תרגיל 3 – מסדי נתונים – אופק קלי 314621897, אלדר עמר 214728131

*חלק א'-*

*2. א) השאילתה של* ChatGPT *מחשבת בפועל את הדברים הבאים:*

* *תחילה עושה מכפלה קרטזית של הטבלה* donors *עם עצמה.*
* *לאחר מכן יש סינון כך שנוצרים רק זוגות של תורמים שהשם של הראשון (d1.name) קטן אלפביתית מהשם של השני (d2.name), וגם הם תרמו לאותה העמותה (d1.cause = d2.cause). במילים אחרות נשאיר רק את השורות שבהן שני התורמים תרמו לאותה עמותה ועמותה אחת לפחות כזו.*
* *לאחר מכן עושים קיבוץ (*group by*) לפי השמות* (n1,n2) *– מצב זה גורם לכך שכל זוג יוצג פעם אחת, גם אם הם תרמו יחד ל־3 עמותות שונות (ייווצרו קודם 3 שורות ב־JOIN, ואז יתאחדו לקבוצה אחת).*
* HAVING *שבודק ש־* COUNT(DISTINCT d1.cause) = COUNT(DISTINCT d2.cause)*:*

*בגלל שהצירוף ב־*WHERE *כבר מגביל לשורות שבהן* d1.cause = d2.cause *כפי שראינו לעיל- הספירה של* d1.cause *ו־*d2.cause *בתוך אותה קבוצה תמיד תהיה זהה (הן סופרות בדיוק את מספר העמותות המשותפות). כלומר כל זוג ששיתף פעולה בתרומה לאותה עמותה יעמוד בתנאי הזה, גם אם אחד מהם תרם לעמותות נוספות שהשני אינו תרם אליהן.*

* *ולבסוף ישנו מיון לפי* n1,n2*.*

*לסיכום השאילתה פשוט מחזירה כל זוג של תורמים (שם1 < שם2 לפי אלפבית) שיש להם לפחות עמותה משותפת אחת לה הם תרמו.*

*התנאי ב-*HAVING *לא מבדיל בין מי שתורם בדיוק לאותן עמותות ובין מי שתורם גם לעמותות נוספות, ולכן השאילתה לא מוודאת שהקבוצות זהות בדיוק (כלומר שהתורמים עונים להגדרת זהה במטרותיו).*

*חלק ב':*

2. *ישנם 1000 שורות בטבלה, וכל שורה תופסת 64 בייטים. לכן ישנם בלוקים.  
   בנוסף יש שורות בבלוק.  
   בהנחה שאין אינדקסים, צריך לעבור על הטבלה עד שניתקל בערך בודד שגדול מ-1990.  
   מאחר ש-birthYear מתפלג אחיד על גבי יש 100 שורות ב-members*
3. *.*
4. *השלב הראשון, בו אנו יורדים בעץ, עלותו היא   
   בשלב השני יש לקרוא עלה אחד על מנת לוודא אם קיים כזה, עלותו   
   סה"כ 3.*
5. *ראשית יש לרדת שוב בעץ, 2. כעת מספר השורות הרלוונטיות הוא 100 ולכן הערכים יתפרשו על פני לכל היותר . סה"כ 4.*
7. *מכיוון שגודל מפתח החיפוש לא השתנה, וכן שאר הערכים, נקבל שוב*
8. *שוב ייקח לנו 2 פעולות לרדת בעץ ושוב עלה 1 לקרוא.  
   אחר כך על מנת לשלוף את name עלינו לקרוא את הבלוק הרלוונטי, עוד פעולה.  
   סה"כ 4 פעולות.*
10. *שלב ראשון ייקח פעולות.  
    שלב שני ייקח לקרוא את כל העלים הרלוונטיים. יש 200 שורות כאלה ולכן עלים.  
    שלב שלישי יהיה לקרוא את כל הבלוקים, נבחין כי ולכן נקרא את כל הבלוקים.  
    סה"כ פעולות.*
11. *כעת כל המידע הנדרש לשאילתה נמצא באינדקס ולכן לא נצטרך לקרוא את הטבלה.  
    שלב ראשון ייקח פעולות.  
    שלב שני יהיה לעבור על כל העלים הרלוונטים. יש 200 שורות כאלה ולכן עלים.  
    סה"כ 11 פעולות .*