Valtakunnallinen henkilöliikennetutkimus Seutuotokset – haarukointia otoskoon vaikutuksesta

WSP Finland Oy *26.3.2019*

Virpi Pastinen, Hannu Lehto, Annika Rantala

Käytännössä havainto-määrätarve on samaa suuruusluokkaa mille tahansa alueelle (seutu, kunta, osa-alue), jonka tuloksia halutaan tarkastella, riippumatta sen väestömäärästä.

Seuturaportin taso, seudun 6 kuntaa (+ jotain varmaan Orivedeltä ja Vesilahdestakin):
- 3000 – 6000 vastausta
- Otos 9 000 – 18 000 henkeä

-> 7000 hengen otos ei vielä mahdollista tätä, 9000 alkaisi olla niillä rajoilla että voi onnistua jos vastausaste hyvä

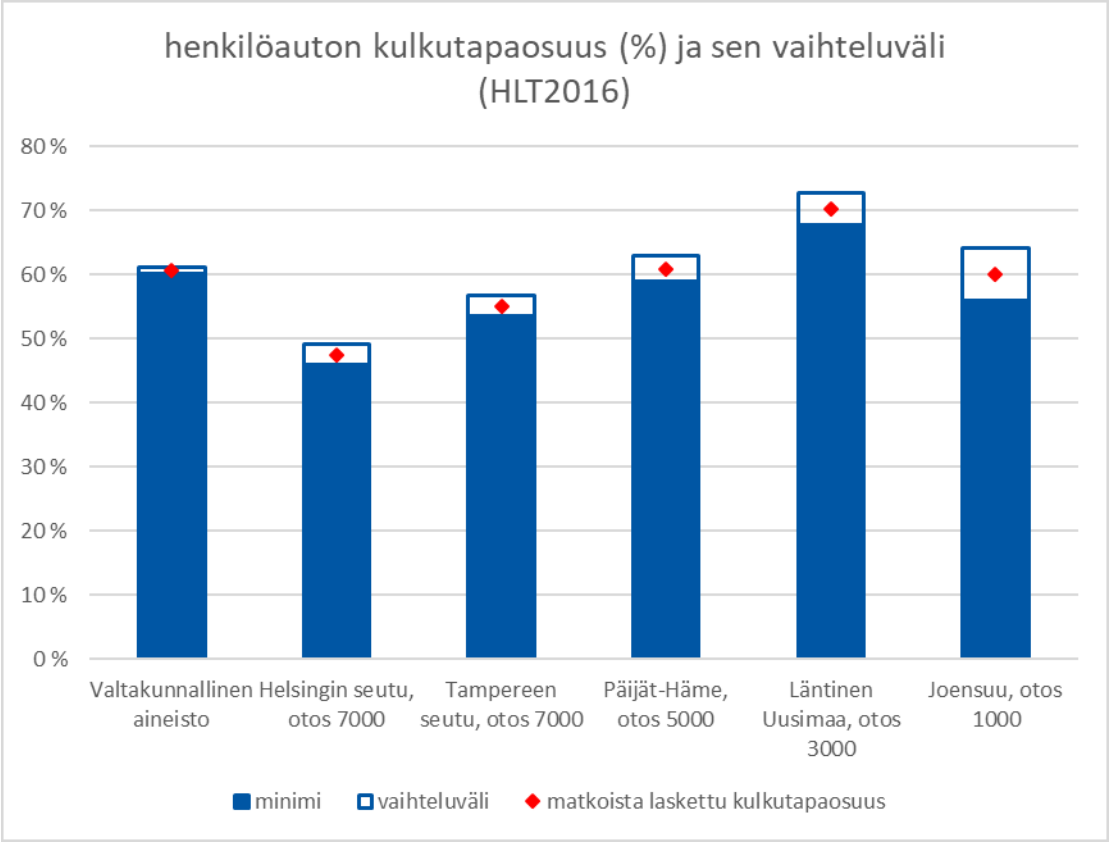
Seuturaporttien tapaiset raportit esim. kunnittain

- Jokaisesta kunnasta, josta vastaava raportti haluttaisiin tehdä, tarvitaan vähintään 500-700, mutta mielellään noin 1000 vastausta (luotettavuus on aina sitä parempi, mitä enemmän vastauksia)
- **Otoskoko** täytyy arvioida oletetun vastausasteen perusteella.
  - *HLT2016: vastausaste 45 prosenttia*
  - *HLS:n liikkumistutkimus 2019: vastausaste noin 30 % (mutta eri metodi)*
  - *Tuntuma on, että vastausasteet laskevat koko ajan. Tämä on yleismaailmallinen haaste.*
  - *→ Kannattaa käyttää varmuudeksi 35 prosentin vastausasteolettamaa*
- → tämä tarkoittaisi karkeasti 1500-3000 hengen otosta / kunta
- Vastaavasti jos jostain kunnasta halutaan tuloksia osa-alueittain (kuten esim. Tampereen kaupunkikohtaisessa raportissa), olisi vastauksia hyvä olla vähintään 500 osa-aluetta kohti. Tämä tarkoittaisi siis n. 1500 hengen otosta per osa-alue.

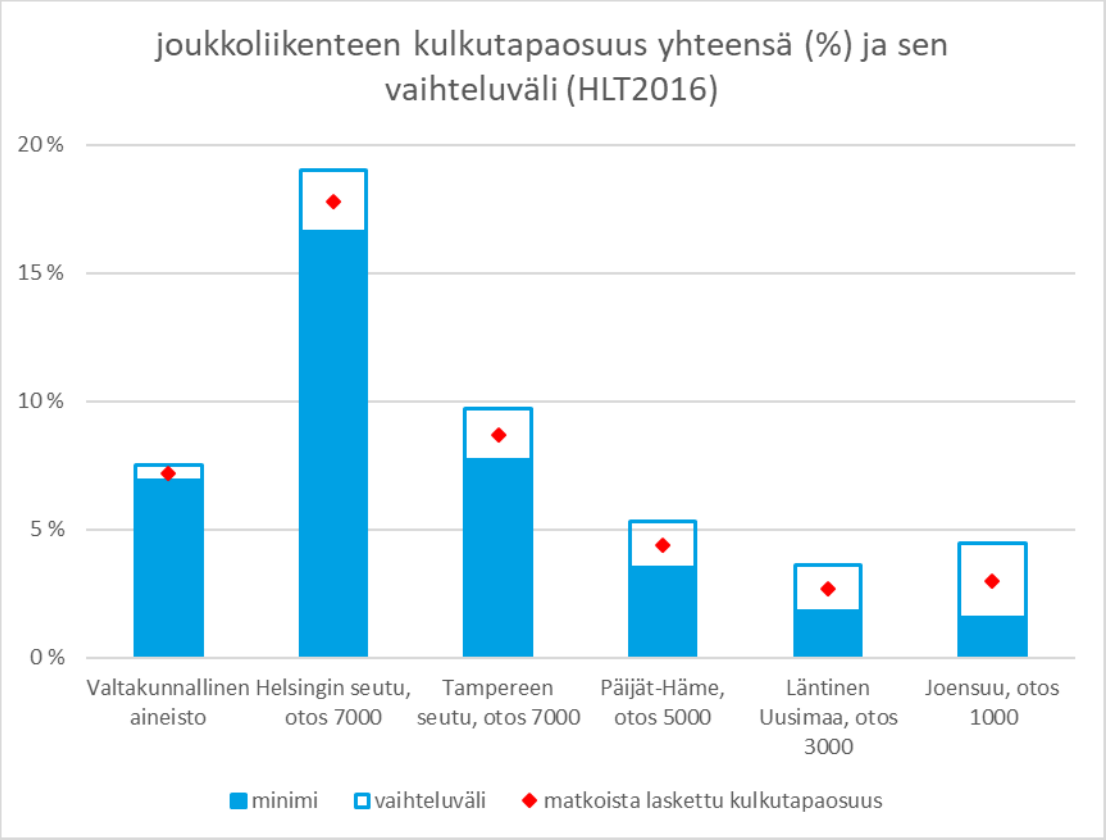
HUOM! Tässä annetut vastausmäärät ovat karkeita arvioita, eivätkä perustu todellisiin laskelmiin. Suuruusluokka tästä kuitenkin jonkin verran saa. Virhemarginaalit jäävät vielä näilläkin vastausmääräarvioilla melko suuriksi; nämä on arvioitu lähinnä siitä näkökulmasta, millä vastausmäärällä vielä uskaltaisi ainakin näyttää aluekohtaisia tuloksia.

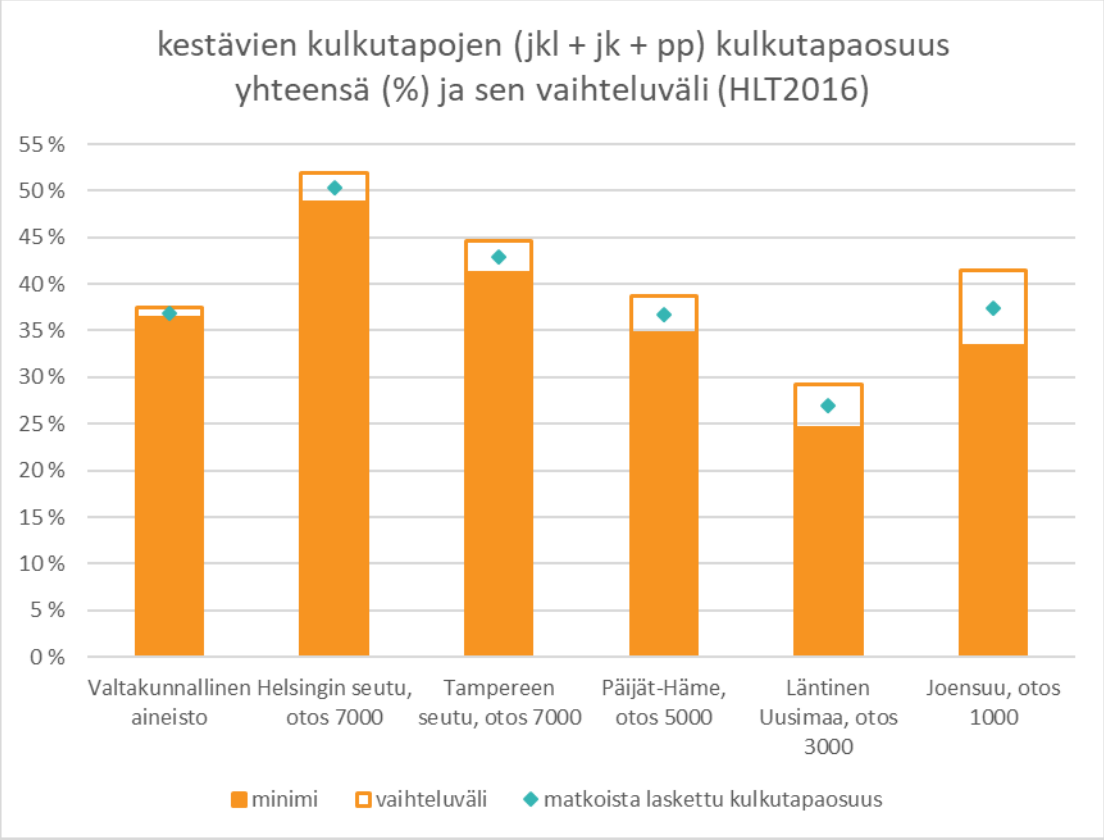
Otoskoon vaikutus luottamusväliin

Otoskoon vaikutusta tulosten luotettavuuteen on havainnollistettu vertaamalla HLT2016:n seuturaporteissa esitettyjä kulkutapaosuuksien 95 %:n luottamusvälejä seuduilla, joilla oli erikokoiset otokset



Otoskoon vaikutus luottamusväliin

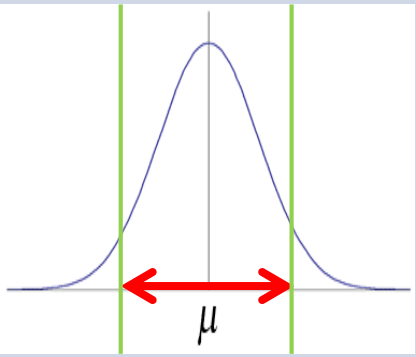
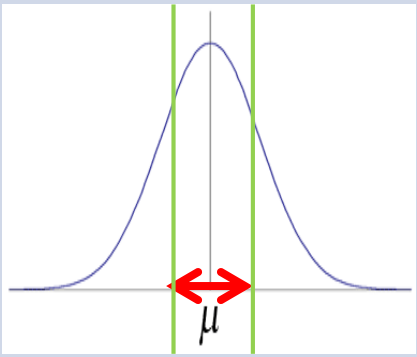




HLT2016:n tutkimus-suunnitelmavaiheessa tehtiin tällaisia laskelmia ja havainnollistuksia otoskoon vaikutuksesta

Tunnuslukujen luotettavuus

- Henkilöliikennetutkimus on otantatutkimus
- Otantatutkimusta käytetään kustannussyistä
- Koska ihmisten liikkumisessa on eroja, aiheutuu otannasta virhemahdollisuus tuloksiin.

PIENI OTOS	SUURI OTOS
	
Suuri virhemarginaali -> epätarkka tulos	**Pieni virhemarginaali -> tarkka tulos**

Esimerkki I

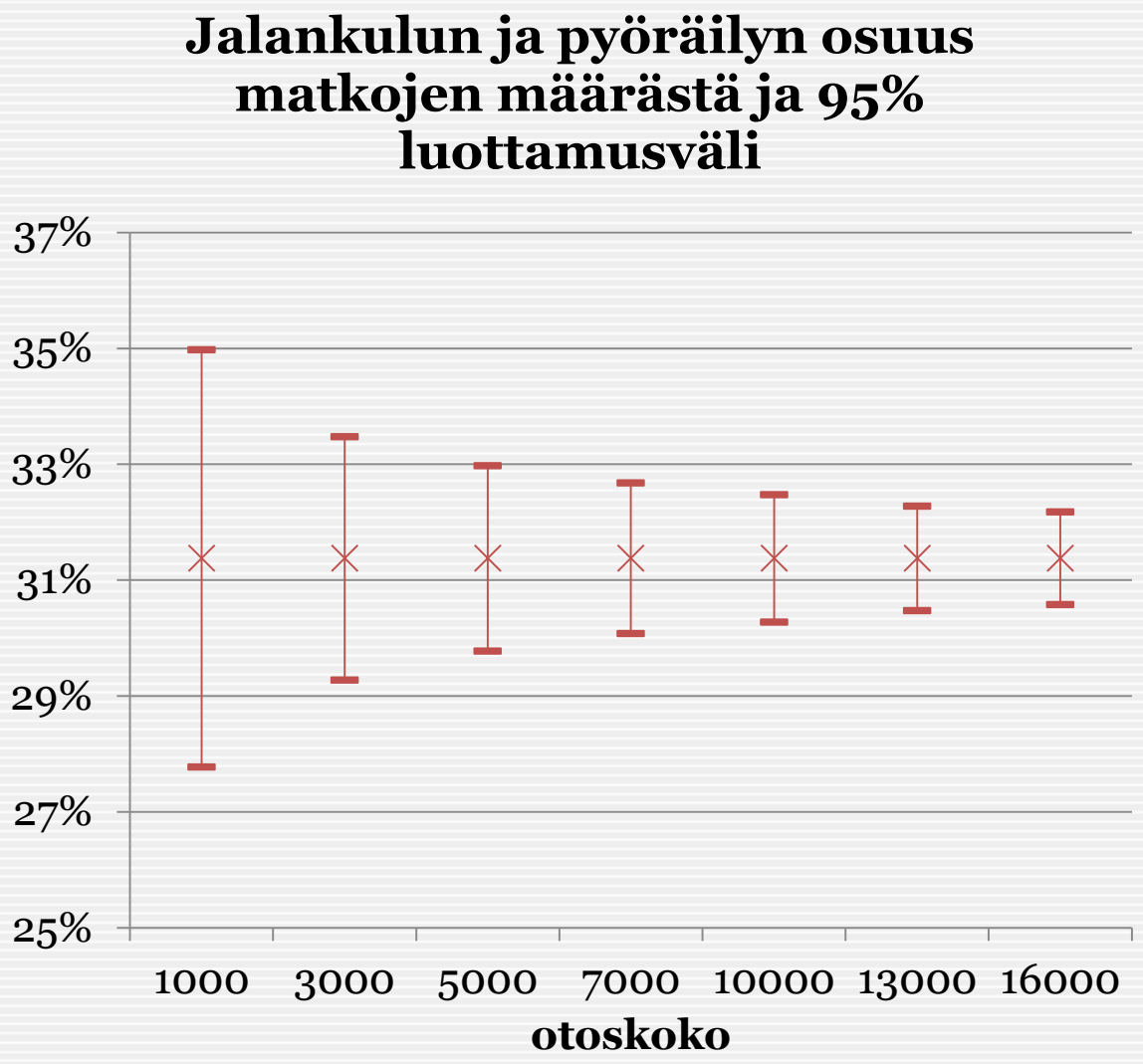
Keskisuuret kaupungit

Koko vuosi keskimäärin

Kotimaanmatkat

Valtakunnallinen henkilö-liikennetutkimus 2010-2011

Oletettu 50% vastausaste



Esimerkki II

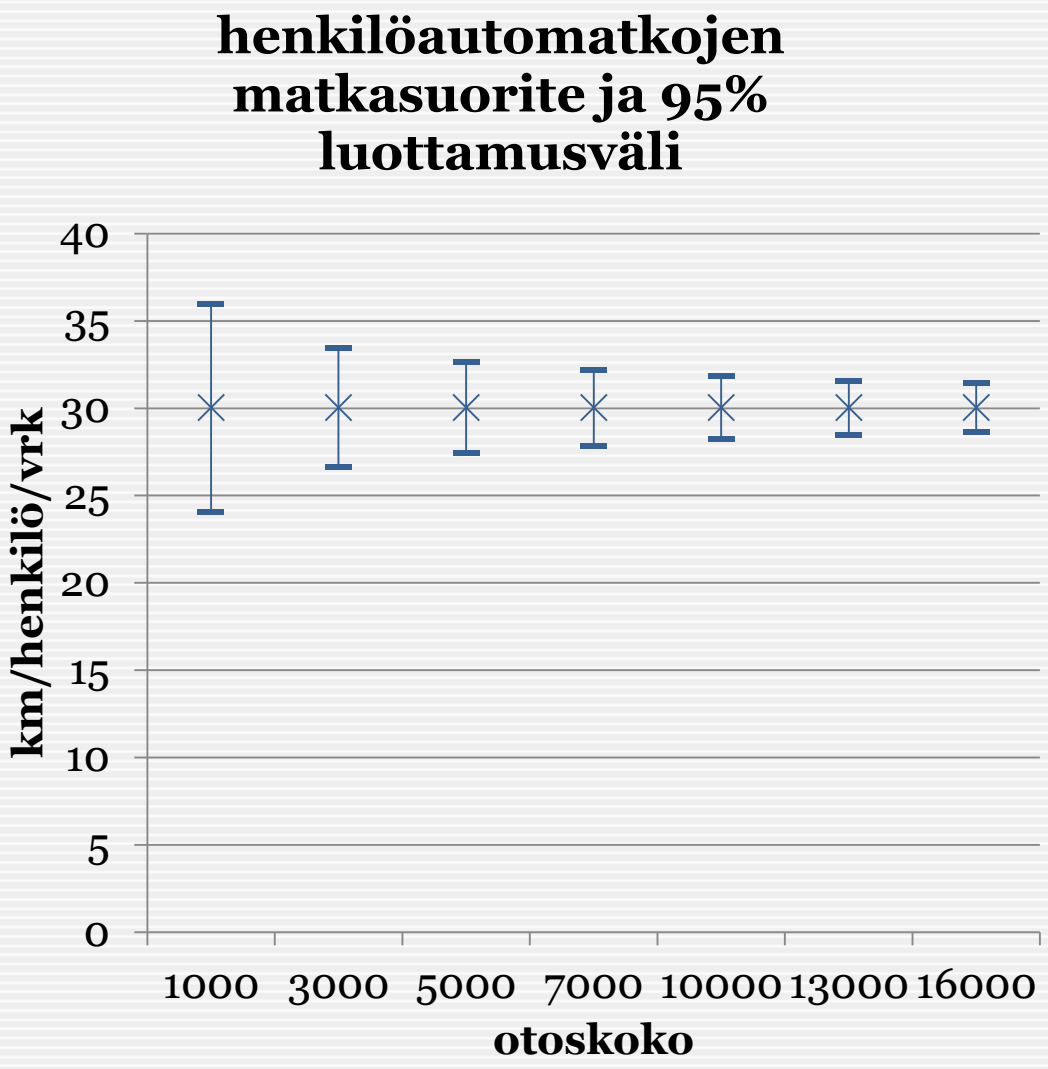
Keskisuuret kaupungit

Koko vuosi keskimäärin

Kotimaanmatkat

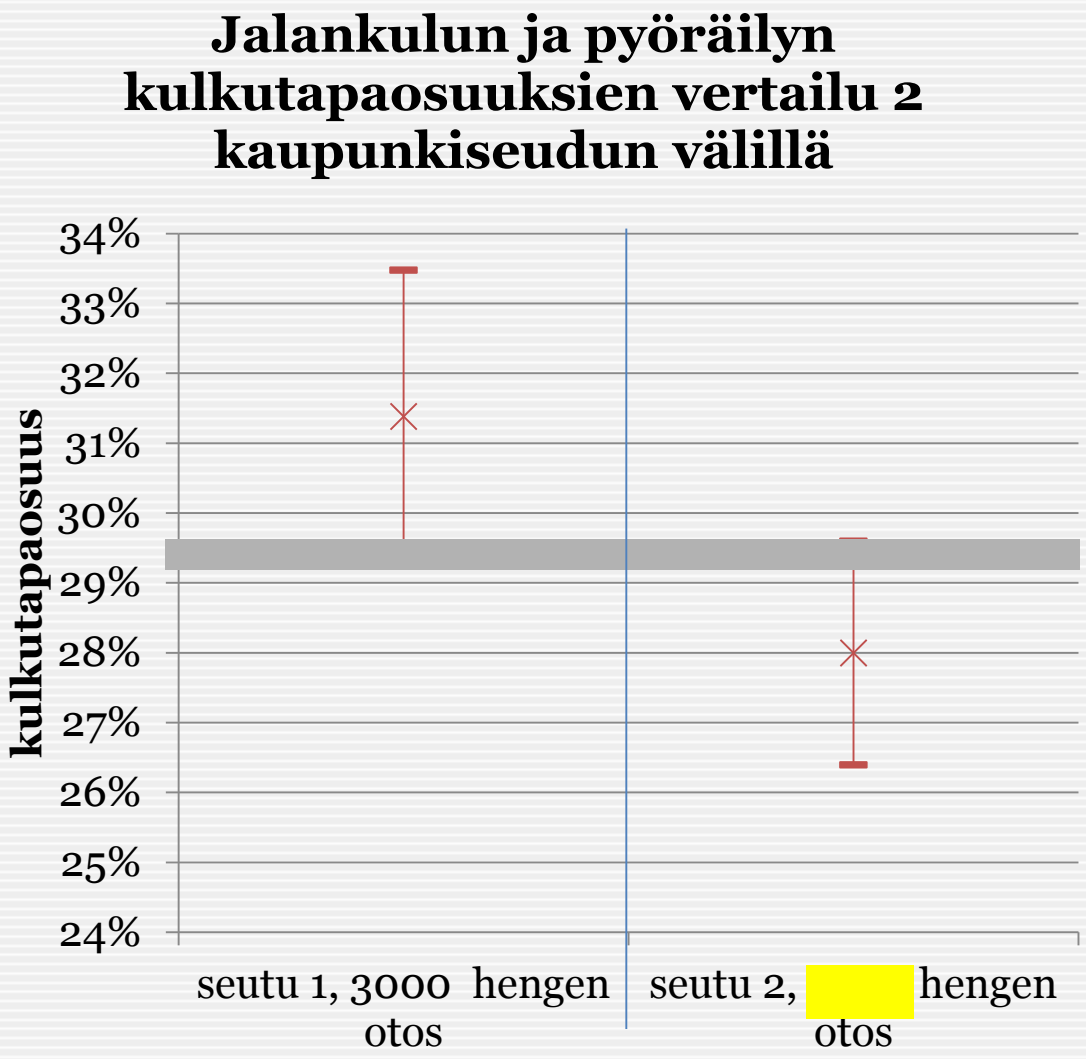
Valtakunnallinen henkilö-liikennetutkimus 2010-2011

Oletettu 50% vastausaste



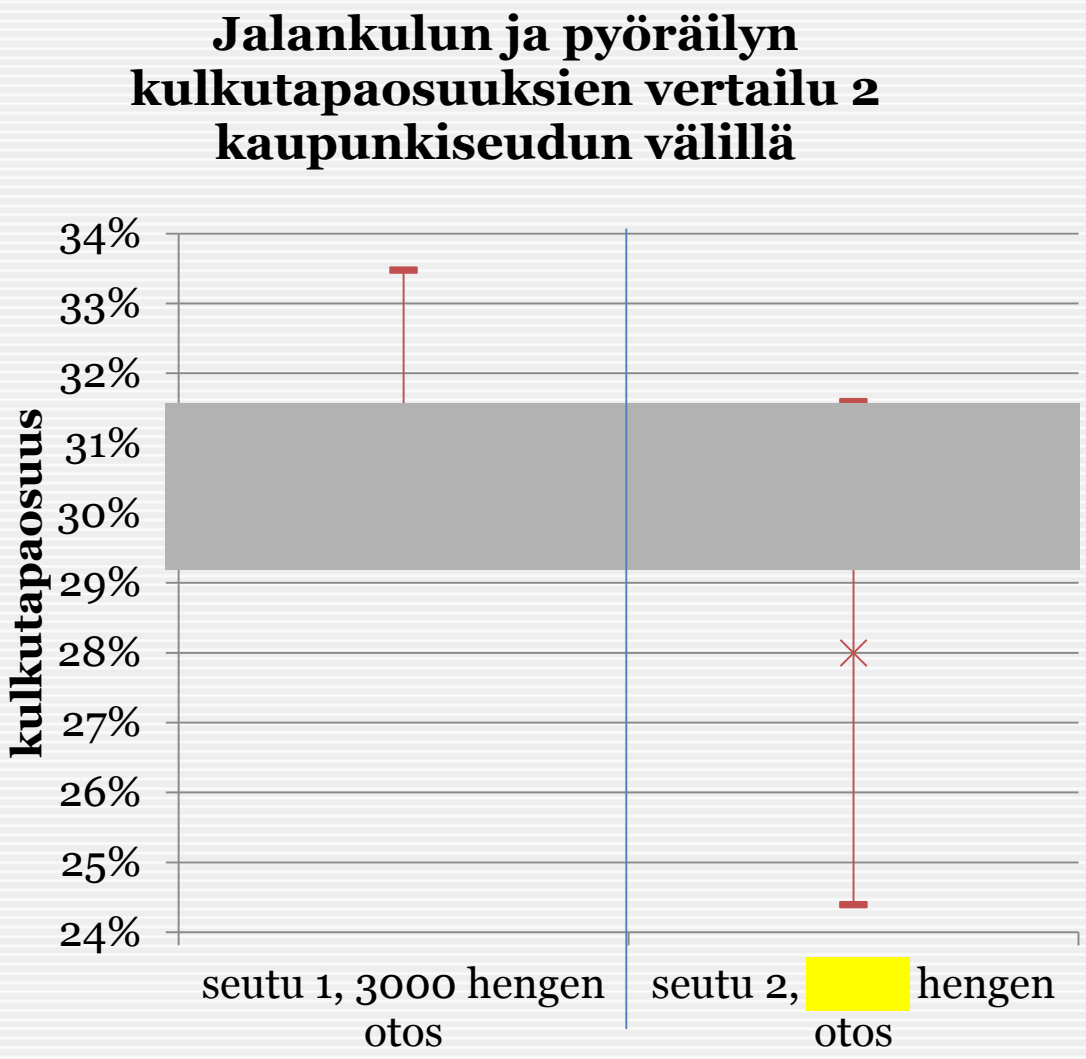
Luotettavuuden merkitys vertailuissa

Vertailu kaupunki-seutujen välillä



Luotettavuuden merkitys vertailuissa

Vertailu kaupunki-seutujen välillä



Luotettavuuden merkitys aikasarjoissa

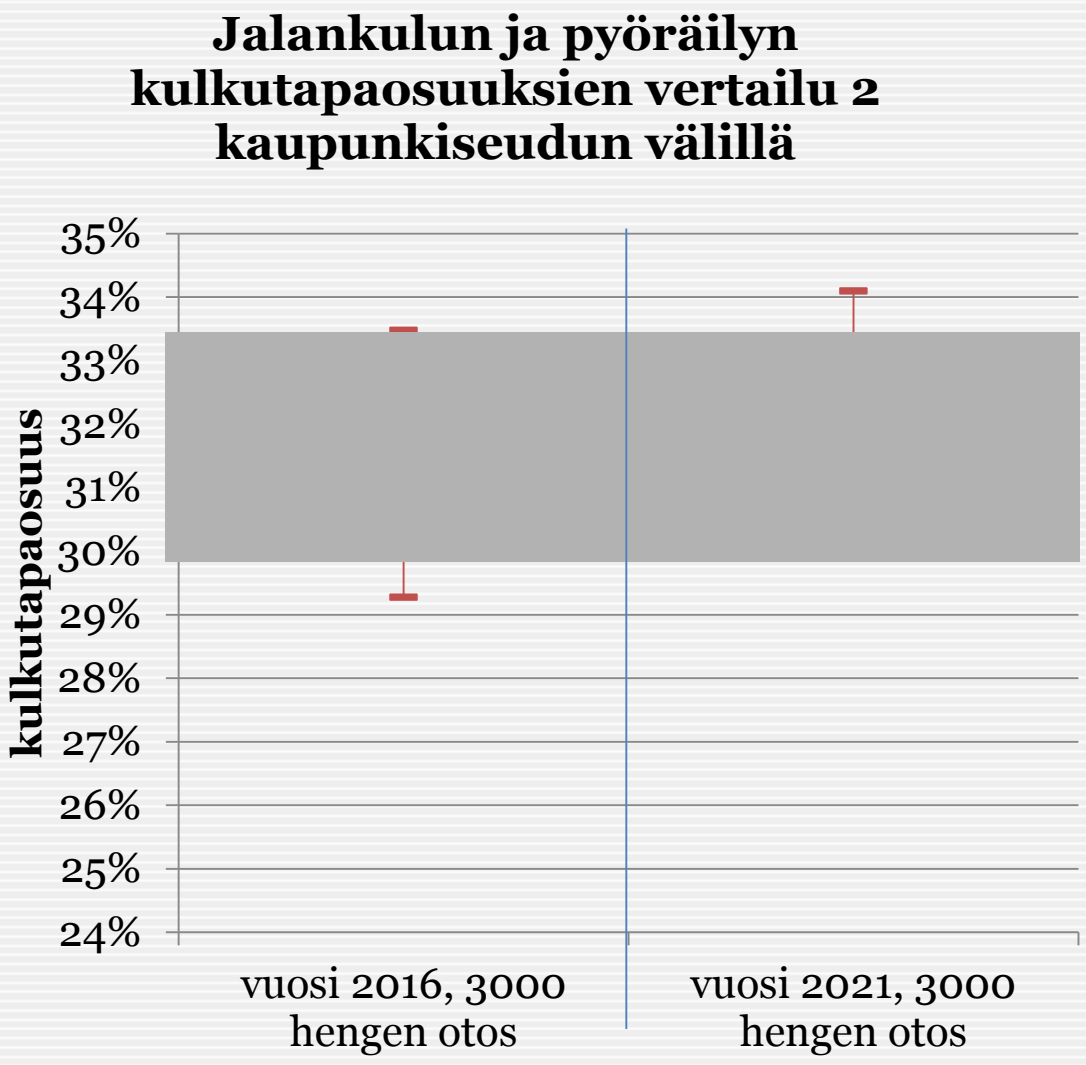
Kaksi eri ajankohtaa

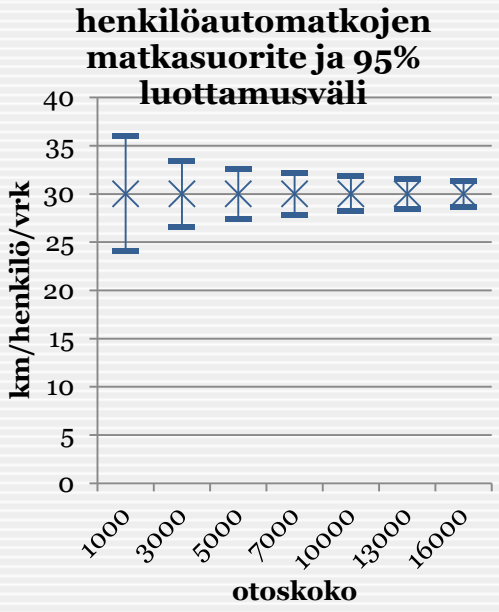
2016 ja 2021

Yksi seutu

Vuonna 2016 kävelyn ja pyöräilyn kulkutapaosuus 31%, ja vuonna 2021 32 %.

Onko kulkutapaosuus kasvanut?





Sama asia tiivistetysti

	arvo HLT	Virhemarginaalit ja näitä vastaavat otoskoot, kun vastausasteeksi oletetaan 50%						
	2010-2011	1000	3000	5000	7000	10000	13000	16000
matkaluku (matkaa/henkilö/vrk)								
henkilöauto	1,93	±0,2	±0,11	±0,09	±0,07	±0,06	±0,05	±0,05
jalankulku ja pyöräily	0,79	±0,12	±0,07	±0,05	±0,04	±0,04	±0,03	±0,03
joukkoliikenne	0,10	±0,04	±0,02	±0,02	±0,01	±0,01	±0,01	±0,01
kotimaanmatkat yhteensä	**3,00**	**±0,21**	**±0,12**	**±0,09**	**±0,08**	**±0,06**	**±0,05**	**±0,05**
matkasuorite (km/henkilö/vrk)								
henkilöauto	32,1	±6	±3,4	±2,6	±2,2	±1,8	±1,6	±1,4
jalankulku ja pyöräily	1,7	±0,4	±0,2	±0,2	±0,1	±0,1	±0,1	±0,1
joukkoliikenne	6,5	±4,5	±2,6	±2	±1,7	±1,4	±1,2	±1
kotimaanmatkat yhteensä	**45,7**	**±8,6**	**±4,9**	**±3,8**	**±3,1**	**±2,6**	**±2,2**	**±2**
kulkutapaosuus (% kotimaanmatkoista, virhemarginaali % -yksikköä)								
henkilöauto	64,4	±3,8	±2,2	±1,6	±1,4	±1,1	±1	±0,9
jalankulku ja pyöräily	26,4	±3,4	±1,9	±1,5	±1,2	±1	±0,9	±0,8
joukkoliikenne	3,5	±1,2	±0,7	±0,5	±0,4	±0,4	±0,3	±0,3

Mitä saavutetaan suuremmalla otoskoolla?

- Lisää **luotettavuutta** tuloksiin (vrt. edelliset kalvot): todelliset tunnusluvut ovat lähellä sitä, mitä tutkimus antaakin
- Vertailu on jonkin verran luotettavampaa em. syystä
  - *Tässä on huomioitava myös esim. menetelmämuutokset, jotka heikentävät vertailtavuutta, sekä sen tutkimuksen havaintomäärä ja luotettavuus, johon halutaan verrata*
- Paremmat mahdollisuudet **liikenteen mallintamiseen**
  - *Tarve riippuu kysyntämallien rakenteesta ja yksityiskohtaisuudesta, mutta 8 000 – 10 000 vastausta riittänee varsin pitkälle (tarkoittaisi otoksena 25 000 – 30 000 hengen otosta; tällöin saadaan myös kuntakohtaiset raportit)*
  - *Mallintamisessa voidaan usein myös hyödyntää vanhempia aineistoja tai yhdistää uusia ja vanhoja aineistoja*

Mitä haittaa on pienemmästä otoskoosta?

- **Luotettavuus huononee** (vrt. edelliset kalvot): suurempi epävarmuus siitä, kuinka hyvin tulokset kuvaavat todellista liikkumista (toki myös vastausteen lasku lisää tätä epävarmuutta)
- **Yksittäisistä kunnista** on vaikeampi sanoa yhtään mitään (Esimerkiksi läntisellä Uudellamaalla otoskoko oli viimeksi 3000 (suuruusluokaltaan 1500 vastausta), ja siellä oltiin hieman pettyneitä, kun edes perustunnuslukuja ei ollut saatavilla kaikista kunnista).
- Vastaavasti **aineiston ristiintaulukointi** ja muu jako pienempiin tarkasteluosiin (esim. ikäryhmät, muut maantieteelliset jaot tms. tarkasteluperusteet) vaikeutuu
- **Vertailu vaikeutuu** – luottamusvälien kasvaessa on vaikeampi sanoa luotettavasti, onko tapahtunut muutosta (Toki tässä on huomioitava myös menetelmämuutokset, jotka heikentävät vertailtavuutta.)

Karkea arvio kustannustasosta

- Kustannukset (yhdelle seudulle) riippuvat todellisuudessa ainakin seuraavista asioista:
  - *Miten tutkimus toteutetaan (mm. vastaustavat ja laatutekijät)*
  - *Kustannusten jakautuminen useamman seudun ja valtakunnallisen tutkimuksen välillä*
- Siten kustannusarviota on mahdotonta antaa tietämättä tarkemmin tutkimuksen toteutustapaa
- HLT2016:n perusteella hyvin karkea arvio suuruusluokasta (sitoumuksetta) on olla (vuoden 2014 tasossa, alv 0%):
  - *Otos 1000: n. 12 000 - 15 000 €*
  - *Otos 3000: n. 30 000€ - 38 000€*
  - *Otos 5000: n. 50 000€ - 63 000€*
  - *Otos 7000: n. 65 000€ - 81 000€*
  - *→ näitä voinee pitää jonkinlaisena alarajana sen tyyppiselle tutkimukselle kuin HLT2016 oli*
  - *Jonkinlaisen karkean ylärajan ison otoksen hinnalle saa tuplaamalla tämän – pienessä otoksessa taas kiinteiden kustannusten osuus on suurempi (riippuen toki jyvityksestä em. maksajien välillä).*