MongoDB -Aggregation Exercises

mongoimport --uri mongodb+srv://DB_USER13:Pass123@samplecluster.2pkxf.mongodb.net/population --collection zipcodes --type json --file F:\zip.json

```
C:\Users\PriyankaY Workspace>mongoimport --uri mongodb+srv://DB_USER13:Pass123@samplecluster.2pkxf.mongodb.net/population --collection zipcodes --type json --file F:\zip.json
2021-10-24T20:04:08.576+0530    connected to: mongodb+srv://[**REDACTED**]@samplecluster.2pkxf.mongodb.net/population
2021-10-24T20:04:11.578+0530    [####....................] population.zipcodes  539KB/3.03MB (17.3%)
2021-10-24T20:04:14.587+0530    [######..................] population.zipcodes  859KB/3.03MB (27.6%)
2021-10-24T20:04:17.579+0530    [#########...............] population.zipcodes  1.15MB/3.03MB (37.9%)
2021-10-24T20:04:20.581+0530    [###########.............] population.zipcodes  1.46MB/3.03MB (48.2%)
2021-10-24T20:04:23.582+0530    [##############..........] population.zipcodes  1.77MB/3.03MB (58.4%)
2021-10-24T20:04:26.580+0530    [################........] population.zipcodes  2.08MB/3.03MB (68.5%)
2021-10-24T20:04:29.588+0530    [##################......] population.zipcodes  2.39MB/3.03MB (78.7%)
2021-10-24T20:04:32.578+0530    [####################....] population.zipcodes  2.70MB/3.03MB (88.8%)
2021-10-24T20:04:35.580+0530    [######################..] population.zipcodes  3.00MB/3.03MB (98.9%)
2021-10-24T20:04:37.526+0530    [########################] population.zipcodes  3.03MB/3.03MB (100.0%)
2021-10-24T20:04:37.527+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({$and: [{ city: ”ATLANTA” } , { state: ”GA” }]})

```
C:\Users\PriyankaY Workspace>mongosh "mongodb+srv://samplecluster.2pkxf.mongodb.net/population" --username DB_USER13
Enter password: *******
Current Mongosh Log ID: 61765706ac5a5df1ddfc1698
Connecting to:          mongodb+srv://samplecluster.2pkxf.mongodb.net/population
Using MongoDB:          4.4.10
Using Mongosh:          1.1.0

For mongosh info see: https://docs.mongodb.com/mongodb-shell/

Atlas atlas-1483qj-shard-0 [primary] population> 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: '30310',
    city: 'ATLANTA',
    loc: [ -84.423173, 33.727849 ],
    pop: 34017,
    state: 'GA'
  },
  {
    _id: '30309',
```

2. use db.zipcodes.aggregate with $match to do the same as above.
   db.zipcodes.aggregate([{$match: {city: "ATLANTA"}},{$match: {state: "GA"}}])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate([{$match: {city: "ATLANTA"}},{$match: {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,
```

3. use $group to count the number of zip codes in Atlanta.
   db.zipcodes.aggregate([{$match: {city: "ATLANTA"}},{ $group: { _id:null, ZipCount: { $sum: 1 } } }])
4. use $group to find the total population in Atlanta.
   db.zipcodes.aggregate([{$match: {city: "ATLANTA"}},{ $group: { _id:null, TotalPopulation: { $sum: "$pop" } } }])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{$match: {city: "ATLANTA"}},{ $group: { _id:null, ZipCount: { $sum: 1 }
 } }] )
[ { _id: null, ZipCount: 41 } ]
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{$match: {city: "ATLANTA"}},{ $group: { _id:null, TotalPopulation: { $s
$sum: "$pop" } } }] )
[ { _id: null, TotalPopulation: 630046 } ]
Atlas atlas-1483qj-shard-0 [primary] population>
```

Populations By State

1. use aggregate to calculate the total population for each state
   db.zipcodes.aggregate([{ $group: { _id:{state: '$state'}, StatePopulation: { $sum: "$pop" } } }])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{ $group: { _id:{state: '$state'}, StatePopulation: { $sum: "$pop" } }
}] )
[
  { _id: { state: 'MD' }, StatePopulation: 4781379 },
  { _id: { state: 'ND' }, StatePopulation: 638272 },
  { _id: { state: 'MO' }, StatePopulation: 5110648 },
  { _id: { state: 'MA' }, StatePopulation: 6016425 },
  { _id: { state: 'NE' }, StatePopulation: 1578139 },
  { _id: { state: 'ID' }, StatePopulation: 1006749 },
  { _id: { state: 'ME' }, StatePopulation: 1226648 },
  { _id: { state: 'CA' }, StatePopulation: 29754890 },
  { _id: { state: 'NH' }, StatePopulation: 1109252 },
  { _id: { state: 'MN' }, StatePopulation: 4372982 },
  { _id: { state: 'OR' }, StatePopulation: 2842321 },
  { _id: { state: 'PA' }, StatePopulation: 11881643 },
  { _id: { state: 'NC' }, StatePopulation: 6628637 },
  { _id: { state: 'LA' }, StatePopulation: 4217595 },
  { _id: { state: 'KY' }, StatePopulation: 3675484 },
  { _id: { state: 'DC' }, StatePopulation: 606900 },
  { _id: { state: 'SC' }, StatePopulation: 3486703 },
  { _id: { state: 'OH' }, StatePopulation: 10846517 },
  { _id: { state: 'MI' }, StatePopulation: 9295297 },
  { _id: { state: 'IL' }, StatePopulation: 11427576 }
]
Type "it" for more
Atlas atlas-1483qj-shard-0 [primary] population>
```

2. sort the results by population, highest first
   db.zipcodes.aggregate([{ $group: { _id:{state: '$state'}, StatePopulation: { $sum: "$pop" } } }, { $sort: {StatePopulation: -1} }])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{ $group: { _id:{state: '$state'}, StatePopulation: { $sum: "$pop" } }
}, { $sort: {StatePopulation: -1} }] )
[
  { _id: { state: 'CA' }, StatePopulation: 29754890 },
  { _id: { state: 'NY' }, StatePopulation: 17990402 },
  { _id: { state: 'TX' }, StatePopulation: 16984601 },
  { _id: { state: 'FL' }, StatePopulation: 12686644 },
  { _id: { state: 'PA' }, StatePopulation: 11881643 },
  { _id: { state: 'IL' }, StatePopulation: 11427576 },
  { _id: { state: 'OH' }, StatePopulation: 10846517 },
  { _id: { state: 'MI' }, StatePopulation: 9295297 },
  { _id: { state: 'NJ' }, StatePopulation: 7730188 },
  { _id: { state: 'NC' }, StatePopulation: 6628637 },
  { _id: { state: 'GA' }, StatePopulation: 6478216 },
  { _id: { state: 'VA' }, StatePopulation: 6181479 },
  { _id: { state: 'MA' }, StatePopulation: 6016425 },
  { _id: { state: 'IN' }, StatePopulation: 5544136 },
  { _id: { state: 'MO' }, StatePopulation: 5110648 },
  { _id: { state: 'WI' }, StatePopulation: 4891769 },
  { _id: { state: 'TN' }, StatePopulation: 4876457 },
  { _id: { state: 'WA' }, StatePopulation: 4866692 },
  { _id: { state: 'MD' }, StatePopulation: 4781379 },
  { _id: { state: 'MN' }, StatePopulation: 4372982 }
]
Type "it" for more
Atlas atlas-1483qj-shard-0 [primary] population>
```

3. limit the results to just the first 3 results. What are the top 3 states in population?
   db.zipcodes.aggregate([{ $group: { _id:{state: '$state'}, StatePopulation: { $sum: "$pop" } } }, { $sort: {StatePopulation: -1} }, { $limit: 3}])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{ $group: { _id:{state: '$state'}, StatePopulation: { $sum: "$pop" } }
}, { $sort: {StatePopulation: -1} }, { $limit: 3}] )
[
  { _id: { state: 'CA' }, StatePopulation: 29754890 },
  { _id: { state: 'NY' }, StatePopulation: 17990402 },
  { _id: { state: 'TX' }, StatePopulation: 16984601 }
]
Atlas atlas-1483qj-shard-0 [primary] population>
```

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'}, CityPopulation: { $sum: "$pop" } } }])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{ $group: { _id:{ city: '$city', state: '$state'}, CityPopulation: { $s
$sum: "$pop" } } }] )
[
  { _id: { city: 'BOISE CITY', state: 'OK' }, CityPopulation: 2064 },
  { _id: { city: 'GREENVILLE', state: 'VA' }, CityPopulation: 1975 },
  { _id: { city: 'LITTLERIVER', state: 'CA' }, CityPopulation: 1119 },
  { _id: { city: 'FARMINGTON', state: 'WA' }, CityPopulation: 355 },
  { _id: { city: 'BONAPARTE', state: 'IA' }, CityPopulation: 903 },
  { _id: { city: 'DIMONDALE', state: 'MI' }, CityPopulation: 5131 },
  { _id: { city: 'RAKE', state: 'IA' }, CityPopulation: 435 },
  { _id: { city: 'MILO', state: 'IA' }, CityPopulation: 1705 },
  { _id: { city: 'HORTONVILLE', state: 'NY' }, CityPopulation: 59 },
  { _id: { city: 'NATIONAL PARK', state: 'NJ' }, CityPopulation: 3398 },
  { _id: { city: 'GROTON', state: 'NY' }, CityPopulation: 5940 },
  { _id: { city: 'NUNDA', state: 'NY' }, CityPopulation: 2439 },
  { _id: { city: 'BASSETT', state: 'VA' }, CityPopulation: 14101 },
  { _id: { city: 'BOLIVAR', state: 'OH' }, CityPopulation: 4225 },
  { _id: { city: 'GRAPEVINE', state: 'AR' }, CityPopulation: 591 },
  { _id: { city: 'PEABODY', state: 'KS' }, CityPopulation: 1927 },
  {
    _id: { city: 'NINE MILE FALLS', state: 'WA' },
    CityPopulation: 4872
  },
  {
    _id: { city: 'GILMANTON IRON W', state: 'NH' },
    CityPopulation: 1301
  },
  { _id: { city: 'BOWLING GREEN', state: 'VA' }, CityPopulation: 559 },
  { _id: { city: 'JEFFERSON', state: 'NC' }, CityPopulation: 3080 }
]
Type "it" for more
Atlas atlas-1483qj-shard-0 [primary] population>
```

2. sort the results by population, highest first
   db.zipcodes.aggregate([{ $group: { _id:{ city: '$city', state: '$state'}, CityPopulation: { $sum: "$pop" } } }, {$sort:{CityPopulation: -1}}])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{ $group: { _id:{ city: '$city', state: '$state'}, CityPopulation: { $s
$sum: "$pop" } } }, {$sort:{CityPopulation: -1}}] )
[
  { _id: { city: 'CHICAGO', state: 'IL' }, CityPopulation: 2452177 },
  { _id: { city: 'BROOKLYN', state: 'NY' }, CityPopulation: 2300504 },
  {
    _id: { city: 'LOS ANGELES', state: 'CA' },
    CityPopulation: 2102295
  },
  { _id: { city: 'HOUSTON', state: 'TX' }, CityPopulation: 2095918 },
  {
    _id: { city: 'PHILADELPHIA', state: 'PA' },
    CityPopulation: 1610956
  },
  { _id: { city: 'NEW YORK', state: 'NY' }, CityPopulation: 1476790 },
  { _id: { city: 'BRONX', state: 'NY' }, CityPopulation: 1209548 },
  { _id: { city: 'SAN DIEGO', state: 'CA' }, CityPopulation: 1049298 },
  { _id: { city: 'DETROIT', state: 'MI' }, CityPopulation: 963243 },
  { _id: { city: 'DALLAS', state: 'TX' }, CityPopulation: 940191 },
  { _id: { city: 'PHOENIX', state: 'AZ' }, CityPopulation: 890853 },
  { _id: { city: 'MIAMI', state: 'FL' }, CityPopulation: 825232 },
  { _id: { city: 'SAN JOSE', state: 'CA' }, CityPopulation: 816653 },
  { _id: { city: 'SAN ANTONIO', state: 'TX' }, CityPopulation: 811792 },
  { _id: { city: 'BALTIMORE', state: 'MD' }, CityPopulation: 733081 },
  {
    _id: { city: 'SAN FRANCISCO', state: 'CA' },
    CityPopulation: 723993
  },
  { _id: { city: 'MEMPHIS', state: 'TN' }, CityPopulation: 632837 },
  { _id: { city: 'SACRAMENTO', state: 'CA' }, CityPopulation: 628279 },
  {
Atlas atlas-1483qj-shard-0 [primary] population>
```

3. 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'}, CityPopulation: { $sum: "$pop" } } }, {$sort:{CityPopulation: -1}}, {$limit: 3}])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{ $group: { _id:{ city: '$city', state: '$state'}, CityPopulation: { $s
$sum: "$pop" } } }, {$sort:{CityPopulation: -1}}, {$limit: 3}] )
[
  { _id: { city: 'CHICAGO', state: 'IL' }, CityPopulation: 2452177 },
  { _id: { city: 'BROOKLYN', state: 'NY' }, CityPopulation: 2300504 },
  {
    _id: { city: 'LOS ANGELES', state: 'CA' },
    CityPopulation: 2102295
  }
]
Atlas atlas-1483qj-shard-0 [primary] population>
```

4. What are the top 3 cities in population in Texas?
   db.zipcodes.aggregate([{$match: {state: "TX"}}, { $group: { _id:{ city: '$city', state: '$state'}, CityPopulation: { $sum: "$pop" } } }, {$sort:{CityPopulation: -1}}, {$limit: 3}])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{$match: {state: "TX"}}, { $group: { _id:{ city: '$city', state: '$stat
e'}, CityPopulation: { $sum: "$pop" } } }, {$sort:{CityPopulation: -1}},
 {$limit: 3}] )
[
  { _id: { city: 'HOUSTON', state: 'TX' }, CityPopulation: 2095918 },
  { _id: { city: 'DALLAS', state: 'TX' }, CityPopulation: 940191 },
  { _id: { city: 'SAN ANTONIO', state: 'TX' }, CityPopulation: 811792 }
]
Atlas atlas-1483qj-shard-0 [primary] population>
```

Bonus

1. Write a query to get the average city population for each state.
   db.zipcodes.aggregate([{ $group: { _id:{ state: '$state'},AvgStateCitiesPop: { $avg: '$pop' } } }])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{ $group: { _id:{ state: '$state'},AvgStateCitiesPop: { $avg: '$pop' }
} }] )
[
  { _id: { state: 'IL' }, AvgStateCitiesPop: 9238.137429264349 },
  { _id: { state: 'LA' }, AvgStateCitiesPop: 9089.644396551725 },
  { _id: { state: 'NM' }, AvgStateCitiesPop: 5489.380434782609 },
  { _id: { state: 'WI' }, AvgStateCitiesPop: 6832.079608938548 },
  { _id: { state: 'NV' }, AvgStateCitiesPop: 11556.086538461539 },
  { _id: { state: 'AL' }, AvgStateCitiesPop: 7126.255731922399 },
  { _id: { state: 'OK' }, AvgStateCitiesPop: 5367.892491467577 },
  { _id: { state: 'VT' }, AvgStateCitiesPop: 2315.8765432098767 },
  { _id: { state: 'TN' }, AvgStateCitiesPop: 8378.792096219931 },
  { _id: { state: 'GA' }, AvgStateCitiesPop: 10201.914960629922 },
  { _id: { state: 'MS' }, AvgStateCitiesPop: 7088.749311294766 },
  { _id: { state: 'DE' }, AvgStateCitiesPop: 12569.207547169812 },
  { _id: { state: 'IN' }, AvgStateCitiesPop: 8201.384615384615 },
  { _id: { state: 'SD' }, AvgStateCitiesPop: 1810.9296875 },
  { _id: { state: 'MT' }, AvgStateCitiesPop: 2544.420382165605 },
  { _id: { state: 'KS' }, AvgStateCitiesPop: 3461.937062937063 },
  { _id: { state: 'UT' }, AvgStateCitiesPop: 8404.146341463415 },
  { _id: { state: 'CT' }, AvgStateCitiesPop: 12498.539923954373 },
  { _id: { state: 'NY' }, AvgStateCitiesPop: 11279.248902821317 },
  { _id: { state: 'NJ' }, AvgStateCitiesPop: 14315.162962962962 }
]
Type "it" for more
Atlas atlas-1483qj-shard-0 [primary] population>
```

2. What are the top 3 states in terms of average city population?
   db.zipcodes.aggregate([{ $group: { _id:{ state: '$state'},AvgStateCitiesPop: { $avg: '$pop' } } },{$sort: {AvgStateCitiesPop: -1}}, {$limit:3}])

```
Atlas atlas-1483qj-shard-0 [primary] population> db.zipcodes.aggregate(
[{ $group: { _id:{ state: '$state'},AvgStateCitiesPop: { $avg: '$pop' }
} },{$sort: {AvgStateCitiesPop: -1}}, {$limit:3}] )
[
  { _id: { state: 'DC' }, AvgStateCitiesPop: 25287.5 },
  { _id: { state: 'CA' }, AvgStateCitiesPop: 19627.236147757256 },
  { _id: { state: 'FL' }, AvgStateCitiesPop: 15779.407960199005 }
]
Atlas atlas-1483qj-shard-0 [primary] population>
```