

Πρότζεκτ ομάδα 3 Quatreto

Σας δίδεται το αρχείο seismoι.dat το οποίο είναι ένα αρχείο κειμένου και περιέχει τους σεισμούς που έχουν γίνει στον Ελλαδικό χώρο από το 550 π.Χ. έως και το 2010 μ.Χ.

Θα δημιουργήσετε ένα πρόγραμμα στην ργthon το οποίο θα διαβάζει αυτό το αρχείο και θα μας βρείτε κάποιες πληροφορίες.

Θα δουλέψετε για τις Περιφέρειες Αττικής και Θεσσαλίας.

Ο Επιστήμονας Δεδομένων (ΕΔ) θα βρει τις συντεταγμένες που περικλείουν την Περιφέρεια Πελοποννήσου (δείτε τα ανοιχτά δεδομένα εδώ: <https://geodata.gov.gr/en/dataset/periphereies-elladas>). Προσοχή: Κατεβάστε το φορμά WGS84. Το αρχείο αυτό έχει τις περιφέρειες μία ανά γραμμή. Για κάθε περιφέρεια υπάρχουν τρεις στήλες χωρισμένες με κόμμα. Αν το εισάγετε στο excel το αρχείο θα φαίνεται κάπως έτσι:

Column1	Column2	Column3
FID	PER	the_geom
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.0	Π. ΑΝΑΤΟΛΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ	MULTIPOLYGON (((23.623583209430933 41.35454572346803, 23.6236119264058 41
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.1	Π. ΚΕΝΤΡΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ	MULTIPOLYGON (((21.70422722305368 40.86338986707295, 21.704252695450897 4
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.2	Π. ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ	MULTIPOLYGON (((20.776478140435636 40.35294042426391, 20.776522072398766
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.3	Π. ΗΠΕΙΡΟΥ	MULTIPOLYGON (((20.008821405368835 39.69129652706123, 20.00883738123856 3
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.4	Π. ΘΕΣΣΑΛΙΑΣ	MULTIPOLYGON (((21.120422292894457 39.67887745931233, 21.120435241616182
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.5	Π. ΒΟΡΕΙΟΥ ΑΙΓΑΙΟΥ	MULTIPOLYGON (((24.97063986924611 39.50330066303183, 24.97071177733071 39
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.6	Π. ΝΟΤΙΟΥ ΑΙΓΑΙΟΥ	MULTIPOLYGON (((24.129093160805155 36.5077423830122, 24.129093280551235 3
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.7	Π. ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ	MULTIPOLYGON (((21.374638040833617 39.01889826370559, 21.374711159059725
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.8	Π. ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ	MULTIPOLYGON (((21.104826798957973 37.85062499628159, 21.105108433751695
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.9	Π. ΠΕΛΟΠΟΝΝΗΣΟΥ	MULTIPOLYGON (((21.5388617829156 37.05523051141978, 21.538862430117256 37
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.10	Π. ΙΟΝΙΩΝ ΝΗΣΩΝ	MULTIPOLYGON (((19.63426181326684 39.729383291004986, 19.63446091891052 3
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.11	Π. ΚΡΗΤΗΣ	MULTIPOLYGON (((24.036505739930025 34.8563644493727455, 24.03664491555135
d7f50467-e5ef-49ac-a7ce-15df3e2ed738.12	Π. ΑΤΤΙΚΗΣ	MULTIPOLYGON (((22.89876308843879 36.19484877951042, 22.8987819232714 36

Η στήλη multi-polygon περιέχει τα πολύγωνα σε συντεταγμένες συμβατές με το google maps. Μια περιφέρεια μπορεί να περιέχει περισσότερα από ένα πολύγωνα. Αυτά είναι στην τρίτη στήλη του excel στη μορφή:

MULTIPOLYGON(((συντεταγμένες 1^{ου} πολύγωνου)), ((συντεταγμένες 2^{ου} πολύγωνου)), ..., ((συντεταγμένες ν^{ου} πολύγωνου)))

Ο Επιστήμονας Δεδομένων (ΕΔ) θα διαμορφώσει κατάλληλα τα δεδομένα ώστε να μπορείτε να τα διαβάσετε από την Python (μπορεί να τα επεξεργαστεί με την Python ή με το excel). Εάν η εργασία με το παραπάνω αρχείο σας φαίνεται πολύ δύσκολη μπορείτε να πάρετε συντεταγμένες οι οποίες περικλείουν την περιοχή που σας ενδιαφέρει από το google maps με το χέρι.

Κάθε πολύγωνο ο επιστήμονας δεδομένων θα το παραδώσει σε μορφή αρχείου ώστε να μπορούμε να το ανεβάσουμε στο google maps και να ελέγξουμε τις συντεταγμένες της περιοχής. Αυτό το αρχείο μπορείτε και εσείς να το ανεβάσετε στο Google My Maps (<https://mymaps.google.com/>) για να δείτε αν έχει γίνει σωστά η δουλειά. Επειδή αυτά τα αρχεία θα είναι περισσότερα από ένα (π.χ. N), θα τα ονομάσετε region1.csv, region2.csv, ..., regionN.csv

Επίσης, ο επιστήμονας δεδομένων θα επεξεργαστεί το αρχείο `seismoι.dat` ώστε να μπορεί να το διαβάσει εύκολα η `rython`. Ακολούθως θα δημιουργήσει μια συνάρτηση (***openSeismoιFile***) που θα διαβάζει το αρχείο `seismoι.dat` και θα επιστρέφει μια λίστα της οποίας κάθε στοιχείο θα είναι μία λίστα με τα δεδομένα ενδιαφέροντος κάθε γραμμής του αρχείου.

Ακολούθως, οι προγραμματιστές θα δημιουργήσουν συνάρτηση (***findSeismoιPoints***) η οποία καλεί την προηγούμενη συνάρτηση, ανοίγει και τα αρχεία που περιέχουν τα πολύγωνα των περιοχών ενδιαφέροντος και θα επιστρέφει δύο διακριτά αρχεία, ένα για κάθε περιφέρεια, το καθένα εκ των οποίων θα περιέχει μόνο το κομμάτι της λίστας που επέστρεψε η *openSeismoιFile*, το οποίο αντιστοιχεί στην υπό διερεύνηση περιφέρεια. Υπόδειξη: Χρησιμοποιήστε το πακέτο `shapely` (`import shapely.geometry`), δείτε τα `Point`, `Polygon` και `MultiPolygon` οι θαρραλέοι. Δείτε ένα παράδειγμα:

```
>>> from shapely import Point, Polygon, contains
>>> area = Polygon([(0, 0), (1, 0), (1, 1), (0, 1), (0, 0)])
>>> contains(area, Point(0.5, 0.5))
True
```

Οι Προγραμματιστές θα εκτυπώνουν στην οθόνη τα στοιχεία (συντεταγμένες, ένταση, μήνας, έτος και όνομα Περιφέρειας στην οποία ανήκει ο σεισμός), με την σειρά που δίνονται, των σεισμών που έχουν πραγματοποιηθεί από το 1960 έως και το 2010. Ακολούθως θα αποθηκεύει τις συντεταγμένες, τον μήνα, το έτος και την ένταση σε δύο διακριτά αρχεία `output1.csv`, `output2.csv` ένα για κάθε περιφέρεια, της μορφής:

Latitude,Longitude,month,year,Magnitude
34.856364493727455,24.036505739930025,03,2001,4.2
34.856418557665,24.03664491555135,08,1995,5.5
34.85633504515916,24.036973394443212,12,2003,3.3
...

Ακολούθως ο αναλυτής θα δημιουργήσει την συνάρτηση ***visualizeData*** στην οποία θα χρησιμοποιήσει την `matplotlib` (δείτε ένα σχετικό [βίντεο](#) εδώ) για να ανοίξει καθένα από τα αρχεία που αποθήκευσαν οι Προγραμματιστές `output1.csv`, `output2.csv` και να δημιουργήσει ένα γράφημα που αναπαριστά το πλήθος των σεισμών που έχουν πραγματοποιηθεί ανά δεκαετία το χρονικό διάστημα από 1960 έως και 2010 για κάθε μία από τις περιοχές ενδιαφέροντος.

Παραδοτέα:

- Σύνδεσμος στο `google maps` όπου θα έχετε βάλει ένα επίπεδο με τις πληροφορίες του αρχείου `output.csv`
- Κώδικας `python`
- Αναφορά
- `output.csv`
- Αρχεία `region1.csv`, `region2.csv`, ..., `regionN.csv`