

# עבודה 2

הגורם האנושי באיסוף נתונים

רון דגני 318170917 , יקטרינה סופרפין 321170383

## 1. הצגת המוצר והשוק

המוצר הנבחר הינו טאבלט.

שימושיו האפשריים הינם: למידה, משחק, ציור, עבודה ועוד.

הפונקציות שהינו מספק הינן מגוונות מאוד וכוללת כמעט את כל הפונקציות הקיימות במחשב – החל מהקלדת וכתיבת מסמכים, צפיה בסרטונים, משחק וכדומה.

ייחודו המרכזי מהמחשב הינו העובדה כי המכשיר קומפקטי וקל, ניתן לכתוב עליו באמצעות עט ולאפשר אפשרויות גרפיות רחבות.

הטאבלט הינו שילוב של מחשב, פלאפון ומחברת.

## 2. מפרט המוצר:

- שם
- מחיר
- זיכרון RAM
- גודל מסך
- נפח אחסון פנימי
- יצרן
- משקל
- תאריך כניסה לאתר
- מערכת הפעלה

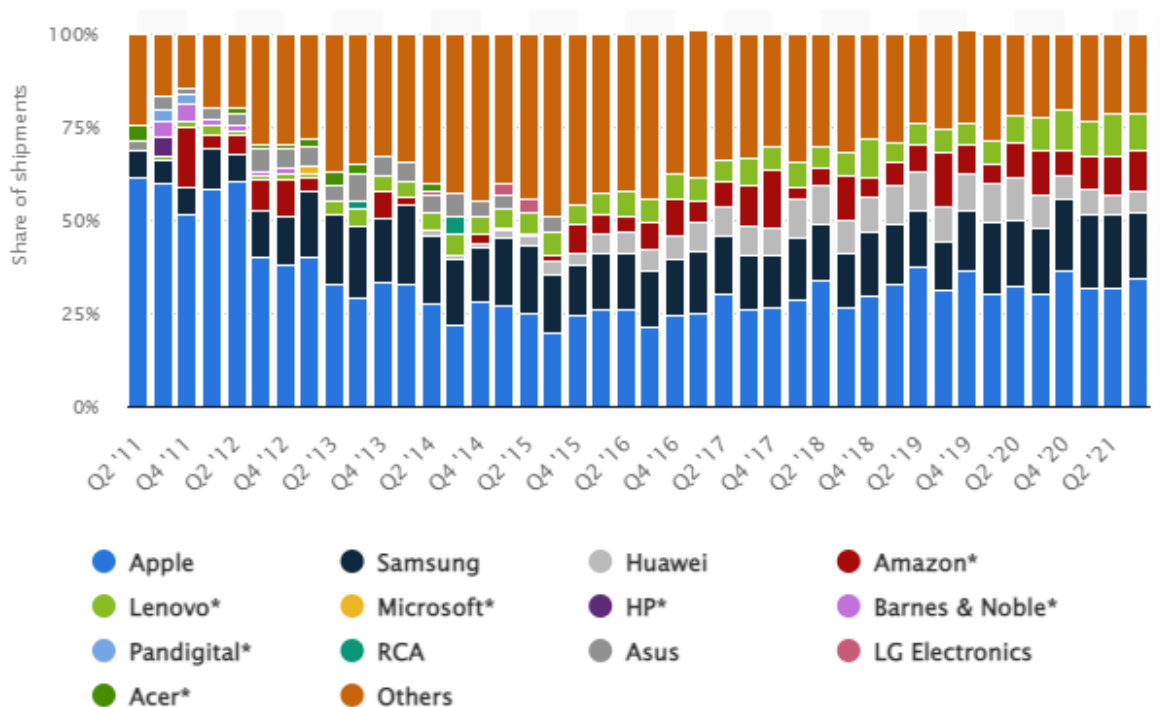
### 3. נתוני השוק:

- כמות משתמשים בעולם לפי שנת 2017 הינה 1.14 מיליארד.
- יצרנים (בסוגריים מצוינות כמות הגרסאות ליצרן):

Apple(199), Samsung(166), Lenovo(86), Xiaomi(5),Microsoft(101),  
Amazon(39),Victorio(6), Silver Line(15), Asus(24), Huawei(22), Packard Bell(10),  
Sansui(2),BDK(4), Crown(2), Sony(1), Panasonic(16), Toshiba(1),  
Viewsonic(1),Getac(16), Acer(5), Blaupunkt(3), Dell(3), Hyundai(2),  
Google(1),ECO(4), HP(1), ZTE(1)

- ניתוח סטטיסטי:

בגרף להלן (קישור מצורף בביבליוגרפיה) מוצגים נתונים סטטיסטיים על שוק הטאבלטים בעולם.



ניתן לראות את פילוח השוק בין החברות נכון לרבעון האחרון בשנת 2021. לפי פילוח שוק זה נוכל להסיק שחברת Apple מובילה בשוק הנ"ל אך גם לחברות כמו Samsung ו-Lenovo יש נציגות מספקת וגם לחברות נוספות יש ייצוג קטן יותר אך עדיין נוכח בשוק.

בנוסף, ניתן לראות את פילוח השוק לאורך זמן, החל מהרבעון השני של שנת 2011. כך לדוגמה נוכל להסיק שהנתח היחסי של חברת Apple קטן לאורך השנים, לעומת חברות

אחרות, למרות שבהתחלה היה יותר מ-50% של השוק. לעומת זאת חברות אחרות הגדילו את נתח השוק שלהן והתפתחו עם הזמן וגרמו לשינויים בשוק.

באופן דומה נוכל לסקור גם את השינויים של כל חברה בפני עצמה. כך לדוגמה ניתן לראות שחברת Samsung גדלה והרחיבה את הנתח שלה בתחום הטאבלטים.

הערה - נשים לב שאמנם הגרף המוצג לעיל מעט עמוס, אך הוא מאוד אינפורמטיבי וברור יחסית לגרפים אחרים ברשת. בנוסף הוא מסכם בתוכו הרבה אינפורמציה שלא מצאנו בגרפים מפוצלים ופשוטים יותר או בנתונים גולמיים.

#### • ניתוח דמוגרפי של קהל היעד:

- אחוז הנשים שמשתמשות בטאבלטים הינו 50.8 בעוד שאחוז הגברים הינו 49.2.
- קהל היעד המרכזי שמשתמש בטאבלטים בינו בגילאים 25-44 (מהווה נתח של 45.8 מהמשתמשים)
- כמעט 60 אחוז מהמשתמשים מגיעים ממשקי בית עם הכנסה של \$75,000 או יותר.

המקורות מהם לקחנו את המידע לעיל מצוינים בחלק הביבליוגרפיה בנספח א'.

#### 4. איסוף הנתונים ועיבוד ראשוני:

אישור שימוש באתר לפי התקנון של האתר עצמו:

תכנים שנמסרו לפרסום על ידי המשתמש

1. החברה מאפשרת למשתמש לכתוב סקירה על מוצר וכן לעיין בסקירות שנכתבו על ידי משתמשים אחרים באתר (להלן: "סקירות").

קישור לתנאי השימוש של האתר: <https://www.wisebuy.co.il/termsfuse.aspx>

קובץ הקוד שבעזרתו ביצענו את שלב ה-web-scraping מצורף לעבודה.

## 5. עיבוד הנתונים:

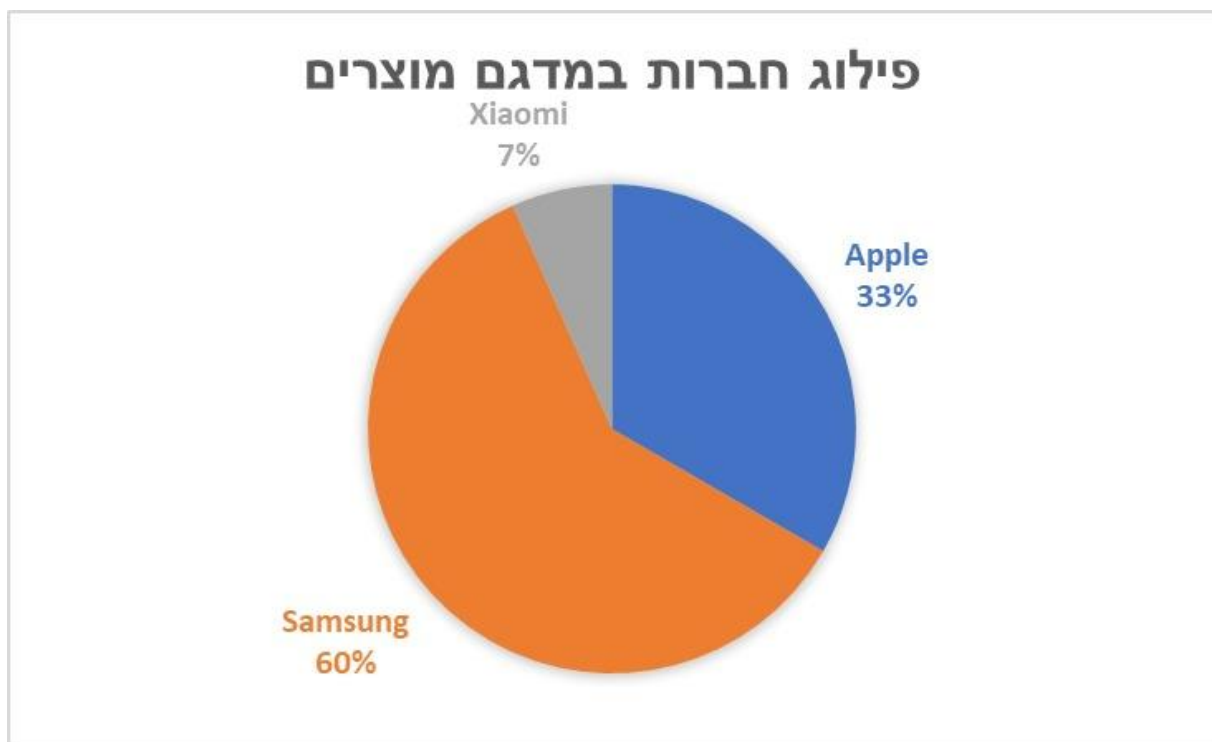
באמצעות ספריית Selenium ביצענו web-scraping לאתר: <https://www.wisebuy.co.il>. המטרה שלנו הייתה לאסוף מידע על 15 פריטים. לשם כך, אספנו מידע על טאבלטים שהופיעו באתר.

מתוך רשימה זו, סיננו את הפריטים אשר לא היה להם דירוג בכלל, הורדנו קטגוריות שלא היו רלוונטיות לנו והורדנו קטגוריות בהם היה חסר מידע בחלק מהפריטים. לדוגמה, היו פרמטרים שהורדנו מכיוון שלשיקולנו אדם ממוצע פחות מתעניין בהם או מבין את משמעותם. בנוסף כדי שתנאי הניסוי יהיו אחידים לכל הטאבלטים, הורדנו פרמטרים שלא מופיעים לכל הטאבלטים. מהפריטים שעברו את הסינון לקחנו 15 טאבלטים. 9 פרמטרים נבחרים, אשר לכלל הטאבלטים היו נתונים עליהם.

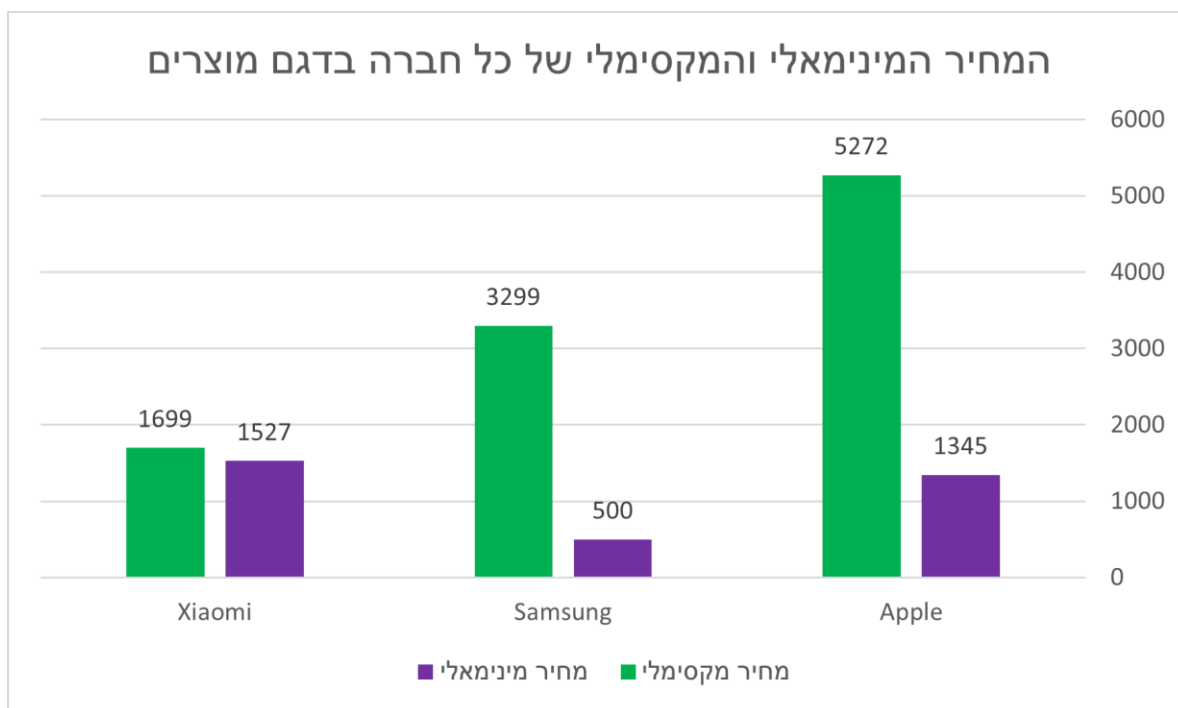
## 6. אקספלורציה:

לאחר קבלת מדגם הפריטים המסוננים מצאנו מספר קשרים מעניינים שנציג בגרפים הבאים:

א. שמנו לב שמדגם המוצרים שלנו שונה מהמדגם בשוק היום. על מנת להבין את המדגם שהתקבל, בחנו את פילוג החברות. לפי הניתוח הסטטיסטי בסעיף 3 החברה הדומיננטית בשוק כיום היא Apple ולעומת זאת במדגם שלנו קיבלנו פילוג חברות שונה- כך שדווקא חברת Samsung מובילה.

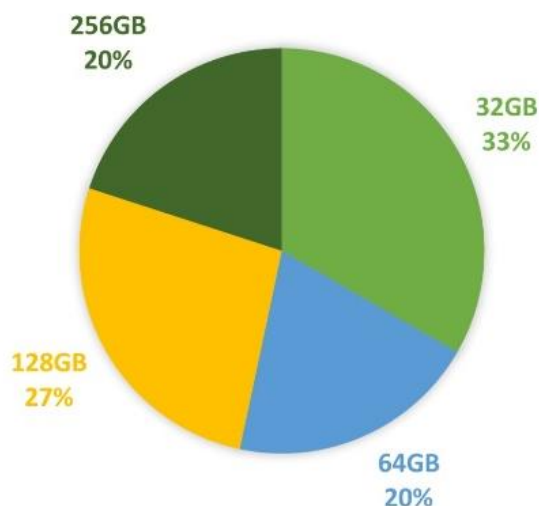


ב. באופן דומה, חשוב להבחין בטווחי המחירים של כל חברה שיכולים להוות היבט מעניין בנושא ובנוסף ישמשו לניתוח התוצאות בהמשך.  
לדוגמה ניתן לראות כי מחירי חברת Apple גבוהים מאוד משאר המוצרים.



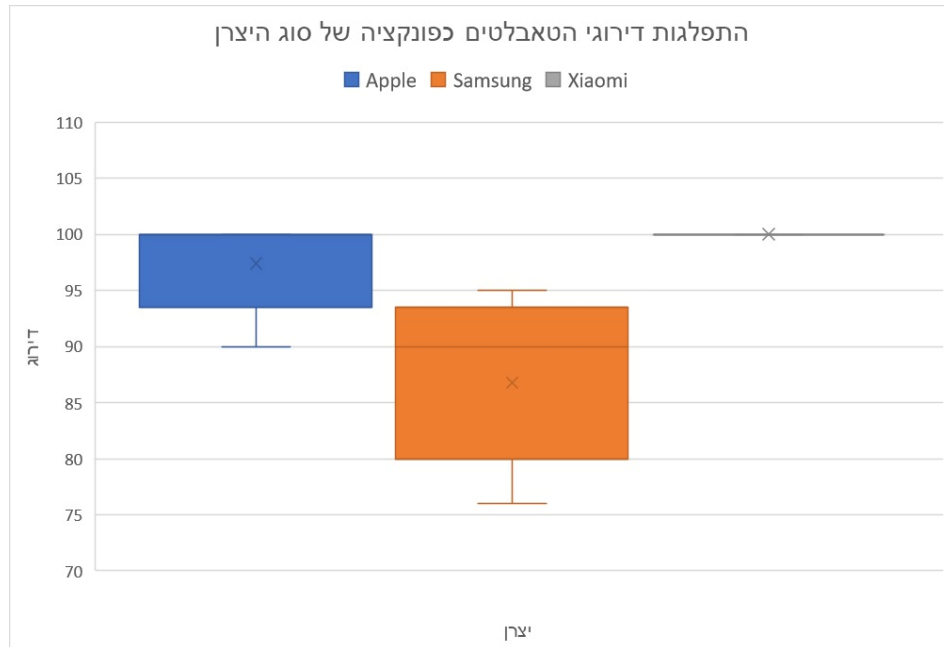
ג. פרט נוסף ששמנו לב אליו הוא פילוח המדגם שלנו לפי נפחי אחסון.  
ניתן להבחין כי ההתפלגות היא כמעט זהה בין נפחי האחסון השונים, זאת בניגוד להשערה שלנו כי רוב הטאבלטים הם בעלי נפח אחסון קטן מסתמכים על שירותי ענן. ניתן להסביר פער זה בכך שהמדגם שלנו קטן ולא מייצג.

### פילוח נפח אחסון פנימי במדגם

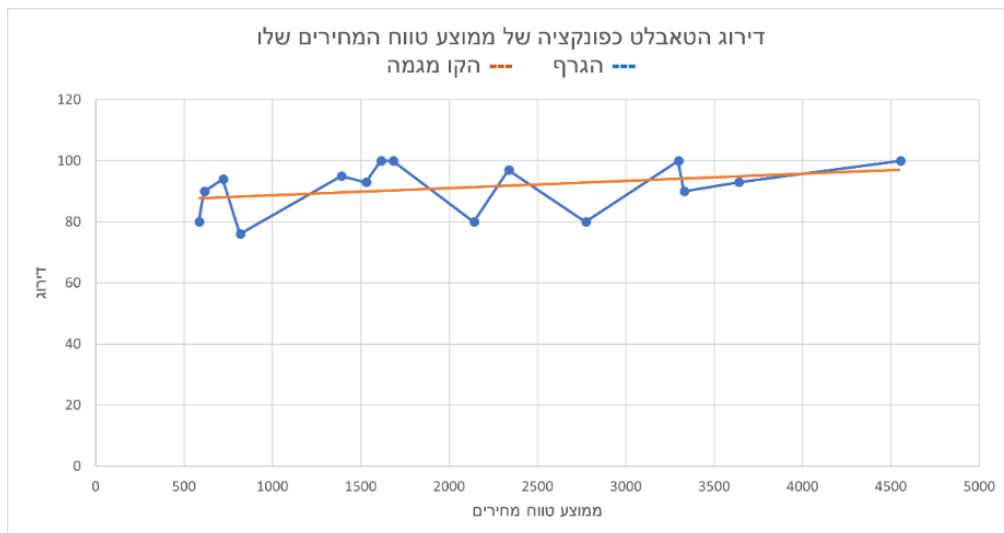


## 7. נציג 4 מתאמים שונים בין המפרט לבין הדירוגים:

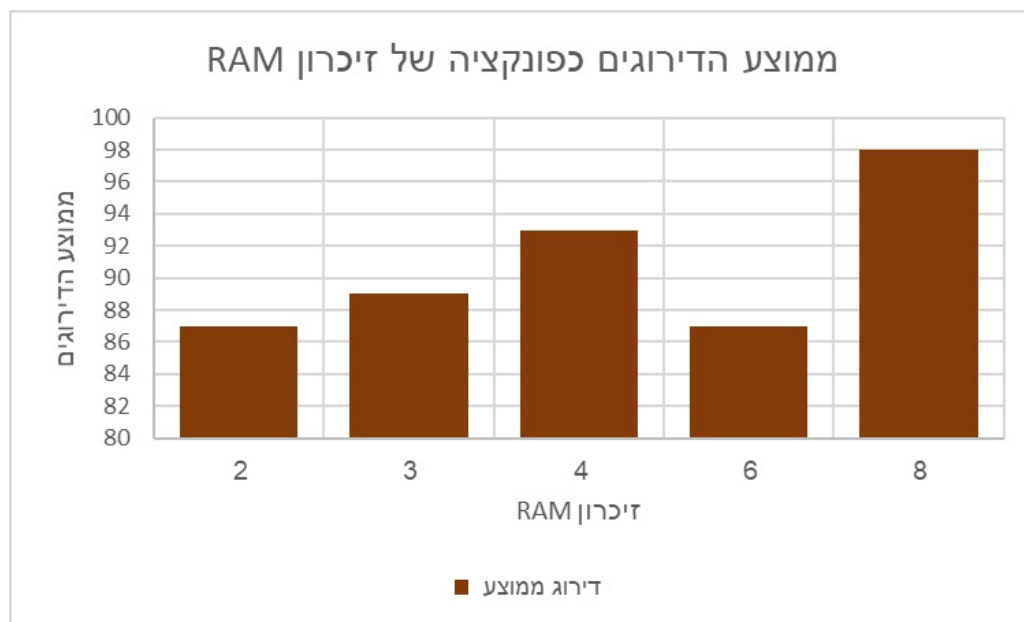
א. ניתן לראות מתאם בין היצרן לדירוג המוצר – נראה שדירוגי Apple גבוהים בממוצע בכ-10 נקודות מדירוגי Samsung, וכי טווחי המינימום והמקסימום שלהם גם הם מתואמים בדיוק באותו אופן. לכן, ישנה קורלציה בין זהות היצרן לדירוג המוצר באתר. Xiaomi מיוצג ע"י מכשיר אחד ולכן מידע אודותיו אינו רלוונטי ואינפורמטיבי בגרף זה.



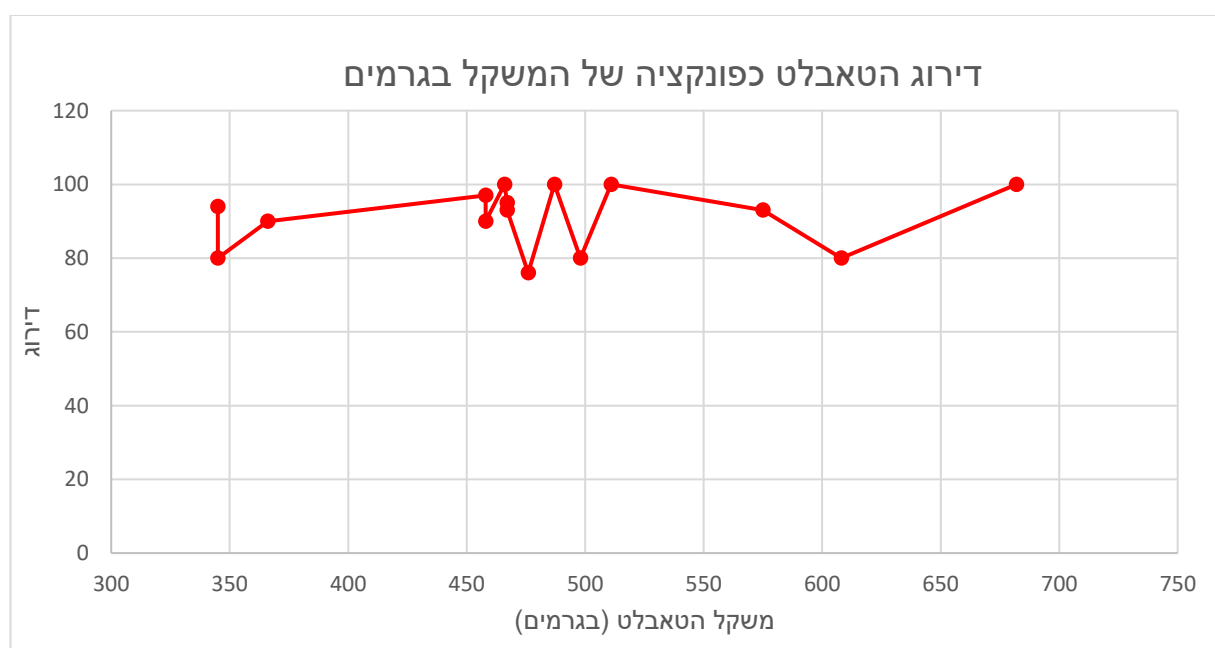
ב. היבט מעניין נוסף הוא הקשר שבין הדירוג למחיר במדגם הספציפי שלנו. לכל מוצר ישנו טווח מחירים (משתנה בין חנויות) ולא מחיר בודד, לכן נציג קשר זה כגרף של הדירוג של הטאבלט כפונקציה של ממוצע טווח המחירים שלו. נראה כי ככל שהמחיר גבוה יותר כך גם הדירוג עולה. המגמה אינה חדה אך יש מתאם כלשהו.  $R \approx 0.4$  לפי מבחן פירסון. על מנת להדגיש את הנקודה הוספנו קו מגמה שמדגיש את הקורלציה (הערה – לא מצאנו מתאם מעניין יותר).



ג. מתאם נוסף בין המפרט לבין הדירוג הוא גרף של דירוג הטאבלט כפונקציה של זיכרון ה-RAM (ב-GB). כפי שניתן לראות, לקריטריון זה של המפרט אכן יש השפעה על הדירוג של המוצר.



ד. כמו כן מצאנו גורם שאין לו קורלציה עם הדירוג של הטאבלט. כפי שניתן לראות בגרף של דירוג הטאבלט כפונקציה משקלו (בגרמים), אין קשר בין שני המשתנים.  $R=0.16$  לפי מבחן פירסון. כלומר, מכאן נוכל להסיק שקריטריון זה של המפרט לא משפיע על הדירוגים של הלקוחות (בשונה ממה שציפינו).



## 8. הסיווג:

הסיווג שבחרנו הינו האם המוצר הוא מוצר מותרות או לא? את הסיווג הזה לא ניתן להסיק מהמפרט הטכני והוא סובייקטיבי לכל אדם על פי צרכיו וראות עיניו.

חשוב לציין שמדגם הנבדקים שלנו הוא בפרט סטודנטים מהטכניון, ולכן גם אם ממבט ראשון ניתן לשער שהמוצר הינו מוצר מותרות, אנו כסטודנטים יכולים להעיד שמוצר זה לא נתפש כמוצר מותרות בקרב אוכלוסיית יעד זו.

השאלה האם המוצר הוא מוצר מותרות היא מאוד סובייקטיבית ומתבססת על פרמטרים ושיקולים אישיים של כל בן אדם, כגון: מספר השעות שישתמש במוצר ולאילו מטרות, חיוניות, נוחות שנובעת ממערכת טעמים והעדפות, תועלת אישית, חיסכון בזמן, שיקולי אקולוגיה וכדומה. כל זאת תוך שקלול המפרט הטכני.

המטרה בשאלון היא שכל נבדק יכריע עבור כל מוצר (15 מוצרים בסה"כ) האם הוא מחשיב אותו כמוצר מותרות או לא וזאת לפי המפרט הטכני שנציג לגבי כל מוצר.

מהשאלון אנו צופים לקבל את המידע הבא:

- מידע לגבי האופן בו האדם החליט איך לסווג את המוצר (על מנת להבין את ביצוע הבחירה באופן טוב יותר) לפי איך שהנבדק העיד על עצמו .
- ניתוח הגורמים שמשפיעים על הסיווג ורמת ההשפעה שלהם ממה שנובע מהתשובות לשאלות הסיווג.
- השוואה בין הגורמים שהנבדקים טענו שהשפיעו עליהם לבין הגורמים שאכן השפיעו עליהם לפי התוצאות.
- הבנת הקריטריונים והספים שגורמים לאדם לסווג טאבלט כמוצר מותרות.

9. העברת השאלון: קובץ השאלון מצורף לעבודה. על השאלון ענו 20 נבדקים מאוכלוסיית היעד שהוגדרה – סטודנטים בטכניון.

הערה- לשאלון לא צורפו תמונות של הטאבלטים, על מנת למנוע השפעה מסוג זה - הבדל בעיצוב, בתאורה, בצבעים באופן פוטנציאלי יכולים להשפיע על הדעה ואין אנו חושבים שתמונה אחת בפרט מייצגת מוצר כלשהו ויכולה להכנס כ'מפרט המוצר'.

## 10. תוצאות השאלון:

בסקירת תוצאות השאלון ניתחנו בין היתר דיאגרמות עוגה לכל שאלת סיווג. לפי דיאגרמות אלו יכולנו להתרשם שישנה הסכמה גבוהה מאוד לגבי הקטגוריה של רוב הטאבלטים, למעט שלושה מהם- שבהם הסיווגים בין הקטגוריות התחלקו כמעט שווה.



מבדיקת מידת החשיבות של הפרמטרים הקיימים לסיווג נראה כי הפרמטרים של המפרט הטכני שהצגנו לנבדקים עזר ל-80% מהנבדקים לקבל החלטה. מידת החשיבות של כל פרמטר שהצגנו בעיני הנבדקים: על מנת לבדוק את מידת החשיבות של כל פרמטר, התבקשו הנבדקים בשאלון לדרג עד כמה קריטריון זה או אחר טרם להכרעתם. בטבלה הבאה נציג לגבי כל קריטריון של המפרט הטכני את ממוצע הדירוגים:

קריטריון	נפח אחסון פנימי	מחיר	זיכרון RAM	גודל מסך	משקל (גרם)	תאריך כניסה לאתר	מערכת הפעלה	יצרן
דירוג ממוצע	5.45	8.65	5.6	6.25	3.8	2.9	4.8	7.15

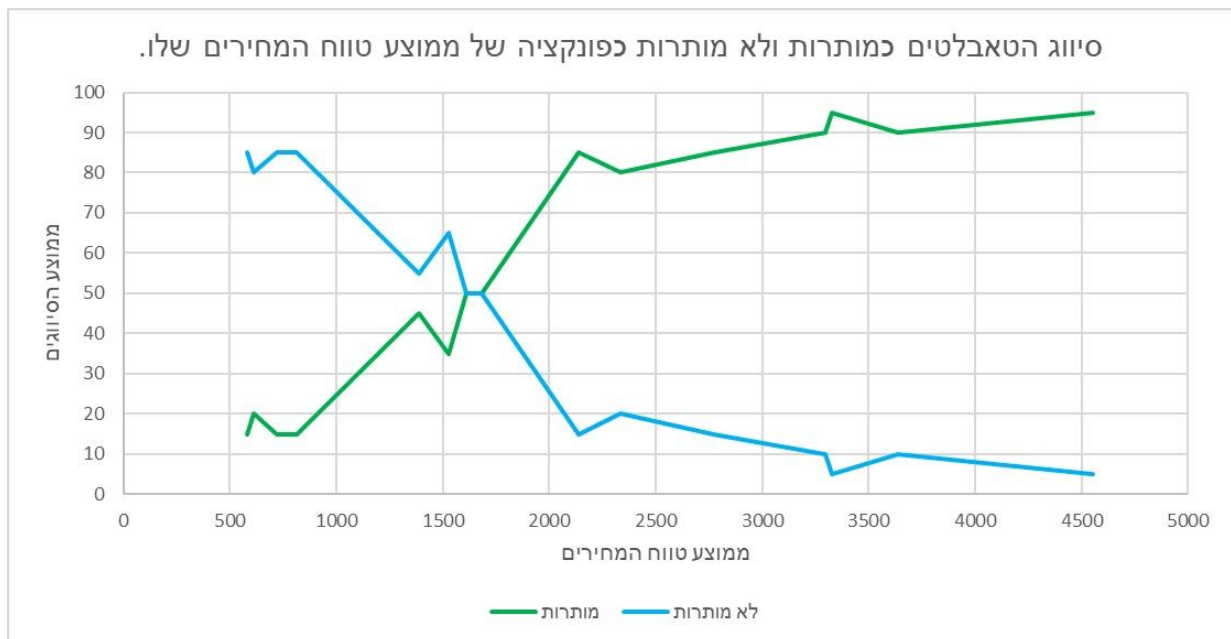
לפי טבלה זו נבחין כי החשיבות הגדולה ביותר בעיני הנבדקים ניתנה למחיר, לאחר מכן ליצרן והשלישי בחשיבותו הוא גודל המסך. כמו כן ניתן לראות שהמשקל ותאריך הכניסה לאתר (שמדגיש את חדישות המוצר) היו הכי פחות חשובים לנבדקים. נשים לב שהנבדקים נתנו דירוג ממוצע (קרוב ל-5) לפרמטרים כגון: מערכת הפעלה, נפח אחסון פנימי וזיכרון RAM, למרות שציפינו שפרמטרים חשובים כאלו יהיו יותר חשובים בעיני הנבדקים בהכרעה. פרמטרים נוספים שלדברי הנבדקים היו עוזרים להם לסווג- זמן חיים של הבטרייה, פיצ'רים (לא פירוט), רזולוציית מסך, איכות הכתיבה על המסך וכו'.

## ניתוח תוצאות שאלות הסיווג:

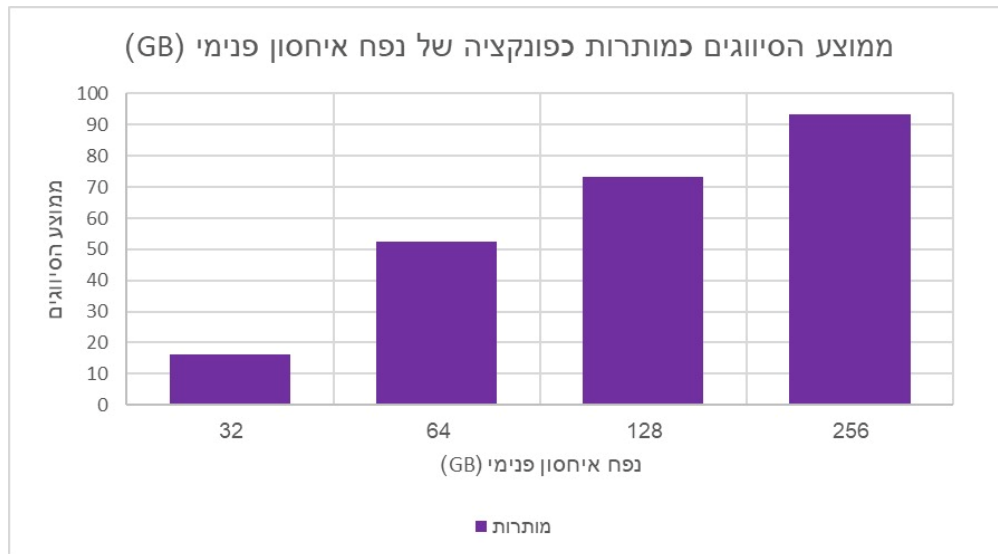
א. נבחן את הקורלציה בין המחיר לסיווג:

לפי מבחן פירסון קיבלנו כי ערכו של  $R=0.93$ . נזכר כי לפי מבחן פירסון, ערך הקרוב ל 1 מעיד על קורלציה חיובית בין המשתנים, כלומר ככל שאחר המשתנים עולה כך גם השני. במקרה זה בדקנו את המשתנה מחיר והמשתנה מותרות.

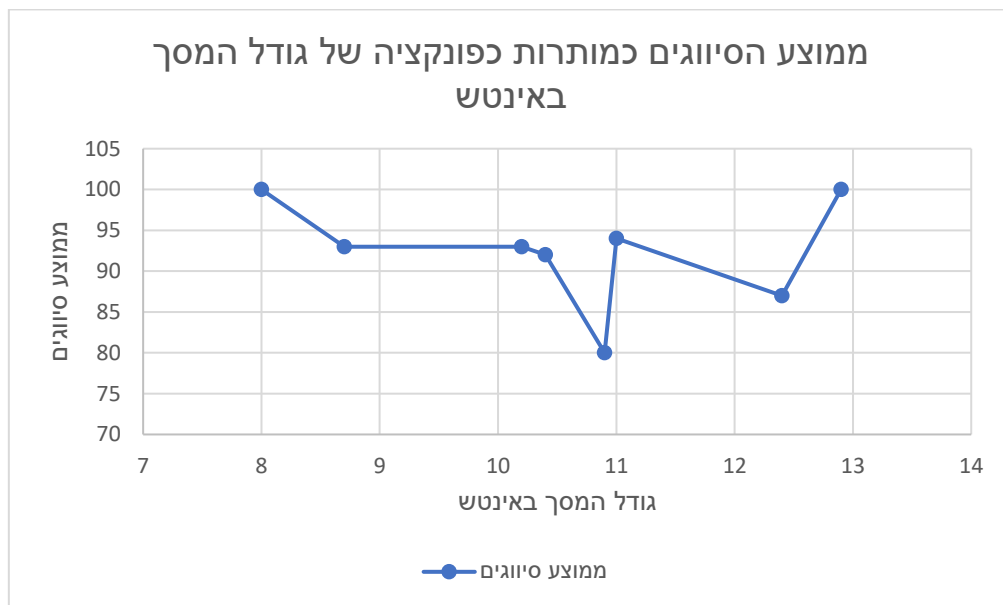
כלומר, כצפוי, קיבלנו קורלציה חזקה בין פרמטר המחיר של המוצר לבין ממוצע הסיווגים שלו כמותרות וכלא מותרות על פני כל מדגם הנבדקים. בנוסף שמנו לב כי ישנו מחיר סף מסוים, המגדיר את סיווג הטאבלט. כלומר, כל טאבלט שמחירו מתחתיו מוגדר כלא מותרות, וכל טאבלט שמחירו מעליו מוגדר כמותרות. מחיר זה הוא 1613-1682. ניתן לראות זאת בגרף זה:



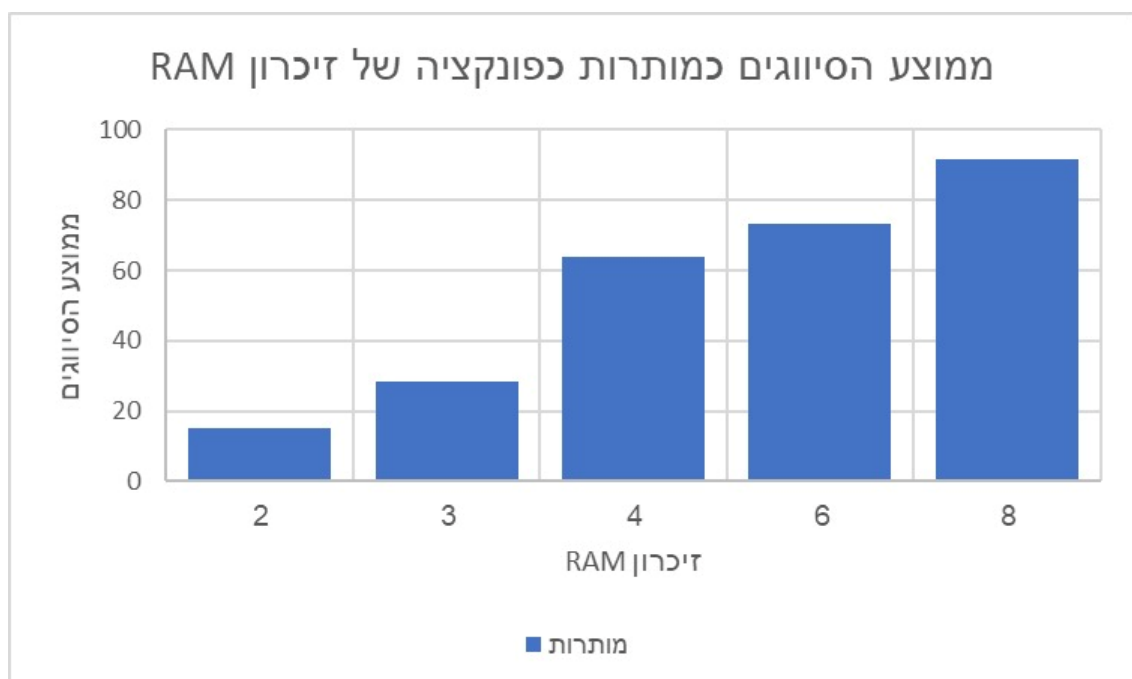
ב. לא שיערנו למצוא מתאם בין נפח האחסון הפנימי של המכשיר לבין הסיווג שלו כמוצר מותרות, מכיוון שמהכרתנו עם קהל היעד, הוא משתמש בעיקר בשירותי ענן למטרת שמירת מידע. למרות זאת, מצאנו כי ישנה קורלציה משמעותית בין המשתנים, ככל שנפח האחסון הפנימי גדל, כך סיווגו הממוצע כמוצר מותרות עולה (נשים לב שמכיוון שבסעיף 6 ג מצאנו שמדגם הטאבלטים מתפלג זהה בין נפחי האחסון הפנימי, נוכל להסיק את המסקנות מבלי חשש להטיה על בסיס המדגם).



ג. מתאם שלא ציפינו למצוא ואכן לא מצאנו הינו הקשר בין גודל המסך הטאבלט לסיווגו. לפי מבחן פירסון נקבל כי  $R=-0.21$  דבר המצביע על חוסר קורלציה. לא ציפינו למצוא מתאם זה מכיוון שההבדלים בגודל המסך בין הטאבלטים אינם גדולים ולכן אין אנו חושבים שזה יהווה שיקול מרכזי לעומת תכונות אחרות של הטאבלט. כפי שניתן לראות בגרף זה אכן אין שום מתאם בין שתי התכונות.



ד. ציפינו שתהיה קורלציה בין הסיווג לזיכרון ה-RAM מכיוון ששערנו כי קהל היעד שלנו מבין בתחום ומבין את חשיבות הפרמטר. אכן לפי תוצאות השאלון התקבלה תוצאה זו.



## 11. השוואה בין תוצאות השאלון לדירוגים:

### דומה:

- במהלך ההשוואה גילינו דמיון רב בין השניים, בכל הפרמטרים שנבדקו, להלן כמה קווי דמיון בולטים:
- הן בשאלון והן בדירוגי האתר מצאנו כי המחיר השפיע על הסיווג\דירוג (סעיף 7.ב, 10.א). בשאלון הקשר נמצא בצורה חדה יותר, כלומר, ככל שהמחיר עלה יותר, כמות המסווגים את המוצר כמוצר מותרות עלה באופן משמעותי.
  - הן בשאלון והן בדירוגי האתר מצאנו כי יצרן המוצר משפיע מאוד על הציון – ראינו זאת בתוצאות השאלון בציון הפרמטר (אשר נמצא השני בחשיבותו לסיווג וקיבל ציון - 7.15 (סעיף 10)) ובהתאמה נמצא בעל קורלציה לדירוגים באתר (סעיף 7.א).
  - כפי שהוצג בסעיפים הקודמים, הן בשאלון והן בדירוגי האתר (10 ד. ו 7.ג) ראינו כי ישנה קורלציה בין זיכרון RAM לסיווג\דירוג המוצר.

### שונה:

נבחין כי הסיווגים שסווגו בשאלון ודירוגי האתר עונים על שאלות שונות: בשאלון התבקשו הנבדקים לסווג סובייקטיבית ל-2 קטגוריות בלבד בעוד שהדירוג באתר הינו כללי ומספרי. הנוסף, קהל היעד אשר ענה על השאלון וקהל היעד לדירוגים באתר שונה. למרות העובדה שנבדקו דברים שונים (דירוג כללי לעומת סיווג בינארי), נמצא דמיון מלא בפרמטרים שנבדקו במכשירים שנבחרו למדגם. כלומר, לא מצאנו שוני בין השאלון לאתר ברמת הפרמטרים שנבדקו בניסוי, אלא מצאנו שוני בסוג השאלה הנבדקת, קהל היעד, אופן הניתוח, כמות הפרמטרים, כמות המכשירים המשוויים.

## **12. מסקנות:**

מהעברת השאלון למשתתפים למדנו כי אכן כפי שציפינו, סטודנטים בטכניון מסכימים כמעט בפה אחד כי טאבלטים באופן כללי הם אינם מוצר מותרות אלא כמוצר חיוני. למדנו איזה פרמטרים גורמים לאוכלוסיית היעד שלנו לחשוב כי מדובר בטאבלט שהוא מוצר מותרות ואילו פרמטרים כלל לא ישפיעו על החלטתם.

בנוסף, ראינו קורלציה בין הפרמטרים עליהם העידו הנבדקים כי עליהם הם התבססו לבין הפרמטרים בהם היה מתאם בינם לבין הדירוגים. לדוגמה, ראינו כי הייתה קורלציה בין המחיר לסווג הטאבלט, ובהתאמה, ראינו כי החשיבות אשר הנבדקים נתנו לפרמטר זה הינה הגבוהה ביותר. קיבלנו מידע איכותי מהשאלות הנוספות בהן למדנו על האופן בו סיווגו הנבדקים את המוצרים, האם הם בכלל התבססו על המידע שקיבלו, על איזה מידע הם התבססו יותר, על איזה פחות ועל אילו פרמטרים היה חסר להם מידע וכמובן משאלות הסיווג. לשיפור השאלות היינו מספקים את המידע שהנבדקים ביקשו בשאלה הפתוחה, שלטענתם היה עוזר להם במשימת הסיווג- פרמטרים שהינם אובייקטיביים וטכניים בלבד.

ביבליוגרפיה:

1. [/https://www.tgspot.co.il/tablet-sales-down-but-dont-tell-apple-2019](https://www.tgspot.co.il/tablet-sales-down-but-dont-tell-apple-2019)
2. [tps://tabletzona.es/iw/los-fabricantes-de-tablets-economicas-rascan-cuota-de-mercado-al-lider-apple](https://tabletzona.es/iw/los-fabricantes-de-tablets-economicas-rascan-cuota-de-mercado-al-lider-apple)
3. <https://gs.statcounter.com/vendor-market-share/tablet/worldwide>
4. [/https://www.statista.com/statistics/276635/market-share-held-by-tablet-vendors](https://www.statista.com/statistics/276635/market-share-held-by-tablet-vendors)
5. <https://www.marketingcharts.com/uncategorized-22427#:~:text=Women%2C%20who%20made%20up%2051,months%2C%2052%25%20are%20female>
6. <https://www.gadgety.co.il/6556/%D7%A1%D7%98%D7%98%D7%99%D7%A1%D7%98%D7%99%D7%A7%D7%95%D7%AA-%D7%A2%D7%9C-%D7%9E%D7%A9%D7%AA%D7%9E%D7%A9%D7%99-%D7%90%D7%A0%D7%93%D7%A8%D7%95%D7%90%D7%99%D7%93>
7. <https://thinkmonsters.com/speakinghuman/media/who-are-tablet-users-why-do-they-matter-to-your-organization>