

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
```

MongoDB Compass - cluster0.f1gpe.mongodb.net/population

Connect View Help

Local

6 DBS 7 COLLECTIONS

☆ FAVORITE

HOSTS

- cluster0-shard-00-00.f1gp...
- cluster0-shard-00-01.f1gp...
- cluster0-shard-00-02.f1gp...

CLUSTER

Replica Set (atlas-bjw9t-s...

3 Nodes

EDITION

MongoDB 5.0.6 Enterprise

Q Filter your data

- admin
- local
- mongo_practice
 - MOvies
- nehabhatt1221
- population + -
- zipcodes
- test

+

Collections

Create collection View

Sort by Collection Name

zipcodes

Storage size:	Documents:	Avg. document size:	Indexes:	Total index size:
1.28 MB	29 K	94.00 B	1	626.69 kB

MongoDB Compass - cluster0.flgpe.mongodb.net/population.zipcodes

Connect View Help

Local

6 DBS 7 COLLECTIONS

☆ FAVORITE

HOSTS

cluster0-shard-00-00.f1gp...
cluster0-shard-00-01.f1gp...
cluster0-shard-00-02.f1gp...

CLUSTER

Replica Set (atlas-bjw9t-s...
3 Nodes

EDITION

MongoDB 5.0.6 Enterprise

Q Filter your data

admin
local
mongo_practice
MOVies
nehabhatt1221
population
zipcodes
test

population.zipcodes Documents

DOCUMENTS 29.4k STORAGE SIZE 1.3MB AVG. SIZE 94B INDEXES 1 TOTAL SIZE 626.7KB AVG. SIZE 626.7KB

Documents Aggregations Schema Explain Plan Indexes Validation

FILTER { field: 'value' } OPTIONS FIND RESET ↺ ...

ADD DATA VIEW {}

Displaying documents 1 - 20 of 29353 REFRESH

```
{
  "_id": "01001",
  "city": "AGAWAM",
  "loc": Array,
  "pop": 15338,
  "state": "MA"
}
```

```
{
  "_id": "01002",
  "city": "CUSHMAN",
  "loc": Array,
  "pop": 36963,
  "state": "MA"
}
```

```
{
  "_id": "01005",
  "city": "BARRE",
  "loc": Array,
  "pop": 4546,
  "state": "MA"
}
```

```
{
  "_id": "01007",
  "city": "BELCHERTOWN",
  "loc": Array,
  "pop": 4546,
  "state": "MA"
}
```

Atlanta Population

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

```
cmd mongosh mongodb+srv://cluster0.f1gpe.mongodb.net/nehabhatter1221
```

```
Microsoft Windows [Version 10.0.22000.493]
```

```
(c) Microsoft Corporation. All rights reserved.
```

```
C:\Users\neha_\Downloads\mongosh-1.1.9-win32-x64\mongosh-1.1.9-win32-x64\bin>mongosh "mongodb+srv://cluster0.f1gpe.mongodb.net/nehabhatter1221"
nehabhatter1221
```

```
Enter password: *****
```

```
Current Mongosh Log ID: 6219b15fe07e90234cd18f63
```

```
Connecting to: mongodb+srv://cluster0.f1gpe.mongodb.net/nehabhatter1221?appName=mongosh+1.1.9
```

```
Using MongoDB: 5.0.6 (API Version 1)
```

```
Using Mongosh: 1.1.9
```

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

```
Atlas atlas-bjyw9t-shard-0 [primary] nehabhatt1221> show dbs
```

```
mongo_practice 73.7 kB
```

```
nehabhatter1221 221 kB
```

```
population 2.42 MB
```

```
test 49.2 kB
```

```
admin 340 kB
```

```
local 4.25 GB
```

```
Atlas atlas-bjyw9t-shard-0 [primary] nehabhatt1221> use population
```

```
switched to db population
```

```
Atlas atlas-bjyw9t-shard-0 [primary] population> show collections
```

```
zipcodes
```

```
Atlas atlas-bjyw9t-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',
```

mongosh mongodb+srv://cluster0.f1gpe.mongodb.net/nehabh1221

```
{
  _id: '30318',
  city: 'ATLANTA',
  loc: [ -84.445432, 33.786454 ],
  pop: 53894,
  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: '30327',
  city: 'ATLANTA',
  loc: [ -84.419966, 33.862723 ],
  pop: 18467,
  state: 'GA'
},
{
  _id: '30329',
  city: 'ATLANTA',
  loc: [ -84.321402, 33.823555 ],
  pop: 17013,
  state: 'GA'
}
}
```

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

mongosh mongodb+srv://cluster0.f1gpe.mongodb.net/nehabh1221

```
    state: 'GA'
  }
]
Type "it" for more
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate({$match:{$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'
  },
]
```

3. use \$group to count the number of zip codes in Atlanta.
4. use \$group to find the total population in Atlanta

mongosh mongodb+srv://cluster0.f1gpe.mongodb.net/nehabh1221

```
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: '30327',
  city: 'ATLANTA',
  loc: [ -84.419966, 33.862723 ],
  pop: 18467,
  state: 'GA'
},
{
  _id: '30329',
  city: 'ATLANTA',
  loc: [ -84.321402, 33.823555 ],
  pop: 17013,
  state: 'GA'
}
]
Type "it" for more
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$match:{city:"ATLANTA"}},{ $group:{_id:"city",count:{$sum:1}}}]
[ { _id: 'city', count: 41 } ]
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$match:{city:"ATLANTA"}},{ $group:{_id:"city",count:{$sum:"$pop"}}}]
[ { _id: 'city', count: 630046 } ]
Atlas atlas-bjyw9t-shard-0 [primary] population>
```

Populations By State

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

```
❏ mongosh mongodb+srv://cluster0.f1gpe.mongodb.net/nehabhatt1221
```

```
]
Type "it" for more
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$match:{city:"ATLANTA"}},{$group:{_id:"city",count:{$sum:1}}}}])
[ { _id: 'city', count: 41 } ]
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$match:{city:"ATLANTA"}},{$group:{_id:"city",count:{$sum:"$pop"}}}}])
[ { _id: 'city', count: 630046 } ]
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:"state",TotalPopulation:{$sum:"$pop"}}}}])
Uncaught:
SyntaxError: Unexpected token, expected ",", (1:30)

> 1 | db.zipcodes.aggregate([{$group:{_id:"state",TotalPopulation:{$sum:"$pop"}}}}])
    |                               ^
    2 |

Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:"state",TotalPopulation:{$sum:"$pop"}}}}])
[ { _id: 'state', TotalPopulation: 248408400 } ]
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:"state",TotPop:{$sum:"$pop"}}}}])
[ { _id: 'state', TotPop: 248408400 } ]
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:"$state",StatePop:{$sum:"$pop"}}}}])
[
  { _id: 'ND', StatePop: 638272 },
  { _id: 'MO', StatePop: 5110648 },
  { _id: 'VA', StatePop: 6181479 },
  { _id: 'OK', StatePop: 3145585 },
  { _id: 'CO', StatePop: 3293755 },
  { _id: 'AZ', StatePop: 3665228 },
  { _id: 'NM', StatePop: 1515069 },
  { _id: 'OR', StatePop: 2842321 },
  { _id: 'MD', StatePop: 4781379 },
  { _id: 'CT', StatePop: 3287116 },
  { _id: 'ID', StatePop: 1006749 },
  { _id: 'WV', StatePop: 1793146 },
  { _id: 'NY', StatePop: 17990402 },
  { _id: 'NJ', StatePop: 7730188 },
  { _id: 'NE', StatePop: 1578139 },
  { _id: 'MI', StatePop: 9295297 },
  { _id: 'TX', StatePop: 16984601 },
  { _id: 'WY', StatePop: 453528 },
  { _id: 'MA', StatePop: 6016425 },
  { _id: 'DC', StatePop: 606900 }
]
```

2. sort the results by population, highest first
3. limit the results to just the first 3 results. What are the top 3 states in population?

```
mongosh mongodb+srv://cluster0.f1gpe.mongodb.net/nehabh1221
```

```
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:"$state",StatePop:{$sum:"$pop"}}},{ $sort:{StatePop:-1}}])
[
  { _id: 'CA', StatePop: 29754890 },
  { _id: 'NY', StatePop: 17990402 },
  { _id: 'TX', StatePop: 16984601 },
  { _id: 'FL', StatePop: 12686644 },
  { _id: 'PA', StatePop: 11881643 },
  { _id: 'IL', StatePop: 11427576 },
  { _id: 'OH', StatePop: 10846517 },
  Atlas atlas-bjyw9t-shard-0 [primary] population>
Atlas atlas-bjyw9t-shard-0 [primary] population>
[
  { _id: 'NC', StatePop: 6628637 },
  { _id: 'GA', StatePop: 6478216 },
  { _id: 'VA', StatePop: 6181479 },
  { _id: 'MA', StatePop: 6016425 },
  { _id: 'IN', StatePop: 5544136 },
  { _id: 'MO', StatePop: 5110648 },
  { _id: 'WI', StatePop: 4891769 },
  { _id: 'TN', StatePop: 4876457 },
  { _id: 'WA', StatePop: 4866692 },
  { _id: 'MD', StatePop: 4781379 },
  { _id: 'MN', StatePop: 4372982 }
]
Type "it" for more
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:"$state",StatePop:{$sum:"$pop"}}},{ $sort:{StatePop:-1}},{ $limit:3}])
[
  { _id: 'CA', StatePop: 29754890 },
  { _id: 'NY', StatePop: 17990402 },
  { _id: 'TX', StatePop: 16984601 }
]
Atlas atlas-bjyw9t-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' }


```
❏ mongosh mongodb+srv://cluster0.f1gpe.mongodb.net/nehabh1221
```

```
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:{city:"$city",state:"$state"},population:{$sum:"$pop"}}}])
```

```
[
  { _id: { city: 'ISMAY' }, population: 317 },
  { _id: { city: 'SHARON HILL' }, population: 9893 },
  { _id: { city: 'VANLEER' }, population: 1531 },
  { _id: { city: 'BECCARIA' }, population: 1571 },
  { _id: { city: 'WILLSALL' }, population: 1293 },
  { _id: { city: 'GLENMONT' }, population: 6206 },
  { _id: { city: 'ELKO' }, population: 5080 },
  { _id: { city: 'TENNYSON' }, population: 1378 },
  { _id: { city: 'FAIRLAND' }, population: 8685 },
  { _id: { city: 'GRANGERLAND' }, population: 5767 },
  { _id: { city: 'CHERRY HILLS VIL' }, population: 60456 },
  { _id: { city: 'JIM FALLS' }, population: 974 },
  { _id: { city: 'BRINSMADE' }, population: 168 },
  { _id: { city: 'STILLWELL' }, population: 6761 },
  { _id: { city: 'MONETTA' }, population: 2068 },
  { _id: { city: 'DAY' }, population: 1567 },
  { _id: { city: 'DELMAR' }, population: 24398 },
  { _id: { city: 'MC CONNELLS' }, population: 1431 },
  { _id: { city: 'RURAL VALLEY' }, population: 3499 },
  { _id: { city: 'SYOSSET' }, population: 24949 }
]
```

```
Type "it" for more
```

```
Atlas atlas-bjyw9t-shard-0 [primary] population> _
```

2. sort the results by population, highest first

```
❏ mongosh mongodb+srv://cluster0.f1gpe.mongodb.net/nehabhatt1221
```

```
nehabhatt1221
```

```
Enter password: *****
```

```
Current Mongosh Log ID: 6219c1cc8b7950b1c65b3b90
```

```
Connecting to:      mongodb+srv://cluster0.f1gpe.mongodb.net/nehabhatt1221?appName=mongosh+1.1.9
```

```
Using MongoDB:      5.0.6 (API Version 1)
```

```
Using Mongosh:      1.1.9
```

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

```
Atlas atlas-bjyw9t-shard-0 [primary] nehabhatt1221> show dbs
```

```
mongo_practice  73.7 kB
```

```
nehabhatt1221   221 kB
```

```
population      2.42 MB
```

```
test            49.2 kB
```

```
admin           340 kB
```

```
local           4.25 GB
```

```
Atlas atlas-bjyw9t-shard-0 [primary] nehabhatt1221> use population
```

```
switched to db population
```

```
Atlas atlas-bjyw9t-shard-0 [primary] population> show collections
```

```
zipcodes
```

```
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:{city:"$city",state:"$state"},population:{$sum:"$pop"}}},{ $sort:{population:
```

```
[
```

```
{ _id: { city: 'CHICAGO' }, population: 2452177 },
```

```
{ _id: { city: 'BROOKLYN' }, population: 2341387 },
```

```
{ _id: { city: 'HOUSTON' }, population: 2123053 },
```

```
{ _id: { city: 'LOS ANGELES' }, population: 2102295 },
```

```
{ _id: { city: 'PHILADELPHIA' }, population: 1639862 },
```

```
{ _id: { city: 'NEW YORK' }, population: 1476790 },
```

```
{ _id: { city: 'BRONX' }, population: 1209548 },
```

```
{ _id: { city: 'SAN DIEGO' }, population: 1054316 },
```

```
{ _id: { city: 'DALLAS' }, population: 999042 },
```

```
{ _id: { city: 'DETROIT' }, population: 967468 },
```

```
{ _id: { city: 'PHOENIX' }, population: 902249 },
```

```
{ _id: { city: 'MIAMI' }, population: 848436 },
```

```
{ _id: { city: 'COLUMBUS' }, population: 825448 },
```

```
{ _id: { city: 'SAN JOSE' }, population: 817497 },
```

```
{ _id: { city: 'SAN ANTONIO' }, population: 813188 },
```

```
{ _id: { city: 'WASHINGTON' }, population: 780954 },
```

```
{ _id: { city: 'BALTIMORE' }, population: 738846 },
```

```
{ _id: { city: 'JACKSONVILLE' }, population: 735505 },
```

```
{ _id: { city: 'SAN FRANCISCO' }, population: 723993 },
```

```
{ _id: { city: 'CLEVELAND' }, population: 687451 }
```

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

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

```
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:{city:"$city",state:"$state"},population:{$sum:"$pop"}}},{$sort:{population:-1}},{$limit:3}])
[
  { _id: { city: 'CHICAGO' }, population: 2452177 },
  { _id: { city: 'BROOKLYN' }, population: 2341387 },
  { _id: { city: 'HOUSTON' }, population: 2123053 }
]
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$match:{state:"TX"}},{$population:{$sum:"$pop"}},{$sort:{population:-1}},{$limit:3}])
Uncaught:
SyntaxError: Unexpected token, expected ",", (1:72)

> 1 | db.zipcodes.aggregate([{$match:{state:"TX"}},{$population:{$sum:"$pop"}},{$sort:{population:-1}},{$limit:3}])
    |
    2 |
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$match:{state:"TX"}},{$group:{_id:{city:"$city",state:"$state"},population:{$sum:"$pop"}},{$sort:{population:-1}},{$limit:3}])
[
  { _id: { city: 'HOUSTON', state: 'TX' }, population: 2095918 },
  { _id: { city: 'DALLAS', state: 'TX' }, population: 940191 },
  { _id: { city: 'SAN ANTONIO', state: 'TX' }, population: 811792 }
]
Atlas atlas-bjyw9t-shard-0 [primary] population>
```

Bonus

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

mongosh mongodb+srv://cluster0.f1gpe.mongodb.net/nehabh1221

```
{ _id: 'MI', AveragePopulation: 10611.069634703197 },
{ _id: 'OH', AveragePopulation: 10771.119165839125 },
{ _id: 'SC', AveragePopulation: 9962.00857142857 },
{ _id: 'MA', AveragePopulation: 12692.879746835442 },
{ _id: 'DC', AveragePopulation: 25287.5 },
{ _id: 'RI', AveragePopulation: 14539.391304347826 },
{ _id: 'TN', AveragePopulation: 8378.792096219931 },
{ _id: 'KY', AveragePopulation: 4543.243510506799 },
{ _id: 'IL', AveragePopulation: 9238.137429264349 },
{ _id: 'IA', AveragePopulation: 3011.301518438178 },
{ _id: 'CO', AveragePopulation: 7955.929951690821 },
{ _id: 'VA', AveragePopulation: 7575.341911764706 },
{ _id: 'MO', AveragePopulation: 5141.496981891348 },
{ _id: 'OR', AveragePopulation: 7401.877604166667 },
{ _id: 'NM', AveragePopulation: 5489.380434782609 },
{ _id: 'CT', AveragePopulation: 12498.539923954373 },
{ _id: 'WD', AveragePopulation: 11384.235714285714 },
{ _id: 'ID', AveragePopulation: 4126.020491803279 }
]
Type "it" for more
Atlas atlas-bjyw9t-shard-0 [primary] population> db.zipcodes.aggregate([{$group:{_id:"$city",AveragePopulation:{$avg:"$pop"}}}])
[
  { _id: 'LANSDALE', AveragePopulation: 41034 },
  { _id: 'HEALDSBURG', AveragePopulation: 16884 },
  { _id: 'STIOUX RAPIDS', AveragePopulation: 1542 },
  { _id: 'REESEVILLE', AveragePopulation: 2027 },
  { _id: 'THIBODAUX', AveragePopulation: 37831 },
  { _id: 'TONAWANDA', AveragePopulation: 53410 },
  { _id: 'WOLF CREEK', AveragePopulation: 995 },
  { _id: 'CASEY CREEK', AveragePopulation: 1102 },
  { _id: 'LUMBER CITY', AveragePopulation: 2036 },
  { _id: 'CLEAR LAKE SHORE', AveragePopulation: 4101 },
  { _id: 'CAMPBELLTON', AveragePopulation: 608 },
  { _id: 'GREENBUSH', AveragePopulation: 988 },
  { _id: 'FABER', AveragePopulation: 958 },
  { _id: 'GROVEOAK', AveragePopulation: 1098 },
  { _id: 'KISSEE MILLS', AveragePopulation: 826 },
  { _id: 'GOWEN CITY', AveragePopulation: 646 },
  { _id: 'STONEY FORK', AveragePopulation: 387 },
  { _id: 'COLMAR', AveragePopulation: 1796.5 },
  { _id: 'GLASSPORT', AveragePopulation: 5582 },
  { _id: 'SEBEWAING', AveragePopulation: 3203 }
```

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

