OPEN API 개요

- Open Application Programming Interface
- 특정 웹 사이트에서 자신이 가진 기능을 이용할 수 있도록 공개한 프로그래밍 인 터페이스
- 누구나 사용할 수 있도록 공개된 API로 응용 소프트웨어나 웹 서비스의 프로그 래밍적인 권한을 제공
- 대한민국 정부의 공공데이터포털(<u>www.data.go.kr</u>)을 통해 제공되는 API 서비스, 도로명 주소 조회 서비스, 동네예보정보조회서비스
- 네이버 지도, 구글맵, 코로나 맵 등
- JSON, XML 등으로 제공됨

- JavaScript Object Notation
- 키와 값으로 구성된 자바스크립트의 데이타 형태
- 파이썬의 JSON 레퍼런스 https://docs.python.org/ko/3/library/json.html
- JSON 예시

```
{
  "이름": "서태웅",
  "성별": "남",
  "학교": "북산 고등학교 1학년 10반 22번",
  "특기": "농구",
  "포지션": "스몰 포워드(SF)",
  "별명": "여우",
  "체격": {"키": "187 cm", "몸무게": " 75 kg", "혈액형": " AB형(Rh+)"},
  "라이벌": ["윤대협", "정우성"]
}
```

- JavaScript Object Notation
- 키와 값으로 구성된 자바스크립트의 데이타 형태
- 파이썬의 JSON 레퍼런스 https://docs.python.org/ko/3/library/json.html
- JSON 예시

```
{
  "이름": "서태웅",
  "성별": "남",
  "학교": "북산 고등학교 1학년 10반 22번",
  "특기": "농구",
  "포지션": "스몰 포워드(SF)",
  "별명": "여우",
  "체격": {"키": "187 cm", "몸무게": " 75 kg", "혈액형": " AB형(Rh+)"},
  "라이벌": ["윤대협", "정우성"]
}
```

- json 샘플 사이트 : http://jsonplaceholder.typicode.com/
- url 요청 => json 데이타 => 리스트 딕셔너리 => 데이타프레임

```
1 import os
2 import requests
3 import pandas as pd
4 import json
```

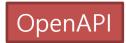
```
1 res = requests.get("https://jsonplaceholder.typicode.com/todos")
2 print(res)
3 json_str = res.text
4 print(json_str)
```

```
1 todos = json.loads(json_str)
2 print(type(todos), len(todos))
3 print(todos[0])
4 print(todos[-1])
5 print(todos[0:5])
```

```
1 cnt = 0
2 for i in range(0,len(todos)):
3    if todos[i]['userId']==1:
4        print(todos[i])
5        cnt += 1
6 print(f'총 갯수 - {cnt}')
7 print('-'*20)
```

```
1 cnt = 0
2 for i in range(0,len(todos)):
3    if todos[i]['userId']==1 and todos[i]['completed']==True:
4        print(todos[i])
5        cnt += 1
6    print('-' * 20)
7    print(f'총 갯수 - {cnt}')
```

```
1 df_todos = pd.DataFrame(todos, columns=['userId', 'id', 'title', 'completed'])
2 df_todos
```



- 1 df_todos.to_csv('output/todos.csv', index=False)
 1 ls output

userId, id, title, completed

- 1,1,delectus aut autem,False
- 1,2,quis ut nam facilis et officia qui,False
- 1,3,fugiat veniam minus,False
- 1,4,et porro tempora,True
- 1,5,laboriosam mollitia et enim quasi adipisci quia provident illum,False

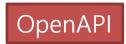
JSON 퀴즈

1) json 샘플 사이트를 이용하여 딕셔너리 리스트 형태로 변경하고 첫번째 회원의 데이타만 출력하여라

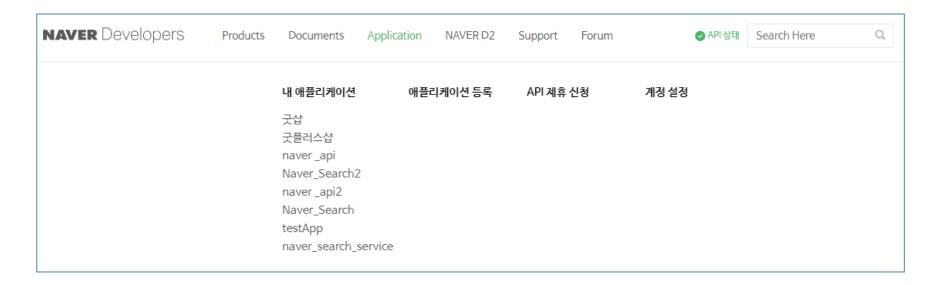
http://jsonplaceholder.typicode.com/users

- 2) 모두 몇명인가?
- 3) id, name, email 만 출력하여라
- 4) id, city, phone만 출력하여라
- 5) phone 이 '210.067.6132'인 회원은 누구인가?
- 6) name이 'C'로 시작하는 회원목록만 아래와 같은 형태로 출력하여라
- 7) 데이타프레임으로 변경하여라

	id	name	username	email	address	phone	website	company
0	1	Leanne Graham	Bret	Sincere@april.biz	{'street': 'Kulas Light', 'suite': 'Apt. 556',	1-770-736-8031 x56442	hildegard.org	{'name': 'Romaguera-Crona', 'catchPhrase': 'Mu
1	2	Ervin Howell	Antonette	Shanna@melissa.tv	<pre>{'street': 'Victor Plains',</pre>	010-692-6593 x09125	anastasia.net	<pre>{'name': 'Deckow-Crist', 'catchPhrase': 'Proac</pre>



■ 네이버 개발자 센터 사이트에서 오픈API 신청 - 네이버 아이디로 로그인 https://developers.naver.com/apps/#/register?defaultScope=search



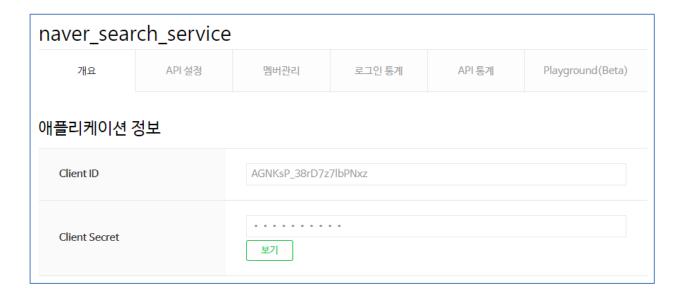


■ 어플리케이션 등록 : 검색/Web 설정 으로 신청





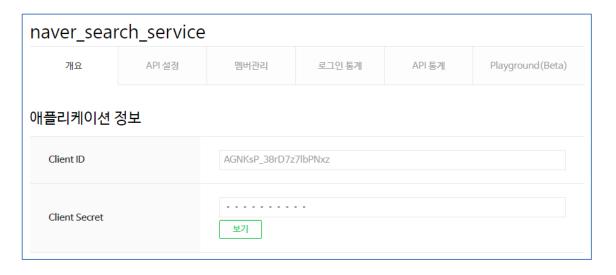
■ 네이버 서비스키와 비번 정보 확인 [Application]-[내 애플리케이션]





■ 파이썬 예시 소스 참조

https://developers.naver.com/docs/serviceapi/search/blog





Python 👄

```
# 네이버 검색 API 예제 - 블로그 검색
import os
import sys
import urllib.request
client_id = "YOUR_CLIENT_ID"
client_secret = "YOUR_CLIENT_SECRET"
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)
```

■ 블로그 검색 서비스 이용하기

```
1 import json
2 import pandas as pd
3 import requests
4 import urllib.request
```

```
1 # 본인 아이디와 시크릿 키 설정|
2 client_id = "본인 아이디"
3 client_secret = "시크릿 키"
```

```
1 encText = urllib.parse.quote("송도 맛집")
2 url = "https://openapi.naver.com/v1/search/blog?display=50&query=" + encText
3 print(url)
```

https://openapi.naver.com/v1/search/blog?display=50&query=%EC%86%A1%EB%8F%84%20%EB%A7%9B%EC%A7%91

```
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)
```



```
1 print(type(response body))
 2 blog_dict = json.loads(response_body.decode('utf-8'))
 3 print(type(blog_dict))
 4 blog dict
<class 'bytes'>
<class 'dict'>
{'lastBuildDate': 'Tue, 25 Oct 2022 23:08:50 +0900',
 'total': 396866.
 'start': 1,
 'display': 50.
 'items': [{'title': '<b>송도 맛집</b> 추천 푸짐한 송도 한정식 밥상편지',
  'link': 'https://blog.naver.com/joojoo5623/222901759267',
  'description': '제이유입니다 :) 오늘 소개해 드릴 <b>송도 맛집</b> 밥상편지는 송도 한정식을 맛볼 수 있
는 곳이에요. 이전에 한번 다녀온 적이 있는데 맛있게 먹었던 기억이 있어 부모님 두 분을 모시고 다녀왔어요 밥상
편지 주소... '.
  'bloggername': '바삭바삭한 일상그리기\',
  'bloggerlink': 'blog.naver.com/joojoo5623',
  'postdate': '20221017'}.
 1 blog_list = blog_dict['items']
                                        Ten(blog list)
 2 print(type(blog_list))
                                    50
<class 'list'>
```



```
1 blog list[0]
{'title': '<b>송도 맛집</b> 추천 푸짐한 송도 한정식 밥상편지'.
 'link': 'https://blog.naver.com/joojoo5623/222901759267',
'description': '제이유입니다 :) 오늘 소개해 드릴 <b>송도 맛집</b> 밥상편지는 송도 한정식을 맛볼 수 있는 곳
이에요. 이전에 한번 다녀온 적이 있는데 맛있게 먹었던 기억이 있어 부모님 두 분을 모시고 다녀왔어요 밥상편지 주
소... '.
 'bloggername': '바삭바삭한 일상그리기\',
 'bloggerlink': 'blog.naver.com/joojoo5623',
 'postdate': '20221017'}
 1 df = pd.DataFrame(blog_list)
 2 df.head()
    title
                                      link
                                             description bloggername
                                                                           bloggerlink postdate
   <b>송
  도 맛집
                                           제이유입니다:)
  </b> 추
                                           오늘 소개해 드
  천 푸짐
                                           릴 <b>송도 맛 바삭바삭한 일
         https://blog.naver.com/joojoo5623/222901759267
                                                                 blog.naver.com/joojoo5623 20221017
  한 송도
                                           집</b> 밥상편
                                                        상그리기₹
   한정식
                                           지는 송도 한정
   밥상편
                                                식을...
     지
```

```
result list = []
2 for item in blog_list:
      title = item['title'].replace('<b>','').replace('</b>','')
      description = item['description'].replace('<b>','').replace('</b>','')
4
      bloggername = item['bloggername']
      postdate = item['postdate']
      result_list.append([title, description, bloggername, postdate])
1 df = pd.DataFrame(result_list, columns=['제목', '내용', '블로거', '등록날짜'])
2 df.head()
                      제목
                                                                  내용
                                                                                블로거 등록날짜
 송도 맛집 추천 푸짐한 송도 한정식 밥
                           제이유입니다:) 오늘 소개해 드릴 송도 맛집 밥상편지는 송도 한정식
                                                                                     20221017
                     상편지
                                                            을 맛볼 수 있...
   마담 샤브칼국수 막국수 송도맛집 탕 면을 좋아하시는 분들이라면 꼭 드시면 너무 좋을 것 같아요! #송도샤 건강한 하루를 보내는
                                                                                     20221023
                   수육맛집
                                                          브샤브 #송도생...
                            친구들과 평일 휴가를 맞춰 현지인이 추천해준 부산 송도 맛집에 다
        호감이었던 부산 송도 맛집
2
                                                                           고테네 사진관 20220716
                                                        녀왔어요. 칼칼하면...
                            며칠전 친구와 함께 송도로 나들이를 떠났다가 미리 찾아둔 부산 송
3
      부산 송도 맛집 제대로 찾은 곳
                                                                           Ranco Holic 20221020
                                                        도 맛집에 방문했어...
```

df.to csv('output/송도맛집.csv')

161



■ 지식인 검색 서비스 이용하기

https://developers.naver.com/docs/serviceapi/search/kin/kin.md#%EC%A7

%80%EC%8B%9DiN



```
1 import json
2 import pandas as pd
3 import requests
4 import urllib.request
```

```
1 # 본인 아이디와 시크릿 키 설정 |
2 client_id = "본인 아이디"
3 client_secret = "시크릿 키"
```

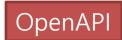
```
encText = urllib.parse.quote("해운대 코스")
url = "https://openapi.naver.com/v1/search/kin.json?sort=date&display=50&query=" + encText
print(url)
```

```
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)
```

```
know_dict = json.loads(response_body.decode('utf-8'))
print(type(know_dict))
know_list = know_dict['items']
print(type(know_list))
```

```
know list[0]
{'title': '부산 오렛만에 가서',
 'link': 'https://kin.naver.com/qna/detail.naver?d1id=9&dirld=90111&docld=431220652&qb=7Z\07Jq064yAl
0y910yKpA==&enc=utf8&section=kin.gna&rank=1&search_sort=2&spg=0'.
 'description': '부산여행 당일 <b>코스</b> 좀 잡아주세요 . 오전10시부터 오후6시까지 있지 싶습니다. 당일치
기면 기장까지는 무리이실거같고 광안리.<b>해운대</b> / 남포. 영도 이쪽중에 고르셔야 할거같네요 광안리를 간다하
면 광안리 바닷가, 광안대교... '}
  1 df = pd.DataFrame(know_list)
  2 df.sample(2)
                      title
                                                       link
                                                                                        description
     2박 3일 부산여행 <b>코스</b>
                               https://kin.naver.com/gna/detail.naver?
                                                            ... 부산의 대표적인 <b>해운대</b>, 광안리, 송정 해수욕
               질문드립니다 ㅎ
                                                   d1id=9&...
                                                                                     장 <b>코스</b...
                               https://kin.naver.com/gna/detail.naver?
                                                            ... 생각나요 직장을 부산에서 잡고 정착하면 되죠~~ 환상
         부산 환상 좀 깨주세요ㅠ
  7
                                                   d1id=9&...
                                                                              을 현실로 만드는겁니다~...
     result_list = []
  2 for item in know list:
        title = item['title'].replace('<b>','').replace('</b>','')
        description = item['description'].replace('<b>','').replace('</b>','')
  4
        link = item['link']
  5
        result list.append([title, description, link])
```



1 2	<pre>df = pd.DataFrame(resul df.sample(2)</pre>	'내용', '링크'])		
	제목		내용	링크
40	cj택배배송 언제 올까요?	04 16:38:40 보내시는 고객님으 	로부터 상품을 인수받 았습니다 부산해운대	https://kin.naver.com/qna/detail.naver? d1id=8&
39	부산 맛집 (갑각류맛집 위주로) 추천부탁드려요!!	해운대랍스터맛집 덴포라 추천드 인데	려요. !!! 코스요리맛집 모임맛집으로도 유	https://kin.naver.com/qna/detail.naver? d1id=12
1	df.to_csv('output/해운다	#지식인검색결과.csv')		
1	Is output₩*.csv			



공공데이타포탈의 API 서비스

- https://www.data.go.kr/
- 회원 가입후에야 API서비스 신청 가능





- 환경부 국립환경과학원
- 조회날짜, 항목코드, 측정소코드, 시간구분을 기준으로 중금속 성분 측정 결과를
 2시간 평균, 24시간 평균 측정 수치 자료로 제공하는 서비스
- https://www.data.go.kr/tcs/dss/selectApiDataDetailView.do?publicDataPk=15016368
- [활용신청] 클릭후 서비스키 정보 확인 : [마이페이지]-[오픈 API]-[개발 계정]

환경부 국립환경과학원_미세먼지(금속성분) 실시간 정보 국립환경과학원 대기연구분야 데이터 입니다.(대기중 중금속 성분 2시간, 24시간 평균 측정 자료) ☑ 3 ☑ ☑ 0 ☐ 교심



요청변수(Request Parameter)

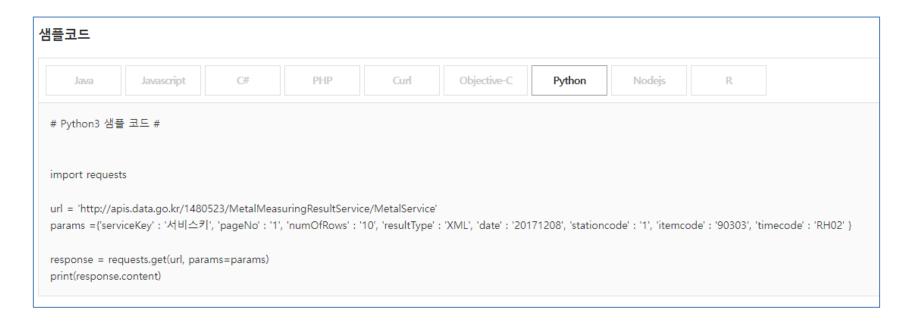
항목명(국문)	항목명(영문)	항목크기	항목구분	샘플데이터	항목설명		
서비스키	ServiceKey	4	필수	-	공공데이터포털에서 받은 인증키		
페이지 번호	pageNo	4	옵션	1	페이지번호		
한 페이지 결과 수	numOfRows	4	옵션	10	한 페이지 결과 수		
결과형식	resultType	4	옵션	XML	결과형식(XML/JSON)		
검색조건 날짜	date	8	옵션	20171208	검색조건 날짜 (예시 : 20171208)		
검색조건 측정소 코드	stationcode	6	옵션	1	검색조건 측정소코드		
검색조건 항목코드	itemcode	5	옵션	90303	검색조건 항목코드		
검색조건 시간구분	timecode	4	옵션	RH02	검색조건 시간구분		



출력결과(Response Element)

항목명(국문)	항목명(영문)	항목크기	항목구분	샘플데이터	항목설명		
결과코드	resultCode	2	필수	00	결과코드		
결과메시지	resultMsg	50	필수	OK	결과메시지		
한 페이지 결과 수	numOfRows	4	필수	10	한 페이지 결과 수		
페이지 번호	pageNo	4	필수	1	페이지번호		
전체 결과 수	totalCount	4	필수	3	전체 결과 수		
날짜	sdate	14	필수	20171208040000	날짜		
측정소 코드	stationcode	1	필수	1	측정소 코드		
측정항목 코드	itemcode	5	필수	90303	측정항목 코드		
측정데이터 시간 구분	timecode	4	필수	RH02	측정데이터 시간 구분		
측정수치	value	8	필수	7.5	측정수치(단위:ng/m3)		





```
import json
import pandas as pd
import requests
import urllib.request
```

```
1 url = 'http://apis.data.go.kr/1480523/MetalMeasuringResultService/MetalService'
2 service_key = 시크릿키값
   params ={'serviceKey' : service_key,
             'pageNo' : '1',
            'numOfRows' : '12',
             'resultType' : 'JSON',
             'date' : '20220101',
            'stationcode' : '3',
             'itemcode' : '90303',
             'timecode' : 'RH02'}
10
11
   response = requests.get(url, params=params)
12
   print(response)
   print(response.text)
```



- 1 dust = json.loads(response.content)
- 2 dust
- 1 dust['MetalService']['item']



```
df_dust = pd.DataFrame(dust['MetalService']['item'], columns=['SDATE', 'STATIONCODE', 'ITEMCODE', 'TIMECODE', 'VALUE'])
df_dust
```

0 202201010000000 3 90303 RH02 1 20220101020000 3 90303 RH02 2 20220101040000 3 90303 RH02 3 20220101060000 3 90303 RH02 4 20220101080000 3 90303 RH02	VALUE
2 20220101040000 3 90303 RH02 3 20220101060000 3 90303 RH02	3.2933
3 202201010600000 3 90303 RH02	2.3088
	3.7324
4 20220101080000 3 90303 RH02	6.3342
	6.3293
5 202201011000000 3 90303 RH02	6.4359
6 202201011200000 3 90303 RH02	4.2071
7 20220101140000 3 90303 RH02	3.4410
8 202201011600000 3 90303 RH02	2.9664
9 202201011800000 3 90303 RH02	48.9368
10 202201012000000 3 90303 RH02	112.6491

station code 1~10 레코드 저장 1 수도권 2 백령도 3 호남권 4 중부권 5 제주권 6 영남권 7 경기권 8 충청권 9 전북권 10 강원권

```
1 url = 'http://apis.data.go.kr/1480523/MetalMeasuringResultService/MetalService'
2 service_key = 서비스키
   dust_list = []
   for i in range(1, 11):
       params ={'serviceKey' : service_key,
 6
                 'pageNo' : '1'.
                 'numOfRows' : '12',
 8
                 'resultType' : 'JSON',
                 'date' : '20220510',
10
                 'stationcode' : str(i),
11
                 'itemcode' : '90303',
12
                 'timecode' : 'RH02'}
13
14
15
       response = requests.get(url, params=params)
16
       try:
           temp = json.loads(response.text)
17
18
           dust_list += temp['MetalService']['item']
19
       except:
20
            pass
21
       print(i)
```



53 20220510100000

6

90303

RH02 4.6156

퀴 즈

- 고속도로 휴게소별 날씨 정보를 제공하는 API 서비스 신청하기 https://www.data.go.kr/data/15076661/openapi.do
 http://data.ex.co.kr/openapi/basicinfo/openApiInfoM?apiId=0508
- 신청한 api 서비스를 이용하여 특정 날짜(예시-2022년10월15일)의 날씨 데이타를 리스트 딕셔너리 형태로 저장하여라.
- 경부선에 위치한 휴게소의 날씨 데이터만 출력하여라
- 판다스의 데이프레임 구조로 저장하고 출력하여라
- 2022년 10월 1~5 일간의 데이타를 API 서비스를 이용하여 추출하고 데이타프레임으로 저장하여라

웹페이지 명세서 (인증키 확인)

http://data.ex.co.kr/openapi/basicinfo/openApiInfoM?apiId=0508&pn=-1



퀴즈

	unitName	unitCode	sdate	stdHour	routeNo	routeName	updownTypeCode	xValue	yValue
0	죽전휴게 소	002	20221001	10	0010	경부선	E	127.104165	37.332651
1	안성휴게 소(서울)	004	20221001	10	0010	경부선	E	127.132081	37.076107
2	옥산휴게 소(서울)	031	20221001	10	0010	경부선	E	127.371859	36.657168
3	천안(삼) 휴게소(서 울)	033	20221001	10	0010	경부선	E	127.173092	36.787388
933	벌곡휴게 소(논산)	043	20221005	10	2510	호남지선	s	127.273793	36.217488
934	현풍휴게 소(현풍)	090	20221005	10	4510	중부내륙지 선	s	128.436431	35.701990

935 rows × 38 columns



XML 개요

- eXtensible Markup Language
- 다목적의 성격을 가진 마크업언어
- XML은 미리 정의된 태그 집합이나 문법을 규정하고 있지 않기 때문에 완전한 확장성을 가진다.
- XML 문서 작성자는 원하는 태그를 사용할 수 있으며, 태그에 원하는 속성을 지정할 수 있으며, 원하는 형태로 태그를 중첩시킬 수 있다.
- https://www.w3schools.com/xml/xml_tree.asp



XML 예시

```
# 짝이 있는 태그 형태

<태그 속성=값> # 예시 - 노트

내용 <?xml version="1.0" encoding="UTF-8"?>

</태그> <note>

# 짝이 없는 태그 형태 : 내용이 없음 <from>Jani</from>
</태그 /> <heading>Reminder</heading>
<body>Don't forget me this weekend!</body>
</note>
```



BS4를 이용한 XML 파싱

```
1 import requests
2 from bs4 import BeautifulSoup
3 import pandas as pd
```

```
with open('bookstore.xml', 'r') as f:
    xml_str = f.read()
    print(type(xml_str))
    print()
    print(xml_str)
```

```
1 soup = BeautifulSoup(xml_str, 'xml')
2 soup
```

```
<?xml version="1.0" encoding="UTF-8"?>
<bookstore>
  <book category="cooking">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <vear>2005</vear>
    <price>30.00</price>
 </hook>
  <book category="children">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
  <book category="web">
    <title lang="en">Learning XML</title>
    <author>Erik T. Rav</author>
    <vear>2003</vear>
    <price>39.95</price>
 </hnnk>
</bookstore>
```



BS4를 이용한 XML 파싱

```
1 target_tag = soup.find('book')
2 print(target_tag)
3 print()
4 print(target_tag.text)
5 print()
6 print(target_tag['category'])
7 print(target_tag.get('category'))
```

```
print(target_tag.title)
print(target_tag.title.text)
print(target_tag.title['lang'])
```

```
<title lang="en">Everyday Italian</title>
Everyday Italian
en
```

```
<book category="cooking">
<title lang="en">Everyday Italian</title>
<author>Giada De Laurentiis</author>
<year>2005</year>
<price>30.00</price>
</book>

Everyday Italian
Giada De Laurentiis
2005
30.00

cooking
cooking
```

```
print(target_tag.find('title'))
print(target_tag.find('title').text)
print(target_tag.find('title')['lang'])
```

```
<title lang="en">Everyday Italian</title>
Everyday Italian
en
```



BS4를 이용한 XML 파싱

```
for book in target list:
   target_list = soup.find_all('book')
                                                          print('category => ', book['category'])
   len(target_list)
                                                    3
                                                          print('title => ', book.title.text)
                                                          print('author => ', book.author.text)
                                                    4
                                                          print('year => ', book.year.text)
                                                    5
   target_list[0]
                                                          print('price => ', book.price.text)
                                                    6
                                                          print('='*50)
                                                    7
<book category="cooking">
<title lang="en">Everyday Italian</title>
<author>Giada De Laurentiis</author>
<vear>2005</vear>
```

```
book_list = []
for book in target_list:
    category = book['category']
    title = book.title.text
    author = book.author.text
    year = book.year.text
    price = book.price.text
    book_list.append([category, title, author, year, price])

book_list
```

<price>30.00</price>

</book>

```
[['cooking', 'Everyday Italian', 'Giada De Laurentiis', '2005', '30.00'], ['children', 'Harry Potter', 'J K. Rowling', '2005', '29.99'], ['web', 'Learning XML', 'Erik T. Ray', '2003', '39.95']]
```



BS4를 이용한 XML 파싱

	category	title	author	year	price
0	cooking	Everyday Italian	Giada De Laurentiis	2005	30.00
1	children	Harry Potter	J K. Rowling	2005	29.99
2	web	Learning XML	Erik T. Ray	2003	39.95

```
print(soup.find(lang="en"))
print(soup.find_all(lang="en"))
print()
print(soup.find('title', {'lang':'en'}))
print(soup.find('book', {'lang':'en'}))

<title lang="en">Everyday Italian</title>
[<title lang="en">Everyday Italian</title>, <title lang="en">Harry Potter</title>, <title lang="en">Learning XML</title>]
</title lang="en">Everyday Italian</title></title>
None
```



country.xml 파일을 open() 함수를 이용하여 불러오기한 후 파싱 과정을 거쳐 아래와 같은 데이타프레임으로 변경하여라 neighbor 컬럼의 값은 리스트안의 튜플구조를 이용한다

neighbor	gdppc	year	rank	name	
[(Austria, E), (Switzerland, W)]	141100	2008	1	Liechtenstein	0
[(Malaysia, N)]	59900	2011	4	Singapore	1
[(Costa Rica, W), (Colombia, E)]	13600	2011	68	Panama	2

xml 소스

https://docs.python.org/3/library/xml.etree.elementtree.html



food_menu.xml파일을 open() 함수를 이용하여 불러오기한 후 파싱 과정을 거쳐 아래와 같은 데이타프레임으로 변경하여라

	메뉴명	가격	칼로리	메뉴 설명
0	Belgian Waffles	\$5.95	650 cal	Two of our famous Belgian Waffles with plenty
1	Strawberry Belgian Waffles	\$7.95	900 cal	Light Belgian waffles covered with strawberrie
2	Berry-Berry Belgian Waffles	\$8.95	900 cal	Light Belgian waffles covered with an assortme
3	French Toast	\$4.50	600 cal	Thick slices made from our homemade sourdough
4	Homestyle Breakfast	\$6.95	950 cal	Two eggs, bacon or sausage, toast, and our eve



기상청 전국날씨 RSS

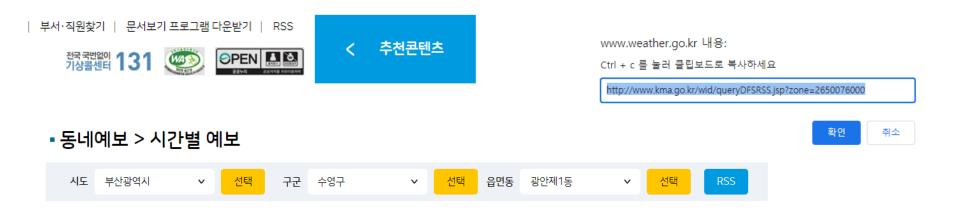
RSS(Really Simple Syndication, Rich Site Summary)란?

블로그처럼 컨텐츠 업데이트가 자주 일어나는 웹사이트에서, 업데이트된 정보를 쉽게 구독자들에게 제공하기 위해 XML을 기초 로 만들어진 데이터 형식



- https://www.weather.go.kr/w/index.do 하단 [RSS] 메뉴 클릭
- [동네예보 > 시간별 예보] 에서 지역 지정 후 [RSS] 클릭 예시) 부산광역시 수영구 광안1동
- 주소 복사: http://www.kma.go.kr/wid/queryDFSRSS.jsp?zone=2650076000
- XML 가이드 =>

https://www.kma.go.kr/images/weather/lifenindustry/timeseries_XML.pdf





```
import requests
  from bs4 import BeautifulSoup
   import pandas as pd
   url = 'http://www.kma.go.kr/wid/quervDFSRSS.isp?zone=2650076000'
   res = requests.get(url)
   xml_str = res.text
   xml_str
    soup = BeautifulSoup(xml_str, 'xml')
    soup
    target list = soup.find all('data')
    len(target list)
17
```

```
target list[0]
<data seg="0">
<hour>24</hour>
<day>0</day>
<temp>15.0</temp>
< tmx > -999.0 < /tmx >
< tmn > -999.0 < /tmn >
\langle sky \rangle 4 \langle /sky \rangle
<pty>0</pty>
<wfKor>흐림</wfKor>
<wfEn>Cloudy</wfEn>
<pop>30</pop>
<r12>0.0</r12>
<s12>0.0</s12>
<ws>3.7</ws>
< wd > 0 < / wd >
<wdKor>북</wdKor>
<wdEn>N</wdEn>
<reh>75</reh>
<r06>0.0</r06>
<s06>0.0</s06>
</data>
```



```
1 day_dict = {'0':'오늘', '1':'내일', '2':'모레'}
2 day = day_dict[target_list[0].day.text]
3 hour = target_list[0].hour.text
4 temp = target_list[0].temp.text

5 
6 
7 sky_dict = {'1':'맑음', '2':'구름조금', '3':'구름많음', '4':'흐림'}
8 sky = sky_dict[target_list[0].sky.text]

9 
10 wf = target_list[0].wfKor.text + '/' + target_list[0].wfEn.text

11 
12 pty_dict = {'0':'없음', '1':'비', '2':'비/눈', '3':'눈/비', '4':'눈'}
13 pty = pty_dict[target_list[0].pty.text]

14 
15 pop = target_list[0].pop.text + ' %'
16 ws = target_list[0].ws.text
```

```
1 print(soup.category.text)
2 print()
3 print('날짜 시간 기온 하늘상태 날씨상태 강수상태 강수확률 풍속(m/s)')
4 print('='*80)
5 print(f'{day} {hour:>3} {temp:>5} {sky:>5} {wf:>11} {pty:>5} {pop:>7}
6 {float(ws):10.2f}')
```

부산광역시 수영구 광안제1동



```
wether list = []
2 day dict = {'0':'오늘', '1':'내일', '2':'모레'}
3 sky_dict = {'1':'맑음', '2':'구름조금', '3':'구름많음', '4':'흐림'}
   pty dict = {'0':'없음', '1':'비', '2':'비/눈', '3':'눈/비', '4':'눈'}
  for w in target_list:
       dav = dav dict[w.dav.text]
       hour = w.hour.text
8
       temp = w.temp.text
       sky = sky dict[w.sky.text]
10
       wf = w.wfKor.text + '/' + w.wfEn.text
11
12
       pty = pty_dict[w.pty.text]
       pop = w.pop.text + ' %'
13
       ws = f'\{f|oat(w.ws.text):.2f\}'
14
       wether_list.append([day, hour, temp, sky, wf, pty, pop, ws])
```

```
1 df = pd.DataFrame(wether_list,
2 columns=['날짜', '시간','기온', '하늘상태',
3 '날씨상태','강수상태','강수률','풍속(m/s)'])
4 df
```

	날짜	시간	기온	하늘상태	날씨상태	강수상태	강수률	풍속(m/s)
0	오늘	24	15.0	흐림	흐림/Cloudy	없음	30 %	3.70
1	내일	3	15.0	흐림	흐림/Cloudy	없음	30 %	4.00
2	내일	6	15.0	흐림	흐림/Cloudy	없음	30 %	3.90
3	내일	9	17.0	흐림	흐림/Cloudy	없음	30 %	3.70



- https://www.weather.go.kr/w/index.do 하단 [RSS] 메뉴 클릭
- [동네예보 > 중기 예보] 에서 해당 지역의 [RSS] 클릭
- 주소 복사 : 서울.경기도 예시

https://www.kma.go.kr/weather/forecast/mid-term-rss3.jsp?stnId=109

- XML 가이드

https://www.weather.go.kr/w/resources/pdf/midtermforecast_rss.pdf

- 동네예보 > 중기예보

	전국	RSS	전라북도	RSS
	서울·경기도	RSS	전라남도	RSS
중기 예보	강원도	RSS	경상북도	RSS
	충청북도	RSS	경상남도	RSS
	충청남도	RSS	제주특별자치도	RSS



```
1 url = 'http://www.kma.go.kr/weather/forecast/mid-term-rss3.jsp?stnld=109'
           2 res = requests.get(url)
           3 \times ml_str = res.text
             soup = BeautifulSoup(xml_str, 'xml')
            soup
              target_list = soup.find_all('location')
              len(target list)
         35
             target_list[0]
                                                                   target_data_list[0]
    target_data_list = target_list[0].find_all('data')
                                                               <data>
                                                               <mode>A02</mode>
    len(target_data_list)
                                                               <tmEf>2022-10-29 00:00</tmEf>
13
                                                               <wf>맑음</wf>
                                                               <tmn>10</tmn>
                                                               < tmx > 18 < / tmx >
                                                               <reliability/>
                                                               <rnSt>0</rnSt>
                                                               </data>
```



```
province = target_list[0].province.text
city = target_list[0].city.text

tmEf = target_list[0].find_all('data')[0].tmEf.text
wf = target_list[0].find_all('data')[0].wf.text

tmn = target_list[0].find_all('data')[0].tmn.text
tmx = target_list[0].find_all('data')[0].tmx.text
rnSt = target_list[0].find_all('data')[0].rnSt.text

print('지역 : ', province)
print('도시 : ', city)
print('날짜 : ', tmEf)
print('날짜 : ', tmEf)
print('실지온도 : ', tmn)
print('청고온도 : ', tmx)
print('강수확률 : ', rnSt, '%')
```

지역 : 서울・인천・경기도

도시 : 서울

날짜 : 2022-10-29 00:00

날씨 : 맑음 최저온도 : 10 최고온도 : 18 강수확률 : 0 %



```
province = target_list[0].province.text
city = target_list[0].city.text
tmEf = target_list[0].find_all('data')[1].tmEf.text
wf = target_list[0].find_all('data')[1].wf.text
tmn = target_list[0].find_all('data')[1].tmn.text
tmx = target_list[0].find_all('data')[1].tmx.text
rnSt = target_list[0].find_all('data')[1].rnSt.text

print('지역 : ', province)
print('도시 : ', city)
print('날짜 : ', tmEf)
print('날짜 : ', tmEf)
print('날짜 : ', tmn)
print('참고온도 : ', tmn)
print('참고온도 : ', tmx)
print('강수확률 : ', rnSt, '%')
```

지역 : 서울・인천・경기도

도시 : 서울

날짜 : 2022-10-29 12:00

날씨 : 구름많음 최저온도 : 10 최고온도 : 18 강수확률 : 20 %



2 3

기상청 전국날씨 RSS - 중기 예보

```
midterm forecast list = []
       for location in target_list:
           province = location.province.text
           city = location.city.text
     6
            target data list = location.find all('data')
           for data in target data list:
    8
    9
                tmEf = data.tmEf.text
                wf = data.wf.text
    10
                tmn = data.tmn.text
    11
    12
                tmx = data.tmx.text
    13
                rnSt = data.rnSt.text
    14
    15
                midterm_forecast_list.append([province, city, tmEf, wf, tmn, tmx, rnSt])
  df = pd.DataFrame(midterm_forecast_list,
                   columns=['지역', '도시', '날짜', '날씨', '최저기온', '최고기온', '강수률'])
4 df
```

	지역	도시	날짜	날씨	최저기온	최고기온	강수률
0 서울・인천	・경기도	서울	2022-10-29 00:00	맑음	10	18	0
1 서울ㆍ인천	・경기도	서울	2022-10-29 12:00	구름많음	10	18	20
2 서울・인천	・경기도	서울	2022-10-30 00:00	맑음	9	18	0



1) 동네예보 > 중기예보 > 전국 에서 도시명(city 태그 이용)만 출력하여라 [중기예보]-[전국]

http://www.kma.go.kr/weather/forecast/mid-term-rss3.jsp?stnId=108

['서울', '인천', '수원', '파주', '이천', '평택', '춘천', '원주', '강릉', '대전', '세종', '홍성', '청주', '충주', '영동', '광주', '목포', '여수', '순천', '광양', '나주', '전주', '군산', '정읍', '남원', '고창', '무주', '부산', '울산', '창원', '진주', '거창', '통영', '대구', '안동', '포항', '경주', '울진', '울릉도', '제주', '서귀포']

2) '동네예보 > 중기예보 > 전국' 의 날씨 데이타를 아래와 같은 형태로 데이타프레임

으로 변경하여라

	도시	날짜	날씨	최저기온	최고기온	강수률
0	서울	2022-06-10 00:00	구름많음	18	28	40
1	서울	2022-06-10 12:00	구름많음	18	28	30
2	서울	2022-06-11 00:00	구름많음	19	29	30
3	서울	2022-06-11 12:00	구름많음	19	29	40
4	서울	2022-06-12 00:00	구름많음	19	28	20



3) '동네예보 > 중기예보 > 전국' 에서 부산, 대구의 날씨 데이타만 아래와 같은 형태로 데이타프레임으로 변경하여라

	도시	날짜	날씨	최저기온	최고기온	강수률
0	부산	2022-06-10 00:00	구름많음	18	24	20
1	부산	2022-06-10 12:00	맑음	18	24	10
2	부산	2022-06-11 00:00	구름많음	18	25	20
3	부산	2022-06-11 12:00	구름많음	18	25	20
4	부산	2022-06-12 00:00	구름많음	19	25	20
5	부산	2022-06-12 12:00	구름많음	19	25	30
6	부산	2022-06-13 00:00	구름많음	18	25	30

19	대구	2022-06-13 00:00	구름많음	18	28	30
20	대구	2022-06-13 12:00	흐림	18	28	40
21	대구	2022-06-14 00:00	흐림	18	26	40
22	대구	2022-06-14 12:00	흐림	18	26	40
23	대구	2022-06-15 00:00	흐림	19	27	40
24	대구	2022-06-16 00:00	구름많음	18	29	30
25	대구	2022-06-17 00:00	구름많음	20	29	30



- 부산관광명소의 이름, 주소, 홈페이지, 전화번호, 좌표, 상세정보를 조회하는 서비스
- https://www.data.go.kr/data/15063481/openapi.do
- [활용신청] 클릭후 서비스키 정보 확인 [마이페이지]-[오픈 API]-[개발 계정]
- · 활용승인 절차 개발단계 : 자동승인 / 운영단계 :
- · 신청가능 트래픽 개발계정: 10,000 / 운영계정: 활용사례 등록시 신청하면 트래픽 증가 가능
- ・요청주소 http://apis.data.go.kr/6260000/AttractionService/getAttractionKr
- ・서비스URL http://apis.data.go.kr/6260000/AttractionService

요청변수(Request Parameter)

항목명(국문)	항목명(영문)	항목크기	항목구분	샘플데이터	항목설명
서비스키	ServiceKey	4	필	-	공공데이터포털에서 받은 인증키
페이지 번호	pageNo	4	필	1	페이지번호
한 페이지 결과 수	numOfRows	4	필	10	한 페이지 결과 수
JSON방식 호출	resultType	4	옵		JSON방식으로 호출 시 파라미터 resultType=json 입력
콘텐츠 ID	UC_SEQ	11	옵		콘텐츠 ID



출력결과(Response Element)

항목명(국문)	항목명(영문)	항목크기	항목구분	샘플데이터	항목설명
결과코드	resultCode	2	필	00	결과코드
결과메시지	resultMsg	50	필	OK	결과메시지
한 페이지 결과 수	numOfRows	4	필	10	한 페이지 결과 수
페이지 번호	pageNo	4	필	1	페이지번호
전체 결과 수	totalCount	4	필	107	전체 결과 수
콘텐츠ID	UC_SEQ	11	필	255	콘텐츠ID
콘텐츠명	MAIN_TITLE	300	옵	흰여울문화마을	콘텐츠명
구군	GUGUN_NM	100	옵	영도구	구군
위도	LAT	20	옵	35.07885	위도
경도	LNG	20	옵	129.04402	경도
여행지	PLACE	300	옵	흰여울문화마을	여행지



제목	TITLE	300	옵	가파른 절벽 끝에 흰여울문화마을	제목
부제목	SUBTITLE	300	옵	흰여울길에서 만난 느림의 미학	부제목
주소	ADDR1	200	옵	부산광역시 영도구 흰여울길	주소
연락처	CNTCT_TEL	200	옵	051-419-4067	연락처
홈페이지	HOMEPAGE_URL	200	옵	http://huinnyeoul.co.kr	홈페이지
교통정보	TRFC_INFO	500	오급	도시철도 1호선 남포역 6번 출구 → 7, 71, 508 버스 환승 → 영선동 백련사 하차 버스 7,71,508 영선동 백련사 하차 부산시티투어버스 부 산역(점보버스) → 흰여울문화마을 하차 주차 절영해안산책로 입구 공 영 주차장 (유료)	교통정보
운영일	USAGE_DAY	500	옵		운영일
휴무일	HLDY_INFO	500	옵		휴무일
운영 및 시간	USAGE_DAY_WEEK_A ND_TIME	500	옵		운영 및 시간
이용요금	USAGE_AMOUNT	500	옵	무료	이용요금



편의시설	MIDDLE_SIZE_RM1	500	옵	점보버스 휠체어 리프트 장착 차량 운행, 요금 할인 50%, 장애인 주차 장(절영해안산책로 공영주차장, 유 료)	편의시설
이미지URL	MAIN_IMG_NORMAL	500	옵	/uploadImgs/files/cntnts/2019122 2164810529_ttiel	이미지URL
썸네일이미지URL	MAIN_IMG_THUMB	500	옵	/uploadImgs/files/cntnts/2019122 2164810529_thumbL	썸네일이미지URL
상세내용	ITEMCNTNTS	1000	옵		



```
import requests
from bs4 import BeautifulSoup
import pandas as pd
```

```
1 xml_str = response.text
2 soup = BeautifulSoup(xml_str, 'xml')
3 print(soup)
```



```
target_list = soup.find_all('item')
                                                   target list[-1]
 2 len(target_list)
1 title = target_list[0].TITLE.text
2 | addr = target_list[0].ADDR1.text
3 | usage_amount = target_list[0].USAGE_AMOUNT.text
4 cntct_tel = target_list[0].CNTCT_TEL.text
   homepage_url = target_list[0].HOMEPAGE_URL.text
   trfc info = target list[0].TRFC INFO.text
   print('제목 :', title)
   print('주소 :', addr)
10 print('요금 :', usage_amount)
11 print('연락처 :', cntct_tel)
12 print('홈페이지 :', homepage_url)
13 print('교통정보 :', trfc_info)
                                   제목 : 가파른 절벽 끝에 흰여울문화마을
                                   주소 : 부산광역시 영도구 흰여울길
                                   요금 : 무료
                                   연락처: 051-419-4067
                                   홈페이지 : http://www.ydculture.com/huinnyeoulculturetown/
                                   교통정보 : 도시철도 1호선 남포역 6번 출구 → 7, 71, 508 버스 환승 → 영선동 백련사 하차
                                   버스 7,71,508 영선동 백련사 하차
                                   부산시티투어버스 부산역(점보버스) → 흰여울문화마을 하차
                                   주차 점영해안산책로 입구 공영 주차장 (유료)
```



1 df = pd.DataFrame(data_list)

2 df

	제목	주소	요금	연락처	홈페이지	교통정보
0	가파른 절벽 끝에 흰여울 문화마을	부산광역시 영 도구 흰여울길	무료	051-419-4067	http://www.ydculture.com/huinnyeoulculturetown/	도시철도 1호선 남포역 6번 출구 → 7, 71, 508 버 스 환승 → 영선동 백련
1	역사가 살아 숨 쉬는 깡 깡이예술마 을	부산시 영도구 대평북로 36 깡 깡이 안내센터	프로그램별 상이	051-418-3336	http://kangkangee.com	도시철도 1호선 남포역 (6번출구) 도보 18분 \n마을 버스 영도구2 대 평동 하차
2	해양 문화의 꽃 국립해양 박물관	부산시 영도구 해양로301번길 45	관람료 무 료(단, 4D 영상관, 유 료특별전시 제외)	051-309-1900	https://www.mmk.or.kr/	도시철도 1호선 남포역 6번 출구 → 186, 66번 버스 환승 → 국립해양 박물관



국토교통부_아파트매매 실거래 상세 자료 API 신청하기

https://www.data.go.kr/tcs/dss/selectApiDataDetailView.do?publicDataPk=15057511

https://www.data.go.kr/iim/api/selectAPIAcountView.do

'부산광역시 수영구 광안동'의 2022년 5월 아파트 매매 데이타를 API로 신청하고 데이타프레

임으로 출력하여라



	apartment_name	build_year	area	floor	dong	road_name	road_name_bonbun	jibun	deal_amount	deal_day	dealer_lawdnm
0	산호맨션	1986	84.15	5	망미 동	연수로249번길	00020	834-3	28,300	2022.5.10	부산 수영구
1	삼성	1991	84.59	4	망미 동	망미로30번길	00023	777	37,900	2022.5.16	부산 수영구
2	망미한신	1989	84.975	2	망미 동	망미배산로70번길	00045	880-84	30,000	2022.5.27	부산 수영구
3	수영협성르네상스타 운	2004	84.9468	6	수영 동	수영로741번길	00046	432-1	90,000	2022.5.4	부산 수영구