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

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 ( { $and : [ { "city" : "ATLANTA" } , { "state" : "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 ( {
           $match : { "city" : "ATLANTA" } },
           { $group : { _id : { city : "$city" },count : { $sum : 1 } }
   } )
   ```

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

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

Populations By State

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

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

2. sort the results by population, highest first

```
db.zipcodes.aggregate( [
    { $group : { _id : { state : "$state" }, totalPop : { $sum : "$pop" } } },
    { $sort : { "totalPop" : -1 } }
] )
```

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

```
db.zipcodes.aggregate( [
    { $group : { _id : { state : "$state" }, totalPop : { $sum : "$pop" } } },
    { $sort : { "totalPop" : -1 } },
    { $limit : 3 }
] )
```

Top 3 cities in population are :

1. CA
2. NY
3. TX

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 ( [
        { $group : { _id : { state : "$state",city : "$city" },
        totalPop : { $sum : "$pop" } } }
] )
```

2. sort the results by population, highest first

```
db.zipcodes.aggregate ( [
        { $group : { _id : { state : "$state",city : "$city" },
        totalPop : { $sum : "$pop" } } },
        { $sort : { totalPop : -1 } }
] )
```

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

```
db.zipcodes.aggregate ( [
        { $group : { _id : { state : "$state",city : "$city" },
        totalPop : { $sum : "$pop" } } },
        { $sort : { totalPop : -1 } },
        { $limit : 3 }
] )
```

    Top 3 cities in Population are :
        1. CHICAGO
        2. BROOKLYN
        3. LOS ANGELES

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

```
db.zipcodes.aggregate ( [
        { $ match : { state : "TX" } } ,
        {$group:{_id:{ state : "$state" ,city : "$city" } , totalPop : { $sum : "$pop" } } },
        { $ sort : { totalPop : -1 } } ,
        { $limit : 3 }
] )
```

    Top 3 cities in population in Texas :
        1. HOUSTON
        2. DALLAS
        3. SAN ANTONIO

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( [
   { $group: { _id: { state: "$state", city: "$city" }, pop: { $sum: "$pop" } } },
   { $group: { _id: "$_id.state", avgCityPop: { $avg: "$pop" } } },
   { $sort : { "avgCityPop" : -1 } } ,
   { $limit : 3 }
   ] )
   ```

   Top 3 states in terms of avg city population :

   1. DC
   2. CA
   3. FL