

ניתוח נתוני שירותי סטרימינג

חקר נתונים, התפלגויות וקורלציות בשוק הסטרימינג העולמי





רקע על הדאטה (Dataset Background)

תיאור הדאטה

הדאטה עוסק בשירותי סטרימינג וכולל נתונים ברמת שירות/פלטפורמה, כגון:

- זמינות במדינות שונות (Countries Available)
- מחירים חודשיים ושנתיים (Monthly/Annual Price)
- מספר מנויים (Subscribers)
- מדדים נוספים התומכים בניתוח עסקי והתנהגותי

מטרת הניתוח

לבצע חקר נתונים (EDA), לבחון התפלגויות מרכזיות ולבדוק קשרים סטטיסטיים (קורלציות) בין משתנים עסקיים—במטרה להפיק תובנות תפעוליות ועסקיות.

גרנולריות (Granularity)

רמת הגרנולריות גבוהה: כל שורה בטבלה מפרטת על שירות סטרימינג אחד



ניתוח מדינות

בעמודת המדינות (countries_available)
ייתכן ערך מרובה מדינות לשירות אחד—לכן
לצורך ניתוח התפלגות מדינות ניתן לפרק
את השדה לרמת "מדינה-שירות"



חישובי התפלגות

מאפשר חישובי התפלגות, דירוגים וחישובי
קורלציה בין מאפייני שירות שונים



השוואה בין שירותים

מאפשר השוואה בין שירותים
(Benchmarking)

חקר ראשוני, ניקוי ובדיקות איכות (EDA + Data Validation)

הנדסת פיצ'רים / עמודות עזר

הוספת עמודת ייחוס לממוצע מחיר חודשי:

```
df_streaming["month price avg"] =  
df_streaming["monthly_price_usd"].mean()
```

עמודה זו שימשה כקו השוואה/Benchmark בניתוחים וויזואליזציות.

בדיקות איכות שבוצעו (כפי שבוצע בקוד)

- בדיקת ערכים חסרים: `df_streaming.isnull().values.any() +`
`df_streaming.isnull().sum`
- בדיקת מבנה וטיפוסי נתונים: `df_streaming.info`
- בדיקת כפילויות: `df_streaming.duplicated().any`

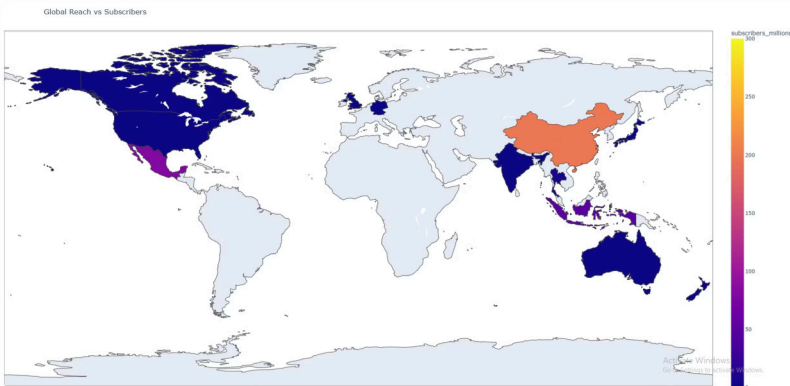


סיכום איכות הדאטה: לאחר הבדיקות והסקירה הראשונית, הדאטה נמצא **תקין לניתוח** (ללא ערכים חסרים/כפילויות לפי בדיקותינו), ולכן המשכנו לשאלות המחקר והקורלציות.

התפלגות שירותי סטרימינג לפי מדינה

שאלת מחקר

מהי התפלגות המדינות בהן קיימים שירותי סטרימינג, ואיזו מדינה מובילה במספר השירותים?



33%

הודו

המדינה המובילה במספר שירותי סטרימינג

ממצא מרכזי

המדינה עם הכי הרבה שירותי סטרימינג בדאטה היא **הודו**, עם כמעט **שליש** מכלל השירותים בטבלה.

מסקנה

הודו מהווה שוק סטרימינג גדול ודומיננטי, עם ריבוי פלטפורמות וזמינות גבוהה.

שירות הסטרימינג הפופולרי ביותר לפי מספר מנויים

שאלת מחקר

מי שירות הסטרימינג הפופולרי ביותר מבחינת מספר מנויים?

300M

MX Player

מנויים רשומים

ממצא מרכזי

השירות הפופולרי ביותר הוא **MX Player** עם כ-**300 מיליון** מנויים רשומים.

מסקנה

מוביל בצורה בולטת בהיקף משתמשים, ומייצג שירות בעל אימוץ גבוה וחדירה רחבה לשוק MX Player.

האם מחיר חודשי משפיע על מספר המנויים?

שאלת מחקר

האם יש קשר/קורלציה בין המחיר החודשי לבין מספר המנויים?

השערה

ציפינו לקשר (לדוגמה: מחיר גבוה → פחות מנויים), ולכן ציפינו לקורלציה משמעותית.

מסקנה

אין קשר מתאמי מובהק בין מחיר חודשי למספר מנויים בדאטה שלנו.

0.05

קורלציה

ממצא מרכזי

קורלציה חיובית: מחיר חודשי מול מחיר שנתי

שאלת מחקר

האם המחיר החודשי "משפיע" על המחיר השנתי (כלומר, האם יש ביניהם קורלציה)?



ממצא מרכזי

נמצאה קורלציה חיובית חזקה מאוד.

הסבר

המחיר השנתי נגזר מהמחיר החודשי, ולכן הקורלציה צפויה ואף "כמעט מושלמת".

מסקנה

הנתונים עקביים מבחינה לוגית: תמחור שנתי מתיישר באופן ישיר עם התמחור החודשי.

תכנים חדשים מול מחיר חודשי

שאלת מחקר

האם יש קשר בין מספר המנויים לבין שיעור/סיכוי נטישה?

השערה

שיעורנו שכל שיש יותר מנויים, ייתכן גם יותר נטישה (במונחים מוחלטים), ולכן ציפינו לקשר כלשהו.

ממצא מרכזי

נמצאה **קורלציה שלילית** בין מספר המנויים לבין שיעור הנטישה.

מסקנה

ככל שמספר המנויים גבוה יותר, שיעור הנטישה נוטה להיות נמוך יותר – מה שעשוי להצביע על יציבות גבוהה יותר בפלטפורמות גדולות.

ומנויים מול אחוזי נטישה

שאלת מחקר

האם "חדשנות בתוכן" (סרטים/סדרות חדשים) קשורה למחיר החודשי? האם שירותים עם יותר תכנים חדשים יקרים יותר?

השערה

ציפינו לקורלציה חיובית גבוהה (יותר חדש → יותר יקר).

ממצא מרכזי

נמצאה **קורלציה חיובית אך נמוכה**.

מסקנה

למרות הציפייה, לא נמצאה עדות לקשר חזק בין מדד התוכן החדש לבין המחיר החודשי בדאטה.

סיכום תובנות (Key Insights)

MX Player

MX Player (~300M) הוא השירות המוביל במספר המנויים MX Player.

הודו מובילה

הודו מובילה במספר שירותי הסטרימינג (כמעט שליש מהדאטה).

מחיר חודשי ושנתי

יש קשר חיובי חזק בין מחיר חודשי למחיר שנתי (נגזרת ישירה).

מחיר ומנויים

אין קשר מובהק בין מחיר חודשי למספר מנויים ($r \approx 0.05$).

יציבות פלטפורמות

נמצא קשר שלילי בין מספר מנויים לשיעור נטישה (פלטפורמות גדולות יציבות יותר).

תכנים חדשים

לתכנים חדשים לא נמצאה השפעה מתאמת חזקה על המחיר החודשי.