

Data Science Klub

3.

Data Science verseny

azaz hogyan legyen 2000 előfizető helyett 5000?

A teleprompter.com célja, hogy innovatív és professzionális teleprompter-megoldásokat kínáljon különböző platformokon, lehetővé téve a felhasználók számára a szkriptek egyszerű kezelését és a magabiztos előadásokat.

Jelen feladat keretében azt vizsgáljuk, hogy hogyan tudjuk növelni az előfizetők számát, kiindulva a vállalat által biztosított log adatbázisból.



2025 január 1 - 2025 február 28 közötti időszak

28 545 487 esemény
névvel, időponttal, országgóddal, és a használt operációs
rendszer verziószámával

191 777 felhasználó

741 egyedi esemény

289 féle elkülönített országgód

120 nevesített operációs rendszer verzió

Az adatokban fellelt problémák 1

Januári időszak:

A keresztény kultúrkörben és a világ egyéb részein is történelmi hagyományok és társadalmi szokások alapján jellemző az újévi fogadalomtétel.

Mindannyiunknak vannak barátai, ismerősei akik megfogadták már új évkor, hogy lefognak, kevesebbet/egyáltalán nem isznak, dohányoznak, kedvesebbek lesznek embertársaikhoz, többet tanulnak stb...

Láttuk és tanulmányokból is követhető, hogy ezek a fogadalmakat nem minden esetben sikerül betartani, de mindenesetre januárban elkezdődik egyes emberek életében "valami más".

A fenti okfejtés miatt, a januári adatokból nem feltétlenül szerencsés konzekvenciákat levonni, de legalábbis nem teljeskörűen.

Jelenleg nem tehetjük meg, hogy a januári adatokat kiemeljük, ezért eltekintünk ettől a problémától!

<https://www.pewresearch.org/short-reads/2024/01/29/new-years-resolutions-who-makes-them-and-why>
<https://ng.24.hu/kultura/2024/01/06/az-ujevi-fogadalmak-pszichologiaja>

Az adatokban fellelt problémák 2

Elemzendő időszak/vizsgálandó felhasználók kiválasztása:

Az adatokból kiderül, hogy a trial időszak elkezdése előtt is van érdemi tevékenység egyeseknél. Amennyiben teljeskörű elemzést szeretnénk, akkor ezen időszak alatti eseményeket is figyelembe kellene venni, hiszen, ezen tapasztalatok ugyanúgy befolyásolhatják az előfizetési hajlandóságot.

Az adatokból nem állapítható meg egyes userek összes tevékenysége, nincs/nem definiálható belépési pont, amely előtt nincs érdemi tevékenység.

Probléma az is, hogy egyes userek "trial_cancelled_event" és 'trial_converted_event' eseménnyel is rendelkeznek, ezen usereket nem tudjuk egyértelműen kategóriába sorolni.

További probléma, hogy egyes userek esetén nincs a trial időszak elkezdése, és a 'trial_converted_event' esemény között érdemi tevékenység, gyakorlatilag látszólag minden tapasztalat és kipróbálás nélkül előfizetővé váltak.

Vélhetően az ilyen felhasználók egyéb, nem azonosított eszközökön ismerték meg az applikáció lehetőségeit. E problémakör jelenléte megkérdőjelezheti a következtetéseket, hiszen az "egyéb" eszközön olyan események történhetnek, amelyek nem elemezhetők jelenleg.

Az adatokban fellelt problémák 2

a feloldás

Elemezendő időszak/vizsgálandó felhasználók kiválasztása:

A felsorolt problémák miatt csak azon userek lesznek kiválasztva, akiknek:

- van 'trial_started_event' eseménye és
- van "trial_cancelled_event" eseménye vagy van 'trial_converted_event' eseménye de csak az egyik és
- legalább 3 eseménnyel rendelkezik

Azon eseményeket vizsgáljuk,

- amelyek a 'trial_started_event' eseménnyel egy időben ill. utána történtek és
- 'trial_cancelled_event'/'trial_converted_event' eseménnyel egy időben, vagy előtte történtek.

Az adatokban fellelt problémák 3:

Vizsgálandó események köre/információtartalma:

a) 741 egyedi esemény fordul elő az adatállományban, de a mellékelt leírásban csak 614 esemény van értelmezve. A kapott információ szerint, a nem definiált események, és a definiált események egy része 'nem érdekes'. Nyilván a domain tudás birtokában ezt ki lehet jelenteni, de előfordulhat, hogy ezt az állítást érdemes felülvizsgálni.

Mivel a feladat kiírója, így/ilyen körülmények között kívánja az adatelemzést végrehajtani, így a fenti felvetésektől eltekintünk, elfogadjuk, hogy a hiányzó események "értéktelenek" az adatok feldolgozásának folyamatában.

b) Az események leírása nem teljeskörű, a fejlesztők számára nyilván egyértelmű, de a programot nem ismerők számára kissé szűk információmennyiség.

Az események között duplum-ok fordulnak elő, de nem egyértelmű, hogy mely eseménynek melyik a párja.

A fentiek probléma feloldása jelenleg úgy lehetséges, hogy az eseményeket kategóriákba rendezzük, így az egymásnak körülbelül megfelelő események egy kategóriába kerülnek.

A vizsgálat feléptése

Adatszűrés

Elemezendő időszak/vizsgálandó felhasználók kiválasztása:

A felsorolt problémák miatt csak azon felhasználók lesznek kiválasztva, akiknek:

- van 'trial_started_event' eseménye és
- van "trial_cancelled_event" eseménye vagy van 'trial_converted_event' eseménye de csak az egyik és
- legalább 3 eseménnyel rendelkezik

Azon eseményeket vizsgáljuk,

- amelyek a 'trial_started_event' eseménnyel egy időben ill. utána történtek és
- 'trial_cancelled_event'/'trial_converted_event' eseménnyel egy időben, vagy előtte történtek.

Az eseményeket nem egyesével vizsgáljuk, hanem kategorizáljuk őket, azaz a valamilyen szempont szerint egybetartozó kategóriák azonos kódot kapnak, azonos csoportba tartozónak tekintjük őket.

A vizsgálat feléptése

Adattárolás

A kategorizálás után az adatállományt átrendezzük úgy, hogy az eseményalapú megközelítés helyett, felhasználó-központúvá alakítjuk, azaz:

- minden felhasználó 1 record
- a felhasználó alapadatait tároljuk:
 - előfizető-e vagy felmondó
 - a felhasználó által használt operációs rendszer (ha nem egyedi akkor a leggyakoribbat)
 - a felhasználó országkódját (ha nem egyedi akkor a leggyakoribbat)
 - az országkód alapján a kontinensek elkülönítése
 - a kategóriakódokat tároljuk, az előfordulási darabszámmal
 - változót kap az események darabszáma
 - és változót kap a rendszerben töltött idő, illetve ennek jellemzői
 - medián
 - módusz
 - átlag

A vizsgálat felépítése

Elemzés tervezett felépítése

1. Előfizetők és Felmondók csoportjának hasonlítása a kategorizált változók alapján
2. Középértékek, maximumok minimumok, szórások
3. korrelációk
4. vajon bármely kategorizált változó és a 2 alcsoportunk szignifikánsan összefüggenek-e? t-teszt
5. adatállomány clusterekre bontása (k-means/dbscan/hierarhikus klaszter)
 - a. cél az olyan alcsoportok elkülönítése, amelyben az egyes alcsoportok dominánsak
 - b. a klaszterek alapján a kódolt eseménykategóriák vizsgálata,
 - c. a fontos kódolt eseménykategóriák bementei eseményeinek vizsgálata
6. A bemeneti események alapján javaslatok kialakítása, mit kell tenni, hogy a kliens előfizetővé váljon
7. + 1: a bemeneti események fontossága alapján egy előrejelző modellt készíteni, amely jelzi, hogy a user várhatóan előfizetővé válik-e

Az események csoportosítása

Az események többszörös átnézése után a következő csoportok alakultak ki:

- .8 - egyéb felhasználói aktivitás
- .9 - felhasználói aktivitás
- .10 - idegesítő/zavaró interakciók
- .11 - programhibák
- .12 - előfizetés rejtett szándék
- .13 - lemondás rejtett szándék
- .14 - speciális szolgáltatás (felhasználó szemszögéből)
- .141 - Facebook-al kapcsolatos interakciók
- .142 - Youtube-al kapcsolatos interakciók
- .15 - felhasználónak nyújtott extra (teleprompter szemszög)

A csoportosítások értelmezése 1

15 - felhasználónak nyújtott extra (Teleprompter szemszög)

Olyan események vannak e csoportban, amelyeket a Teleprompter cég indukálja, az ügyfélmegtartás érdekében. pl AssignedToNewUpgrade - A felhasználó egy új frissítési fizetőfal tesztelésére vagy megtekintésére van kijelölve.

14 - speciális szolgáltatás (felhasználó szemszögéből)

141 - Facebook-al kapcsolatos interakciók

142 - Youtube-al kapcsolatos interakciók

Ezen események speciális események, azaz AI segítség igénybevétele, vagy Facebook/Youtube -al kapcsolatos interakciók. A 141-es és a 142-es kategória, az alacsonyabb elemszám miatt összevonásra kerül a 14-es kategóriával, 140-es kódon. Ezen események körébe tartozik pl. a AIGenerateTap - A felhasználó egy gombra koppintva tartalmat (pl. szkripteket) generálhat mesterséges intelligenciával.

A csoportosítások értelmezése 2

13 - lemondás rejtett szándék

A lemondás rejtett szándék olyan események csoportja, amelyek alapján arra lehet következtetni, hogy a felhasználó nem tervezi hosszabb távon használni az applikációt. Ilyen pl: RegistrationScreenSkipTap - A felhasználó kihagyja a regisztrációt.

12 - előfizetés rejtett szándék

Az előfizetés rejtett szándék olyan események csoportja, amelyek alapján arra lehet következtetni, hogy a felhasználó hosszabb távon tervezi használni az applikációt. Ilyen pl: RegistrationViewControllerNew - Új vagy frissített regisztrációs folyamat.

A csoportosítások értelmezése 3

11 - programhibák

Ezen események halmaza, akár objektív alapokon is nyugodhat, de mindenképpen egy hibajel. Az események az adott pillanatban nem feltétlenül jelentkeznek a felhasználónál, de a jövőbeli működést mindenképpen valamelyest befolyásolják, ezért fontosak. Ilyen pl. RecordingError - Általános hiba a felvétel során.

10 - idegesítő/zavaró interakciók

Az idegesítő/zavaró interakciók a felhasználó számára nem várt felbukkanó üzenetek, figyelmeztetések. Nem álltható, hogy szükségtelenek, de jelentőségük abban van, hogy a felhasználó normál, tervezett munkamenetét megszakítja/megakasztja, adott esetben zavart kelt. OnboardingUpgradeScreen - A felhasználó frissítési kérést lát az előkészítés során.

A csoportosítások értelmezése 4

9 - felhasználói aktivitás

Normál felhasználói aktivitások gyűjteménye. A csoport olyan események gyűjteménye, amelyekből már az extra események (14-es kategória) ki vannak emelve, és közvetlen felhasználói interakciók is (lsd. lent - 8. kategória).
pl. FontChanged - A felhasználó megváltoztatta a betűstílust.

8 - egyéb felhasználói aktivitás

Nem közvetlen interakciók. Ezen események valamilyen események szükséges következményei, mit az adott felhasználási területhez tartozó képernyő bezárása. pl. `ConnectAccountViewController<ConnectYouTubeAccountViewModel>.dismiss()` - A felhasználó bezárja a csatlakozási YT képernyőt.

0 az applikáció belső működésének logolása

Jelenleg fel nem használt kódcsoporthoz, mivel nem világos, hogy ezen események, hogyan befolyásolják a felhasználói tevékenységet

Feltáró elemzések

Az előfizetők és a felmondók összehasonlítása 1

Category	1 - subscriber	2 - cancelled	3 - subscriber event_rate	4 - cancelled event_rate	5 - subscriber user_rate	6 - cancelled user_rate
db_08 egyéb felhasználói aktivitás	50 175,00	160 419,00	14,48	14,34	24,79	13,91
db_09 felhasználói aktivitás	273 599,00	862 759,00	78,93	77,14	135,18	74,79
db_10 bosszantó	4 492,00	23 534,00	1,30	2,10	2,22	2,04
db_11 hiba	2 069,00	9 571,00	0,60	0,86	1,02	0,83
db_12 előfizetés rejtett szándék	4 647,00	26 429,00	1,34	2,36	2,30	2,29
db_13 felmondás rejtett szándék	598,00	6 080,00	0,17	0,54	0,30	0,53
db_140 speciális szolgáltatások	5 492,00	14 725,00	1,58	1,32	2,71	1,28
db_15 TP extra	57,00	164,00	0,02	0,01	0,03	0,01

A fenti táblázat

- 1 - subscriber és 2 - cancelled oszlopa események darabszámát tartalmazza,
- 3 - subscriber event_rate és a 4 - cancelled event_rate az események darabszámár elosztva az összes feldolgozott esemény darabszámával
- 5 - subscriber user_rate és 6 - cancelled user_rate pedig az események darabszáma eloszva az összes kategóriába kerülő felhasználó elemszámával

Tehát a

- a db_08 és a 2 - cancelled keresztben a mező tartalma azt jelenti, hogy a felmondó felhasználók 160 419 olyan eseményt generáltak, amely a 8as kódot kapta
- a db_09 és a 3 - subscriber event_rate keresztben a mező tartalma azt jelenti, hogy az előfizető felhasználók eseményeinek 78,93 %ban fordul a 9es kódú esemény
- a db_10 és a 6 - cancelled user_rate keresztben a mező tartalma azt jelenti, hogy az előfizető felhasználók elemszámának 2,04 %ban fordul a 10es kódú esemény

Feltáró elemzések

Az előfizetők és a felmondók összehasonlítása 2

Category	1 - subscriber	2 - cancelled	3 - subscriber event_rate	4 - cancelled event_rate	5 - subscriber user_rate	6 - cancelled user_rate
db_08 egyéb felhasználói aktivitás	50 175,00	160 419,00	14,48	14,34	24,79	13,91
db_09 felhasználói aktivitás	273 599,00	862 759,00	78,93	77,14	135,18	74,79
db_10 bosszantó	4 492,00	23 534,00	1,30	2,10	2,22	2,04
db_11 hiba	2 069,00	9 571,00	0,60	0,86	1,02	0,83
db_12 előfizetés rejtett szándék	4 647,00	26 429,00	1,34	2,36	2,30	2,29
db_13 felmondás rejtett szándék	598,00	6 080,00	0,17	0,54	0,30	0,53
db_140 speciális szolgáltatások	5 492,00	14 725,00	1,58	1,32	2,71	1,28
db_15 TP extra	57,00	164,00	0,02	0,01	0,03	0,01

Az események vizsgálatánál a következő szempontok lehetnek fontosak:

- nettó elemszám: az adott eseménykategória a 2 alcsoport szempontjából vizsgálatra alkalmas elemszámmal rendelkezik-e
- a user_rate oszlopok azt mutatják, hogy egy-egy átlagos felhasználó esetén mennyiszer fordult elő az egyes kategóriák
- az event_rate azt jelenti, hogy az összes esemény szempontjából az adott kategóriának mekkora a gyakorisága, azaz a felhasználóban milyen erős érzékelést kelt.

Feltáró elemzések

Az előfizetők és a felmondók összehasonlítása 3

Category	1 - subscriber	2 - cancelled	3 - subscriber event_rate	4 - cancelled event_rate	5 - subscriber user_rate	6 - cancelled user_rate
db_08 egyéb felhasználói aktivitás	50 175,00	160 419,00	14,48	14,34	24,79	13,91
db_09 felhasználói aktivitás	273 599,00	862 759,00	78,93	77,14	135,18	74,79
db_10 bosszantó	4 492,00	23 534,00	1,30	2,10	2,22	2,04
db_11 hiba	2 069,00	9 571,00	0,60	0,86	1,02	0,83
db_12 előfizetés rejtett szándék	4 647,00	26 429,00	1,34	2,36	2,30	2,29
db_13 felmondás rejtett szándék	598,00	6 080,00	0,17	0,54	0,30	0,53
db_140 speciális szolgáltatások	5 492,00	14 725,00	1,58	1,32	2,71	1,28
db_15 TP extra	57,00	164,00	0,02	0,01	0,03	0,01

Megfigyelések:

- db_08/db_09: az előfizetőknel átlagosan nagyságrendileg 2x annyi eseménye van, mint a felmondóknak (135 <-> 75) , de az összes eseményeiknek nagyságrendileg ugyanakkora részét képezik ezen események (79<->77)
- db_10/db_11/db_12 az előfizetőkre átlagosan kicsit több ilyen esemény jut, de kevésbé érzékelik, hiszen az összes esemény arányában kisebb az előfordulás mint a felmondók esetén.
- db_13 : a felmondás rejtett szándék, mind a felhasználó darabszám arány esetén mind a esemény darabszám arány esetén erősebb a felmondóknál mint az előfizetőknel
- db_140: az előfizetők és a felmondók extra szolgáltatás használati aránya kb 2x-es az előfizetők javára
- db_15 kb 3x annyi előfizetőre jut TP motiváció, mint felmondóra (arányaiban)

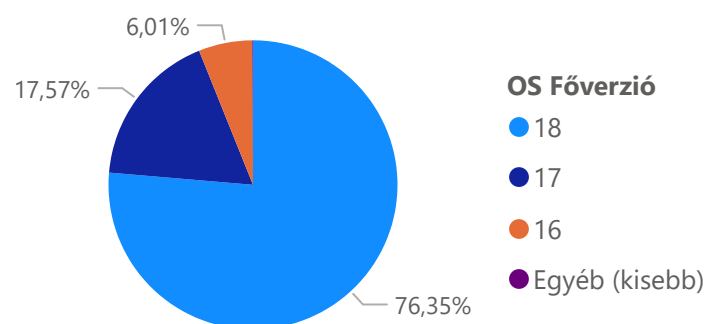
Feltáró elemzések

Használt operációs rendszer főverziók

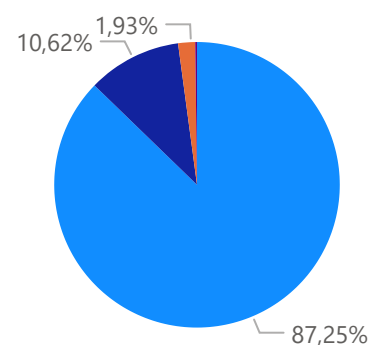
os_main_version	1 - subscriber	2 - cancelled	3 - subscriber - user_rate	4 - cancelled user_rate
12	2,00	4,00	0,05	0,03
14	1,00	1,00	0,05	0,01
15	1,00	1,00	0,05	0,01
16	39,00	776,00	1,93	6,73
17	215,00	2 167,00	10,63	18,79
18	1 765,00	8 585,00	87,25	74,43

Megfigyelhető, hogy az előfizetők esetén a modernebb rendszer a jellemzőbb

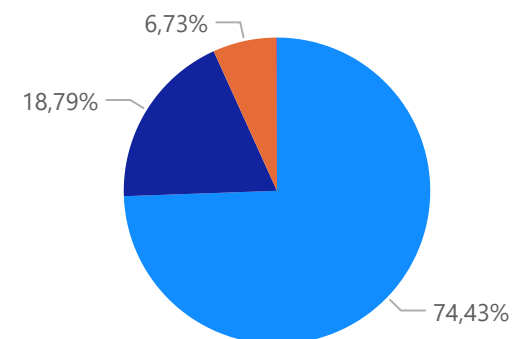
Összes



Előfizetők



Felmondók



Feltáró elemzések

Felhasználók lakóhelye: Felhasználók száma alapján

2024
Előfizetők száma
14,9%
Előfizetők aránya

11534
Felmondók száma
85,1%
Felmondók aránya

13558
Felhasználók száma

17,5%
Előfizetők/felmondók



☐ Előfizetők
☐ Felmondók

Continent ▾
☐ Africa
☐ Asia
☐ Europe
☐ North America
☐ Oceania
☐ South America

Top 10%user ▾
☐ Egyebek
☐ Top 10%

Top 10%:
Az összes országot
felhasználószám alapján
sorba rendezzük és a
felső 10%-ot kiválasztjuk.

Feltáró elemzések

Felhasználók lakóhelye : Milliő lakosra jutó felhasználó

2024
Előfizetők száma
14,9%
Előfizetők aránya

11534
Felmondók száma
85,1%
Felmondók aránya

13558
Felhasználók száma

17,5%
Előfizetők/felmondók



- ☐ Előfizetők
- ☐ Felmondók

Continent

- ☐ Africa
- ☐ Asia
- ☐ Europe
- ☐ North America
- ☐ Oceania
- ☐ South America

Top10%

- ☐ Egyéb
- ☐ Top 10%

Top 10%:
Az összes országot millió lakosra jutó felhasználószám alapján sorba rendezzük és a felső 10%-ot kiválasztjuk.

Feltáró elemzések

Felhasználás helyszíne : Összefoglaló

Érdekességek:

- 133 országban vannak felhasználók A világon 195 elismert ország van, tehát az országok közel 70%-ban jelen van a cég
- Afrikán kívül minden lakott kontinens képviselteti magát a TOP 10%-ban
- A legtöbb felhasználó az Egyesült Államokból való, őt közvetlenül Indonézia követi
- A Top 10 országban a felhasználók kb 80%-a koncentrálódik, kb ez igaz az előfizetőkre is.
- Észak Amerikában a legjobb az Előfizető/Felmondó arány, ez 44,3%
- Ázsiában a legrosszabb az Előfizető/Felmondó arány, ez 4,5%
- A legjobb és a legrosszabb Előfizető/Felmondó arány között 10 szerez szorzó van!
- 47 olyan ország van, ahol kipróbálták, de senki sem fizetett elő
- 2 olyan ország van, ahol minden felhasználó előfizető: Uruguay és Timor-Leste (1-1 user)
- 15 olyan ország van, ahol 33%-nál erősebb a koverziós ráta
- Lakosságarányosan a legtöbb előfizető Nuaru-ban van, a 9122 lakosra 193 felhasználó jut és ebből 59 előfizető!
- Oroszországban és Indonéziában lesz a legkevesebb előfizető az összes userből (azon országok közül, ahol lett Előfizető is)

Feltáró elemzések

Az applikációban eltöltött aktivitás vizsgálata

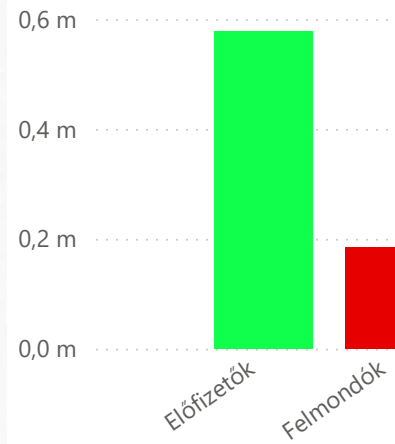
Képzett változók:

- `darab`: felhasználónként a kódolt interakciók száma
- `valodi_darab`: felhasználónként a kódolt interakciók száma közül a valódi interakciók darabszáma. Nem valódi interakció a 8/10/11/15ösre kódolt események, ezek vagy valódi események következménye (bezárja az ablakot), vagy az applikációnak valamilyen reakciója (hibaüzenet)
- `total_time`: a trial időszak megkezdése és az előfizetés illetve a felmondás között eltelt idő
- `time2`: az interakciókkal eltöltött idő. Az `app_open` és az `app_close` fogyelembevételével a 2 interakció közötti idő. Amennyiben 10 percnél több ez az idő, akkor feltételezzük az inaktivtást és ezt az időt már nem vesszük figyelembe
- `time_per_darab`: a teljes időszakra hány darab kódolt esemény jut
- `time_per_valodi_darab`: a valóban eltöltött idő (`time2`) és a valódi darabszám (`valodi_darab`) hányadosa
- `average_time`: felhasználónként az átlagos eltöltött idő
- `medián_all`: a "középső" idő
- `módus_all`: felhasználónként az 1 interakcióra jutó leggyakoribb felhasznált idő

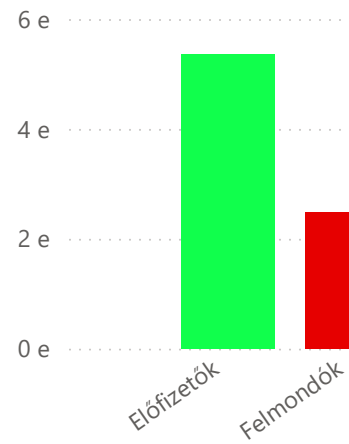
Feltáró elemzések

Az applikációban eltöltött aktivitás vizsgálata

total_time átlaga



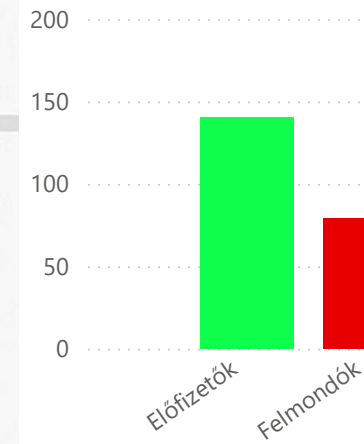
time2 átlaga



darab átlag



db_valodi átlaga



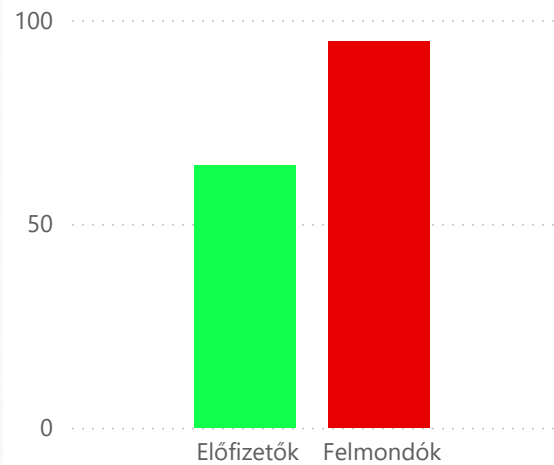
Megfigyelések:

- Megfigyelhető, hogy minden felsorolt mutató esetében igaz a bk 2x-es szorzó az előfizetők javára
- Az előfizetővé váláshoz hosszabb megfontolás szükséges, mint a felmondáshoz
- Az előfizetők több időt töltenek a rendszerben mint a felmondók
- Az előfizetők több interakciót hajtanak végre, mint a Felmondók

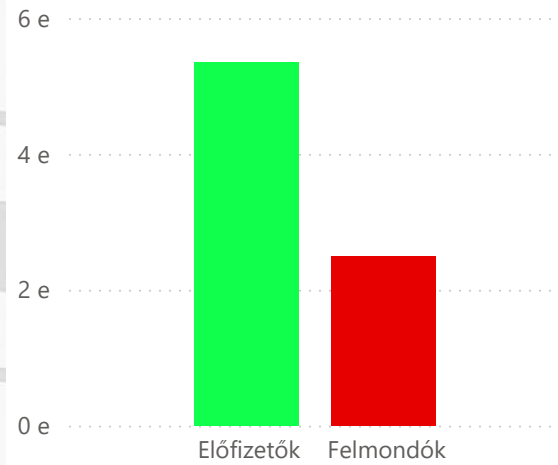
Feltáró elemzések

Az applikációban eltöltött aktivitás vizsgálata

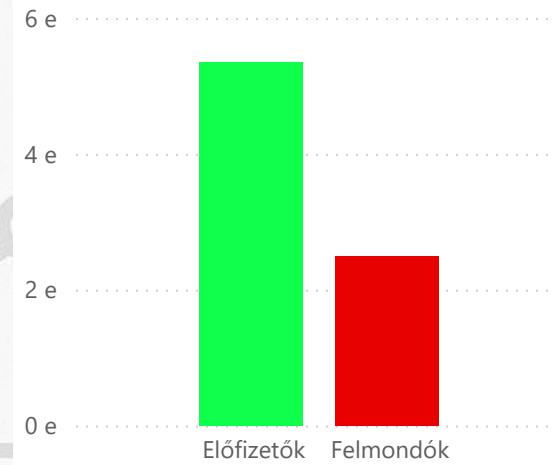
time2_per_valodi_darab



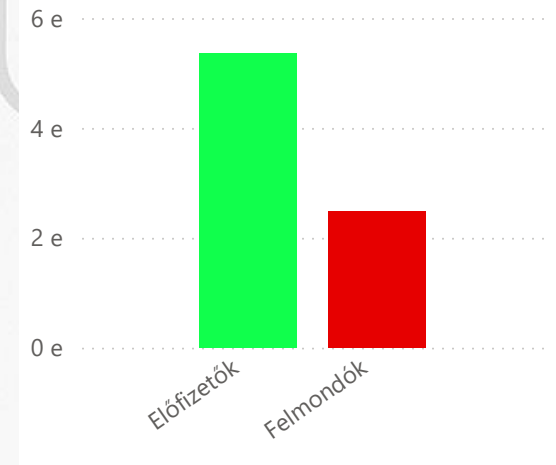
average_time átlaga



median_all átlaga



modus_all átlag



Megfigyelések:

- Megfigyelhető, hogy az average_time, a median_all és a modus_all esetében igaz a kb 2x-es szorzó az előfizetők javára
- Az előfizetők időegység alatt kb 1,5x annyi tevékenységet végeznek el.

Feltáró elemzések

Összefoglaló:

- Az Előfizetői aktivitás sokkal nagyobb mint a Felmondók aktivitása (db_08)
- A bosszantó dolgok és a hibák az Előfizetőket kevésbé befolyásolják mint a Felmondókat (db_10/db_11)
- Bár az alacsony elemszám miatt további vizsgálat lenne szükséges, de látszik, hogy az extra szolgáltatásokat az Előfizetők közül kb 2x annyian használják, mint a felmondók közül
- Bár az alacsony elemszám miatt további vizsgálat lenne szükséges, de látszik, hogy az Teleprompter motivációs lépések száma kb 3x annyi a Előfizetőknél, mint a Felmondóknál
- Minél frissebb az használt operációs rendszer, annál nagyobb az előfizetés valószínűsége
- A felhasználás helyének vizsgálatakor kiderül, hogy az applikációt a lakott világ nagyobbik részében használják
- A legtöbb felhasználó az Egyesült Államokból és Indonéziából használja a felületet
- Földrész szerint megfigyelhető, hogy a legtöbb felhasználó és a legtöbb előfizető Észak Amerikából, majd Európából érkezik
- Észak Amerikában a legjobb az Előfizető/Felmondó arány, ez 44,3%
- Ázsiában a legrosszab az Előfizető/Felmondó arány, ez 4,5%
- 2 országban van 100%os konverziós ráta: Uruguay és Timor-Leste (1-1 user)
- Az Előfizetők nem sajnálták az időt a kipróbálásra, szinte minden mutató szerint több időt szántak a programra, és több interakciót hajtottak végre, mint a felmondók.
- Az előfizetők átlagosan 1-1 interakciót kevesebb idő alatt teljesítettek

Feltáró elemzések

Összefoglaló

Következtetések:

Annál nagyobb az előfizetés valószínűsége:

- Minél kevesebb a "Bosszantó" dolog
- Minél kevesebb a "Hiba"
- Minél több extra szolgáltatást használ valaki
- Minél több motívációt kap a Teleprompter részéről a user
- Minél frissebb az használt operációs rendszer
- Ha a felhasználó Észak Amerikából, azon belül Egyesült Államokból tevékenykedik
- Ha a felhasználók gyorsabban meg tudnak tenni egy-egy teendőt ("Az előfizetők átlagosan 1-1 interakciót kevesebb idő alatt teljesítettek")

• Mit lehet tenni a konverziós ráta növeléséhez:

- Átgondolni a felületet, hogy a "Bosszantó" információk másként jelenjenek meg, vagy valóban megjelenjenek-e
- Tökéletesíteni a programot, hogy minél kevesebb legyen benne a hiba
- Minél szélesebb körben publikálni az extra szolgáltatásokat, úgymint AI segítség, 3rd party kapcsolatok. Kiterjeszteni a szolgáltatási kört Instagram, TikTok, stb. platformokra is.
- Érdekes Teleprompter kedvezményeket kitalálni, és szélesebb körben alkalmazni, amelyek motiválhatják a felhasználókat.
- A területi elemzésnél látszik, hogy a felhasználók a világ minden pontján jelen vannak, érdemes lenne koncentráltabb piacok felé fordulni, és belátni, hogy nem lehet mindenkinek megfelelni
- Kiemelten fontos, hogy a felhasználókat tanítani kell a program használatára, hogy a potenciális előfizetők is hatékonyabban meg tudjanak oldani egy-egy feladatot

Klaszterek kialakítása

Általános megközelítés

A további feldolgozás célja, hogy pontosabban próbáljuk meghatározni az Előfizetők és a Felmondók közötti különbségeket. Ennek érdekében a csoportképzés módszerét alkalmazzuk.

A csoportképzés/klaszterezés lényege, hogy egy adathalmazban automatikusan megtaláljuk az egymáshoz hasonló adatpontok csoportjait (klasztereket). A cél az adatok belső struktúrájának feltárása, ahol az egy klaszteren belüli adatpontok közötti hasonlóság magasabb, mint a különböző klaszterekben lévő pontok közötti hasonlóság - homogén csoportokat alakítunk ki.

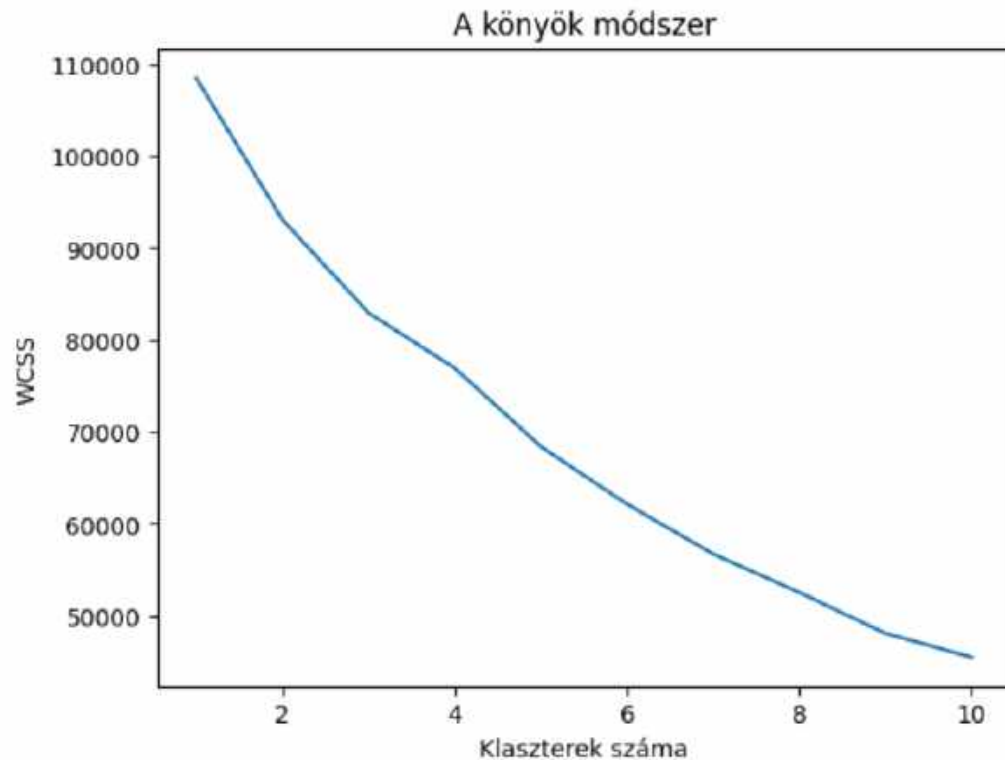
A különböző csoportképzési eljárások közül a K-Means módszert alkalmazzuk.

A klaszterben a felhasznált változók az előre kódolt valamilyen szempontból extra események (darabszámai), tehát db_10/db_11/db_12/db_13/db_140/db_15 változók, illetve a rendszerben eltöltött idő - total_time- és a munkával töltött aktív idő, azaz a time2.

Klaszterek kialakítása

Klaszterek száma, a könyök módszer

A módszer hátránya, hogy előre definiálni kell, hogy hány darab csoportot kívánunk képezni. A darabszám kialakításához az "Elbow method" azaz a könyökmódszer a bejáratott technika.



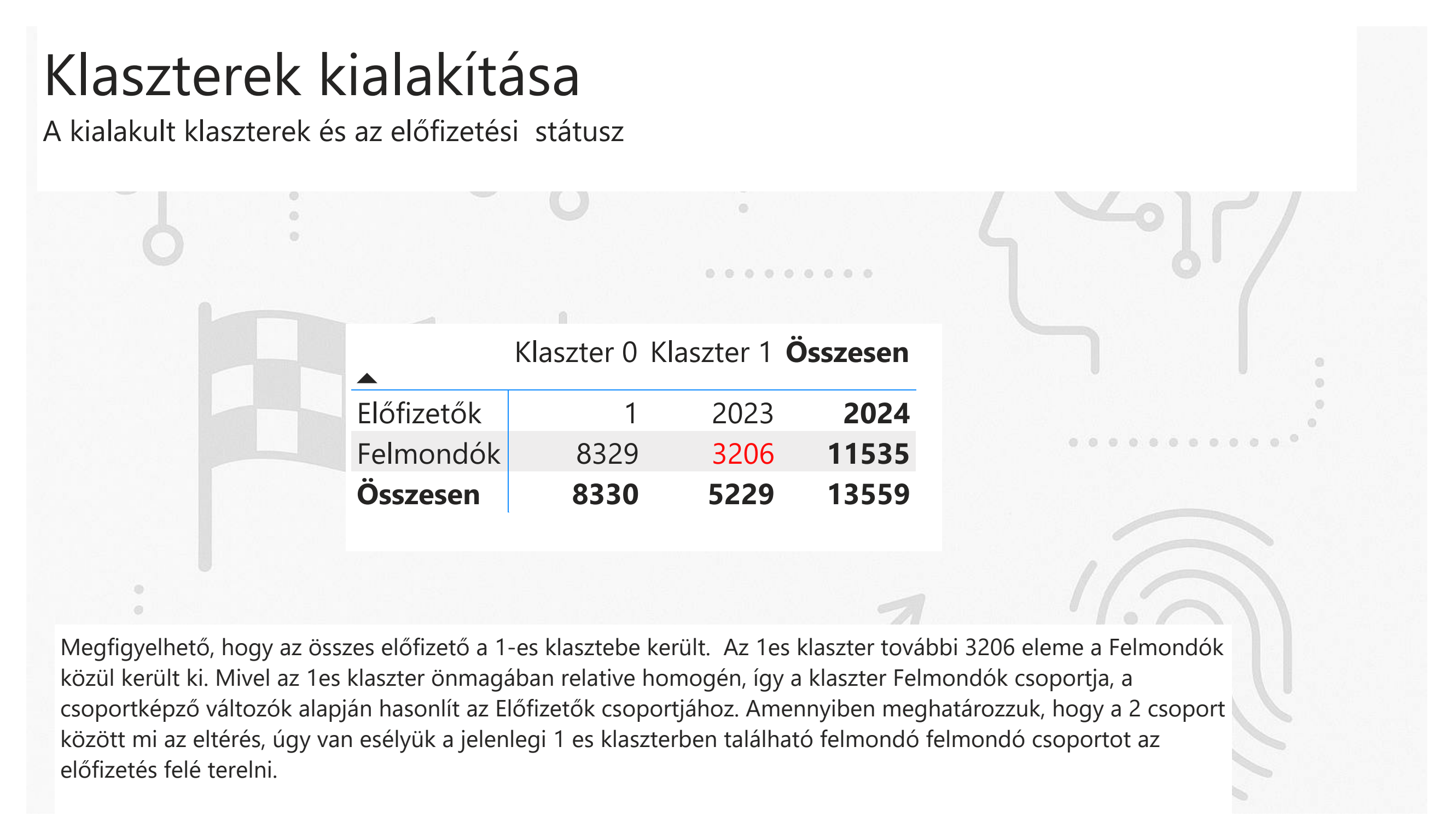
A könyökmódszer közérthetően:

Megnézzük, hogy ha egyre több csoportot képzünk, mennyire "javul" a csoportosítás minősége. Eleinte a minőség sokat javul, de egy pont után a javulás lelassul, mintha egy könyök lenne a grafikonon. Ez a "könyök" pont jelzi azt a csoportszámot, ami már elég jó ahhoz, hogy a dolgok ne zsúfolódjanak össze túlságosan, de még ne is legyen feleslegesen sok csoport.

Jelen esetben látszik, hogy a módszer nem ad tanácsot arra, hogy mi a megfelelő darabszám, ezért a 2 klasztert alakítunk ki.

Klaszterek kialakítása

A kialakult klaszterek és az előfizetési státusz



	Klaszter 0	Klaszter 1	Összesen
▲ Előfizetők	1	2023	2024
Felmondók	8329	3206	11535
Összesen	8330	5229	13559

Megfigyelhető, hogy az összes előfizető a 1-es klaszterbe került. Az 1-es klaszter további 3206 eleme a Felmondók közül került ki. Mivel az 1-es klaszter önmagában relative homogén, így a klaszter Felmondók csoportja, a csoportképző változók alapján hasonlít az Előfizetők csoportjához. Amennyiben meghatározzuk, hogy a 2 csoport között mi az eltérés, úgy van esélyük a jelenlegi 1-es klaszterben található felmondó felmondó csoportot az előfizetés felé terelni.

Klaszterek kialakítása

A kialakult klaszterek és az előfizetési státusz

	Klaszter 0	Klaszter 1
Előfizetők	1	2023
Felmondók	8329	3206

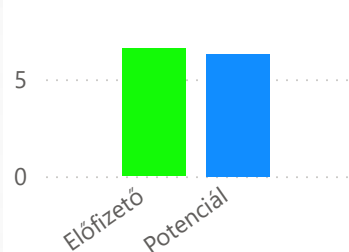
	Klaszter 0	Klaszter 1
Előfizetők	1 -> Kakukktojás	2023 -> Előfizető
Felmondók	8329 -> Elveszett	3206 -> Potenciál

Megfigyelhető, hogy az összes előfizető a 1-es klaszterbe került. Az 1-es klaszter további 3206 eleme a Felmondók közül került ki. Mivel az 1-es klaszter önmagában relative homogén, így a klaszter Felmondók csoportja, a csoportképző változók alapján hasonlít az Előfizetők csoportjához. Amennyiben meghatározzuk, hogy az 1-es klaszter 2 csoportja között mi az eltérés, úgy van esélyük a jelenlegi 1-es klaszterben található felmondó felmondó csoportot az előfizetés felé terelni.

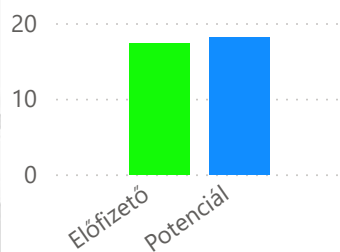
Előfizető + potenciális előfizető heti bontás

A kialakult klaszterek és az előfizetési státusz

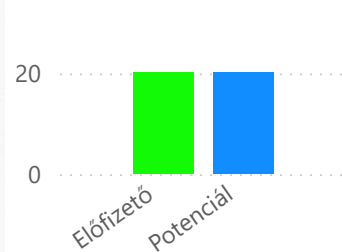
1. hét - események darabszáma szerint



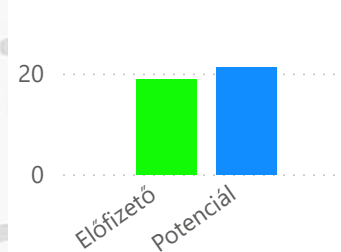
2. hét - események darabszáma szerint



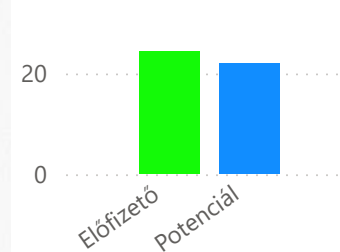
3. hét - események darabszáma szerint



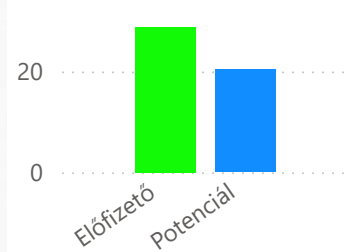
4. hét - események darabszáma szerint



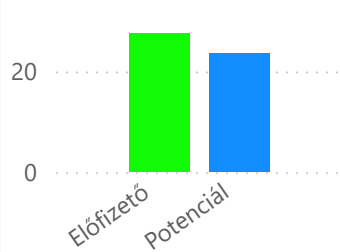
5. hét - események darabszáma szerint



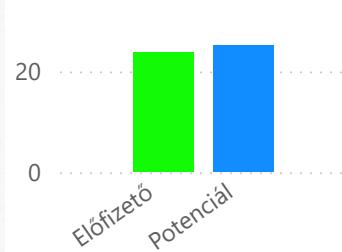
6. hét - események darabszáma szerint



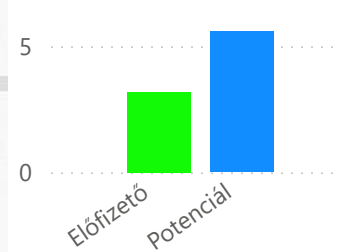
7. hét - események darabszáma szerint



8. hét - események darabszáma szerint



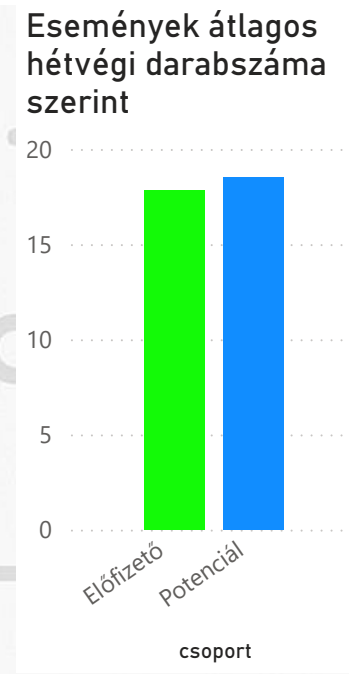
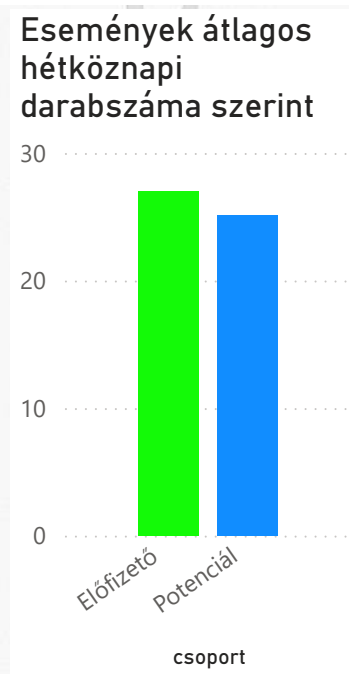
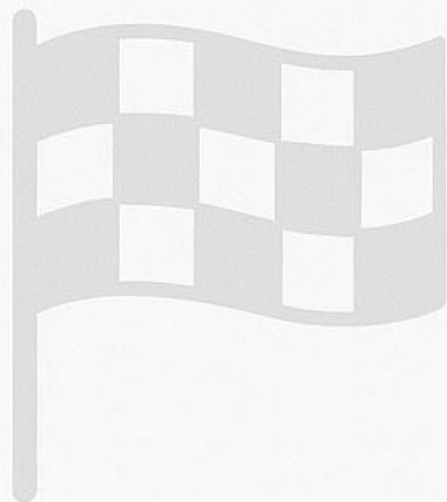
9. hét - események darabszáma szerint



Az ábrákon látszik, hogy az előfizetők átlagos esemény darabszáma az utolsó hét kivételével viszonylag egyforma. Az utolsó héten az előfizetők átlagosan 57%-al kevesebb eseményt generáltak, mint a Potenciális előfizetők. Mivel a 9. hét csonka hét, és ott csak a hétköznapok vannak, ezért a hétvége/hétköznap viszonyát érdemes vizsgálni.

Előfizető + potenciális előfizető

hétköznap/hétvége vizsgálata:

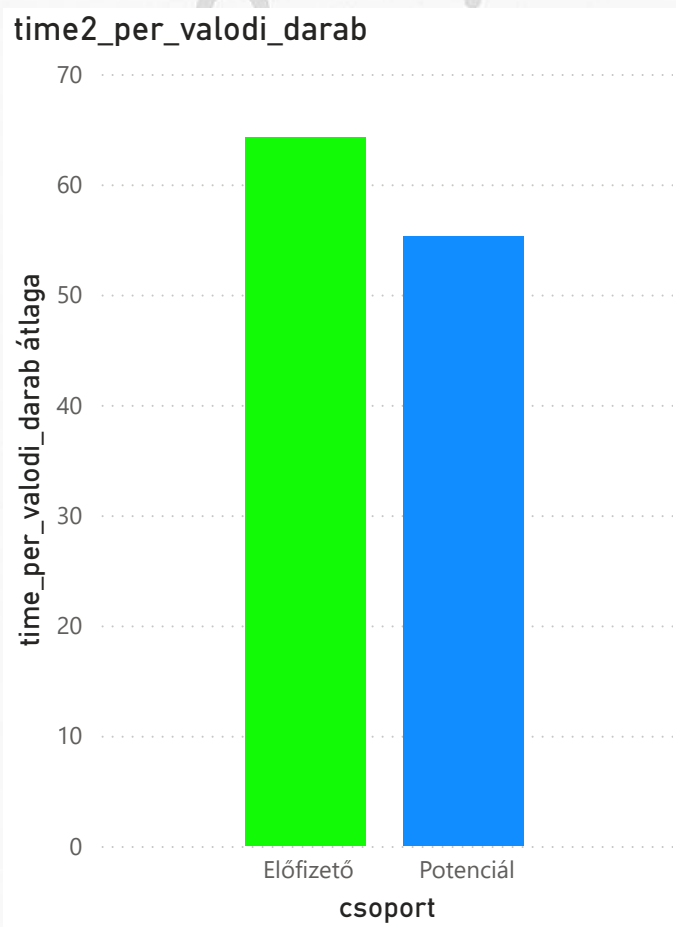


Eltérő felhasználási szokás sejlik fel, ha a hétköznap - hétvége bontásban vizsgáljuk a felhasználókat. A jelenlegi előfizetők aktívabbak a hétköznap, a Potenciál csoportja hétvégén.

Következtetés: A célcsoportunkat érdemes eltolni a hétköznapi használók felé. Itt a jövőben érdemes megvizsgálni, hogy az előfizetőink mekkora része profi, azaz a munkához használja az applikációt.

Előfizető + potenciális előfizető

időegység alatti események darabszáma



Megfigyelések:

A mellékelt ábrán az applikációban eltöltött "valódi" idő-re jutó "valódi" darabszám van. Látható, hogy az előfizetők esetén egy darab esemény átlagos ideje hosszabb, mint a Potenciál csoportnál.

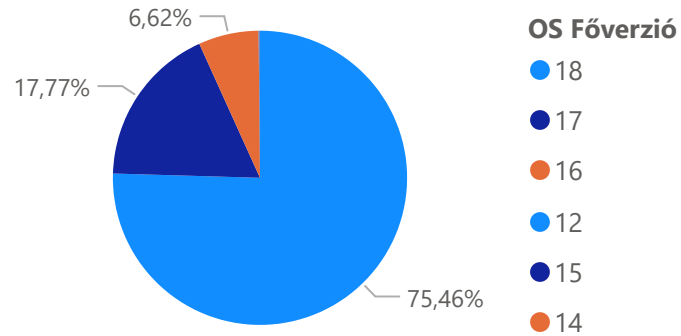
Következtetés:

- Az Előfizetők nagyobb energiát hajlandóak beleölni egy-egy esemény elkészítésébe, mint a Potenciál csoport.

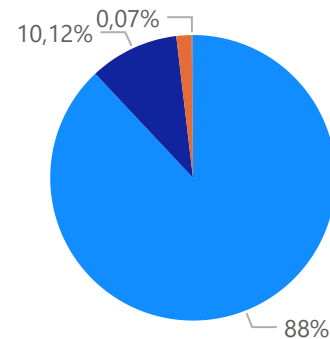
Előfizető + potenciális előfizető

Operációs rendszer

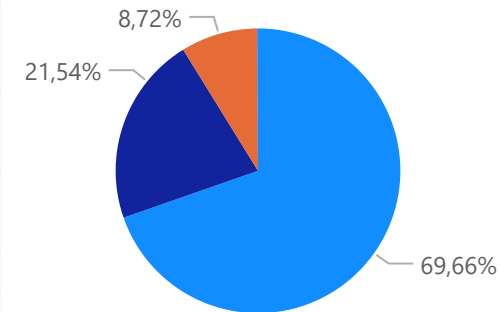
Összes



Előfizetők



Potenciál



Megfigyelhető, hogy az előfizetők esetén a modernebb rendszer a jellemzőbb

Következtetés:

Valami miatt a minél frissebb operációs rendszert használók nagyobb eséllyel válnak előfizetővé.

Érdemes belső teszteket végezni, és megvizsgálni a különböző rendszerek közötti eltéréseket, másrészt érdemes lehet marketingeszközöket bevetni, hogy a felhasználók frissítsék a rendszerüket.

Elemzések

Felhasználók lakóhelye: Felhasználók száma alapján

2023
Előfizetők

3206
Potenciál

63,1%

Előfizető/Potenciál



- ☐ Előfizető
- ☐ Potenciál

- Continent ▾
- ☐ Africa
 - ☐ Asia
 - ☐ Europe
 - ☐ North America
 - ☐ Oceania
 - ☐ South America

- Top 10%user2 ▾
- ☐ Egyebek
 - ☐ Top 10%

Top 10%:
Az összes országot
felhasználószám alapján
sorba rendezzük és a
felső 10%-ot kiválasztjuk.

Elemzések

Felhasználás helyszíne : Összefoglaló

Érdekességek:

- A legnagyobb lehetőség Ázsiában van, közel 10szer akkora a potenciál, mint az előfizetők száma
- Európa és Dél-Amerika esetén, a Potenciál körülbelül a jelenlegi előfizetők kétszerese
- Észak-Amerika 2,5 , Óceánia 1,5ször annyi előfizetővel rendelkezik, mint amekkora potenciál van még benne
- Indonéziában egyedül több mint 1300 lehetséges előfizető van
- Az Egyesült-Államokban ugyan 2,6szor annyi előfizető van, mint Potenciál, mégis a potenciális előfizetők száma az országok között a legnagyobb
- További erős potenciállal rendelkező országok: Brazília, Spanyolország, Franciaország és Németország
- Egyes Szovjet utódállamokban is erős potenciállal rendelkeznek: Oroszország, Ukrajna, Kazasztán
- Önálló kutatást érdemel Indonézia, ahol a legnagyobb a potenciális felhasználók száma
- Önálló kutatást érdemel Nauru ahol 9122 fős lakossághoz képest van 59 előfizető, és még 32 potenciál

Tanulságok és teendők

Első benyomások

A weboldal/youtube csatorna átnézésakor nem látszik, hogy a profi felhasználókat, avagy lelkes amatőröket akarjuk elérni.

A világ "minden sarkából" rendelkezik az appikáció felhasználókkal.

Az applikáció alverziókkal együtt több mint 59 iOS verziót támogat, és nyilván rengeteg android verziót is.

Nem látszik a célcsoport!

A jelenlegi stratégia rendkívüli mértékben pazarló az erőforrásokkal szemben, mindenképpen koncentrálni kellene kiválasztott tulajdonságokkal rendelkező felhasználói csoportokra!

Tanulságok és teendők

Széttöredezettség

Megfigyelhető, hogy a felhasználók a világ minden táján, nyelvén jelen vannak, "szegény" és "gazdag" országokban egyaránt.

Nem lehet mindenkinek megfelelni.

Célcsoportokat, jelen esetben célországokat kell kijelölni, és a rendelkezésre álló marketing és támogatási budgetet ehhez rendelni. Nem kell feltétlenül nagy dolgokra gondolni, de látszik, hogy egyes nyelveken (az angolon kívül) nagy a potenciál, és ha a weboldal és az applikáció támogatja a világnyelveket - német, spanyol, francia, portugál, orosz - akkor már a jelenlegi potenciális felhasználók nagyságrendileg harmada irányában nyitottunk.

Célországok: UK, USA, Franciaország, Németország, Spanyolország, Brazília, Mexico.

Oroszország és Ukrajna is sok potenciális lehetőséget tartalmaz, de a jelenlegi helyzet miatt ez kétszer átgondolandó (pl a nemzetközi bankrendszerrel való leválás).

Tanulságok és teendők

Egyedi vizsgálat: Indonézia és Nauru

Mind Indonézia, mind Nauru különleges helyet foglal el a feldolgozott adatok között.

Indonézia: 15 Előfizető és 1375 Potenciál

Indonéziában egyedül annyi lehetőség van, mit a teljes afrikai, európai és a dél-amerikai kontinensen. Elsősorban adatminőségi vizsgálatot kell tartani, mert ennyire kiugró számoknál ez kötelező feladat.

Amennyiben valóban ennyi lehetőség van az adott országban, akkor a jelenleg aktív felhasználókat meg kell keresni egy jól struktúrált kérdőívvel, tisztázandó, hogy milyen programot keresnek, mik a céljaik, és milyen szempontok alapján döntenek egyik vagy másik program mellett. Indonéziában elég sok nyelvet beszélnek, fontos, hogy a kérdések az angolon kívül az ország gyakoribb nyelvein legyenek feltéve.

Tanulságok és teendők

Egyedi vizsgálat: Indonézia és Nauru

Mind Indonézia, mind Nauru különleges helyet foglal el a feldolgozott adatok között.

Nauru: 95 Előfizető, 32 Potenciál és 9122 lakos (2021 évi népszámlálás)

A tökéletes célország!

Minden 100. lakos kipróbálta az applikációt, és ebből minden második előfizetővé vált.

Csak játék a számokkal, de ez az arány Magyarországon 50.000 előfizetőt jelentene....

Számoljunk csak havi 1000 Ft-al... Ugye? Melyik applikáció nem rakná össze a két kezét ezen számoktól??

Külön meg kell vizsgálni, hogy vajon hogyan sikerült ilyen komoly részesedést elérni a Nauru piacon, mi az ami miatt annyian értesültek a programról, és mitől felelt meg nekik annyira?

Ismét egy jól átgondolt kérdőívet javaslok....

Tanulságok és teendők

Extra szolgáltatások

Láttuk, hogy az Előfizetők gyakrabban használták az extra szolgáltatásokat mint a többiek.

Youtube, Facebook, TikTok, Instagram, X

Hangsúlyozni kell a program 3rd-party lehetőségeit, export - import funkciókat.

Az AI mint hívószó itt is erős lehet!

A szolgáltatáspalettát ki kell terjeszteni a jelenlegi Youtube-on és Facebook-on (mint a logokban könyvelt kapcsolatok) kívüli, népszerű gyakran használt platformokra is mint TikTok, Instagram, X.

Pontosabb célcsoportmeghatározás után esetleg a speciálisabb platformok felé is érdemes lenne fordulni mint a TED, a Udemy, Coursera. Sok esetben ha a leírásokban megjelennek ezek a kifejezések, már az eredményezheti az előfizetők számának növekedését.

Tanulságok és teendők

Idő és eseménydarabszámok

Az Előfizetők több időt fordítottak a programra, és több eseményt is generáltak. Megfigyelhető volt az is, hogy az Előfizetők egységnyi idő alatt több eseményt generáltak.

Ha megismeri, akkor használja! OKTATÁS!

Érdemes motiválni a próbálkozókat, hogy minél több időt töltsenek az applikációval. Aki a trial időszak alatt publikál videót Teleprompter logóval, hivatkozással, annak elengedni az első havidíját!

Az adatokból az derül ki, hogy a jobb megismerés nagyobb eséllyel indukál előfizetést.

Előfizetőknél egységnyi idő alatt több esemény!

Oktatóanyagokkal kell megtölteni a weboldalt és a kapcsolódó youtube csatornát!

Az applikációban érdemes linkeket elhelyezni, ami egy egy probléma megoldására vezet. (Apró nyomógomb, pl egy kérdőjel).

Tanulságok és teendők

Operációs rendszer, UX

A felhasználóknak hangsúlyozni, a friss operációs rendszer fontosságát.
A friss operációs rendszer használója nagyobb eséllyel lesz előfizető!

Nyilván minden programban van hiba, de ki kell dolgozni az ellenőrzési rendszer következő szintjét, hogy a meglévő kevés hibajelenséget is tovább csökkentsük.

Az előfizetők kevesebb hibát érzékeltek mint a felmondók!

Nemcsak az a hiba, ahol a program botlik, a felhasználó már tárterület betelítődést is annak érzékel. Ez az a pont ahol érdemes lehet a használati körülményeket/folyamatokat és a UX-t is megvizsgálni.

Amennyiben a felhasználónak bármilyen negatív élménye van, az kihat a előfizetési hajlandóságra is. (Isd korábban "Bosszantó" dolgok)

Tanulságok és teendők

Professzionális verzió a jelenlegi mellett

Az Előfizetők és a Potenciális előfizetők vizsgálatánál kiderült, hogy az előfizetők több időt dolgoznak hétköznap. Ugyan nem tisztázott, hogy helyi időben vagy UTC-ben voltak az adatok, de érdemes átgondolni:

Professzionális verzió!

Amennyiben a jelenlegi verzió mellett (amely a hétvégi, lelkes ámde amatőr felhasználók terepe lenne), a jelenlegire építve kidolgozásra kerülne egy professzionális verzió az újabb felhasználói csoportot vonna be, és drágábban lehetne értékesíteni.

A professzionális verzió célcsoportja a nagy hírszolgáltatók, televíziós csatornák, nagyvállalatok, amelyek munkához, hétköznap használják az alkalmazást.

A professzionális verzióban API kapcsolatokat lehetne kiépíteni az együttműködő partnerek saját hálózatához. Mivel ezen API-k kimenetét lehet standardizálni ez nem tűnik irreális befektetésnek a várható haszonhoz képest!

Tanulságok és teendők

Sine qua non: Márkanév weboldal és youtube csatorna

Saját másodlagos kvalitatív elemzésem során a következőre jutottam:

Új oldal, új logó, új elvek a Youtube csatornánál!

A weboldal, a márkanev, a Youtube csatorna egy olyan startup látszatát kelti, ahol a programozók összeálltak, és egy minél jobb applikációt csinálnak, ezt támogatják marketinggel, de olyan marketinggel amit a programozók tudnak... Nem eléggé szakmai!

A **weboldal**on sok a reklám, kevés az információ. Nem derül ki, hogy a program, hány nyelvet támogat, amely információ ez esetben minimálisan szükséges. Próbáltam leírást keresni, hogy hogyan is kell használni az alkalmazást, de az derült ki gyorsan, hogy magabiztos leszek, életérzést sugároz, de kevés az információ!

A sötét háttér inkább negatív érzetet sugároz!

Az a **márkanév** választás, ahol meg kell említeni, hogy ez nem az, mint a hasonló nevek, nem szerencsés választás. Nyilván egy ismert márkanev lecserélése nem kis feladat, de esetleg egy elő/utótaggal sikert lehetne elérni. (emlékszik még valaki más is, hogy milyen jó húzás volt amikor a Word-ből Microsoft Word lett?)

A **logó** sötét jellegtelen, és komoly gondolkodás után derül ki, hogy mire is utal.

Profi információdús weboldal, oktató videók tömtelege a Youtube-on, és pl Telepropter Studio név!



Nem lehet befejezni, csak abbahagyni...