

How can I select a random sample of documents in MongoDB?

Authored by
stats writer

June 30, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I select a random sample of documents in MongoDB?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=162100>

MongoDB is a popular NoSQL database system that stores data in a document format. In order to select a random sample of documents from a MongoDB database, there are a few steps that need to be followed. Firstly, the database needs to be queried using the `find()` command to retrieve all the documents. Then, a random number generator can be used to generate a random number within the range of the total number of documents in the database. This random number can be used to specify the number of documents that need to be selected from the database. Finally, the `limit()` command can be used to limit the number of documents returned by the `find()` query to the randomly generated number. This process allows for the selection of a random sample of documents from a MongoDB database.

MongoDB: Select a Random Sample of Documents

You can use the following syntax to select a random sample of documents from a collection in MongoDB:

```
db.myCollection.aggregate()
```

This particular example selects a random sample of 4 documents in the collection titled `myCollection`.

To choose a random sample of a different size, simply change the value in the `size` argument.

The following example shows how to use this syntax in practice with a collection `teams` with the following seven documents:

```
db.teams.insertOne({team: "Mavs", points: 31})
```

```
db.teams.insertOne({team: "Spurs", points: 22})
```

```
db.teams.insertOne({team: "Rockets", points: 19})
db.teams.insertOne({team: "Warriors", points: 26})
db.teams.insertOne({team: "Cavs", points: 33})
db.teams.insertOne({team: "Hornets", points: 30})
db.teams.insertOne({team: "Nets", points: 14})
```

Example: Select Random Sample of Documents in MongoDB

The following code shows how to select a random sample of 4 documents from the teams collection:

```
db.teams.aggregate()
```

This query returns the following documents:

```
{ _id: ObjectId("6203ee711e95a9885e1e765d"),
  team: 'Cavs',
  points: 33 }
{ _id: ObjectId("6203ee711e95a9885e1e765b"),
  team: 'Rockets',
  points: 19 }
{ _id: ObjectId("6203ee711e95a9885e1e7659"),
  team: 'Mavs',
  points: 31 }
{ _id: ObjectId("6203ee711e95a9885e1e765f"),
```

```
team: 'Nets',  
points: 14 }
```

Notice that the following four teams are included in this random sample:

CavsRocketsMavsNets

If we use the `$sample` function again, it will select another random sample of documents which means there is no guarantee that the same set of documents will be chosen.

For example, suppose we select another random sample of 4 documents from the `teams` collection:

```
db.teams.aggregate()
```

This query returns the following documents:

```
{ _id: ObjectId("6203ee711e95a9885e1e765b"),  
team: 'Rockets',  
points: 19 }  
{ _id: ObjectId("6203ee711e95a9885e1e765f"),  
team: 'Nets',
```

```
points: 14 }
{ _id: ObjectId("6203ee711e95a9885e1e765e"),
team: 'Hornets',
points: 30 }
{ _id: ObjectId("6203ee711e95a9885e1e765c"),
team: 'Warriors',
points: 26 }
```

The following four teams are included in this random sample:

RocketsNetsHornetsWarriors

Notice that this random sample does not perfectly match the random sample from the previous example.

Note: You can find the complete documentation for the `$sample` function .

Additional Resources

The following tutorials explain how to perform other common operations in MongoDB:

[MongoDB: How to Query for "NOT NULL" in Specific Field](#)