



오늘의 학습

학습목표

- JSON이해
- 정규 표현식 이해
- 네이버 API 사용법



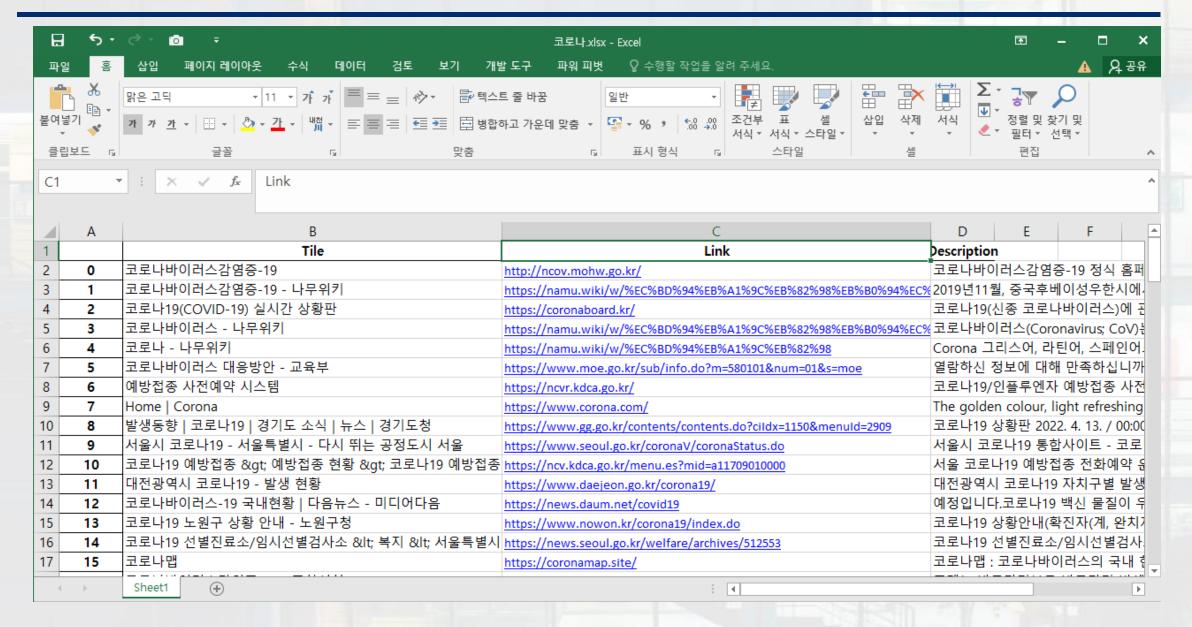
[출력물] 네이버 API를 통해 코로나 관련 웹 문서 결과를 Pandas로 출력하기

검색 질의=>교로나

| | Tile | Link | Description |
|-----|---|---|--|
| 0 | 코로나바이러스감염증-19 | http://ncov.mohw.go.kr/ | 코로나바이러스감염증-19 정식 홈페이지로 발생현황, 국 내발생현황, 국외발생현황, 시 |
| 1 | 코로나바이러스감염증-19 - 나무위키 | https://namu.wiki/w/%EC%BD%94%EB%A1%9C%EB%82%9 | 2019년11월, 중국후베이성우한시에서 처음으로 발생하 여 보고된 새로운 유형의 변종 |
| 2 | 코로나19(COVID-19) 실시간 상황판 | https://coronaboard.kr/ | 코로나19(신종 코로나바이러스)에 관한 세계 각 국가들의 통계 및 뉴스 등을 취합하 |
| 3 | 코로나바이러스 - 나무위키 | https://namu.wiki/w/%EC%BD%94%EB%A1%9C%EB%82%9 | 코로나바이러스(Coronavirus; CoV)는 사람을 포함한 동 물에 광범위한 흐흡 |
| 4 | 코로나 - 나무위키 | https://namu.wiki/w/%EC%BD%94%EB%A1%9C%EB%82%98 | Corona 그리스어, 라틴어, 스페인어로 왕관을 뜻하는 단 어다. 태양 상충부 대기 |
| | | | |
| 115 | 코로나19 방역수칙 위반 신고센터<민원신청< 용답소 - 서울특별시 용답소 | https://eungdapso.seoul.go.kr/Req/Req02/Req02 | 코로나19 방역수칙 위반 신고센터 홈 > 민원신청 > 코로나19 방역수칙 |
| 116 | 코로나19 알림내역 코로나19 대용현황 코로나 19알림내역 코로나19 | https://www.gangdong.go.kr/web/coronamain/bbs/ | 등코로나19가 종식되는 그 날까지 강동구는 구민 여러분 의 생명과 등 코로나19 |
| 117 | #코로나 - YouTube | https://www.youtube.com/hashtag/%EC%BD%94%EB%A | 의료체계도 독감처럼코로나 확진자 격리 의무 사라진 다 뉴스A거리두기도 이제 끝? |
| 118 | <코로나 자가격리 지원금> 코로나 재택치료 지원금 (대상자, 신청방법, | https://dingdo.tistory.com/351 | 코로나 자가격리 지원금 (+재택치료 지원금) 코로나 자가 격리 생활지원금(신청방법 및 |
| 119 | 고용노동부> 뉴스·소식> 코로나19 대용 | https://www.moel.go.kr/policy/corona19/main.do | 4차 긴급고용안정지원금 Q&A(신규신청자)외국인 근로자 다수 사업장 코로나1 |
| | | | |



[출력물] 네이버 API를 통해 코로나 관련 웹 문서 결과를 엑셀로 출력하기







JSON(Javascript Object Notation)으로 자바스크립트 객체 표기법으로 작성된 텍스트 Python에는 JSON 데이터 작업에 사용할 수 있는 이라는 내장 패키지가 있습니다

{}를 사용

예>표현

{"name": "홍길동", "age":20}



파이썬으로 JSON 다루기

json 모듈을 가져오기 import json

```
import json
   # JSON:
   |x = '{ "name":"홍길동", "age":20, "city":"서울"}'
                                        JSON객체는 키값으로
 6 print(x)
                                         접근 불가
 8 print(x["age"])
{ "name":"홍길돔", "age":20, "city":"서울"}
TypeError
                                   Traceback (most recent call last)
C:\Users\Public\Documents\ESTsoft\CreatorTemp/ipykernel_3312/2367281842.py in <module>
     6 print(x)
TypeError: string indices must be integers
```



파이썬으로 JSON 다루기

```
json 모듈을 가져오기 import json
```

JSON 구문 분석: JSON에서 Python으로 변환

{'name': '홍일동', 'age': 20, 'city': '서울'}

json.loads(): JSON 문자열이 있는 경우 메서드를 사용하여 구문 분석

JSON

20

파싱(parsing):

구문 분석이라고 한다. 문장이



[참고] 인코딩과 디코딩 (Encoding & Decoding)]

문자 코드를 기준으로 문자를 코드로 변환하는 것을 **문자 인코딩(encoding)** 이라하고 코드를 문자로 변환하는 것을 **문자 디코딩(decoding)**

UTF-8:하나의 문자를 1~4바이트의 가변길이로 표현. 1바이트 영역은 ASCII코드와 하위 호환되며 ASCII코드의 128개 문자 집합은 UTF-8과 동일하게 호환됨. 현재 인터넷에서 가장 많이 쓰이는 인코딩이며 뛰어난 크로스플랫폼 호환성도 갖고 있음.

정규표현식



- 정규식 모듈
- re Python에는 정규식으로 작업하는 데 사용할 수 있는 내장 패키지
- 정규식은 검색 패턴을 형성하는 일련의 문자
- 정규표현식은 복잡한 문자열을 처리하는데 사용하는 기법으로 문자열을 처리하는 곳에서 모두 사용



정규표현식의 기초, 메타문자

.^*+?{}[]\|()

문자클래스:[]

- 문자 클래스로 만들어진 정규식은 []안에 들어있는 문자들과 매치라는 의미
- []안의 두 문자 사이에 하이픈(-)이 사용되면 두 문자 사이의 범의를 의미(예[a-d]는[abcd]와 동일

<예>정규표현식이 [abc]라면 'a,b,c'중 한 개의 문자와 매치

| 정규식 | 문자열 | 매치여부 | 설명 |
|-------|------|------|--|
| [abc] | а | Yes | 'a'는 정규식과 일치하는 "a"가 존재함으로 매치 |
| | bee | Yes | 'bee'는 정규식과 일치하는 "b"가 존재함으로 매치 |
| | dude | no | 'dude'는 정규식과 일치하는 "a,b,c"중 하나도 포함하 지 않기 때문에 매치되지 않음 |

- [a-zA-Z]:알파벳 모두
- [0-9]: 숫자



정규표현식의 기초, 메타문자

.^*+?{}[]\|()

Dot(.)

줄바꿈 문자인 \n를 제외한 모든 문자와 매치

<예> a.b

설명 : a와 b사이에 줄바꿈 문자를 제외한 어떤 문자가 들어가도 모두 매치

| 정규식 | 문자열 | 매치여부 | 설명 |
|-----|-----|------|--|
| a.b | aab | Yes | 'aab'는 가운데 문자"a"가 모든 문자를 의미함으로 정 규식과 매치 |
| | a3b | Yes | 'a3b'는 가운데 문자"3"가 모든 문자를 의미함으로 정 규식과 매치 |
| | abd | no | 'abd' 가운데 문자"a"문자와 "b"문자 사이에 어떤 문자라도 하나가 있어야 하는 정규식과 일치하지 않기 때문에 매치하지 않음 |



정규표현식의 기초, 메타문자

.^*+?{}[]\|()

반복(*)

- 반복의미, *바로 앞에 있는 문자가 0부터 무한 반복

| 정규식 | 문자열 | 매치여부 | 설명 |
|------|---------|------|---------------------------|
| be*t | bt | Yes | "b"가 0번 반복되어 매치 |
| | beet | Yes | "b"가 0번 이상 반복되어 매치(2번 반복) |
| | bbbbbbt | yes | "b"가 0번 이상 반복되어 매치(5번 반복) |



정규표현식의 기초, 메타문자

.^*+?{}[]\|()

반복(+)

- 반복의미, *바로 앞에 있는 문자가 1부터 무한 반복

| 정규식 | 문자열 | 매치여부 | 설명 |
|------|---------|------|---------------------------|
| be*t | bt | No | "b"가 0번 반복되어 매치되지 않음 |
| | beet | Yes | "b"가 0번 이상 반복되어 매치(2번 반복) |
| | bbbbbbt | yes | "b"가 0번 이상 반복되어 매치(5번 반복) |



정규표현식의 기초, 메타문자

.^*+?{}[]\|()

반복({m,n},?) {} 사용법

| 정규식 | 문자열 | 매치여부 | 설명 |
|--------|------|------|----------------------|
| be{2}t | bet | no | "e"가 1번 반복되어 매치되지 않음 |
| | beet | Yes | "e"가 2번 반복되어 매치 |

반복({m,n},?)

- 반복 횟수가 m부터 n까지, m과 n생략 가능(m은 생략하면 0, n은 생략하면 무한대)

| 정규식 | 문자열 | 매치여부 | 설명 |
|----------|---------|------|----------------------|
| be{2,5}t | bet | no | "e"가 1번 반복되어 매치되지 않음 |
| | beet | Yes | "e"가 2번 반복되어 매치 |
| | beeeeet | yes | "e"가 5번 반복되어 매치 |

remove_tag = re.compile('<.*?>')



정규표현식의 기초, 메타문자

.^*+?{}[]\|()

:

- {0,1}의미(0~1번 사용되면 매치)

| 정규식 | 문자열 | 매치여부 | 설명 |
|------|-----|------|-----------------|
| ab?c | ac | yes | "b"가 0번 사용되어 매치 |
| | abc | Yes | "b"가 1번 사용되어 매치 |



Python에서 정규 표현식 사용하기

- 정규식 모듈
- re(regular expression) Python에는 정규식으로 작업하는 데 사용할 수 있는 이라는 내장 패키?
- 정규식은 검색 패턴을 형성하는 일련의 문자
- 정규표현식은 복잡한 문자열을 처리하는데 사용하는 기법으로 문자열을 처리하는 곳에서 모두 사용

```
re모듈 가져오기
import re
p = re.compile('abc*') # re.compile을 이용하여 표현식('abc*)을 컴파일 한다.
```



Python에서 정규 표현식을 이용한 문자열 검색

- 컴파일된 패턴 객체는 문자열 검색을 위한 4가지 메서드 제공

| 메서드 | |
|------------|---------------------------------|
| match() | 문자열의 처음부터 정규식과 매치되는지 조사 |
| search() | 문자열 전체를 검색하여 정규식과 매치되는지 조사 |
| findall() | 정규식과 매치되는 모든 문자열을 리스트로 리턴 |
| finditer() | 정규식과 매치되는 모든 문자열을 반복 가능한 객체로 리턴 |
| | |

match, search는 정규식과 매치될때 match객체를 리턴하고, 매치가 되지 않으면 none리턴

match 객체의 매서드

| 메서드 | |
|---------|-----------------------------|
| group() | 매치된 문자열을 리턴 |
| start() | 매치된 문자열의 시작위치를 리턴 |
| end() | 매치된 문자열의 끝 위치 리턴 |
| span() | 매치된 문자열의(시작,끝)에 해당되는 튜플을 리턴 |

[실습하기]

정규식 표현 사용

```
|#1. 폐턴 만들기
 2 import re
 3 p = re.compile('[a-z]+') # IHE P = 7
 1 #2. 매치하기
 2 m = p.match("python") # match()메서드는 문자열의 처음부처 정규식과 매치되는지 조사.
                      # "python" 문자열은 [a-z]+ 정규식에 부할되므로 matct 객체 리턴
 3 | print(m)
<re.Match object; span=(0, 6), match='python'>
  m = p.match("7 python")
                       # "7 python" 문자열은 [a-z]+ 정규식에 부활되지 않으므로 Mone 리턴
 2 | print(m)
None
 1 |# 1. 폐턴 만들기
 2 p2 = re.compile('[0-9a-z]+') # INE DE 7
```

1 # 2. 매치하기 2 m = p2.match("7 python") 3 print(m) # "7 python" 문자열은 [0-0a-z]+ 정규식에 부활되지 않으므로 matct 객체 리턴 <re.Match object; span=(0, 1), match='7'>

메타문자 [참고]

| Character | Description | Example |
|-----------|--|---------------|
| [] | []안에 있는 모든 문자 | "[a-m]" |
| \ | 특수 시퀀스 신호(특수 문자를 이스케이프하는 데 사용할 수도 있음) | "\d" |
| | 모든 문자(except newline character) | "heo" |
| ۸ | 시작 | "^hello" |
| \$ | 끝 | "planet\$" |
| * | 0회 이상 발생 | "he.*o" |
| + | 하나 이상의 발생 | "he.+o" |
| ? | 0또는 1회 발생 | "he.?o" |
| {} | 정확히 지정된 발생 횟수 | "he.{2}o" |
| I | 하나 또는 | "falls stays" |
| () | 캡처 및 그룹화 | |

네이버API이용하기



네이버 포털 사이트 API 이용하기

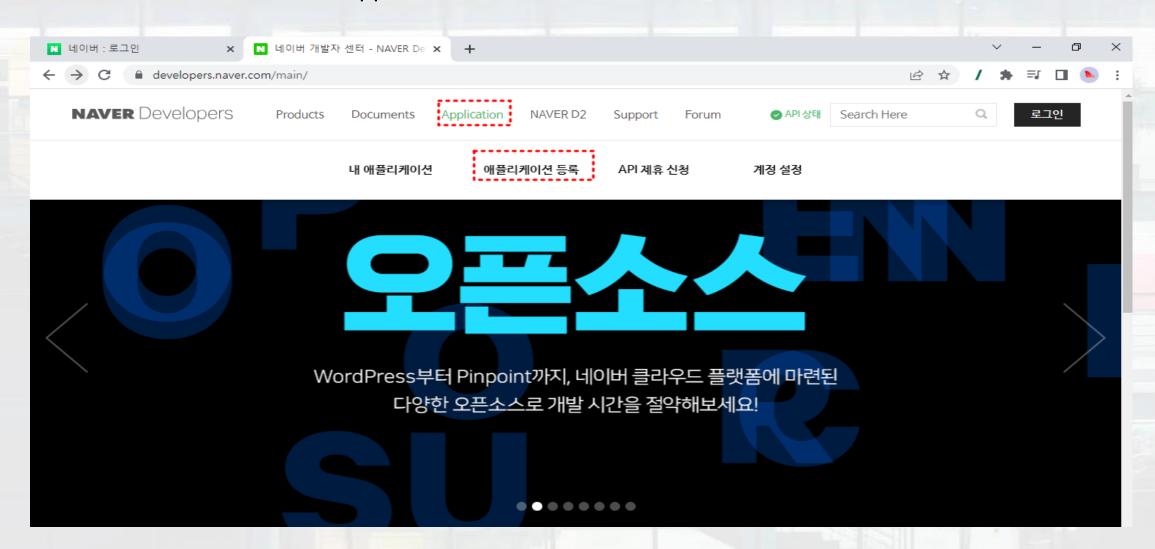
API(application programming interface 애플리케이션 프로그래밍 인터페이스, 응용 프로그램 프로그래밍 인터페이스)는 컴퓨터나 컴퓨터 프로그램 사이의 연결

네이버 API를 호출하기 위해서는 API 토큰과 시크릿이 있어야 합니다.



네이버 개발자 가입하기 developers.naver.com

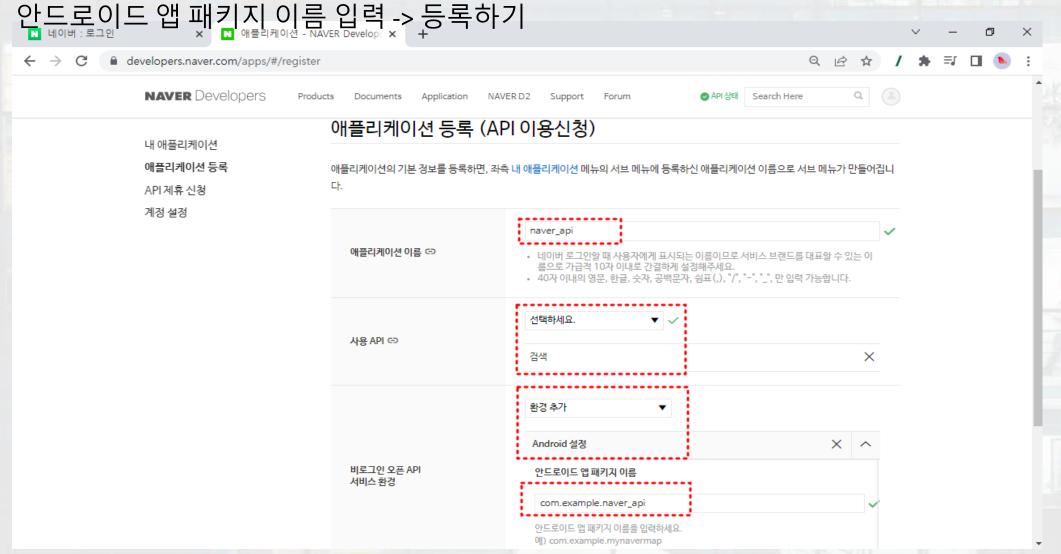
1. 네이버 개발자센터에서 Application-> 애플리케이션 등록





네이버 개발자 가입하기

2. 애플리케이션 이름입력/ 사용API('검색'선택)/비로그인 오픈 API(Android 설정)/





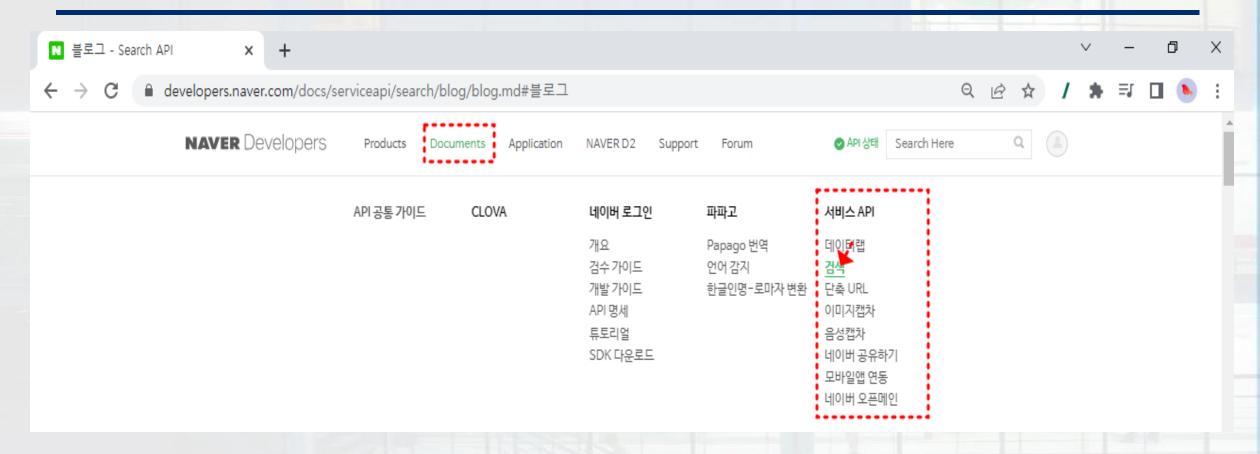
네이버 개발자 가입하기

3. Client ID/ ClientSecret는 남에게 노출되지 않게 합니다. 혹시 노출되면 초기화 합니다.

| → G | | nyapps/lbzft0g8Ql4mcLqZZPL6/ov | | | | Q 总 ☆ | / : | ≱ ≡ | _ |
|-----|------------------------|--------------------------------|---------------------|--------|-------------------|-------------------|-----|------------|---|
| | NAVER Developers | Products Documents Application | on NAVER D2 Support | Forum | ◆ API 상태 Search H | lere Q | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | 내 애플리케이션 | naver_api | | | | | | | |
| | eunPthon | 개요 AF | 기설정 멤버관리 | 로그인 통계 | API 통계 | Playground (Beta) | | | |
| | dBig API_test | 에프리케이너 캠브 | | | | | | | |
| | h_data | 애플리케이션 정보 | | | | | | | |
| | naver_api | Client ID | II. | | | | | | |
| | 애플리케이션 등록 API 제휴 신청 | | | | | | | | |
| | 계정 설정 | Client Secret | 보기 | | | | | | |
| | | | | | | | | | |
| | | ADI = 3 OUII | | | | | | | |
| | | API 호출 안내 | | | | | | | |

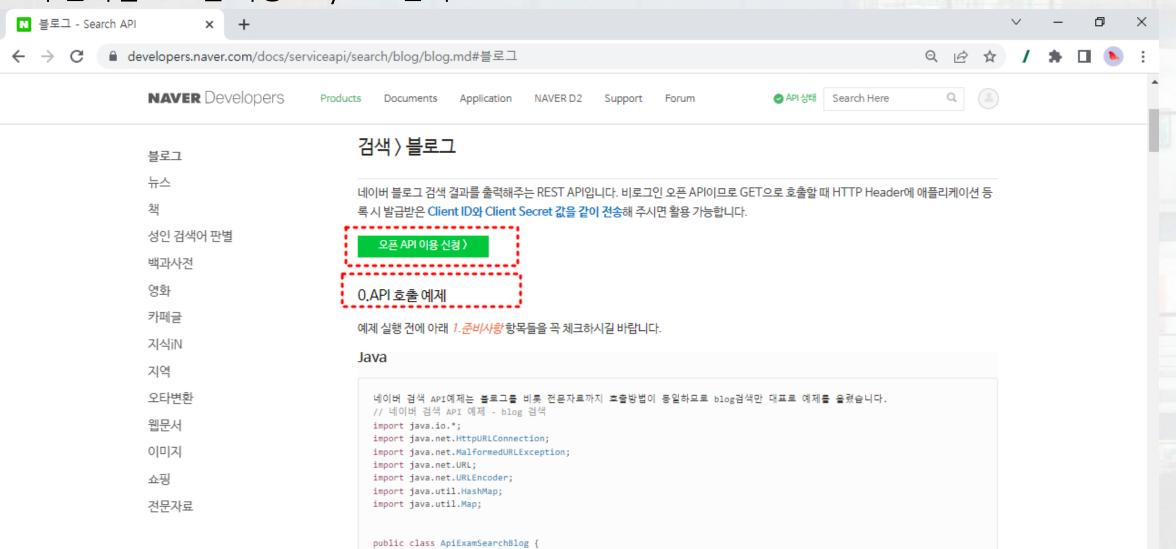


Documents -> 서비스 API -> '검색'



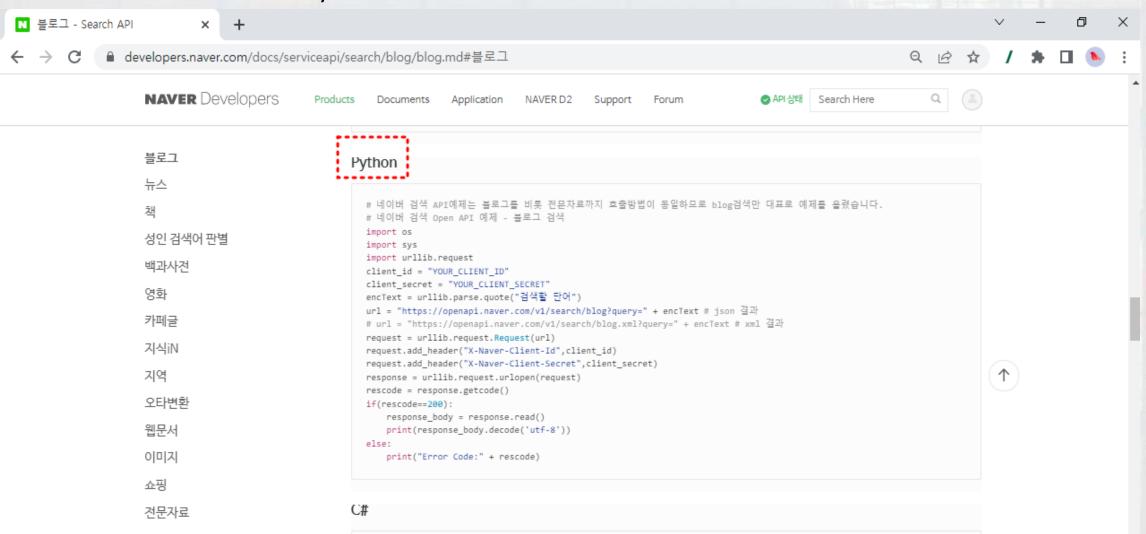
-- API 주소 알아보기

각 언어별로 호출 가능->'Python'선택



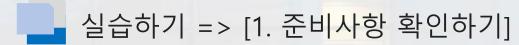
__ API 주소 알아보기

각 언어별로 호출 가능->'Python'선택



▲ 네이버에서 제공하는 API호출 'Python' 예제 코드 (처음 복사한 상태)

```
# 네이버 검색 Open API 예제 - 블로그 검색
import os
Import sys
                                      Client_id와 client_secret를 애플리케이션 등록하기에서 받은
Import urllib.request
                                      자신의 값을 넣어줍니다.
Client id = "YOUR CLIENT ID"
                                            urllib.parse.quote() => 인자로 주어진 문자열에서 특수 문자를
Client_secret = "YOUR_CLIENT_SECRET"
                                           문자열로 변환해서 반환
encText = urllib.parse.quote("검색할 단어")
url = "https://openapi.naver.com/v1/search/blog?query=" + encText # json 결과
# url = "https://openapi.naver.com/v1/search/blog.xml?query=" + encText # xml 결과
Request = urllib.request.Request(url)
Request.add header("X-Naver-Client-Id", client id)
Request.add_header("X-Naver-Client-Secret",client_secret)
Response = urllib.request.urlopen(request)
Rescode = response.getcode()
if(rescode==200):
  response body = response.read()
  print(response body.decode('utf-8'))
else:
  print("Error Code:" + rescode)
```



검색 〉 웹문서

네이버 웹 문서 검색 결과를 출력해주는 REST API입니다. 비로그인 오픈 API이므로 GET으로 호출할 때 HTTP Header에 애플리케이션 등록 시 발급받은 Client ID와 Client Secret 값을 같이 전송해 주시면 활용 가능합니다.

오픈 API 이용 신청 〉

1. 준비사항

 애플리케이션 등록: 네이버 오픈 API로 개발하시려면 먼저 'Application-애플리케이션 등록' 메뉴에서 애플리케이션을 등록하셔야 합니다.

[자세한 방법 보기] 〉

- · 클라이언트 ID와 secret 확인: '내 애플리케이션'에서 등록한 애플리케이션을 선택하면 Client ID와 Client Secret 값을 확인할 수 있습니다.
- · API 권한 설정: '내 애플리케이션'의 'API 권한관리' 탭에서 사용하려는 API가 체크되어 있는지 확인합니다. 체크되어 있지 않을 경우 403 에러 (API 권한 없음)가 발생하니 주의하시기 바랍니다.



→ 실습하기 => 2. API기본정보

2. API 기본 정보

| 메서드 | 인증 | 요청 URL | 출력 포맷 |
|-----|----|--|-------|
| GET | - | https://openapi.naver.com/v1/search/webkrx ml | XML |
| GET | - | https://openapi.naver.com/v1/search/webkr.js on | JSON |



→ 실습하기 => 3. 요청정보

3. 요청 변수

| 요청 변수명 | 타입 | 필수 여부 | 기본값 | 설명 |
|---------|---------|-------|------------------------|-----------------------------|
| query | string | Υ | _ | 검색을 원하는 문자열로서 UTF-8로 인코딩한다. |
| display | integer | N | 10(기본값), 10 0(최대) | 검색 결과 출력 건수 지정 |
| start | integer | N | 1 (기본값), 100 0 (최대) | 검색 시작 위치로 최대 1000까지 가능 |



→ 실습하기 => 4 출력결과

| 필드 | 타입 | 설명 | |
|---------------|----------|--|------------|
| rss | - | 디버그를 쉽게 하고 RSS 리더기만으로 이용할 수 있게 하기 위해 만든 RSS 포맷의 컨 테이너이며 그 외의 특별한 의미는 없다. | |
| channel | - | 검색 결과를 포함하는 컨테이너이다. 이 안에 있는 title, link, description 등의 항목은 참고용으로 무시해도 무방하다. | |
| lastBuildDate | datetime | 검색 결과를 생성한 시간이다. | |
| total | integer | 검색 결과 문서의 총 개수를 의미한다. | |
| start | integer | 검색 결과 문서 중, 문서의 시작점을 의미한다. | \uparrow |
| display | integer | 검색된 검색 결과의 개수이다. | |
| item/items | - | XML 포멧에서는 item 태그로, JSON 포멧에서는 items 속성으로 표현된다. 개별 검 색 결과이며 title, link, description을 포함한다. | |
| title | string | 검색 결과 문서의 제목을 나타낸다. 제목에서 검색어와 일치하는 부분은 태그로 감싸져 있다. | |
| link | string | 검색 결과 문서의 하이퍼텍스트 link를 나타낸다. | |
| description | string | 검색 결과 문서의 내용을 요약한 패시지 정보이다. 문서 전체의 내용은 link를 따라가면 읽을 수 있다. 패시지에서 검색어와 일치하는 부분은 태그로 감싸져 있다. | |

웹문서 결과보기

```
[[astBuildDate]: "Thu, 28 Apr 2022 12:35:55 +0900",
 "total": 7016158.
 start 1,
"display": 10.
 items:
"#ICED: "웹페이지에서 <b>단어</b> <b>검색</b>하는 방법: 7 단계 (이미지 포함) - wikiHow".
┇ 🗓 🖟 📜 "https://ko.wikihow.com/%EC%9B%B9%ED%8E%98%EC%9D%B4%EC%A7%8O%EC%97%9O%EC%84%9C-%EB%8B%A8%EC%96%B4-%EA%B2%8O%EC%83%
. "89%ED%95%98%EB%84%94-%EB%B0%49%EB%B2%95"
『description』: "웹페이지에서 <b>단어</b> <b>검색</b>하는 방법. 거의 모든 웹브라우저에는 <b>검색</b> 도구가 있어 페이지 내
'에서 얼치하는 <b>단어</b>나 문구를 찾을 수 있다. 그리고 구글의 고급 <b>검색</b>을 미용해 원하는 <b>단어</b>나 문구가 포함
【되어 있는 모든 웹페이지를 <b>검색할</b> 수 있다. <b>검색</b> 도구를 함께 사용해서 인터넷 상에서 원하는 <b>단어</b>를 찾아
낸다. . (윈도우즈) 또는 (맥) 누르기. 어느 브라우저에...
"title": "번뜻 (카메라로 <b>단어</b>를 <b>검색</b>하세요) - Google Play 앱",
"link": "https://plav.google.com/store/apps/details?id=com.beanzsoft.outsourcing.bunddet".
```

```
import os
  import sys
  import urllib.request
 4 import pandas as pd
  import json
   import re
  client_id = "lbzftOg8Ql4mcLqZZPL6"
   client_secret = "2208drQ854"
   query= urllib.parse.quote(input("검색 질의=>")) #검색질의
   lidx = 0
                  # 인덱스 변수(몇개인지 셀수 있는 변수)
   display = 100
                  ₩ 최대 100개 단위로 가져오기
  start = 1
16
   end = 1000
   web_df = pd.DataFrame(columns = ("Tile","Link", "Description")) # 컬럼 수 3개짜리 데이터프레일
19
20
   #반복 수행
21
   for start_index in range(start,end,display):
22
23
       url = "https://openapi.naver.com/v1/search/webkr?query=" +query #
24
             + "&display=" + str(display) #
25
             + "&start=" + str(start_index)
26
27
       request = urllib.request.Request(url)
28
       request.add_header("X-Naver-Client-Id",client_id)
29
       request.add_header("X-Naver-Client-Secret",client_secret)
30
       response = urllib.request.urlopen(request)
31
       rescode = response.getcode()
```

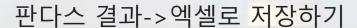
```
32
33.
       if(rescode==200):
                                                        # 정상 응답을 받는경우
34
          response_body = response.read()
35
36
          response_dict = json.loads(response_body.decode('utf-8')) # 되셔더리로 발아오기
37
38
          items = response_dict['items']
                                                      -#itams키에 해욧하는 값 가져오기
39
40.
          for item_index in range(0, len(items)):
              remove_tag = re.compile('<.*?>') # 테그 제외하기 위한 페틴 지점하기()
41
42
43.
              title = re.sub(remove_tag,'',items[item_index]['title']) #점규표현식을 ''로 없애기
44
              link = items[item_index]['link']
              description = re.sub(remove_tag, '', items[item_index]['description'])
45
46
47
              web_df.loc[idx] = [title,link,description] #리스트 활태로 인덱스에 넣기
48
              idx += 1
49
50
      else
51
          print("Error Code:" + rescode)
52
   web_df
```

- 판다스 결과보기

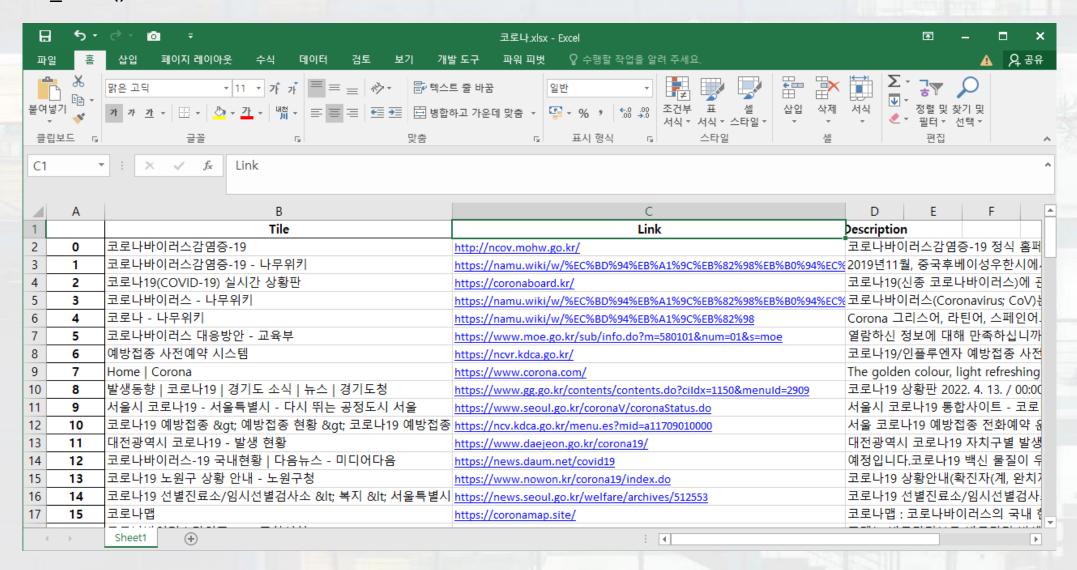
| 검색 | 질의=>코로나 |
|----|---------|
| | |

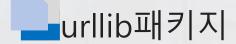
| | Tile | Link | Description |
|-------|---|---|--|
| 0 | 코로나바이러스감염증-19 | http://ncov.mohw.go.kr/ | 코로나바이러스감염증-19 정식 홈페이지로 발생현황, 국 내발생현황, 국외발생현황, 시 |
| 1 | 코로나바이러스감염증-19 - 나무위키 | https://namu.wiki/w/%EC%BD%94%EB%A1%9C%EB%82%9 | 2019년11월, 중국후베이성우한시에서 처음으로 발생하 여 보고된 새로운 유형의 변종 |
| 2 | 코로나19(COVID-19) 실시간 상황판 | https://coronaboard.kr/ | 코로나19(신종 코로나바이러스)에 관한 세계 각 국가들의 통계 및 뉴스 등을 취합하 |
| 3 | 코로나바이러스 - 나무위키 | https://namu.wiki/w/%EC%BD%94%EB%A1%9C%EB%82%9 | 코로나바이러스(Coronavirus; CoV)는 사람을 포함한 동 물에 광범위한 흐흡 |
| 4 | 코로나 - 나무위키 | https://namu.wiki/w/%EC%BD%94%EB%A1%9C%EB%82%98 | Corona 그리스어, 라틴어, 스페인어로 왕관을 뜻하는 단 어다. 태양 상충부 대기 |
| | | | |
| 115 | 코로나19 방역수칙 위반 신고센터<민원신청< 용답소 - 서울특별시 용답소 | https://eungdapso.seoul.go.kr/Req/Req02/Req02 | 코로나19 방역수칙 위반 신고센터 홈 > 민원신청 > 코로나19 방역수칙 |
| 116 | 코로나19 알림내역 코로나19 대용현황 코로나 19알림내역 코로나19 | https://www.gangdong.go.kr/web/coronamain/bbs/ | 등코로나19가 종식되는 그 날까지 강동구는 구민 여러분 의 생명과 등 코로나19 |
| 117 | #코로나 - YouTube | https://www.youtube.com/hashtag/%EC%BD%94%EB%A | 의료체계도 독감처럼코로나 확진자 격리 의무 사라진 다 뉴스A거리두기도 이제 끝? |
| 118 | <코로나 자가격리 지원금> 코로나 재택치료 지원금 (대상자, 신청방법, | https://dingdo.tistory.com/351 | 코로나 자가격리 지원금 (+재택치료 지원금) 코로나 자가 격리 생활지원금(신청방법 및 |
| 119 | 고용노동부> 뉴스·소식> 코로나19 대용 | https://www.moel.go.kr/policy/corona19/main.do | 4차 긴급고용안정지원금 Q&A(신규신청자)외국인 근로자 다수 사업장 코로나1 |
| 120 r | ows × 3 columns | | |

120 rows × 3 columns



to excel() 메서드를 사용





urllib 는 파이썬에서 제공하는 URL 작업을 위한 여려 모듈을 모은 패키지 urllib.request는 URL을 가져오기 위한 파이썬 모듈 urlopen()함수를 통해 결과를 가져올 수 있다.

문자열 바꾸기 : 파이써 sub메서드

Sub(substitute)

- sub메서드를 이용하면 정규식과 매치되는 부분을 다른 문자로 변경할 수 있다.

re.sub(찾을 패턴, 대체할 문자, 찾을 문자열)

- sub메서드를 첫 번째 입력 인수는 '바꿀 문자열', 두 번째 입력 인수는 '대상 문자열'
- <.*?> 가능한 한 가장 최소한의 반복 수행

```
for item_index in_range(0, len(items)):
    remove_tag = re.compile('<.*?>') # 明그 제외하기 위한 패턴 지정하기()

    title = re.sub(remove_tag, ",items[item_index]['title']) #정규표현식을 ''로 없에기
    link = items[item_index]['link']
    description = re.sub(remove_tag, ",items[item_index]['description'])

    web_df.loc[idx] = [title,link,description] #리스트 형태로 인덱스에 넣기
    idx += 1
```

