

Εργαλεία Ανάπτυξης Λογισμικού και Προγραμματισμός Συστημάτων

Εργαστήριο 2

LAB21143027

- Παναγιώτης Σταυριανός
- ΑΜ: 2012030076
- email: pstavrianos@isc.tuc.gr

Δεν έχω άλλα άτομα στην ομάδα μου. Παρακάτω θα βρείτε όλη τη διαδικασία που ακολούθησα, καθώς και συνδέσμους για τα αρχεία κώδικα που χρησιμοποίησα.

Πρόγραμμα computeSales.py

Δημιουργία δεδομένων εισόδου

Για τη δημιουργία δεδομένων εισόδου ένα python script δημιουργεί τυχαίες αποδείξεις, κάνοντας τυχαία λάθη.

computeSales.py

Διαδικασία διαβάσματος απο το αρχείο

read_file

1. Διαβάζω γραμμή-γραμμή το αρχείο
 1. Αποθηκεύω σε ένα προσωρινό string την κάθε γραμμή (concat)
2. Εάν η γραμμή ξεκινάει με - καλώ τη συνάρτηση save_item και συνεχίζω τη διαδικασία

save_item

Η save_item παίρνει ως παράμετρο ένα string με μια πιθανή απόδειξη (χωρίς τη διαχωριστική γραμμή).

1. Εάν υπάρχουν χαρακτήρες tab \t τους αντικαθιστώ με κενά \s
2. Χρησιμοποιώντας regex r'\s +' διαγράφω τα περιττά κενά που μπορεί να υπάρχουν στην απόδειξη
3. Χρησιμοποιώντας την split('\n') χωρίζω την απόδειξη σε ανα γραμμές
4. Χρησιμοποιώντας filter διαγράφουμε ότι κενές γραμμές μπορεί να υπάρχουν
5. Χρησιμοποιώντας s.strip() διαγράφουμε whitespaces που μπορεί να έχουν μείνει
6. Εάν οι γραμμές που έχουν μείνει είναι λιγότερες από τρεις, η απόδειξη είναι εσφαλμένη και την απορρίπτουμε
7. Διαβάζουμε την πρώτη γραμμή που βρίσκεται το ΑΦΜ και τη διαγράφουμε(prop)
8. Έλεγχος του ΑΦΜ. Σύμφωνα με τους ευρωπαϊούς κανόνες το ΑΦΜ μπορεί να είναι ανάμεσα σε 8 και 16 αλφαριθμητικούς χαρακτήρες(δηλαδή χωρίς κενά και ειδικά σύμβολα)
9. Διαβάζοντας την τελευταία γραμμή βάζουμε το ΣΥΝΟΛΟ στη μεταβλητή sum_order και τη διαγράφουμε(prop)
10. Η κάθε γραμμή που έχει απομείνει περιέχει το προϊόν, την ποσότητα, την τιμή μονάδος και την συνολική τιμή. Κάθε μια τη βάζουμε ως είσοδο στη συνάρτηση parse_item όπου μας επιστρέφει ένα tuple με το ΑΦΜ, το όνομα του προϊόντος και το συνολικό ποσό. Επίσης, αν

υπάρχει κάποιο λάθος -πχ. το σύνολο είναι διαφορετικό(ποσότητα * τιμή μονάδος)- θα μας επιστρέψει σφάλμα έτσι ώστε να απορήσουμε τη συγκεκριμένη απόδειξη

11. Ελέγχουμε εαν το σύνολο της απόδειξης είναι το ίδιο με αυτό που υπολογίσαμε κρατώντας δυο δεκαδικά ψηφία
12. Τέλος, εφόσον δεν έχει υπάρξει κάποιο σφάλμα στην απόδειξη μπορούμε να την αποθηκεύσουμε

Αποθήκευση:

Αποφάσισα οτι η καλύτερη δομή δεδομένων είναι μια βάση δεδομένων sqlite, καθώς δε γνωρίζουμε ούτε το μέγεθος των δεδομένων, ούτε τους διαθέσιμους πόρους.

Οι εναλλακτικές που απορρίφτηκαν:

1. **Dictionary** Ενω αυτή η υλοποίηση θα ήταν η γρηγορότερη δυνατή θα χρησιμοποιούσαμε τουλάχιστον διπλάσια μνήμη απο το μέγεθος του αρχείου καθώς, θα είχαμε δυο dictionaries στη μνήμη, ενα με κλειδιά τα προϊόντα και ενα με κλειδιά τα ΑΦΜ
2. **Αρχείο** Εδω θα χρειάζεται να διαβάζουμε κάθε φάρα απο μεγάλα αρχεία χωρίς κάποια βελτιστοποίηση στις αναζητήσεις