

6.6.2018

מבני נתונים  
פתרון תרגיל מס' 9

1. שאלה זו הופיעה במבחן מועד א' 2017

במערכת המחשוב של חברת eBay Inc. (חברה אמריקאית שמנהלת את אתר המכירות eBay), מנהלים מעקב אחרי פעולות המשתמשים באתר המכירות eBay.

לכל משתמש שומרים: שם המשתמש (משמש לזיהוי המשתמש), התאריך שבו הצטרף לאתר, כתובת המגורים של המשתמש שכוללת את המדינה אליה שיך המשתמש, וכן נתונים על כל פעולות הרכישה/מכירה/החזרה שביצע המשתמש באתר.

לכל רכישה שומרים: מאפני הרכישה שכוללים את שם המשתמש של הקונה, שם המשתמש של המוכר, תאריך הרכישה, ושם המוצר שנרכש. בנוסף לכך שומרים את מחיר המוצר שנרכש.

שימו לב שרכישה של מוצר על ידי הקונה היא למעשה גם מכירה של המוצר על ידי המוכר.

לצורך השאלה יש להניח שבכל רכישה נרכש מוצר אחד בלבד.

לכל החזרה שומרים: מאפני ההחזרה שכוללים את מאפני הרכישה של המוצר שהוחזר ואת תאריך ההחזרה.

נגדיר שמכירה היא מכירה מוצלחת אם המוצר שנמכר במכירה זו לא הוחזר.

נגדיר את אחוז המכירות המוצלחות של משתמש כסך כל המכירות המוצלחות של המשתמש לחלק בסך כל המכירות של המשתמש כפול 100.

הערה: האופן בו נשמרים הנתונים הנ"ל אינו מפורט ויהיה עליך לצין אותו כחלק מפתרון השאלה.

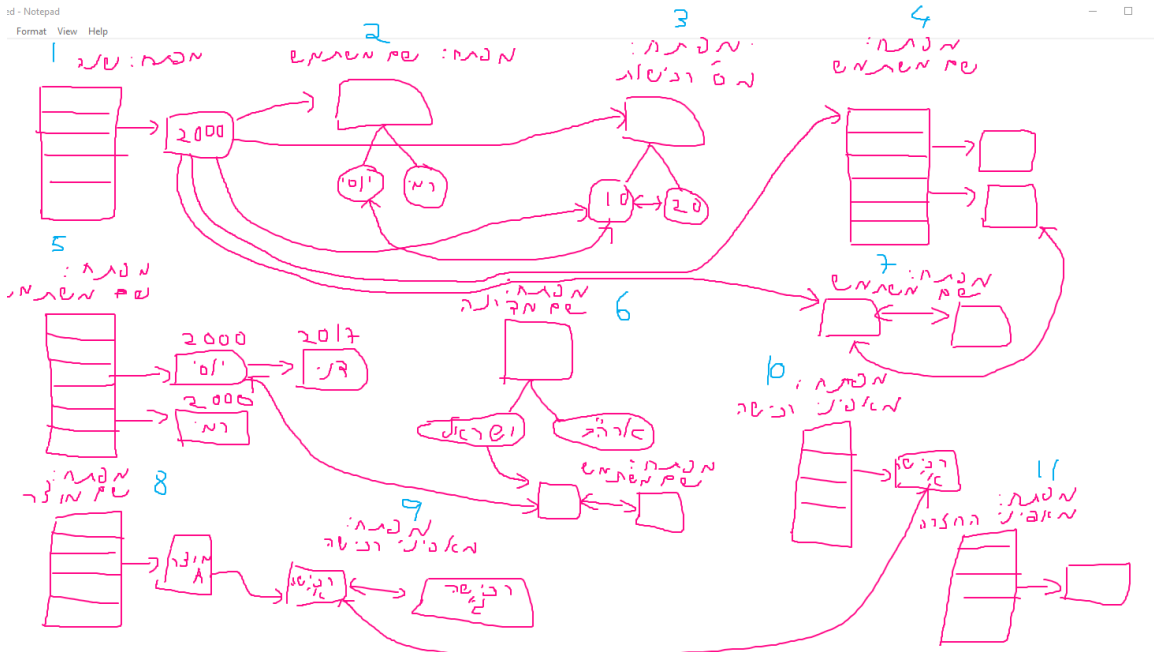
המשך בעמוד הבא...

- הצע/הציעי מבנה נתונים עבור המערכת הנ"ל ששומר את הנתונים הנ"ל ותומך בפעולות הבאות:
- הוספת/הוצאת הנתון שהתבצעה רכישה מסוימת מתבצעת בזמן  $O(\log c + \log p)$  בממוצע כאשר  $c$  מציין את מספר המדינות שיש להן משתמשים באתר eBay ו  $p$  מציין את מספר המשתמשים באתר eBay.
  - הוספת/הוצאת הנתון שהתבצעה החזרה מסוימת, מתבצעת בזמן  $O(1)$  בממוצע. שימו לב שבמקרה של החזרה יש לעדכן עבור המכירה המתאימה להחזרה זו, שהמכירה היתה לא מוצלחת. מצד שני במקרה של החזרה, לא מעדכנים את סך כל המכירות של המשתמש שמכר את המוצר, ולא מעדכנים את סך כל הרכישות של המשתמש שרכש את המוצר.
  - בהינתן שם משתמש, הדפסת כל המשתמשים שהצטרפו לאתר באותה שנה בה הצטרף המשתמש, ממוינים לפי סך כל המוצרים שרכשו המשתמשים, מתבצעת בזמן  $O(m)$  בממוצע כאשר  $m$  מציין את מספר המשתמשים ברשימה שתודפס.
  - בהינתן שם מדינה, הדפסת שמות כל המשתמשים ששיכים למדינה זו ומכרו לפחות מוצר אחד, ממוינים לפי סך כל המוצרים שמכרו, מתבצעת בזמן  $O(\log c + x)$  במקרה הגרוע, כאשר  $c$  מציין את מספר המדינות שיש להן משתמשים באתר eBay ו-  $x$  מציין את מספר המשתמשים ברשימה שתודפס.
  - בהינתן שנה ומספר  $k$ , הדפסת  $k$  המשתמשים עם אחוזי המכירות המוצלחות הגבוהים ביותר בשנה זו, מתבצעת בזמן  $O(q + k \log k)$  בממוצע, כאשר  $q$  מציין את מספר המשתמשים שמכרו מוצרים בשנה זו.
  - בהינתן שם מוצר, הדפסת כל הרכישות של מוצר זה (ללא חשיבות לסדר הרכישות), מתבצעת בזמן  $O(x)$  בממוצע, כאשר  $x$ , מציין את מספר הרכישות ברשימה שתודפס.

בנוסף לתאור מבנה הנתונים שהצעת, תאר/י באופן מילולי איך מתבצעות שלושת הפעולות האחרונות.

# פתרון שאלה 1

## להלן ציור של המבנה המוצע:



להלן תוספות לתאור של המבנה המוצע שלא נראות בציור. ההתייחסות לכל מבנה היא לפי המספר שלו בציור.

מבנה 4 הוא טבלת hash של משתמשים שמכרו מוצרים בשנה מסוימת (שנת 2000 בדוגמה שבציור), לכל משתמש בטבלה זו שומרים את השדות הנוספים הבאים: סך כל המכירות של המשתמש בשנה זו, סך כל המכירות המוצלחות של המשתמש בשנה זו.

מבנה 5 הוא טבלת hash של משתמשים. לכל משתמש שומרים שני שדות נוספים: השנה שבה המשתמש הצטרף לאתר eBay (מתואר בציור, למשל שנת 2000 עבור המשתמש יוסי), סך כל המכירות של המשתמש (לא מתואר באתר).

להלן תיאור ביצוע 3 הפעולות האחרונות.

בהינתן שם מדינה, הדפסת שמות כל המשתמשים ששייכים למדינה זו ומכרו לפחות מוצר אחד, ממוינים לפי סך כל המוצרים שמכרו, מתבצעת באופן הבא: מחפשים את המדינה בעץ המדינות (מבנה 6) בציור  $O(\log c)$  במקרה הגרוע), ובעזרת המצביע מדפיסים את הרשמה המקושרת של כל המשתמשים שמכרו מוצר אחד לפחות ושייכים למדינה זו  $O(x)$  במקרה הגרוע), בסך הכל הפעולה לוקחת  $O(\log c + x)$  במקרה הגרוע.

בהינתן שנה ומספר  $k$ , הדפסת  $k$  המשתמשים עם אחוזי המכירות המוצלחות הגבוהים ביותר בשנה זו מתבצעת באופן הבא: מחפשים את השנה בטבלת hash של השנים (מבנה 1 בצירוף) ( $O(1)$  בממוצע) בעזרת המצביע מגיעים לרשימה המקושרת של המשתמשים שמכרו מוצרים בשנה זו (מבנה 7 בצירוף), עוברים על הרשימה המקושרת ובונים ערימה שמכילה את אחוזי המכירות המוצלחות של כל המשתמשים ( $O(q)$  במקרה הגרוע). את אחוז המכירות המוצלחות של כל משתמש מחשבים על ידי חלוקה של סך כל המכירות המוצלחות בסך כל המכירות של המשתמש ששמורים במבנה 4. לבסוף מוצאים את  $k$  המשתמשים בעלי אחוז מכירות גבוה ביותר מתוך הערימה, בעזרת האלגוריתם שהראינו בכיתה ( $O(k \log k)$  במקרה הגרוע). בסך הכל הפעולה לוקחת  $O(q+k \log k)$  בממוצע.

בהינתן שם מוצר, הדפסת כל הרכישות של מוצר זה (ללא חשיבות לסדר הרכישות), מתבצעת באופן הבא: מחפשים את המוצר בטבלה 8 ( $O(1)$  בממוצע), ובעזרת המצביע מדפיסים את כל הרכישות ברשימה המקושרת ( $O(x)$  במקרה הגרוע), בסך הכל הפעולה לוקחת  $O(x)$  בממוצע.