# 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({$and:[{city:'ATLANTA'},{state:'GA'}]})

[

{

\_id: '30303',

city: 'ATLANTA',

loc: [ -84.388846, 33.752504 ],

pop: 1845,

state: 'GA'

},

{

\_id: '30305',

city: 'ATLANTA',

loc: [ -84.385145, 33.831963 ],

pop: 19122,

state: 'GA'

},

{

\_id: '30306',

city: 'ATLANTA',

loc: [ -84.351418, 33.786027 ],

pop: 20081,

state: 'GA'

},

{

\_id: '30307',

city: 'ATLANTA',

loc: [ -84.335957, 33.769138 ],

pop: 16330,

state: 'GA'

},

{

\_id: '30308',

city: 'ATLANTA',

loc: [ -84.375744, 33.771839 ],

pop: 8549,

state: 'GA'

},

{

\_id: '30309',

city: 'ATLANTA',

loc: [ -84.388338, 33.798407 ],

pop: 14766,

state: 'GA'

},

{

\_id: '30310',

city: 'ATLANTA',

loc: [ -84.423173, 33.727849 ],

pop: 34017,

state: 'GA'

},

{

\_id: '30311',

city: 'ATLANTA',

loc: [ -84.470219, 33.722957 ],

pop: 34880,

state: 'GA'

},

{

\_id: '30312',

city: 'ATLANTA',

loc: [ -84.378125, 33.746749 ],

pop: 17683,

state: 'GA'

},

{

\_id: '30313',

city: 'ATLANTA',

loc: [ -84.39352, 33.76825 ],

pop: 8038,

state: 'GA'

},

{

\_id: '30314',

city: 'ATLANTA',

loc: [ -84.425546, 33.756103 ],

pop: 26649,

state: 'GA'

},

{

\_id: '30315',

city: 'ATLANTA',

loc: [ -84.380771, 33.705062 ],

pop: 41061,

state: 'GA'

},

{

\_id: '30316',

city: 'ATLANTA',

loc: [ -84.333913, 33.721686 ],

pop: 34668,

state: 'GA'

},

{

\_id: '30317',

city: 'ATLANTA',

loc: [ -84.31685, 33.749788 ],

pop: 16395,

state: 'GA'

},

{

\_id: '30318',

city: 'ATLANTA',

loc: [ -84.445432, 33.786454 ],

pop: 53894,

state: 'GA'

},

{

\_id: '30319',

city: 'ATLANTA',

loc: [ -84.335091, 33.868728 ],

pop: 32138,

state: 'GA'

},

{

\_id: '30324',

city: 'ATLANTA',

loc: [ -84.354867, 33.820609 ],

pop: 15044,

state: 'GA'

},

{

\_id: '30326',

city: 'ATLANTA',

loc: [ -84.358232, 33.848168 ],

pop: 125,

state: 'GA'

},

{

\_id: '30327',

city: 'ATLANTA',

loc: [ -84.419966, 33.862723 ],

pop: 18467,

state: 'GA'

},

{

\_id: '30329',

city: 'ATLANTA',

loc: [ -84.321402, 33.823555 ],

pop: 17013,

state: 'GA'

}

]

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

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

[

{

\_id: '30303',

city: 'ATLANTA',

loc: [ -84.388846, 33.752504 ],

pop: 1845,

state: 'GA'

},

{

\_id: '30305',

city: 'ATLANTA',

loc: [ -84.385145, 33.831963 ],

pop: 19122,

state: 'GA'

},

{

\_id: '30306',

city: 'ATLANTA',

loc: [ -84.351418, 33.786027 ],

pop: 20081,

state: 'GA'

},

{

\_id: '30307',

city: 'ATLANTA',

loc: [ -84.335957, 33.769138 ],

pop: 16330,

state: 'GA'

},

{

\_id: '30308',

city: 'ATLANTA',

loc: [ -84.375744, 33.771839 ],

pop: 8549,

state: 'GA'

},

{

\_id: '30309',

city: 'ATLANTA',

loc: [ -84.388338, 33.798407 ],

pop: 14766,

state: 'GA'

},

{

\_id: '30310',

city: 'ATLANTA',

loc: [ -84.423173, 33.727849 ],

pop: 34017,

state: 'GA'

},

{

\_id: '30311',

city: 'ATLANTA',

loc: [ -84.470219, 33.722957 ],

pop: 34880,

state: 'GA'

},

{

\_id: '30312',

city: 'ATLANTA',

loc: [ -84.378125, 33.746749 ],

pop: 17683,

state: 'GA'

},

{

\_id: '30313',

city: 'ATLANTA',

loc: [ -84.39352, 33.76825 ],

pop: 8038,

state: 'GA'

},

{

\_id: '30314',

city: 'ATLANTA',

loc: [ -84.425546, 33.756103 ],

pop: 26649,

state: 'GA'

},

{

\_id: '30315',

city: 'ATLANTA',

loc: [ -84.380771, 33.705062 ],

pop: 41061,

state: 'GA'

},

{

\_id: '30316',

city: 'ATLANTA',

loc: [ -84.333913, 33.721686 ],

pop: 34668,

state: 'GA'

},

{

\_id: '30317',

city: 'ATLANTA',

loc: [ -84.31685, 33.749788 ],

pop: 16395,

state: 'GA'

},

{

\_id: '30318',

city: 'ATLANTA',

loc: [ -84.445432, 33.786454 ],

pop: 53894,

state: 'GA'

},

{

\_id: '30319',

city: 'ATLANTA',

loc: [ -84.335091, 33.868728 ],

pop: 32138,

state: 'GA'

},

{

\_id: '30324',

city: 'ATLANTA',

loc: [ -84.354867, 33.820609 ],

pop: 15044,

state: 'GA'

},

{

\_id: '30326',

city: 'ATLANTA',

loc: [ -84.358232, 33.848168 ],

pop: 125,

state: 'GA'

},

{

\_id: '30327',

city: 'ATLANTA',

loc: [ -84.419966, 33.862723 ],

pop: 18467,

state: 'GA'

},

{

\_id: '30329',

city: 'ATLANTA',

loc: [ -84.321402, 33.823555 ],

pop: 17013,

state: 'GA'

}

]

1. use $group to count the number of zip codes in Atlanta.
2. use $group to find the total population in Atlanta.

# Populations By State

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

Atlas atlas-pac47u-shard-0 [primary] population>

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

1. sort the results by population, highest first

Atlas atlas-pac47u-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{\_id:"state",totalpop:{$sum:"$pop"}}},

{$sort:{totalpop:-1}}])

[

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

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

1. sort the results by population, highest first

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

1. 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"},pop:{$sum:"$pop"}}},{$sort:{pop:-1}},{$limit:3}])
2. What are the top 3 cities in population in Texas?

# 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"}}}])

[

1. 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}])

[

{ \_id: 'DC', avgCityPop: 303450 },

{ \_id: 'CA', avgCityPop: 27756.42723880597 },

{ \_id: 'FL', avgCityPop: 27400.958963282937 }

]