

## מבוא:

מטרת המסמך היא להציג תמונת מצב תיאורית ונקייה של נתוני סקר, תוך הדגשת דפוסים מרכזיים והבדלים בין תתי-אוכלוסיות. שלב ההכנה שבוצע במחברת Jupyter (Python) שימש לבניית **קובץ נתונים קוהרנטי ומוכן לניתוח**: איחוד עמודות כפולות/מפוצלות, תרגום שמות עמודות לעברית, טיפול שיטתי בערכים חסרים, ואחידות סמנטית של שדות דומים. על בסיס נתונים זה הופקו בהמשך ב-R-תרשימים טבלאות וסיכומי תובנות.

**מקור הנתונים וטעינה.** הנתונים נקראו ישירות מ-GitHub ונבדקו באופן ראשוני (דגימה, טיפוסים ושגיאות שכיחות). גישה זו מבטיחה עקביות ויכולת שחזור מלאה של כל צעד בעיבוד.

**איחוד מבני של עמודות.** במהלך איסוף הסקר נוצרו מצבים של עמודות כפולות או שאלות רב-ברירה שפוצלו למספר עמודות. במחברת בוצעו פעולות איחוד ממוקדות:

- איחוד שתי עמודות של **אותה שאלה** (למשל איחוד "שאלה 5-6" לכדי שדה אחד מסודר).
  - צמצום שאלות מרובות-בחירה לייצוג טקסטואלי אחיד (לדוגמה צירוף 12/9/8 אופציות נפרדות לשדה מרוכז אחד).
  - הסרת עמודות כפולות וריקות, וכן הורדת שדות לא נדרשים (כגון-Q8).
  - טיפול במקרים של **אותו שדה באותו שונה** (למשל "ספק טלוויזיה בינלאומי" מול "ספק טלוויזיה בינלאומי") באמצעות מיזוג ערכים ושימור העמודה האחודה בלבד.
- תרגום ואחידות שמות עמודות.** לטובת קריאות ותמיכה בעברית בכל שלבי ההמשך, הופעל מילון מיפוי שהחליף שמות טכניים בשמות עבריים ברורים ועקביים (למשל: סוג חיבור האינטרנט, ספק האינטרנט, ספק טלוויזיה ישראלי, ספק טלוויזיה בינלאומי, צריכת שירותי מוזיקה וכד'). שלב זה קריטי כדי שכל הגרפים והטבלאות יהיו נהירים ללא תיווך.
- ניתוח ערכים חסרים.** בוצעה סקירה שיטתית של שיעורי חסר לכל עמודה (מנייה ואחוזים), כדי לזהות מוקדי סיכון בניתוח ולבחור אסטרטגיית השלמה מותאמת לשדה:

- הכנסה מעל/מתחת לממוצע:** פותחה פונקציה שממירה טווחי הכנסה למספר מייצג (ממוצע טווח) ומסווגת באופן עקבי כ"מעט מעל הממוצע"/"מעט מתחת לממוצע" ביחס לערך ייחוס ממוצע במשק. כך מולאו ערכים חסרים בצורה מוסברת ועקבית גם כאשר לא ניתן היה להסתמך על תשובה מילולית.
- שדות כפולים/מקבילים:** כאשר הופיעו שני שדות המתארים את אותה ישות (כמו "ספק טלוויזיה בינלאומי" באיותים שונים), הוחל איחוד לוגי (coalesce) ובוטלו עמודות העודף.
- לשדות אחרים נשמרה גישה שמרנית: הצגה גלויה של "חסר-" או "לא יודע/ת" כקטגוריות מוצהרות, כדי לשמר שקיפות ולמנוע הטיות עקב מילוי-יתר.

**בקורות ותיעוד.** המחברת כוללת הדפסות ביניים של ספירות מעודכנות לאחר כל שינוי (למשל, כמה ערכים מולאו בעמודת ההכנסה, כמה חסרים נותרו, והתפלגויות חדשות לשדות שאוחדו). בתום העיבוד נשמר קובץ CSV מעודכן ב-UTF-8-SIG-כך שניתן לצרף/לשתף אותו ולהטעינו בכלים נוספים ללא בעיות קידוד.

**מסגרת הניתוח התיאורי (בחלק הבא).** לאחר שהדאטה סודר והושלם, עוברים להצגה תיאורית עקבית לכל משתנה מרכזי:

- התפלגות כללית-** זיהוי קטגוריות מובילות והערכת הפיזור בין חלופות.
- פירוק דמוגרפי-** השוואת אותה התפלגות לפי **מגדר, קבוצת גיל, לאום והגדרה דתית** לצורך איתור הבדלים נראים לעין בין תתי-אוכלוסיות.
- בקרת משמעותיות כללית-** אימות ברמת מאקרו שהפערים הנצפים אינם תוצר של רעש מדגמי בלבד (הצגה תמציתית, ללא העמקה טכנית).  
כאשר מספר הקטגוריות רב, תוצג במקביל גרסת "**Top-N + אחר**" לשיפור הקריאות, לצד הטבלה המלאה לשקיפות.

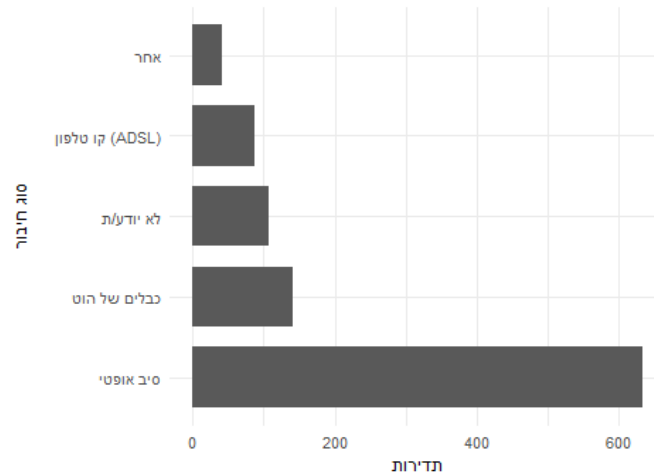
## חלק 1, ניתוח תיאורי:

חלק זה מציג ניתוח תיאורי של שאלות המפתח בסקר. לכל משתנה מוצגת תחילה ההתפלגות הכללית במדגם, ולאחר מכן פירוק דמוגרפי (לפי מגדר, קבוצת גיל, לאום והגדרה דתית). בכל סעיף אנו בוחנים האם בולטת קטגוריה מובילה, עד כמה הפיזור רחב, והאם נראים הבדלים עקביים בין תתי-אוכלוסיות. המטרה: לתרגם את הנתונים לתובנות ישימות על השוק ועל התנהגות המשתמשים.

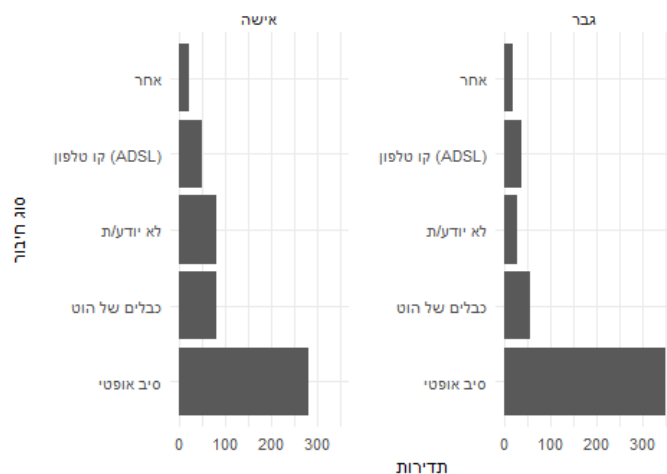
### שלב 1: סוג חיבור האינטרנט

הגצת פריסת סוגי החיבור במדגם כדי להבין מהו ה"סטנדרט" בפועל, ובדיקה האם התמונה משתנה בין קבוצות.

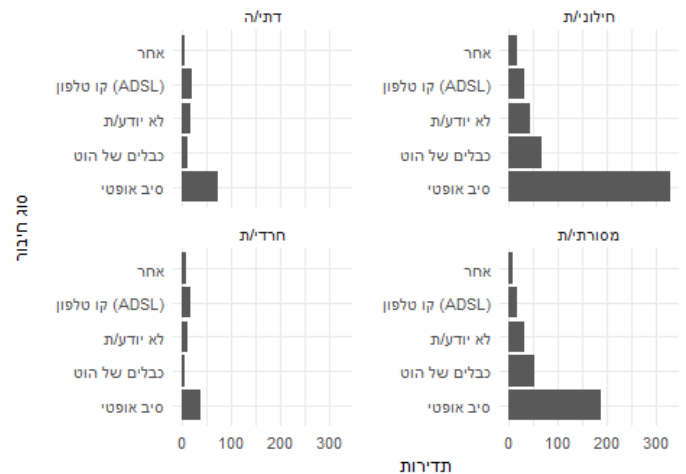
התפלגות סוג חיבור האינטרנט



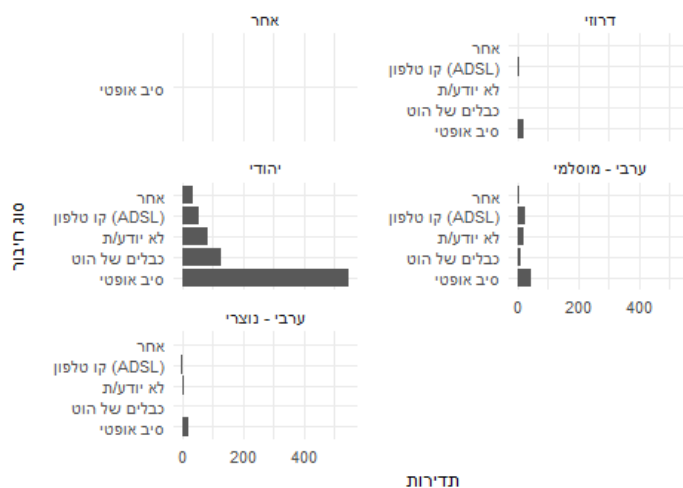
סוג חיבור האינטרנט לפי מגדר



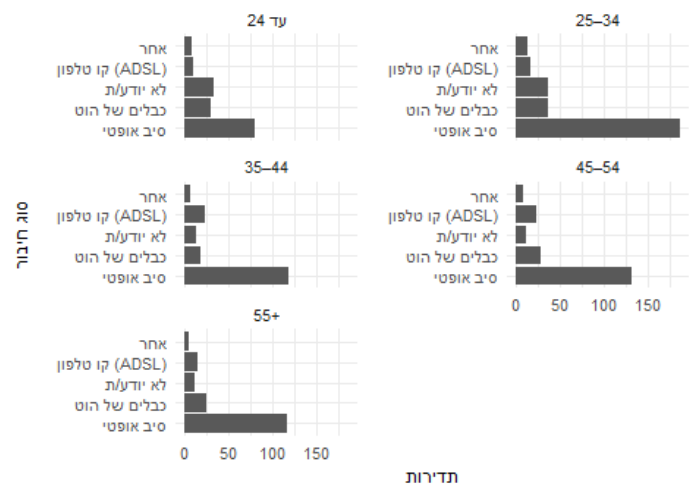
סוג חיבור האינטרנט לפי הגדרה דתית



סוג חיבור האינטרנט לפי לאום



סוג חיבור האינטרנט לפי קבוצת גיל

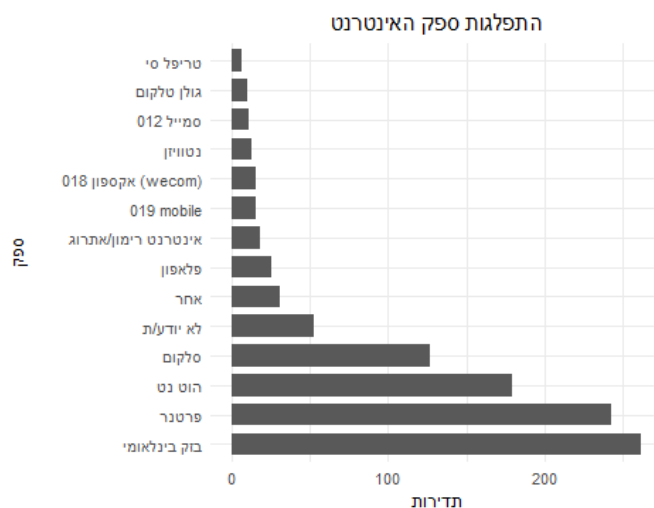


## סיכום ממצאים:

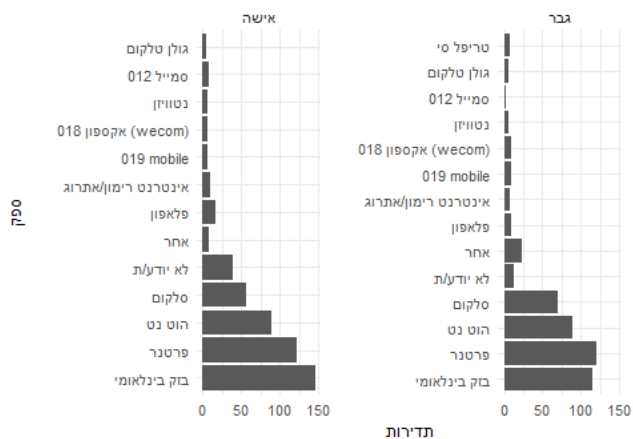
- מגדר: דפוס דומה בין נשים וגברים; אם קיימים פערים, הם נקודתיים בעיקר בקטגוריות המשניות וב"לא יודעת/ת".
- קבוצת גיל: צפויות שונות מתונה-שכיחות גבוהה יותר של טכנולוגיות חדשות בקבוצות הצעירות לעומת הוותיקות.
- לאום: התמונה הכללית נשמרת; הבדלים (אם ישנם) מופיעים במותגים/טכנולוגיות ספציפיים.
- הגדרה דתית: ברוב המקרים הדפוסים דומים; ייתכנו הבדלים מקומיים, בעיקר בקטגוריות עם שכיחות נמוכה.

## שלב 2: ספק האינטרנט

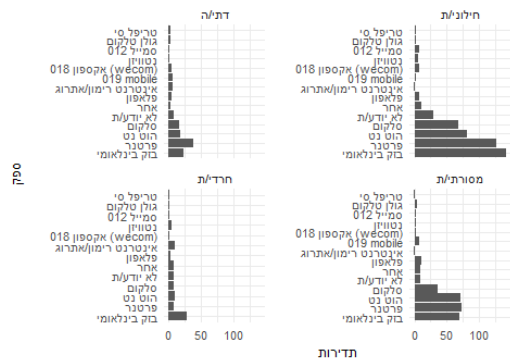
מיפוי מי הספקים הבולטים במדגם וזיהוי ריכוזיות לעומת פיזור בין מותגים; לאחר מכן בחינה של הפערים בין הקבוצות.



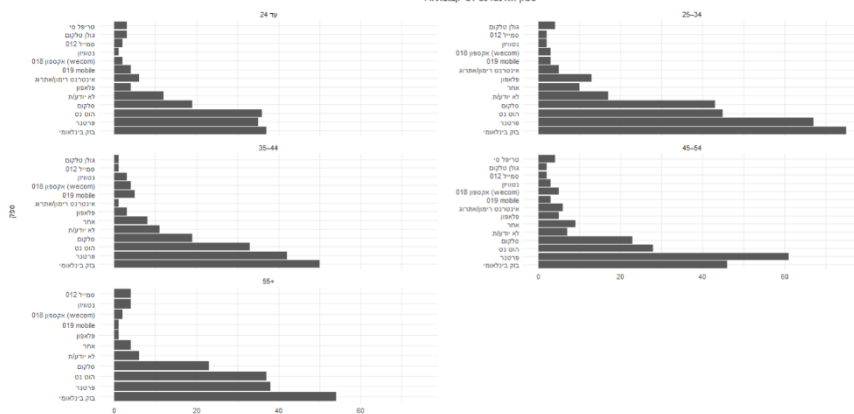
ספק האינטרנט לפי מגדר



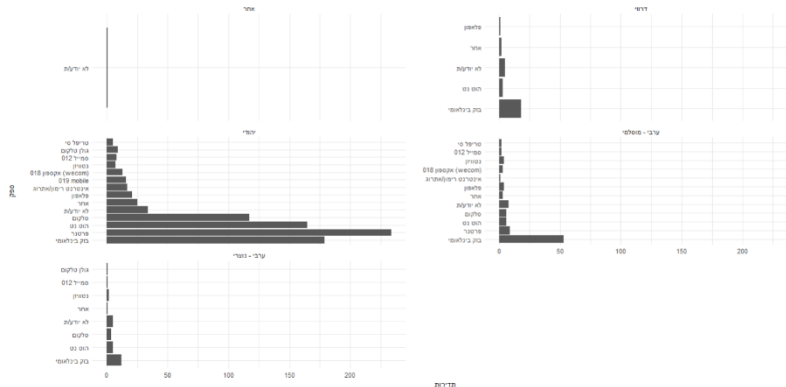
ספק האינטרנט לפי הגדרה דתית



ספק האינטרנט לפי קבוצת גיל



ספק האינטרנט לפי לאום



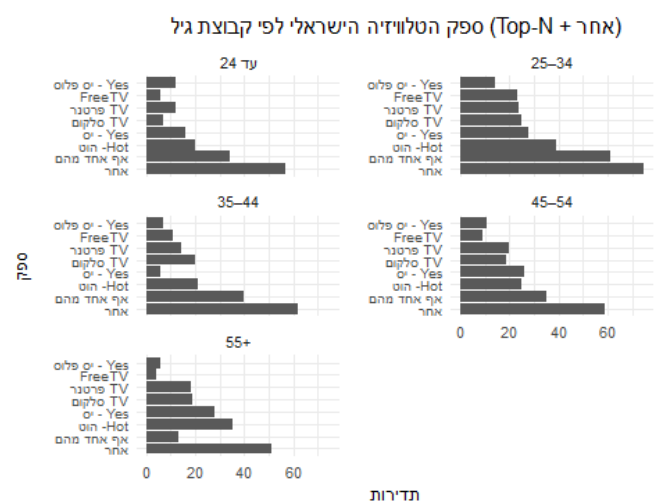
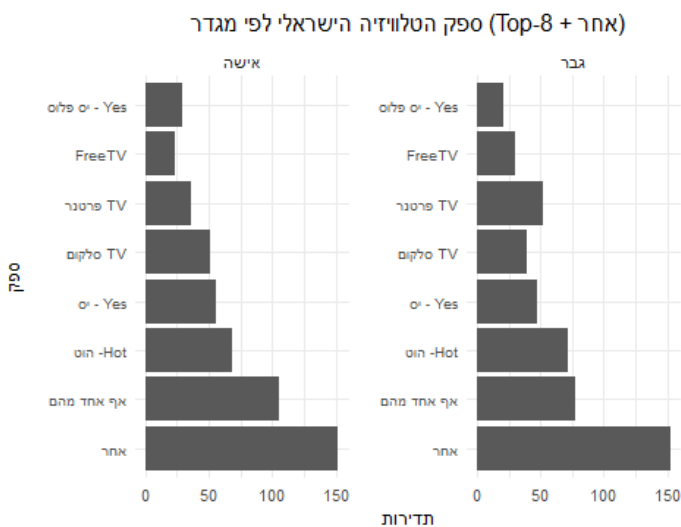
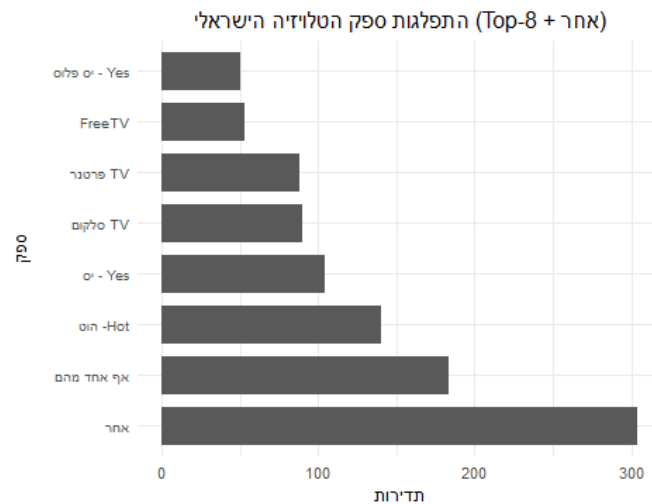
## סיכום ממצאים:

- מגדר: ההרכב דומה בין המגדרים; סטיות קלות עשויות להופיע בספקים משניים.

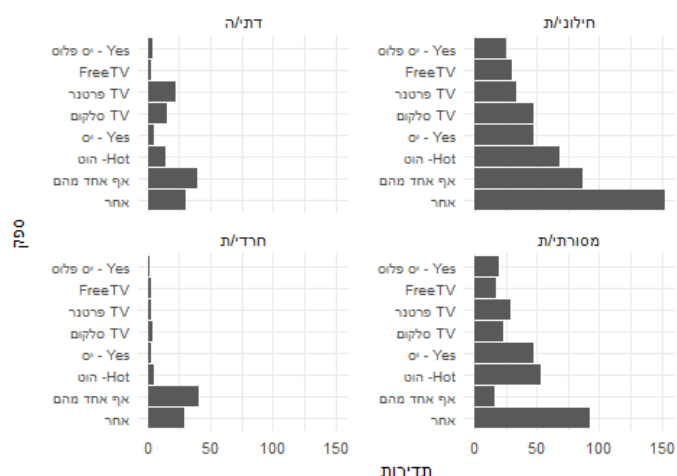
- קבוצת גיל: פערים אפשריים בהעדפת מותגים בין צעירים לבוגרים, אך התמונה הכללית נשמרת.
- לאום: ייתכנו הבדלים נקודתיים בבחירת ספקים מסוימים; השוק בכללותו דומה.
- הגדרה דתית: מגמות דומות; הבדלים (אם קיימים) מצטמצמים לקטגוריות פחות שכיחות.

### שלב 3: ספק טלוויזיה ישראלי

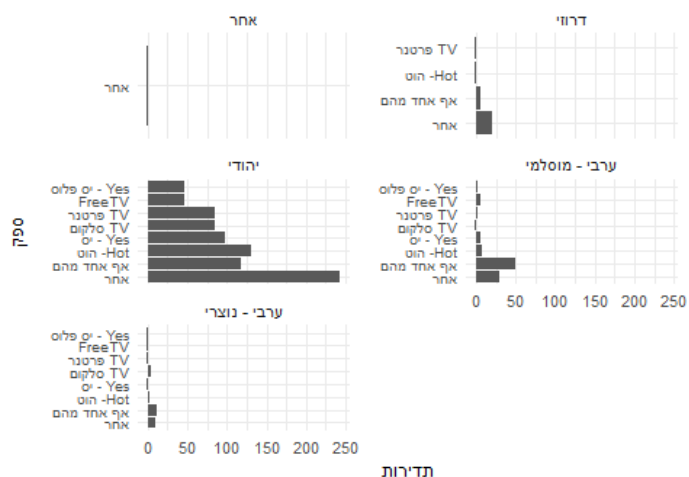
שוק מרובה-מותגים. כדי לשמור על קריאות מוצגת גם גרסת "Top-N + אחר", לצד התמונה המלאה.



ספק הטלוויזיה הישראלי לפי הגדרה דתית (Top-N + אחר)



ספק הטלוויזיה הישראלי לפי לאום (Top-N + אחר)



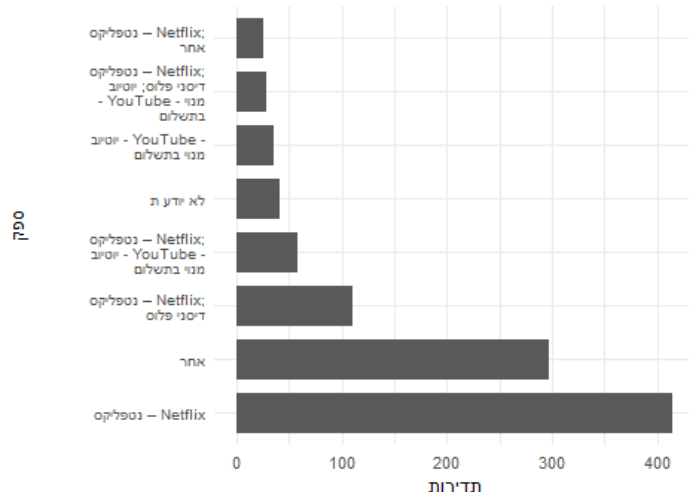
## סיכום ממצאים:

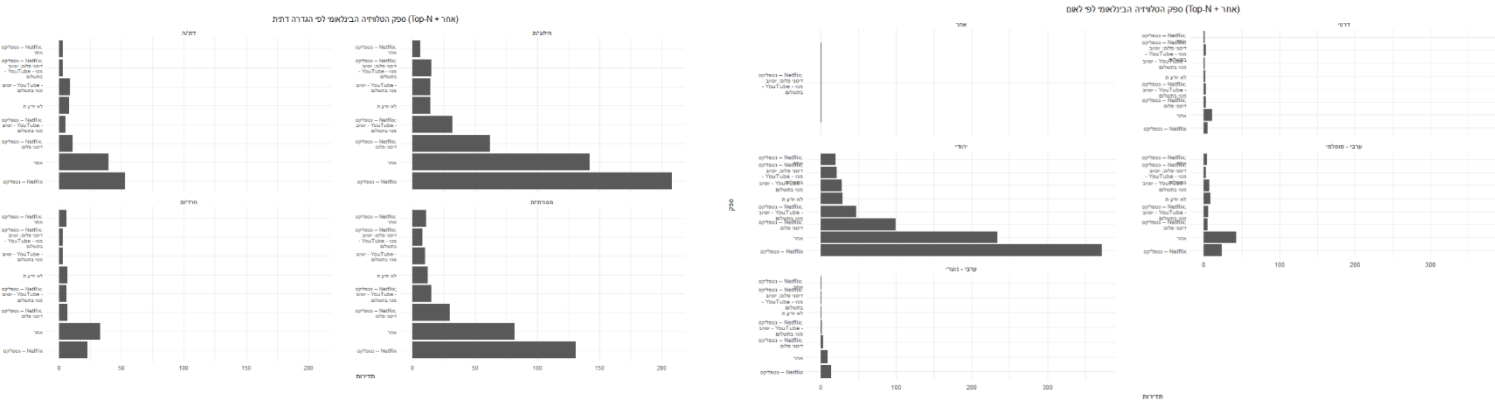
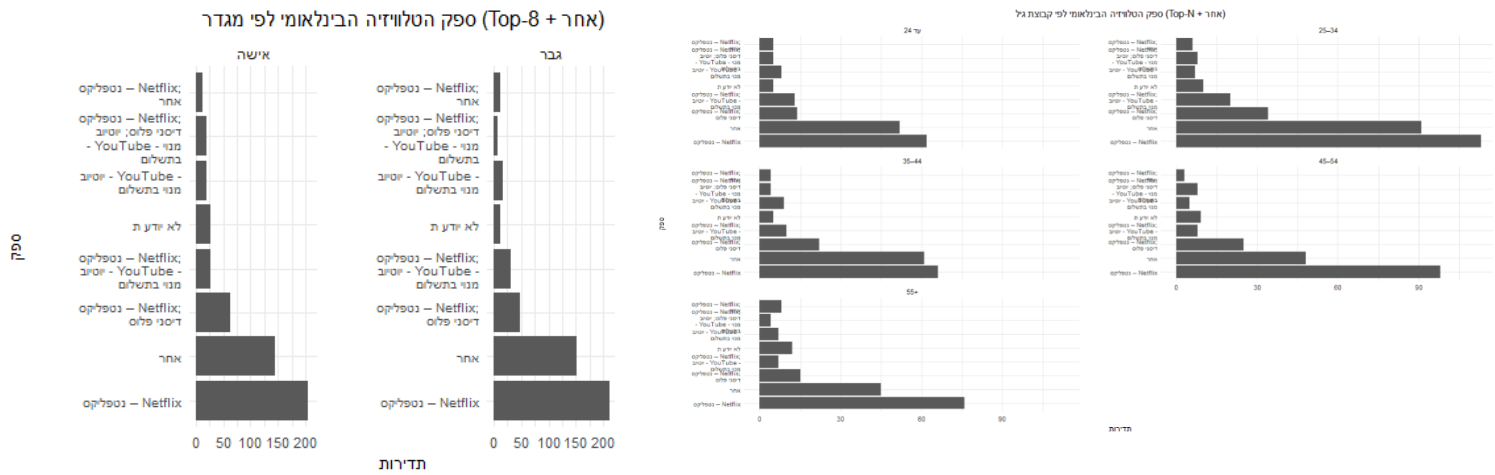
- מגדר: תבנית כללית דומה; שונות קלה אפשרית בהעדפת מותגים ספציפיים.
- קבוצת גיל: ניכרת רגישות גבוהה יותר- צעירים נוטים לשירותים חלופיים/חבילות גמישות יותר לעומת בוגרים.
- לאום: ייתכנו העדפות ממוקדות למותגים מסוימים, אך ללא שינוי מהותי בתמונה הכללית.
- הגדרה דתית: דפוסים דומים בכלל הקבוצות; השונות (אם קיימת) מתמקדת בשוליים ובקטגוריית "אחר".

## שלב 4: ספק טלוויזיה בינלאומי (גם כאן לרוב הרבה קטגוריות)

בדיקה אילו שירותי OTT או סטרימינג בינלאומיים מובילים במדגם, ומיד לאחר מכן את הפילוח הדמוגרפי.

ספק הטלוויזיה הבינלאומי (Top-8 + אחר)



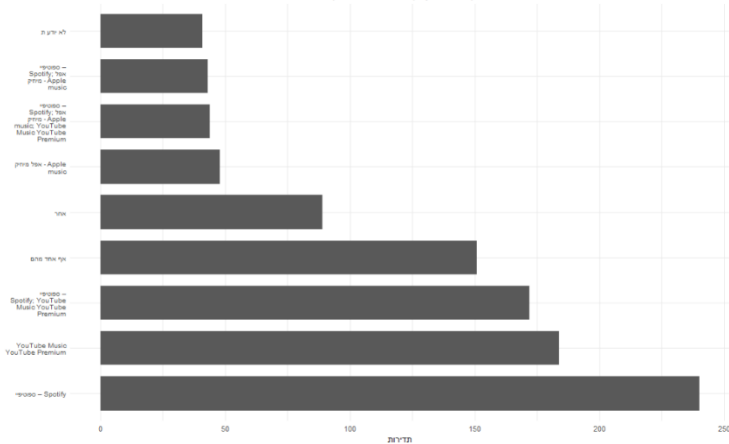
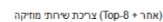
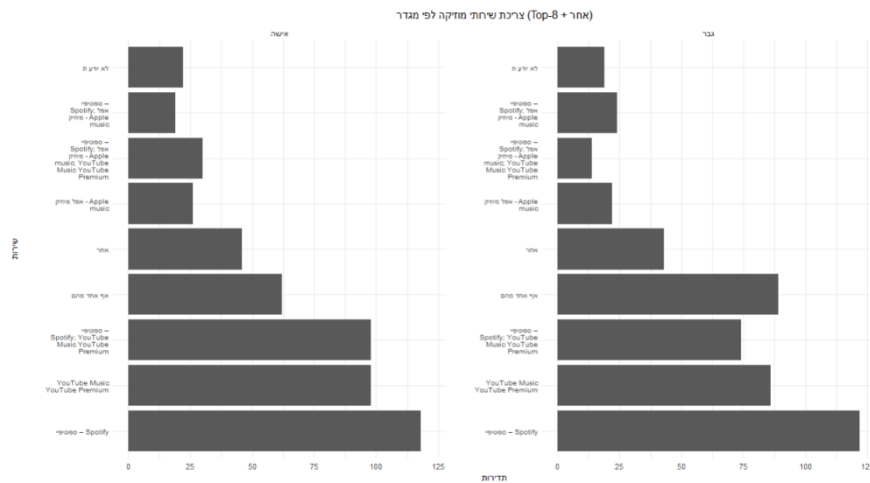


## סיכום ממצאים:

- מגדר: ההובלה של שירותים מרכזיים נשמרת בשני המגדרים; הבדלים קלים בקטגוריות המשנה.
- קבוצת גיל: שונות בולטת יחסית- קבוצות צעירות נוטות לשירותים דיגיטליים מובילים; בקבוצות בוגרות נראית היקשיות מעט גבוהה יותר ב"אחר"/שירותי משנה.
- לאום: הבדלים נקודתיים בהעדפות, אך לא שינוי יסודי במבנה ההובלה.
- הגדרה דתית: דפוס דומה בכלל הקבוצות, עם שונות קלה בשירותים המשניים.

## שלב 5: צריכת שירותי מוזיקה

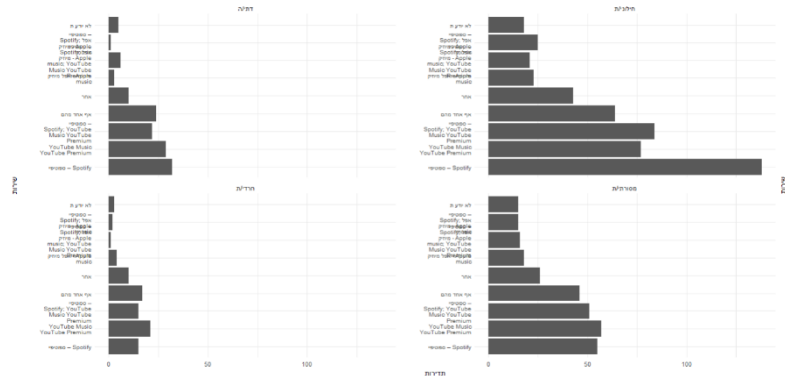
מיפוי השימוש בשירותי מוזיקה דיגיטליים (ספרימינג/פלטפורמות אחרות) וזיהוי דפוסים בין תתי-אוכלוסיות.



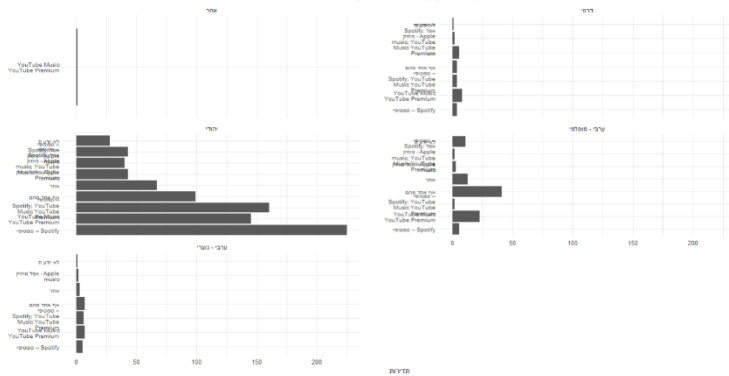
(אחר + Top-N) צרכת שיחת מוזיקה לפי קבוצת גיל



(אחר + Top-N) צרכת שיחת מחזקה לפי הגדרה דתית



(אחר + Top-N) צרכת שיחת מוזיקה לפי לאם



## סיכום ממצאים:

- מגדר: התפלגות דומה; אם יש הבדלים, הם בעיקר בשירותי נישואים.
- קבוצת גיל: שונות משמעותית יותר- צעירים מרוכזים בשירותי סטרימינג מובילים; בוגרים מראים פיזור רחב יותר ואף שיעור "לא משתמש/ת" גבוה מעט יותר.
- לאום: ייתכנו הבדלים נקודתיים בהעדפת פלטפורמות, אך התמונה הכוללת יציבה.
- הגדרה דתית: מגמות דומות; הבדלים (אם קיימים) מתמקדים בשירותים משניים ובקטגוריית "אחר".

**לינק לדשבורד אינטרקטיבי:**

[https://public.tableau.com/views/\\_17559015604400/sheet3?:language=en-US&publish=yes&:sid=&:redirect=auth&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/_17559015604400/sheet3?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link)