

NEULINGER ÁGNES

## *A hivatalos statisztikai adatok üzleti értéke, avagy hogyan kapcsolódik a népszámlálás az üzleti kutatásokhoz?*

### *Absztrakt*

A rendszerint tízévente zajló népszámlálás megkerülhetetlen bázisát jelenti az üzleti célú kutatásoknak, hiszen teljes körű adatot szolgáltat a lakosság alapvető jellemzőiről, az életmódot befolyásoló demográfiai tényezőkről és lakhatási körülményekről. Minden olyan üzleti vállalkozás, amely valaha rendelt vagy saját maga készített piackutatást, jó eséllyel hasznosította a népszámlálás adatait is. Ezzel együtt egy értékesítési vezetőnek, marketingesnek vagy cégvezetőnek sokszor nem jelent sokat a census, és bár kétségtelen, hogy a népszámlálás nem üzleti célú adatgyűjtés, kérdés lehet, hogy van-e a kapcsolódásuknak gazdagabb értelme. Jelen tanulmány ennek a kérdésnek a vizsgálatára vállalkozik és mutatja be hazai és nemzetközi források segítségével a népszámlálás és üzleti kutatás kapcsolódási pontjait és ezek jellemzőit. Kulcsszavak: népszámlálás, piackutatás, irodalmi áttekintés

### *Abstract*

Census taking is usually carried out once every ten years and gives fundamental information for marketing research on the main characteristics of the entire population with the help of general demographic variables and data on living arrangements. There is a good chance that every company has ever ordered or conducted a research study has also used Census data. Nevertheless, sales representatives, marketing managers and executive managers often underestimate the value of Census. It is a fact, that Census data do not primarily serve business purposes, yet the question on the business application of Census statistics is relevant. This study presents the business potential of Census data through the introduction of relevant national and international literature.

Key words: census, market research, literature review

## **Bevezetés**

A népszámlálásból elérhető információk alapinformációknak tekinthetők a társadalom jellemzőinek vonatkozásában, és ahogy Dugmore és szerzőtársai megfogalmazták: „a census fontosságát nem lehet eltúloz-

ni”.<sup>1</sup> Eszerint nem kérdés, hogy a népszámlálás nagy jelentőségű minden mintavételes kutatás számára, így az üzleti célú kutatások számára is. Ezzel együtt az üzleti kutatás mindennapi gyakorlatában a népszámláláshoz való kapcsolódás minimálisnak, rutinszerűnek és inkább kiegészítő jellegűnek tűnik. Ugyanez igaz a kutatómódszertani oktatásra az üzleti képzések esetében. A tanórákon a mintavételes kutatáshoz képest – ami az üzleti célú kutatások alapvető megközelítése – a népszámlálás mint egzotikus ellenpont jelenik meg, ahol adott terület minden egyes személyére kiterjedő, teljes körű adatfelvétel valósul meg. Tapasztalataim szerint a piackutatói, marketingszakmai környezetben a census jellemzően az alapadatai kapcsán érdekes, úgymond, ez az „alapvető háttér”, és így „van mihez képest súlyozni”, valamint a lassú és az elavult jelzők szoktak társulni hozzá. De mi van mindezekén túl? Hogyan kapcsolódik össze a népszámlálás és az üzleti kutatás területe? Jelen tanulmány erről a kapcsolódásról szól elsődlegesen a hazai népszámlálás jellemzőiből kiindulva, amely kiegészítésre kerül nemzetközi tapasztalatok és megoldások ismertetésével.

## A népszámlálások jellemzői és értelmezésük az üzleti kutatások kontextusában

A Központi Statisztikai Hivatal (KSH) honlapján található megfogalmazás szerint „a népszámlálás olyan rendszeresen, általában 10 évente ismétlődő, teljes körű állapotfelvétel, amely egy adott terület (ország) népességének egészére kiterjed és egy előre meghatározott időpontra vonatkozik. A népszámlálás eszmei időponthoz kötődik, teljes körű és igen sokoldalú a felhasználása. Az összeírásra az egész országban egyidejűleg, azonos tartalommal és egységes módszertani alapon, minden lakásra és személyre kiterjedően kerül sor.”<sup>2</sup> Magyarországon 1869 óta tartanak rendszeresen népszámlálásokat, és ezeknek a lebonyolítása a 60-as évek óta a nemzet-

1 Dugmore et al. 2011.

2 Lásd a KSH 2011-es népszámlálási honlapján: [http://www.ksh.hu/nepszamlalas/a\\_nepszamlalas\\_celja\\_sb](http://www.ksh.hu/nepszamlalas/a_nepszamlalas_celja_sb)

közi követelményeket is figyelembe veszi.<sup>3</sup> A következő népszámlálásra az eredeti tervek szerint 2021 május–júniusában került volna sor, a koronavírus-járvány miatt azonban elhalasztották. A halasztást követően a következő népszámlálást – több más országhoz hasonlóan – 2022-ben fogják megtartani.

A népszámlálás kérdőívében a népességre és a lakásállományra vonatkozó kérdések szerepelnek, ezzel a kutatás képet ad a lakosság demográfiai jellemzőiről, valamint annak változásáról.<sup>4</sup> A hosszú távú adatsorokból származó eredmények alapvetőnek tekinthetők, ezeket jellemzően senki nem vitatja, így mindenki által elfogadott alapot, viszonyítást adnak. A népszámlálás fontossága abban rejlik, hogy ez az egyetlen adatfelvétel, amely az ország teljes lakosságára és életkörülményeire kiterjed, pontos adatokat ad a különböző gazdasági és társadalmi szférák döntéshozóinak, továbbá adatait nemzetközi szinten is elfogadják.<sup>5</sup> Az üzleti felhasználásához érdemes tételesen számba venni azt, hogy a népszámlálás (1) miről, (2) mikor, (3) kiknek és milyen feltételek mentén szolgáltat adatot. Ezen szempontok mentén először bemutatásra kerülnek a népszámlálás hazai sajátosságai, amit az üzleti kutatások nézőpontjából történő értékelés kísér.

## Miről?

Adat szintjén a népszámlálás a következőket tartalmazza a kérdőívei alapján:<sup>6</sup>

- A népességre vonatkozó személyi kérdések, úgymint nem, születés helye, ideje és születéskori lakhely, ideiglenes és állandó lakcím, állampolgárság, nyelvismeret, családi állapot és a családban, háztar-

3 További részletekért lásd: [http://www.ksh.hu/nepszamlalas/a\\_magyarorszagi\\_nepszamlalások\\_tortenete\\_sb](http://www.ksh.hu/nepszamlalas/a_magyarorszagi_nepszamlalások_tortenete_sb)

4 A népszámlálás 1960 óta magában foglalja az intézetekben élők, úgymint kollégiumok, munkásszállók, szociális otthonok, kórházak, szállodák adatainak a számbavételét is.

5 A KSH honlapján szereplő, *A népszámlálás fontossága* cím alatt jegyzett tényezők alapján, részletesen lásd: [http://www.ksh.hu/nepszamlalas/a\\_nepszamlalas\\_celja\\_sb](http://www.ksh.hu/nepszamlalas/a_nepszamlalas_celja_sb)

6 A népszámlálás 2001, 2011 és 2022. évi adatfelvételek kérdőíveinek forrásai a KSH honlapjai szerint: [http://www.nepszamlalas2001.hu/hun/\(2001\)](http://www.nepszamlalas2001.hu/hun/(2001)), [http://www.ksh.hu/nepszamlalas/a\\_nepszamlalas\\_kerdesei\\_sb](http://www.ksh.hu/nepszamlalas/a_nepszamlalas_kerdesei_sb) (2011) és [https://www.ksh.hu/nepszam2021-kerdoivteszt-menu-kerdoiv\(2021\)](https://www.ksh.hu/nepszam2021-kerdoivteszt-menu-kerdoiv(2021)).

tásban betöltött szerep, házasságkötés és bejegyzett élettársi kapcsolat ideje, élve született gyermekek születési adatai, iskolába járás, iskolai végzettség, gazdasági aktivitás, munkavégzés vagy tanulás helye, munkavégzéssel, tanulással összefüggő napi közlekedés. Önkéntes kérdésként jelenik meg a nemzetiség, anyanyelv, családi, baráti körben beszélt nyelv, vallás, tartós betegség és fogyatékoság.

- A lakásállomány kapcsán feltett kérdések, úgymint a lakás címe, rendeltetése (típusa), építési éve, lakóegység típusa, falazata, lakáshasználat formája és rendszeressége, tulajdoni jellege, helyiségeinek száma, alapterülete, kommunális ellátottsága, felszereltsége, lakásban lakó háztartások és személyek száma.

A népszámlálás során feltett kérdések a lakásállomány kapcsán alapvetően stabilak, egyetlen változás 2022-ben az internet-hozzáférés megjelenése a kérdőívben. Ezzel szemben a népességre vonatkozó kérdések esetében több változás látható a legutóbbi három kérdőív (beleértve 2022-es évet) alapján. Eszerint a 2011-es adatfelvételben a családi állapot kérdés mellett a családban, háztartásban betöltött szerep rögzítése (úgymint férj/feleség, élettárs, gyermekével egyedül élő szülő, gyermek, más rokon, nem rokon) a korábbi táblázat helyett a személyi kérdőív egyik kérdéseként jelent meg. Majd a lista 2022-ben kiegészül – a 2016-os mikrocenzushoz hasonlóan – a bejegyzett (azonos nemű) élettárs kategóriával. Ugyancsak a mikrocenzusban szerepelt először a külföldön élés, külföldről hazaköltözés témája, amely a 2022-es népszámlálás kérdőívében is helyet kap.

Praktikusan az említett adatok hasznosak lehetnek az üzleti világ azon szereplőinek, akik célcsoportjukat keresik, és azt területi alapon szeretnék megtalálni. Ilyen lehet egy lakásfelújítással foglalkozó vállalkozó, aki információt szeretne arról, hogy mely környék lakásai vannak leginkább előregedve (a lakóépület építési ideje alapján), egy magánóvodát nyitni kívánó óvodapedagógus, hogy megtudja, mely városban/városrészben élnek leginkább hat év alatti gyermekes családok, vagy bármely vállalkozó, aki azt szeretné tudni, hogy működési környezetében hány év a lakosok átlagéletkora. Ezek az adatok egyéni adatkéréssel elérhetők a hazai Központi Statisztikai Hivataltól, és amennyiben lekérésre kerülnek a foglalkozási és iskolázottsági jellemzők is – amelyek segítségével a társadalmi tagozódás megbecsülhető –, akkor azzal a piaci potenciálok már közelíthetők lesznek.

Ha a népszámlálásból elérhető adatokat a napi üzleti gyakorlat szempontjából nézzük, akkor ezek hasznossága érintőlegesnek tűnik. A népszámlálásban felvett demográfiai kérdések mind fontosak ugyan, de üzleti szempontból önmagukban csak korlátozottan érdekesek a menedzserek számára. Erelles, Fukawa és Swayne ennek kapcsán úgy fogalmaz, a cenzust egy női ruházati kiskereskedelmi üzletlánc vásárlói adataival összehasonlítva, hogy „mindkét adathalmaz gazdag, nagy és *insight*ot ad, de csak az utóbbi teszi lehetővé a marketingvezetőnek, hogy aktuális és tényalapú döntéseket hozzon”.<sup>7</sup> Eszerint a mindennapi marketing- és általában üzleti jellegű döntéshozatalt a népszámlálási adatok csak közvetve tudják támogatni.

### *Mikor?*

Az adatfelvétel – a fentebb jelzett meghatározás szerint – általában tízévente történik, az adatok publikálása pedig az adatfelvételt követő évek alatt, jellemzően több hullámban. Ezért szokták a népszámlálás egyik legfőbb kritikájaként említeni a lassúságot. A legutóbbi, 2011-es népszámlálási eredmények publikálásának időbeli jellemzőit – annak tematikus tanulmányai szerint – az 1. táblázat foglalja össze. A táblázatból látható, hogy a népszámlálás adatgazdagsága évek alatt dolgozható fel, gyakorlatilag a mikrocenzusig<sup>8</sup> folyamatos munkát adva a kutatóknak. Ez az időbeli lépték a stratégiai döntéshozatali kérdésekhez illik inkább a napi operatív üzleti döntésekkel szemben.

7 Erelles–Fukawa–Swayne 2016.

8 A mikrocenzus két teljes körű népszámlálás között mintavételes kutatással gyűjt adatot a társadalomról. Legutóbb 2016-ban került lebonyolításra mikrocenzus.

*1. táblázat. A népszámlálás adatainak publikálása  
a megjelenés ideje szerint*

	<i>Megjelenés éve</i>
Előzetes adatok	2012
Demográfiai adatok	2013
Háztartások, családok életkörülményei	2013
Iskolázottsági adatok	2013
A lakások és lakóik	2013
Országos és területi adatok	2013
Foglalkoztatás, munkanélküliség, ingázás	2014
Lakásviszonyok	2014
Fogyatékosággal élők	2014
Nemzetiségi adatok	2014
Vallás, felekezet	2014
A népesség gazdasági aktivitása	2014
A népesség iskolázottsága (azon belül szakképzettség)	2014
A fogyatékosággal élők helyzete és szociális ellátásuk	2015
A család típusok jellemzői	2015
A foglalkoztatottak napi ingázása és közlekedése	2015
A társadalom rétegződése	2015
Műhelytanulmányok	2016
Módszertani kötet	2016

### *Kinek és milyen feltételek mentén elérhetően?*

A népszámlálás adatait elsődlegesen az ország társadalmi-gazdasági és demográfiai kérdéseivel foglalkozó szakemberei használják, az állami, a magán- és a civil szféra képviselői egyaránt. Ahogy Collins és Wallace megállapítják, a népszámlálási adatok hagyományosan a helyi és országos kormányzati munkában fontosak, amely mellett a kilencvenes évekre megjelent az üzleti célú felhasználás iránti érdeklődés is. Angliai tapasztal-

talataik alapján jelzik, hogy a cenzusadatok értékét egyre inkább felismerte az üzleti világ, amely belső használatra és külső értékesítésre is elkezdett a népszámlálási adatokból célpiacon meghatározására, célzására alkalmas, kombinált adatokat képezni.<sup>9</sup>

A teljes lakosságra vonatkozó adatokat gyakran használják a tudományos élet kutatói a társadalmi és területi fókuszú kutatásaikhoz. Továbbá az adatok hasznosulhatnak az üzleti döntéshozatal során, amennyiben az a lakossági jellemzőkre és területi sajátosságokra irányul.

A fenti 1. táblázatból látható, hogy a részletes adatok és az elkészült elemzések folyamatosan kerülnek közlésre bárki számára elérhető módon, a KSH oldaláról letölthető formában. Emellett a 2011. évi népszámlálás és a 2016. évi mikrocenzus esetében a KSH kutatószobájában az adatbázisok is hozzáférhetőek kutatók számára, de kizárólag tudományos kutatási célra. Az adatbázisokat tekintve a mikrocenzusok teljes adatállománya, míg a népszámlálások esetében a népszámlálásból vett tízszázalékos minta áll a tudományos célú kutatók rendelkezésére. A kutatói hozzáférés ezen adatokhoz a kutatásra előzetesen összeállított adatállományok esetében térítésmentes, egyedi adat-összeállítás esetében térítésköteles. Nem az akadémiai szférából érkező kutatók is igényelhetnek egyedi táblázatos adatokat; ezeket a kéréseket a KSH megvizsgálja abból a szempontból, hogy az igényelt táblázatos adatok a kért bontásban rendelkezésre bocsáthatók-e, illetve adatvédelmi szempontból kiadhatók-e. Az adatállomány előállításáért a KSH költségtérítést kérhet.

## Hogyan használják fel a népszámlálás adatait az üzleti kutatások?

Hazai és idegen nyelvű publikációkban – vegyes tudományterületi jellemzőkkel – rendszeresen a cenzusok adatait hivatkozó tanulmányok. Ezen elemzések többsége egyszerűen csak felhasználja a cenzus adatait saját munkájának a támogatására, a minta értékelésére és a kutatás eredményeinek ellenőrzésére, míg egy kisebb részük üzleti, fogyasztói né-

9 Collins–Wallace 1995.

zópontból értékeli a népszámlálások adatait, eljárásait. A téma kapcsán áttekintett tanulmányok tanúsága szerint tipikusan a lakossági demográfiai információk és a területi, földrajzi célzashoz szükséges adatok érdekesek az üzlet számára. Igaz, a népszámlálás üzleti alkalmazhatóságát érintő munkákban visszatérő gondolatként jelenik meg a lassú és túlságosan aggregált minősítés, amely így korlátozott betekintést ad a lakosság fogyasztási és életmódjellemzőibe.<sup>10</sup>

Ezzel együtt számos olyan terület van, ahol üzleti jellegű kutatási kérdésekkel összefüggésben megjelenik a népszámlálás adatainak a felhasználása, ezek az alábbiakban kerülnek bemutatásra.

### **A census mint viszonyítási pont**

Az egyik leggyakoribb megoldás mind a nemzetközi, mind a hazai publikációkban. Adott célcsoportra vonatkozó kutatási eredmények értelmezéskor megjelenik, hogy mi tudható a tágabb kontextusú censusból, és ahhoz képest egy adott kutatás mire jutott.<sup>11</sup>

### **Mintavételes kutatások támogatása**

A másik meghatározó terület a népszámlálás eredményeinek felhasználásában a mintavételhez, a kutatási minta kialakításához kötődik. A lakosságra vonatkozó kutatások rendszeresen hivatkozzák a mintájuk kialakításához a népszámlálási adatokat,<sup>12</sup> vagy ahogy a KSH munkatársa a 2011-es népszámlálás kapcsán fogalmazott: „a piackutatók ehhez viszonyítva választják ki az úgynevezett reprezentatív mintáikat”.<sup>13</sup>

### **Szegmentálás cenzusadatok alapján**

A népszámlálás adatai alkalmasak lehetnek a lakosság szegmentálására is. Campbell–Ryley–Thring<sup>14</sup> kutatásukban népszámlálási adatokon

10 Lásd például Birkin 2019.

11 Lásd például Brown 2003.

12 A legtöbb kérdőíves kutatás hivatkozható ehhez, így például McCarty és Shrum 1993 nemzetközi munkája, vagy Neulinger és Radó 2015 hazai cikke, amelyek a mintavételi súlyokhoz használták a census adatait.

13 Virágh 2012.

14 Campbell–Ryley–Thring 2012.



végeztek klaszterelemzést annak azonosítására, hogy egy adott angliai nagyváros területén belül hol élhetnek azok, akik alternatív üzemanyagot használnak. Ehhez először leírták a potenciális felhasználó profilját a témában elérhető korábbi kutatások alapján, majd a helyi népszámlálási adatok segítségével területileg meghatározták, hogy adott profilú potenciális felhasználók hol élhetnek. Igaz, ehhez szükség volt arra a földrajzi alapú klasszifikációs rendszerre, amelyet az Egyesült Királyság cenzusa során határoz meg az angol statisztikai hivatal. Ugyancsak szegmentációt végzett Clark, aki az autóbirtoklás tényét a népszámlálás lakossági, háztartási adatai alapján modellezte, és arra a következtetésre jutott, hogy a becslés a népszámlálási adatokból megbízhatóan megtehető.<sup>15</sup> Továbbá, ha földrajzi alapú szegmensekben gondolkodunk, akkor a területi lakossági demográfiai adatok segíthetnek abban, hogy hová érdemes kiskereskedelmi egységeket telepíteni,<sup>16</sup> és általában a marketingkommunikációs tevékenységek tervezésében is.<sup>17</sup>

### **Adatkombinációk**

Amennyiben a népszámlálás adatai összeköthetők más adatbázisokból származó adatokkal, az gazdag elemzési lehetőséget biztosíthat piackutatási, üzleti célokra. Waddell és szerzőtársai dohányipari promóciókra fókuszáló elemzése egy dohányipari kereskedelmi hirdetések mérő kérdőív adatait kötötte össze az Egyesült Államok cenzusának demográfiai adataival egy New York államra vonatkozó elemzés esetében. Elemzésük összefüggést talált a vizsgált földrajzi körzetekben élő fiatalok aránya és a mentolos dohánytermékek promóciója között.<sup>18</sup> Hasonlóképpen Trivedi számos adatot kapcsolt össze elemzése során, úgymint egy üzletlanc vásárlói adatai, az üzletek jellemzőire vonatkozó adatbázis és az amerikai cenzus demográfiai adatai. Ezek segítségével meg tudta határozni a földrajzi lokáció és termékkategória jelentőségét az egészséges termékek vásárlása során.<sup>19</sup>

15 Clark 2009.

16 Dugmore et al. 2011.

17 Mitchell–McGoldrick 1994.

18 Waddell et al. 2016.

19 Trivedi 2011.

## Kiváltható a népszámlálás más adatgyűjtési megoldással?

A Big Data idején, amikor számos adatforrásból szinte végtelen adat elérhető, felmerül a kérdés, hogy szükség van-e a népszámlálásra továbbra is. A kérdés az üzleti kutatások szempontjából is releváns: kell továbbra is census, vagy a technológiai fejlődésnek köszönhetően a vállalatok saját maguk is hozzáférnek olyan adatokhoz, amelyek kiválthatják azt? Összességében, a kérdésben megszólalók véleménye alapján a válasz inkább nemleges. Struijs és szerzőtársai számos olyan – egymással összefüggő – problémát megneveznek, amelyek kérdéssé teszik a census helyettesíthetőségét, úgymint a módszertani megbízhatóság, az adatminőség és az adatok megbízható értékelése. Meglátásuk szerint a Big Data használhatóságának egyik legnagyobb kihívását a módszertanhoz fűződő aggodalmak jelentik. Nevezetesen az, hogy a keletkező adatok nagy része (például a közösségi média adatai) nem kutatási célból keletkezik elsődlegesen, így célsokaságuk, struktúrájuk és minőségük nem előzetesen meghatározott, így a mintavételes eljárásokhoz kapcsolódó hagyományos statisztikai elemzések nem mindig lehetségesek ezeken az adatokon. Ráadásul az adatok sokszor nem könnyen értelmezhetők.<sup>20</sup> Shearmur egyértelműen ki is mondja, hogy a népszámlálás a technológiai fejlődés által lehetővé vált adatrögzítésekkel és ezen adatok elemzésével nem kiváltható. Ahogy megfogalmazta: „a Big Data üdvözölhető és érdekes új eszköz, amely számos *insight*ot hoz”, főleg ha annak adataihoz kutatók is hozzáférhetnek. Ezzel együtt a társadalom struktúrájának és jellemzőinek megértéséhez a censzustípusú adatok megkerülhetetlenek. Érvelése szerint a népszámlálás éppen azt az alapvető információt adja meg, amihez képest a Big Data-adatok is értelmezhetők, azaz census nélkül nem áll rendelkezésre a teljes lakosságot leíró információ, amihez képest a mintavételek elvégezhetők, és a reprezentativitás értelmezhető.<sup>21</sup>

Egy másik nézőpontot képviselve Dugmore és szerzőtársai – a census értékét és szükségességét elismerve – kiemelik annak fontosságát, hogy a censusok adatainak relevánsnak kell tudniuk maradni a gyors társa-

20 Struijs et al. 2014.

21 Shearmur 2015.

dalmi és gazdasági változások közepette. A szerzők megkerülhetetlennek tartják a változás és fejlődés szükségességét, hiszen egy gyorsan változó világban hiába megbízható egy adat, ha nem érvényes, mert a turbulens környezetben hamar értékét veszti. Megfogalmazásuk szerint szükséges, hogy rövidebb idő teljen el két adatfelvétel között annak érdekében, hogy „alkalmazkodjon a lakosság megnövekedett mobilitásához, a határok közötti mozgásához és a gazdasági bevándorláshoz”. Tekintettel a kérdőívvel történő teljes körű adatfelvétel magas költségére, elkerülhetetlennek tartják a közigazgatásban keletkező adatokra épülő elemzéseket és gyakoribb mintavételes kutatásokat.<sup>22</sup>

Erre a javaslatra rímelve a KSH is foglalkozik a más céllal regisztrált adatok felhasználhatóságával. A hivatal 2017-ben megjelent műhelytanulmánya közel 160 oldalon foglalkozik az ún. másodlagos adatforrások használatával a statisztikában. Ebben a dokumentumban jelzik, hogy 2016. júniusi állapot szerint a hivatal közel kétszáz adminisztratív adatforrást vesz át, továbbá hozzáfér nyolcvan olyan statisztikai adatgyűjtésnek az adataihoz, amelyet más intézmény gyűjt. Ezzel együtt a népszámlálás kapcsán a műhelytanulmány hangsúlyozza, hogy az adminisztratív adatforrások használata együtt jár azzal, hogy az információ az ezekben a rendszerekben elérhető adatokra korlátozódik, ami nem felel meg teljeskörűen az eddigi népszámlálásokban felvett adatoknak.<sup>23</sup>

A népszámláláson túlmenően vizsgálta a Big Data felhasználhatóságát a hivatalos statisztikában Giczi és Szőke tanulmánya. Az elemzés áttekinti a mobilkommunikációs, az internetes, a szenzoros és a folyamatgenerált tranzakciók alkalmazási lehetőségeit, és megállapítja, hogy ezeket sokszor megnehezítik vagy nem teszik lehetővé az adatvédelmi szempontok, a jelenlegi adatátadási gyakorlat jellemzői vagy éppen annak hiánya. Véggökövetkeztetésük szerint a Big Data „megtalálja majd a helyét a hivatalos statisztikában anélkül, hogy a hagyományos adatgyűjtési eljárásokat feleslegessé tenné”.<sup>24</sup>

22 Dugmore et al. 2011.

23 KSH 2017.

24 Giczi–Szőke 2017.

## Üzleti adatokból hivatalos statisztika

Részben a Big Data területét érinti, így átfed az előző témakörrel a kérdés, hogy az üzleti szférából származó adatok miként használhatók fel hivatalos statisztikai célokra, népszámlálási témakörökben és azokon túlmutatóan is. Giczi és Szőke elemzése bemutat egy olyan nemzetközi példát, amikor vállalati adatok segítik a nemzeti adatgyűjtést. Konkrétan a mobilszolgáltató vállalatoknál keletkező és helymeghatározásra alkalmas adatokról van szó, amelyeket az érintett vállalatok anonim módon bocsátanak a hivatalos statisztika rendelkezésére Észtországban. Ez az információmegosztás lehetővé teszi az országhatár-átlépésekre vonatkozó adatok megbízható gyűjtését. A KSH adatgyűjtési gyakorlatához kapcsolódóan azt az elképzelést fogalmazzák meg a kutatók, hogy a vásárlási folyamat során keletkező adatok (úgy mint bankkártyahasználat, online pénztárgép adatok) támogathatják a fogyasztásstatisztikai célból megvalósuló adatgyűjtést, hiszen ezzel a háztartások fogyasztási jellemzőit és kiadásait pontosabban lehetne megismerni.<sup>25</sup>

A kereskedelmi vállalatoknál keletkező adatok hivatalos statisztikai felhasználásáról értekezik Dugmore tanulmánya az Egyesült Királyság esetében. Érvelése szerint a vállalati adatok ugyan nem reprezentatív mintái a teljes lakosságnak, azaz torzítottak lehetnek területi és státusz-alapon, valamint összetételüket a különböző marketingakciók is befolyásolják, mégis kiterjedt méretük és aktualitásuk miatt trendek előrejelzésére alkalmasak lehetnek. Dugmore nevesíti azokat az adatforrásokat, amelyek a statisztikai hivatal demográfiai adatgyűjtését támogathatják, ezzel a népszámlálási adatfelvételt is megkönnyíthetik. Ilyen lehet például a lakossági címlista ellenőrzése a szolgáltatói vállalati oldalon megtalálható címlistákkal (például villamosművek). Elemzésében azt is megállapítja, hogy az adatmegosztás akadálya lehet az adatbiztonság és adatvédelem kérdése, miközben a vállalati szféra részvételét ösztönözheti, hogy az együttműködés jó hatással lehet a kormányzattal való viszonyra.<sup>26</sup>

Ugyancsak az Egyesült Királyságban foglalkozik a census fogyasztói adatokkal való támogatásával Newing és szerzőtársai munkája. Cikkük-

25 Giczi–Szőke 2017.

26 Dugmore 2010.

ben a lakossági áramfogyasztás intelligens eszközökkel való mérésének statisztikai lehetőségeit tárgyalják. Az intelligens fogyasztásmérés a gáz- és villamosenergia-felhasználás esetében azt jelenti, hogy a háztartási fogyasztásmérés adata automatikusan beküldésre kerül a szolgáltatónak, így nincs szükség leolvasásra, bediktálásra és becslés alapján történő fizetésre, mivel a tényleges fogyasztási adatok azonnal rendelkezésre állnak. Ebből az adatból a háztartás jellemzőire vonatkozóan származtatható információ, amely így közel valós idejű hozzáféréssel támogathatja a demográfiai célú adatgyűjtést.<sup>27</sup>

Mindezek a javaslatok értékesek és fontos lehetőséget kínálnak a statisztikai adatok gyorsabb és költséghatékonyabb gyűjtésére. A kérdés, ami minden esetben felmerül, az a jogi megfelelés, és hogy adott mérések milyen hatókörűek, azaz képesek-e teljes körű információt adni vagy jól lehatárolhatóan kiegészítők lenni, hiszen a hivatalos statisztika számára így lehetnek igazán értékesek.

## Zárógondolatok

A népszámlálás során keletkező adatokat a vizsgált szakirodalmi források mind alapvetően fontosnak tartják, miközben visszatérően megjelenő gondolat a más adatforrások, illetve a Big Data alkalmazása a census kiváltására vagy kiegészítésére. Elmondható az áttekintett tanulmányok alapján, hogy a népszámlálás olyan információkat ad a teljes lakosságról, ami viszonyítási pontnak számít, így megkerülhetetlen. Ezzel együtt az üzleti hasznossága ezen adatoknak – az időbeli frissességen túl – erősen összefügg azzal, hogy az adatfelvétel milyen változókra terjed ki, illetve területi alapon milyen részletezettséggel közlik azokat. Nagy-Britanniában a 2001-es népszámlálás újdonsága volt, hogy a felvett adatokból ún. társadalmi tagolódást számoltak egy algoritmus segítségével, amely olyan változókat vett ehhez figyelembe, mint például a foglalkozás, gazdasági aktivitás, lakáshasználat. Mivel ez a mutató hasonló a gyakorlati marketinggel és piackutatással foglalkozók körében jól ismert ABCDE (Esomar)

27 Newing 2016.

státuszbesoroláshoz, a társadalmi pozícióra vonatkozó becslés értékesebbé tette a népszámlálás adatait az üzleti felhasználás számára is. Továbbá ebben a vizsgálatban határoztak meg először olyan kisebb régiókat<sup>28</sup> – ezek Angliában és Walesben jellemzően egyenként 125 háztartást foglalnak magukba –, amelyek így pontosabb szegmentációt és célzást tettek lehetővé felhasználói oldalról.<sup>29</sup> A társadalmi tagozódás mutatójának bevezetése azzal együtt fontos újdonságot jelentett, hogy a tapasztalatok szerint ez a számítás elsősorban a 16 és 64 év közötti lakosság esetében bizonyult megbízhatónak, míg az idősebb és nyugdíjas, illetve inaktív népesség körében kevésbé működött jól.<sup>30</sup>

Magyarországon a népszámlálás adataiból a foglalkozási rendszer segítségével kerül meghatározásra a magyar társadalom rétegződése. Az adatok területi szempontból elsődlegesen településtípus-, megye- és településszintű formában érhetőek el. A nagy-britanniai példa azt jelzi, hogy a mutatószám szintű feldolgozás és a finomabb területi adat növeli az adatok értékét. Hazai viszonylatban Nagy Péter Tibor elemzése példa arra, hogy miként lehet a népszámlálási adatokból komplexebb társadalmi tagolódást meghatározni. A kutató a cenzust az iskolai egyenlőtlenségek bemutatásához használta fel, és ehhez határozta meg a társadalmi tagoltságot.<sup>31</sup>

Továbbá megállapítható, hogy korunk folyamatos technológiai fejlődése és a keletkező adatok hatalmas mennyisége mellett a nemzeti statisztika sem tud elmenni. A Big Data egyrészt versenytársa a hivatalos statisztikának, másrészt segítője is lehet. Üzleti célú elemzések esetében állandó kritika a népszámlálás adatainak relevanciája és az elavulás problémája.<sup>32</sup> Ennek kapcsán is lényegesnek tűnik, hogy a statisztikai hivatalok együttműködjenek más adatszolgáltatókkal, hiszen ez lehet az alkalmazkodás és lépéstartás útja. Ennek megfelelően a magyar Központi Statisztikai Hivatal – más nemzetközi társához hasonlóan – aktívan foglalkozik a más forrásból elérhető adatok hasznosíthatóságának lehetőségével, illetve a másodlagos adatforrások, közigazgatási regiszterek felhasználásával. Erre mind a gyorsan változó társadalmi-gazdasági környezet, mind

28 *Output areas*, vagy rövidítve OAs.

29 Leventhal 2003.

30 Meyer–Moy 2004.

31 Nagy 2017.

32 Lásd Dugmore et al. 2011.

a költségek nyomása ösztönző hatással van. Ráadásul a statisztikai hivatalok, köszönhetően módszertani megbízhatóságuknak, stabilitásuknak, a jogszabályi jellemzőknek és az üzleti érdekek fölött álló státuszuknak, egyfajta „megbízható harmadik félnek”<sup>33</sup> is számítanak adatmegosztás esetében. Ez azt jelenti, hogy a nagytömegű, sokszor bizonytalan minőségű adatok világának adhatnak hitelesítést, valamint a máshol keletkezett érzékeny adatokat rajtuk keresztül könnyebben megoszthatják azok az üzleti versenytársak is, akik egyébként más módon nem engednék hozzáférést adataikhoz.

Másrészt a népszámlálásra nemcsak úgy tekinthetünk, hogy az mennyiben képes hasznos adatokat eredményezni üzleti célú felhasználásra, hanem úgy is, hogy az üzleti adatok mennyiben tudják támogatni a cenzust. Ebben a tanulmányban több olyan megoldás is bemutatásra került, amelyek a piaci működéshez, üzleti tranzakciókhoz, fogyasztáshoz köthető adatok statisztikai felhasználhatóságát érintették. Ezen megoldások alkalmazásának sokszor vannak jogi vagy technikai akadályai, amelyek azonban sikeresen feloldhatók lehetnek, ahogy azt az észtországi mobilszolgáltatók által átadott adatok példája is jelzi.

Jelen tanulmány abból a célból jött létre, hogy bemutassa a népszámlálás és az üzleti jellegű kutatások kapcsolatát. Az áttekintett kutatási jellemzők, gyakorlatok és példák alapján elmondható, hogy a triviálisnak tűnő válasz mellett, miszerint a népszámlálás az üzleti kutatások szempontjából alapvető, és egyfajta origónak számít, kiderült, hogy a két terület kapcsolata messze túlmutat mintavételi és alapidemográfiai kérdéseken.

## Irodalom

- Birkin, M. 2019: Spatial data analytics of mobility with consumer data. *Journal of Transport Geography*, 76, 245–253.
- Brown, C. 2003: Consumers' preferences for locally produced food: A study in southeast Missouri. *American Journal of Alternative Agriculture*, 18, 14, 213–224.
- Campbell, A. R. – Ryley, T. – Thring, R. 2012: Identifying the early adopters of alternative fuel vehicles: A case study of Birmingham, United Kingdom. *Transportation Research Part A: Policy and Practice*, 46, 8, 1318–1327.

33 Struijs 2014.

- Clark, S. D. 2009: The determinants of car ownership in England and Wales from anonymous 2001 census data. *Transportation Research Part C: Emerging Technologies*, 17, 5, 526–540.
- Collins, J. – Wallace, J. 1995: The use of Census data for target marketing—the way forward for the tourism industry. *Journal of Vacation Marketing*, 1, 3, 273–280.
- Dugmore, K. 2010: Information collected by Commercial Companies: What might be of value to Official Statistics? The case of the UK Office for National Statistics. *RatSWD Working Paper*, 151. Berlin, Rat für Sozial- und Wirtschaftsdaten (RatSWD).
- Dugmore, K. – Furness, P. – Leventhal, B. – Moy, C. 2011: Beyond the 2011 census in the United Kingdom: with an international perspective. *International Journal of Market Research*, 53, 5, 619–650.
- Erevelles, S. – Fukawa, N. – Swayne, L. 2016: Big Data consumer analytics and the transformation of marketing. *Journal of business research*, 69, 2, 897–904.
- Giczi, J. – Szőke, K. 2017: Hivatalos statisztika és a Big Data. *Statisztikai Szemle*, 95, 5, 461–490.
- KSH 2017: Másodlagos adatforrások használata a statisztikában. Általános ismeretek és a hazai gyakorlat. *Műhelytanulmányok 11*. Budapest, [k. n.]
- Leventhal, B. 2003: Leveraging the census for customer analysis. *Journal of Targeting, Measurement and Analysis for Marketing*, 12, 1, 11–19.
- McCarty, J. A. – Shrum, L. J. 1993: The role of personal values and demographics in predicting television viewing behavior: Implications for theory and application. *Journal of Advertising*, 22, 4, 77–101.
- Mitchell, V. – McGoldrick, P. J. 1994: The Role of Geodemographics in Segmenting and Targeting Consumer Markets: A Delphi Study. *European Journal of Marketing*, 28, 5, 54–72. <https://doi.org/10.1108/03090569410062032>
- Meier, E. – Moy, C. 2004: Social grading and the census. *International Journal of Market Research*, 46, 2, 141–170.
- Nagy, P. T. 2017: Az iskolázottsági egyenlőtlenségek és a 18 év alatti népesség a 2011–2016-os időszakban. *Educatio*, 26, 4, 540–556.
- Neulinger, Á. – Radó, M. 2015: A gyermeknevelés hatása a háztartások kiadási szerkezetére. *Statisztikai Szemle*, 93, 7, 662–688.
- Newing, A. – Anderson, B. – Bahaj, A. – James, P. 2016: The role of digital trace data in supporting the collection of population statistics—the case for smart metered electricity consumption data. *Population, Space and Place*, 22, 8, 849–863.
- Shearmur, R. 2015: Dazzled by data: Big Data, the census and urban geography. *Urban Geography*, 36, 7, 965–968. DOI: 10.1080/02723638.2015.1050922
- Struijs, P. – Braaksma, B. – Daas, P. J. 2014: Official statistics and big data. *Big Data & Society*, 1, 1, 1–6. <https://doi.org/10.1177/2053951714538417>
- Trivedi, M. 2011: Regional and categorical patterns in consumer behavior: revealing trends. *Journal of Retailing*, 87, 1, 18–30.
- Virágh, E. 2012: Interjú. In *A statisztikák adják a viszonyítási keretet a piackutatáshoz*. @m, 4(3), 12–14. [https://issuu.com/atm\\_interaktiv/docs/\\_m\\_2012\\_maj\\_jun\\_ok](https://issuu.com/atm_interaktiv/docs/_m_2012_maj_jun_ok)
- Waddell, E. N. – Sacks, R. – Farley, S. M. – Johns, M. 2016: Point-of-sale tobacco marketing to youth in New York State. *Journal of Adolescent Health*, 59, 3, 365–367.