

MongoDB -Aggregation Exercises

Import the zips.json file into your MongoDB. Database name is "population" and collection name is "zipcodes". `mongoimport --db population --collection zipcodes --file zips.json`

Atlanta Population

1. use `db.zipcodes.find()` to filter results to only the results where city is ATLANTA and state is GA.

```
db.zipcodes.find({$text:{$search:'ATLANTA GA'}})
```

2. use `db.zipcodes.aggregate` with `$match` to do the same as above.

```
db.zipcodes.aggregate([{$match:{$and:[{city:'ATLANTA'},{state:'GA'}}]})
```

3. use `$group` to count the number of zip codes in Atlanta.

```
db.zipcodes.aggregate([{$group:({_id:{state:'$state',city:'$ATLANTA'},pop:{$sum: '$pop'}})])
```

4. use `$group` to find the total population in Atlanta.

```
db.zipcodes.aggregate([{$group: {_id:{city:'$ATLANTA'},pop:{$sum:'$pop'}}]})
```

Populations By State

1. use `aggregate` to calculate the total population for each state

```
db.zipcodes.aggregate([{$group: {_id:{state:'$state'},pop:{$sum:'$pop'}}]})
```

2. sort the results by population, highest first

```
db.zipcodes.aggregate([{$sort:{pop:-1}}])
```

3. limit the results to just the first 3 results. What are the top 3 states in population?

```
db.zipcodes.aggregate([{$sort:{pop:-1}},{$limit:3}])
```

Populations by City

1. use aggregate to calculate the total population for each city (you have to use city/state combination). You can use a combination for the `_id` of the `$group`: { city: '\$city', state: '\$state' }

```
db.zipcodes.aggregate([{$group:{_id:{city:'$city'},pop:{sum:'$pop'}}}])
```

2. sort the results by population, highest first

```
db.zipcodes.aggregate([{$sort:{pop:-1}}])
```

3. limit the results to just the first 3 results. What are the top 3 cities in population?

```
db.zipcodes.aggregate([{$sort:{pop:-1}},{$limit:3}])
```

4. What are the top 3 cities in population in Texas?

```
db.zipcodes.aggregate([{$group:{_id:{city:'$city'},pop:{sum:'$pop'}}}])
```

Bonus

1. Write a query to get the average city population for each state.

```
db.zipcodes.aggregate([
  { $group: { _id: { state: "$state", city: "$city" }, pop: { $sum: "$pop" } } },
  { $group: { _id: "$_id.state", avgCityPop: { $avg: "$pop" } } }
])
```

2. What are the top 3 states in terms of average city population?

```
db.zipcodes.aggregate([{$sort:{avg:-1}},{$limit:3}])
```