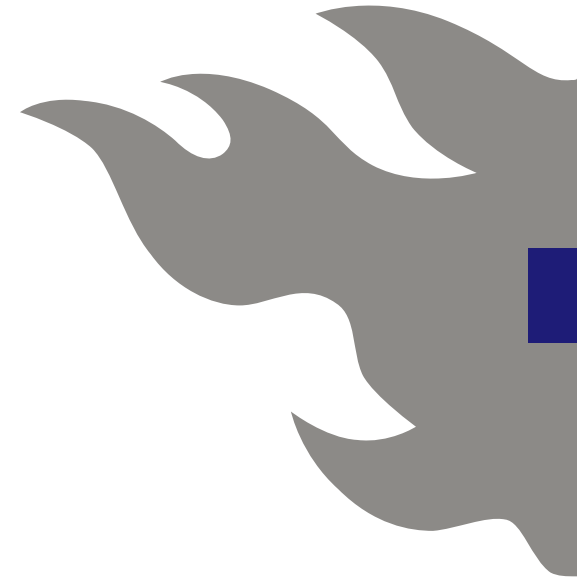




HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Sosiaalitutkimuksen tilastolliset menetelmät Osa 1 – Diat 1 Otanta-asetelmat ja survey-aineiston käsittely

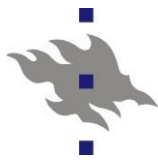
Risto Lehtonen, Helsingin yliopisto
risto.lehtonen@helsinki.fi





Osa 1 Otanta-asetelmat ja survey- aineiston käsittely

- Kohteena **empiirisen kvantitatiivisen yhteiskunta-**
tutkimuksen eli **survey-tutkimuksen** aineiston keruu-
ja käsittelyvaiheiden tilastolliset menetelmät
- **Keskeisiä aiheita**
 - Yhteiskuntatutkimuksen aineistolähteet ja
tiedonkeruumenetelmät
 - Aineiston esikäsittely tilastollista analyysia varten
- Alan keskeiset tilastolliset ohjelmistot
 - SPSS, SAS, R
- **Pääasiallinen esimerkkiaineisto**
 - European Social Survey ESS
Suomen ESS-aineisto ESS2010



ESS – Kaksi esimerkkiä

- Esimerkki1
- Esimerkki2
- Millaisia menetelmällisiä / sisällöllisiä kysymyksiä herää näistä ESS-esimerkeistä?
- Tutkimusasetelmat?
- Tiedonkeruuasetelmat?
- Aineistot?
- Tutkimuskysymykset?
- Analyysit?
- Päätelmät / Yleistykset?
- Tulosten vertailtavuus?

Tutkimus: Suomi on Euroopan neljänneksi onnellisin maa

KOTIMA 5.11.2013, klo 17:21



kuva: iStockphoto.com/John Salsano, Kuvapaja / Maja-Ko

Suomi on eurooppalaisen yhteiskuntatutkimuksen mukaan Euroopan neljänneksi onnellisin maa. Suomalaisten onnellisuus on hieman noussut kahdessa vuodessa.

Onnellisuuden keskiarvolla mitattuna Tanska on paras, Islanti toinen ja Norja kolmas. Suomen kanssa samalle tasolle yltää Sveitsi. Sen sijaan ruotsalaisten onnellisuuskasvun on laskenut ja jää nyt kuudenneksi.

European Social Survey (ESS) -tutkimusta varten tiedusteltiin yli 15-vuotiaiden onnellisuudesta 29 maassa viime ja tänä vuonna.

Tutkimuksen otantaa valvovaan ryhmään kuului tilastotieteen professori Seppo Laaksonen Helsingin yliopistosta. Hän kertoo, että vastaajilta kysyttiin onnellisuudesta asteikolla 1–10, jossa 10 on hyvin onnellinen.

– Tässä vaiheessa uusimmat tulokset ovat käytettävissä suuresta osasta ESS-tutkimusmaista, mutta esimerkiksi Liettua ja Italia eivät vielä ole mukana. Loput tulokset saadaan ensi vuonna, Laaksonen kertoo.

Onnellisuustilaston peränpitäjäksi jää Bulgaria, jossa onnellisuus on vähentynyt entisestään. Onnellisuuden keskiarvo on noussut selvimmän Saksassa.

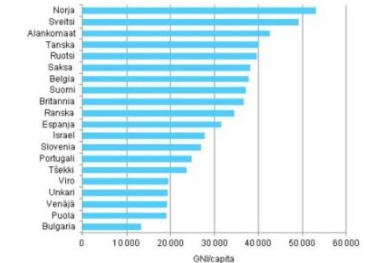
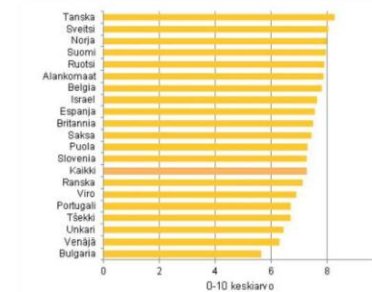
STT

Tilastojen ABC Harjoitukset 1

Onnellisuus ja vauraus 🗳️
Vertaile onnellisuutta kuvaavaa kuviota eri maiden vauraudesta kertovaan kuvioon. Mitä samankaltaisuuksia, entä mitä eroja havaitset?

Kuvio. European Social Survey'n tulokset, kun yli 15-vuotiaita kysyttiin vuonna 2010 kuinka onnellisia he ovat.

Kysymys: Kun arvioidaan elämää kokonaisuutena, kuinka onnellisena pidätte itseänne? Arvioikaa asteikolla 0–10, kuinka onnellinen olette. 0 on "Erittäin onneton", 10 on "Erittäin onnellinen"

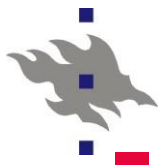


Lähteet

GM-Index: <http://dataresources.worldbank.org/DATASTAT/STATSCSR/resources/GMIIPC.pdf>

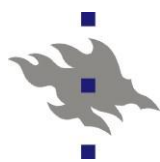
Onnellisuusindeksi: European Social Survey 2010 [raportti 4.1.2012] Saatavuus: <http://ess.nor.uit.no/essfound/>

Kuvio. Maailman pankin tiedot eri maiden vauraudesta kansantulolla mitattuna (Gross National Income/capita) vuonna 2010



Kirjallisuutta

- Lehtonen Risto & Pahkinen Erkki (2004). [Practical Methods for Design and Analysis of Complex Surveys](#). John Wiley & Sons. Ladattavissa [dawsoneran](#) kautta
- Pahkinen Erkki & Lehtonen Risto (1989). [Otanta-asetelmat ja tilastollinen analyysi](#), Gaudeamus.
- Laaksonen Seppo (2013). [Surveymetodiikka](#). 2. painos. bookboon.com ([pdf](#))
- De Vaus D. (2013). [Surveys in Social Research](#) □6th Ed. Routledge.
- Tilastokeskus (2007). [Laatua tilastoissa](#), 2. uudistettu painos, Tilastokeskus, Käsikirjoja 43.
- Eurostat (2008). [Survey Sampling Reference Guidelines](#)
- Huppert F.A. et al. (2009). Measuring well-being across Europe: Description of the ESS well-being module and preliminary findings. Social Indicators Research 91, p. 301-315.
<http://link.springer.com/article/10.1007%2Fs11205-008-9346-0>



Survey: Empiirinen kvantitatiivinen (yhteiskunta)tutkimus: Vaiheet

I Suunnittelu ja testaus

1. Tutkimusongelman muotoilu
2. Tutkimusasetelman laadinta
3. Tiedonkeruuasetelman laadinta
 - a) Otosperusteinen aineisto
 - Valmis aineisto (Tietoarkistot)
 - Itse kerätty aineisto
 - b) Rekisteriperusteinen aineisto
 - c) Yhdistelmäaineisto

II Tiedonkeruuoperaatiot

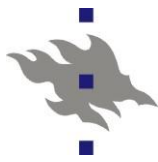
4. Tiedonkeruun toteutus valitun tiedonkeruuasetelman mukaisesti
5. Tiedostonmuodostus, esim:
 - aineistojen yhdistely
 - editointi, imputointi
 - puuttuneisuuden analyysi
 - painokertoimien muodostus
 - metatietokuvaukset

III Tilastollinen analyysi

6. Eksplorointi ja kuvailu
 - tunnuslukujen laskenta,
 - taulukointi
 - graafiset kuvailut
 - piste-estimointi
 - väliestimointi
7. Analyysi ja tulkinta
 - tilastollinen mallinnus

IV Raportointi ja jälkihoito

8. Julkaisut ja artikkelit
9. Opinnäytetyöt
10. Esitelmät
11. Sähköiset tuotteet
12. Dokumentointi ja arkistointi



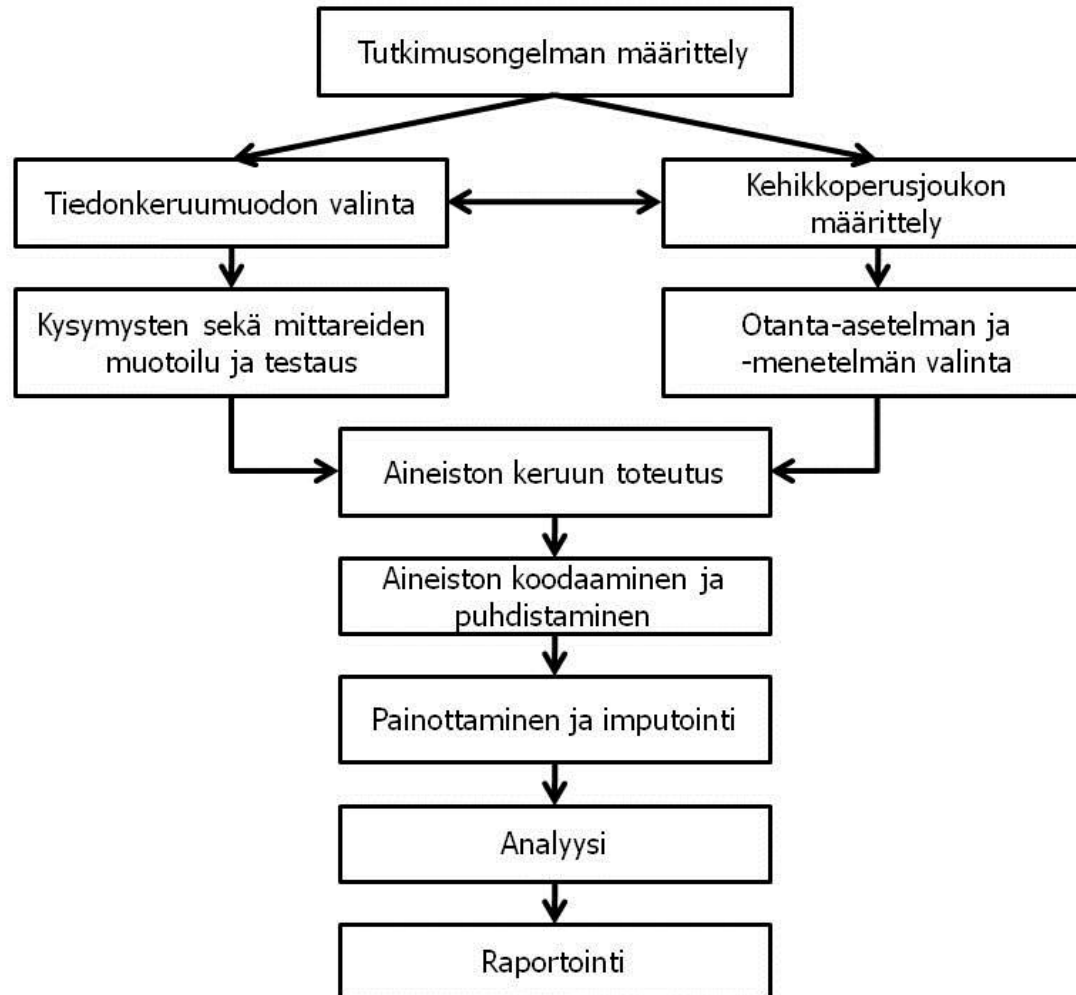
Tiedonkeruuasetelmia

- Suoraan tiedonkeruuseen perustuvat kysely- ja haastatteluaineistot
- **Otosperusteiset** = Käytetään **tilastollisia otantamenetelmiä**
 - Esim: ESS European Social Survey
 - Yleisesti käytössä tieteellisessä tutkimuksessa
- **Ei-otosperusteiset** = Aineisto kerätään jollain **ei-satunnaisella** menetelmällä
 - Kiintiöpoiminta (TNS Gallup, Taloustutkimus,...)
 - Verkkokyselyt , web-kyselyt
 - Esim. *Self-selection web survey*
 - Ei käytetä yleisesti tieteellisessä tutkimuksessa



Esim: Tyypillinen otosperusteinen tutkimusprosessi: Itse kerätty aineisto

Lähde: Jani Miettinen, pro gradu (2011)



Kuvio 2.2. Kyselytutkimuksen prosessi (Groves et al. 2009, s. 149).



Tutkimusaineistojen lähteitä I: Tietoarkistot

- Pääasiassa **otosperusteisia** kotimaisia ja kansainvälisiä **valmiiksi kerättyjä** aineistoja
 - Perustuvat suoraan tiedonkeruuseen
 - Kyselyaineistot, haastatteluaineistot
- Yhteiskuntatieteellinen tietoarkisto [FSD](#)
Finnish Social Science Data Archive
 - Tampereen yliopiston yhteydessä
- Council of European Social Science Data Archives [CESSDA](#)
- Esimerkki
 - European Social Survey [ESS](#) (2002-2012)
 - Riippumattomia poikkileikkausaineistoja



FSD-tietoarkiston menetelmäopiskelun tietovaranto

- **KvantiMOTV:** Kvantitatiivisten tutkimusmenetelmien oppimisympäristö
- Analyysimenetelmien ohella MOTV käsittelee:
 - Tutkimuksen teon keskeisiä kysymyksiä
 - Sisältää eritasoisia SPSS-harjoituksia
 - Tarjoaa vapaasti käytettäviä harjoitusaineistoja itseopiskeluun ja menetelmäkursseille
- FSD on hyvä esimerkki **Open Data** –periaatteen toteuttamisesta tutkimusympäristöissä



Ei-otosperusteiset suorat tiedonkeruut

Esimerkki: Itsevalikoituva verkkokysely

■ *Self-selection web survey*

■ Valtaosa haluaa alle 25 oppilaan luokat

”Yle Uutisten Suora linja kysyi netinkäyttäjiltä, minkä kokoinen on hyvä koululuokka...”

■ Joka kolmas nuori haluaisi juoda vähemmän

”A-klinikkasäätiön Nuortenlinkki-verkkopalvelun kyselyssä kolmasosa nuorista toivoi, että kaveripiirissä käytettäisiin vähemmän alkoholia.”

■ Nettikysely on nykyaikaa...

Seppo Laaksonen (2013) Kansantaloudellinen aikakauskirja

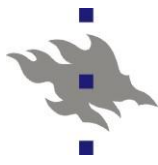
■ ...mutta nettikyselyssä on myös ikäviä sudenkuoppia

Petro Poutanen (2013) Nettikyselyjen sudenkuopat



Nettikysely: Ongelmat ja vaihtoehdot

- Millaisia **yleistyksiä** *itsevalikoituvan nettikyselyn* perusteella voidaan tehdä?
- Mitä **ongelmia** menetelmään liittyy?
- Miten nettikyselyn yleistettävyyttä voidaan parantaa tilastotieteen menetelmillä?
 - Jani Miettinen, HY tilastotieteen [gradu](#) (2011)
- *How accurate are self-selection web surveys?*
Jelke Bethlehem, Statistics Netherlands,
[Discussion paper](#) (2008)
- Nettikyselyjen vaihtoehdot?



Esimerkki: Nettikyselyt (Jani Miettinen, pro gradu)

Taulukko 3.1. Verkkokyselytutkimuksen asetelmat (Couper 2000).

Otosperusteiset menetelmät	Ei-satunnaiset menetelmät
(1) Verkkosivujen käyttäjäkyselyt	(6) Itsevalikoituneet verkkokyselyt
(2) Listauksista kerätyt otokset	(7) Vapaaehtoiset paneelitutkimukset
(3) Vaihtoehto vastata verkon välityksellä	(8) Viihdegallupit
(4) Paneeli etukäteen värvätyistä Internetin käyttäjistä	
(5) Paneeli etukäteen värvätyistä populaation edustajista	



Tutkimusaineistojen lähteitä II: Rekisteriaineistot

■ Hallinnollinen rekisteri

- Hallinnollisen prosessin oheistuote
- Päivittyy jatkuvasti
- Kela: Sosiaalivakuutuksen tietokannat
- Verohallitus: Verotietokanta
- Väestörekisterikeskus: Väestön keskusrekisteri

■ Tilastorekisteri (Tilastokeskus)

- Usean hallinnollisen rekisterin yhdistelmä
- Rekisteriseloste: [Tulonjakotilasto](#)
- StatFin – [Tilastotietokannat](#)

- Rekistereitä käytetään **tutkimustiedon lähteinä ja tutkimusten otantakehikkoina**



Rekisteritutkimuksen tukikeskus ReTKi

- ReTKi on tutkimusinfrastruktuuri, jonka tavoitteena on edistää kansallisten rekisterien tutkimuskäyttöä.
- Sijaitsee Kansallisarkiston yhteydessä
- **Tehtävät**
 - tarjota tietoa rekistereistä ja niiden tutkimuskäytöstä
 - järjestää koulutusta rekisteritutkimuksesta
 - neuvoa rekisteriaineistojen tutkimuskäyttöön liittyvissä asioissa
 - ylläpitää rekisteriviranomaisten ja tutkimuslaitosten yhdyshenkilöiden verkostoa



Tutkimusaineistojen lähteitä III: Aineistotyyppien yhdistelmät

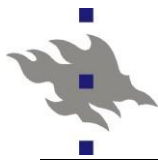
- Suoralla tiedonkeruulla kerättyjen otosperusteisten aineistojen ja rekisteriperusteisten aineistojen yhdistelmät
 - Yleistyvä trendi - erinomaiset mahdollisuudet suomalaisessa yhteiskuntatutkimuksessa
- Esim: Tilastokeskuksen SILC-tutkimus
 - *Statistics on Income and Living Conditions*
 - Osa tiedoista kerätään haastattelemalla (CAPI tai CATI) ja osa otetaan tilastorekistereistä (verotuksen tulotiedot, koulutustiedot,...)
 - Tiedot yhdistetään henkilötunnusten avulla



Esimerkki: Aineistotyyppien yhdistelmät

YHTEENVETO 1. Aineisto-optiot tiedonkeruun tavan ja kattavuuden mukaan.

TIEDONKERUUTAPA	KATTAVUUS PERUSJOUKON SUHTEEN	
	A. OSITTAINEN KATTAVUUS: OTOSTUTKIMUS	B. TÄYSI KATTAVUUS: KOKONAISTUTKIMUS
1. SUORA TIEDONKERUU Tietolähde Haastattelututkimus Tietokoneavusteinen käyntihaastattelu <i>Computer Assisted Personal Interview</i> CAPI Tietokoneavusteinen puhelinhaastattelu <i>Computer Assisted Telephone Interview</i> CATI Tietokoneavusteinen kysely <i>Computer Assisted Self-interview</i> CASI <i>Computer Assisted Web Survey</i> CAWI Tiedonkeruu kynä- ja paperi -menetelmällä <i>Paper-and-Pencil Interview</i> PAPI Postikysely Internet-kysely, Web-kysely, eSurvey	Optio 1a. Suoraan tiedonkeruuseen perustuva otostutkimus Perinteinen otostutkimuksen tyyppi Tilastokeskuksen tutkimuksia ja tilastoja <input type="checkbox"/> Työvoimatutkimus <input type="checkbox"/> Kulutustutkimus Kelan tutkimuksia ja selvityksiä <input type="checkbox"/> Terveysturvan väestötutkimukset Monikansallisia tutkimuksia <input type="checkbox"/> European Social Survey ESS <input type="checkbox"/> PISA	Optio 1b. Suoraan tiedonkeruuseen perustuva kokonaistutkimus Perinteinen kokonaistutkimuksen tyyppi <input type="checkbox"/> Tilastokeskuksen väestölaskennat (vuoteen 1985 saakka)
2. EPÄSUORA TIEDONKERUU Tietolähde: Rekisteri Kattaa kohdeperusjoukon Päivitetään säännöllisesti Hallinnollinen rekisteri Hallinnollisen proseduurin oheistuote Tilastorekisteri Usean hallinnollisen rekisterin yhdistelmä	Optio 2a. Hallinnolliseen rekisteriaineistoon perustuva otostutkimus Puhtaana muotona harvinainen <input type="checkbox"/> Poikkeuksena Tilastokeskuksesta saatavat tilastorekistereiden otosaineistot	Optio 2b. Hallinnolliseen rekisteriin tai tilastorekisteriin perustuva kokonaistutkimus Tämä surveyn tyyppi on yleistymässä Aineistolähteet <input type="checkbox"/> Rekisteriperusteiset väestölaskennat <input type="checkbox"/> Sosiaalivakuutuksen rekisterit <input type="checkbox"/> Väestörekisteri <input type="checkbox"/> Yritysrekisteri <input type="checkbox"/> Verotusrekisterit <input type="checkbox"/> Kelan lääketutkimukset
3. TIEDONKERUUTAPOJEN YHDISTELMÄ Tietolähde: Suoran ja epäsuoran tiedonkeruun yhdistelmä	Optio 3. Otostutkimus, joka perustuu suoran tiedonkeruun ja rekisteriaineiston yhdistelyyn Tämä surveyn tyyppi on yleistymässä <input type="checkbox"/> KTL:n Terveys 2000 ja Terveys 2010 <input type="checkbox"/> Kelan Mini-Suomi-terveystutkimus <input type="checkbox"/> Tilastokeskuksen Tulonjakotutkimus <input type="checkbox"/> EU:n European Community Household Panel ECHP <input type="checkbox"/> EU SILC (Statistics on Income and Living Conditions)	



Eri aineistotyyppien etuja ja rajoituksia

	ETUJA	RAJOITUKSIA
Otosperusteiset kysely- ja haastatteluaineistot		
Tietoarkistojen valmiit otosaineistot	Maksuttomuus, hyvä saatavuus, monipuolisuus Tulosten yleistettävyys	Ei voi räätälöidä omia tutkimustarpeita varten Aineistojen ikääntyminen Metatietojen puutteet
Erikseen kerättävä otosaineisto	Joustavuus, ajankohtaisuus Tulosten yleistettävyys Räätälöinti omia tutkimustarpeita varten	Kallista ja aikaavievää Menetelmällisesti vaativaa Mahdollinen vastauskato
El-otosperusteiset aineistot		
Itsevalikoituva verkkokysely	Pienet kustannukset Räätälöinti mahdollista Nopea kerätä	Valikoitumisharha Tuloksia ei voi yleistää tilastollisesti (ilman erityistoimia tai ei lainkaan)
Rekisteriaineistot		
Tilastorekisterit	Saatavuus, kattavuus, laatu Yhdisteltävyys	Ei voi räätälöidä omia tutkimustarpeita varten
Muita huomioita	Nykytrendi: Otosperusteisten kysely- ja haastatteluaineistojen ja rekisteriaineistojen yhdistelmät	



ESS – European Social Survey

- ESS: tausta ja tavoitteet
- The **European Social Survey** (the ESS) is an academically-driven social survey designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behaviour patterns of its diverse populations.
- ESS survey rounds: 2002, 2004, 2006, 2008, 2010, 2012



ESS – European Social Survey

- Aineistot ja dokumentaatio:
Norjan yhteiskuntatieteellinen tietoaarkisto
<http://ess.nsd.uib.no/>
- [EduNet](#)
- **European Social Survey Education Net**
- ESS EduNet is a training resource mainly developed for use in higher education. The ambition is to create a social science laboratory where theoretical questions can be explored using high quality empirical data. The resource is based on the European Social Survey.



ESS – 2010

- ESS vaihe 5 (2010)

Kv. tietovarannot

FSD:n tietovarannot

- Tiedonkeruu

- Mukana 28 maata

- Tyypillisesti käyntihaastattelu (CAPI)

- Keskimääräinen vastausprosentti 70 %

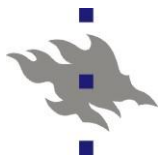
- Vastauskato (*Unit nonresponse*): 30 %

- Vaihtelee maittain

- EU-tasoinen tutkimusaineisto

- Kaikkiaan noin 39 000 henkilöä

- Noin 600 muuttujaa



ESS 2010 – ”Trust”-esimerkki

■ Luottamus yhteiskunnallisiin instituutioihin

■ Kertokaa asteikolla nollasta kymmeneen, kuinka paljon henkilökohtaisesti luotatte seuraavaksi luettelemiini tahoihin. Nolla tarkoittaa sitä, että ette luota ollenkaan kyseiseen tahoon ja 10 sitä, että luotatte erittäin vahvasti kyseiseen tahoon:

■ B4. Eduskunta?

■ B5. Oikeusjärjestelmä?

■ B6. Poliisi?

■ B7. Poliitikot?

■ B8. Poliittiset puolueet?

■ B9. Euroopan parlamentti?

■ B10. YK eli Yhdistyneet Kansakunnat?

EN LUOTA OLLENKAAN 0 ————— 10 LUOTAN TÄYSIN



ESS 2010 – Tiedonkeruulomake

CARD 8 Using this card, please tell me on a score of 0-10 how much you personally trust each of the institutions I read out. 0 means you do not trust an institution at all, and 10 means you have complete trust. Firstly...**READ OUT...**

		<i>No trust at all</i>										<i>Complete trust (Don't know)</i>	
B4	...[country]'s parliament?	00	01	02	03	04	05	06	07	08	09	10	88
B5	...the legal system?	00	01	02	03	04	05	06	07	08	09	10	88
B6	...the police?	00	01	02	03	04	05	06	07	08	09	10	88
B7	...politicians?	00	01	02	03	04	05	06	07	08	09	10	88
B8	...political parties?	00	01	02	03	04	05	06	07	08	09	10	88
B9	...the European Parliament?	00	01	02	03	04	05	06	07	08	09	10	88
B10	...the United Nations?	00	01	02	03	04	05	06	07	08	09	10	88



ESS 2010 – Luottamus instituutioihin taustamuuttujien mukaan

■ Tulosmuuttujat

trstprl	Trust in country's parliament
trstlgl	Trust in the legal system
trstplc	Trust in the police
trstplt	Trust in politicians
trstprrt	Trust in political parties
trstep	Trust in the European Parliament
trstun	Trust in the United Nations

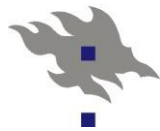
■ Mittaus

0 (ei luottamusta), 0, 1, ..., 9, 10 (täydellinen luottamus)

Mitta-asteikko?

Mittaamisesta lisää: **OSA 2** (Kimmo Vehkalahti)

■ Tunnusluvut: Keskiarvot



ESS 2010 – Luottamus instituutioihin sukupuolen mukaan – keskiarvot (Suomen aineisto)

Gender	Trust in country's parliament	Trust in the legal system	Trust in the police	Trust in politicians	Trust in political parties	Trust in the European Parliament	Trust in the United Nations
Mies	5.4	6.9	7.9	4.3	4.4	4.8	6.5
Nainen	5.4	6.9	8.1	4.6	4.7	5.3	6.6
Kaikki	5.4	6.9	8.0	4.4	4.5	5.1	6.6

Aineistossa havaintoja kaikkiaan n = 1878

Puuttuvia vastauksia (item nonresponse): Vaihtelee muuttujittain!



ESS 2010 – Luottamus instituutioihin iän mukaan – keskiarvot (Suomen aineisto)

Age group	Trust in country's parliament	Trust in the legal system	Trust in the police	Trust in politicians	Trust in political parties	Trust in the European Parliament	Trust in the United Nations
-25	5.8	6.8	7.8	4.9	5.2	5.9	6.6
26-45	5.4	6.9	8.1	4.3	4.4	5.0	6.7
46-65	5.3	6.9	8.0	4.3	4.3	4.8	6.5
66-	5.2	7.0	8.2	4.5	4.7	5.0	6.4
Kaikki	5.4	6.9	8.0	4.4	4.5	5.1	6.6

Aineistossa havaintoja kaikkiaan n = 1878

Puuttuvia vastauksia (item nonresponse): Vaihtelee muuttujittain!



ESS 2010 – Luottamus instituutioihin taustamuuttujien mukaan: Analyysi

- **Tilastollinen päättely**
- Voidaanko havaitut sukupuolien väliset tai ikäryhmien väliset erot yleistää perusjoukkoon, josta otosaineisto on kerätty?
- **Keskiarvojen keskivirheet** (Std Error of Mean)
- **95 % luottamusvälit** (95% CL for Mean)
- Erojen tilastollisen merkitsevyyden testaus
Tilastolliset testit
- **Tilastollinen mallinnus**
OSA 3 (Pekka Pere), **OSA 4** (Jyrki Möttönen)



ESS 2010 – Luottamus instituutioihin sukupuolen mukaan: Analyysi

Domain Analysis: Gender					
Gender	N	Mean	Std Error of Mean	95% CL for Mean	
Trust in country's parliament trstpri					
mies	909	5.390539	0.077844	5.23786820	5.54320991
nainen	957	5.378265	0.069874	5.24122683	5.51530399
Trust in the police trstplc					
mies	911	7.941822	0.060737	7.82270301	8.06094134
nainen	958	8.116910	0.047721	8.02331801	8.21050245
Trust in the European Parliament trstep					
mies	885	4.836158	0.077710	4.68374682	4.98856957
nainen	921	5.326819	0.066514	5.19636680	5.45727055



ESS 2010 – Luottamus instituutioihin

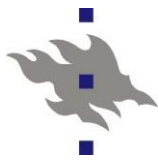
ANOVA-esimerkki

Dependent Variable trstep			
Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	1	23.00	<.0001
Intercept	1	9866.03	<.0001
gndr	1	23.00	<.0001

Dependent Variable trstplc			
Tests of Model Effects			
Effect	Num DF	F Value	Pr > F
Model	1	5.14	0.0236
Intercept	1	43200.3	<.0001
gndr	1	5.14	0.0236

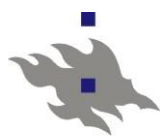
- trstep: Trust in the European Parliament
- Selvä asenne-ero miesten ja naisten välillä
gndr: F value 23.00, $p < 0.0001$
- (ote ohjelmatulostuksesta, SPSS, ANOVA table)

- trstplc: Trust in the police
- (Lievä) asenne-ero miesten ja naisten välillä
gndr: F value 5.14, $p < 0.05$)
- Perusteellisempi tarkastelu: OSA 3, OSA 4



ESS 2010 – Luottamus instituutioihin taustamuuttujien mukaan

- **Tilastollinen päättely**
- Keskiarvojen erojen testaus (OSA 4)
 - yksi (diskreetti, luokiteltu) selittävä muuttuja
 - 2 ryhmää (sukupuoli): t-testi
 - >2 ryhmää (ikäryhmät): 1-suuntainen ANOVA (*Analysis of variance*)
 - Kaksi (diskreettiä, luokiteltua) selittävää muuttujaa (sukupuoli ja ikäryhmä): 2-suuntainen ANOVA
 - Yhdysvaikutusten tarkastelu (interaktiot)
- **Regressioanalyysi (OSA 3)**
 - Selittävät muuttujat jatkuvatyyppisiä
 - Esim: Ikä vuosina



Tekninen liite 1: ESS 2010

Luottamus instituutioihin: Estimaattorit

Keskiarvo $\bar{y} = \sum_{k=1}^n y_k / n$ (Mean)

Varianssi $s^2 = \sum_{k=1}^n (y_k - \bar{y})^2 / (n - 1)$

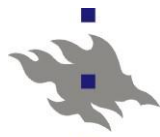
Keskihajonta $s = \sqrt{\sum_{k=1}^n (y_k - \bar{y})^2 / (n - 1)}$

Keskivirhe $s.e = s / \sqrt{n}$ (Standard error)

Keskiarvon 95% luottamusväli

- alaraja $\bar{y} - 1.96 \times s.e$

- yläraja $\bar{y} + 1.96 \times s.e$



Tekninen liite 2: ESS 2010

Laskentaesimerkki (Naisten osajoukko)

trstep - Trust in the European Parliament

Naisten osajoukko $n = 921$

Keskiarvo $\bar{y} = 5.326819$

Keskivirhe $s.e = 0.066514$

Keskiarvon 95% luottamusväli

- alaraja

$$\bar{y} - 1.96 \times s.e = 5.326819 - 1.96 \times 0.066514 = 5.196$$

- yläraja

$$\bar{y} + 1.96 \times s.e = 5.326819 + 1.96 \times 0.066514 = 5.457$$