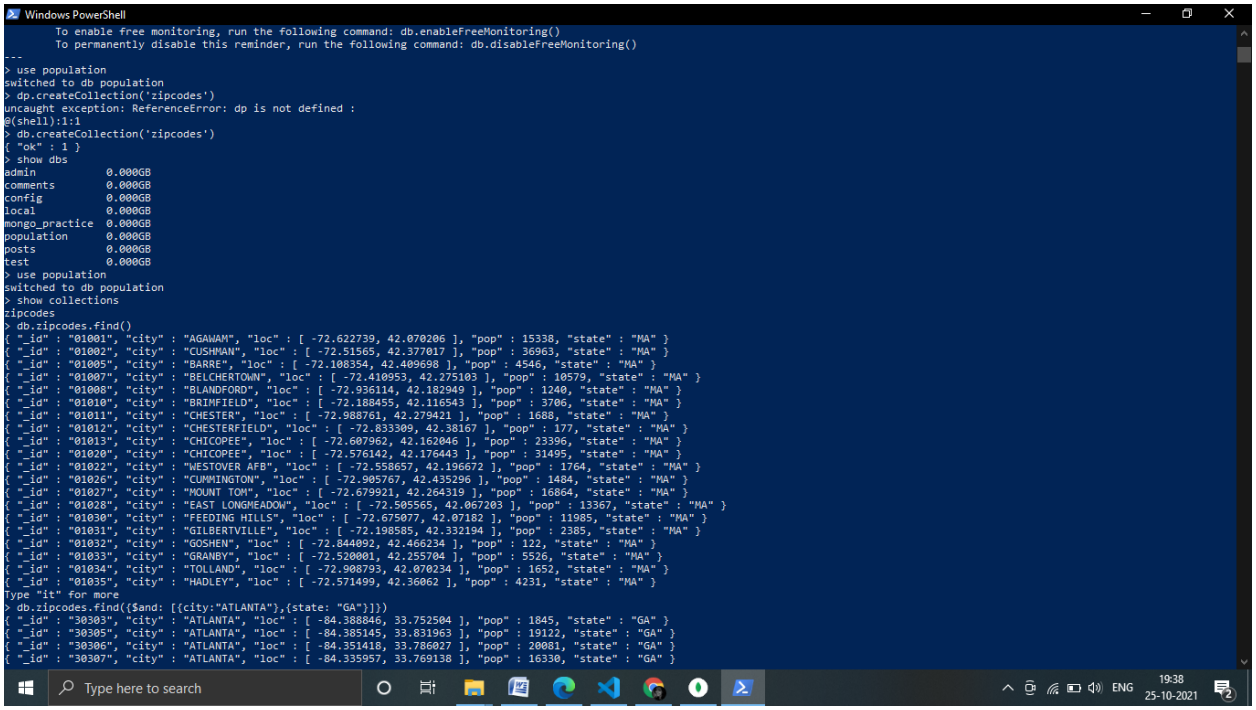
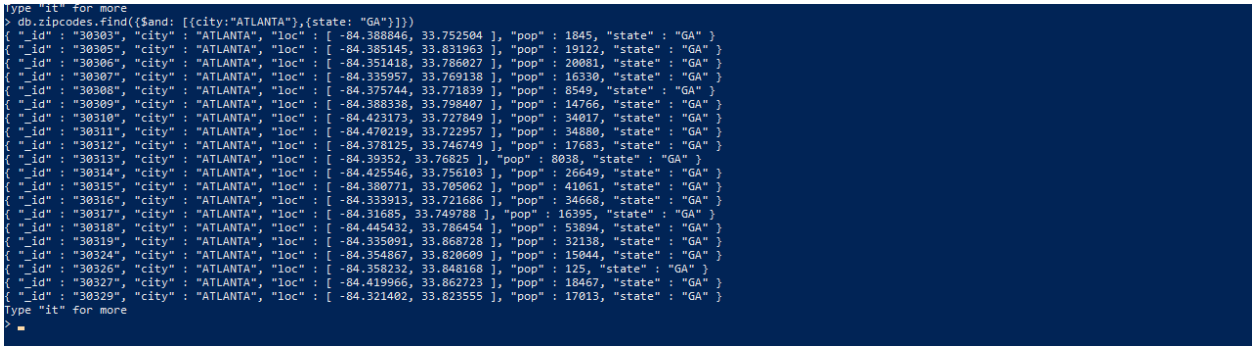
ASSIGNMENT2:

MongoDB -Aggregation Exercises

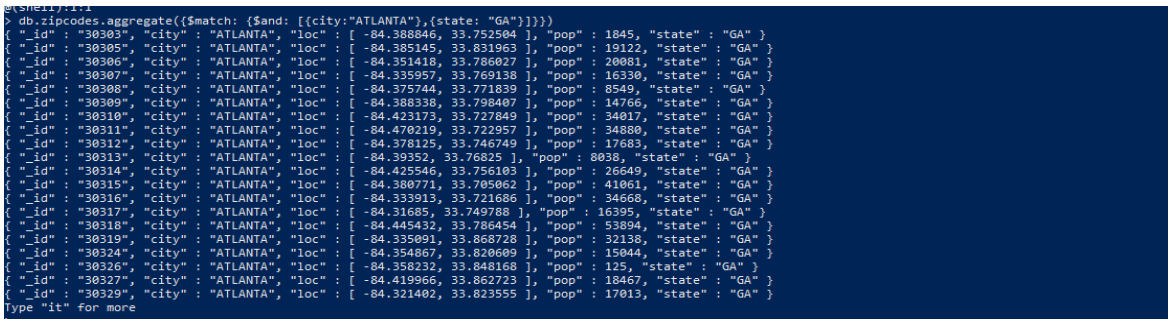


Atlanta Population:

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



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



3. use $group to count the number of zip codes in Atlanta.

```
{ "_id" : "01034", "city" : "TOLLAND", "loc" : [ -72.908793, 42.070234 ], "pop" : 1652, "state" : "MA" }
{ "_id" : "01035", "city" : "HADLEY", "loc" : [ -72.571499, 42.36062 ], "pop" : 4231, "state" : "MA" }
Type "it" for more
> db.zipcodes.aggregate({$match:{city:"ATLANTA"}
... }, {$group:{_id:"$city",
...  zipcodes:{$sum:1}}})
{ "_id" : "ATLANTA", "zipcodes" : 41 }
>
```

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

```
{ "_id" : "ATLANTA", "zipcodes" : 41 }
> db.zipcodes.aggregate([{$group: {_id: "$ATLANTA",count: {$sum:"$pop"}}}])
{ "_id" : null, "count" : 248408400 }
>
```

Populations By State

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

```
>
> db.zipcodes.aggregate({$group:{_id:{state:'$state'},pop:{$sum:'$pop'}}})
{ "_id" : { "state" : "IL" }, "pop" : 11427576 }
{ "_id" : { "state" : "KS" }, "pop" : 2475285 }
{ "_id" : { "state" : "MT" }, "pop" : 798948 }
{ "_id" : { "state" : "ME" }, "pop" : 1226648 }
{ "_id" : { "state" : "IN" }, "pop" : 5544136 }
{ "_id" : { "state" : "NY" }, "pop" : 17990402 }
{ "_id" : { "state" : "CT" }, "pop" : 3287116 }
{ "_id" : { "state" : "WI" }, "pop" : 4891769 }
{ "_id" : { "state" : "NE" }, "pop" : 1578139 }
{ "_id" : { "state" : "MD" }, "pop" : 4781379 }
{ "_id" : { "state" : "GA" }, "pop" : 6478216 }
{ "_id" : { "state" : "AL" }, "pop" : 4040587 }
{ "_id" : { "state" : "AR" }, "pop" : 2350725 }
{ "_id" : { "state" : "UT" }, "pop" : 1722850 }
{ "_id" : { "state" : "HI" }, "pop" : 1108229 }
{ "_id" : { "state" : "RI" }, "pop" : 1003218 }
{ "_id" : { "state" : "CO" }, "pop" : 3293755 }
{ "_id" : { "state" : "MO" }, "pop" : 5110648 }
{ "_id" : { "state" : "AZ" }, "pop" : 3665228 }
{ "_id" : { "state" : "WA" }, "pop" : 4866692 }
Type "it" for more
```

2.sort the results by population, highest first

```
> db.zipcodes.aggregate({$group:{ _id:{state:'$state'},pop:{$sum:'$pop'}}},{$sort:{population:-1}})
{ "_id" : { "state" : "IL" }, "pop" : 11427576 }
{ "_id" : { "state" : "KS" }, "pop" : 2475285 }
{ "_id" : { "state" : "MT" }, "pop" : 798948 }
{ "_id" : { "state" : "ME" }, "pop" : 1226648 }
{ "_id" : { "state" : "IN" }, "pop" : 5544136 }
{ "_id" : { "state" : "NY" }, "pop" : 17990402 }
{ "_id" : { "state" : "CT" }, "pop" : 3287116 }
{ "_id" : { "state" : "WI" }, "pop" : 4891769 }
{ "_id" : { "state" : "NE" }, "pop" : 1578139 }
{ "_id" : { "state" : "MD" }, "pop" : 4781379 }
{ "_id" : { "state" : "GA" }, "pop" : 6478216 }
{ "_id" : { "state" : "AL" }, "pop" : 4040587 }
{ "_id" : { "state" : "AR" }, "pop" : 2350725 }
{ "_id" : { "state" : "UT" }, "pop" : 1722850 }
{ "_id" : { "state" : "HI" }, "pop" : 1108229 }
{ "_id" : { "state" : "RI" }, "pop" : 1003218 }
{ "_id" : { "state" : "CO" }, "pop" : 3293755 }
{ "_id" : { "state" : "MO" }, "pop" : 5110648 }
{ "_id" : { "state" : "AZ" }, "pop" : 3665228 }
{ "_id" : { "state" : "WA" }, "pop" : 4866692 }
Type "it" for more
```

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:{$max:"$pop"}}},{$sort:{state:-1}},{$limit: 3}])
{ "_id" : "IL", "state" : 112047 }
{ "_id" : "NY", "state" : 111396 }
{ "_id" : "CA", "state" : 99568 }
```

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' } 2. sort the results by population, highest first 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?

```
> db.zipcodes.aggregate({$group:{_id:{city:'$city',state:'$state'},population:{$sum:'$pop'}}})
{ "_id" : { "city" : "LOMAN", "state" : "MN" }, "population" : 198 }
{ "_id" : { "city" : "MONTICELLO", "state" : "IN" }, "population" : 13599 }
{ "_id" : { "city" : "PYRMID", "state" : "KY" }, "population" : 568 }
{ "_id" : { "city" : "ROCHESTER", "state" : "NH" }, "population" : 3793 }
{ "_id" : { "city" : "MADEIRA", "state" : "OH" }, "population" : 14999 }
{ "_id" : { "city" : "MIDDLEPORT", "state" : "OH" }, "population" : 3433 }
{ "_id" : { "city" : "GARRETT", "state" : "IN" }, "population" : 6459 }
{ "_id" : { "city" : "PISEK", "state" : "ND" }, "population" : 350 }
{ "_id" : { "city" : "SAVAGE", "state" : "MN" }, "population" : 7764 }
{ "_id" : { "city" : "YAKIMA", "state" : "WA" }, "population" : 39091 }
{ "_id" : { "city" : "KEARNEYSVILLE", "state" : "WV" }, "population" : 3468 }
{ "_id" : { "city" : "BOGART", "state" : "GA" }, "population" : 7245 }
{ "_id" : { "city" : "EL MONTE", "state" : "CA" }, "population" : 84237 }
{ "_id" : { "city" : "MONROE", "state" : "SD" }, "population" : 239 }
{ "_id" : { "city" : "ARBON", "state" : "ID" }, "population" : 121 }
{ "_id" : { "city" : "FARMINGTON", "state" : "NM" }, "population" : 44020 }
{ "_id" : { "city" : "EAGLE BAY", "state" : "NY" }, "population" : 1416 }
{ "_id" : { "city" : "CENTREVILLE", "state" : "MI" }, "population" : 3606 }
{ "_id" : { "city" : "MEADVIEW", "state" : "AZ" }, "population" : 118 }
{ "_id" : { "city" : "KIRBY", "state" : "WV" }, "population" : 303 }
Type "it" for more
>
```

2. sort the results by population, highest first

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

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:{$max:"$pop"}}},{$sort:{city:-1}},{$limit: 3}])
{ "_id" : "CHICAGO", "city" : 112047 }
{ "_id" : "BROOKLYN", "city" : 111396 }
{ "_id" : "NEW YORK", "city" : 106564 }
>
```

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

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

Bonus

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

```
> db.zipcodes.aggregate({$group:{_id:{city:'$city', state:'$state'},population:{$avg:'$pop'}}},{$sort:{population:-1}})
{ "_id" : { "city" : "BELL GARDENS", "state" : "CA" }, "population" : 99568 }
{ "_id" : { "city" : "NORWALK", "state" : "CA" }, "population" : 94188 }
{ "_id" : { "city" : "ARLETA", "state" : "CA" }, "population" : 88114 }
{ "_id" : { "city" : "SOUTH GATE", "state" : "CA" }, "population" : 87026 }
{ "_id" : { "city" : "RIDGEWOOD", "state" : "NY" }, "population" : 85732 }
{ "_id" : { "city" : "WESTLAND", "state" : "MI" }, "population" : 84712 }
{ "_id" : { "city" : "HOLLY PARK", "state" : "CA" }, "population" : 78511 }
{ "_id" : { "city" : "WESTMINSTER", "state" : "CA" }, "population" : 77965 }
{ "_id" : { "city" : "INDUSTRY", "state" : "CA" }, "population" : 77114 }
{ "_id" : { "city" : "COAST GUARD ISLA", "state" : "CA" }, "population" : 76110 }
{ "_id" : { "city" : "GWYNN OAK", "state" : "MD" }, "population" : 76002 }
{ "_id" : { "city" : "WOODSIDE", "state" : "NY" }, "population" : 75894 }
{ "_id" : { "city" : "CORONA", "state" : "NY" }, "population" : 75746 }
{ "_id" : { "city" : "GLEN BURNIE", "state" : "MD" }, "population" : 75692 }
{ "_id" : { "city" : "RIALTO", "state" : "CA" }, "population" : 75341 }
{ "_id" : { "city" : "JACKSON HEIGHTS", "state" : "NY" }, "population" : 72983.5 }
{ "_id" : { "city" : "HUNTINGTON PARK", "state" : "CA" }, "population" : 72139 }
{ "_id" : { "city" : "LAKE LOS ANGELES", "state" : "CA" }, "population" : 71024 }
{ "_id" : { "city" : "TAYLOR", "state" : "MI" }, "population" : 70811 }
{ "_id" : { "city" : "STAR CITY", "state" : "WV" }, "population" : 70185 }
```

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

```
> db.zipcodes.aggregate([{$group:{"_id":"$state", state:{$avg:"$pop"}}},{$sort:{state:-1}},{$limit:3}])
{ "_id" : "DC", "state" : 25287.5 }
{ "_id" : "CA", "state" : 19627.236147757256 }
{ "_id" : "FL", "state" : 15779.407960199005 }
>
```