

웹개발 종합반 정리 -4

- 3주차: Python 패키지 Pymongo, Dnspython / 실습) genie뮤직 스크래핑

링크

- ※ 스파르타코딩클럽 <https://online.spartacodingclub.kr/classroom>
- ※ 웹개발 강의 doc <https://teamsparta.notion.site/1-da6d8595945d4b868e5fcfa933c56011>

※ Python패키지: Pymongo, Dnspython 설치하기

Pymongo: py+mongo, pymongo로 mongodb에 연결 // mongo_client 아니고 **MongoClient**

"mongodb+srv://아이디:비번@cluster0.rliere2.mongodb.net/?retryWrites=true&w=majority"

기본
상차림
3줄

```
from pymongo import MongoClient
client = MongoClient('mongodb+srv:// ~ ~')
db = client.데이터베이스명
```

```
doc = {
    'name': 'bibim-bob',
    'age': 2
}
db.컬렉션.insert_one(doc)
```

pymongo에서 MongoClient를 import
db는 접속한 client의 데이터베이스명
{ document }
db의 컬렉션에 document를 insert_one

insert 한 줄로 db.users.insert_one({'name': 'sam-gyup', 'age': 3})

find({}) 전체출력

```
users = db.users.find({})
for user in users:
    print(user['name'])
```

users = db.users.find({}, {"_id": False}) # _id값 제외하고 찾기
<pymongo.cursor.Cursor object at 0x0000028C593A1D80>

find_one({키:값}) evil = db.users.find_one({'name': 'sam-gyup'})
print(evil)

update_one({키:값}, {\$set: {키:값}}) db.users.update_one({'name': 'sam-gyup'}, {'\$set': {'age': 4}})

delete_one({키:값}) no callback

delete_many({})

실습) genie뮤직 스크래핑	import requests from bs4 import BeautifulSoup	# bs4 import
	headers = {'User-Agent' : 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36(KHTML, like Gecko) Chrome/73.0.3683.86 Safari/537.36'}	# 템플릿에 주소 입력
순위, 제목, 가수	data= requests.get('https://www.genie.co.kr/chart/top200?ditc=M&rtm=N&ynd=20210701', headers=headers)	# soup변수에 html을 parser
	soup = BeautifulSoup(data.text, 'html.parser') tr = soup.select("#body-content > div.newest-list > div > table > tbody > tr") for td in tr: rank = td.select_one('td.number') title = td.select_one('td.info > a.title.ellipsis') artist = td.select_one('td.info > a.artist.ellipsis')	# 요소 잡기 strip() .text [0:2] 자르기
	rank = rank.text[0:2].strip() title = title.text.strip() artist = artist.text print(rank, title, artist)	

예러A: 원지 모를 때 다른 파일을 새로고침하고 있었음 => 당황하지말고 파일 이름 확인하기~

주의A: update_one 업데이트할 데이터의 자료형 먼저 확인하기

주의B: \$set '\$set'에 따옴표 채기기