

monogdb duplicate

```
cursor = WEB_DATA_DB.aggregate([
    {'$group': {
        '_id': {'txt': {'$concat': ["$p_tag", " ", " ", "$p_class", " ", " ",
"$tag", " ", " ", "$class", " ", " ", "$text"]}},
        'uniqueIds': {'$addToSet': "$_id"},
        'count': {'$sum': 1}
    }},
    {'$match': {
        'count': {"$gt": 1}
    }}
]);
for doc in cursor:
    print(doc)
#     for doc_id in doc['uniqueIds']:
#         WEB_DATA_DB.delete_one({'_id': doc_id})
```

Plugin Backlinks:

From:

<https://moro.kr/> - **Various Ways**

Permanent link:

<https://moro.kr/open/mongodb-duplicate>

Last update: **2020/07/08 09:02**

