

ΕΘΝΙΚΟ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ



ΣΥΣΤΗΜΑΤΑ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ (M149)
Project 02

ΦΛΟΥΡΗΣ ΔΗΜΗΤΡΗΣ cs2180023
ΧΡΙΣΤΟΔΟΥΛΟΥ ΚΥΡΙΑΚΟΣ cs2180019

- Η εργασία έχει υλοποιηθεί πλήρως πανω σε Springboot framework. Για τα requests χρησιμοποιήσαμε το Postman.
- Παραδοτέα:
 - dbms2 (gradle project)
 - myFaker (gradle project)
 - ReadMe.pdf
- Σχεδίαση Βάσης Δεδομένων:
Έχουν υλοποιηθεί δύο collections για τη εργασία.
 - Incidents: δημιουργήσαμε ένα μοναδικό id και μετα ανάλογα με τα χαρακτηριστικά που είχε το κάθε incident έμπαιναν και στο JSON Object του collection. Επίσης υπάρχει ένα array με τα ids των citizens οι οποίοι έκαναν upvote στο συγκεκριμένο incident. Η συγκεκριμένη επιλογή έγινε ούτως ώστε να γίνει πιο άμεσα η υλοποίηση του query 7, καθώς βρίσκουμε αμέσως το πλήθος των votes για κάθε incident.
 - Citizens: Ένας citizen έχει τα εξής χαρακτηριστικά.
 - x _id
 - x name
 - x telephone
 - x address
 - x upVotes (=είναι ένας πίνακας απο JSON Objects που περιέχουν τα ids των incidents και το ward τους στα οποία έχει κάνει upvote ο συγκεκριμένος citizen).

Η συγκεκριμένη επιλογή έγινε για να διευκολυνθεί η υλοποίηση των queries 8-11.
- Εισαγωγή δεδομένων στη βάση:
Για την εισαγωγή των δεδομένων απο τα csv αρχεία χρησιμοποιήσαμε parsers υλοποιημένους σε Java.
Για τη δημιουργία και εισαγωγή των δεδομένων για το collection των citizens τρέξαμε το πρόγραμμα myfaker.java όπου δημιούργησε 2000 citizens και έκαναν 1000 upvotes ο καθένας σε διαφορετικά incidents. Στη συνέχεια προσθέσαμε και δικά μας upvotes σε τυχαία incidents.
- Υλοποίηση Εφαρμογής:
Έχουμε ένα UserController.java στον οποίο έχουμε μέσα το κατάλληλο mapping για τα queries για request απο το Postman.
Συνδεόμαστε με τη βάση μέσω τις εφαρμογής και δημιουργούμε το ανάλογο query με τη χρήση της εντολής aggregate. Φορτώνουμε τα aggregates μέσα στη λίστα pipeline και κτίζουμε την εντολή για τη mongodb.
Τα αποτελέσματα επιστρέφονταν στο Cursor όπου τα παίρναμε και τα φορτώνουμε σε μια λίστα για να τα στείλουμε πίσω στο postman (response).
- Εισαγωγή Incident:
Μέσα στο body του request περιέχεται ένα JSON Object απο το οποίο εξάγουμε την creation και completion date έτσι ώστε να τις μετατρέψουμε σε ISODATES και να τις κάνουμε replace στο JSON που πήραμε απο το request και τέλος το κάνουμε εισαγωγή στη βάση.
- Casting Upvotes
Με την ίδια διαδικασία όπως πιο πάνω παίρνουμε το JSON απο το οποίο παίρνουμε το όνομα του citizen και το id του incident για το οποίο θέλει να κάνει upvote. Στη συνέχεια βρίσκουμε τα συγκεκριμένα JSON απο τη βάση και ελέγχουμε τον πίνακα upVotes του citizen για να δούμε αν υπάρχει ήδη upvote για το συγκεκριμένο incident οπου και θα

τερματίσει αλλιώς προσθέτουμε στο πίνακα upVotes του συγκεκριμένου citizen το id και το ward του incident και στο πίνακα upvotes του incidents το id του citizen.

- Indexes

Για τα incidents έχουμε κάνει index στο creation date και στο type.
Ενώ στους citizens έχουμε στο name και στο telephone.

- Queries:

Query 01:

```
db.incidents.aggregate([{"$match": {"cr_date": {"$gt": {"$date": "2011-01-01T00:00:00.000Z"}, "$lt": {"$date": "2012-01-01T00:00:00.000Z"}}}}, {"$project": {"type": 1}}, {"$group": {"_id": "$type", "count": {"$sum": 1}}}, {"$sort": {"count": -1}}]);
```

```
{
  "results": [
    {
      "Graffiti Removal": 154974
    },
    {
      "Street Lights - All/Out": 80559
    },
    {
      "Pot Hole in Street": 80104
    },
    {
      "Garbage Cart Black Maintenance/Replacement": 49040
    },
    {
      "Alley Light Out": 46833
    },
    {
      "Tree Trim": 36146
    },
    {
      "Street Light - 1/Out": 35398
    },
    {
      "Rodent Baiting/Rat Complaint": 34321
    },
    {
      "Abandoned Vehicle Complaint": 23156
    },
    {
      "Tree Debris": 21132
    },
    {
      "Sanitation Code Violation": 17905
    },
    {
      "Street Light Out": 4
    }
  ]
}
```

Query 02:

```
db.incidents.aggregate([{"$match": {"type": "Abandoned Vehicle Complaint", "cr_date": {"$gt": {"$date": "2008-12-31T00:00:00.000Z"}, "$lt": {"$date": "2009-09-30T00:00:00.000Z"}}}}, {"$project": {"cr_date": 1}}, {"$group": {"_id": "$cr_date", "count": {"$sum": 1}}}]
```

```
{
  "results": [
    {
      "Tue Sep 22 03:00:00 EEST 2009": 1
    },
    {
      "Tue Aug 25 03:00:00 EEST 2009": 1
    },
    {
      "Tue Aug 11 03:00:00 EEST 2009": 2
    },
    {
      "Tue Jul 14 03:00:00 EEST 2009": 2
    },
    {
      "Tue Sep 01 03:00:00 EEST 2009": 1
    },
    {
      "Thu Mar 19 02:00:00 EET 2009": 1
    },
    {
      "Mon Jul 13 03:00:00 EEST 2009": 1
    },
    {
      "Wed Aug 19 03:00:00 EEST 2009": 1
    },
    {
      "Tue Feb 03 02:00:00 EET 2009": 1
    },
    {
      "Wed May 13 03:00:00 EEST 2009": 2
    }
  ]
}
```

Query 03:

```
db.incidents.aggregate([ { "$match" : { "cr_date" : { "$date" : "2011-01-01T00:00:00.000Z" } , "zip" : { "$ne" : null } } }, { "$group" : { "_id" : { "zip" : "$zip" , "type" : "$type" } , "count" : { "$sum" : 1 } } }, { "$sort" : { "count" : -1 } }, { "$group" : { "_id" : "$_id.zip" , "types" : { "$push" : { "Type" : "$_id.type" , "Count" : "$count" } } } }, { "$project" : { "_id" : 0 , "ZipCode" : "$_id" , "types" : { "$slice" : [ "$types.Type" , 3 ] } } }, { "$sort" : { "_id" : 1 } } ] );
```

```
{
  "results": [
    {
      "ZipCode": 60654,
      "types": [
        "Sanitation Code Violation"
      ]
    },
    {
      "ZipCode": 60659,
      "types": [
        "Sanitation Code Violation",
        "Alley Light Out"
      ]
    },
    {
      "ZipCode": 60611,
      "types": [
        "Pot Hole in Street"
      ]
    },
    {
      "ZipCode": 60625,
      "types": [
        "Graffiti Removal",
        "Street Light - 1/Out",
        "Garbage Cart Black Maintenance/Replacement"
      ]
    },
    {
      "ZipCode": 60637,
      "types": [
        "Street Lights - All/Out"
      ]
    },
    {
      "ZipCode": 60621,
      "types": [
        "Pot Hole in Street",
        "Street Lights - All/Out",
        "Rodent Baiting/Rat Complaint"
      ]
    },
    {
      "ZipCode": 60642
    }
  ]
}
```

Query 04:

```
db.incidents.aggregate([ { "$match" : { "type" : "Abandoned Vehicle Complaint" ,
"ward" : { "$exists" : "true" , "$ne" : null } } }, { "$project" : { "ward" : 1 } },
{ "$group" : { "_id" : "$ward" , "count" : { "$sum" : 1 } } }, { "$sort" : { "count" : 1 } },
{ "$limit" : 3 } ] );
```

```
{
  "results": [
    {
      "0": 168
    },
    {
      "43": 1355
    },
    {
      "44": 1536
    }
  ]
}
```

Query 05:

```
db.incidents.aggregate([ { "$match" : { "compl_date" : { "$ne" : null } , "cr_date" :
{ "$gte" : { "$date" : "2010-01-01T00:00:00.000Z" } , "$lte" : { "$date" : "2011-01-
01T00:00:00.000Z" } } } }, { "$group" : { "_id" : "$type" , "avg" : { "$avg" :
{ "$subtract" : [ "$compl_date" , "$cr_date" ] } } } } ] );
```

```
{
  "results": [
    {
      "Alley Light Out": 81.0817217716781
    },
    {
      "Street Lights - All/Out": 21.250605326876514
    },
    {
      "Tree Debris": 202
    },
    {
      "Garbage Cart Black Maintenance/Replacement": 99.90034448818898
    },
    {
      "Abandoned Vehicle Complaint": 92.73531989482909
    },
    {
      "Graffiti Removal": 79.67898832684826
    },
    {
      "Sanitation Code Violation": 25.034682080924853
    },
    {
      "Pot Hole in Street": 46.13963963963964
    },
    {
      "Rodent Baiting/Rat Complaint": 44.06565656565657
    },
    {
      "Tree Trim": 451.06056001736493
    },
    {
      "Street Light Out": 2888.6666666666665
    },
    {
      "Street Light - 1/Out": 64.03075291622481
    }
  ]
}
```

Query 06:

```
db.incidents.aggregate([{"$match": {"cr_date": {"$date": "2011-01-01T00:00:00.000Z"}}, {"lat": {"$gt": -90.0, "$lt": 90.0}, "lon": {"$gt": -91.0, "$lt": 91.0}}}, {"$project": {"type": 1}}, {"$group": {"_id": "$type", "count": {"$sum": 1}}}, {"$sort": {"count": -1}}, {"$limit": 1}]);
```

```
{
  "results": "Street Lights - All/Out"
}
```

Query 07:

```
db.incidents.aggregate([{"$match": {"cr_date": {"$date": "2011-01-01T00:00:00.000Z"}}}, {"$unwind": "$supvotes"}, {"$group": {"_id": "$_id", "count": {"$sum": 1}}}, {"$sort": {"count": -1}}, {"$limit": 50}]);
```

```
{
  "results": [
    {
      "5c4bc96c0fa391507cb64d30": 1
    },
    {
      "5c4bc7df0fa3914f444ef90d": 1
    },
    {
      "5c4bc8370fa3914fbe541f73": 1
    },
    {
      "5c4bc7df0fa3914f444ef97a": 1
    },
    {
      "5c4bc8370fa3914fbe541f6f": 1
    },
    {
      "5c4bc8c60fa3915003bc4251": 1
    },
    {
      "5c4bc7df0fa3914f444ef90f": 1
    },
    {
      "5c4bc8c60fa3915003bc4259": 1
    },
    {
      "5c4bc7df0fa3914f444ef909": 1
    },
    {
      "5c4bc8c60fa3915003bc4257": 1
    },
    {
      "5c4bc7df0fa3914f444ef97e": 1
    },
    {
      "5c4bc8370fa3914fbe541f65": 1
    },
    {
      "5c4bc8c60fa3915003bc4253": 1
    },
    {
      "5c4bc7df0fa3914f444ef978": 1
    }
  ]
}
```

Query 08:

```
db.citizens.aggregate([{"$unwind": "$supVotes"}, {"$group": {"_id": {"ID": "$_id", "Name": "$name"}, "count": {"$sum": 1}}}, {"$sort": {"count": -1}}, {"$limit": 50}]);
```

```
{
  "results": [
    {
      "Phillip Guzman": 1000
    },
    {
      "Glen Jordan": 1000
    },
    {
      "Heidi Velez": 1000
    },
    {
      "Carl Cash": 1000
    },
    {
      "Alex Walls": 1000
    },
    {
      "Teresa Drake": 1000
    },
    {
      "Ted Newman": 1000
    },
    {
      "Catherine Lamb": 1000
    },
    {
      "Bethany Tucker": 1000
    },
    {
      "Howard Bates": 1000
    },
    {
      "Bryan Herring": 1000
    },
    {
      "Tom McConnell": 1000
    },
    {
      "Bryan Frederick": 1000
    },
    {
      "Wendy Fowler": 1000
    },
    {
      "Shirley Noble": 1000
    }
  ]
}
```

Query 09:

```
db.citizens.aggregate([ { "$unwind" : "$SupVotes" }, { "$group" : { "_id" : "$_id" ,
"wards" : { "$addToSet" : "$SupVotes.uv_ward" } } }, { "$project" : { "size" : { "$size" :
{ "$ifNull" : [ "$wards" , null ] } } } }, { "$sort" : { "size" : -1 } }, { "$project" :
{ "size" : 0 } } ] );
```

```
{
  "results": [
    {
      "ID": "5c4bd0ff0fa391533cdbec8e"
    },
    {
      "ID": "5c4bd0ea0fa391533cdbec6a"
    },
    {
      "ID": "5c4bcfc80fa391533cdbea20"
    },
    {
      "ID": "5c4bcfdb0fa391533cdbea4c"
    },
    {
      "ID": "5c4bcfce0fa391533cdbea2d"
    },
    {
      "ID": "5c4bcfbd0fa391533cdbea03"
    },
    {
      "ID": "5c4bcef50fa391533cdbe797"
    },
    {
      "ID": "5c4bcfb30fa391533cdbe9ec"
    },
    {
      "ID": "5c4bceef0fa391533cdbe77a"
    },
    {
      "ID": "5c4bcfb30fa391533cdbe9eb"
    },
    {
      "ID": "5c4bcf100fa391533cdbe800"
    },
    {
      "ID": "5c4bcf0f0fa391533cdbe7fe"
    },
    {
      "ID": "5c4bcf0e0fa391533cdbe7fa"
    },
    {
      "ID": "5c4bd03b0fa391533cdbeb1f"
    },
    {
      "ID": "5c4bceee0fa391533cdbe778"
    }
  ]
}
```

Query 10:

```
db.citizens.aggregate([ { "$group" : { "_id" : "$telephone" , "incid" : { "$addToSet" :
"$SupVotes.uv_id" } , "count" : { "$sum" : 1 } } } , { "$match" : { "_id" : { "$ne" :
null } , "count" : { "$gt" : 1 } } } , { "$project" : { "incid" : 1 , "_id" : 0 } } ] );
```

```
1 {
2   "results": [
3     "5c4bbf91d5490951871ce704",
4     "5c4bbfa3d5490951871d2bff"
5   ]
6 }
```

Query 11:

```
db.citizens.aggregate([ { "$match" : { "name" : "Lindsey Craft" } } , { "$unwind" :
"$SupVotes" } , { "$group" : { "_id" : "$SupVotes.uv_ward" } } ] );
```



```
{
  "results": [
    {
      "_id": null
    },
    {
      "_id": 28
    },
    {
      "_id": 0
    },
    {
      "_id": 3
    },
    {
      "_id": 20
    },
    {
      "_id": 7
    },
    {
      "_id": 48
    },
    {
      "_id": 50
    },
    {
      "_id": 40
    },
    {
      "_id": 32
    },
    {
      "_id": 22
    },
    {
      "_id": 35
    },
    {
      "_id": 13
    },
    {
      "_id": 30
    }
  ]
}
```