

Άσκηση 2 (Παράδοση μέχρι 23:55, 29/11/2019)

Συνολικό Βάρος άσκησης: 32% του τελικού βαθμού εργαστηρίου

Πρόγραμμα computeSales.py

Γράψτε ένα πρόγραμμα **computeSales.py** σε Python που να διαβάζει αρχεία με αποδείξεις πώλησης προϊόντων, να μπορεί να κάνει έλεγχο ορθότητας και να μπορεί να εκτυπώνει μετά στατιστικά στοιχεία που του ζητούνται. Πιο αναλυτικά:

Αρχεία Εισόδου:

Κάθε αρχείο εισόδου θα είναι ένα ASCII αρχείο της μορφής:

```
-----
ΑΦΜ: AAAAAAAAAA
ΟΝΟΜΑ_ΠΡΟΪΟΝ: ΠΟΣΟΤΗΤΑ ΤΙΜΗ_ΜΟΝΑΔΑΣ ΣΥΝΟΛΙΚΗ_ΤΙΜΗ
...
ΟΝΟΜΑ_ΠΡΟΪΟΝ: ΠΟΣΟΤΗΤΑ ΤΙΜΗ_ΜΟΝΑΔΑΣ ΣΥΝΟΛΙΚΗ_ΤΙΜΗ
ΣΥΝΟΛΟ: ΠΟΣΟ_ΜΕ_ΣΥΝΟΛΟ_ΑΠΟΔΕΙΞΗΣ
-----
ΑΦΜ: BBBB BBBB
ΟΝΟΜΑ_ΠΡΟΪΟΝ: ΠΟΣΟΤΗΤΑ ΤΙΜΗ_ΜΟΝΑΔΑΣ ΣΥΝΟΛΙΚΗ_ΤΙΜΗ
...
ΟΝΟΜΑ_ΠΡΟΪΟΝ: ΠΟΣΟΤΗΤΑ ΤΙΜΗ_ΜΟΝΑΔΑΣ ΣΥΝΟΛΙΚΗ_ΤΙΜΗ
ΣΥΝΟΛΟ: ΠΟΣΟ_ΜΕ_ΣΥΝΟΛΟ_ΑΠΟΔΕΙΞΗΣ
-----
```

Παράδειγμα μίας απόδειξης είναι η ακόλουθη:

```
-----
ΑΦΜ: 0123456789
ΧΩΡΙΑΤΙΚΗ: 3 7.00 21.00
ΤΖΑΤΖΙΚΙ: 1 3.50 3.50
ΠΟΙΚΙΛΙΑ: 2 14.20 28.40
ΣΥΝΟΛΟ: 52.90
-----
```

- 1) Κάθε αρχείο περιέχει 1 ή περισσότερες αποδείξεις (αλλά τουλάχιστον 1)
- 2) Η κάθε απόδειξη ξεκινάει και τελειώνει με γραμμές που περιέχουν μόνο το χαρακτήρα '-'.
- 3) Οι αποδείξεις διαχωρίζονται μεταξύ τους από γραμμές που περιέχουν μόνο το χαρακτήρα '-'.
- 4) Η πρώτη γραμμή κάθε απόδειξης περιέχει το ΑΦΜ της εταιρείας (10ψήφιο νούμερο).
- 5) Η τελευταία γραμμή κάθε απόδειξης περιέχει το σύνολο των αγορών για την απόδειξη αυτή.
- 6) Σε ένα αρχείο μπορούν να υπάρχουν αποδείξεις διαφορετικών εταιρειών (ΑΦΜ).
- 7) Κάθε απόδειξη περιλαμβάνει την πώληση πολλών προϊόντων (1 ή περισσότερα).
- 8) Ένα προϊόν μπορεί να εμφανίζεται σε μία απόδειξη παραπάνω από 1 φορά. Πχ, στην απόδειξη του παραδείγματος, το προϊόν ΧΩΡΙΑΤΙΚΗ μπορεί να εμφανιστεί ξανά αργότερα στην ίδια απόδειξη.
- 9) Ο αριθμός κενών στις γραμμές μεταξύ των πεδίων είναι μεταβλητός και το πρόγραμμά σας πρέπει να μην κάνει υποθέσεις για τον αριθμό των κενών χαρακτήρων.
- 10) Κάθε απόδειξη που περιέχει σφάλμα θα πρέπει να παραλείπεται, αλλά οι υπόλοιπες αποδείξεις του αρχείου θα πρέπει να διαβάζονται και να αποθηκεύονται. Σφάλματα στις αποδείξεις είναι πολλά, όπως το να μην υπάρχει το ΑΦΜ στην 1^η γραμμή της απόδειξης, να

υπάρχει λανθασμένο ΑΦΜ, το να μην υπάρχουν όλα τα πεδία σε κάθε προϊόν, το να μην έχει γίνει σωστά ο υπολογισμός της συνολικής τιμής ενός προϊόντος (ο πολλαπλασιασμός ποσότητα * τιμή_μονάδας), να μην υπάρχει ως τελευταία γραμμή η γραμμή ΣΥΝΟΛΟ, να μην είναι σωστό το σύνολο, κτλ).

ΛΕΙΤΟΥΡΓΙΚΟΤΗΤΑ ΠΡΟΓΡΑΜΜΑΤΟΣ

Το πρόγραμμά σας πρέπει να τυπώνει επαναληπτικά το ακόλουθο μενού χρήστη:

Give your preference: (1: read new input file, 2: print statistics for a specific product, 3: print statistics for a specific AFM, 4: exit the program)

- Αν ο χρήστης πατήσει 1, τότε πρέπει να ζητείται το όνομα του αρχείου. Αν δοθεί μη έγκυρο αρχείο, τότε επιστρέφουμε στο αρχικό μενού (Give your preference...).
- Αν ο χρήστης πατήσει 2, τότε πρέπει να ζητείται το όνομα του προϊόντος και θα τυπώνεται το **συνολικό ποσό** πωλήσεων του προϊόντος ανά ΑΦΜ. Θα πρέπει να τυπώνεται 1 γραμμή για κάθε ΑΦΜ (και μόνο για αυτά) στο οποίο υπάρχει τουλάχιστον μία πώληση του προϊόντος. **Τα ΑΦΜ θα πρέπει να εμφανίζονται σε αύξουσα σειρά.** Η κάθε γραμμή θα περιέχει 2 αριθμούς μόνο (χωρίς κανένα κείμενο): το ΑΦΜ και το συνολικό ποσό **πωλήσεων** (χωρισμένα με κενό). **Αν δεν έχει υπάρξει πώληση του συγκεκριμένου προϊόντος από κανένα ΑΦΜ, δε θα πρέπει να τυπώνεται τίποτα (ούτε μήνυμα σφάλματος).**
- Αν ο χρήστης πατήσει 3, τότε πρέπει να ζητείται το ΑΦΜ και θα τυπώνεται το άθροισμα πωλήσεων ανά προϊόν για το συγκεκριμένο ΑΦΜ. Θα πρέπει να τυπώνεται 1 γραμμή για κάθε προϊόν που έχει πωληθεί από το συγκεκριμένο ΑΦΜ. **Τα προϊόντα θα πρέπει να εμφανίζονται σε αύξουσα αλφαβητική σειρά.** Η κάθε γραμμή θα περιέχει 2 αριθμούς μόνο (χωρίς κανένα κείμενο): το ΟΝΟΜΑ_ΠΡΟΪΟΝΤΟΣ και το συνολικό ποσό (χωρισμένα με κενό). **Αν δεν έχει υπάρξει πώληση κανενός προϊόντος από το συγκεκριμένο ΑΦΜ, δε θα πρέπει να τυπώνεται τίποτα (ούτε μήνυμα σφάλματος).**
- Αν ο χρήστης πατήσει 4, τότε θα προκαλείται έξοδος από το πρόγραμμα.
- Αν ο χρήστης πατήσει οτιδήποτε άλλο, θα πρέπει να ζητείται ξανά η είσοδός του.

ΣΗΜΑΝΤΙΚΕΣ ΣΗΜΕΙΩΣΕΙΣ:

1. Τα δεδομένα από διαφορετικά αρχεία συνυπολογίζονται. Όταν πατήσουμε την επιλογή 1, τότε οι αποδείξεις κάθε εταιρείας συνυπολογίζονται σε αυτές που είχαμε διαβάσει από προηγούμενα αρχεία.
2. Θα πρέπει να σκεφτείτε ποιες δομές δεδομένων θα χρησιμοποιήσετε για να απαντήσετε όσο πιο γρήγορα γίνεται τα ερωτήματα στις επιλογές 2 και 3 του μενού χρήστη. Σημασία δεν έχει μόνο η ορθότητα των αποτελεσμάτων, αλλά και το αν έγινε αποδοτική η υλοποίηση. Κανένας δε χρησιμοποιεί ένα πρόγραμμα που είναι πολύ αργό. Ο έλεγχος στο πρόγραμμά σας θα γίνει με πολύ μεγάλα αρχεία που θα περιέχουν εκατοντάδες χιλιάδες γραμμές. Θα πρέπει να μην έχετε προβλήματα μνήμης και να απαντάτε αποδοτικά τα ερωτήματα αυτά.
3. Τα δεδομένα εισόδου και τα δεδομένα στις αποδείξεις ΔΕΝ είναι case-sensitive. Το προϊόν **ΧΩΡΙΑΤΙΚΗ** είναι το ίδιο με το προϊόν **χωριατική**. Προτείνω να μετατρέπετε όλα τα προϊόντα σε κεφαλαία γράμματα.
4. Μέσα στην εβδομάδα θα σας δοθεί ένα μικρό αρχείο με λίγες αποδείξεις, που δεν περιέχουν σφάλματα, καθώς και ένα αρχείο με είσοδο από το χρήστη. Αν το πρόγραμμά σας δεν τρέχει ούτε στη συγκεκριμένη απλή και δεδομένη είσοδο χρήστη για το συγκεκριμένο αρχείο, τότε θα βαθμολογηθεί με μέγιστο βαθμό το 20% αυτής της εργασίας (δηλαδή με μέγιστο βαθμό το 20% του 32% του συνολικού βαθμού όλων των projects).
5. **Για το διάβασμα ελληνικών χαρακτήρων από το αρχείο, πρέπει να το ανοίξετε κατάλληλα, δηλώνοντας το encoding. Πχ, open(filename, read_write_flags, encoding='utf-8')**

ΣΗΜΕΙΩΣΗ: Το πρόγραμμά σας θα πρέπει στην εκτύπωση να ακολουθεί **ΑΠΟΛΥΤΑ** τη μορφή του παραπάνω μενού και να τυπώνει τα αποτελέσματα με το συγκεκριμένο τρόπο, χωρίς άλλες γραμμές ή strings.
ΟΔΗΓΙΕΣ ΥΠΟΒΟΛΗΣ

A) Δημιουργήστε ομάδα εργασίας στο μάθημα. Έστω **LAB211XXXXX** η ομάδα σας

B) Δημιουργήστε ένα φάκελο με όνομα ίδιο με το όνομα της ομάδας σας (**LAB211XXXXX**)

Γ) Ο φάκελος μέσα θα πρέπει να περιέχει **ακριβώς 2 αρχεία**: το πρόγραμμα **computeSales.py**, και το αρχείο **report.pdf** με την αναφορά της άσκησης. Είναι **ΣΗΜΑΝΤΙΚΟ** να είναι με την παραπάνω ονομασία ο φάκελος και τα αρχεία σας.

Δ) Για να δημιουργήσετε το αρχείο υποβολής του project σας κάντε cd στο γονικό φάκελο του project σας (βγείτε έξω από το φάκελο **LAB211XXXXX** του project) και δημιουργήστε ένα συμπιεσμένο αρχείο με την εντολή

tar zcvf LAB211XXXXX.tar.gz LAB211XXXXX

Εργασίες που δεν έχουν ακολουθήσει τις παραπάνω οδηγίες θα έχουν σημαντική ποινή στη βαθμολογία.

Τα διαδικαστικά του μαθήματος περιγράφουν τον τρόπο βαθμολόγησης και τον κώδικα δεοντολογίας. Υπενθυμίζω ότι στην αρχή της αναφοράς σας ΠΡΕΠΕΙ να περιγράψετε το ποιο κομμάτι της άσκησης υλοποίησε το κάθε μέλος της ομάδας. **Η αναφορά βαθμολογείται.**

Η αναφορά πρέπει ΣΙΓΟΥΡΑ να περιέχει αναλυτική εξήγηση για το (α) Πώς διαβάσετε κάθε απόδειξη, (β) Πώς ελέγχετε αν η απόδειξη είναι σωστή, (γ) Τι δομές δεδομένων χρησιμοποιείτε για να απαντάτε γρήγορα τα ερωτήματα του χρήστη και γιατί τις επιλέξατε, (δ) Πώς εξασφαλίζετε ότι μπορείτε να διαβάσετε πάρα πολύ μεγάλα αρχεία.

Καλή επιτυχία!