

## MongoDB Exercise in mongo shell - Solutions

Import the zip.json file into your MongoDB. Database name is "population" and collection name is "zipcodes".

```
C:\Windows\System32\cmd.exe
C:\Program Files\MongoDB\Server\4.0\bin>mongoimport --db population --collection zipcodes --file "C:\Users\LENOVO\Downloads\iTransform Assignments\zips.json"
connected to: localhost
[.....] population.zipcodes 1.94KB/3.03MB (0.1%)
[#####] population.zipcodes 3.03MB/3.03MB (100.0%)
imported 29353 documents
```

### Atlanta Population

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

```
C:\Windows\System32\cmd.exe - mongo
> 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" : "30310", "city" : "ATLANTA", "loc" : [ -84.423173, 33.727849 ], "pop" : 34017, "state" : "GA" }
{ "_id" : "30309", "city" : "ATLANTA", "loc" : [ -84.388338, 33.798407 ], "pop" : 14766, "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" : "30324", "city" : "ATLANTA", "loc" : [ -84.354867, 33.820609 ], "pop" : 15044, "state" : "GA" }
{ "_id" : "30319", "city" : "ATLANTA", "loc" : [ -84.335091, 33.868728 ], "pop" : 32138, "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" : "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.

```
C:\Windows\System32\cmd.exe - mongo
> db.zipcodes.aggregate([{'$match': {'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" : "30310", "city" : "ATLANTA", "loc" : [ -84.423173, 33.727849 ], "pop" : 34017, "state" : "GA" }
{ "_id" : "30309", "city" : "ATLANTA", "loc" : [ -84.388338, 33.798407 ], "pop" : 14766, "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" : "30324", "city" : "ATLANTA", "loc" : [ -84.354867, 33.820609 ], "pop" : 15044, "state" : "GA" }
{ "_id" : "30319", "city" : "ATLANTA", "loc" : [ -84.335091, 33.868728 ], "pop" : 32138, "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" : "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.

```
C:\Windows\System32\cmd.exe - mongo
> db.zipcodes.aggregate([{$match: {city:"ATLANTA"}}, {$group: {_id: '$_id'}}]).itcount()
41
> db.zipcodes.aggregate([{$match: {city:"ATLANTA"}}, {$group: {_id: '$_id'}}]).toArray().length
41
> db.zipcodes.aggregate([{$match: {city:"ATLANTA"}}, {$group: {_id: '$_id'}}, {$count: "cityCount"}])
{ "cityCount" : 41 }
```

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

```
C:\Windows\System32\cmd.exe - mongo
> db.zipcodes.aggregate([{$match: {city:"ATLANTA"}}, {$group: {_id:"city", "totalPopulation": {$sum: "$pop"}}}])
{ "_id" : "city", "totalPopulation" : 630046 }
> db.zipcodes.aggregate([{$match: {city:"ATLANTA"}}, {$group: {_id:0, "totalPopulation": {$sum: "$pop"}}}])
{ "_id" : 0, "totalPopulation" : 630046 }
```

## Populations By State

---

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

```
C:\Windows\System32\cmd.exe - mongo
> db.zipcodes.aggregate([{$group: {_id: '$state', totalPopulation:{$sum:'$pop'}}}])
{ "_id" : "AK", "totalPopulation" : 544698 }
{ "_id" : "AZ", "totalPopulation" : 3665228 }
{ "_id" : "UT", "totalPopulation" : 1722850 }
{ "_id" : "WV", "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

```
C:\Windows\System32\cmd.exe - mongo
> db.zipcodes.aggregate([{$group: {_id: '$state', totalPopulation:{$sum:'$pop'}}}, {$sort: {totalPopulation:-1}}])
{ "_id" : "CA", "totalPopulation" : 29754890 }
{ "_id" : "NV", "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
```

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

```
C:\Windows\System32\cmd.exe - mongo
> 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

- 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' }

```
C:\Windows\System32\cmd.exe - mongo
> db.zipcodes.aggregate([{$group: {_id: {city: '$city', state: '$state'}, totalPopulation: {$sum: '$pop'}}}])
{ "_id": { "city": "POINT BAKER", "state": "AK" }, "totalPopulation": 426 }
{ "_id": { "city": "CRAIG", "state": "AK" }, "totalPopulation": 1398 }
{ "_id": { "city": "HYDABURG", "state": "AK" }, "totalPopulation": 891 }
{ "_id": { "city": "KETCHIKAN", "state": "AK" }, "totalPopulation": 14308 }
{ "_id": { "city": "SITKA", "state": "AK" }, "totalPopulation": 8638 }
{ "_id": { "city": "PETERSBURG", "state": "AK" }, "totalPopulation": 4253 }
{ "_id": { "city": "GUSTAVUS", "state": "AK" }, "totalPopulation": 258 }
{ "_id": { "city": "CHALKYITSIK", "state": "AK" }, "totalPopulation": 99 }
{ "_id": { "city": "ANGOOK", "state": "AK" }, "totalPopulation": 1002 }
{ "_id": { "city": "AMBLER", "state": "AK" }, "totalPopulation": 8 }
{ "_id": { "city": "JUNEAU", "state": "AK" }, "totalPopulation": 24947 }
{ "_id": { "city": "NUIQSUT", "state": "AK" }, "totalPopulation": 354 }
{ "_id": { "city": "SHUNGNAK", "state": "AK" }, "totalPopulation": 0 }
{ "_id": { "city": "WHITE MOUNTAIN", "state": "AK" }, "totalPopulation": 194 }
{ "_id": { "city": "TELLER", "state": "AK" }, "totalPopulation": 260 }
{ "_id": { "city": "SHAKTOOLIK", "state": "AK" }, "totalPopulation": 183 }
{ "_id": { "city": "RUBY", "state": "AK" }, "totalPopulation": 172 }
{ "_id": { "city": "NULATO", "state": "AK" }, "totalPopulation": 492 }
{ "_id": { "city": "NOATAK", "state": "AK" }, "totalPopulation": 395 }
{ "_id": { "city": "POINT LAV", "state": "AK" }, "totalPopulation": 139 }
Type "it" for more
```

- sort the results by population, highest first.

```
C:\Windows\System32\cmd.exe - mongo
> 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
```

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

```
C:\Windows\System32\cmd.exe - mongo
> 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 the top 3 cities in population in Texas?

```
C:\Windows\System32\cmd.exe - mongo
> 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. Write a query to get the average city population for each state.

```
C:\Windows\System32\cmd.exe - mongo
> db.zipcodes.aggregate([
... {$group: {_id: {state: "$state", city: "$city"}, totalPopulation: {$sum: '$pop'}}},
... {$group: {_id: "$_id.state", avgCityPopulation: {$avg: "$totalPopulation"}}}
... ])
{ "_id" : "NH", "avgCityPopulation" : 5232.320754716981 }
{ "_id" : "MA", "avgCityPopulation" : 14855.37037037037 }
{ "_id" : "ME", "avgCityPopulation" : 3006.4901960784314 }
{ "_id" : "NY", "avgCityPopulation" : 13131.680291970803 }
{ "_id" : "VT", "avgCityPopulation" : 2315.8765432098767 }
{ "_id" : "PA", "avgCityPopulation" : 8679.067202337472 }
{ "_id" : "DC", "avgCityPopulation" : 303450 }
{ "_id" : "DE", "avgCityPopulation" : 14481.91304347826 }
{ "_id" : "VA", "avgCityPopulation" : 8526.177931034483 }
{ "_id" : "SC", "avgCityPopulation" : 11139.626198083068 }
{ "_id" : "FL", "avgCityPopulation" : 27400.958963282937 }
{ "_id" : "AL", "avgCityPopulation" : 7907.2152641878665 }
{ "_id" : "NJ", "avgCityPopulation" : 15775.89387755102 }
{ "_id" : "WV", "avgCityPopulation" : 2771.4775888717154 }
{ "_id" : "TN", "avgCityPopulation" : 9656.350495049504 }
{ "_id" : "OH", "avgCityPopulation" : 12700.839578454332 }
{ "_id" : "MD", "avgCityPopulation" : 12615.775725593667 }
{ "_id" : "MN", "avgCityPopulation" : 5372.21375921376 }
{ "_id" : "ND", "avgCityPopulation" : 1645.0309278350514 }
{ "_id" : "NC", "avgCityPopulation" : 10622.815705128205 }
Type "it" for more
```

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

```
C:\Windows\System32\cmd.exe - mongo
> db.zipcodes.aggregate([
... {$group: {_id: {state: "$state", city: "$city"}, totalPopulation: {$sum: '$pop'}}},
... {$group: {_id: "$_id.state", avgCityPopulation: {$avg: "$totalPopulation"}}},
... {$sort: {avgCityPopulation: -1}},
... {$limit: 3}
... ])
{ "_id" : "DC", "avgCityPopulation" : 303450 }
{ "_id" : "CA", "avgCityPopulation" : 27756.42723880597 }
{ "_id" : "FL", "avgCityPopulation" : 27400.958963282937 }
```