

פרויקט: מנתה לוגים של תעבורת רשת

שלב 2 - בניית מיליוןים

מטרת השלב

בשלב הקודם למדנוلسן וchlץ נתוני מהЛОג. בשלב זה נארגן את הנתונים במלוניים כדי לאפשר גישה מהירה וניתוח עיל.

נשתמש ב-**Dict Comprehension** ליצירת מיליוןים בצורה קצרה ויעילה.

דרישות שלב 2

לכל דריש יש לכתוב פונקציה נפרדת.

1 ספירת פניות לפי IP

כתבו פונקציה שמקבלת את הנתונים ומהזירה מילון: כתובות IP מקור → מספר ההפניות שלה.

דוגמה לפلت:

```
{ "192.168.1.100": 156, "45.33.32.156": 23,  
"10.0.0.50": 89 }
```

2 מיפוי פורט לפרוטוקול

כתבו פונקציה שמקבלת את הנתונים ומחזירה מילון: מספר פורט → שם הפרוטוקול.

דוגמה לפלט:

```
{ 443: "HTTPS", 80: "HTTP", 22: "SSH" }
```

3 זיהוי חדשות לכל IP

כתבו פונקציה שמקבלת את הנתונים ומחזירה מילון: כתובות IP → רשימת סוגי החשדות שלה.

הfonקציה תבדוק את 4 סוגי החשדות שהגדכנו:

- EXTERNAL_IP - כתובות חיצונית
- SENSITIVE_PORT - פורט רגיש (22,23,3389)
- LARGE_PACKET - חבילה מעל 5000 ביט
- NIGHT_ACTIVITY - פעילות בין 00:00 ל-06:00

דוגמה לפלט:

```
{ "45.33.32.156": [ "EXTERNAL_IP", "LARGE_PACKET" ],  
"87.120.5.22": [ "EXTERNAL_IP", "NIGHT_ACTIVITY",  
"SENSITIVE_PORT" ], "203.0.113.50": [ "EXTERNAL_IP" ] }
```

4 סינון מילון החשדות

כתבו פונקציה שמקבלת את מילון החשדות ומחזירה מילון חדש רק עם כתובות שיש להן לפחות 2 סוגי חשדות.

או עם חשב אחד עדין יכול להיות תקין. או עם כמה חשבות יחד - סביר יותר שמדובר
באיום אמיתי.

זכרו: המבנה צריך להיות דינامي. כනוסיף סוג חשב חמישי או שישי בעתיד, השינוי בקוד
צריך להיות מינימלי.