# Name: Lutika Kolhe

# 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

Ans- C:\Users\hp\Lutika\zip.json

# Atlanta Population

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

Atlas atlas-ojy0k0-shard-0 [primary] Population> 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.

Atlas atlas-ojy0k0-shard-0 [primary] Population> 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.

Atlas atlas-ojy0k0-shard-0 [primary] Population>

[ { ATLANTA: 16584 } ]

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

Atlas atlas-ojy0k0-shard-0 [primary] Population> db.zipcodes.aggregate([{$group:{\_id:'$city', totalpop: {$sum:'$pop'}}},Atlas atlas-ojy0k0-shard-0 [primary] Population>

[ { \_id: 'ATLANTA', totalpop: 630046 } ]

# Populations By State

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

Atlas atlas-ojy0k0-shard-0 [primary] Population> db.zipcodes.aggregate([{$group:{\_id:"$state", totalpop:{ $sum: "$pop"}}}]) \_id: 'ATLANTA', totalpop: 630046 } ]

[

{ \_id: 'CT', totalpop: 3287116 },

{ \_id: 'MD', totalpop: 4781379 },

{ \_id: 'DE', totalpop: 666168 },

{ \_id: 'MA', totalpop: 6016425 },

{ \_id: 'NE', totalpop: 1578139 },

{ \_id: 'TN', totalpop: 4876457 },

{ \_id: 'MT', totalpop: 798948 },

{ \_id: 'LA', totalpop: 4217595 },

{ \_id: 'AR', totalpop: 2350725 },

{ \_id: 'ID', totalpop: 1006749 },

{ \_id: 'UT', totalpop: 1722850 },

{ \_id: 'GA', totalpop: 6478216 },

{ \_id: 'AL', totalpop: 4040587 },

{ \_id: 'WI', totalpop: 4891769 },

{ \_id: 'CO', totalpop: 3293755 },

{ \_id: 'NM', totalpop: 1515069 },

{ \_id: 'HI', totalpop: 1108229 },

{ \_id: 'WA', totalpop: 4866692 },

{ \_id: 'AK', totalpop: 544698 },

{ \_id: 'AZ', totalpop: 3665228 }

]

Type "it" for more

1. sort the results by population, highest first

Atlas atlas-ojy0k0-shard-0 [primary] Population> db.zipcodes.aggregate([{$group:{\_id:"$state", totalpop:{ $sum: "$pop"}}},{$sort: {totalpop:-1}}])

[

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

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

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

{ \_id: 'FL', totalpop: 12686644 },

{ \_id: 'PA', totalpop: 11881643 },

{ \_id: 'IL', totalpop: 11427576 },

{ \_id: 'OH', totalpop: 10846517 },

{ \_id: 'MI', totalpop: 9295297 },

{ \_id: 'NJ', totalpop: 7730188 },

{ \_id: 'NC', totalpop: 6628637 },

{ \_id: 'GA', totalpop: 6478216 },

{ \_id: 'VA', totalpop: 6181479 },

{ \_id: 'MA', totalpop: 6016425 },

{ \_id: 'IN', totalpop: 5544136 },

{ \_id: 'MO', totalpop: 5110648 },

{ \_id: 'WI', totalpop: 4891769 },

{ \_id: 'TN', totalpop: 4876457 },

{ \_id: 'WA', totalpop: 4866692 },

{ \_id: 'MD', totalpop: 4781379 },

{ \_id: 'MN', totalpop: 4372982 }

]

Type "it" for more

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

Atlas atlas-ojy0k0-shard-0 [primary] Population> db.zipcodes.aggregate([{$group:{\_id:"$state", totalpop:{ $sum: "$pop"}}},{$sort: {totalpop:-1}},{$limit:3}])

[

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

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

{ \_id: 'TX', totalpop: 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' }

Atlas atlas-ojy0k0-shard-0 [primary] Population> db.zipcodes.aggregate([{$group: {\_id:{state:"$state",city:"$city"},pop:{$sum:"$pop"}}}])

[

{ \_id: { state: 'OR', city: 'FALL CREEK' }, pop: 1581 },

{ \_id: { state: 'ME', city: 'EASTPORT' }, pop: 2514 },

{ \_id: { state: 'WI', city: 'OMRO' }, pop: 5678 },

{ \_id: { state: 'KS', city: 'BURRTON' }, pop: 1149 },

{ \_id: { state: 'PA', city: 'PEQUEA' }, pop: 1800 },

{ \_id: { state: 'NC', city: 'MOUNT ULLA' }, pop: 683 },

{ \_id: { state: 'KS', city: 'PRINCETON' }, pop: 650 },

{ \_id: { state: 'OR', city: 'SCOTTS MILLS' }, pop: 1326 },

{ \_id: { state: 'IL', city: 'MERRIONETTE PARK' }, pop: 29847 },

{ \_id: { state: 'SC', city: 'PENDLETON' }, pop: 12288 },

{ \_id: { state: 'CO', city: 'BELLVUE' }, pop: 1982 },

{ \_id: { state: 'MN', city: 'FEDERAL DAM' }, pop: 347 },

{ \_id: { state: 'NY', city: 'LEWISTON' }, pop: 10540 },

{ \_id: { state: 'MA', city: 'WINTHROP' }, pop: 18907 },

{ \_id: { state: 'WI', city: 'WILLIAMS BAY' }, pop: 2208 },

{ \_id: { state: 'VT', city: 'COLCHESTER' }, pop: 14731 },

{ \_id: { state: 'NH', city: 'MEREDITH' }, pop: 5959 },

{ \_id: { state: 'GA', city: 'ODUM' }, pop: 2851 },

{ \_id: { state: 'KS', city: 'RUSSELL SPRINGS' }, pop: 243 },

{ \_id: { state: 'OK', city: 'CARNEGIE' }, pop: 3907 }

]

Type "it" for more

1. sort the results by population, highest first

Atlas atlas-ojy0k0-shard-0 [primary] Population> db.zipcodes.aggregate([{ $group: { \_id: { state: "$state", city: "$city" }, pop: { $sum: "$pop" } } }, { $sort:{pop:-1} }])

[

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

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

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

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

{ \_id: { state: 'PA', city: 'PHILADELPHIA' }, pop: 1610956 },

{ \_id: { state: 'NY', city: 'NEW YORK' }, pop: 1476790 },

{ \_id: { state: 'NY', city: 'BRONX' }, pop: 1209548 },

{ \_id: { state: 'CA', city: 'SAN DIEGO' }, pop: 1049298 },

{ \_id: { state: 'MI', city: 'DETROIT' }, pop: 963243 },

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

{ \_id: { state: 'AZ', city: 'PHOENIX' }, pop: 890853 },

{ \_id: { state: 'FL', city: 'MIAMI' }, pop: 825232 },

{ \_id: { state: 'CA', city: 'SAN JOSE' }, pop: 816653 },

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

{ \_id: { state: 'MD', city: 'BALTIMORE' }, pop: 733081 },

{ \_id: { state: 'CA', city: 'SAN FRANCISCO' }, pop: 723993 },

{ \_id: { state: 'TN', city: 'MEMPHIS' }, pop: 632837 },

{ \_id: { state: 'CA', city: 'SACRAMENTO' }, pop: 628279 },

{ \_id: { state: 'FL', city: 'JACKSONVILLE' }, pop: 610160 },

{ \_id: { state: 'GA', city: 'ATLANTA' }, pop: 609591 }

]

Type "it" for more

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

Atlas atlas-ojy0k0-shard-0 [primary] Population> db.zipcodes.aggregate([{ $group: { \_id: { state: "$state", city: "$city" },pop: { $sum: "$pop" } } }, { $sort:{pop:-1} },{$limit:3}])

[

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

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

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

]

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

# Bonus

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

Atlas atlas-ojy0k0-shard-0 [primary] Population> db.zipcodes.aggregate( [{ $group: { \_id: { state: "$state", city: "$city" }, pop: { $sum: "$pop" } } },{ $group: { \_id: "$\_id.state", avgCityPop: { $avg: "$pop" } } }] )

[

{ \_id: 'OH', avgCityPop: 12700.839578454332 },

{ \_id: 'SC', avgCityPop: 11139.626198083068 },

{ \_id: 'PA', avgCityPop: 8679.067202337472 },

{ \_id: 'LA', avgCityPop: 10465.496277915632 },

{ \_id: 'VA', avgCityPop: 8526.177931034483 },

{ \_id: 'IL', avgCityPop: 9954.334494773519 },

{ \_id: 'NM', avgCityPop: 5872.360465116279 },

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

{ \_id: 'AR', avgCityPop: 4175.355239786856 },

{ \_id: 'ME', avgCityPop: 3006.4901960784314 },

{ \_id: 'IA', avgCityPop: 3123.0821147356583 },

{ \_id: 'WV', avgCityPop: 2771.4775888717154 },

{ \_id: 'NE', avgCityPop: 3034.882692307692 },

{ \_id: 'ND', avgCityPop: 1645.0309278350514 },

{ \_id: 'NC', avgCityPop: 10622.815705128205 },

{ \_id: 'NY', avgCityPop: 13131.680291970803 },

{ \_id: 'HI', avgCityPop: 15831.842857142858 },

{ \_id: 'TX', avgCityPop: 13775.02108678021 },

{ \_id: 'KS', avgCityPop: 3819.884259259259 },

{ \_id: 'AL', avgCityPop: 7907.2152641878665 }

]

Type "it" for more

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

Atlas atlas-ojy0k0-shard-0 [primary] 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 }