niinkään parempien algoritmien ansiota, vaan keskeistä on ollut saatavilla olevan datan määrän eksponentiaalinen kasvu yhdistettynä laskentakapasiteettiin, jolla tätä dataa on mahdollista käsitellä. Erityisesti internet-palveluiden yleistymisen 2000-luvun ensimmäisen vuosikymmenen aikana on vaikuttanut tähän.

Viitteet uudesta koneoppimisen noususta saatiin jälleen Googelta. Jokainen 1990-luvun lopulla Altavista-hakukonetta käyttänyt luultavasti muistaa, kuinka taianomaisen tarkoilta Googlen

löytämät hakutulokset tuntuivat. Google todella tuntui ymmärtävän mitä olit etsimässä.

Käyttäjä saattoi jopa kirjoittaa hakusanansa väärin ja silti hakukone osasi ehdottaa oikeaa kirjoitusasua ("did you mean?").

Ensimmäiset iso data -pohjaiset koneoppimistulokset ilmaantuivatkin kielen käsittelyyn liittyvissä sovelluksissa. MapReduce-ohjelmat ovat ideaalisia erilaisten sanojen ja sanayhdistelmien esiintymistodennäköisyyksien laskentaan. Sekä edellä kuvattu "did you mean?" -toiminnallisuus että erityisesti kieliteknologian asiantuntijoiden piirissä

Ison datan suoraviivainen käyttö voi toimia tilanteissa, joissa tarkoitus ei niinkään ole ymmärtää käsillä olevaa ilmiötä, vaan tuottaa mahdollisimman tarkkoja ennusteita.

huomiota herättänyt Googlen automaattinen kielenkäännöspalvelu ovat hyviä esimerkkejä tästä.

Kumpikaan edellä mainittu palvelu ei niinkään perustu kehittyneempiin algoritmeihin vaan yksinkertaisesti siihen, että Googella on käytössään huomattavan suuri aineisto eri kielisiä dokumentteja. Sama oivallus pätee myös moniin muihin mallinnusongelmiin: yksinkertainen malli ja paljon dataa voi joissain tapauksissa toimia paremmin kuin monimutkainen malli ja vähän dataa.

Lisää kohua koneoppimisen ympärille syntyi, kun Geoffrey Hinton ja Ruslan Salakhutdinov esittelivät uudenaikaisilla keinotekoisilla neuroverkoilla saamia tuloksia (Salakhutdinov & Hinton 2009). Näistä menetelmistä alettiin yleisesti käyttämään nimitystä "deep learning", ja niiden tuoma ajattelu muistutti myös paljon yllä kuvattua datalähtöistä ratkaisumallia.

Sovellusalueen syvällisen teoreettisen ymmärtämisen sijaan "deep learning" -verkoille syötetään ainoastaan paljon raakadataa sekä laskentakapasiteettia tulosten aikaansaamiseksi.

Deep learning on johtanut merkittäviin edistysaskeliin ainakin joissakin koneoppimisen tärkeissä sovelluksissa. Vuonna 2010 Microsoftilla kehitetty deep learning -tekniikoihin perustuva akustinen malli vähensi englanninkielisen puheentunnistuksen virheitä 30 prosenttia aikaisempiin tuloksiin verrattuna (ks. esim. Hinton ym. 2012). Tämä on yksi dramaattisimmista pa-

rannuksista automaattisen kielenkäsittelyn alalla vuosikymmeniin.

Pari vuotta myöhemmin Google toi vastaavan deep learning -äänemalliin ja iso data -pohjaisiin kielimalleihin perustuvan puheentunnistusteknologian Android-käyttöjärjestelmään.

On myös huomattava, että deep learning -menetelmät tarvitsevat lähtökohtaisesti erittäin paljon dataa toimiakseen, sillä malleissa on usein miljoonia parametreja. Pienillä aineistoilla niiden tuomat hyödyt voivat olla vähäisemmät. Ne eivät myöskään ole selittäviä malleja eivätkä näin ollen auta ymmärtämään tarkasteltavaa ilmiötä paremmin.

Iso data haastaa ja hämmentää

Iso data on eräillä aloilla haastanut voimakkaasti tavanomaisia toimintatapoja. Datan käsittelyyn liittyvissä teknologioissa se on haastanut vallalla olevat käsitykset kerättävän datan tarkasta rakenteesta sekä skaalautuvan laskentakapasiteetin luonteesta. Koneoppimisen puolella se on haastanut perinteiset käsitykset datan esikäsittelystä sekä sovellusalaokohtaisen monimutkaisen teoreettisen mallintamisen tarpeellisuudesta.

Tulokset puhuvat puolestaan, mutta uudenlainen ajattelutapa on tuottanut myös paljon sekaannusta. Hyvä esimerkki liittyy Googlella työskentelevän tekoälymenetelmiin erikoistuneen tutkija Peter Norvigin vuonna 2008 pitämään esitelmään O'Reillyn ETech -konferenssissa.

Esityksessään Norvig pohdiskeli datan ja teorian suhdetta. Hän viittasi tilastotieteilijä George Boxin 30 vuotta vanhaan toteamukseen "Kaik-

ki mallit ovat virheellisiä, mutta muutamat mallit ovat hyödyllisempiä kuin toiset."

Norvig totesi esityksessään: "Jos kaikki mallit ovat vääriä, ehkä monimutkaisten mallien kehittämisen sijaan voisimme ottaa yksinkertaisempia malleja ja antaa datan tehdä hankala työ."

Tästä lausunnosta syntyi iso hämmennys, kun tunnettu amerikkalainen tietokonekulttuuria käsittelevä lehti *Wired* viittasi Norvigiin artikkelissa "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete" (Anderson 2008).

Monien tutkijoiden mielestä Norvig hyökkäsi tieteellistä menetelmää vastaan, ja he pitivät tätä aivan ymmärrettävästi vaarallisena väitteenä. Asiaa ei kuitenkaan pidä ottaa kirjaimellisesti. Myöhemmin Norvig julkaisi vastineen (Norvig 2008), jossa hän sanoi *Wired*-lehden viittausten olleen tahallisen harhaanjohtavaa.

Epäilemättä datan enemmän tai vähemmän suoraviivainen käyttö voi toimia monissa sellaisissa tilanteissa, joissa tarkoitus ei niinkään ole ymmärtää käsillä olevaa ilmiötä vaan ainoastaan tuottaa mahdollisimman tarkkoja ennusteita.

Esimerkiksi verkkokaupan tavoitteena voi olla arvata edes likimain oikein, minkälaisista tuotteista tietty käyttäjä voisi olla kiinnostunut. Sen ei välttämättä tarvitse selittää tähän kiinnostukseen liittyviä tekijöitä vaan ainoastaan antaa käyttäjille riittävän hyödyllisiä suosituksia.

Jos ennustettavaan asiaan vaikuttavia tekijöitä pitää pystyä datan avulla ymmärtämään paremmin, ovat esimerkiksi yllä kuvatut deep learning -menetelmät väärä työkalu. Niin ikään pitää tiedostaa, että pelkkä datan määrä ei ole automaattisesti tae sen hyödyllisyydestä.

"Roskaa sisään, roskaa ulos" -periaate pätee, vaikka aineisto olisikin iso ja malli hieno ja monimutkainen.

Lisäksi, jos malli on eräänlainen "musta laatikko", josta on vaikea sanoa miksi se ylipäättään toimii, voi olla vaikea löytää perusteltuja tapoja parantaa datan laatua kun mallin suorituskykyä halutaan parantaa.

Jos ilmiöstä ei ole tarkempaa ymmärrystä ja malli perustuu vain havaintoihin, on seuraavista toimenpiteistä hankalaa tehdä järkeviä päätöksiä. Iso data ja koneoppiminen eivät siis automaattisesti voi korvata esimerkiksi satunnaistettuja vertailukokeita tai muita tutkimusmenetelmiä, jos on tarve mitata tutkittavan ilmiön syy-seuraussuhteita.

Ovatko yritykset valmiita soveltamaan ison datan teknologioita?

Artikkelimme keskeinen viesti on, että useat ison datan teknologiat ovat olleet ratkaisuja niihin melko tarkkaan rajattuihin ongelmiin, joi-

Iso data ja koneoppiminen eivät korvaa perinteisiä tutkimusmenetelmiä, jos on tarve mitata tutkittavan ilmiön syy-seuraussuhteita.

ta ensimmäiset globaalit internet-yritykset kuten Google, Yahoo, Facebook, Amazon ja Twitter ovat joutuneet ratkaisemaan.

Suomessa vielä melko harvan yrityksen tietojenkäsittelyhaasteet ovat mittakaavaltaan samantaisia, joskin maailmanlaajuisen menestyksen myötä myös osa suomalaisista yrityksistä kohtaa vastaavanlaisia ongelmia ja mahdollisuuksia. Digitalisaatio, esineiden internet (Internet of things), ja yleistyvä datan keruu erilaisissa järjestelmissä tuovat nämä haasteet internet-palveluiden maailmasta myös perinteisen teollisuuden piiriin.

Mikään yksittäinen ison datan teknologia ei kuitenkaan ratkaise kaikkia tietojenkäsittelyn ongelmia. Sopivia ratkaisuja etsittäessä on tiedostettava kunkin yrityksen erityiset haasteet, toimitatavat, heikkoudet ja vahvuudet. On järjestelmällisesti rakennettava näihin sopivia ratkaisuja.

Teknisessä mielessä liiketoimintaa on nykyisin helppo mitata ja kehittää systemaattisesti ja datalähtöisesti. Parhaimmillaan datalähtöisen ajattelun kautta päätöksiä on mahdollista tehdä ilman liiketoiminnan ytimessä ja ruohonjuuritasolla, toisin sanoen nopeuttaa, parantaa ja automatisoida toimintoja juuri siellä, missä asiakas on.

Kun teknologia ei ole este, ovat varsinainen ongelma ennemminkin organisaatioiden vanhat, hitaat, byrokraattiset ja hierarkkiset päätöksentekorakenteet. ■

Hannes Heikinheimo on tekniikan tohtori ja asiantuntija Reaktor Innovations Oy:ssä. Antti Ukkonen on tekniikan tohtori ja tutkija Työterveyslaitoksella.

Lähteet:

Anderson, Chris 2008. *The end of theory*. *Wired magazine* 16, no. 7 (2008).

Brin, Sergey & Page, Lawrence 1998. *The anatomy of a large-scale hypertextual Web search engine*. *Computer networks and ISDN systems* 30, no. 1 (1998).

Chang, Fay & Dean, Jeffery & Ghemawat, Sanjay & Hsieh, Wilson C. & Wallach, Deborah A & Burrows, Mike & Chandra, Tushar & Fikes, Andrew & Gruber, Robert E. 2008. *Bigtable: A distributed storage system for structured data*. *ACM Transactions on Computer Systems (TOCS)* 26, no. 2 (2008).

Dean, Jeffrey & Ghemawat, Sanjay 2008. *MapReduce: simplified data processing on large clusters*. *Communications of the ACM* 51, no. 1 (2008).

Gates, Alan F. & Natkovich, Olga & Chopra, Shubham & Kamath, Pradeep & Narayanamurthy, Shravan M. & Olston, Christopher & Reed, Benjamin & Srinivasan, Santhosh & Srivastava, Utkarsh 2009. *Building a high-level dataflow system on top of MapReduce: the Pig experience*. *Proceedings of the VLDB Endowment* 2, no. 2 (2009).

Ginsberg, Jeremy & Mohebbi, Matthew H. & Patel, Rajan S. & Brammer, Lynnette & Smolinski, Mark S. & Brilliant, Larry 2014. *Detecting influenza epidemics using search engine query data*. *Nature* 457, no. 7232 (2009).

Hinton, Geoffrey & Deng, Li & Yu, Dong & Dahl, George E. & Mohamed, Abdel-rahman & Jaitly, Navdeep & Senior, Andrew 2012. *Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups*. *Signal Processing Magazine, IEEE* 29, no. 6 (2012).

Lazer, David & Kennedy, Ryan & King, Gary & Vespignani, Alessandro 2014. *The Parable of Google Flu: Traps in Big Data Analysis*. *Science* 343, no. 6176 (2014).

Norvig, Peter 2008. *All we want are the facts, ma'am*. <http://norvig.com/fact-check.html>.

Salakhutdinov, Ruslan & Hinton, Geoffrey E. 2009. *Deep boltzmann machines*. Konferenssipaperi. *International Conference on Artificial Intelligence and Statistics*.

Shvachko, Konstantin & Kuang, Hairong & Radia, Sanjay & Chansler, Robert 2010. *The hadoop distributed file system*. *IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST) 2010*, pp. 1-10.

Thusoo, Ashish & Sen Sarma, Joydeep & Jain, Namit & Shao, Zheng & Chakka, Prasad & Anthony, Suresh & Liu, Hao & Wyckoff, Pete & Murthy, Raghotham 2009. *Hive: a warehousing solution over a map-reduce framework*. *Proceedings of the VLDB Endowment* 2, no. 2 (2009).

Tuhkuri, Joonas 2014. *Big Data: Google-haut ennustavat työttömyyttä Suomessa*. ETLA-raportti nro 31. <http://www.etla.fi/julkaisut/33195/>.

TARJA HATAKKA

Soveltuvatko kaupan kanta-asiakasrekisterit kulutuksen tilastointiin?

Kauppojen kanta-asiakasrekisterit ovat yksi mahdollinen uusi tietolähde kotitalouksien kulutuksen tilastointiin. Edellytyksenä on, että monia käytännön ongelmia ja tietosuojaan liittyviä tekijöitä saadaan ratkaistua. Tulevaisuudessa tiedot tullaan todennäköisimmin keräämään monesta eri lähteestä. Tällöin on kiinnitettävä huomiota erityisesti tietojen ajalliseen vertailtavuuteen.

Tilastokeskuksen kulutustutkimuksen tietoista lähes puolet perustuu päivittäistavaraostoihin, joista haastatteluun mukaan valitut kotitaloudet keräävät kuitit talteen kahden viikon ajalta.

Pohdin artikkelissani suurten kauppaketjujen kanta-asiakasjärjestelmien mahdollisuutta toimia päivittäistavarahankintojen uutena tietolähteenä tulevaisuuden kulutustutkimuksissa sekä arvioin näiden käytettävyyttä tilastoinnissa.

Kulutustutkimuksen tehtävä on tuottaa tietoa yksityisen kulutuksen rakenteesta väestöryhmittäin. Kulutustutkimus on itse asiassa Suomessa ainoa tietolähde, jonka avulla voidaan tarkastella väestöryhmien ja alueiden välisiä kulutuseroja. Toistaiseksi tietoa kotitalouksien kulutusmenoista ei ole ollut mahdollista saada muuten kuin kysymällä asiaa kotitalouksilta suoraan.

Tietoa kotitalouksien kulutuksesta on kerätty Suomessa vuodesta 1966 lähtien noin viiden vuoden välein käyttäen kahta tiedonkeruumenetelmää: haastattelua ja kulutuspäiväkirjaa. Haastattelussa on kysytty kotitalouksien suuret hankinnat (esim. kodinkoneiden, huonekalujen ja autojen ostot) ja säännölliset kuukausimenot. Sen sijaan päivittäistavarojen ostot on pyydetty kirjaamaan kahden viikon ajalta päiväkirjaan. Näin tiedot kerätään perinteisesti myös muissa maissa. (Hatakka & Kajantie 2013.)

Pisimmällä kotitalouksien kulutustilastoinnin uudistamisessa tällä hetkellä lienee Norjan tilastovirasto. Norjassa vuoden 2017 kulutustutkimuksessa on tarkoitus yhdistää digitaalista pankki- ja luottokorttien transaktiotietoa kyselyin saataviin tietoihin (Haraldsen ym. 2015).

Suomessa päivittäistavaraostot on 2000-luvulla kerätty pääosin siten, että kotitalous on kerännyt ostokuitit talteen kahden viikon ajalta. Tämä uudistus vähensi kotitalouden tiedonantorasitusta merkittävästi päiväkirjan pitoon verrattuna. Lisäksi Tilastokeskuksessa tehtävä manuaalinen työ väheni huomattavasti, kun kuittitiedot skannattiin ja digitointiin.



Uudistuksesta huolimatta kriittiselle tasolle laskenut vastausaste ja tiedonkeruun suuret kustannukset ovat suuria haasteita kulutuksen tilastoinnille. Kulutustutkimus on tutkimukseen osallistuvalla kotitaloudella työläämpi kuin moni muu tutkimus, jossa kotitalouden työ rajoittuu haastatteluun vastaamiseen. Vaikka Tilastokeskuksessa käytetyt kulutustutkimuksen tiedonkeruumenetelmät ja -välineet ovat kansainvälisesti hyvää tasoa myös kustannustehokkuuden näkökulmasta, Tilastokeskuksen toiminnan rahoituksen kiristyminen rajoittaa kulutustutkimusten tapaisten harvoin tehtävien ei-lakisääteisten tutkimusten tekemistä. Vuodeksi 2015 suunniteltu kulutustutkimus peruttiin taloudellisista syistä. Harkittavana on, että se tehtäisiin vuonna 2016.

Uudet luovat tiedonkeruuratkaisut ovat tulevaisuudessa välttämättömiä, sillä sekä kansallisesti että kansainvälisesti kulutustutkimuksen tiedot ovat erittäin kysytyjä ja tarpeellisia. Kulutustutkimuksen aikasarja-aineistoa käyttävät mm. lukuisat tutkijat ja yhteiskunnallisen päätöksenteon valmistelijat yhteiskunnan muutosten – kuten eriarvoisuuden kehitys, verotusmuutosten vaikutukset eri väestöryhmiin tai säästämisen ja kulutuksen kehitys – analysointiin. Tiedot ovat pohjana myös mm. sosiaaliturvan riittävyyden arvioinnissa ja kulluttajahintaindeksin laadinnassa.

Toistaiseksi kulutustutkimuksen toteuttamiselle otostutkimuksena ei ole nähty realistista vaihtoehtoa, mutta uusia mahdollisuuksia on syytä tutkia. Yksi tällainen mahdollisuus saattaisi

olla suurten kaupparyhmien kanta-asiakasjärjestelmien hyödyntäminen. Tätä asiaa on alustavasti selvitetty Tilastokeskuksessa, mutta vain siten, että asiakasrekisterin tiedoilla korvattaisiin kotitalouksien 14 päivän kuittienkeruu. Tästä on vielä pitkä askel uusien aineistojen hyödyntämiseen.

Kanta-asiakasrekisterit eivät nykyisellään täytä tilastoinnin tarpeita

Suomessa päivittäistavarakauppa on varsin keskittynyttä. Kahden suurimman kauppaketjun markkinaosuus päivittäistavarakaupasta oli noin 80 prosenttia vuonna 2013 (taulukko).

Päivittäistavarakaupan markkinaosuudet 2013

KAUPPARYHMÄ	MARKKINAOSUUS 2013, %
S-ryhmä	45,7
K-ryhmä	34,0
Suomen Lähikauppa Oy	7,0
Lidl Suomi Ky	6,6
Muut	6,8
Yhteensä	100,0

Lähde: Nielsen 2014



Tilastokeskuksen syksyllä 2013 tekemän selvityksen mukaan vain 1.2.2012 jälkeen K-Plus-asiakkaiksi liittyneiden ostotapahtumat olisivat olleet kulutustutkimuksen tarpeisiin riittävät, sillä ne rekisteröitiin tuotetasolla. Kaikkien tätä aiemmin K-ryhmän kanta-asiakkaiksi liittyneiden ostotiedot rekisteröitiin vain kuitin loppusumma- tai tuoteryhmätasolla, mikä ei riitä kulutusmenojen ryhmittelyyn kansainvälisen COICOP-luokituksen¹ mukaisesti. K-ryhmän osalta tilanne on tänä keväänä muuttumassa, sillä 1.4.2015 alkaen kaikkien K-Plus-asiakkaiden ostotapahtumat rekisteröidään tuotetasolla. (Leppänen 2015; K-ryhmä 2012a; K-ryhmä 2012b). Näin ollen K-ryhmän osalta tietosisältö olisi jo vuonna 2016 riittävä kulutustutkimuksen tarpeisiin.

S-ryhmä teki samansuuntaisen muutoksen 1.5.2014 alkaen, mutta kuitenkin siten, että siirtymä oli pikemminkin kuitin loppusummatasosta tuoteryhmätasolle kuin tuotetasolle. Suuri osa päivittäistavaraostoista (esimerkiksi ostokset S-Marketissa, Prismassa, Kodin Terrassa ja ABC-huoltamoilla) kerätään rekisteriin vain tuoteryhmätasoisena. Tuotetasolla tiedot kertyvät esimerkiksi Sokos-, Emotion- ja verkkokauppaostoista. (S-ryhmä 2014a; S-ryhmä 2014b). Kulutustutkimuksen tavoitteisiin S-ryhmän tiedot eivät siis edelleenkään ole riittäviä tuoteryhmätasoisien tietojen osalta.

Ruokakaupoista Lidlillä ei ole lainkaan kanta-asiakasjärjestelmää. Suomen Lähikaupalla oli oma kanta-asiakasohjelma YkkösBonus, mutta keväällä 2014 se vaihtui PINS-ohjelmaan. PINS-ohjelmassa on mukana Suomen Lähikaupan lisäksi useita verkkokauppoja kuten suomalainen Verkkokauppa.com ja useita kansainvälisiä verkkokauppoja. (PINS 2014a.) PINS-jäsenrekisteriin kerätään ostotapahtuman tiedot tuoteryhmä- tai tuotetasolla (PINS 2014b). Tiedot saatvaisivat näin ollen osittain palvella kulutustutkimuksen tavoitteita.

Tämänhetkisten tietojen perusteella vain K-ryhmän kanta-asiakasrekisteri täyttäisi kulutustutkimuksen tietotarpeet, jos seuraava tiedonkeruu toteutetaan vuonna 2016. Käytännössä kulutustutkimuksen nykyisten tietotarpeiden täyttämiseksi on vielä muitakin teknisiä vaatimuksia. Tuotteiden on oltava sellaisessa muodossa, että niiden luokittelu COICOP-luokituksen mukaisiin ryhmiin on mahdollista. Lisäksi kulutustutkimuksessa on tuotettu tietoa elintarvikkeiden kulutusmääristä. Tämän tietotarpeen täyttäminen

edellyttää, että ostetuista elintarvikkeista on saatavissa paitsi tuotekohtainen hintatieto myös tuotekohtainen määrätieto.

Kauppojen tietojen käyttö edellyttää kotitalouksien lupaa

Kuten edellä todettiin, vain K-ryhmän tiedot täyttävät tällä hetkellä (1.4.2015 alkaen) kulutustutkimuksen tietotarpeet. Jos kuitenkin oletetaan, että kaikkien kaupparyhmien tiedot palvelisivat kulutustutkimuksen tietotarpeita, pohditavaksi jäävät vielä mm. tietosuojaa-asiat sekä käytännön toimintatavat.

Tilastolain (Tilastokeskus 2013) mukaan tilastoihin tulee ensi sijassa käyttää muussa yhte-

Tiedot kotitalouksien kulutuksesta ja sen jakautumisesta ovat kysytyjä kansallisesti ja kansainvälisesti.



¹ Yksilöllisen kulutuksen käyttötarkoituksen mukainen luokittelu (COICOP-luokitus). Kts. <http://tilastokeskus.fi/meta/luokitukset/coicop/011-2013/index.html>.

ydessä kerättyjä tietoja. Suurin osa tilastojen tekoa varten tarvittavista tiedoista saadaankin erilaisista rekistereistä, ja valtion viranomaisilla on yleinen tiedonantovelvollisuus hallussaan olevista tiedoista. Sen sijaan yrityksillä on velvollisuus antaa tietoa vain laissa erikseen määritellyissä asioissa, jotka koskevat lähinnä yrityksen omaa toimintaa. Näin ollen lähtökohtana on, että lupa kotitalouden ostotietojen saamiseen kaupparyhmillä tulee saada kotitaloudelta itseltään.

Lupa pitäisi saada myös rinnakkaiskorttien käyttäjiltä, koska tutkimuskohteena on koko kotitalous. Olettavasti osa tutkimukseen valikoiduista henkilöistä kieltäytyisi tutkimuksesta lupamenettelyn vuoksi, vaikka se olisi vapaaehtoinen järjestely. Norjan tilastovirasto kysyi kesällä 2014 kotitalouksilta halukkuutta antaa pankki- ja luottokorttitietoja tilastovirastolle tilastointitarkoitusta varten. Hieman alle 40 prosenttia vastanneista suhtautui asiaan myönteisesti (Haraldsen ym. 2015). Suhtautumista asiaan pitäisi selvittää etukäiteistillä myös Suomessa.

Lupamenettely tarkoittaisi, että Tilastokeskus toimittaisi allekirjoitetut luvat ja bonuskorttien numerot sekä kuhunkin korttiin liittyvän poimintajakson tiedot kauppaketjuille, jotka tekisivät tarvittavat poiminnat ja toimittaisivat ko. kotitalouksien ostotiedot sähköisessä muodossa Tilastokeskukselle. Sekä tietosuojaan että byrokratian kannalta yksinkertaisempi vaihtoehto olisi, jos kotitalous saisi kauppaketjulta helposti omat ostotietonsa sähköisessä muodossa edelleen Tilastokeskukselle toimitettavaksi.

Mahdollinen säästö riippuu tiedonkeruun toteutustavasta

Vuonna 2012 kulutustutkimuksen tiedonkeruussa kuitteja tuli kahden viikon keräysjaksolta keskimäärin yksi kuitti päivää kohden eli yhteensä 14 kuittia kotitaloutta kohden. Kuittien kerääminen ei ole siis kovin suuri vaiva. On tietysti toivottavaa, että kotitalouksia ei vaivata turhaan, jos tieto on saatavissa muuten. Kuittien keräämiseen verrattuna lupien antaminen on kotitalouksille kuitenkin lähes yhtä suuri vaiva – etenkin jos kotitaloudella on useamman kaupparyhmän bonuskortteja.

Bonuskorttien käyttölupien käsittely lisäisi Tilastokeskuksessa haastattelijoiden ja kulutustilaston työntekijöiden työtä. Lisäksi kotitalouskohtaisten ostorivitietojen poiminta tietyltä ajanjaksolta vaatii ohjelmointi- ja muuta työtä myös kaupparyhmässä. Kokonaisuutena kustannukset jakautuisivat hieman eri tavalla kuin nykyisin, mutta eivät todennäköisesti juurikaan vähenisi verrattuna nykyiseen menettelytapaan. Jos kotitalous voisi itse suoraan toimittaa kauppaketjun sähköisestä järjes-

telmästä saamansa tiedot Tilastokeskukselle, käyttölupien käsittelytyötä ei tarvittaisi. Tämä vaihtoehto olisi jo varsin kustannustehokas sekä kotitalouden että tilastoviraston kannalta.

Kokonaisaineiston käyttöönotto edellyttää huolellista pohdintaa ja selvitystyötä

Mikäli kanta-asiakasjärjestelmät tulevaisuudessa riittäisivät palvelemaan kulutustilastoinnin tarpeita, on selvää, että aineiston käyttöönotto edellyttäisi huolellista pohdintaa. Tilastoettisten normien kunnioittaminen on Suomen virallisen tilastotuotannon suuri vahvuus, josta ei missään tilanteessa tingitä. Toisaalta myös tilastojen luotettavuus on ensiarvoisen tärkeää. Kulutustutkimuksessa – ja muissa otostutkimuksissa – alhainen vastausaste on uhka tilaston luotettavuudelle, joten uusia tilastontuottamistapoja on etsittävä.

Aineistona kauppojen asiakasrekisterit eivät olennaisesti eroa julkisen hallinnon rekistereistä, joita Suomen tilastotuotannossa käytetään erittäin laajasti (ks. Jari Niemisen artikkeli tässä numerossa). Vuonna 2014 Tilastokeskuksessa toteutettiin Euroopan tilastojen käytännesääntöjen vertaisarviointi (Eurostat 2015). Arvioinnin tulokset olivat varsin myönteiset, ja yhtenä vahvuutena arviointiraportissa mainitaan laaja rekisteriaineistojen käyttö. Hallinnollisten aineistojen ansiosta yritysten ja kotitalouksien tiedonantotaakka pysyy kohtuullisena, ja menettely on myös kustannustehokas.

Toisaalta arvioinnissa huomautettiin, että hallinnollisten aineistojen käytössä on myös riskejä. Koska aineiston ensisijainen tarkoitus ei ole tuottaa tilastoja, aineiston tietosisältö voi nopeastikin muuttua ja aiheuttaa ongelmia aineistosta riippuvaisen tilaston tuottamiseen. Yksityisten toimijoiden hallussa olevien aineistojen edut ja riskit olisivat luonnollisesti samat. Kuten Suomen Lähi-kaupan esimerkki osoittaa, kauppa voi nopeastikin vaihtaa kanta-asiakasohjelmaansa.

Voisivatko kauppojen kanta-asiakasrekisterit kokonaisuutena toimia kulutustutkimuksen tietolähteenä? Toisin sanoen, olisiko kulutustutkimus mahdollista tehdä kokonaan rekisteripohjaisena käyttäen kokonaisaineistoja? Jotta tämä olisi mahdollista, tarvittaisiin kokonaisaineistoja paitsi päivittäistavarakaupasta myös erikoiskaupasta eli esimerkiksi huonekalu-, kodinkone-, vaate- ja kenkäkaupasta. Erikoiskauppa muodostaa lähes puolet koko vähittäiskaupasta, ilman autokauppaa (Erikoiskaupan Liitto).

Vaikka osa erikoiskaupasta on ketjuuntunutta, suuri osa on yhä pieniä yksityisyrittäjiä. Käytännössä on vaikea nähdä, miten erikoiskaupan tietoja saisi kerättyä edustavasti ilman haastatte- ►

lututkimusta. Tavaroiden lisäksi tarvitaan tietoa myös erilaisten palveluiden ostoista, joten kaupan kokonaisaineistot eivät ”täydellisinäkään” riittäisi täyttämään kulutustutkimuksen tarpeita.

Toistaiseksi on realistista perustaa kulutuksen tilastointi haastattelumenetelmään ja kuittien keräämiseen.

Entä voisiko kulutustutkimuksen tehdä tulonjakotilaston otokselle käyttäen ns. rekisterimenetelmää, kuten Tilastokeskuksen varallisuustutkimuksessa tehdään? Tulonjakotilasto sisältää jo nyt paljon yhteisiä osia kulutustutkimuksen kanssa. Tällaisia ovat ainakin tiedot kotitalouden rakenteesta ja toiminnasta, tuloista, asumisesta, asumismenoista ja päivähoitomaksuista. Haastattelut tehdään keväisin, joten käytännössä tämä lähestymistapa voisi tarkoittaa, että määrävuosin tulonjakotilaston haastatteluun liitettäisiin kauppojen kanta-asiakasrekisterien kokonaisaineistot koko edelliseltä kalenterivuodelta. Haastatteluun tarvittaisiin toki lisäkysymyksiä ainakin kestotavaroiden ostoista, vakuutusmenoista, matkustamiseen käytetystä rahasta ja muutamista säännöllisistä kulutusmenoista (lehdet, puhelin, netti).

Tällainen lähestymistapa mullistaisi nykykäytännön ja luonnollisesti myös tarkoittaisi, että kulutustutkimuksen aikasarja katkeaisi. Todennäköisesti tämä menettely tarkoittaisi myös kysytävien tietojen ja tuotettavien tulosten tarkkuustason heikkenemistä.

Toisaalta koko kalenterivuodelta saatavat varsin kattavat tiedot kotitalouksien päivittäistavaraostoista parantaisivat tietopohjaa otoksen kahden viikon jaksoilta kerättyihin tietoihin verrattuna. Tämä visio voisi olla jossain määrin realistinen olettaen, että kauppojen aineistot olisivat käytettävissä. Mutta toistaiseksi näin ei ole – vielä tällä vuosikymmenellä lienee realistista perustaa kotitalouksien kulutuksen tilastointi nykyiseen haastattelumenetelmään ja kuittien keräämiseen. ■

Kirjoittaja on yliaktuaari Tilastokeskuksen väestö- ja elinoloilastot -yksikössä.

Lähteet:

Durr, Jean-Michel & de Pourbaix, Isabelle & Redmond, Adrian 2014. Peer Review Report on Compliance with the Code of Practice and the Coordination Role of

the National Statistical Institute Finland. Eurostat, December 2014. <http://ec.europa.eu/eurostat/documents/64157/4372828/2015-FI-Report/688a9a11-741e-4311-b2f8-a0943104c933>.

Erikoiskaupan liitto. http://www.erikoiskaupanliitto.fi/cms/media/tilastot/Erikoisk_vk_osuudet_2012.pdf.

Haraldsen, Gustav & Amdam, Sverre & Zhang, Li-Chun 2015. Towards an integrated Consumer Expenditure Survey — Combining Multi-mode Data Collection and Big Data Extracts. http://www.cros-portal.eu/sites/default/files/Haraldsen-et-al_NTTTS2015abstract-20140715-unblinded_0.pdf.

Hatakka, Tarja & Kajantie, Mira 2013. Mitä ostit tänään? – vaihtoehtoja kulutuksen tutkimiseen. Hyvinvointikatsaus 4/2013. http://www.tilastokeskus.fi/artikkelit/2013/art_2013-12-09_007.html.

K-ryhmä 2012a. K-Plussa-järjestelmän säännöt 1.2.2012 lähtien. <https://www.plussa.com/Mika-on-K-Plussa/Sopimusehdot/K-Plussan-saannot/>.

K-ryhmä 2012b. K-Plussa-asiakastiedoston rekisteriseloste 17.12.2012. Henkilötietolain (523 /1999)10 §:n mukainen rekisteriseloste, K-Plussa-asiakastiedosto.

Leppänen, Marko 2015. Haluatko parempia etuja? Pirkka 1–2/2015. <http://www.digipaper.fi/pirkka/126861/index.php?pgnumb=41>.

Nielsen 2014. Päivittäistavara-kaupan menestyjä vuonna 2013 oli Lidl. <http://www.nielsen.com/fi/fi/press-room/2014/paivittaeistavara-kaupan-menestyjae-vuonna-2013-oli-lidl.html>.

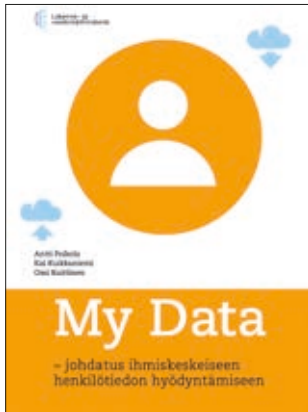
PINS 2014a. PINS jyllää Suomen markkinoille kymmenien yritysten voimalla. <http://www.pins.fi/fi/news/2014/05/pins-jyllaae-suomen-markkinoille-kymmenien-yritysten-voimalla>.

PINS 2014b. PINS-ohjelman tietosuojaseloste. <http://www.pins.fi/fi/privacy-policy>.

S-ryhmä 2014a. S-ryhmän asiakasomistaja- ja asiakasrekisterin rekisteriseloste. <https://www.s-kanava.fi/web/s/s-kanavan-rekisteriseloste>.

S-ryhmä 2014b. Asiakasomistaja- ja asiakasrekisterin tietosisältö tarkentuu. <https://www.s-kanava.fi/web/s/asiakasomistaja-ja-asiakasrekisterin-tietosisalto-tarkentuu>.

Tilastokeskus 2013. Tilastolaki. <http://tilastokeskus.fi/meta/lait/tilastolaki.html>.



[JULKAISU]

Oman datan hallinta on perusoikeus

My Data eli omadata on henkilötietojen hyödyntämismalli, jonka perusajatukseen on tarjota yksilölle pääsy häntä koskevaan digitaaliseen tietoon. Liikenne- ja viestintäministeriön syksyllä 2014 ilmestyneessä raportissa korostetaan kansalaisten oikeutta olla perillä heitä koskevista aineistoista ja vaikuttaa siihen, miten henkilötietoja hyödynnetään.

Henkilötietoja kertyy julkishallinnon rekistereihin, yritysten asiakasjärjestelmiin ja erilaisiin verkkopalveluihin mm. osto-, tele-, terveys- ja taloustietojen muodossa. Omadata-ajattelussa kansalaisen katsotaan hallitsevan henkilötietojaan. Hänellä on oikeus jakaa, vaihtaa ja myydä tietojaan sopivaksi katsomallaan tavalla.

Digitaalinen jalanjälki kasvaa myös älypuhelinsovellusten, puettavien sensorien ja muiden teknisten apuneuvojen myötä. Sovellusten käyttäjistä kertyy dataa, joka liittyy esimerkiksi ruokailuun, mielialoihin, fyysiseen ja henkiseen suorituskykyyn, ajan- ja rahankäyttöön, sosiaaliseen vuorovaikutukseen, liikkumiseen ja nukkumiseen. Näiden asioiden seuraamisessa on usein kyse oman terveyden, hyvinvoinnin ja elämänlaadun parantamisesta tiedon avulla. Puhutaan itsensä mittaamisen (*quantified self*) ilmiöstä (ks. oheinen taulukko).

Oman datan haltuunottoon on herätty eritoten Yhdysvalloissa, jossa lukuisat yritykset (*data brokers*)

<p>Kulutus</p> <ul style="list-style-type: none"> • Kalorinkulutus ja ravinto • Ravintoaineet • Alkoholinkulutus • Tupakointi • Kahvi • Vesi • Lääkkeet 	<p>Kehon toiminta</p> <ul style="list-style-type: none"> • Kehon PH • Kuukautiskierto • Raskaus • Ulostaminen 	<p>Fyysinen aktiivisuus</p> <ul style="list-style-type: none"> • Liikuntasuoritukset • Nukkuminen, unen laatu • Matkustaminen • Seksuaalinen aktiivisuus • Hampaiden pesu • Asento • Askelmäärä ja aktiivisuus
<p>Oireet</p> <ul style="list-style-type: none"> • Päänsäryt • Kivut • Astma-kohtaukset • Allergiat 	<p>Tila ja aika</p> <ul style="list-style-type: none"> • Sijainti • Korkeus • Aika • Mitä näet 	<p>Fysiologiset suureet</p> <ul style="list-style-type: none"> • Syke • Verenokeri • Kehon lämpö • Verenpaine • Paino ja kehon koostumus • Hengitys
<p>Mielen hyvinvointi</p> <ul style="list-style-type: none"> • Mieliala • Stressi • Vireystila 	<p>Ajankäyttö</p> <ul style="list-style-type: none"> • Median kulutus • Tietokoneohjelmien käyttö • Työaika • Sähköpostiviestien määrä • Puhelut ja muu viestintä 	<p>Muuta</p> <ul style="list-style-type: none"> • Rahan käyttö • Ostokset • Sosiaaliset verkostot • Oppimistulokset • Liikenne • Energiankulutus

Esimerkkejä datasta jota itsensä mittamisen työkaluilla ja sovelluksilla voi kerätä

Lähde: My Data -selvitys, s. 49

keräävät ja myyvät eteenpäin tietoja kuluttajista ilman, että nämä itse tietävät asiasta. Tietovarantoihin on kerätty tietoja käytännössä kaikista kuluttajista. Paikallisen kuluttajasuojaviranomaisen (Federal Trade Commission) mukaan yritysten toiminta voi vaarantaa kansalaisten yksityisyydensuojan.

Omadata-ajattelun ympärille on syntymässä uutta teknologiaa ja liiketoimintaa eri puolilla maailmaa. My Data -selvitys on yleiskatsaus omadata-mallin mukaisesti henkilötietoja hyödyntäviin kotimaisiin ja kansainvälisiin toimijoihin ja toimintatapoihin. Raportissa pohditaan myös, mitä haasteita tai ongelmia mallin käyttöönottoon liittyy. ■

Lähde:

Antti Poikola & Kai Kuikkaniemi & Ossi Kuittinen 2014. My Data. Johdatus ihmiskeeseen henkilötiedon hyödyntämiseen. Liikenne- ja viestintäministeriö sekä Open Knowledge Finland ry.

Selvitys on saatavilla liikenne- ja viestintäministeriön verkkosivuilta.

Katso myös netissä Antti Poikolan esitys My Datasta Otavan opistolla lokakuussa 2014 (<https://www.youtube.com/watch?v=tv-9j6o4sTU>).

REIJO SUND

Miksi isoon dataan hukutaan?

Ison datan jalostaminen hyödylliseksi tiedoksi on vaikeaa. Data ei muunnu informaatioksi ilman tietoa ja ymmärrystä tutkittavan ilmiön taustatekijöistä.

Iso data on noussut yhdeksi aikamme muoti-termiksi, jota ryöstöviljellään miltei alalla kuin alalla. Aina ei ole selvää mitä sillä oikein loppujen lopuksi tarkoitetaan.

Hakusukupolven tyypillinen edustaja hyvin todennäköisesti etsisi aiheesta lisätietoa vaikkapa Googlen hakukoneella ja päätyisi esimerkiksi vapaan tietosanakirja Wikipedian aihetta käsittelevään artikkeliin.

Kirjoitushetkellä suomenkielisen Wikipedian mukaan iso data on ”erittäin suurten, järjestelemättömien, jatkuvasti lisääntyvien tietomassojen keräämistä, säilyttämistä, jakamista, etsimistä, analysointia sekä esittämistä tilastotiedettä ja tietotekniikkaa hyödyntäen” (Wikipedia 2015).

Samassa lähteessä todetaan myös, että isolle datalle on tyypillistä sen määritelmän muuttuminen ajan myötä, kun data-analyysin mahdollisuuksien parantuessa aikaisemmin vaikeuksia tuottaneita aineistoja pystytään hyödyntämään jo niin helposti, ettei moisia enää voida kutsua isoksi dataksi (Wikipedia 2015).

Väitän, ettei isosta datasta olisi tullut muoti-termiä, jollei siihen liittyisi mahdollisuutta tavoitella huomattavan suurta taloudellista voittoa. Ilmiöön liittyy olennaisesti hehkus, jossa jälleen kerran (yritysten) upeat tietovarannot luvataan taianomaisesti muuttua (liiketoimintaa tukevaksi) hyödylliseksi informaatioksi.

Ennen 2010-luvulla läpilyönyttä isoa dataa käytännössä samaa ilosanomaa jaettiin ainakin businessälykkyyden (business intelligence, 2000-lu-

ku), tiedonlouhinnan (data mining, 1990-luku) ja asiantuntijajärjestelmien (expert systems, 1980-luku) nimillä, mutta näiden nimitysten hohdotto on hiipunut, kun ne eivät täyttäneetkään kaikkia suuria lupauksiaan.

Jää nähtäväksi, missä vaiheessa massadatan jälleen avaamat rahahanat sulkeutuvat seuraavan vastaavan villityksen ja uusien termien myötä.

Artikkelini tarkoitus on luoda kriittinen katsaus ison datan lupauksiin.

Data ei ole informaatiota

Data ei automaattisesti ole tai sisällä informaatiota. On syytä erotella toisistaan käsitteet data, informaatio ja tieto tai tietämys (knowledge).

Usein nämä käsitteet esitetään hierarkkisesti siten, että datasta tuotetaan informaatiota, joka sitten lisää tietämystä. Tämä on kuitenkin harhaanjohtava yksinkertaistus, eikä se ole sopuisoinnussa empiirisen tutkimuksen periaatteiden kanssa, varsinkin jos havaintoaineistot ja ilmiöt erotellaan toisistaan niin, että teoriat selittävät ilmiöitä eivätkä suinkaan havaintoaineistoa.

Itse asiassa on tarkoituksenmukaista erotella kolmekin niin sanotun kontekstuaalisen riippuvuuden tasoa: teoriapohjainen havainnointi, operationalisoitu data ja datan teoreettinen tulkinta (Faber & Scheper 2003).

Pragmaattisessa mittaamisteoriassa oletetaan, että jostain ilmiöstä pystytään tekemään havainnot ja havainnoista edelleen systemaattisia mitauksia, jotka voidaan sitten koota yhteen dataksi (Hand 2004). Tässä mielessä järkevää dataa eikä varsinkaan informaatiota voi edes olla olemassa ilman alustavaa taustatietämystä asioista.

Tulkintaan tarvittavan taustatietämyksen määrä voi kasvaa vielä enemmän kuin itse data.

Ruotsalainen tietojenkäsittelyn edelläkävijä Börje Langefors esitti 1970-luvulla niin sanotun infologisen yhtälön $I = i(D, S, t)$, jonka mukaan informaatiota I tuotetaan funktiolla i datasta D ja taustatietämyksestä S ajanhetkellä t (Langefors 1993). Jo tällainen puolihermeneuttinen lähestymistapa tarjoaa selityksiä vaikeuksiin tuottaa informaatiota datasta.

Myöhemmin on esimerkiksi varsin osuvasti todettu, että mitä monimutkaisemmasta datasta on kyse, sitä enemmän tarvitaan taustatietämystä sen

hyödyntämiseen (Berg & Goorman 1999). Toisin sanoen entistä suurempi määrä entistä tarkempaa dataa ei tule ratkaisemaan kaikkia ongelmia datan muuntamisessa informaatioksi. On otettava huomioon, että tulkintaan tarvittavan taustatietämyksen määrä voi kasvaa vielä enemmän kuin itse data.

Empiirisen tutkimuksen ideaalitalanteessa on olemassa riittävästi taustatietämystä tutkittavasta ilmiöstä vakiintuneen tieteellisen teorian muodossa. Silloin tästä tarkasteltavasta ilmiöstä P voidaan havainnoinnin ja mittaamisen keinoin tuottaa dataa D juuri sillä tavalla, kuin tutkija näkee sen tarkoituksenmukaiseksi.

Käytännössä siis todellisuus operationalisoidaan (hyvin suppeaksi) dataksi, joka sitten muodostaa tutkijan oman "leikkimaailman". Data-analyysin keinoin tästä leikkimaailmasta pyritään löytämään säännönmukaisuuksia.

Tutkimuksen onnistumisen kannalta on oleellista, että operationalisoinnin käännteistoimenpiteen eli tulkinnan avulla pystytään nostamaan datan heijastelemassa leikkimaailmassa havaittu säännönmukaisuus takaisin teoreettiselle tasolle tai jopa todellisuuteen. Tämä on helpointa tilanteissa, joissa data koostuu suhteellisen pysyvän todellisuuden suoraan havaittavista asioista.

Asiat monimutkaistuvat huomattavasti, jos kiinnostuksen kohteena oleva ilmiö ei ole suoraan havaittavissa tai jos ilmiötä määrittelevä todellisuus tai siihen oleellisesti vaikuttava konteksti vaihtelee rajusti¹.

Monissa yhteyksissä on tarkoituksenmukaista eritellä ilmiön tarkasteluun tarvittavia käsitteitä ja niiden välisiä yhteyksiä. Erityisen tärkeää on pyrkiä hahmottamaan yksittäiseen käsitteeseen liittyvää epävarmuutta, joka on usein kohtuullisen suoraan linjassa sen hyödyntämiseen tarvittavan taustatietämyksen suhteen.

Mikäli jo itse käsitteeseenkin liittyy epävarmuutta, ollaan hyvin perustavanlaatuisten tietoteoreettisten kysymysten äärellä. Onko käsitteen taustalla joku todella yksiselitteisesti olemassa oleva ja ilmenevä asia, jolloin vain sen havainnoiminen, operationalisointi ja mittaaminen ovat ongelmallisia ja sisältävät ehkä mittausvirhettä, vai liittyykö jo itse käsitteeseen epävarmuutta tai vaihtelua, jolloin kyseessä ei oikein voi olla mittausvirhe vaan käsitteeseen liittyvän ilmiön ominaisuus? (Sund 2005.)

Käytännön kannalta tällä on suurta merkitystä siinä suhteessa, että kaikki data ei ole perus-

¹ Tämä on tyypillisempää ihmis- ja yhteiskuntatieteellisten kuin luonnontieteellisten kysymystenasetteluiden kohdalla ja onkin johtanut tutkimusta ohjaavien metodologisten paradigmojen jonkinasteiseen eriytymiseen, joka vaikeuttaa asioista kommunikointia ja yhteisen ymmärryksen saavuttamista. ►

Miten havainnot jalostetaan tiedoksi

Aikojen alusta asti aistihavainnoilla ja niistä tehdyillä tulkinnoilla on ollut erittäin merkittävä rooli niin selviytymisen kuin lisääntymisenkin kannalta. Ihmisten erityispiirteeksi on kehittynyt symbolinen viestintä, joka mahdollistaa (hyödyllisen) informaation epäsuoran välittämisen ja kollektiivisen tietämyksen kertymisen; kaikkea ei tarvitse kokea omakohtaisesti.

Voimme kirjata havaintomme sellaiseen pysyvään muotoon, että myös muut pystyvät tulkitsemaan näitä havaintoja riittävän samalla tavalla. Käytännössä tämä tarkoittaa numeeristen merkintöjen ja varsinaisen kirjoitustaidon omaksumista.

Ensimmäistä näyttöä lukumäärien merkitsemisestä on yli 30 000 vuoden takaa (Hand 2004). Lukumääriä ilmeisesti merkittiin ylös erityisesti kaupankäyntiin sekä hallintoon liittyvistä asioista. Tämän otaksutaan johtaneen noin 5 000 vuotta sitten myös varsinaisen kirjoitustaidon syntymiseen.

Lukumääriä voidaan pitää myös ensimmäisinä ilmentyminä siitä, että havaittava asia tai ilmiö tarkoituksellisesti operationalisoidaan muotoon, jossa se esitetään symbolisesti numeroita käyttäen. Tällaiset kaupankäyntiin ja hallintoon liittyvät lukumääräkokelmat ovat

samalla myös ensimmäisiä ihmisen tarkoituksellisesti tuottamia systemaattisia havaintoaineistoja, ensimmäisiä datoja.

Varhainen massadata ja tieteen kehittyminen

Varsin pian on varmasti törmätty tilanteeseen, jossa tuotettu data ei enää sellaisenaan ole edustanut hyödyllistä informaatiota, vaan sitä on pitänyt alkaa jäsentelemään edelleen vaikkapa ajankohtien, hyödykkeiden tai alueiden suhteen.

On alettu laskea yksittäisten lukumäärien tarkastelun sijaan summia, jotta saataisiin hyödyllinen kokonaiskäsitys kiinnostuksen kohteena olevasta asiasta. Edellä mainittu tilanne voisi hyvin olla esimerkki varhaisesta massadatasta ja sen analysoinnista.

Manuaalisesti kerättävän aineiston tuottaminen on ollut niin työlästä, että käytännön syistä suurin osa varsinaisista historiallisista havaintoaineistoista heijastelee suoraan havaittavia ominaisuuksia tärkeiksi koetuista asioista. Pelkkää aineistojen keräämistä tärkeämpää on ehkä toiminnan käsittämisen systemaattisen tiedonhankinnan keinoksi, joka pikkuhiljaa johti tieteen kehittymiseen.

Antiikin ajoista 1700-luvulle tieteellinen

tutkimus ja filosofia olivat käytännössä erottamaton osa toisiaan. Galilein, Keplerin ja Newtonin töiden ansiosta luonnontieteet saivat oman erityisroolinsa todellisuuden kuvaajina, ja 1900-luvun alkuun mennessä oli muotoutunut nykyaikainen käsitys tieteestä. Sen mukaan tiedettä tekevien tieteilijöiden tavoitteena on tuottaa systemaattisesti uutta tutkimustietoa todellisuudesta käyttäen kriittistä tieteellistä menetelmää. Tämä antaa tieteellisesti tuotetulle tiedolle erityisen statuksen.

Käytännössä empiirisen tutkimuksen nousu on luonut tarvetta havaintoaineistojen systemaattiselle ja entistä monipuolisemmalta analysoinnille. Tähän data-analyysin tarpeeseen kehitettiin tilastomatemattisia menetelmiä, joiden avulla pystytään tekemään todennäköisyyslaskentaan perustuvia päätelmiä myös tilanteissa, joissa aineistoihin ja havaintoihin liittyy jonkunlaista epävarmuutta ja vaihtelua.

Tietokoneet ovat tehostaneet tietojen käsittelyä

Vielä 1900-luvun alkupuolella analyysien tekeminen oli haastavaa ja vaati tyypillisesti käsityönä tehtävää manuaalista laskemista.

luonteeltaan samanlaista eikä siis ole yhtä tapaa muuntaa dataa informaatioksi.

Sekundaarisen datan käytöstä lisää ongelmia

Edellä mainitut ongelmat on kohdattava jo silloin, kun empiiristä tutkimusta päästään tekemään ihanneolosuhteissa. Jos tilanne ei ole näin hyvä, vaan halutaan esimerkiksi välttyä uuden datan tuottamiselta ja käyttää jo jotain olemassa olevaa, tilanne ymmärrettävästi vain hankaloituu.

Jos siis päätetään käyttää tutkimuksessa sekundaarista dataa eli dataa, joka on alun perin tuotettu jotain muuta tarkoitusta varten (kuten yleensä aina massadatojen kohdalla), asettaa se lisärajoituksia koko tutkimusprosessille (Sund 2003). Tässä tapauksessa ei voida rätätelöidä datankeruuta tutkimusongelman mukaisesti vaan joudutaan käyttämään valmista dataa valmiine operationalisointineen.

Datan tuottamiseen käytetty teoria ja käsitteet voivat poiketa oleellissalakin tavalla tehtävän tutkimuksen vastaavista, ja tämä johtaa hankalaan tilanteeseen: miten tehdä tulkintoja sellaisesta ilmiöstä, jota ei voida havaita kuin toivotavasti eräänlaisina heijastuksina valmiiksi kerätystä datasta.

Käytännössä joudutaan usein tyytymään kompromissiin, jossa sekundaarisen datan mahdollisuudet ja teoreettinen kysymyksenasettelu jollain lailla saatetaan vastaamaan toisiaan. Tämä tekee tutkimuksesta usein opportunistista eli joudutaan tai tyydytään tarkastelemaan asioita aineistojen ehdoilla.

Datan käyttökelpoistaminen

Hyvin usein on tarpeen rikastaa dataa ylimääräisellä taustatietämyksellä niin, että siitä voidaan esikäsitteilyllä saada irti jotain tarkasteltavan ilmi-

1950-luvulta lähtien tietokoneet alkoivat kuitenkin yleistyä ja niiden myötä muuttuivat sekä aineistojen tallentamisen että analysoinnin mahdollisuudet. Monimutkaisiakin analyysejä pystyttiin tekemään aina isommille aineistoille, kun tietokoneet hoitivat yksitoikkoisen ja virhealttiin työn ihmistä paremmin.

Laskennallisten mahdollisuuksien paraneminen johti tilanteeseen, jossa taitamattomakin tutkijat pystyivät teknisesti viemään läpi monimutkaisia analyysejä. Esimerkiksi faktori-analyysejä suorastaan väärinkäytettiin niin, että sen suosio alkoi lopulta hiipua.

Rekisteritiedot ovat alikäytössä

Tietokoneet ovat mahdollistaneet myös merkittävästi entistä suurempien aineistojen tarkoituksenmukaisen taltioinnin. Suomessa on yhtenä ensimmäisistä maailmassa ryhtytty keräämään tapahtumatasoista tietoa valtakunnallisiin rekistereihin. Esimerkiksi kaikista sairaalapoistoista koko maassa kerättiin tiedot vuonna 1960. Tiedonkeruu ja taltiointi tietokantoihin on jatkunut yhtäjaksoisesti vuodesta 1967 lähtien (Härö 1969).

Samoihin aikoihin käyttöön otettu henkilötunnus tekee suomalaisista aineistoista vie-

läkin ainutlaatuisempia, sillä henkilötunnuksen avulla henkilö pystytään identifioimaan ja muodostamaan elämäntapahtumia kuvaavia tietopolkuja.

Idea samaa kohdetta koskevien havaintojen yhdistelemisestä on vanha (ks. esim. Dunn 1946), mutta monessa maassa tällaisista valtakunnallisista aineistoista ei ole edes usallettu unelmoida kuin kymmeniä vuosia Suomen jälkeen.

Ainutlaatuisia tietovarantojamme ei kuitenkaan ole hyödynnetty niin paljon kuin ehkä olisi kannattanut tai voinut. Tyypilliseen tapaan resursseja ja tahtoa on kyllä löytynyt tiedonkeruuseen, mutta varsinaiseen aineiston analysointiin niitä on ollut jaossa vain hyvin rajallisesti.

Tiedon tuottamisen pullonkaula

Laskentakapasiteetin lisääntyminen ei siis ole pystynyt ratkaisemaan ongelmaa muuntaa aineistoa hyödylliseksi informaatioksi ja edelleen tiedoksi, vaan sama iänikuinen pullonkaula vaivaa edelleen.

Toki aineistojen koot ovat edelleen kasvaneet ja dataa jatkuvasti tuottavien tietolähteiden määrä on digitalisoinnin ja teknologisen kehityksen myötä radikaalisti lisäänty-

nyt, mutta se ei ole ainoa ongelma. Käsitykset datojen luonteesta ja mahdollisuuksista ovat edelleen monilta osin yksinkertaisesti väärin ymmärrettyjä ja ylimitoitettuja.

Mitkään ”uudet” menetelmät, jotka itse asiassa monilta osin ovat tilastotieteessä jo pitkään tunnettujen menetelmien reinkarnaatioita toisilla nimillä, eivät pysty automaattisesti ratkaisemaan tai kiertämään empiiriseen tutkimukseen välttämättä kuuluvia reunaehjoja.

Vain tieteellisesti tuotettu tieto on uskottavaa

Tieteenfilosofiassa on pohdittu paljon näitä tiedon tuottamiseen liittyviä asioita. Käytännön kannalta voidaan keskeinen sanoma tiivistää toteamukseen, että jos tietoa ei tuoteta tieteellisen menetelmän periaatteita käyttäen, ei sillä ole uskottavuutta. Tässä mielessä siis vain tieteellisesti tuotettua tietoa voidaan käytännössä pitää nykypäivänä jollain yleisesti hyväksytyllä tavalla hyödyllisenä.

ön kannalta kiinnostavaa. Tällöin puhutaan niin sanotusta datan abstraktioinnista, joka tuo tutkimukseen välttämättä hyvinkin subjektiivisia elementtejä, kun tutkimusprosessiin tulee rikastamisen myötä ylimääräinen tulkinta-operationalisointi -kierros (Sund 2003).

Toinen tapa on pyrkiä käyttämään vain sellaisia osia datasta, jotka ovat käsitteellisiltä ominaisuuksiltaan mahdollisimman vakaita eli joiden käsittelyyn tarvitaan vain vähän erikoistunutta taustatietämystä (Sund ym. 2004).

Ihmisen biologisesta taustasta johtuen monet aikaan ja paikkaan kiinnitettävissä olevat asiat ymmärretään varsin samalla tavalla ja ovat siis sieltä helpoimmasta päästä tarvittavan lisätaustatietämyksen suhteen – tällaiset osat dataa voivat edustaa suorastaan faktoja siinä määrin kuin se ylipäättään on mahdollista ja joskus jo niistä voidaan saada irti tutkimusongelman ratkaisun kannalta kiinnostavaa ja uutta tietoa.

Aineiston ehdoilla eteneminen voi johtaa tulosten shoppailuun

Empiirisen tutkimuksen perusasetelma on usein selittämiseen ja mallintamiseen tähtäävä, mutta se ei ole ainoa vaihtoehto vaan voidaan edetä myös enemmän aineiston ehdoilla.

Aineistolähtöiseen analyysiin liittyvät omat sudenkuoppansa ja tulosten kalastelu tai shoppailu isosta aineistosta ei tietenkään ole tieteellisesti perusteltua.

Toisaalta aineistoa pitäisi osata ”kuunnella” tarpeeksi ja olla herkkä ymmärtämään, mitä se milloinkin kertoo. Tietyissä yhteyksissä ilmiön selittämisen ja ymmärtämisen sijaan voi olla aivan riittävää, että saadaan aikaiseksi mahdollisimman hyviä ennusteita johonkin suoraan havaittavaan asiaan. Tämä on itse asiassa hyvinkin perusteltu lähestymistapa tuottaa hyödyllistä informaatiota, vaikka silloin varsinaisen data-ana-

lyysin toimintamekanismit näyttäytyvät vain mustana laatikkona (Breiman 2000).

Jos näitä yleisiä periaatteita suhteutetaan isoon dataan, on selvää, että aineistoon liittyvät ongelmat ja mahdollisuudet eivät ole lainkaan uusia. Harva massadataksi katsottava aineisto on oikeastaan mitään erityistä isoa dataa, vaan se on yksinkertaisesti (sekundaarista) dataa, jota on tarkoitettu analysoida ja toivottavasti muuntaa jollain tavalla hyödylliseksi informaatioksi. Tässä mielessä ”vanhanaikaisetkin” datat voivat siis olla hyvinkin arvokkaita. Massadataan liitettäviä asioita on tehty kauan, vaikka onkin tyypillisesti käytetty jotain muuta ilmaisua kuin ”iso data” (Huber 1999).

Monia keskeisiä data-analyysin ongelmia on siis jo huolellisesti ja kriittisesti pohdittu tilastotieteilijöiden keskuudessa, mutta nämä ajatukset ovat jääneet yllättävänkin vähälle huomiolle, kun muiden alojen toimijat ovat pyrkineet ottamaan data-analyysin kenttää enenevässä määrin haltuunsa. Moiseen toimintaan liittyy usein pyörän uudestaan keksimistä ja myös pettymyksiä, kun samat perusongelmat välttämättä jossain vaiheessa näyttävät todelliset karvansa kaikista suurista uskomuksista huolimatta.

Isojen datojen parissa tiimityön merkitys kasvaa

Muutamii lähinnä aineiston kokoon liittyviin ongelmiin pystyttäneen melko helposti vastaamaan, kunhan tutkimuksen tekemisen kulttuuri ja eri alojen metodologiset paradigmat mukautuvat tilanteeseen tarkoituksenmukaisella tavalla.

Isojen datojen analysointi vaatii tavanomaista huolellisempaa suunnittelua ja vaihtoehtoisia algoritmeja.

Isojen datojen kanssa tiimityön merkitys kasvaa ja datan heijasteleminen asioihin liittyvästä taustatietämyksestä tulee entistä tärkeämpää – ehkä myös pienehkö hyvin analysoitu laadullinen tekstiaineisto voi osoittautua yllättävänkin hyödylliseksi isojen tekstimassojen analyysin tukena.

Toisaalta isoissa datoina kaikkia niissä välttämättä olevia ja piileviä virheitä ei mitenkään voi alkaa manuaalisesti tarkastamaan. On hyväksyttävä, että datassa voi olla virheitä, ja että virheettömätkin isot kokonaisaineistot sisältäisivät varmasti aikamoisia yllätyksiä, kun kaikki ilmiön oimuisimmatkin – mutta todelliset – ilmentymät voivat löytyä sieltä.

Isojen datojen analysointi vaatii tavanomaisia analyyseja huolellisempaa suunnittelua ja vaihtoehtoisia algoritmeja, kun laskennalliset seikat on oikeasti otettava huomioon. Kriittisen maalaisjärjen merkitystä ei voi liikaa korostaa järkevien lopputulosten saamiseksi.

Tilastotieteellisessä mielessä on varsin kiinnostavaa, että isot aineistot ovat usein kokonaisaineistoja eikä niihin siinä mielessä liity lainkaan otantavirhettä. Tämä muuttaa tilastollisen päätelyn luonnetta, kun epävarmuuden on oletettava syntyvän jostain muusta.

Isojen aineistojen kanssa myös käytännössä merkityksettömistä eroista tulee perinteisessä päätelyssä helposti tilastollisesti merkitseviä. Tilastollinen merkitsevyys ei siis välttämättä ole aina merkki hyödyllisestä informaatiosta.

Todennäköisyyslaskentaan perustuvat data-analyysin menetelmät ovat kuitenkin edelleen se ykkösvaihtoehto epävarmuuden hallintaan, vaikka niitä pitääkin pystyä käyttämään toisella tavalla kuin mihin tiukkarajaisimpien metodologisten paradigmojen rajat yltyvät. Joskus myös datan koon pienentäminen otannalla saattaa olla yllättävän toimiva ratkaisu.

Onnistuneet iso data -hankkeet pohjaavat useimmiten faktadataan

Nykypäivänä on aikaisempaan nähden erilaisia ja siinä mielessä uusia datan lähteitä. Historialliset dokumentit eivät sinänsä ole uutta dataa, mutta uutta on se, että iso osa nykypäiviin asti säilyneestä ihmisen tuottamasta kirjallisesta tai muusta symbolisesta materiaalista on digitoitu ja helposti saatavilla.

Erilaisilla sensoreilla voidaan tuottaa huimia määriä dataa myös asioista, jotka eivät ole normaalein aistein havaittavissa. Myös erilaiset sosiaaliset mediat, kuva- ja videoaineistot, kauppojen ja maksujärjestelmien rekisteröimät tiedot ja vaikkapa gps-paikallistamisella saatavat reaaliaikaiset koordinaatit sijainneista ovat muutamia esimerkkejä lukemattomista kiinnostavista sekundaarisista aineis-

tolähteistä, joita ehkä haluttaisiin analysoida.

Tällaisten massadata-aineistojen analysoinnista on verrattain helppo löytää hyviä esimerkkejä toimivista tavoista saada irti hyödyllistä informaatiota. Valitettavasti on nähty myös karkeita ylilyöntejä ja väärin perustein hehkutusta mahdollisuuksista.

Käytetyimpiä esimerkkejä ison datan mahdollisuuksista olivat esimerkiksi Googlen hakutuloksiin perustuvat ennusteet influenssa-aaltojen ilmaantumisesta. Tulokset olivat yksittäisen kerran yllättävän hyviä, mutta varsinaiseksi seurantajärjestelmäksi lähestymistavasta ei ole, koska löydökseen johtanut käyttäytymismalli ei ole pysyvä, vaan muuttuu kaiken aikaa – yksinkertaisillakin aikasarjamalleilla saadaan paljon parempia ennusteita (Lazer ym. 2014).

Hyödyllisimmät ja käytännössä onnistuneimmat iso data -hankkeet perustuvat yleensä dataan, joka heijastelee suoraan havaittavia asioita tai jota voidaan käyttää kuin se edustaisi totuutta. Silloinkaan datan käyttö ei ole helppoa, ja teoriataidon taustatietämystä voidaan tarvita hämmästyttävän paljon. Monimutkaisempia mittauksia sisältävissä tilanteissa empiirisen tutkimuksen rajat ja ilmiöiden epävakaisuus tulevat kuitenkin paljon aikaisemmin vastaan.

Sama data voi olla yhdessä yhteydessä suoraa faktadataa ja toisessa taas välillisiä heijastumia jostain muusta asiasta. Tällaisia piirteitä voidaan havaita myös esimerkiksi internetin keskustelupalstojen hyödyntämisessä tutkimuksessa: kielitieteilijälle teksti on kielen faktuaalista ilmentymistä keskustelupalstojen kontekstista, mutta yhteiskuntatieteilijä havaitsee samasta aineistosta välillisiä heijastumia jostain yhteiskunnallisista ilmiöistä.

Vaikka teknologian kehitys mahdollistaakin yhä isompien datamäärien analysoinnin, ei teknologia automaattisesti ratkaise analysointiin liittyviä keskeisiä tietoteoreettisia ongelmia.

Tällä kirjoituksellani olen halunnut peräänkuuluttaa riittävän kriittistä ja tieteellistä lähestymistapaa, joka yleensä johtaa parhaisiin lopputuloksiin myös isojen datojen analysoinnissa ja auttaa välttämään niihin hukkumista.

Toisaalta on myös hyväksyttävä, että liiallinen skeptisyys voi estää havaitsemasta sellaisia tilanteita, joissa iso data on tutkittavan ilmiön kannalta oikeasti hyödyllistä. Joka tapauksessa havaintojen jalostaminen hyödylliseksi informaatioksi on historian saatossa johtanut huikeisiin saavutuksiin. Tarvitaanko sellaiseen jatkossa ihmistä tai isoa dataa on sitten ihan eri kysymys. ■

Kirjoittaja on valtiotieteiden tohtori ja dosentti sekä työskentelee Yhteiskuntatieteiden menetelmäkeskuksessa Helsingin yliopistossa.

Lähteet:

Berg, Marc & Goorman, Els 1999. *The contextual nature of medical information. International Journal of Medical Informatics* 56.

Breiman, Leo 2000. *Statistical Modeling: The Two Cultures. Statistical Science* 16.

Dunn, Halbert L. 1946. *Record Linkage. American Journal of Public Health* 36.

Faber, Jan & Scheper, Willem J. 2003. *Social scientific explanations? On Quines's legacy and contextual fallacies. Quality and Quantity* 37.

Hand, David J. 2004. *Measurement Theory and Practice: The World Through Quantification. Lontoo: Hodder Arnold.*

Huber, Peter J. 1999. *Massive datasets workshop – four years after. Journal of Computational and Graphical Statistics* 8.

Härö, A. Sakari 1969. *The Centralized Statistics of Hospital Utilization in Finland. Methods of Information in Medicine. Supplement* 4.

Langefors, Börje 1993. *Essays on Infology: Summing Up and Planning for the Future. Göteborg: Department of Information Systems, University of Göteborg.*

Lazer, David & Kennedy, Ryan & King, Gary & Vespignani, Alessandro 2014. *The Parable of Google Flu: Traps in Big Data Analysis. Science* 343.

Sund, Reijo 2003. *Utilisation of administrative registers using scientific knowledge discovery. Intelligent Data Analysis* 7.

Sund, Reijo 2005. *Huono-osaisuus tiedollisena haasteena. Teoksessa: Sakari Hänninen & Jouko Karjalainen & Tuukka Lahti (toim.). Toinen tieto. Kirjoituksia huono-osaisuuden tunnistamisesta. Helsinki: Sosiaali- ja terveysalan tutkimus- ja kehittämiskeskus.*

Sund, Reijo & Nylander, Olli & Palonen, Tuula 2004. *Raa'asta rekisteriaineistosta terveystieteellisesti relevanttiin informaatioon. Yhteiskuntapolitiikka* 69.

Wikipedia 2015. *Big Data.*
http://fi.wikipedia.org/wiki/Big_data.



U. Östlund

JARI NIEMINEN

Tilastovirastot vahvoilla erilaisen tiedon yhdistämisessä

Hallinnollisten rekisteriaineistojen, haastattelututkimusten ja iso data -aineistojen yhdistäminen on yksi tapa uudistaa ja tehostaa tilastointia. Suomen ja muiden Pohjoismaiden kokemus rekisteriaineistoista tarjoaa kehitystyössä etulyöntiaseman verrattuna maihin, joissa ollaan riippuvaisia kyselyaineistoista.

so data on ollut viime vuosina esillä tilastontekijöiden keskusteluissa perinteistä tiedonkeruumenetelmistä poikkeavana uudenlaisena tietolähteenä, jonka mahdollisuuksia ei ole osattu tai haluttu lähteä hyödyntämään virallisen tilaston laatimisessa. Rekisterejä laajasti hyödynnettävälle tilastontekijälle nousee mieleen väistämättä kysymys, eivätkö laajat ja jatkuvasti päivittyvät rekisterit ole iso dataa?

Suomessa ja muissa Pohjoismaissa on jo kauan käytetty rekisteriaineistoja tilastojen tietolähteenä. Sen sijaan muulla maailmassa, etenkin Yhdysvalloissa ja Isossa-Britanniassa myös rekisteriaineistot ovat uusia tilastoaineistoja, joiden mahdollisuuksia virallisen tilaston tuottamisessa on lähdetty tutkimaan vasta viime vuosina.

Yhtenä syynä suureen innostukseen ison datan hyödyntämiseksi on oletettavasti juuri rekisteri-

aineistojen puute etenkin anglosaksisessa maailmassa. Niin Iso-Britannia, USA, Kanada, Australia kuin Uusi-Seelantikin ovat kaikki maita, joissa ei ole ollut poliittisesti mahdollista virallisten koko väestöä kattavien rekisterien kuten väestörekisterin perustaminen.

Tilastoinnissa otostutkimuksia ja kyselyjä on perinteisesti totuttu pitämään ainoina oikeina tapoina tuottaa laadukasta tilastotietoa. Kaikki muut tavat on määritelty epäluotettaviksi ja menetelmiltään kyseenalaisiksi. Tilastokeskuksella on monivuotinen kokemus kansainvälisestä yhteistyöstä, jossa rekisterien käyttöä tilastointiin on pitänyt yhä uudelleen perustella. On pitänyt todistaa, että rekistereihin perustuvat tilastotiedot täyttävät kansainväliset laatukriteerit.

Kansallisissa tilastoissa on jo pitkään ollut itsestään selvää, että joidenkin tietojen osalta re-

kisteritiedot ovat vähintään yhtä laadukkaita kun kyselyillä kerätyt vastaavat tiedot. Tällaisia ovat esimerkiksi verottajan verotettavia tuloja kuvaavat tiedot ja väestörekisterin demografiset tiedot.

Ison datan käyttöön tilastoinnissa liittyy samoja ennakkoluuloja kuin aikoinaan rekisteriaineistoonkin. Rekisterien käyttöönoton seurauksena tilastojen tietosisällöt laajenivat merkittävästi, ja tietoja voidaan julkaista aiempaa useammin. Samalla myös tietojen käyttö on lisääntynyt. Nykyiset tiedon käyttäjät eivät varmasti hyväksyisi, että esimerkiksi kunntasaisia tietoja työssäkäynnistä ja asumisesta tuotettaisiin vain viiden tai kymmenen vuoden välein.

Iso data mahdollistaa usein lähes ajantasaisen tiedon seurannan, kunhan ratkaistaan edustavuuden ja jatkuvasti muuttuvien teknologioiden haasteet. Esimerkiksi osa nopeista suhdan- netilastoista voitaisiin tuottaa yhdistelemällä rekisteritietoa, isoa dataa, suoraa tiedonkeruuta ja kaikkea muuta tietoa, jota jo aikaisemmin ker- tyneen datan analysointi voi tarjota luotettavuuden parantamiseen. Näitä välineitä ei juurikaan ole hallussa muilla organisaatioilla kuin tilastovirastoilla. Tästä syystä myös tilastovirastojen pitää nähdä roolinsa ison datan käytön edistämässä.

Rekisteritieto kerätään, iso data syntyy

Rekisteriaineistoissa ja isossa datassa on yhteisiä ominaisuuksia ja eroja. Rekisteriaineistot ovat yleensä suurivolyymisiä ja jatkuvasti päivittyviä aineistoja. Tässä suhteessa rekisterit ovat ainakin nykyisin hyvin lähellä transaktiodataa. Myös rekistereissä päivytystapahtumat koostuvat useista pienistä tapahtumista, joista osa voi olla vain tietojärjestelmän sisäiseen ylläpitoon liittyvää loki- tietoa. Ison datan ja rekisterin erona puolestaan on se, että iso data koostuu juuri näistä tapahtu- mista, joiden lopputulos on esimerkiksi rekisteriin tehtävä ylläpito. Siinä mielessä myös rekisterin ylläpidosta syntyy koko ajan uutta isoa dataa, jota ei ole ehkä toistaiseksi osattu hyödyntää riittävästi.

Erottavana piirteenä ison datan ja rekisteridatan välillä on mm. se, että rekisterit ovat tyypillisesti jul- kisen sektorin perustamia ja ylläpitämiä tietojärjes- telmiä, kun iso data on usein yksityistä tietoa. Re- kisterit ovat yleensä syntyneet hallinnon vaatimuk- sesta, ja nykyiset sähköiset rekisterit on useimmiten perustettu siirtämällä kortistoja tai muita alun pe- rin kirjallisia aineistoja sähköiseen muotoon. Yleen- sä tässä vaiheessa on tehty suuri järjestelmäuudistus, jonka päämääränä on ollut nimenomaan kerätä ja ylläpitää tärkeiksi katsottuja tietoja ja yleensä vielä kattavasti koko rekisterin kohdejoukosta.

Iso data on puolestaan kertynyt jonkin toimin- nan sivutuotteena, ja tarkoituksena ei ole ollut tiedon keruu kyseisestä ilmiöstä, vaan ylipäätään

järjestelmän toiminnan ylläpito. Esimerkiksi sosi- aalisen median päivitykset tehdään aivan muista lähtökohdista käsin kuin tietojen saamiseksi vi- rallisiin tietokantoihin. Tämä on ison datan kes- keinen heikkous, mutta samalla sen suurin vah- vuus. Iso data kertyy henkilöiden omasta aloit- teesta ilman varsinaista keruuoperaatiota. Datan kertyminen on usein vain sivutuotteena itse tie- don välitykselle.

Tilastontekijän kannalta datan kertyminen il- man varsinaista tiedonkeruuta vaikuttaa ihanne- tilanteelta. Tiedonkeruu on helppoa ja tapahtuu kaiken aikaa ja vielä siten, että kohde ei edes tiedä olevansa tilastoinnin kohteena eikä koe aikaansa tuhlattavan. Tähän seikkaan liittyy kuitenkin tietosuojaongelmia ja eettisiä kysymyksiä. Tilaston- tekijät ovat pitäneet erittäin tärkeänä, että tutki- muksessa mukana oleva tietää, miksi kysymys teh- dään ja mihin hänen antamaansa tietoa käytetään.

On tärkeää kiinnittää huomiota ison datan laa- tuun, tietosuojaan ja edustavuuteen. Esimerkik- si sosiaalisen median päivityksistä kerättävä tie- to saadaan vain niiltä, jotka ovat kyseisen median käyttäjiä. Ainakin toistaiseksi osa väestöstä jää ko- konaan tällaisen tiedonkeruun ulkopuolelle.

Erilaisen tiedon yhteiskäyttö on Suomen vahvuus

Suomessa on muiden Pohjoismaiden tavoin ra- kennettu tehokas hallinnollisia rekistereitä laa- jasti hyödyntävä tilastojärjestelmä, jota muualta maailmasta tullaan tänne opiskelemaan ja mah- dollisuuksien mukaan myös jäljittelemään. Se, et- tä rekisterien käyttö on ollut aikanaan tärkeä ja kansainvälisesti suuria epäilyksiä herättävä tie- donkeruutapa, antaa meille kuitenkin jonkinlais- ta pohjaa uusien tiedonkeruutapojen ennako- luulottomalle käyttöönotolle.

Rekisterien ja ison datan käytössä on paljon yhtäläisiä piirteitä. Etenkin aineistojen tulkintaan, niiden yhdistelyyn sekä otosaineiston täydentä- miseen liittyy paljon menetelmällistä kehittämis- työtä, jota voimme hyödyntää varmasti nopeam- min kuin perinteiset, vain suoraa tiedonkeruuta käyttävät tilastovirastot muualla maailmassa.

Onneksi meillä Suomessa on sekä isoa dataa että rekisteriaineistoja. Tulevaisuudessa meidän on syytä edelleen ennakkoluulottomasti jatkaa ti- lastotoimessa näiden hyödyntämistä yhdessä suo- ran tiedonkeruun kanssa. Tilastoaineistot tulevat yhä enemmän koostumaan kaikkien eri tietoläh- teiden yhdistämisestä, ja jopa yksittäisten muut- tujen tiedot voivat olla peräisin monista eri tieto- lähteistä. Tietojen laatu voidaan varmistaa myös menetelmällisesti ja dokumentoida tietojen käyt- täjälle ymmärrettävässä muodossa. ▶

Voisiko iso data olla ratkaisu vastauskadon kanssa kamppaileville haastattelututkimuksille? Yhä suurempi ongelma haastatteluissa on, että kohteita ei enää tavoiteta puhelimella. Vaikka lähes kaikilla on mobiililaitte, osa meistä ei vastaa tuntemattoman soittajan puheluihin. Nykyisin haastattelutietoja täydennetään rekisteriaineistoista. Ison datan osalta suuri puute on se, että sitä ei voida suoraan yhdistää olemassa olevaan tietoaaineistoon. Esimerkiksi henkilön tietoja ei ole yhdistettävissä henkilötunnuksella.

Survey-tiedonkeruussa on lähdeyhtiä siitä oleuksesta, että otoksen tulee olla mahdollisimman pieni, koska haastattelu on kallista; ison datan kohdalla on mahdollista analysoida suurempaa populaatiota kuin perinteinen otos. Jos iso data sisältää tietoa, jota haastatteluaineistossa on tarkoitus kerätä, voidaan haastattelutieto korvata tai sitä täydentää ison datan käytöllä. Tämä vaatii paljon menetelmätyötä etenkin siksi, että tällaisissa muutoksissa menetetään usein vertailu aikaisempiin tutkimuksiin tai joudutaan jopa tinkimään kansainvälisten sopimusten velvoitteista.

Rekisteriaineiston avulla voidaan tutkia ison datan edustavuutta ja jakaamaa. Vinon jakauman korjaamiseen on olemassa tilastollisia menetelmiä kuten uudelleenpainotus, mutta ison datan kohdalla saattaa olla vaikeuksia havaita, miltä osin aineisto on vino, tai mikä osajoukko puuttuu kokonaan.

Aineiston vinouden tutkimiseen ja menetelmien kehittämiseen on Suomessa erinomaiset mahdollisuudet hyvien rekisteriaineistojemme vuoksi. Kun esimerkiksi Isossa-Britanniassa tehtiin arvioita maassa asuvien ulkomaalaisten määristä eri kielisten internethakujen avulla, havaittiin mm. suuri lisäys puolankielisten hakujen määrässä. Hakujen määrää kuvaava käyrä seurasi melko tarkasti puolaalaisten maahanmuuttajien kasvua. Ongelmana oli kuitenkin, että tietoa ei voitu suhteuttaa väestötilastojen aineistoon, koska edelliset tiedot olivat vuoden 2000 väestölaskennasta ja siten vanhentuneet.

Hallinnolliset esteet ison datan käytölle

Iso data on nostettu esille keskeisenä kilpailutekijänä Euroopan unionin piirissä (Euroopan komissio 2014). Tilastotoimen osalta voi todeta, että Euroopan unionin nykyinen säädöspohja tilastojen tietosisältöjen toimitusten ja joskus jopa yksittäisten kysymysten muotoilun osalta hidastaa tai jopa estää ison datan hyödyntämisen virallisessa tilastoissa.

Tilanne on samankaltainen kuin joskus aiemmin oli rekisteriaineistojen käytön kohdalla, vaikka Euroopan unionin sosiaalitalustostrategia pyrkii uudelleenlaiseen eri tilastojärjestelmiä yhdistävään lähestymistapaan. Nähtäväksi jää onko kehitys tarpeeksi nopeaa, jotta virallinen EU-tilasto säilyttää

kilpailukykyä ja ennen kaikkea uskottavuutensa kansalaisten silmissä, kun muut toimijat julkaisevat samoja tietoja julkista tilastointia huomattavasti nopeammin.

Jos ison datan ilmiötä tarkastellaan kansalaisen kannalta, herää kysymys, onko isosta datasta puhuminen keino välttää tosiasiaa, että ison datan kertyminen on tietojen rekisteröintiä enemmän tai vähemmän epävirallisiin tietokantoihin. Esimerkiksi liikenteen valvontakameroiden ja GPS-paikannuksen käyttö liikennevalvonnan tai vaikkapa kehittäillä olevat tiemaksujärjestelmät luovat uuden täysin rekisteriin verrattavan ja virallisen tietoaaineiston, jonka käyttö ei sinällään poikke mitenkään esimerkiksi verotuksen rekistereistä.

Avointa dataa tulee lisätä

Isosta datasta puhutaan tällä hetkellä usein avoimen datan yhteydessä; ikään kuin kaikki iso data olisi avointa ja avoin data isoa dataa. Tällaista johdopäätöstä ei pidä kuitenkaan tehdä automaattisesti. Luultavasti jatkossa suurin osa isosta datasta, jota maailmanlaajuiset tietojärjestelmät tallentavat ja käsittelevät kaiken aikaa, ei tule suuren yleisön tietoon saati käyttöön avoimena datana.

Tilastoviranomaisten olisi syytä varmistaa pääsy myös niihin yhteiskunnan kannalta merkittäviin ison datan tietolähteisiin, jotka eivät ole avointa dataa. Tilastotoimen tehtäväksi tulee mieltää näiden datojen jatkojalostus, jotta niistä voidaan saada käyttäjälle hyödyllistä avointa dataa.

Ison datan ja rekistereissä olevan tiedon välillä on tilastoinnin kannalta sekä eroja että yhtäläisyyksiä. Olennaista on, että tilastotoimessa tulee seurata yhteiskunnan kehitystä niin tilastojen tietosisällön kuin niiden lähdeaineistojenkin osalta. Muuten päädytään tilanteeseen, jossa tehdään perinteistä tilastoa perinteisin menetelmin ilmiöistä, joita yhteiskunnassa ei enää esiinny, ja uudet tärkeätkin ilmiöt jäävät tilastoinnin ulkopuolelle, koska niistä ei enää saada laadukasta tietoa perinteisin tiedonkeruumenetelmin. ■

Kirjoittaja on kehittämisspällikkö Tilastokeskuksen väestö- ja elinoloilastot -yksikössä.

Lähde:

Euroopan komissio 2014. Komission tiedonanto Euroopan parlamentille, neuvostolle, Euroopan talous- ja sosiaalikomitealle ja alueiden komitealle. Kohti menestyvää datavetoista taloutta. COM(2014)442. final. Bryssel.

Reittiopas taipuu myös työmatkalaskelmiin

Pääkaupunkiseudulla tehdään päivittäin yli miljoona työmatkaa. HSL:n Reittioppaan ja Tilastokeskuksen työssäkäyntitilaston avulla on mahdollista laskea työmatkojen kesto julkisilla liikennevälineillä. Tyypillinen työmatka kestää julkisilla liikennevälineillä 34 minuuttia.

Tilastokeskuksessa on tehty laskelmia työmatkojen pituuksista sekä niihin kuluva ajasta henkilöautolla ja pyörällä (Huhta & Pyykkönen 2013; Piela 2014; ks. myös Pasi Pielan kirjoitus tämän numeron lopussa). Matkojen laskemisessa on hyödynnetty Liikenneviraston Digiroad-aineistoja sekä Tilastokeskuksen työssäkäyntitilaston aineistoa, joka sisältää lähes kaikkien työllisten koti- ja työpaikan sijaintitiedot.

Pyöräilyn, kävelyn ja henkilöauton ohella työmatka taittuu usein julkisilla liikennevälineillä. Helsingin seudun liikkumistottumuksia koskevan tutkimuksen mukaan joukkoliikenteen osuus oli 44 prosenttia kaikista pääkaupunkiseudun sisäisistä matkoista (ei siis pelkästään työmatkoista) vuonna 2012 (HSL 2013, 49).

Vaikka julkisen liikenteen osuus matkoista on pääkaupunkiseudulla melko suuri, ei matkoista tai niihin kuluva ajasta ole juurikaan tietoa.

Pääkaupunkiseudulla on yli puoli miljoonaa työllistä. He tekevät päivittäin yli miljoona yhdensuuntaista työmatkaa.

Tätä artikkelia varten laskin pääkaupunkiseudun työllisille matkan ja matka-ajan julkisilla liikennevälineillä hyödyntämällä Helsingin Seudun Liikenne-kuntayhtymän (HSL) Reittiopas-palvelua ja Tilastokeskuksen työssäkäyntitilastoa. Työssäkäyntitilastosta ei käy ilmi, millä kulkuvälineellä kukin matkansa kulkee. Asuin- ja työpaikkatie-

tojen pohjalta voidaan kuitenkin arvioida miten kauan matkaan menisi, jos kukin työllinen kulki julkisilla.

Käytännössä laskenta tapahtui hyödyntämällä Reittiopas-palvelua. Tarkoituksena oli selvittää, miten Reittiopas-palvelun avoin rajapintaratkaisu soveltuu tilastointiin. Samalla syntyi lisätietoa pääkaupunkiseudun työmatkoista.

Esitän ensiksi millä oletuksilla ja miten matkat on laskettu. Toiseksi esitän alustavia tunnuslukuja työmatkoista ja ajoista pääkaupunkiseudulla sekä pohdin niiden valossa, missä julkinen liikenne palvelee työmatkaajia tehokkaimmin. Lopuksi puntaroin avointen rajapintojen hyödyntämistä tilastoinnissa sekä mahdollisuutta laskea työmatkat ja -ajat koko maan tasolla.

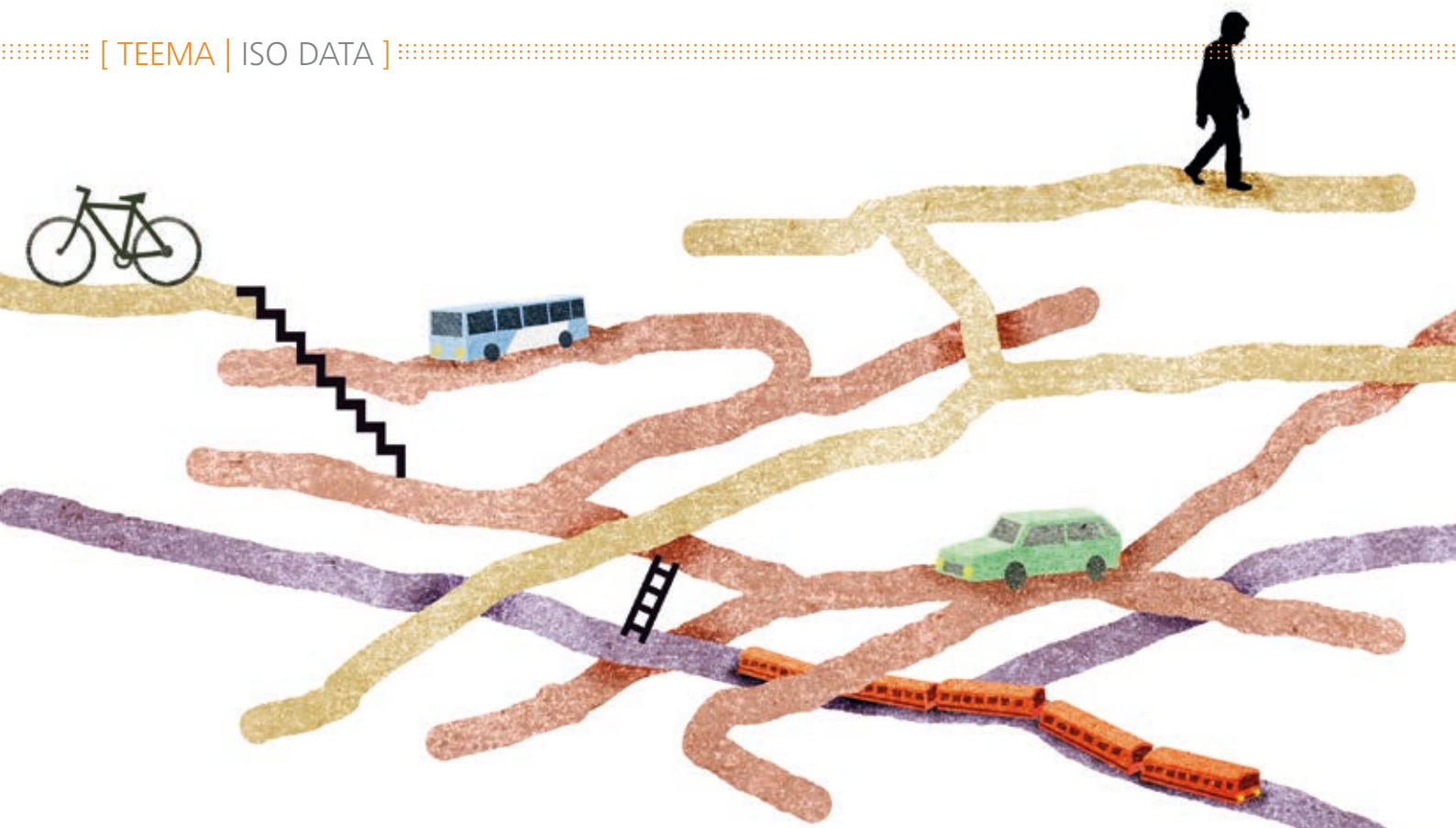
Töissä kello yhdeksän

Reittioppaan avoimen rajapinnan kautta on saatavilla kaikki aikataulu- ja pysäkkitiedot. Rajapinnan kautta on myös mahdollista käyttää Reittioppaan palveluita. Yksittäisen työmatkan pituus ja kesto voidaan selvittää lähettämällä rajapintaan osoiteparin (kotiosoite-työpaikan osoite) sisältävä kysely, jonka seurauksena rajapinnasta saadaan reitti osoitteiden välille.

Kaikkien pääkaupunkiseudun työllisten työmatkojen laskeminen edellyttäisi yli puolen mil- ▶



U. Östlund



joonan kyselyn tekemistä rajapintaan. Ajan ja palvelinkapasiteetin säästämiseksi tein osoitetun otannan pääkaupunkiseudun työllisistä osa-alueittain ja laskin työmatkan ainoastaan otantaan osuneille työllisille.

Otannan koko oli neljäsosa (114 000 henkeä) niistä pääkaupunkiseudun työllisistä, joille voitiin laskea työmatka. Noin 10 000 työlliselle ei saatu laskettua matkaa, koska Reittiopas ei löytänyt joko asuin- tai työpaikan osoitetta.

Suuresta otoskoosta johtuen tuloksiin ei sisälly juurikaan sattumanvaraisuutta. Tästä huolimatta on muistettava, että kaikki esitetyt tunnusluvut ovat vain estimaatteja. Todelliset työmatkat ja työmatka-ajat saattavat olla hieman suurempia tai pienempiä.

Aiempien työmatkalaskentojen pohjalta kuitenkin tiedetään, ettei otantavirhe ole erityisen suuri. Esimerkiksi helsinkiläisten työ- ja asuinpaikan lineaarisen etäisyyden keskiarvojen erotus totaali- ja otosaineistojen välillä on kymmenen metriä.

HSL:n tutkimuksessa pääkaupunkiseudulla julkisilla liikennevälineillä kulkevien työmatka-ajan keskiarvo oli 32 minuuttia, kun otoksen keskiarvo oli 34,8 minuuttia. Mittakaava on siis oikeasuuntainen.

Matkat on laskettu sillä oletuksella, että kukin työmatkaaja saapuisi työpaikalleen kello yhdeksäksi. Oletus perustuu HSL:n tutkimukseen, jonka mukaan suurin osa työmatkoista tapahtuu kello seitsemän ja yhdeksän välillä.

Laskenta on tehty Reittioppaan oletusarvoilla (esim. kävelynopeus on oletusarvon mukainen).

Aineistoon pätevät samat rajaukset kuin edellä mainittuihin työmatkalaskelmiin; koska tiedot on laadittu rekistereiden perusteella, ei aineistosta käy ilmi erilaiset työjärjestelyt, kuten etätöskentely. Otokseen on siis kelpuutettu kuka tahansa henkilö työpaikasta tai työsuhteesta riippumatta.

Laskelma on siinä mielessä teoreettinen, että julkisen liikenteen aikataulutiedot ovat tämän vuoden maaliskuun ensimmäiseltä maanantailta, mutta asuin- ja työpaikat ovat vuodenvaihteesta 2012/2013. Näin ollen luvut kuvastavat matkoja, jos nykyiset aikataulut olisivat olleet voimassa joulukuussa 2012.

Teoreettisuutta ei ole kuitenkaan syytä korostaa liikaa; aikataulut tai väestön ja työpaikkojen sijainti eivät muutu merkittävästi muutamassa vuodessa.

Pyörä on nopeampi kuin bussi

Pääkaupunkiseudun kunnissa tyypillinen työmatka teitä pitkin oli vuonna 2012 lähes kymmenen kilometriä (työmatkalaskelmista ks. esim. Piela 2014). Julkisilla liikennevälineillä työmatkat olivat noin puoli kilometriä pidempiä (mediaanipituus 10,3 km).

Kenties hieman yllättäen pyöräily on osoittautunut julkista liikennettä nopeammaksi tavaksi taittaa työmatka koko pääkaupunkiseudulla. Työmatka-ajan mediaani julkisilla kulkien oli 34 minuuttia, kun pyörällä se oli 30 minuuttia (ks. Pasi Pielan kirjoitus tämän numeron lopussa).

Koska pääkaupunkiseudun työpaikat ovat keskittyneet voimakkaasti, kannattaa työmatkojen

TAULUKKO 1. Työmatkan ja -ajan mediaani julkisilla liikennevälineillä yhteen suuntaan pääkaupunkiseudulla asuin- ja työkunnittain 2012

		TYÖKUNTA				
	AIKA, MIN.	Espoo	Helsinki	Vantaa	Kauniainen	Yhteensä
ASUINKUNTA	Espoo	27	46	65	33	38
	Helsinki	43	28	45	52	31
	Vantaa	61	46	27	63	40
	Kauniainen	31	45	64	10	40
	Yhteensä	34	34	37	29	34

		TYÖKUNTA				
	PITUUS, KM	Espoo	Helsinki	Vantaa	Kauniainen	Yhteensä
ASUINKUNTA	Espoo	8	17	27	7	13
	Helsinki	15	7	16	18	8
	Vantaa	25	18	7	21	14
	Kauniainen	8	17	26	1	14
	Yhteensä	11	10	13	6	10

Lähteet: Tilastokeskus, työssäkäyntitilasto; HSL, Reittiopas

kestoja tarkastella vähintään kuntatasolla ja mieluiten sitäkin tarkemmalla aluejaolla (taulukko 1).

Kuten arvata saattaa, julkisilla liikennevälineillä kulkevat espoolaiset ja vantaalaiset käyttävät työmatkoihinsa helsinkiläisiä enemmän aikaa. Vantaalaisen tyypillinen työmatka kestää yhdeksän minuuttia helsinkiläistä kauemmin.

Erot korostuvat entisestään, kun tarkastellaan pelkästään Helsinkiin kohdistuvia julkisilla kuljettuja työmatkoja.

Helsingissä työssäkäyvät espoolaiset ja vantaalaiset käyttävät työmatkoihinsa tyypillisesti noin 12 minuuttia enemmän aikaa kuin muut pääkaupunkiseutulaiset. Päivittäin matkoihin kuluu siis noin 24 minuuttia keskivertopääkaupunkiseutulaista enemmän aikaa.

Vantaalla tai Espoossa asuminen ei kuitenkaan välttämättä tarkoita ajallisesti pitkää työmatkaa, sillä kuntien sisäisissä työmatkoissa espoolaiset ja vantaalaiset eivät juurikaan poikkea helsinkiläisistä.

Kaikkia pääkaupunkiseudun kunnissa sisäiset työmatkat taittuisivat julkisilla liikennevälineillä keskimäärin puolessa tunnissa. Sekä Vantaalla että Espoossa hieman alle puolet työmatkoista kohdistui kotikuntaan, kun Helsingissä neljä viidestä työmatkasta kohdistui Helsinkiin.

Vantaalaisten ja espoolaisten kannalta on lohdullista, että vaikka työmatka on tyypillisesti pidempi kuin helsinkiläisillä, on matkojen keskinopeus suurempi kuin helsinkiläisillä.

Kaikkien pääkaupunkiseutulaiden työmatkojen mediaaninopeus julkisilla liikennevälineillä oli 18 kilometriä tunnissa, mutta vantaalaisten työmatka

Minkä espoolaiset ja vantaalaiset häviävät etäisyydessä, he voittavat vauhdissa.

taittui hiukan muuta pääkaupunkiseutua nopeammin (21 km/h) (ks. taulukko 2). Osan siitä, minkä espoolaiset ja vantaalaiset häviävät etäisyydessä, he voittavat vauhdissa.

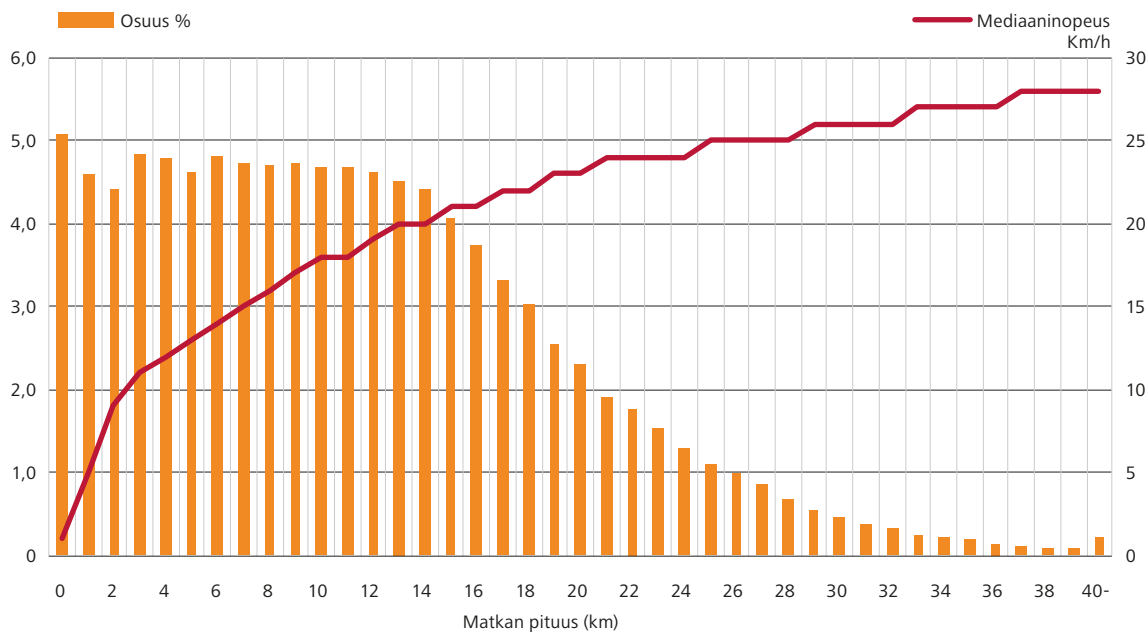
Keskinopeudet vaihtelevat myös työmatkan kohteen mukaan. Helsingissä työssäkäyvä espoolainen taittaa matkansa tyypillisesti noin 22 km/h vauhdilla, kun taas Vantaalla työskentelevä espoolainen kulkee hieman nopeammin, noin 25 km/h vauhdilla.

Hitaimmat matkat taittavat kuntien sisällä: esimerkiksi Espoossa työssäkäyvä espoolainen kulkee julkisilla 17 kilometriä tunnissa ja helsinkiläinen kotikunnassaan työskentelevä 14 kilometriä tunnissa. Tätä selittää se, että nopeudet kasvavat matkan pituuden kasvaessa.

Työmatkan pituuden kasvaessa myös matkan keskinopeus kasvaa

Työmatkojen pituuksia koskeville jakaumille on tyypillistä, että ne ovat vinoja (vrt. kuvio 1a; jakaumista ks. esim. Huhta & Pyykkönen 2013). Tästä syystä suosin mediaania enkä keskiarvoa ku- ➤

KUVIO 1A. Työmatkat ja työmatkan mediaaninopeus matkan pituuden mukaan julkisilla liikennevälineillä 2012



vatessani työmatkojen pituuksia. Kuviossa 1a erotettava työmatkojen pituuden vino jakauma johtuu siitä, että alle kilometrin matkoja on verrattain paljon, lähes kymmenesosa kaikista havainnoista.

Matka-ajat ovat sen sijaan jakautuneet normaalisti (vrt. kuvio 1b).

Kuviosta 1a käy myös ilmi miten työmatkan pituuden kasvaessa työmatkan keskinopeus kasvaa. Alle 10 kilometrin matkoilla keskinopeus jää alle 20 kilometriin tunnissa. Suhdetta selittänee se, että pitkät matkat taitetaan usein suurilla ajoväylillä, junalla tai metrolla. Vaihtoihin kuluvan ajan osuus koko matkasta on verraten pieni.

Toinen selitys on, että lyhyitä matkoja tehdään etenkin tiheästi asutuilla alueilla, joissa ruuhkat ovat yleisiä, pysäkkiväli tiheä ja kulkuvälineet verraten hitaita (esimerkiksi raitiovaunu tai bussi).

Nopeuden hajonnassa on kaksi piikkiä, joita selittää lyhyiden matkojen osuus ja Reittioppaan oletuskävelyvauhti (70 metriä minuutissa eli 4,2 kilometriä tunnissa) (vrt. kuvio 1c).

Oletus näkyy nopeuden hajonnassa piikkinä neljän kilometrin kohdalla. Toisen piikin muodostavat nollamatkat eli kotiosoitteessaan tai sen välittömässä läheisyydessä työskentelevät henkilöt.

Missä julkinen liikenne palvelee tehokkaimmin?

Koska kuntarajojen sisään mahtuu monenlaisia alueita, työmatkoja kannattaa tarkastella esimerkiksi osa-alueittain tai postinumeroalueittain.

Seuraavassa keskityn osa-alueisiin, tarkemmin sanottuna suuralueisiin.

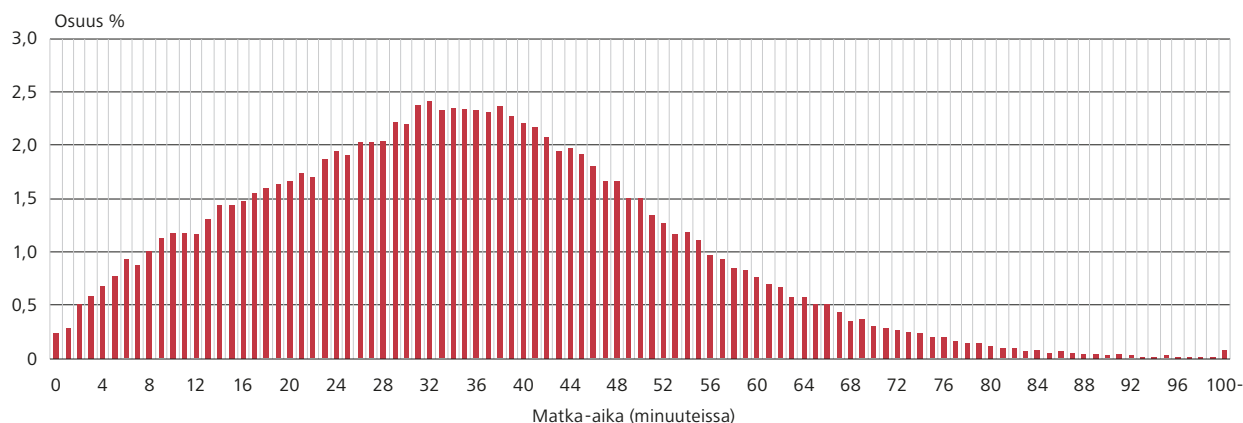
Suuraluetasolla tulos ei sinänsä ole yllättävä: työmatkat kasvavat siirryttäessä pois päin kaupungin keskustasta. Helsingin eteläisessä suurpiirissä puolet työmatkoista kesti alle 22 minuuttia julkisilla liikennevälineillä, kun itäisessä suurpiirissä joka toinen työmatkaaja käytti matkoihinsa yli 37 minuuttia. Koska pääkaupunkiseudun työpaikat ovat keskittyneet voimakkaasti kaupungin keskustaan (eteläinen suurpiiri), ei tulos yllätä.

Julkisen liikenteen kannalta mielenkiintoisempaa on, miten paljon ja kuinka nopeasti työmatkaan kulunut aika kasvaa etäisyyden kasvaessa. Oletettavasti alueilla, joissa työpaikan saavutettavuus julkisilla on muita alueita heikompi, yksi lisäkilometri työmatkaan kasvattaisi matkaan kuluvaa aikaa enemmän kuin paremman saavutettavuuden alueilla.

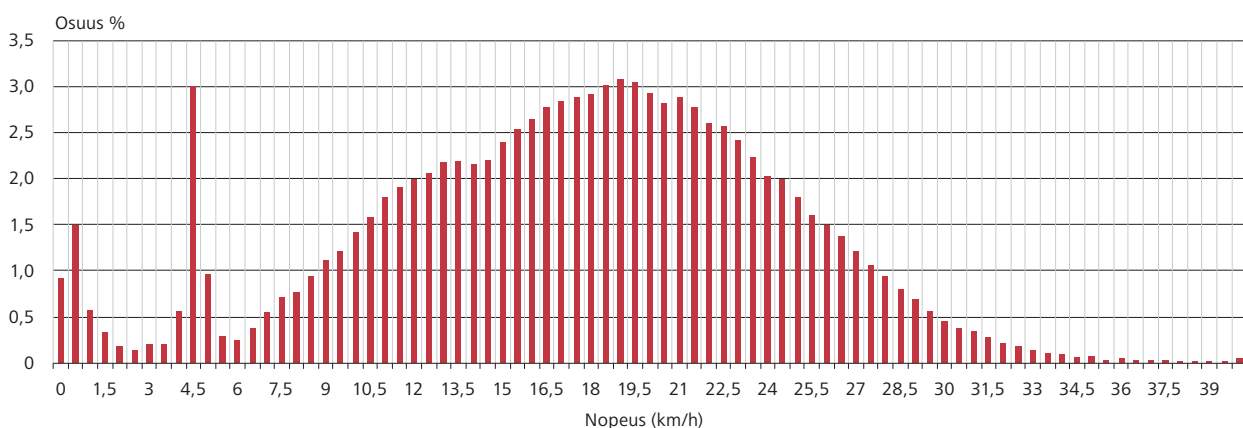
Oletusta voi testata mallintamalla ajan ja matkan välistä riippuvuutta regressiomallilla. Mallinnuksen tuloksena saadaan regressiosuoran kulmakerroin, joka tässä tapauksessa kertoo, paljonko yksi kilometri etäisyydessä kasvattaa matka-aikaa. Toisin sanoen: mikäli alueiden välillä on eroja saavutettavuudessa, tulisi kulmakerrointen poiketa toisistaan.

Yksinkertaistaen voi sanoa, että mitä suurempi kulmakerroin, sitä heikommin julkinen liikenne vastaa asukkaiden työmatkoihin liittyviin tarpeisiin. On tärkeää muistaa, että tunnusluku ei kuvaa, kuinka nopeasti alueelta pääsee esimerkiksi Helsingin keskustaan, vaan luku kuvaa alueen

KUVIO 1B. Työmatkat matka-ajan mukaan julkisilla liikennevälineillä pääkaupunkiseudulla 2012



KUVIO 1C. Työmatkat mediaaninopeuden mukaan julkisilla liikennevälineillä pääkaupunkiseudulla 2012



Kuvioiden lähteet: Tilastokeskus, työssäkäyntitilasto; HSL, Reittiopas

asukkaiden työmatkojen ja niihin kuluvan ajan suhdetta. Toisin sanoen myös työmatkojen kohde vaikuttaa saavutettavuuteen.

Taulukkoon 2 on koottu osa-alueittain tunnuslukuja työmatkoista ja matka-ajoista sekä kulmakerroin regressiomallille, joka tutkii matka-ajan ja kodin ja työpaikan etäisyyden välistä riippuvuutta. Kulmakerroin kertoo, monellako minuutilla matka-aika kasvaa, kun tietä pitkin kuljettava matka työpaikalle kasvaa kilometrin.

Voi tuntua oudolta puhua minuuteista desimaalin tarkkuudella. Erot alueiden välillä mitataan kuitenkin kymmenyksissä, ei kokonaisluvuissa.

Kulmakertoimen tulkintaa ja alueellisia eroja havainnollistaa kulmakertoimen kertominen luvulla kymmenen. Tulos kertoo kuinka kauan alueen asukkailla kestää kymmenen kilometrin työmatkassa.

Koska pääkaupunkiseudulla yksi lisäkilometri työmatkaan tarkoittaa kahta lisäminuuttia matka-

aikaan, kuluu kymmenen kilometrin matkaan pääkaupunkiseudulla mallin mukaan 20 minuuttia.

Alueiden välillä ei ole juurikaan eroja saavutettavuudessa. Espoolaisen matka-aika kasvaa lisäkilometrin myötä kahdella minuutilla, helsinkiläisen 2,1 minuutilla ja vantaalaisen 1,9 minuutilla.

Jos mallin pohjalta pitäisi veikata, voittaako espoolainen, vantaalainen vai helsinkiläinen kymmenen kilometrin työmatkalla, voittaisi vantaalainen minuutilla. Pienet erot selittynevät osaksi sillä, että nopeus kasvaa etäisyyden myötä: pitkällä matkalla nopeus paikkaa mahdollisia puutteita saavutettavuudessa.

Kuntien sisäiset erot saavutettavuudessa ovat suurempia kuin kuntien väliset erot, mutta vain hieman. Esimerkiksi Helsingissä matkustusai-ka kasvaa eniten etäisyyden kasvaessa Helsingin pohjoisessa suurpiirissä, jossa sijaitsevat mm. Maunula, Pakila ja Paloheinä.



Taulukko 2. Julkisen liikenteen työmatkojen ajan, pituuden ja nopeuden mediaanit ja regressiomallin kulmakertoimet osaluueittain 2012

ALUE	AIKA, MIN	PITUUS, KM	NOPEUS, KM/H	KULMAKERROIN
Pääkaupunkiseutu yhteensä	34	10	18	2,0
Espoo yhteensä	38	13	20	2,0
Suur-Leppävaara	35	11	18	2,2
Suur-Tapiola	33	10	18	2,2
Suur-Matinkylä	36	13	20	2,2
Suur-Espoonlahti	40	17	24	2,0
Suur-Kauklahti	51	19	22	1,7
Vanha Espoo	43	16	21	2,0
Pohjois-Espoo	55	19	20	2,4
Helsinki yhteensä	31	12	16	2,1
Eteläinen suurpiiri	24	4	11	2,3
Läntinen suurpiiri	30	7	15	2,2
Keskinen suurpiiri	24	5	13	2,1
Pohjoinen suurpiiri	33	9	16	2,4
Koillinen suurpiiri	36	11	18	2,1
Kaakkoinen suurpiiri	35	10	17	2,3
Itäinen suurpiiri	39	14	21	2,1
Östersundomin suurpiiri	60	26	24	1,8
Vantaa yhteensä	40	14	21	1,9
Myyrmäki	37	13	19	2,2
Kivistö	55	21	22	2,2
Aviapolis	35	12	20	2,2
Tikkurila	35	12	20	1,9
Koivukylä	43	18	22	1,8
Korso	52	22	24	1,8
Hakunila	40	16	23	1,9
Kauniainen	40	14	19	2,1

Lähteet: Tilastokeskus, työssäkäyntitilasto; HSL, Reittiopas

Tätä suotuisampi matka-ajan ja etäisyyden suhde on esimerkiksi itäisessä suurpiirissä, joka koostuu mm. Vuosaaren, Mellunkylän ja Myllypuron alueista. Ilmeisin tulkinta alueiden väliselle erolle on metroyhteys Itä-Helsingistä keskustaan.

Alustavasti voi sanoa, että vaikka esimerkiksi Helsingin pohjoisen suurpiirin asukkailla on tyypillisesti itäisen suurpiiriin asukkaita lyhyempi matka työpaikalle, kuluttavat pohjoisen suurpiirin asukkaat yhteen kilometriin itäisen suurpiiriin asukkaita enemmän aikaa. Itäisen suurpiiriin asukkaiden näkökulmasta työpaikan saavutettavuus julkisilla on siis parempi.

Tulosten tulkintaa hankaloittaa kuitenkin kärkeä aluetaso ja lyhyiden matkojen tulkinta. Esimerkiksi Helsingin kaakkoinen suurpiiriin alueella sijaitsevat Herttoniemi ja Laajasalo. Herttoniemestä pääsee metrolla keskustaan, Laajasalosta ei.

Tässä mielessä tulokset eivät ole lopullinen

totuus asiasta; myös suuralueiden sisään mahtuu monenlaisia alueita.

Lisäksi alueilla, joilla on paljon alle kahden kilometrin matkoja, kulmakertoiminta kasvattaa se, että reittiopas saattaa tarjota kävelyratkaisua.

Alueiden väliset erot kasvaisivat selvästi, jos pitkiä ja lyhyitä työmatkoja ei otettaisi huomioon ja kulmakertoimet laskettaisiin vain työmatkojen keskipitkälle kvartiilivälille (5–16 kilometriä). Helsingissä asuvien työmatkan kulmakertoimeksi saataisiin 1,8, kun puolestaan Espoon kulmakertoimeksi tulisi 2,2.

Näin laskettuna Helsingin pohjoisessa suurpiirissä asuvien kulmakeroimien kasvaksi 2,7:ään ja keskinen suurpiiriin kulmakeroimien laskiksi 1,5:een. Kymmenen kilometrin matka pohjoisessa suurpiirissä kestäisi 27 minuuttia ja keskisessä suurpiirissä vartin.

Ääriarvojen sivuuttaminen saattaisi antaa paremmin kuvan matka-ajoista ja saavutettavuudes-

ta, jos esimerkiksi kävelymahdollisuutta ei tarvitsisi ottaa huomioon. Toisaalta voi ajatella, että jos tarkoituksena on kuvata alueen todellisia matka-aikoja, lienee perusteltua laskea mukaan kaikki matkat.

Koko maan tietoa suunnittelun tueksi

Vaikuttaa siltä, että Reittioppaan kaltaisia palveluita voisi hyödyntää työmatkojen ja -aikojen tilastoinnissa. Tulosten luotettavuutta voi arvioida vertailemalla tuloksia aiempiin työmatkalaskelmiin. Tässä esitetyt laskelmat istuvat hyvin yhteen sen kanssa, mitä työmatkoista ja -ajoista tiedetään. Matkat – ja sitä myötä matka-ajat – kasvavat siirryttäessä kauemmaksi kaupungin keskustasta. Lisäksi laskemani työmatka-ajan mediaani osuu hyvin lähelle HSL:n kyselytutkimuksessa saatua työmatka-ajan mediaania. Aiemman tiedon valossa tulokset ovat oikeansuuntaisia.

Vaikka tulokset ovat oikeansuuntaisia, on tulkinnaissa otettava huomioon, että laskelma on luonteeltaan teoreettinen. Työmatka-ajan mediaani olisi 34 minuuttia, jos kaikki matkustaisivat töihin julkisilla liikennevälineillä. Koska käytettävissä ei ole tietoa siitä miten työmatkat tosiasiaa kuljetaan, ei todellista matka-aikaakaan tiedetä. Tunnusluvut kertovat enemmän saavutettavuudesta kuin tosiasiallisista työmatkoista ja niihin kuluva ajasta.

Laskelman teoreettisuus korostuu etenkin lyhyiden matkojen kohdalla, sillä todennäköisesti suuri osa alle kilometrin matkoista kuljetaan kävellen eikä julkisilla kulkuvälineillä.

Kelpuutin mukaan tähän tarkasteluuni kaikki julkisen liikenteen matkat ja matka-ajat sellaiseenaan, riippumatta siitä ehdottiko Reittiopas vaihtoehtoksi kävelyä. Voi perustellusti kysyä, pitäisikö tunnuslukuja laskiessa olettaa, että esimerkiksi kaikki alle kilometrin tai puolen kilometrin matkat kuljetaan kävellen.

Ratkaisu riippuu oleellisesti siitä mitä pyritään kuvaamaan. Jos tahdotaan kuvata esimerkiksi liikuntarajoitteisten mahdollisuuksia käyttää julkista liikennettä, ei kannata olettaa, että alle kahden kilometrin matkat sujuisivat kävellen.

Reittioppaan rajapinta taipuu kuitenkin erilaisiin tarpeisiin tarjoamalla mahdollisuuden muokata reitin valintaan liittyviä oletuksia (mm. kävelynopeus, liikenneväline ja vaihtoihin käytettävä maksimiaika).

Tilastojen näkökulmasta Reittioppaan suurin puute on alueellinen rajoite – se ei kata kuin pienen osan Suomea. Vaikka pääkaupunkiseudulla julkisen liikenteen käyttö lienee aivan toista luokkaa kuin muualla Suomessa, se ei riitä perusteluksi keskittyä pelkästään pääkaupunkiseutuun. Päinvastoin: julkisen liikenteen vähäinen käyttö ja heikko peittävyys tekevät tilastoinnista entistä

tärkeämpää. Koko maan tasolla tuotettu vertailtava tieto olisi valttia liikenne- ja aluesuunnittelussa.

Liikennevirasto onkin koonnut yhteen palveluun (matka.fi) eri kaupunkien joukkoliikenneaikatauluja, VR:n juna-aikatauluja sekä pitkän matkan bussilinjoja. Tässäkin palvelussa on avoin rajapinta. Teoriassa työmatka- ja aikalaskenta julkisilla liikennevälineillä voitaisiin siis toteuttaa koko maan tasolla. Käytännössä laskennan voisi keskittää suurimpiin kaupunkeihin, joissa joukkoliikenteen käyttö on yleisintä.

Suurimpien kaupunkien asukkaiden työmatkojen kartoittaminen edellyttäisi suurta määrää kyselyitä edellä mainittuun rajapintaan. Kyselyiden teko ei juurikaan vie työaikaa, mutta kylläkin jonkin verran kalenteriaikaa ja palvelinresursseja. Tämä pitää ottaa huomioon aineistojen koostamisesta suunniteltaessa

Mikäli matka-aika ja pituus selvitettäisiin kaikille työllisille, pystyttäisiin tekemään hyvinkin tarkkoja alueellisia tilastoja työmatkoihin kuluvista ajoista. Työmatkatietoihin on myös mahdollista yhdistää taustatietoja työmatkaajista (esim. kotitalouden autojen lukumäärä, tulot jne.). Näin syntyisi kullannarvoinen tietopaketti joukkoliikenteestä ja yhdyskuntasuunnittelusta kiinnostuneille tutkijoille. ■

Kirjoittaja on yliaktuaari Tilastokeskuksen väestö- ja elinoloilastot -yksikössä.

Lähteet:

HSL. Reittiopas. <http://www.reittiopas.fi/>.

HSL 2013. HLJ 2015. Liikku mistottumukset Helsingin seudulla 2012. HSL:n julkaisuja 27/2013. Helsingin Seudun Liikenne. https://www.hsl.fi/sites/default/files/uploads/liikku mistottumukset_helsingin_seudulla_2012_hlj2015_raportti_0.pdf.

Huhta, Jaana & Pyykkönen, Topias 2013. Linnuntieltä oikeille teille. Hyvinvointikatsaus 4/2013.

Piela, Pasi 2014. Commuting time for every employed: combining traffic sensors and many other data sources for population statistics. European Forum for Geography and Statistics (EFGS) Krakow Conference 2014. Puolan tilastovirasto. <http://geo.stat.gov.pl/efgs/programme>.

Piela, Pasi 2015. Entä jos kaikki polkisivat töihin? Tieto&trendit -blogi. <http://tietotrenditblogi.stat.fi/enta-kaikki-polkisivat-toihin/>.

Tilastokeskus. Työssäkäyntitilasto. <http://www.stat.fi/til/tyokay/>.

[Kommentoi verkossa tietotrendit.stat.fi](#)

JUHA SAUKKONEN

Kirjoittaja on lehtori (johtaminen)
Jyväskylän ammattikorkeakoulussa.

Johtaminen edellyttää taitoa ja tietoa

Johtamisen guruja, opettajia ja harjoittajien suhde tietoon on ollut kompleksinen ja intensiteitiltään kovin vaihteleva. Muutama vuosi sitten vietettiin taylorismin 100-vuotista taivalta. **Frederick Taylorin** teos *Principles of Scientific Management* julkaistiin melko tarkasti 200 vuotta sen jälkeen, kun termi johtaja (manager) ilmestyi kielenkäyttöön tuolloin johtavassa maailmanmahdissa Isossa-Britanniassa. Taylorin opit kohensivat tuottavuutta ja levisivät nopeasti läpi teollisuusalojen. Suunnittelu, tehokkuus, mittaaminen ja palkitseminen ovat kaikki edelleen läsnä johtamisen opeissa, puheissa ja käytännöissä.

Samaan aikaan, kun Taylorin jalostettuja oppeja otettiin käyttöön, teknologinen ja sosiologinen murros loi kokonaisia uusia tuote- ja palvelukategorioita, joihin ei olisi ikinä ylletty analyyttisellä otteella. Muistettiin jälleen **Joseph Schumpeterin** opit luovasta tuhosta, saatiin strategiapalaverien vakiosanastoon ilmaisuja kuten ”thinking outside the box” ja ”uusitutuminen”.

Clayton M. Christensen pohti maailman menoa kirjassaan *Innovator's dilemma* ja eritteli syitä, miksi johtavat yritykset menettivät asemansa alansa murroskohdissa, tuli murros sitten teknologian tai uuden liiketoimintamallien tuomana. Näitä Christensenin mainitsemia ”halvaantumisen” syitä olivat mm. se, että suurimmilla toimijoilla on suurimmat asiakkaat, toimittajat ja omat resurssit eli eniten epärelevantiksi muuttuvaa tietoa. Yksi Christensenin havainto oli se, että markkinasta, jota ei vielä ole, ei voi olla myöskään tietoa. Ja ilman tietoa ei voi olla päätöksiä. Tilausta oli uudelle näkemykselle: mitä tietoa yritysjohto tarvitsee, mistä sitä voidaan kerätä, miten sitä voidaan prosessoida?

Big data ilmaantui vuonna 2011 johtavan ICT-alan konsulttiyhtiö Gartnerin vuosittain julkaisemalle kuumien ICT-teknologioiden hype-käyrälle. Tuolloin sen kehitysvaiheeksi luonnehdittiin *teknologinen käynnistäjä*. Tietokoneiden ja tietokantaohjelmien suorituskyvyn kautta tiedettiin, että big data on teknisesti mahdollista. Gartner arvioi, että suurin osa yrityksistä ottaa ison datan käyttöön operatiivisessa toiminnassaan 2–5 vuoden kuluessa.

Tultiin viime vuoteen eli heinäkuuhun 2014 – iso data oli ehtinyt kolmessa vuodessa Gartnerin

terminologialla jo ylittää ylikuumenneiden odotusten huipun ja liukui kohti pettymysten laaksoa. Vaikka kyseinen kehityskulku onkin Gartner-käyrän perusluonteen mukainen, niin samalla odotusaika yritysten valtavirran hyppäämisestä ison datan kelkkaan on kasvanut 5–10 vuoteen. Ollaan siis 2020-luvun puolella ennen kuin saamme nauttia päivittäin ison datan hedelmistä.

Mikä meni vikaan – vai menikö mikään? Johtajuuden asiantuntijajulkaisu *Harvard Business Review* omisti vuonna 2012 kaksi laajaa artikkelia aiheelle: ”Johtajuuden kehitys algoritmin aikakaudella” (Buckingham) ja ”Miten saat edistyneen analytiikan tekemään työtä hyväksesi” (Barton & Court). Tarkoitus oli sivistää lukijakuntaa big datan olemuksesta sekä opastaa sen käytäntöön soveltamisessa osana johtamisen järjestelmiä ja käytäntöjä. Itse uskon, että luetellessaan datalla johtamisen etuja kirjoittajat samalla kertovat sen heikkoudet. Kuuluisat ”3 V”-tekijää: volyyymi, vaihtelevuus (variety), vauhti (velocity) sekä houkuttavat että pelottavat. Data on ikiliikkua, se ei pysy paikallaan eikä muodossaan. Ja johtamisen kannalta olennaisin on neljäs V: valinta.

Johtamisen muuttumaton peruspilari on, että kaikki päätökset tehdään epätäydellisen informaation varassa. Kyse on siis informaatiosta, ei datasta. Yleisesti hyväksyttyä taksonomiaa seuraillen (data, informaatio, tietäminen) sillä joka päättää, mitä ja miten dataa kerätään, milloin sitä on riittävästi ja milloin se on relevanttia, on oltava tietämystä alasta ja päätöksistä mitä aiotaan tehdä, mikä on tarvittavaa informaatiota, mikä jätetään pois. Teknisten toteuttajien on sitten tuotettava informaatio kaikkia ”3+1 V” tekijöitä hyödyntäen.

Ei siis ihme, että hämmennys uuden asian edessä on suuri. **Henry Mintzbergin** mukaan tulevaisuudessa tulisi opettaa johtamisen taitoa – Art of Management – ei vain valikoimaa analyyttisiä työkaluja. On kuitenkin mahdotonta nähdä, että ison datan aalto olisi pysäytettävissä, vaikka vesi vetäytyykin juuri nyt rannasta. Datan avulla johtaminen tulee varmasti olemaan osa johtamisen taitoa.

Käsi sydämelle: Kuka haluaa työskennellä organisaatiossa, joka ei tiedä mitä tietää? Kuka haluaa työskennellä organisaatiossa, joka uskoo kaiken tietämyksensä olevan olemassa datana? ■

Tilastokeskus kerää vuosittain väestön elinoloihin liittyviä tietoja noin 200 000 kansalaiselta. Tietojen pohjalta laaditaan sosiaali-tilastoja, jotka muodostavat kattavan kuvan suomalaisten elämästä ja olosuhteista. Kansalaisilta kysyttäviin haastattelu-tietoihin perustuvat muun muassa työllisyys- ja tulonjakotilastot sekä väestön tietotekniikan käyttöä, kulutusta, ajan-käyttöä ja vapaa-ajan toimintaa kuvaavat tilastot. Näitä tiedonkeruita kutsutaan henkilötiedonkeruiksi, ja ne kuuluvat Suomen virallisen tilastoinnin piiriin.

Nykyisin virallisten tilastojen tiedot kerätään kansalaisilta pääosin puhelinhaastattelujen avulla. Myös käyntihaastatteluja käytetään eräissä isohkoissa tiedonkeruissa. Haastattelujen määrä on valtava; yksinomaan työvoimatutkimuksessa tehdään joka kuukausi noin 10 000 puhelinhaastattelua. Tiedonkeruuta on tekemässä jatkuvasti lähes 200 haastattelijaa.

Tietotekniikka ja internetin käyttö alkavat olla arkipäivää jo suurimmassa osassa suomalaiskoteja. Alle 65-vuotiaista suurin osa käyttää internetiä lähes päivittäin, ja aktiivikäyttäjää löytyy yhä runsaammin myös vanhimmissa ikäryhmissä (kuvio 1). Erilaiset nettikyselyt ovat kansalaisille tuttuja, ja myös monet viranomaisasiat voidaan hoitaa internetin välityksellä. Netissä voi esimerkiksi tilata verokortin, hakea Kelasta tukia ja tehdä päivähoitohakemuksen, ja suuri osa kansalaisista myös käyttää näitä mahdollisuuksia (kuvio 2). Tämän myötä tuntuu luontevalta, että myös virallisen tilaston tiedonkeruuseen olisi mahdollista vastata verkon kautta. Tilastokeskuksen tavoitteena onkin mahdollistaa verkossa vastaaminen kansalaisille vaiheittain vuodesta 2016 alkaen. Yritystilastoissa ja ulkopuolisella rahoituksella toteutetuissa tilaustiedonkeruissa verkkovastaaminen on ollut käytössä jo pitkään.

Viimeksi kuluneiden parin vuoden aikana verkkotiedonkeruuta on testattu pilottitutkimusten avulla kahdessa erityyppisessä tiedonkeruussa: kansalaisten talouteen liittyviä mielikuvia kartoittavassa kuluttajabarometrissa sekä työllisyys- ja työttömyyslukujen pohjana olevassa työvoimatutkimuksessa. Pilottitutkimuksista saadut kokemukset osoittavat, että virallisten henkilöperustisten tilastojen tiedonkeruusiin liittyy erityisiä haasteita.



Henkilötiedonkeruuta tehdään pian myös verkossa

Tilastokeskuksen tavoitteena on mahdollistaa kansalaisille verkossa vastaaminen vuodesta 2016 alkaen. Verkkovastaaminen edellyttää kuitenkin paljon tiedonkeruusiin liittyvää menetelmätyötä ja testaamista. Eri menetelmin saatujen tietojen tulee olla keskenään vertailukelpoisia, vastauskato ei saa nousta, tietoturva tulee huolehtia ja kansalaisia tulee kohdella tasapuolisesti.

Yhdistelmätiedonkeruu käyttää useita menetelmiä

Haastatteluja ei täysin korvata verkkolomakkeilla, vaan suurimmassa osassa henkilötiedonkeruuta tullaan käyttämään kahden rinnakkaista tiedonkeruumenetelmää. Kyse on siis niin kutsutusta yhdistelmätiedonkeruusta (mixed mode data collection), jossa osa tiedoista saadaan verkkotiedonkeruun kautta, osa haastattelemalla.

Kansalaisten tasapuolisten vastaamismahdollisuuksien turvaaminen on Tilastokeskukselle tärkeää, joten esimerkiksi tietokoneen puuttuminen ei saa olla esteenä tutkimukseen osallistumiselle. Ja vaikka välineissä ja taidoissa löytyisikin, vastaaja voi muusta syystä kokea verkossa vastaamisen epämieluisaksi.

Haastatteluja tarvitaan paikkaamaan katoa, sillä aiempien pilottitutkimusten perusteella pelkällä verkkotiedonkeruulla

ei päästä riittäviin vastausosuuksiin. Kun kuluttajabarometrissa testattiin yhdistelmätiedonkeruuta, saatiin vastauksista 42 prosenttia verkkotiedonkeruun kautta, ja työvoimatutkimuksen verkkotiedonkeruupilotissa vastausaste oli 30 prosenttia. Puhelimitse tehtävän työvoimatutkimuksen vastausosuus on noin 70 prosenttia.

Tiedonkeruilla on usein tiukka aikataulu, eikä vastauksia voida pitkään odotella. Toisaalta väljemmstä aikataulustakaan ei välttämättä olisi merkittävää hyötyä. Esimerkiksi työvoimatutkimuksen verkkotiedonkeruupilotissa suurin osa verkkovastauksista saatiin heti tiedonkeruun käynnistyttyä tai muistutusviestien jälkeen. Jos vastaaminen jää tuonemmaksi, se jää helposti kokonaan, vaikka vastaaja suhtautuisi kyselyyn periaatteessa myönteisesti. Tällöin haastattelijan yhteydenotto on tarpeellinen vastausten saamiseksi.

Haastattelemalla saatujen vastaus- ➤

ten osuus tulee siis ainakin toistaiseksi olemaan tiedonkeruissa suurempi kuin verkkovastausten osuus. Ajan myötä tämä suhde voi toki muuttua.

Tietojen laadusta on huolehdittava

Uuden tiedonkeruumenetelmän käyttöönotto ja usean tiedonkeruutavan yhdistäminen tuo mukanaan tietojen laatuun ja vertailtavuuteen liittyviä haasteita. Lopullisten tietojen tulee olla vertailukelpoisia tiedonkeruutavasta riippumatta, ja aikasarjojen jatkuvuudesta on huolehdittava.

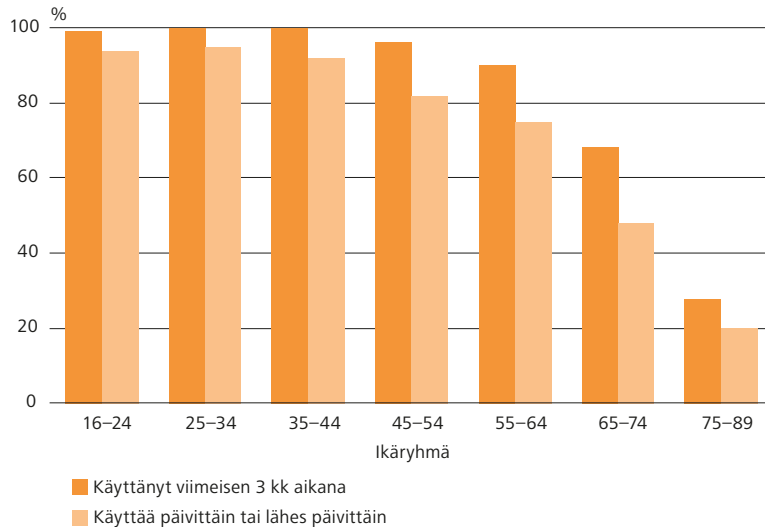
Henkilötiedonkeruiden pohjalta tuetaan monia poliittisesti keskeisiä indikaattoreita, kuten kuluttajabarometrin perusteella laskettava kuluttajien luottamusindikaattori, työvoimatutkimukseen perustuvat työttömyys- ja työllisyysasteet sekä tulo- ja elinolotutkimukseen pohjautuva pienituloisuusaste. Erityisesti taloudellisesti epävakaina aikoina epäily siitä, että muutokset näissä luvuissa johtuisivat uudesta tiedonkeruutavasta, vaarantaisi tilastotietojen luotettavuuden vakavasti.

Verkossa ja haastattelussa käytettävät tiedonkeruulomakkeet eivät voi olla ominaisuuksiltaan täysin samanlaisia, sillä vastaaminen on erilaista eri menetelmillä. Verkkolomake nojaa visuaalisiin ratkaisuihin, joiden avulla vastaaminen yritetään tehdä mahdollisimman helpoksi, kun taas haastattelussa haastattelija voi motivoida ja auttaa vastaajaa.

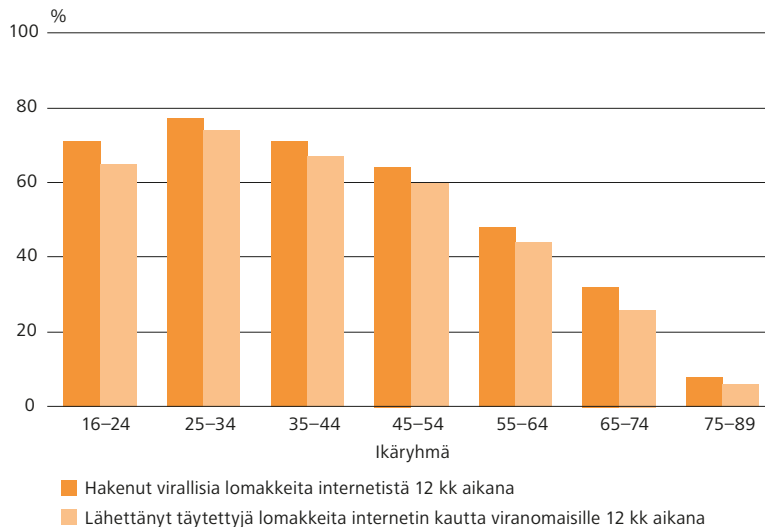
Tiedonkeruu on kuitenkin muutoksille herkkää, ja pieniltäkin tuntuvat muutokset lomakkeella saattavat vaikuttaa tuloksiin yllättävällä tavalla. Yksi esimerkki tästä on ”en osaa sanoa” eli eos -vastausvaihtoehdon esittämistapa lomakkeella. Tutkimukset ovat osoittaneet, että eos-vastauksen tarjoaminen yhtenä näkyvänä vastausvaihtoehtona lisää eos-vastausten osuutta erityisesti tieto-, mielipide- ja asennekysymyksissä ja siten heikentää tiedonkeruun laatua (ks. Jokinen & Järvensivu 2014).

Tämä nähtiin selvästi kuluttajabarometrin pilottitutkimuksessa, jossa eos-vaihtoehto oli näkyvissä verkkolomakkeella, mutta sitä ei puhelinhaastattelussa luettu ääneen. Eos-vastausten osuus oli selvästi suurempi verkkolomakkeella kuin puhelimesta vastanneilla, mikä heikentää eri keruutapojen vertailukelpoisuutta.

KUVIO 1. Suomalaisen internetin käyttö ikäryhmittäin vuonna 2014



KUVIO 2. Suomalaisen internetin käyttö viranomaisasiointiin vuonna 2014



Kuvioiden lähde: Tilastokeskus, väestön tieto- ja viestintätekniiikan käyttö

Keruutapa voi vaikuttaa selvästi tuloksiin myös kysymyksissä, joissa vastaus valitaan usean vaihtoehdon joukosta. Kun haastattelija lukee vaihtoehdot ääneen, vastaajan mieleen jäävät listalta helpoimmin viimeksi mainitut, ja vastaus valitaan herkemmin näiden joukosta. Verkkolomakkeella annetut vastaukset puolestaan painottuvat listan alkupäähän, sillä vastaaja valitsee listalta herkästi ensimmäisen, joka tuntuu sopivalta. Näin ollen samaan kysymykseen

voidaan saada haastattelussa ja verkkolomakkeella erilainen vastausjakauma. (Bethlehem & Biffignandi 2012.)

Tiedonkeruuta testataan monin tavoin

Tietojen laadunvarmistukseen on onneksi keinoja. Tärkein keino on ammattitaitoisen ja huolellisen lomakesuunnittelun ja huolellisen lomakesuunnittelun, joka Tilastokeskuksessa osataan. Olennaista on myös lomakkeiden testaus aidoilla vas-

taajilla. Testaus voi olla joko laadullista tai kvantitatiivista, ja molempia tarvitaan näin laajamittaisen uudistuksen yhteydessä.

Laadullisessa testauksessa mukana on pieni joukko vastaajia, joita pyydetään vastatessaan kertomaan siitä, miten he lomakkeen kysymykset ymmärtävät ja miltä vastaamiskokemus tuntuu. Laadullinen testaus on joustavaa ja kustannustehokasta, ja sen avulla saadaan arvokasta tietoa lomakkeen kehittämiseksi. Erityisen hyödyllistä laadullinen testaus on silloin, kun halutaan tutkia jonkin rajatun, juuri kyseiselle tiedonkeruulle kriittisen kohderyhmän vastaamista.

Kvantitatiivisessa testauksessa kyse on pilottitutkimuksesta, jossa tutkimusasetelma on lähellä aitoa tiedonkeruuta ja vastaajajoukot melko laajoja. Pilottitutkimus on usein eräänlainen kentraaliharjoitus, jonka avulla voidaan varmistua tiedonkeruun toimivuudesta ennen muutosten käyttöönottoa varsinaisessa tiedonkeruussa. Pilottitutkimuksen tuloksia voidaan tutkia tilastollisin menetelmin ja siten selvittää esimerkiksi muutoksia keskeisissä indikaattoreissa.

Tänä vuonna Tilastokeskus toteuttaa yhdistelmä tiedonkeruun käyttöönottoon liittyvää laadullista testausta työvoimatutkimuksessa ja asuntojen vuokratiedustelussa sekä kvantitatiiviset pilottitutkimukset väestön tieto- ja viestintätekniikan käyttö - tutkimuksessa ja työvoimatutkimuksessa. Testikierrosten tuloksia saadaan käyttöön vaiheittain vuosien 2015–2016 aikana, ja niitä hyödynnetään yhdistelmä tiedonkeruun käyttöönoton suunnittelussa.

Vastauskokemus vaikuttaa katoon

Lomakesuunnittelu on tärkeää paitsi tietojen laadun kannalta, myös kadon hillitsemisen keinona. Vastauskadon kasvu on yksi keskeisimmistä huolenaiheista henkilötiedonkeruissa.

Yrityksille ja yhteisöille tilastotiedonkeruisiin vastaaminen on usein pakollista, mutta kansalaisten osallistuminen perustuu vapaaehtoisuuteen. Haastattelujen yhteydessä haastattelija voi suostutella vastahakoistakin vastaajaa jatkamaan kyselyn loppuun, mutta verkossa vastatessaan vastaaja on omillaan. Jos vastaaminen tuntuu hankalalta tai epämiellyttävältä, verkossa

kyselyyn vastaaminen on helppo lopettaa.

Tiedonkeruulomakkeiden tekniseen toimivuuteen ja käytettävyyteen sekä sisällön ymmärrettävyyteen on siis panostettava, jotta vastaaminen olisi mahdollisimman miellyttävä kokemus. On myös kiinnitettävä erityistä huomiota vastaajien motivointiin tiedonkeruumateriaaleissa (kuten tiedonkeruusta kertovissa saatekirjeissä ja Tilastokeskuksen verkkosivuilla).

Erityisen tärkeää motivointi on toistuvien tiedonkeruiden kohdalla; esimerkiksi työvoimatutkimuksessa saman vastaajan pitäisi pysyä mukana kaikkiaan viisi ja tulo- ja elinolotutkimuksessa neljä tiedonkeruukertaa. Vastaamiskokemuksen laatu vaikuttaa ilman muuta siihen, miten innokkaasti vastaaja seuraavalla kerralla lähete lomaketta täyttämään.

Tietoturva on varmistettava

Vastaajan tietoturva huolehtiminen on ensiarvoisen tärkeää kaikissa tiedonkeruissa. Yksittäisen vastaajan antamat tiedot ovat ehdottoman luottamuksellisia, ja myös verkkotiedonkeruun puolella on varmistuttava siitä, ettei vastauksia pääse vuotamaan tiedonkeruujärjestelmän ulkopuolelle. Vastaajien luottamuksen säilyttäminen on Tilastokeskukselle ensiarvoisen tärkeää, eikä tietoturva saa vaarantua missään olosuhteissa.

Tietoturvaan liittyy myös se, että varmistetaan vastaajan henkilöllisyydestä. Tilastokeskuksen henkilötiedonkeruut perustuvat väestötietojärjestelmästä poimituun satunnaisotokseen. Vastaajaksi käy vain otokseen poimittu henkilö, eikä alkuperäistä vastaajaa voi korvata toisella henkilöllä. Usein lomakkeelle on viety valmiiksi tietoja, jotka ovat peräisin saman tiedonkeruun edelliseltä vastauskerralta, joten luonnollisesti on tärkeää, että tiedot ovat vain vastaajan itsensä nähtävissä.

On siis rakennettava tunnistautumisjärjestelmä, jonka kautta otokseen valittu pääsee kirjautumaan lomakkeelle. Järjestelmä varmistaa, että kyse on juuri oikeasta henkilöstä. Pankkitunnukset ovat varsin yleisesti käytössä viranomaisten internet-palveluissa. Pankkitunnusten kautta kirjautuminen on yksi vaihtoehto myös Tilastokeskuksen verkkotiedonkeruissa.

Toinen vaihtoehto on käyttää erikseen määriteltäviä kirjautumistunnuksia, jotka lähetetään vastaajalle henkilökohtai-

sen saatekirjeen mukana. Molemmissa ratkaisuissa on hyvät ja huonot puolet, ja ne soveltuvat vaihtelevasti eri tiedonkeruiden tarpeisiin. Lopullinen käyttöön otettava järjestelmä tulee todennäköisesti olemaan jonkinlainen yhdistelmä eri ratkaisuja.

Verkkovastaaminen ei tuo nopeita säästöjä

Verkkotiedonkeruuta kehitetään parhailaan myös muissa Euroopan tilastovirastoissa. Työ on edennyt varovaisin askelin, eikä täysin valmista ratkaisua ole tarjolla missään. Tiedonkeruiden ja maiden lähtötilanteen erilaisuudesta johtuen samat ratkaisut eivät toimi kaikissa maissa. Samassa tilanteessa Suomen kanssa ovat vain muut Pohjoismaat ja muutama muu Euroopan maa kuten Hollanti, Viro ja Itävalta.

Selvää kuitenkin on, että verkkotiedonkeruun käyttöönotto on lähitulevaisuuden tavoitteena kaikissa Euroopan tilastovirastoissa. Tähän vaikuttavat paitsi tietoteknisten välineiden läpimurto, myös toive säästää tiedonkeruukustannuksissa.

Uusi tiedonkeruumenetelmä ei tuo ihmepeleistä tilastovirastojen budjetteihin. Kehittämistyö vaatii resursseja ja kustannushyödyt realisoituvat vasta pitkällä tähtäimellä, kun yhdistelmä tiedonkeruu saadaan käyttöön täysimittaisesti. Ajan myötä vastaajat oppivat kuitenkin varmasti hyödyntämään verkkovastaamisen hyviä puolia, ja verkkotiedonkeruiden suunnittelusta tulee rutiinia myös tilastoviranomaiselle. ■

Kirjoittaja toimii yhdistelmä tiedonkeruusta vastaavana projektipäällikkönä Tilastokeskuksen tiedonhankinta-yksikössä.

Lähteet:

Bethlehem, Jelke & Biffignandi, Silvia 2012. Handbook of Web Surveys. Hoboken, New Jersey: John Wiley & Sons.

Jokinen, Marika & Järvensivu, Marjaana 2014. En osaa sanoa -vastaaminen verkkokyselyissä. Hyvinvointikatsaus 4/2014.



MARI YLÄ-JARKKO

Mari Ylä-Jarkko vetää Tilastokeskuksen talous- ja ympäristötilastojen tuke-vastuualueetta. Hän on osallistunut kuluttajahintaindeksin kehittämistoi-mintaan Suomessa sekä maailmalla.

Negatiivinen inflaatio, hintavakaus vai deflaatio?

Kuluttajahintaindeksin vuosimuutos on ollut muutaman kuukauden miinuksella. Kuitenkin Tilastokeskuksen julkistuksessa puhutaan inflaatiosta – pitäisikö puhua jo deflaatiosta?

Inflaatio käsitetään taloustieteessä yleisesti kuluttajahintojen nousuksi ja deflaatio tämän vastakohdaksi. Hintavakaudesta puhutaan silloin, kun hinnat eivät merkittävästi nouse tai laske. Euroopan keskuspankin EKP:n neuvosto on määritellyt hintavakaudeksi euroalueen yhdenmukaistetun kuluttajahintaindeksin alle kahden prosentin vuotuisen nousun.

Jotta voidaan puhua deflaatiosta kansantalouden näkökulmasta, tulee talouden kehitystä tarkastella laajemmin kuin vain kuluttajahintaindeksin mukaan. Deflaatioissa yleinen hintataso laskee taloudellisen kehityksen seurauksena, ku-

lutus vähenee ja palkat alenevat.

Itseään ruokkivan kierteen synnyttää helposti se, että nekin kuluttajat, joilla olisi varaa kuluttaa, odottavat hintojen alentumista. Tällöin kysynnän pieneminen johtaa tuotannon pienentymiseen ja työttömyyden kasvuun ja kulutettavissa olevan rahamäärän vähentymiseen.

Ansiot nousseet hintoja enemmän

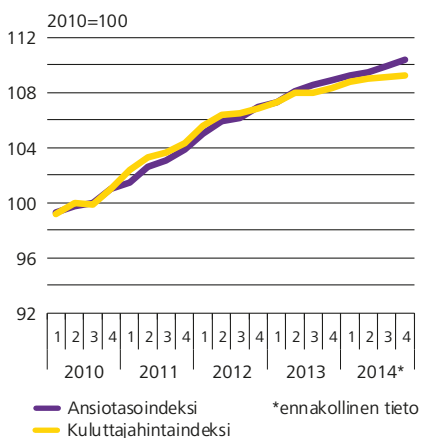
Tarkasteltaessa nykyistä tilannetta voidaan havaita, että kaikki hinnat eivät ole laskeneet eikä hintojen laskun syynä ole

palkansaajien ostovoiman väheneminen. Vuosina 2011–12 reaaliensiot tosin hieman laskivat, mutta pidemmällä aikavälillä kuluttajien reaaliensiot ovat nousseet huomattavasti (Kuvio 1).

Sen sijaan Suomen kuluttajahintoja ovat pitkään hillinneet ulkoisista tekijöistä johtuvat syyt: korkotason lasku sekä kulutuselektronikan ja polttonesteiden halpeneminen.

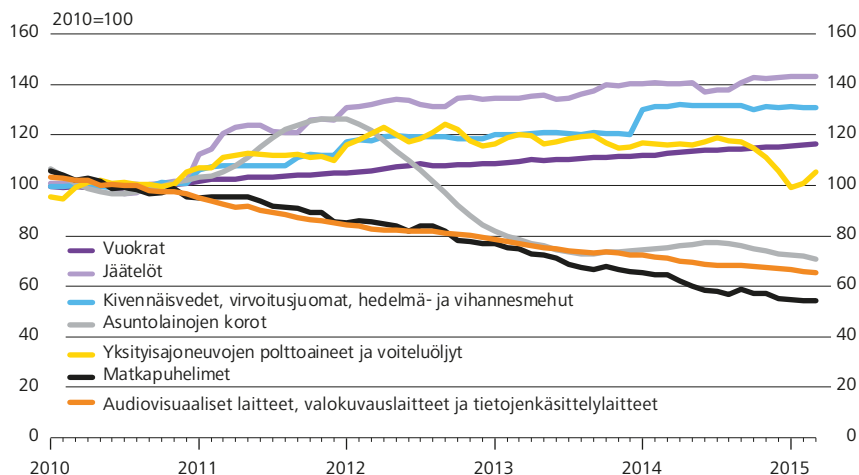
Kulutuselektronikan osalta kehittyneemmät tuotantotavat ja sitä kautta säävutettu tuottavuuden nousu ovat laske-neet hintoja, jopa yli 40 prosenttia vuodes-ta 2010 (Kuvio 2). Öljyn maailmanmark-

KUVIO 1. Ansiotason ja kuluttajahintojen kehitys 2010–2014



Lähde: Tilastokeskus, Ansiotasoindeksi & Kuluttajahintaindeksi

KUVIO 2. Eräiden kulutushyödykkeiden hintakehitys 2010–2015



Lähde: Tilastokeskus, Kuluttajahintaindeksi

50 000 hintatietoa joka kuukausi

Kuluttajahintaindeksin paikkansapitävyyttä kyseenalaistetaan jatkuvasti, kun on vilkaistu muutamia hintoja lähikaupassa tai kahvilassa. Ihan niin helppoa inflaation mittaaminen ei kuitenkaan ole.

”Miten se inflaatio lasketaan? Maailmassa myytäviä tuotteita on miljoonittain josten lukuja voidaan siirtää tikkarista nostokurkeen. Vuokrat? Keinotekoinen hinnan muutos.”

Lainaus on Taloussanomien nettikeskustelusta, mutta sama epäily tulee eri yhteyksissä ja muodoissa vastaan: hinnat ovat varmasti nousseet enemmän kuin Tilastokeskus kertoo...

Mitä on inflaation takana, miten se lasketaan?

Yleisimmän inflaatiomittarin, kuluttajahintaindeksin pohjana on ensiksikin Suomessa tapahtuva kotitalouksien kulutus. Kulutusrakenteen selvittämistä varten toteutetaan kulutustutkimus, jossa otokseen valikoidut kotitaloudet pitävät kirjaa kulutusmenoistaan. Otoksessa huolehditaan siitä, että mukana on kaiken kokoisia kotitalouksia ympäri Suomea siinä suhteessa kuin niitä on kaikista kotitalouksista.

Näitä tietoja korjataan ja tarkistetaan. Korjauksia tehdään eritoten alkoholin kuluutukseen ja tupakkaan – näitä kun harvoin ilmoitetaan ihan niin paljoa kulutetuiksi kuin niitä ostetaan.

Kulutustietojen pohjalta indeksiin laaditaan painorakenne ja hyödykeotos. Otoksen laadinnassa apuna käytetään muutakin myyntitietoa, jotta mittaukseen tulisi oikeanlaisia hyödykkeitä – siis niitä joita tavalliset ihmiset tavallisissa kotitalouksissa ostavat.

Suomalainen kotitalous laittaa suurimman osan rahoistaan asumiseen, noin 21 prosenttia, ja ruokaan, 14 prosenttia. Vaihdelua on hieman alueittain, siksi kuluttajahintaindeksissä on oma painorakenteensa kuudelle eri alueelle.

Kun on valittu kulutusta hyvin edustavat hyödykkeet – lähes 500 tuotetta ja palvelua – valitaan kaupat, joista hinnat kerätään. Osa kaupoista valitaan otoksena (ruokakaupat), osan Tilastokeskuksen haastattelijat ympäri maata valitsevat paikallistuntemuksensa perusteella. Ruoan ja muiden päivittäistavaroiden hintoja havainnoidaan 200 liikkeestä ympäri Suomea, muita liikkeitä hintakeruussa on noin 2 500.

Lisäksi osa hinnoista kerätään netistä tai vieranomaisilta, esimerkiksi valmismatkojen hintoja tai apteekkihintoja. Myös muita tilastoja kuten vuokratiedustelua käytetään hyödyksi. Vuokrat kysytään suoraan henkilöiltä: paljonko on vuokra ja minkä kokoinen asunto.

Kun on tiedossa, mitä tuotteita ja palveluita ja mistä liikkeistä halutaan seurata, on raa’an työn aika. Kuukaudesta toiseen haastatteli ja menee samaan liikkeeseen, valitsee saman tuotteen ja lähettää hinnan laskentajärjestelmään. Aitoja hintoja laskentajärjestelmään kertyy noin 50 000 joka kuukausi.

Kaikista näistä hinnoista lasketaan jokaiselle 500 hyödykkeelle keskimääräinen kuukausimuutos. Vaikkapa rasvattomalle maidolle lasketaan keskimääräinen muutos maaliskuusta huhtikuuhun. Muutokset lasketaan jokaiselle indeksilaskennan kuudelle alueelle, jotka kattavat yhdessä koko Suomen.

Jos samaa tuotetta ei enää löydy hyllystä, etsitään laadullisesti vastaava tuote. Mikäli vastaavaa ei löydy, hinnanmuutos huomioidaan vasta seuraavassa kuussa.

Tavoitteena kuluttajahintaindeksissä on mitata pelkkää hinnanmuutosta. Maitopurkki on yksinkertainen tapaus, kännykkä haastavampi esimerkki. Kännykän indek-

si on laskenut, vaikka kuluttaja pulittaa entistä enemmän uudesta puhelimesta. Hinnan nousu johtuu siitä, että kännykkään on tullut paljon uusia ominaisuuksia. (Oma ensimmäinen matkapuhelimeni maksoi käytettynä 6 000 markkaa vuonna 1994, karkeasti muuntaen lähes tonnin euroissa; nyt laadullisesti vastaavan puhelimen saa parilla kymppillä.)

Painorakenne ratkaisee, paljonko kukin hinnanmuutos kuluttajahintaindeksiin vaikuttaa. Mitä suurempi paino, sitä enemmän pienikin muutos näkyy indeksissä.

Maidonjuojakansana suomalaiset tietävät, että maidon hinta on nousnut viime aikoina – maaliskuusta 2012 maaliskuuhun 2015 esimerkiksi 14,9 prosenttia. Vastavana aikana vuokrat ovat nousseet 9,9 prosenttia ja kokonaisindeksi 2,8 prosenttia.

Suomalaisten kulutukseen laittamista euroista maitoon menee 0,7 prosenttia ja vuokraan 6,9 prosenttia. Maidon hinnan nousu vaikutti koko hintatason nousuun 0,1 prosenttiyksikköä ja vuokran 0,7 prosenttiyksikköä. Eli jos kaikki muut hinnat olisivat pysyneet ennallaan, olisi tuon ajanjakson indeksimuutos ollut maidon ja vuokrien kallistumisen takia 0,8 prosenttia.

Jokapäiväisten ostosten hinnat muistetaan parhaiten. Toisaalta pääsee unohtumaan, että osa hinnoista laskee painaen inflaatiota alaspäin. Esimerkiksi asuntolainojen korot ovat laskeneet 2012 maaliskuusta 2015 maaliskuuhun peräti 42 prosenttia, ja niihin Suomessa kuluu enemmän rahaa kuin maitoon. Toisaalta vuokra-asujilta on jäänyt tämä etu saamatta. Kunkin kotitalouden kohtaama inflaatio on erilainen.

kinahintojen romahdus viime vuoden lopulla ja Euroopan taloustilanteen aiheuttama korkojen lasku vuoden 2012 alusta ovat myös hyödyttäneet suomalaisia kotitalouksia ja parantaneet ostovoimaa.

Toisessa päässä hintakehitystä löytyy elintarvikkeita: juomatiivistä ja jäätelöpaketit, joiden hinnat ovat mm. makeisveron korotusten myötä nousseet yli 60 prosent-

tia vuodesta 2010 maaliskuuhun 2015.

Nyt kolmen kuukauden negatiiviset kuluttajahintojen muutokset eivät ole mitenkään poikkeuksellisia lähihistoriaa tarkasteltaessa. Viimeksi kesäkuusta 2009 tammikuulle 2010 nähtiin miinusmerkkejä ja samoin 2004 maaliskesäkuussa oltiin miinuksien puolella.

Nykytilanteesta on näin ollen oi-

keampaa puhua negatiivisena inflaationa tai hintavakautena, sillä deflaation merkkejä ei taloudessa laajemmin tarkasteltuna tällä hetkellä ole näkyvissä. Ekonomistien näkemys on, että nykyinen hintavakaustilanne säilyy ja siten pitää yllä palkansaajien ostovoimaa. ■

Maaliskuun inflaatiosta enemmän sivulla 85.

Mari Ylä-jarkko



Hiljalleen avautuva Myanmar parantaa inflaationsa mittaamista. Hinnoiksi indeksiin pyritään ottamaan kuluttajan oikeasti maksama hinta eikä listahintaa, kirjoittaa konsulttina toimiva Mari Ylä-Jarkko.

MARI YLÄ-JARKKO

Indeksilaskentaa Myanmarin markkinoilla

Thaimaan naapurissa on pitkään muulta maailmalta suljettuna ollut Myanmar (ent. Burma), jota asuttaa noin 52-miljoonainen kansa. Pikkuhiljaa rajat ovat avautuneet ja markkinat kasvavat, Suomestakin tulee jo lomamatkalaisia.

Myanmarin bruttokansantuote on 1 700 USA:n dollaria asukasta kohden, maa on tässä suhteessa 201. sijalla maailmassa eli kaikkein köyhimpiä. Kansantaloudesta suuri osa muodostuu palveluista (42 %) ja maataloudesta (38 %).

Kansainvälinen valuuttarahasto on muutaman vuoden tehnyt työtä maassa edistääkseen tilastoinnin kehittämistä yleisen inflaatiomittarin parantamiseksi. Tarkoituksena on saada kuluttajahintaindeksi ennen pitkää täyttämään kansainväliset standardit. Tässä työssä olen vieraillut IMF:n lyhytaikaisena asiantuntijana Myanmarin tilastovirastossa maan pääkaupungissa Nay Pyi Tawssa kahdesti, lokakuussa 2014 ja helmikuussa 2015.

Ensimmäisellä kerralla käytiin läpi paikallista indeksilaskentaa ja keskusteltiin siitä, mitä muutoksia pitäisi tehdä. Enimmäkseen ollaan "käden savessa" eli tehdään yhdessä laskentakavioita ja keskustellaan eteen tulevista ongelmakohtista.

Vielä matkaa edustavaan mittariin

Periaatteessa inflaation mittaamistapa on Myanmarissa varsin kaukana hyvistä tilastokäytännöistä. Koko maan indeksi saadaan alueiden mediaanihintojen painottamattomasta keskiarvosta, jolloin suuri kaupunki (township) saa yhtä suuren painon kuin piskuisinkin.

Hinnat kerätään kolmesta satunnaisesta liikkeestä niissä kaupungeissa, joissa on tehty kulutustutkimus. Samaiseen kulutustutkimukseen perustuvat painot lasketaan yhteen – eli siis vain kaupunkilaisten osalta. (Myös Suomen ensimmäinen kulutustutkimus vuonna 1908 koski vain kaupunkien teollisuusväestöä;

maaseudulla elettiin pääosin luontaistaloudessa.)

Muutokset koskevat tältä osin hintojen keruuta ja liikeotosta: uusia keruupaikkoja tulee lisätä suuriin kaupunkeihin ja pienet kaupungit voi jättää pois. Samoin tuotteista tulee kerätä useampia hintoja, nykyisin kerätään vain esimerkiksi kokonaisen kanan hintaa, vaikka kanaa myydään myös pilkottuna ja eri osien hinnat vaihtelevat.

Kaikkia hintoja ei tarvitse kerätä kaikista kaupungeista: autoja myydään vain muutamassa suuremmissa kaupungissa, eikä kaikkialla ole takseja tai riksoja.

Periaatteena noudatettaisiin siis yksinkertaisesti sitä, että kerätään ihmisten kuluttamien hyödykkeiden hintoja sieltä, mistä niitä ostetaan.

Mitä kala maksaa tinkimisen jälkeen?

Tietokoneet ovat jo laskennassa käytössä, mutta luottamus perustaulukkolasken-

taan on vahvaa, joskin osaaminen ei ole kovin korkealla tasolla siinäkin. Perustaulukkolaskennassa on puolensa: kaikki on näkyvissä ja helposti ymmärrettävissä, mutta työ on varsin riskialtista. Tämän kaiken ymmärtää, sillä tietotekniikan läpimurto ei ole vaikeasti saavutettavassa ja sulkeutuneessa maassa vielä tapahtunut.

Toisaalta indeksilaskennassa ollaan varsin lähellä länsimaista peruskäytäntöä. Hinnat kerätään kauppapaikoilta ympäri maata, ja hinnoiksi pyritään ottamaan kuluttajan oikeasti maksama hinta eikä listahintaa. Perinteinen markkinatoiminta tinkimisineen on voimissaan, vain kolmessa suurimmassa kaupungissa on muutama market kiinteillä hinnoilla.

Keskusteluissa tuli ilmi, että hintatietojen kerääjät pyrkivät nyt selvittämään tingittyä hintaa keskustelemalla myyjän kanssa. Myös ns. listahinnan voi kerätä markkinoilta. Kiinteään hintaan myyviestä supermarketista hintakeruu on helpompaa, joskin harvalla paikallisella on varaa asioida niissä.

Hinnat kerätään siis sieltä, mistä kuluttajat tuotteita ostavat, kuten pitääkin. Maaseudun ihmiset, joita suurin osa väestöstä on, jäävät indeksin ulkopuolelle. Hinnat toimitetaan sitten paikalliseen tilastovirastoon, jossa lasketaan keskihinnat ja niistä indeksit.

Valtaosa tuloista ruokaan

Hyödykevalikoima peruselintarvikkeissa on runsasta ja suomalaisesta poikkeavaa. Kulutuksesta yli 60 prosenttia koostuu ruoasta (aivan kuten Suomessakin vuonna 1908). Kaloja ja vihanneksia on hyödykevalikoimassa mukana lukuisia, samoin mausteseoksia.

Nykyinen kuluttajahintaindeksi kattaa varsin vähän muita hyödykkeitä kuin ruokaa. Tähän on uudistusten myötä tulossa muutos, sillä yhteiskunta on Myanmarissa modernistumassa, joskin hitaasti.

Sähköä saa pistorasiasta ja vettä hanasta yksityiskotitalouksiin vain suurimmissa kaupungeissa, eikä julkista liikennettä maan syrjäisemmissä osissa juuri ole. Ihmiset asuvat edelleen bambumajoissa, ja energialähteenä toimii dieselgeneraattori. Niitä on usein vain yksi koko kylällä. Viihdepuolella löytyy hyödykkeeksi elokuvien katsominen, masiina käy ulkoilmassa ja sähkö saadaan generaattorista.

Indeksiin on valikoitunut myös hytysverkko, hätälamppu ja suuri vesiastia, joissa kotien juomavesi säilötään. Vaatepuolelta löytyy länsimaisittain katsoen eksoottinen longyi, hameen tapainen, niin miehille kuin naisillekin omansa.

Jalkine on sandaalit, tavallisiin kenkiin ei niinkään rahaa kulu eikä sukkia näy kaupassa lainkaan. Käytännönläheinen valinta lämpimässä maassa, jossa kengät jätetään huoneen ulkopuolelle, myös töissä. Helppoa siis nähdä, missä huoneessa on kokous menossa.

Työ etenee, ministeriön hyväksyntää odotetaan

Konsulttina pääsee käyttämään osamistaan laajalti. Projektityön käytäntöjä on opeteltu, Gantt-kaavioon on listattu tehtävät ja niille aikataulu visuaaliseen muotoon sekä tekijät ja toteumatilanne.

Tiedostojen yhtenäistä nimeämiskäytäntöä on myös opeteltu.

Paikallinen kulutustutkimus on tehty koskien vain marraskuuta ja ilman kunnan otossuunnittelua, joten myös kulutustutkimuksen ja yleisesti otostutkimuksen perusteiden opastus on kuulunut työn kuvaan.

Kansainväliset standardit täyttävä indeksi on saavutettavissa tekemällä hie-man kaavojen muutoksia ja muutamia tarkistuksia hintakeruukäytäntöihin. Työ edistyy hitaasti, mutta kuitenkin edistyy. Laajassa maassa on laskettava indeksi monelle alueelle ja käsipelillä tehtävillä taulukkolaskelmilla tämä vie aikaa.

Koska projekti on nyt hyvin suunniteltu, puuttuu enää toteutus. Niin, ja sen lisäksi puuttuu edelleen Kansallisen suunnittelu- ja talouskehitysministeriön hyväksyntä uudelle indeksille, joka pitäisi julkistaa jo loppukesästä. ■

Tilastokeskus kouluttaa

Kevään 2015 kurssitarjontaa

MITÄ TILASTOT KERTOVAT TYÖMARKKINOIDEN MUUTOKSISTA?
7.5.2015

TULONJAON KEHITYS SUOMESSA JA TAVAT MITATA TULONJAKOA
27.5.2015

AJANKOHTAISTA ASUNTOMARKKINOILTA 28.5.2015

GLOBALISAATION VAIKUTUKSET TALOUTEEN JA TILASTOINTIIN
9.6.2015

Kurssit järjestetään Tilastokeskuksessa osoitteessa
Työpajankatu 13, 00580 Helsinki.

Kurssien ohjelmat ja hintatiedot löytyvät kotisivuiltamme:
TILASTOKESKUS.FI/KOULUTUSPALVELUT

Lisätietoja: koulutuspalvelut@tilastokeskus.fi
puh. 029 551 2345 tai 029 551 3681

PERUSTIETOJA TILASTOISTA JA NIIDEN
KÄYTTÖMAHDOLLISUUKSISTA –
TUTUSTU TILASTOKOULUUN **TILASTOKOULU.STAT.FI**

Tilastokeskus 

TAB. III.
GENERAL-GOUVERNEURENS i FINNLAND
i underdanighet afgifne
NÖDIGE UNDERRÄTTELSE

ÅR
1751.

Om STOR-FURSTENDÖMETS Underfåares Tilväxt, genom hälsosamma inrättningar, eller Aftagande genom Sjukdomar, Hung-
näringsjång, befunnen efter följande fördelning af Alder, Kön och Gütermål, samt Stånd och Omst.

ÅLDER.		STÅND.	
	MANKÖN.	QVINKÖN.	
Yngre	7106	7810	1. a. Ridderskap och Adel 320. 917
a. 1. och 2. år	14055	14255	b. Ungdom öfver 15. år 121. 299
b. 3. och 4. år	12875	13442	c. Barn under 15. år 234. 306
c. 5. och 6. år	23486	23206	2. a. Prästestapel, Academi och Schola- Stater 737. 610
d. 7. och 8. år	22005	21516	b. Ungdom öfver 15. år 310. 348
e. 9. och 10. år	20264	21615	c. Barn under 15. år 1062. 713
f. 11. och 12. år	19192	21273	3. a. Sålunda- Personer och öfriga och- liga 1021. 1100
g. 13. och 14. år	15392	16327	b. Kruka-Resjörer, Fogdar, Länsmän &c. 441. 407
h. 15. och 16. år	11429	12302	c. Ungdom öfver 15. år 362. 629
i. 17. och 18. år	10119	10824	d. Barn under 15. år 950. 985
k. 19. och 20. år	9576	11162	4. a. Sålunda behörigare Resjörer 258. 240
l. 21. och 22. år	8412	9380	b. Dito Lapsuvar och tjänsfolk öf- ver 15. år 236. 2965
m. 23. och 24. år	9761	10886	c. Dito under 15. år 411. 504
n. 25. och 26. år	5895	6902	5. a. Magistret i Städerna 113. 117
o. 27. och 28. år	5409	7125	b. Grossjörer 10. 11
p. 29. och 30. år	4091	5571	c. Kramare 149. 163
q. 31. och 32. år	3162	3883	d. Hållare och Smidsamare 372. 349
r. 33. och 34. år	1460	2177	e. Besjörer och Ungdom öfver 15. år 228. 176
s. 35. och 36. år	1137	1721	f. Barn under 15. år 338. 407
t. 37. och 38. år	761	1009	6. a. Fäbriqvarer 8. 5
u. öfver 39. år	205635	224227	b. Ungdom och Arbetare öfver 15. år 56. 88
Summa	429912	429912	c. Barn under 15. år 37. 28
GIFTERMÅL			7. a. Ambets- och Handtricks-folk 614. 689
	MANKÖN.	QVINKÖN.	b. Gefäller 242. 95
A. Gift Folk	20345	20505	c. Lärlingar öfver 15. år 320. 74
B. Anklingar och Ankor	4075	19770	d. Dito under 15. år 205. 15
C. Orifte öfver 15. år	4142	43866	8. Ringare, Borgerstap, Fäbriqvar- och Fördels-Kalar 1602. 2220
D. Barn under 15. år	80125	80286	9. Rätters, Seters, Kyrka- och Stads- Betjörer 1103. 1062
Summa cfr. Kön, Summa.	429912	429912	10. a. Frikt inhyrt folk i Städerna 47. 275
			b. Dito fräckligt, dock ej Hofstalls- Hjon 62. 283
			11. Resande och Främmande 15. 17
			12. Slappare och siffant folk 330. 253
			13. a. Handtricks och öfriga Borgerstap- pets, samt Betjörarns barn öfver 15. år 651. 985
			b. Dito under 15. år 2172. 2473
			14. a. Samtliga Borgerstapets och Betjör- arns tjänsfolk öfver 15. år 629. 1367
			b. Dito under 15. år 136. 244
			15. Ställe och mindre Seminarier på Länder 40640. 38161
			16. a. Tjäre med afslä 7179. 3730
			b. Tjäre utan afslä 150
			17. a. Frikt inhyrt och Gatubas-Folk på Länder 180
			b. Dito fräckligt, dock ej Almof-Hjon 310
			18. a. Grottsmän och Svär-Handtricksare 63
			b. Andra Handtricksare på Länder 65
			19. Ryttare och Dragooner, Söldater och Bå- män 929
			20. Svandfjörare och siffant friskt folk, dock ej Hemmansbaser 16
			21. a. Handtricksare vid Rikspolisken 14
			b. Kruksfolk dito 38
			22. Månare 15
			23. a. Menige Almqvists Barn och Träng- folk, öfver 15. år 5933
			b. Dito under 15. år 74356
			24. Fattiga Präst-Ankor 214
			25. Gammalt afskädat Krigsfolk utan Krigs- mansbas 4
			26. a. Färdigen uttagne i Hofstaller 30
			b. Färdigen uttagne i Fattighus 98
			27. a. Eldande utom Hofstaller eller Fattighus 71
			b. Geloe och svfvanige utan Hofstäl 15
			28. Bräddige af följande st och smutt. Sjuk- dom 29
			29. Lägrade Qvinnfolk 9
			30. Rofs- och Spinnas-Folk 67
			31. Lifsbogor på Fästningar och Slätt 2
			32. Främmande Religions-Förvander 2
			Andra handtricksare deras barn 2
			Tjäre 1
			öfriga öfriga 1
			OMSTÄNDIGHE
			I. Hushåll eller Matlag i Städerna, dock undantagne the Gemene af Krigsfolk, alla sfattiga och eländige
			II. Caffehus i Städerna
			III. Källare i Städerna
			IV. Krogar i Städerna
			V. Hushåll, Matlag eller Rökstäl på Länder, undantagne the Gemene af Krigsfolk, alla sfattiga och eländige
			VI. Gästgiverier på Länder
			VII. Krogar på Länder

Suomea koskeva tilastotalukko vuodelta
1751. Ruotsin valtionarkisto.

MAAILMAN ENSIMMÄINEN VÄESTÖTILASTO

Vuonna 1750 Ruotsi-Suomessa aloitettiin väestötilasto. Se on maailman vanhin vuosittain kerätty väestötilasto. Aluksi tietoja oli väestönmäärästä, syntyneistä, kuolleista ja avioliiton solmineista. Vähin erin tilastot ovat täydentyneet monipuolisiksi tietolähteiksi.

Väestötietojen keruu yleistyi Euroopan kaupungeissa jo keskiajan loppupuolella. Tosin tätä ennen löytyy jo mainintoja väestöluetteloiden pidosta aina ennen ajanlaskun alkua. Synnä väestötietojen keruuseen oli verotus. Kuuluisin verollepanosta johtunut väestölaskenta lienee ajoittunut ajanlaskumme alkuun Palestiinan alueelle.

Väestötietojen keruu yleistyi katolisen kirkon piirissä, kun vuonna 1563 Tridentin kirkolliskokouksessa määrättiin luetteloiden pitämisestä kastetuista, haudatuista ja avioliiton solmineista. Kiinnostus varsinkin kuolleiden lukumäärän tutkimiseen kasvoi Euroopassa jo 1600-luvulla. Tautiepidemioiden leviämistä tutkittiin haudattujen luetteloiden avulla ja ensimmäiset tilastollista elinajanodotetta käsittelevät taulukot – Bills of Mortality – julkaistiin Lontoossa 1600-luvun puolessa välissä (Nieminen 1999).

Maallinen ja hengellinen valta luetteloi kansan

Ruotsi-Suomessa ensimmäiset väestöluettelot ovat 1500-luvulta. Ne syntyivät, kun Kustaa Vaasan toimesta ns. maakirjoihin merkittiin luettelo talonpojista ja heidän maksamistaan veroista. Näihin maakirjoihin kerättiin myöhemmin myös tietoja asekuutoisista miehistä sekä yli 12-vuotiaasta väestöstä. Henkikirjoi-

er och Wankötsel, samt brott på nyttige ändigheter.

N.	QVINKON.
0.	106.
8.	3832.
6.	1105.
2.	537.
7.	524.
0.	7551.
3.	97.
1.	125.
7.	378.
5.	141.
24.	60743.
6.	74565.
200.	200.
29.	29.
4.	4.
6.	1224.
2751.	2751.
120.	120.
141.	141.
2178.	2178.
26.	26.
6.	6.
620.	620.
37.	37.
6.	6.
5.	5.
17.	17.

ANMÄRKNINGAR.

I. GENERAL-GOUVERNEUREN i FINNLAND adderar på thenne sin 3:de Tabell, ruta från ruta, alla LANDSHÖFDINGARNES Sittre-Summor, uti sin Tabell öfver hela STOR-FURSTENDÖMET: Utaf hvilken äfven som af the 2:ne förra, et Ark ärligen inländes, in medio Aprilis, til Vårt CANCELLEI-COLLEGIUM, sedan en lika lydande Afkrift uti STOR-FURSTENDÖMETS Tabell-Bok blifvit aftagen, som af GENERAL-GOUVERNEUREN och thes efterträdare förvaras.

II. Hela befväret, at förfärdiga så väl alla 3. Kyrko-Tabellerne, som then Methodiska, hvilken endast tjenar til thes beqvämliga uprättande för PASTORES, består blott therutinnan, at man drager några få strek, och adderar the samma. Sättes at vinna thetta så nödige ändemål, både til GUDS ÄRAS utvidgande, genom ovedertäjelige härat flytande och beviste Grund-Sanningar, sålöm ock til Rikets sanskylliga välstånd och styrka, är jämväl med sitt gjordt så lätt och enfaldigt, at ock et Barn, som alenast lärde addera någon liten Summa, kan nyttjas til at afluta alla thesse Tabeller, utur hvilka thenna så nödige kundskapen, til alla delar, endast kan framledas, hvilken genom vittra Utlänningars rön, och mer än 100. års stadgad utöfning, gifvit anledning til at bringa the mäktigaste Riken i Europa til märkeligit anseende, samt tillskyndat thes underfattare mång befynnerlig lifa och osörmodelige förmoner, jämte tidige hjälpe-medel uti Epidemiske och Chroniske sjukdomar.

Stockholm d. 1. Augusti 1753.

J. F. V. Rosen

T E R.

44 5/1.

42517.

Jacob Beckrook

tuksen tärkeimpänä tehtävänä oli määrittellä henkilön kotipaikka ja verotuskunta. Henkikirjoitus suoritettiin vuosittain aina vuoteen 1989 saakka, jolloin väestötietojärjestelmä oli jo niin luotettava, ettei vuosittaista tarkistusta tarvittu.

Myös kirkon toimesta alettiin 1600-luvun alkupuolella kiinnostua väestötiedoista ja useissa seurakunnissa aloitettiin luetteloiden pitäminen kastetuista, haudatuista ja avioliiton solmineista hallinnollisia tarpeita varten.

Väestötilaston kehittämisen kannalta merkittävin uudistus oli vuoden 1686 kirkkolaki, jossa kaikki seurakunnat veloitettiin pitämään edellä mainittujen luetteloiden lisäksi ns. rippikirjaa, johon merkittiin seurakuntalaisten kristinopin tuntemus ja lukutaito. Lisäksi näin muodostuneisiin kirkonkirjoihin merkittiin tietoja ”mistä tullut, miten käyttäytynyt ja mihin mennyt”. Tämän lisäksi niihin merkittiin tiedot henkilön säädystä ja kuolemansyystä. Näin kirkonkirjat loivat arvokkaan pohjan myöhemmälle historian tutkimukselle sekä tulevalle väestötilastolle (Tilastollinen Päätoimisto 1948; Nieminen 1999).

Pisimmät sarjat väestönmuutoksista

Kuten muualla Euroopassa myös Ruotsi-Suomessa oltiin kiinnostuneita väestöilmöiden tutkimisesta, ja 1700-luvun alkupuolella muutamat papit tekivät yhteenvetoja kastettujen luetteloista. Ensimmäinen koko maata koskeva selvitys kastettujen ja haudattujen luetteloista tehtiin vuonna 1736, mutta se jäi puutteelliseksi.

Ruotsi-Suomessa oli jo 1700-luvun alkupuolella koko maan kattava kirkon toimesta ylläpidettävä väestörekisteröinti, ja se sai monet niin maallisen vallan kuin papistonkin edustajat ehdottamaan koko maan kattavan tilaston laatimista väestön määrästä sekä syntyneiden, kuolleiden ja avioliittojen määristä.

Vuonna 1748 valtiopäivät päättivät väestötilastojen laadinnan aloittamisesta ja tehtävä annettiin erityiselle Tukholmassa toimivalle taulustotoimikunnalle (Tabellverket). Ensimmäiset väestötiedot kerättiin vuodelta 1749 ja sen jälkeen väestötilastot on kerätty vuosittain. Näin Ruotsissa ja Suomessa on tällä hetkellä käytössä maailman pisim-

mät vuosittain kerätyt tiedot niin väkiluvusta kuin väestönmuutoksistakin, jotka perustuvat todellisiin väestöluetteluihin (Tilastollinen Päätoimisto 1948).

Tiedot kerättiin kirkonkirjoista ja papiston tehtäväksi jäi tilastolomakkeiden täyttäminen. Tilastolomakkeita oli kolme joista yhdellä kerättiin tiedot väkiluvusta ja ikärakenteesta ja kahdella muulla kerättiin tiedot väestönmuutoksista, säädystä ja kuolemansyistä.

Tietojen keruuseen liittyi alusta alkaen puutteita, jotka vaikuttivat tietojen luotettavuuteen, mutta niitä pyrittiin ajan myötä korjaamaan. Esimerkiksi syntyneiden ja kuolleiden lukumäärät kerättiin aluksi kastetuista ja haudatuista. Näistä tuli tietenkin virheitä ti-

lastoon, mutta jo vuonna 1775 pappeja pyydettiin ilmoittamaan tiedot syntyneistä ja kuolleista.

Väestötilaston laadinta ei keskeytynyt, kun Suomi vuonna 1809 liitettiin keisarillisen Venäjän suuriruhtinaskunnaksi. Porvoon valtiopäivät antoivat väestötilaston yhteenvedon laadinnan Senaatin Kamaritoimikunnalle (Tilastollinen Päätoimisto 1948; Nieminen 1999).

Tilastollinen Päätoimisto perustetaan

Vuonna 1865 eri puolilla valtionhallintoa olevien tilastojen keruu ja laadinta keskitettiin perustetulle tilastovirastolle, josta myöhemmin kehittyi Tilastolli-

VÄESTÖTILASTON LAADINTA EI KESKEYTYNYT, KUN SUOMI VUONNA 1809 LIITETTIIN KEISARILLISEN VENÄJÄN SUURIRUHTINASKUNNANKSI.



Tilastollinen päätoimisto vuonna 1906. Tilastokeskuksen arkisto.

nen Päätoimisto (Tilastokeskus vuodesta 1971 alkaen). Sadan vuoden aikana väestötilaston sisältöä oli laajennettu yhteiskunnan tarpeita vastaavaksi ja niistä oli tehty monia uudistusehdotuksia. Viraston perustaminen loi uuden pohjan tilastojen kehittämiseksi.

Uudessa tilastovirastossa alettiin julkaista säännöllisesti Suomen virallisen tilaston (SVT) sarjaa ja väestötilastot julkaistiin sarjassa VI. Ensimmäinen virallinen väestötilastojulkaisu oli ”Suomen väestö Joulukuun 31 p. 1865”. Sen jälkeen alkoikin säännöllinen väestötilaston julkaiseminen. Tosin usein vuosijulkaisut sisälsivät useamman vuoden väestötilastot samassa julkaisussa. Sarja VI vastaa nykyistä SVT:n Väestö-sarjaa.

Henkilömakepohjaiset laskennat käynnistyivät vuonna 1870

Aina Tilastollisen Päätoimiston perustamisesta alkaen yhteiskunnan tietotarve

lisääntyi vuosi vuodelta. Ripeä teollistuminen 1870-luvulla alkoi muuttaa yhteiskuntaa, ja sen myötä myös tilastojen tarve yhteiskunnan kuvaajana lisääntyi.

Yli sata vuotta oli kirkonkirjoista laadittu väestötilastoa, mutta jatkuvasti nousi esiin seurakuntien täyttämien tilastolomakkeiden puutteellisuudet, eikä aina oltu tyytyväisiä väkilukua kuvaavien tietojen tarkkuuteen. Kirkonkirjoissa pidettiin usein henkilöitä, jotka olivat muuttaneet seurakunnista pois, ja näin annetut tiedot olivat epätarkkoja.

Kirkonkirjoista lasketut väkiluvut hahutettiin korvata koko maassa suoritettavalla väestölaskennalla, jossa jokaisesta henkilöstä papit täyttäsivät tilastolomakkeen. Uudistus kariutui koko maan tasolla papiston vastustukseen.

Kansainvälinen yhteistyö tilastoalalla alkoi 1800-luvun loppupuolella lisääntyä, ja monissa maissa olikin suoritettu henkilömakepohjainen väestölaskenta. Suomessa jo vuonna 1870 tehtiin ensimmäiset henkilömakeksiin perustuvat laskennat maamme suurimmissa kaupungeissa. Kunkin kaupungin jokainen ruokakunta täytti tilastolomakkeen. Kymmenvuosittaisten väkilukutietojen lisäksi kaupunkien väestölaskenta sisälsi asumista ja ruokakuntia koskevaa tietoa, jota normaaleilla tilastolomakkeilla ei saatu. Tilasto sisälsi tietoa äidinkielestä, sivistystasosta (=lukutaito), uskontokunnasta, ammatista ja elinkeinosta. Lisäksi kerättiin tietoja sokeista, kuuroista ja mielisairaista. Aluksi mukana olivat kaupungeista Helsinki, Turku, Viipuri ja Oulu, mutta myöhemmin kaupunkien lukumäärä kasvoi.

Kaupunkien henkilömakeksiin perustuvat väestölaskennat tehtiin joka kymmenes vuosi aina vuoteen 1930 saakka, ja niistä laadittiin erilliset tilastojulkaisut.

Väestölaskentalaki säädettiin vuonna 1938

Aina itsenäistymisen alusta lähtien oli yritetty henkilömakepohjaisen väes-

tölaskennan suorittamista koko maassa, mutta aina se tyrehtyi rahapulaan. 1930-luvulla koko maan kattavat henkilömakepohjaiset väestölaskennat oli tehty Euroopan maista kaikkialla muualla paitsi Albaniassa ja Suomessa.

Henkilömakepohjaisen väestölaskennan toteuttamiseksi annettiin vuonna 1938 väestölaskentalaki (154/38). Siinä todettiin, että väestölaskenta suoritetaan joka kymmenes vuosi. Väkilukutietojen lisäksi väestölaskennassa kerättäisiin tietoja perheiden ja ruokakuntien määrän lisäksi asunnoista, rakennuksista ja kiinteistöistä.

Väestö- ja asuntolaskenta (VAL) oli määrä toteuttaa vuonna 1940, mutta syntyneen sodan takia se jäi toteuttamatta. Heti sodan jälkeen YK laati kansainväliset suositukset eri maissa toteutettaville väestölaskennoille, ja niin Suomessakin suoritettiin 31.12.1950 ensimmäinen väestölaskenta, jossa koko maan väestöltä kerättiin tilastolomakkeella tietoja.

Seuraavat henkilömakepohjaiset väestölaskennat suoritettiin vuosina 1960, 1970 ja 1980. Tämän lisäksi vuonna 1975 tehtiin niin sanottu asunto- ja elinkeinotutkimus (AET), joka tietosisällöltään vastasi varsinaista väestölaskentaa. Vuonna 1985 tehtiin vielä osittain perinteinen väestölaskenta. Enää ei kuitenkaan ollut käytössä raskasta alueittaista väestölaskentaorganisaatiota.

Kuolinsyytodistukset paransivat tilastointia

Kuolemansyitä on tilastoitu aina väestötilaston laatimisen alkamisesta lähtien vuonna 1749. Kuolemansyytiedot olivat pitkään ehkä kaikkein puutteellisin väestötilasto, sillä papit joutuivat merkitsemään kirkonkirjoihin usein vain omaisten ilmoittaman kuolemansyyn. Näin ryhmä ”muut kuolemansyyt” olivat pitkään suurin kuolemansyyryhmä.

Vuosien mittaan kuolemansyyloki- tusta uudistettiin, ja tautien määrittelys- ➤

RIPEÄ TEOLLISTUMINEN 1870-LUVULLA ALKOI MUUTTA YHTEISKUNTA, JA MYÖS TILASTOJEN TARVE YHTEISKUNNAN KUVAAJANA LISÄÄNTYI.





Väestönlaskentaa vuonna 1960. Tilastokeskuksen arkisto.

HENKILÖTUNNUKSEN KÄYTTÖNOTTO ANTOI TILASTONLAATIJOILLE JA TUTKIJOILLE MAHDOLLISUUDEN HYÖDYNTÄÄ VÄESTÖTIETOJA SEKÄ MUITA HENKILÖÄ KOSKEVIA SOSIOEKONOMISIA TIETOJA.

tä annettiin ohjeita. Suurin uudistus tapahtui vasta vuonna 1936, jolloin otettiin käyttöön kuolintodistus. Tilastoa alettiin laatia suoraan tältä lomakkeelta, ja 200-vuotinen käytäntö kirkonkirjoihin perustuvan kuolemansyyn tilastoinnista lopetettiin (Pitkänen 1999).

Tällä hetkellä kuolemansyytilasto vastaa täysin kansainvälistä huipputasoa perustanaan uusin WHO:n hyväksymä tauluokitus.

Henkilötunnus koko kansalle – väestörekisteri perustetaan

Väestökirjanpidon uudistamiseksi oli vuosien mittaan tehty uudistusehdotuksia, sillä väestötietojen rekisteröinti oli jo yli 200 vuoden ajan ollut osittain päällekkäinen. Uskonnolliset yhdyskunnat pitivät omia rekistereitään, joista tiedot saatiin väestötilastoa varten. Toisaalta joka vuosi pidettiin henkikirjoitus, jossa määriteltiin henkilöiden virallinen asuinpaikka. Henkikirjoitusviranomaiset pitivät omaa rekisteriään (siviilirekisteri) niistä henkilöistä, jotka eivät kuuluneet mihinkään uskonnolliseen yhdyskuntaan. Siviilirekisteriin kuuluvista henkilöistä toimitettiin tiedot väestötilastoa varten.

Vuonna 1964 otettiin käyttöön Kansaneläkelaitoksen sosiaaliturvarekisteri ja kaikki Suomessa vakituaisesti asuvat henkilöt saivat sosiaaliturvatunnuksen – nykyisen henkilötunnuksen.

Väestörekisteröinnin uudistamista selvittävä toimikunta ehdotti vuonna

1966 väestön henkilörekisterin perustamista ja sille omaa virastoa Väestörekisterikeskusta (VRK), joka aloitti toimintansa vuonna 1969. Vuosien mittaan sinne kehittyi laaja väestötietojärjestelmä niin väestöstä, asunnoista, rakennuksista kuin kiinteistöistäkin (Väestörekisterikomitean mietintö 1967).

Väestötietojärjestelmän myötä perustettiin perhetilasto

Vuoden 1970 väestönlaskennassa oli jo käytössä henkilötunnus, ja myös väestötilaston käsittelyssä otettiin käyttöön henkilötunnus. Tämä merkitsi väestömuutosaineiston yhdistelyn mahdollisuutta kaikkiin niihin aineistoihin, joissa oli käytössä henkilötunnus.

Henkilötunnuksen käyttöönotto antoi tilastonlaatiijoille ja tutkijoille mahdollisuuden hyödyntää väestötietoja sekä muita henkilöä koskevia sosioekonomisia tietoja. Erillisten paperisten lomakkeiden täyttämistä paikallisrekistereissä väestötilastoja varten luovuttiin, ja vuodesta 1976 alkaen kaikki väestötiedot saatiin sähköisesti VRK:n väestötietojärjestelmästä.

Samalla väestötilaston sisältöä laajennettiin aloittamalla perhetilaston laadinta. Tämän mahdollisti VRK:n väestörekisterin tietojen sisältämät tiedot sukulaisuussuhteista.

Ensimmäinen laajempi tutkimus, jossa väestötilaston ja väestönlaskennan sosioekonomisia tietoja käytettiin hyväk-

si, oli elinoloja ja kuolemansyitä koskeva tutkimus (EKSY70). Siinä vuoden 1970 väestönlaskennan tiedot yhdistettiin vuosina 1971–75 kuolleiden tietojen kanssa ja tutkittiin muun muassa eri ammattiryhmien välistä kuolevuutta. Tämä tutkimus olikin alkuna Tilastokeskuksen laajalle yhteistyölle ensin Helsingin yliopiston väestötutkijoiden ja myöhemmin myös muiden tutkijoiden kanssa.

Myöhemmin laajimpia väestötilastossa tehtyjä tutkimuksia oli kaksi otantatutkimusta, joissa toisessa selvitettiin maassamuuton syitä 1970-luvun lopulla ja toisessa osallistuttiin laajaan YK:n Euroopan talouskomission (ECE) organisoimaan kansainväliseen hedelmällisyystutkimukseen 1980-luvun lopulla. Molemmat tutkimukset olivat haastattelututkimuksia (Nieminen 1999).

Rekisteripohjainen väestönlaskenta tehtiin vuonna 1990

Viranomaisilla oli 1980-luvulla käytössään kattavat rekisterit, joita alettiin hyödyntää myös tilastojen teossa. Samalla yksityisille ihmisille lähetettävät kyselyt vähenivät.

Vuonna 1990 tilastontekijöiden pitkäaikainen unelma toteutettiin, kun väestönlaskenta suoritettiin pelkästään ns. rekisterilaskentana, johon tiedot kerättiin eri viranomaisten rekistereistä. Tärkeimmät rekisterit olivat VRK:n väestötietojärjestelmä (väestörekisteri), verohallinnon rekisterit, eläketurvakeskuk-

sen työsuhderekisteri, valtiokonttorin rekisterit, kuntien eläkelaitoksen rekisterit, kansaneläkelaitoksen eläkerekisteri, työministeriön työnhakijarekisteri sekä pääesikunnan varusmiesrekisteri. Lisäksi Tilastokeskuksella oli käytössään oma väestön tutkintorekisteri ja yritys- ja toimipaikkarekisteri. Rekisterien yhdistely mahdollisti kaikkien perinteisten väestölaskentatietojen vuosittaisen laadinnan.

Väestöennuste huolen ja kritiikin aiheena

Tilastollinen Päätoimisto julkaisi vuonna 1934 ensimmäisen koko maata koskevan väestöennusteen. Väestöennuste laskettiin menneen kehityksen perusteella käyttäen apuna ikäryhmittäisiä hedelmällisyys- ja kuolevuuslukuja (Modeen 1934).

Väestöennuste laskettiin vuoteen 1980 saakka ja sen tulos – väkiluku ei saavuta koskaan neljää miljoonaa – herätti sen ajan päättäjissä huolta. Muutamaa vuotta myöhemmin laskettiin uusi väestöennuste samalla menetelmällä ja tulos oli lähes sama. Neljä miljoonaa ylitettiin, mutta sen jälkeen väkiluku alkasi laskea (Modeen & Fougstedt 1939). Kuitenkin toisen maailmansodan syttyminen muutti kaiken aikaisemmin ennustetun kehityksen.

1950-luvulla alettiin yhä enemmän kiinnittää huomiota tulevan väestönkehityksen merkitykseen suunnittelun ja päätöksenteon pohjana (Komiteanmietintö 1956:6). Ensimmäinen alueittainen väestöennuste laadittiin väestötilastossa kuitenkin vasta vuonna 1964 ja alueyksikkönä käytettiin tilastoalueita. Tämän väestöennusteen pohjalta Valtakunnansuunnittelutoimisto julkaisi kunnittaisen väestöennusteen samana vuonna (Nieminen 1999).

Kunnittaisen väestöennusteen säännöllinen laadinta alkoi vuonna 1969. Nykyisin kunnittainen väestöennuste uusitaan joka kolmas vuosi.

Kunnittainen väestöennuste on tyyppillinen trendiennuste, joka lasketaan ikäryhmittäin ottamalla huomioon syntyvyyden, kuolevuuden, maan sisäisen muuttoliikkeen ja siirtolaisuuden mennyt kehitys ja arvioimalla väestömuutosten tulevaa kehitystä. Varsinkin kuntapäätäjien parissa väestöennusteen tulokset ovat

usein saaneet kritiikkiä, mutta ennuste on kuitenkin täyttänyt useilla suunnittelun aloilla sille vaaditun tehtävän erityisesti väestön ikärakenteen muuttumisen ennustamisesta ja maassamuuton suuntautumisesta asutuskeskuksiin.

Vuonna 1983 julkaistiin myös Suomen ruotsinkielisen väestön ennuste. Tämä ennuste oli tyyppillinen trendiennuste, jossa mennyttä kehitystä projisoitiin tulevaisuuteen ottamalla huomioon ruotsinkielisen väestön ikä- ja sukupuolirakenne (Nieminen 1983).

Maastamuuttomaasta maahanmuuttomaaksi

1800-luvun loppupuolella siirtolaisuus Amerikkaan alkoi saada sellaiset mittasuhteet, että myös väestötilaston laatijat kiinnittivät tähän huomiota. Kirkonkirjoista ei tietoja toimitettu tilaston laadintaa varten, ja näin useiden kuntien väkiluvut jäivät liian suuriksi.

Ongelma oli erityisen suuri Pohjanmaan kunnissa, joista tiedettiin siirtolaisuuden 1880-luvulla voimistuneen. Tällöin tehtiin aloite, että tiedot siirtolaisiksi lähteivistä annettiin lääninhallituksen passiluetteloista tilastoa varten.

Siirtolaisuustilasto koski aluksi vain Vaasan ja Oulun läänenejä, mutta vuonna 1893 tilasto laajeni koko maan kattavaksi. Lisäksi siirtolaisuustilasto laajennettiin käsittämään myös muut Euroopan ulkopuoliset maat kuin Pohjois-Amerikan.

Passiluetteloista laadittu siirtolaisuustilasto jatkui muista väestömuutoksista erillisenä tilastona aina 1980-luvun puoleen väliin saakka, jolloin se lopetettiin. Lähes sadan vuoden ajan tilastoa tehtiin erillisenä eikä se vaikuttanut mitenkään esimerkiksi kuntien väkilukuihin. Varsinkin 1954 voimaan astunut yhteispuolisoimainen työmarkkinasopimus heikensi passiluetteloista laaditun tilaston laatua, kun muuttajilta ei vaadittu passia Pohjoismaihin.

Siirtolaisuustilastoon liittyi jatkuvasti puutteita ja sen luotettavuus oli kyseenalainen. Monet siirtolaisuustutkijat ovatkin tutkimuksissaan täydentäneet siirtolaisuustilaston tietoja käyttämällä mm. laivayhtiöiden laatimia luetteloita matkustajista.

Kun maahanmuutto Suomeen alkoi 1980-luvun lopulla kasvaa, ja tulevi-

na vuosina oli näköpiirissä ikärakenteen vanhenemisesta johtuva työvoimapula, päätettiin aloittaa erityinen siirtolaisuutta koskeva tilasto. Tärkeimmät maahan- ja maastamuuttotilastot oli julkaistu väestömuutoksia koskevan julkaisun yhteydessä. Ensimmäinen ”Ulkomaalaiset ja siirtolaisuus”-julkaisu laadittiin vuodelta 1994 ja sen jälkeen sitä on julkaistu joka vuosi. Julkaisuun kerättiin keskeiset tiedot maahan- ja maastamuutosta sekä tietoja Suomessa vakituisesti asuvista ulkomaan kansalaisista. Julkaisu sisälsi vuositilaston lisäksi pitempiä sarjoja siirtolaisten ja ulkomaalaisten demografisista ominaisuuksista (Tilastokeskus). ■

Kirjoittaja on eläkkeellä oleva Tilastokeskuksen entinen kehittämisspääällikkö.

Lähteet:

Komiteanmietintö 1956:6. Suomen yhteiskuntatilasto ja sen kehittäminen.

Modeen, Gunnar 1934. Suomen väkiluvun vastainen kehitys. Tilastokatsauksia 1934:10. Helsinki.

Modeen, Gunnar & Fougstedt, Gunnar 1939. Laskelmia Suomen vastaisesta väestönkehityksestä. SVT VI:89. Helsinki.

Nieminen, Mauri 1983. Suomen ruotsinkielinen väestö vuosina 1980–2020. Tutkimuksia Nro 98. Helsinki: Tilastokeskus.

Nieminen, Mauri 1999. Väestötilastoja 250 vuotta. Väestö 1999:8. Helsinki: Tilastokeskus.

Pitkänen, Kari 1999. Kuolemansyyt. Teoksessa: Kristiina Andreasson & Vesa Helin (toim.). Suomen vuosisata. Helsinki: Tilastokeskus.

Tilastollinen Päätoimisto 1948. Taulustolaitos 1748–1948.

Tilastokeskus. Ulkomaalaiset ja siirtolaisuus. Väestö-sarja 1995–.

Väestörekisterikomitean mietintö 1966:A15. Helsinki.

EEVA HAMUNEN

Kansantalouden tilinpitoa on kehitetty etenkin kriisiaikoina

Nykyinen kansantalouden tilinpidon järjestelmä on kehittynyt pääosin 1900-luvun jälkipuoliskolla, mutta sen juuret ovat vuosisatojen takaisessa kansantuloestimaattien arvioinnissa. Kansantalouden tilinpitoa pidetään eräänä 1900-luvun merkittävimmistä yhteiskuntatieteen menetelmäkeksinnöistä. Tilinpidolla on vahva kytkentä talouspolitiikkaan. On sanottu, että kansantalouden tilinpito on rokote talouspoliittista toiveajattelua vastaan.

Maailmanlaajuisen finanssikriisin puhjetessa viimeisin kansantalouden tilinpidon suositus (System of National Accounts, SNA2008) oli juuri saatu valmiiksi. Suositusta esiteltäessä kysyttiin, pystyykö tilinpito kuvaamaan maailmantaloutta kohdannutta kriisiä. SNA2008:n toimittajan Anne Harrisonin vastaus oli, että SNA pystyy siihen edellyttäen, että tilijärjestelmää osataan soveltaa ja tarkastella kokonaisuudessaan.

Kansantalouden tilinpidon laadintaa voidaan verrata palapelin kokoamiseen: kaikkien palojen on sovittava toisiinsa ja kuva täydentyy, mitä enemmän paloja saadaan yhdisteltyä. Kansantalouden tilinpito on makrotilasto, ”tilastojen tilasto”; sen laadinnassa käytetään lähes kaikkia muita taloustilastoja ja useita sosiaalitulastoja. Tilinpidon luonteeseen kahden- tai nelinkertaisena kirjanpitoa kuuluu, että sitä pystytään laatimaan, vaikka kaikista siinä käsiteltävistä ilmiöistä ei olekaan lähdetilastoa.

Tilinpidon esi-isät pohtivat verotuksen kohdentamista

Kansantalouden tilinpidon juuret ovat 1600-luvulla: englantilainen lääkäri William Petty teki arvion Englannin kansantalosta. Laskelmat julkaistiin vuonna 1690 kirjassa *Political Arithmetick*. Poliittista aritmetiikkaa pidetään yhtenä nykyisen tilastotieteen lähtökohtana. Laskelmat tehtiin verotuksen tasapainoista kohdentamista silmällä pitäen. Gregory King vertasi Englannin taloutta Ranskan ja Hollannin talouteen vuosilta 1688 ja 1695 tehdyissä laskelmissa. King loi edelleenkin käytössä olevat tuotannon ja tulon käsitteet. Ranskassa varhaisia kansantuloelaskelmia verotuksen uudistamista varten tekivät Pierre de Boisguillebert ja Sebastien de Vauban.

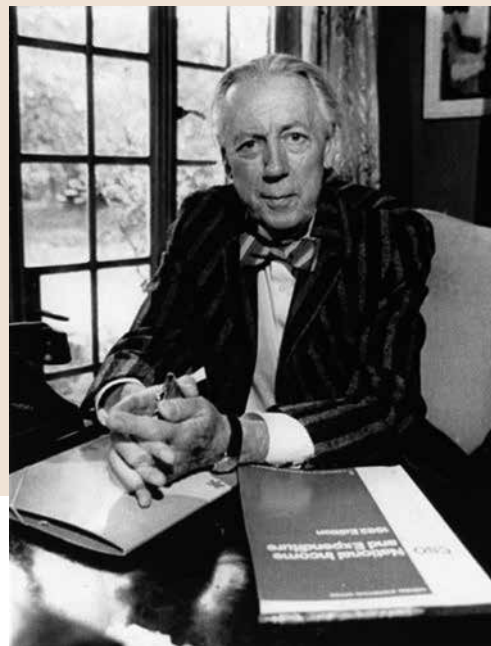
Ranskalainen lääkäri ja fysiokraatti Francois Quesnay laati vuonna 1758 *Tableau Economiquen*, jossa kuvataan talouden kiertokulkua. Idean perustana oli venkiertojärjestelmä. Quesnay toi myös esiin tuottavan toiminnan käsitteen.

Yksittäiset tutkijat tekivät 1700-luvulta eteenpäin kansantuloelaskelmia satunnaisesti mm. Venäjällä, Hollannissa ja Saksassa. 1800-luvun puolivälissä laskelmia tehtiin myös Yhdysvalloissa. Yksittäisten tutkijoiden tuloksia ei välttämättä hyväksytty, niitä pidettiin jopa harhaoppisina. (Varhaisesta historiasta ks. esim. Bos 2009; Vanoli 2005; Aulin-Ahmavaara 2007.)

Vuoden 1929 talouskriisi lisäsi kansantuloelaskelmien kysyntää

Kansantuloestimaattien laatiminen oli systemaattisempaa 1900-luvulta lähtien. Yhdysvaltain kansallinen taloustutkimusvirasto NBER julkaisi ensimmäisen maailmansodan jälkeen kansantuloelaskelmat vuosilta 1909–1919. Vuoden 1929 talouskriisi ja lama merkitsivät selvää käännekohtaa kansantuloelaskelmien kysynnässä. Simon Kuznetsin johdolla laadittiin Yhdysvaltain senaattile vuonna 1934 esitellyt laskelmat vuo-

Monet kansantalouden tilinpidon laadintaan myötävaikuttaneet taloustieteilijät on palkittu taloustieteen Nobelpalkinnolla. Tärkein heistä on Richard Stone, joka palkittiin vuonna 1984 työstään kansantalouden tilinpidon kehittämisessä. Norjalainen Ragnar Frisch sai vuonna 1969 yhdessä hollantilaisen Jan Tinbergenin kanssa ensimmäisen taloustieteen Nobelin ekonometristen mallien kehittämisestä. Simon Kuznets sai palkinnon vuonna 1971 taloudellisen kasvun empiirisestä tutkimuksesta, James Meade vuonna 1977 kansainvälistä kauppaa ja kansainvälisiä pääomaliikkeitä koskevasta tutkimuksesta (yhdessä Bertil Olinin kanssa) ja Wassily Leontief vuonna 1973 panos-tuotos-menettelyn kehittämisestä.



Richard Stone loi perustan nykyaikaiselle kansantalouden tilinpidolle.

silta 1929–1932. Kuznets kehitti termin GNP (Gross National Product).

Hollannissa julkaistiin ensimmäiset viralliset kansantuloestimaatit vuonna 1933 Jan Tinbergenin johdolla. Ruotsissa laadittiin Erik Lindahlin ohjaamana Tukholman yliopistossa pitkät aikasarjat vuosilta 1861–1930. Tanskassa Viggo Kampmann laski Ruotsin laskelmissa käytetyn menetelmän pohjalta panos-tuotostaulukot vuosilta 1930–1939. (Vanoli 2005.) Norjassa Ragnar Frisch kehitti taloudellisen kiertokulun mallin 1930-luvulla ja antoi perustan useiden tilinpidon keskeisten käsitteiden määrittelylle (Eco circulation system). (Fløttum 2012.) Englannissa Cambridgen yliopistossa Colin Clark julkaisi vuonna 1932 omat estimaattinsa vuosilta 1924–1931 ja täydensi niitä myöhemmin. Nämä laskelmat koskivat nykyisen kansantalouden tilinpidon keskeisimpiä estimaatteja: tuloja, tuotosta, kulutusmenoja, valtion tuloja ja menoja, pääoman muodostusta, säästämistä, ulkomaankauppaa ja vaihtotaset-

ta. (Vanoli 2005, 19; Stone 1984, 121.)

Wassili Leontief kehitti panos-tuotos -systeemiä irrallaan em. kansantulolaskelmista 1930-luvulla Yhdysvalloissa, Harvardin yliopistossa. U.S. Bureau of Labour Statistics jatkoi työtä Leontiefin johdolla ja vuonna 1945 julkaistiin panos-tuotostaulukot vuodelta 1939. (Vanoli 2005, 23.)

Keynes korosti julkisen sektorin mittaamisen tärkeyttä

Ensimmäisen maailmansodan jälkeen 1920 rauhanneuvotteluihin osallistunut John Maynard Keynes julkaisi poleemisen kirjoituksensa *The economic consequences of the peace*, jossa hän arvioi eri maiden tilastotietojen perusteella, että Saksa ei tule selviämään sille langetetuista ankarista sotakorvauksista (Bos 2009, 20). Keynesin *Yleisessä teoriassa* vuodelta 1936 määritellyt makrotalouden keskeiset käsitteet kuten tulot, kulutus ja investoinnit edellyttivät näiden käsitteiden mittaamisme-

**KANSANTALouden
TILINPIDON JUURET OVAT
1600-LUVULLA.**

1930-LUVUN TALOUSLAMA TOI VALTION ROOLIN AIEMPAA NÄKYVÄMMÄKSI.

netelmien kehittämistä. 1930-luvun talouslama toi myös valtion roolin aiempaa näkyvämmäksi, ja talouden analyysi edellytti julkisen sektorin tulojen ja menojen mittaamista. Kansantalouden tilinpidon kaltaisen tilastovälineen tarve talousanalyysiä varten kasvoi.

Toinen maailmansota loi tarpeen arvioida sodan aiheuttamia kustannuksia kansakunnalle. Keynes toimi Britannian hallituksen neuvonantajana ja julkaisi alkuvuodesta 1940 kirjoituksensa *How to pay for the war* käyttäen hyväkseen Colin Clarkin laskelmia (Skidelsky 2005).

Britannian sotakabinetissa työskennellyt James Meade laati Colin Clarkin laskelmien pohjalta uudet kansantuloestimaatit ja yhdessä Richard Stonen kanssa he laativat raportin *National Income, Saving and Consumption*. Keynes esitteli paperin valtiovarainministerille ja ministeriön budjettikomitealle, ja se julkaisiin huhtikuussa 1941 budjetin liitteenä.

Meade ja Stone kehittivät edelleen estimaattien taulukointia, ja tulos julkaistiin *Economic Journal*issa kesällä 1941 ensimmäisenä nykyisen kansantalouden tilinpidon elementit sisältävänä järjestelmänä. (Skidelsky 2005; Vanoli 2005.)

Toisen maailmansodan jälkeen käynnistyi tilinpidon kansainvälinen kehittäminen

Richard Stone palkittiin vuonna 1984 taloustieteen Nobelilla kansantalouden tilinpidon kehittämisestä. Juuri Stone loi perustan nykyaikaiselle kansantalouden tilinpidolle.

Toisen maailmansodan jälkeen kansantalouksien toiminnan analyysiin tarvittavan tilastovälineen tarve tuli tunnistetuksi. Kaivattiin myös kansainvälisesti vertailukelpoisia tietoja. YK julkaisi vuonna 1947 ensimmäisen raportin kansantalouden tilinpidon suositukseksi. Raportissa oli Stonen laatima yksityiskohtainen ja varsin kunnianhimoinen kansantalouden tilinpidon järjestelmä.

Toisen maailmansodan jälkeiseen jälleerakennukseen osoitettu Yhdysvaltain Marshall-apu tarvitsi avun kohdentamiseksi aiempaa parempia makrotilastovälineitä. OECD:n edeltäjä OEEC oli tässä aloitteellinen. Richard Stone oli laatimassa OEEC:n versiota. Cambridgen yliopistoon oli perustettu kansantalouden tilinpidon tutkimusyksikkö, jota Stone johdatti. Maailmanlaajuiseksi tarkoitettua tilinpidon suositusta kehitettiin Stonen myötävaikutuksella niin, että vuonna 1953 YK julkaisi ensimmäisen varsinaisen tilinpidon suosituksen *A System of National Accounts and Supporting Tables*. Suositus oli laadittu ottamaan huomioon myös kehitysmaiden olosuhteita, mm. tuotantorajan määrittelyssä, mutta oli varsin suppea ja aggregatiivinen, koska Stonen vuonna 1947 laatiman kunnianhimoisen järjestelmän edellyttämät perustiedot ei ollut saatavilla. (Vanoli 2005; Bos 2009.)

Niinpä tätä SNA53:a ei monessakaan länsimaassa otettu käyttöön. Skandinavian maissa noudatettiin pidemmälle viettyä tilinpitoa omien tutkimusten pohjalta. Skandinaviassa kansantalouden tilinpitoa kehitettiin ottaen vaikutteita 1930-luvun Frischin ja Lindahlin töistä sekä Leontiefin panos-tuotoksesta. Hyödykevirtamenetelmä oli vallitseva näkökulma, jos-

kin Ruotsi luopui siitä 1950-luvulla. (Aukrust 1994; Fløttum 2012.)

Ranska kehitti 1950-luvulla omaa systeemiä kritisoiden YK:n järjestelmää liian aggregatiiviseksi. Yhdysvalloissa tilinpitoa on laadittu omista lähtökohdista pitkälle näihin päiviin ja vasta viimeisintä, SNA2008:aa on alettu soveltaa myös siellä. 1950-luvulla niin Ranskassa kuin Yhdysvalloissakin painotettiin myös rahoitustaloustoimia, jotka YK:n järjestelmästä puuttuivat. (Vanoli 2005.)

Vuonna 1968 julkaistua SNA68:aa tehtiin edelleen Richard Stonen johdolla ja YK:n vastuulla. Vuoden 1968 SNA oli jo selvästi ensimmäistä versiota laajempi. Sen laadintaa koskeviin keskusteluihin osallistui maailmanlaajuinen joukko asiantuntijoita ja mm. Skandinavian maiden, esimerkiksi Norjan Odd Aukrustin hyödykevirtamenetelmää koskevat ehdotukset otettiin siinä huomioon. SNA68:ssa suositus laajeni sisältämään panos-tuotostaulukot ja rahoitustilit, myös kiinteähintaiset estimaatit ohjeistettiin. (Fløttum 2012.)

Socialistisissa maissa kehitettiin oma tilinpitojärjestelmä MPS, material product system, joka perustui marxilaiseen tuotantokäsitteeseen ja kuvasi vain tavaratuotantoa.

Nykyisessä suosituksessa tutkimus- ja kehittämisenot mukana investointeina

SNA:n seuraava versio oli SNA93. Sitä kirjoittivat Peter Hill ja Andre Vanoli. YK:n panos väheni, koska järjestön rahoitus oli heikentynyt, ja suosituksen laadinnasta vastasi kansainvälisten järjestöjen yhteistyöryhmä ISWGNA, jossa YK:n lisäksi ovat mukana Kansainvälinen valuuttarahasto (IMF), Eurostat, OECD ja Maailmanpankki. Erityisesti IMF:n ja Eurostatin panos uudistuksessa oli merkittävä. Tässä versiossa tilinpito kansantalouden kuvausjärjestelmänä saavutti täyden laajuutensa: sektoreiden tilit tuotannosta varallisuuden muodostukseen ja rahoitukseen tulivat mukaan, panos-tuotoskehikon rinnalla. Tiliasetelma uudistui, kansantalouden varallisuuden taseista tuli olennainen osa järjestelmää. Myös harmonisointi SNA:n ja IMF:n maksutasemanuaalin (BPM) välillä eteni. (Vanoli 2005.)

Seuraavassa SNA-uudistuksessa oli tar-

koitus alun perin vain päivittää SNA93:a, eikä olennaisia muutoksia tilinpidon rakenteeseen tehtykään. Suosituksen laadintaa ohjasi jälleen ISWGNA, käsikirjan toimittajana oli Anne Harrison. Muutoksia tehtiin kuitenkin siinä määrin, että lyhenteeksi tuli SNA2008. Eräät SNA93:ssa keskeneräisiksi jääneet asiat kuten tutkimus- ja kehittämismenonjen käsitteleminen pääomanmuodostuksena otettiin mukaan. SNA:han vietiin myös joukko EU:ssa jo hyväksytyjä ja käytössä olevia menetelmäohjeita, erityisesti julkista taloutta koskevia. Myös kaikkien eläkevastuiden kuvaaminen sisältyy SNA2008:n tietosisältöön, jotta eri maiden erilaisista piilevistä eläkevastuista saataisiin vertailukelpoista tietoa.

EU:n politiikalle tilinpidon tiedot välttämättömiä

Euroopan Yhteisön tarpeita varten laadittiin vuonna 1970 oma suositus ESA1970 (European System of Accounts), joka on SNA68:aa yksityiskohtaisempi ja eurooppalaiseen oloihin soveltuvampi, mutta peruskäsitteiltään sama. Tätä päivitetiin ja tarkennettiin vuosikymmenen lopulla, ja ESA79 oli voimassa Suomen liitetyssä Euroopan unioniin 1995.

SNA93:n pohjalta uudistetusta Euroopan kansantalouden tilinpitäjärjestelmästä EKT1995:stä tuli myös osa EU-lainsäädäntöä ja sellaisena jäsenmaita velvoittava. EKT-asetus sisältää menetelmä-käsikirjan lisäksi myös tietojen lähettämishjelman, jota jäsenmaiden on noudatettava. Vuonna 2014 voimaan tullut EKT2010 noudattaa SNA2008:aa. Eurooppalainen vaikutus SNA:n päivittämisessä oli merkittävä, koska Euroopan unionissa oli jouduttu ratkaisemaan erityisesti julkista sektoria koskevia tilinpidon soveltamiskysymyksiä vakaus- ja kasvusopimuksen seuranta varten.

Kansantalouden tilinpidon tiedot ovat EU:n hallinnossa hyvin keskeisessä asemassa, ja yhdenmukaisten tietojen tuottamista jäsenmaissa seurataan tarkkaan. Kansantalouden tilinpidon tietoja tai siihen pohjautuvia käsitteitä käytetään perinteisen talousanalyysin lisäksi määrättäessä EU:n budjetin kokoa ja rahoitusta (bruttokansantulo), seurattaessa julkisen talouden tilaa ja kehitystä (julkisyhteisöjen alijäämä ja velka) ja

makrotalouden vakautta (nk. MIP-indikaattorit) sekä aluetukien kohdentamisessa (alueellinen bruttokansantuote) ja EU:n maatalouspolitiikassa (maatalouden taloustilit). Etenkin julkista alijäämää ja velkaa kuvaavista tilastotiedoista on tullut varsin mittava, omaan asetukseen perustuva raportointikonaisuus.

Suomen kansantulolaskelmista EU:n asetuksen mukaiseen kansantalouden tilinpitoon

K. F. Ignatius, joka oli Suomen tilastoviraston päällikkö vuosina 1868–1885, arvioi veroluetteloon liitetyn tulotiedon perusteella Suomen kansantulon vuodelle 1881. Työ jäi kuitenkin tuolloin kertaluonteiseksi, koska kyseistä tulotietoa ei enää saatu. Vasta Valter Lindberg, tuleva Tilastollisen Päätoimiston ylijohdaja, käytti Ignatiuksen metodia arvioidessaan Suomen kansantulon vuodelle 1924. (Luther 1993, 174.)

1920-luvun lopulla alkaneen suuren laman jälkeen myös Suomessa havahduttiin valtion aiempaa aktiivisemmän suhdannepolitiikan tarpeeseen ja makrotalouden tilaston kehittämiseen sitä varten. Valter Lindberg jatkoi laskelmiaan ja Suomen kansantulo vuosilta 1926–38 julkaistiin vuonna 1943 Suomen Pankin suhdannetutkimusosastolla (Luther 1993, 174). Sota-ajan ja sen jälkeisen jälleenrakennuskauden talouspolitiikan harjoittaminen nosti edelleen tarvetta kansantalouden rakenteen ja kehityksen analysointiin. Valtiovarainministeriöön perustettiin kansantalousosasto vuonna 1942 sekä kansantaloudellinen neuvottelukunta. Sodan päättyttyä 1946 perustettiin talousneuvosto laatimaan Suomelle talousohjelmaa. (Luther 1993, 244.)

Vuonna 1947 budjetinuudistuskomitea kiinnitti huomiota virallisen kansantulon laskennan välttämättömyyteen budjetin tarpeita varten. Tämän seurauksena vuonna 1948 valtion budjetissa perustettiin yliaktuaarin ja kahden apulaisen virat. Ensimmäinen yliaktuaari oli Eino H. Laurila, josta myöhemmin tuli Tilastollisen Päätoimiston ylijohdaja.

Ensimmäiset laskelmat – ennakoarvot vuosien 1945–47 kansantulosta ja ennuste vuodelle 1948 liitettiin vasta perustetun valtiovarainministeriön taloudelli-



Eino H. Laurila oli kansantalouden tilinpidon merkittävä edistäjä ja kehittäjä Suomessa. Tilastokeskuksen arkisto.

sen katsauksen perustaksi. Siitä lähtien kansantalouden tilinpidon tiedot ovat olleet keskeinen osa tuota katsausta. Aikasarjat vuosilta 1926–49 julkaistiin vuoden 1950 lopussa.

Ensimmäisten virallisten kansantulolaskelmien esikuvana olivat tilinpidon kehittämisessä aktiiviset muut Pohjoismaat. Eino H. Laurila haki oppia myös kauempaa. Hän osallistui vasta perustetun kansantulo- ja talonjakotutkijoiden kansainvälisen järjestön (International Association for Research in Income and Wealth, IARIW) ensimmäiseen konferenssiin Cambridgesssä vuonna 1949 ja seurasi YK:n työtä tilinpidon kehittämisen alueella. Tilasto laajenikin kansainvälisten suositusten mukaisesti pelkän kansantulon arvioinnista talouden monipuoliseksi kuvaukseksi, kansantalouden tilinpidoksi. (Luther 1993.)

Jo Valter Lindberg oli kiinnittänyt huomiota, että kansantulotilaston kehittäminen edellytti lähteenä käytettävän tilas-



Kirjoittaja on toiminut kehittämisspäälikkänä Tilastokeskuksessa ja tehnyt pitkän työuran kansantalouden tilinpidon parissa.

VOIKO YKSITTÄISELLE KANSANTALOUDELLE LAATIA TILINPITOA GLOBAALISSA MAAILMANTALOUDESSA?

toaineiston kehittämistä. Vähitellen näin tapahtuikin. 1940-luvun lopulta lähtien kehitettiin mm. kiinteähintaisia laskelmia varten tarvittava teollisuuden volyyymi-indeksi, saatiin aikaan liikeyrityslaskenta ja luotiin ansiotasoindeksi. (Luther 1993.)

1950-luvulla kansantulotutkimusta – tai pikemminkin kansantalouden tilinpidon tutkimusta – sekä talustilastojen kehittämistä Suomessa edisti ja harjoitti ns. O-ryhmä. Ryhmässä tutkittiin mm. kiinteähintaisia laskelmia, trendi-, suhdanne- ja kausivaihteluiden mittaamista, säästämistä sekä tuottavuuden ja hyvinvoinnin mittaamista. Ryhmän jäsenistä voidaan erityisesti mainita O.E. Niitamo, josta tuli Tilastokeskuksen pääjohtaja vuosiksi 1979–1992, Jouko Paunio ja Jussi Linnamo. Ryhmän jäsen oli myös Mauno Koi-

visto. (Suvanto & Vesikansa 2002.)

Tutkimuksellinen ote jatkui 1960-luvulta alkaen Suomen Pankin käynnistämässä kasvututkimuksissa, joissa tilinpidon menetelmillä tutkittiin eri toimialojen ja talouden ilmiöiden kehitystä Suomessa (Suur-Kujala 2000, 247).

Vuonna 1958 perustettiin vuositilinpidon yhteyteen kansantuotteen ja -tulon neljännesvuositilasto ja aloitettiin talouden rakennetta kuvaavien panos-tuotostaulukoiden laadinta. Ensimmäiset vuotta 1956 koskeneet taulukot julkaistiin vuonna 1960.

Aluetilinpidosta tehtiin koelaskelmat vuodelta 1962, mutta kehittäminen käynnistyi 1970-luvun alussa, kun lääneittäisiä tietoja tarvittiin juuri säädettyjen aluepoliittisten lakien toteutu-

misen seuranta varten. Aluetilinpidosta ei ollut tuolloin kansainvälisiä suosituksia, joten järjestelmää kehitettiin pitkälti kansalliselta pohjalta. Ensimmäiset tiedot julkaistiin vuonna 1975, ja sen jälkeen tietoja on julkaistu säännöllisesti. (Suur-Kujala 2000.) Nykyään myös aluetilinpito on osa EKT-asetusta.

SNA68:aan siirryttiin Suomessa vuonna 1979, pitkän ja vaivalloisen SKT75-kehittämiprojektin tuloksena. Sektoritilit, joissa mitataan päätöksiä tekevien talousyksiköiden kuten yritysten, julkisyhteisöjen ja kotitalouksien tuloja, menoja ja rahoitusasemaa saatiin tuolloin mukaan, mutta panos-tuotoslaskelmat, jotka alun perin oli tässä projektissa tarkoitus yhdistää vuositilinpitoon, jäivät vielä erillisiksi, vuositilinpidon jälkeen laadittaviksi. Uudistusta varten käynnistettiin myös mitattava yritystilastojen uudistus, joka käytännössä osoittautui yrityksille liian hankalaksi, koska yrityskyselyssä käytettiin yrityksille vierasta kansantalouden tilinpidon käsitteistöä. Kyselylomaketta jouduttiinkin supistamaan. Rahoitustilinpito, josta oli tehty laskelmia tilastovuodesta 1970 lähtien, liitettiin muuhun vuositilinpitoon vuodesta 1985 alkaen. Neljännesvuositilinpito eriytettiin omaksi tilastokseen vuodesta 1978 alkaen. Nopeaa suhdannetietoa varten aloitettiin vuonna 1986 bruttokansantuotteen kehitystä ennakoiva kokonaistuotannon kuukausikuvaaja, nykyään tuotannon suhdannekuvaaja. (Suur-Kujala 2000.)

Suomen EU-jäsenyyden myötä siirryttiin EKT95:een EU-asetuksen mukaisesti vuonna 1999. EKT95:n mukaiset panos-tuotostaulukot ja nk. tarjontaja käyttötaulukot oli myös tuotettava, ja näin jo 1970-luvulta lähtien tavoitteena ollut panos-tuotoksen integrointi vuositilinpitoon toteutui, ensin käypähintaisena (2003) ja hiukan myöhemmin kiinteähintaisena (2006). EU:n talouspolitiikan tarpeita varten melko pitkällä viiveellä julkaistavat vuositiedot eivät ole riittäneet, joten EU:ssa on panostettu nopeiden neljännesvuosiestimaattien kehittämiseen. Perinteisen bruttokansantuotteen ja kansantulon kehitystä kuvaavan neljännesvuositilinpidon rinnalle ovat tulleet ensin julkisyhteisöjen neljännesvuositilit ja myöhemmin myös muita sektoreita koskevat neljännesvuositiedot.

Kesällä 2014 siirryttiin EKT 2010:een.

Tällä hetkellä Suomen tilinpito täyttää EU-vaatimukset, mutta kehitettävää on edelleen vielä erityisesti varallisuustilinpidoissa, joka EU-asetuksessakin on toteutettava vasta vuoteen 2017 mennessä. Julkisen talouden tilastointia joudutaneen myös edelleen kehittämään EU:n hallinnollisten tarpeiden vuoksi. (Ks. Sainio & Koistinen-Jokiniemi 2014.)

Kansantalouden tilinpidon tulevaisuuden kysymyksiä

Jo Simon Kuznetsin ajoista lähtien on keskusteltu siitä, onko bruttokansantuote tai kansantulo hyvinvoinnin mittari. Kansantalouden tilinpidon kehittäjät ja laatijat ovat olleet varsin yksimielisiä siitä, että nämä eivät ole hyvinvoinnin mittareita, korkeintaan kuvaavat yhtä hyvinvoinnin osatekijää. Mielenkiintoisena yksityiskohtana em. O-ryhmän työstä mainittakoon Jouko Paunioin artikkeli kansantulolaskelmista hyvinvoinnin ja tuottavuuden muutosten mittana. Hän käyttää kansakunnan hyvinvoinnin mittarina reaalisista käytettävissä olevaa kansantuloa. (Paunio 1957.) Perusteellisesti kansantalouden tilinpidon hyvinvointitulottuuksia on käsitellyt nk. Stiglitzin komitea koottaessaan hyvinvointinäkökulmaan vaikuttavia tilinpidon ominaisuuksia ja mahdollisia laajennuksia, erityisesti koskien kotitaloussektoria (Stiglitz 2009).

Kansantalouden tilinpito on makrotalasto eikä pureudu kovin syvälle kaikkien ilmiöalueiden yksityiskohtiin. Niinpä sen rinnalle on syntynyt ns. satelliittitilejä, joissa jotakin ilmiötä kuvataan joko varsinaista tilinpitoa yksityiskohtaisemmin tai laajentamalla tilinpidon mukaista käsitteistöä tai rajauksia. Satelliittitilien käsite ja ohjeistus ovat nykyään myös SNA:ssa ja EKT:ssä. Suomessa on laadittu kotitaloustuotantoa, kulttuuria ja matkailua koskevia tilinpitoja sekä metsätilinpitoa. Tilastotoimen resurssien supistumisen takia näiden tulevaisuus ei juuri nyt näytä kovin valoisalta. Ympäristötilinpidon kehittämisellä on kuitenkin EU-asetuksella säädettyä turvattu tulevaisuus.

Kansantalouden tilinpitoon ja sen tulevaisuuteen liittyy tiettyjä uhkakuvia. Seuraavankaltaisia kysymyksiä on esitetty:

- Onko tilinpidon tietojen laajentunut hallinnollinen käyttö pirstomassa ko-

ko talouden kuvaukseen ja analyysiin tarkoitettua järjestelmää?

- Ovatko nykyisen talouspolitiikan tarvitsemat nopeat tiedot syrjäyttämässä tutkimuksen tarvitsemien syvällisempien rakenteellisten tietojen laadintaa?
- Pystyykö laaja järjestelmä uudistumaan riittävän nopeasti talouden kiihtyvän muutoksen vauhdissa?
- Voiko yksittäiselle kansantaloudelle ylipäänsä laatia tilinpitoa globaalissa maailmantaloudessa?

Näiden kysymysten ratkaisemisessa riittää haastetta. Kansantalouden tilinpito ei ole koskaan täysin valmis. ■

Lähteet:

Aukrust, Odd 1994. *The Scandinavian contribution to national accounting*.

Teoksessa: Zoltan Kenessey (toim.). *The Accounts of Nations*. Amsterdam: IOS Press.

Aulin-Ahmavaara, Pirkko 2007. *Kansantulon mittaamisesta kansantalouden tilinpitoon*.

Teoksessa: Jukka Jalava & Jari Eloranta & Jari Ojala (toim.). *Muutoksen merkit. Kvantitatiivisia perspektiivejä Suomen taloushistoriaan*. Tutkimuksia 245. Helsinki: Tilastokeskus.

Bos, Frits 2009. *The National Accounts as a Tool for Analysis and Policy*. Saarbrücken: VDM Verlag Dr. Muller.

Fløttum, Erling Joar & Halvorsen, Tore & Hobbeltstad Simpson, Liv & Skoglund, Tor 2012. *History of National Accounts in Norway. From free research to statistics regulated by law*. http://www.ssb.no/en/nasjonalregnskap-og-konjunkturer/artikler-og-publikasjoner/_attachment/79796?_ts=13a306a8068.

Kaynes, John Maynard 1936. *The General Theory of Employment Interest and Money*. London: Macmillian.

Luther, Georg 1993. *Suomen tilastotoimen historia vuoteen 1970*. Helsinki: WSOY.

Paunio, Jouko 1957. *Kansantulolaskemat hyvinvoinnin ja tuottavuuden muutosten*

mittana. Teoksessa: Antti Suvanto & Jyrki Vesikansa 2002 (toim.). *Modernismi taloustieteessä ja talouspolitiikassa. O-ryhmän kirjoituksia ja kirjoituksia O-ryhmästä*. Helsinki: Gaudeamus.

Sainio, Mika & Koistinen-Jokiniemi, Paula 2014. *Kansantalouden tilinpito uudelle vuosikymmenelle*. *Kansantaloudellinen Aikakauskirja* 3/2014.

Skidelsky, Robert 2005. *John Maynard Keynes 1883–1946: Economist, Philosopher, Statesman*. London: Penguin Books.

Stiglitz, Joseph E. & Sen, Amartya & Fitoussi, Jean-Paul 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress*. http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf.

Stone, Richard 1984. *The Accounts of the Society*. Nobel memorial lecture. http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1984/stone-lecture.pdf.

Suur-Kujala, Markku 2000. *Kansantalouden tilinpito eurooppalaistuu*. Teoksessa: Pertti Marjomaa (toim.). *Tilastokeskus 1971–2000*. Helsinki: Tilastokeskus.

Suvanto, Antti & Vesikansa, Jyrki (toim.). 2002. *Modernismi taloustieteessä ja talouspolitiikassa. O-ryhmän kirjoituksia ja kirjoituksia O-ryhmästä*. Helsinki: Gaudeamus.

Vanoli, Andre 2005. *A History of National Accounting*. Amsterdam: IOS Press.

Kansantalouden tilinpidon suositukset ja käsikirjat:

Historic versions of the System of National Accounts. <http://unstats.un.org/unsd/nationalaccount/hsna.asp>.

European system of accounts ESA2010, 2013. Luxembourg: Eurostat European Commission.

List of the main methodological references on national accounts, Eurostat. <http://ec.europa.eu/eurostat/web/national-accounts/methodology>.



KLAUS HELKAMA: SUOMALAISTEN ARVOT. MIKÄ MEILLE ON OIKEASTI TÄRKEÄÄ? HELSINKI: SUOMALAISEN KIRJALLISUUDEN SEURA 2015. 252 S.

Arvaamattomat arvot

Klaus Helkama on kirjoittanut helppolukuisen kirjan suomalaisten arvoista. Hän kertoo arvojen perusteista, suomalaisten arvoista verrattuna muiden kansojen arvoihin ja miten arvot kuuluvat suomalaisten omakuvaaan. Hän käsittelee talouden, politiikan ja globalisaation yhteyksiä arvomuutoksiin sekä pohtii, miten ikä, sukupuoli ja koulutus vaikuttavat arvoihin. Kirjassa on jopa testi, jolla lukija voi kartoittaa omat arvonsa.

Erityisen tärkeää kirjassa on, että Helkama on kuvannut perusteellisesti suomalaisen arvotutkimuksen lähtökohtia, historiaa ja kiinteitä kansainvälisiä yhteyksiä. Kirjaa lukiessa avautuu, miten laajaa, pitkäaikaista ja kansainvälistä arvotutkimus Suomessa on ollut. Helkama valottaa asiaa myös kirjan loppuun kirjoittamisensa kiitossanoissa, joissa hän korostaa kirjan pohjana olevan tutkimustiedon syntyneen monien ihmisten pitkäaikaisista yhteisistä ponnisteluista. Lukija oppii tuntemaan kansainvälistä ja suomalaista arvotutkimusta sekä suomalaisia arvotutkijoita.

Kiinnostavan näkökulman kirjaan tuo Helkaman käyttämä kaunokirjallinen aineisto. Hän ku-

vaa mm. mitä Aleksis Kiven *Seitsemän veljestä* kertoo suomalaisista arvoista. Hän tulkitsee Kiven teosta kehitysromaanina, ”tarinana sosiaalisuudesta, siirtymisestä suullisen tarinanperinteen, taikauskon ja metsästyksen maailmasta kirjoitetun kielen, valistuksen ja maanviljelyksen maailmaan”. Kiven romaanin aihe ilmaisee Helkaman mukaan sivistyksen ja koulutuksen arvostusta Suomessa. Myös tasa-arvon ja työn teemat ovat teoksessa keskeisesti esillä.

Arvotutkimus ei ole ollut kiistatonta. Helkama käsittelee – joskin melko lyhyesti – yhteiskuntatieteilijöiden, humanistien ja historioitsijoiden epäilyksiä arvotutkimusta kohtaan. Arvostelu kohdistuu kulttuurien välisten erojen tutkimiseen. Arvoilla on kulttuuriin ja historiaan sidotut merkityksensä. Abstrakteihin arvoihin viittaavien sanojen merkitykset poikkeavat eri kielissä ja kulttuureissa.

Helkama vastaa tähän kritiikkiin esittelemällä, miten arvotutkimuksessa on asiaa käsitelty ja miten siinä on pyritty kehittämään ilmauksia, jotka ovat merkitykseltään ”suhteellisesti samoja” karsien ilmauksia, joiden merkitykset poikkeavat kulttuurista toiseen. Hän korostaa, että nykyisin kulttuurienvälisessä psykologiassa käytettävät mittarit perustuvat huolelliseen pohdintaan ja vertailuun sekä käyvät läpi kielestä toiseen takaisin-kääntämisen. Helkama muistuttaa kuitenkin, että arvotutkimus on suhteellista ja tulokset riippuvat siitä arvojen joukosta, jota kulloinkin arvioidaan sekä siitä, miten arvo sanallisesti määritellään. Eri tavoin toteutetuista arvotutkimuksista voidaan saada erilaisia tuloksia.

Pyhtää ja sen asukkaiden arvot ovat olleet Helsingin yliopiston arvotutkijoiden tutkimuskohteena vuosina 1975–2007. Pyhtää on edustanut arvotutkijoille suomalaista keskivertopaikkakuntaa, jonka elinkeinorakenne ja poliittinen rakenne ovat olleet hyvin lähellä Suomen keskiarvoja. Helkama toteaa: ”Mikä pitää paikkansa pyhtääläisistä, pitää paikkansa myös kaikista suomalaisista.” Pyhtää symboloi hänen mukaansa myös pysyvyyden ja muutoksen suhdetta mm. pitkän historian omaavana maantieteellisenä läpikulkupaikkana idän ja lännen välillä. Pyhtäältä on kerätty arvojen lisäksi työhön, vapaa-aikaan ja terveyteen liittyviä tietoja, joten aineisto on varsin rikas ja monipuolinen.

Teoksessaan Helkama esittelee kolme kansainvälistä tutkimusohjelmaa, joiden kautta suomalaisten arvoja voidaan vertailla muiden maiden kansalaisten arvoihin. Amerikkalainen politiikan tutkija Ronald Inglehart käynnisti 1970-luvulla aluksi poliittisia ja myöhemmin muitakin arvoja tutkivan ohjelman. 1980-luvulla hän muotoili teesinsä äänettömästä tai hiljaisesta vallankumouksesta, jonka mukaan vauriissa maissa postmaterialistiset arvot korvaavat materialistisia arvoja.

Alankomaalainen insinööri ja sosiaalipsykologi Geert Hofstede käynnisti tutkimuksensa niin ikään 1970-luvulla ja selvitti IBM:n työntekijöiden työhön liittyviä arvoja. Hän kiinnitti huomiota erityisesti kulttuurien välisiin eroihin.

Amerikkalais-israelilainen sosiaalipsykologi Shalom Schwartz aloitti 1980-luvulla opettajien ja opiskelijoiden henkilökohtaisia arvoja käsitelleen tutkimuksen, joka on laajentunut kattamaan suuren osan maailman maita. Suomalaiset arvotutkijat ovat tehneet paljon yhteistyötä Schwartzin kanssa. Hän osoitti, että arvot muodostavat psykologisesti kehämäisen kaksiosaisen järjestelmän, jolla on omat säännönmukaisuutensa.

Vuonna 1992 julkaistussa Schwartzin arvomittarissa oli 40 eri kulttuureissa samaa merkittävää arvo-osiota. Nämä sijoittuivat 10 arvotyypin, jotka muodostivat Schwartzin mallin kehämäisen rakenteen. Mallissa on kaksi pääulottuvuutta: itsensä ylittäminen – itsensä korostaminen sekä säilyttäminen – muutosvalmius. Itsensä ylittäminen viittaa universalistiseen suvaitsevaiseen ja altruistiseen arvomaailmaan, kun itsensä korostamisessa painottuvat kunnianhimo ja menestys. Säilyttäminen viittaa perinnettä kunnioittavaan ja turvallisuushakuiseen arvomaailmaan. Muutosvalmius on haasteiden ja toiminnan arvostamista.

Vuonna 2002 Schwartzin arvomittari otettiin yksinkertaistetussa muodossa käyttöön Euroopassa. Kaikkien EU:n jäsenmaiden kansalaisille suunnattu European Social Survey osoitti, että keskimääräiset arvot ovat varsin samankaltaisia eri maissa. Suomalaiset erottautuivat kahdessa suhteessa: he olivat Euroopan ykkösiä universalismiarvojen kannatuksessa. Heille kaikkien ihmisten hyvinvointi ja suvaitsevaisuus olivat tärkeämpiä kuin muille.

Toinen arvotyyppi, jossa suomalaiset erottautuivat, oli yhdenmukaisuus. Suomessa oli Länsi-Euroopan maista Norjan ohella kaikkein vahvin yhdenmukaisuuden kannatus, mikä lähentää meitä Itä-Euroopan maihin.

Helkama käsittelee teoksensa omissa luvuissa joitakin erityisesti suomalaisten omakuvaan kuuluvia arvoja. Tällaisia ovat luontoarvot ja tasa-arvo, suomalainen työ ja kunnia, rehellisyys, maanpuolustustahto, sivistys sekä koulutus ja viisaus. Hän toteaa, että nämä arvot ovat säilyneet omakuvassamme vuosikymmeniä, vaikka yksilöiden arvot ovatkin muuttuneet.

Kirjan päättää luku, jossa Helkama pohtii, mihin suuntaan arvot muuttuvat. Hän korostaa ensinnäkin, että arvot muuttuvat hitaasti, mutta muistuttaa toisaalta, että pyhtääläisten arvoissa tapahtui vuosien 1975 ja 2007 välillä merkittäviä muutoksia. Niin Pyhtäällä kuin koko Suomessakin yksilökeskeisyys ja minäkeskeiset arvot kuten

”JOKAISELLA SUKUPOLVELLA
TUNTUU OLEVAN TAIPUMUS NÄHDÄ OMA
AIKANSÄ ARVOMURROKSEN AIKAKAUTENA,
MUTTA ARVOMUUTOS ON TODELLISUUDESSA
USEIMMITEN VÄHITTÄISTÄ JA HIDASTA.”

itsekkunnioitus, rohkeus, vapaus, terveys ja hyvä toimeentulo ovat lisääntyneet mm. universalismiarvojen kuten maailmanrauhan, nälän poistamisen ja luonnon säilyttämisen vähentyessä. Helkama tulkitsee muutosta tukeutuen Inglehartin näkemykseen, jonka mukaan vaurastuminen muuttaa arvomaailmaa aiempaa yksilökeskeisempään suuntaan turvallisuutta korostavien arvojen jäädessä vähemmälle merkitykselle.

Helkama esittelee mielenkiintoisen tuloksen 1990-luvun laman vaikutuksesta pyhtääläisten arvoihin. Lama ilmeni universalismiarvojen sovellusalan kapeutumisenä ja moraalisen universumin supistumisena. Tämä tarkoittaa sitä, miten laajasti henkilö tulkitsee sellaisia arvoja kuin tasa-arvo, suvaitsevuus, sosiaalinen oikeudenmukaisuus ja maailmanrauha. Kyse on siitä, mieltäkö hän ne koskemaan kaikkia maailman ihmisiä vai vain omaa kansaa tai muuta omaa ryhmää.

Yksilökeskeisyyden kasvu on kuitenkin viime aikoina taittunut: Yksilökeskeisyyden kasvu suomalaisten arvoissa tapahtui 1970-luvun ja 1990-luvun alun välillä. Sen jälkeen arvomaailma pysyi vakaana yli kymmenen vuotta, mutta vuonna 2005 yksilökeskeistyminen taittui ja suomalaisten arvot kääntyivät jälleen yhteisöllisempään suuntaan. Perinteiden, turvallisuuden ja yhdenmukaisuuden arvostus on lisääntynyt. Myös tasa-arvon tärkeys on vahvistunut sitä mukaa, kun suomalaisten epävarmuus on lisääntynyt.

Helkaman ennustus suomalaisista arvoista kahdenkymmenen vuoden päähän on, että pohjoismaiset arvot kuten tasa-arvo, suvaitsevaisuus ja yhteisöllisyys tulevat pitämään pintansa, mihin vaikuttaa valtion ja kansalaisyhteiskunnan välinen laaja luottamus. ”Mutta tulevaisuus on arvaamaton.” ■

KIRSTI AHLQVIST

Kirjoittaja on erikoistutkija Tilastokeskuksen väestö- ja elinoloilastot -yksikössä.



KAI ERIKSSON (TOIM.). VERKOSTOT YHTEISKUNTATUTKIMUKSESSA. HELSINKI: GAUDEAMUS 2015. 355 S.

Verkostojen voima



ALEX PENTLAND. SOCIAL PHYSICS. HOW GOOD IDEAS SPREAD – THE LESSONS FROM A NEW SCIENCE. NEW YORK: THE PENGUIN PRESS 2014. 300 S.

Tietotekniikan kehittyminen on mahdollistanut yhä suurempien tietoa-aineistojen mallintamisen. Laskentakapasiteetin kasvaessa asioiden, ihmisten ja ilmiöiden välisten yhteysien selvittämisestä on tullut aiempaa nopeampaa ja edullisempää. Rinnan tämän teknologisen kehityksen kanssa on todellisuutta alettu hahmottaa enenevässä määrin verkosto-käsitteen kautta.

”Ajatus verkostosta yhteiskuntaa ja sen toimintapaa kuvaavana muotona on yleistynyt varsinkin sosiologisessa puheessa niin, että 2010-luvulla yhteiskuntaa ei juuri voi ajatella verkoston kä-

sitteestä irrallaan”, kirjoittaa Kai Eriksson tänä keväänä ilmestyneen *Verkostot yhteiskuntatutkimuksessa* -teoksen johdantoluvussa.

Mistä verkostotutkimuksessa on kysymys? Tähän vastaa Erikssonin toimittama kotimainen kirja. Entä miten verkostotutkimuksen tuomaa uutta tietoa voisi hyödyntää? Tähän tarjoaa ideoita amerikkalaisen Alex ”Sandy” Pentlandin viime vuonna julkaistua teos *Social Physics*.

Sosiaalifysiikka mallintaa tiedon ja ideoiden liikettä

Alex Pentland on tietojenkäsittelytieteilijä, MIT-yliopiston professori, yrittäjä ja yksi datavallankumouksen merkittävimmistä visionääreistä. Pentland ja hänen tuottelias tutkimusryhmänsä ovat tutkineet ja luoneet erilaisia iso data -aineistoja, kehittäneet aineistojen käsittelyyn soveltuvia ohjelmistoja sekä perustaneet yrityksiä, jotka tuoteistavat tietojenkäsittelyä.

Pentland on propellihattu, jota tietojenkäsittelyn avaamat mahdollisuudet innostavat teknisessä mielessä mutta myös maailmanparantajana. Hän uskoo ja toivoo, että tekniikka voidaan valjastaa ihmiskunnan hyväksi.

Pentlandin kirja *Social Physics* on vapaasti suomennettuna Sosiaalifysiikka. Kirjan perusidea on, että iso data -aineistot ja uusin verkostotutkimus ovat luoneet tai luomassa aivan uuden tieteenalan, sosiaalifysiikan.

Pentland määrittelee sosiaalifysiikan määrälliseksi yhteiskuntatieteeksi, joka pyrkii matemaattisesti mallintamaan tiedon ja ideoiden liikettä sekä yksilöiden ja yhteisöjen käyttäytymistä. Jos fysiikka tutkii aineen ja energian virtaamisen lakeja ja vaikutuksia liikkeeseen, pyrkii sosiaalifysiikka tutkimaan ajatusten ja informaation virtauksen vaikutuksia ihmisten ja organisaatioiden käyttäytymiseen – sekä merkitystä yhteiskunnallisessa murroksessa.

Teos jakaantuu neljään osaan: Ensimmäisessä osassa Pentland selittää mitä tarkoittaa sosiaalifysiikalla. Toisessa osassa hän esittää tapoja käsitteellistää ja mallintaa ihmisten kanssakäymistä. Kolmas osa on omistettu yhteiskuntasuunnittelulle ja kertoo siitä, miten tietotekniikkaa voidaan hyödyntää ns. älykkään kaupungin rakentamisessa. Neljäs osa on otsikoitu *Data-Driven Society* ja kertoo datavallankumouksen mahdollisuuksista ja vaikutuksista yhteiskunnille.

Menestyminen edellyttää ajatusvirtaa

Kirjan verkostotutkimuksesta ponnistava lähtökohta on, että ihmiset muistuttavat mitä suurimmassa määrin mehiläisiä ja muita sosiaalisia hyönteisiä. Ihmissuhteemme ja ryhmäjäsenyytemme ohjaa-

vat arkisia tapojamme, kuten syömistä ja muuta kulutusta, makua ja mielipiteitä. Usein ryhmä vaikuttaa yksilön toimiin alitajuisesti.

Keskeinen käsite Pentlandin kirjassa on *idea flow*, jonka voisi suomentaa vaikkapa ajatusvirraksi. Pentlandin mukaan menestyville yksilöille, ryhmille ja organisaatioille on tyypillistä uusien ajatusten aktiivinen etsiminen (*exploration*). Menestyjät pyrkivät kartoittamaan mahdollisimman monia erilaisia näkemyksiä ja kokemuksia ja mm. saatavilla olevia palveluita ja muita kulutuskohteita. Kysymys ei ole siitä, että he tietoisesti tai tarkoituksella etsisivät ”parhaita” ideoita, vaan pikemminkin pyrkivät keräämään ja testaamaan mahdollisimman monia ja erilaisia ajatuksia. Tilastotieteen termeiksi tiivistettynä voisi sanoa, että pyrkimyksenä on laatia sopiva otos todellisuudesta.

Toisinaan – ja ehkä hyvinkin tyypillisesti – saatetaan kuitenkin päätyä vastakkaiseen olotilaan, josta Pentland käyttää sanaa *echo chamber* (kaikekammio). Tällöin yksilö tai ryhmä liikkuu yksipuolisissa piireissä ja altistuu vain samoille keskinertaisille tai jopa vahingollisille ajatuksille. Uusia vaikutteita ei keräänny eikä vanhoja oleksia jouduta haastamaan.

Pentland on optimistinen sen suhteen, että paremmat tiedot innostavat meitä tekemään parempia valintoja ja elämään paremmin, vaikka hän toisaalla mainitseekin ohimennen, että olisi tärkeää vakuuttaa ihmiset tiedon hyödyllisyydestä ja innostaa heidät toimimaan omaksi parhaakseen.

Hajautettu päätöksenteko on keskitettyä parempi

Pentland pohtii kirjassaan, miten yhteiskuntien tulisi hyödyntää uutta tietoa verkostojen voimasta ja merkityksestä ihmisyyhteisöissä. Voitaisiinko tietoa soveltaa esimerkiksi yhteiskuntasuunnittelussa? Yhteiskuntasuunnittelun pitäisi Pentlandin mielestä pohjautua kolmen tekijän – yhteiskunnan tehokkuuden (*social efficiency*), toiminnallisen tehokkuuden (*operational efficiency*) ja resilienssin (*resilience*) lisäämiselle.

Yhteiskunnallisen tehokkuuden mitta on karkeistaen, että yksilön menestyminen johtaa koko yhteiskunnan menestymiseen. Toiminnallinen tehokkuus viittaa yhteiskunnan infrastruktuuriin, jonka tulisi olla luotettavaa ja taloudellista. Pentlandin mielestä nykyisellään (Yhdysvalloissa) niin rahamarkkinoiden, liikenteen, terveydenhuollon, energiatalouden kuin poliittisen järjestelmänkin toimivuus on kyseenalaista. Pentland arvelee, että järjestelmien halvaantuneisuus johtuu siitä, että niiden perusrakenteet ovat peräisin 1800-luvulta, hevostäryjen aikakaudelta.

Resilienssillä Pentland tarkoittaa käsittääksee-

ni suunnilleen samaa asiaa kuin mitä filosofi ja tutkija Nassim Nicholas Taleb peräänkuuluttaa kirjoissaan *Musta Joutsen: Erittäin epätodennäköisen vaikutus* ja *Antihauras: Asioita jotka hyötyvät epäjärjestyksestä* (suom. Kimmo Pietiläinen): Järjestelmistä – olivatpa ne sitten talouteen, hallintoon tai työhön liittyviä – ei pidä rakentaa liian isoja, vaan tulisi sallia järjestelmien moninaisuus.

Liian iso tai yhden kortin varaan rakennettu

IHMISET MUISTUTTAVAT MITÄ SUURIMMASSA MÄÄRIN MEHILÄISIÄ JA MUITA SOSIAALISIA HYÖNTEISIÄ.

järjestelmä on haavoittuvainen katastrofin iskiessä. Olisi tärkeää, että yhden osan pettäessä löytyisi korvaavia voimia. Esimerkkinä Pentland mainitsee hurrikaani Katrinan vaikutukset New Orleansissa vuonna 2005, kun terveydenhuoltojärjestelmä uhkasi kaatua kaupungin puhelin-yhteyksien katkettua. Radioamatöörit tulivat apuun ja koordinoivat lääkkeitä ja hoitovälineiden toimitusta. Talebin tavoin Pentland uskoo hajautettuun, paikallistason päätöksentekoon.

Kuka ylläpitää rekisterejä?

Kirjan tekee kiinnostavaksi Pentlandin pyrkimys luoda eräänlainen yhteiskuntateoria, joka perustuu verkostotutkimukseen (ja sen luomaan ihmiskuvaan), tietoaineistoihin ja -teknologioihin, kuten digitaaliseen henkilötietoon. Havaintojensa pohjalta Pentland sommittelee ohjekirjaa siitä, mihin suuntaan yhteiskuntaa pitäisi kehittää.

Pentlandin teknologinen optimismi, edistysusko ja luottamus yhteiskuntien järkipäiseen hallitsemiseen tuovat mieleen valistusajan kirjailijat. Maailmanparannuspuheista huolimatta Pentland on kuitenkin pohjimmiltaan ennen kaikkea nörtti eikä yhteiskuntafilosofi. Kirjasta puuttuvat viitaukset merkittävimpiin tietoyhteiskuntateorioihin ja verkostotutkimuksen klassikkoihin. Yhteiskunnalliseen tutkimukseen perehtyneelle lukijalle kirjan näkökulma ja perustelut näyttävätkin paikoin ohuina. Tarina loppuu jotenkin kesken.

Moniin isoihin kysymyksiin ei Pentlandilla ole ►

”VERKOSTOSTA ON TULLUT YLEINEN NIMITYS NYKYISELLE YHTEISKUNTAMUODOLLE JA SOSIAALISELLE JÄRJESTYKSELLE.”

suoraa vastausta. Mikä motivoi yksilöitä ja perheitä tekemään hyviä valintoja arjessaan? Miten luoda tehokas ja hyvä hallintojärjestelmä, joka paitisi hyödyntää uutta tietoa myös aidosti pyrkii parantamaan kansalaisten hyvinvointia? Ketkä lopulta ovat vastuussa tietoon perustuvan yhteiskunnan luomisesta tai ihmisten kannustamisesta ottamaan vastuuta omasta elämästään?

On kiinnostavaa, että Pentland näkee järjestelmien kehityksen varteenotettavimpana lupauksena erilaisten rekisterien luonnin, joihin kootaan hyödyllinen tieto (esim. eri lääkeaineiden tehokkuudesta ja haittavaikutuksista). Pentland käyttää rekistereistä nimeä *public data commons*, muttei ota kantaa siihen kuka tällaisia rekisterejä voisi ylläpitää. Koska Pentland suhtautuu amerikkalaisittain kriittisesti niin julkiseen kuin yritysvaltaankin, on tämä ollut hänelle todennäköisesti vaikea kysymys. Olisiko tässä nähtävissä eräänlainen sikäläisen yhteiskunnan kaikukammio-olotila?

Kirjan lopussa Pentland toteaa heitonomaisesti, että tieteellisen metodin avulla ei pystytä vastaamaan yhteiskunnallisiin haasteisiin – ainakaan yhtä hyvin kuin ison datan avulla. Tässä hän jättää mainitsematta ison datan keräämiseen ja tulkitsemiseen liittyvät ongelmat. Tulkitsemisen Pentlandin pettymyksen olemassa oleviin metodeihin johtuvan ainakin osittain siitä, että hänen saatavillaan oleva survey- ja rekisteritieto ei ole erityisen korkealaatuista. Pentlandia harmittaa mm. se, että Yhdysvalloissa (kuten myös monissa muissa maissa) väestönlaskenta suoritetaan vain joka kymmenes vuosi. Jokohan meikäläisestä rekisteriosaamisesta olisi vientituotteeksi?

Kohti teorian ja käytännön synteisiä

Siinä missä Pentlandin kirja keskittyy enemmän ison datan ja verkostotutkimuksen käytännön sovelluksiin, tarjoaa Kai Erikssonin toimittama *Verkostot yhteiskuntatutkimuksessa* -teos harvinaisen

perusteellisen kirjallisuuskatsauksen verkostotutkimuksen teorioihin. Verkostoajattelun aatteelliset ja historialliset yhteydet käyvät kirjasta hyvin ilmi, vaikka paikoin lukijalle saattaa syntyä vaikutelma, että lähes kaikki sosiologinen ajattelu viimeisen sadan vuoden ajalta on tavalla tai toisella kytköksissä verkostoajatteluun.

Verkostoajattelun kehittymisen kannalta keskeiseksi paikoiksi nostetaan Pariisi ranskalaisen yhteiskuntafilosofian keskuksena, Chicago amerikkalaisen yhteiskunta- ja taloustutkimuksen keskuksena sekä Harvard, jossa verkostotutkimuksen kannalta merkittävä Pareto-piiri veti puoleensa eri tieteenalojen edustajia.

Verkostot yhteiskuntatutkimuksessa on rakenteeltaan onnistunut ja selkeä. Ensimmäisen osan artikkeleissa käydään läpi verkostoajattelun (länsi-)eurooppalaista perinnettä ja toisessa osassa angloamerikkalaista perinnettä. Osien alussa on yleiskatsaus kumpaisenkin verkostoajattelun juuriin, minkä jälkeen on erikoistuneempia artikkeleita.

Eurooppa-osiossa käsitellään mm. rihmaston, dispositiivin ja triadin käsitteitä sekä toimijaverkoston ja oikeuttamisen verkoston ideoita. Amerikka-osiossa käsitellään pragmatismia, verkostoyhteiskuntaa, heikkojen suhteiden ja uppoutuneisuuden käsitteitä, sosiaalista pääomaa, verkostotutkimusta, verkostohallintaa sekä verkostotutkimustiedettä.

Artikkelit vilisevät viittauksia sosiologian klassikkoihin, filosofisiin koulukuntiin sekä näiden välisiin suhteisiin ja vastakkainasetteluihin. Empiiristä tutkimusta ei juurikaan esitellä. Jos Pentlandin esimerkit ja *Verkostot yhteiskuntatutkimuksessa* -kirjan teoreettiset pohdinnat voisi jotenkin yhdistää, saattaisi verkostojen voiman kuvaus olla harvinaisen vakuuttava.

Verkostot yhteiskuntatutkimuksessa -teoksella on yhteensä 12 kirjoittajaa. He ovat tutkijoita, dosentteja ja professoreita eri yliopistoista ja tutkimuslaitoksista. Kirjoittajien joukossa ei ole yhtään naista. Sangan vähän naisia on myöskään verkostotutkimuksen klassikoiden joukossa. Tai iso data -asiantuntijoiden joukossa. Mikä tähän lienee syynä? Saataisiinko vastaus – verkostotutkimuksesta? ■

KAISA SAARENMAA

Kirjoittaja on Tieto&trendit – talous- ja hyvinvointikatsaus -julkaisun toimittaja.

Vaalipulssi.fi seurasi vaalien somekohinaa

Euro, sananvapaus, vihapuhe, järjestöjohto, keskiolut, osanotto, koulutuslupaus, opintotuki, sateenkaarisuomi, energiaremontti, tasa-arvo, ennakkoääni...

Siinäpä joukko aiheita, jotka kukin vuorollaan nousivat sosiaalisessa mediassa kevään kuluessa käydyin vaalikeskustelun kärkeen. Missä määrin ja milloin aiheet liittyivät eri puolueisiin ja ehdokkaisiin, selviää sosiaalista mediaa vaalien alla monitoroineesta vaalipulssi.fi -palvelusta.

Palvelun avulla sosiaalisessa mediasa käytyä vaalikeskustelua saattoi seurata interaktiivisesti reaaliajassa. Vaalipulssi.fi poimi vaalikeskustelussa esiin nousseita aiheita paitsi yleisellä myös puolue- ja ehdokastasolla. Päivän polttavien aiheiden lisäksi keskustelua saattoi seurata viikkokohtaisesti ja tarkkailla, mitä aiheita tai teemoja seurattaviin puolueisiin ja ehdokkaisiin somekeskustelussa kulloinkin liitettiin.

Seurannan kohteina olivat kahdeksan suurinta puoluetta ja kunkin puolueen kymmenen ajankohtaisinta poliitikkoa – eli heidän itsensä käymä ja heistä käytävä sosiaalisen median keskustelu. Si-

vuston tulokset muodostettiin monitorimalla valittuihin ehdokkaisiin sosiaalisessa mediassa liitettäviä sanoja ja niiden esiintymistiheyttä.

Datalähteinä toimivat Twitter, Facebook, YouTube, Instagram, Reddit, noin 500 suomalaista keskustelupalstaa ja kymmeniä tuhansia suomalaisia blogereita. Vaalipulssi.fi -sivusto havainnollistaa puhutuimmat aiheet kunkin puolueen tai poliitikon osalta selkeässä näkyvässä.

Vaalikeskustelu on tullut jäädäkseen sosiaaliseen mediaan. Kyselytutkimusten rinnalle sosiaalinen media tarjoaa uuden tietolähteen kansan syvien rivien tunteiden analysointiin. Useat tahot ovat tart-

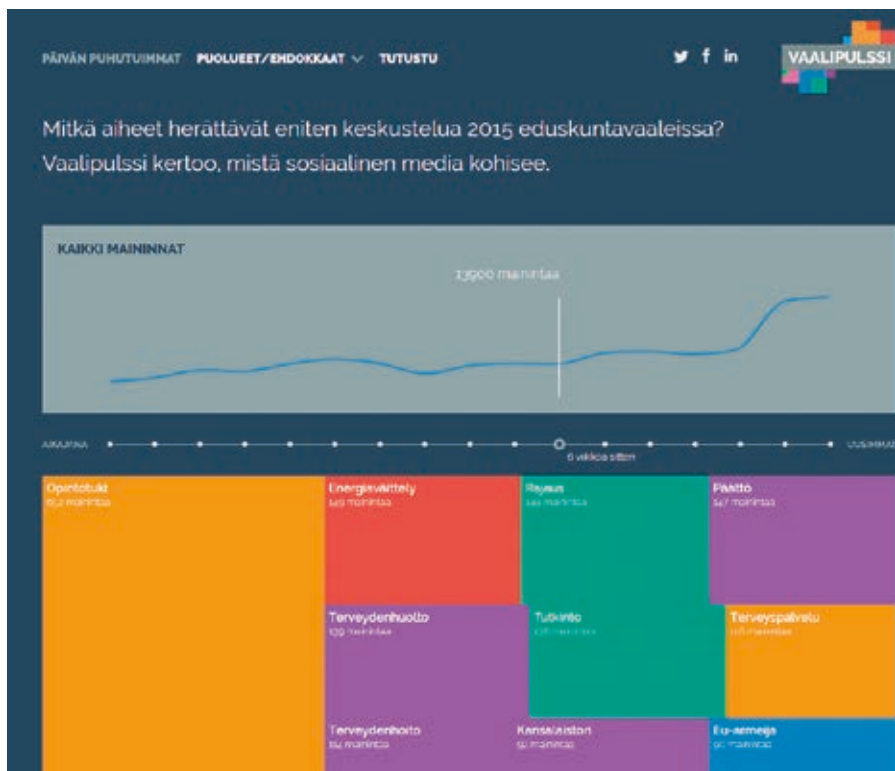
tuneet somedatan mahdollisuuksiin. Esimerkiksi Tampereen yliopiston COMET -tutkimuskeskus analysoi hankkeessaan vaalilähetysten aikaisten livetwiittien sisältöä. SN4 vuorostaan toteutti Iltasanomien juttusarjaan sentimenttianalyysia puolueita ja puoluejohtajia koskevista somekeskustelusta.

Vaalipulssi.fi -palvelun tuottajan, Big Data Solutions Oy:n, mukaan palvelu havainnollistaa, miten helppoa ja nopeaa isojen tietomassojen hyödyntäminen voi olla.

Tulosten tulkinta ja johtopäätösten teko some-datan pohjalta lienee jo astetta vaikeampaa. ■

Vaalipulssi.fi on Big Data Solutions Oy:n hanke, joka hyödyntää IBM-teknologiaa. Vaalipulssin analysoima tieto kootaan automatisoidusti sosiaalisesta mediasta ja valikoiduista keskustelupalstoilta. Vaalipulssi hakee ennalta määritellyille ehdokaslistalle kotimaisen somekeskustelun volyymin ja esiin nousseet aiheet.

Tieto tuotetaan Futusome Oy:n tarjoaman sosiaalisen median API:n ja sähköisen rajapinnan kautta.



Tuotanto sakkaa edelleen

Suomen kansantalouden kausitasoitettu tuotanto laski helmikuussa 0,7 prosenttia edellisluokauksesta.

Työpäiväkorjattuna tuotanto oli 0,5 prosenttia alemmalla tasolla kuin vuoden 2014 helmikuussa. Alkutuotanto kasvoi 1,6 prosenttia ja palvelut 0,7 prosenttia vuoden 2014 helmikuusta. Jalostus laski neljä prosenttia.

Jalostus sisältää teollisuuden ja rakentamisen. Palvelut sisältää kaupan, majoitus- ja ravitsemistoiminnan, liikenteen, liike-elämän palvelut sekä kiinteistö-, vuokraus- ja tutkimuspalvelut, rahoitus- ja vakuustoittoiminnan ja julkiset palvelut. Alkutuotannolla tarkoitetaan maariista-, metsä- ja kalataloutta.

Lähde: Tuotannon suhdannekuvaaja 2015 helmikuu, Tilastokeskus

Teollisuus taas syöksyyn

Teollisuustuotannon alamäen loiveneminen viime vuonna kääntyi uudelleen romahdukseksi alkuvuodesta. Koko teollisuuden työpäiväkorjattu tuotanto oli helmikuussa 5,1 prosenttia pienempi kuin vuoden 2014 helmikuussa. Tammi-helmikuun aikana teollisuustuotanto väheni 4,9 prosenttia edellisvuodesta.

Kausitasoitettu tuotanto väheni helmikuussa 0,4 prosenttia tammikuuhun verrattuna. Tammikuussa kausitasoitettu tuotanto väheni 2 prosenttia. Joulukuussa kausitasoitettu tuotanto väheni 0,5 prosenttia edelliseen kuukauteen verrattuna.

Kaivokset, kemia, metalli...

Teollisuustuotanto väheni helmikuussa kaikilla päätoimialoilla. Eniten tuotanto väheni kaivostoiminnassa runsaat 18 prosenttia. Kemianteollisuuden tuotanto väheni 5,3 prosenttia. Metalliteollisuuden tuotanto väheni helmikuussa 4,7 prosenttia.

Helmikuussa teollisuuden kapasiteetin käyttöaste oli 78,9 prosenttia, mikä oli 0,3 prosent-

tiyksikköä suurempi kuin vuotta aiemmin. Metsäteollisuuden kapasiteetista oli helmikuussa käytössä 84,3 prosenttia eli 2,9 prosenttiyksikköä vähemmän kuin vuoden 2014 helmikuussa. Metalliteollisuuden kapasiteetin käyttöaste oli helmikuussa 79,9 prosenttia, mikä oli 2,3 prosenttiyksikköä vuodentakaista suurempi.

Ei iloa tilauksistakaan

Teollisuuden uusien tilausten arvo oli helmikuussa 5,5 prosenttia pienempi kuin vuotta aikaisemmin. Tammi-helmikuun aikana tilaukset vähenivät 3,3 prosenttia edellisvuodesta.

Yritykset saivat helmikuussa uusia tilauksia vuodentakaista enemmän vain paperiteollisuudessa. Paperin, paperi- ja kartonkituotteiden valmistuksen toimialalla tilaukset kasvoivat 3,9 prosenttia vuodentakaisesta. Kemianteollisuudessa tilaukset vähenivät 6,6 prosenttia, metalliteollisuudessa 7,7 prosenttia ja tekstiiliteollisuudessa 9,8 prosenttia edellisvuodesta.

Lähde: Tilastokeskus, Teollisuustuotannon volyyymi-indeksi 2015, helmikuu & Teollisuuden uudet tilaukset 2015, helmikuu

Tuotannon suhdannekuvaaja*

	Syysk.	Lokak.	Marrask.	Jouluk.	Tammik.	Helmik.
	2014			2015		
Muutos edell. kk-sta, kausitasoitettu, ja työpäiväkorjattu, %	0,0	0,2	-0,6	-1,0	1,2	-0,7
Muutos edell. kk-sta, trendi, %	0,0	0,0	-0,1	-0,1	-0,1	-0,1
Vuosimuutos, alkuperäinen, %	2,0	0,1	-1,7	1,1	-2,0	-1,6
Vuosimuutos, työpäiväkorjattu, %	0,4	0,8	0,4	-1,3	0,1	-0,5

* Tilaston nimi oli aikaisemmin Kokonaistuotannon kuukausikuvaaja.

Kansantalous

	Volyymin muutos edellisestä vuodesta, %				
	2013*	2014*			
TOL 2008	4. nelj.	1. nelj.	2. nelj.	3. nelj.	4. nelj.
Huoltotase					
Bkt markkinahintaan	0,0	-0,1	0,0	-0,2	-0,2
Tavaroiden ja palvelujen tuonti	-0,7	-5,8	4,1	-2,8	-0,9
Kokonaistarjonta	-0,2	-1,4	0,8	-0,9	-0,4
Tavaroiden ja palvelujen vienti	-0,9	-3,8	1,5	0,5	0,5
Kulutusmenot	0,1	0,2	-0,5	0,0	0,1
julkiset	1,2	0,4	-0,7	0,4	0,6
yksityiset	-0,7	0,3	-0,3	-0,3	-0,4
Investoinnit	-5,2	-6,9	-4,6	-3,4	-5,5
julkiset	8,8	-0,1	2,0	-0,1	0,7
yksityiset	-8,3	-8,4	-6,3	-4,3	-6,9
Kokonaiskysyntä **	-0,3	-1,0	0,5	-1,2	-1,0
Toimialojen arvonlisäykset					
Maa-, metsä- ja kalatalous	5,9	4,4	-0,9	1,7	0,3
Koko teollisuus	1,1	-1,0	-0,2	-2,0	-2,1
Rakentaminen	-3,3	-3,2	-3,0	-4,8	-3,6
Kauppa	0,3	0,1	0,6	-0,1	-2,0
Kuljetus ja varastointi	-3,0	-0,4	-2,2	-1,1	-2,2
Kiinteistöalan toiminta	-0,4	1,7	1,6	1,8	1,9

* Ennakkotieto (kausitasoitettu ja työpäiväkorj. sarja)

** Pl. tilastollinen ero

Teollisuustuotannon volyyymi-indeksi

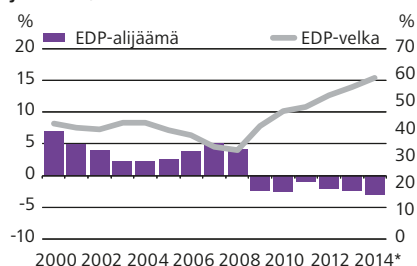
	Kausitasoitettu Helmik. 2015/ Helmik. 2014*	Työpäiväkorjattu Helmikuu 2015/ Helmikuu 2014
TOL 2008		
Volyymin muutos, %		
Koko teollisuus (B,C,D,E)	-0,4	-5,1
Kaivostoiminta ja louhinta (B)	3,7	-18,3
Teollisuus (C)	-0,2	-4,8
Elintarvikkeiden valmistus	1,6	-5,3
Juomien valmistus	5,3	0,0
Sahatavaran ja puu- ja korkkituotteiden valmistus	2,8	-2,9
Paperin, paperi- ja kartonkituotteiden valmistus	-2,4	-6,1
Painaminen ja tallenteiden jäljentäminen	-4,1	-18,2
Muiden ei-metallisten mineraalituotteiden valmistus	4,1	-4,7
Metallien jalostus	3,1	3,1
Metallituotteiden valmistus (pl. koneet ja laitteet)	-3,8	-6,9
Muiden koneiden ja laitteiden valmistus	1,1	-12,9
Huonekalujen valmistus	6,3	-6,1
Sähkö-, kaasu- ja lämpö- ja ilmastointihuolto (D)	-4,3	-5,3
Vesihuolto, viemäri- ja jätevesihuolto (E)	0,7	-0,4
Käyttötarkoituseräluokitus		
Raaka-aineet ja tuotantohyödykkeet	-0,4	-2,8
Investointitavarat	0,0	-7,3
Kestokulutustavarat	11,9	16,9
Muut kulutustavarat	-1,9	-12,1
Erikoisindeksi		
Elintarviketeollisuus	1,5	-4,6
Metsäteollisuus	-1,2	-5,3
Kemianteollisuus	0,5	-5,3
Sähkö- ja elektroniikkateollisuus	-3,7	-5,6
Kone- ja metallituoteteollisuus	1,6	-6,3

* Ennakkotieto

Alijäämä ylitti, velka vielä alitti viitearvon

Suomen julkisyhteisöjen alijäämä oli 3,2 prosenttia suhteessa bruttokansantuotteeseen vuonna 2014. Ennakkotietojen mukaan alijäämä ylitti näin EU:n vakaus- ja kasvusopimuksen mukaisen viitearvon, joka on kolme prosenttia suhteessa bkt:hen.

Suomen julkisyhteisöjen alijäämä (-) ja velka, suhteessa BKT:hen



Julkisyhteisöjen EDP-velka eli sulautettu bruttovelka sen sijaan pysyi täpärästi 60 prosentin viitearvon alapuolella. Velkaa oli 59,3 prosenttia suhteessa bruttokansantuotteeseen vuonna 2014.

Vuonna 2014 julkisyhteisöjen alijäämä eli kansantalouden tilinpidon mukainen nettoluotonotto oli 6,4 miljardia euroa. Vuoteen 2013 verrattuna julkisyhteisöjen alijäämä kasvoi viime vuonna 1,3 miljardia euroa. Valtionhallinnon alijäämä kasvoi edellisvuodesta 173 miljoonaa euroa ja paikallishallinnon 279 miljoonaa euroa.

Sosiaaliturvarahastot jakaantuvat työelä-

kelaitoksiin ja muihin sosiaaliturvarahastoihin. Työeläkelaitosten ylijäämä pieni edellisvuosien tapaan, nyt 252 miljoonalla eurola. Muut sosiaaliturvarahastot olivat alijäämisiä ensimmäistä kertaa vuoden 2009 jälkeen. Niiden rahoitusasema heikkeni 609 miljoonaa euroa edellisvuodesta ja oli -584 miljoonaa euroa vuonna 2014. Tähän vaikutti eniten työttömyysmenojen kasvu.

Lisää velkaa 8,4 miljardia

Julkisyhteisöjen sulautettu bruttovelka (EDP-velka) oli 121,1 miljardia euroa vuoden 2014 lopussa. Velka kasvoi vuoden aikana 8,4 miljardilla eurolla. Valtionhallinnon osuus sulautetun velan kasvusta oli 6,6 miljardia euroa, paikallishallinnon 1,5 miljardia ja sosiaaliturvarahastojen 0,3 miljardia.

Velkatiedot tarkentuvat seuraavan kerran kesäkuussa. Julkisyhteisöjen yli-/alijäämällä tarkoitetaan kansantalouden tilinpidon sektoritilien mukaista nettoluotonottoa/-ottoa, joka on kokonaistulojen ja kokonaismenojen erotus. Esimerkiksi paikallishallinnon tapauksessa käsite poikkeaa sektoriin kuuluvien kuntien ja kuntayhtymien tuloslaskelmien mukaisesta tilikauden yli-/alijäämästä. Keskeinen ero liittyy investointeihin, jotka kirjataan kansantalouden tilinpidossa menoiksi sellaisenaan.

Takaukset osana velkaa

Julkisyhteisöjen tilinpidossa Euroopan väliaikainen rahoitusvakausväline ERVV käsitellään Eurostatin päätöksen perusteella siten, että

ERVV:n velanotto tilastoidaan osaksi takauksia myöntäneiden maiden julkista velkaa. Avunsaajamaan ERVV:itä saama laina tilastoidaan ikään kuin se olisi saatu takaukset myöntäneiltä euromailta, ja nämä maat puolestaan ovat vastaavan summan velkaa ERVV:lle. Tämän seurauksena takaukset myöntäneillä euromailta EDP-bruttovelka kasvaa, mutta nettovelka pysyy ennallaan, koska maille syntyy yhtä iso saaminen avunsaajamaalta. Kunkin maan julkisyhteisöjen velkaan kirjattava osuus ERVV:n myöntämistä lainoista lasketaan jakamalla avunsaajamaille annettu laina sovitulla jakosuudella (perustuu kuhunkin tukipakettiin osallistuvien jäsenmaiden osuuksiin EKP:n pääomasta).

Vuoden 2014 lopussa Suomen julkisyhteisöjen velkaan on kirjattu 3 611 miljoonaa euroa velkaa ERVV:n myöntämien lainojen perusteella. Vastaavaa käsittelyä ei Euroopan vakausmekanismin (EVM) kohdalla tehdä.

Valtionhallinnon EDP-velka eroaa käsitteellisesti Valtiokonttorin julkaisemasta valtionvelasta (95,1 mrd euroa 2014). Valtion EDP-velkaan luetaan mukaan lisäksi väliaikaisen rahoitusvakausväline ERVV:n myöntämät lainat avunsaajamaille, johdannaisopimuksiin liittyvät saadut vakuustalletukset, ydinjätehuoltorahaston pääoma, valtion elinkaarihankkeiden investointeista syntyneet velat ja kierrossa olevat kolikot. Kansantalouden tilinpidossa valtionhallinto on myös käsitteellisesti laajempi kuin budjetti- ja rahastotalous, joiden lisäksi siihen luetaan mm. yliopistot, Solidium, Yleisradio ja Senaatti-kiinteistöt. Arvostusperiaatteena molemmissa velkakäsitteissä on nimellisarvo, jossa valuutanvaihtosopimusten vaikutus on huomioitu.

Lähde: Julkisyhteisöjen alijäämä ja velka, Tilastokeskus

Liiketoiminnan suhdannekuvaajat

	Liikevaihtokuvaaja		Palkkasummakuvaaja	
	%-muutos	%-muutos	%-muutos	%-muutos
2010=100				
TOL 2008	Loka-jouluk. 2014/2013	Tammi-jouluk. 2014/2013	Marrask-tammik. 2015/2014	Tammik-jouluk. 2014/2013
Teollisuus	-2,7	-1,7	-2,0	-0,7
Rakentaminen	4,0	2,1	0,4	1,0
Kauppa	-1,9	-0,7	0,6	0,6
Muut palvelut	4,0	2,8	2,1	1,3

Kaupan liikevaihto

	Tammikuu 2015/2014	Tammi-jouluk. 2014/2013
TOL 2008	%-muutos	%-muutos
Kauppa yhteensä (G)	-8,8	-0,7
Moottoriajoneuvojen kauppa ja korjaus (45)	-11,3	2,3
Tukkukauppa (46)*	-12,0	-1,2
Vähittäiskauppa (47)*	-2,2	-0,9
tavaratalokauppa	-1,3	-2,1
päivittäistavarakauppa	-1,4	0,4

* Pl. Moottoriajoneuvojen kauppa

Konkurssit

TOL 2008	Tammi-helmik. 2015	Tammi-helmik. 2014	Vuosi-muutos, %
Vireille pantuja konkurssseja yht.¹⁾	440	565	-22,1
Maa-, metsä- ja kalatalous	16	11	45,5
Teollisuus, kaivostoiminta, energia- ja vesihuolto	60	69	-13,0
Rakentaminen	97	117	-17,1
Kauppa	69	119	-42,0
Kuljetus ja varastointi	38	55	-30,9
Majoitus- ja ravitsemistoiminta	35	37	-5,4
Muut palvelut	121	151	-19,9
Toimiala tuntematon	4	6	-33,3
Henkilökunta yhteensä¹⁾	2 042	2 886	-29,2

1) Sisältää yhtiöt, yhteisöt ja yrittäjät

Enemmän työttömiä

Työttömiä oli helmikuussa työvoimatutkimuksen mukaan 268 000, mikä oli 27 000 enemmän kuin vuosi sitten. Työttömiä miehiä oli 142 000 ja naisia 126 000 henkeä.

Työttömyysaste oli 10,1 prosenttia, kun se edellisvuoden helmikuussa oli 9,1 prosenttia. Miesten työttömyysaste oli 10,5 ja naisten 9,7 prosenttia.

Työllisiä oli 13 000 vähemmän kuin edellisen vuoden helmikuussa. Työllisten määrä väheni yksityisellä sektorilla ja kasvoi julkisella sektorilla vuoden 2014 helmikuuhun verrattuna.

Kaksi kolmesta työikäisestä töissä

Työllisyysaste eli työllisten osuus 15–64-vuotiaista oli helmikuussa 66,6 prosenttia, kun se vuotta aiemmin oli 66,9 prosenttia. Miesten työllisyysaste laski edellisen vuoden helmikuusta 0,6 prosenttiyksikköä 66,7 prosenttiin. Naisten työllisyysaste pysyi lähes ennallaan edellisvuoden helmikuuhun verrattuna ollen 66,6 prosenttia.

Työvoiman ulkopuolella olevia oli vuoden 2015 helmikuussa 1 441 000 henkeä eli 5 000 vähemmän kuin vuotta aiemmin. Työvoiman ulkopuolella olevista oli piilotyöttömiä 121 000, mikä oli 14 000 vähemmän vuoden 2014 helmikuuhun verrattuna.

Lähde: Työvoimatutkimus 2015, helmikuu. Tilastokeskus

Opiskelijoiden työssäkäynti väheni

Työssäkäyvien opiskelijoiden osuus väheni vajaa kaksi prosenttiyksikköä vuonna 2013 edellisestä vuodesta. Opiskelijoista kävi opintojen ohella töissä hieman yli puolet. Vuoden 2004 jälkeen osuus on ollut korkeimmillaan vuonna 2008, jolloin lähes kuusi kymmenestä opiskelijasta kävi töissä opintojen ohessa.

Työssäkäynti oli yleisintä yliopisto- ja ammattikorkeakouluopintojen yhteydessä. Yliopisto-opiskelijoista oli opintojen ohessa työsuhde 58 prosentilla ja ammattikorkeakoulu-opiskelijoista 56 prosentilla. Eniten vuonna 2013 väheni työssäkäyvien osuus toisen asteen ammatillisen koulutuksen opiskelijoista, runsas 3 prosenttiyksikköä edellisvuodesta. Toisen asteen ammatillisen koulutuksen opiskelijoista kävi työssä opintojen ohessa joka toinen opiskelija.

Naiset kävivät miehiä yleisemmin opintojen ohella työssä, naisista 56 prosentilla ja miehistä 48 prosentilla oli työsuhde opintojen ohessa.

Keskeyttäminen lisääntyi vain yliopistoissa

Tutkintoon johtavan koulutuksen opiskelijoista 5,6 prosenttia keskeytti opinnot eikä jatkanut missään tutkintoon johtavassa koulutuksessa lukuvuoden 2012/2013 aikana. Lukio-koulutuksessa keskeyttämisprosentti oli 3,4, nuorille suunnatussa ammatillisessa koulutuk-

nessa 8,5, ammattikorkeakoulukoulutuksessa 8,3 ja yliopistokoulutuksessa 6,9 prosenttia.

Keskeyttäminen väheni edellisvuoteen verrattuna kaikilla koulutussektoreilla paitsi yliopistokoulutuksessa, jossa se lisääntyi.

Jos koulutussektoria vaihtaneita, esimerkiksi lukiokoulutuksesta ammatilliseen koulutukseen siirtyneitä, ei lasketa keskeyttäneiksi, keskeyttämisprosentit pienenevät. Useimmin koulutussektoria vaihtoivat ammattikorkeakoulu- ja lukiokoulutuksen opiskelijat.

Miehet keskeyttivät koulutuksen omassa koulutussektorissaan naisia useammin lukuun ottamatta ammatillista koulutusta, jossa naiset keskeyttivät useammin kuin miehet. Ammatillisessa koulutuksessa opiskelleet miehet sen sijaan keskeyttivät tutkintoon johtavan koulutuksen kokonaan useammin kuin naiset.

Lähde: Koulustilastot. Tilastokeskus

Öljy ja asunnot pitivät inflaation miinuksella

Kuluttajahintojen vuosimuutos pysyi maaliskuussa -0,1 prosentissa. Inflaatio pysyi negatiivisena edellisvuoden maaliskuuhun verrattuna erityisesti polttonesteiden sekä omistus-asumisen hintojen laskusta johtuen.

Maaliskuussa kuluttajahintoja nosti vuodessa eniten vuokrien, tupakan ja kerrostalojen kunnossapitopalveluiden kallistuminen. Laskua edellisvuoteen verrattuna eniten oli polttonesteiden, omistus-asumisen sekä maito-

Työttömyys

	Helmikuu 2014/2015		4. nelj. 2013/2014	
	Helmikuu 2015	%-muutos	4. nelj. 2014	%-muutos
Työttömät, 1 000 henkeä	268	11,1	223	9,7
Miehet	142	2,9	126	15,2
Naiset	126	22,1	96	3,1
15–24-vuotiaat	75	6,4	55	16,6
Työttömyysaste, %	10,1	1,0 yks.	8,4	0,7 yks.
Miehet	10,5	0,4 yks.	9,2	1,1 yks.
Naiset	9,7	1,6 yks.	7,6	0,2 yks.
Työttömyysaste ikäryhmittäin, %				
15–24-vuotiaat	25,5	2,6 yks.	18,8	2,7 yks.
25–34-vuotiaat	10,8	2,2 yks.	9,1	0,3 yks.
35–44-vuotiaat	7,1	0,7 yks.	6,1	0,1 yks.
45–54-vuotiaat	7,5	0,5 yks.	6,6	0,7 yks.
55–64-vuotiaat	8,4	0,0 yks.	7,7	1,0 yks.
15–64-vuotiaat	10,4	1,0 yks.	8,6	0,7 yks.
Työttömyysaste aluehallintovirastojen (AVI) mukaan, %				
Etelä-Suomen AVI	9,4	1,2 yks.	7,8	1,0 yks.
Lounais-Suomen AVI	10,5	1,0 yks.	9,3	2,2 yks.
Itä-Suomen AVI	8,9	-1,9 yks.	9,1	-0,2 yks.
Länsi- ja Sisä-Suomen AVI	10,0	0,2 yks.	8,1	0,0 yks.
Pohjois-Suomen AVI	13,4	2,4 yks.	10,5	0,4 yks.
Lapin AVI yks.	9,0	-1,0 yks.

Työllisyys

	Helmikuu 2014/2015		4. nelj. 2013/2014	
	Helmikuu 2015	%-muutos	4. nelj. 2014	%-muutos
1 000 henkeä				
15–74-vuotias väestö	4 101	0,2	4 099	0,2
Työvoima yhteensä	2 660	0,5	2 640	0,6
Työvoimaosuus, %	64,9	0,2 yks.	64,4	0,2 yks.
Työllisyysaste (15–64-vuotiaat), %	66,6	-0,3 yks.	67,4	-0,2 yks.
Miehet	66,7	-0,6 yks.	68,0	-0,1 yks.
Naiset	66,6	0,1 yks.	66,8	-0,2 yks.
Työlliset	2 392	-0,5	2 418	-0,2
Miehet	1 215	-1,2	1 244	0,1
Naiset	1 177	0,1	1 174	-0,5
Työlliset ammattiaseman mukaan				
Yrittäjät ja yrittäjäperheenjäsenet	330	2,1	346	5,3
Palkansaajat	2 062	-1,0	2 072	-1,0
Palkansaajat työsuhteen mukaan				
Jatkuva työsuhde	1 758	-1,8	1 778	-0,5
Määräaikainen työsuhde	304	4,0	294	-4,2
Palkansaajat työajan mukaan				
Kokoaikatyölliset	1 734	-0,8	1 751	-1,2
Osa-aikatyölliset	329	-1,5	321	-0,2
Työvoiman ulkopuolella olevat	1 441	-0,4	1 459	-0,5

Lisätietoja: www.tilastokeskus.fi/tyovoimatutkimus

taloustuotteiden kuluttajahinnoissa. Suurimmat nousut ja laskut kohdentuivat maaliskuussa samoihin hyödykkeisiin kuin helmikuussa.

Helmikuusta maaliskuuhun kuluttajahinnat nousivat 0,3 prosenttia, mikä johtui lähinnä vihannesten ja hedelmien sekä bensiinin hintojen noususta.

Myös euroalueella inflaatio pysyi pakkasen puolella. Yhdenmukaistetun kuluttajahintaindeksin ennakkotietojen mukaan euroalueen inflaatio oli maaliskuussa -0,1 prosenttia. Helmikuussa se oli -0,3 prosenttia. Suomen vastaava inflaatio oli maaliskuussa 0,0 prosenttia.

Yhdenmukaistetussa kuluttajahintaindeksissä ei ole mukana omistusasumista, rahapelejä, kulutus- ja muiden luottojen korkoja, omakotitalon palovakuutusta eikä ajoneuvovero. Yhdenmukaistettuun kuluttajahintaindeksiin kuuluvat kulutuserät ja laadintasäänöt on määritelty EU-asetuksiin.

Eurostatin arvio euroalueen inflaatiosta perustuu jäsenmaiden ennakkotietoihin ja energian hinnankehitykseen.

Lähde: Kuluttajahintaindeksi, Tilastokeskus

Eläkeläistalouksien tuloissa suuret erot

Eläkeläistalouksien tulot ovat keskimäärin 65 prosenttia palkansaajatalouksien tuloista. Eläkeläistalouksista selvästi suurituloisimpia ovat entiset ylempien toimihenkilöiden taloudet, joiden tulot ylittävät palkansaajatalouksien tulotason.

Heikoimmassa asemassa ovat ne eläkeläistaloudet, joiden pääasiallisesta toimeentulosta vastaa kansaneläkeläinen tai entinen maanviljelijä. Näissä ryhmissä kotitalouksien tulot ovat vain noin puolet keskivertopalkansaajatalouden tuloista.

Eläkkeellä olevien ylempien toimihenkilöiden tulot olivat vielä 1990-luvulla ja 2000-luvun alussa selvästi palkansaajatalouksien tuloja suuremmat, mutta palkansaajatalouksien tuloaso kohosi pitkään jatkuneen talouskasvun aikana 2000-luvun puolivälissä muutamaksi vuodeksi eläkkeellä olevien ylempien toimihenkilöiden tulotason yläpuolelle. Muiden eläkeläistalouksien tuloaso jää selvästi palkansaajatalouksien tulojen alapuolelle.

Eläkeläiskotitalouksien tulot eivät ole yhtä herkkiä suhdannevaihteluille kuin työikäisten tulot. Sekä 1990-luvun että viime vuosien taloustaantumien aikoina on eläkeläistalouksien suhteellinen asema parantunut palkansaajatalouksiin verrattuna. Eläkkeisiin ei ole taloudellisesti vaikeinaan aikoina kohdistettu suoranaista leikkauksia.

Lähde: Tulonjakotilasto 2013, Tilastokeskus

Vanhoiden asuntojen ja omakotitalojen hinnat*

Kerrostaloasuntojen hinnat ja hintaindeksi 2005=100	4. nelj. 2013	1. nelj. 2014**	2. nelj. 2014**	3. nelj. 2014**	4. nelj. 2014**
Koko maa, €/m ²	2 409	2 407	2 454	2 437	2 410
Nimellishintaindeksi	138,1	137,7	138,6	137,7	137
vuosimuutos, %	2,4	1,4	0,4	-0,6	-0,8
Pääkaupunkiseutu, €/m ²	3 721	3 699	3 792	3 743	3 725
Nimellishintaindeksi	148,5	148,8	150,3	148,1	148,2
vuosimuutos, %	3,3	2,3	1,6	-0,7	-0,2
Muu maa, €/m ²	1 682	1 691	1 713	1 714	1 682
Nimellishintaindeksi	127,7	126,8	127,1	127,3	125,9
vuosimuutos, %	1,5	0,3	-0,9	-0,4	-1,4
Omakotitalojen nimellishintaindeksi 2005=100					
Koko maa	120,8	121,7	121,7	120,9	119,3
vuosimuutos, %	-2,9	-3,3	-2,1	-0,2	-1,2
Omakotitalotonttien nimellishintaindeksi 2005=100					
Koko maa	143,0	147,9	156,4	146,5	153,6
vuosimuutos, %	-3,2	-0,6	7,5	2,2	7,4

* Vanhalla asunnolla tarkoitetaan asuntoa, joka ei ole valmistunut tarkasteluvuonna tai sitä edeltävänä vuonna

** Ennakkotieto

Kuluttajahintaindeksi (2010=100)

	Maalisk. 2015	Kuukausi- muutos, %	Vuosi- muutos, %
Kokonaisindeksi	108,9	0,3	-0,1
Elintarvikkeet ja alkoholittomat juomat	116,8	0,0	-1,6
Alkoholijuomat, tupakka	115,9	-0,1	2,0
Vaatetus ja jalkineet	105,5	7,0	0,6
Asuminen, vesi, sähkö, kaasu ja muut polttoaineet	110,3	0,1	0,9
Kalusteet, kotitalouskoneet ja yleinen kodinhoito	106,6	0,3	-0,1
Terveys	108,6	0,1	4,3
Liikenne	106,4	0,7	-2,4
Viestintä	83,4	-1,7	-5,3
Kulttuuri ja vapaa-aika	100,8	-0,7	-0,9
Koulutus	112,3	0,0	-0,5
Ravintolat ja hotellit	115,1	0,0	1,4
Muut tavarat ja palvelut	110,9	-0,3	0,7
Elinkustannusindeksi 1951:10=100	1908		

Palkansaajien ansiotasoindeksi ja ansioiden vuosimuutokset

Palkansaajaryhmä / sektori	2010=100	Vuosi- muutos	Kuukausiansiot ¹⁾ , €**
	4. nelj. 2014	%	4. nelj. 2014*
Yhteensä	110,4	1,3	3 327
Tuntipalkkaiset	109,8	1,4	2 697
Kuukausipalkkaiset	110,5	1,3	3 463
Miehet	109,7	1,3	3 637
Naiset	110,9	1,3	3 021
Sektori			
Yksityinen ja muut	110,3	1,4	3 410
Kunnat	110,0	1,1	3 001
Valtio	111,8	1,1	3 732
Muut	111,0	1,4	..

¹⁾ Palkansaajien keskimääräiset kuukausiansiot

* Ennakkotieto

** Reaaliansioiden vuosimuutos oli ennakkotietojen mukaan 4. neljänneksellä 0,5 prosenttia.

Tuottajahintaindeksit (2010=100)

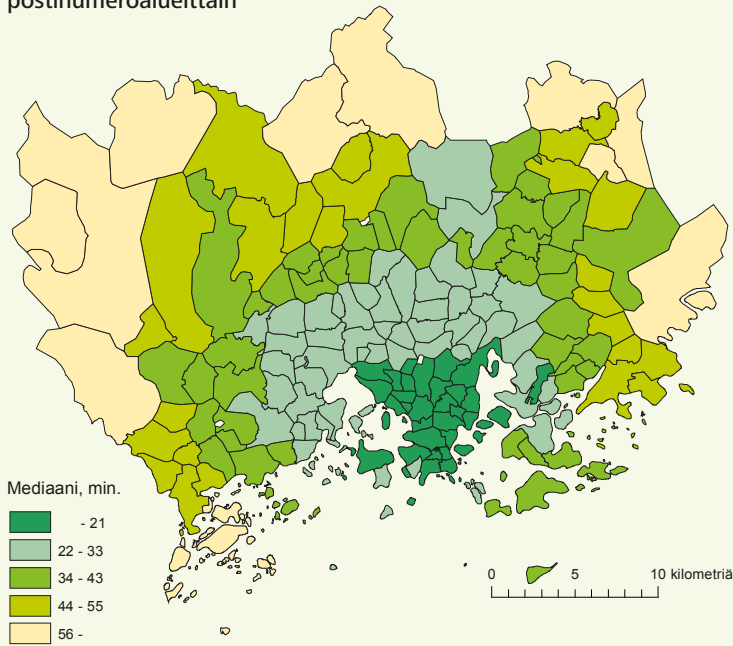
	Helmikuu 2015	Kuukausi- muutos, %	Vuosi- muutos, %
Teollisuuden tuottajahintaindeksi	104,0	1,0	-1,8
Teollisuuden tuottajahintaindeksi, kotimaiset tavarat	106,4	1,1	-2,3
Teollisuuden tuottajahintaindeksi, vientitavarat	101,1	1,0	-1,1
Vientihintaindeksi	101,2	1,0	-1,2
Tuontihintaindeksi	101,5	1,9	-5,3
Kotimarkkinoiden perushintaindeksi	105,7	1,1	-3,0
Verollinen kotimarkkinoiden perushintaindeksi	108,8	1,1	-2,5

Rakennuskustannusindeksi (2010=100)

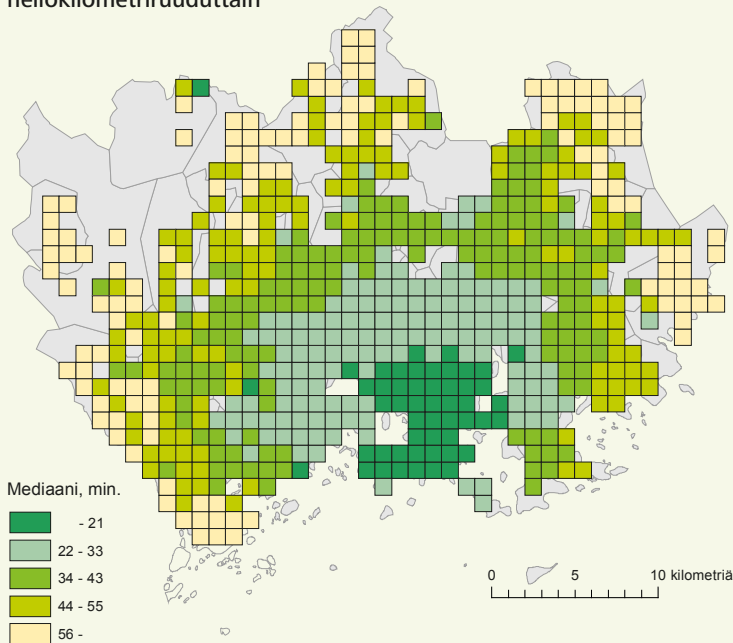
	Maaliskuu 2015	Kuukausi- muutos, %	Vuosi- muutos, %
Ammattimainen uudisrakentaminen			
Kokonaisindeksi	108,3	-0,1	0,5
Työ	106,7	0,0	1,2
Tarvikkeet	107,9	-0,1	0,2
Palvelut	115,2	0,1	-0,1

Entä jos kaikki polkisivat töihin?

Pyöräillen kotoa töihin kuluva matka-aika pääkaupunkiseudulla, postinumeroalueittain



Pyöräillen kotoa töihin kuluva matka-aika pääkaupunkiseudulla, neliökilometriruuduttain



Kuvioiden lähde: Tilastokeskus

PASI PIELA

Tilastokeskuksessa on viime aikoina tutustuttu erilaisiin saavutettavuuslaskelmiin Liikenneviraston digitaalista tieverkkoa Digiroad hyödyntäen. Keskeisenä soveltamisalueena ovat olleet työmatkat matka-aikoinen kaikille työssäkäyville.

Useiden hallinnollisten aineistojen muodostamaa väestötietojen tietovarastoa käytettäessä on tyyppioletuksena yksityisauton käyttö kotoa päätoimiselle työpaikalle. Todellisuus on monen kohdalla kuitenkin toinen: Julkinen liikenne palvelee eri tavoin eri osissa maata, ja auto voi muistakin syistä olla tarpeeton. Yhä useammilla vaakakupissa painavat myös elämänlaatu ja vihreämmät liikumismuodot etätyöstä puhumattakaan.

Pyöräilytalous, cycling economy, työllistää Euroopassa jo noin 655 000 ihmistä, kertoo The Guardian -lehdessä vii-



me marraskuussa julkaistu Euroopan pyöräilynedistämisyjärjestöjen kattojärjestön ECF:n tutkimus. Se on enemmän kuin Euroopan kaivos- ja louhintateollisuudessa ja jopa 87 prosenttia enemmän kuin metalliteollisuuden suoraan työllistämät. Jos pyöräilyn ”kolmen prosentin osuus matkan teosta” tuplaantuisi, pyöräilyn työllistämien lukumäärä nousisi Euroopassa yli miljoonaan vuoteen 2020 mennessä.

Entä jos kaikki pyöräilisivät töihin? Miten paljon siihen kuluisi aikaa?

Digiroadia – ja sen sisältämiä tien toiminnallisuutta koskevia luokituksia – hyödyntäen oheisiin karttakuviin on kuvattu pääkaupunkiseudun asukkaiden työmatkat tilanteessa, missä kaikki käyttäisivät polkupyörää työmatkailuunsa ja valitsisivat lyhimmän reitin huomioiden luonnollisesti myös kevyenliikenteen väylät.

Kohteina ovat siis espoolaiset, helsinkiläiset, kauniaislaiset ja vantaalaiset työlliset. Tarkastelussa on käytetty HSL:n Reittioppaan pyöräilylle tarjoamaa oletusnopeutta, joka on 17 kilometriä tunnissa.

Tilastokeskuksen uusimmat työpaikatiedot ovat vuoden 2012 lopulta. Ne

[Kommentoi blogissa:](#)

tietotrenditblogi.stat.fi

JULKAISTU BLOGISSA 10.3.2015

PÄÄKAUPUNKISEUDULLA
JOKA TOINEN PYÖRÄILISI TÖIHIN
PUOLESSA TUNNISSA.



Tuula Roos / Vastavalo.fi

kattavat 91,2 prosenttia kaikista maamme työssäkävivistä. Työskentelytapojen muutokset kuten vuokratyö voivat vaikeuttaa tulevaisuudessa päätoimisen työpaikan määrittelyä entisestään.

Työmatkajakaumat ovat tyypillisesti erittäin vinoja. Sen vuoksi niiden tarkastelussa suositaan fraktiilien kuten mediaanin käyttöä.

Niiden pääkaupunkiseutulastaisten, joiden työmatkan pituus lyhintä reittiä pitkin on enintään 200 kilometriä, polkupyörämatkaan kuluu keskimäärin aikaa noin 41 minuuttia. Vastaava mediaani on puolestaan 30 minuuttia ja ilman edellä mainittua rajausta samoin noin 30 minuuttia (0,3 minuuttia korkeampi).

Puolella työssä käyvistä väestöstä matka-aika on siis alle puoli tuntia. 66 prosentilla se on edelleen alle 41 minuutin keskiarvon.

Ikäryhmien väliset erot ovat pääkaupunkiseudulla vähäisiä: alle 35-vuotiailla pyöräillen työmatkaan kuluvan ajan mediaani on noin 28 minuuttia ja seuraavissa ikäryhmissä (35-44- ja yli 44-vuotiaat) vastaavasti vain 3–4 minuuttia korkeampi. Miehillä on postinumeroalueittain tarkasteltuna tyypillisesti muutamaa minuuttia pidemmät mediaanimatkat kuin naisilla.

Saariston osalta tarkastelu on tavalista haasteellisempää. Suomenlinnassa asuvien tai työssäkävien kohdalla on huomioitu lauttamatkaan kuluvaksi ajaksi 15 minuuttia HSL:n reittioppaan arvion mukaisesti. Santahaminasta työssäkävien osalta ei ole voitu huomioida puolustusvoimien mahdollisesti tarjoamia meriteitse kuljettavia työmatkoja.

Karttakuvassa 1 on postinumeroalueittain ja karttakuvassa 2 neliökilometrin ruuduille yleistettynä mediaanimatkaajat (ilman edellä mainittua rajausta pitkille työmatkoille). Luokitus on valittu postinumeroalueille tilastollisesti optimoituna. Alle 10 työssäkävää sisältävät ruudut on poistettu.

Karttakuvista näemme, miten matkaajat selvästi pitenevät, kun etäisyys Helsingin keskustasta kasvaa. Jätkäsaarella on uutena asuinalueena ollut tilastointihetkellä naapurialueitaan pidemmät työmatkat. ■



TIETO&TRENDIT – TALOUS- JA HYVINVOINTIKATSAUS

- taustoittaa, tulkitsee ja sukeltaa numeroiden taakse
- tarjoaa vahvaa tilasto-osaamista yhdistettynä kiinnostavaan analyysiin
- tarttuu talouden, hyvinvoinnin ja ympäristön teemoihin

SEURAA MEITÄ VERKOSSA: [TIETOTRENDIT.STAT.FI](https://tietotrendit.stat.fi)

LÖYDÄT KIINNOSTAVIA NÄKÖKULMIA JA LUKIJAKESKUSTELUJA

ASiantuntijablogissa: [TIETOTRENDITBLOGI.STAT.FI](https://tietotrenditblogi.stat.fi)



Voit tilata julkaisun joko tilauslomakkeella (qr-koodi)
tai soittamalla numeroon 020 450 05, Edita Publishing Oy.

Kestotilauksen vuosihinta on 98 euroa.

Tilastokeskus 