

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.

```
Command Prompt - mongo "mongodb://madhu:ahghp.mongodb.net/madhu" --username Madhu-Manepalli --password madhu
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> show dbs
admin            0.000GB
local            4.547GB
population       0.002GB
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> use population
switched to db population
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> show collections
zipcodes
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.find({'$and':{'city':'ATLANTA','state':'GA'}})
{"_id":"30303","city":"ATLANTA","loc":[{"x":84.388848,"y":33.752504},"pop":1845,"state":"GA"}
{"_id":"30305","city":"ATLANTA","loc":[{"x":84.385145,"y":33.831963},"pop":19122,"state":"GA"}
{"_id":"30306","city":"ATLANTA","loc":[{"x":84.351418,"y":33.786827},"pop":20081,"state":"GA"}
{"_id":"30307","city":"ATLANTA","loc":[{"x":84.335957,"y":33.769138},"pop":16330,"state":"GA"}
{"_id":"30308","city":"ATLANTA","loc":[{"x":84.375744,"y":33.771839},"pop":8549,"state":"GA"}
{"_id":"30309","city":"ATLANTA","loc":[{"x":84.388338,"y":33.798407},"pop":14766,"state":"GA"}
{"_id":"30310","city":"ATLANTA","loc":[{"x":84.423173,"y":33.727849},"pop":34017,"state":"GA"}
{"_id":"30311","city":"ATLANTA","loc":[{"x":84.470219,"y":33.722957},"pop":34880,"state":"GA"}
{"_id":"30312","city":"ATLANTA","loc":[{"x":84.378125,"y":33.746749},"pop":17683,"state":"GA"}
{"_id":"30313","city":"ATLANTA","loc":[{"x":84.39352,"y":33.76825},"pop":8038,"state":"GA"}
{"_id":"30314","city":"ATLANTA","loc":[{"x":84.425546,"y":33.756103},"pop":26649,"state":"GA"}
{"_id":"30315","city":"ATLANTA","loc":[{"x":84.380771,"y":33.705062},"pop":41061,"state":"GA"}
{"_id":"30316","city":"ATLANTA","loc":[{"x":84.333913,"y":33.721086},"pop":34668,"state":"GA"}
{"_id":"30317","city":"ATLANTA","loc":[{"x":84.31685,"y":33.749788},"pop":16395,"state":"GA"}
{"_id":"30318","city":"ATLANTA","loc":[{"x":84.445432,"y":33.786454},"pop":53894,"state":"GA"}
{"_id":"30319","city":"ATLANTA","loc":[{"x":84.335091,"y":33.868728},"pop":32138,"state":"GA"}
{"_id":"30324","city":"ATLANTA","loc":[{"x":84.354867,"y":33.820049},"pop":15044,"state":"GA"}
{"_id":"30326","city":"ATLANTA","loc":[{"x":84.358232,"y":33.848168},"pop":125,"state":"GA"}
{"_id":"30327","city":"ATLANTA","loc":[{"x":84.419966,"y":33.862723},"pop":18467,"state":"GA"}
{"_id":"30328","city":"ATLANTA","loc":[{"x":84.321402,"y":33.823555},"pop":17013,"state":"GA"}
Type "it" for more
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY>
```

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

```
Command Prompt - mongo "mongodb://madhu:ahghp.mongodb.net/madhu" --username Madhu-Manepalli --password madhu
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate({'$match':{'$and':{'city':'ATLANTA','state':'GA'}}})
{"_id":"30303","city":"ATLANTA","loc":[{"x":84.388848,"y":33.752504},"pop":1845,"state":"GA"}
{"_id":"30305","city":"ATLANTA","loc":[{"x":84.385145,"y":33.831963},"pop":19122,"state":"GA"}
{"_id":"30306","city":"ATLANTA","loc":[{"x":84.351418,"y":33.786827},"pop":20081,"state":"GA"}
{"_id":"30307","city":"ATLANTA","loc":[{"x":84.335957,"y":33.769138},"pop":16330,"state":"GA"}
{"_id":"30308","city":"ATLANTA","loc":[{"x":84.375744,"y":33.771839},"pop":8549,"state":"GA"}
{"_id":"30309","city":"ATLANTA","loc":[{"x":84.388338,"y":33.798407},"pop":14766,"state":"GA"}
{"_id":"30310","city":"ATLANTA","loc":[{"x":84.423173,"y":33.727849},"pop":34017,"state":"GA"}
{"_id":"30311","city":"ATLANTA","loc":[{"x":84.470219,"y":33.722957},"pop":34880,"state":"GA"}
{"_id":"30312","city":"ATLANTA","loc":[{"x":84.378125,"y":33.746749},"pop":17683,"state":"GA"}
{"_id":"30313","city":"ATLANTA","loc":[{"x":84.39352,"y":33.76825},"pop":8038,"state":"GA"}
{"_id":"30314","city":"ATLANTA","loc":[{"x":84.425546,"y":33.756103},"pop":26649,"state":"GA"}
{"_id":"30315","city":"ATLANTA","loc":[{"x":84.380771,"y":33.705062},"pop":41061,"state":"GA"}
{"_id":"30316","city":"ATLANTA","loc":[{"x":84.333913,"y":33.721086},"pop":34668,"state":"GA"}
{"_id":"30317","city":"ATLANTA","loc":[{"x":84.31685,"y":33.749788},"pop":16395,"state":"GA"}
{"_id":"30318","city":"ATLANTA","loc":[{"x":84.445432,"y":33.786454},"pop":53894,"state":"GA"}
{"_id":"30319","city":"ATLANTA","loc":[{"x":84.335091,"y":33.868728},"pop":32138,"state":"GA"}
{"_id":"30324","city":"ATLANTA","loc":[{"x":84.354867,"y":33.820049},"pop":15044,"state":"GA"}
{"_id":"30326","city":"ATLANTA","loc":[{"x":84.358232,"y":33.848168},"pop":125,"state":"GA"}
{"_id":"30327","city":"ATLANTA","loc":[{"x":84.419966,"y":33.862723},"pop":18467,"state":"GA"}
{"_id":"30328","city":"ATLANTA","loc":[{"x":84.321402,"y":33.823555},"pop":17013,"state":"GA"}
Type "it" for more
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY>
```

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

```
Command Prompt - mongo "mongodb://madhu:ahghp.mongodb.net/madhu" --username Madhu-Manepalli --password madhu
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate({'$match':{'city':'ATLANTA'}},{'$group':{'_id':{'$id'}}},{'$count':'Count of zipcodes'})
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate({'$match':{'city':'ATLANTA'}},{'$group':{'_id':{'$id'}}},{'$count':'Count of zipcodes'})
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate({'$match':{'city':'ATLANTA'}},{'$group':{'_id':{'$id'}}},{'$count':'Count of zipcodes'})
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> show dbs
admin            0.000GB
local            4.547GB
population       0.002GB
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> use population
switched to db population
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> show collections
zipcodes
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate({'$match':{'city':'ATLANTA'}},{'$group':{'_id':{'$id'}}},{'$count':'count of zipcodes'})
{"_id":"count of zipcodes":41}
```

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

```
count of zipcodes: 41
MongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate({'$match':{'city':'ATLANTA'}},{'$group':{'_id':'city','population':{'$sum':'$pop'}}})
{"_id":"ATLANTA","population":68846}
```

Populations By State

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

```
mongo0 Enterprise atlas-zibv5-shard-0:PRIMARY> db.zipcodes.aggregate([{$group: {_id: "$state", Statepopulation: {$sum: "$pop"}}}])
{"_id": "IN", "Statepopulation": 5544136 }
{"_id": "MI", "Statepopulation": 9295297 }
{"_id": "MS", "Statepopulation": 2522216 }
{"_id": "NJ", "Statepopulation": 7730188 }
{"_id": "CA", "Statepopulation": 29754890 }
{"_id": "TX", "Statepopulation": 16984601 }
{"_id": "VT", "Statepopulation": 562758 }
{"_id": "VA", "Statepopulation": 6381479 }
{"_id": "SD", "Statepopulation": 495397 }
{"_id": "OR", "Statepopulation": 2842321 }
{"_id": "OK", "Statepopulation": 3345585 }
{"_id": "AR", "Statepopulation": 2358725 }
{"_id": "NM", "Statepopulation": 1515869 }
{"_id": "RI", "Statepopulation": 1083218 }
{"_id": "NC", "Statepopulation": 6628637 }
{"_id": "LA", "Statepopulation": 4217595 }
{"_id": "PA", "Statepopulation": 11881643 }
{"_id": "NH", "Statepopulation": 1189252 }
{"_id": "WV", "Statepopulation": 1793146 }
{"_id": "ME", "Statepopulation": 4372982 }
```

2. sort the results by population, highest first

```
mongo0 Enterprise atlas-zibv5-shard-0:PRIMARY> db.zipcodes.aggregate([{$group: {_id: "$state", Statepopulation: {$sum: "$pop"}}}, {$sort: {Statepopulation: -1}}])
{"_id": "CA", "Statepopulation": 29754890 }
{"_id": "NY", "Statepopulation": 17998402 }
{"_id": "TX", "Statepopulation": 16984601 }
{"_id": "FL", "Statepopulation": 12686644 }
{"_id": "PA", "Statepopulation": 11881643 }
{"_id": "IL", "Statepopulation": 11427576 }
{"_id": "OH", "Statepopulation": 10846517 }
{"_id": "MI", "Statepopulation": 9295297 }
{"_id": "NJ", "Statepopulation": 7730188 }
{"_id": "NC", "Statepopulation": 6628637 }
{"_id": "GA", "Statepopulation": 6478216 }
{"_id": "VA", "Statepopulation": 6381479 }
{"_id": "WA", "Statepopulation": 6016425 }
{"_id": "IN", "Statepopulation": 5544136 }
{"_id": "MO", "Statepopulation": 5118648 }
{"_id": "HI", "Statepopulation": 4881760 }
{"_id": "TN", "Statepopulation": 4876457 }
{"_id": "MA", "Statepopulation": 4866692 }
{"_id": "ND", "Statepopulation": 4781379 }
{"_id": "WY", "Statepopulation": 4372982 }
Type "it" for more
mongo0 Enterprise atlas-zibv5-shard-0:PRIMARY>
```

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

```
Command Prompt - mongo "mongodb+srv://madhu:ahgbb.mongodb.net/madhu?username=Madhu-Manepalli&password=madhu"
Type "it" for more
mongo0 Enterprise atlas-zibv5-shard-0:PRIMARY> db.zipcodes.aggregate([{$group: {_id: "$state", Statepopulation: {$sum: "$pop"}}}, {$sort: {Statepopulation: -1}}, {$limit: 3}])
{"_id": "CA", "Statepopulation": 29754890 }
{"_id": "NY", "Statepopulation": 17998402 }
{"_id": "TX", "Statepopulation": 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' }

```
mongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate([{$group:{_id:{city:'$city',state:'$state'},population:{$sum:'$pop'}}}])
{"_id":{"city":"SUFFOLK","state":"VA"},"population":52141}
{"_id":{"city":"PIERSON","state":"FL"},"population":7341}
{"_id":{"city":"PERRIS","state":"CA"},"population":1543}
{"_id":{"city":"TWIN BRIDGES","state":"CA"},"population":0}
{"_id":{"city":"PARISHVILLE","state":"MO"},"population":300}
{"_id":{"city":"GULFPORT","state":"MS"},"population":78008}
{"_id":{"city":"SANDUSKY","state":"MI"},"population":4198}
{"_id":{"city":"LEXINGTON","state":"TX"},"population":3130}
{"_id":{"city":"KENNETT SQUARE","state":"PA"},"population":10663}
{"_id":{"city":"PLAINVIEW","state":"TX"},"population":27037}
{"_id":{"city":"BLUMSBURG","state":"PA"},"population":5443}
{"_id":{"city":"DELOIT","state":"IA"},"population":618}
{"_id":{"city":"MOUNTAIN HOME","state":"WI"},"population":236}
{"_id":{"city":"BURAS","state":"LA"},"population":6406}
{"_id":{"city":"ATASCOSA","state":"TX"},"population":2943}
{"_id":{"city":"GARLAND","state":"TX"},"population":4133}
{"_id":{"city":"HIGHLAND HEIGHTS","state":"CA"},"population":40511}
{"_id":{"city":"POWHER","state":"MO"},"population":1410}
{"_id":{"city":"COVINGTON","state":"GA"},"population":3154}
{"_id":{"city":"AUBREY","state":"AR"},"population":901}
type "it" for more
```

2. sort the results by population, highest first

```
type "it" for more
mongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate([{$group:{_id:{city:'$city',state:'$state'},population:{$sum:'$pop'}}},{sort:{population:-1}}])
{"_id":{"city":"CHICAGO","state":"IL"},"population":2652172}
{"_id":{"city":"BROOKLYN","state":"NY"},"population":2300504}
{"_id":{"city":"LOS ANGELES","state":"CA"},"population":2102295}
{"_id":{"city":"HOUSTON","state":"TX"},"population":2095910}
{"_id":{"city":"PHILADELPHIA","state":"PA"},"population":1610956}
{"_id":{"city":"NEW YORK","state":"NY"},"population":1457000}
{"_id":{"city":"BROWN","state":"NY"},"population":1209548}
{"_id":{"city":"SAN DIEGO","state":"CA"},"population":1040298}
{"_id":{"city":"DETROIT","state":"MI"},"population":903343}
{"_id":{"city":"DALLAS","state":"TX"},"population":940191}
{"_id":{"city":"PHOENIX","state":"AZ"},"population":890853}
{"_id":{"city":"MIAMI","state":"FL"},"population":825223}
{"_id":{"city":"SAN JOSE","state":"CA"},"population":816653}
{"_id":{"city":"SAN ANTONIO","state":"TX"},"population":811792}
{"_id":{"city":"BALTIMORE","state":"MD"},"population":723003}
{"_id":{"city":"SAN FRANCISCO","state":"CA"},"population":723993}
{"_id":{"city":"MEMPHIS","state":"TN"},"population":632837}
{"_id":{"city":"SACRAMENTO","state":"CA"},"population":632179}
{"_id":{"city":"JACKSONVILLE","state":"FL"},"population":610160}
{"_id":{"city":"ATLANTA","state":"GA"},"population":609591}
type "it" for more
mongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY>
```

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

```
Command Prompt - mongo "mongodb+srv://madhu.ahgph.mongodb.net/madhu" --username Madhu-Manepalli --password madhu
type "it" for more
mongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate([{$group:{_id:{city:'$city',state:'$state'},population:{$sum:'$pop'}}},{sort:{population:-1}},{$limit:3}])
{"_id":{"city":"CHICAGO","state":"IL"},"population":2652172}
{"_id":{"city":"BROOKLYN","state":"NY"},"population":2300504}
{"_id":{"city":"LOS ANGELES","state":"CA"},"population":2102295}
```

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

```
mongoDB Enterprise atlas-zsbv5-shard-0:PRIMARY> db.zipcodes.aggregate([{$match:{state:"TX"}},{_id:{city:'$city',state:'$state'},population:{$sum:'$pop'}},{sort:{population:-1}},{$limit:3}])
{"_id":{"city":"HOUSTON","state":"TX"},"population":2095910}
{"_id":{"city":"DALLAS","state":"TX"},"population":940191}
{"_id":{"city":"SAN ANTONIO","state":"TX"},"population":811792}
```

Bonus

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

```
mongoDB Enterprise atlas-zibvss-shard-0-PRIMARY> db.zipcodes.aggregate([{$group: {_id: '$city', AvgPopulation: {$avg: '$pop'}}}])
{"_id": "ABERDEEN", "AvgPopulation": 1338 }
{"_id": "WICK", "AvgPopulation": 479 }
{"_id": "BORDER LAKE", "AvgPopulation": 8481 }
{"_id": "HANCEVILLE", "AvgPopulation": 10186 }
{"_id": "ORANGE PARK", "AvgPopulation": 28264.5 }
{"_id": "TSAILI", "AvgPopulation": 1193 }
{"_id": "PONCE DE LEON", "AvgPopulation": 1317.5 }
{"_id": "HOLLY BLUFF", "AvgPopulation": 218 }
{"_id": "LANGSTON", "AvgPopulation": 1041 }
{"_id": "HEATH", "AvgPopulation": 10316 }
{"_id": "HERICHO CENTER", "AvgPopulation": 3449 }
{"_id": "MIDPRANCH", "AvgPopulation": 7213 }
{"_id": "SHELIXTA", "AvgPopulation": 7854 }
{"_id": "ELLSWORTH AFB", "AvgPopulation": 6355 }
{"_id": "ROCKS RIVER", "AvgPopulation": 236 }
{"_id": "JEFFERSON", "AvgPopulation": 4236.588235294118 }
{"_id": "JOLLETTE", "AvgPopulation": 907 }
{"_id": "COAL MOUNTAIN", "AvgPopulation": 850 }
{"_id": "CRESCENT MILLS", "AvgPopulation": 109 }
{"_id": "POMACH", "AvgPopulation": 1542 }
```

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

```
mongoDB Enterprise atlas-zibvss-shard-0-PRIMARY> db.zipcodes.aggregate([{$group: {_id: '$city', AvgPopulation: {$avg: '$pop'}}}, {$sort: (AvgPopulation: -1)}, {$limit: 3}])
{"_id": "WELL GARDENS", "AvgPopulation": 99568 }
{"_id": "ARLETA", "AvgPopulation": 88114 }
{"_id": "SOUTH GATE", "AvgPopulation": 87826 }
```