



데이터 처리 프로그래밍

Data Processing Programming

10
데이터 수집

목차

1. 오픈 데이터
2. CSV 파일
3. HTML 파일
4. XML 파일
5. JSON 파일



1. 오픈 데이터터

- 많은 정부 기관이 마이크로소프트 엑셀과 같은 스프레드시트로 정보 관리
- 스프레드시트는 서로 다른 컬럼의 의미를 올바르게 기술하기 위해 사용 가능
- 스프레드시트안에 다루기 어려운 매크로와 공식이 있기 때문에, 사용자가 읽기 쉽도록 해당 스프레드시트와 함께 계산 공식을 제공하는 것이 바람직



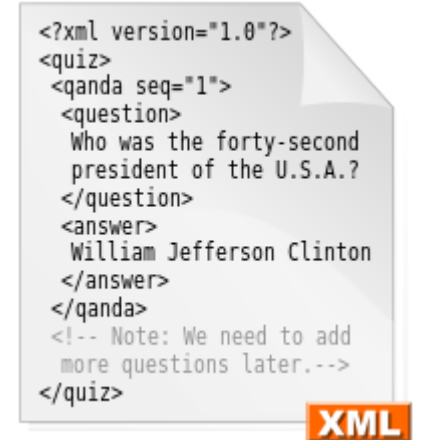
- CSV 파일은 간결하고 동일한 구조로 대용량 데이터 집합을 전송하기 위해 적합하기 때문에 매우 유용한 형식
- CSV로 표현된 데이터는 별도의 문서가 없을 경우 서로 다른 컬럼의 중요성을 추측하기 어려움 (유용성이 떨어질 수 있음)
- 콤마로 구분된 형식에 따라 개별 필드에 대한 정확한 문서화를 만드는 것이 중요
- 파일 구조가 침해받지 않아야 됨
- CSV 파일에서 하나의 필드가 삭제되면 나머지 데이터를 해석하기 위한 방법이 없기 때문에, 파일 수정 없이는 남은 데이터의 해석에 문제가 있음



- 웹 페이지를 위한 지배적인 마크업 언어
- 제목, 단락, 목록 등과 같은 본문을 위한 구조적 의미를 나타내는 것뿐만 아니라 링크, 인용과 그 밖의 항목으로 구조적 문서를 만들 수 있는 방법 제공
- 이미지와 객체를 내장하고 대화형 양식을 생성하는 데 사용
- HTML은 웹 페이지 콘텐츠 안의 <> 괄호에 둘러싸인 "태그"로 되어있는 HTML 요소 형태로 작성
- HTML은 웹 브라우저와 같은 HTML 처리 장치의 행동에 영향을 주는 자바스크립트와 본문과 그 밖의 항목의 외관과 배치를 정의하는 CSS 같은 스크립트를 포함하거나 불러올 수 있음



- W3C에서 개발
- 다른 특수한 목적을 갖는 마크업 언어를 만드는데 사용하도록 권장하는 다목적 마크업 언어
- XML은 SGML의 단순화된 부분집합
- 다른 많은 종류의 데이터를 기술하는 데 사용할 수 있음
- XML은 주로 다른 종류의 시스템, 특히 인터넷에 연결된 시스템끼리 데이터를 쉽게 주고 받을 수 있게 하여 HTML의 한계를 극복할 목적으로 만들어짐
- 데이터 상호교환을 위해 폭넓게 사용되는 포맷
- 데이터 구조를 유지하는데 유리하며 개발자는 파일을 읽는 간섭없이 데이터에 문서의 일부 작성 가능



- 속성-값 쌍(attribute-value pairs and array data types (or any other serializable value)) 또는 "키-값 쌍"으로 이루어진 데이터 오브젝트를 전달하기 위해 인간이 읽을 수 있는 텍스트를 사용하는 개방형 표준 포맷
- 비동기 브라우저/서버 통신 (AJAX)을 위해, 넓게는 XML(AJAX가 사용)을 대체하는 주요 데이터 포맷
- 특히, 인터넷에서 자료를 주고 받을 때 그 자료를 표현하는 방법으로 알려져 있음
- 자료의 종류에 큰 제한은 없으며, 특히 컴퓨터 프로그램의 변수값을 표현하는 데 적합
- 어떤 프로그래밍 언어로도 읽을 수 있는 단순한 파일 포맷
- XML 같은 포맷에 비해 컴퓨터가 처리하기 쉬움



- W3C가 권고하는 포맷
- 다수의 자료원으로부터 데이터를 쉽게 조합할 수 있는 형식으로 데이터 표현 가능
- RDF 데이터는 여러 가지 직렬화 방법이 있으며, XML과 JSON으로 저장 가능
- RDF는 웹에 있는 {term:열린 데이터} 이니셔티브를 서로 연결하는데 편리한 방법을 제공 (식별자로 URL 사용을 장려)
- RDF는 폭넓게 확산되어 있지 않지만, 영국과 스페인 정부의 Linked Data 프로젝트를 포함한 열린 정부 이니셔티브 사이에서 트렌드
- 웹의 창시자인 팀 버너스리는 최근 5개 별 스키마를 제안하였는데, 이는 오픈데이터 이니셔티브들이 추구하는 목표로서 연결된 RDF 데이터를 포함



- Word, ODF, OOXML, PDF와 같은 포맷의 표준적인 문서로 일정한 종류의 데이터를 보여주는데 활용
- 이 포맷은 일관적으로 구조를 유지하지 않으며, 자동적으로 데이터를 입력하기 어려움
- 데이터를 재사용할 수 있도록 문서 기준의 템플릿을 사용한다면, 최소한 정보를 문서 밖으로 추출 가능
- 가능한 많은 활자 마크업을 이용하여 기계가 콘텐츠로부터 표제를 구분하기 쉽게 되어있는 것은 데이터의 추가 사용을 이끌어낼 수 있음
- 일반적으로 다른 포맷으로 데이터가 있다면, 워드 프레세싱 포맷에 표현하는 것을 추천하지 않음



- 일반 텍스트 문서 (.txt) 는 컴퓨터가 읽기 매우 쉬움
- 문서 내에 구조화된 메타데이터를 갖지 않음
- 개발자는 개별 문서를 처리할 수 있는 파서를 개발할 필요가 있음
- 운영체제 사이에서 일반 텍스트 문서를 전환할 때 문제가 생길 수 있음
(MS 윈도우, Mac OSX, Unix 시스템은 라인 끝을 처리하는 서로 다른 방식을 갖고 있음)



- 아마도 대부분의 데이터에서 가장 적합한 형식은 문서의 전체 텍스트와 문서의 이미지를 함께 표시하는 것
- TIFF와 JPEG-2000 포맷은 전자적으로 만들어지지 않은 데이터를 이미지로 표현하는데 적합
- 오래된 교회 기록과 기록 사진은 자료가 없는 것보다 나음



- 전용 시스템은 데이터를 저장하고 추출할 수 있는 독립적인 데이터 포맷을 갖고 있음
- 이와 같은 포맷으로 데이터를 노출시키는 경우가 종종 있음
- 특히 데이터를 유사한 시스템에서 사용할 것으로 기대된다면, 전용 포맷에서만 찾을 수 있는 정보는 항상 표시해야 함
- 예를 들어 공급자의 웹사이트에 링크를 제공하는 것 같이 전용 포맷이 아닌 방법으로 표시하는 것을 추천





2. CSV 파일

표준데이터

문화관광

·데이터셋 유형

공통

·키워드

여행, 관광단지, 관광명소

·분류

문화체육관광 > 관광

·보유근거

·수집방법

전국관광지정보표준데이터

ENGLISH

전국 관광지정보(편의시설, 관광지소개 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)는 [공공데이터 개방 표준]고시를 참고하시기 바랍니다.(게시판>자료실>법령&지침) 전국표준데이터는 각 기관에서 등록한 표준데이터를 취합하여 제공합니다. 표준데이터의 갱신주기는 개별 파일마다 다릅니다.(기관에서 등록한 데이터를 취합한 것으로 개별 파일별 갱신시점이 다름) 전국표준데이터의 취합시점은 월말 일괄 취합됩니다.

XLS

전국관광지정보표준데이터

다운로드

오류신고

총 111 건

+ 더보기

경상남도_김해시_관광지정보_20180218

바로가기

강원도_양구군_관광지정보_20180228

바로가기

전라남도_해남군_관광지정보_20180219

바로가기

그리드

OPEN API

지도

차트

XLS

XML

JSON

CSV

RDF

※ XLS 이외의 파일은 다운로드시 다소 시간이 걸릴 수 있습니다.

※ 지도는 그리드에 조회된 데이터 중 최대 5,000건까지 표시됩니다.

전국관광지정보표준데이터 보기(CSV)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	관광지명	관광지구번호	소재지도	소재지지번	위도	경도	면적	공공편익시설	숙박시설	운동및오락	휴양및문화	접객시설	지원시설	지정일자	수용인원수	주차가능수
2	오수의견	관광지	전라북도	전라북도	35.54533	127.333	0	관리사무소+화장실						#####	0	0
3	사선대관	관광지	전라북도	전라북도	35.67236	127.2745	0	관리사무소+화장실등						#####	0	0
4	최참판택	관광지	경상남도	경상남도	35.15596	127.6884	9529	화장실(4동)+주차장						#####	5000	230
5	화개장터	관광지	경상남도	경상남도	35.188	127.6244	8226	화장실(2동)+주차장						#####	3000	300
6	삼성궁	관광지	경상남도	경상남도	35.23887	127.7054	43967	화장실(2동)+주차장						#####	3000	100
7	이병주문	관광지	경상남도	경상남도	35.09736	127.8942	13535	화장실(2동)+주차장						#####	500	40
8	지리산역	관광지	경상남도	경상남도	35.28994	127.6465	4311	화장실(1동)+주차장						#####	100	15
9	소요산관	관광지	경기도	경기도	37.94649	127.0693	2472000	관리사무소+주차장+화장실			자유수호	평야외음악당		#####	100000	500
10	경천대관	관광지	경상북도	경상북도	36.45713	128.2472	209000	관리사무소	경천대펜션	카누장+눈	수영장+오	관광식당		#####	3000	213
11	회상나루	관광지	경상북도	경상북도	36.44204	128.2656	165000	관리사무소	객주촌(펜	학(철새)	전문학관	주막촌(식당)		#####	1500	159
12	김해가야	관광지	경상남도	경상남도	35.25107	128.893	165078	일반음식점	카라반 캠프	가야무사	철광산공연장			#####	5000	281
13	김해낙동	관광지	경상남도	경상남도	35.37206	128.8209	313.87	전망대, 열	없음	낙동강레	일와인동굴, 열차카페			#####	1000	210
14	가야랜드	관광지	경상남도	경상남도	35.2612	128.9015	1352.91	음식점, 화	카라반, 글	12종의 유	기기구			#####	1000	398
15	김해천문	관광지	경상남도	경상남도	35.25307	128.8874	6600	사무실, 관	없음		전시동, 관	측동		#####	500	13
16	롯데위터	관광지	경상남도	경상남도	35.1799	128.8281	122397.3	음식점, 화	없음	25종의 유	기기구			#####	500	732
17	클레이아	관광지	경상남도	경상남도	35.25126	128.7448	41952	전시관, 체	없음		돔하우스, 큐	빅하우스, 체험관		#####	500	203
18	김해한옥	관광지	경상남도	경상남도	35.2344	128.8771	4132	음식점, 화	13개 객실					#####	30	10
19	국립김해	관광지	경상남도	경상남도	35.24311	128.8735	16783.83	전시관, 화	없음					#####	1000	194
20	노무현대	관광지	경상남도	경상남도	35.31473	128.7712	305	화장실, 주	없음					#####	5000	240
21	수로왕릉	관광지	경상남도	경상남도	35.23553	128.8784	49488	화장실, 주	없음					#####	5000	38
22	부경동물	관광지	경상남도	경상남도	35.22059	128.8198	250	음식점, 호	없음					#####	500	30
23	도림사	관광지	전라남도	전라남도	35.26706	127.2576	242.55	공중화장	실도림사	오토캠핑장	도림사	오토캠핑장		#####	1000	100

- csv 모듈 사용
- csv.reader()를 통해 csv 파일 읽기
- 각 line 마다 출력

```
1 import csv
2
3 with open('data/전국관광지정보표준데이터.csv', 'r') as f:
4     r = csv.reader(f)
5     for row in r:
6         print(row)
```

D:\Python\venv\Scripts\python.exe D:/Python/csv_file.py

['관광지명', '관광지구분', '소재지도로명주소', '소재지지번주소', '위도', '경도', '면적', '공공편의시설정보', '숙박시설정보', '운동및오락시
['오수익견관광지', '관광지', '전라북도 임실군 오수면 충효로 2096-16', '전라북도 임실군 오수면 오수리 33 임실군문화체육센터', '35.545326',
['사선대관광지', '관광지', '전라북도 임실군 관촌면 사선2길 68-7', '전라북도 임실군 관촌면 관촌리 222 사선대 관리 사무소', '35.672357', '1
['최참판댁', '관광지', '경상남도 하동군 악양면 평사리길 66-7', '경상남도 하동군 악양면 평사리 483', '35.155961', '127.688408', '9529',
['화개장터', '관광지', '경상남도 하동군 화개면 쌍계로 15', '경상남도 하동군 화개면 탑리 726-46', '35.187995', '127.624366', '8226', '호
['삼성궁', '관광지', '경상남도 하동군 청암면 삼성궁길 2', '경상남도 하동군 청암면 묵계리 1738-3', '35.238866', '127.705394', '43967', '
['이병주문학관', '관광지', '경상남도 하동군 북천면 이명골길 14-28', '경상남도 하동군 북천면 직전리 231', '35.097358', '127.894176', '13
['지리산역사관', '관광지', '경상남도 하동군 화개면 화개로 1438', '경상남도 하동군 화개면 대성리 1389', '35.289936', '127.646527', '4311
['소요산관광지', '관광지', '경기도 동두천시 평화로2910번길 145', '경기도 동두천시 상봉암동 산 1-1', '37.9464858', '127.069317', '247206
['경천대관광지', '관광지', '경상북도 상주시 사벌면 경천로 652', '경상북도 상주시 사벌면 삼덕리 1-13', '36.457131', '128.247205', '20900
['회상나루관광지', '관광지', '경상북도 상주시 중동면 갯다불길 97', '경상북도 상주시 중동면 회상리 764', '36.442044', '128.265614', '1650

```
1 import csv
2
3 with open('data/전국관광지정보표준데이터.csv', 'r') as f:
4     r = csv.reader(f,
5                     delimiter=',',
6                     quotechar='"',
7                     quoting=csv.QUOTE_ALL)
8     for row in r:
9         print(row)
```

- delimiter: 데이터 구분자
- quotechar: 데이터를 묶는 문자
- quoting: 데이터를 묶는 기준
 - QUOTE_ALL: 모든 데이터를 자료형에 상관없이 묶음 (문자열 처리)
 - QUOTE_MINIMAL: 최소 데이터만 묶음
 - QUOTE_NONNUMERIC: 숫자 데이터가 아닌 경우만 묶음
 - QUOTE_NONE: 데이터를 묶지 않음

■ 관광지명 컬럼 데이터만 출력

관광지명
오수의견관광지
사선대관광지
최참판댁
화개장터
삼성궁
이병주문학관
지리산역사관
소요산관광지
경천대관광지
회상나루관광지
김해가야테마파크
김해낙동강레일파크
가야랜드
김해천문대
롯데워터파크
...

```
1 import csv
2
3 row_index = 0
4
5 with open('data/전국관광지정보표준데이터.csv', 'r') as f:
6     r = csv.reader(f,
7                     delimiter=',',
8                     quotechar='"',
9                     quoting=csv.QUOTE_ALL)
10
11     for row in r:
12         print(row[0])
```

- 첫번째 행은 csv의 헤더 정보

```
1 import csv
2
3 row_index = 0
4 header = []
5
6 with open('data/전국관광지정보표준데이터.csv', 'r') as f:
7     r = csv.reader(f,
8                     delimiter=',',
9                     quotechar='"',
10                    quoting=csv.QUOTE_ALL)
11
12     for row in r:
13         if row_index == 0:
14             header = row
15
16         row_index += 1
17
18 print(header)
```

['관광지명', '관광지구분', '소재지도로명주소', '소재지지번주소', '위도', '경도', '면적', '공공편의시설정보', '숙박시설정보', '운동및오락시설정보', '휴양및문화시설정보', '접객시설정보', '지원시설정보', '지정일자', '수용인원수', '주차가능수', '관광지소개', '관리기관전화번호', '관리기관명', '데이터기준일자', '제공기관코드', '제공기관명', '']

- csv 모듈 사용
- csv.writer()를 통해 csv 파일 쓰기
- writerow()를 이용해 한 행씩 쓰기

```
1 import csv
2
3 with open('data/주소록.csv', 'w') as f:
4     w = csv.writer(f)
5     w.writerow([1, "이수안", "010-1234-1234"])
6     w.writerow([2, "홍길동", "010-4321-4321"])
```

Lab. 다른 CSV 파일 가져오기

파일데이터 26,136건을 찾았습니다.

기관별 검색

파일데이터 [26,136건]	정확도 ▾	날 짜	제 목	조회수	다운로드
-----------------	-------	-----	-----	-----	------

강원도 양양군_소규모공공시설위험지정정보 조회수 : 8 다운로드수 : 19

수정일 : 2018.10.05 기관 : 강원도 양양군 서비스유형 : 다운로드

강원도 양양군_소규모공공시설위험지정정보표준데이터

CSV

충청북도 보은군_지진해일대피소 조회수 : 6 다운로드수 : 31

수정일 : 2017.12.01 기관 : 충청북도 보은군 서비스유형 : 다운로드

충청북도 보은군_지진해일대피소표준데이터

CSV

인천광역시 동구_육교정보 조회수 : 4 다운로드수 : 22

수정일 : 2018.08.14 기관 : 인천광역시 동구 서비스유형 : 다운로드

인천광역시 동구_육교정보표준데이터

CSV

인천광역시 동구_음식물쓰레기납부필증가격정보 조회수 : 5 다운로드수 : 25

수정일 : 2018.09.28 기관 : 인천광역시 동구 서비스유형 : 다운로드

인천광역시 동구_음식물쓰레기납부필증가격정보표준데이터

CSV

경상북도 성주군_민방위대피시설 조회수 : 5 다운로드수 : 18

수정일 : 2018.10.05 기관 : 경상북도 성주군 서비스유형 : 다운로드

경상북도 성주군_민방위대피시설표준데이터

CSV

표준데이터 108건을 찾았습니다.

표준데이터 [108건]	정확도	날 짜	제 목	조회수 ▾
--------------	-----	-----	-----	-------

전국도시공원표준데이터 조회수 1,942

수정일 : 2015.12.19 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 도시공원정보(공원유형, 보유시설 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)는 [공공...

CSV JSON RDF XLS XML

전국평생학습강좌표준데이터 조회수 1,272

수정일 : 2015.12.30 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 평생학습강좌 정보(강사, 교육일정 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)는 [공공...

CSV JSON RDF XLS XML

전국CCTV표준데이터 조회수 1,258

수정일 : 2015.12.30 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 CCTV 정보(관리기관명, 소재지도로명주소 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)...

CSV JSON RDF XLS XML

전국금연구역표준데이터 조회수 1,172

수정일 : 2016.01.21 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 금연구역 정보(금연구역 지역, 면적, 위반과태료 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)...

CSV JSON RDF XLS XML ZIP

전국무더위쉼터표준데이터 조회수 980

수정일 : 2016.01.21 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 무더위쉼터 정보(쉼터위치, 냉방시설 개수 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)...

CSV JSON RDF XLS XML



3. HTML 파일

HTML 구조

- `<html> ~ </html>`: HTML 문서의 시작과 끝
- `<head> ~ </head>`: HTML 문서 메타데이터 정의
- `<title> ~ </title>`: HTML 문서 제목
- `<body> ~ </body>`: HTML 문서 내용
- `<h1> ~ </h1>`: 제목^{heading}
- `<p> ~ </p>`: 단락^{paragraph}
- `<a>`: 문서 링크

```
<!DOCTYPE html>
<html>
<head>
  <title>Title</title>
</head>
<body>
  <h1>Heading 1</h1>
  <h2>Heading 2</h2>
  <p>HTML Document</p>
</body>
</html>
```

BeautifulSoup 모듈

- 파싱 기능이 있는 다른 라이브러리를 쉽게 사용할 수 있도록 만든 래퍼^{wrapper}
- 전통적인 파이썬 XML 파서(XML parser)에는 lxml과 html5lib 등이 존재
- BeautifulSoup 모듈은 이를 래핑하여 데이터를 쉽고 빠르게 처리
- 설치 명령어
 - `pip install beautifulsoup4`
 - `pip install lxml`



```
1 from bs4 import BeautifulSoup
2
3 bs = BeautifulSoup(open("../data/basic.html"), "html.parser")
4
5 print(bs.get_text)
6 print(bs.find_all("h1"))
```

- BeautifulSoup 모듈을 이용하여 basic.html 파일을 파싱

```
<bound method Tag.get_text of <html>
<head>
<title>Title</title>
</head>
<body>
<h1>Heading 1</h1>
<h2>Heading 2</h2>
<p>HTML Document</p>
</body>
</html>>
[<h1>Heading 1</h1>]
```

- URL을 이용해 웹에 접근하여 HTML 파일 파싱
- <a> 태그에 href 속성 값만 출력

```
1  from urllib.request import urlopen
2  from bs4 import BeautifulSoup
3
4  with urlopen("http://suanlab.com") as html:
5      bs = BeautifulSoup(html)
6
7  for link in bs.find_all("a"):
8      if link.has_attr("href"):
9          print(link["href"])
```



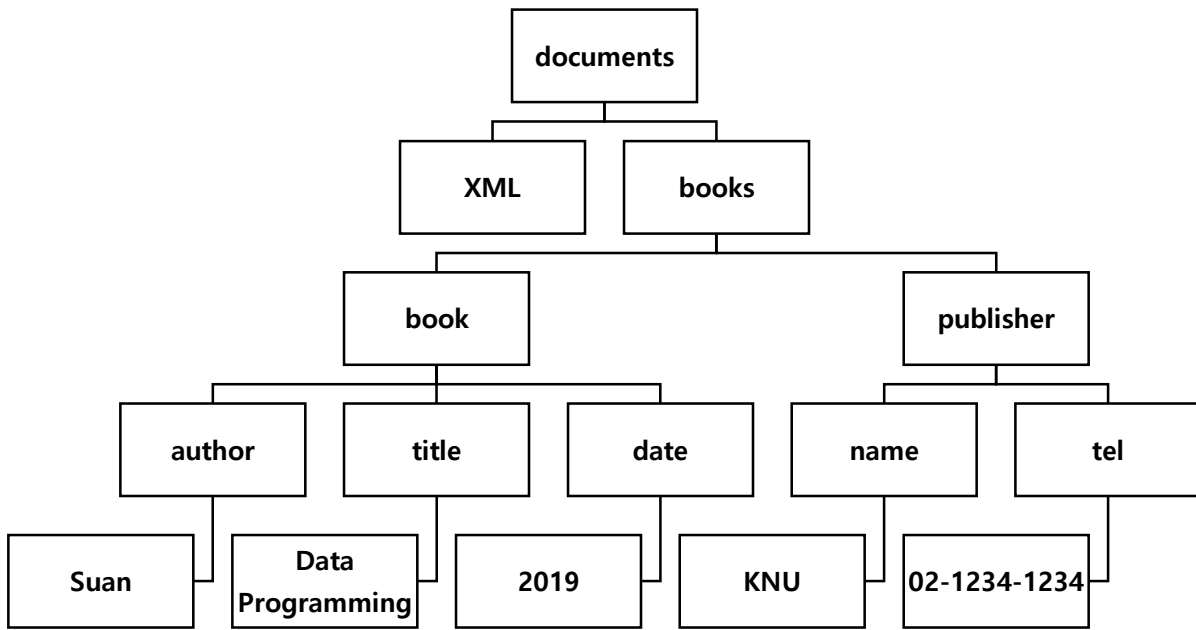
4. XML 파일

XML 구조

- 데이터의 구조와 의미를 설명하는 태그
- 어떤 데이터의 속성과 값을 표현

```
<?xml version="1.0"?>  
  <학생>  
    <학번>201821512</학번>  
    <학과>데이터사이언스학과</학과>  
    <이름>이수안</이름>  
    <성별>남자</성별>  
  </학생>
```

XML 문서



```
<?xml version="1.0"?>
<books>
    <book>
        <author>Suan</author>
        <title>Data Programming</title>
        <date>2019</date>
    </book>
    <publisher>
        <name>KNU</name>
        <tel>02-1234-1234</tel>
    </publisher>
</books>
```

- BeautifulSoup 모듈을 이용하여 books.xml 파일을 파싱

```
1 from bs4 import BeautifulSoup
2
3 with open("../data/books.xml", "r") as books_file:
4     books_xml = books_file.read()
5
6 bs = BeautifulSoup(books_xml, "lxml")
7
8 for book in bs.find_all("books"):
9     print(book)
10
```


Lab. 다른 XML 파일 가져오기

파일데이터 26,136건을 찾았습니다.

기관별 검색

파일데이터 [26,136건]	정확도 ▾	날 짜	제 목	조회수	다운로드
-----------------	-------	-----	-----	-----	------

강원도 양양군_소규모공공시설위험지정정보 조회수 : 8 다운로드수 : 19

수정일 : 2018.10.05 기관 : 강원도 양양군 서비스유형 : 다운로드

강원도 양양군_소규모공공시설위험지정정보표준데이터

CSV

충청북도 보은군_지진해일대피소 조회수 : 6 다운로드수 : 31

수정일 : 2017.12.01 기관 : 충청북도 보은군 서비스유형 : 다운로드

충청북도 보은군_지진해일대피소표준데이터

CSV

인천광역시 동구_육교정보 조회수 : 4 다운로드수 : 22

수정일 : 2018.08.14 기관 : 인천광역시 동구 서비스유형 : 다운로드

인천광역시 동구_육교정보표준데이터

CSV

인천광역시 동구_음식물쓰레기납부필증가격정보 조회수 : 5 다운로드수 : 25

수정일 : 2018.09.28 기관 : 인천광역시 동구 서비스유형 : 다운로드

인천광역시 동구_음식물쓰레기납부필증가격정보표준데이터

CSV

경상북도 성주군_민방위대피시설 조회수 : 5 다운로드수 : 18

수정일 : 2018.10.05 기관 : 경상북도 성주군 서비스유형 : 다운로드

경상북도 성주군_민방위대피시설표준데이터

CSV

표준데이터 108건을 찾았습니다.

표준데이터 [108건]	정확도	날 짜	제 목	조회수 ▾
--------------	-----	-----	-----	-------

전국도시공원표준데이터 조회수 1,942

수정일 : 2015.12.19 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 도시공원정보(공원유형, 보유시설 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)는 [공공...

CSV JSON RDF XLS XML

전국평생학습강좌표준데이터 조회수 1,272

수정일 : 2015.12.30 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 평생학습강좌 정보(강사, 교육일정 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)는 [공공...

CSV JSON RDF XLS XML

전국CCTV표준데이터 조회수 1,258

수정일 : 2015.12.30 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 CCTV 정보(관리기관명, 소재지도로명주소 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)...

CSV JSON RDF XLS XML

전국금연구역표준데이터 조회수 1,172

수정일 : 2016.01.21 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 금연구역 정보(금연구역 지역, 면적, 위반과태료 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)...

CSV JSON RDF XLS XML ZIP

전국무더위쉼터표준데이터 조회수 980

수정일 : 2016.01.21 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 무더위쉼터 정보(쉼터위치, 냉방시설 개수 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)...

CSV JSON RDF XLS XML



5. JSON 파일

JSON 구조

- JSON은 XML보다 데이터 용량이 적고 코드로의 전환이 쉽다는 측면에서 XML의 대체재로 가장 많이 활용
- JSON은 파이썬의 딕셔너리형과 매우 비슷하여, 키-값의 쌍으로 구성
- XML과 비교할 때 JSON의 장점은 일단 코드가 간결하고, 코드의 전환이 쉬우며, 코드의 간결함 때문에 용량 절약

```
{  
    "students": [  
        {  
            "id": "20181512",  
            "department": "data science",  
            "name": "Suan",  
            "sex": "Male"  
        }  
    ]  
}
```

JSON 모듈

- 파이썬에서 JSON을 사용하기 위해서는 json 모듈 이용
- JSON 데이터 포맷은 데이터 저장 및 읽기가 딕셔너리형과 완벽히 상호 호환
- 딕셔너리형에 익숙한 사용자가 매우 쉽게 사용할 수 있다는 장점
- 설치 명령어
 - `pip install json`

- json 모듈을 이용하여
students.json 파일 읽기

```
1 import json
2
3 with open("students.json", "r") as json_file:
4     json_doc = json_file.read()
5     json_data = json.loads(json_doc)
6     print(json_data["students"])
```

- json 모듈을 이용하여 students.json 파일 쓰기

```
1 import json
2
3 student_data = {"id": "20191234", "department": "data science", "name": "Susan", "sex": "Female"}
4
5 with open("students.json", "w") as json_file:
6     json.dump(student_data, json_file)
7
```

Lab. 다른 JSON 파일 가져오기

파일데이터 26,136건을 찾았습니다.

기관별 검색

파일데이터 [26,136건]	정확도 ▾	날 짜	제 목	조회수	다운로드
-----------------	-------	-----	-----	-----	------

강원도 양양군_소규모공공시설위험지정정보 조회수 : 8 다운로드수 : 19

수정일 : 2018.10.05 기관 : 강원도 양양군 서비스유형 : 다운로드

강원도 양양군_소규모공공시설위험지정정보표준데이터

CSV

충청북도 보은군_지진해일대피소 조회수 : 6 다운로드수 : 31

수정일 : 2017.12.01 기관 : 충청북도 보은군 서비스유형 : 다운로드

충청북도 보은군_지진해일대피소표준데이터

CSV

인천광역시 동구_육교정보 조회수 : 4 다운로드수 : 22

수정일 : 2018.08.14 기관 : 인천광역시 동구 서비스유형 : 다운로드

인천광역시 동구_육교정보표준데이터

CSV

인천광역시 동구_음식물쓰레기납부필증가격정보 조회수 : 5 다운로드수 : 25

수정일 : 2018.09.28 기관 : 인천광역시 동구 서비스유형 : 다운로드

인천광역시 동구_음식물쓰레기납부필증가격정보표준데이터

CSV

경상북도 성주군_민방위대피시설 조회수 : 5 다운로드수 : 18

수정일 : 2018.10.05 기관 : 경상북도 성주군 서비스유형 : 다운로드

경상북도 성주군_민방위대피시설표준데이터

CSV

표준데이터 108건을 찾았습니다.

표준데이터 [108건]	정확도	날 짜	제 목	조회수 ▾
--------------	-----	-----	-----	-------

전국도시공원표준데이터 조회수 1,942

수정일 : 2015.12.19 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 도시공원정보(공원유형, 보유시설 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)는 [공공...

CSV JSON RDF XLS XML

전국평생학습강좌표준데이터 조회수 1,272

수정일 : 2015.12.30 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 평생학습강좌 정보(강사, 교육일정 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)는 [공공...

CSV JSON RDF XLS XML

전국CCTV표준데이터 조회수 1,258

수정일 : 2015.12.30 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 CCTV 정보(관리기관명, 소재지도로명주소 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)...

CSV JSON RDF XLS XML

전국금연구역표준데이터 조회수 1,172

수정일 : 2016.01.21 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 금연구역 정보(금연구역 지역, 면적, 위반과태료 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)...

CSV JSON RDF XLS XML ZIP

전국무더위쉼터표준데이터 조회수 980

수정일 : 2016.01.21 기관 : 공공데이터활용지원센터 서비스유형 : 그리드 차트 지도 다운로드

전국 무더위쉼터 정보(쉼터위치, 냉방시설 개수 등)를 제공합니다. 공공데이터 개방 표준데이터 속성정보(표현형식/단위 등)...

CSV JSON RDF XLS XML

Q & A