

A LAKOSSÁG MINT MUNKAERŐ – A LAKOSSÁG MINT FOGYASZTÓ

ALTERNATÍV ADATGYŰJTÉSI TECHNIKÁK ADATFORRADALOM IDEJÉRE

Dr. Kárpáti József PhD

Neumann János Egyetem

Magyar Regionális Tudományi Társaság XVI. Vándorgyűlése

Kecskemét, 2018. október 19.

16. szekció

MIÉRT KELL MEGFIGYELNI A LAKOSSÁGOT?

- A társadalmi és gazdasági folyamatok mérésének fontos, máig ki nem váltható eszköze maradt a közvetlen információgyűjtés.
- A természetes személyekből álló háztartások, mint **mikroszintű gazdasági egységek** egyrészt munkaerejük „eladásával”, másrészt a megszerzett jövedelem elfogyasztásával vesznek részt a gazdasági folyamatokban. (ld. GPD számítás - háztartások szektora, makroökonómia...)
- Demográfiai sajátosságaik mellett lakóhelyük, ingázási, utazási, költési szokásaik alapvetően befolyásolják a gazdaság működésének **területi vetületét** is.



A HÁZTARTÁSOK VIZSGÁLATÁNAK ALAPVETŐ FORRÁSAI, MÓDJAI

- Szektorális kérdésekben piaci alapon működő közvéleménykutatók, piackutatók megkeresései
- Különböző piaci szolgáltatók big data adatbázisai – tranzakciós rendszerekből szerzett információk (a vállalat szellemi tőkéje...)
- Hivatalos statisztikai adatfelvételek, országos vagy területi reprezentatív szinten.
- Adminisztratív nyilvántartások adatai.

? Hogyan állíthatunk elő a közjavak körébe tartozó, tudományosan is kutatható információt mindezekből a forrásokból?

HÁZTARTÁSI / LAKOSSÁGI FELVÉTELEK ALTERNATÍVÁI

- **Mixed-mode surveys**

Több adatgyűjtési csatornán, de közvetlenül valahogy a háztartásoktól szerzünk be adatokat.

- Kérdőíves felvétel (interjúval) papíron vagy mobileszközzel (PAPI / CAPI)
- Telefonos interjú a háztartással (CATI)
- Kérdőíves önkitöltés (papíron), jobb esetben ma már webes felületen (CAWI)

- **Mixed source surveys**

A háztartások megkérdezése mellett (vagy helyett) alternatív forrásokat kutatunk fel és hasznosítunk.

- Adminisztratív nyilvántartások használata
- A vásárlói magatartás mintázatainak vizsgálata passzív módon („kenyérmorzsák”)
- Szenzoros adatok begyűjtése
- Big data a piacról

MIXED-MODE SURVEY PROBLÉMÁK

- Az interjú kitöltés esetén segítség áll rendelkezésre, míg önkitöltésnél a válaszadó magára marad.
 - Hosszú kérdőívbe belefásul, nincs aki tovább motiválja
 - Besorolási, módszertani kérdésekben pontatlan lehet az önkitöltés – ismerethiányból fakadó hibák
 - Torzítja a válaszokat az önkitöltés, főleg ha pl. saját képességek megítéléséről van szó (l. IKT – PIAAC felvételek példája)
 - A szubjektív megítélést igénylő kérdésekben a személyes visszajelzés pozitívan torzít, a személytelen negatívan

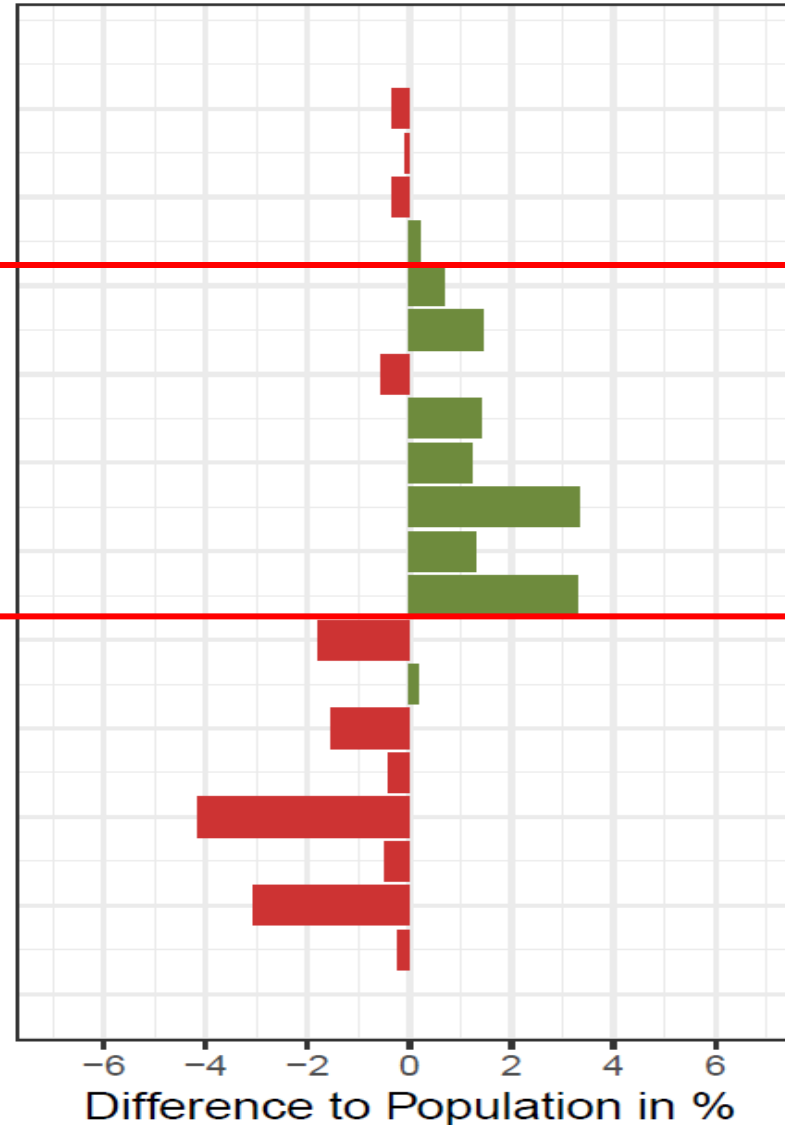
MIXED-MODE SURVEY PROBLÉMÁK (2.)

- A webes önkitöltés külön problémákat eredményez
- Mintakiválasztás és felvétel-módszertani -
 - Mintavételes felvétel esetén mintatorzulás pusztán az adatgyűjtési csatornából fakadóan (pl. csak webes lehetőség a válaszadásra...)
 - Életkor vagy más sajátosságok alapján eltérő válaszadói rétegek azonosíthatóak (pl. telefonos vs. webes)

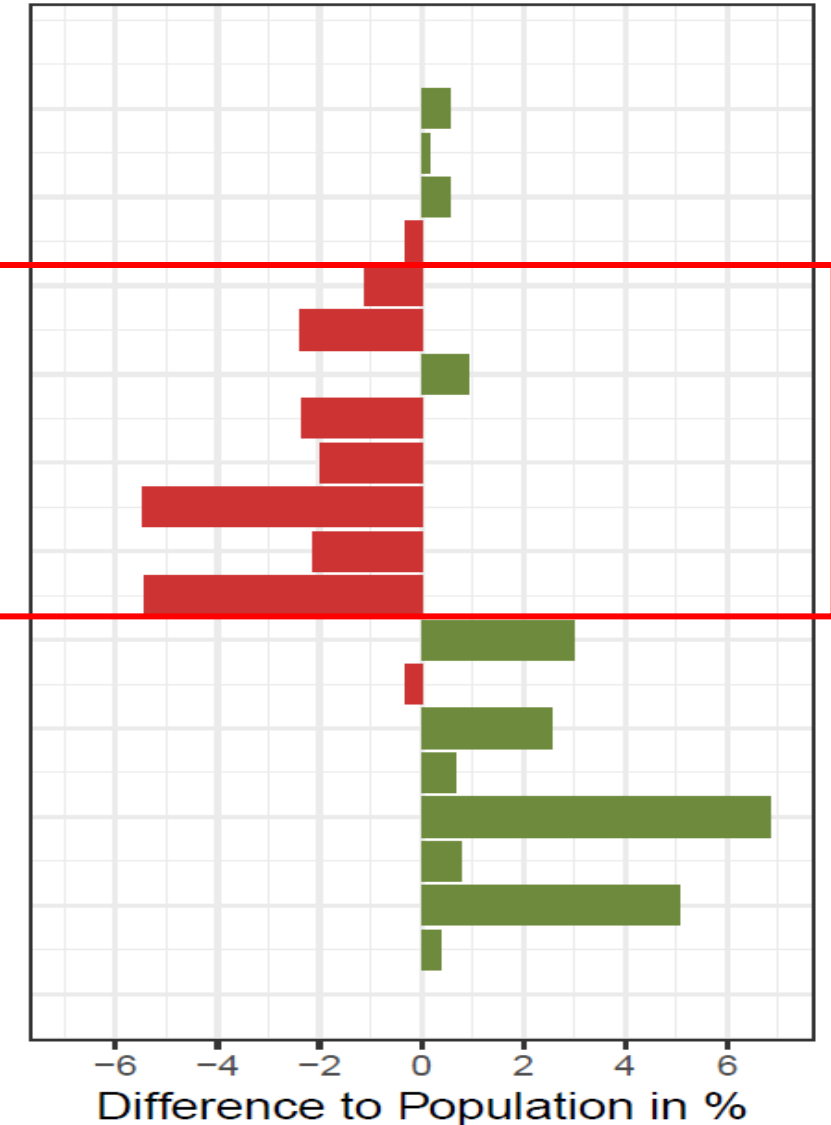
SAMPLE COMPOSITION LFS 2017 LUXEMBOURG

Age	Sex	Nat.	Pop.
15–19	Female	LU	1.9%
		NotLU	1.4%
	Male	LU	2.0%
		NotLU	1.5%
20–29	Female	LU	4.5%
		NotLU	3.6%
	Male	LU	4.5%
		NotLU	3.7%
30–49	Female	LU	7.5%
		NotLU	11.0%
	Male	LU	7.6%
		NotLU	11.3%
50–64	Female	LU	6.6%
		NotLU	4.8%
	Male	LU	6.5%
		NotLU	5.6%
65+	Female	LU	6.2%
		NotLU	2.4%
	Male	LU	5.2%
		NotLU	2.3%

Web Sample



Telephone Sample



MIXED-MODE SURVEY PROBLÉMÁK (3.)

- Platformok közötti technikai átjárás problémája



- A webböngészős felület önmagában nem megoldás!

MIXED-SOURCE SURVEY KIHÍVÁSOK, KÍSÉRLETEK (1.)

Finnország - adminisztratív adatösszekapcsolás

- népesség-nyilvántartás minden egyedi adata összekapcsolható a statisztikai adatfelvétellel
- az oktatási nyilvántartásból a legmagasabb befejezett iskolai végzettség adata (igaz, két évvel korábbról), továbbá
- a nyilvántartott álláskeresők adatbázisa is (havi frissítés).
- Ilyen megoldásokat jelenleg az európai országok mindössze 10%-a alkalmaz.

MIXED-SOURCE SURVEY KIHÍVÁSOK, KÍSÉRLETEK (2.)

Portugália - bankkártya adatok felhasználása (kenyérmorzsák)

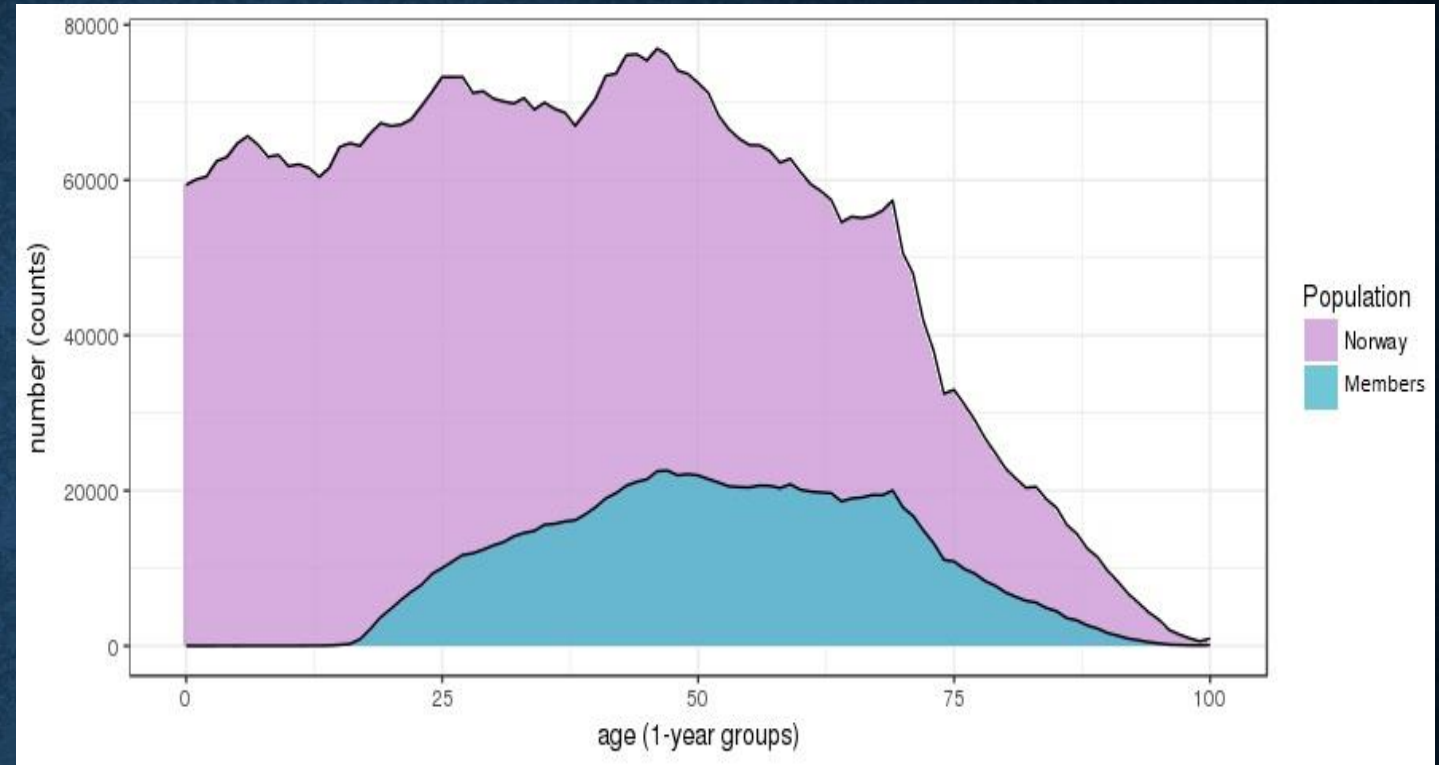
- A háztartások fogyasztási szerkezete és a szegénység mérése céljára
- A bankok legféltettebb kincse – ritka a sikeres adatigénylés
- Vonzó az adatok részletessége és a kártyatulajdonos anonim, de demográfiai jellemzőit részletesen bemutató adatainak elméleti lehetősége
- Egyértelmű lefedettségi problémák, jövedelmi decilisek nincsenek összhangban
- Konklúzió: a társadalom felső 50%-át jobban jellemzi

MIXED-SOURCE SURVEY KIHÍVÁSOK, KÍSÉRLETEK (3.)

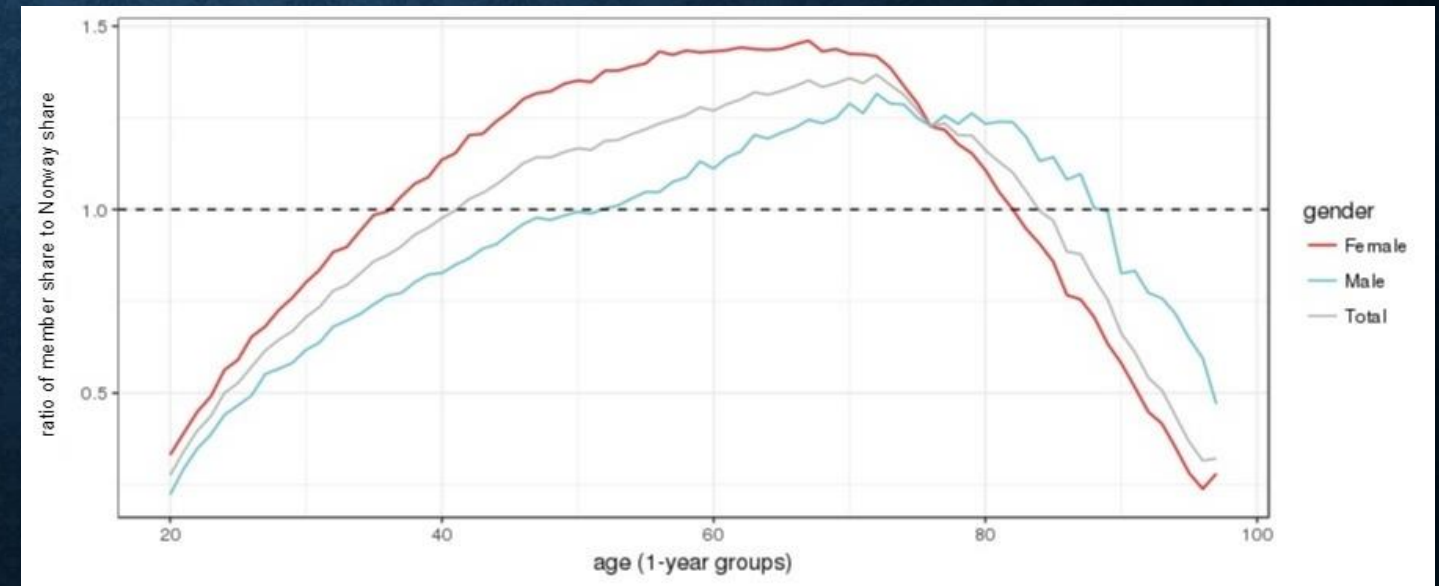
Norvégia - vásárlói pontgyűjtő kártya adatok felhasználása
(kenyérmorzsák)

- A vásárlásokhoz hozzá lehet társítani a vásárló nemét, életkorát és postai irányítószámát.
- A vásárlásról magáról pedig teljes részletességű adatsor állt rendelkezésre (termék, mennyiség, egységár, érték, vásárlás időpontja).
- A megvásárolt termékek 55%-át, az összes költség 56%-át, a tranzakciók 42%-át fedte le.

- *Distribution of population in Norway and of loyalty members by age*



- *Over- and underrepresentation of men and women among loyalty members compared to the total population*



SZENZOROS ADATOK ÉS BIG DATA ÖTLETEK

Mobiltelefonok, okoseszközök (okosóra, karkötő)

- Egészségi állapot (pulzus)
- Sport tevékenység (lépésszámláló, úszás, futás mérése stb.)
- Ingázás (sebesség, jellemző tartózkodási hely) + cellainformációk
- A Big Data adatok az appok és a hálózat tulajdonosának az adatai
- Személyes adatok jogi kérdései
- Egyéb big data források, web scraping stb. ára, további jogi kérdései

DR. KÁRPÁTI JÓZSEF

Neumann János Egyetem GTK – Központi Statisztikai Hivatal

E-mail: karpati.jozsef@gtk.uni-neumann.hu