**MongoDB -Aggregation Exercises**

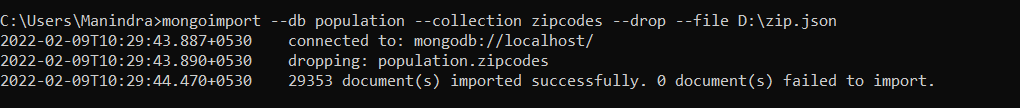
Import the zips.json file into your MongoDB. Database name is "population" and collection name is "zipcodes". mongoimport --db population --collection zipcodes --file zips.json

C:\Users\Manindra>mongoimport --db population --collection zipcodes --drop --file D:\zip.json

2022-02-09T10:29:43.887+0530 connected to: mongodb://localhost/

2022-02-09T10:29:43.890+0530 dropping: population.zipcodes

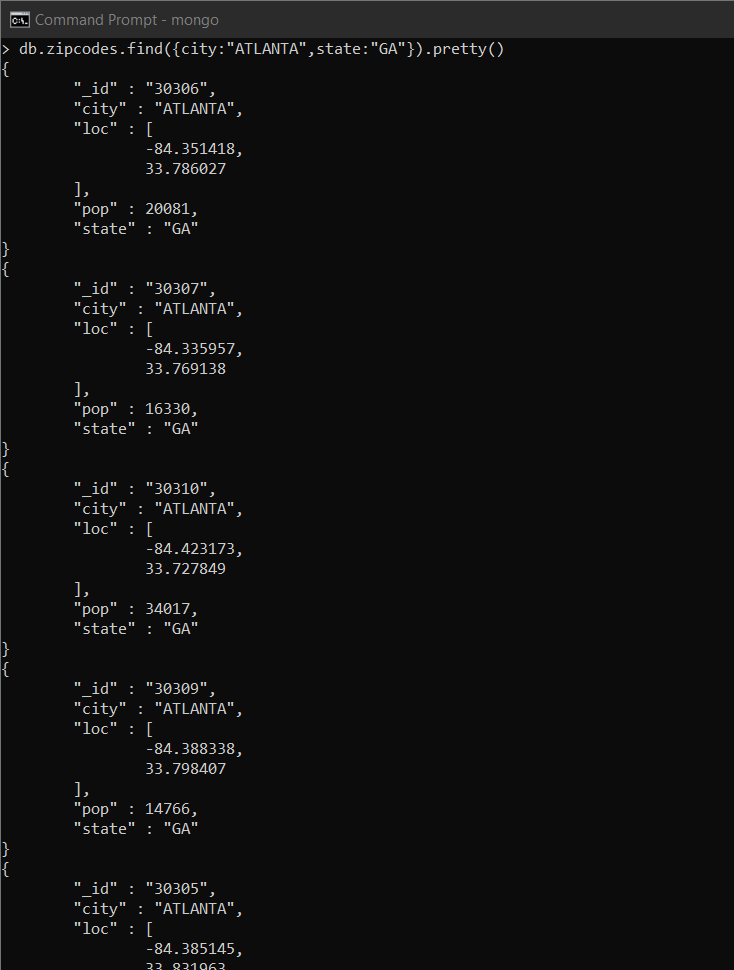
2022-02-09T10:29:44.470+0530 29353 document(s) imported successfully. 0 document(s) failed to import



# Atlanta Population

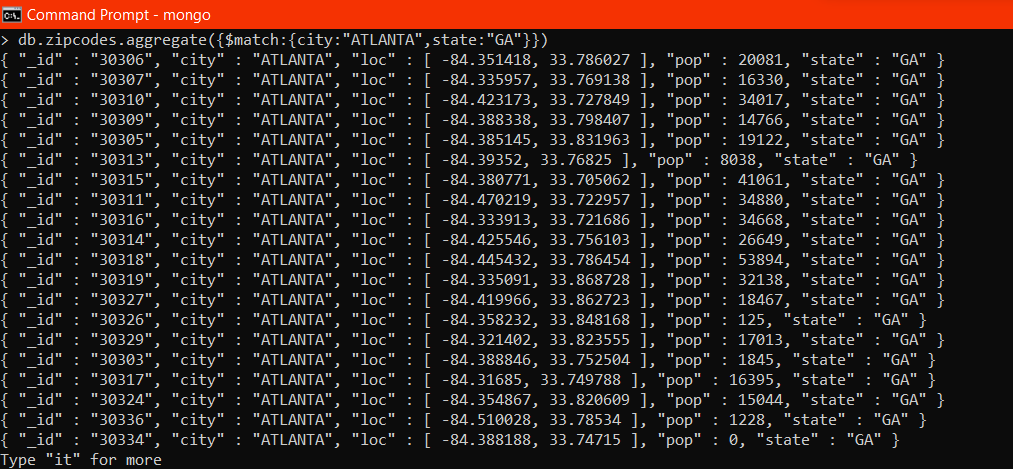
1. use db.zipcodes.find() to filter results to only the results where city is ATLANTA and state is GA.

> db.zipcodes.find({city:"ATLANTA",state:"GA"}).pretty()



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

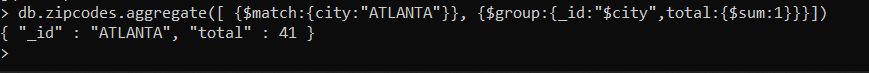
>db.zipcodes.aggregate({$match:{city:"ATLANTA",state:"GA"}})



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

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

{ "\_id" : "ATLANTA", "total" : 41 }

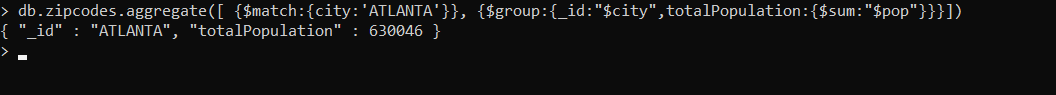


1. use $group to find the total population in Atlanta.

> db.zipcodes.aggregate([ {$match:{city:'ATLANTA'}}, {$group:{\_id:"$city",totalPopulation:{$sum:"$pop"}}}])

{ "\_id" : "ATLANTA", "totalPopulation" : 630046 }

>



# Populations By State

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

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

{ "\_id" : "AL", "totalPopulation" : 4040587 }

{ "\_id" : "SC", "totalPopulation" : 3486703 }

{ "\_id" : "OH", "totalPopulation" : 10846517 }

{ "\_id" : "FL", "totalPopulation" : 12686644 }

{ "\_id" : "MA", "totalPopulation" : 6016425 }

{ "\_id" : "ME", "totalPopulation" : 1226648 }

{ "\_id" : "IL", "totalPopulation" : 11427576 }

{ "\_id" : "AR", "totalPopulation" : 2350725 }

{ "\_id" : "PA", "totalPopulation" : 11881643 }

{ "\_id" : "WY", "totalPopulation" : 453528 }

{ "\_id" : "WV", "totalPopulation" : 1793146 }

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

{ "\_id" : "VA", "totalPopulation" : 6181479 }

{ "\_id" : "MN", "totalPopulation" : 4372982 }

{ "\_id" : "NE", "totalPopulation" : 1578139 }

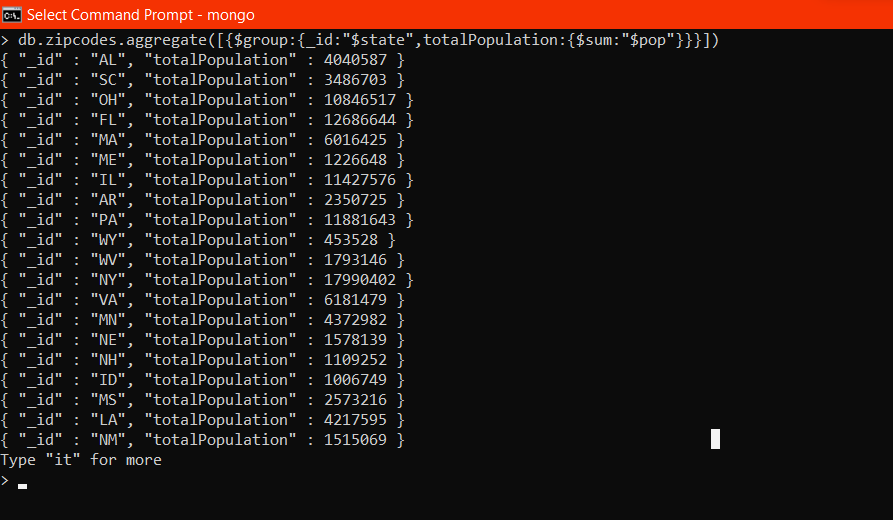
{ "\_id" : "NH", "totalPopulation" : 1109252 }

{ "\_id" : "ID", "totalPopulation" : 1006749 }

{ "\_id" : "MS", "totalPopulation" : 2573216 }

{ "\_id" : "LA", "totalPopulation" : 4217595 }

{ "\_id" : "NM", "totalPopulation" : 1515069 }



1. sort the results by population, highest first

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

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

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

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

{ "\_id" : "FL", "totalPopulation" : 12686644 }

{ "\_id" : "PA", "totalPopulation" : 11881643 }

{ "\_id" : "IL", "totalPopulation" : 11427576 }

{ "\_id" : "OH", "totalPopulation" : 10846517 }

{ "\_id" : "MI", "totalPopulation" : 9295297 }

{ "\_id" : "NJ", "totalPopulation" : 7730188 }

{ "\_id" : "NC", "totalPopulation" : 6628637 }

{ "\_id" : "GA", "totalPopulation" : 6478216 }

{ "\_id" : "VA", "totalPopulation" : 6181479 }

{ "\_id" : "MA", "totalPopulation" : 6016425 }

{ "\_id" : "IN", "totalPopulation" : 5544136 }

{ "\_id" : "MO", "totalPopulation" : 5110648 }

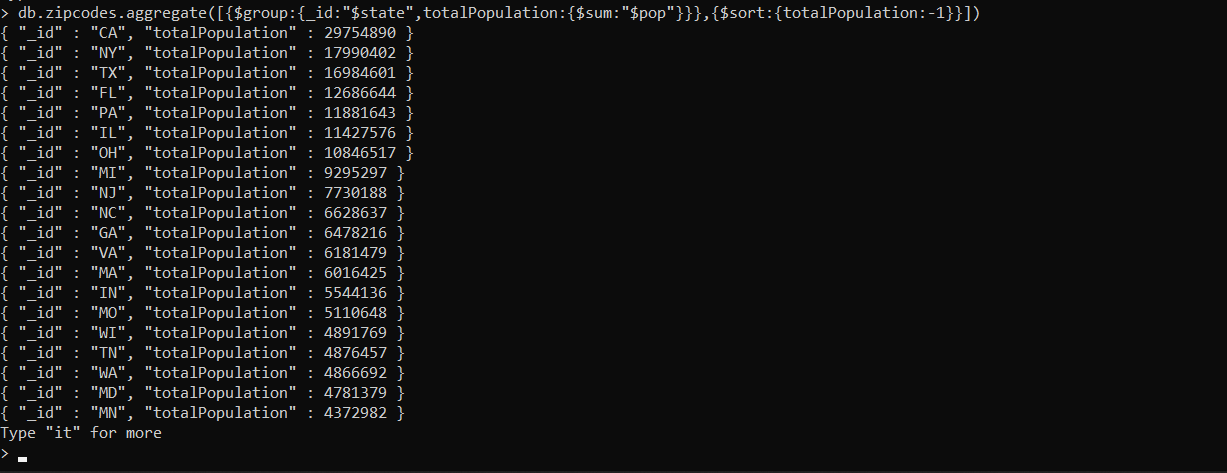
{ "\_id" : "WI", "totalPopulation" : 4891769 }

{ "\_id" : "TN", "totalPopulation" : 4876457 }

{ "\_id" : "WA", "totalPopulation" : 4866692 }

{ "\_id" : "MD", "totalPopulation" : 4781379 }

{ "\_id" : "MN", "totalPopulation" : 4372982 }



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

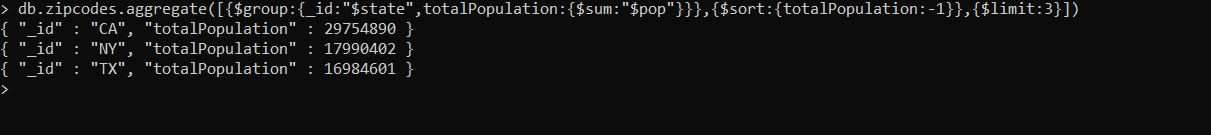
> db.zipcodes.aggregate([{$group:{\_id:"$state",totalPopulation:{$sum:"$pop"}}},{$sort:{totalPopulation:-1}},{$limit:3}])

{ "\_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' }

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

{ "\_id" : { "city" : "UNION GROVE", "state" : "AL", "totalPopulation" : 4921 } }

{ "\_id" : { "city" : "COVENTRY", "state" : "CT", "totalPopulation" : 10776 } }

{ "\_id" : { "city" : "DREYFUS", "state" : "KY", "totalPopulation" : 562 } }

{ "\_id" : { "city" : "GUERNEVILLE", "state" : "CA", "totalPopulation" : 5060 } }

{ "\_id" : { "city" : "MIDLAND", "state" : "NC", "totalPopulation" : 4012 } }

{ "\_id" : { "city" : "LAKEHILLS", "state" : "TX", "totalPopulation" : 3187 } }

{ "\_id" : { "city" : "KNIPPA", "state" : "TX", "totalPopulation" : 618 } }

{ "\_id" : { "city" : "LITTLE VALLEY", "state" : "NY", "totalPopulation" : 2311 } }

{ "\_id" : { "city" : "MILLERSBURG", "state" : "OH", "totalPopulation" : 20350 } }

{ "\_id" : { "city" : "NEW POINT", "state" : "VA", "totalPopulation" : 153 } }

{ "\_id" : { "city" : "FOREST", "state" : "MS", "totalPopulation" : 6588 } }

{ "\_id" : { "city" : "BERLIN CENTER", "state" : "OH", "totalPopulation" : 2771 } }

{ "\_id" : { "city" : "CENTER", "state" : "MO", "totalPopulation" : 971 } }

{ "\_id" : { "city" : "OKLAHOMA CITY", "state" : "OK", "totalPopulation" : 23130 } }



1. sort the results by population, highest first

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

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 112047 } }

{ "\_id" : { "city" : "BROOKLYN", "state" : "NY", "totalPopulation" : 111396 } }

{ "\_id" : { "city" : "NEW YORK", "state" : "NY", "totalPopulation" : 106564 } }

{ "\_id" : { "city" : "NEW YORK", "state" : "NY", "totalPopulation" : 100027 } }

{ "\_id" : { "city" : "BELL GARDENS", "state" : "CA", "totalPopulation" : 99568 } }

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 98612 } }

{ "\_id" : { "city" : "LOS ANGELES", "state" : "CA", "totalPopulation" : 96074 } }

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 95971 } }

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 94317 } }

{ "\_id" : { "city" : "NORWALK", "state" : "CA", "totalPopulation" : 94188 } }

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 92005 } }

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 91814 } }

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 89762 } }

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 88377 } }

{ "\_id" : { "city" : "JACKSON HEIGHTS", "state" : "NY", "totalPopulation" : 88241 } }

{ "\_id" : { "city" : "ARLETA", "state" : "CA", "totalPopulation" : 88114 } }

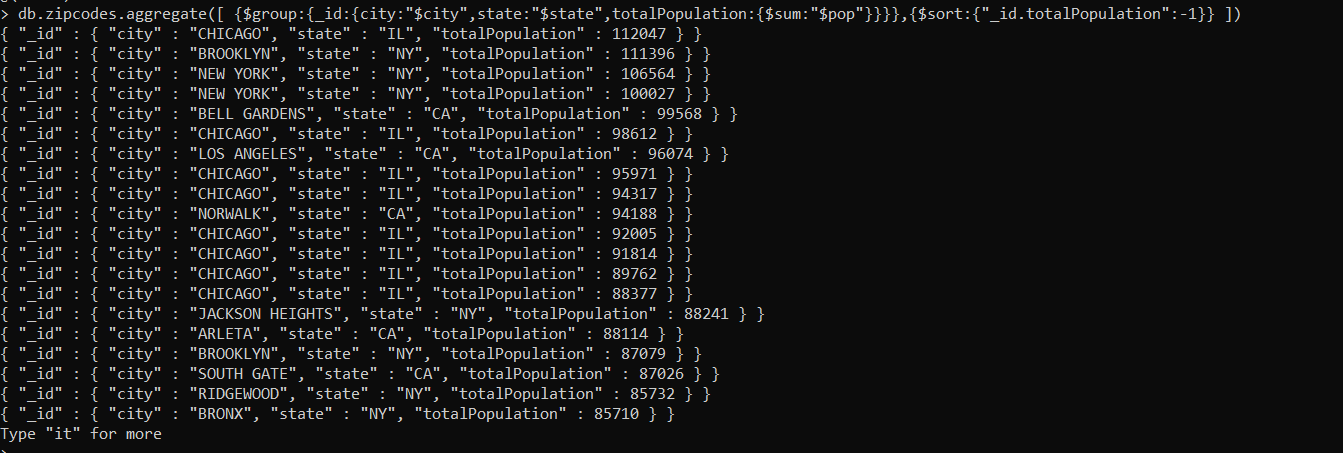
{ "\_id" : { "city" : "BROOKLYN", "state" : "NY", "totalPopulation" : 87079 } }

{ "\_id" : { "city" : "SOUTH GATE", "state" : "CA", "totalPopulation" : 87026 } }

{ "\_id" : { "city" : "RIDGEWOOD", "state" : "NY", "totalPopulation" : 85732 } }

{ "\_id" : { "city" : "BRONX", "state" : "NY", "totalPopulation" : 85710 } }

Type "it" for more



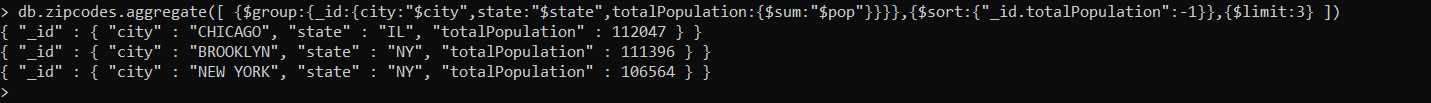
1. limit the results to just the first 3 results. What are the top 3 cities in population?

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

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 112047 } }

{ "\_id" : { "city" : "BROOKLYN", "state" : "NY", "totalPopulation" : 111396 } }

{ "\_id" : { "city" : "NEW YORK", "state" : "NY", "totalPopulation" : 106564 } }



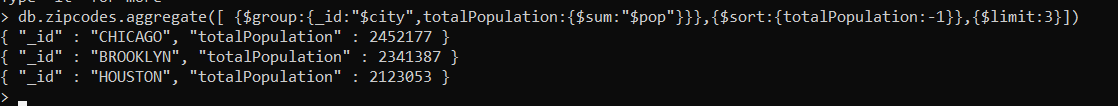
1. What are the top 3 cities in population in Texas?

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

{ "\_id" : "CHICAGO", "totalPopulation" : 2452177 }

{ "\_id" : "BROOKLYN", "totalPopulation" : 2341387 }

{ "\_id" : "HOUSTON", "totalPopulation" : 2123053 }



# Bonus

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

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

{ "\_id" : { "city" : "PORT EWEN", "state" : "NY", "totalPopulation" : 7283 } }

{ "\_id" : { "city" : "COURTENAY", "state" : "ND", "totalPopulation" : 166 } }

{ "\_id" : { "city" : "COLORADO SPRINGS", "state" : "CO", "totalPopulation" : 2550 } }

{ "\_id" : { "city" : "LONE PINE", "state" : "CA", "totalPopulation" : 2257 } }

{ "\_id" : { "city" : "EATON", "state" : "IN", "totalPopulation" : 3637 } }

{ "\_id" : { "city" : "OHATCHEE", "state" : "AL", "totalPopulation" : 3369 } }

{ "\_id" : { "city" : "ABERDEEN", "state" : "OH", "totalPopulation" : 2176 } }

{ "\_id" : { "city" : "STARBUCK", "state" : "MN", "totalPopulation" : 1964 } }

{ "\_id" : { "city" : "DOWNTOWN", "state" : "PA", "totalPopulation" : 1763 } }

{ "\_id" : { "city" : "KERNVILLE", "state" : "CA", "totalPopulation" : 812 } }

{ "\_id" : { "city" : "GRASS CREEK", "state" : "WY", "totalPopulation" : 4809 } }

{ "\_id" : { "city" : "HOLYOKE", "state" : "MA", "totalPopulation" : 43704 } }

{ "\_id" : { "city" : "FARLINGTON", "state" : "KS", "totalPopulation" : 519 } }

{ "\_id" : { "city" : "GIANT FOREST", "state" : "CA", "totalPopulation" : 132 } }

{ "\_id" : { "city" : "JOHNSTOWN", "state" : "PA", "totalPopulation" : 13249 } }

{ "\_id" : { "city" : "POULAN", "state" : "GA", "totalPopulation" : 3119 } }

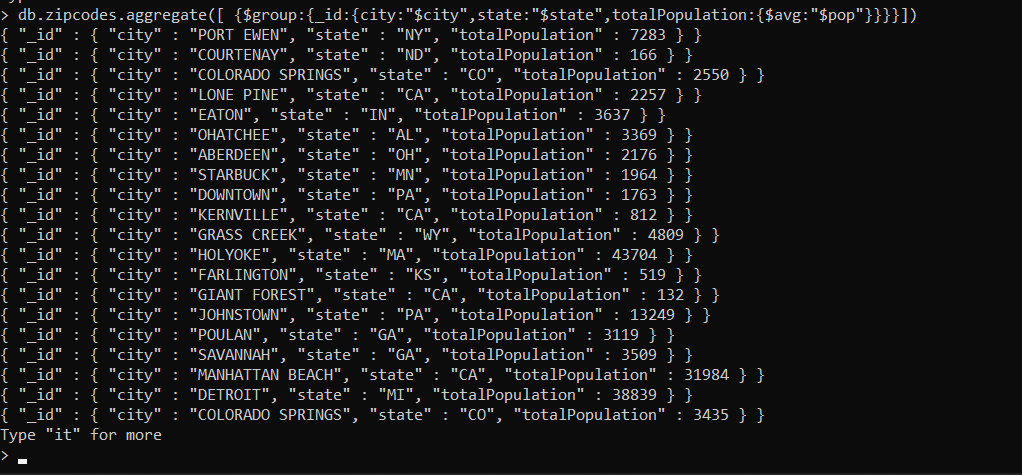
{ "\_id" : { "city" : "SAVANNAH", "state" : "GA", "totalPopulation" : 3509 } }

{ "\_id" : { "city" : "MANHATTAN BEACH", "state" : "CA", "totalPopulation" : 31984 } }

{ "\_id" : { "city" : "DETROIT", "state" : "MI", "totalPopulation" : 38839 } }

{ "\_id" : { "city" : "COLORADO SPRINGS", "state" : "CO", "totalPopulation" : 3435 } }

Type "it" for more



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

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

{ "\_id" : { "city" : "CHICAGO", "state" : "IL", "totalPopulation" : 112047 } }

{ "\_id" : { "city" : "BROOKLYN", "state" : "NY", "totalPopulation" : 111396 } }

{ "\_id" : { "city" : "NEW YORK", "state" : "NY", "totalPopulation" : 106564 } }

