

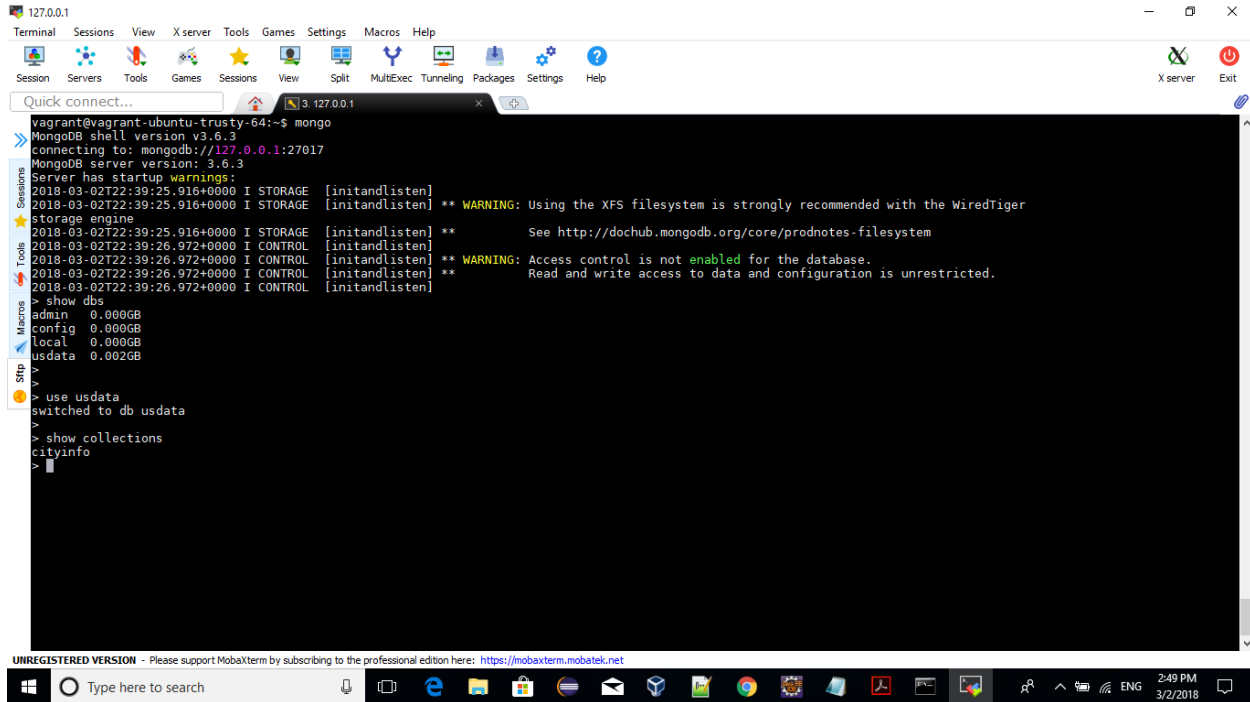
Task 1 - wget <http://media.mongodb.org/zips.json>

Task 2- mongoimport -d usdata -c cityinfo --file /home/vagrant/zips.json

Task 3 – Screen 1 - shows all collections of the database usdata

use usdata

show collections



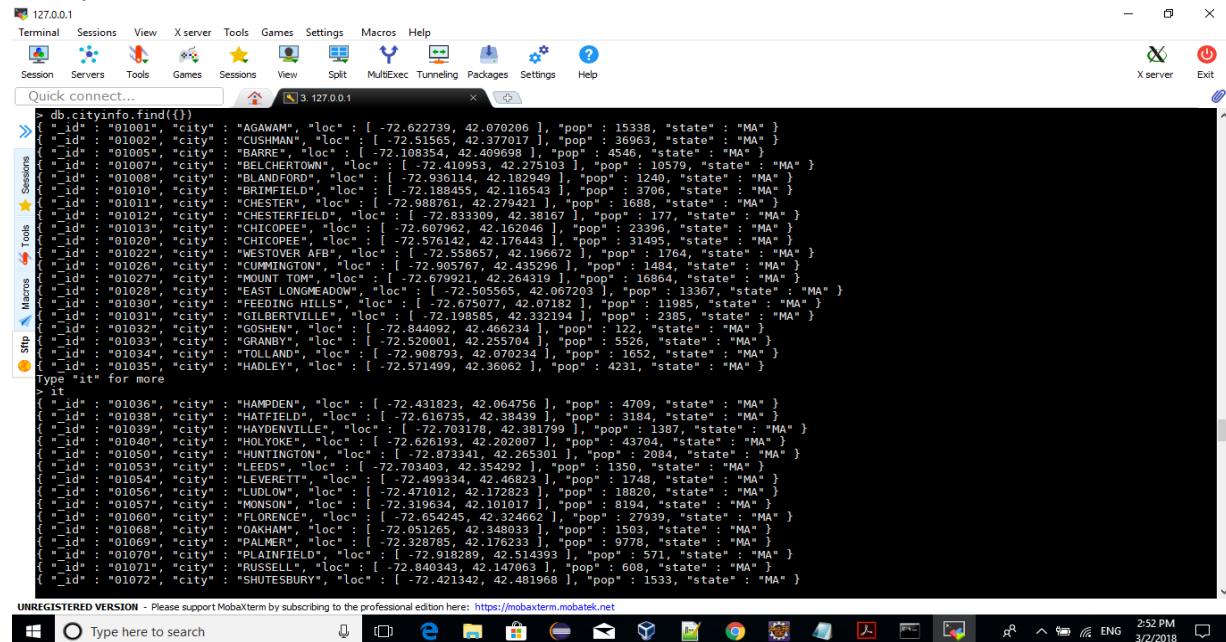
The screenshot shows a MobaXterm terminal window titled '127.0.0.1'. The terminal displays the following content:

```
vagrant@vagrant-ubuntu-trusty-64:~$ mongo
MongoDB shell version v3.6.3
connecting to: mongodb://127.0.0.1:27017
MongoDB server version: 3.6.3
Server has startup warnings:
2018-03-02T22:39:25.916+0000 I STORAGE [initandlisten] ** WARNING: Using the XFS filesystem is strongly recommended with the WiredTiger
storage engine
2018-03-02T22:39:25.916+0000 I STORAGE [initandlisten] ** See http://dochub.mongodb.org/core/prodnotes-filesystem
2018-03-02T22:39:26.972+0000 I CONTROL [initandlisten] ** WARNING: Access control is not enabled for the database.
2018-03-02T22:39:26.972+0000 I CONTROL [initandlisten] ** Read and write access to data and configuration is unrestricted.
2018-03-02T22:39:26.972+0000 I CONTROL [initandlisten]
> show dbs
admin 0.000GB
config 0.000GB
local 0.000GB
usdata 0.002GB
> use usdata
switched to db usdata
> show collections
cityinfo
```

The terminal window has a sidebar with icons for Sessions, Servers, Tools, Games, Sessions, View, Split, MultiExec, Tunneling, Packages, Settings, and Help. The bottom status bar shows 'UNREGISTERED VERSION - Please support MobaXterm by subscribing to the professional edition here: <https://mobaxterm.mobatek.net>' and the system clock '2:49 PM 3/2/2018'.

## Task 4 – Screen 2- Find all documents of cityinfo collection.

db.cityinfo.find()

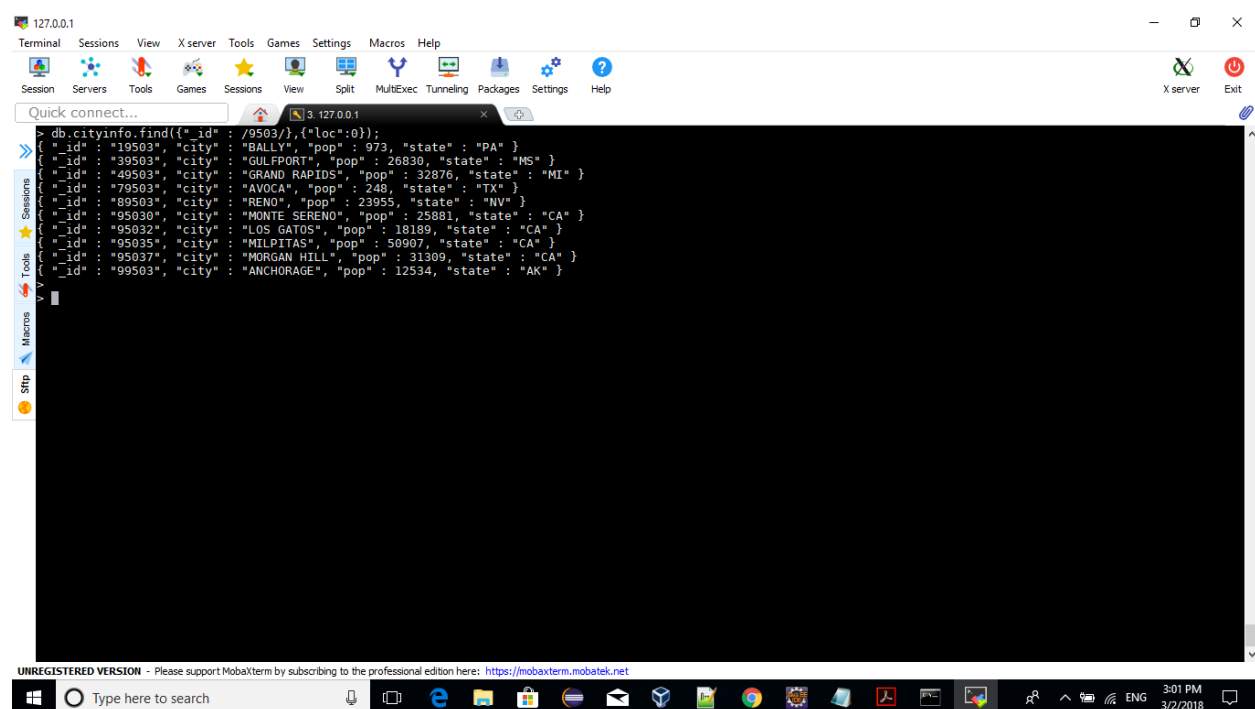


The screenshot shows a MobaXterm terminal window with the command `db.cityinfo.find()` entered. The output displays a list of city documents from the 'cityinfo' collection, each containing fields like `_id`, `city`, `loc`, `pop`, and `state`. The cities listed include AGAWAM, CUSHMAN, BARRE, BELCHERTOWN, BLANDFORD, BRIMFIELD, CHESTER, CHESTERFIELD, CHICOPPEE, CHICOPPEE, WESTOVER AFB, CUMMINGTON, MOUNT TOM, EAST LONGMEADOW, FEEDING HILLS, GILBERTVILLE, GOSHEN, GRANBY, TOLLAND, HADLEY, HAMPDEN, HATFIELD, HAYDENVILLE, HOLYOKE, HUNTINGTON, LEEDS, LEVERETT, LUDLOW, MONSON, FLORENCE, OAKHAM, PALMER, PLAINFIELD, RUSSELL, and SHUTESBURY. The terminal window has a title bar '127.0.0.1' and a menu bar with options like Terminal, Sessions, View, X server, Tools, Games, Settings, Macros, and Help. A 'Quick connect...' dialog is visible at the top.

```
db.cityinfo.find()
{"_id": "01001", "city": "AGAWAM", "loc": [ -72.622739, 42.070206 ], "pop": 15338, "state": "MA" }
{"_id": "01002", "city": "CUSHMAN", "loc": [ -72.51565, 42.377017 ], "pop": 36963, "state": "MA" }
{"_id": "01005", "city": "BARRE", "loc": [ -72.108354, 42.489698 ], "pop": 4546, "state": "MA" }
{"_id": "01007", "city": "BELCHERTOWN", "loc": [ -72.410953, 42.275163 ], "pop": 10579, "state": "MA" }
{"_id": "01008", "city": "BLANDFORD", "loc": [ -72.936114, 42.182949 ], "pop": 1240, "state": "MA" }
{"_id": "01010", "city": "BRIMFIELD", "loc": [ -72.188455, 42.116543 ], "pop": 3706, "state": "MA" }
{"_id": "01011", "city": "CHESTER", "loc": [ -72.988761, 42.279421 ], "pop": 1688, "state": "MA" }
{"_id": "01012", "city": "CHESTERFIELD", "loc": [ -72.833309, 42.38167 ], "pop": 177, "state": "MA" }
{"_id": "01013", "city": "CHICOPPEE", "loc": [ -72.607962, 42.162046 ], "pop": 23396, "state": "MA" }
{"_id": "01020", "city": "CHICOPPEE", "loc": [ -72.576142, 42.176443 ], "pop": 31495, "state": "MA" }
{"_id": "01022", "city": "WESTOVER AFB", "loc": [ -72.558657, 42.196672 ], "pop": 1764, "state": "MA" }
{"_id": "01026", "city": "CUMMINGTON", "loc": [ -72.905767, 42.435296 ], "pop": 1484, "state": "MA" }
{"_id": "01027", "city": "MOUNT TOM", "loc": [ -72.679921, 42.264319 ], "pop": 16864, "state": "MA" }
{"_id": "01028", "city": "EAST LONGMEADOW", "loc": [ -72.505565, 42.067203 ], "pop": 13367, "state": "MA" }
{"_id": "01030", "city": "FEEDING HILLS", "loc": [ -72.675977, 42.07182 ], "pop": 11985, "state": "MA" }
{"_id": "01031", "city": "GILBERTVILLE", "loc": [ -72.198585, 42.332194 ], "pop": 2385, "state": "MA" }
{"_id": "01032", "city": "GOSHEN", "loc": [ -72.844092, 42.468234 ], "pop": 122, "state": "MA" }
{"_id": "01033", "city": "GRANBY", "loc": [ -72.520801, 42.255704 ], "pop": 5526, "state": "MA" }
{"_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" }
{"_id": "01036", "city": "HAMPDEN", "loc": [ -72.431823, 42.064756 ], "pop": 4709, "state": "MA" }
{"_id": "01038", "city": "HATFIELD", "loc": [ -72.616735, 42.38439 ], "pop": 3184, "state": "MA" }
{"_id": "01039", "city": "HAYDENVILLE", "loc": [ -72.703178, 42.381799 ], "pop": 1387, "state": "MA" }
{"_id": "01040", "city": "HOLYOKE", "loc": [ -72.626193, 42.202007 ], "pop": 43704, "state": "MA" }
{"_id": "01050", "city": "HUNTINGTON", "loc": [ -72.873241, 42.265301 ], "pop": 2084, "state": "MA" }
{"_id": "01053", "city": "LEEDS", "loc": [ -72.703403, 42.354292 ], "pop": 1350, "state": "MA" }
{"_id": "01054", "city": "LEVERETT", "loc": [ -72.499334, 42.46823 ], "pop": 1748, "state": "MA" }
{"_id": "01056", "city": "LUDLOW", "loc": [ -72.471012, 42.172823 ], "pop": 18820, "state": "MA" }
{"_id": "01057", "city": "MONSON", "loc": [ -72.319634, 42.101017 ], "pop": 8194, "state": "MA" }
{"_id": "01060", "city": "FLORENCE", "loc": [ -72.654245, 42.324662 ], "pop": 27939, "state": "MA" }
{"_id": "01068", "city": "OAKHAM", "loc": [ -72.051265, 42.348033 ], "pop": 1503, "state": "MA" }
{"_id": "01069", "city": "PALMER", "loc": [ -72.328785, 42.176233 ], "pop": 9778, "state": "MA" }
{"_id": "01070", "city": "PLAINFIELD", "loc": [ -72.918289, 42.514393 ], "pop": 571, "state": "MA" }
{"_id": "01071", "city": "RUSSELL", "loc": [ -72.840343, 42.147063 ], "pop": 608, "state": "MA" }
{"_id": "01072", "city": "SHUTESBURY", "loc": [ -72.421342, 42.481968 ], "pop": 1533, "state": "MA" }
```

## Task 5 – Screen 3 - Find all documents with \_id that contains 9503 in it. Do not include "loc" in the output. For example, expected documents in the output may include a document with "\_id": "19503" and a document with "\_id": "95037".

db.cityinfo.find({"\_id" : /9503/},{ "loc":0})

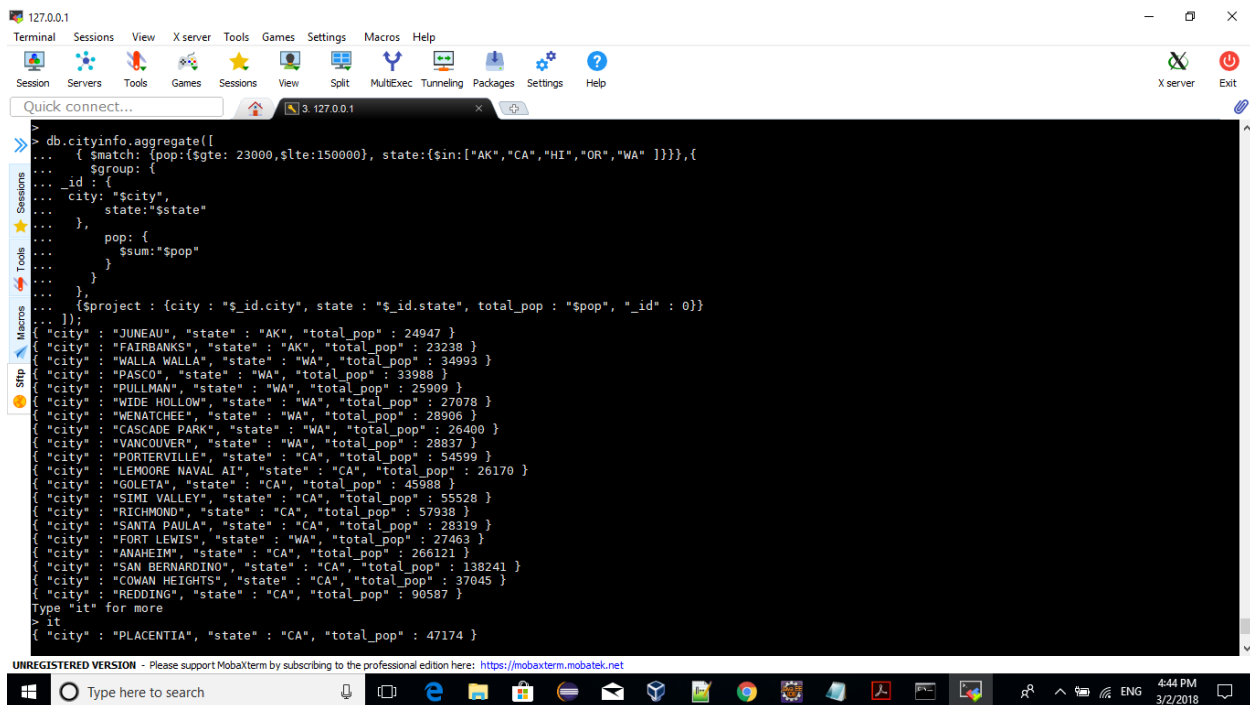


The screenshot shows a MobaXterm terminal window with the command `db.cityinfo.find({'_id' : /9503/},{ 'loc':0})` entered. The output displays a list of city documents where the 'loc' field is excluded. The cities listed include BALLY, GULFPORT, GRAND RAPIDS, AVOCADO, RENO, MONTE SERENO, LOS GATOS, MILPITAS, MORGAN HILL, and ANCHORAGE. The terminal window has a title bar '127.0.0.1' and a menu bar with options like Terminal, Sessions, View, X server, Tools, Games, Settings, Macros, and Help. A 'Quick connect...' dialog is visible at the top.

```
db.cityinfo.find({'_id' : /9503/},{ 'loc':0})
{"_id": "19503", "city": "BALLY", "pop": 973, "state": "PA" }
{"_id": "39503", "city": "GULFPORT", "pop": 26830, "state": "MS" }
{"_id": "49503", "city": "GRAND RAPIDS", "pop": 32876, "state": "MI" }
{"_id": "79503", "city": "AVOCADO", "pop": 248, "state": "TX" }
{"_id": "89503", "city": "RENO", "pop": 23955, "state": "NV" }
{"_id": "95030", "city": "MONTE SERENO", "pop": 25881, "state": "CA" }
{"_id": "95032", "city": "LOS GATOS", "pop": 18189, "state": "CA" }
{"_id": "95035", "city": "MILPITAS", "pop": 50907, "state": "CA" }
{"_id": "95037", "city": "MORGAN HILL", "pop": 31309, "state": "CA" }
{"_id": "99503", "city": "ANCHORAGE", "pop": 12534, "state": "AK" }
```

Task 6 -Screen 4 - Find all cities with populations between 23,000 and 150,000 where the state they are in borders the Pacific Ocean. In the output, a city should be unique for a given state and it should come with the total population of the city. For example, there should be one line of output for LOS ANGELES, CA with total population of LOS ANGELES.

```
db.cityinfo.aggregate([
  { $match: { pop: { $gte: 23000, $lte: 150000 }, state: { $in: ["AK", "CA", "HI", "OR", "WA" ] } } }, {
    $group: { _id : { city: "$city", state: "$state" }, pop: { $sum: "$pop" } } },
  { $project: { city : "$_id.city", state : "$_id.state", total_pop : "$pop", "_id" : 0 } }
]);
```



```
127.0.0.1
Terminal Sessions View X server Tools Games Settings Macros Help
Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help
Quick connect... 3 127.0.0.1
> db.cityinfo.aggregate([
... { $match: { pop: { $gte: 23000, $lte: 150000 }, state: { $in: ["AK", "CA", "HI", "OR", "WA" ] } } }, {
...   $group: {
...     _id : {
...       city: "$city",
...       state: "$state"
...     },
...     pop: {
...       $sum: "$pop"
...     }
...   },
...   { $project: { city : "$_id.city", state : "$_id.state", total_pop : "$pop", "_id" : 0 } }
... ]});
{"city": "JUNEAU", "state": "AK", "total_pop": 24947 }
{"city": "FAIRBANKS", "state": "AK", "total_pop": 23238 }
{"city": "WALLA WALLA", "state": "WA", "total_pop": 34993 }
{"city": "PASCO", "state": "WA", "total_pop": 33988 }
{"city": "PULLMAN", "state": "WA", "total_pop": 25909 }
{"city": "WIDE HOLLOW", "state": "WA", "total_pop": 27078 }
{"city": "MENACHEE", "state": "WA", "total_pop": 28906 }
{"city": "CASCADE PARK", "state": "WA", "total_pop": 26400 }
{"city": "VANCOUVER", "state": "WA", "total_pop": 28837 }
{"city": "PORTERVILLE", "state": "CA", "total_pop": 54599 }
{"city": "LEMOORE NAVAL AI", "state": "CA", "total_pop": 26170 }
{"city": "GOLETA", "state": "CA", "total_pop": 45088 }
{"city": "SIMI VALLEY", "state": "CA", "total_pop": 55528 }
{"city": "RICHMOND", "state": "CA", "total_pop": 57938 }
{"city": "SANTA PAULA", "state": "CA", "total_pop": 28319 }
{"city": "FORT LEWIS", "state": "WA", "total_pop": 27463 }
{"city": "ANAHEIM", "state": "CA", "total_pop": 266121 }
{"city": "SAN BERNARDINO", "state": "CA", "total_pop": 138241 }
{"city": "COWAN HEIGHTS", "state": "CA", "total_pop": 37045 }
{"city": "REDDING", "state": "CA", "total_pop": 90587 }
Type "it" for more
> it
{"city": "PLACENTIA", "state": "CA", "total_pop": 47174 }
```

UNREGISTERED VERSION - Please support Mobaxterm by subscribing to the professional edition here: <https://mobaxterm.mobatek.net>

Type here to search

4:44 PM 3/2/2018

Task 7 – screen 5 - Find all zip code in San Jose, New York, or Washington that have a population between 6,000 and 11,000.

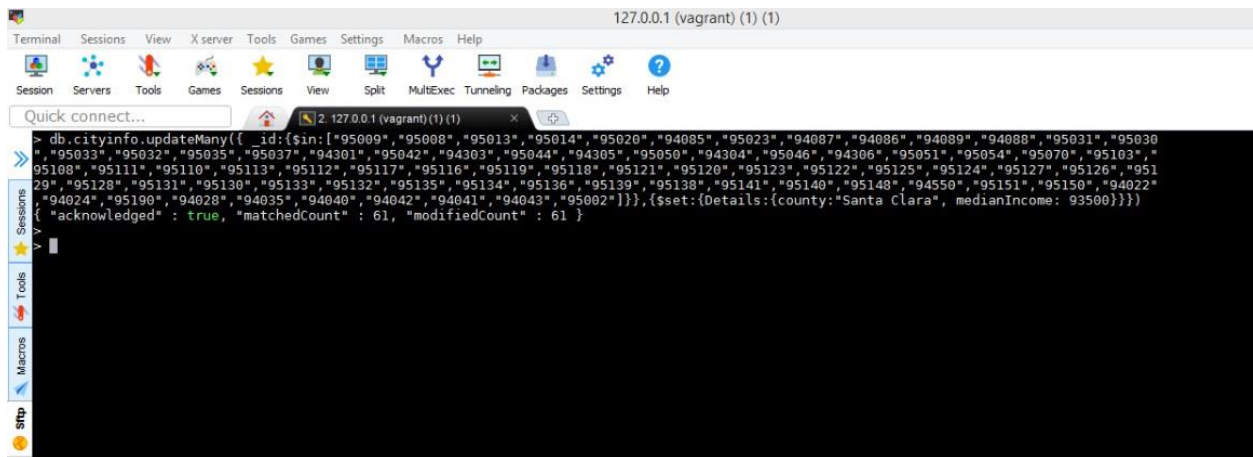
```
db.cityinfo.aggregate([{$match: {pop:{$gte: 6000,$lte:11000}, city:{$in:["SAN JOSE","NEW YORK","WASHINGTON" ]}}},{ $project : { zipcode:"$_id" , "_id" : 0}}])
```

The screenshot shows a MobaXterm terminal window with the following content:

```
> use usdata
switched to db usdata
> db.cityinfo.aggregate([{$match: {pop:{$gte: 6000,$lte:11000}, city:{$in:["SAN JOSE","NEW YORK","WASHINGTON" ]}}},{ $project : { zipcode:"$_id" , "_id" : 0}}])
{"zipcode": "10044" }
{"zipcode": "20005" }
{"zipcode": "20036" }
{"zipcode": "30673" }
{"zipcode": "48094" }
{"zipcode": "52353" }
{"zipcode": "95119" }
{"zipcode": "95135" }
{"zipcode": "95139" }
```

Task 8 – screen 6 - Add an embedded document called "Details" into all documents with a Santa Clara County zip code. In Details, add the following fields with names and values: {county:"Santa Clara", medianIncome: 93500}

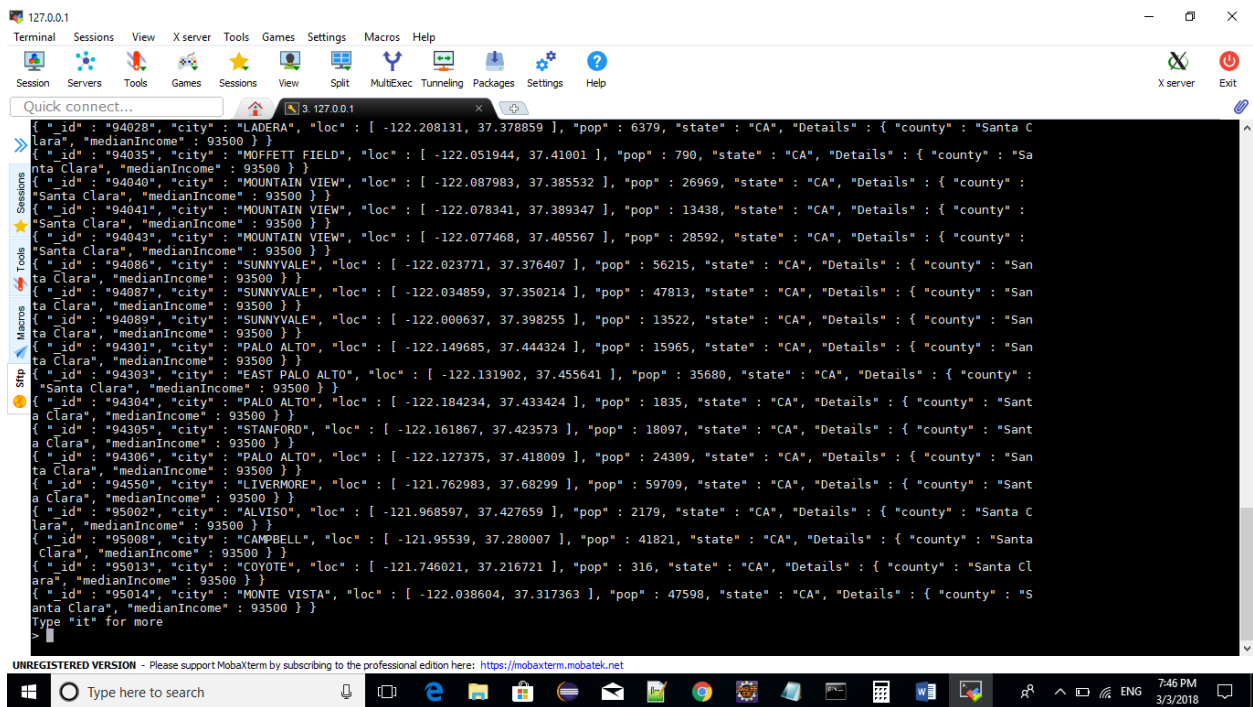
```
db.cityinfo.updateMany(
  {
    _id :
    {$in:["95009","95008","95013","95014","95020","94085","95023","94087","94086","94089","94088","95031","95030","95033","95032","95035","95037","94301","95042","94303","95044","94305","95050","94304","95046","9306","95051","95054","95070","95103","95108","95111","95110","95113","95112","95117","95116","95119","95118","95121","95120","95123","95122","95125","95124","95127","95126","95129","95128","95131","95130","95133","95132","95135","95134","95136","95139","95138","95141","95140","95148","94550","95151","95150","94022","94024","95190","94028","94035","94040","94042","94041","94043","95002"]}},{$set : {Details: {county:"Santa Clara", medianIncome: 93500}}}
)
```



```
127.0.0.1 (vagrant) (1) (1)
Terminal Sessions View X server Tools Games Settings Macros Help
Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help
Quick connect...
>> db.cityinfo.updateMany({_id:{$in:["95000","95006","95013","95014","95020","94085","95023","94087","94086","94089","94088","95031","95030","95033","95032","95035","95037","94301","95042","94303","95044","94305","95050","94304","95046","94306","95051","95054","95070","95103","95108","95111","95110","95113","95112","95117","95116","95118","95121","95120","95123","95122","95125","95124","95127","95126","95129","95128","95131","95130","95133","95132","95135","95134","95136","95139","95138","95141","95140","95148","94550","95151","95150","94022","94024","95190","94028","94035","94040","94042","94041","94043","95002"]}},{$set:{Details:{county:"Santa Clara",medianIncome: 93500}}})
"acknowledged" : true, "matchedCount" : 61, "modifiedCount" : 61
```

Task 9 – screen 7 - Find all documents that have an embedded document named Details without using the zip code.

```
db.cityinfo.find({Details:{$exists:true}})
```



```
127.0.0.1
Terminal Sessions View X server Tools Games Settings Macros Help
Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help
Quick connect...
>> db.cityinfo.find({Details:{$exists:true}})
{"_id": "94028", "city": "LADERA", "loc": [ -122.208131, 37.378859 ], "pop": 6379, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94035", "city": "MOFFETT FIELD", "loc": [ -122.051944, 37.41001 ], "pop": 790, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94040", "city": "MOUNTAIN VIEW", "loc": [ -122.087983, 37.385532 ], "pop": 26969, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94041", "city": "MOUNTAIN VIEW", "loc": [ -122.078341, 37.389347 ], "pop": 13438, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94043", "city": "MOUNTAIN VIEW", "loc": [ -122.077468, 37.405567 ], "pop": 28592, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94086", "city": "SUNNYVALE", "loc": [ -122.023771, 37.376407 ], "pop": 56215, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94087", "city": "SUNNYVALE", "loc": [ -122.034859, 37.350214 ], "pop": 47813, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94089", "city": "SUNNYVALE", "loc": [ -122.000637, 37.398255 ], "pop": 13522, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94301", "city": "PALO ALTO", "loc": [ -122.149685, 37.444324 ], "pop": 15965, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94303", "city": "EAST PALO ALTO", "loc": [ -122.131902, 37.455641 ], "pop": 35680, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94304", "city": "PALO ALTO", "loc": [ -122.184234, 37.433424 ], "pop": 1835, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94305", "city": "STANFORD", "loc": [ -122.161867, 37.423573 ], "pop": 18097, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94306", "city": "PALO ALTO", "loc": [ -122.127375, 37.418009 ], "pop": 24309, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "94550", "city": "LIVERMORE", "loc": [ -121.762983, 37.68299 ], "pop": 59709, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "95002", "city": "ALVISO", "loc": [ -121.968597, 37.427659 ], "pop": 2179, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "95008", "city": "CAMPBELL", "loc": [ -121.95539, 37.280007 ], "pop": 41821, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "95013", "city": "COVOTE", "loc": [ -121.746021, 37.216721 ], "pop": 316, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
{"_id": "95014", "city": "MONTE VISTA", "loc": [ -122.038604, 37.317363 ], "pop": 47598, "state": "CA", "Details": { "county": "Santa Clara", "medianIncome": 93500 } }
Type "it" for more
>
```

Task 10 – screen 8 - Use a MongoDB mapreduce function to find the population of every city and show the populations in the output

The screenshot shows a terminal window titled '127.0.0.1' with a menu bar (Terminal, Sessions, View, X server, Tools, Games, Settings, Macros, Help) and a toolbar. A 'Quick connect...' bar is at the top. The terminal displays the following commands and output:

```
>> db.cityinfo.mapReduce(
  function() {
    emit(this.city, this.pop);
  },
  function(keyCity, valuesPop) {
    return Array.sum(valuesPop);
  },
  { out: "map_reduce_population" }
)

{
  "result": "map_reduce_population",
  "timeMillis": 2201,
  "counts": {
    "input": 29353,
    "emit": 29353,
    "reduce": 6212,
    "output": 16584
  },
  "ok": 1
}

db.map_reduce_population.find()
{ "_id": "AARON", "value": 270 }
{ "_id": "AARONSBURG", "value": 100 }
{ "_id": "ABAC", "value": 27906 }
{ "_id": "ABBEVILLE", "value": 23400 }
{ "_id": "ABBOT VILLAGE", "value": 1193 }
{ "_id": "ABBOTSFORD", "value": 2480 }
{ "_id": "ABBOTT", "value": 577 }
{ "_id": "ABBOTT PARK", "value": 26542 }
{ "_id": "ABBOTTSTOWN", "value": 1777 }
{ "_id": "ABBYVILLE", "value": 267 }
{ "_id": "ABELL", "value": 601 }
{ "_id": "ABERDEEN", "value": 96073 }
{ "_id": "ABERDEEN PROVING", "value": 5294 }
{ "_id": "ABERNATHY", "value": 3482 }
{ "_id": "ABIE", "value": 282 }
{ "_id": "ABILENE", "value": 112711 }
```

At the bottom, a message reads: 'UNREGISTERED VERSION - Please support Mobaxterm by subscribing to the professional edition here: <https://mobaxterm.mobatek.net>'.

Task 11 – Screen 9 - Use a MongoDB aggregate function to find the population of every city and show the populations in the output.

The screenshot shows a terminal window titled '127.0.0.1' with a menu bar (Terminal, Sessions, View, X server, Tools, Games, Settings, Macros, Help) and a toolbar. A 'Quick connect...' bar is at the top. The terminal displays the following command:

```
>> var timeStart=new Date(); db.cityinfo.aggregate([{$group:{"_id":{"city":"$city"},"total_pop":{"$sum":"$pop"}}, {$project:{"city":"$_id.city","total_pop":"$total_pop","_id":0}}]); var timeEnd=new Date(); timeEnd-timeStart;
```

The command is followed by a cursor. At the bottom, a message reads: 'UNREGISTERED VERSION - Please support Mobaxterm by subscribing to the professional edition here: <https://mobaxterm.mobatek.net>'.

Task 12 - Compare the execution times of the mapreduce and the aggregate functions to find the population of every city.

Map reduce function takes more time than aggregation.

Map reduce takes more time here because we have not split the job into multiple threads. Additionally, there is also a need to convert Mongo db's BSON document into JSON document because map reduce uses java script. However, in the case of aggregation pipeline, it is a compiled code while map reduce's code is interpreted code (java script).

Also, map reduce takes more time because there is a write to the database included.