

יסודות מערכות פתוחות
פתרון תרגיל מס' 10

1. שאלה זו הופיעה במבחן מועד א 2015

ברשת הסניפים של רמי לוי שומרים נתונים על רכישות מוצרים כפי שמתואר להלן.

עבור כל סניף ברשת שומרים בקובץ (ששמו כשם הסניף) את הנתונים הבאים:

מספר קופה, מספר פריט, כמות פריטים, מחיר ליחידה, תאריך מכירה.

לדוגמה נניח שיש שורה בקובץ ששמו `zichron` שמכילה:

```
3 1234 2 30 20/2/2014
```

אז המשמעות היא שבקופה מספר 3 בתאריך 20/2/2014 התבצעה מכירה של 2 יחידות של פריט שמספרו 1234 כאשר מחיר כל פריט היה 30 ש"ח.

כתוב/כתבי תוכנית ב- `awk` בשם `p10.1` שמקבלת כפרמטר שם קובץ שמתאר את הנתונים של סניף מסוים ברשת הסניפים של רמי לוי הפורמט של שורות הקובץ כפי שתואר לעיל וכפי שמפורט בדוגמה שלהלן.

התכנית `p10.1` מדפיסה לפלט שורה אחת שמכילה את מספר הפריט שהסכום שהתקבל מכל המכירות שלו הוא הגדול ביותר, לאחר מכן תו רווח אחד ולאחר מכן מספר שמציין את סך כל המכירות של הפריט.

בתכנית `p10.1` מותר להשתמש לכל היותר ב- 4 פקודות `system`.

ראה/י דוגמה בעמוד הבא.

לדוגמה:

נניח שתוכן הקובץ zichron הוא:

מספר קופה	מספר פריט	מספר יחידות	מחיר יחידה	תאריך
3	1234	2	30	20/2/2014
4	1111	4	20	21/1/2014
2	2222	2	40	10/3/2014
3	1234	3	30	20/2/2014
1	1234	2	30	20/2/2014
2	1111	1	20	21/1/2014
3	1234	2	30	20/2/2014
1	1111	3	20	21/1/2014
3	3333	2	50	21/1/2014
6	4444	2	60	20/2/2014
3	2222	2	40	21/1/2014

לאחר הפעלת התכנית P10.1 ע"י הפקודה:

P10.1 zichron

מתקבל הפלט:

1234 270

סכום המכירות הכולל הגדול ביותר של פריט בקובץ הוא של הפריט 1234 והסכום הוא 270. ולכן הודפסה השורה הנ"ל.

החישוב של הסכום הכולל של 270 ש"ח שהתקבל עבור הפריט שמספרו 1234 הוא כתוצאה מהחישוב של סך כל המכירות הבאות:

2 פריטים מקופה 3 (שורה ראשונה בקובץ)
3 פריטים מקופה 3 (שורה רביעית בקובץ)
2 פריטים מקופה 1 (שורה חמישית בקובץ)
2 פריטים מקופה 3 (שורה שביעית בקובץ)

סה"כ נמכרו 9 פריטים כאשר המחיר ליחידה של כל פריט היה 30 ש"ח ולכן הסכום הכולל שהתקבל הוא 270 ש"ח.

פתרון שאלה 1

```
#!/bin/awk -f
{
  P[$2]=P[$2]+$3*$4
}
END {
  for ( b in P ){
    print P[b]" "b > "tmp"
  }
  system("sort -nr tmp | head -1 >| tmp1 ")
  while(getline < "tmp1"){
    print $2" "$1
  }
  close("tmp")
}
```

2. שאלה זו הופיעה במבחן מועד א 2015

כתוב/כתבי תוכנית ב-awk בשם P10.2 שמקבלת כפרמטרים רשימת שמות קבצים שמתארים את הנתונים של סניפים ברשת הסניפים של רמי לוי.

הפורמט של שורות הקבצים ברשימה הוא כפי שתואר לעיל וכפי שמפורט בדוגמה שלהלן.

התכנית P10.2 מדפיסה לפלט עבור כל קובץ ברשימה: שם הקובץ (בשורה נפרדת) ולאחריו תו רווח אחד ולאחריו הפלט שמתקבל עבור קובץ זה כפי שתואר בתרגיל 1.

על סדר הקבצים בפלט להיות לפי סך כל המכירות של הפריטים שמודפסים בסדר מספרי עולה.

בתכנית P10.2 מותר להשתמש לכל היותר ב-4 פקודות system.

לדוגמה,

נניח שתוכן הקובץ netanya הוא:

תאריך	מחיר יחידה	מספר יחידות פריט	מספר קופה
20/2/2014	30	1	1234
21/1/2014	20	15	1111
20/2/2014	30	3	1234
21/1/2014	20	1	1111

נניח שתוכן הקובץ pardes-hana הוא:

```
1      6666      1      70      20/2/2014
4      1234      3      20      21/1/2014
```

לאחר הפעלת התכנית P10.2 ע"י הפקודה:

```
P10.2 zichron netanya pardes-hana
```

מתקבל הפלט:

```
pardes-hana 6666 70
zichron 1234 270
netanya 111 320
```

פתרון שאלה 2

```
#!/bin/awk -f
BEGIN {
    printf "" > "tmp"
    for (i=1; i < length(ARGV) ; i++){
        system("echo $( P10.1 "ARGV[i]" ) "ARGV[i]" >> tmp")
    }
    while (getline < "tmp"){
        print $2" "$1" "$3 > "tmp1"
    }
    system("sort -n tmp1 >| tmp2")
    while (getline < "tmp2"){
        print $3" "$2" "$1
    }
}
```

כתוב תכנית ב- awk בשם P10.3 המקבלת כפרמטרים רשימת שמות קבצים שמכילים מטריצות (לא בהכרח ריבועיות ולא בהכרח עם מספר שווה של עמודות בכל שורה ולא בהכרח של מספרים בלבד) ומדפיסה שורה אחת עבור כל קובץ שמכילה את שם הקובץ, לאחריו תו רווח אחד ולאחריו רשימת מספרים שמציינים את סכומי העמודות של המטריצה (כאשר בין המספרים ברשימה ישנו רווח אחד בדיוק).

בחישוב של סכומי העמודות במטריצה יש להתחשב בכלל הבא: אם בעמודה מסוימת מופיעה מחרוזת שאינה מספר (דהינו המחרוזת אינה מורכבת מספרות בלבד) אז ערך המחרוזת יהיה 1.

על סדר שורות הפלט להיות לפי סדר הקבצים ברשימת הפרמטרים לתכנית.

הנח/הניחי שהקבצים ששמם מועבר בפרמטרים קיימים ותוכנם חוקי (דהינו השורות שלהם מכילות מספרים עם תו רווח אחד או יותר בין המספרים).

דוגמה

נניח ש-

תוכן קובץ F2 הוא:

```
6 3a 9
8
1 13 8 10
```

תוכן קובץ F3 הוא:

```
-200 202 cd abc def
200 8
8 -200 4
```

תוכן קובץ G הוא:

```
-60 10 -1 ab2d -2
10
50 -11 1 -2 3
```

תוכן קובץ A הוא:

```
60
```

לאחר הפעלת התכנית ע"י הפקודה P10.3 F2 F3 G A יתקבל הפלט:

```
F2 15 14 17 10
F3 8 10 5 1 1
G 0 -1 0 -1 1
A 60
```

פתרון שאלה 3

```
#!/bin/awk -f
{
  for(j=1; j<=NF; j++){
    if (sub("(^-$)|(^-.*[^0-9])|(^[-0-9])|(^[0-9].*[^0-9])","&",$j)>0) {
      A[FILENAME,FNR,j]=1
    }
    else {
      A[FILENAME,FNR,j]=$j
    }
  }
  B[FILENAME]=FNR
  if(NF > C[FILENAME]) {
    C[FILENAME] = NF
  }
}
END{
  for(x=1; x<ARGC; x++){
    s=ARGV[x]
    for(j=1; j<=C[ARGV[x]]; j++){
      sum_col=0
      for(i=1; i<=B[ARGV[x]]; i++){
        sum_col+=A[ARGV[x],i,j]
      }
      s=s" "sum_col
    }
    print s
  }
}
```

הסבר לביטוי הרגולארי

הביטוי הרגולארי בפקודה sub בודק שהמילה היא מספר באופן הבא. נבדקים 4 תנאים, אם אחד מהתנאים יתקיים אז זה סימן שהמילה שנבדקת אינה מספר. אחרת, זה אומר שכל 4 התנאים לא מתקיימים ואז המילה היא מספר.

4 התנאים הם:

1. המילה היא - בלבד
2. המילה מתחילה ב - ולאחר מכן יש תו שאינו ספרה
3. המילה מתחילה בתו שאינו ספרה ואינו -
4. המילה מתחילה בספרה ולאחר מכן בהמשך ישנו תו שאינו ספרה.