



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
& ΠΛΗΡΟΦΟΡΙΚΗΣ

ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ
ΜΕΡΟΣ Α΄
ΓΙΑ ΤΟ ΜΑΘΗΜΑ
ΑΡΧΕΣ ΓΛΩΣΣΩΝ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ ΚΑΙ
ΜΕΤΑΦΡΑΣΤΩΝ

Προαιρετική Εργαστηριακή Άσκηση

ΜΑΝΤΑΣ ΕΛΕΥΘΕΡΙΟΣ

A.M. 1047128

Έτος Σπουδών: 5^ο

e-mail: up1047128@upnet.gr

ΠΑΤΡΑ 2020

Εισαγωγή

Η εργασία αυτή υλοποιήθηκε σε γλώσσα προγραμματισμού **Python**.

Χρησιμοποιήθηκαν οι παρακάτω βιβλιοθήκες:

- requests (<https://requests.readthedocs.io/en/master/>)
- Sqlite3 (<https://www.sqlite.org/index.html>)
- Pandas (<https://pandas.pydata.org/>)
- xlrd (<https://xlrd.readthedocs.io/en/latest/>)
- os (<https://docs.python.org/3/library/os.html>)
- unicode (<https://docs.python.org/3/howto/unicode.html>), για την μετάφραση της ελληνικής γλώσσας σε αγγλική καθώς δημιουργούσε πρόβλημα στη βάση δεδομένων.

Δεδομένα και κώδικας

Μαζί με την αναφορά επισυνάπτονται τα δεδομένα, στο φάκελο **data**, που εξάχθηκαν και κατέβηκαν με τη χρήση Python και της συνάρτησης request.

Η βάση δεδομένων είναι το αρχείο **database.db** που παραδίδεται μαζί με τα υπόλοιπα απαιτούμενα.

Το βασικό script που γράφτηκε **statistics.py** περιέχει τις παρακάτω συναρτήσεις:

- **download_excels()** η οποία είναι υπεύθυνη για το κατέβασμα των αρχείων από τη σελίδα <https://www.statistics.gr/> που αφορούν το τουρισμό στην Ελλάδα μεταξύ των ετών 2011-2015, καθώς και για την αποθήκευση όλων των αρχείων σε ξεχωριστά αρχεία δεδομένων.
- **create_connection()** η οποία δημιουργεί τη σύνδεση με την απαραίτητη βάση δεδομένων στην οποία θα αποθηκευτούν αργότερα όλα τα απαραίτητα δεδομένα για να απαντηθούν τα σχετικά ερωτήματα.
- **tables_creation()** η οποία δημιουργεί τους κατάλληλους πίνακες (tables) της βάσης δεδομένων στους οποίους αποθηκεύονται τα δεδομένα.

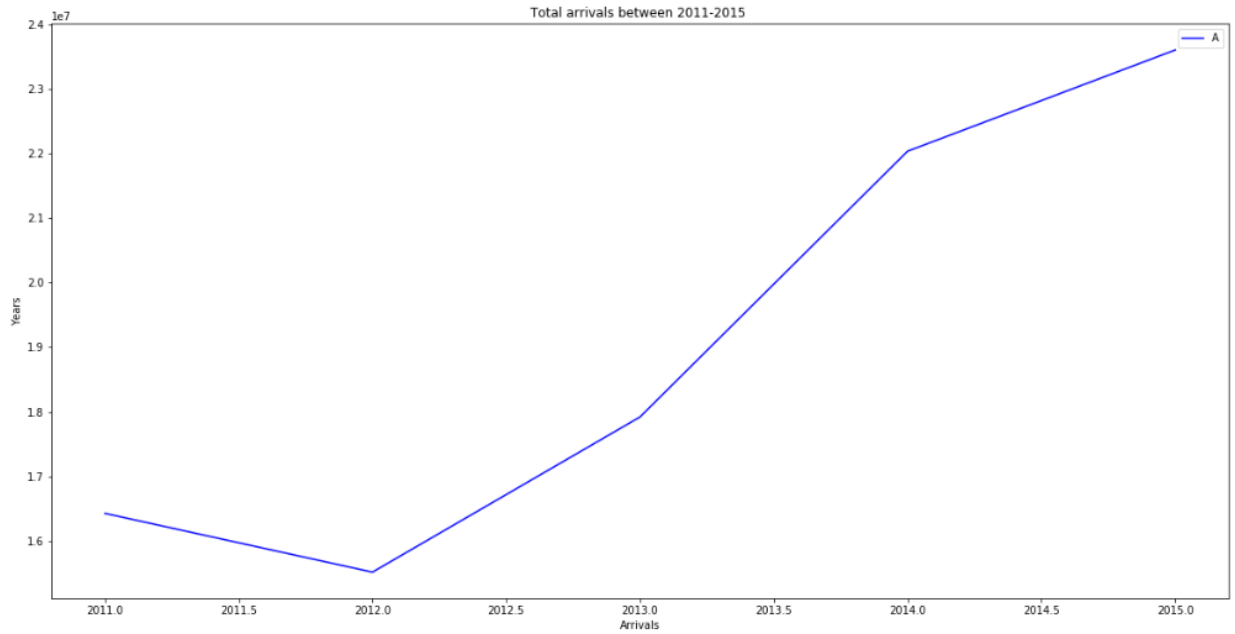
- **insert_data()** η οποία κάνει τη κατάλληλη εισαγωγή και ανανέωση των πινάκων παίρνοντας δεδομένα από τα αρχεία excel με τα στοιχεία περί τουρισμού.
- **queries()** η οποία εκτελεί τα ζητούμενα ερωτήματα στη βάση δεδομένων και κρατάει τα δεδομένα.
- **hasNumbers()** η οποία αναγνωρίζει εάν ένα string εισόδου περιέχει αριθμούς. Απαραίτητη για την επόμενη κατά σειρά συνάρτηση.
- **open_file()** η οποία δέχεται τη σύνδεση με τη βάση δεδομένων που έχει γίνει και ανοίγοντας το κατάλληλο φύλλο του κάθε αρχείου excel παίρνει τα απαραίτητα δεδομένα από αυτό.
- **plot_results()** και **plot2()** οι οποίες δημιουργούν τις κατάλληλες γραφικές παραστάσεις.
- **write_results_to_csv()** η οποία δημιουργεί τα ζητούμενα αρχεία csv και αποθηκεύει τα απαραίτητα δεδομένα σε αυτά.
- **main()** η οποία καλεί με κατάλληλη σειρά τις συναρτήσεις που αναπτύχθηκαν παραπάνω.

Αποτελέσματα

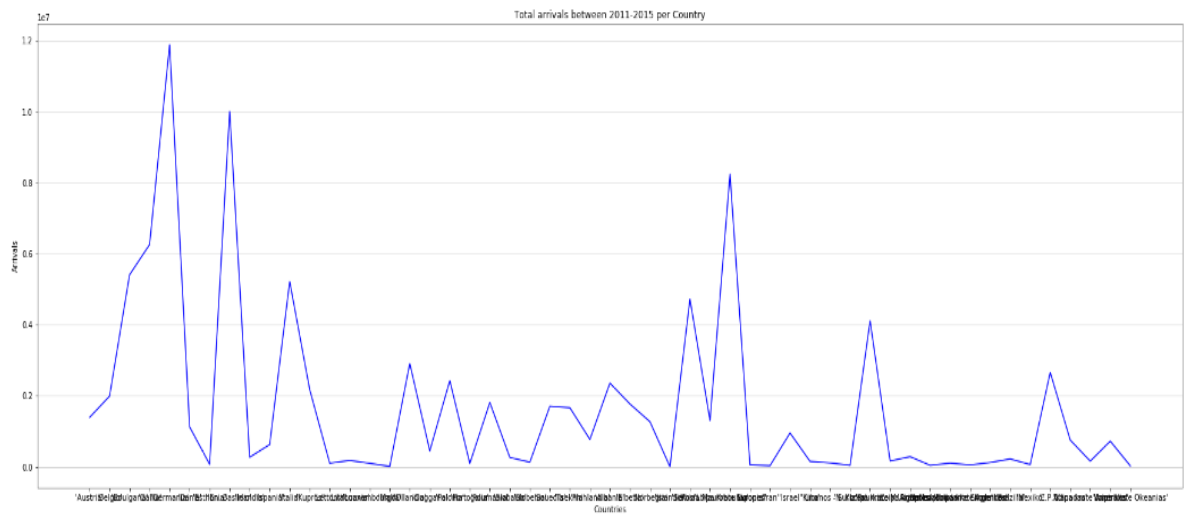
Τα αποτελέσματα των στατιστικών αναλύσεων υπάρχουν σε μορφή αρχείων .csv στο φάκελο **results**. Υπάρχουν 4 αρχεία ένα για κάθε ζητούμενο ερώτημα.

Επίσης, υπάρχει ο φάκελος **images** στον οποίο υπάρχουν τα ζητούμενα γραφήματα για τα ερωτήματα 1 έως 4. Τα γραφήματα φαίνονται και παρακάτω, όμως περισσότερη ευκρίνεια υπάρχει στα αντίστοιχα αρχεία που αναφέρθηκαν προηγουμένως.

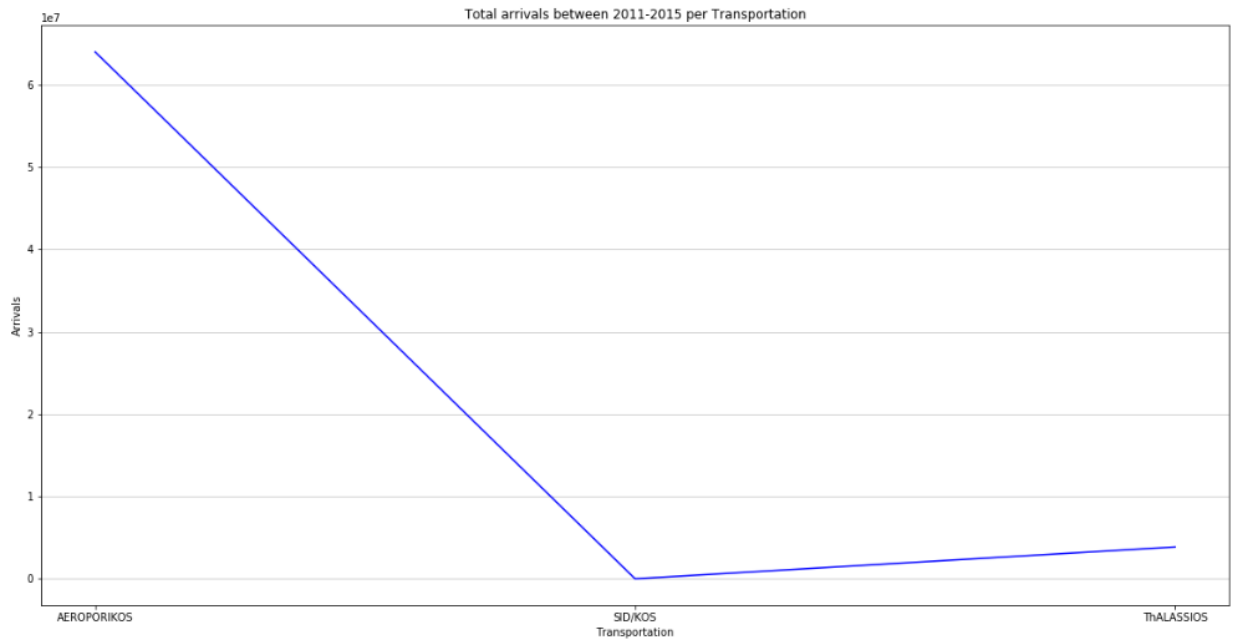
1. Συνολικές αφίξεις τουριστών στην Ελλάδα για την τετραετία 2011-2015



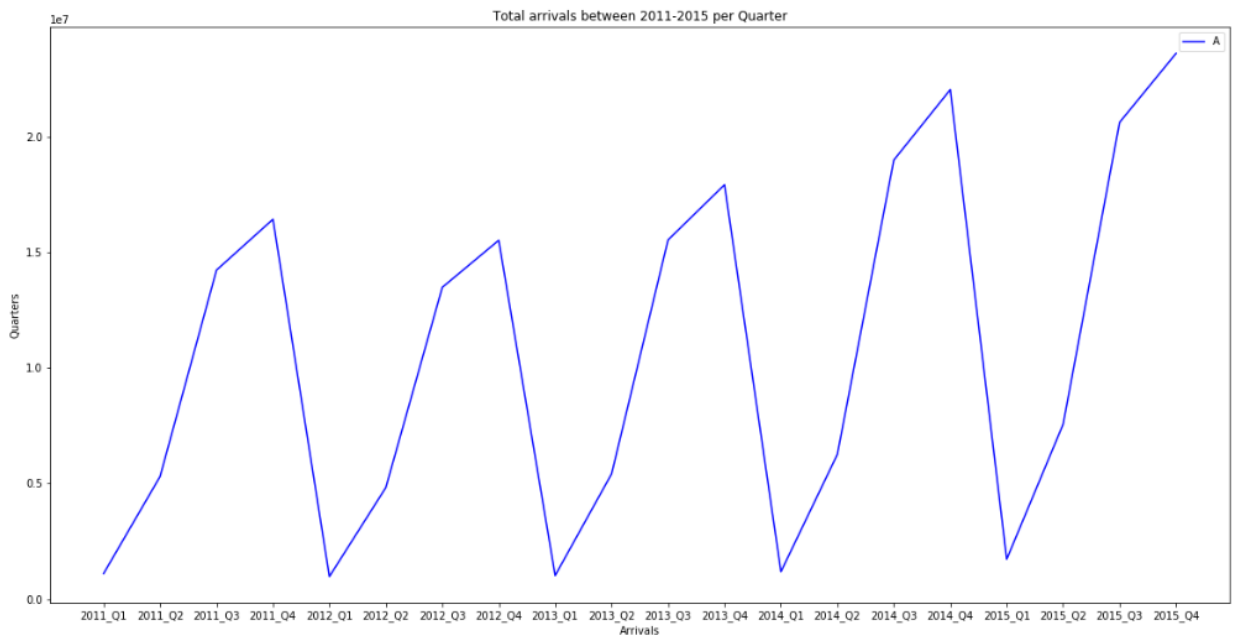
2. Χώρες καταγωγής με το μεγαλύτερο μερίδιο στις αφίξεις τουριστών στην Ελλάδα για την τετραετία 2011-2015.



3. Αφίξεις τουριστών στην Ελλάδα ανά μέσο μεταφοράς για την τετραετία 2011-2015.



4. Αφίξεις τουριστών στην Ελλάδα ανά τρίμηνο για την τετραετία 2011-2015.



Ο συνολικός κώδικας βρίσκεται στο αρχείο **stat_analysis.py** που έχει επισυναπτεί στο φάκελο zip και περιέχει αρκετά σχόλια ώστε να είναι κατανοητός.