

카카오 이미지 검색결과 수집하기

패션 잡지에 넣을 봄을 주제로 한 옷차림을 알아보고자 한다. 해당 주제로 검색된 이미지를 수집하여라

```
# 필요한 모듈 참조
import requests
import json
from pandas import DataFrame
import os
```

접속할 URL

```
urlFmt = "https://dapi.kakao.com/v2/search/image?query={query}&page={page}&size={size}"
```

```
# 요청변수
query = "봄 패션"
page = 1
size = 80
key = "b636ff83df45f778874b079dcb979102"
```

브라우저 버전정보

```
userAgent = "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like
```

```
# 접속객체 생성
session = requests.Session()

# 접속객체에 부가정보(header) 삽입하기
session.headers.update({
    "Authorization": "KakaoAK %s" % key
})
```

```
# 다음 페이지가 존재하는지 여부 --> False인 동안은 while문 반복
isEnd = False
```

```
# 1페이지부터 시작
page = 1
```

```
# 저장된 이미지 카운트
count = 0
```

```
while (not isEnd):
    # -----
    # 이전 예제 코드 시작
    # -----
    # 접속할 URL
    url = urlFmt.format(query=query, page=page, size=size)
    url
```

```

# 생성한 접속객체를 활용하여 API에 접속
r = session.get(url)

# 접속에 실패한 경우
if r.status_code != 200:
    # 에러코드와 에러메시지 출력
    msg = "[%d Error] %s 에러가 발생함" % (r.status_code, r.reason)
    # 에러를 강제로 생성시킴
    raise Exception(msg)

# 인코딩 형식 지정
r.encoding = "utf-8"
# 가져온 문자열을 딕셔너리로 변환
data = json.loads(r.text)
#data

#-----
# 이전 예제 코드 끝
#-----

# 검색 결과를 갖고 있는 딕셔너리를 포함하는 리스트 추출
searchList = data['documents']

# 증감식에 해당하는 isEnd변수 추출
isEnd = bool(data['meta']['is_end'])
page += 1
print("isEnd=", isEnd, " / 다음페이지=", page)

# 이미지가 저장될 폴더 생성
if os.path.exists(query) == False:
    os.mkdir(query)

# 이미지 저장
for i, v in enumerate(searchList):
    # 숫자 카운트
    count += 1

    # 이미지 URL 추출
    docUrl = v['doc_url']
    imageUrl = v['image_url']
    print(imageUrl)

    # 접속객체에 부가정보(header) 삽입하기
    session.headers.update({
        "Authorization": "KakaoAK %s" % key,
        "Referer": docUrl,
        "User-Agent": userAgent
    })

# 텍스트 이외의 파일 형식을 가져오기 위해서는 stream 파라미터 추가
try:
    r = session.get(imageUrl, stream=True)
except:
    continue

```

```
# 다운로드에 실패한 경우 다음 이미지로 이동
if r.status_code  $\neq$  200:
    continue

fileName = "%s/%05d.png" % (query, count)

with open(fileName, 'wb') as f:
    # 통신으로 가져온 내용의 바이너리(기계어)를 직접 추출
    f.write(r.raw.read())
    print(fileName, "(이)가 저장되었습니다.")

print("종료")
```