

크롬 network 탭 분석

18번째 세션

NEXT X LIKELION 전해주

| 크롬 개발자도구

[Elements]

HTML 코드를 분석하고, 실시간으로 수정할 수 있는 도구

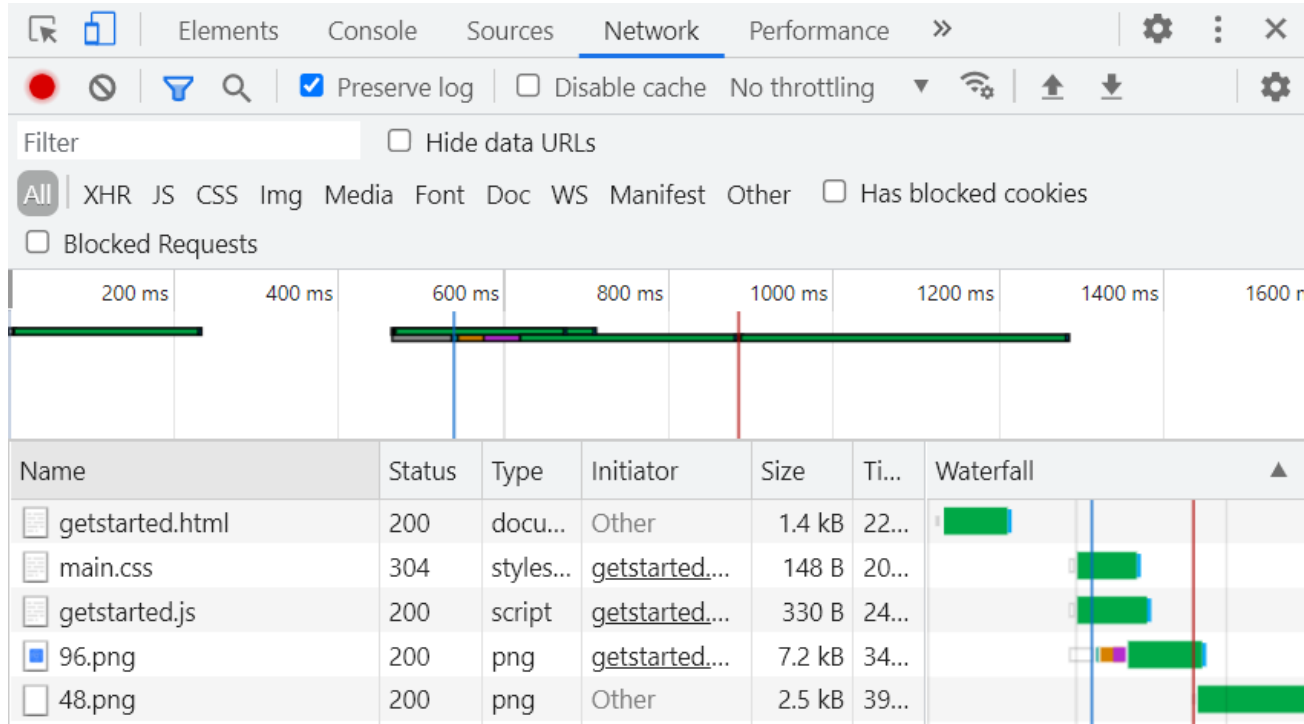
[Sources]

현재 로딩된 웹페이지에서 사용된
리소스(이미지, 스크립트, 데이터)를 열람 할 수 있는 도구

[Network]

서버와의 통신 내역을 보여주는 도구

크롬 개발자도구 Network탭



name : 이름

status : 상태 200이면 정상

type : js, font, css 종류

size : 불러온 파일의 사이즈

time : 해당 파일이 로드되는데 시간

| 크롬 개발자도구 Network탭

서버와의 통신 내역



브라우저와 서버사이 요청과 응답



HTTP 통신 과정

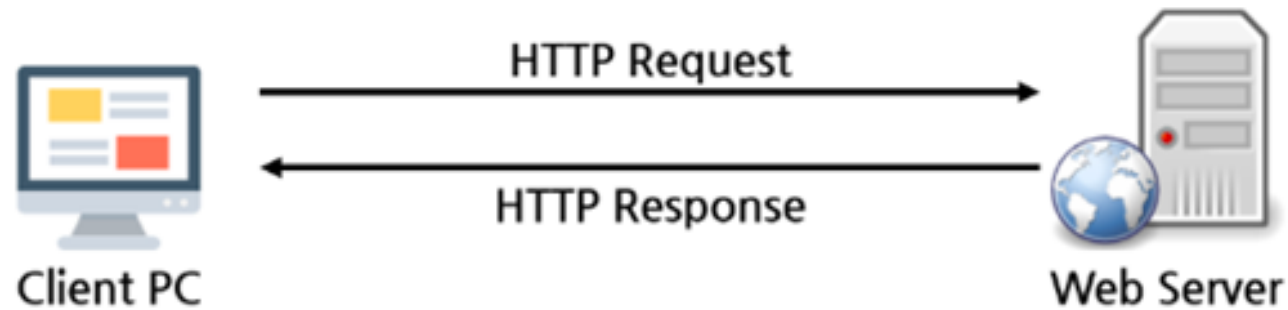
| HTTP란

하이퍼텍스트(HTML) 문서를 교환하기 위해 만들어진 **protocol(통신 규약)**

즉 웹상에서 네트워크로 서버끼리 통신을 할 때
어떠한 형식으로 서로 통신을 하자고 규정해 놓은 약속이라고 보면 된다.

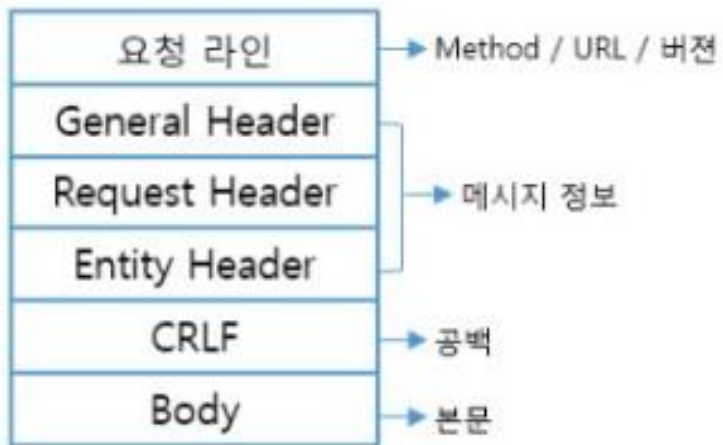
프론트엔드 서버와 클라이언트간의 통신에 사용된다.
또한 백엔드와 프론트엔드 서버간의 통신에도 사용된다.

| HTTP의 요청/응답



| HTTP 구조

HTTP Request Message



HTTP Response Message



[참고]

General Header : 요청/응답이 생성된 날짜 및 시간 등 전송되는 Body 콘텐츠와 관련 없는 HTTP 통신에 대한 일반적인 정보. HTTP 요청과 응답 메시지에 공통으로 사용된다.

Request/Response Header : 요청한 URL, 메소드 (GET, POST, HEAD), 요청 생성에 사용 된 브라우저 및 기타 정보와 같은 요청에 대한 정보.

Entity Header : 실제 메시지 또는 전송중인 HTTP 본문에 대한 정보. 콘텐츠 길이, 콘텐츠 언어, 인코딩, 만료 날짜 및 기타 중요한 정보.

| HTTP 구조

HTTP 구조를 직접 눈으로 확인해봅시다!

<https://devtools.glitch.me/network/getstarted.html>



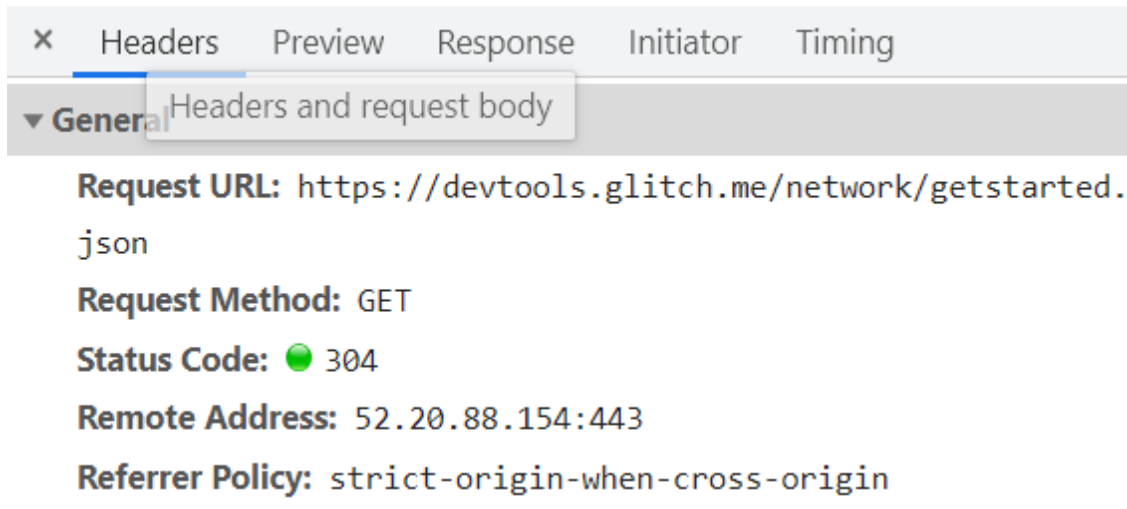
Inspect Network Activity Demo

This is the companion demo for the [Inspect Network Activity In Chrome DevTools](#) tutorial.

Get Data

자주 쓰이는 Header 정보

General Header



[Request URL]

요청을 보내는 곳

[Request Method]

request 방식

[Status Code]

요청이 잘 갔는지

[Remote Address]

host(request가 가는 곳)의 IP주소. : 뒤에는 포트번호

[Referrer Policy]

이 페이지를 방문하기 전에 어디에 있었냐에 따라

사용자 접속 제어 가능

<https://m.blog.naver.com/PostView.naver?isHttpsRedirect=true&blogId=opensellerbiz&logNo=222116598063>

자주 쓰이는 Header 정보

Request Header

×

Headers

Preview

Response

Initiator

Timing

▼ Request Headers

:authority:

devtools.glitch.me

:method:

GET

:path:

/network/getstarted.json

:scheme:

https

accept:

/

accept-encoding:

gzip, deflate, br

accept-language:

ko-KR,ko;q=0.9,en-US;q=0.8,en;q=0.7

if-modified-since:

Fri, 01 Feb 2019 20:58:50 GMT

if-none-match:

W/"52-168aadb5f10"

referer:

https://devtools.glitch.me/network/getstarted.html

sec-ch-ua:

" Not;A Brand";v="99", "Google Chrome";v="91", "Ch

[user-agent]

요청을 보내는 쪽(클라이언트)의 대한 정보

[accept]

이러한 형식으로 응답을 보내달라

[connection]

해당 요청이 끝난 후에 클라이언트와 서버가 계속해서 네트워크 커넥션을 유지 할 것인지 아니면 끊을 것인지.

[content-type]

해당 요청이 보내는 메시지 body의 타입.
ex)JSON을 보내면 application/json.

[content-length]

메세지 body의 길이.

자주 쓰이는 Header 정보

Request Header

[cookie 와 session]

공통점 : 데이터를 임시로 계속 저장해두는 역할. 비효율을 줄이고 생산성을 높이는 것이 목적

차이점 : 쿠키는 사용자에게 저장되고(자동로그인 , 오늘그만보기 팝업창 등)

세션은 서버에 저장되므로 보안이 우수하나 느림(로그인 한 정보들).

[cache]

브라우저를 사용하는 환경(로컬 컴퓨터)에 서버에서 받은 데이터를 저장한 파일 (쿠키와 동일)

이미지 등 재사용될 것 같거나 용량이 큰 리소스를 임시로 저장해서 렌더링 속도를 높이는 것이 목적

쿠키와는 비슷하지만 목적성에 차이가 있다

| 지금까지 정리

지금까지 웹이 어떻게 동작하는지
즉, 서버로 어떻게 요청을 보내고, 어떻게 응답을 받는지에 대해 알아보았습니다

| 지금까지 정리

그렇다면...

HTTP 형식만 맞춰준다면

웹이 동작하는 것처럼 요청을 **대신** 보낼 수 있지 않을까?

| 대신 요청하기



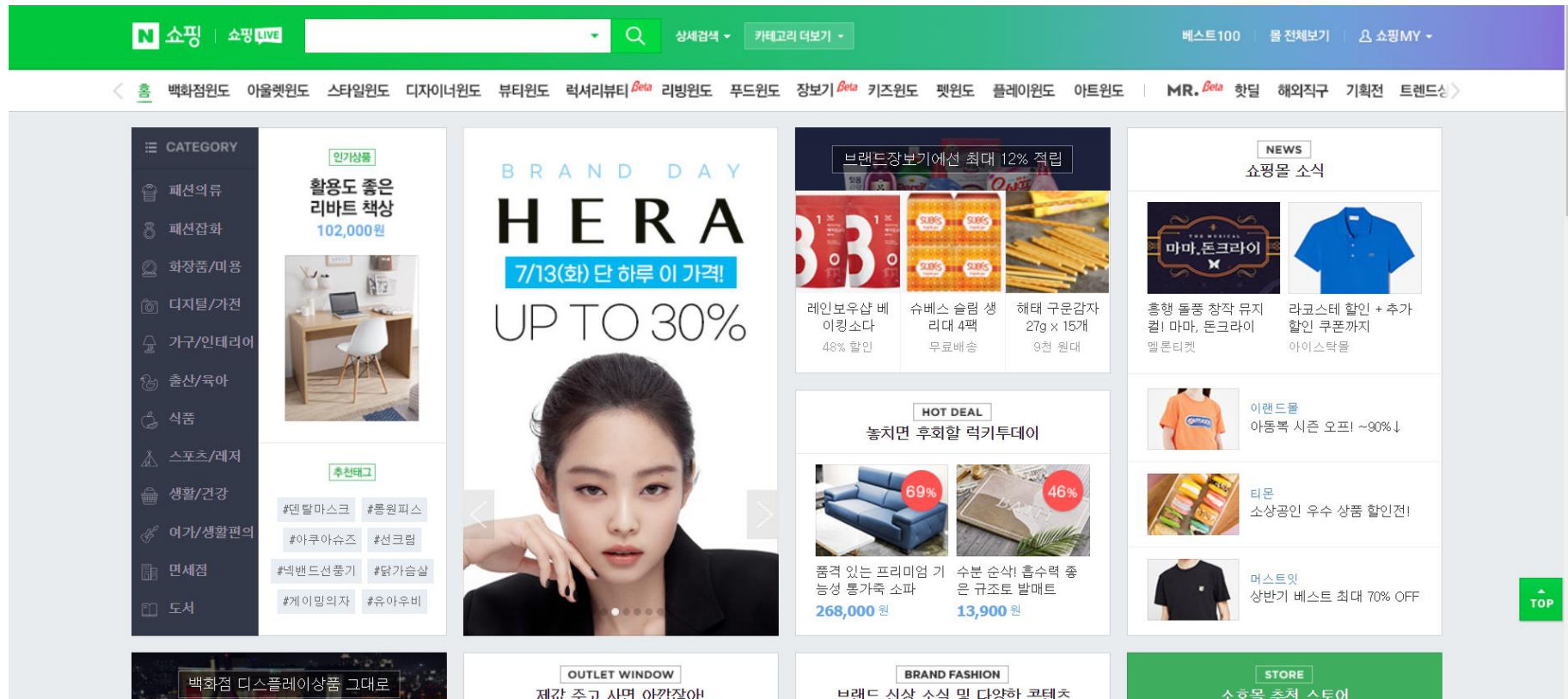
| 대신 요청하기

일부 사이트들에서 크롤링이 안되었던 이유는
자바스크립트가 웹 브라우저에서 작동하며 방해하기 때문!
하지만 이제는...

자바스크립트의 방해를 받지 않고
바로 서버로 요청을 전송해 크롤링을 할 수 있습니다.

네이버 스마트 스토어 크롤링

<https://search.shopping.naver.com>



네이버 스마트 스토어 크롤링

전체 815,760	가격비교 4,063	네이버페이 226,373	백화점/홈쇼핑 80,965	쇼핑원도 547	해외직구 199,836
---------------	---------------	------------------	-------------------	-------------	-----------------

· 네이버 랭킹순 · 낮은 가격순 · 높은 가격순 · 등록일순 · 리뷰 많은순

쇼핑물선택 상품타입(전체) 40개씩 보기

추천 리뷰 많은순 으로 정렬되었습니다. 적용기준 ON



삼성전자 갤럭시북 이온 NT950XCR-G58A

최저 1,199,000원 판매처 95

디지털/가전 > 노트북

화면크기 : 15인치(37~39cm) | 무게 : 1.26kg | 종류 : 코어i5 10세대 | 운영체제 : 미포함(FreeDos)
CPU : 코어i5-10210U | 칩셋 제조사 : 인텔 | 코드명 : 코멧레이크 | CPU속도 : 1.6GHz |

리뷰 ★★★★★ 7,094 · 등록일 2019.12. · ♥ 찜하기 417 · 📄 정보 수정요청

· 기본구성 SSD 256GB 1,199,000원 | 41개


· SSD 1TB 1,343,990원 | 10개

· SSD 250GB + SSD 256GB 1,635,510원 | 2개

· SSD 500GB 1,315,900원 | 21개

· SSD 2TB 1,899,000원 | 2개

· SSD 256GB + SSD 256GB 1,368,090원 | 10개



[삼성전자] 삼성전자 인기가전 모음_대형가전/생활가전/컴퓨터/휴대폰

82,980원

디지털/가전 > 노트북

[삼성카드5%할인:78,830원]

리뷰 6,650 · 등록일 2019.12. · ♥ 찜하기 0 · 📄 신고하기

롯데홈쇼핑 정보

플래티넘

N Pay 1%적립

배송비 무료

적립 | 할인 구매정보

상품명을 리뷰 많은 순으로
추출해보겠습니다.

| 네이버 스마트 스토어 크롤링












pipenv shell

pipenv install requests

```
import json
import requests

keyword = input('검색어를 입력하세요 : \n')
```

네이버 스마트 스토어 크롤링

Name	× Headers Preview Response Initiator Timing Cookies
 cc?a=ops.review&bw=990&px=399&py=559&sx=39...	
 click?a=ops.review&&ssc=svc.shopping.v3&&ts=162...	
 all?sort=review&pagingIndex=1&pagingSize=40&vi...	General Request URL: https://search.shopping.naver.com/search/all?sort=review&pagingIndex=1&pagingSize=40&viewType=list&productSet=total&deliveryFee=&deliveryTypeValue=&frm=NVSHATC&query=%EB%85%B8%ED%8A%B8%EB%B6%81&origQuery=%EB%85%B8%ED%8A%B8%EB%B6%81&iq=&eq=&xq= Request Method: GET Status Code:  200 Remote Address: 117.52.137.136:443 Referrer Policy: unsafe-url
 40eebacb-52bf-47f2-ad1d-5f34a7b53e82.jpg	
 platinum_v1.png	
 premium_v1.png	
 window_9_42.png	
 products?nvMid=21490861969,21491096190,184653...	
 21490861969.20200327125127.jpg?type=f140	
 21491096190.1.jpg?type=f140	Response Headers cache-control: no-cache, no-store, must-revalidate content-encoding: gzip content-type: application/json; charset=utf-8

네이버 스마트 스토어 크롤링

```
# all?sort=review&의 request headers
headers = {
    'authority': 'search.shopping.naver.com',
    'method': 'GET',
    'path': '/search/all?sort=review&pagingIndex=1&pagingSize=40&viewType=list&productSet=total&deliveryFee=&deliveryTypeValue=&frm=NVSHAKW&query=%EB%85%B8%ED%8A%B8%EB%B6%81&origQuery=%EB%85%B8%ED%8A%B8%EB%B6%81&iq=&eq=&xq=',
    'scheme': 'https',
    'accept': 'application/json, text/plain, */*',
    'accept-encoding': 'gzip, deflate, br',
    'accept-language': 'ko-KR,ko;q=0.9,en-US;q=0.8,en;q=0.7',
    'cookie': 'NNB=KLHGSTAVMDRGA; nid_inf=336029027; NID_AUT=blnkehevwonUM+LFdt7GL6+mG+oySu37FJ8/F6WYUIm1NaFvDIpLKct8epVLPwiw; NID_JKL=jnr9V95/mip/g3X+bnVlBrRmiR2BH+MaQUjSiYxVz0=; nx_ssl=2; page_uid=hMF0lsp0JXossf8eZkKsssstKR-253471; AD_SHP_BID=26; SHP_BID=4; BMR=s=1626030675086&r=https%3A%2F%2Fm.blog.naver.com%2FPostView.naver%3FisHttpsRedirect%3Dtrue%26blogId%3Dbestmember10%26logNo%3D220272365850&r2=https%3A%2F%2Fwww.google.com%2F; NID_SES=AAABkruVIptYPbXAF0sg1XBdkjnaDnCVE1nPvIB0HKQBC522qZ+KWW3A/w0cgFx64G44yJn70d3cFbhwPy2yi84clz8g4zVZ1drxshpKNLuvbzN3bgwFMSK2kth3QeGIgCoZuFaWmr3Zkeu+jT7Ba27HWLGkYbP8JP0AUGdTWjK0q9X+fbGAqib18LHGCasD4SdtCv7TvQG/ZrJIEbv0Spn1x8zKE0rmotsR7fW6e8lW6ryT6MMX+azr5vs/qbnyJLD9eZ0JUzq165Xw+/CUV0S5I9rn2tomLpqW2826MLCaXMDQSN1+ht0Bz53lyKKuQ+1sNGF+YDwb6BeVHjWZ6ktb0Gm2tXoT6mxaY/goEpTTmbfsw3ochP0ycgnZrKtoyiiZ7FDk2q8oL/HwEFpxhM/hY9Y7e9DN6+oEE mzT94KQEITbJ0t70crU01dfEW0zDtpdMnykAvQ1a+9Bye2XEraqAubo8i9gogcBeUdS+S0/+9nUzhrGZIJ+18MsS1Xp2fFRihZWN0HHnsbXFl9lGucIRRa101+3+7aiBzcXyctV/VtT; spage_uid=hMF0lsp0JXossf8eZkKsssstKR-253471; sus_val=EJYVC7PCR2tGkwbGk1er4cjG',
    'logic': 'PART',
    'referer': 'https://search.shopping.naver.com/search/all?query=%EB%85%B8%ED%8A%B8%EB%B6%81&cat_id=&frm=NVSHAKW',
    'sec-ch-ua': '" Not;A Brand";v="99", "Google Chrome";v="91", "Chromium";v="91"',
    'sec-ch-ua-mobile': '?1',
    'sec-fetch-dest': 'empty',
    'sec-fetch-mode': 'cors',
    'sec-fetch-site': 'same-origin',
    'urlprefix': '/api',
    'user-agent': 'Mozilla/5.0 (Linux; Android 6.0; Nexus 5 Build/MRA58N) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/91.0.4472.124 Mobile Safari/537.36'
}
```

| 네이버 스마트 스토어 크롤링

```
# all?sort=review& 의 query string parameters
params = (
    ('sort', 'review'),
    ('pagingIndex', '1'),
    ('pagingSize', '40'),
    ('viewType', 'list'),
    ('productSet', 'total'),
    ('frm', 'NVSHAKW'),
    ('query', keyword),
    ('origQuery', keyword),
    ('iq', ''),
    ('eq', ''),
    ('xq', ''),
)
```

네이버 스마트 스토어 크롤링


```
# all?sort=review &의 request url
# response = requests.get('https://search.shopping.naver.com/search/all?sort=rel&pagingIndex=2&pagingSize=40&viewType=list&productSet=total&frm=NVSHATC&query=%EB%85%B8%ED%8A%B8&origQuery=%EB%85%B8%ED%8A%B8&iq=&eq=&xq=', headers=headers)
response = requests.get('https://search.shopping.naver.com/search/all', headers=headers, params=params)


# json 디코딩 (json to python)
result = json.loads(response.text)
# print(list(result.keys()))
# print(result['shoppingResult'].keys())
result_contents = result['shoppingResult']['products']


# 상품명만 출력
for content in result_contents:
    print(content['productTitle'])
```


Youtube Premium 만들기


<https://www.youtube.com/>


YouTube^{KR}


홈


탐색


구독


보관함

시청 기록




내 동영상

나중에 볼 동영상


좋아요 표시한 동영상

더보기

아이유




곡 필터




1. 다들봐	13. 좋은 날
2. Coin	14. 너와 나
3. Celebrity	15. 너의 의미
4. 내 손을 잡아	16. 사랑이 잘
5. 내맘	17. 스물셋
6. Blueming	18. 미치 매치 크리스마스
7. 마음을 드려요	19. 무릎
8. Love poem	20. 가을바람
9. 밤영지	21. 분홍선
10. 팔레트	22. 마음
11. 비비	23. 나만 몰랐던 이야기
12. 금요일에 만나요	24. 잠 못 드는 밤

[PLAYLIST] 아이유노래모음

조회수 209만회 · 3개월 전


 나 들으려고 올리는 채널

[PLAYLIST] 아이유노래모음 저는 광고를 붙이지 않습니다. 유튜브 정책상 중간 광고가 나타날 수 있습니다. 모든 광고수익은 ...



아이유가 유인나를 희롱하는 이휘재에게 날린 돌직구

조회수 13만회 · 11시간 전

 드라마갤러리

Session 18

NEXT X LIKELION

Youtube Premium 만들기

The screenshot displays the 'Headers' tab of a web browser's developer tools. On the left, a list of network requests is shown, with the first request, 'results?search_query=%EC%95%84%EC%9D%B4%EC...', selected. The right pane shows the details for this request. Under the 'General' section, the 'Request URL' is 'https://m.youtube.com/results?search_query=%EC%95%84%EC%9D%B4%EC%9C%A0&pbj=1', the 'Request Method' is 'GET', the 'Status Code' is '200' (indicated by a green circle), the 'Remote Address' is '172.217.175.110:443', and the 'Referrer Policy' is 'strict-origin-when-cross-origin'. Under the 'Response Headers' section, several headers are listed: 'alt-svc' with a complex value including 'h3=:443'; ma=2592000, h3-29=:443; ma=2592000, h3-T051=:443; ma=2592000, h3-Q050=:443; ma=2592000, h3-Q046=:443; ma=2592000, h3-Q043=:443; ma=2592000, quic=:443; ma=2592000; v="46,43"', 'cache-control' with 'no-cache, no-store, max-age=0, must-revalidate', 'content-disposition' with 'attachment', 'content-encoding' with 'br', and 'content-type' with 'application/json; charset=utf-8'.

Name
results?search_query=%EC%95%84%EC%9D%B4%EC...
paralleladview?ai=CyKiswknrYI7GKsuligaGjL2YDQDQ...
mqdefault.webp
mqdefault.webp
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..
log_event?alt=json&key=AlzaSyAO_FJ2SlqU8Q4STE..

× Headers Preview Response Initiator Timing Cookies

▼ General

Request URL: https://m.youtube.com/results?search_query=%EC%95%84%EC%9D%B4%EC%9C%A0&pbj=1

Request Method: GET

Status Code: 200

Remote Address: 172.217.175.110:443

Referrer Policy: strict-origin-when-cross-origin

▼ Response Headers

alt-svc: h3=":443"; ma=2592000, h3-29=":443"; ma=2592000, h3-T051=":443"; ma=2592000, h3-Q050=":443"; ma=2592000, h3-Q046=":443"; ma=2592000, h3-Q043=":443"; ma=2592000, quic=":443"; ma=2592000; v="46,43"

cache-control: no-cache, no-store, max-age=0, must-revalidate

content-disposition: attachment

content-encoding: br

content-type: application/json; charset=utf-8

| Youtube Premium 만들기

```
import json
import requests

# results?search_query=% 의 request headers 파트에서 확인
headers = {
    'x-youtube-client-name': '1',
    'x-youtube-client-version': '2.20210708.06.00'
}

# url ? 뒤에 들어가는 parameter
# results?search_query=% 의 query string parameters 파트에서 확인
params = (
    ('search_query', input('검색어를 입력하세요 : \n')),
    ('pbj', '1')
)

# 검색어 입력했을 때 url
# results?search_query=% 의 general 파트의 request url
response = requests.get('https://www.youtube.com/results', headers=headers, params=params)
```

| Youtube Premium 만들기

```
# json 디코딩 (json to python)
result = json.loads(response.text)
contents = result[1]['response']['contents']['twoColumnSearchResultsRenderer']['primaryContents']['sectionListRenderer']['contents'][0]['itemSectionRenderer']['contents']

# 검색된 영상들의 videoId 값들을 videoId_list에 저장
videoId_list = []
for content in contents:
    keys = list(content.keys())

    if 'videoRenderer' in keys:
        videoId_list.append(content['videoRenderer']['videoId'])

print(videoId_list)
```

Youtube Premium 만들기

The screenshot displays the Chrome DevTools Network tab. The left sidebar shows a list of resources, with the selected resource being a video player key request. The right sidebar shows the details for this request.

Name	Headers	Preview	Response	Initiator	>>
00CMxzZ5o_oNA8E6LMbbDmtxK-SUboRvQf9tkrftkQzL4pjREZ...3TX_	▼ General				
sw.js	Request URL: https://www.youtube.com/youtubei/v				
serviceworker-kevlar-appshell.js	player?key=AIzaSyAO_FJ2SlqU8Q4STEHLGCilw_Y9_11				
sw.js_data	8				
app_shell	Request Method: POST				
webcomponents-all-noPatch.js	Status Code: 200				
www-player-2x.css	Remote Address: 172.217.31.142:443				
next?key=AlzaSyAO_FJ2SlqU8Q4STEHLGCilw_Y9_11qcW8	Referrer Policy: strict-origin-when-cross-origin				
player?key=AlzaSyAO_FJ2SlqU8Q4STEHLGCilw_Y9_11qcW8	▼ Response Headers				
generate_204	alt-svc: h3=":443"; ma=2592000,h3-29=":443"; ma=				
favicon.ico	2000,h3-T051=":443"; ma=2592000,h3-Q050=":443"				
endscreen.js	a=2592000,h3-Q046=":443"; ma=2592000,h3-Q043="				
annotations_module.js	3"; ma=2592000,quic=":443"; ma=2592000; v="46,				
data:image/png;base...	3"				
generate_204	cache-control: private				
annotations_module.js	content-encoding: br				
data:image/svg+xml;...	content-length: 24946				
qoe?event=streamingstats&fmt=399&afmt=251&cpn=2R_h...14332					

Youtube Premium 만들기

```
for videoId in videoId_list:

    print("-----")
    # player?key= 의 query string parameters 파트 복붙
    # 심포 주의
    video_params = (
        ('key', 'AIzaSyAO_FJ2SlqU8Q4STEHLGCilw_Y9_11qcW8'),
    )

    # player?key= 의 request payload 파트 복붙 (view source)
    # videoId 값만 변수로 변경
    data = {
        "videoId":videoId,
        "context":{
            "client":{
                # 너무 길어서 중간 생략합니다.
            },
            "user " :{},
            "request " :{
                "useSNI " :True,
                "internalExperimentFlags " :[],
                "consistencyTokenJars " :[]
            },
            "adSignalsInfo " :{
                "bid":"ANyPxKrzWuvB5JF2P_ysQQY6TNWhGJZpzLsE_PrEdoIqst846L72mFhr68xv8iHwOg09MIG1do5A_fi781dGhyWeI24dXFlBAQ"
            }
        },
        "playbackContext":{
            "contentPlaybackContext":{
                # 너무 길어서 중간 생략합니다.
            }
        },
        "cpn":"9wbWDn8Kn3B5Yv4s","captionParams":{}
    }
```

Youtube Premium 만들기

```
# json 인코딩 (python to json)
data = json.dumps(data)

# player?key= 의 general 파트의 request url
response = requests.post('https://m.youtube.com/youtubei/v1/player',
                        headers=video_headers, params=video_params, data=data)

# json 디코딩 (json to python)
result = json.loads(response.text)
# print(list(result['videoDetails']['videoId']))
player_response = result['streamingData']['formats'][1]

keys = list(player_response.keys())

if 'url' in keys :
    print(player_response['url'])
```

| 과제

과제는 따로 없습니다!

오늘 배운 내용 복습



[추가 기능 참고]

크롬 개발자도구를 이용해서 자바스크립트 디버깅하는 법

(vscode와 크롬 연동까지)

<https://subicura.com/2018/02/14/javascript-debugging.html>