

**Αρχές Γλωσσών Προγραμματισμού & Μεταφραστών****Τμήμα Μηχανικών Η/Υ & Πληροφορικής****Πανεπιστήμιο Πατρών****Εαρινό Εξάμηνο 2020****Διδάσκοντες: Ι. Γαροφαλάκης, Σ. Σιούτας****Προαιρετική Εργαστηριακή Άσκηση**

Το οικοσύστημα Python που περιλαμβάνει τη γλώσσα προγραμματισμού καθώς και μια σειρά από εξωτερικά πακέτα και βιβλιοθήκες λογισμικού είναι ένα ισχυρό εργαλείο ανοικτού κώδικα, το οποίο έχει μεγάλη χρησιμότητα εκτός από την αυστηρά οριοθετημένη περιοχή της επιστήμης των υπολογιστών, και στην έρευνα κάθε ειδικότητας αναφορικά με την ανάλυση και την επεξεργασία των δεδομένων.

Η γλώσσα Python είναι στην ουσία μια γλώσσα προγραμματισμού γενικού σκοπού, αντίθετα με άλλες γλώσσες οι οποίες επικεντρώνονται αυστηρά σε συγκεκριμένους τομείς όπως π.χ. η στατιστική ανάλυση. Το μεγάλο της πλεονέκτημα αφορά τη δυνατότητα της συγγραφής κώδικα στον οποίο μπορεί να γίνεται κάθε φορά χρήση των βιβλιοθηκών που είναι σχετικές με τις ζητούμενες επιστημονικές περιοχές του εκάστοτε project.

Ειδικότερα τώρα για τη συγκεκριμένη εργαστηριακή άσκηση θα χρησιμοποιηθούν δεδομένα σχετικά με τον τουρισμό στην Ελλάδα για την τετραετία 2011-2015. Τα συγκεκριμένα δεδομένα παρέχονται από την ιστοσελίδα της Ελληνικής Στατιστικής Αρχής ([www.statistics.gr](http://www.statistics.gr)) στη περιοχή Στατιστικές -> Βιομηχανία, Εμπόριο, Υπηρεσίες, Μεταφορές -> Τουρισμός σε μορφή αρχείων excel.

**Ζητούμενα :**

Αρχικά να υλοποιηθεί python script που θα κατεβάζει συνολικά τα απαραίτητα αρχεία excel από την ιστοσελίδα της ΕΛΣΤΑΤ. Στη συνέχεια να κατασκευαστεί πρόγραμμα σε python που θα επεξεργάζεται τα απαραίτητα δεδομένα από τα αρχεία excel για τα ζητούμενα της άσκησης (θα αναφερθούν στη συνέχεια στα σχετικά γραφήματα που θα εξαγάγετε κάνοντας χρήση της python) και θα τα φορτώσει σε αντίστοιχους πίνακες μιας βάσης SQLite ή MySQL. Στο τέλος τα συγκεκριμένα δεδομένα να εξαχθούν σε αντίστοιχα αρχεία .csv (που θα αποτελούν και αυτά παραδοτέα της άσκησης).

**Προσοχή όλες οι απαραίτητες ενέργειες θα πρέπει να γίνουν σε γλώσσα Python. Η βάση δεδομένων θα χρησιμοποιηθεί **ΜΟΝΟ** για αποθήκευση.**

Τέλος να εξαχθούν τα γραφήματα για τις παρακάτω περιπτώσεις:

- Συνολικές αφίξεις τουριστών στην Ελλάδα για την τετραετία 2011-2015
- Χώρες καταγωγής με το μεγαλύτερο μερίδιο στις αφίξεις τουριστών στην Ελλάδα για την τετραετία 2011-2015
- Αφίξεις τουριστών στην Ελλάδα ανά μέσο μεταφοράς για την τετραετία 2011-2015
- Αφίξεις τουριστών στην Ελλάδα ανά τρίμηνο για την τετραετία 2011-2015

**Υπόδειξη:** Χρησιμοποιείτε τις απαραίτητες βιβλιοθήκες της Python όπως pandas, xlrd, matplotlib ή όποιες άλλες κρίνετε εσείς απαραίτητο.

### Παραδοτέα

1. **Γραπτή Αναφορά** που περιλαμβάνει:
  - a. Τον κώδικα σε γλώσσα python εμπλουτισμένο με σχόλια
  - b. Screenshots παραδειγμάτων εφαρμογής
  - c. Τα ζητούμενα γραφήματα
2. Συμπιεσμένα σε **ένα αρχείο zip**:
  - Την πιο πάνω γραπτή αναφορά
  - Τον **ΤΕΛΙΚΟ** κώδικα σε python.
  - Τα αρχεία .csv που θα περιέχουν τα τελικά δεδομένα

***Το αρχείο zip πρέπει να έχει όνομα τον αριθμό μητρώου του φοιτητή (π.χ. 3972.zip), και να ανεβεί (ΥΠΟΧΡΕΩΤΙΚΑ) στο e-class. Σε ξεχωριστό αρχείο word μέσα στο zip να αναφέρεται το ονοματεπώνυμο, το έτος, ο αριθμός μητρώου και η e-mail διεύθυνση του φοιτητή.***

### Διευκρινήσεις

1. Η άσκηση είναι **ατομική**
2. Η άσκηση είναι **προαιρετική με bonus έως 1 μονάδα στον τελικό βαθμό, εφόσον ο βαθμός της είναι  $\geq 5$**
3. Τελικές ημερομηνίες παράδοσης είναι οι ημερομηνίες γραπτών εξετάσεων της τρέχουσας περιόδου **ΜΟΝΟ!**
4. Η άσκηση **ΔΕΝ ΘΑ ΚΡΑΤΗΘΕΙ** για το επόμενο ΑΚΑΔΗΜΑΙΚΟ ΕΤΟΣ!
5. Για τυχόν απορίες ή υποδείξεις μπορείτε να απευθύνεστε με e-mail στο [mvonitsanos@ceid.upatras.gr](mailto:mvonitsanos@ceid.upatras.gr) ή στο ανάλογο thread στις συζητήσεις στο e-class