**MongoDB -Aggregation Exercises**

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

* **mongoimport --db population --collection zipcodes --file <filepath>/zips.json**

**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'}]})**

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

* **db.zipcodes.aggregate([{$match:{city:'ATLANTA',state:'GA'}}])**

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

* **db.zipcodes.aggregate([{$match:{city:'ATLANTA'}},{$group:{\_id:"$city",count:{$sum:1}}])**

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

* **db.zipcodes.aggregate([{$match:{city:'ATLANTA'}},{$group:{\_id:"$city","Total Population":{$sum:"$pop"}}}])**

**Populations By State**

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

* **db.zipcodes.aggregate([{$group:{ \_id: "$state", totalPop:{ $sum:"$pop" }}}])**

1. sort the results by population, highest first

* **db.zipcodes.aggregate({$group:{\_id:{state:"$state"},population:{$sum:"$pop"}}},{$sort:{population:-1}})**

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

* **db.zipcodes.aggregate({$group:{\_id:{state:"$state"},population:{$sum:"$pop"}}},{$sort:{population:-1}},{$limit:3})**
* TOP 3 STATES IN POPULATION:

{ "\_id" : { "state" : "**CA**" }, "population" : 29754890 }

{ "\_id" : { "state" : "**NY**" }, "population" : 17990402 }

{ "\_id" : { "state" : "**TX**" }, "population" : 16984601 }

**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' }, totalPop:{ $sum:"$pop" }}}])**

1. sort the results by population, highest first

* **db.zipcodes.aggregate({$group:{\_id:{city:"$city",state:"$state"},population:{$sum:"$pop"}}},{$sort:{population:-1}})**

1. 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:"$state"},population:{$sum:"$pop"}}},{$sort:{population:-1}},{$limit:3})**
* TOP 3 CITIES IN POPULATION:

{ "\_id" : { "city" : "**CHICAGO**", "state" : "IL" }, "population" : 2452177 }

{ "\_id" : { "city" : "**BROOKLYN**", "state" : "NY" }, "population" : 2300504 }

{ "\_id" : { "city" : "**LOS** **ANGELES**", "state" : "CA" }, "population" : 2102295 }

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

* **db.zipcodes.aggregate({$group:{\_id:{city:"$city",state:"Texas"},population:{$sum:"$pop"}}},{$sort:{population:-1}},{$limit:3})**
* TOP 3 CITIES IN POPULATION IN TEXAS:

{ "\_id" : { "city" : "**CHICAGO**", "state" : "Texas" }, "population" : 2452177 }

{ "\_id" : { "city" : "**BROOKLYN**", "state" : "Texas" }, "population" : 2341387 }

{ "\_id" : { "city" : "**HOUSTON**", "state" : "Texas" }, "population" : 2123053 }

**Bonus**

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

* **db.zipcodes.aggregate( [ { $group: { \_id: { state: "$state", city: "$city" }, pop: { $sum: "$pop" } } }, { $group: { \_id: "$\_id.state", avgCityPop: { $avg: "$pop" } } }] )**

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

* **db.zipcodes.aggregate( [ { $group: { \_id: { state: "$state", city: "$city" }, pop: { $sum: "$pop" } } },{$sort:{pop:-1}}, { $group: { \_id: "$\_id.state", avgCityPop: { $avg: "$pop" } } },{$limit:3}] )**