

Tartalomjegyzék

Összefoglaló.....	3
Adatgyűjtések és adatátvételek száma.....	5
Adatgyűjtések és adatátvételek megoszlása	7
Az OSAP-ban szereplő adatfelvételek megvalósulása a teljesülési kérdőívek alapján	11
Adatszolgáltatói terhek.....	12
Megvalósulás, az adatok hasznosulása.....	19
Adatvédelem.....	28
Felhasználók	29
Minőség értékelése és javítása	30
Megfelelő módszertani információk rendelkezésre állása és visszacsatolás az adatszolgáltatók/adatgazdák felé.....	34

A Központi Statisztikai Hivatal (KSH) 2014 tárgyév óta évente összefoglaló elemzést készít az Országos Statisztikai Adatgyűjtési Program (idéntől: Országos Statisztikai Adatfelvételi Program) (OSAP) teljesüléséről. A tájékoztató jelentősége, hogy nem csak a KSH statisztikai tevékenységének főbb adatait tartalmazza, hanem a Hivatalos Statisztikai Szolgálat (HSSz) egészének statisztikai feladatellátásáról igyekszik számot adni. A 2017. január 1-től hatályba lépett új statisztikai törvény¹ előírja az összefoglaló kötelező elkészítését és a Nemzeti Statisztikai Koordinációs Testület (NSKT), valamint az Országos Statisztikai Tanács (OST) véleményezését követően a nyilvános közzétételét.

Az információk forrása a KSH nyilvántartása az OSAP tartalmáról, továbbá az adatfelvételek² felelősei által kitöltött értékelő kérdőívek. A kérdőívek tartalma 2016 vonatkozási évre bővült, mert a korábbi tapasztalatok alapján javítottuk, így a 2016 vonatkozási időszakról már szélesebb áttekintést tudunk nyújtani, mint korábban. A Hivatalos Statisztikai Szolgálat működésének összehangolása az új statisztikai törvény értelmében a KSH kiemelt feladata, amelyhez a hasonló, összefogott képet nyújtó elemzésekkel is hozzá tudunk járulni, a rendszerszintű fejlesztések érdekében.

A KSH a 2016 tárgyévre vonatkozó elemzést a törvényi előírásnak megfelelően 2017 szeptember hónapban egyeztette az NSKT-val és az OST-vel.

Alábbi tájékoztatónk a beérkezett kérdőíves válaszok és a testületek észrevételei alapján alakult ki.

A Központi Statisztikai Hivatal az Országos Statisztikai Adatgyűjtési Programban, a 2016. évi vonatkozási időszakra összesen 272 adatgyűjtést és 102 adatátvételt (összefoglaló néven: adatfelvételt) nevesített. Elemzésünk az ezekről begyűjtött információkat tartalmazza.

Összefoglaló

A 2016 tárgyévre szóló Országos Statisztikai Adatgyűjtési Program a terveknek megfelelően valósult meg. Strukturálisan a Program az elmúlt években nem ment keresztül jelentős változásokon, azonban a felhasználói igények változásait egy-egy adatfelvétel megszüntetésével vagy új adatfelvételek indításával a Hivatalos Statisztikai Szolgálat igyekszik követni. Az elemzés nem tér ki rá külön, de az elmúlt évek során a legjelentősebb alapvető változás az OSAP adatfelvételeinek megvalósítási módjában következett be: a KSH és a HSSz további tagjai szinte kivétel nélkül az összes közvetlen adatgyűjtést elektronikus

¹ 2016. évi CLV. törvény a hivatalos statisztikáról, 31.§ (6) bekezdés

² Statisztikai adatfelvétel: egy adott sokaság adott időszakra vagy időpontra vonatkozó ismérveinek gyűjtése egy társadalmi, gazdasági vagy környezeti jelenség statisztikai megfigyelése céljából, különböző adatforrások felhasználásával. Ezen belül lehet adatgyűjtés: a statisztikai adatfelvétel által meghatározott sokaság adott időszakra vagy időpontra vonatkozó ismérveinek összegyűjtése a kiválasztott adatszolgáltatók megkérdezésével vagy közvetlen megfigyeléssel; adatátvétel: a statisztikai adatfelvétel által meghatározott sokaság adott időszakra vagy időpontra vonatkozó ismérveinek meghatározásához más szervezet által nyilvántartott vagy összegyűjtött adatok átvétele, felhasználása.

eszközzel vagy internetes kérdőívek formájában hajtják végre. 5 év alatt lényegében 100%-ban elektronikussá vált a hivatalos statisztikai adatgyűjtés minden csatornája.

Az adatszolgáltatói terhek vizsgálata tekintetében még jelentős fejlesztésekre van szükség. Az OSAP változásai – az adatfelvételek számát tekintve minimális növekedést látunk – nem feltétlenül tudják jól igazolni azt a törekvésünket, hogy az adatszolgáltatók terhei is folyamatosan csökkenjenek, bár a statisztika, mint közfeladat társadalmi terhelése objektíven mérve vélhetően kisebb, mint azt az egyedi adatszolgáltató szintjén jelentkező teherként a gazdaság és a társadalom szereplői érzik és visszajelzik azt pl. a médiában. Ezt a jelenséget évek óta nagy nehézségekkel, kevés eredménnyel kezeli a Hivatalos Statisztikai Szolgálat.

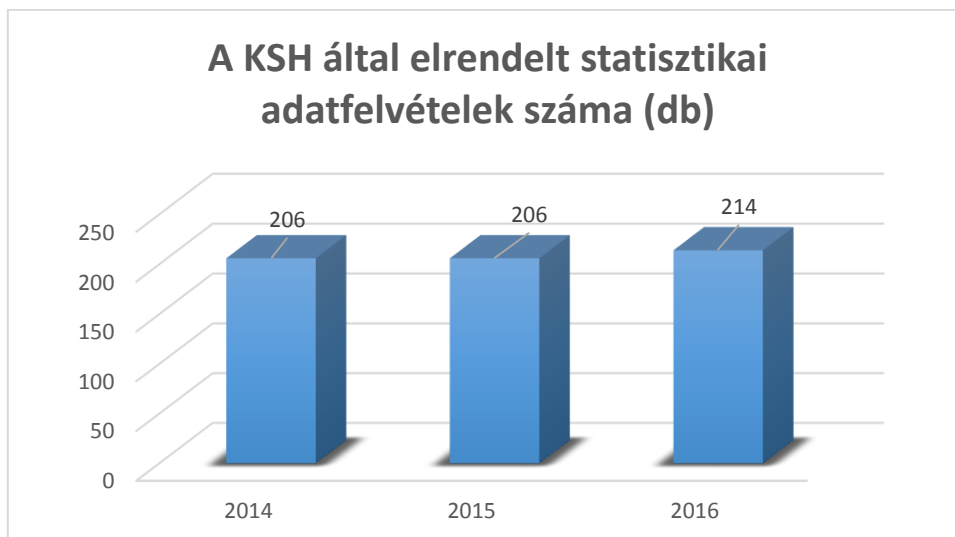
A minőségjavítás érdekében tett lépések tekintetében 2015 után az adatok jelentős javulást mutatnak a kérdőív kitöltőinek saját megítélése alapján, különösen a statisztikák módszertanának dokumentálása terén. A KSH, mint szakmai koordináló szervezet egyik legfőbb célja, hogy a HSSz minden tagja megfelelően dokumentálja hivatalos statisztikai tevékenységét, így az ezen a téren tapasztalt fejlődés elismerésre méltó, ám az részletes vizsgálatokat igényel, hogy ezek a dokumentációk minőségükben, részletességükben, tartalmi elemeikben megfelelnek-e a hazai és európai statisztika minőségi elvárásainak. Ugyancsak fejlődő képet mutat a kérdőíves visszajelzések szerint az adatvédelem érdekében tett intézkedések általánossá válása a KSH-n kívüli HSSz-tagok esetében is. Az adatszolgáltatók által biztosított információk a statisztikai célú felhasználás körében a hatályos adatvédelmi előírásoknak megfelelően kerülnek kezelésre, a HSSz tagjai folyamatosan fejlesztik informatikai és statisztikai eljárásaikat is az adatvédelem terén.

Az adatfelvételek eredményeit a HSSz tagjai ma már jellemzően, elsősorban interneten publikálják, azonban számos esetben kevés nyilvánosságot kap, nehezebben található meg az általuk előállított információ. Míg a KSH aktívan és rendszeresen fejleszti internetes kommunikációját, a HSSz további tagjainál kifejezetten a statisztikáik kommunikációjára kevés fejlesztési forrás jut. 2017-ben ezért egyik célunk a koordinációs tevékenység keretében az, hogy közös megjelenési felületet, nagyobb nyilvánosságot biztosítsunk az interneten a Hivatalos Statisztikai Szolgálat tagjainak statisztikai tevékenységéhez, ezzel is segítve az általuk előállított információk könnyebb elérését és a felhasználói párbeszédet. Ennek érdekében a KSH 2017 őszén elindítja a Hivatalos Statisztikai Szolgálat weboldalát, amelyen a HSSz tagjainak statisztikai tevékenysége és az általuk publikált adatok elérhetősége összefoglalásra kerül.

Az elemzés az adatfelvételek felelősei által kitöltött, részben önértékelő kérdéseket tartalmazó kérdőív válaszain alapul, amelyet a 374-ből 367 adatfelvételtől sikerült összegyűjtenünk. A kérdőívek esetleges hiányos kitöltését, vélhető hibáit a kitöltőkkel egyeztetjük. A 367 visszaküldött kérdőívből 360 esetben kaptunk teljes körűnek tekinthető információt, hét esetben néhány kérdés megválaszolatlanul maradt. A felhasználók igényeinek elemzése és az előállított statisztikák hasznosulásának vizsgálata további fejlesztéseket igényel a statisztikák előállítóinak oldalán. Összeállításunk célja a felhasználók egyre fejlettebb tájékoztatása a Hivatalos Statisztikai Szolgálat tevékenységéről. Ezért az elemzésünk bővítése érdekében további konzultációkat folytatunk a HSSz tagjaival és javaslataikat a következő évben készülő tájékoztatónkba beépítjük.

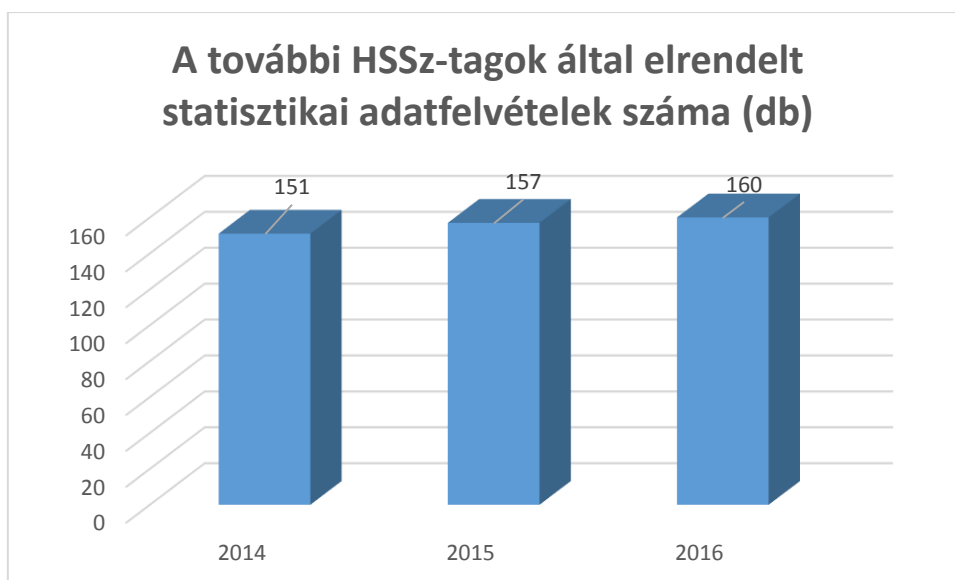
Adatgyűjtések és adatátvételek száma

Az adatgyűjtésekről illetve adatátvételekről (összefoglaló néven: adatfelvételekről) szóló teljesülési kérdőívet a Központi Statisztikai Hivatal adatfelvételekért felelős dolgozói illetve a HSSz tagjai saját adatfelvételeikről töltötték ki. A 2016 tárgyévi kérdőív a megelőzőekhez képest részletesebb, így nagyobb betekintést kaphatunk az adatfelvételek legfontosabb jellemzőibe. A kérdőívekre adott válaszok elemzése előtt, ebben a fejezetben az elrendelt OSAP alapinformációt mutatjuk be.



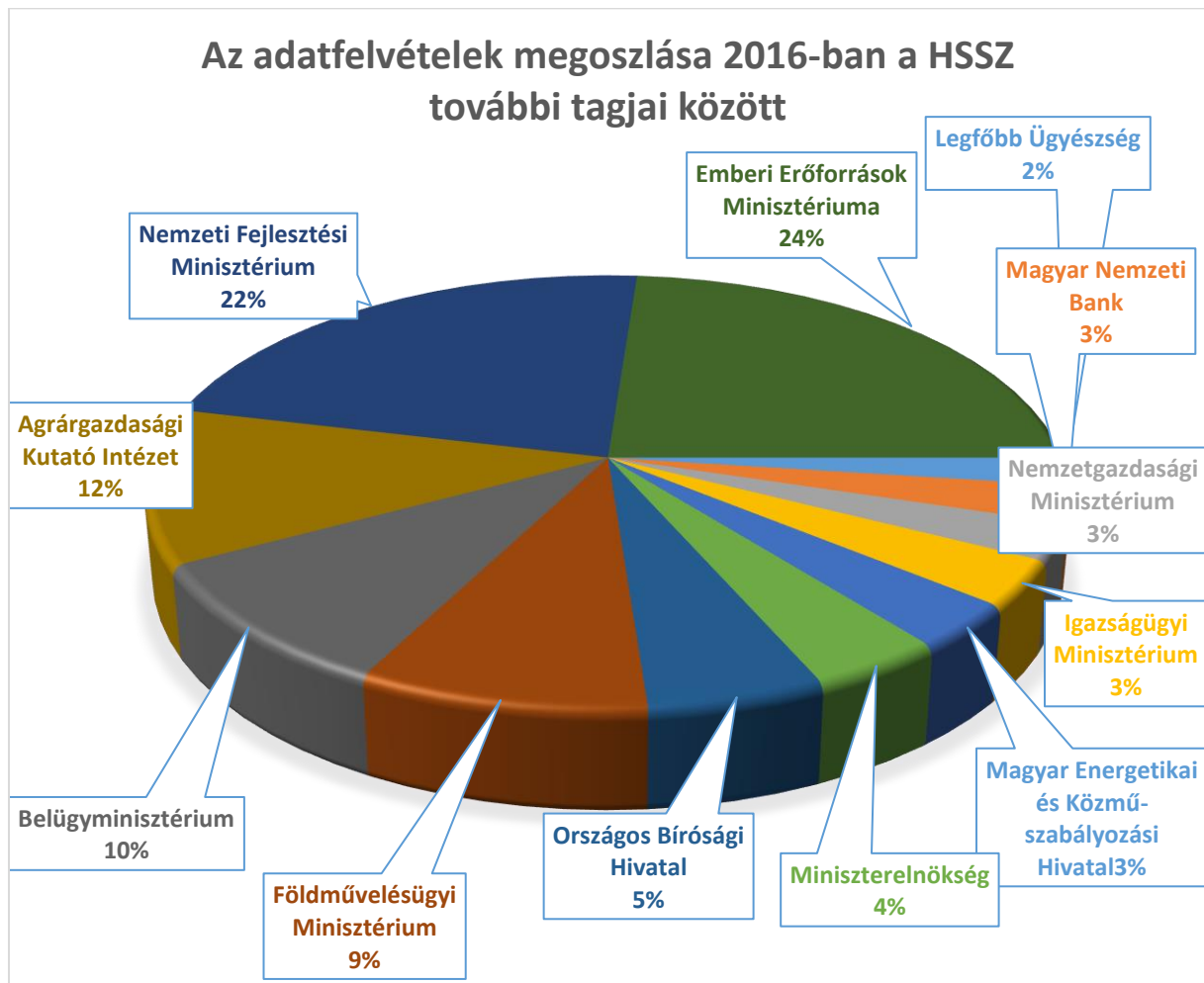
1. sz. ábra

A 2016 vonatkozó évre elrendelt programban az adatfelvételek során a 374-ből 214 adatfelvétel a KSH fejezetében szerepelt.



2. sz. ábra

A Hivatalos Statisztikai Szolgálat KSH-n kívüli tagjai esetében sincs lényeges változás az adatfelvételek számában az elmúlt évekhez képest (lásd 2. sz. ábra). A felvételek számának csekély változása önmagában nem jelez problémát, mivel a minimális változások mögött a felhasználói igényeknek megfelelően megszűnő vagy újonnan elinduló adatfelvételek állnak. A HSSz tagok adatfelvételeinek száma 2014 óta folyamatosan kismértékben növekszik.



3. sz. ábra

A 2016. vonatkozási évre az Emberi Erőforrások Minisztériumához (EMMI) tartozott a legtöbb adatfelvétel, a HSSz-tagok által végrehajtott adatfelvételek mintegy negyede, s ezt követi a Nemzeti Fejlesztési Minisztérium (NFM) az adatfelvételek 22%-ával (lásd 3. sz. ábra).

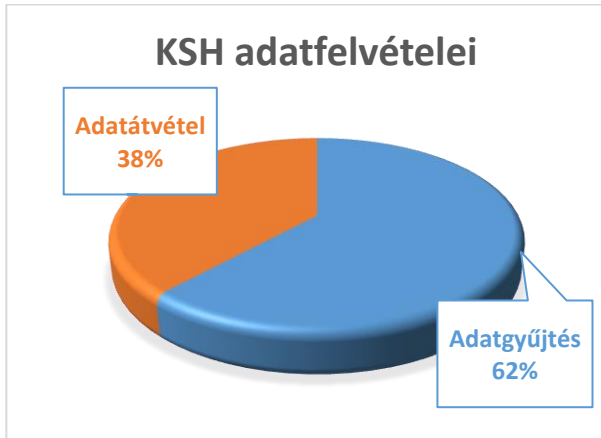
Az Emberi Erőforrások Minisztériumának fejezete az OSAP-ban 7 db adatátvételtől és 28 db adatgyűjtésből állt, a Nemzeti Fejlesztési Minisztérium fejezete 27 db adatgyűjtésből és 2 db adatátvételtől tevődött össze.

Az Emberi Erőforrások Minisztériuma adatfelvételeinek a főbb témái oktatással kapcsolatos ügyeket vizsgálnak, valamint kulturális tevékenységekről, illetve az egészségügyi ágazat jelentéseiről szólnak.

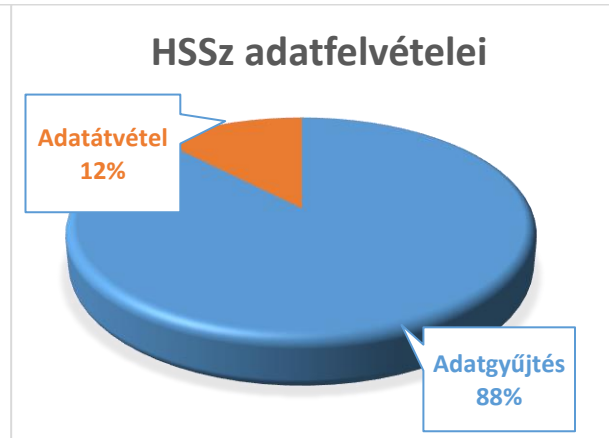
Az NFM statisztikai adatfelvételeinek nagy része a légi, vízi, vasúti illetve közúti közlekedési adatok és az azokkal kapcsolatos tárgyi eszközök, műszaki berendezések, engedélyek beszerzéséről illetve az ezekhez kapcsolódó forgalmi adatokról szólnak. Ezen felül postai hírközlési témában is készítenek adatfelvételeket. Az NFM adatgyűjtéseit a KSH hajtotta végre. Az adatfelvételekben begyűjtött adathelyek mennyiségének megoszlása hasonló sorrendet eredményez a HSSZ tagjai között, azzal, hogy az EMMI a nagy ellátórendszerekből felhasznált adminisztratív adatok mennyisége miatt ebben a megközelítésben is első helyre kerül, de az NFM-et a harmadik és negyedik helyen az Országos Bírósági Hivatal és a Magyar Nemzeti Bank követi.

Adatgyűjtések és adatátvételek megoszlása

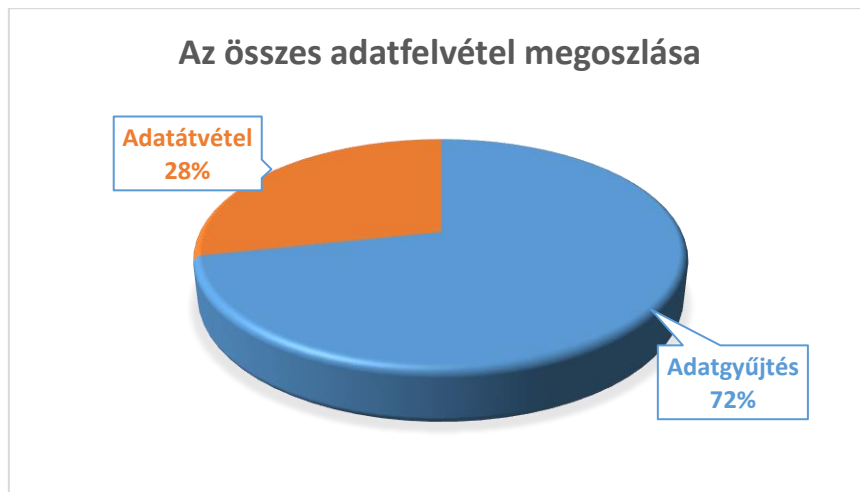
Mind a közvetlen adatszolgáltatói körben végrehajtott adatgyűjtések, mind az adminisztratív adatátvételek hangsúlyosak a Központi Statisztikai Hivatal tevékenysége során. Az összes adatfelvételből 38% adatátvétel és 62% adatgyűjtés. Az adatátvételek száma kis mértékben növekszik, az adatgyűjtéseké csökken – ez a tendencia összhangban van az adminisztratív terhek csökkentésével a közigazgatásban (lásd 4-6. sz. ábrák).



4. sz. ábra

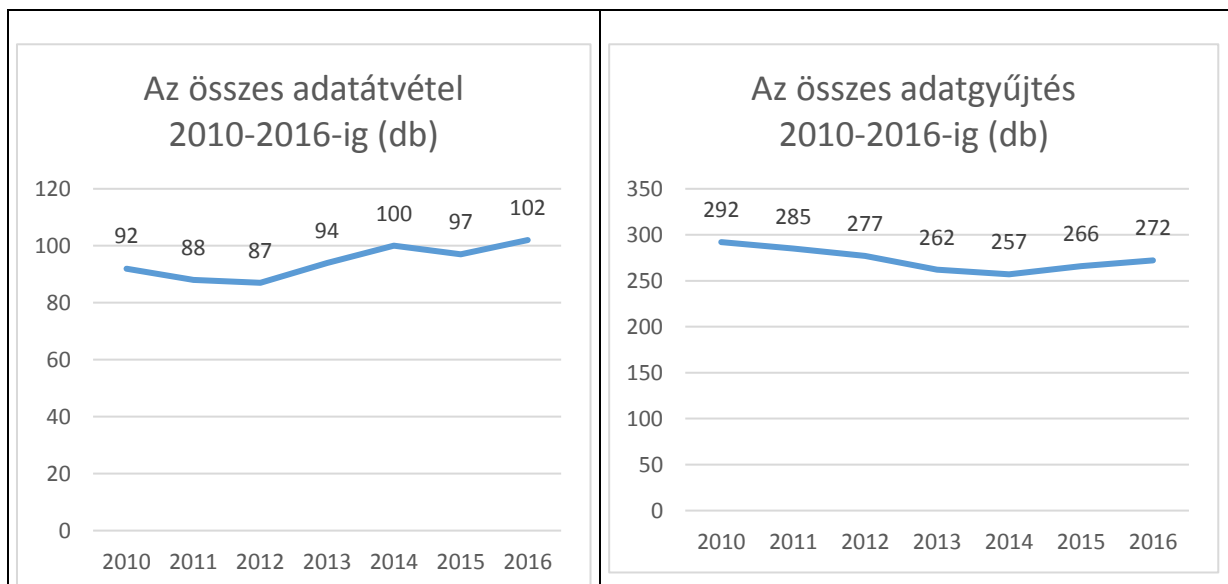


5.sz. ábra



6. sz. ábra

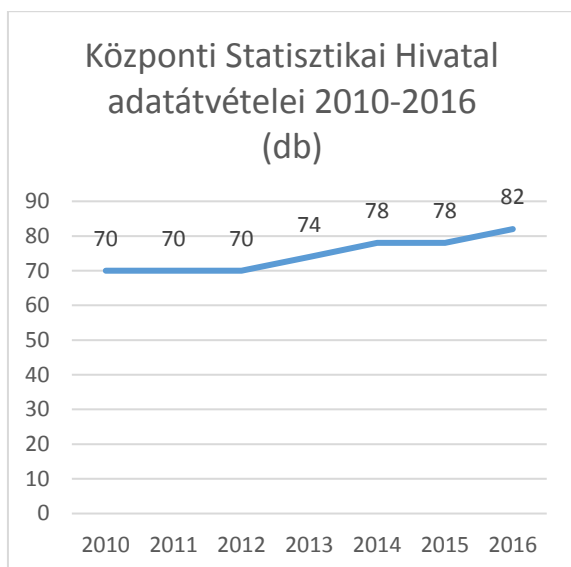
A HSSz további tagjai esetében jóval hangsúlyosabbak az adatgyűjtések, mint az adatátvételek. A HSSz tagjai esetében az összes adatfelvételből 88%-ot tesz ki az adatgyűjtések száma. A szervezeti alapon tekintve két igen eltérő összetételű rész átlagaként 2016 tárgyévben a teljes Országos Statisztikai Adatgyűjtési Programban az adatfelvételek 28%-a adatátvétel, és 72%-a adatgyűjtés. Ez az arány az OSAP-ban elrendelt adatfelvételek számának kisebb (384-ről 374-re történő) csökkenése mellett nehezen, lassan, három százalékponttal mozdult el az adatátvételek hangsúlyosabb irányába, ugyanis 2010-ben az adatgyűjtések és adatátvételek aránya 76% (292/384) és 24% (92/384) volt.



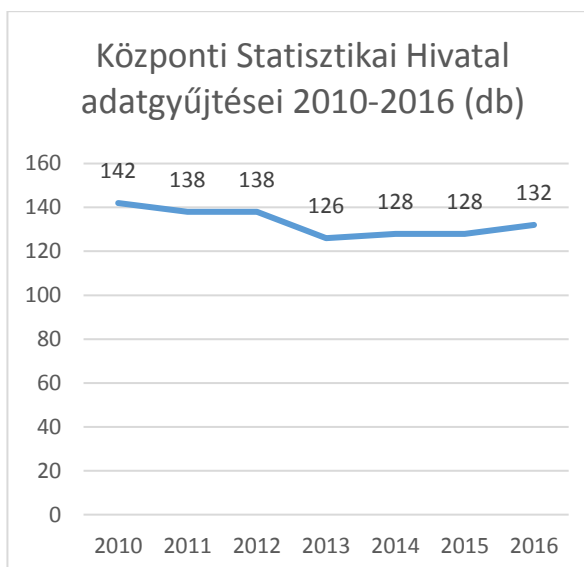
7. sz. ábra

8. sz. ábra

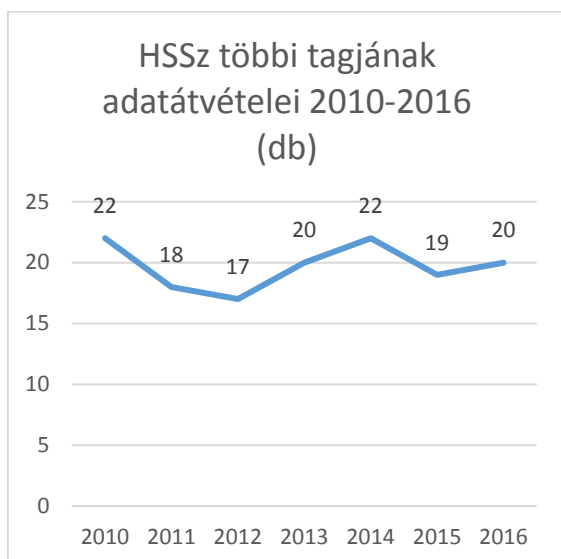
2010 és 2014 között az adatgyűjtések tekintetében egy lassú csökkenést figyelhetünk meg az összesítő diagramokon és a HSSz diagramjain egyaránt (7-12. sz. ábra). 2014 óta viszont enyhe növekedést tapasztalhatunk. A KSH részéről az adatgyűjtések száma összességében kevesebb volt ugyan 2010-hez képest, de a legalacsonyabb számot mutató 2013-hoz viszonyítva kis mértékben nő az adatgyűjtések száma. Az adatgyűjtések csökkenésével párhuzamosan az adatátvételek száma növekszik. Mind az adatgyűjtések, mind pedig az adatátvételek száma tekintetében minimális változás zajlott le 2010-től 2016-ig.



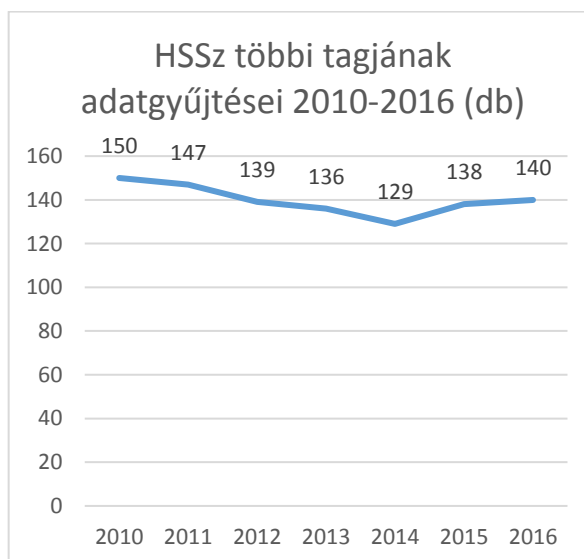
9. sz. ábra



10. sz. ábra

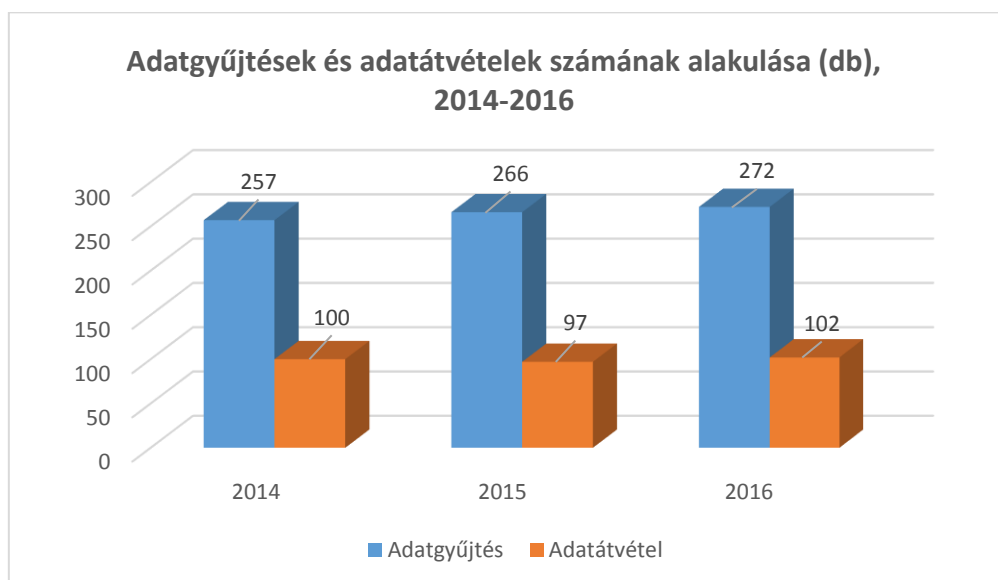


11. sz. ábra



12. sz. ábra

Az utóbbi három év összevetésben hasonlóképp megállapíthatjuk, hogy csak kismértékű változások következtek be az adatfelvételek típus szerinti összetételében. Az adatgyűjtések száma 2014-ben 257 volt, ez 2015-re 266-ra nőtt, 2016-ra pedig ez az érték tovább növekedett 272-re. Az adatátvételek tekintetében megfigyelhető, hogy 2014-ben 100 db-ot számláltak, 2015-ben 97-et és most pedig 102-t, ami csekély növekedés.



13. sz. ábra

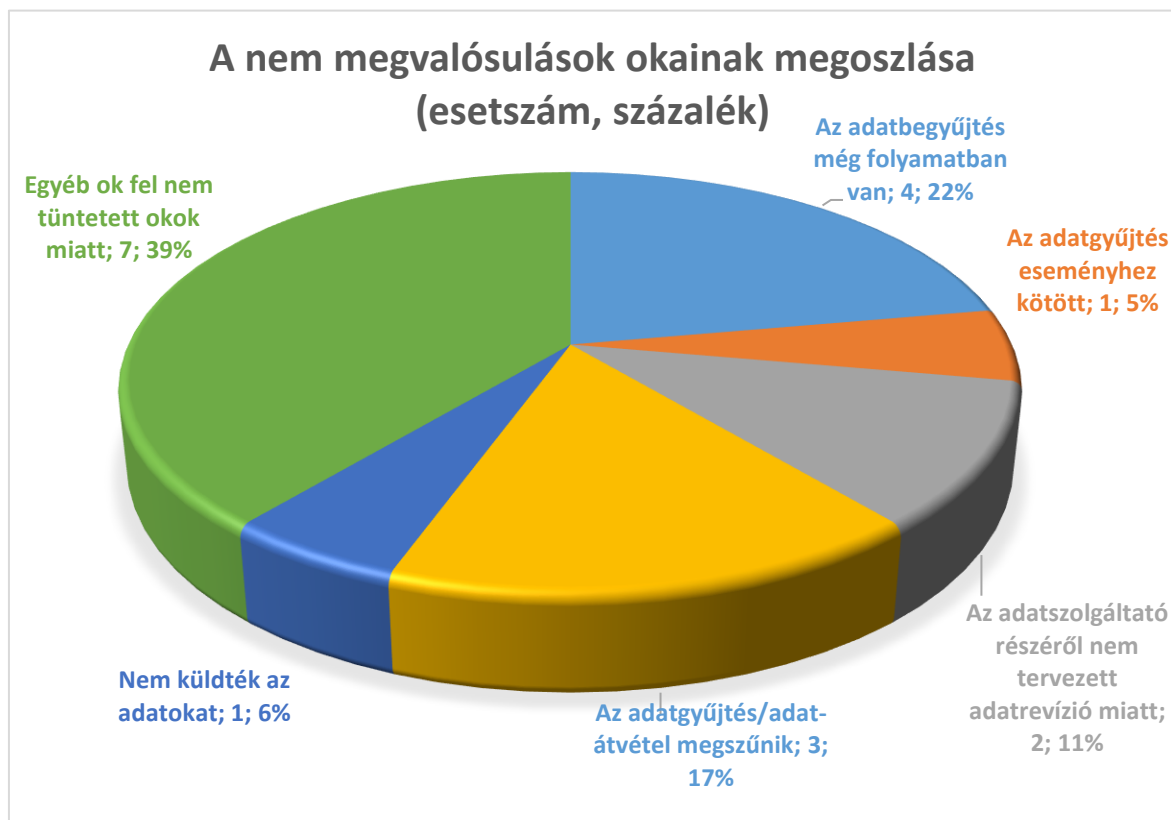
A minimális számszerű változás mögött strukturálisan azonban jelentősebb különbségek lehetnek, főképp az egyes felvételek adatszolgáltatói terhekre gyakorolt változásai terén, azonban ezekre jelen vizsgálat keretében még nem volt lehetőségünk részletesen kitérni. Az elmúlt években a kismértékű csökkenés elsődleges oka, hogy adatgyűjtéseket váltottak ki adatátvételekkel, de előfordult tévesen kategorizált elem is az OSAP-ban, így az átsorolások esetenként nem jártak az adatfelvételi program érdemi módosulásával.

Ezekre a kérdésekre a jövőben az adatszolgáltatói terhek részletesebb elemzési lehetőségének megteremtésével kell visszatérni.

A 2016. vonatkozó év 374 adatfelvételéből idén 367 adatgyűjtésről és adatátvételről érkezett vissza a teljesülési kérdőív, ami 98,12%-os beérkezési arányt jelent. A továbbiakban összefoglalónk következő fejezeteiben az adatfelvételekről beérkezett adatok megoszlását mutatjuk be, különböző szempontok szerint.

Az OSAP-ban szereplő adatfelvételek megvalósulása a teljesülési kérdőívek alapján

A beérkezett 367 kérdőívből megállapítható, hogy az azok által lefedett körben 95%-uk (349/367) esetén az adatbegyűjtés folyamata a 2016-os megfigyelési időszakról lezárult 2017. júliusra. Abban a 18 esetben, ahol ez nem történt meg, többnyire kétféle okcsoportot fedezhettünk fel. Vagy adatátadási problémák fordultak elő illetve technikai rendszerfejlesztés zajlott, vagy az adatgyűjtés természete többévenkénti vagy eseményhez kötődő volt, illetve beérkezési határideje a megvalósulási adatok összegyűjtési időpontja utáni. Elszigetelten akadt olyan adatfelvétel, mely valóban nem került végrehajtásra, azonban a be nem fejeződött adatbegyűjtési folyamatok egyharmadánál nem kaptunk érdemi indokolást a kérdőíven.



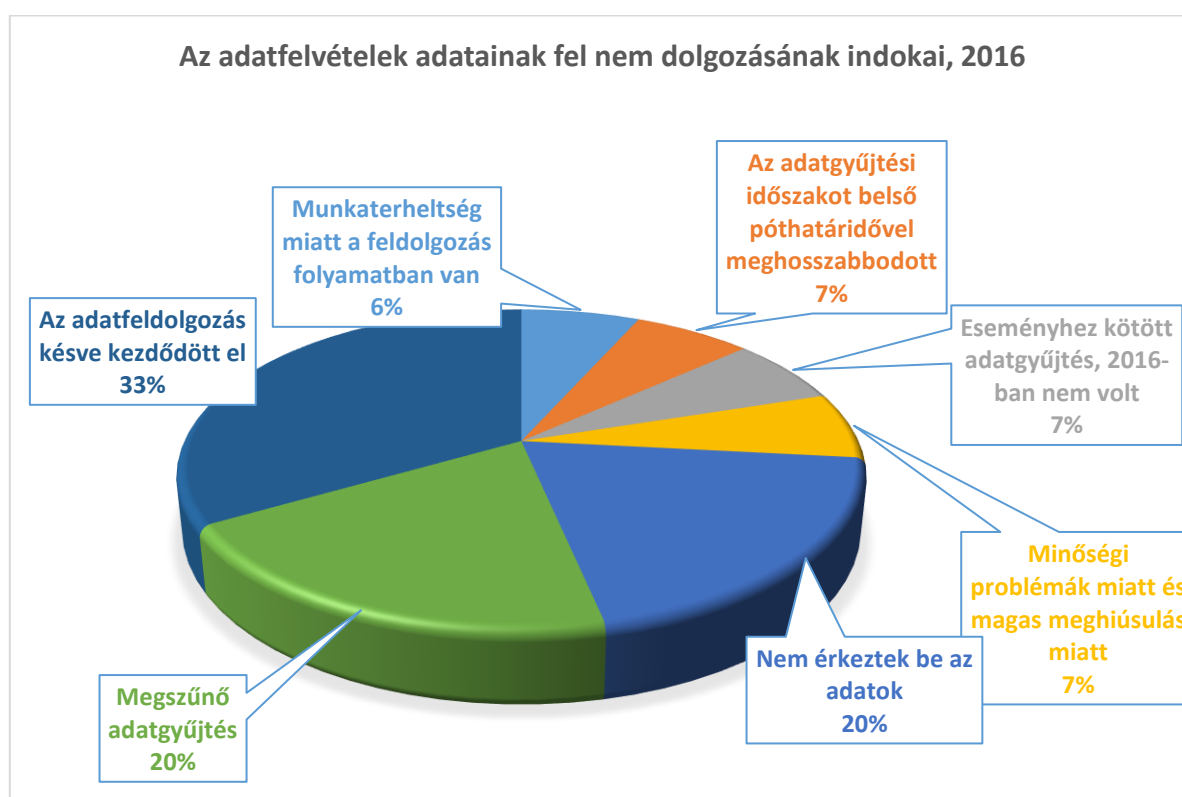
14. sz. ábra

Az előző évhez képest tehát a tárgyévét követő júliusban még be nem fejeződött adatgyűjtési folyamatok száma hasonlóan alakult, hiszen tavaly ez a szám 12 darab volt.

A begyűjtött/átvett adatok feldolgozása szintén nagyon előrehaladott állapotban volt az értékelő kérdőív kitöltésekor. A feldolgozottságról szóló kérdést is pontosítani szükséges a kérdőíven a visszajelzések alapján, mivel esetenként előfordul, hogy a kérdőív kitöltése időpontjában az előzetes adatok közzétételéhez szükséges feldolgozás már megtörténik, de több esetben a végleges adatok közzétételéhez a feldolgozás az értékelő kérdőív összegyűjtését követően esedékes. 2017 augusztusában a beérkezett kérdőívek által lefedett 367-ből 324 adatfelvétel esetében megtörtént az adatok feldolgozása is. A továbbiak szinte mindegyike

esetén a feldolgozási folyamat a kérdőív kitöltése idején is zajlott, általában ezeknél a tervezett feldolgozási határidők is későbbre esnek, mint az értékelő kérdőív szokásos beérkezési ideje. Időszerűséget tekintve tehát kijelenthető, hogy a 2016 vonatkozású időszak adatainak feldolgozása a t+8. hónapra döntő többségében megvalósult.

A vizsgálat idején még fel nem dolgozott 43 adatfelvétel esetében 65%-uknál, azaz 28 felvétel esetében az eredetileg kitűzött befejezési határidő általában szeptemberben, októberben jár le. Ilyen feldolgozási határidőkkel találkozhatunk a Belügyminisztériumnál, a Magyar Energetikai és Közmű-szabályozási Hivatalnál, az Emberi Erőforrások Minisztériumánál és esetenként magánál a Központi Statisztikai Hivatalnál. A további okokról a következő kördiagram szól (15. számú ábra). A megszűnéseken túl a szervezetek munkaterheltségéből eredő késedelme a fő oka annak, hogy késik a feldolgozás lépése a folyamatban.



15. sz. ábra

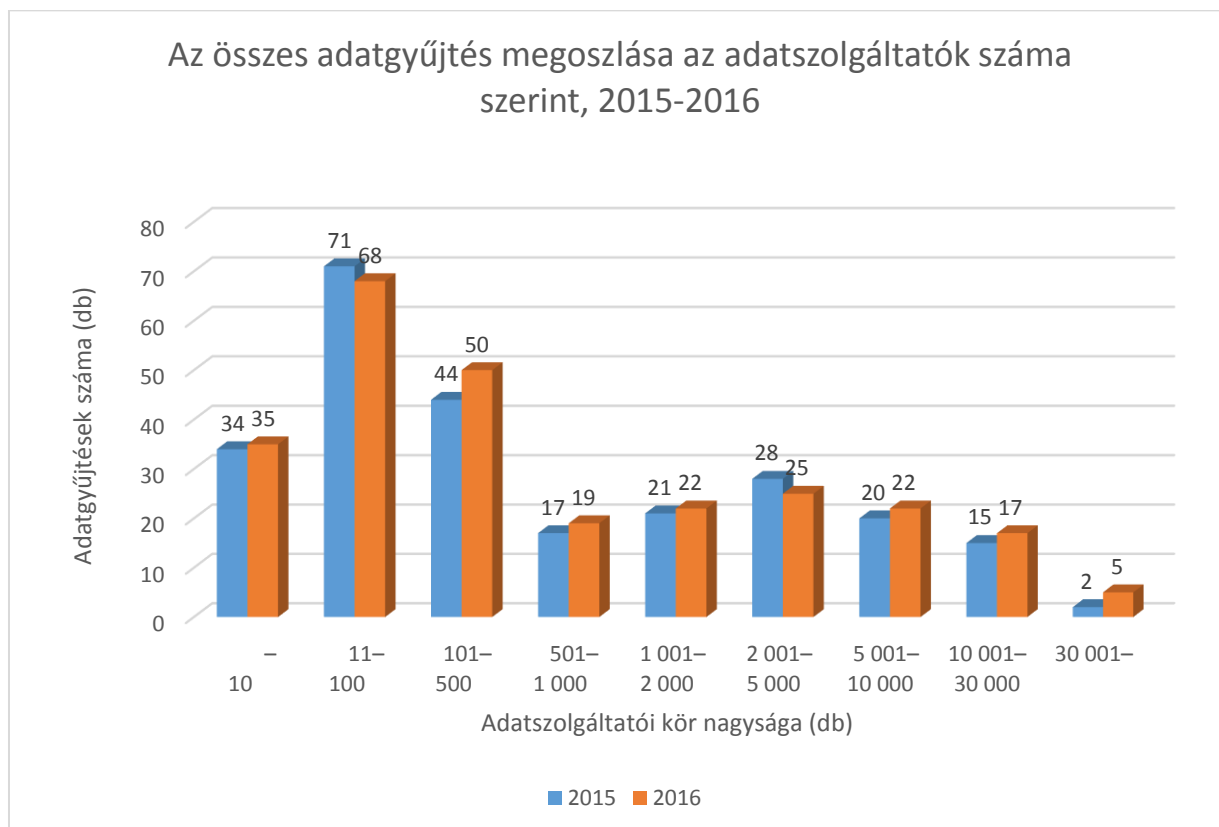
Adatszolgáltatói terhek

Az adatszolgáltatói terhek mérésére, azok megítélésére az értékelő kérdőíven két kérdés szolgált. Mind az adatszolgáltatók tényleges számát, mind pedig az átlagosan kitöltött adathelyek számát célszerű külön vizsgálni adatgyűjtések és adatátvételek körében, mivel a két típus között ezekből a szempontokból is jelentős a különbség.

Az adatszolgáltatói terhek mérésével 2014 óta foglalkozik a teljesülési kérdőív, ugyanakkor 2014-ben még nem érkeztek érdemben értékelhető válaszok, s azok alapján nem születtek

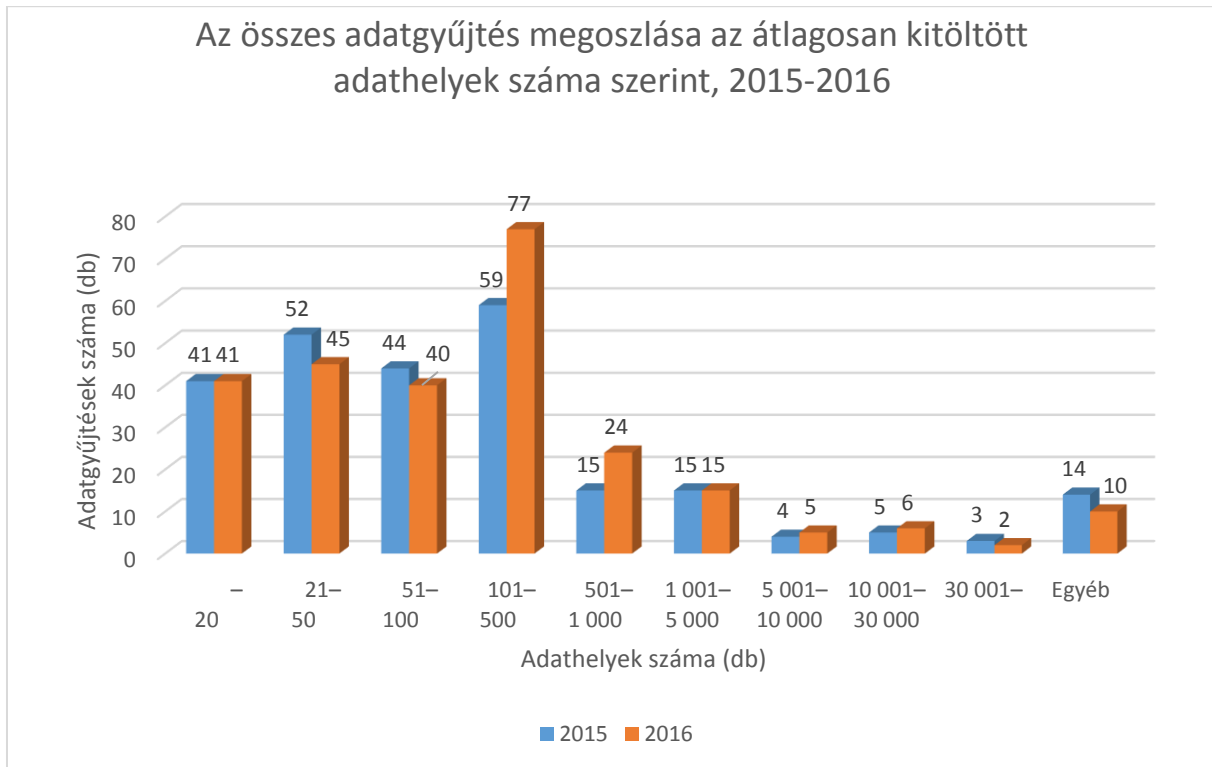
érdemleges adatok. Ehhez képest a 2015-ös és 2016-os vonatkozó évre szóló adatokból már ki tudtunk nyerni információkat és a kérdőív módosításával folytatjuk ezt a törekvést. Nem kaptunk ugyanakkor minden beérkezett kérdőíven információt az adatszolgáltatói terhekről, ezért az ábrázolt kategóriák összesen értéke néhány darabbal alacsonyabb.

A téma jelentőségét jelzi, hogy a KSH 2017-től az adatszolgáltatói terhek csökkentése, mérhetővé tételének javítása érdekében külön fejlesztési intézkedéseket indít.



16. sz. ábra

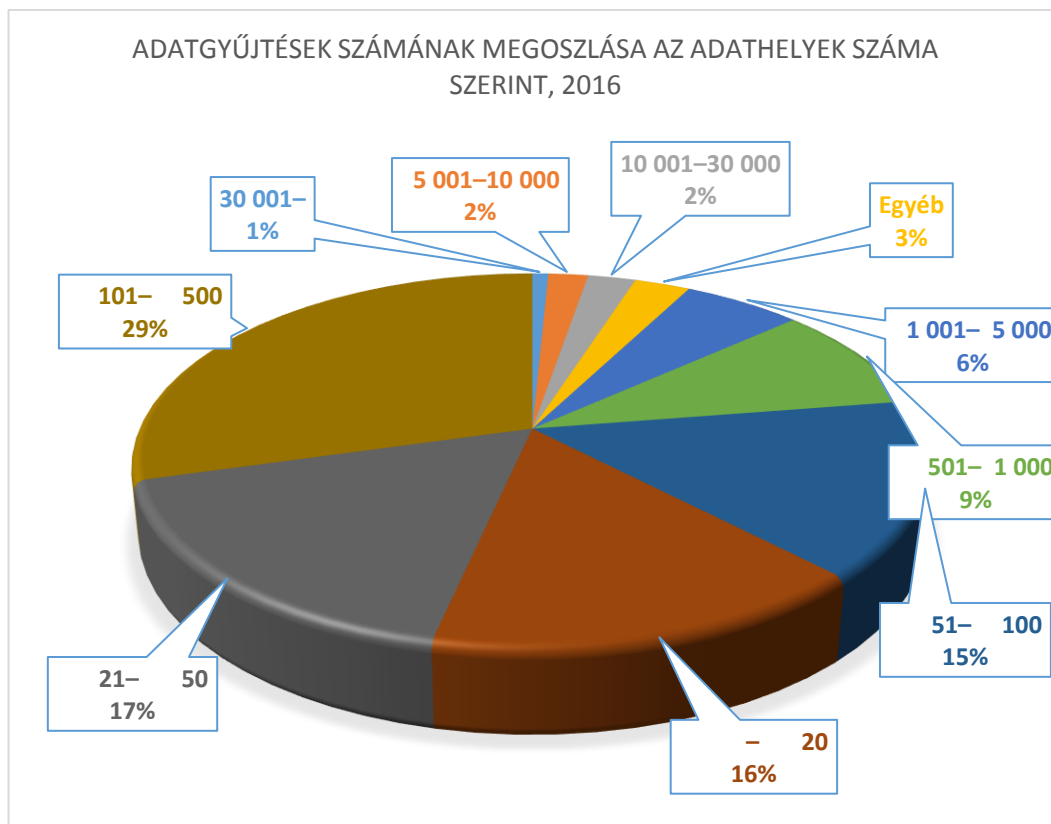
Az adatgyűjtések megoszlása az adatszolgáltatók száma szerint csoportosítva hasonlóan alakult a 2016-os évben, mint 2015-ben (16. számú ábra). Sok adatgyűjtés viszonylag kis adatszolgáltatói körrel dolgozik. 2015-ben az adatgyűjtések 41,66%-ára volt legfeljebb 100 vagy annál kevesebb adatszolgáltató kijelölve, ám 2016-ban ez minimálisan, 38,86%-ra csökkent, miközben a 101-2000 közötti adatszolgáltatót érintő felvételek száma hasonlóan minimális mértékben nőtt.



17. sz. ábra

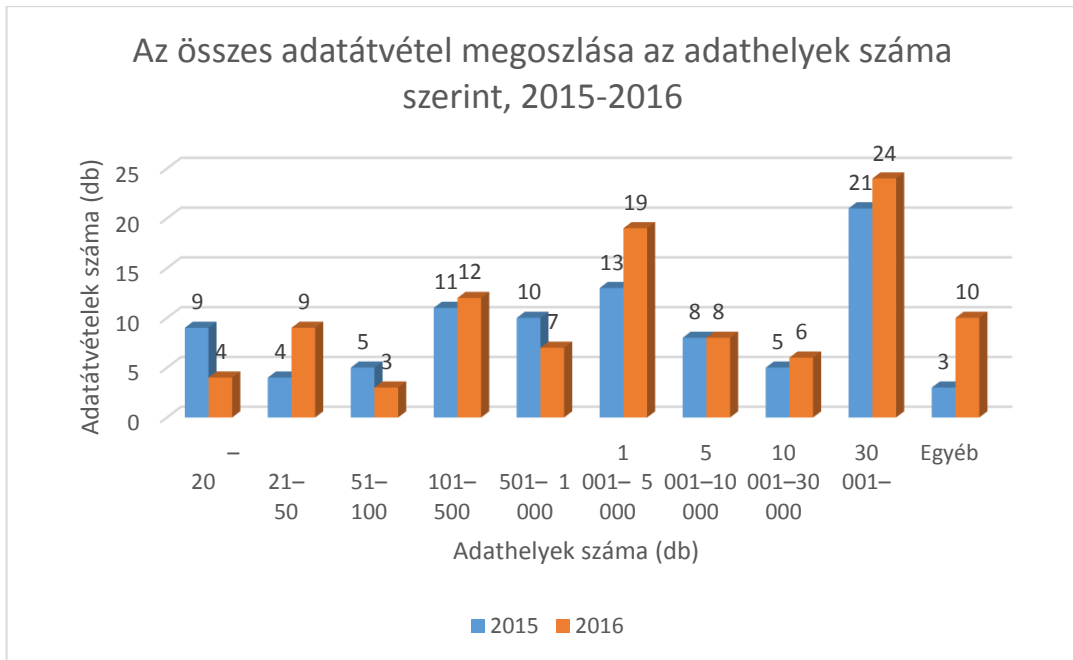
126 db olyan adatgyűjtés van, amely maximum 100 adathellyel rendelkezik. Ez gyakorlatilag az adatgyűjtések majdnem felét teszi ki és csökkenést mutat ezek száma 2015-höz képest (lásd 17. sz. ábra). Mindeközben a 101-500 illetve 501-1000 közötti adathelyekkel rendelkező adatgyűjtések száma némileg növekedett az előző évhez képest. Ez azt mutatja, hogy az adatszolgáltatói terhek az adatgyűjtések kérdőívének terjedelmi szempontjából vizsgálva összességükben makroszinten lényegében nem változhattak. Ezen felül 28 olyan adatgyűjtés volt 2016-ban, amelyben az átlagos adathelyek száma elérte a 1001-et, ezek többnyire forgalmi adatok, köznevelési, oktatási, hatósági, ügyforgalmi illetve pénzügyi témákban készültek.

A 2016-os adatok megoszlását jobban szemlélteti az alábbi tortadiagram:

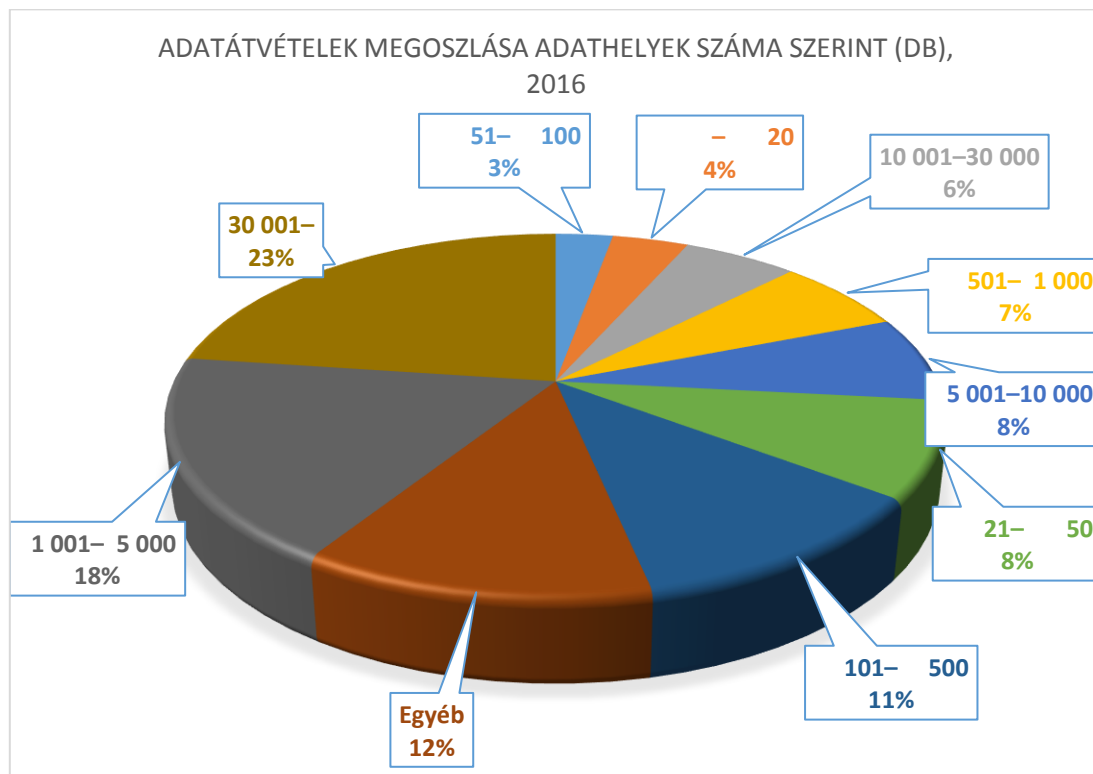


18. sz. ábra

Az adatátvételekre vonatkozóan, az egy adatszolgáltatóra jutó átlagos adathelyek számát vizsgálva az látszik, hogy az adatgyűjtésekhez képest itt nem a kisszámú adathelyek a tipikusak. Az adatátvételek esetében az átvett adathelyek számát kaptuk meg a kitöltőktől. A legtöbb adatátvételnél a 30001 feletti (23%) adathely-szám a tipikus mind 2015-ben, mind 2016-ban. Ez azért lehetséges, mert ezek nem egyedi adat-összeállítási terhet jelentenek az adatgazdáknak. Az adatátvételeknél egy-egy szabályozó hatóság több adatot gyűjt adminisztratív célból, és ezek nagytömegű átadása informatikai eszközökkel jellemzően nem jelent kiugró terhelést.



19. sz. ábra



20. sz. ábra

A 2016-os adatszolgáltatói terhekről komplexebb képet akkor kapunk, ha az adatszolgáltatók számát és az átlagos adathelyeket (adatgyűjtés esetén kitöltött, adatátvétel esetén az átadott állományban szereplő mennyiségben) egy-egy keresztábrába foglalva szemléltetjük.

Az adatgyűjtések számának (db) megoszlása a statisztikai adathelyek száma és a tényleges adatszolgáltatók száma szerint

	Statisztikai adathelyek száma										
		– 10	11– 100	101– 500	501– 1000	1001–2 000	2 001– 5000	5001–1 0000	10001– 30000	30001–	Összesen
Tényleges adatszolgáltatók száma	– 10	6	13	9	3	2	1	3	2	0	39
	11– 100	2	27	20	7	6	1	0	3	2	68
	101– 500	5	18	16	8	0	0	1	0	0	48
	501– 1000	0	8	8	1	1	1	0	0	0	19
	1001– 2000	4	7	10	0	0	1	0	0	0	22
	2001– 5000	3	10	8	2	2	0	0	1	0	26
	5001–10000	1	12	4	3	0	0	1	0	0	21
	10001–30000	5	9	3	0	0	0	0	0	0	17
	30001–	1	3	1	0	0	0	0	0	0	5
	Összesen	27	107	79	24	11	4	5	6	2	265

1. sz. tábla

Az adatgyűjtések fele általában 500 vagy annál kevesebb adathellyel és legfeljebb 1000 adatszolgáltatóval dolgozik (zöld mezővel jelöltük a tábla celláit), illetve a zöld területbe tartozónak jelöltük még azokat az adatgyűjtéseket, amelyek viszonylag több adatszolgáltatóval dolgoznak kevés adathellyel, valamint azokat is, amelyek kevés adatszolgáltatót kérdeznek, de több adathellyel dolgoznak (lásd 1. sz. tábla). Habár a terhelés szempontjából egyértelmű határvonalak nem állapíthatók meg az alacsony - közepes - magas szintek között (ezt a színátmenetekkel is igyekeztünk ábrázolni), kijelenthető, hogy az adatgyűjtések fele egyértelműen a két szempont szerinti mátrix alacsony terhelést jelentő bal felső negyedében

van. Érdemes kiemelni még a 10 alatti, speciális adatszolgáltatói számúakat, ilyen adatgyűjtés például a repülőterek forgalmi adatai, illetve hajózási forgalommal foglalkozó adatgyűjtések.

Tehát komplexen értelmezve az adatszolgáltatói terhet, pusztán e két paraméter vizsgálatával az adatgyűjtések fele esetében relatíve alacsony statisztikai terhelést látunk. A citromsárgával jelölt területen a HSSz nagyobb horderejű adatgyűjtéseit láthatjuk, ilyen például Gazdaságszerkezeti Összeírás, 2016. (kiemelt egyéni gazdaságok) vagy a Munkaerőköltség-felvétel. Ezek jellemzően nem minden évben végrehajtott közvetlen adatgyűjtések. Látható az is, hogy az igazán sok adatszolgáltatót érintő (több, mint 10 ezer adatszolgáltatóval működő) adatgyűjtések esetében az adathelyek száma jellemzően nem magas, 11 és 100 adathely közötti, ilyen adatgyűjtés például „A lakosság utazásai külföldön”, a gazdálkodó szervezetek körében végrehajtott Éves beruházás statisztikai jelentés vagy az Egyéni gazdaságok decemberi mezőgazdasági összeírása. Összességében megállapítható, hogy elenyésző az olyan adatgyűjtések száma az OSAP-ban, amelyek egyszerre sok adatszolgáltatót érintenek és nagy mennyiségű az adatigényük is. További vizsgálat tárgyává lehet még tenni a jövőben az egy adatszolgáltatót párhuzamosan terhelő, különböző részletességű adatgyűjtések dimenzióját.

Az adatátvételek számának megoszlása az adatállományok mezőinek száma és a ténylegesen nyilvántartott egységek száma szerint

	Adatállomány mezőinek száma									
	– 10	11– 100	101– 500	501– 1000	1001–2 000	2001– 5000	5001–1 0000	10001– 30000	30001–	Összesen
– 10	9	5	2	4	2	3	2	4	4	35
11– 100	0	0	2	1	1	1	0	0	3	8
101– 500	0	0	1	1	0	6	2	0	1	11
501– 1000	1	1	1	0	0	0	2	0	0	5
1 001– 2 000	1	0	0	0	0	1	0	0	4	6
2 001– 5 000	0	3	1	0	1	1	1	0	1	8
5 001–10000	1	0	0	1	0	0	0	1	1	4
10 001–30000	1	0	0	0	0	0	1	1	2	5
30 001–	2	4	3	0	3	0	0	0	8	20
Összesen	15	13	10	7	7	12	8	6	24	102

2. sz. tábla

Az adatgyűjtésekkel ellentétben, az adatátvételek körében természetes módon a ténylegesen nyilvántartott egységek³ száma és az adatállományok mezőinek száma esetében jelentősen gyakoribb a mindkét dimenzió szerint nagyobb állományok átadása (lásd 2. sz. tábla).

A zöld mezőben olyan adatátvételeket találunk, mint a Levegőtisztaság-védelmi adatok vagy az országos hálózaton történő vasúti szállításhoz szükséges engedéllyel rendelkező vállalkozások listája. A sárga sávban oktatással kapcsolatos adatátvételek találhatunk, ilyen például a jelentés a tanév első évfolyamára jelentkezzék és felvett pályázók állománya. És a piros jellemzően olyan állományokat tartalmaz, amelyek a nyugdíjban, ellátásban, járadékban és egyéb járandóságban részesülők adatait tartalmazzák, azaz a nagy ellátórendszerek nyilvántartásaiból statisztikai célra átvett adatbázisokat.

Megvalósulás, az adatok hasznosulása

Fontos része az adatelőállítási folyamatnak a felhasználók azonosítása, a publikációk, majd a továbbfelhasználás kérdése. Ezt több oldalról vizsgáltuk, előbb az OSAP adatforrásainak publikációkban és adatátadásokban történő hasznosulása vonatkozásában, később pedig a felhasználók azonosításával próbálunk pontosabb képet kapni az adatgyűjtések vagy adatátvételek relevanciájáról.

A publikációkra vonatkozóan a válaszadók a következő lehetőségeket jelölhették meg: online tájékoztatás, nyomtatott kiadványok, média, biztonságos környezetben (pl. KSH Kutatószoba keretében), egyéb.

Mivel nagyságrendileg nem azonos az adatgyűjtések és az adatátvételek száma, fontosnak tartjuk megjeleníteni a publikálással kapcsolatos arányokat (21. sz. ábra) és darabszámokat (22. sz. ábra) is. A 265 adatgyűjtésből, valamint a 102 adatátvételtől ezért külön-külön kimutatjuk, mekkora százalékban történt publikálás az egyes csatornákon (lásd 21. sz. ábra). Az adatgyűjtések és adatátvételek esetén is vezető szerepet töltött be az online publikálás, az adatgyűjtések 84,4%-át, az adatátvételek 65,4%-át tették közzé ilyen módon. A HSSz azonban 2016-ban még nem rendelkezett semmilyen közös publikálást vagy a felhasználók jobb tájékozódását segítő koordinált megjelenési lehetőséggel, azaz az arányát tekintve döntő fontosságú internetes publikálás szervezetenként elszigetelten, jórészt saját korábbi gyakorlatuknak megfelelő formában, összehangolt fellépés nélkül történt. Ezen a jövőben célszerű változtatni.

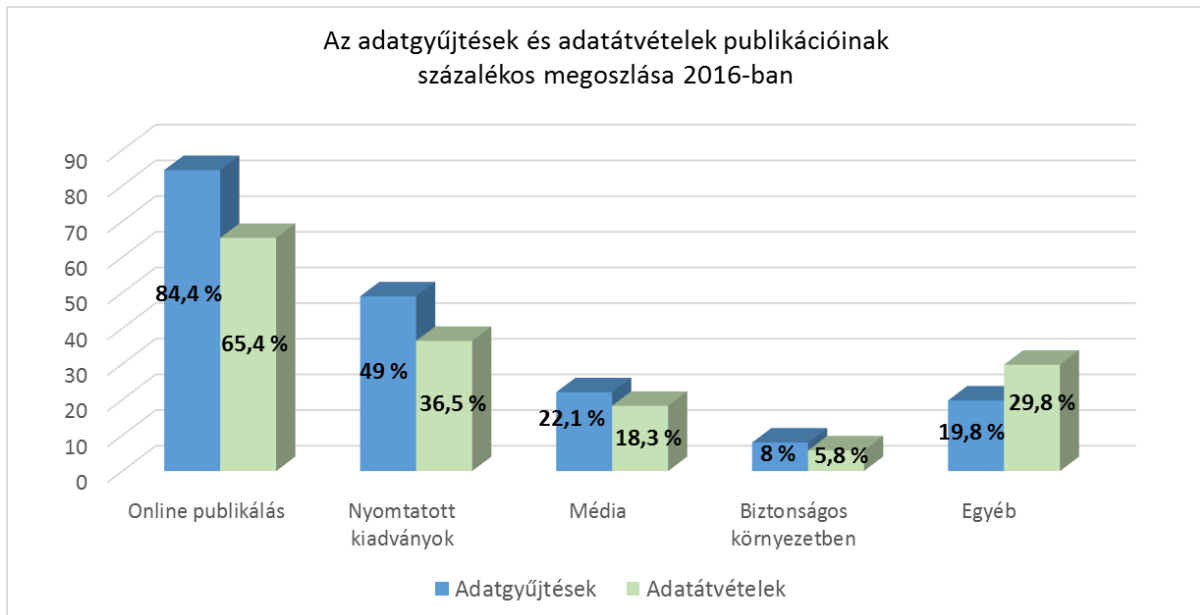
Az internetes megjelenést a nyomtatott kiadványok követik, arányukban szintén jelentős mértékben, adatgyűjtések esetén azok 49%-ában, míg az adatátvételeknél is 36,5%-ban. A nyomtatásban kiadásra kerülő adatokkal zömmel hagyományos évkönyvekben, zsebkönyvekben, illetve különböző más kiadványokban találkozhatunk.

A biztonságos környezetben (itt: szűken értelmezve a KSH Kutatószobájában) történő adathozzáférés biztosítása közvetlen azonosításra alkalmatlanná tett állományokhoz

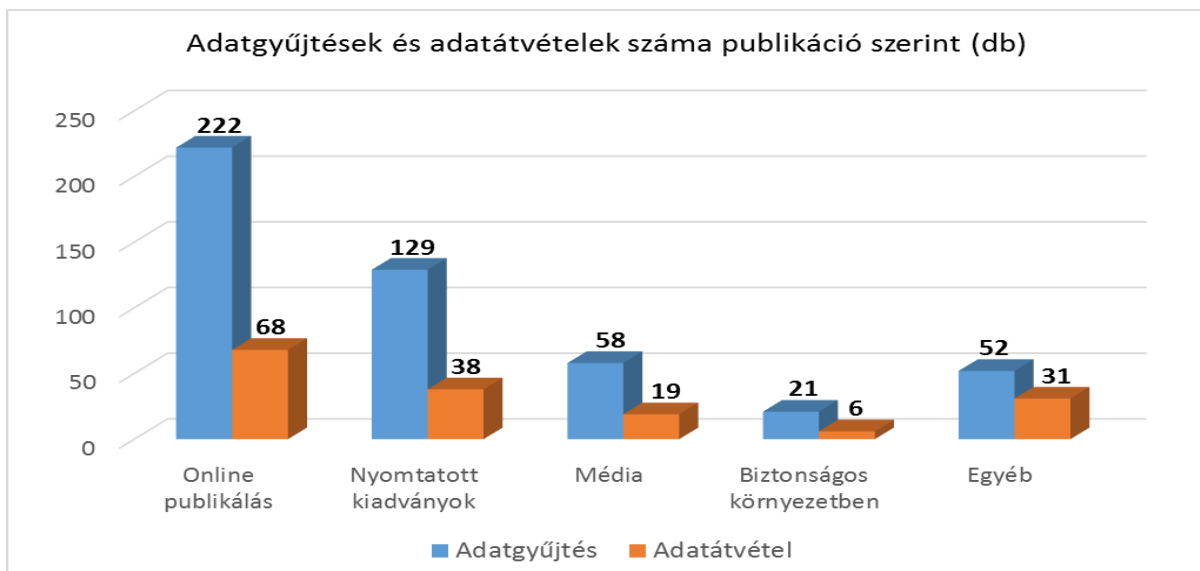
³ Nyilvántartási egység: személy, gazdasági szervezet vagy tárgy, amire az eredeti adatgyűjtés (nyilvántartás) adatai vonatkoznak. Például: természetes személyek, gazdasági társaságok, nonprofit szervezetek, gépjárművek stb. A nyilvántartás tárgyát képező sokaság egyede, egysége, az őt jellemző azonosító, elérési és egyéb tulajdonságokat leíró információkkal együtt.

(jellemzően: mikroadatokhoz) a legkevésbé alkalmazott publikációs módszer egyelőre akár az adatgyűjtések (8%), akár az adatátvételek (5,8%) körében.

Az adatgyűjtések 22,1%-a, míg az adatátvételek 18,3%-a esetében kerül publikálásra a belőlük származó adat a médián keresztül. Az összes kategóriát összehasonlítva, az adatátvételeknél a publikációk aránya csak az egyéb kategóriában magasabb az adatgyűjtésekhez képest.



21. sz. ábra



22. sz. ábra

Az egyéb kategóriánál felsoroltak alapján, az alábbi területeken történik a publikálás mind az adatgyűjtések, mind az adatátvételek esetén: megjelenés szakmai folyóiratokban, közvetlen adatkiadás megkeresésre kutatásokhoz, oktatási célra, gyorstájékoztatókba, beépül a Statatba, a kormányzati és nemzetgazdasági adatokba, valamint a negyedéves GDP flash becslés során.

Az egyéb hasznosulásként megjelölt területek is igen változatosak: szakmai konferenciák, nemzeti számlák összeállítása, kormányzati döntéselőkészítés, valamint előzetes és utólagos hatásvizsgálatok, miniszteri beszámolók, más adatállományok validálása, adatszolgáltatói kör kijelölése, jogszabályalkotás döntéselőkészítése, regiszterek karbantartása.

Azok az adatok, amelyek nem kerülnek publikálásra, az egyéb hasznosulásként megjelölt területeken felhasználásra kerülnek, valamint számos esetben részadatokként megjelennek egy későbbi publikációban vagy tovább feldolgozott adatokban, akár a KSH-n belül, szervezeten kívül, illetve az Eurostatnál, tehát ilyen formán elmondható, hogy átadásra kerülnek. Esetenként regiszterek karbantartásához készülnek adatok, illetve adatminőség javítása céljából, közvetlen publikálás nélkül. Néhány esetben nem jelentkezett igény a közzétételére. Viszont hasznosulnak döntéshozatal esetén, tárgyalásokon, szakdolgozatokban, egyedi igények kielégítésekor.

Egy adatgyűjtés esetén jelölték meg mind az öt tájékoztatási formát ('Jelentés a települések villamosenergia-ellátásáról'). Összesen 8 olyan adatgyűjtés volt 2016-ban, amelynél mind a négy konkrétan megnevezett publikációs csatornát megjelölték. Ezek különféle területekhez tartoznak és több tekintetben kiemelt jelentőségűek: 'Statiztikai adatszolgáltató lap a teresség megszakításáról', 'Kimutatás a pénzben és természetben nyújtható támogatások adatairól', 'Éves gazdaságstatisztikai jelentés', 'Jelentés a lakossági lakáshitelezési tevékenységről', 'Éves teljesítménystatisztikai jelentés', 'Éves munkaügyi jelentés', 'Gazdaságszerkezeti összeírás, 2016. (gazdasági szervezetek)', 'Gazdaságszerkezeti összeírás, 2016. (egyéni gazdaságok)'.

Az adatgyűjtéseknél jelentős arányban fordul elő a két, illetve három tájékoztatási csatorna megjelölése. Az adatgyűjtések körében a 263 visszaérkezett kérdőívből összesen 163 esetén jelölték meg ennyit, két csatornán 107, míg három csatornán 56 adatgyűjtés esetén. Egy féle publikálási forma 63 adatgyűjtés esetén fordult elő. 28 adatgyűjtésből a kérdőív adatai alapján nem publikáltak.

Az adatátvételeknél mind a négy publikációs csatornát 5 adatátvétel esetében jelölték meg a kitöltők („Személyesüléses közúti közlekedési balesetek”, „Egészségbiztosítási ellátások éves és területi adatai”, „Nyugdíjban, ellátásban, járadékban és egyéb járandóságban részesülők”, „Családtámogatási ellátások éves és területi adatai”, „Állománystatisztikai adatok a nyugdíjban, ellátásban, járadékban és egyéb járandóságban részesülőkről (január-emelés után)”).

Három csatornán 18, míg két csatornán 27 adatátvétel esetén publikáltak. Jelentős az egy féle publikálással érintett adatátvételek száma, mely 34 darab volt. 20 adatátvétel esetén a kérdőívek tanúsága szerint nem történt publikálás. Az utóbbiakat mind adatgyűjtés, mind adatátvétel esetében célszerű tovább vizsgálni, a másodlagos hasznosulás azonosítása érdekében.

Az adatok hasznosulása témakörében a publikálás mellett a statisztikai és a tudományos célú adathozzáférésre vonatkozóan is kértük a kérdőív kitöltését. A statisztikai célú adatátadásokra vonatkozó kérdésnél próbáltunk széles skálát biztosítani a válaszadók számára.

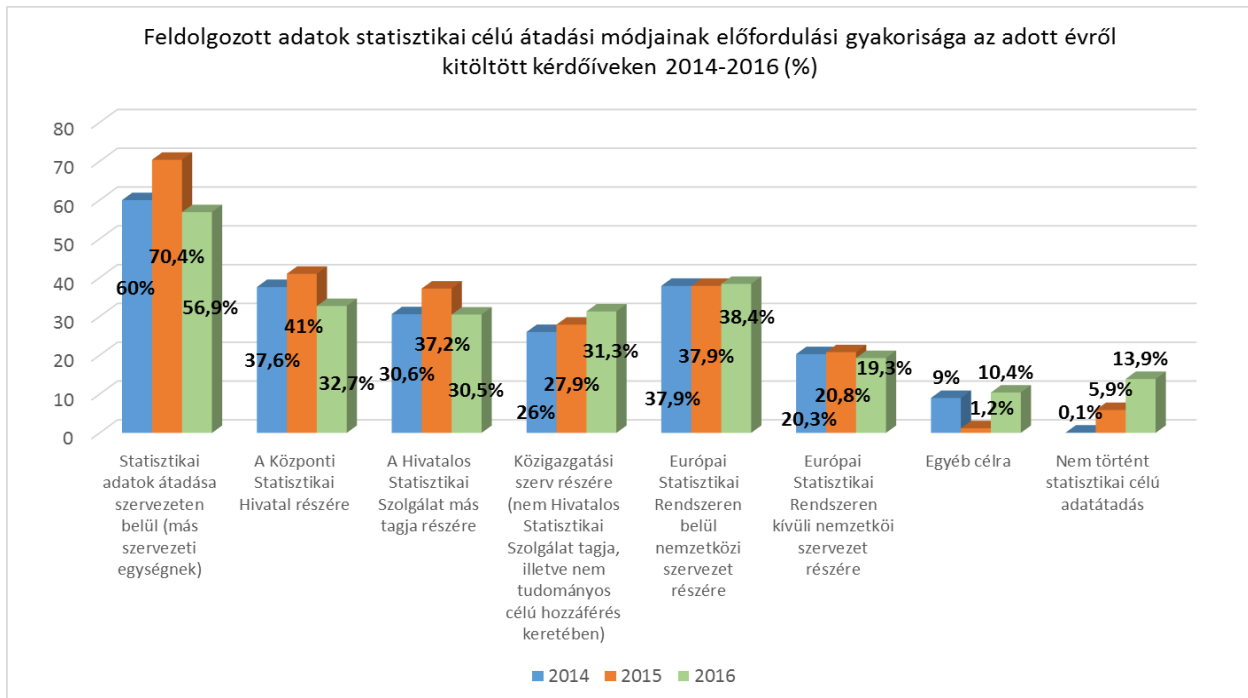
Megvizsgáltuk az adatátadásoknak az egyes felhasználói típusonkénti megoszlását (lásd 3. sz. táblázat). Az előző kérdéshez hasonlóan itt is több válasz párhuzamos megjelölésére volt lehetőség.

Feldolgozott adatok statisztikai célú adatátadásainak előfordulási gyakorisága a kérdőívben megjelölt módok szerint (db)

Adatátadások felhasználói típusai	2014. év	2015. év	2016. év
Statisztikai adatok átadása szervezeten belül (más szervezeti egységnek)	218	240	209
A Központi Statisztikai Hivatal részére	124	140	120
A Hivatalos Statisztikai Szolgálat más tagja részére	101	127	112
Közigazgatási szerv részére (nem Hivatalos Statisztikai Szolgálat tagja, illetve nem tudományos célú hozzáférés keretében)	86	95	115
Európai Statisztikai Rendszeren belül nemzetközi szervezet részére	125	126	141
Európai Statisztikai Rendszeren kívüli nemzetközi szervezet részére	67	71	71
Egyéb célra	30	4	38
Nem történt statisztikai célú adatátadás	3	20	51
Összes adatfelvétel, amelyről beérkezett értékelő kérdőív az adott évről (technikai összesen)	330	341	367

3. sz. tábla

Az adatokból megállapítható, hogy a feldolgozott statisztikák közigazgatáson belüli felhasználása érzékelhetően növekszik. Ebből arra lehet következtetni, hogy a közigazgatási döntéshozatali mechanizmusokban a korábbinál szélesebb körben használják fel a tényekre alapozott döntéshozatal támogatására a statisztikai adatokat. Ez a pozitív tendencia megfelel a hivatalos statisztika céljának. Az is látható, hogy a közvetlen adatküldések mellett egyre több adatforrás a feldolgozás után (is) hasznosul az Európai Statisztikai Rendszerben, ami jól szemlélteti a magyar hivatalos statisztika beágyazottságát az európai statisztikába. Fenti adatokat egymáshoz viszonyított százalékos megoszlásban is szemléltetjük (23. számú ábra).



23. sz. ábra

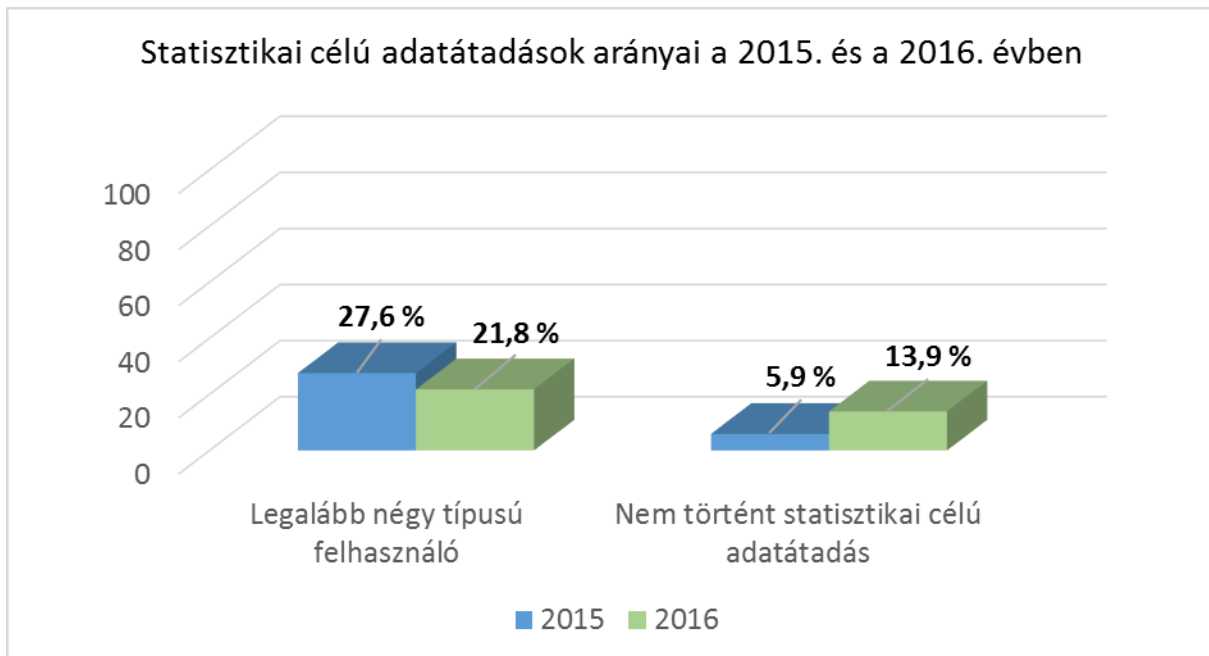
2016-ról a korábitól eltérően már külön is vizsgáltuk a statisztikai célú adatátadásokat KSH és HSSZ szerinti bontásban (4. számú tábla).

Adatátadások felhasználói típusai	Válaszok előfordulása	
	KSH	HSSZ
Statisztikai adatátadás - szervezeten belül (más szervezeti egységnek)	123	86
Statisztikai adatátadás - KSH részére	45*	75
Statisztikai adatátadás - HSSZ-nek	81	31
Statisztikai adatátadás - közigazgatási szervnek (nem HSSZ)	68	47
Statisztikai adatátadás - Európai Statisztikai Rendszeren belüli nemzetközi szervezetnek	104	37
Statisztikai adatátadás - Európai Statisztikai Rendszeren kívüli nemzetközi szervezetnek	47	24
Egyéb	32	6

4. sz. tábla (*: más HSSZ-tag fejezetében szereplő adatfelvételtől, amelyet a KSH hajt végre)

Megállapítható, hogy a KSH és a HSSZ között a feldolgozott adatok cseréje az üzemszerű működés részét képezi, illetve az is érzékelhető, hogy a nemzetközi adatátadások közel háromnegyedét a KSH teljesíti.

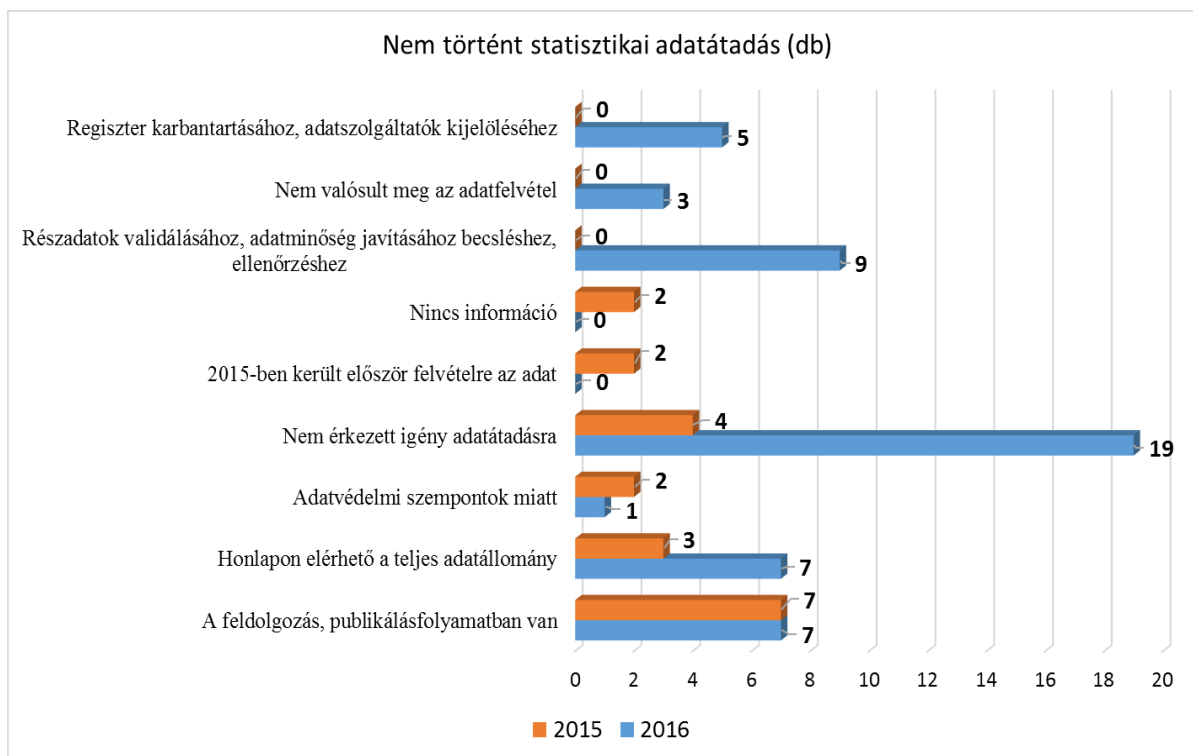
2016-ban az adatgyűjtések és adatátvételek 21,8%-a legalább négy típusú felhasználónak szolgáltatott adatot. Ez az arány 2015-ben 27,6% volt (lásd 24. sz. ábra). 2015-ben 13,9%-ban nem történt adatátadás az adatok feldolgozása után. 2016-ban ez az arány még csak 5,9% volt.



24. sz. ábra

2016-ban a kitöltők 1 adatfelvétel esetén jelölték meg mind a hét felkínált felhasználói típust a kérdőíven ('Jelentés a vállalkozások kutatási, fejlesztési adatairól'). Emellett 15 adatfelvétel esetén hat, 14 esetén öt felhasználói típust jelöltek.

A 25. sz. ábra azokat a válaszokat és előfordulásuk gyakoriságát mutatja, ahol a kitöltő nem jelölt meg egyetlen felhasználói csoportot sem.



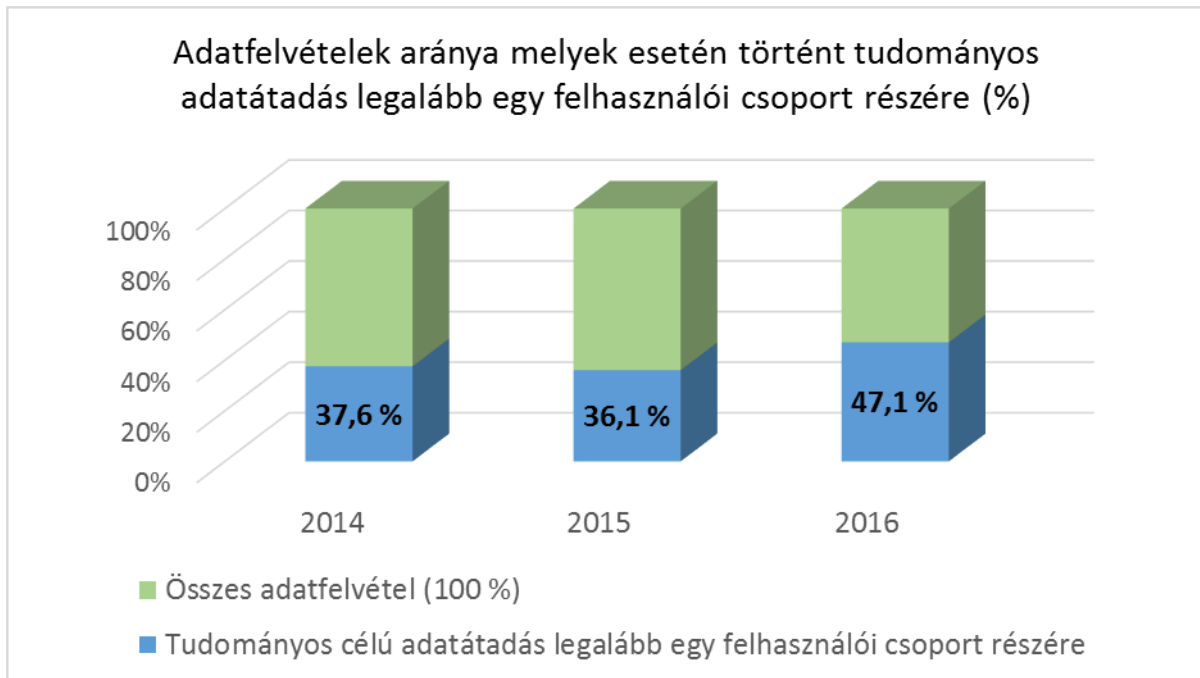
25. sz. ábra

Az indokok között a legtöbb egyszerűen indokolható, azonban figyelemre méltó, hogy 19 esetben a 2016 tárgyév vonatkozásában nem érkezett igény a feldolgozott adatok átadására. Az ezen adatfelvételekből származó statisztikai adatok további felhasználási lehetőségeit érdemes megvizsgálni.

A statisztikai célú adatátadások mellett vizsgáltuk a tudományos célú adatátadások mértékét is 2014-től 2016-ig (lásd 5. számú tábla és 26. számú ábra). A táblázatban és az ábrán is szerepel minden adatfelvétel, ahol legalább egy felhasználónak történt átadás tudományos célból.

Adatfelvételek száma, melyek esetén történt tudományos célú adathozzáférés biztosítása			
	2014	2015	2016
Tudományos célú hozzáférés biztosításában érintett adatfelvételek száma	124	123	173
Összes adatfelvétel száma, amelyről értékelő kérdőívet töltöttek ki	330	341	367

5. sz. tábla

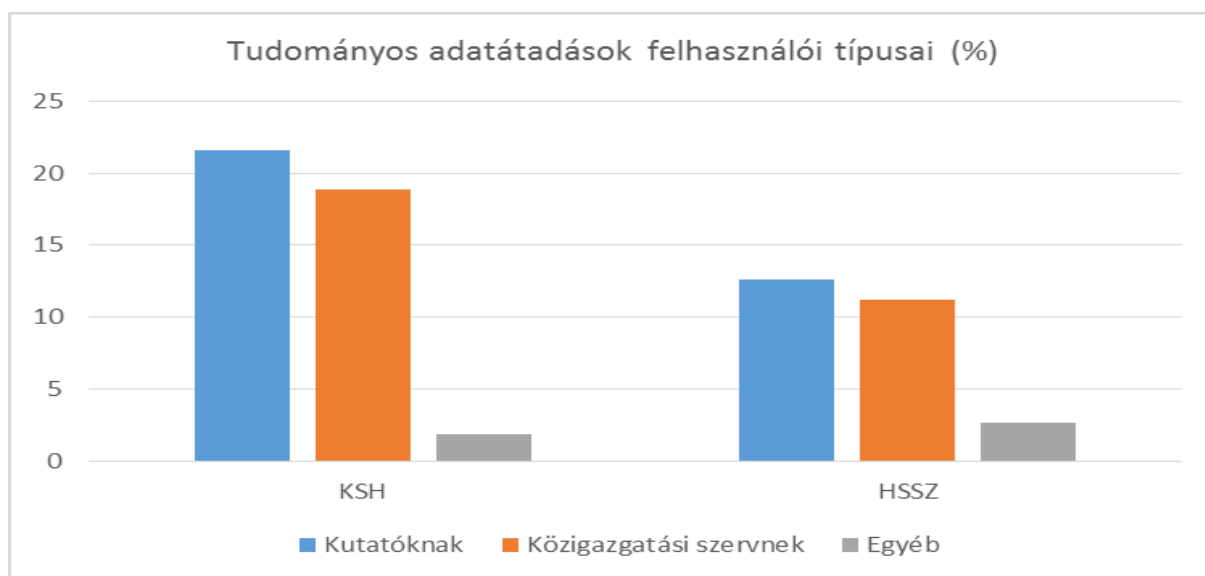


26. sz. ábra

A 2016-os évre vonatkozóan külön vizsgáltuk a tudományos élet résztvevőinek történt adatátadások formáit. Összesen 173 adatfelvétel esetén került legalább egy felhasználói csoport részére tudományos célú adathozzáférés biztosításra. Egy-egy adatfelvétel több sorban is szerepelhet annak megfelelően, hogy egy, vagy több felhasználói típusnak adták át az adatokat. (lásd 6. számú táblázat)

Tudományos célú adathozzáférés felhasználói típusai	Válaszok elfordulási gyakorisága	
	KSH	HSSZ
Tudományos célú adathozzáférés – kutatók részére	79	46
Tudományos célú adathozzáférés – közigazgatási szerv részére	69	41
Tudományos célú adathozzáférés – egyéb	7	10

6. sz. tábla



27. sz. ábra

A feldolgozott adatok továbbfelhasználásáról informál az adatátadások formáira vonatkozó kérdés, melyről a 2015. és a 2016. év vonatkozásában van adatunk.

Az adathozzáférés formája, előfordulás gyakorisága (db)		
	2015	2016
Adatátadás anonimizált mikroadatként	107	87
Adatátadás táblázatos adatként	302	315
Mindkettő	88	72
Egyik sem	20	37

7. sz. tábla

2015-ről 2016-ra csökkent azoknak az adatfelvételeknek az aránya, melyek statisztikai információiból anonimizált mikroadatként adatok kerültek átadásra statisztikai célra, az aggregált adatok átadása ugyanakkor növekedett, ám a mindkét formában történő adatátadások előfordulása ugyancsak csökkent. (7.sz. tábla)

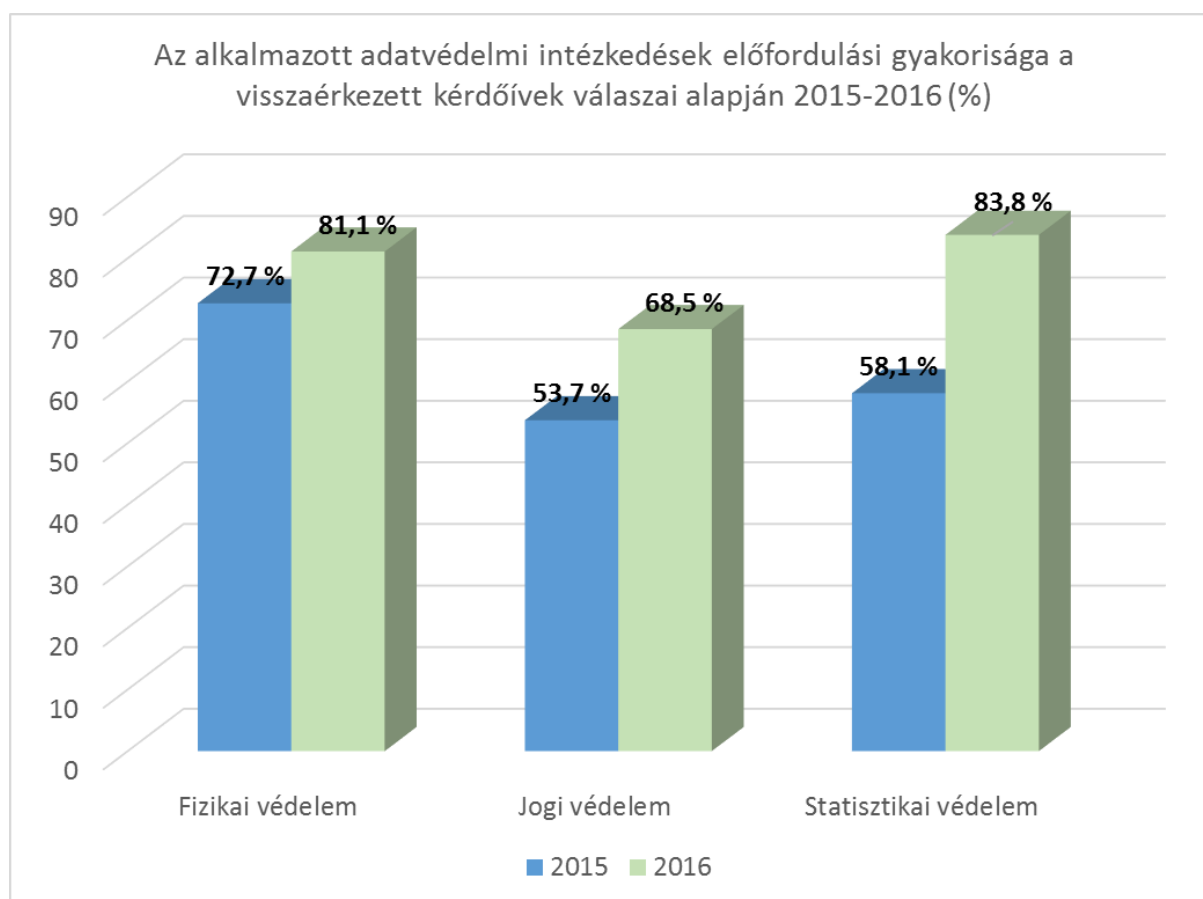
Adatvédelem

Az adatok védelméről szóló kérdésnél arra voltunk kíváncsiak, hogy az adatgazdák milyen fizikai-, jogi-, illetve statisztikai felfedés elleni védelmet alkalmaznak az illetéktelen adathozzáférések megakadályozására.

A rendelkezésünkre álló adatok alapján a végrehajtó szervezetek 2014-ben és 2015-ben is az adatfelvételek több mint 90%-ánál alkalmaztak valamilyen adatvédelmi technikát, 2016-ban ez az arány már elérte a 94,5%-ot.

2016-ban mindhárom típusú adatvédelmi technikát az esetek 63,3%-ban alkalmazták a statisztikai adatok védelmére. 20 esetben (5,5%-ban) nem kaptunk erről információt.

Az egyes adatvédelmi intézkedések alkalmazását tekintve 2015-ben, illetve 2016-ban egyértelmű növekedés volt tapasztalható mindhárom típus esetében (lásd 28. sz. ábra).



28. sz. ábra

Fizikai adatvédelmet a kitöltők 81,1%-a jelölt meg az értékelő kérdőíveken. Ez az arány 11,6%-kal magasabb a 2015. évi teljesüléshez képest. Itt zömében a szervezetek Informatikai Biztonsági Szabályzatai (IBSZ), egyéb belső szabályzatok, elektronikus adatgyűjtő és – továbbító rendszerek (pl.: KARÁT, ELEKTRA, EDAMIS) kerültek megjelölésre. Emellett az

adatgazdák szeparáltan tárolják az adatokat és az azonosítókat, a papír kérdőívek zárt helyiségekbe kerülnek, az IBSZ-ekkel összhangban az informatikai rendszerek zártak, az adatokhoz szigorú hozzáférési engedélyek szükségesek, melyeket jelszavas védelemmel tesznek biztonságossá.

Jogi védelem kategóriában a kitöltők legtöbbször hazai jogszabályokat, belső szabályzatokat jelöltek meg. A konkrétumok között leginkább a titoktartási nyilatkozatok, a szerződéses feltételek szerepeltek az összesen 250 kitöltött válaszban, mely 2016-ban az adatfelvételek 68,5%-át jelentette.

A statisztikai felfedés elleni védelemre fordítanak a legtöbb gondot az adatgazdák a felhasznált eszközök kiterjedtségét alapul véve. A válaszadók ezen a téren jelentős javulásról számoltak be 2015-höz képest. A felfedés elleni védelmet az adatgyűjtések és adatátvételek 83,8%-ánál használják az adatgazdák. Itt szinte a teljes paletta megjelent a válaszok között. A hármasszabály mellett sok esetben alkalmaznak aggregálást, kódolást, anonimizálást, vagy rekordcserét annak érdekében, hogy azonosításra alkalmatlan adat kerülhessen csak a felhasználók kezébe.

A kérdőívet tovább szükséges pontosítani annak érdekében, hogy az adatvédelem alkalmazásának esetleges hiányáról, indokairól pontos képet kapjunk.

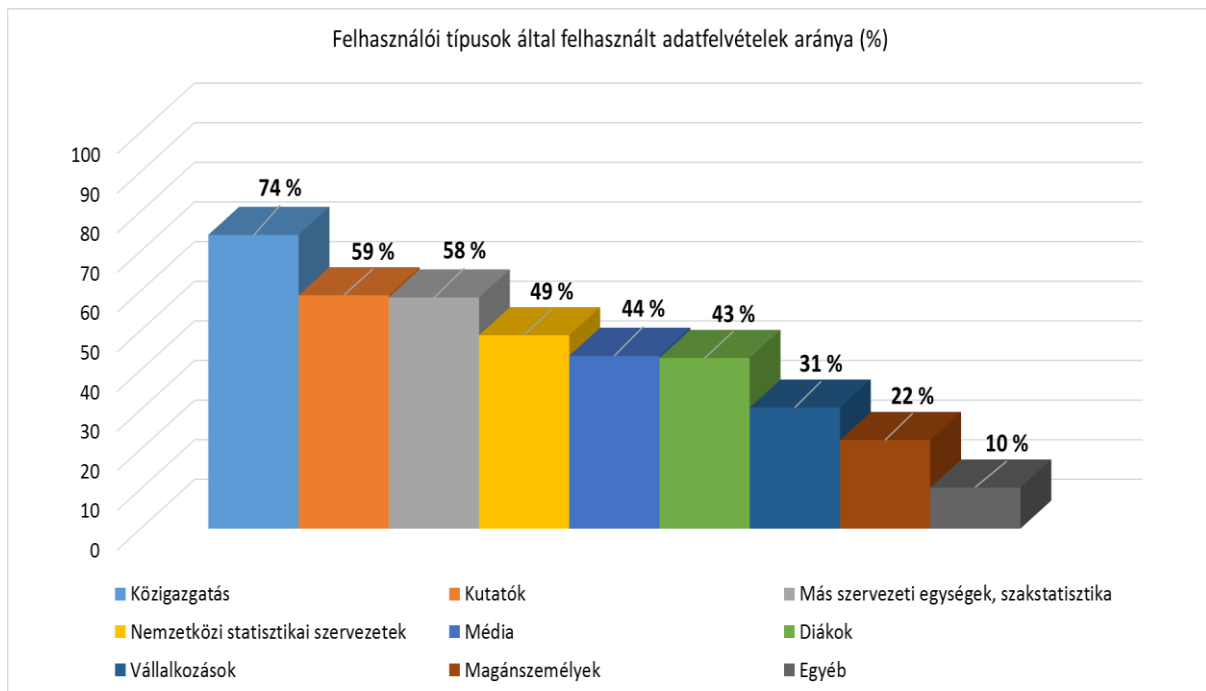
Felhasználók

A felhasználói csoportok megjelölésénél az előre meghatározott legjellemzőbb válaszkategóriákon túl lehetőség volt az egyéb felhasználókat szabadszövegesen is feltüntetni.

A begyűjtött és átvett adatokat az alábbi felhasználói csoportok használják: közigazgatás, kutatók, más szervezeti egységek, szakstatisztika, nemzetközi statisztikai szervezetek, média, diákok, vállalkozások, magánszemélyek, egyéb felhasználók.

Az összes adatfelvétel tekintetében a legnagyobb arányú felhasználás a közigazgatási szervek részéről mutatkozott, ők az OSAP végrehajtott adatgyűjtéseinek és adatátvételeinek 74%-ából használtak fel információkat, mely a 367 kitöltött kérdőívből 272 adatfelvételt jelent. A korábbi évekhez képest lényeges változás nem mutatható ki, 2014-ben és 2015-ben is 72% volt ez az arány.

További kiemelkedő felhasználói körök: a kutatók, a szakstatisztika, valamint a nemzetközi statisztikai szervezetek.



29. sz. ábra

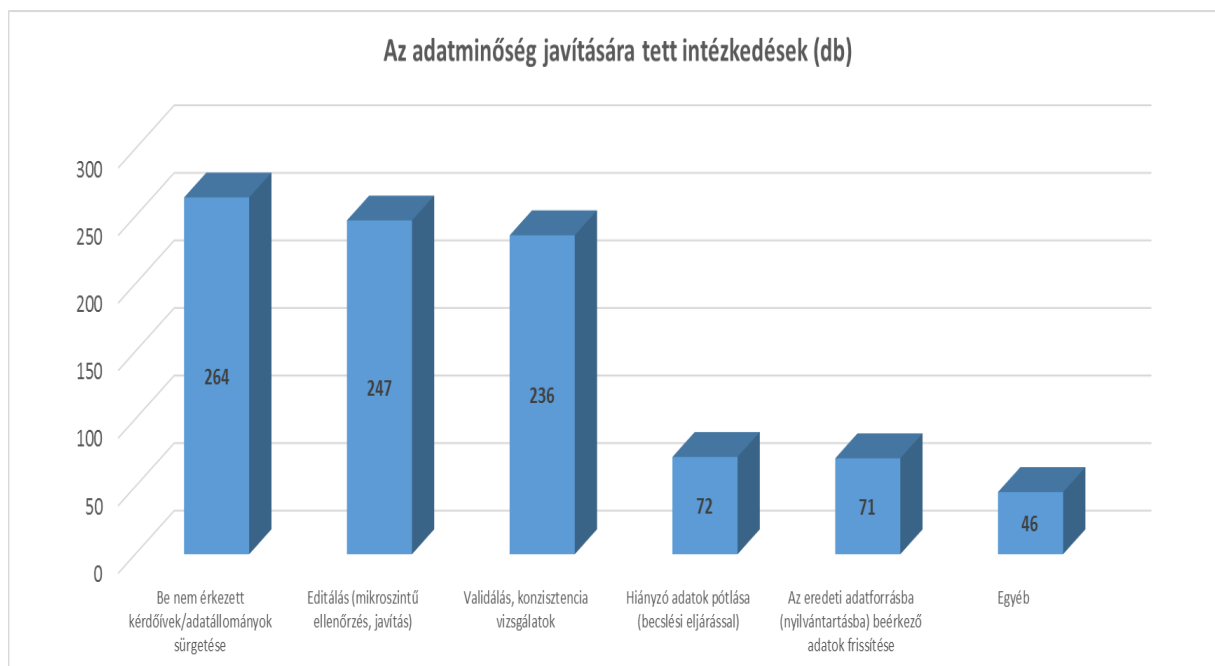
A felhasználói csoportok adatigényének egymáshoz viszonyított mértéke a korábbi évekkel azonos volt, a legnagyobb a közigazgatásé, a legkisebb a magánszemélyeket követően az egyéb kategória. Utóbbi 43 adatfelvétel esetén fordult elő. Az egyéb kategória magában foglalja például a szakmai szervezeteket, érdekképviselőket, illetve az adatok a Tájékoztatói Adatbázisban jelennek meg, ahol bármely felhasználói csoport hozzáfér, továbbá közvetett felhasználásra is kerülnek például más adatfelvételek pontosítása céljából.

2014-ben 7, 2015-ben 1, 2016-ban pedig 11 esetben jelöltek meg olyan adatfelvételt, mely nem került felhasználásra egyik felhasználói csoport által sem. A 2016. évi 11 adatfelvételből hat nem valósult meg, melyek közül három 2017. évtől megszűnik, mert más adatfelvétel helyettesíti, illetve a KSH-hoz kerül az eddig más szerv által kezelt adattár. A fennmaradó esetekben vizsgálatra javasolt a felvétel hasznosulása, hiszen a hivatalos statisztika, mint tevékenység legfontosabb célja a felhasználók tájékoztatása.

A 367 vizsgált adatfelvételből 56,1%-ban, 206 esetén legalább négy felhasználói csoport használta az átvett és gyűjtött adatokat. 2015-ben ez az arány még csak 54,3%-ot mutatott, ami arra utal, hogy kissé nőtt a HSZ adatainak kiterjedt használata.

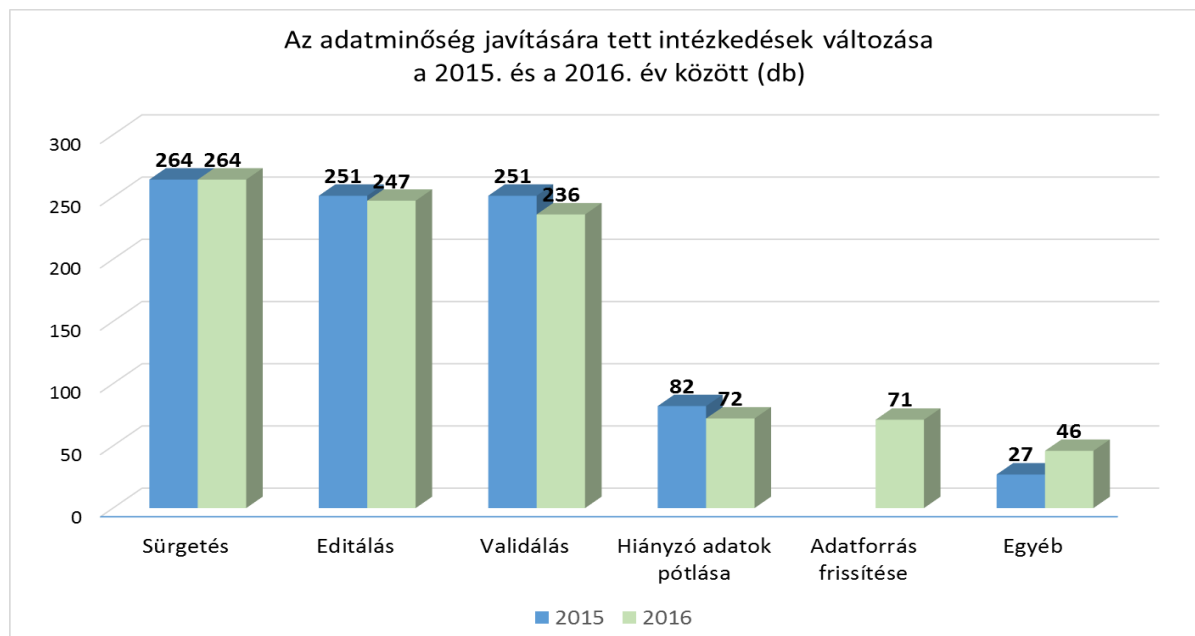
Minőség értékelése és javítása

Az adatminőség javítására tett intézkedésekre vonatkozó kérdéseknél is a legjellemzőbb válaszkategóriákon túl lehetőség volt egyéb, szabad szöveges válaszadásra.



30. sz. ábra

A 2015. és 2016. vonatkozási évben is a leginkább alkalmazott minőségjavítási intézkedések a sürgetés, editálás és validálás voltak, mely az adatfelvételek körülbelül 72%-át érintik. Az alábbi ábrán látható, hogy 2015-ben és 2016-ban is közel azonos mértékben volt szükség e három minőségjavítási intézkedésre. Az eredeti adatforrásba (nyilvántartásba) beérkező adatok frissítéséről nincs számszerű adatunk 2015-re.



31. sz. ábra

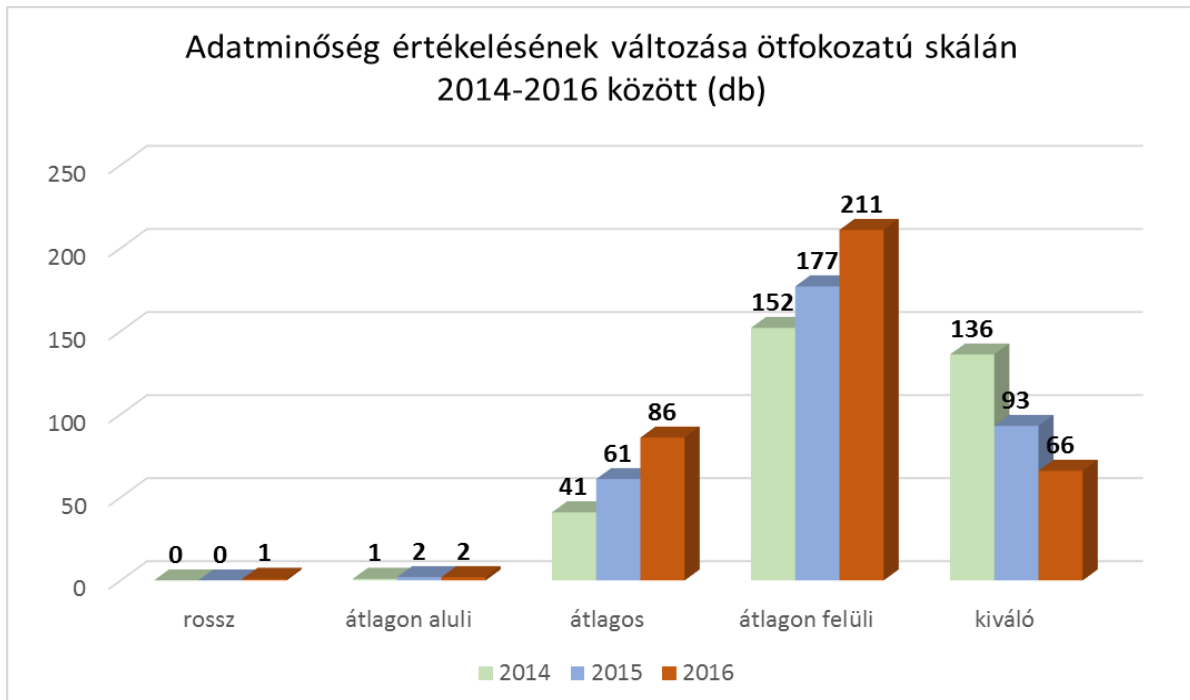
Az egyéb kategóriában a 2015. évben legtöbbször a trendelemzést, az idősor összehasonlítást, az outlier-szűrést, a szoros adatszolgáltatói kapcsolatartást, a szűrőpróbaszerű ellenőrzéseket, valamint az intézmények közötti rendszeres egyeztetéseket jelölték meg a kitöltők. Ezeket a 2016. évben kiegészítik az alábbiak: teljességellenőrzés, egyedi szakértői becslések, mezovalidálás, súlyozás, adminisztratív adatforrásokból inputálás, duplikációk kiszűrése.

2014-ben az adatfelvételek több mint 40%-a esetén alkalmazták egyidejűleg a három leggyakoribb minőségjavítási intézkedést (sürgetés, editálás, validálás). Ez 2016-ra sem változott, mert ebben az évben az adatfelvételek 44,7%-ában továbbra is mindhárom intézkedést használták egyidejűleg.

2015-ben az adatfelvételek 55,1%-ában használtak legalább három minőségjavítási intézkedést. 2016-ban ez az arány 52,9% volt.

2015-ben 7 esetben nem történt semmilyen minőségjavítási intézkedés, amelyet azzal indokoltak a válaszadók, hogy ezeknél az adatfelvételeknél már korábban megtörtént az adatminőség felülvizsgálata és javítása. Szükséges azonban megjegyeznünk, hogy az adatminőség javítása fogalmilag folyamatos tevékenység, így ezek a válaszok vizsgálatra szorulnak. 2016-ban 15 adatfelvételnél nem történt minőségjavítási intézkedés, melyből 2 esetben az adatgazda korábban ellenőrizte az adatokat, 6 adatfelvétel nem valósult meg, 3 esetben a KSH rendszereibe (ELEKTRA, ADÉL) beépített ellenőrzések biztosítják a gyűjtött adatok minőségét, illetve 3 esetben a 2016-ban átlagos minőségűnek minősített adatgyűjtés/adatátvétel megszűnik 2017-től. 1 esetben pedig átlagon felülinek értékelte az adatok minőségét a kitöltő.

A minőségmérés kérdésköre különösen aktuális a hivatalos statisztika fejlesztéseinek körében. Az adatok minőségének és pontosságának értékelését jelen kérdőívben egy öt fokozatú skálán is meg kellett adniuk az adatgazdáknak. Egyelőre nincsenek objektív eszközeink az egész Hivatalos Statisztikai Szolgálatra nézve egy teljes adatminőség-értékeléshez, ezért a bemutatott információ az adatgazdák önértékelésén alapul. A 2014-2016-ig tartó időszakban történt változások az adatgazdák szubjektív értékeléséből az alábbi ábrán láthatóak (32. számú ábra).



32. sz. ábra

Elmondható, hogy az átlagon felüli és kiváló minőségű adatok összessége a vizsgált három évben csekély mértékben változott, 2014-ben összesen 288, 2015-ben 270, 2016-ban 277 adatfelvételt soroltak e két kategóriába az adatgazdák.

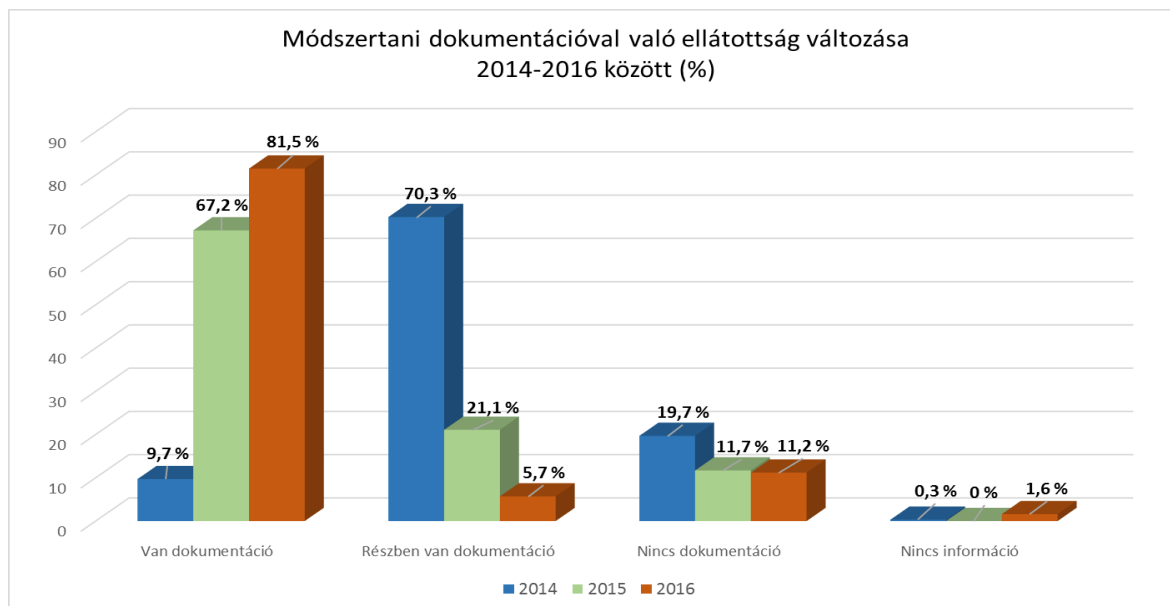
Azonban, ahogy a fenti ábrán is látszik, az átlagon felülinek értékelt adatfelvételek száma növekedett, a kiválónak minősítettek ugyanakkor a felükre csökkentek. Ezt a tendenciát úgy értelmezzük, hogy az adatgazdák több szakmai ismeret birtokában a korábbinál kritikusabban ítélik meg saját adataik minőségét.

A rossz és átlagon aluli minősítésű adatfelvételek részleges adatokkal foglalkoznak, kettő közülük meg is szűnik.

A korábbi évekhez hasonlóan az adatok minőségének javítására tett célkitűzéseket nem csak az átlagos, hanem az átlagon felüli és kiváló adatminőségű adatfelvételek esetén is megfogalmazták a kitöltők.

Megfelelő módszertani információk rendelkezésre állása és visszacsatolás az adatszolgáltatók/adatgazdák felé

Az adatfelvételek módszertani dokumentációval (fogalmak, osztályozások, módszerek, stb.) való ellátottságának változása 2014-2016 években (33. számú ábra).



33. sz. ábra

Jelentős javulás mutatkozott ilyen rövid idő alatt, hiszen míg a 2014. vonatkozási évben az adatfelvételeknek mindössze 9,7%-ánál rendelkeztek teljes módszertani dokumentációval, addig 2016-ra már 81,5%-nál jelezték ezt az állapotot az adatgazdák.

Ennek megfelelően 2014-ről 2016-ra a részleges dokumentációk száma hasonló nagyságrendben csökkent, 70,3%-ról 5,7%-ra.

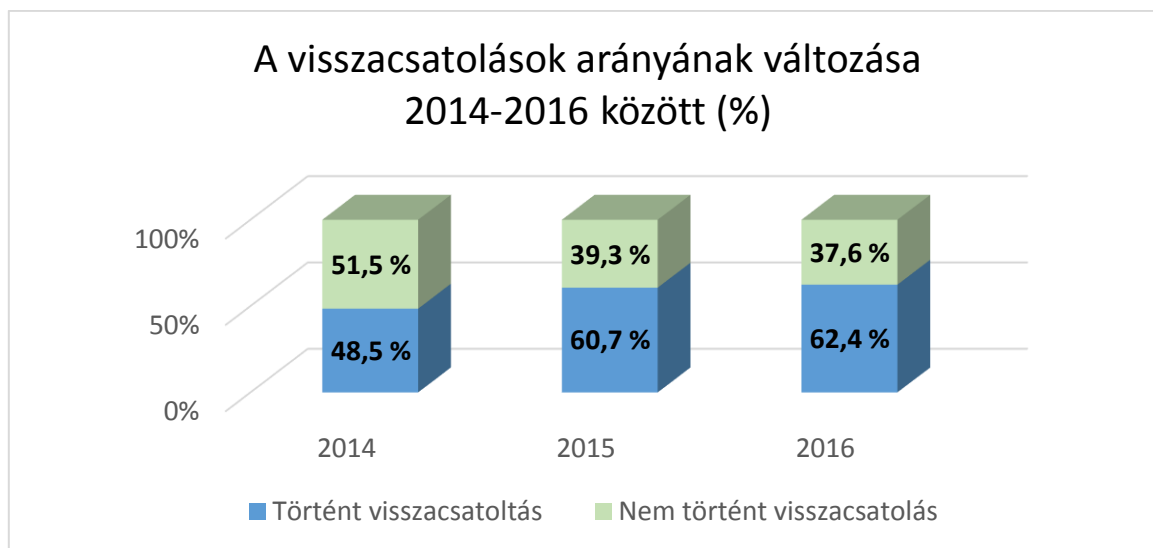
Javulás mutatkozott a korábban dokumentáció nélküli adatfelvételek tekintetében is. 2014-ben 19,7%-uknál, 2016-ban már csak 11,2%-uknál nem volt dokumentáció.

2016-ban 299 adatfelvétel rendelkezik teljes módszertani dokumentációval. A 41 részleges dokumentációval rendelkező adatfelvétel esetén jellemzően a dokumentáció fejlesztését tűzték ki célul a kitöltők. A fennmaradó 41 adatfelvétel nem rendelkezik módszertani dokumentációval.

Négy alkalommal a 2017., hat alkalommal a 2018., két alkalommal a 2019., egy alkalommal a 2020. évben szeretnék bevezetni a dokumentációt. Tizenhárom adatfelvétel esetén az adatgyűjtő szervezet úgy nyilatkozott, hogy nem tartotta szükségesnek a módszertani dokumentáció bevezetését, mert informatikai rendszerből, adminisztratív adatbázisból származnak az adatok, valamint az adatszolgáltatói kör meghatározásához, illetve más adatfelvétel validációjához használják az adatokat. Ki kell emelnünk azonban, hogy az adatfelvételi folyamat módszertanának dokumentálására ettől még szükség van. Három

esetben megszűnő adatfelvétélről töltötték ki a kérdőívet. Két-két esetben jelezték, hogy bizonytalan időpontban, valamint ha a technikai háttér rendelkezésre áll, akkor tervezik a módszertani dokumentáció elkészítését és bevezetését.

Míg a 2014. vonatkozási év esetében az adatfelvételek 48,5%-ában történt visszacsatolás az adatszolgáltatók/adatgazdák részére az általuk benyújtott adatokról, addig 2016-ban már 62,4%-ban történt visszajelzés (lásd 34. sz. ábra). A nagyobb változás 2014-ről 2015-re figyelhető meg, azonban 2015 és 2016 között is lassú javulás mutatkozott.



34. sz. ábra

2016-ban összesen 138 kérdőív szerint nem történt visszacsatolás, elsősorban azért, mert az adatszolgáltatók/adatgazdák részéről nem merült fel ilyen igény. Emellett ugyanakkor jelentős a publikációkra, mint visszacsatolásra való hivatkozás, illetve a folyamatos kapcsolattartás az adatszolgáltatókkal, amely eszközök mellett az adatfelvételek felelősei további célzott visszacsatolást nem tartottak szükségesnek. Ebben a tekintetben a Hivatalos Statisztikai Szolgálat részéről több proaktív kezdeményezésre lenne szükség, amely rávezetheti a partnereket a visszacsatolás fontosságára, felhasználási lehetőségeire.

2016-ban összesen 229 adatfelvétélről érkezett a kérdőíven visszajelzés arról, hogy valamilyen visszacsatolás történt az adatszolgáltatók/adatgazdák részére. A visszacsatolás eszközei publikációk (nyomtatott és elektronikus kiadványok), évkönyvek, szakstatisztika, minőségjelentés, KARÁT és ELEKTRA rendszer visszaigazolások, jelentések, riportok, teljes adatállományok visszaküldése, beszámolók. E válaszok mellett került megjelölésre a folyamatos kapcsolattartás gyakorlata, valamint az egyéni adatkérés esetén is kapnak visszajelzést az adatszolgáltatók/adatgazdák.

Budapest, 2017. szeptember 15.

KSH Statisztikai koordinációs főosztály