GCD Week 6 – Spark

Kevin Jetten

i318843

Inhoud

[Assignment 1: Indexes 3](#_Toc496277733)

[**Exercises** 5](#_Toc496277734)

[Assignment 2: GeoNear 8](#_Toc496277735)

[Assignment 3: Mini-project 10](#_Toc496277736)

[1. Design a MongoDB scheme 10](#_Toc496277737)

[2. Develop and create the database. Create an ERD to make things clear. 10](#_Toc496277738)

[3. Insert a few documents and sub documents, make your database work. 11](#_Toc496277739)

[4. Show data using simple queries on the collections. Tip: use the .pretty() method. 11](#_Toc496277740)

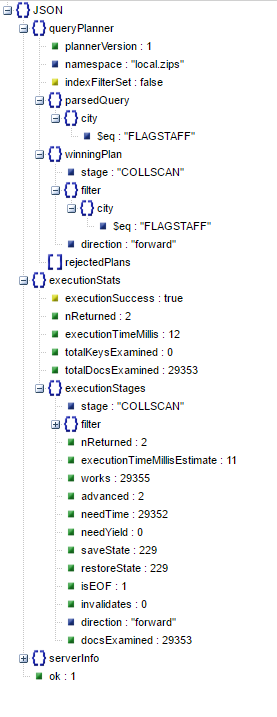
## Assignment 1: Indexes

1. Run find query on the data (without index):

a. db.zips.find().count()

b. db.zips.find({"city":"FLAGSTAFF"}).count()

c. db.zips.find({"city":"FLAGSTAFF"}).explain("executionStats")



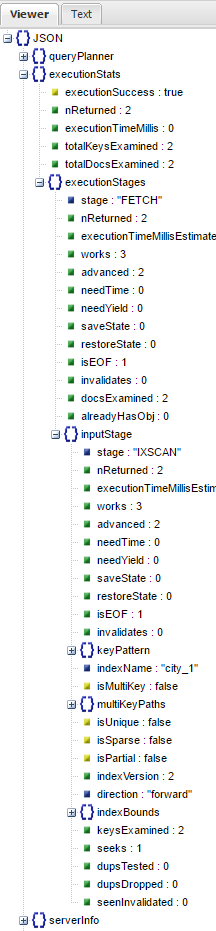
1. Run find query on the data (with index):

**db.zips.createIndex({"city" : 1})**

a. db.zips.find().count()

b. db.zips.find({"city":"FLAGSTAFF"}).count()

c. db.zips.find({"city":"FLAGSTAFF"}).explain("executionStats")

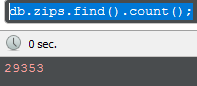


**2. Explain the difference.**

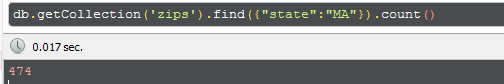
De execution time is van 12 naar 2 milliseconden gegaan. Dit is logisch aangezien er op het city-veld een index is geplaatst en vervolgens wordt er op het city-veld een query uitgevoerd.

## **Exercises**

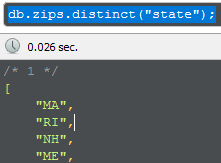
1. How many records does the zips collection consist of?



2. How many zipcodes are there in the state of Massachusetts?



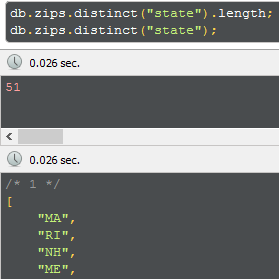
3. Give all state names from the zips collection.



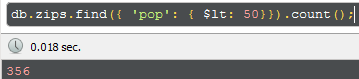
4. Sort the state names alphabetically.



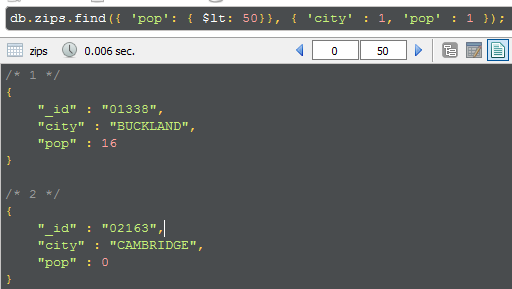
5. How many states are there in the zips collection?



6. How many cities have population of less than 50?



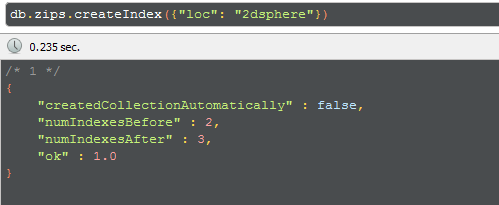
7. Which cities have population of less than 50?



## Assignment 2: GeoNear

1. Create a 2D index on location (loc field in de zips database).

**db.zips.createIndex({"loc": "2dsphere"})**

****

2. Find locations within 50 000 meters (50km) from Flagstaff

Query:

db.zips.find(

{ loc : {

$near : {

$geometry : {

type : "Point" ,

coordinates : db.zips.findOne({'city' : 'FLAGSTAFF'}, { '\_id' : 0, 'loc': 1 }).loc

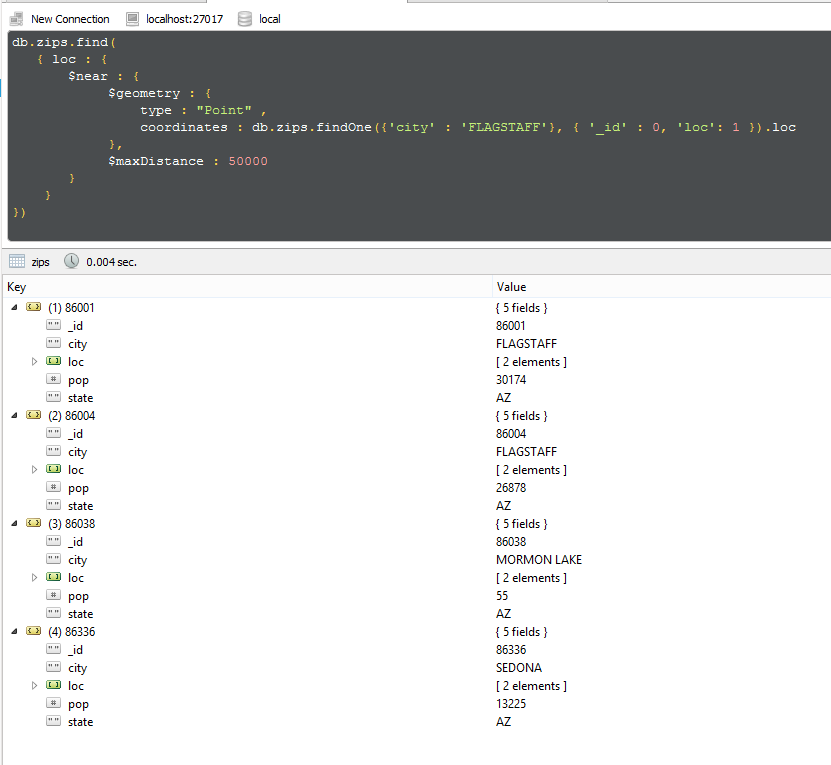
},

$maxDistance : 50000

}

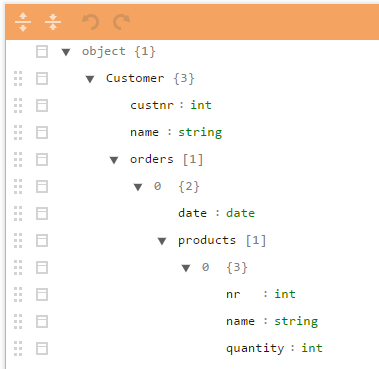
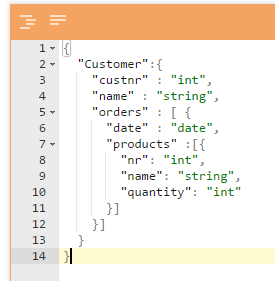
}

})

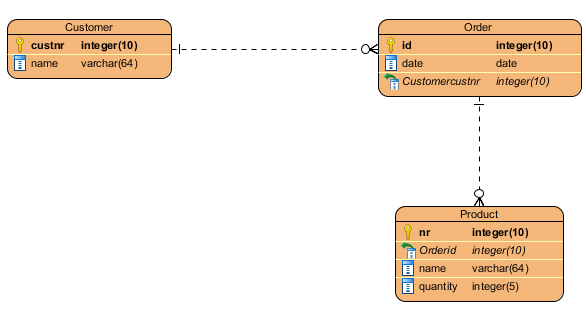


## Assignment 3: Mini-project

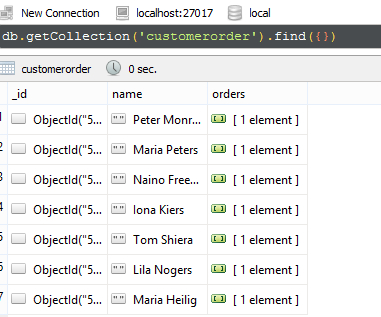
### 1. Design a MongoDB scheme



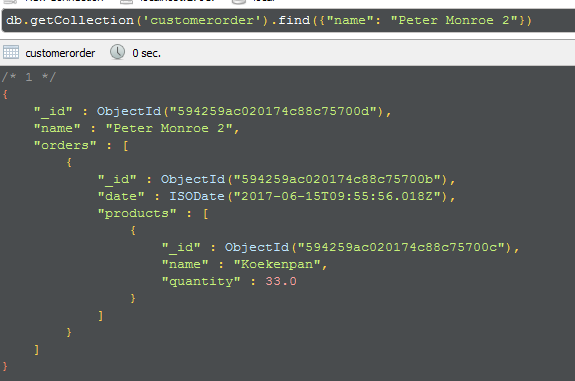
### 2. Develop and create the database. Create an ERD to make things clear.



## 3. Insert a few documents and sub documents, make your database work.



## 4. Show data using simple queries on the collections. Tip: use the .pretty() method.



*Ophalen van de persoon die Peter Monroe 2 heet.*

Het aanroepen van de .pretty() methode is niet nodig in Robomongo, zoals hierboven te zien is wordt de JSON-data al op een leesbare manier gepresenteerd.



*Ophalen van de personen die een jodenkoek in hun products heeft staan.*