**MongoDB -Aggregation Exercises**

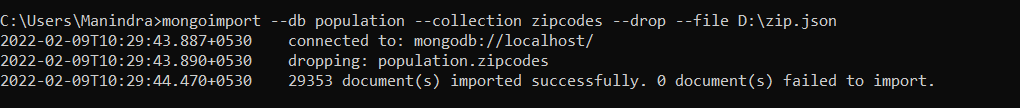
Import the zips.json file into your MongoDB. Database name is "population" and collection name is "zipcodes". mongoimport --db population --collection zipcodes --file zips.json

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

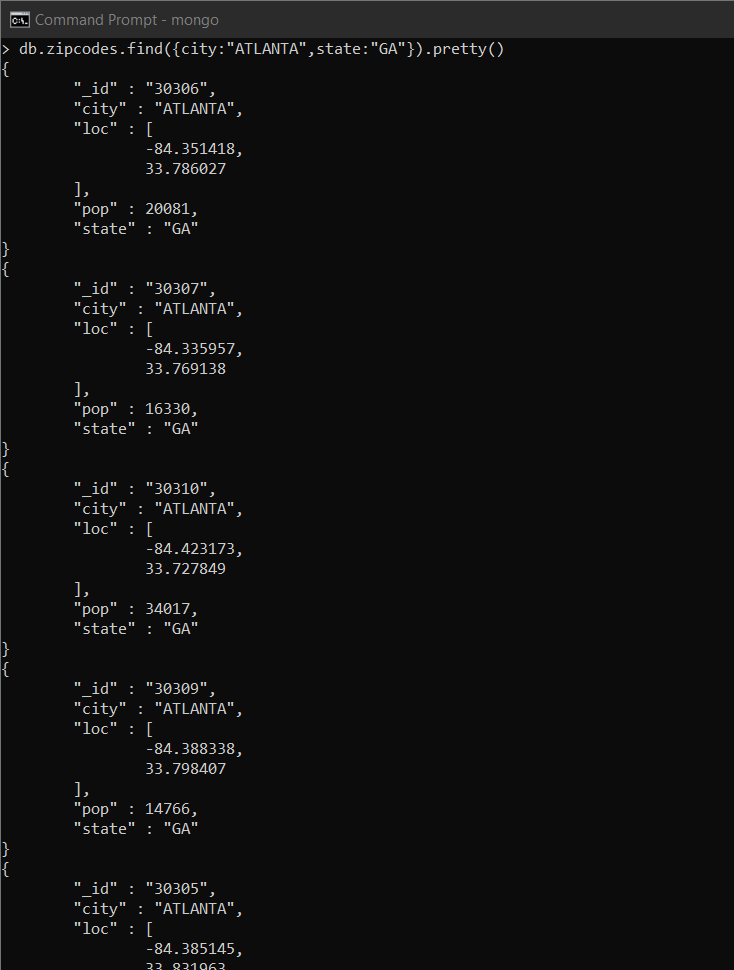
2022-02-09T10:29:44.470+0530 29353 document(s) imported successfully. 0 document(s) failed to import



# Atlanta Population

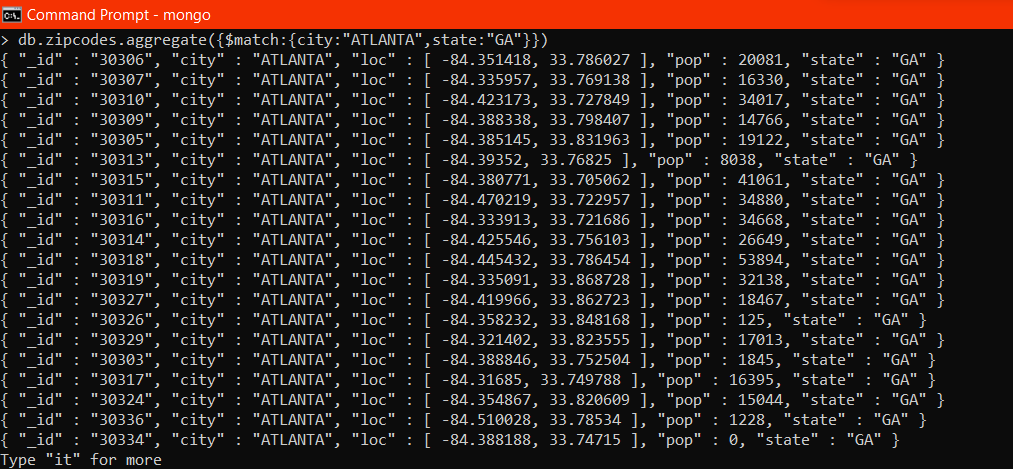
1. use db.zipcodes.find() to filter results to only the results where city is ATLANTA and state is GA.

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



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

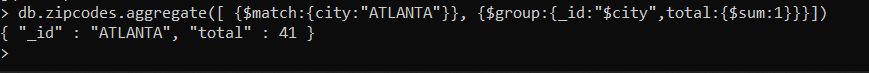
>db.zipcodes.aggregate({$match:{city:"ATLANTA",state:"GA"}})



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 }

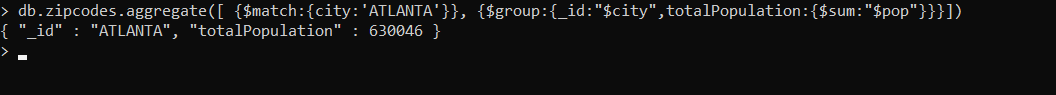


1. use $group to find the total population in Atlanta.

> 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 }

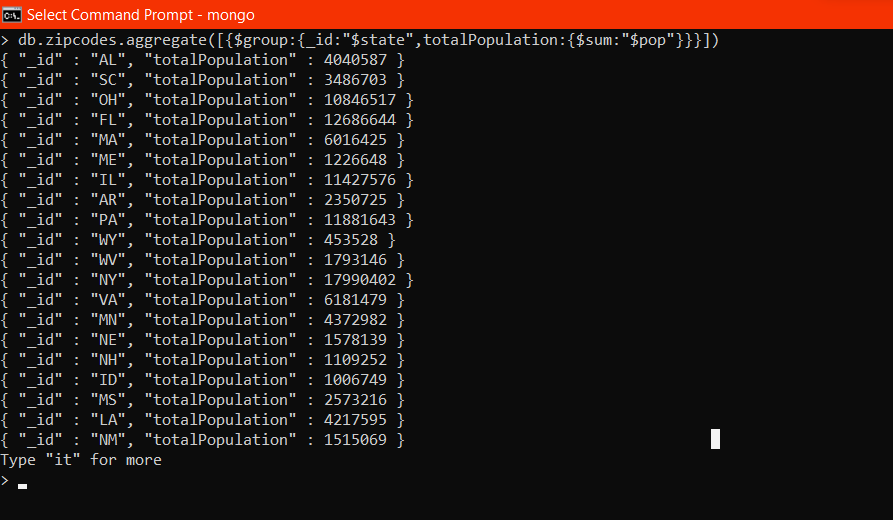
{ "\_id" : "NH", "totalPopulation" : 1109252 }

{ "\_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 }

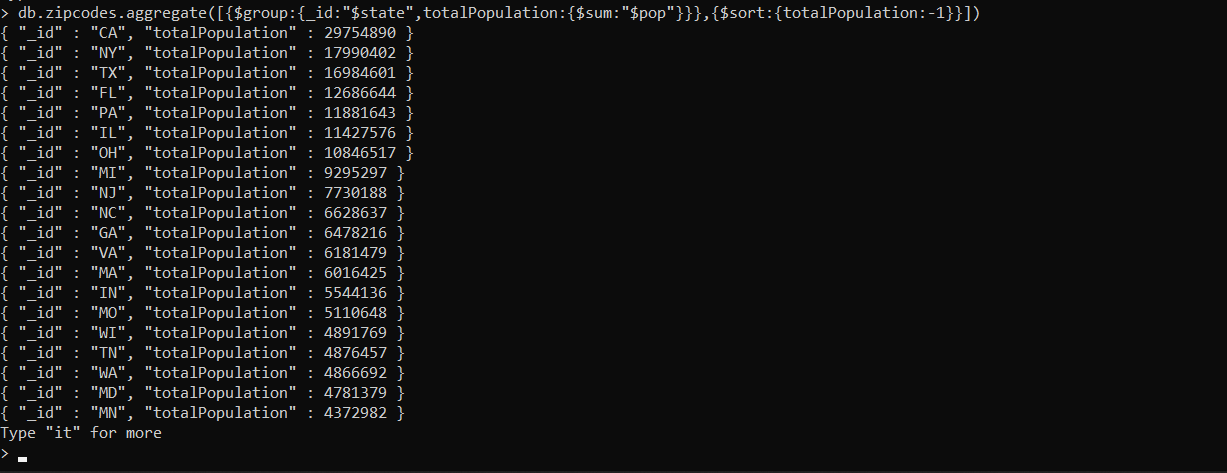
{ "\_id" : "WI", "totalPopulation" : 4891769 }

{ "\_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?

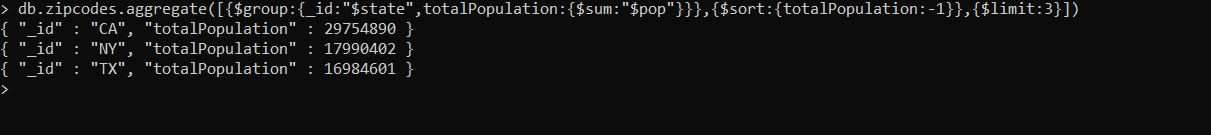
> db.zipcodes.aggregate([{$group:{\_id:"$state",totalPopulation:{$sum:"$pop"}}},{$sort:{totalPopulation:-1}},{$limit:3}])

{ "\_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:"$state",totalPpulation:{$sum:"$pop"}}}}])

{ "\_id" : { "city" : "GREAT MEADOWS", "state" : "NJ", "totalPpulation" : 2303 } }

{ "\_id" : { "city" : "MOUNTAIN LAKE", "state" : "VA", "totalPpulation" : 3087 } }

{ "\_id" : { "city" : "MIFFLIN", "state" : "PA", "totalPpulation" : 1189 } }

{ "\_id" : { "city" : "ELEVA", "state" : "WI", "totalPpulation" : 2291 } }

{ "\_id" : { "city" : "DRAPER", "state" : "SD", "totalPpulation" : 304 } }

{ "\_id" : { "city" : "LETART", "state" : "WV", "totalPpulation" : 4478 } }

{ "\_id" : { "city" : "RANDLETT", "state" : "UT", "totalPpulation" : 78 } }

{ "\_id" : { "city" : "MANAKIN SABOT", "state" : "VA", "totalPpulation" : 5177 } }

{ "\_id" : { "city" : "SIERRA MADRE", "state" : "CA", "totalPpulation" : 10560 } }

{ "\_id" : { "city" : "NORTH HUNTINGDON", "state" : "PA", "totalPpulation" : 36415 } }

{ "\_id" : { "city" : "SAINT AGATHA", "state" : "ME", "totalPpulation" : 919 } }

{ "\_id" : { "city" : "BROADVIEW HEIGHT", "state" : "OH", "totalPpulation" : 11951 } }

{ "\_id" : { "city" : "BROOKWOOD", "state" : "AL", "totalPpulation" : 2319 } }

{ "\_id" : { "city" : "TIMBER LAKE", "state" : "SD", "totalPpulation" : 939 } }

{ "\_id" : { "city" : "GILMORE CITY", "state" : "IA", "totalPpulation" : 604 } }

{ "\_id" : { "city" : "HITCHCOCK", "state" : "SD", "totalPpulation" : 501 } }

{ "\_id" : { "city" : "HAMMOND", "state" : "IN", "totalPpulation" : 12384 } }

{ "\_id" : { "city" : "TRENTON", "state" : "SC", "totalPpulation" : 2825 } }

{ "\_id" : { "city" : "WESTPORT", "state" : "CT", "totalPpulation" : 24705 } }

{ "\_id" : { "city" : "EMPORIUM", "state" : "PA", "totalPpulation" : 5219 } }

Type "it" for more



1. sort the results by population, highest first

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

{ "\_id" : { "city" : "ADDISON", "state" : "IL", "totalPpulation" : 35140 } }

{ "\_id" : { "city" : "HOWARDS GROVE", "state" : "WI", "totalPpulation" : 17004 } }

{ "\_id" : { "city" : "JACKSONVILLE", "state" : "FL", "totalPpulation" : 18244 } }

{ "\_id" : { "city" : "BRENTON", "state" : "WV", "totalPpulation" : 1908 } }

{ "\_id" : { "city" : "HANOVER", "state" : "IL", "totalPpulation" : 1559 } }

{ "\_id" : { "city" : "CLAYTON", "state" : "DE", "totalPpulation" : 4906 } }

{ "\_id" : { "city" : "LOS ANGELES", "state" : "CA", "totalPpulation" : 23530 } }

{ "\_id" : { "city" : "NEWARK", "state" : "CA", "totalPpulation" : 37861 } }

{ "\_id" : { "city" : "MOUNT VERNON", "state" : "TX", "totalPpulation" : 5948 } }

{ "\_id" : { "city" : "RED BANK", "state" : "TN", "totalPpulation" : 21549 } }

{ "\_id" : { "city" : "ELIZABETHTOWN", "state" : "NY", "totalPpulation" : 1274 } }

{ "\_id" : { "city" : "HIGHLAND SPRINGS", "state" : "VA", "totalPpulation" : 9373 } }

{ "\_id" : { "city" : "DASSEL", "state" : "MN", "totalPpulation" : 4818 } }

{ "\_id" : { "city" : "BATCHTOWN", "state" : "IL", "totalPpulation" : 579 } }

{ "\_id" : { "city" : "CAMDEN", "state" : "NJ", "totalPpulation" : 26375 } }

{ "\_id" : { "city" : "JUNCTION CITY", "state" : "GA", "totalPpulation" : 472 } }

{ "\_id" : { "city" : "BUFFALO", "state" : "NY", "totalPpulation" : 10962 } }

{ "\_id" : { "city" : "MADDEN", "state" : "MS", "totalPpulation" : 1586 } }

{ "\_id" : { "city" : "NORCROSS", "state" : "MN", "totalPpulation" : 303 } }

{ "\_id" : { "city" : "WASHINGTON", "state" : "ME", "totalPpulation" : 1261 } }

Type "it" for more



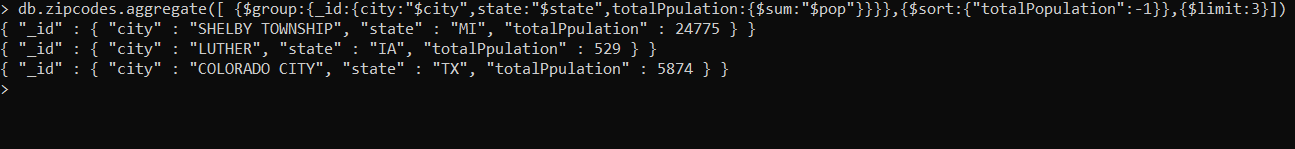
1. limit the results to just the first 3 results. What are the top 3 cities in population?

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

{ "\_id" : { "city" : "SHELBY TOWNSHIP", "state" : "MI", "totalPpulation" : 24775 } }

{ "\_id" : { "city" : "LUTHER", "state" : "IA", "totalPpulation" : 529 } }

{ "\_id" : { "city" : "COLORADO CITY", "state" : "TX", "totalPpulation" : 5874 } }



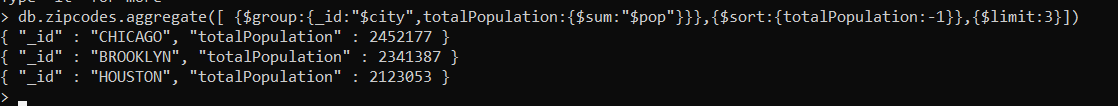
1. What are the top 3 cities in population in Texas?

> 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:"$pop"}}}},{$sort:{"totalPopulation":-1}}])

{ "\_id" : { "city" : "BROOKS", "state" : "OR", "totalPopulation" : 28239 } }

{ "\_id" : { "city" : "SAN JOSE", "state" : "CA", "totalPopulation" : 24778 } }

{ "\_id" : { "city" : "TRENTON", "state" : "TN", "totalPopulation" : 9011 } }

{ "\_id" : { "city" : "WACO", "state" : "GA", "totalPopulation" : 2297 } }

{ "\_id" : { "city" : "BILLINGS", "state" : "MO", "totalPopulation" : 2408 } }

{ "\_id" : { "city" : "STARLIGHT", "state" : "PA", "totalPopulation" : 431 } }

{ "\_id" : { "city" : "LAS VEGAS", "state" : "NV", "totalPopulation" : 5610 } }

{ "\_id" : { "city" : "FRIANT", "state" : "CA", "totalPopulation" : 871 } }

{ "\_id" : { "city" : "FAIR GROVE", "state" : "MO", "totalPopulation" : 2079 } }

{ "\_id" : { "city" : "EL PASO", "state" : "TX", "totalPopulation" : 46537 } }

{ "\_id" : { "city" : "TWIN CITY", "state" : "GA", "totalPopulation" : 5531 } }

{ "\_id" : { "city" : "POWAY", "state" : "CA", "totalPopulation" : 43490 } }

{ "\_id" : { "city" : "SUMITON", "state" : "AL", "totalPopulation" : 3066 } }

{ "\_id" : { "city" : "FRANKLIN", "state" : "NC", "totalPopulation" : 16689 } }

{ "\_id" : { "city" : "WASHINGTON", "state" : "MO", "totalPopulation" : 15437 } }

{ "\_id" : { "city" : "GUALALA", "state" : "CA", "totalPopulation" : 1806 } }

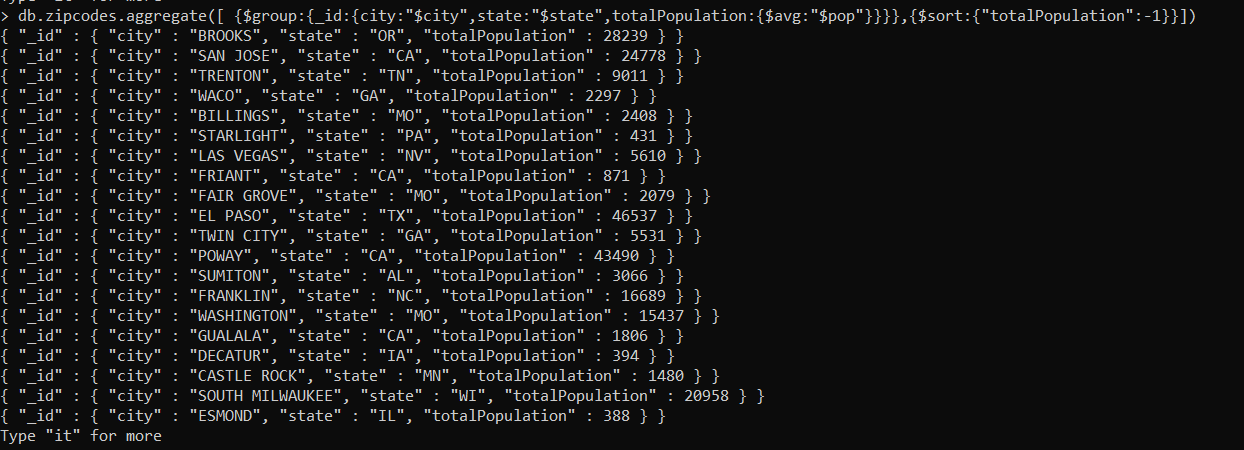
{ "\_id" : { "city" : "DECATUR", "state" : "IA", "totalPopulation" : 394 } }

{ "\_id" : { "city" : "CASTLE ROCK", "state" : "MN", "totalPopulation" : 1480 } }

{ "\_id" : { "city" : "SOUTH MILWAUKEE", "state" : "WI", "totalPopulation" : 20958 } }

{ "\_id" : { "city" : "ESMOND", "state" : "IL", "totalPopulation" : 388 } }

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:{"totalPopulation":-1}},{$limit:3}])

{ "\_id" : { "city" : "SHELBY TOWNSHIP", "state" : "MI", "totalPopulation" : 24775 } }

{ "\_id" : { "city" : "LUTHER", "state" : "IA", "totalPopulation" : 529 } }

{ "\_id" : { "city" : "COLORADO CITY", "state" : "TX", "totalPopulation" : 5874 } }

>

