



# **Telepromter adatok elemzése – 3. Házi Data Science Verseny**

Készítette: Jákob Tamás

Rendező: [dataklub.hu](http://dataklub.hu)



# Az elemzés célja

A Telepromter csapatának, mint megrendelőnek olyan következtetések szállítása, amelyek üzleti előnyként hasznosíthatóak:

- Megérteni a felhasználók döntéseinek okait
- Optimalizálhatják a marketing költségeket
- Megadhatják a terjeszkedés irányát
- Az egyes funkciók fejlesztésének fontosságát irányíthatja





# Tartalom

01

## Terület alapú elemzés

Az egyes felhasználók előhelye szerinti elemzések.

02

## Viselkedés elemzés I.

A felhasználókhöz köthető események elemzése statisztikai módszerekkel.

03

## Viselkedés elemzés II.

A felhasználókhöz köthető események elemzése gépi tanulási módszerrel.

04

## Következtetések összefoglalása

Lehetséges üzleti előnyökkel járó információk összegyűjtése.





01

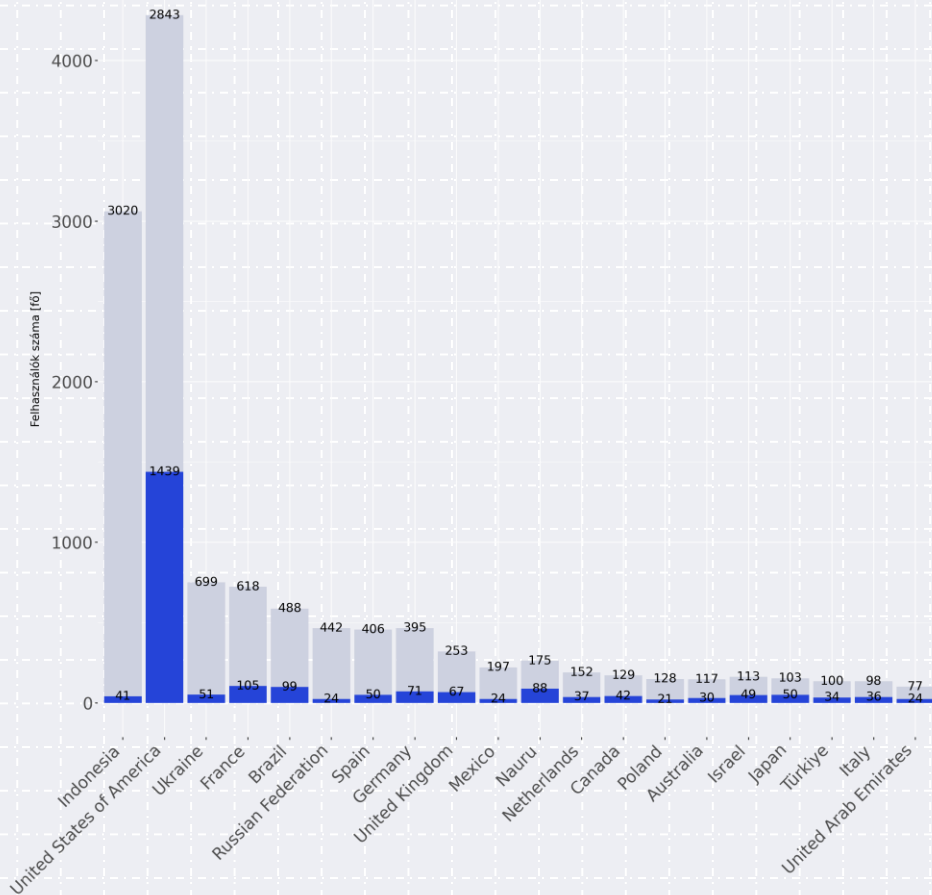
# Terület alapú elemzés

Milyen üzleti előnyök következnek az adatokból?



# A felhasználók terület alapú eloszlásának vizsgálata

Országok szerinti cancel vs. convert az ismert országokra



Feladat: az eltérő konverziós rátára magyarázatot találni.

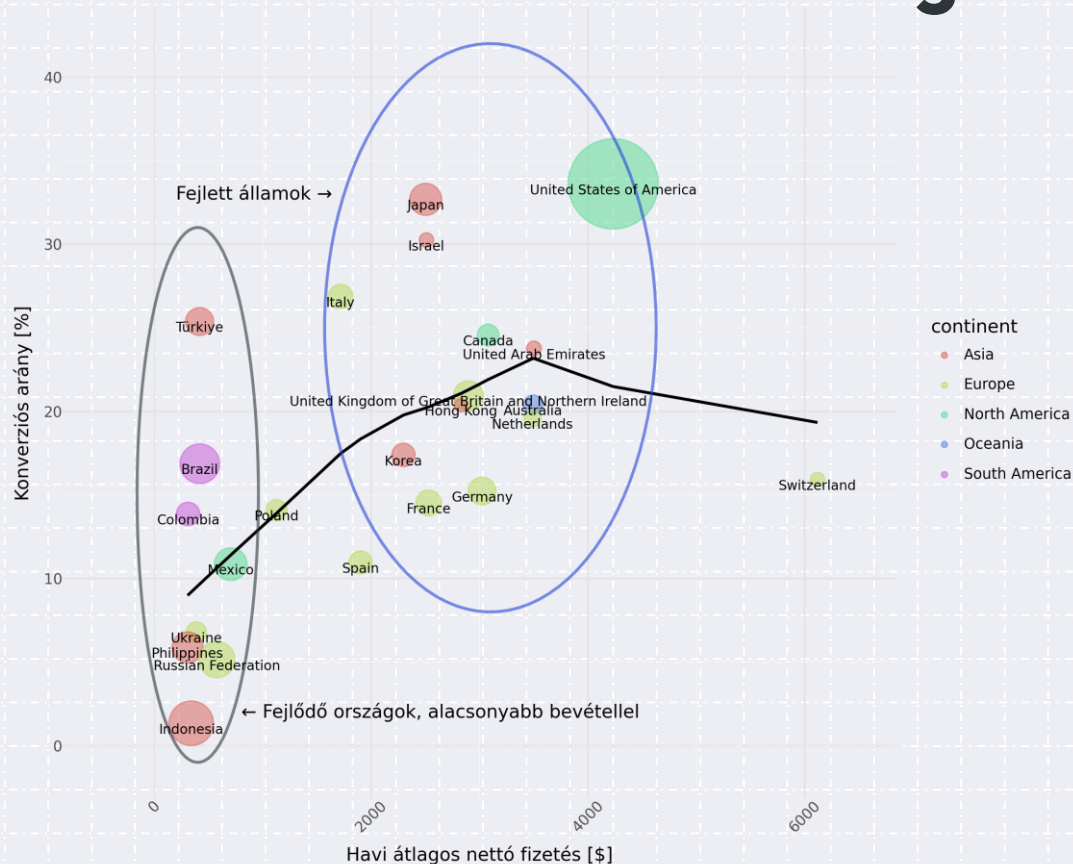
Hipotézis: az előfizetési hajlandóságnak köze van az adott felhasználó jövedelméhez?

**Felhasználók, akik lemondták az előfizetést**

**Előfizető felhasználók**

# Hipotézis ellenőrzése nem paraméteres regresszióval

Országokénti konverziós arány az egy főre eső  
jövedelm függvényében



Korreláció figyelhető meg a fizetések  
és a konverziós arány között.

**Európa nem jól „teljesít”.**

**Észak-Amerika átlagon  
felüli bevételt hoz!**

# Szélsőséges értékek az adatokban



## USA

A legnagyobb felhasználói bázissal rendelkező ország, további fejlesztések és hirdetések leendő célja.

## Törökország

Pozitív példa, sok előfizető felhasználó.

## Indonézia

Nagyon sok felhasználó próbálta ki az alkalmazást, de itt a legalacsonyabb a konverziós arány.

## Svájc, Németország, Franciaország

Az életszínvonalhoz képest nagyon alacsony felhasználói bázis.





02

# Viselkedés elemzés I

Milyen megfigyelések vonhatóak le a felhasználók viselkedéséből?

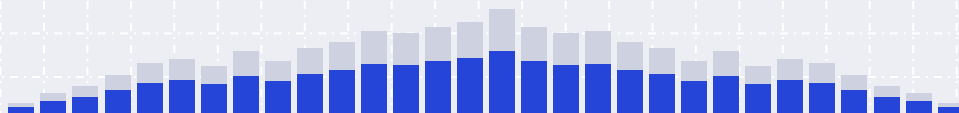




# Userflow a fő eseményekre

Esemény	Trial_started	Billing_issue	Trial_cancelled	Trial_converted
Trial_started	4	961	13893	2779
Billing_issue	708	7	<u>1336</u>	<u>12</u>
Trial_cancelled	2604	1329	17	20
Trial_converted	2020	105	287	56

Legnagyobb töréspont: billing\_issue\_event!  
Technikai probléma?





# Előfizetők és lemondók közötti különbségek



## Hipotézis:

Eltérés van a két csoport viselkedési mintázata között.

**Nehezítő körülmény:** A lemondók nagy része ki sem próbálta az alkalmazást, így az eseményeik mérőszámaik nagy része 0.

**Legjelentősebb szűrő:** Billing\_issue\_event!



Mann-Whitney U-teszt eredménye:

## Előfizetők esetében domináns események:

Dyslexic fontok használata, felirat stílusok, AI funkciók, hangfelismerés

## Lemondók és előfizetők között nem eltérő események:

4K felvétel, mirroring, remote funkciók, és minden általános videófelvételi funkció





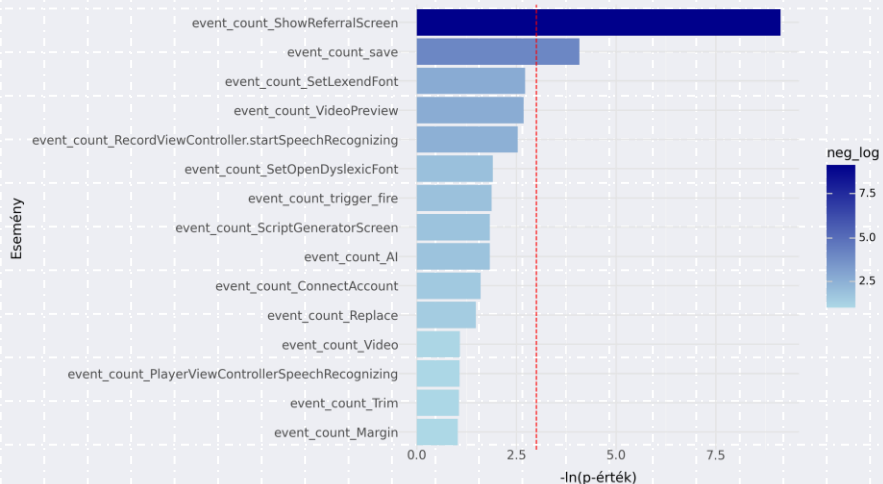
# Egyesült Államok és Indonézia: különbségek

**Megfigyelés:** a két ország konverziós aránya a legszélsőségesebb

## Előfizetők összehasonlítása:

(nincs különbség)

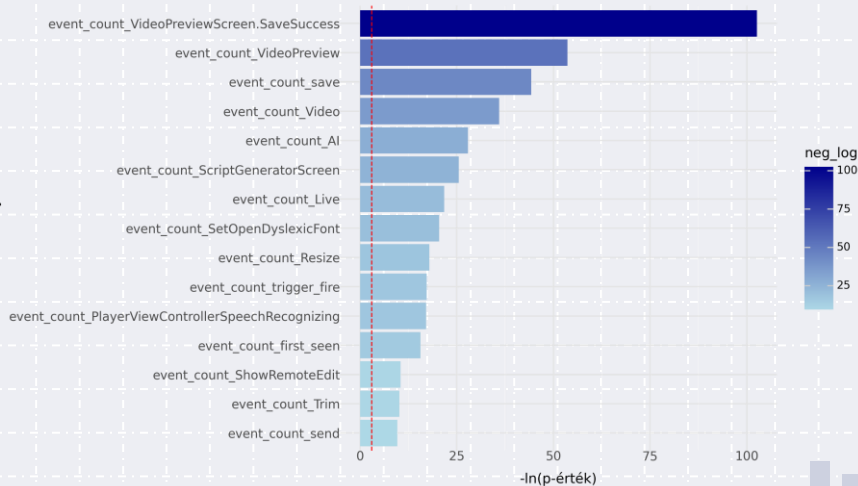
Top 15 eltérő esemény az előfordulási mennyiség szerint  
USA vs Indonézia előfizetői között.

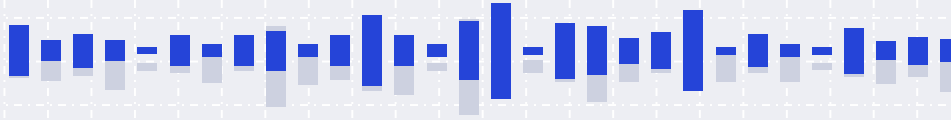


## Lemondók összehasonlítása:

(nagy különbség van!)

Top 15 eltérő esemény az előfordulási mennyiség szerint  
USA vs. Indonézia lemondói között





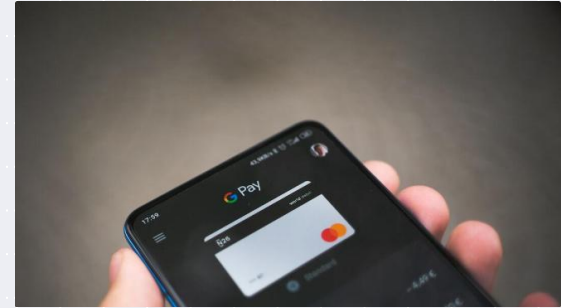
### **Előfizetők:**

Nincs lényegi különbség az előfizetők között

### **Lemondók:**

Az indonéz lemondók sokat használták az alkalmazás fő funkcióit.

Rendszeres próbaverzió újraindítás, vásárlás helyett?

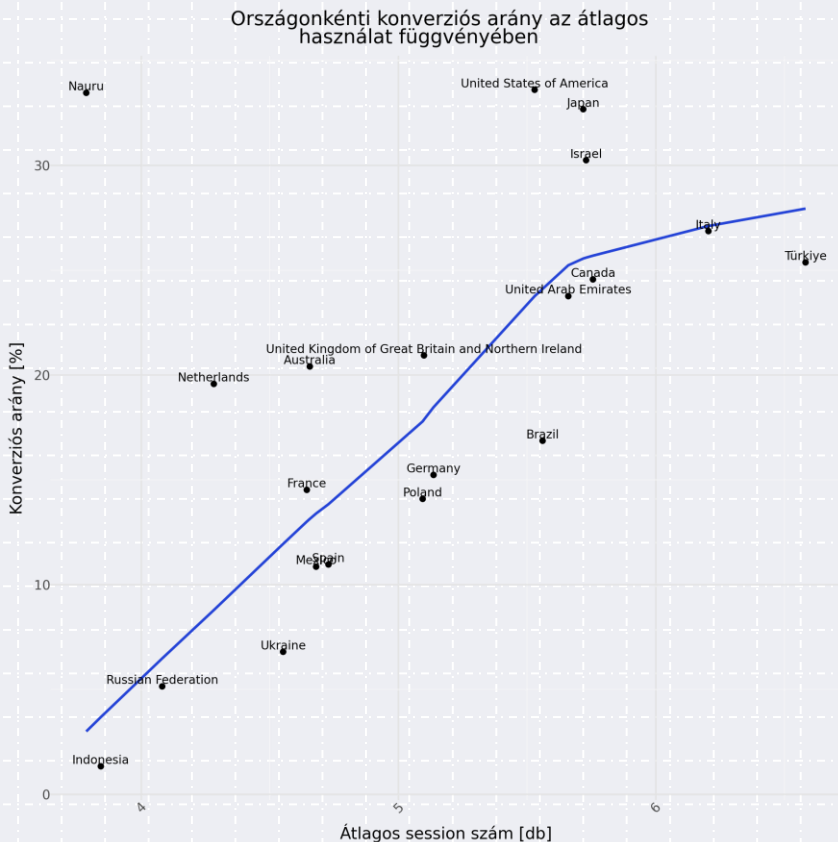


### **Javaslat:**

- Jelentős árcsökkentés.
- GDPR miatti adattörlés „kiszedése” az ugyanazon felhasználók ismételt, sorozatos próbaidőszakjainak megakadályozására.



# A felhasználók terület alapú eloszlásának vizsgálata a használat mennyisége szerint



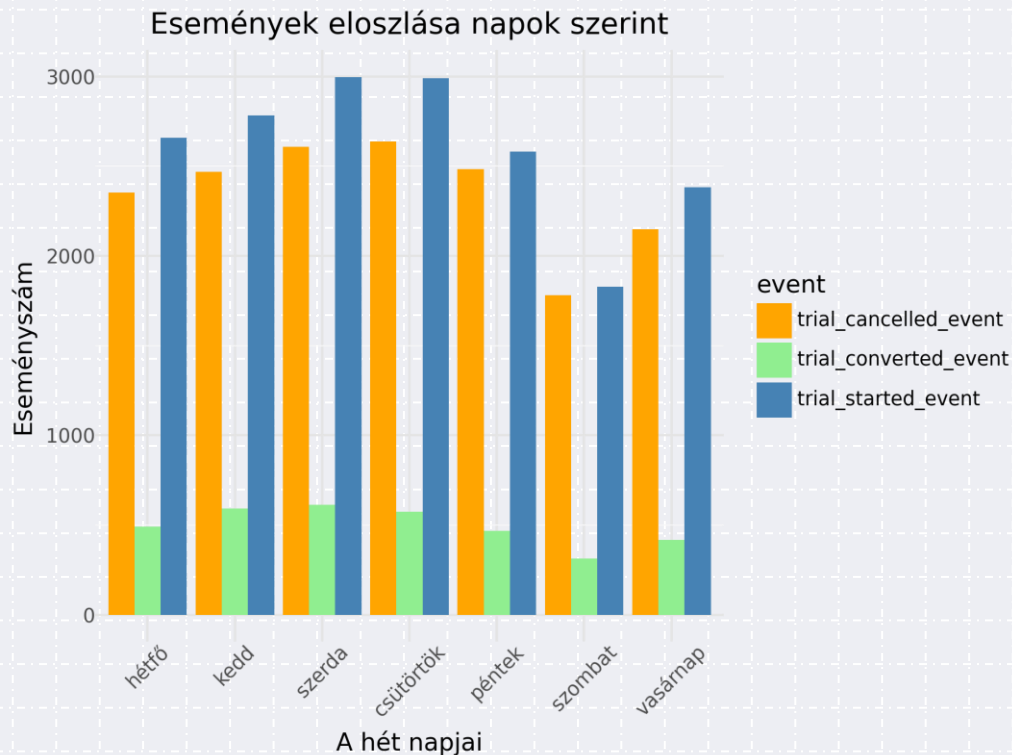
Új session:

30 perc tétlenség után

**Erős pozitív korreláció !**

**A konverziós hajlandóság  
függ a használat számától.**

# Felhasználók előfizetési hajlama a használat napja alapján



**Nagyon enyhe hullámmzés.**

Befolyásolható a felhasználó marketing eszközökkel, hogy mikor kezdje el használni az alkalmazást?

**Kedd-szerda-csütörtök a legoptimálisabb kezdés!**



03

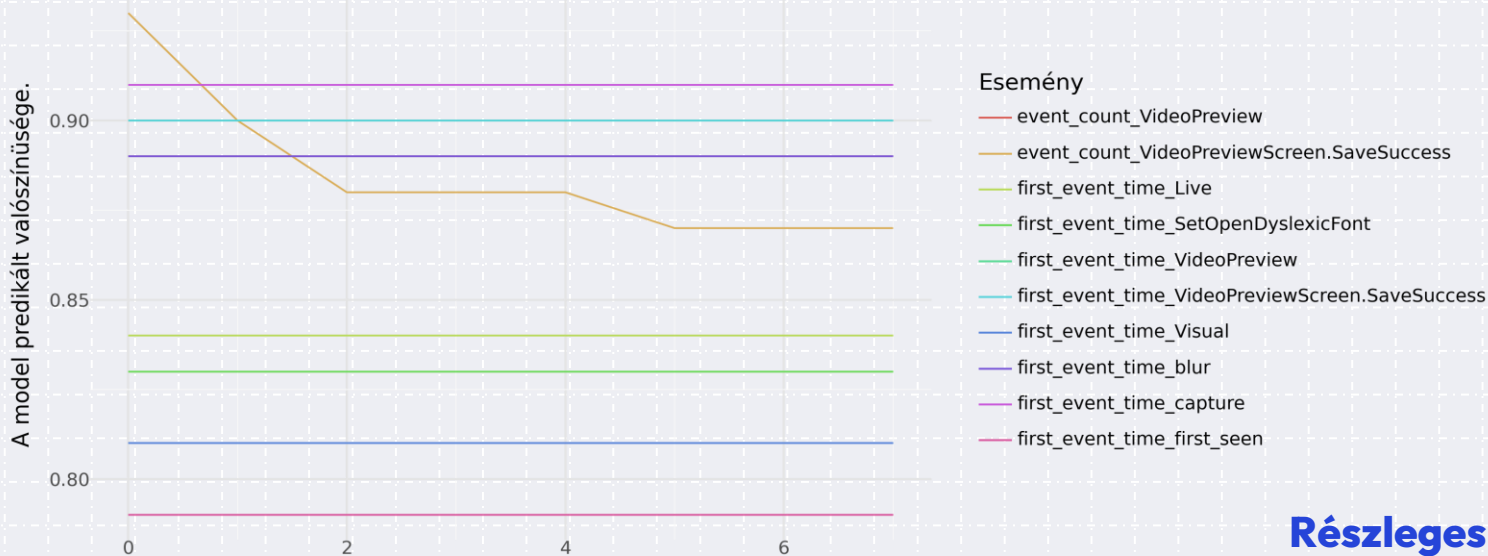
# Viselkedés elemzés II

Milyen megfigyelések vonhatóak le a felhasználók viselkedéséből?



# Felhasználók viselkedésének elemzése gépi tanulási módszerrel I.

A top10 esemény időbeli előfordulásának  
hatása a konverziós rátára.



Az esemény hányszor történt meg,  
vagy hanyadik napon történt meg először.

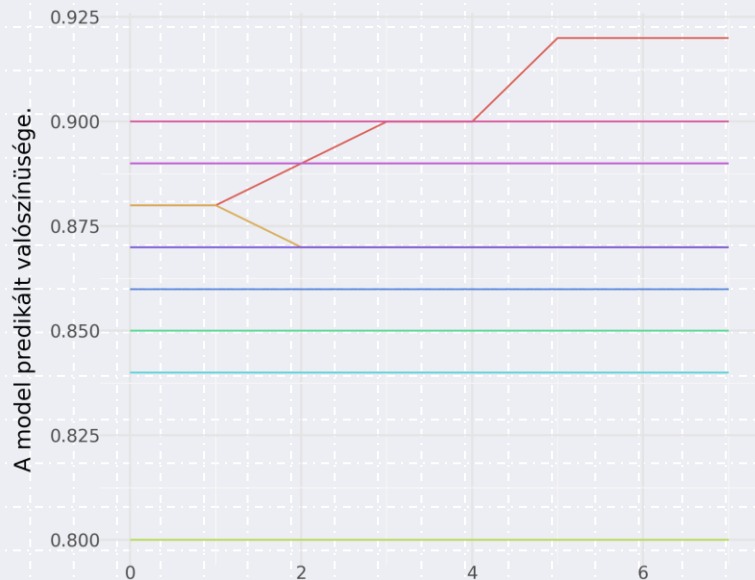
**Részleges hatásdiagramm**

Sem a használat száma, sem ez egyes funkciók utáni  
érdeklődés mértéke (első használat ideje) nem befolyásolja  
az előfizetés esélyét.



# Felhasználók viselkedésének elemzése gépi tanulási módszerrel II.

A top 10-20 esemény időbeli előfordulásának hatása a konverziós rátára.



Az esemény hányszor történt meg, vagy hanyadik napon történt meg először.

Esemény

- event\_count\_SetOpenDyslexicFont
- event\_count\_save
- first\_event\_time\_AI
- first\_event\_time\_Intercom
- first\_event\_time\_PlayerViewControllerSpeechRecognizing
- first\_event\_time\_RecordViewController.startSpeechRecognizing
- first\_event\_time\_Remote
- first\_event\_time\_ScriptGeneratorScreen
- first\_event\_time\_SubtitleStyleSelectorViewController.styleSelected
- first\_event\_time\_save

**Részleges hatásdiagramm**

Sem a használat száma, sem ez egyes funkciók utáni érdeklődés mértéke (első használat ideje) nem befolyásolja az előfizetés esélyét.

# Felhasználói események Spearman korrelációi

values	event_name
1.000000	target
0.111320	event_count_VideoPreview
0.097001	first_event_time_SetOpenDyslexicFont
0.093540	event_count_SetOpenDyslexicFont
0.091435	event_count_first_seen
0.091262	first_event_time_first_seen
0.088217	event_count_SubtitleStyleSelectorViewControll...
0.087227	first_event_time_SubtitleStyleSelectorViewCont...
0.082314	event_count_AI
0.082034	event_count_ScriptGeneratorScreen
0.081946	first_event_time_Onboarding
0.081699	event_count_Onboarding
0.081517	first_event_time_ScriptGeneratorScreen
0.081396	first_event_time_AI
0.080193	first_event_time_Live
0.079034	first_event_time_Visual
0.077480	event_count_Live
0.068379	first_event_time_SetLexendFont
0.067248	event_count_Visual
0.066457	event_count_SetLexendFont

**Események számának  
(funkcióknak) vagy  
használati idejüknek  
gyakorlatilag nincs  
befolyásuk az előfizetői  
hajlandóságra.**



04

# Összefoglalás

Az adatokból levonható következtetések összegzése.



# Konklúziók

## Legfontosabb felhasználói jellemző:



Lokáció és pénzügyi helyzet.  
Marketing költségek optimalizálhatóak.



Svájc, Kanada, Ny-Eu országok



Indonézia

## Funkciók az alkalmazásban



Nem emelhető ki egyetlen funkció sem.



AI, beszédfelismerés



4K felvétel

## „Billing\_issue\_event”



Legnagyobb lemorzsolódás.

## Dolgos hétköznapiak



Keddtől csütörtökig hajlamosabbak előfizetni a felhasználók.

The slide features a light blue background with a white dashed grid. On the left and right sides, there are decorative elements consisting of numerous horizontal bars of varying lengths and shades of blue and gray, creating a layered, architectural effect.

# Köszönöm a figyelmet!

**Készítette:**

**Jákob Tamás**

tamas.jakob@gmail.com