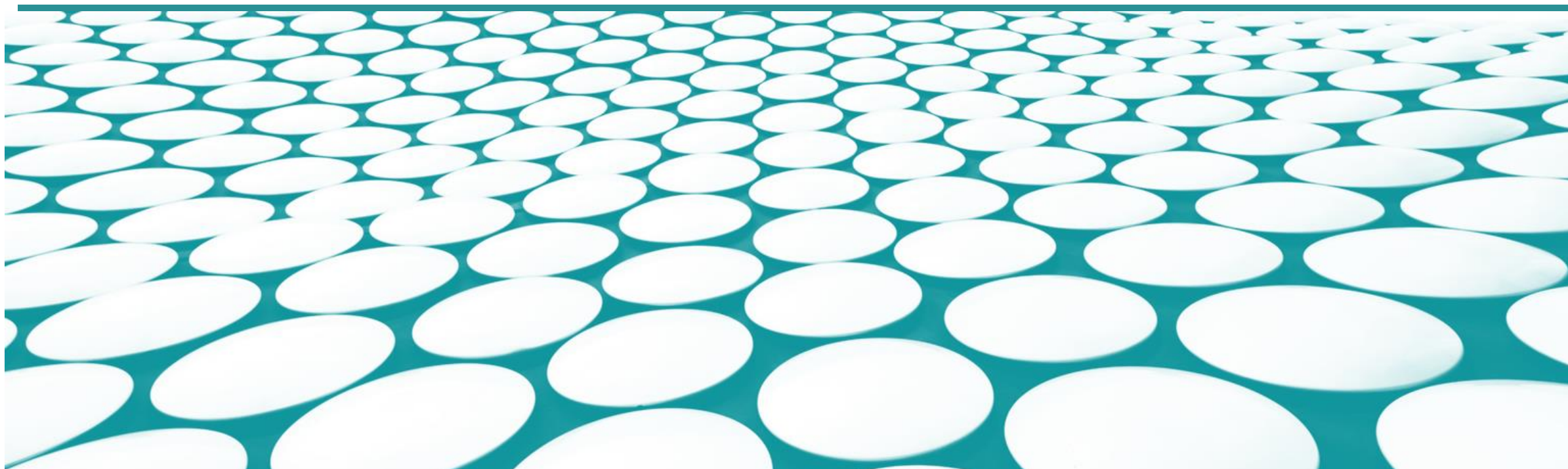


„SZÁMLA SZOKATLANSÁG” SZÁMLA ÉRTÉKÉHEZ KÖTŐDŐ ANOMÁLIÁK FELTÁRÁSA

FARKAS ALFRÉD FAJKUSZ ÁDÁM

2025. MÁJUS 28.



TARTALOM

- ❖ Célok áttekintése
- ❖ Elkészült megoldások
 - ✓ Statisztikai megoldás
 - ✓ Regressziós modell
 - ✓ Klasszifikáció
- ❖ Vizsgamérés
- ❖ Tovább lépési lehetőségek

CÉLOK ÁTTEKINTÉSE

Számla szokatlanság

Gazdaságilag nem
racionális viselkedés

Ritka
események

Adminisz-
tratív hiba

Tevékeny-
ségbe nem
illő
események

Fiktív
események

Beruházá-
sok

Tevékeny-
ség/profil-
váltás

Szokatlan lehet:

- ❖ Érték
- ❖ Tétel megnevezés
- ❖ Számlaszám
- ❖ Dátum
- ❖ Partner viszonylat
- ❖ Stb.

STATISZTIKAI MEGOLDÁS CÉLOK

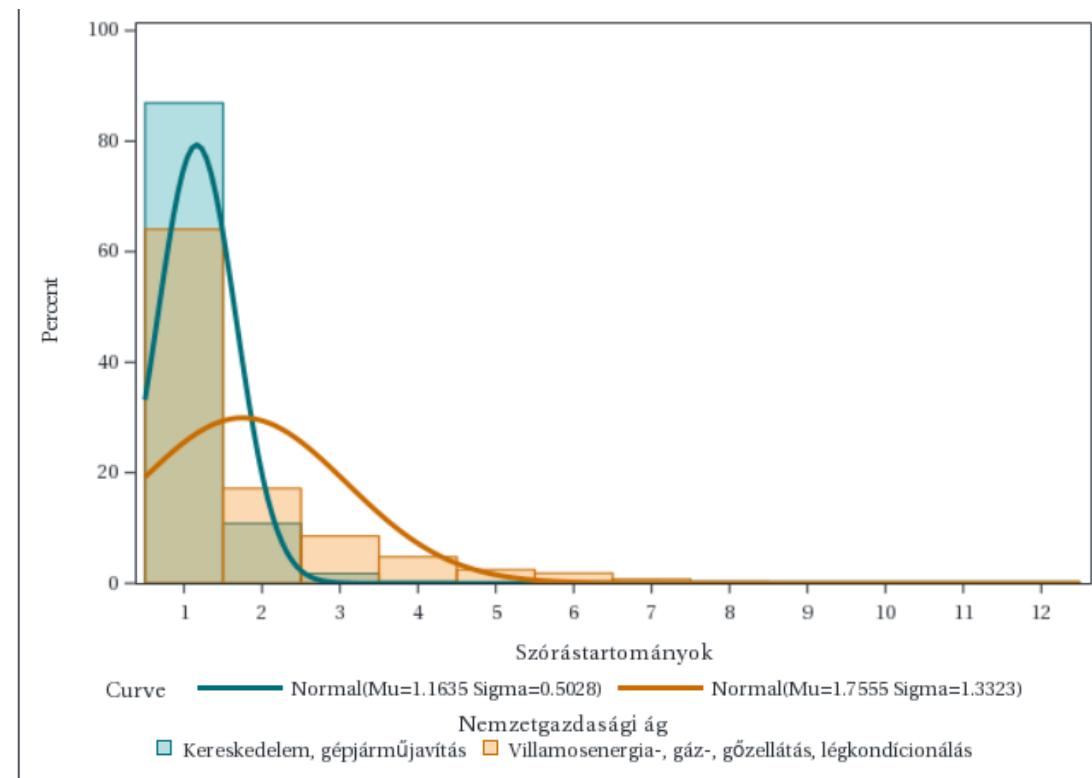
Jelenlegi kiugró érték képzési algoritmus:

- Pillér Kft. fejlesztésében egy robusztus módszer az extrém outlierok kiszűrésére,
- Z-score alapú módszer, középértéktől való távolság,
- Adózói méret alapú szegmensekhez igazított fix limitértékekkel.

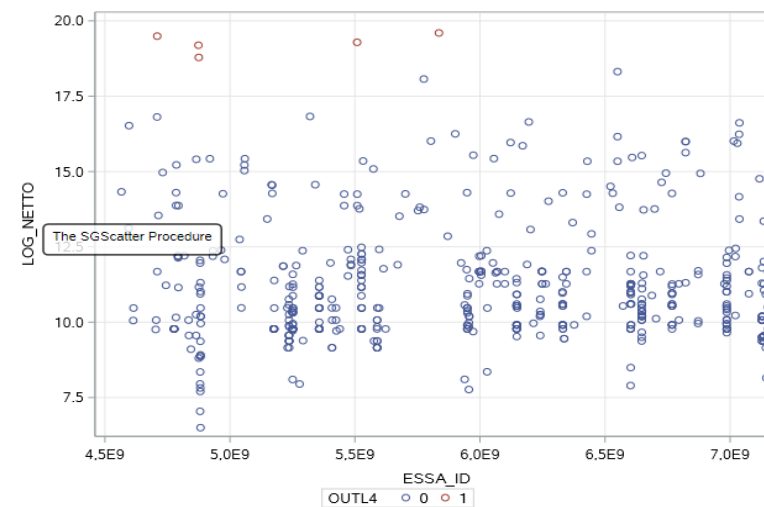
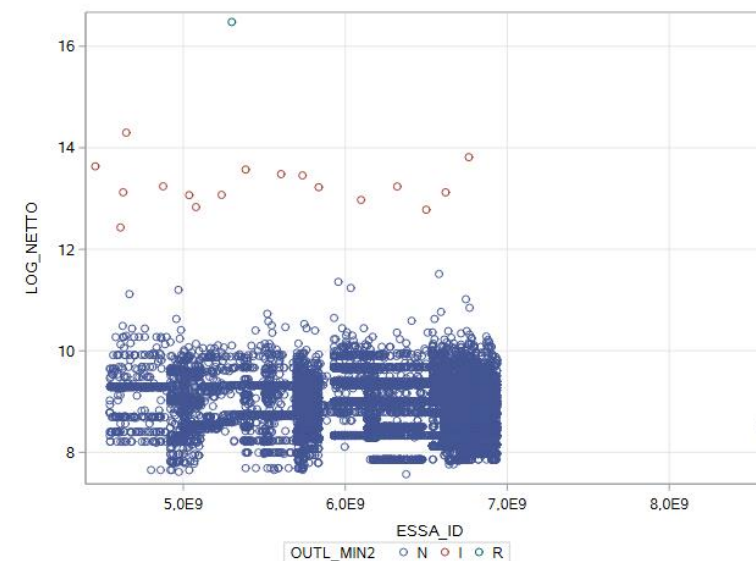
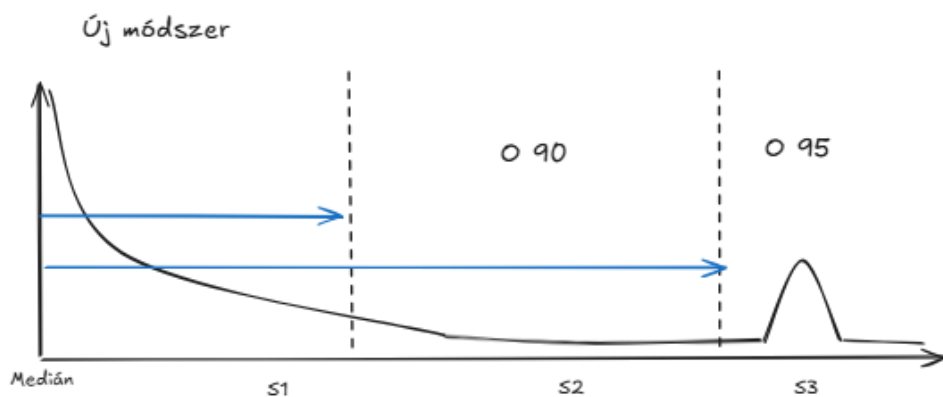
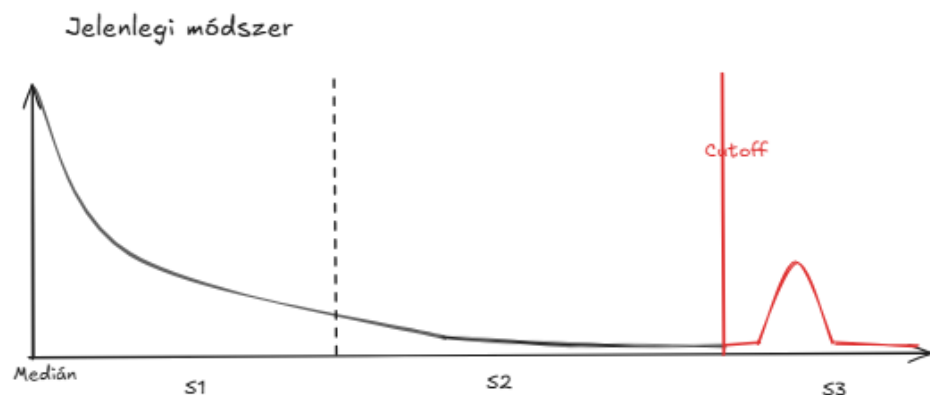
Az új megoldás fejlesztésénél cél volt a

- kiugró érték jelöléséhez egy olyan dinamikus, normalizált mutató létrehozása, amely
- az adózó korábbi számlázási viselkedését hangsúlyosabban figyelembe veszi,
- ugyanakkor a hasonló tulajdonságú adózókkal történő összevetést is cizelláltabban valósítja meg.

STATISZTIKAI MEGOLDÁS PROBLÉMA



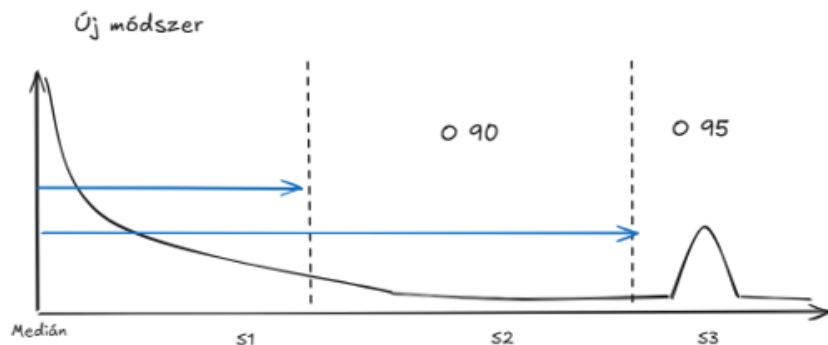
STATISZTIKAI MEGOLDÁS ÖSSZEVETÉS



PÉLDA - KÖNYVKERESKEDÉS

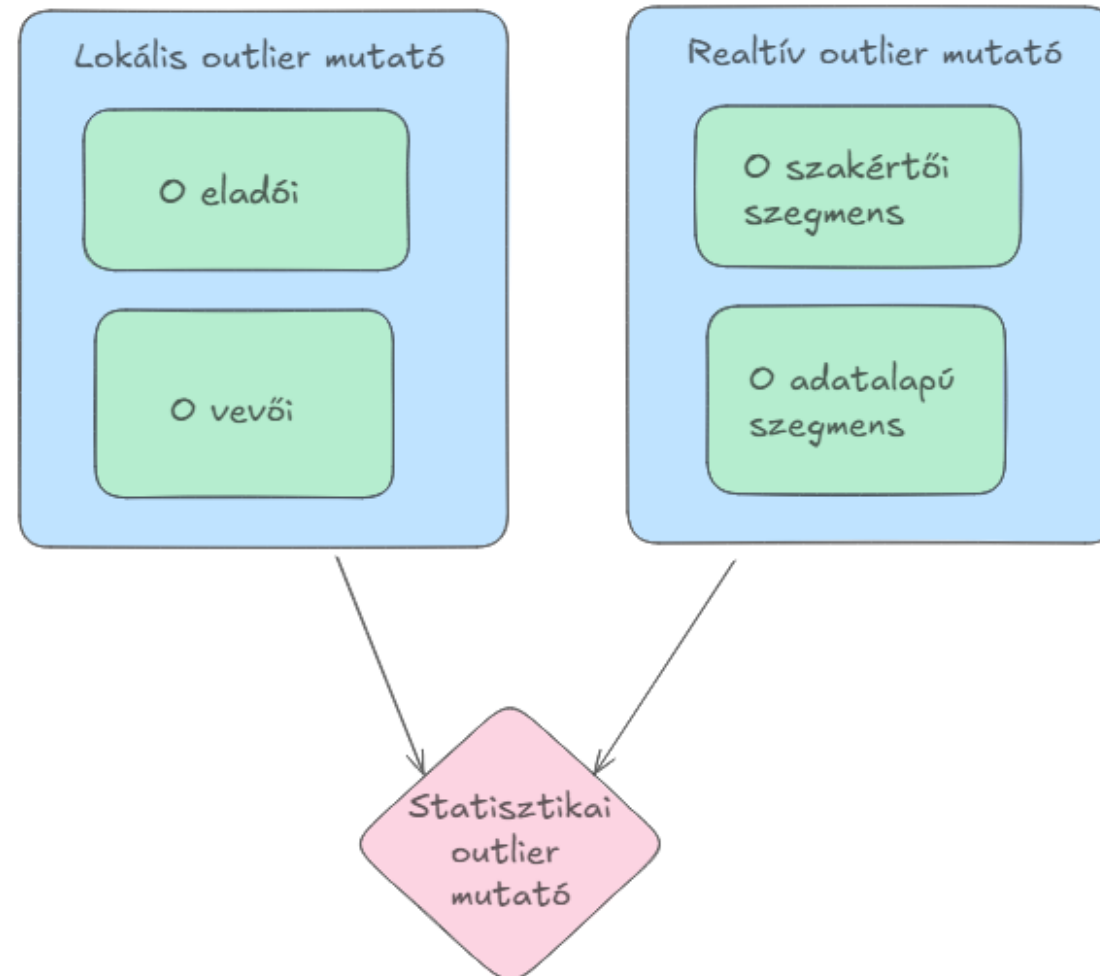
ESSA_ID	Számla tétel összes	Mennyiség összesen	Nettó összeg összesen	Megnevezések
4612568088	2	2	250 000	"Könyvcímek"
6500535065	7	174	354 960	"Könyvcímek"
6097704941	7	211	430 440	"Könyvcímek"
5036615700	6	232	473 280	"Könyvcímek"
5237454568	6	233	475 320	"Könyvcímek"
6619674913	7	245	499 800	"Könyvcímek"
5836000670	7	271	552 840	"Könyvcímek"
6321440177	7	275	561 000	"Könyvcímek"
4876000458	6	276	563 040	"Könyvcímek"
5604140281	7	351	716 040	"Könyvcímek"
5385441404	7	384	783 360	"Könyvcímek"
4457033394	6	409	834 360	"Könyvcímek"
6763268983	1	1	1 000 000	Előadás
4648427004	6	794	1 619 760	"Könyvcímek"
5299876585	1	1	14 330 709	Autóértékesítés AABW-776 / WDD2573591A010933

STATISZTIKAI MEGOLDÁS MÓDSZER

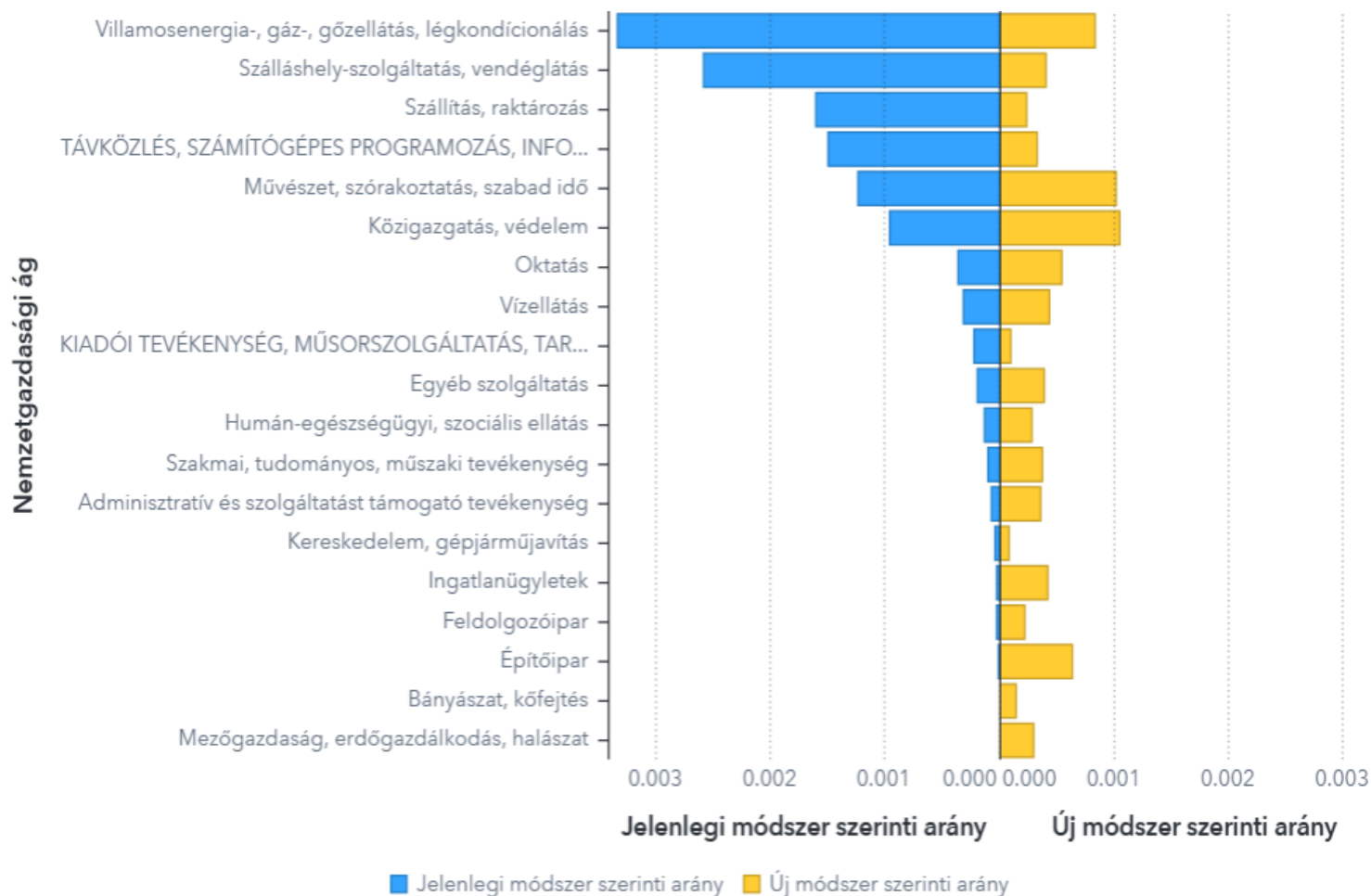


- 1 $Z_i = \frac{x_i - \mu}{\sigma}$ - Z score értékek kiszámítása
- 2 $s_i = \text{round}(Z_i)$ - szórástartományok meghatározása
- 3 $f(s_i) = \frac{N_{s_i}}{N} \times 100$ - szórástartományokra jutó gyakoriságok meghat
- 4 $F(s_i) = \sum_{i=1}^k f(s_i)$ - kumulált gyakoriságok meghatározása
- 5 $O(s_i) = \frac{4 * \text{nettó összeg} + 2 * \text{F}(s_{i-1}) \text{ mennyi ség}}{6}$ - outlier mutató meghatározása

6



STATISZTIKAI MEGOLDÁS - EREDMÉNYEK I.



STATISZTIKAI MEGOLDÁS

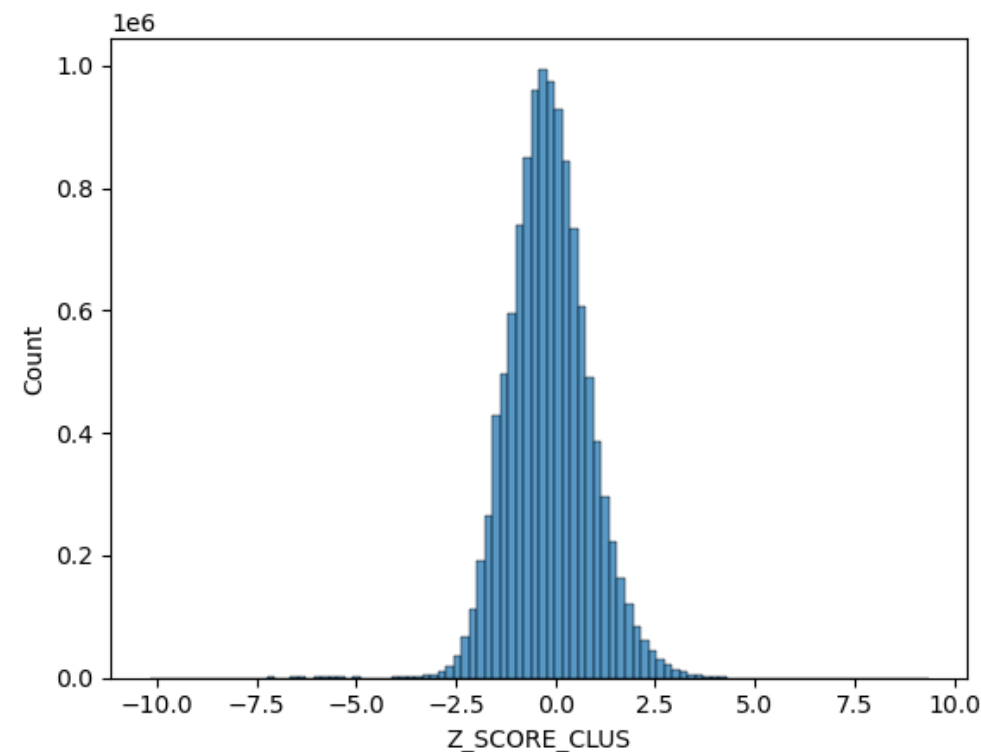
EREDMÉNYEK II.

ESSA_ID	Számla tétel összes	Mennyiség összesen	Nettó összeg összesen	Megnevezések	O saját	O relatív	O átlagos
4612568088	2	2	250 000	"Könyvcímek"	65,31%	25,41%	38,71%
6500535065	7	174	354 960	"Könyvcímek"	65,31%	33,33%	43,99%
6097704941	7	211	430 440	"Könyvcímek"	73,31%	33,33%	46,66%
5036615700	6	232	473 280	"Könyvcímek"	73,31%	54,74%	60,93%
5237454568	6	233	475 320	"Könyvcímek"	73,31%	54,74%	60,93%
6619674913	7	245	499 800	"Könyvcímek"	73,31%	54,74%	60,93%
5836000670	7	271	552 840	"Könyvcímek"	84,90%	54,74%	64,79%
6321440177	7	275	561 000	"Könyvcímek"	73,31%	54,74%	60,93%
4876000458	6	276	563 040	"Könyvcímek"	84,90%	54,74%	64,79%
5604140281	7	351	716 040	"Könyvcímek"	84,90%	54,74%	64,79%
5385441404	7	384	783 360	"Könyvcímek"	84,90%	54,74%	64,79%
4457033394	6	409	834 360	"Könyvcímek"	85,95%	71,77%	80,06%
6763268983	1	1	1 000 000	Előadás	92,97%	62,53%	77,75%
4648427004	6	794	1 619 760	"Könyvcímek"	83,33%	65,12%	71,19%
5299876585	1	1	14 330 709	Autóértékesítés AABW-776 / WDD2573591A010933	100,00%	98,41%	98,94%

REGRESSZIÓS MODELL CÉLOK

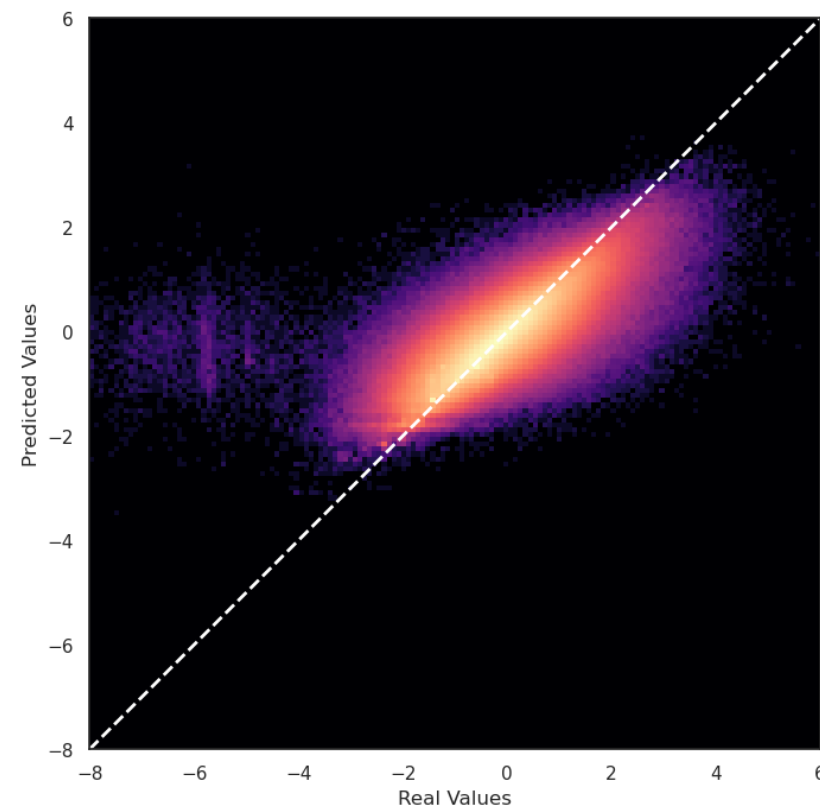
- ❖ Cél: számlaérték becslése és kiugró számlák meghatározása
 - ✓ Cégszintű változók
 - ✓ Számlaszintű változók
- ❖ Célváltozó: Z-Score
 - ✓ Számlaérték helyett
 - ✓ Saját csoporthoz képest számított

$$Z_SCORE_CLUS = \frac{SZÁMLAÉRTÉK - SZÁMLAÉRTÉK_MEAN_{CLUS}}{SZÁMLAÉRTÉK_STD_{CLUS}}$$



REGRESSZIÓS MODELL TANÍTÁS

- ❖ Tanítás ~12 millió számlán
- ❖ Validáció ~30 millió számlán
- ❖ $R^2 = 0,63$
- ❖ Legerősebb változók:
 - ✓ Számlán szereplő tételek mennyiségének összege
 - ✓ Előző 1 év fajlagos számlaértékeinek átlaga
 - ✓ Kiállító & befogadó TEÁOR
 - ✓ Pénznem
 - ✓ Előző időszakokra vonatkozó beszerzési és kiadási adatok

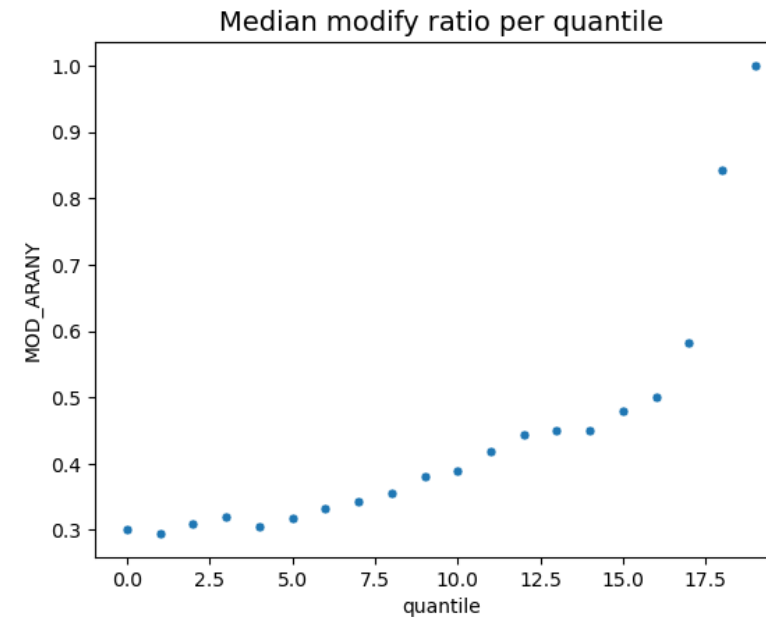
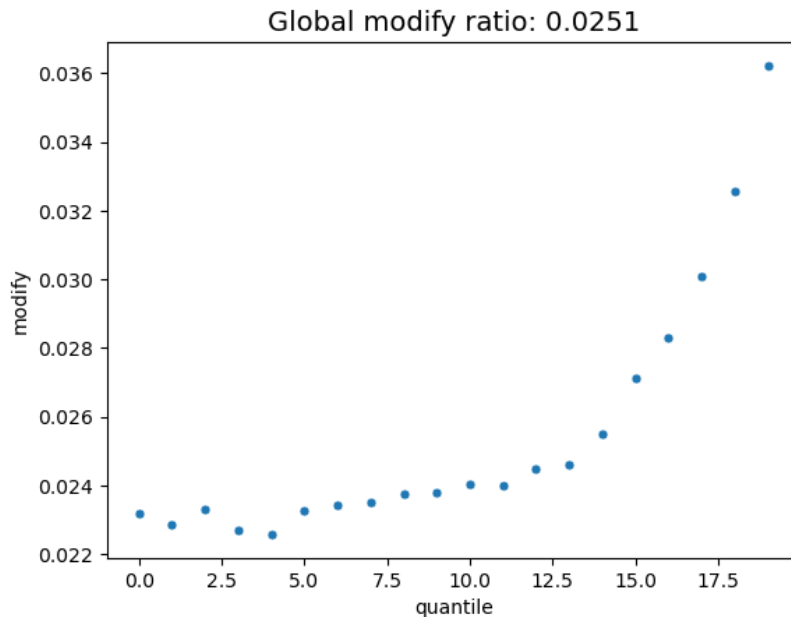


REGRESSZIÓS MODELL VALIDÁCIÓ

❖ Hipotézis:

- ✓ Az adózók többsége „racionálisan” viselkedik -> A modell a jellemző viselkedésre tanul rá
- ✓ Nagyobb hiba -> nagyobb modify valószínűség

❖ Az eredmény igazolja (majdnem teljesen)



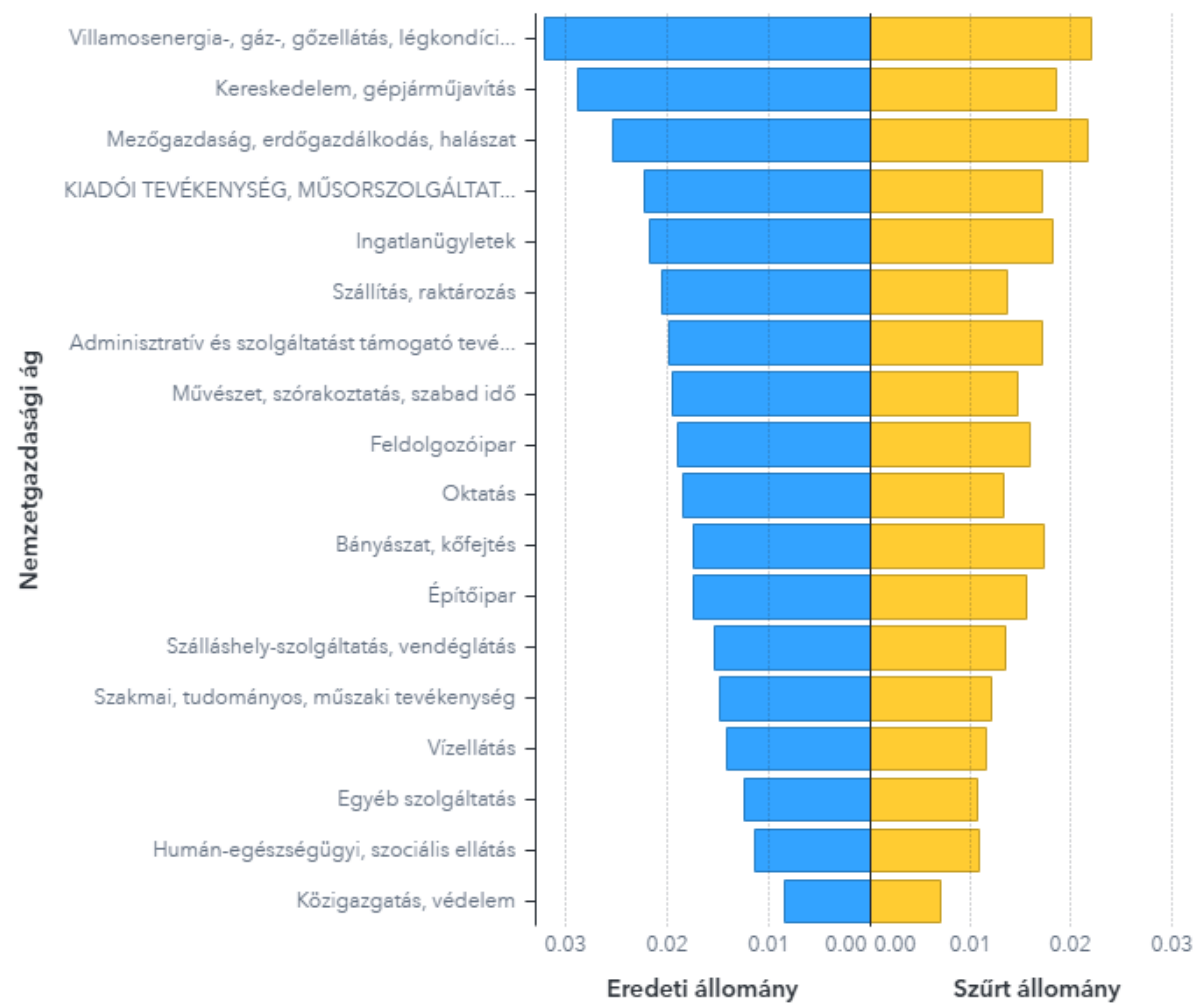
KLASSZIFIKÁCIÓ CÉLVÁLTOZÓ

Cél:

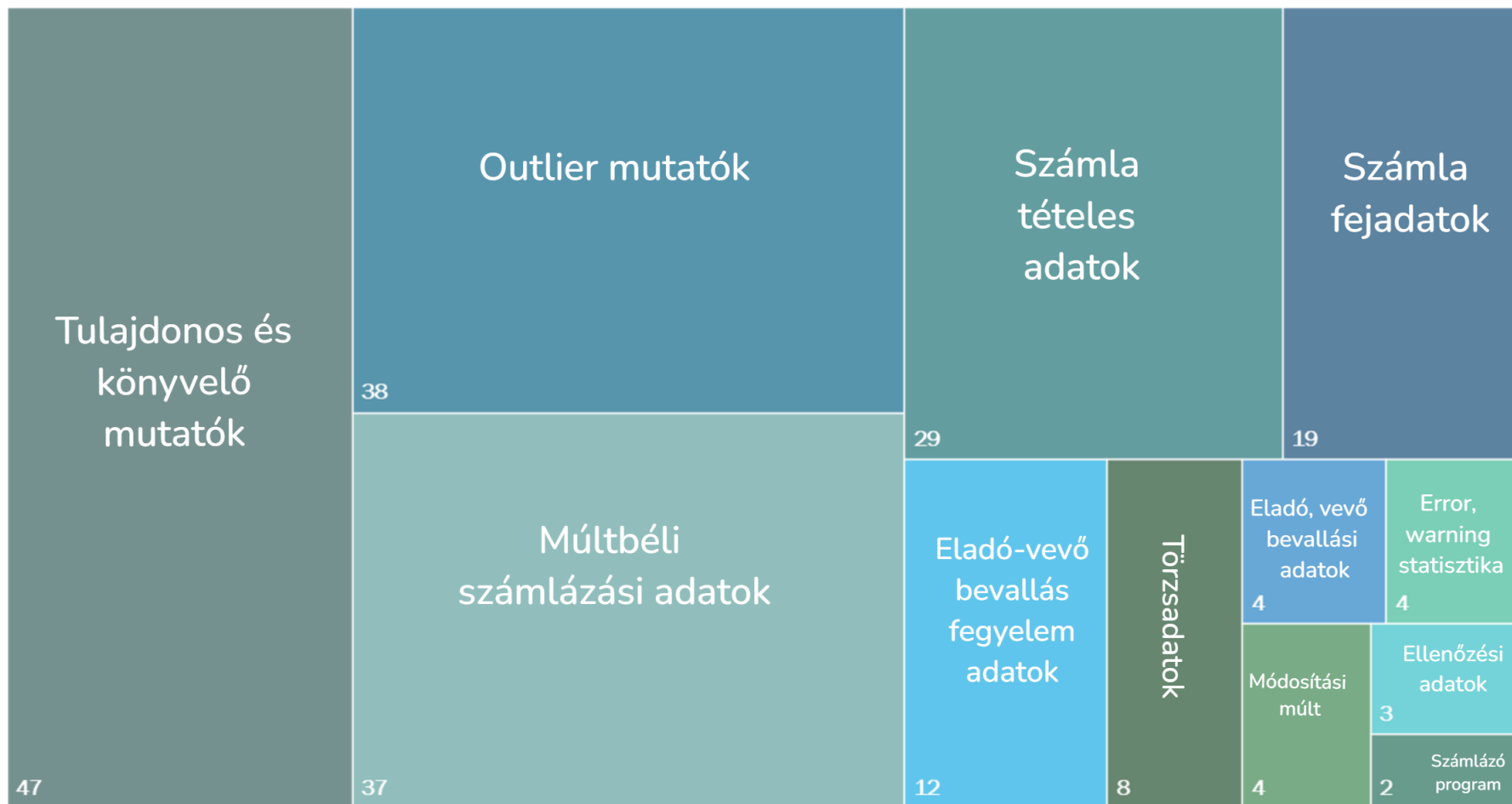
- ❖ Adminisztrációs hiba előrejelzése

Célváltozó:

- ❖ számlák értéke módosítással csökkentésre került
- ❖ a vizsgált időszakban adózó számláinak kevesebb mint 10%-át módosította
- ❖ Legalább az eredeti érték 10%-val



KLASSZIFIKÁCIÓ MAGYARÁZÓ VÁLTOZÓK



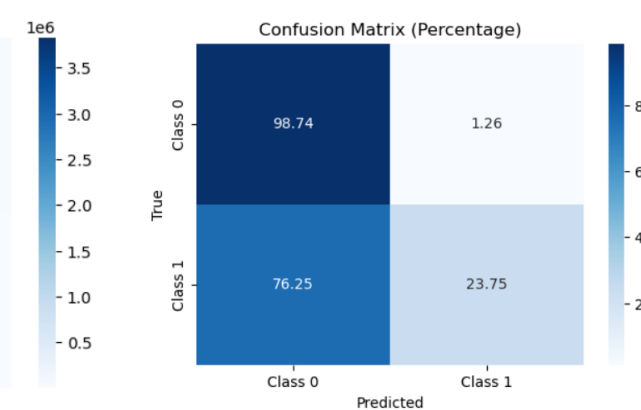
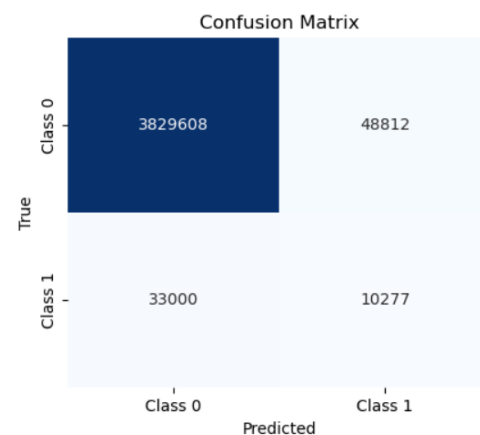
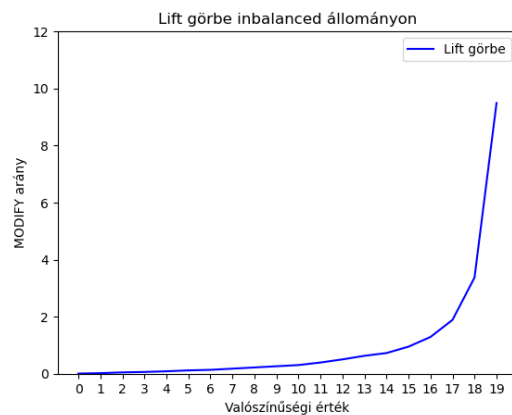
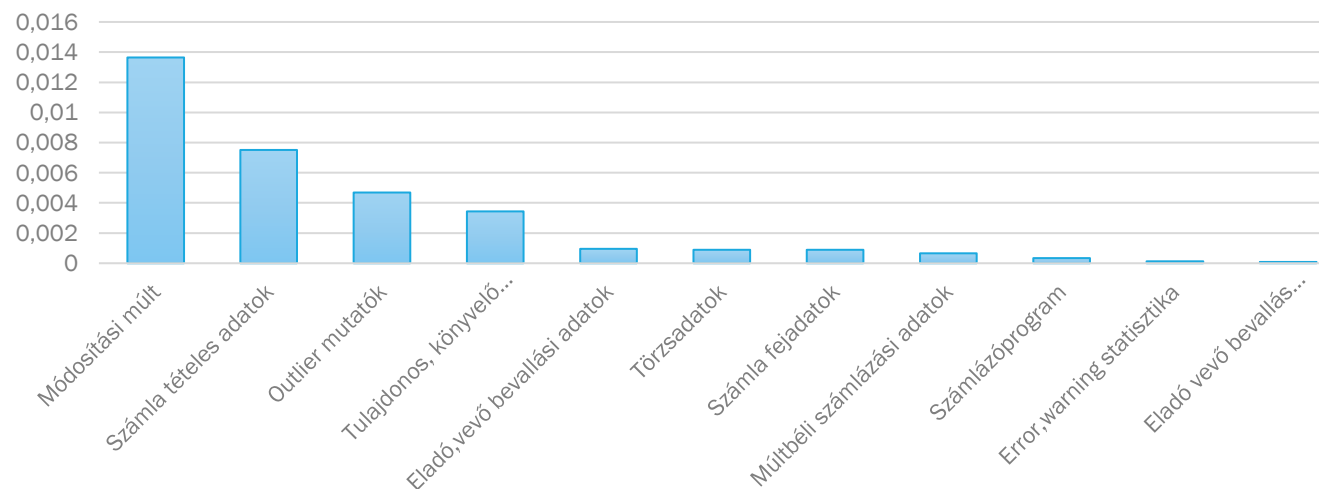
Összesen 207 mutató

KLASSZIFIKÁCIÓ PERFORMANCIA

Fejlesztési állomány

- ❖ A múltbéli adatok időigényes futása miatt csak 2023. évre tanultunk
- ❖ Célváltozó eredeti aránya 1,6% volt – extrém tanulóállományban undersampling módszerrel arány 10%-ra dúsítása
- ❖ Tanulóállomány – 484 697 db számla

Mutatócsoportok modell szerinti fontossága



VISSZAMÉRÉS

- ❖ Meglévő csalásfelderítési modellek által kiválasztott bevallások esetén bejövő és kimenő számlák megjelölése (közel 1000 ellenőrzés)
 - ✓ több kijelölési módszer tesztelése (metszet, unió, súlyozott mutatók)
 - ✓ módszerek finomhangolása az eredmények alapján
- ❖ Adategyeztetési eljárás keretében történő számlaszintű kiajánlások
- ❖ Sikeres visszamérés esetén külső-belső szolgáltatások fejlesztése

HASZNOSÍTÁS

Belső szolgáltatások:

- ❖ Számlaszintű jelölők elhelyezése a KOCKA adatbázisokban
- ❖ Intézkedések vezérlése

Külső szolgáltatások:

- ❖ Online számla - INFO
- ❖ eÁFA - INFO
- ❖ ÜPO csempe

TOVÁBBI FEJLESZTÉSEK

Számla szokatlanság

Gazdaságilag nem
racionális viselkedés

Ritka
események

Adminisz-
tratív hiba

Tevékeny-
ségbe nem
illő
események

Fiktív
események

Beruházá-
sok

Tevékeny-
ség/profil-
váltás



KÖSZÖNJÜK A FIGYELMET!

A kutatást végzők:

- ❖ **Fajkus Ádám Tibor**
- ❖ Borkó Károly
- ❖ Rikli Szabolcs

Közreműködők:

- ❖ Farkas Imre – Borsod-Abaúj-Zemplén
- ❖ Bolla Orsolya – Kelet-Budapest
- ❖ Kapás Eszter – Kelet-Budapest
- ❖ Kónya Nándor – Baranya
- ❖ Décsi Ivett – Baranya
- ❖ Jancsó Zsuzsanna – Baranya/KAF