

הסבר למי שמתקשה באנגלית, לתרגיל שמתואר בקובץ wordcount.py (בתיקה הנוכחית) זהו תירגול של עבודה עם מילון.

השלד כבר נתון לכם. התוכנית מקבלת 2 פרמטרים מבחוץ (sys.argv)

הפרמטר הראשון יכול להיות אחד משתי אפשרויות: -count - או -topcount - (זה 2 מינוסים צמודים אחד לשני וצמודים למילה שאחרי).

הפרמטר השני הוא שם הקובץ שגם נימצא בתיקה הנוכחית - alice.txt

כפי שאתם רואים בקוד של wordcount.py הוא קורא לשתי פונקציות שונות בהתאם לפרמטר הראשון. print_words או print_top עם פרמטר שהוא שם הקובץ שקיבלנו מ- sys.argv (כפרמטר שני עבור התוכנית).

אתם צריכים לכתוב את print_count ואת print_top.

שתיהן עושות דבר מאד דומה עד שלב מסוים. הן קוראות את הקובץ וסופרות כמה פעמים מופיעה כל מילה בו (בעזרת מילון, כפי שתיארנו בכיתה). לאחר מכן, אחת (print_words) מדפיסה את המילים ואת מספר המופעים של כל מילה ליד, בצורה ממוינת על פי המילים בסדר אלפביתי עולה. לדוגמה: (לא מילים מהקובץ, רק להראות איך יראה)

```
apron 10
bye 6
column 30
```

השניה, print_top, גם תדפיס בצורה דומה, אבל הפעם זה יהיה ממויין בסדר יורד של מספר המופעים. לדוגמה:

```
is 1700
the 600
yes 250
```

כדי לא להסתבך עם אותיות גדולות וקטנות (למשל The זה לא אותו דבר כמו the) יש להסב את כל המילים לאותיות קטנות.

מומלץ לעבוד בשלבים, כלומר בכל פעם שביצעתם צעד, להדפיס את התוצאה עד כה ולסיים את התוכנית עם sys.exit(0). כדי להפריד מילים אפשר לעשות split על רווח. נסו להיפטר מסימני הפיסוק שצמודים לחלק מהמילים, כדי שלמשל: is ו- is לא תיספרנה כמילים ניפרדות. במצגת 12 ישנם כל ההסברים הדרושים בקשר למילון והמיון של התוצאות.

בהצלחה!