



A 2021. ÉVI ORSZÁGOS STATISZTIKAI ADATFELVÉTELI PROGRAM (OSAP) TELJESÜLÉSE

2022. november

Bevezetés

A Központi Statisztikai Hivatal (KSH) 2014 óta évente, a hivatalos statisztikáról szóló 2016. évi CLV. törvény (a továbbiakban: Stt.) hatálybalépését követően azzal összhangban készít összefoglaló elemzést az Országos Statisztikai Adatfelvételi Program (OSAP) teljesüléséről. Ez nemcsak a KSH statisztikai tevékenységének főbb adatait tartalmazza, hanem a Hivatalos Statisztikai Szolgálat (HSSz) egészének statisztikai feladatellátásáról is számot ad. A tájékoztatót a KSH a Nemzeti Statisztikai Koordinációs Testület (NSKT), valamint az Országos Statisztikai Tanács (OST) egyetértésével teszi közzé.

Az elemzés forrása a KSH nyilvántartása az OSAP tartalmáról, továbbá az adatfelvételek¹ felelősei által kitöltött értékelő kérdőívek, illetve az egyéb, rendelkezésre álló nyilvántartások. A kérdőívek a 2021-ben végrehajtott vagy végrehajtani tervezett adatfelvételeket vizsgálják, kivéve azokat az adatfelvételeket, amelyeknél az adatok – azok határidejére tekintettel – a kérdőívek kitöltésekor még nem érkeztek be. Ezekben az esetekben az elemzés adatai a 2020-ban végrehajtott felvételekre vonatkozó adatokat tükrözik.

Összességében megállapítható, hogy a 2021-es vonatkozási évre szóló OSAP a terveknek megfelelően valósult meg. Ki kell emelni, hogy az OSAP az Stt.-ben foglaltak alapján, valamint az annak végrehajtását szolgáló 184/2017. (VII. 5.) Korm. rendelet hatálybalépésével 2018 végére lényeges átalakuláson ment át, tartalma a statisztikai adatfelvételeknek az új jogszabályon alapuló, egységes kezelése révén jelentősen bővült. Növekedett a statisztikai adatátvételek száma, ezzel párhuzamosan a statisztikai adatgyűjtések lassú csökkenése figyelhető meg. A folyamat összhangban van az Stt.-ben előírtakkal, miszerint csakis akkor lehet elsődleges adatforrást használni, ha nincs olyan másodlagos adatforrás, amelyből a kívánt adatok előállíthatók.

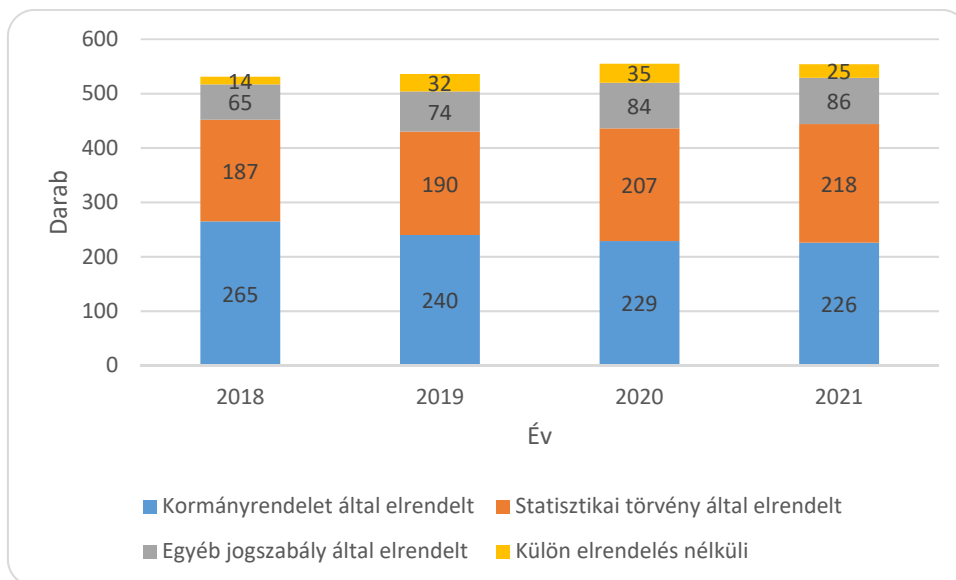
Az új szemléletű OSAP egységes eszként nyilvántartja a hivatalos statisztika összes adatfelvételét, egységes struktúrában és egységes, a Központi Statisztikai Hivatal által kialakított metaadat-nyilvántartási és -dokumentálási rend szerint.

Az adatgyűjtések és adatátvételek alakulása

2021-ben 555 adatfelvétel végrehajtását rendelte el az OSAP, ebből 246 adatgyűjtés és 309 adatátvétel volt. Ahogy az 1. ábrán is láthatjuk, az előző évhez képest az adatfelvételek száma változatlan, ugyanakkor darabszámuk az elrendelésük jogalapja szerint módosult. Ennek az az oka, hogy a HSSz tagjainak egyre inkább az a céljuk, hogy adminisztratív nyilvántartásokból származó adatokból állítsanak elő hivatalos statisztikai adatokat.

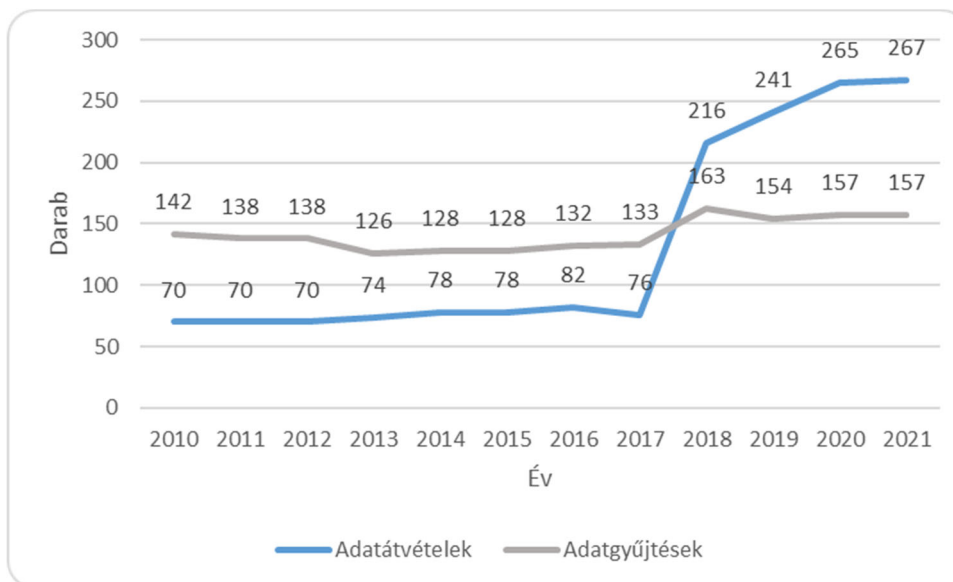
¹ Statisztikai adatfelvétel: egy adott sokaság adott időszakra vagy időpontra vonatkozó ismérveinek gyűjtése egy társadalmi, gazdasági vagy környezeti jelenség statisztikai megfigyelése céljából, különböző adatforrások felhasználásával. Ezen belül adatgyűjtés a statisztikai adatfelvétel által meghatározott 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.

1. ábra Az OSAP tartalma az előző évihez képest

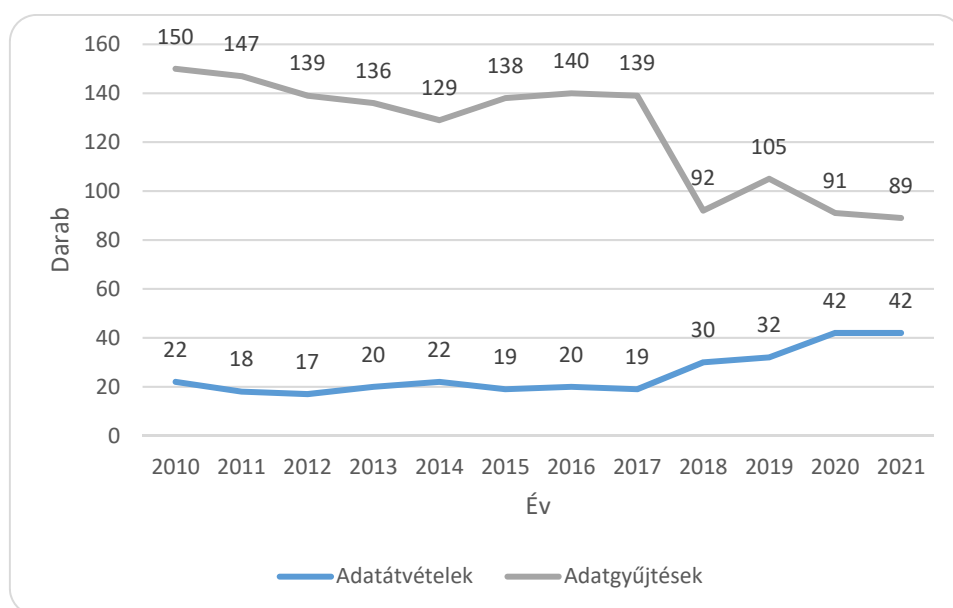


A **2. és a 3. ábrán** összesítve látható az elmúlt 10 évben az OSAP-ban elrendelt adatgyűjtéseknek és adatátvételeknek alakulása, külön a KSH-ra és a KSH-n kívüli HSSz-tagokra, szemléltetve a 2021. évi változásokat az előző évekhez képest. 2021-ben az egy évvel korábbihoz viszonyítva a KSH adatátvételeinek száma tovább nőtt, a KSH-n kívüli HSSz-tagok adatátvételeinek száma változatlan maradt, míg az adatgyűjtéseké a KSH-ban és a HSSz többi tagjánál egyaránt csökkent.

2. ábra A KSH adatfelvételeinek alakulása



3. ábra A KSH-n kívüli HSSz-tagok adatfelvételeinek alakulása

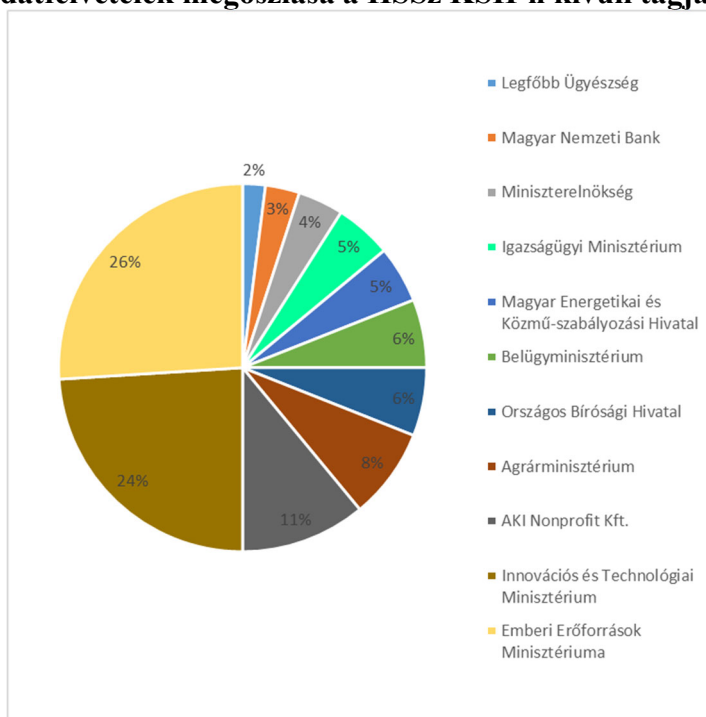


2021-ben a kormányrendeleti részben 226 adatfelvétel elrendelése történt meg. Ez a szám 2018 óta szinte folyamatosan csökken, mivel az adatátvételek elrendelése nem kormányrendeletben történik, és az elmúlt években a HSSz tagjai folyamatosan kutatják azon adminisztratív adatforrásokat, amelyek kiválthatják az adatgyűjtéseket, ezáltal csökkentve az adatszolgáltatói terheket is.

1. tábla A kormányrendeletben elrendelt adatfelvételek száma a végrehajtó szerint

Adatfelvétel végrehajtója	2016	2017	2018	2019	2020	2021
KSH	214	209	142	135	139	135
A HSSz KSH-n kívüli tagjai	160	158	123	105	90	91
Összesen	374	367	265	240	229	226

4. ábra Az adatfelvételek megoszlása a HSSz KSH-n kívüli tagjai között, 2021



A HSSz KSH-n kívüli tagjai esetében 2021-ben az Emberi Erőforrások Minisztériumához (EMMI) tartozott a legtöbb adatfelvétel, (26%), ezt követte az Innovációs és Technológiai Minisztérium (ITM, 24%). (Az arányok hasonlóak a 2020. évihez.).

A kiküldött 555 teljesülési kérdőívből 550 érkezett vissza a KSH részére. Az elemzéshez beérkezett kérdőívek alapján 2021-ben 529 adatfelvétel valósult meg, az összes elrendelt adatfelvétel 96%-a. 14 esetben még nem történt meg az adatfelvételek végrehajtása. A meg nem valósulások fő oka az volt, hogy az adatszolgáltató teljesítésre előírt határideje későbbi időpontra esik, mint a teljesülési kérdőív visszaküldésének határideje, illetve a beérkezett adatok feldolgozása még nem zárult le. 12 esetben nem érkezett válasz a kérdőív 1.1.-es kérdésére („Az adatgyűjtés, adatátvétel végrehajtása megvalósult-e?”).

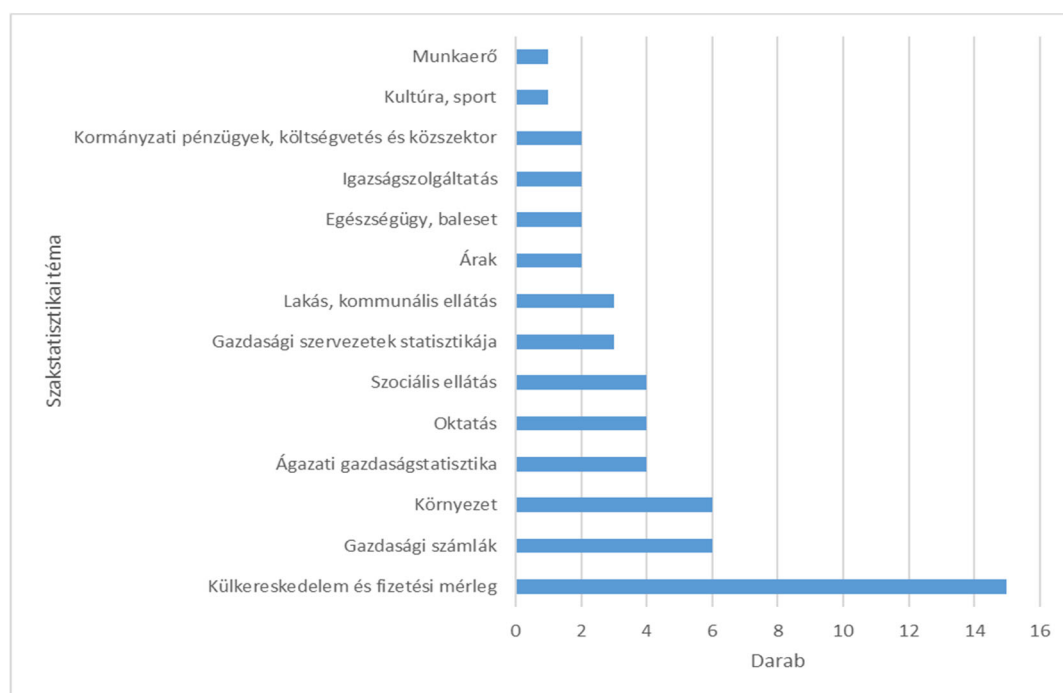
A megvalósult 529 adatfelvételtől 474 esetben a teljesülési kérdőív kitöltéséig a feldolgozás is befejeződött, ez az összes, 555 darab adatfelvétel 85%-át tette ki. Szinte az összes ilyen esetben a már előállt előzetes adatokról készítették el a teljesülési kérdőíveket.

A fentiek alapján az időszerűség tekintetében megállapítható, hogy összességében a 2021-es tárgyév mint vonatkozási időszak adatainak feldolgozása a t+8. hónapra döntő többségében megvalósult (hasonlóan az előző évekhez).

A feldolgozás elmaradásának további okairól a leggyakrabban azt a visszajelzést adták, hogy az előzetes adatok már rendelkezésre álltak, viszont az adatok teljes körű feldolgozása még nem fejeződött be, előzetes és utólagos ellenőrzések és adattisztítások voltak még folyamatban az érintett adatfelvételek vonatkozásában. Ez jellemzően nem késedelemből következett, hanem az eltérések az egyes adatfelvételeknél megfigyelhető adat-előállítási folyamat természetéből fakadtak.

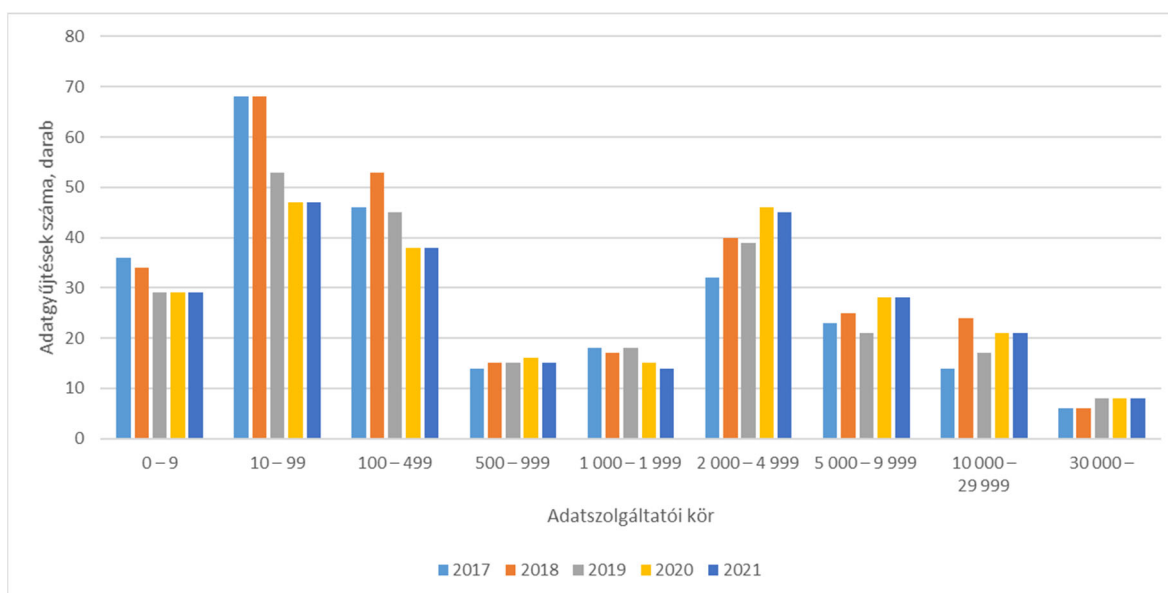
A HSSz-tagok által még fel nem dolgozott adatfelvételeket – ami az összes adatfelvétel 10%-át érintette (55 darab) – az 5. ábra mutatja be szakstatistikai téma szerint. Az adatfelvételek beérkezési határidejének különbözősége miatt jelen értékelésnél a teljesülési kérdőívek kitöltési határideje került figyelembe vételre a fel nem dolgozottság vizsgálatakor. A legnagyobb arányt a külkereskedelem és fizetési mérleg, valamint a környezetstatisztika témájú adatfelvételek képviselték, ugyanakkor ezen adatfelvételek esetében a feldolgozás a kialakított ütemterv szerint haladt, illetve a felvételek teljes körű jellegére vezethető vissza a pillanatnyi fel nem dolgozottság ténye. (A kijelölt összes adatszolgáltatótól szükséges összegyűjteni a kitöltött kérdőíveket. Ez okozhat a feldolgozásban időbeli elcsúszást.) A legtöbb esetben a kérdőív kitöltésekor az előzetes adatok publikációja már megtörtént.

5. ábra A még fel nem dolgozott adatfelvételek szakstatistikai témák szerinti alakulása, 2021

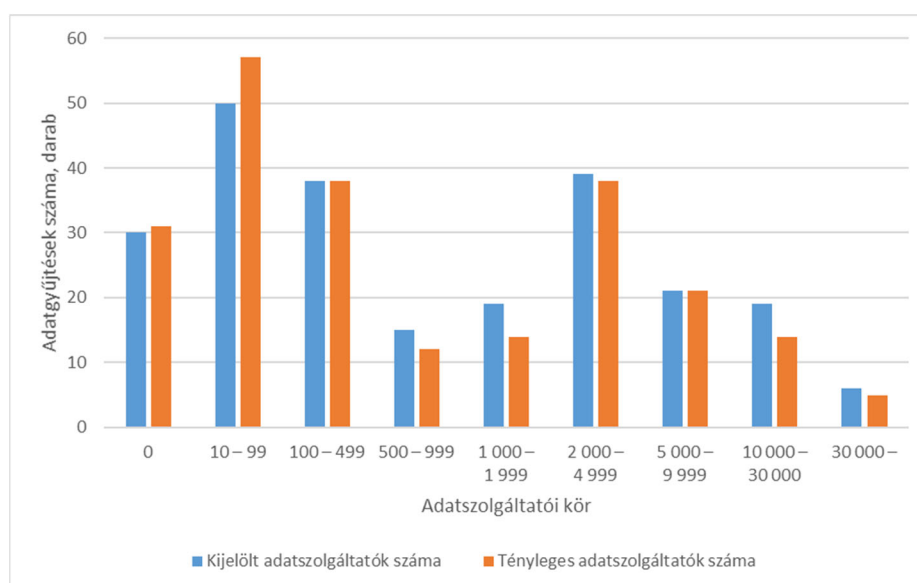


Az adatszolgáltatói terhek mérésére, azok megítélésére az értékelő kérdőívben két kérdésblokk szolgált. Az adatszolgáltatók tényleges számát célszerű külön vizsgálni az adatgyűjtések körében (6. és 7. ábra). Az adatátvételek körében az elemzés a teljesítendő feladatok számát, valamint a feladatok határidőre történő teljesítését, illetve nem teljesítését vizsgálta. Ennek oka, hogy az adatátvételek jellemzően adminisztratív forrásokból származnak, amelyek adatátadásoként több 10 ezer adathelyet is tartalmazhatnak, viszont ez az adminisztratív jelleg miatt nem jelent kifejezett terhet az adatszolgáltatóknak.

6. ábra Kijelölt adatszolgáltatók száma adatgyűjtések esetén



7. ábra Kijelölt és tényleges adatszolgáltatók száma adatgyűjtések esetén, 2021



Az adatgyűjtések adatszolgáltatói számának eloszlása ebben az évben is hasonlóan alakult, mint a megelőző években; kevéssel több mint fele (51%) legfeljebb 500 adatszolgáltatót érintett.

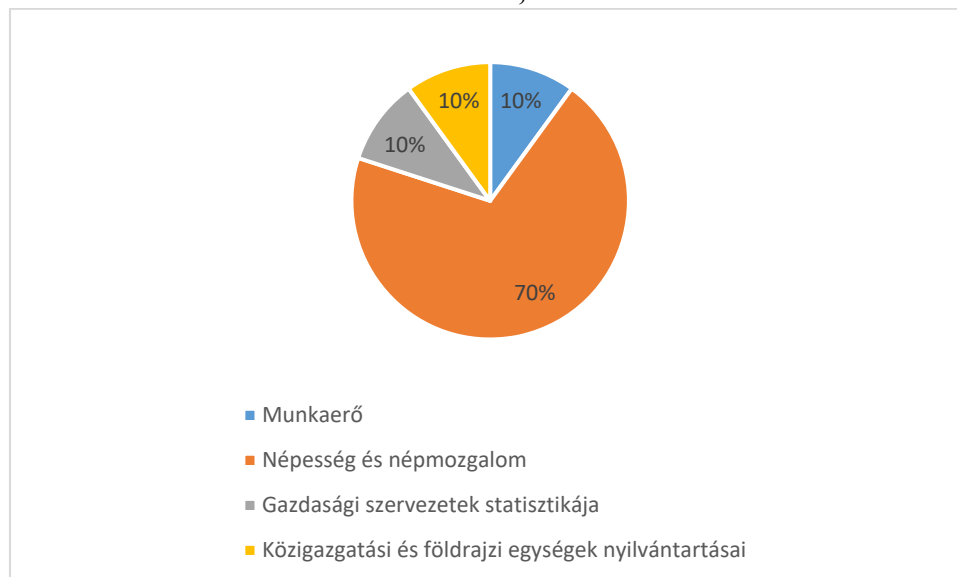
2021-ben a 10-nél kevesebb adatszolgáltatói kategóriában továbbra is a csővezetékes, a vízi, közúti, légi és vasúti szállítás, az infrastruktúra, valamint az egészségügyi adatokkal kapcsolatos adatgyűjtések voltak a legjellemzőbbek.

A 10–99 kijelölt adatszolgáltatóval rendelkező adatfelvételeknél sokkal szerteágazóbb témák fordultak elő, de legfőképpen itt is az ágazati gazdaságstatisztikai témákban végrehajtott adatgyűjtések voltak jellemzők (posta és távközlés, információs és kommunikációs eszközök felhasználása, elektronikus kereskedelem). Emellett az igazságszolgáltatás témában készült adatgyűjtések (bíróságok ügyforgalma, vádlottak, jogerősen elítéltek) is itt jelentek meg.

A következő csoportban (100–499 adatszolgáltató), hasonlóan a 2020-as évhez, a mezőgazdasághoz, gazdasághoz kapcsolódó adatgyűjtések dominálnak (állatállomány és állati termékek, halászat, mezőgazdasági számlarendszer), valamint ugyancsak találunk ebben a csoportban is társadalmi témákat, a kulturális intézményekre és tevékenységekre (könyvtár, levéltár, múzeum, színház, mozi, hangverseny, közművelődési intézmények), illetve a szórakoztatásra, a kultúrára (könyv- és lapkiadás, tömegkommunikáció, tv, rádió) vonatkozott.

A kijelölt és tényleges adatszolgáltatók számának eltéréséből látható, hogy az adatfelvételek többségében kevesebben válaszolnak az adott adatgyűjtésre, mint ahány adatszolgáltató ki van jelölve. Ennek az is lehet a magyarázata, hogy egyes adatszolgáltatók státuszában a megadott időszakban változás történik, a kijelölést követően valamilyen ok miatt megszűnik az adatszolgáltatási kötelezettsége. Néhány esetben az előző ellentétje is előfordulhat, vagyis a tényleges adatszolgáltatók száma magasabb, mint a kijelölteké (pl. az 1654-es, az 1824-es, az 1846-os, az 1852-es, a 2009-es számú adatgyűjtés). A szakstatisztikák szintjén ugyanakkor nincs szignifikáns különbség a nemválaszolási mintázatok között, egyenletesen érintett minden felvétel.

8. ábra A kizárólag postai úton érkező adatgyűjtések megoszlása szakstatisztikai téma szerint, 2021



A 8. ábrán a kizárólag postai úton érkező adatfelvételeket láthatjuk szakstatisztikai bontásban, ami 2021-ben 10 darab adatgyűjtést érintett – ez némi csökkenést mutat a 2020-as értékhez képest. Az adatgyűjtések jelentős része (70%-a) a népmozgalmi statisztikák közé tartozik.

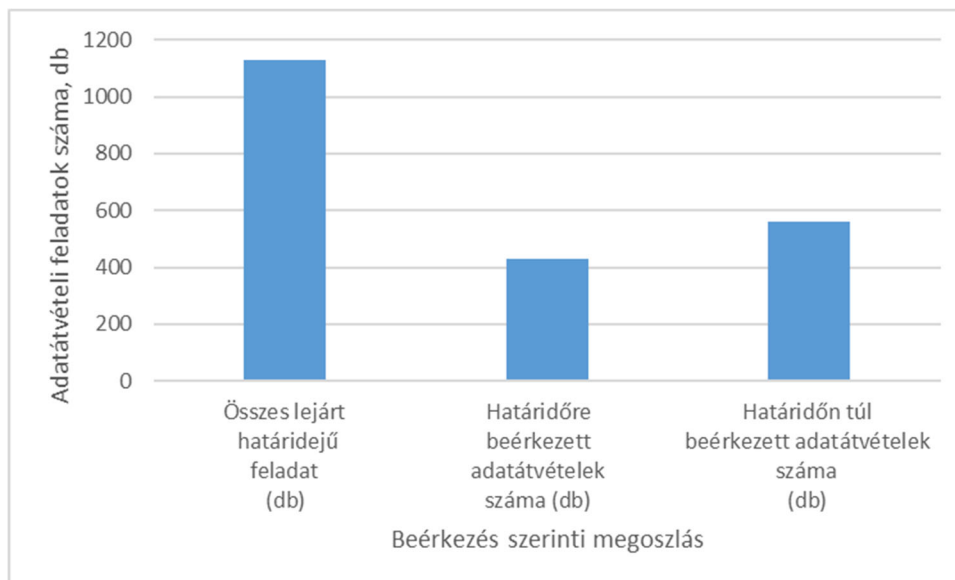
Adatfelvételenként megvizsgáltuk, hány darab érkezett be 2021-ben határidőre, határidőn túl vagy egyáltalán nem. A 2. táblában látható, hogy az adatátvételek adatátadási feladatai 2906 esetben határidőn belül beérkeztek a HSSz elrendelő szerveihez.

Adatfelvételi szinten az összes adatátvétel 99%-a határidőre beérkezett, ami emelkedést mutat a 2020-as 94%-os arányhoz képest, emellett mindössze 1% azon adatfelvételek aránya, amelyeknél volt olyan adatátadási feladat, ami nem érkezett be határidőre a HSSz szerveihez. Elenyésző azon adatátvételek száma, amelyeket egyáltalán nem teljesítettek.

2. tábla Adatátvételek évenkénti esedékesség szerinti alakulása, 2021

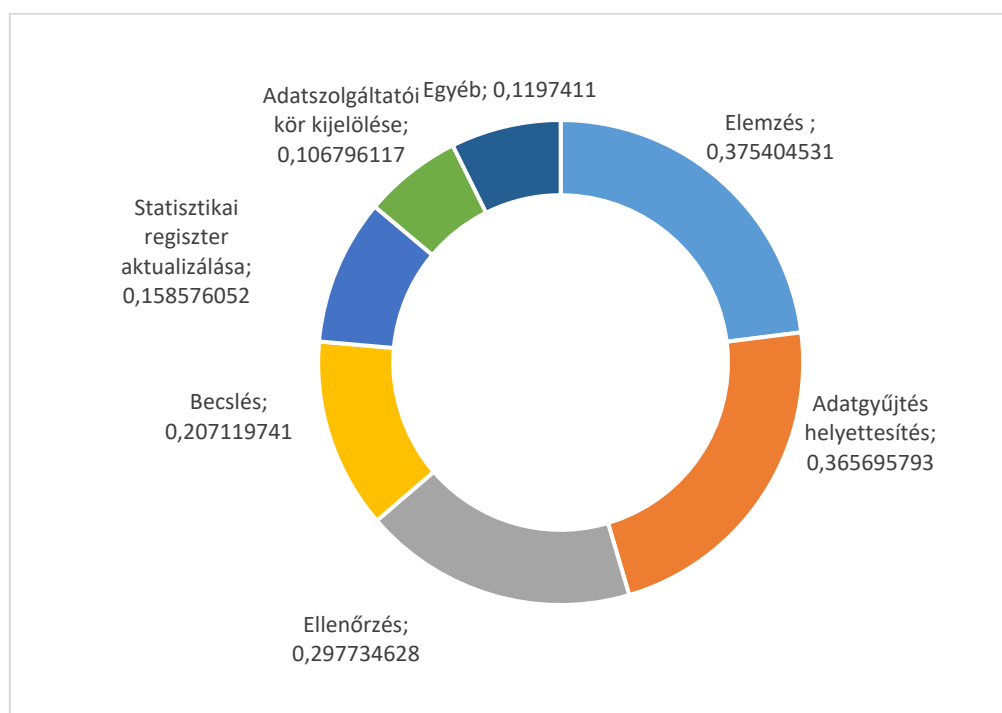
Megnevezés	HSSz	KSH	Összesen
Határidőre teljesített	105	2801	2906
Határidőre nem teljesített	3	24	27
Egyáltalán nem teljesített	0	5	5
Teljesítendő feladatok száma összesen	108	2842	2950

9. ábra A KSH KARÁT-rendszerében teljesített adatátvételi feladatok számai, 2021



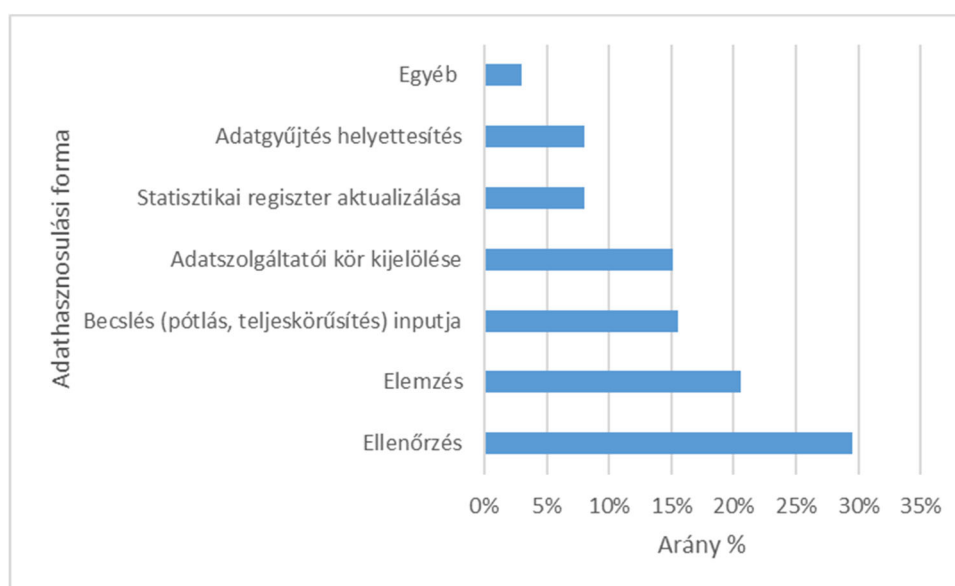
A 9. ábrán láthatók a KSH elrendelésében lévő adatátvételek teljesítésére vonatkozó információk a KSH KARÁT-rendszeréből. A KARÁT-rendszerben az összes beérkezett adatátvételek (429 darab) aránya a lejárt határidejű adatátvételi feladatokhoz viszonyítva 88% volt.

10. ábra Adminisztratív adatok felhasználásának célja az összes adatátvétel számához viszonyítva, 2021



A 10. ábrán az adatátvételek felhasználásának céljait láthatjuk a HSSz-szervek által elrendelt adatfelvételek esetében. A százalékos értékek az ábrán az egyes felhasználási célok tekintetében az összes adatátvétel számához viszonyítva mutatják az adott felhasználási cél megoszlását. Egy adatátvétel esetében több választ is megjelölhettek a kollégák. Jellemzően nagy arányú az elemzési célú felhasználása az adatátvételeknek, emellett az adatgyűjtés-helyettesítésre az adatátvételek hasonlóan nagy arányát használják a HSSz-ben, ezt követte az ellenőrzési célú felhasználás. Jellemző még a becslésre és pótlásra történő alkalmazás, valamint a statisztikai regiszter aktualizálása is magas arányban.

11. ábra Az egyéb adatforrásból származó adatok hasznosulása, 2021



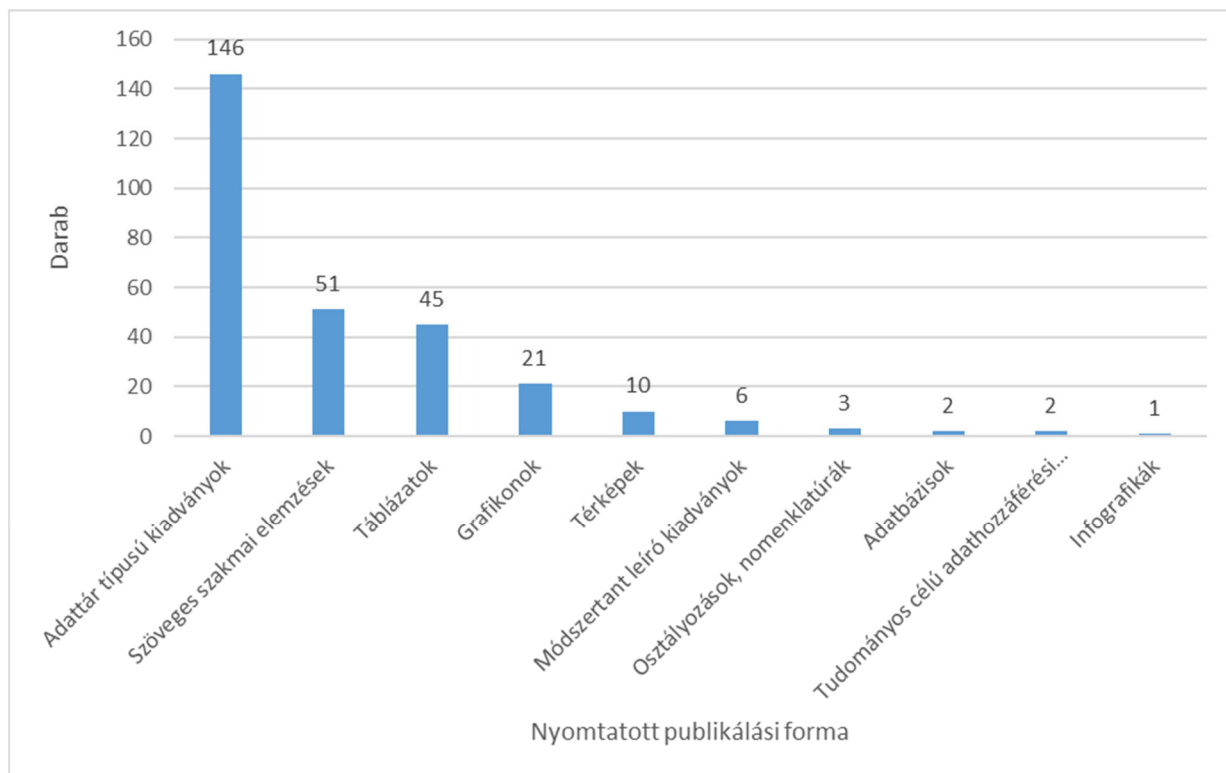
2021-ben a kérdésre érkező válaszok alapján összesen 692 esetben² jelölték azt a kollégák, hogy egyéb adatforrásokat is felhasználnak az érintett adatfelvétel feldolgozási folyamatában (lásd 11. ábra). Az adatfelvételek 33%-ához nem használtak egyéb adatforrásból származó adatokat a HSSz tagjai.

Az egyéb felhasználási cél válaszlehetőségeire jellemzően a súlyozást, súlyképzést, illetve a validálást jelölték meg a kollégák.

Kérdőívünk a publikálás három formáját vizsgálta (nyomtatott, online és egyéb). A begyűjtött adatokból az adat-előállítási folyamat után kapott eredmények publikálása legnagyobb részben – a tavalyi elemzésben bemutatottakhoz hasonlóan – online felületeken történik. A kapott eredmények alapján a nyomtatott és az egyéb felületeken megjelenő tartalmak az online publikációk 20%-át teszik ki. A nyomtatott formátumok között kiemelkedő helyen vannak az évkönyvek, de jelentős még az aggregált adatokat tartalmazó publikációk és az elemzések szerepe is.

²A kérdőívén ennél a kérdésnél egy adatfelvétel esetén több lehetőség közül lehetett választani.

12. ábra A statisztikai termékek nyomtatott formában történő megjelenései, 2021*

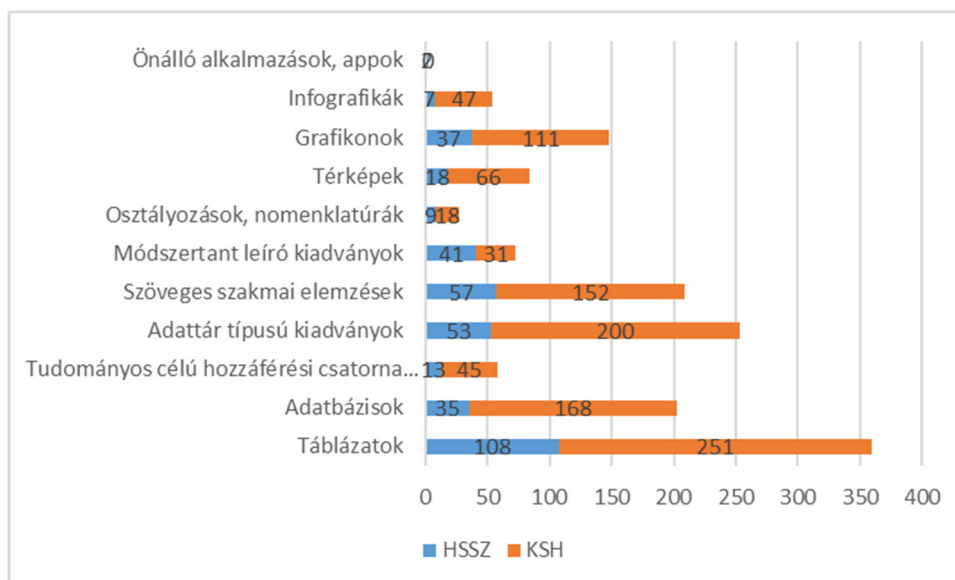


* Nyomtatott publikálási formák: adattár típusú kiadványok, szöveges szakmai elemzések, táblák, grafikonok, térképek, módszertant leíró kiadványok, osztályozások, nomenklatúrák, adatbázisok, tudományos célú adathozzáférési csatornákon kutatható állományok, infografikák, önálló alkalmazások, appok.

Az adattár típusú kiadványok publikálása a legjellemzőbb nyomtatott publikálási forma, ezt követte a szöveges szakmai elemzések megjelenítés.

A következőkben a beérkezett kérdőívek összesített eredményei alapján az egyes kategóriák online megjelenéseit ábrázoljuk.

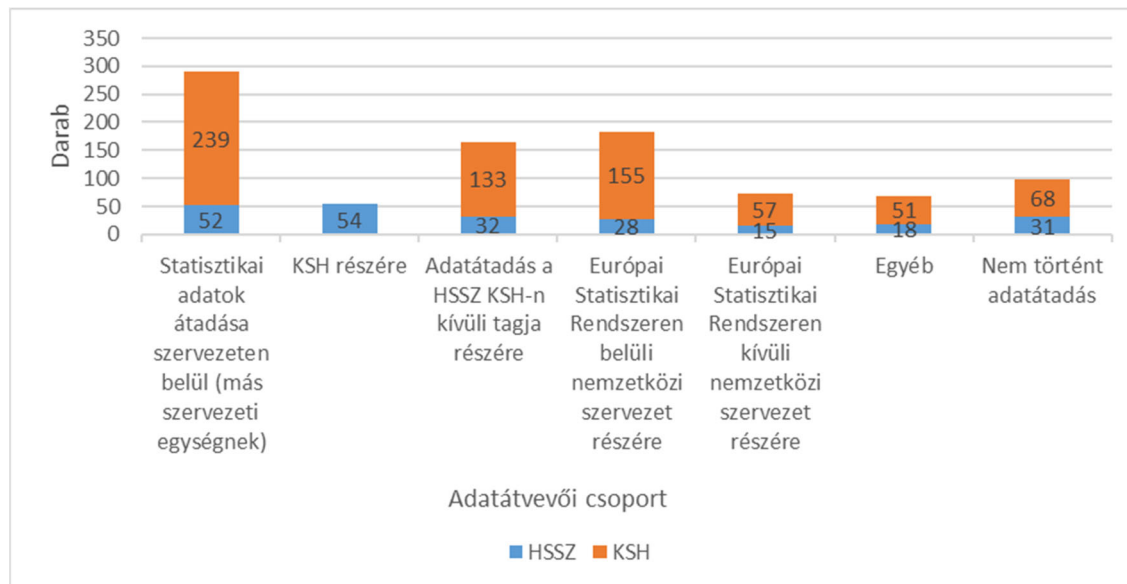
13. ábra Az online előállított statisztikai termékek publikálásának leggyakoribb módjai, 2021



Az online megjelenítés esetében érkezett válaszok az előző évhez hasonlóan alakultak. Jellemzően online is táblás adatokat publikálnak, de az adattár típusú kiadványok és az adatbázisok is elterjedtek.

A vizualizációs eszközök közül az online felületeken való megjelenést tekintve továbbra is vezető helyen van a grafikonos megoldásokat használó publikálási forma, ezt követte a térképes megjelenítés. Infografikát 54 esetben alkalmaztak az adatok megjelenítéséhez.

14. ábra A statisztikai célú * adatátadások átvevői, 2021



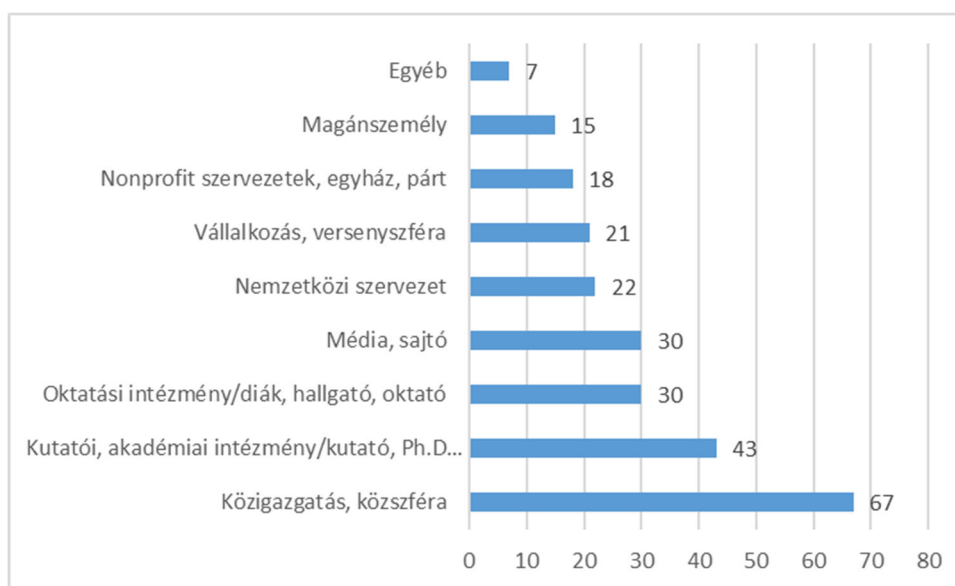
* Statisztikai cél: hivatalos statisztikai adat statisztikai adat-előállítási folyamatban való felhasználása.

A statisztikai célú adatátadások átvevőinek szám szerinti megoszlását mind a KSH, mind a HSSZ KSH-n kívüli tagjai által a kérdőívre adott válaszaikból összesítettük (14. ábra). A leggyakoribb kategória a korábbi évekhez hasonlóan a szervezeten belüli átadások volt, ennek száma (291) ez az érték nőtt a 2020. évihez (250) képest. Az ezt követő kategóriák, a HSSZ-tagok részére történő átadások, illetve az Európai Statisztikai Rendszer tagjai, valamint az Európai Unió Bizottságának szervei részére történő átadások (165, illetve 183) a 2020. évihez (158, illetve 168) ugyancsak hasonlóan egymáshoz közeli értékeket mutatnak, azonban az elmondható, hogy 2021-ben összeségében nőtt az átadások száma. Az adatok összesítéséből az is kiderült, hogy a kitöltött kérdőívek 89%-ában történt adatátadás az adatátvételek/adatgyűjtések feldolgozott adataiból.

A tudományos vagy tájékoztatási célú adatkiadásokra vonatkozó kérdés idén is csak a HSSZ KSH-n kívüli tagjainak kérdőívében szerepelt, az adatok összesítése után az alábbi eredményt kaptuk (15. ábra).

Az ábrán kiemelkedik a közigazgatás kategória, ami alapján elmondható, hogy a feldolgozott adatokkal kapcsolatos adatkiadások a leggyakrabban a minisztériumok és egyéb közigazgatási szervek részére teljesülnek, vagyis az adatokat kiterjedten a kormányzati döntéshozatalban hasznosítják.

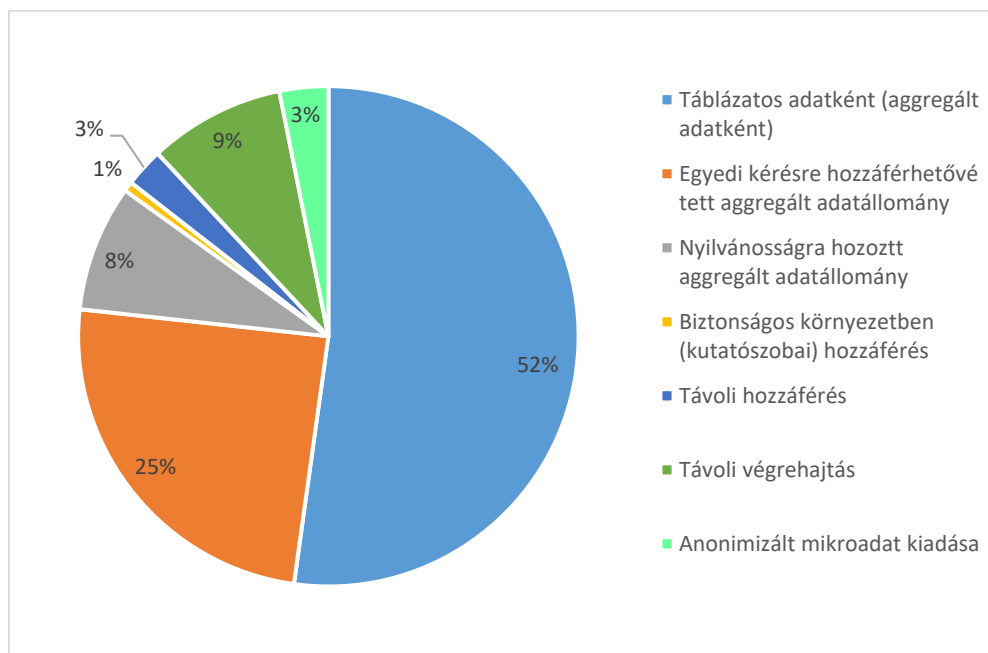
15. ábra Tudományos vagy tájékoztatási célú adatkiadások átvevői a HSSz további szerveinél, 2021*



* Kutató: olyan természetes személy, aki tudományos kutatási célból, kutatással foglalkozó intézménnyel való együttműködésben vagy anélkül kíván hozzáférni statisztikai adatokhoz, ideértve a közoktatásban, illetve a felsőoktatásban részt vevő hallgatókat, diákokat is. Egyéb: azon felhasználói csoport megjelölése, amelyeket a fentiekben nem jelöl a kérdőív.

A HSSz esetében vizsgált adatkiadások a megvalósulás formáját tekintve az alábbiak szerint alakultak.

16. ábra Adatátadások és adatkiadások megvalósulási formái a HSSz KSH-n kívüli tagjainál, minden felhasználói csoportot figyelembe véve, 2021



Az adatkiadások több mint fele (52%-a) táblás adatként, 25%-a egyedi kérésre összeállított táblás formában valósult meg. Ebből látható, hogy az összes adatkiadás 77%-a táblákba rendezve jelenik meg, ami 1 százalékpontos növekedés a 2020. évi 76%-hoz képest. Az adatkiadások mindössze 16%-át adják ki az egyéb módozatok (kutatószobai hozzáférés, távoli hozzáférés/végrehajtás, anonimizált mikroadat³ kiadása). A táblás adatok kiadásának százalékos aránya az elmúlt években így alakult: 2017-ben 60, 2018-ban 53, 2019-ben 90, 2020-ban 53, 2021-ben 52%, tehát idén újra kisebb részben volt jelen. Az egyedi kérésre összeállított adatok 2017-ben 30, 2018-ban 31, 2019-ben 35, 2020-ban 23, 2021-ben 25%-ot tettek ki.

Ezek alapján elmondható, hogy 2021-ben a két kategória – egyedi kérésre összeállított táblás adatok, illetve táblás adatként (aggregált adatként) – összevont aránya hasonló volt az előző évihez. A kérdőívek idén is tartalmazták az adatok kezelése során felmerülő adatvédelmi megoldásokat tartalmazó kérdést.

³ Mikroadat: olyan, rekordok sorozatából álló állomány, amely megfigyelési egységek adatait tartalmazza. Az aggregált adatok előállításának elsődleges forrása. Anonimizált mikroadat: olyan mikroadat, amelyeket – megfelelő felfedés elleni védelmi módszerek alkalmazásával – annak érdekében módosítottak, hogy az azonosításhoz használható legjobb eljárásokkal összhangban minimális mértékűre csökkenjen az érintett statisztikai egységek felfedési kockázata. Az így létrejött adatállományok már nem hozhatók kapcsolatba az adatszolgáltatókkal, ezért – megfelelő szerződéses garanciák mellett – átadhatók az adatkérők részére.

A **jogi védelem**⁴ terén elmondható, hogy a HSSz KSH-n kívüli tagjai rendelkeznek az általános adatvédelmi rendelkezéseiken túl kifejezetten a statisztikai adatvédelemről rendelkező belső szabályzatokkal, titoktartási nyilatkozatokkal.

A **fizikai védelem**⁵ esetében az alábbiak voltak a leggyakrabban említett eszközök:

- kártyás belépőrendszer, kamerarendszer, portaszolgálat;
- az adatlapok zárható termekben, szekrényekben történő tárolása;
- jogosultság alapú rendszer, jelszavas védelem;
- az adatokról biztonsági másolat készül.

A **statisztikai felfedés elleni védelem**⁶ területén a leggyakrabban alkalmazott megoldások:

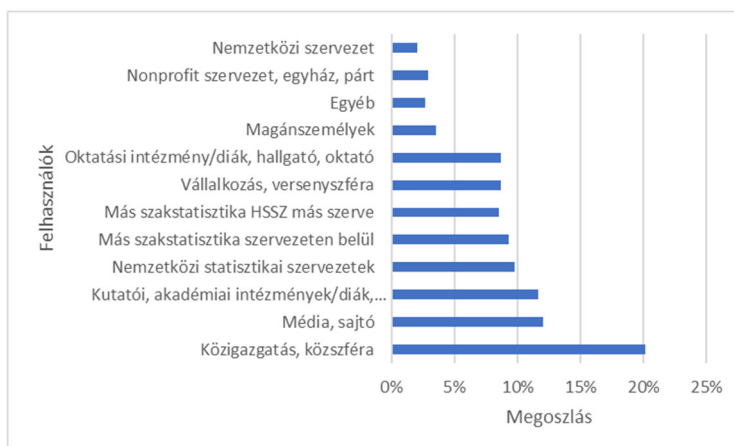
- az adatok aggregált kezelése;
- az ún. „hármasszabály”, dominanciaszabály, elsődleges és másodlagos cellaelnyomás.

További eszközöket is megjelöltek a válaszokban: anonim feldolgozás, egyedi kódolás (egyedi kulcs alkalmazása).

Az előállított statisztikai adatállományok jellemző felhasználói

A gyűjtött/átvett adatokból előállított statisztikai adatállományok felhasználói oldalát tekintve a következő következtetések vonhatók le az alábbi kördiagram (17. ábra) segítségével.

17. ábra Az előállított statisztikai adatállományok jellemző felhasználói a HSSz-nek a KSH-n kívüli tagjai esetében, 2021



⁴ Jogi adatvédelem: olyan jogi adatvédelmi eszközök összessége, amelyek egyértelműen meghatározzák az érintettek jogait, kötelezettségeit és felelősségét (pl. szerződéskötés adatkérővel, felhasználási feltételek kikötése, titoktartási nyilatkozatok kikötése stb.).

⁵ Fizikai adatvédelem: olyan adatvédelmi megoldások összessége, melyek célja az adatok külső, fizikai behatásoktól, veszélyektől való megóvása, valamint az eszközök és adatok közelébe jutásának megnehezítése (pl. adatokat tároló szerverek szeparált helyiségben való tárolása, ahová a belépés szigorúan ellenőrzött, biztonsági másolatok készítése, jogosultságkezelési rendszer bevezetése, jelszavak, azonosítási metódusok stb.).

⁶ Statisztikai felfedés elleni védelem: olyan, statisztikai célra összeállított adatállományok módosítását célzó módszerek összessége, amelyek a lehető legnagyobb mértékben csökkentik annak kockázatát, hogy valamely statisztikai egységet vagy rá vonatkozó új információt beazonosítsunk az állomány alapján (pl. mikroadatokat esetén azonosításhoz/felfedéshez szükséges kulcsváltozók meghatározása, és azok alapján a kulcsokra való találatok mesterséges növelése, pl. zajosítás, rekordcsere, aggregálás, kerekítés, alsó-felső kódolás stb.), táblás adatok esetén elsődlegesen elnyomandó cellák vizsgálata (pl. „hármasszabály”, dominanciaszabály stb.), illetve másodlagos cellaelnyomás szükségességének vizsgálata, és szükség esetén másodlagos cellaelnyomási akciók elvégzése.

A felhasználók ötödét a közigazgatás adta (2020-ban 20%). A második legjelentősebb kategória a HSSz-tagoknál mért eredmények alapján a kutatók, illetve a média, sajtó (egyaránt 12%). Ezeket követte a nemzetközi statisztikai szervek (10%). Ezzel szemben a legkisebb arányt a magánszemélyek, a nonprofit szervezetek, az egyéb felhasználók (mindhárom kategória 3%-ot tett ki) és a nemzetközi szervezetek (2%) jelentették. Az adatokban kisebb eltolódások láthatók a 2020. évi adatokhoz képest.

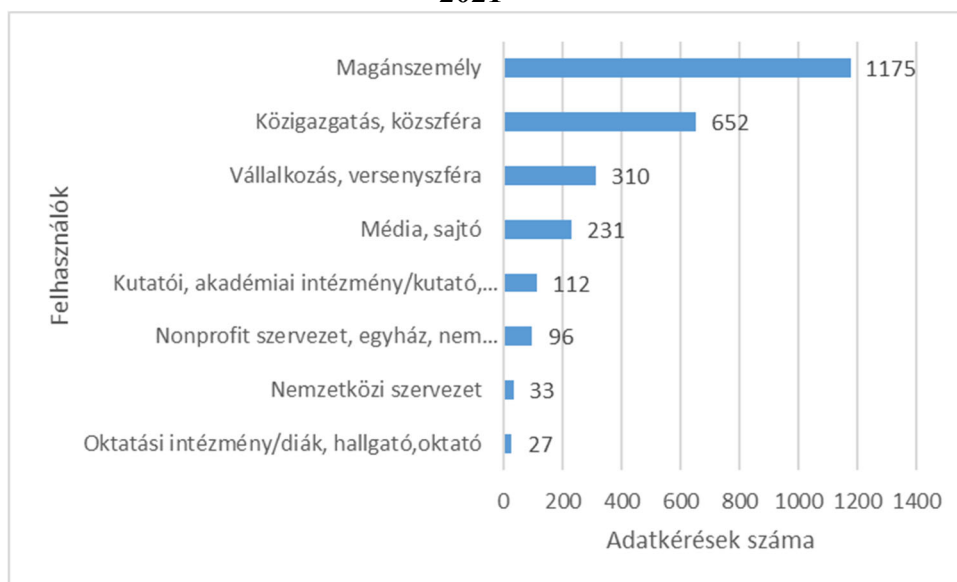
A felhasználók az egyéb kategóriáján belül jellemzően olyan további szakmai szervezeteket jelöltek meg, amelyekkel szorosan együttműködnek.

A KSH „Lépjen velünk kapcsolatba!” rendszerében regisztrált adatkérésekre részletes statisztika áll rendelkezésre 2021-re vonatkozóan.

A következőkben röviden áttekintjük a felhasználók szerinti, a leggyakrabban érintett SzOR (Szakstatisztikák Osztályozási Rendszere) -elemek szerinti, illetve a 2017–2021 közötti összesített adatkérések számának alakulását.

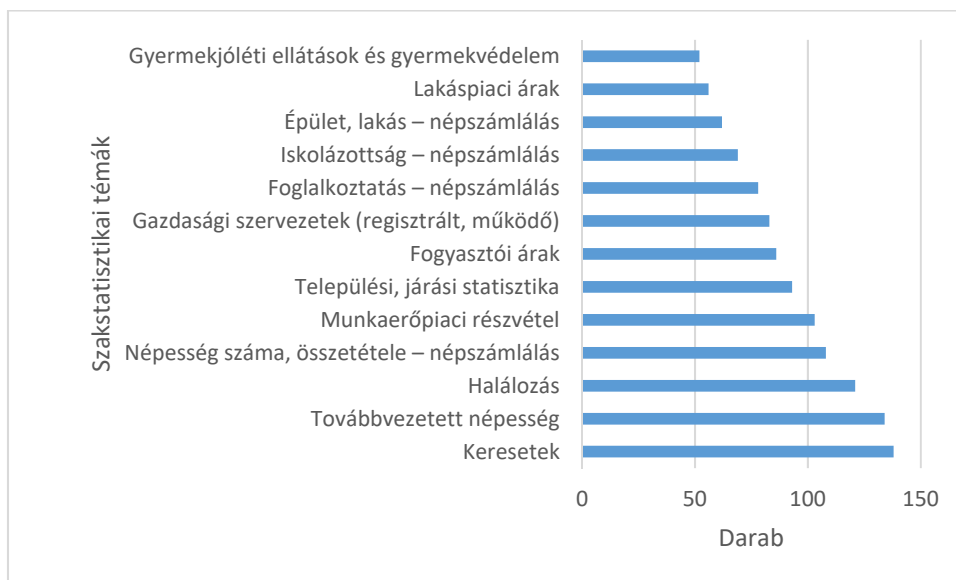
Az adatkérések felhasználói csoportok szerinti megoszlása a következő diagramon (18. ábra) látható. 2021-ben összesen 2636 esetben érkezett a KSH felületén keresztül adatkérés. Ezek közül a magánszemélyek (1175)⁷, a közigazgatás, közszféra (652) és a vállalkozások, versenyszféra (310) területéről érkezett kérések száma emelkedett ki. 2020-ban az első három helyen szintén ez a három kategória állt.

18. ábra A KSH által teljesített adatkérések, felhasználói csoportok szerint, 2021



⁷ A vállalkozások gyakran magánszemélyként nyújtják be a kérdéseiket, ezért olyan magas a magánszemélyek száma.

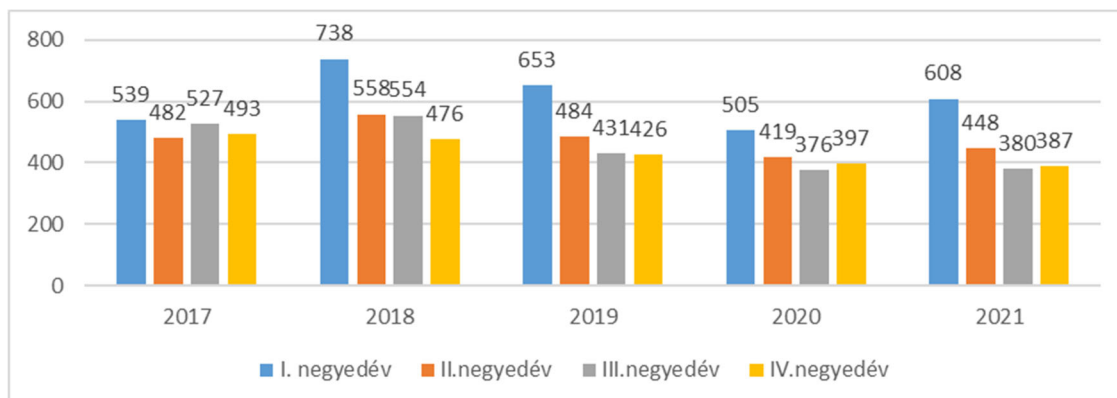
19. ábra A KSH által teljesített adatkérések a legkeresettebb SzOR-témák szerint, 2021



A legkeresettebb elemeket a 2021. évi négy negyedév összesítésével a 19. ábra tartalmazza. Mind közül kiemelkedő a keresetek kategóriája, amely negyedévenkénti bontásban is a legnépszerűbb volt. Az ábrán látható, kiemelt kategóriákon kívül sok adatkérés érkezett még a továbbvezetett népesség, a népesség száma, összetétele – népszámlálás, a foglalkoztatás – népszámlálás, a foglalkoztatás – a népszámlálás és a halálozás elnevezésű SzOR-elemekkel kapcsolatban is.

A 20., az adatkérésekkel kapcsolatos ábrán az elmúlt 3 évben beérkezett adatok idősoros ábrázolása látható. Az adatkérések száma jellemzően az I. negyedévben volt a legmagasabb és a IV. negyedévben a legkisebb. A 2021. évi adatok alakulása követte ezt a trendet.

20. ábra A KSH által teljesített adatkérések negyedévek szerint



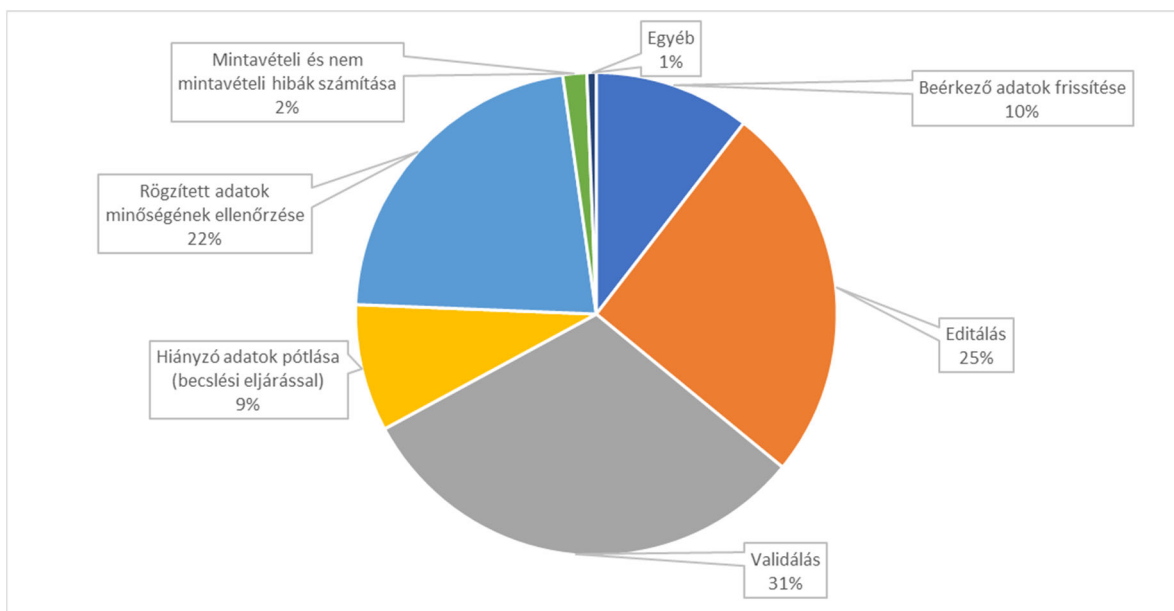
Minőségjelentések

A HSSz-tagok által kitöltött kérdőívek összesítéséből kiderült, hogy 104 esetben nem készítettek nemzetközi minőségjelentést. A kitöltők 31 esetben jelölték meg a „szervezetten belüli minőségirányítási folyamaton keresztül” lehetőséget, 12 esetben a nemzetközi szervezeteknek küldték meg jelentéseiket. Az egyéb esetet 11 kérdőíven jelölték meg a kitöltők. Ezeknek a döntő többsége a DESAP-kérdőív (a „Development of a Self Assessment Programme” projekt keretében, az Eurostat és a tagállamok által kialakított adatfelvételek önértékelési kérdőíve) kitöltését adta meg válaszként.

A KSH által elrendelt, 2021-ben megvalósult adatfelvételeknél (415 adatfelvétel) a kitöltött válaszok alapján elmondható, hogy 191 esetben készült minőségjelentés. A KSH belső gyakorlata szerint szakstatisztikánként készül minőségjelentés és nem adatfelvételi szinten, ugyanakkor időközben elindult a minőség mérésére vonatkozó belső fejlesztésű rendszer (DOQ), amelyben a 2022-es vonatkozási évre már elkészültek a minőségjelentések. A KSH-n belül vagy azért nem készítettek belső minőségjelentést, mert az uniós kötelezettségként is elő van írva ezekre az adatfelvételekre, és annak keretében már teljesítették (115 esetben) vagy az adatfelvétel jellegéből adódóan nem is készült eleve minőségjelentés.

A minőség javítására tett intézkedések megoszlása a következő ábrán (21. ábra) látható, ez a kérdés kizárólag a HSSz KSH-n kívüli tagjai esetében szerepelt a kérdőíven.

21. ábra A minőség javítására tett intézkedések megoszlása, 2021



Ennek alapján a minőség javítása érdekében megtett leggyakoribb intézkedés a 2021. évi adatfelvételekben a *validálás*⁸ (2020-ban 24%, 2019-ben 25%) és az *editálás*⁹ (2020-ban 29%,

⁸ Validálás: az aggregált adatok összehasonlítása, ellenőrzése, javítása azonos, illetve más (statisztikai felvételek, mérlegek stb.) adminisztratív forrásokból származó adatok felhasználásával.

⁹ Editálás: olyan tevékenység, amelynek eredményeként a validálás során megjelölt/azonosított hibás adatok javításra kerülnek, vagy törölődnek és pótlásra jelölődnek.

2019-ben 29%) volt. Előbbi értéke emelkedést mutatott, míg az *adatok frissítése* a tavalyihoz képest csökkent (2020-ban 16%, 2019-ben 15%).

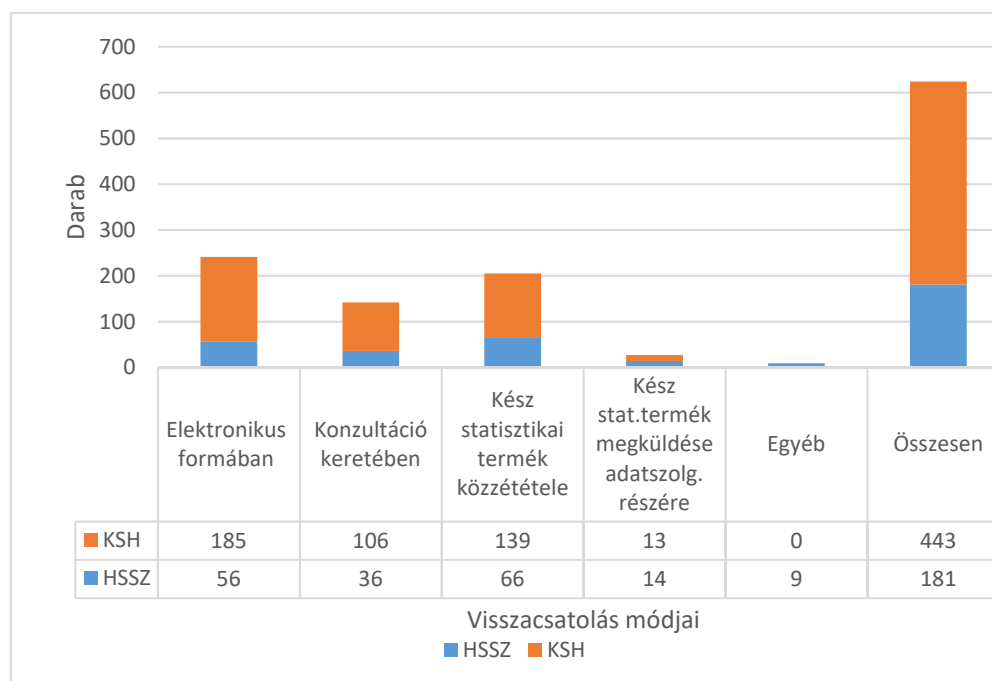
Az előre definiált kategóriákon kívül minőségjavítási eszközként merült még fel

- az adatellenőrzés, adatjavítás feladatainak biztosítása, felügyelete,
- a hiánypótlási kérés, felszólítás küldése,
- a be nem érkezett adatok sürgetése.

Visszacsatolás

2021-ben ismét megtörtént a visszacsatolások módjának mérése. 397 kérdőíven jelölték be, hogy különböző formákban történnek visszacsatolások az adatszolgáltatóknak, ez az összes adatfelvétel 72%-a (22. ábra). A legjellemzőbben ez elektronikus formában (43%) zajlott, ezt követte a kész statisztikai termék megküldése, és az esetek közel negyedében konzultáció történt az adatszolgáltató partnerekkel.

22. ábra A visszacsatolások módjainak megoszlása, 2021



Ha megvizsgáljuk a kérdést a KSH és a HSSz további szervei szerinti bontásban is, láthatjuk a különbségeket: míg a KSH-nál nagy számban volt elektronikus visszajelzés, addig a HSSz egyéb tagjainál a visszacsatolás magasabb számban történt a kész statisztikai termék megküldésével.

Budapest, 2022. november

KSH Statisztikai koordinációs és jogi főosztály