

Practical 5 Aggregation using MongoDB

Class: MSc. DSAI Roll No.:L005

Creating database and importing files

Create Database

Database Name

Database

Collection Name

Database

☐ Time-Series

Time-series collections efficiently store sequences of measurements over a period of time. [Learn More](#)

> Additional preferences

(e.g. Custom collation, Clustered collections)

Cancel

Create Database

Import

To collection Database.Database

Import file:

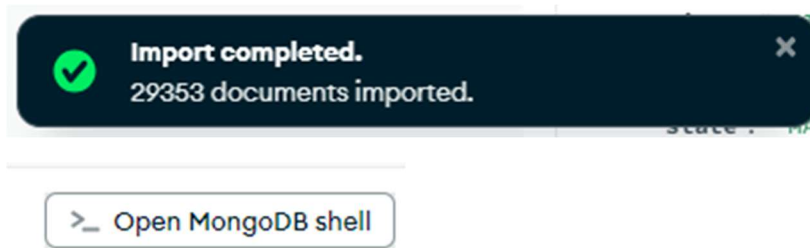
zips.json

Options

☐ Stop on errors

Cancel

Import



Working with a single field

COMPARISON OPERATORS

`$eq` Matches values that are equal to a specified value.

```

> use Database
< switched to db Database
> db["Database"].find({pop:{$eq: 1350}});
< {
  _id: '01053',
  city: 'LEEDS',
  loc: [
    -72.703403,
    42.354292
  ],
  pop: 1350,
  state: 'MA'
}
{
  _id: '15538',
  city: 'GLENCOE',
  loc: [
    -78.827522,
    39.909046
  ],
  pop: 1350,
  state: 'PA'
}
{
  _id: '21654',
  city: 'OXFORD',
  loc: [
    -76.153783,
    38.686358
  ],
  pop: 1350,
  state: 'MD'
}

```

```

{
  _id: '47326',
  city: 'BRYANT',
  loc: [
    -84.911748,
    40.54462
  ],
  pop: 1350,
  state: 'IN'
}

```

\$gt Matches values that are greater than a specified value.

```
> db["Database"].find({'pop':{'$gt: 40117}});  
< {  
  _id: '01040',  
  city: 'HOLYOKE',  
  loc: [  
    -72.626193,  
    42.202007  
  ],  
  pop: 43704,  
  state: 'MA'  
}  
{  
  _id: '01201',  
  city: 'PITTSFIELD',  
  loc: [  
    -73.247088,  
    42.453086  
  ],  
  pop: 50655,  
  state: 'MA'  
}  
{  
  _id: '01420',  
  city: 'FITCHBURG',  
  loc: [  
    -71.803133,  
    42.579563  
  ],  
  pop: 41194,  
  state: 'MA'  
}  
}
```

.....

```
{
  _id: '02840',
  city: 'MIDDLETOWN',
  loc: [
    -71.30348,
    41.504502
  ],
  pop: 47687,
  state: 'RI'
}
```

Type "it" for more

\$ltMatches values that are less than a specified value.

```
> db["Database"].find({pop:{$lt: 20}});
```

```
< {
```

```
  _id: '01338',
```

```
  city: 'BUCKLAND',
```

```
  loc: [
```

```
    -72.764124,
```

```
    42.615174
```

```
  ],
```

```
  pop: 16,
```

```
  state: 'MA'
```

```
}
```

```
{
```

```
  _id: '02163',
```

```
  city: 'CAMBRIDGE',
```

```
  loc: [
```

```
    -71.141879,
```

```
    42.364005
```

```
  ],
```

```
  pop: 0,
```

```
  state: 'MA'
```

```
}
```

```
{
```

```
  _id: '04013',
```

```
  city: 'BUSTINS ISLAND',
```

```
  loc: [
```

```
    -70.042247,
```

```
    43.79602
```

```
  ],
```

```
  pop: 0,
```

```
  state: 'ME'
```

.....

```
{
  _id: '13450',
  city: 'ROSEBOOM',
  loc: [
    -74.74161,
    42.788246
  ],
  pop: 8,
  state: 'NY'
}
```

Type "it" for more

\$lte Matches values that are less than or equal to a specified value.

```
> db["Database"].find({pop:{$lte: 50000}});
```

```
< {
```

```
  _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,
```

```
    42.409698
```

```
  ],
```

```
  pop: 4546,
```

```
  state: 'MA'
```

```
}
```

..... •


```
{
  _id: '01035',
  city: 'HADLEY',
  loc: [
    -72.571499,
    42.36062
  ],
  pop: 4231,
  state: 'MA'
}
Type "it" for more
```

\$in Matches any of the values specified in an array.

```
> db["Database"].find({loc:{$in: [-72.520001,42.255704]}});
< {
  _id: '01033',
  city: 'GRANBY',
  loc: [
    -72.520001,
    42.255704
  ],
  pop: 5526,
  state: 'MA'
}
```

\$ne Matches all values that are not equal to a specified value.

```
> db["Database"].find({state:{$ne: "MA"}});
```

```
< {
```

```
  _id: '02804',
```

```
  city: 'ASHAWAY',
```

```
  loc: [
```

```
    -71.783745,
```

```
    41.423054
```

```
  ],
```

```
  pop: 2472,
```

```
  state: 'RI'
```

```
}
```

```
{
```

```
  _id: '02806',
```

```
  city: 'BARRINGTON',
```

```
  loc: [
```

```
    -71.317497,
```

```
    41.744334
```

```
  ],
```

```
  pop: 15849,
```

```
  state: 'RI'
```

```
}
```

.....

```
{
```

```
  _id: '02835',
```

```
  city: 'JAMESTOWN',
```

```
  loc: [
```

```
    -71.376108,
```

```
    41.516405
```

```
  ],
```

```
  pop: 4999,
```

```
  state: 'RI'
```

```
}
```

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

LOGICAL OPERATORS

`$and` Joins query clauses with a logical AND returns all documents that match the conditions of both clauses.

```
> db["Database"].find({$and:[{state: {$eq: "MA"}}, {city: {$eq: "HADLEY"}}]});  
< {  
  _id: '01035',  
  city: 'HADLEY',  
  loc: [  
    -72.571499,  
    42.36062  
  ],  
  pop: 4231,  
  state: 'MA'  
}
```

`$or` Joins query clauses with a logical OR returns all documents that match the conditions of either clause.

```
> db["Database"].find({$or:[{state: {$eq: "MA"}}, {city: {$eq: "HADLEY"}}]});
< {
  _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: '01035',
  city: 'HADLEY',
  loc: [
    -72.571499,
    42.36062
  ],
  pop: 4231,
  state: 'MA'
}
```

Type "it" for more

\$not Inverts the effect of a query expression and returns documents that do not match the query expression.

```
> db["Database"].find({$and: [{state: {$not: { $eq: "MA" } } }, {city: {$not: { $eq: "HADLEY" }}}]});  
< {  
  _id: '02804',  
  city: 'ASHAWAY',  
  loc: [  
    -71.783745,  
    41.423054  
  ],  
  pop: 2472,  
  state: 'RI'  
}  
{  
  _id: '02806',  
  city: 'BARRINGTON',  
  loc: [  
    -71.317497,  
    41.744334  
  ],  
  pop: 15849,  
  state: 'RI'  
}
```

```
{  
  _id: '02835',  
  city: 'JAMESTOWN',  
  loc: [  
    -71.376108,  
    41.516405  
  ],  
  pop: 4999,  
  state: 'RI'  
}  
Type "it" for more
```

`$nor` Joins query clauses with a logical NOR returns all documents that fail to match both clauses.

```
> db["Database"].find({$nor:[{state: {$eq: "MA"}}, {city: {$eq: "HADLEY"}}]});
< {
  _id: '02804',
  city: 'ASHAWAY',
  loc: [
    -71.783745,
    41.423054
  ],
  pop: 2472,
  state: 'RI'
}
{
  _id: '02806',
  city: 'BARRINGTON',
  loc: [
    -71.317497,
    41.744334
  ],
  pop: 15849,
  state: 'RI'
}
```

```
{
  _id: '02835',
  city: 'JAMESTOWN',
  loc: [
    -71.376108,
    41.516405
  ],
  pop: 4999,
  state: 'RI'
}
Type "it" for more
```

ELEMENT OPERATORS

\$exists Matches documents that have the specified field.

```
> db["Database"].find({state: { $exists: true }});
```

```
< {  
  _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: '01035',  
  city: 'HADLEY',  
  loc: [  
    -72.571499,  
    42.36062  
  ],  
  pop: 4231,  
  state: 'MA'  
}
```

Type "it" for more

\$type Selects documents if a field is of the specified type.

```
> db["Database"].find({city: { $type: "string" }});  
< {  
  _id: '01001',  
  city: 'AGAWAM',  
  loc: [  
    -72.622739,  
    42.070206  
  ],  
  pop: 15338,  
  state: 'MA'  
}
```

..... . .

```
{  
  _id: '01035',  
  city: 'HADLEY',  
  loc: [  
    -72.571499,  
    42.36062  
  ],  
  pop: 4231,  
  state: 'MA'  
}  
Type "it" for more
```

ARRAY OPERATORS

\$all Matches arrays that contain all elements specified in the query.

```
> db["Database"].find({loc: {$all: [-72.571499, 42.36062]}});  
< {  
  _id: '01035',  
  city: 'HADLEY',  
  loc: [  
    -72.571499,  
    42.36062  
  ],  
  pop: 4231,  
  state: 'MA'  
}
```

\$elemMatch Selects documents if
element in the array field matches all the
specified \$elemMatch conditions.

```
> db["Database"].find({loc: {$elemMatch: { $gt: 42, $gt:50 }}});
< {
  _id: '99501',
  city: 'ANCHORAGE',
  loc: [
    -149.876077,
    61.211571
  ],
  pop: 14436,
  state: 'AK'
}
{
  _id: '99502',
  city: 'ANCHORAGE',
  loc: [
    -150.093943,
    61.096163
  ],
  pop: 15891,
  state: 'AK'
}
```

```
{
  _id: '99557',
  city: 'CHUATHBALUK',
  loc: [
    -157.758502,
    61.691648
  ],
  pop: 352,
  state: 'AK'
}
Type "it" for more
```

\$size Selects documents if the array field is a specified size.

```
> db["Database"].find({loc:{$size: 2}});
< {
  _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: '01035',
  city: 'HADLEY',
  loc: [
    -72.571499,
    42.36062
  ],
  pop: 4231,
  state: 'MA'
}
```

Type "it" for more

Advanced Queries

The "\$group" operator

```
> db.Database.aggregate([
  {
    "$group": {
      "_id": "$state",
      "totalPopulation": { "$sum": "$pop" },
      "cityCount": { "$sum": 1 }
    }
  }
])
< {
  _id: 'OK',
  totalPopulation: 3145585,
  cityCount: 586
}
{
  _id: 'AK',
  totalPopulation: 544698,
  cityCount: 195
}
```

..... . .

```
{
  _id: 'MO',
  totalPopulation: 5110648,
  cityCount: 994
}
{
  _id: 'LA',
  totalPopulation: 4217595,
  cityCount: 464
}
Type "it" for more
```

The "\$match" operator

```
> db.Database.aggregate([
  {
    "$match": {
      "state": "MA",
      "pop": { "$gt": 10000 }
    }
  })
< {
  _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: '01108',
  city: 'SPRINGFIELD',
  loc: [
    -72.558432,
    42.085314
  ],
  pop: 25519,
  state: 'MA'
}
Type "it" for more
```

The "\$sort" operator

```
> db.Database.aggregate([
  {
    "$sort": { "pop": -1 }
  })
< {
  _id: '60623',
  city: 'CHICAGO',
  loc: [
    -87.7157,
    41.849015
  ],
  pop: 112047,
  state: 'IL'
}
{
  _id: '11226',
  city: 'BROOKLYN',
  loc: [
    -73.956985,
    40.646694
  ],
  pop: 111396,
  state: 'NY'
}
```



```
{
  _id: '11385',
  city: 'RIDGEWOOD',
  loc: [
    -73.896122,
    40.703613
  ],
  pop: 85732,
  state: 'NY'
}
{
  _id: '10467',
  city: 'BRONX',
  loc: [
    -73.871242,
    40.873671
  ],
  pop: 85710,
  state: 'NY'
}
Type "it" for more
```

The "\$unwind" operator

```
> db.Database.aggregate([
  {
    "$unwind": "$loc"
  })
< {
  _id: '01001',
  city: 'AGAWAM',
  loc: -72.622739,
  pop: 15338,
  state: 'MA'
}
{
  _id: '01001',
  city: 'AGAWAM',
  loc: 42.070206,
  pop: 15338,
  state: 'MA'
}
{
  _id: '01002',
  city: 'CUSHMAN',
  loc: -72.51565,
  pop: 36963,
  state: 'MA'
}
```

```
{  
  _id: '01108',  
  city: 'SPRINGFIELD',  
  loc: [  
    -72.558432,  
    42.085314  
  ],  
  pop: 25519,  
  state: 'MA'  
}
```

Type "it" for more

The Combining Operators

```
> db.Database.aggregate([
  {
    "$match": { "state": "MA" }
  },
  {
    "$group": {
      "_id": "$city",
      "totalPopulation": { "$sum": "$pop" }
    },
    {
      "$sort": { "totalPopulation": -1 }
    }
  ])
< {
  _id: 'WORCESTER',
  totalPopulation: 169856
}
{
  _id: 'SPRINGFIELD',
  totalPopulation: 148062
}
{
  _id: 'DORCHESTER',
  totalPopulation: 126821
}
```

```
{
  _id: 'CHICOPEE',
  totalPopulation: 54891
}
{
  _id: 'MALDEN',
  totalPopulation: 54114
}
{
  _id: 'WEYMOUTH',
  totalPopulation: 54063
}
```

Type "it" for more