כתבתי את התרגיל בשפת C# visual studio- console app (.NET core)

השקעתי הרבה זמן ומחשבה ע"מ שהמימוש יהיה בסיבוכיות זמן נמוכה ככל האפשר,

תאור האלגוריתם:

עוברים על הרשימה של זוגות השמות הזהים (יצרתי אוביקט שמכיל שם 1, שם 2) במטרה לצור קבוצות של שמות זהים ומכניסים את השמות לטבלת HASH ולמערך שכל תא מכיל רשימת שמות ,באופן הבא:

* אם שתי השמות אינם נמצאים עדיין בטבלה מכניסים את שתיהם עם ערך זהה לטבלה ,הערך יהיה האינדקס במערך שם נכניס רשימה של שתי השמות .
* אם שם אחד נמצא , נכניס את השם השני לטבלת הHASH עם אותו ערך של השם שנמצא בטבלת הHASH,נגש למערך (במקום של הערך שמצאנו בטבלת הHASH) ונוסיף את השם לרשימה .
* אם שתי השמות נמצאים אך הערכים שונים, ניקח את הערך הנמוך ונעבור על רשימת השמות בערך הגבוה כל שם נכניס לרשימה של הערך הנמוך ונשנה את ערכו בטבלת הHASH

(יכול להתרחש במקרה שבו ניתן לכתוב את אותו השם ב5 אופנים שונים ,

לדוג' הזוג הראשון יהיה yakov-yaakov

הזוג השני יהיה yacov-yaacov

ורק בזוג העשירי נגלה ש yakov -yacov

ובעצם כל השמות באותה קבוצה)

במקרה ששתי השמות נמצאים עם ערכים זהים אין צורך לבצע פעולה.

לאחר מכן יוצרים רשימה חדשה של שמות ותדרויות (יצרתי אוביקט שמכיל שם ותדירות)– רשימה סופית,

עוברים על הרשימה המקורית ועבור כל שם בודקים האם קיים בטבלת הHASH , אם כן מכניסים לרשימה החדשה באותו באינדקס של ערך השם בטבלה , אם קיים איבר במערך – הכנסנו כבר שם זהה מוספים לתדירות את הסכום של השם הנוכחי.

אם השם לא נמצא בטבלת הHASH ,( אין לו שם דומה) הכנסתי באידקס הגבוה מאורך רשימת השמות הזהים .

ע"מ לקרוא מהקבצים ולכתוב לקובץ את הרשימה הסופית השתמשתי בספריה **IO.File**

(בקבצים עצמם בין שם לשם או שם לתדירות יש מקף ,ובין זוג לזןג מפריד פסיק ,

יצאתי מתוך נקודת הנחה שהקלט תקין)

בשביל רשימת השמות ניתן להכניס את הנתיב הבא :

..\..\..\names.txt

בשביל רשימת הזוגות ניתן להכניס את הנתיב הבא :

..\..\..\pairs.txt

הקובץ שנוצר נמצא באותו מיקום- בתוך תיקית הפרויקט עם השם final

(הדפסתי את תוכן הקבצים על המסך , ליתר נוחות )

מקווה שתהנו...