**MongoDB -Aggregation Exercises**

**Name:** Sumit Hanmant Gaikwad

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.**

Ans: db.zipcodes.find({$and:[{city:"ATLANTA"},{state:"GA"}]})

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

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

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

Ans: db.zipcodes.aggregate([ { $match: { city: "ATLANTA" } }, { $group: { \_id: "$city", count: { $sum: 1 } } }])

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

Ans: db.zipcodes.aggregate([{$match: {city:"ATLANTA"}},{$group:{\_id:"$city",total\_population:{$sum:"$pop"}}}])

**Populations By State:**

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

Ans: db.zipcodes.aggregate([{$group:{\_id:"$state",totalPopulation:{$sum:"$pop"}}}])

**2. sort the results by population, highest first**

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

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

**population?**

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

[

{ \_id: 'CA', totalPopulation: 29754890 },

{ \_id: 'NY', totalPopulation: 17990402 },

{ \_id: 'TX', totalPopulation: 16984601 }

]

**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' }**

Ans: db.zipcodes.aggregate([{$group:{\_id:{city:"$city",state:"$state"},totalPopulation:{$sum:"$pop"}}}])

**2. sort the results by population, highest first**

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

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

**population?**

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

[

{ \_id: { city: 'CHICAGO', state: 'IL' }, totalPopulation: 2452177 },

{ \_id: { city: 'BROOKLYN', state: 'NY' }, totalPopulation: 2300504 },

{ \_id: { city: 'LOS ANGELES', state: 'CA' },totalPopulation: 2102295 }

]

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

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

[

{ \_id: { city: 'HOUSTON', state: 'TX' }, totalPopulation: 2095918 },

{ \_id: { city: 'DALLAS', state: 'TX' }, totalPopulation: 940191 },

{ \_id: { city: 'SAN ANTONIO', state: 'TX' }, totalPopulation: 811792 }

]

**Bonus**

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

Ans: db.zipcodes.aggregate([{$group:{\_id:"$state",avgPopulation:{$avg:"$pop"}}}])

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

Ans: db.zipcodes.aggregate([{$group:{\_id:"$state",avgPopulation:{$avg:"$pop"}}},{$sort:{avgPopulation:-1}},{$limit:3}])

[

{ \_id: 'DC', avgPopulation: 25287.5 },

{ \_id: 'CA', avgPopulation: 19627.236147757256 },

{ \_id: 'FL', avgPopulation: 15779.407960199005 }

]