



Adattárház, RADAR


Egy kockázatelemző szemszögéből

Vikárus Gabriella
osztályvezető, projektvezető (2001-2013)

Budapest, 2023. október 13.



Miről lesz szó?

- 90-es évek - SAS alapú központi adattár
- 2005 - 2006 - ATAR - Adattárház Oracle alapokon
- 2009 -  adattárház (Oracle platform, SAS analitika)
- Eredmények
 - Kiválasztás országos állományokon
 - ÁFA csalás becslése mesterséges intelligencia algoritmusokkal
 - Hálózatelemzés, hálózatok sorrendezése, tömeges szűrése, vizualizálása



90-es évek közepe - SAS alapú központi adattár

➤ Adattárház elődje, forrásadatok a helyi rendszerekből fájl szinten kerültek átvételre

- **Bevallások, adófolyószámla, adóalanyi törzs**
 - Bevallás mezők META kódolása - évfüggetlen

➤ **Cél: külső felhasználás**

- Központi adatszolgáltatások
- PM adótervezési munkájának támogatása
- Házi fejlesztésű, SAS alapú önkiszolgáló leválogató eszköz (Lakó Ferenc)
 - használata szűk körben

➤ **Központi Hivatalban és a PM-ben**

- Csak az ellenőrzések típusa, darabszáma és a feltárt adókülönbözetek nagysága ismert
 - ún. VIR táblák
- A vizsgált adózók, ezek jellemzői, a vizsgált adónemek és időszakok, a feltárt szabálytalanságok típusai nem ismertek

➤ **Ellenőrzésre történő kiválasztás**

- Megyei igazgatóságokon, a helyi bevallási adatbázisokból

➤ **Adónem- és bevallás centrikus szűrések**

- Fókusz egy-egy időszaki bevallás, egy-egy adónem
- Ellenőrzési vagy tervezési szakterület egyedi igénye alapján a helyi informatikusok leválogatásai
- Igazgatósági fejlesztés (Borsod)
 - 2001-ben országosan is elterjesztettük
- Kiutalást kérő bevallásokra
 - Központi fejlesztésű, de a helyi osztott adatbázisokon működő paraméterezzhető szűrő rendszer

➤ **Legkockázatosabb adózók nemigen kerültek látókörbe**

- Megyek között vándorlók
- Bevallást be nem nyújtók
- Bevallott adót be nem fizetők



2000: Központi kockázatkezelési koncepció (Söllei János)

➤ Holisztikus megközelítés

- Az adózókról széles körből, országos szinten gyűjtött adatok
- Az adózói populáció nagyság és kockázati szempontú szegmentálása

➤ 2001: Kockázatkezelési osztályok létrehozása

- VAK rendszer, Gyorsított ÁFA Kiutalási rendszer (GYÁK) – utóbbi a SAS adattárnak köszönhetően 2 hónap alatt !!!
- Egységes mérőszámon alapuló ügyfélérték számítás
 - **Ellenőrzött időszakok, adónemek és megállapított adókülönbözetek** SAS adattárba integrálása – célváltozó (Koczkáné Korompay Katalin, Székely Józsefné)
 - 17 lehetséges mérőszám és a megállapított adókülönbözetek összevetése Spearman-féle rangkorrelációval (Raig Mária Magdolna)

➤ Kockázati mutatók bővítése a SAS adattárban (400 db)

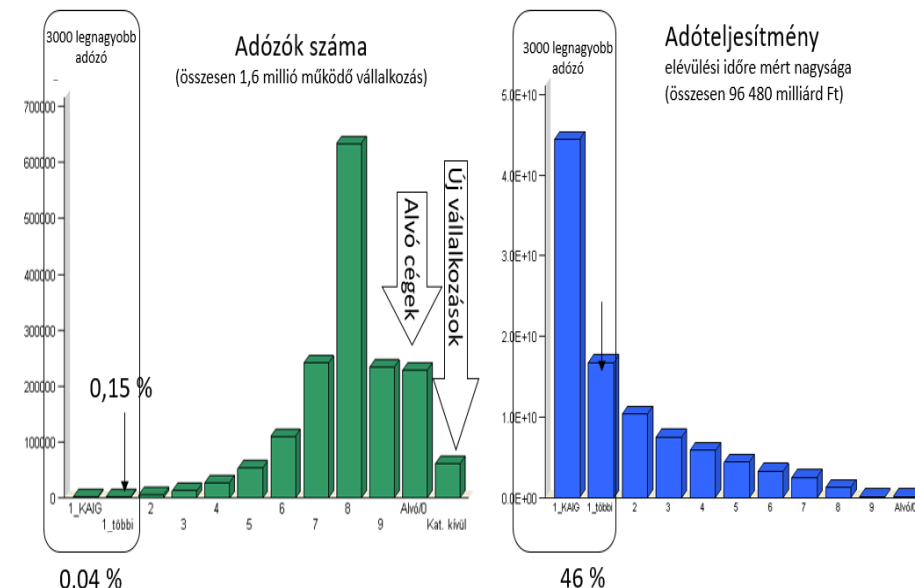
- Az igazgatóságok által jelzett revizori tapasztalatok alapján
- 2003: Központi Oracle Forms szűrő felület (KAT rendszer)

➤ 2003: Adóelkerülés eseteinek kódolt rögzítése

- R+ és KKR Oracle alapú rendszerek (pl. karusszel csalás szerepköreit) – célváltozók (Szentés Boldizsár, Udvardy Magdolna)

➤ Központi megrendelések - specifikációk, tesztelések

Működő adószámos adózók megoszlása



- Adóteljesítmény mutató - adóeljárasi tv.
- Adózási módtól, vállalkozási formától független nagyság szerinti összehasonlítás lehetősége
- Ellenőrzött/még vizsgálható adóteljesítmény

Ellenőrzési igazgató-helyettes megjegyzése:

- Ezentúl látjátok majd a Központban, hogy mit teszünk az igazgatóságokon!



2005-2006 Oracle adattárház (ATAR)

dr. Futó Iván
Szilágyiné Debreceni Erika
Fejes Erzsébet
Szentés Boldizsár

➤ SAS adattárban elérhető adatok és riportok

○ Adózási életút mutatóink újbóli megképzése

- Az új környezet sajátosságai okán (Parditkáné Jakubovics Erzsébet)

➤ Adattárház Portál a felhasználóknak

➤ További adatkörök integrálása 1TB → 40 TB

○ Kontroll adatok (pl. EU-s forgalom ütköztetése, VISKO, JAROK)

○ Megyei revíziós rendszerek (Kovács Attila), R+, KKR rendszerek

- Mely adózó, mely adónem, mely időszakát vizsgálták, melyekre tettek és milyen okból megállapítást
- Célváltozók a gépi tanuláshoz és további kockázati mutatók
- Pl. az adózó csalási láncolatban betöltött szerepe kódok

➤ 2008: Integrációs és Kiaknázó adatbázisok

➤ 2006: vagyonosodási vizsgálatra kiválasztást támogató adatpiacok - 3 adatkör együtt

- Pl. cégtulajdonosok szja bevallásaiban vallott jövedelmek és a társasági adóbevallásaikban jelzett tagi kölcsönök
- Még ez is megrendelés és IT fejlesztés

➤ Önkiszolgáló analitika

ORACLE®

DISCOVERER

➤ Minden adatot elér mindenki

- Igazgatóságok és központ is
- Országos adatok

➤ Lekérdező „munkafüzetek” önálló kidolgozása

- Országos elérhetővé tétele, oktatása

➤ Technikai korlátok

- Mely táblákat kapcsolt össze az IT és hogyan
- Műveleti sorrend nem befolyásolható

➤ Szűrés egy-két adattáblából, majd Excel export

- További helyi feldolgozás (FKERES)
- Központban erről mit sem láttunk

➤ Adattárház adattábláinak megismerése

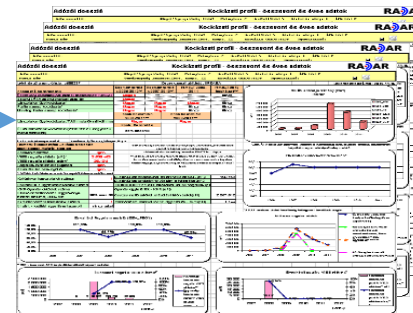


2008. XI. – 2009. XI. RAAR adattárház

Varga Lászlóné
Söllei János

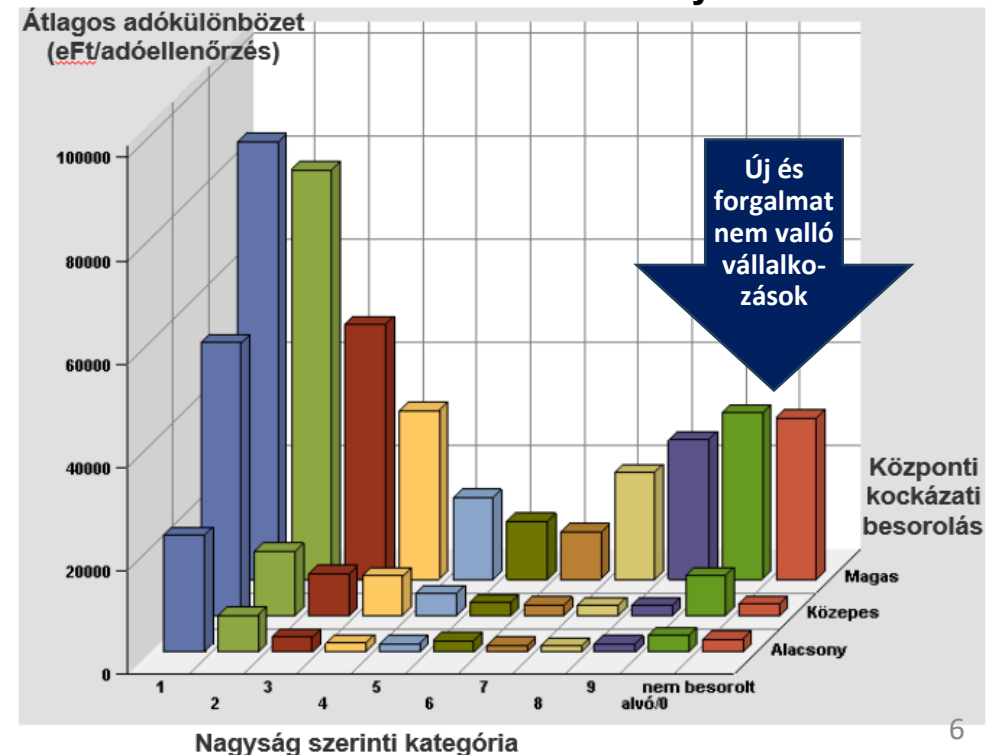
Rugalmas Adóellenőrzési Döntéstámogató és Analitikai Rendszer

- 2. adattárház, szintén Oracle platformon
 - ATAR adattárház, Tranzakciós rendszerek (ÜZEM) adataiból
- Adózói dosszié - 14 db Excel munkalap
 - Egy-egy adózó 5 éves életútja - kimutatások, diagramok
- Kiaknázás SAS Enterprise Guide-dal
 - Új kockázati mutatók, kiválasztási állomány bármely szempont szerint
- Csalási mintázatok modellezése felügyelt tanulással
 - 1 db lineáris regressziós, 7 db logisztikus regressziós modellszabály
 - 550 ezer ÁFA alany kockázati sorrendezése és szegmentálása
- Kockázati besorolások visszajuttatása a tranzakciós rendszerekbe
 - Kiutalás előtti kiválasztó rendszer
 - Revíziós rendszer – a megbízólevél kiállítás időpillanatában érvényes



Székely Józsefné
Vas György
Lakó Ferenc
Sárközi Margit
Angeli Károlyné
László László

Adatbányászati modellszabályok becslési pontossága a következő évi ellenőrzések eredményein visszamérve

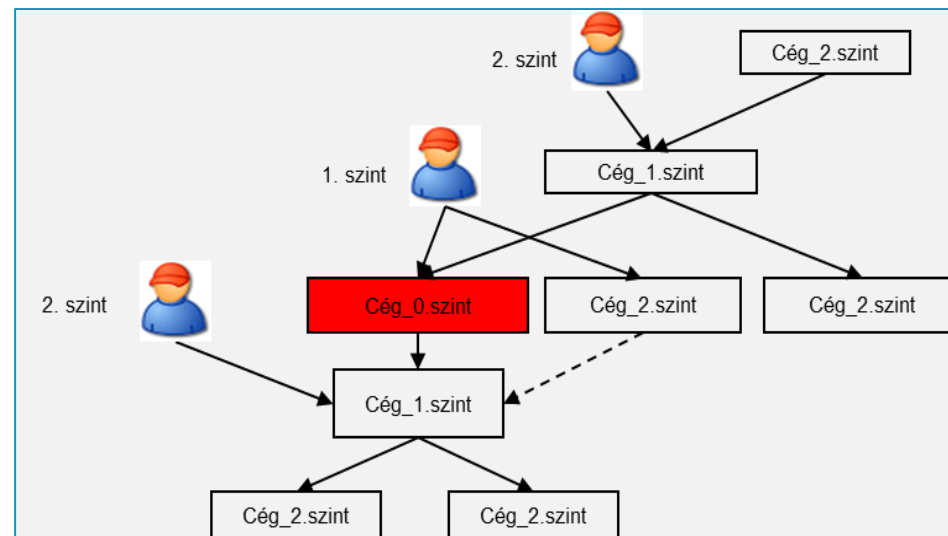




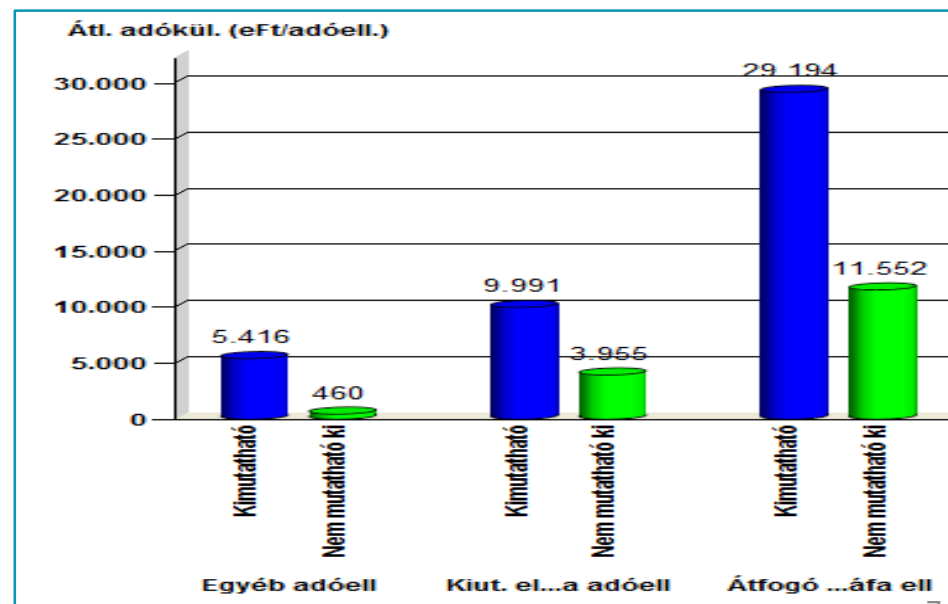
IT vezetés: miért SAS kiaknázás?

Szándékos adóelkerülő cégek hierarchia kapcsolati (kapcsolatszám, összes elcsalt összeg)

- Éves licenszelés! IT területen Oracle kompetencia!
- Szabad adattranszformáció programozási ismeret nélkül
 - IT közreműködés nem szükséges – megyei IT sem hiányzik
 - Elvi alapon (írásban) nehezen definiálható, bonyolult feldolgozások, lekérdezések összeállítása
 - Táblák szabad összekapcsolása, műveleti lépésenként ideiglenes eredménytábla
 - Hetek, hónapok helyett napok alatt eredmény
 - Adózoói dosszié mögötti adatpiacokat már a projektben a kockázatelemzők állították elő
 - Importálási lehetőség (Excel, Dbase, CSV)
 - Betölti, összekapcsolhatja az adattárház tábláival, elmentheti az adattárházba
- **Gyanús hálózatok adatpiaca** (Lakó Ferenc, Vas György)
 - A cégek tulajdonosain, vezetőin keresztül (később székhelycímeik is)
 - KAIN - minden kapcsolat 2 szintig, KoKaIn – adóelkerülők kapcsolatai
 - 7 szakrendszer adatköreinek összekapcsolása
 - SAS eljárással - 2 kockázatelemző kollégám, hetente frissül



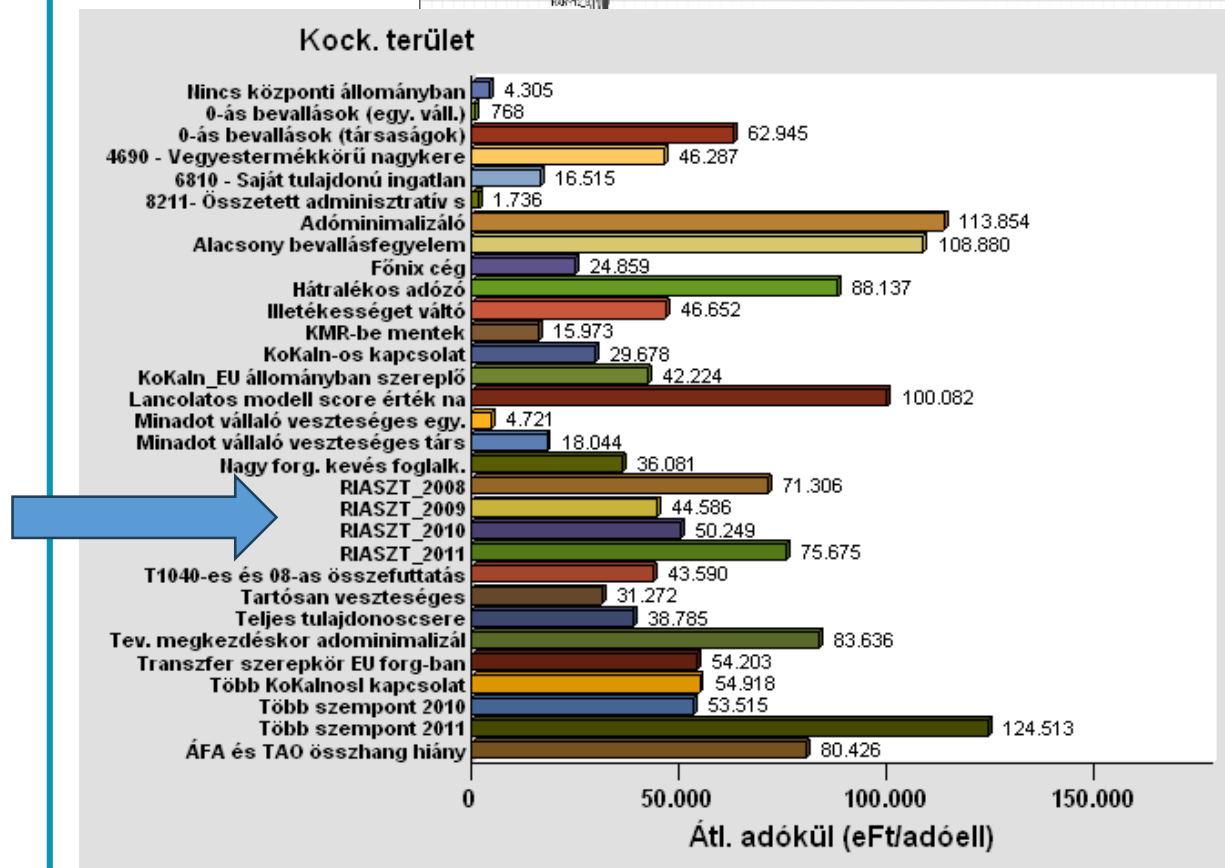
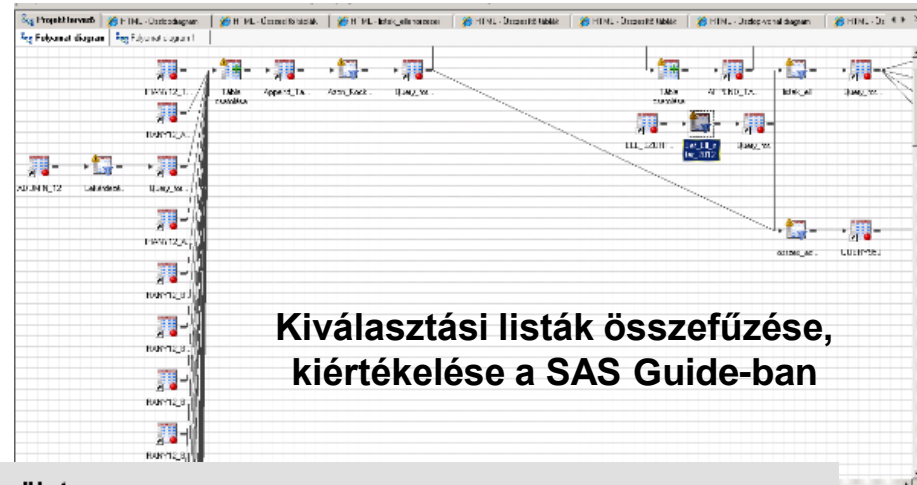
Csalárd adózók érdekkörébe tartozás relevanciája





SAS Guide további lehetőségek

- Adatpiaci táblák és központi kiválasztási állományok előállítás, adattárházba írása
- Közös munka támogatása
 - Központ publikációs szerver könyvtár, igazgatóságok saját szerver könyvtárai
 - Központi kiválasztási állományok továbbfejlesztése
- Előzetes és utólagos kontroll lehetősége
 - Kiválasztási listák hatékonyságmérése kiküldés előtt
- Gyors és pontos reagálás lehetősége
 - Központi ún. **RIASZT listák** az előző negyedévi ÁFA csalások kiszűrésére
 - Szakértői szabályok és döntési fa algoritmus alapján
 - Lakó Ferenc
- Együttműködő adattárházak
 - Nagyobb állományok már nem kerültek áttöltésre
 - A SAS-sal mindkét adattárházat elérjük



Kiválasztási listák hatékonysága az utólagos adóellenőrzéseknél



Revizori kapacitás átrendezése

➤ Szembesülés:

- Az ellenőrzési portfólióban a kisebb és alacsonyabb kockázatú adózók vannak mégis túlsúlyban
- Főként a fővárosi igazgatóságokon
- Revizorok kiválasztási szempontja:
 - Melyik van közelebb a lakásomhoz és tudok-e parkolni?
- Az igazgatóságok túlzottan őrizték a vadászterületüket
 - Inkább a múltat (az elévülési idő kezdetéhez közeli időszakokat) vizsgálják

➤ Az ellenőrzési portfólió szerkezetének módosítása

- Magas-közepes kockázatú szegmensekben elvárt darabszám
- 2010-es előzetes számítás szerint 90 milliárd Ft adókülönbözet növekmény várható
- Jelen közeli időszakokra koncentráció
- RIASZT listákról minimum ellenőrzöttségi szintek, de azon belül helyi döntés
- Az átrendezés hatása 110 milliárd Ft növekményt eredményezett 2011-ben

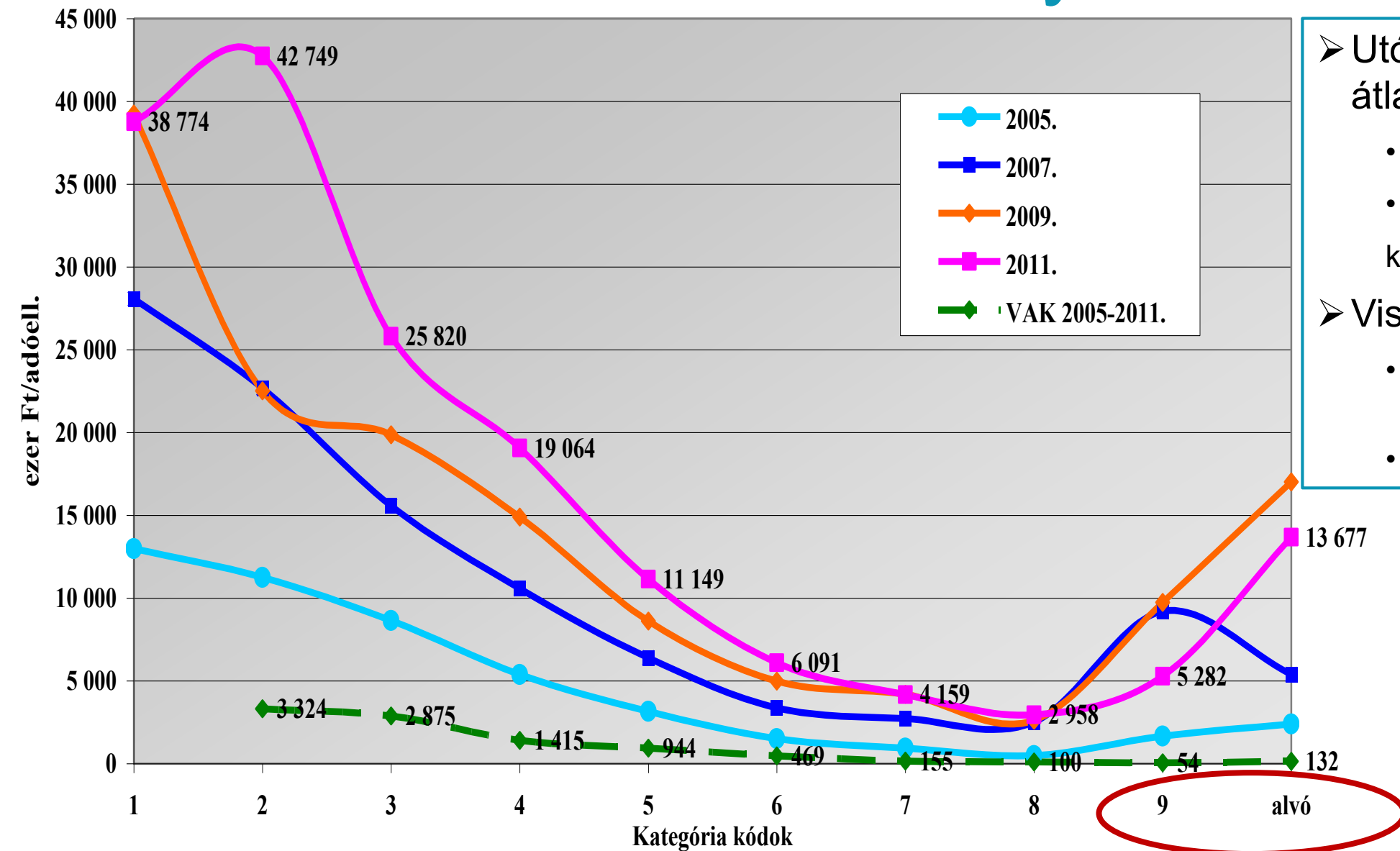
➤ Szemléletváltás

- A költségvetés védelme, a csalárd tevékenység lerövidítése a prioritás
- Rendszeres képzések – kockázatelemzők és ellenőrzési vezetők

Kategória kód	Központi Kockázati Besorolás				Összesen db
	Magas	Közepes	Alacsony	Nincs besorolva	
	db	db	db	db	
1	935	1 061	1 059	20	3 075
2	1 802	2 005	2 586	27	6 420
3	3 320	4 105	5 393	25	12 843
4	5 607	6 320	10 543	61	25 731
5	8 741	9 663	20 292	164	52 361
6	12 706	13 820	39 694	911	103 688
7	17 368	18 482	79 319	3 690	172 631
8	12 800	41 440	40 647	7 216	102 137
9	3 833	13 763	9 689	2 304	29 589
alvó/0	7 405	12 239	3 804	3 458	26 906
nem besorolt	283	1 199	6 989	4 909	13 380
Összesen	74 801	231 160	220 015	22 785	548 761



Eredmények



➤ Utólagos adóellenőrzések
átlagos adókülönbözete

- a célzott és
- a véletlen
kiválasztásoknál

➤ Viszonyítási alap:

- Kategóriánként véletlen
minta
- Az ellenőrzések 5 %-a



Köszönöm a figyelmüket !