# חלק א'

**סעיף א' משימה 5**

סיבוכיות זמן ומקום:

L – מספר השורות בקובץ.

S – אורך שורה בקובץ (ניתן להתייחס כקבוע).

E – מספר שגיאות ייחודיות.

num – מספר הקבצים כמספר הליבות.

הפונקצייה MakeDirs שמחלקת את הקובץ לחלקים קטנים –

* זמן: קוראת את כל הקובץ פעם אחת כדי לספור כמה שורות ואז כותבת כל שורה לקובץ אחד, פעמיים O(L). סה"כ לינארית, כמספר השורות בקובץ **O(L)**.
* מקום: גודל הרשימה של שמות הקבצים, יוצרים קבצים כמספר התהליכונים – **O(num)**.

הפונקצייה CountErrorsFromFile שסופרת כמה מופעים יש לשגיאה מסוימת בקובץ –

* זמן: עבור כל שורה עוברים עליה וחותכים את שם השגיאה **O(L\* S)**. סה"כ כגודל הקובץ.
* מקום: כל פעם קוראים שורה אחת פלוס מילון של השגיאות היחודיות. סה"כ – **O(S+E)**.

הפונקצייה CountErrorsAndMergeWithTPL שמריצה את CountErrorsFromFile על כל הקבצים שנוצרו במקביל ומאחדת את הספירות –

* זמן: קוראת לכל קובץ בנפרד ועוברת על כל השורות שלו. סה"כ כמספר השורות של הקובץ הגדול כפול אורך השורה **O(L\*S)**. בקיצור גודל הקובץ.
* מקום: כמספר השגיאות הייחודיות – **O(E)**.

**סעיף ב' משימה 3**

לבדוק תקינות לשורה, לבצע המרות פורמטים, לוודא שאין ערכים NaN או NaT, בדומה לבדיקות בתחילת התרגיל.

מבנה נתונים שאשתמש בו:

מילון שהמפתח שלו הוא שעה התחלה בתאריך מסוים ובכל תא במילון יש:

* count מאותחל ב- 1.
* average מאותחל ב- value של ה- stream.

עבור כל שורה שמגיעה מה- stream ועברה את כל הבדיקות: לעגל את השעה כלפי מטה ולבדוק:

* האם זאת שעה שלא קיימת מילון – להוסיף אותה למילון כמפתח –

stream.timestamp ולאתחל את הערך: count=1 average=stream.value

* האם זאת שורה שקיימת במילון –

count+=1 average=(average\*(count-1)+stream.value)/count

כשרוצים לגשת לממוצע שעתי מסוים – פשוט ניגשים למקום במילון עם שעת ההתחלה ושולפים את average.

**סעיף ב' משימה 4**

בקוד – שינויים בטעינה ושמירה של הקובץ.

הבדלים בין הפורמטים:

csv פורמט טקסט פשוט שכל שורה היא רשומה, קל לעריכה ידנית. טבלה שטוחה, שומר את כל הערכים כמחרוזות.

parquet הוא דחוס יותר, שמור לפי עמודות ועם חיסכון בנפח אחסון. נשמר מידע על טיפוסי הנתונים לפי עמודה – אין צורך להמיר פורמטים וכד' כמו ב- csv.

ב- csv חייב לקרא את כל הקובץ מהדיסק ואז לבחור עמודות לזיכרון. לעומת זאת ב- parquet העמודות שמורות כבלוקים נפרדים ככה שאפשר מראש לדלג על עמודות לא רלוונטיות.

# חלק ג'

1. 2
   1. השלט פולט פולסים של אור אינפרה אדום – IR – שבנ"א לא יכולים לראות. המזגן כן מסוגל לקלוט. בשלטים מודרנים הבהוב האור שולח סדרה של מספרים בינאריים
   2. שלט: כפתורים, סוללות, דיודת אינפרא אדום.

מזגן: מקלט אינפרה אדום

* 1. כל כפתור המנורה מאירה תדר קצת שונה וככה המזגן יודע מה נלחץ (למשל 2 הבהובים בשניה עבור פעולה מסוימת). הלחיצה על הכפתור סוגרת מעגל חשמלי משתנה בשלט בהתאם לכל כפתור כך שנשלח אות שונה