

## חלק 1 שאלה 2

א. השאילתה מחשבת את זוגות התורמים שתרמו לעמותה משותפת כלשהי (having תמיד יתן true), מסודר לפי ה-1n ואז 2n.

## חלק ב

### שאלה 1

א. העלות הינה קריאת כל הבלוקים מהזיכרון. נחשב כפי שראינו בהרצאה:

$$B(\text{members}) = \left\lceil \frac{1000}{\left\lceil \frac{1024}{64} \right\rceil} \right\rceil = \left\lceil \frac{1000}{16} \right\rceil = 63$$

לכן העלות **63** קריאות IO.

ב.

$$d \leq \frac{1024 + 4}{4 + 4} = 128.5$$

לכן  $d=128$  היא האופטימלית.

ג. העלות תכלול ירידה לעלים וקריאת עלה בודד שכן יש distinct ולכן מספיק שיש ערך יחיד.

$$\left\lceil \log_{\left\lceil \frac{128}{2} \right\rceil} 1000 \right\rceil + 1 = 2 + 1 = 3$$

לכן העלות **3** קריאות IO.

### שאלה 2

העלות תכלול ירידה לעלים, קריאה לעלה וקריאה לעלים מימין שיתכן שבהם ערכים שגדולים מ-1990, שכן אנו רוצים לקחת avg שלהם.

הפרמטר k הוא כמות השורות שאנו מעוניינים בהם, הערכים מתפלגים אחיד על [1900,2000] ולכן

$$k = 1000 * \frac{10}{100} = 100$$

$$\left\lceil \log_{\left\lceil \frac{128}{2} \right\rceil} 1000 \right\rceil + \left\lceil \frac{100}{\left\lceil \frac{128}{2} \right\rceil - 1} \right\rceil = 2 + 2 = 4$$

לכן העלות **4** קריאות IO.

### שאלה 3

א.

$$d \leq \frac{1024 + 4}{4 + 4} = 128.5$$

לכן  $d=128$  היא האופטימלית.

ב. העלות תכלול ירידה בעץ, קריאת העלה וגישה אליו כדי לקבל את name. כיוון שערך זה הינו

מפתח, אז יש לנו בטבלה רק שורה אחת מתאימה שהיא תהיה בעלה בודד שבו ערך בודד אליו צריך לגשת.

$$\left\lceil \log_{\left\lceil \frac{128}{2} \right\rceil} 1000 \right\rceil + 1 + 1 = 2 + 2 = 4$$

לכן העלות **4** קריאות IO.

#### שאלה 4

א.

$$d \leq \frac{1024 + 10}{10 + 4} = 73.8$$

לכן  $d=73$  היא האופטימלית.

ב. העלות כוללת ירידה בעץ, קריאת כל העלים בהם יש ערך Hebrew וגישה אליהם כדי לקבל את birthyear.

$k = \frac{1000}{5} = 200$  שכן כל 1000 השורות הערכים מתפלגים אחיד על 5. **כלומר צריך לגשת ל-200 בלוקים בזיכרון, אבל הטבלה מאוחסנת ב-63 בלוקים, לכן בהכרח הערכים בכולם. כלומר 63.**

$$\left\lceil \log_{\left\lceil \frac{73}{2} \right\rceil} 1000 \right\rceil + \left\lceil \frac{200}{\left\lceil \frac{73}{2} \right\rceil - 1} \right\rceil + 63 = \left\lceil \frac{3}{\log(37)} \right\rceil + \left\lceil \frac{200}{36} \right\rceil + 63 = 2 + 6 + 63 = 71$$

לכן העלות 71 קריאות IO. **נשים לב שבמקרה זה היה עדיף לקרוא את כל הבלוקים ולחפש ללא שימוש במבנה אינדקס.**

ג. הפעם מה שהשתנה מסעיף א זה שכל קודקוד בעל ערך נוסף, יחד birthYear language תופסות 14 ביטים.

$$d \leq \frac{1024 + 14}{14 + 4} = 57.6$$

לכן  $d=57$  היא האופטימלית.

ד. יורדים בעץ לעלים ואז עבור  $k=200$  שורות שהערך של שפה היא עברית, נקרא את כל העלים ללא צורך לגשת לטבלה שכן ערך birthyear שמור כבר בעץ.

$$\left\lceil \log_{\left\lceil \frac{57}{2} \right\rceil} 1000 \right\rceil + \left\lceil \frac{200}{\left\lceil \frac{57}{2} \right\rceil - 1} \right\rceil = \left\lceil \frac{3}{\log(29)} \right\rceil + \left\lceil \frac{200}{28} \right\rceil = 3 + 8 = 11$$

לכן העלות 11 קריאות IO.