

תשובות מילוליות

חלק א' סעיף א, שאלה 5:

קריאת וחיתוך האקסל- $O(n)$

ספירת שגיאות בקבצים- $O(n)$

מיון השגיאות- $O(e \log e)$

כאשר:

n – מספר השורות בקובץ האקסל המקורי

e – מספר סוגי השגיאות השונים (בדרך כלל קטן)

סך הכול – סיבוכיות זמן הריצה: $O(n + e \log e)$

חלק א' סעיף ב, שאלה 3:

בזרימה של נתונים, אי אפשר לשמור את כל המידע בזיכרון, לכן צריך לעדכן את הממוצעים השעתיים בזמן אמת בעזרת שמירת סכום הערכים ומספרם לכל שעה. בכל שורה חדשה שמתקבלת, צריך לזהות את השעה המתאימה (למשל 14:00), לעדכן את הסכום והמונה, ולחשב ממוצע לפי סכום חלקי כמות. כך ניתן לנהל ממוצעים שעתיים ללא צורך לשמור את כל הנתונים.

חלק א' סעיף ב, שאלה 4:

לפורמט Parquet יש כמה יתרונות לאחסון נתונים:

- דחיסה יעילה – חיסכון בגודל הקובץ בזכות אחסון עמודתי.
 - ביצועים גבוהים – קריאה של חלק מהנתונים בלבד, מה שמייצל את הזמן.
 - אחסון עמודתי – מאפשר ביצוע שאילתות מהירות וממוקדות.
 - שילוב עם כלים מתקדמים – תמיכה במערכות כמו Apache Spark ו-AWS Athena.
 - תמיכה בטיפוסים שונים – שומר מידע על סוגי נתונים, מה שמקל על ניתוחם.
- סיכום: Parquet הוא פורמט אידיאלי לנתונים גדולים – קומפקטי, מהיר ומודרני יותר מ-CSV ו-Excel.

חלק ג':

1. פחות מעוניית, הייתי מדרגת כ-2.

2. אופן השידור בין השלט למזגן:

a. השלט והמזגן מתקשרים בדרך כלל דרך שידור אינפרה אדום (IR) או, בשלטים יותר חדשים, דרך בלוטות'. אינפרה אדום דורש שדה ראייה ישיר בין השלט למזגן, בעוד שבלוטות' יכול לעבוד במרחקים גדולים יותר.

b. רכיבים בצד השלט ובצד המזגן:

בשלט יהיו רכיבים כמו כפתורים, משדר אינפרה אדום או בלוטות', ומעבד שמחבר את כל המידע. במזגן יהיה מקלט אינפרה אדום או בלוטות', שמקבל את השידור ומעבד אותו, יחד עם מעבד שמבצע את הפעולה המתאימה (למשל שינוי טמפרטורה).

c. איך המזגן יודע איזה כפתור נלחץ בשלט: כשהשלט שולח שדרים למזגן, כל כפתור בשלט יכול לשלוח קוד שונה. לדוגמה, אם לוחצים על כפתור "הפחתת טמפרטורה", השלט שולח קוד מסוים. המזגן מקבל את הקוד הזה ומבצע את הפעולה המתאימה. אפשרות נוספת היא שכל כפתור בשלט משדר גל בתדר שונה, והמזגן מזהה את התדר ומתאים לו פעולה.