

Học phần Quản trị Dữ liệu lớn: Bài thực hành MongoDB

Phạm Tiến Lâm, Đặng Văn Báu

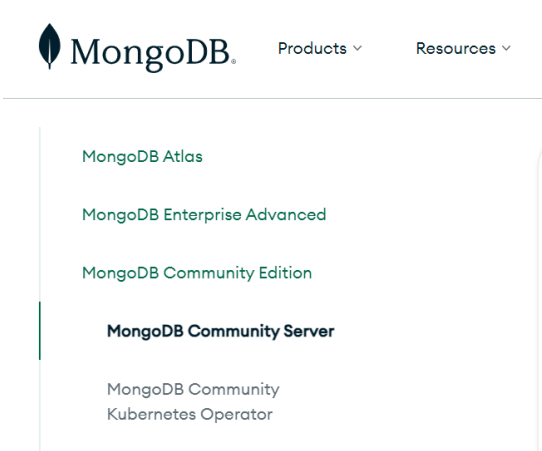
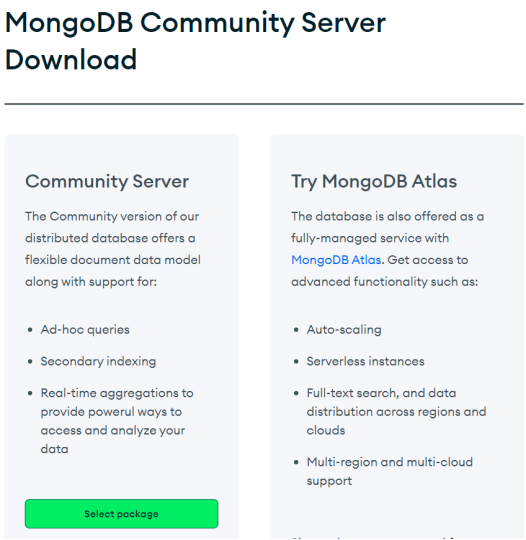

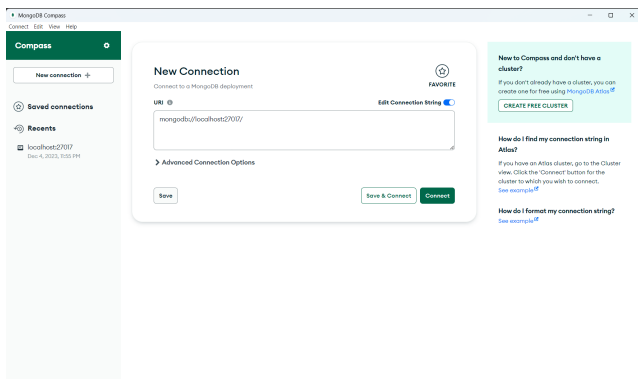
Cài đặt MongoDB - Ubuntu

Link hướng dẫn cài đặt:

<https://www.youtube.com/watch?v=3wNvKybVyaI&list=LL&index=2&t=164s>

Cài đặt MongoDB - Windows

Link download: <https://www.mongodb.com/try/download/community>

<p>Bước 1: Truy cập vào link download:</p> 	<p>Bước 2: Chọn Select Package:</p> 
<p>Bước 3: Chọn Download:</p> 	<p>Bước 5: Mở ứng dụng MongoDB Compass sau khi download:</p> 

Cài đặt MongoDB với Python:

`python -m pip install pymongo`

Hoặc

`Pip install pymongo`

Activity 1: Làm quen với Pymongo

Khai báo thư viện:

```
import pymongo
from bson import ObjectId
```

Kết nối đến MongoDB Server

```
myclient = pymongo.MongoClient("mongodb://localhost:27017/")
```

Chọn/Tạo cơ sở dữ liệu

```
mydb = myclient["test"]
```

Tạo Collection

```
mycol = mydb["customers"]
```

Liệt kê tên tất cả các cơ sở dữ liệu đã khởi tạo trên máy chủ MongoDB

```
print(myclient.list_database_names())
```

Activity 2: Các thao tác cơ bản với MongoDB

- Thêm/Tạo 1 Document trong database: **insert_one**

```
# insert 1 document
mydict = { "name": "Bau", "address": "HaNoi", 'age': 24 }

x = mycol.insert_one(mydict)
print(x.inserted_id)
for x in mycol.find():
    print(x)
```

- Thêm/Tạo nhiều Document trong database: **insert_many**

```
# insert many document
mydict = [
    { "name": "Bau1", "address": "SaiGon", 'age': 29 },
    { "name": "Bau2", "address": "ThaiNguyen", 'age': 25 }
]

x = mycol.insert_many(mydict)
print(x.inserted_ids)
for x in mycol.find():
    print(x)
```

- Cập nhật 1 Document: **update_one**

```
# update one
myquery = { "name": "Bau1" }
newvalues = { "$set": { "name": "Bau 123" } }

mycol.update_one(myquery, newvalues)

# print "customers" after the update:
for x in mycol.find():
    print(x)
```

- Cập nhật nhiều Document: **update_many**

```
# Update many
myquery = { }
newvalues = { "$set": { "name": "value123", 'age': 30 } }

x = mycol.update_many(myquery, newvalues)

print(x.modified_count, "documents updated.")
#print "customers" after the update:
for x in mycol.find():
    print(x)
```

- Cập nhật có điều kiện: **update_many**

```
# Update selection
myquery = {"_id": {"$in": [ObjectId('656f62aac8a00fe5426f744e'),
                           ObjectId('656f62adc8a00fe5426f744f')]}}
```

newvalues = {"\$set": {"name": "New Name", "age": 25}}

x = mycol.update_many(myquery, newvalues)

print(x.modified_count, "documents updated.")

for x in mycol.find():

print(x)

- Xóa 1 Document: **delete_one**

(Trong trường hợp có nhiều Document giống nhau thì mặc định xóa Document đầu tiên)

```
# Delete one
myquery = { 'name': 'New Name' }
```

x = mycol.delete_one(myquery)

print(x.deleted_count, "documents delete.")

for x in mycol.find():

print(x)

- Xóa nhiều Document: **delete_many**

```
# Delete many
myquery = { "name": {"$regex": "^N"} }
```

x = mycol.delete_many(myquery)

print(x.deleted_count, " documents deleted.")

for x in mycol.find():

print(x)

- Xóa toàn bộ Document: **delete_many**

```
# Delete all
x = mycol.delete_many({})

print(x.deleted_count, " documents deleted.")
for x in mycol.find():
    print(x)
```

- Tìm kiếm 1 Document: **find_one**

```
# find one

# Find the document with the specified _id
target_id = ObjectId('656f537325ac558c1f7bbbd0')
result1 = mycol.find_one({"_id": target_id})

# Find the document with the name
result2 = mycol.find_one({"name": 'Bau'})

# Display the result
print(result1)
print(result2)
```

- Tìm kiếm tất cả Document: **find**

```
# Find all
for x in mycol.find():
    print(x)
```

- Tìm kiếm Document theo truy vấn (Filter):

```
# Filter
# Muốn lấy 1 vài field trong câu query thì có thể làm như sau
for x in mycol.find({}, {"_id": 0, "name": 1, "age": 1 }):
    print(x)
```

- Sắp xếp các Document theo truy vấn: **sort**

```
# Sort

mydoc = mycol.find().sort("age")

for x in mydoc:
    print(x)
```

- Truy vấn Document có giới hạn đầu ra: **limit**

```
# limit
myresult = mycol.find().limit(10)

#print the result:
for x in myresult:
    print(x)
```

Activity 2: Bài tập

1. Hãy viết một đoạn mã Python để chèn một bản ghi mới vào một collection có tên là "students". Khởi tạo 10 bản ghi (Document) bao gồm các trường thông tin "name", "age", và "major".
2. Hãy viết một đoạn mã để cập nhật thông tin của sinh viên có tên là "student1" trong collection "students". Bạn có thể cập nhật bất kỳ trường nào bạn muốn, ví dụ: thay đổi age hoặc major.
3. Tạo một đoạn mã để xóa sinh viên có tên là "student5" và "student9" khỏi collection "students".
4. Viết một câu truy vấn để tìm một sinh viên bằng tên (ví dụ: "student4") trong collection "students" và hiển thị thông tin của sinh viên đó.
5. Hãy viết một đoạn mã để tìm tất cả sinh viên có độ tuổi lớn hơn 20 và sắp xếp kết quả theo tên tăng dần. Đồng thời, giới hạn kết quả chỉ hiển thị 5 sinh viên đầu tiên.