

MongoDB-Aggregation Exercises

Import the zips.json , run database name “population” collection name is “zipcodes”

```
C:\Users\HP\mongodb-win32-x86_64-2008plus-ssl-3.6.23\bin>mongoimport --db population --collection zipcodes --file D:\Downloads\zip.json
2021-08-02T21:52:13.528+0530    connected to: localhost
2021-08-02T21:52:14.475+0530    [#####.....] population.zipcodes  1.67MB/3.03MB (55.0%)
2021-08-02T21:52:14.894+0530    [#####] population.zipcodes  3.03MB/3.03MB (100.0%)
2021-08-02T21:52:14.894+0530    imported 29353 documents
```

Atlanta Population

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

```
> use population
switched to db population
> db.zipcodes.find({city:"ATLANTA"})
{ "_id" : "14808", "city" : "ATLANTA", "loc" : [ -77.480091, 42.556269 ], "pop" : 472, "state" : "NY" }
{ "_id" : "30305", "city" : "ATLANTA", "loc" : [ -84.385145, 33.831963 ], "pop" : 19122, "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" : "30313", "city" : "ATLANTA", "loc" : [ -84.39352, 33.76825 ], "pop" : 8038, "state" : "GA" }
{ "_id" : "30312", "city" : "ATLANTA", "loc" : [ -84.378125, 33.746749 ], "pop" : 17683, "state" : "GA" }
{ "_id" : "30314", "city" : "ATLANTA", "loc" : [ -84.425546, 33.756103 ], "pop" : 26649, "state" : "GA" }
{ "_id" : "30307", "city" : "ATLANTA", "loc" : [ -84.335957, 33.769138 ], "pop" : 16330, "state" : "GA" }
{ "_id" : "30315", "city" : "ATLANTA", "loc" : [ -84.380771, 33.705062 ], "pop" : 41061, "state" : "GA" }
{ "_id" : "30308", "city" : "ATLANTA", "loc" : [ -84.375744, 33.771839 ], "pop" : 8549, "state" : "GA" }
{ "_id" : "30317", "city" : "ATLANTA", "loc" : [ -84.31685, 33.749788 ], "pop" : 16395, "state" : "GA" }
{ "_id" : "30316", "city" : "ATLANTA", "loc" : [ -84.333913, 33.721686 ], "pop" : 34668, "state" : "GA" }
{ "_id" : "30306", "city" : "ATLANTA", "loc" : [ -84.351418, 33.786027 ], "pop" : 20081, "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" : "30318", "city" : "ATLANTA", "loc" : [ -84.445432, 33.786454 ], "pop" : 53894, "state" : "GA" }
{ "_id" : "30327", "city" : "ATLANTA", "loc" : [ -84.419966, 33.862723 ], "pop" : 18467, "state" : "GA" }
{ "_id" : "30331", "city" : "ATLANTA", "loc" : [ -84.520468, 33.72241 ], "pop" : 38185, "state" : "GA" }
{ "_id" : "30330", "city" : "ATLANTA", "loc" : [ -84.434735, 33.70645 ], "pop" : 643, "state" : "GA" }
Type "it" for more
```

- 2) Use db.zipcodes.aggregate with \$match to do the same as above.

```
> db.zipcodes.aggregate({$match : {city : "ATLANTA"}})
{ "_id" : "14808", "city" : "ATLANTA", "loc" : [ -77.480091, 42.556269 ], "pop" : 472, "state" : "NY" }
{ "_id" : "30305", "city" : "ATLANTA", "loc" : [ -84.385145, 33.831963 ], "pop" : 19122, "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" : "30313", "city" : "ATLANTA", "loc" : [ -84.39352, 33.76825 ], "pop" : 8038, "state" : "GA" }
{ "_id" : "30312", "city" : "ATLANTA", "loc" : [ -84.378125, 33.746749 ], "pop" : 17683, "state" : "GA" }
{ "_id" : "30314", "city" : "ATLANTA", "loc" : [ -84.425546, 33.756103 ], "pop" : 26649, "state" : "GA" }
{ "_id" : "30307", "city" : "ATLANTA", "loc" : [ -84.335957, 33.769138 ], "pop" : 16330, "state" : "GA" }
{ "_id" : "30315", "city" : "ATLANTA", "loc" : [ -84.380771, 33.705062 ], "pop" : 41061, "state" : "GA" }
{ "_id" : "30308", "city" : "ATLANTA", "loc" : [ -84.375744, 33.771839 ], "pop" : 8549, "state" : "GA" }
{ "_id" : "30317", "city" : "ATLANTA", "loc" : [ -84.31685, 33.749788 ], "pop" : 16395, "state" : "GA" }
{ "_id" : "30316", "city" : "ATLANTA", "loc" : [ -84.333913, 33.721686 ], "pop" : 34668, "state" : "GA" }
{ "_id" : "30306", "city" : "ATLANTA", "loc" : [ -84.351418, 33.786027 ], "pop" : 20081, "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" : "30318", "city" : "ATLANTA", "loc" : [ -84.445432, 33.786454 ], "pop" : 53894, "state" : "GA" }
{ "_id" : "30327", "city" : "ATLANTA", "loc" : [ -84.419966, 33.862723 ], "pop" : 18467, "state" : "GA" }
{ "_id" : "30331", "city" : "ATLANTA", "loc" : [ -84.520468, 33.72241 ], "pop" : 38185, "state" : "GA" }
{ "_id" : "30330", "city" : "ATLANTA", "loc" : [ -84.434735, 33.70645 ], "pop" : 643, "state" : "GA" }
Type "it" for more
```

- 3) Use \$group to count the number of zip codes in Atlanta.

```
> db.zipcodes.aggregate([{$match:{city: "ATLANTA"}},{ $group:{_id:"$ _id"}},{ $count:"zip-codes"}])
{ "zip-codes" : 41 }
>
```

- 4) Use \$group to find the total population in Atlanta.

```
> db.zipcodes.aggregate([
... {$group: {_id:"city.ATLANTA",totalpop:{$sum: "$pop"}}}]])
{ "_id" : "city.ATLANTA", "totalpop" : 248408400 }
>
```

Population By State

- 1) Use aggregate to calculate total population of each state

```
> db.zipcodes.aggregate([{$group: {_id:"$state", totalpopulation:{$sum: "$pop"}}}])
{ "_id" : "AK", "totalpopulation" : 544698 }
{ "_id" : "AZ", "totalpopulation" : 3665228 }
{ "_id" : "UT", "totalpopulation" : 1722850 }
{ "_id" : "WY", "totalpopulation" : 453528 }
{ "_id" : "CO", "totalpopulation" : 3293755 }
{ "_id" : "ID", "totalpopulation" : 1006749 }
{ "_id" : "OK", "totalpopulation" : 3145585 }
{ "_id" : "AR", "totalpopulation" : 2350725 }
{ "_id" : "LA", "totalpopulation" : 4217595 }
{ "_id" : "NV", "totalpopulation" : 1201833 }
{ "_id" : "NE", "totalpopulation" : 1578139 }
{ "_id" : "KS", "totalpopulation" : 2475285 }
{ "_id" : "MO", "totalpopulation" : 5110648 }
{ "_id" : "IL", "totalpopulation" : 11427576 }
{ "_id" : "OH", "totalpopulation" : 10846517 }
{ "_id" : "KY", "totalpopulation" : 3675484 }
{ "_id" : "IN", "totalpopulation" : 5544136 }
{ "_id" : "WI", "totalpopulation" : 4891769 }
{ "_id" : "HI", "totalpopulation" : 1108229 }
{ "_id" : "MS", "totalpopulation" : 2573216 }
Type "it" for more
>
```

- 2) 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 }
Type "it" for more
>
```

- 3) Limit the results to just 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 }
>
```

Population by City

- 1) Use agg to cal the total population for each city

```
> db.zipcodes.aggregate([{$group:{_id:{city:"$city", state:"$state"}, totpop:{$sum:"$pop"}}}])
{ "_id" : { "city" : "POINT BAKER", "state" : "AK" }, "totpop" : 426 }
{ "_id" : { "city" : "HYDABURG", "state" : "AK" }, "totpop" : 891 }
{ "_id" : { "city" : "CRAIG", "state" : "AK" }, "totpop" : 1398 }
{ "_id" : { "city" : "KETCHIKAN", "state" : "AK" }, "totpop" : 14308 }
{ "_id" : { "city" : "SITKA", "state" : "AK" }, "totpop" : 8638 }
{ "_id" : { "city" : "PETERSBURG", "state" : "AK" }, "totpop" : 4253 }
{ "_id" : { "city" : "GUSTAVUS", "state" : "AK" }, "totpop" : 258 }
{ "_id" : { "city" : "ANGOON", "state" : "AK" }, "totpop" : 1002 }
{ "_id" : { "city" : "NUIQSUT", "state" : "AK" }, "totpop" : 354 }
{ "_id" : { "city" : "JUNEAU", "state" : "AK" }, "totpop" : 24947 }
{ "_id" : { "city" : "CHALKYITSIK", "state" : "AK" }, "totpop" : 99 }
{ "_id" : { "city" : "WHITE MOUNTAIN", "state" : "AK" }, "totpop" : 194 }
{ "_id" : { "city" : "AMBLER", "state" : "AK" }, "totpop" : 8 }
{ "_id" : { "city" : "TELLER", "state" : "AK" }, "totpop" : 260 }
{ "_id" : { "city" : "SHUNGNAK", "state" : "AK" }, "totpop" : 0 }
{ "_id" : { "city" : "SHAKTOOLIK", "state" : "AK" }, "totpop" : 183 }
{ "_id" : { "city" : "RUBY", "state" : "AK" }, "totpop" : 172 }
{ "_id" : { "city" : "KOYUK", "state" : "AK" }, "totpop" : 231 }
{ "_id" : { "city" : "NULATO", "state" : "AK" }, "totpop" : 492 }
{ "_id" : { "city" : "MANLEY HOT SPRIN", "state" : "AK" }, "totpop" : 122 }
Type "it" for more
>
```

- 2) Sort the results by population, highest first

```
> db.zipcodes.aggregate([{$group: { _id:{city:"$city",state:"$state"}, totalpopulation:{$sum:"$pop"}},{$sort:{totalpopulation:-1}}])
{ "_id" : { "city" : "CHICAGO", "state" : "IL" }, "totalpopulation" : 2452177 }
{ "_id" : { "city" : "BROOKLYN", "state" : "NY" }, "totalpopulation" : 2300504 }
{ "_id" : { "city" : "LOS ANGELES", "state" : "CA" }, "totalpopulation" : 2102295 }
{ "_id" : { "city" : "HOUSTON", "state" : "TX" }, "totalpopulation" : 2095918 }
{ "_id" : { "city" : "PHILADELPHIA", "state" : "PA" }, "totalpopulation" : 1610956 }
{ "_id" : { "city" : "NEW YORK", "state" : "NY" }, "totalpopulation" : 1476790 }
{ "_id" : { "city" : "BRONX", "state" : "NY" }, "totalpopulation" : 1209548 }
{ "_id" : { "city" : "SAN DIEGO", "state" : "CA" }, "totalpopulation" : 1049298 }
{ "_id" : { "city" : "DETROIT", "state" : "MI" }, "totalpopulation" : 963243 }
{ "_id" : { "city" : "DALLAS", "state" : "TX" }, "totalpopulation" : 940191 }
{ "_id" : { "city" : "PHOENIX", "state" : "AZ" }, "totalpopulation" : 890853 }
{ "_id" : { "city" : "MIAMI", "state" : "FL" }, "totalpopulation" : 825232 }
{ "_id" : { "city" : "SAN JOSE", "state" : "CA" }, "totalpopulation" : 816653 }
{ "_id" : { "city" : "SAN ANTONIO", "state" : "TX" }, "totalpopulation" : 811792 }
{ "_id" : { "city" : "BALTIMORE", "state" : "MD" }, "totalpopulation" : 733081 }
{ "_id" : { "city" : "SAN FRANCISCO", "state" : "CA" }, "totalpopulation" : 723993 }
{ "_id" : { "city" : "MEMPHIS", "state" : "TN" }, "totalpopulation" : 632837 }
{ "_id" : { "city" : "SACRAMENTO", "state" : "CA" }, "totalpopulation" : 628279 }
{ "_id" : { "city" : "JACKSONVILLE", "state" : "FL" }, "totalpopulation" : 610160 }
{ "_id" : { "city" : "ATLANTA", "state" : "GA" }, "totalpopulation" : 609591 }
Type "it" for more
```

- 3) Limit the results to just the first 3 results. What are top 3 cities in population.

```
> db.zipcodes.aggregate([{$group: { _id:{city:"$city", state:"$state"}, totalpopulation:{$sum:"$pop"}},{$sort:{totalpopulation:-1}},{$limit:3}])
{ "_id" : { "city" : "CHICAGO", "state" : "IL" }, "totalpopulation" : 2452177 }
{ "_id" : { "city" : "BROOKLYN", "state" : "NY" }, "totalpopulation" : 2300504 }
{ "_id" : { "city" : "LOS ANGELES", "state" : "CA" }, "totalpopulation" : 2102295 }
>
```

4) What are top 3 cities in population in Texas?

```
> db.zipcodes.aggregate([{$match:{state:"TX"}},{$group:{_id:{city:"$city", state:"$state"},totalpopulation:{$sum:"$pop"}},{$sort:{totalpopulation:-1}},{$limit:3}}]
{ "_id" : { "city" : "HOUSTON", "state" : "TX" }, "totalpopulation" : 2095918 }
{ "_id" : { "city" : "DALLAS", "state" : "TX" }, "totalpopulation" : 940191 }
{ "_id" : { "city" : "SAN ANTONIO", "state" : "TX" }, "totalpopulation" : 811792 }
```

Bonus

1) Query to get the average city population for each state

```
> db.zipcodes.aggregate([{$group:{_id:{city:"$city", state:"$state"},avgpop:{$avg:"$pop"}}}])
{ "_id" : { "city" : "POINT BAKER", "state" : "AK" }, "avgpop" : 426 }
{ "_id" : { "city" : "HYDABURG", "state" : "AK" }, "avgpop" : 891 }
{ "_id" : { "city" : "CRAIG", "state" : "AK" }, "avgpop" : 1398 }
{ "_id" : { "city" : "KETCHIKAN", "state" : "AK" }, "avgpop" : 7154 }
{ "_id" : { "city" : "SITKA", "state" : "AK" }, "avgpop" : 8638 }
{ "_id" : { "city" : "PETERSBURG", "state" : "AK" }, "avgpop" : 4253 }
{ "_id" : { "city" : "GUSTAVUS", "state" : "AK" }, "avgpop" : 258 }
{ "_id" : { "city" : "ANGOON", "state" : "AK" }, "avgpop" : 1002 }
{ "_id" : { "city" : "NUIQSUT", "state" : "AK" }, "avgpop" : 354 }
{ "_id" : { "city" : "JUNEAU", "state" : "AK" }, "avgpop" : 24947 }
{ "_id" : { "city" : "CHALKYITSIK", "state" : "AK" }, "avgpop" : 99 }
{ "_id" : { "city" : "WHITE MOUNTAIN", "state" : "AK" }, "avgpop" : 194 }
{ "_id" : { "city" : "AMBLER", "state" : "AK" }, "avgpop" : 8 }
{ "_id" : { "city" : "TELLER", "state" : "AK" }, "avgpop" : 260 }
{ "_id" : { "city" : "SHUNGNAK", "state" : "AK" }, "avgpop" : 0 }
{ "_id" : { "city" : "SHAKTOOLIK", "state" : "AK" }, "avgpop" : 183 }
{ "_id" : { "city" : "RUBY", "state" : "AK" }, "avgpop" : 172 }
{ "_id" : { "city" : "KOYUK", "state" : "AK" }, "avgpop" : 231 }
{ "_id" : { "city" : "NULATO", "state" : "AK" }, "avgpop" : 492 }
{ "_id" : { "city" : "MANLEY HOT SPRING", "state" : "AK" }, "avgpop" : 122 }
Type "it" for more
>
```

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

```
> db.zipcodes.aggregate([{$group:{_id:{city:"$city", state:"$state"},avgpop:{$avg:"$pop"}},{$sort:{avgpop:-1}},{$limit:3}}]
{ "_id" : { "city" : "BELL GARDENS", "state" : "CA" }, "avgpop" : 99568 }
{ "_id" : { "city" : "NORWALK", "state" : "CA" }, "avgpop" : 94188 }
{ "_id" : { "city" : "ARLETA", "state" : "CA" }, "avgpop" : 88114 }
```