

Lab Exercise (Filtering Data in MongoDB)

```
In [59]: ▶ #Imports MongoDB into Python if not already imported
import pymongo
from pymongo import MongoClient
client = MongoClient('mongodb://localhost:27017/')
```

```
In [60]: ▶ #This will create a Collection/Database called books if it doesn't already exist
db = client['books']
```

```
In [61]: ▶ #Names the collection
books = db.books
```

```
In [62]: ▶ #Inserts the book data
db.books.insert_many([{"book": "The Ranch", "author": "Danielle Steel"},
                      {"book": "The Maniac", "author": "Rishabh Mishra"},
                      {"book": "The runaway jury", "author": "John Grisham"},
                      {"book": "Time to kill", "author": "John Grisham"},
                      {"book": "The testament", "author": "John Grisham"},
                      {"book": "Tell me your Dreams", "author": "Sidney Sheldon"}
                      ])
```

```
Out[62]: <pymongo.results.InsertManyResult at 0x28fbc8ac140>
```

```
In [63]: ▶ #Finds and list all of the books
for book in books.find():
    print(book)
```

```
{ '_id': ObjectId('60e8481e5b0d30bf54d0d1cd'), 'book': 'The Ranch', 'author': 'Danielle Steel' }
{ '_id': ObjectId('60e8481e5b0d30bf54d0d1ce'), 'book': 'The Maniac', 'author': 'Rishabh Mishra' }
{ '_id': ObjectId('60e8481e5b0d30bf54d0d1cf'), 'book': 'The runaway jury', 'author': 'John Grisham' }
{ '_id': ObjectId('60e8481e5b0d30bf54d0d1d0'), 'book': 'Time to kill', 'author': 'John Grisham' }
{ '_id': ObjectId('60e8481e5b0d30bf54d0d1d1'), 'book': 'The testament', 'author': 'John Grisham' }
{ '_id': ObjectId('60e8481e5b0d30bf54d0d1d2'), 'book': 'Tell me your Dreams', 'author': 'Sidney Sheldon' }
```

```
In [64]: #Selects all of the books by a single author  
books.find_one({'author': "Danielle Steel"})
```

```
Out[64]: {'_id': ObjectId('60e8481e5b0d30bf54d0d1cd'),  
          'book': 'The Ranch',  
          'author': 'Danielle Steel'}
```

```
In [65]: #Selects all of the books by a single author  
for book in books.find({'author': "Sidney Sheldon"}):  
    print(book)
```

```
{'_id': ObjectId('60e8481e5b0d30bf54d0d1d2'), 'book': 'Tell me your Dreams', 'author': 'Sidney Sheldon'}
```

```
In [66]: #Selects and lists all the books by a single author by just the title  
for book in books.find({'author': "John Grisham"}):  
    print(book['book'])
```

```
The runaway jury  
Time to kill  
The testament
```

```
In [67]: #Orders and lists all the books in ASCENDING order by the previous authors  
for book in books.find().sort([("author", pymongo.ASCENDING)]):  
    print(book)
```

```
{'_id': ObjectId('60e8481e5b0d30bf54d0d1cd'), 'book': 'The Ranch', 'author': 'Danielle Steel'}  
{'_id': ObjectId('60e8481e5b0d30bf54d0d1cf'), 'book': 'The runaway jury', 'author': 'John Grisham'}  
{'_id': ObjectId('60e8481e5b0d30bf54d0d1d0'), 'book': 'Time to kill', 'author': 'John Grisham'}  
{'_id': ObjectId('60e8481e5b0d30bf54d0d1d1'), 'book': 'The testament', 'author': 'John Grisham'}  
{'_id': ObjectId('60e8481e5b0d30bf54d0d1ce'), 'book': 'The Maniac', 'author': 'Rishabh Mishra'}  
{'_id': ObjectId('60e8481e5b0d30bf54d0d1d2'), 'book': 'Tell me your Dreams', 'author': 'Sidney Sheldon'}
```

```
In [68]: #Groups and counts the number of books from each of the two authors and places the data in a pipeline.  
pipeline = [  
    {"$group": {"_id": "$author", "count": {"$sum": 1}}}]  
  
grp_books = db.books.aggregate(pipeline)
```

```
In [69]: ► #Prints the authors name and the number of the books in the collection  
for book in grp_books:  
    print(book)
```

```
{'_id': 'Rishabh Mishra', 'count': 1}  
{'_id': 'Sidney Sheldon', 'count': 1}  
{'_id': 'Danielle Steel', 'count': 1}  
{'_id': 'John Grisham', 'count': 3}
```

```
In [ ]: ►
```