

파이썬 활용 – REST API 요청 및 데이터 처리

# Session 4

NEXT X LIKELION 김지욱

# || 목차

1. 파이썬 활용
2. 크롤링
  - 2-1. 크롤링 관련 개념
    - 2-1. Data
    - 2-1. REST API
  - 2-2. 공동실습
3. 미니실습

# < 파이썬 활용 />

\* 활용 : 잘 이용하는 것

# | 파이썬 활용



# | 파이썬 활용



# | 파이썬 활용

상하수도

벽지

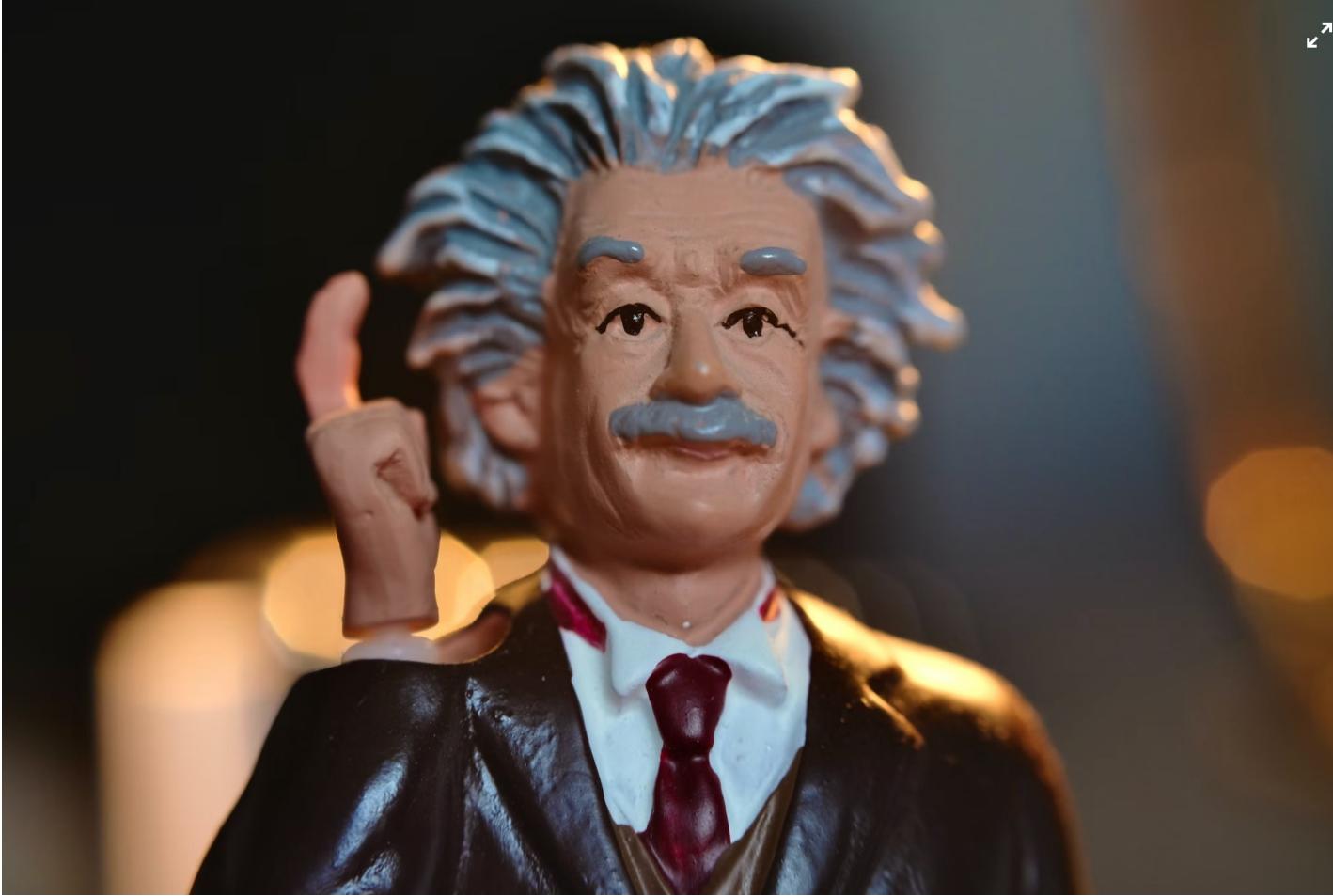
바닥재

현관문

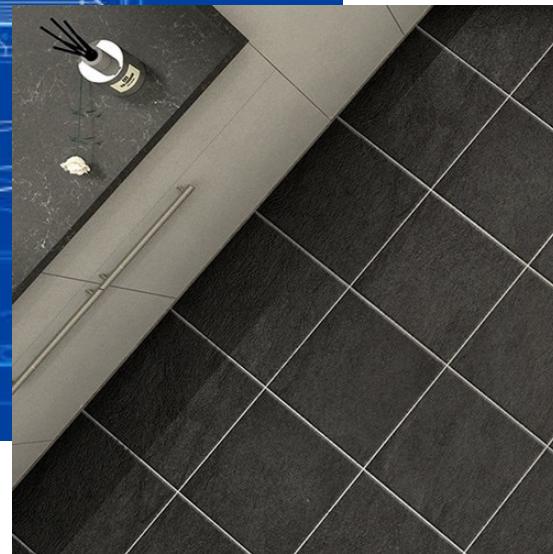
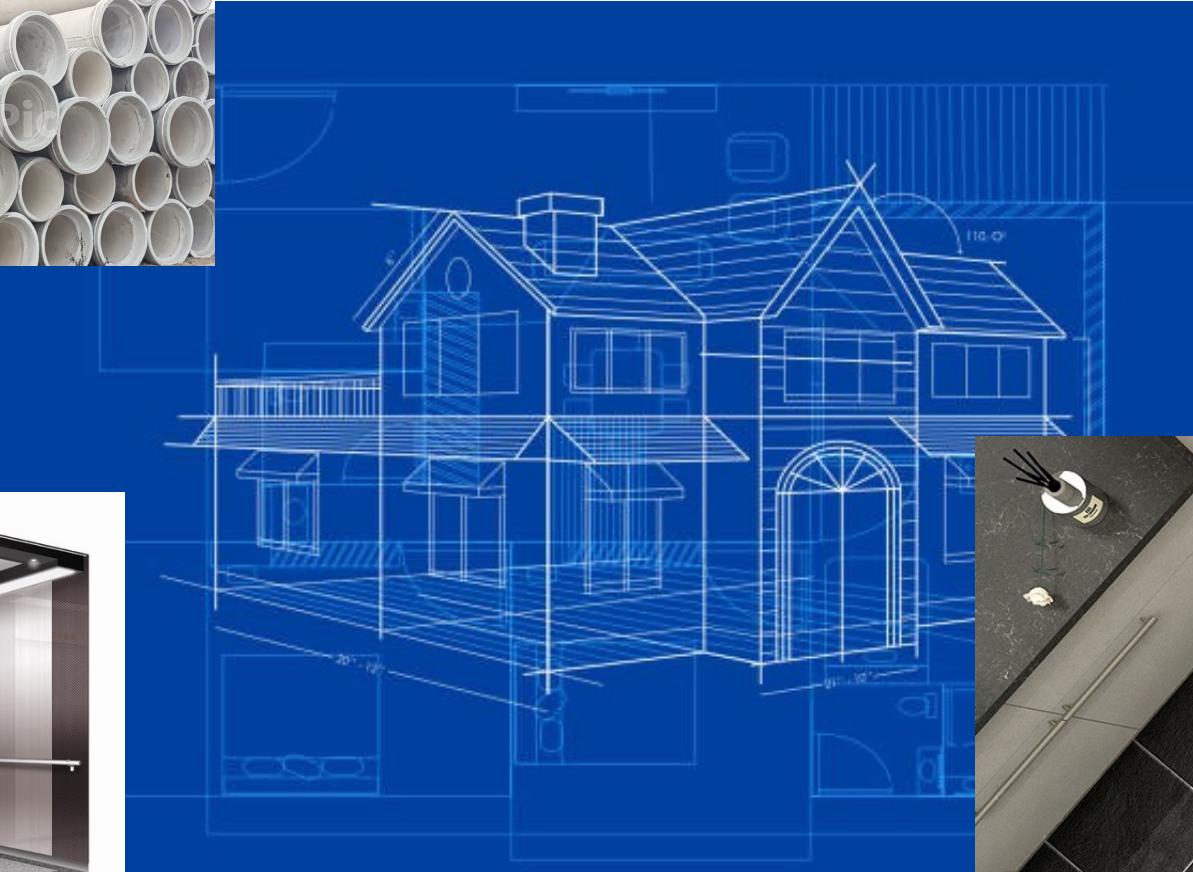
엘리베이터



# | 파이썬 활용



# | 파이썬 활용



# | 파이썬 활용



# | 파이썬 활용

어떤 버전으로 만든 것이지?

누가 만든 것이지?

언제 만든 것이지?

어떤 맥락에서 만든 것이지?



어떻게 사용하는 것이지?

# | 파이썬 활용

module 

# | 파이썬 활용

pip

Python Install Package

# | 파이썬 활용



# | 파이썬 활용



**Requests**  
*http for humans*

**Beautiful**Soup

 **pandas**

# < Crawling />

\* A powerful technique to collect data from the web

# Before Crawling

# | Crawling – why

Why now?

## 목적

\* 목적 : 이루고자 하는 것

# | Crawling - why

1. Data

2. REST API

# | Data process

## Data

- \* a collection of discrete values



# Data process

만족도가 높은! 강의

2019년 여름학기 **교양** 강의로 이동 >  
회계학원리(영강)  
이상호 교수님  
강의 만족도 **4.8**/5.0

2019년 2학기 **전공** 강의로 이동 >  
러시아작가론II  
석영중 교수님  
강의 만족도 **4.3**/5.0

2019년 2학기 **전공** 강의로 이동 >  
모바일포렌식  
한재혁 교수님  
강의 만족도 **4.5**/5.0

학점평가가 좋은! 강의

2019년 1학기 **전공** 강의로 이동 >  
라틴아메리카단편소설강독  
배지완 교수님  
학점 평가 **4.2**/5.0

2020년 1학기 **전공** 강의로 이동 >  
영어교과논리및논술(영강)  
Zoe Ksan Rubadeau 교수님  
학점 평가 **4.1**/5.0

2020년 1학기 **교양** 강의로 이동 >  
한국의신분과사회상(담당교수 변경/2....  
이송순 교수님  
학점 평가 **4.4**/5.0

# Data process



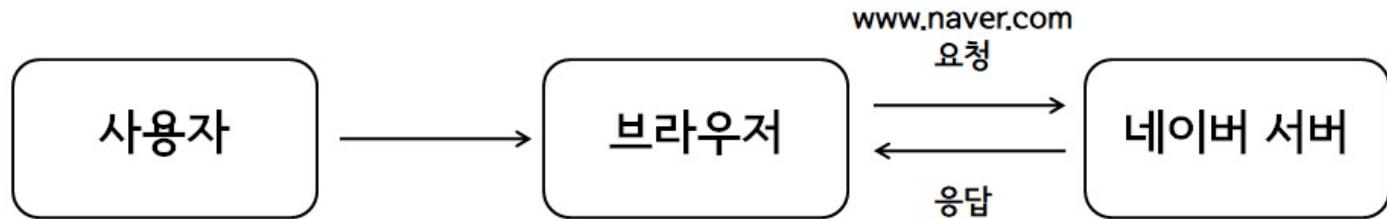
	userId	id	title	body
0	1	1	sunt aut facere repellat provident occaecati e...	quia et suscipit\nsuscipit recusandae consequu...
1	1	2		qui est esse est rerum tempore vitae\nsequi sint nihil repr...
2	1	3	ea molestias quasi exercitationem repellat qui...	et iusto sed quo iure\nvoluptatem occaecati om...
3	1	4		eum et est occaecati ullam et saepe reiciendis voluptatem adipisci...
4	1	5		nesciunt quas odio repudiandae veniam quaerat sunt sed\nalias aut...
...	...	...		...
95	10	96	quaerat velit veniam amet cupiditate aut numqu...	in non odio excepturi sint eum\nlabore volupta...
96	10	97	quas fugiat ut perspiciatis vero provident	eum non blanditiis soluta porro quibusdam volu...
97	10	98		laboriosam dolor voluptates doloremque ex facilis sit sint culpa\nsoluta a...
98	10	99	temporibus sit alias delectus eligendi possimu...	quo deleniti praesentium dicta non quod\naut e...
99	10	100	at nam consequatur ea labore ea harum	cupiditate quo est a modi nesciunt soluta\nips...

# | Data process

## REST API

\* 자원을 정의하고 자원에 대한 주소를 지정하는 방법

# REST API



1. HTML
2. CSS
3. JS

# REST API



1. 인스타 앱을 실행하기.
2. 게시물을 업로드하기.
3. 오타가 있어서 수정하기.
4. 게시물 삭제하기.



# REST API

1. 인스타 앱을 실행하기.

사진을 서버자원에서 불러와.

GET method 200



Instagram

# REST API



2. 게시물을 업로드하기.

사진을 서버자원에서 등록해.

POST method 201



Instagram

# REST API



3. 오타가 있어서 수정하기.

사진을 서버자원에서 업데이트해.

PATCH method 200



Instagram

# REST API



4. 게시물 삭제하기.

사진을 서버자원에서 삭제해.

DELETE method 200



Instagram

# | REST API

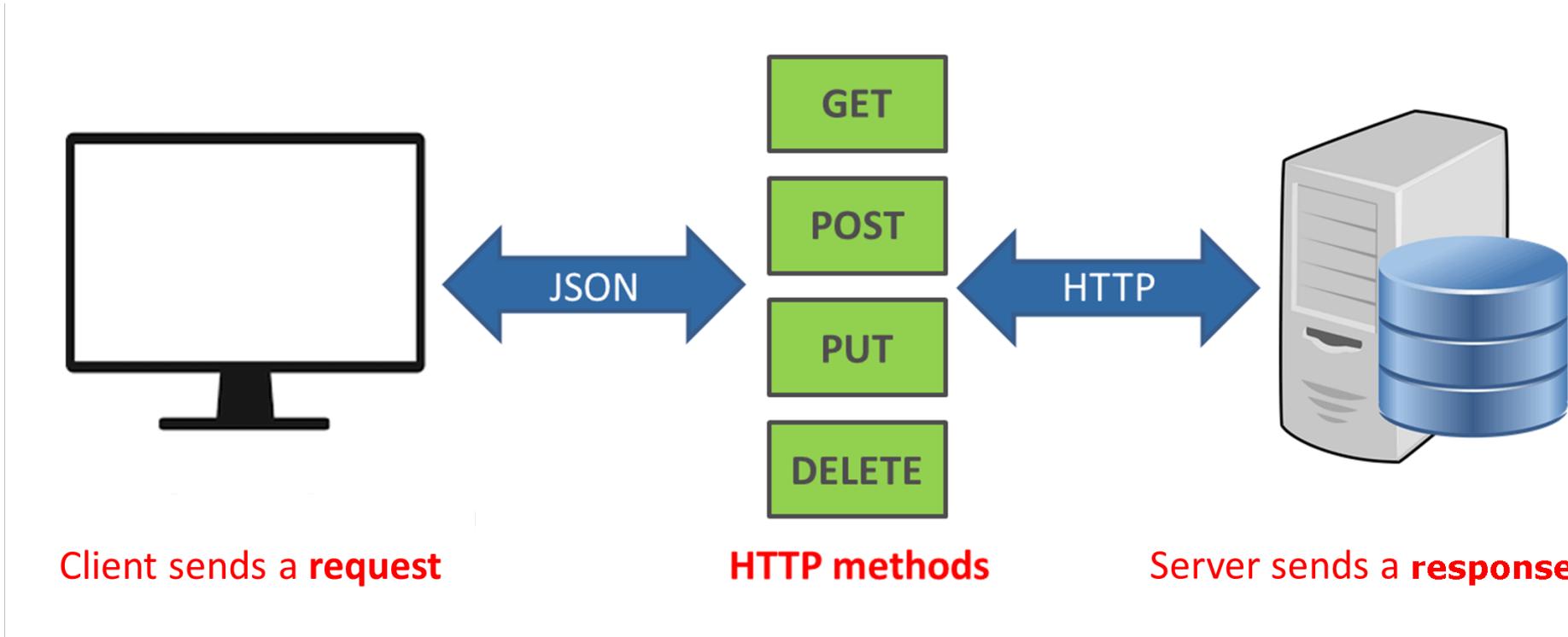
데이터를 서버자원에서 어떻게 해!

Representation

Resource

verb

# REST API



# REST API

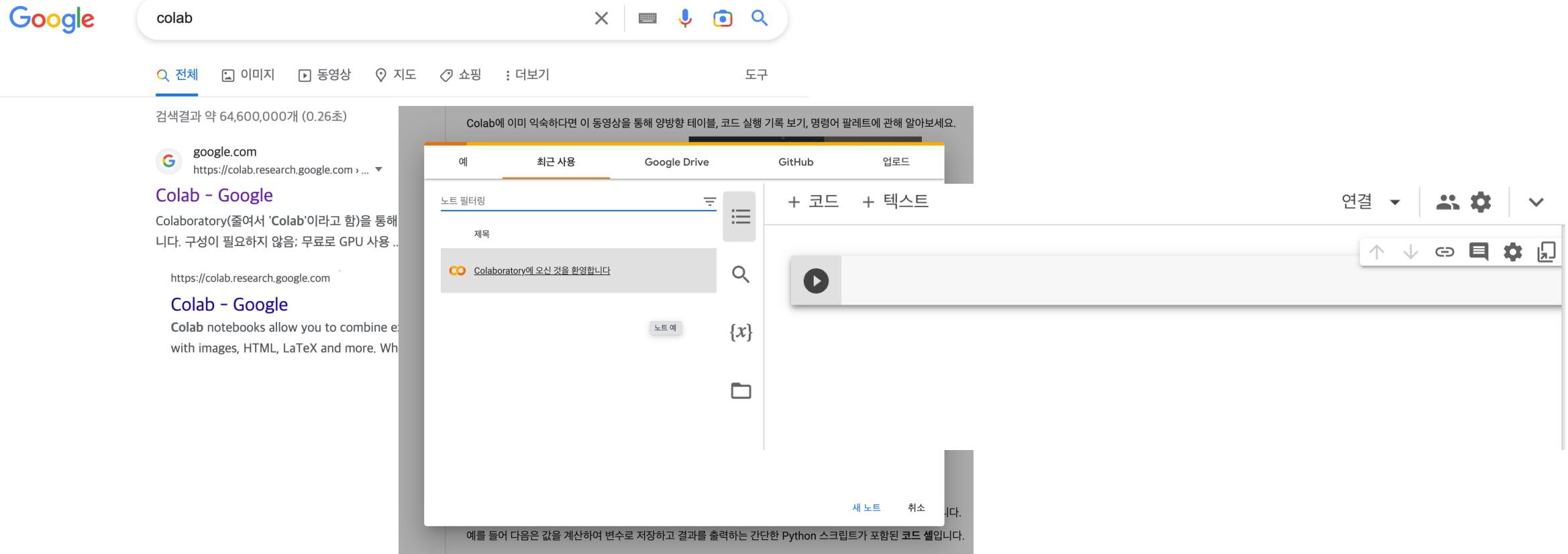
# “REST”

verb	Resource	Representation
GET	/movies	Get list of movies
GET	/movies/:id	Find a movie by its ID
POST	/movies	Create a new movie
PUT	/movies	Update an existing movie
DELETE	/movies	Delete an existing movie



< 같이 실습 />

# REST API



# REST API

공공데이터포털

JSONPlaceholder

Guide Sponsor this project Blog My JSON Server

<https://jsonplaceholder.typicode.com>

## {JSON} Placeholder

Free fake API for testing and prototyping.

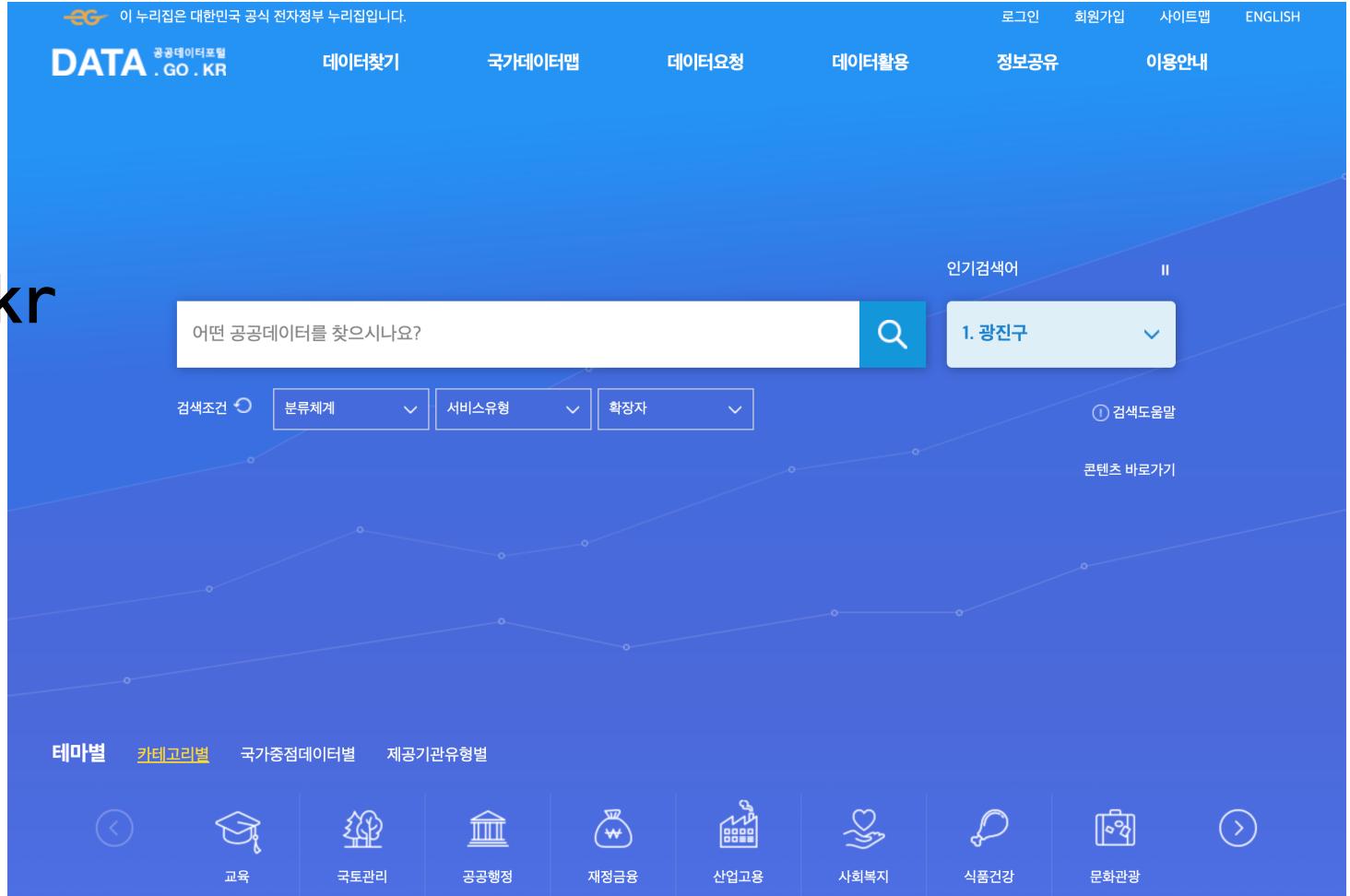
Powered by JSON Server + LowDB. Tested with XV.

Serving ~2 billion requests each month.

# REST API

공공데이터포털

<https://www.data.go.kr>



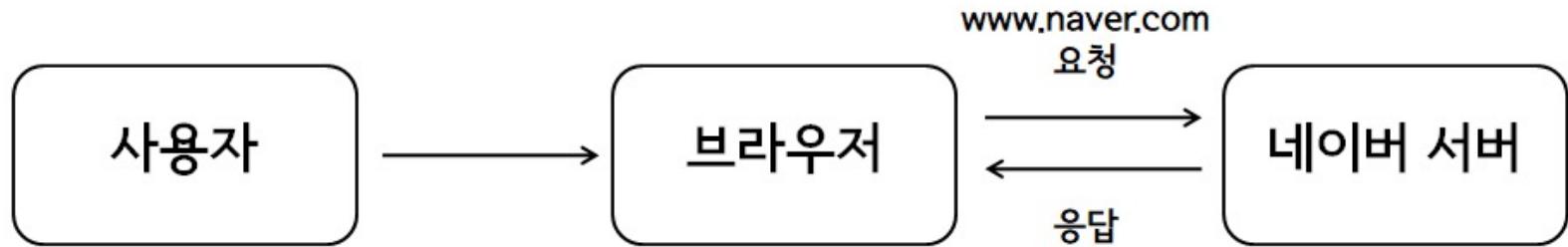
# < Crawling />

\* A powerful technique to collect data from the web

# | Crawling



# REST API



1. HTML
2. CSS
3. JS

# Data process

만족도가 높은! 강의

2019년 여름학기 **교양** 강의로 이동 >  
회계학원리(영강)  
이상호 교수님  
강의 만족도 **4.8**/5.0

2019년 2학기 **전공** 강의로 이동 >  
러시아작가론II  
석영중 교수님  
강의 만족도 **4.3**/5.0

2019년 2학기 **전공** 강의로 이동 >  
모바일포렌식  
한재혁 교수님  
강의 만족도 **4.5**/5.0

학점평가가 좋은! 강의

2019년 1학기 **전공** 강의로 이동 >  
라틴아메리카단편소설강독  
배지완 교수님  
학점 평가 **4.2**/5.0

2020년 1학기 **전공** 강의로 이동 >  
영어교과논리및논술(영강)  
Zoe Ksan Rubadeau 교수님  
학점 평가 **4.1**/5.0

2020년 1학기 **교양** 강의로 이동 >  
한국의신분과사회상(담당교수 변경/2....  
이송순 교수님  
학점 평가 **4.4**/5.0

< div />



< img />

< p />

파이썬 : ??????

# Parsing

\* the process of analyzing a string of symbols

< div />



< img

< p />

파이썬 : 😊

# | Crawling

# Beautifulsoup

< 같이 실습 />

# 과제 (3월 9일 전까지)

교수님 살려주세요

Opentutorials.org

후원 로그인 도움말

생활코딩  
Coding Everybody

WEB | DATABASE | 언어 | 클라이언트 | 서버 | 데이터 과학 | 개발도구 | 프로젝트 관리 | 기타

생활코딩



hello world



iamBeen22 11일 전  
2023-03-09



왜곡 19일 전  
2023-02-28

생활코딩의 세계에 오신 것을 환영합니다. 생활코딩은 일반인들에게 프로그래밍을 공부할 것인가를 생각해보기 전에 왜 프로그래밍을 공부하는 이유에 대한 이유를 힘 보시죠.



kosta 24일 전  
감사합니다 청준생을 위한 교육과정을 운영중인데  
생활코딩의 강의와 tutorial을 개강전 선수 학습용으로 제공하고자 합니다  
게시판 형식으로 영상 중심으로 제시할 예정이고  
출처사이트를 표시하도록 하겠습니다



101proof 28일 전  
2023-02-20

답글

< 과제 />

# 과제 (3월 23일 전까지)

opentutorials.org

후원 로그인 도움말

## 생활코딩 Coding Everybody

WEB | DATABASE | 언어 | 클라이언트 | 서버 | 데이터 과학 | 개발도구 | 프로젝트 관리 | 기타

생활코딩



iamBeen22 11일 전  
2023-03-09



왜곡 19일 전  
2023-02-28



kosta 24일 전  
감사합니다 취준생을 위한 교육과정을 운영중인데  
생활코딩의 강의와 tutorial을 개강전 선수 학습용으로 제공하고자 합니다  
게시판 형식으로 영상 중심으로 제시할 예정이고

출처사이트를 표시하도록 하겠습니다



101proof 28일 전  
2023-02-20

	name	time	comments
0	iamBeen22	11일 전	2023-03-09
1	왜곡	19일 전	2023-02-28
2	kosta	24일 전	감사합니다 취준생을 위한 교육과정을 운영중인데\n생활코딩의 강의와 tutorial을...
3	101proof	28일 전	2023-02-20
4	까마귀튀김	한 달 전	2023-02-18
5	Judynnn	한 달 전	2023-02-14
6	호떡이	한 달 전	23년 2월 12일 시작합니다. 화이팅!
7	탱	한 달 전	2023.02.08 우선 시작합니다!
8	소윤호	한 달 전	로컬호스트, 127.0.0.1 이 안열리는데 따로 설정을 해야하나요?
9	답답	한 달 전	23. 02. 01 가벼운 마음으로 시작합니다. 끝까지 와주할 수 있길!

# 과제 (3월 23일 전까지)

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

target_url = 'https://opentutorials.org/course/1

res = req.get(target_url)
soup = BeautifulSoup(res.text, 'html.parser')

li_comments = soup.find_all(class_='comment_content')
li_comments_cleaned = [e.text.strip() for e in li_comments]
li_name = soup.select('div.name.time > strong')
li_name_cleaned = [e.text for e in li_name]
li_time = soup.select('div.name.time > a > time')
li_time_cleaned = [e.text for e in li_time]

container = {'name':li_name_cleaned, 'time':li_time_cleaned, 'comments':li_comments_cleaned}
df_table = pd.DataFrame(container)
df_table
```