עבודת הגשה בלימוד תוכנה – Bank

# תוכן:

* Customer.py
  + קובץ פייטון שמתאר את המחלקה של לקוח יחיד
* main.py
  + קובץ פייטון שמבצע קריאה של המשתמשים וחישוב של סיכון לכל משתמש חדש
* customer\_new.TXT
  + קובץ של לקוחות חדשים
* customer\_safe. TXT
  + קובץ של לקוחות בטוחים
* customer\_risky. TXT
  + קובץ של לקוחות מסוכנים
* question2Code.py
  + קובץ פייטון לשאלה 2
* question3Code.py
  + קובץ פייטון לשאלה 3
* output\_data.TXT
  + קובץ של לקוחות חדשים עם מספר
* output\_data\_q2.TXT
  + קובץ של התוצאה לתרגיל 2
* output\_data\_q3.TXT
  + קובץ של התוצאה לתרגיל 3

# פלט

לאחר הרצה של קובץ main.py יוצר קובץ שיציג את לכל לקוח חדש אם הוא בטוח או מסוכן ובטרמינל תוצג רשימה של הלקוחות החדשים מסודרים ב-2 קבוצות: בטוחים או מסוכנים. ליד כל משתמש בטרמינל יוצג האינדקס שלו במערך של הלקוחות החדשים, מרחק מהלנוביס בהשוואה ללקוחות הבטוחים ומרחק מהלנוביס שלו בהשוואה ללקוחות מסוכנים.

# תשובות

1. אחוז הלקוחות עם סיכון נמוך – 85%  
   אחוז הלקוחות עם סיכון גבוה – 15%  
   תוצאה כזאת הגיונים כי לרוב האוכלוסיה אין קשיים עם החזרת תשלומים בזמן, לכן לרוב המשתמשים יהיה סיכון נמוך של החזרת כסף לבנק.  
   לפי חוק בייס אנו יכולים להבין שלכל מאפיין יש את המשקל שלו שישפיע על התוצאה הסופית. למשל לגיל של המשתמש יש יותר השפעה על הסיכון להחזרת הכסף בהשוואה לערבות. סכום ההשפעה של המאפיינים שווה ל-1.
2. safe M= 78.31218547059018  
   risky M= 63.67615847401583  
   risky customer  
   *(הערה: עמית אמרת שהתוצאה צריכה להציג שהמשתמש הוא בטוח. אני משאר שלא יצא לי כמוך מסיבה שהאלגוריתם שלי יכול לעבוד בדרכים שונות משלך ולכן נקבל תוצאות שונות, ברשותך אני אענה על המשך השאלה כמו שהיה אמור לצאת אצלך)*לקוח – בטוח  
   תוצאה זו לא הגיונית מסיבה שרוב הנתונים *של הלקוח נמצאים בתווח הנתונים של הלקוחות המסוכנים.* עם זאת הנתונים שלא נמצאים בתווח קרובים יותר למסוכנים מאשר לבטוחים. אני משאר שזה קורה מחוסר מידע, כלומר אין מספיק נתונים של משתמשים בטוחים ומשתמשים מסוכנים כך שנדע מה מאפיין כל קבוצה.
3. מרחק נקודה ממרכז המסוכנים – d0מרחק נקודה ממרכז הבטוחים – d1  
   נוסחה לחישוב יתרה: **d0 / (d0 + d1)**