



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΥΠΡΟΥ

Τμήμα Πληροφορικής

ΕΠΛ 646 – Προχωρημένα Θέματα Βάσεων Δεδομένων

ΑΣΚΗΣΗ 4 – NoSQL Βάσεις Δεδομένων

(Δημιουργία Εφαρμογής για Επερωτήσεις σε Web 2.0 APIs)

Ημερομηνία Ανάθεσης: Τρίτη, 26/3/2024

Ημερομηνία Παράδοσης: Τρίτη, 2/4/2024 (7 ημέρες)

Διδάσκων: Δημήτρης Ζεϊναλιπούρ

(Να υποβληθούν όλα τα Design Documents μέσω του Moodle)

I. Στόχος Άσκησης

Ο στόχος της άσκησης είναι η δημιουργία μιας υποθετικής εφαρμογής στη βάση εγγράφων CouchDB και η εισαγωγή σε αυτή δεδομένων JSON τα οποία θα έχετε ανακτήσει από το Foursquare API (<https://developer.foursquare.com/>). Τελικός στόχος της άσκησης είναι η δημιουργία ανάλογων Design Views για την εξαγωγή των δεδομένων που θα ζητηθούν πιο κάτω από τη CouchDB μέσω curl (<http://curl.haxx.se/>) και η προσομοίωση του δικού σας Foursquare API με επαυξημένες δυνατότητες (enhanced Foursquare API).

Επίσης, μέρος της άσκησης έχει σαν στόχο τη δημιουργία Master-Master replicas της βάσης σας στην ίδια τοποθεσία με τον διαθέτη σας.

II. Προεργασία

Εγκαταστήστε την CouchDB και το curl αν δεν είναι ήδη εγκατεστημένα στο Virtual Machine (VM) που σας έχει δοθεί ή στον προσωπικό σας Η/Υ. Στη συνέχεια σιγουρευτείτε ότι εγκαταστήσατε επιτυχημένα την CouchDB ανοίγοντας το τερματικό στο VM και εκτελώντας την εντολή:

- `curl http://127.0.0.1:5984/`

Η απάντηση η οποία θα πάρετε θα πρέπει να είναι JSON μορφής και να μοιάζει:

```
{"couchdb": "Welcome", "version": "3.0.0", "git_sha": "03a77db6c", "uuid": "433bc90648ae4f3cf8ac714d110b255a", "features": ["access-ready", "partitioned", "pluggable-storage-engines", "reshard", "scheduler"], "vendor": {"name": "The Apache Software Foundation"}}
```

Σημείωση: Μην ξεχνάτε ότι αν δεν έχετε ρυθμίσει το CouchDB service να ξεκινάει με την εκκίνηση του συστήματός σας θα πρέπει να το κάνετε αλλιώς δεν θα μπορείτε να πάρετε καμία απάντηση. Για να το ξεκινήσετε χρησιμοποιείτε την εντολή: **service couchdb start** σαν root χρήστης.

III. Εξαγωγή δεδομένων από το Foursquare και εισαγωγή δεδομένων στη δική σας βάση

Χρησιμοποιώντας το curl και μέσω του Foursquare API (γενικές οδηγίες στον σύνδεσμο: <https://location.foursquare.com/developer/reference/places-api-overview> - λεπτομερείς οδηγίες για το curl API: <https://location.foursquare.com/developer/reference/search-data>) προσπαθήστε να εξάγετε όλες τις διαθέσιμες τοποθεσίες (venues) στην περιοχή σας μέσω του Foursquare (παράδειγμα δίνεται στο τέλος της εκφώνησης). Θα πρέπει να δημιουργήσετε λογαριασμό Foursquare (δείτε τις οδηγίες εδώ: <https://location.foursquare.com/developer/reference/places-api-get-started>) τον οποίο θα χρησιμοποιήσετε στην άσκηση αυτή.

Δημιουργήστε τη δική σας βάση στην CouchDB την οποία είχατε εγκαταστήσει στο προηγούμενο βήμα και εισάγετε σε αυτή τα δεδομένα τα οποία εξάγατε από το Foursquare API.

Σημείωση: Για να συμπληρώσετε σωστά τη βάση δεδομένων πρέπει να δημιουργήσετε ένα έγγραφο με τον πίνακα “venues” και να προσθέσετε "{" docs ":" στην αρχή του εγγράφου και "}" στο τέλος του.

IV. Δημιουργία Design Documents και MAP/REDUCE ερωτημάτων

Χρησιμοποιώντας τη βάση την οποία θα έχετε σε αυτό το στάδιο προσπαθήστε να δημιουργήσετε τα ανάλογα Design Documents με τα ανάλογα Views και τις ανάλογες Map/Reduce συναρτήσεις με την βοήθεια των οποίων θα προσπαθήσετε να δώσετε απαντήσεις στις πιο κάτω ερωτήσεις δημιουργώντας έτσι τη δική σας έκδοση του Foursquare API (enhanced Foursquare API).

Κατά τη διάρκεια αυτή της άσκησης θα πρέπει να δημιουργήσετε ένα Design Document ως ακολούθως:

- `_design_both`

Και τα οποία παράγουν τον πιο κάτω σύνδεσμο:

- http://127.0.0.1:5984/foursquare/_design/both

Στο πιο πάνω Design Document θα πρέπει να προσθέσετε τα Views (Q1, Q2, ..., Q12) έτσι ώστε να καταφέρετε να απαντήσετε τα παρακάτω ερωτήματα (queries). Οι επί μέρους όψεις πρέπει να είναι διαθέσιμες μέσω: http://127.0.0.1:5984/foursquare/_design/both/_view/QID ή με επιπλέον ορίσματα στο URL (π.χ., http://127.0.0.1:5984/foursquare/_design/both/_view/QID?group=true,ascending,descending,startkey,endkey,limit,κτλ.,βλέπεhttps://docs.couchdb.org/en/stable/api/index.html)

IV. MAP/REDUCE ερωτήματα

- Q1. Επιλέξτε όλα τα ονόματα (**name**) μαζί με το όνομα της πρώτης κατηγορίας (**categories**) των διαθέσιμων τοποθεσιών.
- Q2. Επιλέξτε όλα τα **name** των διαθέσιμων τοποθεσιών που ανήκουν στην κατηγορία **Bakery**.
- Q3. Επιλέξτε την πιο κοντινή τοποθεσία από **όλες** τις τοποθεσίες.
- Q4. Επιλέξτε την πιο μακρινή τοποθεσία από **όλες** τις τοποθεσίες.
- Q5. Επιλέξτε τις **δύο** τοποθεσίες με τα περισσότερα ratings.
- Q6. Επιλέξτε τις **δύο** τοποθεσίες με τις περισσότερες φωτογραφίες.
- Q7. Επιστρέψτε τον αριθμό των **τοποθεσιών που δεν έχουν φωτογραφία**.

- Q8. Επιστρέψετε τον αριθμό των **συνολικών φωτογραφιών** για όλες τις τοποθεσίες.
- Q9. Επιστρέψετε τον αριθμό των **συνολικών ratings** ανά κατηγορία.
- Q10. Επιστρέψετε τον αριθμό των **συνολικών τοποθεσιών** ανά κατηγορία.
- Q11. Επιστρέψετε τον αριθμό των **συνολικών τοποθεσιών που έχουν rating>5**.
- Q12. Επιλέξτε όλα τα **name** και **rating** των διαθέσιμων τοποθεσιών που δεν έχουν φωτογραφίες.

Πρέπει να χρησιμοποιήσετε τις εξής παραμέτρους:

- ll=35.144737,33.410966 (καθορίζει την περιοχή γύρω από την πανεπιστημιούπολη)
- limit=50
- fields= fsq_id,name,location,categories,chains,distance,rating,stats

Παραδοτέα

Παραδώσετε **όλα** τα πηγαία αρχεία σας μέσω του Moodle σε 1 συμπίεσμένο αρχείο (as4-myid.zip):

A) `download.txt`: καταγράψετε τις εντολές που χρησιμοποιήσατε για ανάκτηση και προσθήκη των δεδομένων από το Foursquare στην CouchDB.

B) `queries.txt`: καταγράψετε για κάθε επί μέρους Query τα πιο κάτω:

- Map/Reduce συνάρτηση
- URL για εκτέλεση του Query
- Απάντηση

Παράδειγμα διαλογικής επικοινωνίας με `developer.foursquare.com`

```
curl --request GET \
  --url
'https://api.foursquare.com/v3/places/search?ll=35.144737%2C33.410966&fields=
fsq_id%2Cname%2Clocation%2Ccategories%2Cchains%2Cdistance%2Crating%2Cstats&li
mit=2' \
  --header 'Authorization: <<PUT YOUR API KEY HERE>>' \
  --header 'accept: application/json'
```

Αποτέλεσμα (κατά τη χρονική στιγμή της ερώτησης **με όριο 2**):

```
{
  "results": [
    {
      "fsq_id": "4d10762f6331a093d6bc5694",
      "categories": [
        {
          "id": 16032,
          "name": "Park",
          "short_name": "Park",
          "plural_name": "Parks",
          "icon": {
            "prefix": "https://ss3.4sqi.net/img/categories_v2/parks_outdoors/park_",
            "suffix": ".png"
          }
        }
      ]
    },
    {
      "chains": [],
      "distance": 2897,
      "location": {
        "country": "CY",

```

```

    "cross_street": "",
    "formatted_address": "Λευκωσία",
    "locality": "Λευκωσία",
    "region": "Επαρχία Λευκωσίας"
  },
  "name": "Athalassa National Park (Εθνικό Πάρκο Αθαλάσσης)",
  "rating": 8.9,
  "stats": {
    "total_photos": 104,
    "total_ratings": 67,
    "total_tips": 15
  }
},
{
  "fsq_id": "54818471498e7c64503a88bd",
  "categories": [
    {
      "id": 13035,
      "name": "Coffee Shop",
      "short_name": "Coffee Shop",
      "plural_name": "Coffee Shops",
      "icon": {
        "prefix": "https://ss3.4sqi.net/img/categories_v2/food/coffeeshop_",
        "suffix": ".png"
      }
    }
  ],
  "chains": [],
  "distance": 4710,
  "location": {
    "address": "Pindarou 29d",
    "country": "CY",
    "cross_street": "",
    "formatted_address": "Pindarou 29d, 1060 Λευκωσία",
    "locality": "Λευκωσία",
    "postcode": "1060",
    "region": "Επαρχία Λευκωσίας"
  },
  "name": "a kxoffee project",
  "rating": 9.3,
  "stats": {
    "total_photos": 30,
    "total_ratings": 89,
    "total_tips": 12
  }
},
],
"context": {
  "geo_bounds": {
    "circle": {
      "center": {
        "latitude": 35.144737,
        "longitude": 33.410966
      },
      "radius": 22000
    }
  }
}
}
}

```

Σημείωση για Foursquare unverified account:

Rate Limiting

You may make 1,000 requests per 24 hour period.

Καλή Επιτυχία!