3. אם הנתונים מגיעים בזרימה (stream) במקום מקובץ, כיצד תתכנני את הפתרון כדי

לעדכן את הממוצעים השעתיים בזמן אמת?

כל פעם שמגיע נתון חדש, נעבד אותו במקום לממוצע השעתי- נגדיר אוביקט שיעקוב אחרי ממוצעי השעות בו כל מפתח יהיה השעה והערך יהיה רשימה של נתונים.  
כל פעם שמגיע נתון חדש נעדכן את הרשומה המתאימה לשעה  
לאחר העידכון נחשב את הממוצע עבור כל שעה על ידי חישוב

4. לתרגיל מצורף קובץ נוסף בפורמט שונה המכיל את אותו מידע, התאימי את החלקים

הרלוונטיים בקוד על מנת לתמוך בפורמט החדש, וכתבי תיעוד המתאר את

היתרונות באחסון המידע בפורמט הנתון.

יתרונות:

* ביצועים ויעילות בזיכרון: **Parquet** הוא פורמט עמודתי- שומר את הנתונים בצורה עמודתית תוכנה זו מאפרת גישה מהירה ויעילה יותר לנתונים
* דחיסת נתונים: **Parquet** תומך בדחיסה מה שמצמצם משמעותית את גודל הקובץ כמו כן, טעינת נתונים דחוסים היא הרבה יותר מהירה
* תמיכה בסכמות: **Parquet** שומר את המידע על הסכמות כך שהוא יודע את סוגי הנתונים של כל עמודה. בכך נמנע שגיאות במהלך הקריאה או השמירה של הנתונים
* תמיכה בקבצים גדולים: **Parquet** מתאים יותר לעבודה עם קבצים גדולים מאוד