

How can I find duplicates in MongoDB?

Authored by
stats writer

July 2, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I find duplicates in MongoDB?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=165501>

Finding duplicates in MongoDB can be achieved by using the aggregation framework. This involves grouping the data by a specific field and then counting the number of documents in each group. If the count is more than one, it indicates the presence of duplicates. Additionally, the `$group` and `$match` operators can be used to further refine the search and narrow down the specific field in which the duplicates are present. Another method is to use the `$distinct` operator, which returns a list of unique values for a given field. By comparing this list with the original data, any duplicates can be identified. Overall, the use of aggregation and specific operators in MongoDB can effectively help in finding and managing duplicates in a database.

Find Duplicates in MongoDB

You can use the following syntax to find documents with duplicate values in MongoDB:

```
db.collection.aggregate()
```

Here's what this syntax does:

Group all documents having the same value in field1
Match the groups that have more than one document
Project all groups that have more than one document

This particular query finds duplicate values in the field1 column. Simply change this value to change the field to look in.

The following example shows how to use this syntax

with a collection teams with the following documents:

```
db.teams.insertOne({team: "Mavs", position: "Guard",  
points: 31})db.teams.insertOne({team: "Mavs", position:  
"Guard", points: 22})db.teams.insertOne({team:  
"Rockets", position: "Center", points:  
19})db.teams.insertOne({team: "Rockets", position:  
"Forward", points: 26})db.teams.insertOne({team:  
"Cavs", position: "Guard", points: 33})
```

Example: Find Documents with Duplicate Values

We can use the following code to find all of the duplicate values in the 'team' column:

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

This query returns the following results:

```
{ name: 'Rockets' }  
{ name: 'Mavs' }
```

This tells us that the values 'Rockets' and 'Mavs' occur multiple times in the 'team' field.

Note that we can simply change \$team to \$position to instead search for duplicate values in the 'position' field:

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

This query returns the following results:

```
{ name: 'Guard' }
```

Additional Resources

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