קצת על המטלה:

חלק א

סעיף א : סיבוכיות זמן n+NlogN

n לקריאת השורות בקובץ (n שורות)

ו NlogNלמיון השכיחים הגדולים.

סיבוכיות מקום, O(N) לשמירת השכיחויות+ הקבצים שנוצרו.

סעיף ב:

1. זמן ריצה: O(N) מספר השורות.

סיבוכיות מקום: 12 מערכים של חודש, שכל אחד מהם מערך של 31 ימים, שכל אחד מהם מערך של 24 שעות. שכל זה בעצם: 12\*24\*30, עדיין מספר קבוע, אבל זה כפול השנים כי זה שמור במפה ששומרת שנה ותאריך. אז אם אפשר להגביל את מספר השנים זה עדיין יהיה קבוע. אם לא זה יהיה O(N) כאשר N מייצג את מספר השנים.

חוץ מזה יש לנו טבלה ששומרת את כל התאריכים בשביל למנוע כפיליות, זה לוקח גם O(N).

סיבוכיות המקום אכן גדולה אבל זה עוזר לסיבוכיות של זמן הריצה כי מעדכנים כל פעם את ערכי הממוצעים תוך כדי קריאת השורות מהקובץ.

2. כאשר יש T תהליכים בו זמנית N קבצים וM שורות בכל קובץ:

זמן ריצה: קודם O(N) כדי לחלק לקבוצות.

ואז לכל תהליך זמן הריצה הוא O(N\*M) אבל כיוון שכל פעם עובדים לנו במקביל T תהליכים אז הסיבוכיות זמן יכולה להתחלק בT.

סיבוכיות מקום: כעת נצטרך רק מערך של 24 שעות עבור כל תהליך כלומר O(T). + מפה לשמירת כל התאריכים כדי למנוע כפיליות O(T\*M)<n.

3. אם הסעיפים היו מגיעים בזרימה הייתי משתמשת בשיטה הראשונה כי זה מעדכן את הממוצעים בתוך המערכים בזמן אמת.